Basic Analysis II

Introduction to Real Analysis, Volume II

by Jiří Lebl

May 23, 2025
(version 6.2)

Typeset in LaTeX.

During the writing of these notes, the author was in part supported by NSF grant DMS-1362337.

The date is the main identifier of version. The major version / edition number is raised only if there have been substantial changes. From 6th edition onwards, both volumes share the same version number.

See https://www.jirka.org/ra/ for more information (including contact information, possible updates, and errata).

The LaTeX source for the book is available for possible modification and customization at github: https://github.com/jirilebl/ra

Contents

Introduction

About this book

This second volume of “Basic Analysis” is meant to be a seamless continuation. The chapter numbers start where the first volume left off. The book started with my notes for a second-semester undergraduate analysis at University of Wisconsin—Madison in 2012, which I taught more or less with Rudin’s book. Some of the material and some of the proofs are similar to Rudin, though I try to provide more detail and context. In 2016, I taught a second-semester undergraduate analysis at Oklahoma State University, modifying and cleaning up the notes, this time using them as the main text. I have since taught the course several more times, adding chapter 11 (originally written for the Wisconsin course), and making many other smaller improvments.

I plan to eventually add a few more topics. I will try to preserve the numbering in subsequent editions as always. The new topics planned would add chapters onto the end of the book, or add sections to end of existing chapters, and I will try as hard as possible to leave exercise numbers unchanged.

For the most part, this second volume depends on the non-optional parts of volume I, while some of the optional parts are also used. Higher order derivatives (but not Taylor’s theorem itself) are used in 8.6, 9.3, 10.6. Exponentials, logarithms, and improper integrals are used in a few examples and exercises, and they are heavily used in chapter 11.

An alternate plan for a two-semester course is that some bits of the first volume, such as metric spaces, are covered in the second semester, while some of the optional topics of volume I are covered in the first semester. Leaving metric spaces for the second semester makes the second semester the “multivariable” part of the course.

Several possibilities for things to cover after metric spaces, depending on time are:

1) 8.1–8.5, 10.1–10.5, 10.7 (multivariable calculus, focus on multivariable integral).
2) Chapter 8, chapter 9, 10.1 and 10.2 (multivariable calculus, focus on path integrals).
3) Chapters 8, 9, and 10 (multivariable calculus, path integrals, multivariable integrals).
4) Chapters 8, (maybe 9), and 11 (multivariable differential calculus, some advanced analysis).
5) Chapter 8, chapter 9, 11.1, 11.6, 11.7 (a simpler variation of the above).

Chapter 8

Several Variables and Partial Derivatives

8.1 Vector spaces, linear mappings, and convexity

Note: 3 lectures

8.1.1 Vector spaces

The euclidean space $\mathbb{R}^n$ has already made an appearance in the metric space chapter. In this chapter, we extend the differential calculus we created for one variable to several variables. The key idea in differential calculus is to approximate differentiable functions by linear functions (approximating the graph by a straight line). In several variables, we must introduce a little bit of linear algebra before we can move on. We start with vector spaces and linear mappings of vector spaces.

It is common to use $\vec{v}$ or the bold $\mathbf{v}$ for elements of $\mathbb{R}^n$, especially in the applied sciences. We use just plain old v, which is common in mathematics. That is, $v \in \mathbb{R}^n$ is a *vector*, which means $v = (v_1, v_2, \ldots, v_n)$ is an n-tuple of real numbers.* At times, it is useful to write and treat vectors as *column vectors*, that is, n-by-1 matrices:

$$v = (v_1, v_2, \ldots, v_n) = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$$

We do so when convenient. We call real numbers *scalars* to distinguish them from vectors.

We often think of vectors as a direction and a magnitude and draw the vector as an arrow. The vector $(v_1, v_2, \ldots, v_n)$ is represented by an arrow from the origin to the point $(v_1, v_2, \ldots, v_n)$. When we think of vectors as arrows, they are not based at the origin necessarily; a vector is simply the direction and the magnitude, and it does not know where it starts. There is a natural algebraic structure when thinking of vectors as arrows. We can add vectors as arrows by following one vector and then the other. And we can take scalar multiples of vectors as arrows by rescaling the magnitude. See Figure 8.1.

*Subscripts are used for many purposes, so sometimes we may have several vectors that may also be identified by subscript, such as a finite or infinite sequence of vectors $y_1, y_2, \ldots$.

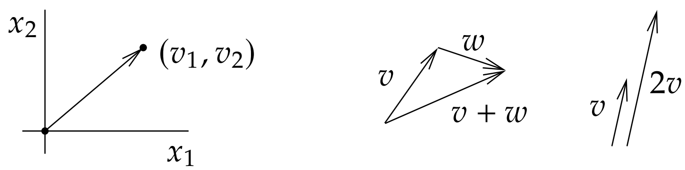

Figure 8.1: Vector as an arrow in $\mathbb{R}^2$, and the meaning of addition and scalar multiplication.

A vector may also represent a point in $\mathbb{R}^n$. Usually, we think of $v \in \mathbb{R}^n$ as a point if we are thinking of $\mathbb{R}^n$ as a metric space, and we think of it as an arrow if we think of the so-called *vector space structure* on $\mathbb{R}^n$ (addition and scalar multiplication). Let us define the abstract notion of a vector space, as there are many other vector spaces than just $\mathbb{R}^n$.

Definition 8.1.1. Let X be a set together with the operations of addition, $+ \colon X \times X \to X$, and multiplication, $\cdot \colon \mathbb{R} \times X \to X$, (we usually write ax instead of $a \cdot x$). The set X is called a *vector space* (or a *real vector space*) if the following conditions are satisfied:

(i) (Addition is associative) If $u, v, w \in X$, then $u + (v + w) = (u + v) + w$.

(ii) (Addition is commutative) If $u, v \in X$, then $u + v = v + u$.

(iii) (Additive identity) There is a $0 \in X$ such that $v + 0 = v$ for all $v \in X$.

(iv) (Additive inverse) For each $v \in X$, there is a $-v \in X$, such that $v + (-v) = 0$.

(v) (Distributive law) If $a \in \mathbb{R}$, and $u, v \in X$, then $a(u + v) = au + av$.

(vi) (Distributive law) If $a, b \in \mathbb{R}$, and $v \in X$, then $(a + b)v = av + bv$.

(vii) (Multiplication is associative) If $a, b \in \mathbb{R}$, and $v \in X$, then $(ab)v = a(bv)$.

(viii) (Multiplicative identity) $1v = v$ for all $v \in X$.

Elements of a vector space are usually called *vectors*, even if they are not elements of $\mathbb{R}^n$ (vectors in the "traditional" sense). If $Y \subset X$ is a subset that is a vector space itself using the same operations, then Y is called a *subspace* or a *vector subspace* of X.

Multiplication by scalars works as one would expect. For example, $2v = (1 + 1)v = 1v + 1v = v + v$, similarly $3v = v + v + v$, and so on. One particular fact we often use is that $0v = 0$, where the zero on the left is $0 \in \mathbb{R}$ and the zero on the right is $0 \in X$. To see this, start with $0v = (0 + 0)v = 0v + 0v$, and add $-(0v)$ to both sides to obtain $0 = 0v$. Similarly, $-v = (-1)v$, which follows by $(-1)v + v = (-1)v + 1v = (-1 + 1)v = 0v = 0$. Such algebraic facts which follow quickly from the definition will be taken for granted from now on.

Example 8.1.2: The set $\mathbb{R}^n$ is a vector space; the zero vector is $(0, \ldots, 0)$, addition and multiplication by a scalar is done componentwise: If $a \in \mathbb{R}$, $v = (v_1, v_2, \ldots, v_n) \in \mathbb{R}^n$, and $w = (w_1, w_2, \ldots, w_n) \in \mathbb{R}^n$, then

$$\begin{aligned} v + w &:= (v_1, v_2, \ldots, v_n) + (w_1, w_2, \ldots, w_n) = (v_1 + w_1, v_2 + w_2, \ldots, v_n + w_n), \\ av &:= a(v_1, v_2, \ldots, v_n) = (av_1, av_2, \ldots, av_n). \end{aligned}$$

We will mostly deal with "finite-dimensional" vector spaces that can be regarded as subsets of $\mathbb{R}^n$, but other vector spaces are useful in analysis. It is better to think of even such simpler vector spaces abstractly abstract notion rather than as $\mathbb{R}^n$.

Example 8.1.3: A trivial example of a vector space is $X := \{0\}$. The operations are defined in the obvious way: $0 + 0 := 0$ and $a0 := 0$. A zero vector must always exist, so all vector spaces are nonempty sets, and this X is the smallest possible vector space.

Example 8.1.4: The space $C([0,1],\mathbb{R})$ of continuous functions on the interval $[0,1]$ is a vector space. For two functions f and g in $C([0,1],\mathbb{R})$ and $a \in \mathbb{R}$, we make the obvious definitions of $f + g$ and af:

$$(f+g)(x) := f(x) + g(x), \qquad (af)(x) := a\bigl(f(x)\bigr).$$

The 0 is the function that is identically zero. We leave it as an exercise to check that all the vector space conditions are satisfied. The space $C^1([0,1],\mathbb{R})$ of continuously differentiable functions is a subspace of $C([0,1],\mathbb{R})$.

Example 8.1.5: The space of polynomials $c_0 + c_1 t + c_2 t^2 + \cdots + c_m t^m$ (of arbitrary degree m) is a vector space, denoted by $\mathbb{R}[t]$ (coefficients are real and the variable is t). The operations are defined in the same way as for functions above. Suppose there are two polynomials, one of degree m and one of degree n. Assume $n \geq m$ for simplicity. Then

$$\begin{aligned}(c_0 + c_1 t + c_2 t^2 + \cdots + c_m t^m) + (d_0 + d_1 t + d_2 t^2 + \cdots + d_n t^n) = \\ (c_0 + d_0) + (c_1 + d_1) t + (c_2 + d_2) t^2 + \cdots + (c_m + d_m) t^m + d_{m+1} t^{m+1} + \cdots + d_n t^n\end{aligned}$$

and

$$a(c_0 + c_1 t + c_2 t^2 + \cdots + c_m t^m) = (ac_0) + (ac_1) t + (ac_2) t^2 + \cdots + (ac_m) t^m.$$

Despite what it looks like, $\mathbb{R}[t]$ is not equivalent to $\mathbb{R}^n$ for any n. In particular, it is not "finite-dimensional." We will make this notion precise in just a little bit. One can make a finite-dimensional vector subspace by restricting the degree. For example, if $\mathcal{P}_n$ is the set of polynomials of degree n or less, then $\mathcal{P}_n$ is a finite-dimensional vector space, and we could identify it with $\mathbb{R}^{n+1}$.

Above, the variable t is really just a formal placeholder. By setting t equal to a real number, we obtain a function. So the space $\mathbb{R}[t]$ can be thought of as a subspace of $C(\mathbb{R},\mathbb{R})$. If we restrict the range of t to $[0,1]$, $\mathbb{R}[t]$ can be identified with a subspace of $C([0,1],\mathbb{R})$.

Proposition 8.1.6. *For $S \subset X$ to be a vector subspace of a vector space X, we only need to check*:

1) $0 \in S$.
2) *S is closed under addition*: *If $x, y \in S$, then $x + y \in S$.*
3) *S is closed under scalar multiplication*: *If $x \in S$ and $a \in \mathbb{R}$, then $ax \in S$.*

Items 2) and 3) ensure that addition and scalar multiplication are indeed defined on S. Item 1) is required to fulfill item (iii) from the definition of vector space. Existence of additive inverse $-v$, item (iv), follows because $-v = (-1)v$ and item 3) says that $-v \in S$ if $v \in S$. All other properties are certain equalities that are already satisfied in X and thus must be satisfied in a subset.

It is possible to use other fields than $\mathbb{R}$ in the definition (for example, it is common to use the complex numbers $\mathbb{C}$), but let us stick with the real numbers*.

8.1.2 Linear combinations and dimension

Definition 8.1.7. Suppose X is a vector space, $x_1, x_2, \ldots, x_m \in X$ are vectors, and $a_1, a_2, \ldots, a_m \in \mathbb{R}$ are scalars. Then

$$a_1 x_1 + a_2 x_2 + \cdots + a_m x_m$$

is called a *linear combination* of the vectors $x_1, x_2, \ldots, x_m$.

For a subset $Y \subset X$, let $\operatorname{span}(Y)$, or the *span* of Y, be the set of all linear combinations of all finite subsets of Y. We say Y *spans* $\operatorname{span}(Y)$. By convention, define $\operatorname{span}(\emptyset) := \{0\}$.

Example 8.1.8: Let $Y := \bigl\{(1,1)\bigr\} \subset \mathbb{R}^2$. Then

$$\operatorname{span}(Y) = \bigl\{(x,x) \in \mathbb{R}^2 : x \in \mathbb{R}\bigr\}.$$

That is, $\operatorname{span}(Y)$ is the line through the origin and the point $(1,1)$.

Example 8.1.9: Let $Y := \bigl\{(1,1),(0,1)\bigr\} \subset \mathbb{R}^2$. Then

$$\operatorname{span}(Y) = \mathbb{R}^2,$$

as every point $(x,y) \in \mathbb{R}^2$ can be written as a linear combination

$$(x,y) = x(1,1) + (y-x)(0,1).$$

Example 8.1.10: Let $Y := \{1, t, t^2, t^3, \ldots\} \subset \mathbb{R}[t]$ and $E := \{1, t^2, t^4, t^6, \ldots\} \subset \mathbb{R}[t]$. The span of Y is all polynomials,

$$\operatorname{span}(Y) = \mathbb{R}[t].$$

The span of E is the set of polynomials with even powers of t only.

Suppose we have two linear combinations of vectors from Y. One linear combination uses the vectors $\{x_1, x_2, \ldots, x_m\}$, and the other uses $\{\widetilde{x}_1, \widetilde{x}_2, \ldots, \widetilde{x}_n\}$. We can write their sum using vectors from the union $\{x_1, x_2, \ldots, x_m\} \cup \{\widetilde{x}_1, \widetilde{x}_2, \ldots, \widetilde{x}_n\}$:

$$\begin{aligned}(a_1 x_1 + a_2 x_2 + \cdots + a_m x_m) &+ (b_1 \widetilde{x}_1 + b_2 \widetilde{x}_2 + \cdots + b_n \widetilde{x}_n) \\ &= a_1 x_1 + a_2 x_2 + \cdots + a_m x_m + b_1 \widetilde{x}_1 + b_2 \widetilde{x}_2 + \cdots + b_n \widetilde{x}_n.\end{aligned}$$

*If you want a very funky vector space over a different field, $\mathbb{R}$ itself is a vector space over the field $\mathbb{Q}$.

So the sum is also a linear combination of vectors from Y. Similarly, a scalar multiple of a linear combination of vectors from Y is a linear combination of vectors from Y:

$$b(a_1x_1 + a_2x_2 + \cdots + a_mx_m) = ba_1x_1 + ba_2x_2 + \cdots + ba_mx_m.$$

Finally, $0 \in \operatorname{span}(Y)$; if Y is nonempty, $0 = 0v$ for some $v \in Y$. We have proved the following proposition.

Proposition 8.1.11. *Let X be a vector space and $Y \subset X$ is a subset. Then the set* $\operatorname{span}(Y)$ *is a vector subspace of X.*

Every linear combination of elements in a subspace is an element of that subspace. So $\operatorname{span}(Y)$ is the smallest subspace that contains Y. In particular, if Y is already a vector subspace, then $\operatorname{span}(Y) = Y$.

Definition 8.1.12. A set of vectors $\{x_1, x_2, \ldots, x_m\} \subset X$ is *linearly independent** if the equation

$$a_1x_1 + a_2x_2 + \cdots + a_mx_m = 0 \tag{8.1}$$

has only the trivial solution $a_1 = a_2 = \cdots = a_m = 0$. By convention, $\emptyset$ is linearly independent. A set that is not linearly independent is *linearly dependent*. A linearly independent set of vectors $B \subset X$ such that $\operatorname{span}(B) = X$ is called a *basis* of X. We generally consider the basis as not just a set, but as an ordered m-tuple: $x_1, x_2, \ldots, x_m$.

Suppose d is largest integer for which X contains a set of d linearly independent vectors. We then say d is the *dimension* of X, and we write $\dim\, X := d$. If X contains a set of d linearly independent vectors for arbitrarily large d, we say X is infinite-dimensional and write $\dim\, X := \infty$. For the trivial vector space $\{0\}$, we define $\dim\,\{0\} := 0$.

A subset of a linearly independent set is clearly linearly independent. So if a set contains d linearly independent vectors, it also contains a set of m linearly independent vectors for all $m \leq d$. Moreover, if a set does not have $d+1$ linearly independent vectors, no set of more than $d+1$ vectors is linearly independent. So X is of dimension is d if there is a set of d linearly independent vectors, but no set of $d+1$ vectors is linearly independent.

No element of a linearly independent set can be zero, and a set with one nonzero element is always linearly independent. In particular, $\{0\}$ is the only vector space of dimension 0. Every other vector space has a positive dimension or is infinite-dimensional. As the empty set is linearly independent, it is a basis of $\{0\}$.

As an example, the set Y of the two vectors in Example 8.1.9 is a basis of $\mathbb{R}^2$, and so $\dim\, \mathbb{R}^2 \geq 2$. We will see in a moment that every vector subspace of $\mathbb{R}^n$ has a finite dimension, and that dimension is less than or equal to n. So every set of 3 vectors in $\mathbb{R}^2$ is linearly dependent, and $\dim\, \mathbb{R}^2 = 2$.

If a set is linearly dependent, then one of the vectors is a linear combination of the others. In (8.1), if $a_k \neq 0$, then we solve for x_k:

$$x_k = \frac{-a_1}{a_k}x_1 + \cdots + \frac{-a_{k-1}}{a_k}x_{k-1} + \frac{-a_{k+1}}{a_k}x_{k+1} + \cdots + \frac{-a_m}{a_m}x_m.$$

*For an infinite set $Y \subset X$, we say Y is linearly independent if every finite subset of Y is linearly independent in the sense given. However, this situation only comes up in infinitely many dimensions.

The vector x_k has at least two different representations as linear combinations of the vectors $\{x_1, x_2, \ldots, x_m\}$. The one above and x_k itself. For instance, the set $\{(0,1),(2,3),(5,0)\}$ in $\mathbb{R}^2$ is linearly dependent:

$$3(0,1) - (2,3) + 2(1,0) = 0, \qquad \text{so} \qquad (2,3) = 3(0,1) + 2(1,0).$$

Proposition 8.1.13. *Suppose a vector space X has a basis $B = \{x_1, x_2, \ldots, x_n\}$. Then every $y \in X$ has a unique representation of the form*

$$y = \sum_{k=1}^{n} a_k \, x_k$$

for some scalars $a_1, a_2, \ldots, a_n$.

Proof. As X is the span of B, every $y \in X$ is a linear combination of elements of B. Suppose

$$y = \sum_{k=1}^{n} a_k \, x_k = \sum_{k=1}^{n} b_k \, x_k.$$

Then

$$\sum_{k=1}^{n} (a_k - b_k) x_k = 0.$$

By linear independence of the basis, $a_k = b_k$ for all k, and so the representation is unique. □

For $\mathbb{R}^n$, we define the *standard basis* of $\mathbb{R}^n$:

$$e_1 := (1,0,0,\ldots,0), \quad e_2 := (0,1,0,\ldots,0), \quad \ldots, \quad e_n := (0,0,0,\ldots,1).$$

We use the same letters e_k for any $\mathbb{R}^n$, and which space $\mathbb{R}^n$ we are working in is understood from context. A direct computation shows that $\{e_1, e_2, \ldots, e_n\}$ really is a basis of $\mathbb{R}^n$; it spans $\mathbb{R}^n$ and is linearly independent. In fact,

$$a = (a_1, a_2, \ldots, a_n) = \sum_{k=1}^{n} a_k \, e_k.$$

Proposition 8.1.14. *Let X be a vector space and d a nonnegative integer.*

(*i*) *If X is spanned by d vectors, then* $\dim X \leq d$.

(*ii*) *If T is a linearly independent set and $v \in X \setminus \operatorname{span}(T)$, then $T \cup \{v\}$ is linearly independent.*

(*iii*) $\dim X = d$ *if and only if X has a basis of d vectors. In particular,* $\dim \mathbb{R}^n = n$.

(*iv*) *If $Y \subset X$ is a vector subspace and* $\dim X = d$, *then* $\dim Y \leq d$.

(*v*) *If* $\dim X = d$ *and a set T of d vectors spans X, then T is linearly independent.*

(*vi*) *If* $\dim X = d$ *and a set T of m vectors is linearly independent, then there is a set S of $d - m$ vectors such that $T \cup S$ is a basis of X.*

In particular, the last item says that if $\dim X = d$ and T is a set of d linearly independent vectors, then T spans X. Another thing to note is that item (iii) implies that every basis of a finite dimensional vector space has the same number of elements.

Proof. All statements hold trivially when $d = 0$, so assume $d \geq 1$.

We start with (i). Suppose $S := \{x_1, x_2, \ldots, x_d\}$ spans X, and $T := \{y_1, y_2, \ldots, y_m\}$ is a linearly independent subset of X. We wish to show that $m \leq d$. As S spans X, write

$$y_1 = \sum_{k=1}^{d} a_{k,1} x_k,$$

for some numbers $a_{1,1}, a_{2,1}, \ldots, a_{d,1}$. One of the $a_{k,1}$ is nonzero, otherwise y_1 would be zero. Without loss of generality, suppose $a_{1,1} \neq 0$. Solve

$$x_1 = \frac{1}{a_{1,1}} y_1 - \sum_{k=2}^{d} \frac{a_{k,1}}{a_{1,1}} x_k.$$

In particular, $\{y_1, x_2, \ldots, x_d\}$ spans X, since x_1 can be obtained from $\{y_1, x_2, \ldots, x_d\}$. Therefore, there are some numbers for some numbers $a_{1,2}, a_{2,2}, \ldots, a_{d,2}$, such that

$$y_2 = a_{1,2} y_1 + \sum_{k=2}^{d} a_{k,2} x_k.$$

As T is linearly independent—and so $\{y_1, y_2\}$ is linearly independent—one of the $a_{k,2}$ for $k \geq 2$ must be nonzero. Without loss of generality suppose $a_{2,2} \neq 0$. Solve

$$x_2 = \frac{1}{a_{2,2}} y_2 - \frac{a_{1,2}}{a_{2,2}} y_1 - \sum_{k=3}^{d} \frac{a_{k,2}}{a_{2,2}} x_k.$$

In particular, $\{y_1, y_2, x_3, \ldots, x_d\}$ spans X.

We continue this procedure. If $m < d$, we are done. Suppose $m \geq d$. After d steps, we obtain that $\{y_1, y_2, \ldots, y_d\}$ spans X. Any other vector v in X is a linear combination of $\{y_1, y_2, \ldots, y_d\}$ and hence cannot be in T as T is linearly independent. So $m = d$.

We continue with (ii). Suppose $T = \{x_1, x_2, \ldots, x_m\}$ is linearly independent, does not span X, and $v \in X \setminus \operatorname{span}(T)$. Suppose $a_1 x_1 + a_2 x_2 + \cdots + a_m x_m + a_{m+1} v = 0$ for some scalars $a_1, a_2, \ldots, a_{m+1}$. If $a_{m+1} \neq 0$, then v would be a linear combination of T, so $a_{m+1} = 0$. Then, as T is linearly independent, $a_1 = a_2 = \cdots = a_m = 0$. So $T \cup \{v\}$ is linearly independent.

We move to (iii). If $\dim X = d$, then there must exist some linearly independent set T of d vectors, and T must span X, otherwise we could choose a larger set of linearly independent vectors via (ii). So we have a basis of d vectors. On the other hand, if we have a basis of d vectors, the dimension is at least d as a basis is linearly independent. A basis also spans X, and so by (i) we know that dimension is at most d. Hence the dimension of X must equal d. The "in particular" follows by noting that $\{e_1, e_2, \ldots, e_n\}$ is a basis of $\mathbb{R}^n$.

To see (iv), suppose $Y \subset X$ is a vector subspace, where $\dim\, X = d$. As X cannot contain $d+1$ linearly independent vectors, neither can Y.

For (v), suppose T is a set of m vectors that is linearly dependent and spans X. We will show that $m > d$. One of the vectors is a linear combination of the others. If we remove it from T, we obtain a set of $m-1$ vectors that still span X. Hence $d = \dim\, X \leq m-1$ by (i).

For (vi) suppose $T = \{x_1, x_2, \ldots, x_m\}$ is a linearly independent set. First, $m \leq d$ by definition of dimension. If $m = d$, the set T must span X as in the proof of (iii), otherwise we could add another vector to T. If $m < d$, T cannot span X by (iii). So find v not in the span of T. Via (ii), the set $T \cup \{v\}$ is a linearly independent set of $m+1$ elements. Therefore, we repeat this procedure $d-m$ times to find a set of d linearly independent vectors. Again, they must span X, otherwise we could add yet another vector. □

8.1.3 Linear mappings

When $Y \neq \mathbb{R}$, a function $f \colon X \to Y$ is often called a *mapping* or a *map* rather than a *function*.

Definition 8.1.15. A map $A \colon X \to Y$ of vector spaces X and Y is *linear* (we also say A is a *linear transformation* or a *linear operator*) if for all $a \in \mathbb{R}$ and all $x, y \in X$,

$$A(ax) = aA(x) \qquad \text{and} \qquad A(x+y) = A(x) + A(y).$$

We usually write Ax instead of $A(x)$ if A is linear. If A is one-to-one and onto, then we say A is *invertible*, and we denote the inverse by A^{-1}. If $A \colon X \to X$ is linear, then we say A is a *linear operator on* X.

We write $L(X,Y)$ for the set of linear maps from X to Y, and $L(X)$ for the set of linear operators on X. If $a \in \mathbb{R}$ and $A, B \in L(X,Y)$, define the maps aA and $A+B$ by

$$(aA)(x) := aAx, \qquad (A+B)(x) := Ax + Bx.$$

If $A \in L(Y,Z)$ and $B \in L(X,Y)$, define the map $AB \colon X \to Z$ as the composition $A \circ B$,

$$ABx := A(Bx).$$

Finally, denote by $I \in L(X)$ the *identity*: the linear operator such that $Ix = x$ for all x.

Proposition 8.1.16. *Let X, Y, and Z be vector spaces.*

(i) *If $A \in L(X,Y)$, then $A0 = 0$.*

(ii) *If $A, B \in L(X,Y)$, then $A + B \in L(X,Y)$.*

(iii) *If $A \in L(X,Y)$ and $a \in \mathbb{R}$, then $aA \in L(X,Y)$.*

(iv) *If $A \in L(Y,Z)$ and $B \in L(X,Y)$, then $AB \in L(X,Z)$.*

(v) *If $A \in L(X,Y)$ is invertible, then $A^{-1} \in L(Y,X)$.*

In particular, $L(X,Y)$ is a vector space, where $0 \in L(X,Y)$ is the linear map that takes everything to 0. As $L(X)$ is not only a vector space, but also admits a product (composition), it is called an *algebra*.

Proof. We leave the first four items as a quick exercise, Exercise 8.1.20. Let us prove the last item. Let $a \in \mathbb{R}$ and $y \in Y$. As A is onto, then there is an $x \in X$ such that $y = Ax$. As it is also one-to-one, $A^{-1}(Az) = z$ for all $z \in X$. So

$$A^{-1}(ay) = A^{-1}(aAx) = A^{-1}\big(A(ax)\big) = ax = aA^{-1}(y).$$

Similarly, let $y_1, y_2 \in Y$ and $x_1, x_2 \in X$ be such that $Ax_1 = y_1$ and $Ax_2 = y_2$, then

$$A^{-1}(y_1 + y_2) = A^{-1}(Ax_1 + Ax_2) = A^{-1}\big(A(x_1 + x_2)\big) = x_1 + x_2 = A^{-1}(y_1) + A^{-1}(y_2). \qquad \square$$

Proposition 8.1.17. *If $A \in L(X,Y)$ is linear, then it is completely determined by its values on a basis of X. Furthermore, if B is a basis of X, then every function $\widetilde{A} \colon B \to Y$ extends to a linear function A on X.*

We only prove this proposition for finite-dimensional spaces, as we do not need infinite-dimensional spaces.*

Proof. Let $\{x_1, x_2, \ldots, x_n\}$ be a basis of X, and let $y_k \coloneqq Ax_k$. Every $x \in X$ has a unique representation

$$x = \sum_{k=1}^{n} b_k \, x_k$$

for some numbers $b_1, b_2, \ldots, b_n$. By linearity,

$$Ax = A \sum_{k=1}^{n} b_k x_k = \sum_{k=1}^{n} b_k \, Ax_k = \sum_{k=1}^{n} b_k \, y_k.$$

The "furthermore" follows by setting $y_k \coloneqq \widetilde{A}(x_k)$, and then for $x = \sum_{k=1}^{n} b_k \, x_k$, defining the extension as $A(x) \coloneqq \sum_{k=1}^{n} b_k \, y_k$. The function is well-defined by uniqueness of the representation of x. We leave it to the reader to check that A is linear. $\square$

For a linear map, it is sufficient to check injectivity at the origin. That is, if the only x such that $Ax = 0$ is $x = 0$, then A is one-to-one, because if $Ay = Az$, then $A(y - z) = 0$. For this reason, one often studies the *nullspace* of A, that is, $\{x \in X : Ax = 0\}$. For finite-dimensional vector spaces (and only in finitely many dimensions) we have the following special case of the so-called rank-nullity theorem from linear algebra.

Proposition 8.1.18. *If X is a finite-dimensional vector space and $A \in L(X)$, then A is one-to-one if and only if it is onto.*

Proof. Let $\{x_1, x_2, \ldots, x_n\}$ be a basis for X. First suppose A is one-to-one. Let $c_1, c_2, \ldots, c_n$ be scalars such that

$$0 = \sum_{k=1}^{n} c_k \, Ax_k = A \sum_{k=1}^{n} c_k \, x_k.$$

*For infinite-dimensional spaces, the proof is essentially the same, but a little trickier to write. Moreover, we haven't even defined what a basis is for infinite-dimensional spaces.

As A is one-to-one, the only vector that is taken to 0 is 0 itself. Hence,

$$0 = \sum_{k=1}^{n} c_k \, x_k,$$

and so $c_k = 0$ for all k as $\{x_1, x_2, \dots, x_n\}$ is a basis. So $\{Ax_1, Ax_2, \dots, Ax_n\}$ is linearly independent. By Proposition 8.1.14 and the fact that the dimension is n, we conclude $\{Ax_1, Ax_2, \dots, Ax_n\}$ spans X. Consequently, A is onto, as any $y \in X$ can be written as

$$y = \sum_{k=1}^{n} a_k \, Ax_k = A \sum_{k=1}^{n} a_k \, x_k.$$

For the other direction, suppose A is onto. Suppose that for some $c_1, c_2, \dots, c_n$,

$$0 = A \sum_{k=1}^{n} c_k \, x_k = \sum_{k=1}^{n} c_k \, Ax_k.$$

As A is determined by the action on the basis, $\{Ax_1, Ax_2, \dots, Ax_n\}$ spans X. So by Proposition 8.1.14, the set is linearly independent, and $c_k = 0$ for all k. In other words, if $Ax = 0$, then $x = 0$. Thus, A is one-to-one. □

We leave the proof of the next proposition as an exercise.

Proposition 8.1.19. *If X and Y are finite-dimensional vector spaces, then $L(X, Y)$ is also finite-dimensional.*

We can identify a finite-dimensional vector space X of dimension n with $\mathbb{R}^n$, provided we fix a basis $\{x_1, x_2, \dots, x_n\}$ in X. That is, we define a bijective linear map $A \in L(X, \mathbb{R}^n)$ by $Ax_k := e_k$, where $\{e_1, e_2, \dots, e_n\}$ is the standard basis in $\mathbb{R}^n$. We have the correspondence

$$\sum_{k=1}^{n} c_k \, x_k \in X \quad \stackrel{A}{\mapsto} \quad (c_1, c_2, \dots, c_n) \in \mathbb{R}^n.$$

8.1.4 Convexity

A subset U of a vector space is *convex* if whenever $x, y \in U$, the line segment from x to y lies in U. That is, if the *convex combination* $(1-t)x + ty$ is in U for all $t \in [0, 1]$. We write $[x, y]$ for this line segment. See Figure 8.2.

In $\mathbb{R}$, convex sets are precisely the intervals, which are also precisely the connected sets. In two or more dimensions there are lots of nonconvex connected sets. For example, the set $\mathbb{R}^2 \setminus \{0\}$ is connected, but not convex—for any $x \in \mathbb{R}^2 \setminus \{0\}$ where $y := -x$, we find $(1/2)x + (1/2)y = 0$, which is not in the set. Balls (in the standard metric) in $\mathbb{R}^n$ are convex. It is a useful enough result to state as a proposition, but we leave its proof as an exercise.

Proposition 8.1.20. *Let $x \in \mathbb{R}^n$ and $r > 0$. The ball $B(x, r) \subset \mathbb{R}^n$ is convex.*

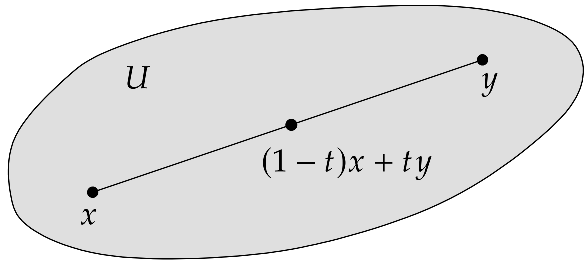

Figure 8.2: Convexity.

Example 8.1.21: A convex combination is, in particular, a linear combination. So every vector subspace V of a vector space X is convex.

Example 8.1.22: Let $C([0,1],\mathbb{R})$ be the vector space of continuous real-valued functions on $\mathbb{R}$. Let $V \subset C([0,1],\mathbb{R})$ be the set of those f such that

$$\int_0^1 f(x)\,dx \leq 1 \qquad \text{and} \qquad f(x) \geq 0 \ \text{ for all } x \in [0,1].$$

Then V is convex. Take $t \in [0,1]$, and note that if $f,g \in V$, then $(1-t)f(x)+tg(x) \geq 0$ for all x. Furthermore,

$$\int_0^1 \big((1-t)f(x)+tg(x)\big)\,dx = (1-t)\int_0^1 f(x)\,dx + t\int_0^1 g(x)\,dx \leq 1.$$

Note that V is not a vector subspace of $C([0,1],\mathbb{R})$. The function $f(x) := 1$ is in V, but $2f$ and $-f$ is not.

Proposition 8.1.23. *The intersection of two convex sets is convex. In fact, if $\{C_\lambda\}_{\lambda\in I}$ is an arbitrary collection of convex sets in a vector space, then*

$$C := \bigcap_{\lambda \in I} C_\lambda \qquad \textit{is convex.}$$

Proof. If $x,y \in C$, then $x,y \in C_\lambda$ for all $\lambda \in I$, and hence if $t \in [0,1]$, then $(1-t)x+ty \in C_\lambda$ for all $\lambda \in I$. Therefore, $(1-t)x+ty \in C$ and C is convex. □

A useful construction using intersections of convex sets is the *convex hull*. Given a subset S of a vector space X, define the convex hull of S as the intersection of all convex sets containing S:

$$\operatorname{co}(S) := \bigcap \{C \subset X : S \subset C, \text{ and } C \text{ is convex}\}.$$

That is, the convex hull is the smallest convex set containing S. By Proposition 8.1.23, the intersection of convex sets is convex. Hence the convex hull is convex.

Example 8.1.24: The convex hull of $\{0,1\}$ in $\mathbb{R}$ is $[0,1]$. Proof: A convex set containing 0 and 1 must contain $[0,1]$, so $[0,1] \subset \operatorname{co}(\{0,1\})$. The set $[0,1]$ is convex and contains $\{0,1\}$, so $\operatorname{co}(\{0,1\}) \subset [0,1]$.

Linear mappings preserve convex sets. So in some sense, convex sets are the right sort of sets when considering linear mappings or changes of coordinates.

Proposition 8.1.25. *Let X, Y be vector spaces, $A \in L(X,Y)$, and let $C \subset X$ be convex. Then $A(C)$ is convex.*

Proof. Take two points $p, q \in A(C)$. Pick $u, v \in C$ such that $Au = p$ and $Av = q$. As C is convex, then $(1-t)u + tv \in C$ for all $t \in [0,1]$, so

$$(1-t)p + tq = (1-t)Au + tAv = A\big((1-t)u + tv\big) \in A(C). \qquad \square$$

8.1.5 Exercises

***Exercise* 8.1.1:** *Show that in $\mathbb{R}^n$ (with the standard euclidean metric), for every $x \in \mathbb{R}^n$ and every $r > 0$, the ball $B(x,r)$ is convex.*

***Exercise* 8.1.2:** *Verify that $\mathbb{R}^n$ is a vector space.*

***Exercise* 8.1.3:** *Let X be a vector space. Prove that a finite set of vectors $\{x_1, x_2, \ldots, x_n\} \subset X$ is linearly independent if and only if for every $k = 1, 2, \ldots, n$*

$$\operatorname{span}\big(\{x_1, \ldots, x_{k-1}, x_{k+1}, \ldots, x_n\}\big) \subsetneq \operatorname{span}\big(\{x_1, x_2, \ldots, x_n\}\big).$$

That is, the span of the set with one vector removed is strictly smaller.

***Exercise* 8.1.4:** *Show that the set $X \subset C([0,1],\mathbb{R})$ of those functions such that $\int_0^1 f = 0$ is a vector subspace. Compare Exercise 8.1.16.*

***Exercise* 8.1.5** (Challenging)**:** *Prove $C([0,1],\mathbb{R})$ is an infinite-dimensional vector space where the operations are defined in the obvious way: $s = f + g$ and $m = af$ are defined as $s(x) := f(x) + g(x)$ and $m(x) := af(x)$. Hint: For the dimension, think of functions that are only nonzero on the interval $(1/n+1, 1/n)$.*

***Exercise* 8.1.6:** *Let $k \colon [0,1]^2 \to \mathbb{R}$ be continuous. Show that $L \colon C([0,1],\mathbb{R}) \to C([0,1],\mathbb{R})$ defined by*

$$Lf(y) := \int_0^1 k(x,y) f(x)\,dx$$

is a linear operator. That is, first show that L is well-defined by showing that Lf is continuous whenever f is, and then showing that L is linear.

***Exercise* 8.1.7:** *Let $\mathcal{P}_n$ be the vector space of polynomials in one variable of degree n or less. Show that $\mathcal{P}_n$ is a vector space of dimension $n+1$.*

***Exercise* 8.1.8:** *Let $\mathbb{R}[t]$ be the vector space of polynomials in one variable t. Let $D \colon \mathbb{R}[t] \to \mathbb{R}[t]$ be the derivative operator (derivative in t). Show that D is a linear operator.*

Exercise* 8.1.9*: *Let us show that Proposition 8.1.18 only works in finite dimensions. Take the space of polynomials $\mathbb{R}[t]$ and define the operator $A \colon \mathbb{R}[t] \to \mathbb{R}[t]$ by $A\bigl(P(t)\bigr) := tP(t)$. Show that A is linear and one-to-one, but show that it is not onto.*

Exercise* 8.1.10*: *Finish the proof of Proposition 8.1.17 in the finite-dimensional case. That is, suppose $\{x_1, x_2, \ldots x_n\}$ is a basis of X, $\{y_1, y_2, \ldots y_n\} \subset Y$, and define a function*

$$A(x) := \sum_{k=1}^{n} b_k y_k, \qquad \textit{if} \quad x = \sum_{k=1}^{n} b_k x_k.$$

Prove that $A \colon X \to Y$ is linear.

Exercise* 8.1.11*: *Prove Proposition 8.1.19. Hint: A linear transformation is determined by its action on a basis. So given two bases $\{x_1, \ldots, x_n\}$ and $\{y_1, \ldots, y_m\}$ for X and Y respectively, consider the linear operators A_{jk} that send $A_{jk}x_j = y_k$, and $A_{jk}x_\ell = 0$ if $\ell \neq j$.*

Exercise* 8.1.12** (Easy):*** *Suppose X and Y are vector spaces and $A \in L(X,Y)$ is a linear operator.*

a) Show that the nullspace $N := \{x \in X : Ax = 0\}$ is a vector space.

b) Show that the range $R := \{y \in Y : Ax = y \text{ for some } x \in X\}$ is a vector space.

Exercise* 8.1.13** (Easy):*** *Show by example that a union of convex sets need not be convex.*

Exercise* 8.1.14*: *Compute the convex hull of the set of 3 points $\bigl\{(0,0), (0,1), (1,1)\bigr\}$ in $\mathbb{R}^2$.*

Exercise* 8.1.15*: *Show that the set $\bigl\{(x,y) \in \mathbb{R}^2 : y > x^2\bigr\}$ is a convex set.*

Exercise* 8.1.16*: *Show that the set $X \subset C([0,1],\mathbb{R})$ of those functions such that $\int_0^1 f = 1$ is a convex set, but not a vector subspace. Compare Exercise 8.1.4.*

Exercise* 8.1.17*: *Show that every convex set in $\mathbb{R}^n$ is connected using the standard topology on $\mathbb{R}^n$.*

Exercise* 8.1.18*: *Suppose $K \subset \mathbb{R}^2$ is a convex set such that the only point of the form $(x,0)$ in K is the point $(0,0)$. Further suppose that $(0,1) \in K$ and $(1,1) \in K$. Show that if $(x,y) \in K$ and $x \neq 0$, then $y > 0$.*

Exercise* 8.1.19*: *Prove that an arbitrary intersection of vector subspaces is a vector subspace. That is, if X is a vector space and $\{V_\lambda\}_{\lambda \in I}$ is an arbitrary collection of vector subspaces of X, then $\bigcap_{\lambda \in I} V_\lambda$ is a vector subspace of X.*

Exercise* 8.1.20** (Easy):*** *Finish the proof of Proposition 8.1.16, that is, prove the first four items of the proposition.*

8.2 Analysis with vector spaces

Note: 3 lectures

8.2.1 Norms

Let us start measuring the size of vectors and hence distance.

Definition 8.2.1. If X is a vector space, then we say a function $\|\cdot\| : X \to \mathbb{R}$ is a *norm* if

(i) $\|x\| \geq 0$, with $\|x\| = 0$ if and only if $x = 0$.

(ii) $\|cx\| = |c| \, \|x\|$ for all $c \in \mathbb{R}$ and $x \in X$.

(iii) $\|x + y\| \leq \|x\| + \|y\|$ for all $x, y \in X$ (*triangle inequality*).

A vector space equipped with a norm is called a *normed vector space*.

Given a norm (any norm) on a vector space X, define a distance $d(x, y) \coloneqq \|x - y\|$, which makes X into a metric space (exercise). So what you know about metric spaces applies to normed vector spaces. Before defining the standard norm on $\mathbb{R}^n$, we define the standard scalar *dot product* on $\mathbb{R}^n$. For $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ and $y = (y_1, y_2, \ldots, y_n) \in \mathbb{R}^n$ define

$$x \cdot y \coloneqq \sum_{k=1}^{n} x_k \, y_k .$$

Dot product is linear in each variable separately—in more fancy language, it is *bilinear*. That is, if y is fixed, the map $x \mapsto x \cdot y$ is a linear map from $\mathbb{R}^n$ to $\mathbb{R}$. Similarly, if x is fixed, $y \mapsto x \cdot y$ is linear. It is *symmetric* in the sense that $x \cdot y = y \cdot x$. Define the *euclidean norm* as

$$\|x\| \coloneqq \|x\|_{\mathbb{R}^n} \coloneqq \sqrt{x \cdot x} = \sqrt{(x_1)^2 + (x_2)^2 + \cdots + (x_n)^2}.$$

We will normally write $\|x\|$, only in the rare instance when it is necessary to emphasize that we are talking about the euclidean norm will we write $\|x\|_{\mathbb{R}^n}$. Unless otherwise stated, if we talk about $\mathbb{R}^n$ as a normed vector space, we mean the standard euclidean norm. It is easy to see that the euclidean norm satisfies (i) and (ii). To prove that (iii) holds, the key inequality is the so-called Cauchy–Schwarz inequality we saw before. As this inequality is so important, we state and prove a slightly stronger version using the notation of this chapter.

Theorem 8.2.2 (Cauchy–Schwarz inequality)**.** *Let $x, y \in \mathbb{R}^n$, then*

$$|x \cdot y| \leq \|x\| \, \|y\| = \sqrt{x \cdot x}\, \sqrt{y \cdot y},$$

with equality if and only if $x = \lambda y$ or $y = \lambda x$ for some $\lambda \in \mathbb{R}$.

Proof. If $x = 0$ or $y = 0$, then the theorem holds trivially. So assume $x \neq 0$ and $y \neq 0$.

If x is a scalar multiple of y, that is, $x = \lambda y$ for some $\lambda \in \mathbb{R}$, then the theorem holds with equality:

$$|x \cdot y| = |\lambda y \cdot y| = |\lambda|\,|y \cdot y| = |\lambda|\,\|y\|^2 = \|\lambda y\|\,\|y\| = \|x\|\,\|y\|.$$

Fixing x and y, $\|x + ty\|^2$ is a quadratic polynomial as a function of t:

$$\|x + ty\|^2 = (x + ty) \cdot (x + ty) = x \cdot x + x \cdot ty + ty \cdot x + ty \cdot ty = \|x\|^2 + 2t(x \cdot y) + t^2\|y\|^2.$$

If x is not a scalar multiple of y, then $\|x + ty\|^2 > 0$ for all t. So the polynomial $\|x + ty\|^2$ is never zero. Elementary algebra says that the discriminant must be negative:

$$4(x \cdot y)^2 - 4\|x\|^2\|y\|^2 < 0.$$

In other words, $(x \cdot y)^2 < \|x\|^2\|y\|^2$. □

Item (iii), the triangle inequality in $\mathbb{R}^n$, follows from:

$$\|x + y\|^2 = x \cdot x + y \cdot y + 2(x \cdot y) \leq \|x\|^2 + \|y\|^2 + 2\big(\|x\|\,\|y\|\big) = \big(\|x\| + \|y\|\big)^2.$$

The distance $d(x, y) := \|x - y\|$ is the standard distance (standard metric) on $\mathbb{R}^n$ that we used when we talked about metric spaces.

Definition 8.2.3. Let $A \in L(X, Y)$. Define

$$\|A\| := \sup\big\{\|Ax\| : x \in X \text{ with } \|x\| = 1\big\}.$$

The number $\|A\|$ (possibly ∞) is called the *operator norm*. We will see below that it is indeed a norm on $L(X, Y)$ for finite-dimensional spaces. Again, when necessary to emphasize which norm we are talking about, we may write it as $\|A\|_{L(X,Y)}$.

For example, if $X = \mathbb{R}^1$ with norm $\|x\| = |x|$, elements of $L(X)$ are multiplication by scalars, $x \mapsto ax$, and we identify $a \in \mathbb{R}$ with the corresponding element of $L(X)$. If $\|x\| = |x| = 1$, then $|ax| = |a|$, so the operator norm of a is $|a|$.

By linearity, $\left\|A\frac{x}{\|x\|}\right\| = \frac{\|Ax\|}{\|x\|}$ for all nonzero $x \in X$. The vector $\frac{x}{\|x\|}$ is of norm 1. Therefore,

$$\|A\| = \sup\big\{\|Ax\| : x \in X \text{ with } \|x\| = 1\big\} = \sup_{\substack{x \in X\\ x \neq 0}} \frac{\|Ax\|}{\|x\|}.$$

This implies, assuming $\|A\| \neq \infty$ to avoid a technicality when $x = 0$, that for every $x \in X$,

$$\|Ax\| \leq \|A\|\|x\|.$$

Conversely, if one shows $\|Ax\| \leq C\|x\|$ for all x, then $\|A\| \leq C$.

It is not hard to see from the definition that $\|A\| = 0$ if and only if $A = 0$, where $A = 0$ means that A takes every vector to the zero vector. It is also not difficult to compute the operator norm of the identity operator:

$$\|I\| = \sup_{\substack{x \in X \\ x \neq 0}} \frac{\|Ix\|}{\|x\|} = \sup_{\substack{x \in X \\ x \neq 0}} \frac{\|x\|}{\|x\|} = 1.$$

The operator norm is not always so easy to compute using the definition alone, nor is it easy to read off the form of the operator. Consider $\mathbb{R}^2$ and the operator $A \in L(\mathbb{R}^2)$ that takes (x, y) to $(x + y, 2x)$. Unit norm vectors can be written as $\bigl(\pm t, \pm\sqrt{1-t^2}\bigr)$ for $t \in [0,1]$ (or perhaps $\bigl(\cos(\theta), \sin(\theta)\bigr)$). One then maximizes

$$\|A(x, y)\| = \sqrt{\left(t \pm \sqrt{1-t^2}\right)^2 + 4t^2}$$

to find $\|A\| = \sqrt{3 + \sqrt{5}}$. More generally, one often does two steps. For instance, consider the operator $B \in L\bigl(C([0,1], \mathbb{R}), \mathbb{R}\bigr)$ taking a continuous f to $f(0)$. If $\|f\| = 1$ (the uniform norm), then clearly $|f(0)| \leq 1$, so $|Bf| \leq 1$, meaning $\|B\| \leq 1$. To prove it is equal to 1, note that the constant function 1 has norm 1, so $B1 = 1$, meaning $\|B\| \geq 1$. So $\|B\| = 1$.

The operator norm is not always a norm on $L(X, Y)$, in particular, $\|A\|$ is not always finite for $A \in L(X, Y)$. We prove below that $\|A\|$ is finite when X is finite-dimensional. The operator norm being finite is equivalent to A being continuous. For infinite-dimensional spaces, neither statement needs to be true. For an example, consider the vector space of continuously differentiable functions on $[0, 2\pi]$ using the uniform norm. The functions $t \mapsto \sin(nt)$ have norm 1, but their derivatives have norm n. So differentiation, which is a linear operator valued in the space of continuous functions, has infinite operator norm on this space. We will stick to finite-dimensional spaces.

Given a finite-dimensional vector space X, we often think of $\mathbb{R}^n$, although if we have a norm on X, the norm might not be the standard euclidean norm. In the exercises, you can prove that every norm on $\mathbb{R}^n$ is "equivalent" to the euclidean norm in that the topology it generates is the same. For simplicity, we only prove the following proposition for euclidean spaces, and the proof for general finite-dimensional spaces is left as an exercise.

Proposition 8.2.4. *Let X and Y be normed vector spaces, $A \in L(X, Y)$, and X is finite-dimensional. Then $\|A\| < \infty$, and A is uniformly continuous (Lipschitz with constant $\|A\|$).*

Proof. As we said we only prove the proposition for euclidean spaces, so suppose that $X = \mathbb{R}^n$ and the norm is the standard euclidean norm. The general case is left as an exercise.

Let $\{e_1, e_2, \ldots, e_n\}$ be the standard basis of $\mathbb{R}^n$. Write $x \in \mathbb{R}^n$, with $\|x\| = 1$, as

$$x = \sum_{k=1}^{n} c_k \, e_k.$$

Since $e_k \cdot e_\ell = 0$ whenever $k \neq \ell$ and $e_k \cdot e_k = 1$, we have $c_k = x \cdot e_k$. By Cauchy–Schwarz,

$$|c_k| = |x \cdot e_k| \leq \|x\| \, \|e_k\| = 1.$$

Then

$$\|Ax\| = \left\| \sum_{k=1}^{n} c_k \, Ae_k \right\| \leq \sum_{k=1}^{n} |c_k| \, \|Ae_k\| \leq \sum_{k=1}^{n} \|Ae_k\|.$$

The right-hand side does not depend on x. We found a finite upper bound for $\|Ax\|$ independent of x, so $\|A\| < \infty$.

Take normed vector spaces X and Y, and $A \in L(X,Y)$ with $\|A\| < \infty$. For $v, w \in X$,

$$\|Av - Aw\| = \|A(v-w)\| \leq \|A\| \, \|v-w\|.$$

As $\|A\| < \infty$, then the inequality above says that A is Lipschitz with constant $\|A\|$. ◻

Proposition 8.2.5. *Let X, Y, and Z be finite-dimensional normed vector spaces*.*

(i) *If $A, B \in L(X,Y)$ and $c \in \mathbb{R}$, then*

$$\|A+B\| \leq \|A\| + \|B\|, \qquad \|cA\| = |c| \, \|A\|.$$

In particular, the operator norm is a norm on the vector space $L(X,Y)$.

(ii) *If $A \in L(X,Y)$ and $B \in L(Y,Z)$, then*

$$\|BA\| \leq \|B\| \, \|A\|.$$

Proof. First, since all the spaces are finite-dimensional, then all the operator norms are finite, and the statements make sense to begin with.

For (i), let $x \in X$ be arbitrary. Then

$$\|(A+B)x\| = \|Ax + Bx\| \leq \|Ax\| + \|Bx\| \leq \|A\| \, \|x\| + \|B\| \, \|x\| = \bigl(\|A\| + \|B\|\bigr)\|x\|.$$

So $\|A+B\| \leq \|A\| + \|B\|$. Similarly,

$$\|(cA)x\| = |c| \, \|Ax\| \leq \bigl(|c| \, \|A\|\bigr)\|x\|.$$

Thus $\|cA\| \leq |c| \, \|A\|$. Next,

$$|c| \, \|Ax\| = \|cAx\| \leq \|cA\| \, \|x\|.$$

Hence $|c| \, \|A\| \leq \|cA\|$.

For (ii), write

$$\|BAx\| \leq \|B\| \, \|Ax\| \leq \|B\| \, \|A\| \, \|x\|.$$ ◻

A norm defines a metric, giving a metric space topology on $L(X,Y)$ for finite-dimensional vector spaces. So, we can talk about open/closed sets, continuity, convergence, etc.

---

*If we strike the "In particular" part and interpret the algebra with infinite operator norms properly, namely decree that 0 times ∞ is 0, then this result also holds for infinite-dimensional spaces.

Proposition 8.2.6. *Let X be a finite-dimensional normed vector space. Let $GL(X) \subset L(X)$ be the set of invertible linear operators.**

(i) *If $A \in GL(X)$, $B \in L(X)$, and*
$$\|A - B\| < \frac{1}{\|A^{-1}\|}, \tag{8.2}$$
then $B \in GL(X)$, that is, B is invertible. In particular, $GL(X)$ is open.

(ii) *$A \mapsto A^{-1}$ is a continuous function on $GL(X)$.*

We illustrate this proposition on a simple example. Consider $X = \mathbb{R}^1$, where linear operators are just numbers a and the operator norm of a is $|a|$. The operator a is invertible ($a^{-1} = 1/a$) whenever $a \neq 0$. The condition $|a - b| < \frac{1}{|a^{-1}|}$ indeed implies that b is not zero. Moreover, $a \mapsto 1/a$ is a continuous function. When the dimension is 2 or higher, there are other noninvertible operators than just zero, and things are a bit more difficult.

