Mathematical Analysis

by

Tom Lindstrøm

Department of Mathematics
University of Oslo
2015

Preface

These notes were first written as an emergency measure when the textbook for the course MAT2400 failed to show up in the spring of 2011. The present version is corrected, updated and extended, but still far from perfect. I would like to thank everybody who has pointed out mistakes and weaknesses in the previous versions, in particular Geir Ellingsrud, Erik Løw, Per Størset, and Daniel Aubert.

Blindern, January 7th, 2014

Tom Lindstrøm

Contents

Chapter 1

Preliminaries: Proofs, Sets, and Functions

Chapters with the word "preliminaries" in the title are never much fun, but they are useful — they provide the reader with the background information necessary to enjoy the main body of the text. This chapter is no exception, but I have tried to keep it short and to the point; everything you find here will be needed at some stage, and most of the material will show up throughout the book.

Mathematical analysis is a continuation of calculus, but it is more abstract and therefore in need of a larger vocabulary and more precisely defined concepts. You have undoubtedly dealt with proofs, sets, and functions in your previous mathematics courses, but probably in a rather casual way. Now they become the centerpiece of the theory, and there is no way to understand what is going on if you don't have a good grasp of them: The subject matter is so abstract that you can no longer rely on drawings and intuition; you simply have to be able to understand the concepts and to read, create and write proofs. Fortunately, this is not as difficult as it may sound if you have never tried to take proofs and formal definitions seriously before.

1.1 Proofs

There is nothing mysterious about mathematical proofs; they are just chains of logically irrefutable arguments that bring you from things you already know to whatever you want to prove. Still there are a few tricks of the trade that are useful to know about.

Many mathematical statements are of the form "If A, then B". This simply means that whenever statement A holds, statement B also holds, but not necessarily vice versa. A typical example is: "If $n \in \mathbb{N}$ is divisible by 14, then n is divisible by 7". This is a true statement since any natural number that is divisible by 14, is also divisible by 7. The opposite statement is not true as there are numbers that are divisible by 7, but not by 14 (e.g. 7 and 21).

Instead of "If A, then B", we often say that "A implies B" and write $A \Longrightarrow B$. As already observed, $A \Longrightarrow B$ and $B \Longrightarrow A$ mean two different things. If they are both true, A and B hold in exactly the same cases, and we say that A and B are *equivalent*. In words, we say "A if and only if B", and in symbols we write $A \Longleftrightarrow B$. A typical example is:

"A triangle is equilateral if and only if all three angels are 60°"

When we want to prove that $A \Longleftrightarrow B$, it is often convenient to prove $A \Longrightarrow B$ and $B \Longrightarrow A$ separately.

If you think a little, you will realize that "$A \Longrightarrow B$" and "not-$B \Longrightarrow$ not-A" mean exactly the same thing — they both say that whenever A happens, so does B. This means that instead of proving "$A \Longrightarrow B$", we might just a well prove "not-$B \Longrightarrow$ not-A". This is called a *contrapositive proof*, and is convenient when the hypothesis "not-B" gives us more to work on than the hypothesis "A". Here is a typical example.

Proposition 1.1.1 *If n^2 is an even number, so is n.*

Proof: We prove the contrapositive statement: "If n is odd, so is n^2": If n is odd, it can be written as $n = 2k+1$ for a nonnegative integer k. But then

$$n^2 = (2k+1)^2 = 4k^2 + 4k + 1 = 2(2k^2 + 2k) + 1$$

which is clearly odd. □

It should be clear why a contrapositive proof is best in this case: The hypothesis "n is odd" is much easier to work with than the original hypothesis "n^2 is even".

A related method of proof is *proof by contradiction* or *reductio ad absurdum.* In these proofs, we assume the *opposite* of what we want to show, and prove that this leads to a contradiction. Hence our assumption must be false, and the original claim is established. Here is a well-known example.

Proposition 1.1.2 *$\sqrt{2}$ is an irrational number.*

Proof: We assume for contradiction that $\sqrt{2}$ is rational. This means that

$$\sqrt{2} = \frac{m}{n}$$

for natural numbers m and n. By cancelling as much as possible, we may assume that m and n have no common factors.

If we square the equality above and multiply by n^2 on both sides, we get

$$2n^2 = m^2$$

This means that m^2 is even, and by the previous proposition, so is m. Hence $m = 2k$ for some natural number k, and if we substitute this into the last formula above and cancel a factor 2, we see that

$$n^2 = 2k^2$$

This means that n^2 is even, and by the previous proposition n is even. Thus we have proved that both m and n are even, which is impossible as we assumed that they have no common factors. This means that the assumption that $\sqrt{2}$ is rational leads to a contradiction, and hence $\sqrt{2}$ must be irrational. □

Let me end this section by reminding you of a technique you have certainly seen before, *proof by induction.* We use this technique when we want to prove that a certain statement $P(n)$ holds for all natural numbers $n = 1, 2, 3, \ldots$. A typical statement one may want to prove in this way, is

$$P(n) : 1 + 2 + 3 + \cdots + n = \frac{n(n+1)}{2}$$

The basic observation behind the technique is:

1.1.3 (Induction Principle) *Assume that $P(n)$ is a statement about natural numbers $n = 1, 2, 3, \ldots$. Assume that the following two conditions are satisfied:*

(i) $P(1)$ is true

(ii) If $P(k)$ is true for a natural number k, then $P(k+1)$ is also true.

Then $P(n)$ holds for all natural numbers n.

Let us see how we can use the principle to prove that

$$P(n) : 1 + 2 + 3 + \cdots + n = \frac{n(n+1)}{2}$$

holds for all natural numbers n.

First we check that the statement holds for $n = 1$: In this case the formula says

$$1 = \frac{1 \cdot (1+1)}{2}$$

which is obviously true. Assume now that $P(k)$ holds for some natural number k, i.e.

$$1 + 2 + 3 + \cdots + k = \frac{k(k+1)}{2}$$

We then have

$$1+2+3+\cdots+k+(k+1)=\frac{k(k+1)}{2}+(k+1)=\frac{(k+1)(k+2)}{2}$$

which means that $P(k+1)$ is true. By the Induction Principle, $P(n)$ holds for all natural numbers n.

Exercises for Section 1.1

1. Assume that the product of two integers x and y is even. Show that at least one of the numbers is even.
2. Assume that the sum of two integers x and y is even. Show that x and y are either both even or both odd.
3. Show that if n is a natural number such that n^2 is divisible by 3, then n is divisible by 3. Use this to show that $\sqrt{3}$ is irrational.

1.2 Sets and boolean operations

In the systematic development of mathematics, *set* is usually taken as the fundamental notion from which all other concepts are developed. We shall not be so ambitious, we shall just think naively of a set as a collection of mathematical objects. A set may be finite, such as the set

$$\{1,2,3,4,5,6,7,8,9\}$$

of all natural numbers less than 10, or infinite as the set $(0,1)$ of all real numbers between 0 and 1.

We shall write $x \in A$ to say that x is an *element* of the set A, and $x \notin A$ to say that x is not an element of A. Two sets are *equal* if they have exactly the same elements, and we say that A is *subset* of B (and write $A \subseteq B$) if all elements of A are elements of B, but not necessarily vice versa. Note that there is no requirement that A is *strictly* included in B, and hence it is correct to write $A \subseteq B$ when $A = B$ (in fact, a standard technique for showing that $A = B$ is first to show that $A \subseteq B$ and then that $B \subseteq A$). By $\emptyset$ we shall mean the *empty set*, i.e. the set with no elements (you may feel that a set with no elements is a contradiction in terms, but mathematical life would be much less convenient without the empty set).

Many common sets have a standard name and notation such as

$\mathbb{N} = \{1,2,3,\ldots\}$, the set of natural numbers

$\mathbb{Z} = \{\ldots -3,-2,-1,0,1,2,3,\ldots\}$, the set of all integers

$\mathbb{Q}$, the set of all rational numbers

$\mathbb{R}$, the set of all real numbers

$\mathbb{C}$, the set of all complex numbers

$\mathbb{R}^n$, the set of all real n-tuples

To specify other sets, we shall often use expressions of the kind

$$A = \{a \mid P(a)\}$$

which means the set of all objects satisfying condition P. Often it is more convenient to write

$$A = \{a \in B \mid P(a)\}$$

which means the set of all elements in B satisfyíng the condition P. Examples of this notation are

$$[-1, 1] = \{x \in \mathbb{R} \mid -1 \leq x \leq 1\}$$

and

$$A = \{2n - 1 \mid n \in \mathbb{N}\}$$

where A is the set of all odd numbers. To increase readability I shall occasionally replace the vertical bar $|$ by a colon : and write $A = \{a : P(a)\}$ and $A = \{a \in B : P(a)\}$ instead of $A = \{a \mid P(a)\}$ and $A = \{a \in B \mid P(a)\}$, e.g. in expressions like $\{\|\alpha\mathbf{x}\| : |\alpha| < 1\}$ where there are lots of vertical bars already.

If $A_1, A_2, \ldots, A_n$ are sets, their *union* and *intersection* are given by

$$A_1 \cup A_2 \cup \ldots \cup A_n = \{a \mid a \text{ belongs to } \textit{at least one of} \text{ the sets } A_1, A_2, \ldots, A_n\}$$

and

$$A_1 \cap A_2 \cap \ldots \cap A_n = \{a \mid a \text{ belongs to } \textit{all} \text{ the sets } A_1, A_2, \ldots, A_n\},$$

respectively. Two sets are called *disjoint* if they do not have elements in common, i.e. if $A \cap B = \emptyset$.

When we calculate with numbers, the *distributive law* tells us how to move common factors in and out of parentheses:

$$b(a_1 + a_2 + \cdots + a_n) = ba_1 + ba_2 + \cdots ba_n$$

Unions and intersections are distributive both ways, i.e. we have:

Proposition 1.2.1 *For all sets* $B, A_1, A_2, \ldots, A_n$

$$B \cap (A_1 \cup A_2 \cup \ldots \cup A_n) = (B \cap A_1) \cup (B \cap A_2) \cup \ldots \cup (B \cap A_n) \quad (1.2.1)$$

and

$$B \cup (A_1 \cap A_2 \cap \ldots \cap A_n) = (B \cup A_1) \cap (B \cup A_2) \cap \ldots \cap (B \cup A_n) \quad (1.2.2)$$

Proof: We prove the first formula and leave the second as an exercise. The proof is in two steps: first we prove that the set on the left is a subset of the one on the right, and then we prove that the set on the right is a subset of the one on the left.

Assume first that x is an element of the set on the left, i.e. $x \in B \cap (A_1 \cup A_2 \cup \ldots \cup A_n)$. Then x must be in B and at least one of the sets A_i. But then $x \in B \cap A_i$, and hence $x \in (B \cap A_1) \cup (B \cap A_2) \cup \ldots \cup (B \cap A_n)$. This proves that

$$B \cap (A_1 \cup A_2 \cup \ldots \cup A_n) \subseteq (B \cap A_1) \cup (B \cap A_2) \cup \ldots \cup (B \cap A_n)$$

To prove the opposite inclusion, assume that $x \in (B \cap A_1) \cup (B \cap A_2) \cup \ldots \cup (B \cap A_n)$. Then $x \in B \cap A_i$ for at least one i, and hence $x \in B$ and $x \in A_i$. But if $x \in A_i$ for some i, then $x \in A_1 \cup A_2 \cup \ldots \cup A_n$, and hence $x \in B \cap (A_1 \cup A_2 \cup \ldots \cup A_n)$. This proves that

$$B \cap (A_1 \cup A_2 \cup \ldots \cup A_n) \supseteq (B \cap A_1) \cup (B \cap A_2) \cup \ldots \cup (B \cap A_n)$$

As we now have inclusion in both directions, (1.2.1) follows. □

Remark: It is possible to prove formula (1.2.1) in one sweep by noticing that all steps in the argument are equivalences and not only implications, but most people are more prone to making mistakes when they work with chains of equivalences than with chains of implications.

There are also other algebraic rules for unions and intersections, but most of them are so obvious that we do not need to state them here (an exception is De Morgan's laws which we shall return to in a moment).

The *set theoretic difference* $A \setminus B$ (also written $A - B$) is defined by

$$A \setminus B = \{a \mid a \in A, a \notin B\}$$

In many situations we are only interested in subsets of a given set U (often referred to as the *universe*). The *complement* A^c of a set A with respect to U is defined by

$$A^c = U \setminus A = \{a \in U \mid a \notin A\}$$

We can now formulate *De Morgan's laws*:

Proposition 1.2.2 (De Morgan's laws) *Assume that $A_1, A_2, \ldots, A_n$ are subsets of a universe U. Then*

$$(A_1 \cup A_2 \cup \ldots \cup A_n)^c = A_1^c \cap A_2^c \cap \ldots \cap A_n^c \tag{1.2.3}$$

and

$$(A_1 \cap A_2 \cap \ldots \cap A_n)^c = A_1^c \cup A_2^c \cup \ldots \cup A_n^c \tag{1.2.4}$$

(These rules are easy to remember if you observe that you can distribute the c outside the parentheses on the individual sets provided you turn all $\cup$'s into $\cap$'s and all $\cap$'s into $\cup$'s).

Proof of De Morgan's laws: We prove the first part and leave the second as an exercise. The strategy is as indicated above; we first show that any element of the set on the left must also be an element of the set on the right, and then vice versa.

Assume that $x \in (A_1 \cup A_2 \cup \ldots \cup A_n)^c$. Then $x \notin A_1 \cup A_2 \cup \ldots \cup A_n$, and hence for all i, $x \notin A_i$. This means that for all i, $x \in A_i^c$, and hence $x \in A_1^c \cap A_2^c \cap \ldots \cap A_n^c$.

Assume next that $x \in A_1^c \cap A_2^c \cap \ldots \cap A_n^c$. This means that $x \in A_i^c$ for all i, in other words: for all i, $x \notin A_i$. Thus $x \notin A_1 \cup A_2 \cup \ldots \cup A_n$ which means that $x \in (A_1 \cup A_2 \cup \ldots \cup A_n)^c$. □

We end this section with a brief look at cartesian products. If we have two sets, A and B, the *cartesian product* $A \times B$ consists of all pairs (a, b) where $a \in A$ and $b \in B$. If we have more sets $A_1, A_2, \ldots, A_n$, the cartesian product $A_1 \times A_2 \times \cdots \times A_n$ consists of all n-tuples $(a_1, a_2, \ldots, a_n)$ where $a_1 \in A_1, a_2 \in A_2, \ldots, a_n \in A_n$. If all the sets are the same (i.e. $A_i = A$ for all i), we usually write A^n instead of $A \times A \times \cdots \times A$. Hence $\mathbb{R}^n$ is the set of all n-tuples of real numbers, just as you are used to, and $\mathbb{C}^n$ is the set of all n-tuples of complex numbers.

Exercises for Section 1.2

1. Show that $[0, 2] \cup [1, 3] = [0, 3]$ and that $[0, 2] \cap [1, 3] = [1, 2]$
2. Let $U = \mathbb{R}$ be the universe. Explain that $(-\infty, 0)^c = [0, \infty)$
3. Show that $A \setminus B = A \cap B^c$.
4. The *symmetric difference* $A \triangle B$ of two sets A, B consists of the elements that belong to *exactly one* of the sets A, B. Show that

$$A \triangle B = (A \setminus B) \cup (B \setminus A)$$

5. Prove formula (1.2.2).
6. Prove formula (1.2.4).
7. Prove that $A_1 \cup A_2 \cup \ldots \cup A_n = U$ if and only if $A_1^c \cap A_2^c \cap \ldots \cap A_n^c = \emptyset$.
8. Prove that $(A \cup B) \times C = (A \times C) \cup (B \times C)$ and $(A \cap B) \times C = (A \times C) \cap (B \times C)$.

1.3 Families of sets

A collection of sets is usually called a *family*. An example is the family

$$\mathcal{A} = \{[a, b] \mid a, b \in \mathbb{R}\}$$

of all closed and bounded intervals on the real line. Families may seem abstract, but you have to get used to them as they appear in all parts of higher mathematics. We can extend the notions of union and intersection to families in the following way: If $\mathcal{A}$ is a family of sets, we define

$$\bigcup_{A\in\mathcal{A}} A = \{a \mid a \text{ belongs to at least one set } A \in \mathcal{A}\}$$

and

$$\bigcap_{A\in\mathcal{A}} A = \{a \mid a \text{ belongs to all sets } A \in \mathcal{A}\}$$

The distributive laws extend to this case in the obvious way. i.e.,

$$B \cap (\bigcup_{A\in\mathcal{A}} A) = \bigcup_{A\in\mathcal{A}} (B \cap A) \quad \text{and} \quad B \cup (\bigcap_{A\in\mathcal{A}} A) = \bigcap_{A\in\mathcal{A}} (B \cup A)$$

and so do the laws of De Morgan:

$$(\bigcup_{A\in\mathcal{A}} A)^c = \bigcap_{A\in\mathcal{A}} A^c \quad \text{and} \quad (\bigcap_{A\in\mathcal{A}} A)^c = \bigcup_{A\in\mathcal{A}} A^c$$

Families are often given as *indexed sets*. This means we we have a basic set I, and that the family consists of one set A_i for each element in I. We then write the family as

$$\mathcal{A} = \{A_i \mid i \in I\},$$

and use notation such as

$$\bigcup_{i\in I} A_i \qquad \text{and} \qquad \bigcap_{i\in I} A_i$$

or alternatively

$$\bigcup\{A_i \,:\, i \in I\} \qquad \text{and} \qquad \bigcap\{A_i \,:\, i \in I\}$$

for unions and intersections

A rather typical example of an indexed set is $\mathcal{A} = \{B_r \mid r \in [0, \infty)\}$ where $B_r = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = r^2\}$. This is the family of all circles in the plane with centre at the origin.

Exercises for Section 1.3

1. Show that $\bigcup_{n\in\mathbb{N}}[-n, n] = \mathbb{R}$
2. Show that $\bigcap_{n\in\mathbb{N}}(-\frac{1}{n}, \frac{1}{n}) = \{0\}$.
3. Show that $\bigcup_{n\in\mathbb{N}}[\frac{1}{n}, 1] = (0, 1]$
4. Show that $\bigcap_{n\in\mathbb{N}}(0, \frac{1}{n}] = \emptyset$

5. Prove the distributive laws for families. i.e.,

$$B \cap (\bigcup_{A\in\mathcal{A}} A) = \bigcup_{A\in\mathcal{A}} (B \cap A) \quad \text{and} \quad B \cup (\bigcap_{A\in\mathcal{A}} A) = \bigcap_{A\in\mathcal{A}} (B \cup A)$$

6. Prove De Morgan's laws for families:

$$(\bigcup_{A\in\mathcal{A}} A)^c = \bigcap_{A\in\mathcal{A}} A^c \quad \text{and} \quad (\bigcap_{A\in\mathcal{A}} A)^c = \bigcup_{A\in\mathcal{A}} A^c$$

1.4 Functions

Functions can be defined in terms of sets, but for our purposes it suffices to think of a function $f : X \to Y$ from X to Y as a *rule* which to each element $x \in X$ assigns an element $y = f(x)$ in Y. If $f(x) \neq f(y)$ whenever $x \neq y$, we call the function *injective* (or *one-to-one*). If there for each $y \in Y$ is an $x \in X$ such that $f(x) = y$, the function is called *surjective* (or *onto*). A function which is both injective and surjective, is called *bijective* — it establishes a one-to-one correspondence between the elements of X and Y.

If A is subset of X, the set $f(A) \subseteq Y$ defined by

$$f(A) = \{f(a) \mid a \in A\}$$

is called the *image of A under f*. If B is subset of Y, the set $f^{-1}(B) \subseteq X$ defined by

$$f^{-1}(B) = \{x \in X \mid f(x) \in B\}$$

is called the *inverse image of B under f*. In analysis, images and inverse images of sets play important parts, and it is useful to know how these operations relate to the boolean operations of union and intersection. Let us begin with the good news.

Proposition 1.4.1 *Let $\mathcal{B}$ be a family of subset of Y. Then for all functions $f : X \to Y$ we have*

$$f^{-1}(\bigcup_{B\in\mathcal{B}} B) = \bigcup_{B\in\mathcal{B}} f^{-1}(B) \quad \text{and} \quad f^{-1}(\bigcap_{B\in\mathcal{B}} B) = \bigcap_{B\in\mathcal{B}} f^{-1}(B)$$

We say that inverse images commute *with arbitrary unions and intersections.*

Proof: I prove the first part; the second part is proved similarly. Assume first that $x \in f^{-1}(\bigcup_{B\in\mathcal{B}} B)$. This means that $f(x) \in \bigcup_{B\in\mathcal{B}} B$, and consequently there must be at least one $B \in \mathcal{B}$ such that $f(x) \in B$. But then $x \in f^{-1}(B)$, and hence $x \in \bigcup_{B\in\mathcal{B}} f^{-1}(B)$. This proves that $f^{-1}(\bigcup_{B\in\mathcal{B}} B) \subseteq \bigcup_{B\in\mathcal{B}} f^{-1}(B)$.

To prove the opposite inclusion, assume that $x \in \bigcup_{B\in\mathcal{B}} f^{-1}(B)$. There must be at least one $B \in \mathcal{B}$ such that $x \in f^{-1}(B)$, and hence $f(x) \in B$.

This implies that $f(x) \in \bigcup_{B\in\mathcal{B}} B$, and hence $x \in f^{-1}(\bigcup_{B\in\mathcal{B}} B)$. □

For forward images the situation is more complicated:

Proposition 1.4.2 *Let $\mathcal{A}$ be a family of subset of X. Then for all functions $f : X \to Y$ we have*

$$f(\bigcup_{A\in\mathcal{A}} A) = \bigcup_{A\in\mathcal{A}} f(A) \quad \textit{and} \quad f(\bigcap_{A\in\mathcal{A}} A) \subseteq \bigcap_{A\in\mathcal{A}} f(A)$$

In general, we do not *have equality in the last case. Hence forward images commute with unions, but not always with intersections.*

Proof: To prove the statement about unions, we first observe that since $A \subseteq \bigcup_{A\in\mathcal{A}} A$ for all $A \in \mathcal{A}$, we have $f(A) \subseteq f(\bigcup_{A\in\mathcal{A}} A)$ for all such A. Since this inclusion holds for all A, we must also have $\bigcup_{A\in\mathcal{A}} f(A) \subseteq f(\bigcup_{A\in\mathcal{A}})$. To prove the opposite inclusion, assume that $y \in f(\bigcup_{A\in\mathcal{A}} A)$. This means that there exists an $x \in \bigcup_{A\in\mathcal{A}} A$ such that $f(x) = y$. This x has to belong to at least one $A \in \mathcal{A}$, and hence $y \in f(A) \subseteq \bigcup_{A\in\mathcal{A}} f(A)$.

To prove the inclusion for intersections, just observe that since $\bigcap_{A\in\mathcal{A}} A \subseteq A$ for all $A \in \mathcal{A}$, we must have $f(\bigcap_{A\in\mathcal{A}} A) \subseteq f(A)$ for all such A. Since this inclusion holds for all A, it follows that $f(\bigcap_{A\in\mathcal{A}} A) \subseteq \bigcap_{A\in\mathcal{A}} f(A)$. The example below shows that the opposite inclusion does not always hold. □

Example 1: Let $X = \{x_1, x_2\}$ and $Y = \{y\}$. Define $f : X \to Y$ by $f(x_1) = f(x_2) = y$, and let $A_1 = \{x_1\}, A_2 = \{x_2\}$. Then $A_1 \cap A_2 = \emptyset$ and consequently $f(A_1 \cap A_2) = \emptyset$. On the other hand $f(A_1) = f(A_2) = \{y\}$, and hence $f(A_1) \cap f(A_2) = \{y\}$. This means that $f(A_1 \cap A_2) \neq f(A_1) \cap f(A_2)$. ♣

The problem in this example stems from the fact that y belongs to both $f(A_1)$ and $f(A_2)$, but only as the image of two *different* elements $x_1 \in A_1$ og $x_2 \in A_2$; there is no *common* element $x \in A_1 \cap A_2$ which is mapped to y. This problem disappears if f is injective:

Corollary 1.4.3 *Let $\mathcal{A}$ be a family of subset of X. Then for all* injective *functions $f : X \to Y$ we have*

$$f(\bigcap_{A\in\mathcal{A}} A) = \bigcap_{A\in\mathcal{A}} f(A)$$

Proof: The easiest way to show this is probably to apply Proposition 2 to the inverse function of f, but I choose instead to prove the missing inclusion $f(\bigcap_{A\in\mathcal{A}} A) \supseteq \bigcap_{A\in\mathcal{A}} f(A)$ directly.

Assume $y \in \bigcap_{A\in\mathcal{A}} f(A)$. For each $A \in \mathcal{A}$ there must be an element $x_A \in A$ such that $f(x_A) = y$. Since f is injective, all these $x_A \in A$ must be the same element x, and hence $x \in A$ for all $A \in \mathcal{A}$. This means that $x \in \bigcap_{A\in\mathcal{A}} A$, and since $y = f(x)$, we have proved that $y \in f(\bigcap_{A\in\mathcal{A}} A)$. □

Taking complements is another operation that commutes with inverse images, but not (in general) with forward images.

Proposition 1.4.4 *Assume that $f : X \to Y$ is a function and that $B \subseteq Y$. Then $f^{-1}(B^c)) = (f^{-1}(B))^c$. (Here, of course, $B^c = Y \setminus B$ is the complement with respect to the universe Y, while $(f^{-1}(B))^c = X \setminus f^{-1}(B)$ is the complemet with respect to the universe X).*

Proof: An element $x \in X$ belongs to $f^{-1}(B^c)$ if and only if $f(x) \in B^c$. On the other hand, it belongs to $(f^{-1}(B))^c$ if and only if $f(x) \notin B$, i.e. iff $f(x) \in B^c$. □

Finally, let us just observe that being disjoint is also a property that is conserved under inverse images; if $A \cap B = \emptyset$, then $f^{-1}(A) \cap f^{-1}(B) = \emptyset$. Again the corresponding property for forward images does not hold in general.

Exercises for Section 1.4

1. Let $f : \mathbb{R} \to \mathbb{R}$ be the function $f(x) = x^2$. Find $f([-1, 2])$ and $f^{-1}([-1, 2])$.
2. Let $g : \mathbb{R}^2 \to \mathbb{R}$ be the function $g(x, y) = x^2 + y^2$. Find $f([-1, 1] \times [-1, 1])$ and $f^{-1}([0, 4])$.
3. Show that a strictly increasing function $f : \mathbb{R} \to \mathbb{R}$ is injective. Does it have to be surjective?
4. Prove the second part of Proposition 1.4.1.
5. Find a function $f : X \to Y$ and a set $A \subseteq X$ such that we have neither $f(A^c) \subseteq f(A)^c$ nor $f(A)^c \subseteq f(A^c)$.
6. Show that if $f : X \to Y$ and $g : Y \to Z$ are injective, then $g \circ f : X \to Z$ is injective.
7. Show that if $f : X \to Y$ and $g : Y \to Z$ are surjective, then $g \circ f : X \to Z$ is surjective.

1.5 Relations and partitions

In mathematics there are lots of relations between objects; numbers may be smaller or larger than each other, lines may be parallell, vectors may be orthogonal, matrices may be similar and so on. Sometimes it is convenient to have an abstract definition of what we mean by a relation.

Definition 1.5.1 *By a* relation *on a set X, we mean a subset R of the cartesian product $X \times X$. We usually write xRy instead of $(x, y) \in R$ to denote that x and y are related. The symbols $\sim$ and $\equiv$ are often used to denote relations, and we then write $x \sim y$ and $x \equiv y$.*

At first glance this definition may seem strange as very few people think of relations as subsets of $X \times X$, but a little thought will convince you that it gives us a convenient starting point, especially if I add that in practice relations are rarely arbitrary subsets of $X \times X$, but have much more structure than the definition indicates. We shall take a look at one such class of relations, the *equivalence relations.* Equivalence relations are used to partition sets into subsets, and from a pedagogical point of view, it is probably better to start with the related notion of a partition.

Informally, a partition is what we get if we divide a set into non-overlapping pieces. More precisely, If X is a set, a *partition* $\mathcal{P}$ of X is a family of subset of X such that each element in x belongs to exactly one set $P \in \mathcal{P}$. The elements P of $\mathcal{P}$ are called *partition classes* of $\mathcal{P}$.

Given a partition of X, we may introduce a relation $\sim$ on X by

$$x \sim y \iff x \text{ and } y \text{ belong to the same set } P \in \mathcal{P}$$

It is easy to check that $\sim$ has the following three properties:

(i) $x \sim x$ for all $x \in X$.

(ii) If $x \sim y$, then $y \sim x$.

(iii) If $x \sim y$ and $y \sim z$, then $x \sim z$.

We say that $\sim$ is the relation *induced by* the partition $\mathcal{P}$.

Let us now turn the tables around and start with a relation on X satisfying conditions (i)-(iii):

Definition 1.5.2 *An* equivalence relation *on X is a relation $\sim$ satisfying the following conditions:*

(i) Reflexivity*: $x \sim x$ for all $x \in X$,*

(ii) Symmetry*: If $x \sim y$, then $y \sim x$,*

(iii) Transitivity*: If $x \sim y$ and $y \sim z$, then $x \sim z$.*

Given an equivalence relation $\sim$ on X, we may for each $x \in X$ define the *equivalence class* $[x]$ of x by:

$$[x] = \{y \in X \mid x \sim y\}$$

The following result tells us that there is a one-to-one correspondence between partitions and equivalence relations — all partitions induce an equivalence relation, and all equivalence relations define a partition

Proposition 1.5.3 *If $\sim$ is an equivalence relation on X, the collection of equivalence classes*

$$\mathcal{P} = \{[x] : x \in X\}$$

is a partition of X.

Proof: We have to prove that each x in X belongs to exactly one equivalence class. We first observe that since $x \sim x$ by (i), $x \in [x]$ and hence belongs to at least one equivalence class. To finish the proof, we have to show that if $x \in [y]$ for some other element $y \in X$, then $[x] = [y]$.

We first prove that $[y] \subseteq [x]$. To this end assume that $z \in [y]$. By definition, this means that $y \sim z$. On the other hand, the assumption that $x \in [y]$ means that $y \sim x$, which by (ii) implies that $x \sim y$. We thus have $x \sim y$ and $y \sim z$, which by (iii) means that $x \sim z$. Thus $z \in [x]$, and hence we have proved that $[y] \subseteq [x]$.

The opposite inclusion $[x] \subseteq [y]$ is proved similarly: Assume that $z \in [x]$. By definition, this means that $x \sim z$. On the other hand, the assumption that $x \in [y]$ means that $y \sim x$. We thus have $y \sim x$ and $x \sim z$, which by (iii) implies that $y \sim z$. Thus $z \in [y]$, and we have proved that $[x] \subseteq [y]$. □

The main reason why this theorem is useful is that it is often more natural to describe situations through equivalence relations than through partitions. The following example assumes that you remember a little linear algebra:

Example 1.5.3: Let V be a vector space and U a subspace. Define a relation on V by

$$x \sim y \Longleftrightarrow x - y \in U$$

Let us show that $\sim$ is an equivalence relation by checking the three conditions (i)-(iii) in the definition:

(i) Since $x - x = 0 \in U$, we see that $x \sim x$ for all $x \in V$.

(ii) Assume that $x \sim y$. This means that $x - y \in U$, and hence $y - x = -(x - y) \in U$ since subspaces are closed under multiplication by scalars. This means that $y \sim x$.

(iii) If $x \sim y$ and $y \sim z$, then $x - y \in U$ and $y - z \in U$. Since subspaces are closed under addition, this means that $x - z = (x - y) + (y - z) \in U$, and hence $x \sim z$.

As we have now proved that $\sim$ is an equivalence relation, the equivalence classes of $\sim$ form a partition of V. ♣

.

If $\sim$ is an equivalence relation on X, we let $X/\sim$ denote the set of all equivalence classes of $\sim$. Such *quotient constructions* are common in all parts of mathematics, and you will see a few examples in this book.

Exercises to Section 1.5

1. Let $\mathcal{P}$ be a partition of a set A, and define a relation $\sim$ on A by

   $$x \sim y \Longleftrightarrow x \text{ and } y \text{ belong to the same set } P \in \mathcal{P}$$

   Check that $\sim$ is an equivalence relation.

2. Assume that $\mathcal{P}$ is the partition defined by an equivilance relation $\sim$. Show that $\sim$ is the equivalence relation induced by $\mathcal{P}$.

3. Let $\mathcal{L}$ be the collection of all lines in the plane. Define a relation on $\mathcal{L}$ by saying that two lines are equivalent if and only if they are parallel. Show that this an equivalence relation on $\mathcal{L}$.

4. Define a relation on $\mathbb{C}$ by

   $$z \sim y \Longleftrightarrow |z| = |w|$$

   Show that $\sim$ is an equivalence relation. What does the equivalence classes look like?

5. Define a relation $\sim$ on $\mathbb{R}^3$ by

   $$(x, y, z) \sim (x', y', z') \iff 3x - y + 2z = 3x' - y' + 2z'$$

   Show that $\sim$ is an equivalence relation and describe the equivalence classes of $\sim$.

6. Let m be a natural number. Define a relation $\equiv$ on $\mathbb{Z}$ by

   $$x \equiv y \Longleftrightarrow x - y \text{ is divisible by } m$$

   Show that $\equiv$ is an equivalence relation on $\mathbb{Z}$. How many partition classes are there, and what do they look like?

7. Let $\mathcal{M}$ be the set of all $n \times n$ matrices. Define a relation on $\sim$ on $\mathcal{M}$ by

   $$A \sim B \Longleftrightarrow \text{if there exists an invertible matrix } P \text{ such that } A = P^{-1}BP$$

   Show that $\sim$ is an equivalence relation.

1.6 Countability

A set A is called *countable* if it possible to make a list $a_1, a_2, \ldots, a_n, \ldots$ which contains all elements of A. Finite sets $A = \{a_1, a_2, \ldots, a_m\}$ are obviously countable[1] as they can be listed

$$a_1, a_2, \ldots, a_m, a_m, a_m, \ldots$$

[1] Some books exclude the finite sets from the countable and treat them as a separate category, but that would be impractical for our purposes.

(you may list the same elements many times). The set $\mathbb{N}$ of all natural numbers is also countable as it is automatically listed by

$$1, 2, 3, \ldots$$

It is a little less obvious that the set $\mathbb{Z}$ of all integers is countable, but we may use the list

$$0, 1, -1, 2, -2, 3, -3 \ldots$$

It is also easy to see that a subset of a countable set must be countable, and that the image $f(A)$ of a countable set is countable (if $\{a_n\}$ is a listing of A, then $\{f(a_n)\}$ is a listing of $f(A)$).

The next result is perhaps more surprising:

Proposition 1.6.1 *If the sets A, B are countable, so is the cartesian product $A \times B$.*

Proof: Since A and B are countable, there are lists $\{a_n\}$, $\{b_n\}$ containing all the elements of A and B, respectively. But then

$$\{(a_1, b_1), (a_2, b_1), (a_1, b_2), (a_3, b_1), (a_2, b_2), (a_1, b_3), (a_4, b_1), (a_3, b_2), \ldots, \}$$

is a list containing all elements of $A \times B$ (observe how the list is made; first we list the (only) element (a_1, b_1) where the indicies sum to 2, then we list the elements $(a_2, b_1), (a_1, b_2)$ where the indicies sum to 3, then the elements $(a_3, b_1), (a_2, b_2), (a_1, b_3)$ where the indicies sum to 4 etc.) □

Remark If $A_1, A_2, \ldots, A_n$ is a finite collection of countable sets, then the cartesian product $A_1 \times A_2 \times \cdots \times A_n$ is countable. This can be proved by induction from the Proposition above, using that $A_1 \times \cdots \times A_k \times A_{k+1}$ is essentially the same set as $(A_1 \times \cdots \times A_k) \times A_{k+1}$.

The same trick we used to prove Proposition 1.6.1, can also be used to prove the next result:

Proposition 1.6.2 *If the sets $A_1, A_2, \ldots, A_n, \ldots$ are countable, so is their union $\bigcup_{n\in\mathbb{N}} A_n$. Hence a countable union of countable sets is itself countable.*

Proof: Let $A_i = \{a_{i1}, a_{i2}, \ldots, a_{in}, \ldots\}$ be a listing of the i-th set. Then

$$\{a_{11}, a_{21}, a_{12}, a_{31}, a_{22}, a_{13}, a_{41}, a_{32}, \ldots\}$$

is a listing of $\bigcup_{i\in\mathbb{N}} A_i$. □

Proposition 1.6.1 can also be used to prove that the rational numbers are countable:

Proposition 1.6.3 *The set $\mathbb{Q}$ of all rational numbers is countable.*

Proof: According to Proposition 1.6.1, the set $\mathbb{Z} \times \mathbb{N}$ is countable and can be listed $(a_1, b_1), (a_2, b_2), (a_3, b_3), \ldots$. But then $\frac{a_1}{b_1}, \frac{a_2}{b_2}, \frac{a_3}{b_3}, \ldots$ is a list of all the elements in $\mathbb{Q}$ (due to cancellations, all rational numbers will appear infinitely many times in this list, but that doesn't matter). □

Finally, we prove an important result in the opposite direction:

Theorem 1.6.4 *The set $\mathbb{R}$ of all real numbers is* not *countable.*

Proof: (Cantor's diagonal argument) Assume for contradiction that $\mathbb{R}$ is countable and can be listed $r_1, r_2, r_3, \ldots$. Let us write down the decimal expansions of the numbers on the list:

$$\begin{array}{rcl} r_1 & = & w_1.a_{11}a_{12}a_{13}a_{14}\ldots \\ r_2 & = & w_2.a_{21}a_{22}a_{23}a_{24}\ldots \\ r_3 & = & w_3.a_{31}a_{32}a_{33}a_{34}\ldots \\ r_4 & = & w_4.a_{41}a_{42}a_{43}a_{44}\ldots \\ \vdots & \vdots & \qquad \vdots \end{array}$$

(w_i is the integer part of r_i, and $a_{i1}, a_{i2}, a_{i3}, \ldots$ are the decimals). To get our contradiction, we introduce a new decimal number $c = 0.c_1c_2c_3c_4\ldots$ where the decimals are defined by:

$$c_i = \begin{cases} 1 & \text{if } a_{ii} \neq 1 \\ \\ 2 & \text{if } a_{ii} = 1 \end{cases}$$

This number has to be different from the i-th number r_i on the list as the decimal expansions disagree on the i-th place (as c has only 1 and 2 as decimals, there are no problems with nonuniqueness of decimal expansions). This is a contradiction as we assumed that all real numbers were on the list. □

Exercises to Section 1.6

1. Show that a subset of a countable set is countable.
2. Show that if $A_1, A_2, \ldots A_n$ are countable, then $A_1 \times A_2 \times \cdots A_n$ is countable.
3. Show that the set of all finite sequences $(q_1, q_2, \ldots, q_k)$, $k \in \mathbb{N}$, of rational numbers is countable.
4. Show that if A is an *infinite*, countable set, then there is a list $a_1, a_2, a_3, \ldots$ which only contains elements in A and where each element in A appears only once. Show that if A and B are two infinite, countable sets, there is a bijection (i.e. an injective and surjective function) $f : A \to B$.

5. Show that the set of all subsets of $\mathbb{N}$ is *not* countable (*Hint:* Try to modify the proof of Theorem 1.6.4.)

Chapter 2

Metric Spaces

Many of the arguments you have seen in several variable calculus are almost identical to the corresponding arguments in one variable calculus, especially arguments concerning convergence and continuity. The reason is that the notions of convergence and continuity can be formulated in terms of distance, and that the notion of distance between numbers that you need in the one variable theory, is very similar to the notion of distance between points or vectors that you need in the theory of functions of severable variables. In more advanced mathematics, we need to find the distance between more complicated objects than numbers and vectors, e.g. between sequences, sets and functions. These new notions of distance leads to new notions of convergence and continuity, and these again lead to new arguments suprisingly similar to those we have already seen in one and several variable calculus.

After a while it becomes quite boring to perform almost the same arguments over and over again in new settings, and one begins to wonder if there is general theory that covers all these examples — is it possible to develop a general theory of distance where we can prove the results we need once and for all? The answer is yes, and the theory is called the theory of metric spaces.

A metric space is just a set X equipped with a function d of two variables which measures the distance between points: $d(x, y)$ is the distance between two points x and y in X. It turns out that if we put mild and natural conditions on the function d, we can develop a general notion of distance that covers distances between numbers, vectors, sequences, functions, sets and much more. Within this theory we can formulate and prove results about convergence and continuity once and for all. The purpose of this chapter is to develop the basic theory of metric spaces. In later chapters we shall meet some of the applications of the theory.

2.1 Definitions and examples

As already mentioned, a metric space is just a set X equipped with a function $d : X \times X \to \mathbb{R}$ which measures the distance $d(x, y)$ beween points $x, y \in X$. For the theory to work, we need the function d to have properties similar to the distance functions we are familiar with. So what properties do we expect from a measure of distance?

First of all, the distance $d(x, y)$ should be a nonnegative number, and it should only be equal to zero if $x = y$. Second, the distance $d(x, y)$ from x to y should equal the distance $d(y, x)$ from y to x. Note that this is not always a reasonable assumption — if we, e.g., measure the distance from x to y by the time it takes to walk from x to y, $d(x, y)$ and $d(y, x)$ may be different — but we shall restrict ourselves to situations where the condition is satisfied. The third condition we shall need, says that the distance obtained by going directly from x to y, should always be less than or equal to the distance we get when we go via a third pont z, i.e.

$$d(x, y) \leq d(x, z) + d(z, x)$$

It turns out that these conditions are the only ones we need, and we sum them up in a formal definition.

Definition 2.1.1 *A* metric space (X, d) *consists of a non-empty set* X *and a function* $d : X \times X \to [0, \infty)$ *such that:*

(i) (Positivity) For all $x, y \in X$, $d(x, y) \geq 0$ *with equality if and only if* $x = y$.

(ii) (Symmetry) For all $x, y \in X$, $d(x, y) = d(y, x)$.

(iii) (Triangle inequality) For all $x, y, z \in X$

$$d(x, y) \leq d(x, z) + d(z, y)$$

A function d *satisfying conditions (i)-(iii), is called a* metric *on* X.

Comment: When it is clear – or irrelevant – which metric d we have in mind, we shall often refer to "the metric space X" rather than "the metric space (X, d)".

Let us take a look at some examples of metric spaces.

Example 1: If we let $d(x, y) = |x-y|$, $(\mathbb{R}, d)$ is a metric space. The first two conditions are obviously satisfied, and the third follows from the ordinary triangle inequality for real numbers:

$$d(x, y) = |x - y| = |(x - z) + (z - y)| \leq |x - z| + |z - y| = d(x, z) + d(z, y)$$

Example 2: If we let $d(\mathbf{x}, \mathbf{y}) = |\mathbf{x} - \mathbf{y}|$, then $(\mathbb{R}^n, d)$ is a metric space. The first two conditions are obviously satisfied, and the third follows from the triangle inequality for vectors the same way as above :

$$d(\mathbf{x}, \mathbf{y}) = |\mathbf{x} - \mathbf{y}| = |(\mathbf{x} - \mathbf{z}) + (\mathbf{z} - \mathbf{y})| \leq |\mathbf{x} - \mathbf{z}| + |\mathbf{z} - \mathbf{y}| = d(\mathbf{x}, \mathbf{z}) + d(\mathbf{z}, \mathbf{y})$$

Example 3: Assume that we want to move from one point $\mathbf{x} = (x_1, x_2)$ in the plane to another $\mathbf{y} = (y_1, y_2)$, but that we are only allowed to move horizontally and vertically. If we first move horizontally from (x_1, x_2) to (y_1, x_2) and then vertically from (y_1, x_2) to (y_1, y_2), the total distance is

$$d(\mathbf{x}, \mathbf{y}) = |y_1 - x_1| + |y_2 - x_2|$$

This gives us a metric on $\mathbb{R}^2$ which is different from the usual metric in Example 2. It is ofte referred to as the *Manhattan metric* or the *taxi cab metric*.

Also in this cas the first two conditions of a metric space are obviously satisfied. To prove the triangle inequality, observe that for any third point $\mathbf{z} = (z_1, z_2)$, we have

$$d(\mathbf{x}, \mathbf{y}) = |y_1 - x_1| + |y_2 - x_1| =$$

$$= |(y_1 - z_1) + (z_1 - x_1)| + |(y_2 - z_2) + (z_2 - x_2)| \leq$$

$$\leq |y_1 - z_1| + |z_1 - x_1| + |y_2 - z_2| + |z_2 - x_2| =$$

$$= |z_1 - x_1| + |z_2 - x_2| + |y_1 - z_1| + |y_2 - z_2| =$$

$$= d(x, z) + d(z, y)$$

where we have used the ordinary triangle inequality for real numbers to get from the second to the third line. ♣

Example 4: We shall now take a look at an example of a different kind. Assume that we want to send messages in a language with N symbols (letters, numbers, punctuation marks, space, etc.) We assume that all messages have the same length K (if they are too short or too long, we either fill them out or break them into pieces). We let X be the set of all messages, i.e. all sequences of symbols from the language of length K. If $\mathbf{x} = (x_1, x_2, \ldots, x_K)$ and $\mathbf{y} = (y_1, y_2, \ldots, y_K)$ are two messages, we define

$$d(\mathbf{x}, \mathbf{y}) = \text{the number of indices } n \text{ such that } x_n \neq y_n$$

It is not hard to check that d is a metric. It is usually referred to as the *Hamming-metric*, and is much used in coding theory where it serves as a measure of how much a message gets distorted during transmission. ♣

Example 5: There are many ways to measure the distance between functions, and in this example we shall look at some. Let X be the set of all continuous functions $f : [a, b] \to \mathbb{R}$. Then

$$d_1(f, g) = \sup\{|f(x) - g(x)| \ : \ x \in [a, b]\}$$

is a metric on X. This metric determines the distance beween two functions by measuring the distance at the x-value where the graphs are most apart. This means that the distance between two functions may be large even if the functions in average are quite close. The metric

$$d_2(f, g) = \int_a^b |f(x) - g(x)|\, dx$$

instead sums up the distance between $f(x)$ og $g(x)$ at all points. A third popular metric is

$$d_3(f, g) = \left(\int_a^b |f(x) - g(x)|^2\, dx\right)^{\frac{1}{2}}$$

This metric is a generalization of the usual (*euclidean*) metric in $\mathbb{R}^n$:

$$d(\mathbf{x}, \mathbf{y}) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2} = \left(\sum_{i=1}^{n}(x_i - y_i)^2\right)^{\frac{1}{2}}$$

(think of the integral as a generalized sum). That we have more than one metric on X, doesn't mean that one of them is "right" and the others "wrong", but that they are useful for different purposes. ♣

Example 6: The metrics in this example may seem rather strange. Although they are not very useful in applications, they are important to know about as they are totally different from the metrics we are used to from $\mathbb{R}^n$ and may help sharpen our intuition of how a metric can be. Let X be any non-empty set, and define:

$$d(x, y) = \begin{cases} 0 & \text{if } x = y \\ \\ 1 & \text{if } x \neq y \end{cases}$$

It is not hard to check that d is a metric on X, usually referred to as the *discrete* metric. ♣

Example 7: There are many ways to make new metric spaces from old. The simplest is the subspace metric: If (X, d) is a metric space and A is a non-empty subset of X, we can make a metric d_A on A by putting

$d_A(x,y) = d(x,y)$ for all $x, y \in A$ — we simply restrict the metric to A. It is trivial to check that d_A is a metric on A. In practice, we rarely bother to change the name of the metric and refer to d_A simply as d, but remember in the back of our head that d is now restricted to A. ♣

There are many more types of metric spaces than we have seen so far, but the hope is that the examples above will give you a certain impression of the variety of the concept. In the next section we shall see how we can define convergence and continuity for sequences and functions in metric spaces. When we prove theorems about these concepts, they automatically hold in all metric spaces, saving us the labor of having to prove them over and over again each time we introduce a new class of spaces.

An important question is when two metric spaces (X, d_X) and (Y, d_Y) are the same. The easy answer is to say that we need the sets X, Y and the functions d_X, d_Y to be equal. This is certainly correct if one interprets "being the same" in the strictest sense, but it is often more appropriate to use a looser definition — in mathematics we are usually not interested in what the elements of a set are, but only in the relationship between them (you may, e.g., want to ask yourself what the natural number 3 "is").

An *isometry* between two metric spaces is a bijection which preserves what is important for metric spaces: the distance between points. More precisely:

Definition 2.1.2 *Assume that (X, d_X) and (Y, d_Y) are metric spaces. An* isometry *from (X, d_X) to (Y, d_Y) is a bijection $i : X \to Y$ such that $d_X(x,y) = d_Y(i(x), i(y))$ for all $x, y \in X$. We say that (X, d_X) and (Y, d_Y) are* isometric *if there exists an isometry from (X, d_X) to (Y, d_Y).*

In many situations it is convenient to think of two metric spaces as "the same" if they are isometric. Note that if i is an isometry from (X, d_X) to (Y, d_Y), then the inverse i^{-1} is an isometry from (Y, d_Y) to (X, d_X), and hence being isometric is a symmetric relation.

A map which preserves distance, but does not necessarily hit all of Y, is called an *embedding*:

Definition 2.1.3 *Assume that (X, d_X) and (Y, d_Y) are metric spaces. An* embedding *of (X, d_X) into (Y, d_Y) is an injection $i : X \to Y$ such that $d_X(x,y) = d_Y(i(x), i(y))$ for all $x, y \in X$.*

Note that an embedding i can be regarded as an isometry between X and its image $i(X)$.

We end this section with an important consequence of the triangle inequality.

Proposition 2.1.4 (Inverse Triangle Inequality) *For all elements x, y, z in a metric space (X, d), we have*

$$|d(x, y) - d(x, z)| \leq d(y, z)$$

Proof: Since the absolute value $|d(x, y) - d(x, z)|$ is the largest of the two numbers $d(x, y) - d(x, z)$ and $d(x, z) - d(x, y)$, it suffices to show that they are both less than or equal to $d(y, z)$. By the triangle inequality

$$d(x, y) \leq d(x, z) + d(z, y)$$

and hence $d(x, y) - d(x, z) \leq d(z, y) = d(y, z)$. To get the other inequality, we use the triangle inequality again,

$$d(x, z) \leq d(x, y) + d(y, z)$$

and hence $d(x, z) - d(x, y) \leq d(y, z)$. □

Exercises for Section 2.1

1. Show that (X, d) in Example 4 is a metric space.
2. Show that (X, d_1) in Example 5 is a metric space.
3. Show that (X, d_2) in Example 5 is a metric space.
4. Show that (X, d) in Example 6 is a metric space.
5. A sequence $\{x_n\}_{n\in\mathbb{N}}$ of real numbers is called *bounded* if there is a number $M \in \mathbb{R}$ such that $|x_n| \leq M$ for all $n \in \mathbb{N}$. Let X be the set of all bounded sequences. Show that

$$d(\{x_n\}, \{y_n\}) = \sup\{|x_n - y_n| \ : n \in \mathbb{N}\}$$

   is a metric on X.
6. If V is a (real) vector space, a function $\| \cdot \| : V \to \mathbb{R}$ is called a *norm* if the following conditions are satisfied:

   (i) For all $x \in V$, $\|x\| \geq 0$ with equality if and only if $x = 0$.

   (ii) $\|\alpha x\| = |\alpha|\|x\|$ for all $\alpha \in \mathbb{R}$ and all $x \in V$.

   (iii) $\|x + y\|| \leq \|x\| + \|y\|$ for all $x, y \in V$.

   Show that if $\| \cdot \|$ is a norm, then $d(x, y) = \|x - y\|$ defines a metric on V.
7. Show that for vectors $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^m$,

$$|\,|\mathbf{x} - \mathbf{y}| - |\mathbf{x} - \mathbf{z}|\,| \leq |\mathbf{y} - \mathbf{z}|$$

8. Assume that d_1 og d_2 are two metrics on X. Show that

$$d(x, y) = d_1(x, y) + d_2(x, y)$$

   is a metric on X.

9. Assume that (X, d_X) and (Y, d_Y) are two metric spaces. Define a function
$$d : (X \times Y) \times (X \times Y) \to \mathbb{R}$$
by
$$d((x_1, y_1), (x_2, y_2)) = d_X(x_1, x_2) + d_Y(y_1, y_2)$$
Show that d is a metric on $X \times Y$.

10. Let X be a non-empty set, and let $\rho : X \times X \to \mathbb{R}$ be a function satisfying:
   (i) $\rho(x, y) \geq 0$ with equality if and only if $x = y$.
   (ii) $\rho(x, y) \leq \rho(x, z) + \rho(z, y)$ for all $x, y, z \in X$.

   Define $d : X \times X \to \mathbb{R}$ by
$$d(x, y) = \max\{\rho(x, y), \rho(y, x)\}$$
Show that d is a metric on X.

11. Let $a \in \mathbb{R}$. Show that the function $f(x) = x + a$ is an isometry from $\mathbb{R}$ to $\mathbb{R}$.

12. Recall that an $n \times n$ matrix U is *orthogonal* if $U^{-1} = U^T$. Show that if U is orthogonal and $\mathbf{b} \in \mathbb{R}^n$, then the mapping $i : \mathbb{R}^n \to \mathbb{R}^n$ given by $i(\mathbf{x}) = U\mathbf{x} + \mathbf{b}$ is an isometry.

2.2 Convergence and continuity

We begin our study of metric spaces by defining convergence. A sequence $\{x_n\}$ in a metric space X is just an ordered collection $\{x_1, x_2, x_3, \ldots, x_n, \ldots\}$ of elements in X enumerated by the natural numbers.

Definition 2.2.1 *Let (X, d) be a metric space. A sequencee $\{x_n\}$ in X* converges *to a point $a \in X$ if there for every $\epsilon > 0$ exists an $N \in \mathbb{N}$ such that $d(x_n, a) < \epsilon$ for all $n \geq N$. We write $\lim_{n\to\infty} x_n = a$ or $x_n \to a$.*

Note that this definition exactly mimics the definition of convergence in $\mathbb{R}$ og $\mathbb{R}^n$. Here is an alternative formulation.

Lemma 2.2.2 *A sequence $\{x_n\}$ in a metric space (X, d) converges to a if and only if $\lim_{n\to\infty} d(x_n, a) = 0$.*

Proof: The distances $\{d(x_n, a)\}$ form a sequence of nonnegative numbers. This sequence converges to 0 if and only if there for every $\epsilon > 0$ exists an $N \in \mathbb{N}$ such that $d(x_n, a) < \epsilon$ when $n \geq N$. But this is exactly what the definition above says. □

May a sequence converge to more than one point? We know that it cannot in $\mathbb{R}^n$, but some of these new metric spaces are so strange that we can not be certain without a proof.

Proposition 2.2.3 *A sequence in a metric space can not converge to more than one point.*

Proof: Assume that $\lim_{n\to\infty} x_n = a$ and $\lim_{n\to\infty} x_n = b$. We must show that this is only possible if $a = b$. According to the triangle inequality

$$d(a,b) \leq d(a,x_n) + d(x_n,b)$$

Taking limits, we get

$$d(a,b) \leq \lim_{n\to\infty} d(a,x_n) + \lim_{n\to\infty} d(x_n,b) = 0 + 0 = 0$$

Consequently, $d(a,b) = 0$, and according to point (i) (positivity) in the definition of metric spaces, $a = b$. □

Note how we use the conditions in Definition 2.1.1 in the proof above. So far they are all we know about metric spaces. As the theory develops, we shall get more and more tools to work with.

We can also phrase the notion of convergence in more geometric terms. If a is an element of a metric space X, and r is a positive number, the (open) *ball centered at a with radius r* is the set

$$\mathrm{B}(a;r) = \{x \in X \mid d(x,a) < r\}$$

As the terminology suggests, we think of $\mathrm{B}(a;r)$ as a ball around a with radius r. Note that $x \in \mathrm{B}(a;r)$ means exactly the same as $d(x,a) < r$.

The definition of convergence can now be rephrased by saying that $\{x_n\}$ converges to a if the elements of the sequence $\{x_n\}$ eventually end up inside any ball $\mathrm{B}(a;\epsilon)$ around a.

Let us now see how we can define continuity in metric spaces.

Definition 2.2.4 *Assume that (X,d_X), (Y,d_Y) are two metric spaces. A function $f : X \to Y$ is* continuous *at a point $a \in X$ if for every $\epsilon > 0$ there is a $\delta > 0$ such that $d_Y(f(x),f(a)) < \epsilon$ whenever $d_X(x,a) < \delta$.*

This definition says exactly the same as as the usual definitions of continuity for functions of one or several variables; we can get the distance between $f(x)$ and $f(a)$ smaller than ϵ by choosing x such that the distance between x and a is smaller than δ. The only difference is that we are now using the metrics d_X og d_Y to measure the distances.

A more geometric formulation of the definition is to say that for any open ball $\mathrm{B}(f(a);\epsilon)$ around $f(a)$, there is an open ball $\mathrm{B}(a,\delta)$ around a such that $f(\mathrm{B}(a;\delta)) \subseteq \mathrm{B}(f(a);\epsilon)$ (make a drawing!).

There is a close connection between continuity and convergence which reflects our intuitive feeling that f is continuous at a point a if $f(x)$ approaches $f(a)$ whenever x approaches a.

Proposition 2.2.5 *The following are equivalent for a function $f : X \to Y$ between metric spaces:*

(i) *f is continuous at a point $a \in X$.*

(ii) *For all sequences $\{x_n\}$ converging to a, the sequence $\{f(x_n)\}$ converges to $f(a)$.*

Proof: $(i) \implies (ii)$: We must show that for any $\epsilon > 0$, there is an $N \in \mathbb{N}$ such that $d_Y(f(x_n), f(a)) < \epsilon$ when $n \geq N$. Since f is continuous at a, there is a $\delta > 0$ such that $d_Y(f(x_n), f(a)) < \epsilon$ whenever $d_X(x, a) < \delta$. Since x_n converges to a, there is an $N \in \mathbb{N}$ such that $d_X(x_n, a) < \delta$ when $n \geq N$. But then $d_Y(f(x_n), f(a)) < \epsilon$ for all $n \geq N$.

$(ii) \implies (i)$ We argue contrapositively: Assume that f is not continuous at a. We shall show that there is a sequence $\{x_n\}$ converging to a such that $\{f(x_n)\}$ does *not* converge to $f(a)$. That f is not continuous at a, means that there is an $\epsilon > 0$ such that no matter how small we choose $\delta > 0$, there is an x such that $d_X(x, a) < \delta$, but $d_Y(f(x), f(a)) \geq \epsilon$. In particular, we can for each $n \in \mathbb{N}$ find an x_n such that $d_X(x_n, a) < \frac{1}{n}$, but $d_Y(f(x_n), f(a)) \geq \epsilon$. Then $\{x_n\}$ converges to a, but $\{f(x_n)\}$ does not converge to $f(a)$. □

The composition of two continuous functions is continuous.

Proposition 2.2.6 *Let (X, d_X), (Y, d_Y), (Z, d_Z) be three metric spaces. Assume that $f : X \to Y$ and $g : Y \to Z$ are two functions, and let $h : X \to Z$ be the composition $h(x) = g(f(x))$. If f is continuous at the point $a \in X$ and g is continuous at the point $b = f(a)$, then h is continuous at a.*

Proof: Assume that $\{x_n\}$ converges to a. Since f is continuous at a, the sequence $\{f(x_n)\}$ converges to $f(a)$, and since g is continuous at $b = f(a)$, the sequence $\{g(f(x_n))\}$ converges to $g(f(a))$, i.e $\{h(x_n)\}$ converges to $h(a)$. By the proposition above, h is continuous at a. □

As in calculus, a function is called continuous if it is continuous at all points:

Definition 2.2.7 *A function $f : X \to Y$ between two metrics spaces is called* continuous *if it continuous at all points x in X.*

Occasionally, we need to study functions which are only defined on a subset A of our metric space X. We define continuity of such functions by restricting the conditions to elements in A:

Definition 2.2.8 *Assume that (X, d_X), (Y, d_Y) are two metric spaces and that A is a subset of X. A function $f : A \to Y$ is* continuous *at a point*

$a \in A$ if for every $\epsilon > 0$ there is a $\delta > 0$ such that $d_Y(f(x), f(a)) < \epsilon$ whenever $x \in A$ and $d_X(x, a) < \delta$. We say that f is continuous *if it is continuous at all $a \in A$.*

There is another way of formulating this definition that is often useful: We can think of f as a function from the metric space (A, d_A) (recall Example 7 in Section 2.1) to (Y, d_Y) and use the original definition of continuity in 2.2.4. By just writing it out, it is easy to see that this definition says exactly the same as the one above. The advantage of the second definition is that it makes it easier to transfer results from the full to the restricted setting, e.g., it is now easy to see that Proposition 2.2.5 can be generalized to:

Proposition 2.2.9 *Assume that (X, d_X) and (Y, d_Y) are metric spaces and that $A \subseteq X$. Then the following are equivalent for a function $f : A \to Y$:*

(i) f is continuous at a point $a \in A$.

(ii) For all sequences $\{x_n\}$ in A converging to a, the sequence $\{f(x_n)\}$ converges to $f(a)$.

Exercises to Section 2.2

1. Assume that (X, d) is a discrete metric space (recall Example 6 in Section 2.1). Show that the sequence $\{x_n\}$ converges to a if and only if there is an $N \in \mathbb{N}$ such that $x_n = a$ for all $n \geq N$.
2. Prove Proposition 2.2.6 without using Proposition 2.2.5, i.e. use only the definition of continuity.
3. Prove Proposition 2.2.9.
4. Assume that (X, d) is a metric space, and let $\mathbb{R}$ have the usual metric $d_{\mathbb{R}}(x, y) = |x - y|$. Assume that $f, g : X \to \mathbb{R}$ are continuous functions.
    a) Show that cf is continuous for all constants $c \in \mathbb{R}$.
    b) Show that $f + g$ is continuous.
    c) Show that fg is continuous.
5. Let (X, d) be a metric space and choose a point $a \in X$. Show that the function $f : X \to \mathbb{R}$ given by $f(x) = d(x, a)$ is continuous (we are using the usual metric $d_{\mathbb{R}}(x, y) = |x - y|$ on $\mathbb{R}$).
6. Let (X, d_X) and (Y, d_Y) be two metric spaces. A function $f : X \to Y$ is said to be a *Lipschitz function* if there is a constant $K \in \mathbb{R}$ such that $d_Y(f(u), f(v)) \leq K d_X(u, v)$ for all $u, v \in X$. Show that all Lipschitz functions are continuous.
7. Let $d_{\mathbb{R}}$ be the usual metric on $\mathbb{R}$ and let d_{disc} be the discrete metric on $\mathbb{R}$. Let $id : \mathbb{R} \to \mathbb{R}$ be the identity function $id(x) = x$. Show that

$$id : (\mathbb{R}, d_{\text{disc}}) \to (\mathbb{R}, d_{\mathbb{R}})$$

is continuous, but that

$$id : (\mathbb{R}, d_{\mathbb{R}}) \to (\mathbb{R}, d_{\text{disc}})$$

is not continuous. Note that this shows that the inverse of a bijective, continuous function is not necessarily continuous.

