Texts and Readings in Mathematics 38

Terence Tao

Analysis II

Fourth Edition

HINDUSTAN BOOK AGENCY

Springer

Texts and Readings in Mathematics

The **Texts and Readings in Mathematics** series publishes high-quality textbooks, research-level monographs, lecture notes and contributed volumes. Undergraduate and graduate students of mathematics, research scholars and teachers would find this book series useful. The volumes are carefully written as teaching aids and highlight characteristic features of the theory. Books in this series are co-published with Hindustan Book Agency, New Delhi, India.

Terence Tao

Analysis II

Fourth Edition

HINDUSTAN BOOK AGENCY

Springer

Terence Tao
Department of Mathematics
University of California Los Angeles
Los Angeles, CA, USA

ISSN 2366-8717 ISSN 2366-8725 (electronic)
Texts and Readings in Mathematics
ISBN ISBN 978-981-19-7284-3 (eBook)
https://doi.org/10.1007/978-981-19-7284-3

Jointly published with Hindustan Book Agency
This work is a co-publication with Hindustan Book Agency, New Delhi, licensed for sale in all countries in electronic form only. Sold and distributed in print across the world by Hindustan Book Agency, P-19 Green Park Extension, New Delhi 110016, India.

This Springer imprint is published by the registered company Springer Nature Singapore Pte Ltd.
The registered company address is: 152 Beach Road, #21-01/04 Gateway East, Singapore 189721, Singapore

To my parents, for everything

Preface to the First Edition

This text originated from the lecture notes I gave teaching the honours undergraduate-level real analysis sequence at the University of California, Los Angeles, in 2003. Among the undergraduates here, real analysis was viewed as being one of the most difficult courses to learn, not only because of the abstract concepts being introduced for the first time (e.g., topology, limits, measurability, etc.), but also because of the level of rigour and proof demanded of the course. Because of this perception of difficulty, one was often faced with the difficult choice of either reducing the level of rigour in the course in order to make it easier, or to maintain strict standards and face the prospect of many undergraduates, even many of the bright and enthusiastic ones, struggling with the course material.

Faced with this dilemma, I tried a somewhat unusual approach to the subject. Typically, an introductory sequence in real analysis assumes that the students are already familiar with the real numbers, with mathematical induction, with elementary calculus, and with the basics of set theory, and then quickly launches into the heart of the subject, for instance the concept of a limit. Normally, students entering this sequence do indeed have a fair bit of exposure to these prerequisite topics, though in most cases the material is not covered in a thorough manner. For instance, very few students were able to actually *define* a real number, or even an integer, properly, even though they could visualize these numbers intuitively and manipulate them algebraically. This seemed to me to be a missed opportunity. Real analysis is one of the first subjects (together with linear algebra and abstract algebra) that a student encounters, in which one truly has to grapple with the subtleties of a truly rigorous mathematical proof. As such, the course offered an excellent chance to go back to the foundations of mathematics, and in particular the opportunity to do a proper and thorough construction of the real numbers.

Thus the course was structured as follows. In the first week, I described some well-known "paradoxes" in analysis, in which standard laws of the subject (e.g., interchange of limits and sums, or sums and integrals) were applied in a non-rigorous way to give nonsensical results such as $0 = 1$. This motivated the need to go back to the very beginning of the subject, even to the very definition of the natural numbers, and check all the foundations from scratch. For instance, one of the first homework assignments was to check (using only the Peano axioms) that addition was associative for natural numbers (i.e., that $(a + b) + c = a + (b + c)$ for all natural numbers a, b, c: see Exercise 2.2.1). Thus even in the first week, the students had to write rigorous proofs using mathematical induction. After we had derived all the basic properties of the natural numbers, we then moved on to the integers (initially defined as formal differences of natural numbers); once the students had verified all the basic properties of the integers, we moved on to the rationals (initially defined as formal quotients of integers); and then from there we moved on (via formal limits of Cauchy sequences) to the reals. Around the same time, we covered the basics of set theory, for instance demonstrating the uncountability of the reals. Only then (after about ten lectures) did we begin what one normally considers the heart of undergraduate real analysis—limits, continuity, differentiability, and so forth.

The response to this format was quite interesting. In the first few weeks, the students found the material very easy on a conceptual level, as we were dealing only with the basic properties of the standard number systems. But on an intellectual level it was very challenging, as one was analyzing these number systems from a foundational viewpoint, in order to rigorously derive the more advanced facts about these number systems from the more primitive ones. One student told me how difficult it was to explain to his friends in the non-honours real analysis sequence (a) why he was still learning how to show why all rational numbers are either positive, negative, or zero (Exercise 4.2.4), while the non-honours sequence was already distinguishing absolutely convergent and convergent series, and (b) why, despite this, he thought his homework was significantly harder than that of his friends. Another student commented to me, quite wryly, that while she could obviously *see* why one could always divide a natural number n into a positive integer q to give a quotient a and a remainder r less than q (Exercise 2.3.5), she still had, to her frustration, much difficulty in writing down a proof of this fact. (I told her that later in the course she would have to prove statements for which it would not be as obvious to see that the statements were true; she did not seem to be particularly consoled by this.) Nevertheless, these students greatly enjoyed the homework, as when they did perservere and obtain a rigorous proof of an intuitive fact, it solidified the link in their minds between the abstract manipulations of formal mathematics and their informal intuition of mathematics (and of the real world), often in a very satisfying way. By the time they were assigned the task of giving the infamous "epsilon and delta" proofs in real analysis, they had already had so much experience with formalizing intuition, and in discerning the subtleties of mathematical logic (such as the distinction between the "for all" quantifier and the "there exists" quantifier), that the transition to these proofs was fairly smooth, and we were able to cover material both thoroughly and rapidly. By the tenth week, we had caught up with the non-honours class, and the students were verifying the change of variables formula for Riemann–Stieltjes integrals, and showing that piecewise continuous functions were Riemann integrable. By the conclusion of the sequence in the twentieth week, we had covered (both in lecture and in homework) the convergence theory of Taylor and Fourier series, the inverse and implicit function theorem for continuously differentiable functions of several variables, and established the dominated convergence theorem for the Lebesgue integral.

In order to cover this much material, many of the key foundational results were left to the student to prove as homework; indeed, this was an essential aspect of the course, as it ensured the students truly appreciated the concepts as they were being introduced. This format has been retained in this text; the majority of the exercises consist of proving lemmas, propositions and theorems in the main text. Indeed, I would strongly recommend that one do as many of these exercises as possible—and this includes those exercises proving "obvious" statements—if one wishes to use this text to learn real analysis; this is not a subject whose subtleties are easily appreciated just from passive reading. Most of the chapter sections have a number of exercises, which are listed at the end of the section.

To the expert mathematician, the pace of this book may seem somewhat slow, especially in early chapters, as there is a heavy emphasis on rigour (except for those discussions explicitly marked "Informal"), and justifying many steps that would ordinarily be quickly passed over as being self-evident. The first few chapters develop (in painful detail) many of the "obvious" properties of the standard number systems, for instance that the sum of two positive real numbers is again positive (Exercise 5.4.1), or that given any two distinct real numbers, one can find rational number between them (Exercise 5.4.5). In these foundational chapters, there is also an emphasis on *non-circularity*—not using later, more advanced results to prove earlier, more primitive ones. In particular, the usual laws of algebra are not used until they are derived (and they have to be derived separately for the natural numbers, integers, rationals, and reals). The reason for this is that it allows the students to learn the art of abstract reasoning, deducing true facts from a limited set of assumptions, in the friendly and intuitive setting of number systems; the payoff for this practice comes later, when one has to utilize the same type of reasoning techniques to grapple with more advanced concepts (e.g., the Lebesgue integral).

The text here evolved from my lecture notes on the subject, and thus is very much oriented towards a pedagogical perspective; much of the key material is contained inside exercises, and in many cases I have chosen to give a lengthy and tedious, but instructive, proof instead of a slick abstract proof. In more advanced textbooks, the student will see shorter and more conceptually coherent treatments of this material, and with more emphasis on intuition than on rigour; however, I feel it is important to know how to do analysis rigorously and "by hand" first, in order to truly appreciate the more modern, intuitive and abstract approach to analysis that one uses at the graduate level and beyond.

The exposition in this book heavily emphasizes rigour and formalism; however this does not necessarily mean that lectures based on this book have to proceed the same way. Indeed, in my own teaching I have used the lecture time to present the intuition behind the concepts (drawing many informal pictures and giving examples), thus providing a complementary viewpoint to the formal presentation in the text. The exercises assigned as homework provide an essential bridge between the two, requiring the student to combine both intuition and formal understanding together in order to locate correct proofs for a problem. This I found to be the most difficult task for the students, as it requires the subject to be genuinely *learnt*, rather than merely memorized or vaguely absorbed. Nevertheless, the feedback I received from the students was that the homework, while very demanding for this reason, was also very rewarding, as it allowed them to connect the rather abstract manipulations of formal mathematics with their innate intuition on such basic concepts as numbers, sets, and functions. Of course, the aid of a good teaching assistant is invaluable in achieving this connection.

With regard to examinations for a course based on this text, I would recommend either an open-book, open-notes examination with problems similar to the exercises given in the text (but perhaps shorter, with no unusual trickery involved), or else a take-home examination that involves problems comparable to the more intricate exercises in the text. The subject matter is too vast to force the students to memorize the definitions and theorems, so I would not recommend a closed-book examination, or an examination based on regurgitating extracts from the book. (Indeed, in my own examinations I gave a supplemental sheet listing the key definitions and theorems which were relevant to the examination problems.) Making the examinations similar to the homework assigned in the course will also help motivate the students to work through and understand their homework problems as thoroughly as possible (as opposed to, say, using flash cards or other such devices to memorize material), which is good preparation not only for examinations but for doing mathematics in general.

Some of the material in this textbook is somewhat peripheral to the main theme and may be omitted for reasons of time constraints. For instance, as set theory is not as fundamental to analysis as are the number systems, the chapters on set theory (Chapters 3, 8) can be covered more quickly and with substantially less rigour, or be given as reading assignments. The appendices on logic and the decimal system are intended as optional or supplemental reading and would probably not be covered in the main course lectures; the appendix on logic is particularly suitable for reading concurrently with the first few chapters. Also, Chapter 5 (on Fourier series) is not needed elsewhere in the text and can be omitted.

For reasons of length, this textbook has been split into two volumes. The first volume is slightly longer, but can be covered in about thirty lectures if the peripheral material is omitted or abridged. The second volume refers at times to the first, but can also be taught to students who have had a first course in analysis from other sources. It also takes about thirty lectures to cover.

I am deeply indebted to my students, who over the progression of the real analysis course corrected several errors in the lectures notes from which this text is derived, and gave other valuable feedback. I am also very grateful to the many anonymous referees who made several corrections and suggested many important improvements to the text. I also thank Adam, James Ameril, Quentin Batista, Biswaranjan Behara, José Antonio Lara Benítez, Dingjun Bian, Petrus Bianchi, Phillip Blagoveschensky, Tai-Danae Bradley, Brian, Eduardo Buscicchio, Carlos, cebismellim, Matheus Silva Costa, Gonzales Castillo Cristhian, Ck, William Deng, Kevin Doran, Lorenzo Dragani, EO, Florian, Gyao Gamm, Evangelos Georgiadis, Aditya Ghosh, Elie Goudout, Ti Gong, Ulrich Groh, Gökhan Güçlü, Yaver Gulusoy, Christian Gz., Kyle Hambrook, Minyoung Jeong, Bart Kleijngeld, Erik Koelink, Brett Lane, David Latorre, Matthis Lehmkühler, Bin Li, Percy Li, Ming Li, Mufei Li, Zijun Liu, Rami Luisto, Jason M., Manoranjan Majji, Mercedes Mata, Simon Mayer, Geoff Mess, Pieter Naaijkens, Vineet Nair, Jorge Peña-Vélez, Cristina Pereyra, Huaying Qiu, David Radnell, Tim Reijnders, Issa Rice, Eric Rodriquez, Pieter Roffelsen, Luke Rogers, Feras Saad, Gabriel Salmerón, Vijay Sarthak, Leopold Schlicht, Marc Schoolderman, SkysubO, Rainer aus dem Spring, Sundar, Rafał Szlendak, Karim Taya, Chaitanya Tappu, Winston Tsai, Kent Van Vels, Andrew Verras, Murtaza Wani, Daan Wanrooy, John Waters, Yandong Xiao, Sam Xu, Xueping, Hongjiang Ye, Luqing Ye, Muhammad Atif Zaheer, Zelin, and the students of Math 401/501 and Math 402/502 at the University of New Mexico for corrections to the first, second, and third editions.

Terence Tao

Preface to Subsequent Editions

Since the publication of the first edition, many students and lecturers have communicated a number of minor typos and other corrections to me. There was also some demand for a hardcover edition of the texts. Because of this, the publishers and I have decided to incorporate the corrections and issue a hardcover second edition of the textbooks. The layout, page numbering, and indexing of the texts have also been changed; in particular the two volumes are now numbered and indexed separately. However, the chapter and exercise numbering, as well as the mathematical content, remains the same as the first edition, and so the two editions can be used more or less interchangeably for homework and study purposes.

The third edition contains a number of corrections that were reported for the second edition, together with a few new exercises, but are otherwise essentially the same text. The fourth edition similarly incorporates a large number of additional corrections reported since the release of the third edition, as well as some additional exercises.

Los Angeles, USA — Terence Tao

Contents

1 Metric Spaces 1
1.1 Definitions and Examples 1
1.2 Some Point-Set Topology of Metric Spaces 9
1.3 Relative Topology 13
1.4 Cauchy Sequences and Complete Metric Spaces 14
1.5 Compact Metric Spaces 18

2 Continuous Functions on Metric Spaces 23
2.1 Continuous Functions 23
2.2 Continuity and Product Spaces 25
2.3 Continuity and Compactness 28
2.4 Continuity and Connectedness 30
2.5 Topological Spaces (Optional) 33

3 Uniform Convergence 39
3.1 Limiting Values of Functions 40
3.2 Pointwise and Uniform Convergence 42
3.3 Uniform Convergence and Continuity 46
3.4 The Metric of Uniform Convergence 48
3.5 Series of Functions; the Weierstrass M-Test 51
3.6 Uniform Convergence and Integration 53
3.7 Uniform Convergence and Derivatives 55
3.8 Uniform Approximation by Polynomials 58

4 Power Series 65
4.1 Formal Power Series 65
4.2 Real Analytic Functions 67
4.3 Abel's Theorem 72
4.4 Multiplication of Power Series 75
4.5 The Exponential and Logarithm Functions 78
4.6 A Digression on Complex Numbers 81
4.7 Trigonometric Functions 88

5 Fourier Series . . . 93
5.1 Periodic Functions . . . 94
5.2 Inner Products on Periodic Functions . . . 96
5.3 Trigonometric Polynomials . . . 99
5.4 Periodic Convolutions . . . 102
5.5 The Fourier and Plancherel Theorems . . . 106

6 Several Variable Differential Calculus . . . 113
6.1 Linear Transformations . . . 113
6.2 Derivatives in Several Variable Calculus . . . 119
6.3 Partial and Directional Derivatives . . . 122
6.4 The Several Variable Calculus Chain Rule . . . 129
6.5 Double Derivatives and Clairaut's Theorem . . . 131
6.6 The Contraction Mapping Theorem . . . 133
6.7 The Inverse Function Theorem in Several Variable Calculus . . . 136
6.8 The Implicit Function Theorem . . . 141

7 Lebesgue Measure . . . 145
7.1 The Goal: Lebesgue Measure . . . 146
7.2 First Attempt: Outer Measure . . . 148
7.3 Outer Measure Is not Additive . . . 156
7.4 Measurable Sets . . . 158
7.5 Measurable Functions . . . 163

8 Lebesgue Integration . . . 167
8.1 Simple Functions . . . 167
8.2 Integration of Non-negative Measurable Functions . . . 172
8.3 Integration of Absolutely Integrable Functions . . . 179
8.4 Comparison with the Riemann Integral . . . 184
8.5 Fubini's Theorem . . . 185

Index . . . 191

About the Author

Terence Tao has been a professor of Mathematics at the University of California Los Angeles (UCLA), USA, since 1999, having completed his Ph.D. under Prof. Elias Stein at Princeton University, USA, in 1996. Tao's areas of research include harmonic analysis, partial differential equations, combinatorics, and number theory. He has received a number of awards, including the Salem Prize in 2000, the Bochner Prize in 2002, the Fields Medal in 2006, the MacArthur Fellowship in 2007, the Waterman Award in 2008, the Nemmers Prize in 2010, the Crafoord Prize in 2012, and the Breakthrough Prize in Mathematics in 2015. Terence Tao also currently holds the James and Carol Collins chair in Mathematics at UCLA and is a fellow of the Royal Society, the Australian Academy of Sciences (the corresponding member), the National Academy of Sciences (a foreign member), and the American Academy of Arts and Sciences. He was born in Adelaide, Australia, in 1975.

Chapter 1
Metric Spaces

1.1 Definitions and Examples

In Definition 6.1.5 we defined what it meant for a sequence $(x_n)_{n=m}^{\infty}$ of real numbers to converge to another real number x; indeed, this meant that for every $\varepsilon > 0$, there exists an $N \geq m$ such that $|x - x_n| \leq \varepsilon$ for all $n \geq N$. When this is the case, we write $\lim_{n\to\infty} x_n = x$.

Intuitively, when a sequence $(x_n)_{n=m}^{\infty}$ converges to a limit x, this means that somehow the elements x_n of that sequence will eventually be as close to x as one pleases. One way to phrase this more precisely is to introduce the *distance function* $d(x, y)$ between two real numbers by $d(x, y) := |x - y|$. (Thus for instance $d(3, 5) = 2$, $d(5, 3) = 2$, and $d(3, 3) = 0$.) Then we have

Lemma 1.1.1 *Let $(x_n)_{n=m}^{\infty}$ be a sequence of real numbers, and let x be another real number. Then $(x_n)_{n=m}^{\infty}$ converges to x if and only if $\lim_{n\to\infty} d(x_n, x) = 0$.*

Proof See Exercise 1.1.1. □

One would now like to generalize this notion of convergence, so that one can take limits not just of sequences of real numbers, but also sequences of complex numbers, or sequences of vectors, or sequences of matrices, or sequences of functions, even sequences of sequences. One way to do this is to redefine the notion of convergence each time we deal with a new type of object. As you can guess, this will quickly get tedious. A more efficient way is to work *abstractly*, defining a very general class of spaces—which includes such standard spaces as the real numbers, complex numbers, vectors, etc.—and define the notion of convergence on this entire class of spaces at once. (A *space* is just the set of all objects of a certain type—the space of all real numbers, the space of all 3×3 matrices, etc. Mathematically, there is not much distinction between a space and a set, except that spaces tend to have much more structure than what a random set would have. For instance, the space of real numbers comes with operations such as addition and multiplication, while a general set would not.)

It turns out that there are two very useful classes of spaces which do the job. The first class is that of *metric spaces*, which we will study here. There is a more general

T. Tao, *Analysis II*, Texts and Readings in Mathematics,
https://doi.org/10.1007/978-981-19-7284-3_1

class of spaces, called *topological spaces*, which is also very important, but we will only deal with this generalization briefly, in Sect. 2.5.

Roughly speaking, a metric space is any space X which has a concept of *distance* $d(x, y)$—and this distance should behave in a reasonable manner. More precisely, we have

Definition 1.1.2 (*Metric spaces*) A *metric space* (X, d) is a space X of objects (called *points*), together with a *distance function* or *metric* $d : X \times X \to [0, +\infty)$, which associates to each pair x, y of points in X a non-negative real number $d(x, y) \geq 0$. Furthermore, the metric must satisfy the following four axioms:

(a) For any $x \in X$, we have $d(x, x) = 0$.
(b) (Positivity) For any *distinct* $x, y \in X$, we have $d(x, y) > 0$.
(c) (Symmetry) For any $x, y \in X$, we have $d(x, y) = d(y, x)$.
(d) (Triangle inequality) For any $x, y, z \in X$, we have $d(x, z) \leq d(x, y) + d(y, z)$.

In many cases it will be clear what the metric d is, and we shall abbreviate (X, d) as just X.

Remark 1.1.3 The conditions (a) and (b) can be rephrased as follows: for any $x, y \in X$ we have $d(x, y) = 0$ if and only if $x = y$. (Why is this equivalent to (a) and (b)?)

Example 1.1.4 (The real line) Let $\mathbf{R}$ be the real numbers, and let $d : \mathbf{R} \times \mathbf{R} \to [0, \infty)$ be the metric $d(x, y) := |x - y|$ mentioned earlier. Then $(\mathbf{R}, d)$ is a metric space (Exercise 1.1.2). We refer to d as the *standard metric* on $\mathbf{R}$, and if we refer to $\mathbf{R}$ as a metric space, we assume that the metric is given by the standard metric d unless otherwise specified.

Example 1.1.5 (Induced metric spaces) Let (X, d) be any metric space, and let Y be a subset of X. Then we can restrict the metric function $d : X \times X \to [0, +\infty)$ to the subset $Y \times Y$ of $X \times X$ to create a restricted metric function $d|_{Y \times Y} : Y \times Y \to [0, +\infty)$ of Y; this is known as the metric on Y *induced* by the metric d on X. The pair $(Y, d|_{Y \times Y})$ is a metric space (Exercise 1.1.4) and is known the *subspace* of (X, d) induced by Y. Thus for instance the metric on the real line in the previous example induces a metric space structure on any subset of the reals, such as the integers $\mathbf{Z}$, or an interval $[a, b]$.

Example 1.1.6 (Euclidean spaces) Let $n \geq 1$ be a natural number, and let $\mathbf{R}^n$ be the space of n-tuples of real numbers:

$$\mathbf{R}^n = \{(x_1, x_2, \ldots, x_n) : x_1, \ldots, x_n \in \mathbf{R}\}.$$

We define the *Euclidean metric* (also called the *l^2 metric*) $d_{l^2} : \mathbf{R}^n \times \mathbf{R}^n \to \mathbf{R}$ by

$$\begin{aligned} d_{l^2}((x_1, \ldots, x_n), (y_1, \ldots, y_n)) &:= \sqrt{(x_1 - y_1)^2 + \ldots + (x_n - y_n)^2} \\ &= \left(\sum_{i=1}^{n} (x_i - y_i)^2 \right)^{1/2}. \end{aligned}$$

Thus for instance, if $n = 2$, then $d_{l^2}((1, 6), (4, 2)) = \sqrt{3^2 + 4^2} = 5$. This metric corresponds to the geometric distance between the two points $(x_1, x_2, \ldots, x_n)$, $(y_1, y_2, \ldots, y_n)$ as given by Pythagoras' theorem. (We remark however that while geometry does give some very important examples of metric spaces, it is possible to have metric spaces which have no obvious geometry whatsoever. Some examples are given below.) The verification that $(\mathbf{R}^n, d)$ is indeed a metric space can be seen geometrically (for instance, the triangle inequality now asserts that the length of one side of a triangle is always less than or equal to the sum of the lengths of the other two sides), but can also be proven algebraically (see Exercise 1.1.6). We refer to $(\mathbf{R}^n, d_{l^2})$ as the *Euclidean space* of *dimension* n. Extending the convention from Example 1.1.4, if we refer to $\mathbf{R}^n$ as a metric space, we assume that the metric is given by the Euclidean metric unless otherwise specified.

Example 1.1.7 (Taxicab metric) Again let $n \geq 1$, and let $\mathbf{R}^n$ be as before. But now we use a different metric d_{l^1}, the so-called *taxicab metric* (or l^1 *metric*), defined by

$$\begin{aligned} d_{l^1}((x_1, x_2, \ldots, x_n), (y_1, y_2, \ldots, y_n)) &:= |x_1 - y_1| + \cdots + |x_n - y_n| \\ &= \sum_{i=1}^{n} |x_i - y_i|. \end{aligned}$$

Thus for instance, if $n = 2$, then $d_{l^1}((1, 6), (4, 2)) = 3 + 4 = 7$. This metric is called the taxicab metric, because it models the distance a taxicab would have to traverse to get from one point to another if the cab was only allowed to move in cardinal directions (north, south, east, west) and not diagonally. As such it is always at least as large as the Euclidean metric, which measures distance "as the crow flies", as it were. We claim that the space $(\mathbf{R}^n, d_{l^1})$ is also a metric space (Exercise 1.1.7). The metrics are not quite the same, but we do have the inequalities

$$d_{l^2}(x, y) \leq d_{l^1}(x, y) \leq \sqrt{n} d_{l^2}(x, y) \tag{1.1}$$

for all x, y (see Exercise 1.1.8).

Remark 1.1.8 The taxicab metric is useful in several places, for instance in the theory of error correcting codes. A string of n binary digits can be thought of as an element of $\mathbf{R}^n$, for instance the binary string 10010 can be thought of as the point $(1, 0, 0, 1, 0)$ in $\mathbf{R}^5$. The taxicab distance between two binary strings is then the number of bits in the two strings which do not match, for instance $d_{l^1}(10010, 10101) = 3$. The goal of error-correcting codes is to encode each piece of information (e.g., a letter of the alphabet) as a binary string in such a way that all the binary strings are as far away in the taxicab metric from each other as possible; this minimizes the chance that any distortion of the bits due to random noise can accidentally change one of the coded binary strings to another and also maximizes the chance that any such distortion can be detected and correctly repaired.

Example 1.1.9 (Sup norm metric) Again let $n \geq 1$, and let $\mathbf{R}^n$ be as before. But now we use a different metric d_{l^∞}, the so-called *sup norm metric* (or l^∞ *metric*), defined by

$$d_{l^\infty}((x_1, x_2, \ldots, x_n), (y_1, y_2, \ldots, y_n)) := \sup\{|x_i - y_i| : 1 \leq i \leq n\}.$$

Thus for instance, if $n = 2$, then $d_{l^\infty}((1, 6), (4, 2)) = \sup(3, 4) = 4$. The space $(\mathbf{R}^n, d_{l^\infty})$ is also a metric space (Exercise 1.1.9) and is related to the l^2 metric by the inequalities

$$\frac{1}{\sqrt{n}} d_{l^2}(x, y) \leq d_{l^\infty}(x, y) \leq d_{l^2}(x, y) \qquad (1.2)$$

for all x, y (see Exercise 1.1.10).

Remark 1.1.10 The l^1, l^2, and l^∞ metrics are special cases of the more general l^p *metrics*, where $p \in [1, +\infty]$, but we will not discuss these more general metrics in this text.

Example 1.1.11 (Discrete metric) Let X be an arbitrary set (finite or infinite), and define the *discrete metric* d_{disc} by setting $d_{\text{disc}}(x, y) := 0$ when $x = y$, and $d_{\text{disc}}(x, y) := 1$ when $x \neq y$. Thus, in this metric, all points are equally far apart. The space (X, d_{disc}) is a metric space (Exercise 1.1.11). Thus every set X has at least one metric on it.

Example 1.1.12 (Geodesics) (Informal) Let X be the sphere $\{(x, y, z) \in \mathbf{R}^3 : x^2 + y^2 + z^2 = 1\}$, and let $d((x, y, z), (x', y', z'))$ be the length of the shortest curve in X which starts at (x, y, z) and ends at (x', y', z'). (This curve turns out to be an arc of a great circle; we will not prove this here, as it requires *calculus of variations*, which is beyond the scope of this text.) This makes X into a metric space; the reader should be able to verify (without using any geometry of the sphere) that the triangle inequality is more or less automatic from the definition.

Example 1.1.13 (Shortest paths) (Informal) Examples of metric spaces occur all the time in real life. For instance, X could be all the computers currently connected to the internet, and $d(x, y)$ is the shortest number of connections it would take for a packet to travel from computer x to computer y; for instance, if x and y are not directly connected, but are both connected to z, then $d(x, y) = 2$. Assuming that all computers in the internet can ultimately be connected to all other computers (so that $d(x, y)$ is always finite), then (X, d) is a metric space (why?). Games such as "six degrees of separation" are also taking place in a similar metric space (what is the space, and what is the metric, in this case?). Or, X could be a major city, and $d(x, y)$ could be the shortest time it takes to drive from x to y (although this space might not satisfy axiom (c) in real life!).

Now that we have metric spaces, we can define convergence in these spaces.

Definition 1.1.14 (*Convergence of sequences in metric spaces*) Let m be an integer, (X, d) be a metric space, and let $(x^{(n)})_{n=m}^{\infty}$ be a sequence of points in X (i.e., for every natural number $n \geq m$, we assume that $x^{(n)}$ is an element of X). Let x be a point in X. We say that $(x^{(n)})_{n=m}^{\infty}$ *converges to x with respect to the metric d*, if and only if the limit $\lim_{n\to\infty} d(x^{(n)}, x)$ exists and is equal to 0. In other words, $(x^{(n)})_{n=m}^{\infty}$ converges to x with respect to d if and only if for every $\varepsilon > 0$, there exists an $N \geq m$ such that $d(x^{(n)}, x) \leq \varepsilon$ for all $n \geq N$. (Why are these two definitions equivalent?)

Remark 1.1.15 In view of Lemma 1.1.1 we see that this definition generalizes our existing notion of convergence of sequences of real numbers. In many cases, it is obvious what the metric d is, and so we shall often just say "$(x^{(n)})_{n=m}^{\infty}$ converges to x" instead of "$(x^{(n)})_{n=m}^{\infty}$ converges to x with respect to the metric d" when there is no chance of confusion. We also sometimes write "$x^{(n)} \to x$ as $n \to \infty$" instead.

Remark 1.1.16 There is nothing special about the superscript n in the above definition; it is a dummy variable. Saying that $(x^{(n)})_{n=m}^{\infty}$ converges to x is exactly the same statement as saying that $(x^{(k)})_{k=m}^{\infty}$ converges to x, for example; and sometimes it is convenient to change superscripts, for instance if the variable n is already being used for some other purpose. Similarly, it is not necessary for the sequence $x^{(n)}$ to be denoted using the superscript (n); the above definition is also valid for sequences x_n, or functions $f(n)$, or indeed of any expression which depends on n and takes values in X. Finally, from Exercises 6.1.3 and 6.1.4 we see that the starting point m of the sequence is unimportant for the purposes of taking limits; if $(x^{(n)})_{n=m}^{\infty}$ converges to x, then $(x^{(n)})_{n=m'}^{\infty}$ also converges to x for any $m' \geq m$.

Example 1.1.17 We work in the Euclidean space $\mathbf{R}^2$ with the standard Euclidean metric d_{l^2}. Let $(x^{(n)})_{n=1}^{\infty}$ denote the sequence $x^{(n)} := (1/n, 1/n)$ in $\mathbf{R}^2$, i.e., we are considering the sequence $(1, 1), (1/2, 1/2), (1/3, 1/3), \ldots$. Then this sequence converges to $(0, 0)$ with respect to the Euclidean metric d_{l^2}, since

$$\lim_{n\to\infty} d_{l^2}(x^{(n)}, (0, 0)) = \lim_{n\to\infty} \sqrt{\frac{1}{n^2} + \frac{1}{n^2}} = \lim_{n\to\infty} \frac{\sqrt{2}}{n} = 0.$$

The sequence $(x^{(n)})_{n=1}^{\infty}$ also converges to $(0, 0)$ with respect to the taxicab metric d_{l^1}, since

$$\lim_{n\to\infty} d_{l^1}(x^{(n)}, (0, 0)) = \lim_{n\to\infty} \frac{1}{n} + \frac{1}{n} = \lim_{n\to\infty} \frac{2}{n} = 0.$$

Similarly the sequence converges to $(0, 0)$ in the sup norm metric d_{l^∞} (why?). However, the sequence $(x^{(n)})_{n=1}^{\infty}$ does *not* converge to $(0, 0)$ in the discrete metric d_{disc}, since

$$\lim_{n\to\infty} d_{\text{disc}}(x^{(n)}, (0, 0)) = \lim_{n\to\infty} 1 = 1 \neq 0.$$

Thus the convergence of a sequence can depend on what metric one uses.[1]

[1] For a somewhat whimsical real-life example, one can give a city an "automobile metric", with $d(x, y)$ defined as the time it takes for a car to drive from x to y, or a "pedestrian metric", where $d(x, y)$ is the time it takes to walk on foot from x to y. (Let us assume for sake of argument that

In the case of the above four metrics—Euclidean, taxicab, sup norm, and discrete—it is in fact rather easy to test for convergence.

Proposition 1.1.18 (Equivalence of l^1, l^2, l^∞) *Let* $\mathbf{R}^n$ *be a Euclidean space, and let* $(x^{(k)})_{k=m}^\infty$ *be a sequence of points in* $\mathbf{R}^n$. *We write* $x^{(k)} = (x_1^{(k)}, x_2^{(k)}, \dots, x_n^{(k)})$, *i.e., for* $j = 1, 2, \dots, n$, $x_j^{(k)} \in \mathbf{R}$ *is the* j*th co-ordinate of* $x^{(k)} \in \mathbf{R}^n$. *Let* $x = (x_1, \dots, x_n)$ *be a point in* $\mathbf{R}^n$. *Then the following four statements are equivalent:*

(a) $(x^{(k)})_{k=m}^\infty$ *converges to* x *with respect to the Euclidean metric* d_{l^2}.
(b) $(x^{(k)})_{k=m}^\infty$ *converges to* x *with respect to the taxicab metric* d_{l^1}.
(c) $(x^{(k)})_{k=m}^\infty$ *converges to* x *with respect to the sup norm metric* d_{l^∞}.
(d) *For every* $1 \le j \le n$, *the sequence* $(x_j^{(k)})_{k=m}^\infty$ *converges to* x_j. *(Notice that this is a sequence of real numbers, not of points in* $\mathbf{R}^n$.*)*

Proof See Exercise 1.1.12. □

In other words, a sequence converges in the Euclidean, taxicab, or sup norm metric if and only if each of its components converges individually. Because of the equivalence of (a), (b), and (c), we say that the Euclidean, taxicab, and sup norm metrics on $\mathbf{R}^n$ are *equivalent*. (There are infinite-dimensional analogues of the Euclidean, taxicab, and sup norm metrics which are *not* equivalent, see for instance Exercise 1.1.15.)

For the discrete metric, convergence is much rarer: the sequence must be eventually constant in order to converge.

Proposition 1.1.19 (Convergence in the discrete metric) *Let* X *be any set, and let* d_{disc} *be the discrete metric on* X. *Let* $(x^{(n)})_{n=m}^\infty$ *be a sequence of points in* X, *and let* x *be a point in* X. *Then* $(x^{(n)})_{n=m}^\infty$ *converges to* x *with respect to the discrete metric* d_{disc} *if and only if there exists an* $N \ge m$ *such that* $x^{(n)} = x$ *for all* $n \ge N$.

Proof See Exercise 1.1.13. □

We now prove a basic fact about converging sequences; they can only converge to at most one point at a time.

Proposition 1.1.20 (Uniqueness of limits) *Let* (X, d) *be a metric space, and let* $(x^{(n)})_{n=m}^\infty$ *be a sequence in* X. *Suppose that there are two points* $x, x' \in X$ *such that* $(x^{(n)})_{n=m}^\infty$ *converges to* x *with respect to* d, *and* $(x^{(n)})_{n=m}^\infty$ *also converges to* x' *with respect to* d. *Then we have* $x = x'$.

Proof See Exercise 1.1.14. □

Because of the above proposition, it is safe to introduce the following notation: if $(x^{(n)})_{n=m}^\infty$ converges to x in the metric d, then we write $d - \lim_{n\to\infty} x^{(n)} = x$, or simply $\lim_{n\to\infty} x^{(n)} = x$ when there is no confusion as to what d is. For instance, in the example of $(\frac{1}{n}, \frac{1}{n})$, we have

---

these metrics are symmetric, though this is not always the case in real life.) One can easily imagine examples where two points are close in one metric but not another.

$$d_{l^2} - \lim_{n\to\infty}\left(\frac{1}{n}, \frac{1}{n}\right) = d_{l^1} - \lim_{n\to\infty}\left(\frac{1}{n}, \frac{1}{n}\right) = (0, 0),$$

but $d_{\mathrm{disc}} - \lim_{n\to\infty}(\frac{1}{n}, \frac{1}{n})$ is undefined. Thus the meaning of $d - \lim_{n\to\infty} x^{(n)}$ can depend on what d is; however Proposition 1.1.20 assures us that once d is fixed, there can be at most one value of $d - \lim_{n\to\infty} x^{(n)}$. (Of course, it is still possible that this limit does not exist; some sequences are not convergent.) Note that by Lemma 1.1.1, this definition of limit generalizes the notion of limit in Definition 6.1.8.

Remark 1.1.21 It is possible for a sequence to converge to one point using one metric, and another point using a different metric, although such examples are usually quite artificial. For instance, let $X := [0, 1]$, the closed interval from 0 to 1. Using the usual metric d, we have $d - \lim_{n\to\infty} \frac{1}{n} = 0$. But now suppose we "swap" the points 0 and 1 in the following manner. Let $f: [0, 1] \to [0, 1]$ be the function defined by $f(0) := 1$, $f(1) := 0$, and $f(x) := x$ for all $x \in (0, 1)$, and then define $d'(x, y) := d(f(x), f(y))$. Then (X, d') is still a metric space (why?), but now $d' - \lim_{n\to\infty} \frac{1}{n} = 1$. Thus changing the metric on a space can greatly affect the nature of convergence (also called the *topology*) on that space; see Sect. 2.5 for a further discussion of topology.

— Exercises —

Exercise 1.1.1 Prove Lemma 1.1.1.

Exercise 1.1.2 Show that the real line with the metric $d(x, y) := |x - y|$ is indeed a metric space. (*Hint:* you may wish to review your proof of Proposition 4.3.3.)

Exercise 1.1.3 Let X be a set, and let $d : X \times X \to [0, \infty)$ be a function.

(a) Give an example of a pair (X, d) which obeys axioms (bcd) of Definition 1.1.2, but not (a). (*Hint:* modify the discrete metric.)
(b) Give an example of a pair (X, d) which obeys axioms (acd) of Definition 1.1.2, but not (b).
(c) Give an example of a pair (X, d) which obeys axioms (abd) of Definition 1.1.2, but not (c).
(d) Give an example of a pair (X, d) which obeys axioms (abc) of Definition 1.1.2, but not (d). (*Hint:* try examples where X is a finite set.)

Exercise 1.1.4 Show that the pair $(Y, d|_{Y\times Y})$ defined in Example 1.1.5 is indeed a metric space.

Exercise 1.1.5 Let $n \geq 1$, and let $a_1, a_2, \ldots, a_n$ and $b_1, b_2, \ldots, b_n$ be real numbers. Verify the identity

$$\left(\sum_{i=1}^{n} a_i b_i\right)^2 + \frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n}(a_i b_j - a_j b_i)^2 = \left(\sum_{i=1}^{n} a_i^2\right)\left(\sum_{j=1}^{n} b_j^2\right),$$

and conclude the *Cauchy–Schwarz inequality*

$$\left|\sum_{i=1}^{n} a_i b_i\right| \leq \left(\sum_{i=1}^{n} a_i^2\right)^{1/2} \left(\sum_{j=1}^{n} b_j^2\right)^{1/2}. \tag{1.3}$$

Then use the Cauchy–Schwarz inequality to prove the *triangle inequality*

$$\left(\sum_{i=1}^{n} (a_i + b_i)^2\right)^{1/2} \leq \left(\sum_{i=1}^{n} a_i^2\right)^{1/2} + \left(\sum_{j=1}^{n} b_j^2\right)^{1/2}.$$

Exercise 1.1.6 Show that $(\mathbf{R}^n, d_{l^2})$ in Example 1.1.6 is indeed a metric space. (*Hint:* use Exercise 1.1.5.)

Exercise 1.1.7 Show that the pair $(\mathbf{R}^n, d_{l^1})$ in Example 1.1.7 is indeed a metric space.

Exercise 1.1.8 Prove the two inequalities in (1.1). (*Hint:* For the first inequality, square both sides. For the second inequality, use Exercise (1.1.5).)

Exercise 1.1.9 Show that the pair $(\mathbf{R}^n, d_{l^\infty})$ in Example 1.1.9 is indeed a metric space.

Exercise 1.1.10 Prove the two inequalities in (1.2).

Exercise 1.1.11 Show that the discrete metric (X, d_{disc}) in Example 1.1.11 is indeed a metric space.

Exercise 1.1.12 Prove Proposition 1.1.18.

Exercise 1.1.13 Prove Proposition 1.1.19.

Exercise 1.1.14 Prove Proposition 1.1.20. (*Hint:* modify the proof of Proposition 6.1.7.)

Exercise 1.1.15 Let

$$X := \left\{ (a_n)_{n=0}^{\infty} : \sum_{n=0}^{\infty} |a_n| < \infty \right\}$$

be the space of absolutely convergent sequences. Define the l^1 and l^∞ metrics on this space by

$$d_{l^1}((a_n)_{n=0}^{\infty}, (b_n)_{n=0}^{\infty}) := \sum_{n=0}^{\infty} |a_n - b_n|;$$

$$d_{l^\infty}((a_n)_{n=0}^{\infty}, (b_n)_{n=0}^{\infty}) := \sup_{n \in \mathbf{N}} |a_n - b_n|.$$

Show that these are both metrics on X, but show that there exist sequences $x^{(1)}, x^{(2)}, \ldots$ of elements of X (i.e., sequences of sequences) which are convergent with respect to the d_{l^∞} metric but not with respect to the d_{l^1} metric. Conversely, show that any sequence which converges in the d_{l^1} metric automatically converges in the d_{l^∞} metric.

Exercise 1.1.16 Let $(x_n)_{n=1}^\infty$ and $(y_n)_{n=1}^\infty$ be two sequences in a metric space (X, d). Suppose that $(x_n)_{n=1}^\infty$ converges to a point $x \in X$, and $(y_n)_{n=1}^\infty$ converges to a point $y \in X$. Show that $\lim_{n\to\infty} d(x_n, y_n) = d(x, y)$. (*Hint:* use the triangle inequality several times.)

1.2 Some Point-Set Topology of Metric Spaces

Having defined the operation of convergence on metric spaces, we now define a couple other related notions, including that of open set, closed set, interior, exterior, boundary, and adherent point. The study of such notions is known as *point-set topology*, which we shall return to in Sect. 2.5.

We first need the notion of a *metric ball*, or more simply a *ball*.

Definition 1.2.1 (*Balls*) Let (X, d) be a metric space, let x_0 be a point in X, and let $r > 0$. We define the *ball* $B_{(X,d)}(x_0, r)$ in X, centered at x_0, and with radius r, in the metric d, to be the set

$$B_{(X,d)}(x_0, r) := \{x \in X : d(x, x_0) < r\}.$$

When it is clear what the metric space (X, d) is, we shall abbreviate $B_{(X,d)}(x_0, r)$ as just $B(x_0, r)$.

Example 1.2.2 In $\mathbf{R}^2$ with the Euclidean metric d_{l^2}, the ball $B_{(\mathbf{R}^2,d_{l^2})}((0, 0), 1)$ is the open disc

$$B_{(\mathbf{R}^2,d_{l^2})}((0, 0), 1) = \{(x, y) \in \mathbf{R}^2 : x^2 + y^2 < 1\}.$$

However, if one uses the taxicab metric d_{l^1} instead, then we obtain a diamond:

$$B_{(\mathbf{R}^2,d_{l^1})}((0, 0), 1) = \{(x, y) \in \mathbf{R}^2 : |x| + |y| < 1\}.$$

If we use the discrete metric, the ball is now reduced to a single point:

$$B_{(\mathbf{R}^2,d_{\text{disc}})}((0, 0), 1) = \{(0, 0)\},$$

although if one increases the radius to be larger than 1, then the ball now encompasses all of $\mathbf{R}^2$. (Why?)

Example 1.2.3 In $\mathbf{R}$ with the usual metric d, the open interval $(3, 7)$ is also the metric ball $B_{(\mathbf{R},d)}(5, 2)$.

Remark 1.2.4 Note that the smaller the radius r, the smaller the ball $B(x_0, r)$. However, $B(x_0, r)$ always contains at least one point, namely the center x_0, as long as r stays positive, thanks to Definition 1.1.2(a). (We don't consider balls of zero radius or negative radius since they are rather boring, being just the empty set.)

Using metric balls, one can now take a set E in a metric space X and classify three types of points in X: interior, exterior, and boundary points of E.

Definition 1.2.5 (*Interior, exterior, boundary*) Let (X, d) be a metric space, let E be a subset of X, and let x_0 be a point in X. We say that x_0 is an *interior point of* E if there exists a radius $r > 0$ such that $B(x_0, r) \subseteq E$. We say that x_0 is an *exterior point of* E if there exists a radius $r > 0$ such that $B(x_0, r) \cap E = \emptyset$. We say that x_0 is a *boundary point of* E if it is neither an interior point nor an exterior point of E.

The set of all interior points of E is called the *interior* of E and is sometimes denoted $\text{int}(E)$. The set of exterior points of E is called the *exterior* of E and is sometimes denoted $\text{ext}(E)$. The set of boundary points of E is called the *boundary* of E and is sometimes denoted ∂E.

Remark 1.2.6 If x_0 is an interior point of E, then x_0 must actually be an element of E, since balls $B(x_0, r)$ always contain their center x_0. Conversely, if x_0 is an exterior point of E, then x_0 cannot be an element of E. In particular it is not possible for x_0 to simultaneously be an interior and an exterior point of E. If x_0 is a boundary point of E, then it could be an element of E, but it could also not lie in E; we give some examples below.

Example 1.2.7 We work on the real line $\mathbf{R}$ with the standard metric d. Let E be the half-open interval $E = [1, 2)$. The point 1.5 is an interior point of E, since one can find a ball (for instance $B(1.5, 0.1)$) centered at 1.5 which lies in E. The point 3 is an exterior point of E, since one can find a ball (for instance $B(3, 0.1)$) centered at 3 which is disjoint from E. The points 1 and 2, however, are neither interior points nor exterior points of E and are thus boundary points of E. Thus in this case $\text{int}(E) = (1, 2)$, $\text{ext}(E) = (-\infty, 1) \cup (2, \infty)$, and $\partial E = \{1, 2\}$. Note that in this case one of the boundary points is an element of E, while the other is not.

Example 1.2.8 When we give a set X the discrete metric d_{disc}, and E is any subset of X, then every element of E is an interior point of E, every point not contained in E is an exterior point of E, and there are no boundary points; see Exercise 1.2.1.

Definition 1.2.9 (*Closure*) Let (X, d) be a metric space, let E be a subset of X, and let x_0 be a point in X. We say that x_0 is an *adherent point* of E if for every radius $r > 0$, the ball $B(x_0, r)$ has a non-empty intersection with E. The set of all adherent points of E is called the *closure* of E and is denoted $\overline{E}$.

Note that these notions are consistent with the corresponding notions on the real line defined in Definitions 9.1.8 and 9.1.10 (why?).

The following proposition links the notions of adherent point with interior and boundary point and also to that of convergence.

Proposition 1.2.10 *Let (X, d) be a metric space, let E be a subset of X, and let x_0 be a point in X. Then the following statements are logically equivalent.*

(a) *x_0 is an adherent point of E.*
(b) *x_0 is either an interior point or a boundary point of E.*
(c) *There exists a sequence $(x_n)_{n=1}^{\infty}$ in E which converges to x_0 with respect to the metric d.*

Proof See Exercise 1.2.2. □

From the equivalence of Proposition 1.2.10(a) and (b) we obtain an immediate corollary:

Corollary 1.2.11 *Let (X, d) be a metric space, and let E be a subset of X. Then* $\overline{E} = \text{int}(E) \cup \partial E = X \backslash \text{ext}(E)$.

As remarked earlier, the boundary of a set E may or may not lie in E. Depending on how the boundary is situated, we may call a set open, closed, or neither:

Definition 1.2.12 (*Open and closed sets*) Let (X, d) be a metric space, and let E be a subset of X. We say that E is *closed* if it contains all of its boundary points, i.e., $\partial E \subseteq E$. We say that E is *open* if it contains none of its boundary points, i.e., $\partial E \cap E = \emptyset$. If E contains some of its boundary points but not others, then it is neither open nor closed.

Example 1.2.13 We work in the real line $\mathbf{R}$ with the standard metric d. The set $(1, 2)$ does not contain either of its boundary points 1, 2 and is hence open. The set $[1, 2]$ contains both of its boundary points 1, 2 and is hence closed. The set $[1, 2)$ contains one of its boundary points 1, but does not contain the other boundary point 2, so is neither open nor closed.

Remark 1.2.14 It is possible for a set to be simultaneously open and closed, if it has no boundary. For instance, in a metric space (X, d), the whole space X has no boundary (every point in X is an interior point—why?), and so X is both open and closed. The empty set $\emptyset$ also has no boundary (every point in X is an exterior point—why?), and so $\emptyset$ is both open and closed. In many cases these are the only sets that are simultaneously open and closed, but there are exceptions. For instance, using the discrete metric d_{disc}, *every* set is both open and closed! (why?)

From the above two remarks we see that the notions of being open and being closed are *not* negations of each other; there are sets that are both open and closed, and there are sets which are neither open nor closed. Thus, if one knew for instance that E was not an open set, it would be erroneous to conclude from this that E was a closed set, and similarly with the rôles of open and closed reversed. The correct relationship between open and closed sets is given by Proposition 1.2.15(e) below.

Now we list some more properties of open and closed sets.

Proposition 1.2.15 (Basic properties of open and closed sets) *Let (X, d) be a metric space.*

(a) *Let E be a subset of X. Then E is open if and only if $E = \text{int}(E)$. In other words, E is open if and only if for every $x \in E$, there exists an $r > 0$ such that $B(x, r) \subseteq E$.*

(b) *Let E be a subset of X. Then E is closed if and only if E contains all its adherent points. In other words, E is closed if and only if for every convergent sequence $(x_n)_{n=m}^{\infty}$ in E, the limit $\lim_{n\to\infty} x_n$ of that sequence also lies in E.*

(c) *For any $x_0 \in X$ and $r > 0$, then the ball $B(x_0, r)$ is an open set. The set $\{x \in X : d(x, x_0) \le r\}$ is a closed set. (This set is sometimes called the* closed ball *of radius r centered at x_0.)*

(d) *Any singleton set $\{x_0\}$, where $x_0 \in X$, is automatically closed.*

(e) *If E is a subset of X, then E is open if and only if the complement $X\backslash E := \{x \in X : x \notin E\}$ is closed.*

(f) *If $E_1, \dots, E_n$ is a finite collection of open sets in X, then $E_1 \cap E_2 \cap \dots \cap E_n$ is also open. If $F_1, \dots, F_n$ is a finite collection of closed sets in X, then $F_1 \cup F_2 \cup \dots \cup F_n$ is also closed.*

(g) *If $\{E_\alpha\}_{\alpha\in I}$ is a collection of open sets in X (where the index set I could be finite, countable, or uncountable), then the union $\bigcup_{\alpha\in I} E_\alpha := \{x \in X : x \in E_\alpha$ for some $\alpha \in I\}$ is also open. If $\{F_\alpha\}_{\alpha\in I}$ is a collection of closed sets in X, then the intersection $\bigcap_{\alpha\in I} F_\alpha := \{x \in X : x \in F_\alpha$ for all $\alpha \in I\}$ is also closed.*

(h) *If E is any subset of X, then $\text{int}(E)$ is the largest open set which is contained in E; in other words, $\text{int}(E)$ is open, and given any other open set $V \subseteq E$, we have $V \subseteq \text{int}(E)$. Similarly $\overline{E}$ is the smallest closed set which contains E; in other words, $\overline{E}$ is closed, and given any other closed set $K \supset E$, $K \supseteq \overline{E}$.*

Proof See Exercise 1.2.3. □

— Exercises —

Exercise 1.2.1 Verify the claims in Example 1.2.8.

Exercise 1.2.2 Prove Proposition 1.2.10. (*Hint:* for some of the implications one will need the axiom of choice, as in Lemma 8.4.5.)

Exercise 1.2.3 Prove Proposition 1.2.15. (*Hint:* you can use earlier parts of the proposition to prove later ones.)

Exercise 1.2.4 Let (X, d) be a metric space, x_0 be a point in X, and $r > 0$. Let B be the open ball $B := B(x_0, r) = \{x \in X : d(x, x_0) < r\}$, and let C be the closed ball $C := \{x \in X : d(x, x_0) \le r\}$.

(a) Show that $\overline{B} \subseteq C$.

(b) Give an example of a metric space (X, d), a point x_0, and a radius $r > 0$ such that $\overline{B}$ is *not* equal to C.

1.3 Relative Topology

When we defined notions such as open and closed sets, we mentioned that such concepts depended on the choice of metric one uses. For instance, on the real line **R**, if one uses the usual metric $d(x, y) = |x - y|$, then the set $\{1\}$ is not open, however if instead one uses the discrete metric d_{disc}, then $\{1\}$ is now an open set (why?).

However, it is not just the choice of metric which determines what is open and what is not—it is also the choice of *ambient space* X. Here are some examples.

Example 1.3.1 Consider the plane $\mathbf{R}^2$ with the Euclidean metric d_{l^2}. Inside the plane, we can find the x-axis $X := \{(x, 0) : x \in \mathbf{R}\}$. The metric d_{l^2} can be restricted to X, creating a subspace $(X, d_{l^2}|_{X\times X})$ of $(\mathbf{R}^2, d_{l^2})$. (This subspace is essentially the same as the real line $(\mathbf{R}, d)$ with the usual metric; the precise way of stating this is that $(X, d_{l^2}|_{X\times X})$ is *isometric* to $(\mathbf{R}, d)$. We will not pursue this concept further in this text, however.) Now consider the set

$$E := \{(x, 0) : -1 < x < 1\}$$

which is both a subset of X and of $\mathbf{R}^2$. Viewed as a subset of $\mathbf{R}^2$, it is not open, because the point $(0, 0)$, for instance, lies in E but is not an interior point of E. (Any ball $B_{\mathbf{R}^2,d_{l^2}}(0, r)$ will contain at least one point that lies outside of the x-axis, and hence outside of E.) On the other hand, if viewed as a subset of X, it is open; every point of E is an interior point of E *with respect to the metric space* $(X, d_{l^2}|_{X\times X})$. For instance, the point $(0, 0)$ is now an interior point of E, because the ball $B_{X,d_{l^2}|_{X\times X}}(0, 1)$ is contained in E (in fact, in this case it *is* E).

Example 1.3.2 Consider the real line **R** with the standard metric d, and let X be the interval $X := (-1, 1)$ contained inside **R**; we can then restrict the metric d to X, creating a subspace $(X, d|_{X\times X})$ of $(\mathbf{R}, d)$. Now consider the set $[0, 1)$. This set is not closed in **R**, because the point 1 is adherent to $[0, 1)$ but is not contained in $[0, 1)$. However, when considered as a subset of X, the set $[0, 1)$ now becomes closed; the point 1 is not an element of X and so is no longer considered an adherent point of $[0, 1)$, and so now $[0, 1)$ contains all of its adherent points.

To clarify this distinction, we make a definition.

Definition 1.3.3 (*Relative topology*) Let (X, d) be a metric space, let Y be a subset of X, and let E be a subset of Y. We say that E is *relatively open with respect to* Y if it is open in the metric subspace $(Y, d|_{Y\times Y})$. Similarly, we say that E is *relatively closed with respect to* Y if it is closed in the metric space $(Y, d|_{Y\times Y})$.

The relationship between open (or closed) sets in X, and relatively open (or relatively closed) sets in Y, is the following.

Proposition 1.3.4 *Let (X, d) be a metric space, let Y be a subset of X, and let E be a subset of Y.*

(a) *E is relatively open with respect to Y if and only if $E = V \cap Y$ for some set $V \subseteq X$ which is open in X.*
(b) *E is relatively closed with respect to Y if and only if $E = K \cap Y$ for some set $K \subseteq X$ which is closed in X.*

Proof We just prove (a) and leave (b) to Exercise 1.3.1. First suppose that E is relatively open with respect to Y. Then, E is open in the metric space $(Y, d|_{Y\times Y})$. Thus, for every $x \in E$, there exists a radius $r > 0$ such that the ball $B_{(Y,d|_{Y\times Y})}(x, r)$ is contained in E. This radius r depends on x; to emphasize this we write r_x instead of r, thus for every $x \in E$ the ball $B_{(Y,d|_{Y\times Y})}(x, r_x)$ is contained in E. (Note that we have used the axiom of choice, Proposition 8.4.7, to do this.)

Now consider the set

$$V := \bigcup_{x\in E} B_{(X,d)}(x, r_x).$$

This is a subset of X. By Proposition 1.2.15(c) and (g), V is open. Now we prove that $E = V \cap Y$. Certainly any point x in E lies in $V \cap Y$, since it lies in Y and it also lies in $B_{(X,d)}(x, r_x)$, and hence in V. Now suppose that y is a point in $V \cap Y$. Then $y \in V$, which implies that there exists an $x \in E$ such that $y \in B_{(X,d)}(x, r_x)$. But since y is also in Y, this implies that $y \in B_{(Y,d|_{Y\times Y})}(x, r_x)$. But by definition of r_x, this means that $y \in E$, as desired. Thus we have found an open set V for which $E = V \cap Y$ as desired.

Now we do the converse. Suppose that $E = V \cap Y$ for some open set V; we have to show that E is relatively open with respect to Y. Let x be any point in E; we have to show that x is an interior point of E in the metric space $(Y, d|_{Y\times Y})$. Since $x \in E$, we know $x \in V$. Since V is open in X, we know that there is a radius $r > 0$ such that $B_{(X,d)}(x, r)$ is contained in V. Strictly speaking, r depends on x, and so we could write r_x instead of r, but for this argument we will only use a single choice of x (as opposed to the argument in the previous paragraph) and so we will not bother to subscript r here. Since $E = V \cap Y$, this means that $B_{(X,d)}(x, r) \cap Y$ is contained in E. But $B_{(X,d)}(x, r) \cap Y$ is exactly the same as $B_{(Y,d|_{Y\times Y})}(x, r)$ (why?), and so $B_{(Y,d|_{Y\times Y})}(x, r)$ is contained in E. Thus x is an interior point of E in the metric space $(Y, d|_{Y\times Y})$, as desired. □

— Exercises —

Exercise 1.3.1 Prove Proposition 1.3.4(b).

1.4 Cauchy Sequences and Complete Metric Spaces

We now generalize much of the theory of limits of sequences from Chap. 6 to the setting of general metric spaces. We begin by generalizing the notion of a *subsequence* from Definition 6.6.1:

Definition 1.4.1 (*Subsequences*) Suppose that $(x^{(n)})_{n=m}^{\infty}$ is a sequence of points in a metric space (X, d). Suppose that $n_1, n_2, n_3, \ldots$ is an increasing sequence of integers which are at least as large as m, thus

$$m \le n_1 < n_2 < n_3 < \cdots .$$

Then we call the sequence $(x^{(n_j)})_{j=1}^{\infty}$ a *subsequence* of the original sequence $(x^{(n)})_{n=m}^{\infty}$.

Example 1.4.2 The sequence $\left(\left(\frac{1}{j^2}, \frac{1}{j^2}\right)\right)_{j=1}^{\infty}$ in $\mathbf{R}^2$ is a subsequence of the sequence $\left(\left(\frac{1}{n}, \frac{1}{n}\right)\right)_{n=1}^{\infty}$ (in this case, $n_j := j^2$). The sequence $1, 1, 1, 1, \ldots$ is a subsequence of $1, 0, 1, 0, 1, \ldots$.

If a sequence converges, then so do all of its subsequences:

Lemma 1.4.3 *Let $(x^{(n)})_{n=m}^{\infty}$ be a sequence in (X, d) which converges to some limit x_0. Then every subsequence $(x^{(n_j)})_{j=1}^{\infty}$ of that sequence also converges to x_0.*

Proof See Exercise 1.4.1. □

On the other hand, it is possible for a subsequence to be convergent without the sequence as a whole being convergent. For example, the sequence $1, 0, 1, 0, 1, \ldots$ is not convergent, even though certain subsequences of it (such as $1, 1, 1, \ldots$) converge. To quantify this phenomenon, we generalize Definition 6.4.1 as follows:

Definition 1.4.4 (*Limit points*) Suppose that $(x^{(n)})_{n=m}^{\infty}$ is a sequence of points in a metric space (X, d), and let $L \in X$. We say that L is a *limit point* of $(x^{(n)})_{n=m}^{\infty}$ iff for every $N \ge m$ and $\varepsilon > 0$ there exists an $n \ge N$ such that $d(x^{(n)}, L) \le \varepsilon$.

Proposition 1.4.5 *Let $(x^{(n)})_{n=m}^{\infty}$ be a sequence of points in a metric space (X, d), and let $L \in X$. Then the following are equivalent:*

- *L is a limit point of $(x^{(n)})_{n=m}^{\infty}$.*
- *There exists a subsequence $(x^{(n_j)})_{j=1}^{\infty}$ of the original sequence $(x^{(n)})_{n=m}^{\infty}$ which converges to L.*

Proof See Exercise 1.4.2. □

Next, we review the notion of a *Cauchy sequence* from Definition 6.1.3 (see also Definition 5.1.8).

Definition 1.4.6 (*Cauchy sequences*) Let $(x^{(n)})_{n=m}^{\infty}$ be a sequence of points in a metric space (X, d). We say that this sequence is a *Cauchy sequence* iff for every $\varepsilon > 0$, there exists an $N \ge m$ such that $d(x^{(j)}, x^{(k)}) < \varepsilon$ for all $j, k \ge N$.

Lemma 1.4.7 (Convergent sequences are Cauchy sequences) *Let $(x^{(n)})_{n=m}^{\infty}$ be a sequence in (X, d) which converges to some limit x_0. Then $(x^{(n)})_{n=m}^{\infty}$ is also a Cauchy sequence.*

Proof See Exercise 1.4.3. □

It is also easy to check that subsequence of a Cauchy sequence is also a Cauchy sequence (why?). However, not every Cauchy sequence converges:

Example 1.4.8 (Informal) Consider the sequence

$$3, 3.1, 3.14, 3.141, 3.1415, \ldots$$

in the metric space $(\mathbf{Q}, d)$ (the rationals $\mathbf{Q}$ with the usual metric $d(x, y) := |x - y|$). While this sequence is convergent in $\mathbf{R}$ (it converges to π), it does not converge in $\mathbf{Q}$ (since $\pi \notin \mathbf{Q}$, and a sequence cannot converge to two different limits).

So in certain metric spaces, Cauchy sequences do not necessarily converge. However, if even part of a Cauchy sequence converges, then the entire Cauchy sequence must converge (to the same limit):

Lemma 1.4.9 *Let $(x^{(n)})_{n=m}^{\infty}$ be a Cauchy sequence in (X, d). Suppose that there is some subsequence $(x^{(n_j)})_{j=1}^{\infty}$ of this sequence which converges to a limit x_0 in X. Then the original sequence $(x^{(n)})_{n=m}^{\infty}$ also converges to x_0.*

Proof See Exercise 1.4.4. □

In Example 1.4.8 we saw an example of a metric space which contained Cauchy sequences which did not converge. However, in Theorem 6.4.18 we saw that in the metric space $(\mathbf{R}, d)$, every Cauchy sequence did have a limit. This motivates the following definition.

Definition 1.4.10 (*Complete metric spaces*) A metric space (X, d) is said to be *complete* iff every Cauchy sequence in (X, d) is in fact convergent in (X, d).

Example 1.4.11 By Theorem 6.4.18, the reals $(\mathbf{R}, d)$ are complete; by Example 1.4.8, the rationals $(\mathbf{Q}, d)$, on the other hand, are not complete.

Complete metric spaces have some nice properties. For instance, they are *intrinsically closed*: no matter what space one places them in, they are always closed sets. More precisely:

Proposition 1.4.12 *(a) Let (X, d) be a metric space, and let $(Y, d|_{Y\times Y})$ be a subspace of (X, d). If $(Y, d|_{Y\times Y})$ is complete, then Y must be closed in X.*
(b) Conversely, suppose that (X, d) is a complete metric space, and Y is a closed subset of X. Then the subspace $(Y, d|_{Y\times Y})$ is also complete.

Proof See Exercise 1.4.7. □

In contrast, an incomplete metric space such as $(\mathbf{Q}, d)$ may be considered closed in some spaces (for instance, $\mathbf{Q}$ is closed in $\mathbf{Q}$) but not in others (for instance, $\mathbf{Q}$ is not closed in $\mathbf{R}$). Indeed, it turns out that given any incomplete metric space (X, d), there exists a *completion* $(\overline{X}, \overline{d})$, which is a larger metric space containing (X, d) which is complete, and such that X is not closed in $\overline{X}$ (indeed, the closure of X in $(\overline{X}, \overline{d})$ will be all of $\overline{X}$); see Exercise 1.4.8. For instance, one possible completion of **Q** is **R**.

— Exercises —

Exercise 1.4.1 Prove Lemma 1.4.3. (*Hint:* review your proof of Proposition 6.6.5.)

Exercise 1.4.2 Prove Proposition 1.4.5. (*Hint:* review your proof of Proposition 6.6.6.)

Exercise 1.4.3 Prove Lemma 1.4.7. (*Hint:* review your proof of Proposition 6.1.12.)

Exercise 1.4.4 Prove Lemma 1.4.9.