Proof. Let us prove (i). We know something about A^{-1} and $A - B$; they are linear operators. So apply them to a vector:
$$A^{-1}(A - B)x = x - A^{-1}Bx.$$

Therefore,
$$\begin{aligned} \|x\| &= \|A^{-1}(A - B)x + A^{-1}Bx\| \\ &\leq \|A^{-1}\| \, \|A - B\| \, \|x\| + \|A^{-1}\| \, \|Bx\|. \end{aligned}$$

Assume $x \neq 0$ and so $\|x\| \neq 0$. Using (8.2), we obtain
$$\|x\| < \|x\| + \|A^{-1}\| \, \|Bx\|.$$

Thus $\|Bx\| \neq 0$ for all $x \neq 0$, and consequently $Bx \neq 0$ for all $x \neq 0$. So B is one-to-one; if $Bx = By$, then $B(x - y) = 0$, so $x = y$. As B is a one-to-one linear mapping from X to X, which is finite-dimensional, it is also onto by Proposition 8.1.18. Therefore, B is invertible. It follows that, in particular, $GL(X)$ is open.

Let us prove (ii). We must show that the inverse is continuous. Fix a $A \in GL(X)$. Let B be near A, specifically $\|A - B\| < \frac{1}{2\|A^{-1}\|}$. Then (8.2) is satisfied and B is invertible. A similar computation as above (using $B^{-1}y$ instead of x) gives
$$\|B^{-1}y\| \leq \|A^{-1}\| \, \|A - B\| \, \|B^{-1}y\| + \|A^{-1}\| \, \|y\| \leq \frac{1}{2}\|B^{-1}y\| + \|A^{-1}\| \, \|y\|,$$

or
$$\|B^{-1}y\| \leq 2\|A^{-1}\| \, \|y\|.$$

So $\|B^{-1}\| \leq 2\|A^{-1}\|$.

Now
$$A^{-1}(A - B)B^{-1} = A^{-1}(AB^{-1} - I) = B^{-1} - A^{-1},$$

*$GL(X)$ is called the *general linear group*, that is where the acronym GL comes from.

and

$$\|B^{-1} - A^{-1}\| = \|A^{-1}(A - B)B^{-1}\| \leq \|A^{-1}\| \, \|A - B\| \, \|B^{-1}\| \leq 2\|A^{-1}\|^2 \|A - B\|.$$

Therefore, as B tends to A, $\|B^{-1} - A^{-1}\|$ tends to 0, and so the inverse operation is a continuous function at A. □

8.2.2 Matrices

Once we fix a basis in a finite-dimensional vector space X, we can represent a vector of X as an n-tuple of numbers—a vector in $\mathbb{R}^n$. Same can be done with $L(X, Y)$, bringing us to matrices, which are a convenient way to represent finite-dimensional linear transformations. Suppose $\{x_1, x_2, \ldots, x_n\}$ and $\{y_1, y_2, \ldots, y_m\}$ are bases for vector spaces X and Y respectively. A linear operator is determined by its values on the basis. Given $A \in L(X, Y)$, Ax_j is an element of Y. Define the numbers $a_{i,j}$ via

$$Ax_j = \sum_{i=1}^{m} a_{i,j} \, y_i, \tag{8.3}$$

and write them as a *matrix*, which we, by slight abuse of notation, also call A,

$$A = \begin{bmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{bmatrix}.$$

We sometimes write A as $[a_{i,j}]$. We say A is an m-by-n matrix. The jth *column* of the matrix contains precisely the coefficients that represent Ax_j in terms of the basis $\{y_1, y_2, \ldots, y_m\}$. Given the numbers $a_{i,j}$, then via the formula (8.3), we find the corresponding linear operator, as it is determined by the action on a basis. Hence, once we fix bases on X and Y, we have a one-to-one correspondence between $L(X, Y)$ and the m-by-n matrices. When

$$z = \sum_{j=1}^{n} z_j \, x_j,$$

then

$$Az = \sum_{j=1}^{n} z_j \, Ax_j = \sum_{j=1}^{n} z_j \left(\sum_{i=1}^{m} a_{i,j} \, y_i \right) = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{i,j} \, z_j \right) y_i,$$

which gives rise to the familiar rule for matrix multiplication, thinking of z as a column vector, that is, an n-by-1 matrix. More generally, if B is an n-by-r matrix with entries $b_{j,k}$, then the matrix for $C = AB$ is an m-by-r matrix whose (i, k)th entry $c_{i,k}$ is

$$c_{i,k} = \sum_{j=1}^{n} a_{i,j} \, b_{j,k}.$$

A way to remember it is if you order the indices as we do—*row, column*—and put the elements in the same order as the matrices, then the "middle index" is "summed-out."

There is a one-to-one correspondence between matrices and linear operators in $L(X, Y)$, once we fix bases in X and Y. If we choose different bases, we get different matrices. This is an important distinction. The operator A acts on elements of X, while the matrix is something that works with n-tuples of numbers, that is, vectors of $\mathbb{R}^n$. By convention, we use standard bases in $\mathbb{R}^n$ unless otherwise specified, and we identify $L(\mathbb{R}^n, \mathbb{R}^m)$ with the set of m-by-n matrices.

A linear mapping changing one basis to another is represented by a square matrix in which the columns represent vectors of the second basis in terms of the first basis. We call such a linear mapping a *change of basis*. So for two choices of a basis in an n-dimensional vector space, there is a linear mapping (a change of basis) taking one basis to the other, and this corresponds to an n-by-n matrix which does the corresponding operation on $\mathbb{R}^n$.

Suppose $X = \mathbb{R}^n$, $Y = \mathbb{R}^m$, and all the bases are just the standard bases. Using the Cauchy–Schwarz inequality, with $c = (c_1, c_2, \ldots, c_n) \in \mathbb{R}^n$, compute

$$\|Ac\|^2 = \sum_{i=1}^{m} \left(\sum_{j=1}^{n} a_{i,j}\, c_j \right)^2 \leq \sum_{i=1}^{m} \left(\left(\sum_{j=1}^{n} (a_{i,j})^2 \right) \left(\sum_{j=1}^{n} (c_j)^2 \right) \right) = \left(\sum_{i=1}^{m} \sum_{j=1}^{n} (a_{i,j})^2 \right) \|c\|^2.$$

In other words, we have a bound on the operator norm (note that equality rarely happens)

$$\|A\| \leq \sqrt{\sum_{i=1}^{m} \sum_{j=1}^{n} (a_{i,j})^2}.$$

The right hand side is the euclidean norm on $\mathbb{R}^{nm}$, the space of all the entries of the matrix. If the entries go to zero, then $\|A\|$ goes to zero. Conversely,

$$\sum_{i=1}^{m} \sum_{j=1}^{n} (a_{i,j})^2 = \sum_{j=1}^{n} \|Ae_j\|^2 \leq \sum_{j=1}^{n} \|A\|^2 = n\|A\|^2.$$

So if the operator norm of A goes to zero, so do the entries. In particular, if A is fixed and B is changing, then the entries of B go to the entries of A if and only if B goes to A in operator norm ($\|A - B\|$ goes to zero). We have proved:

Proposition 8.2.7. *The topology (the set of open sets) on $L(\mathbb{R}^n, \mathbb{R}^m)$ is the same whether we consider $L(\mathbb{R}^n, \mathbb{R}^m)$ as a metric space using the operator norm, or the euclidean metric of $\mathbb{R}^{nm}$.*

In particular, let S be a metric space and let $\pi \colon L(\mathbb{R}^n, \mathbb{R}^m) \to \mathbb{R}^{nm}$ identify an operator with the nm-tuple of entries of the corresponding matrix. Then $f \colon S \to L(\mathbb{R}^n, \mathbb{R}^m)$ is continuous if and only if $\pi \circ f \colon S \to \mathbb{R}^{nm}$ is continuous. Similarly for $g \colon L(\mathbb{R}^n, \mathbb{R}^m) \to S$ and $g \circ \pi^{-1} \colon \mathbb{R}^{nm} \to S$.

8.2.3 Determinants

A certain number can be assigned to square matrices that measures how the corresponding linear mapping stretches space. In particular, this number, called the determinant, can be used to test for invertibility of a matrix.

Define the symbol $\operatorname{sgn}(x)$ (read "sign of x") for a number x by

$$\operatorname{sgn}(x) := \begin{cases} -1 & \text{if } x < 0, \\ 0 & \text{if } x = 0, \\ 1 & \text{if } x > 0. \end{cases}$$

A *permutation* $\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_n)$ is a reordering of $(1, 2, \ldots, n)$. Define

$$\operatorname{sgn}(\sigma) = \operatorname{sgn}(\sigma_1, \ldots, \sigma_n) := \prod_{p<q} \operatorname{sgn}(\sigma_q - \sigma_p). \tag{8.4}$$

Here $\prod$ stands for multiplication, similarly to how $\sum$ stands for summation.

Every permutation can be obtained by a sequence of transpositions (switchings of two elements). A permutation is *even* (resp. *odd*) if it takes an even (resp. odd) number of transpositions to get from $(1, 2, \ldots, n)$ to σ. For instance, $(2, 4, 3, 1)$ is two transpositions away from $(1, 2, 3, 4)$ and is therefore even: $(1, 2, 3, 4) \to (2, 1, 3, 4) \to (2, 4, 3, 1)$. Being even or odd is well-defined: $\operatorname{sgn}(\sigma)$ is 1 if σ is even and -1 if σ is odd (exercise). This fact follows since applying a transposition changes the sign and $\operatorname{sgn}(1, 2, \ldots, n) = 1$.

Let S_n be the set of all permutations on n elements (the *symmetric group*). Let $A = [a_{i,j}]$ be a square n-by-n matrix. Define the *determinant* of A as

$$\det(A) := \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \prod_{i=1}^{n} a_{i,\sigma_i}.$$

Proposition 8.2.8.

(*i*) $\det(I) = 1$.

(*ii*) *For every* $j = 1, 2, \ldots, n$, *the function* $x_j \mapsto \det\bigl([x_1 \quad x_2 \quad \cdots \quad x_n]\bigr)$ *is linear.*

(*iii*) *If two columns of a matrix are interchanged, then the determinant changes sign.*

(*iv*) *If two columns of* A *are equal, then* $\det(A) = 0$.

(*v*) *If a column is zero, then* $\det(A) = 0$.

(*vi*) $A \mapsto \det(A)$ *is a continuous function on* $L(\mathbb{R}^n)$.

(*vii*) $\det\left(\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]\right) = ad - bc$, *and* $\det\bigl([a]\bigr) = a$.

In fact, the determinant is the unique function that satisfies (i), (ii), and (iii), but we digress. By (ii), we mean that if we fix all the vectors $x_1, \ldots, x_n$ except for x_j, and let $v, w \in \mathbb{R}^n$ be two vectors, and $a, b \in \mathbb{R}$ be scalars, then

$$\begin{aligned} &\det\bigl([x_1 \quad \cdots \quad x_{j-1} \quad (av + bw) \quad x_{j+1} \quad \cdots \quad x_n]\bigr) = \\ &\quad a \det\bigl([x_1 \quad \cdots \quad x_{j-1} \quad v \quad x_{j+1} \quad \cdots \quad x_n]\bigr) + b \det\bigl([x_1 \quad \cdots \quad x_{j-1} \quad w \quad x_{j+1} \quad \cdots \quad x_n]\bigr). \end{aligned}$$

Proof. We go through the proof quickly, as you have likely seen it before. Item (i) is trivial. For (ii), note that each term in the definition of the determinant contains exactly one factor from each column. Item (iii) follows as switching two columns is switching the two corresponding numbers in every element in S_n. Hence, all the signs are changed. Item (iv) follows because if two columns are equal, and we switch them, we get the same matrix back. So item (iii) says the determinant must be 0. Item (v) follows because the product in each term in the definition includes one element from the zero column. Item (vi) follows as det is a polynomial in the entries of the matrix and hence continuous (as a function of the entries of the matrix). A function defined on matrices is continuous in the operator norm if and only if it is continuous as a function of the entries (Proposition 8.2.7). Finally, item (vii) is a direct computation. □

The determinant tells us about areas and volumes, and how they change. For example, in the 1-by-1 case, a matrix is just a number, and the determinant is exactly this number. It says how the linear mapping "stretches" the space. Similarly, suppose $A \in L(\mathbb{R}^2)$ is a linear transformation. It can be checked directly that the area of the image of the unit square $A\big([0,1]^2\big)$ is $|\det(A)|$, see Figure 8.3 for an example. This works with arbitrary figures, not just the unit square: The absolute value of the determinant tells us the stretch in the area. The sign of the determinant tells us if the image is flipped (changes orientation) or not. In $\mathbb{R}^3$ it tells us about the 3-dimensional volume, and in n dimensions about the n-dimensional volume. We claim this without proof.

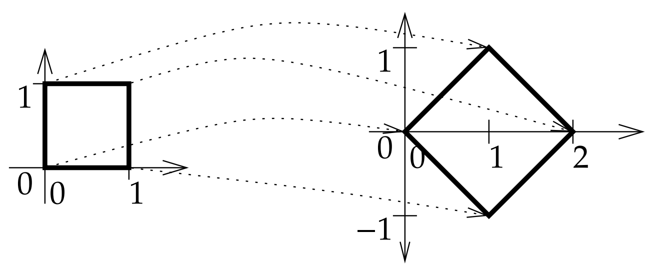

Figure 8.3: Image of the unit square $[0,1]^2$ via the matrix $\left[\begin{smallmatrix} 1 & 1 \\ -1 & 1 \end{smallmatrix}\right]$. The image is a square of side $\sqrt{2}$, thus of area 2, and the determinant of the matrix is 2.

Proposition 8.2.9. *If A and B are n-by-n matrices, then* $\det(AB) = \det(A)\det(B)$*. Furthermore, A is invertible if and only if* $\det(A) \neq 0$ *and in this case,* $\det(A^{-1}) = \frac{1}{\det(A)}$.

Proof. Let $b_1, b_2, \ldots, b_n$ be the columns of B. Then

$$AB = [Ab_1 \quad Ab_2 \quad \cdots \quad Ab_n].$$

That is, the columns of AB are $Ab_1, Ab_2, \ldots, Ab_n$.

Let $b_{j,k}$ denote the elements of B and a_j the columns of A. By linearity of the determinant,

$$\begin{aligned}
\det(AB) = \det\bigl([Ab_1 \quad Ab_2 \quad \cdots \quad Ab_n]\bigr) &= \det\left(\left[\sum_{j=1}^{n} b_{j,1}a_j \quad Ab_2 \quad \cdots \quad Ab_n\right]\right) \\
&= \sum_{j=1}^{n} b_{j,1} \det\bigl([a_j \quad Ab_2 \quad \cdots \quad Ab_n]\bigr) \\
&= \sum_{1 \leq j_1, j_2, \ldots, j_n \leq n} b_{j_1,1} b_{j_2,2} \cdots b_{j_n,n} \det\bigl([a_{j_1} \quad a_{j_2} \quad \cdots \quad a_{j_n}]\bigr) \\
&= \left(\sum_{(j_1, j_2, \ldots, j_n) \in S_n} b_{j_1,1} b_{j_2,2} \cdots b_{j_n,n} \operatorname{sgn}(j_1, j_2, \ldots, j_n) \right) \det\bigl([a_1 \quad a_2 \quad \cdots \quad a_n]\bigr).
\end{aligned}$$

In the last equality, we sum over the elements of S_n instead of all n-tuples for integers between 1 and n, because when two columns in the determinant are the same, then the determinant is zero. Reordering the columns to the original ordering to obtains the sgn.

The conclusion that $\det(AB) = \det(A)\det(B)$ follows by recognizing that the expression in parentheses above is the determinant of B. We obtain this by plugging in $A = I$. The expression we get for the determinant of B has rows and columns swapped, so as a bonus, we have also just proved that the determinant of a matrix and its transpose are equal.

Let us prove the "Furthermore." If A is invertible, then $A^{-1}A = I$. Consequently $\det(A^{-1})\det(A) = \det(A^{-1}A) = \det(I) = 1$. If A is not invertible, then it is not one-to-one, and so A takes some nonzero vector to zero. In other words, the columns of A are linearly dependent. Suppose

$$\sum_{k=1}^{n} \gamma_k \, a_k = 0,$$

where not all γ_k are equal to 0. Without loss of generality, suppose $\gamma_1 \neq 0$. Take

$$B := \begin{bmatrix} \gamma_1 & 0 & 0 & \cdots & 0 \\ \gamma_2 & 1 & 0 & \cdots & 0 \\ \gamma_3 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \gamma_n & 0 & 0 & \cdots & 1 \end{bmatrix}.$$

Using the definition of the determinant (there is only a single permutation σ for which $\prod_{i=1}^{n} b_{i,\sigma_i}$ is nonzero) we find $\det(B) = \gamma_1 \neq 0$. Then $\det(AB) = \det(A)\det(B) = \gamma_1 \det(A)$. The first column of AB is zero, and hence $\det(AB) = 0$. We conclude $\det(A) = 0$. □

Proposition 8.2.10. *Determinant is independent of the basis: If A and B are n-by-n matrices and B is invertible, then*

$$\det(A) = \det(B^{-1}AB).$$

Proof. $\det(B^{-1}AB) = \det(B^{-1})\det(A)\det(B) = \frac{1}{\det(B)}\det(A)\det(B) = \det(A)$. □

If in one basis A is the matrix representing a linear operator, then for another basis we can find a matrix B such that the matrix $B^{-1}AB$ takes us to the first basis, applies A in the first basis, and takes us back to the basis we started with. Let X be a finite-dimensional vector space. Let $\Phi \in L(X, \mathbb{R}^n)$ take a basis $\{x_1, \ldots, x_n\}$ to the standard basis $\{e_1, \ldots, e_n\}$ and let $\Psi \in L(X, \mathbb{R}^n)$ take another basis $\{y_1, \ldots, y_n\}$ to the standard basis. Let $T \in L(X)$ be a linear operator and let a matrix A represent the operator in the basis $\{x_1, \ldots, x_n\}$. Then B would be such that we have the following diagram*:

$$\begin{tikzcd}
\mathbb{R}^n \arrow[r, "B^{-1}AB"] \arrow[dd, bend right=60, "B"'] & \mathbb{R}^n \arrow[d, "\Psi^{-1}"'] \\
X \arrow[u, "\Psi"'] \arrow[r, "T"] \arrow[d, "\Phi"] & X \\
\mathbb{R}^n \arrow[r, "A"] & \mathbb{R}^n \arrow[u, "\Phi^{-1}"'] \arrow[uu, bend right=60, "B^{-1}"']
\end{tikzcd}$$

The two $\mathbb{R}^n$s on the bottom row represent X in the first basis, and the $\mathbb{R}^n$s on top represent X in the second basis.

If we compute the determinant of the matrix A, we obtain the same determinant if we use any other basis; in the other basis the matrix would be $B^{-1}AB$. Consequently,

$$\det : L(X) \to \mathbb{R}$$

is a well-defined function without the need to fix a basis. That is, det is defined on $L(X)$, not just on matrices.

There are three types of so-called *elementary matrices*. Let $e_1, e_2, \ldots, e_n$ be the standard basis on $\mathbb{R}^n$ as usual. First, for $j = 1, 2, \ldots, n$ and $\lambda \in \mathbb{R}$, $\lambda \neq 0$, define the first type of an elementary matrix, an n-by-n matrix E by

$$Ee_i := \begin{cases} e_i & \text{if } i \neq j, \\ \lambda e_i & \text{if } i = j. \end{cases}$$

Given any n-by-m matrix M the matrix EM is the same matrix as M except with the jth row multiplied by λ. It is an easy computation (exercise) that $\det(E) = \lambda$.

Next, for j, k with $j \neq k$ and $\lambda \in \mathbb{R}$, define the second type of an elementary matrix E by

$$Ee_i := \begin{cases} e_i & \text{if } i \neq j, \\ e_i + \lambda e_k & \text{if } i = j. \end{cases}$$

Given any n-by-m matrix M the matrix EM is the same matrix as M except with λ times the kth row added to the jth row. It is an easy computation (exercise) that $\det(E) = 1$.

*This is a so-called *commutative diagram*. Following arrows in any way should end up with the same result.

Finally, for j and k with $j \neq k$, define the third type of an elementary matrix E by

$$Ee_i := \begin{cases} e_i & \text{if } i \neq j \text{ and } i \neq k, \\ e_k & \text{if } i = j, \\ e_j & \text{if } i = k. \end{cases}$$

Given any n-by-m matrix M the matrix EM is the same matrix with jth and kth rows swapped. It is an easy computation (exercise) that $\det(E) = -1$.

Proposition 8.2.11. *Let T be an n-by-n invertible matrix. Then there exists a finite sequence of elementary matrices $E_1, E_2, \ldots, E_k$ such that*

$$T = E_1 E_2 \cdots E_k,$$

and

$$\det(T) = \det(E_1)\det(E_2)\cdots\det(E_k).$$

The proof is left as an exercise. The proposition says we can compute the determinant via elementary row operations. We do not have to factor the matrix into a product of elementary matrices completely. It is sufficient to do row operations until we find an *upper triangular matrix*, that is, a matrix $[a_{i,j}]$ where $a_{i,j} = 0$ if $i > j$. Computing determinant of such a matrix is not difficult (exercise).

Factorization into elementary matrices (or variations on elementary matrices) is useful in proofs involving an arbitrary linear operator, by reducing to a proof for an elementary matrix, similarly as the computation of the determinant.

8.2.4 Exercises

Exercise* 8.2.1*: *For a vector space X with a norm $\|\cdot\|$, show that $d(x,y) := \|x-y\|$ makes X a metric space.*

Exercise* 8.2.2** (Easy):*** *Show that for square matrices A and B, $\det(AB) = \det(BA)$.*

Exercise* 8.2.3*: *For $x \in \mathbb{R}^n$, define*

$$\|x\|_\infty := \max\{|x_1|, |x_2|, \ldots, |x_n|\},$$

sometimes called the sup or the max norm.

a) *Show that $\|\cdot\|_\infty$ is a norm on $\mathbb{R}^n$ (defining a different distance).*

b) *What is the unit ball $B(0,1)$ in this norm?*

Exercise* 8.2.4*: *For $x \in \mathbb{R}^n$, define*

$$\|x\|_1 := \sum_{k=1}^{n} |x_k|,$$

sometimes called the 1-norm (or L^1 norm).

a) *Show that $\|\cdot\|_1$ is a norm on $\mathbb{R}^n$ (defining a different distance, sometimes called the taxicab distance).*

b) *What is the unit ball $B(0,1)$ in this norm? Think about what it is in $\mathbb{R}^2$ and $\mathbb{R}^3$. Hint: It is, for example, a convex hull of a finite number of points.*

Exercise* 8.2.5*: *Using the euclidean norm on $\mathbb{R}^2$, compute the operator norm of the operators in $L(\mathbb{R}^2)$ given by the matrices:*

a) $\begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$ *b)* $\begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$ *c)* $\begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$ *d)* $\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}$

Exercise* 8.2.6*: *Using the standard euclidean norm $\mathbb{R}^n$, show:*

a) Suppose $A \in L(\mathbb{R}, \mathbb{R}^n)$ is defined for $x \in \mathbb{R}$ by $Ax := xa$ for a vector $a \in \mathbb{R}^n$. Then the operator norm $\|A\|_{L(\mathbb{R},\mathbb{R}^n)} = \|a\|_{\mathbb{R}^n}$. (That is, the operator norm of A is the euclidean norm of a.)

b) Suppose $B \in L(\mathbb{R}^n, \mathbb{R})$ is defined for $x \in \mathbb{R}^n$ by $Bx := b \cdot x$ for a vector $b \in \mathbb{R}^n$. Then the operator norm $\|B\|_{L(\mathbb{R}^n,\mathbb{R})} = \|b\|_{\mathbb{R}^n}$.

Exercise* 8.2.7*: *Suppose $\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_n)$ is a permutation of $(1, 2, \ldots, n)$.*

a) Show that we can make a finite number of transpositions (switching of two elements) to get to $(1, 2, \ldots, n)$.

b) Using the definition (8.4) show that σ is even if $\operatorname{sgn}(\sigma) = 1$ and σ is odd if $\operatorname{sgn}(\sigma) = -1$. In particular, showing that being odd or even is well-defined.

Exercise* 8.2.8*: *Verify the computation of the determinant for the three types of elementary matrices.*

Exercise* 8.2.9*: *Prove Proposition 8.2.11.*

Exercise* 8.2.10*:

a) Suppose $D = [d_{i,j}]$ is an n-by-n diagonal matrix*, that is, $d_{i,j} = 0$ whenever $i \neq j$. Show that* $\det(D) = d_{1,1}d_{2,2} \cdots d_{n,n}$.

b) Suppose A is a diagonalizable matrix. That is, there exists a matrix B such that $B^{-1}AB = D$ for a diagonal matrix $D = [d_{i,j}]$. Show that $\det(A) = d_{1,1}d_{2,2} \cdots d_{n,n}$.

Exercise* 8.2.11*: *Take the vector space of polynomials $\mathbb{R}[t]$ and let $D \in L(\mathbb{R}[t])$ be differentiation (we proved in an earlier exercise that D is a linear operator). Given $P(t) = c_0 + c_1 t + \cdots + c_n t^n \in \mathbb{R}[t]$ define* $\|P\| := \sup\{|c_j| : j = 0, 1, 2, \ldots, n\}$.

a) Show that $\|\cdot\|$ is a norm on $\mathbb{R}[t]$.

b) Prove $\|D\| = \infty$. Hint: Consider the polynomials t^n as n tends to infinity.

Exercise* 8.2.12*: *We finish the proof of Proposition 8.2.4. Let X be a finite-dimensional normed vector space with basis $\{x_1, x_2, \ldots, x_n\}$. Denote by $\|\cdot\|_X$ the norm on X, by $\|\cdot\|_{\mathbb{R}^n}$ the standard euclidean norm on $\mathbb{R}^n$, and by $\|\cdot\|_{L(X,Y)}$ the operator norm.*

a) Define $f : \mathbb{R}^n \to \mathbb{R}$,

$$f(c_1, c_2, \ldots, c_n) := \|c_1 x_1 + c_2 x_2 + \cdots + c_n x_n\|_X.$$

Show f is continuous.

b) Show that there exist numbers m and M such that if $c = (c_1, c_2, \ldots, c_n) \in \mathbb{R}^n$ with $\|c\|_{\mathbb{R}^n} = 1$, then $m \leq \|c_1 x_1 + c_2 x_2 + \cdots + c_n x_n\|_X \leq M$.

c) Show that there exists a number B such that if $\|c_1 x_1 + c_2 x_2 + \cdots + c_n x_n\|_X = 1$, then $|c_j| \leq B$.

d) Use part c) to show that if X is a finite-dimensional vector space and $A \in L(X, Y)$, then $\|A\|_{L(X,Y)} < \infty$.

***Exercise* 8.2.13:** *Let X be a finite-dimensional vector space with basis $\{x_1, x_2, \ldots, x_n\}$.*

a) Let $\|\cdot\|_X$ be a norm on X, $c = (c_1, c_2, \ldots, c_n) \in \mathbb{R}^n$, and $\|\cdot\|_{\mathbb{R}^n}$ the standard euclidean norm on $\mathbb{R}^n$. Prove that there exist numbers $m, M > 0$ such that for all $c \in \mathbb{R}^n$,

$$m\|c\|_{\mathbb{R}^n} \leq \|c_1x_1 + c_2x_2 + \cdots + c_nx_n\|_X \leq M\|c\|_{\mathbb{R}^n}.$$

Hint: See the previous exercise.

b) Use part a) to show that if $\|\cdot\|_1$ and $\|\cdot\|_2$ are two norms on X, then there exist numbers $m, M > 0$ (perhaps different from above) such that for all $x \in X$,

$$m\|x\|_1 \leq \|x\|_2 \leq M\|x\|_1.$$

c) Show that $U \subset X$ is open in the metric defined by $\|x - y\|_1$ if and only if U is open in the metric defined by $\|x - y\|_2$. So convergence of sequences and continuity of functions is the same in either norm.

***Exercise* 8.2.14:** *Let A be an upper triangular matrix. Find a formula for the determinant of A in terms of the diagonal entries, and prove that your formula works.*

***Exercise* 8.2.15:** *Given an n-by-n matrix A, prove that $|\det(A)| \leq \|A\|^n$ (the norm on A is the operator norm). Hint: One way to do it is to first prove it in the case $\|A\| = 1$, which means that all columns are of norm 1 or less, then prove that this means that $|\det(A)| \leq 1$ using linearity.*

***Exercise* 8.2.16:** *Consider Proposition 8.2.6 where $X = \mathbb{R}^n$ (for all n) using the euclidean norm.*

a) Prove that the estimate $\|A - B\| < \frac{1}{\|A^{-1}\|}$ is the best possible: For every $A \in GL(\mathbb{R}^n)$, find a B where equality is satisfied and B is not invertible. Hint: Difficulty is that $\|A\|\|A^{-1}\|$ is not always 1. Prove that a vector x_1 can be completed to a basis $\{x_1, \ldots, x_n\}$ such that $x_1 \cdot x_j = 0$ for $j \geq 2$. For the right x_1, make it so that $(A - B)x_j = 0$ for $j \geq 2$.

b) For every fixed $A \in GL(\mathbb{R}^n)$, let $\mathcal{M}$ denote the set of matrices B such that $\|A - B\| < \frac{1}{\|A^{-1}\|}$. Prove that while every $B \in \mathcal{M}$ is invertible, $\|B^{-1}\|$ is unbounded as a function of B on $\mathcal{M}$.

Let A be an n-by-n matrix. A $\lambda \in \mathbb{C}$ (possibly complex even for a real matrix) is an *eigenvalue* of A if there is a nonzero (possibly complex) vector $x \in \mathbb{C}^n$ such that $Ax = \lambda x$ (the multiplication by complex vectors is the same as for real vectors; if $x = a + ib$ for real vectors a and b, and A is a real matrix, then $Ax = Aa + iAb$). The number

$$\rho(A) := \sup\{|\lambda| : \lambda \text{ is an eigenvalue of } A\}$$

is the *spectral radius* of A. Here $|\lambda|$ is the complex modulus. We state without proof that at least one eigenvalue always exists, and there are no more than n distinct eigenvalues of A. You can therefore assume that $0 \leq \rho(A) < \infty$. The exercises below hold for complex matrices, but feel free to assume they are real matrices.

***Exercise* 8.2.17:** *Let A, S be n-by-n matrices, where S is invertible. Prove that λ is an eigenvalue of A, if and only if it is an eigenvalue of $S^{-1}AS$. Then prove that $\rho(S^{-1}AS) = \rho(S)$. In particular, ρ is a well-defined function on $L(X)$ for every finite-dimensional vector space X.*

***Exercise* 8.2.18:** *Let A be an n-by-n matrix A.*

a) Prove $\rho(A) \leq \|A\|$. (See above for definition of ρ.)

b) For every $k \in \mathbb{N}$, prove $\rho(A) \leq \|A^k\|^{1/k}$.

c) Suppose $\lim\limits_{k\to\infty} A^k = 0$ (limit in the operator norm). Prove that $\rho(A) < 1$.

***Exercise* 8.2.19:** *We say a set $C \subset \mathbb{R}^n$ is* symmetric *if $x \in C$ implies $-x \in C$.*

a) Let $\|\cdot\|$ be any given norm on $\mathbb{R}^n$. Show that the closed unit ball $C(0,1)$ (using the metric induced by this norm) is a compact symmetric convex set.

b) (Challenging) Let $C \subset \mathbb{R}^n$ be a compact, but note symmetric convex set and $0 \in C$. Show that

$$\|x\| := \inf \left\{ \lambda : \lambda > 0 \text{ and } \frac{x}{\lambda} \in C \right\}$$

is a norm on $\mathbb{R}^n$, and $C = C(0,1)$ (the closed unit ball) in the metric induced by this norm.

Hint: Feel free to the result of Exercise 8.2.13 part c). In particular, whether a set is "compact" is independent of the norm.

8.3 The derivative

Note: 2–3 lectures

8.3.1 The derivative

For a function $f : \mathbb{R} \to \mathbb{R}$, we defined the derivative at x as

$$\lim_{h\to 0} \frac{f(x+h) - f(x)}{h}.$$

In other words, there is a number a (the derivative of f at x) such that

$$\lim_{h\to 0} \left| \frac{f(x+h) - f(x)}{h} - a \right| = \lim_{h\to 0} \left| \frac{f(x+h) - f(x) - ah}{h} \right| = \lim_{h\to 0} \frac{\lvert f(x+h) - f(x) - ah \rvert}{\lvert h \rvert} = 0.$$

Multiplying by a is a linear map in one dimension: $h \mapsto ah$. Namely, we think of $a \in L(\mathbb{R}^1, \mathbb{R}^1)$, which is the best linear approximation of how f changes near x. We use this interpretation to extend differentiation to more variables.

Definition 8.3.1. Let $U \subset \mathbb{R}^n$ be open and $f : U \to \mathbb{R}^m$ a function. We say f is *differentiable* at $x \in U$ if there exists an $A \in L(\mathbb{R}^n, \mathbb{R}^m)$ such that

$$\lim_{\substack{h\to 0\\ h\in\mathbb{R}^n}} \frac{\lVert f(x+h) - f(x) - Ah \rVert}{\lVert h \rVert} = 0.$$

We will show momentarily that A, if it exists, is unique. We write $Df(x) := A$, or $f'(x) := A$, and we say A is the *derivative* of f at x. When f is differentiable at every $x \in U$, we say simply that f is *differentiable*. See Figure 8.4 for an illustration.

For a differentiable function, the derivative of f is a function from U to $L(\mathbb{R}^n, \mathbb{R}^m)$. Compare to the one-dimensional case, where the derivative is a function from U to $\mathbb{R}$, but we really want to think of $\mathbb{R}$ here as $L(\mathbb{R}^1, \mathbb{R}^1)$. As in one dimension, the idea is that a differentiable mapping is "infinitesimally close" to a linear mapping, and this linear mapping is the derivative.

Notice the norms in the definition. The norm in the numerator is on $\mathbb{R}^m$, and the norm in the denominator is on $\mathbb{R}^n$ where h lives. Normally it is understood that $h \in \mathbb{R}^n$ from context (the formula makes no sense otherwise). We will not explicitly say so from now on. Let us prove, as promised, that the derivative is unique.

Proposition 8.3.2. *Let $U \subset \mathbb{R}^n$ be an open subset and $f : U \to \mathbb{R}^m$ a function. Suppose $x \in U$ and there exist $A, B \in L(\mathbb{R}^n, \mathbb{R}^m)$ such that*

$$\lim_{h\to 0} \frac{\lVert f(x+h) - f(x) - Ah \rVert}{\lVert h \rVert} = 0 \qquad \textit{and} \qquad \lim_{h\to 0} \frac{\lVert f(x+h) - f(x) - Bh \rVert}{\lVert h \rVert} = 0.$$

Then $A = B$.

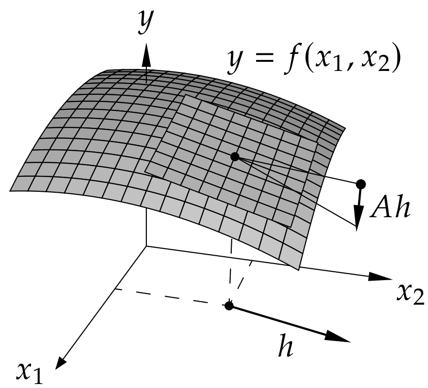

Figure 8.4: Illustration of a derivative for a function $f\colon \mathbb{R}^2 \to \mathbb{R}$. The vector h is shown in the x_1x_2-plane based at (x_1, x_2), and the vector $Ah \in \mathbb{R}^1$ is shown along the y direction.

Proof. Suppose $h \in \mathbb{R}^n$, $h \neq 0$. Compute

$$\begin{aligned}\frac{\lVert (A-B)h \rVert}{\lVert h \rVert} &= \frac{\lVert -\bigl(f(x+h)-f(x)-Ah\bigr) + f(x+h)-f(x)-Bh \rVert}{\lVert h \rVert} \\ &\leq \frac{\lVert f(x+h)-f(x)-Ah \rVert}{\lVert h \rVert} + \frac{\lVert f(x+h)-f(x)-Bh \rVert}{\lVert h \rVert}.\end{aligned}$$

So $\frac{\lVert (A-B)h \rVert}{\lVert h \rVert} \to 0$ as $h \to 0$. Given $\epsilon > 0$, for all nonzero h in some δ-ball around the origin we have

$$\epsilon > \frac{\lVert (A-B)h \rVert}{\lVert h \rVert} = \left\lVert (A-B)\frac{h}{\lVert h \rVert} \right\rVert.$$

For any given $v \in \mathbb{R}^n$ with $\lVert v \rVert = 1$, if $h = (\delta/2)\,v$, then $\lVert h \rVert < \delta$ and $\frac{h}{\lVert h \rVert} = v$. So $\lVert (A-B)v \rVert < \epsilon$. Taking the supremum over all v with $\lVert v \rVert = 1$, we get the operator norm $\lVert A-B \rVert \leq \epsilon$. As $\epsilon > 0$ was arbitrary, $\lVert A-B \rVert = 0$, or in other words $A = B$. □

Example 8.3.3: If $f(x) = Ax$ for a linear mapping A, then $f'(x) = A$:

$$\frac{\lVert f(x+h)-f(x)-Ah \rVert}{\lVert h \rVert} = \frac{\lVert A(x+h)-Ax-Ah \rVert}{\lVert h \rVert} = \frac{0}{\lVert h \rVert} = 0.$$

Example 8.3.4: Let $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ be defined by

$$f(x,y) = \bigl(f_1(x,y), f_2(x,y)\bigr) := (1 + x + 2y + x^2, 2x + 3y + xy).$$

Let us show that f is differentiable at the origin and compute the derivative directly using the definition. If the derivative exists, it is in $L(\mathbb{R}^2, \mathbb{R}^2)$, so it can be represented by a 2-by-2 matrix $\left[\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right]$. Suppose $h = (h_1, h_2)$. We need the following expression to go to zero.

$$\frac{\left\lVert f(h_1,h_2) - f(0,0) - (ah_1 + bh_2, ch_1 + dh_2)\right\rVert}{\lVert (h_1,h_2)\rVert} = \\ \frac{\sqrt{\bigl((1-a)h_1 + (2-b)h_2 + h_1^2\bigr)^2 + \bigl((2-c)h_1 + (3-d)h_2 + h_1h_2\bigr)^2}}{\sqrt{h_1^2+h_2^2}}.$$

If we choose $a = 1$, $b = 2$, $c = 2$, $d = 3$, the expression becomes

$$\frac{\sqrt{h_1^4 + h_1^2h_2^2}}{\sqrt{h_1^2+h_2^2}} = \lvert h_1\rvert \frac{\sqrt{h_1^2+h_2^2}}{\sqrt{h_1^2+h_2^2}} = \lvert h_1\rvert .$$

This expression does indeed go to zero as $h \to 0$. The function f is differentiable at the origin and the derivative $f'(0)$ is represented by the matrix $\left[\begin{smallmatrix} 1 & 2 \\ 2 & 3 \end{smallmatrix}\right]$.

Proposition 8.3.5. *Let $U \subset \mathbb{R}^n$ be open and $f \colon U \to \mathbb{R}^m$ be differentiable at $p \in U$. Then f is continuous at p.*

Proof. Another way to write the differentiability of f at p is to consider

$$r(h) \coloneqq f(p+h) - f(p) - f'(p)h.$$

The function f is differentiable at p if $\frac{\lVert r(h)\rVert}{\lVert h\rVert}$ goes to zero as $h \to 0$, so $r(h)$ itself goes to zero. The mapping $h \mapsto f'(p)h$ is a linear mapping between finite-dimensional spaces, hence continuous and $f'(p)h \to 0$ as $h \to 0$. Thus, $f(p+h)$ must go to $f(p)$ as $h \to 0$. That is, f is continuous at p. □

Differentiation is a linear operator on the space of differentiable functions.

Proposition 8.3.6. *Suppose $U \subset \mathbb{R}^n$ is open, $f \colon U \to \mathbb{R}^m$ and $g \colon U \to \mathbb{R}^m$ are differentiable at $p \in U$, and $\alpha \in \mathbb{R}$. Then the functions $f + g$ and αf are differentiable at p,*

$$(f+g)'(p) = f'(p) + g'(p), \qquad \text{and} \qquad (\alpha f)'(p) = \alpha f'(p).$$

Proof. Let $h \in \mathbb{R}^n$, $h \neq 0$. Then

$$\frac{\bigl\lVert f(p+h) + g(p+h) - \bigl(f(p) + g(p)\bigr) - \bigl(f'(p) + g'(p)\bigr)h\bigr\rVert}{\lVert h\rVert} \\ \leq \frac{\bigl\lVert f(p+h) - f(p) - f'(p)h\bigr\rVert}{\lVert h\rVert} + \frac{\bigl\lVert g(p+h) - g(p) - g'(p)h\bigr\rVert}{\lVert h\rVert},$$

and

$$\frac{\bigl\lVert \alpha f(p+h) - \alpha f(p) - \alpha f'(p)h\bigr\rVert}{\lVert h\rVert} = \lvert\alpha\rvert \frac{\bigl\lVert f(p+h)) - f(p) - f'(p)h\bigr\rVert}{\lVert h\rVert}.$$

The limits as h goes to zero of the right-hand sides are zero by hypothesis. The result follows. □

If $A \in L(\mathbb{R}^n, \mathbb{R}^m)$ and $B \in L(\mathbb{R}^m, \mathbb{R}^k)$ are linear maps, then they are their own derivative. The composition $BA \in L(\mathbb{R}^n, \mathbb{R}^k)$ is also its own derivative, and so the derivative of the composition is the composition of the derivatives. As differentiable maps are "infinitesimally close" to linear maps, they have the same property:

Theorem 8.3.7 (Chain rule)**.** *Let $U \subset \mathbb{R}^n$ and $V \subset \mathbb{R}^m$ be open sets, $f : U \to \mathbb{R}^m$ be differentiable at $p \in U$, $f(U) \subset V$, and let $g : V \to \mathbb{R}^\ell$ be differentiable at $f(p)$. Then $F : U \to \mathbb{R}^\ell$ defined by*

$$F(x) := g\bigl(f(x)\bigr)$$

is differentiable at p, and

$$F'(p) = g'\bigl(f(p)\bigr) f'(p).$$

Without the points where things are evaluated, we write $F' = (g \circ f)' = g' f'$. The derivative of the composition $g \circ f$ is the composition of the derivatives of g and f: If $f'(p) = A$ and $g'\bigl(f(p)\bigr) = B$, then $F'(p) = BA$, just as for linear maps.

Proof. Let $A := f'(p)$ and $B := g'\bigl(f(p)\bigr)$. Take a nonzero $h \in \mathbb{R}^n$ and write $q := f(p)$, $k := f(p+h) - f(p)$. Let

$$r(h) := f(p+h) - f(p) - Ah.$$

Then $r(h) = k - Ah$ or $Ah = k - r(h)$, and $f(p+h) = q + k$. We look at the quantity we need to go to zero:

$$\begin{aligned}
\frac{\lVert F(p+h) - F(p) - BAh \rVert}{\lVert h \rVert} &= \frac{\lVert g\bigl(f(p+h)\bigr) - g\bigl(f(p)\bigr) - BAh \rVert}{\lVert h \rVert} \\
&= \frac{\lVert g(q+k) - g(q) - B\bigl(k - r(h)\bigr) \rVert}{\lVert h \rVert} \\
&\leq \frac{\lVert g(q+k) - g(q) - Bk \rVert}{\lVert h \rVert} + \lVert B \rVert \frac{\lVert r(h) \rVert}{\lVert h \rVert} \\
&= \frac{\lVert g(q+k) - g(q) - Bk \rVert}{\lVert k \rVert} \frac{\lVert f(p+h) - f(p) \rVert}{\lVert h \rVert} + \lVert B \rVert \frac{\lVert r(h) \rVert}{\lVert h \rVert}.
\end{aligned}$$

First, $\lVert B \rVert$ is a constant and f is differentiable at p, so the term $\lVert B \rVert \frac{\lVert r(h) \rVert}{\lVert h \rVert}$ goes to 0. Next, because f is continuous at p, k goes to 0 as h goes to 0. Thus $\frac{\lVert g(q+k) - g(q) - Bk \rVert}{\lVert k \rVert}$ goes to 0, because g is differentiable at q. Finally,

$$\frac{\lVert f(p+h) - f(p) \rVert}{\lVert h \rVert} \leq \frac{\lVert f(p+h) - f(p) - Ah \rVert}{\lVert h \rVert} + \frac{\lVert Ah \rVert}{\lVert h \rVert} \leq \frac{\lVert f(p+h) - f(p) - Ah \rVert}{\lVert h \rVert} + \lVert A \rVert.$$

As f is differentiable at p, for small enough h, the quantity $\frac{\lVert f(p+h) - f(p) - Ah \rVert}{\lVert h \rVert}$ is bounded. Hence, the term $\frac{\lVert f(p+h) - f(p) \rVert}{\lVert h \rVert}$ stays bounded as h goes to 0. Therefore, $\frac{\lVert F(p+h) - F(p) - BAh \rVert}{\lVert h \rVert}$ goes to zero, and $F'(p) = BA$, which is what was claimed. $\square$

8.3.2 Partial derivatives

There is another way to generalize the derivative from one dimension. We hold all but one variable constant and take the regular one-variable derivative.

Definition 8.3.8. Let $f: U \to \mathbb{R}$ be a function on an open set $U \subset \mathbb{R}^n$. If the following limit exists, we write

$$\frac{\partial f}{\partial x_j}(x) := \lim_{h\to 0} \frac{f(x_1,\ldots,x_{j-1},x_j+h,x_{j+1},\ldots,x_n)-f(x)}{h} = \lim_{h\to 0} \frac{f(x+he_j)-f(x)}{h}.$$

We call $\frac{\partial f}{\partial x_j}(x)$ the *partial derivative* of f with respect to x_j. See Figure 8.5. Here h is a number, not a vector.

For a mapping $f: U \to \mathbb{R}^m$, we write $f = (f_1, f_2, \ldots, f_m)$, where f_k are real-valued functions. We then take partial derivatives of the components, $\frac{\partial f_k}{\partial x_j}$.

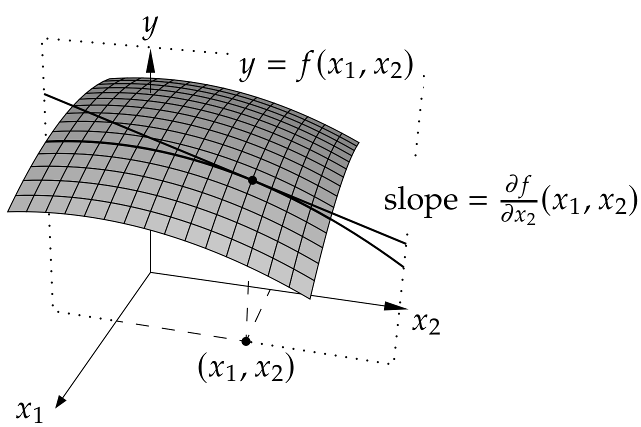

Figure 8.5: Illustration of a partial derivative for a function $f: \mathbb{R}^2 \to \mathbb{R}$. The yx_2-plane where x_1 is fixed is marked in dotted line, and the slope of the tangent line in the yx_2-plane is $\frac{\partial f}{\partial x_2}(x_1, x_2)$.

Partial derivatives are easier to compute with all the machinery of calculus, and they provide a way to compute the derivative of a function.

Proposition 8.3.9. *Let $U \subset \mathbb{R}^n$ be open and let $f: U \to \mathbb{R}^m$ be differentiable at $p \in U$. Then all the partial derivatives at p exist and, in terms of the standard bases of $\mathbb{R}^n$ and $\mathbb{R}^m$, $f'(p)$ is represented by the matrix*

$$\begin{bmatrix} \frac{\partial f_1}{\partial x_1}(p) & \frac{\partial f_1}{\partial x_2}(p) & \ldots & \frac{\partial f_1}{\partial x_n}(p) \\ \frac{\partial f_2}{\partial x_1}(p) & \frac{\partial f_2}{\partial x_2}(p) & \ldots & \frac{\partial f_2}{\partial x_n}(p) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(p) & \frac{\partial f_m}{\partial x_2}(p) & \ldots & \frac{\partial f_m}{\partial x_n}(p) \end{bmatrix}.$$

In other words,

$$f'(p)\, e_j = \sum_{k=1}^{m} \frac{\partial f_k}{\partial x_j}(p)\, e_k.$$

If $v = \sum_{j=1}^{n} c_j\, e_j = (c_1, c_2, \dots, c_n)$, then

$$f'(p)\, v = \sum_{j=1}^{n} \sum_{k=1}^{m} c_j \frac{\partial f_k}{\partial x_j}(p)\, e_k = \sum_{k=1}^{m} \left(\sum_{j=1}^{n} c_j \frac{\partial f_k}{\partial x_j}(p) \right) e_k.$$

Proof. Fix a j and note that for nonzero h,

$$\begin{aligned} \left\lVert \frac{f(p+he_j)-f(p)}{h} - f'(p)\, e_j \right\rVert &= \left\lVert \frac{f(p+he_j)-f(p)-f'(p)\, he_j}{h} \right\rVert \\ &= \frac{\lVert f(p+he_j)-f(p)-f'(p)\, he_j \rVert}{\lVert he_j \rVert}. \end{aligned}$$

As h goes to 0, the right-hand side goes to zero by differentiability of f. Hence,

$$\lim_{h\to 0} \frac{f(p+he_j)-f(p)}{h} = f'(p)\, e_j.$$

The limit is in $\mathbb{R}^m$. Represent f in components $f = (f_1, f_2, \dots, f_m)$. Taking a limit in $\mathbb{R}^m$ is the same as taking the limit in each component separately. So for every k, the partial derivative

$$\frac{\partial f_k}{\partial x_j}(p) = \lim_{h\to 0} \frac{f_k(p+he_j)-f_k(p)}{h}$$

exists and is equal to the kth component of $f'(p)\, e_j$, which is the jth column of $f'(p)$, and we are done. $\square$

The converse of the proposition is not true. Just because the partial derivatives exist, does not mean that the function is differentiable. See the exercises. However, when the partial derivatives are continuous, we will prove that the converse holds. One of the consequences of the proposition above is that if f is differentiable on U, then $f' \colon U \to L(\mathbb{R}^n, \mathbb{R}^m)$ is a continuous function if and only if all the $\frac{\partial f_k}{\partial x_j}$ are continuous functions.

8.3.3 Gradients, curves, and directional derivatives

Let $U \subset \mathbb{R}^n$ be open and $f \colon U \to \mathbb{R}$ a differentiable function. We define the *gradient* as

$$\nabla f(x) := \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(x)\, e_j.$$

The gradient gives a way to represent the action of the derivative as a dot product: $f'(x)\, v = \nabla f(x) \cdot v$.

Suppose $\gamma \colon (a,b) \subset \mathbb{R} \to \mathbb{R}^n$ is differentiable. Such a function and its image is sometimes called a *curve*, or a *differentiable curve*. Write $\gamma = (\gamma_1, \gamma_2, \ldots, \gamma_n)$. For the purposes of computation, we identify $L(\mathbb{R}^1)$ and $\mathbb{R}$ as we did when we defined the derivative in one variable. We also identify $L(\mathbb{R}^1, \mathbb{R}^n)$ with $\mathbb{R}^n$. We treat $\gamma'(t)$ both as an operator in $L(\mathbb{R}^1, \mathbb{R}^n)$ and the vector $\bigl(\gamma_1'(t), \gamma_2'(t), \ldots, \gamma_n'(t)\bigr)$ in $\mathbb{R}^n$. Using Proposition 8.3.9, if $v \in \mathbb{R}^n$ is $\gamma'(t)$ acting as a vector, then $h \mapsto h\,v$ (for $h \in \mathbb{R}^1 = \mathbb{R}$) is $\gamma'(t)$ acting as an operator in $L(\mathbb{R}^1, \mathbb{R}^n)$. We often use this slight abuse of notation when dealing with curves. The vector $\gamma'(t)$ is called a *tangent vector*. See Figure 8.6.

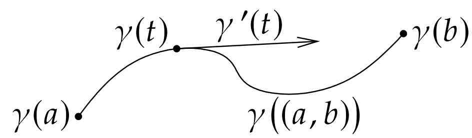

Figure 8.6: Differentiable curve and its derivative as a vector (for clarity assuming γ defined on $[a,b]$). The tangent vector $\gamma'(t)$ points along the curve.

Suppose $\gamma\bigl((a,b)\bigr) \subset U$ and let

$$g(t) := f\bigl(\gamma(t)\bigr).$$

The function g is differentiable. Treating $g'(t)$ as a number,

$$g'(t) = f'\bigl(\gamma(t)\bigr)\gamma'(t) = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}\bigl(\gamma(t)\bigr)\frac{d\gamma_j}{dt}(t) = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}\frac{d\gamma_j}{dt}.$$

For convenience, we often leave out the points where we are evaluating, such as above on the far right-hand side. With the notation of the gradient and the dot product the equation becomes

$$g'(t) = (\nabla f)\bigl(\gamma(t)\bigr) \cdot \gamma'(t) = \nabla f \cdot \gamma'.$$

We use this idea to define derivatives in a specific direction. A direction is simply a vector pointing in that direction. Pick a vector $u \in \mathbb{R}^n$ such that $\lVert u \rVert = 1$, and fix $x \in U$. We define the *directional derivative* as

$$D_u f(x) := \frac{d}{dt}\Big|_{t=0}\bigl[f(x+tu)\bigr] = \lim_{h \to 0} \frac{f(x+hu) - f(x)}{h},$$

where the notation $\frac{d}{dt}\big|_{t=0}$ represents the derivative evaluated at $t = 0$. When $u = e_j$ is a standard basis vector, we find $\frac{\partial f}{\partial x_j} = D_{e_j} f$. For this reason, sometimes the notation $\frac{\partial f}{\partial u}$ is used instead of $D_u f$.

Define γ by

$$\gamma(t) := x + tu.$$

Then $\gamma'(t) = u$ for all t. Let us see what happens to f when we travel along γ:

$$D_u f(x) = \frac{d}{dt}\Big|_{t=0} \big[f(x+tu)\big] = (\nabla f)\big(\gamma(0)\big) \cdot \gamma'(0) = (\nabla f)(x) \cdot u.$$

In fact, this computation holds whenever γ is any curve such that $\gamma(0) = x$ and $\gamma'(0) = u$.

Suppose $(\nabla f)(x) \neq 0$. By the Cauchy–Schwarz inequality,

$$\lvert D_u f(x) \rvert \leq \lVert (\nabla f)(x) \rVert .$$

Equality is achieved when u is a scalar multiple of $(\nabla f)(x)$. That is, when

$$u = \frac{(\nabla f)(x)}{\lVert (\nabla f)(x) \rVert},$$

we get $D_u f(x) = \lVert (\nabla f)(x) \rVert$. The gradient points in the direction in which the function grows fastest, in other words, in the direction in which $D_u f(x)$ is maximal.

8.3.4 The Jacobian

Definition 8.3.10. Let $U \subset \mathbb{R}^n$ and $f : U \to \mathbb{R}^n$ be a differentiable mapping. Define the *Jacobian determinant**, or simply the *Jacobian*†, of f at x as

$$J_f(x) := \det\big(f'(x)\big).$$

Sometimes J_f is written as

$$\frac{\partial(f_1, f_2, \ldots, f_n)}{\partial(x_1, x_2, \ldots, x_n)}.$$

This last piece of notation may seem somewhat confusing, but it is quite useful when we need to specify the exact variables and function components used, as we will do, for example, in the implicit function theorem.