8. Assume that d_1 and d_2 are two metrics on the same space X. We say that d_1 and d_2 are *equivalent* if there are constants K and M such that $d_1(x, y) \leq Kd_2(x, y)$ and $d_2(x, y) \leq Md_1(x, y)$ for all $x, y \in X$.

   a) Assume that d_1 and d_2 are equivalent metrics on X. Show that if $\{x_n\}$ converges to a in one of the metrics, it also converges to a in the other metric.

   b) Assume that d_1 and d_2 are equivalent metrics on X, and that (Y, d) is a metric space. Show that if $f : X \to Y$ is continuous when we use the d_1-metric on X, it is also continuous when we use the d_2-metric.

   c) We are in the same setting as i part b), but this time we have a function $g : Y \to X$. Show that if g is continuous when we use the d_1-metric on X, it is also continuous when we use the d_2-metric.

   d Assume that d_1, d_2 and d_3 are three metrics on X. Show that if d_1 and d_2 are equivalent, and d_2 and d_3 are equivalent, then d_1 and d_3 are equivalent.

   e) Show that

$$d_1(\mathbf{x}, \mathbf{y}) = |x_1 - y_1| + |x_2 - y_2| + \ldots + |x_n - y_n|$$

$$d_2(\mathbf{x}, \mathbf{y}) = \max\{|x_1 - y_1|, |x_2 - y_2|, \ldots, |x_n - y_n[\}$$

$$d_3(\mathbf{x}, \mathbf{y}) = \sqrt{|x_1 - y_1|^2 + |x_2 - y_2|^2 + \ldots + |x_n - y_n|^2}$$

   are equivalent metrics on $\mathbb{R}^n$.

2.3 Open and closed sets

In this and the following sections, we shall study some of the most important classes of subsets of metric spaces. We begin by recalling and extending the definition of balls in a metric space:

Definition 2.3.1 *Let a be a point in a metric space (X, d), and assume that r is a positive, real number. The (open)* ball centered at a with radius r *is the set*

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

The closed ball centered ar a with radius r *is the set*

$$\overline{\mathrm{B}}(a; r) = \{x \in X : d(x, a) \leq r\}$$

In many ways, balls in metric spaces behave just the way we are used to, but geometrically they may look quite different from ordinary balls. A ball in the Manhattan metric (Example 3 in Section 2.1) looks like an ace of diamonds, while a ball in the discrete metric (Example 6 i Section 2.1) consists either of only one point or the entire space X.

If A is a subset of X and x is a point in X, there are three possibilities:

(i) There is a ball $\mathrm{B}(x; r)$ around x which is contained in A. In this case x is called an *interior* point of A.

(ii) There is a ball $\mathrm{B}(x; r)$ around x which is contained in the complement A^c. In this case x is called an *exterior point* of A.

(iii) All balls $\mathrm{B}(x; r)$ around x contain points in A as well as points in the complement A^c. In this case x is a *boundary point* of A.

Note that an interior point *always* belongs to A, while an exterior point *never* belongs to A. A boundary point will some times belong to A, and some times to A^c.

We now define the important concepts of open and closed sets:

Definition 2.3.2 *A subset A of a metric space is* open *if it does not contain any of its boundary points, and it is* closed *if it contains all its boundary points.*

Most sets contain some, but not all of their boundary points, and are hence neither open nor closed. The empty set $\emptyset$ and the entire space X are both open and closed as they do not have any boundary points. Here is an obvious, but useful reformulation of the definition of an open set.

Proposition 2.3.3 *A subset A of a metric space X is open if and only if it only consists of interior points, i.e. for all $a \in A$, there is a ball $\mathrm{B}(a; r)$ around a which is contained in A.*

Observe that a set A and its complement A^c have exactly the same boundary points. This leads to the following useful result.

Proposition 2.3.4 *A subset A of a metric space X is open if and only if its complement A^c is closed.*

Proof: If A is open, it does not contain any of the (common) boundary points. Hence they all belong to A^c, and A^c must be closed.

Conversely, if A^c is closed, it contains all boundary points, and hence A can not have any. This means that A is open. $\square$

The following observation may seem obvious, but needs to be proved:

Lemma 2.3.5 *All open balls* $\mathrm{B}(a;r)$ *are open sets, while all closed balls* $\overline{\mathrm{B}}(a;r)$ *are closed sets.*

Proof: We prove the statement about open balls and leave the other as an exercise. Assume that $x \in \mathrm{B}(a;r)$; we must show that there is a ball $\mathrm{B}(x;\epsilon)$ around x which is contained in $\mathrm{B}(a;r)$. If we choose $\epsilon = r - d(x,a)$, we see that if $y \in \mathrm{B}(x;\epsilon)$ then by the triangle inequality

$$d(y,a) \leq d(y,x) + d(x,a) < \epsilon + d(x,a) = (r - d(x,a)) + d(x,a) = r$$

Thus $d(y,a) < r$, and hence $\mathrm{B}(x;\epsilon) \subseteq \mathrm{B}(a;r)$ □

The next result shows that closed sets are indeed closed as far as sequences are concerned:

Proposition 2.3.6 *Assume that* F *is a subset of a metric space* X*. The following are equivalent:*

(i) F *is closed.*

(ii) If $\{x_n\}$ *is a convergent sequence of elements in* F*, then the limit* $a = \lim_{n\to\infty} x_n$ *always belongs to* F*.*

Proof: Assume that F is closed and that a does not belong to F. We must show that a sequence from F cannot converge to a. Since F is closed and contains all its boundary points, a has to be an exterior point, and hence there is a ball $\mathrm{B}(a;\epsilon)$ around a which only contains points from the complement of F. But then a sequence from F can never get inside $\mathrm{B}(a,\epsilon)$, and hence cannot converge to a.

Assume now that that F is *not* closed. We shall construct a sequence from F that converges to a point outside F. Since F is not closed, there is a boundary point a that does not belong to F. For each $n \in \mathbb{N}$, we can find a point x_n from F in $\mathrm{B}(a;\frac{1}{n})$. Then $\{x_n\}$ is a sequence from F that converges to a point a which is not in F. □

An open set containing x is called a *neighborhood* of x[1]. The next result is rather silly, but also quite useful.

Lemma 2.3.7 *Let* U *be a subset of the metric space* X*, and assume that each* $x_0 \in U$ *has a neighborhood* $U_{x_0} \subseteq U$*. Then* U *is open.*

Proof: We must show that any $x_0 \in U$ is an interior point. Since U_{x_0} is open, there is an $r > 0$ such that $\mathrm{B}(x_0,r) \subseteq U_{x_0}$. But then $\mathrm{B}(x_0,r) \subseteq U$,

[1] In some books, a *neighborhood* of x is not necessarily open, but does contain a ball centered at x. What we have defined, is the then referred to as an *open neighborhood*

which shows that x_0 is an interior point of U. □

In Proposition 2.2.5 we gave a characterization of continuity in terms of sequences. We shall now prove three characterizations in terms of open and closed sets. The first one characterizes continuity at a point.

Proposition 2.3.8 *Let $f : X \to Y$ be a function between metric spaces, and let x_0 be a point in X. Then the following are equivalent:*

(i) f is continuous at x_0.

(ii) For all neighborhoods V of $f(x_0)$, there is a neighborhood U of x_0 such that $f(U) \subseteq V$.

Proof: $(i) \Longrightarrow (ii)$: Assume that f is continuous at x_0. If V is a neighborhood of $f(x_0)$, there is a ball $B_Y(f(x_0), \epsilon)$ centered at $f(x_0)$ and contained in V. Since f is continuous at x_0, there is a $\delta > 0$ such that $d_Y(f(x), f(x_0)) < \epsilon$ whenever $d_X(x, x_0) < \delta$. But this means that $f(B_X(x_0, \delta)) \subseteq B_Y(f(x_0), \epsilon) \subseteq V$. Hence (ii) is satisfied if we choose $U = \mathrm{B}(x_0, \delta)$.

$(ii) \Longrightarrow (i)$ We must show that for any given $\epsilon > 0$, there is a $\delta > 0$ such that $d_Y(f(x), f(x_0)) < \epsilon$ whenever $d_X(x, x_0) < \delta$. Since $V = B_Y(f(x_0), \epsilon)$ is a neighbohood of $f(x_0)$, there must be a neighborhood U of x_0 such that $f(U) \subseteq V$. Since U is open, there is a ball $B(x_0, \delta)$ centered at x_0 and contained in U. Assume that $d_X(x, x_0) < \delta$. Then $x \in B_X(x_0, \delta) \subseteq U$, and hence $f(x) \in V = B_Y(f(x_0), \epsilon)$, which means that $d_Y(f(x), f(x_0)) < \epsilon$. Hence we have found a $\delta > 0$ such that $d_Y(f(x), f(x_0)) < \epsilon$ whenever $d_X(x, x_0) < \delta$, and thus f is continuous at x_0. □

We can also use open sets to characterize global continuity of functions:

Proposition 2.3.9 *The following are equivalent for a function $f : X \to Y$ between two metric spaces:*

(i) f is continuous.

(ii) Whenever V is an open subset of Y, the inverse image $f^{-1}(V)$ is an open set in X.

Proof: $(i) \Longrightarrow (ii)$: Assume that f is continuous and that $V \subseteq Y$ is open. We shall prove that $f^{-1}(V)$ is open. For any $x_0 \in f^{-1}(V)$, $f(x_0) \in V$, and we know from the previous theorem that there is a neighborhood U_{x_0} of x_0 such that $f(U_{x_0}) \subseteq V$. But then $U_{x_0} \subseteq f^{-1}(V)$, and by Lemma 2.3.7, $f^{-1}(V)$ is open.

$(ii) \Longrightarrow (i)$ Assume that the inverse images of open sets are open. To prove that f is continuous at an arbitrary point x_0, Proposition 2.3.6 tells us that it suffices to show that for any neighborhood V of $f(x_0)$, there is a neighborhood U of x_0 such that $f(U) \subseteq V$. But this easy: Since the inverse image of an open set is open, we can simply choose $U = f^{-1}(V)$. □

The description above is useful in many situations. Using that inverse images commute with complements, and that closed sets are the complements of open, we can translate it into a statement about closed sets:

Proposition 2.3.10 *The following are equivalent for a function $f : X \to Y$ between two metric spaces:*

(i) f is continuous.

(ii) Whenever F is a closed subset of Y, the inverse image $f^{-1}(F)$ is a closed set in X.

Proof: $(i) \implies (ii)$: Assume that f is continuous and that $F \subseteq Y$ is closed. Then F^c is open, and by the previous proposition, $f^{-1}(F^c)$ is open. Since inverse images commute with complements, $(f^{-1}(F))^c = f^{-1}(F^c)$. This means that $f^{-1}(F)$ has an open complement and hence is closed.

$(ii) \implies (i)$ Assume that the inverse images of closed sets are closed. According to the previous proposition, it suffices to show that the inverse image of any open set $V \subseteq Y$ is open. But if V is open, the complement V^c is closed, and hence by assumption $f^{-1}(V^c)$ is closed. Since inverse images commute with complements, $(f^{-1}(V))^c = f^{-1}(V^c)$. This means that the complement of $f^{-1}(V)$ is closed, and hence $f^{-1}(V)$ is open. □

Mathematicians usually sum up the last two theorems by saying that openness and closedness are preserved under inverse, continuous images. Be aware that these properties are *not* preserved under continuous, *direct* images; even if f is continuous, the image $f(U)$ of an open set U need not be open, and the image $f(F)$ of a closed F need not be closed:

Example 1: Let $f, g : \mathbb{R} \to \mathbb{R}$ be the continuous functions defined by

$$f(x) = x^2 \qquad \text{and} \qquad g(x) = \arctan x$$

The set $\mathbb{R}$ is both open and closed, but $f(\mathbb{R})$ equals $[0, \infty)$ which is not open, and $g(\mathbb{R})$ equals $(-\frac{\pi}{2}, \frac{\pi}{2})$ which is not closed. Hence the continuous image of an open set need not be open, and the continuous image of a closed set need not be closed. ♣

We end this section with two simple but useful observations on open and closed sets.

Proposition 2.3.11 *Let (X, d) be a metric space.*

a) *If $\mathcal{G}$ is a (finite or infinite) collection of open sets, then the union $\bigcup_{G\in\mathcal{G}} G$ is open.*

b) *If $G_1, G_2, \ldots, G_n$ is a* finite *collection of open sets, then the intersection $G_1 \cap G_2 \cap \ldots \cap G_n$ is open.*

Proof: Left to the reader (see Exercise 12, where you are also asked to show that the intersection of infinitely many open sets is not necessarily open). □

Proposition 2.3.12 *Let (X, d) be a metric space.*

a) *If $\mathcal{F}$ is a (finite or infinite) collection of closed sets, then the intersection $\bigcap_{F\in\mathcal{F}} F$ is closed.*

b) *If $F_1, F_2, \ldots, F_n$ is a* finite *collection of closed sets, then the union $F_1 \cup F_2 \cup \ldots \cup F_n$ is closed.*

Proof: Left to the reader (see Exercise 13, where you are also asked to show that the union of infinitely many closed sets is not necessarily closed). □

Propositions 2.3.11 and 2.3.12 are the starting points for *topology*, an even more abstract theory of nearness.

Exercises to Section 2.3

1. Assume that (X, d) is a discrete metric space.
   a) Show that an open ball in X is either a set with only one element (a *singleton*) or all of X.
   b) Show that all subsets of X are both open and closed.
   c) Assume that (Y, d_Y) is another metric space. Show that all functions $f : X \to Y$ are continuous.
2. Give a geometric description of the ball $\mathrm{B}(a; r)$ in the Manhattan metric (see Example 3 in Section 2.1). Make a drawing of a typical ball. Show that the Manhattan metric and the usual metric in $\mathbb{R}^2$ have exactly the same open sets.
3. Assume that F is a non-empty, closed and bounded subset of $\mathbb{R}$ (with the usual metric $d(x, y) = |y - x|$). Show that $\sup F \in F$ and $\inf F \in F$. Give an example of a bounded, but not closed set F such that $\sup F \in F$ and $\inf F \in F$.
4. Prove the second part of Lemma 2.3.5, i.e. prove that a closed ball $\overline{\mathrm{B}}(a; r)$ is always a closed set.
5. Assume that $f : X \to Y$ and $g : Y \to Z$ are continuous functions. Use Proposition 2.3.9 to show that the composition $g \circ f : X \to Z$ is continuous.

6. Assume that A is a subset of a metric space (X, d). Show that the interior points of A are the exterior points of A^c, and that the exterior points of A are the interior points of A^c. Check that the boundary points of A are the boundary points of A^c.

7. Assume that A is a subset of a metric space X. The *interior* A° of A is the set consisting of all interior points of A. Show that A° is open.

8. Assume that A is a subset of a metric space X. The *closure* $\overline{A}$ of A is the set consisting of all interior points plus all boundary points of A.

   a) Show that $\overline{A}$ is closed.

   b) Let $\{a_n\}$ be a sequence from A converging to a point a. Show that $a \in \overline{A}$.

9. Let (X, d) be a metric space, and let A be a subset of X. We shall consider A with the subset metric d_A.

   a) Assume that $G \subseteq A$ is open in (X, d). Show that G is open in (A, d_A).

   b) Find an example which shows that although $G \subseteq A$ is open in (A, d_A) it need not be open in (X, d_X).

   c) Show that if A is an open set in (X, d_X), then a subset G of A is open in (A, d_A) if and only if it is open in (X, d_X)

10. Let (X, d) be a metric space, and let A be a subset of X. We shall consider A with the subset metric d_A.

    a) Assume that $F \subseteq A$ is closed in (X, d). Show that F is closed in (A, d_A).

    b) Find an example which shows that although $F \subseteq A$ is closed in (A, d_A) it need not be closed in (X, d_X).

    c) Show that if A is a closed set in (X, d_X), then a subset F of A is open in (A, d_A) if and only if it is closed in (X, d_X)

11. Let (X, d) be a metric space and give $\mathbb{R}$ the usual metric. Assume that $f : X \to \mathbb{R}$ is continuous.

    a) Show that the set
    $$\{x \in X \mid f(x) < a\}$$
    is open for all $a \in \mathbb{R}$.

    a) Show that the set
    $$\{x \in X \mid f(x) \leq a\}$$
    is closed for all $a \in \mathbb{R}$.

12. Prove Proposition 2.3.11. Find an example of an infinite collection of open sets $G_1, G_2, \ldots$ whose intersection is *not* open.

13. Prove Proposition 2.3.12. Find an example of an infinite collection of closed sets $F_1, F_2, \ldots$ whose union is *not* closed.

2.4 Complete spaces

One of the reasons why calculus in $\mathbb{R}^n$ is so successful, is that $\mathbb{R}^n$ is a complete space. We shall now generalize this notion to metric spaces. The key concept is that of a Cauchy sequence:

Definition 2.4.1 *A sequence $\{x_n\}$ in a metric space (X, d) is a* Cauchy sequence *if for each $\epsilon > 0$ there is an $N \in \mathbb{N}$ such that $d(x_n, x_m) < \epsilon$ whenever $n, m \geq N$.*

We begin by a simple observation:

Proposition 2.4.2 *Every convergent sequence is a Cauchy sequence.*

Proof: If a is the limit of the sequence, there is for any $\epsilon > 0$ a number $N \in \mathbb{N}$ such that $d(x_n, a) < \frac{\epsilon}{2}$ whenever $n \geq N$. If $n, m \geq N$, the triangle inequality tells us that

$$d(x_n, x_m) \leq d(x_n, a) + d(a, x_m) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

and consequently $\{x_n\}$ is a Cauchy sequence. □

The converse of the proposition above does not hold in all metric spaces, and we make the following definition:

Definition 2.4.3 *A metric space is called* complete *if all Cauchy sequences converge.*

We know from MAT1110 that $\mathbb{R}^n$ is complete, but that $\mathbb{Q}$ is not when we use the usual metric $d(x, y) = |x - y|$. The complete spaces are in many ways the "nice" metric spaces, and we shall spend much time studying their properties. We shall also spend some time showing how we can make non-complete spaces complete. Example 5 in Section 2.1 (where X is the space of all continuous $f : [a, b] \to \mathbb{R}$) shows some interesting cases; X with the metric d_1 is complete, but not X with the metrics d_2 and d_3. By introducing a stronger notion of integral (the Lebesgue integral) we can extend d_2 and d_3 to complete metrics by making them act on richer spaces of functions. In Section 2.7, we shall study an abstract method for making incomplete spaces complete by adding new points.

The following proposition is quite useful. Remember that if A is a subset of X, then d_A is the subspace metric obtained by restricting d to A (see Example 7 in Section 2.1).

Proposition 2.4.4 *Assume that (X, d) is a complete metric space. If A is a subset of X, (A, d_A) is complete if and only if A is closed.*

Proof: Assume first that A is closed. If $\{a_n\}$ is a Cauchy sequence in A, $\{a_n\}$ is also a Cauchy sequence in X, and since X is complete, $\{a_n\}$ converges to a point $a \in X$. Since A is closed, Proposition 2.3.6 tells us that $a \in A$. But then $\{a_n\}$ converges to a in (A, d_A), and hence (A, d_A) is complete.

If A is not closed, there is a boundary point a that does not belong to A. Each ball $\mathrm{B}(a, \frac{1}{n})$ must contain an element a_n from A. In X, the sequence $\{a_n\}$ converges to a, and must be a Cauchy sequence. However, since $a \notin A$, the sequence $\{a_n\}$ does *not* converge to a point in A. Hence we have found a Cauchy sequence in (A, d_A) that does not converge to a point in A, and hence (A, d_A) is incomplete. □

The nice thing about complete spaces is that we can prove that sequences converge to a limit without actually constructing or specifying the limit — all we need is to prove that the sequence is a Cauchy sequence. To prove that a sequence has the Cauchy property, we only need to work with the given terms of the sequence and not the unknown limit, and this often makes the arguments much easier. As an example of this technique, we shall now prove an important theorem that will be useful later in the book, but first we need some definitions.

A function $f : X \to X$ is called a *contraction* if there is a positive number $s < 1$ such that

$$d(f(x), f(y)) \leq s\, d(x, y) \quad \text{for all } x, y \in X$$

We call s a *contraction factor* for f. All contractions are continuous (prove this!), and by induction it is easy to see that

$$d(f^{\circ n}(x), f^{\circ n}(y)) \leq s^n d(x, y)$$

where $f^{\circ n}(x) = f(f(f(\ldots f(x))))$ is the result of iterating f exactly n times. If $f(a) = a$, we say that a is a *fixed point* for f.

Theorem 2.4.5 (Banach's Fixed Point Theorem) *Assume that (X, d) is a complete metric space and that $f : X \to X$ is a contraction. Then f has a unique fixed point a, and no matter which starting point $x_0 \in X$ we choose, the sequence*

$$x_0,\ x_1 = f(x_0),\ x_2 = f^{\circ 2}(x_0), \ldots,\ x_n = f^{\circ n}(x_0), \ldots$$

converges to a.

Proof: Let us first show that f can not have more than one fixed point. If a and b are two fixed points, and s is a contraction factor for f, we have

$$d(a, b) = d(f(a), f(b)) \leq s\, d(a, b)$$

Since $0 < s < 1$, this is only possible if $d(a, b) = 0$, i.e. if $a = b$.

To show that f has a fixed point, choose a starting point x_0 in X and consider the sequence

$$x_0,\ x_1 = f(x_0),\ x_2 = f^{\circ 2}(x_0), \ldots,\ x_n = f^{\circ n}(x_0), \ldots$$

Assume, for the moment, that we can prove that this is a Cauchy sequence. Since (X, d) is complete, the sequence must converge to a point a. To prove that a is a fixed point, observe that we have $x_{n+1} = f(x_n)$ for all n, and taking the limit as $n \to \infty$, we get $a = f(a)$. Hence a is a fixed point of f, and the theorem must hold. Thus it suffices to prove our assumption that $\{x_n\}$ is a Cauchy sequence.

Choose two elements x_n and x_{n+k} of the sequence. By repeated use of the triangle inequality, we get

$$d(x_n, x_{n+k}) \leq d(x_n, x_{n+1}) + d(x_{n+1}, x_{n+2}) + \ldots + d(x_{n+k-1}, x_{n+k}) =$$

$$= d(f^{\circ n}(x_0), f^{\circ n}(x_1)) + d(f^{\circ(n+1)}(x_0), f^{\circ(n+1)}(x_1)) + \ldots$$

$$\ldots + d(f^{\circ(n+k-1)}(x_0), f^{\circ(n+k-1)}(x_1)) \leq$$

$$\leq s^n d(x_0, x_1) + s^{n+1} d(x_0, x_1) + \ldots + s^{n+k-1} d(x_0, x_1) =$$

$$= \frac{s^n(1 - s^k)}{1 - s}\, d(x_0, x_1) \leq \frac{s^n}{1 - s}\, d(x_0, x_1)$$

where we have summed a geometric series to get to the last line. Since $s < 1$, we can get the last expression as small as we want by choosing n large enough. Given an $\epsilon > 0$, we can in particular find an N such that $\frac{s^N}{1-s}\, d(x_0, x_1) < \epsilon$. For $n, m = n + k$ larger than or equal to N, we thus have

$$d(x_n, x_m) \leq \frac{s^n}{1 - s}\, d(x_0, x_1) < \epsilon$$

and hence $\{x_n\}$ is a Cauchy sequence. □

In Section 3.4 we shall use Banach's Fixed Point Theorem to prove the existence of solutions to quite general differential equations.

Exercises to Section 2.4

1. Show that the discrete metric is always complete.
2. Assume that (X, d_X) and (Y, d_Y) are complete spaces, and give $X \times Y$ the metric d defined by

$$d((x_1, y_1), (x_2, y_2)) = d_X(x_1, x_2) + d_Y(y_1, y_2)$$

   Show that $(X \times Y, d)$ is complete.

3. If A is a subset of a metric space (X, d), the *diameter* $\text{diam}(A)$ of A is defined by
$$\text{diam}(A) = \sup\{d(x, y) \mid x, y \in A\}$$
Let $\{A_n\}$ be a collection of subsets of X such that $A_{n+1} \subseteq A_n$ and $\text{diam}(A_n) \to 0$, and assume that $\{a_n\}$ is a sequence such that $a_n \in A_n$ for each $n \in \mathbb{N}$. Show that if X is complete, the sequence $\{a_n\}$ converges.

4. Assume that d_1 and d_2 are two metrics on the same space X. We say that d_1 and d_2 are *equivalent* if there are constants K and M such that $d_1(x, y) \leq K d_2(x, y)$ and $d_2(x, y) \leq M d_1(x, y)$ for all $x, y \in X$. Show that if d_1 and d_2 are equivalent, and one of the spaces (X, d_1), (X, d_2) is complete, then so is the other.

5. Assume that $f : [0, 1] \to [0, 1]$ is a differentiable function and that there is a number $s < 1$ such that $|f'(x)| < s$ for all $x \in (0, 1)$. Show that there is exactly one point $a \in [0, 1]$ such that $f(a) = a$.

6. You are standing with a map in your hand inside the area depicted on the map. Explain that there is exactly one point on the map that is vertically above the point it depicts.

7. Assume that (X, d) is a complete metric space, and that $f : X \to X$ is a function such that $f^{\circ n}$ is a contraction for some $n \in \mathbb{N}$. Show that f has a unique fixed point.

8. A subset D of a metric space X is *dense* if for all $x \in X$ and all $\epsilon \in \mathbb{R}_+$ there is an element $y \in D$ such that $d(x, y) < \epsilon$. Show that if all Cauchy sequence $\{y_n\}$ from a dense set D converge in X, then X is complete.

2.5 Compact sets

We now turn to the study of compact sets. These sets are related both to closed sets and to the notion of completeness, and they are extremely useful in many applications.

Assume that $\{x_n\}$ is a sequence in a metric space X. If we have a strictly increasing sequence of natural numbers

$$n_1 < n_2 < n_3 < \ldots < n_k < \ldots$$

we call the sequence $\{y_k\} = \{x_{n_k}\}$ a *subsequence* of $\{x_n\}$. A subsequence contains infinitely many of the terms in the original sequence, but usually not all.

I leave the first result as an exercise:

Proposition 2.5.1 *If the sequence $\{x_n\}$ converges to a, so does all subsequences.*

We are now ready to define compact sets:

Definition 2.5.2 *A subset K of a metric space (X, d) is called* compact *if every sequence in K has a subsequence converging to a point in K. The space (X, d) is* compact *if X a compact set, i.e. if all sequences in X has a convergent subsequence.*

Compactness is a rather complex notion that it takes a while to get used to. We shall start by relating it to other concepts we have already introduced. First a definition:

Definition 2.5.3 *A subset A of a metric space (X, d) is* bounded *if there is a point $b \in X$ and a constant $K \in \mathbb{R}$ such that $d(a, b) \leq K$ for all $a \in A$ (it does not matter which point $b \in X$ we use in this definition).*

Here is our first result on compact sets:

Proposition 2.5.4 *Every compact set K in a metric space (X, d) is closed and bounded.*

Proof: We argue contrapositively. First we show that if a set K is not closed, then it can not be compact, and then we show that if K is not bounded, it can not be compact.

Assume that K is not closed. Then there is a boundary point a that does not belong to K. For each $n \in \mathbb{N}$, there is an $x_n \in K$ such that $d(x_n, a) < \frac{1}{n}$. The sequence $\{x_n\}$ converges to $a \notin K$, and so does all its subsequences, and hence no subsequence can converge to a point in K.

Assume now that K is not bounded. For every $n \in \mathbb{N}$ there is an element $x_n \in K$ such that $d(x_n, b) > n$. If $\{y_k\}$ is a subsequence of x_n, clearly $\lim_{k \to \infty} d(y_k, b) = \infty$. It is easy to see that $\{y_k\}$ can not converge to any element $y \in X$: According to the triangle inequality

$$d(y_k, b) \leq d(y_k, y) + d(y, b)$$

and since $d(y_k, b) \to \infty$, we must have $d(y_k, y) \to \infty$. Hence $\{x_n\}$ has no convergent subsequences, and K can not be compact. □

In $\mathbb{R}^n$ the converse of the result above holds: All closed and bounded subsets of $\mathbb{R}^n$ are compact (this is just a reformulation of Bolzano-Weierstrass' Theorem in MAT1110). The following example shows that this is not the case for all metric space.

Example 1: Consider the metric space $(\mathbb{N}, d)$ where d is the discrete metric. Then $\mathbb{N}$ is complete, closed and bounded, but the sequence $\{n\}$ does not have a convergent subsequence.

We shall later see how we can strengthen the boundedness condition (to something called *total boundedness*) to get a characterization of compactness.

We next want to take a look at the relationship between completeness and compactness. Not all complete spaces are compact ($\mathbb{R}$ is complete but not compact), but it turns out that all compact spaces are complete. To prove this, we need a lemma on subsequences of Cauchy sequences that is useful also in other contexts.

Lemma 2.5.5 *Assume that $\{x_n\}$ is a Cauchy sequence in a (not necessarily complete) metric space (X, d). If there is a subsequence $\{x_{n_k}\}$ converging to a point a, then the original sequence $\{x_n\}$ also converges to a*

Proof: We must show that for any given $\epsilon > 0$, there is an $N \in \mathbb{N}$ such that $d(x_n, a) < \epsilon$ for all $n \geq N$. Since $\{x_n\}$ is a Cauchy sequence, there is an $N \in \mathbb{N}$ such that $d(x_n, x_m) < \frac{\epsilon}{2}$ for all $n, m \geq N$. Since $\{x_{n_k}\}$ converges to a, there is a K such that $n_K \geq N$ and $d(x_{n_K}, a) \leq \frac{\epsilon}{2}$. For all $n \geq N$ we then have

$$d(x_n, a) \leq d(x_n, x_{n_K}) + d(x_{n_K}, a) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

by the triangle inequality. □

Proposition 2.5.6 *Every compact metric space is complete.*

Proof: Let $\{x_n\}$ be a Cauchy sequence. Since X is compact, there is a subsequence $\{x_{n_k}\}$ converging to a point a. By the lemma above, $\{x_n\}$ also converges to a. Hence all Cauchy sequences converge, and X must be complete. □

Here is another useful result:

Proposition 2.5.7 *A closed subset F of a compact set K is compact.*

Proof: Assume that $\{x_n\}$ is a sequence in F — we must show that $\{x_n\}$ has a subsequence converging to a point in F. Since $\{x_n\}$ is also a sequence in K, and K is compact, there is a subsequence $\{x_{n_k}\}$ converging to a point $a \in K$. Since F is closed, $a \in F$, and hence $\{x_n\}$ has a subsequence converging to a point in F. □

We have previously seen that if f is a continuous function, the inverse images of open and closed sets are open and closed, respectively. The inverse image of a compact set need not be compact, but it turns out that the (direct) image of a compact set under a continuous function is always compact.

Proposition 2.5.8 *Assume that $f : X \to Y$ is a continuous function between two metric spaces. If $K \subseteq X$ is compact, then $f(K)$ is a compact subset of Y.*

Proof: Let $\{y_n\}$ be a sequence in $f(K)$; we shall show that $\{y_n\}$ has subsequence converging to a point in $f(K)$. Since $y_n \in f(K)$, we can for each n find an element $x_n \in K$ such that $f(x_n) = y_n$. Since K is compact, the sequence $\{x_n\}$ has a subsequence $\{x_{n_k}\}$ converging to a point $x \in K$. But then by Proposition 2.2.5, $\{y_{n_k}\} = \{f(x_{n_k})\}$ is a subsequence of $\{y_n\}$ converging to $y = f(x) \in f(K)$. □

So far we have only proved technical results about the nature of compact sets. The next result gives the first indication why these sets are useful.

Theorem 2.5.9 (The Extreme Value Theorem) *Assume that K is a non-empty, compact subset of a metric space (X, d) and that $f : K \to \mathbb{R}$ is a continuous function. Then f has maximum and minimum points in K, i.e. there are points $c, d \in K$ such that*

$$f(d) \leq f(x) \leq f(c)$$

for all $x \in K$.

Proof: There is a quick way of proving this theorem by using the previous proposition (see the remark below), but I choose a slightly longer proof as I think it gives a better feeling for what is going on and how compactness argumentness are used in practice. I only prove the maximum part and leave the minimum as an exercise.

Let

$$M = \sup\{f(x) \mid x \in K\}$$

(if F is unbounded, we put $M = \infty$) and choose a sequence $\{x_n\}$ in K such that $\lim_{n\to\infty} f(x_n) = M$. Since K is compact, $\{x_n\}$ has a subsequence $\{x_{n_k}\}$ converging to a point $c \in K$. Then on the one hand $\lim_{k\to\infty} f(x_{n_k}) = M$, and on the other $\lim_{k\to\infty} f(x_{n_k}) = f(c)$ according to Proposition 2.2.9. Hence $f(c) = M$, and since $M = \sup\{f(x) \mid x \in K\}$, we see that c is a maximum point for f on K. □.

Remark: As already mentioned, it is possible to give a shorter proof of the Extreme Value Theorem by using Proposition 2.5.7. According to it, the set $f(K)$ is compact and thus closed and bounded. This means that $\sup f(K)$ and $\inf f(K)$ belong to $f(K)$, and hence there are points $c, d \in K$ such that $f(c) = \sup f(K)$ and $f(d) = \inf f(K)$. Clearly, c is a maximum and d a minimum point for f.

Let us finally turn to the description of compactness in terms of total boundedness.

Definition 2.5.10 *A subset A of a metric space X is called* totally bounded *if for each $\epsilon > 0$ there is a finite number* $\mathrm{B}(a_1, \epsilon), \mathrm{B}(a_2, \epsilon), \ldots, \mathrm{B}(a_n, \epsilon)$ *of balls with centers in A and radius ϵ that cover A (i.e. $A \subseteq \mathrm{B}(a_1, \epsilon) \cup \mathrm{B}(a_2, \epsilon) \cup \ldots \cup \mathrm{B}(a_n, \epsilon)$).*

We first observe that a compact set is always totally bounded.

Proposition 2.5.11 *Let K be a compact subset of a metric space X. Then K is totally bounded.*

Proof: We argue contrapositively: Assume that A is *not* totally bounded. Then there is an $\epsilon > 0$ such that no finite collection of ϵ-balls cover A. We shall construct a sequence $\{x_n\}$ in A that does not have a convergent subsequence. We begin by choosing an arbitrary element $x_1 \in A$. Since $\mathrm{B}(x_1, \epsilon)$ does not cover A, we can choose $x_2 \in A \setminus \mathrm{B}(x_1, \epsilon)$. Since $\mathrm{B}(x_1, \epsilon)$ and $\mathrm{B}(x_2, \epsilon)$ do not cover A, we can choose $x_3 \in A \setminus \big(\mathrm{B}(x_1, \epsilon) \cup \mathrm{B}(x_2, \epsilon)\big)$. Continuing in this way, we get a sequence $\{x_n\}$ such that

$$x_n \in A \setminus \big(\mathrm{B}(x_1, \epsilon) \cup \mathrm{B}(x_2, \epsilon) \cup \ldots \cup (\mathrm{B}(x_{n-1}, \epsilon)\big)$$

This means that $d(x_n, x_m) \geq \epsilon$ for all $n, m \in \mathbb{N}$, $n > m$, and hence $\{x_n\}$ has no convergent subsequence. □

We are now ready for the final theorem. Note that we have now added the assumption that X is complete — without this condition, the statement is false.

Theorem 2.5.12 *A subset A of a complete metric space X is compact if and only if it is closed and totally bounded.*

Proof: As we already know that a compact set is closed and totally bounded, it suffices to prove that a closed and totally bounded set A is compact. Let $\{x_n\}$ be a sequence in A. Our aim is to construct a convergent subsequence $\{x_{n_k}\}$. Choose balls $B_1^1, B_2^1, \ldots, B_{k_1}^1$ of radius one that cover A. At least one of these balls must contain infinitely many terms from the sequence. Call this ball S_1 (if there are more than one such ball, just choose one). We now choose balls $B_1^2, B_2^2, \ldots, B_{k_2}^2$ of radius $\frac{1}{2}$ that cover A. At least one of these ball must contain infinitely many of the terms from the sequence that lies in S_1. If we call this ball S_2, $S_1 \cap S_2$ contains infinitely many terms from the sequence. Continuing in this way, we find a sequence of balls S_k of radius $\frac{1}{k}$ such that

$$S_1 \cap \ S_2 \cap \ldots \cap S_k$$

always contains infinitely many terms from the sequence.

We can now construct a convergent subsequence of $\{x_n\}$. Choose n_1 to be the first number such that x_{n_1} belongs to S_1. Choose n_2 to be first number larger that n_1 such that x_{n_2} belongs to $S_1 \cap S_2$, then choose n_3 to be the first number larger than n_2 such that x_{n_3} belongs to $S_1 \cap S_2 \cap S_3$. Continuing in this way, we get a subsequence $\{x_{n_k}\}$ such that

$$x_{n_k} \in S_1 \cap S_2 \cap \ldots \cap S_k$$

for all k. Since the S_k's are shrinking, $\{x_{n_k}\}$ is a Cauchy sequence, and since X is complete, $\{x_{n_k}\}$ converges to a point a. Since A is closed, $a \in A$. Hence we have proved that any sequence in A has a subsequence converging to a point in A, and thus A is compact. $\square$

Problems to Section 2.5

1. Show that a space (X, d) with the discrete metric is compact if and only if X is a finite set.
2. Prove Proposition 2.5.1.
3. Prove the minimum part of Theorem 2.5.9.
4. Let b and c be two points in a metric space (X, d), and let A be a subset of X. Show that if there is a number $K \in \mathbb{R}$ such that $d(a, b) \leq K$ for all $a \in A$, then there is a number $M \in \mathbb{R}$ such that $d(a, c) \leq M$ for all $a \in A$. Hence it doesn't matter which point $b \in X$ we use in Definition 2.5.3.
5. Assume that (X, d) is metric space and that $f : X \to [0, \infty)$ is a continuous function. Assume that for each $\epsilon > 0$, there is a compact $K_\epsilon \subseteq X$ such that $f(x) < \epsilon$ when $x \notin K_\epsilon$. Show that f has a maximum point.
6. Let (X, d) be a compact metric space, and assume that $f : X \to \mathbb{R}$ is continuous when we give $\mathbb{R}$ the usual metric. Show that if $f(x) > 0$ for all $x \in X$, then there is a positive, real number a such that $f(x) > a$ for all $x \in X$.
7. Assume that $f : X \to Y$ is a continuous function between metric spaces, and let K be a compact subset of Y. Show that $f^{-1}(K)$ is closed. Find an example which shows that $f^{-1}(K)$ need not be compact.
8. Show that a totally bounded subset of a metric space is always bounded. Find an example of a bounded set in a metric space that is not totally bounded.
9. The Bolzano-Weierstrass' Theorem says that any bounded sequence in $\mathbb{R}^n$ has a convergent subsequence. Use it to prove that a subset of $\mathbb{R}^n$ is compact if and only if it is closed and bounded.
10. Let (X, d) be a metric space.
    a) Assume that $K_1, K_2, \ldots, K_n$ is a finite collection of compact subsets of X. Show that the union $K_1 \cup K_2 \cup \ldots \cup K_n$ is compact.
    b) Assume that $\mathcal{K}$ is a collection of compact subset of X. Show that the intersection $\bigcap_{K \in \mathcal{K}} K$ is compact.
11. Let (X, d) be a metric space. Assume that $\{K_n\}$ is a sequence of non-empty, compact subsets of X such that $K_1 \supseteq K_2 \supseteq \ldots \supseteq K_n \supseteq \ldots$. Prove that $\bigcap_{n \in \mathbb{N}} K_n$ is non-empty.

12. Let (X, d_X) and (Y, d_Y) be two metric spaces. Assume that (X, d_X) is compact, and that $f : X \to Y$ is bijective and continuous. Show that the inverse function $f^{-1} : Y \to X$ is continuous.

13. Assume that C and K are disjoint, compact subsets of a metric space (X, d), and define
$$a = \inf\{d(x, y) \mid x \in C, y \in K\}$$
Show that a is strictly positive and that there are points $x_0 \in C$, $y_0 \in K$ such that $d(x_0, y_0) = a$. Show by an example that the result does not hold if we only assume that one of the sets C and K is compact and the other one closed.

14. Assume that (X, d) is compact and that $f : X \to X$ is continuous.
    a) Show that the function $g(x) = d(x, f(x))$ is continuous and has a minimum point.
    b) Assume in addition that $d(f(x), f(y)) < d(x, y)$ for all $x, y \in X$, $x \neq y$. Show that f has a unique fixed point. (*Hint:* Use the minimum from a))

2.6 An alternative description of compactness

The descriptions of compactness we studied in the previous section, suffice for most purposes in this book, but for some of the more advanced proofs there is another description that is more convenient. This alternative description is also the right one to use if one wants to extend the concept of compactness to even more general spaces, so-called *topological spaces.* In such spaces, sequences are not always an efficient tool, and it is better to have a description of compactness in terms of coverings by open sets.

To see what this means, assume that K is a subset of a metric space X. An *open covering* of X is simply a (finite or infinite) collection $\mathcal{O}$ of open sets whose union contains K, i.e.

$$K \subseteq \bigcup\{O : O \in \mathcal{O}\}$$

The purpose of this section is to show that in metric spaces, the following property is equivalent to compactness.

Definition 2.6.1 (Open Covering Property) *Let K be a subset of a metric space X. Assume that for each open covering $\mathcal{O}$ of K, there is a* finite *number of elements $O_1, O_2, \ldots, O_n$ in $\mathcal{O}$ such that*

$$K \subseteq O_1 \cup O_2 \cup \ldots \cup O_n$$

(we say that each open covering of K has a finite subcovering*). Then the set K is said to have the* open covering property*.*

The open covering property is quite abstract and may take some time to get used to, but it turns out to be a very efficient tool. Note that the term "open covering property" is not standard terminology, and that it will more or less disappear once we have proved that it is equivalent to compactness.

Let us first prove that a set with the open covering property is necessarily compact. Before we begin, we need a simple observation: Assume that x is a point in our metric space X, and that no subsequence of a sequence $\{x_n\}$ converges to x. Then there must be an open ball $\mathrm{B}(x;r)$ around x which only contains finitely many terms from $\{x_n\}$ (because if all balls around x contained infinitely many terms, we could use these terms to construct a subsequence converging to x).

Proposition 2.6.2 *If a subset K of a metric space X has the open covering property, then it is compact.*

Proof: We argue contrapositively, i.e., we assume that K is *not* compact and prove that it does not have the open covering property. Since K is not compact, there is a sequence $\{x_n\}$ which does not have any subsequences converging to points in K. By the observation above, this means that for each element $x \in K$, there is an open ball $\mathrm{B}(x;r_x)$ around x which only contains finitely many terms of the sequence. The family $\{\mathrm{B}(x,r_x) : x \in K\}$ is an open covering of K, but it cannot have a finite subcovering since any such subcovering $\mathrm{B}(x_1,r_{x_1}), \mathrm{B}(x_2,r_{x_2}), \ldots, \mathrm{B}(x_m,r_{x_m})$ can only contain finitely many of the infinitely many terms in the sequence. □

To prove the opposite implication, we shall use an elegant trick based on the Extreme Value Theorem, but first we need a lemma (the strange cut-off at 1 in the definition of $f(x)$ below is just to make sure that the function is finite):

Lemma 2.6.3 *Let $\mathcal{O}$ be an open covering of a subset A of a metric spece X. Define a function $f : A \to \mathbb{R}$ by*

$$f(x) = \sup\{r \in \mathbb{R} \,|\, r < 1 \text{ and } \mathrm{B}(x;r) \subseteq O \text{ for some } O \in \mathcal{O}\}$$

Then f is continuous and strictly positive (i.e. $f(x) > 0$ for all $x \in A$).

Proof: The strict positivity is easy: Since $\mathcal{O}$ is a covering of A, there is a set $O \in \mathcal{O}$ such that $x \in O$, and since O is open, there is an r, $0 < r < 1$, such that $\mathrm{B}(x;r) \subseteq O$. Hence $f(x) \geq r > 0$.

To prove the continuity, it suffices to show that $|f(x) - f(y)| \leq d(x,y)$ as we can then choose $\delta = \epsilon$ in the definition of continuity. Observe first that if $f(x), f(y) \leq d(x,y)$, there is nothing to prove. Assume therefore that at least one of these values is larger than $d(x,y)$. Without out loss of generality, we may assume that $f(x)$ is the larger of the two. There must then be an $r > d(x,y)$ and an $O \in \mathcal{O}$ such that $\mathrm{B}(x,r) \subseteq O$. For any such r, $\mathrm{B}(y, r - d(x,y)) \subseteq O$ since $\mathrm{B}(y, r - d(x,y)) \subset \mathrm{B}(x,r)$. This means that $f(y) \geq f(x) - d(x,y)$. Since by assumption $f(x) \geq f(y)$, we have $|f(x) - f(y)| \leq d(x,y)$ which is what we set out to prove. □

We are now ready for the main theorem:

Theorem 2.6.4 *A subset K of a metric space is compact if and only if it has the open covering property.*

Proof: It remains to prove that if K is compact and $\mathcal{O}$ is an open covering of K, then $\mathcal{O}$ has a finite subcovering. By the Extremal Value Theorem, the function f in the lemma attains a minimal value r on K, and since f is strictly positive, $r > 0$. This means that for all $x \in K$, the ball $\mathcal{B}(x,r)$ is contained in a set $O \in \mathcal{B}$. Since K is compact, it is totally bounded, and hence there is a finite collection of balls $\mathcal{B}(x_1, r_1)$, $\mathcal{B}(x_2, r_2)$, $\ldots$, $\mathcal{B}(x_n, r_n)$ that cover K. Each ball $\mathcal{B}(x_i, r_i)$ is contained in a set $O_i \in \mathcal{O}$, and hence $O_1, O_2, \ldots, O_n$ is a finite subcovering of $\mathcal{O}$. □

As usual, there is a reformulation of the theorem above in terms of closed sets. Let us first agree to say that a collection $\mathcal{F}$ of sets has the *finite intersection property over K* if

$$K \cap F_1 \cap F_2 \cap \ldots \cap F_n \neq \emptyset$$

for all finite collections $F_1, F_2, \ldots, F_n$ of sets from $\mathcal{F}$.

Corollary 2.6.5 *Assume that K is a subset of a metric space X. Then the following are equivalent:*

(i) K is compact.

(ii) If a collection $\mathcal{F}$ of closed sets has the finite intersection property over K, then

$$K \cap \left(\bigcap_{F \in \mathcal{F}} F \right) \neq \emptyset$$

Proof: Left to the reader (see Exercise 8). □

Problems to Section 2.6

1. Assume that $\mathcal{I}$ is a collection of open intervals in $\mathbb{R}$ whose union contains $[0,1]$. Show that there exists a finite collection $I_1, I_2, \ldots, I_n$ of sets from $\mathcal{I}$ such that
$$[0,1] \subseteq I_1 \cup I_1 \cup \ldots \cup I_n$$

2. Let $\{K_n\}$ be a decrasing sequence (i.e., $K_{n+1} \subseteq K_n$ for all $n \in \mathbb{N}$) of nonempty, compact sets. Show that $\bigcap_{n\in\mathbb{N}} K_n \neq \emptyset$.

3. Assume that $f : X \to Y$ is a continuous function between two metric spaces. Use the open covering property to show that if K is a compact subset of X, then $f(K)$ is a compact subset of Y.

4. Assume that $K_1, K_2, \ldots, K_n$ are compact subsets of a metric space X. Use the open covering property to show that $K_1 \cup K_2 \cup \ldots \cup K_n$ is compact.

5. Use the open covering property to show that a closed subset of a compact set is compact.

6. Assume that $f : X \to Y$ is a continuous function between two metric spaces, and assume that K is a compact subset of X. We shall prove that f is *uniformly continuous*, i.e. that for each $\epsilon > 0$, there exists a $\delta > 0$ such that whenever $x, y \in K$ and $d_X(x, y) < \delta$, then $d_Y(f(x), f(y)) < \epsilon$ (this looks very much like ordinary continuity, but the point is that we can use the *same* δ at all points $x, y \in K$).

   a) Given $\epsilon > 0$, explain that for each $x \in K$ there is a $\delta(x) > 0$ such that $d_Y(f(x), f(y)) < \frac{\epsilon}{2}$ for all y with $d(x, y) < \delta(x)$.

   b) Explain that $\{\mathrm{B}(x, \frac{\delta(x)}{2})\}_{x\in K}$ is an open cover of X, and that it has a finite subcover $\mathrm{B}(x_1, \frac{\delta(x_1)}{2})$, $\mathrm{B}(x_2, \frac{\delta(x_2)}{2})$, $\ldots$, $\mathrm{B}(x_n, \frac{\delta(x_n)}{2})$.

   c) Put $\delta = \min\{\frac{\delta(x_1)}{2}, \frac{\delta(x_2)}{2}, \ldots, \frac{\delta(x_n)}{2}\}$, and show that if $x, y \in K$ with $d_X(x, y) < \delta$, then $d_Y(f(x), f(y)) < \epsilon$.

2.7 The completion of a metric space

Completeness is probably the most important notion in this book as most of the deep and important theorems about metric spaces only hold when space is complete. In this section we shall see that it is always possible to make an incomplete space complete by adding new elements, but before we turn to this, we need to take a look at a concept that will be important in many different contexts throughout the book.

Definition 2.7.1 *Let (X, d) be a metric space and assume that D is a subset of X. We say that D is* dense *in X if for each $x \in X$ there is a sequence $\{y_n\}$ from D converging to x.*

We know that $\mathbb{Q}$ is dense in $\mathbb{R}$ — we may, e.g., approximate a real number by longer and longer parts of its decimal expansion. For $x = \sqrt{2}$ this would mean the approximating sequence

$$y_1 = 1.4 = \frac{14}{10},\ \ y_2 = 1.41 = \frac{141}{100},\ \ y_3 = 1.414 = \frac{1414}{1000},\ \ y_4 = 1.4142 = \frac{14142}{10000}, \ldots$$

There is an alternative description of dense that we shall also need.

Proposition 2.7.2 *A subset D of a metric space X is dense if and only if for each $x \in X$ and each $\delta > 0$, there is a $y \in D$ such that $d(x, y) \leq \delta$.*

Proof: Left as an exercise. □

We are now ready to return to our initial problem: How do we extend an incomplete metric space to a complete one? The following definition describes what we are looking for.

Definition 2.7.3 *If (X, d_X) is a metric space, a* completion *of (X, d_X) is a metric space $(\bar{X}, d_{\bar{X}})$ such that:*

(i) (X, d_X) is a subspace of $(\bar{X}, d_{\bar{X}})$; i.e. $X \subseteq \bar{X}$ and $\bar{d}(x, y)) = d(x, y)$ for all $x, y \in X$.

(ii) X is dense $(\bar{X}, d_{\bar{X}})$.

The canonical example of a completion is that $\mathbb{R}$ is the completion $\mathbb{Q}$. We also note that a complete metric space is its own (unique) completion.

An incomplete metric space will have more than one completion, but as they are all isometric[2], they are the same for all practical purposes, and we usually talk about *the* completion of a metric space.

Proposition 2.7.4 *Assume that (Y, d_Y) and (Z, d_Z) are completions of the metric space (X, d_X). Then (Y, d_Y) and (Z, d_Z) are isometric.*

Proof: We shall construct an isometry $i : Y \to Z$. Since X is dense in Y, there is for each $y \in Y$ a sequence $\{x_n\}$ from X converging to y. This sequence must be a Cauchy sequence in X and hence in Z. Since Z is complete, $\{x_n\}$ converges to an element $z \in Z$. The idea is to define i by letting $i(y) = z$. For the definition to work properly, we have to check that if $\{\hat{x}_n\}$ is another sequence in X converging to y, then $\{\hat{x}_n\}$ converges to z in Z. This is the case since $d_Z(x_n, \hat{x}_n) = d_X(x_n, \hat{x}_n) = d_Y(x_n, \hat{x}_n) \to 0$ as $n \to \infty$.

To prove that i preserves distances, assume that $y, \hat{y}$ are two points in Y, and that $\{x_n\}, \{\hat{x}_n\}$ are sequences in X converging to y and $\hat{y}$, respectively. Then $\{x_n\}, \{\hat{x}_n\}$ converges to $i(y)$ and $i(\hat{y})$, respectively, in Z, and we have

$$d_Z(i(y), i(\hat{y})) = \lim_{n\to\infty} d_Z(x_n, \hat{x}_n) = \lim_{n\to\infty} d_X(x_n, \hat{x}_n) =$$

$$= \lim_{n\to\infty} d_Y(x_n, \hat{x}_n) = d_Y(y, \hat{y})$$

[2]Recall from Section 2.1 that an *isometry* from (X, d_X) to (Y, d_Y) is a bijection $i : X \to Y$ such that $d_Y(i(x), i(y)) = d_X(x, y)$ for all $x, y \in X$. Two metric spaces are often considered "the same" when they are isomorphic; i.e. when there is an isomorphism between them.

It remains to prove that i is a bijection. Injectivity follows immediately from distance preservation: If $y \neq \hat{y}$, then $d_Z(i(y), i(\hat{y})) = d_Y(y, \hat{y}) \neq 0$, and hence $i(y) \neq i(\hat{y})$. To show that i is surjective, consider an arbitrary element $z \in Z$. Since X is dense in Z, there is a sequence $\{x_n\}$ from X converging to z. Since Y is complete, $\{x_n\}$ is also converging to an element y in Y. By construction, $i(y) = z$, and hence i is surjective. □

We shall use the rest of the section to show that all metric spaces (X, d) have a completion. The construction is longer and more complicated than most others in this book, but also quite instructive as it is typical of a type of construction that is very common in mathematics. As what we want to construct is a space where all Cauchy sequences from X has a limit, it is not unnatural to start with the set $\mathcal{X}$ of all Cauchy sequences and see if we can turn it into a metric space containing X.

The first lemma gives us the information we need to construct a metric.

Lemma 2.7.5 *Assume that $\{x_n\}$ and $\{y_n\}$ are two Cauchy sequences in a metric space (X, d). Then $\lim_{n\to\infty} d(x_n, y_n)$ exists.*

Proof: As $\mathbb{R}$ is complete, it suffices to show that $\{d(x_n, y_n)\}$ is a Cauchy sequence. We have

$$|d(x_n, y_n) - d(x_m, y_m)| = |d(x_n, y_n) - d(x_m, y_n) + d(x_m, y_n) - d(x_m, y_m)| \leq$$

$$\leq |d(x_n, y_n) - d(x_m, y_n)| + |d(x_m, y_n) - d(x_m, y_m)| \leq d(x_n, x_m) + d(y_n, y_m)$$

where we have used the inverse triangle inequality (Proposition 2.1.4) in the final step. Since $\{x_n\}$ and $\{y_n\}$ are Cauchy sequences, we can get $d(x_n, x_m)$ and $d(y_n, y_m)$ as small as we wish by choosing n and m sufficiently large, and hence $\{d(x_n, y_n)\}$ is a Cauchy sequence. □

As already mentioned, we let $\mathcal{X}$ be the set of all Cauchy sequences on the metric space (X, d_X). We want to turn $\mathcal{X}$ into a metric space by using the "metric" $\bar{d}(\{x_n\}, \{y_n\}) = \lim_{n\to\infty} d(x_n, y_n)$ to measure the distance between the sequences $\{x_n\}$ and $\{y_n\}$, but before we can do this, we have to identify Cauchy sequences that will converge to the same point in any completion. To this end we introduce a relation $\sim$ on $\mathcal{X}$ by

$$\{x_n\} \sim \{y_n\} \iff \lim_{n\to\infty} d(x_n, y_n) = 0$$

Lemma 2.7.6 $\sim$ *is an equivalence relation.*

Proof: We have to check the three properties in Definition 1.5.2:
Reflexivity: Since $\lim_{n\to\infty} d(x_n, x_n) = 0$, the relation is reflexiv.
Symmetry: Since $\lim_{n\to\infty} d(x_n, y_n) = \lim_{n\to\infty} d(y_n, x_n)$, the relation is symmetric.

Transitivity: Assume that $\{x_n\} \sim \{y_n\}$ og $\{y_n\} \sim \{z_n\}$. Then $\lim_{n\to\infty} d(x_n, y_n) = \lim_{n\to\infty} d(y_n, z_n) = 0$, and consequently

$$0 \leq \lim_{n\to\infty} d(x_n, z_n) \leq \lim_{n\to\infty} \big(d(x_n, y_n) + d(y_n, z_n)\big) =$$

$$= \lim_{n\to\infty} d(x_n, y_n) + \lim_{n\to\infty} d(y_n, z_n) = 0$$

which shows that $\{x_n\} = \{y_n\}$. □

We shall denote the equivalence class of $\{x_n\}$ by $[x_n]$, and we let $\bar{X}$ be the set of all equivalence classes. The next lemma will allow us to define a natural metric on $\bar{X}$.

Lemma 2.7.7 *If* $\{x_n\} \sim \{\hat{x}_n\}$ *and* $\{y_n\} \sim \{\hat{y}_n\}$, *then* $\lim_{n\to\infty} d(x_n, y_n) = \lim_{n\to\infty} d(\hat{x}_n, \hat{y}_n)$.

Proof: Since $d(x_n, y_n) \leq d(x_n, \hat{x}_n) + d(\hat{x}_n, \hat{y}_n) + d(\hat{y}_n, y_n)$ by the triangle inequality, and $\lim_{n\to\infty} d(x_n, \hat{x}_n) = \lim_{n\to\infty} d(\hat{y}_n, y_n) = 0$, we get

$$\lim_{n\to\infty} d(x_n, y_n) \leq \lim_{n\to\infty} d(\hat{x}_n, \hat{y}_n)$$

By reversing the roles of elements with and without hats, we get the opposite inequality. □

We may now define a function $\bar{d} : \bar{X} \times \bar{X} \to [0, \infty)$ by

$$\bar{d}([x_n], [y_n]) = \lim_{n\to\infty} d(x_n, y_n)$$

Note that by the previous lemma $\bar{d}$ is *well-defined*; i.e. the value of $\bar{d}([x_n], [y_n])$ does not depend on which representatives $\{x_n\}$ and $\{y_n\}$ we choose from the equivalence classes $[x_n]$ and $[y_n]$.

Lemma 2.7.8 $(\bar{X}, \bar{d})$ *is a metric space.*

Proof: We need to check the three conditions in the definition of a metric space.
Positivity: Clearly $\bar{d}([x_n], [y_n]) = \lim_{n\to\infty} d(x_n, y_n) \geq 0$, and by definition of the equivalence relation, we have equality if and only if $[x_n] = [y_n]$.
Symmetry: Since the underlying metric d is symmetric, we have

$$\bar{d}([x_n], [y_n]) = \lim_{n\to\infty} d(x_n, y_n) = \lim_{n\to\infty} d(y_n, x_n) = \bar{d}([y_n], [x_n])$$

Triangle inequality: For all equivalence classes $[x_n], [y_n], [z_n]$, we have

$$\bar{d}([x_n], [z_n]) = \lim_{n\to\infty} d(x_n, z_n) \leq \lim_{n\to\infty} d(x_n, y_n) + \lim_{n\to\infty} d(y_n, z_n) =$$

$$= \bar{d}([x_n],[y_n]) + \bar{d}([y_n],[z_n])$$

□

For each $x \in X$, let $\bar{x}$ be the equivalence class of the constant sequence $\{x, x, x, \ldots\}$. Since $\bar{d}(\bar{x}, \bar{y}) = \lim_{n\to\infty} d(x, y) = d(x, y)$, the mapping $x \to \bar{x}$ is an embedding[3] of X into $\bar{X}$. Hence $\bar{X}$ contains a copy of X, and the next lemma shows that this copy is dense in $\bar{X}$.

Lemma 2.7.9 *The set*

$$D = \{\bar{x} \,:\, x \in X\}$$

is dense in $\bar{X}$*.*

Proof: Assume that $[x_n] \in \bar{X}$. It suffices to show (see Problem 4) that for each $\epsilon > 0$ there is an $\bar{x} \in D$ such that $\bar{d}(\bar{x}, [x_n]) < \epsilon$. Since $\{x_n\}$ is a Cauchy sequence, there is an $N \in \mathbb{N}$ such that $d(x_n, x_N) < \frac{\epsilon}{2}$ for all $n \geq N$. Put $x = x_N$. Then $\bar{d}([x_n], \bar{x}) = \lim_{n\to\infty} d(x_n, x_N) \leq \frac{\epsilon}{2} < \epsilon$. □

It still remains to prove that $(\bar{X}, \bar{d})$ is complete. The next lemma is the first step in this direction.

Lemma 2.7.10 *All Cauchy sequences in D converges to an element in $\bar{X}$.*

Proof: Let $\{\bar{u}_k\}$ be a Cauchy sequence in D. Since $d(u_n, u_m) = \bar{d}(\bar{u}_n, \bar{u}_m)$, $\{u_n\}$ is a Cauchy sequence in X, and gives rise to an element $[u_n]$ in $\bar{X}$. To see that $\{\bar{u}_k\}$ converges to $[u_n]$, note that $\bar{d}(\bar{u}_k, [u_n]) = \lim_{n\to\infty} d(u_k, u_n)$. Since $\{u_n\}$ is a Cauchy sequence, this limit decreases to 0 as k goes to infinity. □

We are now ready to prove completeness:

Lemma 2.7.11 $(\bar{X}, \bar{d})$ *is complete.*

Proof: Let $\{x_n\}$ be a Cauchy sequence in $\bar{X}$. Since D is dense in $\bar{X}$, there is for each n a $y_n \in D$ such that $\bar{d}(x_n, y_n) < \frac{1}{n}$. It is easy to check that since $\{x_n\}$ is a Cauchy sequence, so is $\{y_n\}$. By the previous lemma, $\{y_n\}$ converges to an element in $\bar{X}$, and by construction $\{x_n\}$ must converge to the same element. Hence $(\bar{X}, \bar{d})$ is complete. □

We have reached the main theorem.

Theorem 2.7.12 *Every metric space (X, d) has a completion.*

[3] Recall Definition 2.1.3

Proof: We have already proved that $(\bar{X}, \bar{d})$ is a complete metric space that contains $D = \{\bar{x} : x \in X\}$ as a dense subset. In addition, we know that D is a copy of X (more precisel, $x \to \bar{x}$ is an isometry from X to D). All we have to do, is to replace the elements $\bar{x}$ in D by the original elements x in X, and we have found a completion of X. □

Remark: The theorem above doesn't solve all problems with incomplete spaces as there may be additional structure we want the completion to reflect. If, e.g., the original space consists of functions, we may want the completion also to consist of functions, but there is nothing in the construction above that guarantees that this is possible. We shall return to this question in later chapters.

Problems to Section 2.7

1. Prove Proposition 2.7.2.
2. Let us write $(X, d_X) \sim (Y, d_Y)$ to indicate that the two spaces are isometric. Show that
   (i) $(X, d_X) \sim (X, d_X)$
   (ii) If $(X, d_X) \sim (Y, d_Y)$, then $(Y, d_Y) \sim (X, d_X)$
   (iii) If $(X, d_X) \sim (Y, d_Y)$ and $(Y, d_Y) \sim (Z, d_Z)$, then $(X, d_X) \sim (Z, d_Z)$.
3. Show that the only completion of a complete metric space is the space itself.
4. Show that $\mathbb{R}$ is the completion of $\mathbb{Q}$ (in the usual metrics).
5. Assume that $i : X \to Y$ is an isometry between two metric spaces (X, d_X) and (Y, d_Y).
   (i) Show that a sequence $\{x_n\}$ converges in X if and only if $\{i(x_n)\}$ converges in Y.
   (ii) Show that a set $A \subseteq X$ is open/closed/compact if and only if $i(A)$ is open/closed/compact.

Chapter 3

Spaces of continuous functions

In this chapter we shall apply the theory we developed in the previous chapter to spaces where the elements are continuous functions. We shall study completeness and compactness of such spaces and take a look at some applications.