Exercise 1.4.5 Let $(x^{(n)})_{n=m}^{\infty}$ be a sequence of points in a metric space (X, d), and let $L \in X$. Show that if L is a limit point of the sequence $(x^{(n)})_{n=m}^{\infty}$, then L is an adherent point of the set $\{x^{(n)} : n \geq m\}$. Is the converse true?

Exercise 1.4.6 Show that every Cauchy sequence can have at most one limit point.

Exercise 1.4.7 Prove Proposition 1.4.12.

Exercise 1.4.8 The following construction generalizes the construction of the reals from the rationals in Chap. 5, allowing one to view any metric space as a subspace of a complete metric space. In what follows we let (X, d) be a metric space.

(a) Given any Cauchy sequence $(x_n)_{n=1}^{\infty}$ in X, we introduce the *formal limit* $\mathrm{LIM}_{n\to\infty} x_n$. We say that two formal limits $\mathrm{LIM}_{n\to\infty} x_n$ and $\mathrm{LIM}_{n\to\infty} y_n$ are equal if $\lim_{n\to\infty} d(x_n, y_n)$ is equal to zero. Show that this equality relation obeys the reflexive, symmetry, and transitive axioms.

(b) Let $\overline{X}$ be the space of all formal limits of Cauchy sequences in X, with the above equality relation. Define a metric $d_{\overline{X}} \colon \overline{X} \times \overline{X} \to [0, +\infty)$ by setting

$$d_{\overline{X}}(\mathrm{LIM}_{n\to\infty} x_n, \mathrm{LIM}_{n\to\infty} y_n) := \lim_{n\to\infty} d(x_n, y_n).$$

Show that this function is well-defined (this means not only that the limit $\lim_{n\to\infty} d(x_n, y_n)$ exists, but also that the axiom of substitution is obeyed; cf. Lemma 5.3.7) and gives $\overline{X}$ the structure of a metric space.

(c) Show that the metric space $(\overline{X}, d_{\overline{X}})$ is complete.

(d) We identify an element $x \in X$ with the corresponding formal limit $\mathrm{LIM}_{n\to\infty} x$ in $\overline{X}$; show that this is legitimate by verifying that $x = y \iff \mathrm{LIM}_{n\to\infty} x = \mathrm{LIM}_{n\to\infty} y$. With this identification, show that $d(x, y) = d_{\overline{X}}(x, y)$, and thus (X, d) can now be thought of as a subspace of $(\overline{X}, d_{\overline{X}})$.

(e) Show that the closure of X in $\overline{X}$ is $\overline{X}$ (which explains the choice of notation $\overline{X}$).

(f) Show that the formal limit agrees with the actual limit, thus if $(x_n)_{n=1}^{\infty}$ is any Cauchy sequence in X, then we have $\lim_{n\to\infty} x_n = \mathrm{LIM}_{n\to\infty} x_n$ in $\overline{X}$.

1.5 Compact Metric Spaces

We now come to one of the most useful notions in point-set topology, that of *compactness*. Recall the Heine–Borel theorem (Theorem 9.1.24), which asserted that every sequence in a closed and bounded subset X of the real line $\mathbf{R}$ had a convergent subsequence whose limit was also in X. Conversely, only the closed and bounded sets have this property. This property turns out to be so useful that we give it a name.

Definition 1.5.1 (*Compactness*) A metric space (X, d) is said to be *compact* iff every sequence in (X, d) has at least one convergent subsequence. A subset Y of a metric space X is said to be *compact* if the subspace $(Y, d|_{Y\times Y})$ is compact.

Remark 1.5.2 The notion of a set Y being compact is *intrinsic*, in the sense that it only depends on the metric function $d|_{Y\times Y}$ restricted to Y, and not on the choice of the ambient space X. The notions of completeness in Definition 1.4.10, and of boundedness below in Definition 1.5.3, are also intrinsic, but the notions of open and closed are not (see the discussion in Sect. 1.3).

Thus, Theorem 9.1.24 shows that in the real line $\mathbf{R}$ with the usual metric, every closed and bounded set is compact, and conversely every compact set is closed and bounded.

Now we investigate how the Heine–Borel extends to other metric spaces.

Definition 1.5.3 (*Bounded sets*) Let (X, d) be a metric space, and let Y be a subset of X. We say that Y is *bounded* iff for every $x \in X$ there exists a ball $B(x, r)$ in X of some finite radius r which contains Y. We call the metric space (X, d) bounded if X is bounded.

Remark 1.5.4 This definition is compatible with the definition of a bounded set in Definition 9.1.22 (Exercise 1.5.1).

Proposition 1.5.5 *Let (X, d) be a compact metric space. Then (X, d) is both complete and bounded.*

Proof See Exercise 1.5.2. □

From this proposition and Proposition 1.4.12(a) we obtain one half of the Heine–Borel theorem for general metric spaces:

Corollary 1.5.6 (Compact sets are closed and bounded) *Let (X, d) be a metric space, and let Y be a compact subset of X. Then Y is closed and bounded.*

The other half of the Heine–Borel theorem is true in Euclidean spaces:

Theorem 1.5.7 (Heine–Borel theorem) *Let $(\mathbf{R}^n, d)$ be a Euclidean space with either the Euclidean metric, the taxicab metric, or the sup norm metric. Let E be a subset of $\mathbf{R}^n$. Then E is compact if and only if it is closed and bounded.*

Proof See Exercise 1.5.3. □

However, the Heine–Borel theorem is not true for more general metrics. For instance, the integer **Z** with the discrete metric is closed (indeed, it is complete) and bounded, but not compact, since the sequence $1, 2, 3, 4, \ldots$ is in **Z** but has no convergent subsequence (why?). Another example is in Exercise 1.5.8. However, a version of the Heine–Borel theorem is available if one is willing to replace closedness with the stronger notion of completeness and boundedness with the stronger notion of *total boundedness*; see Exercise 1.5.10.

One can characterize compactness topologically via the following, rather strange-sounding statement: every open cover of a compact set has a finite subcover.

Theorem 1.5.8 *Let (X, d) be a metric space, and let Y be a compact subset of X. Let $(V_\alpha)_{\alpha \in A}$ be a collection of open sets in X, and suppose that*

$$Y \subseteq \bigcup_{\alpha \in A} V_\alpha.$$

(i.e., the collection $(V_\alpha)_{\alpha \in A}$ covers *Y). Then there exists a* finite *subset F of A such that*

$$Y \subseteq \bigcup_{\alpha \in F} V_\alpha.$$

Proof We assume for sake of contradiction that there does not exist any finite subset F of A for which $Y \subseteq \bigcup_{\alpha \in F} V_\alpha$.

Let y be any element of Y. Then y must lie in at least one of the sets V_α. Since each V_α is open, there must therefore be an $r > 0$ such that $B_{(X,d)}(y, r) \subseteq V_\alpha$. Now let $r(y)$ denote the quantity

$$r(y) := \sup\{r \in (0, \infty) : B_{(X,d)}(y, r) \subseteq V_\alpha \text{ for some } \alpha \in A\}.$$

By the above discussion, we know that $r(y) > 0$ for all $y \in Y$. Now, let r_0 denote the quantity

$$r_0 := \inf\{r(y) : y \in Y\}.$$

Since $r(y) > 0$ for all $y \in Y$, we have $r_0 \geq 0$. There are three cases: $r_0 = 0$, $0 < r_0 < \infty$, and $r_0 = \infty$.

- **Case 1: $r_0 = 0$.** Then for every integer $n \geq 1$, there is at least one point y in Y such that $r(y) < 1/n$ (why?). We thus choose, for each $n \geq 1$, a point $y^{(n)}$ in Y such that $r(y^{(n)}) < 1/n$ (we can do this because of the axiom of choice, see Proposition 8.4.7). In particular we have $\lim_{n\to\infty} r(y^{(n)}) = 0$, by the squeeze test. The sequence $(y^{(n)})_{n=1}^\infty$ is a sequence in Y; since Y is compact, we can thus find a subsequence $(y^{(n_j)})_{j=1}^\infty$ which converges to a point $y_0 \in Y$.
  As before, we know that there exists some $\alpha \in A$ such that $y_0 \in V_\alpha$, and hence (since V_α is open) there exists some $\varepsilon > 0$ such that $B(y_0, \varepsilon) \subseteq V_\alpha$. Since $y^{(n_j)}$

converges to y_0, there must exist an $N \geq 1$ such that $y^{(n_j)} \in B(y_0, \varepsilon/2)$ for all $n \geq N$. In particular, by the triangle inequality we have $B(y^{(n_j)}, \varepsilon/2) \subseteq B(y_0, \varepsilon)$, and thus $B(y^{(n_j)}, \varepsilon/2) \subseteq V_\alpha$. By definition of $r(y^{(n_j)})$, this implies that $r(y^{(n_j)}) \geq \varepsilon/2$ for all $n \geq N$. But this contradicts the fact that $\lim_{n\to\infty} r(y^{(n)}) = 0$.

- **Case 2:** $0 < r_0 < \infty$. In this case we now have $r(y) > r_0/2$ for all $y \in Y$. This implies that for every $y \in Y$ there exists an $\alpha \in A$ such that $B(y, r_0/2) \subseteq V_\alpha$ (why?).
  We now construct a sequence $y^{(1)}, y^{(2)}, \ldots$ by the following recursive procedure. We let $y^{(1)}$ be any point in Y. The ball $B(y^{(1)}, r_0/2)$ is contained in one of the V_α and thus cannot cover all of Y, since we would then obtain a finite cover, a contradiction. Thus there exists a point $y^{(2)}$ which does not lie in $B(y^{(1)}, r_0/2)$, so in particular $d(y^{(2)}, y^{(1)}) \geq r_0/2$. Choose such a point $y^{(2)}$. The set $B(y^{(1)}, r_0/2) \cup B(y^{(2)}, r_0/2)$ cannot cover all of Y, since we would then obtain two sets V_{α_1} and V_{α_2} which covered Y, a contradiction again. So we can choose a point $y^{(3)}$ which does not lie in $B(y^{(1)}, r_0/2) \cup B(y^{(2)}, r_0/2)$, so in particular $d(y^{(3)}, y^{(1)}) \geq r_0/2$ and $d(y^{(3)}, y^{(2)}) \geq r_0/2$. Continuing in this fashion we obtain a sequence $(y^{(n)})_{n=1}^\infty$ in Y with the property that $d(y^{(k)}, y^{(j)}) \geq r_0/2$ for all $k > j$. In particular the sequence $(y^{(n)})_{n=1}^\infty$ is not a Cauchy sequence, and in fact no subsequence of $(y^{(n)})_{n=1}^\infty$ can be a Cauchy sequence either. But this contradicts the assumption that Y is compact (by Lemma 1.4.7).
- **Case 3:** $r_0 = \infty$. For this case we argue as in Case 2, but replacing the role of $r_0/2$ by (say) 1.

□

It turns out that Theorem 1.5.8 has a converse: if Y has the property that every open cover has a finite subcover, then it is compact (Exercise 1.5.11). In fact, this property is often considered the more fundamental notion of compactness than the sequence-based one. (For metric spaces, the two notions, that of compactness and sequential compactness, are equivalent, but for more general *topological spaces*, the two notions are slightly different, though we will not show this here.)

Theorem 1.5.8 has an important corollary: that every nested sequence of non-empty compact sets is still non-empty.

Corollary 1.5.9 *Let (X, d) be a metric space, and let $K_1, K_2, K_3, \ldots$ be a sequence of non-empty compact subsets of X such that*

$$K_1 \supseteq K_2 \supseteq K_3 \supseteq \cdots .$$

Then the intersection $\bigcap_{n=1}^\infty K_n$ is non-empty.

Proof See Exercise 1.5.6. □

We close this section by listing some miscellaneous properties of compact sets.

Theorem 1.5.10 *Let (X, d) be a metric space.*

(a) *If Y is a compact subset of X, and $Z \subseteq Y$, then Z is compact if and only if Z is closed.*
(b) *If $Y_1, \ldots, Y_n$ are a finite collection of compact subsets of X, then their union $Y_1 \cup \ldots \cup Y_n$ is also compact.*
(c) *Every finite subset of X (including the empty set) is compact.*

Proof See Exercise 1.5.7. □

— Exercises —

Exercise 1.5.1 Show that Definitions 9.1.22 and 1.5.3 match when talking about subsets of the real line with the standard metric.

Exercise 1.5.2 Prove Proposition 1.5.5. (*Hint:* prove the completeness and boundedness separately. For both claims, use proof by contradiction. You will need the axiom of choice, as in Lemma 8.4.5.)

Exercise 1.5.3 Prove Theorem 1.5.7. (*Hint:* use Proposition 1.1.18 and Theorem 9.1.24.)

Exercise 1.5.4 Let $(\mathbf{R}, d)$ be the real line with the standard metric. Give an example of a continuous function $f : \mathbf{R} \to \mathbf{R}$, and an open set $V \subseteq \mathbf{R}$, such that the image $f(V) := \{f(x) : x \in V\}$ of V is *not* open.

Exercise 1.5.5 Let $(\mathbf{R}, d)$ be the real line with the standard metric. Give an example of a continuous function $f : \mathbf{R} \to \mathbf{R}$, and a closed set $F \subseteq \mathbf{R}$, such that $f(F)$ is *not* closed.

Exercise 1.5.6 Prove Corollary 1.5.9. (*Hint:* work in the compact metric space $(K_1, d|_{K_1 \times K_1})$, and consider the sets $V_n := K_1 \backslash K_n$, which are open on K_1. Assume for sake of contradiction that $\bigcap_{n=1}^{\infty} K_n = \emptyset$, and then apply Theorem 1.5.8.)

Exercise 1.5.7 Prove Theorem 1.5.10. (*Hint:* for part (c), you may wish to use (b), and first prove that every singleton set is compact.)

Exercise 1.5.8 Let (X, d_{l^1}) be the metric space from Exercise 1.1.15. For each natural number n, let $e^{(n)} = (e_j^{(n)})_{j=0}^{\infty}$ be the sequence in X such that $e_j^{(n)} := 1$ when $n = j$ and $e_j^{(n)} := 0$ when $n \neq j$. Show that the set $\{e^{(n)} : n \in \mathbf{N}\}$ is a closed and bounded subset of X, but is not compact. (This is despite the fact that (X, d_{l^1}) is even a complete metric space—a fact which we will not prove here. The problem is that not that X is incomplete, but rather that it is "infinite-dimensional", in a sense that we will not discuss here.)

Exercise 1.5.9 Show that a metric space (X, d) is compact if and only if every sequence in X has at least one limit point.

Exercise 1.5.10 A metric space (X, d) is called *totally bounded* if for every $\varepsilon > 0$, there exists a natural number n and a finite number of balls $B(x^{(1)}, \varepsilon), \ldots, B(x^{(n)}, \varepsilon)$ which cover X (i.e., $X = \bigcup_{i=1}^{n} B(x^{(i)}, \varepsilon)$.

(a) Show that every totally bounded space is bounded.
(b) Show the following stronger version of Proposition 1.5.5: if (X, d) is compact, then complete and totally bounded. (*Hint:* if X is not totally bounded, then there is some $\varepsilon > 0$ such that X cannot be covered by finitely many ε-balls. Then use Exercise 8.5.20 to find an infinite sequence of balls $B(x^{(n)}, \varepsilon/2)$ which are disjoint from each other. Use this to then construct a sequence which has no convergent subsequence.)
(c) Conversely, show that if X is complete and totally bounded, then X is compact. (*Hint:* if $(x^{(n)})_{n=1}^{\infty}$ is a sequence in X, use the total boundedness hypothesis to recursively construct a sequence of subsequences $(x^{(n;j)})_{n=1}^{\infty}$ of $(x^{(n)})_{n=1}^{\infty}$ for each positive integer j, such that for each j, the elements of the sequence $(x^{(n;j)})_{n=1}^{\infty}$ are contained in a single ball of radius $1/j$, and also that each sequence $(x^{(n;j+1)})_{n=1}^{\infty}$ is a subsequence of the previous one $(x^{(n;j)})_{n=1}^{\infty}$. Then show that the "diagonal" sequence $(x^{(n;n)})_{n=1}^{\infty}$ is a Cauchy sequence, and then use the completeness hypothesis.)

Exercise 1.5.11 Let (X, d) have the property that every open cover of X has a finite subcover. Show that X is compact. (*Hint:* if X is not compact, then by Exercise 1.5.9, there is a sequence $(x^{(n)})_{n=1}^{\infty}$ with no limit points. Then for every $x \in X$ there exists a ball $B(x, \varepsilon)$ containing x which contains at most finitely many elements of this sequence. Now use the hypothesis.)

Exercise 1.5.12 Let (X, d_{disc}) be a metric space with the discrete metric d_{disc}.

(a) Show that X is always complete.
(b) When is X compact, and when is X not compact? Prove your claim. (*Hint:* the Heine–Borel theorem will be useless here since that only applies to Euclidean spaces.)

Exercise 1.5.13 Let E and F be two compact subsets of $\mathbf{R}$ (with the standard metric $d(x, y) = |x - y|$). Show that the Cartesian product $E \times F := \{(x, y) : x \in E, y \in F\}$ is a compact subset of $\mathbf{R}^2$ (with the Euclidean metric d_{l^2}).

Exercise 1.5.14 Let (X, d) be a metric space, let E be a non-empty compact subset of X, and let x_0 be a point in X. Show that there exists a point $x \in E$ such that

$$d(x_0, x) = \inf\{d(x_0, y) : y \in E\},$$

i.e., x is the closest point in E to x_0. (*Hint:* let R be the quantity $R := \inf\{d(x_0, y) : y \in E\}$. Construct a sequence $(x^{(n)})_{n=1}^{\infty}$ in E such that $d(x_0, x^{(n)}) \leq R + \frac{1}{n}$, and then use the compactness of E.)

Exercise 1.5.15 Let (X, d) be a compact metric space. Suppose that $(K_\alpha)_{\alpha \in I}$ is a collection of closed sets in X with the property that any finite subcollection of these sets necessarily has non-empty intersection, thus $\bigcap_{\alpha \in F} K_\alpha \neq \emptyset$ for all finite $F \subseteq I$. (This property is known as the *finite intersection property*.) Show that the *entire* collection has non-empty intersection, thus $\bigcap_{\alpha \in I} K_\alpha \neq \emptyset$. Show by counterexample that this statement fails if X is not compact.

Chapter 2
Continuous Functions on Metric Spaces

2.1 Continuous Functions

In the previous chapter we studied a single metric space (X, d), and the various types of sets one could find in that space. While this is already quite a rich subject, the theory of metric spaces becomes even richer, and of more importance to analysis, when one considers not just a single metric space, but rather *pairs* (X, d_X) and (Y, d_Y) of metric spaces, as well as *continuous functions* $f: X \to Y$ between such spaces. To define this concept, we generalize Definition 9.4.1 as follows:

Definition 2.1.1 (*Continuous functions*) Let (X, d_X) be a metric space, and let (Y, d_Y) be another metric space, and let $f: X \to Y$ be a function. If $x_0 \in X$, we say that f is *continuous at* x_0 iff for every $\varepsilon > 0$, there exists a $\delta > 0$ such that $d_Y(f(x), f(x_0)) < \varepsilon$ whenever $d_X(x, x_0) < \delta$. We say that f is *continuous* iff it is continuous at every point $x \in X$.

Remark 2.1.2 Continuous functions are also sometimes called *continuous maps*. Mathematically, there is no distinction between the two terminologies.

Remark 2.1.3 If $f: X \to Y$ is continuous, and K is any subset of X, then the restriction $f|_K : K \to Y$ of f to K is also continuous (why?).

We now generalize much of the discussion in Chap. 9. We first observe that continuous functions preserve convergence:

Theorem 2.1.4 (Continuity preserves convergence) *Suppose that (X, d_X) and (Y, d_Y) are metric spaces. Let $f: X \to Y$ be a function, and let $x_0 \in X$ be a point in X. Then the following three statements are logically equivalent:*

(a) *f is continuous at x_0.*
(b) *Whenever $(x^{(n)})_{n=1}^{\infty}$ is a sequence in X which converges to x_0 with respect to the metric d_X, the sequence $(f(x^{(n)}))_{n=1}^{\infty}$ converges to $f(x_0)$ with respect to the metric d_Y.*

T. Tao, *Analysis II*, Texts and Readings in Mathematics,
https://doi.org/10.1007/978-981-19-7284-3_2

(c) *For every open set $V \subseteq Y$ that contains $f(x_0)$, there exists an open set $U \subseteq X$ containing x_0 such that $f(U) \subseteq V$.*

Proof See Exercise 2.1.1. □

Another important characterization of continuous functions involves open sets.

Theorem 2.1.5 *Let (X, d_X) be a metric space, and let (Y, d_Y) be another metric space. Let $f : X \to Y$ be a function. Then the following four statements are equivalent:*

(a) *f is continuous.*
(b) *Whenever $(x^{(n)})_{n=1}^{\infty}$ is a sequence in X which converges to some point $x_0 \in X$ with respect to the metric d_X, the sequence $(f(x^{(n)}))_{n=1}^{\infty}$ converges to $f(x_0)$ with respect to the metric d_Y.*
(c) *Whenever V is an open set in Y, the set $f^{-1}(V) := \{x \in X : f(x) \in V\}$ is an open set in X.*
(d) *Whenever F is a closed set in Y, the set $f^{-1}(F) := \{x \in X : f(x) \in F\}$ is a closed set in X.*

Proof See Exercise 2.1.2. □

Remark 2.1.6 It may seem strange that continuity ensures that the *inverse* image of an open set is open. One may guess instead that the reverse should be true, that the *forward* image of an open set is open; but this is not true; see Exercises 1.5.4, 1.5.5.

As a quick corollary of the above two theorems we obtain

Corollary 2.1.7 (Continuity preserved by composition) *Let (X, d_X), (Y, d_Y), and (Z, d_Z) be metric spaces.*

(a) *If $f : X \to Y$ is continuous at a point $x_0 \in X$, and $g : Y \to Z$ is continuous at $f(x_0)$, then the composition $g \circ f : X \to Z$, defined by $g \circ f(x) := g(f(x))$, is continuous at x_0.*
(b) *If $f : X \to Y$ is continuous, and $g : Y \to Z$ is continuous, then $g \circ f : X \to Z$ is also continuous.*

Proof See Exercise 2.1.3. □

Example 2.1.8 If $f : X \to \mathbf{R}$ is a continuous function, then the function $f^2 : X \to \mathbf{R}$ defined by $f^2(x) := f(x)^2$ is automatically continuous also. This is because we have $f^2 = g \circ f$, where $g : \mathbf{R} \to \mathbf{R}$ is the squaring function $g(x) := x^2$, and g is a continuous function.

— Exercises —

Exercise 2.1.1 Prove Theorem 2.1.4. (*Hint:* review your proof of Proposition 9.4.7.)

Exercise 2.1.2 Prove Theorem 2.1.5. (*Hint:* Theorem 2.1.4 already shows that (a) and (b) are equivalent.)

Exercise 2.1.3 Use Theorem 2.1.4 and Theorem 2.1.5 to prove Corollary 2.1.7.

Exercise 2.1.4 Give an example of functions $f: \mathbf{R} \to \mathbf{R}$ and $g: \mathbf{R} \to \mathbf{R}$ such that

(a) f is not continuous, but g and $g \circ f$ are continuous.
(b) g is not continuous, but f and $g \circ f$ are continuous.
(c) f and g are not continuous, but $g \circ f$ is continuous.

Explain briefly why these examples do not contradict Corollary 2.1.7.

Exercise 2.1.5 Let (X, d) be a metric space, and let $(E, d|_{E\times E})$ be a subspace of (X, d). Let $\iota_{E\to X} : E \to X$ be the inclusion map, defined by setting $\iota_{E\to X}(x) := x$ for all $x \in E$. Show that $\iota_{E\to X}$ is continuous.

Exercise 2.1.6 Let $f: X \to Y$ be a function from one metric space (X, d_X) to another (Y, d_Y). Let E be a subset of X (which we give the induced metric $d_X|_{E\times E}$), and let $f|_E : E \to Y$ be the restriction of f to E, thus $f|_E(x) := f(x)$ when $x \in E$. If $x_0 \in E$ and f is continuous at x_0, show that $f|_E$ is also continuous at x_0. (Is the converse of this statement true? Explain.) Conclude that if f is continuous, then $f|_E$ is continuous. Thus restriction of the domain of a function does not destroy continuity. (*Hint:* use Exercise 2.1.5.)

Exercise 2.1.7 Let $f: X \to Y$ be a function from one metric space (X, d_X) to another (Y, d_Y). Suppose that the image $f(X)$ of X is contained in some subset $E \subseteq Y$ of Y. Let $g: X \to E$ be the function which is the same as f but with the codomain restricted from Y to E, thus $g(x) = f(x)$ for all $x \in X$. We give E the metric $d_Y|_{E\times E}$ induced from Y. Show that for any $x_0 \in X$, that f is continuous at x_0 if and only if g is continuous at x_0. Conclude that f is continuous if and only if g is continuous. (Thus the notion of continuity is not affected if one restricts the codomain of the function.)

2.2 Continuity and Product Spaces

Given two functions $f: X \to Y$ and $g: X \to Z$, one can define their *pairing* $(f, g): X \to Y \times Z$ defined by $(f, g)(x) := (f(x), g(x))$, i.e., this is the function taking values in the Cartesian product $Y \times Z$ whose first coordinate is $f(x)$ and whose second coordinate is $g(x)$ (cf. Exercise 3.5.7). For instance, if $f: \mathbf{R} \to \mathbf{R}$ is the function $f(x) := x^2 + 3$, and $g: \mathbf{R} \to \mathbf{R}$ is the function $g(x) = 4x$, then $(f, g): \mathbf{R} \to \mathbf{R}^2$ is the function $(f, g)(x) := (x^2 + 3, 4x)$. The pairing operation preserves continuity:

Lemma 2.2.1 *Let $f: X \to \mathbf{R}$ and $g: X \to \mathbf{R}$ be functions, and let $(f, g): X \to \mathbf{R}^2$ be their direct sum. We give $\mathbf{R}^2$ the Euclidean metric.*

(a) If $x_0 \in X$, then f and g are both continuous at x_0 if and only if (f, g) is continuous at x_0.
(b) f and g are both continuous if and only if (f, g) is continuous.

Proof See Exercise 2.2.1. □

To use this, we first need another continuity result:

Lemma 2.2.2 *The addition function* $(x, y) \mapsto x + y$*, the subtraction function* $(x, y) \mapsto x - y$*, the multiplication function* $(x, y) \mapsto xy$*, the maximum function* $(x, y) \mapsto \max(x, y)$*, and the minimum function* $(x, y) \mapsto \min(x, y)$ *are all continuous functions from* $\mathbf{R}^2$ *to* $\mathbf{R}$*. The division function* $(x, y) \mapsto x/y$ *is a continuous function from* $\mathbf{R} \times (\mathbf{R}\backslash\{0\}) = \{(x, y) \in \mathbf{R}^2 : y \neq 0\}$ *to* $\mathbf{R}$*. For any real number c, the function* $x \mapsto cx$ *is a continuous function from* $\mathbf{R}$ *to* $\mathbf{R}$*.*

Proof See Exercise 2.2.2. □

Combining these lemmas we obtain

Corollary 2.2.3 *Let* (X, d) *be a metric space, and let* $f : X \to \mathbf{R}$ *and* $g : X \to \mathbf{R}$ *be functions. Let c be a real number.*

(a) If $x_0 \in X$ *and f and g are continuous at* x_0*, then the functions* $f + g : X \to \mathbf{R}$*,* $f - g : X \to \mathbf{R}$*,* $fg : X \to \mathbf{R}$*,* $\max(f, g) : X \to \mathbf{R}$*,* $\min(f, g) : X \to \mathbf{R}$*, and* $cf : X \to \mathbf{R}$ *(see Definition 9.2.1 for definitions) are also continuous at* x_0*. If* $g(x) \neq 0$ *for all* $x \in X$*, then* $f/g : X \to \mathbf{R}$ *is also continuous.*
(b) If f and g are continuous, then the functions $f + g : X \to \mathbf{R}$*,* $f - g : X \to \mathbf{R}$*,* $fg : X \to \mathbf{R}$*,* $\max(f, g) : X \to \mathbf{R}$*,* $\min(f, g) : X \to \mathbf{R}$*, and* $cf : X \to \mathbf{R}$ *are also continuous at* x_0*. If* $g(x) \neq 0$ *for all* $x \in X$*, then* $f/g : X \to \mathbf{R}$ *is also continuous at* x_0*.*

Proof We first prove (a). Since f and g are continuous at x_0, then by Lemma 2.2.1 $(f, g) : X \to \mathbf{R}^2$ is also continuous at x_0. On the other hand, from Lemma 2.2.2 the function $(x, y) \mapsto x + y$ is continuous at every point in $\mathbf{R}^2$ and in particular is continuous at $(f, g)(x_0)$. If we then compose these two functions using Corollary 2.1.7 we conclude that $f + g : X \to \mathbf{R}$ is continuous. A similar argument gives the continuity of $f - g$, fg, $\max(f, g)$, $\min(f, g)$, and cf. To prove the claim for f/g, we first use Exercise 2.1.7 to restrict the codomain of g from $\mathbf{R}$ to $\mathbf{R}\backslash\{0\}$, and then one can argue as before. The claim (b) follows immediately from (a). □

This corollary allows us to demonstrate the continuity of a large class of functions; we give some examples below.

— Exercises —

Exercise 2.2.1 Prove Lemma 2.2.1. (*Hint:* use Proposition 1.1.18 and Theorem 2.1.4.)

Exercise 2.2.2 Prove Lemma 2.2.2. (*Hint:* use Theorem 2.1.5 and limit laws (Theorem 6.1.19).)

Exercise 2.2.3 Show that if $f : X \to \mathbf{R}$ is a continuous function, so is the function $|f| : X \to \mathbf{R}$ defined by $|f|(x) := |f(x)|$.

Exercise 2.2.4 Let $\pi_1 \colon \mathbf{R}^2 \to \mathbf{R}$ and $\pi_2 : \mathbf{R}^2 \to \mathbf{R}$ be the functions $\pi_1(x, y) := x$ and $\pi_2(x, y) := y$ (these two functions are sometimes called the *coordinate functions* on $\mathbf{R}^2$). Show that π_1 and π_2 are continuous. Conclude that if $f \colon \mathbf{R} \to X$ is any continuous function into a metric space (X, d), then the functions $g_1 \colon \mathbf{R}^2 \to X$ and $g_2 : \mathbf{R}^2 \to X$ defined by $g_1(x, y) := f(x)$ and $g_2(x, y) := f(y)$ are also continuous.

Exercise 2.2.5 Let $n, m \geq 0$ be integers. Suppose that for every $0 \leq i \leq n$ and $0 \leq j \leq m$ we have a real number c_{ij}. Form the function $P \colon \mathbf{R}^2 \to \mathbf{R}$ defined by

$$P(x, y) := \sum_{i=0}^{n} \sum_{j=0}^{m} c_{ij} x^i y^j.$$

(Such a function is known as a *polynomial of two variables*; a typical example of such a polynomial is $P(x, y) = x^3 + 2xy^2 - x^2 + 3y + 6$.) Show that P is continuous. (*Hint:* use Exercise 2.2.4 and Corollary 2.2.3.) Conclude that if $f \colon X \to \mathbf{R}$ and $g \colon X \to \mathbf{R}$ are continuous functions, then the function $P(f, g) : X \to \mathbf{R}$ defined by $P(f, g)(x) := P(f(x), g(x))$ is also continuous.

Exercise 2.2.6 Let $\mathbf{R}^m$ and $\mathbf{R}^n$ be Euclidean spaces. If $f \colon X \to \mathbf{R}^m$ and $g \colon X \to \mathbf{R}^n$ are continuous functions, show that $(f, g) \colon X \to \mathbf{R}^{m+n}$ is also continuous, where we have identified $\mathbf{R}^m \times \mathbf{R}^n$ with $\mathbf{R}^{m+n}$ in the obvious manner. Is the converse statement true?

Exercise 2.2.7 Let $k \geq 1$, let I be a finite subset of $\mathbf{N}^k$, and let $c : I \to \mathbf{R}$ be a function. Form the function $P \colon \mathbf{R}^k \to \mathbf{R}$ defined by

$$P(x_1, \ldots, x_k) := \sum_{(i_1, \ldots, i_k) \in I} c(i_1, \ldots, i_k) x_1^{i_1} \ldots x_k^{i_k}.$$

(Such a function is known as a *polynomial of k variables*; a typical example of such a polynomial is $P(x_1, x_2, x_3) = 3x_1^3 x_2 x_3^2 - x_2 x_3^2 + x_1 + 5$.) Show that P is continuous. (*Hint:* use induction on k, Exercise 2.2.6, and either Exercise 2.2.5 or Lemma 2.2.2.)

Exercise 2.2.8 Let (X, d_X) and (Y, d_Y) be metric spaces. Define the metric $d_{X \times Y} : (X \times Y) \times (X \times Y) \to [0, \infty)$ by the formula

$$d_{X \times Y}((x, y), (x', y')) := d_X(x, x') + d_Y(y, y').$$

Show that $(X \times Y, d_{X \times Y})$ is a metric space, and deduce an analogue of Proposition 1.1.18 and Lemma 2.2.1.

Exercise 2.2.9 Let $f \colon \mathbf{R}^2 \to \mathbf{R}$ be a function from $\mathbf{R}^2$ to $\mathbf{R}$. Let (x_0, y_0) be a point in $\mathbf{R}^2$. If f is continuous at (x_0, y_0), show that

$$\lim_{x \to x_0} \limsup_{y \to y_0} f(x, y) = \lim_{y \to y_0} \limsup_{x \to x_0} f(x, y) = f(x_0, y_0)$$

and

$$\lim_{x\to x_0} \liminf_{y\to y_0} f(x, y) = \lim_{y\to y_0} \liminf_{x\to x_0} f(x, y) = f(x_0, y_0).$$

(Recall that $\limsup_{x\to x_0} f(x) := \inf_{r>0} \sup_{|x-x_0|<r} f(x)$ and $\liminf_{x\to x_0} f(x) := \sup_{r>0} \inf_{|x-x_0|<r} f(x)$.) In particular, we have

$$\lim_{x\to x_0} \lim_{y\to y_0} f(x, y) = \lim_{y\to y_0} \lim_{x\to x_0} f(x, y)$$

whenever the limits on both sides exist. (Note that the limits do not necessarily exist in general; consider for instance the function $f: \mathbf{R}^2 \to \mathbf{R}$ such that $f(x, y) = y \sin \frac{1}{x}$ when $xy \neq 0$ and $f(x, y) = 0$ otherwise.) Discuss the comparison between this result and Example 1.2.7.

Exercise 2.2.10 Let $f: \mathbf{R}^2 \to \mathbf{R}$ be a continuous function. Show that for each $x \in \mathbf{R}$, the function $y \mapsto f(x, y)$ is continuous on $\mathbf{R}$, and for each $y \in \mathbf{R}$, the function $x \mapsto f(x, y)$ is continuous on $\mathbf{R}$. Thus a function $f(x, y)$ which is jointly continuous in (x, y) is also continuous in each variable x, y separately.

Exercise 2.2.11 Let $f: \mathbf{R}^2 \to \mathbf{R}$ be the function defined by $f(x, y) := \frac{xy}{x^2+y^2}$ when $(x, y) \neq (0, 0)$, and $f(x, y) = 0$ otherwise. Show that for each fixed $x \in \mathbf{R}$, the function $y \mapsto f(x, y)$ is continuous on $\mathbf{R}$, and that for each fixed $y \in \mathbf{R}$, the function $x \mapsto f(x, y)$ is continuous on $\mathbf{R}$, but that the function $f: \mathbf{R}^2 \to \mathbf{R}$ is not continuous on $\mathbf{R}^2$. This shows that the converse to Exercise 2.2.10 fails; it is possible to be continuous in each variable separately without being jointly continuous.

Exercise 2.2.12 Let $f: \mathbf{R}^2 \to \mathbf{R}$ be the function defined by $f(x, y) := x^2/y$ when $y \neq 0$, and $f(x, y) := 0$ when $y = 0$. Show that $\lim_{t\to 0} f(tx, ty) = f(0, 0)$ for every $(x, y) \in \mathbf{R}^2$, but that f is not continuous at the origin. Thus being continuous on every line through the origin is not enough to guarantee continuity at the origin!

2.3 Continuity and Compactness

Continuous functions interact well with the concept of compact sets defined in Definition 1.5.1.

Theorem 2.3.1 (Continuous maps preserve compactness) *Let $f: X \to Y$ be a continuous map from one metric space (X, d_X) to another (Y, d_Y). Let $K \subseteq X$ be any compact subset of X. Then the image $f(K) := \{f(x) : x \in K\}$ of K is also compact.*

Proof See Exercise 2.3.1. □

This theorem has an important consequence. Recall from Definition 9.6.5 the notion of a function $f: X \to \mathbf{R}$ attaining a maximum or minimum at a point. We may generalize Proposition 9.6.7 as follows:

Proposition 2.3.2 (Maximum principle) *Let (X, d) be a compact metric space, and let $f : X \to \mathbf{R}$ be a continuous function. Then f is bounded. Furthermore, if X is non-empty, then f attains its maximum at some point $x_{max} \in X$ and also attains its minimum at some point $x_{min} \in X$.*

Proof See Exercise 2.3.2. □

Remark 2.3.3 As was already noted in Exercise 9.6.1, this principle can fail if X is not compact. This proposition should be compared with Lemma 9.6.3 and Proposition 9.6.7.

Another advantage of continuous functions on compact sets is that they are *uniformly continuous*. We generalize Definition 9.9.2 as follows:

Definition 2.3.4 (*Uniform continuity*) Let $f : X \to Y$ be a map from one metric space (X, d_X) to another (Y, d_Y). We say that f is *uniformly continuous* if, for every $\varepsilon > 0$, there exists a $\delta > 0$ such that $d_Y(f(x), f(x')) < \varepsilon$ whenever $x, x' \in X$ are such that $d_X(x, x') < \delta$.

Every uniformly continuous function is continuous, but not conversely (Exercise 2.3.3). But if the domain X is compact, then the two notions are equivalent:

Theorem 2.3.5 *Let (X, d_X) and (Y, d_Y) be metric spaces, and suppose that (X, d_X) is compact. If $f : X \to Y$ is function, then f is continuous if and only if it is uniformly continuous.*

Proof If f is uniformly continuous then it is also continuous by Exercise 2.3.3. Now suppose that f is continuous. Fix $\varepsilon > 0$. For every $x_0 \in X$, the function f is continuous at x_0. Thus there exists a $\delta(x_0) > 0$, depending on x_0, such that $d_Y(f(x), f(x_0)) < \varepsilon/2$ whenever $d_X(x, x_0) < \delta(x_0)$. In particular, by the triangle inequality this implies that $d_Y(f(x), f(x')) < \varepsilon$ whenever $x \in B_{(X,d_X)}(x_0, \delta(x_0)/2)$ and $d_X(x', x) < \delta(x_0)/2$ (why?).

Now consider the (possibly infinite) collection of balls

$$\{B_{(X,d_X)}(x_0, \delta(x_0)/2) : x_0 \in X\}.$$

Each ball in this collection is of course open, and the union of all these balls covers X, since each point x_0 in X is contained in its own ball $B_{(X,d_X)}(x_0, \delta(x_0)/2)$. Hence, by Theorem 1.5.8, there exist a finite number of points $x_1, \ldots, x_n$ such that the balls $B_{(X,d_X)}(x_j, \delta(x_j)/2)$ for $j = 1, \ldots, n$ cover X:

$$X \subseteq \bigcup_{j=1}^{n} B_{(X,d_X)}(x_j, \delta(x_j)/2).$$

Now let $\delta := \min_{j=1}^{n} \delta(x_j)/2$. Since each of the $\delta(x_j)$ is positive, and there are only a finite number of j, we see that $\delta > 0$. Now let x, x' be any two points in X such that

$d_X(x, x') < \delta$. Since the balls $B_{(X,d_X)}(x_j, \delta(x_j)/2)$ cover X, we see that there must exist $1 \leq j \leq n$ such that $x \in B_{(X,d_X)}(x_j, \delta(x_j)/2)$. Since $d_X(x, x') < \delta$, we have $d_X(x, x') < \delta(x_j)/2$, and so by the previous discussion we have $d_Y(f(x), f(x')) < \varepsilon$. We have thus found a δ such that $d_Y(f(x), f(x')) < \varepsilon$ whenever $d(x, x') < \delta$, and this proves uniform continuity as desired. □

— Exercises —

Exercise 2.3.1 Prove Theorem 2.3.1.

Exercise 2.3.2 Prove Proposition 2.3.2. (*Hint:* modify the proof of Proposition 9.6.7.)

Exercise 2.3.3 Show that every uniformly continuous function is continuous, but give an example that shows that not every continuous function is uniformly continuous.

Exercise 2.3.4 Let (X, d_X), (Y, d_Y), (Z, d_Z) be metric spaces, and let $f: X \to Y$ and $g: Y \to Z$ be two uniformly continuous functions. Show that $g \circ f: X \to Z$ is also uniformly continuous.

Exercise 2.3.5 Let (X, d_X) be a metric space, and let $f: X \to \mathbf{R}$ and $g: X \to \mathbf{R}$ be uniformly continuous functions. Show that the pairing $(f, g): X \to \mathbf{R}^2$ defined by $(f, g)(x) := (f(x), g(x))$ is uniformly continuous.

Exercise 2.3.6 Show that the addition function $(x, y) \mapsto x + y$ and the subtraction function $(x, y) \mapsto x - y$ are uniformly continuous from $\mathbf{R}^2$ to $\mathbf{R}$, but the multiplication function $(x, y) \mapsto xy$ is not. Conclude that if $f: X \to \mathbf{R}$ and $g: X \to \mathbf{R}$ are uniformly continuous functions on a metric space (X, d), then $f + g: X \to \mathbf{R}$ and $f - g: X \to \mathbf{R}$ are also uniformly continuous. Give an example to show that $fg: X \to \mathbf{R}$ need not be uniformly continuous. What is the situation for $\max(f, g)$, $\min(f, g)$, f/g, and cf for a real number c?

2.4 Continuity and Connectedness

We now describe another important concept in metric spaces, that of *connectedness*.

Definition 2.4.1 (*Connected spaces*) Let (X, d) be a metric space. We say that X is *disconnected* iff there exist disjoint non-empty open sets V and W in X such that $V \cup W = X$. (Equivalently, X is disconnected if and only if X contains a non-empty proper subset which is simultaneously closed and open.) We say that X is *connected* iff it is non-empty and not disconnected.

We declare the empty set $\emptyset$ as being special—it is neither connected nor disconnected; one could think of the empty set as "unconnected".

Example 2.4.2 Consider the set $X := [1, 2] \cup [3, 4]$, with the usual metric. This set is disconnected because the sets $[1, 2]$ and $[3, 4]$ are open relative to X (why?).

Intuitively, a disconnected set is one which can be separated into two disjoint open sets; a connected set is one which cannot be separated in this manner. We defined what it means for a metric space to be connected; we can also define what it means for a set to be connected.

Definition 2.4.3 (*Connected sets*) Let (X, d) be a metric space, and let Y be a subset of X. We say that Y is *connected* iff the metric space $(Y, d|_{Y\times Y})$ is connected, and we say that Y is *disconnected* iff the metric space $(Y, d|_{Y\times Y})$ is disconnected.

Remark 2.4.4 This definition is intrinsic; whether a set Y is connected or not depends only on what the metric is doing on Y, but not on what ambient space X one placing Y in.

On the real line, connected sets are easy to describe.

Theorem 2.4.5 *Let X be a non-empty subset of the real line* **R**. *Then the following statements are equivalent.*

(a) X is connected.
(b) Whenever $x, y \in X$ and $x < y$, the interval $[x, y]$ is also contained in X.
(c) X is an interval (in the sense of Definition 9.1.1).

Proof First we show that (a) implies (b). Suppose that X is connected, and suppose for sake of contradiction that we could find points $x < y$ in X such that $[x, y]$ is *not* contained in X. Then there exists a real number $x < z < y$ such that $z \notin X$. Thus the sets $(-\infty, z) \cap X$ and $(z, \infty) \cap X$ will cover X. But these sets are non-empty (because they contain x and y, respectively) and are open relative to X, and so X is disconnected, a contradiction.

Now we show that (b) implies (a). Let X be a set obeying the property (b). Suppose for sake of contradiction that X is disconnected. Then there exist disjoint non-empty sets V, W which are open relative to X, such that $V \cup W = X$. Since V and W are non-empty, we may choose an $x \in V$ and $y \in W$. Since V and W are disjoint, we have $x \neq y$; without loss of generality we may assume $x < y$. By property (b), we know that the entire interval $[x, y]$ is contained in X.

Now consider the set $[x, y] \cap V$. This set is both bounded and non-empty (because it contains x). Thus it has a supremum

$$z := \sup([x, y] \cap V).$$

Clearly $z \in [x, y]$, and hence $z \in X$. Thus either $z \in V$ or $z \in W$. Suppose first that $z \in V$. Then $z \neq y$ (since $y \in W$ and V is disjoint from W). But V is open relative to X, which contains $[x, y]$, so there is some ball $B_{([x,y],d)}(z, r)$ which is contained in V. But this contradicts the fact that z is the supremum of $[x, y] \cap V$. Now suppose that $z \in W$. Then $z \neq x$ (since $x \in V$ and V is disjoint from W). But W is open relative to X, which contains $[x, y]$, so there is some ball $B_{([x,y],d)}(z, r)$ which is contained in W. But this again contradicts the fact that z is the supremum of $[x, y] \cap V$. Thus in either case we obtain a contradiction, which means that X cannot be disconnected and must therefore be connected.

It remains to show that (b) and (c) are equivalent; we leave this to Exercise 2.4.3. □

Continuous functions map connected sets to connected sets:

Theorem 2.4.6 (Continuity preserves connectedness) *Let $f: X \to Y$ be a continuous map from one metric space (X, d_X) to another (Y, d_Y). Let E be any connected subset of X. Then $f(E)$ is also connected.*

Proof See Exercise 2.4.4. □

An important corollary of this result is the intermediate value theorem, generalizing Theorem 9.7.1.

Corollary 2.4.7 (Intermediate value theorem) *Let $f: X \to \mathbf{R}$ be a continuous map from one metric space (X, d_X) to the real line. Let E be any connected subset of X, and let a, b be any two elements of E. Let y be a real number between $f(a)$ and $f(b)$, i.e., either $f(a) \leq y \leq f(b)$ or $f(a) \geq y \geq f(b)$. Then there exists $c \in E$ such that $f(c) = y$.*

Proof See Exercise 2.4.5. □

— Exercises —

Exercise 2.4.1 Let (X, d_{disc}) be a metric space with the discrete metric. Let E be a subset of X which contains at least two elements. Show that E is disconnected.

Exercise 2.4.2 Let $f: X \to Y$ be a function from a connected metric space (X, d) to a metric space (Y, d_{disc}) with the discrete metric. Show that f is continuous if and only if it is constant. (*Hint:* use Exercise 2.4.1.)

Exercise 2.4.3 Prove the equivalence of statements (b) and (c) in Theorem 2.4.5.

Exercise 2.4.4 Prove Theorem 2.4.6. (*Hint:* the formulation of continuity in Theorem 2.1.5(c) is the most convenient to use.)

Exercise 2.4.5 Use Theorem 2.4.6 to prove Corollary 2.4.7.

Exercise 2.4.6 Let (X, d) be a metric space, and let $(E_\alpha)_{\alpha \in I}$ be a collection of connected sets in X with I non-empty. Suppose also that $\bigcap_{\alpha \in I} E_\alpha$ is non-empty. Show that $\bigcup_{\alpha \in I} E_\alpha$ is connected.

Exercise 2.4.7 Let (X, d) be a metric space, and let E be a subset of X. We say that E is *path-connected* iff, for every $x, y \in E$, there exists a continuous function $\gamma : [0, 1] \to E$ from the unit interval $[0, 1]$ to E such that $\gamma(0) = x$ and $\gamma(1) = y$. Show that every non-empty path-connected set is connected. (The converse is false, but is a bit tricky to show and will not be detailed here.)

Exercise 2.4.8 Let (X, d) be a metric space, and let E be a subset of X. Show that if E is connected, then the closure $\overline{E}$ of E is also connected. Is the converse true?

Exercise 2.4.9 Let (X, d) be a metric space. Let us define a relation $x \sim y$ on X by declaring $x \sim y$ iff there exists a connected subset of X which contains both x and y. Show that this is an equivalence relation (i.e., it obeys the reflexive, symmetric, and transitive axioms). Also, show that the equivalence classes of this relation (i.e., the sets of the form $\{y \in X : y \sim x\}$ for some $x \in X$) are all closed and connected. (*Hint:* use Exercise 2.4.6 and Exercise 2.4.8.) These sets are known as the *connected components* of X.

Exercise 2.4.10 Combine Proposition 2.3.2 and Corollary 2.4.7 to deduce a theorem for continuous functions on a compact connected domain which generalizes Corollary 9.7.4.

2.5 Topological Spaces (Optional)

The concept of a metric space can be generalized to that of a *topological space*. The idea here is not to view the metric d as the fundamental object; indeed, in a general topological space there is no metric at all. Instead, it is the collection of *open sets* which is the fundamental concept. Thus, whereas in a metric space one introduces the metric d first, and then uses the metric to define first the concept of an open ball and then the concept of an open set, in a topological space one starts just with the notion of an open set. As it turns out, starting from the open sets, one cannot necessarily reconstruct a usable notion of a ball or metric (thus not all topological spaces will be metric spaces), but remarkably one can still define many of the concepts in the preceding sections.

We will not use topological spaces at all in this text, and so we shall be rather brief in our treatment of them here. A more complete study of these spaces can of course be found in any topology textbook or a more advanced analysis text.

Definition 2.5.1 (*Topological spaces*) A *topological space* is a pair $(X, \mathcal{F})$, where X is a set and $\mathcal{F} \subseteq 2^X$ is a collection of subsets of X, whose elements are referred to as *open sets*. Furthermore, the collection $\mathcal{F}$ must obey the following properties:

- The empty set $\emptyset$ and the whole set X are open; in other words, $\emptyset \in \mathcal{F}$ and $X \in \mathcal{F}$.
- Any finite intersection of open sets is open. In other words, if $V_1, \ldots, V_n$ is elements of $\mathcal{F}$, then $V_1 \cap \ldots \cap V_n$ is also in $\mathcal{F}$.
- Any arbitrary union of open sets is open (including infinite unions). In other words, if $(V_\alpha)_{\alpha \in I}$ is a family of sets in $\mathcal{F}$, then $\bigcup_{\alpha \in I} V_\alpha$ is also in $\mathcal{F}$.

In many cases, the collection $\mathcal{F}$ of open sets can be deduced from context, and we shall refer to the topological space $(X, \mathcal{F})$ simply as X.

From Proposition 1.2.15 we see that every metric space (X, d) is automatically also a topological space (if we set $\mathcal{F}$ equal to the collection of sets which are open in (X, d)). However, there do exist topological spaces which do not arise from metric spaces (see Exercise 2.5.1, 2.5.6).

We now develop the analogues of various notions in this chapter and the previous chapter for topological spaces. The notion of a ball must be replaced by the notion of a *neighbourhood*.

Definition 2.5.2 (*Neighborhoods*) Let $(X, \mathcal{F})$ be a topological space, and let $x \in X$. A *neighborhood of* x is defined to be any open set in $\mathcal{F}$ which contains x.

Example 2.5.3 If (X, d) is a metric space, $x \in X$, and $r > 0$, then $B(x, r)$ is a neighborhood of x.

Definition 2.5.4 (*Topological convergence*) Let m be an integer, $(X, \mathcal{F})$ be a topological space and let $(x^{(n)})_{n=m}^{\infty}$ be a sequence of points in X. Let x be a point in X. We say that $(x^{(n)})_{n=m}^{\infty}$ *converges to* x if and only if, for every neighborhood V of x, there exists an $N \geq m$ such that $x^{(n)} \in V$ for all $n \geq N$.

This notion is consistent with that of convergence in metric spaces (Exercise 2.5.2). One can then ask whether one has the basic property of uniqueness of limits (Proposition 1.1.20). The answer turns out to usually be yes—if the topological space has an additional property known as the *Hausdorff property*—but the answer can be no for other topologies; see Exercise 2.5.4.

Definition 2.5.5 (*Interior, exterior, boundary*) Let $(X, \mathcal{F})$ be a topological space, let E be a subset of X, and let x_0 be a point in X. We say that x_0 is an *interior point of* E if there exists a neighborhood V of x_0 such that $V \subseteq E$. We say that x_0 is an *exterior point of* E if there exists a neighborhood V of x_0 such that $V \cap E = \emptyset$. We say that x_0 is a *boundary point of* E if it is neither an interior point nor an exterior point of E.

This definition is consistent with the corresponding notion for metric spaces (Exercise 2.5.3).

Definition 2.5.6 (*Closure*) Let $(X, \mathcal{F})$ be a topological space, let E be a subset of X, and let x_0 be a point in X. We say that x_0 is an *adherent point* of E if every neighborhood V of x_0 has a non-empty intersection with E. The set of all adherent points of E is called the *closure* of E and is denoted $\overline{E}$.

There is a partial analogue of Theorem 1.2.10, see Exercise 2.5.9.

We define a set K in a topological space $(X, \mathcal{F})$ to be *closed* iff its complement $X \backslash K$ is open; this is consistent with the metric space definition, thanks to Proposition 1.2.15(e). Some partial analogues of that proposition are true (see Exercise 2.5.10).

To define the notion of a relative topology, we cannot use Definition 1.3.3 as this requires a metric function. However, we can instead use Proposition 1.3.4 as our starting point:

Definition 2.5.7 (*Relative topology*) Let $(X, \mathcal{F})$ be a topological space, and Y be a subset of X. Then we define $\mathcal{F}_Y := \{V \cap Y : V \in \mathcal{F}\}$ and refer this as the topology on Y *induced* by $(X, \mathcal{F})$. We call $(Y, \mathcal{F}_Y)$ a *topological subspace* of $(X, \mathcal{F})$. This is indeed a topological space, see Exercise 2.5.11.

From Proposition 1.3.4 we see that this notion is compatible with the one for metric spaces.

Next we define the notion of continuity.

Definition 2.5.8 (*Continuous functions*) Let $(X, \mathcal{F}_X)$ and $(Y, \mathcal{F}_Y)$ be topological spaces, and let $f : X \to Y$ be a function. If $x_0 \in X$, we say that f is *continuous at* x_0 iff for every neighborhood V of $f(x_0)$, there exists a neighborhood U of x_0 such that $f(U) \subseteq V$. We say that f is *continuous* iff it is continuous at every point $x \in X$.

This definition is consistent with that in Definition 2.1.1 (Exercise 2.5.14). Partial analogues of Theorems 2.1.4 and 2.1.5 are available (Exercise 2.5.15). In particular, a function is continuous iff the pre-images of every open set are open.

There is unfortunately no notion of a Cauchy sequence, a complete space, or a bounded space, for general topological spaces. However, there is certainly a notion of a compact space, as we can see by taking Theorem 1.5.8 as our starting point:

Definition 2.5.9 (*Compact topological spaces*) Let $(X, \mathcal{F})$ be a topological space. We say that this space is *compact* if every open cover of X has a finite subcover. If Y is a subset of X, we say that Y is compact if the topological space on Y induced by $(X, \mathcal{F})$ is compact.

Many basic facts about compact metric spaces continue to hold true for compact topological spaces, notably Theorem 2.3.1 and Proposition 2.3.2 (Exercise 2.5.16). However, there is no notion of uniform continuity, and so there is no analogue of Theorem 2.3.5.

We can also define the notion of connectedness by repeating Definition 2.4.1 verbatim and also repeating Definition 2.4.3 (but with Definition 2.5.7 instead of Definition 1.3.3). Many of the results and exercises in Sect. 2.4 continue to hold for topological spaces (with almost no changes to any of the proofs!).

— Exercises —

Exercise 2.5.1 Let X be an arbitrary set, and let $\mathcal{F} := \{\emptyset, X\}$. Show that $(X, \mathcal{F})$ is a topology (called the *trivial topology* on X). If X contains more than one element, show that the trivial topology cannot be obtained from by placing a metric d on X. Show that this topological space is both compact and connected.

Exercise 2.5.2 Let (X, d) be a metric space (and hence a topological space). Show that the two notions of convergence of sequences in Definition 1.1.14 and Definition 2.5.4 coincide.

Exercise 2.5.3 Let (X, d) be a metric space (and hence a topological space). Show that the two notions of interior, exterior, and boundary in Definition 1.2.5 and Definition 2.5.5 coincide.

Exercise 2.5.4 A topological space $(X, \mathcal{F})$ is said to be *Hausdorff* if given any two distinct points $x, y \in X$, there exists a neighborhood V of x and a neighborhood W of y such that $V \cap W = \emptyset$. Show that any topological space coming from a metric space is Hausdorff, and show that the trivial topology is not Hausdorff. Show that the analogue of Proposition 1.1.20 holds for Hausdorff topological spaces, but give an example of a non-Hausdorff topological space in which Proposition 1.1.20 fails. (In practice, most topological spaces one works with are Hausdorff; non-Hausdorff topological spaces tend to be so pathological that it is not very profitable to work with them.)

Exercise 2.5.5 Given any totally ordered set X with order relation $\leq$, declare a set $V \subseteq X$ to be *open* if for every $x \in V$ there exists a set I which is an interval $\{y \in X : a < y < b\}$ for some $a, b \in X$, a ray $\{y \in X : a < y\}$ for some $a \in X$, the ray $\{y \in X : y < b\}$ for some $b \in X$, or the whole space X, which contains x and is contained in V. Let $\mathcal{F}$ be the set of all open subsets of X. Show that $(X, \mathcal{F})$ is a topology (this is the *order topology* on the totally ordered set $(X, \leq)$) which is Hausdorff in the sense of Exercise 2.5.4. Show that on the real line $\mathbf{R}$ (with the standard ordering $\leq$), the order topology matches the standard topology (i.e., the topology arising from the standard metric). If instead one applies this to the extended real line $\mathbf{R}^*$, show that $\mathbf{R}$ is an open set with boundary $\{-\infty, +\infty\}$. If $(x_n)_{n=1}^{\infty}$ is a sequence of numbers in $\mathbf{R}$ (and hence in $\mathbf{R}^*$), show that x_n converges to $+\infty$ if and only if $\liminf_{n\to\infty} x_n = +\infty$, and x_n converges to $-\infty$ if and only if $\limsup_{n\to\infty} x_n = -\infty$.

Exercise 2.5.6 Let X be an uncountable set, and let $\mathcal{F}$ be the collection of all subsets E in X which are either empty or cofinite (which means that $X \backslash E$ is finite). Show that $(X, \mathcal{F})$ is a topology (this is called the *cofinite topology* on X) which is not Hausdorff in the sense of Exercise 2.5.4 and is compact and connected. Also, show that if $x \in X$ $(V_n)_{n=1}^{\infty}$ is any countable collection of open sets containing x, then $\bigcap_{n=1}^{\infty} V_n \neq \{x\}$. Use this to show that the cofinite topology cannot be obtained by placing a metric d on X. (*Hint:* what is the set $\bigcap_{n=1}^{\infty} B(x, 1/n)$ equal to in a metric space?)

Exercise 2.5.7 Let X be an uncountable set, and let $\mathcal{F}$ be the collection of all subsets E in X which are either empty or cocountable (which means that $X \backslash E$ is at most countable). Show that $(X, \mathcal{F})$ is a topology (this is called the *cocountable topology* on X) which is not Hausdorff in the sense of Exercise 2.5.4, and connected, but cannot arise from a metric space and is not compact.

Exercise 2.5.8 Let $(X, \mathcal{F})$ be a compact topological space. Assume that this space is *first countable*, which means that for every $x \in X$ there exists a countable collection $V_1, V_2, \ldots$ of neighborhoods of x, such that every neighborhood of x contains one of the V_n. Show that every sequence in X has a convergent subsequence, by modifying Exercise 1.5.11.

Exercise 2.5.9 Prove the following partial analogue of Proposition 1.2.10 for topological spaces: (c) implies both (a) and (b), which are equivalent to each other. Show that in the cocountable topology in Exercise 2.5.7, it is possible for (a) and (b) to hold without (c) holding.

Exercise 2.5.10 Let E be a subset of a topological space $(X, \mathcal{F})$. Show that E is open if and only if every element of E is an interior point, and show that E is closed if and only if E contains all of its adherent points. Prove analogues of Proposition 1.2.15(e)-(h) (some of these are automatic by definition). If we assume in addition that X is Hausdorff, prove an analogue of Proposition 1.2.15(d) also, but give an example to show that (d) can fail when X is not Hausdorff.

Exercise 2.5.11 Show that the pair $(Y, \mathcal{F}_Y)$ defined in Definition 2.5.7 is indeed a topological space.

Exercise 2.5.12 Generalize Corollary 1.5.9 to compact sets in a Hausdorff topological space.

Exercise 2.5.13 Generalize Theorem 1.5.10 to compact sets in a Hausdorff topological space.

Exercise 2.5.14 Let (X, d_X) and (Y, d_Y) be metric spaces (and hence a topological space). Show that the two notions continuity (both at a point, and on the whole domain) of a function $f : X \to Y$ in Definition 2.1.1 and Definition 2.5.8 coincide.

Exercise 2.5.15 Show that when Theorem 2.1.4 is extended to topological spaces, that (a) implies (b). (The converse is false, but constructing an example is difficult.) Show that when Theorem 2.1.5 is extended to topological spaces, that (a), (c), (d) are all equivalent to each other and imply (b). (Again, the converse implications are false, but difficult to prove.)

Exercise 2.5.16 Generalize both Theorem 2.3.1 and Proposition 2.3.2 to compact sets in a topological space.

Chapter 3
Uniform Convergence

In the previous two chapters we have seen what it means for a sequence $(x^{(n)})_{n=1}^{\infty}$ of points in a metric space (X, d_X) to converge to a limit x; it means that $\lim_{n\to\infty} d_X(x^{(n)}, x) = 0$, or equivalently that for every $\varepsilon > 0$ there exists an $N > 0$ such that $d_X(x^{(n)}, x) < \varepsilon$ for all $n > N$. (We have also generalized the notion of convergence to topological spaces $(X, \mathcal{F})$, but in this chapter we will focus on metric spaces.)

In this chapter, we consider what it means for a sequence of *functions* $(f^{(n)})_{n=1}^{\infty}$ from one metric space (X, d_X) to another (Y, d_Y) to converge. In other words, we have a sequence of functions $f^{(1)}, f^{(2)}, \ldots$, with each function $f^{(n)}: X \to Y$ being a function from X to Y, and we ask what it means for this sequence of functions to converge to some limiting function f.

It turns out that there are several different concepts of convergence of functions; here we describe the two most important ones, *pointwise convergence* and *uniform convergence*. (There are other types of convergence for functions, such as L^1 convergence, L^2 convergence, convergence in measure, almost everywhere convergence, and so forth, but these are beyond the scope of this text.) The two notions are related, but not identical; the relationship between the two is somewhat analogous to the relationship between continuity and uniform continuity.

Once we work out what convergence means for functions, and thus can make sense of such statements as $\lim_{n\to\infty} f^{(n)} = f$, we will then ask how these limits interact with other concepts. For instance, we already have a notion of limiting values of functions: $\lim_{x\to x_0; x\in X} f(x)$. Can we interchange limits, i.e.,

$$\lim_{n\to\infty} \lim_{x\to x_0; x\in X} f^{(n)}(x) = \lim_{x\to x_0; x\in X} \lim_{n\to\infty} f^{(n)}(x)?$$

As we shall see, the answer depends on what type of convergence we have for $f^{(n)}$. We will also address similar questions involving interchanging limits and integrals, or limits and sums, or sums and integrals.

T. Tao, *Analysis II*, Texts and Readings in Mathematics,
https://doi.org/10.1007/978-981-19-7284-3_3

3.1 Limiting Values of Functions

Before we talk about limits of sequences of functions, we should first discuss a similar, but distinct, notion, that of limiting values of functions. We shall focus on the situation for metric spaces, but there are similar notions for topological spaces (Exercise 3.1.3).

Definition 3.1.1 (*Limiting value of a function*) Let (X, d_X) and (Y, d_Y) be metric spaces, let E be a subset of X, and let $f : E \to Y$ be a function. If $x_0 \in X$ is an adherent point of E, and $L \in Y$, we say that $f(x)$ *converges to* L *in* Y *as* x *converges to* x_0 *in* E, or write $\lim_{x\to x_0; x\in E} f(x) = L$, if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that $d_Y(f(x), L) < \varepsilon$ for all $x \in E$ such that $d_X(x, x_0) < \delta$.

Remark 3.1.2 Some authors exclude the case $x = x_0$ from the above definition, thus requiring $0 < d_X(x, x_0) < \delta$. In our current notation, this would correspond to removing x_0 from E, thus one would consider $\lim_{x\to x_0; x\in E\setminus\{x_0\}} f(x)$ instead of $\lim_{x\to x_0; x\in E} f(x)$. See Exercise 3.1.1 for a comparison of the two concepts.