The Jacobian determinant J_f is a real-valued function, and when $n = 1$ it is simply the derivative. From the chain rule and the fact that $\det(AB) = \det(A)\det(B)$, it follows that:

$$J_{f \circ g}(x) = J_f\big(g(x)\big) J_g(x).$$

The determinant of a linear mapping tells us what happens to area/volume under the mapping. Similarly, the Jacobian determinant measures how much a differentiable mapping stretches things locally, and if it flips orientation. In particular, if the Jacobian determinant is non-zero than we would assume that locally the mapping is invertible (and we would be correct as we will later see).

---

*Named after the Italian mathematician Carl Gustav Jacob Jacobi (1804–1851).

†The matrix from Proposition 8.3.9 representing $f'(x)$ is called the *Jacobian matrix*, or sometimes confusingly also called just “the Jacobian.”

8.3.5 Exercises

***Exercise* 8.3.1:** *Suppose $\gamma\colon (-1,1) \to \mathbb{R}^n$ and $\alpha\colon (-1,1) \to \mathbb{R}^n$ are two differentiable curves such that $\gamma(0) = \alpha(0)$ and $\gamma'(0) = \alpha'(0)$. Suppose $F\colon \mathbb{R}^n \to \mathbb{R}$ is a differentiable function. Show that*

$$\frac{d}{dt}\Big|_{t=0} F\big(\gamma(t)\big) = \frac{d}{dt}\Big|_{t=0} F\big(\alpha(t)\big).$$

***Exercise* 8.3.2:** *Let $f\colon \mathbb{R}^2 \to \mathbb{R}$ be given by $f(x,y) := \sqrt{x^2+y^2}$, see Figure 8.7. Show that f is not differentiable at the origin.*

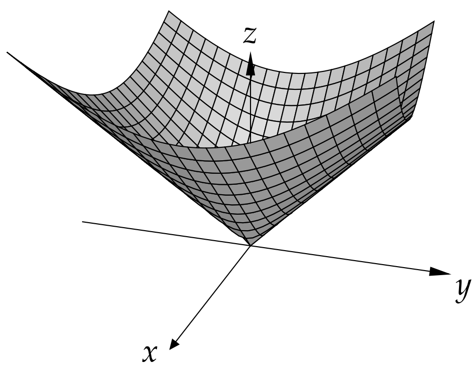

Figure 8.7: Graph of $\sqrt{x^2+y^2}$.

***Exercise* 8.3.3:** *Using only the definition of the derivative, show that the following $f\colon \mathbb{R}^2 \to \mathbb{R}^2$ are differentiable at the origin and find their derivative.*

a) $f(x,y) := (1+x+xy, x)$,

b) $f(x,y) := \big(y - y^{10}, x\big)$,

c) $f(x,y) := \big((x+y+1)^2, (x-y+2)^2\big)$.

***Exercise* 8.3.4:** *Suppose $f\colon \mathbb{R} \to \mathbb{R}$ and $g\colon \mathbb{R} \to \mathbb{R}$ are differentiable functions. Using only the definition of the derivative, show that $h\colon \mathbb{R}^2 \to \mathbb{R}^2$ defined by $h(x,y) := \big(f(x), g(y)\big)$ is a differentiable function, and find the derivative, at all points (x,y).*

***Exercise* 8.3.5:** *Define a function $f\colon \mathbb{R}^2 \to \mathbb{R}$ by (see Figure 8.8)*

$$f(x,y) := \begin{cases} \frac{xy}{x^2+y^2} & \text{if } (x,y) \neq (0,0), \\ 0 & \text{if } (x,y) = (0,0). \end{cases}$$

a) Show that the partial derivatives $\frac{\partial f}{\partial x}$ and $\frac{\partial f}{\partial y}$ exist at all points (including the origin).

b) Show that f is not continuous at the origin (and hence not differentiable).

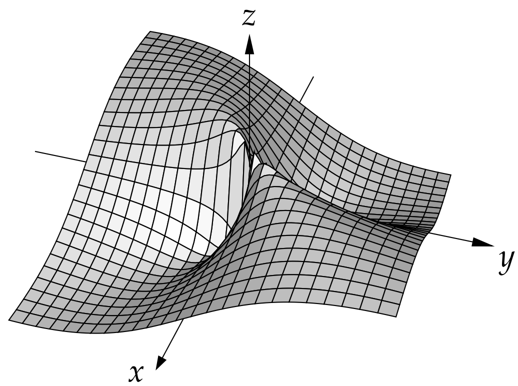

Figure 8.8: Graph of $\frac{xy}{x^2+y^2}$.

***Exercise* 8.3.6:** *Define a function* $f\colon \mathbb{R}^2 \to \mathbb{R}$ *by (see Figure 8.9)*

$$f(x,y) := \begin{cases} \frac{x^2y}{x^2+y^2} & \text{if } (x,y) \neq (0,0), \\ 0 & \text{if } (x,y) = (0,0). \end{cases}$$

a) *Show that the partial derivatives* $\frac{\partial f}{\partial x}$ *and* $\frac{\partial f}{\partial y}$ *exist at all points.*

b) *Show that for all* $u \in \mathbb{R}^2$ *with* $\|u\| = 1$, *the directional derivative* $D_u f$ *exists at all points.*

c) *Show that* f *is continuous at the origin.*

d) *Show that* f *is not differentiable at the origin.*

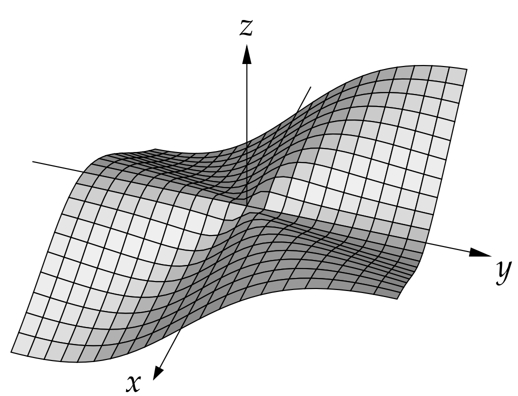

Figure 8.9: Graph of $\frac{x^2y}{x^2+y^2}$.

***Exercise* 8.3.7:** *Suppose $f: \mathbb{R}^n \to \mathbb{R}^n$ is one-to-one, onto, differentiable at all points, and such that f^{-1} is also differentiable at all points.*

a) *Show that $f'(p)$ is invertible at all points p and compute $(f^{-1})'\big(f(p)\big)$. Hint: Consider $x = f^{-1}\big(f(x)\big)$.*

b) *Let $g: \mathbb{R}^n \to \mathbb{R}^n$ be a function differentiable at $q \in \mathbb{R}^n$ and such that $g(q) = q$. Suppose $f(p) = q$ for some $p \in \mathbb{R}^n$. Show $J_g(q) = J_{f^{-1}\circ g \circ f}(p)$ where J_g is the Jacobian determinant.*

***Exercise* 8.3.8:** *Suppose $f: \mathbb{R}^2 \to \mathbb{R}$ is differentiable and such that $f(x,y) = 0$ if and only if $y = 0$ and such that $\nabla f(0,0) = (0,1)$. Prove that $f(x,y) > 0$ whenever $y > 0$, and $f(x,y) < 0$ whenever $y < 0$.*

As for functions of one variable, $f: U \to \mathbb{R}$ has a *relative maximum* at $p \in U$ if there exists a $\delta > 0$ such that $f(q) \leq f(p)$ for all $q \in B(p,\delta) \cap U$. Similarly for *relative minimum*.

***Exercise* 8.3.9:** *Suppose $U \subset \mathbb{R}^n$ is open and $f: U \to \mathbb{R}$ is differentiable. Suppose f has a relative maximum at $p \in U$. Show that $f'(p) = 0$, that is, the zero mapping in $L(\mathbb{R}^n, \mathbb{R})$. Namely, p is a* critical point *of f.*

***Exercise* 8.3.10:** *Suppose $f: \mathbb{R}^2 \to \mathbb{R}$ is differentiable and $f(x,y) = 0$ whenever $x^2 + y^2 = 1$. Prove that there exists at least one point (x_0, y_0) such that $\frac{\partial f}{\partial x}(x_0,y_0) = \frac{\partial f}{\partial y}(x_0,y_0) = 0$.*

***Exercise* 8.3.11:** *Define $f(x,y) := (x - y^2)(2y^2 - x)$. The graph of f is called the* Peano surface.*

a) *Show that $(0,0)$ is a critical point, that is, $f'(0,0) = 0$, the zero linear map in $L(\mathbb{R}^2, \mathbb{R})$.*

b) *Show that for every direction the restriction of f to a line through the origin in that direction has a relative maximum at the origin. In other words, for every (x,y) such that $x^2 + y^2 = 1$, the function $g(t) := f(tx,ty)$, has a relative maximum at $t = 0$.*
   Hint: While not necessary §4.3 of volume I makes this part easier.

c) *Show that f does not have a relative maximum at $(0,0)$.*

***Exercise* 8.3.12:** *Suppose $f: \mathbb{R} \to \mathbb{R}^n$ is differentiable and $\lVert f(t) \rVert = 1$ for all t (that is, we have a curve in the unit sphere). Show that $f'(t) \cdot f(t) = 0$ (treating $f'(t)$ as a vector) for all t.*

***Exercise* 8.3.13:** *Define $f: \mathbb{R}^2 \to \mathbb{R}^2$ by $f(x,y) := \big(x, y + \varphi(x)\big)$ for some differentiable function φ of one variable. Show f is differentiable and find f'.*

***Exercise* 8.3.14:** *Suppose $U \subset \mathbb{R}^n$ is open, $p \in U$, and $f: U \to \mathbb{R}$, $g: U \to \mathbb{R}$, $h: U \to \mathbb{R}$ are functions such that $f(p) = g(p) = h(p)$, f and h are differentiable at p, $f'(p) = h'(p)$, and*

$$f(x) \leq g(x) \leq h(x) \qquad \text{for all } x \in U$$

Show that g is differentiable at p and $g'(p) = f'(p) = h'(p)$.

***Exercise* 8.3.15:** *Prove a version of* mean value theorem *for functions of several variables. That is, suppose $U \subset \mathbb{R}^n$ is open, $f: U \to \mathbb{R}$ differentiable, $p, q \in U$, and the segment $[p,q] \in U$. Prove that there exists an $x \in [p,q]$ such that $\nabla f(x) \cdot (q - p) = f(q) - f(p)$.*

---

*Named after the Italian mathematician Giuseppe Peano (1858–1932).

8.4 Continuity and the derivative

Note: 1–2 lectures

8.4.1 Bounding the derivative

Let us prove a "mean value theorem" for vector-valued functions.

Lemma 8.4.1. *If $\varphi \colon [a,b] \to \mathbb{R}^n$ is differentiable on (a,b) and continuous on $[a,b]$, then there exists a $t_0 \in (a,b)$ such that*

$$\lVert \varphi(b) - \varphi(a) \rVert \leq (b-a) \lVert \varphi'(t_0) \rVert .$$

Proof. By the mean value theorem on the scalar-valued function $t \mapsto \bigl(\varphi(b)-\varphi(a)\bigr) \cdot \varphi(t)$, where the dot is the dot product, we obtain a $t_0 \in (a,b)$ such that

$$\begin{aligned} \lVert \varphi(b) - \varphi(a) \rVert^2 &= \bigl(\varphi(b)-\varphi(a)\bigr) \cdot \bigl(\varphi(b)-\varphi(a)\bigr) \\ &= \bigl(\varphi(b)-\varphi(a)\bigr) \cdot \varphi(b) - \bigl(\varphi(b)-\varphi(a)\bigr) \cdot \varphi(a) \\ &= (b-a)\bigl(\varphi(b)-\varphi(a)\bigr) \cdot \varphi'(t_0), \end{aligned}$$

where we treat φ' as a vector in $\mathbb{R}^n$ by the abuse of notation we mentioned in the previous section. If we think of $\varphi'(t)$ as a vector, then by Exercise 8.2.6, $\lVert \varphi'(t) \rVert_{L(\mathbb{R},\mathbb{R}^n)} = \lVert \varphi'(t) \rVert_{\mathbb{R}^n}$. That is, the euclidean norm of the vector is the same as the operator norm of $\varphi'(t)$.

By the Cauchy–Schwarz inequality

$$\lVert \varphi(b) - \varphi(a) \rVert^2 = (b-a)\bigl(\varphi(b)-\varphi(a)\bigr) \cdot \varphi'(t_0) \leq (b-a) \lVert \varphi(b)-\varphi(a) \rVert \, \lVert \varphi'(t_0) \rVert . \qquad \square$$

Recall that a set U is convex if whenever $p,q \in U$, the line segment from p to q lies in U.

Proposition 8.4.2. *Let $U \subset \mathbb{R}^n$ be a convex open set, $f \colon U \to \mathbb{R}^m$ be a differentiable function, and an M be such that*

$$\lVert f'(p) \rVert \leq M \qquad \text{for all } p \in U .$$

Then f is Lipschitz with constant M, that is,

$$\lVert f(p) - f(q) \rVert \leq M \lVert p-q \rVert \qquad \text{for all } p,q \in U .$$

Proof. Fix p and q in U and note that $(1-t)p+tq \in U$ for all $t \in [0,1]$ by convexity. Next

$$\frac{d}{dt} \Bigl[f\bigl((1-t)p+tq\bigr) \Bigr] = f'\bigl((1-t)p+tq\bigr)(q-p).$$

By Lemma 8.4.1, there is some $t_0 \in (0,1)$ such that

$$\begin{aligned} \lVert f(p) - f(q) \rVert &\leq \left\lVert \frac{d}{dt}\Big|_{t=t_0} \Bigl[f\bigl((1-t)p+tq\bigr) \Bigr] \right\rVert \\ &\leq \bigl\lVert f'\bigl((1-t_0)p+t_0q\bigr) \bigr\rVert \, \lVert q-p \rVert \leq M \lVert q-p \rVert . \end{aligned} \qquad \square$$

Example 8.4.3: If U is not convex the proposition is not true: Consider the set

$$U := \{(x, y) : 0.5 < x^2 + y^2 < 2\} \setminus \{(x, 0) : x < 0\}.$$

For $(x, y) \in U$, let $f(x, y)$ be the angle that the line from the origin to (x, y) makes with the positive x axis. We even have a formula for f:

$$f(x, y) = 2 \arctan\left(\frac{y}{x + \sqrt{x^2 + y^2}}\right).$$

Think a spiral staircase with room in the middle. See Figure 8.10.

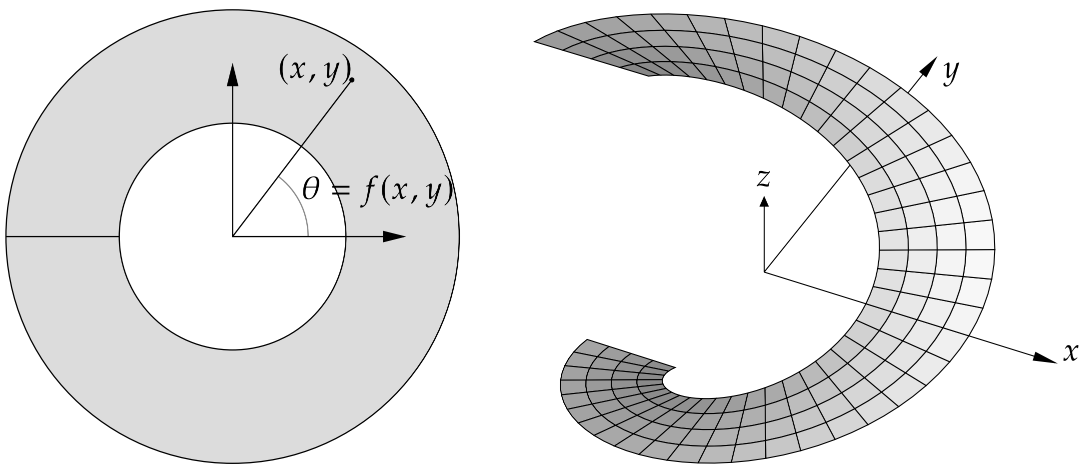

Figure 8.10: A non-Lipschitz function with uniformly bounded derivative.

The function is differentiable, and the derivative is bounded on U, which is not hard to see. Now think of what happens near where the negative x-axis cuts the annulus in half. As we approach this cut from positive y, $f(x, y)$ approaches π. From negative y, $f(x, y)$ approaches $-\pi$. So for small $\epsilon > 0$, $|f(-1, \epsilon) - f(-1, -\epsilon)|$ approaches 2π, but $\|(-1, \epsilon) - (-1, -\epsilon)\| = 2\epsilon$, which is arbitrarily small. The conclusion of the proposition does not hold for this nonconvex U.

Let us solve the differential equation $f' = 0$.

Corollary 8.4.4. *If $U \subset \mathbb{R}^n$ is open and connected, $f : U \to \mathbb{R}^m$ is differentiable, and $f'(x) = 0$ for all $x \in U$, then f is constant.*

Proof. For any given $x \in U$, there is a ball $B(x, \delta) \subset U$. The ball $B(x, \delta)$ is convex. Since $\|f'(y)\| \leq 0$ for all $y \in B(x, \delta)$, then by the proposition, $\|f(x) - f(y)\| \leq 0\|x - y\| = 0$. So $f(x) = f(y)$ for all $y \in B(x, \delta)$. Therefore, $f^{-1}(c)$ is open for all $c \in \mathbb{R}^m$.

Suppose $c_0 \in \mathbb{R}^m$ is such that $f^{-1}(c_0) \neq \emptyset$. As f is also continuous, the two sets

$$U' = f^{-1}(c_0), \qquad U'' = f^{-1}\big(\mathbb{R}^m \setminus \{c_0\}\big)$$

are open and disjoint, and further $U = U' \cup U''$. As U' is nonempty and U is connected, then $U'' = \emptyset$. So $f(x) = c_0$ for all $x \in U$. $\square$

8.4.2 Continuously differentiable functions

Definition 8.4.5. Let $U \subset \mathbb{R}^n$ be open. We say $f : U \to \mathbb{R}^m$ is *continuously differentiable*, or $C^1(U)$, if f is differentiable and $f' : U \to L(\mathbb{R}^n, \mathbb{R}^m)$ is continuous.

Proposition 8.4.6. *Let $U \subset \mathbb{R}^n$ be open and $f : U \to \mathbb{R}^m$. The function f is continuously differentiable if and only if the partial derivatives $\frac{\partial f_k}{\partial x_j}$ exist for all k and j and are continuous.*

Without continuity the theorem does not hold. Just because partial derivatives exist does not mean that f is differentiable, in fact, f may not even be continuous. See the exercises for the last section and also for this section.

Proof. We proved that if f is differentiable, then the partial derivatives exist. The partial derivatives are the entries of the matrix representing $f'(x)$. If $f' : U \to L(\mathbb{R}^n, \mathbb{R}^m)$ is continuous, then the entries are continuous, and hence the partial derivatives are continuous.

To prove the opposite direction, suppose the partial derivatives exist and are continuous. Fix $x \in U$. If we show that $f'(x)$ exists we are done, because the entries of the matrix representing $f'(x)$ are the partial derivatives and if the entries are continuous functions, the matrix-valued function f' is continuous.

We do induction on dimension. First, the conclusion is true when $n = 1$ (exercise, note that f is vector-valued). In this case, $f'(x)$ is essentially the derivative of chapter 4. Suppose the conclusion is true for $\mathbb{R}^{n-1}$. That is, if we restrict to the first $n-1$ variables, the function is differentiable. When taking the partial derivatives in x_1 through x_{n-1}, it does not matter if we consider f or f restricted to the set where x_n is fixed. In the following, by a slight abuse of notation, we think of $\mathbb{R}^{n-1}$ as a subset of $\mathbb{R}^n$, that is, the set in $\mathbb{R}^n$ where $x_n = 0$. In other words, we identify the vectors $(x_1, x_2, \ldots, x_{n-1})$ and $(x_1, x_2, \ldots, x_{n-1}, 0)$.

Fix $p \in U$ and let

$$A := \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(p) & \ldots & \frac{\partial f_1}{\partial x_n}(p) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(p) & \ldots & \frac{\partial f_m}{\partial x_n}(p) \end{bmatrix}, \qquad A' := \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(p) & \ldots & \frac{\partial f_1}{\partial x_{n-1}}(p) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(p) & \ldots & \frac{\partial f_m}{\partial x_{n-1}}(p) \end{bmatrix}, \qquad v := \begin{bmatrix} \frac{\partial f_1}{\partial x_n}(p) \\ \vdots \\ \frac{\partial f_m}{\partial x_n}(p) \end{bmatrix}.$$

Let $\epsilon > 0$ be given. By the induction hypothesis, there is a $\delta > 0$ such that for every $h' \in \mathbb{R}^{n-1}$ with $\|h'\| < \delta$, we have

$$\frac{\|f(p+h') - f(p) - A'h'\|}{\|h'\|} < \epsilon.$$

By continuity of the partial derivatives, suppose δ is small enough so that

$$\left| \frac{\partial f_k}{\partial x_n}(p+h) - \frac{\partial f_k}{\partial x_n}(p) \right| < \epsilon$$

for all k and all $h \in \mathbb{R}^n$ with $\|h\| < \delta$.

Suppose $h = h' + te_n$ is a vector in $\mathbb{R}^n$, where $h' \in \mathbb{R}^{n-1}$, $t \in \mathbb{R}$, such that $\|h\| < \delta$. Then $\|h'\| \leq \|h\| < \delta$. Note that $Ah = A'h' + tv$.

$$\begin{aligned}
\|f(p+h) - f(p) - Ah\| &= \|f(p+h'+te_n) - f(p+h') - tv + f(p+h') - f(p) - A'h'\| \\
&\leq \|f(p+h'+te_n) - f(p+h') - tv\| + \|f(p+h') - f(p) - A'h'\| \\
&\leq \|f(p+h'+te_n) - f(p+h') - tv\| + \epsilon\|h'\|.
\end{aligned}$$

As all the partial derivatives exist, by the mean value theorem, for each k there is some $\theta_k \in [0,t]$ (or $[t,0]$ if $t < 0$), such that

$$f_k(p+h'+te_n) - f_k(p+h') = t\frac{\partial f_k}{\partial x_n}(p+h'+\theta_k e_n).$$

We have $\|h' + \theta_k e_n\| \leq \|h\| < \delta$, and so we can finish the estimate

$$\begin{aligned}
\|f(p+h) - f(p) - Ah\| &\leq \|f(p+h'+te_n) - f(p+h') - tv\| + \epsilon\|h'\| \\
&\leq \sqrt{\sum_{k=1}^{m}\left(t\frac{\partial f_k}{\partial x_n}(p+h'+\theta_k e_n) - t\frac{\partial f_k}{\partial x_n}(p)\right)^2} + \epsilon\|h'\| \\
&\leq \sqrt{m}\,\epsilon|t| + \epsilon\|h'\| \\
&\leq (\sqrt{m}+1)\epsilon\|h\|.
\end{aligned}$$

□

A common application is to prove that a certain function is differentiable. For example, we can show that all polynomials are differentiable, and in fact continuously differentiable, by computing the partial derivatives.

Corollary 8.4.7. *A polynomial $p\colon \mathbb{R}^n \to \mathbb{R}$ in several variables*

$$p(x_1, x_2, \ldots, x_n) = \sum_{0 \leq j_1 + j_2 + \cdots + j_n \leq d} c_{j_1, j_2, \ldots, j_n} \, x_1^{j_1} x_2^{j_2} \cdots x_n^{j_n}$$

is continuously differentiable.

Proof. Consider the partial derivative of p in the x_n variable. Write p as

$$p(x) = \sum_{j=0}^{d} p_j(x_1, \ldots, x_{n-1}) \, x_n^j,$$

where p_j are polynomials in one less variable. Then

$$\frac{\partial p}{\partial x_n}(x) = \sum_{j=1}^{d} p_j(x_1, \ldots, x_{n-1}) \, j x_n^{j-1},$$

which is again a polynomial. So the partial derivatives of polynomials exist and are again polynomials. By the continuity of algebraic operations, polynomials are continuous functions. Therefore p is continuously differentiable. □

8.4.3 Exercises

Exercise* 8.4.1*: *Define* $f : \mathbb{R}^2 \to \mathbb{R}$ *as*

$$f(x,y) := \begin{cases} (x^2+y^2)\sin\bigl((x^2+y^2)^{-1}\bigr) & \text{if } (x,y) \neq (0,0), \\ 0 & \text{if } (x,y) = (0,0). \end{cases}$$

Show that f *is differentiable at the origin, but that it is not continuously differentiable.*
Note: Feel free to use what you know about sine and cosine from calculus.

Exercise* 8.4.2*: *Let* $f : \mathbb{R}^2 \to \mathbb{R}$ *be the function from Exercise 8.3.5, that is,*

$$f(x,y) := \begin{cases} \frac{xy}{x^2+y^2} & \text{if } (x,y) \neq (0,0), \\ 0 & \text{if } (x,y) = (0,0). \end{cases}$$

Compute the partial derivatives $\frac{\partial f}{\partial x}$ *and* $\frac{\partial f}{\partial y}$ *at all points and show that these are not continuous functions.*

Exercise* 8.4.3*: *Let* $B(0,1) \subset \mathbb{R}^2$ *be the unit ball, that is, the set given by* $x^2 + y^2 < 1$*. Suppose* $f : B(0,1) \to \mathbb{R}$ *is a differentiable function such that* $|f(0,0)| \leq 1$*, and* $\left|\frac{\partial f}{\partial x}\right| \leq 1$ *and* $\left|\frac{\partial f}{\partial y}\right| \leq 1$ *for all points in* $B(0,1)$*.*

a) Find an $M \in \mathbb{R}$ *such that* $\|f'(x,y)\| \leq M$ *for all* $(x,y) \in B(0,1)$*.*

b) Find a $B \in \mathbb{R}$ *such that* $|f(x,y)| \leq B$ *for all* $(x,y) \in B(0,1)$*.*

Exercise* 8.4.4*: *Define* $\varphi : [0,2\pi] \to \mathbb{R}^2$ *by* $\varphi(t) = \bigl(\sin(t), \cos(t)\bigr)$*. Compute* $\varphi'(t)$ *for all* t*. Compute* $\|\varphi'(t)\|$ *for all* t*. Notice that* $\varphi'(t)$ *is never zero, yet* $\varphi(0) = \varphi(2\pi)$*, therefore, Rolle's theorem is not true in more than one dimension.*

Exercise* 8.4.5*: *Let* $f : \mathbb{R}^2 \to \mathbb{R}$ *be a function such that* $\frac{\partial f}{\partial x}$ *and* $\frac{\partial f}{\partial y}$ *exist at all points and there exists an* $M \in \mathbb{R}$ *such that* $\left|\frac{\partial f}{\partial x}\right| \leq M$ *and* $\left|\frac{\partial f}{\partial y}\right| \leq M$ *at all points. Show that* f *is continuous.*

Exercise* 8.4.6*: *Let* $f : \mathbb{R}^2 \to \mathbb{R}$ *be a function and* $M \in \mathbb{R}$*, such that for every* $(x,y) \in \mathbb{R}^2$*, the function* $g(t) := f(xt, yt)$ *is differentiable and* $|g'(t)| \leq M$ *for all* t*.*

a) Show that f *is continuous at* $(0,0)$*.*

b) Find an example of such an f *that is discontinuous at every other point of* $\mathbb{R}^2$*.*
Hint: Think back to how we constructed a nowhere continuous function on $[0,1]$*.*

Exercise* 8.4.7*: *Suppose* $r : \mathbb{R}^n \setminus X \to \mathbb{R}$ *is a rational function, that is,* $p : \mathbb{R}^n \to \mathbb{R}$ *and* $q : \mathbb{R}^n \to \mathbb{R}$ *are polynomials,* q *is not identically zero,* $X = q^{-1}(0)$*, and* $r = \frac{p}{q}$*. Show that* r *is continuously differentiable.*

Exercise* 8.4.8*: *Suppose* $f : \mathbb{R}^n \to \mathbb{R}$ *and* $h : \mathbb{R}^n \to \mathbb{R}$ *are two differentiable functions such that* $f'(x) = h'(x)$ *for all* $x \in \mathbb{R}^n$*. Prove that if* $f(0) = h(0)$*, then* $f(x) = h(x)$ *for all* $x \in \mathbb{R}^n$*.*

Exercise* 8.4.9*: *Prove the base case in Proposition 8.4.6. That is, prove that if* $n = 1$ *and "the partials exist and are continuous," then the function is continuously differentiable. Note that* f *is vector-valued.*

Exercise* 8.4.10*: *Suppose that* $U \subset \mathbb{R}^n$ *is open,* $f : U \to \mathbb{R}^m$ *is differentiable, there is an* M *such that* $\|f'(p)\| \leq M$ *for all* $p \in U$*, and* $K \subset U$ *is a compact set. Prove that there exists an* M' *(where* $M' \geq M$*), such that for all* $p, q \in K$ *we have* $\|f(p) - f(q)\| \leq M'\|p - q\|$*. Compare to Proposition 8.4.2.*

8.5 Inverse and implicit function theorems

Note: 2–3 lectures

Intuitively, if a function is continuously differentiable, then it locally "behaves like" the derivative (which is a linear function). The idea of the inverse function theorem is that if a function is continuously differentiable and the derivative is invertible, the function is (locally) invertible.

Theorem 8.5.1 (Inverse function theorem)**.** *Let $U \subset \mathbb{R}^n$ be an open set and let $f : U \to \mathbb{R}^n$ be a continuously differentiable function. Suppose $p \in U$ and $f'(p)$ is invertible (that is, $J_f(p) \neq 0$). Then there exist open sets $V, W \subset \mathbb{R}^n$ such that $p \in V \subset U$, $f(V) = W$, and $f|_V$ is one-to-one. Hence a function $g : W \to V$ exists such that $g(y) := (f|_V)^{-1}(y)$. Furthermore, g is continuously differentiable and*

$$g'(y) = \bigl(f'(x)\bigr)^{-1}, \qquad \text{for all } x \in V, y = f(x).$$

See Figure 8.11.

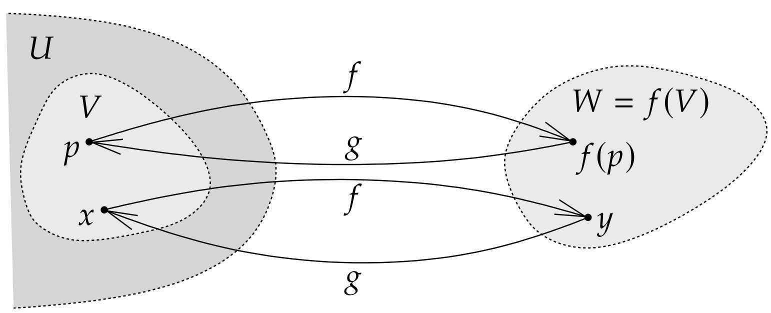

Figure 8.11: Setup of the inverse function theorem in $\mathbb{R}^n$.

To prove the theorem, we use the contraction mapping principle from chapter 7, where we used it to prove Picard's theorem. Recall that a mapping $f : X \to Y$ between metric spaces (X, d_X) and (Y, d_Y) is a contraction if there exists a $k < 1$ such that

$$d_Y\bigl(f(p), f(q)\bigr) \leq k\, d_X(p, q) \qquad \text{for all } p, q \in X.$$

The contraction mapping principle says that if $f : X \to X$ is a contraction and X is a complete metric space, then there exists a unique fixed point, that is, there exists a unique $x \in X$ such that $f(x) = x$.

Proof. Write $A = f'(p)$. As f' is continuous, there is an open ball V centered at p such that

$$\|A - f'(x)\| < \frac{1}{2\|A^{-1}\|} \qquad \text{for all } x \in V.$$

Consequently, the derivative $f'(x)$ is invertible for all $x \in V$ by Proposition 8.2.6.

Given $y \in \mathbb{R}^n$, define $\varphi_y : V \to \mathbb{R}^n$ by

$$\varphi_y(x) := x + A^{-1}\big(y - f(x)\big).$$

As A^{-1} is one-to-one, $\varphi_y(x) = x$ (x is a fixed point) if only if $y - f(x) = 0$, or in other words $f(x) = y$. Using the chain rule we obtain

$$\varphi_y'(x) = I - A^{-1} f'(x) = A^{-1}\big(A - f'(x)\big).$$

So for $x \in V$, we have

$$\|\varphi_y'(x)\| \leq \|A^{-1}\| \, \|A - f'(x)\| < 1/2.$$

As V is a ball, it is convex. Hence

$$\|\varphi_y(x_1) - \varphi_y(x_2)\| \leq \frac{1}{2}\|x_1 - x_2\| \qquad \text{for all } x_1, x_2 \in V.$$

In other words, φ_y is a contraction defined on V, though we so far do not know what is the range of φ_y. We cannot yet apply the fixed point theorem, but we can say that φ_y has at most one fixed point in V: If $\varphi_y(x_1) = x_1$ and $\varphi_y(x_2) = x_2$, then $\|x_1 - x_2\| = \|\varphi_y(x_1) - \varphi_y(x_2)\| \leq \frac{1}{2}\|x_1 - x_2\|$, so $x_1 = x_2$. That is, there exists at most one $x \in V$ such that $f(x) = y$, and so $f|_V$ is one-to-one.

Let $W := f(V)$ and let $g : W \to V$ be the inverse of $f|_V$. We need to show that W is open. Take a $y_0 \in W$. There is a unique $x_0 \in V$ such that $f(x_0) = y_0$. Let $r > 0$ be small enough such that the closed ball $C(x_0, r) \subset V$ (such $r > 0$ exists as V is open).

Suppose y is such that

$$\|y - y_0\| < \frac{r}{2\|A^{-1}\|}.$$

If we show that $y \in W$, then we have shown that W is open. If $x_1 \in C(x_0, r)$, then

$$\begin{aligned}
\|\varphi_y(x_1) - x_0\| &\leq \|\varphi_y(x_1) - \varphi_y(x_0)\| + \|\varphi_y(x_0) - x_0\| \\
&\leq \frac{1}{2}\|x_1 - x_0\| + \|A^{-1}(y - y_0)\| \\
&\leq \frac{1}{2}r + \|A^{-1}\| \, \|y - y_0\| \\
&< \frac{1}{2}r + \|A^{-1}\| \frac{r}{2\|A^{-1}\|} = r.
\end{aligned}$$

So φ_y takes $C(x_0, r)$ into $B(x_0, r) \subset C(x_0, r)$. It is a contraction on $C(x_0, r)$ and $C(x_0, r)$ is complete (closed subset of $\mathbb{R}^n$ is complete). Apply the contraction mapping principle to obtain a fixed point x, i.e. $\varphi_y(x) = x$. That is, $f(x) = y$, and $y \in f\big(C(x_0, r)\big) \subset f(V) = W$. Therefore, W is open.

Next we need to show that g is continuously differentiable and compute its derivative. First, let us show that it is differentiable. Let $y \in W$ and $k \in \mathbb{R}^n$, $k \neq 0$, such that $y + k \in W$. Because $f|_V$ is a one-to-one and onto mapping of V onto W, there are unique $x \in V$ and

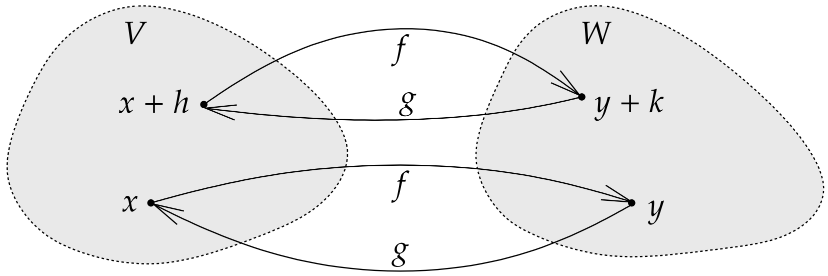

Figure 8.12: Proving that g is differentiable.

$h \in \mathbb{R}^n$, $h \neq 0$ and $x+h \in V$, such that $f(x) = y$ and $f(x+h) = y+k$. In other words, $g(y) = x$ and $g(y+k) = x+h$. See Figure 8.12.

We can still squeeze some information from the fact that φ_y is a contraction.

$$\varphi_y(x+h) - \varphi_y(x) = h + A^{-1}\bigl(f(x) - f(x+h)\bigr) = h - A^{-1}k.$$

So

$$\|h - A^{-1}k\| = \|\varphi_y(x+h) - \varphi_y(x)\| \leq \frac{1}{2}\|x+h-x\| = \frac{\|h\|}{2}.$$

By the inverse triangle inequality, $\|h\| - \|A^{-1}k\| \leq \frac{1}{2}\|h\|$. So

$$\|h\| \leq 2\|A^{-1}k\| \leq 2\|A^{-1}\|\,\|k\|.$$

In particular, as k goes to 0, so does h.

As $x \in V$, then $f'(x)$ is invertible. Let $B := \bigl(f'(x)\bigr)^{-1}$, which is what we think the derivative of g at y is. Then

$$\begin{aligned}
\frac{\|g(y+k) - g(y) - Bk\|}{\|k\|} &= \frac{\|h - Bk\|}{\|k\|} \\
&= \frac{\|h - B\bigl(f(x+h) - f(x)\bigr)\|}{\|k\|} \\
&= \frac{\|B\bigl(f(x+h) - f(x) - f'(x)h\bigr)\|}{\|k\|} \\
&\leq \|B\|\frac{\|h\|}{\|k\|}\,\frac{\|f(x+h) - f(x) - f'(x)h\|}{\|h\|} \\
&\leq 2\|B\|\,\|A^{-1}\|\frac{\|f(x+h) - f(x) - f'(x)h\|}{\|h\|}.
\end{aligned}$$

As k goes to 0, so does h. So the right-hand side goes to 0 as f is differentiable, and hence the left-hand side also goes to 0. And B is precisely what we wanted $g'(y)$ to be.

We have g is differentiable, let us show it is $C^1(W)$. The function $g\colon W \to V$ is continuous (it is differentiable), f' is a continuous function from V to $L(\mathbb{R}^n)$, and $X \mapsto X^{-1}$ is a continuous function on the set of invertible operators. As $g'(y) = \Bigl(f'\bigl(g(y)\bigr)\Bigr)^{-1}$ is the composition of these three continuous functions, it is continuous. □

Corollary 8.5.2. *Suppose $U \subset \mathbb{R}^n$ is open and $f : U \to \mathbb{R}^n$ is a continuously differentiable mapping such that $f'(x)$ is invertible for all $x \in U$. Then for every open set $V \subset U$, the set $f(V)$ is open (f is said to be an open mapping).*

Proof. Without loss of generality, suppose $U = V$. For each $y \in f(V)$, pick $x \in f^{-1}(y)$ (there could be more than one such point), then by the inverse function theorem there is a neighborhood of x in V that maps onto a neighborhood of y. Hence $f(V)$ is open. □

Example 8.5.3: The theorem, and the corollary, is not true if $f'(x)$ is not invertible for some x. For example, the map $f(x, y) := (x, xy)$, maps $\mathbb{R}^2$ onto the set $\mathbb{R}^2 \setminus \{(0, y) : y \neq 0\}$, which is neither open nor closed. In fact, $f^{-1}(0, 0) = \{(0, y) : y \in \mathbb{R}\}$. This bad behavior only occurs on the y-axis, everywhere else the function is locally invertible. If we avoid the y-axis, f is even one-to-one.

Example 8.5.4: Just because $f'(x)$ is invertible everywhere does not mean that f is one-to-one. It is "locally" one-to-one, but perhaps not "globally." Consider $f : \mathbb{R}^2 \setminus \{(0, 0)\} \to \mathbb{R}^2 \setminus \{(0, 0)\}$ defined by $f(x, y) := (x^2 - y^2, 2xy)$. It is left to the reader to verify the following statements. The map f is differentiable and the derivative is invertible. On the other hand, f is 2-to-1 globally: For every (a, b) that is not the origin, there are exactly two solutions to $x^2 - y^2 = a$ and $2xy = b$ (f is also onto). Notice that once you show that there is at least one solution, replacing x and y with $-x$ and $-y$ we obtain another solution.

The invertibility of the derivative is not a necessary condition, just sufficient, for having a continuous inverse and for being an open mapping. For example, the function $f(x) := x^3$ is an open mapping from $\mathbb{R}$ to $\mathbb{R}$ and is globally one-to-one with a continuous inverse, although the inverse is not differentiable at $x = 0$.

As a side note, there is a related famous, and as yet unsolved, problem called the *Jacobian conjecture*. If $F : \mathbb{R}^n \to \mathbb{R}^n$ is polynomial (each component is a polynomial) and J_F (the Jacobian determinant) is a nonzero constant, does F have a polynomial inverse? The inverse function theorem gives a local C^1 inverse, but can one always find a global polynomial inverse is the question.

8.5.1 Implicit function theorem

The inverse function theorem is a special case of the implicit function theorem, which we prove next. Although somewhat ironically we prove the implicit function theorem using the inverse function theorem. In the inverse function theorem we showed that the equation $x - f(y) = 0$ is solvable for y in terms of x if the derivative with respect to y is invertible, that is, if $f'(y)$ is invertible. Then there is (locally) a function g such that $x - f\big(g(x)\big) = 0$.

In general, the equation $f(x, y) = 0$ is not solvable for y in terms of x in every case. For instance, there is generally no solution when $f(x, y)$ does not actually depend on y. For a more interesting example, notice that $x^2 + y^2 - 1 = 0$ defines the unit circle, and we can locally solve for y in terms of x when 1) we are near a point on the unit circle and 2) we are not at a point where the circle has a vertical tangency, that is, where $\frac{\partial f}{\partial y} = 0$.

We fix some notation. Let $(x, y) \in \mathbb{R}^{n+m}$ denote the coordinates $(x_1, \ldots, x_n, y_1, \ldots, y_m)$. We can then write a linear map $A \in L(\mathbb{R}^{n+m}, \mathbb{R}^m)$ as $A = [A_x\ A_y]$ so that $A(x, y) = A_x x + A_y y$, where $A_x \in L(\mathbb{R}^n, \mathbb{R}^m)$ and $A_y \in L(\mathbb{R}^m)$. First, the linear version of the theorem.

Proposition 8.5.5. *Let $A = [A_x\ A_y] \in L(\mathbb{R}^{n+m}, \mathbb{R}^m)$ and suppose A_y is invertible. If $B = -(A_y)^{-1}A_x$, then*

$$0 = A(x, Bx) = A_x x + A_y Bx.$$

Furthermore, $y = Bx$ is the unique $y \in \mathbb{R}^m$ such that $A(x, y) = 0$.

The proof is immediate: We solve and obtain $y = Bx$. Another way to solve is to "complete the basis," that is, add rows to the matrix until we have an invertible matrix: The operator in $L(\mathbb{R}^{n+m})$ given by $(x, y) \mapsto (x, A_x x + A_y y)$ is invertible, and the map B can be read off from the inverse. Let us show that the same can be done for C^1 functions.

Theorem 8.5.6 (Implicit function theorem)**.** *Let $U \subset \mathbb{R}^{n+m}$ be an open set and let $f : U \to \mathbb{R}^m$ be a $C^1(U)$ mapping. Let $(p, q) \in U$ be a point such that $f(p, q) = 0$ and such that*

$$\frac{\partial(f_1, \ldots, f_m)}{\partial(y_1, \ldots, y_m)}(p, q) \neq 0.$$

Then there exists an open set $W \subset \mathbb{R}^n$ with $p \in W$, an open set $W' \subset \mathbb{R}^m$ with $q \in W'$, where $W \times W' \subset U$, and a $C^1(W)$ map $g : W \to W'$, with $g(p) = q$, and for all $x \in W$, the point $g(x)$ is the unique point in W' such that

$$f\bigl(x, g(x)\bigr) = 0.$$

Furthermore, if $A = [A_x\ A_y] = f'(p, q)$, then

$$g'(p) = -(A_y)^{-1}A_x.$$

The condition $\frac{\partial(f_1,\ldots,f_m)}{\partial(y_1,\ldots,y_m)}(p, q) = \det(A_y) \neq 0$ simply means that A_y is invertible. If $n = m = 1$, the condition is $\frac{\partial f}{\partial y}(p, q) \neq 0$, and W and W' are open intervals. See Figure 8.13.

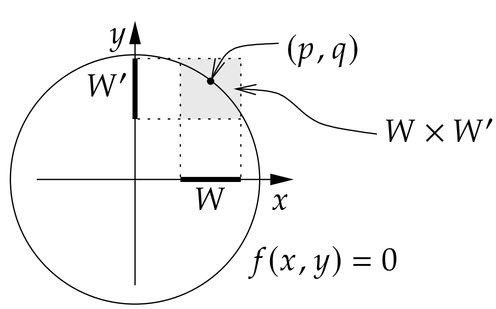

Figure 8.13: Implicit function theorem for $f(x, y) = x^2 + y^2 - 1$ in $U = \mathbb{R}^2$ and (p, q) in the first quadrant.

Proof. Define $F: U \to \mathbb{R}^{n+m}$ by $F(x,y) := \bigl(x, f(x,y)\bigr)$. It is clear that F is C^1, and we want to show that its derivative at (p,q) is invertible. Let us compute the derivative. The quotient

$$\frac{\|f(p+h,q+k) - f(p,q) - A_x h - A_y k\|}{\|(h,k)\|}$$

goes to zero as $\|(h,k)\| = \sqrt{\|h\|^2 + \|k\|^2}$ goes to zero. But then so does

$$\begin{aligned}
&\frac{\|F(p+h,q+k) - F(p,q) - (h, A_x h + A_y k)\|}{\|(h,k)\|} \\
&\qquad\qquad = \frac{\|\bigl(h, f(p+h,q+k) - f(p,q)\bigr) - (h, A_x h + A_y k)\|}{\|(h,k)\|} \\
&\qquad\qquad = \frac{\|f(p+h,q+k) - f(p,q) - A_x h - A_y k\|}{\|(h,k)\|}.
\end{aligned}$$

So the derivative of F at (p,q) takes (h,k) to $(h, A_x h + A_y k)$. In block matrix form, it is $\left[\begin{smallmatrix} I & 0 \\ A_x & A_y \end{smallmatrix}\right]$. If $(h, A_x h + A_y k) = (0,0)$, then $h = 0$, and so $A_y k = 0$. As A_y is one-to-one, $k = 0$. Thus $F'(p,q)$ is one-to-one, and hence invertible. We apply the inverse function theorem.

That is, there exists an open set $V \subset \mathbb{R}^{n+m}$ with $F(p,q) = (p,0) \in V$, and a C^1 mapping $G: V \to \mathbb{R}^{n+m}$, such that $F\bigl(G(x,s)\bigr) = (x,s)$ for all $(x,s) \in V$, G is one-to-one, and $G(V)$ is open. Write $G = (G_1, G_2)$ (the first n and the next m components of G). Then

$$F\bigl(G_1(x,s), G_2(x,s)\bigr) = \Bigl(G_1(x,s), f\bigl(G_1(x,s), G_2(x,s)\bigr)\Bigr) = (x,s).$$

So $x = G_1(x,s)$ and $f\bigl(G_1(x,s), G_2(x,s)\bigr) = f\bigl(x, G_2(x,s)\bigr) = s$. Plugging in $s = 0$, we obtain

$$f\bigl(x, G_2(x,0)\bigr) = 0.$$

As the set $G(V)$ is open and $(p,q) \in G(V)$, there exist some open sets $\widetilde{W}$ and W' such that $\widetilde{W} \times W' \subset G(V)$ with $p \in \widetilde{W}$ and $q \in W'$. Take $W := \bigl\{x \in \widetilde{W} : G_2(x,0) \in W'\bigr\}$. The function that takes x to $G_2(x,0)$ is continuous and therefore W is open. Define $g: W \to \mathbb{R}^m$ by $g(x) := G_2(x,0)$, which is the g in the theorem. The fact that $g(x)$ is the unique point in W' follows because $W \times W' \subset G(V)$ and G is one-to-one.

Next, differentiate

$$x \mapsto f\bigl(x, g(x)\bigr)$$

at p, which is the zero map, so its derivative is zero. Using the chain rule,

$$0 = A\bigl(h, g'(p)h\bigr) = A_x h + A_y g'(p) h$$

for all $h \in \mathbb{R}^n$, and we obtain the desired derivative for g. $\square$

In other words, in the context of the theorem, we have m equations in $n+m$ unknowns:

$$\begin{aligned} f_1(x_1,\ldots,x_n,y_1,\ldots,y_m) &= 0, \\ f_2(x_1,\ldots,x_n,y_1,\ldots,y_m) &= 0, \\ &\vdots \\ f_m(x_1,\ldots,x_n,y_1,\ldots,y_m) &= 0. \end{aligned}$$

The theorem guarantees a solution if $f=(f_1,f_2,\ldots,f_m)$ is a C^1 map (the components are C^1: partial derivatives in all variables exist and are continuous) and the matrix

$$\begin{bmatrix} \frac{\partial f_1}{\partial y_1} & \frac{\partial f_1}{\partial y_2} & \cdots & \frac{\partial f_1}{\partial y_m} \\ \frac{\partial f_2}{\partial y_1} & \frac{\partial f_2}{\partial y_2} & \cdots & \frac{\partial f_2}{\partial y_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial y_1} & \frac{\partial f_m}{\partial y_2} & \cdots & \frac{\partial f_m}{\partial y_m} \end{bmatrix}$$

is invertible at (p,q).

Example 8.5.7: Consider the set given by $x^2+y^2-(z+1)^3=-1$ and $e^x+e^y+e^z=3$ near the point $(0,0,0)$. It is the zero set of the mapping

$$f(x,y,z)=\bigl(x^2+y^2-(z+1)^3+1, e^x+e^y+e^z-3\bigr),$$

whose derivative is

$$f'=\begin{bmatrix} 2x & 2y & -3(z+1)^2 \\ e^x & e^y & e^z \end{bmatrix}.$$

The matrix

$$\begin{bmatrix} 2(0) & -3(0+1)^2 \\ e^0 & e^0 \end{bmatrix} = \begin{bmatrix} 0 & -3 \\ 1 & 1 \end{bmatrix}$$

is invertible. Hence near $(0,0,0)$, we can solve for y and z as C^1 functions of x such that for x near 0,

$$x^2+y(x)^2-\bigl(z(x)+1\bigr)^3=-1, \qquad e^x+e^{y(x)}+e^{z(x)}=3.$$

In other words, near the origin the set of solutions is a smooth curve in $\mathbb{R}^3$ that goes through the origin. The theorem does not tell us how to find $y(x)$ and $z(x)$ explicitly, it just tells us they exist.

An interesting, and sometimes useful, observation from the proof is that we solved the equation $f\bigl(x,g(x)\bigr)=s$ for all s in some neighborhood of 0, not just $s=0$.

Remark 8.5.8. There are versions of the theorem for arbitrarily many derivatives: If f has k continuous derivatives (see the next section), then the solution has k continuous derivatives as well.

8.5.2 Exercises

Exercise* 8.5.1*: *Let* $C := \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$.

a) *Solve for y in terms of x near $(0,1)$ (that is, find the function g from the implicit function theorem for a neighborhood of the point $(p,q) = (0,1)$).*

b) *Solve for y in terms of x near $(0,-1)$.*

c) *Solve for x in terms of y near $(-1,0)$.*

Exercise* 8.5.2*: *Define* $f : \mathbb{R}^2 \to \mathbb{R}^2$ *by* $f(x,y) := \bigl(x, y + h(x)\bigr)$ *for some continuously differentiable function h of one variable.*

a) *Show that f is one-to-one and onto.*

b) *Compute f'. (Make sure to argue why f' exists.)*

c) *Show that f' is invertible at all points, and compute its inverse.*

Exercise* 8.5.3*: *Define* $f : \mathbb{R}^2 \to \mathbb{R}^2 \setminus \{(0,0)\}$ *by* $f(x,y) := \bigl(e^x \cos(y), e^x \sin(y)\bigr)$.

a) *Show that f is onto.*

b) *Show that f' is invertible at all points.*

c) *Show that f is not one-to-one, in fact for every $(a,b) \in \mathbb{R}^2 \setminus \{(0,0)\}$, there exist infinitely many different points $(x,y) \in \mathbb{R}^2$ such that $f(x,y) = (a,b)$.*

Therefore, invertible derivative at every point does not mean that f is invertible globally.
Note: Feel free to use what you know about sine and cosine from calculus.

Exercise* 8.5.4*: *Find a map $f : \mathbb{R}^n \to \mathbb{R}^n$ that is one-to-one, onto, continuously differentiable, but $f'(0) = 0$.*
Hint: Generalize $f(x) = x^3$ from one to n dimensions.

Exercise* 8.5.5*: *Consider $z^2 + xz + y = 0$ in $\mathbb{R}^3$. Find an equation $D(x,y) = 0$, such that if $D(x_0,y_0) \neq 0$ and $z^2 + x_0 z + y_0 = 0$ for some $z \in \mathbb{R}$, then for points near (x_0,y_0) there exist exactly two distinct continuously differentiable functions $r_1(x,y)$ and $r_2(x,y)$ such that $z = r_1(x,y)$ and $z = r_2(x,y)$ solve $z^2 + xz + y = 0$. Do you recognize the expression D from algebra?*

Exercise* 8.5.6*: *Suppose $f : (a,b) \to \mathbb{R}^2$ is continuously differentiable and the first component (the x component) of $\nabla f(t)$ is not equal to 0 for all $t \in (a,b)$. Prove that there exists an open interval $I \subset \mathbb{R}$ and a continuously differentiable function $g : I \to \mathbb{R}$ such that $(x,y) \in f\bigl((a,b)\bigr)$ if and only if $x \in I$ and $y = g(x)$. In other words, the set $f\bigl((a,b)\bigr)$ is a graph of g.*

Exercise* 8.5.7*: *Define* $f : \mathbb{R}^2 \to \mathbb{R}^2$

$$f(x,y) := \begin{cases} \bigl(x^2 \sin(1/x) + x/2, y\bigr) & \text{if } x \neq 0, \\ (0,y) & \text{if } x = 0. \end{cases}$$

a) *Show that f is differentiable everywhere.*

b) *Show that $f'(0,0)$ is invertible.*

c) *Show that f is not one-to-one in every neighborhood of the origin (it is not locally invertible, that is, the inverse function theorem does not work).*

d) *Show that f is not continuously differentiable.*

Note: Feel free to use what you know about sine and cosine from calculus.

***Exercise* 8.5.8** (Polar coordinates)*: Define a mapping $F(r,\theta) := \bigl(r\cos(\theta), r\sin(\theta)\bigr)$.*

a) *Show that F is continuously differentiable (for all $(r,\theta) \in \mathbb{R}^2$).*
b) *Compute $F'(0,\theta)$ for all θ.*
c) *Show that if $r \neq 0$, then $F'(r,\theta)$ is invertible, therefore an inverse of F exists locally as long as $r \neq 0$.*
d) *Show that $F\colon \mathbb{R}^2 \to \mathbb{R}^2$ is onto, and for each point $(x,y) \in \mathbb{R}^2$, the set $F^{-1}(x,y)$ is infinite.*
e) *Show that $F\colon \mathbb{R}^2 \to \mathbb{R}^2$ is not an open mapping. Note that $F|_{(0,\infty)\times\mathbb{R}}$ is an open mapping via Corollary 8.5.2. Hint: Where does a small open rectangle such as $(-\epsilon,\epsilon) \times (-\epsilon,\epsilon)$ go?*
f) *Show that $F|_{(0,\infty)\times[0,2\pi)}$ is one-to-one and onto $\mathbb{R}^2 \setminus \bigl\{(0,0)\bigr\}$.*

Note: Feel free to use what you know about sine and cosine from calculus.