3.1 Modes of continuity

If (X, d_X) and (Y, d_Y) are two metric spaces, the function $f : X \to Y$ is continuous at a point a if for each $\epsilon > 0$ there is a $\delta > 0$ such that $d_Y(f(x), f(a)) < \epsilon$ whenever $d_X(x, a) < \delta$. If f is also continuous at another point b, we may need a different δ to match the same ϵ. A question that often comes up is when we can use the *same* δ for *all* points x in the space X. The function is then said to be *uniformly continuous* in X. Here is the precise definition:

Definition 3.1.1 *Let $f : X \to Y$ be a function between two metric spaces. We say that f is* uniformly continuous *if for each $\epsilon > 0$ there is a $\delta > 0$ such that $d_Y(f(x), f(y)) < \epsilon$ for all points $x, y \in X$ such that $d_X(x, y) < \delta$.*

A function which is continuous at all points in X, but not uniformly continuous, is often called *pointwise continuous* when we want to emphasize the distinction.

Example 1 The function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = x^2$ is pointwise continuous, but not uniformly continuous. The reason is that the curve becomes steeper and steeper as $|x|$ goes to infinity, and that we hence need increasingly smaller δ's to match the same ϵ (make a sketch!) See Exercise 1 for a more detailed discussion. ♣

If the underlying space X is compact, pointwise continuity and uniform continuity are the same. This means that a continuous function defined on a closed and bounded subset of $\mathbb{R}^n$ is always uniformly continuous.

Proposition 3.1.2 *Assume that X and Y are metric spaces. If X is compact, all continuous functions $f : X \to Y$ are uniformly continuous.*

Proof: We argue contrapositively: Assume that f is *not* uniformly continuous; we shall show that f is not continuous.

Since f fails to be uniformly continuous, there is an $\epsilon > 0$ we cannot match; i.e. for each $\delta > 0$ there are points $x, y \in X$ such that $d_X(x, y) < \delta$, but $d_Y(f(x), f(y)) \geq \epsilon$. Choosing $\delta = \frac{1}{n}$, there are thus points $x_n, y_n \in X$ such that $d_X(x_n, y_n) < \frac{1}{n}$ and $d_Y(f(x_n), f(y_n)) \geq \epsilon$. Since X is compact, the sequence $\{x_n\}$ has a subsequence $\{x_{n_k}\}$ converging to a point a. Since $d_X(x_{n_k}, y_{n_k}) < \frac{1}{n_k}$, the corresponding sequence $\{y_{n_k}\}$ of y's must also converge to a. We are now ready to show that f is not continuous at a: Had it been, the two sequences $\{f(x_{n_k})\}$ and $\{f(y_{n_k})\}$ would both have converged to $f(a)$, something they clearly can not since $d_Y(f(x_n), f(y_n)) \geq \epsilon$ for all $n \in \mathbb{N}$. □

There is an even more abstract form of continuity that will be important later. This time we are not considering a single function, but a whole collection of functions:

Definition 3.1.3 *Let (X, d_X) and (Y, d_Y) be metric spaces, and let $\mathcal{F}$ be a collection of functions $f : X \to Y$. We say that $\mathcal{F}$ is* equicontinuous *if for all $\epsilon > 0$, there is a $\delta > 0$ such that for all $f \in \mathcal{F}$ and all $x, y \in X$ with $d_X(x, y) < \delta$, we have $d_Y(f(x), f(y)) < \epsilon$.*

Note that in the case, the same δ should not only hold at all points $x, y \in X$, but also for all functions $f \in \mathcal{F}$.

Example 2 Let $\mathcal{F}$ be the set of all contractions $f : X \to X$. Then $\mathcal{F}$ is equicontinuous, since we can can choose $\delta = \epsilon$. To see this, just note that if $d_X(x, y) < \delta = \epsilon$, then $d_X(f(x), f(y)) \leq d_X(x, y) < \epsilon$ for all $x, y \in X$ and all $f \in \mathcal{F}$. ♣

Equicontinuous families will be important when we study compact sets of continuous functions in Section 3.5.

Exercises for Section 3.1

1. Show that the function $f(x) = x^2$ is not uniformly continuous on $\mathbb{R}$. (*Hint:* You may want to use the factorization $f(x) - f(y) = x^2 - y^2 = (x+y)(x-y)$).

2. Prove that the function $f : (0, 1) \to \mathbb{R}$ given by $f(x) = \frac{1}{x}$ is not uniformly continuous.
3. A function $f : X \to Y$ between metric spaces is said to be *Lipschitz-continuous with Lipschitz constant* K if $d_Y(f(x), f(y)) \leq K d_X(x, y)$ for all $x, y \in X$. Asume that $\mathcal{F}$ is a collection of functions $f : X \to Y$ with Lipschitz constant K. Show that $\mathcal{F}$ is equicontinuous.
4. Let $f : \mathbb{R} \to \mathbb{R}$ be a differentiable function and assume that the derivative f' is bounded. Show that f is uniformly continuous.

3.2 Modes of convergence

In this section we shall study two ways in which a sequence $\{f_n\}$ of continuous functions can converge to a limit function f: *pointwise convergence* and *uniform convergence*. The distinction is rather simililar to the distinction between pointwise and uniform continuity in the previous section — in the pointwise case, a condition can be satisfied in different ways for different x's; in the uniform, case it must be satisfied in the same way for all x. We begin with pointwise convergence:

Definition 3.2.1 *Let (X, d_X) and (Y, d_Y) be two metric spaces, and let $\{f_n\}$ be a sequence of functions $f_n : X \to Y$. We say that $\{f_n\}$* converges pointwise *to a function $f : X \to Y$ if $f_n(x) \to f(x)$ for all $x \in X$. This means that for each x and each $\epsilon > 0$, there is an $N \in \mathbb{N}$ such that $d_Y(f_n(x), f(x)) < \epsilon$ when $n \geq N$.*

Note that the N in the last sentence of the definition depends on x — we may need a much larger N for some x's than for others. If we can use the *same* N for all $x \in X$, we have uniform convergence. Here is the precise definition:

Definition 3.2.2 *Let (X, d_X) and (Y, d_Y) be two metric spaces, and let $\{f_n\}$ be a sequence of functions $f_n : X \to Y$. We say that $\{f_n\}$* converges uniformly *to a function $f : X \to Y$ if for each $\epsilon > 0$, there is an $N \in \mathbb{N}$ such that if $n \geq N$, then $d_Y(f_n(x), f(x)) < \epsilon$ for all $x \in X$.*

At first glance, the two definitions may seem confusingly similar, but the difference is that in the last one, the *same* N should work simultaneously for all x, while in the first we can adapt N to each individual x. Hence uniform convergence implies pointwise convergence, but a sequence may converge pointwise but not uniformly. Before we look at an example, it will be useful to reformulate the definition of uniform convergence.

Proposition 3.2.3 *Let (X, d_X) and (Y, d_Y) be two metric spaces, and let $\{f_n\}$ be a sequence of functions $f_n : X \to Y$. For any function $f : X \to Y$ the following are equivalent:*

(i) $\{f_n\}$ *converges uniformly to* f.

(ii) $\sup\{d_Y(f_n(x), f(x)) \,|\, x \in X\} \to 0$ *as* $n \to \infty$.

Hence uniform convergence means that the "maximal" distance between f *and* f_n *goes to zero.*

Proof: $(i) \implies (ii)$ Assume that $\{f_n\}$ converges uniformly to f. For any $\epsilon > 0$, we can find an $N \in \mathbb{N}$ such that $d_Y(f_n(x), f(x)) < \epsilon$ for all $x \in X$ and all $n \geq N$. This means that $\sup\{d_Y(f_n(x), f(x)) \,|\, x \in X\} \leq \epsilon$ for all $n \geq N$ (note that we may have unstrict inequality $\leq$ for the supremum although we have strict inequality $<$ for each $x \in X$), and since ϵ is arbitrary, this implies that $\sup\{d_Y(f_n(x), f(x)) \,|\, x \in X\} \to 0$.

$(ii) \implies (i)$ Assume that $\sup\{d_Y(f_n(x), f(x)) \,|\, x \in X\} \to 0$ as $n \to \infty$. Given an $\epsilon > 0$, there is an $N \in \mathbb{N}$ such that $\sup\{d_Y(f_n(x), f(x)) \,|\, x \in X\} < \epsilon$ for all $n \geq N$. But then we have $d_Y(f_n(x), f(x)) < \epsilon$ for all $x \in X$ and all $n \geq N$, which means that $\{f_n\}$ converges uniformly to f. □

Here is an example which shows clearly the distinction between pointwise and uniform convergence:

Example 1 Let $f_n : [0, 1] \to \mathbb{R}$ be the function in Figure 1. It is constant zero except on the interval $[0, \frac{1}{n}]$ where it looks like a tent of height 1.

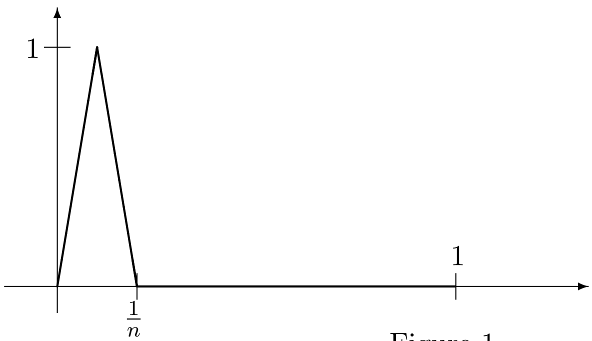


Figure 1

If you insist, the function is defined by

$$f_n(x) = \begin{cases} 2nx & \text{if } 0 \leq x < \frac{1}{2n} \\ -2nx + 2 & \text{if } \frac{1}{2n} \leq x < \frac{1}{n} \\ 0 & \text{if } \frac{1}{n} \leq x \leq 1 \end{cases}$$

but it is much easier just to work from the picture.

The sequence $\{f_n\}$ converges pointwise to 0, because at every point $x \in [0, 1]$ the value of $f_n(x)$ eventually becomes 0 (for $x = 0$, the value is always

0, and for $x > 0$ the "tent" will eventually pass to the left of x.) However, since the maximum value of all f_n is 1, $\sup\{d_Y(f_n(x), f(x)) \,|\, x \in [0,1]\} = 1$ for all n, and hence $\{f_n\}$ does not converge uniformly to 0. ♣

When we are working with convergent sequences, we would often like the limit to inherit properties from the elements in the sequence. If, e.g., $\{f_n\}$ is a sequence of *continuous* functions converging to a limit f, we are often interested in showing that f is also continuous. The next example shows that this is not always the case when we are dealing with pointwise convergence.

Example 2: Let $f_n : \mathbb{R} \to \mathbb{R}$ be the function in Figure 2.

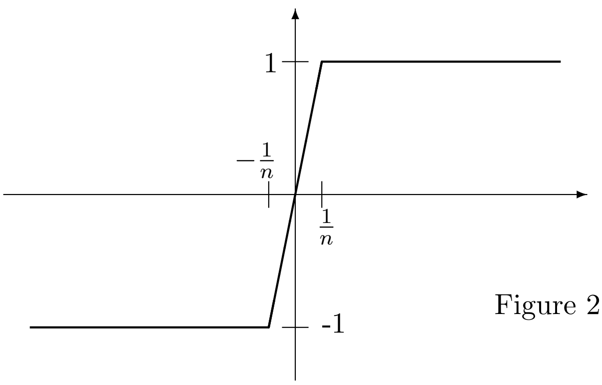


Figure 2

It is defined by

$$f_n(x) = \begin{cases} -1 & \text{if } x \leq -\frac{1}{n} \\ \\ nx & \text{if } -\frac{1}{n} < x < \frac{1}{n} \\ \\ 1 & \text{if } \frac{1}{n} \leq x \end{cases}$$

The sequence $\{f_n\}$ converges pointwise to the function, f defined by

$$f(x) = \begin{cases} -1 & \text{if } x < 0 \\ \\ 0 & \text{if } x = 0 \\ \\ 1 & \text{if } x > 0 \end{cases}$$

but although all the functions $\{f_n\}$ are continuous, the limit function f is not. ♣

If we strengthen the convergence from pointwise to uniform, the limit of a sequence of continuous functions is always continuous.

Proposition 3.2.4 *Let (X, d_X) and (Y, d_Y) be two metric spaces, and assume that $\{f_n\}$ is a sequence of continuous functions $f_n : X \to Y$ converging uniformly to a function f. Then f is continuous.*

Proof: Let $a \in X$. Given an $\epsilon > 0$, we must find a $\delta > 0$ such that $d_Y(f(x), f(a)) < \epsilon$ whenever $d_X(x, a) < \delta$. Since $\{f_n\}$ converges uniformly to f, there is an $N \in \mathbb{N}$ such that when $n \geq N$, $d_Y(f(x), f_n(x)) < \frac{\epsilon}{3}$ for *all* $x \in X$. Since f_N is continuous at a, there is a $\delta > 0$ such that $d_Y(f_N(x), f_N(a)) < \frac{\epsilon}{3}$ whenever $d_X(x, a) < \delta$. If $d_X(x, a) < \delta$, we then have

$$d_Y(f(x), f(a)) \leq d_Y(f(x), f_N(x)) + d_Y(f_N(x), f_N(a)) + d_Y(f_N(a), f(a)) <$$
$$< \frac{\epsilon}{3} + \frac{\epsilon}{3} + \frac{\epsilon}{3} = \epsilon$$

and hence f is continuous at a. □

The technique in the proof above is quite common, and arguments of this kind are often referred to as $\frac{\epsilon}{3}$-arguments.

Exercises for Section 3.2

1. Let $f_n : \mathbb{R} \to \mathbb{R}$ be defined by $f_n(x) = \frac{x}{n}$. Show that $\{f_n\}$ converges pointwise, but not uniformly to 0.
2. Let $f_n : (0, 1) \to \mathbb{R}$ be defined by $f_n(x) = x^n$. Show that $\{f_n\}$ converges pointwise, but not uniformly to 0.
3. The function $f_n : [0, \infty) \to \mathbb{R}$ is defined by $f_n(x) = e^{-x} \left(\frac{x}{n}\right)^{ne}$.
   a) Show that $\{f_n\}$ converges pointwise.
   b) Find the maximum value of f_n. Does $\{f_n\}$ converge uniformly?
4. The function $f_n : (0, \infty) \to \mathbb{R}$ is defined by
   $$f_n(x) = n(x^{1/n} - 1)$$
   Show that $\{f_n\}$ converges pointwise to $f(x) = \ln x$. Show that the convergence is uniform on each interval $(\frac{1}{k}, k)$, $k \in \mathbb{N}$, but not on $(0, \infty)$.
5. Let $f_n : \mathbb{R} \to \mathbb{R}$ and assume that the sequence $\{f_n\}$ of continuous functions converges uniformly to $f : \mathbb{R} \to \mathbb{R}$ on all intervals $[-k, k]$, $k \in \mathbb{N}$. Show that f is continuous.
6. Assume that X is a metric space and that f_n, g_n are functions from X to $\mathbb{R}$. Show that if $\{f_n\}$ and $\{g_n\}$ converge uniformly to f and g, respectively, then $\{f_n + g_n\}$ converges uniformly to $f + g$.
7. Assume that $f_n : [a, b] \to \mathbb{R}$ are continuous functions converging uniformly to f. Show that
   $$\int_a^b f_n(x)\, dx \to \int_a^b f(x)\, dx$$
   Find an example which shows that this is not necessarily the case if $\{f_n\}$ only converges pointwise to f.

8. Let $f_n : \mathbb{R} \to \mathbb{R}$ be given by $f_n(x) = \frac{1}{n}\sin(nx)$. Show that $\{f_n\}$ converges uniformly to 0, but that the sequence $\{f'_n\}$ of derivates does not converge. Sketch the graphs of f_n to see what is happening.

9. Let (X,d) be a metric space and assume that the sequence $\{f_n\}$ of continuous functions converges uniformly to f. Show that if $\{x_n\}$ is a sequence in X converging to x, then $f_n(x_n) \to f(x)$. Find an example which shows that this is not necessarily the case if $\{f_n\}$ only converges pointwise to f.

10. Assume that the functions $f_n : X \to Y$ converges uniformly to f, and that $g : Y \to Z$ is uniformly continuous. Show that the sequence $\{g \circ f_n\}$ converges uniformly. Find an example which shows that the conclusion does not necessarily hold if g is only pointwise continuous.

11. Assume that $\sum_{n=0}^{\infty} M_n$ is a convergent series of positive numbers. Assume that $f_n : X \to \mathbb{R}$ is a sequence of continuous functions defined on a metric space (X,d). Show that if $|f_n(x)| \leq M_n$ for all $x \in X$ and all $n \in N$, then the partial sums $s_N(x) = \sum_{n=0}^{N} f_n(x)$ converge uniformly to a continuous function $s : X \to \mathbb{R}$ as $N \to \infty$. (This is called *Weierstrass' M-test*).

12. Assume that (X,d) is a compact space and that $\{f_n\}$ is a decreasing sequence of continuous functions converging pointwise to a continuous function f. Show that the convergence is uniform (this is called *Dini's theorem*).

3.3 The spaces $C(X,Y)$

If (X,d_X) and (Y,d_Y) are metric spaces, we let

$$C(X,Y) = \{f : X \to Y \mid f \text{ is continuous}\}$$

be the collection of all continuous functions from X to Y. In this section we shall see how we can turn $C(X,Y)$ into a metric space. To avoid certain technicalities, we shall restrict ourselves to the case where X is compact as this is sufficient to cover most interesting applications (see Exercise 4 for one possible way of extending the theory to the non-compact case).

The basic idea is to measure the distance between two functions by looking at the point they are the furthest apart; i.e. by

$$\rho(f,g) = \sup\{d_Y(f(x),g(x)) \mid x \in X\}$$

Our first task is to show that ρ is a metric on $C(X,Y)$. But first we need a lemma:

Lemma 3.3.1 *Let (X,d_X) and (Y,d_Y) be metric spaces, and assume that X is compact. If $f,g : X \to Y$ are continuous functions, then*

$$\rho(f,g) = \sup\{d_Y(f(x),g(x)) \mid x \in X\}$$

is finite, and there is a point $x \in X$ such that $d_Y(f(x),g(x)) = \rho(f,g)$.

Proof: The result will follow from the Extreme Value Theorem (Theorem 2.5.9) if we can only show that the function

$$h(x) = d_Y(f(x), g(x))$$

is continiuous. By the triangle inequality for numbers and the inverse triangle inequality 2.1.4, we get

$$|h(x) - h(y)| = |d_Y(f(x), g(x)) - d_Y(f(y), g(y))| =$$
$$= |d_Y(f(x), g(x)) - d_Y(f(x), g(y)) + d_Y(f(x), g(y)) - d_Y(f(y), g(y))| \leq$$
$$\leq |d_Y(f(x), g(x)) - d_Y(f(x), g(y))| + |d_Y(f(x), g(y)) - d_Y(f(y), g(y))| \leq$$
$$\leq d_Y(g(x), g(y)) + d_Y(f(x), f(y))$$

To prove that h is continuous at x, just observe that since f and g are continuous at x, there is for any given $\epsilon > 0$ a $\delta > 0$ such that $d_Y(f(x), f(y)) < \frac{\epsilon}{2}$ and $d_Y(g(x), g(y)) < \frac{\epsilon}{2}$ when $d_X(x, y) < \delta$. But then

$$|h(x) - h(y)| \leq d_Y(f(x), f(y)) + d_Y(g(y), g(x)) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

whenever $d_X(x, y) < \delta$, and hence h is continuous. □

We are now ready to prove that ρ is a metric on $C(X, Y)$:

Proposition 3.3.2 *Let (X, d_X) and (Y, d_Y) be metric spaces, and assume that X is compact. Then*

$$\rho(f, g) = \sup\{d_Y(f(x), g(x)) \mid x \in X\}$$

defines a metric on $C(X, Y)$.

Proof: By the lemma, $\rho(f, g)$ is always finite, and we only have to prove that ρ satisfies the three properties of a metric: positivity, symmetry, and the triangle inequality. The first two are more or less obvious, and we concentrate on the triangle inequality:

Assume that f, g, h are three functions in $C(X, Y)$; we must show that

$$\rho(f, g) \leq \rho(f, h) + \rho(h, g)$$

According to the lemma, there is a point $x \in X$ such that $\rho(f, g) = d_Y(f(x), g(x))$. But then

$$\rho(f, g) = d_Y(f(x), g(x)) \leq d_Y(f(x), h(x)) + d_Y(h(x), g(x)) \leq \rho(f, h) + \rho(h, g)$$

where we have used the triangle inequality in Y and the definition of ρ. □

Not surprisingly, convergence in $C(X, Y)$ is exactly the same as uniform convergence.

Proposition 3.3.3 *A sequence $\{f_n\}$ converges to f in $(C(X, Y), \rho)$ if and only if it converges uniformly to f.*

Proof: According to Proposition 3.2.3, $\{f_n\}$ converges uniformly to f if and only if

$$\sup\{d_Y(f_n(x), f(x)) \,|\, x \in X\} \to 0$$

This just means that $\rho(f_n, f) \to 0$, which is to say that $\{f_n\}$ converges to f in $(C(X, Y), \rho)$. □

The next result is the starting point for many applications; it shows that $C(X, Y)$ is complete if Y is.

Theorem 3.3.4 *Assume that (X, d_X) is a compact and (Y, d_Y) a complete metric space. Then $C(X, Y), \rho)$ is complete.*

Proof: Assume that $\{f_n\}$ is a Cauchy sequence in $C(X, Y)$. We must prove that f_n converges to a function $f \in C(X, Y)$.

Fix an element $x \in X$. Since $d_Y(f_n(x), f_m(x)) \leq \rho(f_n, f_m)$ and $\{f_n\}$ is a Cauchy sequence in $(C(X, Y), \rho)$, the function values $\{f_n(x)\}$ form a Cauchy sequence in Y. Since Y is complete, $\{f_n(x)\}$ converges to a point $f(x)$ in Y. This means that $\{f_n\}$ converges *pointwise* to a function $f : X \to Y$. We must prove that $f \in C(X, Y)$ and that $\{f_n\}$ converges to f in the ρ-metric.

Since $\{f_n\}$ is a Cauchy sequence, we can for any $\epsilon > 0$ find an $N \in \mathbb{N}$ such that $\rho(f_n, f_m) < \frac{\epsilon}{2}$ when $n, m \geq N$. This means that all $x \in X$ and all $n, m \geq N$, $d_Y(f_n(x), f_m(x)) < \frac{\epsilon}{2}$. If we let $m \to \infty$, we see that for all $x \in X$ and all $n \geq N$

$$d_Y(f_n(x), f(x)) = \lim_{m\to\infty} d_Y(f_n(x), f_m(x)) \leq \frac{\epsilon}{2} < \epsilon$$

This means that $\{f_n\}$ converges uniformly to f. According to Proposition 3.2.4, f is continuos and belongs to $C(X, Y)$, and according to the proposition above, $\{f_n\}$ converges to f in $(C(X, Y), \rho)$. □

In the next section we shall combine the result above with Banach's Fixed Point Theorem to obtain our first real application.

Exercises to Section 3.3

1. Let $f, g : [0, 1] \to \mathbb{R}$ be given by $f(x) = x$, $g(x) = x^2$. Find $\rho(f, g)$.
2. Let $f, g : [0, 2\pi] \to \mathbb{R}$ be given by $f(x) = \sin x$, $g(x) = \cos x$. Find $\rho(f, g)$.
3. Complete the proof of Proposition 3.3.2 by showing that ρ satisfies the first two conditions of a metric (positivity and symmetry).

4. The main reason why we have restricted the theory above to the case where X is compact, is that if not,

$$\rho(f, g) = \sup\{d_Y(f(x), g(x)) \mid x \in X\}$$

may be infinite, and then ρ is not a metric. In this problem we shall sketch a way to avoid this problem.

A function $f : X \to Y$ is called *bounded* if there is a point $a \in Y$ and a constant $K \in \mathbb{R}$ such that $d_Y(a, f(x)) \leq K$ for all $x \in X$ (it doesn't matter which point a we use in this definition). Let $C_0(X, Y)$ be the set of all bounded, continuous functions $f : X \to Y$, and define

$$\rho(f, g) = \sup\{d_Y(f(x), g(x)) \mid x \in X\}$$

a) Show that $\rho(f, g) < \infty$ for all $f, g \in C_0(X, Y)$.
b) Show by an example that there need not be a point x in X such that $\rho(f, g) = d_Y(f(x), g(x))$.
c) Show that ρ is a metric on $C_0(X, Y)$.
d) Show that if a sequence $\{f_n\}$ of functions in $C_0(X, Y)$ converges uniformly to a function f, then $f \in C_0(X, Y)$.
e) Assume that (Y, d_Y) is complete. Show that $(C_0(X, Y), \rho)$ is complete.
f) Let c_0 be the set of all bounded sequences in $\mathbb{R}$. If $\{x_n\}$, $\{y_n\}$ are in c_0, define

$$\rho(\{x_n\}, \{y_n\}) = \sup(|x_n - y_n| : n \in \mathbb{N}\}$$

Prove that (c_0, ρ) is a complete metric space. (*Hint:* You may think of c_0 as $C_0(\mathbb{N}, \mathbb{R})$ where $\mathbb{N}$ has the discrete metric).

3.4 Applications to differential equations

Consider a system of differential equations

$$\begin{array}{rcl}
y_1'(t) & = & f_1(t, y_1(t), y_2(t), \ldots, y_n(t)) \\
y_2'(t) & = & f_2(t, y_1(t), y_2(t), \ldots, y_n(t)) \\
\vdots & \vdots & \quad\vdots \qquad\qquad \vdots \\
y_n'(t) & = & f_n(t, y_1(t), y_2(t), \ldots, y_n(t))
\end{array}$$

with initial conditions $y_1(0) = Y_1$, $y_2(0) = Y_2$, $\ldots$, $y_n(0) = Y_n$. In this section we shall use Banach's Fixed Point Theorem 2.4.5 and the completeness of $C([0, a], \mathbb{R}^n)$ to prove that under reasonable conditions such systems have a unique solution.

We begin by introducing vector notation to make the formulas easier to read:

$$\mathbf{y}(t) = \begin{pmatrix} y_1(t) \\ y_2(t) \\ \vdots \\ y_n(t) \end{pmatrix}$$

$$\mathbf{y}_0 = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{pmatrix}$$

and

$$\mathbf{f}(t, \mathbf{y}(t)) = \begin{pmatrix} f_1(t, y_1(t), y_2(t), \ldots, y_n(t)) \\ f_2(t, y_1(t), y_2(t), \ldots, y_n(t)) \\ \vdots \\ f_n(t, y_1(t), y_2(t), \ldots, y_n(t)) \end{pmatrix}$$

In this notation, the system becomes

$$\mathbf{y}'(t) = \mathbf{f}(t, \mathbf{y}(t)), \qquad \mathbf{y}(0) = \mathbf{y}_0 \tag{3.4.1}$$

The next step is to rewrite the differential equation as an integral equation. If we integrate on both sides of (3.4.1), we get

$$\mathbf{y}(t) - \mathbf{y}(0) = \int_0^t \mathbf{f}(s, \mathbf{y}(s))\, ds$$

i.e.

$$\mathbf{y}(t) = \mathbf{y}_0 + \int_0^t \mathbf{f}(s, \mathbf{y}(s))\, ds \tag{3.4.2}$$

On the other hand, if we start with a solution of (3.4.2) and differentiate, we arrive at (3.4.1). Hence solving (3.4.1) and (3.4.2) amounts to exactly the same thing, and for us it will be convenient to concentrate on (3.4.2).

Let us begin by putting an arbitrary, continuous function $\mathbf{z}$ into the right hand side of (3.4.2). What we get out is another function $\mathbf{u}$ defined by

$$\mathbf{u}(t) = y_0 + \int_0^t \mathbf{f}(s, \mathbf{z}(s))\, ds$$

We can think of this as a function F mapping continuous functions $\mathbf{z}$ to continuous functions $\mathbf{u} = F(\mathbf{z})$. From this point of view, a solution $\mathbf{y}$ of the integral equation (3.4.2) is just a fixed point for the function F — we are looking for a $\mathbf{y}$ such that $\mathbf{y} = F(\mathbf{y})$. (Don't worry if you feel a little dizzy; that's just normal at this stage! Note that F is a function acting on a function $\mathbf{z}$ to produce a new function $\mathbf{u} = F(\mathbf{z})$ — it takes some time to get used to such creatures!)

Our plan is to use Banach's Fixed Point Theorem to prove that F has a unique fixed point, but first we have to introduce a crucial condition. We say that the function $\mathbf{f} : [a, b] \times \mathbb{R}^n \to \mathbb{R}^n$ is *uniformly Lipschitz with Lipschitz constant K on the interval* $[a, b]$ if K is a real number such that

$$|\mathbf{f}(t, \mathbf{y}) - \mathbf{f}(t, \mathbf{z})| \leq K|\mathbf{y} - \mathbf{z}|$$

for all $t \in [a, b]$ and all $\mathbf{y}, \mathbf{z} \in \mathbb{R}^n$. Here is the key observation in our argument.

Lemma 3.4.1 *Assume that* $\mathbf{y}_0 \in \mathbb{R}^n$ *and that* $\mathbf{f} : [0,\infty) \times \mathbb{R}^n \to \mathbb{R}^n$ *is continuous and uniformly Lipschitz with Lipschitz constant* K *on* $[0,\infty)$*. If* $a < \frac{1}{K}$*, the map*

$$F : C([0,a],\mathbb{R}^n) \to C([0,a],\mathbb{R}^n)$$

defined by

$$F(\mathbf{z})(t) = \mathbf{y}_0 + \int_0^t \mathbf{f}(t,\mathbf{z}(t))\,dt$$

is a contraction.

Remark: The notation here is rather messy. Remember that $F(\mathbf{z})$ is a function from $[0,a]$ to $\mathbb{R}^n$. The expression $F(\mathbf{z})(t)$ denotes the value of this function at point $t \in [0,a]$.

Proof: Let $\mathbf{v},\mathbf{w}$ be two elements in $C([0,a],\mathbb{R}^n)$, and note that for any $t \in [0,a]$

$$|F(\mathbf{v})(t) - F(\mathbf{w})(t)| = |\int_0^t \big(\mathbf{f}(s,\mathbf{v}(s)) - \mathbf{f}(s,\mathbf{w}(s))\big)\,ds| \leq$$

$$\leq \int_0^t |\mathbf{f}(s,\mathbf{v}(s)) - \mathbf{f}(s,\mathbf{w}(s))|\,ds \leq \int_0^t K|\mathbf{v}(s) - \mathbf{w}(s)|\,ds \leq$$

$$\leq K\int_0^t \rho(\mathbf{v},\mathbf{w})\,ds \leq K\int_0^a \rho(\mathbf{v},\mathbf{w})\,ds = Ka\,\rho(\mathbf{v},\mathbf{w})$$

Taking the supremum over all $t \in [0,a]$, we get

$$\rho(F(\mathbf{v}),F(\mathbf{w})) \leq Ka\,\rho(\mathbf{v},\mathbf{w}).$$

Since $Ka < 1$, this means that F is a contraction. □

We are now ready for the main theorem.

Theorem 3.4.2 *Assume that* $\mathbf{y}_0 \in \mathbb{R}^n$ *and that* $\mathbf{f} : [0,\infty) \times \mathbb{R}^n \to \mathbb{R}^n$ *is continuous and uniformly Lipschitz on* $[0,\infty)$*. Then the initial value problem*

$$\mathbf{y}'(t) = \mathbf{f}(t,\mathbf{y}(t)), \qquad \mathbf{y}(0) = \mathbf{y}_0 \tag{3.4.3}$$

has a unique solution $\mathbf{y}$ *on* $[0,\infty)$*.*

Proof: Let K be the uniform Lipschitz constant, and choose a number $a < 1/K$. According to the lemma, the function

$$F : C([0,a],\mathbb{R}^n) \to C([0,a],\mathbb{R}^n)$$

defined by

$$F(\mathbf{z})(t) = \mathbf{y}_0 + \int_0^t \mathbf{f}(t,\mathbf{z}(t))\,dt$$

is a contraction. Since $C([0,a],\mathbb{R}^n)$ is complete by Theorem 3.3.4, Banach's Fixed Point Theorem tells us that F has a unique fixed point $\mathbf{y}$. This means that the integral equation

$$\mathbf{y}(t) = \mathbf{y}_0 + \int_0^t \mathbf{f}(s, \mathbf{y}(s))\, ds \tag{3.4.4}$$

has a unique solution on the interval $[0, a]$. To extend the solution to a longer interval, we just repeat the argument on the interval $[a, 2a]$, using $\mathbf{y}(a)$ as initial value. The function we then get, is a solution of the integral equation (3.4.4) on the extended interval $[0, 2a]$ as we for $t \in [a, 2a]$ have

$$\mathbf{y}(t) = \mathbf{y}(a) + \int_a^t \mathbf{f}(s, \mathbf{y}(s))\, ds =$$

$$= \mathbf{y}_0 + \int_0^a \mathbf{f}(s, \mathbf{y}(s))\, ds + \int_a^t \mathbf{f}(s, \mathbf{y}(s))\, ds = \mathbf{y}_0 + \int_0^t \mathbf{f}(s, \mathbf{y}(s))\, ds$$

Continuing this procedure to new intervals $[2a, 3a]$, $[3a, 4a]$, we see that the integral equation (3.4.3) has a unique solution on all of $[0, \infty)$. As we have already observed that equation (3.4.3) has exactly the same solutions as equation (3.4.4), the theorem is proved. □

In the exercises you will see that the conditions in the theorem are important. If they fail, the equation may have more than one solution, or a solution defined only on a bounded interval.

Exercises to Section 3.4

1. Solve the initial value problem

$$y' = 1 + y^2, \qquad y(0) = 0$$

   and show that the solution is only defined on the interval $[0, \pi/2)$.

2. Show that the functions

$$y(t) = \begin{cases} 0 & \text{if } 0 \le t \le a \\ (t-a)^{\frac{3}{2}} & \text{if } t > a \end{cases}$$

   where $a \geq 0$ are all solutions of the initial value problem

$$y' = \frac{3}{2} y^{\frac{1}{3}}, \qquad y(0) = 0$$

   Remember to check that the differential equation is satisfied at $t = a$.

3. In this problem we shall sketch how the theorem in this section can be used to study higher order systems. Assume we have a second order initial value problem

$$u''(t) = g(t, u(t), u'(t)) \qquad u(0) = a, u'(0) = b \qquad (*)$$

where $g : [0, \infty) \times \mathbb{R}^2 \to \mathbb{R}$ is a given function. Define a function $\mathbf{f} : [0, \infty) \times \mathbb{R}^2 \to \mathbb{R}^2$ by

$$\mathbf{f}(t, u, v) = \begin{pmatrix} v \\ g(t, u, v) \end{pmatrix}$$

Show that if

$$\mathbf{y}(t) = \begin{pmatrix} u(t) \\ v(t) \end{pmatrix}$$

is a solution of the initial value problem

$$\mathbf{y}'(t) = \mathbf{f}(t, \mathbf{y}(t)). \qquad \mathbf{y}(0) = \begin{pmatrix} a \\ b \end{pmatrix},$$

then u is a solution of the original problem $(*)$.

3.5 Compact subsets of $C(X, \mathbb{R}^m)$

The compact subsets of $\mathbb{R}^m$ are easy to describe — they are just the closed and bounded sets. This characterization is extremely useful as it is much easier to check that a set is closed and bounded than to check that it satisfies the definition of compactness. In the present section we shall prove a similar kind of characterization of compact sets in $C(X, \mathbb{R}^m)$ — we shall show that a subset of $C(X, \mathbb{R}^m)$ is compact if and only if it it closed, bounded and equicontinuous. This is known as the Arzelà-Ascoli Theorem. But before we turn to it, we have a question of independent interest to deal with. We have already encountered the notion of a dense set in Section 2.7, but repeat it here:

Definition 3.5.1 *Let (X, d) be a metric space and assume that A is a subset of X. We say that A is* dense *in X if for each $x \in X$ there is a sequence from A converging to x.*

Recall (Proposition 2.7.2) that dense sets can also be described in a slightly different way: A subset D of a metric space X is dense if and only if for each $x \in X$ and each $\delta > 0$, there is a $y \in D$ such that $d(x, y) \leq \delta$.

We know that $\mathbb{Q}$ is dense in $\mathbb{R}$ — we may, e.g., approximate a real number by longer and longer parts of its decimal expansion. For $x = \sqrt{2}$ this would mean the approximating sequence

$$a_1 = 1.4 = \frac{14}{10}, \ \ a_2 = 1.41 = \frac{141}{100}, \ \ a_3 = 1.414 = \frac{1414}{1000}, \ \ a_4 = 1.4142 = \frac{14142}{10000}, \ldots$$

Recall that $\mathbb{Q}$ is countable, but that $\mathbb{R}$ is not. Still every element in the uncountable set $\mathbb{R}$ can be approximated arbitrarily well by elements in the much smaller set $\mathbb{Q}$. This property turns out to be so useful that it deserves a name.

Definition 3.5.2 *A metric set* (X, d) *is called* separable *if it has a countable, dense subset* A.

Our first result is a simple, but rather surprising connection between separability and compactness.

Proposition 3.5.3 *All compact metric* (X, d) *spaces are separable. We can choose the countable dense set* A *in such a way that for any* $\delta > 0$*, there is a finite subset* A_δ *of* A *such that all elements of* X *are within distance less than* δ *of* A_δ*, i.e. for all* $x \in X$ *there is an* $a \in A_\delta$ *such that* $d(x, a) < \delta$.

Proof: We use that a compact space X is totally bounded (recall Theorem 2.5.12). This mean that for all $n \in \mathbb{N}$, there is a finite number of balls of radius $\frac{1}{n}$ that cover X. The centers of all these balls form a countable subset A of X (to get a listing of A, first list the centers of the balls of radius 1, then the centers of the balls of radius $\frac{1}{2}$ etc.). We shall prove that A is dense in X.

Let x be an element of X. To find a sequence $\{a_n\}$ from A converging to x, we first pick the center a_1 of (one of) the balls of radius 1 that x belongs to, then we pick the center a_2 of (one of) the balls of radius $\frac{1}{2}$ that x belong to, etc. Since $d(x, a_n) < \frac{1}{n}$, $\{a_n\}$ is a sequence from A converging to x.

To find the set A_δ, just choose $m \in \mathbb{N}$ so big that $\frac{1}{m} < \delta$, and let A_δ consist of the centers of the balls of radius $\frac{1}{m}$. □

We are now ready to turn to $C(X, \mathbb{R}^m)$. First we recall the definition of equicontinuous sets of functions from Section 3.1.

Definition 3.5.4 *Let* (X, d_X) *and* (Y, d_Y) *be metric spaces, and let* $\mathcal{F}$ *be a collection of functions* $f : X \to Y$. *We say that* $\mathcal{F}$ *is* equicontinuous *if for all* $\epsilon > 0$*, there is a* $\delta > 0$ *such that for all* $f \in \mathcal{F}$ *and all* $x, y \in X$ *with* $d_X(x, y) < \delta$*, we have* $d_Y(f(x), f(y)) < \epsilon$.

We begin with a lemma that shows that for equicontinuous sequences, it suffices to check convergence on dense sets of the kind described above.

Lemma 3.5.5 *Assume that* (X, d_X) *is a compact and* (Y, d_Y) *a complete metric space, and let* $\{g_k\}$ *be an equicontinuous sequence in* $C(X, Y)$. *Assume that* $A \subseteq X$ *is a dense set as described in Proposition 3.5.3 and that* $\{g_k(a)\}$ *converges for all* $a \in A$. *Then* $\{g_k\}$ *converges in* $C(X, Y)$.

Proof: Since $C(X, Y)$ is complete, it suffices to prove that $\{g_k\}$ is a Cauchy sequence. Given an $\epsilon > 0$, we must thus find an $N \in \mathbb{N}$ such that $\rho(g_n, g_m) < \epsilon$ when $n, m \geq N$. Since the sequence is equicontinuous, there exists a $\delta > 0$ such that if $d_X(x, y) < \delta$, then $d_Y(g_k(x), g_k(y)) < \frac{\epsilon}{4}$ for all k. Choose a finite subset A_δ of A such that any element in X is within less than δ of an element in A_δ. Since the sequences $\{g_k(a)\}$, $a \in A_\delta$, converge, they are all Cauchy sequences, and we can find an $N \in \mathbb{N}$ such that when $n, m \geq N$, $d_Y(g_n(a), g_m(a)) < \frac{\epsilon}{4}$ for all $a \in A_\delta$ (here we are using that A_δ is finite).

For any $x \in X$, we can find an $a \in A_\delta$ such that $d_X(x, a) < \delta$. But then for all $n, m \geq N$,

$$d_Y(g_n(x), g_m(x)) \leq$$
$$\leq d_Y(g_n(x), g_n(a)) + d_Y(g_n(a), g_m(a)) + d_Y(g_m(a), g_m(x)) <$$
$$< \frac{\epsilon}{4} + \frac{\epsilon}{4} + \frac{\epsilon}{4} = \frac{3\epsilon}{4}$$

Since this holds for any $x \in X$, we must have $\rho(g_n, g_m) \leq \frac{3\epsilon}{4} < \epsilon$ for all $n, m \geq N$, and hence $\{g_k\}$ is a Cauchy sequence and converges in the complete space $C(X, Y)$. □

We are now ready to prove the hard part of the Arzelà-Ascoli Theorem.

Proposition 3.5.6 *Assume that (X, d) is a compact metric space, and let $\{f_n\}$ be a bounded and equicontinuous sequence in $C(X, \mathbb{R}^m)$. Then $\{f_n\}$ has a subsequence converging in $C(X, \mathbb{R}^m)$.*

Proof: Since X is compact, there is a countable, dense subset

$$A = \{a_1, a_2 \ldots, a_n, \ldots\}$$

as in Proposition 3.5.3. According to the lemma, it suffices to find a subsequence $\{g_k\}$ of $\{f_n\}$ such that $\{g_k(a)\}$ converges for all $a \in A$.

We begin a little less ambitiously by showing that $\{f_n\}$ has a subsequence $\{f_n^{(1)}\}$ such that $\{f_n^{(1)}(a_1)\}$ converges (recall that a_1 is the first element in our listing of the countable set A). Next we show that $\{f_n^{(1)}\}$ has a subsequence $\{f_n^{(2)}\}$ such that both $\{f_n^{(2)}(a_1)\}$ and $\{f_n^{(2)}(a_2)\}$ converge. Continuing taking subsequences in this way, we shall for each $j \in \mathbb{N}$ find a sequence $\{f_n^{(j)}\}$ such that $\{f_n^{(j)}(a)\}$ converges for $a = a_1, a_2, \ldots, a_j$. Finally, we shall construct the sequence $\{g_k\}$ by combining all the sequences $\{f_n^{(j)}\}$ in a clever way.

Let us start by constructing $\{f_n^{(1)}\}$. Since the sequence $\{f_n\}$ is bounded, $\{f_n(a_1)\}$ is a bounded sequence in $\mathbb{R}^m$, and by Bolzano-Weierstrass' Theorem, it has a convergent subsequence $\{f_{n_k}(a_1)\}$. We let $\{f_n^{(1)}\}$ consist of the functions appearing in this subsequence. If we now apply $\{f_n^{(1)}\}$ to a_2, we get a new bounded sequence $\{f_n^{(1)}(a_2)\}$ in $\mathbb{R}^m$ with a convergent subsequence. We let $\{f_n^{(2)}\}$ be the functions appearing in this subsequence. Note that $\{f_n^{(2)}(a_1)\}$ still converges as $\{f_n^{(2)}\}$ is a subsequence of $\{f_n^{(1)}\}$. Continuing in this way, we see that we for each $j \in \mathbb{N}$ have a sequence $\{f_n^{(j)}\}$ such that $\{f_n^{(j)}(a)\}$ converges for $a = a_1, a_2, \ldots, a_j$. In addition, each sequence $\{f_n^{(j)}\}$ is a subsequence of the previous ones.

We are now ready to construct a sequence $\{g_k\}$ such that $\{g_k(a)\}$ converges for *all* $a \in A$. We do it by a diagonal argument, putting g_1 equal to the first element in the first sequence $\{f_n^{(1)}\}$, g_2 equal to the second element in the second sequence $\{f_n^{(2)}\}$ etc. In general, the k-th term in the g-sequence equals the k-th term in the k-th f-sequence $\{f_n^k\}$, i.e. $g_k = f_k^{(k)}$. Note that except for the first few elements, $\{g_k\}$ is a subsequence of *any* sequence $\{f_n^{(j)}\}$. This means that $\{g_k(a)\}$ converges for all $a \in A$, and the proof is complete. □

As a simple consequence of this result we get:

Corollary 3.5.7 *If (X, d) is a compact metric space, all bounded, closed and equicontinuous sets $\mathcal{K}$ in $C(X, \mathbb{R}^m)$ are compact.*

Proof: According to the proposition, any sequence in $\mathcal{K}$ has a convergent subsequence. Since $\mathcal{K}$ is closed, the limit must be in $\mathcal{K}$, and hence $\mathcal{K}$ is compact. □

As already mentioned, the converse of this result is also true, but before we prove it, we need a technical lemma that is quite useful also in other situations:

Lemma 3.5.8 *Assume that (X, d_X) and (Y, d_Y) are metric spaces and that $\{f_n\}$ is a sequence of continuous function from X to Y which converges uniformly to f. If $\{x_n\}$ is a sequence in X converging to a, then $\{f_n(x_n)\}$ converges to $f(a)$.*

Remark: This lemma is not as obvious as it may seem — it is not true if we replace uniform convergence by pointwise!

Proof of Lemma 3.5.8: Given $\epsilon > 0$, we must show how to find an $N \in \mathbb{N}$ such that $d_Y(f_n(x_n), f(a)) < \epsilon$ for all $n \geq N$. Since we know from Proposition 3.2.4 that f is continuous, there is a $\delta > 0$ such that $d_Y(f(x), f(a)) < \frac{\epsilon}{2}$ when $d_X(x, a) < \delta$. Since $\{x_n\}$ converges to x, there is an $N_1 \in \mathbb{N}$ such that $d_X(x_n, a) < \delta$ when $n \geq N_1$. Also, since $\{f_n\}$ converges uniformly to f, there is an $N_2 \in \mathbb{N}$ such that if $n \geq N_2$, then $d_Y(f_n(x), f(x)) < \frac{\epsilon}{2}$ for all $x \in X$. If we choose $N = \max\{N_1, N_2\}$, we see that if $n \geq N$,

$$d_Y(f_n(x_n), f(a)) \leq d_Y(f_n(x_n), f(x_n)) + d_Y(f(x_n), f(a)) < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

and the lemma is proved. □

We are finally ready to prove the main theorem:

Theorem 3.5.9 (Arzelà-Ascoli's Theorem) *Let* (X, d_X) *be a compact metric space. A subset* $\mathcal{K}$ *of* $C(X, \mathbb{R}^m)$ *is compact if and only if it is closed, bounded and equicontinuous.*

Proof: It remains to prove that a compact set $\mathcal{K}$ in $C(X, \mathbb{R}^m)$ is closed, bounded and equicontinuous. Since compact sets are always closed and bounded according to Proposition 2.5.4, if suffices to prove that $\mathcal{K}$ is equicontinuous. We argue by contradiction: We assume that the compact set $\mathcal{K}$ is *not* equicontinuous and show that this leads to a contradiction.

Since $\mathcal{K}$ is not equicontinuous, there must be an $\epsilon > 0$ which can not be matched by any δ; i.e. for any $\delta > 0$, there is a function $f \in \mathcal{K}$ and points $x, y \in X$ such that $d_X(x, y) < \delta$, but $d_{\mathbb{R}^m}(f(x), f(y)) \geq \epsilon$. If we put $\delta = \frac{1}{n}$, we get at function $f_n \in \mathcal{K}$ and points $x_n, y_n \in X$ such that $d_X(x_n, y_n) < \frac{1}{n}$, but $d_{\mathbb{R}^m}(f_n(x_n), f_n(y_n)) \geq \epsilon$. Since $\mathcal{K}$ is compact, there is a subsequence $\{f_{n_k}\}$ of $\{f_n\}$ which converges (uniformly) to a function $f \in \mathcal{K}$. Since X is compact, the corresponding subsequence $\{x_{n_k}\}$ of $\{x_n\}$, has a subsequence $\{x_{n_{k_j}}\}$ converging to a point $a \in X$. Since $d_X(x_{n_{k_j}}, y_{n_{k_j}}) < \frac{1}{n_{k_j}}$, the corresponding sequence $\{y_{n_{k_j}}\}$ of y's also converges to a.

Since $\{f_{n_{k_j}}\}$ converges uniformly to f, and $\{x_{n_{k_j}}\}$, $\{y_{n_{k_j}}\}$ both converge to a, the lemma tells us that

$$f_{n_{k_j}}(x_{n_{k_j}}) \to f(a) \quad \text{and} \quad f_{n_{k_j}}(y_{n_{k_j}}) \to f(a)$$

But this is impossible since $d_{\mathbb{R}^m}(f(x_{n_{k_j}}), f(y_{n_{k_j}})) \geq \epsilon$ for all j. Hence we have our contradiction, and the theorem is proved. □

Exercises for Section 3.5

1. Show that $\mathbb{R}^n$ is separable for all n.
2. Show that a subset A of a metric space (X, d) is dense if and only if all open balls $\mathrm{B}(a, r)$, $a \in X$, $r > 0$, contain elements from A.
3. Assume that (X, d) is a complete metric space, and that A is a dense subset of X. We let A have the subset metric d_A.
   a) Assume that $f : A \to \mathbb{R}$ is uniformly continuous. Show that if $\{a_n\}$ is a sequence from A converging to a point $x \in X$, then $\{f(a_n)\}$ converges. Show that the limit is the same for all such sequences $\{a_n\}$ converging to the same point x.
   b) Define $\bar{f} : X \to \mathbb{R}$ by putting $\bar{f}(x) = \lim_{n\to\infty} f(a_n)$ where $\{a_n\}$ is a sequence from a converging to x. We call f the *continuous extension of f to X*. Show that $\bar{f}$ is uniformly continuous.
   c) Let $f : \mathbb{Q} \to \mathbb{R}$ be defined by

$$f(q) = \begin{cases} 0 & \text{if } q < \sqrt{2} \\ \\ 1 & \text{if } q > \sqrt{2} \end{cases}$$

Show that f is continuous on $\mathbb{Q}$ (we are using the usual metric $d_{\mathbb{Q}}(q,r) = |q-r|$). Is f uniformly continuous?

d) Show that f does not have a continuous extension to $\mathbb{R}$.

4. Let K be a compact subset of $\mathbb{R}^n$. Let $\{f_n\}$ be a sequence of contractions of K. Show that $\{f_n\}$ has uniformly convergent subsequence.

5. A function $f : [-1,1] \to \mathbb{R}$ is called *Lipschitz continuous with Lipschitz constant* $K \in \mathbb{R}$ if
$$|f(x) - f(y)| \le K|x-y|$$
for all $x, y \in [-1,1]$. Let $\mathcal{K}$ be the set of all Lipschitz continuous functions with Lipschitz constant K such that $f(0) = 0$. Show that $\mathcal{K}$ is a compact subset of $C([-1,1],\mathbb{R})$.

6. Assume that (X, d_X) and (Y, d_Y) are two metric spaces, and let $\sigma : [0,\infty) \to [0,\infty)$ be a nondecreasing, continuous function such that $\sigma(0) = 0$. We say that σ is a *modulus of continuity* for a function $f : X \to Y$ if
$$d_Y(f(u), f(v)) \le \sigma(d_X(u,v))$$
for all $u, v \in X$.

   a) Show that a family of functions with the same modulus of continuity is equicontinuous.

   b) Assume that (X, d_X) is compact, and let $x_0 \in X$. Show that if σ is a modulus of continuity, then the set
$$\mathcal{K} = \{f : X \to \mathbb{R}^n \ : \ f(x_0) = \mathbf{0} \text{ and } \sigma \text{ is modulus of continuity for } f\}$$
   is compact.

   c) Show that all functions in $C([a,b],\mathbb{R}^m)$ has a modulus of continuity.

7. A metric space (X,d) is called *locally compact* if for each point $a \in X$, there is a *closed* ball $\overline{\mathrm{B}}(a;r)$ centered at a that is compact. (Recall that $\overline{\mathrm{B}}(a;r) = \{x \in X \ : \ d(a,x) \le r\}$). Show that $\mathbb{R}^m$ is locally compact, but that $C([0,1],\mathbb{R})$ is not.

3.6 Differential equations revisited

In Section 3.4, we used Banach's Fixed Point Theorem to study initial value problems of the form
$$\mathbf{y}'(t) = \mathbf{f}(t, \mathbf{y}(t)), \quad \mathbf{y}(0) = \mathbf{y}_0 \tag{3.6.1}$$
or equivalently
$$\mathbf{y}(t) = \mathbf{y}_0 + \int_0^t \mathbf{f}(s, \mathbf{y}(s))\, ds \tag{3.6.2}$$
In this section we shall see how Arzelà-Ascoli's Theorem can be used to prove existence of solutions under weaker conditions than before. But in the new

approach we shall also lose something — we can only prove that the solutions exist in small intervals, and we can no longer guarantee uniqueness.

The starting point is Euler's method for finding approximate solutions to differential equations. If we want to approximate the solution starting at $\mathbf{y}_0$ at time $t = 0$, we begin by partitioning time into discrete steps of length Δt; hence we work with the time line

$$T = \{t_0, t_1, t_2, t_3 \ldots\}$$

where $t_0 = 0$ and $t_{i+1} - t_i = \Delta t$. We start the approximate solution $\hat{\mathbf{y}}$ at $\mathbf{y}_0$ and move in the direction of the derivative $\mathbf{f}(t_0, \mathbf{y}_0)$, i.e. we put

$$\hat{\mathbf{y}}(t) = \mathbf{y}_0 + \mathbf{f}(t_0, \mathbf{y}_0)(t - t_0)$$

for $t \in [t_0, t_1]$. Once we reach t_1, we change directions and move in the direction of the new derivative $\mathbf{f}(t_1, \hat{\mathbf{y}}(t_1))$ so that we have

$$\hat{\mathbf{y}}(t) = \hat{\mathbf{y}}(t_1) + \mathbf{f}(t_0, \hat{\mathbf{y}}(t_1))(t - t_1)$$

for $t \in [t_1, t_2]$. If we insert the expression for $\hat{\mathbf{y}}(t_1)$, we get:

$$\hat{\mathbf{y}}(t) = \mathbf{y}_0 + \mathbf{f}(t_0, \mathbf{y}_0)(t_1 - t_0) + \mathbf{f}(t_1, \hat{\mathbf{y}}(t_1))(t - t_1)$$

If we continue in this way, changing directions at each point in T, we get

$$\hat{\mathbf{y}}(t) = \mathbf{y}_0 + \sum_{i=0}^{k-1} \mathbf{f}(t_i, \hat{\mathbf{y}}(t_i))(t_{i+1} - t_i) + \mathbf{f}(t_k, \hat{\mathbf{y}}(t_k))(t - t_k)$$

for $t \in [t_k, t_{k+1}]$. If we observe that

$$\mathbf{f}(t_i, \hat{\mathbf{y}}(t_i))(t_{i+1} - t_i) = \int_{t_i}^{t_{i+1}} \mathbf{f}(t_i, \hat{\mathbf{y}}(t_i)\, ds\,,$$

we can rewrite this expression as

$$\hat{\mathbf{y}}(t) = \mathbf{y}_0 + \sum_{i=0}^{k-1} \int_{t_i}^{t_{i+1}} \mathbf{f}(t_i, \hat{\mathbf{y}}(t_i)\, ds + \int_{t_k}^{t} \mathbf{f}(t_k, \hat{\mathbf{y}}(t_k)\, ds$$

If we also introduce the notation

$$\underline{s} = \text{the largest } t_i \in T \text{ such that } t_i \leq s,$$

we may express this more compactly as

$$\hat{\mathbf{y}}(t) = \mathbf{y}_0 + \int_0^t \mathbf{f}(\underline{s}, \hat{\mathbf{y}}(\underline{s}))\, ds$$

Note that we can also write this as

$$\hat{\mathbf{y}}(t) = \mathbf{y}_0 + \int_0^t \mathbf{f}(s, \hat{\mathbf{y}}(s))\, ds + \int_0^t \big(\mathbf{f}(\underline{s}, \hat{\mathbf{y}}(\underline{s})) - \mathbf{f}(s, \hat{\mathbf{y}}(s))\big)\, ds$$

(observe that there is one s and one $\underline{s}$ term in the last integral) where the last term measures how much $\hat{\mathbf{y}}$ "deviates" from being a solution of equation (3.6.2).

Intuitively, one would think that the approximate solution $\hat{\mathbf{y}}$ will converge to a real solution $\mathbf{y}$ when the step size Δt goes to zero. To be more specific, if we let $\hat{\mathbf{y}}_n$ be the approximate solution we get when we choose $\Delta t = \frac{1}{n}$, we would expext the squence $\{\hat{\mathbf{y}}_n\}$ to converge to a solution of (2). It turns out that in the most general case we can not quite prove this, but we can instead use the Arzelà-Ascoli Theorem to find a *subsequence* converging to a solution.

Before we turn to the proof, it will useful to see how intergals of the form

$$I_k(t) = \int_0^t \mathbf{f}(s, \hat{\mathbf{y}}_k(s))\, ds$$

behave when the functions $\hat{\mathbf{y}}_k$ converge uniformly to a limit $\mathbf{y}$.

Lemma 3.6.1 *Let $\mathbf{f} : [0,\infty) \times \mathbb{R}^m \to \mathbb{R}^m$ be a continuous function, and assume that $\{\hat{\mathbf{y}}_k\}$ is a sequence of continuous functions $\hat{\mathbf{y}}_k : [0,a] \to \mathbb{R}^m$ converging uniformly to a function $\mathbf{y}$. Then the integral functions*

$$I_k(t) = \int_0^t \mathbf{f}(s, \hat{\mathbf{y}}_k(s))\, ds$$

converge uniformly to

$$I(t) = \int_0^t \mathbf{f}(s, \mathbf{y}(s))\, ds$$

on $[0,a]$.

Proof: Since the sequence $\{\hat{\mathbf{y}}_k\}$ converges uniformly, it is bounded, and hence there is a constant K such that $|\hat{\mathbf{y}}_k(t)| \leq K$ for all $k \in \mathbb{N}$ and all $t \in [0,a]$ (prove this!). The continuous function $\mathbf{f}$ is uniformly continuous on the compact set $[0,a] \times [-K,K]^m$, and hence for every $\epsilon > 0$, there is a $\delta > 0$ such that if $|\mathbf{y} - \mathbf{y}'| < \delta$, then $|\mathbf{f}(s,\mathbf{y}) - \mathbf{f}(s,\mathbf{y}')| < \frac{\epsilon}{a}$ for all $s \in [0,a]$. Since $\{\hat{\mathbf{y}}_k\}$ converges uniformly to $\mathbf{y}$, there is an $N \in \mathbb{N}$ such that if $n \geq N$, $|\hat{\mathbf{y}}_n(s) - \mathbf{y}(s)| < \delta$ for all $s \in [0,a]$. But then

$$|I_n(t) - I(t)| = |\int_0^t \big(\mathbf{f}(s, \hat{\mathbf{y}}_n(s)) - \mathbf{f}(s, \mathbf{y}(s))\big)\, ds| \leq$$

$$\leq \int_0^t \big|\mathbf{f}(s, \hat{\mathbf{y}}_n(s)) - \mathbf{f}(s, \mathbf{y}(s))\big|\, ds < \int_0^a \frac{\epsilon}{a}\, ds = \epsilon$$

for all $t \in [0, a]$, and hence $\{I_k\}$ converges uniformly to I. □

We are now ready for the main result.

Theorem 3.6.2 *Assume that* $\mathbf{f} : [0, \infty) \times \mathbb{R}^m \to \mathbb{R}^m$ *is a continuous function and that* $\mathbf{y}_0 \in \mathbb{R}^m$. *Then there exists a positive real number* a *and a function* $\mathbf{y} : [0, a] \to \mathbb{R}^m$ *such that* $\mathbf{y}(0) = \mathbf{y}_0$ *and*

$$y'(t) = \mathbf{f}(t, \mathbf{y}(t)) \quad \textit{for all } t \in [0, a]$$

Remark: Note that there is no uniqueness statement (the problem may have more than one solution), and that the solution is only guaranteed to exist on a bounded intervall (it may disappear to infinity after finite time).

Proof of Theorem 3.6.2: Choose a big, compact subset $C = [0, R] \times [-R, R]^m$ of $[0, \infty) \times \mathbb{R}^m$ containing $(0, \mathbf{y}_0)$ in its interior. By the Extreme Value Theorem, the components of $\mathbf{f}$ have a maximum value on C, and hence there exists a number $M \in \mathbb{R}$ such that $|f_i(t, \mathbf{y})| \leq M$ for all $(t, \mathbf{y}) \in C$ and all $i = 1, 2, \ldots, m$. If the initial value has components

$$\mathbf{y}_0 = \begin{pmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_m \end{pmatrix}$$

we choose $a \in \mathbb{R}$ so small that the set

$$A = [0, a] \times [Y_1 - Ma, Y_1 + Ma] \times [Y_2 - Ma, Y_2 + Ma] \times \cdots \times [Y_m - Ma, Y_m + ma]$$

is contained in C. This may seem mysterious, put the point is that our approximate solutions of the differential equation can never leave the area

$$[Y_1 - Ma, Y_1 + Ma] \times [Y_2 - Ma, Y_2 + Ma] \times \cdots \times [Y_m - Ma, Y + ma]$$

while $t \in [0, a]$ since all the derivatives are bounded by M.

Let $\hat{\mathbf{y}}_n$ be the approximate solution obtained by using Euler's method on the interval $[0, a]$ with time step $\frac{a}{n}$. The sequence $\{\hat{\mathbf{y}}_n\}$ is bounded since $(t, \hat{\mathbf{y}}_n(t)) \in A$, and it is equicontinuous since the components of $\mathbf{f}$ are bounded by M. By Proposition 3.5.6, $\hat{\mathbf{y}}_n$ has a subsequence $\{\hat{\mathbf{y}}_{n_k}\}$ converging uniformly to a function $\mathbf{y}$. If we can prove that $\mathbf{y}$ solves the integral equation

$$\mathbf{y}(t) = \mathbf{y}_0 + \int_0^t \mathbf{f}(s, \mathbf{y}(s))\, ds$$

for all $t \in [0, a]$, we shall have proved the theorem.