Comparing this with Definition 2.1.1, we see that f is continuous at x_0 if and only if

$$\lim_{x\to x_0; x\in X} f(x) = f(x_0).$$

Thus f is continuous on X if we have

$$\lim_{x\to x_0; x\in X} f(x) = f(x_0) \text{ for all } x_0 \in X.$$

Example 3.1.3 If $f : \mathbf{R} \to \mathbf{R}$ is the function $f(x) = x^2 - 4$, then

$$\lim_{x\to 1} f(x) = f(1) = 1 - 4 = -3$$

since f is continuous.

Remark 3.1.4 Often we shall omit the condition $x \in X$, and abbreviate $\lim_{x\to x_0; x\in X} f(x)$ as simply $\lim_{x\to x_0} f(x)$ when it is clear what space x will range in.

One can rephrase Definition 3.1.1 in terms of sequences:

Proposition 3.1.5 *Let (X, d_X) and (Y, d_Y) be metric spaces, let E be a subset of X, and let $f : E \to Y$ be a function. Let $x_0 \in X$ be an adherent point of E and $L \in Y$. Then the following four statements are logically equivalent:*

(a) $\lim_{x\to x_0; x\in E} f(x) = L$.
(b) For every sequence $(x^{(n)})_{n=1}^\infty$ in E which converges to x_0 with respect to the metric d_X, the sequence $(f(x^{(n)}))_{n=1}^\infty$ converges to L with respect to the metric d_Y.

(c) *For every open set $V \subseteq Y$ which contains L, there exists an open set $U \subseteq X$ containing x_0 such that $f(U \cap E) \subseteq V$.*

(d) *If one defines the function $g\colon E \cup \{x_0\} \to Y$ by defining $g(x_0) := L$, and $g(x) := f(x)$ for $x \in E \backslash \{x_0\}$, then g is continuous at x_0. Furthermore, if $x_0 \in E$, then $f(x_0) = L$.*

Proof See Exercise 3.1.2. □

Remark 3.1.6 Observe from Propositions 3.1.5(b) and 1.1.20 that a function $f(x)$ can converge to at most one limit L as x converges to x_0. In other words, if the limit

$$\lim_{x \to x_0; x \in E} f(x)$$

exists at all, then it can only take at most one value.

Remark 3.1.7 The requirement that x_0 be an adherent point of E is necessary for the concept of limiting value to be useful, otherwise x_0 will lie in the exterior of E, the notion that $f(x)$ converges to L as x converges to x_0 in E is vacuous (for δ sufficiently small, there are no points $x \in E$ so that $d(x, x_0) < \delta$).

Remark 3.1.8 Strictly speaking, we should write

$$d_Y - \lim_{x \to x_0; x \in E} f(x) \text{ instead of } \lim_{x \to x_0; x \in E} f(x),$$

since the convergence depends on the metric d_Y. However in practice it will be obvious what the metric d_Y is and so we will omit the d_Y– prefix from the notation.

— Exercises —

Exercise 3.1.1 Let (X, d_X) and (Y, d_Y) be metric spaces, let E be a subset of X, let $f\colon E \to Y$ be a function, and let x_0 be an element of E. Assume that x_0 is an adherent point of $E \backslash \{x_0\}$ (or equivalently, that x_0 is not an *isolated point* of E). Show that the limit $\lim_{x \to x_0; x \in E} f(x)$ exists if and only if the limit $\lim_{x \to x_0; x \in E \backslash \{x_0\}} f(x)$ exists and is equal to $f(x_0)$. Also, show that if the limit $\lim_{x \to x_0; x \in E} f(x)$ exists at all, then it must equal $f(x_0)$.

Exercise 3.1.2 Prove Proposition 3.1.5. (*Hint:* review your proof of Theorem 2.1.4.)

Exercise 3.1.3 Use Proposition 3.1.5(c) to define a notion of a limiting value of a function $f\colon E \to Y$ from one topological space $(X, \mathcal{F}_X)$ to another $(Y, \mathcal{F}_Y)$, with E a subset of X. If X is a topological space and Y is a Hausdorff topological space (see Exercise 2.5.4), prove the equivalence of Proposition 3.1.5(c) and (d), as well as an analogue of Remark 3.1.6. What happens to these statements if Y is not assumed to be Hausdorff?

Exercise 3.1.4 Recall from Exercise 2.5.5 that the extended real line $\mathbf{R}^*$ comes with a standard topology (the order topology). We view the natural numbers $\mathbf{N}$ as a subspace of this topological space, and $+\infty$ as an adherent point of $\mathbf{N}$ in $\mathbf{R}^*$. Let $(a_n)_{n=0}^{\infty}$ be a sequence taking values in a topological space $(Y, \mathcal{F}_Y)$, and let $L \in Y$. Show that $\lim_{n \to +\infty; n \in \mathbf{N}} a_n = L$ (in the sense of Exercise 3.1.3) if and only if $\lim_{n \to \infty} a_n = L$ (in the sense of Definition 2.5.4). This shows that the notions of limiting values of a sequence, and limiting values of a function, are compatible.

Exercise 3.1.5 Let (X, d_X), (Y, d_Y), (Z, d_Z) be metric spaces, let E be a subset of X, and let $x_0 \in X$, $y_0 \in Y$, $z_0 \in Z$. Let $f : E \to Y$ and $g : Y \to Z$ be functions, and let E be a set. If we have $\lim_{x \to x_0; x \in E} f(x) = y_0$ and $\lim_{y \to y_0; y \in f(E)} g(y) = z_0$, conclude that $\lim_{x \to x_0; x \in E} g \circ f(x) = z_0$.

Exercise 3.1.6 State and prove an analogue of the limit laws in Proposition 9.3.14 when X is now a metric space rather than a subset of $\mathbf{R}$. (*Hint:* use Corollary 2.2.3.)

3.2 Pointwise and Uniform Convergence

The most obvious notion of convergence of functions is *pointwise convergence*, or convergence at each point of the domain:

Definition 3.2.1 (*Pointwise convergence*) Let $(f^{(n)})_{n=1}^{\infty}$ be a sequence of functions from one metric space (X, d_X) to another (Y, d_Y), and let $f : X \to Y$ be another function. We say that $(f^{(n)})_{n=1}^{\infty}$ *converges pointwise to f on X* if we have

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

for all $x \in X$, i.e.,

$$\lim_{n \to \infty} d_Y(f^{(n)}(x), f(x)) = 0.$$

Or in other words, for every x and every $\varepsilon > 0$ there exists $N > 0$ such that $d_Y(f^{(n)}(x), f(x)) < \varepsilon$ for every $n > N$. We call the function f the *pointwise limit* of the functions $f^{(n)}$.

Remark 3.2.2 Note that $f^{(n)}(x)$ and $f(x)$ are points in Y, rather than functions, so we are using our prior notion of convergence in metric spaces to determine convergence of functions. Also note that we are not really using the fact that (X, d_X) is a metric space (i.e., we are not using the metric d_X); for this definition it would suffice for X to just be a plain old set with no metric structure. However, later on we shall want to restrict our attention to *continuous* functions from X to Y, and in order to do so we need a metric on X (and on Y), or at least a topological structure. Also when we introduce the concept of *uniform convergence*, then we will definitely need a metric structure on X and Y; there is no comparable notion for topological spaces.

Example 3.2.3 Consider the functions $f^{(n)}: \mathbf{R} \to \mathbf{R}$ defined by $f^{(n)}(x) := x/n$, while $f: \mathbf{R} \to \mathbf{R}$ is the zero function $f(x) := 0$. Then $f^{(n)}$ converges pointwise to f, since for each fixed real number x we have $\lim_{n\to\infty} f^{(n)}(x) = \lim_{n\to\infty} x/n = 0 = f(x)$.

From Proposition 1.1.20 we see that a sequence $(f^{(n)})_{n=1}^{\infty}$ of functions from one metric space (X, d_X) to another (Y, d_Y) can have at most one pointwise limit f (this explains why we can refer to f as *the* pointwise limit). However, it is of course possible for a sequence of functions to have no pointwise limit (can you think of an example?), just as a sequence of points in a metric space do not necessarily have a limit.

Pointwise convergence is a very natural concept, but it has a number of disadvantages: it does not preserve continuity, derivatives, limits, or integrals, as the following three examples show.

Example 3.2.4 Consider the functions $f^{(n)}: [0, 1] \to \mathbf{R}$ defined by $f^{(n)}(x) := x^n$, and let $f: [0, 1] \to \mathbf{R}$ be the function defined by setting $f(x) := 1$ when $x = 1$ and $f(x) := 0$ when $0 \le x < 1$. Then the functions $f^{(n)}$ are continuous, and converge pointwise to f on $[0, 1]$ (why? Treat the cases $x = 1$ and $0 \le x < 1$ separately), however the limiting function f is not continuous. Note that the same example shows that pointwise convergence does not preserve differentiability either.

Example 3.2.5 If $\lim_{x\to x_0; x\in E} f^{(n)}(x) = L$ for every n, and $f^{(n)}$ converges pointwise to f, we cannot always take limits conclude that $\lim_{x\to x_0; x\in E} f(x) = L$. The previous example is also a counterexample here: observe that $\lim_{x\to 1; x\in[0,1)} x^n = 1$ for every n, but x^n converges pointwise to the function f defined in the previous paragraph, and $\lim_{x\to 1; x\in[0,1)} f(x) = 0$. In particular, we see that

$$\lim_{n\to\infty} \lim_{x\to x_0; x\in X} f^{(n)}(x) \neq \lim_{x\to x_0; x\in X} \lim_{n\to\infty} f^{(n)}(x).$$

(cf. Example 1.2.8). Thus pointwise convergence does not preserve limits.

Example 3.2.6 Suppose that $f^{(n)}: [a, b] \to \mathbf{R}$ a sequence of Riemann-integrable functions on the interval $[a, b]$. If $\int_{[a,b]} f^{(n)} = L$ for every n, and $f^{(n)}$ converges pointwise to some new function f, this does not mean that $\int_{[a,b]} f = L$. An example comes by setting $[a, b] := [0, 1]$, and letting $f^{(n)}$ be the function $f^{(n)}(x) := 2n$ when $x \in [1/2n, 1/n]$, and $f^{(n)}(x) := 0$ for all other values of x. Then $f^{(n)}$ converges pointwise to the zero function $f(x) := 0$ (why?). On the other hand, $\int_{[0,1]} f^{(n)} = 1$ for every n, while $\int_{[0,1]} f = 0$. In particular, we have an example where

$$\lim_{n\to\infty} \int_{[a,b]} f^{(n)} \neq \int_{[a,b]} \lim_{n\to\infty} f^{(n)}.$$

One may think that this counterexample has something to do with the $f^{(n)}$ being discontinuous, but one can easily modify this counterexample to make the $f^{(n)}$ continuous (can you see how?).

Another example in the same spirit is the "moving bump" example. Let $f^{(n)}: \mathbf{R} \to \mathbf{R}$ be the function defined by $f^{(n)}(x) := 1$ if $x \in [n, n+1]$ and $f^{(n)}(x) := 0$ otherwise. Then $\int_{\mathbf{R}} f^{(n)} = 1$ for every n (where $\int_{\mathbf{R}} f$ is defined as the limit of $\int_{[-N,N]} f$ as N goes to infinity). On the other hand, $f^{(n)}$ converges pointwise to the zero function 0 (why?), and $\int_{\mathbf{R}} 0 = 0$. In both of these examples, functions of area 1 have somehow "disappeared" to produce functions of area 0 in the limit. See also Example 1.2.9.

These examples show that pointwise convergence is too weak a concept to be of much use. The problem is that while $f^{(n)}(x)$ converges to $f(x)$ for each x, the *rate* of that convergence varies substantially with x. For instance, consider the first example where $f^{(n)}: [0, 1] \to \mathbf{R}$ was the function $f^{(n)}(x) := x^n$, and $f: [0, 1] \to \mathbf{R}$ was the function such that $f(x) := 1$ when $x = 1$, and $f(x) := 0$ otherwise. Then for each x, $f^{(n)}(x)$ converges to $f(x)$ as $n \to \infty$; this is the same as saying that $\lim_{n\to\infty} x^n = 0$ when $0 \le x < 1$, and that $\lim_{n\to\infty} x^n = 1$ when $x = 1$. But the convergence is much slower near 1 than far away from 1. For instance, consider the statement that $\lim_{n\to\infty} x^n = 0$ for all $0 \le x < 1$. This means, for every $0 \le x < 1$, that for every ε, there exists an $N \ge 1$ such that $|x^n| < \varepsilon$ for all $n \ge N$—or in other words, the sequence $1, x, x^2, x^3, \ldots$ will eventually get less than ε, after passing some finite number N of elements in this sequence. But the number of elements N one needs to go out to depends very much on the location of x. For instance, take $\varepsilon := 0.1$. If $x = 0.1$, then we have $|x^n| < \varepsilon$ for all $n \ge 2$—the sequence gets underneath ε after the second element. But if $x = 0.5$, then we only get $|x^n| < \varepsilon$ for $n \ge 4$—you have to wait until the fourth element to get within ε of the limit. And if $x = 0.9$, then one only has $|x^n| < \varepsilon$ when $n \ge 22$. Clearly, the closer x gets to 1, the longer one has to wait until $f^{(n)}(x)$ will get within ε of $f(x)$, although it still will get there eventually. (Curiously, however, while the convergence gets worse and worse as x approaches 1, the convergence suddenly becomes perfect when $x = 1$.)

To put things another way, the convergence of $f^{(n)}$ to f is not *uniform* in x—the N that one needs to get $f^{(n)}(x)$ within ε of f depends on x as well as on ε. This motivates a stronger notion of convergence.

Definition 3.2.7 (*Uniform convergence*) Let $(f^{(n)})_{n=1}^{\infty}$ be a sequence of functions from one metric space (X, d_X) to another (Y, d_Y), and let $f: X \to Y$ be another function. We say that $(f^{(n)})_{n=1}^{\infty}$ *converges uniformly to* f *on* X if for every $\varepsilon > 0$ there exists $N > 0$ such that $d_Y(f^{(n)}(x), f(x)) < \varepsilon$ for every $n > N$ and $x \in X$. We call the function f the *uniform limit* of the functions $f^{(n)}$.

Remark 3.2.8 Note that this definition is subtly different from the definition for pointwise convergence in Definition 3.2.1. In the definition of pointwise convergence, N was allowed to depend on x; now it is not. The reader should compare this distinction to the distinction between continuity and uniform continuity (i.e., between Definitions 2.1.1 and 2.3.4). A more precise formulation of this analogy is given in Exercise 3.2.1.

It is easy to see that if $f^{(n)}$ converges uniformly to f on X, then it also converges pointwise to the same function f (see Exercise 3.2.2); thus when the uniform limit and pointwise limit both exist, then they have to be equal. However, the converse is not true; for instance the functions $f^{(n)}: [0, 1] \to \mathbf{R}$ defined earlier by $f^{(n)}(x) := x^n$ converge pointwise, but do not converge uniformly (see Exercise 3.2.2).

Example 3.2.9 Let $f^{(n)}: [0, 1] \to \mathbf{R}$ be the functions $f^{(n)}(x) := x/n$, and let $f: [0, 1] \to \mathbf{R}$ be the zero function $f(x) := 0$. Then it is clear that $f^{(n)}$ converges to f pointwise. Now we show that in fact $f^{(n)}$ converges to f uniformly. We have to show that for every $\varepsilon > 0$, there exists an N such that $|f^{(n)}(x) - f(x)| < \varepsilon$ for every $x \in [0, 1]$ and every $n \geq N$. To show this, let us fix an $\varepsilon > 0$. Then for any $x \in [0, 1]$ and $n \geq N$, we have

$$|f^{(n)}(x) - f(x)| = |x/n - 0| = x/n \leq 1/n \leq 1/N.$$

Thus if we choose N such that $N > 1/\varepsilon$ (note that this choice of N does not depend on what x is), then we have $|f^{(n)}(x) - f(x)| < \varepsilon$ for all $n \geq N$ and $x \in [0, 1]$, as desired.

We make one trivial remark here: if a sequence $f^{(n)}: X \to Y$ of functions converges pointwise (or uniformly) to a function $f: X \to Y$, then the restrictions $f^{(n)}|_E: E \to Y$ of $f^{(n)}$ to some subset E of X will also converge pointwise (or uniformly) to $f|_E$. (Why?)

— Exercises —

Exercise 3.2.1 The purpose of this exercise is to demonstrate a concrete relationship between continuity and pointwise convergence, and between uniform continuity and uniform convergence. Let $f : \mathbf{R} \to \mathbf{R}$ be a function. For any $a \in \mathbf{R}$, let $f_a : \mathbf{R} \to \mathbf{R}$ be the shifted function $f_a(x) := f(x - a)$.

(a) Show that f is continuous if and only if, whenever $(a_n)_{n=0}^\infty$ is a sequence of real numbers which converges to zero, the shifted functions f_{a_n} converge pointwise to f.
(b) Show that f is uniformly continuous if and only if, whenever $(a_n)_{n=0}^\infty$ is a sequence of real numbers which converges to zero, the shifted functions f_{a_n} converge uniformly to f.

Exercise 3.2.2 (a) Let $(f^{(n)})_{n=1}^\infty$ be a sequence of functions from one metric space (X, d_X) to another (Y, d_Y), and let $f: X \to Y$ be another function from X to Y. Show that if $f^{(n)}$ converges uniformly to f, then $f^{(n)}$ also converges pointwise to f.
(b) For each integer $n \geq 1$, let $f^{(n)}: (-1, 1) \to \mathbf{R}$ be the function $f^{(n)}(x) := x^n$. Prove that $f^{(n)}$ converges pointwise to the zero function 0, but does not converge uniformly to any function $f: (-1, 1) \to \mathbf{R}$.
(c) Let $g: (-1, 1) \to \mathbf{R}$ be the function $g(x) := x/(1 - x)$. With the notation as in (b), show that the partial sums $\sum_{n=1}^N f^{(n)}$ converge pointwise as $N \to \infty$ to g, but does not converge uniformly to g, on the open interval $(-1, 1)$. (*Hint:* use Lemma 7.3.3.) What would happen if we replaced the open interval $(-1, 1)$ with the closed interval $[-1, 1]$?

Exercise 3.2.3 Let (X, d_X) a metric space, and for every integer $n \geq 1$, let $f_n : X \to \mathbf{R}$ be a real-valued function. Suppose that f_n converges pointwise to another function $f : X \to \mathbf{R}$ on X (in this question we give $\mathbf{R}$ the standard metric $d(x, y) = |x - y|$). Let $h : \mathbf{R} \to \mathbf{R}$ be a continuous function. Show that the functions $h \circ f_n$ converge pointwise to $h \circ f$ on X, where $h \circ f_n : X \to \mathbf{R}$ is the function $h \circ f_n(x) := h(f_n(x))$, and similarly for $h \circ f$.

Exercise 3.2.4 Let $f_n : X \to Y$ be a sequence of bounded functions from one metric space (X, d_X) to another metric space (Y, d_Y). Suppose that f_n converges uniformly to another function $f : X \to Y$. Suppose that f is a bounded function; i.e., there exists a ball $B_{(Y,d_Y)}(y_0, R)$ in Y such that $f(x) \in B_{(Y,d_Y)}(y_0, R)$ for all $x \in X$. Show that the sequence f_n is *uniformly bounded*; i.e., there exists a ball $B_{(Y,d_Y)}(y_0, R)$ in Y such that $f_n(x) \in B_{(Y,d_Y)}(y_0, R)$ for all $x \in X$ and all positive integers n.

3.3 Uniform Convergence and Continuity

We now give the first demonstration that uniform convergence is significantly better than pointwise convergence. Specifically, we show that the uniform limit of continuous functions is continuous.

Theorem 3.3.1 (Uniform limits preserve continuity I) *Suppose $(f^{(n)})_{n=1}^{\infty}$ is a sequence of functions from one metric space (X, d_X) to another (Y, d_Y), and suppose that this sequence converges uniformly to another function $f : X \to Y$. Let x_0 be a point in X. If the functions $f^{(n)}$ are continuous at x_0 for each n, then the limiting function f is also continuous at x_0.*

Proof See Exercise 3.3.1. □

This has an immediate corollary:

Corollary 3.3.2 (Uniform limits preserve continuity II) *Let $(f^{(n)})_{n=1}^{\infty}$ be a sequence of functions from one metric space (X, d_X) to another (Y, d_Y), and suppose that this sequence converges uniformly to another function $f : X \to Y$. If the functions $f^{(n)}$ are continuous on X for each n, then the limiting function f is also continuous on X.*

This should be contrasted with Example 3.2.4. There is a slight variant of Theorem 3.3.1 which is also useful:

Proposition 3.3.3 (Interchange of limits and uniform limits) *Let (X, d_X) and (Y, d_Y) be metric spaces, with Y complete, and let E be a subset of X. Let $(f^{(n)})_{n=1}^{\infty}$ be a sequence of functions from E to Y, and suppose that this sequence converges uniformly in E to some function $f : E \to Y$. Let $x_0 \in X$ be an adherent point of E, and suppose that for each n the limit $\lim_{x \to x_0; x \in E} f^{(n)}(x)$ exists. Then* *the limit* $\lim_{x \to x_0; x \in E} f(x)$ *also exists, and is equal to the limit of the sequence* $(\lim_{x \to x_0; x \in E} f^{(n)}(x))_{n=1}^{\infty}$*; in other words we have the interchange of limits*

$$\lim_{n \to \infty} \lim_{x \to x_0; x \in E} f^{(n)}(x) = \lim_{x \to x_0; x \in E} \lim_{n \to \infty} f^{(n)}(x).$$

Proof See Exercise 3.3.2. □

This should be contrasted with Example 3.2.5. Finally, we have a version of these theorems for sequences:

Proposition 3.3.4 *Let* $(f^{(n)})_{n=1}^{\infty}$ *be a sequence of continuous functions from one metric space* (X, d_X) *to another* (Y, d_Y)*, and suppose that this sequence converges uniformly to another function* $f : X \to Y$*. Let* $x^{(n)}$ *be a sequence of points in* X *which converge to some limit* x*. Then* $f^{(n)}(x^{(n)})$ *converges (in* Y*) to* $f(x)$*.*

Proof See Exercise 3.3.4. □

A similar result holds for bounded functions:

Definition 3.3.5 (*Bounded functions*) A function $f : X \to Y$ from one metric space (X, d_X) to another (Y, d_Y) is *bounded* if $f(X)$ is a bounded set, i.e., there exists a ball $B_{(Y,d_Y)}(y_0, R)$ in Y such that $f(x) \in B_{(Y,d_Y)}(y_0, R)$ for all $x \in X$.

Proposition 3.3.6 (Uniform limits preserve boundedness) *Let* $(f^{(n)})_{n=1}^{\infty}$ *be a sequence of functions from one metric space* (X, d_X) *to another* (Y, d_Y)*, and suppose that this sequence converges uniformly to another function* $f : X \to Y$*. If the functions* $f^{(n)}$ *are bounded on* X *for each* n*, then the limiting function* f *is also bounded on* X*.*

Proof See Exercise 3.3.6. □

Remark 3.3.7 The above propositions sound very reasonable, but one should caution that it only works if one assumes uniform convergence; pointwise convergence is not enough. (See Exercises 3.3.3, 3.3.5 and 3.3.7.)

— Exercises —

Exercise 3.3.1 Prove Theorem 3.3.1. Explain briefly why your proof requires uniform convergence, and why pointwise convergence would not suffice. (*Hints:* it is easiest to use the "epsilon-delta" definition of continuity from Definition 2.1.1. You may find the triangle inequality

$$\begin{aligned} d_Y(f(x), f(x_0)) \leq\ & d_Y(f(x), f^{(n)}(x)) + d_Y(f^{(n)}(x), f^{(n)}(x_0)) \\ & + d_Y(f^{(n)}(x_0), f(x_0)) \end{aligned}$$

useful. Also, you may need to divide ε as $\varepsilon = \varepsilon/3 + \varepsilon/3 + \varepsilon/3$. Finally, it is possible to prove Theorem 3.3.1 from Proposition 3.3.3, but you may find it easier conceptually to prove Theorem 3.3.1 first.)

Exercise 3.3.2 Prove Proposition 3.3.3. (*Hint:* this is very similar to Theorem 3.3.1. Theorem 3.3.1 cannot be used to prove Proposition 3.3.3, however it is possible to use Proposition 3.3.3 to prove Theorem 3.3.1.)

Exercise 3.3.3 Compare Proposition 3.3.3 with Example 1.2.8. Can you now explain why the interchange of limits in Example 1.2.8 led to a false statement, whereas the interchange of limits in Proposition 3.3.3 is justified?

Exercise 3.3.4 Prove Proposition 3.3.4. (*Hint:* again, this is similar to Theorem 3.3.1 and Proposition 3.3.3, although the statements are slightly different, and one cannot deduce this directly from the other two results.)

Exercise 3.3.5 Give an example to show that Proposition 3.3.4 fails if the phrase "converges uniformly" is replaced by "converges pointwise". (*Hint:* some of the examples already given earlier will already work here.)

Exercise 3.3.6 Prove Proposition 3.3.6. Discuss how this proposition differs from Exercise 3.2.4.

Exercise 3.3.7 Give an example to show that Proposition 3.3.6 fails if the phrase "converges uniformly" is replaced by "converges pointwise". (*Hint:* some of the examples already given earlier will already work here.)

Exercise 3.3.8 Let (X, d) be a metric space, and for every positive integer n, let $f_n : X \to \mathbf{R}$ and $g_n : X \to \mathbf{R}$ be functions. Suppose that $(f_n)_{n=1}^\infty$ converges uniformly to another function $f : X \to \mathbf{R}$, and that $(g_n)_{n=1}^\infty$ converges uniformly to another function $g : X \to \mathbf{R}$. Suppose also that the functions $(f_n)_{n=1}^\infty$ and $(g_n)_{n=1}^\infty$ are uniformly bounded, i.e., there exists an $M > 0$ such that $|f_n(x)| \leq M$ and $|g_n(x)| \leq M$ for all $n \geq 1$ and $x \in X$. Prove that the functions $f_n g_n : X \to \mathbf{R}$ converge uniformly to $fg : X \to \mathbf{R}$.

3.4 The Metric of Uniform Convergence

We have now developed at least four, apparently separate, notions of limit in this text:

(a) limits $\lim_{n\to\infty} x^{(n)}$ of sequences of points in a metric space (Definition 1.1.14; see also Definition 2.5.4);
(b) limiting values $\lim_{x\to x_0; x\in E} f(x)$ of functions at a point (Definition 3.1.1);
(c) pointwise limits f of functions $f^{(n)}$ (Definition 3.2.1); and
(d) uniform limits f of functions $f^{(n)}$ (Definition 3.2.7).

This proliferation of limits may seem rather complicated. However, we can reduce the complexity slightly by observing that (d) can be viewed as a special case of (a), though in doing so it should be cautioned that because we are now dealing with functions instead of points, the convergence is not in X or in Y, but rather in a new space, the space of functions from X to Y.

Remark 3.4.1 If one is willing to work in topological spaces instead of metric spaces, we can also view (a) as a special case of (b), see Exercise 3.1.4, and (c) is also a special case of (a), see Exercise 3.4.4. Thus the notion of convergence in a topological space can be used to unify all the notions of limits we have encountered so far.

Definition 3.4.2 (*Metric space of bounded functions*) Suppose (X, d_X) and (Y, d_Y) are metric spaces. We let $B(X \to Y)$ denote the space[1] of bounded functions from X to Y:

$$B(X \to Y) := \{f \,\big|\, f : X \to Y \text{ is a bounded function}\}.$$

If X is non-empty, we define a metric $d_\infty : B(X \to Y) \times B(X \to Y) \to [0, +\infty)$ by defining

$$d_\infty(f, g) := \sup_{x \in X} d_Y(f(x), g(x)) = \sup\{d_Y(f(x), g(x)) : x \in X\}$$

for all $f, g \in B(X \to Y)$. This metric is sometimes known as the *uniform metric*, *sup norm metric* or the L^∞ *metric*. We will also use $d_{B(X \to Y)}$ as a synonym for d_∞. If X is empty, we instead define $d_\infty(f, g) = 0$.

Notice that the distance $d_\infty(f, g)$ is always finite because f and g are assumed to be bounded on X.

Example 3.4.3 Let $X := [0, 1]$ and $Y = \mathbf{R}$. Let $f : [0, 1] \to \mathbf{R}$ and $g : [0, 1] \to \mathbf{R}$ be the functions $f(x) := 2x$ and $g(x) := 3x$. Then f and g are both bounded functions and thus live in $B([0, 1] \to \mathbf{R})$. The distance between them is

$$d_\infty(f, g) = \sup_{x \in [0,1]} |2x - 3x| = \sup_{x \in [0,1]} |x| = 1.$$

This space turns out to be a metric space (Exercise 3.4.1). Convergence in this metric turns out to be identical to uniform convergence:

Proposition 3.4.4 *Let (X, d_X) and (Y, d_Y) be metric spaces. Let $(f^{(n)})_{n=1}^\infty$ be a sequence of functions in $B(X \to Y)$, and let f be another function in $B(X \to Y)$. Then $(f^{(n)})_{n=1}^\infty$ converges to f in the metric $d_{B(X \to Y)}$ if and only if $(f^{(n)})_{n=1}^\infty$ converges uniformly to f.*

Proof See Exercise 3.4.2. □

Now let $C(X \to Y)$ be the space of bounded continuous functions from X to Y:

$$C(X \to Y) := \{f \in B(X \to Y) \,\big|\, f \text{ is continuous}\}.$$

[1] Note that this is a set, thanks to the power set axiom (Axiom 3.11) and the axiom of specification (Axiom 3.6).

This set $C(X \to Y)$ is clearly a subset of $B(X \to Y)$. Corollary 3.3.2 asserts that this space $C(X \to Y)$ is closed in $B(X \to Y)$ (why?). Actually, we can say a lot more:

Theorem 3.4.5 (The space of continuous functions is complete) *Let* (X, d_X) *be a metric space, and let* (Y, d_Y) *be a* complete *metric space. The space* $(C(X \to Y), d_{B(X\to Y)}|_{C(X\to Y)\times C(X\to Y)})$ *is a complete subspace of* $(B(X \to Y), d_{B(X\to Y)})$. *In other words, every Cauchy sequence of functions in* $C(X \to Y)$ *converges to a function in* $C(X \to Y)$.

Proof See Exercise 3.4.3. □

— Exercises —

Exercise 3.4.1 Let (X, d_X) and (Y, d_Y) be metric spaces. Show that the space $B(X \to Y)$ defined in Definition 3.4.2, with the metric $d_{B(X\to Y)}$, is indeed a metric space.

Exercise 3.4.2 Prove Proposition 3.4.4.

Exercise 3.4.3 Prove Theorem 3.4.5. (*Hint:* this is similar, but not identical, to the proof of Theorem 3.3.1).

Exercise 3.4.4 Let (X, d_X) and (Y, d_Y) be metric spaces, and let $Y^X := \{f \big| f : X \to Y\}$ be the space of all functions from X to Y (cf. Axiom 3.11). If $x_0 \in X$ and V is an open set in Y, let $V^{(x_0)} \subseteq Y^X$ be the set

$$V^{(x_0)} := \{f \in Y^X : f(x_0) \in V\}.$$

If E is a subset of Y^X, we say that E is *open* if for every $f \in E$, there exists a finite number of points $x_1, \dots, x_n \in X$ and open sets $V_1, \dots, V_n \subseteq Y$ such that

$$f \in V_1^{(x_1)} \cap \dots \cap V_n^{(x_n)} \subseteq E.$$

(a) Show that if $\mathcal{F}$ is the collection of open sets in Y^X, then $(Y^X, \mathcal{F})$ is a topological space.
(b) For each natural number n, let $f^{(n)} : X \to Y$ be a function from X to Y, and let $f : X \to Y$ be another function from X to Y. Show that $f^{(n)}$ converges to f in the topology $\mathcal{F}$ (in the sense of Definition 2.5.4) if and only if $f^{(n)}$ converges to f pointwise (in the sense of Definition 3.2.1).

The topology $\mathcal{F}$ is known as the *topology of pointwise convergence*, for obvious reasons; it is also known as the *product topology*. It shows that the concept of pointwise convergence can be viewed as a special case of the more general concept of convergence in a topological space.

3.5 Series of Functions; the Weierstrass *M*-Test

Having discussed sequences of functions, we now discuss infinite series $\sum_{n=1}^{\infty} f_n$ of functions. Now we shall restrict our attention to functions $f: X \to \mathbf{R}$ from a metric space (X, d_X) to the real line $\mathbf{R}$ (which we of course give the standard metric); this is because we know how to add two real numbers, but don't necessarily know how to add two points in a general metric space Y. Functions whose codomain is $\mathbf{R}$ are sometimes called *real-valued* functions.

Finite summation is, of course, easy: given any finite collection $f^{(1)}, \ldots, f^{(N)}$ of functions from X to $\mathbf{R}$, we can define the finite sum $\sum_{i=1}^{N} f^{(i)}: X \to \mathbf{R}$ by

$$\left(\sum_{i=1}^{N} f^{(i)}\right)(x) := \sum_{i=1}^{N} f^{(i)}(x).$$

Example 3.5.1 If $f^{(1)}: \mathbf{R} \to \mathbf{R}$ is the function $f^{(1)}(x) := x$, $f^{(2)}: \mathbf{R} \to \mathbf{R}$ is the function $f^{(2)}(x) := x^2$, and $f^{(3)}: \mathbf{R} \to \mathbf{R}$ is the function $f^{(3)}(x) := x^3$, then $f := \sum_{i=1}^{3} f^{(i)}$ is the function $f: \mathbf{R} \to \mathbf{R}$ defined by $f(x) := x + x^2 + x^3$.

It is easy to show that finite sums of bounded functions are bounded, and finite sums of continuous functions are continuous (Exercise 3.5.1).

Now to add infinite series.

Definition 3.5.2 (*Infinite series*) Let (X, d_X) be a metric space. Let $(f^{(n)})_{n=1}^{\infty}$ be a sequence of functions from X to $\mathbf{R}$, and let f be another function from X to $\mathbf{R}$. If the partial sums $\sum_{n=1}^{N} f^{(n)}$ converge pointwise to f on X as $N \to \infty$, we say that the infinite series $\sum_{n=1}^{\infty} f^{(n)}$ *converges pointwise* to f, and write $f = \sum_{n=1}^{\infty} f^{(n)}$. If the partial sums $\sum_{n=1}^{N} f^{(n)}$ converge uniformly to f on X as $N \to \infty$, we say that the infinite series $\sum_{n=1}^{\infty} f^{(n)}$ *converges uniformly* to f, and again write $f = \sum_{n=1}^{\infty} f^{(n)}$. (Thus when one sees an expression such as $f = \sum_{n=1}^{\infty} f^{(n)}$, one should look at the context to see in what sense this infinite series converges.)

Remark 3.5.3 A series $\sum_{n=1}^{\infty} f^{(n)}$ converges pointwise to f on X if and only if $\sum_{n=1}^{\infty} f^{(n)}(x)$ converges to $f(x)$ for *every* $x \in X$. (Thus if $\sum_{n=1}^{\infty} f^{(n)}$ does not converge pointwise to f, this does not mean that it diverges pointwise; it may just be that it converges for some points x but diverges at other points x.)

If a series $\sum_{n=1}^{\infty} f^{(n)}$ converges uniformly to f, then it also converges pointwise to f; but not vice versa, as the following example shows.

Example 3.5.4 Let $f^{(n)}: (-1, 1) \to \mathbf{R}$ be the sequence of functions $f^{(n)}(x) := x^n$. Then $\sum_{n=1}^{\infty} f^{(n)}$ converges pointwise, but not uniformly, to the function $x/(1-x)$ (see Exercise 3.2.2 and Example 3.5.8).

It is not always clear when a series $\sum_{n=1}^{\infty} f^{(n)}$ converges or not. However, there is a very useful test that gives at least one test for uniform convergence.

Definition 3.5.5 (*Sup norm*) If $f: X \to \mathbf{R}$ is a bounded real-valued function, and X is non-empty, we define the *sup norm* $\|f\|_\infty$ of f to be the number

$$\|f\|_\infty := \sup\{|f(x)| : x \in X\}.$$

In other words, $\|f\|_\infty = d_\infty(f, 0)$, where $0 : X \to \mathbf{R}$ is the zero function $0(x) := 0$, and d_∞ was defined in Definition 3.4.2. (Why is this the case?) If X is empty, we instead define $\|f\|_\infty := 0$.

Example 3.5.6 Thus, for instance, if $f: (-2, 1) \to \mathbf{R}$ is the function $f(x) := 2x$, then $\|f\|_\infty = \sup\{|2x| : x \in (-2, 1)\} = 4$ (why?). Notice that when f is bounded then $\|f\|_\infty$ will always be a non-negative real number.

Theorem 3.5.7 (Weierstrass M-test) *Let (X, d) be a metric space, and let $(f^{(n)})_{n=1}^\infty$ be a sequence of bounded real-valued continuous functions on X such that the series $\sum_{n=1}^\infty \|f^{(n)}\|_\infty$ is convergent. (Note that this is a series of plain old real numbers, not of functions.) Then the series $\sum_{n=1}^\infty f^{(n)}$ converges uniformly to some function f on X, and that function f is also continuous.*

Proof See Exercise 3.5.2. □

To put the Weierstrass M-test succinctly: absolute convergence of sup norms implies uniform convergence of functions.

Example 3.5.8 Let $0 < r < 1$ be a real number, and let $f^{(n)}: [-r, r] \to \mathbf{R}$ be the series of functions $f^{(n)}(x) := x^n$. Then each $f^{(n)}$ is continuous and bounded, and $\|f^{(n)}\|_\infty = r^n$ (why?). Since the series $\sum_{n=1}^\infty r^n$ is absolutely convergent (e.g., by the root test, Theorem 7.5.1 from *Analysis I*), we thus see that $\sum_{n=1}^\infty f^{(n)}$ converges uniformly in $[-r, r]$ to some continuous function; in Exercise 3.2.2(c) we see that this function must in fact be the function $f: [-r, r] \to \mathbf{R}$ defined by $f(x) := x/(1 - x)$. In other words, the series $\sum_{n=1}^\infty x^n$ is pointwise convergent, but not uniformly convergent, on $(-1, 1)$, but is uniformly convergent on the smaller interval $[-r, r]$ for any $0 < r < 1$.

The Weierstrass M-test is especially useful in relation to *power series*, which we will encounter in the next chapter.

— Exercises —

Exercise 3.5.1 Let $f^{(1)}, \ldots, f^{(N)}$ be a finite sequence of bounded functions from a metric space (X, d_X) to $\mathbf{R}$. Show that $\sum_{i=1}^N f^{(i)}$ is also bounded. Prove a similar claim when "bounded" is replaced by "continuous". What if "continuous" was replaced by "uniformly continuous"?

Exercise 3.5.2 Prove Theorem 3.5.7. (*Hint:* first show that the sequence $\sum_{i=1}^N f^{(i)}$ is a Cauchy sequence in $C(X \to \mathbf{R})$. Then use Theorem 3.4.5.)

3.6 Uniform Convergence and Integration

We now connect uniform convergence with Riemann integration (which was discussed in Chap. 11), by showing that uniform limits can be safely interchanged with integrals.

Theorem 3.6.1 *Let $[a, b]$ be an interval, and for each integer $n \geq 1$, let $f^{(n)} \colon [a, b] \to \mathbf{R}$ be a Riemann-integrable function. Suppose $f^{(n)}$ converges uniformly on $[a, b]$ to a function $f \colon [a, b] \to \mathbf{R}$. Then f is also Riemann integrable, and*

$$\lim_{n\to\infty} \int_{[a,b]} f^{(n)} = \int_{[a,b]} f.$$

Proof We first show that f is Riemann integrable on $[a, b]$. This is the same as showing that the upper and lower Riemann integrals of f match: $\underline{\int}_{[a,b]} f = \overline{\int}_{[a,b]} f$.

Let $\varepsilon > 0$. Since $f^{(n)}$ converges uniformly to f, we see that there exists an $N > 0$ such that $|f^{(n)}(x) - f(x)| < \varepsilon$ for all $n > N$ and $x \in [a, b]$. In particular we have

$$f^{(n)}(x) - \varepsilon < f(x) < f^{(n)}(x) + \varepsilon$$

for all $x \in [a, b]$. Integrating this on $[a, b]$ we obtain

$$\underline{\int}_{[a,b]} (f^{(n)} - \varepsilon) \leq \underline{\int}_{[a,b]} f \leq \overline{\int}_{[a,b]} f \leq \overline{\int}_{[a,b]} (f^{(n)} + \varepsilon).$$

Since $f^{(n)}$ is assumed to be Riemann integrable, we thus see

$$\left(\int_{[a,b]} f^{(n)}\right) - \varepsilon(b-a) \leq \underline{\int}_{[a,b]} f \leq \overline{\int}_{[a,b]} f \leq \left(\int_{[a,b]} f^{(n)}\right) + \varepsilon(b-a).$$

In particular, we see that

$$0 \leq \overline{\int}_{[a,b]} f - \underline{\int}_{[a,b]} f \leq 2\varepsilon(b-a).$$

Since this is true for every $\varepsilon > 0$, we obtain $\underline{\int}_{[a,b]} f = \overline{\int}_{[a,b]} f$ as desired.

The above argument also shows that for every $\varepsilon > 0$ there exists an $N > 0$ such that

$$\left|\int_{[a,b]} f^{(n)} - \int_{[a,b]} f\right| \leq \varepsilon(b-a)$$

for all $n \geq N$. Since ε is arbitrary, we see that $\int_{[a,b]} f^{(n)}$ converges to $\int_{[a,b]} f$ as desired. □

To rephrase Theorem 3.6.1: we can rearrange limits and integrals (on compact intervals $[a, b]$),

$$\lim_{n\to\infty} \int_{[a,b]} f^{(n)} = \int_{[a,b]} \lim_{n\to\infty} f^{(n)},$$

provided that the convergence is uniform. This should be contrasted with Examples 1.2.9 and 3.2.5.

There is an analogue of this theorem for series:

Corollary 3.6.2 *Let $[a, b]$ be an interval, and let $(f^{(n)})_{n=1}^{\infty}$ be a sequence of Riemann-integrable functions on $[a, b]$ such that the series $\sum_{n=1}^{\infty} f^{(n)}$ is uniformly convergent. Then we have*

$$\sum_{n=1}^{\infty} \int_{[a,b]} f^{(n)} = \int_{[a,b]} \sum_{n=1}^{\infty} f^{(n)}.$$

Proof See Exercise 3.6.1. □

This corollary works particularly well in conjunction with the Weierstrass M-test (Theorem 3.5.7):

Example 3.6.3 (Informal) From Lemma 7.3.3 of *Analysis I* we have the geometric series identity

$$\sum_{n=1}^{\infty} x^n = \frac{x}{1-x}$$

for $x \in (-1, 1)$, and the convergence is uniform (by the Weierstrass M-test) on $[-r, r]$ for any $0 < r < 1$. By adding 1 to both sides we obtain

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}$$

and the converge is again uniform. We can thus integrate on $[0, r]$ and use Corollary 3.6.2 to obtain

$$\sum_{n=0}^{\infty} \int_{[0,r]} x^n \, dx = \int_{[0,r]} \frac{1}{1-x} \, dx.$$

The left-hand side is $\sum_{n=0}^{\infty} r^{n+1}/(n+1)$. If we accept for now the use of logarithms (we will justify this use in Sect. 4.5), the anti-derivative of $1/(1-x)$ is $-\log(1-x)$, and so the right-hand side is $-\log(1-r)$. We thus obtain the formula

$$-\log(1-r) = \sum_{n=0}^{\infty} r^{n+1}/(n+1)$$

for all $0 < r < 1$.

— Exercises —

Exercise 3.6.1 Use Theorem 3.6.1 to prove Corollary 3.6.2.

3.7 Uniform Convergence and Derivatives

We have already seen how uniform convergence interacts well with continuity, with limits, and with integrals. Now we investigate how it interacts with derivatives.

The first question we can ask is: if f_n converges uniformly to f, and the functions f_n are differentiable, does this imply that f is also differentiable? And does f_n' also converge to f'?

The answer to the second question is, unfortunately, no. To see a counterexample, we will use without proof some basic facts about trigonometric functions (which we will make rigorous in Sect. 4.7). Consider the functions $f_n \colon [0, 2\pi] \to \mathbf{R}$ defined by $f_n(x) := n^{-1/2} \sin(nx)$, and let $f \colon [0, 2\pi] \to \mathbf{R}$ be the zero function $f(x) := 0$. Then, since sin takes values between -1 and 1, we have $d_\infty(f_n, f) \le n^{-1/2}$, where we use the uniform metric $d_\infty(f, g) := \sup_{x \in [0,2\pi]} |f(x) - g(x)|$ introduced in Definition 3.4.2. Since $n^{-1/2}$ converges to 0, we thus see by the squeeze test that f_n converges uniformly to f. On the other hand, $f_n'(x) = n^{1/2} \cos(nx)$, and so in particular $|f_n'(0) - f'(0)| = n^{1/2}$. Thus f_n' does not converge pointwise to f', and so in particular does not converge uniformly either. In particular we have

$$\frac{d}{dx} \lim_{n\to\infty} f_n(x) \ne \lim_{n\to\infty} \frac{d}{dx} f_n(x).$$

The answer to the first question is also no. An example is the sequence of functions $f_n \colon [-1, 1] \to \mathbf{R}$ defined by $f_n(x) := \sqrt{\frac{1}{n^2} + x^2}$. These functions are differentiable (why?). Also, one can easily check that

$$|x| \le f_n(x) \le |x| + \frac{1}{n}$$

for all $x \in [-1, 1]$ (why? square both sides), and so by the squeeze test f_n converges uniformly to the absolute value function $f(x) := |x|$. But this function is not differentiable at 0 (why?). Thus, the uniform limit of differentiable functions need not be differentiable. (See also Example 1.2.10.)

So, in summary, uniform convergence of the functions f_n says nothing about the convergence of the derivatives f_n'. However, the converse is true, as long as f_n converges at at least one point:

Theorem 3.7.1 *Let* $[a, b]$ *be an interval, and for every integer* $n \geq 1$, *let* $f_n\colon [a, b] \to \mathbf{R}$ *be a differentiable function whose derivative* $f_n'\colon [a, b] \to \mathbf{R}$ *is continuous. Suppose that the derivatives* f_n' *converge uniformly to a function* $g\colon [a, b] \to \mathbf{R}$. *Suppose also that there exists a point* x_0 *such that the limit* $\lim_{n\to\infty} f_n(x_0)$ *exists. Then the functions* f_n *converge uniformly to a differentiable function* f, *and the derivative of* f *equals* g.

Informally, the above theorem says that if f_n' converges uniformly, and $f_n(x_0)$ converges for some x_0, then f_n also converges uniformly, and $\frac{d}{dx}\lim_{n\to\infty} f_n(x) = \lim_{n\to\infty}\frac{d}{dx} f_n(x)$.

Proof We only give the beginning of the proof here; the remainder of the proof will be an exercise (Exercise 3.7.1).

Since f_n' is continuous, we see from the fundamental theorem of calculus (Theorem 11.9.4) that

$$f_n(x) - f_n(x_0) = \int_{[x_0,x]} f_n'$$

when $x \in [x_0, b]$, and

$$f_n(x) - f_n(x_0) = -\int_{[x,x_0]} f_n'$$

when $x \in [a, x_0]$.

Let L be the limit of $f_n(x_0)$ as $n \to \infty$:

$$L := \lim_{n\to\infty} f_n(x_0).$$

By hypothesis, L exists. Now, since each f_n' is continuous on $[a, b]$, and f_n' converges uniformly to g, we see by Corollary 3.3.2 that g is also continuous. Now define the function $f\colon [a, b] \to \mathbf{R}$ by setting

$$f(x) := L - \int_{[a,x_0]} g + \int_{[a,x]} g$$

for all $x \in [a, b]$. To finish the proof, we have to show that f_n converges uniformly to f, and that f is differentiable with derivative g; this shall be done in Exercise 3.7.1. □

Remark 3.7.2 It turns out that Theorem 3.7.1 is still true when the functions f_n' are not assumed to be continuous, but the proof is more difficult; see Exercise 3.7.2.

By combining this theorem with the Weierstrass M-test, we obtain

Corollary 3.7.3 *Let $[a, b]$ be an interval, and for every integer $n \geq 1$, let $f_n : [a, b] \to \mathbf{R}$ be a differentiable function whose derivative $f_n' : [a, b] \to \mathbf{R}$ is continuous. Suppose that the series $\sum_{n=1}^{\infty} \|f_n'\|_\infty$ is absolutely convergent, where*

$$\|f_n'\|_\infty := \sup_{x \in [a,b]} |f_n'(x)|$$

is the sup norm of f_n', as defined in Definition 3.5.5. Suppose also that the series $\sum_{n=1}^{\infty} f_n(x_0)$ is convergent for some $x_0 \in [a, b]$. Then the series $\sum_{n=1}^{\infty} f_n$ converges uniformly on $[a, b]$ to a differentiable function, and in fact

$$\frac{d}{dx} \sum_{n=1}^{\infty} f_n(x) = \sum_{n=1}^{\infty} \frac{d}{dx} f_n(x)$$

for all $x \in [a, b]$.

Proof See Exercise 3.7.3. □

We now pause to give an example of a function which is continuous everywhere, but differentiable nowhere (this particular example was discovered by Weierstrass). Again, we will presume knowledge of the trigonometric functions, which will be covered rigorously in Sect. 4.7.

Example 3.7.4 Let $f : \mathbf{R} \to \mathbf{R}$ be the function

$$f(x) := \sum_{n=1}^{\infty} 4^{-n} \cos(32^n \pi x).$$

Note that this series is uniformly convergent, thanks to the Weierstrass M-test, and since each individual function $4^{-n} \cos(32^n \pi x)$ is continuous, the function f is also continuous. However, it is not differentiable (Exercise 4.7.10); in fact it is a *nowhere differentiable function*, one which is not differentiable at *any* point, despite being continuous everywhere!

— Exercises —

Exercise 3.7.1 Complete the proof of Theorem 3.7.1. Compare this theorem with Example 1.2.10, and explain why this example does not contradict the theorem.

Exercise 3.7.2 Prove Theorem 3.7.1 without assuming that f_n' is continuous. (This means that you cannot use the fundamental theorem of calculus. However, the mean value theorem (Corollary 10.2.9) is still available. Use this to show that if $d_\infty(f_n', f_m') \leq \varepsilon$, then $|(f_n(x) - f_m(x)) - (f_n(x_0) - f_m(x_0))| \leq \varepsilon |x - x_0|$ for all $x \in [a, b]$, and then use this to complete the proof of Theorem 3.7.1.)

Exercise 3.7.3 Prove Corollary 3.7.3.

3.8 Uniform Approximation by Polynomials

As we have just seen, continuous functions can be very badly behaved, for instance they can be nowhere differentiable (Example 3.7.4). On the other hand, functions such as polynomials are always very well behaved, in particular being always differentiable. Fortunately, while most continuous functions are not as well behaved as polynomials, they can always be *uniformly approximated* by polynomials; this important (but difficult) result is known as the *Weierstrass approximation theorem*, and is the subject of this section.

Definition 3.8.1 Let $[a, b]$ be an interval. A *polynomial on* $[a, b]$ is a function $f : [a, b] \to \mathbf{R}$ of the form $f(x) := \sum_{j=0}^{n} c_j x^j$, where $n \geq 0$ is an integer and $c_0, \ldots, c_n$ are real numbers. If $c_n \neq 0$, then n is called the *degree* of f.

Example 3.8.2 The function $f : [1, 2] \to \mathbf{R}$ defined by $f(x) := 3x^4 + 2x^3 - 4x + 5$ is a polynomial on $[1, 2]$ of degree 4.

Theorem 3.8.3 (Weierstrass approximation theorem) *If* $[a, b]$ *is an interval,* $f : [a, b] \to \mathbf{R}$ *is a continuous function, and* $\varepsilon > 0$*, then there exists a polynomial* P *on* $[a, b]$ *such that* $d_\infty(P, f) \leq \varepsilon$ *(i.e.,* $|P(x) - f(x)| \leq \varepsilon$ *for all* $x \in [a, b]$*).*

Another way of stating this theorem is as follows. Recall that $C([a, b] \to \mathbf{R})$ was the space of continuous functions from $[a, b]$ to $\mathbf{R}$, with the uniform metric d_∞. Let $P([a, b] \to \mathbf{R})$ be the space of all polynomials on $[a, b]$; this is a subspace of $C([a, b] \to \mathbf{R})$, since all polynomials are continuous (Exercise 9.4.7). The Weierstrass approximation theorem then asserts that every continuous function is an adherent point of $P([a, b] \to \mathbf{R})$; or in other words, that the closure of the space of polynomials is the space of continuous functions:

$$\overline{P([a, b] \to \mathbf{R})} = C([a, b] \to \mathbf{R}).$$

In particular, every continuous function on $[a, b]$ is the uniform limit of polynomials. Another way of saying this is that the space of polynomials is *dense* in the space of continuous functions, in the *uniform topology*.

The proof of the Weierstrass approximation theorem is somewhat complicated and will be done in stages. We first need the notion of an *approximation to the identity*.

Definition 3.8.4 (*Compactly supported functions*) Let $[a, b]$ be an interval. A function $f : \mathbf{R} \to \mathbf{R}$ is said to be *supported* on $[a, b]$ if $f(x) = 0$ for all $x \notin [a, b]$. We say that f is *compactly supported* if it is supported on some interval $[a, b]$. If f is continuous and supported on $[a, b]$, we define the improper integral $\int_{-\infty}^{\infty} f$ to be $\int_{-\infty}^{\infty} f := \int_{[a,b]} f$.

Note that a function can be supported on more than one interval, for instance a function which is supported on $[3, 4]$ is also automatically supported on $[2, 5]$ (why?). In principle, this might mean that our definition of $\int_{-\infty}^{\infty} f$ is not well defined, however this is not the case:

Lemma 3.8.5 *If* $f : \mathbf{R} \to \mathbf{R}$ *is continuous and supported on an interval* $[a, b]$*, and is also supported on another interval* $[c, d]$*, then* $\int_{[a,b]} f = \int_{[c,d]} f$.

Proof See Exercise 3.8.1. □

Definition 3.8.6 *Approximation to the identity*) Let $\varepsilon > 0$ and $0 < \delta < 1$. A function $f : \mathbf{R} \to \mathbf{R}$ is said to be an (ε, δ)*-approximation to the identity* if it obeys the following three properties:

(a) f is supported on $[-1, 1]$, and $f(x) \geq 0$ for all $-1 \leq x \leq 1$.
(b) f is continuous, and $\int_{-\infty}^{\infty} f = 1$.
(c) $|f(x)| \leq \varepsilon$ for all $\delta \leq |x| \leq 1$.

Remark 3.8.7 For those of you who are familiar with the Dirac delta function, approximations to the identity are ways to approximate this (very discontinuous) delta function by a continuous function (which is easier to analyze). We will not however discuss the Dirac delta function in this text.

Our proof of the Weierstrass approximation theorem relies on three key facts. The first fact is that polynomials can be approximations to the identity:

Lemma 3.8.8 (Polynomials can approximate the identity) *For every* $\varepsilon > 0$ *and* $0 < \delta < 1$ *there exists an* (ε, δ)*-approximation to the identity which is a polynomial* P *on* $[-1, 1]$.

Proof See Exercise 3.8.2. □

We will use these polynomial approximations to the identity to approximate continuous functions by polynomials. We will need the following important notion of a *convolution*.

Definition 3.8.9 (*Convolution*) Let $f : \mathbf{R} \to \mathbf{R}$ and $g : \mathbf{R} \to \mathbf{R}$ be continuous, compactly supported functions. We define the *convolution* $f * g : \mathbf{R} \to \mathbf{R}$ of f and g to be the function

$$(f * g)(x) := \int_{-\infty}^{\infty} f(y)g(x - y)\, dy.$$

Note that if f and g are continuous and compactly supported, then for each x the function $f(y)g(x - y)$ (thought of as a function of y) is also continuous and compactly supported, so the above definition makes sense.

Remark 3.8.10 Convolutions play an important rôle in Fourier analysis and in partial differential equations (PDE), and are also important in physics, engineering, and signal processing. An in-depth study of convolution is beyond the scope of this text; only a brief treatment will be given here.

Proposition 3.8.11 (Basic properties of convolution) *Let* $f : \mathbf{R} \to \mathbf{R}$, $g : \mathbf{R} \to \mathbf{R}$, *and* $h : \mathbf{R} \to \mathbf{R}$ *be continuous, compactly supported functions. Then the following statements are true.*

(a) *The convolution* $f * g$ *is also a continuous, compactly supported function.*
(b) *(Convolution is commutative) We have* $f * g = g * f$.
(c) *(Convolution is linear) We have* $f * (g + h) = f * g + f * h$. *Also, for any real number* c, *we have* $f * (cg) = (cf) * g = c(f * g)$.

Proof See Exercise 3.8.4. □

Remark 3.8.12 There are many other important properties of convolution, for instance it is associative, $(f * g) * h = f * (g * h)$, and it commutes with derivatives, $(f * g)' = f' * g = f * g'$, when f and g are differentiable. The Dirac delta function δ mentioned earlier is an identity for convolution: $f * \delta = \delta * f = f$. These results are slightly harder to prove than the ones in Proposition 3.8.11, however, and we will not need them in this text.

As mentioned earlier, the proof of the Weierstrass approximation theorem relies on three facts. The second key fact is that convolution with polynomials produces another polynomial:

Lemma 3.8.13 *Let* $f : \mathbf{R} \to \mathbf{R}$ *be a continuous function supported on* $[0, 1]$, *and let* $g : \mathbf{R} \to \mathbf{R}$ *be a continuous function supported on* $[-1, 1]$ *which is a polynomial on* $[-1, 1]$. *Then* $f * g$ *is a polynomial on* $[0, 1]$. *(Note however that it may be non-polynomial outside of* $[0, 1]$.*)*

Proof Since g is polynomial on $[-1, 1]$, we may find an integer $n \geq 0$ and real numbers $c_0, c_1, \ldots, c_n$ such that

$$g(x) = \sum_{j=0}^{n} c_j x^j \text{ for all } x \in [-1, 1].$$

On the other hand, for all $x \in [0, 1]$, we have

$$f * g(x) = \int_{-\infty}^{\infty} f(y)g(x - y)\, dy = \int_{[0,1]} f(y)g(x - y)\, dy$$

since f is supported on $[0, 1]$. Since $x \in [0, 1]$ and the variable of integration y is also in $[0, 1]$, we have $x - y \in [-1, 1]$. Thus we may substitute in our formula for g to obtain

$$f * g(x) = \int_{[0,1]} f(y) \sum_{j=0}^{n} c_j (x - y)^j\, dy.$$

We expand this using the binomial formula (Exercise 7.1.4) to obtain

$$f * g(x) = \int_{[0,1]} f(y) \sum_{j=0}^{n} c_j \sum_{k=0}^{j} \frac{j!}{k!(j - k)!} x^k (-y)^{j-k}\, dy.$$

We can interchange the two summations (by Corollary 7.1.14) to obtain

$$f * g(x) = \int_{[0,1]} f(y) \sum_{k=0}^{n} \sum_{j=k}^{n} c_j \frac{j!}{k!(j-k)!} x^k (-y)^{j-k}\, dy$$

(why did the limits of summation change? It may help to plot j and k on a graph). Now we interchange the k summation with the integral, and observe that x^k is independent of y, to obtain

$$f * g(x) = \sum_{k=0}^{n} x^k \int_{[0,1]} f(y) \sum_{j=k}^{n} c_j \frac{j!}{k!(j-k)!} (-y)^{j-k}\, dy.$$

If we thus define

$$C_k := \int_{[0,1]} f(y) \sum_{j=k}^{n} c_j \frac{j!}{k!(j-k)!} (-y)^{j-k}\, dy$$

for each $k = 0, \ldots, n$, then C_k is a number which is independent of x, and we have

$$f * g(x) = \sum_{k=0}^{n} C_k x^k$$

for all $x \in [0, 1]$. Thus $f * g$ is a polynomial on $[0, 1]$. □

The third key fact is that if one convolves a uniformly continuous function with an approximation to the identity, we obtain a new function which is close to the original function (which explains the terminology "approximation to the identity"):

Lemma 3.8.14 *Let $f: \mathbf{R} \to \mathbf{R}$ be a continuous function supported on $[0, 1]$, which is bounded by some $M > 0$ (i.e., $|f(x)| \le M$ for all $x \in \mathbf{R}$), and let $\varepsilon > 0$ and $0 < \delta < 1$ be such that one has $|f(x) - f(y)| < \varepsilon$ whenever $x, y \in \mathbf{R}$ and $|x - y| < \delta$. Let g be any (ε, δ)-approximation to the identity. Then we have*

$$|f * g(x) - f(x)| \le (1 + 4M)\varepsilon$$

for all $x \in [0, 1]$.

Proof See Exercise 3.8.6. □

Combining these together, we obtain a preliminary version of the Weierstrass approximation theorem:

Corollary 3.8.15 (Weierstrass approximation theorem I) *Let $f: \mathbf{R} \to \mathbf{R}$ be a continuous function supported on $[0, 1]$. Then for every $\varepsilon > 0$, there exists a function*

$P : \mathbf{R} \to \mathbf{R}$ which is polynomial on $[0, 1]$ and such that $|P(x) - f(x)| \le \varepsilon$ for all $x \in [0, 1]$.

Proof See Exercise 3.8.7. □

Now we perform a series of modifications to convert Corollary 3.8.15 into the actual Weierstrass approximation theorem. We first need a simple lemma.

Lemma 3.8.16 *Let $f : [0, 1] \to \mathbf{R}$ be a continuous function which equals 0 on the boundary of $[0, 1]$, i.e., $f(0) = f(1) = 0$. Let $F : \mathbf{R} \to \mathbf{R}$ be the function defined by setting $F(x) := f(x)$ for $x \in [0, 1]$ and $F(x) := 0$ for $x \notin [0, 1]$. Then F is also continuous.*

Proof See Exercise 3.8.9. □

Remark 3.8.17 The function F obtained in Lemma 3.8.16 is sometimes known as the *extension of f by zero*.

From Corollary 3.8.15 and Lemma 3.8.16 we immediately obtain

Corollary 3.8.18 (Weierstrass approximation theorem II) *Let $f : [0, 1] \to \mathbf{R}$ be a continuous function such that $f(0) = f(1) = 0$. Then for every $\varepsilon > 0$ there exists a polynomial $P : [0, 1] \to \mathbf{R}$ such that $|P(x) - f(x)| \le \varepsilon$ for all $x \in [0, 1]$.*

Now we strengthen Corollary 3.8.18 by removing the assumption that $f(0) = f(1) = 0$.

Corollary 3.8.19 (Weierstrass approximation theorem III) *Let $f : [0, 1] \to \mathbf{R}$ be a continuous function. Then for every $\varepsilon > 0$ there exists a polynomial $P : [0, 1] \to \mathbf{R}$ such that $|P(x) - f(x)| \le \varepsilon$ for all $x \in [0, 1]$.*

Proof Let $F : [0, 1] \to \mathbf{R}$ denote the function

$$F(x) := f(x) - f(0) - x(f(1) - f(0)).$$

Observe that F is also continuous (why?), and that $F(0) = F(1) = 0$. By Corollary 3.8.18, we can thus find a polynomial $Q : [0, 1] \to \mathbf{R}$ such that $|Q(x) - F(x)| \le \varepsilon$ for all $x \in [0, 1]$. But

$$Q(x) - F(x) = Q(x) + f(0) + x(f(1) - f(0)) - f(x),$$

so the claim follows by setting P to be the polynomial $P(x) := Q(x) + f(0) + x(f(1) - f(0))$. □

Finally, we can prove the full Weierstrass approximation theorem.

Proof of Theorem 3.8.3 Let $f : [a, b] \to \mathbf{R}$ be a continuous function on $[a, b]$. Let $g : [0, 1] \to \mathbf{R}$ denote the function

$$g(x) := f(a + (b - a)x) \text{ for all } x \in [0, 1]$$

Observe then that

$$f(y) = g((y - a)/(b - a)) \text{ for all } y \in [a, b].$$

The function g is continuous on $[0, 1]$ (why?), and so by Corollary 3.8.19 we may find a polynomial $Q : [0, 1] \to \mathbf{R}$ such that $|Q(x) - g(x)| \leq \varepsilon$ for all $x \in [0, 1]$. In particular, for any $y \in [a, b]$, we have

$$|Q((y - a)/(b - a)) - g((y - a)/(b - a))| \leq \varepsilon.$$

If we thus set $P(y) := Q((y - a)/(b - a))$, then we observe that P is also a polynomial (why?), and so we have $|P(y) - f(y)| \leq \varepsilon$ for all $y \in [a, b]$, as desired.

Remark 3.8.20 Note that the Weierstrass approximation theorem only works on bounded intervals $[a, b]$; continuous functions on $\mathbf{R}$ cannot be uniformly approximated by polynomials. For instance, the exponential function $f : \mathbf{R} \to \mathbf{R}$ defined by $f(x) := e^x$ (which we shall study rigorously in Sect. 4.5) cannot be approximated by any polynomial, because exponential functions grow faster than any polynomial (Exercise 4.5.9) and so there is no way one can even make the sup metric between f and a polynomial finite.