***Exercise* 8.5.9:** *Let $H := \bigl\{(x,y) \in \mathbb{R}^2 : y > 0\bigr\}$, and for $(x,y) \in H$ define*

$$F(x,y) := \left(\frac{x^2+y^2-1}{x^2+2y+y^2+1}, \frac{-2x}{x^2+2y+y^2+1} \right).$$

Prove that F is a bijective mapping from H to $B(0,1)$, it is continuously differentiable on H, and its inverse is also continuously differentiable.

***Exercise* 8.5.10:** *Suppose $U \subset \mathbb{R}^2$ is open and $f\colon U \to \mathbb{R}$ is a C^1 function such that $\nabla f(x,y) \neq 0$ for all $(x,y) \in U$. Show that every level set is a C^1 smooth curve. That is, for every $(x,y) \in U$, there exists a C^1 function $\gamma\colon (-\delta,\delta) \to \mathbb{R}^2$ with $\gamma'(0) \neq 0$ such that $f\bigl(\gamma(t)\bigr)$ is constant for all $t \in (-\delta,\delta)$.*

***Exercise* 8.5.11:** *Suppose $U \subset \mathbb{R}^2$ is open and $f\colon U \to \mathbb{R}$ is a C^1 function such that $\nabla f(x,y) \neq 0$ for all $(x,y) \in U$. Show that for every (x,y) there exists a neighborhood V of (x,y) an open set $W \subset \mathbb{R}^2$, a bijective C^1 function with a C^1 inverse $g\colon W \to V$ such that the level sets of $f \circ g$ are horizontal lines in W, that is, the set given by $(f \circ g)(s,t) = c$ for a constant c is a set of the form $\bigl\{(s,t_0) \in \mathbb{R}^2 : s \in \mathbb{R}, (s,t_0) \in W\bigr\}$, where t_0 is fixed. That is, the level curves can be locally "straightened."*

8.6 Higher order derivatives

Note: less than 1 lecture, optional, see also the optional §4.3 of volume I

Let $U \subset \mathbb{R}^n$ be an open set and $f : U \to \mathbb{R}$ a function. Denote our coordinates by $x = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$. Suppose $\frac{\partial f}{\partial x_\ell}$ exists everywhere in U, then it is also a function $\frac{\partial f}{\partial x_\ell} : U \to \mathbb{R}$. Therefore, it makes sense to talk about its partial derivatives. We denote the partial derivative of $\frac{\partial f}{\partial x_\ell}$ with respect to x_m by

$$\frac{\partial^2 f}{\partial x_m \partial x_\ell} := \frac{\partial \left(\frac{\partial f}{\partial x_\ell} \right)}{\partial x_m}.$$

If $m = \ell$, then we write $\frac{\partial^2 f}{\partial x_\ell^2}$ for simplicity.

We define higher order derivatives inductively. Suppose $\ell_1, \ell_2, \ldots, \ell_k$ are integers between 1 and n, and suppose

$$\frac{\partial^{k-1} f}{\partial x_{\ell_{k-1}} \partial x_{\ell_{k-2}} \cdots \partial x_{\ell_1}}$$

exists and is differentiable in the variable x_{ℓ_k}, then the partial derivative with respect to that variable is denoted by

$$\frac{\partial^k f}{\partial x_{\ell_k} \partial x_{\ell_{k-1}} \cdots \partial x_{\ell_1}} := \frac{\partial \left(\frac{\partial^{k-1} f}{\partial x_{\ell_{k-1}} \partial x_{\ell_{k-2}} \cdots \partial x_{\ell_1}} \right)}{\partial x_{\ell_k}}.$$

Such a derivative is called a *partial derivative of order k*.

Sometimes the notation $f_{x_\ell x_m}$ is used for $\frac{\partial^2 f}{\partial x_m \partial x_\ell}$. This notation swaps the order in which we write the derivatives, which may be important.

Definition 8.6.1. Suppose $U \subset \mathbb{R}^n$ is an open set and $f : U \to \mathbb{R}$ is a function. We say f is *k-times continuously differentiable function*, or a C^k function, if all partial derivatives of all orders up to and including order k exist and are continuous.

So a continuously differentiable, or C^1, function is one where all first order partial derivatives exist and are continuous, which agrees with our previous definition due to Proposition 8.4.6. We could have required only that the kth order partial derivatives exist and are continuous, as the existence of lower order partial derivatives is clearly necessary to even define kth order partial derivatives, and these lower order partial derivatives are continuous as they are (continuously) differentiable functions.

When the partial derivatives are continuous, we can swap their order.

Proposition 8.6.2. *Suppose $U \subset \mathbb{R}^n$ is open and $f : U \to \mathbb{R}$ is a C^2 function, and ℓ and m are two integers from 1 to n. Then*

$$\frac{\partial^2 f}{\partial x_m \partial x_\ell} = \frac{\partial^2 f}{\partial x_\ell \partial x_m}.$$

Proof. Fix a $p \in U$, and let e_ℓ and e_m be the standard basis vectors. Pick two positive numbers s and t small enough so that $p + s_0 e_\ell + t_0 e_m \in U$ whenever $0 < s_0 \leq s$ and $0 < t_0 \leq t$. Any small enough s and t work as U is open and so contains a small open ball (or a box if you wish) around p.

Use the mean value theorem on the function

$$\tau \mapsto f(p + se_\ell + \tau e_m) - f(x + \tau e_m),$$

on the interval $[0, t]$ to find a $t_0 \in (0, t)$ such that

$$\frac{f(p + se_\ell + te_m) - f(p + te_m) - f(p + se_\ell) + f(p)}{t} = \frac{\partial f}{\partial x_m}(p + se_\ell + t_0 e_m) - \frac{\partial f}{\partial x_m}(p + t_0 e_m).$$

Similarly, there exists a number $s_0 \in (0, s)$ such that

$$\frac{\frac{\partial f}{\partial x_m}(p + se_\ell + t_0 e_m) - \frac{\partial f}{\partial x_m}(p + t_0 e_m)}{s} = \frac{\partial^2 f}{\partial x_\ell \partial x_m}(p + s_0 e_\ell + t_0 e_m).$$

In other words,

$$g(s, t) := \frac{f(p + se_\ell + te_m) - f(p + te_m) - f(p + se_\ell) + f(p)}{st} = \frac{\partial^2 f}{\partial x_\ell \partial x_m}(p + s_0 e_\ell + t_0 e_m).$$

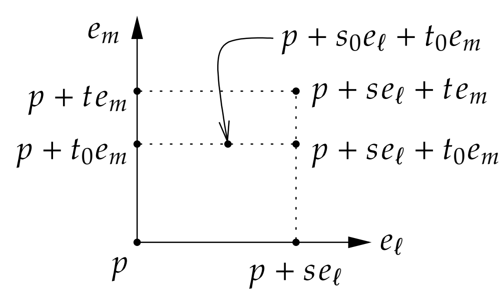

Figure 8.14: Using the mean value theorem to estimate a second order partial derivative by a certain difference quotient.

See Figure 8.14. The s_0 and t_0 depend on s and t, but $0 < s_0 < s$ and $0 < t_0 < t$. Let the domain of the function g be the set $(0, \epsilon) \times (0, \epsilon)$ for some small $\epsilon > 0$. As $(s, t) \in (0, \epsilon) \times (0, \epsilon)$ goes to $(0, 0)$, the point (s_0, t_0) also goes to $(0, 0)$. By continuity of the second partial derivatives,

$$\lim_{(s,t)\to(0,0)} g(s, t) = \frac{\partial^2 f}{\partial x_\ell \partial x_m}(p).$$

Now reverse the roles of s and t (and ℓ and m). Start with the function $\sigma \mapsto f(p + \sigma e_\ell + te_m) - f(p + \sigma e_\ell)$ find an $s_1 \in (0, s)$ such that

$$\frac{f(p + se_\ell + te_m) - f(p + se_\ell) - f(p + te_m) + f(p)}{s} = \frac{\partial f}{\partial x_\ell}(p + s_1 e_\ell + te_m) - \frac{\partial f}{\partial x_\ell}(p + s_1 e_\ell).$$

Find a $t_1 \in (0, t)$ such that

$$\frac{\frac{\partial f}{\partial x_\ell}(p + s_1 e_\ell + t e_m) - \frac{\partial f}{\partial x_\ell}(p + s_1 e_\ell)}{t} = \frac{\partial^2 f}{\partial x_m \partial x_\ell}(p + s_1 e_\ell + t_1 e_m).$$

So $g(s,t) = \frac{\partial^2 f}{\partial x_m \partial x_\ell}(p + s_1 e_\ell + t_1 e_m)$ for the same g as above. As before,

$$\lim_{(s,t)\to(0,0)} g(s,t) = \frac{\partial^2 f}{\partial x_m \partial x_\ell}(p).$$

Therefore, the two partial derivatives are equal. $\square$

The proposition does not hold if the derivatives are not continuous. See Exercise 8.6.2. Notice also that we did not really need a C^2 function, we only needed the two second order partial derivatives involved to be continuous functions.

8.6.1 Exercises

***Exercise* 8.6.1:** *Suppose $f: U \to \mathbb{R}$ is a C^2 function for some open $U \subset \mathbb{R}^n$ and $p \in U$. Use the proof of Proposition 8.6.2 to find an expression in terms of just the values of f (analogue of the difference quotient for the first derivative), whose limit is $\frac{\partial^2 f}{\partial x_\ell \partial x_m}(p)$.*

***Exercise* 8.6.2:** *Define*

$$f(x,y) := \begin{cases} \frac{xy(x^2-y^2)}{x^2+y^2} & \text{if } (x,y) \neq (0,0), \\ 0 & \text{if } (x,y) = (0,0). \end{cases}$$

Show that

a) The first order partial derivatives exist and are continuous.

b) The partial derivatives $\frac{\partial^2 f}{\partial x \partial y}$ and $\frac{\partial^2 f}{\partial y \partial x}$ exist, but are not continuous at $(0,0)$, and $\frac{\partial^2 f}{\partial x \partial y}(0,0) \neq \frac{\partial^2 f}{\partial y \partial x}(0,0)$.

***Exercise* 8.6.3:** *Let $f: U \to \mathbb{R}$ be a C^k function for some open $U \subset \mathbb{R}^n$ and $p \in U$. Suppose $\ell_1, \ell_2, \ldots, \ell_k$ are integers between 1 and n, and $\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_k)$ is a permutation of $(1, 2, \ldots, k)$. Prove*

$$\frac{\partial^k f}{\partial x_{\ell_k} \partial x_{\ell_{k-1}} \cdots \partial x_{\ell_1}}(p) = \frac{\partial^k f}{\partial x_{\ell_{\sigma_k}} \partial x_{\ell_{\sigma_{k-1}}} \cdots \partial x_{\ell_{\sigma_1}}}(p).$$

***Exercise* 8.6.4:** *Suppose $\varphi: \mathbb{R}^2 \to \mathbb{R}$ is a C^k function such that $\varphi(0,\theta) = \varphi(0,\psi)$ for all $\theta, \psi \in \mathbb{R}$ and $\varphi(r,\theta) = \varphi(r, \theta + 2\pi)$ for all $r, \theta \in \mathbb{R}$. Let $F(r,\theta) := \big(r\cos(\theta), r\sin(\theta)\big)$ from Exercise 8.5.8. Show that a function $g: \mathbb{R}^2 \to \mathbb{R}$, given $g(x,y) := \varphi\big(F^{-1}(x,y)\big)$ is well-defined (notice that $F^{-1}(x,y)$ can only be defined locally), and when restricted to $\mathbb{R}^2 \setminus \{0\}$ it is a C^k function.*
Note: Feel free to use what you know about sine and cosine from calculus.

***Exercise* 8.6.5:** *Suppose $f: \mathbb{R}^2 \to \mathbb{R}$ is a C^2 function. For all $(x,y) \in \mathbb{R}^2$, compute*

$$\lim_{t\to 0} \frac{f(x+t,y) + f(x-t,y) + f(x,y+t) + f(x,y-t) - 4f(x,y)}{t^2}$$

in terms of the partial derivatives of f.

***Exercise* 8.6.6:** *Suppose $f : \mathbb{R}^2 \to \mathbb{R}$ is a function such that all first and second order partial derivatives exist. Furthermore, suppose that all second order partial derivatives are bounded functions. Prove that f is continuously differentiable.*

***Exercise* 8.6.7:** *Follow the strategy below to prove the following simple version of the second derivative test for functions defined on $\mathbb{R}^2$ (using (x, y) as coordinates):* Suppose $f : \mathbb{R}^2 \to \mathbb{R}$ is a twice continuously differentiable function with a critical point at the origin, $f'(0,0) = 0$. If

$$\frac{\partial^2 f}{\partial x^2}(0,0) > 0 \qquad \text{and} \qquad \frac{\partial^2 f}{\partial x^2}(0,0)\frac{\partial^2 f}{\partial y^2}(0,0) - \left(\frac{\partial^2 f}{\partial x \partial y}(0,0)\right)^2 > 0,$$

then f has a (strict) local minimum at $(0,0)$. *Use the following technique: First suppose without loss of generality that $f(0,0) = 0$. Then prove:*

a) *There exists an $A \in L(\mathbb{R}^2)$ such that $g = f \circ A$ is such that $\frac{\partial^2 g}{\partial x \partial y}(0,0) = 0$, and $\frac{\partial^2 g}{\partial x^2}(0,0) = \frac{\partial^2 g}{\partial y^2}(0,0) = 1$.*

b) *For every $\epsilon > 0$, there exists a $\delta > 0$ such that $\left|g(x,y) - x^2 - y^2\right| < \epsilon(x^2 + y^2)$ for all $(x, y) \in B\big((0,0), \delta\big)$.*
Hint: You can use Taylor's theorem in one variable.

c) *This means that g, and therefore f, has a strict local minimum at $(0,0)$.*

Note: You must avoid the temptation to just apply the one variable second derivative test along lines through the origin, see Exercise 8.3.11.

Chapter 9

One-dimensional Integrals in Several Variables

9.1 Differentiation under the integral

Note: less than 1 lecture

Let $f(x, y)$ be a function of two variables, and define

$$g(y) := \int_a^b f(x, y)\, dx.$$

If f is continuous on the compact rectangle $[a, b] \times [c, d]$, then Proposition 7.5.12 from volume I says that g is continuous on $[c, d]$.

Suppose f is differentiable in y. When can we "differentiate under the integral"? That is, when is it true that g is differentiable and its derivative is

$$g'(y) \stackrel{?}{=} \int_a^b \frac{\partial f}{\partial y}(x, y)\, dx.$$

Differentiation is a limit and therefore we are really asking when do the two limiting operations of integration and differentiation commute. This is not always possible and some extra hypothesis is necessary. The first question we would face is the integrability of $\frac{\partial f}{\partial y}$, but the formula above can fail even if $\frac{\partial f}{\partial y}$ is integrable as a function of x for every fixed y.

We prove a simple, but perhaps the most useful version of this kind of result.

Theorem 9.1.1 (Leibniz integral rule)**.** *Suppose $f : [a, b] \times [c, d] \to \mathbb{R}$ is a continuous function, such that $\frac{\partial f}{\partial y}$ exists for all $(x, y) \in [a, b] \times [c, d]$ and is continuous. Define $g : [c, d] \to \mathbb{R}$ by*

$$g(y) := \int_a^b f(x, y)\, dx.$$

Then g is continuously differentiable and

$$g'(y) = \int_a^b \frac{\partial f}{\partial y}(x, y)\, dx.$$

The hypotheses on f and $\frac{\partial f}{\partial y}$ can be weakened, see e.g. Exercise 9.1.8, but not dropped outright. The main point in the proof requires that $\frac{\partial f}{\partial y}$ exists and is continuous for all x up to the endpoints, but we only need a small interval in the y direction. In applications, we often make $[c,d]$ a small interval around the point where we need to differentiate.

Proof. Fix $y \in [c,d]$ and let $\epsilon > 0$ be given. As $\frac{\partial f}{\partial y}$ is continuous on $[a,b] \times [c,d]$ it is uniformly continuous. In particular, there exists $\delta > 0$ such that whenever $y_1 \in [c,d]$ with $\lvert y_1 - y \rvert < \delta$ and all $x \in [a,b]$, we have

$$\left\lvert \frac{\partial f}{\partial y}(x,y_1) - \frac{\partial f}{\partial y}(x,y) \right\rvert < \epsilon.$$

Suppose h is such that $y+h \in [c,d]$ and $\lvert h \rvert < \delta$. Fix x for a moment and apply the mean value theorem to find a y_1 between y and $y+h$ such that

$$\frac{f(x,y+h)-f(x,y)}{h} = \frac{\partial f}{\partial y}(x,y_1).$$

As $\lvert y_1 - y \rvert \leq \lvert h \rvert < \delta$,

$$\left\lvert \frac{f(x,y+h)-f(x,y)}{h} - \frac{\partial f}{\partial y}(x,y) \right\rvert = \left\lvert \frac{\partial f}{\partial y}(x,y_1) - \frac{\partial f}{\partial y}(x,y) \right\rvert < \epsilon.$$

The argument worked for every $x \in [a,b]$ (different y_1 may have been used). Thus, as a function of x

$$x \mapsto \frac{f(x,y+h)-f(x,y)}{h} \qquad \text{converges uniformly to} \qquad x \mapsto \frac{\partial f}{\partial y}(x,y) \qquad \text{as } h \to 0.$$

We defined uniform convergence for sequences although the idea is the same. You may replace h with a sequence of nonzero numbers $\{h_n\}_{n=1}^{\infty}$ converging to 0 such that $y + h_n \in [c,d]$ and let $n \to \infty$.

Consider the difference quotient of g,

$$\frac{g(y+h)-g(y)}{h} = \frac{\int_a^b f(x,y+h)\,dx - \int_a^b f(x,y)\,dx}{h} = \int_a^b \frac{f(x,y+h)-f(x,y)}{h}\,dx.$$

Uniform convergence implies the limit can be taken underneath the integral. So

$$\lim_{h\to 0} \frac{g(y+h)-g(y)}{h} = \int_a^b \lim_{h\to 0} \frac{f(x,y+h)-f(x,y)}{h}\,dx = \int_a^b \frac{\partial f}{\partial y}(x,y)\,dx.$$

Then g' is continuous on $[c,d]$ by Proposition 7.5.12 from volume I mentioned above. □

Example 9.1.2: Let

$$f(y) = \int_0^1 \sin(x^2 - y^2)\, dx.$$

Then

$$f'(y) = \int_0^1 -2y \cos(x^2 - y^2)\, dx.$$

Example 9.1.3: Consider

$$\int_0^1 \frac{x-1}{\ln(x)}\, dx.$$

The function under the integral extends to be continuous on $[0,1]$, and hence the integral exists, see Exercise 9.1.1. Trouble is finding it. We introduce a parameter y and define a function:

$$g(y) := \int_0^1 \frac{x^y - 1}{\ln(x)}\, dx.$$

The function $\frac{x^y-1}{\ln(x)}$ also extends to a continuous function of x and y for $(x,y) \in [0,1] \times [0,1]$ (also part of the exercise). See Figure 9.1.

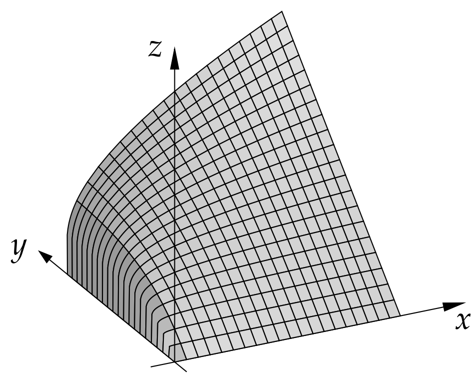

Figure 9.1: The graph $z = \frac{x^y-1}{\ln(x)}$ on $[0,1] \times [0,1]$.

Hence, g is a continuous function on $[0,1]$ and $g(0) = 0$. For every $\epsilon > 0$, the y derivative of the integrand, x^y, is continuous on $[0,1] \times [\epsilon, 1]$. Therefore, for $y > 0$, we may differentiate under the integral sign,

$$g'(y) = \int_0^1 \frac{\ln(x) x^y}{\ln(x)}\, dx = \int_0^1 x^y\, dx = \frac{1}{y+1}.$$

We need to figure out $g(1)$ given that $g'(y) = \frac{1}{y+1}$ and $g(0) = 0$. Elementary calculus says that $g(1) = \int_0^1 g'(y)\, dy = \ln(2)$. Thus,

$$\int_0^1 \frac{x-1}{\ln(x)}\, dx = \ln(2).$$

9.1.1 Exercises

***Exercise* 9.1.1:** *Prove the two statements that were asserted in Example 9.1.3:*

a) *Prove* $\frac{x-1}{\ln(x)}$ *extends to a continuous function of* $[0,1]$. *That is, there exists a continuous function on* $[0,1]$ *that equals* $\frac{x-1}{\ln(x)}$ *on* $(0,1)$.

b) *Prove* $\frac{x^y-1}{\ln(x)}$ *extends to a continuous function on* $[0,1]\times[0,1]$.

***Exercise* 9.1.2:** *Suppose* $h\colon \mathbb{R}\to\mathbb{R}$ *is continuous and* $g\colon\mathbb{R}\to\mathbb{R}$ *is continuously differentiable and compactly supported. That is, there exists some* $M>0$, *such that* $g(x)=0$ *whenever* $|x|\geq M$. *Define*

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

Show that f *is differentiable.*

***Exercise* 9.1.3:** *Suppose* $f\colon\mathbb{R}\to\mathbb{R}$ *is infinitely differentiable (derivatives of all orders exist) and* $f(0)=0$. *Show that there exists an infinitely differentiable function* $g\colon\mathbb{R}\to\mathbb{R}$ *such that* $f(x)=x\,g(x)$. *Show also that if* $f'(0)\neq 0$, *then* $g(0)\neq 0$.
Hint: Write $f(x)=\int_0^x f'(s)\,ds$ *and then rewrite the integral to go from* 0 *to* 1.

***Exercise* 9.1.4:** *Compute* $\int_0^1 e^{tx}\,dx$. *Derive the formula for* $\int_0^1 x^n e^x\,dx$ *not using integration by parts, but by differentiation underneath the integral.*

***Exercise* 9.1.5:** *Let* $U\subset\mathbb{R}^n$ *be open and suppose* $f(x,y_1,y_2,\ldots,y_n)$ *is a continuous function defined on* $[0,1]\times U\subset\mathbb{R}^{n+1}$. *Suppose* $\frac{\partial f}{\partial y_1},\frac{\partial f}{\partial y_2},\ldots,\frac{\partial f}{\partial y_n}$ *exist and are continuous on* $[0,1]\times U$. *Prove that* $F\colon U\to\mathbb{R}$ *defined by*

$$F(y_1,y_2,\ldots,y_n) := \int_0^1 f(x,y_1,y_2,\ldots,y_n)\,dx$$

is continuously differentiable.

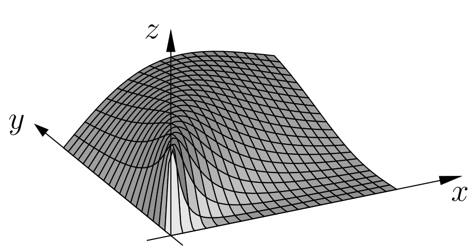

Figure 9.2: The graph $z=\frac{xy^3}{(x^2+y^2)^2}$ on $[0,1]\times[0,1]$.

***Exercise* 9.1.6:** *Work out the following counterexample: Let*

$$f(x,y) := \begin{cases} \frac{xy^3}{(x^2+y^2)^2} & \text{if } x \neq 0 \text{ or } y \neq 0, \\ 0 & \text{if } x = 0 \text{ and } y = 0. \end{cases}$$

See Figure 9.2.

a) *Prove that for every fixed y, the function $x \mapsto f(x,y)$ is Riemann integrable on $[0,1]$, and*

$$g(y) := \int_0^1 f(x,y)\,dx = \frac{y}{2y^2+2}.$$

Therefore, $g'(y)$ exists and its derivative is the continuous function

$$g'(y) = \frac{d}{dy}\int_0^1 f(x,y)\,dx = \frac{1-y^2}{2(y^2+1)^2}.$$

b) *Prove $\frac{\partial f}{\partial y}$ exists at all x and y and compute it.*

c) *Show that for all y*

$$\int_0^1 \frac{\partial f}{\partial y}(x,y)\,dx$$

exists, but

$$g'(0) \neq \int_0^1 \frac{\partial f}{\partial y}(x,0)\,dx.$$

***Exercise* 9.1.7:** *Work out the following counterexample: Let*

$$f(x,y) := \begin{cases} x \, \sin\left(\frac{y}{x^2+y^2}\right) & \text{if } (x,y) \neq (0,0), \\ 0 & \text{if } (x,y) = (0,0). \end{cases}$$

a) *Prove f is continuous on all of $\mathbb{R}^2$. Therefore the following function is well-defined for every $y \in \mathbb{R}$:*

$$g(y) := \int_0^1 f(x,y)\,dx.$$

b) *Prove $\frac{\partial f}{\partial y}$ exists for all (x,y), but is not continuous at $(0,0)$.*

c) *Show that $\int_0^1 \frac{\partial f}{\partial y}(x,0)\,dx$ does not exist even if we take improper integrals, that is, that the limit $\lim\limits_{h\to 0^+} \int_h^1 \frac{\partial f}{\partial y}(x,0)\,dx$ does not exist.*

Note: Feel free to use what you know about sine and cosine from calculus.

***Exercise* 9.1.8:** *Strengthen the Leibniz integral rule in the following way. Suppose* $f : (a, b) \times (c, d) \to \mathbb{R}$ *is a bounded continuous function, such that* $\frac{\partial f}{\partial y}$ *exists for all* $(x, y) \in (a, b) \times (c, d)$ *and is continuous and bounded. Define* $g : (c, d) \to \mathbb{R}$ *by*

$$g(y) := \int_a^b f(x, y)\, dx.$$

Then g *is continuously differentiable and*

$$g'(y) = \int_a^b \frac{\partial f}{\partial y}(x, y)\, dx.$$

Hint: See also Exercise 7.5.18 and Theorem 6.2.10 from volume I.

***Exercise* 9.1.9:** *Suppose* $g : \mathbb{R} \to \mathbb{R}$ *is continuously differentiable,* $h : \mathbb{R}^2 \to \mathbb{R}$ *is continuous,* $\frac{\partial h}{\partial x}$ *exists and is continuous at all points. Show that*

$$F(x, y) := g(x) + \int_0^y h(x, s)\, ds$$

is continuously differentiable, and that it is the solution of the partial differential equation $\frac{\partial F}{\partial y} = h$*, with the initial condition* $F(x, 0) = g(x)$ *for all* $x \in \mathbb{R}$*.*

9.2 Path integrals

Note: 2–3 lectures

9.2.1 Piecewise smooth paths

Let $\gamma\colon [a,b] \to \mathbb{R}^n$ be a function and write $\gamma = (\gamma_1, \gamma_2, \ldots, \gamma_n)$. Suppose γ is *continuously differentiable*, meaning it is differentiable and the derivative is continuous. In other words, there exists a continuous function $\gamma'\colon [a,b] \to \mathbb{R}^n$ such that for every $t \in [a,b]$, we have $\lim\limits_{h\to 0} \frac{\lVert \gamma(t+h)-\gamma(t)-\gamma'(t)\,h\rVert}{\lvert h\rvert} = 0$. We treat $\gamma'(t)$ either as a linear operator (an $n \times 1$ matrix) or a vector, $\gamma'(t) = \bigl(\gamma_1'(t), \gamma_2'(t), \ldots, \gamma_n'(t)\bigr)$. Equivalently, γ_j is a continuously differentiable function on $[a,b]$ for every $j = 1, 2, \ldots, n$. By Exercise 8.2.6, the operator norm of the operator $\gamma'(t)$ equals the euclidean norm of the corresponding vector, which allows us to write $\lVert\gamma'(t)\rVert$ without any confusion.

Definition 9.2.1. A continuously differentiable function $\gamma\colon [a,b] \to \mathbb{R}^n$ is called a *smooth path* or a *continuously differentiable path** if γ is continuously differentiable and $\gamma'(t) \neq 0$ for all $t \in [a,b]$.

The function $\gamma\colon [a,b] \to \mathbb{R}^n$ is called a *piecewise smooth path* or a *piecewise continuously differentiable path* if there exist finitely many points $t_0 = a < t_1 < t_2 < \cdots < t_k = b$ such that the restriction $\gamma|_{[t_{j-1},t_j]}$ is smooth path for every $j = 1, 2, \ldots, k$.

A path γ is a *closed path* if $\gamma(a) = \gamma(b)$, that is, the path starts and ends in the same point. A path γ is a *simple path* if either 1) γ is a one-to-one function, or 2) $\gamma|_{[a,b)}$ is one-to-one and $\gamma(a) = \gamma(b)$ (γ is a simple closed path).

Example 9.2.2: Let $\gamma\colon [0,4] \to \mathbb{R}^2$ be defined by

$$\gamma(t) := \begin{cases} (t,0) & \text{if } t \in [0,1], \\ (1,t-1) & \text{if } t \in (1,2], \\ (3-t,1) & \text{if } t \in (2,3], \\ (0,4-t) & \text{if } t \in (3,4]. \end{cases}$$

The path γ is the unit square traversed counterclockwise. See Figure 9.3. It is a piecewise smooth path. For example, $\gamma|_{[1,2]}(t) = (1,t-1)$ and so $(\gamma|_{[1,2]})'(t) = (0,1) \neq 0$. Similarly for the other 3 sides. Notice that $(\gamma|_{[1,2]})'(1) = (0,1)$, $(\gamma|_{[0,1]})'(1) = (1,0)$, but $\gamma'(1)$ does not exist. At the corners γ is not differentiable. The path γ is a simple closed path, as $\gamma|_{[0,4)}$ is one-to-one and $\gamma(0) = \gamma(4)$.

The definition of a piecewise smooth path as we have given it implies continuity (exercise). For general functions, many authors also allow finitely many discontinuities, when they use the term *piecewise smooth*, and so one may say that we defined a piecewise

*The word "smooth" can sometimes mean "infinitely differentiable" in the literature.

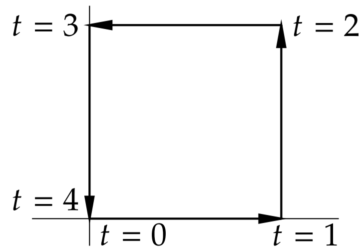

Figure 9.3: The path γ traversing the unit square.

smooth path to be a *continuous piecewise smooth* function. While one may get by with smooth paths, for computations, the simplest paths to write down are often piecewise smooth.

Generally, we are interested in the direct image $\gamma\bigl([a,b]\bigr)$, rather than the specific parametrization, although that is also important to some degree. When we informally talk about a path or a curve, we often mean the set $\gamma\bigl([a,b]\bigr)$, depending on context.

Example 9.2.3: The condition $\gamma'(t) \neq 0$ means that the image $\gamma\bigl([a,b]\bigr)$ has no "corners" where γ is smooth. Consider

$$\gamma(t) := \begin{cases} (t^2, 0) & \text{if } t < 0, \\ (0, t^2) & \text{if } t \geq 0. \end{cases}$$

See Figure 9.4. It is left for the reader to check that γ is continuously differentiable, yet the image $\gamma(\mathbb{R}) = \bigl\{(x,y) \in \mathbb{R}^2 : (x,y) = (s,0) \text{ or } (x,y) = (0,s) \text{ for some } s \geq 0\bigr\}$ has a "corner" at the origin. And that is because $\gamma'(0) = (0,0)$. More complicated examples with, say, infinitely many corners exist, see the exercises.

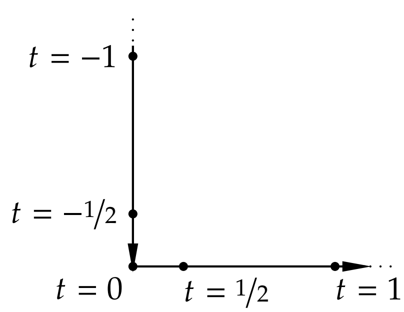

Figure 9.4: "Smooth" path with a corner if we allow zero derivative. The points corresponding to several values of t are marked with dots.

The condition $\gamma'(t) \neq 0$ even at the endpoints of every subinterval guarantees not only no corners, but also that the path ends nicely, that is, it can extend a little bit past the endpoints if needed. Again, see the exercises.

Example 9.2.4: A graph of a continuously differentiable function $f : [a,b] \to \mathbb{R}$ is a smooth path. Define $\gamma : [a,b] \to \mathbb{R}^2$ by

$$\gamma(t) := \bigl(t, f(t)\bigr).$$

Then $\gamma'(t) = \bigl(1, f'(t)\bigr)$, which is never zero, and $\gamma\bigl([a,b]\bigr)$ is the graph of f.

There are other ways of parametrizing the path. That is, there are different paths with the same image. The function $t \mapsto (1-t)a + tb$, takes the interval $[0,1]$ to $[a,b]$. Define $\alpha : [0,1] \to \mathbb{R}^2$ by

$$\alpha(t) := \bigl((1-t)a + tb, f((1-t)a + tb)\bigr).$$

Then $\alpha'(t) = \bigl(b-a,\ (b-a)f'((1-t)a+tb)\bigr)$, which is never zero. As sets, $\alpha\bigl([0,1]\bigr) = \gamma\bigl([a,b]\bigr) = \bigl\{(x,y) \in \mathbb{R}^2 : x \in [a,b] \text{ and } f(x) = y\bigr\}$, which is just the graph of f.

The last example leads us to a definition.

Definition 9.2.5. Let $\gamma : [a,b] \to \mathbb{R}^n$ be a smooth path and $h : [c,d] \to [a,b]$ a continuously differentiable bijective function such that $h'(t) \neq 0$ for all $t \in [c,d]$. Then the composition $\gamma \circ h$ is called a *smooth reparametrization* of γ.

Let γ be a piecewise smooth path, and h a piecewise smooth bijective function with nonzero one-sided limits of h'. The composition $\gamma \circ h$ is called a *piecewise smooth reparametrization* of γ.

If h is strictly increasing, then h is said to *preserve orientation*. If h does not preserve orientation, then h is said to *reverse orientation*.

A reparametrization is another path for the same set. That is, $(\gamma \circ h)\bigl([c,d]\bigr) = \gamma\bigl([a,b]\bigr)$.

The conditions on the piecewise smooth h mean that there is some partition $t_0 = c < t_1 < t_2 < \cdots < t_k = d$, such that $h|_{[t_{j-1},t_j]}$ is continuously differentiable and $\bigl(h|_{[t_{j-1},t_j]}\bigr)'(t) \neq 0$ for all $t \in [t_{j-1}, t_j]$. Since h is bijective, it is either strictly increasing or strictly decreasing. So either $\bigl(h|_{[t_{j-1},t_j]}\bigr)'(t) > 0$ for all t or $\bigl(h|_{[t_{j-1},t_j]}\bigr)'(t) < 0$ for all t.

Proposition 9.2.6. *If $\gamma : [a,b] \to \mathbb{R}^n$ is a piecewise smooth path, and $\gamma \circ h : [c,d] \to \mathbb{R}^n$ is a piecewise smooth reparametrization, then $\gamma \circ h$ is a piecewise smooth path.*

Proof. Assume that h preserves orientation, that is, h is strictly increasing. If $h : [c,d] \to [a,b]$ gives a piecewise smooth reparametrization, then for some partition $r_0 = c < r_1 < r_2 < \cdots < r_\ell = d$, the restriction $h|_{[r_{j-1},r_j]}$ is continuously differentiable with a positive derivative.

Let $t_0 = a < t_1 < t_2 < \cdots < t_k = b$ be the partition from the definition of piecewise smooth for γ together with the points $\{h(r_0), h(r_1), h(r_2), \ldots, h(r_\ell)\}$. Let $s_j := h^{-1}(t_j)$. Then $s_0 = c < s_1 < s_2 < \cdots < s_k = d$ is a partition that includes (is a refinement of) the $\{r_0, r_1, \ldots, r_\ell\}$. If $\tau \in [s_{j-1}, s_j]$, then $h(\tau) \in [t_{j-1}, t_j]$ since $h(s_{j-1}) = t_{j-1}$, $h(s_j) = t_j$, and h is strictly increasing. Also $h|_{[s_{j-1},s_j]}$ is continuously differentiable, and $\gamma|_{[t_{j-1},t_j]}$ is also continuously differentiable. Then

$$(\gamma \circ h)|_{[s_{j-1},s_j]}(\tau) = \gamma|_{[t_{j-1},t_j]}\bigl(h|_{[s_{j-1},s_j]}(\tau)\bigr).$$

The function $(\gamma \circ h)|_{[s_{j-1},s_j]}$ is therefore continuously differentiable and by the chain rule

$$\big((\gamma \circ h)|_{[s_{j-1},s_j]}\big)'(\tau) = \big(\gamma|_{[t_{j-1},t_j]}\big)'\big(h(\tau)\big)\big(h|_{[s_{j-1},s_j]}\big)'(\tau) \neq 0.$$

Consequently, $\gamma \circ h$ is a piecewise smooth path. The proof for orientation reversing h is left as an exercise. □

If two paths are simple and their images are the same, it is left as an exercise that there exists a reparametrization. Here is where our assumption that γ' is never zero is important.

9.2.2 Path integral of a one-form

Definition 9.2.7. Let $(x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$ be our coordinates. Given n real-valued continuous functions $\omega_1, \omega_2, \ldots, \omega_n$ defined on a set $S \subset \mathbb{R}^n$, we define a *one-form* to be an object of the form

$$\omega = \omega_1 \, dx_1 + \omega_2 \, dx_2 + \cdots + \omega_n \, dx_n .$$

We could represent ω as a continuous function from S to $\mathbb{R}^n$, although it is better to think of it as a different object.

Example 9.2.8:

$$\omega(x,y) := \frac{-y}{x^2+y^2} \, dx + \frac{x}{x^2+y^2} \, dy$$

is a one-form defined on $\mathbb{R}^2 \setminus \{(0,0)\}$.

Definition 9.2.9. Let $\gamma \colon [a,b] \to \mathbb{R}^n$ be a smooth path and let

$$\omega = \omega_1 \, dx_1 + \omega_2 \, dx_2 + \cdots + \omega_n \, dx_n$$

be a one-form defined on the direct image $\gamma\big([a,b]\big)$. Write $\gamma = (\gamma_1, \gamma_2, \ldots, \gamma_n)$. Define:

$$\begin{aligned}
\int_\gamma \omega &:= \int_a^b \Big(\omega_1\big(\gamma(t)\big)\gamma_1'(t) + \omega_2\big(\gamma(t)\big)\gamma_2'(t) + \cdots + \omega_n\big(\gamma(t)\big)\gamma_n'(t) \Big) dt \\
&= \int_a^b \left(\sum_{j=1}^n \omega_j\big(\gamma(t)\big)\gamma_j'(t) \right) dt .
\end{aligned}$$

To remember the definition note that x_j is $\gamma_j(t)$, so dx_j becomes $\gamma_j'(t)\, dt$.

If γ is piecewise smooth, take the corresponding partition $t_0 = a < t_1 < t_2 < \ldots < t_k = b$, and assume the partition is minimal in the sense that γ is not differentiable at $t_1, t_2, \ldots, t_{k-1}$. As each $\gamma|_{[t_{j-1},t_j]}$ is a smooth path, define

$$\int_\gamma \omega := \int_{\gamma|_{[t_0,t_1]}} \omega \, + \int_{\gamma|_{[t_1,t_2]}} \omega \, + \cdots + \int_{\gamma|_{[t_{k-1},t_k]}} \omega .$$

The notation makes sense from the formula you remember from calculus, let us state it somewhat informally: If $x_j(t) = \gamma_j(t)$, then $dx_j = \gamma_j'(t)\,dt$.

Paths can be cut up or concatenated. The proof is a direct application of the additivity of the Riemann integral, and is left as an exercise. The proposition justifies why we defined the integral over a piecewise smooth path in the way we did, and it justifies that we may as well have taken any partition not just the minimal one in the definition.

Proposition 9.2.10. *Let $\gamma \colon [a,c] \to \mathbb{R}^n$ be a piecewise smooth path, and $b \in (a,c)$. Define the piecewise smooth paths $\alpha := \gamma|_{[a,b]}$ and $\beta := \gamma|_{[b,c]}$. Let ω be a one-form defined on $\gamma\bigl([a,c]\bigr)$. Then*

$$\int_\gamma \omega = \int_\alpha \omega + \int_\beta \omega.$$

Example 9.2.11: Let the one-form ω and the path $\gamma \colon [0,2\pi] \to \mathbb{R}^2$ be defined by

$$\omega(x,y) := \frac{-y}{x^2+y^2}\,dx + \frac{x}{x^2+y^2}\,dy, \qquad \gamma(t) := \bigl(\cos(t),\sin(t)\bigr).$$

Then

$$\begin{aligned} \int_\gamma \omega &= \int_0^{2\pi} \left(\frac{-\sin(t)}{\bigl(\cos(t)\bigr)^2 + \bigl(\sin(t)\bigr)^2} \bigl(-\sin(t)\bigr) + \frac{\cos(t)}{\bigl(\cos(t)\bigr)^2 + \bigl(\sin(t)\bigr)^2} \bigl(\cos(t)\bigr) \right) dt \\ &= \int_0^{2\pi} 1\,dt = 2\pi. \end{aligned}$$

Next, parametrize the same curve as $\alpha \colon [0,1] \to \mathbb{R}^2$ defined by $\alpha(t) := \bigl(\cos(2\pi t),\sin(2\pi t)\bigr)$, that is, α is a smooth reparametrization of γ. Then

$$\begin{aligned} \int_\alpha \omega = \int_0^1 \Biggl(& \frac{-\sin(2\pi t)}{\bigl(\cos(2\pi t)\bigr)^2 + \bigl(\sin(2\pi t)\bigr)^2} \bigl(-2\pi\sin(2\pi t)\bigr) \\ & + \frac{\cos(2\pi t)}{\bigl(\cos(2\pi t)\bigr)^2 + \bigl(\sin(2\pi t)\bigr)^2} \bigl(2\pi\cos(2\pi t)\bigr) \Biggr) dt \\ = \int_0^1 & 2\pi\,dt = 2\pi. \end{aligned}$$

Finally, reparametrize with $\beta \colon [0,2\pi] \to \mathbb{R}^2$ as $\beta(t) := \bigl(\cos(-t),\sin(-t)\bigr)$. Then

$$\begin{aligned} \int_\beta \omega &= \int_0^{2\pi} \left(\frac{-\sin(-t)}{\bigl(\cos(-t)\bigr)^2 + \bigl(\sin(-t)\bigr)^2} \bigl(\sin(-t)\bigr) + \frac{\cos(-t)}{\bigl(\cos(-t)\bigr)^2 + \bigl(\sin(-t)\bigr)^2} \bigl(-\cos(-t)\bigr) \right) dt \\ &= \int_0^{2\pi} (-1)\,dt = -2\pi. \end{aligned}$$

The path α is an orientation preserving reparametrization of γ, and the integrals are the same. The path β is an orientation reversing reparametrization of γ and the integral is minus the original. See Figure 9.5.

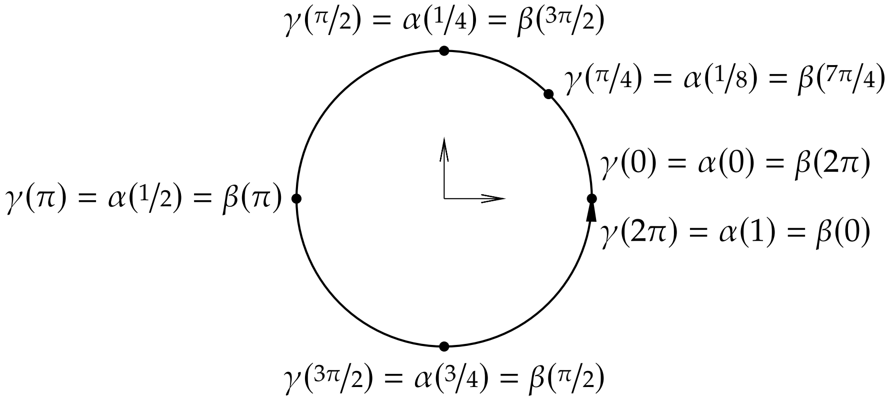

Figure 9.5: A circular path reparametrized in two different ways. The arrow indicates the orientation of γ and α. The path β traverses the circle in the opposite direction.

The previous example is not a fluke. The path integral does not depend on the parametrization of the curve, the only thing that matters is the direction in which the curve is traversed.

Proposition 9.2.12. *Let $\gamma\colon [a,b] \to \mathbb{R}^n$ be a piecewise smooth path and $\gamma \circ h\colon [c,d] \to \mathbb{R}^n$ a piecewise smooth reparametrization. Suppose ω is a one-form defined on the set $\gamma\big([a,b]\big)$. Then*

$$\int_{\gamma\circ h} \omega = \begin{cases} \int_\gamma \omega & \text{if } h \text{ preserves orientation,} \\ -\int_\gamma \omega & \text{if } h \text{ reverses orientation.} \end{cases}$$

Proof. Assume first that γ and h are both smooth. Write $\omega = \omega_1\, dx_1 + \omega_2\, dx_2 + \cdots + \omega_n\, dx_n$. Suppose that h is orientation preserving. Use the change of variables formula for the Riemann integral:

$$\begin{aligned} \int_\gamma \omega &= \int_a^b \left(\sum_{j=1}^n \omega_j\big(\gamma(t)\big)\gamma_j'(t) \right) dt \\ &= \int_c^d \left(\sum_{j=1}^n \omega_j\Big(\gamma\big(h(\tau)\big)\Big)\gamma_j'\big(h(\tau)\big) \right) h'(\tau)\, d\tau \\ &= \int_c^d \left(\sum_{j=1}^n \omega_j\Big(\gamma\big(h(\tau)\big)\Big)(\gamma_j \circ h)'(\tau) \right) d\tau = \int_{\gamma\circ h} \omega. \end{aligned}$$

If h is orientation reversing, it swaps the order of the limits on the integral and introduces a minus sign. The details, along with finishing the proof for piecewise smooth paths, is left as Exercise 9.2.4. □

Due to this proposition (and the exercises), if $\Gamma \subset \mathbb{R}^n$ is the image of a simple piecewise smooth path $\gamma([a,b])$, then as long as we somehow indicate the orientation, that is, the direction in which we traverse the curve, we can write

$$\int_\Gamma \omega,$$

without mentioning the specific γ. Furthermore, for a simple closed path, it does not even matter where we start the parametrization. See the exercises.

Recall that *simple* means that γ is one-to-one except perhaps at the endpoints, in particular it is one-to-one when restricted to $[a,b)$. We may relax the condition that the path is simple a little bit. For example, it is enough to suppose that $\gamma\colon [a,b] \to \mathbb{R}^n$ is one-to-one except at finitely many points. See Exercise 9.2.14. But we cannot remove the condition completely as is illustrated by the following example.

Example 9.2.13: Take $\gamma\colon [0,2\pi] \to \mathbb{R}^2$ given by $\gamma(t) \coloneqq \bigl(\cos(t), \sin(t)\bigr)$, and $\beta\colon [0,2\pi] \to \mathbb{R}^2$ by $\beta(t) \coloneqq \bigl(\cos(2t), \sin(2t)\bigr)$. Notice that $\gamma\bigl([0,2\pi]\bigr) = \beta\bigl([0,2\pi]\bigr)$; we travel around the same curve, the unit circle. But γ goes around the unit circle once in the counter clockwise direction, and β goes around the unit circle twice (in the same direction). See Figure 9.6.

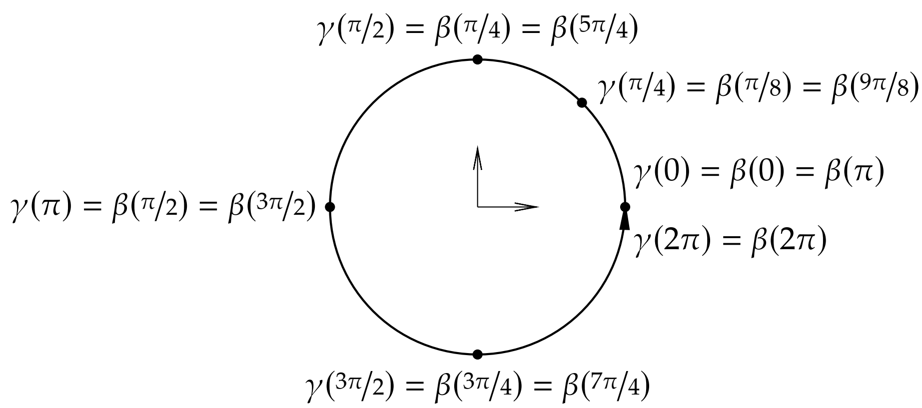

Figure 9.6: Circular path traversed once by $\gamma\colon [0,2\pi] \to \mathbb{R}^2$ and twice by $\beta\colon [0,2\pi] \to \mathbb{R}^2$.

Compute

$$\int_\gamma -y\,dx + x\,dy = \int_0^{2\pi} \Bigl(\bigl(-\sin(t)\bigr)\bigl(-\sin(t)\bigr) + \cos(t)\cos(t)\Bigr)\,dt = 2\pi,$$
$$\int_\beta -y\,dx + x\,dy = \int_0^{2\pi} \Bigl(\bigl(-\sin(2t)\bigr)\bigl(-2\sin(2t)\bigr) + \cos(t)\bigl(2\cos(t)\bigr)\Bigr)\,dt = 4\pi.$$

It is sometimes convenient to define a path integral over $\gamma\colon [a,b] \to \mathbb{R}^n$ that is not a path. Define

$$\int_\gamma \omega \coloneqq \int_a^b \left(\sum_{j=1}^n \omega_j\bigl(\gamma(t)\bigr)\gamma_j'(t) \right) dt$$

for every continuously differentiable γ. A case that comes up naturally is when γ is constant. Then $\gamma'(t) = 0$ for all t, and $\gamma([a,b])$ is a single point, which we regard as a "curve" of length zero. Then, $\int_\gamma \omega = 0$ for every ω.

9.2.3 Path integral of a function

Next, we integrate a function against the so-called *arc-length measure* ds. The geometric picture we have in mind is the area under the graph of the function over a path. Imagine a fence erected over γ with height given by the function and the integral is the area of the fence. See Figure 9.7.

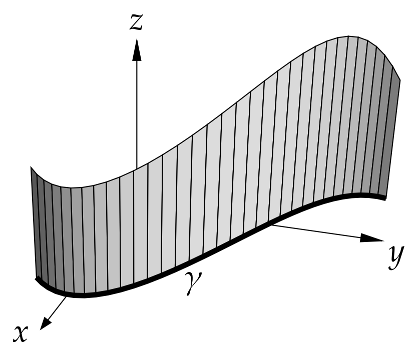

Figure 9.7: A path $\gamma\colon [a,b] \to \mathbb{R}^2$ in the xy-plane (bold curve), and a function $z = f(x,y)$ graphed above it in the z direction. The integral is the shaded area depicted.

Definition 9.2.14. Suppose $\gamma\colon [a,b] \to \mathbb{R}^n$ is a smooth path, and f is a continuous function defined on the image $\gamma([a,b])$. Then define

$$\int_\gamma f\, ds := \int_a^b f\bigl(\gamma(t)\bigr)\lVert\gamma'(t)\rVert\, dt.$$

To emphasize the variables we may use

$$\int_\gamma f(x)\, ds(x) := \int_\gamma f\, ds.$$

The definition for a piecewise smooth path is similar as before and is left to the reader.

The path integral of a function is also independent of the parametrization, and in this case, the orientation does not matter.

Proposition 9.2.15. *Let $\gamma\colon [a,b] \to \mathbb{R}^n$ be a piecewise smooth path and $\gamma \circ h\colon [c,d] \to \mathbb{R}^n$ a piecewise smooth reparametrization. Suppose f is a continuous function defined on the set $\gamma([a,b])$. Then*

$$\int_{\gamma\circ h} f\, ds = \int_\gamma f\, ds.$$

Proof. Suppose h is orientation preserving and that γ and h are both smooth. Then

$$\begin{aligned}
\int_\gamma f\,ds &= \int_a^b f\big(\gamma(t)\big)\|\gamma'(t)\|\,dt \\
&= \int_c^d f\Big(\gamma\big(h(\tau)\big)\Big)\|\gamma'\big(h(\tau)\big)\|\,h'(\tau)\,d\tau \\
&= \int_c^d f\Big(\gamma\big(h(\tau)\big)\Big)\|\gamma'\big(h(\tau)\big)h'(\tau)\|\,d\tau \\
&= \int_c^d f\big((\gamma\circ h)(\tau)\big)\|(\gamma\circ h)'(\tau)\|\,d\tau \\
&= \int_{\gamma\circ h} f\,ds.
\end{aligned}$$

If h is orientation reversing it swaps the order of the limits on the integral, but you also have to introduce a minus sign in order to take h' inside the norm. The details, along with finishing the proof for piecewise smooth paths is left to the reader as Exercise 9.2.5. □

As before, due to this proposition (and the exercises), if γ is simple, it does not matter which parametrization we use. Therefore, if $\Gamma = \gamma\big([a,b]\big)$, we can simply write

$$\int_\Gamma f\,ds.$$

In this case we do not need to worry about orientation, either way we get the same integral.

Example 9.2.16: Let $f(x,y) := x$. Let $C \subset \mathbb{R}^2$ be half of the unit circle for $x \geq 0$. We wish to compute

$$\int_C f\,ds.$$

Parametrize the curve C via $\gamma\colon [{-\pi}/2, \pi/2] \to \mathbb{R}^2$ defined as $\gamma(t) := \big(\cos(t), \sin(t)\big)$. Then $\gamma'(t) = \big(-\sin(t), \cos(t)\big)$, and

$$\int_C f\,ds = \int_\gamma f\,ds = \int_{-\pi/2}^{\pi/2} \cos(t)\sqrt{\big(-\sin(t)\big)^2 + \big(\cos(t)\big)^2}\,dt = \int_{-\pi/2}^{\pi/2} \cos(t)\,dt = 2.$$

Definition 9.2.17. Suppose $\Gamma \subset \mathbb{R}^n$ is parametrized by a simple piecewise smooth path $\gamma\colon [a,b] \to \mathbb{R}^n$, that is, $\gamma\big([a,b]\big) = \Gamma$. We define the *length* by

$$\ell(\Gamma) := \int_\Gamma ds = \int_\gamma ds.$$

If γ is smooth,

$$\ell(\Gamma) = \int_a^b \|\gamma'(t)\|\,dt.$$

This may be a good time to mention that it is common to write $\int_a^b \lVert \gamma'(t) \rVert \, dt$ even if the path is only piecewise smooth. That is because $\lVert \gamma'(t) \rVert$ is defined and continuous at all but finitely many points and is bounded, and so the integral exists.