From the calculations at the beginning of the section, we know that

$$\hat{\mathbf{y}}_{n_k}(t) = \mathbf{y}_0 + \int_0^t \mathbf{f}(s, \hat{\mathbf{y}}_{n_k}(s))\, ds + \int_0^t \big(\mathbf{f}(\underline{s}, \hat{\mathbf{y}}_{n_k}(\underline{s})) - \mathbf{f}(s, \hat{\mathbf{y}}_{n_k}(s))\big)\, ds \quad (3.6.3)$$

and according to the lemma

$$\int_0^t \mathbf{f}(s, \hat{\mathbf{y}}_{n_k}(s))\, ds \to \int_0^t \mathbf{f}(s, \mathbf{y}(s))\, ds \quad \text{uniformly for } t \in [0, a]$$

If we can only prove that

$$\int_0^t \big(\mathbf{f}(\underline{s}, \hat{\mathbf{y}}_{n_k}(\underline{s})) - \mathbf{f}(s, \hat{\mathbf{y}}_{n_k}(s))\big)\, ds \to 0 \quad (3.6.4)$$

we will get

$$\mathbf{y}(t) = \mathbf{y}_0 + \int_0^t \mathbf{f}(s, \mathbf{y}(s))\, ds$$

as $k \to \infty$ in (3.6.3), and the theorem will be proved

To prove (3.6.4), observe that since A is a compact set, $\mathbf{f}$ is uniformly continuous on A. Given an $\epsilon > 0$, we thus find a $\delta > 0$ such that $|\mathbf{f}(s, \mathbf{y}) - \mathbf{f}(s', \mathbf{y}')| < \frac{\epsilon}{a}$ when $|(s, \mathbf{y}) - (s', \mathbf{y})| < \delta$ (we are measuring the distance in the ordinary $\mathbb{R}^{m+1}$-metric). Since

$$|(\underline{s}, \hat{\mathbf{y}}_{n_k}(\underline{s})) - (s, \hat{\mathbf{y}}_{n_k}(s))| \leq |(\Delta t, M\Delta t, \dots, M\Delta t)| = \sqrt{1 + nM^2}\, \Delta t\,,$$

we can clearly get $|(\underline{s}, \hat{\mathbf{y}}_{n_k}(\underline{s})) - (s, \hat{\mathbf{y}}_{n_k}(s))| < \delta$ by choosing k large enough (and hence Δt small enough). For such k we then have

$$|\int_0^t \big(\mathbf{f}(\underline{s}, \hat{\mathbf{y}}_{n_k}(\underline{s})) - \mathbf{f}(s, \hat{\mathbf{y}}_{n_k}(s))\big| < \int_0^a \frac{\epsilon}{a}\, ds = \epsilon$$

and hence

$$\int_0^t \big(\mathbf{f}(\underline{s}, \hat{\mathbf{y}}_{n_k}(\underline{s})) - \mathbf{f}(s, \hat{\mathbf{y}}_{n_k}(s))\big)\, ds \to 0$$

as $k \to \infty$. As already observed, this completes the proof. □

Remark: An obvious question at this stage is why didn't we extend our solution beyond the interval $[0, a]$ as we did in the proof of Theorem 3.4.2? The reason is that in the present case we do not have control over the length of our intervals, and hence the second interval may be very small compared to the first one, the third one even smaller, and so one. Even if we add an infinite number of intervals, we may still only cover a finite part of the real line. There are good reasons for this: the differential equation may only have solutions that survive for a finite amount of time. A typical example is the equation

$$y' = (1 + y^2), \quad y(0) = 0$$

where the (unique) solution $y(t) = \tan t$ goes to infinity when $t \to \frac{\pi}{2}^-$.

The proof above is a simple, but typical example of a wide class of compactness arguments in the theory of differential equations. In such arguments one usually starts with a sequence of approximate solutions and then uses compactness to extract a subsequence converging to a solution. Compactness methods are strong in the sense that they can often prove local existence of solutions under very general conditions, but they are weak in the sense that they give very little information about the nature of the solution. But just knowing that a solution exists, is often a good starting point for further explorations.

Exercises for Section 3.6

1. Prove that if $\mathbf{f}_n : [a,b] \to \mathbb{R}^m$ are continuous functions converging uniformly to a function $\mathbf{f}$, then the sequence $\{\mathbf{f}_n\}$ is bounded in the sense that there is a constant $K \in \mathbb{R}$ such that $|\mathbf{f}_n(t)| \leq K$ for all $n \in \mathbb{N}$ and all $t \in [a,b]$ (this property is used in the proof of Lemma 3.6.1).
2. Go back to exercises 1 and 2 in Section 3.4. Show that the differential equations satisfy the conditions of Theorem 3.6.2. Comment.
3. It is occasionally useful to have a slightly more general version of Theorem 3.6.2 where the solution doesn't just start a given point, but passes through it:

   Teorem *Assume that $\mathbf{f} : \mathbb{R} \times \mathbb{R}^m \to \mathbb{R}^m$ is a continuous function. For any $t_0 \in \mathbb{R}$ and $\mathbf{y}_0 \in \mathbb{R}^m$, there exists a positive real number a and a function $\mathbf{y} : [t_0 - a, t_0 + a] \to \mathbb{R}^m$ such that $\mathbf{y}(t_0) = \mathbf{y}_0$ and*

   $$y'(t) = \mathbf{f}(t, \mathbf{y}(t)) \quad \text{for all } t \in [t_0 - a, t_0 + a]$$

   Prove this theorem by modifying the proof of Theorem 3.6.2 (run Euler's method "backwards" on the interval $[t_0 - a, t_0]$).

3.7 Polynomials are dense in $C([a,b],\mathbb{R})$

From calculus we know that many continuous functions can be approximated by their Taylor polynomials, but to have Taylor polynomials of all orders, a function f has to be infinitely differentiable, i.e. the higher order derivatives $f^{(k)}$ have to exist for all k. Most continuous functions are not differentiable at all, and the question is whether they still can be approximated by polynomials. In this section we shall prove:

Theorem 3.7.1 (Weierstrass' Theorem) *The polynomials are dense in $C([a,b],\mathbb{R})$ for all $a, b \in \mathbb{R}$, $a < b$. In other words, for each continuous function $f : [a,b] \to \mathbb{R}$, there is a sequence of polynomials $\{p_n\}$ converging uniformly to f.*

The proof I shall give (due to the Russian mathematician Sergei Bernstein (1880-1968)) is quite surprising; it uses probability theory to establish the result for the interval $[0, 1]$, and then a straight forward scaling argument to extend it to all closed and bounded intervals.

The idea is simple: Assume that you are tossing a biased coin which has probability x of coming up "heads". If you toss it more and more times, you expect the proportion of times it comes up "heads" to stabilize around x. If somebody has promised you an award of $f(X)$ dollars, where X is the actually proportion of "heads" you have had during your (say) 1000 first tosses, you would expect your award to be close to $f(x)$. If the number of tosses was increased to 10 000, you would feel even more certain.

Let us fomalize this: Let Y_i be the outcome of the i-th toss in the sense that Y_i has the value 0 if the coin comes up "tails" and 1 if it comes up "heads". The proportion of "heads" in the first N tosses is then given by

$$X_N = \frac{1}{N}(Y_1 + Y_2 + \cdots + Y_N)$$

Each Y_i is binomially distributed with mean $\mathrm{E}(Y_i) = x$ and variance $\mathrm{Var}(Y_i) = x(1-x)$. We thus have

$$\mathrm{E}(X_N) = \frac{1}{N}(\mathrm{E}(Y_1) + \mathrm{E}(Y_2) + \cdots \mathrm{E}(Y_N)) = x$$

and (using that the Y_i's are independent)

$$\mathrm{Var}(X_N) = \frac{1}{N^2}(\mathrm{Var}(Y_1) + \mathrm{Var}(Y_2) + \cdots + \mathrm{Var}(Y_N)) = \frac{1}{N}x(1-x)$$

(if you don't remember these formulas from probability theory, we shall derive them by analytic methods in the exercises). As N goes to infinity, we would expect X_N to converge to x with probability 1. If the "award function" f is continuous, we would also expect our average award $\mathrm{E}(f(X_N))$ to converge to $f(x)$.

To see what this has to do with polynomials, let us compute the average award $\mathrm{E}(f(X_N))$. Since the probability of getting exactly k heads in N tosses is $\binom{N}{k}x^k(1-x)^{n-k}$, we get

$$\mathrm{E}(f(X_N)) = \sum_{k=0}^{N} f(\frac{k}{N})\binom{N}{k}x^k(1-x)^{N-k}$$

Our expectation that $\mathrm{E}(f(X_N)) \to f(x)$ as $N \to \infty$, can therefore be rephrased as

$$\sum_{k=0}^{N} f(\frac{k}{N})\binom{N}{k}x^k(1-x)^{N-k} \to f(x) \quad N \to \infty$$

If we expand the parentheses $(1-x)^{N-k}$, we see that the expressions on the right hand side are just polynomials in x, and hence we have arrived at the hypothesis that the polynomials

$$p_N(x) = \sum_{k=0}^{N} f(\frac{k}{N})\binom{N}{k}x^k(1-x)^{N-k}$$

converge to $f(x)$. We shall prove that this is indeed the case, and that the convergence is uniform.

Before we turn to the proof, we need some notation and a lemma. For any random variable X with expectation x and any $\delta > 0$, we shall write

$$\mathbf{1}_{\{|x-X|<\delta\}} = \begin{cases} 1 & \text{if } |x-X| < \delta \\ \\ 0 & \text{otherwise} \end{cases}$$

and oppositely for $\mathbf{1}_{\{|x-X_|\geq\delta\}}$.

Lemma 3.7.2 (Chebyshev's Inequality) *For a bounded random variable X with mean x*

$$\mathrm{E}(\mathbf{1}_{\{|x-X|\geq\delta\}}) \leq \frac{1}{\delta^2}\mathrm{Var}(X)$$

Proof: Since $\delta^2\mathbf{1}_{\{|x-X|\geq\delta\}}) \leq (x-X)^2$, we have

$$\delta^2\mathrm{E}(\mathbf{1}_{\{|x-X|\geq\delta\}}) \leq \mathrm{E}((x-X)^2) = \mathrm{Var}(X)$$

Dividing by δ^2, we get the lemma. □

We are now ready to prove that the Bernstein polynomials converge.

Proposition 3.7.3 *If $f : [0,1] \to \mathbb{R}$ is a continuous function, the* Bernstein polynomials

$$p_N(x) = \sum_{k=0}^{N} f(\frac{k}{N})\binom{N}{k}x^k(1-x)^{N-k}$$

converge uniformly to f on $[0,1]$.

Proof: Given $\epsilon > 0$, we must show how to find an N such that $|f(x) - p_n(x)| < \epsilon$ for all $n \geq N$ and all $x \in [0,1]$. Since f is continuous on the compact set $[0,1]$, it has to be uniformly continuous, and hence we can find a $\delta > 0$ such that $|f(u) - f(v)| < \frac{\epsilon}{2}$ whenever $|u - v| < \delta$. Since $p_n(x) = \mathrm{E}(f(X_n))$, we have

$$|f(x)-p_n(x)| = |f(x)-\mathrm{E}(f(X_n))| = |\mathrm{E}(f(x)-f(X_n))| \leq \mathrm{E}(|f(x)-f(X_n)|)$$

We split the last expectation into two parts: the cases where $|x - X_n| < \delta$ and the rest:

$$\mathrm{E}(|f(x)-f(X_n)|) = \mathrm{E}(\mathbf{1}_{\{|x-X_n|<\delta\}}|f(x)-f(X_n)|)+\mathrm{E}(\mathbf{1}_{\{|x-X_n|\geq\delta\}}|f(x)-f(X_n)|)$$

The idea is that the first term is always small due to the choice of δ and that the second part will be small when N is large because X_N then is unlikely to deviate much from x. Here are the details:

By choice of δ, we have for the first term

$$\mathrm{E}(\mathbf{1}_{\{|x-X_n|<\delta\}}|f(x) - f(X_n)|) \leq \mathrm{E}\left(\mathbf{1}_{\{|x-X_n|<\delta\}}\frac{\epsilon}{2}\right) \leq \frac{\epsilon}{2}$$

For the second term, we first note that since f is a continuous function on a compact interval, it must be bounded by a constant M. Hence by Chebyshev's inequality

$$\mathrm{E}(\mathbf{1}_{\{|x-X_n|\geq\delta\}}|f(x) - f(X_n)|) \leq 2M\mathrm{E}(\mathbf{1}_{\{|x-X_n|\geq\delta\}}) \leq$$

$$\leq \frac{2M}{\delta^2}\mathrm{Var}(X_n) = \frac{2Mx(1-x)}{\delta^2 n} \leq \frac{M}{2\delta^2 n}$$

where we in the last step used that $\frac{1}{4}$ is the maximal value of $x(1-x)$ on $[0, 1]$. If we now choose $N \geq \frac{M}{\delta^2\epsilon}$, we see that we get

$$\mathrm{E}(\mathbf{1}_{\{|x-X_n|\geq\delta\}}|f(x) - f(X_n)|) < \frac{\epsilon}{2}$$

for all $n \geq N$. Combining all the inequalities above, we see that if $n \geq N$, we have for all $x \in [0, 1]$

$$|f(x) - p_n(x)| \leq \mathrm{E}(|f(x) - f(X_n)|) =$$

$$= \mathrm{E}(\mathbf{1}_{\{|x-X_n|<\delta\}}|f(x) - f(X_n)|) + \mathrm{E}(\mathbf{1}_{\{|x-X_n|\geq\delta\}}|f(x) - f(X_n)|) <$$

$$< \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

and hence the Bernstein polynomials p_n converge uniformly to f. □

To get Weierstrass' result, we just have to move functions from arbitrary intervals $[a, b]$ to $[0, 1]$ and back. The function

$$T(x) = \frac{x-a}{b-a}$$

maps $[a, b]$ bijectively to $[0, 1]$, and the inverse function

$$T^{-1}(y) = a + (b-a)y$$

maps $[0, 1]$ back to $[a, b]$. If f is a continuous function on $[a, b]$, the function $\hat{f} = f \circ T^{-1}$ is a continuous function on $[0, 1]$ taking exactly the same values in the same order. If $\{q_n\}$ is a sequence of pynomials converging uniformly to $\hat{f}$ on $[0,1]$, then the functions $p_n = q_n \circ T$ converge uniformly to f on $[a,b]$. Since

$$p_n(x) = q_n(\frac{x-a}{b-a})$$

the p_n's are polynomials, and hence Weierstrass' theorem is proved.

Remark: Weierstrass' theorem is important because many mathematical arguments are easier to perform on polynomials than on continuous functions in general. If the property we study is preserved under uniform limits (i.e. if the if the limit function f of a uniformly convergent sequence of functions $\{f_n\}$ always inherits the property from the f_n's), we can use Weierstrass' Theorem to extend the argument from polynomials to all continuous functions. There is an extension of the result called the Stone-Weierstrass Theorem which generalizes the result to much more general settings.

Exercises for Section 3.7

1. Show that there is no sequence of polynomials that converges uniformly to the continuous function $f(x) = \frac{1}{x}$ on $(0,1)$.
2. Show that there is no sequence of poynomials that converges uniformly to the function $f(x) = e^x$ on $\mathbb{R}$.
3. In this problem

$$f(x) = \begin{cases} e^{-1/x^2} & \text{if } x \neq 0 \\ \\ 0 & \text{if } x = 0 \end{cases}$$

   a) Show that if $x \neq 0$, then the n-th derivative has the form

$$f^{(n)}(x) = e^{-1/x^2}\frac{P_n(x)}{x^{N_n}}$$

   where P_n is a polynomial and $N_n \in \mathbb{N}$.

   b) Show that $f^{(n)}(0) = 0$ for all n.

   c) Show that the Taylor polynomials of f at 0 do not converge to f except in the point 0.

4. Assume that $f : [a,b] \to \mathbb{R}$ is a continuous function such that $\int_a^b f(x)x^n\,dx = 0$ for all $n = 0, 1, 2, 3, \ldots$.

   a) Show that $\int_a^b f(x)p(x)\,dx = 0$ for all polynomials p.

   b) Use Weierstrass' theorem to show that $\int_a^b f(x)^2\,dx = 0$. Conclude that $f(x) = 0$ for all $x \in [a,b]$.

5. In this exercise we shall show that $C([a,b],\mathbb{R})$ is a separable metric space.

   a) Assume that (X,d) is a metric space, and that $S \subseteq T$ are subsets of X. Show that if S is dense in (T, d_T) and T is dense in (X,d), then S is dense in (X,d).

b) Show that for any polynomial p, there is a sequence $\{q_n\}$ of polynomials with rational coefficients that converges uniformly to p on $[a,b]$.

c) Show that the polynomials with rational coefficients are dense in $C([a,b],\mathbb{R})$.

d) Show that $C([a,b],\mathbb{R})$ is separable.

6. In this problem we shall reformulate Bernstein's proof in purely analytic terms, avoiding concepts and notation from probability theory. You should keep the Binomial Formula

$$(a+b)^N = \sum_{k=0}^{N} \binom{n}{k} a^k b^{N-k}$$

and the definition $\binom{N}{k} = \frac{N(N-1)(N-2)\cdot\ldots\cdot(N-k+1)}{1\cdot 2\cdot 3\cdot\ldots\cdot k}$ in mind.

a) Show that $\sum_{k=0}^{N} \binom{N}{k} x^k (1-x)^{N-k} = 1$.

b) Show that $\sum_{k=0}^{N} \frac{k}{N}\binom{N}{k} x^k (1-x)^{N-k} = x$ (this is the analytic version of $\mathrm{E}(X_N) = \frac{1}{N}(\mathrm{E}(Y_1) + \mathrm{E}(Y_2) + \cdots \mathrm{E}(Y_N)) = x$)

c) Show that $\sum_{k=0}^{N} \left(\frac{k}{N} - x\right)^2 \binom{N}{k} x^k (1-x)^{N-k} = \frac{1}{N}x(1-x)$ (this is the analytic version of $\mathrm{Var}(X_N) = \frac{1}{N}x(1-x)$). *Hint:* Write $(\frac{k}{N} - x)^2 = \frac{1}{N^2}\left(k(k-1) + (1-2xN)k + N^2x^2\right)$ and use points b) and a) on the second and third term in the sum.

d) Show that if p_n is the n-th Bernstein polynomial, then

$$|f(x) - p_n(x)| \leq \sum_{k=0}^{n} |f(x) - f(k/n)| \binom{n}{k} x^n (1-x)^{n-k}$$

e) Given $\epsilon > 0$, explain why there is a $\delta > 0$ such that $|f(u) - f(v)| < \epsilon/2$ for all $u, v \in [0,1]$ such that $|u - v| < \delta$. Explain why

$$|f(x) - p_n(x)| \leq \sum_{\{k: |\frac{k}{n} - x| < \delta\}} |f(x) - f(k/n)| \binom{n}{k} x^n (1-x)^{n-k} +$$

$$+ \sum_{\{k: |\frac{k}{n} - x| \geq \delta\}} |f(x) - f(k/n)| \binom{n}{k} x^n (1-x)^{n-k} \leq$$

$$< \frac{\epsilon}{2} + \sum_{\{k: |\frac{k}{n} - x| \geq \delta\}} |f(x) - f(k/n)| \binom{n}{k} x^n (1-x)^{n-k}$$

f) Show that there is a constant M such that $|f(x)| \leq M$ for all $x \in [0,1]$. Explain all the steps in the calculation:

$$\sum_{\{k: |\frac{k}{n} - x| \geq \delta\}} |f(x) - f(k/n)| \binom{n}{k} x^n (1-x)^{n-k} \leq$$

$$\leq 2M \sum_{\{k: |\frac{k}{n} - x| \geq \delta\}} \binom{n}{k} x^n (1-x)^{n-k} \leq$$

$$\leq 2M \sum_{k=0}^{n} \left(\frac{\frac{k}{n} - x}{\delta}\right)^2 \binom{n}{k} x^n (1-x)^{n-k} \leq \frac{2M}{n\delta^2} x(1-x) \leq \frac{M}{2n\delta^2}$$

g) Explain why we can get $|f(x) - p_n(x)| < \epsilon$ by chosing n large enough, and explain why this proves Proposition 3.7.2.

3.8 Baire's Category Theorem

Recall that a subset A of a metric space (X, d) is *dense* if for all $x \in X$ there is a sequence from A converging to x. An equivalent definition is that all balls in X contain elements from A. To show that a set S is *not* dense, we thus have to find an open ball that does not intersect S. Obviously, a set can fail to be dense in parts of X, and still be dense in other parts. If G is a nonempty, open subset of X, we say that A is *dense in G* if every ball $\mathrm{B}(x; r) \subseteq G$ contains elements from A. The following definition catches our intiution of a set set that is not dense anywhere.

Definition 3.8.1 *A subset S of a metric space (X, d) is said to be* nowhere dense *if it isn't dense in* any *nonempty, open set G. In other words, for all nonempty, open sets $G \subseteq X$, there is a ball $\mathrm{B}(x; r) \subseteq G$ that does not intersect S.*

This definition simply says that no matter how much we restrict our attention, we shall never find an area in X where S is dense.

Example 1. $\mathbb{N}$ is nowhere dense in $\mathbb{R}$. ♣

Nowhere dense sets are sparse in an obvious way. The following definition indicates that even countable unions of nowhere dense sets are unlikely to be very large.

Definition 3.8.2 *A set is called* meager *if it is a countable union of nowhere dense sets. The complement of a meager set is called* comeager.[1]

Example 2. $\mathbb{Q}$ is a meager set in $\mathbb{R}$ as it can be written as a countable union $\mathbb{Q} = \bigcup_{a \in \mathbb{Q}} \{a\}$ of the nowhere dense singletons $\{a\}$. By the same argument, $\mathbb{Q}$ is also meager in $\mathbb{Q}$.

The last part of the example shows that a meager set can fill up a metric space. However, in *complete* spaces the meager sets are always "meager" in the following sense:

[1]Most books refer to meager sets as "sets of first category" while comeager sets are called "residual sets". Sets that are not of first category, are said to be of "second category". Although this is the original terminology of René-Louis Baire (1874-1932) who introduced the concepts, it is in my opinion so nondescriptive that it should be abandoned in favor of more evocative terms.

Theorem 3.8.3 (Baire's Category Theorem) *Assume that M is a meager subset of a complete metric space (X, d). Then M does not contain any open balls, i.e. M^c is dense in X.*

Proof: Since M is meager, it can be written as a union $M = \bigcup_{k\in\mathbb{N}} N_k$ of nowhere dense sets N_k. Given a ball $\mathrm{B}(a; r)$, our task is to find an element $x \in \mathrm{B}(a; r)$ which does not belong to M.

We first observe that since N_1 is nowhere dense, there is a ball $\mathrm{B}(a_1; r_1)$ inside $\mathrm{B}(a; r)$ which does not intersect N_1. By shrinking the radius r_1 slightly if necessary, we may assume that the *closed* ball $\overline{\mathrm{B}}(a_1; r_1)$ is contained in $\mathrm{B}(a; r)$, does not intersect N_1, and has radius less than 1. Since N_2 is nowhere dense, there is a ball $\mathrm{B}(a_2; r_2)$ inside $\mathrm{B}(a_1; r_1)$ which does not intersect N_2. By shrinking the radius r_2 if necessary, we may assume that the closed ball $\overline{\mathrm{B}}(a_2; r_2)$ does not intersect N_2 and has radius less than $\frac{1}{2}$. Continuing in this way, we get a sequence $\{\overline{\mathrm{B}}(a_k; r_k)\}$ of closed balls, each contained in the previous, such that $\overline{\mathrm{B}}(a_k; r_k)$ has radius less than $\frac{1}{k}$ and does not intersect N_k.

Since the balls are nested and the radii shrink to zero, the centers a_k form a Cauchy sequence. Since X is complete, the sequence converges to a point x. Since each ball $\overline{\mathrm{B}}(a_k; r_k)$ is closed, and the "tail" $\{a_n\}_{n=k}^{\infty}$ of the sequence belongs to $\overline{\mathrm{B}}(a_k; r_k)$, the limit x also belongs to $\overline{\mathrm{B}}(a_k; r_k)$. This means that for all k, $x \notin N_k$, and hence $x \notin M$. Since $\overline{\mathrm{B}}(a_1; r_1) \subseteq \mathrm{B}(a; r)$, we see that $x \in \mathrm{B}(a; r)$, and the theorem is proved. □

As an immediate consequence we have:

Corollary 3.8.4 *A complete metric space is not a countable union of nowhere dense sets.*

Baire's Category Theorem is a surprisingly strong tool for proving theorems about sets and families of functions. Before we take a look at some examples, we shall prove the following lemma which gives a simpler description of *closed*, nowhere dense sets.

Lemma 3.8.5 *A closed set F is nowhere dense if and only if it does not contain any open balls.*

Proof: If F contains an open ball, it obviously isn't nowhere dense. We therefore assume that F does *not* contain an open ball, and prove that it is nowhere dense. Given a nonempty, open set G, we know that F cannot contain all of G as G contains open balls and F does not. Pick an element x in G that is not in F. Since F is closed, there is a ball $\mathrm{B}(x; r_1)$ around x that does not intersect F. Since G is open, there is a ball $\mathrm{B}(x; r_2)$ around x that is contained in G. If we choose $r = \min\{r_1, r_2\}$, the ball $\mathrm{B}(x; r)$ is contained in G and does not intersect F, and hence F is nowhere dense. □

Remark: Without the assumption that F is closed, the lemma is false, but it is still possible to prove a related result: A (general) set S is nowhere dense if and only if its closure $\bar{S}$ doesn't contain any open balls. See Exercise 5.

We are now ready to take a look at our first application.

Definition 3.8.6 *Let (X, d) be a metric space. A family $\mathcal{F}$ of functions $f : X \to \mathbb{R}$ is called* pointwise bounded *if for each $x \in X$, there is a constant $M_x \in \mathbb{R}$ such that $|f(x)| \leq M_x$ for all $f \in \mathcal{F}$.*

Note that the constant M_x may vary from point to point, and that there need not be a constant M such that $|f(x)| \leq M$ for all f and all x (a simple example is $\mathcal{F} = \{f : \mathbb{R} \to \mathbb{R} \,|\, f(x) = kx$ for $k \in [-1, 1\}$, where $M_x = |x|$). The next result shows that although we cannot guarantee boundedness on all X, we can under reasonable assumptions guarantee boundedness on a part of X.

Proposition 3.8.7 *Let (X, d) be a complete metric space, and assume that $\mathcal{F}$ is a pointwise bounded family of continuous functions $f : X \to \mathbb{R}$. Then there exists an open, nonempty set G and a constant $M \in \mathbb{R}$ such that $|f(x)| \leq M$ for all $f \in \mathcal{F}$ and all $x \in G$.*

Proof: For each $n \in \mathbb{N}$ and $f \in \mathcal{F}$, the set $f^{-1}([-n, n])$ is closed as it is the inverse image of a closed set under a continuous function (recall Proposition 2.3.10). As intersections of closed sets are closed (Proposition 2.3.12)

$$A_n = \bigcap_{f \in \mathcal{F}} f^{-1}([-n, n])$$

is also closed. Since $\mathcal{F}$ is pointwise bounded, $X = \bigcup_{n \in \mathbb{N}} A_n$, and Corollary 3.8.4 tells us that not all A_n can be nowhere dense. If A_{n_0} is not nowhere dense, it contains an open set G by the lemma above. By definition of A_{n_0}, we see that $|f(x)| \leq n_0$ for all $f \in \mathcal{F}$ and all $x \in A_{n_0}$ (and hence all $x \in G$). □

You may doubt the usefulness of this theorem as we only know that the result holds for *some* open set G, but the point is that if we have extra information on the the family $\mathcal{F}$, the sole existence of such a set may be exactly what we need to pull through a more complex argument. In functional analysis, there is a famous (and most useful) example of this called the *Banach-Steinhaus Theorem* (see Exercise 4.7.11).

For our next application, we first observe that although $\mathbb{R}^n$ is not compact, it can be written as a countable union of compact sets:

$$\mathbb{R}^n = \bigcup_{k \in \mathbb{N}} [-k, k]^n$$

We shall show that this is *not* the case for $C([0,1],\mathbb{R})$ — this space can not be written as a countable union of compact sets. We need a lemma.

Lemma 3.8.8 *A compact subset K of $C([0,1],\mathbb{R})$ is nowhere dense.*

Proof: Since compact sets are closed, it suffices (by the previous lemma) to show that each ball $\mathrm{B}(f;\epsilon)$ contains elements that are not in K. By Arzelà-Ascoli's Theorem, we know that compact sets are equicontinuous, and hence we need only prove that $\mathrm{B}(f;\epsilon)$ contains a family of functions that is not equicontinuous. We shall produce such a family by perturbing f by functions that are very steep on small intervals.

For each $n \in \mathbb{N}$, let g_n be the function

$$g_n(x) = \begin{cases} nx & \text{for } x \leq \frac{\epsilon}{2n} \\ \\ \frac{\epsilon}{2} & \text{for } x \geq \frac{\epsilon}{2n} \end{cases}$$

Then $f + g_n$ is in $B(f,\epsilon)$, but since $\{f + g_n\}$ is not equicontinuous (see Exercise 9 for help to prove this), all these functions can not be in K, and hence $\mathrm{B}(f;\epsilon)$ contains elements that are not in K. □

Proposition 3.8.9 *$C([0,1],\mathbb{R})$ is not a countable union of compact sets.*

Proof: Since $C([0,1],\mathbb{R})$ is complete, it is not the countable union of nowhere dense sets by Corollary 3.8.4. Since the lemma tells us that all compact sets are nowhere dense, the theorem follows. □

Remark: The basic idea in the proof above is that the compact sets are nowhere dense since we can obtain arbitrarily steep functions by perturbing a given function just a little. The same basic idea can be used to prove more sophisticated results, e.g. that the set of nowhere differentiable functions is comeager in $C([0,1],\mathbb{R})$. The key idea is that starting with any continuous function, we can perturb it into functions with arbitrarily large derivatives by using small, but rapidly oscillating functions. With a little bit of technical work, this implies that the set of functions that are differentiable at at least one point, is meager.

Exercises for Section 3.8

1. Show that $\mathbb{N}$ is a nowhere dense subset of $\mathbb{R}$.
2. Show that the set $A = \{g \in C([0,1],\mathbb{R}) \mid g(0) = 0\}$ is nowhere dense in $C([0,1],\mathbb{R})$.
3. Show that a subset of a nowhere dense set is nowhere dense and that a subset of a meager set is meager.
4. Show that a subset S of a metric space X is nowhere dense if and only if for each open ball $\mathrm{B}(a_0;r_0) \subseteq X$, there is a ball $\mathrm{B}(x;r) \subseteq \mathrm{B}(a_0;r_0)$ that does not intersect S.
5. Recall that the closure $\overline{N}$ of a set N consist of N plus all its boundary points.
   a) Show that if N is nowhere dense, so is $\overline{N}$.
   b) Find an example of a meager set M such that $\overline{M}$ is not meager.
   c) Show that a set is nowhere dense if and only if $\overline{N}$ does not contain any open balls.
6. Show that a countable union of meager sets is meager.
7. Show that if $N_1, N_2, \ldots, N_k$ are nowhere dense, so is $N_1 \cup N_2 \cup \ldots N_k$.
8. Prove that S is nowhere dense if and only if S^c contains an open, dense subset.
9. In this problem we shall prove that the set $\{f + g_n\}$ in the proof of Lemma 3.8.8 is not equicontinuous.
   a) Show that the set $\{g_n : n \in \mathbb{N}\}$ is not equicontinuous.
   b) Show that if $\{h_n\}$ is an equicontinous family of functions $h_n : [0,1] \to \mathbb{R}$ and $k : [0,1] \to \mathbb{R}$ is continuous, then $\{h_n + k\}$ is equicontinuous.
   c) Prove that the set $\{f + g_n\}$ in the lemma is not equicontinuous. (*Hint:* Assume that the sequence is equicontinuous, and use part b) with $h_n = f + g_n$ and $k = -f$ to get a contradiction with a)).
10. Let $\mathbb{N}$ have the discrete metric. Show that $\mathbb{N}$ is complete and that $\mathbb{N} = \bigcup_{n\in\mathbb{N}}\{n\}$. Why doesn't this contradict Baire's Category Theorem?
11. Show that in a complete space, a closed set is meager if and only if it is nowhere dense.
12. Let (X,d) be a metric space.
    a) Show that if $G \subseteq X$ is open and dense, then G^c is nowhere dense.
    b) Assume that (X,d) is complete. Show that if $\{G_n\}$ is a countable collection of open, dense subsets of X, then $\bigcap_{n\in\mathbb{N}} G_n$ is dense in X
13. Assume that a sequence $\{f_n\}$ of continuous functions $f_n : [0,1] \to \mathbb{R}$ converges pointwise to f. Show that f must be bounded on a subinterval of $[0,1]$. Find an example which shows that f need not be bounded on all of $[0,1]$.

14. In this problem we shall study sequences $\{f_n\}$ of functions converging pointwise to 0.

    a) Show that if the functions f_n are continuous, then there exists a nonempty subinterval (a, b) of $[0, 1]$ and an $N \in \mathbb{N}$ such that for $n \geq N$, $|f_n(x)| \leq 1$ for all $x \in (a, b)$.

    b) Find a sequence of functions $\{f_n\}$ converging to 0 on $[0, 1]$ such that for each nonempty subinterval (a, b) there is for each $N \in \mathbb{N}$ an $x \in (a, b)$ such that $f_N(x) > 1$.

15. Let (X, d) be a metric space. A point $x \in X$ is called *isolated* if there is an $\epsilon > 0$ such that $\mathrm{B}(x; \epsilon) = \{x\}$.

    a) Show that if $x \in X$, the singleton $\{x\}$ is nowhere dense if and only if x is not an isolated point.

    b) Show that if X is a complete metric space without isolated points, then X is uncountable.

    We shall now prove:

    Theorem: The unit interval $[0, 1]$ can *not* be written as a countable, disjoint union of closed, proper subintervals $I_n = [a_n, b_n]$.

    c) Assume for contradictions that $[0, 1]$ can be written as such a union. Show that the set of all endpoints, $F = \{a_n, b_n \mid n \in \mathbb{N}\}$ is a closed subset of $[0, 1]$, and that so is $F_0 = F \setminus \{0, 1\}$. Explain that since F_0 is countable and complete in the subspace metric, F_0 must have an isolated point, and use this to force a contradiction.

Chapter 4

Series of functions

In this chapter we shall see how the theory in the previous chapters can be used to study functions. We shall be particularly interested in how general functions can be written as sums of series of simple functions such as power functions and trigonometric functions. This will take us to the theories of power series and Fourier series.

4.1 $\limsup$ and $\liminf$

In this section we shall take a look at a useful extension of the concept of limit. Many sequences do not converge, but still have a rather regular asymptotic behavior as n goes to infinity — they may, for instance, oscillate between an upper set of values and a lower set. The notions of *limit superior*, $\limsup$, and *limit inferior*, $\liminf$, are helpful to describe such behavior. They also have the advantage that they always exist (provided we allow them to take the values $\pm\infty$).

We start with a sequence $\{a_n\}$ of real numbers, and define two new sequences $\{M_n\}$ and $\{m_n\}$ by

$$M_n = \sup\{a_k \mid k \geq n\}$$

and

$$m_n = \inf\{a_k \mid k \geq n\}$$

We allow that $M_n = \infty$ and that $m_n = -\infty$ as may well occur. Note that the sequence $\{M_n\}$ is decreasing (as we are taking suprema over smaller and smaller sets), and that $\{m_n\}$ is increasing (as we are taking infima over increasingly smaller sets). Since the sequences are monotone, the limits

$$\lim_{n\to\infty} M_n \qquad \text{and} \qquad \lim_{n\to\infty} m_n$$

clearly exist, but they may be $\pm\infty$. We now define the *limit superior* of the original sequence $\{a_n\}$ to be

$$\limsup_{n\to\infty} a_n = \lim_{n\to\infty} M_n$$

and the *limit inferior* to be

$$\liminf_{n\to\infty} a_n = \lim_{n\to\infty} m_n$$

The intuitive idea is that as n goes to infinity, the sequence $\{a_n\}$ may oscillate and not converge to a limit, but the oscillations will be asymptotically bounded by $\limsup a_n$ above and $\liminf a_n$ below.

The following relationship should be no surprise:

Proposition 4.1.1 *Let $\{a_n\}$ be a sequence of real numbers. Then*

$$\lim_{n\to\infty} a_n = b$$

if and only if

$$\limsup_{n\to\infty} a_n = \liminf_{n\to\infty} a_n = b$$

(we allow b to be a real number or $\pm\infty$.)

Proof: Assume first that $\limsup_{n\to\infty} a_n = \liminf_{n\to\infty} a_n = b$. Since $m_n \leq a_n \leq M_n$, and

$$\lim_{n\to\infty} m_n = \liminf_{n\to\infty} a_n = b\,,$$

$$\lim_{n\to\infty} M_n = \limsup_{n\to\infty} a_n = b\,,$$

we clearly have $\lim_{n\to\infty} a_n = b$ by "squeezing".

We now assume that $\lim_{n\to\infty} a_n = b$ where $b \in \mathbb{R}$ (the cases $b = \pm\infty$ are left to the reader). Given an $\epsilon > 0$, there exists an $N \in \mathbb{N}$ such that $|a_n - b| < \epsilon$ for all $n \geq N$. In other words

$$b - \epsilon < a_n < b + \epsilon$$

for all $n \geq N$. But then

$$b - \epsilon \leq m_n < b + \epsilon$$

and

$$b - \epsilon < M_n \leq b + \epsilon$$

for $n \geq N$. Since this holds for all $\epsilon > 0$, we have $\limsup_{n\to\infty} a_n = \liminf_{n\to\infty} a_n = b$ □

Exercises for section 4.1

1. Let $a_n = (-1)^n$. Find $\limsup_{n\to\infty} a_n$ and $\liminf_{n\to\infty} a_n$.
2. Let $a_n = \cos\frac{n\pi}{2}$. Find $\limsup_{n\to\infty} a_n$ and $\liminf_{n\to\infty} a_n$.
3. Let $a_n = \arctan(n)\sin\left(\frac{n\pi}{2}\right)$. Find $\limsup_{n\to\infty} a_n$ and $\liminf_{n\to\infty} a_n$.
4. Complete the proof of Proposition 4.1.1 for the case $b = \infty$.
5. Show that
$$\limsup_{n\to\infty}(a_n + b_n) \leq \limsup_{n\to\infty} a_n + \limsup_{n\to\infty} b_n$$
and
$$\liminf_{n\to\infty}(a_n + b_n) \geq \liminf_{n\to\infty} a_n + \liminf_{n\to\infty} b_n$$
and find examples which show that we do not in general have equality. State and prove a similar result for the product $\{a_n b_n\}$ of two *positive* sequences.
6. Assume that the sequence $\{a_n\}$ is nonnegative and converges to a, and that $b = \limsup b_n$ is finite and positive. Show that $\limsup_{n\to\infty} a_n b_n = ab$ (the result holds without the condition that b is positive, but the proof becomes messy). What happens if the sequence $\{a_n\}$ is negative?
7. We shall see how we can define lim sup and lim inf for functions $f : \mathbb{R} \to \mathbb{R}$. Let $a \in \mathbb{R}$, and define
$$M_\epsilon = \sup\{f(x) \mid x \in (a-\epsilon, a+\epsilon)\}$$
$$m_\epsilon = \inf\{f(x) \mid x \in (a-\epsilon, a+\epsilon)\}$$
for $\epsilon > 0$ (we allow $M_\epsilon = \infty$ and $m_\epsilon = -\infty$).
   a) Show that M_ϵ decreases and m_ϵ increases as $\epsilon \to 0$.
   b) Show that $\limsup_{x\to a} f(x) = \lim_{\epsilon\to 0^+} M_\epsilon$ and $\liminf_{x\to a} f(x) = \lim_{\epsilon\to 0^+} m_\epsilon$ exist (we allow $\pm\infty$ as values).
   c) Show that $\lim_{x\to a} f(x) = b$ if and only if $\limsup_{x\to a} f(x) = \liminf_{x\to a} f(x) = b$
   d) Find $\liminf_{x\to 0} \sin\frac{1}{x}$ and $\limsup_{x\to 0} \sin\frac{1}{x}$

4.2 Integrating and differentiating sequences

Assume that we have a sequence of functions $\{f_n\}$ converging to a limit function f. If we integrate the functions f_n, will the integrals converge to the integral of f? And if we differentiate the f_n's, will the derivatives converge to f'?

In this section, we shall see that without any further restrictions, the answer to both questions are no, but that it is possible to put conditions on the sequences that turn the answers into yes.

Let us start with integration and the following example.

Example 1: Let $f_n : [0,1] \to \mathbb{R}$ be the function in the figure.

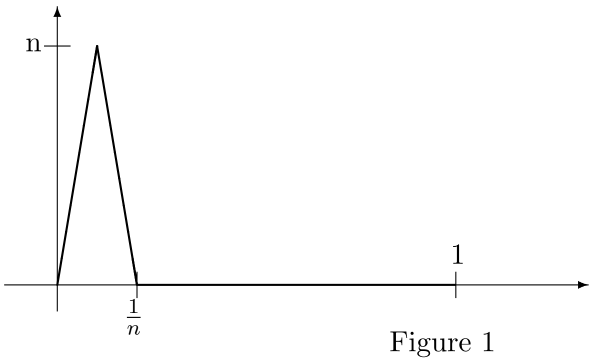


Figure 1

It is given by the formula

$$f_n(x) = \begin{cases} 2n^2x & \text{if } 0 \leq x < \frac{1}{2n} \\ \\ -2n^2x + 2n & \text{if } \frac{1}{2n} \leq x < \frac{1}{n} \\ \\ 0 & \text{if } \frac{1}{n} \leq x \leq 1 \end{cases}$$

but it is much easier just to work from the picture. The sequence $\{f_n\}$ converges pointwise to 0, but the integrals do not not converge to 0. In fact, $\int_0^1 f_n(x)\,dx = \frac{1}{2}$ since the value of the integral equals the area under the function graph, i.e. the area of a triangle with base $\frac{1}{n}$ and height n. ♣

The example above shows that if the functions f_n converge *pointwise* to a function f on an interval $[a,b]$, the integrals $\int_a^b f_n(x)\,dx$ need not converge to $\int_a^b f(x)\,dx$. The reason is that with pointwise convergence, the difference between f and f_n may be very large on small sets — so large that the integrals of f_n do not converge to the integral of f. If the convergence is *uniform*, this can not happen (note that the result below is actually a special case of Lemma 3.6.1):

Proposition 4.2.1 *Assume that $\{f_n\}$ is a sequence of continuous functions converging uniformly to f on the interval $[a,b]$. Then the functions*

$$F_n(x) = \int_a^x f_n(t)\,dt$$

converge uniformly to

$$F(x) = \int_a^x f(t)\,dt$$

on $[a,b]$.

Proof: We must show that for a given $\epsilon > 0$, we can always find an $N \in \mathbb{N}$ such that $|F(x) - F_n(x)| < \epsilon$ for all $n \geq N$ and all $x \in [a,b]$. Since $\{f_n\}$

converges uniformly to f, there is an $N \in \mathbb{N}$ such that $|f(t) - f_n(t)| < \frac{\epsilon}{b-a}$ for all $t \in [a,b]$. For $n \geq N$, we then have for all $x \in [a,b]$:

$$|F(x) - F_n(x)| = |\int_a^x (f(t) - f_n(t))\,dt\,| \leq \int_a^x |f(t) - f_n(t)|\,dt \leq$$

$$\leq \int_a^x \frac{\epsilon}{b-a}\,dt \leq \int_a^b \frac{\epsilon}{b-a}\,dt = \epsilon$$

This shows that $\{F_n\}$ converges uniformly to F on $[a,b]$. □

In applications it is often useful to have the result above with a flexible lower limit.

Corollary 4.2.2 *Assume that $\{f_n\}$ is a sequence of continuous functions converging uniformly to f on the interval $[a,b]$. For any $x_0 \in [a,b]$, the functions*

$$F_n(x) = \int_{x_0}^x f_n(t)\,dt$$

converge uniformly to

$$F(x) = \int_{x_0}^x f(t)\,dt$$

on $[a,b]$.

Proof: Recall that

$$\int_a^x f_n(t)\,dt = \int_a^{x_0} f_n(t)\,dt + \int_{x_0}^x f_n(t)\,dt$$

regardless of the order of the numbers a, x_0, x, and hence

$$\int_{x_0}^x f_n(t)\,dt = \int_a^x f_n(t)\,dt - \int_a^{x_0} f_n(t)\,dt$$

The first integral on the right converges uniformly to $\int_a^x f(t)\,dt$ according to the proposition, and the second integral converges (as a sequence of numbers) to $\int_a^{x_0} f(t)\,dt$. Hence $\int_{x_0}^x f_n(t)\,dt$ converges uniformly to

$$\int_a^x f(t)\,dt - \int_a^{x_0} f(t)\,dt = \int_{x_0}^x f(t)\,dt$$

as was to be proved. □

Let us reformulate this result in terms of series. Recall that a series of functions $\sum_{n=0}^{\infty} v_n(x)$ converges pointwise/unifomly to a function f on an interval I if an only if the sequence $\{s_n\}$ of partial sum $s_n(x) = \sum_{k=0}^{n} v_k(x)$ converges pointwise/uniformly to f on I.

Corollary 4.2.3 *Assume that* $\{v_n\}$ *is a sequence of continuous functions such that the series* $\sum_{n=0}^{\infty} v_n(x)$ *converges uniformly on the interval* $[a,b]$*. Then for any* $x_0 \in [a,b]$*, the series* $\sum_{n=0}^{\infty} \int_{x_0}^{x} v_n(t)\,dt$ *converges uniformly and*

$$\sum_{n=0}^{\infty} \int_{x_0}^{x} v_n(t)\,dt = \int_{x_0}^{x} \sum_{n=0}^{\infty} v_n(t)\,dt$$

The corollary tell us that if the series $\sum_{n=0}^{\infty} v_n(x)$ converges uniformly, we can integrate it term by term to get

$$\int_{x_0}^{x} \sum_{n=0}^{\infty} v_n(t)\,dt = \sum_{n=0}^{\infty} \int_{x_0}^{x} v_n(t)\,dt$$

This formula may look obvious, but it does not in general hold for series that only converge pointwise. As we shall see later, interchanging integrals and infinite sums is quite a tricky business.

To use the corollary efficiently, we need to be able to determine when a series of functions converges uniformly. The following simple test is often helpful:

Proposition 4.2.4 (Weierstrass' M-test) *Let* $\{v_n\}$ *be a sequence of continuous functions on the interval* $[a,b]$*, and assume that there is a convergent series* $\sum_{n=0}^{\infty} M_n$ *of positive numbers such that* $|v_n(x)| \leq M_n$ *for all* $n \in \mathbb{N}$ *and all* $x \in [a,b]$*. Then series* $\sum_{n=0}^{\infty} v_n(x)$ *converges uniformly on* $[a,b]$*.*

Proof: Since $(C([a,b],\mathbb{R}),\rho)$ is complete, we only need to check that the partial sums $s_n(x) = \sum_{k=0}^{n} v_k(x)$ form a Cauchy sequence. Since the series $\sum_{n=0}^{\infty} M_n$ converges, we know that its partial sums $S_n = \sum_{k=0}^{n} M_k$ form a Cauchy sequence. Since for all $x \in [a,b]$ and all $m > n$,

$$|s_m(x) - s_n(x)| = \left|\sum_{k=n+1}^{m} v_k(x)\right| \leq \sum_{k=n+1}^{m} |v_k(x)| \leq \sum_{k=n+1}^{m} M_k = |S_m - S_n|,$$

this implies that $\{s_n\}$ is a Cauchy sequence. □

Example 1: Consider the series $\sum_{n=1}^{\infty} \frac{\cos nx}{n^2}$. Since $|\frac{\cos nx}{n^2}| \leq \frac{1}{n^2}$, and $\sum_{n=0}^{\infty} \frac{1}{n^2}$ converges, the original series $\sum_{n=1}^{\infty} \frac{\cos nx}{n^2}$ converges uniformly to a function f on any closed and bounded interval $[a,b]$. Hence we may intergrate termwise to get

$$\int_0^x f(t)\,dt = \sum_{n=1}^{\infty} \int_x \frac{\cos nt}{n^2}\,dt = \sum_{n=1}^{\infty} \frac{\sin nx}{n^3}$$

♣

Let us now turn to differentiation of sequences. This is a much trickier business than integration as integration often helps to smoothen functions while differentiation tends to make them more irregular. Here is a simple example.

Example 2: The sequence (not series!) $\{\frac{\sin nx}{n}\}$ obviously converges uniformly to 0, but the sequence of derivatives $\{\cos nx\}$ does not converge at all. ♣

The example shows that even if a sequence $\{f_n\}$ of differentiable functions converges uniformly to a differentiable function f, the derivatives f_n' need not converge to the derivative f' of the limit function. If you draw the graphs of the functions f_n, you will see why — although they live in an increasingly narrower strip around the x-axis, they all wriggle equally much, and the derivatives do not converge.

To get a theorem that works, we have to put the conditions on the derivatives. The following result may look ugly and unsatisfactory, but it gives us the information we shall need.

Proposition 4.2.5 *Let $\{f_n\}$ be a sequence of differentiable functions on the interval $[a,b]$. Assume that the derivatives f_n' are continuous and that they converge uniformly to a function g on $[a,b]$. Assume also that there is a point $x_0 \in [a,b]$ such that the sequence $\{f(x_0)\}$ converges. Then the sequence $\{f_n\}$ converges uniformly on $[a,b]$ to a differentiable function f such that $f' = g$.*

Proof: The proposition is just Corollary 4.2.2 in a convenient disguise. If we apply that proposition to the sequence $\{f_n'\}$, we se that the integrals $\int_{x_0}^x f_n'(t)\,dt$ converge uniformly to $\int_{x_0}^x g(t)\,dt$. By the Fundamental Theorem of Calculus, we get

$$f_n(x) - f_n(x_0) \to \int_{x_0}^x g(t)\,dt \qquad \text{uniformly on } [a,b]$$

Since $f_n(x_0)$ converges to a limit b, this means that $f_n(x)$ converges uniformly to the function $f(x) = b + \int_{x_0}^x g(t)\,dt$. Using the Fundamental Theorem of Calculus again, we see that $f'(x) = g(x)$. □

Also in this case it is useful to have a reformulation in terms of series:

Corollary 4.2.6 *Let $\sum_{n=0}^\infty u_n(x)$ be a series where the functions u_n are differentiable with continuous derivatives on the interval $[a,b]$. Assume that the series of derivatives $\sum_{n=0}^\infty u_n'(x)$ converges uniformly on $[a,b]$. Assume also that there is a point $x_0 \in [a,b]$ where the series $\sum_{n=0}^\infty u_n(x_0)$ converges.*

Then the series $\sum_{n=0}^{\infty} u_n(x)$ *converges uniformly on* $[a, b]$*, and*

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

The corollary tells us that under rather strong conditions, we can differentiate the series $\sum_{n=0}^{\infty} u_n(x)$ term by term.

Example 3: Summing a geometric series, we see that

$$\frac{1}{1-e^{-x}} = \sum_{n=0}^{\infty} e^{-nx} \qquad \text{for } x > 0 \tag{4.2.1}$$

If we can differentiate term by term on the right hand side, we shall get

$$\frac{e^{-x}}{(1-e^{-x})^2} = \sum_{n=1}^{\infty} ne^{-nx} \qquad \text{for } x > 0 \tag{4.2.2}$$

To check that this is correct, we must check the convergence of the differentiated series (4.2.2). Choose an interval $[a, b]$ where $a > 0$, then $ne^{-nx} \leq ne^{-na}$ for all $x \in [a, b]$. Using, e.g., the ratio test, it is easy to see that the series $\sum_{n=0}^{\infty} ne^{-na}$ converges, and hence $\sum_{n=0}^{\infty} ne^{-nx}$ converges uniformly on $[a, b]$ by Weierstrass' M-test. The corollary now tells us that the sum of the sequence (4.2.2) is the derivative of the sum of the sequence (4.2.1), i.e.

$$\frac{e^{-x}}{(1-e^{-x})^2} = \sum_{n=1}^{\infty} ne^{-nx} \qquad \text{for } x \in [a, b]$$

Since $[a, b]$ is an arbitrary subinterval of $(0, \infty)$, we have

$$\frac{e^{-x}}{(1-e^{-x})^2} = \sum_{n=1}^{\infty} ne^{-nx} \qquad \text{for all } x > 0$$

♣

Exercises for Section 4.2

1. Show that $\sum_{n=0}^{\infty} \frac{\cos(nx)}{n^2+1}$ converges uniformly on $\mathbb{R}$.
2. Does the series $\sum_{n=0}^{\infty} ne^{-nx}$ in Example 3 converge uniformly on $(0, \infty)$?
3. Let $f_n : [0, 1] \to \mathbb{R}$ be defined by $f_n(x) = nx(1-x^2)^n$. Show that $f_n(x) \to 0$ for all $x \in [0, 1]$, but that $\int_0^1 f_n(x)\, dx \to \frac{1}{2}$.
4. Explain in detail how Corollary 4.2.3 follows from Corollary 4.2.2.
5. Explain in detail how Corollary 4.2.6 follows from Proposition 4.2.5.

6. a) Show that series $\sum_{n=1}^{\infty} \frac{\cos \frac{x}{n}}{n^2}$ converges uniformly on $\mathbb{R}$.

   b) Show that $\sum_{n=1}^{\infty} \frac{\sin \frac{x}{n}}{n}$ converges to a continuous function f, and that

$$f'(x) = \sum_{n=1}^{\infty} \frac{\cos \frac{x}{n}}{n^2}$$

7. One can show that

$$x = \sum_{n=1}^{\infty} \frac{2(-1)^{n+1}}{n} \sin(nx) \qquad \text{for } x \in (-\pi, \pi)$$

   If we differentiate term by term, we get

$$1 = \sum_{n=1}^{\infty} 2(-1)^{n+1} \cos(nx) \qquad \text{for } x \in (-\pi, \pi)$$

   Is this a correct formula?

8. a) Show that the sequence $\sum_{n=1}^{\infty} \frac{1}{n^x}$ converges uniformly on all intervals $[a, \infty)$ where $a > 1$.

   b) Let $f(x) = \sum_{n=1}^{\infty} \frac{1}{n^x}$ for $x > 1$. Show that $f'(x) = -\sum_{n=1}^{\infty} \frac{\ln x}{n^x}$.

4.3 Power series

Recall that a power series is a function of the form

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

where a is a real number and $\{c_n\}$ is a sequence of real numbers. It is defined for the x-values that make the series converge. We define the *radius of convergence* of the series to be the number R such that

$$\frac{1}{R} = \limsup_{n\to\infty} \sqrt[n]{|c_n|}$$

with the interpretation that $R = 0$ if the limit is infinite, and $R = \infty$ if the limit is 0. To justify this terminology, we need the the following result.

Proposition 4.3.1 *If R is the radius of convergence of the power series $\sum_{n=0}^{\infty} c_n (x-a)^n$, the series converges for $|x-a| < R$ and diverges for $|x-a| > R$. If $0 < r < R$, the series converges uniformly on $[a-r, a+r]$.*

Proof: Let us first assume that $|x-a| > R$. This means that $\frac{1}{|x-a|} < \frac{1}{R}$, and since $\limsup_{n\to\infty} \sqrt[n]{|c_n|} = \frac{1}{R}$, there must be arbitrarily large values of n such that $\sqrt[n]{|c_n[} > \frac{1}{|x-a|}$. Hence $|c_n(x-a)^n| > 1$, and consequently the series must diverge as the terms do not decrease to zero.

To prove the (uniform) convergence, assume that r is a number between 0 and R. Since $\frac{1}{r} > \frac{1}{R}$, we can pick a positive number $b < 1$ such that $\frac{b}{r} > \frac{1}{R}$. Since $\limsup_{n\to\infty} \sqrt[n]{|c_n|} = \frac{1}{R}$, there must be an $N \in \mathbb{N}$ such that $\sqrt[n]{|c_n|} < \frac{b}{r}$ when $n \geq N$. This means that $|c_n r^n| < b^n$ for $n \geq N$, and hence that $|c_n(x-a)|^n < b^n$ for all $x \in [a-r, a+r]$. Since $\sum_{n=N}^{\infty} b^n$ is a convergent, geometric series, Weierstrass' M-test tells us that the series $\sum_{n=N}^{\infty} c_n(x-a)^n$ converges uniformly on $[a-r, a+r]$. Since only the tail of a sequence counts for convergence, the full series $\sum_{n=0}^{\infty} c_n(x-a)^n$ also converges uniformly on $[a-r, a+r]$. Since r is an arbitrary number less than R, we see that the series must converge on the open interval $(a-R, a+R)$, i.e. whenever $|x-a| < R$. □

Remark: When we want to find the radius of convergence, it is occasionally convenient to compute a slightly different limit such as $\lim_{n\to\infty} \sqrt[n+1]{c_n}$ or $\lim_{n\to\infty} \sqrt[n-1]{c_n}$ instead of $\lim_{n\to\infty} \sqrt[n]{c_n}$. This corresponds to finding the radius of convergence of the power series we get by either multiplying or dividing the original one by $(x-a)$, and gives the correct answer as multiplying or dividing a series by a non-zero number doesn't change its convergence properties.

The proposition above does not tell us what happens at the endpoints $a \pm R$ of the interval of convergence, but we know from calculus that a series may converge at both, one or neither endpoint. Although the convergence is uniform on all subintervals $[a-r, a+r]$, it is not in general uniform on $(a-R, a+R)$.

Corollary 4.3.2 *Assume that the power series $f(x) = \sum_{n=0}^{\infty} c_n(x-a)^n$ has radius of convergence R larger than 0. Then the function f is continuous and differentiable on the open interval $(a-R, a+R)$ with*

$$f'(x) = \sum_{n=1}^{\infty} nc_n(x-a)^{n-1} = \sum_{n=0}^{\infty}(n+1)c_{n+1}(x-a)^n \qquad \text{for } x \in (a-R, a+R)$$

and

$$\int_a^x f(t)\,dt = \sum_{n=0}^{\infty} \frac{c_n}{n+1}(x-a)^{n+1} = \sum_{n=1}^{\infty} \frac{c_{n-1}}{n}(x-a)^n \qquad \text{for } x \in (a-R, a+R)$$

Proof: Since the power series converges uniformly on each subinterval $[a-r, a+r]$, the sum is continuous on each such interval according to Proposition 3.2.4. Since each x in $(a-R, a+R)$ is contained in the interior of some of the subintervals $[a-r, a+r]$, we see that f must be continuous on the full interval $(a-R, a+R)$. The formula for the integral follows immediately by applying Corollary 4.2.3 on each subinterval $[a-r, a+r]$ in a similar way.

To get the formula for the derivative, we shall apply Corollary 4.2.6. To use this result, we need to know that the differentiated series $\sum_{n=1}^{\infty}(n+1)c_{n+1}(x-a)^n$ has the same radius of convergence as the original series; i.e. that

$$\limsup_{n\to\infty} \sqrt[n+1]{|(n+1)c_{n+1}|} = \limsup_{n\to\infty} \sqrt[n]{|c_n|} = \frac{1}{R}$$

(note that by the remark above, we may use the $n+1$-st root on the left hand side instead of the n-th root). Since $\lim_{n\to\infty} \sqrt[n+1]{(n+1)} = 1$, this is not hard to show (see Exercise 6). Applying Corollary 4.2.6 on each subinterval $[a-r, a+r]$, we now get the formula for the derivative at each point $x \in [a-r, a+r]$. Since each point in $(a-R, a+R)$ belongs to the interior of some of the subintervals, the formula for the derivative must hold at all points $x \in (a-R, a+R)$. □

A function that is the sum of a power series, is called a *real analytic function.* Such functions have derivatives of all orders.

Corollary 4.3.3 *Let $f(x) = \sum_{n=0}^{\infty} c_n(x-a)^n$ for $x \in (a-R, a+R)$. Then f is k times differentiable in $(a-R, a+R)$ for any $k \in \mathbb{N}$, and $f^{(k)}(a) = k!c_k$. Hence $\sum_{n=0}^{\infty} c_n(x-a)^n$ is the Taylor series*

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

Proof: Using the previous corollary, we get by induction that $f^{(k)}$ exists on $(a-R, a+R)$ and that

$$f^{(k)}(x) = \sum_{n=k}^{\infty} n(n-1)\cdot \ldots \cdot (n-k+1)c_n(x-a)^{n-k}$$

Putting $x = a$, we get $f^{(k)}(a) = k!c_k$, and the corollary follows. □

Exercises for Section 4.3

1. Find power series with radius of convergence 0, 1, 2, and ∞.

2. Find power series with radius of convergence 1 that converge at both, one and neither of the endpoints.

3. Show that for any polynomial P, $\lim_{n\to\infty} \sqrt[n]{|P(n)|} = 1$.

4. Use the result in Exercise 3 to find the radius of convergence:

   a) $\sum_{n=0}^{\infty} \frac{2^n x^n}{n^3+1}$

b) $\sum_{n=0}^{\infty} \frac{2n^2+n-1}{3n+4} x^n$

c) $\sum_{n=0}^{\infty} nx^{2n}$

5. a) Explain that $\frac{1}{1-x^2} = \sum_{n=0}^{\infty} x^{2n}$ for $|x| < 1$,

   b) Show that $\frac{2x}{(1-x^2)^2} = \sum_{n=0}^{\infty} 2nx^{2n-1}$ for $|x| < 1$.

   c) Show that $\frac{1}{2}\ln\left|\frac{1+x}{1-x}\right| = \sum_{n=0}^{\infty} \frac{x^{2n+1}}{2n+1}$ for $|x| < 1$.

6. Let $\sum_{n=0}^{\infty} c_n(x-a)^n$ be a power series.

   a) Show that the radius of convergence is given by

$$\frac{1}{R} = \limsup_{n\to\infty} \sqrt[n+k]{|c_n|}$$

   for any integer k.

   b) Show that $\lim_{n\to\infty} \sqrt[n+1]{n+1} = 1$ (write $\sqrt[n+1]{n+1} = (n+1)^{\frac{1}{n+1}}$).

   c) Prove the formula

$$\limsup_{n\to\infty} \sqrt[n+1]{|(n+1)c_{n+1}|} = \limsup_{n\to\infty} \sqrt[n]{|c_n|} = \frac{1}{R}$$

   in the proof of Corollary 4.3.2.

4.4 Abel's Theorem

We have seen that the sum $f(x) = \sum_{n=0}^{\infty} c_n(x-a)^n$ of a power series is continuous in the interior $(a-R, a+R)$ of its interval of convergence. But what happens if the series converges at an endpoint $a \pm R$? It turns out that the sum is also continuous at the endpoint, but that this is surprisingly intricate to prove.

Before we turn to the proof, we need a lemma that can be thought of as a discrete version of integration by parts.

Lemma 4.4.1 (Abel's Summation Formula) *Let $\{a_n\}_{n=0}^{\infty}$ and $\{b_n\}_{n=0}^{\infty}$ be two sequences of real numbers, and let $s_n = \sum_{k=0}^{n} a_k$. Then*

$$\sum_{n=0}^{N} a_n b_n = s_N b_N + \sum_{n=0}^{N-1} s_n(b_n - b_{n+1}).$$

If the series $\sum_{n=0}^{\infty} a_n$ converges, and $b_n \to 0$ as $n \to \infty$, then

$$\sum_{n=0}^{\infty} a_n b_n = \sum_{n=0}^{\infty} s_n(b_n - b_{n+1})$$

in the sense that either the two series both diverge or they converge to the same limit.