Remark 3.8.21 There is a generalization of the Weierstrass approximation theorem to higher dimensions: if K is any compact subset of $\mathbf{R}^n$ (with the Euclidean metric d_{l^2}), and $f : K \to \mathbf{R}$ is a continuous function, then for every $\varepsilon > 0$ there exists a polynomial $P : K \to \mathbf{R}$ of n variables $x_1, \ldots, x_n$ such that $d_\infty(f, P) < \varepsilon$. This general theorem can be proven by a more complicated variant of the arguments here, but we will not do so. (There is in fact an even more general version of this theorem applicable to an arbitrary metric space, known as the *Stone-Weierstrass theorem*, but this is beyond the scope of this text.)

— Exercises —

Exercise 3.8.1 Prove Lemma 3.8.5.

Exercise 3.8.2 (a) Prove that for any real number $0 \leq y \leq 1$ and any natural number $n \geq 0$, that $(1 - y)^n \geq 1 - ny$. (*Hint:* induct on n. Alternatively, differentiate with respect to y.)

(b) Show that $\int_{-1}^{1} (1 - x^2)^n \, dx \geq \frac{1}{\sqrt{n}}$. (*Hint:* for $|x| \leq 1/\sqrt{n}$, use part (a); for $|x| \geq 1/\sqrt{n}$, just use the fact that $(1 - x^2)$ is positive. It is also possible to proceed via trigonometric substitution, but I would not recommend this unless you know what you are doing.)

(c) Prove Lemma 3.8.8. (*Hint:* choose $f(x)$ to equal $c(1-x^2)^N$ for $x \in [-1, 1]$ and to equal zero for $x \notin [-1, 1]$, where N is a large number N, where c is chosen so that f has integral 1, and use (b).)

Exercise 3.8.3 Let $f : \mathbf{R} \to \mathbf{R}$ be a compactly supported, continuous function. Show that f is bounded and uniformly continuous. (*Hint:* the idea is to use Proposition 2.3.2 and Theorem 2.3.5, but one must first deal with the issue that the domain $\mathbf{R}$ of f is non-compact.)

Exercise 3.8.4 Prove Proposition 3.8.11. (*Hint:* to show that $f * g$ is continuous, use Exercise 3.8.3.)

Exercise 3.8.5 Let $f : \mathbf{R} \to \mathbf{R}$ and $g : \mathbf{R} \to \mathbf{R}$ be continuous, compactly supported functions. Suppose that f is supported on the interval $[0, 1]$, and g is constant on the interval $[0, 2]$ (i.e., there is a real number c such that $g(x) = c$ for all $x \in [0, 2]$). Show that the convolution $f * g$ is constant on the interval $[1, 2]$.

Exercise 3.8.6 (a) Let g be an (ε, δ) approximation to the identity. Show that $1 - 2\varepsilon \leq \int_{[-\delta,\delta]} g \leq 1$.

(b) Prove Lemma 3.8.14. (*Hint:* begin with the identity

$$f * g(x) = \int f(x-y)g(y)\,dy = \int_{[-\delta,\delta]} f(x-y)g(y)\,dy + \int_{[\delta,1]} f(x-y)g(y)\,dy + \int_{[-1,-\delta]} f(x-y)g(y)\,dy.$$

The idea is to show that the first integral is close to $f(x)$, and that the second and third integrals are very small. To achieve the former task, use (a) and the fact that $f(x)$ and $f(x-y)$ are within ε of each other; to achieve the latter task, use property (c) of the approximation to the identity and the fact that f is bounded.)

Exercise 3.8.7 Prove Corollary 3.8.15. (*Hint:* combine Exercise 3.8.3 and Lemmas 3.8.8, 3.8.13, 3.8.14.)

Exercise 3.8.8 Let $f : [0, 1] \to \mathbf{R}$ be a continuous function, and suppose that $\int_{[0,1]} f(x)x^n\,dx = 0$ for all non-negative integers $n = 0, 1, 2, \ldots$. Show that f must be the zero function $f \equiv 0$. (*Hint:* first show that $\int_{[0,1]} f(x)P(x)\,dx = 0$ for all polynomials P. Then, using the Weierstrass approximation theorem, show that $\int_{[0,1]} f(x)f(x)\,dx = 0$.)

Exercise 3.8.9 Prove Lemma 3.8.16.

Chapter 4
Power Series

4.1 Formal Power Series

We now discuss an important subclass of series of functions, that of *power series*. As in earlier chapters, we begin by introducing the notion of a formal power series and then focus in later sections on when the series converges to a meaningful function and what one can say about the function obtained in this manner.

Definition 4.1.1 (*Formal power series*) Let a be a real number. A *formal power series centered at a* is any series of the form

$$\sum_{n=0}^{\infty} c_n (x-a)^n$$

where $c_0, c_1, \ldots$ is a sequence of real numbers (not depending on x); we refer to c_n as the n^{th} *coefficient* of this series. Note that each term $c_n(x-a)^n$ in this series is a function of a real variable x.

Example 4.1.2 The series $\sum_{n=0}^{\infty} n!(x-2)^n$ is a formal power series centered at 2. The series $\sum_{n=0}^{\infty} 2^x (x-3)^n$ is not a formal power series, since the coefficients 2^x depend on x.

We call these power series *formal* because we do not yet assume that these series converge for any x. However, these series are automatically guaranteed to converge when $x = a$ (why?). In general, the closer x gets to a, the easier it is for this series to converge. To make this more precise, we need the following definition.

Definition 4.1.3 (*Radius of convergence*) Let $\sum_{n=0}^{\infty} c_n(x-a)^n$ be a formal power series. We define the *radius of convergence* R of this series to be the quantity

$$R := \frac{1}{\limsup_{n\to\infty} |c_n|^{1/n}}$$

T. Tao, *Analysis II*, Texts and Readings in Mathematics,
https://doi.org/10.1007/978-981-19-7284-3_4

where we adopt the convention that $\frac{1}{0} = +\infty$ and $\frac{1}{+\infty} = 0$.

Remark 4.1.4 Each number $|c_n|^{1/n}$ is non-negative, so the limit $\limsup_{n\to\infty} |c_n|^{1/n}$ can take on any value from 0 to $+\infty$, inclusive. Thus R can also take on any value between 0 and $+\infty$ inclusive (in particular it is not necessarily a real number). Note that the radius of convergence always exists, even if the sequence $|c_n|^{1/n}$ is not convergent, because the lim sup of any sequence always exists (though it might be $+\infty$ or $-\infty$).

Example 4.1.5 The series $\sum_{n=0}^{\infty} n(-2)^n (x-3)^n$ has radius of convergence

$$\frac{1}{\limsup_{n\to\infty} |n(-2)^n|^{1/n}} = \frac{1}{\limsup_{n\to\infty} 2n^{1/n}} = \frac{1}{2}.$$

The series $\sum_{n=0}^{\infty} 2^{n^2} (x+2)^n$ has radius of convergence

$$\frac{1}{\limsup_{n\to\infty} |2^{n^2}|^{1/n}} = \frac{1}{\limsup_{n\to\infty} 2^n} = \frac{1}{+\infty} = 0.$$

The series $\sum_{n=0}^{\infty} 2^{-n^2} (x+2)^n$ has radius of convergence

$$\frac{1}{\limsup_{n\to\infty} |2^{-n^2}|^{1/n}} = \frac{1}{\limsup_{n\to\infty} 2^{-n}} = \frac{1}{0} = +\infty.$$

The significance of the radius of convergence is the following.

Theorem 4.1.6 *Let $\sum_{n=0}^{\infty} c_n (x-a)^n$ be a formal power series, and let R be its radius of convergence.*

(a) (Divergence outside of the radius of convergence) If $x \in \mathbf{R}$ is such that $|x-a| > R$, then the series $\sum_{n=0}^{\infty} c_n (x-a)^n$ is divergent for that value of x.
(b) (Convergence inside the radius of convergence) If $x \in \mathbf{R}$ is such that $|x-a| < R$, then the series $\sum_{n=0}^{\infty} c_n (x-a)^n$ is absolutely convergent for that value of x.

For parts (c)-(e) we assume that $R > 0$ (i.e., the series converges at at least one other point than $x = a$). Let $f\colon (a-R, a+R) \to \mathbf{R}$ be the function $f(x) := \sum_{n=0}^{\infty} c_n (x-a)^n$; this function is guaranteed to exist by (b).

(c) (Uniform convergence on compact sets) For any $0 < r < R$, the series $\sum_{n=0}^{\infty} c_n (x-a)^n$ converges uniformly to f on the compact interval $[a-r, a+r]$. In particular, f is continuous on $(a-R, a+R)$.
(d) (Differentiation of power series) The function f is differentiable on $(a-R, a+R)$, and for any $0 < r < R$, the series $\sum_{n=1}^{\infty} nc_n (x-a)^{n-1}$ converges uniformly to f' on the interval $[a-r, a+r]$.
(e) (Integration of power series) For any closed interval $[y, z]$ contained in $(a-R, a+R)$, we have

$$\int_{[y,z]} f = \sum_{n=0}^{\infty} c_n \frac{(z-a)^{n+1} - (y-a)^{n+1}}{n+1}.$$

Proof See Exercise 4.1.1. □

Theorem 4.1.6 (a) and (b) of the above theorem give another way to find the radius of convergence, by using your favorite convergence test to work out the range of x for which the power series converges:

Example 4.1.7 Consider the power series $\sum_{n=0}^{\infty} n(x-1)^n$. The ratio test shows that this series converges when $|x-1| < 1$ and diverges when $|x-1| > 1$ (why?). Thus the only possible value for the radius of convergence is $R = 1$ (if $R < 1$, then we have contradicted Theorem 4.1.6(a); if $R > 1$, then we have contradicted Theorem 4.1.6(b)).

Remark 4.1.8 Theorem 4.1.6 is silent on what happens when $|x-a| = R$, i.e., at the points $a - R$ and $a + R$. Indeed, one can have either convergence or divergence at those points; see Exercise 4.1.2.

Remark 4.1.9 Note that while Theorem 4.1.6(b) assures us that the power series $\sum_{n=0}^{\infty} c_n(x-a)^n$ will converge pointwise on the interval $(a-R, a+R)$, it need not converge uniformly on that interval (see Exercise 4.1.2(e)). On the other hand, Theorem 4.1.6(c) assures us that the power series will converge on any smaller interval $[a-r, a+r]$. In particular, being uniformly convergent on every closed subinterval of $(a-R, a+R)$ is not enough to guarantee being uniformly convergent on all of $(a-R, a+R)$.

—Exercise—

Exercise 4.1.1 Prove Theorem 4.1.6. (*Hints:* for (a) and (b), use the root test (Theorem 7.5.1). For (c), use the Weierstrass M-test (Theorem 3.5.7). For (d), use Theorem 3.7.1. For (e), use Corollary 3.6.2.)

Exercise 4.1.2 Give examples of a formal power series $\sum_{n=0}^{\infty} c_n x^n$ centered at 0 with radius of convergence 1, which

(a) diverges at both $x = 1$ and $x = -1$;
(b) diverges at $x = 1$ but converges at $x = -1$;
(c) converges at $x = 1$ but diverges at $x = -1$;
(d) converges at both $x = 1$ and $x = -1$.
(e) converges pointwise on $(-1, 1)$, but does not converge uniformly on $(-1, 1)$.

4.2 Real Analytic Functions

A function $f(x)$ which is lucky enough to be representable as a power series has a special name; it is a *real analytic* function.

Definition 4.2.1 (*Real analytic functions*) Let E be a subset of $\mathbf{R}$, and let $f : E \to \mathbf{R}$ be a function. If a is an interior point of E, we say that f is *real analytic at* a if there exists an open interval $(a - r, a + r)$ in E for some $r > 0$ such that there exists a power series $\sum_{n=0}^{\infty} c_n (x - a)^n$ centered at a which has a radius of convergence greater than or equal to r and which converges to f on $(a - r, a + r)$. If E is an open set, and f is real analytic at every point a of E, we say that f is *real analytic on* E.

Example 4.2.2 Consider the function $f : \mathbf{R}\backslash\{1\} \to \mathbf{R}$ defined by $f(x) := 1/(1 - x)$. This function is real analytic at 0 because we have a power series $\sum_{n=0}^{\infty} x^n$ centered at 0 which converges to $1/(1 - x) = f(x)$ on the interval $(-1, 1)$. This function is also real analytic at 2 because we have a power series $\sum_{n=0}^{\infty} (-1)^{n+1} (x - 2)^n$ which converges to $\frac{-1}{1-(-(x-2))} = \frac{1}{1-x} = f(x)$ on the interval $(1, 3)$ (why? Use Lemma 7.3.3). In fact this function is real analytic on all of $\mathbf{R}\backslash\{1\}$; see Exercise 4.2.2.

Remark 4.2.3 The notion of being real analytic is closely related to another notion, that of being *complex analytic*, but this is a topic for complex analysis, and will not be discussed here.

We now discuss which functions are real analytic. From Theorem 4.1.6(c) and (d) we see that if f is real analytic at a point a, then f is both continuous and differentiable on $(a - r, a + r)$ for some $r > 0$. We can in fact say more:

Definition 4.2.4 (*k-times differentiability*) Let E be a subset of $\mathbf{R}$ with the property that every element of E is a limit point of E. We say a function $f : E \to \mathbf{R}$ is *once differentiable on* E iff it is differentiable (so in particular $f' : E \to \mathbf{R}$ is also a function on E. More generally, for any $k \geq 2$ we say that $f : E \to \mathbf{R}$ is *k times differentiable on* E, or just *k times differentiable*, iff f is differentiable, and f' is $k - 1$ times differentiable. If f is k times differentiable, we define the k^{th} derivative $f^{(k)} : E \to \mathbf{R}$ by the recursive rule $f^{(1)} := f'$, and $f^{(k)} := (f^{(k-1)})'$ for all $k \geq 2$. We also define $f^{(0)} := f$ (this is f differentiated 0 times), and we allow every function to be zero times differentiable (since clearly $f^{(0)}$ exists for every f). A function is said to be *infinitely differentiable* (or *smooth*) iff it is k times differentiable for every $k \geq 0$.

Example 4.2.5 The function $f(x) := |x|^3$ is twice differentiable on $\mathbf{R}$, but not three times differentiable (why?). Indeed, $f^{(2)} = f'' = 6|x|$, which is not differentiable, at 0.

Proposition 4.2.6 (Real analytic functions are k-times differentiable) *Let E be a subset of $\mathbf{R}$, let a be an interior point of E, and and let f be a function which is real analytic at a, thus there is an $r > 0$ for which we have the power series expansion*

$$f(x) = \sum_{n=0}^{\infty} c_n (x - a)^n$$

for all $x \in (a - r, a + r)$. Then for every $k \geq 0$, the function f is k-times differentiable on $(a - r, a + r)$, and for each $k \geq 0$ the k^{th} derivative is given by

$$f^{(k)}(x) = \sum_{n=0}^{\infty} c_{n+k}(n+1)(n+2)\dots(n+k)(x-a)^n$$
$$= \sum_{n=0}^{\infty} c_{n+k}\frac{(n+k)!}{n!}(x-a)^n$$

for all $x \in (a-r, a+r)$.

Proof See Exercise 4.2.3. □

Corollary 4.2.7 (Real analytic functions are infinitely differentiable) *Let E be an open subset of* $\mathbf{R}$*, and let* $f: E \to \mathbf{R}$ *be a real analytic function on* E*. Then* f *is infinitely differentiable on* E*. Also, all derivatives of* f *are also real analytic on* E*.*

Proof For every point $a \in E$ and $k \geq 0$, we know from Proposition 4.2.6 that f is k-times differentiable at a (we will have to apply Exercise 10.1.1 k times here, why?). Thus f is k-times differentiable on E for every $k \geq 0$ and is hence infinitely differentiable. Also, from Proposition 4.2.6 we see that each derivative $f^{(k)}$ of f has a convergent power series expansion at every $x \in E$ and thus $f^{(k)}$ is real analytic. □

Example 4.2.8 Consider the function $f: \mathbf{R} \to \mathbf{R}$ defined by $f(x) := |x|$. This function is not differentiable at $x = 0$ and hence cannot be real analytic at $x = 0$. It is however real analytic at every other point $x \in \mathbf{R}\backslash\{0\}$ (why?).

Remark 4.2.9 The converse statement to Corollary 4.2.7 is not true; there are infinitely differentiable functions which are not real analytic. See Exercise 4.5.4.

Proposition 4.2.6 has an important corollary, due to Brook Taylor (1685–1731).

Corollary 4.2.10 (Taylor's formula) *Let* E *be a subset of* $\mathbf{R}$*, let* a *be an interior point of* E*, and let* $f: E \to \mathbf{R}$ *be a function which is real analytic at* a *and has the power series expansion*

$$f(x) = \sum_{n=0}^{\infty} c_n(x-a)^n$$

for all $x \in (a-r, a+r)$ *and some* $r > 0$*. Then for any integer* $k \geq 0$*, we have*

$$f^{(k)}(a) = k!c_k,$$

where $k! := 1 \times 2 \times \dots \times k$ *(and we adopt the convention that* $0! = 1$*). In particular, we have* Taylor's formula

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n$$

for all x *in* $(a-r, a+r)$.

Proof See Exercise 4.2.4. □

The power series $\sum_{n=0}^{\infty} \frac{f^{(n)}(a)}{n!}(x-a)^n$ is sometimes called the *Taylor series* of f around a. Taylor's formula thus asserts that if a function is real analytic, then it is equal to its Taylor series.

Remark 4.2.11 Note that Taylor's formula only works for functions which are real analytic; there are examples of functions which are infinitely differentiable but for which Taylor's theorem fails (see Exercise 4.5.4).

Another important corollary of Taylor's formula is that a real analytic function can have at most one power series at a point:

Corollary 4.2.12 (Uniqueness of power series) *Let E be a subset of $\mathbf{R}$, let a be an interior point of E, and let $f : E \to \mathbf{R}$ be a function which is real analytic at a. Suppose that f has two power series expansions*

$$f(x) = \sum_{n=0}^{\infty} c_n (x-a)^n$$

and

$$f(x) = \sum_{n=0}^{\infty} d_n (x-a)^n$$

centered at a, each with a nonzero radius of convergence. Then $c_n = d_n$ for all $n \geq 0$.

Proof By Corollary 4.2.10, we have $f^{(k)}(a) = k!c_k$ for all $k \geq 0$. But we also have $f^{(k)}(a) = k!d_k$, by similar reasoning. Since $k!$ is never zero, we can cancel it and obtain $c_k = d_k$ for all $k \geq 0$, as desired. □

Remark 4.2.13 While a real analytic function has a unique power series around any given point, it can certainly have different power series at different points. For instance, the function $f(x) := \frac{1}{1-x}$, defined on $\mathbf{R} - \{1\}$, has the power series

$$f(x) := \sum_{n=0}^{\infty} x^n$$

around 0, on the interval $(-1, 1)$, but also has the power series

$$\begin{aligned} f(x) &= \frac{1}{1-x} = \frac{2}{1-2(x-\frac{1}{2})} \\ &= \sum_{n=0}^{\infty} 2\left(2\left(x-\frac{1}{2}\right)\right)^n = \sum_{n=0}^{\infty} 2^{n+1}\left(x-\frac{1}{2}\right)^n \end{aligned}$$

around 1/2, on the interval $(0, 1)$ (note that the above power series has a radius of convergence of $1/2$, thanks to the root test; see also Exercise 4.2.8).

—Exercise—

Exercise 4.2.1 Let $n \geq 0$ be an integer, let c, a be real numbers, and let f be the function $f(x) := c(x-a)^n$. Show that f is infinitely differentiable, and that $f^{(k)}(x) = c\frac{n!}{(n-k)!}(x-a)^{n-k}$ for all integers $0 \leq k \leq n$. What happens when $k > n$?

Exercise 4.2.2 Show that the function f defined in Example 4.2.2 is real analytic on all of $\mathbf{R}\backslash\{1\}$.

Exercise 4.2.3 Prove Proposition 4.2.6. (*Hint:* induct on k and use Theorem 4.1.6(d).)

Exercise 4.2.4 Use Proposition 4.2.6 and Exercise 4.2.1 to prove Corollary 4.2.10.

Exercise 4.2.5 Let a, b be real numbers, and let $n \geq 0$ be an integer. Prove the identity

$$(x-a)^n = \sum_{m=0}^{n} \frac{n!}{m!(n-m)!}(b-a)^{n-m}(x-b)^m$$

for any real number x. (*Hint:* use the binomial formula, Exercise 7.1.4.) Explain why this identity is consistent with Taylor's theorem and Exercise 4.2.1. (Note however that Taylor's theorem cannot be rigorously applied until one verifies Exercise 4.2.6 below.)

Exercise 4.2.6 Using Exercise 4.2.5, show that every polynomial $P(x)$ of one variable is real analytic on $\mathbf{R}$.

Exercise 4.2.7 Let $m \geq 0$ be a positive integer, and let $0 < x < r$ be real numbers. Use Lemma 7.3.3 to establish the identity

$$\frac{r}{r-x} = \sum_{n=0}^{\infty} x^n r^{-n}$$

for all $x \in (-r, r)$. Using Proposition 4.2.6, conclude the identity

$$\frac{r}{(r-x)^{m+1}} = \sum_{n=m}^{\infty} \frac{n!}{m!(n-m)!} x^{n-m} r^{-n}$$

for all integers $m \geq 0$ and $x \in (-r, r)$. Also explain why the series on the right-hand side is absolutely convergent.

Exercise 4.2.8 Let E be a subset of $\mathbf{R}$, let a be an interior point of E, and let $f: E \to \mathbf{R}$ be a function which is real analytic at a and has a power series expansion

$$f(x) = \sum_{n=0}^{\infty} c_n (x-a)^n$$

at a which converges on the interval $(a-r, a+r)$. Let $(b-s, b+s)$ be any subinterval of $(a-r, a+r)$ for some $s > 0$.

(a) Prove that $|a - b| \leq r - s$, so in particular $|a - b| < r$.
(b) Show that for every $0 < \varepsilon < r$, there exists a $C > 0$ such that $|c_n| \leq C(r - \varepsilon)^{-n}$ for all integers $n \geq 0$. (*Hint:* what do we know about the radius of convergence of the series $\sum_{n=0}^{\infty} c_n(x - a)^n$?)
(c) Show that the numbers $d_0, d_1, \ldots$ given by the formula

$$d_m := \sum_{n=m}^{\infty} \frac{n!}{m!(n-m)!}(b-a)^{n-m} c_n \text{ for all integers } m \geq 0$$

are well-defined, in the sense that the above series is absolutely convergent. (*Hint:* use (b) and the comparison test, Corollary 7.3.2, followed by Exercise 4.2.7.)
(d) Show that for every $0 < \varepsilon < s$ there exists a $C > 0$ such that

$$|d_m| \leq C(s - \varepsilon)^{-m}$$

for all integers $m \geq 0$. (*Hint:* use the comparison test, and Exercise 4.2.7.)
(e) Show that the power series $\sum_{m=0}^{\infty} d_m(x - b)^m$ is absolutely convergent for $x \in (b - s, b + s)$ and converges to $f(x)$. (You may need Fubini's theorem for infinite series, Theorem 8.2.2 of *Analysis I*, as well as Exercise 4.2.5. One may also need to study a variant of the d_m in which the c_n are replaced by $|c_n|$.)
(f) Conclude that f is real analytic at every point in $(a - r, a + r)$.

4.3 Abel's Theorem

Let $f(x) = \sum_{n=0}^{\infty} c_n(x - a)^n$ be a power series centered at a with a radius of convergence $0 < R < \infty$ strictly between 0 and infinity. From Theorem 4.1.6 we know that the power series converges absolutely whenever $|x - a| < R$ and diverges when $|x - a| > R$. However, at the boundary $|x - a| = R$ the situation is more complicated; the series may either converge or diverge (see Exercise 4.1.2). However, if the series does converge at the boundary point, then it is reasonably well-behaved; in particular, it is continuous at that boundary point.

Theorem 4.3.1 (Abel's theorem) *Let $f(x) = \sum_{n=0}^{\infty} c_n(x - a)^n$ be a power series centered at a with radius of convergence $0 < R < \infty$. If the power series converges at $a + R$, then f is continuous at $a + R$, i.e.,*

$$\lim_{x \to a+R : x \in (a-R, a+R)} \sum_{n=0}^{\infty} c_n(x - a)^n = \sum_{n=0}^{\infty} c_n R^n.$$

Similarly, if the power series converges at $a - R$, then f is continuous at $a - R$, i.e.,

$$\lim_{x\to a-R: x\in(a-R,a+R)} \sum_{n=0}^{\infty} c_n(x-a)^n = \sum_{n=0}^{\infty} c_n(-R)^n.$$

Before we prove Abel's theorem, we need the following lemma.

Lemma 4.3.2 (Summation by parts formula) *Let* $(a_n)_{n=0}^{\infty}$ *and* $(b_n)_{n=0}^{\infty}$ *be sequences of real numbers which converge to limits A and B, respectively, i.e.,* $\lim_{n\to\infty} a_n = A$ *and* $\lim_{n\to\infty} b_n = B$. *Suppose that the sum* $\sum_{n=0}^{\infty}(a_{n+1} - a_n)b_n$ *is convergent. Then the sum* $\sum_{n=0}^{\infty} a_{n+1}(b_{n+1} - b_n)$ *is also convergent, and*

$$\sum_{n=0}^{\infty}(a_{n+1} - a_n)b_n = AB - a_0b_0 - \sum_{n=0}^{\infty} a_{n+1}(b_{n+1} - b_n).$$

Proof See Exercise 4.3.1. □

Remark 4.3.3 One should compare this formula with the more well-known *integration by parts formula*

$$\int_0^{\infty} f'(x)g(x)\,dx = f(x)g(x)|_0^{\infty} - \int_0^{\infty} f(x)g'(x)\,dx,$$

see Proposition 11.10.1.

Proof of Abel's theorem It will suffice to prove the first claim, i.e., that

$$\lim_{x\to a+R: x\in(a-R,a+R)} \sum_{n=0}^{\infty} c_n(x-a)^n = \sum_{n=0}^{\infty} c_n R^n$$

whenever the sum $\sum_{n=0}^{\infty} c_n R^n$ converges; the second claim will then follow (why?) by replacing c_n by $(-1)^n c_n$ in the above claim. If we make the substitutions $d_n := c_n R^n$ and $y := \frac{x-a}{R}$, then the above claim can be rewritten as

$$\lim_{y\to 1: y\in(-1,1)} \sum_{n=0}^{\infty} d_n y^n = \sum_{n=0}^{\infty} d_n$$

whenever the sum $\sum_{n=0}^{\infty} d_n$ converges. (Why is this equivalent to the previous claim?)

Write $D := \sum_{n=0}^{\infty} d_n$, and for every $N \geq 0$ write

$$S_N := \left(\sum_{n=0}^{N-1} d_n\right) - D$$

so in particular $S_0 = -D$. Then observe that $\lim_{N\to\infty} S_N = 0$, and that $d_n = S_{n+1} - S_n$. Thus for any $y \in (-1, 1)$ we have

$$\sum_{n=0}^{\infty} d_n y^n = \sum_{n=0}^{\infty} (S_{n+1} - S_n) y^n.$$

Applying the summation by parts formula (Lemma 4.3.2), and noting that $\lim_{n\to\infty} y^n = 0$, we obtain

$$\sum_{n=0}^{\infty} d_n y^n = -S_0 y^0 - \sum_{n=0}^{\infty} S_{n+1}(y^{n+1} - y^n).$$

Observe that $-S_0 y^0 = +D$. Thus to finish the proof of Abel's theorem, it will suffice to show that

$$\lim_{y\to 1: y\in(-1,1)} \sum_{n=0}^{\infty} S_{n+1}(y^{n+1} - y^n) = 0.$$

Since y converges to 1, we may as well restrict y to $[0, 1)$ instead of $(-1, 1)$; in particular we may take y to be positive.

From the triangle inequality for series (Proposition 7.2.9), we have

$$\left|\sum_{n=0}^{\infty} S_{n+1}(y^{n+1} - y^n)\right| \leq \sum_{n=0}^{\infty} |S_{n+1}(y^{n+1} - y^n)|$$
$$= \sum_{n=0}^{\infty} |S_{n+1}|(y^n - y^{n+1}),$$

so by the squeeze test (Corollary 6.4.14) it suffices to show that

$$\lim_{y\to 1: y\in[0,1)} \sum_{n=0}^{\infty} |S_{n+1}|(y^n - y^{n+1}) = 0.$$

The expression $\sum_{n=0}^{\infty} |S_{n+1}|(y^n - y^{n+1})$ is clearly non-negative, so it will suffice to show that

$$\limsup_{y\to 1: y\in[0,1)} \sum_{n=0}^{\infty} |S_{n+1}|(y^n - y^{n+1}) = 0.$$

Let $\varepsilon > 0$. Since S_n converges to 0, there exists an N such that $|S_n| \leq \varepsilon$ for all $n > N$. Thus we have

$$\sum_{n=0}^{\infty} |S_{n+1}|(y^n - y^{n+1}) \leq \sum_{n=0}^{N} |S_{n+1}|(y^n - y^{n+1}) + \sum_{n=N+1}^{\infty} \varepsilon(y^n - y^{n+1}).$$

The last summation is a telescoping series, which sums to εy^{N+1} (See Lemma 7.2.14, recalling from Lemma 6.5.2 that $y^n \to 0$ as $n \to \infty$), and thus

$$\sum_{n=0}^{\infty} |S_{n+1}|(y^n - y^{n+1}) \leq \sum_{n=0}^{N} |S_{n+1}|(y^n - y^{n+1}) + \varepsilon y^{N+1}.$$

Now take limits as $y \to 1$. Observe that $y^n - y^{n+1} \to 0$ as $y \to 1$ for every $n \in 0, 1, \ldots, N$. Since we can interchange limits and *finite* sums (Exercise 7.1.5), we thus have

$$\limsup_{y \to 1: y \in [0,1)} \sum_{n=0}^{\infty} |S_{n+1}|(y^n - y^{n+1}) \leq \varepsilon.$$

But $\varepsilon > 0$ was arbitrary, and thus we must have

$$\limsup_{y \to 1: y \in [0,1)} \sum_{n=0}^{\infty} |S_{n+1}|(y^n - y^{n+1}) = 0$$

since the left-hand side must be non-negative. The claim follows. □

—Exercise—

Exercise 4.3.1 Prove Lemma 4.3.2. (*Hint:* first work out the relationship between the partial sums $\sum_{n=0}^{N}(a_{n+1} - a_n)b_n$ and $\sum_{n=0}^{N} a_{n+1}(b_{n+1} - b_n)$.)

4.4 Multiplication of Power Series

We now show that the product of two real analytic functions is again real analytic.

Theorem 4.4.1 *Let $f : (a - r, a + r) \to \mathbf{R}$ and $g : (a - r, a + r) \to \mathbf{R}$ be functions analytic on $(a - r, a + r)$, with power series expansions*

$$f(x) = \sum_{n=0}^{\infty} c_n (x - a)^n$$

and

$$g(x) = \sum_{n=0}^{\infty} d_n (x - a)^n$$

, respectively. Then $fg : (a - r, a + r) \to \mathbf{R}$ is also analytic on $(a - r, a + r)$, with power series expansion

$$f(x)g(x) = \sum_{n=0}^{\infty} e_n (x - a)^n$$

where $e_n := \sum_{m=0}^{n} c_m d_{n-m}$.

Remark 4.4.2 The sequence $(e_n)_{n=0}^\infty$ is sometimes referred to as the *convolution* of the sequences $(c_n)_{n=0}^\infty$ and $(d_n)_{n=0}^\infty$; it is closely related (though not identical) to the notion of convolution introduced in Definition 3.8.9.

Proof We have to show that the series $\sum_{n=0}^\infty e_n(x-a)^n$ converges to $f(x)g(x)$ for all $x \in (a-r, a+r)$. Now fix x to be any point in $(a-r, a+r)$. By Theorem 4.1.6, we see that both f and g have radii of convergence at least r. In particular, the series $\sum_{n=0}^\infty c_n(x-a)^n$ and $\sum_{n=0}^\infty d_n(x-a)^n$ are absolutely convergent. Thus if we define

$$C := \sum_{n=0}^\infty |c_n(x-a)^n|$$

and

$$D := \sum_{n=0}^\infty |d_n(x-a)^n|$$

then C and D are both finite.

For any $N \geq 0$, consider the partial sum

$$\sum_{n=0}^N \sum_{m=0}^\infty |c_m(x-a)^m d_n(x-a)^n|.$$

We can rewrite this as

$$\sum_{n=0}^N |d_n(x-a)^n| \sum_{m=0}^\infty |c_m(x-a)^m|,$$

which by definition of C is equal to

$$\sum_{n=0}^N |d_n(x-a)^n| C,$$

which by definition of D is less than or equal to DC. Thus the above partial sums are bounded by DC for every N. In particular, the series

$$\sum_{n=0}^\infty \sum_{m=0}^\infty |c_m(x-a)^m d_n(x-a)^n|$$

is convergent, which means that the sum

$$\sum_{n=0}^\infty \sum_{m=0}^\infty c_m(x-a)^m d_n(x-a)^n$$

is absolutely convergent.

Let us now compute this sum in two ways. First of all, we can pull the $d_n(x-a)^n$ factor out of the $\sum_{m=0}^{\infty}$ summation, to obtain

$$\sum_{n=0}^{\infty} d_n(x-a)^n \sum_{m=0}^{\infty} c_m(x-a)^m.$$

By our formula for $f(x)$, this is equal to

$$\sum_{n=0}^{\infty} d_n(x-a)^n f(x);$$

by our formula for $g(x)$, this is equal to $f(x)g(x)$. Thus

$$f(x)g(x) = \sum_{n=0}^{\infty}\sum_{m=0}^{\infty} c_m(x-a)^m d_n(x-a)^n.$$

Now we compute this sum in a different way. We rewrite it as

$$f(x)g(x) = \sum_{n=0}^{\infty}\sum_{m=0}^{\infty} c_m d_n(x-a)^{n+m}.$$

By Fubini's theorem for series (Theorem 8.2.2), because the series was absolutely convergent, we may rewrite it as

$$f(x)g(x) = \sum_{m=0}^{\infty}\sum_{n=0}^{\infty} c_m d_n(x-a)^{n+m}.$$

Now make the substitution $n' := n+m$, to rewrite this as

$$f(x)g(x) = \sum_{m=0}^{\infty}\sum_{n'=m}^{\infty} c_m d_{n'-m}(x-a)^{n'}.$$

If we adopt the convention that $d_j = 0$ for all negative j, then this is equal to

$$f(x)g(x) = \sum_{m=0}^{\infty}\sum_{n'=0}^{\infty} c_m d_{n'-m}(x-a)^{n'}.$$

Applying Fubini's theorem again, we obtain

$$f(x)g(x) = \sum_{n'=0}^{\infty} \sum_{m=0}^{\infty} c_m d_{n'-m} (x-a)^{n'},$$

which we can rewrite as

$$f(x)g(x) = \sum_{n'=0}^{\infty} (x-a)^{n'} \sum_{m=0}^{\infty} c_m d_{n'-m}.$$

Since d_j was 0 when j is negative, we can rewrite this as

$$f(x)g(x) = \sum_{n'=0}^{\infty} (x-a)^{n'} \sum_{m=0}^{n'} c_m d_{n'-m},$$

which by definition of e is

$$f(x)g(x) = \sum_{n'=0}^{\infty} e_{n'} (x-a)^{n'},$$

as desired. □

4.5 The Exponential and Logarithm Functions

We can now use the machinery developed in the last few sections to develop a rigorous foundation for many standard functions used in mathematics. We begin with the exponential function.

Definition 4.5.1 (*Exponential function*) For every real number x, we define the *exponential function* $\exp(x)$ to be the real number

$$\exp(x) := \sum_{n=0}^{\infty} \frac{x^n}{n!}.$$

Theorem 4.5.2 (Basic properties of exponential)

(a) *For every real number* x*, the series* $\sum_{n=0}^{\infty} \frac{x^n}{n!}$ *is absolutely convergent. In particular,* $\exp(x)$ *exists and is real for every* $x \in \mathbf{R}$*, the power series* $\sum_{n=0}^{\infty} \frac{x^n}{n!}$ *has an infinite radius of convergence, and* exp *is a real analytic function on* $(-\infty, \infty)$.

(b) exp *is differentiable on* $\mathbf{R}$*, and for every* $x \in \mathbf{R}$, $\exp'(x) = \exp(x)$.

(c) exp *is continuous on* $\mathbf{R}$*, and for every interval* $[a, b]$*, we have* $\int_{[a,b]} \exp(x)\, dx = \exp(b) - \exp(a)$.

(d) *For every* $x, y \in \mathbf{R}$*, we have* $\exp(x + y) = \exp(x) \exp(y)$.

(e) *We have* $\exp(0) = 1$. *Also, for every* $x \in \mathbf{R}$, $\exp(x)$ *is positive, and* $\exp(-x) = 1/\exp(x)$.
(f) $\exp$ *is strictly monotone increasing: in other words, if* x, y *are real numbers, then we have* $\exp(y) > \exp(x)$ *if and only if* $y > x$.

Proof See Exercise 4.5.1. □

One can write the exponential function in a more compact form, introducing famous *Euler's number* $e = 2.71828183\ldots$, also known as the *base of the natural logarithm*:

Definition 4.5.3 (*Euler's number*) The number e is defined to be

$$e := \exp(1) = \sum_{n=0}^{\infty} \frac{1}{n!} = \frac{1}{0!} + \frac{1}{1!} + \frac{1}{2!} + \frac{1}{3!} + \ldots.$$

Proposition 4.5.4 *For every real number* x, *we have* $\exp(x) = e^x$.

Proof See Exercise 4.5.3. □

In light of this proposition we can and will use e^x and $\exp(x)$ interchangeably.

Since $e > 1$ (why?), we see that $e^x \to +\infty$ as $x \to +\infty$, and $e^x \to 0$ as $x \to -\infty$. From this and the intermediate value theorem (Theorem 9.7.1) we see that the range of the function exp is $(0, \infty)$. Since exp is strictly increasing, it is injective, and hence exp is a bijection from $\mathbf{R}$ to $(0, \infty)$ and thus has an inverse from $(0, \infty) \to \mathbf{R}$. This inverse has a name:

Definition 4.5.5 (*Logarithm*) We define the *natural logarithm function* $\log : (0, \infty) \to \mathbf{R}$ (also called ln) to be the inverse of the exponential function. Thus $\exp(\log(x)) = x$ and $\log(\exp(x)) = x$.

Since exp is continuous and strictly monotone increasing, we see that log is also continuous and strictly monotone increasing (see Proposition 9.8.3). Since exp is also differentiable, and the derivative is never zero, we see from the inverse function theorem (Theorem 10.4.2) that log is also differentiable. We list some other properties of the natural logarithm below.

Theorem 4.5.6 (Logarithm properties)

(a) *For every* $x \in (0, \infty)$, *we have* $\ln'(x) = \frac{1}{x}$. *In particular, by the fundamental theorem of calculus, we have* $\int_{[a,b]} \frac{1}{x}\, dx = \ln(b) - \ln(a)$ *for any interval* $[a, b]$ *in* $(0, \infty)$.
(b) *We have* $\ln(xy) = \ln(x) + \ln(y)$ *for all* $x, y \in (0, \infty)$.
(c) *We have* $\ln(1) = 0$ *and* $\ln(1/x) = -\ln(x)$ *for all* $x \in (0, \infty)$.
(d) *For any* $x \in (0, \infty)$ *and* $y \in \mathbf{R}$, *we have* $\ln(x^y) = y \ln(x)$.

(e) For any $x \in (-1, 1)$*, we have*

$$\ln(1 - x) = -\sum_{n=1}^{\infty} \frac{x^n}{n}.$$

In particular, ln *is analytic at* 1*, with the power series expansion*

$$\ln(x) = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}(x - 1)^n$$

for $x \in (0, 2)$*, with radius of convergence 1.*

Proof See Exercise 4.5.5. □

Example 4.5.7 We now give a modest application of Abel's theorem (Theorem 4.3.1): from the alternating series test we see that $\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}$ is convergent. By Abel's theorem we thus see that

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} = \lim_{x \to 2} \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}(x - 1)^n$$

$$= \lim_{x \to 2} \ln(x) = \ln(2),$$

thus we have the formula

$$\ln(2) = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \ldots.$$

—Exercise—

Exercise 4.5.1 Prove Theorem 4.5.2. (*Hints:* for part (a), use the ratio test. For parts (bc), use Theorem 4.1.6. For part (d), use Theorem 4.4.1. For part (e), use part (d). For part (f), use part (d), and prove that $\exp(x) > 1$ when x is positive. You may find the binomial formula from Exercise 7.1.4 to be useful.)

Exercise 4.5.2 Show that for every integer $n \geq 3$, we have

$$0 < \frac{1}{(n+1)!} + \frac{1}{(n+2)!} + \ldots < \frac{1}{n!}.$$

(*Hint:* first show that $(n + k)! > 2^k n!$ for all $k = 1, 2, 3, \ldots$.) Conclude that $n!e$ is not an integer for every $n \geq 3$. Deduce from this that e is irrational. (*Hint:* prove by contradiction.)

Exercise 4.5.3 Prove Proposition 4.5.4. (*Hint:* first prove the claim when x is a natural number. Then prove it when x is an integer. Then prove it when x is a rational number. Then use the fact that real numbers are the limits of rational numbers to prove it for all real numbers. You may find the exponent laws (Proposition 6.7.3) to be useful.)

Exercise 4.5.4 Let $f: \mathbf{R} \to \mathbf{R}$ be the function defined by setting $f(x) := \exp(-1/x)$ when $x > 0$, and $f(x) := 0$ when $x \leq 0$. Prove that f is infinitely differentiable, and $f^{(k)}(0) = 0$ for every integer $k \geq 0$, but that f is not real analytic at 0.

Exercise 4.5.5 Prove Theorem 4.5.6. (*Hints:* for part (a), use the inverse function theorem (Theorem 10.4.2) or the chain rule (Theorem 10.1.15). For parts (bcd), use Theorem 4.5.2 and the exponent laws (Proposition 6.7.3). For part (e), start with the geometric series formula (Lemma 7.3.3) and integrate using Theorem 4.1.6).

Exercise 4.5.6 Prove that the natural logarithm function is real analytic on $(0, +\infty)$.

Exercise 4.5.7 Let $f: \mathbf{R} \to (0, \infty)$ be a positive, real analytic function such that $f'(x) = f(x)$ for all $x \in \mathbf{R}$. Show that $f(x) = Ce^x$ for some positive constant C; justify your reasoning. (*Hint:* there are basically three different proofs available. One proof uses the logarithm function, another proof uses the function e^{-x}, and a third proof uses power series. Of course, you only need to supply one proof.)

Exercise 4.5.8 Let $m > 0$ be an integer. Show that

$$\lim_{x \to +\infty} \frac{e^x}{x^m} = +\infty.$$

(*Hint:* what happens to the ratio between $e^{x+1}/(x+1)^m$ and e^x/x^m as $x \to +\infty$?)

Exercise 4.5.9 Let $P(x)$ be a polynomial, and let $c > 0$. Show that there exists a real number $N > 0$ such that $e^{cx} > |P(x)|$ for all $x > N$; thus an exponentially growing function, no matter how small the growth rate c, will eventually overtake any given polynomial $P(x)$, no matter how large. (*Hint:* use Exercise 4.5.8.)

Exercise 4.5.10 Let $f: (0, +\infty) \times \mathbf{R} \to \mathbf{R}$ be the exponential function $f(x, y) := x^y$. Show that f is continuous. (*Hint:* note that Propositions 9.4.10, 9.4.11 only show that f is continuous in each variable, which is insufficient, as Exercise 2.2.11 shows. The easiest way to proceed is to write $f(x, y) = \exp(y \ln x)$ and use the continuity of exp() and ln(). For an extra challenge, try proving this exercise without using the logarithm function.)

4.6 A Digression on Complex Numbers

To proceed further we need the complex number system $\mathbf{C}$, which is an extension of the real number system $\mathbf{R}$. A full discussion of this important number system (and in particular the branch of mathematics known as *complex analysis*) is beyond the scope of this text; here, we need the system primarily because of a very useful mathematical operation, the *complex exponential function* $z \mapsto \exp(z)$, which generalizes the real exponential function $x \mapsto \exp(x)$ introduced in the previous section.

Informally, we could define the complex numbers as

Definition 4.6.1 (*Informal definition of complex numbers*) The complex numbers $\mathbf{C}$ are the set of all numbers of the form $a + bi$, where a, b are real numbers and i is a square root of -1, $i^2 = -1$.

However, this definition is a little unsatisfactory as it does not explain how to add, multiply, or compare two complex numbers. To construct the complex numbers rigorously we will first introduce a *formal* version of the complex number $a + bi$, which we shall temporarily denote as (a, b); this is similar to how in Chap. 4, when constructing the integers $\mathbf{Z}$, we needed a formal notion of subtraction $a \text{---} b$ before the actual notion of subtraction $a - b$ could be introduced, or how when constructing the rational numbers, a formal notion of division $a//b$ was needed before it was superceded by the actual notion a/b of division. It is also similar to how, in the construction of the real numbers, we defined a formal limit $\text{LIM}_{n\to\infty} a_n$ before we defined a genuine limit $\lim_{n\to\infty} a_n$.

Definition 4.6.2 (*Formal definition of complex numbers*) A *complex number* is any pair of the form (a, b), where a, b are real numbers, thus for instance $(2, 4)$ is a complex number. Two complex numbers (a, b), (c, d) are said to be equal iff $a = c$ and $b = d$, thus for instance $(2 + 1, 3 + 4) = (3, 7)$, but $(2, 1) \neq (1, 2)$ and $(2, 4) \neq (2, -4)$. The set of all complex numbers is denoted $\mathbf{C}$.

At this stage the complex numbers $\mathbf{C}$ are indistinguishable from the Cartesian product $\mathbf{R}^2 = \mathbf{R} \times \mathbf{R}$ (also known as the *Cartesian plane*). However, we will introduce a number of operations on the complex numbers, notably that of *complex multiplication*, which are not normally placed on the Cartesian plane $\mathbf{R}^2$. Thus one can think of the complex number system $\mathbf{C}$ as the Cartesian plane $\mathbf{R}^2$ equipped with a number of additional structures. We begin with the notion of addition and negation. Using the informal definition of the complex numbers, we expect

$$(a, b) + (c, d) = (a + bi) + (c + di) = (a + c) + (b + d)i = (a + c, b + d)$$

and similarly

$$-(a, b) = -(a + bi) = (-a) + (-b)i = (-a, -b).$$

As these derivations used the informal definition of the complex numbers, these identities have not yet been rigorously proven. However we shall simply *encode* these identities into our complex number system by defining the notion of addition and negation by the above rules:

Definition 4.6.3 (*Complex addition, negation, and zero*) If $z = (a, b)$ and $w = (c, d)$ are two complex numbers, we define their *sum* $z + w$ to be the complex number $z + w := (a + c, b + d)$. Thus for instance $(2, 4) + (3, -1) = (5, 3)$. We also define the *negation* $-z$ of z to be the complex number $-z := (-a, -b)$, thus for instance $-(3, -1) = (-3, 1)$. We also define the *complex zero* $0_{\mathbf{C}}$ to be the complex number $0_{\mathbf{C}} = (0, 0)$.

It is easy to see that notion of addition is well-defined in the sense that if $z = z'$ and $w = w'$ then $z + w = z' + w'$. Similarly for negation. The complex addition, negation, and zero operations obey the usual laws of arithmetic:

Lemma 4.6.4 (The complex numbers are an additive group) *If z_1, z_2, z_3 are complex numbers, then we have the commutative property $z_1 + z_2 = z_2 + z_1$, the associative property $(z_1 + z_2) + z_3 = z_1 + (z_2 + z_3)$, the identity property $z_1 + 0_{\mathbf{C}} = 0_{\mathbf{C}} + z_1 = z_1$, and the inverse property $z_1 + (-z_1) = (-z_1) + z_1 = 0_{\mathbf{C}}$.*

Proof See Exercise 4.6.1. □

Next, we define the notion of complex multiplication and reciprocal. The informal justification of the complex multiplication rule is

$$\begin{aligned}(a, b) \cdot (c, d) &= (a + bi)(c + di)\\ &= ac + adi + bic + bidi\\ &= (ac - bd) + (ad + bc)i\\ &= (ac - bd, ad + bc)\end{aligned}$$

since i^2 is supposed to equal -1. Thus we define

Definition 4.6.5 (*Complex multiplication*) If $z = (a, b)$ and $w = (c, d)$ are complex numbers, then we define their *product* zw to be the complex number $zw := (ac - bd, ad + bc)$. We also introduce the *complex identity* $1_{\mathbf{C}} := (1, 0)$.

This operation is easily seen to be well-defined, and also obeys the usual laws of arithmetic:

Lemma 4.6.6 *If z_1, z_2, z_3 are complex numbers, then we have the commutative property $z_1 z_2 = z_2 z_1$, the associative property $(z_1 z_2) z_3 = z_1 (z_2 z_3)$, the identity property $z_1 1_{\mathbf{C}} = 1_{\mathbf{C}} z_1 = z_1$, and the distributivity properties $z_1(z_2 + z_3) = z_1 z_2 + z_1 z_3$ and $(z_2 + z_3) z_1 = z_2 z_1 + z_3 z_1$.*

Proof See Exercise 4.6.2. □

The above lemma can also be stated more succinctly, as the assertion that $\mathbf{C}$ is a commutative ring. As is usual, we now write $z - w$ as shorthand for $z + (-w)$.

We now identify the real numbers $\mathbf{R}$ with a subset of the complex numbers $\mathbf{C}$ by identifying any real number x with the complex number $(x, 0)$, thus $x \equiv (x, 0)$. Note that this identification is consistent with equality (thus $x = y$ iff $(x, 0) = (y, 0)$), with addition ($x_1 + x_2 = x_3$ iff $(x_1, 0) + (x_2, 0) = (x_3, 0)$), with negation ($x = -y$ iff $(x, 0) = -(y, 0)$), and multiplication ($x_1 x_2 = x_3$ iff $(x_1, 0)(x_2, 0) = (x_3, 0)$), so we will no longer need to distinguish between "real addition" and "complex addition", and similarly for equality, negation, and multiplication. For instance, we can compute $3(2, 4)$ by identifying the real number 3 with the complex number $(3, 0)$ and then computing $(3, 0)(2, 4) = (3 \times 2 - 0 \times 4, 3 \times 4 + 0 \times 2) = (6, 12)$. Note also that $0 \equiv 0_{\mathbf{C}}$ and $1 \equiv 1_{\mathbf{C}}$, so we can now drop the **C** subscripts from the zero 0 and the identity 1.

We now define i to be the complex number $i := (0, 1)$. We can now reconstruct the informal definition of the complex numbers as a lemma:

Lemma 4.6.7 *Every complex number $z \in \mathbf{C}$ can be written as $z = a + bi$ for exactly one pair a, b of real numbers. Also, we have $i^2 = -1$, and $-z = (-1)z$.*

Proof See Exercise 4.6.3. □

Because of this lemma, we will now refer to complex numbers in the more usual notation $a + bi$ and discard the formal notation (a, b) henceforth.

Definition 4.6.8 (*Real and imaginary parts*) If z is a complex number with the representation $z = a + bi$ for some real numbers a, b, we shall call a the *real part* of z and denote $\Re(z) := a$, and call b the *imaginary part* of z and denote $\Im(z) := b$, thus for instance $\Re(3 + 4i) = 3$ and $\Im(3 + 4i) = 4$, and in general $z = \Re(z) + i\Im(z)$. Note that z is real iff $\Im(z) = 0$. We say that z is *imaginary* iff $\Re(z) = 0$, thus for instance $4i$ is imaginary, while $3 + 4i$ is neither real nor imaginary, and 0 is both real and imaginary. We define the *complex conjugate* $\overline{z}$ of z to be the complex number $\overline{z} := \Re(z) - i\Im(z)$, thus for instance $\overline{3 + 4i} = 3 - 4i$, $\overline{i} = -i$, and $\overline{3} = 3$.

The operation of complex conjugation has several nice properties:

Lemma 4.6.9 (Complex conjugation is an involution) *Let z, w be complex numbers, then $\overline{z + w} = \overline{z} + \overline{w}$, $\overline{-z} = -\overline{z}$, and $\overline{zw} = \overline{z}\,\overline{w}$. Also $\overline{\overline{z}} = z$. Finally, we have $\overline{z} = \overline{w}$ if and only if $z = w$, and $\overline{z} = z$ if and only if z is real.*

Proof See Exercise 4.6.4. □

The notion of absolute value $|x|$ was defined for rational numbers x in Definition 4.3.1, and this definition extends to real numbers in the obvious manner. However, we cannot extend this definition directly to the complex numbers, as most complex numbers are neither positive nor negative. (For instance, we do not classify i as either a positive or negative number; see Exercise 4.6.11 for some reasons why.) However, we can still define absolute value by generalizing the formula $|x| = \sqrt{x^2}$ from Exercise 5.6.4:

Definition 4.6.10 (*Complex absolute value*) If $z = a + bi$ is a complex number, we define the *absolute value* $|z|$ of z to be the real number $|z| := \sqrt{a^2 + b^2} = (a^2 + b^2)^{1/2}$.

From Exercise 5.6.4 we see that this notion of absolute value generalizes the notion of real absolute value. The absolute value has a number of other good properties:

Lemma 4.6.11 (Properties of complex absolute value) *Let z, w be complex numbers. Then $|z|$ is a non-negative real number, and $|z| = 0$ if and only if $z = 0$. Also we have the identity $z\overline{z} = |z|^2$, and so $|z| = \sqrt{z\overline{z}}$. As a consequence we have $|zw| = |z||w|$ and $|\overline{z}| = |z|$. Finally, we have the inequalities*

$$-|z| \leq \Re(z) \leq |z|; \quad -|z| \leq \Im(z) \leq |z|; \quad |z| \leq |\Re(z)| + |\Im(z)|$$

as well as the triangle inequality $|z + w| \leq |z| + |w|$.

Proof See Exercise 4.6.6. □

Using the notion of absolute value, we can define a notion of reciprocal:

Definition 4.6.12 (*Complex reciprocal*) If z is a nonzero complex number, we define the *reciprocal* z^{-1} of z to be the complex number $z^{-1} := |z|^{-2}\overline{z}$ (note that $|z|^{-2}$ is well-defined as a positive real number because $|z|$ is positive real, thanks to Lemma 4.6.11). Thus for instance $(1 + 2i)^{-1} = |1 + 2i|^{-2}(1 - 2i) = (1^2 + 2^2)^{-1}(1 - 2i) = \frac{1}{5} - \frac{2}{5}i$. If z is zero, $z = 0$, we leave the reciprocal 0^{-1} undefined.

From the definition and Lemma 4.6.11, we see that

$$zz^{-1} = z^{-1}z = |z|^{-2}\overline{z}z = |z|^{-2}|z|^2 = 1,$$

and so z^{-1} is indeed the reciprocal of z. We can thus define a notion of quotient z/w for any two complex numbers z, w with $w \neq 0$ in the usual manner by the formula $z/w := zw^{-1}$.

The complex numbers can be given a distance by defining $d(z, w) = |z - w|$.

Lemma 4.6.13 *The complex numbers* **C** *with the distance d form a metric space. If $(z_n)_{n=1}^\infty$ is a sequence of complex numbers, and z is another complex number, then we have* $\lim_{n\to\infty} z_n = z$ *in this metric space if and only if* $\lim_{n\to\infty} \Re(z_n) = \Re(z)$ *and* $\lim_{n\to\infty} \Im(z_n) = \Im(z)$.

Proof See Exercise 4.6.9. □

Observe that with our choice of definitions, the space **C** of complex numbers is identical (as a metric space) to the Euclidean plane $\mathbf{R}^2$, since the complex distance between two complex numbers (a, b), (a', b') is exactly the same as the Euclidean distance $\sqrt{(a - a')^2 + (b - b')^2}$ between these points. Thus, every metric property that $\mathbf{R}^2$ satisfies is also obeyed by **C**; for instance, **C** is complete and connected, but not compact.

We also have the usual limit laws:

Lemma 4.6.14 (Complex limit laws) *Let $(z_n)_{n=1}^\infty$ and $(w_n)_{n=1}^\infty$ be convergent sequences of complex numbers, and let c be a complex number. Then the sequences $(z_n + w_n)_{n=1}^\infty$, $(z_n - w_n)_{n=1}^\infty$, $(cz_n)_{n=1}^\infty$, $(z_n w_n)_{n=1}^\infty$, and $(\overline{z_n})_{n=1}^\infty$ are also convergent, with*

$$\lim_{n\to\infty} z_n + w_n = \lim_{n\to\infty} z_n + \lim_{n\to\infty} w_n$$

$$\lim_{n\to\infty} z_n - w_n = \lim_{n\to\infty} z_n - \lim_{n\to\infty} w_n$$

$$\lim_{n\to\infty} cz_n = c \lim_{n\to\infty} z_n$$

$$\lim_{n\to\infty} z_n w_n = \left(\lim_{n\to\infty} z_n\right)\left(\lim_{n\to\infty} w_n\right)$$

$$\lim_{n\to\infty} \overline{z_n} = \overline{\lim_{n\to\infty} z_n}$$

Also, if the w_n *are all nonzero and* $\lim_{n\to\infty} w_n$ *is also nonzero, then* $(z_n/w_n)_{n=1}^{\infty}$ *is also a convergent sequence, with*

$$\lim_{n\to\infty} z_n/w_n = \left(\lim_{n\to\infty} z_n)/(\lim_{n\to\infty} w_n\right).$$

Proof See Exercise 4.6.10. □

Observe that the real and complex number systems are in fact quite similar; they both obey similar laws of arithmetic, and they have similar structure as metric spaces. Indeed many of the results in this textbook that were proven for real-valued functions are also valid for complex-valued functions, simply by replacing "real" with "complex" in the proofs but otherwise leaving all the other details of the proof unchanged. Alternatively, one can always split a complex-valued function f into real and imaginary parts $\Re(f)$, $\Im(f)$, thus $f = \Re(f) + i\Im(f)$, and then deduce results for the complex-valued function f from the corresponding results for the real-valued functions $\Re(f)$, $\Im(f)$. For instance, the theory of pointwise and uniform convergence from Chapter 3, or the theory of power series from this chapter, extends without any difficulty to complex-valued functions. In particular, we can define the complex exponential function in exactly the same manner as for real numbers:

Definition 4.6.15 (*Complex exponential*) If z is a complex number, we define the function $\exp(z)$ by the formula

$$\exp(z) := \sum_{n=0}^{\infty} \frac{z^n}{n!}.$$

Inspired by Proposition 4.5.4, we shall use $\exp(z)$ and e^z interchangeably. It is also possible to define a^z for complex z and other real numbers $a > 0$, but we will not need to do so in this text.

One can state and prove the ratio test for complex series and use it to show that $\exp(z)$ converges for every z. It turns out that many of the properties from Theorem 4.5.2 still hold: we have that $\exp(z + w) = \exp(z)\exp(w)$, for instance; see Exercise 4.6.12. (The other properties require complex differentiation and complex integration, but these topics are beyond the scope of this text.) Another useful observation is that $\overline{\exp(z)} = \exp(\overline{z})$; this can be seen by conjugating the partial sums $\sum_{n=0}^{N} \frac{z^n}{n!}$ and taking limits as $N \to \infty$.

The complex logarithm turns out to be somewhat more subtle, mainly because exp is no longer invertible, and also because the various power series for the logarithm only have a finite radius of convergence (unlike exp, which has an infinite radius of convergence). This rather delicate issue is beyond the scope of this text and will not be discussed here.

—Exercise—

Exercise 4.6.1 Prove Lemma 4.6.4.

Exercise 4.6.2 Prove Lemma 4.6.6.

Exercise 4.6.3 Prove Lemma 4.6.7.

Exercise 4.6.4 Prove Lemma 4.6.9.

Exercise 4.6.5 If z is a complex number, show that $\Re(z) = \frac{z+\overline{z}}{2}$ and $\Im(z) = \frac{z-\overline{z}}{2i}$.

Exercise 4.6.6 Prove Lemma 4.6.11. (*Hint:* to prove the triangle inequality, first prove that $\Re(z\overline{w}) \leq |z||w|$, and hence (from Exercise 4.6.5) that $z\overline{w} + \overline{z}w \leq 2|z||w|$. Then add $|z|^2 + |w|^2$ to both sides of this inequality.)

Exercise 4.6.7 Show that if z, w are complex numbers with $w \neq 0$, then $|z/w| = |z|/|w|$.

Exercise 4.6.8 Let z, w be nonzero complex numbers. Show that $|z + w| = |z| + |w|$ if and only if there exists a positive real number $c > 0$ such that $z = cw$.

Exercise 4.6.9 Prove Lemma 4.6.13.

Exercise 4.6.10 Prove Lemma 4.6.14. (*Hint:* split z_n and w_n into real and imaginary parts and use the usual limit laws, Lemma 6.1.19, combined with Lemma 4.6.13.)

Exercise 4.6.11 The purpose of this exercise is to explain why we do not try to organize the complex numbers into positive and negative parts. Suppose that there was a notion of a "positive complex number" and a "negative complex number" which obeyed the following reasonable axioms (cf. Proposition 4.2.9):

- (Trichotomy) For every complex number z, exactly one of the following statements is true: z is positive, z is negative, z is zero.
- (Negation) If z is a positive complex number, then $-z$ is negative. If z is a negative complex number, then $-z$ is positive.
- (Additivity) If z and w are positive complex numbers, then $z + w$ is also positive.
- (Multiplicativity) If z and w are positive complex numbers, then zw is also positive.

Show that these four axioms are inconsistent,, i.e., one can use these axioms to deduce a contradiction. (*Hints:* first use the axioms to deduce that 1 is positive, and then conclude that -1 is negative. Then apply the Trichotomy axiom to $z = i$ and obtain a contradiction in any one of the three cases.)

Exercise 4.6.12 Prove the ratio test for complex series, and use it to show that the series used to define the complex exponential is absolutely convergent. Then prove that $\exp(z + w) = \exp(z)\exp(w)$ for all complex numbers z, w.

4.7 Trigonometric Functions

We now discuss the next most important class of special functions, after the exponential and logarithmic functions, namely the trigonometric functions. (There are several other useful special functions in mathematics, such as the hyperbolic trigonometric functions and hypergeometric functions, the gamma and zeta functions, and elliptic functions, but they occur more rarely and will not be discussed here.)

Trigonometric functions are often defined using geometric concepts, notably those of circles, triangles, and angles. However, it is also possible to define them using more analytic concepts and in particular the (complex) exponential function.

Definition 4.7.1 (*Trigonometric functions*) If z is a complex number, then we define

$$\cos(z) := \frac{e^{iz} + e^{-iz}}{2}$$

and

$$\sin(z) := \frac{e^{iz} - e^{-iz}}{2i}.$$

We refer to cos and sin as the *cosine* and *sine* functions, respectively.

These formulae were discovered by Leonhard Euler (1707–1783) in 1748, who recognized the link between the complex exponential and the trigonometric functions. Note that since we have defined the sine and cosine for complex numbers z, we automatically have defined them also for real numbers x. In fact in most applications one is only interested in the trigonometric functions when applied to real numbers.

From the power series definition of exp, we have

$$e^{iz} = 1 + iz - \frac{z^2}{2!} - \frac{iz^3}{3!} + \frac{z^4}{4!} + \ldots$$

and

$$e^{-iz} = 1 - iz - \frac{z^2}{2!} + \frac{iz^3}{3!} + \frac{z^4}{4!} - \ldots$$

and so from the above formulae we have

$$\cos(z) = 1 - \frac{z^2}{2!} + \frac{z^4}{4!} - \ldots = \sum_{n=0}^{\infty} \frac{(-1)^n z^{2n}}{(2n)!}$$

and

$$\sin(z) = z - \frac{z^3}{3!} + \frac{z^5}{5!} - \ldots = \sum_{n=0}^{\infty} \frac{(-1)^n z^{2n+1}}{(2n+1)!}.$$

In particular, $\cos(x)$ and $\sin(x)$ are always real whenever x is real. From the ratio test we see that the two power series $\sum_{n=0}^{\infty} \frac{(-1)^n x^{2n}}{(2n)!}$, $\sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{(2n+1)!}$ are absolutely convergent for every x, thus $\sin(x)$ and $\cos(x)$ are real analytic at 0 with an infinite radius of convergence. From Exercise 4.2.8 we thus see that the sine and cosine functions are real analytic on all of $\mathbf{R}$. (They are also complex analytic on all of $\mathbf{C}$, but we will not pursue this matter in this text.) In particular the sine and cosine functions are continuous and differentiable.

We list some basic properties of the sine and cosine functions below.