Example 9.2.18: Let $x, y \in \mathbb{R}^n$ be two points and write $[x, y]$ as the straight line segment between the two points x and y. Parametrize $[x, y]$ by $\gamma(t) := (1-t)x + ty$ for t running between 0 and 1. See Figure 9.8. Then $\gamma'(t) = y - x$, and therefore

$$\ell\big([x,y]\big) = \int_{[x,y]} ds = \int_0^1 \lVert y - x \rVert \, dt = \lVert y - x \rVert .$$

The length of $[x, y]$ is the standard euclidean distance between x and y, justifying the name.

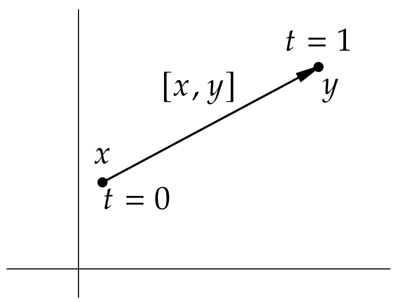

Figure 9.8: Straight path between x and y parametrized by $(1-t)x + ty$.

A simple piecewise smooth path $\gamma : [0, r] \to \mathbb{R}^n$ is said to be an *arc-length parametrization* if for all $t \in [0, r]$, we have

$$\ell\big(\gamma([0,t])\big) = t .$$

If γ is smooth, then

$$\int_0^t d\tau = t = \ell\big(\gamma([0,t])\big) = \int_0^t \lVert \gamma'(\tau) \rVert \, d\tau$$

for all t. The fundamental theorem of calculus then says that $\lVert \gamma'(t) \rVert = 1$ for all t. Similarly for piecewise smooth γ, we get $\lVert \gamma'(t) \rVert = 1$ for all t where the derivative exists. Think of such a parametrization as moving around your curve at speed 1. If $\gamma : [0, r] \to \mathbb{R}^n$ is an arclength parametrization, it is common to use s as the variable as $\int_\gamma f \, ds = \int_0^r f\big(\gamma(s)\big) \, ds$.

9.2.4 Exercises

***Exercise* 9.2.1:** *Show that if $\varphi : [a, b] \to \mathbb{R}^n$ is a piecewise smooth path as we defined it, then φ is a continuous function.*

***Exercise* 9.2.2:** *Finish the proof of Proposition 9.2.6 for orientation reversing reparametrizations.*

Exercise* 9.2.3*: *Prove Proposition 9.2.10.*

Exercise* 9.2.4*: *Finish the proof of Proposition 9.2.12 for*

a) orientation reversing reparametrizations, and

b) piecewise smooth paths and reparametrizations.

Exercise* 9.2.5*: *Finish the proof of Proposition 9.2.15 for*

a) orientation reversing reparametrizations, and

b) piecewise smooth paths and reparametrizations.

Exercise* 9.2.6*: *Suppose $\gamma\colon [a,b] \to \mathbb{R}^n$ is a piecewise smooth path, and f is a continuous function defined on the image $\gamma\bigl([a,b]\bigr)$. Provide a definition of $\int_\gamma f\,ds$.*

Exercise* 9.2.7*: *Directly using the definitions compute:*

a) The arc-length of the unit square from Example 9.2.2 using the given parametrization.

b) The arc-length of the unit circle using the parametrization $\gamma\colon [0,1] \to \mathbb{R}^2$, $\gamma(t) := \bigl(\cos(2\pi t), \sin(2\pi t)\bigr)$.

c) The arc-length of the unit circle using the parametrization $\beta\colon [0,2\pi] \to \mathbb{R}^2$, $\beta(t) := \bigl(\cos(t), \sin(t)\bigr)$.

Note: Feel free to use what you know about sine and cosine from calculus.

Exercise* 9.2.8*: *Suppose $\gamma\colon [0,1] \to \mathbb{R}^n$ is a smooth path, and ω is a one-form defined on the image $\gamma\bigl([a,b]\bigr)$. For $r \in [0,1]$, let $\gamma_r\colon [0,r] \to \mathbb{R}^n$ be defined as simply the restriction of γ to $[0,r]$. Show that the function $h(r) := \int_{\gamma_r} \omega$ is a continuously differentiable function on $[0,1]$.*

Exercise* 9.2.9*: *Suppose $\gamma\colon [a,b] \to \mathbb{R}^n$ is a smooth path. Show that there exists an $\epsilon > 0$ and a smooth function $\widetilde{\gamma}\colon (a-\epsilon, b+\epsilon) \to \mathbb{R}^n$ with $\widetilde{\gamma}(t) = \gamma(t)$ for all $t \in [a,b]$ and $\widetilde{\gamma}\,'(t) \neq 0$ for all $t \in (a-\epsilon, b+\epsilon)$. That is, prove that a smooth path extends some small distance past the end points.*

Exercise* 9.2.10*: *Suppose $\alpha\colon [a,b] \to \mathbb{R}^n$ and $\beta\colon [c,d] \to \mathbb{R}^n$ are piecewise smooth paths such that $\Gamma := \alpha\bigl([a,b]\bigr) = \beta\bigl([c,d]\bigr)$. Show that there exist finitely many points $\{p_1, p_2, \ldots, p_k\} \in \Gamma$, such that the sets $\alpha^{-1}\bigl(\{p_1, p_2, \ldots, p_k\}\bigr)$ and $\beta^{-1}\bigl(\{p_1, p_2, \ldots, p_k\}\bigr)$ are partitions of $[a,b]$ and $[c,d]$ such that on every subinterval the paths are smooth (that is, they are partitions as in the definition of piecewise smooth path).*

Exercise* 9.2.11*:

a) Suppose $\gamma\colon [a,b] \to \mathbb{R}^n$ and $\alpha\colon [c,d] \to \mathbb{R}^n$ are two smooth paths that are one-to-one and $\gamma\bigl([a,b]\bigr) = \alpha\bigl([c,d]\bigr)$. Then there exists a smooth reparametrization $h\colon [a,b] \to [c,d]$ such that $\gamma = \alpha \circ h$.
Hint 1: It is not hard to show h exists. The trick is to prove it is continuously differentiable with a nonzero derivative. Apply the implicit function theorem though it may at first seem the dimensions are wrong.
Hint 2: Worry about derivative of h in (a,b) first.

b) Prove the same thing as part a, but now for simple closed paths with the further assumption that $\gamma(a) = \gamma(b) = \alpha(c) = \alpha(d)$.

c) Prove parts a) and b) but for piecewise smooth paths, obtaining piecewise smooth reparametrizations.
Hint: The trick is to find two partitions such that when restricted to a subinterval of the partition both paths have the same image and are smooth, see the exercise above.

Exercise* 9.2.12*: *Suppose $\alpha\colon [a,b] \to \mathbb{R}^n$ and $\beta\colon [b,c] \to \mathbb{R}^n$ are piecewise smooth paths with $\alpha(b) = \beta(b)$. Let $\gamma\colon [a,c] \to \mathbb{R}^n$ be defined by*

$$\gamma(t) := \begin{cases} \alpha(t) & \text{if } t \in [a,b], \\ \beta(t) & \text{if } t \in (b,c]. \end{cases}$$

Show that γ is a piecewise smooth path, and that if ω is a one-form defined on the curve given by γ, then

$$\int_\gamma \omega = \int_\alpha \omega + \int_\beta \omega.$$

Exercise* 9.2.13*: *Suppose $\gamma\colon [a,b] \to \mathbb{R}^n$ and $\beta\colon [c,d] \to \mathbb{R}^n$ are two simple closed piecewise smooth paths. That is, $\gamma(a) = \gamma(b)$ and $\beta(c) = \beta(d)$ and the restrictions $\gamma|_{[a,b)}$ and $\beta|_{[c,d)}$ are one-to-one. Suppose $\Gamma = \gamma\big([a,b]\big) = \beta\big([c,d]\big)$ and ω is a one-form defined on $\Gamma \subset \mathbb{R}^n$. Show that either*

$$\int_\gamma \omega = \int_\beta \omega, \qquad \text{or} \qquad \int_\gamma \omega = -\int_\beta \omega.$$

In particular, the notation $\int_\Gamma \omega$ makes sense if we indicate the direction in which the integral is evaluated. Hint: See previous three exercises.

Exercise* 9.2.14*: *Suppose $\gamma\colon [a,b] \to \mathbb{R}^n$ and $\beta\colon [c,d] \to \mathbb{R}^n$ are two piecewise smooth paths which are one-to-one except at finitely many points. That is, there exist finite sets $S \subset [a,b]$ and $T \subset [c,d]$ such that $\gamma|_{[a,b]\setminus S}$ and $\beta|_{[c,d]\setminus T}$ are one-to-one. Suppose $\Gamma = \gamma\big([a,b]\big) = \beta\big([c,d]\big)$ and ω is a one-form defined on $\Gamma \subset \mathbb{R}^n$. Show that either*

$$\int_\gamma \omega = \int_\beta \omega, \qquad \text{or} \qquad \int_\gamma \omega = -\int_\beta \omega.$$

In particular, the notation $\int_\Gamma \omega$ makes sense if we indicate the direction in which the integral is evaluated. Hint: Same hint as the last exercise.

Exercise* 9.2.15*: *Define $\gamma\colon [0,1] \to \mathbb{R}^2$ by $\gamma(t) := \Big(t^3 \sin(1/t),\ t\big(3t^2\sin(1/t) - t\cos(1/t)\big)^2\Big)$ for $t \neq 0$ and $\gamma(0) = (0,0)$. Show that*

a) γ is continuously differentiable on $[0,1]$.

b) Show that there exists an infinite sequence $\{t_n\}_{n=1}^\infty$ in $[0,1]$ converging to 0, such that $\gamma\,'(t_n) = (0,0)$.

c) Show that the points $\gamma(t_n)$ lie on the line $y = 0$ and such that the x-coordinate of $\gamma(t_n)$ alternates between positive and negative (if they do not alternate you only found a subsequence, you need to find them all).

d) Show that there is no piecewise smooth α whose image equals $\gamma\big([0,1]\big)$. Hint: Look at part c) and show that α' must be zero where it reaches the origin.

e) (Computer) If you know a plotting software that allows you to plot parametric curves, make a plot of the curve, but only for t in the range $[0,0.1]$ otherwise you will not see the behavior. In particular, you should notice that $\gamma\big([0,1]\big)$ has infinitely many "corners" near the origin.

Note: Feel free to use what you know about sine and cosine from calculus.

9.3 Path independence

Note: 2 lectures

9.3.1 Path independent integrals

Let $U \subset \mathbb{R}^n$ be a set and ω a one-form defined on U. The integral of ω is said to be *path independent* if for every pair of points $x, y \in U$ and every pair of piecewise smooth paths $\gamma\colon [a,b] \to U$ and $\beta\colon [c,d] \to U$ such that $\gamma(a) = \beta(c) = x$ and $\gamma(b) = \beta(d) = y$, we have

$$\int_\gamma \omega = \int_\beta \omega.$$

In this case, we simply write

$$\int_x^y \omega := \int_\gamma \omega = \int_\beta \omega.$$

Not every one-form gives a path independent integral. Most do not.

Example 9.3.1: Let $\gamma\colon [0,1] \to \mathbb{R}^2$ be the path $\gamma(t) := (t, 0)$ going from $(0,0)$ to $(1,0)$. Let $\beta\colon [0,1] \to \mathbb{R}^2$ be the path $\beta(t) := \bigl(t, (1-t)t\bigr)$ also going between the same points. Then

$$\begin{aligned}
\int_\gamma y\,dx &= \int_0^1 \gamma_2(t)\gamma_1'(t)\,dt = \int_0^1 0(1)\,dt = 0,\\
\int_\beta y\,dx &= \int_0^1 \beta_2(t)\beta_1'(t)\,dt = \int_0^1 (1-t)t(1)\,dt = \frac{1}{6}.
\end{aligned}$$

The integral of $y\,dx$ is not path independent. In particular, $\int_{(0,0)}^{(1,0)} y\,dx$ does not make sense.

Definition 9.3.2. Let $U \subset \mathbb{R}^n$ be an open set and $f\colon U \to \mathbb{R}$ a continuously differentiable function. The one-form

$$df := \frac{\partial f}{\partial x_1}\,dx_1 + \frac{\partial f}{\partial x_2}\,dx_2 + \cdots + \frac{\partial f}{\partial x_n}\,dx_n$$

is called the *total derivative* of f.

An open set $U \subset \mathbb{R}^n$ is said to be *path connected*[*] if for every two points x and y in U, there exists a piecewise smooth path starting at x and ending at y.

We leave as an exercise that every connected open set is path connected.

[*]Normally only a continuous path is used in this definition, but for open sets the two definitions are equivalent. See the exercises.

Proposition 9.3.3. *Let $U \subset \mathbb{R}^n$ be a path connected open set and ω a one-form defined on U. Then $\int_x^y \omega$ is path independent (for all $x, y \in U$) if and only if there exists a continuously differentiable $f : U \to \mathbb{R}$ such that $\omega = df$.*

In fact, if such an f exists, then for every pair of points $x, y \in U$

$$\int_x^y \omega = f(y) - f(x).$$

In other words, if we fix $p \in U$, then $f(x) = C + \int_p^x \omega$ for some constant C.

Proof. First suppose that the integral is path independent. Pick $p \in U$. Since U is path connected, there exists a path from p to every $x \in U$. Define

$$f(x) := \int_p^x \omega.$$

Write $\omega = \omega_1\, dx_1 + \omega_2\, dx_2 + \cdots + \omega_n\, dx_n$. We wish to show that for every $j = 1, 2, \ldots, n$, the partial derivative $\frac{\partial f}{\partial x_j}$ exists and is equal to ω_j.

Let e_j be an arbitrary standard basis vector, and h a nonzero real number. Compute

$$\frac{f(x+he_j) - f(x)}{h} = \frac{1}{h}\left(\int_p^{x+he_j} \omega - \int_p^x \omega\right) = \frac{1}{h}\int_x^{x+he_j} \omega,$$

which follows by Proposition 9.2.10 and path independence as $\int_p^{x+he_j} \omega = \int_p^x \omega + \int_x^{x+he_j} \omega$, because we pick a path from p to $x + he_j$ that also happens to pass through x, and then we cut this path in two, see Figure 9.9.

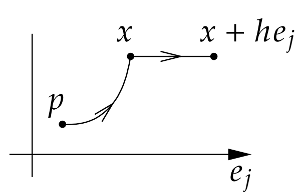

Figure 9.9: Using path independence in computing the partial derivative.

Since U is open, suppose h is so small so that all points of distance $|h|$ or less from x are in U. As the integral is path independent, pick the simplest path possible from x to $x + he_j$, that is, $\gamma(t) := x + the_j$ for $t \in [0, 1]$. The path is in U. Notice $\gamma'(t) = he_j$ has only one nonzero component and that is the jth component, which is h. Therefore,

$$\frac{1}{h}\int_x^{x+he_j} \omega = \frac{1}{h}\int_\gamma \omega = \frac{1}{h}\int_0^1 \omega_j(x + the_j)h\, dt = \int_0^1 \omega_j(x + the_j)\, dt.$$

We wish to take the limit as $h \to 0$. The function ω_j is continuous at x. Given $\epsilon > 0$, suppose h is small enough so that $\left|\omega_j(x) - \omega_j(y)\right| < \epsilon$ whenever $\|x - y\| \leq |h|$. Thus, $\left|\omega_j(x + the_j) - \omega_j(x)\right| < \epsilon$ for all $t \in [0,1]$, and we estimate

$$\left| \int_0^1 \omega_j(x + the_j)\,dt - \omega_j(x) \right| = \left| \int_0^1 \bigl(\omega_j(x + the_j) - \omega_j(x)\bigr)\,dt \right| \leq \epsilon.$$

That is,

$$\lim_{h \to 0} \frac{f(x + he_j) - f(x)}{h} = \omega_j(x).$$

All partials of f exist and are equal to ω_j, which are continuous functions. Thus, f is continuously differentiable, and furthermore $df = \omega$.

For the other direction, suppose a continuously differentiable f exists such that $df = \omega$. Take a smooth path $\gamma\colon [a,b] \to U$ such that $\gamma(a) = x$ and $\gamma(b) = y$. Then

$$\begin{aligned}
\int_\gamma df &= \int_a^b \left(\frac{\partial f}{\partial x_1}\bigl(\gamma(t)\bigr)\gamma_1'(t) + \frac{\partial f}{\partial x_2}\bigl(\gamma(t)\bigr)\gamma_2'(t) + \cdots + \frac{\partial f}{\partial x_n}\bigl(\gamma(t)\bigr)\gamma_n'(t) \right) dt \\
&= \int_a^b \frac{d}{dt}\Bigl[f\bigl(\gamma(t)\bigr) \Bigr]\,dt \\
&= f(y) - f(x).
\end{aligned}$$

The value of the integral only depends on x and y, not the path taken. Therefore the integral is path independent. We leave checking this fact for a piecewise smooth path as an exercise. □

Path independence can be stated more neatly in terms of integrals over closed paths.

Proposition 9.3.4. *Let $U \subset \mathbb{R}^n$ be a path connected open set and ω a one-form defined on U. Then $\omega = df$ for some continuously differentiable $f\colon U \to \mathbb{R}$ if and only if*

$$\int_\gamma \omega = 0 \qquad \textit{for every piecewise smooth closed path } \gamma\colon [a,b] \to U.$$

Proof. Suppose $\omega = df$ and let γ be a piecewise smooth closed path. Since $\gamma(a) = \gamma(b)$ for a closed path, the previous proposition says

$$\int_\gamma \omega = f\bigl(\gamma(b)\bigr) - f\bigl(\gamma(a)\bigr) = 0.$$

Now suppose that for every piecewise smooth closed path γ, $\int_\gamma \omega = 0$. Let x, y be two points in U and let $\alpha\colon [0,1] \to U$ and $\beta\colon [0,1] \to U$ be two piecewise smooth paths with $\alpha(0) = \beta(0) = x$ and $\alpha(1) = \beta(1) = y$. See Figure 9.10.

Define $\gamma\colon [0,2] \to U$ by

$$\gamma(t) := \begin{cases} \alpha(t) & \text{if } t \in [0,1], \\ \beta(2-t) & \text{if } t \in (1,2]. \end{cases}$$

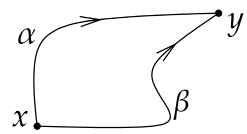

Figure 9.10: Two paths from x to y.

This path is piecewise smooth. This is due to the fact that $\gamma|_{[0,1]}(t) = \alpha(t)$ and $\gamma|_{[1,2]}(t) = \beta(2-t)$ (note especially $\gamma(1) = \alpha(1) = \beta(2-1)$). It is also closed as $\gamma(0) = \alpha(0) = \beta(0) = \gamma(2)$. So

$$0 = \int_\gamma \omega = \int_\alpha \omega - \int_\beta \omega.$$

This follows first by Proposition 9.2.10, and then noticing that the second part is β traveled backwards so that we get minus the β integral. Thus the integral of ω on U is path independent. $\square$

However one states path independence, it is often a difficult criterion to check, you have to check something "for all paths." There is a local criterion, a differential equation, that guarantees path independence, or in other words it guarantees an *antiderivative* f whose total derivative is the given one-form ω. Since the criterion is local, we generally only find the function f locally. We can find the antiderivative in every so-called *simply connected* domain, which informally is a connected open set where every path between two points can be "continuously deformed" into any other path between those two points. But to make matters simple, we prove the result for so-called star-shaped domains, which is often good enough. As a bonus the proof in the star-shaped case constructs the antiderivative explicitly. As balls are star-shaped we then have the result locally.

Definition 9.3.5. Let $U \subset \mathbb{R}^n$ be an open set and $p \in U$. We say U is a *star-shaped domain* with respect to p if for every other point $x \in U$, the line segment $[p, x]$ is in U, that is, if $(1-t)p + tx \in U$ for all $t \in [0, 1]$. If we say simply *star-shaped*, then U is star-shaped with respect to some $p \in U$. See Figure 9.11.

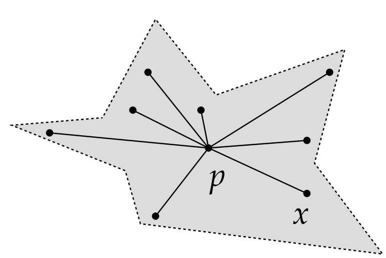

Figure 9.11: A star-shaped domain with respect to p.

Notice the difference between star-shaped and convex. Convex implies star-shaped, but a star-shaped domain need not be convex.

Theorem 9.3.6 (Poincaré lemma)**.** *Let $U \subset \mathbb{R}^n$ be a star-shaped domain and ω a continuously differentiable one-form defined on U. That is, if*

$$\omega = \omega_1 \, dx_1 + \omega_2 \, dx_2 + \cdots + \omega_n \, dx_n,$$

then $\omega_1, \omega_2, \ldots, \omega_n$ are continuously differentiable functions. Suppose that for every j and k

$$\frac{\partial \omega_j}{\partial x_k} = \frac{\partial \omega_k}{\partial x_j},$$

then there exists a twice continuously differentiable function $f \colon U \to \mathbb{R}$ such that $df = \omega$.

The condition on the derivatives of ω is precisely the condition that the second partial derivatives commute. That is, if $df = \omega$, and f is twice continuously differentiable, then

$$\frac{\partial \omega_j}{\partial x_k} = \frac{\partial^2 f}{\partial x_k \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_k} = \frac{\partial \omega_k}{\partial x_j}.$$

The condition is clearly necessary. The Poincaré lemma says that it is sufficient for a star-shaped U.

Proof. Suppose U is a star-shaped domain with respect to $p = (p_1, p_2, \ldots, p_n) \in U$. Given $x = (x_1, x_2, \ldots, x_n) \in U$, define the path $\gamma \colon [0,1] \to U$ as $\gamma(t) := (1-t)p + tx$, so $\gamma'(t) = x - p$. Let

$$f(x) := \int_\gamma \omega = \int_0^1 \left(\sum_{k=1}^n \omega_k\bigl((1-t)p + tx\bigr)\,(x_k - p_k) \right) dt.$$

We differentiate in x_j under the integral, which is allowed as everything, including the partials, is continuous:

$$\begin{aligned}
\frac{\partial f}{\partial x_j}(x) &= \int_0^1 \left(\left(\sum_{k=1}^n \frac{\partial \omega_k}{\partial x_j}\bigl((1-t)p + tx\bigr)\, t(x_k - p_k) \right) + \omega_j\bigl((1-t)p + tx\bigr) \right) dt \\
&= \int_0^1 \left(\left(\sum_{k=1}^n \frac{\partial \omega_j}{\partial x_k}\bigl((1-t)p + tx\bigr)\, t(x_k - p_k) \right) + \omega_j\bigl((1-t)p + tx\bigr) \right) dt \\
&= \int_0^1 \frac{d}{dt}\Bigl[t\omega_j\bigl((1-t)p + tx\bigr) \Bigr] \, dt \\
&= \omega_j(x).
\end{aligned}$$

That is precisely what we wanted. □

Example 9.3.7: Without some hypothesis on U, the theorem is not true. On $\mathbb{R}^2 \setminus \{(0,0)\}$, let

$$\omega(x,y) := \frac{-y}{x^2+y^2}\,dx + \frac{x}{x^2+y^2}\,dy$$

Then

$$\frac{\partial}{\partial y}\left[\frac{-y}{x^2+y^2}\right] = \frac{y^2-x^2}{(x^2+y^2)^2} = \frac{\partial}{\partial x}\left[\frac{x}{x^2+y^2}\right].$$

However, there is no $f\colon \mathbb{R}^2\setminus\{(0,0)\} \to \mathbb{R}$ such that $df = \omega$. In Example 9.2.11, we integrated from $(1,0)$ to $(1,0)$ along the unit circle counterclockwise, that is, $\gamma(t) = \bigl(\cos(t), \sin(t)\bigr)$ for $t \in [0,2\pi]$, and we found the integral to be 2π. We would have gotten 0 if the integral was path independent, or in other words if there would exist an f such that $df = \omega$.

9.3.2 Vector fields

A common object to integrate is a so-called vector field.

Definition 9.3.8. Let $U \subset \mathbb{R}^n$ be a set. A continuous function $v\colon U \to \mathbb{R}^n$ is called a *vector field*. Write $v = (v_1, v_2, \ldots, v_n)$.

Given a smooth path $\gamma\colon [a,b] \to \mathbb{R}^n$ with $\gamma\bigl([a,b]\bigr) \subset U$ we define the path integral of the vectorfield v as

$$\int_\gamma v \cdot d\gamma := \int_a^b v\bigl(\gamma(t)\bigr) \cdot \gamma'(t)\,dt,$$

where the dot in the definition is the standard dot product. The definition for a piecewise smooth path is, again, done by integrating over each smooth interval and adding the results.

Unraveling the definition, we find that

$$\int_\gamma v \cdot d\gamma = \int_\gamma v_1\,dx_1 + v_2\,dx_2 + \cdots + v_n\,dx_n.$$

What we know about integration of one-forms carries over to the integration of vector fields. For example, path independence for integration of vector fields is simply that

$$\int_x^y v \cdot d\gamma$$

is path independent if and only if $v = \nabla f$, that is, v is the gradient of a function. The function f is then called a *potential* for v.

A vector field v whose path integrals are path independent is called a *conservative vector field*. The rationale for the naming is that such vector fields arise in physical systems where a certain quantity, the energy, is conserved.

9.3.3 Exercises

Exercise* 9.3.1*: *Find an* $f : \mathbb{R}^2 \to \mathbb{R}$ *such that* $df = xe^{x^2+y^2}\, dx + ye^{x^2+y^2}\, dy$.

Exercise* 9.3.2*: *Find an* $\omega_2 : \mathbb{R}^2 \to \mathbb{R}$ *such that there exists a continuously differentiable* $f : \mathbb{R}^2 \to \mathbb{R}$ *for which* $df = e^{xy}\, dx + \omega_2\, dy$.

Exercise* 9.3.3*: *Finish the proof of Proposition 9.3.3, that is, we only proved the second direction for a smooth path, not a piecewise smooth path.*

Exercise* 9.3.4*: *Show that if* $U \subset \mathbb{R}^n$ *is a star-shaped domain, then it is path connected.*

Exercise* 9.3.5*: *Show that* $U := \mathbb{R}^2 \setminus \{(x, y) \in \mathbb{R}^2 : x \leq 0, y = 0\}$ *is a star-shaped domain and find all points* $(x_0, y_0) \in U$ *such that* U *is star-shaped with respect to* (x_0, y_0).

Exercise* 9.3.6*: *Suppose* U_1 *and* U_2 *are two open sets in* $\mathbb{R}^n$ *with* $U_1 \cap U_2$ *nonempty and path connected. Suppose there exists an* $f_1 : U_1 \to \mathbb{R}$ *and* $f_2 : U_2 \to \mathbb{R}$, *both twice continuously differentiable such that* $df_1 = df_2$ *on* $U_1 \cap U_2$. *Then there exists a twice differentiable function* $F : U_1 \cup U_2 \to \mathbb{R}$ *such that* $dF = df_1$ *on* U_1 *and* $dF = df_2$ *on* U_2.

Exercise* 9.3.7** (Hard):*** *Let* $\gamma : [a, b] \to \mathbb{R}^n$ *be a simple nonclosed piecewise smooth path (so* γ *is one-to-one). Suppose* ω *is a continuously differentiable one-form defined on some open set* V *with* $\gamma\big([a, b]\big) \subset V$ *and* $\frac{\partial \omega_j}{\partial x_k} = \frac{\partial \omega_k}{\partial x_j}$ *for all* j *and* k. *Prove that there exists an open set* U *with* $\gamma\big([a, b]\big) \subset U \subset V$ *and a twice continuously differentiable function* $f : U \to \mathbb{R}$ *such that* $df = \omega$.
Hint 1: $\gamma\big([a, b]\big)$ *is compact.*
Hint 2: Show that you can cover the curve by finitely many balls in sequence so that the kth ball only intersects the $(k - 1)$*th ball.*
Hint 3: See previous exercise.

Exercise* 9.3.8*:

a) *Show that a connected open set* $U \subset \mathbb{R}^n$ *is path connected. Hint: Start with a point* $x \in U$, *and let* $U_x \subset U$ *is the set of points that are reachable by a path from* x. *Show that* U_x *and* $U \setminus U_x$ *are both open, and since* U_x *is nonempty (*$x \in U_x$*) it must be that* $U_x = U$.
b) *Prove the converse, that is, an open* path connected set* $U \subset \mathbb{R}^n$ *is connected. Hint: For contradiction assume there exist two open and disjoint nonempty open sets and then assume there is a piecewise smooth (and therefore continuous) path between a point in one to a point in the other.*

Exercise* 9.3.9*: *Usually path connectedness is defined using continuous paths rather than piecewise smooth paths. Prove that for open subsets of* $\mathbb{R}^n$ *the definitions are equivalent, in other words prove:*
Suppose $U \subset \mathbb{R}^n$ *is open and for every* $x, y \in U$, *there exists a continuous function* $\gamma : [a, b] \to U$ *such that* $\gamma(a) = x$ *and* $\gamma(b) = y$. *Then* U *is path connected, that is, there is a piecewise smooth path in* U *from* x *to* y.

---

*If the definition of "path connected" is as in the next exercise, "open" would not be needed for this part.

Exercise* 9.3.10** (Hard):*** *Take*

$$\omega(x, y) := \frac{-y}{x^2 + y^2}\, dx + \frac{x}{x^2 + y^2}\, dy$$

defined on $\mathbb{R}^2 \setminus \{(0,0)\}$. *Let* $\gamma \colon [a, b] \to \mathbb{R}^2 \setminus \{(0,0)\}$ *be a closed piecewise smooth path. Let* $R := \{(x, y) \in \mathbb{R}^2 : x \leq 0 \text{ and } y = 0\}$. *Suppose* $R \cap \gamma\big([a, b]\big)$ *is a finite set of* k *points. Prove that*

$$\int_\gamma \omega = 2\pi\ell$$

for some integer ℓ *with* $|\ell| \leq k$.
Hint 1: First prove that for a path β *that starts and end on* R *but does not intersect it otherwise, you find that* $\int_\beta \omega$ *is* -2π*, 0, or* 2π.
Hint 2: You proved above that $\mathbb{R}^2 \setminus R$ *is star-shaped.*
Note: The number ℓ *is called the* winding number *it measures how many times does* γ *wind around the origin in the clockwise direction.*

Chapter 10

Multivariable Integral

10.1 Riemann integral over rectangles

Note: 2–3 lectures

As in chapter 5, we define the Riemann integral using the Darboux upper and lower integrals. The ideas in this section are very similar to integration in one dimension. The complication is mostly notational. The differences between one and several dimensions will grow more pronounced in the sections following.

10.1.1 Rectangles and partitions

Definition 10.1.1. Let $(a_1, a_2, \ldots, a_n)$ and $(b_1, b_2, \ldots, b_n)$ be such that $a_k \leq b_k$ for all k. The set $[a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n]$ is called a *closed rectangle*. It is sometimes useful to allow $a_k = b_k$, and we think of $[a_k, b_k] = \{a_k\}$. If $a_k < b_k$ for all k, then the set $(a_1, b_1) \times (a_2, b_2) \times \cdots \times (a_n, b_n)$ is called an *open rectangle*.

For an open or closed rectangle $R \coloneqq [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n] \subset \mathbb{R}^n$ or $R \coloneqq (a_1, b_1) \times (a_2, b_2) \times \cdots \times (a_n, b_n) \subset \mathbb{R}^n$, we define the *n-dimensional volume* by

$$V(R) \coloneqq (b_1 - a_1)(b_2 - a_2) \cdots (b_n - a_n).$$

A *partition* P of the closed rectangle $R = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n]$ is given by partitions $P_1, P_2, \ldots, P_n$ of the intervals $[a_1, b_1], [a_2, b_2], \ldots, [a_n, b_n]$. We write $P = (P_1, P_2, \ldots, P_n)$. That is, for every $k = 1, 2, \ldots, n$ there is an integer ℓ_k and a finite set of numbers $P_k = \{x_{k,0}, x_{k,1}, x_{k,2}, \ldots, x_{k,\ell_k}\}$ such that

$$a_k = x_{k,0} < x_{k,1} < x_{k,2} < \cdots < x_{k,\ell_k - 1} < x_{k,\ell_k} = b_k.$$

Picking a set of n integers $j_1, j_2, \ldots, j_n$ where $j_k \in \{1, 2, \ldots, \ell_k\}$ we get the *subrectangle*

$$[x_{1,j_1-1}\ ,\ x_{1,j_1}] \times [x_{2,j_2-1}\ ,\ x_{2,j_2}] \times \cdots \times [x_{n,j_n-1}\ ,\ x_{n,j_n}].$$

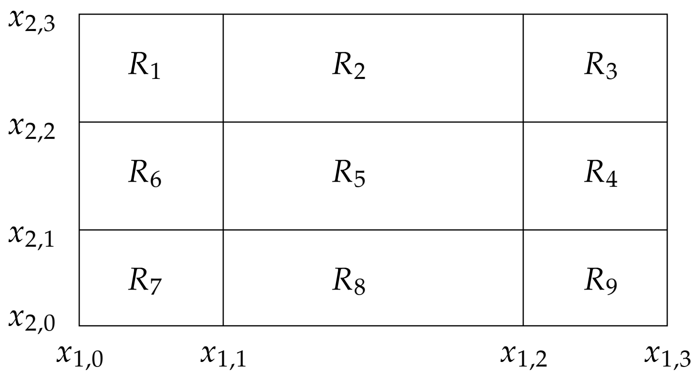

Figure 10.1: Example partition of a rectangle in $\mathbb{R}^2$. The order of the subrectangles is not important.

We order the subrectangles somehow and we say $\{R_1, R_2, \ldots, R_N\}$ are the subrectangles corresponding to the partition P of R, or more simply, subrectangles of P. In other words, we subdivided the original rectangle into many smaller subrectangles. See Figure 10.1.

Let $R \subset \mathbb{R}^n$ be a closed rectangle and let $f\colon R \to \mathbb{R}$ be a bounded function. Let P be a partition of R with N subrectangles $R_1, R_2, \ldots, R_N$. Define

$$m_i := \inf\{f(x) : x \in R_i\}, \qquad M_i := \sup\{f(x) : x \in R_i\},$$
$$L(P,f) := \sum_{i=1}^{N} m_i V(R_i), \qquad U(P,f) := \sum_{i=1}^{N} M_i V(R_i).$$

We call $L(P,f)$ the *lower Darboux sum* and $U(P,f)$ the *upper Darboux sum*.

To see the relationship to the Δ notation from the one-variable definition, note that when

$$R_i = [x_{1,j_1-1}\,,\, x_{1,j_1}] \times [x_{2,j_2-1}\,,\, x_{2,j_2}] \times \cdots \times [x_{n,j_n-1}\,,\, x_{n,j_n}],$$

then

$$V(R_i) = (x_{1,j_1} - x_{1,j_1-1})(x_{2,j_2} - x_{2,j_2-1}) \cdots (x_{n,j_n} - x_{n,j_n-1}) = \Delta x_{1,j_1} \Delta x_{2,j_2} \cdots \Delta x_{n,j_n}.$$

It is not difficult to see (left to reader) that the subrectangles of P cover our original R, and their volumes sum to that of R. That is,

$$R = \bigcup_{k=1}^{N} R_k, \qquad \text{and} \qquad V(R) = \sum_{k=1}^{N} V(R_k).$$

The indexing in the definition may be complicated, but fortunately we do not need to go back directly to the definition often.

Proposition 10.1.2. *Suppose $R \subset \mathbb{R}^n$ is a closed rectangle and $f : R \to \mathbb{R}$ is a bounded function. Let $m, M \in \mathbb{R}$ be such that for all $x \in R$, we have $m \leq f(x) \leq M$. Then for every partition P of R,*

$$m\,V(R) \leq L(P, f) \leq U(P, f) \leq M\,V(R).$$

Proof. Let P be a partition of R. For all i, we have $m \leq m_i \leq M_i \leq M$. Also $\sum_{i=1}^{N} V(R_i) = V(R)$. Therefore,

$$\begin{aligned} m\,V(R) = m\left(\sum_{i=1}^{N} V(R_i)\right) = \sum_{i=1}^{N} m\,V(R_i) \leq \sum_{i=1}^{N} m_i\,V(R_i) \leq \\ \leq \sum_{i=1}^{N} M_i\,V(R_i) \leq \sum_{i=1}^{N} M\,V(R_i) = M\left(\sum_{i=1}^{N} V(R_i)\right) = M\,V(R). \end{aligned} \qquad \square$$

10.1.2 Upper and lower integrals

By Proposition 10.1.2, the sets of upper and lower Darboux sums are bounded sets, and we can take their infima and suprema. As in one variable, we define the lower and upper integrals.

Definition 10.1.3. Let $f : R \to \mathbb{R}$ be a bounded function on a closed rectangle $R \subset \mathbb{R}^n$. Define

$$\underline{\int_R} f := \sup \left\{ L(P, f) : P \text{ a partition of } R \right\}, \qquad \overline{\int_R} f := \inf \left\{ U(P, f) : P \text{ a partition of } R \right\}.$$

We call $\underline{\int}$ the *lower Darboux integral* and $\overline{\int}$ the *upper Darboux integral*.

To refine a partition of a rectangle, we refine the underlying partitions of intervals.

Definition 10.1.4. Let $R \subset \mathbb{R}^n$ be a closed rectangle. Let $P = (P_1, P_2, \ldots, P_n)$ and $\widetilde{P} = (\widetilde{P}_1, \widetilde{P}_2, \ldots, \widetilde{P}_n)$ be partitions of R. We say $\widetilde{P}$ a *refinement* of P if, as sets, $P_k \subset \widetilde{P}_k$ for all $k = 1, 2, \ldots, n$.

If $\widetilde{P}$ is a refinement of P, then subrectangles of P are unions of subrectangles of $\widetilde{P}$. Simply put, in a refinement, we take the subrectangles of P, and we cut them into smaller subrectangles and call that $\widetilde{P}$. See Figure 10.2.

Proposition 10.1.5. *Suppose $R \subset \mathbb{R}^n$ is a closed rectangle, P is a partition of R, and $\widetilde{P}$ is a refinement of P. If $f : R \to \mathbb{R}$ is bounded, then*

$$L(P, f) \leq L(\widetilde{P}, f) \qquad \text{and} \qquad U(\widetilde{P}, f) \leq U(P, f).$$

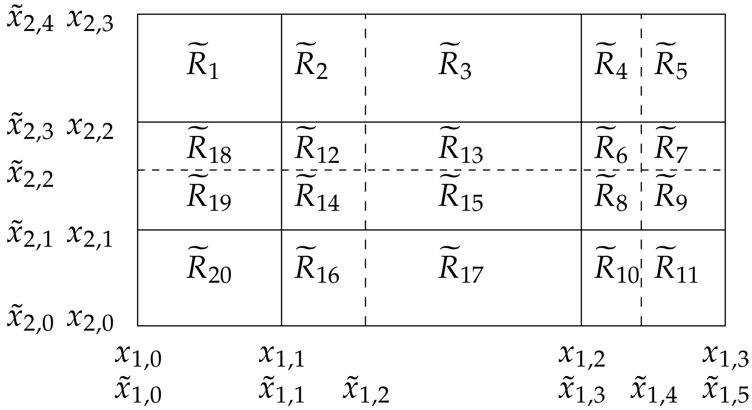

Figure 10.2: Example refinement of the partition from Figure 10.1. New "cuts" are marked in dashed lines. The exact order of the new subrectangles does not matter.

Proof. We prove the first inequality, and the second follows similarly. Let $R_1, R_2, \ldots, R_N$ be the subrectangles of P and $\widetilde{R}_1, \widetilde{R}_2, \ldots, \widetilde{R}_{\widetilde{N}}$ be the subrectangles of $\widetilde{R}$. Let I_k be the set of all indices j such that $\widetilde{R}_j \subset R_k$. For example, in figures 10.1 and 10.2, $I_4 = \{6,7,8,9\}$ as $R_4 = \widetilde{R}_6 \cup \widetilde{R}_7 \cup \widetilde{R}_8 \cup \widetilde{R}_9$. Then,

$$R_k = \bigcup_{j \in I_k} \widetilde{R}_j, \qquad V(R_k) = \sum_{j \in I_k} V(\widetilde{R}_j).$$

Let $m_j := \inf\{f(x) : x \in R_j\}$, and $\widetilde{m}_j := \inf\{f(x) :\in \widetilde{R}_j\}$ as usual. If $j \in I_k$, then $m_k \leq \widetilde{m}_j$. Then

$$L(P,f) = \sum_{k=1}^{N} m_k V(R_k) = \sum_{k=1}^{N} \sum_{j \in I_k} m_k V(\widetilde{R}_j) \leq \sum_{k=1}^{N} \sum_{j \in I_k} \widetilde{m}_j V(\widetilde{R}_j) = \sum_{j=1}^{\widetilde{N}} \widetilde{m}_j V(\widetilde{R}_j) = L(\widetilde{P}, f). \qquad \square$$

The key point of this next proposition is that the lower Darboux integral is less than or equal to the upper Darboux integral.

Proposition 10.1.6. *Let $R \subset \mathbb{R}^n$ be a closed rectangle and $f \colon R \to \mathbb{R}$ a bounded function. Let $m, M \in \mathbb{R}$ be such that for all $x \in R$, we have $m \leq f(x) \leq M$. Then*

$$m\, V(R) \leq \underline{\int_R} f \leq \overline{\int_R} f \leq M\, V(R). \tag{10.1}$$

Proof. For every partition P, via Proposition 10.1.2,

$$m\, V(R) \leq L(P,f) \leq U(P,f) \leq M\, V(R).$$

Taking supremum of $L(P,f)$ and infimum of $U(P,f)$ over all partitions P, we obtain the first and the last inequality in (10.1).

The key inequality in (10.1) is the middle one. Let $P = (P_1, P_2, \dots, P_n)$ and $Q = (Q_1, Q_2, \dots, Q_n)$ be partitions of R. Define $\widetilde{P} = (\widetilde{P}_1, \widetilde{P}_2, \dots, \widetilde{P}_n)$ by letting $\widetilde{P}_k := P_k \cup Q_k$ for every k. Then $\widetilde{P}$ is a partition of R, and $\widetilde{P}$ is a refinement of P and also a refinement of Q. By Proposition 10.1.5, $L(P, f) \leq L(\widetilde{P}, f)$ and $U(\widetilde{P}, f) \leq U(Q, f)$. Therefore,

$$L(P, f) \leq L(\widetilde{P}, f) \leq U(\widetilde{P}, f) \leq U(Q, f).$$

In other words, for two arbitrary partitions P and Q, we have $L(P, f) \leq U(Q, f)$. Via Proposition 1.2.7 from volume I, we obtain

$$\sup \{L(P, f) : P \text{ a partition of } R\} \leq \inf \{U(P, f) : P \text{ a partition of } R\}.$$

In other words, $\underline{\int_R} f \leq \overline{\int_R} f$. □

10.1.3 The Riemann integral

We have all we need to define the Riemann integral in n-dimensions over rectangles. As in one dimension, the Riemann integral is only defined on a certain class of functions, called the Riemann integrable functions.

Definition 10.1.7. Let $R \subset \mathbb{R}^n$ be a closed rectangle and $f\colon R \to \mathbb{R}$ a bounded function such that

$$\underline{\int_R} f(x)\, dx = \overline{\int_R} f(x)\, dx.$$

Then f is said to be *Riemann integrable*, and we sometimes say simply *integrable*. We denote the set of Riemann integrable functions on R by $\mathcal{R}(R)$. For $f \in \mathcal{R}(R)$ define the *Riemann integral*

$$\int_R f := \underline{\int_R} f = \overline{\int_R} f.$$

When the variable $x \in \mathbb{R}^n$ needs to be emphasized, we write

$$\int_R f(x)\, dx, \qquad \int_R f(x_1, \dots, x_n)\, dx_1 \cdots dx_n, \qquad \text{or} \qquad \int_R f(x)\, dV.$$

If $R \subset \mathbb{R}^2$, then we usually say area instead of volume, and we write

$$\int_R f(x)\, dA.$$

Proposition 10.1.6 immediately implies the following proposition.

Proposition 10.1.8. *Let $f\colon R \to \mathbb{R}$ be a Riemann integrable function on a closed rectangle $R \subset \mathbb{R}^n$. Let $m, M \in \mathbb{R}$ be such that $m \leq f(x) \leq M$ for all $x \in R$. Then*

$$m\, V(R) \leq \int_R f \leq M\, V(R).$$

Example 10.1.9: A constant function is Riemann integrable. Proof: Suppose $f(x) = c$ for all $x \in R$. Then

$$c\,V(R) \leq \underline{\int_R} f \leq \overline{\int_R} f \leq c\,V(R).$$

So f is integrable, and furthermore $\int_R f = c\,V(R)$.

The proofs of linearity and monotonicity are almost completely identical to the proofs from one variable. We leave the next two propositions as exercises.

Proposition 10.1.10 (Linearity). *Let $R \subset \mathbb{R}^n$ be a closed rectangle and let f and g be in $\mathcal{R}(R)$ and $\alpha \in \mathbb{R}$.*

(*i*) *αf is in $\mathcal{R}(R)$ and*

$$\int_R \alpha f = \alpha \int_R f.$$

(*ii*) *$f + g$ is in $\mathcal{R}(R)$ and*

$$\int_R (f+g) = \int_R f + \int_R g.$$

Proposition 10.1.11 (Monotonicity). *Let $R \subset \mathbb{R}^n$ be a closed rectangle, let f and g be in $\mathcal{R}(R)$, and suppose $f(x) \leq g(x)$ for all $x \in R$. Then*

$$\int_R f \leq \int_R g.$$

Checking for integrability using the definition often involves the following technique, as in the single variable case.

Proposition 10.1.12. *Let $R \subset \mathbb{R}^n$ be a closed rectangle and $f \colon R \to \mathbb{R}$ a bounded function. Then $f \in \mathcal{R}(R)$ if and only if for every $\epsilon > 0$, there exists a partition P of R such that*

$$U(P,f) - L(P,f) < \epsilon.$$

Proof. First, if f is integrable, then the supremum of $L(P,f)$ and infimum of $U(Q,f)$ over all partitions P and Q are equal and hence the infimum of $U(P,f) - L(Q,f)$ is zero. Taking a common refinement $\widetilde{P}$ of P and Q we find $U(\widetilde{P},f) - L(\widetilde{P},f) \leq U(P,f) - L(Q,f)$. Hence the infimum of $U(P,f) - L(P,f)$ over all partitions P is zero, and so for every $\epsilon > 0$, there must be some partition P such that $U(P,f) - L(P,f) < \epsilon$.

For the other direction, given an $\epsilon > 0$ find P such that $U(P,f) - L(P,f) < \epsilon$.

$$\overline{\int_R} f - \underline{\int_R} f \leq U(P,f) - L(P,f) < \epsilon.$$

As $\overline{\int_R} f \geq \underline{\int_R} f$ and the above holds for every $\epsilon > 0$, we conclude $\overline{\int_R} f = \underline{\int_R} f$ and $f \in \mathcal{R}(R)$. ◻

Suppose $f \colon S \to \mathbb{R}$ is a function and $R \subset S$ is a closed rectangle. If the restriction $f|_R$ is integrable, then for simplicity we say f is *integrable on* R, or $f \in \mathscr{R}(R)$, and we write

$$\int_R f := \int_R f|_R .$$

Proposition 10.1.13. *Let $S \subset \mathbb{R}^n$ be a closed rectangle. If $f \colon S \to \mathbb{R}$ is integrable and $R \subset S$ is a closed rectangle, then f is integrable on R.*

Proof. Given $\epsilon > 0$, find a partition $P = (P_1, \ldots, P_n)$ of S such that $U(P, f) - L(P, f) < \epsilon$. By making a refinement of P if necessary, assume that the endpoints of R are in P. That is, if $R = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n]$, then $a_i, b_i \in P_i$. Let $\widetilde{P} = (\widetilde{P}_1, \ldots, \widetilde{P}_n)$ be the partition of R given by $\widetilde{P}_i = P_i \cap [a_i, b_i]$. Subrectangles of $\widetilde{P}$ are subrectangles of P, that is, R is a union of subrectangles of P. Divide the subrectangles of P into two collections: Let $R_1, R_2 \ldots, R_K$ be the subrectangles of P that are also subrectangles of $\widetilde{P}$ and let $R_{K+1}, \ldots, R_N$ be the rest. See Figure 10.3. Let m_k and M_k be the infimum and supremum of f on R_k as usual. Then,

$$\begin{aligned}
\epsilon > U(P, f) - L(P, f) &= \sum_{k=1}^{K} (M_k - m_k) V(R_k) + \sum_{k=K+1}^{N} (M_k - m_k) V(R_k) \\
&\geq \sum_{k=1}^{K} (M_k - m_k) V(R_k) = U(\widetilde{P}, f|_R) - L(\widetilde{P}, f|_R).
\end{aligned}$$

Therefore, $f|_R$ is integrable. □

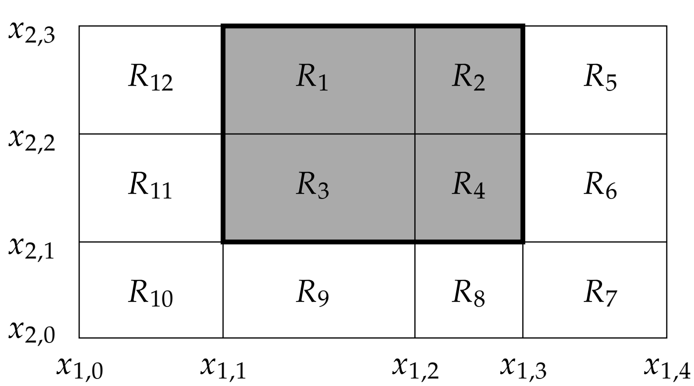

Figure 10.3: A partition of a large rectangle S, that also gives a partition of a smaller rectangle (shaded and outlined) $R \subset S$. The subrectangles R_1, R_2, R_3, R_4 are the subrectangles of $\widetilde{P} = \bigl(\{x_{1,1}, x_{1,2}, x_{1,3}\}, \{x_{2,1}, x_{2,2}, x_{2,3}\}\bigr)$.

10.1.4 Integrals of continuous functions

Although we will prove a more general result later, it is useful to start with integrability of continuous functions. To do so, we wish to measure the fineness of partitions. In one variable, we measure the length of a subinterval. In several variables, we measure the sides of a subrectangle. We say a rectangle $R = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n]$ has *longest side at most* α if $b_k - a_k \leq \alpha$ for all $k = 1, 2, \ldots, n$.

Proposition 10.1.14. *If a rectangle $R \subset \mathbb{R}^n$ has longest side at most α, then for all $x, y \in R$,*

$$\|x - y\| \leq \sqrt{n}\, \alpha.$$

Proof.

$$\begin{aligned} \|x - y\| &= \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2 + \cdots + (x_n - y_n)^2} \\ &\leq \sqrt{(b_1 - a_1)^2 + (b_2 - a_2)^2 + \cdots + (b_n - a_n)^2} \\ &\leq \sqrt{\alpha^2 + \alpha^2 + \cdots + \alpha^2} = \sqrt{n}\, \alpha. \end{aligned}$$

□

Theorem 10.1.15. *Let $R \subset \mathbb{R}^n$ be a closed rectangle. If $f \colon R \to \mathbb{R}$ is continuous, then $f \in \mathcal{R}(R)$.*

Proof. The proof is analogous to the one-variable proof with some complications. The set R is a closed and bounded subset of $\mathbb{R}^n$, and hence compact. So f is uniformly continuous by Theorem 7.5.11 from volume I. Let $\epsilon > 0$ be given. Find a $\delta > 0$ such that $\|x - y\| < \delta$ implies $|f(x) - f(y)| < \frac{\epsilon}{V(R)}$.

Let P be a partition of R, where the longest side of every subrectangle is strictly less than $\frac{\delta}{\sqrt{n}}$. If $x, y \in R_k$ for a subrectangle R_k of P, then, by the proposition, $\|x - y\| < \sqrt{n} \frac{\delta}{\sqrt{n}} = \delta$. Therefore,

$$f(x) - f(y) \leq |f(x) - f(y)| < \frac{\epsilon}{V(R)}.$$

As f is continuous on R_k, which is compact, f attains a maximum and a minimum on this subrectangle. Let x be a point where f attains the maximum and y be a point where f attains the minimum. Then $f(x) = M_k$ and $f(y) = m_k$ in the notation from the definition of the integral. Thus,

$$M_k - m_k = f(x) - f(y) < \frac{\epsilon}{V(R)}.$$

And so

$$\begin{aligned} U(P, f) - L(P, f) &= \left(\sum_{k=1}^{N} M_k V(R_k) \right) - \left(\sum_{k=1}^{N} m_k V(R_k) \right) \\ &= \sum_{k=1}^{N} (M_k - m_k) V(R_k) \\ &< \frac{\epsilon}{V(R)} \sum_{k=1}^{N} V(R_k) = \epsilon. \end{aligned}$$

Proposition 10.1.12 then says that $f \in \mathcal{R}(R)$. □

10.1.5 Integration of functions with compact support

Let $U \subset \mathbb{R}^n$ be an open set and $f : U \to \mathbb{R}$ be a function. The *support* of f is the set

$$\operatorname{supp}(f) := \overline{\{x \in U : f(x) \neq 0\}},$$

where the closure is with respect to the subspace topology on U. Taking the closure with respect to the subspace topology is the same as $\overline{\{x \in U : f(x) \neq 0\}} \cap U$, where the closure is with respect to the ambient euclidean space $\mathbb{R}^n$. In particular, $\operatorname{supp}(f) \subset U$. The support is the closure (in U) of the set of points where the function is nonzero. Its complement in U is open. If $x \in U$ and x is not in the support of f, then f is constantly zero in a whole neighborhood of x.

A function f is said to have *compact support* if $\operatorname{supp}(f)$ is a compact set.

Example 10.1.16: The function $f : \mathbb{R}^2 \to \mathbb{R}$, defined by

$$f(x,y) := \begin{cases} -x(x^2+y^2-1)^2 & \text{if } \sqrt{x^2+y^2} \leq 1, \\ 0 & \text{else,} \end{cases}$$

is continuous and its support is the closed unit disc $C(0,1) = \bigl\{(x,y) : \sqrt{x^2+y^2} \leq 1\bigr\}$, which is a compact set, so f has compact support. Note that the function is zero on the entire y-axis and on the unit circle, but all such points that lie in the closed unit disc are still within the support as they are in the closure of points where f is nonzero. See Figure 10.4.

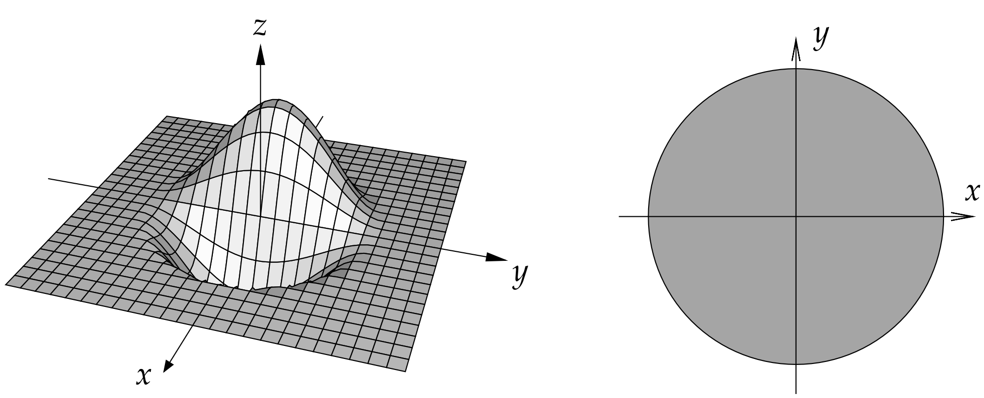

Figure 10.4: Function with compact support (left), the support is the closed unit disc (right).