Proof: Note that $a_n = s_n - s_{n-1}$ for $n \geq 1$, and that this formula even holds for $n = 0$ if we define $s_{-1} = 0$. Hence

$$\sum_{n=0}^{N} a_n b_n = \sum_{n=0}^{N} (s_n - s_{n-1}) b_n = \sum_{n=0}^{N} s_n b_n - \sum_{n=0}^{N} s_{n-1} b_n$$

Changing the index of summation and using that $s_{-1} = 0$, we see that $\sum_{n=0}^{N} s_{n-1} b_n = \sum_{n=0}^{N-1} s_n b_{n+1}$. Putting this into the formula above, we get

$$\sum_{n=0}^{N} a_n b_n = \sum_{n=0}^{N} s_n b_n - \sum_{n=0}^{N-1} s_n b_{n+1} = s_N b_N + \sum_{n=0}^{N-1} s_n (b_n - b_{n+1})$$

and the first part of the lemma is proved. The second follows by letting $N \to \infty$. □

We are now ready to prove:

Theorem 4.4.2 (Abel's Theorem) *The sum of a power series $f(x) = \sum_{n=0}^{\infty} c_n (x-a)^n$ is continuous in its entire interval of convergence. This means in particular that if R is the radius of convergence, and the power series converges at the right endpoint $a+R$, then $\lim_{x \uparrow a+R} f(x) = f(a+R)$, and if the power series converges at the left endpoint $a-R$, then $\lim_{x \downarrow a-R} f(x) = f(a-R)$.*[1]

Proof: We already know that f is continuous in the open interval $(a-R, a+R)$, and that we only need to check the endpoints. To keep the notation simple, we shall assume that $a = 0$ and concentrate on the right endpoint R. Thus we want to prove that $\lim_{x \uparrow R} f(x) = f(R)$.

Note that $f(x) = \sum_{n=0}^{\infty} c_n R^n \left(\frac{x}{R}\right)^n$. If we assume that $|x| < R$, we may apply the second version of Abel's summation formula with $a_n = c_n R^n$ and $b_n = \left(\frac{x}{n}\right)^n$ to get

$$f(x) = \sum_{n=0}^{\infty} f_n(R) \left(\left(\frac{x}{R}\right)^n - \left(\frac{x}{R}\right)^{n+1} \right) = \left(1 - \frac{x}{R}\right) \sum_{n=0}^{\infty} f_n(R) \left(\frac{x}{R}\right)^n$$

where $f_n(R) = \sum_{k=0}^{n} c_k R^k$. Summing a geometric series, we see that we also have

$$f(R) = \left(1 - \frac{x}{R}\right) \sum_{n=0}^{\infty} f(R) \left(\frac{x}{R}\right)^n$$

Hence

$$|f(x) - f(R)| = \left| \left(1 - \frac{x}{R}\right) \sum_{n=0}^{\infty} (f_n(R) - f(R)) \left(\frac{x}{R}\right)^n \right|$$

[1] I use $\lim_{x \uparrow b}$ and $\lim_{x \downarrow b}$ for one-sided limits, also denoted by $\lim_{x \to b^-}$ and $\lim_{x \to b^+}$.

Given an $\epsilon > 0$, we must find a $\delta > 0$ such that this quantity is less than ϵ when $R - \delta < x < R$. This may seem obvious due to the factor $(1 - x/R)$, but the problem is that the infinite series may go to infinity when $x \to R$. Hence we need to control the tail of the sequence before we exploit the factor $(1 - x/R)$. Fortunately, this is not difficult: Since $f_n(R) \to f(R)$, we first pick an $N \in \mathbb{N}$ such that $|f_n(R) - f(R)| < \frac{\epsilon}{2}$ for $n \geq N$. Then

$$|f(x) - f(R)| \leq \left(1 - \frac{x}{R}\right) \sum_{n=0}^{N-1} |f_n(R) - f(R)| \left(\frac{x}{R}\right)^n +$$

$$+ \left(1 - \frac{x}{R}\right) \sum_{n=N}^{\infty} |f_n(R) - f(R)| \left(\frac{x}{R}\right)^n \leq$$

$$\leq \left(1 - \frac{x}{R}\right) \sum_{n=0}^{N-1} |f_n(R) - f(R)| \left(\frac{x}{R}\right)^n + \left(1 - \frac{x}{R}\right) \sum_{n=0}^{\infty} \frac{\epsilon}{2} \left(\frac{x}{R}\right)^n =$$

$$= \left(1 - \frac{x}{R}\right) \sum_{n=0}^{N-1} |f_n(R) - f(R)| \left(\frac{x}{R}\right)^n + \frac{\epsilon}{2}$$

where we have summed a geometric series. Now the sum is finite, and the first term clearly converges to 0 when $x \uparrow R$. Hence there is a $\delta > 0$ such that this term is less than $\frac{\epsilon}{2}$ when $R - \delta < x < R$, and consequently $|f(x) - f(R)| < \epsilon$ for such values of x. □

Let us take a look at a famous example.

Example 1: Summing a geometric series, we clearly have

$$\frac{1}{1 + x^2} = \sum_{n=0}^{\infty} (-1)^n x^{2n} \qquad \text{for } |x| < 1$$

Integrating, we get

$$\arctan x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{2n+1} \qquad \text{for } |x| < 1$$

Using the Alternating Series Test, we see that the series converges even for $x = 1$. By Abel's Theorem

$$\frac{\pi}{4} = \arctan 1 = \lim_{x \uparrow 1} \arctan x = \lim_{x \uparrow 1} \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{2n+1} = \sum_{n=0}^{\infty} (-1)^n \frac{1}{2n+1}$$

Hence we have proved

$$\frac{\pi}{4} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \ldots$$

This is often called Leibniz' or Gregory's formula for π, but it was actually first discovered by the Indian mathematician Madhava (ca. 1350 – ca. 1425). ♣

This example is rather typical; the most interesting information is often obtained at an endpoint, and we need Abel's Theorem to secure it.

It is natural to think that Abel's Theorem must have a converse saying that if $\lim_{x\uparrow a+R}\sum_{n=0}^{\infty} c_n x^n$ exists, then the sequence converges at the right endpoint $x = a + R$. This, however, is not true as the following simple example shows.

Example 2: Summing a geometric series, we have

$$\frac{1}{1+x} = \sum_{n=0}^{\infty}(-x)^n \qquad \text{for } |x| < 1$$

Obviously, $\lim_{x\uparrow 1}\sum_{n=0}^{\infty}(-x)^n = \lim_{x\uparrow 1}\frac{1}{1+x} = \frac{1}{2}$, but the series does not converge for $x = 1$. ♣

It is possible to put extra conditions on the coefficients of the series to ensure convergence at the endpoint, see Exercise 2.

Exercises for Section 4.4

1. a) Explain why $\frac{1}{1+x} = \sum_{n=0}^{\infty}(-1)^n x^n$ for $|x| < 1$.

   b) Show that $\ln(1+x) = \sum_{n=0}^{\infty}(-1)^n \frac{x^{n+1}}{n+1}$ for $|x| < 1$.

   c) Show that $\ln 2 = \sum_{n=0}^{\infty}(-1)^n \frac{1}{n+1}$.

2. In this problem we shall prove the following partial converse of Abel's Theorem:

   Tauber's Theorem Assume that $s(x) = \sum_{n=0}^{\infty} c_n x^n$ is a power series with radius of convergence 1. Assume that $s = \lim_{x\uparrow 1}\sum_{n=0}^{\infty} c_n x^n$ is finite. If in addition $\lim_{n\to\infty} nc_n = 0$, then the power series converges for $x = 1$ and $s = s(1)$.

   a) Explain that if we can prove that the power series converges for $x = 1$, then the rest of the theorem will follow from Abel's Theorem.

   b) Show that $\lim_{N\to\infty}\frac{1}{N}\sum_{n=0}^{N} n|c_n| = 0$.

   c) Let $s_N = \sum_{n=0}^{N} c_n$. Explain that

   $$s(x) - s_N = -\sum_{n=0}^{N} c_n(1-x^n) + \sum_{n=N+1}^{\infty} c_n x^n$$

   d) Show that $1 - x^n \leq n(1-x)$ for $|x| < 1$.

e) Let N_x be the integer such that $N_x \leq \frac{1}{1-x} < N_x + 1$ Show that

$$\sum_{n=0}^{N_x} c_n(1-x^n) \leq (1-x)\sum_{n=0}^{N_x} n|c_n| \leq \frac{1}{N_x}\sum_{n=0}^{N_x} n|c_n| \to 0$$

as $x \uparrow 1$.

f) Show that

$$\left|\sum_{n=N_x+1}^{\infty} c_n x^n\right| \leq \sum_{n=N_x+1}^{\infty} n|c_n|\frac{x^n}{n} = \frac{d_x}{N_x}\sum_{n=0}^{\infty} x^n$$

where $d_x \to 0$ as $x \uparrow 1$. Show that $\sum_{n=N_x+1}^{\infty} c_n x^n \to 0$ as $x \uparrow 1$.

g) Prove Tauber's theorem.

4.5 Normed spaces

In a later chapter we shall continue our study of how general functions can be expressed as series of simpler functions. This time the "simple functions" will be trigonometric functions and not power functions, and the series will be called Fourier series and not power series. Before we turn to Fourier series, we shall take a look at normed spaces and inner product spaces. Strictly speaking, it is not necessary to know about such spaces to study Fourier series, but a basic understanding will make it much easier to appreciate the basic ideas and put them into a wider framework.

In Fourier analysis, one studies vector spaces of functions, and let me begin by reminding you that a vector space is just a set where you can add elements and multiply them by numbers in a reasonable way. More precisely:

Definition 4.5.1 *Let $\mathbb{K}$ be either $\mathbb{R}$ or $\mathbb{C}$, and let V be a nonempty set. Assume that V is equipped with two operations:*

- Addition *which to any two elements $\mathbf{u}, \mathbf{v} \in V$ assigns an element $\mathbf{u} + \mathbf{v} \in V$.*
- Scalar multiplication *which to any element $\mathbf{u} \in V$ and any number $\alpha \in \mathbb{K}$ assigns an element $\alpha\mathbf{u} \in V$.*

We call V a vector space over $\mathbb{K}$ *if the following axioms are satisfied:*

(i) $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$ *for all* $\mathbf{u}, \mathbf{v} \in V$.

(ii) $(\mathbf{u} + \mathbf{v}) + \mathbf{w} = \mathbf{u} + (\mathbf{v} + \mathbf{w})$ *for all* $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$.

(iii) There is a zero vector $\mathbf{0} \in V$ *such that* $\mathbf{u} + \mathbf{0} = \mathbf{u}$ *for all* $\mathbf{u} \in V$.

(iv) For each $\mathbf{u} \in V$*, there is an element* $-\mathbf{u} \in V$ *such that* $\mathbf{u} + (-\mathbf{u}) = \mathbf{0}$.

(v) $\alpha(\mathbf{u}+\mathbf{v}) = \alpha\mathbf{u} + \alpha\mathbf{v}$ *for all* $\mathbf{u}, \mathbf{v} \in V$ *and all* $\alpha \in \mathbb{K}$.

(vi) $(\alpha+\beta)\mathbf{u} = \alpha\mathbf{u} + \beta\mathbf{u}$ *for all* $\mathbf{u} \in V$ *and all* $\alpha, \beta \in \mathbb{K}$*:*

(vii) $\alpha(\beta\mathbf{u}) = (\alpha\beta)\mathbf{u}$ *for all* $\mathbf{u} \in V$ *and all* $\alpha, \beta \in \mathbb{K}$*:*

(viii) $1\mathbf{u} = \mathbf{u}$ *for all* $\mathbf{u} \in V$.

To make it easier to distinguish, we sometimes refer to elements in V *as* vectors *and elements in* $\mathbb{K}$ *as* scalars.

I'll assume that you are familar with the basic consequences of these axioms as presented in a course on linear algebra. Recall in particular that a subset $U \subseteq V$ is a vector space if it closed under addition and scalar multiplication, i.e. that whenever $\mathbf{u}, \mathbf{v} \in U$ and $\alpha \in \mathbb{K}$, then $\mathbf{u}+\mathbf{v}, \alpha\mathbf{u} \in U$.

To measure the size of an element in a vector space, we introduce norms:

Definition 4.5.2 *If* V *is a vector space over* $\mathbb{K}$*, a* norm *on* V *is a function* $\|\cdot\| : V \to \mathbb{R}$ *such that:*

(i) $\|\mathbf{u}\| \geq 0$ *with equality if and only if* $\mathbf{u} = \mathbf{0}$.

(ii) $\|\alpha\mathbf{u}\| = |\alpha|\|\mathbf{u}\|$ *for all* $\alpha \in \mathbb{K}$ *and all* $\mathbf{u} \in V$.

(iii) $\|\mathbf{u}+\mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$ *for all* $\mathbf{u}, \mathbf{v} \in V$.

Example 1: The classical example of a norm on a real vector space, is the *euclidean norm* on $\mathbb{R}^n$ given by

$$\|\mathbf{x}\| = \sqrt{x_1^2 + x_2^2 + \cdots + x_n^2}$$

where $\mathbf{x} = (x_1, x_2. \ldots, x_n)$. The corresponding norm on the complex vector space $\mathbb{C}^n$ is

$$\|\mathbf{z}\| = \sqrt{|z_1|^2 + |z_2|^2 + \cdots + |z_n|^2}$$

where $\mathbf{z} = (z_1, z_2. \ldots, z_n)$. ♣

The spaces above are the most common vector spaces and norms in linear algebra. More relevant for our purposes in this chapter are:

Example 2: Let (X, d) be a compact metric space, and let $V = C(X, \mathbb{R})$ be the set of all continuous, real valued functions on X. Then V is a vector space over $\mathbb{R}$ and

$$\|f\| = \sup\{|f(x)| \mid x \in X\}$$

is a norm on V. To get a complex example, let $V = C(X, \mathbb{C})$ and define the norm by the same formula as before. ♣

From a norm we can always get a metric in the following way:

Proposition 4.5.3 *Assume that V is a vector space over $\mathbb{K}$ and that $\|\cdot\|$ is a norm on V. Then*

$$d(\mathbf{u},\mathbf{v}) = \|\mathbf{u}-\mathbf{v}\|$$

is a metric on V.

Proof: We have to check the three properties of a metric:
Positivity: Since $d(\mathbf{u},\mathbf{v}) = \|\mathbf{u}-\mathbf{v}\|$, we see from part (i) of the definition above that $d(\mathbf{u},\mathbf{v}) \geq 0$ with equality if and only if $\mathbf{u}-\mathbf{v} = \mathbf{0}$, i.e. if and only if $\mathbf{u} = \mathbf{v}$.
Symmetry: Since

$$\|\mathbf{u}-\mathbf{v}\| = \|(-1)(\mathbf{v}-\mathbf{u})\| = |(-1)|\|\mathbf{v}-\mathbf{u}\| = \|\mathbf{v}-\mathbf{u}\|$$

by part (ii) of the definition above, we see that $d(\mathbf{u},\mathbf{v}) = d(\mathbf{v},\mathbf{u})$.
Triangle inequality: By part (iii) of the definition above, we see that for all $\mathbf{u},\mathbf{v},\mathbf{w} \in V$:

$$d(\mathbf{u},\mathbf{v}) = \|\mathbf{u}-\mathbf{v}\| = \|(\mathbf{u}-\mathbf{w}) + (\mathbf{w}-\mathbf{v})\| \leq$$

$$\leq \|\mathbf{u}-\mathbf{w}\| + \|\mathbf{w}-\mathbf{v}\| = d(\mathbf{u},\mathbf{w}) + d(\mathbf{w},\mathbf{v})$$

□

Whenever we refer to notions such as convergence, continuity, openness, closedness, completeness, compactness etc. in a normed vector space, we shall be refering to these notions with respect to the metric defined by the norm. In practice, this means that we continue as before, but write $\|\mathbf{u}-\mathbf{v}\|$ instead of $d(\mathbf{u},\mathbf{v})$ for the distance between the points $\mathbf{u}$ and $\mathbf{v}$.

Remark: The inverse triangle inequality (recall Proposition 2.1.4)

$$|d(x,y) - d(x,z)| \leq d(y,z) \tag{4.5.1}$$

is a useful tool in metric spaces. In normed spaces, it is most conveniently expressed as

$$|\,\|\mathbf{u}\| - \|\mathbf{v}\|\,| \leq \|\mathbf{u}-\mathbf{v}\| \tag{4.5.2}$$

(use formula (4.5.1) with $x = \mathbf{0}$, $y = \mathbf{u}$ and $z = \mathbf{v}$).

Note that if $\{\mathbf{u}_n\}_{n=1}^{\infty}$ is a sequence of elements in a normed vector space, we define the infinite sum $\sum_{n=1}^{\infty} \mathbf{u}_n$ as the limit of the partial sums $\mathbf{s}_n = \sum_{k=1}^{n} \mathbf{u}_k$ provided this limit exists; i.e.

$$\sum_{n=1}^{\infty} \mathbf{u}_n = \lim_{n\to\infty} \sum_{k=1}^{n} \mathbf{u}_k$$

When the limit exists, we say that the series *converges.*

Remark: The notation $\mathbf{u} = \sum_{n=1}^{\infty} \mathbf{u}_n$ is rather treacherous — it seems to be a purely algebraic relationship, but it does, in fact, depend on which norm we are using. If we have a two different norms $\|\cdot\|_1$ and $\|\cdot\|_2$ on the same space V, we may have $\mathbf{u} = \sum_{n=1}^{\infty} \mathbf{u}_n$ with respect to $\|\cdot\|_1$, but not with respect to $\|\cdot\|_2$, as $\|\mathbf{u} - \mathbf{s}_n\|_1 \to 0$ does not necesarily imply $\|\mathbf{u} - \mathbf{s}_n\|_2 \to 0$. This phenomenon is actually quite common, and we shall meet it on several occasions later in the book.

Recall from linear algebra that at vector space V is *finite dimensional* if there is a finite set $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e_n}$ of elements in V such that each element $\mathbf{x} \in V$ can be written as a linear combination $\mathbf{x} = \alpha_1 \mathbf{e}_1 + \alpha_2 \mathbf{e}_2 + \cdots + \alpha_n \mathbf{e}_n$ in a unique way. We call $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ a *basis* for V, and say that V has dimension n. A space that is not finite dimensional is called *infinte dimensional.* Most of the spaces we shall be working with are infinite dimensional, and we shall now extend the notion of basis to (some) such spaces.

Definition 4.5.4 *Let $\{\mathbf{e}_n\}_{n=1}^{\infty}$ be a sequence of elements in a normed vector space V. We say that $\{\mathbf{e}_n\}$ is a* basis[2] *for V if for each $\mathbf{x} \in V$ there is a unique sequence $\{\alpha_n\}_{n=1}^{\infty}$ from $\mathbb{K}$ such that*

$$\mathbf{x} = \sum_{n=1}^{\infty} \alpha_n \mathbf{e}_n$$

Not all normed spaces have a basis; there are, e.g., spaces so big that not all elements can be reached from a countable set of basis elements. Let us take a look at an infinite dimensional space with a basis.

Example 3: Let c_0 be the set of all sequences $\mathbf{x} = \{x_n\}_{n\in\mathbb{N}}$ of real numbers such that $\lim_{n\to\infty} x_n = 0$. It is not hard to check that $\{c_0\}$ is a vector space and that

$$\|\mathbf{x}\| = \sup\{|x_n| : n \in \mathbb{N}\}$$

is a norm on c_0. Let $\mathbf{e}_n = (0, 0, \ldots, 0, 1, 0, \ldots)$ be the sequence that is 1 on element number n and 0 elsewhere. Then $\{\mathbf{e}_n\}_{n\in\mathbb{N}}$ is a basis for c_0 with $\mathbf{x} = \sum_{n=1}^{\infty} x_n \mathbf{e}_n$. ♣

If a normed vector space is complete, we call it a *Banach space.* The next theorem provides an efficient method for checking that a normed space

[2] Strictly speaking, there are two notions of basis for an infinite dimensional space. The type we are introducing here is sometimes called a *Schauder basis* and only works in normed spaces where we can give meaning to infinite sums. There ia another kind of basis called a *Hamel basis* which does not require the space to be normed, but which is less practical for applications.

is complete. We say that a series $\sum_{n=1}^{\infty} \mathbf{u}_n$ in V *converges absolutely* if $\sum_{n=1}^{\infty} \|\mathbf{u}_n\|$ converges (note that $\sum_{n=1}^{\infty} \|\mathbf{u}_n\|$ is a series of positive *numbers*).

Proposition 4.5.5 *A normed vector space V is complete if and only if every absolutely convergent series converges.*

Proof: Assume first that V is complete and that the series $\sum_{n=0}^{\infty} \mathbf{u}_n$ converges absolutely. We must show that the series converges in the ordinary sense. Let $S_n = \sum_{k=0}^{n} \|\mathbf{u}_k\|$ and $\mathbf{s}_n = \sum_{k=0}^{n} \mathbf{u}_k$ be the partial sums of the two series. Since the series converges absolutely, the sequence $\{S_n\}$ is a Cauchy sequence, and given an $\epsilon > 0$, there must be an $N \in \mathbb{N}$ such that $|S_n - S_m| < \epsilon$ when $n, m \geq N$. Without loss of generality, we may assume that $m > n$. By the triangle inequality

$$\|\mathbf{s}_m - \mathbf{s}_n\| = \|\sum_{k=n+1}^{m} \mathbf{u}_k\| \leq \sum_{k=n+1}^{m} \|\mathbf{u}_k\| = |S_m - S_n| < \epsilon$$

when $n, m \geq N$, and hence $\{\mathbf{s}_n\}$ is a Cauchy sequence. Since V is complete, the series $\sum_{n=0}^{\infty} \mathbf{u}_n$ converges.

For the converse, assume that all absolutely convergent series converge, and let $\{\mathbf{x}_n\}$ be a Cauchy sequence. We must show that $\{\mathbf{x}_n\}$ converges. Since $\{\mathbf{x}_n\}$ is a Cauchy sequence, we can find an increasing sequence $\{n_i\}$ in $\mathbb{N}$ such that $\|\mathbf{x}_n - \mathbf{x}_m\| < \frac{1}{2^i}$ for all $n, m \geq n_i$. In particular $\|\mathbf{x}_{n_{i+1}} - \mathbf{x}_{n_i}\| < \frac{1}{2^i}$, and clearly $\sum_{i=1}^{\infty} \|\mathbf{x}_{n_{i+1}} - \mathbf{x}_{n_i}\|$ converges. This means that the series $\sum_{i=1}^{\infty} (\mathbf{x}_{n_{i+1}} - \mathbf{x}_{n_i})$ converges absolutely, and by assumption it converges in the ordinary sense to some element $\mathbf{s} \in V$. The partial sums of this sequence are

$$\mathbf{s}_N = \sum_{i=1}^{N} (\mathbf{x}_{n_{i+1}} - \mathbf{x}_{n_i}) = \mathbf{x}_{n_{N+1}} - \mathbf{x}_{n_1}$$

(the sum is "telescoping" and almost all terms cancel), and as they converge to $\mathbf{s}$, we see that $\mathbf{x}_{n_{N+1}}$ must converge to $\mathbf{s} + \mathbf{x}_{n_1}$. This means that a subsequence of the Cauchy sequence $\{\mathbf{x}_n\}$ converges, and thus the sequence itself converges according to Lemma 2.5.5. □

Exercises for Section 4.5

1. Check that the norms in Example 1 really are norms (i.e. that they satisfy the conditions in Definition 4.5.2).

2. Check that the norms in Example 2 really are norms (i.e. that they satisfy the conditions in Definition 4.5.2).

3. Let V be a normed vector space over $\mathbb{K}$. Assume that $\{\mathbf{u}_n\}, \{\mathbf{v}_n\}$ are sequences in V converging to $\mathbf{u}$ og $\mathbf{v}$, respectively, and that $\{\alpha_n\}, \{\beta_n\}$ are sequences in $\mathbb{K}$ converging to α og β, respectively.

a) Show that $\{\mathbf{u}_n + \mathbf{v}_n\}$ converges to $\mathbf{u} + \mathbf{v}$.

b) Show that $\{\alpha_n \mathbf{u}_n\}$ converges to $\alpha\mathbf{u}$

c) Show that $\{\alpha_n \mathbf{u}_n + \beta_n \mathbf{v}_n\}$ converges to $\alpha\mathbf{u} + \beta\mathbf{v}$.

4. Let V be a normed vector space over $\mathbb{K}$.

   a) Prove the inverse triangle inequality $|\|\mathbf{u}\| - \|\mathbf{v}\|| \leq \|\mathbf{u} - \mathbf{v}\|$ for all $\mathbf{u}, \mathbf{v} \in V$.

   b) Assume that $\{\mathbf{u}_n\}$ is a sequence in V converging to $\mathbf{u}$. Show that $\{\|\mathbf{u}_n\|\}$ converges to $\|\mathbf{u}\|$

5. Show that
$$\|f\| = \int_0^1 |f(t)|\, dt$$
is a norm on $C([0,1], \mathbb{R})$.

6. Prove that the set $\{\mathbf{e}_n\}_{n\in\mathbb{N}}$ in Example 3 really is a basis for c_0.

7. Let $V \neq \{\mathbf{0}\}$ be a vector space, and let d be the discrete metric on V. Show that d is *not* generated by a norm (i.e. there is no norm on V such that $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$).

8. Let $V \neq \{\mathbf{0}\}$ be a normed vector space. Show that V is complete if and only if the unit sphere $S = \{\mathbf{x} \in V : \|\mathbf{x}\| = 1\}$ is complete.

9. Show that if a normed vector space V has a basis (as defined in Definition 4.5.4), then it is separable (i.e. it has a countable, dense subset).

10. l_1 is the set of all sequences $\mathbf{x} = \{x_n\}_{n\in\mathbb{N}}$ of real numbers such that $\sum_{n=1}^{\infty} |x_n|$ converges.

    a) Show that
$$\|\mathbf{x}\| = \sum_{n=1}^{\infty} |x_n|$$
    is a norm on l_1.

    b) Show that the set $\{\mathbf{e}_n\}_{n\in\mathbb{N}}$ in Example 3 is a basis for l_1.

    c) Show that l_1 is complete.

4.6 Inner product spaces

The usual (euclidean) norm in $\mathbb{R}^n$ can be defined in terms of the scalar (dot) product:
$$\|\mathbf{x}\| = \sqrt{\mathbf{x} \cdot \mathbf{x}}$$
This relationship is extremely important as it connects length (defined by the norm) and orthogonality (defined by the scalar product), and it is the key to many generalizations of geometric arguments from $\mathbb{R}^2$ and $\mathbb{R}^3$ to $\mathbb{R}^n$. In this section we shall see how we can extend this generalization to certain infinite dimensional spaces called inner product spaces.

The basic observation is that some norms on infinite dimensional spaces can be defined in terms of an inner product just as the euclidean norm is defined in terms of the scalar product. Let us begin by taking a look at such products. As in the previous section, we assume that all vector spaces are over $\mathbb{K}$ which is either $\mathbb{R}$ or $\mathbb{C}$. As we shall be using complex spaces in our study of Fourier series, it is important that you don't neglect the complex case.

Definition 4.6.1 *An* inner product $\langle\cdot,\cdot\rangle$ *on a vector space* V *over* $\mathbb{K}$ *is a function* $\langle\cdot,\cdot\rangle : V \times V \to \mathbb{K}$ *such that:*

(i) $\langle \mathbf{u}, \mathbf{v}\rangle = \overline{\langle \mathbf{v}, \mathbf{u}\rangle}$ *for all* $\mathbf{u}, \mathbf{v} \in V$ *(the bar denotes complex conjugation; if the vector space is real, we just have* $\langle \mathbf{u}, \mathbf{v}\rangle = \langle \mathbf{v}, \mathbf{u}\rangle$*).*

(ii) $\langle \mathbf{u}+\mathbf{v}, \mathbf{w}\rangle = \langle \mathbf{u}, \mathbf{w}\rangle + \langle \mathbf{v}, \mathbf{w}\rangle$ *for all* $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$*.*

(iii) $\langle \alpha\mathbf{u}, \mathbf{v}\rangle = \alpha\langle \mathbf{u}, \mathbf{v}\rangle$ *for all* $\alpha \in \mathbb{K}$*,* $\mathbf{u}, \mathbf{v} \in V$*.*

(iv) For all $\mathbf{u} \in V$*,* $\langle \mathbf{u}, \mathbf{u}\rangle \geq 0$ *with equality if and only if* $\mathbf{u} = \mathbf{0}$ *(by (i),* $\langle \mathbf{u}, \mathbf{u}\rangle$ *is always a real number).*[3]

As immediate consequences of (i)-(iv), we have

(v) $\langle \mathbf{u}, \mathbf{v}+\mathbf{w}\rangle = \langle \mathbf{u}, \mathbf{v}\rangle + \langle \mathbf{u}, \mathbf{w}\rangle$ for all $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$.

(vi) $\langle \mathbf{u}, \alpha\mathbf{v}\rangle = \overline{\alpha}\langle \mathbf{u}, \mathbf{v}\rangle$ for all $\alpha \in \mathbb{K}$, $\mathbf{u}, \mathbf{v} \in V$ (note the complex conjugate!).

(vii) $\langle \alpha\mathbf{u}, \alpha\mathbf{v}\rangle = |\alpha|^2\langle \mathbf{u}, \mathbf{v}\rangle$ (combine (i) and(vi) and recall that for complex numbers $|\alpha|^2 = \alpha\overline{\alpha}$).

Example 1: The classical examples are the dot products in $\mathbb{R}^n$ and $\mathbb{C}^n$. If $\mathbf{x} = (x_1, x_2, \dots, x_n)$ and $\mathbf{y} = (y_1, y_2, \dots, y_n)$ are two real vectors, we define

$$\langle \mathbf{x}, \mathbf{y}\rangle = \mathbf{x}\cdot\mathbf{y} = x_1y_1 + x_2y_2 + \ldots + x_ny_n$$

If $\mathbf{z} = (z_1, z_2, \dots, z_n)$ and $\mathbf{w} = (w_1, w_2, \dots, w_n)$ are two complex vectors, we define

$$\langle \mathbf{z}, \mathbf{w}\rangle = \mathbf{z}\cdot\mathbf{w} = z_1\overline{w_1} + z_2\overline{w_2} + \ldots + z_n\overline{w_n}$$

Before we look at the next example, we need to extend integration to complex valued functions. If $a, b \in \mathbb{R}$, $a < b$, and $f, g : [a, b] \to \mathbb{R}$ are continuous functions, we get a complex valued function $h : [a, b] \to \mathbb{C}$ by letting

$$h(t) = f(t) + i\,g(t)$$

[3]Strictly speaking, we are defining *positive definite* inner products, but they are the only inner products we have use for.

We define the integral of h in the natural way:

$$\int_a^b h(t)\,dt = \int_a^b f(t)\,dt + i\int_a^b g(t)\,dt$$

i.e., we integrate the real and complex parts separately.

Example 2: Again we look at the real and complex case separately. For the real case, let V be the set of all continuous functions $f : [a, b] \to \mathbb{R}$, and define the inner product by

$$\langle f, g\rangle = \int_a^b f(t)g(t)\,dt$$

For the complex case, let V be the set of all continuous, complex valued functions $h : [a, b] \to \mathbb{C}$ as descibed above, and define

$$\langle h, k\rangle = \int_a^b h(t)\overline{k(t)}\,dt$$

Then $\langle\cdot,\cdot\rangle$ is an inner product on V.

Note that these inner products may be thought of as natural extensions of the products in Example 1; we have just replaced discrete sums by continuous products.

Given an inner product $\langle\cdot,\cdot\rangle$, we define $\|\cdot\| : V \to [0,\infty)$ by

$$\|\mathbf{u}\| = \sqrt{\langle\mathbf{u},\mathbf{u}\rangle}$$

in analogy with the norm and the dot product in $\mathbb{R}^n$ and $\mathbb{C}^n$. For simplicity, we shall refer to $\|\cdot\|$ as a *norm*, although at this stage it is not at all clear that it is a norm in the sense of Definition 4.5.2.

On our way to proving that $\|\cdot\|$ really is a norm, we shall pick up a few results of a geometric nature that will be useful later. We begin by defining two vectors $\mathbf{u}, \mathbf{v} \in V$ to be *orthogonal* if $\langle\mathbf{u},\mathbf{v}\rangle = 0$. Note that if this is the case, we also have $\langle\mathbf{v},\mathbf{u}\rangle = 0$ since $\langle\mathbf{v},\mathbf{u}\rangle = \overline{\langle\mathbf{u},\mathbf{v}\rangle} = \overline{0} = 0$.

With these definitions, we can prove the following generalization of the Pythagorean theorem:

Proposition 4.6.2 (Pythagorean Theorem) *For all orthogonal* $\mathbf{u}_1$, $\mathbf{u}_2$, *...,* $\mathbf{u}_n$ *in* V*,*

$$\|\mathbf{u}_1 + \mathbf{u}_2 + \ldots + \mathbf{u}_n\|^2 = \|\mathbf{u}_1\|^2 + \|\mathbf{u}_2\|^2 + \ldots + \|\mathbf{u}_n\|^2$$

Proof: We have

$$\|\mathbf{u}_1 + \mathbf{u}_2 + \ldots + \mathbf{u}_n\|^2 = \langle\mathbf{u}_1 + \mathbf{u}_2 + \ldots + \mathbf{u}_n, \mathbf{u}_1 + \mathbf{u}_2 + \ldots + \mathbf{u}_n\rangle =$$

$$= \sum_{1\leq i,j\leq n} \langle \mathbf{u}_i, \mathbf{u}_j \rangle = \|\mathbf{u}_1\|^2 + \|\mathbf{u}_2\|^2 + \ldots + \|\mathbf{u}_n\|^2$$

where we have used that by orthogonality, $\langle \mathbf{u}_i, \mathbf{u}_j \rangle = 0$ whenever $i \neq j$. □

Two nonzero vectors $\mathbf{u}$, $\mathbf{v}$ are said to be *parallel* if there is a number $\alpha \in \mathbb{K}$ such that $\mathbf{u} = \alpha\mathbf{v}$. As in $\mathbb{R}^n$, the *projection* of $\mathbf{u}$ on $\mathbf{v}$ is the vector $\mathbf{p}$ parallel with $\mathbf{v}$ such that $\mathbf{u} - \mathbf{p}$ is orthogonal to $\mathbf{v}$. Figure 1 shows the idea.

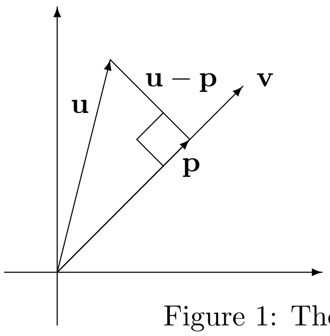


Figure 1: The projection $\mathbf{p}$ of $\mathbf{u}$ on $\mathbf{v}$

Proposition 4.6.3 *Assume that* $\mathbf{u}$ *and* $\mathbf{v}$ *are two nonzero elements of* V*. Then the projection* $\mathbf{p}$ *of* $\mathbf{u}$ *on* $\mathbf{v}$ *is given by:*

$$\mathbf{p} = \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\|\mathbf{v}\|^2}\mathbf{v}$$

The norm of the projection is $\|\mathbf{p}\| = \frac{|\langle \mathbf{u},\mathbf{v} \rangle|}{\|\mathbf{v}\|}$

Proof: Since $\mathbf{p}$ is parallel to $\mathbf{v}$, it must be of the form $\mathbf{p} = \alpha\mathbf{v}$. To determine α, we note that in order for $\mathbf{u} - \mathbf{p}$ to be orthogonal to $\mathbf{v}$, we must have $\langle \mathbf{u} - \mathbf{p}, \mathbf{v} \rangle = 0$. Hence α is determined by the equation

$$0 = \langle \mathbf{u} - \alpha\mathbf{v}, \mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle - \langle \alpha\mathbf{v}, \mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{v} \rangle - \alpha\|\mathbf{v}\|^2$$

Solving for α, we get $\alpha = \frac{\langle \mathbf{u},\mathbf{v} \rangle}{\|\mathbf{v}\|^2}$, and hence $\mathbf{p} = \frac{\langle \mathbf{u},\mathbf{v} \rangle}{\|\mathbf{v}\|^2}\mathbf{v}$.

To calculate the norm, note that

$$\|\mathbf{p}\|^2 = \langle \mathbf{p}, \mathbf{p} \rangle = \langle \alpha\mathbf{v}, \alpha\mathbf{v} \rangle = |\alpha|^2 \langle \mathbf{v}, \mathbf{v} \rangle = \frac{|\langle \mathbf{u}, \mathbf{v} \rangle|^2}{\|\mathbf{v}\|^4}\langle \mathbf{v}, \mathbf{v} \rangle = \frac{|\langle \mathbf{u}, \mathbf{v} \rangle|^2}{\|\mathbf{v}\|^2}$$

(recall property (vi) just after Definition 4.6.1). □

We can now extend Cauchy-Schwarz' inequality to general inner products:

Proposition 4.6.4 (Cauchy-Schwarz' inequality) *For all* $\mathbf{u}, \mathbf{v} \in V$*,*

$$|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\| \|\mathbf{v}\|$$

with equality if and only if $\mathbf{u}$ *and* $\mathbf{v}$ *are parallel or at least one of them is zero.*

Proof: The proposition clearly holds with equality if one of the vectors is zero. If they are both nonzero, we let $\mathbf{p}$ be the projection of $\mathbf{u}$ on $\mathbf{v}$, and note that by the pythagorean theorem

$$\|\mathbf{u}\|^2 = \|\mathbf{u} - \mathbf{p}\|^2 + \|\mathbf{p}\|^2 \geq \|\mathbf{p}\|^2$$

with equality only if $\mathbf{u} = \mathbf{p}$, i.e. when $\mathbf{u}$ and $\mathbf{v}$ are parallel. Since $\|\mathbf{p}\| = \frac{|\langle \mathbf{u}, \mathbf{v} \rangle|}{\|\mathbf{v}\|}$ by Proposition 4.6.3, we have

$$\|\mathbf{u}\|^2 \geq \frac{|\langle \mathbf{u}, \mathbf{v} \rangle|^2}{\|\mathbf{v}\|^2}$$

and the proposition follows. □

We may now prove:

Proposition 4.6.5 (Triangle inequality for inner products) *For all* $\mathbf{u}$*,* $\mathbf{v} \in V$

$$\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$$

Proof: We have (recall that $\mathrm{Re}(z)$ refers to the real part a of a complex number $z = a + ib$):

$$\|\mathbf{u} + \mathbf{v}\|^2 = \langle \mathbf{u} + \mathbf{v}, \mathbf{u} + \mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{u} \rangle + \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{v}, \mathbf{u} \rangle + \langle \mathbf{v}, \mathbf{v} \rangle =$$

$$= \langle \mathbf{u}, \mathbf{u} \rangle + \langle \mathbf{u}, \mathbf{v} \rangle + \overline{\langle \mathbf{u}, \mathbf{v} \rangle} + \langle \mathbf{v}, \mathbf{v} \rangle = \langle \mathbf{u}, \mathbf{u} \rangle + 2\mathrm{Re}(\langle \mathbf{u}, \mathbf{v} \rangle) + \langle \mathbf{v}, \mathbf{v} \rangle \leq$$

$$\leq \|\mathbf{u}\|^2 + 2\|\mathbf{u}\|\|\mathbf{v}\| + \|\mathbf{v}\|^2 = (\|\mathbf{u}\| + \|\mathbf{v}\|)^2$$

where we have used that according to Cauchy-Schwarz' inequality, we have $\mathrm{Re}(\langle \mathbf{u}, \mathbf{v} \rangle) \leq |\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\|\|\mathbf{v}\|$. □

We are now ready to prove that $\| \cdot \|$ really is a norm:

Proposition 4.6.6 *If* $\langle \cdot, \cdot \rangle$ *is an inner product on a vector space* V*, then*

$$\|\mathbf{u}\| = \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle}$$

defines a norm on V*, i.e.*

(i) $\|\mathbf{u}\| \geq 0$ *with equality if and only if* $\mathbf{u} = \mathbf{0}$*.*

(ii) $\|\alpha\mathbf{u}\| = |\alpha|\|\mathbf{u}\|$ *for all* $\alpha \in \mathbb{C}$ *and all* $\mathbf{u} \in V$*.*

(iii) $\|\mathbf{u}+\mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$ *for all* $\mathbf{u}, \mathbf{v} \in V$*.*

Proof: (i) follows directly from the definition of inner products, and (iii) is just the triangle inequality. We have actually proved (ii) on our way to Cauchy-Scharz' inequality, but let us repeat the proof here:

$$\|\alpha\mathbf{u}\|^2 = \langle \alpha\mathbf{u}, \alpha\mathbf{u}\rangle = |\alpha|^2\|\mathbf{u}\|^2$$

where we have used property (vi) just after Definition 4.6.1. □

The proposition above means that we can think of an inner product space as a metric space with metric defined by

$$d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x}-\mathbf{y}\| = \sqrt{\langle \mathbf{x}-\mathbf{y}, \mathbf{x}-\mathbf{y}\rangle}$$

Example 3: Returning to Example 2, we see that the metric in the real as well as the complex case is given by

$$d(f, g) = \left(\int_a^b |f(t)-g(t)|^2\, dt\right)^{\frac{1}{2}}$$

The next proposition tells us that we can move limits and infinite sums in and out of inner products.

Proposition 4.6.7 *Let* V *be an inner product space.*

(i) If $\{\mathbf{u}_n\}$ *is a sequence in* V *converging to* $\mathbf{u}$*, then the sequence* $\{\|\mathbf{u}_n\|\}$ *of norms converges to* $\|\mathbf{u}\|$*.*

(ii) If the series $\sum_{n=0}^{\infty} \mathbf{w}_n$ *converges in* V*, then*

$$\|\sum_{n=0}^{\infty} \mathbf{w}_n\| = \lim_{N\to\infty} \|\sum_{n=0}^{N} \mathbf{w}_n\|$$

(iii) If $\{\mathbf{u}_n\}$ *is a sequence in* V *converging to* $\mathbf{u}$*, then the sequence* $\langle \mathbf{u}_n, \mathbf{v}\rangle$ *of inner products converges to* $\langle \mathbf{u}, \mathbf{v}\rangle$ *for all* $\mathbf{v} \in V$*. In symbols,* $\lim_{n\to\infty}\langle \mathbf{u}_n, \mathbf{v}\rangle = \langle \lim_{n\to\infty} \mathbf{u}_n, \mathbf{v}\rangle$ *for all* $\mathbf{v} \in V$*.*

(iv) If the series $\sum_{n=0}^{\infty} \mathbf{w}_n$ *converges in* V*, then*

$$\langle \sum_{n=1}^{\infty} \mathbf{w}_n, \mathbf{v}\rangle = \sum_{n=1}^{\infty} \langle \mathbf{w}_n, \mathbf{v}\rangle$$

.

Proof: (i) follows directly from the inverse triangle inequality

$$|\,\|\mathbf{u}\| - \|\mathbf{u}_n\|\,| \leq \|\mathbf{u} - \mathbf{u}_n\|$$

(ii) follows immediately from (i) if we let $\mathbf{u}_n = \sum_{k=0}^{n} \mathbf{w}_k$

(iii) Assume that $\mathbf{u}_n \to \mathbf{u}$. To show that $\langle \mathbf{u}_n, \mathbf{v}\rangle \to \langle \mathbf{u}, \mathbf{v}\rangle$, is suffices to prove that $\langle \mathbf{u}_n, \mathbf{v}\rangle - \langle \mathbf{u}, \mathbf{v}\rangle = \langle \mathbf{u}_n - \mathbf{u}, \mathbf{v}\rangle \to 0$. But by Cauchy-Schwarz' inequality

$$|\langle \mathbf{u}_n - \mathbf{u}, \mathbf{v}\rangle| \leq \|\mathbf{u}_n - \mathbf{u}\|\|\mathbf{v}\| \to 0$$

since $\|\mathbf{u}_n - \mathbf{u}\| \to 0$ by assumption.

(iv) We use (iii) with $\mathbf{u} = \sum_{n=1}^{\infty} \mathbf{w}_n$ and $\mathbf{u}_n = \sum_{k=1}^{n} \mathbf{w}_k$. Then

$$\langle \sum_{n=1}^{\infty} \mathbf{w}_n, \mathbf{v}\rangle = \langle \mathbf{u}, \mathbf{v}\rangle = \lim_{n\to\infty} \langle \mathbf{u}_n, \mathbf{v}\rangle = \lim_{n\to\infty} \langle \sum_{k=1}^{n} \mathbf{w}_k, \mathbf{v}\rangle =$$

$$= \lim_{n\to\infty} \sum_{k=1}^{n} \langle \mathbf{w}_k, \mathbf{v}\rangle = \sum_{n=1}^{\infty} \langle \mathbf{w}_n, \mathbf{v}\rangle$$

□

We shall now generalize some notions from linear algebra to our new setting. If $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ is a finite set of elements in V, we define the *span*

$$\mathrm{Sp}\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$$

of $\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ to be the set of all linear combinations

$$\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u}_2 + \ldots + \alpha_n\mathbf{u}_n, \qquad \text{where } \alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{K}$$

A set $A \subseteq V$ is said to be *orthonormal* if it consists of orthogonal elements of length one, i.e. if for all $\mathbf{a}, \mathbf{b} \in A$, we have

$$\langle \mathbf{a}, \mathbf{b}\rangle = \begin{cases} 0 & \text{if } \mathbf{a} \neq \mathbf{b} \\ \\ 1 & \text{if } \mathbf{a} = \mathbf{b} \end{cases}$$

If $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ is an orthonormal set and $\mathbf{u} \in V$, we define the *projection of* $\mathbf{u}$ on $\mathrm{Sp}\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ by

$$\mathrm{P}_{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n}(\mathbf{u}) = \langle \mathbf{u}, \mathbf{e}_1\rangle\mathbf{e}_1 + \langle \mathbf{u}, \mathbf{e}_2\rangle\mathbf{e}_2 + \cdots + \langle \mathbf{u}, \mathbf{e}_n\rangle\mathbf{e}_n$$

This terminology is justified by the following result.

Proposition 4.6.8 *Let* $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ *be an orthonormal set in* V*. For every* $\mathbf{u} \in V$*, the projection* $\mathrm{P}_{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n}(\mathbf{u})$ *is the element in* $\mathrm{Sp}\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ *closest to* $\mathbf{u}$*. Moreover,* $\mathbf{u} - \mathrm{P}_{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n}(\mathbf{u})$ *is orthogonal to all elements in* $\mathrm{Sp}\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$.

Proof: We first prove the orthogonality. It suffices to prove that

$$\langle \mathbf{u} - \mathrm{P}_{\mathbf{e}_1,\mathbf{e}_2,\dots,\mathbf{e}_n}(\mathbf{u}), \mathbf{e}_i\rangle = 0 \tag{4.6.1}$$

for each $i = 1, 2, \dots, n$, as we then have

$$\langle \mathbf{u} - \mathrm{P}_{\mathbf{e}_1,\mathbf{e}_2,\dots,\mathbf{e}_n}(\mathbf{u}), \alpha_1\mathbf{e}_1 + \cdots + \alpha_n\mathbf{e}_n\rangle =$$

$$= \overline{\alpha}_1\langle \mathbf{u} - \mathrm{P}_{\mathbf{e}_1,\mathbf{e}_2,\dots,\mathbf{e}_n}(\mathbf{u}), \mathbf{e}_1\rangle + \ldots + \overline{\alpha}_n\langle \mathbf{u} - \mathrm{P}_{\mathbf{e}_1,\mathbf{e}_2,\dots,\mathbf{e}_n}(\mathbf{u}), \mathbf{e}_n\rangle = 0$$

for all $\alpha_1\mathbf{e}_1 + \cdots + \alpha_n\mathbf{e}_n \in \mathrm{Sp}\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$. To prove formula (4.6.1), just observe that for each $\mathbf{e}_i$

$$\langle \mathbf{u} - \mathrm{P}_{\mathbf{e}_1,\mathbf{e}_2,\dots,\mathbf{e}_n}(\mathbf{u}), \mathbf{e}_i\rangle = \langle \mathbf{u}, \mathbf{e}_i\rangle - \langle \mathrm{P}_{\mathbf{e}_1,\mathbf{e}_2,\dots,\mathbf{e}_n}(\mathbf{u}), \mathbf{e}_i\rangle$$

$$= \langle \mathbf{u}, \mathbf{e}_i\rangle - \big(\langle \mathbf{u}, \mathbf{e}_i\rangle\langle \mathbf{e}_1, \mathbf{e}_i\rangle + \langle \mathbf{u}, \mathbf{e}_2\rangle\langle \mathbf{e}_2, \mathbf{e}_i\rangle + \cdots + \langle \mathbf{u}, \mathbf{e}_n\rangle\langle \mathbf{e}_n, \mathbf{e}_i\rangle\big) =$$

$$= \langle \mathbf{u}, \mathbf{e}_i\rangle - \langle \mathbf{u}, \mathbf{e}_i\rangle = 0$$

To prove that the projection is the element in $\mathrm{Sp}\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$ closest to $\mathbf{u}$, let $\mathbf{w} = \alpha_1\mathbf{e}_1 + \alpha_2\mathbf{e}_2 + \cdots + \alpha_n\mathbf{e}_n$ be another element in $\mathrm{Sp}\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$. Then $\mathrm{P}_{\mathbf{e}_1,\mathbf{e}_2,\dots,\mathbf{e}_n}(\mathbf{u}) - \mathbf{w}$ is in $\mathrm{Sp}\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n\}$, and hence orthogonal to $\mathbf{u} - \mathrm{P}_{\mathbf{e}_1,\mathbf{e}_2,\dots,\mathbf{e}_n}(\mathbf{u})$ by what we have just proved. By the Pythagorean theorem

$$\|\mathbf{u}-\mathbf{w}\|^2 = \|\mathbf{u}-\mathrm{P}_{\mathbf{e}_1,\mathbf{e}_2,\dots,\mathbf{e}_n}(\mathbf{u})\|^2 + \|\mathrm{P}_{\mathbf{e}_1,\mathbf{e}_2,\dots,\mathbf{e}_n}(\mathbf{u})-\mathbf{w}\|^2 > \|\mathbf{u}-\mathrm{P}_{\mathbf{e}_1,\mathbf{e}_2,\dots,\mathbf{e}_n}(\mathbf{u})\|^2$$

□

As an immediate consequence of the proposition above, we get:

Corollary 4.6.9 (Bessel's inequality) *Let* $\{\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n, \dots\}$ *be an orthonormal sequence in* V*. For any* $\mathbf{u} \in V$*,*

$$\sum_{i=1}^{\infty} |\langle \mathbf{u}, \mathbf{e}_i\rangle|^2 \leq \|\mathbf{u}\|^2$$

Proof: Since $\mathbf{u} - \mathrm{P}_{\mathbf{e}_1,\mathbf{e}_2,\dots,\mathbf{e}_n}(\mathbf{u})$ is orthogonal to $\mathrm{P}_{\mathbf{e}_1,\mathbf{e}_2,\dots,\mathbf{e}_n}(\mathbf{u})$, we get by the Pythagorean theorem that for any n

$$\|\mathbf{u}\|^2 = \|\mathbf{u} - \mathrm{P}_{\mathbf{e}_1,\mathbf{e}_2,\dots,\mathbf{e}_n}(\mathbf{u})\|^2 + \|\mathrm{P}_{\mathbf{e}_1,\mathbf{e}_2,\dots,\mathbf{e}_n}(\mathbf{u})\|^2 \geq \|\mathrm{P}_{\mathbf{e}_1,\mathbf{e}_2,\dots,\mathbf{e}_n}(\mathbf{u})\|^2$$

Using the Pythagorean Theorem again, we see that

$$\|\mathrm{P}_{\mathbf{e}_1,\mathbf{e}_2,\dots,\mathbf{e}_n}(\mathbf{u})\|^2 = \|\langle \mathbf{u}, \mathbf{e}_1\rangle\mathbf{e}_1 + \langle \mathbf{u}, \mathbf{e}_2\rangle\mathbf{e}_2 + \cdots + \langle \mathbf{u}, \mathbf{e}_n\rangle\mathbf{e}_n\|^2 =$$

$$= \|\langle \mathbf{u}, \mathbf{e}_1\rangle\mathbf{e}_1\|^2 + \|\langle \mathbf{u}, \mathbf{e}_2\rangle\mathbf{e}_2\|^2 + \cdots + \|\langle \mathbf{u}, \mathbf{e}_n\rangle\mathbf{e}_n\|^2 =$$

$$= |\langle \mathbf{u}, \mathbf{e}_1\rangle|^2 + |\langle \mathbf{u}, \mathbf{e}_2\rangle|^2 + \cdots + |\langle \mathbf{u}, \mathbf{e}_n\rangle|^2$$

and hence

$$\|\mathbf{u}\|^2 \geq |\langle \mathbf{u}, \mathbf{e}_1 \rangle|^2 + |\langle \mathbf{u}, \mathbf{e}_2 \rangle|^2 + \cdots + |\langle \mathbf{u}, \mathbf{e}_n \rangle|^2$$

for all n. Letting $n \to \infty$, the corollary follows. □

We have now reached the main result of this section. Recall from Definition 4.5.4 that $\{\mathbf{e}_i\}$ is a *basis* for V if any element u in V can be written as a linear combination $\mathbf{u} = \sum_{i=1}^{\infty} \alpha_i \mathbf{e}_i$ in a unique way. The theorem tells us that if the basis is orthonormal, the coeffisients α_i are easy to find; they are simply given by $\alpha_i = \langle \mathbf{u}, \mathbf{e}_i \rangle$.

Theorem 4.6.10 (Parseval's Theorem) *If $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n, \ldots\}$ is an orthonormal basis for V, then for all $\mathbf{u} \in V$, we have $\mathbf{u} = \sum_{i=1}^{\infty} \langle \mathbf{u}, \mathbf{e}_i \rangle \mathbf{e}_i$ and $\|\mathbf{u}\|^2 = \sum_{i=1}^{\infty} |\langle \mathbf{u}, \mathbf{e}_i \rangle|^2$.*

Proof: Since $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n, \ldots\}$ is a basis, we know that there is a unique sequence $\alpha_1, \alpha_2, \ldots, \alpha_n, \ldots$ from $\mathbb{K}$ such that $\mathbf{u} = \sum_{n=1}^{\infty} \alpha_n \mathbf{e}_n$. This means that $\|\mathbf{u} - \sum_{n=1}^{N} \alpha_n \mathbf{e}_n\| \to 0$ as $N \to \infty$. Since the projection $\mathrm{P}_{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_N}(\mathbf{u}) = \sum_{n=1}^{N} \langle \mathbf{u}, \mathbf{e}_n \rangle \mathbf{e}_n$ is the element in $\mathrm{Sp}\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_N\}$ closest to $\mathbf{u}$, we have

$$\|\mathbf{u} - \sum_{n=1}^{N} \langle \mathbf{u}, \mathbf{e}_n \rangle \mathbf{e}_n\| \leq \|\mathbf{u} - \sum_{n=1}^{N} \alpha_n \mathbf{e}_n\| \to 0 \qquad \text{as } N \to \infty$$

and hence $\mathbf{u} = \sum_{n=1}^{\infty} \langle \mathbf{u}, \mathbf{e}_n \rangle \mathbf{e}_n$. To prove the second part, observe that since $\mathbf{u} = \sum_{n=1}^{\infty} \langle \mathbf{u}, \mathbf{e}_n \rangle \mathbf{e}_n = \lim_{N \to \infty} \sum_{n=1}^{N} \langle \mathbf{u}, \mathbf{e}_n \rangle \mathbf{e}_n$, we have (recall Proposition 4.6.7(ii))

$$\|\mathbf{u}\|^2 = \lim_{N \to \infty} \|\sum_{n=1}^{N} \langle \mathbf{u}, \mathbf{e}_n \rangle \mathbf{e}_n\|^2 = \lim_{N \to \infty} \sum_{n=1}^{N} |\langle \mathbf{u}, \mathbf{e}_n \rangle|^2 = \sum_{n=1}^{\infty} |\langle \mathbf{u}, \mathbf{e}_n \rangle|^2$$

□

The coefficients $\langle \mathbf{u}, \mathbf{e}_n \rangle$ in the arguments above are often called (abstract) *Fourier coefficients*. By Parseval's theorem, they are *square summable* in the sense that $\sum_{n=1}^{\infty} |\langle \mathbf{u}, \mathbf{e}_n \rangle|^2 < \infty$. A natural question is whether we can reverse this procedure: Given a square summable sequence $\{\alpha_n\}$ of elements in $\mathbb{K}$, does there exist an element $\mathbf{u}$ in V with Fourier coefficients α_n, i.e. such that $\langle \mathbf{u}, \mathbf{e}_n \rangle = \alpha_n$ for all n? The answer is affirmative provided V is complete.

Proposition 4.6.11 *Let V be a complete inner product space over $\mathbb{K}$ with an orthonormal basis $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n, \ldots\}$. Assume that $\{\alpha_n\}_{n \in \mathbb{N}}$ is a sequence from $\mathbb{K}$ which is square summable in the sense that $\sum_{n=1}^{\infty} |\alpha_n|^2$ converges. Then the series $\sum_{n=1}^{\infty} \alpha_n \mathbf{e}_n$ converges to an element $\mathbf{u} \in V$, and $\langle \mathbf{u}, \mathbf{e}_n \rangle = \alpha_n$ for all $n \in \mathbb{N}$.*

Proof: We must prove that the partial sums $\mathbf{s}_n = \sum_{k=1}^{n} \alpha_k \mathbf{e}_k$ form a Cauchy sequence. If $m > n$, we have

$$\|\mathbf{s}_m - \mathbf{s}_n\|^2 = \|\sum_{k=n+1}^{m} \alpha_n \mathbf{e}_n\|^2 = \sum_{k=n+1}^{m} |\alpha_n|^2$$

Since $\sum_{n=1}^{\infty} |\alpha_n|^2$ converges, we can get this expression less than any $\epsilon > 0$ by choosing n, m large enough. Hence $\{\mathbf{s}_n\}$ is a Cauchy sequence, and the series $\sum_{n=1}^{\infty} \alpha_n \mathbf{e}_n$ converges to some element $\mathbf{u} \in V$. By Proposition 4.6.7,

$$\langle \mathbf{u}, \mathbf{e}_i \rangle = \langle \sum_{n=1}^{\infty} \alpha_n \mathbf{e}_n, \mathbf{e}_i \rangle = \sum_{n=1}^{\infty} \langle \alpha_n \mathbf{e}_n, \mathbf{e}_i \rangle = \alpha_i$$

□

Completeness is necessary in the proposition above — if V is *not* complete, there will always be a square summable sequence $\{\alpha_n\}$ such that $\sum_{n=1}^{\infty} \alpha_n \mathbf{e}_n$ does *not* converge (see exercise 13).

A complete inner product space is called a *Hilbert space*.

Exercises for Section 4.6

1. Show that the inner products in Example 1 really are inner products (i.e. that they satisfy Definition 4.6.1).

2. Show that the inner products in Example 2 really are inner products.

3. Prove formula (v) just after Definition 4.6.1.

4. Prove formula (vi) just after Definition 4.6.1.

5. Prove formula (vii) just after Definition 4.6.1.

6. Show that if A is a symmetric (real) matrix with strictly positive eigenvalues, then
$$\langle \mathbf{u}, \mathbf{v} \rangle = (A\mathbf{u}) \cdot \mathbf{v}$$
is an inner product on $\mathbb{R}^n$.

7. If $h(t) = f(t) + i\, g(t)$ is a complex valued function where f and g are differentiable, define $h'(t) = f'(t) + i\, g'(t)$. Prove that the integration by parts formula
$$\int_a^b u(t)v'(t)\, dt = \Big[\ u(t)v(t)\Big]_a^b - \int_a^b u'(t)v(t)\, dt$$
holds for complex valued functions.

8. Assume that $\{\mathbf{u}_n\}$ and $\{\mathbf{v}_n\}$ are two sequences in an inner product space converging to $\mathbf{u}$ and $\mathbf{v}$, respectively. Show that $\langle \mathbf{u}_n, \mathbf{v}_n \rangle \to \langle \mathbf{u}, \mathbf{v} \rangle$.

9. Show that if the norm $\|\cdot\|$ is defined from an inner product by $\|\mathbf{u}\| = \langle \mathbf{u}, \mathbf{u} \rangle^{\frac{1}{2}}$, we have the *parallelogram law*

$$\|\mathbf{u}+\mathbf{v}\|^2 + \|\mathbf{u}-\mathbf{v}\|^2 = 2\|\mathbf{u}\|^2 + 2\|\mathbf{v}\|^2$$

for all $\mathbf{u}, \mathbf{v} \in V$. Show that the norms on $\mathbb{R}^2$ defined by $\|(x,y)\| = \max\{|x|, |y|\}$ and $\|(x,y)\| = |x| + |y|$ do not come from inner products.

10. Let $\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n\}$ be an orthonormal set in an inner product space V. Show that the projection $P = \mathrm{P}_{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n}$ is linear in the sense that $P(\alpha\mathbf{u}) = \alpha P(\mathbf{u})$ and $P(\mathbf{u}+\mathbf{v}) = P(\mathbf{u}) + P(\mathbf{v})$ for all $\mathbf{u}, \mathbf{v} \in V$ and all $\alpha \in \mathbb{K}$.

11. In this problem we prove the *polarization identities* for real and complex inner products. These identities are useful as they express the inner product in terms of the norm.

   a) Show that if V is an inner product space over $\mathbb{R}$, then

$$\langle \mathbf{u}, \mathbf{v} \rangle = \frac{1}{4}\left(\|\mathbf{u}+\mathbf{v}\|^2 - \|\mathbf{u}-\mathbf{v}\|^2\right)$$

   b) Show that if V is an inner product space over $\mathbb{C}$, then

$$\langle \mathbf{u}, \mathbf{v} \rangle = \frac{1}{4}\left(\|\mathbf{u}+\mathbf{v}\|^2 - \|\mathbf{u}-\mathbf{v}\|^2 + i\|\mathbf{u}+i\mathbf{v}\|^2 - i\|\mathbf{u}-i\mathbf{v}\|^2\right)$$

12. If S is a nonempty subset of an inner product space, let

$$S^\perp = \{\mathbf{u} \in V \,:\, \langle \mathbf{u}, \mathbf{s} \rangle = 0 \text{ for all } \mathbf{s} \in S\}$$

   a) Show that $S^\perp$ is a closed subspace of S.

   b) Show that if $S \subseteq T$, then $S^\perp \supseteq T^\perp$.

13. Let l_2 be the set of all real sequences $\mathbf{x} = \{x_n\}_{n\in\mathbb{N}}$ such that $\sum_{n=1}^\infty x_n^2 < \infty$.

   a) Show that if $\mathbf{x} = \{x_n\}_{n\in\mathbb{N}}$ and $\mathbf{y} = \{y_n\}_{n\in\mathbb{N}}$ are in l_2, then the series $\sum_{n=1}^\infty x_n y_n$ converges. (*Hint:* For each N,

$$\sum_{n=1}^N x_n y_n \leq \left(\sum_{n=1}^N x_n^2\right)^{\frac{1}{2}} \left(\sum_{n=1}^N y_n^2\right)^{\frac{1}{2}}$$

   by Cauchy-Schwarz' inequality)

   b) Show that l_2 is a vector space.

   c) Show that $\langle \mathbf{x}, \mathbf{y} \rangle = \sum_{n=1}^\infty x_n y_n$ is an inner product on l_2.

   d) Show that l_2 is complete.

   e) Let $\mathbf{e}_n$ be the sequence where the n-th component is 1 and all the other components are 0. Show that $\{\mathbf{e}_n\}_{n\in\mathbb{N}}$ is an orthonormal basis for l_2.

   f) Let V be an inner product space with an orthonormal basis $\{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_n, \ldots\}$. Assume that for every square summable sequence $\{\alpha_n\}$, there is an element $\mathbf{u} \in V$ such that $\langle \mathbf{u}, \mathbf{v}_i \rangle = \alpha_i$ for all $i \in \mathbb{N}$. Show that V is complete.