Theorem 4.7.2 (Trigonometric identities) *Let x, y be real numbers.*

(a) We have $\sin(x)^2 + \cos(x)^2 = 1$. In particular, we have $\sin(x) \in [-1, 1]$ and $\cos(x) \in [-1, 1]$ for all $x \in \mathbf{R}$.
(b) We have $\sin'(x) = \cos(x)$ and $\cos'(x) = -\sin(x)$.
(c) We have $\sin(-x) = -\sin(x)$ and $\cos(-x) = \cos(x)$.
(d) We have $\cos(x + y) = \cos(x)\cos(y) - \sin(x)\sin(y)$ and $\sin(x + y) = \sin(x)\cos(y) + \cos(x)\sin(y)$.
(e) We have $\sin(0) = 0$ and $\cos(0) = 1$.
(f) We have $e^{ix} = \cos(x) + i\sin(x)$ and $e^{-ix} = \cos(x) - i\sin(x)$. In particular $\cos(x) = \Re(e^{ix})$ and $\sin(x) = \Im(e^{ix})$.

Proof See Exercise 4.7.1. □

Now we describe some other properties of sin and cos.

Lemma 4.7.3 *There exists a positive number x such that $\sin(x)$ is equal to 0.*

Proof Suppose for sake of contradiction that $\sin(x) \neq 0$ for all $x \in (0, \infty)$. Observe that this would also imply that $\cos(x) \neq 0$ for all $x \in (0, \infty)$, since if $\cos(x) = 0$ then $\sin(2x) = 0$ by Theorem 4.7.2(d) (why?). Since $\cos(0) = 1$, this implies by the intermediate value theorem (Theorem 9.7.1) that $\cos(x) > 0$ for all $x > 0$ (why?). Also, since $\sin(0) = 0$ and $\sin'(0) = 1 > 0$, we see that sin increasing near 0, hence is positive to the right of 0. By the intermediate value theorem again we conclude that $\sin(x) > 0$ for all $x > 0$ (otherwise sin would have a zero on $(0, \infty)$).

In particular if we define the cotangent function $\cot(x) := \cos(x)/\sin(x)$, then $\cot(x)$ would be positive and differentiable on all of $(0, \infty)$. From the quotient rule (Theorem 10.1.13(h)) and Theorem 4.7.2 we see that the derivative of $\cot(x)$ is $-1/\sin(x)^2$ (why?). In particular, we have $\cot'(x) \leq -1$ for all $x > 0$. By the fundamental theorem of calculus (Theorem 11.9.1) this implies that $\cot(x + s) \leq \cot(x) - s$ for all $x > 0$ and $s > 0$. But letting $s \to \infty$ we see that this contradicts our assertion that cot is positive on $(0, \infty)$ (why?). □

Let E be the set $E := \{x \in (0, +\infty) : \sin(x) = 0\}$, i.e., E is the set of roots of sin on $(0, +\infty)$. By Lemma 4.7.3, E is non-empty. Since $\sin'(0) > 0$, there exists a $c > 0$ such that $E \subseteq [c, +\infty)$ (see Exercise 4.7.2). Also, since sin is continuous in $[c, +\infty)$, E is closed in $[c, +\infty)$ (why? Use Theorem 2.1.5(d)). Since $[c, +\infty)$ is closed in $\mathbf{R}$, we conclude that E is closed in $\mathbf{R}$. Thus E contains all its adherent points, and thus contains $\inf(E)$. Thus if we make the definition

Definition 4.7.4 We define π to be the number

$$\pi := \inf\{x \in (0, \infty) : \sin(x) = 0\}$$

then we have $\pi \in E \subseteq [c, +\infty)$ (so in particular $\pi > 0$) and $\sin(\pi) = 0$. By definition of π, sin cannot have any zeroes in $(0, \pi)$, and so in particular must be positive on $(0, \pi)$, (cf. the arguments in Lemma 4.7.3 using the intermediate value theorem). Since $\cos'(x) = -\sin(x)$, we thus conclude that $\cos(x)$ is strictly decreasing on $(0, \pi)$. Since $\cos(0) = 1$, this implies in particular that $\cos(\pi) < 1$; since $\sin^2(\pi) + \cos^2(\pi) = 1$ and $\sin(\pi) = 0$, we thus conclude that $\cos(\pi) = -1$.

In particular we have Euler's famous formula

$$e^{\pi i} = \cos(\pi) + i \sin(\pi) = -1.$$

We now conclude with some other properties of sine and cosine.

Theorem 4.7.5 (Periodicity of trigonometric functions) *Let x be a real number.*

(a) We have $\cos(x + \pi) = -\cos(x)$ and $\sin(x + \pi) = -\sin(x)$. In particular we have $\cos(x + 2\pi) = \cos(x)$ and $\sin(x + 2\pi) = \sin(x)$, i.e., sin and cos are periodic with period 2π.
(b) We have $\sin(x) = 0$ if and only if x/π is an integer.
(c) We have $\cos(x) = 0$ if and only if x/π is an integer plus 1/2.

Proof See Exercise 4.7.3. □

We can of course define all the other trigonometric functions: tangent, cotangent, secant, and cosecant, and develop all the familiar identities of trigonometry; some examples of this are given in the exercises.

—Exercise—

Exercise 4.7.1 Prove Theorem 4.7.2. (*Hint:* write everything in terms of exponentials whenever possible.)

Exercise 4.7.2 Let $f : \mathbf{R} \to \mathbf{R}$ be a function which is differentiable at x_0, with $f(x_0) = 0$ and $f'(x_0) \neq 0$. Show that there exists a $c > 0$ such that $f(y)$ is nonzero whenever $0 < |x_0 - y| < c$. Conclude in particular that there exists a $c > 0$ such that $\sin(x) \neq 0$ for all $0 < x < c$.

Exercise 4.7.3 Prove Theorem 4.7.5. (*Hint:* for (c), you may wish to first compute $\sin(\pi/2)$ and $\cos(\pi/2)$, and then link $\cos(x)$ to $\sin(x + \pi/2)$.)

Exercise 4.7.4 Let x, y be real numbers such that $x^2 + y^2 = 1$. Show that there is exactly one real number $\theta \in (-\pi, \pi]$ such that $x = \sin(\theta)$ and $y = \cos(\theta)$. (*Hint:* you may need to divide into cases depending on whether x, y are positive, negative, or zero.)

Exercise 4.7.5 Show that if $r, s > 0$ are positive real numbers, and θ, α are real numbers such that $re^{i\theta} = se^{i\alpha}$, then $r = s$ and $\theta = \alpha + 2\pi k$ for some integer k.

Exercise 4.7.6 Let z be a nonzero complex number. Using Exercise 4.7.4, show that there is exactly one pair of real numbers r, θ such that $r > 0, \theta \in (-\pi, \pi]$, and $z = re^{i\theta}$. (This is sometimes known as the *standard polar representation* of z.)

Exercise 4.7.7 For any real number θ and integer n, prove the *de Moivre identities*

$$\cos(n\theta) = \Re((\cos\theta + i\sin\theta)^n); \quad \sin(n\theta) = \Im((\cos\theta + i\sin\theta)^n).$$

Exercise 4.7.8 Let $\tan\colon (-\pi/2, \pi/2) \to \mathbf{R}$ be the tangent function $\tan(x) := \sin(x)/\cos(x)$. Show that tan is differentiable and monotone increasing, with $\frac{d}{dx}\tan(x) = 1 + \tan(x)^2$, and that $\lim_{x\to\pi/2}\tan(x) = +\infty$ and $\lim_{x\to-\pi/2}\tan(x) = -\infty$. Conclude that tan is in fact a bijection from $(-\pi/2, \pi/2) \to \mathbf{R}$, and thus has an inverse function $\tan^{-1}\colon \mathbf{R} \to (-\pi/2, \pi/2)$ (this function is called the *arctangent function*). Show that $\tan^{-1}$ is differentiable and $\frac{d}{dx}\tan^{-1}(x) = \frac{1}{1+x^2}$.

Exercise 4.7.9 Recall the arctangent function $\tan^{-1}$ from Exercise 4.7.8. By modifying the proof of Theorem 4.5.6(e), establish the identity

$$\tan^{-1}(x) = \sum_{n=0}^{\infty} \frac{(-1)^n x^{2n+1}}{2n+1}$$

for all $x \in (-1, 1)$. Using Abel's theorem (Theorem 4.3.1) to extend this identity to the case $x = 1$, conclude in particular the identity

$$\pi = 4 - \frac{4}{3} + \frac{4}{5} - \frac{4}{7} + \ldots = 4\sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1}.$$

(Note that the series converges by the alternating series test, Proposition 7.2.11.) Conclude in particular that $4 - \frac{4}{3} < \pi < 4$. (One can of course compute $\pi = 3.1415926\ldots$ to much higher accuracy, though if one wishes to do so it is advisable to use a different formula than the one above, which converges very slowly.)

Exercise 4.7.10 Let $f\colon \mathbf{R} \to \mathbf{R}$ be the function

$$f(x) := \sum_{n=1}^{\infty} 4^{-n}\cos(32^n \pi x).$$

(a) Show that this series is uniformly convergent, and that f is continuous.
(b) Show that for every integer j and every integer $m \geq 1$, we have

$$|f\left(\frac{j+1}{32^m}\right) - f\left(\frac{j}{32^m}\right)| \geq 4^{-m}.$$

(*Hint:* use the identity

$$\sum_{n=1}^{\infty} a_n = \left(\sum_{n=1}^{m-1} a_n\right) + a_m + \sum_{n=m+1}^{\infty} a_n$$

for certain sequences a_n. Also, use the fact that the cosine function is periodic with period 2π, as well as the geometric series formula $\sum_{n=0}^{\infty} r^n = \frac{1}{1-r}$ for any $|r| < 1$. Finally, you will need the inequality $|\cos(x) - \cos(y)| \leq |x - y|$ for any real numbers x and y; this can be proven by using the mean value theorem (Corollary 10.2.9), or the fundamental theorem of calculus (Theorem 11.9.4).)

(c) Using (b), show that for every real number x_0, the function f is not differentiable at x_0. (*Hint:* for every x_0 and every $m \geq 1$, there exists an integer j such that $j \leq 32^m x_0 \leq j + 1$, thanks to Exercise 5.4.3.)

(d) Explain briefly why the result in (c) does not contradict Corollary 3.7.3.

Chapter 5
Fourier Series

In the previous two chapters, we discussed the issue of how certain functions (for instance, compactly supported continuous functions) could be approximated by polynomials. Later, we showed how a different class of functions (real analytic functions) could be written exactly (not approximately) as an infinite polynomial, or more precisely a power series.

Power series are already immensely useful, especially when dealing with special functions such as the exponential and trigonometric functions discussed earlier. However, there are some circumstances where power series are not so useful, because one has to deal with functions (e.g., $\sqrt{x}$) which are not real analytic, and so do not have power series.

Fortunately, there is another type of series expansion, known as *Fourier series*, which is also a very powerful tool in analysis (though used for slightly different purposes). Instead of analyzing compactly supported functions, it instead analyzes *periodic functions*; instead of decomposing into polynomials, it decomposes into *trigonometric polynomials*. Roughly speaking, the theory of Fourier series asserts that just about every periodic function can be decomposed as an (infinite) sum of sines and cosines.

Remark 5.0.1 Jean-Baptiste Fourier (1768–1830) was, among other things, an administrator accompanying Napoleon on his invasion of Egypt, and then a Prefect in France during Napoleon's reign. After the Napoleonic wars, he returned to mathematics. He introduced Fourier series in an important 1807 paper in which he used them to solve what is now known as the heat equation. At the time, the claim that every periodic function could be expressed as a sum of sines and cosines was extremely controversial, even such leading mathematicians as Euler declared that it was impossible. Nevertheless, Fourier managed to show that this was indeed the case, although the proof was not completely rigorous and was not totally accepted for almost another hundred years.

T. Tao, *Analysis II*, Texts and Readings in Mathematics,
https://doi.org/10.1007/978-981-19-7284-3_5

There will be some similarities between the theory of Fourier series and that of power series, but there are also some major differences. For instance, the convergence of Fourier series is usually not uniform (i.e., not in the L^∞ metric), but instead we have convergence in a different metric, the L^2-metric. Also, we will need to use complex numbers heavily in our theory, while they played only a tangential rôle in power series.

The theory of Fourier series (and of related topics such as Fourier integrals and the Laplace transform) is vast, and deserves an entire course in itself. It has many, many applications, most directly to differential equations, signal processing, electrical engineering, physics, and analysis, but also to algebra and number theory. We will only give the barest bones of the theory here, however, and almost no applications.

5.1 Periodic Functions

The theory of Fourier series has to do with the analysis of *periodic functions*, which we now define. It turns out to be convenient to work with complex-valued functions rather than real-valued ones.

Definition 5.1.1 Let $L > 0$ be a real number. A function $f : \mathbf{R} \to \mathbf{C}$ is *periodic with period L*, or *L-periodic*, if we have $f(x + L) = f(x)$ for every real number x.

Example 5.1.2 The real-valued functions $f(x) = \sin(x)$ and $f(x) = \cos(x)$ are 2π-periodic, as is the complex-valued function $f(x) = e^{ix}$. These functions are also 4π-periodic, 6π-periodic, etc. (why?). The function $f(x) = x$, however, is not periodic. The constant function $f(x) = 1$ is L-periodic for every L.

Remark 5.1.3 If a function f is L-periodic, then we have $f(x + kL) = f(x)$ for every integer k (why? Use induction for the positive k, and then use a substitution to convert the positive k result to a negative k result. The $k = 0$ case is of course trivial). In particular, if a function f is 1-periodic, then we have $f(x + k) = f(x)$ for every $k \in \mathbf{Z}$. Because of this, 1-periodic functions are sometimes also called **Z**-*periodic* (and L-periodic functions called $L\mathbf{Z}$-periodic).

Example 5.1.4 For any integer n, the functions $\cos(2\pi nx)$, $\sin(2\pi nx)$, and $e^{2\pi inx}$ are all **Z**-periodic. (What happens when n is not an integer?) Another example of a **Z**-periodic function is the function $f : \mathbf{R} \to \mathbf{C}$ defined by $f(x) := 1$ when $x \in [n, n + \frac{1}{2})$ for some integer n, and $f(x) := 0$ when $x \in [n + \frac{1}{2}, n + 1)$ for some integer n. This function is an example of a *square wave*.

Henceforth, for simplicity, we shall only deal with functions which are **Z**-periodic (for the Fourier theory of L-periodic functions, see Exercise 5.5.6). Note that in order to completely specify a **Z**-periodic function $f : \mathbf{R} \to \mathbf{C}$, one only needs to specify its values on the interval $[0, 1)$, since this will determine the values of f everywhere else. This is because every real number x can be written in the form $x = k + y$

where k is an integer (called the *integer part* of x, and sometimes denoted $[x]$) and $y \in [0, 1)$ (this is called the *fractional part* of x, and sometimes denoted $\{x\}$); see Exercise 5.1.1. Because of this, sometimes when we wish to describe a **Z**-periodic function f we just describe what it does on the interval $[0, 1)$, and then say that it is *extended periodically* to all of **R**. This means that we define $f(x)$ for any real number x by setting $f(x) := f(y)$, where we have decomposed $x = k + y$ as discussed above. (One can in fact replace the interval $[0, 1)$ by any other half-open interval of length 1, but we will not do so here.)

The space of complex-valued continuous **Z**-periodic functions is denoted $C(\mathbf{R}/\mathbf{Z}; \mathbf{C})$. (The notation $\mathbf{R}/\mathbf{Z}$ comes from algebra, and denotes the quotient group of the additive group **R** by the additive group **Z**; more information in this can be found in any algebra text.) By "continuous" we mean continuous at all points on **R**; merely being continuous on an interval such as $[0, 1]$ will not suffice, as there may be a discontinuity between the left and right limits at 1 (or at any other integer). Thus for instance, the functions $\sin(2\pi nx)$, $\cos(2\pi nx)$, and $e^{2\pi inx}$ are all elements of $C(\mathbf{R}/\mathbf{Z}; \mathbf{C})$, as are the constant functions, however the square wave function described earlier is not in $C(\mathbf{R}/\mathbf{Z}; \mathbf{C})$ because it is not continuous. Also the function $\sin(x)$ would also not qualify to be in $C(\mathbf{R}/\mathbf{Z}; \mathbf{C})$ since it is not **Z**-periodic.

Lemma 5.1.5 (Basic properties of $C(\mathbf{R}/\mathbf{Z}; \mathbf{C})$)

(a) (Boundedness) If $f \in C(\mathbf{R}/\mathbf{Z}; \mathbf{C})$, then f is bounded (i.e., there exists a real number $M > 0$ such that $|f(x)| \le M$ for all $x \in \mathbf{R}$).

(b) (Vector space and algebra properties) If $f, g \in C(\mathbf{R}/\mathbf{Z}; \mathbf{C})$, then the functions $f + g$, $f - g$, and fg are also in $C(\mathbf{R}/\mathbf{Z}; \mathbf{C})$. Also, if c is any complex number, then the function cf is also in $C(\mathbf{R}/\mathbf{Z}; \mathbf{C})$.

(c) (Closure under uniform limits) If $(f_n)_{n=1}^{\infty}$ is a sequence of functions in $C(\mathbf{R}/\mathbf{Z}; \mathbf{C})$ which converges uniformly to another function $f : \mathbf{R} \to \mathbf{C}$, then f is also in $C(\mathbf{R}/\mathbf{Z}; \mathbf{C})$.

Proof See Exercise 5.1.2. □

One can make $C(\mathbf{R}/\mathbf{Z}; \mathbf{C})$ into a metric space by re-introducing the now familiar sup norm metric

$$d_\infty(f, g) = \sup_{x \in \mathbf{R}} |f(x) - g(x)| = \sup_{x \in [0,1)} |f(x) - g(x)|$$

of uniform convergence. (Why is the first supremum the same as the second?) See Exercise 5.1.3.

— Exercise —

Exercise 5.1.1 Show that every real number x can be written in exactly one way in the form $x = k + y$, where k is an integer and $y \in [0, 1)$. (Hint: to prove existence of such a representation, set $k := \sup\{l \in \mathbf{Z} : l \le x\}$.)

Exercise 5.1.2 Prove Lemma 5.1.5. (Hint: for (a), first show that f is bounded on $[0, 1]$.)

Exercise 5.1.3 Show that $C(\mathbf{R}/\mathbf{Z}; \mathbf{C})$ with the sup norm metric d_∞ is a metric space. Furthermore, show that this metric space is complete.

5.2 Inner Products on Periodic Functions

From Lemma 5.1.5 we know that we can add, subtract, multiply, and take limits of continuous periodic functions. We will need a couple more operations on the space $C(\mathbf{R}/\mathbf{Z}; \mathbf{C})$, though. The first one is that of *inner product*.

Definition 5.2.1 (*Inner product*) If $f, g \in C(\mathbf{R}/\mathbf{Z}; \mathbf{C})$, we define the *inner product* $\langle f, g\rangle$ to be the quantity

$$\langle f, g\rangle = \int_{[0,1]} f(x)\overline{g(x)}\,\mathrm{d}x.$$

Remark 5.2.2 In order to integrate a complex-valued function, $f(x) = g(x) + ih(x)$, we use the definition that $\int_{[a,b]} f := \int_{[a,b]} g + i\int_{[a,b]} h$; i.e., we integrate the real and imaginary parts of the function separately. Thus for instance $\int_{[1,2]}(1 + ix)\,\mathrm{d}x = \int_{[1,2]} 1\,\mathrm{d}x + i\int_{[1,2]} x\,\mathrm{d}x = 1 + \frac{3}{2}i$. It is easy to verify that all the standard rules of calculus (integration by parts, fundamental theorem of calculus, substitution, etc.) still hold when the functions are complex-valued instead of real-valued.

Example 5.2.3 Let f be the constant function $f(x):=1$, and let $g(x)$ be the function $g(x):=e^{2\pi ix}$. Then we have

$$\begin{aligned}\langle f, g\rangle &= \int_{[0,1]} 1\overline{e^{2\pi ix}}\,\mathrm{d}x\\ &= \int_{[0,1]} e^{-2\pi ix}\,\mathrm{d}x\\ &= \frac{e^{-2\pi ix}}{-2\pi i}\Big|_{x=0}^{x=1}\\ &= \frac{e^{-2\pi i} - e^0}{-2\pi i}\\ &= \frac{1-1}{-2\pi i}\\ &= 0.\end{aligned}$$

Remark 5.2.4 In general, the inner product $\langle f, g\rangle$ will be a complex number. (Note that $f(x)\overline{g(x)}$ will be Riemann integrable since both functions are bounded and continuous.)

Roughly speaking, the inner product $\langle f, g\rangle$ is to the space $C(\mathbf{R}/\mathbf{Z}; \mathbf{C})$ what the dot product $x \cdot y$ is to Euclidean spaces such as $\mathbf{R}^n$. We list some basic properties of the inner product below; a more in-depth study of inner products on vector spaces can be found in any linear algebra text but is beyond the scope of this text.

Lemma 5.2.5 *Let $f, g, h \in C(\mathbf{R}/\mathbf{Z}; \mathbf{C})$.*

(a) (Hermitian property) We have $\langle g, f\rangle = \overline{\langle f, g\rangle}$.
(b) (Positivity) We have $\langle f, f\rangle \geq 0$. Furthermore, we have $\langle f, f\rangle = 0$ if and only if $f = 0$ (i.e., $f(x) = 0$ for all $x \in \mathbf{R}$).
(c) (Linearity in the first variable) We have $\langle f + g, h\rangle = \langle f, h\rangle + \langle g, h\rangle$. For any complex number c, we have $\langle cf, g\rangle = c\langle f, g\rangle$.
(d) (Antilinearity in the second variable) We have $\langle f, g + h\rangle = \langle f, g\rangle + \langle f, h\rangle$. For any complex number c, we have $\langle f, cg\rangle = \overline{c}\langle f, g\rangle$.

Proof See Exercise 5.2.1. □

From the positivity property, it makes sense to define the L^2 *norm* $\|f\|_2$ of a function $f \in C(\mathbf{R}/\mathbf{Z}; \mathbf{C})$ by the formula

$$\|f\|_2 := \sqrt{\langle f, f\rangle} = \left(\int_{[0,1]} f(x)\overline{f(x)}\, dx\right)^{1/2} = \left(\int_{[0,1]} |f(x)|^2\, dx\right)^{1/2}.$$

Thus $\|f\|_2 \geq 0$ for all f. The norm $\|f\|_2$ is sometimes called the *root mean square* of f.

Example 5.2.6 If $f(x)$ is the function $e^{2\pi ix}$, then

$$\|f\|_2 = \left(\int_{[0,1]} e^{2\pi ix}e^{-2\pi ix}\, dx\right)^{1/2} = \left(\int_{[0,1]} 1\, dx\right)^{1/2} = 1^{1/2} = 1.$$

This L^2 norm is related to, but is distinct from, the L^∞ norm $\|f\|_\infty := \sup_{x\in\mathbf{R}} |f(x)|$. For instance, if $f(x) = \sin(2\pi x)$, then $\|f\|_\infty = 1$ but $\|f\|_2 = \frac{1}{\sqrt{2}}$. In general, the best one can say is that $0 \leq \|f\|_2 \leq \|f\|_\infty$; see Exercise 5.2.3.

Some basic properties of the L^2 norm are given below.

Lemma 5.2.7 *Let $f, g \in C(\mathbf{R}/\mathbf{Z}; \mathbf{C})$.*

(a) (Non-degeneracy) We have $\|f\|_2 = 0$ if and only if $f = 0$.
(b) (Cauchy–Schwarz inequality) We have $|\langle f, g\rangle| \leq \|f\|_2\|g\|_2$.
(c) (Triangle inequality) We have $\|f + g\|_2 \leq \|f\|_2 + \|g\|_2$.

(d) (Pythagoras' theorem) If $\langle f, g\rangle = 0$*, then* $\|f+g\|_2^2 = \|f\|_2^2 + \|g\|_2^2$.
(e) (Homogeneity) We have $\|cf\|_2 = |c|\|f\|_2$ *for all* $c \in \mathbf{C}$.

Proof See Exercise 5.2.2. □

In light of Pythagoras' theorem, we sometimes say that f and g are *orthogonal* iff $\langle f, g\rangle = 0$.

We can now define the L^2 *metric* d_{L^2} on $C(\mathbf{R}/\mathbf{Z}; \mathbf{C})$ by defining

$$d_{L^2}(f, g) := \|f-g\|_2 = \left(\int_{[0,1]} |f(x) - g(x)|^2 \, dx\right)^{1/2}.$$

Remark 5.2.8 One can verify that d_{L^2} is indeed a metric (Exercise 5.2.4). Indeed, the L^2 metric is very similar to the l^2 metric on Euclidean spaces $\mathbf{R}^n$, which is why the notation is deliberately chosen to be similar; you should compare the two metrics yourself to see the analogy.

Note that a sequence f_n of functions in $C(\mathbf{R}/\mathbf{Z}; \mathbf{C})$ will *converge in the* L^2 *metric* to $f \in C(\mathbf{R}/\mathbf{Z}; \mathbf{C})$ if $d_{L^2}(f_n, f) \to 0$ as $n \to \infty$, or in other words that

$$\lim_{n\to\infty} \int_{[0,1]} |f_n(x) - f(x)|^2 \, dx = 0.$$

Remark 5.2.9 The notion of convergence in L^2 metric is different from that of uniform or pointwise convergence; see Exercise 5.2.6.

Remark 5.2.10 The L^2 metric is not as well-behaved as the L^∞ metric. For instance, it turns out the space $C(\mathbf{R}/\mathbf{Z}; \mathbf{C})$ is not complete in the L^2 metric, despite being complete in the L^∞ metric; see Exercise 5.2.5.

— Exercise —

Exercise 5.2.1 Prove Lemma 5.2.5. (Hint: the last part of (b) is a little tricky. You may need to prove by contradiction, assuming that f is not the zero function, and then show that $\int_{[0,1]} |f(x)|^2$ is strictly positive. You will need to use the fact that f, and hence $|f|$, is continuous, to do this.)

Exercise 5.2.2 Prove Lemma 5.2.7. (Hint: use Lemma 5.2.5 frequently. For the Cauchy–Schwarz inequality, begin with the positivity property $\langle f, f\rangle \geq 0$, but with f replaced by the function $f\|g\|_2^2 - \langle f, g\rangle g$, and then simplify using Lemma 5.2.5. You may have to treat the case $\|g\|_2 = 0$ separately. Use the Cauchy–Schwarz inequality to prove the triangle inequality.)

Exercise 5.2.3 If $f \in C(\mathbf{R}/\mathbf{Z}; \mathbf{C})$ is a non-zero function, show that $0 < \|f\|_2 \leq \|f\|_{L^\infty}$. Conversely, if $0 < A \leq B$ are real numbers, show that there exists a non-zero function $f \in C(\mathbf{R}/\mathbf{Z}; \mathbf{C})$ such that $\|f\|_2 = A$ and $\|f\|_\infty = B$. (Hint: let g

be a non-constant non-negative real-valued function in $C(\mathbf{R}/\mathbf{Z}; \mathbf{C})$, and consider functions f of the form $f = (c + dg)^{1/2}$ for some constant real numbers $c, d > 0$.)

Exercise 5.2.4 Prove that the L^2 metric d_{L^2} on $C(\mathbf{R}/\mathbf{Z}; \mathbf{C})$ does indeed turn $C(\mathbf{R}/\mathbf{Z}; \mathbf{C})$ into a metric space. (cf. Exercise 1.1.6).

Exercise 5.2.5 Find a sequence of continuous periodic functions which converge in L^2 to a discontinuous periodic function. (Hint: try converging to the square wave function.)

Exercise 5.2.6 Let $f \in C(\mathbf{R}/\mathbf{Z}, \mathbf{C})$, and let $(f_n)_{n=1}^{\infty}$ be a sequence of functions in $C(\mathbf{R}/\mathbf{Z}; \mathbf{C})$.

(a) Show that if f_n converges uniformly to f, then f_n also converges to f in the L^2 metric.
(b) Give an example where f_n converges to f in the L^2 metric, but does not converge to f uniformly. (Hint: take $f = 0$. Try to make the functions f_n large in sup norm.)
(c) Give an example where f_n converges to f in the L^2 metric, but does not converge to f pointwise. (Hint: take $f = 0$. Try to make the functions f_n large at one point.)
(d) Give an example where f_n converges to f pointwise, but does not converge to f in the L^2 metric. (Hint: take $f = 0$. Try to make the functions f_n large in L^2 norm.)

5.3 Trigonometric Polynomials

We now define the concept of a *trigonometric polynomial*. Just as polynomials are combinations of the functions x^n (sometimes called *monomials*), trigonometric polynomials are combinations of the functions $e^{2\pi inx}$ (sometimes called *characters*).

Definition 5.3.1 (*Characters*) For every integer n, we let $e_n \in C(\mathbf{R}/\mathbf{Z}; \mathbf{C})$ denote the function

$$e_n(x) := e^{2\pi inx}.$$

This is sometimes referred to as the *character with frequency n*.

Definition 5.3.2 (*Trigonometric polynomials*) A function f in $C(\mathbf{R}/\mathbf{Z}; \mathbf{C})$ is said to be a *trigonometric polynomial* if we can write $f = \sum_{n=-N}^{N} c_n e_n$ for some integer $N \geq 0$ and some complex numbers $(c_n)_{n=-N}^{N}$.

Example 5.3.3 The function $f = 4e_{-2} + ie_{-1} - 2e_0 + 0e_1 - 3e_2$ is a trigonometric polynomial; it can be written more explicitly as

$$f(x) = 4e^{-4\pi ix} + ie^{-2\pi ix} - 2 - 3e^{4\pi ix}.$$

Example 5.3.4 For any integer n, the function $\cos(2\pi nx)$ is a trigonometric polynomial, since

$$\cos(2\pi nx) = \frac{e^{2\pi inx} + e^{-2\pi inx}}{2} = \frac{1}{2}e_{-n} + \frac{1}{2}e_n.$$

Similarly the function $\sin(2\pi nx) = \frac{-1}{2i}e_{-n} + \frac{1}{2i}e_n$ is a trigonometric polynomial. In fact, any linear combination of sines and cosines is also a trigonometric polynomial, for instance $3 + i\cos(2\pi x) + 4i\sin(4\pi x)$ is a trigonometric polynomial.

The Fourier theorem will allow us to write any function in $C(\mathbf{R}/\mathbf{Z};\mathbf{C})$ as a Fourier series, which is to trigonometric polynomials what power series is to polynomials. To do this we will use the inner product structure from the previous section. The key computation is

Lemma 5.3.5 (Characters are an orthonormal system) *For any integers n and m, we have $\langle e_n, e_m\rangle = 1$ when $n = m$ and $\langle e_n, e_m\rangle = 0$ when $n \neq m$. Also, we have $\|e_n\| = 1$.*

Proof See Exercise 5.3.2. □

As a consequence, we have a formula for the coefficients of a trigonometric polynomial.

Corollary 5.3.6 *Let $f = \sum_{n=-N}^{N} c_n e_n$ be a trigonometric polynomial. Then we have the formula*

$$c_n = \langle f, e_n\rangle$$

for all integers $-N \leq n \leq N$. Also, we have $0 = \langle f, e_n\rangle$ whenever $n > N$ or $n < -N$. Also, we have the identity

$$\|f\|_2^2 = \sum_{n=-N}^{N} |c_n|^2.$$

Proof See Exercise 5.3.3. □

We rewrite the conclusion of this corollary in a different way.

Definition 5.3.7 (*Fourier transform*) For any function $f \in C(\mathbf{R}/\mathbf{Z};\mathbf{R})$, and any integer $n \in \mathbf{Z}$, we define the n^{th} *Fourier coefficient of* f, denoted $\hat{f}(n)$, by the formula

$$\hat{f}(n) := \langle f, e_n\rangle = \int_{[0,1]} f(x)e^{-2\pi inx}\,dx.$$

The function $\hat{f}: \mathbf{Z} \to \mathbf{C}$ is called the *Fourier transform* of f.

From Corollary 5.3.6, we see that whenever $f = \sum_{n=-N}^{N} c_n e_n$ is a trigonometric polynomial, we have

$$f = \sum_{n=-N}^{N} \langle f, e_n \rangle e_n = \sum_{n=-\infty}^{\infty} \langle f, e_n \rangle e_n$$

and in particular we have the *Fourier inversion formula*

$$f = \sum_{n=-\infty}^{\infty} \hat{f}(n) e_n$$

or in other words

$$f(x) = \sum_{n=-\infty}^{\infty} \hat{f}(n) e^{2\pi i n x}.$$

The right-hand side is referred to as the *Fourier series* of f. Also, from the second identity of Corollary 5.3.6 we have the *Plancherel formula*

$$\|f\|_2^2 = \sum_{n=-\infty}^{\infty} |\hat{f}(n)|^2.$$

Remark 5.3.8 We stress that at present we have only proven the Fourier inversion and Plancherel formulae in the case when f is a trigonometric polynomial. Note that in this case that the Fourier coefficients $\hat{f}(n)$ are mostly zero (indeed, they can only be non-zero when $-N \leq n \leq N$), and so this infinite sum is really just a finite sum in disguise. In particular there are no issues about what sense the above series converge in; they both converge pointwise, uniformly, and in L^2 metric, since they are just finite sums.

In the next few sections we will extend the Fourier inversion and Plancherel formulae to general functions in $C(\mathbf{R}/\mathbf{Z}; \mathbf{C})$, not just trigonometric polynomials. (It is also possible to extend the formula to discontinuous functions such as the square wave, but we will not do so here.) To do this we will need a version of the Weierstrass approximation theorem, this time requiring that a continuous periodic function be approximated uniformly by *trigonometric* polynomials. Just as convolutions were used in the proof of the polynomial Weierstrass approximation theorem, we will also need a notion of convolution tailored for periodic functions.

— Exercise —

Exercise 5.3.1 Show that the sum or product of any two trigonometric polynomials is again a trigonometric polynomial.

Exercise 5.3.2 Prove Lemma 5.3.5.

Exercise 5.3.3 Prove Corollary 5.3.6. (Hint: use Lemma 5.3.5. For the second identity, either use Pythagoras' theorem and induction, or substitute $f = \sum_{n=-N}^{N} c_n e_n$ and expand everything out.)

5.4 Periodic Convolutions

The goal of this section is to prove the Weierstrass approximation theorem for trigonometric polynomials:

Theorem 5.4.1 *Let $f \in C(\mathbf{R}/\mathbf{Z}; \mathbf{C})$, and let $\varepsilon > 0$. Then there exists a trigonometric polynomial P such that $\|f - P\|_\infty \le \varepsilon$.*

This theorem asserts that any continuous periodic function can be uniformly approximated by trigonometric polynomials. To put it another way, if we let $P(\mathbf{R}/\mathbf{Z}; \mathbf{C})$ denote the space of all trigonometric polynomials, then the closure of $P(\mathbf{R}/\mathbf{Z}; \mathbf{C})$ in the L^∞ metric is $C(\mathbf{R}/\mathbf{Z}; \mathbf{C})$.

It is possible to prove this theorem directly from the Weierstrass approximation theorem for polynomials (Theorem 3.8.3), and both theorems are a special case of a much more general theorem known as the *Stone-Weierstrass theorem*, which we will not discuss here. However we shall instead prove this theorem from scratch, in order to introduce a couple of interesting notions, notably that of periodic convolution. The proof here, though, should strongly remind you of the arguments used to prove Theorem 3.8.3.

Definition 5.4.2 (*Periodic convolution*) Let $f, g \in C(\mathbf{R}/\mathbf{Z}; \mathbf{C})$. Then we define the *periodic convolution* $f * g : \mathbf{R} \to \mathbf{C}$ of f and g by the formula

$$f * g(x) := \int_{[0,1]} f(y) g(x - y)\, dy.$$

Remark 5.4.3 Note that this formula is slightly different from the convolution for compactly supported functions defined in Definition 3.8.9, because we are only integrating over $[0, 1]$ and not on all of $\mathbf{R}$. Thus, in principle we have given the symbol $f * g$ two conflicting meanings. However, in practice there will be no confusion, because it is not possible for a non-zero function to both be periodic and compactly supported (Exercise 5.4.1).

Lemma 5.4.4 (Basic properties of periodic convolution) *Let $f, g, h \in C(\mathbf{R}/\mathbf{Z}; \mathbf{C})$.*

(a) (Closure) The convolution $f * g$ is continuous and $\mathbf{Z}$-periodic. In other words, $f * g \in C(\mathbf{R}/\mathbf{Z}; \mathbf{C})$.
(b) (Commutativity) We have $f * g = g * f$.
(c) (Bilinearity) We have $f * (g + h) = f * g + f * h$ and $(f + g) * h = f * h + g * h$. For any complex number c, we have $c(f * g) = (cf) * g = f * (cg)$.

Proof See Exercise 5.4.2. □

Now we observe an interesting identity: for any $f \in C(\mathbf{R}/\mathbf{Z}; \mathbf{C})$ and any integer n, we have

$$f * e_n = \hat{f}(n)e_n.$$

To prove this, we compute

$$f * e_n(x) = \int_{[0,1]} f(y)e^{2\pi in(x-y)}\, \mathrm{d}y$$

$$= e^{2\pi inx} \int_{[0,1]} f(y)e^{-2\pi iny}\, \mathrm{d}y = \hat{f}(n)e^{2\pi inx} = \hat{f}(n)e_n$$

as desired.

More generally, we see from Lemma 5.4.4(iii) that for any trigonometric polynomial $P = \sum_{n=-N}^{n=N} c_n e_n$, we have

$$f * P = \sum_{n=-N}^{n=N} c_n(f * e_n) = \sum_{n=-N}^{n=N} \hat{f}(n)c_n e_n.$$

Thus the periodic convolution of any function in $C(\mathbf{R}/\mathbf{Z}; \mathbf{C})$ with a trigonometric polynomial, is again a trigonometric polynomial. (Compare with Lemma 3.8.13.)

Next, we introduce the periodic analogue of an approximation to the identity.

Definition 5.4.5 (*Periodic approximation to the identity*) Let $\varepsilon > 0$ and $0 < \delta < 1/2$. A function $f \in C(\mathbf{R}/\mathbf{Z}; \mathbf{C})$ is said to be a *periodic* (ε, δ) *approximation to the identity* if the following properties are true:

(a) $f(x) \geq 0$ for all $x \in \mathbf{R}$, and $\int_{[0,1]} f = 1$.
(b) We have $f(x) < \varepsilon$ for all $\delta \leq |x| \leq 1 - \delta$.

Now we have an analogue of Lemma 3.8.8:

Lemma 5.4.6 *For every* $\varepsilon > 0$ *and* $0 < \delta < 1/2$*, there exists a trigonometric polynomial* P *which is an* (ε, δ) *approximation to the identity.*

Proof We sketch the proof of this Lemma here, and leave the completion of it to Exercise 5.4.3. Let $N \geq 1$ be an integer. We define the *Fejér kernel* F_N to be the function

$$F_N = \sum_{n=-N}^{N} \left(1 - \frac{|n|}{N}\right) e_n.$$

Clearly F_N is a trigonometric polynomial. We observe the identity

$$F_N = \frac{1}{N}\left|\sum_{n=0}^{N-1} e_n\right|^2$$

(why?). But from the geometric series formula (Lemma 7.3.3) we have

$$\sum_{n=0}^{N-1} e_n(x) = \frac{e_N - e_0}{e_1 - e_0} = \frac{e^{\pi i(N-1)x}\sin(\pi N x)}{\sin(\pi x)}$$

when x is not an integer, (why?) and hence we have the formula

$$F_N(x) = \frac{\sin(\pi N x)^2}{N\sin(\pi x)^2}.$$

When x is an integer, the geometric series formula does not apply, but one has $F_N(x) = N$ in that case, as one can see by direct computation. In either case we see that $F_N(x) \geq 0$ for any x. Also, we have

$$\int_{[0,1]} F_N(x)\,\mathrm{d}x = \sum_{n=-N}^{N}\left(1 - \frac{|n|}{N}\right)\int_{[0,1]} e_n = \left(1 - \frac{|0|}{N}\right)1 = 1$$

(why?). Finally, since $\sin(\pi N x) \leq 1$, we have

$$F_N(x) \leq \frac{1}{N\sin(\pi x)^2} \leq \frac{1}{N\sin(\pi\delta)^2}$$

whenever $\delta < |x| < 1-\delta$ (this is because sin is increasing on $[0, \pi/2]$ and decreasing on $[\pi/2, \pi]$). Thus by choosing N large enough, we can make $F_N(x) \leq \varepsilon$ for all $\delta < |x| < 1-\delta$. □

Proof of Theorem 5.4.1 Let f be any element of $C(\mathbf{R}/\mathbf{Z}; \mathbf{C})$; we know that f is bounded, so that we have some $M > 0$ such that $|f(x)| \leq M$ for all $x \in \mathbf{R}$.

Let $\varepsilon > 0$ be arbitrary. Since f is uniformly continuous, there exists a $\delta > 0$ such that $|f(x) - f(y)| \leq \varepsilon$ whenever $|x-y| \leq \delta$. Now use Lemma 5.4.6 to find a trigonometric polynomial P which is a (ε, δ) approximation to the identity. Then $f * P$ is also a trigonometric polynomial. We now estimate $\|f - f * P\|_\infty$.

Let x be any real number. We have

$$\begin{aligned}
|f(x) - f * P(x)| &= |f(x) - P * f(x)| \\
&= \left| f(x) - \int_{[0,1]} f(x-y)P(y)\,\mathrm{d}y \right| \\
&= \left| \int_{[0,1]} f(x)P(y)\,\mathrm{d}y - \int_{[0,1]} f(x-y)P(y)\,\mathrm{d}y \right| \\
&= \left| \int_{[0,1]} (f(x) - f(x-y))P(y)\,\mathrm{d}y \right| \\
&\le \int_{[0,1]} |f(x) - f(x-y)|P(y)\,\mathrm{d}y.
\end{aligned}$$

The right-hand side can be split as

$$\begin{aligned}
\int_{[0,\delta]} |f(x) - f(x-y)|P(y)\,\mathrm{d}y &+ \int_{[\delta,1-\delta]} |f(x) - f(x-y)|P(y)\,\mathrm{d}y \\
&+ \int_{[1-\delta,1]} |f(x) - f(x-y)|P(y)\,\mathrm{d}y
\end{aligned}$$

which we can bound from above by

$$\begin{aligned}
&\le \int_{[0,\delta]} \varepsilon P(y)\,\mathrm{d}y + \int_{[\delta,1-\delta]} 2M\varepsilon\,\mathrm{d}y \\
&\quad + \int_{[1-\delta,1]} |f(x-1) - f(x-y)|P(y)\,\mathrm{d}y \\
&\le \int_{[0,\delta]} \varepsilon P(y)\,\mathrm{d}y + \int_{[\delta,1-\delta]} 2M\varepsilon\,\mathrm{d}y + \int_{[1-\delta,1]} \varepsilon P(y)\,\mathrm{d}y \\
&\le \varepsilon + 2M\varepsilon + \varepsilon \\
&= (2M+2)\varepsilon.
\end{aligned}$$

Thus we have $\|f - f * P\|_\infty \le (2M+2)\varepsilon$. Since M is fixed and ε is arbitrary, we can thus make $f * P$ arbitrarily close to f in sup norm, which proves the periodic Weierstrass approximation theorem. □

— Exercise —

Exercise 5.4.1 Show that if $f : \mathbf{R} \to \mathbf{C}$ is both compactly supported and $\mathbf{Z}$-periodic, then it is identically zero.

Exercise 5.4.2 Prove Lemma 5.4.4. (Hint: to prove that $f * g$ is continuous, you will have to do something like use the fact that f is bounded, and g is uniformly continuous, or vice versa. To prove that $f * g = g * f$, you will need to use the periodicity to "cut and paste" the interval [0, 1].)

Exercise 5.4.3 Fill in the gaps marked (why?) in Lemma 5.4.6. (Hint: for the first identity, use the identities $|z|^2 = z\overline{z}$, $\overline{e_n} = e_{-n}$, and $e_n e_m = e_{n+m}$.)

5.5 The Fourier and Plancherel Theorems

Using the Weierstrass approximation theorem (Theorem 5.4.1), we can now generalize the Fourier and Plancherel identities to arbitrary continuous periodic functions.

Theorem 5.5.1 (Fourier theorem) *For any $f \in C(\mathbf{R}/\mathbf{Z}; \mathbf{C})$, the series $\sum_{n=-\infty}^{\infty} \hat{f}(n)e_n$ converges in L^2 metric to f. In other words, we have*

$$\lim_{N\to\infty} \left\| f - \sum_{n=-N}^{N} \hat{f}(n)e_n \right\|_2 = 0.$$

Proof Let $\varepsilon > 0$. We have to show that there exists an N_0 such that $\|f - \sum_{n=-N}^{N} \hat{f}(n)e_n\|_2 \le \varepsilon$ for all sufficiently large N.

By the Weierstrass approximation theorem (Theorem 5.4.1), we can find a trigonometric polynomial $P = \sum_{n=-N_0}^{N_0} c_n e_n$ such that $\|f - P\|_\infty \le \varepsilon$, for some $N_0 > 0$. In particular we have $\|f - P\|_2 \le \varepsilon$.

Now let $N > N_0$, and let $F_N := \sum_{n=-N}^{n=N} \hat{f}(n)e_n$. We claim that $\|f - F_N\|_2 \le \varepsilon$. First observe that for any $|m| \le N$, we have

$$\langle f - F_N, e_m\rangle = \langle f, e_m\rangle - \sum_{n=-N}^{N} \hat{f}(n)\langle e_n, e_m\rangle = \hat{f}(m) - \hat{f}(m) = 0,$$

where we have used Lemma 5.3.5. In particular we have

$$\langle f - F_N, F_N - P\rangle = 0$$

since we can write $F_N - P$ as a linear combination of the e_m for which $|m| \le N$. By Pythagoras' theorem we therefore have

$$\|f - P\|_2^2 = \|f - F_N\|_2^2 + \|F_N - P\|_2^2$$

and in particular

$$\|f - F_N\|_2 \le \|f - P\|_2 \le \varepsilon$$

as desired. □

Remark 5.5.2 Note that we have only obtained convergence of the Fourier series $\sum_{n=-\infty}^{\infty} \hat{f}(n)e_n$ to f in the L^2 metric. One may ask whether one has convergence in the uniform or pointwise sense as well, but it turns out (perhaps somewhat surprisingly) that the answer is no to both of those questions. However, if one assumes that the function f is not only continuous, but is also differentiable, then one can recover pointwise convergence; if one assumes continuously differentiable, then one gets uniform convergence as well. These results are beyond the scope of this text and will not be proven here. However, we will prove one theorem about when one can improve the L^2 convergence to uniform convergence.

Theorem 5.5.3 *Let $f \in C(\mathbf{R}/\mathbf{Z}; \mathbf{C})$, and suppose that the series $\sum_{n=-\infty}^{\infty} |\hat{f}(n)|$ is absolutely convergent. Then the series $\sum_{n=-\infty}^{\infty} \hat{f}(n)e_n$ converges uniformly to f. In other words, we have*

$$\lim_{N\to\infty} \left\| f - \sum_{n=-N}^{N} \hat{f}(n)e_n \right\|_{\infty} = 0.$$

Proof By the Weierstrass M-test (Theorem 3.5.7), we see that $\sum_{n=-\infty}^{\infty} \hat{f}(n)e_n$ converges to *some* function F, which by Lemma 5.1.5(iii) is also continuous and $\mathbf{Z}$-periodic. (Strictly speaking, the Weierstrass M test was phrased for series from $n = 1$ to $n = \infty$, but also works for series from $n = -\infty$ to $n = +\infty$; this can be seen by splitting the doubly infinite series into two pieces.) Thus

$$\lim_{N\to\infty} \left\| F - \sum_{n=-N}^{N} \hat{f}(n)e_n \right\|_{\infty} = 0$$

which implies that

$$\lim_{N\to\infty} \left\| F - \sum_{n=-N}^{N} \hat{f}(n)e_n \right\|_{2} = 0$$

since the L^2 norm is always less than or equal to the L^∞ norm. But the sequence $\sum_{n=-N}^{N} \hat{f}(n)e_n$ is already converging in L^2 metric to f by the Fourier theorem, so can only converge in L^2 metric to F if $F = f$ (cf. Proposition 1.1.20). Thus $F = f$, and so we have

$$\lim_{N\to\infty} \left\| f - \sum_{n=-N}^{N} \hat{f}(n)e_n \right\|_{\infty} = 0$$

as desired. □

As a corollary of the Fourier theorem, we obtain

Theorem 5.5.4 (Plancherel theorem) *For any $f \in C(\mathbf{R}/\mathbf{Z}; \mathbf{C})$, the series $\sum_{n=-\infty}^{\infty} |\hat{f}(n)|^2$ is absolutely convergent, and*

$$\|f\|_2^2 = \sum_{n=-\infty}^{\infty} |\hat{f}(n)|^2.$$

This theorem is also known as *Parseval's theorem*.

Proof Let $\varepsilon > 0$. By the Fourier theorem we know that

$$\left\| f - \sum_{n=-N}^{N} \hat{f}(n) e_n \right\|_2 \le \varepsilon$$

if N is large enough (depending on ε). In particular, by the triangle inequality this implies that

$$\|f\|_2 - \varepsilon \le \left\| \sum_{n=-N}^{N} \hat{f}(n) e_n \right\|_2 \le \|f\|_2 + \varepsilon.$$

On the other hand, by Corollary 5.3.6 we have

$$\left\| \sum_{n=-N}^{N} \hat{f}(n) e_n \right\|_2 = \left(\sum_{n=-N}^{N} |\hat{f}(n)|^2 \right)^{1/2}$$

and hence

$$(\|f\|_2 - \varepsilon)^2 \le \sum_{n=-N}^{N} |\hat{f}(n)|^2 \le (\|f\|_2 + \varepsilon)^2.$$

Taking lim sup, we obtain

$$(\|f\|_2 - \varepsilon)^2 \le \limsup_{N\to\infty} \sum_{n=-N}^{N} |\hat{f}(n)|^2 \le (\|f\|_2 + \varepsilon)^2.$$

Since ε is arbitrary, we thus obtain by the squeeze test that

$$\limsup_{N\to\infty} \sum_{n=-N}^{N} |\hat{f}(n)|^2 = \|f\|_2^2$$

and the claim follows. □

There are many other properties of the Fourier transform, but we will not develop them here. In the exercises you will see a small number of applications of the Fourier and Plancherel theorems.

— Exercise —

Exercise 5.5.1 Let f be a function in $C(\mathbf{R}/\mathbf{Z}; \mathbf{C})$, and define the *trigonometric Fourier coefficients* a_n, b_n for $n = 0, 1, 2, 3, \ldots$ by

$$a_n := 2 \int_{[0,1]} f(x)\cos(2\pi nx)\, dx; \quad b_n := 2 \int_{[0,1]} f(x)\sin(2\pi nx)\, dx.$$

(a) Show that the series

$$\frac{1}{2}a_0 + \sum_{n=1}^{\infty}(a_n \cos(2\pi nx) + b_n \sin(2\pi nx))$$

converges in L^2 metric to f. (*Hint:* use the Fourier theorem, and break up the exponentials into sines and cosines. Combine the positive n terms with the negative n terms.)

(b) Show that if $\sum_{n=1}^{\infty} a_n$ and $\sum_{n=1}^{\infty} b_n$ are absolutely convergent, then the above series actually converges uniformly to f, and not just in L^2 metric. (*Hint:* use Theorem 5.5.3.)

Exercise 5.5.2 Let $f(x)$ be the function defined by $f(x) = (1 - 2x)^2$ when $x \in [0, 1)$, and extended to be $\mathbf{Z}$-periodic for the rest of the real line.

(a) Using Exercise 5.5.1, show that the series

$$\frac{1}{3} + \sum_{n=1}^{\infty} \frac{4}{\pi^2 n^2} \cos(2\pi nx)$$

converges uniformly to f.

(b) Conclude that $\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$. (Hint: evaluate the above series at $x = 0$.)

(c) Conclude that $\sum_{n=1}^{\infty} \frac{1}{n^4} = \frac{\pi^4}{90}$. (Hint: expand the cosines in terms of exponentials, and use Plancherel's theorem.)

Exercise 5.5.3 If $f \in C(\mathbf{R}/\mathbf{Z}; \mathbf{C})$ and P is a trigonometric polynomial, show that

$$\widehat{f * P}(n) = \hat{f}(n)c_n = \hat{f}(n)\hat{P}(n)$$

for all integers n. More generally, if $f, g \in C(\mathbf{R}/\mathbf{Z}; \mathbf{C})$, show that

$$\widehat{f * g}(n) = \hat{f}(n)\hat{g}(n)$$

for all integers n. (A fancy way of saying this is that the Fourier transform *intertwines* convolution and multiplication.)

Exercise 5.5.4 Let $f \in C(\mathbf{R}/\mathbf{Z}; \mathbf{C})$ be a function which is differentiable, and whose derivative f' is also continuous (where we define derivatives of complex-valued functions in exactly the same way as for their real-valued counterparts). Show that f' also lies in $C(\mathbf{R}/\mathbf{Z}; \mathbf{C})$, and that $\widehat{f'}(n) = 2\pi i n \hat{f}(n)$ for all integers n.

Exercise 5.5.5 Let $f, g \in C(\mathbf{R}/\mathbf{Z}; \mathbf{C})$. Prove the *Parseval identity*

$$\Re \int_0^1 f(x)\overline{g(x)}\, \mathrm{d}x = \Re \sum_{n \in \mathbf{Z}} \hat{f}(n)\overline{\hat{g}(n)}.$$

(Hint: apply the Plancherel theorem to $f + g$ and $f - g$, and subtract the two.) Then conclude that the real parts can be removed, thus

$$\int_0^1 f(x)\overline{g(x)}\, \mathrm{d}x = \sum_{n \in \mathbf{Z}} \hat{f}(n)\overline{\hat{g}(n)}.$$

(Hint: apply the first identity with f replaced by if.)

Exercise 5.5.6 In this exercise we shall develop the theory of Fourier series for functions of any fixed period L.

Let $L > 0$, and let $f : \mathbf{R} \to \mathbf{C}$ be a complex-valued function which is continuous and L-periodic. Define the numbers c_n for every integer n by

$$c_n := \frac{1}{L} \int_{[0,L]} f(x) e^{-2\pi i n x/L}\, \mathrm{d}x.$$

(a) Show that the series

$$\sum_{n=-\infty}^{\infty} c_n e^{2\pi i n x/L}$$

converges in L^2 metric to f. More precisely, show that

$$\lim_{N \to \infty} \int_{[0,L]} |f(x) - \sum_{n=-N}^{N} c_n e^{2\pi i n x/L}|^2\, \mathrm{d}x = 0.$$

(Hint: apply the Fourier theorem to the function $f(Lx)$.)

(b) If the series $\sum_{n=-\infty}^{\infty} |c_n|$ is absolutely convergent, show that

$$\sum_{n=-\infty}^{\infty} c_n e^{2\pi i n x/L}$$

converges uniformly to f.

(c) Show that

$$\frac{1}{L}\int_{[0,L]} |f(x)|^2 \,\mathrm{d}x = \sum_{n=-\infty}^{\infty} |c_n|^2.$$

(Hint: apply the Plancherel theorem to the function $f(Lx)$.)

Chapter 6
Several Variable Differential Calculus

6.1 Linear Transformations

We shall now switch to a different topic, namely that of differentiation in several variable calculus. More precisely, we shall be dealing with maps $f : \mathbf{R}^n \to \mathbf{R}^m$ from one Euclidean space to another, and trying to understand what the derivative of such a map is.

Before we do so, however, we need to recall some notions from linear algebra, most importantly that of a linear transformation and a matrix. We shall be rather brief here; a more thorough treatment of this material can be found in any linear algebra text.

Definition 6.1.1 (*Row vectors*) Let $n \geq 1$ be an integer. We refer to elements of $\mathbf{R}^n$ as *n-dimensional row vectors*. A typical n-dimensional row vector may take the form $x = (x_1, x_2, \ldots, x_n)$, which we abbreviate as $(x_i)_{1\leq i\leq n}$; the quantities $x_1, x_2, \ldots, x_n$ are of course real numbers. If $(x_i)_{1\leq i\leq n}$ and $(y_i)_{1\leq i\leq n}$ are n-dimensional row vectors, we can define their vector sum by

$$(x_i)_{1\leq i\leq n} + (y_i)_{1\leq i\leq n} = (x_i + y_i)_{1\leq i\leq n},$$

and also if $c \in \mathbf{R}$ is any scalar, we can define the scalar product $c(x_i)_{1\leq i\leq n}$ by

$$c(x_i)_{1\leq i\leq n} := (cx_i)_{1\leq i\leq n}.$$

Of course one has similar operations on $\mathbf{R}^m$ as well. However, if $n \neq m$, then we do not define any operation of vector addition between vectors in $\mathbf{R}^n$ and vectors in $\mathbf{R}^m$ (e.g., $(2, 3, 4) + (5, 6)$ is undefined). We also refer to the vector $(0, \ldots, 0)$ in $\mathbf{R}^n$ as the *zero vector* and also denote it by 0. (Strictly speaking, we should denote the zero vector of $\mathbf{R}^n$ by $0_{\mathbf{R}^n}$, as they are technically distinct from each other and from the number zero, but we shall not take care to make this distinction.) We abbreviate $(-1)x$ as $-x$.

T. Tao, *Analysis II*, Texts and Readings in Mathematics,
https://doi.org/10.1007/978-981-19-7284-3_6

The operations of vector addition and scalar multiplication obey a number of basic properties:

Lemma 6.1.2 ($\mathbf{R}^n$ is a vector space) *Let x, y, z be vectors in $\mathbf{R}^n$, and let c, d be real numbers. Then we have the commutativity property $x + y = y + x$, the additive associativity property $(x + y) + z = x + (y + z)$, the additive identity property $x + 0 = 0 + x = x$, the additive inverse property $x + (-x) = (-x) + x = 0$, the multiplicative associativity property $(cd)x = c(dx)$, the distributivity properties $c(x + y) = cx + cy$ and $(c + d)x = cx + dx$, and the multiplicative identity property $1x = x$.*

Proof See Exercise 6.1.1. □

Definition 6.1.3 (*Transpose*) If $(x_i)_{1\le i\le n} = (x_1, x_2, \ldots, x_n)$ is an n-dimensional row vector, we can define its *transpose* $(x_i)^T_{1\le i\le n}$ by

$$(x_i)^T_{1\le i\le n} = (x_1, x_2, \ldots, x_n)^T := \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

We refer to objects such as $(x_i)^T_{1\le i\le n}$ as *n-dimensional column vectors.*

Remark 6.1.4 There is no functional difference between a row vector and a column vector (e.g., one can add and scalar multiply column vectors just as well as we can row vectors); however we shall (rather annoyingly) need to transpose our row vectors into column vectors in order to be consistent with the conventions of matrix multiplication, which we will see later. Note that we view row vectors and column vectors as residing in different spaces; thus for instance we will not define the sum of a row vector with a column vector, even when they have the same number of elements.

Definition 6.1.5 (*Standard basis row vectors*) We identify n special vectors in $\mathbf{R}^n$, the *standard basis row vectors* $e_1, \ldots, e_n$. For each $1 \le j \le n$, e_j is the vector which has 0 in all entries except for the j-th entry, which is equal to 1.

For instance, in $\mathbf{R}^3$, we have $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$, and $e_3 = (0, 0, 1)$. Note that if $x = (x_j)_{1\le j\le n}$ is a vector in $\mathbf{R}^n$, then

$$x = x_1e_1 + x_2e_2 + \ldots + x_ne_n = \sum_{j=1}^{n} x_je_j,$$

or in other words every vector in $\mathbf{R}^n$ is a *linear combination* of the standard basis vectors $e_1, \ldots, e_n$. (The notation $\sum_{j=1}^n x_je_j$ is unambiguous because the operation of vector addition is both commutative and associative). Of course, just as every row vector is a linear combination of standard basis row vectors, every column vector is a linear combination of standard basis column vectors:

$$x^T = x_1 e_1^T + x_2 e_2^T + \ldots + x_n e_n^T = \sum_{j=1}^{n} x_j e_j^T.$$

There are (many) other ways to create a basis for $\mathbf{R}^n$, but this is a topic for a linear algebra text and will not be discussed here.

Definition 6.1.6 (*Linear transformations*) A *linear transformation* $T\colon \mathbf{R}^n \to \mathbf{R}^m$ is any function from one Euclidean space $\mathbf{R}^n$ to another $\mathbf{R}^m$ which obeys the following two axioms:

(a) (Additivity) For every $x, x' \in \mathbf{R}^n$, we have $T(x + x') = Tx + Tx'$.
(b) (Homogeneity) For every $x \in \mathbf{R}^n$ and every $c \in \mathbf{R}$, we have $T(cx) = cTx$.

Example 6.1.7 The *dilation operator* $T_1\colon \mathbf{R}^3 \to \mathbf{R}^3$ defined by $T_1 x := 5x$ (i.e., it dilates each vector x by a factor of 5) is a linear transformation, since $5(x + x') = 5x + 5x'$ for all $x, x' \in \mathbf{R}^3$ and $5(cx) = c(5x)$ for all $x \in \mathbf{R}^3$ and $c \in \mathbf{R}$.

Example 6.1.8 The *rotation operator* $T_2\colon \mathbf{R}^2 \to \mathbf{R}^2$ defined by a counterclockwise rotation by $\pi/2$ radians around the origin (so that $T_2(1, 0) = (0, 1)$, $T_2(0, 1) = (-1, 0)$, etc.) is a linear transformation; this can best be seen geometrically rather than analytically.

Example 6.1.9 The *projection operator* $T_3 : \mathbf{R}^3 \to \mathbf{R}^2$ defined by $T_3(x, y, z) := (x, y)$ is a linear transformation (why?). The *inclusion operator* $T_4 : \mathbf{R}^2 \to \mathbf{R}^3$ defined by $T_4(x, y) := (x, y, 0)$ is also a linear transformation (why?). Finally, the *identity operator* $I_n : \mathbf{R}^n \to \mathbf{R}^n$, defined for any n by $I_n x := x$ is also a linear transformation (why?).

As we shall shortly see, there is a connection between linear transformations and matrices.

Definition 6.1.10 (*Matrices*) An $m \times n$ *matrix* is an object A of the form

$$A = \begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix};$$

we shall abbreviate this as

$$A = (a_{ij})_{1 \le i \le m; 1 \le j \le n}.$$

In particular, n-dimensional row vectors are $1 \times n$ matrices, while n-dimensional column vectors are $n \times 1$ matrices.

Definition 6.1.11 (*Matrix product*) Given an $m \times n$ matrix A and an $n \times p$ matrix B, we can define the *matrix product AB* to be the $m \times p$ matrix defined as

$$(a_{ij})_{1\le i\le m;1\le j\le n}(b_{jk})_{1\le j\le n;1\le k\le p} := \left(\sum_{j=1}^{n} a_{ij}b_{jk}\right)_{1\le i\le m;1\le k\le p}.$$

In particular, if $x^T = (x_j)^T_{1\le j\le n}$ is an n-dimensional column vector, and $A = (a_{ij})_{1\le i\le m;1\le j\le n}$ is an $m \times n$ matrix, then Ax^T is an m-dimensional column vector:

$$Ax^T = \left(\sum_{j=1}^{n} a_{ij}x_j\right)^T_{1\le i\le m}.$$

We now relate matrices to linear transformations. If A is an $m \times n$ matrix, we can define the transformation $L_A : \mathbf{R}^n \to \mathbf{R}^m$ by the formula

$$(L_A x)^T := Ax^T.$$

Example 6.1.12 If A is the matrix

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix},$$

and $x = (x_1, x_2, x_3)$ is a 3-dimensional row vector, then $L_A x$ is the 2-dimensional row vector defined by

$$(L_A x)^T = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 + 2x_2 + 3x_3 \\ 4x_1 + 5x_2 + 6x_3 \end{pmatrix}$$

or in other words

$$L_A(x_1, x_2, x_3) = (x_1 + 2x_2 + 3x_3, 4x_1 + 5x_2 + 6x_3).$$

More generally, if

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

then we have

$$L_A(x_j)_{1\le j\le n} = \left(\sum_{j=1}^{n} a_{ij}x_j\right)_{1\le i\le m}.$$

For any $m \times n$ matrix A, the transformation L_A is automatically linear; one can easily verify that $L_A(x + y) = L_A x + L_A y$ and $L_A(cx) = c(L_A x)$ for any n-dimensional row vectors x, y and any scalar c. (Why?)