If $U \neq \mathbb{R}^n$, then you must be careful to take the closure in U. Consider the following two examples.

Example 10.1.17: Let $B(0,1) \subset \mathbb{R}^2$ be the unit disc. The function $f : B(0,1) \to \mathbb{R}$, defined by

$$f(x,y) := \begin{cases} 0 & \text{if } \sqrt{x^2+y^2} > 1/2, \\ 1/2 - \sqrt{x^2+y^2} & \text{if } \sqrt{x^2+y^2} \leq 1/2, \end{cases}$$

is continuous on $B(0,1)$ and its support is the smaller closed ball $C(0,1/2)$. As that is a compact set, f has compact support.

The function $g\colon B(0,1) \to \mathbb{R}$, defined by

$$g(x,y) := \begin{cases} 0 & \text{if } x \leq 0, \\ x & \text{if } x > 0, \end{cases}$$

is continuous on $B(0,1)$, but its support is the set $\{(x,y) \in B(0,1) : x \geq 0\}$. In particular, g is not compactly supported.

We really only need to consider the case when $U = \mathbb{R}^n$. In light of Exercise 10.1.1, which says every continuous function on an open $U \subset \mathbb{R}^n$ with compact support can be extended to a continuous function with compact support on $\mathbb{R}^n$, considering $U = \mathbb{R}^n$ is not an oversimplification.

Example 10.1.18: The continuous function $f\colon B(0,1) \to \mathbb{R}$ given by $f(x,y) := \sin\bigl(\frac{1}{1-x^2-y^2}\bigr)$ does not have compact support; as f is not constantly zero on any neighborhood of every point in $B(0,1)$, the support is the entire disc $B(0,1)$. The function does not extend as above to a continuous function on $\mathbb{R}^2$. In fact, it is not difficult to show that f cannot be extended in any way whatsoever to be continuous on all of $\mathbb{R}^2$ (the boundary of the disc is the problem).

Proposition 10.1.19. *Suppose $f\colon \mathbb{R}^n \to \mathbb{R}$ is a continuous function with compact support. If R and S are closed rectangles such that $\operatorname{supp}(f) \subset R$ and $\operatorname{supp}(f) \subset S$, then*

$$\int_S f = \int_R f.$$

Proof. As f is continuous, it is automatically integrable on the rectangles R, S, and $R \cap S$. Then Exercise 10.1.7 says $\int_S f = \int_{S\cap R} f = \int_R f$. □

Because of this proposition, when $f\colon \mathbb{R}^n \to \mathbb{R}$ has compact support and is integrable on a rectangle R containing the support, we write

$$\int f := \int_R f \qquad \text{or} \qquad \int_{\mathbb{R}^n} f := \int_R f.$$

For example, if f is continuous and of compact support, then $\int_{\mathbb{R}^n} f$ exists.

10.1.6 Exercises

***Exercise* 10.1.1:** *Suppose $U \subset \mathbb{R}^n$ is open and $f\colon U \to \mathbb{R}$ is continuous and of compact support. Show that the function $\widetilde{f}\colon \mathbb{R}^n \to \mathbb{R}$, where*

$$\widetilde{f}(x) := \begin{cases} f(x) & \text{if } x \in U, \\ 0 & \text{otherwise,} \end{cases}$$

is continuous.

Exercise* 10.1.2*: *Prove Proposition 10.1.10.*

Exercise* 10.1.3*: *Suppose R is a closed rectangle with the length of one of the sides equal to 0. For every bounded function $f\colon R \to \mathbb{R}$, show that $f \in \mathcal{R}(R)$ and $\int_R f = 0$.*

Exercise* 10.1.4*: *Suppose R is a closed rectangle with the length of one of the sides equal to 0, and suppose S is a closed rectangle with $R \subset S$. If $f\colon S \to \mathbb{R}$ is a bounded function such that $f(x) = 0$ for $x \in S \setminus R$, show that $f \in \mathcal{R}(S)$ and $\int_S f = 0$.*

Exercise* 10.1.5*: *Suppose $f\colon \mathbb{R}^n \to \mathbb{R}$ is such that $f(x) := 0$ if $x \neq 0$ and $f(0) := 1$. Show that f is integrable on $R := [-1,1] \times [-1,1] \times \cdots \times [-1,1]$ directly using the definition, and find $\int_R f$.*

Exercise* 10.1.6*: *Suppose R is a closed rectangle and $h\colon R \to \mathbb{R}$ is a bounded function such that $h(x) = 0$ if $x \notin \partial R$ (the boundary of R). Let S be a closed rectangle. Show that $h \in \mathcal{R}(S)$ and*

$$\int_S h = 0.$$

Hint: Write h as a sum of functions as in Exercise 10.1.4.

Exercise* 10.1.7*: *Suppose R and R' are two closed rectangles with $R' \subset R$. Suppose $f\colon R \to \mathbb{R}$ is in $\mathcal{R}(R')$ and $f(x) = 0$ for $x \in R \setminus R'$. Show that $f \in \mathcal{R}(R)$ and*

$$\int_{R'} f = \int_R f.$$

Do this in the following steps.

a) First do the proof assuming that furthermore $f(x) = 0$ whenever $x \in \overline{R \setminus R'}$.

b) Write $f(x) = g(x) + h(x)$ where $g(x) = 0$ whenever $x \in \overline{R \setminus R'}$, and $h(x)$ is zero except perhaps on $\partial R'$. Then show $\int_R h = \int_{R'} h = 0$ (see Exercise 10.1.6).

c) Show $\int_{R'} f = \int_R f$.

Exercise* 10.1.8*: *Suppose $R' \subset \mathbb{R}^n$ and $R'' \subset \mathbb{R}^n$ are two rectangles such that $R = R' \cup R''$ is a rectangle, and $R' \cap R''$ is rectangle with one of the sides having length 0 (that is, $V(R' \cap R'') = 0$). Let $f\colon R \to \mathbb{R}$ be a function such that $f \in \mathcal{R}(R')$ and $f \in \mathcal{R}(R'')$. Show that $f \in \mathcal{R}(R)$ and*

$$\int_R f = \int_{R'} f + \int_{R''} f.$$

Hint: See previous exercise.

Exercise* 10.1.9*: *Prove a stronger version of Proposition 10.1.19. Suppose $f\colon \mathbb{R}^n \to \mathbb{R}$ is a function with compact support but not necessarily continuous. Prove that if R is a closed rectangle such that $\operatorname{supp}(f) \subset R$ and f is integrable on R, then for every other closed rectangle S with $\operatorname{supp}(f) \subset S$, the function f is integrable on S and $\int_S f = \int_R f$. Hint: See Exercise 10.1.7.*

Exercise* 10.1.10*: *Suppose R and S are closed rectangles of $\mathbb{R}^n$. Define $f\colon \mathbb{R}^n \to \mathbb{R}$ as $f(x) := 1$ if $x \in R$, and $f(x) := 0$ otherwise. Prove f is integrable on S and compute $\int_S f$. Hint: Consider $S \cap R$.*

Exercise* 10.1.11*: *Let* $R := [0,1] \times [0,1] \subset \mathbb{R}^2$.

a) *Suppose* $f : R \to \mathbb{R}$ *is defined by*
$$f(x,y) := \begin{cases} 1 & \textit{if } x = y, \\ 0 & \textit{else.} \end{cases}$$
Show that $f \in \mathcal{R}(R)$ *and compute* $\int_R f$.

b) *Suppose* $f : R \to \mathbb{R}$ *is defined by*
$$f(x,y) := \begin{cases} 1 & \textit{if } x \in \mathbb{Q} \textit{ or } y \in \mathbb{Q}, \\ 0 & \textit{else.} \end{cases}$$
Show that $f \notin \mathcal{R}(R)$.

Exercise* 10.1.12*: *Suppose R is a closed rectangle, and suppose S_j are closed rectangles such that $S_j \subset R$ and $S_j \subset S_{j+1}$ for all j. Suppose $f : R \to \mathbb{R}$ is bounded and $f \in \mathcal{R}(S_j)$ for all j. Show that $f \in \mathcal{R}(R)$ and*
$$\lim_{j\to\infty} \int_{S_j} f = \int_R f.$$

Exercise* 10.1.13*: *Suppose $f : [-1,1] \times [-1,1] \to \mathbb{R}$ is a Riemann integrable function such $f(x) = -f(-x)$. Using the definition prove*
$$\int_{[-1,1]\times[-1,1]} f = 0.$$

10.2 Iterated integrals and Fubini theorem

Note: 1–2 lectures

The Riemann integral in several variables is hard to compute via the definition. For one-dimensional Riemann integral, we have the fundamental theorem of calculus, which allows computing many integrals without having to appeal to the definition of the integral. We will rewrite a Riemann integral in several variables into several one-dimensional Riemann integrals by iterating. However, if $f \colon [0,1]^2 \to \mathbb{R}$ is a Riemann integrable function, it is not immediately clear if the three expressions

$$\int_{[0,1]^2} f, \qquad \int_0^1 \int_0^1 f(x,y)\,dx\,dy, \qquad \text{and} \qquad \int_0^1 \int_0^1 f(x,y)\,dy\,dx$$

are equal, or if the last two are even well-defined.

Example 10.2.1: Define

$$f(x,y) := \begin{cases} 1 & \text{if } x = 1/2 \text{ and } y \in \mathbb{Q}, \\ 0 & \text{otherwise.} \end{cases}$$

Then f is Riemann integrable on $R := [0,1]^2$ and $\int_R f = 0$. Moreover, $\int_0^1 \int_0^1 f(x,y)\,dx\,dy = 0$. However,

$$\int_0^1 f(1/2, y)\,dy$$

does not exist, so strictly speaking, $\int_0^1 \int_0^1 f(x,y)\,dy\,dx$ does not make sense. See Figure 10.5.

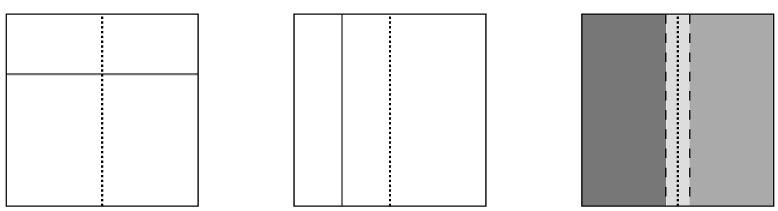

Figure 10.5: Left: $[0,1]^2$ with the line $x = 1/2$ marked dotted and $\int_0^1 f(x,y)\,dx$ marked as gray solid line for a generic y. Center: Similar picture but $\int_0^1 f(x,y)\,dy$ marked for some $x \neq 1/2$. Right: The three different rectangles in the partition used to integrate f in different grays.

Proof: We start with integrability of f. Consider the partition of $[0,1]^2$ where the partition in the x direction is $\{0, 1/2 - \epsilon, 1/2 + \epsilon, 1\}$ and in the y direction $\{0,1\}$. The corresponding subrectangles are

$$R_1 := [0, 1/2 - \epsilon] \times [0,1], \qquad R_2 := [1/2 - \epsilon, 1/2 + \epsilon] \times [0,1], \qquad R_3 := [1/2 + \epsilon, 1] \times [0,1].$$

We have $m_1 = M_1 = 0$, $m_2 = 0$, $M_2 = 1$, and $m_3 = M_3 = 0$. Therefore,

$$L(P, f) = m_1 V(R_1) + m_2 V(R_2) + m_3 V(R_3) = 0(1/2 - \epsilon) + 0(2\epsilon) + 0(1/2 - \epsilon) = 0,$$

and

$$U(P, f) = M_1 V(R_1) + M_2 V(R_2) + M_3 V(R_3) = 0(1/2 - \epsilon) + 1(2\epsilon) + 0(1/2 - \epsilon) = 2\epsilon.$$

The upper and lower sums are arbitrarily close and the lower sum is always zero, so the function is integrable and $\int_R f = 0$.

For every fixed y, the function that takes x to $f(x, y)$ is zero except perhaps at a single point $x = 1/2$. Such a function is integrable and $\int_0^1 f(x, y)\,dx = 0$. Therefore, $\int_0^1 \int_0^1 f(x, y)\,dx\,dy = 0$. However, if $x = 1/2$, the function that takes y to $f(1/2, y)$ is the nonintegrable function that is 1 on the rationals and 0 on the irrationals. See Example 5.1.4 from volume I.

We solve this problem of undefined inside integrals by using the upper and lower integrals, which are always defined for any bounded function.

Split the coordinates of $\mathbb{R}^{n+m}$ into two parts: Write the coordinates on $\mathbb{R}^{n+m} = \mathbb{R}^n \times \mathbb{R}^m$ as (x, y) where $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$. For a function $f(x, y)$, write

$$f_x(y) := f(x, y)$$

when x is fixed and we want a function of y. Write

$$f^y(x) := f(x, y)$$

when y is fixed and we want a function of x.

Theorem 10.2.2 (Fubini version A*). *Let $R \times S \subset \mathbb{R}^n \times \mathbb{R}^m$ be a closed rectangle and $f : R \times S \to \mathbb{R}$ be integrable. The functions $g : R \to \mathbb{R}$ and $h : R \to \mathbb{R}$ defined by*

$$g(x) := \underline{\int_S} f_x \qquad \textit{and} \qquad h(x) := \overline{\int_S} f_x$$

are integrable on R and

$$\int_R g = \int_R h = \int_{R \times S} f.$$

In other words,

$$\int_{R \times S} f = \int_R \left(\underline{\int_S} f(x, y)\,dy \right) dx = \int_R \left(\overline{\int_S} f(x, y)\,dy \right) dx.$$

If f_x is integrable for all x, for example when f is continuous, we obtain the more familiar

$$\int_{R \times S} f = \int_R \int_S f(x, y)\,dy\,dx.$$

*Named after the Italian mathematician Guido Fubini (1879–1943).

Proof. A partition of $R \times S$ is a concatenation of a partition of R and a partition of S. That is, write a partition of $R \times S$ as $(P, P') = (P_1, P_2, \ldots, P_n, P'_1, P'_2, \ldots, P'_m)$, where $P = (P_1, P_2, \ldots, P_n)$ and $P' = (P'_1, P'_2, \ldots, P'_m)$ are partitions of R and S respectively. Let $R_1, R_2, \ldots, R_N$ be the subrectangles of P and $R'_1, R'_2, \ldots, R'_K$ be the subrectangles of P'. The subrectangles of (P, P') are $R_i \times R'_j$ where $1 \leq i \leq N$ and $1 \leq j \leq K$.

Let

$$m_{i,j} := \inf_{(x,y)\in R_i \times R'_j} f(x, y).$$

Notice that $V(R_i \times R'_j) = V(R_i)V(R'_j)$ and hence

$$L\big((P, P'), f\big) = \sum_{i=1}^{N} \sum_{j=1}^{K} m_{i,j}\, V(R_i \times R'_j) = \sum_{i=1}^{N} \left(\sum_{j=1}^{K} m_{i,j}\, V(R'_j) \right) V(R_i).$$

Define

$$m_j(x) := \inf_{y\in R'_j} f(x, y) = \inf_{y\in R'_j} f_x(y).$$

For $x \in R_i$, we have $m_{i,j} \leq m_j(x)$, and therefore,

$$\sum_{j=1}^{K} m_{i,j}\, V(R'_j) \leq \sum_{j=1}^{K} m_j(x)\, V(R'_j) = L(P', f_x) \leq \underline{\int_S} f_x = g(x).$$

The inequality holds for all $x \in R_i$, and so

$$\sum_{j=1}^{K} m_{i,j}\, V(R'_j) \leq \inf_{x\in R_i} g(x).$$

We obtain

$$L\big((P, P'), f\big) \leq \sum_{j=1}^{N} \left(\inf_{x\in R_j} g(x) \right) V(R_j) = L(P, g).$$

Similarly, $U\big((P, P'), f\big) \geq U(P, h)$, and the proof of this inequality is left as an exercise. Putting the two inequalities together with the fact that $g(x) \leq h(x)$ for all x,

$$L\big((P, P'), f\big) \leq L(P, g) \leq U(P, g) \leq U(P, h) \leq U\big((P, P'), f\big).$$

Since f is integrable, it must be that g is integrable as

$$U(P, g) - L(P, g) \leq U\big((P, P'), f\big) - L\big((P, P'), f\big),$$

and we can make the right-hand side arbitrarily small. As for any partition we have $L\big((P, P'), f\big) \leq L(P, g) \leq U\big((P, P'), f\big)$, we have $\int_R g = \int_{R\times S} f$.

Likewise,

$$L\big((P, P'), f\big) \leq L(P, g) \leq L(P, h) \leq U(P, h) \leq U\big((P, P'), f\big),$$

and hence

$$U(P,h) - L(P,h) \leq U\big((P,P'),f\big) - L\big((P,P'),f\big).$$

As f is integrable, so is h. Moreover, $L\big((P,P'),f\big) \leq L(P,h) \leq U\big((P,P'),f\big)$ implies $\int_R h = \int_{R\times S} f$. □

We can also do the iterated integration in the opposite order. The proof of this version is almost identical to version A (or follows quickly from version A). We leave it as an exercise.

Theorem 10.2.3 (Fubini version B)**.** *Let $R \times S \subset \mathbb{R}^n \times \mathbb{R}^m$ be a closed rectangle and $f \colon R \times S \to \mathbb{R}$ be integrable. The functions $g \colon S \to \mathbb{R}$ and $h \colon S \to \mathbb{R}$ defined by*

$$g(y) := \underline{\int_R} f^y \qquad \text{and} \qquad h(y) := \overline{\int_R} f^y$$

are integrable on S and

$$\int_S g = \int_S h = \int_{R\times S} f.$$

That is,

$$\int_{R\times S} f = \int_S \left(\underline{\int_R} f(x,y)\,dx \right) dy = \int_S \left(\overline{\int_R} f(x,y)\,dx \right) dy.$$

Next suppose f_x and f^y are integrable. For example, suppose f is continuous. By putting the two versions together we obtain the familiar

$$\int_{R\times S} f = \int_R \int_S f(x,y)\,dy\,dx = \int_S \int_R f(x,y)\,dx\,dy.$$

Often the Fubini theorem is stated in two dimensions for a continuous function $f \colon R \to \mathbb{R}$ on a rectangle $R = [a,b] \times [c,d]$. Then the Fubini theorem states that

$$\int_R f = \int_a^b \int_c^d f(x,y)\,dy\,dx = \int_c^d \int_a^b f(x,y)\,dx\,dy.$$

The Fubini theorem is commonly thought of as *the* theorem that allows us to swap the order of iterated integrals, although there are many variations on Fubini, and we have seen but two of them.

Repeatedly applying Fubini theorem gets us the following corollary: Let $R := [a_1,b_1] \times [a_2,b_2] \times \cdots \times [a_n,b_n] \subset \mathbb{R}^n$ be a closed rectangle and let $f \colon R \to \mathbb{R}$ be continuous. Then

$$\int_R f = \int_{a_1}^{b_1} \int_{a_2}^{b_2} \cdots \int_{a_n}^{b_n} f(x_1,x_2,\ldots,x_n)\,dx_n\,dx_{n-1} \cdots dx_1.$$

We may switch the order of integration to any order we please. We may relax the continuity requirement by making sure that all the intermediate functions are integrable, or by using upper or lower integrals appropriately.

10.2.1 Exercises

***Exercise* 10.2.1:** *Compute $\int_0^1 \int_{-1}^1 xe^{xy}\,dx\,dy$ in a simple way.*

***Exercise* 10.2.2:** *Prove the assertion $U\big((P,P'),f\big) \geq U(P,h)$ from the proof of Theorem 10.2.2.*

***Exercise* 10.2.3** (Easy)**:** *Prove Theorem 10.2.3.*

***Exercise* 10.2.4:** *Let $R := [a,b]\times[c,d]$ and $f(x,y)$ is an integrable function on R such that for every fixed y, the function that takes x to $f(x,y)$ is zero except at finitely many points. Show*

$$\int_R f = 0.$$

***Exercise* 10.2.5:** *Let $R := [a,b]\times[c,d]$ and $f(x,y) := g(x)h(y)$ for continuous functions $g\colon [a,b]\to\mathbb{R}$ and $h\colon [c,d]\to\mathbb{R}$. Prove*

$$\int_R f = \left(\int_a^b g\right)\left(\int_c^d h\right).$$

***Exercise* 10.2.6:** *Compute (using calculus)*

$$\int_0^1\int_0^1 \frac{x^2-y^2}{(x^2+y^2)^2}\,dx\,dy \qquad \text{and} \qquad \int_0^1\int_0^1 \frac{x^2-y^2}{(x^2+y^2)^2}\,dy\,dx.$$

You will need to interpret the integrals as improper, that is, the limit of $\int_\epsilon^1$ as $\epsilon\to 0^+$.

***Exercise* 10.2.7:** *Suppose $f(x,y) := g(x)$ where $g\colon [a,b]\to\mathbb{R}$ is Riemann integrable. Show that f is Riemann integrable for every $R = [a,b]\times[c,d]$ and*

$$\int_R f = (d-c)\int_a^b g.$$

***Exercise* 10.2.8:** *Define $f\colon [-1,1]\times[0,1]\to\mathbb{R}$ by*

$$f(x,y) := \begin{cases} x & \text{if } y \in \mathbb{Q}, \\ 0 & \text{else.}\end{cases}$$

a) *Show $\int_0^1\int_{-1}^1 f(x,y)\,dx\,dy$ exists, but $\int_{-1}^1\int_0^1 f(x,y)\,dy\,dx$ does not.*

b) *Compute $\int_{-1}^1 \overline{\int_0^1} f(x,y)\,dy\,dx$ and $\int_{-1}^1 \underline{\int_0^1} f(x,y)\,dy\,dx$.*

c) *Show f is not Riemann integrable on $[-1,1]\times[0,1]$ (use Fubini).*

***Exercise* 10.2.9:** *Define $f\colon [0,1]\times[0,1]\to\mathbb{R}$ by*

$$f(x,y) := \begin{cases} 1/q & \text{if } x\in\mathbb{Q},\ y\in\mathbb{Q}, \text{ and } y = p/q \text{ in lowest terms,} \\ 0 & \text{else.}\end{cases}$$

a) *Show f is Riemann integrable on $[0,1]\times[0,1]$.*

b) *Find $\overline{\int_0^1} f(x,y)\,dx$ and $\underline{\int_0^1} f(x,y)\,dx$ for all $y\in[0,1]$, and show they are unequal for all $y\in\mathbb{Q}$.*

c) *Show $\int_0^1\int_0^1 f(x,y)\,dy\,dx$ exists, but $\int_0^1\int_0^1 f(x,y)\,dx\,dy$ does not.*

Note: By Fubini, $\int_0^1\overline{\int_0^1} f(x,y)\,dy\,dx$ and $\int_0^1\underline{\int_0^1} f(x,y)\,dy\,dx$ do exist and equal the integral of f on R.

10.3 Outer measure and null sets

Note: 2 lectures

10.3.1 Outer measure and null sets

Before we characterize all Riemann integrable functions, we need to make a slight detour. We introduce a way of measuring the size of sets in $\mathbb{R}^n$.

Definition 10.3.1. Define the *outer measure* of a set $S \subset \mathbb{R}^n$ as

$$m^*(S) := \inf \sum_{j=1}^{\infty} V(R_j),$$

where the infimum is taken over all sequences $\{R_j\}_{j=1}^{\infty}$ of open rectangles such that $S \subset \bigcup_{j=1}^{\infty} R_j$, and we are allowing both the sum and the infimum to be ∞. See Figure 10.6. In particular, S is of *measure zero* or a *null set* if $m^*(S) = 0$.

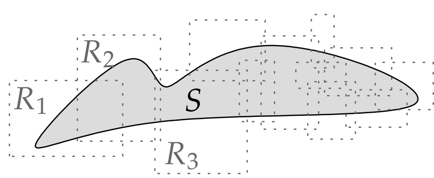

Figure 10.6: Outer measure construction, in this case $S \subset R_1 \cup R_2 \cup R_3 \cup \cdots$, so $m^*(S) \leq V(R_1) + V(R_2) + V(R_3) + \cdots$.

An immediate consequence (Exercise 10.3.2) of the definition is that if $A \subset B$, then $m^*(A) \leq m^*(B)$. It is also not difficult to show (Exercise 10.3.13) that we obtain the same number $m^*(S)$ if we also allow both finite and infinite sequences of rectangles in the definition. It is not enough, however, to allow only finite sequences.

The theory of measures on $\mathbb{R}^n$ is a very complicated subject. We will only require measure-zero sets and so we focus on these. A set S is of measure zero if for every $\epsilon > 0$, there exists a sequence of open rectangles $\{R_j\}_{j=1}^{\infty}$ such that

$$S \subset \bigcup_{j=1}^{\infty} R_j \qquad \text{and} \qquad \sum_{j=1}^{\infty} V(R_j) < \epsilon. \tag{10.2}$$

If S is of measure zero and $S' \subset S$, then S' is of measure zero. We can use the same exact rectangles.

It is sometimes more convenient to use balls instead of rectangles. Furthermore, we can choose balls no bigger than a fixed radius.

Proposition 10.3.2. *Let $\delta > 0$ be given. A set $S \subset \mathbb{R}^n$ is of measure zero if and only if for every $\epsilon > 0$, there exists a sequence of open balls $\{B_k\}_{k=1}^\infty$, where the radius of B_k is $r_k < \delta$, and such that*

$$S \subset \bigcup_{k=1}^\infty B_k \qquad \textit{and} \qquad \sum_{k=1}^\infty r_k^n < \epsilon.$$

Note that the "volume" of B_k is proportional to r_k^n.

Proof. If C is a closed cube (rectangle with all sides equal) of side s, then C is contained in a closed ball of radius $\sqrt{n}\, s$ by Proposition 10.1.14, and hence in an open ball of radius $2\sqrt{n}\, s$.

Suppose R is a rectangle of positive volume. Let $s > 0$ be a number less than the smallest side of R and such that $2\sqrt{n}\, s < \delta$. If each side of R is an integer multiple of s, then R is contained in a union of closed cubes $C_1, C_2, \ldots, C_m$ of side s such that $\sum_{k=1}^m V(C_k) = V(R)$. So suppose the sides of R are not integer multiples of s. Consider a side of length $(\ell + \alpha)s$, for an integer ℓ and $0 \leq \alpha < 1$. As s is less than the smallest side, $\ell \geq 1$, and so $(\ell+\alpha)s \leq 2\ell s$. Increasing this side to $2\ell s$, and similarly increasing every side of R, we obtain a new larger rectangle of volume at most 2^n times larger, whose sides are multiples of s. See Figure 10.7. Thus R is contained in a union of closed cubes $C_1, C_2, \ldots, C_m$ of side s such that

$$\sum_{k=1}^m V(C_k) \leq 2^n V(R).$$

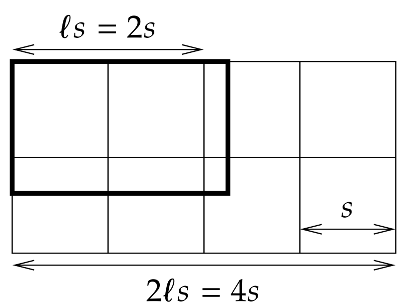

Figure 10.7: Covering a rectangle by cubes of total size at most $2^n V(R)$.

So suppose that S is a null set and there exist open rectangles $\{R_j\}_{j=1}^\infty$ whose union contains S and such that (10.2) is true. Choose closed cubes $\{C_k\}_{k=1}^\infty$ with C_k of side s_k as above that cover all the rectangles $\{R_j\}_{j=1}^\infty$ and so that

$$\sum_{k=1}^\infty s_k^n = \sum_{k=1}^\infty V(C_k) \leq 2^n \sum_{j=1}^\infty V(R_j) < 2^n \epsilon.$$

Covering each C_k with a ball B_k of radius $r_k = 2\sqrt{n}\, s_k < \delta$, we obtain

$$\sum_{k=1}^{\infty} r_k^n = \sum_{k=1}^{\infty} (2\sqrt{n})^n s_k^n < (4\sqrt{n})^n \epsilon.$$

As $S \subset \bigcup_j R_j \subset \bigcup_k C_k \subset \bigcup_k B_k$ and $(4\sqrt{n})^n \epsilon$ can be arbitrarily small, the forward direction follows.

For the other direction, suppose S is covered by balls B_j of radii r_j, such that $\sum_{j=1}^{\infty} r_j^n < \epsilon$, as in the statement of the proposition. Each B_j is contained in an open cube R_j of side $2r_j$. So $V(R_j) = (2r_j)^n = 2^n r_j^n$. Therefore,

$$S \subset \bigcup_{j=1}^{\infty} R_j \qquad \text{and} \qquad \sum_{j=1}^{\infty} V(R_j) \leq \sum_{j=1}^{\infty} 2^n r_j^n < 2^n \epsilon. \qquad \square$$

The definition of outer measure (not just null sets) could have been done with open balls as well. We leave this generalization to the reader.

10.3.2 Examples and basic properties

Example 10.3.3: The set $\mathbb{Q}^n \subset \mathbb{R}^n$ of points with rational coordinates is of measure zero.

Proof: The set $\mathbb{Q}^n$ is countable, so write it as a sequence $q_1, q_2, \ldots$. For each q_j, find an open rectangle R_j with $q_j \in R_j$ and $V(R_j) < \epsilon 2^{-j}$. Then

$$\mathbb{Q}^n \subset \bigcup_{j=1}^{\infty} R_j \qquad \text{and} \qquad \sum_{j=1}^{\infty} V(R_j) < \sum_{j=1}^{\infty} \epsilon 2^{-j} = \epsilon.$$

The example points to a more general result.

Proposition 10.3.4. *A countable union of measure zero sets is of measure zero.*

Proof. Suppose

$$S = \bigcup_{j=1}^{\infty} S_j,$$

where S_j are all measure zero sets. Let $\epsilon > 0$ be given. For each j, there exists a sequence of open rectangles $\{R_{j,k}\}_{k=1}^{\infty}$ such that

$$S_j \subset \bigcup_{k=1}^{\infty} R_{j,k} \qquad \text{and} \qquad \sum_{k=1}^{\infty} V(R_{j,k}) < 2^{-j}\epsilon.$$

Then

$$S \subset \bigcup_{j=1}^{\infty} \bigcup_{k=1}^{\infty} R_{j,k}.$$

All $V(R_{j,k})$ are nonnegative, so the sum over all j and k can be done by summing first over the k and then over the j, see Exercise 2.6.15 in volume I. In particular, as

$$\sum_{j=1}^{\infty}\sum_{k=1}^{\infty} V(R_{j,k}) < \sum_{j=1}^{\infty} 2^{-j}\epsilon = \epsilon. \qquad \square$$

The next example is not just interesting, it will be useful later.

Example 10.3.5: Suppose $n \in \mathbb{N}$, $k = 1, 2, \ldots, n$, and $c \in \mathbb{R}$. Then $P := \{x \in \mathbb{R}^n : x_k = c\}$ is of measure zero. Note that if $n \geq 2$, then P is uncountable.

Proof: First fix $s \in \mathbb{N}$ and consider

$$P_s := \left\{ x \in \mathbb{R}^n : x_k = c \text{ and } |x_j| \leq s \text{ for all } j \neq k \right\}.$$

Given any $\epsilon > 0$ define the open rectangle

$$R := \left\{ x \in \mathbb{R}^n : c - \epsilon < x_k < c + \epsilon \text{ and } |x_j| < s + 1 \text{ for all } j \neq k \right\}.$$

Clearly, $P_s \subset R$. Furthermore,

$$V(R) = 2\epsilon \big(2(s+1)\big)^{n-1}.$$

As s is fixed, $V(R)$ can be arbitrarily small by picking ϵ small enough. So P_s is of measure zero.

Next

$$P = \bigcup_{j=1}^{\infty} P_j$$

and a countable union of measure zero sets is of measure zero.

Example 10.3.6: If $a < b$, then $m^*\big([a, b]\big) = b - a$.

Proof: In $\mathbb{R}$, open rectangles are open intervals. Since $[a, b] \subset (a - \epsilon, b + \epsilon)$ for all $\epsilon > 0$, we have $m^*\big([a, b]\big) \leq b - a$.

The other inequality is harder. Suppose $\big\{(a_j, b_j)\big\}_{j=1}^{\infty}$ are open intervals such that

$$[a, b] \subset \bigcup_{j=1}^{\infty} (a_j, b_j).$$

We wish to bound $\sum_{j=1}^{\infty}(b_j - a_j)$ from below. Since $[a, b]$ is compact, finitely many of the open intervals still cover $[a, b]$. As throwing out some of the intervals only makes the sum smaller, we only need to consider the finite number of intervals covering $[a, b]$. If $(a_i, b_i) \subset (a_j, b_j)$, then we throw out (a_i, b_i) as well. The intervals that are left have distinct left endpoints, and whenever $a_j < a_i < b_j$, then $b_j < b_i$. Therefore, $[a, b] \subset \bigcup_{j=1}^{k}(a_j, b_j)$ for some k, and we assume that the intervals are sorted such that $a_1 < a_2 < \cdots < a_k$. As

(a_2, b_2) is not contained in (a_1, b_1), since $a_j > a_2$ for all $j > 2$, and since the intervals must contain every point in $[a, b]$, we find that $a_2 < b_1$, or in other words $a_1 < a_2 < b_1 < b_2$. Similarly $a_j < a_{j+1} < b_j < b_{j+1}$ for all j. Furthermore, $a_1 < a$ and $b_k > b$. See Figure 10.8 for a sample configuration. As $b_j - a_j > a_{j+1} - a_j$, we obtain

$$\sum_{j=1}^{k} (b_j - a_j) \geq \sum_{j=1}^{k-1} (a_{j+1} - a_j) + (b_k - a_k) = b_k - a_1 > b - a.$$

So $m^*\big([a, b]\big) \geq b - a$.

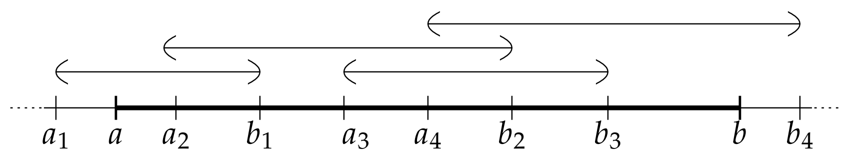

Figure 10.8: Open intervals covering $[a, b]$ which satisfy $a_j < a_{j+1} < b_j < b_{j+1}$ for all j.

Proposition 10.3.7. *Suppose $E \subset \mathbb{R}^n$ is a compact set of measure zero. Then for every $\epsilon > 0$, there exist finitely many open rectangles $R_1, R_2, \ldots, R_k$ such that*

$$E \subset R_1 \cup R_2 \cup \cdots \cup R_k \qquad \textit{and} \qquad \sum_{j=1}^{k} V(R_j) < \epsilon.$$

Moreover, for every $\epsilon > 0$ and every $\delta > 0$, there exist finitely many open balls $B_1, B_2, \ldots, B_\ell$ of radii $r_1, r_2, \ldots, r_\ell < \delta$ such that

$$E \subset B_1 \cup B_2 \cup \cdots \cup B_\ell \qquad \textit{and} \qquad \sum_{j=1}^{\ell} r_j^n < \epsilon.$$

Proof. As E is of measure zero, there exists a sequence of open rectangles $\{R_j\}_{j=1}^{\infty}$ such that

$$E \subset \bigcup_{j=1}^{\infty} R_j \qquad \text{and} \qquad \sum_{j=1}^{\infty} V(R_j) < \epsilon.$$

By compactness, there are finitely many of these rectangles that still contain E. That is, there is some k such that $E \subset R_1 \cup R_2 \cup \cdots \cup R_k$. Hence

$$\sum_{j=1}^{k} V(R_j) \leq \sum_{j=1}^{\infty} V(R_j) < \epsilon.$$

The proof that we can choose balls instead of rectangles is left as an exercise. □

Example 10.3.8: So that the reader is not under the impression that there are only few measure zero sets and that these sets are uncomplicated, here is an uncountable, compact, measure zero subset of $[0,1]$, which contains no intervals. Any $x \in [0,1]$ can be expanded in ternary:

$$x = \sum_{n=1}^{\infty} d_n 3^{-n}, \qquad \text{where } d_n = 0, 1, \text{ or } 2.$$

See §1.5 in volume I, in particular Exercise 1.5.4. Define the *Cantor set* C as

$$C := \left\{ x \in [0,1] : x = \sum_{n=1}^{\infty} d_n 3^{-n}, \text{ where } d_n = 0 \text{ or } d_n = 2 \text{ for all } n \right\}.$$

That is, x is in C if it has a ternary expansion in only 0s and 2s. If x has two expansions, as long as one of them does not have any 1s, then x is in C. Define $C_0 := [0,1]$ and

$$C_k := \left\{ x \in [0,1] : x = \sum_{n=1}^{\infty} d_n 3^{-n}, \text{ where } d_n = 0 \text{ or } d_n = 2 \text{ for all } n = 1, 2, \ldots, k \right\}.$$

Clearly,

$$C = \bigcap_{k=1}^{\infty} C_k .$$

See Figure 10.9.

We leave as an exercise to prove:

(i) Each C_k is a finite union of closed intervals. It is obtained by taking C_{k-1}, and from each closed interval removing the "middle third."

(ii) Each C_k is closed, and so C is closed.

(iii) $m^*(C_k) = 1 - \sum_{n=1}^{k} \frac{2^n}{3^{n+1}}$.

(iv) Hence, $m^*(C) = 0$.

(v) The set C is in one-to-one correspondence with $[0,1]$, in other words, C is uncountable.

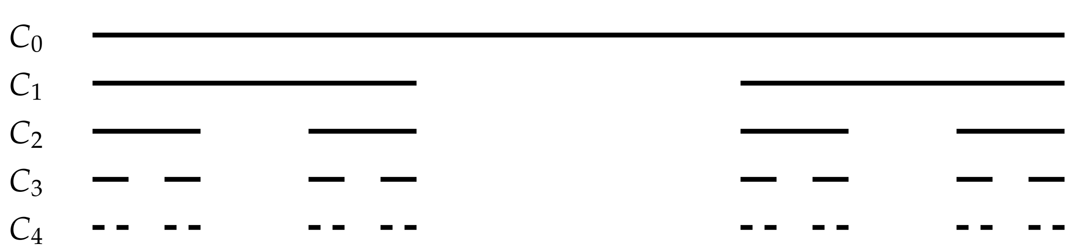

Figure 10.9: Cantor set construction.

10.3.3 Images of null sets under differentiable functions

Before we look at images of measure zero sets, let us see what a continuously differentiable function does to a ball.

Lemma 10.3.9. *Suppose $U \subset \mathbb{R}^n$ is an open set, $B \subset U$ is an open (resp. closed) ball of radius at most r, $f : U \to \mathbb{R}^n$ is continuously differentiable, and suppose $\|f'(x)\| \leq M$ for all $x \in B$. Then $f(B) \subset B'$, where B' is an open (resp. closed) ball of radius at most Mr.*

Proof. Suppose B is open. As the ball B is convex, Proposition 8.4.2 says that $\|f(x) - f(y)\| \leq M\|x - y\|$ for all $x, y \in B$. So if $\|x - y\| < r$, then $\|f(x) - f(y)\| < Mr$. In other words, if $B = B(y, r)$, then $f(B) \subset B\bigl(f(y), Mr\bigr)$. If B is closed, then $\overline{B(y, r)} = B$. As f is continuous, $f(B) = f\bigl(\overline{B(y,r)}\bigr) \subset \overline{f\bigl(B(y,r)\bigr)} \subset \overline{B\bigl(f(y), Mr\bigr)}$, as $f(\bar{A}) \subset \overline{f(A)}$ for any set A. ◻

The image of a measure zero set using a continuous map is not necessarily a measure zero set, although this takes some work to show (see the exercises). However, if the mapping is continuously differentiable, then it cannot "stretch" the set that much.

Proposition 10.3.10. *Suppose $U \subset \mathbb{R}^n$ is open and $f : U \to \mathbb{R}^n$ is continuously differentiable. If $E \subset U$ is a measure zero set, then $f(E)$ is measure zero.*

Proof. We prove the proposition for a compact E and leave the general case as an exercise. Suppose E is compact and of measure zero. First, we will replace U by a smaller open set to make $\|f'(x)\|$ bounded. At each point $x \in E$ pick an open ball $B(x, r_x)$ such that the closed ball $C(x, r_x) \subset U$. By compactness, we only need to take finitely many points $x_1, x_2, \ldots, x_q$ to cover E with the balls $B(x_j, r_{x_j})$. Define

$$U' := \bigcup_{j=1}^{q} B(x_j, r_{x_j}), \qquad K := \bigcup_{j=1}^{q} C(x_j, r_{x_j}).$$

We have $E \subset U' \subset K \subset U$. The set K, being a finite union of compact sets, is compact. The function that takes x to $\|f'(x)\|$ is continuous, and therefore there exists an $M > 0$ such that $\|f'(x)\| \leq M$ for all $x \in K$. So without loss of generality, we may replace U by U' and from now on suppose that $\|f'(x)\| \leq M$ for all $x \in U$.

At each $x \in E$, take the maximum radius δ_x such that $B(x, \delta_x) \subset U$ (we may assume $U \neq \mathbb{R}^n$). Let $\delta := \inf_{x \in E} \delta_x$. We want to show that $\delta > 0$. Take a sequence $\{x_j\}_{j=1}^{\infty}$ in E so that $\delta_{x_j} \to \delta$. As E is compact, we can pick the sequence to be convergent to some $y \in E$. Once $\|x_j - y\| < \frac{\delta_y}{2}$, then $\delta_{x_j} > \frac{\delta_y}{2}$ by the triangle inequality. Thus, $\delta > 0$.

Given $\epsilon > 0$, there exist balls $B_1, B_2, \ldots, B_k$ of radii $r_1, r_2, \ldots, r_k < \delta/2$ such that

$$E \subset B_1 \cup B_2 \cup \cdots \cup B_k \qquad \text{and} \qquad \sum_{j=1}^{k} r_j^n < \epsilon.$$

We can assume that each ball contains a point of E and so the balls are contained in U. Suppose $B'_1, B'_2, \ldots, B'_k$ are the balls of radius $Mr_1, Mr_2, \ldots, Mr_k$ from Lemma 10.3.9, such that $f(B_j) \subset B'_j$ for all j. Then,

$$\begin{aligned} f(E) &\subset f(B_1) \cup f(B_2) \cup \cdots \cup f(B_k) \\ &\subset B'_1 \cup B'_2 \cup \cdots \cup B'_k \end{aligned} \qquad \text{and} \qquad \sum_{j=1}^{k} (Mr_j)^n < M^n \epsilon. \qquad \square$$

10.3.4 Exercises

***Exercise* 10.3.1:** *Finish the proof of Proposition 10.3.7: Show that you can use balls instead of rectangles.*

***Exercise* 10.3.2:** *If $A \subset B$, then $m^*(A) \leq m^*(B)$.*

***Exercise* 10.3.3:** *Suppose $X \subset \mathbb{R}^n$ is a set such that for every $\epsilon > 0$, there exists a set Y such that $X \subset Y$ and $m^*(Y) \leq \epsilon$. Prove that X is a measure zero set.*

***Exercise* 10.3.4:** *Show that if $R \subset \mathbb{R}^n$ is a closed rectangle, then $m^*(R) = V(R)$.*

***Exercise* 10.3.5:** *The closure of a measure zero set can be quite large. Find an example set $S \subset \mathbb{R}^n$ that is of measure zero, but whose closure $\overline{S} = \mathbb{R}^n$.*

***Exercise* 10.3.6:** *Prove the general case of Proposition 10.3.10 without using compactness:*

a) *Mimic the proof to prove that the proposition holds if E is* relatively compact*; a set $E \subset U$ is relatively compact if the closure of E in the subspace topology on U is compact, or in other words if there exists a compact set K with $K \subset U$ and $E \subset K$.*
   Hint: The bound on the size of the derivative still holds, but you need to use countably many balls in the second part of the proof. Be careful as the closure of E need no longer be measure zero.

b) *Now prove it for every null set E.*
   Hint: First show that $\{x \in U : \|x - y\| \geq 1/m$ for all $y \notin U$ and $\|x\| \leq m\}$ is compact for every $m > 0$.

***Exercise* 10.3.7:** *Let $U \subset \mathbb{R}^n$ be an open set and let $f : U \to \mathbb{R}$ be a continuously differentiable function. Let $G := \{(x, y) \in U \times \mathbb{R} : y = f(x)\}$ be the graph of f. Show that G is of measure zero.*

***Exercise* 10.3.8:** *Given a closed rectangle $R \subset \mathbb{R}^n$, show that for every $\epsilon > 0$, there exists a number $s > 0$ and finitely many open cubes $C_1, C_2, \ldots, C_k$ of side s such that $R \subset C_1 \cup C_2 \cup \cdots \cup C_k$ and*

$$\sum_{j=1}^{k} V(C_j) \leq V(R) + \epsilon.$$

***Exercise* 10.3.9:** *Show that there exists a number $k = k(n, r, \delta)$ depending only on n, r and δ such the following holds: Given $B(x, r) \subset \mathbb{R}^n$ and $\delta > 0$, there exist k open balls $B_1, B_2, \ldots, B_k$ of radius at most δ such that $B(x, r) \subset B_1 \cup B_2 \cup \cdots \cup B_k$. Note that you can find k that only depends on n and the ratio δ/r.*

***Exercise* 10.3.10** (Challenging)**:** *Prove the statements of Example 10.3.8. That is, prove:*

a) *Each C_k is a finite union of closed intervals, and so C is closed.*

b) *$m^*(C_k) = 1 - \sum_{n=1}^{k} \frac{2^n}{3^{n+1}}$.*

c) *$m^*(C) = 0$.*

d) *The set C is in one-to-one correspondence with $[0, 1]$.*

***Exercise* 10.3.11:** *Prove that the Cantor set of Example 10.3.8 contains no interval. That is, whenever $a < b$, there exists a point $x \notin C$ such that $a < x < b$.*
Note a consequence of this statement. While every open set in $\mathbb{R}$ is a countable disjoint union of intervals, a closed set (despite being the complement of an open set) need not be a countable union of intervals or points.

***Exercise* 10.3.12** (Challenging)**:** *Let us construct the so-called* Cantor function *or the* Devil's staircase. *Let C be the Cantor set and let C_k be as in Example 10.3.8. Write $x \in [0,1]$ in ternary representation $x = \sum_{n=1} d_n 3^{-n}$. If $d_n \neq 1$ for all n, then let $c_n := \frac{d_n}{2}$ for all n. Otherwise, let k be the smallest integer such that $d_k = 1$. Let $c_n := \frac{d_n}{2}$ if $n < k$, $c_k := 1$, and $c_n := 0$ if $n > k$. Define*

$$\varphi(x) := \sum_{n=1}^{\infty} c_n \, 2^{-n}.$$

a) Prove that φ is continuous and increasing (see Figure 10.9).

b) Prove that for $x \notin C$, φ is differentiable at x and $\varphi'(x) = 0$. (Notice that φ' exists and is zero except for a set of measure zero, yet the function manages to climb from 0 to 1.)

c) Define $\psi : [0,1] \to [0,2]$ by $\psi(x) := \varphi(x) + x$. Show that ψ is continuous, strictly increasing, and bijective.

d) Prove that while $m^*(C) = 0$, $m^*\bigl(\psi(C)\bigr) \neq 0$. That is, continuous functions need not take measure zero sets to measure zero sets. Hint: $m^*\bigl(\psi([0,1] \setminus C)\bigr) = 1$, but $m^*\bigl([0,2]\bigr) = 2$.

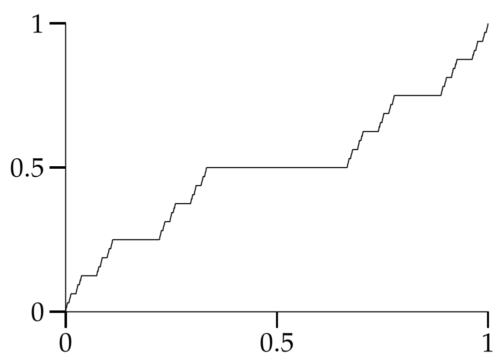

Figure 10.10: Cantor function or Devil's staircase (the function φ from the exercise).

***Exercise* 10.3.13:** *Prove that we obtain the same outer measure if we allow both finite and infinite sequences in the definition. That is, define $\mu^*(S) := \inf \sum_{j \in I} V(R_j)$ where the infimum is taken over all countable (finite or infinite) sets of open rectangles $\{R_j\}_{j \in I}$ such that $S \subset \bigcup_{j \in I} R_j$. Prove that for every $S \subset \mathbb{R}^n$, $\mu^*(S) = m^*(S)$.*

***Exercise* 10.3.14:** *Prove that for any two subsets $A, B \subset \mathbb{R}^n$, we have $m^*(A \cup B) \leq m^*(A) + m^*(B)$.*

***Exercise* 10.3.15:** *Suppose $A, B \subset \mathbb{R}^n$ are such that $m^*(B) = 0$. Prove that $m^*(A \cup B) = m^*(A)$.*

***Exercise* 10.3.16** (Challenging)**:** *Suppose $R_1, R_2, \ldots, R_n$ are pairwise disjoint open rectangles. Prove that $m^*(R_1 \cup R_2 \cup \cdots \cup R_n) = m^*(R_1) + m^*(R_2) + \cdots + m^*(R_n)$. Hint: Some of the exercises above may prove very useful.*

10.4 The set of Riemann integrable functions

Note: 1 lecture

10.4.1 Oscillation and continuity

Consider $D \subset \mathbb{R}^n$ and $f : D \to \mathbb{R}$. Instead of just saying that f is or is not continuous at a point $x \in D$, we want to quantify how discontinuous is f at x. For every $\delta > 0$, define the *oscillation* of f on the δ-ball in subspace topology, $B_D(x, \delta) = B_{\mathbb{R}^n}(x, \delta) \cap D$, as

$$o(f, x, \delta) := \sup_{y \in B_D(x,\delta)} f(y) - \inf_{y \in B_D(x,\delta)} f(y) = \sup_{y_1, y_2 \in B_D(x,\delta)} \bigl(f(y_1) - f(y_2)\bigr).$$

That is, $o(f, x, \delta)$ is the length of the smallest interval that contains the image $f\bigl(B_D(x, \delta)\bigr)$. The definition makes sense for unbounded functions, where the oscillation can be ∞, although we will mainly consider bounded functions. Clearly $o(f, x, \delta) \geq 0$ and $o(f, x, \delta) \leq o(f, x, \delta')$ whenever $\delta < \delta'$. Therefore, the limit as $\delta \to 0$ from the right exists, and we define the *oscillation* of f at x as

$$o(f, x) := \lim_{\delta \to 0^+} o(f, x, \delta) = \inf_{\delta > 0} o(f, x, \delta).$$

We will prove that function is continuous at x if and only if $o(f, x) = 0$. Fox example, if $f : \mathbb{R} \to \mathbb{R}$ is the Dirichlet function where $f(x) = 1$ if $x \in \mathbb{Q}$ and $f(x) = 0$ otherwise, then $o(f, x) = 1$ for every x, as any interval contains both rational and irrational numbers. Accordingly, f is not continuous at any x. For another example, which is perhaps the origin of the terminology, let $g : \mathbb{R} \to \mathbb{R}$ be given by $g(x) = \sin(1/x)$ for $x \neq 0$ and $g(0) = 0$, see Figure 10.11. Then at the discontinuity at $x = 0$, we find $o(g, 0) = 2$, as in any neighborhood of 0, the function takes both values 1 and -1. For all $x \neq 0$, the function is continuous and so, as we will see, $o(g, x) = 0$.

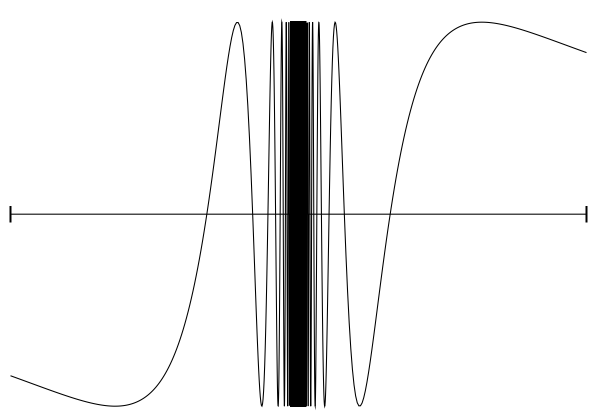

Figure 10.11: Graph of $\sin(1/x)$.

Proposition 10.4.1. *A function $f : D \to \mathbb{R}$ is continuous at $x \in D$ if and only if $o(f, x) = 0$.*

Proof. First suppose that f is continuous at $x \in D$. Given $\epsilon > 0$, there exists a $\delta > 0$ such that for $y \in B_D(x, \delta)$, we have $|f(x) - f(y)| < \epsilon$. Therefore, if $y_1, y_2 \in B_D(x, \delta)$, then

$$f(y_1) - f(y_2) = \bigl(f(y_1) - f(x)\bigr) - \bigl(f(y_2) - f(x)\bigr) < \epsilon + \epsilon = 2\epsilon.$$

Take the supremum over y_1 and y_2 to find

$$o(f, x, \delta) = \sup_{y_1, y_2 \in B_D(x,\delta)} \bigl(f(y_1) - f(y_2)\bigr) \leq 2\epsilon.$$

As $o(x, f) \leq o(f, x, \delta) \leq 2\epsilon$, and $\epsilon > 0$ was arbitrary, $o(x, f) = 0$.

On the other hand, suppose $o(x, f) = 0$. Given $\epsilon > 0$, find a $\delta > 0$ such that $o(f, x, \delta) < \epsilon$. If $y \in B_D(x, \delta)$, then

$$|f(x) - f(y)| \leq \sup_{y_1, y_2 \in B_D(x,\delta)} \bigl(f(y_1) - f(y_2)\bigr) = o(f, x, \delta) < \epsilon. \qquad \square$$

Proposition 10.4.2. *Let $D \subset \mathbb{R}^n$ be closed, $f : D \to \mathbb{R}$, and $\epsilon > 0$. The set $\{x \in D : o(f, x) \geq \epsilon\}$ is closed.*

Proof. Equivalently, we want to show that $G := \{x \in D : o(f, x) < \epsilon\}$ is open in the subspace topology. Consider $x \in G$. As $\inf_{\delta > 0} o(f, x, \delta) < \epsilon$, find a $\delta > 0$ such that

$$o(f, x, \delta) < \epsilon.$$

Take any $\xi \in B_D(x, \delta/2)$. Notice that $B_D(\xi, \delta/2) \subset B_D(x, \delta)$. Therefore,

$$o(f, \xi, \delta/2) = \sup_{y_1, y_2 \in B_D(\xi,\delta/2)} \bigl(f(y_1) - f(y_2)\bigr) \leq \sup_{y_1, y_2 \in B_D(x,\delta)} \bigl(f(y_1) - f(y_2)\bigr) = o(f, x, \delta) < \epsilon.$$

So $o(f, \xi) < \epsilon$ as well. As this is true for all $\xi \in B_D(x, \delta/2)$, we get that G is open in the subspace topology, and $D \setminus G$ is closed as claimed. $\square$

10.4.2 The set of Riemann integrable functions

We have seen that continuous functions are Riemann integrable, but we also know that certain kinds of discontinuities are allowed. It turns out that as long as the discontinuities happen on a set of measure zero, the function is integrable, and vice versa.