4.7 Linear operators

In linear algebra the important functions are the linear maps. The same holds for infinitely dimensional spaces, but the maps are then usually referred to as linear operators or linear transformations:

Definition 4.7.1 *Assume that V and W are two vector spaces over $\mathbb{K}$. A function $A : V \to W$ is called a* linear operator *if it satisfies:*

(i) $A(\alpha\mathbf{u}) = \alpha A(\mathbf{u})$ for all $\alpha \in \mathbb{K}$ and $\mathbf{u} \in V$.

(ii) $A(\mathbf{u} + \mathbf{v}) = A(\mathbf{u}) + A(\mathbf{v})$ for all $\mathbf{u}, \mathbf{v} \in V$.

Combining (i) and (ii), we see that

$$A(\alpha\mathbf{u} + \beta\mathbf{v}) = \alpha A(\mathbf{u}) + \beta A(\mathbf{v})$$

Using induction, this can be generalized to

$$A(\alpha_1\mathbf{u}_1 + \alpha_2\mathbf{u_2} + \cdots + \alpha_n\mathbf{u}_n) = \alpha_1 A(\mathbf{u}_1) + \alpha_2 A(\mathbf{u_2}) + \cdots + \alpha_n A(\mathbf{u}_n) \quad (4.7.1)$$

It is also useful to observe that since $A(\mathbf{0}) = A(0\mathbf{0}) = 0A(\mathbf{0}) = \mathbf{0}$, we have $A(\mathbf{0}) = \mathbf{0}$ for all linear operators.

As $\mathbb{K}$ may be regarded as a vector space over itself, the definition above covers the case where $W = \mathbb{K}$. The map is then usually referred to as a *(linear) functional.*

Example 1: Let $V = C([a, b], \mathbb{R})$ be the space of continuous functions from the interval $[a, b]$ to $\mathbb{R}$. The function $A : V \to \mathbb{R}$ defined by

$$A(u) = \int_a^b u(x)\, dx$$

is a linear functional, while the function $B : V \to V$ defined by

$$B(u)(x) = \int_a^x u(t)\, dt$$

is a linear operator. ♣

Example 2: Just as integration, differentiation is a linear operation, but as the derivative of a differentiable function is not necessarily differentiable, we have to be careful which spaces we work with. A function $f : (a, b) \to \mathbb{R}$ is said to be *infinitely differentiable* if it has derivatives of all orders at all points in (a, b), i.e. if $f^{(n)}(x)$ exists for all $n \in \mathbb{N}$ and all $x \in (a, b)$. Let U be the space of all infinitely differentiable functions, and define $D : U \to U$ by $Du(x) = u'(x)$. Then D is a linear operator. ♣

We shall mainly be interested in linear operators between normed spaces, and the following notion is of central importance:

Definition 4.7.2 *Assume that* $(V, \|\cdot\|_v)$ *and* $(W, \|\cdot\|_W)$ *are two normed spaces. A linear operator* $A : V \to W$ *is* bounded *if there is a constant* $M \in \mathbb{R}$ *such that* $\|A(\mathbf{u})\|_W \leq M\|\mathbf{u}\|_V$ *for all* $\mathbf{u} \in V$.

Remark: The terminology here is rather treacherous as a bounded operator is *not* a bounded function in the sense of, e.g., the Extremal Value Theorem. To see this, note that if $A(\mathbf{u}) \neq \mathbf{0}$, we can get $\|A(\alpha\mathbf{u})\|_W = |\alpha|\|A(\mathbf{u})\|_W$ as large as we want by increasing the size of α.

The best (i.e. smallest) value of the constant M in the definition above is denoted by $\|A\|$ and is given by

$$\|A\| = \sup\left\{\frac{\|A(\mathbf{u})\|_W}{\|\mathbf{u}\|_V} \ : \ \mathbf{u} \neq \mathbf{0}\right\}$$

An alternative formulation (see Exercise 4) is

$$\|A\| = \sup\{\|A(\mathbf{u})\|_W \ : \ \|\mathbf{u}\|_V = 1\} \tag{4.7.2}$$

We call $\|A\|$ the *operator norm* of A. The name is justified in Exercise 7.

It's instructive to take a new look at the operators in Examples 1 and 2:

Example 3: The operators A and B in Example 1 are bounded if we use the (usual) supremum norm on V. To see this for B, note that

$$|B(u)(x)| = |\int_a^x u(t)\,dt| \leq \int_a^x |u(t)|\,dt \leq \int_a^x \|u\|\,du = \|u\|(x-a) \leq \|u\|(b-a)$$

which implies that $\|B(u)\| \leq (b-a)\|u\|$ for all $u \in V$. ♣

Example 4: If we let U have the supremum norm, the operator D in Example 2 is *not* bounded. If we let $u_n = \sin nx$, we have $\|u_n\| = 1$, but $\|D(u_n)\| = \|n\cos nx\| \to \infty$ as $n \to \infty$. That D is an unbounded operator is the source of a lot of trouble, e,g. the rather unsatisfactory conditions we had to enforce in our treatment of differentiation of series in Proposition 4.2.5. ♣

We shall end this section with a brief study of the connection between boundedness and continuity. One way is easy:

Lemma 4.7.3 *A bounded linear operator A is uniformly continuous.*

Proof: If $\|A\| = 0$, A is constant zero and there is nothing to prove. If $\|A\| \neq 0$, we may for a given $\epsilon > 0$, choose $\delta = \frac{\epsilon}{\|A\|}$. For $\|\mathbf{u} - \mathbf{v}\|_V < \delta$, we then have

$$\|A(\mathbf{u}) - A(\mathbf{v})\|_W = \|A(\mathbf{u} - \mathbf{v})\|_W \leq \|A\|\|\mathbf{u} - \mathbf{v}\|_V < \|A\| \cdot \frac{\epsilon}{\|A\|} < \epsilon$$

which shows that A is uniformly continuous. $\square$

The result in the opposite direction is perhaps more surprising:

Lemma 4.7.4 *If a linear map A is continuous at $\mathbf{0}$, it is bounded.*

Proof: We argue contrapositively; i.e. we assume that A is *not* bounded and prove that A is *not* continuous at $\mathbf{0}$. Since A is not bounded, there must for each $n \in \mathbb{N}$ exist a $\mathbf{u_n}$ such that $\frac{\|A\mathbf{u}_n\|_W}{\|\mathbf{u}_n\|_V} = M_n \geq n$. If we put $\mathbf{v}_n = \frac{\mathbf{u}_n}{M_n\|\mathbf{u}_n\|_V}$, we see that $\mathbf{v}_n \to \mathbf{0}$, while $A(\mathbf{v}_n)$ does not converge to $A(\mathbf{0}) = \mathbf{0}$ since $\|A(\mathbf{v}_n)\|_W = \|A(\frac{\mathbf{u}_n}{M_n\|\mathbf{u}_n\|_V})\| = \frac{\|A(\mathbf{u}_n)\|_W}{M_n\|\mathbf{u}_n\|_V} = \frac{M_n\|\mathbf{u}_n\|_V}{M_n\|\mathbf{u}_n\|_V} = 1$. By Proposition 2.2.5, this means that A is not contiuous at $\mathbf{0}$. $\square$

Let us sum up the two lemmas in a theorem:

Theorem 4.7.5 *For linear operators $A : V \to W$ between normed spaces, the following are equivalent:*

(i) A is bounded.

(ii) A is uniformly continuous.

(iii) A is continuous at $\mathbf{0}$.

Proof: It suffices to prove (i)$\Longrightarrow$(ii)$\Longrightarrow$(iii)$\Longrightarrow$(i). As (ii)$\Longrightarrow$(iii) is obvious, we just have to observe that (i)$\Longrightarrow$(ii) by Lemma 4.7.3 and (iii)$\Longrightarrow$(i) by Lemma 4.7.4. $\square$

Exercises for Section 4.7

1. Prove Formula (4.7.1).
2. Check that the operator A in Example 1 is a linear functional and that B is a linear operator.
3. Check that the operator D in Example 2 is a linear operator.
4. Prove formula (4.7.2).
5. Define $F : C([0,1],\mathbb{R}) \to \mathbb{R}$ by $F(u) = u(0)$. Show that F is a linear functional. Is F continuous?
6. Assume that $(U, \|\cdot\|_U)$, $(V, \|\cdot\|_V)$ and $(W, \|\cdot\|_W)$ are three normed vector spaces over $\mathbb{R}$. Show that if $A : U \to V$ and $B : V \to W$ are bounded, linear operators, then $C = B \circ A$ is a bounded, linear operator. Show that $\|C\| \leq \|A\|\|B\|$ and find an example where we have strict inequality (it is possible to find simple, finite dimensional examples)
7. Assume that $(V, \|\cdot\|_V)$ and $(W, \|\cdot\|_W)$ are two normed vector spaces over $\mathbb{R}$, and let $B(V,W)$ be the set of all bounded, linear operators from V to W.

a) Show that if $A, B \in B(V, W)$, then $A + B \in B(V, W)$.

b) Show that if $A \in B(V, W)$ and $\alpha \in \mathbb{R}$, then $\alpha A \in B(V, W)$.

c) Show that $B(V, W)$ is a vector space.

d) Show that

$$\|A\| = \inf \left\{ \frac{\|A(\mathbf{u})\|_W}{\|\mathbf{u}\|_V} \ : \ \mathbf{u} \neq \mathbf{0} \right\}$$

is a norm on $B(V, W)$.

8. Assume that $(W, \|\cdot\|_W)$ is a normed vector space. Show that all linear operators $A : \mathbb{R}^d \to W$ are bounded.

9. In this problem we shall give another characterization of boundedness for functionals. We assume that V is a normed vector space over $\mathbb{K}$ and let $A : V \to \mathbb{K}$ be a linear functional. The *kernel* of A is defined by

$$\ker(A) = \{\mathbf{v} \in V \ : \ A(\mathbf{v}) = \mathbf{0}\} = A^{-1}(\{\mathbf{0}\})$$

a) Show that if A is bounded, $\ker(A)$ is closed. (*Hint:* Recall Proposition 2.3.10)

We shall use the rest of the problem to prove the converse: If $\ker A$ is closed, then A is bounded. As this is obvious when A is identically zero, we may assume that there is an element $\mathbf{a}$ in $\ker(A)^c$. Let $\mathbf{b} = \frac{\mathbf{a}}{A(\mathbf{a})}$ (since $A(a)$ is a number, this makes sense).

b) Show that $A(\mathbf{b}) = 1$ and that there is a ball $\mathrm{B}(\mathbf{b}; r)$ around $\mathbf{b}$ contained in $\ker A^c$.

c) Show that if $\mathbf{u} \in \mathrm{B}(\mathbf{0}; r)$ (where r is as in b) above), then $\|A(\mathbf{u})\|_W \leq 1$. (*Hint:* Assume for contradiction that $\mathbf{u} \in \mathrm{B}(\mathbf{0}, r)$, but $\|A(\mathbf{u})\|_W > 1$, and show that $A(\mathbf{b} - \frac{\mathbf{u}}{A(\mathbf{u})}) = 0$ although $\mathbf{b} - \frac{\mathbf{u}}{A(\mathbf{u})} \in \mathrm{B}(\mathbf{b}; r)$.)

d) Use a) and c) to prove:

Teorem: *Assume that $(V, \|\cdot\|_V)$ is a normed spaces over $\mathbb{K}$. A linear functional $A : V \to \mathbb{K}$ is bounded if and only if $\ker(A)$ is closed.*

10. Let $(V, \langle \cdot, \cdot \rangle)$ be a complete inner product space over $\mathbb{R}$ with an orthonormal basis $\{\mathbf{e}_n\}$.

a) Show that for each $\mathbf{y} \in V$, the map $B(\mathbf{x}) = \langle \mathbf{x}, \mathbf{y} \rangle$ is a bounded linear functional.

b) Assume now that $A : V \to \mathbb{R}$ is a bounded linear functional, and let $\beta_n = A(\mathbf{e}_n)$. Show that $A(\sum_{i=1}^n \beta_i \mathbf{e}_i) = \sum_{i=1}^n \beta_i^2$ and conclude that $\left(\sum_{i=1}^\infty \beta_i^2\right)^{\frac{1}{2}} \leq \|A\|$.

c) Show that the series $\sum_{i=1}^\infty \beta_i \mathbf{e}_i$ converges in V.

d) Let $\mathbf{y} = \sum_{i=1}^\infty \beta_i \mathbf{e}_i$. Show that $A(\mathbf{x}) = \langle \mathbf{x}, \mathbf{y} \rangle$ for all $\mathbf{x} \in V$, and that $\|A\| = \|\mathbf{y}\|_V$. (*Note:* This is a special case of the *Riesz-Fréchet Representation Theorem* which says that all linear functionals A on a Hilbert space H is of the form $A(\mathbf{x}) = \langle \mathbf{x}, \mathbf{y} \rangle$ for some $\mathbf{y} \in H$. The assumption that V has an orthonormal basis is not needed for the theorem to be true).

11. Assume that $(V, \|\cdot\|_V)$ and $(W, \|\cdot\|_W)$ are two normed vector spaces over $\mathbb{R}$, and let $\{A_n\}$ be a sequence of bounded, linear operators from V to W. Assume that $\lim_{n\to\infty} A_n(\mathbf{v})$ exists for all $\mathbf{v} \in V$, and define $A(\mathbf{v}) = \lim_{n\to\infty} A_n(\mathbf{v})$.

    a) Show that A is a linear operator.

    b) Assume from now on that U is complete and show that there is a closed ball $\overline{\mathrm{B}}(\mathbf{a}; r)$, $r > 0$, and a constant $M \in \mathbb{R}$ such that $\|A_n(\mathbf{u})\|_W \leq M$ for all $u \in \overline{\mathrm{B}}(\mathbf{a}; r)$ and all $n \in \mathbb{N}$. (*Hint:* Use Proposition 3.8.7).

    c) Show that there is a number $K \in \mathbb{R}$ such that $\|A_n(\mathbf{u})\|_W \leq K\|\mathbf{u}\|_V$ for all $u \in V$ and all $n \in \mathbb{N}$. (*Hint:* $\mathbf{a} + \frac{\mathbf{r}}{\|u\|_U}\mathbf{u} \in \overline{\mathrm{B}}(\mathbf{a}; r)$ for all nonzero $\mathbf{u} \in V$).

    d) Show that the linear operator A is bounded. (*Note:* This result is often referred to as the *Banach-Steinhaus Theorem.*)

4.8 Complex exponential functions

Our next task is to apply the ideas in the previous sections to spaces of functions. In particular, we shall see how the abstract Fourier analysis of inner product spaces in Section 4.6, can be turned into concrete Fourier analysis of functions on the real line. Before we do so, it will be convenient to take a brief look at the functions that will serve as elements of our orthonormal basis. Recall that for a complex number $z = x + iy$, the exponential e^z is defined by

$$e^z = e^x(\cos y + i \sin y)$$

We shall mainly be interested in purely imaginary exponents:

$$e^{iy} = \cos y + i \sin y \tag{4.8.1}$$

Since we also have

$$e^{-iy} = \cos(-y) + i \sin(-y) = \cos y - i \sin y$$

we may add and subtract to get

$$\cos y = \frac{e^{iy} + e^{-iy}}{2} \tag{4.8.2}$$

$$\sin y = \frac{e^{iy} - e^{-iy}}{2i} \tag{4.8.3}$$

Formulas (4.8.1)-(4.8.3) give us important connections between complex exponetials and trigonometric functions that we shall exploit in the next sections.

We shall be interested in functions $f : \mathbb{R} \to \mathbb{C}$ of the form

$$f(x) = e^{(a+ib)x} = e^{ax} \cos bx + i e^{ax} \sin bx, \qquad \text{where } a \in \mathbb{R}$$

If we differentiate f by differentiating the real and complex parts separately, we get

$$f'(x) = ae^{ax}\cos bx - be^{ax}\sin bx + iae^{ax}\sin bx + ibe^{ax}\cos bx =$$

$$= ae^{ax}(\cos bx + i\sin bx) + ibe^{ax}(\cos bx + i\sin bx) = (a+ib)e^{(a+ib)x}$$

and hence we have the formula

$$\left(e^{(a+ib)x}\right)' = (a+ib)e^{(a+ib)x} \tag{4.8.4}$$

that we would expect from the real case. Antidifferentiating, we see that

$$\int e^{(a+ib)x}\,dx = \frac{e^{(a+ib)x}}{a+ib} + C \tag{4.8.5}$$

where $C = C_1 + iC_2$ is an arbitrary, complex constant.

Note that if we multiply by the conjugate $a - ib$ in the numerator and the denominator, we get

$$\frac{e^{(a+ib)x}}{a+ib} = \frac{e^{(a+ib)x}(a-ib)}{(a+ib)(a-ib)} = \frac{e^{ax}}{a^2+b^2}(\cos bx + i\sin bx)(a-ib) =$$

$$= \frac{e^{ax}}{a^2+b^2}\big(a\cos bx + b\sin bx + i(a\sin bx - b\cos bx)\big)$$

Hence (4.8.5) may also be written

$$\int \big(e^{ax}\cos bx \quad + \quad ie^{ax}\sin bx\big)\,dx =$$

$$= \quad \frac{e^{ax}}{a^2+b^2}\big(a\cos bx + b\sin bx + i(a\sin bx - b\cos bx)\big)$$

Separating the real and the imaginary parts, we get

$$\int e^{ax}\cos bx\,dx = \frac{e^{ax}}{a^2+b^2}\big(a\cos bx + b\sin bx\big) \tag{4.8.6}$$

and

$$\int e^{ax}\sin bx\,dx = \frac{e^{ax}}{a^2+b^2}\big(a\sin bx - b\cos bx\big) \tag{4.8.7}$$

In calculus, these formulas are usually proved by two times integration by parts, but in our complex setting they follow more or less immediately from the basic integration formula (4.8.5).

We shall be particularly interested in the functions

$$e_n(x) = e^{inx} = \cos nx + i\sin nx \quad \text{where } n \in \mathbb{Z}$$

Observe first that these functions are 2π-periodic in the sense that

$$e_n(x+2\pi) = e^{in(x+2\pi)} = e^{inx}e^{2n\pi i} = e^{inx}\cdot 1 = e_n(x)$$

This means in particular that $e_n(-\pi) = e_n(\pi)$ (they are both equal to $(-1)^n$ as is easily checked). Integrating, we see that for $n \neq 0$, we have

$$\int_{-\pi}^{\pi} e_n(x)\,dx = \left[\frac{e^{inx}}{in}\right]_{-\pi}^{\pi} = \frac{e_n(\pi) - e_n(-\pi)}{in} = 0$$

while we for $n = 0$ have

$$\int_{-\pi}^{\pi} e_0(x)\,dx = \int_{-\pi}^{\pi} 1\,dx = 2\pi$$

This leads to the following orthogonality relation.

Proposition 4.8.1 *For all $n, m \in \mathbb{Z}$ we have*

$$\int_{-\pi}^{\pi} e_n(x)\overline{e_m(x)}\,dx = \begin{cases} 0 & \text{if } n \neq m \\ \\ 2\pi & \text{if } n = m \end{cases}$$

Proof: Since

$$e_n(x)\overline{e_m(x)} = e^{inx}e^{-imx} = e^{i(n-m)x}$$

the lemma follows from the formulas above. □

The proposition shows that the family $\{e_n\}_{n\in\mathbb{Z}}$ is almost orthonormal with respect to the inner product

$$\langle f, g\rangle = \int_{-\pi}^{\pi} f(x)\overline{g(x)}\,dx.$$

The only problem is that $\langle e_n, e_n\rangle$ is 2π and not 1. We could fix this by replacing e_n by $\frac{e_n}{\sqrt{2\pi}}$, but instead we shall choose to change the inner product to

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

Abusing terminology slightly, we shall refer to this at the L_2-*inner product on* $[-\pi, \pi]$. The norm it induces will be called the L_2-*norm.*

The Fourier coefficients of a function f with respect to $\{e_n\}_{n\in\mathbb{Z}}$ are defined by

$$\langle f, e_n\rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\overline{e_n(x)}\,dx = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-inx}\,dx$$

From the previous section we know that $f = \sum_{n=-\infty}^{\infty}\langle f, e_n\rangle e_n$ (where the series converges in L_2-norm) provided f belongs to a space where $\{e_n\}_{n\in\mathbb{Z}}$ is a basis. We shall study this question in detail in the next sections. For the time being, we look at an example of how to compute Fourier coefficients.

Example 1: We shall compute the Fourier coefficients α_n of the function $f(x) = x$. By definition

$$\alpha_n = \langle f, e_n \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} x e^{-inx}\, dx$$

It is easy to check that $\alpha_0 = \int_{-\pi}^{\pi} x\, dx = 0$. For $n \neq 0$, we use integration by parts (see Exercise 4.8.7) with $u = x$ and $v' = e^{-inx}$. We get $u' = 1$ and $v = \frac{e^{-inx}}{-in}$, and:

$$\alpha_n = -\frac{1}{2\pi} \left[x \frac{e^{-inx}}{in} \right]_{-\pi}^{\pi} + \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{e^{-inx}}{in}\, dx =$$

$$= \frac{(-1)^{n+1}}{in} - \frac{1}{2\pi} \left[\frac{e^{-inx}}{n^2} \right]_{-\pi}^{\pi} = \frac{(-1)^{n+1}}{in}$$

The Fourier series becomes

$$\sum_{n=-\infty}^{\infty} \alpha_n e_n = \sum_{n=-\infty}^{-1} \frac{(-1)^{n+1}}{in} e^{inx} + \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{in} e^{inx} =$$

$$= \sum_{n=1}^{\infty} \frac{2(-1)^{n+1}}{n} \sin(nx)$$

We would like to conclude that $x = \sum_{n=1}^{\infty} \frac{2(-1)^{n+1}}{n} \sin(nx)$ for $x \in (-\pi, \pi)$, but we don't have the theory to take that step yet.

Exercises for Section 4.8

1. Show that $\langle f, g \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x)\overline{g(x)}\, dx$ is an inner product on $C([-\pi, \pi], \mathbb{C})$.

2. Deduce the formulas for $\sin(x + y)$ and $\cos(x + y)$ from the rule $e^{i(x+y)} = e^{ix}e^{iy}$.

3. In this problem we shall use complex exponentials to prove some trigonometric identities.

   a) Use the complex expressions for sin and cos to show that

   $$\sin(u)\sin(v) = \frac{1}{2}\cos(u - v) - \frac{1}{2}\cos(u + v)$$

   b) Integrate $\int \sin 4x \sin x\, dx$.

   c) Find a similar expression for $\cos u \cos v$ and use it to compute the integral $\int \cos 3x \cos 2x\, dx$.

   d) Find an expression for $\sin u \cos v$ and use it to integrate $\int \sin x \cos 4x\, dx$.

4. Find the Fourier series of $f(x) = e^x$.

5. Find the Fourier series of $f(x) = x^2$.

6. Find the Fourier sries of $f(x) = \sin \frac{x}{2}$.

7. a) Let $s_n = a_0 + a_0 r + a_0 r^2 + \cdots + a_0 r^n$ be a geometric series of complex numbers. Show that if $r \neq 1$, then

$$s_n = \frac{a_0(1 - r^{n+1})}{1 - r}$$

(*Hint:* Subtract rs_n from s_n.)

b) Explain that $\sum_{k=0}^{n} e^{ikx} = \frac{1-e^{i(n+1)x}}{1-e^{ix}}$ when x is not a multiplum of 2π.

c) Show that $\sum_{k=0}^{n} e^{ikx} = e^{i\frac{nx}{2}} \frac{\sin(\frac{n+1}{2}x)}{\sin(\frac{x}{2})}$ when x is not a multiplum of 2π.

d) Use the result in c) to find formulas for $\sum_{k=0}^{n} \cos(kx)$ and $\sum_{k=0}^{n} \sin(kx)$.

8. Show that the integration by parts formula

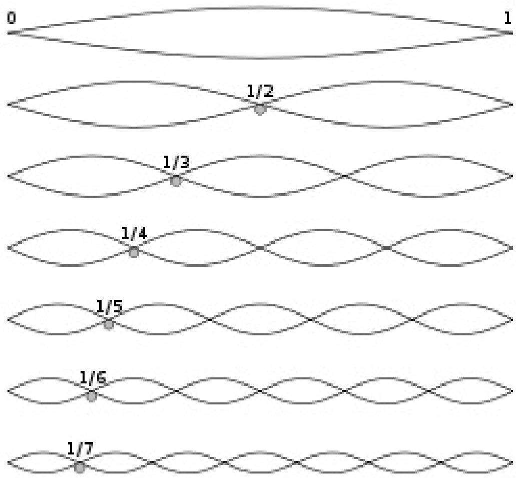
$$\int f(x)g'(x)\,dx = f(x)g(x) - \int f'(x)g(x)\,dx$$

holds for complex valued functions f, g.

4.9 Fourier series

In the middle of the 18th century, mathematicians and physicists started to study the motion of a vibrating string (think of the strings of a violin or a guitar). If you pull the string out and then let it go, how will it vibrate? To make a mathematical model, assume that at rest the string is stretched along the x-axis from 0 to 1 and fastened at both ends.

The figure below shows some possibilities. If we start with a simple sine curve $f_1(x) = C_1 \sin(\pi x)$, the string will oscillate up an down between the two curves shown in the top line of the picture (we are neglecting air resistance and other frictional forces). The frequency of the oscillation is called the *fundamental harmonic* of the string. If we start from a position where the string is pinched in the middle as on the second line of the figure (i.e. we use a starting position of the form $f_2(x) = C_2 \sin(2\pi x)$), the curve will oscillate with a node in the middle. The frequency will be twice the fundamental harmonic. This is the first overtone of the string. Pinching the string at more and more ponts (i.e. using starting positions of the form $f_n(x) = C_n \sin(n\pi x)$ for bigger and bigger integers n), we introduce more and more nodes and more and more overtones (the frequency of f_n will be n times the fundamental harmonic). If the string is vibrating in air, the frequencies (the fundamental harmonic and its overtones) can be heard as tones of different pitch.



Imagine now that we start with a mixture

$$f(x) = \sum_{n=1}^{\infty} C_n \sin(n\pi x) \tag{4.9.1}$$

of the starting positions above. The motion of the string will now be a superposition of the motions created by each individual f_n. The sound produced will be a mixture of the fundamental harmonic and the different overtones, and the size of the constant C_n will determine how much overtone number n contributes to the sound.

This is a nice description, but the problem is that a function is usually not of the form (4.9.1). Or – perhaps it is? Perhaps any reasonable starting position for the string can be written in the form (4.9.1)? But if so, how do we prove it, and how do we find the coefficients C_n? There was a heated discussion on these questions around 1750, but nobody at the time was able to come up with a satisfactory solution.

The solution came with a memoir published by Joseph Fourier in 1807. To understand Fourier's solution, we need to generalize the situation a little. Since the string is fastened at both ends of the interval, a starting position for the string must always satisfy $f(0) = f(1) = 0$. Fourier realized that if he were to include general functions that did not satisfy these boundary conditions in his theory, he needed to allow constant terms and cosine functions in his series. Hence he looked for representations of the form

$$f(x) = A + \sum_{n=1}^{\infty} \big(C_n \sin(n\pi x) + D_n \cos(n\pi x)\big) \tag{4.9.2}$$

with $A, C_n, D_n \in \mathbb{R}$. The big breakthrough was that Fourier managed to find simple formulas to compute the coefficients A, C_n, D_n of this series. This turned trigonometric series into a useful tool in applications (Fourier himself was mainly interested in heat propagation).

When we now begin to develop the theory, we shall change the setting slightly. We shall replace the interval $[0,1]$ by $[-\pi,\pi]$ (it is easy to go from one interval to another by scaling the functions, and $[-\pi,\pi]$ has certain notational advantages), and we shall replace sin and cos by complex exponentials e^{inx}. Not only does this reduce the types of functions we have to work with from two to one, but it also makes many of our arguments easier and more transparent. The formulas in the previous section makes it easy to get back to the cos/sin-setting when one needs to.

Recall from the previous section that the functions

$$e_n(x) = e^{inx}, \quad n \in \mathbb{Z}$$

form an orthonormal set with respect to the L_2-inner product

$$\langle f, g\rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\overline{g(x)}\,dx$$

The Fourier coefficients of a continuous function $f : [-\pi,\pi] \to \mathbb{C}$ with respect to this set are given by

$$\alpha_n = \langle f, e_n\rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\overline{e_n(x)}\,dx$$

From Parseval's theorem, we know that if $\{e_n\}$ is a basis, then

$$f(x) = \sum_{n=-\infty}^{\infty} \alpha_n e_n(x)$$

where the series converges in the L_2-norm, i.e.

$$\lim_{N\to\infty} \|f - \sum_{n=-N}^{N} \alpha_n e_n\|_2 = 0$$

where $\|\cdot\|_2$ denotes the norm induced by the L_2-inner product (we shall refer to it as the L_2-norm).

At this stage, life becomes complicated in two ways. First, we don't know yet that $\{e_n\}_{n\in\mathbb{Z}}$ is a basis for $C([-\pi,\pi],\mathbb{C})$, and second, we don't really know what L_2-convergence means. It turns out that L_2-convergence is quite weak, and that a sequence may converge in L_2-norm without actually converging at any point! This means that we would also like to investigate other forms for convergence (pointwise, uniform etc.).

Let us begin by observing that since $e_n(-\pi) = e_n(\pi)$ for all $n \in \mathbb{Z}$, any function that is the pointwise limit of a series $\sum_{n=-\infty}^{\infty} \alpha_n e_n$ must also satisfy this periodicity assumption. Hence it is natural to introduce the following class of functions:

Definition 4.9.1 *Let C_P be the set of all continuous functions $f : [-\pi, \pi] \to \mathbb{C}$ such that $f(-\pi) = f(\pi)$. A function in C_P is called a* trigonometric polynomial *if it is of the form $\sum_{n=-N}^{N} \alpha_n e_n$ where $N \in \mathbb{N}$ and each $\alpha_n \in \mathbb{C}$.*

To distinguish it from the L_2-norm, we shall denote the supremum norm on $C([-\pi, \pi], \mathbb{C})$ by $\|\cdot\|_\infty$, i.e.

$$\|f\|_\infty = \sup\{|f(x) : x \in [-\pi.\pi]\}$$

Note that the metric generated by $\|\cdot\|_\infty$ is the metric ρ that we studied in Chapter 3. Hence convergence with respect to $\|\cdot\|_\infty$ is the same as uniform convergence.

Theorem 4.9.2 *The trigonometric polynomials are dense in C_P in the $\|\cdot\|_\infty$-norm. Hence for any $f \in C_P$ there is a sequence $\{p_n\}$ of trigonometric polynomials which converges uniformly to f.*

It is possible to prove this result from Weierstrass' Approximation Theorem 3.7.1, but the proof is technical and not very informative. In the next section, we shall get a more informative proof from ideas we have to develop anyhow, and we postpone the proof till then. In the meantime we look at some consequences.

Corollary 4.9.3 *For all $f \in C_P$, the Fourier series $\sum_{n=-\infty}^{\infty} \langle f, e_n \rangle e_n$ converges to f in L_2-norm, i.e. $\lim_{n\to\infty} \|f - \sum_{n=-N}^{N} \langle f, e_n \rangle e_n\|_2 = 0$.*

Proof: Given $\epsilon > 0$, we must show that there is an $N \in \mathbb{N}$ such that $\|f - \sum_{n=-M}^{M} \langle f, e_n \rangle e_n\|_2 < \epsilon$ when $M \geq N$. According to the theorem above, there is a trigonometric polynomial $p(x) = \sum_{n=-N}^{N} \alpha_n e_n$ such that $\|f - p\|_\infty < \epsilon$. Hence

$$\|f - p\|_2 = \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} |f(x) - p(x)|^2 \, dx \right)^{\frac{1}{2}} < \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} \epsilon^2 \, dx \right)^{\frac{1}{2}} = \epsilon$$

According to Proposition 4.6.8, $\|f - \sum_{n=-M}^{M} \langle f, e_n \rangle\|_2 \leq \|f - p\|_2$ for all $M \geq N$, and the corollary follows. □

The corollary above is rather unsatisfactory. It is particularly inconvenient that it only applies to periodic functions such that $f(-\pi) = f(\pi)$ (although we can not have *pointwise convergence* to functions violating this condition, we may well have L_2-convergence as we soon shall see). To get a better result, we introduce a bigger space D of piecewise continuous functions.

Definition 4.9.4 *A function* $f : [-\pi, \pi] \to \mathbb{C}$ *is said to be* piecewise continuous with one sided limits *if there exists a finite set of points*

$$-\pi = a_0 < a_1 < a_2 < \ldots < a_{n-1} < a_n = \pi$$

such that:

(i) f *is continuous on each interval* (a_i, a_{i+1}).

(ii) f *have one sided limits at each point* a_i, *i.e.* $f(a_i^-) = \lim_{x\uparrow a_i} f(x)$ *and* $f(a_i^+) = \lim_{x\downarrow a_i} f(x)$ *both exist, but need not be equal (at the endpoints* $a_0 = -\pi$ *and* $a_n = \pi$ *we do, of course, only require limits from the appropriate side).*

(iii) The value of f *at each jump point* a_i *is the avarage of the one-sided limits, i.e.* $f(a_i) = \frac{1}{2}(f(a_i^-) + f(a_i^+))$. *At the endpoints, this is interpreted as* $f(a_0) = f(a_n) = \frac{1}{2}(f(a_n^-) + f(a_0^+))$

The collection of all such functions will be denoted by D.

Remark: Part (iii) is only included for technical reasons (we must specify the values at the jump points to make D an inner product space), but it reflects how Fourier series behave — at jump points they always choose the average value. The treatment of the end points may seem particularly strange; why should we enforce the average rule even here? The reason is that since the trigonometric polynomials are 2π-periodic, they regard 0 and 2π as the "same" point, and hence it is natural to compare the right limit at 0 to the left limit at 2π.

Note that the functions in D are bounded and integrable, that the sum and product of two functions in D are also in D, and that D is a inner product space over $\mathbb{C}$ with the L_2-inner product. The next lemma will allow us to extend the corollary above to D.

Lemma 4.9.5 C_P *is dense in* D *in the* L_2*-norm, i.e. for each* $f \in D$ *and each* $\epsilon > 0$, *there is a* $g \in C_P$ *such that* $\|f - g\|_2 < \epsilon$.

Proof: I only sketch the main idea of the proof, leaving the details to the reader. Assume that $f \in D$ and $\epsilon > 0$ are given. To construct g, choose a very small $\delta > 0$ (it is your task to figure out how small) and construct g as follows: Outside the (nonoverlapping) intervals $(a_i - \delta, a_i + \delta)$, we let g agree with f, but in each of these intervals, g follows the straight line connecting the points $(a_i - \delta, f(a_i - \delta))$ and $(a_i + \delta, f(a_i + \delta))$ on f's graph. Check that if we choose δ small enough, $\|f - g\|_2 < \epsilon$ (In making your choice, you have to take $M = \sup\{|f(x)| : x \in [-\pi, \pi]\}$ into account, and you also have to figure ut what to do at the endpoints $-\pi, \pi$ of the interval). □

We can now extend the corollary above from C_P to D.

Theorem 4.9.6 *For all $f \in D$, the Fourier series $\sum_{n=-\infty}^{\infty}\langle f, e_n\rangle e_n$ converges to f in L_2-norm, i.e.* $\lim_{n\to\infty}\|f - \sum_{n=-N}^{N}\langle f, e_n\rangle e_n\|_2 = 0$.

Proof: Assume that $f \in D$ and $\epsilon > 0$ are given. By the lemma, we know that there is a $g \in C_P$ such that $\|f - g\|_2 < \frac{\epsilon}{2}$, and by the corollary above, there is a trigonometric polynomial $p = \sum_{n=-N}^{N}\alpha_n e_n$ such that $\|g - p\|_\infty < \frac{\epsilon}{2}$. By the same argument as in the proof of the corollary, we get $\|g - p\|_2 < \frac{\epsilon}{2}$. The triangle inequality now tells us that

$$\|f - p\|_2 \leq \|f - g\|_2 + \|g - p\|_2 < \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon$$

Invoking Proposition 4.6.8 again, we see that for $M \geq N$, we have

$$\|f - \sum_{n=-M}^{M}\langle f, \mathbf{e}_n\rangle e_n\|_2 \leq \|f - p\|_2 < \epsilon$$

and the theorem is proved. □

The theorem above is satisfactory in the sense that we know that the Fourier series of f converges to f for a reasonably wide class of functions. However, we still have things to attend to: We haven't proved Theorem 4.9.2 yet, and we would really like to prove that Fourier series converge pointwise (or even uniformly) for a reasonable class of functions. We shall take a closer look at these questions in the next sections.

Exercises for Section 4.9

1. Show that C_P is a closed subset of $C([-\pi, \pi], \mathbb{C})$
2. In this problem we shall prove some properties of the space D.
   a) Show that if $f, g \in D$, then $f + g, fg \in D$.
   b) Show that D is a vector space.
   c) Show that all functions in D are bounded.
   d) Show that all functions in D are integrable on $[-\pi, \pi]$.
   e) Show that $\langle f, g\rangle = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\overline{g(x)}\,dx$ is an inner product on D.
3. In this problem we shall show that if $f : [-\pi, \pi] \to \mathbb{R}$ is a *realvalued* function, then the Fourier series $\sum_{n=-\infty}^{\infty}\alpha_n e_n$ can be turned into a sine/cosine-series of the form (4.9.2).
   a) Show that if $\alpha_n = a_n + ib_n$ are Fourier coefficients of f, then $\alpha_{-n} = \overline{\alpha}_n = a_n - ib_n$.
   b) Show that $a_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\cos(nx)\,dx$ and $b_n = -\frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)\sin(nx)\,dx$.
   c) Show that the Fourier series can be written
   $$\alpha_0 + \sum_{n=0}^{\infty}\big(2a_n\cos(nx) - 2b_n\sin(nx)\big)$$
4. Complete the proof of Lemma 4.9.5.

4.10 The Dirichlet kernel

Our arguments so far have been entirely abstract — we have not really used any properties of the functions $e_n(x) = e^{inx}$ except that they are orthonormal. To get better results, we need to take a closer look at these functions. In some of our arguments, we shall need to change variables in integrals, and such changes may take us outside our basic interval $[-\pi, \pi]$, and hence outside the region where our functions are defined. To avoid these problems, we extend our functions $f \in D$ periodically outside the basic interval such that $f(x + 2\pi) = f(x)$ for all $x \in \mathbb{R}$. The figure shows the extension graphically; in part a) we have the original function, and in part b) (a part of) the periodic extension. As there is no danger of confusion, we shall denote the original function and the extension by the same symbol f.

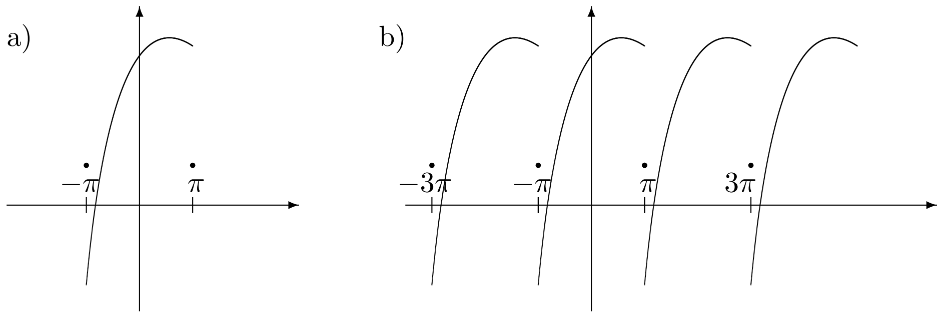


Figure 1

Remark: Here is a way of thinking that is often useful: Assume that we take our interval $[-\pi, \pi]$ and bend it into a circle such that the points $-\pi$ and π become the same. If we think of our trigonometric polynomials p as being defined on the circle instead of on the interval $[-\pi, \pi]$, it becomes quite logical that $p(-\pi) = p(\pi)$. When we are extending functions $f \in D$ the way we did above, we can imagine that we are wrapping the entire real line up around the circle such that the the points x and $x + 2\pi$ on the real line always become the same point on the circle. Mathematicians often say they are "doing Fourier analysis on the unit circle".

Let us begin by looking at the partial sums

$$s_N(x) = \sum_{n=-N}^{N} \langle f, e_n \rangle e_n(x)$$

of the Fourier series. Since

$$\alpha_n = \langle f, e_n \rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) e^{-int} \, dt$$

we have

$$s_N(x) = \frac{1}{2\pi} \sum_{n=-N}^{N} \left(\int_{-\pi}^{\pi} f(t) e^{-int}\, dt \right) e^{inx} = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \sum_{n=-N}^{N} e^{in(x-t)}\, dt =$$

$$= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x-u) \sum_{n=-N}^{N} e^{inu}\, du$$

where we in the last step has substituted $u = x - t$ and used the periodicity of the functions to remain in the interval $[-\pi, \pi]$. If we introduce the *Dirichlet kernel*

$$D_N(u) = \sum_{n=-N}^{N} e^{inu}$$

we may write this as

$$s_N(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x-u) D_N(u)\, du$$

Note that the sum $\sum_{n=-N}^{N} e^{inu} = \sum_{n=-N}^{N} (e^{iu})^n$ is a geomtric series. For $u = 0$, all the terms are 1 and the sum is $2N+1$. For $u \neq 0$, we use the sum formula for a finite geometric series to get:

$$D_N(u) = \frac{e^{-iNu} - e^{i(N+1)u}}{1 - e^{iu}} = \frac{e^{-i(N+\frac{1}{2})u} - e^{i(N+\frac{1}{2})u}}{e^{-i\frac{u}{2}} - e^{i\frac{u}{2}}} = \frac{\sin((N+\frac{1}{2})u)}{\sin\frac{u}{2}}$$

where we have used the formula $\sin x = \frac{e^{ix} - e^{-ix}}{2i}$ twice in the last step. This formula gives us a nice, compact expression for $D_N(u)$. If we substitute it into the formula above, we get

$$s_N(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x-u) \frac{\sin((N+\frac{1}{2})u)}{\sin\frac{u}{2}}\, du$$

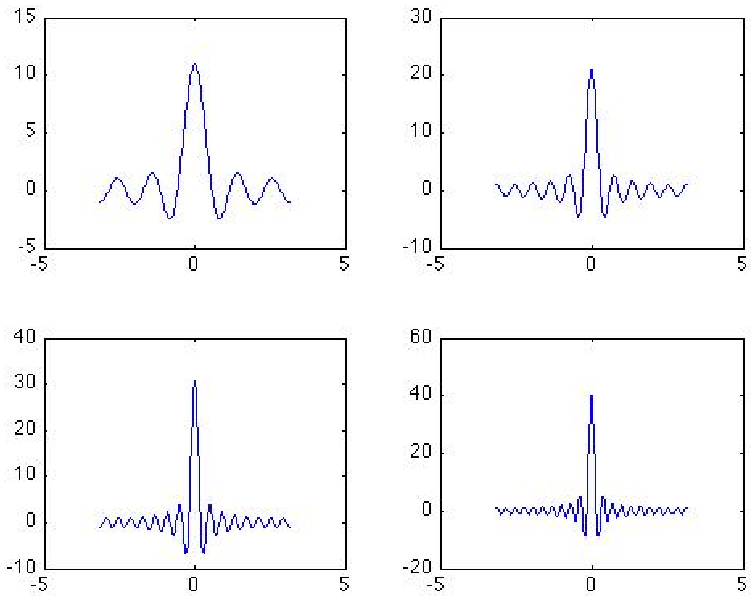
If we want to prove that partial sums $s_N(x)$ converge to $f(x)$ (i.e. that the Fourier series converges pointwise to f), the obvious strategy is to prove that the integral above converges to f. In 1829, Dirichlet used this approach to prove:

Theorem 4.10.1 (Dirichlet's Theorem) *If $f \in D$ has only a finite number of local minima and maxima, then the Fourier series of f converges pointwise to f.*

Dirichlet's result must have come as something of a surprise; it probably seemed unlikely that a theorem should hold for functions with jumps, but not for continuous functions with an infinite number of extreme points.

Through the years that followed, a number of mathematicians tried — and failed — to prove that the Fourier series of a periodic, continuous function always converges to the function. In 1873, the German mathematician Paul Du Bois-Reymond explained why they failed by constructing a periodic, continuous function whose Fourier series diverges at a dense set of points.

It turns out that the theory for pointwise convergence of Fourier series is quite complicated, and we shall not prove Dirichlet's theorem here. Instead we shall prove a result known as *Dini's test* which allows us to prove convergence for many of the functions that appear in practice. But before we do that, we shall take a look at a different notion of convergence which is easier to handle, and which will also give us some tools that are useful in the proof of Dini's test. This alternative notion of convergence is called *Cesaro convergence* or *convergence in Cesaro mean*. But first of all we shall collect some properties of the Dirichlet kernels that will be useful later.



Let us first see what they look like. The figure above shows Dirichlet's kernel D_n for $n = 5, 10, 15, 20$. Note the changing scale on the y-axis; as we have alresdy observed, the maximum value of D_n is $2n + 1$. As n grows, the graph becomes more and more dominated by a sharp peak at the origin. The smaller peaks and valleys shrink in size relative to the big peak, but the problem with the Dirichlet kernel is that they do not shrink in absolute terms — as n goes to infinity, the area between the curve and the x-axis (measured in absolute value) goes to infinity. This makes the Dirichlet kernel quite difficult to work with. When we turn to Cesaro convergence in the next section, we get another set of kernels — the *Fejér kernels* — and they turn out not to have this problem. This is the main reason why Cesaro convergence works much better than ordinary convergence for Fourier series.

Let us now prove some of the crucial properties of the Dirichlet kernel. Recall that a function g is even if $g(t) = g(-t)$ for all t in the domain:

Lemma 4.10.2 *The Dirichlet kernel $D_n(t)$ is an even, realvalued function such that $|D_n(t)| \leq D_n(0) = 2n + 1$ for all t. For all n,*

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} D_n(t)\,dt = 1$$

but

$$\lim_{n\to\infty}\int_{-\pi}^{\pi} |D_n(t)|\,dt \to \infty$$

Proof: That D_n is realvalued and even, follows immediately from the formula $D_n(t) = \frac{\sin((n+\frac{1}{2})t)}{\sin\frac{t}{2}}$ To prove that $|D_n(t)| \leq D_n(0) = 2n+1$, we just observe that

$$D_n(t) = |\sum_{k=-n}^{n} e^{ikt}| \leq \sum_{k=-n}^{n} |e^{ikt}| = 2n + 1 = D_n(0)$$

Similarly for the integral

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} D_n(t)\,dt = \sum_{k=-n}^{n}\frac{1}{2\pi}\int_{-\pi}^{\pi} \mathbf{e}^{ikt}\,dt = 1$$

as all integrals except the one for $k = 0$ is zero. To prove the last part, we observe that since $|\sin u| \leq |u|$ for all u, we have

$$|D_n(t)| = \frac{|\sin((n+\frac{1}{2})t)|}{|\sin\frac{t}{2}|} \geq \frac{2|\sin((n+\frac{1}{2})t)|}{|t|}$$

Using the symmetry and the substitution $z = (n + \frac{1}{2})t$, we see that

$$\int_{-\pi}^{\pi} |D_n(t)|\,dt = \int_0^{\pi} 2|D_n(t)|\,dt \geq \int_0^{\pi} \frac{4|\sin((n+\frac{1}{2})t)|}{|t|}\,dt =$$

$$= \int_0^{(n+\frac{1}{2})\pi} \frac{4|\sin z|}{z}\,dz \geq \sum_{k=1}^{n}\int_{(k-1)\pi}^{k\pi} \frac{4|\sin z|}{k\pi}\,dz = \frac{8}{\pi}\sum_{k=1}^{n}\frac{1}{k}$$

The expression on the right goes to infinity since the series diverges. □

The last part of the lemma is bad news. It tells us that when we are doing calculations with the Dirichlet kernel, we have to be very careful in putting in absolute values as the integrals are likely to diverge. For this reason we shall now introduce another kernel — the *Fejér kernel* — where this problem does not occur.

Exercises for Section 4.10

1. Let $f : [-\pi, \pi] \to \mathbb{C}$ be the function $f(x) = x$. Draw the periodic extension of f. Do the same with the function $g(x) = x^2$.
2. Check that $D_n(0) = 2n + 1$ by computing $\lim_{t\to 0} \frac{\sin((n+\frac{1}{2})t)}{\sin\frac{t}{2}}$.
3. Work out the details of the substitution $u = x - t$ in the derivation of the formula $s_N(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-u)\sum_{n=-N}^{N} e^{inu}\,du$.
4. Explain the details in the last part of the proof of Lemma 4.10.2 (the part that proves that $\lim_{n\to\infty}\int_{-\pi}^{\pi} |D_n(t)|\,dt = \infty$).

4.11 The Fejér kernel

Before studying the Fejér kernel, we shall take a look at a generalized notion of convergence for sequences. Certain sequences such at

$$0, 1, 0, 1, 0, 1, 0, 1, \ldots$$

do not converge in the ordinary sense, but they do converge "in average" in the sense that the average of the first n elements approaches a limit as n goes to infinity. In this sense, the sequence above obviously converges to $\frac{1}{2}$. Let us make this notion precise:

Definition 4.11.1 *Let $\{a_k\}_{k=0}^{\infty}$ be a sequence of complex numbers, and let $S_n = \frac{1}{n}\sum_{k=0}^{n-1} a_k$. We say that the sequence converges to $a \in \mathbb{C}$ in* Cesaro mean *if*

$$a = \lim_{n\to\infty} S_n = \lim_{n\to\infty} \frac{a_0 + a_1 + \cdots + a_{n-1}}{n}$$

We shall write $a = C\text{-}\lim_{n\to\infty} a_n$.

The sequence at the beginning of the section converges to $\frac{1}{2}$ in Cesaro mean, but diverges in the ordinary sense. Let us prove that the opposite can not happen:

Lemma 4.11.2 *If $\lim_{n\to\infty} a_n = a$, then $C\text{-}\lim_{n\to\infty} a_n = a$.*

Proof: Given an $\epsilon > 0$, we must find an N such that

$$|S_n - a| < \epsilon$$

when $n \geq N$. Since $\{a_n\}$ converges to a, there is a $K \in \mathbb{N}$ such that $|a_n - a| < \frac{\epsilon}{2}$ when $n \geq K$. If we let $M = \max\{|a_k - a| : k = 0, 1, 2, \ldots\}$, we have for any $n \geq K$:

$$|S_n - a| = \left|\frac{(a_0 - a) + (a_1 - a) + \cdots + (a_{K-1} - a) + (a_K - a) + \cdots (a_{n-1} - a)}{n}\right| \leq$$

$$\leq \left| \frac{(a_0 - a) + (a_1 - a) + \cdots + (a_{K-1} - a)}{n} \right| + \left| \frac{(a_K - a) + \cdots (a_{n-1} - a)}{n} \right| \leq \frac{MK}{n} + \frac{\epsilon}{2}$$

Choosing n large enough, we get $\frac{MK}{n} < \frac{\epsilon}{2}$, and the lemma follows. □

The idea behind the Fejér kernel is to show that the partial sums $s_n(x)$ converge to $f(x)$ in Cesaro mean; i.e. that the sums

$$S_n(x) = \frac{s_0(x) + s_1(x) + \cdots + s_{n-1}(x)}{n}$$

converge to $f(x)$. Since

$$s_k(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x-u) D_k(u) \, du$$

where D_k is the Dirichlet kernel, we get

$$S_n(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x-u) \left(\frac{1}{n} \sum_{k=0}^{n-1} D_k(u) \right) du = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x-u) F_n(u) \, du$$

where $F_n(u) = \frac{1}{n} \sum_{k=0}^{n-1} D_k(u)$ is the *Fejér kernel.*

We can find a closed expression for the Fejér kernel as we did for the Dirichlet kernel, but the arguments are a little longer:

Lemma 4.11.3 *The Fejér kernel is given by*

$$F_n(u) = \frac{\sin^2(\frac{nu}{2})}{n \sin^2(\frac{u}{2})}$$

for $u \neq 0$*, and* $F_n(0) = n$*.*

Proof: Since

$$F_n(u) = \frac{1}{n} \sum_{k=0}^{n-1} D_k(u) = \frac{1}{n \sin(\frac{u}{2})} \sum_{k=0}^{n-1} \sin((k + \frac{1}{2})u)$$

we have to find

$$\sum_{k=0}^{n-1} \sin((k + \frac{1}{2})u) = \frac{1}{2i} \left(\sum_{k=0}^{n-1} e^{i(k+\frac{1}{2})u} - \sum_{k=0}^{n-1} e^{-i(k+\frac{1}{2})u} \right)$$

The series are geometric and can easily be summed:

$$\sum_{k=0}^{n-1} e^{i(k+\frac{1}{2})u} = e^{i\frac{u}{2}} \sum_{k=0}^{n-1} e^{iku} = e^{i\frac{u}{2}} \frac{1 - e^{inu}}{1 - e^{iu}} = \frac{1 - e^{inu}}{e^{-i\frac{u}{2}} - e^{i\frac{u}{2}}}$$

and

$$\sum_{k=0}^{n-1} e^{-i(k+\frac{1}{2})u} = e^{-i\frac{u}{2}} \sum_{k=0}^{n-1} e^{-iku} = e^{-i\frac{u}{2}} \frac{1-e^{-inu}}{1-e^{-iu}} = \frac{1-e^{-inu}}{e^{i\frac{u}{2}} - e^{-i\frac{u}{2}}}$$

Hence

$$\sum_{k=0}^{n-1} \sin((k+\frac{1}{2})u) = \frac{1}{2i}\left(\frac{1-e^{inu}+1-e^{-inu}}{e^{-i\frac{u}{2}} - e^{i\frac{u}{2}}}\right) = \frac{1}{2i}\left(\frac{e^{inu}-2+e^{-inu}}{e^{i\frac{u}{2}} - e^{-i\frac{u}{2}}}\right) =$$

$$= \frac{1}{2i} \cdot \frac{(e^{i\frac{nu}{2}} - e^{-\frac{nu}{2}})^2}{e^{i\frac{u}{2}} - e^{-i\frac{u}{2}}} = \frac{\left(\frac{e^{i\frac{nu}{2}} - e^{-\frac{nu}{2}})}{2i}\right)^2}{\frac{e^{i\frac{u}{2}} - e^{-i\frac{u}{2}}}{2i}} = \frac{\sin^2(\frac{nu}{2})}{\sin\frac{u}{2}}$$

and thus

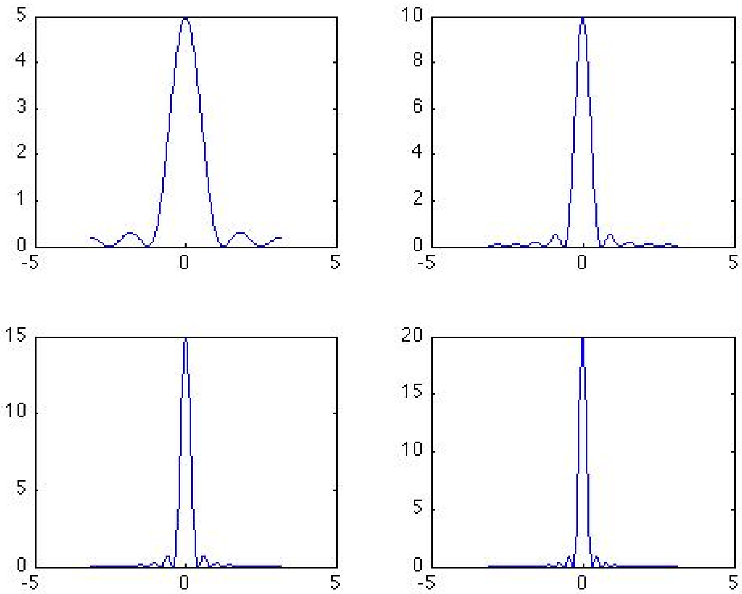
$$F_n(u) = \frac{1}{n\sin(\frac{u}{2})} \sum_{k=0}^{n-1} \sin((k+\frac{1}{2})u) = \frac{\sin^2(\frac{nu}{2})}{n\sin^2\frac{u}{2}}$$

To prove that $F_n(0) = n$, we just have to sum an arithmetic series

$$F_n(0) = \frac{1}{n}\sum_{k=0}^{n-1} D_k(0) = \frac{1}{n}\sum_{k=0}^{n-1}(2k+1) = n$$

□

The figure below shows the Fejer kernels F_n for $n = 5, 10, 15, 20$.



At first glance they look very much like the Dirichlet kernels in the previous section. The peak in the middle is growing slower than before in absolute terms (the maximum value is n compared to $2n+1$ for the Dirichlet kernel), but relative to the smaller peaks and values, it is much more dominant. The functions are now positive, and the area between their graphs and the x-axis is always equal to one. As n gets big, almost all this area belongs to the dominant peak in the middle. The positivity and the concentration of all the area in the center peak make the Fejér kernels much easier to handle than their Dirichlet counterparts.

Let us now prove some of the properties of the Fejér kernels.

Proposition 4.11.4 *For all n, the Fejér kernel F_n is an even, positive function such that*

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} F_n(x)\,dx = 1$$

For all nonzero $x \in [-\pi, \pi]$

$$0 \leq F_n(x) \leq \frac{\pi^2}{nx^2}$$

Proof: That F_n is even and positive follows directly from the formula in the lemma. By Proposition 4.10.2, we have

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} F_n(x)\,dx = \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{1}{n}\sum_{k=0}^{n-1} D_k\,dx = \frac{1}{n}\sum_{k=0}^{n-1}\frac{1}{2\pi}\int_{-\pi}^{\pi} D_k\,dx = \frac{1}{n}\sum_{k=0}^{n-1} 1 = 1$$

For the last formula, observe that for $u \in [-\frac{\pi}{2}, \frac{\pi}{2}]$, we have $\frac{2}{\pi}|u| \leq |\sin u|$ (make a drawing). Thus

$$F_n(x) = \frac{\sin^2(\frac{nx}{2})}{n\sin^2\frac{x}{2}} \leq \frac{1}{n(\frac{2}{\pi}\frac{x}{2})^2} \leq \frac{\pi^2}{nx^2}$$

□

We shall now show that $S_n(x)$ converges to $f(x)$, i.e. that the Fourier series converges to f in Cesaro mean. We have already observed that

$$S_n(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-u)F_n(u)\,du$$

If we introduce a new variable $t = -u$ and use that F_n is even, we get

$$S_n(x) = \frac{1}{2\pi}\int_{\pi}^{-\pi} f(x+t)F_n(-t)\,(-dt) =$$

$$= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x+t)F_n(t)\,dt = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x+u)F_n(u)\,du$$

If we combine the two expressions we now have for $S_n(x)$, we get

$$S_n(x) = \frac{1}{4\pi}\int_{-\pi}^{\pi} (f(x+u) + f(x-u))\, F_n(u)\, du$$

Since $\frac{1}{2\pi}\int_{-\pi}^{\pi} F_n(u)\, du = 1$, we also have

$$f(x) = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)F_n(u)\, du$$

Hence

$$S_n(x) - f(x) = \frac{1}{4\pi}\int_{-\pi}^{\pi} (f(x+u) + f(x-u) - 2f(x))\, F_n(u)\, du$$

To prove that $S_n(x)$ converges to $f(x)$, we only need to prove that the integral goes to 0 as n goes to infinity. The intuitive reason for this is that for large n, the kernel $F_n(u)$ is extremely small except when u is close to 0, but when u is close to 0, the other factor in the integral, $f(x+u)+f(x-u)-2f(x)$, is very small. Here are the technical details.

Theorem 4.11.5 *If $f \in D$, then S_n converges to f on $[-\pi, \pi]$, i.e. the Fourier series converges in Cesaro mean. The convergence is uniform on each subinterval $[a, b] \subseteq [-\pi, \pi]$ where f is continuous.*

Proof: Given $\epsilon > 0$, we must find an $N \in \mathbb{N}$ such that $|S_n(x) - f(x)| < \epsilon$ when $n \geq N$. Since f is in D, there is a $\delta > 0$ such that

$$|f(x+u) - f(x-u) - 2f(x)| < \epsilon$$

when $|u| < \delta$ (keep in mind that since $f \in D$, $f(x) = \frac{1}{2}\lim_{u\uparrow 0}(f(x+u) - f(x-u))$). We have

$$|S_n(x) - f(x)| = \frac{1}{4\pi}\int_{-\pi}^{\pi} |f(x+u) + f(x-u) - 2f(x)|\, F_n(u)\, du =$$

$$= \frac{1}{4\pi}\int_{-\delta}^{\delta} |f(x+u) + f(x-u) - 2f(x)|\, F_n(u)\, du +$$

$$+ \frac{1}{4\pi}\int_{-\pi}^{-\delta} |f(x+u) + f(x-u) - 2f(x)|\, F_n(u)\, du +$$

$$+ \frac{1}{4\pi}\int_{-\delta}^{\pi} |f(x+u) + f(x-u) - 2f(x)|\, F_n(u)\, du$$

For the first integral we have

$$\frac{1}{4\pi}\int_{-\delta}^{\delta} |f(x+u) + f(x-u) - 2f(x)|\, F_n(u)\, du \leq$$

$$\leq \frac{1}{4\pi}\int_{-\delta}^{\delta} \epsilon F_n(u)\,du \leq \frac{1}{4\pi}\int_{-\pi}^{\pi} \epsilon F_n(u)\,du = \frac{\epsilon}{2}$$

For the second integral we get

$$\frac{1}{4\pi}\int_{-\pi}^{-\delta} |f(x+u)+f(x-u)-2f(x)|\,F_n(u)\,du \leq$$

$$\leq \frac{1}{4\pi}\int_{-\pi}^{-\delta} 4\|f\|_\infty \frac{\pi^2}{n\delta^2}\,du = \frac{\pi^2\|f\|_\infty}{n\delta^2}$$

Exactly the same estimate holds for the third integral, and by choosing $N > \frac{4\pi^2\|f\|_\infty}{\epsilon\delta^2}$, we get the sum of the last two integrals less than $\frac{\epsilon}{2}$. But then $|S_n(x) - f(x)| < \epsilon$, and the convergence is proved.

So what about the uniform convergence? We need to check that we can choose the same N for all $x \in [a,b]$. Note that N only depends on x through the choice of δ, and hence it suffices to show that we can use the same δ for all $x \in [a,b]$. One might think that this follows immediately from the fact that a continuous function on a compact interval $[a,b]$ is uniformly continuous, but we need to be a little careful as $x+u$ or $x-u$ may be outside the interval $[a,b]$ even if x is inside. The quickest way to fix this, is to observe that since f is in D, it must be continuous — and hence uniformly continuous — on a slightly larger interval $[a-\eta, b+\eta]$. This means that we can use the same $\delta < \eta$ for all x and $x \pm u$ in $[a-\eta, b+\eta]$, and this clinches the argument.□

We have now finally proved Theorem 4.9.2 which we restate here:

Corollary 4.11.6 *The trigonometric polynomials are dense in C_p in $\|\cdot\|_\infty$-norm, i.e. for any $f \in C_P$ there is a sequence of trigonometric polynomials converging uniformly to f.*

Proof: According to the theorem, the sums $S_N(x) = \frac{1}{N}\sum_{n=0}^{N-1} s_n(x)$ converge uniformly to f. Since each s_n is a trigonometric polynomial, so are the S_N's. □

Exercises to Section 4.11

1. Let $\{a_n\}$ be the sequence $1,0,1,0,1,0,1,0,\ldots$. Prove that C-$\lim_{n\to\infty} a_n = \frac{1}{2}$.
2. Show that C-$\lim_{n\to\infty}(a_n+b_n) = $ C$-\lim_{n\to\infty} a_n +$ C$-\lim_{n\to\infty} b_n$
3. Check that $F_n(0) = n$ by computing $\lim_{u\to 0} \frac{\sin^2(\frac{nu}{2})}{n\sin^2\frac{u}{2}}$.
4. Show that $S_N(x) = \sum_{n=-(N-1)}^{N-1} \alpha_n(1-\frac{|n|}{N})e_n(x)$, where $\alpha_n = \langle f, e_n\rangle$ is the Fourier coefficient.
5. Assume that $f \in C_P$. Work through the details of the proof of Theorem 4.11.5 and check that S_n converges uniformly to f.