Perhaps surprisingly, the converse is also true, i.e., every linear transformation from $\mathbf{R}^n$ to $\mathbf{R}^m$ is given by a matrix:

Lemma 6.1.13 *Let $T: \mathbf{R}^n \to \mathbf{R}^m$ be a linear transformation. Then there exists exactly one $m \times n$ matrix A such that $T = L_A$.*

Proof Suppose $T: \mathbf{R}^n \to \mathbf{R}^m$ is a linear transformation. Let $e_1, e_2, \ldots, e_n$ be the standard basis row vectors of $\mathbf{R}^n$. Then $Te_1, Te_2, \ldots, Te_n$ are vectors in $\mathbf{R}^m$. For each $1 \le j \le n$, we write Te_j in co-ordinates as

$$Te_j = (a_{1j}, a_{2j}, \ldots, a_{mj}) = (a_{ij})_{1\le i\le m},$$

i.e., we define a_{ij} to be the i^{th} component of Te_j. Then for any n-dimensional row vector $x = (x_1, \ldots, x_n)$, we have

$$Tx = T\left(\sum_{j=1}^{n} x_j e_j\right),$$

which (since T is linear) is equal to

$$\begin{aligned}
&= \sum_{j=1}^{n} T(x_j e_j)\\
&= \sum_{j=1}^{n} x_j Te_j\\
&= \sum_{j=1}^{n} x_j (a_{ij})_{1\le i\le m}\\
&= \sum_{j=1}^{n} (a_{ij}x_j)_{1\le i\le m}\\
&= \left(\sum_{j=1}^{n} a_{ij}x_j\right)_{1\le i\le m}.
\end{aligned}$$

But if we let A be the matrix

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

then the previous vector is precisely $L_A x$. Thus $Tx = L_A x$ for all n-dimensional vectors x, and thus $T = L_A$.

Now we show that A is unique, i.e., there does not exist any other matrix

$$B = \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1n} \\ b_{21} & b_{22} & \dots & b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ b_{m1} & b_{m2} & \dots & b_{mn} \end{pmatrix}$$

for which T is equal to L_B. Suppose for sake of contradiction that we could find such a matrix B which was different from A. Then we would have $L_A = L_B$. In particular, we have $L_A e_j = L_B e_j$ for every $1 \leq j \leq n$. But from the definition of L_A we see that

$$L_A e_j = (a_{ij})_{1 \leq i \leq m}$$

and

$$L_B e_j = (b_{ij})_{1 \leq i \leq m}$$

and thus we have $a_{ij} = b_{ij}$ for every $1 \leq i \leq m$ and $1 \leq j \leq n$, thus A and B are equal, a contradiction. □

Remark 6.1.14 Lemma 6.1.13 establishes a one-to-one correspondence between linear transformations and matrices, and is one of the fundamental reasons why matrices are so important in linear algebra. One may ask then why we bother dealing with linear transformations at all, and why we don't just work with matrices all the time. The reason is that sometimes one does not want to work with the standard basis $e_1, \dots, e_n$, but instead wants to use some other basis. In that case, the correspondence between linear transformations and matrices changes, and so it is still important to keep the notions of linear transformation and matrix distinct. More discussion on this somewhat subtle issue can be found in any linear algebra text.

Remark 6.1.15 If $T = L_A$, then A is sometimes called the *matrix representation of* T and is sometimes denoted $A = [T]$. We shall avoid this notation here, however.

The composition $T \circ S$ of two linear transformations T, S is again a linear transformation (Exercise 6.1.2). It is customary in linear algebra to abbreviate such compositions $T \circ S$ simply as TS. The next lemma shows that the operation of composing linear transformations is connected to that of matrix multiplication.

Lemma 6.1.16 *Let A be an $m \times n$ matrix, and let B be an $n \times p$ matrix. Then $L_A L_B = L_{AB}$.*

Proof See Exercise 6.1.3. □

— Exercise —

Exercise 6.1.1 Prove Lemma 6.1.2.

Exercise 6.1.2 If $T: \mathbf{R}^n \to \mathbf{R}^m$ is a linear transformation, and $S: \mathbf{R}^p \to \mathbf{R}^n$ is a linear transformation, show that the composition $TS: \mathbf{R}^p \to \mathbf{R}^m$ of the two transforms, defined by $TS(x) := T(S(x))$, is also a linear transformation. (*Hint:* expand $TS(x + y)$ and $TS(cx)$ carefully, using plenty of parentheses.)

Exercise 6.1.3 Prove Lemma 6.1.16.

Exercise 6.1.4 Let $T: \mathbf{R}^n \to \mathbf{R}^m$ be a linear transformation. Show that there exists a number $M > 0$ such that $\|Tx\| \le M \|x\|$ for all $x \in \mathbf{R}^n$. (*Hint:* use Lemma 6.1.13 to write T in terms of a matrix A, and then set M to be the sum of the absolute values of all the entries in A. Use the triangle inequality often—it's easier than messing around with square roots, etc.) Conclude in particular that every linear transformation from $\mathbf{R}^n$ to $\mathbf{R}^m$ is continuous.

6.2 Derivatives in Several Variable Calculus

Now that we've reviewed some linear algebra, we turn now to our main topic of this chapter, which is that of understanding differentiation of functions of the form $f: \mathbf{R}^n \to \mathbf{R}^m$, i.e., functions from one Euclidean space to another. For instance, one might want to differentiate the function $f: \mathbf{R}^3 \to \mathbf{R}^4$ defined by

$$f(x, y, z) = (xy, yz, xz, xyz).$$

In single-variable calculus, when one wants to differentiate a function $f: E \to \mathbf{R}$ at a point x_0, where E is a subset of $\mathbf{R}$ that contains x_0, this is given by

$$f'(x_0) := \lim_{x \to x_0; x \in E \setminus \{x_0\}} \frac{f(x) - f(x_0)}{x - x_0}.$$

One could try to mimic this definition in the several variable case $f: E \to \mathbf{R}^m$, where E is now a subset of $\mathbf{R}^n$; however we encounter a difficulty in this case: the quantity $f(x) - f(x_0)$ will live in $\mathbf{R}^m$, and $x - x_0$ lives in $\mathbf{R}^n$, and we do not know how to divide an m-dimensional vector by an n-dimensional vector.

To get around this problem, we first rewrite the concept of derivative (in one dimension) in a way which does not involve division of vectors. Instead, we view differentiability at a point x_0 as an assertion that a function f is "approximately linear" near x_0.

Lemma 6.2.1 *Let E be a subset of $\mathbf{R}$, $f : E \to \mathbf{R}$ be a function, and $L \in \mathbf{R}$. Let x_0 be a limit point of E. Then the following two statements are equivalent.*

(a) f is differentiable at x_0, and $f'(x_0) = L$.
(b) We have $\lim_{x \to x_0; x \in E - \{x_0\}} \frac{|f(x) - (f(x_0) + L(x - x_0))|}{|x - x_0|} = 0$.

Proof See Exercise 6.2.1. □

In light of the above lemma, we see that the derivative $f'(x_0)$ can be interpreted as the number L for which $|f(x) - (f(x_0) + L(x - x_0))|$ is small, in the sense that it tends to zero as x tends to x_0, even if we divide out by the very small number $|x - x_0|$. More informally, the derivative is the quantity L such that we have the approximation $f(x) - f(x_0) \approx L(x - x_0)$.

This does not seem too different from the usual notion of differentiation, but the point is that we are no longer explicitly dividing by $x - x_0$. (We are still dividing by $|x - x_0|$, but this will turn out to be OK.) When we move to the several variable case $f : E \to \mathbf{R}^m$, where $E \subseteq \mathbf{R}^n$, we shall still want the derivative to be some quantity L such that $f(x) - f(x_0) \approx L(x - x_0)$. However, since $f(x) - f(x_0)$ is now an m-dimensional vector and $x - x_0$ is an n-dimensional vector, we no longer want L to be a scalar; we want it to be a linear transformation. More precisely:

Definition 6.2.2 (*Differentiability*) Let E be a subset of $\mathbf{R}^n$, $f : E \to \mathbf{R}^m$ be a function, $x_0 \in E$ be a limit point of E, and let $L : \mathbf{R}^n \to \mathbf{R}^m$ be a linear transformation. We say that f is *differentiable at x_0 with derivative L* if we have

$$\lim_{x \to x_0; x \in E - \{x_0\}} \frac{\|f(x) - (f(x_0) + L(x - x_0))\|}{\|x - x_0\|} = 0.$$

Here $\|x\|$ is the length of x (as measured in the l^2 metric):

$$\|(x_1, x_2, \ldots, x_n)\| = (x_1^2 + x_2^2 + \ldots + x_n^2)^{1/2}.$$

Example 6.2.3 Let $f : \mathbf{R}^2 \to \mathbf{R}^2$ be the map $f(x, y) := (x^2, y^2)$, let x_0 be the point $x_0 := (1, 2)$, and let $L : \mathbf{R}^2 \to \mathbf{R}^2$ be the map $L(x, y) := (2x, 4y)$. We claim that f is differentiable at x_0 with derivative L. To see this, we compute

$$\lim_{(x,y) \to (1,2) : (x,y) \neq (1,2)} \frac{\|f(x, y) - (f(1, 2) + L((x, y) - (1, 2)))\|}{\|(x, y) - (1, 2)\|}.$$

Making the change of variables $(x, y) = (1, 2) + (a, b)$, this becomes

$$\lim_{(a,b) \to (0,0) : (a,b) \neq (0,0)} \frac{\|f(1 + a, 2 + b) - (f(1, 2) + L(a, b))\|}{\|(a, b)\|}.$$

Substituting the formula for f and for L, this becomes

$$\lim_{(a,b) \to (0,0) : (a,b) \neq (0,0)} \frac{\|((1 + a)^2, (2 + b)^2) - (1, 4) - (2a, 4b))\|}{\|(a, b)\|},$$

which simplifies to

$$\lim_{(a,b)\to(0,0):(a,b)\neq(0,0)} \frac{\|(a^2,b^2)\|}{\|(a,b)\|}.$$

We use the squeeze test. The expression $\frac{\|(a^2,b^2)\|}{\|(a,b)\|}$ is clearly non-negative. On the other hand, we have by the triangle inequality

$$\|(a^2,b^2)\| \leq \|(a^2,0)\| + \|(0,b^2)\| = a^2 + b^2$$

and hence

$$\frac{\|(a^2,b^2)\|}{\|(a,b)\|} \leq \sqrt{a^2+b^2}.$$

Since $\sqrt{a^2+b^2} \to 0$ as $(a,b) \to 0$, we thus see from the squeeze test that the above limit exists and is equal to 0. Thus f is differentiable at x_0 with derivative L.

As you can see, verifying that a function is differentiable from first principles can be somewhat tedious. Later on we shall find better ways to verify differentiability, and to compute derivatives.

Before we proceed further, we have to check a basic fact, which is that a function can have at most one derivative at any *interior* point of its domain:

Lemma 6.2.4 (Uniqueness of derivatives) *Let E be a subset of $\mathbf{R}^n$, $f : E \to \mathbf{R}^m$ be a function, $x_0 \in E$ be an* interior point *of E, and let $L_1 : \mathbf{R}^n \to \mathbf{R}^m$ and $L_2 : \mathbf{R}^n \to \mathbf{R}^m$ be linear transformations. Suppose that f is differentiable at x_0 with derivative L_1, and also differentiable at x_0 with derivative L_2. Then $L_1 = L_2$.*

Proof See Exercise 6.2.2. □

Because of Lemma 6.2.4, we can now talk about *the* derivative of f at interior points x_0, and we will denote this derivative by $f'(x_0)$. Thus $f'(x_0)$ is the unique linear transformation from $\mathbf{R}^n$ to $\mathbf{R}^m$ such that

$$\lim_{x\to x_0; x\neq x_0} \frac{\|f(x) - (f(x_0) + f'(x_0)(x-x_0))\|}{\|x-x_0\|} = 0.$$

Informally, this means that the derivative $f'(x_0)$ is the linear transformation such that we have

$$f(x) - f(x_0) \approx f'(x_0)(x-x_0)$$

or equivalently

$$f(x) \approx f(x_0) + f'(x_0)(x-x_0)$$

(this is known as *Newton's approximation*; compare with Proposition 10.1.7).

Another consequence of Lemma 6.2.4 is that if you know that $f(x) = g(x)$ for all $x \in E$, and f, g are differentiable at x_0, then you also know that $f'(x_0) = g'(x_0)$ at every *interior* point of E. However, this is not necessarily true if x_0 is a boundary point of E; for instance, if E is just a single point $E = \{x_0\}$, merely knowing that $f(x_0) = g(x_0)$ does not imply that $f'(x_0) = g'(x_0)$. We will not deal with these boundary issues here and only compute derivatives on the interior of the domain.

We will sometimes refer to f' as the *total derivative* of f, to distinguish this concept from that of partial and directional derivatives below. The total derivative f is also closely related to the *derivative matrix* Df, which we shall define in the next section.

— Exercise —

Exercise 6.2.1 Prove Lemma 6.2.1.

Exercise 6.2.2 Prove Lemma 6.2.4. (*Hint:* prove by contradiction. If $L_1 \neq L_2$, then there exists a vector v such that $L_1 v \neq L_2 v$; this vector must be nonzero (why?). Now apply the definition of derivative, and try to specialize to the case where $x = x_0 + tv$ for some scalar t, to obtain a contradiction.)

6.3 Partial and Directional Derivatives

We now connect the notion of differentiability with that of partial and directional derivatives, which we now introduce.

Definition 6.3.1 (*Directional derivative*) Let E be a subset of $\mathbf{R}^n$, $f : E \to \mathbf{R}^m$ be a function, let x_0 be an interior point of E, and let v be a vector in $\mathbf{R}^n$. If the limit

$$\lim_{t\to 0; t>0, x_0+tv\in E} \frac{f(x_0 + tv) - f(x_0)}{t}$$

exists, we say that f is *differentiable in the direction v at x_0*, and we denote the above limit by $D_v f(x_0)$:

$$D_v f(x_0) := \lim_{t\to 0; t>0} \frac{f(x_0 + tv) - f(x_0)}{t}.$$

Remark 6.3.2 One should compare this definition with Definition 6.2.2. Note that we are dividing by a scalar t, rather than a vector, so this definition makes sense, and $D_v f(x_0)$ will be a vector in $\mathbf{R}^m$. It is sometimes possible to also define directional derivatives on the boundary of E, if the vector v is pointing in an "inward" direction (this generalizes the notion of left derivatives and right derivatives from single-variable calculus); but we will not pursue these matters here.

Example 6.3.3 If $f : \mathbf{R} \to \mathbf{R}$ is a function, then $D_{+1} f(x)$ is the same as the right derivative of $f(x)$ (if it exists), and similarly $D_{-1} f(x)$ is the same as the negative of the left derivative of $f(x)$ (if it exists).

Example 6.3.4 We use the function $f : \mathbf{R}^2 \to \mathbf{R}^2$ defined by $f(x, y) := (x^2, y^2)$ from before, and let $x_0 := (1, 2)$ and $v := (3, 4)$. Then

$$\begin{aligned} D_v f(x_0) &= \lim_{t \to 0; t > 0} \frac{f(1 + 3t, 2 + 4t) - f(1, 2)}{t} \\ &= \lim_{t \to 0; t > 0} \frac{(1 + 6t + 9t^2, 4 + 16t + 16t^2) - (1, 4)}{t} \\ &= \lim_{t \to 0; t > 0} (6 + 9t, 16 + 16t) = (6, 16). \end{aligned}$$

Directional derivatives are connected with total derivatives as follows:

Lemma 6.3.5 *Let E be a subset of $\mathbf{R}^n$, $f : E \to \mathbf{R}^m$ be a function, x_0 be an interior point of E, and let v be a vector in $\mathbf{R}^n$. If f is differentiable at x_0, then f is also differentiable in the direction v at x_0, and*

$$D_v f(x_0) = f'(x_0) v.$$

Proof See Exercise 6.3.1. □

Remark 6.3.6 One consequence of this lemma is that total differentiability implies directional differentiability. However, the converse is not true; see Exercise 6.3.3.

Closely related to the concept of directional derivative is that of *partial derivative*:

Definition 6.3.7 (*Partial derivative*) Let E be a subset of $\mathbf{R}^n$, let $f : E \to \mathbf{R}^m$ be a function, let x_0 be an interior point of E, and let $1 \leq j \leq n$. Then the *partial derivative of f with respect to the x_j variable* at x_0, denoted $\frac{\partial f}{\partial x_j}(x_0)$, is defined by

$$\frac{\partial f}{\partial x_j}(x_0) := \lim_{t \to 0; t \neq 0, x_0 + te_j \in E} \frac{f(x_0 + te_j) - f(x_0)}{t} = \frac{d}{dt} f(x_0 + te_j)|_{t=0}$$

provided of course that the limit exists. (If the limit does not exist, we leave $\frac{\partial f}{\partial x_j}(x_0)$ undefined.)

We say that f is *continuously differentiable* if the partial derivatives $\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}$ exist and are continuous on E.

Informally, the partial derivative can be obtained by holding all the variables other than x_j fixed and then applying the single-variable calculus derivative in the x_j variable. Note that if f takes values in $\mathbf{R}^m$, then so will $\frac{\partial f}{\partial x_j}$. Indeed, if we write f in components as $f = (f_1, \ldots, f_m)$, it is easy to see (why?) that

$$\frac{\partial f}{\partial x_j}(x_0) = \left(\frac{\partial f_1}{\partial x_j}(x_0), \ldots, \frac{\partial f_m}{\partial x_j}(x_0) \right),$$

i.e., to differentiate a vector-valued function one just has to differentiate each of the components separately.

We sometimes replace the variables x_j in $\frac{\partial f}{\partial x_j}$ with other symbols. For instance, if we are dealing with the function $f(x, y) = (x^2, y^2)$, then we might refer to $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ instead of $\frac{\partial f}{\partial x_1}$ and $\frac{\partial f}{\partial x_2}$. (In this case, $\frac{\partial f}{\partial x}(x, y) = (2x, 0)$ and $\frac{\partial f}{\partial y}(x, y) = (0, 2y)$.) One should caution however that one should only relabel the variables if it is absolutely clear which symbol refers to the first variable, which symbol refers to the second variable, etc.; otherwise one may become unintentionally confused. For instance, in the above example, the expression $\frac{\partial f}{\partial x}(x, x)$ is just $(2x, 0)$; however one may mistakenly compute

$$\frac{\partial f}{\partial x}(x, x) = \frac{\partial}{\partial x}(x^2, x^2) = (2x, 2x);$$

the problem here is that the symbol x is being used for more than just the first variable of f. (On the other hand, it is true that $\frac{\mathrm{d}}{\mathrm{d}x}f(x, x)$ is equal to $(2x, 2x)$; thus the operation of total differentiation $\frac{\mathrm{d}}{\mathrm{d}x}$ is not the same as that of partial differentiation $\frac{\partial}{\partial x}$.)

From Lemma 6.3.5 (and Proposition 9.5.3 from *Analysis I*), we know that if a function is differentiable at a point x_0, then all the partial derivatives $\frac{\partial f}{\partial x_j}$ exist at x_0, and that

$$\frac{\partial f}{\partial x_j}(x_0) = D_{e_j}f(x_0) = -D_{-e_j}f(x_0) = f'(x_0)e_j.$$

Also, if $v = (v_1, \ldots, v_n) = \sum_j v_j e_j$, then we have

$$D_v f(x_0) = f'(x_0)\sum_j v_j e_j = \sum_j v_j f'(x_0)e_j$$

(since $f'(x_0)$ is linear) and thus

$$D_v f(x_0) = \sum_j v_j \frac{\partial f}{\partial x_j}(x_0).$$

Thus one can write directional derivatives in terms of partial derivatives, *provided that* the function is actually differentiable at that point.

Just because the partial derivatives exist at a point x_0, we cannot conclude that the function is differentiable there (Exercise 6.3.3). However, if we know that the partial derivatives not only exist, but are continuous, then we can in fact conclude differentiability, thanks to the following handy theorem:

Theorem 6.3.8 *Let E be a subset of $\mathbf{R}^n$, $f : E \to \mathbf{R}^m$ be a function, F be a subset of E, and x_0 be an interior point of F. If all the partial derivatives $\frac{\partial f}{\partial x_j}$ exist on F and are continuous at x_0, then f is differentiable at x_0, and the linear transformation $f'(x_0) : \mathbf{R}^n \to \mathbf{R}^m$ is defined by*

$$f'(x_0)(v_j)_{1\le j\le n} = \sum_{j=1}^n v_j \frac{\partial f}{\partial x_j}(x_0).$$

Proof Let $L : \mathbf{R}^n \to \mathbf{R}^m$ be the linear transformation

$$L(v_j)_{1\le j\le n} := \sum_{j=1}^{n} v_j \frac{\partial f}{\partial x_j}(x_0).$$

We have to prove that

$$\lim_{x\to x_0; x\in E-\{x_0\}} \frac{\|f(x) - (f(x_0) + L(x - x_0))\|}{\|x - x_0\|} = 0.$$

Let $\varepsilon > 0$. It will suffice to find a radius $\delta > 0$ such that

$$\frac{\|f(x) - (f(x_0) + L(x - x_0))\|}{\|x - x_0\|} \le \varepsilon$$

for all $x \in B(x_0, \delta)\backslash\{x_0\}$. Equivalently, we wish to show that

$$\|f(x) - f(x_0) - L(x - x_0)\| \le \varepsilon\|x - x_0\|$$

for all $x \in B(x_0, \delta)\backslash\{x_0\}$.

Because x_0 is an interior point of F, there exists a ball $B(x_0, r)$ which is contained inside F. Because each partial derivative $\frac{\partial f}{\partial x_j}$ exists on F and is continuous at x_0, there thus exists an $0 < \delta_j < r$ such that $\|\frac{\partial f}{\partial x_j}(x) - \frac{\partial f}{\partial x_j}(x_0)\| \le \varepsilon/nm$ for every $x \in B(x_0, \delta_j)$. If we take $\delta = \min(\delta_1, \ldots, \delta_n)$, then we thus have $\|\frac{\partial f}{\partial x_j}(x) - \frac{\partial f}{\partial x_j}(x_0)\| \le \varepsilon/nm$ for every $x \in B(x_0, \delta)$ and every $1 \le j \le n$.

Let $x \in B(x_0, \delta)$. We write $x = x_0 + v_1e_1 + v_2e_2 + \ldots + v_ne_n$ for some scalars $v_1, \ldots, v_n$. Note that

$$\|x - x_0\| = \sqrt{v_1^2 + v_2^2 + \ldots + v_n^2}$$

and in particular we have $|v_j| \le \|x - x_0\|$ for all $1 \le j \le n$. Our task is to show that

$$\left\| f(x_0 + v_1e_1 + \ldots + v_ne_n) - f(x_0) - \sum_{j=1}^{n} v_j \frac{\partial f}{\partial x_j}(x_0) \right\| \le \varepsilon\|x - x_0\|.$$

Write f in components as $f = (f_1, f_2, \ldots, f_m)$ (so each f_i is a function from E to $\mathbf{R}$). From the mean value theorem in the x_1 variable, we see that

$$f_i(x_0 + v_1e_1) - f_i(x_0) = \frac{\partial f_i}{\partial x_1}(x_0 + t_ie_1)v_1$$

for some t_i between 0 and v_1. But we have

$$\left|\frac{\partial f_i}{\partial x_j}(x_0 + t_i e_1) - \frac{\partial f_i}{\partial x_j}(x_0)\right| \leq \left\|\frac{\partial f}{\partial x_j}(x_0 + t_i e_1) - \frac{\partial f}{\partial x_j}(x_0)\right\| \leq \varepsilon/nm$$

and hence

$$\left|f_i(x_0 + v_1 e_1) - f_i(x_0) - \frac{\partial f_i}{\partial x_1}(x_0)v_1\right| \leq \varepsilon|v_1|/nm.$$

Summing this over all $1 \leq i \leq m$ (and noting that $\|(y_1, \ldots, y_m)\| \leq |y_1| + \ldots + |y_m|$ from the triangle inequality) we obtain

$$\left\|f(x_0 + v_1 e_1) - f(x_0) - \frac{\partial f}{\partial x_1}(x_0)v_1\right\| \leq \varepsilon|v_1|/n;$$

since $|v_1| \leq \|x - x_0\|$, we thus have

$$\left\|f(x_0 + v_1 e_1) - f(x_0) - \frac{\partial f}{\partial x_1}(x_0)v_1\right\| \leq \varepsilon\|x - x_0\|/n.$$

A similar argument gives

$$\left\|f(x_0 + v_1 e_1 + v_2 e_2) - f(x_0 + v_1 e_1) - \frac{\partial f}{\partial x_2}(x_0)v_2\right\| \leq \varepsilon\|x - x_0\|/n$$

and so forth up to

$$\begin{aligned}&\|f(x_0 + v_1 e_1 + \cdots + v_n e_n) - f(x_0 + v_1 e_1 + \cdots + v_{n-1} e_{n-1})\\ &-\frac{\partial f}{\partial x_n}(x_0)v_n\Big\| \leq \varepsilon \, \|x - x_0\| \, /n.\end{aligned}$$

If we sum these n inequalities and use the triangle inequality $\|x + y\| \leq \|x\| + \|y\|$, we obtain a telescoping series which simplifies to

$$\left\|f(x_0 + v_1 e_1 + \ldots + v_n e_n) - f(x_0) - \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x_0)v_j\right\| \leq \varepsilon\|x - x_0\|$$

as desired. □

From Theorem 6.3.8 and Lemma 6.3.5 we see that if the partial derivatives of a function $f : E \to \mathbf{R}^m$ exist and are continuous on some set F, then all the directional derivatives also exist at every interior point x_0 of F, and we have the formula

$$D_{(v_1,\ldots,v_n)}f(x_0) = \sum_{j=1}^{n} v_j \frac{\partial f}{\partial x_j}(x_0).$$

In particular, if $f : E \to \mathbf{R}$ is a real-valued function, and we define the *gradient* $\nabla f(x_0)$ of f at x_0 to be the n-dimensional row vector $\nabla f(x_0) := (\frac{\partial f}{\partial x_1}(x_0), \ldots, \frac{\partial f}{\partial x_n}(x_0))$, then we have the familiar formula

$$D_v f(x_0) = v \cdot \nabla f(x_0)$$

whenever x_0 is in the interior of the region where the gradient exists and is continuous.

More generally, if $f : E \to \mathbf{R}^m$ is a function taking values in $\mathbf{R}^m$, with $f = (f_1, \ldots, f_m)$, and x_0 is in the interior of the region where the partial derivatives of f exist and are continuous, then we have from Theorem 6.3.8 that

$$f'(x_0)(v_j)_{1 \le j \le n} = \sum_{j=1}^{n} v_j \frac{\partial f}{\partial x_j}(x_0)$$

$$= \left(\sum_{j=1}^{n} v_j \frac{\partial f_i}{\partial x_j}(x_0) \right)_{1 \le i \le m},$$

which we can rewrite as

$$L_{Df(x_0)}(v_j)_{1 \le j \le n}$$

where $Df(x_0)$ is the $m \times n$ matrix

$$Df(x_0) := \left(\frac{\partial f_i}{\partial x_j}(x_0) \right)_{1 \le i \le m; 1 \le j \le n}$$

$$= \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x_0) & \frac{\partial f_1}{\partial x_2}(x_0) & \ldots & \frac{\partial f_1}{\partial x_n}(x_0) \\ \frac{\partial f_2}{\partial x_1}(x_0) & \frac{\partial f_2}{\partial x_2}(x_0) & \ldots & \frac{\partial f_2}{\partial x_n}(x_0) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(x_0) & \frac{\partial f_m}{\partial x_2}(x_0) & \ldots & \frac{\partial f_m}{\partial x_n}(x_0) \end{pmatrix}.$$

Thus we have

$$(D_v f(x_0))^T = (f'(x_0)v)^T = Df(x_0)v^T.$$

The matrix $Df(x_0)$ is sometimes also called the *derivative matrix* or *differential matrix* of f at x_0 and is closely related to the total derivative $f'(x_0)$. One can also write Df as

$$Df(x_0) = \left(\frac{\partial f}{\partial x_1}(x_0)^T, \frac{\partial f}{\partial x_2}(x_0)^T, \ldots, \frac{\partial f}{\partial x_n}(x_0)^T \right),$$

i.e., each of the columns of $Df(x_0)$ is one of the partial derivatives of f, expressed as a column vector. Or one could write

$$Df(x_0) = \begin{pmatrix} \nabla f_1(x_0) \\ \nabla f_2(x_0) \\ \vdots \\ \nabla f_m(x_0) \end{pmatrix}$$

i.e., the rows of $Df(x_0)$ are the gradient of various components of f. In particular, if f is scalar-valued (i.e., $m = 1$), then Df is the same as ∇f.

Example 6.3.9 Let $f : \mathbf{R}^2 \to \mathbf{R}^2$ be the function $f(x, y) = (x^2 + xy, y^2)$. Then $\frac{\partial f}{\partial x} = (2x + y, 0)$ and $\frac{\partial f}{\partial y} = (x, 2y)$. Since these partial derivatives are continuous on $\mathbf{R}^2$, we see that f is differentiable on all of $\mathbf{R}^2$, and

$$Df(x, y) = \begin{pmatrix} 2x + y & x \\ 0 & 2y \end{pmatrix}.$$

Thus for instance, the directional derivative in the direction (v, w) is

$$D_{(v,w)}f(x, y) = ((2x + y)v + xw, 2yw).$$

— Exercise —

Exercise 6.3.1 Prove Lemma 6.3.5. (This will be similar to Exercise 6.2.1).

Exercise 6.3.2 Let E be a subset of $\mathbf{R}^n$, let $f : E \to \mathbf{R}^m$ be a function, let x_0 be an interior point of E, and let $1 \leq j \leq n$. Show that $\frac{\partial f}{\partial x_j}(x_0)$ exists if and only if $D_{e_j}f(x_0)$ and $D_{-e_j}f(x_0)$ exist and are negatives of each other (thus $D_{e_j}f(x_0) = -D_{-e_j}f(x_0)$); furthermore, one has $\frac{\partial f}{\partial x_j}(x_0) = D_{e_j}f(x_0)$ in this case.

Exercise 6.3.3 Let $f : \mathbf{R}^2 \to \mathbf{R}$ be the function defined by $f(x, y) := \frac{x^3}{x^2+y^2}$ when $(x, y) \neq (0, 0)$, and $f(0, 0) := 0$. Show that f is not differentiable at $(0, 0)$, despite being differentiable in every direction $v \in \mathbf{R}^2$ at $(0, 0)$. Explain why this does not contradict Theorem 6.3.8.

Exercise 6.3.4 Let $f : \mathbf{R}^n \to \mathbf{R}^m$ be a differentiable function such that $f'(x) = 0$ for all $x \in \mathbf{R}^n$. Show that f is constant. (*Hint*: you may use the mean value theorem or fundamental theorem of calculus for one-dimensional functions, but bear in mind that there is no direct analogue of these theorems for several variable functions. I would not advise proceeding via first principles.) For a tougher challenge, replace the domain $\mathbf{R}^n$ by an open connected subset Ω of $\mathbf{R}^n$.

6.4 The Several Variable Calculus Chain Rule

We are now ready to state the several variable calculus chain rule. Recall that if $f : X \to Y$ and $g : Y \to Z$ are two functions, then the composition $g \circ f : X \to Z$ is defined by $g \circ f(x) := g(f(x))$ for all $x \in X$.

Theorem 6.4.1 (Several variable calculus chain rule) *Let E be a subset of $\mathbf{R}^n$, and let F be a subset of $\mathbf{R}^m$. Let $f : E \to F$ be a function, and let $g : F \to \mathbf{R}^p$ be another function. Let x_0 be a point in the interior of E. Suppose that f is differentiable at x_0, and that $f(x_0)$ is in the interior of F. Suppose also that g is differentiable at $f(x_0)$. Then $g \circ f : E \to \mathbf{R}^p$ is also differentiable at x_0, and we have the formula*

$$(g \circ f)'(x_0) = g'(f(x_0))f'(x_0).$$

Proof See Exercise 6.4.3. □

One should compare this theorem with the single-variable chain rule, Theorem 10.1.15; indeed one can easily deduce the single-variable rule as a consequence of the several variable rule.

Intuitively, one can think of the several variable chain rule as follows. Let x be close to x_0. Then Newton's approximation asserts that

$$f(x) - f(x_0) \approx f'(x_0)(x - x_0)$$

and in particular $f(x)$ is close to $f(x_0)$. Since g is differentiable at $f(x_0)$, we see from Newton's approximation again that

$$g(f(x)) - g(f(x_0)) \approx g'(f(x_0))(f(x) - f(x_0)).$$

Combining the two, we obtain

$$g \circ f(x) - g \circ f(x_0) \approx g'(f(x_0))f'(x_0)(x - x_0)$$

which then should give $(g \circ f)'(x_0) = g'(f(x_0))f'(x_0)$. This argument however is rather imprecise; to make it more precise one needs to manipulate limits rigorously; see Exercise 6.4.3.

As a corollary of the chain rule and Lemma 6.1.16 (and Lemma 6.1.13), we see that

$$D(g \circ f)(x_0) = Dg(f(x_0))Df(x_0);$$

i.e., we can write the chain rule in terms of matrices and matrix multiplication, instead of in terms of linear transformations and composition.

Example 6.4.2 Let $f : \mathbf{R}^n \to \mathbf{R}$ and $g : \mathbf{R}^n \to \mathbf{R}$ be differentiable functions. We form the combined function $h : \mathbf{R}^n \to \mathbf{R}^2$ by defining $h(x) := (f(x), g(x))$. Now let $k : \mathbf{R}^2 \to \mathbf{R}$ be the multiplication function $k(a, b) := ab$. Note that

$$Dh(x_0) = \begin{pmatrix} \nabla f(x_0) \\ \nabla g(x_0) \end{pmatrix}$$

while

$$Dk(a, b) = (b, a)$$

(why?). By the chain rule, we thus see that

$$D(k \circ h)(x_0) = (g(x_0), f(x_0)) \begin{pmatrix} \nabla f(x_0) \\ \nabla g(x_0) \end{pmatrix} = g(x_0)\nabla f(x_0) + f(x_0)\nabla g(x_0).$$

But $k \circ h = fg$ (why?), and $D(fg) = \nabla(fg)$. We have thus proven the *product rule*

$$\nabla(fg) = g\nabla f + f\nabla g.$$

A similar argument gives the sum rule $\nabla(f + g) = \nabla f + \nabla g$, or the difference rule $\nabla(f - g) = \nabla f - \nabla g$, as well as the quotient rule (Exercise 6.4.4). As you can see, the several variable chain rule is quite powerful and can be used to deduce many other rules of differentiation.

We record one further useful application of the chain rule. Let $T : \mathbf{R}^n \to \mathbf{R}^m$ be a linear transformation. From Exercise 6.4.1 we observe that T is continuously differentiable at every point, and in fact $T'(x) = T$ for every x. (This equation may look a little strange, but perhaps it is easier to swallow if you view it in the form $\frac{\mathrm{d}}{\mathrm{d}x}(Tx) = T$.) Thus, for any differentiable function $f : E \to \mathbf{R}^n$, we see that $Tf : E \to \mathbf{R}^m$ is also differentiable, and hence by the chain rule

$$(Tf)'(x_0) = T(f'(x_0)).$$

This is a generalization of the single-variable calculus rule $(cf)' = c(f')$ for constant scalars c.

Another special case of the chain rule which is quite useful is the following: if $f : \mathbf{R}^n \to \mathbf{R}^m$ is some differentiable function, and $x_j : \mathbf{R} \to \mathbf{R}$ are differentiable functions for each $j = 1, \ldots n$, then

$$\frac{\mathrm{d}}{\mathrm{d}t} f(x_1(t), x_2(t), \ldots, x_n(t)) = \sum_{j=1}^{n} x_j'(t) \frac{\partial f}{\partial x_j}(x_1(t), x_2(t), \ldots, x_n(t)).$$

(Why is this a special case of the chain rule?).

— Exercise —

Exercise 6.4.1 Let $T : \mathbf{R}^n \to \mathbf{R}^m$ be a linear transformation. Show that T is continuously differentiable at every point, and in fact $T'(x) = T$ for every x. What is DT?

Exercise 6.4.2 Let E be a subset of $\mathbf{R}^n$. Prove that if a function $f : E \to \mathbf{R}^m$ is differentiable at an interior point x_0 of E, then it is also continuous at x_0. (*Hint:* use Exercise 6.1.4.)

Exercise 6.4.3 Prove Theorem 6.4.1. (*Hint:* you may wish to review the proof of the ordinary chain rule in single-variable calculus, Theorem 10.1.15. The easiest way to proceed is by using the sequence-based definition of limit (see Proposition 3.1.5(b)), and use Exercise 6.1.4.)

Exercise 6.4.4 State and prove some version of the quotient rule for functions of several variables (i.e., functions of the form $f : E \to \mathbf{R}$ for some subset E of $\mathbf{R}^n$). In other words, state a rule which gives a formula for the gradient of f/g; compare your answer with Theorem 10.1.13(h). Be sure to make clear what all your assumptions are.

Exercise 6.4.5 Let $\mathbf{x} : \mathbf{R} \to \mathbf{R}^3$ be a differentiable function, and let $r : \mathbf{R} \to \mathbf{R}$ be the function $r(t) := \|\mathbf{x}(t)\|$, where $\|\mathbf{x}\|$ denotes the length of $\mathbf{x}$ as measured in the usual l^2 metric. Let t_0 be a real number. Show that if $r(t_0) \neq 0$, then r is differentiable at t_0, and

$$r'(t_0) = \frac{\mathbf{x}'(t_0) \cdot \mathbf{x}(t_0)}{r(t_0)}.$$

(*Hint*: use Theorem 6.4.1.)

6.5 Double Derivatives and Clairaut's Theorem

We now investigate what happens if one differentiates a function twice.

Definition 6.5.1 (*Twice continuous differentiability*) Let E be an open subset of $\mathbf{R}^n$, and let $f : E \to \mathbf{R}^m$ be a function. We say that f is *twice continuously differentiable* if it is continuously differentiable, and the partial derivatives $\frac{\partial f}{\partial x_1}, \ldots, \frac{\partial f}{\partial x_n}$ are themselves continuously differentiable.

Remark 6.5.2 Continuously differentiable functions are sometimes called C^1 functions; twice continuously differentiable functions are sometimes called C^2 functions. One can also define C^3, C^4, etc., but we shall not do so here.

Example 6.5.3 Let $f : \mathbf{R}^2 \to \mathbf{R}^2$ be the function $f(x, y) = (x^2 + xy, y^2)$. Then f is continuously differentiable because the partial derivatives $\frac{\partial f}{\partial x}(x, y) = (2x + y, 0)$ and $\frac{\partial f}{\partial y}(x, y) = (x, 2y)$ exist and are continuous on all of $\mathbf{R}^2$. It is also twice continuously differentiable, because the double partial derivatives $\frac{\partial}{\partial x}\frac{\partial f}{\partial x}(x, y) = (2, 0)$, $\frac{\partial}{\partial y}\frac{\partial f}{\partial x}(x, y) = (1, 0)$, $\frac{\partial}{\partial x}\frac{\partial f}{\partial y}(x, y) = (1, 0)$, $\frac{\partial}{\partial y}\frac{\partial f}{\partial y}(x, y) = (0, 2)$ all exist and are continuous.

Observe in the above example that the double derivatives $\frac{\partial}{\partial y}\frac{\partial f}{\partial x}$ and $\frac{\partial}{\partial x}\frac{\partial f}{\partial y}$ are the same. This is in fact a general phenomenon:

Theorem 6.5.4 (Clairaut's theorem) *Let E be an open subset of $\mathbf{R}^n$, and let $f : E \to \mathbf{R}^m$ be a twice continuously differentiable function on E. Then we have $\frac{\partial}{\partial x_j}\frac{\partial f}{\partial x_i}(x_0) = \frac{\partial}{\partial x_i}\frac{\partial f}{\partial x_j}(x_0)$ for all $1 \le i, j \le n$.*

Proof By working with one component of f at a time we can assume that $m = 1$. The claim is trivial if $i = j$, so we shall assume that $i \ne j$. We shall prove the theorem for $x_0 = 0$; the general case is similar. (Actually, once one proves Clairaut's theorem for $x_0 = 0$, one can immediately obtain it for general x_0 by applying the theorem with $f(x)$ replaced by $f(x + x_0)$.)

Let a be the number $a := \frac{\partial}{\partial x_j}\frac{\partial f}{\partial x_i}(0)$, and a' denote the quantity $a' := \frac{\partial}{\partial x_i}\frac{\partial f}{\partial x_j}(0)$. Our task is to show that $a' = a$.

Let $\varepsilon > 0$. Because the double derivatives of f are continuous, we can find a $\delta > 0$ such that

$$\left| \frac{\partial}{\partial x_j}\frac{\partial f}{\partial x_i}(x) - a \right| \le \varepsilon$$

and

$$\left| \frac{\partial}{\partial x_i}\frac{\partial f}{\partial x_j}(x) - a' \right| \le \varepsilon$$

whenever $\|x\| \le 2\delta$.

Now we consider the quantity

$$X := f(\delta e_i + \delta e_j) - f(\delta e_i) - f(\delta e_j) + f(0).$$

From the fundamental theorem of calculus in the e_i variable, we have

$$f(\delta e_i + \delta e_j) - f(\delta e_j) = \int_0^\delta \frac{\partial f}{\partial x_i}(x_i e_i + \delta e_j)\,\mathrm{d}x_i$$

and

$$f(\delta e_i) - f(0) = \int_0^\delta \frac{\partial f}{\partial x_i}(x_i e_i)\,\mathrm{d}x_i$$

and hence

$$X = \int_0^\delta \left(\frac{\partial f}{\partial x_i}(x_i e_i + \delta e_j) - \frac{\partial f}{\partial x_i}(x_i e_i) \right) \mathrm{d}x_i.$$

But by the mean value theorem, for each x_i we have

$$\frac{\partial f}{\partial x_i}(x_i e_i + \delta e_j) - \frac{\partial f}{\partial x_i}(x_i e_i) = \delta \frac{\partial}{\partial x_j} \frac{\partial f}{\partial x_i}(x_i e_i + x_j e_j)$$

for some $0 \le x_j \le \delta$. By our construction of δ, we thus have

$$\left| \frac{\partial f}{\partial x_i}(x_i e_i + \delta e_j) - \frac{\partial f}{\partial x_i}(x_i e_i) - \delta a \right| \le \varepsilon \delta.$$

Integrating this from 0 to δ, we thus obtain

$$|X - \delta^2 a| \le \varepsilon \delta^2.$$

We can run the same argument with the rôle of i and j reversed (note that X is symmetric in i and j), to obtain

$$|X - \delta^2 a'| \le \varepsilon \delta^2.$$

From the triangle inequality we thus obtain

$$|\delta^2 a - \delta^2 a'| \le 2\varepsilon \delta^2,$$

and thus

$$|a - a'| \le 2\varepsilon.$$

But this is true for all $\varepsilon > 0$, and a and a' do not depend on ε, and so we must have $a = a'$, as desired. □

One should caution that Clairaut's theorem fails if we do not assume the double derivatives to be continuous; see Exercise 6.5.1.

— Exercise —

Exercise 6.5.1 Let $f : \mathbf{R}^2 \to \mathbf{R}$ be the function defined by $f(x, y) := \frac{xy^3}{x^2+y^2}$ when $(x, y) \neq (0, 0)$, and $f(0, 0) := 0$. Show that f is continuously differentiable, and the double derivatives $\frac{\partial}{\partial y}\frac{\partial f}{\partial x}$ and $\frac{\partial}{\partial x}\frac{\partial f}{\partial y}$ exist, but are not equal to each other at $(0, 0)$. Explain why this does not contradict Clairaut's theorem.

6.6 The Contraction Mapping Theorem

Before we turn to the next topic—namely the inverse function theorem—we need to develop a useful fact from the theory of complete metric spaces, namely the contraction mapping theorem.

Definition 6.6.1 (*Contraction*) Let (X, d) be a metric space, and let $f : X \to X$ be a map. We say that f is a *contraction* if we have $d(f(x), f(y)) \le d(x, y)$ for all

$x, y \in X$. We say that f is a *strict contraction* if there exists a constant $0 < c < 1$ such that $d(f(x), f(y)) \leq cd(x, y)$ for all $x, y \in X$; we call c the *contraction constant* of f.

Examples 6.6.2 The map $f : \mathbf{R} \to \mathbf{R}$ defined by $f(x) := x + 1$ is a contraction but not a strict contraction. The map $f : \mathbf{R} \to \mathbf{R}$ defined by $f(x) := x/2$ is a strict contraction. The map $f : [0, 1] \to [0, 1]$ defined by $f(x) := x - x^2$ is a contraction but not a strict contraction. (For justifications of these statements, see Exercise 6.6.5.)

Definition 6.6.3 (*Fixed points*) Let $f : X \to X$ be a map, and $x \in X$. We say that x is a *fixed point* of f if $f(x) = x$.

Contractions do not necessarily have any fixed points; for instance, the map $f : \mathbf{R} \to \mathbf{R}$ defined by $f(x) = x + 1$ does not. However, it turns out that *strict* contractions always do, at least when X is complete:

Theorem 6.6.4 (Contraction mapping theorem) *Let* (X, d) *be a metric space, and let* $f : X \to X$ *be a strict contraction. Then* f *can have at most one fixed point. Moreover, if we also assume that* X *is non-empty and complete, then* f *has* exactly *one fixed point.*

Proof See Exercise 6.6.7. □

Remark 6.6.5 The contraction mapping theorem is one example of a *fixed point theorem*—a theorem which guarantees, assuming certain conditions, that a map will have a fixed point. There are a number of other fixed point theorems which are also useful. One amusing one is the so-called *hairy ball theorem*, which (among other things) states that any continuous map $f : S^2 \to S^2$ from the sphere $S^2 := \{(x, y, z) \in \mathbf{R}^3 : x^2 + y^2 + z^2 = 1\}$ to itself, must contain either a fixed point, or an anti-fixed point (a point $x \in S^2$ such that $f(x) = -x$). A proof of this theorem can be found in any topology text; it is beyond the scope of this text.

We shall give one consequence of the contraction mapping theorem which is important for our application to the inverse function theorem. Basically, this says that any map f on a ball which is a "small" perturbation of the identity map, remains one-to-one and cannot create any internal holes in the ball.

Lemma 6.6.6 *Let* $B(0, r)$ *be a ball in* $\mathbf{R}^n$ *centered at the origin, and let* $g : B(0, r) \to \mathbf{R}^n$ *be a map such that* $g(0) = 0$ *and*

$$\|g(x) - g(y)\| \leq \frac{1}{2}\|x - y\|$$

for all $x, y \in B(0, r)$ *(here* $\|x\|$ *denotes the length of* x *in* $\mathbf{R}^n$*). Then the function* $f : B(0, r) \to \mathbf{R}^n$ *defined by* $f(x) := x + g(x)$ *is one-to-one, and furthermore the image* $f(B(0, r))$ *of this map contains the ball* $B(0, r/2)$.

Proof We first show that f is one-to-one. Suppose for sake of contradiction that we had two different points $x, y \in B(0, r)$ such that $f(x) = f(y)$. But then we would have $x + g(x) = y + g(y)$, and hence

$$\|g(x) - g(y)\| = \|x - y\|.$$

The only way this can be consistent with our hypothesis $\|g(x) - g(y)\| \leq \frac{1}{2}\|x - y\|$ is if $\|x - y\| = 0$, i.e., if $x = y$, a contradiction. Thus f is one-to-one.

Now we show that $f(B(0, r))$ contains $B(0, r/2)$. Let y be any point in $B(0, r/2)$; our objective is to find a point $x \in B(0, r)$ such that $f(x) = y$, or in other words that $x = y - g(x)$. So the problem is now to find a fixed point of the map $x \mapsto y - g(x)$.

Let $F: B(0, r) \to B(0, r)$ denote the function $F(x) := y - g(x)$. Observe that if $x \in B(0, r)$, then

$$\|F(x)\| \leq \|y\| + \|g(x)\| \leq \frac{r}{2} + \|g(x) - g(0)\| \leq \frac{r}{2} + \frac{1}{2}\|x - 0\| < \frac{r}{2} + \frac{r}{2} = r,$$

so F does indeed map $B(0, r)$ to itself. The same argument shows that for a sufficiently small $\varepsilon > 0$, F maps the closed ball $\overline{B(0, r - \varepsilon)}$ to itself. Also, for any x, x' in $B(0, r)$ we have

$$\|F(x) - F(x')\| = \|g(x') - g(x)\| \leq \frac{1}{2}\|x' - x\|$$

so F is a strict contraction on $B(0, r)$, and hence on the complete space $\overline{B(0, r - \varepsilon)}$. By the contraction mapping theorem, F has a fixed point, i.e., there exists an x such that $x = y - g(x)$. But this means that $f(x) = y$, as desired. □

— Exercise —

Exercise 6.6.1 Let $f: [a, b] \to [a, b]$ be a differentiable function of one variable such that $|f'(x)| \leq 1$ for all $x \in [a, b]$. Prove that f is a contraction. (*Hint:* use the mean value theorem, Corollary 10.2.9.) If in addition $|f'(x)| < 1$ for all $x \in [a, b]$ and f' is continuous, show that f is a strict contraction.

Exercise 6.6.2 Show that if $f: [a, b] \to \mathbf{R}$ is differentiable and is a contraction, then $|f'(x)| \leq 1$.

Exercise 6.6.3 Give an example of a function $f: [a, b] \to \mathbf{R}$ which is continuously differentiable and such that $|f(x) - f(y)| < |x - y|$ for all distinct $x, y \in [a, b]$, but such that $|f'(x)| = 1$ for at least one value of $x \in [a, b]$.

Exercise 6.6.4 Given an example of a function $f: [a, b] \to \mathbf{R}$ which is a strict contraction but which is not differentiable for at least one point x in $[a, b]$.

Exercise 6.6.5 Verify the claims in Examples 6.6.2.

Exercise 6.6.6 Show that every contraction on a metric space X is necessarily continuous.

Exercise 6.6.7 Prove Theorem 6.6.4. (*Hint*: to prove that there is at most one fixed point, argue by contradiction. To prove that there is at least one fixed point, pick any $x_0 \in X$ and define recursively $x_1 = f(x_0)$, $x_2 = f(x_1)$, $x_3 = f(x_2)$, etc. Prove inductively that $d(x_{n+1}, x_n) \leq c^n d(x_1, x_0)$, and conclude (using the geometric series formula, Lemma 7.3.3) that the sequence $(x_n)_{n=0}^{\infty}$ is a Cauchy sequence. Then prove that the limit of this sequence is a fixed point of f.)

Exercise 6.6.8 Let (X, d) be a complete metric space, and let $f: X \to X$ and $g: X \to X$ be two strict contractions on X with contraction coefficients c and c', respectively. From Theorem 6.6.4 we know that f has some fixed point x_0, and g has some fixed point y_0. Suppose we know that there is an $\varepsilon > 0$ such that $d(f(x), g(x)) \leq \varepsilon$ for all $x \in X$ (i.e., f and g are within ε of each other in the uniform metric). Show that $d(x_0, y_0) \leq \varepsilon/(1 - \min(c, c'))$. Thus nearby contractions have nearby fixed points.

6.7 The Inverse Function Theorem in Several Variable Calculus

We recall the inverse function theorem in single-variable calculus (Theorem 10.4.2), which asserts that if a function $f: \mathbf{R} \to \mathbf{R}$ is invertible, differentiable, and $f'(x_0)$ is nonzero, then f^{-1} is differentiable at $f(x_0)$, and

$$(f^{-1})'(f(x_0)) = \frac{1}{f'(x_0)}.$$

In fact, one can say something even when f' is not invertible, as long as we know that f is *continuously* differentiable. If $f'(x_0)$ is nonzero, then $f'(x_0)$ must be either strictly positive or strictly negative, which implies (since we are assuming f' to be continuous) that $f'(x)$ is either strictly positive for x near x_0, or strictly negative for x near x_0. In particular, f must be either strictly increasing near x_0, or strictly decreasing near x_0. In either case, f will become invertible if we restrict the domain and codomain of f to be sufficiently close to x_0 and to $f(x_0)$, respectively. (The technical terminology for this is that f is *locally invertible near* x_0.)

The requirement that f be continuously differentiable is important; see Exercise 6.7.1.

It turns out that a similar theorem is true for functions $f: \mathbf{R}^n \to \mathbf{R}^n$ from one Euclidean space to the same space. However, the condition that $f'(x_0)$ is nonzero must be replaced with a slightly different one, namely that $f'(x_0)$ is *invertible*. We first remark that the inverse of a linear transformation is also linear:

Lemma 6.7.1 *Let $T: \mathbf{R}^n \to \mathbf{R}^n$ be a linear transformation which is also invertible. Then the inverse transformation $T^{-1}: \mathbf{R}^n \to \mathbf{R}^n$ is also linear.*

Proof See Exercise 6.7.2. □

We can now prove an important and useful theorem, arguably one of the most important theorems in several variable differential calculus.

Theorem 6.7.2 (Inverse function theorem) *Let E be an open subset of $\mathbf{R}^n$, and let $f : E \to \mathbf{R}^n$ be a function which is continuously differentiable on E. Suppose $x_0 \in E$ is such that the linear transformation $f'(x_0) : \mathbf{R}^n \to \mathbf{R}^n$ is invertible. Then there exists an open set U in E containing x_0, and an open set V in $\mathbf{R}^n$ containing $f(x_0)$, such that f is a bijection from U to V. In particular, there is an inverse map $f^{-1} : V \to U$. Furthermore, this inverse map is differentiable at $f(x_0)$, and*

$$(f^{-1})'(f(x_0)) = (f'(x_0))^{-1}.$$

Proof We first observe that once we know the inverse map f^{-1} is differentiable, the formula $(f^{-1})'(f(x_0)) = (f'(x_0))^{-1}$ is automatic. This comes from starting with the identity

$$I = f^{-1} \circ f$$

on U, where $I : \mathbf{R}^n \to \mathbf{R}^n$ is the identity map $Ix := x$, and then differentiating both sides using the chain rule at x_0 to obtain

$$I'(x_0) = (f^{-1})'(f(x_0))f'(x_0).$$

Since $I'(x_0) = I$, we thus have $(f^{-1})'(f(x_0)) = (f'(x_0))^{-1}$ as desired.

We remark that this argument shows that if $f'(x_0)$ is *not* invertible, then there is no way that an inverse f^{-1} can exist and be differentiable at $f(x_0)$.

Next, we observe that it suffices to prove the theorem under the additional assumption $f(x_0) = 0$. The general case then follows from the special case by replacing f by a new function $\tilde{f}(x) := f(x) - f(x_0)$ and then applying the special case to $\tilde{f}$ (note that V will have to shift by $f(x_0)$). Note that $f^{-1}(y) = \tilde{f}^{-1}(y - f(x_0))$ (why?). Henceforth we will always assume $f(x_0) = 0$.

In a similar manner, one can make the assumption $x_0 = 0$. The general case then follows from this case by replacing f by a new function $\tilde{f}(x) := f(x + x_0)$ and applying the special case to $\tilde{f}$ (note that E and U will have to shift by x_0). Note that $f^{-1}(y) = \tilde{f}^{-1}(y) + x_0$ - why? Henceforth we will always assume $x_0 = 0$. Thus we now have that $f(0) = 0$ and that $f'(0)$ is invertible.

Finally, one can assume that $f'(0) = I$, where $I : \mathbf{R}^n \to \mathbf{R}^n$ is the identity transformation $Ix = x$. The general case then follows from this case by replacing f with a new function $\tilde{f} : E \to \mathbf{R}^n$ defined by $\tilde{f}(x) := f'(0)^{-1} f(x)$, and applying the special case to this case. Note from Lemma 6.7.1 that $f'(0)^{-1}$ is a linear transformation. In particular, we note that $\tilde{f}(0) = 0$ and that

$$\tilde{f}'(0) = f'(0)^{-1} f'(0) = I,$$

so by the special case of the inverse function theorem we know that there exists an open set U' containing 0, and an open set V' containing 0, such that $\tilde{f}$ is a bijection from U' to V', and that $\tilde{f}^{-1} : V' \to U'$ is differentiable at 0 with derivative I. But we have $f(x) = f'(0)\tilde{f}(x)$, and hence f is a bijection from U' to $f'(0)(V')$ (note that $f'(0)$ is also a bijection). Since $f'(0)$ and its inverse are both continuous, $f'(0)(V')$ is open, and it certainly contains 0. Now consider the inverse function $f^{-1} : f'(0)(V') \to U'$. Since $f(x) = f'(0)\tilde{f}(x)$, we see that $f^{-1}(y) = \tilde{f}^{-1}(f'(0)^{-1}y)$ for all $y \in f'(0)(V')$ (why? use the fact that $\tilde{f}$ is a bijection from U' to V'). In particular we see that f^{-1} is differentiable at 0.

So all we have to do now is prove the inverse function theorem in the special case, when $x_0 = 0$, $f(x_0) = 0$, and $f'(x_0) = I$. Let $g: E \to \mathbf{R}^n$ denote the function $g(x) := f(x) - x$. Then $g(0) = 0$ and $g'(0) = 0$. In particular

$$\frac{\partial g}{\partial x_j}(0) = 0$$

for $j = 1, \ldots, n$. Since g is continuously differentiable, there thus exists a ball $B(0, r)$ in E such that

$$\left\| \frac{\partial g}{\partial x_j}(x) \right\| \leq \frac{1}{2n^2}$$

for all $x \in B(0, r)$. (There is nothing particularly special about $\frac{1}{2n^2}$, we just need a nice small number here.) In particular, for any $x \in B(0, r)$ and $v = (v_1, \ldots, v_n)$ we have

$$\|D_v g(x)\| = \left\| \sum_{j=1}^{n} v_j \frac{\partial g}{\partial x_j}(x) \right\|$$

$$\leq \sum_{j=1}^{n} |v_j| \left\| \frac{\partial g}{\partial x_j}(x) \right\|$$

$$\leq \sum_{j=1}^{n} \|v\| \frac{1}{2n^2} \leq \frac{1}{2n} \|v\|.$$

But now for any $x, y \in B(0, r)$, we have by the fundamental theorem of calculus

$$g(y) - g(x) = \int_0^1 \frac{\mathrm{d}}{\mathrm{d}t} g(x + t(y - x))\, \mathrm{d}t$$

$$= \int_0^1 D_{y-x} g(x + t(y - x))\, \mathrm{d}t,$$

where the integral of a vector-valued function is defined by integrating each component separately. By the previous remark, the vectors $D_{y-x}g(x+t(y-x))$ have a magnitude of at most $\frac{1}{2n}\|y-x\|$. Thus every component of these vectors has magnitude at most $\frac{1}{2n}\|y-x\|$. Thus every component of $g(y)-g(x)$ has magnitude at most $\frac{1}{2n}\|y-x\|$, and hence $g(y)-g(x)$ itself has magnitude at most $\frac{1}{2}\|y-x\|$ (actually, it will be substantially less than this, but this bound will be enough for our purposes). In other words, g is a contraction. By Lemma 6.6.6, the map $f = g + I$ is thus one-to-one on $B(0, r)$, and the image $f(B(0, r))$ contains $B(0, r/2)$. In particular we have an inverse map $f^{-1} : B(0, r/2) \to B(0, r)$ defined on $B(0, r/2)$.

Applying the contraction bound with $y = 0$ we obtain in particular that

$$\|g(x)\| \leq \frac{1}{2}\|x\|$$

for all $x \in B(0, r)$, and so by the triangle inequality

$$\frac{1}{2}\|x\| \leq \|f(x)\| \leq \frac{3}{2}\|x\|$$

for all $x \in B(0, r)$.

Now we set $V{:=}B(0, r/2)$ and $U{:=}f^{-1}(V) \cap B(0, r)$. Then by construction f is a bijection from U to V. V is clearly open, and U is also open since f is continuous. (Notice that if a set is open relative to $B(0, r)$, then it is open in $\mathbf{R}^n$ as well.) Now we want to show that $f^{-1} : V \to U$ is differentiable at 0 with derivative $I^{-1} = I$. In other words, we wish to show that

$$\lim_{x\to 0; x\in V\backslash\{0\}} \frac{\|f^{-1}(x) - f^{-1}(0) - I(x-0)\|}{\|x\|} = 0.$$

Since $f(0) = 0$, we have $f^{-1}(0) = 0$, and the above simplifies to

$$\lim_{x\to 0; x\in V\backslash\{0\}} \frac{\|f^{-1}(x) - x\|}{\|x\|} = 0.$$

Let $(x_n)_{n=1}^{\infty}$ be any sequence in $V\backslash\{0\}$ that converges to 0. By Proposition 3.1.5(b), it suffices to show that

$$\lim_{n\to\infty} \frac{\|f^{-1}(x_n) - x_n\|}{\|x_n\|} = 0.$$

Write $y_n{:=}f^{-1}(x_n)$. Then $y_n \in B(0, r)$ and $x_n = f(y_n)$. In particular we have

$$\frac{1}{2}\|y_n\| \leq \|x_n\| \leq \frac{3}{2}\|y_n\|$$

and so since $\|x_n\|$ goes to 0, $\|y_n\|$ goes to zero also, and their ratio remains bounded. It will thus suffice to show that

$$\lim_{n\to\infty} \frac{\|y_n - f(y_n)\|}{\|y_n\|} = 0.$$

But since y_n is going to 0, and f is differentiable at 0, we have

$$\lim_{n\to\infty} \frac{\|f(y_n) - f(0) - f'(0)(y_n - 0)\|}{\|y_n\|} = 0$$

as desired (since $f(0) = 0$ and $f'(0) = I$). □

The inverse function theorem gives a useful criterion for when a function is (locally) invertible at a point x_0 - all we need is for its derivative $f'(x_0)$ to be invertible (and then we even get further information, for instance we can compute the derivative of f^{-1} at $f(x_0)$). Of course, this begs the question of how one can tell whether the linear transformation $f'(x_0)$ is invertible or not. Recall that we have $f'(x_0) = L_{Df(x_0)}$, so by Lemmas 6.1.13 and 6.1.16 we see that the linear transformation $f'(x_0)$ is invertible if and only if the matrix $Df(x_0)$ is. There are many ways to check whether a matrix such as $Df(x_0)$ is invertible; for instance, one can use determinants, or alternatively Gaussian elimination methods. We will not pursue this matter here, but refer the reader to any linear algebra text.

If $f'(x_0)$ exists but is non-invertible, then the inverse function theorem does not apply. In such a situation it is not possible for f^{-1} to exist and be differentiable at $f(x_0)$; this was remarked in the above proof. But it is still possible for f to be invertible. For instance, the single-variable function $f : \mathbf{R} \to \mathbf{R}$ defined by $f(x) = x^3$ is invertible despite $f'(0)$ not being invertible.

— Exercise —

Exercise 6.7.1 Let $f : \mathbf{R} \to \mathbf{R}$ be the function defined by $f(x) := x + x^2 \sin(1/x^4)$ for $x \neq 0$ and $f(0) := 0$. Show that f is differentiable and $f'(0) = 1$, but f is not increasing on any open set containing 0 (*Hint:* show that the derivative of f can turn negative arbitrarily close to 0. Drawing a graph of f may aid your intuition.)

Exercise 6.7.2 Prove Lemma 6.7.1.

Exercise 6.7.3 Let $f : \mathbf{R}^n \to \mathbf{R}^n$ be a continuously differentiable function such that $f'(x)$ is an invertible linear transformation for every $x \in \mathbf{R}^n$. Show that whenever V is an open set in $\mathbf{R}^n$, that $f(V)$ is also open. (*Hint*: use the inverse function theorem.)

Exercise 6.7.4 Let the notation and hypotheses be as in Theorem 6.7.2. Show that after shrinking the open sets U, V as necessarily (while still keeping x_0 in U and $f(x_0)$ in V), the derivative map $f'(x)$ is invertible for all $x \in U$, and that the inverse map f^{-1} is differentiable at every point of V with $(f^{-1})'(f(x) = (f'(x))^{-1}$ for all $x \in U$. Finally, show that f^{-1} is continuously differentiable on V.

6.8 The Implicit Function Theorem

Recall (from Exercise 3.5.10) that a function $f : \mathbf{R} \to \mathbf{R}$ gives rise to a *graph*

$$\{(x, f(x)) : x \in \mathbf{R}\}$$

which is a subset of $\mathbf{R}^2$, usually looking like a curve. However, not all curves are graphs, they must obey the *vertical line test*, that for every x there is exactly one y such that (x, y) is in the curve. For instance, the circle $\{(x, y) \in \mathbf{R}^2 : x^2 + y^2 = 1\}$ is not a graph, although if one restricts to a semicircle such as $\{(x, y) \in \mathbf{R}^2 : x^2 + y^2 = 1, y > 0\}$ then one again obtains a graph. Thus while the entire circle is not a graph, certain local portions of it are. (The portions of the circle near $(1, 0)$ and $(-1, 0)$ are not graphs over the variable x, but they are graphs over the variable y).

Similarly, any function $g : \mathbf{R}^n \to \mathbf{R}$ gives rise to a graph $\{(x, g(x)) : x \in \mathbf{R}^n\}$ in $\mathbf{R}^{n+1}$, which in general looks like some sort of n-dimensional surface in $\mathbf{R}^{n+1}$ (the technical term for this is a *hypersurface*). Conversely, one may ask which hypersurfaces are actually graphs of some function, and whether that function is continuous or differentiable.

If the hypersurface is given geometrically, then one can again invoke the vertical line test to work out whether it is a graph or not. But what if the hypersurface is given algebraically, for instance the surface $\{(x, y, z) \in \mathbf{R}^3 : xy + yz + zx = -1\}$? Or more generally, a hypersurface of the form $\{x \in \mathbf{R}^n : g(x) = 0\}$, where $g : \mathbf{R}^n \to \mathbf{R}$ is some function? In this case, it is still possible to say whether the hypersurface is a graph, locally at least, by means of the *implicit function theorem*.