Theorem 10.4.3 (Riemann–Lebesgue or Lebesgue–Vitali*). *Let $R \subset \mathbb{R}^n$ be a closed rectangle and $f : R \to \mathbb{R}$ bounded. Then f is Riemann integrable if and only if the set of discontinuities of f is of measure zero.*

*Giuseppe Vitali (1875–1932) was an Italian mathematician. Note also that the name Riemann–Lebesgue often refers to a result like Exercise 5.2.18 from volume I.

Proof. Let $S \subset R$ be the set of discontinuities of f, that is, $S = \{x \in R : o(f,x) > 0\}$. Suppose S is a measure zero set: $m^*(S) = 0$. The trick to proving that f is integrable is to isolate the bad set into a small set of subrectangles of a partition. A partition has finitely many subrectangles, so we need compactness. If S were closed, then it would be compact and we could cover it by finitely many small rectangles. Unfortunately, S itself is not closed in general, but the following set is. Given $\epsilon > 0$, define

$$S_\epsilon := \{x \in R : o(f,x) \geq \epsilon\}.$$

By Proposition 10.4.2, S_ϵ is closed, and as it is also a subset of the bounded R, S_ϵ is compact. Moreover, $S_\epsilon \subset S$ and S is of measure zero, so S_ϵ is of measure zero. Via Proposition 10.3.7, finitely many open rectangles $O_1, O_2, \ldots, O_k$ cover S_ϵ and $\sum_{j=1}^{\infty} V(O_j) < \epsilon$.

The set $T := R \setminus (O_1 \cup \cdots \cup O_k)$ is closed, bounded, and so compact. As $o(f,x) < \epsilon$ for all $x \in T$, for each $x \in T$, there is a $\delta > 0$ such that $o(f,x,\delta) < \epsilon$, so there exists a small closed rectangle $T_x \subset B(x,\delta)$ with x in the interior of T_x, such that

$$\sup_{y \in T_x} f(y) - \inf_{y \in T_x} f(y) < \epsilon.$$

The interiors of the rectangles T_x cover T. As T is compact, finitely many such rectangles $T_1, T_2, \ldots, T_m$ cover T. Construct a partition P out of the endpoints of the rectangles $T_1, T_2, \ldots, T_m$ and $O_1, O_2, \ldots, O_k$ (ignoring those that are outside the endpoints of R). The subrectangles $R_1, R_2, \ldots, R_p$ of P are such that every R_j is contained in T_ℓ for some ℓ or the closure of O_ℓ for some ℓ. Order the rectangles so that $R_1, R_2, \ldots, R_q$ are those that are contained in some T_ℓ, and $R_{q+1}, R_{q+2}, \ldots, R_p$ are the rest. See Figure 10.12. So

$$\sum_{j=1}^{q} V(R_j) \leq V(R) \qquad \text{and} \qquad \sum_{j=q+1}^{p} V(R_j) \leq \sum_{\ell=1}^{k} V(O_\ell) < \epsilon.$$

The second estimate holds because the R_j that are subsets of $\overline{O}_\ell$ give a partition of $\overline{O}_\ell$ and hence their volumes sum to $V(O_\ell)$. Let m_j and M_j be the inf and sup of f over R_j as usual. If $R_j \subset T_\ell$ for some ℓ, then $M_j - m_j < \epsilon$. Let $B \in \mathbb{R}$ be such that $|f(x)| \leq B$ for all $x \in R$, so $M_j - m_j \leq 2B$ over all rectangles. Then

$$\begin{aligned}
U(P,f) - L(P,f) &= \sum_{j=1}^{p} (M_j - m_j) V(R_j) \\
&= \left(\sum_{j=1}^{q} (M_j - m_j) V(R_j) \right) + \left(\sum_{j=q+1}^{p} (M_j - m_j) V(R_j) \right) \\
&< \left(\sum_{j=1}^{q} \epsilon\, V(R_j) \right) + \left(\sum_{j=q+1}^{p} 2B\, V(R_j) \right) \\
&< \epsilon\, V(R) + 2B\epsilon = \epsilon\big(V(R) + 2B\big).
\end{aligned}$$

We can make the right-hand side as small as we want, and hence f is integrable.

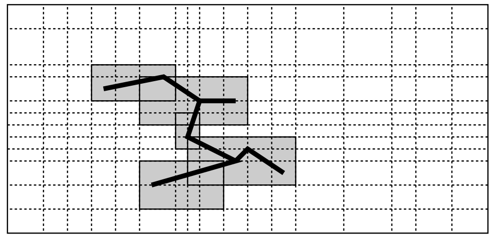

Figure 10.12: A rectangle R with S_ϵ marked as thick black line, and the O_ℓ as shaded rectangles. The partition is given by the dotted lines. Note how the R_j partition the O_ℓ.

For the other direction, suppose f is Riemann integrable on R. Let S be the set of discontinuities of f again. Consider the sequence of sets

$$S_{1/k} = \left\{ x \in R : o(f,x) \geq 1/k \right\}.$$

Fix a $k \in \mathbb{N}$. Given an $\epsilon > 0$, find a partition P with subrectangles $R_1, R_2, \ldots, R_p$ such that

$$U(P,f) - L(P,f) = \sum_{j=1}^{p} (M_j - m_j) V(R_j) < \epsilon.$$

Suppose $R_1, R_2, \ldots, R_p$ are ordered so that the interiors of $R_1, R_2, \ldots, R_q$ intersect $S_{1/k}$, while the interiors of $R_{q+1}, R_{q+2}, \ldots, R_p$ are disjoint from $S_{1/k}$. Let R_j° denote the interior of R_j. Suppose $j \leq q$ and consider $x \in R_j^\circ \cap S_{1/k}$. Let $\delta > 0$ be small enough so that $B(x,\delta) \subset R_j$. As $x \in S_{1/k}$, we get $o(f,x,\delta) \geq o(f,x) \geq 1/k$, which, along with $B(x,\delta) \subset R_j$, implies $M_j - m_j \geq 1/k$. Then

$$\epsilon > \sum_{j=1}^{p} (M_j - m_j) V(R_j) \geq \sum_{j=1}^{q} (M_j - m_j) V(R_j) \geq \frac{1}{k} \sum_{j=1}^{q} V(R_j).$$

In other words, $\sum_{j=1}^{q} V(R_j) < k\epsilon$. Let G be the set of all boundaries of all the subrectangles of P. The set G is of measure zero (it can be covered by finitely many sets from Example 10.3.5). We find

$$S_{1/k} \subset R_1^\circ \cup R_2^\circ \cup \cdots \cup R_q^\circ \cup G.$$

As G can also be covered by open rectangles arbitrarily small volume, $S_{1/k}$ must be of measure zero. As

$$S = \bigcup_{k=1}^{\infty} S_{1/k}$$

and a countable union of measure zero sets is of measure zero, S is of measure zero. $\square$

Corollary 10.4.4. *Let $R \subset \mathbb{R}^n$ be a closed rectangle. Let $\mathscr{R}(R)$ be the set of Riemann integrable functions on R. Then*

(*i*) *$\mathscr{R}(R)$ is a real algebra: If $f, g \in \mathscr{R}(R)$ and $a \in \mathbb{R}$, then $af \in \mathscr{R}(R)$, $f + g \in \mathscr{R}(R)$ and $fg \in \mathscr{R}(R)$.*

(*ii*) *If $f, g \in \mathscr{R}(R)$ and*

$$\varphi(x) := \max\{f(x), g(x)\}, \qquad \psi(x) := \min\{f(x), g(x)\},$$

then $\varphi, \psi \in \mathscr{R}(R)$.

(*iii*) *If $f \in \mathscr{R}(R)$, then $|f| \in \mathscr{R}(R)$, where $|f|(x) := |f(x)|$.*

(*iv*) *If $R' \subset \mathbb{R}^n$ is another closed rectangle, $U \subset \mathbb{R}^n$ and $U' \subset \mathbb{R}^n$ are open sets such that $R \subset U$ and $R' \subset U'$, $g \colon U \to U'$ is continuously differentiable, bijective, g^{-1} is continuously differentiable, $g(R) \subset R'$, and $f \in \mathscr{R}(R')$, then the composition $f \circ g$ is Riemann integrable on R.*

The proof is contained in the exercises.

10.4.3 Exercises

Exercise* 10.4.1*: *Suppose $f \colon (a, b) \times (c, d) \to \mathbb{R}$ is a bounded continuous function. Show that the integral of f over $R = [a, b] \times [c, d]$ makes sense and is uniquely defined. That is, set f to be anything (bounded) on the boundary of R and compute the integral, showing that the values on the boundary are irrelevant.*

Exercise* 10.4.2*: *Suppose $R \subset \mathbb{R}^n$ is a closed rectangle. Show that $\mathscr{R}(R)$, the set of Riemann integrable functions, is an algebra. That is, show that if $f, g \in \mathscr{R}(R)$ and $a \in \mathbb{R}$, then $af \in \mathscr{R}(R)$, $f + g \in \mathscr{R}(R)$, and $fg \in \mathscr{R}(R)$.*

Exercise* 10.4.3*: *Suppose $R \subset \mathbb{R}^n$ is a closed rectangle and $f \colon R \to \mathbb{R}$ is a bounded function which is zero except on a closed set $E \subset R$ of measure zero. Show that $\int_R f$ exists and compute it.*

Exercise* 10.4.4*: *Suppose $R \subset \mathbb{R}^n$ is a closed rectangle and $f \colon R \to \mathbb{R}$ and $g \colon R \to \mathbb{R}$ are two Riemann integrable functions. Suppose $f = g$ except for a closed set $E \subset R$ of measure zero. Show that $\int_R f = \int_R g$.*

Exercise* 10.4.5*: *Suppose $R \subset \mathbb{R}^n$ is a closed rectangle and $f \colon R \to \mathbb{R}$ is a bounded function.*

a) Suppose there exists a closed set $E \subset R$ of measure zero such that $f|_{R \setminus E}$ is continuous. Then $f \in \mathscr{R}(R)$.

b) Find an example where $E \subset R$ is a set of measure zero (not closed) such that $f|_{R \setminus E}$ is continuous and $f \notin \mathscr{R}(R)$.

Exercise* 10.4.6*: *Suppose $R \subset \mathbb{R}^n$ is a closed rectangle and $f \colon R \to \mathbb{R}$ and $g \colon R \to \mathbb{R}$ are Riemann integrable. Show that*

$$\varphi(x) := \max\{f(x), g(x)\} \qquad \text{and} \qquad \psi(x) := \min\{f(x), g(x)\}$$

are Riemann integrable.

***Exercise* 10.4.7:** *Suppose $R \subset \mathbb{R}^n$ is a closed rectangle and $f \colon R \to \mathbb{R}$ is Riemann integrable. Show that $|f|$ is Riemann integrable. Hint: Define $f_+(x) := \max\{f(x), 0\}$ and $f_-(x) := \max\{-f(x), 0\}$, and then write $|f|$ in terms of f_+ and f_-.*

***Exercise* 10.4.8:**

a) Suppose $R \subset \mathbb{R}^n$ and $R' \subset \mathbb{R}^n$ are closed rectangles, $U \subset \mathbb{R}^n$ and $U' \subset \mathbb{R}^n$ are open sets such that $R \subset U$ and $R' \subset U'$, $g \colon U \to U'$ is continuously differentiable, bijective, g^{-1} is continuously differentiable, $g(R) \subset R'$, and $f \in \mathcal{R}(R')$, then the composition $f \circ g$ is Riemann integrable on R.

b) Find a counterexample when g is not one-to-one. Hint: Try $g(x, y) := (x, 0)$ and $R = R' = [0,1]\times[0,1]$.

***Exercise* 10.4.9:** *Suppose $f \colon [0,1]^2 \to \mathbb{R}$ is defined by*

$$f(x,y) := \begin{cases} \frac{1}{kq} & \text{if } x, y \in \mathbb{Q} \text{ and } x = \frac{\ell}{k} \text{ and } y = \frac{p}{q} \text{ in lowest terms,} \\ 0 & \text{else.} \end{cases}$$

Show that $f \in \mathcal{R}\bigl([0,1]^2\bigr)$.

***Exercise* 10.4.10:** *Compute the oscillation $o\bigl(f, (x,y)\bigr)$ for all $(x,y) \in \mathbb{R}^2$ for the function*

$$f(x,y) := \begin{cases} \frac{xy}{x^2+y^2} & \text{if } (x,y) \neq (0,0), \\ 0 & \text{if } (x,y) = (0,0). \end{cases}$$

***Exercise* 10.4.11:** *Consider the popcorn function $f \colon [0,1] \to \mathbb{R}$,*

$$f(x) := \begin{cases} \frac{1}{q} & \text{if } x \in \mathbb{Q} \text{ and } x = \frac{p}{q} \text{ in lowest terms,} \\ 0 & \text{else.} \end{cases}$$

Compute $o(f,x)$ for all $x \in [0,1]$.

***Exercise* 10.4.12:** *Suppose $f \colon [a,b] \to \mathbb{R}$ and $g \colon [c,d] \to \mathbb{R}$ are Riemann integrable. Show that $h \colon [a,b] \times [c,d] \to \mathbb{R}$ defined by $h(x,y) := f(x)g(y)$ is Riemann integrable and*

$$\int_{[a,b]\times[c,d]} h = \left(\int_a^b f \right) \left(\int_c^d g \right).$$

***Exercise* 10.4.13:** *Let $R \subset \mathbb{R}^n$ be a closed rectangle and $f \colon R \to \mathbb{R}$ a Riemann integrable function such that $f(x) \geq 0$ for all $x \in R$. Show that if $\int_R f = 0$, then there is a measure zero set $E \subset R$ such that $f(x) = 0$ for all $x \in R \setminus E$ (one says "$f = 0$ almost everywhere"). Note: This exercise in particular implies the rather subtle statement: If $f(x) > 0$ for all $x \in R$, then $\int_R f > 0$.*

10.5 Jordan measurable sets

Note: 1–1.5 lecture

10.5.1 Volume and Jordan measurable sets

Given a set $S \subset \mathbb{R}^n$, its *characteristic function* or *indicator function* $\chi_S \colon \mathbb{R}^n \to \mathbb{R}$ is defined by

$$\chi_S(x) := \begin{cases} 1 & \text{if } x \in S, \\ 0 & \text{if } x \notin S. \end{cases}$$

A bounded set S is *Jordan measurable*[*] if for some closed rectangle R such that $S \subset R$, the function χ_S is Riemann integrable, that is, $\chi_S \in \mathcal{R}(R)$. Take two closed rectangles R and R' with $S \subset R$ and $S \subset R'$, then $R \cap R'$ is a closed rectangle also containing S. By Proposition 10.1.13 and Exercise 10.1.7, $\chi_S \in \mathcal{R}(R \cap R')$ and so $\chi_S \in \mathcal{R}(R')$. Thus

$$\int_R \chi_S = \int_{R'} \chi_S = \int_{R \cap R'} \chi_S.$$

We define the *n-dimensional volume* of the bounded Jordan measurable set S as

$$V(S) := \int_R \chi_S,$$

where R is any closed rectangle containing S.

Proposition 10.5.1. *A bounded set $S \subset \mathbb{R}^n$ is Jordan measurable if and only if the boundary ∂S is a measure zero set.*

Proof. Suppose R is a closed rectangle such that S is contained in the interior of R. If $x \in \partial S$, then for every $\delta > 0$, the sets $S \cap B(x, \delta)$ (where χ_S is 1) and the sets $(R \setminus S) \cap B(x, \delta)$ (where χ_S is 0) are both nonempty. So χ_S is not continuous at x. If x is either in the interior of S or in the complement of the closure $\overline{S}$, then χ_S is either identically 1 or identically 0 in a whole neighborhood of x and hence χ_S is continuous at x. Therefore, the set of discontinuities of χ_S is precisely the boundary ∂S. The proposition follows. ☐

Proposition 10.5.2. *Suppose S and T are bounded Jordan measurable sets. Then*

(i) *The closure $\overline{S}$ is Jordan measurable.*

(ii) *The interior S° is Jordan measurable.*

(iii) *$S \cup T$ is Jordan measurable.*

(iv) *$S \cap T$ is Jordan measurable.*

(v) *$S \setminus T$ is Jordan measurable.*

[*] Named after the French mathematician Marie Ennemond Camille Jordan (1838–1922).

The proof of the proposition is left as an exercise. Next, the volume that we defined above coincides with the outer measure we defined previously.

Proposition 10.5.3. *If $S \subset \mathbb{R}^n$ is Jordan measurable, then $V(S) = m^*(S)$.*

Proof. Given $\epsilon > 0$, let R be a closed rectangle that contains S. Let P be a partition of R such that

$$U(P, \chi_S) \leq \left(\int_R \chi_S \right) + \epsilon = V(S) + \epsilon \qquad \text{and} \qquad L(P, \chi_S) \geq \left(\int_R \chi_S \right) - \epsilon = V(S) - \epsilon.$$

Let $R_1, R_2, \ldots, R_k$ be all the subrectangles of P such that χ_S is not identically zero on each R_j. That is, there is some point $x \in R_j$ such that $x \in S$ (i.e. $\chi_S(x) = 1$). Let O_j be an open rectangle such that $R_j \subset O_j$ and $V(O_j) < V(R_j) + \epsilon/k$. Notice that $S \subset \bigcup_j O_j$. Then

$$U(P, \chi_S) = \sum_{j=1}^{k} V(R_j) > \left(\sum_{j=1}^{k} V(O_j) \right) - \epsilon \geq m^*(S) - \epsilon.$$

As $U(P, \chi_S) \leq V(S) + \epsilon$, then $m^*(S) - \epsilon \leq V(S) + \epsilon$, or in other words $m^*(S) \leq V(S)$.

Let $R'_1, R'_2, \ldots, R'_\ell$ be all the subrectangles of P such that χ_S is identically one on each R'_j. In other words, these are the subrectangles contained in S. The interiors of the subrectangles R'°_j are disjoint and $V(R'^\circ_j) = V(R'_j)$. Via Exercise 10.3.16,

$$m^*\Bigl(\bigcup_{j=1}^{\ell} R'^\circ_j\Bigr) = \sum_{j=1}^{\ell} V(R'^\circ_j).$$

Hence

$$m^*(S) \geq m^*\Bigl(\bigcup_{j=1}^{\ell} R'_j\Bigr) \geq m^*\Bigl(\bigcup_{j=1}^{\ell} R'^\circ_j\Bigr) = \sum_{j=1}^{\ell} V(R'^\circ_j) = \sum_{j=1}^{\ell} V(R'_j) = L(P, f) \geq V(S) - \epsilon.$$

Therefore $m^*(S) \geq V(S)$ as well. □

10.5.2 Integration over Jordan measurable sets

In $\mathbb{R}$ there is really only one reasonable type of set to integrate over—an interval. In $\mathbb{R}^n$ there are many kinds of very reasonable sets. The ones that work with the Riemann integral are the Jordan measurable sets.

Definition 10.5.4. Let $S \subset \mathbb{R}^n$ be a bounded Jordan measurable set. A bounded function $f \colon S \to \mathbb{R}$ is said to be *Riemann integrable on S*, or $f \in \mathscr{R}(S)$, if for a closed rectangle R such that $S \subset R$, the function $\widetilde{f} \colon R \to \mathbb{R}$ defined by

$$\widetilde{f}(x) := \begin{cases} f(x) & \text{if } x \in S, \\ 0 & \text{otherwise,} \end{cases}$$

is in $\mathscr{R}(R)$. We write

$$\int_S f := \int_R \widetilde{f}.$$

As above, Proposition 10.1.13 and Exercise 10.1.7 imply that the actual rectangle R used is irrelevant and the integral is well-defined.

When f is defined on a larger set and we wish to integrate over S, then we apply the definition to the restriction $f|_S$. As the restriction can be defined by the product $f\chi_S$, and the product of Riemann integrable functions is Riemann integrable, $f|_S$ is automatically Riemann integrable. In particular, if $f : R \to \mathbb{R}$ for a closed rectangle R, and $S \subset R$ is a Jordan measurable subset, then

$$\int_S f = \int_R f\chi_S.$$

Proposition 10.5.5. *If $S \subset \mathbb{R}^n$ is a bounded Jordan measurable set and $f : S \to \mathbb{R}$ is a bounded continuous function, then f is integrable on S.*

Proof. Define the function $\widetilde{f}$ as above for some closed rectangle R with $S \subset R$. If $x \in R \setminus \overline{S}$, then $\widetilde{f}$ is identically zero in a neighborhood of x. Similarly, if x is in the interior of S, then $\widetilde{f} = f$ on a neighborhood of x and f is continuous at x. Therefore, $\widetilde{f}$ is only ever possibly discontinuous at ∂S, which is a set of measure zero, and we are finished. □

We say some property *for almost every x* or *almost everywhere* if it holds for all x except on a set of measure zero. For example, we say $f : S \to \mathbb{R}$ and $g : S \to \mathbb{R}$ are equal almost everywhere if there exists a measure zero set $E \subset S$ such that $f(x) = g(x)$ for all $x \in S \setminus E$.

Many of the standard properties of the integral carry over easily since we are really integrating over a rectangle. Furthermore, we can make some of the statements to be *almost everywhere*. Proofs of the following three propositions left as exercises.

Proposition 10.5.6. *Suppose $S \subset \mathbb{R}^n$ is a bounded Jordan measurable set and $f : S \to \mathbb{R}$ and $g : S \to \mathbb{R}$ are Riemann integrable on S, and $\alpha \in \mathbb{R}$. Then*

(i) *If $f = 0$ almost everywhere, then $\int_S f = 0$.*

(ii) *If $f = g$ almost everywhere, then $\int_S f = \int_S g$.*

(iii) *$f + g$ is Riemann integrable on S and $\int_S (f + g) = \int_S f + \int_S g$.*

(iv) *αf is Riemann integrable on S and $\int_S \alpha f = \alpha \int_S f$.*

(v) *If $f(x) \leq g(x)$ for almost every x, then $\int_S f \leq \int_S g$.*

We also have additivity.

Proposition 10.5.7. *Suppose $A \subset \mathbb{R}^n$ and $B \subset \mathbb{R}^n$ are disjoint bounded Jordan measurable sets and $f : A \cup B \to \mathbb{R}$ is such that the restrictions $f|_A$ and $f|_B$ are Riemann integrable on A and B respectively. Then f is Riemann integrable on $A \cup B$ and*

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

Finally, to integrate over non-rectangular regions using Fubini's theorem, the typical way is to cut the region into simpler pieces that can be described by two graphs. We state the theorem in the plane, but similar statements can be made in more variables. The proof is again left as an exercise.

Proposition 10.5.8. *Let $f\colon [a,b] \to \mathbb{R}$ and $g\colon [a,b] \to \mathbb{R}$ be continuous functions and such that for all $x \in (a,b)$, $f(x) < g(x)$. Let*

$$U := \bigl\{ (x,y) \in \mathbb{R}^2 : a < x < b \text{ and } f(x) < y < g(x) \bigr\}.$$

See Figure 10.13. Then U is Jordan measurable, and if $\varphi\colon U \to \mathbb{R}$ is Riemann integrable on U, then

$$\int_U \varphi = \int_a^b \int_{f(x)}^{g(x)} \varphi(x,y)\, dy\, dx.$$

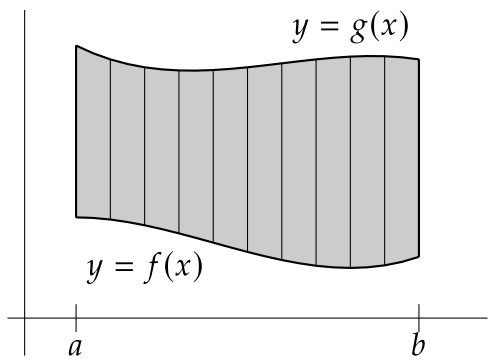

Figure 10.13: Region between two graphs.

10.5.3 Images of Jordan measurable subsets

Finally, images of Jordan measurable sets are Jordan measurable under nice enough mappings. For simplicity, we assume that the Jacobian determinant never vanishes.

Proposition 10.5.9. *Suppose $U \subset \mathbb{R}^n$ is open and $S \subset U$ is a compact Jordan measurable set. Suppose $g\colon U \to \mathbb{R}^n$ is a one-to-one continuously differentiable mapping such that the Jacobian determinant J_g is never zero on S. Then $g(S)$ is bounded and Jordan measurable.*

Proof. Let $T := g(S)$. By Lemma 7.5.5 from volume I, the set T is also compact and so closed and bounded. We claim $\partial T \subset g(\partial S)$. Suppose the claim is proved. As S is Jordan measurable, then ∂S is measure zero. Then $g(\partial S)$ is measure zero by Proposition 10.3.10. As $\partial T \subset g(\partial S)$, then T is Jordan measurable.

It is therefore left to prove the claim. As T is closed, $\partial T \subset T$. Suppose $y \in \partial T$, then there must exist an $x \in S$ such that $g(x) = y$, and by hypothesis $J_g(x) \neq 0$. We use the inverse function theorem (Theorem 8.5.1). We find a neighborhood $V \subset U$ of x and an open set W such that the restriction $f|_V$ is a one-to-one and onto function from V to W with a continuously differentiable inverse. In particular, $g(x) = y \in W$. As $y \in \partial T$, there exists a sequence $\{y_k\}_{k=1}^{\infty}$ in W with $\lim_{k\to\infty} y_k = y$ and $y_k \notin T$. As $g|_V$ is invertible and in particular has a continuous inverse, there exists a sequence $\{x_k\}_{k=1}^{\infty}$ in V such that $g(x_k) = y_k$ and $\lim_{k\to\infty} x_k = x$. Since $y_k \notin T = g(S)$, clearly $x_k \notin S$. Since $x \in S$, we conclude that $x \in \partial S$. The claim is proved, $\partial T \subset g(\partial S)$. □

10.5.4 Exercises

***Exercise* 10.5.1:** *Prove Proposition 10.5.2.*

***Exercise* 10.5.2:** *Prove that a bounded convex set is Jordan measurable. Hint: Induction on dimension.*

***Exercise* 10.5.3:** *Prove Proposition 10.5.8. That is,*

a) *Show that U is Jordan measurable.*

b) *Prove that $\int_U \varphi = \int_a^b \int_{f(x)}^{g(x)} \varphi(x,y)\,dy\,dx$.*

***Exercise* 10.5.4:** *Let us construct an example of a non-Jordan measurable open set. Start in one dimension. Let $\{r_j\}_{j=1}^{\infty}$ be an enumeration of all rational numbers in $(0,1)$. Let (a_j, b_j) be open intervals such that $(a_j, b_j) \subset (0,1)$ for all j, $r_j \in (a_j, b_j)$, and $\sum_{j=1}^{\infty}(b_j - a_j) < 1/2$. Now let $U := \bigcup_{j=1}^{\infty}(a_j, b_j)$.*

a) *Show the open intervals (a_j, b_j) as above actually exist.*

b) *Prove $\partial U = [0,1] \setminus U$.*

c) *Prove ∂U is not of measure zero, and therefore U is not Jordan measurable.*

d) *Show that $W := \big(U \times (0,2)\big) \cup \big((0,1) \times (1,2)\big)$ is a connected bounded open set in $\mathbb{R}^2$ that is not Jordan measurable.*

***Exercise* 10.5.5:** *Suppose $K \subset \mathbb{R}^n$ is a closed measure zero set.*

a) *If K is bounded, prove that K is Jordan measurable.*

b) *If $S \subset \mathbb{R}^n$ is bounded and Jordan measurable, prove that $S \setminus K$ is Jordan measurable.*

c) *Construct a bounded Jordan measurable $S \subset \mathbb{R}^n$ and a bounded $T \subset \mathbb{R}^n$ of measure zero, such that neither T nor $S \setminus T$ is Jordan measurable.*

***Exercise* 10.5.6:** *Suppose $U \subset \mathbb{R}^n$ is open and $K \subset U$ is compact. Find a compact Jordan measurable set S such that $S \subset U$ and $K \subset S^\circ$ (K is in the interior of S).*

***Exercise* 10.5.7:** *Prove a version of Corollary 10.4.4, replacing all closed rectangles with closed and bounded Jordan measurable sets.*

***Exercise* 10.5.8:** *Prove Proposition 10.5.6.*

***Exercise* 10.5.9:** *Prove Proposition 10.5.7.*

10.6 Green's theorem

Note: 1 lecture, requires chapter 9

One of the most important theorems in the calculus of several variables is the so-called generalized Stokes' theorem, a generalization of the fundamental theorem of calculus. The two-dimensional version is called Green's theorem*. We will state the theorem in general, but we will only prove a special, but important, case.

Definition 10.6.1. Let $U \subset \mathbb{R}^2$ be a bounded connected open set. Suppose the boundary ∂U is a disjoint union of (the images of) finitely many simple closed piecewise smooth paths such that every $p \in \partial U$ is in the closure of $\mathbb{R}^2 \setminus \overline{U}$. Then U is called a *bounded domain with piecewise smooth boundary* in $\mathbb{R}^2$.

The condition about points outside the closure says that locally ∂U separates $\mathbb{R}^2$ into an "inside" and an "outside." The condition prevents ∂U from being just a "cut" inside U. As we travel along the path in a certain orientation, there is a well-defined left and a right, and either U is on the left and the complement of U is on the right, or vice versa. The orientation on U is the direction in which we travel along the paths. We can switch orientation if needed by reparametrizing the path.

Definition 10.6.2. Let $U \subset \mathbb{R}^2$ be a bounded domain with piecewise smooth boundary, let ∂U be oriented , and let $\gamma \colon [a, b] \to \mathbb{R}^2$ be a parametrization of ∂U giving the orientation. Write $\gamma(t) = \big(x(t), y(t)\big)$. If the vector $n(t) := \big(-y'(t), x'(t)\big)$ points into the domain, that is, $\epsilon n(t) + \gamma(t)$ is in U for all small enough $\epsilon > 0$, then ∂U is *positively oriented*. See Figure 10.14. Otherwise it is *negatively oriented*.

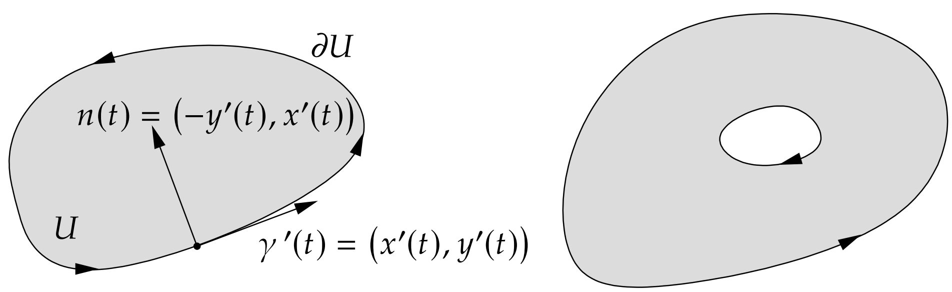

Figure 10.14: Positively oriented domain (left), and a positively oriented domain with a hole (right).

The vector $n(t)$ turns $\gamma\,'(t)$ counterclockwise by 90°, that is, to the left. When we travel along a positively oriented boundary in the direction of its orientation, the domain is "on our left." For example, if U is a bounded domain with "no holes," that is, ∂U is connected, then the positive orientation means we are traveling counterclockwise around ∂U. If we do have "holes," then we travel around them clockwise.

*Named after the British mathematical physicist George Green (1793–1841).

Proposition 10.6.3. *Let $U \subset \mathbb{R}^2$ be a bounded domain with piecewise smooth boundary. Then U is Jordan measurable.*

Proof. We must show that ∂U is a null set. As ∂U is a finite union of piecewise smooth paths, which are finite unions of smooth paths, we need only show that a smooth path in $\mathbb{R}^2$ is a null set. Let $\gamma \colon [a,b] \to \mathbb{R}^2$ be a smooth path. It is enough to show that $\gamma\big((a,b)\big)$ is a null set, as adding the points $\gamma(a)$ and $\gamma(b)$, to a null set still results in a null set. Define

$$f \colon (a,b) \times (-1,1) \to \mathbb{R}^2, \qquad \text{as} \qquad f(x,y) := \gamma(x).$$

The set $(a,b) \times \{0\}$ is a null set in $\mathbb{R}^2$ and $\gamma\big((a,b)\big) = f\big((a,b) \times \{0\}\big)$. By Proposition 10.3.10, $\gamma\big((a,b)\big)$ is a null set in $\mathbb{R}^2$ and so $\gamma\big([a,b]\big)$ is a null set, and so finally ∂U is a null set. □

Theorem 10.6.4 (Green)**.** *Suppose $U \subset \mathbb{R}^2$ is a bounded domain with piecewise smooth boundary with the boundary positively oriented. Suppose P and Q are continuously differentiable functions defined on some open set that contains the closure $\overline{U}$. Then*

$$\int_{\partial U} P\,dx + Q\,dy = \int_U \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right).$$

We stated Green's theorem in general, although we will only prove a special version of it. That is, we will only prove it for a special kind of domain. The general version follows from the special case by application of further geometry, and cutting up the general domain into smaller domains on which to apply the special case.

Let $U \subset \mathbb{R}^2$ be a bounded domain with piecewise smooth boundary. We say U is of *type I* if there exist numbers $a < b$, and continuous functions $f \colon [a,b] \to \mathbb{R}$ and $g \colon [a,b] \to \mathbb{R}$, such that

$$U := \big\{ (x,y) \in \mathbb{R}^2 : a < x < b \text{ and } f(x) < y < g(x) \big\}.$$

Similarly, U is of *type II* if there exist numbers $c < d$, and continuous functions $h \colon [c,d] \to \mathbb{R}$ and $k \colon [c,d] \to \mathbb{R}$, such that

$$U := \big\{ (x,y) \in \mathbb{R}^2 : c < y < d \text{ and } h(y) < x < k(y) \big\}.$$

Finally, $U \subset \mathbb{R}^2$ is of *type III* if it is both of type I and type II. See Figure 10.15.

Common domains to apply Green's theorem to are rectangles and discs, and these are type III domains. We will only prove Green's theorem for type III domains.

Proof of Green's theorem for U of type III. Let f, g, h, k be the functions defined above. Using Proposition 10.5.8, U is Jordan measurable and as U is of type I, then

$$\begin{aligned}
\int_U \left(-\frac{\partial P}{\partial y} \right) &= \int_a^b \int_{g(x)}^{f(x)} \left(-\frac{\partial P}{\partial y}(x,y) \right) dy\,dx \\
&= \int_a^b \Big(-P\big(x, f(x)\big) + P\big(x, g(x)\big) \Big)\,dx \\
&= \int_a^b P\big(x, g(x)\big)\,dx - \int_a^b P\big(x, f(x)\big)\,dx.
\end{aligned}$$

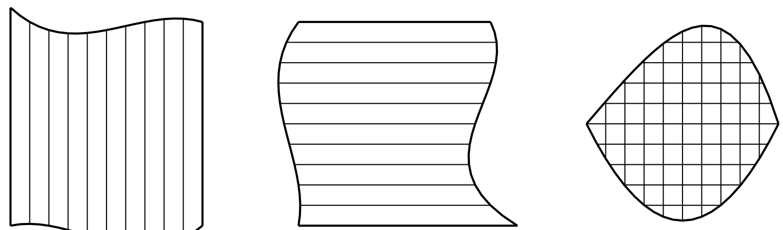

Figure 10.15: Domain types for Green's theorem.

We integrate $P\,dx$ along the boundary. The one-form $P\,dx$ integrates to zero along the straight vertical lines in the boundary. Therefore, it is only integrated along the top and along the bottom. As a parameter, x runs from left to right. If we use the parametrizations that take x to $\big(x, f(x)\big)$ and to $\big(x, g(x)\big)$, we recognize path integrals above. However, the second path integral is in the wrong direction; the top should be going right to left. After switching this orientation, we find

$$\int_{\partial U} P\,dx = \int_a^b P\big(x, g(x)\big)\,dx + \int_b^a P\big(x, f(x)\big)\,dx = \int_U \left(-\frac{\partial P}{\partial y}\right).$$

Similarly, U is also of type II. The form $Q\,dy$ integrates to zero along horizontal lines. So

$$\int_U \frac{\partial Q}{\partial x} = \int_c^d \int_{k(y)}^{h(y)} \frac{\partial Q}{\partial x}(x,y)\,dx\,dy = \int_a^b \Big(Q\big(y, h(y)\big) - Q\big(y, k(y)\big)\Big)\,dx = \int_{\partial U} Q\,dy.$$

Putting the two computations together, we obtain

$$\int_{\partial U} P\,dx + Q\,dy = \int_{\partial U} P\,dx + \int_{\partial U} Q\,dy = \int_U \left(-\frac{\partial P}{\partial y}\right) + \int_U \frac{\partial Q}{\partial x} = \int_U \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right). \qquad \square$$

Let us see how one can use the simple version of Green's (type III domains only) for a more complex path.

Example 10.6.5: Suppose $P(x,y) = \frac{-y}{x^2+y^2}$, $Q(x,y) = \frac{x}{x^2+y^2}$. If we think of (P,Q) as a vector, so that we have a so-called *vector field*, (P,Q) is called the *vortex vector field*, as it gives the velocity of particles traveling in a vortex around the origin. Variations on this vector field come up often in applications. Suppose that γ is a path that goes counterclockwise around a rectangle whose interior contains the origin. We claim

$$\int_\gamma \frac{-y}{x^2+y^2}\,dx + \frac{x}{x^2+y^2}\,dy = 2\pi.$$

First we draw a circle C of radius $r > 0$ centered at the origin such that the entire circle is within γ and oriented clockwise. Consider U to be the domain between γ and C. See

Figure 10.16. The integral around ∂U is the integral around γ plus the integral around C. Now U is not a domain of type III, so we cannot just apply the version of Green's theorem we actually proved. However, if we cut the box along the axis as shown in the figure with dashed lines, the four resulting domains, let us call them U_1, U_2, U_3, U_4, are of type III. The dashed lines are oriented in opposite directions for the two U_j that share them, and so when we integrate along both, the integrals cancel. That is,

$$\int_{\partial U} P\,dx + Q\,dy = \int_{\partial U_1} P\,dx + Q\,dy + \int_{\partial U_2} P\,dx + Q\,dy + \int_{\partial U_3} P\,dx + Q\,dy + \int_{\partial U_4} P\,dx + Q\,dy.$$

Now we are allowed to apply Green's theorem to every U_j. We leave it to the reader to verify that outside of the origin, $\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} = 0$. So for each j, we find

$$\int_{\partial U_j} P\,dx + Q\,dy = \int_{U_j} \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y} \right) = \int_{U_j} 0 = 0,$$

and so $\int_{\partial U} P\,dx + Q\,dy = 0$. As ∂U is C together with γ, we find

$$\int_C P\,dx + Q\,dy + \int_\gamma P\,dx + Q\,dy = \int_{\partial U} P\,dx + Q\,dy = 0.$$

So the integral around C is minus the integral around γ. The integral around C is easy to compute as on C we have $x^2 + y^2 = r^2$, so $P(x,y) = \frac{-y}{r^2}$ and $Q(x,y) = \frac{x}{r^2}$. We leave it to the reader to compute

$$\int_C P\,dx + Q\,dy = \int_C \frac{-y}{r^2}\,dx + \frac{x}{r^2}\,dy = -2\pi.$$

The claim follows.

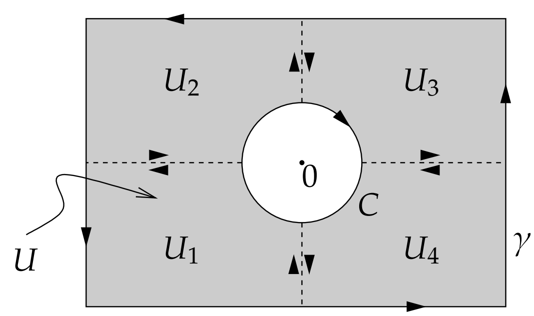

Figure 10.16: Changing the box integral to an integral around a small circle around the origin. The domain U is the entire shaded area between the circle and the box.

We remark that if γ would not contain the origin, $\int_\gamma P\,dx + Q\,dy = 0$, as we could just apply Green's to γ. So this integral can detect whether the origin is inside γ or not.

As a second example, we illustrate the usefulness of Green's theorem on a fundamental result about harmonic functions.

Example 10.6.6: Suppose $U \subset \mathbb{R}^2$ is open and $f : U \to \mathbb{R}$ is harmonic, that is, f is twice continuously differentiable and satisfies the *Laplace equation*, $\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} = 0$. Harmonic functions are, for instance, the steady state heat distribution, or the electric potential between charges. We will prove one of the most fundamental properties of these functions.

Let $D_r := B(p,r)$ be a disc such that its closure $\overline{D_r} = C(p,r) \subset U$. Write $p = (x_0, y_0)$. We orient ∂D_r positively. See Exercise 10.6.1. Then via Green's and differentiation under the integral,

$$
\begin{aligned}
0 &= \frac{1}{2\pi r} \int_{D_r} \left(\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \right) \\
&= \frac{1}{2\pi r} \int_{\partial D_r} -\frac{\partial f}{\partial y}\, dx + \frac{\partial f}{\partial x}\, dy \\
&= \frac{1}{2\pi r} \int_0^{2\pi} \Bigg(-\frac{\partial f}{\partial y}\big(x_0 + r\cos(t), y_0 + r\sin(t)\big)\big(-r\sin(t)\big) \\
&\qquad\qquad + \frac{\partial f}{\partial x}\big(x_0 + r\cos(t), y_0 + r\sin(t)\big)\, r\cos(t) \Bigg)\, dt \\
&= \frac{d}{dr} \left[\frac{1}{2\pi} \int_0^{2\pi} f\big(x_0 + r\cos(t), y_0 + r\sin(t)\big)\, dt \right].
\end{aligned}
$$

Let $g(r) := \frac{1}{2\pi} \int_0^{2\pi} f\big(x_0 + r\cos(t), y_0 + r\sin(t)\big)\, dt$ for $r \geq 0$ (small enough). The function is continuous at $r = 0$ (exercise), and we have just proved that $g'(r) = 0$ for all $r > 0$. Therefore, $g(0) = g(r)$ for all $r > 0$, and

$$
g(r) = g(0) = \frac{1}{2\pi} \int_0^{2\pi} f\big(x_0 + 0\cos(t), y_0 + 0\sin(t)\big)\, dt = f(x_0, y_0).
$$

We proved the *mean value property* of harmonic functions:

$$
f(x_0, y_0) = \frac{1}{2\pi} \int_0^{2\pi} f\big(x_0 + r\cos(t), y_0 + r\sin(t)\big)\, dt = \frac{1}{2\pi r} \int_{\partial D_r} f\, ds.
$$

That is, for a harmonic function, the value at $p = (x_0, y_0)$ equals the average of its values over a circle of any radius r centered at (x_0, y_0), provided r is small enough so that the entire closed disc fits within U.

10.6.1 Exercises

***Exercise* 10.6.1:** *Prove that a disc $B(p,r) \subset \mathbb{R}^2$ is a type III domain, and prove that the orientation given by the parametrization $\gamma(t) = \big(x_0 + r\cos(t), y_0 + r\sin(t)\big)$ where $p = (x_0, y_0)$ is the positive orientation of the boundary $\partial B(p,r)$.*
Note: Feel free to use what you know about sine and cosine from calculus.

***Exercise* 10.6.2:** *Prove that a convex bounded domain with piecewise smooth boundary is a type III domain.*

***Exercise* 10.6.3:** *Suppose $V \subset \mathbb{R}^2$ is a bounded domain with piecewise smooth boundary of type III and suppose that $U \subset \mathbb{R}^2$ is an open set such that $\overline{V} \subset U$. Suppose $f : U \to \mathbb{R}$ is a twice continuously differentiable function. Prove that $\int_{\partial V} \frac{\partial f}{\partial x}\, dx + \frac{\partial f}{\partial y}\, dy = 0$.*

***Exercise* 10.6.4:** *For a disc $B(p,r) \subset \mathbb{R}^2$, orient the boundary $\partial B(p,r)$ positively.*

a) Compute $\displaystyle\int_{\partial B(p,r)} -y\, dx$.

b) Compute $\displaystyle\int_{\partial B(p,r)} x\, dy$.

c) Compute $\displaystyle\int_{\partial B(p,r)} \frac{-y}{2}\, dx + \frac{x}{2}\, dy$.

***Exercise* 10.6.5:** *Using Green's theorem show that the area of a triangle with vertices (x_1, y_1), (x_2, y_2), (x_3, y_3) is $\frac{1}{2}\lvert x_1y_2 + x_2y_3 + x_3y_1 - y_1x_2 - y_2x_3 - y_3x_1\rvert$. Hint: See previous exercise.*

***Exercise* 10.6.6:** *Using the mean value property prove the* maximum principle *for harmonic functions: Suppose $U \subset \mathbb{R}^2$ is a connected open set and $f : U \to \mathbb{R}$ is harmonic. Prove that if f attains a maximum at $p \in U$, then f is constant.*

***Exercise* 10.6.7:** *Let $f(x,y) := \ln \sqrt{x^2 + y^2}$.*

a) Show f is harmonic where defined.

b) Show $\displaystyle\lim_{(x,y)\to 0} f(x,y) = -\infty$.

c) Using a circle C_r of radius r around the origin, compute $\frac{1}{2\pi r}\int_{\partial C_r} f\, ds$. What happens as $r \to 0$?

d) Why can't you use Green's theorem?

10.7 Change of variables

Note: 1 lecture

In one variable, we have the familiar change of variables

$$\int_a^b f\bigl(g(x)\bigr)g'(x)\,dx = \int_{g(a)}^{g(b)} f(x)\,dx.$$

The analogue in higher dimensions is quite a bit more complicated. The first complication is orientation. If we use the definition of integral from this chapter, then we do not have the notion of $\int_a^b$ versus $\int_b^a$. We are simply integrating over an interval $[a,b]$. With this notation, the change of variables becomes

$$\int_{[a,b]} f\bigl(g(x)\bigr)\,|g'(x)|\,dx = \int_{g([a,b])} f(x)\,dx.$$

In this section we will obtain the several-variable analogue of this form.

Let us remark the role of $|g'(x)|$ in the formula. The integral measures volumes in general, so in one dimension it measures length. Notice that $|g'(x)|$ scales the dx and so it scales the lengths. If our g is linear, that is, $g(x) = Lx$, then $g'(x) = L$ and the length of the interval $g([a,b])$ is simply $|L|(b-a)$. That is because $g([a,b])$ is either $[La, Lb]$ or $[Lb, La]$. This property holds in higher dimension with $|L|$ replaced by the absolute value of the determinant.

Proposition 10.7.1. *Suppose $R \subset \mathbb{R}^n$ is a rectangle and $A\colon \mathbb{R}^n \to \mathbb{R}^n$ is linear. Then $A(R)$ is Jordan measurable and $V\bigl(A(R)\bigr) = |\det(A)|\,V(R)$.*

Proof. It is enough to prove for elementary matrices. The proof is left as an exercise. □

Let us prove that absolute value of the Jacobian determinant $|J_g(x)| = \bigl|\det\bigl(g'(x)\bigr)\bigr|$ is the replacement of $|g'(x)|$ for multiple dimensions in the change of variables formula. The following theorem holds in more generality, but this statement is sufficient for many uses.

Theorem 10.7.2. *Suppose $U \subset \mathbb{R}^n$ is open, $S \subset U$ is a compact Jordan measurable set, and $g\colon U \to \mathbb{R}^n$ is a one-to-one continuously differentiable mapping, such that J_g is never zero on S. Suppose $f\colon g(S) \to \mathbb{R}$ is Riemann integrable. Then $f \circ g$ is Riemann integrable on S and*

$$\int_{g(S)} f(x)\,dx = \int_S f\bigl(g(x)\bigr)\,|J_g(x)|\,dx.$$

The set $g(S)$ is Jordan measurable by Proposition 10.5.9, so the left-hand side does make sense. That the right-hand side makes sense follows by Corollary 10.4.4 (actually Exercise 10.5.7).

Proof. The set S can be covered by finitely many closed rectangles $P_1, P_2, \ldots, P_k$, whose interiors do not overlap such that each $P_j \subset U$ (Exercise 10.7.2). Proving the theorem for $P_j \cap S$ instead of S is enough. Define $f(y) := 0$ for all $y \notin g(S)$. The new f is still Riemann integrable since $g(S)$ is Jordan measurable. We can now replace the integrals over S with integrals over the whole rectangle. We therefore assume that S is equal to a rectangle R.

Let $\epsilon > 0$ be given. For every $x \in R$, let

$$W_x := \left\{ y \in U : \|g'(x) - g'(y)\| < \epsilon/2 \right\}.$$

By Exercise 10.7.3, W_x is open. As $x \in W_x$ for every x, it is an open cover. By the Lebesgue covering lemma (Lemma 7.4.10 from volume I), there exists a $\delta > 0$ such that for every $y \in R$, there is an x such that $B(y, \delta) \subset W_x$. In other words, if P is a rectangle of maximum side length less than $\frac{\delta}{\sqrt{n}}$ and $y \in P$, then $P \subset B(y, \delta) \subset W_x$. By triangle inequality, $\|g'(\xi) - g'(\eta)\| < \epsilon$ for all $\xi, \eta \in P$.

Let $R_1, R_2, \ldots, R_N$ be subrectangles partitioning R such that the maximum side of every R_j is less than $\frac{\delta}{\sqrt{n}}$. We also make sure that the minimum side length is at least $\frac{\delta}{2\sqrt{n}}$, which we can do if δ is sufficiently small relative to the sides of R (Exercise 10.7.4).

Consider some R_j and some fixed $x_j \in R_j$. First suppose $x_j = 0$, $g(0) = 0$, and $g'(0) = I$. For any given $y \in R_j$, apply the fundamental theorem of calculus to the function $t \mapsto g(ty)$ to find $g(y) = \int_0^1 g'(ty)y\,dt$. As the side of R_j is at most $\frac{\delta}{\sqrt{n}}$, then $\|y\| \leq \delta$. So

$$\|g(y) - y\| = \left\| \int_0^1 \left(g'(ty)y - y \right) dt \right\| \leq \int_0^1 \|g'(ty)y - y\|\,dt \leq \|y\| \int_0^1 \|g'(ty) - I\|\,dt \leq \delta\epsilon.$$

Therefore, $g(R_j) \subset \widetilde{R}_j$, where $\widetilde{R}_j$ is a rectangle obtained from R_j by extending by $\delta\epsilon$ on all sides. See Figure 10.17.

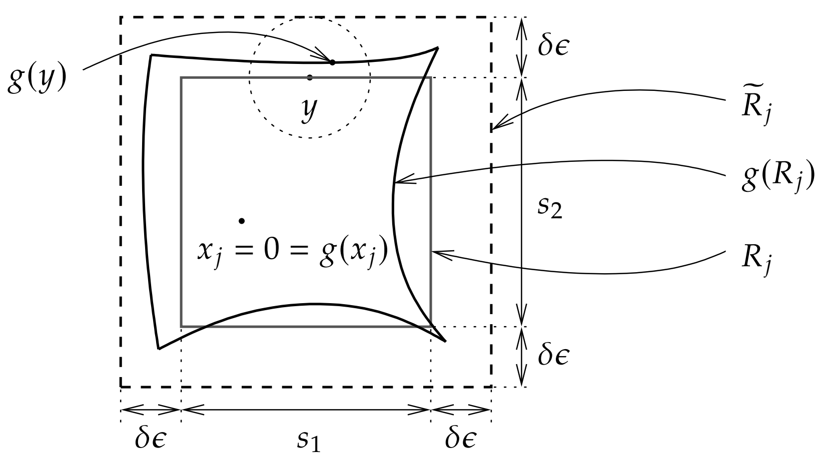

Figure 10.17: Image of R_j under g lies inside $\widetilde{R}_j$. A sample point $y \in R_j$ (on the boundary of R_j in fact) is marked and $g(y)$ must lie within with a radius of $\delta\epsilon$ (also marked).

If the sides of R_j are $s_1, s_2, \ldots, s_n$, then $V(R_j) = s_1 s_2 \cdots s_n$. Recall $\delta \leq 2\sqrt{n}\, s_j$. Thus,

$$\begin{aligned}
V(\widetilde{R}_j) &= (s_1 + 2\delta\epsilon)(s_2 + 2\delta\epsilon)\cdots(s_n + 2\delta\epsilon) \\
&\leq (s_1 + 4\sqrt{n}\, s_1\epsilon)(s_2 + 4\sqrt{n}\, s_2\epsilon)\cdots(s_n + 4\sqrt{n}\, s_n\epsilon) \\
&= s_1(1 + 4\sqrt{n}\,\epsilon)\, s_2(1 + 4\sqrt{n}\,\epsilon)\cdots s_n(1 + 4\sqrt{n}\,\epsilon) = V(R_j)\,(1 + 4\sqrt{n}\,\epsilon)^n.
\end{aligned}$$

In other words,

$$V\big(g(R_j)\big) \leq V(\widetilde{R}_j) \leq V(R_j)\,(1 + 4\sqrt{n}\,\epsilon)^n.$$

Next, suppose $A := g'(0)$ is not necessarily the identity. Write $g = A \circ \widetilde{g}$ where $\widetilde{g}'(0) = I$. By Proposition 10.7.1, $V\big(A(R_j)\big) = |\det(A)|\, V(R_j)$, and hence

$$\begin{aligned}
V\big(g(R_j)\big) &\leq |\det(A)|\, V(R_j)\,(1 + 4\sqrt{n}\,\epsilon)^n \\
&= |J_g(0)|\, V(R_j)\,(1 + 4\sqrt{n}\,\epsilon)^n.
\end{aligned}$$

Translation does not change volume, and therefore for every R_j, and $x_j \in R_j$, including when $x_j \neq 0$ and $g(x_j) \neq 0$, we find

$$V\big(g(R_j)\big) \leq |J_g(x_j)|\, V(R_j)\,(1 + 4\sqrt{n}\,\epsilon)^n.$$

Write f as $f = f_+ - f_-$ for two nonnegative Riemann integrable functions f_+ and f_-:

$$f_+(x) := \max\big\{f(x), 0\big\}, \qquad f_-(x) := \max\big\{-f(x), 0\big\}.$$

So, if we prove the theorem for a nonnegative f, we obtain the theorem for arbitrary f. Therefore, suppose that $f(y) \geq 0$ for all $y \in R$.