4.12 The Riemann-Lebesgue lemma

The Riemann-Lebesgue lemma is a seemingly simple observation about the size of the Fourier coefficients, but it turns out to be a very efficient tool in the study of pointwise convergence.

Theorem 4.12.1 (Riemann-Lebesgue Lemma) *If $f \in D$ and*

$$\alpha_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx}\, dx, \quad n \in \mathbb{Z},$$

are the Fourier coefficients of f, then $\lim_{|n|\to\infty} \alpha_n \to 0$.

Proof: According to Bessel's inequality 4.6.9, $\sum_{n=-\infty}^{\infty} |\alpha_n|^2 \leq \|f\|_2^2 < \infty$, and hence $\alpha_n \to 0$ as $|n| \to \infty$. □

Remark: We are cheating a little here as we only prove the Riemann-Lebesgue lemma for function which are in D and hence square integrable. The lemma holds for integrable functions in general, but even in that case the proof is quite easy.

The Riemann-Lebesgue lemma is quite deceptive. It seems to be a result about the coefficients of certain series, and it is proved by very general and abstract methods, but it is really a theorem about oscillating integrals as the following corollary makes clear.

Corollary 4.12.2 *If $f \in D$ and $[a,b] \subseteq [-\pi, \pi]$, then*

$$\lim_{|n|\to\infty} \int_a^b f(x) e^{-inx}\, dx = 0$$

Also

$$\lim_{|n|\to\infty} \int_a^b f(x) \cos(nx)\, dx = \lim_{|n|\to\infty} \int_a^b f(x) \sin(nx)\, dx = 0$$

Proof: Let g be the function (this looks more horrible than it is!)

$$g(x) = \begin{cases} 0 & \text{if } x \notin [a,b] \\ f(x) & \text{if } x \in (a,b) \\ \frac{1}{2}\lim_{x\downarrow a} f(x) & \text{if } x = a \\ \frac{1}{2}\lim_{x\uparrow b} f(x) & \text{if } x = b \end{cases}$$

then g is in D, and

$$\int_a^b f(x) e^{-inx}\, dx = \int_{-\pi}^{\pi} g(x) e^{-inx}\, dx = 2\pi\alpha_n$$

where α_n is the Fourier coefficient of g. By the Riemann-Lebesgue lemma, $\alpha_n \to 0$. The last two parts follows from what we have just proved and the identities $\sin(nx) = \frac{e^{inx}-e^{-inx}}{2i}$ and $\cos(nx) = \frac{e^{inx}+e^{-inx}}{2}$ □

Let us pause for a moment to discuss why these results hold. The reason is simply that for large values of n, the functions $\sin nx$, $\cos nx$, and e^{inx} (if we consider the real and imaginary parts separately) oscillate between positive and negative values. If the function f is relatively smooth, the positive and negative contributions cancel more and more as n increases, and in the limit there is nothing left. This argument also indicates why rapidly oscillating, continuous functions are a bigger challenge for Fourier analysis than jump discontinuities — functions with jumps average out on each side of the jump, while for wildly oscillating functions "the averaging" procedure may not work.

Since the Dirichlet kernel contains the factor $\sin((n+\frac{1}{2})x)$, the following result will be useful in the next section:

Corollary 4.12.3 *If $f \in D$ and $[a,b] \subseteq [-\pi, \pi]$, then*

$$\lim_{|n|\to\infty} \int_a^b f(x) \sin\big((n+\frac{1}{2})x\big)\, dx = 0$$

Proof: Follows from the corollary above and the identity

$$\sin\big((n+\frac{1}{2})x\big) = \sin(nx)\cos\frac{x}{2} + \cos(nx)\sin\frac{x}{2}$$

□

Exercises to Section 4.12

1. Work out the details of the $\sin(nx)$- and $\cos(nx)$-part of Corollary 4.12.2.
2. Work out the details of the proof of Corollary 4.12.3.
3. a) Show that if p is a trigonometric polynomial, then the Fourier coefficients $\beta_n = \langle p, e_n \rangle$ are zero when $|n|$ is sufficiently large.

   b) Let f be an integrable function, and assume that for each $\epsilon > 0$ there is a trigonometric polynomial such that $\frac{1}{2\pi}\int_{-\pi}^{\pi} |f(t) - p(t)|\, dt < \epsilon$. Show that if $\alpha_n = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)e^{-int}\, dt$ are the Fourier coefficients of f, then $\lim_{|n|\to\infty} \alpha_n = 0$.

4.13 Dini's test

We shall finally take a serious look at pointwise convergence of Fourier series. As aready indicated, this is a rather tricky business, and there is no ultimate theorem, just a collection of scattered results useful in different settings. We shall concentrate on a criterion called *Dini's test* which is relatively easy to prove and sufficiently general to cover a lot of different situations.

Recall from Section 4.10 that if

$$s_N(x) = \sum_{n=-N}^{N} \langle f, e_n \rangle e_n(x)$$

is the partial sum of a Fourier series, then

$$s_N(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x-u) D_N(u)\, du$$

If we change variable in the intergral and use the symmetry of D_N, we see that we also get

$$s_N(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x+u) D_N(u)\, du$$

Combining these two expressions, we get

$$s_N(x) = \frac{1}{4\pi} \int_{-\pi}^{\pi} \big(f(x+u) + f(x-u)\big) D_N(u)\, du$$

Since $\frac{1}{2\pi}\int_{-\pi}^{\pi} D_N(u)\, du = 1$, we also have

$$f(x) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) D_N(u)\, du$$

and hence

$$s_N(x) - f(x) = \frac{1}{4\pi} \int_{-\pi}^{\pi} \left(f(x+u) + f(x-u) - 2f(x)\right) D_N(u)\, du$$

(note that the we are now doing exactly the same to the Dirichlet kernel as we did to the Fejér kernel in Section 4.11). To prove that the Fourier series converges pointwise to f, we just have to prove that the integral converges to 0.

The next lemma simplifies the problem by telling us that we can concentrate on what happens close to the origin:

Lemma 4.13.1 *Let $f \in D$ and assume that there is a $\eta > 0$ such that*

$$\lim_{N\to\infty} \frac{1}{4\pi} \int_{-\eta}^{\eta} \left(f(x+u) + f(x-u) - 2f(x)\right) D_N(u)\, du = 0$$

Then the Fourier series $\{s_N(x)\}$ converges to $f(x)$.

Proof: Note that since $\frac{1}{\sin\frac{x}{2}}$ is a bounded function on $[\eta, \pi]$, Corollary 4.12.3 tells us that

$$\lim_{N\to\infty} \frac{1}{4\pi}\int_{\eta}^{\pi} (f(x+u)+f(x-u)-2f(x))\, D_N(u)\, du =$$

$$= \lim_{N\to\infty} \frac{1}{4\pi}\int_{\eta}^{\pi} \left[(f(x+u)+f(x-u)-2f(x))\, \frac{1}{\sin\frac{u}{2}}\right] \sin\left((N+\frac{1}{2})u\right) du = 0$$

The same obviously holds for the integral from $-\pi$ to $-\eta$, and hence

$$s_N(x) - f(x) = \frac{1}{4\pi}\int_{-\pi}^{\pi} (f(x+u)+f(x-u)-2f(x))\, D_N(u)\, du =$$

$$= \frac{1}{4\pi}\int_{-\pi}^{\eta} (f(x+u)+f(x-u)-2f(x))\, D_N(u)\, du +$$

$$+ \frac{1}{4\pi}\int_{-\eta}^{\eta} (f(x+u)+f(x-u)-2f(x))\, D_N(u)\, du +$$

$$+ \frac{1}{4\pi}\int_{\eta}^{\pi} (f(x+u)+f(x-u)-2f(x))\, D_N(u)\, du$$

$$\to 0+0+0=0$$

□

Theorem 4.13.2 (Dini's test) *Let $x \in [-\pi, \pi]$, and assume that there is a $\delta > 0$ such that*

$$\int_{-\delta}^{\delta} \left|\frac{f(x+u)+f(x-u)-2f(x)}{u}\right| du < \infty$$

Then the Fourier series converges to the function f at the point x, i.e. $s_N(x) \to f(x)$.

Proof: According to the lemma, it suffices to prove that

$$\lim_{N\to\infty} \frac{1}{4\pi}\int_{-\delta}^{\delta} (f(x+u)+f(x-u)-2f(x))\, D_N(u)\, du = 0$$

Given an $\epsilon > 0$, we have to show that if $N \in \mathbb{N}$ is large enough, then

$$\frac{1}{4\pi}\int_{-\delta}^{\delta} (f(x+u)+f(x-u)-2f(x))\, D_N(u)\, du < \epsilon$$

Since the integral in the theorem converges, there is an $\eta > 0$ such that

$$\int_{-\eta}^{\eta} \left|\frac{f(x+u)+f(x-u)-2f(x)}{u}\right| du < \epsilon$$

Since $|\sin v| \geq \frac{2|v|}{\pi}$ for $v \in [-\frac{\pi}{2}, \frac{\pi}{2}]$ (make a drawing), we have $|D_N(u)| = |\frac{\sin((N+\frac{1}{2})u)}{\sin\frac{u}{2}}| \leq \frac{\pi}{|u|}$ for $u \in [-\pi, \pi]$. Hence

$$|\frac{1}{4\pi}\int_{-\eta}^{\eta}(f(x+u)+f(x-u)-2f(x))\,D_N(u)\,du| \leq$$

$$\leq \frac{1}{4\pi}\int_{-\eta}^{\eta}|f(x+u)+f(x-u)-2f(x)|\,\frac{\pi}{|u|}\,du < \frac{\epsilon}{4}$$

By Corollary 4.12.3 we can get

$$\frac{1}{4\pi}\int_{\eta}^{\delta}(f(x+u)+f(x-u)-2f(x))\,D_N(u)\,du$$

as small as we want by choosing N large enough and similarly for the integral from $-\delta$ to $-\eta$. In particular, we can get

$$\frac{1}{4\pi}\int_{-\delta}^{\delta}(f(x+u)+f(x-u)-2f(x))\,D_N(u)\,du =$$

$$= \frac{1}{4\pi}\int_{-\delta}^{-\eta}(f(x+u)+f(x-u)-2f(x))\,D_N(u)\,du +$$

$$+\frac{1}{4\pi}\int_{-\eta}^{\eta}(f(x+u)+f(x-u)-2f(x))\,D_N(u)\,du +$$

$$+\frac{1}{4\pi}\int_{\eta}^{\delta}(f(x+u)+f(x-u)-2f(x))\,D_N(u)\,du$$

less than ϵ, and hence the theorem is proved. □

Dini's test has some immediate consequences that we leave to the reader to prove.

Corollary 4.13.3 *If $f \in D$ is differentiable at a point x, then the Fourier series converges to $f(x)$ at this point.*

We may even extend this result to one-sided derivatives:

Corollary 4.13.4 *Assume $f \in D$ and that the limits*

$$\lim_{u\downarrow 0}\frac{f(x+u)-f(x^+)}{u}$$

and

$$\lim_{u\uparrow 0}\frac{f(x+u)-f(x^-)}{u}$$

exist at a point x. Then the Fourier series $s_N(x)$ converges to $f(x)$ at this point.

Exercises to Section 4.13

1. Show that the Fourier series $\sum_{n=1}^{\infty} \frac{2(-1)^{n+1}}{n} \sin(nx)$ in Example 4.7.1 converges to $f(x) = x$ for $x \in (-\pi, \pi)$. What happens in the endpoints?
2. Prove Corollary 4.13.3
3. Prove Corollary 4.13.4

4.14 Termwise operations

In Section 4.3 we saw that power series can be integrated and differentiated term by term, and we now want to take a quick look at the corresponding questions for Fourier series. Let us begin by integration which is by far the easiest operation to deal with.

The first thing we should observe, is that when we integrate a Fourier series $\sum_{-\infty}^{\infty} \alpha_n e^{inx}$ term by term, we do *not* get a new Fourier series since the constant term α_0 integrates to $\alpha_0 x$, which is not a term in a Fourier series when $\alpha_0 \neq 0$. However, we may, of course, still integrate term by term to get the series

$$\alpha_0 x + \sum_{n \in \mathbb{Z}, n \neq 0} \left(-\frac{i\alpha_n}{n}\right) e^{inx}$$

The question is if this series converges to the integral of f.

Proposition 4.14.1 *Let $f \in D$, and define $g(x) = \int_0^x f(t)\,dt$. If s_n is the partial sums of the Fourier series of f, then the functions $t_n(x) = \int_0^x s_n(t)\,dt$ converge uniformly to g on $[-\pi, \pi]$. Hence*

$$g(x) = \int_0^x f(t)\,dt = \alpha_0 x + \sum_{n \in \mathbb{Z}, n \neq 0} -\frac{i\alpha_n}{n} \left(e^{inx} - 1\right)$$

where the convergence of the series is uniform.

Proof: By Cauchy-Schwarz's inequality we have

$$|g(x) - t_n(x)| = |\int_0^{\pi} (f(t) - s_n(t))\,dt| \leq \int_{-\pi}^{\pi} |f(t) - s_n(t)|\,dt \leq$$

$$\leq 2\pi \left(\frac{1}{2\pi} \int_{-\pi}^{\pi} |f(s) - s_n(s)| \cdot 1\,ds\right) = 2\pi \langle |f - s_n|, 1 \rangle \leq$$

$$\leq 2\pi \|f - s_n\|_2 \|1\|_2 = 2\pi \|f - s_n\|_2$$

By Theorem 4.9.6, we see that $\|f - s_n\|_2 \to 0$, and hence t_n converges uniformly to $g(x)$. □

If we move the term $\alpha_0 x$ to the other side in the formula above, we get

$$g(x) - \alpha_0 x = \sum_{n\in\mathbb{Z}, n\neq 0} \frac{i\alpha_n}{n} - \sum_{n\in\mathbb{Z}, n\neq 0} \frac{i\alpha_n}{n} e^{inx}$$

where the series on the right is the Fourier series of $g(x) - \alpha_0 x$ (the first sum is just the constant term of the series).

As always, termwise differentiation is a much trickier subject. In Example 1 of Section 4.8, we showed that the Fourier series of x is

$$\sum_{n=1}^{\infty} \frac{2(-1)^{n+1}}{n} \sin(nx),$$

and by what we now know, it is clear that the series converges pointwise to x on $(-\pi, \pi)$. However, if we differentiate term by term, we get the hopelessly divergent series

$$\sum_{n=1}^{\infty} 2(-1)^{n+1} \cos(nx)$$

Fortunately, there is more hope when $f \in C_p$, i.e. when f is continuous and $f(-\pi) = f(\pi)$:

Proposition 4.14.2 *Assume that $f \in C_P$ and that f' is continuous on $[-\pi, \pi]$. If $\sum_{n=0}^{\infty} \alpha_n e^{inx}$ is the Fourier series of f, then the differentiated series $\sum_{n=0}^{\infty} in\alpha_n e^{inx}$ is the Fourier series of f', and it converges pointwise to f' at any point x where $f''(x)$ exists.*

Proof: Let β_n be the Fourier coefficient of f'. By integration by parts

$$\beta_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f'(t) e^{-int}\, dt = \frac{1}{2\pi} \left[f(t) e^{-int} \right]_{-\pi}^{\pi} - \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t)(-ine^{-int})\, dt =$$

$$= 0 + in \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) e^{-int}\, dt = in\alpha_n$$

which shows that $\sum_{n=0}^{\infty} in\alpha_n e^{inx}$ is the Fourier series of f'. The convergence follows from Corollary 4.13.3. □

Final remark: In this chapter we have developed Fourier analysis over the interval $[-\pi, \pi]$. If we want to study Fourier series over another interval $[a - r, a + r]$, all we have to do is to move and rescale the functions: The basis now consists of the functions

$$e_n(x) = e^{\frac{in\pi}{r}(x-a)},$$

the inner product is defined by

$$\langle f, g \rangle = \frac{1}{2r} \int_{a-r}^{a+r} f(x)\overline{g(x)}\, dx$$

and the Fourier series becomes

$$\sum_{n=-\infty}^{\infty} \alpha_n e^{\frac{in\pi}{r}(x-a)}$$

Note that when the length r of the interval increases, the frequencies $\frac{in\pi}{r}$ of the basis functions $e^{\frac{in\pi}{r}(x-a)}$ get closer and closer. In the limit, one might imagine that the sum $\sum_{n=-\infty}^{\infty} \alpha_n e^{\frac{in\pi}{r}(x-a)}$ turns into an integral (think of the case $a = 0$):

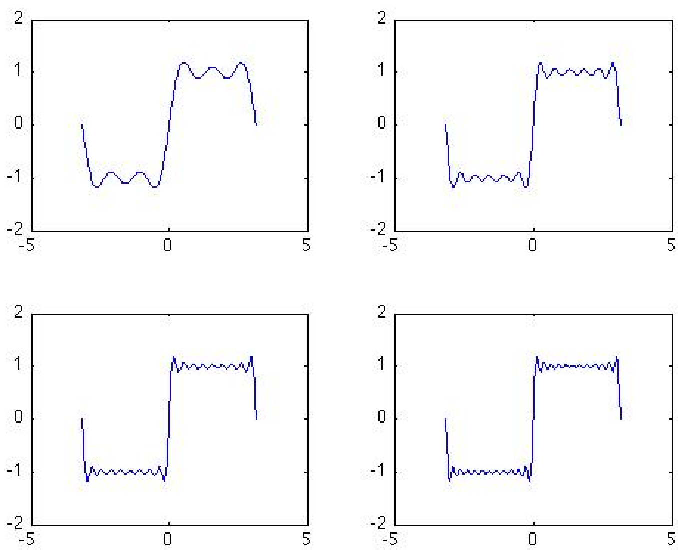
$$\int_{-\infty}^{\infty} \alpha(t) e^{ixt}\, dt$$

This leads to the theory of Fourier integrals and Fourier transforms, but we shall not look into these topics here.

Exercises for Section 4.14

1. Use integration by parts to check that $\sum_{n\in\mathbb{Z},n\neq 0} \frac{i\alpha_n}{n} - \sum_{n\in\mathbb{Z},n\neq 0} \frac{i\alpha_n}{n} e^{inx}$ is the Fourier series of $g(x) - \alpha_0 x$ (see the passage after the proof of Proposition 4.14.1).
2. Show that $\sum_{k=1}^{n} \cos((2k-1)x) = \frac{\sin 2nx}{2\sin x}$.
3. In this problem we shall study a feature of Fourier series known as *Gibbs' phenomenon*. Let $f : [-\pi, \pi] \to \mathbb{R}$ be given by

$$f(x) = \begin{cases} -1 & \text{for } x < 0 \\ 0 & \text{for } x = 0 \\ 1 & \text{for } x > 1 \end{cases}$$



The figure above shows the partial sums $s_n(x)$ of order $n = 5, 11, 17, 23$. We see that although the approximation in general seems to get better and better, the maximal distance between f and s_n remains more or less constant — it seems that the partial sums have "bumps" of more or less constant height near the jump in function values. We shall take a closer look at this phenomenon. Along the way you will need the solution of problem 3.

a) Show that the partial sums can be expressed as

$$s_{2n-1}(x) = \frac{4}{\pi} \sum_{k=1}^{n} \frac{\sin((2k-1)x)}{2k-1}$$

b) Use problem 2 to find a short expression for $s'_{2n-1}(x)$.

c) Show that the local minimum and maxima of s_{2n-1} closest to 0 are $x_- = -\frac{\pi}{2n}$ and $x_+ = \frac{\pi}{2n}$.

d) Show that

$$s_{2n-1}(\pm\frac{\pi}{2n}) = \pm\frac{4}{\pi} \sum_{k=1}^{n} \frac{\sin\frac{(2k-1)\pi}{2n}}{2k-1}$$

e) Show that $s_{2n-1}(\pm\frac{\pi}{2n}) \to \pm\frac{2}{\pi}\int_0^\pi \frac{\sin x}{x}\,dx$ by recognizing the sum above as a Riemann sum.

f) Use a calculator or a computer or whatever you want to show that $\frac{2}{\pi}\int_0^\pi \frac{\sin x}{x}\,dx \approx 1.18$

These calculations show that the size of the "bumps" is 9% of the size of the jump in the function value. Gibbs showed that this number holds in general for functions in D.

Chapter 5

Lebesgue measure and integration

If you look back at what you have learned in your earlier mathematics courses, you will definitely recall a lot about area and volume — from the simple formulas for the areas of rectangles and triangles that you learned in grade school, to the quite sophisticated calculations with double and triple integrals that you had to perform in calculus class. What you have probably never seen, is a systematic theory for area and volume that unifies all the different methods and techniques. In this chapter we shall first study such a unified theory for d-dimensional volume based on the notion of a *measure*, and then we shall use this theory to build a stronger and more flexible theory for integration. You may think of this as a reversal of previous strategies; instead of basing the calculation of volumes on integration, we shall create a theory of integration based on a more fundamental notion of volume.

The theory will cover volume in $\mathbb{R}^d$ for all $d \in \mathbb{N}$, including $d = 1$ and $d = 2$. To get a unified terminology, we shall think of the length of a set in $\mathbb{R}$ and the area of a set in $\mathbb{R}^2$ as one- and two-dimensional volume, respectively.

To get a feeling for what we are aiming for, let us assume that we want to measure the volume of subsets $A \subseteq \mathbb{R}^3$, and that we denote the volume of A by $\mu(A)$. What properties would we expect μ to have?

(i) $\mu(A)$ should be a nonnegative number or ∞. There are subsets of $\mathbb{R}^3$ that have an infinite volume in an intuitive sense (such as $\mathbb{R}^3$ itself), and we capture this intuition by the symbol ∞.

(ii) $\mu(\emptyset) = 0$. It will be convenient to assign a volume to the empty set, and the only reasonable alternative is 0.

(iii) If $A_1, A_2, \ldots, A_n, \ldots$ are disjoint (i.e., non-overlapping) sets, then

$$\mu(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mu(A_n).$$

This means that the volume of the whole is equal to the sum of the volumes of the parts.

(iv) If $A = (a_1, a_2) \times (b_1, b_2) \times (c_1, c_2)$ is a rectangular box, then $\mu(A)$ is equal to the volume of A in the traditional sense, i.e.

$$\mu(A) = (a_2 - a_1)(b_2 - b_1)(c_2 - c_1)$$

It turns out that it is impossible to measure the size of *all* subsets of A such that all these requirements are satisfied; there are sets that are simply too irregular to be measured in a good way. For this reason we shall restrict ourselves to a class of *measurable sets* which behave the way we want. The hardest part of the theory will be to decide which sets are measurable.

We shall use a two step procedure to construct our measure μ: First we shall construct an *outer measure* μ^* which will assign a size $\mu^*(A)$ to *all* subsets $A \in \mathbb{R}^3$, but which will not satisfy all the conditions (i)-(iv) above. Then we shall use μ^* to single out the class of measurable sets, and prove that if we restrict μ^* to this class, our four conditions are satisfied.

5.1 Outer measure in $\mathbb{R}^d$

The first step in our construction is to define outer measure in $\mathbb{R}^d$. The outer measure is built from rectangular boxes, and we begin by intoducing the appropriate notation and teminology.

Definition 5.1.1 *A subset A of $\mathbb{R}^d$ is called an* open box *if there are numbers $a_1^{(1)} < a_2^{(1)}$, $a_1^{(2)} < a_2^{(2)}$, ..., $a_1^{(d)} < a_2^{(d)}$ such that*

$$A = (a_1^{(1)}, a_2^{(1)}) \times (a_1^{(2)}, a_2^{(2)}) \times \ldots \times (a_1^{(d)}, a_2^{(d)})$$

In addition, we count the empty set as a rectangular box. We define the volume $|A|$ *of A to be 0 if A is the empty set, and otherwise*

$$|A| = (a_2^{(1)} - a_1^{(1)})(a_2^{(2)} - a_1^{(2)}) \cdot \ldots \cdot (a_2^{(d)} - a_1^{(d)})$$

Observe that when $d = 1, 2$ and 3, $|A|$ denotes the length, area and volume of A in the usual sense.

If $\mathcal{A} = \{A_1, A_2, \ldots, A_n, \ldots\}$ is a countable collection of open boxes, we define its *size* $|\mathcal{A}|$ by

$$|\mathcal{A}| = \sum_{k=1}^{\infty} |A_k|$$

(we may clearly have $|\mathcal{A}| = \infty$). Note that we can think of a finite collection $\mathcal{A} = \{A_1, A_2, \ldots, A_n\}$ of open boxes as a countable one by putting in the empty set in the missing positions: $\mathcal{A} = \{A_1, A_2, \ldots, A_n, \emptyset, \emptyset, \ldots\}$. This is the main reason for including the empty set among the open boxes. Note also that since the boxes $A_1, A_2, \ldots$ may overlap, the size $|\mathcal{A}|$ need not be closely connected to the volume of $\bigcup_{n=1}^{\infty} A_n$.

A *covering*[1] of a set $B \subseteq \mathbb{R}^d$ is a countable collection

$$\mathcal{A} = \{A_1, A_2, \ldots, A_n, \ldots\}$$

of open boxes such that $B \subseteq \bigcup_{n=1}^{\infty} A_n$. We are now ready to define outer measure.

Definition 5.1.2 *The* outer measure *of a set* $B \in \mathbb{R}^d$ *is defined by*

$$\mu^*(B) = \inf\{|\mathcal{A}| : \mathcal{A} \textit{ is a covering of } B \textit{ by open boxes}\}$$

The idea behind outer measure should be clear – we measure the size of B by approximating it as economically as possible from the outside by unions of open boxes. You may wonder why we use open boxes and not closed boxes

$$A = [a_1^{(1)}, a_2^{(1)}] \times [a_1^{(2)}, a_2^{(2)}] \times \ldots \times [a_1^{(d)}, a_2^{(d)}]$$

in the definition above. The answer is that it does not really matter, but that open boxes are a little more convenient in some arguments. The following lemma tells us that closed boxes would have given us exactly the same result. You may want to skip the proof at the first reading.

Lemma 5.1.3 *For all* $B \subseteq \mathbb{R}^d$*,*

$$\mu^*(B) = \inf\{|\mathcal{A}| : \mathcal{A} \textit{ is a covering of } B \textit{ by } \text{closed} \textit{ boxes}\}$$

Proof: We must prove that

$$\inf\{|\mathcal{A}| : \mathcal{A} \text{ is a covering of } B \text{ by } \textit{open} \text{ boxes}\}$$

$$= \inf\{|\mathcal{A}| : \mathcal{A} \text{ is a covering of } B \text{ by } \textit{closed} \text{ boxes}\}$$

Observe first that if $\mathcal{A}_0 = \{A_1, A_2, \ldots\}$ is a covering of B by *open* boxes, we can get a covering $\mathcal{A} = \{\overline{A}_1, \overline{A}_2, \ldots\}$ of B by *closed* boxes just by closing each box. Since the two coverings have the same size, this means that

$$\mu^*(B) = \inf\{|\mathcal{A}| : \mathcal{A} \text{ is a covering of } B \text{ by } \textit{open} \text{ boxes}\}$$

$$\geq \inf\{|\mathcal{A}| : \mathcal{A} \text{ is a covering of } B \text{ by } \textit{closed} \text{ boxes}\}$$

To prove the opposite inequality, assume that $\epsilon > 0$ is given. If $\mathcal{A} = \{A_1, A_2, \ldots\}$ is a covering of B by closed boxes, we can for each n find

[1] This a related but different notion of covering than the one we used in Section 2.6 – there we used arbitrary covers of open sets, here we use countable covers of open rectangles

an open box $\tilde{A}_n$ containing A_n such that $|\tilde{A}_n| < |A_n| + \frac{\epsilon}{2^n}$. Then $\tilde{\mathcal{A}} = \{\tilde{A}_n\}$ is a covering of B by open boxes, and $|\tilde{\mathcal{A}}| < |\mathcal{A}| + \epsilon$. Since $\epsilon > 0$ is arbitrary, this shows that to any closed covering, there is an open covering arbitrarily close in size, and hence

$$\inf\{|\mathcal{A}| : \mathcal{A} \text{ is a covering of } B \text{ by } \textit{open} \text{ boxes}\}$$
$$\leq \inf\{|\mathcal{A}| : \mathcal{A} \text{ is a covering of } B \text{ by } \textit{closed} \text{ boxes}\}$$

□

Here are some properties of the outer measure:

Proposition 5.1.4 *The outer measure μ^* on $\mathbb{R}^d$ satisfies:*

(i) $\mu^*(\emptyset) = 0$.

(ii) (Monotonicity) If $B \subseteq C$, then $\mu^*(B) \leq \mu^*(C)$.

(iii) (Subadditivity) If $\{B_n\}_{n\in\mathbb{N}}$ is a sequence of subsets of $\mathbb{R}^d$, then

$$\mu^*(\bigcup_{n=1}^{\infty} B_n) \leq \sum_{n=1}^{\infty} \mu^*(B_n)$$

(iv) For all closed boxes

$$B = [b_1^{(1)}, b_2^{(1)}] \times [b_1^{(2)}, b_2^{(2)}] \times \ldots \times [b_1^{(d)}, b_2^{(d)}]$$

we have

$$\mu^*(B) = |B| = (b_2^{(1)} - b_1^{(1)})(b_2^{(2)} - b_1^{(2)}) \cdot \ldots \cdot (b_2^{(d)} - b_1^{(d)})$$

Proof: (i) Since $\mathcal{A} = \{\emptyset, \emptyset, \emptyset, \ldots\}$ is a covering of $\emptyset$, $\mu^*(\emptyset) = 0$.

(ii) Since any covering of C is a covering of B, we have $\mu^*(B) \leq \mu^*(C)$.

(iii) If $\mu^*(B_n) = \infty$ for some $n \in \mathbb{N}$, there is nothing to prove, and we may hence assume that $\mu^*(B_n) < \infty$ for all n. Let $\epsilon > 0$ be given. For each $n \in \mathbb{N}$, we can find a covering $A_1^{(n)}, A_2^{(n)}, \ldots$ of B_n such that

$$\sum_{k=1}^{\infty} |A_k^{(n)}| < \mu^*(B_n) + \frac{\epsilon}{2^n}$$

The collection $\{A_k^{(n)}\}_{k,n\in\mathbb{N}}$ of all sets in all coverings is a countable covering of $\bigcup_{n=1}^{\infty} B_n$, and

$$\sum_{k,n\in\mathbb{N}} |A_k^{(n)}| = \sum_{n=1}^{\infty} \left(\sum_{k=1}^{\infty} |A_k^{(n)}| \right) \leq \sum_{n=1}^{\infty} \left(\mu^*(B_n) + \frac{\epsilon}{2^n} \right) = \sum_{n=1}^{\infty} \mu^*(B_n) + \epsilon$$

(if you are unsure about these manipulation, take a look at exercise 5). This means that

$$\mu^*(\bigcup_{n=1}^{\infty} B_n) \leq \sum_{n=1}^{\infty} \mu^*(B_n) + \epsilon$$

and since ϵ is an arbitrary, positive number, we must have

$$\mu^*(\bigcup_{n=1}^{\infty} B_n) \leq \sum_{n=1}^{\infty} \mu^*(B_n)$$

(iv) Since we can cover B by $\mathcal{B}_\epsilon = \{B_\epsilon, \emptyset, \emptyset, \ldots\}$, where

$$B_\epsilon = (b_1^{(1)} - \epsilon, b_2^{(1)} + \epsilon) \times (b_1^{(2)} - \epsilon, b_2^{(2)} + \epsilon) \times \ldots \times (b_1^{(d)} - \epsilon, b_2^{(d)} + \epsilon),$$

for any $\epsilon > 0$, we se that

$$\mu^*(B) \leq |B| = (b_2^{(1)} - b_1^{(1)})(b_2^{(2)} - b_1^{(2)}) \cdot \ldots \cdot (b_2^{(d)} - b_1^{(d)})$$

The opposite inequality,

$$\mu^*(B) \geq |B| = (b_2^{(1)} - b_1^{(1)})(b_2^{(2)} - b_1^{(2)}) \cdot \ldots \cdot (b_2^{(d)} - b_1^{(d)})$$

may seem obvious (it just says that it is impossible to cover the box B by boxes whose total volume is less than the volume of B), but is actually a little tricky to prove. We shall need a few lemmas to establish it and finish the proof. □

I shall carry out the remaining part of the proof of Proposition 5.1.4(iv) in the three dimensional case. The proof is exactly the same in the d-dimensional case, but the notation becomes so messy that it tends to blur the underlying ideas. Let us begin with a lemma.

Lemma 5.1.5 *Assume that the intervals (a_0, a_K), (b_0, b_N), (c_0, c_M) are partioned*

$$a_0 < a_1 < a_2 < \ldots < a_K$$

$$b_0 < b_1 < b_2 < \ldots < b_N$$

$$c_0 < c_1 < c_2 < \ldots < c_M$$

and let $\Delta a_k = a_{k+1} - a_k$, $\Delta b_n = a_{n+1} - n_n$, $\Delta c_m = c_{m+1} - c_m$. Then

$$(a_K - a_0)(b_N - b_0)(c_m - c_0) = \sum_{k,n,m} \Delta a_k \Delta b_n \Delta c_m$$

where the sum is over all triples (k, n, m) such that $0 \leq k < K$, $0 \leq n < N$, $0 \leq m < M$. In other words, if we partition the box

$$A = (a_0, a_K) \times (b_0, b_N) \times (c_0, c_M)$$

into KNM smaller boxes $B_1, B_2, \ldots, B_{KNM}$, then

$$|A| = \sum_{j=1}^{KNM} |B_j|$$

Proof: If you think geometrically, the lemma seems obvious — it just says that if you divide a big box into smaller boxes, the volume of the big box is equal to the sum of the volumes of the smaller boxes. An algebraic proof is not much harder and has the advantage of working also in higher dimensions: Note that since $a_K - a_0 = \sum_{k=0}^{K-1} \Delta a_k$, $b_N - b_0 = \sum_{n=0}^{N-1} \Delta b_n$, $c_M - c_0 = \sum_{m=0}^{M-1} \Delta c_m$, we have

$$(a_K - a_0)(b_N - b_0)(c_m - c_0) =$$

$$= \left(\sum_{k=0}^{K-1} \Delta a_k\right)\left(\sum_{n=0}^{N-1} \Delta b_n\right)\left(\sum_{m=0}^{M-1} \Delta c_m\right) =$$

$$= \sum_{k,n,m} \Delta a_k \Delta b_n \Delta c_m$$

□

The next lemma reduces the problem from countable coverings to finite ones. It is the main reason why we have chosen to work with open coverings (If you have read section 2.6, you will see that this result is an immediate consequence of Theorem 2.6.4).

Lemma 5.1.6 *Assume that $\mathcal{A} = \{A_1, A_2, \ldots, A_n, \ldots\}$ is a countable covering of a compact set K by open boxes. Then K is covered by a finite number $A_1, A_2, \ldots, A_n$ of elements in $\mathcal{A}$.*

Proof: Assume not, then we can for each $n \in \mathbb{N}$ find an element $x_n \in K$ which does not belong to $\bigcup_{k=1}^{n} A_k$. Since K is compact, there is a subsequence $\{x_{n_k}\}$ converging to an element $x \in K$. Since $\mathcal{A}$ is a covering of K, x must belong to an A_i. Since A_i is open, $x_{n_k} \in A_i$ for all sufficently large k. But this is impossible since $x_{n_k} \notin A_i$ when $n_k \geq i$. □

We are now ready to prove the missing inequality in Proposition 5.1.4(iv).

Lemma 5.1.7 *For all closed boxes*

$$B = [a_1, a_2] \times [b_1, b_2] \times [c_1, c_2]$$

we have

$$\mu^*(B) \geq |B| = (a_2 - a_1)(b_2 - b_1)(c_2 - c_1)$$

Proof: By the lemma above, it suffices to show that if $A_1, A_2, \ldots, A_n$ is a finite covering of B, then

$$|B| \leq |A_1| + |A_2| + \ldots + |A_n|$$

Let

$$B^\circ = (a_1, a_2) \times (b_1, b_2) \times (c_1, c_2)$$

be the *open* box with the same dimensions as B, and let $A'_i = A_i \cap B^\circ$ be the boxes we get by just looking at the parts of $A_1, A_2, \ldots, A_n$ that lie inside B°. The boxes $A'_1\, A'_2, \ldots, A'_n$ are no longer a covering of B, but they are a covering of B°, and since $|A'_i| \leq |A_i|$, it suffices to prove that

$$|B| = |B^\circ| \leq |A'_1| + |A'_2| + \ldots + |A'_n|$$

Let

$$A'_i = (x_1^{(i)}, x_2^{(i)}) \times (y_1^{(i)}, y_2^{(i)}) \times (z_1^{(i)}, z_2^{(i)})$$

We collect all x-coordinates $x_1^{(1)}, x_2^{(1)}, x_1^{(2)}, x_2^{(2)}, \ldots, x_1^{(n)}, x_2^{(n)}$ and rearrange them according to size:

$$x_0 < x_1 < x_2 < \ldots < x_I$$

Doing the same with the y- and the z-coodinates, we get partitions

$$y_0 < y_1 < y_2 < \ldots < y_J$$

$$z_0 < z_1 < z_2 < \ldots < z_K$$

Let $B_1, B_2, \ldots, B_P$ be all boxes of the form $(x_i, x_{i+1}) \times (y_j, y_{j+1}) \times (z_k, z_{k+1})$. Each A'_i is "made up of" a finite number of B_j's, and according to Lemma 5.1.5,

$$|A'_i| = |B_{j_{i_1}}| + |B_{j_{i_2}}| + \ldots + |B_{j_{i_q}}|$$

where $B_{j_{i_1}}, B_{j_{i_2}}, \ldots, B_{j_{i_q}}$ are the small boxes making up A'_i. If we sum over all i (and use that each B_j lie inside at least one A'_i), we see that

$$\sum_{i=1}^{n} |A'_i| \geq \sum_{j=1}^{P} |B_j|$$

(we get an inequality as some of the B_j's may belong to more than one A'_i, and hence are counted twice or more on the left hand side).

On the other hand, since the B_j's make up the whole box B°, Lemma 5.1.5 also tells us that

$$|B^\circ| = \sum_{k=1}^{P} |B_j|$$

Hence

$$|B| = |B^\circ| = \sum_{j=1}^{P} |B_j| \leq \sum_{i=1}^{n} |A_i'|$$

and the lemma is proved. □

We have now finally established all parts of Proposition 5.1.4. and are ready to move on. The problem with the outer measure μ^* is that it fails to be countably additive: If $\{A_n\}$ is a disjoint sequence of sets, we can only guarantee that

$$\mu^*(\bigcup_{n=1}^{\infty} A_n) \leq \sum_{n=1}^{\infty} \mu^*(A_n)$$

not that

$$\mu^*(\bigcup_{n=1}^{\infty} A_n) = \sum_{n=1}^{\infty} \mu^*(A_n) \tag{5.1.1}$$

As it is impossible to change μ^* such that (5.1.1) holds for all disjoint sequences $\{A_n\}$ of subsets of $\mathbb{R}^d$, we shall follow a different strategy: We shall show that there is a large class $\mathcal{M}$ of subsets of $\mathbb{R}^d$ such that (5.1.1) holds whenever $A_n \in \mathcal{M}$ for all $n \in \mathbb{N}$. The sets in $\mathcal{M}$ will be called measurable sets.

Exercises for Section 5.1

1. Show that all countable sets have outer measure zero.
2. Show that the x-axis has outer measure 0 in $\mathbb{R}^2$.
3. If A is a subset of $\mathbb{R}^d$ and $\mathbf{b} \in \mathbb{R}^d$, we define

   $$A + \mathbf{b} = \{\mathbf{a} + \mathbf{b} \mid \mathbf{a} \in A\}$$

   Show that $\mu^*(A + \mathbf{b}) = \mu^*(A)$.
4. If A is a subset of $\mathbb{R}^d$, define $2A = \{2\mathbf{a} \mid \mathbf{a} \in A\}$. Show that $\mu^*(2A) = 2^d \mu^*(A)$.
5. Let $\{a_{n,k}\}_{n,k\in\mathbb{N}}$ be a collection of nonnegative, real numbers, and let a be the supremum over all finite sums of distinct elements in this collection, i.e.

   $$A = \sup\{\sum_{i=1}^{I} a_{n_i,k_i} \ : \ I \in \mathbb{N} \text{ and all pairs } (n_1, k_1), \dots, (n_I, k_I) \text{ are different}\}$$

   a) Assume that $\{b_m\}_{m\in\mathbb{N}}$ is a sequence which contains each element in the set $\{a_{n,k}\}_{n,k\in\mathbb{N}}$ exactly ones. Show that $\sum_{m=1}^{\infty} b_m = a$.
   b) Show that $\sum_{n=1}^{\infty} (\sum_{k=1}^{\infty} a_{n,k}) = a$.
   c) Comment on the proof of Proposition 5.1.4(iii).

5.2 Measurable sets

We shall now begin our study of measurable sets — the sets that can be assigned a "volume" in a coherent way. The definition is rather mysterious:

Definition 5.2.1 *A subset E of X is called* measurable *if*

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

for all $A \subseteq X$. The collection of all measurable sets is denoted by $\mathcal{M}$.

Although the definition above is easy to grasp, it is not easy too see why it captures the essence of the sets that are possible to measure. The best I can say, is that the reason why some sets are impossible to measure, is that they have very irregular boundaries. The definition above says that a set is measurable if we can use it to split any other set in two without introducing any further irregularities, i.e., all parts of its boundary must be reasonably regular. Admittedly, this explanation is vague and not very helpful in understanding why the definition captures exactly the right notion of measurability. The best argument may simply be to show that the definition works, so let us get started.

Let us first of all make a very simple observation. Since $A = (A \cap E) \cup (A \cap E^c)$, subadditivity (recall Proposition 5.1.4(iii)) tells us that we always have

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

Hence to prove that a set is measurable, we only need to prove that

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

Our first observation on measurable sets is simple.

Lemma 5.2.2 *If E has outer measure 0, then E is measurable. In particular, $\emptyset \in \mathcal{M}$.*

Bevis: If E has measure 0, so has $A \cap E$ since $A \cap E \subseteq E$. Hence

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

for all $A \subseteq X$. □

Next we have:

Proposition 5.2.3 *$\mathcal{M}$ is an algebra of sets, i.e.:*

(i) $\emptyset \in \mathcal{M}$.

(ii) *If $E \in \mathcal{M}$, then $E^c \in \mathcal{M}$.*

(iii) *If* $E_1, E_2, \ldots, E_n \in \mathcal{M}$*, then* $E_1 \cup E_2 \cup \ldots \cup E_n \in \mathcal{M}$*.*

(iv) *If* $E_1, E_2, \ldots, E_n \in \mathcal{M}$*, then* $E_1 \cap E_2 \cap \ldots \cap E_n \in \mathcal{M}$*.*

Proof: We have already proved (i), and (ii) is obvious from the definition of measurable sets. Since $E_1 \cup E_2 \cup \ldots \cup E_n = (E_1^c \cap E_2^c \cap \ldots \cap E_n^c)^c$ by De Morgan's laws, (iii) follows from (ii) and (iv). Hence it remains to prove (iv).

To prove (iv) is suffices to prove that if $E_1, E_2 \in \mathcal{M}$, then $E_1 \cap E_2 \in \mathcal{M}$ as we can then add more sets by induction. If we first use the measurability of E_1, we see that for any set $A \subseteq X$

$$\mu^*(A) = \mu^*(A \cap E_1) + \mu^*(A \cap E_1^c)$$

Using the measurability of E_2, we get

$$\mu^*(A \cap E_1) = \mu^*((A \cap E_1) \cap E_2) + \mu^*((A \cap E_1) \cap E_2^c)$$

Combining these two expressions, we have

$$\mu^*(A) = \mu^*((A \cap (E_1 \cap E_2)) + \mu^*((A \cap E_1) \cap E_2^c) + \mu^*(A \cap E_1^c)$$

Observe that (draw a picture!)

$$(A \cap E_1 \cap E_2^c) \cup (A \cap E_1^c) = A \cap (E_1 \cap E_2)^c$$

and hence by subadditivity

$$\mu^*(A \cap E_1 \cap E_2^c) + \mu^*(A \cap E_1^c) \geq \mu^*(A \cap (E_1 \cap E_2)^c)$$

Putting this into the expression for $\mu^*(A)$ above, we get

$$\mu^*(A) \geq \mu^*((A \cap (E_1 \cap E_2)) + \mu^*(A \cap (E_1 \cap E_2)^c)$$

which means that $E_1 \cap E_2 \in \mathcal{M}$. □

We would like to extend parts (iii) and (iv) in the proposition above to countable unions and intersection. For this we need the following lemma:

Lemma 5.2.4 *If* $E_1, E_2, \ldots, E_n$ *is a disjoint collection of measurable sets, then*

$$\mu^*(A \cap (E_1 \cup E_2 \cup \ldots \cup E_n)) = \mu^*(A \cap E_1) + \mu^*(A \cap E_2) + \ldots + \mu^*(A \cap E_n)$$

Proof: It suffices to prove the lemma for two sets E_1 and E_2 as we can then extend it by induction. Using the measurability of E_1, we see that

$$\mu^*(A \cap (E_1 \cup E_2)) = \mu^*((A \cap (E_1 \cup E_2)) \cap E_1) + \mu^*(A \cap (E_1 \cup E_2)) \cap E_1^c) =$$
$$= \mu^*(A \cap E_1) + \mu^*(A \cap E_2)$$

□

We can now prove that $\mathcal{M}$ is closed under countable unions.

Lemma 5.2.5 *If $A_n \in \mathcal{M}$ for each $n \in \mathbb{N}$, then $\bigcup_{n\in\mathbb{N}} A_n \in \mathcal{M}$.*

Proof: Note that since $\mathcal{M}$ is an algebra,

$$E_n = A_n \cap (E_1 \cup E_2 \cup \ldots E_{n-1})^c$$

belongs to $\mathcal{M}$ for $n > 1$ (for $n = 1$, we just let $E_1 = A_1$). The new sequence $\{E_n\}$ is disjoint and have the same union as $\{A_n\}$ (make a drawing!), and hence it suffices to prove that $\bigcup_{n\in\mathbb{N}} E_n \in \mathcal{M}$, i.e.

$$\mu^*(A) \geq \mu^*\big(A \cap \bigcup_{n=1}^{\infty} E_n\big) + \mu^*\big(A \cap \big(\bigcup_{n=1}^{\infty} E_n\big)^c\big)$$

Since $\bigcup_{n=1}^{N} E_n \in \mathcal{M}$ for all $N \in \mathbb{N}$, we have:

$$\mu^*(A) = \mu^*\big(A \cap \bigcup_{n=1}^{N} E_n\big) + \mu^*\big(A \cap \big(\bigcup_{n=1}^{N} E_n\big)^c\big) \geq$$

$$\geq \sum_{n=1}^{N} \mu^*(A \cap E_n) + \mu^*\big(A \cap \big(\bigcup_{n=1}^{\infty} E_n\big)^c\big)$$

where we in the last step have used the lemma above plus the observation that $\big(\bigcup_{n=1}^{\infty} E_n\big)^c \subseteq \big(\bigcup_{n=1}^{N} E_n\big)^c$. Since this inequality holds for all $N \in \mathbb{N}$, we get

$$\mu^*(A) \geq \sum_{n=1}^{\infty} \mu^*(A \cap E_n) + \mu^*\big(A \cap \big(\bigcup_{n=1}^{\infty} E_n\big)^c\big)$$

By subadditivity we have $\sum_{n=1}^{\infty} \mu^*(A \cap E_n) \geq \mu^*(\bigcup_{n=1}^{\infty}(A \cap E_n)) = \mu^*(A \cap \bigcup_{n=1}^{\infty}(E_n))$, and hence

$$\mu^*(A) \geq \mu^*\big(A \cap \bigcup_{n=1}^{\infty} E_n\big) + \mu^*\big(A \cap \big(\bigcup_{n=1}^{\infty} E_n\big)^c\big)$$

□

Let us sum up our results so far.

Theorem 5.2.6 *The measurable sets $\mathcal{M}$ form a σ-algebra, i.e.:*

(i) $\emptyset \in \mathcal{M}$

(ii) If $E \in \mathcal{M}$, then $E^c \in \mathcal{M}$.

(iii) If $E_n \in \mathcal{M}$ for all $n \in \mathbb{N}$, then $\bigcup_{n=1}^{\infty} E_n \in \mathcal{M}$.

(iv) If $E_n \in \mathcal{M}$ for all $n \in \mathbb{N}$, then $\bigcap_{n=1}^{\infty} E_n \in \mathcal{M}$.

Proof: We have proved everything except (iv), which follows from (ii) and (iii) since $\bigcap_{n=1}^{\infty} E_n = \left(\bigcup_{n=1}^{\infty} E_n^c\right)^c$. □

Remark: By definition, a σ-algebra is a collection of subsets satisfying (i)-(iii), but — as we have seen — point (iv) follows from the others.

There is one more thing we have to check: that M contains sufficiently many sets. So far we only know that $\mathcal{M}$ contains the sets of outer measure 0 and their complements!

In the first proof it is convenient to use *closed* coverings as in Lemma 5.1.3 to determine the outer measure.

Lemma 5.2.7 *For each i and each $a \in \mathbb{R}$, the open halfspaces*

$$H = \{(x_1, \ldots, x_i, \ldots, x_d) \in \; X \; : \; x_i < a\}$$

and

$$K = \{(x_1, \ldots, x_i, \ldots, x_d) \in \mathbb{R}^d \; : \; x_i > a\}$$

are measurable.

Proof: We only prove the H-part. We have to check that for any $B \subseteq \mathbb{R}^d$,

$$\mu^*(B) \geq \mu^*(B \cap H) + \mu^*(B \cap H^c)$$

Given a covering $\mathcal{A} = \{A_i\}$ of B by closed boxes, we can create closed coverings $\mathcal{A}^{(1)} = \{A_i^{(1)}\}$ and $\mathcal{A}^{(1)} = \{A_i^{(2)}\}$ of $B \cap H$ and $B \cap H^c$, respectively, by putting

$$A_i^{(1)} = \{(x_1, \ldots, x_i, \ldots, x_d) \in A_i \; : \; x_i \leq a\}$$

$$A_i^{(2)} = \{(x_1, \ldots, x_i, \ldots, x_d) \in A_i \; : \; x_i \geq a\}$$

Hence

$$|\mathcal{A}| = |\mathcal{A}^{(1)}| + |\mathcal{A}^{(2)}| \geq \mu^*(B \cap H) + \mu^*(B \cap H^c)$$

and since this holds for all closed coverings $\mathcal{A}$ of B, we get

$$\mu^*(B) \geq \mu^*(B \cap H) + \mu^*(B \cap H^c)$$

□

The next step is now easy:

Lemma 5.2.8 *All open boxes are measurable.*

Proof: An open box is a finite intersection of open halfspaces. □

The next result tells us that there are many measurable sets:

Theorem 5.2.9 *All open sets in $\mathbb{R}^d$ are countable unions of open boxes. Hence all open and closed sets are measurable.*

Proof: Note first that the measurability of closed sets follows from the measurability of open sets since a closed set is the complement of an open set. To prove the theorem for open sets, let us first agree to call an open box

$$A = (a_1^{(1)}, a_2^{(1)}) \times (a_1^{(2)}, a_2^{(2)}) \times \ldots \times (a_1^{(d)}, a_2^{(d)})$$

rational if all the coordinates $a_1^{(1)}, a_2^{(1)}, a_1^{(2)}, a_2^{(2)}, \ldots, a_1^{(d)}, a_2^{(d)}$ are rational. There are only countably many rationals boxes, and hence we only need to prove that if G is an open set, then

$$G = \bigcup \{B \ : \ B \text{ is a rational box contained in } G\}$$

We leave the details to the reader. □

Exercises for Section 5.2

1. Show that if $A, B \in \mathcal{M}$, then $A \setminus B \in \mathcal{M}$.
2. Explain in detail why 5.2.3(iii) follows from (ii) and (iv).
3. Carry out the induction step in the proof of Proposition 5.2.3(iv).
4. Explain the equality $(A \cap E_1 \cap E_2^c) \cup (A \cap E_1^c) = A \cap (E_1 \cap E_2)^c$ in the proof of Lemma 5.2.3.
5. Carry out the induction step in the proof of Lemma 5.2.4.
6. Explain in detail why (iv) follows from (ii) and (iii) in Theorem 5.2.6.
7. Show that all *closed* halfspaces

   $$H = \{(x_1, \ldots, x_i, \ldots, x_d) \in \mathbb{R}^d \ : \ x_i \leq a\}$$

   and

   $$K = \{(x_1, \ldots, x_i, \ldots, x_d) \in \mathbb{R}^d \ : \ x_i \geq a\}$$

   are measurable
8. Recall that if A is a subset of $\mathbb{R}^d$ and $\mathbf{b} \in \mathbb{R}^d$, then

   $$A + \mathbf{b} = \{\mathbf{a} + \mathbf{b} \mid \mathbf{a} \in A\}$$

   Show that $A + \mathbf{b}$ is measurable if and only if A is.
9. If A is a subset of $\mathbb{R}^d$, define $2A = \{2\mathbf{a} \mid \mathbf{a} \in A\}$. Show that $2A$ is measurable if and only if A is.
10. Fill in the details in the proof of Lemma 5.2.8.
11. Complete the proof of Theorem 5.2.9.

5.3 Lebesgue measure

Having constructed the outer measure μ^* and explored its basic properties, we are now ready to define the measure μ.

Definition 5.3.1 *The* Lebesgue measure μ *is the restriction of the outer measure* μ^* *to the measurable sets, i.e. it is the function*

$$\mu : \mathcal{M} \to [0, \infty]$$

defined by

$$\mu(A) = \mu^*(A)$$

for all $A \in \mathcal{M}$.

Remark: Since μ and μ^* are essentially the same function, you may wonder why we have introduced a new symbol for the Lebesgue measure. The answer is that although it is not going to make much of a difference for us here, it is convenient to distinguish between the two in more theoretical studies of measurability. All you have to remember for this text, is that $\mu(A)$ and $\mu^*(A)$ are defined and equal as long as A is measurable.

We can now prove that μ has the properties we asked for at the beginning of the chapter:

Theorem 5.3.2 *The Lebesgue measure* $\mu : \mathcal{M} \to [0, \infty]$ *has the following properties:*

(i) $\mu(\emptyset) = 0$.

(ii) (Completeness) Assume that $A \in \mathcal{M}$, *and that* $\mu(A) = 0$. *Then all subset* $B \subseteq A$ *are measurable, and* $\mu(B) = 0$.

(iv) (Countable subadditivity) If $\{A_n\}_{n\in\mathbb{N}}$ *is a sequence of measurable sets, then*

$$\mu\Big(\bigcup_{n=1}^{\infty} A_n\Big) \leq \sum_{n=1}^{\infty} \mu(A_n)$$

(iv) (Countable additivity) If $\{E_n\}_{n\in\mathbb{N}}$ *is a disjoint sequence of measurable sets, then*

$$\mu\Big(\bigcup_{n=1}^{\infty} E_n\Big) = \sum_{n=1}^{\infty} \mu(E_n)$$

(v) For all closed boxes

$$B = [b_1^{(1)}, b_2^{(1)}] \times [b_1^{(2)}, b_2^{(2)}] \times \ldots \times [b_1^{(d)}, b_2^{(d)}]$$

we have

$$\mu(B) = |B| = (b_2^{(1)} - b_1^{(1)})(b_2^{(2)} - b_1^{(2)}) \cdot \ldots \cdot (b_2^{(d)} - b_1^{(d)})$$

Proof: (i) and (ii) follow from Lemma 5.2.2, and (iii) follows from part (iii) of Proposisition 5.1.4 since $\mathcal{M}$ is a σ-algebra, and $\bigcup_{n=1}^{\infty} A_n$ hence is measurable. Since we know from Theorem 5.2.9 that closed boxes are measurable, part (v) follows from Proposition 5.1.4(iv).

To prove (iv), we first observe that

$$\mu(\bigcup_{n=1}^{\infty} E_n) \leq \sum_{n=1}^{\infty} \mu(E_n)$$

by (iii). To get the opposite inequality, we use Lemma 5.2.4 with $A = \mathbb{R}^d$ to see that

$$\sum_{n=1}^{N} \mu(E_n) = \mu(\bigcup_{n=1}^{N} E_n) \leq \mu(\bigcup_{n=1}^{\infty} E_n)$$

Since this holds for all $N \in \mathbb{N}$, we must have

$$\sum_{n=1}^{\infty} \mu(E_n) \leq \mu(\bigcup_{n=1}^{\infty} E_n)$$

Hence we have both inequalities, and (iii) is proved. □

In what follows, we shall often need the following simple lemma:

Lemma 5.3.3 *If C, D are measurable sets such that $C \subseteq D$ and $\mu(D) < \infty$, then*

$$\mu(D \setminus C) = \mu(D) - \mu(C)$$

Proof: By additivity

$$\mu(D) = \mu(C) + \mu(D \setminus C)$$

Since $\mu(D)$ is finite, so is $\mu(C)$, and it makes sense to subtract $\mu(C)$ on both sides to get

$$\mu(D \setminus C) = \mu(D) - \mu(C)$$

□

The next properties are often referred to as *continuity of measure*:

Proposition 5.3.4 *Let $\{A_n\}_{n\in\mathbb{N}}$ be a sequence of measurable sets.*

(i) If the sequence is increasing (i.e. $A_n \subseteq A_{n+1}$ for all n), then

$$\mu(\bigcup_{n=1}^{\infty} A_n) = \lim_{n\to\infty} \mu(A_n)$$

(ii) If the sequence is decreasing (i.e. $A_n \supseteq A_{n+1}$ for all n), and $\mu(A_1)$ is finite, then

$$\mu(\bigcap_{n=1}^{\infty} A_n) = \lim_{n\to\infty} \mu(A_n)$$

Proof: (i) If we put $E_1 = A_1$ and $E_n = A_n \setminus A_{n-1}$ for n> 1, the sequence $\{E_n\}$ is disjoint, and $\bigcup_{k=1}^{n} E_k = A_n$ for all N (make a drawing). Hence

$$\mu(\bigcup_{n=1}^{\infty} A_n) = \mu(\bigcup_{n=1}^{\infty} E_n) = \sum_{n=1}^{\infty} \mu(E_n) =$$

$$= \lim_{n\to\infty} \sum_{k=1}^{n} \mu(E_k) = \lim_{n\to\infty} \mu(\bigcup_{k=1}^{n} E_k) = \lim_{n\to\infty} \mu(A_n)$$

where we have used the additivity of μ twice.

(ii) We first observe that $\{A_1 \setminus A_n\}_{n\in\mathbb{N}}$ is an increasing sequence of sets with union $A_1 \setminus \bigcap_{n=1}^{\infty} A_n$. By part (ii), we thus have

$$\mu(A_1 \setminus \bigcap_{n=1}^{\infty} A_n) = \lim_{n\to\infty} \mu(A_1 \setminus A_n)$$

Applying Lemma 5.3.3 on both sides, we get

$$\mu(A_1) - \mu(\bigcap_{n=1}^{\infty} A_n) = \lim_{n\to\infty} (\mu(A_1) - \mu(A_n)) = \mu(A_1) - \lim_{n\to\infty} \mu(A_n)$$

Since $\mu(A_1)$ is finite, we get $\mu(\bigcap_{n=1}^{\infty} A_n) = \lim_{n\to\infty} \mu(A_n)$, as we set out to prove. □

Remark: The finiteness condition in part (ii) may look like an unnecessary consequence of a clumsy proof, but it is actually needed. To see why, let μ be Lebesgue measure in $\mathbb{R}$, and let $A_n = [n, \infty)$. Then $\mu(A_n) = \infty$ for all n, but $\mu(\bigcap_{n=1}^{\infty} A_n) = \mu(\emptyset) = 0$. Hence $\lim_{n\to\infty} \mu(A_n) \neq \mu(\bigcap_{n=1}^{\infty} A_n)$.

Example 1: We know already that *closed* boxes have the "right" measure (Theorem 5.3.2 (iv)), but what about *open* boxes? If

$$B = (b_1^{(1)}, b_2^{(1)}) \times (b_1^{(2)}, b_2^{(2)}) \times \ldots \times (b_1^{(d)}, b_2^{(d)})$$

is an open box, let B_n be the closed box

$$B_n = \left[b_1^{(1)} + \frac{1}{n}, b_2^{(1)} - \frac{1}{n}\right] \times \left[b_1^{(2)} + \frac{1}{n}, b_2^{(2)} - \frac{1}{n}\right] \times \ldots \times \left[b_1^{(d)} + \frac{1}{n}, b_2^{(d)} - \frac{1}{n}\right]$$

obtained by moving all walls a distance $\frac{1}{n}$ inwards. By the proposition,

$$\mu(B) = \lim_{n\to\infty} \mu(B_n)$$

and since the closed boxes B_n have the "right" measure, it follows that so does the open box B. ♣

Example 2: Let

$$K_n = [-n, n]^d$$

be the closed box centered at the origin and with edges of length $2n$. For any measurable set A, it follows from the proposition above that

$$\mu(A) = \lim_{n\to\infty} \mu(A \cap K_n)$$

♣

We shall need one more property of measurable sets. It tells us that measurable sets can be approximated from the outside by open sets and from the inside by closed sets.

Proposition 5.3.5 *Assume that $A \subseteq \mathbb{R}^d$ is a measurable set. For each $\epsilon > 0$, there is an open set $G \supseteq A$ such that $\mu(G \setminus A) < \epsilon$, and a closed set $F \subseteq A$ such that $\mu(A \setminus F) < \epsilon$.*

Proof: We begin with the open sets. Assume first A has finite measure. Then there is a covering $\{B_n\}$ of A by open rectangles such that

$$\sum_{n=1}^{\infty} |B_n| < \mu(A) + \epsilon$$

Since $\mu(\bigcup_{n=1}^{\infty} B_n) \leq \sum_{n=1}^{\infty} \mu(B_n) = \sum_{n=1}^{\infty} |B_n|$, we see that $G = \bigcup_{n=1}^{\infty} B_n$ is an open set such that $A \subseteq G$, and $\mu(G) < \mu(A) + \epsilon$. Hence

$$\mu(G \setminus A) = \mu(G) - \mu(A) < \epsilon$$

by Lemma 5.3.3.

If $\mu(A)$ is infinite, we first use the boxes K_n in Example 2 to slice A into pieces of finite measure. More precisely, we let $A_n = A \cap (K_n \setminus K_{n-1})$, and use what we have already proved to find an open set G_n such that $A_n \subseteq G_n$ and $\mu(G_n \setminus A_n) < \frac{\epsilon}{2^n}$. Then $G = \bigcup_{n=1}^{\infty} G_n$ is an open set which contains A, and since $G \setminus A \subseteq \bigcup_{n=1}^{\infty}(G_n \setminus A_n)$, we get

$$\mu(G \setminus A) \leq \sum_{n=1}^{\infty} \mu(G_n \setminus A_n) < \sum_{n=1}^{\infty} \frac{\epsilon}{2^n} = \epsilon,$$

proving the statement about approximation by open sets.