Theorem 6.8.1 (Implicit function theorem) *Let E be an open subset of $\mathbf{R}^n$, let $f : E \to \mathbf{R}$ be continuously differentiable, and let $y = (y_1, \ldots, y_n)$ be a point in E such that $f(y) = 0$ and $\frac{\partial f}{\partial x_n}(y) \neq 0$. Then there exists an open subset U of $\mathbf{R}^{n-1}$ containing $(y_1, \ldots, y_{n-1})$, an open subset V of E containing y, and a function $g : U \to \mathbf{R}$ such that $g(y_1, \ldots, y_{n-1}) = y_n$, and*

$$\{(x_1, \ldots, x_n) \in V : f(x_1, \ldots, x_n) = 0\}$$

$$= \{(x_1, \ldots, x_{n-1}, g(x_1, \ldots, x_{n-1})) : (x_1, \ldots, x_{n-1}) \in U\}.$$

In other words, the set $\{x \in V : f(x) = 0\}$ is a graph of a function over U. Moreover, g is differentiable at $(y_1, \ldots, y_{n-1})$, and we have

$$\frac{\partial g}{\partial x_j}(y_1, \ldots, y_{n-1}) = -\frac{\partial f}{\partial x_j}(y) / \frac{\partial f}{\partial x_n}(y) \qquad (6.1)$$

for all $1 \le j \le n-1$.

Remark 6.8.2 Equation (6.1) is sometimes derived using *implicit differentiation*. Basically, the point is that if you know that

$$f(x_1, \dots, x_n) = 0$$

then (as long as $\frac{\partial f}{\partial x_n} \neq 0$) the variable x_n is "implicitly" defined in terms of the other $n - 1$ variables, and one can differentiate the above identity in, say, the x_j direction using the chain rule to obtain

$$\frac{\partial f}{\partial x_j} + \frac{\partial f}{\partial x_n} \frac{\partial x_n}{\partial x_j} = 0$$

which is (6.1) in disguise (we are using g to represent the implicit function defining x_n in terms of $x_1, \dots, x_n$). Thus, the implicit function theorem allows one to define a dependence implicitly, by means of a constraint rather than by a direct formula of the form $x_n = g(x_1, \dots, x_{n-1})$.

Proof This theorem looks somewhat fearsome, but actually it is a fairly quick consequence of the inverse function theorem. Let $F : E \to \mathbf{R}^n$ be the function

$$F(x_1, \dots, x_n) := (x_1, \dots, x_{n-1}, f(x_1, \dots, x_n)).$$

This function is continuously differentiable. Also note that

$$F(y) = (y_1, \dots, y_{n-1}, 0)$$

and

$$DF(y) = \left(\frac{\partial F}{\partial x_1}(y)^T, \frac{\partial F}{\partial x_2}(y)^T, \dots, \frac{\partial F}{\partial x_n}(y)^T \right)$$

$$= \begin{pmatrix} 1 & 0 & \dots 0 & 0 \\ 0 & 1 & \dots 0 & 0 \\ \vdots & \vdots & \ddots \vdots & \vdots \\ 0 & 0 & \dots 1 & 0 \\ \frac{\partial f}{\partial x_1}(y) & \frac{\partial f}{\partial x_2}(y) & \dots \frac{\partial f}{\partial x_{n-1}}(y) & \frac{\partial f}{\partial x_n}(y) \end{pmatrix}.$$

Since $\frac{\partial f}{\partial x_n}(y)$ is assumed by hypothesis to be nonzero, this matrix is invertible; this can be seen either by computing the determinant, or using row reduction, or by computing the inverse explicitly, which is

$$DF(y)^{-1} = \begin{pmatrix} 1 & 0 & \dots 0 & 0 \\ 0 & 1 & \dots 0 & 0 \\ \vdots & \vdots & \ddots \vdots & \vdots \\ 0 & 0 & \dots 1 & 0 \\ -\frac{\partial f}{\partial x_1}(y)/a & -\frac{\partial f}{\partial x_2}(y)/a & \dots -\frac{\partial f}{\partial x_{n-1}}(y)/a & 1/a \end{pmatrix},$$

where we have written $a = \frac{\partial f}{\partial x_n}(y)$ for short. Thus the inverse function theorem applies, and we can find an open set V in E containing y, and an open set W in $\mathbf{R}^n$ containing $F(y) = (y_1, \ldots, y_{n-1}, 0)$, such that F is a bijection from V to W, and that F^{-1} is differentiable at $(y_1, \ldots, y_{n-1}, 0)$.

Let us write F^{-1} in co-ordinates as

$$F^{-1}(x) = (h_1(x), h_2(x), \ldots, h_n(x))$$

where $x \in W$. Since $F(F^{-1}(x)) = x$, we have $h_j(x_1, \ldots, x_n) = x_j$ for all $1 \leq j \leq n-1$ and $x \in W$, and

$$f(x_1, \ldots, x_{n-1}, h_n(x_1, \ldots, x_n)) = x_n.$$

Also, h_n is differentiable at $(y_1, \ldots, y_{n-1}, 0)$ since F^{-1} is.

Now we set $U := \{(x_1, \ldots, x_{n-1}) \in \mathbf{R}^{n-1} : (x_1, \ldots, x_{n-1}, 0) \in W\}$. Note that U is open and contains $(y_1, \ldots, y_{n-1})$. Now we define $g\colon U \to \mathbf{R}$ by $g(x_1, \ldots, x_{n-1}) := h_n(x_1, \ldots, x_{n-1}, 0)$. Then g is differentiable at $(y_1, \ldots, y_{n-1})$. Now we prove that

$$\{(x_1, \ldots, x_n) \in V : f(x_1, \ldots, x_n) = 0\}$$

$$= \{(x_1, \ldots, x_{n-1}, g(x_1, \ldots, x_{n-1})) : (x_1, \ldots, x_{n-1}) \in U\}.$$

First suppose that $(x_1, \ldots, x_n) \in V$ and $f(x_1, \ldots, x_n) = 0$. Then we have $F(x_1, \ldots, x_n) = (x_1, \ldots, x_{n-1}, 0)$, which lies in W. Thus $(x_1, \ldots, x_{n-1})$ lies in U. Applying F^{-1}, we see that $(x_1, \ldots, x_n) = F^{-1}(x_1, \ldots, x_{n-1}, 0)$. In particular $x_n = h_n(x_1, \ldots, x_{n-1}, 0)$, and hence $x_n = g(x_1, \ldots, x_{n-1})$. Thus every element of the left-hand set lies in the right-hand set. The reverse inclusion comes by reversing all the above steps and is left to the reader.

Finally, we show the formula for the partial derivatives of g. From the preceding discussion we have

$$f(x_1, \ldots, x_{n-1}, g(x_1, \ldots, x_{n-1})) = 0$$

for all $(x_1, \ldots, x_{n-1}) \in U$. Since g is differentiable at $(y_1, \ldots, y_{n-1})$, and f is differentiable at $(y_1, \ldots, y_{n-1}, g(y_1, \ldots, y_{n-1})) = y$, we may use the chain rule, differentiating in x_j, to obtain

$$\frac{\partial f}{\partial x_j}(y) + \frac{\partial f}{\partial x_n}(y)\frac{\partial g}{\partial x_j}(y_1, \ldots, y_{n-1}) = 0$$

and the claim follows by simple algebra. □

Example 6.8.3 Consider the surface $S := \{(x, y, z) \in \mathbf{R}^3 : xy + yz + zx = -1\}$, which we rewrite as $\{(x, y, z) \in \mathbf{R}^3 : f(x, y, z) = 0\}$, where $f\colon \mathbf{R}^3 \to \mathbf{R}$ is the function $f(x, y, z) := xy + yz + zx + 1$. Clearly f is continuously differentiable, and $\frac{\partial f}{\partial z} = y + x$. Thus for any (x_0, y_0, z_0) in S with $y_0 + x_0 \neq 0$, one can write this surface (near (x_0, y_0, z_0)) as a graph of the form $\{(x, y, g(x, y)) : (x, y) \in U\}$ for some open set U containing (x_0, y_0), and some function g which is differentiable at (x_0, y_0). Indeed one can implicitly differentiate to obtain that

$$\frac{\partial g}{\partial x}(x_0, y_0) = -\frac{y_0 + z_0}{y_0 + x_0} \text{ and } \frac{\partial g}{\partial y}(x_0, y_0) = -\frac{x_0 + z_0}{y_0 + x_0}.$$

In the implicit function theorem, if the derivative $\frac{\partial f}{\partial x_n}$ equals zero at some point, then it is unlikely that the set $\{x \in \mathbf{R}^n : f(x) = 0\}$ can be written as a graph of the x_n variable in terms of the other $n - 1$ variables near that point. However, if some other derivative $\frac{\partial f}{\partial x_j}$ is nonzero, then it would be possible to write the x_j variable in terms of the other $n - 1$ variables, by a variant of the implicit function theorem. Thus as long as the gradient ∇f is not entirely zero, one can write this set $\{x \in \mathbf{R}^n : f(x) = 0\}$ as a graph of *some* variable x_j in terms of the other $n - 1$ variables. (The circle $\{(x, y) \in \mathbf{R}^2 : x^2 + y^2 - 1 = 0\}$ is a good example of this; it is not a graph of y in terms of x, or x in terms of y, but near every point it is one of the two. And this is because the gradient of $x^2 + y^2 - 1$ is never zero on the circle.) However, if ∇f does vanish at some point x_0, then we say that f has a *critical point* at x_0 and the behavior there is much more complicated. For instance, the set $\{(x, y) \in \mathbf{R}^2 : x^2 - y^2 = 0\}$ has a critical point at $(0, 0)$ and there the set does not look like a graph of any sort (it is the union of two lines).

Remark 6.8.4 Sets which look like graphs of continuous functions at every point have a name, they are called *manifolds*. Thus $\{x \in \mathbf{R}^n : f(x) = 0\}$ will be a manifold if it contains no critical points of f. The theory of manifolds is very important in modern geometry (especially differential geometry and algebraic geometry), but we will not discuss it here as it is a graduate level topic.

— Exercise —

Exercise 6.8.1 Let the notation and hypotheses be as in Theorem 6.8.1. Show that, after shrinking the open sets U, V as necessary, that the function g becomes continuously differentiable on all of U, and the Eq. (6.1) holds at all points of U.

Chapter 7
Lebesgue Measure

In the previous chapter we discussed differentiation in several variable calculus. It is now only natural to consider the question of integration in several variable calculus. The general question we wish to answer is this: given some subset Ω of $\mathbf{R}^n$, and some real-valued function $f : \Omega \to \mathbf{R}$, is it possible to integrate f on Ω to obtain some number $\int_\Omega f$? (It is possible to consider other types of functions, such as complex-valued or vector-valued functions, but this turns out not to be too difficult once one knows how to integrate real-valued functions, since one can integrate a complex or vector-valued function, by integrating each real-valued component of that function separately.)

In one dimension we already have developed (in Chap. 11) the notion of a *Riemann integral* $\int_{[a,b]} f$, which answers this question when Ω is an interval $\Omega = [a, b]$, and f is *Riemann integrable*. Exactly what Riemann integrability means is not important here, but let us just remark that every piecewise continuous function is Riemann integrable, and in particular every piecewise constant function is Riemann integrable. However, not all functions are Riemann integrable. It is possible to extend this notion of a Riemann integral to higher dimensions, but it requires quite a bit of effort and one can still only integrate "Riemann integrable" functions, which turn out to be a rather unsatisfactorily small class of functions. (For instance, the pointwise limit of Riemann integrable functions need not be Riemann integrable, and the same goes for an L^2 limit, although we have already seen that uniform limits of Riemann integrable functions remain Riemann integrable.)

Because of this, we must look beyond the Riemann integral to obtain a truly satisfactory notion of integration, one that can handle even very discontinuous functions. This leads to the notion of the *Lebesgue integral*, which we shall spend this chapter and the next constructing. The Lebesgue integral can handle a very large class of functions, including all the Riemann integrable functions but also many others as well; in fact, it is safe to say that it can integrate virtually any function that one actually needs in mathematics, at least if one works on Euclidean spaces and everything is absolutely integrable. (If one assumes the axiom of choice, then there are still some

T. Tao, *Analysis II*, Texts and Readings in Mathematics,
https://doi.org/10.1007/978-981-19-7284-3_7

pathological functions one can construct which cannot be integrated by the Lebesgue integral, but these functions will not come up in real-life applications.)

Before we turn to the details, we begin with an informal discussion. In order to understand how to compute an integral $\int_\Omega f$, we must first understand a more basic and fundamental question: how does one compute the *length/area/volume* of Ω? To see why this question is connected to that of integration, observe that if one integrates the function 1 on the set Ω, then one should obtain the length of Ω (if Ω is one-dimensional), the area of Ω (if Ω is two-dimensional), or the volume of Ω (if Ω is three-dimensional). To avoid splitting into cases depending on the dimension, we shall refer to the *measure* of Ω as either the length, area, volume, (or hypervolume, etc.) of Ω, depending on what Euclidean space $\mathbf{R}^n$ we are working in.

Ideally, to every subset Ω of $\mathbf{R}^n$ we would like to associate a non-negative number $m(\Omega)$, which will be the measure of Ω (i.e., the length, area, volume, etc.). We allow the possibility for $m(\Omega)$ to be zero (e.g., if Ω is just a single point or the empty set) or for $m(\Omega)$ to be infinite (e.g., if Ω is all of $\mathbf{R}^n$). This measure should obey certain reasonable properties; for instance, the measure of the unit cube $(0, 1)^n := \{(x_1, \ldots, x_n) : 0 < x_i < 1\}$ should equal 1, we should have $m(A \cup B) = m(A) + m(B)$ if A and B are disjoint (and similarly that $m\left(\bigcup_{n=1}^{\infty} A_n\right) = \sum_{n=1}^{\infty} m(A_n)$ when the A_n are disjoint), we should have $m(A) \leq m(B)$ whenever $A \subseteq B$, and we should have $m(x + A) = m(A)$ for any $x \in \mathbf{R}^n$ (i.e., if we shift A by the vector x the measure should be the same).

Remarkably, it turns out that such a measure *does not exist*; one cannot assign a non-negative number to *every* subset of $\mathbf{R}^n$ which has the above properties. This is quite a surprising fact, as it goes against one's intuitive concept of volume; we shall prove it later in these notes. (An even more dramatic example of this failure of intuition is the *Banach-Tarski paradox*, in which a unit ball in $\mathbf{R}^3$ is decomposed into five pieces, and then the five pieces are reassembled via translations and rotations to form two complete and disjoint unit balls, thus violating any concept of conservation of volume; however we will not discuss this paradox here.)

What these paradoxes mean is that it is impossible to find a reasonable way to assign a measure to every single subset of $\mathbf{R}^n$. However, we can salvage matters by only measuring a certain class of sets in $\mathbf{R}^n$—the *measurable sets*. These are the only sets Ω for which we will define the measure $m(\Omega)$, and once one restricts one's attention to measurable sets, one recovers all the above properties again. Furthermore, almost all the sets one encounters in real life are measurable (e.g., all open and closed sets will be measurable), and so this turns out to be good enough to do analysis.

7.1 The Goal: Lebesgue Measure

Let $\mathbf{R}^n$ be a Euclidean space. Our goal in this chapter is to define a concept of *measurable set*, which will be a special kind of subset of $\mathbf{R}^n$, and for every such measurable set $\Omega \subseteq \mathbf{R}^n$, we will define the *Lebesgue measure* $m(\Omega)$ to be a certain number in $[0, \infty]$. The concept of measurable set will obey the following properties:

(i) (Borel property) Every open set in $\mathbf{R}^n$ is measurable, as is every closed set.
(ii) (Complementarity) If Ω is measurable, then $\mathbf{R}^n \backslash \Omega$ is also measurable.
(iii) (Boolean algebra property) If $(\Omega_j)_{j \in J}$ is any finite collection of measurable sets (so J is finite), then the union $\bigcup_{j \in J} \Omega_j$ and intersection $\bigcap_{j \in J} \Omega_j$ are also measurable.
(iv) (σ-algebra property) If $(\Omega_j)_{j \in J}$ are any countable collection of measurable sets (so J is countable), then the union $\bigcup_{j \in J} \Omega_j$ and intersection $\bigcap_{j \in J} \Omega_j$ are also measurable.

Note that some of these properties are redundant; for instance, (iv) will imply (iii), and once one knows all open sets are measurable, (ii) will imply that all closed sets are measurable also. The properties (i–iv) will ensure that virtually every set one cares about is measurable; though as indicated in the introduction, there do exist non-measurable sets.

To every measurable set Ω, we associate the *Lebesgue measure* $m(\Omega)$ of Ω, which will obey the following properties:

(v) (Empty set) The empty set $\emptyset$ has measure $m(\emptyset) = 0$.
(vi) (Positivity) We have $0 \leq m(\Omega) \leq +\infty$ for every measurable set Ω.
(vii) (Monotonicity) If $A \subseteq B$, and A and B are both measurable, then $m(A) \leq m(B)$.
(viii) (Finite sub-additivity) If $(A_j)_{j \in J}$ are a finite collection of measurable sets, then $m\left(\bigcup_{j \in J} A_j\right) \leq \sum_{j \in J} m(A_j)$.
(ix) (Finite additivity) If $(A_j)_{j \in J}$ are a finite collection of *disjoint* measurable sets, then $m(\bigcup_{j \in J} A_j) = \sum_{j \in J} m(A_j)$.
(x) (Countable sub-additivity) If $(A_j)_{j \in J}$ are a countable collection of measurable sets, then $m\left(\bigcup_{j \in J} A_j\right) \leq \sum_{j \in J} m(A_j)$.
(xi) (Countable additivity) If $(A_j)_{j \in J}$ are a countable collection of *disjoint* measurable sets, then $m\left(\bigcup_{j \in J} A_j\right) = \sum_{j \in J} m(A_j)$.
(xii) (Normalization) The unit cube $[0, 1]^n = \{(x_1, \ldots, x_n) \in \mathbf{R}^n : 0 \leq x_j \leq 1$ for all $1 \leq j \leq n\}$ has measure $m([0, 1]^n) = 1$.
(xiii) (Translation invariance) If Ω is a measurable set, and $x \in \mathbf{R}^n$, then $x + \Omega := \{x + y : y \in \Omega\}$ is also measurable, and $m(x + \Omega) = m(\Omega)$.

Again, many of these properties are redundant; for instance the countable additivity property can be used to deduce the finite additivity property, which in turn can be used to derive monotonicity (when combined with the positivity property). One can also obtain the sub-additivity properties from the additivity ones. Note that $m(\Omega)$ can be $+\infty$, and so in particular some of the sums in the above properties may also equal $+\infty$; in this chapter we adopt the convention that an infinite sum $\sum_{j \in J} a_j$ of non-negative quantities a_j is equal to $+\infty$ if the sum is not absolutely convergent. (Since everything is non-negative we will never have to deal with indeterminate forms such as $-\infty + +\infty$.)

Our goal for this chapter can then be stated thus:

Theorem 7.1.1 (Existence of Lebesgue measure). *There exists a concept of a measurable set, and a way to assign a number $m(\Omega)$ to every measurable subset $\Omega \subseteq \mathbf{R}^n$, which obeys all of the properties (i)–(xiii).*

It turns out that Lebesgue measure is pretty much unique; any other concept of measurability and measure which obeys axioms (i)–(xiii) will largely coincide with the construction we give. However there are other measures which obey only some of the above axioms; also, we may be interested in concepts of measure for other domains than Euclidean spaces $\mathbf{R}^n$. This leads to *measure theory*, which is an entire subject in itself and will not be pursued here; however we do remark that the concept of measures is very important in modern probability, and in the finer points of analysis (e.g., in the theory of distributions).

7.2 First Attempt: Outer Measure

Before we construct Lebesgue measure, we first discuss a somewhat naive approach to finding the measure of a set—namely, we try to cover the set by boxes, and then add up the volume of each box. This approach will almost work, giving us a concept called *outer measure* which can be applied to every set and obeys all of the properties (v)–(xiii) except for the additivity properties (ix), (xi). Later we will have to modify outer measure slightly to recover the additivity property.

We begin by starting with the notion of an open box.

Definition 7.2.1 (*Open box*) An *open box* (or *box* for short) B in $\mathbf{R}^n$ is any set of the form

$$B = \prod_{i=1}^{n} (a_i, b_i) := \{(x_1, \ldots, x_n) \in \mathbf{R}^n : x_i \in (a_i, b_i) \text{ for all } 1 \le i \le n\},$$

where $b_i \ge a_i$ are real numbers. We define the *volume* $\mathrm{vol}(B)$ of this box to be the number

$$\mathrm{vol}(B) := \prod_{i=1}^{n} (b_i - a_i) = (b_1 - a_1)(b_2 - a_2) \ldots (b_n - a_n).$$

For instance, the unit cube $(0, 1)^n$ is a box, and has volume 1. In one dimension $n = 1$, boxes are the same as open intervals. One can easily check that in general dimension that open boxes are indeed open. Note that if we have $b_i = a_i$ for some i, then the box becomes empty, and has volume 0, but we still consider this to be a box (albeit a rather silly one). Sometimes we will use $\mathrm{vol}_n(B)$ instead of $\mathrm{vol}(B)$ to emphasize that we are dealing with n-dimensional volume, thus for instance $\mathrm{vol}_1(B)$ would be the length of a one-dimensional box B, $\mathrm{vol}_2(B)$ would be the area of a two-dimensional box B, etc.

Remark 7.2.2 We of course expect the measure $m(B)$ of a box to be the same as the volume $\mathrm{vol}(B)$ of that box. This is in fact an inevitable consequence of the axioms (i)–(xiii) (see Exercise 7.2.5).

Definition 7.2.3 (*Covering by boxes*) Let $\Omega \subseteq \mathbf{R}^n$ be a subset of $\mathbf{R}^n$. We say that a collection $(B_j)_{j\in J}$ of boxes *cover* Ω iff $\Omega \subseteq \bigcup_{j\in J} B_j$.

Suppose $\Omega \subseteq \mathbf{R}^n$ can be covered by a finite or countable collection of boxes $(B_j)_{j\in J}$. If we wish Ω to be measurable, and if we wish to have a measure obeying the monotonicity and sub-additivity properties (vii), (viii), (x) and if we wish $m(B_j) = \mathrm{vol}(B_j)$ for every box j, then we must have

$$m(\Omega) \leq m\left(\bigcup_{j\in J} B_j\right) \leq \sum_{j\in J} m(B_j) = \sum_{j\in J} \mathrm{vol}(B_j).$$

We thus conclude

$$m(\Omega) \leq \inf\left\{\sum_{j\in J} \mathrm{vol}(B_j) : (B_j)_{j\in J} \text{ covers } \Omega;\ J \text{ at most countable}\right\}.$$

Inspired by this, we define

Definition 7.2.4 (*Outer measure*) If Ω is a set, we define the *outer measure* $m^*(\Omega)$ of Ω to be the quantity

$$m^*(\Omega) := \inf\left\{\sum_{j\in J} \mathrm{vol}(B_j) : (B_j)_{j\in J} \text{ covers } \Omega;\ J \text{ at most countable}\right\}.$$

Since $\sum_{j=1}^{\infty} \mathrm{vol}(B_j)$ is non-negative, we know that $m^*(\Omega) \geq 0$ for all Ω. However, it is quite possible that $m^*(\Omega)$ could equal $+\infty$. Note that because we are allowing ourselves to use a countable number of boxes, that every subset of $\mathbf{R}^n$ has at least one countable cover by boxes; in fact $\mathbf{R}^n$ itself can be covered by countably many translates of the unit cube $(0, 1)^n$ (how?). We will sometimes write $m_n^*(\Omega)$ instead of $m^*(\Omega)$ to emphasize the fact that we are using n-dimensional outer measure.

Note that outer measure can be defined for every single set (not just the measurable ones), because we can take the infimum of any non-empty set. It obeys several of the desired properties of a measure:

Lemma 7.2.5 (Properties of outer measure) *Outer measure has the following six properties:*

(v) (Empty set) The empty set $\emptyset$ has outer measure $m^*(\emptyset) = 0$.
(vi) (Positivity) We have $0 \leq m^*(\Omega) \leq +\infty$ for every measurable set Ω.
(vii) (Monotonicity) If $A \subseteq B \subseteq \mathbf{R}^n$, then $m^*(A) \leq m^*(B)$.

(viii) *(Finite sub-additivity) If* $(A_j)_{j\in J}$ *are a finite collection of subsets of* $\mathbf{R}^n$*, then* $m^*\left(\bigcup_{j\in J} A_j\right) \le \sum_{j\in J} m^*(A_j)$.

(x) *(Countable sub-additivity) If* $(A_j)_{j\in J}$ *are a countable collection of subsets of* $\mathbf{R}^n$*, then* $m^*\left(\bigcup_{j\in J} A_j\right) \le \sum_{j\in J} m^*(A_j)$.

(xiii) *(Translation invariance) If* Ω *is a subset of* $\mathbf{R}^n$*, and* $x \in \mathbf{R}^n$*, then* $m^*(x+\Omega) = m^*(\Omega)$.

Proof See Exercise 7.2.1. □

The outer measure of a closed box is also what we expect:

Proposition 7.2.6 (Outer measure of closed box) *For any closed box*

$$B = \prod_{i=1}^{n}[a_i, b_i] := \{(x_1, \ldots, x_n) \in \mathbf{R}^n : x_i \in [a_i, b_i] \text{ for all } 1 \le i \le n\},$$

we have

$$m^*(B) = \prod_{i=1}^{n}(b_i - a_i).$$

Proof Clearly, we can cover the closed box $B = \prod_{i=1}^{n}[a_i, b_i]$ by the open box $\prod_{i=1}^{n}(a_i - \varepsilon, b_i + \varepsilon)$ for every $\varepsilon > 0$. Thus we have

$$m^*(B) \le \text{vol}\left(\prod_{i=1}^{n}(a_i - \varepsilon, b_i + \varepsilon)\right) = \prod_{i=1}^{n}(b_i - a_i + 2\varepsilon)$$

for every $\varepsilon > 0$. Taking limits as $\varepsilon \to 0$, we obtain

$$m^*(B) \le \prod_{i=1}^{n}(b_i - a_i).$$

To finish the proof, we need to show that

$$m^*(B) \ge \prod_{i=1}^{n}(b_i - a_i).$$

By the definition of $m^*(B)$, it suffices to show that

$$\sum_{j\in J} \text{vol}(B_j) \ge \prod_{i=1}^{n}(b_i - a_i)$$

whenever $(B_j)_{j\in J}$ is a finite or countable cover of B.

Since B is closed and bounded, it is compact (by the Heine–Borel theorem, Theorem 1.5.7), and in particular every open cover has a finite subcover (Theorem 1.5.8). Thus to prove the above inequality for countable covers, it suffices to do it for finite covers (since if $(B_j)_{j\in J'}$ is a finite subcover of $(B_j)_{j\in J}$ then $\sum_{j\in J} \mathrm{vol}(B_j)$ will be greater than or equal to $\sum_{j\in J'} \mathrm{vol}(B_j)$).

To summarize, our goal is now to prove that

$$\sum_{j\in J} \mathrm{vol}(B^{(j)}) \geq \prod_{i=1}^{n}(b_i - a_i) \tag{7.1}$$

whenever $(B^{(j)})_{j\in J}$ is a finite cover of $\prod_{i=1}^n [a_i, b_i]$; we have changed the subscript B_j to superscript $B^{(j)}$ because we will need the subscripts to denote components.

To prove the inequality (7.1), we shall use induction on the dimension n. First we consider the base case $n = 1$. Here B is just a closed interval $B = [a, b]$, and each box $B^{(j)}$ is just an open interval $B^{(j)} = (a_j, b_j)$. We have to show that

$$\sum_{j\in J}(b_j - a_j) \geq (b - a).$$

To do this we use the Riemann integral. For each $j \in J$, let $f^{(j)} : \mathbf{R} \to \mathbf{R}$ be the function such that $f^{(j)}(x) = 1$ when $x \in (a_j, b_j)$ and $f^{(j)}(x) = 0$ otherwise. Then we have that $f^{(j)}$ is Riemann integrable (because it is piecewise constant, and compactly supported) and

$$\int_{-\infty}^{\infty} f^{(j)} = b_j - a_j.$$

Summing this over all $j \in J$, and interchanging the integral with the finite sum, we have

$$\int_{-\infty}^{\infty} \sum_{j\in J} f^{(j)} = \sum_{j\in J} b_j - a_j.$$

But since the intervals (a_j, b_j) cover $[a, b]$, we have $\sum_{j\in J} f^{(j)}(x) \geq 1$ for all $x \in [a, b]$ (why?). For all other values of x, we have $\sum_{j\in J} f^{(j)}(x) \geq 0$. Thus

$$\int_{-\infty}^{\infty} \sum_{j\in J} f^{(j)} \geq \int_{[a,b]} 1 = b - a$$

and the claim follows by combining this inequality with the previous equality. This proves (7.1) when $n = 1$.

Now assume inductively that $n > 1$, and we have already proven the inequality (7.1) for dimensions $n-1$. We shall use a similar argument to the preceding one. Each box $B^{(j)}$ is now of the form

$$B^{(j)} = \prod_{i=1}^{n}(a_i^{(j)}, b_i^{(j)}).$$

We can write this as

$$B^{(j)} = A^{(j)} \times (a_n^{(j)}, b_n^{(j)})$$

where $A^{(j)}$ is the $n-1$-dimensional box $A^{(j)} := \prod_{i=1}^{n-1}(a_i^{(j)}, b_i^{(j)})$. Note that

$$\mathrm{vol}(B^{(j)}) = \mathrm{vol}_{n-1}(A^{(j)})(b_n^{(j)} - a_n^{(j)})$$

where we have subscripted vol_{n-1} by $n-1$ to emphasize that this is $n-1$-dimensional volume being referred to here. We similarly write

$$B = A \times [a_n, b_n]$$

where $A := \prod_{i=1}^{n-1}[a_i, b_i]$, and again note that

$$\mathrm{vol}(B) = \mathrm{vol}_{n-1}(A)(b_n - a_n).$$

For each $j \in J$, let $f^{(j)}$ be the function such that $f^{(j)}(x_n) = \mathrm{vol}_{n-1}(A^{(j)})$ for all $x_n \in (a_n^{(j)}, b_n^{(j)})$, and $f^{(j)}(x_n) = 0$ for all other x_n. Then $f^{(j)}$ is Riemann integrable and

$$\int_{-\infty}^{\infty} f^{(j)} = \mathrm{vol}_{n-1}(A^{(j)})(b_n^{(j)} - a_n^{(j)}) = \mathrm{vol}(B^{(j)})$$

and hence

$$\sum_{j\in J} \mathrm{vol}(B^{(j)}) = \int_{-\infty}^{\infty} \sum_{j\in J} f^{(j)}.$$

Now let $x_n \in [a_n, b_n]$ and $(x_1, \ldots, x_{n-1}) \in A$. Then $(x_1, \ldots, x_n)$ lies in B, and hence lies in one of the $B^{(j)}$. Clearly we have $x_n \in (a_n^{(j)}, b_n^{(j)})$, and $(x_1, \ldots, x_{n-1}) \in A^{(j)}$. In particular, we see that for each $x_n \in [a_n, b_n]$, the set

$$\{A^{(j)} : j \in J; x_n \in (a_n^{(j)}, b_n^{(j)})\}$$

of $n-1$-dimensional boxes covers A. Applying the inductive hypothesis (7.1) at dimension $n-1$ we thus see that

$$\sum_{j\in J:x_n\in(a_n^{(j)},b_n^{(j)})} \mathrm{vol}_{n-1}(A^{(j)}) \geq \mathrm{vol}_{n-1}(A),$$

or in other words

$$\sum_{j\in J} f^{(j)}(x_n) \geq \mathrm{vol}_{n-1}(A).$$

Integrating this over $[a_n, b_n]$, we obtain

$$\int_{[a_n,b_n]} \sum_{j\in J} f^{(j)} \geq \mathrm{vol}_{n-1}(A)(b_n - a_n) = \mathrm{vol}(B)$$

and in particular

$$\int_{-\infty}^{\infty} \sum_{j\in J} f^{(j)} \geq \mathrm{vol}_{n-1}(A)(b_n - a_n) = \mathrm{vol}(B)$$

since $\sum_{j\in J} f^{(j)}$ is always non-negative. Combining this with our previous identity for $\int_{-\infty}^{\infty} \sum_{j\in J} f^{(j)}$ we obtain (7.1), and the induction is complete. □

Once we obtain the measure of a closed box, the corresponding result for an open box is easy:

Corollary 7.2.7 *For any open box*

$$B = \prod_{i=1}^{n}(a_i, b_i) := \{(x_1, \ldots, x_n) \in \mathbf{R}^n : x_i \in (a_i, b_i) \text{ for all } 1 \leq i \leq n\},$$

we have

$$m^*(B) = \prod_{i=1}^{n}(b_i - a_i).$$

In particular, outer measure obeys the normalization (xii).

Proof We may assume that $b_i > a_i$ for all i, since if $b_i = a_i$ this follows from Lemma 7.2.5(v). Now observe that

$$\prod_{i=1}^{n}[a_i + \varepsilon, b_i - \varepsilon] \subseteq \prod_{i=1}^{n}(a_i, b_i) \subseteq \prod_{i=1}^{n}[a_i, b_i]$$

for all $\varepsilon > 0$, assuming that ε is small enough that $b_i - \varepsilon > a_i + \varepsilon$ for all i. Applying Proposition 7.2.6 and Lemma 7.2.5(vii) we obtain

$$\prod_{i=1}^{n}(b_i - a_i - 2\varepsilon) \le m^*\left(\prod_{i=1}^{n}(a_i, b_i)\right) \le \prod_{i=1}^{n}(b_i - a_i).$$

Sending $\varepsilon \to 0$ and using the squeeze test (Corollary 6.4.14), one obtains the result. □

We now compute some examples of outer measure on the real line $\mathbf{R}$.

Example 7.2.8 Let us compute the one-dimensional measure of $\mathbf{R}$. Since $(-R, R) \subseteq \mathbf{R}$ for all $R > 0$, we have

$$m^*(\mathbf{R}) \ge m^*((-R, R)) = 2R$$

by Corollary 7.2.7. Letting $R \to +\infty$ we thus see that $m^*(\mathbf{R}) = +\infty$.

Example 7.2.9 Now let us compute the one-dimensional measure of $\mathbf{Q}$. From Proposition 7.2.6 we see that for each rational number $\mathbf{Q}$, the point $\{q\}$ has outer measure $m^*(\{q\}) = 0$. Since $\mathbf{Q}$ is clearly the union $\mathbf{Q} = \bigcup_{q\in\mathbf{Q}}\{q\}$ of all these rational points q, and $\mathbf{Q}$ is countable, we have

$$m^*(\mathbf{Q}) \le \sum_{q\in\mathbf{Q}} m^*(\{q\}) = \sum_{q\in\mathbf{Q}} 0 = 0,$$

and so $m^*(Q)$ must equal zero. In fact, the same argument shows that every countable set has measure zero. (This, incidentally, gives another proof that the real numbers are uncountable, Corollary 8.3.4.)

Remark 7.2.10 One consequence of the fact that $m^*(\mathbf{Q}) = 0$ is that given any $\varepsilon > 0$, it is possible to cover the rationals $\mathbf{Q}$ by a countable number of intervals whose total length is less than ε. This fact is somewhat un-intuitive; can you find a more explicit way to construct such a countable covering of $\mathbf{Q}$ by short intervals?

Example 7.2.11 Now let us compute the one-dimensional measure of the irrationals $\mathbf{R}\backslash\mathbf{Q}$. From finite sub-additivity we have

$$m^*(\mathbf{R}) \le m^*(\mathbf{R}\backslash\mathbf{Q}) + m^*(\mathbf{Q}).$$

Since $\mathbf{Q}$ has outer measure 0, and $m^*(\mathbf{R})$ has outer measure $+\infty$, we thus see that the irrationals $\mathbf{R}\backslash\mathbf{Q}$ have outer measure $+\infty$. A similar argument shows that $[0, 1]\backslash\mathbf{Q}$, the irrationals in $[0, 1]$, have outer measure 1 (why?).

Example 7.2.12 By Proposition 7.2.6, the unit interval $[0, 1]$ in $\mathbf{R}$ has one-dimensional outer measure 1, but the unit interval $\{(x, 0) : 0 \le x \le 1\}$ in $\mathbf{R}^2$ has two-dimensional outer measure 0. Thus one-dimensional outer measure and two-dimensional outer measure are quite different. Note that the above remarks and countable sub-additivity imply that the entire x-axis of $\mathbf{R}^2$ has two-dimensional outer measure 0, despite the fact that $\mathbf{R}$ has infinite one-dimensional measure.

— Exercises —

Exercise 7.2.1 Prove Lemma 7.2.5. (*Hint:* you will have to use the definition of inf, and probably introduce a parameter ε. You may have to treat separately the cases when certain outer measures are equal to $+\infty$. (viii) can be deduced from (x) and (v). For (x), label the index set J as $J = \{j_1, j_2, j_3, \ldots\}$, and for each A_j, pick a covering of A_j by boxes whose total volume is no larger than $m^*(A_j) + \varepsilon/2^j$.)

Exercise 7.2.2 Let A be a subset of $\mathbf{R}^n$, and let B be a subset of $\mathbf{R}^m$. Note that the Cartesian product $\{(a, b) : a \in A, b \in B\}$ is then a subset of $\mathbf{R}^{n+m}$. Show that $m^*_{n+m}(A \times B) \leq m^*_n(A)m^*_m(B)$. Here we adopt the convention that $c \times +\infty = +\infty \times c$ is equal to $+\infty$ for any $0 < c \leq +\infty <$ and equal to zero for $c = 0$. (It is in fact true that $m^*_{n+m}(A \times B) = m^*_n(A)m^*_m(B)$, but this is substantially harder to prove.)

In Exercises 7.2.3–7.2.5, we assume that $\mathbf{R}^n$ is a Euclidean space, and we have a notion of measurable set in $\mathbf{R}^n$ (which may or may not coincide with the notion of Lebesgue measurable set) and a notion of measure (which may or may not coincide with Lebesgue measure) which obeys axioms (i)–(xiii).

Exercise 7.2.3 (a) Show that if $A_1 \subseteq A_2 \subseteq A_3 \ldots$ is an increasing sequence of measurable sets (so $A_j \subseteq A_{j+1}$ for every positive integer j), then we have $m\left(\bigcup_{j=1}^{\infty} A_j\right) = \lim_{j\to\infty} m(A_j)$.

(b) Show that if $A_1 \supseteq A_2 \supseteq A_3 \ldots$ is a decreasing sequence of measurable sets (so $A_j \supseteq A_{j+1}$ for every positive integer j), and $m(A_1) < +\infty$, then we have $m\left(\bigcap_{j=1}^{\infty} A_j\right) = \lim_{j\to\infty} m(A_j)$.

Exercise 7.2.4 Show that for any positive integer $q > 1$, that the open box

$$(0, 1/q)^n := \{(x_1, \ldots, x_n) \in \mathbf{R}^n : 0 < x_j < 1/q \text{ for all } 1 \leq j \leq n\}$$

and the closed box

$$[0, 1/q]^n := \{(x_1, \ldots, x_n) \in \mathbf{R}^n : 0 \leq x_j \leq 1/q \text{ for all } 1 \leq j \leq n\}$$

both measure q^{-n}. (*Hint:* first show that $m((0, 1/q)^n) \leq q^{-n}$ for every $q \geq 1$ by covering $(0, 1)^n$ by some translates of $(0, 1/q)^n$. Using a similar argument, show that $m([0, 1/q]^n) \geq q^{-n}$. Then show that $m([0, 1/q]^n \backslash (0, 1/q)^n) \leq \varepsilon$ for every $\varepsilon > 0$, by covering the boundary of $[0, 1/q]^n$ with some very small boxes.)

Exercise 7.2.5 Show that for any box B, that $m(B) = \text{vol}(B)$. (*Hint:* first prove this when the co-ordinates a_j, b_j are rational, using Exercise 7.2.4. Then take limits somehow (perhaps using Q1) to obtain the general case when the co-ordinates are real.)

Exercise 7.2.6 Use Lemma 7.2.5 and Proposition 7.2.6 to furnish another proof that the reals are uncountable (i.e., reprove Corollary 8.3.4 from *Analysis I*).

7.3 Outer Measure Is not Additive

In light of Lemma 7.2.5, it would seem now that all we need to do is to verify the additivity properties (ix), (xi), and we have everything we need to have a usable measure. Unfortunately, these properties fail for outer measure, even in one dimension.

Proposition 7.3.1 (Failure of countable additivity) *There exists a countable collection $(A_j)_{j\in J}$ of disjoint subsets of* $\mathbf{R}$, *such that* $m^*(\bigcup_{j\in J} A_j) \neq \sum_{j\in J} m^*(A_j)$.

Proof We shall need some notation. Let $\mathbf{Q}$ be the rationals, and $\mathbf{R}$ be the reals. We say that a set $A \subseteq \mathbf{R}$ is a *coset* of $\mathbf{Q}$ if it is of the form $A = x + \mathbf{Q}$ for some real number x. For instance, $\sqrt{2} + \mathbf{Q}$ is a coset of $\mathbf{Q}$, as is $\mathbf{Q}$ itself, since $\mathbf{Q} = 0 + \mathbf{Q}$. Note that a coset A can correspond to several values of x; for instance $2 + \mathbf{Q}$ is exactly the same coset as $0 + \mathbf{Q}$. Also observe that it is not possible for two cosets to partially overlap; if $x + \mathbf{Q}$ intersects $y + \mathbf{Q}$ in even just a single point z, then $x - y$ must be rational (why? Use the identity $x - y = (x - z) - (y - z)$), and thus $x + \mathbf{Q}$ and $y + \mathbf{Q}$ must be equal (why?). So any two cosets are either identical or disjoint.

We observe that every coset A of the rationals $\mathbf{Q}$ has a non-empty intersection with $[0, 1]$. Indeed, if A is a coset, then $A = x + \mathbf{Q}$ for some real number x. If we then pick a rational number q in $[-x, 1 - x]$ then we see that $x + q \in [0, 1]$, and thus $A \cap [0, 1]$ contains $x + q$.

Let $\mathbf{R}/\mathbf{Q}$ denote the set of all cosets of $\mathbf{Q}$; note that this is a set whose elements are themselves sets (of real numbers). For each coset A in $\mathbf{R}/\mathbf{Q}$, let us pick an element x_A of $A \cap [0, 1]$. (This requires us to make an infinite number of choices, and thus requires the axiom of choice, see Sect. 8.4.) Let E be the set of all such x_A, i.e., $E := \{x_A : A \in \mathbf{R}/\mathbf{Q}\}$. Note that $E \subseteq [0, 1]$ by construction.

Now consider the set

$$X = \bigcup_{q\in\mathbf{Q}\cap[-1,1]} (q + E).$$

Clearly this set is contained in $[-1, 2]$ (since $q + x \in [-1, 2]$ whenever $q \in [-1, 1]$ and $x \in E \subseteq [0, 1]$). We claim that this set contains the interval $[0, 1]$. Indeed, for any $y \in [0, 1]$, we know that y must belong to some coset A (for instance, it belongs to the coset $y + \mathbf{Q}$). But we also have x_A belonging to the same coset, and thus $y - x_A$ is equal to some rational q. Since y and x_A both live in $[0, 1]$, then q lives in $[-1, 1]$. Since $y = q + x_A$, we have $y \in q + E$, and hence $y \in X$ as desired.

Note that the translates $q + E$ for $q \in \mathbf{Q}$ are all disjoint. For, if there were two distinct $q, q' \in \mathbf{Q}$ with $q + E$ intersecting $q' + E$, then there would be $A, A' \in \mathbf{R}/\mathbf{Q}$ such that $q + x_A = q' + x_{A'}$. But then $A = x_A + \mathbf{Q} = x_{A'} + \mathbf{Q} = A'$ and thus $x_A = x_{A'}$ which implies that $q = q'$, contradicting the hypothesis.

We claim that

$$m^*(X) \neq \sum_{q\in\mathbf{Q}\cap[-1,1]} m^*(q + E),$$

which would prove the claim. To see why this is true, observe that since $[0, 1] \subseteq X \subseteq [-1, 2]$, that we have $1 \leq m^*(X) \leq 3$ by monotonicity and Proposition 7.2.6.

For the right-hand side, observe from translation invariance that

$$\sum_{q \in \mathbf{Q} \cap [-1,1]} m^*(q + E) = \sum_{q \in \mathbf{Q} \cap [-1,1]} m^*(E).$$

The set $\mathbf{Q} \cap [-1, 1]$ is countably infinite (why?). Thus the right-hand side is either 0 (if $m^*(E) = 0$) or $+\infty$ (if $m^*(E) > 0$). Either way, it cannot be between 1 and 3, and the claim follows. □

Remark 7.3.2 The above proof used the axiom of choice. This turns out to be absolutely necessary; one can prove using some advanced techniques in mathematical logic that if one does not assume the axiom of choice, then it is possible to have a mathematical model where outer measure is countably additive.

One can refine the above argument, and show in fact that m^* is not finitely additive either:

Proposition 7.3.3 (Failure of finite additivity) *There exists a finite collection* $(A_j)_{j \in J}$ *of disjoint subsets of* $\mathbf{R}$, *such that*

$$m^*\left(\bigcup_{j \in J} A_j\right) \neq \sum_{j \in J} m^*(A_j).$$

Proof This is accomplished by an indirect argument. Suppose for sake of contradiction that m^* *was* finitely additive. Let E and X be the sets introduced in Proposition 7.3.1. From countable sub-additivity and translation invariance we have

$$m^*(X) \leq \sum_{q \in \mathbf{Q} \cap [-1,1]} m^*(q + E) = \sum_{q \in \mathbf{Q} \cap [-1,1]} m^*(E).$$

Since we know that $1 \leq m^*(X) \leq 3$, we thus have $m^*(E) \neq 0$, since otherwise we would have $m^*(X) \leq 0$, a contradiction.

Since $m^*(E) \neq 0$, there exists a finite integer $n > 0$ such that $m^*(E) > 1/n$. Now let J be a finite subset of $\mathbf{Q} \cap [-1, 1]$ of cardinality 3n. If m^* were finitely additive, then we would have

$$m^*\left(\bigcup_{q \in J} q + E\right) = \sum_{q \in J} m^*(q + E) = \sum_{q \in J} m^*(E) > 3n\frac{1}{n} = 3.$$

But we know that $\bigcup_{q \in J} q + E$ is a subset of X, which has outer measure at most 3. This contradicts monotonicity. Hence m^* cannot be finitely additive. □

Remark 7.3.4 The examples here are related to the *Banach-Tarski paradox*, which demonstrates (using the axiom of choice) that one can partition the unit ball in $\mathbf{R}^3$

into a finite number of pieces which, when rotated and translated, can be reassembled to form *two* complete unit balls! Of course, this partition involves non-measurable sets. We will not present this paradox here as it requires some group theory which is beyond the scope of this text.

7.4 Measurable Sets

In the previous section we saw that certain sets were badly behaved with respect to outer measure, in particular they could be used to contradict finite or countable additivity. However, those sets were rather pathological, being constructed using the axiom of choice and looking rather artificial. One would hope to be able to exclude them and then somehow recover finite and countable additivity. Fortunately, this can be done, thanks to a clever definition of Constantin Carathéodory (1873–1950):

Definition 7.4.1 (*Lebesgue measurability*) Let E be a subset of $\mathbf{R}^n$. We say that E is *Lebesgue measurable*, or *measurable* for short, iff we have the identity

$$m^*(A) = m^*(A \cap E) + m^*(A \backslash E)$$

for every subset A of $\mathbf{R}^n$. If E is measurable, we define the *Lebesgue measure* of E to be $m(E) = m^*(E)$; if E is not measurable, we leave $m(E)$ undefined.

In other words, E being measurable means that if we use the set E to divide up an arbitrary set A into two parts, we keep the additivity property. Of course, if m^* were finitely additive then every set E would be measurable; but we know from Proposition 7.3.3 that not every set is finitely additive. One can think of the measurable sets as the sets for which finite additivity works. We sometimes subscript $m(E)$ as $m_n(E)$ to emphasize the fact that we are using n-dimensional Lebesgue measure.

The above definition is somewhat hard to work with, and in practice one does not verify a set is measurable directly from this definition. Instead, we will use this definition to prove various useful properties of measurable sets (Lemmas 7.4.2–7.4.11), and after that we will rely more or less exclusively on the properties in those lemmas, and no longer need to refer to the above definition.

We begin by showing that a large number of sets are indeed measurable. The empty set $E = \emptyset$ and the whole space $E = \mathbf{R}^n$ are clearly measurable (why?). Here is another example of a measurable set:

Lemma 7.4.2 (Half-spaces are measurable) *The half-space*

$$\{(x_1, \ldots, x_n) \in \mathbf{R}^n : x_n > 0\}$$

is measurable.

Proof See Exercise 7.4.3. □

Remark 7.4.3 A similar argument will also show that any half-space of the form $\{(x_1, \ldots, x_n) \in \mathbf{R}^n : x_j > 0\}$ or $\{(x_1, \ldots, x_n) \in \mathbf{R}^n : x_j < 0\}$ for some $1 \le j \le n$ is measurable.

Now for some more properties of measurable sets.

Lemma 7.4.4 (Properties of measurable sets)

(a) If E is measurable, then $\mathbf{R}^n \backslash E$ is also measurable.
(b) (Translation invariance) If E is measurable, and $x \in \mathbf{R}^n$, then $x + E$ is also measurable, and $m(x + E) = m(E)$.
(c) If E_1 and E_2 are measurable, then $E_1 \cap E_2$ and $E_1 \cup E_2$ are measurable.
(d) (Boolean algebra property) If $E_1, E_2, \ldots, E_N$ are measurable, then $\bigcup_{j=1}^N E_j$ and $\bigcap_{j=1}^N E_j$ are measurable.
(e) Every open box, and every closed box, is measurable.
(f) Any set E of outer measure zero (i.e., $m^*(E) = 0$) is measurable.

Proof See Exercise 7.4.4. □

From Lemma 7.4.4, we have proven properties (ii), (iii), (xiii) on our wish list of measurable sets, and we are making progress toward (i). We also have finite additivity (property (ix) on our wish list):

Lemma 7.4.5 (Finite additivity) *If $(E_j)_{j \in J}$ are a finite collection of* disjoint *measurable sets, then for any set A (not necessarily measurable), we have*

$$m^*\left(A \cap \bigcup_{j \in J} E_j\right) = \sum_{j \in J} m^*(A \cap E_j).$$

Furthermore, we have $m\left(\bigcup_{j \in J} E_j\right) = \sum_{j \in J} m(E_j)$.

Proof See Exercise 7.4.6. □

Remark 7.4.6 Lemma 7.4.5 and Proposition 7.3.3, when combined, imply that there exist non-measurable sets: see Exercise 7.4.5.

Corollary 7.4.7 *If $A \subseteq B$ are two measurable sets, then $B \backslash A$ is also measurable, and*

$$m(B \backslash A) + m(A) = m(B).$$

Proof See Exercise 7.4.7. □

Now we show countable additivity.

Lemma 7.4.8 (Countable additivity) *If $(E_j)_{j \in J}$ are a countable collection of* disjoint *measurable sets, then $\bigcup_{j \in J} E_j$ is measurable, and $m\left(\bigcup_{j \in J} E_j\right) = \sum_{j \in J} m(E_j)$.*

Proof Let $E := \bigcup_{j\in J} E_j$. Our first task will be to show that E is measurable. Thus, let A be an arbitrary set (not necessarily measurable); we need to show that

$$m^*(A) = m^*(A \cap E) + m^*(A \backslash E).$$

Since J is countable, we may write $J = \{j_1, j_2, j_3, \ldots\}$. Note that

$$A \cap E = \bigcup_{k=1}^{\infty} (A \cap E_{j_k})$$

(why?) and hence by countable sub-additivity

$$m^*(A \cap E) \leq \sum_{k=1}^{\infty} m^*(A \cap E_{j_k}).$$

We rewrite this as

$$m^*(A \cap E) \leq \sup_{N\geq 1} \sum_{k=1}^{N} m^*(A \cap E_{j_k}).$$

Let F_N be the set $F_N := \bigcup_{k=1}^{N} E_{j_k}$. Since the $A \cap E_{j_k}$ are all disjoint, and their union is $A \cap F_N$, we see from Lemma 7.4.5 that

$$\sum_{k=1}^{N} m^*(A \cap E_{j_k}) = m^*(A \cap F_N)$$

and hence

$$m^*(A \cap E) \leq \sup_{N\geq 1} m^*(A \cap F_N).$$

Now we look at $A\backslash E$. Since $F_N \subseteq E$ (why?), we have $A\backslash E \subseteq A\backslash F_N$ (why?). By monotonicity, we thus have

$$m^*(A\backslash E) \leq m^*(A\backslash F_N)$$

for all N. In particular, we see that

$$\begin{aligned} m^*(A \cap E) + m^*(A\backslash E) &\leq \sup_{N\geq 1} \left(m^*(A \cap F_N) + m^*(A\backslash E) \right) \\ &\leq \sup_{N\geq 1} \left(m^*(A \cap F_N) + m^*(A\backslash F_N) \right). \end{aligned}$$

But from Lemma 7.4.4(d) we know that F_N is measurable, and hence

$$m^*(A \cap F_N) + m^*(A\backslash F_N) = m^*(A).$$

Putting this all together we obtain

$$m^*(A \cap E) + m^*(A \backslash E) \leq m^*(A).$$

But from finite sub-additivity we have

$$m^*(A \cap E) + m^*(A \backslash E) \geq m^*(A)$$

and the claim follows. This shows that E is measurable.

To finish the lemma, we need to show that $m(E)$ is equal to $\sum_{j \in J} m(E_j)$. We first observe from countable sub-additivity that

$$m(E) \leq \sum_{j \in J} m(E_j) = \sum_{k=1}^{\infty} m(E_{j_k}).$$

On the other hand, by finite additivity and monotonicity we have

$$m(E) \geq m(F_N) = \sum_{k=1}^{N} m(E_{j_k}).$$

Taking limits as $N \to \infty$ we obtain

$$m(E) \geq \sum_{k=1}^{\infty} m(E_{j_k})$$

and thus we have

$$m(E) = \sum_{k=1}^{\infty} m(E_{j_k}) = \sum_{j \in J} m(E_j)$$

as desired. □

This proves property (xi) on our wish list. Next, we do countable unions and intersections.

Lemma 7.4.9 (σ-algebra property) *If $(\Omega_j)_{j \in J}$ are any countable collection of measurable sets (so J is countable), then the union $\bigcup_{j \in J} \Omega_j$ and the intersection $\bigcap_{j \in J} \Omega_j$ are also measurable.*

Proof See Exercise 7.4.8. □

The final property left to verify on our wish list is (a). We first need a preliminary lemma.

Lemma 7.4.10 *Every open set can be written as a countable or finite union of open boxes.*

Proof We first need some notation. Call a box $B = \prod_{i=1}^n (a_i, b_i)$ *rational* if all of its components a_i, b_i are rational numbers. Observe that there are only a countable number of rational boxes (this is since a rational box is described by $2n$ rational numbers, and so has the same cardinality as $\mathbf{Q}^{2n}$. But $\mathbf{Q}$ is countable, and the Cartesian product of any finite number of countable sets is countable; see Corollaries 8.1.14, 8.1.15).

We make the following claim: given any open ball $B(x, r)$, there exists a rational box B which is contained in $B(x, r)$ and which contains x. To prove this claim, write $x = (x_1, \ldots, x_n)$. For each $1 \leq i \leq n$, let a_i and b_i be rational numbers such that

$$x_i - \frac{r}{n} < a_i < x_i < b_i < x_i + \frac{r}{n}.$$

Then it is clear that the box $\prod_{i=1}^n (a_i, b_i)$ is rational and contains x. A simple computation using Pythagoras' theorem (or the triangle inequality) also shows that this box is contained in $B(x, r)$; we leave this to the reader.

Now let E be an open set, and let Σ be the set of all rational boxes B which are subsets of E, and consider the union $\bigcup_{B \in \Sigma} B$ of all those boxes. Clearly, this union is contained in E, since every box in Σ is contained in E by construction. On the other hand, since E is open, we see that for every $x \in E$ there is a ball $B(x, r)$ contained in E, and by the previous claim this ball contains a rational box which contains x. In particular, x is contained in $\bigcup_{B \in \Sigma} B$. Thus we have

$$E = \bigcup_{B \in \Sigma} B$$

as desired; note that Σ is countable or finite because it is a subset of the set of all rational boxes, which is countable. □

Lemma 7.4.11 (Borel property) *Every open set, and every closed set, is Lebesgue measurable.*

Proof It suffices to do this for open sets, since the claim for closed sets then follows by Lemma 7.4.4(a) (i.e., property (ii)). Let E be an open set. By Lemma 7.4.10, E is the countable union of boxes. Since we already know that boxes are measurable, and that the countable union of measurable sets is measurable, the claim follows. □

The construction of Lebesgue measure and its basic properties are now complete. Now we make the next step in constructing the Lebesgue integral—describing the class of functions we can integrate.

— Exercises —

Exercise 7.4.1 If A is an open interval in $\mathbf{R}$, show that $m^*(A) = m^*(A \cap (0, \infty)) + m^*(A \backslash (0, \infty))$.

Exercise 7.4.2 If A is an open box in $\mathbf{R}^n$, and E is the half-plane $E := \{(x_1, \ldots, x_n) \in \mathbf{R}^n : x_n > 0\}$, show that $m^*(A) = m^*(A \cap E) + m^*(A \backslash E)$. (*Hint:* use Exercise 7.4.1.)

Exercise 7.4.3 Prove Lemma 7.4.2. (*Hint:* use Exercise 7.4.2.)

Exercise 7.4.4 Prove Lemma 7.4.4. (*Hints:* for (c), first prove that

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

A Venn diagram may be helpful. Also you may need the finite sub-additivity property. Use (c) to prove (d), and use (bd) and the various versions of Lemma 7.4.2 to prove (e)).

Exercise 7.4.5 Show that the set E used in the proof of Propositions 7.3.1 and 7.3.3 is non-measurable.

Exercise 7.4.6 Prove Lemma 7.4.5.

Exercise 7.4.7 Use Lemma 7.4.5 to prove Corollary 7.4.7.

Exercise 7.4.8 Prove Lemma 7.4.9. (*Hint:* for the countable union problem, write $J = \{j_1, j_2, \ldots\}$, write $F_N := \bigcup_{k=1}^{N} \Omega_{j_k}$, and write $E_N := F_N \backslash F_{N-1}$, with the understanding that F_0 is the empty set. Then apply Lemma 7.4.8. For the countable intersection problem, use what you just did and Lemma 7.4.4(a).)

Exercise 7.4.9 Let $A \subseteq \mathbf{R}^2$ be the set $A := [0, 1]^2 \backslash \mathbf{Q}^2$; i.e., A consists of all the points (x, y) in $[0, 1]^2$ such that x and y are not both rational. Show that A is measurable and $m(A) = 1$, but that A has no interior points. (*Hint:* it's easier to use the properties of outer measure and measure, including those in the exercises above, than to try to do this problem from first principles.)

Exercise 7.4.10 Let $A \subseteq B \subseteq \mathbf{R}^n$. Show that if B is Lebesgue measurable with measure zero, then A is also Lebesgue measurable with measure zero.

7.5 Measurable Functions

In the theory of the Riemann integral, we are only able to integrate a certain class of functions—the Riemann integrable functions. We will now be able to integrate a much larger range of functions—the *measurable functions*. More precisely, we can only integrate those measurable functions which are absolutely integrable—but more on that later.

Definition 7.5.1 (*Measurable functions*) Let Ω be a measurable subset of $\mathbf{R}^n$, and let $f : \Omega \to \mathbf{R}^m$ be a function. A function f is *measurable* iff $f^{-1}(V)$ is measurable for every open set $V \subseteq \mathbf{R}^m$.

As discussed earlier, most sets that we deal with in real life are measurable, so it is only natural to learn that most functions we deal with in real life are also measurable. For instance, continuous functions are automatically measurable:

Lemma 7.5.2 (Continuous functions are measurable) *Let Ω be a measurable subset of $\mathbf{R}^n$, and let $f: \Omega \to \mathbf{R}^m$ be continuous. Then f is also measurable.*

Proof Let V be any open subset of $\mathbf{R}^m$. Then since f is continuous, $f^{-1}(V)$ is open relative to Ω (see Theorem 2.1.5(c)), i.e., $f^{-1}(V) = W \cap \Omega$ for some open set $W \subseteq \mathbf{R}^n$ (see Proposition 1.3.4(a)). Since W is open, it is measurable; since Ω is measurable, $W \cap \Omega$ is also measurable. □

Because of Lemma 7.4.10, we have an easy criterion to test whether a function is measurable or not:

Lemma 7.5.3 *Let Ω be a measurable subset of $\mathbf{R}^n$, and let $f: \Omega \to \mathbf{R}^m$ be a function. Then f is measurable if and only if $f^{-1}(B)$ is measurable for every open box B.*

Proof See Exercise 7.5.1. □

Corollary 7.5.4 *Let Ω be a measurable subset of $\mathbf{R}^n$, and let $f: \Omega \to \mathbf{R}^m$ be a function. Suppose that $f = (f_1, \ldots, f_m)$, where $f_j: \Omega \to \mathbf{R}$ is the jth co-ordinate of f. Then f is measurable if and only if all of the f_j are individually measurable.*

Proof See Exercise 7.5.2. □

Unfortunately, it is not true that the composition of two measurable functions is automatically measurable; however we can do the next best thing: a continuous function applied to a measurable function is measurable.

Lemma 7.5.5 *Let Ω be a measurable subset of $\mathbf{R}^n$, and let W be an open subset of $\mathbf{R}^m$. If $f: \Omega \to W$ is measurable, and $g: W \to \mathbf{R}^p$ is continuous, then $g \circ f: \Omega \to \mathbf{R}^p$ is measurable.*

Proof See Exercise 7.5.3. □

This has an immediate corollary:

Corollary 7.5.6 *Let Ω be a measurable subset of $\mathbf{R}^n$. If $f: \Omega \to \mathbf{R}$ is a measurable function, then so is $|f|$,* $\max(f, 0)$*, and* $\min(f, 0)$.

Proof Apply Lemma 7.5.5 with $g(x) := |x|$, $g(x) := \max(x, 0)$, and $g(x) := \min(x, 0)$. □

A slightly less immediate corollary:

Corollary 7.5.7 *Let Ω be a measurable subset of $\mathbf{R}^n$. If $f: \Omega \to \mathbf{R}$ and $g: \Omega \to \mathbf{R}$ are measurable functions, then so is $f + g$, $f - g$, fg,* $\max(f, g)$*, and* $\min(f, g)$. *If $g(x) \neq 0$ for all $x \in \Omega$, then f/g is also measurable.*

Proof Consider $f + g$. We can write this as $k \circ h$, where $h\colon \Omega \to \mathbf{R}^2$ is the function $h(x) = (f(x), g(x))$, and $k : \mathbf{R}^2 \to \mathbf{R}$ is the function $k(a, b) := a + b$. Since f, g are measurable, then h is also measurable by Corollary 7.5.4. Since k is continuous, we thus see from Lemma 7.5.5 that $k \circ h$ is measurable, as desired. A similar argument deals with all the other cases; the only thing concerning the f/g case is that the space $\mathbf{R}^2$ must be replaced with $\{(a, b) \in \mathbf{R}^2 : b \neq 0\}$ in order to keep the map $(a, b) \mapsto a/b$ continuous and well-defined. □

Another characterization of measurable functions is given by

Lemma 7.5.8 *Let Ω be a measurable subset of $\mathbf{R}^n$, and let $f\colon \Omega \to \mathbf{R}$ be a function. Then f is measurable if and only if $f^{-1}((a, \infty))$ is measurable for every real number a.*

Proof See Exercise 7.5.4. □

Inspired by this lemma, we extend the notion of a measurable function to the extended real number system $\mathbf{R}^* := \mathbf{R} \cup \{+\infty\} \cup \{-\infty\}$:

Definition 7.5.9 (*Measurable functions in the extended reals*) Let Ω be a measurable subset of $\mathbf{R}^n$. A function $f\colon \Omega \to \mathbf{R}^*$ is said to be *measurable* iff $f^{-1}((a, +\infty])$ is measurable for every real number a.

Note that Lemma 7.5.8 ensures that the notion of measurability for functions taking values in the extended reals $\mathbf{R}^*$ is compatible with that for functions taking values in just the reals $\mathbf{R}$.

Measurability behaves well with respect to limits:

Lemma 7.5.10 (Limits of measurable functions are measurable) *Let Ω be a measurable subset of $\mathbf{R}^n$. For each positive integer n, let $f_n : \Omega \to \mathbf{R}^*$ be a measurable function. Then the functions $\sup_{n\geq 1} f_n$, $\inf_{n\geq 1} f_n$, $\limsup_{n\to\infty} f_n$, and $\liminf_{n\to\infty} f_n$ are also measurable. In particular, if the f_n converge pointwise to another function $f\colon \Omega \to \mathbf{R}^*$, then f is also measurable.*

Proof We first prove the claim about $\sup_{n\geq 1} f_n$. Call this function g. We have to prove that $g^{-1}((a, +\infty])$ is measurable for every a. But by the definition of supremum, we have

$$g^{-1}((a, +\infty]) = \bigcup_{n\geq 1} f_n^{-1}((a, +\infty])$$

(why?), and the claim follows since the countable union of measurable sets is again measurable.

A similar argument works for $\inf_{n\geq 1} f_n$. The claim for lim sup and lim inf then follow from the identities

$$\limsup_{n\to\infty} f_n = \inf_{N\geq 1} \sup_{n\geq N} f_n$$

and

$$\liminf_{n\to\infty} f_n = \sup_{N\geq 1} \inf_{n\geq N} f_n$$

(see Definition 6.4.6). □

As you can see, just about anything one does to a measurable function will produce another measurable function. This is basically why almost every function one deals with in mathematics is measurable. (Indeed, the only way to construct non-measurable functions is via artificial means such as invoking the axiom of choice.)

— Exercises —

Exercise 7.5.1 Prove Lemma 7.5.3. (*Hint:* use Lemma 7.4.10 and the σ-algebra property.)

Exercise 7.5.2 Use Lemma 7.5.3 to deduce Corollary 7.5.4.

Exercise 7.5.3 Prove Lemma 7.5.5.

Exercise 7.5.4 Prove Lemma 7.5.8. (*Hint:* use Lemma 7.5.3. As a preliminary step, you may need to show that if $f^{-1}((a,\infty))$ is measurable for all a, then $f^{-1}([a,\infty))$ is also measurable for all a.)

Exercise 7.5.5 Let $f: \mathbf{R}^n \to \mathbf{R}$ be Lebesgue measurable, and let $g: \mathbf{R}^n \to \mathbf{R}$ be a function which agrees with f outside of a set of measure zero, thus there exists a set $A \subseteq \mathbf{R}^n$ of measure zero such that $f(x) = g(x)$ for all $x \in \mathbf{R}^n \backslash A$. Show that g is also Lebesgue measurable. (*Hint:* use Exercise 7.4.10.)

Chapter 8
Lebesgue Integration

In Chap. 11, we approached the Riemann integral by first integrating a particularly simple class of functions, namely the *piecewise constant* functions. Among other things, piecewise constant functions only attain a finite number of values (as opposed to most functions in real life, which can take an infinite number of values). Once one learns how to integrate piecewise constant functions, one can then integrate other Riemann integrable functions by a similar procedure.

We shall use a similar philosophy to construct the Lebesgue integral. We shall begin by considering a special subclass of measurable functions—the *simple* functions. Then we will show how to integrate simple functions, and then from there we will integrate all measurable functions (or at least the absolutely integrable ones).

8.1 Simple Functions

Definition 8.1.1 (*Simple functions*) Let Ω be a measurable subset of $\mathbf{R}^n$, and let $f : \Omega \to \mathbf{R}$ be a measurable function. We say that f is a *simple function* if the image $f(\Omega)$ is finite. In other words, there exists a finite number of real numbers $c_1, c_2, \ldots, c_N$ such that for every $x \in \Omega$, we have $f(x) = c_j$ for some $1 \leq j \leq N$.

Example 8.1.2 Let Ω be a measurable subset of $\mathbf{R}^n$, and let E be a measurable subset of Ω. We define the *characteristic function* $\chi_E : \Omega \to \mathbf{R}$ by setting $\chi_E(x) := 1$ if $x \in E$, and $\chi_E(x) := 0$ if $x \notin E$. (In some texts, χ_E is also written 1_E and is referred to as an *indicator function*.) Then χ_E is a measurable function (why?) and is a simple function, because the image $\chi_E(\Omega)$ is $\{0, 1\}$ (or $\{0\}$ if E is empty, or $\{1\}$ if $E = \Omega$).

We remark on three basic properties of simple functions: that they form a vector space, that they are linear combinations of characteristic functions, and that they approximate measurable functions. More precisely, we have the following three lemmas:

T. Tao, *Analysis II*, Texts and Readings in Mathematics,
https://doi.org/10.1007/978-981-19-7284-3_8

Lemma 8.1.3 *Let Ω be a measurable subset of $\mathbf{R}^n$, and let $f : \Omega \to \mathbf{R}$ and $g : \Omega \to \mathbf{R}$ be simple functions. Then $f + g$ is also a simple function. Also, for any scalar $c \in \mathbf{R}$, the function cf is also a simple function.*

Proof See Exercise 8.1.1. □

Lemma 8.1.4 *Let Ω be a measurable subset of $\mathbf{R}^n$, and let $f : \Omega \to \mathbf{R}$ be a simple function. Then there exists a finite number of real numbers $c_1, \ldots, c_N$, and a finite number of disjoint measurable sets $E_1, E_2, \ldots, E_N$ in Ω, such that $f = \sum_{i=1}^N c_i \chi_{E_i}$.*

Proof See Exercise 8.1.2. □

Lemma 8.1.5 *Let Ω be a measurable subset of $\mathbf{R}^n$, and let $f : \Omega \to [0, +\infty]$ be a measurable function. Then there exists a sequence $f_1, f_2, f_3, \ldots$ of simple functions, $f_n : \Omega \to \mathbf{R}$, such that the f_n are non-negative and increasing,*

$$0 \leq f_1(x) \leq f_2(x) \leq f_3(x) \leq \ldots \text{ for all } x \in \Omega$$

and converge pointwise to f:

$$\lim_{n\to\infty} f_n(x) = f(x) \text{ for all } x \in \Omega.$$

Proof See Exercise 8.1.3. □

We now show how to compute the integral of simple functions.

Definition 8.1.6 (*Lebesgue integral of simple functions*) Let Ω be a measurable subset of $\mathbf{R}^n$, and let $f : \Omega \to \mathbf{R}$ be a simple function which is non-negative; thus f is measurable and the image $f(\Omega)$ is finite and contained in $[0, \infty)$. We then define the *Lebesgue integral* $\int_\Omega f$ of f on Ω by

$$\int_\Omega f := \sum_{\lambda \in f(\Omega); \lambda > 0} \lambda m(\{x \in \Omega : f(x) = \lambda\}).$$

We will also sometimes write $\int_\Omega f$ as $\int_\Omega f\, dm$ (to emphasize the rôle of Lebesgue measure m) or use a dummy variable such as x, e.g., $\int_\Omega f(x)\, dx$.

Example 8.1.7 Let $f : \mathbf{R} \to \mathbf{R}$ be the function which equals 3 on the interval $[1, 2]$, equals 4 on the interval $(2, 4)$, and is zero everywhere else. Then

$$\int_\Omega f := 3 \times m([1, 2]) + 4 \times m((2, 4)) = 3 \times 1 + 4 \times 2 = 11.$$

Or if $g : \mathbf{R} \to \mathbf{R}$ is the function which equals 1 on $[0, \infty)$ and is zero everywhere else, then

$$\int_\Omega g = 1 \times m([0, \infty)) = 1 \times +\infty = +\infty.$$

Thus the simple integral of a simple function can equal $+\infty$. (The reason why we restrict this integral to non-negative functions is to avoid ever encountering the indefinite form $+\infty + (-\infty)$.)

Remark 8.1.8 Note that this definition of integral corresponds to one's intuitive notion of integration (at least of non-negative functions) as the area under the graph of the function (or volume, if one is in higher dimensions).

Another formulation of the integral for non-negative simple functions is as follows.

Lemma 8.1.9 *Let Ω be a measurable subset of $\mathbf{R}^n$, and let $E_1, \ldots, E_N$ br a finite number of disjoint measurable subsets in Ω. Let $c_1, \ldots, c_N$ be non-negative numbers (not necessarily distinct). Then we have*

$$\int_\Omega \sum_{j=1}^N c_j \chi_{E_j} = \sum_{j=1}^N c_j m(E_j).$$

Proof We can assume that none of the c_j are zero, since we can just remove them from the sum on both sides of the equation. Let $f := \sum_{j=1}^N c_j \chi_{E_j}$. Then $f(x)$ is either equal to one of the c_j (if $x \in E_j$) or equal to 0 (if $x \notin \bigcup_{j=1}^N E_j$). Thus f is a simple function, and $f(\Omega) \subseteq \{0\} \cup \{c_j : 1 \leq j \leq N\}$. Thus, by the definition,

$$\begin{aligned}\int_\Omega f &= \sum_{\lambda \in \{c_j : 1 \leq j \leq N\}} \lambda m(\{x \in \Omega : f(x) = \lambda\}) \\ &= \sum_{\lambda \in \{c_j : 1 \leq j \leq N\}} \lambda m\left(\bigcup_{1 \leq j \leq N : c_j = \lambda} E_j \right).\end{aligned}$$

But by the finite additivity property of Lebesgue measure, this is equal to

$$\begin{aligned}&\sum_{\lambda \in \{c_j : 1 \leq j \leq N\}} \lambda \sum_{1 \leq j \leq N : c_j = \lambda} m(E_j) \\ &= \sum_{\lambda \in \{c_j : 1 \leq j \leq N\}} \sum_{1 \leq j \leq N : c_j = \lambda} c_j m(E_j).\end{aligned}$$

Each j appears exactly once in this sum, since c_j is only equal to exactly one value of λ. So the above expression is equal to $\sum_{1 \leq j \leq N} c_j m(E_j)$ as desired. □

Some basic properties of Lebesgue integration of non-negative simple functions:

Proposition 8.1.10 *Let Ω be a measurable set, and let $f : \Omega \to \mathbf{R}$ and $g : \Omega \to \mathbf{R}$ be non-negative simple functions.*

(a) *We have* $0 \leq \int_{\Omega} f \leq \infty$. *Furthermore, we have* $\int_{\Omega} f = 0$ *if and only if* $m(\{x \in \Omega : f(x) \neq 0\}) = 0$.
(b) *We have* $\int_{\Omega}(f+g) = \int_{\Omega} f + \int_{\Omega} g$.
(c) *For any positive number* c*, we have* $\int_{\Omega} cf = c\int_{\Omega} f$.
(d) *If* $f(x) \leq g(x)$ *for all* $x \in \Omega$*, then we have* $\int_{\Omega} f \leq \int_{\Omega} g$.

We make a very convenient notational convention: if a property $P(x)$ holds for all points in Ω, except for a set of measure zero, then we say that P holds for *almost every* point in Ω. Thus (a) asserts that $\int_{\Omega} f = 0$ if and only if f is zero for almost every point in Ω.

Proof From Lemma 8.1.4 or from the formula

$$f = \sum_{\lambda \in f(\Omega)\setminus\{0\}} \lambda \chi_{\{x\in\Omega: f(x)=\lambda\}}$$

we can write f as a combination of characteristic functions, say

$$f = \sum_{j=1}^{N} c_j \chi_{E_j},$$

where $E_1, \ldots, E_N$ are disjoint subsets of Ω and the c_j are positive. Similarly we can write

$$g = \sum_{k=1}^{M} d_k \chi_{F_k}$$

where $F_1, \ldots, F_M$ are disjoint subsets of Ω and the d_k are positive.

(a) Since $\int_{\Omega} f = \sum_{j=1}^{N} c_j m(E_j)$ it is clear that the integral is between 0 and infinity. If f is zero almost everywhere, then all of the E_j must have measure zero (why?) and so $\int_{\Omega} f = 0$. Conversely, if $\int_{\Omega} f = 0$, then $\sum_{j=1}^{N} c_j m(E_j) = 0$, which can only happen when all of the $m(E_j)$ are zero (since all the c_j are positive). But then $\bigcup_{j=1}^{N} E_j$ has measure zero, and hence f is zero almost everywhere in Ω.

(b) Write $E_0 := \Omega \setminus \bigcup_{j=1}^{N} E_j$ and $c_0 := 0$, then we have $\Omega = E_0 \cup E_1 \cup \ldots \cup E_N$ and

$$f = \sum_{j=0}^{N} c_j \chi_{E_j}.$$

Similarly if we write $F_0 := \Omega \setminus \bigcup_{k=1}^{M} F_k$ and $d_0 := 0$ then

$$g = \sum_{k=0}^{M} d_k \chi_{F_k}.$$

Since $\Omega = E_0 \cup \ldots \cup E_N = F_0 \cup \ldots \cup F_M$, we have

$$f = \sum_{j=0}^{N} \sum_{k=0}^{M} c_j \chi_{E_j \cap F_k}$$

and

$$g = \sum_{k=0}^{M} \sum_{j=0}^{N} d_k \chi_{E_j \cap F_k}$$

and hence

$$f + g = \sum_{0 \le j \le N; 0 \le k \le M} (c_j + d_k) \chi_{E_j \cap F_k}.$$

By Lemma 8.1.9, we thus have

$$\int_\Omega (f + g) = \sum_{0 \le j \le N; 0 \le k \le M} (c_j + d_k) m(E_j \cap F_k).$$

On the other hand, we have

$$\int_\Omega f = \sum_{0 \le j \le N} c_j m(E_j) = \sum_{0 \le j \le N; 0 \le k \le M} c_j m(E_j \cap F_k)$$

and similarly

$$\int_\Omega g = \sum_{0 \le k \le M} d_k m(F_k) = \sum_{0 \le j \le N; 0 \le k \le M} d_k m(E_j \cap F_k)$$

and the claim (b) follows.

(c) Since $cf = \sum_{j=1}^N cc_j \chi_{E_j}$, we have $\int_\Omega cf = \sum_{j=1}^N cc_j m(E_j)$. Since $\int_\Omega f = \sum_{j=1}^N c_j m(E_j)$, the claim follows.

(d) Write $h := g - f$. Then h is simple and non-negative and $g = f + h$, hence by (b) we have $\int_\Omega g = \int_\Omega f + \int_\Omega h$. But by (a) we have $\int_\Omega h \ge 0$, and the claim follows.

□

— Exercise —

Exercise 8.1.1 Prove Lemma 8.1.3.

Exercise 8.1.2 Prove Lemma 8.1.4.

Exercise 8.1.3 Prove Lemma 8.1.5. (*Hint*: set

$$f_n(x) := \sup\{\frac{j}{2^n} : j \in \mathbf{Z}, \frac{j}{2^n} \leq \min(f(x), 2^n)\},$$

, i.e., $f_n(x)$ is the greatest integer multiple of 2^{-n} which does not exceed either $f(x)$ or 2^n. You may wish to draw a picture to see how f_1, f_2, f_3, etc., works. Then prove that f_n obeys all the required properties.)

8.2 Integration of Non-negative Measurable Functions

We now pass from the integration of non-negative simple functions to the integration of non-negative measurable functions. We will allow our measurable functions to take the value of $+\infty$ sometimes.

Definition 8.2.1 (*Majorization*) Let $f : \Omega \to \mathbf{R}$ and $g : \Omega \to \mathbf{R}$ be functions. We say that f *majorizes* g, or g *minorizes* f, if we have $f(x) \geq g(x)$ for all $x \in \Omega$.

We sometimes use the phrase "f dominates g" instead of "f majorizes g".

Definition 8.2.2 (*Lebesgue integral for non-negative functions*) Let Ω be a measurable subset of $\mathbf{R}^n$, and let $f : \Omega \to [0, \infty]$ be measurable and non-negative. Then we define the *Lebesgue integral* $\int_\Omega f$ of f on Ω to be

$$\int_\Omega f := \sup\left\{\int_\Omega s : s \text{ is simple and non-negative, and minorizes } f\right\}.$$

Remark 8.2.3 The reader should compare this notion to that of a lower Riemann integral from Definition 11.3.2. Interestingly, we will not need to match this lower integral with an upper integral here.

Remark 8.2.4 Note that if Ω' is any measurable subset of Ω, then we can define $\int_{\Omega'} f$ as well by restricting f to Ω', thus $\int_{\Omega'} f := \int_{\Omega'} f|_{\Omega'}$.

We have to check that this definition is consistent with our previous notion of Lebesgue integral for non-negative simple functions; in other words, if $f : \Omega \to \mathbf{R}$ is a non-negative simple function, then the value of $\int_\Omega f$ given by this definition should be the same as the one given in the previous definition. But this is clear because f certainly minorizes itself, and any other non-negative simple function s which minorizes f will have an integral $\int_\Omega s$ less than or equal to $\int_\Omega f$, thanks to Proposition 8.1.10(d).

Remark 8.2.5 Note that $\int_\Omega f$ is always at least 0, since 0 is simple, non-negative, and minorizes f. Of course, $\int_\Omega f$ could equal $+\infty$.

Some basic properties of the Lebesgue integral on non-negative measurable functions (which supercede Proposition 8.1.10):

Proposition 8.2.6 *Let Ω be a measurable set, and let $f : \Omega \to [0, \infty]$ and $g : \Omega \to [0, \infty]$ be non-negative measurable functions.*

(a) We have $0 \le \int_\Omega f \le \infty$. Furthermore, we have $\int_\Omega f = 0$ if and only if $f(x) = 0$ for almost every $x \in \Omega$.
(b) For any positive number c, we have $\int_\Omega cf = c \int_\Omega f$.
(c) If $f(x) \le g(x)$ for all $x \in \Omega$, then we have $\int_\Omega f \le \int_\Omega g$.
(d) If $f(x) = g(x)$ for almost every $x \in \Omega$, then $\int_\Omega f = \int_\Omega g$.
(e) If $\Omega' \subseteq \Omega$ is measurable, then $\int_{\Omega'} f = \int_\Omega f \chi_{\Omega'} \le \int_\Omega f$.

Proof See Exercise 8.2.1. □

Remark 8.2.7 Proposition 8.2.6(d) is quite interesting; it says that one can modify the values of a function on any measure zero set (e.g., you can modify a function on every rational number), and not affect its integral at all. It is as if no individual point, or even a measure zero collection of points, has any "vote" in what the integral of a function should be; only the collective set of points has an influence on an integral.

Remark 8.2.8 Note that we do not yet try to interchange sums and integrals. From the definition it is fairly easy to prove that $\int_\Omega (f + g) \ge \int_\Omega f + \int_\Omega g$ (Exercise 8.2.2), but to prove equality requires more work and will be done later.

As we have seen in previous chapters, we cannot always interchange an integral with a limit (or with limit-like concepts such as supremum). However, with the Lebesgue integral it is possible to do so if the functions are increasing:

Theorem 8.2.9 (Lebesgue monotone convergence theorem) *Let Ω be a measurable subset of $\mathbf{R}^n$, and let $(f_n)_{n=1}^\infty$ be a sequence of non-negative measurable functions from Ω to $[0, +\infty]$ which are increasing in the sense that*

$$0 \le f_1(x) \le f_2(x) \le f_3(x) \le \dots \text{ for all } x \in \Omega.$$

(Note we are assuming that $f_n(x)$ is increasing with respect to n; this is a different notion from $f_n(x)$ increasing with respect to x.) Then we have

$$0 \le \int_\Omega f_1 \le \int_\Omega f_2 \le \int_\Omega f_3 \le \dots$$

and

$$\int_\Omega \sup_n f_n = \sup_n \int_\Omega f_n.$$

Proof The first conclusion is clear from Proposition 8.2.6(c). Now we prove the second conclusion. From Proposition 8.2.6(c) again we have

$$\int_\Omega \sup_m f_m \ge \int_\Omega f_n$$

for every n; taking suprema in n we obtain

$$\int_\Omega \sup_m f_m \geq \sup_n \int_\Omega f_n$$

which is one half of the desired conclusion. To finish the proof we have to show

$$\int_\Omega \sup_m f_m \leq \sup_n \int_\Omega f_n.$$

From the definition of $\int_\Omega \sup_m f_m$, it will suffice to show that

$$\int_\Omega s \leq \sup_n \int_\Omega f_n$$

for all simple non-negative functions which minorize $\sup_m f_m$.

Fix s. We will show that

$$(1-\varepsilon)\int_\Omega s \leq \sup_n \int_\Omega f_n$$

for every $0 < \varepsilon < 1$; the claim then follows by taking limits as $\varepsilon \to 0$.

Fix ε. By construction of s, we have

$$s(x) \leq \sup_n f_n(x)$$

for every $x \in \Omega$. Hence, for every $x \in \Omega$ there exists an N (depending on x) such that

$$f_N(x) \geq (1-\varepsilon)s(x).$$

Since the f_n are increasing, this will imply that $f_n(x) \geq (1-\varepsilon)s(x)$ for all $n \geq N$. Thus, if we define the sets E_n by

$$E_n := \{x \in \Omega : f_n(x) \geq (1-\varepsilon)s(x)\}$$

then we have $E_1 \subseteq E_2 \subseteq E_3 \subseteq \ldots$ and $\bigcup_{n=1}^\infty E_n = \Omega$.

It is not difficult to check that all the E_n are measurable. From Proposition 8.2.6(bce) we have

$$(1-\varepsilon)\int_{E_n} s = \int_{E_n} (1-\varepsilon)s \leq \int_{E_n} f_n \leq \int_\Omega f_n$$

so to finish the argument it will suffice to show that

$$\sup_n \int_{E_n} s = \int_\Omega s.$$

Since s is a simple function, we may write $s = \sum_{j=1}^N c_j \chi_{F_j}$ for some measurable F_j and positive c_j. Since

$$\int_\Omega s = \sum_{j=1}^N c_j m(F_j)$$

and

$$\int_{E_n} s = \int_{E_n} \sum_{j=1}^N c_j \chi_{F_j \cap E_n} = \sum_{j=1}^N c_j m(F_j \cap E_n)$$

it thus suffices to show that

$$\sup_n m(F_j \cap E_n) = m(F_j)$$

for each j. But this follows from Exercise 7.2.3(a). □

This theorem is extremely useful. For instance, we can now interchange addition and integration:

Lemma 8.2.10 (Interchange of addition and integration) *Let Ω be a measurable subset of $\mathbf{R}^n$, and let $f : \Omega \to [0, \infty]$ and $g : \Omega \to [0, \infty]$ be measurable functions. Then $\int_\Omega (f + g) = \int_\Omega f + \int_\Omega g$.*

Proof By Lemma 8.1.5, there exists a sequence $0 \le s_1 \le s_2 \le \dots \le f$ of simple functions such that $\sup_n s_n = f$, and similarly a sequence $0 \le t_1 \le t_2 \le \ldots \le g$ of simple functions such that $\sup_n t_n = g$. Since the s_n are increasing and the t_n are increasing, it is then easy to check that $s_n + t_n$ is also increasing and $\sup_n (s_n + t_n) = f + g$ (why?). By the monotone convergence theorem (Theorem 8.2.9) we thus have

$$\int_\Omega f = \sup_n \int_\Omega s_n$$

$$\int_\Omega g = \sup_n \int_\Omega t_n$$

$$\int_\Omega (f + g) = \sup_n \int_\Omega (s_n + t_n).$$

But by Proposition 8.1.10(db) we have $\int_\Omega (s_n + t_n) = \int_\Omega s_n + \int_\Omega t_n$. By Proposition 8.1.9(d), $\int_\Omega s_n$ and $\int_\Omega t_n$ are both increasing in n, so

$$\sup_n \left(\int_\Omega s_n + \int_\Omega t_n \right) = \left(\sup_n \int_\Omega s_n \right) + \left(\sup_n \int_\Omega t_n \right)$$

and the claim follows. □

Of course, once one can interchange an integral with a sum of two functions, one can handle an integral and any finite number of functions by induction. More surprisingly, one can handle infinite sums as well of *non-negative* functions:

Corollary 8.2.11 *If Ω is a measurable subset of $\mathbf{R}^n$, and $g_1, g_2, \ldots$ are a sequence of non-negative measurable functions from Ω to $[0, \infty]$, then*

$$\int_\Omega \sum_{n=1}^{\infty} g_n = \sum_{n=1}^{\infty} \int_\Omega g_n.$$

Proof See Exercise 8.2.3. □

Remark 8.2.12 Note that we do not need to assume anything about the convergence of the above sums; it may well happen that both sides are equal to $+\infty$. However, we *do* need to assume non-negativity; see Exercise 8.3.4.

One could similarly ask whether we could interchange limits and integrals; in other words, is it true that

$$\int_\Omega \lim_{n\to\infty} f_n = \lim_{n\to\infty} \int_\Omega f_n.$$

Unfortunately, this is not true, as the following "moving bump" example shows. For each $n = 1, 2, 3 \ldots$, let $f_n : \mathbf{R} \to \mathbf{R}$ be the function $f_n = \chi_{[n,n+1)}$. Then $\lim_{n\to\infty} f_n(x) = 0$ for every x, but $\int_\mathbf{R} f_n = 1$ for every n, and hence $\lim_{n\to\infty} \int_\mathbf{R} f_n = 1 \neq 0$. In other words, the limiting function $\lim_{n\to\infty} f_n$ can end up having significantly smaller integral than any of the original integrals. However, the following very useful lemma of Fatou shows that the reverse cannot happen—there is no way the limiting function has larger integral than the (limit of the) original integrals:

Lemma 8.2.13 (Fatou's lemma) *Let Ω be a measurable subset of $\mathbf{R}^n$, and let $f_1, f_2, \ldots$ be a sequence of non-negative functions from Ω to $[0, \infty]$. Then*

$$\int_\Omega \liminf_{n\to\infty} f_n \leq \liminf_{n\to\infty} \int_\Omega f_n.$$

Proof Recall that

$$\liminf_{n\to\infty} f_n = \sup_n \left(\inf_{m\geq n} f_m \right)$$

and hence by the monotone convergence theorem

$$\int_\Omega \liminf_{n\to\infty} f_n = \sup_n \int_\Omega \left(\inf_{m\geq n} f_m \right).$$

By Proposition 8.2.6(c) we have

$$\int_\Omega \left(\inf_{m\geq n} f_m \right) \leq \int_\Omega f_j$$

for every $j \geq n$; taking infima in j we obtain

$$\int_\Omega \left(\inf_{m\geq n} f_m \right) \leq \inf_{j\geq n} \int_\Omega f_j.$$

Thus

$$\int_\Omega \liminf_{n\to\infty} f_n \leq \sup_n \inf_{j\geq n} \int_\Omega f_j = \liminf_{n\to\infty} \int_\Omega f_n$$

as desired. □

Note that we are allowing our functions to take the value $+\infty$ at some points. It is even possible for a function to take the value $+\infty$ but still have a finite integral; for instance, if E is a measure zero set, and $f : \Omega \to \mathbf{R}$ is equal to $+\infty$ on E but equals 0 everywhere else, then $\int_\Omega f = 0$ by Proposition 8.2.6(a). However, if the integral is finite, the function must be finite almost everywhere:

Lemma 8.2.14 *Let Ω be a measurable subset of $\mathbf{R}^n$, and let $f : \Omega \to [0, \infty]$ be a non-negative measurable function such that $\int_\Omega f$ is finite. Then f is finite almost everywhere (i.e., the set $\{x \in \Omega : f(x) = +\infty\}$ has measure zero).*

Proof See Exercise 8.2.4. □

Form Corollary 8.2.11 and Lemma 8.2.14 one has a useful lemma:

Lemma 8.2.15 (Borel–Cantelli lemma) *Let $\Omega_1, \Omega_2, \ldots$ be measurable subsets of $\mathbf{R}^n$ such that $\sum_{n=1}^\infty m(\Omega_n)$ is finite. Then the set*

$$\{x \in \mathbf{R}^n : x \in \Omega_n \textit{ for infinitely many } n\}$$

is a set of measure zero. In other words, almost every point belongs to only finitely many Ω_n.

Proof See Exercise 8.2.5. □

— Exercise —

Exercise 8.2.1 Prove Proposition 8.2.6. (*Hint:* do not attempt to mimic the proof of Proposition 8.1.10; rather, try to use Proposition 8.1.10 and Definition 8.2.2. For one direction of part (a), start with $\int_\Omega f = 0$ and conclude that $m(\{x \in \Omega : f(x) > 1/n\}) = 0$ for every $n = 1, 2, 3, \ldots$, and then use the countable subadditivity. To prove (e), first prove it for simple functions.)

Exercise 8.2.2 Let Ω be a measurable subset of $\mathbf{R}^n$, and let $f : \Omega \to [0, +\infty]$ and $g : \Omega \to [0, +\infty]$ be measurable functions. Without using Theorem 8.2.9 or Lemma 8.2.10, prove that $\int_\Omega (f + g) \geq \int_\Omega f + \int_\Omega g$.

Exercise 8.2.3 Prove Corollary 8.2.11. (*Hint*: use the monotone convergence theorem with $f_N := \sum_{n=1}^N g_n$.)

Exercise 8.2.4 Prove Lemma 8.2.14.

Exercise 8.2.5 Use Corollary 8.2.11 and Lemma 8.2.14 to prove Lemma 8.2.15. (*Hint:* use the indicator functions χ_{Ω_n}.)

Exercise 8.2.6 Let $p > 2$ and $c > 0$. Using the Borel–Cantelli lemma, show that the set

$$\left\{x \in [0, 1] : |x - \frac{a}{q}| \leq \frac{c}{q^p} \text{ for infinitely many positive integers } a, q\right\}$$

has measure zero. (*Hint*: one only has to consider those integers a in the range $0 \leq a \leq q$ (why?). Use Corollary 11.6.5 to show that the sum $\sum_{q=1}^\infty \frac{c(q+1)}{q^p}$ is finite.)

Exercise 8.2.7 Call a real number $x \in \mathbf{R}$ *diophantine* if there exist real numbers $p, C > 0$ such that $|x - \frac{a}{q}| > C/|q|^p$ for all nonzero integers q and all integers a. Using Exercise 8.2.6, show that almost every real number is diophantine. (*Hint:* first work in the interval $[0, 1]$. Show that one can take p and C to be rational and one can also take $p > 2$. Then use the fact that the countable union of measure zero sets has measure zero.)

Exercise 8.2.8 For every positive integer n, let $f_n : \mathbf{R} \to [0, \infty)$ be a non-negative measurable function such that

$$\int_{\mathbf{R}} f_n \leq \frac{1}{4^n}.$$

Show that for every $\varepsilon > 0$, there exists a set E of Lebesgue measure $m(E) \leq \varepsilon$ such that $f_n(x)$ converges pointwise to zero for all $x \in \mathbf{R} \backslash E$. (*Hint*: first prove that $m(\{x \in \mathbf{R} : f_n(x) > \frac{1}{\varepsilon 2^n}\}) \leq \frac{\varepsilon}{2^n}$ for all $n = 1, 2, 3, \ldots$, and then consider the union of all the sets $\{x \in \mathbf{R} : f_n(x) > \frac{1}{\varepsilon 2^n}\}$.)

Exercise 8.2.9 For every positive integer n, let $f_n \colon [0, 1] \to [0, \infty)$ be a non-negative measurable function such that f_n converges pointwise to zero. Show that for every $\varepsilon > 0$, there exists a set E of Lebesgue measure $m(E) \le \varepsilon$ such that $f_n(x)$ converges *uniformly* to zero for all $x \in [0, 1]\backslash E$. (This is a special case of *Egoroff's theorem*. To prove it, first show that for any positive integer m, we can find an $N > 0$ such that $m(\{x \in [0, 1] : f_n(x) > 1/m \text{ for all } n \ge N\}) \le \varepsilon/2^m$.) Is the claim still true if $[0, 1]$ is replaced by $\mathbf{R}$?

Exercise 8.2.10 Give an example of a bounded non-negative function $f : \mathbf{N} \times \mathbf{N} \to \mathbf{R}^+$ such that $\sum_{m=1}^{\infty} f(n, m)$ converges for every n, and such that $\lim_{n\to\infty} f(n, m)$ exists for every m, but such that

$$\lim_{n\to\infty} \sum_{m=1}^{\infty} f(n, m) \ne \sum_{m=1}^{\infty} \lim_{n\to\infty} f(n, m).$$

(*Hint*: modify the moving bump example. It is even possible to use a function f which only takes the values 0 and 1.) This shows that interchanging limits and infinite sums can be dangerous.

8.3 Integration of Absolutely Integrable Functions

We have now completed the theory of the Lebesgue integral for non-negative functions. Now we consider how to integrate functions which can be both positive and negative. However, we do wish to avoid the indefinite expression $+\infty + (-\infty)$, so we will restrict our attention to a subclass of measurable functions—the *absolutely integrable functions*.

Definition 8.3.1 (*Absolutely integrable functions*) Let Ω be a measurable subset of $\mathbf{R}^n$. A measurable function $f \colon \Omega \to \mathbf{R}^*$ is said to be *absolutely integrable* if the integral $\int_\Omega |f|$ is finite.

Of course, $|f|$ is always non-negative, so this definition makes sense even if f changes sign. Absolutely integrable functions are also known as $L^1(\Omega)$ functions.

If $f \colon \Omega \to \mathbf{R}^*$ is a function, we define the *positive part* $f^+ \colon \Omega \to [0, \infty]$ and *negative part* $f^- \colon \Omega \to [0, \infty]$ by the formulae

$$f^+ := \max(f, 0); \quad f^- := -\min(f, 0).$$

From Corollary 7.5.6 (which can be extended to $\mathbf{R}^*$-valued functions without difficulty) we know that f^+ and f^- are measurable. Observe also that f^+ and f^- are non-negative, that $f = f^+ - f^-$, and $|f| = f^+ + f^-$. (Why?).

Definition 8.3.2 (*Lebesgue integral*) Let $f \colon \Omega \to \mathbf{R}^*$ be an absolutely integrable function. We define the *Lebesgue integral* $\int_\Omega f$ of f to be the quantity

$$\int_\Omega f := \int_\Omega f^+ - \int_\Omega f^-.$$

Note that since f is absolutely integrable, $\int_\Omega f^+$ and $\int_\Omega f^-$ are less than or equal to $\int_\Omega |f|$ and hence are finite. Thus $\int_\Omega f$ is always finite; we are never encountering the indeterminate form $+\infty - (+\infty)$.

Note that this definition is consistent with our previous definition of the Lebesgue integral for non-negative functions, since if f is non-negative then $f^+ = f$ and $f^- = 0$. We also have the useful *triangle inequality*

$$\left|\int_\Omega f\right| \leq \int_\Omega f^+ + \int_\Omega f^- = \int_\Omega |f| \tag{8.1}$$

(Exercise 8.3.1).

Some other properties of the Lebesgue integral:

Proposition 8.3.3 *Let Ω be a measurable set, and let $f: \Omega \to \mathbf{R}$ and $g: \Omega \to \mathbf{R}$ be absolutely integrable functions.*

(a) For any real number c (positive, zero, or negative), we have that cf is absolutely integrable and $\int_\Omega cf = c \int_\Omega f$.
(b) The function $f + g$ is absolutely integrable, and $\int_\Omega (f + g) = \int_\Omega f + \int_\Omega g$.
(c) If $f(x) \leq g(x)$ for all $x \in \Omega$, then we have $\int_\Omega f \leq \int_\Omega g$.
(d) If $f(x) = g(x)$ for almost every $x \in \Omega$, then $\int_\Omega f = \int_\Omega g$.

Proof See Exercise 8.3.2. □

As mentioned in the previous section, one cannot necessarily interchange limits and integrals, $\lim \int f_n = \int \lim f_n$, as the "moving bump example" showed. However, it is possible to exclude the moving bump example and successfully interchange limits and integrals, if we know that the functions f_n are all majorized by a single absolutely integrable function. This important theorem is known as the *Lebesgue dominated convergence theorem* and is extremely useful:

Theorem 8.3.4 (Lebesgue dominated convergence thm) *Let Ω be a measurable subset of $\mathbf{R}^n$, and let $f_1, f_2, \ldots$ be a sequence of measurable functions from Ω to $\mathbf{R}^*$ which converge pointwise. Suppose also that there is an absolutely integrable function $F: \Omega \to [0, \infty]$ such that $|f_n(x)| \leq F(x)$ for all $x \in \Omega$ and all $n = 1, 2, 3, \ldots$. Then*

$$\int_\Omega \lim_{n\to\infty} f_n = \lim_{n\to\infty} \int_\Omega f_n.$$

Proof If F was infinite on a set of positive measure then F would not be absolutely integrable; thus the set where F is infinite has zero measure. We may delete this set from Ω (this does not affect any of the integrals) and thus assume without loss of generality that $F(x)$ is finite for every $x \in \Omega$, which implies the same assertion for the $f_n(x)$.

Let $f : \Omega \to \mathbf{R}^*$ be the function $f(x) := \lim_{n\to\infty} f_n(x)$; this function exists by hypothesis. By Lemma 7.5.10, f is measurable. Also, since $|f_n(x)| \le F(x)$ for all n and all $x \in \Omega$, we see that each f_n is absolutely integrable, and by taking limits we obtain $|f(x)| \le F(x)$ for all $x \in \Omega$, so f is also absolutely integrable. Our task is to show that $\lim_{n\to\infty} \int_\Omega f_n = \int_\Omega f$.

The functions $F + f_n$ are non-negative and converge pointwise to $F + f$. So by Fatou's lemma (Lemma 8.2.13)

$$\int_\Omega F + f \le \liminf_{n\to\infty} \int_\Omega F + f_n$$

and thus

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

But the functions $F - f_n$ are also non-negative and converge pointwise to $F - f$. So by Fatou's lemma again

$$\int_\Omega F - f \le \liminf_{n\to\infty} \int_\Omega F - f_n.$$

Since the right-hand side is $\int_\Omega F - \limsup_{n\to\infty} \int_\Omega f_n$ (why did the lim inf become a lim sup?), we thus have

$$\int_\Omega f \ge \limsup_{n\to\infty} \int_\Omega f_n.$$

Thus the lim inf and lim sup of $\int_\Omega f_n$ are both equal to $\int_\Omega f$, as desired. □

Finally, we record a lemma which is not particularly interesting in itself, but will have some useful consequences later in these notes.

Definition 8.3.5 (*(Upper and lower Lebesgue integral)*) Let Ω be a measurable subset of $\mathbf{R}^n$, and let $f : \Omega \to \mathbf{R}$ be a function (not necessarily measurable). We define the *upper Lebesgue integral* $\overline{\int}_\Omega f$ to be

$$\overline{\int_\Omega} f := \inf\left\{ \int_\Omega g : g \text{ is an absolutely integrable function from } \Omega \text{ to } \mathbf{R} \text{ that majorizes } f \right\}$$

and the *lower Lebesgue integral* $\underline{\int}_\Omega f$ to be

$$\underline{\int}_\Omega f := \sup\Big\{\int_\Omega g : g \text{ is an absolutely integrable function from } \Omega \text{ to } \mathbf{R} \text{ that minorizes } f\Big\}.$$

It is easy to see that $\underline{\int}_\Omega f \leq \overline{\int}_\Omega f$ (why? Use Proposition 8.3.3(c)). When f is absolutely integrable then equality occurs (why?). The converse is also true:

Lemma 8.3.6 *Let Ω be a measurable subset of $\mathbf{R}^n$, and let $f : \Omega \to \mathbf{R}$ be a function (not necessarily measurable). Let A be a real number, and suppose $\overline{\int}_\Omega f = \underline{\int}_\Omega f = A$. Then f is absolutely integrable, and*

$$\int_\Omega f = \overline{\int}_\Omega f = \underline{\int}_\Omega f = A.$$

Proof By definition of upper Lebesgue integral, for every integer $n \geq 1$ we may find an absolutely integrable function $f_n^+ : \Omega \to \mathbf{R}$ which majorizes f such that

$$\int_\Omega f_n^+ \leq A + \frac{1}{n}.$$

Similarly we may find an absolutely integrable function $f_n^- : \Omega \to \mathbf{R}$ which minorizes f such that

$$\int_\Omega f_n^- \geq A - \frac{1}{n}.$$

Let $F^+ := \inf_n f_n^+$ and $F^- := \sup_n f_n^-$. Then F^+ and F^- are measurable (by Lemma 7.5.10) and absolutely integrable (because they are squeezed between the absolutely integrable functions f_1^+ and f_1^-, for instance). Also, F^+ majorizes f and F^- minorizes f. Finally, we have

$$\int_\Omega F^+ \leq \int_\Omega f_n^+ \leq A + \frac{1}{n}$$

for every n, and hence

$$\int_\Omega F^+ \leq A.$$

Similarly we have

$$\int_\Omega F^- \geq A.$$

but F^+ majorizes F^-, and hence $\int_\Omega F^+ \geq \int_\Omega F^-$. Hence we must have

$$\int_\Omega F^+ = \int_\Omega F^- = A.$$

In particular

$$\int_\Omega F^+ - F^- = 0.$$

By Proposition 8.2.6(a), we thus have $F^+(x) = F^-(x)$ for almost every x. But since f is squeezed between F^- and F^+, we thus have $f(x) = F^+(x) = F^-(x)$ for almost every x. In particular, f differs from the absolutely integrable function F^+ only on a set of measure zero and is thus measurable (see Exercise 7.5.5) and absolutely integrable, with

$$\int_\Omega f = \int_\Omega F^+ = \int_\Omega F^- = A$$

as desired. □

— Exercise —

Exercise 8.3.1 Prove (8.1) whenever Ω is a measurable subset of $\mathbf{R}^n$ and f is an absolutely integrable function.

Exercise 8.3.2 Prove Proposition 8.3.3. (*Hint:* for (b), break f, g, and $f+g$ up into positive and negative parts, and try to write everything in terms of integrals of non-negative functions only, using Lemma 8.2.10.)

Exercise 8.3.3 Let $f: \mathbf{R} \to \mathbf{R}$ and $g: \mathbf{R} \to \mathbf{R}$ be absolutely integrable, measurable functions such that $f(x) \leq g(x)$ for all $x \in \mathbf{R}$, and that $\int_{\mathbf{R}} f = \int_R g$. Show that $f(x) = g(x)$ for almost every $x \in \mathbf{R}$ (i.e., that $f(x) = g(x)$ for all $x \in \mathbf{R}$ except possibly for a set of measure zero).

Exercise 8.3.4 For each $n = 1, 2, 3, \ldots$, let $f_n: \mathbf{R} \to \mathbf{R}$ be the function $f_n = \chi_{[n,n+1)} - \chi_{[n+1,n+2)}$; i.e., let $f_n(x)$ equal $+1$ when $x \in [n, n+1)$, equal -1 when $x \in [n+1, n+2)$, and 0 everywhere else. Show that

$$\int_{\mathbf{R}} \sum_{n=1}^{\infty} f_n \neq \sum_{n=1}^{\infty} \int_{\mathbf{R}} f_n.$$

Explain why this does not contradict Corollary 8.2.11.

8.4 Comparison with the Riemann Integral

We have spent a lot of effort constructing the Lebesgue integral, but have not yet addressed the question of how to actually compute any Lebesgue integrals, and whether Lebesgue integration is any different from the Riemann integral (say for integrals in one dimension). Now we show that the Lebesgue integral is a generalization of the Riemann integral. To clarify the following discussion, we shall temporarily distinguish the Riemann integral from the Lebesgue integral by writing the Riemann integral $\int_I f$ as $R.\int_I f$.

Our objective here is to prove

Proposition 8.4.1 *Let $I \subseteq \mathbf{R}$ be a bounded interval, and let $f : I \to \mathbf{R}$ be a Riemann integrable function. Then f is also absolutely integrable, and $\int_I f = R.\int_I f$.*

Proof Write $A := R.\int_I f$. Since f is Riemann integrable, we know that the upper and lower Riemann integrals are equal to A. Thus, for every $\varepsilon > 0$, there exists a partition $\mathbf{P}$ of I into smaller intervals J such that

$$A - \varepsilon \le \sum_{J \in \mathbf{P}} |J| \inf_{x \in J} f(x) \le A \le \sum_{J \in \mathbf{P}} |J| \sup_{x \in J} f(x) \le A + \varepsilon,$$

where $|J|$ denotes the length of J. Note that $|J|$ is the same as $m(J)$, since J is a box.

Let $f_\varepsilon^- : I \to \mathbf{R}$ and $f_\varepsilon^+ : I \to \mathbf{R}$ be the functions

$$f_\varepsilon^-(x) = \sum_{J \in \mathbf{P}} \inf_{x \in J} f(x) \chi_J(x)$$

and

$$f_\varepsilon^+(x) = \sum_{J \in \mathbf{P}} \sup_{x \in J} f(x) \chi_J(x);$$

these are simple functions and hence measurable and absolutely integrable. By Lemma 8.1.9 we have

$$\int_I f_\varepsilon^- = \sum_{J \in \mathbf{P}} |J| \inf_{x \in J} f(x)$$

and

$$\int_I f_\varepsilon^+ = \sum_{J \in \mathbf{P}} |J| \sup_{x \in J} f(x)$$

and hence

$$A - \varepsilon \le \int_I f_\varepsilon^- \le A \le \int_I f_\varepsilon^+ \le A + \varepsilon.$$

Since f_ε^+ majorizes f, and f_ε^- minorizes f, we thus have

$$A - \varepsilon \leq \underline{\int}_I f \leq \overline{\int}_I f \leq A + \varepsilon$$

for every ε, and thus

$$\underline{\int}_I f = \overline{\int}_I f = A$$

and hence by Lemma 8.3.6, f is absolutely integrable with $\int_I f = A$, as desired. □

Thus every Riemann integrable function is also Lebesgue integrable, at least on bounded intervals, and we no longer need the $R. \int_I f$ notation. However, the converse is not true. Take for instance the function $f: [0, 1] \to \mathbf{R}$ defined by $f(x):=1$ when x is rational, and $f(x):=0$ when x is irrational. Then from Proposition 11.7.1 we know that f is not Riemann integrable. On the other hand, f is the characteristic function of the set $\mathbf{Q} \cap [0, 1]$, which is countable and hence measure zero. Thus f is Lebesgue integrable and $\int_{[0,1]} f = 0$. Thus the Lebesgue integral can handle more functions than the Riemann integral; this is one of the primary reasons why we use the Lebesgue integral in analysis. (The other reason is that the Lebesgue integral interacts well with limits, as the Lebesgue monotone convergence theorem, Fatou's lemma, and Lebesgue dominated convergence theorem already attest. There are no comparable theorems for the Riemann integral.)

8.5 Fubini's Theorem

In one dimension we have shown that the Lebesgue integral is connected to the Riemann integral. Now we will try to understand the connection in higher dimensions. To simplify the discussion we shall just study two-dimensional integrals, although the arguments we present here can easily be extended to higher dimensions.

We shall study integrals of the form $\int_{\mathbf{R}^2} f$. Note that once we know how to integrate on $\mathbf{R}^2$, we can integrate on measurable subsets Ω of $\mathbf{R}^2$, since $\int_\Omega f$ can be rewritten as $\int_{\mathbf{R}^2} f \chi_\Omega$.

Let $f(x, y)$ be a function of two variables. In principle, we have three different ways to integrate f on $\mathbf{R}^2$. First of all, we can use the two-dimensional Lebesgue integral, to obtain $\int_{\mathbf{R}^2} f$. Secondly, we can fix x and compute a one-dimensional integral in y, and then take that quantity and integrate in x, thus obtaining $\int_{\mathbf{R}}(\int_{\mathbf{R}} f(x, y)\, \mathrm{d}y)\, \mathrm{d}x$. Thirdly, we could fix y and integrate in x, and then integrate in y, thus obtaining $\int_{\mathbf{R}}(\int_{\mathbf{R}} f(x, y)\, \mathrm{d}x)\, \mathrm{d}y$.

Fortunately, if the function f is absolutely integrable on f, then all three integrals are equal:

Theorem 8.5.1 (Fubini's theorem) *Let $f: \mathbf{R}^2 \to \mathbf{R}$ be an absolutely integrable function. Then there exists absolutely integrable functions $F: \mathbf{R} \to \mathbf{R}$ and $G: \mathbf{R} \to \mathbf{R}$ such that for almost every x, $f(x, y)$ is absolutely integrable in y with*

$$F(x) = \int_{\mathbf{R}} f(x, y)\, dy,$$

and for almost every y, $f(x, y)$ is absolutely integrable in x with

$$G(y) = \int_{\mathbf{R}} f(x, y)\, dx.$$

Finally, we have

$$\int_{\mathbf{R}} F(x)\, dx = \int_{\mathbf{R}^2} f = \int_{\mathbf{R}} G(y)\, dy.$$

Remark 8.5.2 Very roughly speaking, Fubini's theorem says that

$$\int_{\mathbf{R}} \left(\int_{\mathbf{R}} f(x, y)\, \mathrm{d}y \right) \mathrm{d}x = \int_{\mathbf{R}^2} f = \int_{\mathbf{R}} \left(\int_{\mathbf{R}} f(x, y)\, \mathrm{d}x \right) \mathrm{d}y.$$

This allows us to compute two-dimensional integrals by splitting them into two one-dimensional integrals. The reason why we do not write Fubini's theorem this way, though, is that it is possible that the integral $\int_{\mathbf{R}} f(x, y)\, \mathrm{d}y$ does not actually exist for every x, and similarly $\int_{\mathbf{R}} f(x, y)\, \mathrm{d}x$ does not exist for every y; Fubini's theorem only asserts that these integrals only exist for *almost every* x and y. For instance, if $f(x, y)$ is the function which equals 1 when $y > 0$ and $x = 0$, equals -1 when $y < 0$ and $x = 0$, and is zero otherwise, then f is absolutely integrable on $\mathbf{R}^2$ and $\int_{\mathbf{R}^2} f = 0$ (since f equals zero almost everywhere in $\mathbf{R}^2$), but $\int_{\mathbf{R}} f(x, y)\, \mathrm{d}y$ is not absolutely integrable when $x = 0$ (though it is absolutely integrable for every other x).

Proof The proof of Fubini's theorem is quite complicated, and we will only give a sketch here. We begin with a series of reductions.

Roughly speaking (ignoring issues relating to sets of measure zero), we have to show that

$$\int_{\mathbf{R}} \left(\int_{\mathbf{R}} f(x, y)\, \mathrm{d}y \right) \mathrm{d}x = \int_{\mathbf{R}^2} f$$

together with a similar equality with x and y reversed. We shall just prove the above equality, as the other one is very similar.

First of all, it suffices to prove the theorem for non-negative functions, since the general case then follows by writing a general function f as a difference $f^+ - f^-$ of two non-negative functions, and applying Fubini's theorem to f^+ and f^- separately (and using Proposition 8.3.3(a) and (b)). Thus we will henceforth assume that f is non-negative.

Next, it suffices to prove the theorem for non-negative functions f supported on a bounded set such as $[-N, N] \times [-N, N]$ for some positive integer N. Indeed, once one obtains Fubini's theorem for such functions, one can then write a general function f as the supremum of such compactly supported functions as

$$f = \sup_{N>0} f \chi_{[-N,N]\times[-N,N]},$$

apply Fubini's theorem to each function $f \chi_{[-N,N]\times[-N,N]}$ separately, and then take suprema using the monotone convergence theorem. Thus we will henceforth assume that f is supported on $[-N, N] \times [-N, N]$.

By another similar argument, it suffices to prove the theorem for non-negative simple functions supported on $[-N, N] \times [-N, N]$, since one can use Lemma 8.1.5 to write f as the supremum of simple functions (which must also be supported on $[-N, N]$), apply Fubini's theorem to each simple function, and then take suprema using the monotone convergence theorem. Thus we may assume that f is a non-negative simple function supported on $[-N, N] \times [-N, N]$.

Next, we see that it suffices to prove the theorem for characteristic functions supported in $[-N, N] \times [-N, N]$. This is because every simple function is a linear combination of characteristic functions, and so we can deduce Fubini's theorem for simple functions from Fubini's theorem for characteristic functions. Thus we may take $f = \chi_E$ for some measurable $E \subseteq [-N, N] \times [-N, N]$. Our task is then to show (ignoring sets of measure zero) that

$$\int_{[-N,N]} \left(\int_{[-N,N]} \chi_E(x, y)\, \mathrm{d}y \right) \mathrm{d}x = m(E).$$

It will suffice to show the upper Lebesgue integral estimate

$$\overline{\int_{[-N,N]}} \left(\overline{\int_{[-N,N]}} \chi_E(x, y)\, \mathrm{d}y \right) \mathrm{d}x \le m(E). \tag{8.2}$$

We will prove this estimate later. Once we show this for every set E, we may substitute E with $[-N, N] \times [-N, N] \backslash E$ and obtain

$$\overline{\int_{[-N,N]}} \left(\overline{\int_{[-N,N]}} (1 - \chi_E(x, y))\, \mathrm{d}y \right) \mathrm{d}x \le 4N^2 - m(E).$$

But the left-hand side is equal to

$$\overline{\int}_{[-N,N]} (2N - \underline{\int}_{[-N,N]} \chi_E(x, y)\,\mathrm{d}y)\,\mathrm{d}x$$

which is in turn equal to

$$4N^2 - \underline{\int}_{[-N,N]} \left(\underline{\int}_{[-N,N]} \chi_E(x, y)\,\mathrm{d}y \right) \mathrm{d}x$$

and thus we have

$$\underline{\int}_{[-N,N]} \left(\underline{\int}_{[-N,N]} \chi_E(x, y)\,\mathrm{d}y \right) \mathrm{d}x \geq m(E).$$

In particular we have

$$\underline{\int}_{[-N,N]} \left(\overline{\int}_{[-N,N]} \chi_E(x, y)\,\mathrm{d}y \right) \mathrm{d}x \geq m(E)$$

and hence by Lemma 8.3.6 we see that $\overline{\int}_{[-N,N]} \chi_E(x, y)\,\mathrm{d}y$ is absolutely integrable and

$$\int_{[-N,N]} \left(\overline{\int}_{[-N,N]} \chi_E(x, y)\,\mathrm{d}y \right) \mathrm{d}x = m(E).$$

A similar argument shows that

$$\int_{[-N,N]} \left(\underline{\int}_{[-N,N]} \chi_E(x, y)\,\mathrm{d}y \right) \mathrm{d}x = m(E)$$

and hence

$$\int_{[-N,N]} \left(\overline{\int}_{[-N,N]} \chi_E(x, y)\,\mathrm{d}y - \underline{\int}_{[-N,N]} \chi_E(x, y) \right) \mathrm{d}x = 0.$$

Thus by Proposition 8.2.6(a) we have

$$\underline{\int}_{[-N,N]} \chi_E(x, y)\,\mathrm{d}y = \overline{\int}_{[-N,N]} \chi_E(x, y)\,\mathrm{d}y$$

for almost every $x \in [-N, N]$. Thus $\chi_E(x, y)$ is absolutely integrable in y for almost every x, and $\int_{[-N,N]} \chi_E(x, y)$ is thus equal (almost everywhere) to a function $F(x)$ such that

$$\int\limits_{[-N,N]} F(x)\,\mathrm{d}x = m(E)$$

as desired.

It remains to prove the bound (8.2). Let $\varepsilon > 0$ be arbitrary. Since $m(E)$ is the same as the outer measure $m^*(E)$, we know that there exists an at most countable collection $(B_j)_{j \in J}$ of boxes such that $E \subseteq \bigcup_{j \in J} B_j$ and

$$\sum_{j \in J} m(B_j) \leq m(E) + \varepsilon.$$

Each box B_j can be written as $B_j = I_j \times I'_j$ for some intervals I_j and I'_j. Observe that

$$m(B_j) = |I_j||I'_j| = \int\limits_{I_j} |I'_j|\,\mathrm{d}x = \int\limits_{I_j} \left(\int\limits_{I'_j} \mathrm{d}y \right) \mathrm{d}x$$

$$= \int\limits_{[-N,N]} \left(\int\limits_{[-N,N]} \chi_{I_j \times I'_j}(x, y)\,\mathrm{d}x \right) \mathrm{d}y$$

$$= \int\limits_{[-N,N]} \left(\int\limits_{[-N,N]} \chi_{B_j}(x, y)\,\mathrm{d}x \right) \mathrm{d}y.$$

Adding this over all $j \in J$ (using Corollary 8.2.11) we obtain

$$\sum_{j \in J} m(B_j) = \int\limits_{[-N,N]} \left(\int\limits_{[-N,N]} \sum_{j \in J} \chi_{B_j}(x, y)\,\mathrm{d}x \right) \mathrm{d}y.$$

In particular we have

$$\overline{\int\limits_{[-N,N]}} \left(\overline{\int\limits_{[-N,N]}} \sum_{j \in J} \chi_{B_j}(x, y)\,\mathrm{d}x \right) \mathrm{d}y \leq m(E) + \varepsilon.$$

But $\sum_{j\in J} \chi_{B_j}$ majorizes χ_E (why?) and thus

$$\overline{\int_{[-N,N]}} \left(\overline{\int_{[-N,N]}} \chi_E(x, y)\,\mathrm{d}x \right) \mathrm{d}y \le m(E) + \varepsilon.$$

But ε is arbitrary, and so we have (8.2) as desired. This completes the proof of Fubini's theorem. □

Index

A
Abel's theorem, 72
Absolutely integrable, 179
Absolute value
  for complex numbers, 84
Abstraction, 1
Addition
  of complex numbers, 82
(countably) Additive measure, 147, 159
Adherent point
  of sets, 10
Ambient space, 13
Approximation to the identity, 59, 61, 103
Arctangent, *see* trigonometric functions
Associativity
  of addition in **C**, 82
  of scalar multiplication, 117
  of vector addition, 114

B
Ball, 9
Banach-Tarski paradox, 146, 157
Base of the natural logarithm, *see* *e*
Basis
  standard basis of row vectors, 114
Boolean algebra, 147, 159
Borel–Cantelli lemma, 177
Borel property, 147, 162
Boundary (point), 10, 34
Bounded
  function, 47
  set, 18

C
C, C^0, C^1, C^2, C^k, 131
Cauchy–Schwarz inequality, 8, 97
Cauchy sequence, 15
Chain rule
  in higher dimensions, 128, 130
Character, 99
Characteristic function, 167
Clairaut's theorem, *see* interchanging derivatives with derivatives
Closed
  box, 150
  set, 11, 34
Closure, 10, 34
Cocountable topology, 36
Coefficient, 65
Cofinite topology, 36
Column vector, 114
Commutativity, 102
  of addition in **C**, 82
  of addition in vector spaces, 114
  of convolution, 60, 102
Compactness, 18, 35
  compact support, 58
Completeness
  of metric spaces, 16
  of the space of continuous functions, 50
Completion of a metric space, 17
Complex numbers **C**, 82
  complex conjugation, 84
Connectedness, 30
  connected component, 33
Continuity, 23, 35
  and compactness, 28
  and connectedness, 32
  and convergence, 23
Contraction, 133
  mapping theorem, 134
Convergence

T. Tao, *Analysis II*, Texts and Readings in Mathematics,
https://doi.org/10.1007/978-981-19-7284-3

- in L^2, 517, 98
- of a function at a point, 40
- of sequences, 4, 34
- pointwise, *see* pointwise convergence
- uniform, *see* uniform convergence

Convolution, 59, 76, 102
Coset, 156
Cosine, *see* trigonometric functions
Cotangent, *see* trigonometric functions
Cover, 149
- *see also* open cover, 151

Critical point, 144

D
Degree, 58
De Moivre identities, 91
Dense, 58
Derivative
- directional, 122
- in higher dimensions, 120, 123, 127, 130
- matrix, 127
- partial, 123
- total, 122, 130
- uniqueness of, 121

Differentiability
- continuous, 131
- directional, 122
- infinite, 68
- in higher dimensions, 120
- k-fold, 68, 131

Differential matrix, *see* derivative matrix
Dilation, 115
Diophantine, 178
Dirac delta function, 59
Discrete metric, 4
Distance
- in **C**, 85
- in **R**, 2

Distributive law
- for complex numbers, 83

E
e, 79
Egoroff's theorem, 179
Empty
- set, 147, 149

Error-correcting codes, 3
Euclidean metric, 2
Euclidean space, 2
Euler's formula, 88, 90
Euler's number, *see e*
Exponential function, 78, 86
Exterior (point), 10, 34

F
Fatou's lemma, 176
Fejér kernel, 103
Finite intersection property, 22
Fixed point theorem, 134
Fourier
- coefficients, 100
- inversion formula, 101
- series, 101
- series for arbitrary periods, 110
- theorem, 106
- transform, 100

Fractional part, 95
Frequency, 99
Fubini's theorem, 185

G
Geodesic, 4
Geometric series
- formula, 54

Gradient, 127
Graph, 141

H
Hairy ball theorem, 134
Half-space, 158
Hausdorff space, 36
Heine-Borel theorem, 18
Hermitian form, 97
Homogeneity, 98, 115
Hypersurface, 141

I
Identity map (or operator), 115
Imaginary, 84
Implicit differentiation, 141
Implicit function theorem, 141
Inconsistent, 87
Indicator function, *see* characteristic function
Induced
- metric, 2, 13
- topology, 13, 35

Inner product, 96
Integer part, 95
Integration
- by parts, 73

Interchanging
derivatives with derivatives, 132
integrals with integrals, 176, 185
limits with derivatives, 56
limits with integrals, 56, 173, 180
limits with limits, 46
limits with sums, 179
sums with derivatives, 57, 66
sums with integrals, 54, 66, 176, 178
Interior (point), 10, 34
Intermediate value theorem, 32
Intrinsic, 18
Inverse
function theorem, 137
Invertible function, *see* bijection
local, 136
Involution, 84
Isolated point, 41
Isometry, 13

L
$l^1, l^2, l^\infty, L^1, L^2, L^\infty$, 2–4, 97, 179
equivalence of in finite dimensions, 6
Laws of algebra
for complex numbers, 82, 83
Laws of arithmetic, *see* laws of algebra
Laws of exponentiation, 78
Lebesgue dominated convergence theorem, 180
Lebesgue integral
of absolutely integrable functions, 179
of simple functions, 168
of nonnegative functions, 172
upper and lower, 181
vs. the Riemann integral, 184
Lebesgue measurable, 158
Lebesgue measure, 148
motivation of, 146–148
Lebesgue monotone convergence theorem, 173
Leibniz rule, 130
Limit
formal (LIM), 17
laws, 85
limiting values of functions, 40
pointwise, 42
uniform, *see* uniform limit
uniqueness of, 6, 41
Limit point
of sequences, 15
Linear combination, 114
Linearity
approximate, 119
of convolution, 62, 102
of inner product, 97
of integration, 170, 175
of transformations, 115
Logarithm (natural), 79
power series of, 54, 80

M
Majorize, 172
Manifold, 144
Map, *see* function
Matrix, 115
identification with linear transformations, 116–119
Maximum
principle, 29
Measurability
for functions, 163, 164
for sets, 158
motivation of, 146
see also Lebesgue measure, outer measure, 148
Metric, 2
ball, *see* ball
on **C**, 85
on **R**, 2
space, 2
Monomial, 99
Monotone (increasing or decreasing)
convergence, *see* Lebesgue monotone convergence theorem
measure, 147, 150
Moving bump example, 44, 176
Multiplication
of complex numbers, 83
of matrices, 115, 118

N
Negation
of complex numbers, 83
Neighbourhood, 34
Newton's approximation, 121
Non-degenerate, 97
Nowhere differentiable function, 57, 91

O
Of sets
adherent point, 34
Open
box, 148

cover, 19
set, 11
Order topology, 36
Orthogonality, 98
Orthonormal, 100
non-additivity of, 157
Outer measure, 149
non-Additivity of, 156

P
Parseval identity, 101, 110
Path-connected, 32
Periodic, 94
extension, 95
π, 90, 91
Plancherel formula (or theorem), 101, 107
Pointwise convergence, 42
of series, 51
topology of, 50
Polar representation, 91
Polynomial, 58
and convolution, 60
approximation by, 58, 61
Positive
complex number, 83, 87
inner product, 97
measure, 147, 149
Power series, 65
formal, 65
multiplication of, 75
uniqueness of, 70
Pre-image, *see* inverse image
Product topology, 50
Projection, 115
Pythagoras' theorem, 98

Q
Quotient rule, 131

R
Radius of convergence, 65
Real analytic, 68
Real numbers **R**
are uncountable, *see* uncountability of the reals
Real part, 84
Real-valued, 51
Reciprocal
of complex numbers, 85
Relative topology, *see* induced topology
Ring
commutative, 83
Root
mean square, *see* L^2
Row vector, 113

S
Scalar multiplication, 113
Series
of functions, 54
σ-algebra, 147, 161
Simple function, 167
Sine, *see* trigonometric functions
Space, 2
Square wave, 94, 99
Stone-Weierstrass theorem, 63, 102
Sub-additive measure, 147, 150
Subsequence, 15
Summation by parts, 73
Sup norm, *see* supremum as norm
Support, 58
Supremum (and infimum)
as metric, 4
as norm, 4, 52

T
Taxi-cab metric, 3
Taylor series, 69
Taylor's formula, *see* Taylor series
Topological space, 33
Totally bounded, 21
Translation invariance, 147, 150, 159
Transpose, 114
Triangle inequality
for integrals, 180
in **C**, 85
in Euclidean space, 8
in inner product spaces, 97
in metric spaces, 2
Trigonometric functions, 88, 92
and Fourier series, 109
power series, 88, 91
trigonometric polynomials, 99
Trivial topology, 35

U
Uniform continuity, 29
Uniform convergence, 44
and anti-derivatives, 55
and derivatives, 47
and integrals, 54
and limits, 46

and radius of convergence, 65
as a metric, 49, 95
of series, 51
Uniform limit
and Riemann integration, 53
of bounded functions, 47
of continuous functions, 46

V
Vector space, 114
Vertical line test, 59
Volume, 148

W
Weierstrass approximation theorem, 58, 61–62, 101
Weierstrass example, *see* nowhere differentiable function
Weierstrass M-test, 52