To prove the statement about closed sets, just note that if we apply the first part of the theorem to A^c, we get an open set $G \supseteq A^c$ such that $\mu(G \setminus A^c) < \epsilon$. This means that $F = G^c$ is a closed set such that $F \subseteq A$, and since $A \setminus F = G \setminus A^c$, we have $\mu(A \setminus F) < \epsilon$. □

We have now established the basic properties of the Lebesgue measure. For the remainder of the chapter, you may forget about the construction of the measure and concentrate on the results of this section plus the properties of measurable sets summed up in theorems 5.2.6 and 5.2.9 of the previous section.

Exercises for Section 5.3

1. Explain that $A \setminus F = G \setminus A^c$ and the end of the proof of Proposition 5.3.5.
2. Show that if E_1, E_2 are measurable, then
$$\mu(E_1) + \mu(E_2) = \mu(E_1 \cup E_2) + \mu(E_1 \cap E_2)$$
3. The *symmetric difference* $A \triangle B$ of two sets A, B is defined by
$$A \triangle B = (A \setminus B) \cup (B \setminus A)$$
A subset of X is called a $\mathcal{G}_\delta$-set if it is the intersection of countably many open sets, and it is called a $\mathcal{F}_\sigma$-set if it is union of countably many closed set.
   a) Show that if A and B are measurable, then so is $A \triangle B$.
   b) Explain why all $\mathcal{G}_\delta$- and $\mathcal{F}_\sigma$-sets are measurable.
   c) Show that if A is measurable, there is a $\mathcal{G}_\delta$-set G such that $\mu(A \triangle G) = 0$.
   d) Show that if A is measurable, there is a $\mathcal{F}_\sigma$-set F such that $\mu(A \triangle F) = 0$.
4. Assume that $A \in \mathcal{M}$ has finite measure. Show that for every $\epsilon > 0$, there is a compact set $K \subseteq A$ such that $\mu(A \setminus K) < \epsilon$.
5. Assume that $\{A_n\}$ is a countable sequence of measurable sets, and assume that $\sum_{n=1}^{\infty} \mu(A_n) < \infty$. Show that the set
$$A = \{x \in X \,| x \text{ belongs to infinitely many } A_n\}$$
has measure zero.

5.4 Measurable functions

Before we turn to integration, we need to look at the functions we hope to integrate, the *measurable* functions. As functions taking the values $\pm\infty$ will occur naturally as limits of sequences of ordinary functions, we choose to include them from the beginning; hence we shall study functions
$$f : \mathbb{R}^d \to \overline{\mathbb{R}}$$
where $\overline{\mathbb{R}} = \mathbb{R} \cup \{-\infty, \infty\}$ is the set of *extended real numbers*. Don't spend too much effort on trying to figure out what $-\infty$ and ∞ "really" are — they are just convenient symbols for describing divergence.

To some extent we may extend ordinary algebra to $\overline{\mathbb{R}}$, e.g., we shall let
$$\infty + \infty = \infty, \quad -\infty - \infty = -\infty$$
and
$$\infty \cdot \infty = \infty, \quad (-\infty) \cdot \infty = -\infty, \quad (-\infty) \cdot (-\infty) = \infty.$$

If $r \in \mathbb{R}$, we similarly let

$$\infty + r = \infty, \quad -\infty + r = -\infty$$

For products, we have to take the sign of r into account, hence

$$\infty \cdot r = \begin{cases} \infty & \text{if } r > 0 \\ \\ -\infty & \text{if } r < 0 \end{cases}$$

and similarly for $(-\infty) \cdot r$.

All the rules above are natural and intuitive. Expressions that do not have an intuitive interpretation, are usually left undefined, e.g. is $\infty - \infty$ *not* defined. There is one exception to this rule; it turns out that in measure theory (but not in other parts of mathematics!) it is convenient to define $0 \cdot \infty = \infty \cdot 0 = 0$.

Since algebraic expressions with extended real numbers are not always defined, we need to be careful and always check that our expressions make sense.

We are now ready to define measurable functions:

Definition 5.4.1 *Let $(X, \mathcal{A}, \mu)$ be a measure space. A function $f : X \to \overline{\mathbb{R}}$ is* measurable *if*

$$f^{-1}([-\infty, r)) \in \mathcal{A}$$

for all $r \in \mathbb{R}$. In other words, the set

$$\{x \in \mathbb{R}^d \,:\, f(x) < r\}$$

must be measurable for all $r \in \mathbb{R}$.

The half-open intervals in the definition are just a convenient starting point for showing that the inverse images of more complicated sets are measurable:

Proposition 5.4.2 *If $f : \mathbb{R}^d \to \overline{\mathbb{R}}$ is measurable, then $f^{-1}(I) \in \mathcal{M}$ for all intervals $I = (s, r), I = (s, r], I = [s, r), I = [s, r]$ where $s, r \in \overline{\mathbb{R}}$. Indeed, $f^{-1}(A) \in \mathcal{M}$ for all open and closed sets A.*

Proof: We use that inverse images commute with intersections, unions and complements. First observe that for any $r \in \mathbb{R}$

$$f^{-1}\big([-\infty, r]\big) = f^{-1}\big(\bigcap_{n \in \mathbb{N}} [-\infty, r + \frac{1}{n})\big) = \bigcap_{n \in \mathbb{N}} f^{-1}\big([-\infty, r + \frac{1}{n})\big) \in \mathcal{M}$$

which shows that the closed intervals $[-\infty, r]$ are measurable. Taking complements, we see that the intervals $[s, \infty]$ and $(s, \infty]$ are measurable:

$$f^{-1}([s, \infty]) = f^{-1}([-\infty, s)^c) = \big(f^{-1}([-\infty, s))\big)^c \in \mathcal{M}$$

and

$$f^{-1}((s,\infty]) = f^{-1}([-\infty,s]^c) = \big(f^{-1}([-\infty,s])\big)^c \in \mathcal{M}$$

To show that finite intervals are measurable, we just take intersections, e.g.,

$$f^{-1}((s,r)) = f^{-1}([-\infty,r)\cap(s,\infty]) = f^{-1}([-\infty,r))\cap f^{-1}((s,\infty]) \in \mathcal{M}$$

If A is open, we know from Theorem 5.2.9 that it is a countable union $A=\bigcup_{n\in\mathbb{N}} I_n$ of open intervals. Hence

$$f^{-1}(A) = f^{-1}\big(\bigcup_{n\in\mathbb{N}} I_n\big) = \bigcup_{n\in\mathbb{N}} f^{-1}(I_n) \in \mathcal{M}$$

Finally, if A is closed, we use that its complement is open to get

$$f^{-1}(A) = \big(f^{-1}(A^c)\big)^c \in \mathcal{M}$$

□

It is sometimes convenient to use other kinds of intervals than those in the definition to check that a function is measurable:

Proposition 5.4.3 *Consider a function $f:\mathbb{R}^d\to\overline{\mathbb{R}}$. If either*

(i) $f^{-1}([-\infty,r])\in\mathcal{M}$ *for all* $r\in\mathbb{R}$, *or*

(ii) $f^{-1}([r,\infty])\in\mathcal{M}$ *for all* $r\in\mathbb{R}$, *or*

(iii) $f^{-1}((r,\infty])\in\mathcal{M}$ *for all* $r\in\mathbb{R}$,

then f is measurable.

Proof: In either case we just have to check that $f^{-1}([-\infty,r))\in\mathcal{M}$ for all $r\in\mathbb{R}$. This can be done by the techniques in the previous proof. The details are left to the reader. □

The next result tells us that there are many measurable functions:

Proposition 5.4.4 *All continuous functions $f:\mathbb{R}^d\to\mathbb{R}$ are measurable.*

Proof: Since f is continuous and takes values in $\mathbb{R}$,

$$f^{-1}([-\infty,r)) = f^{-1}((-\infty,r))$$

is an open set by Proposition 2.3.9 and thus measurable by Theorem 5.2.9. □

We shall now prove a series of results showing how we can obtain new measurable functions from old ones. These results are not very exciting, but they are necessary for the rest of the theory. Note that the functions in the next two propositions take values in $\mathbb{R}$ and not $\overline{\mathbb{R}}$.

Proposition 5.4.5 *If $f : \mathbb{R}^d \to \mathbb{R}$ is measurable, then $\phi \circ f$ is measurable for all continuous functions $\phi : \mathbb{R} \to \mathbb{R}$. In particular, f^2 is measurable.*

Proof: We have to check that

$$(\phi \circ f)^{-1}((-\infty, r)) = f^{-1}(\phi^{-1}((-\infty, r)))$$

is measurable. Since ϕ is contiunuous, $\phi^{-1}((-\infty, r))$ is open, and consequently $f^{-1}(\phi^{-1}((-\infty, r)))$ is measurable by Proposition 5.4.2. To see that f^2 is measurable, apply the first part of the theorem to the function $\phi(x) = x^2$. □

Proposition 5.4.6 *If the functions $f, g :\to \mathbb{R}$ are measurable, so are $f + g$, $f - g$, and fg.*

Proof: To prove that $f + g$ is measurable, observe first that $f + g < r$ means that $f < r - g$. Since the rational numbers are dense, it follows that there is a rational number q such that $f < q < r - g$. Hence

$$(f + g)^{-1}([-\infty, r)) = \{x \in \mathbb{R}^d \mid f + g < r\}$$

$$= \bigcup_{q \in \mathbb{Q}} \Big(\{x \in \mathbb{R}^d \mid f(x) < q\} \cap \{x \in \mathbb{R}^d \mid g < r - q\} \Big)$$

which is measurable since $\mathbb{Q}$ is countable and a countabe union of measurable sets is measurable. A similar argument proves that $f - g$ is measurable.

To prove that fg is measurable, note that by Proposition 5.4.5 and what we have already proved, f^2, g^2, and $(f + g)^2$ are measurable, and hence

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

is measurable (check the details). □

We would often like to apply the result above to functions that take values in the extended real numbers, but the problem is that the expressions need not make sense. As we shall mainly be interested in functions that are finite except on a set of measure zero, there is a way out of the problem. Let us start with the terminology.

Definition 5.4.7 *We say that a measurable function $f : \mathbb{R}^d \to \overline{\mathbb{R}}$ is* finite almost everywhere *if the set $\{x \in \mathbb{R}^d : f(x) = \pm\infty\}$ has measure zero. We say that two measurable functions $f, g : \mathbb{R}^d \to \overline{\mathbb{R}}$ are* equal almost everywhere *if the set $\{x \in \mathbb{R}^d : f(x) \neq g(x)\}$ has measure zero. We usually abbreviate "almost everywhere" by "a.e.".*

If the measurable functions f and g are finite a.e., we can modify them to get measurable functions f' and g' which take values in $\mathbb{R}$ and are equal a.e. to f and g, respectively (see exercise 11). By the proposition above, $f' + g'$, $f' - g'$ and $f'g'$ are measurable, and for many purposes they are good representatives for $f + g$, $f - g$ and fg.

Let us finally see what happens to limits of sequences.

Proposition 5.4.8 *If $\{f_n\}$ is a sequence of measurable functions, then $\sup_{n\in\mathbb{N}} f_n(x)$, $\inf_{n\in\mathbb{N}} f_n(x)$, $\limsup_{n\to\infty} f_n(x)$ and $\liminf_{n\to\infty} f_n(x)$ are measurable. If the sequence converges pointwise, then $\lim_{n\to\infty} f_n(x)$ is a measurable function.*

Proof: To see that $f(x) = \sup_{n\in\mathbb{N}} f_n(x)$ is measurable, we use Proposition 5.4.3(iii). For any $r \in \mathbb{R}$

$$f^{-1}((r,\infty)) = \{x \in \mathbb{R}^d \, : \sup_{n\in\mathbb{N}} f_n(x) > r\} =$$

$$= \bigcup_{n\in\mathbb{N}} \{\mathbf{x} \in \mathbb{R}^d \; : \; f_n(x) > r\} = \bigcup_{n\in\mathbb{N}} f_n^{-1}((r,\infty]) \in \mathcal{M}$$

and hence f is measurable by Propostion 5.4.3(iii). The argument for $\inf_{n\in\mathbb{N}} f_n(x)$ is similar.

To show that $\limsup_{n\to\infty} f_n(x)$ is measurable, first observe that the functions

$$g_k(x) = \sup_{n\geq k} f_n(x)$$

are measurable by what we have already shown. Since

$$\limsup_{n\to\infty} f_n(x) = \inf_{k\in\mathbb{N}} g_k(x)\big),$$

the measurability of $\limsup_{n\to\infty} f_n(x)$ follows. A similar argument holds for $\liminf_{n\to\infty} f_n(x)$. If the sequence converges pointwise, then $\lim_{n\to\infty} f_n(x) = \limsup_{n\to\infty} f_n(x)$ and is hence measurable. □

Let us sum up what we have done so far in this chapter. We have constructed the Lebesgue measure μ which assigns a d-dimensional volume to a large class of subset of $\mathbb{R}^d$, and we have explored the basic properties of a class of measurable functions which are closely connected to the Lebesgue measure. In the following sections we shall combine the two to create a theory of integration which is stronger and more flexible than the one you are used to.

Exercises for Section 5.4

1. Show that if $f : \mathbb{R}^d \to \overline{\mathbb{R}}$ is measurable, the sets $f^{-1}(\{\infty\})$ and $f^{-1}(\{-\infty\})$ are measurable.
2. Complete the proof of Proposition 5.4.2 by showing that f^{-1} of the intervals $(-\infty, r)$, $(-\infty, r]$, $[r, \infty)$, (r, ∞), $(-\infty, \infty)$, where $r \in \mathbb{R}$, are measurable.
3. Prove Proposition 5.4.3.
5. Show that if $f_1, f_2, \ldots, f_n$ are measurable functions with values in $\mathbb{R}$, then $f_1 + f_2 + \cdots + f_n$ and $f_1 f_2 \cdot \ldots \cdot f_n$ are measurable.
5. The *indicator function* of a set $A \subseteq \mathbb{R}$ is defined by

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

   a) Show that $\mathbf{1}_A$ is a measurable function if and only if $A \in \mathcal{M}$.
   b) A *simple* function is a function $f : \mathbb{R}^d \to \mathbb{R}$ of the form

$$f(x) = \sum_{i=1}^{n} a_i \mathbf{1}_{A_i}(x)$$

   where $a_1, a_2, \ldots, a_n \in \mathbb{R}$ and $A_1, A_2, \ldots, A_n \in \mathcal{M}$. Show that all simple functions are measurable.

6. Let $\{E_n\}$ be a disjoint sequence of measurable sets such that $\bigcup_{n=1}^{\infty} E_n = \mathbb{R}^d$, and let $\{f_n\}$ be a sequence of measurable functions. Show that the function defined by

$$f(x) = f_n(x) \text{ when } x \in E_n$$

   is measurable.
7. Fill in the details of the proof of the fg part of Proposition 5.4.6. You may want to prove first that if $h : \mathbb{R}^d \to \mathbb{R}$ is measurable, then so is $\frac{h}{2}$.
8. Prove the inf- and the lim inf-part of Proposition 5.4.8.
9. Let us write $f \sim g$ to denote that f and g are two measurable functions which are equal a.e.. Show that $\sim$ is an equivalence relation, i.e.:
   (i) $f \sim f$
   (ii) If $f \sim g$, then $g \sim f$.
   (iii) If $f \sim g$ and $g \sim h$, then $f \sim h$.
10. Show that if $f : \mathbb{R}^d \to \overline{\mathbb{R}}$ is measurable and $g : \mathbb{R}^d \to \overline{\mathbb{R}}$ equals f almost everywhere, then g is measurable.
11. Assume that $f : \mathbb{R}^d \to \mathbb{R}$ is finite a.e. Define a new function $f' : \mathbb{R}^d \to \mathbb{R}$ by

$$f'(x) = \begin{cases} f(x) & \text{if } f(x) \text{ is finite} \\ \\ 0 & \text{otherwise} \end{cases}$$

   Show that f' is measurable and equal to f a.e.
12. A sequence $\{f_n\}$ of measurable functions is said to *converge almost everywhere* to f if there is a set A of measure 0 such that $f_n(x) \to f(x)$ for all $x \notin A$. Show that f is measurable.

5.5 Integration of simple functions

If A is a subset of $\mathbb{R}^d$, we define its *indicator function* by

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

The indicator function is measuable if and only if A is measurable.

A measurable function $f : \mathbb{R}^d \to \mathbb{R}$ is called a *simple* function if it takes only finitely many different values $a_1, a_2, \ldots, a_n$. We may then write

$$f(x) = \sum_{i=1}^{n} a_1 \mathbf{1}_{A_i}(x)$$

where the sets $A_i = \{x \in \mathbb{R}^d \mid f(x) = a_i\}$ are disjoint and measurable. Note that if one of the a_i's is zero, the term does not contribute to the sum, and it is occasionally convenient to drop it.

If we instead start with measurable sets $B_1, B_2, \ldots, B_m$ and real numbers $b_1, b_2, \ldots, b_m$, then

$$g(x) = \sum_{i=1}^{m} b_i \mathbf{1}_{B_i}(x)$$

is measurable and takes only finitely many values, and hence is a simple function. The difference between f and g is that the sets $A_1, A_2, \ldots, A_n$ in f are disjoint with union $\mathbb{R}^d$, and that the numbers $a_1, a_2, \ldots, a_n$ are distinct. The same need not be the case for g. We say that the simple function f is on *standard form*, while g is not.

You may think of a simple function as a generalized step function. The difference is that step functions are constant on intervals (in $\mathbb{R}$), rectangles (in $\mathbb{R}^2$), or boxes (in higher dimensions), while simple functions need only be constant on much more complicated (but still measurable) sets.

We can now define the integral of a nonnegative simple function.

Definition 5.5.1 *Assume that*

$$f(x) = \sum_{i=1}^{n} a_i \mathbf{1}_{A_i}(x)$$

is a nonnegative simple function on standard form. Then the (Lebesgue) integral *of f is defined by*

$$\int f \, d\mu = \sum_{i=1}^{n} a_i \mu(A_i)$$

Recall that we are using the convention that $0 \cdot \infty = 0$, and hence $a_i \mu(A_i) = 0$ if $a_i = 0$ and $\mu(A_i) = \infty$.

Note that the integral of a simple function is

$$\int \mathbf{1}_A \, d\mu = \mu(A)$$

To see that the definition is reasonable, assume that you are in $\mathbb{R}^2$. Since $\mu(A_i)$ measures the area of the set A_i, the product $a_i\mu(A_i)$ measures in an intuitive way the volume of the solid with base A_i and height a_i.

We need to know that the formula in the definition also holds when the simple function is not on standard form. The first step is the following, simple lemma

Lemma 5.5.2 *If*

$$g(x) = \sum_{j=1}^{m} b_j \mathbf{1}_{B_j}(x)$$

is a nonnegative simple function where the B_j's are disjoint and $\mathbb{R}^d = \bigcup_{j=1}^m B_j$, then

$$\int g \, d\mu = \sum_{j=1}^{n} b_j \mu(B_j)$$

Proof: The problem is that the values $b_1, b_2, \ldots, b_m$ need not be distinct, but this is easily fixed: If $c_1, c_2, \ldots, c_k$ are the distinct values taken by g, let $b_{i,1}$, $b_{i,2}$,. . . ,b_{i,n_i} be the b_j's that are equal to c_i, and let $C_i = B_{i,1} \cup B_{i,2} \cup \ldots \cup B_{i,n_i}$. Then $\mu(C_i) = \mu(B_{i,1}) + \mu(B_{i,2}) + \ldots + \mu(B_{i,n_i})$, and hence

$$\sum_{j=1}^{n} b_j \mu(B_j) = \sum_{i=1}^{k} c_i \mu(C_i)$$

Since $g(x) = \sum_{i=1}^k c_i \mathbf{1}_{C_i}(x)$ is the standard form representation of g, we have

$$\int g \, d\mu = \sum_{j=1}^{n} c_i \mu(C_i)$$

and the lemma is proved □

The next step is also easy:

Proposition 5.5.3 *Assume that f and g are two nonnegative simple functions, and let c be a nonnnegative, real number. Then*

(i) $\int cf \, d\mu = c \int f \, d\mu$

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

Proof: (i) is left to the reader. To prove (ii), let

$$f(x) = \sum_{i=1}^{n} a_i \mathbf{1}_{A_i}(x)$$

$$g(x) = \sum_{j=1}^{n} b_j \mathbf{1}_{B_j}(x)$$

be standard form representations of f and g, and define $C_{i,j} = A_i \cap B_j$. By the lemma above

$$\int f \, d\mu = \sum_{i,j} a_i \mu(C_{i,j})$$

and

$$\int g \, d\mu = \sum_{i,j} b_j \mu(C_{i,j})$$

and also

$$\int (f+g) \, d\mu = \sum_{i,j} (a_i + b_j) \mu(C_{i,j})$$

since the value of $f + g$ on $C_{i,j}$ is $a_i + b_j$ □

We can now easily prove that the formula in Definition 5.4.1 holds for all positive representations of step functions:

Corollary 5.5.4 *If* $f(x) = \sum_{n=1} a_i \mathbf{1}_{A_i}(x)$ *is a step function with* $a_i \geq 0$ *for all* i*, then*

$$\int f \, d\mu = \sum_{i=1}^{n} a_i \mu(A_i)$$

Proof: By the Proposition

$$\int f \, d\mu = \int \sum_{i=1}^{n} a_i \mathbf{1}_{A_i} \, d\mu = \sum_{i=1}^{n} \int a_i \mathbf{1}_{A_i} \, d\mu = \sum_{i=1}^{n} a_i \int \mathbf{1}_{A_i} \, d\mu = \sum_{i=1}^{n} a_i \mu(A_i)$$

□

We need to prove yet another almost obvious result. We write $g \leq f$ to say that $g(x) \leq f(x)$ for all x.

Proposition 5.5.5 *Assume that* f *and* g *are two nonnegative simple functions. If* $g \leq f$*, then*

$$\int g \, d\mu \leq \int f \, d\mu$$

Proof: Note that $f - g$ is also a nonnegative simple function, and since $f = (f - g) + g$, we get

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

by Proposition 6.5.3(ii). □

We shall end this section with a key result on limits of integrals, but first we need some notation. Observe that if $f = \sum_{i=1}^{n} a_n \mathbf{1}_{A_n}$ is a simple function and B is a measurable set, then $\mathbf{1}_B f = \sum_{i=1}^{n} a_n \mathbf{1}_{A_n \cap B}$ is also a simple function. We shall write

$$\int_B f\,d\mu = \int \mathbf{1}_B f\,d\mu$$

and call this the *integral of f over B*. The lemma below may seem obvious, but it is the key to many later results.

Lemma 5.5.6 *Assume that B is a measurable set, b a positive real number, and $\{f_n\}$ an increasing sequence of nonnegative simple functions such that $\lim_{n\to\infty} f_n(x) \geq b$ for all $x \in B$. Then $\lim_{n\to\infty} \int_B f_n\,d\mu \geq b\mu(B)$.*

Proof: Let a be any positive number less than b, and define

$$A_n = \{x \in B \mid f_n(x) \geq a\}$$

Since $f_n(x)$ increases beyond a for all $x \in B$, we see that the sequence $\{A_n\}$ is increasing and that

$$B = \bigcup_{n=1}^{\infty} A_n$$

By continuity of measure (Proposition 5.3.4(i)), $\mu(B) = \lim_{n\to\infty} \mu(A_n)$, and hence for any positive number m less that $\mu(B)$, we can find an $N \in \mathbb{N}$ such that $\mu(A_n) > m$ when $n \geq N$. Since $f_n \geq a$ on A_n, we thus have

$$\int_B f_n\,d\mu \geq \int_{A_n} a\,d\mu = am$$

whenever $n \geq N$. Since this holds for any number a less than b and any number m less than $\mu(B)$, we must have $\lim_{n\to\infty} \int_B f_n\,d\mu \geq b\mu(B)$ □

To get the result we need, we extend the lemma to simple functions:

Proposition 5.5.7 *Let g be a nonnegative simple function and assume that $\{f_n\}$ is an increasing sequence of nonnegative simple functions such that $\lim_{n\to\infty} f_n(x) \geq g(x)$ for all x. Then*

$$\lim_{n\to\infty} \int f_n\,d\mu \geq \int g\,d\mu$$

Proof: Let $g(x) = \sum_{i=1}^{m} b_i \mathbf{1}_{B_1}(x)$ be the standard form of g. If any of the b_i's is zero, we may just drop that term in the sum, so that we from now on assume that all the b_i's are nonzero. By Corollary 5.4.3(ii), we have

$$\int_{B_1 \cup B_2 \cup \ldots \cup B_m} f_n \, d\mu = \int_{B_1} f_n \, d\mu + \int_{B_2} f_n \, d\mu + \ldots + \int_{B_m} f_n \, d\mu$$

By the lemma, $\lim_{n\to\infty} \int_{B_i} f_n \, d\mu \geq b_i \mu(B_i)$, and hence

$$\lim_{n\to\infty} \int f_n \, d\mu \geq \lim_{n\to\infty} \int_{B_1 \cup B_2 \cup \ldots \cup B_m} f_n \, d\mu \geq \sum_{i=1}^{m} b_i \mu(B_i) = \int g \, d\mu$$

□

We are now ready to extend the Lebesgue integral to all positive, measurable functions. This will be the topic of the next section.

Exercises for Section 5.5

1. Show that if f is a measurable function, then the *level set*

$$A_a = \{x \in \mathbb{R}^d \mid f(x) = a\}$$

is measurable for all $a \in \overline{\mathbb{R}}$.

2. Check that according to Definition 5.5.1, $\int \mathbf{1}_A \, \mathbf{d}\mu = \mu(A)$ for all $A \in \mathcal{M}$.
3. Prove part (i) of Proposition 5.5.3.
4. Show that if $f_1, f_2, \ldots, f_n$ are simple functions, then so are

$$h(x) = \max\{f_1(x), f_2(x), \ldots, f_n(x)\}$$

and

$$h(x) = \min\{f_1(x), f_2(x), \ldots, f_n(x)\}$$

5. Let $A = \mathbb{Q} \cap [0,1]$. This function is not integrable in the Riemann sense. What is $\int \mathbf{1}_A \, d\mu$?

5.6 Integrals of nonnegative functions

We are now ready to define the integral of a general, nonnegative, measurable function.

Definition 5.6.1 *If $f : \mathbb{R}^d \to [0,\infty]$ is measurable, we define*

$$\int f \, d\mu = \sup\{\int g \, d\mu \mid g \textit{ is a nonnegative simple function, } g \leq f\}$$

Remark: Note that if f is a simple function, we now have two definitions of $\int f\,d\mu$; the original one in Definition 5.5.1 and a new one in the definition above. It follows from Proposition 5.5.5 that the two definitions agree.

The definition above is natural, but also quite abstract, and we shall work toward a reformulation that is often easier to handle.

Proposition 5.6.2 *Let $f : \mathbb{R}^d \to [0,\infty]$ be a measurable function, and assume that $\{h_n\}$ is an increasing sequence of simple functions converging pointwise to f. Then*

$$\lim_{n\to\infty} \int h_n\,d\mu = \int f\,d\mu$$

Proof: Since the sequence $\{\int h_n\,d\mu\}$ is increasing by Proposition 5.5.5, the limit clearly exists (it may be ∞), and since $\int h_n\,d\mu \leq \int f\,d\mu$ for all n, we must have

$$\lim_{n\to\infty} \int h_n\,d\mu \leq \int f\,d\mu$$

To get the opposite inequality, it suffices to show that

$$\lim_{n\to\infty} \int h_n\,d\mu \geq \int g\,d\mu$$

for every simple function $g \leq f$, but this follows from Proposition 5.5.7. □

The proposition above would lose much of its power if there weren't any increasing sequences of simple functions converging to f. The next result tells us that there always are. Pay attention to the argument, it is a key to why the theory works.

Proposition 5.6.3 *If $f : \mathbb{R}^d \to [0,\infty)$ is measurable, there is an increasing sequence $\{h_n\}$ of simple functions converging pointwise to f. Moreover, for each n either $f_n(x) - \frac{1}{2^n} < h_n(x) \leq f_n(x)$ or $h_n(x) = 2^n$*

Proof: To construct the simple function h_n, we cut the interval $[0, 2^n)$ into half-open subintervals of length $\frac{1}{2^n}$, i.e. intervals

$$I_k = \left[\frac{k}{2^n}, \frac{k+1}{2^n}\right)$$

where $0 \leq k < 2^{2n}$, and then let

$$A_k = f^{-1}(I_k)$$

We now define

$$h_n(x) = \sum_{k=0}^{2^{2n}-1} \frac{k}{2^n}\mathbf{1}_{A_k}(x) + 2^n\mathbf{1}_{\{x \,|\, f(x)\geq 2^n\}}$$

By definition, h_n is a simple function no greater than f. Since the intervals get narrower and narrower and cover more and more of $[0, \infty)$, it is easy to see that h_n converges pointwise to f. To see why the sequence increases, note that each time we increase n by one, we split each of the former intervals I_k in two, and this will cause the new step function to equal the old one for some x's and jump one step upwards for others (make a drawing).

The last statement follows directly from the construction. □

Remark: You should compare the partitions in the proof above to the partitions you have seen in earlier treatments of integration. When we integrate a function of one variable in calculus, we partition an interval $[a, b]$ on the x-axis and use this partition to approximate the original function by a step function. In the proof above, we instead partitioned the y-axis into intervals and used this partition to approximate the original function by a simple function. The difference is that the latter approach gives us much better control over what is going one; the partition controls the oscillations of the function. The price we have to pay, it that we get simple functions instead of step functions, and to use simple functions for integration, we need measure theory.

Let us combine the last two results in a handy corollary:

Corollary 5.6.4 *If $f : \mathbb{R}^d \to [0, \infty)$ is measurable, there is an increasing sequence $\{h_n\}$ of simple functions converging pointwise to f, and for all such sequences*

$$\int f \, d\mu = \lim_{n \to \infty} \int h_n \, d\mu$$

Let us take a look at some properties of the integral.

Proposition 5.6.5 *Assume that $f, g : \mathbb{R}^d \to [0, \infty]$ are measurable functions and that c is a nonnegative, real number. Then:*

(i) $\int cf \, d\mu = c \int f \, d\mu$.

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

(iii) *If $g \le f$, then* $\int g \, d\mu \le \int f \, d\mu$.

Proof: (iii) is immediate from the definition, and (i) is left to the reader. To prove (ii), let $\{f_n\}$ and $\{g_n\}$ be to increasing sequence of simple functions converging to f and g, respectively. Then $\{f_n + g_n\}$ is an increasing sequence of simple functions converging to $f + g$, and

$$\int (f + g) \, d\mu = \lim_{n \to \infty} \int (f_n + g_n) \, d\mu = \lim_{n \to \infty} \left(\int f_n \, d\mu + \int g_n \, d\mu \right) =$$

$$= \lim_{n\to\infty} \int f_n \, d\mu + \lim_{n\to\infty} \int g_n \, d\mu = \int f \, d\mu + \int g \, d\mu$$

where we have used Proposition 6.5.3(ii) to see that $\int (f_n + g_n) \, d\mu = \int f_n \, d\mu + \int g_n \, d\mu$. □

One of the big advantages of Lebesgue integration over traditional Riemann integration, is that the Lebesgue integral is much better behaved with respect to limits. The next result is our first example:

Theorem 5.6.6 (Monotone Convergence Theorem) *If $\{f_n\}$ is an increasing sequence of nonnegative, measurable functions such that $f(x) = \lim_{n\to\infty} f_n(x)$ for all x, then*

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

In other words,

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

Proof: We know from Proposition 5.4.8 that f is measurable, and hence the integral $\int f \, d\mu$ is defined. Since $f_n \leq f$, we have $\int f_n \, d\mu \leq \int f \, d\mu$ for all n, and hence

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

To prove the opposite inequality, we approximate each f_n by simple functions as in the proof of Proposition 5.6.3; in fact, let h_n be the n-th approximation to f_n. Assume that we can prove that the sequence $\{h_n\}$ converges to f; then

$$\lim_{n\to\infty} \int h_n \, d\mu = \int f \, d\mu$$

by Proposition 5.5.2. Since $f_n \geq h_n$, this would give us the desired inequality

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

It remains to show that $h_n(x) \to f(x)$ for all x. From Proposition 5.6.3 we know that for all n, either $f_n(x) - \frac{1}{2^n} < h_n(x) \leq f_n(x)$ or $h_n(x) = 2^n$. If $h_n(x) = 2^n$ for infinitely many n, then $h_n(x)$ goes to ∞, and hence to $f(x)$. If $h_n(x)$ is *not* equal to 2^n for infinitely many n, then we eventually have $f_n(x) - \frac{1}{2^n} < h_n(x) \leq f_n(x)$, and hence $h_n(x)$ converges to $f(x)$ since $f_n(x)$ does. □

We would really have liked the formula

$$\lim_{n\to\infty} \int f_n \, d\mu = \int \lim_{n\to\infty} f_n \, d\mu \tag{5.6.1}$$

above to hold in general, but as the following example shows, this is not the case.

Example 1: Let $f_n = \mathbf{1}_{[n,n+1]}$. Then $\lim_{n\to\infty} f_n(x) = 0$ for all x, but $\int f_n \, d\mu = 1$. Hence

$$\lim_{n\to\infty} \int f_n \, d\mu = 1$$

but

$$\int \lim_{n\to\infty} f_n \, d\mu = 0$$

♣

There are many results in measure theory giving conditions for (5.6.1) to hold, but there is no ultimate theorem covering all others. There is, however, a simple inequality that always holds.

Theorem 5.6.7 (Fatou's Lemma) *Assume that* $\{f_n\}$ *is a sequence of non-negative, measurable functions. Then*

$$\liminf_{n\to\infty} \int f_n \, d\mu \geq \int \liminf_{n\to\infty} f_n \, d\mu$$

Proof: Let $g_k(x) = \inf_{k\geq n} f_n(x)$. Then $\{g_k\}$ is an increasing sequence of measurable functions, and by the Monotone Convergence Theorem

$$\lim_{k\to\infty} \int g_k \, d\mu = \int \lim_{k\to\infty} g_k \, d\mu = \int \liminf_{n\to\infty} f_n \, d\mu$$

where we have used the definition of lim inf in the last step. Since $f_k \geq g_k$, we have $\int f_k \, d\mu \geq \int g_k \, d\mu$, and hence

$$\liminf_{k\to\infty} \int f_k \, d\mu \geq \lim_{k\to\infty} \int g_k \, d\mu = \int \liminf_{n\to\infty} f_n \, d\mu$$

and the result is proved. □

Fatou's Lemma is often a useful tool in establishing more sophisticated results, see Exercise 14 for a typical example.

Just as for simple functions, we define integrals over measurable subsets A of $\mathbb{R}^d$ by the formula

$$\int_A f \, d\mu = \int \mathbf{1}_A f \, d\mu$$

So far we have allowed our integrals to be infinite, but we are mainly interested in situations where $\int f \, d\mu$ is finite:

Definition 5.6.8 *A function* $f : \mathbb{R}^d \to [0,\infty]$ *is said to be* integrable *if it is measurable and* $\int f \, d\mu < \infty$.

Exercises for Section 5.6

1. Assume $f : \mathbb{R}^d \to [0, \infty]$ is a nonnegative simple function. Show that the two definitions of $\int f \, d\mu$ given in Definitions 5.5.1 and 5.6.1 coincide.
2. Prove Proposition 5.6.5(i).
3. Show that if $f : \mathbb{R}^d \to [0, \infty]$ is measurable, then
$$\mu(\{x \in \mathbb{R}^d \mid f(x) \geq a\}) \leq \frac{1}{a} \int f \, d\mu$$
for all positive, real numbers a.
4. In this problem, $f, g : \mathbb{R}^d \to [0, \infty]$ are measurable functions.
    a) Show that $\int f \, d\mu = 0$ if and only if $f = 0$ a.e.
    b) Show that if $f = g$ a.e., then $\int f \, d\mu = \int g \, d\mu$.
    c) Show that if $\int_E f \, d\mu = \int_E g \, d\mu$ for all measurable sets E, then $f = g$ a.e.
5. In this problem, $f : \mathbb{R}^d \to [0, \infty]$ is a measurable function and A, B are measurable sets.
    a) Show that $\int_A f \, d\mu \leq \int f \, d\mu$
    b) Show that if A, B are disjoint, then $\int_{A \cup B} f \, d\mu = \int_A f \, d\mu + \int_B f \, d\mu$.
    c) Show that in general $\int_{A \cup B} f \, d\mu + \int_{A \cap B} f \, d\mu = \int_A f \, d\mu + \int_B f \, d\mu$.
6. Show that if $f : \mathbb{R}^d \to [0, \infty]$ is integrable, then f is finite a.e.
7. Let $f : \mathbb{R} \to \mathbb{R}$ be the function
$$f(x) = \begin{cases} 1 & \text{if } x \text{ is rational} \\ \\ 0 & \text{otherwise} \end{cases}$$
and for each $n \in \mathbb{N}$, let $f_n : \mathbb{R} \to \mathbb{R}$ be the function
$$f_n(x) = \begin{cases} 1 & \text{if } x = \frac{p}{q} \text{ where } p \in \mathbb{Z},\ q \in \mathbb{N}, q \leq n \\ \\ 0 & \text{otherwise} \end{cases}$$
    a) Show that $\{f_n(x)\}$ is an increasing sequence converging to $f(x)$ for all $x \in \mathbb{R}$.
    b) Show that each f_n is Riemann integrable over $[0, 1]$ with $\int_0^1 f_n(x) \, dx = 0$ (this is integration as taught in calculus courses).
    c) Show that f is *not* Riemann integrable over $[0, 1]$.
    d) Show that the one-dimensional Lebesgue integral $\int_{[0,1]} f \, d\mu$ exists and find its value.
8. a) Let $\{u_n\}$ be a sequence of positive, measurable functions. Show that
$$\int \sum_{n=1}^{\infty} u_n \, d\mu = \sum_{n=1}^{\infty} \int u_n \, d\mu$$

b) Assume that f is a nonnnegative, measurable function and that $\{B_n\}$ is a disjoint sequence of measurable sets with union B. Show that

$$\int_B f\,d\mu = \sum_{n=1}^{\infty} \int_{B_n} f\,d\mu$$

9. Assume that f is a nonnegative, measurable function and that $\{A_n\}$ is an increasing sequence of measurable sets with union A. Show that

$$\int_A f\,d\mu = \lim_{n\to\infty} \int_{A_n} f\,d\mu$$

10. Show the following generalization of the Monotone Convergence Theorem: If $\{f_n\}$ is an increasing sequence of nonnegative, measurable functions such that $f(x) = \lim_{n\to\infty} f_n(x)$ *almost everywhere.* (i.e. for all x outside a set N of measure zero), then

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

11. Find a decreasing sequence $\{f_n\}$ of measurable functions $f_n : \mathbb{R} \to [0,\infty)$ converging pointwise to zero such that $\lim_{n\to\infty} \int f_n\,d\mu \neq 0$

12. Assume that $f : \mathbb{R}^d \to [0,\infty]$ is a measurable function, and that $\{f_n\}$ is a sequence of measurable functions converging pointwise to f. Show that if $f_n \leq f$ for all n,

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

13. Assume that $\{f_n\}$ is a sequence of nonnegative functions converging pointwise to f. Show that if

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

then

$$\lim_{n\to\infty} \int_E f_n\,d\mu = \int_E f\,d\mu$$

for all measurable $E \subseteq \mathbb{R}^d$.

14. Assume that $g : \mathbb{R}^d \to [0,\infty]$ is an *integrable* function, and that $\{f_n\}$ is a sequence of nonnegative, measurable functions converging pointwise to a function f. Show that if $f_n \leq g$ for all n, then

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

Hint: Apply Fatou's Lemma to both sequences $\{f_n\}$ and $\{g - f_n\}$.

5.7 Integrable functions

So far we only know how to integrate nonnegative functions, but it is not difficult to extend the theory to general functions. Given a function $f : \mathbb{R}^d \to \overline{\mathbb{R}}$, we first observe that $f = f_+ - f_-$, where f_+ and f_- are the nonnegative functions

$$f_+(x) = \begin{cases} f(x) & \text{if } f(x) > 0 \\ \\ 0 & \text{otherwise} \end{cases}$$

and

$$f_-(x) = \begin{cases} -f(x) & \text{if } f(x) < 0 \\ \\ 0 & \text{otherwise} \end{cases}$$

Note that f_+ and f_- are measurable if f is. Recall that a nonnegative, measurable function f is integrable if $\int f\,d\mu < \infty$.

Definition 5.7.1 *A function $f : \mathbb{R}^d \to \overline{\mathbb{R}}$ is called* integrable *if it is measurable, and $f+$ and f_- are integrable. We define the* (Lebesgue) integral *of f by*

$$\int f\,d\mu = \int f_+\,d\mu - \int f_-\,d\mu$$

The next lemma gives a useful characterization of integrable functions. The proof is left to the reader (see Exercise 3 for a hint).

Lemma 5.7.2 *A measurable function f is integrable if and only if its absolute value $|f|$ is integrable, i.e. if and only if $\int |f|\,d\mu < \infty$.*

The next lemma is a useful technical tool:

Lemma 5.7.3 *Assume that $g : \mathbb{R}^d \to [0, \infty]$ and $h : \mathbb{R}^d \to [0, \infty]$ are two integrable, nonnegative functions, and that $f(x) = g(x) - h(x)$ at all points where the difference is defined. Then f is integrable and*

$$\int f\,d\mu = \int g\,d\mu - \int h\,d\mu$$

Proof: Note that since g and h are integrable, they are finite a.e., and hence $f = g - h$ a.e. Modifying g and h on a set of measure zero (this will not change their integrals), we may assume that $f(x) = g(x) - h(x)$ for all x. Since $|f(x)| = |g(x) - h(x)| \leq |g(x)| + |h(x)|$, it follows from the lemma above that f is integrable.

As

$$f(x) = f_+(x) - f_-(x) = g(x) - h(x)$$

we have

$$f_+(x) + h(x) = g(x) + f_-(x)$$

where we on both sides have sums of nonnegative functions. By Proposition 5.5.5(ii), we get

$$\int f_+ \, d\mu + \int h \, d\mu = \int g \, d\mu + \int f_- \, d\mu$$

Rearranging the integrals (they are all finite), we get

$$\int f \, d\mu = \int f_+ \, d\mu - \int f_- \, d\mu = \int g \, d\mu - \int h \, d\mu$$

and the lemma is proved. □

We are now ready to prove that the integral behaves the way we expect:

Proposition 5.7.4 *Assume that $f, g : \mathbb{R}^d \to \overline{\mathbb{R}}$ are integrable functions, and that c is a constant. Then $f + g$ and cf are integrable, and*

(i) $\int cf \, d\mu = c \int f \, d\mu$.

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

(iii) If $g \leq f$, *then* $\int g \, d\mu \leq \int f \, d\mu$.

Proof: (i) is left to the reader (treat positive and negative c's separately). To prove (ii), first note that since f and g are integrable, the sum $f(x) + g(x)$ is defined a.e., and by changing f and g on a set of measure zero (this doesn't change their integrals), we may assume that $f(x) + g(x)$ i defined everywhere. Since

$$|f(x) + g(x)| \leq |f(x)| + |g(x)|,$$

$f + g$ is integrable. Obviously,

$$f + g = (f_+ - f_-) + (g_+ - g_-) = (f_+ + g_+) - (f_- + g_-)$$

and hence by the lemma above and Proposition 5.5.5(ii)

$$\int (f + g) \, d\mu = \int (f_+ + g_+) \, d\mu - \int (f_- + g_-) \, d\mu =$$

$$= \int f_+ \, d\mu + \int g_+ \, d\mu - \int f_- \, d\mu - \int g_- \, d\mu =$$

$$= \int f_+ \, d\mu - \int f_- \, d\mu + \int g_+ \, d\mu - \int g_- \, d\mu =$$

$$= \int f\,d\mu + \int g\,d\mu$$

To prove (iii), note that $f - g$ is a nonnegative function and hence by (i) and (ii):

$$\int f\,d\mu - \int g\,d\mu = \int f\,d\mu + \int (-1)g\,d\mu = \int (f-g)\,d\mu \geq 0$$

Consequently, $\int f\,d\mu \geq \int g\,d\mu$ and the proposition is proved. □

We can now extend our limit theorems to integrable functions taking both signs. The following result is probably the most useful of all limit theorems for integrals as it quite strong and at the same time easy to use. It tells us that if a convergent sequence of functions is dominated by an integrable function, then

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

Theorem 5.7.5 (Lebesgue's Dominated Convergence Theorem) *Assume that $g : \mathbb{R}^d \to \overline{\mathbb{R}}$ is a nonnegative, integrable function and that $\{f_n\}$ is a sequence of measurable functions converging pointwise to f. If $|f_n| \leq g$ for all n, then*

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

Proof: First observe that since $|f| \leq g$, f is integrable. Next note that since $\{g - f_n\}$ and $\{g + f_n\}$ are two sequences of nonnegative measurable functions, Fatou's Lemma gives:

$$\liminf_{n\to\infty} \int (g-f_n)\,d\mu \geq \int \liminf_{n\to\infty} (g-f_n)\,d\mu = \int (g-f)\,d\mu = \int g\,d\mu - \int f\,d\mu$$

and

$$\liminf_{n\to\infty} \int (g+f_n)\,d\mu \geq \int \liminf_{n\to\infty} (g+f_n)\,d\mu = \int (g+f)\,d\mu = \int g\,d\mu + \int f\,d\mu$$

On the other hand,

$$\liminf_{n\to\infty} \int (g-f_n)\,d\mu = \int g\,d\mu - \limsup_{n\to\infty} \int f_n\,d\mu$$

and

$$\liminf_{n\to\infty} \int (g+f_n)\,d\mu = \int g\,d\mu + \liminf_{n\to\infty} \int f_n\,d\mu$$

Combining the two expressions for $\liminf_{n\to\infty} \int (g-f_n)\,d\mu$, we see that

$$\int g\,d\mu - \limsup_{n\to\infty} \int f_n\,d\mu \geq \int g\,d\mu - \int f\,d\mu$$

and hence

$$\limsup_{n\to\infty} \int f_n \, d\mu \leq \int f \, d\mu$$

Combining the two expressions for $\liminf_{n\to\infty} \int (g + f_n) \, d\mu$, we similarly get

$$\liminf_{n\to\infty} \int f_n \, d\mu \geq \int f \, d\mu$$

Hence

$$\limsup_{n\to\infty} \int f_n \, d\mu \leq \int f \, d\mu \leq \liminf_{n\to\infty} f_n \, d\mu$$

which means that $\lim_{n\to\infty} \int f_n \, d\mu = \int f \, d\mu$. The theorem is proved. □

Remark: It is easy to check that we can relax the conditions above somewhat: If $f_n(x)$ converges to $f(x)$ a.e., and $|f_n(x)| \leq g(x)$ fails on a set of measure zero, the conclusion still holds (see Exercise 8 for the precise statement).

Let us take a look at a typical application of the theorem:

Proposition 5.7.6 *Assume that* $f : \mathbb{R}^2 \to \mathbb{R}$ *is a continuous function, and assume that there is an integrable function* $g : \mathbb{R} \to [0, \infty]$ *such that* $|f(x, y)| \leq g(y)$ *for all* $x, y \in \mathbb{R}$. *Then the function*

$$h(x) = \int f(x, y) \, d\mu(y)$$

is continuous (the expression $\int f(x, y) \, d\mu(y)$ *means that we for each fixed* x *integrate* $f(x, y)$ *as a function of* y*).*

Proof: According to Proposition 2.2.5 it suffices to prove that if $\{a_n\}$ is a sequence converging to a point a, then $h(a_n)$ converges to $h(a)$. Observe that

$$h(a_n) = \int f(a_n, y) \, d\mu(y)$$

and

$$h(a) = \int f(a, y) \, d\mu(y)$$

Observe also that since f is continuous, $f(a_n, y) \to f(a, y)$ for all y. Hence $\{f(a_n, y)\}$ is a sequence of functions which is dominated by the integrable function g and which converges pointwise to $f(a, y)$. By Lebesgue's Dominated Convergence Theorem,

$$\lim_{n\to\infty} h(a_n) = \lim_{n\to\infty} \int f(a_n, y) \, d\mu = \int f(a, y) \, d\mu = h(a)$$

and the proposition is proved. □

As before, we define $\int_A f \, d\mu = \int f \mathbf{1}_A \, d\mu$ for measurable sets A. We say that f *is integrable over* A if $f\mathbf{1}_A$ is integrable.

Exercises to Section 5.7

1. Show that if f is measurable, so are f_+ and f_-.
2. Show that if an integrable function f is zero a.e., then $\int f\,d\mu = 0$.
3. Prove Lemma 5.7.2. (*Hint:* Observe that $|f| = f_+ + f_-$.)
4. Prove Proposition 5.7.4(i). You may want to treat positive and negative c's separately.
5. Assume that $f : \mathbb{R}^d \to \overline{\mathbb{R}}$ is a measurable function.

   a) Show that if f is integrable over a measurable set A, and A_n is an increasing sequence of measurable sets with union A, then

   $$\lim_{n\to\infty} \int_{A_n} f\,d\mu = \int_A f\,d\mu$$

   b) Assume that $\{B_n\}$ is a decreasing sequence of measurable sets with intersection B. Show that if f is integrable over B_1, then

   $$\lim_{n\to\infty} \int_{B_n} f\,d\mu = \int_B f\,d\mu$$

6. Show that if $f : \mathbb{R}^d \to \mathbb{R}$ is integrable over a measurable set A, and A_n is a disjoint sequence of measurable sets with union A, then

   $$\int_A f\,d\mu = \sum_{n=1}^{\infty} \int_{A_n} f\,d\mu$$

7. Let $f : \mathbb{R} \to \overline{\mathbb{R}}$ be a measurable function, and define

   $$A_n = \{x \in \mathbb{R}^d \,|\, f(x) \geq n\}$$

   Show that

   $$\lim_{n\to\infty} \int_{A_n} f\,d\mu = 0$$

8. Prove the following slight extension of the Dominated Convergence Theorem:

   Theorem: Assume that $g : \mathbb{R}^d \to \overline{\mathbb{R}}$ is a nonnegative, integrable function and that $\{f_n\}$ is a sequence of measurable functions converging a.e. to f. If $|f_n(x)| \leq g(x)$ a.e. for each n, then

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

9. Assume that $g : \mathbb{R}^2 \to \mathbb{R}$ is continuous and that $y \to g(x, y)$ is integrable for each x. Assume also that the partial derivative $\frac{\partial g}{\partial x}(x, y)$ exists for all x and y, and that there is an integrable function $h : \mathbb{R} \to [0, \infty]$ such that

   $$\left|\frac{\partial g}{\partial x}(x, y)\right| \leq h(y)$$

for all x, y. Show that the function

$$f(x) = \int g(x, y)\, d\mu(y)$$

is differentiable at all points x and that

$$f'(x) = \int \frac{\partial g}{\partial x}(x, y)\, d\mu(y)$$

This is usually referred to as *differentiation under the integral sign.*

5.8 $L^1(\mathbb{R}^d)$ and $L^2(\mathbb{R}^d)$

In this final section we shall connect integration theory to the theory of normed spaces in Chapter 4. Recall from Definition 4.5.2 that a norm on a real vector space V is a function $\|\cdot\| : V \to [0, \infty)$ satisfying

(i) $\|\mathbf{u}\| \geq 0$ with equality if and only if $\mathbf{u} = 0$.

(ii) $\|\alpha \mathbf{u}\| = |\alpha| \|\mathbf{u}\|$ for all $\alpha \in \mathbb{R}$ and all $\mathbf{u} \in V$.

(iii) $\|\mathbf{u} + \mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$ for all $\mathbf{u}, \mathbf{v} \in V$.

Let us now put

$$\mathcal{L}^1(\mathbb{R}^d) = \{f : \mathbb{R}^d \to \overline{\mathbb{R}} \,:\, f \text{ is integrable}\}$$

and define $\|\cdot\|_1 : \mathcal{L}^1(\mathbb{R}^d) \to [0, \infty)$ by

$$\|f\|_1 = \int |f|\, d\mu$$

It is not hard to see that $\|\cdot\|_1$ satisfies the three axioms above with one exception; $\|f\|_1$ may be zero even when f is not zero — actually $\|f\|_1 = 0$ if and only if $f = 0$ a.e.

The usual way to fix this is to consider two functions f and g to be equal if they are equal almost everywhere. To be more precise, let us write $f \sim g$ if f and g are equal a.e., and define the *equivalence class* of f to be the set

$$[f] = \{g \in \mathcal{L}^1(\mathbb{R}^d) \,|\, g \sim f\}$$

Note that two such equivalence classes $[f]$ and $[g]$ are either equal (if f equals g a.e.) or disjoint (if f is not equal to g a.e.). If we let $L^1(\mathbb{R}^d)$ be the collection of all equivalence classes, we can organize $L^1(\mathbb{R}^d)$ as a normed vector space by defining

$$\alpha[f] = [\alpha f] \quad \text{and} \quad [f] + [g] = [f + g] \quad \text{and} \quad |[f]|_1 = \|f\|_1$$

The advantage of the space $(L^1(\mathbb{R}^d), |\cdot|_1)$ compared to $(\mathcal{L}^1(\mathbb{R}^d), \|\cdot\|_1)$ is that it is a normed space where all the theorems we have proved about such spaces apply — the disadvantage is that the elements are no longer functions, but equivalence classes of functions. In practice, there is very little difference between $(L^1(\mathbb{R}^d), |\cdot|_1)$ and $(\mathcal{L}^1(\mathbb{R}^d), \|\cdot\|_1)$, and mathematicians tend to blur the distinction between the two spaces: they pretend to work in $L^1(\mathbb{R}^d)$, but still consider the elements as functions. We shall follow this practice here; it is totally harmless as long as you remember that whenever we talk about an element of $L^1(\mathbb{R}^d)$ as a function, we are really choosing a representative from an equivalence class (Exercise 3 gives a more thorough and systematic treatment of $L^1(\mathbb{R}^d)$).

The most important fact about $(L^1(\mathbb{R}^d), |\cdot|_1)$ is that it is complete. In many ways, this is the most impressive success of the theory of Lebesgue integration: We have seen in previous chapters how important completeness is, and it is a great advantage to work with a theory of integration where the space of integrable functions is naturally complete. Before we turn to the proof, you may want to remind yourself of Proposition 4.5.5 which shall be our main tool.

Theorem 5.8.1 *$(L^1(\mathbb{R}^d), |\cdot|_1)$ is complete.*

Proof: Assume that $\{u_n\}$ is a sequence of functions in $L^1(\mathbb{R}^d)$ such that the series $\sum_{n=1}^{\infty} |u_n|_1$ converges. According to Proposition 4.5.5, it suffices to show that the series $\sum_{n=1}^{\infty} u_n(x)$ must converge in $L^1(\mathbb{R}^d)$. Observe that

$$\infty > \sum_{n=1}^{\infty} |u_n|_1 = \lim_{N\to\infty} \sum_{n=1}^{N} |u_n|_1 = \lim_{N\to\infty} \sum_{n=1}^{N} \int |u_n|\, d\mu =$$

$$= \lim_{N\to\infty} \int \sum_{n=1}^{N} |u_n|\, d\mu = \int \lim_{N\to\infty} \sum_{n=1}^{N} |u_n|\, d\mu = \int \sum_{n=1}^{\infty} |u_n|\, d\mu$$

where we have used the Monotone Convergence Theorem to move the limit inside the integral sign. This means that the function

$$g(x) = \sum_{n=1}^{\infty} |u_n(x)|$$

is integrable. We shall use g as the dominating function in the Dominated Convergence Theorem.

Let us first observe that since $g(x) = \sum_{n=1}^{\infty} |u_n(x)|$ is integrable, the series converges a.e. Hence the sequence $\sum_{n=1}^{\infty} u_n(x)$ (without the absolute values) converges absolutely a.e., and hence it converges a.e. in the ordinary sense. Let $f(x) = \sum_{n=1}^{\infty} u_n(x)$ (put $f(x) = 0$ on the null set where the series does not converge). It remains to prove that the series converges in L^1-sense, i.e. that $|f - \sum_{n=1}^N u_n|_1 \to 0$ as $N \to \infty$. By definition of f, we see that $\lim_{N\to\infty}\left(f(x) - \sum_{n=1}^N u_n(x)\right) = 0$ a.e. Since $|f(x) - \sum_{n=1}^N u_n(x)| = |\sum_{n=N+1}^\infty u_n(x)| \leq g(x)$ a.e., it follows from Dominated Convergence Theorem (actually from the slight extension in Exercise 5.7.8) that

$$|f - \sum_{n=1}^N u_n|_1 = \int |f - \sum_{n=1}^N u_n|\, d\mu \to 0$$

The theorem is proved. □

Let us take a brief look at another space of the same kind. Let

$$\mathcal{L}^2(\mathbb{R}^d) = \{f : \mathbb{R}^d \to \overline{\mathbb{R}} : |f|^2 \text{ is integrable}\}$$

and define $\|\cdot\|_2 : \mathcal{L}^2(\mathbb{R}^d) \to [0,\infty)$ by

$$\|f\|_2 = \left(\int |f|^2\, d\mu\right)^{\frac{1}{2}}$$

It turns out (see Exercise 4) that $\mathcal{L}^2(\mathbb{R}^d)$ is a vector space, and that $\|\cdot\|$ is a norm on $\mathcal{L}^2(\mathbb{R}^d)$, except that $\|f\|_2 = 0$ if $f = 0$ a.e. If we consider functions as equal if they are equal a.e., we can turn $(\mathcal{L}^2(\mathbb{R}^d), \|\cdot\|_2)$ into a normed space $(L^2(\mathbb{R}^d), |\cdot|_2)$ just as we did with $\mathcal{L}^1(\mathbb{R}^d)$. One of the advantages of this space, is that it is an inner product space with inner product

$$\langle f, g\rangle = \int fg\, d\mu$$

By almost exactly the same argument as for $L^1(\mathbb{R}^d)$, one may prove:

Theorem 5.8.2 $(L^2(\mathbb{R}^d), |\cdot|_2)$ *is complete.*

Let me finally mention that $L^1(\mathbb{R}^d)$ and $L^2(\mathbb{R}^d)$ are just two representatives of a whole family of spaces. For any $p \in [1,\infty)$, we may let

$$\mathcal{L}^p(\mathbb{R}^d) = \{f : \mathbb{R}^d \to \overline{\mathbb{R}} : |f|^p \text{ is integrable}\}$$

and define $\|\cdot\|_p : \mathcal{L}^p \to [0,\infty)$ by

$$\|f\|_2 = \left(\int |f|^2\, d\mu\right)^{\frac{1}{p}}$$

Proceeding as before, we get complete, normed spaces $(L^p(\mathbb{R}^d), |\cdot|_p)$.

Exercises for Section 5.8

1. Show that $\mathcal{L}^1(\mathbb{R}^d)$ is a vector space. Since the set of *all* functions from $\mathbb{R}^d$ to $\mathbb{R}$ is a vector space, it suffices to show that $\mathcal{L}^1(\mathbb{R}^d)$ is a subspace, i.e. that cf and $f+g$ are in $\mathcal{L}^1(\mathbb{R}^d)$ whenever $f, g \in \mathcal{L}^1(\mathbb{R}^d)$ and $c \in \mathbb{R}$.

2. Show that $\|\cdot\|_1$ satisfies the following conditions:

   (i) $\|f\|_1 \geq 0$ for all f, and $\|\mathbf{0}\|_1 = 0$ (here $\mathbf{0}$ is the function that is constant 0).

   (ii) $\|cf\|_1 = |c|\|f\|_1$ for all $f \in \mathcal{L}^1(\mathbb{R}^d)$ and all $c \in \mathbb{R}$.

   (iii) $\|f+g\|_1 \leq \|f\|_1 + \|g\|_1$ for all $f, g \in \mathcal{L}^1(\mathbb{R}^d)$

   This means that $\|\cdot\|_1$ is a *seminorm.*

3 If $f, g \in \mathcal{L}^1(\mathbb{R}^d)$, we write $f \sim g$ if $f = g$ a.e. Recall that the equivalence class $[f]$ of f is defined by

$$[f] = \{g \in \mathcal{L}(\mathbb{R}^d) \,:\, g \sim f\}$$

   a) Show that two equivalence classes $[f]$ and $[g]$ are either equal or disjoint.

   b) Show that if $f \sim f'$ and $g \sim g'$, then $f + g \sim f' + g'$. Show also that $cf \sim cf'$ for all $c \in \mathbb{R}$.

   c) Show that if $f \sim g$, then $\|f - g\|_1 = 0$ and $\|f\|_1 = \|g\|_1$.

   d) Show that the set $L^1(\mathbb{R}^d)$ of all equivalence classes is a normed space if we define scalar multiplication, addition and norm by:

      (i) $c[f] = [cf]$ for all $c \in \mathbb{R}$, $f \in \mathcal{L}^1(\mathbb{R}^d)$.

      (ii) $[f] + [g] = [f + g]$ for all $f, g \in \mathcal{L}^1(\mathbb{R}^d)$

      (iii) $|[f]|_1 = \|f\|_1$ for all $f \in \mathcal{L}^1(\mathbb{R}^d)$.

      Why do we need to establish the results in (i), (ii), and (iii) before we can make these definitions?

4. a) Show that $\mathcal{L}^2(\mathbb{R}^d)$ is a vector space. Since the set of *all* functions from $\mathbb{R}^d$ to $\mathbb{R}$ is a vector space, it suffices to show that $\mathcal{L}^2(\mathbb{R}^d)$ is a subspace, i.e. that cf and $f+g$ are in $\mathcal{L}^2(\mathbb{R}^d)$ whenever $f, g \in \mathcal{L}^2(\mathbb{R}^d)$ and $c \in \mathbb{R}$. (To show that $f + g \in \mathcal{L}^2(\mathbb{R}^d)$, you may want to use that $|a+b|^2 \leq 2|a|^2 + 2|b|^2$ for all real numbers a, b).

   b) Show that if $f, g \in \mathcal{L}^2(\mathbb{R}^d)$, then fg is integrable. (You may want to use the identity $|fg| = \frac{1}{2}((|f| + |g|)^2 - |f|^2 - |g|^2)$.

   c) Show that the *semi inner product*

$$\langle f, g\rangle = \int fg\, d\mu$$

   on $\mathcal{L}^2(\mathbb{R}^d)$ satisfies:

      (i) $\langle f, g\rangle = \langle g, f\rangle$ for all $f, g \in \mathcal{L}^2(\mathbb{R}^d)$.

      (ii) $\langle f + g, h\rangle = \langle f, h\rangle + \langle g, h\rangle$ for all $f, g, h \in \mathcal{L}^2(\mathbb{R}^d)$.

      (iii) $\langle cf, g\rangle = c\langle f, g\rangle$ for all $c \in \mathbb{R}$, $f, g \in \mathcal{L}^2(\mathbb{R}^d)$.

(iv) For all $f \in \mathcal{L}^2(\mathbb{R}^d)$, $\langle f, f\rangle \geq 0$ with equality if $f = \mathbf{0}$ (here $\mathbf{0}$ is the function that is constant 0).

Show also that $\langle f, f\rangle = 0$ if and only if $f = 0$ a.e.

e) Assume that $f, f', g, g' \in \mathcal{L}^2(\mathbb{R}^d)$, and that $f = f'$, $g = g'$ a.e. Show that $\langle f, g\rangle = \langle f', g'\rangle$

5. Show that $(L^2(\mathbb{R}^d), |\cdot|_2)$ is complete by modifying the proof of Theorem 5.8.1.