For a small enough $\delta > 0$, we have

$$\begin{aligned}
\epsilon + \int_R f\big(g(x)\big)\,|J_g(x)|\,dx &\geq \sum_{j=1}^{N} \left(\sup_{x\in R_j} f\big(g(x)\big)\,|J_g(x)|\right) V(R_j) \\
&\geq \sum_{j=1}^{N} \left(\sup_{x\in R_j} f\big(g(x)\big)\right) |J_g(x_j)|\, V(R_j) \\
&\geq \sum_{j=1}^{N} \left(\sup_{y\in g(R_j)} f(y)\right) V\big(g(R_j)\big) \frac{1}{(1 + 4\sqrt{n}\,\epsilon)^n} \\
&\geq \sum_{j=1}^{N} \left(\int_{g(R_j)} f(y)\,dy\right) \frac{1}{(1 + 4\sqrt{n}\,\epsilon)^n} \\
&= \frac{1}{(1 + 4\sqrt{n}\,\epsilon)^n} \int_{g(R)} f(y)\,dy.
\end{aligned}$$

The last equality follows because the overlaps of the rectangles are their boundaries, which are of measure zero, and hence the image of their boundaries is also measure zero. Let ϵ go to zero to find

$$\int_R f\big(g(x)\big)\,|J_g(x)|\,dx \geq \int_{g(R)} f(y)\,dy.$$

By adding this result for several rectangles covering an S we obtain the result for an arbitrary bounded Jordan measurable $S \subset U$, and nonnegative integrable function f:

$$\int_S f\big(g(x)\big)\,\lvert J_g(x)\rvert\,dx \geq \int_{g(S)} f(y)\,dy.$$

Recall that g^{-1} exists and $g^{-1}\big(g(S)\big) = S$. Also, $1 = J_{g\circ g^{-1}} = J_g\big(g^{-1}(y)\big)\,J_{g^{-1}}(y)$ for $y \in g(S)$. So

$$\begin{aligned}\int_{g(S)} f(y)\,dy &= \int_{g(S)} f\Big(g\big(g^{-1}(y)\big)\Big)\,\big\lvert J_g\big(g^{-1}(y)\big)\big\rvert\,\lvert J_{g^{-1}}(y)\rvert\,dy\\ &\geq \int_{g^{-1}(g(S))} f\big(g(x)\big)\,\lvert J_g(x)\rvert\,dx = \int_S f\big(g(x)\big)\,\lvert J_g(x)\rvert\,dx.\end{aligned}$$

The conclusion of the theorem holds for all nonnegative f and as we mentioned above, it thus holds for all Riemann integrable f. □

10.7.1 Exercises

***Exercise* 10.7.1:** *Prove Proposition 10.7.1.*

***Exercise* 10.7.2:** *Suppose $U \subset \mathbb{R}^n$ is open and $S \subset U$ is a compact Jordan measurable set. Show that there exist finitely many closed rectangles $P_1, P_2, \ldots, P_k$ such that $P_j \subset U$, $S \subset P_1 \cup P_2 \cup \cdots \cup P_k$, and the interiors are mutually disjoint, that is, $P_j^\circ \cap P_\ell^\circ = \emptyset$ whenever $j \neq \ell$.*

***Exercise* 10.7.3:** *Suppose $U \subset \mathbb{R}^n$ is open, $x \in U$, and $g \colon U \to \mathbb{R}^n$ is a continuously differentiable mapping. For every $\epsilon > 0$, show that*

$$W_x := \big\{ y \in U : \lVert g'(x) - g'(y) \rVert < \epsilon/2 \big\}$$

is an open set.

***Exercise* 10.7.4:** *Suppose $R \subset \mathbb{R}^n$ is a closed rectangle. Show that if $\delta' > 0$ is sufficiently small relative to the sides of R, then R can be partitioned into subrectangles where each side of every subrectangle is between $\frac{\delta'}{2}$ and δ'.*

***Exercise* 10.7.5:** *Prove the following version of the theorem:* Suppose $f \colon \mathbb{R}^n \to \mathbb{R}$ is a Riemann integrable compactly supported function. Suppose $K \subset \mathbb{R}^n$ is the support of f, S is a compact set, and $g \colon \mathbb{R}^n \to \mathbb{R}^n$ is a function that when restricted to a neighborhood U of S is one-to-one and continuously differentiable, $g(S) = K$ and J_g is never zero on S (in the formula assume $J_g(x) = 0$ if g not differentiable at x, that is, when $x \notin U$). Then

$$\int_{\mathbb{R}^n} f(x)\,dx = \int_{\mathbb{R}^n} f\big(g(x)\big)\,\lvert J_g(x)\rvert\,dx.$$

Exercise* 10.7.6*: *Prove the following version of the theorem:* Suppose $S \subset \mathbb{R}^n$ is an open bounded Jordan measurable set, $g \colon S \to \mathbb{R}^n$ is a one-to-one continuously differentiable mapping such that J_g is never zero on S, and such that $g(S)$ is bounded and Jordan measurable (it is also open). Suppose $f \colon g(S) \to \mathbb{R}$ is Riemann integrable. Then $f \circ g$ is Riemann integrable on S and

$$\int_{g(S)} f(x)\, dx = \int_S f\bigl(g(x)\bigr)\, \lvert J_g(x) \rvert \, dx.$$

Hint: Write S as an increasing union of compact Jordan measurable sets, then apply the theorem of the section to those. Then prove that you can take the limit.

Chapter 11

Functions as Limits

11.1 Complex numbers

Note: half a lecture

11.1.1 The complex plane

In this chapter we consider approximation of functions, or in other words functions as limits of sequences and series. We will extend some results we already saw to a somewhat more general setting, and we will look at some completely new results. In particular, we consider complex-valued functions. We gave complex numbers as examples before, but let us start from scratch and properly define the complex number field.

A complex number is just a pair $(x, y) \in \mathbb{R}^2$ on which we define multiplication (see below). We call the set the *complex numbers* and denote it by $\mathbb{C}$. We identify $x \in \mathbb{R}$ with $(x, 0) \in \mathbb{C}$. The x-axis is then called the *real axis* and the y-axis is called the *imaginary axis*. As $\mathbb{C}$ is just the plane, we also call the set $\mathbb{C}$ the *complex plane*.

Define:

$$(x, y) + (s, t) := (x + s, y + t), \qquad (x, y)(s, t) := (xs - yt, xt + ys).$$

Under the identification above, we have $0 = (0, 0)$ and $1 = (1, 0)$. These two operations make the plane into a field (exercise). We write a complex number (x, y) as $x + iy$, where we define*

$$i := (0, 1).$$

Notice that $i^2 = (0, 1)(0, 1) = (0 - 1, 0 + 0) = -1$. That is, i is a solution to the polynomial equation

$$z^2 + 1 = 0.$$

From now on, we will not use the notation (x, y) and use only $x + iy$. See Figure 11.1.

*Note that engineers use j instead of i.

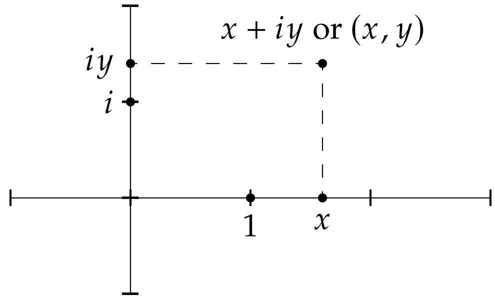

Figure 11.1: The points 1, i, x, iy, and $x + iy$ in the complex plane.

We generally use x, y, r, s, t for real values and z, w, ξ, ζ for complex values, although that is not a hard and fast rule. In particular, z is often used as a third real variable in $\mathbb{R}^3$.

Definition 11.1.1. Suppose $z = x + iy$. We call x the *real part* of z, and we call y the *imaginary part* of z. We write

$$\operatorname{Re} z := x, \qquad \operatorname{Im} z := y.$$

Define *complex conjugate* as

$$\bar{z} := x - iy,$$

and define *modulus* as

$$|z| := \sqrt{x^2 + y^2}.$$

Modulus is the complex analogue of the absolute value and has similar properties. For example, $|zw| = |z|\,|w|$ (exercise). The complex conjugate is a reflection of the plane across the real axis. The real numbers are precisely those numbers for which the imaginary part $y = 0$. In particular, they are precisely those numbers which satisfy the equation

$$z = \bar{z}.$$

As $\mathbb{C}$ is really $\mathbb{R}^2$, we let the metric on $\mathbb{C}$ be the standard euclidean metric on $\mathbb{R}^2$. In particular,

$$|z| = d(z, 0), \qquad \text{and also} \qquad |z - w| = d(z, w).$$

So the topology on $\mathbb{C}$ is the same exact topology as the standard topology on $\mathbb{R}^2$ with the euclidean metric, and $|z|$ is equal to the euclidean norm on $\mathbb{R}^2$. Importantly, since $\mathbb{R}^2$ is a complete metric space, then so is $\mathbb{C}$. As $|z|$ is the euclidean norm on $\mathbb{R}^2$, we have the *triangle inequality* of both flavors:

$$|z + w| \leq |z| + |w| \qquad \text{and} \qquad \big||z| - |w|\big| \leq |z - w|.$$

The complex conjugate and the modulus are even more intimately related:

$$|z|^2 = x^2 + y^2 = (x + iy)(x - iy) = z\bar{z}.$$

Remark 11.1.2. There is no natural ordering on the complex numbers. In particular, no ordering that makes the complex numbers into an ordered field. Ordering is one of the things we lose when we go from real to complex numbers.

11.1.2 Complex numbers and limits

Algebraic operations with complex numbers are continuous because convergence in $\mathbb{R}^2$ is the same as convergence for each component, and we already know that the real algebraic operations are continuous. For example, write $z_n = x_n + i\, y_n$ and $w_n = s_n + i\, t_n$, and suppose that $\lim_{n\to\infty} z_n = z = x + i\, y$ and $\lim_{n\to\infty} w_n = w = s + i\, t$. Let us show

$$\lim_{n\to\infty} z_n w_n = zw.$$

First,

$$z_n w_n = (x_n s_n - y_n t_n) + i(x_n t_n + y_n s_n).$$

The topology on $\mathbb{C}$ is the same as on $\mathbb{R}^2$, and so $x_n \to x$, $y_n \to y$, $s_n \to s$, and $t_n \to t$. Hence,

$$\lim_{n\to\infty} (x_n s_n - y_n t_n) = xs - yt \qquad \text{and} \qquad \lim_{n\to\infty} (x_n t_n + y_n s_n) = xt + ys.$$

As $(xs - yt) + i(xt + ys) = zw$,

$$\lim_{n\to\infty} z_n w_n = zw.$$

Similarly the modulus and the complex conjugate are continuous functions. We leave the remainder of the proof of the following proposition as an exercise.

Proposition 11.1.3. *Suppose $\{z_n\}_{n=1}^{\infty}$, $\{w_n\}_{n=1}^{\infty}$ are sequences of complex numbers converging to z and w respectively. Then*

(*i*) $\lim\limits_{n\to\infty} z_n + w_n = z + w.$

(*ii*) $\lim\limits_{n\to\infty} z_n w_n = zw.$

(*iii*) *Assuming $w_n \neq 0$ for all n and $w \neq 0$,* $\lim\limits_{n\to\infty} \dfrac{z_n}{w_n} = \dfrac{z}{w}.$

(*iv*) $\lim\limits_{n\to\infty} |z_n| = |z|.$

(*v*) $\lim\limits_{n\to\infty} \bar{z}_n = \bar{z}.$

As we have seen above, convergence in $\mathbb{C}$ is the same as convergence in $\mathbb{R}^2$. In particular, a sequence in $\mathbb{C}$ converges if and only if the real and imaginary parts converge. Therefore, feel free to apply everything you have learned about convergence in $\mathbb{R}^2$, as well as applying results about real numbers to the real and imaginary parts.

We also need convergence of complex series. Let $\{z_n\}_{n=1}^{\infty}$ be a sequence of complex numbers. The series

$$\sum_{n=1}^{\infty} z_n$$

converges if the limit of partial sums converges, that is, if

$$\lim_{k\to\infty} \sum_{n=1}^{k} z_n \qquad \text{exists.}$$

A series *converges absolutely* if $\sum_{n=1}^{\infty} |z_n|$ converges.

We say a series is *Cauchy* if the sequence of partial sums is Cauchy. The following two propositions have essentially the same proofs as for real series and we leave them as exercises.

Proposition 11.1.4. *The complex series $\sum_{n=1}^{\infty} z_n$ is Cauchy if for every $\epsilon > 0$, there exists an $M \in \mathbb{N}$ such that for every $n \geq M$ and every $k > n$, we have*

$$\left| \sum_{j=n+1}^{k} z_j \right| < \epsilon.$$

Proposition 11.1.5. *If a complex series $\sum_{n=1}^{\infty} z_n$ converges absolutely, then it converges.*

The series $\sum_{n=1}^{\infty} |z_n|$ is a real series. All the convergence tests (ratio test, root test, etc.) that talk about absolute convergence work with the numbers $|z_n|$, that is, they are really talking about convergence of series of nonnegative real numbers. You can directly apply these tests them without needing to reprove anything for complex series.

11.1.3 Complex-valued functions

When we deal with complex-valued functions $f \colon X \to \mathbb{C}$, what we often do is to write $f = u + i\,v$ for real-valued functions $u \colon X \to \mathbb{R}$ and $v \colon X \to \mathbb{R}$.

Suppose we wish to integrate $f \colon [a,b] \to \mathbb{C}$. We write $f = u + i\,v$ for real-valued u and v. We say that f is *Riemann integrable* if u and v are Riemann integrable, and in this case we define

$$\int_a^b f := \int_a^b u + i \int_a^b v.$$

We make the same definition for every other type of integral (improper, multivariable, etc.).

Similarly when we differentiate, write $f \colon [a,b] \to \mathbb{C}$ as $f = u + i\,v$. Thinking of $\mathbb{C}$ as $\mathbb{R}^2$, we say that f is differentiable if u and v are differentiable. For a function valued in $\mathbb{R}^2$, the derivative is represented by a vector in $\mathbb{R}^2$. Now a vector in $\mathbb{R}^2$ is a complex number. In other words, we write the *derivative* as

$$f'(t) := u'(t) + i\,v'(t).$$

The linear operator representing the derivative is the multiplication by the complex number $f'(t)$, so nothing is lost in this identification.

11.1.4 Exercises

***Exercise* 11.1.1:** *Check that $\mathbb{C}$ is a field.*

***Exercise* 11.1.2:** *Prove that for $z, w \in \mathbb{C}$, we have $|zw| = |z|\,|w|$.*

***Exercise* 11.1.3:** *Finish the proof of Proposition 11.1.3.*

***Exercise* 11.1.4:** *Prove Proposition 11.1.4.*

***Exercise* 11.1.5:** *Prove Proposition 11.1.5.*

***Exercise* 11.1.6:** *Given $x + iy$ define the matrix $\begin{bmatrix} x & -y \\ y & x \end{bmatrix}$. Prove:*

a) *The action of this matrix on a vector (s, t) is the same as the action of multiplying $(x + iy)(s + it)$.*

b) *Multiplying two such matrices is the same multiplying the underlying complex numbers and then finding the corresponding matrix for the product. In other words, the field $\mathbb{C}$ can be identified with a subset of the 2-by-2 matrices.*

c) *The matrix $\begin{bmatrix} x & -y \\ y & x \end{bmatrix}$ has eigenvalues $x + iy$ and $x - iy$. Recall that λ is an eigenvalue of a matrix A if $A - \lambda I$ (a complex matrix in our case) is not invertible, that is, if it has linearly dependent rows: one row is a (complex) multiple of the other.*

***Exercise* 11.1.7:** *Prove the Bolzano–Weierstrass theorem for complex sequences. Suppose $\{z_n\}_{n=1}^{\infty}$ is a bounded sequence of complex numbers, that is, there exists an M such that $|z_n| \leq M$ for all n. Prove that there exists a subsequence $\{z_{n_k}\}_{k=1}^{\infty}$ that converges to some $z \in \mathbb{C}$.*

***Exercise* 11.1.8:**

a) *Prove that there is no simple mean value theorem for complex-valued functions: Find a differentiable function $f \colon [0, 1] \to \mathbb{C}$ such that $f(0) = f(1) = 0$, but $f'(t) \neq 0$ for all $t \in [0, 1]$.*

b) *However, there is a weaker form of the mean value theorem as there is for vector-valued functions. Prove: If $f \colon [a, b] \to \mathbb{C}$ is continuous and differentiable in (a, b), and for some M, $|f'(x)| \leq M$ for all $x \in (a, b)$, then $|f(b) - f(a)| \leq M|b - a|$.*

***Exercise* 11.1.9:** *Prove that there is no simple mean value theorem for integrals for complex-valued functions: Find a continuous function $f \colon [0, 1] \to \mathbb{C}$ such that $\int_0^1 f = 0$ but $f(t) \neq 0$ for all $t \in [0, 1]$.*

11.2 Swapping limits

Note: 2 lectures

11.2.1 Continuity

Let us get back to swapping limits and expand on chapter 6 of volume I. Let $\{f_n\}_{n=1}^{\infty}$ be a sequence of functions $f_n \colon X \to Y$ for a set X and a metric space Y. Let $f \colon X \to Y$ be a function and for every $x \in X$, suppose

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

We say the sequence $\{f_n\}_{n=1}^{\infty}$ *converges pointwise* to f.

For $Y = \mathbb{C}$, a series of functions *converges pointwise* to f if for every $x \in X$, we have

$$f(x) = \lim_{n\to\infty} \sum_{k=1}^{n} f_k(x) = \sum_{k=1}^{\infty} f_k(x).$$

The question is: If f_n are all continuous, is f continuous? Differentiable? Integrable? What are the derivatives or integrals of f? For example, for continuity of the pointwise limit of a sequence of functions $\{f_n\}_{n=1}^{\infty}$, we are asking if

$$\lim_{x\to x_0} \lim_{n\to\infty} f_n(x) \overset{?}{=} \lim_{n\to\infty} \lim_{x\to x_0} f_n(x).$$

A priori, we do not even know if both sides exist, let alone if they equal each other.

Example 11.2.1: The functions $f_n \colon \mathbb{R} \to \mathbb{R}$,

$$f_n(x) := \frac{1}{1+nx^2},$$

are continuous and converge pointwise to the discontinuous function

$$f(x) := \begin{cases} 1 & \text{if } x = 0, \\ 0 & \text{else.} \end{cases}$$

So pointwise convergence is not enough to preserve continuity (nor even boundedness). For that, we need uniform convergence. Let $f_n \colon X \to Y$ be functions. Then $\{f_n\}_{n=1}^{\infty}$ *converges uniformly* to f if for every $\epsilon > 0$, there exists an M such that for all $n \geq M$ and all $x \in X$, we have

$$d\big(f_n(x), f(x)\big) < \epsilon.$$

A series $\sum_{n=1}^{\infty} f_n$ of complex-valued functions converges uniformly if the sequence of partial sums converges uniformly, that is, if for every $\epsilon > 0$, there exists an M such that for all $n \geq M$ and all $x \in X$,

$$\left| \left(\sum_{k=1}^{n} f_k(x) \right) - f(x) \right| < \epsilon.$$

The simplest property preserved by uniform convergence is boundedness. We leave the proof of the following proposition as an exercise. It is almost identical to the proof for real-valued functions.

Proposition 11.2.2. *Let X be a set and (Y, d) a metric space. If $f_n \colon X \to Y$ are bounded functions and converge uniformly to $f \colon X \to Y$, then f is bounded.*

If X is a set and (Y, d) is a metric space, then a sequence $f_n \colon X \to Y$ is *uniformly Cauchy* if for every $\epsilon > 0$, there is an M such that for all $n, m \geq M$ and all $x \in X$, we have

$$d\big(f_n(x), f_m(x)\big) < \epsilon.$$

The notion is the same as for real-valued functions. The proof of the following proposition is again essentially the same as in that setting and is left as an exercise.

Proposition 11.2.3. *Let X be a set, (Y, d) be a metric space, and $f_n \colon X \to Y$ be functions. If $\{f_n\}_{n=1}^{\infty}$ converges uniformly, then $\{f_n\}_{n=1}^{\infty}$ is uniformly Cauchy. Conversely, if $\{f_n\}_{n=1}^{\infty}$ is uniformly Cauchy and (Y, d) is Cauchy-complete, then $\{f_n\}_{n=1}^{\infty}$ converges uniformly.*

For $f \colon X \to \mathbb{C}$, we write

$$\|f\|_X := \sup_{x \in X} |f(x)|.$$

We call $\|\cdot\|_X$ the *supremum norm* or *uniform norm*, and the subscript denotes the set over which the supremum is taken. Then a sequence of functions $f_n \colon X \to \mathbb{C}$ converges uniformly to $f \colon X \to \mathbb{C}$ if and only if

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

The supremum norm satisfies the triangle inequality: For every $x \in X$,

$$|f(x) + g(x)| \leq |f(x)| + |g(x)| \leq \|f\|_X + \|g\|_X.$$

Take a supremum on the left to get

$$\|f + g\|_X \leq \|f\|_X + \|g\|_X.$$

For a compact metric space X, the uniform norm is a norm on the vector space $C(X, \mathbb{C})$. We leave it as an exercise. While we will not need it, $C(X, \mathbb{C})$ is in fact a complex vector space, that is, in the definition of a vector space we can replace $\mathbb{R}$ with $\mathbb{C}$. Convergence in the metric space $C(X, \mathbb{C})$ is uniform convergence.

We will study a couple of types of series of functions, and a useful test for uniform convergence of a series is the so-called *Weierstrass M-test*.

Theorem 11.2.4 (Weierstrass M-test)**.** *Let X be a set. Suppose $f_n : X \to \mathbb{C}$ are functions and $M_n > 0$ numbers such that*

$$|f_n(x)| \leq M_n \quad \textit{for all } x \in X, \qquad \textit{and} \qquad \sum_{n=1}^{\infty} M_n \quad \textit{converges}.$$

Then

$$\sum_{n=1}^{\infty} f_n(x) \quad \textit{converges uniformly}.$$

Another way to state the theorem is to say that if $\sum_{n=1}^{\infty} \|f_n\|_X$ converges, then $\sum_{n=1}^{\infty} f_n$ converges uniformly. Note that the converse of this theorem is not true. Applying the theorem to $\sum_{n=1}^{\infty} |f_n(x)|$, we see that this series also converges uniformly. So the series converges both absolutely and uniformly.

Proof. Suppose $\sum_{n=1}^{\infty} M_n$ converges. Given $\epsilon > 0$, we have that the partial sums of $\sum_{n=1}^{\infty} M_n$ are Cauchy so there is an N such that for all $m, n \geq N$ with $m \geq n$, we have

$$\sum_{k=n+1}^{m} M_k < \epsilon.$$

We estimate a Cauchy difference of the partial sums of the functions

$$\left| \sum_{k=n+1}^{m} f_k(x) \right| \leq \sum_{k=n+1}^{m} |f_k(x)| \leq \sum_{k=n+1}^{m} M_k < \epsilon.$$

The series converges by Proposition 11.1.4. The convergence is uniform, as N does not depend on x. Indeed, for all $n \geq N$,

$$\left| \sum_{k=1}^{\infty} f_k(x) - \sum_{k=1}^{n} f_k(x) \right| \leq \left| \sum_{k=n+1}^{\infty} f_k(x) \right| \leq \epsilon. \qquad \square$$

Example 11.2.5: The series

$$\sum_{n=1}^{\infty} \frac{\sin(nx)}{n^2}$$

converges uniformly on $\mathbb{R}$. See Figure 11.2. This series is a Fourier series, and we will see more of these in a later section. Proof: The series converges uniformly because $\sum_{n=1}^{\infty} \frac{1}{n^2}$ converges and

$$\left| \frac{\sin(nx)}{n^2} \right| \leq \frac{1}{n^2}.$$

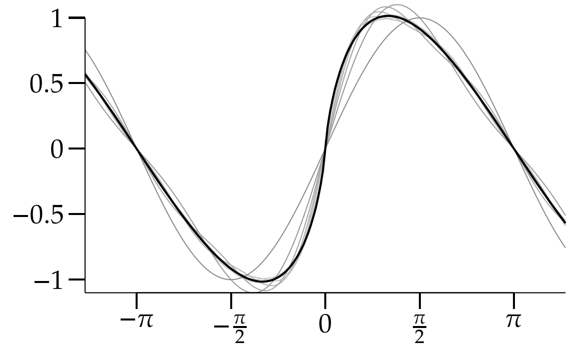

Figure 11.2: Plot of $\sum_{n=1}^{\infty} \frac{\sin(nx)}{n^2}$ including the first 8 partial sums in various shades of gray.

Example 11.2.6: The series

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

converges uniformly on every bounded interval. This series is a power series that we will study shortly. Proof: Take the interval $[-r,r] \subset \mathbb{R}$ (every bounded interval is contained in some $[-r,r]$). The series $\sum_{n=0}^{\infty} \frac{r^n}{n!}$ converges by the ratio test, so $\sum_{n=0}^{\infty} \frac{x^n}{n!}$ converges uniformly on $[-r,r]$ as

$$\left|\frac{x^n}{n!}\right| \leq \frac{r^n}{n!}.$$

Now we would love to say something about the limit. For example, is it continuous?

Proposition 11.2.7. *Let (X,d_X) and (Y,d_Y) be metric spaces, and suppose (Y,d_Y) is Cauchy-complete. Suppose $f_n \colon X \to Y$ converge uniformly to $f \colon X \to Y$. Let $\{x_k\}_{k=1}^{\infty}$ be a sequence in X and $x := \lim_{k\to\infty} x_k$. Suppose*

$$a_n := \lim_{k\to\infty} f_n(x_k)$$

exists for all n. Then $\{a_n\}_{n=1}^{\infty}$ converges and

$$\lim_{k\to\infty} f(x_k) = \lim_{n\to\infty} a_n.$$

In other words,

$$\lim_{k\to\infty} \lim_{n\to\infty} f_n(x_k) = \lim_{n\to\infty} \lim_{k\to\infty} f_n(x_k).$$

Proof. First we show that $\{a_n\}_{n=1}^{\infty}$ converges. As $\{f_n\}_{n=1}^{\infty}$ converges uniformly it is uniformly Cauchy. Let $\epsilon > 0$ be given. There is an M such that for all $m,n \geq M$, we have

$$d_Y\big(f_n(x_k), f_m(x_k)\big) < \epsilon \qquad \text{for all } k.$$

Note that $d_Y(a_n,a_m) \leq d_Y\big(a_n, f_n(x_k)\big) + d_Y\big(f_n(x_k), f_m(x_k)\big) + d_Y\big(f_m(x_k), a_m\big)$ and take the limit as $k \to \infty$ to find

$$d_Y(a_n,a_m) \leq \epsilon.$$

Hence $\{a_n\}_{n=1}^\infty$ is Cauchy and converges since Y is complete. Write $a := \lim_{k\to\infty} a_n$.

Find a $k \in \mathbb{N}$ such that

$$d_Y\bigl(f_k(p), f(p)\bigr) < \epsilon/3$$

for all $p \in X$. Assume k is large enough so that

$$d_Y(a_k, a) < \epsilon/3.$$

Find an $N \in \mathbb{N}$ such that for $m \geq N$,

$$d_Y\bigl(f_k(x_m), a_k\bigr) < \epsilon/3.$$

Then for $m \geq N$,

$$d_Y\bigl(f(x_m), a\bigr) \leq d_Y\bigl(f(x_m), f_k(x_m)\bigr) + d_Y\bigl(f_k(x_m), a_k\bigr) + d_Y\bigl(a_k, a\bigr) < \epsilon/3 + \epsilon/3 + \epsilon/3 = \epsilon. \qquad \square$$

We obtain an immediate corollary about continuity. If f_n are all continuous then $a_n = f_n(x)$ and so $\{a_n\}_{n=1}^\infty$ converges automatically to $f(x)$ and so we do not require completeness of Y.

Corollary 11.2.8. *Let X and Y be metric spaces. If $f_n \colon X \to Y$ are continuous functions such that $\{f_n\}_{n=1}^\infty$ converges uniformly to $f \colon X \to Y$, then f is continuous.*

The converse is not true. Just because the limit is continuous does not mean that the convergence is uniform. For example: $f_n \colon (0,1) \to \mathbb{R}$ defined by $f_n(x) := x^n$ converge to the zero function, but not uniformly. However, if we add extra conditions on the sequence, we can obtain a partial converse such as Dini's theorem, see Exercise 6.2.10 from volume I.

In Exercise 11.2.3 the reader is asked to prove that for a compact X, $C(X,\mathbb{C})$ is a normed vector space with the uniform norm, and hence a metric space. We have just shown that $C(X,\mathbb{C})$ is Cauchy-complete: Proposition 11.2.3 says that a Cauchy sequence in $C(X,\mathbb{C})$ converges uniformly to some function, and Corollary 11.2.8 shows that the limit is continuous and hence in $C(X,\mathbb{C})$.

Corollary 11.2.9. *Let (X,d) be a compact metric space. Then $C(X,\mathbb{C})$ is a Cauchy-complete metric space.*

Example 11.2.10: By Example 11.2.5 the Fourier series

$$\sum_{n=1}^\infty \frac{\sin(nx)}{n^2}$$

converges uniformly and hence is continuous by Corollary 11.2.8 (as is visible in Figure 11.2).

11.2.2 Integration

Proposition 11.2.11. *Suppose $f_n \colon [a,b] \to \mathbb{C}$ are Riemann integrable and suppose that $\{f_n\}_{n=1}^\infty$ converges uniformly to $f \colon [a,b] \to \mathbb{C}$. Then f is Riemann integrable and*

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

Since the integral of a complex-valued function is just the integral of the real and imaginary parts separately, the proof follows directly by the results of chapter 6 of volume I. We leave the details as an exercise.

Corollary 11.2.12. *Suppose $f_n \colon [a,b] \to \mathbb{C}$ are Riemann integrable and suppose that*

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

converges uniformly. Then the series is Riemann integrable on $[a,b]$ and

$$\int_a^b \sum_{n=1}^{\infty} f_n(x)\,dx = \sum_{n=1}^{\infty} \int_a^b f_n(x)\,dx$$

Example 11.2.13: Let us show how to integrate a Fourier series.

$$\int_0^x \sum_{n=1}^{\infty} \frac{\cos(nt)}{n^2}\,dt = \sum_{n=1}^{\infty} \int_0^x \frac{\cos(nt)}{n^2}\,dt = \sum_{n=1}^{\infty} \frac{\sin(nx)}{n^3}$$

The swapping of integral and sum is possible because of uniform convergence, which we have proved before using the Weierstrass M-test (Theorem 11.2.4).

We remark that we can swap integrals and limits under far less stringent hypotheses, but for that we would need a stronger integral than the Riemann integral. E.g. the Lebesgue integral.

11.2.3 Differentiation

Recall that a complex-valued function $f \colon [a,b] \to \mathbb{C}$, where $f(x) = u(x) + i\,v(x)$, is differentiable, if u and v are differentiable and the derivative is

$$f'(x) = u'(x) + i\,v'(x).$$

The proof of the following theorem is to apply the corresponding theorem for real functions to u and v, and is left as an exercise.

Theorem 11.2.14. *Let $I \subset \mathbb{R}$ be a bounded interval and let $f_n \colon I \to \mathbb{C}$ be continuously differentiable functions. Suppose $\{f_n'\}_{n=1}^{\infty}$ converges uniformly to $g \colon I \to \mathbb{C}$, and suppose $\{f_n(c)\}_{n=1}^{\infty}$ is a convergent sequence for some $c \in I$. Then $\{f_n\}_{n=1}^{\infty}$ converges uniformly to a continuously differentiable function $f \colon I \to \mathbb{C}$, and $f' = g$.*

Uniform limits of the functions themselves are not enough, and can make matters even worse. In §11.7 we will prove that continuous functions are uniform limits of polynomials, yet as the following example demonstrates, a continuous function need not be differentiable anywhere.

Example 11.2.15: There exist continuous nowhere differentiable functions. Such functions are often called *Weierstrass functions*, although this particular one, essentially due to Takagi*, is a different example than what Weierstrass gave. Define

$$\varphi(x) := |x| \qquad \text{for } x \in [-1,1].$$

Extend φ to all of $\mathbb{R}$ by making it 2-periodic: Decree that $\varphi(x) = \varphi(x+2)$. The function $\varphi\colon \mathbb{R} \to \mathbb{R}$ is continuous, in fact, $|\varphi(x) - \varphi(y)| \leq |x-y|$ (why?). See Figure 11.3.

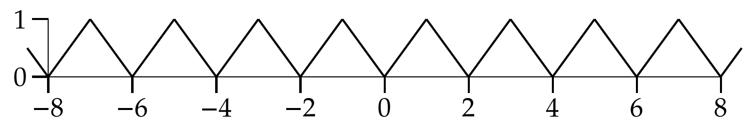

Figure 11.3: The 2-periodic function φ.

As $\sum_{n=0}^{\infty} \left(\frac{3}{4}\right)^n$ converges and $|\varphi(x)| \leq 1$ for all x, by the M-test (Theorem 11.2.4),

$$f(x) := \sum_{n=0}^{\infty} \left(\frac{3}{4}\right)^n \varphi(4^n x)$$

converges uniformly and hence is continuous. See Figure 11.4.

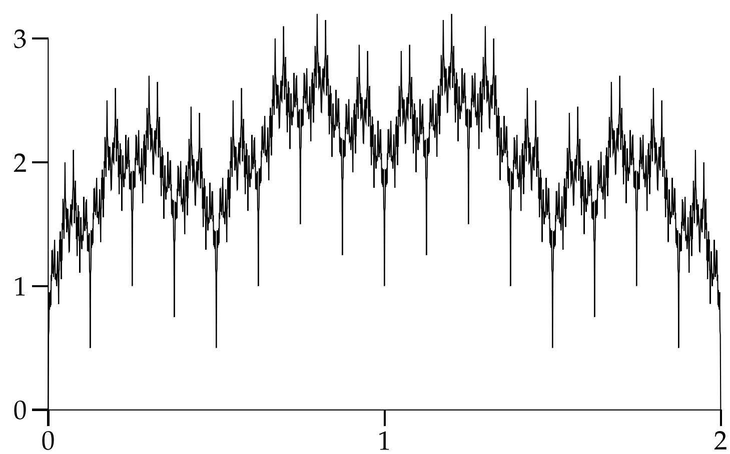

Figure 11.4: Plot of the nowhere differentiable function f.

We claim $f\colon \mathbb{R} \to \mathbb{R}$ is nowhere differentiable. Fix x, and we will show f is not differentiable at x. Define

$$\delta_m := \pm \frac{1}{2} 4^{-m},$$

where the sign is chosen so that there is no integer between $4^m x$ and $4^m(x+\delta_m) = 4^m x \pm \frac{1}{2}$.

*Teiji Takagi (1875–1960) was a Japanese mathematician.

We want to look at the difference quotient

$$\frac{f(x+\delta_m)-f(x)}{\delta_m} = \sum_{n=0}^{\infty} \left(\frac{3}{4}\right)^n \frac{\varphi\big(4^n(x+\delta_m)\big)-\varphi(4^n x)}{\delta_m}.$$

Fix m for a moment. Consider the expression inside the series:

$$\gamma_n := \frac{\varphi\big(4^n(x+\delta_m)\big)-\varphi(4^n x)}{\delta_m}.$$

If $n > m$, then $4^n\delta_m$ is an even integer. As φ is 2-periodic we get that $\gamma_n = 0$.

As there is no integer between $4^m(x+\delta_m) = 4^m x \pm 1/2$ and $4^m x$, then on this interval $\varphi(t) = \pm t + \ell$ for some integer ℓ. In particular, $\left|\varphi\big(4^m(x+\delta_m)\big)-\varphi(4^m x)\right| = |4^m x \pm 1/2 - 4^m x| = 1/2$. Therefore,

$$|\gamma_m| = \left|\frac{\varphi\big(4^m(x+\delta_m)\big)-\varphi(4^m x)}{\pm(1/2)4^{-m}}\right| = 4^m.$$

Similarly, suppose $n < m$. Since $|\varphi(s)-\varphi(t)| \leq |s-t|$,

$$|\gamma_n| = \left|\frac{\varphi\big(4^n x \pm (1/2)4^{n-m}\big)-\varphi(4^n x)}{\pm(1/2)4^{-m}}\right| \leq \left|\frac{\pm(1/2)4^{n-m}}{\pm(1/2)4^{-m}}\right| = 4^n.$$

And so

$$\begin{aligned}
\left|\frac{f(x+\delta_m)-f(x)}{\delta_m}\right| = \left|\sum_{n=0}^{\infty}\left(\frac{3}{4}\right)^n \gamma_n\right| &= \left|\sum_{n=0}^{m}\left(\frac{3}{4}\right)^n \gamma_n\right| \\
&\geq \left|\left(\frac{3}{4}\right)^m \gamma_m\right| - \left|\sum_{n=0}^{m-1}\left(\frac{3}{4}\right)^n \gamma_n\right| \\
&\geq 3^m - \sum_{n=0}^{m-1} 3^n = 3^m - \frac{3^m-1}{3-1} = \frac{3^m+1}{2}.
\end{aligned}$$

As $m \to \infty$, we have $\delta_m \to 0$, but $\frac{3^m+1}{2}$ goes to infinity. So f cannot be differentiable at x.

11.2.4 Exercises

***Exercise* 11.2.1:** *Prove Proposition 11.2.2.*

***Exercise* 11.2.2:** *Prove Proposition 11.2.3.*

***Exercise* 11.2.3:** *Suppose (X,d) is a compact metric space. Prove that the uniform norm $\|\cdot\|_X$ is a norm on the vector space of continuous complex-valued functions $C(X,\mathbb{C})$.*

***Exercise* 11.2.4:**

a) *Prove that* $f_n(x) := 2^{-n}\sin(2^n x)$ *converge uniformly to zero, but there exists a dense set* $D \subset \mathbb{R}$ *such that* $\lim_{n\to\infty} f_n'(x) = 1$ *for all* $x \in D$.

b) *Prove that* $\sum_{n=1}^{\infty} 2^{-n}\sin(2^n x)$ *converges uniformly to a continuous function, and there exists a dense set* $D \subset \mathbb{R}$ *where the derivatives of the partial sums do not converge.*

***Exercise* 11.2.5:** *Prove that* $\|f\|_{C^1} := \|f\|_{[a,b]} + \|f'\|_{[a,b]}$ *is a norm on the vector space of continuously differentiable complex-valued functions* $C^1\big([a,b],\mathbb{C}\big)$.

***Exercise* 11.2.6:** *Prove Theorem 11.2.14.*

***Exercise* 11.2.7:** *Prove Proposition 11.2.11 by reducing to the real result.*

***Exercise* 11.2.8:** *Work through the following counterexample to the converse of the Weierstrass M-test (Theorem 11.2.4). Define* $f_n \colon [0,1] \to \mathbb{R}$ *by*

$$f_n(x) := \begin{cases} \frac{1}{n} & \text{if } \frac{1}{n+1} < x < \frac{1}{n}, \\ 0 & \text{else.} \end{cases}$$

Prove that $\sum_{n=1}^{\infty} f_n$ *converges uniformly, but* $\sum_{n=1}^{\infty} \|f_n\|_{[0,1]}$ *does not converge.*

***Exercise* 11.2.9:** *Suppose* $f_n \colon [0,1] \to \mathbb{R}$ *are monotone increasing functions and suppose that* $\sum_{n=1}^{\infty} f_n$ *converges pointwise. Prove that* $\sum_{n=1}^{\infty} f_n$ *converges uniformly.*

***Exercise* 11.2.10:** *Prove that*

$$\sum_{n=1}^{\infty} e^{-nx}$$

converges for all $x > 0$ *to a differentiable function.*

11.3 Power series and analytic functions

Note: 2–3 lectures

11.3.1 Analytic functions

A (complex) power series is a series of the form

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

for $c_n, z, a \in \mathbb{C}$. We say the series *converges* if the series converges for some $z \neq a$.

Let $U \subset \mathbb{C}$ be an open set and $f : U \to \mathbb{C}$ a function. Suppose that for every $a \in U$ there exists a $\rho > 0$ and a power series convergent to the function

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

for all $z \in B(a, \rho)$. Then we say f is an *analytic* function. Similarly, given an interval $(a, b) \subset \mathbb{R}$, we say that $f : (a, b) \to \mathbb{C}$ is analytic or perhaps *real-analytic* if for each point $c \in (a, b)$ there is a power series around c that converges in some $(c - \rho, c + \rho)$ for some $\rho > 0$. As we will sometimes talk about real and sometimes about complex power series, we will use z to denote a complex number and x a real number. We will always mention which case we are working with.

An analytic function has different expansions around different points. Moreover, convergence does not automatically happen on the entire domain of the function. For example, if $|z| < 1$, then

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

While the left-hand side exists on all of $z \neq 1$, the right-hand side happens to converge only if $|z| < 1$. See a graph of a small piece of $\frac{1}{1-z}$ in Figure 11.5. We cannot graph the function itself, we can only graph its real or imaginary parts for lack of dimensions in our universe.

11.3.2 Convergence of power series

We proved several results for power series of a real variable in §2.6 of volume I. For the most part the convergence properties of power series deal with the series $\sum_{k=0}^{\infty} |c_k| \, |z-a|^k$ and so we have already proved many results about complex power series. In particular, we computed the so-called radius of convergence of a power series.

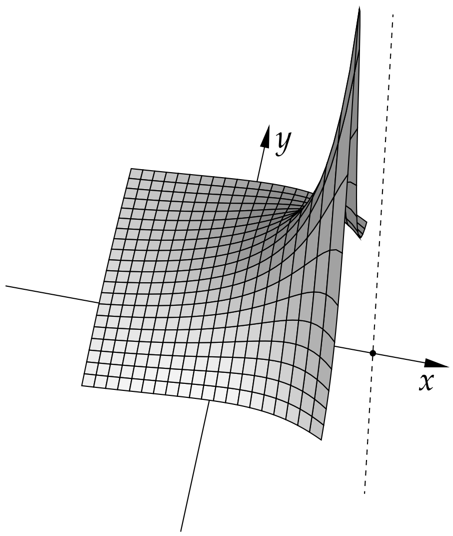

Figure 11.5: Graphs of the real and imaginary parts of $z = x+iy \mapsto \frac{1}{1-z}$ in the square $[-0.8, 0.8]^2$. The singularity at $z = 1$ is marked with a vertical dashed line.

Proposition 11.3.1. *Let $\sum_{n=0}^{\infty} c_n(z-a)^n$ be a power series. There exists a $\rho \in [0,\infty]$ such that*

(*i*) *If $\rho = 0$, then the series diverges.*

(*ii*) *If $\rho = \infty$, then the series converges absolutely for all $z \in \mathbb{C}$.*

(*iii*) *If $0 < \rho < \infty$, then the series converges absolutely on $B(a,\rho)$, and diverges when $|z-a| > \rho$.*

Furthermore, if $0 < r < \rho$, then the series converges uniformly on the closed ball $C(a,r)$.

The number ρ is the *radius of convergence*. See Figure 11.6. The radius of convergence gives a disc around a where the series converges. A power series is convergent if $\rho > 0$.

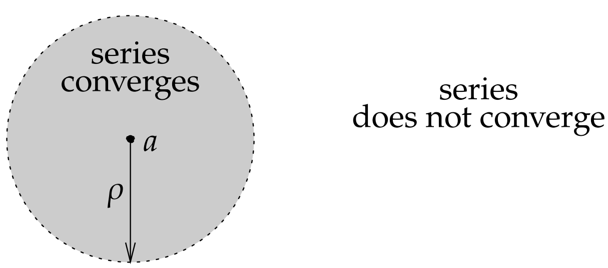

Figure 11.6: Radius of convergence.

Proof. We use the real version of this proposition, Proposition 2.6.10 in volume I. Let

$$R := \limsup_{n\to\infty} \sqrt[n]{|c_n|}.$$

If $R = 0$, then $\sum_{n=0}^{\infty} |c_n| \, |z-a|^n$ converges for all z. If $R = \infty$, then $\sum_{n=0}^{\infty} |c_n| \, |z-a|^n$ converges only at $z = a$. Otherwise, let $\rho := 1/R$ and $\sum_{n=0}^{\infty} |c_n| \, |z-a|^n$ converges when $|z-a| < \rho$, and diverges (in fact the terms of the series do not go to zero) when $|z-a| > \rho$.

To prove the "Furthermore," suppose $0 < r < \rho$ and $z \in C(a,r)$. Then consider the partial sums

$$\left| \sum_{n=0}^{k} c_n (z-a)^n \right| \leq \sum_{n=0}^{k} |c_n| |z-a|^n \leq \sum_{n=0}^{k} |c_n| r^n. \qquad \square$$

If $\sum_{n=0}^{\infty} c_n (z-a)^n$ converges for some z, then

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

converges absolutely whenever $|w-a| < |z-a|$. Conversely, if the series diverges at z, then it must diverge at w whenever $|w-a| > |z-a|$. Hence, to show that the radius of convergence is at least some number, we simply need to show convergence at some point by any method we know.

Example 11.3.2: We list some series we already know:

$$\sum_{n=0}^{\infty} z^n \qquad \text{has radius of convergence 1.}$$

$$\sum_{n=0}^{\infty} \frac{1}{n!} z^n \qquad \text{has radius of convergence } \infty.$$

$$\sum_{n=0}^{\infty} n^n z^n \qquad \text{has radius of convergence 0.}$$

Example 11.3.3: Note the difference between $\frac{1}{1-z}$ and its power series. Let us expand $\frac{1}{1-z}$ as power series around a point $a \neq 1$. Let $c := \frac{1}{1-a}$, then

$$\frac{1}{1-z} = \frac{c}{1-c(z-a)} = c \sum_{n=0}^{\infty} c^n (z-a)^n = \sum_{n=0}^{\infty} \left(\frac{1}{(1-a)^{n+1}} \right) (z-a)^n.$$

The series $\sum_{n=0}^{\infty} c^n (z-a)^n$ converges if and only if the series on the right-hand side converges and

$$\limsup_{n\to\infty} \sqrt[n]{|c^n|} = |c| = \frac{1}{|1-a|}.$$

The radius of convergence of the power series is $|1-a|$, that is, the distance from 1 to a. The function $\frac{1}{1-z}$ has a power series representation around every $a \neq 1$ and so is analytic in $\mathbb{C} \setminus \{1\}$. The domain of the function is bigger than the region of convergence of the power series representing the function at any point.

It turns out that if a function has a power series representation converging to the function on some ball, then it has a power series representation at every point in the ball. We will prove this result later.

11.3.3 Properties of analytic functions

Proposition 11.3.4. *If*

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

is convergent in $B(a,\rho)$ for some $\rho > 0$, then $f\colon B(a,\rho) \to \mathbb{C}$ is continuous. In particular, analytic functions are continuous.

Proof. For $z_0 \in B(a,\rho)$, pick $r < \rho$ such that $z_0 \in B(a,r)$. On $B(a,r)$ the partial sums (which are continuous) converge uniformly, and so the limit $f|_{B(a,r)}$ is continuous. Any sequence converging to z_0 has some tail that is completely in the open ball $B(a,r)$, hence f is continuous at z_0. □

In Corollary 6.2.13 of volume I, we proved that we can differentiate real power series term by term. That is, we proved that if

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

converges for real x in an interval around $a \in \mathbb{R}$, then we can differentiate term by term and obtain a series

$$f'(x) = \sum_{n=1}^{\infty} n c_n (x-a)^{n-1} = \sum_{n=0}^{\infty} (n+1) c_{n+1} (x-a)^n$$

with the same radius of convergence. We only proved this theorem when c_n is real, however, for complex c_n, we write $c_n = s_n + i t_n$, and as x and a are real

$$\sum_{n=0}^{\infty} c_n (x-a)^n = \sum_{n=0}^{\infty} s_n (x-a)^n + i \sum_{n=0}^{\infty} t_n (x-a)^n.$$

We apply the theorem to the real and imaginary part.

By iterating this theorem, we find that an analytic function is infinitely differentiable:

$$f^{(\ell)}(x) = \sum_{n=\ell}^{\infty} n(n-1)\cdots(n-\ell+1) c_k (x-a)^{n-\ell} = \sum_{n=0}^{\infty} (n+\ell)(n+\ell-1)\cdots(n+1) c_{n+\ell} (x-a)^n.$$

In particular,

$$f^{(\ell)}(a) = \ell!\, c_\ell. \tag{11.1}$$

The coefficients are uniquely determined by the derivatives of the function, and vice versa.

On the other hand, just because we have an infinitely differentiable function doesn't mean that the numbers c_n obtained by $c_n = \frac{f^{(n)}(0)}{n!}$ give a convergent power series. There is a theorem, which we will not prove, that given an arbitrary sequence $\{c_n\}_{n=1}^{\infty}$, there exists an infinitely differentiable function f such that $c_n = \frac{f^{(n)}(0)}{n!}$. Moreover, even if the obtained series converges, it may not converge to the function we started with. For an example, see Exercise 5.4.11 in volume I: The function

$$f(x) := \begin{cases} e^{-1/x} & \text{if } x > 0, \\ 0 & \text{if } x \leq 0, \end{cases}$$

is infinitely differentiable, and all derivatives at the origin are zero. So its series at the origin would be just the zero series, and while that series converges, it does not converge to f for $x > 0$.

We can apply an affine transformation $z \mapsto z + a$ that converts a power series at a to a series at the origin. That is, if

$$f(z) = \sum_{n=0}^{\infty} c_n (z-a)^n, \qquad \text{we consider} \qquad f(z+a) = \sum_{n=0}^{\infty} c_n z^n.$$

Therefore, it is usually sufficient to prove results about power series at the origin. From now on, we often assume $a = 0$ for simplicity.

11.3.4 Power series as analytic functions

We need a theorem on swapping limits of series, that is, Fubini's theorem for sums. For real series this was Exercise 2.6.15 in volume I, but we have a slicker argument now.

Theorem 11.3.5 (Fubini for sums). *Let $\{a_{k,m}\}_{k=1,m=1}^{\infty}$ be a double sequence of complex numbers and suppose that for every k the series*

$$\sum_{m=1}^{\infty} \lvert a_{k,m} \rvert \qquad \textit{converges}$$

and furthermore that

$$\sum_{k=1}^{\infty} \left(\sum_{m=1}^{\infty} \lvert a_{k,m} \rvert \right) \qquad \textit{converges}.$$

Then

$$\sum_{k=1}^{\infty} \left(\sum_{m=1}^{\infty} a_{k,m} \right) = \sum_{m=1}^{\infty} \left(\sum_{k=1}^{\infty} a_{k,m} \right),$$

where all the series involved converge.

Proof. Let E be the set $\{1/n : n \in \mathbb{N}\} \cup \{0\}$, and treat it as a metric space with the metric inherited from $\mathbb{R}$. Define the sequence of functions $f_k \colon E \to \mathbb{C}$ by

$$f_k(1/n) := \sum_{m=1}^{n} a_{k,m} \qquad \text{and} \qquad f_k(0) := \sum_{m=1}^{\infty} a_{k,m}.$$

As the series converges, each f_k is continuous at 0 (since 0 is the only cluster point, they are continuous at every point of E, but we don't need that). For all $x \in E$, we have

$$\lvert f_k(x) \rvert \leq \sum_{m=1}^{\infty} \lvert a_{k,m} \rvert .$$

As $\sum_k \sum_m \lvert a_{k,m} \rvert$ converges (and does not depend on x), we know that

$$\sum_{k=1}^{n} f_k(x)$$

converges uniformly on E. Define

$$g(x) := \sum_{k=1}^{\infty} f_k(x),$$

which is, therefore, a continuous function at 0. So

$$\begin{aligned}
\sum_{k=1}^{\infty} \left(\sum_{m=1}^{\infty} a_{k,m} \right) &= \sum_{k=1}^{\infty} f_k(0) = g(0) = \lim_{n\to\infty} g(1/n) \\
&= \lim_{n\to\infty} \sum_{k=1}^{\infty} f_k(1/n) = \lim_{n\to\infty} \sum_{k=1}^{\infty} \sum_{m=1}^{n} a_{k,m} \\
&= \lim_{n\to\infty} \sum_{m=1}^{n} \sum_{k=1}^{\infty} a_{k,m} = \sum_{m=1}^{\infty} \left(\sum_{k=1}^{\infty} a_{k,m} \right).
\end{aligned}$$

□

Now we prove that once we have a series converging to a function in some interval, we can expand the function around every point.

Theorem 11.3.6 (Taylor's theorem for real-analytic functions)**.** *Let*

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

be a power series converging in $(-\rho, \rho)$ *for some* $\rho > 0$*. Given any* $a \in (-\rho, \rho)$*, and* x *such that* $\lvert x - a \rvert < \rho - \lvert a \rvert$*, we have*

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

The power series at a could of course converge in a larger interval, but the one above is guaranteed. It is the largest symmetric interval about a that fits in $(-\rho, \rho)$.

Proof. Given a and x as in the theorem, write

$$f(x) = \sum_{k=0}^{\infty} a_k \big((x-a)+a\big)^k$$
$$= \sum_{k=0}^{\infty} a_k \sum_{m=0}^{k} \binom{k}{m} a^{k-m} (x-a)^m.$$

Define $c_{k,m} := a_k \binom{k}{m} a^{k-m}$ if $m \leq k$ and 0 if $m > k$. Then

$$f(x) = \sum_{k=0}^{\infty} \sum_{m=0}^{\infty} c_{k,m} (x-a)^m. \tag{11.2}$$

Let us show that the double sum converges absolutely.

$$\sum_{k=0}^{\infty} \sum_{m=0}^{\infty} \left| c_{k,m} (x-a)^m \right| = \sum_{k=0}^{\infty} \sum_{m=0}^{k} \left| a_k \binom{k}{m} a^{k-m} (x-a)^m \right|$$
$$= \sum_{k=0}^{\infty} \lvert a_k \rvert \sum_{m=0}^{k} \binom{k}{m} \lvert a \rvert^{k-m} \lvert x-a \rvert^m$$
$$= \sum_{k=0}^{\infty} \lvert a_k \rvert \big(\lvert x-a \rvert + \lvert a \rvert\big)^k,$$

and this series converges as long as $(\lvert x-a\rvert + \lvert a \rvert) < \rho$ or in other words if $\lvert x-a \rvert < \rho - \lvert a \rvert$.

Using Theorem 11.3.5, swap the order of summation in (11.2), and the following series converges when $\lvert x-a \rvert < \rho - \lvert a \rvert$:

$$f(x) = \sum_{k=0}^{\infty} \sum_{m=0}^{\infty} c_{k,m} (x-a)^m = \sum_{m=0}^{\infty} \left(\sum_{k=0}^{\infty} c_{k,m} \right) (x-a)^m.$$

The formula in terms of derivatives at a follows by differentiating the series to obtain (11.1). ☐

Note that if a series converges for real $x \in (a-\rho, a+\rho)$ it also converges for all complex numbers in $B(a,\rho)$. We have the following corollary, which says that functions defined by power series are analytic.

Corollary 11.3.7. *For every $a \in \mathbb{C}$, if $\sum_{k=0}^{\infty} c_k (z-a)^k$ converges to $f(z)$ in $B(a,\rho)$ and $b \in B(a,\rho)$, then there exists a power series $\sum_{k=0}^{\infty} d_k (z-b)^k$ that converges to $f(z)$ in $B(b, \rho - \lvert b-a \rvert)$.*

Proof. Without loss of generality assume that $a = 0$. We can rotate to assume that b is real, but since that is harder to picture, let us do it explicitly. Let $\alpha := \frac{\bar{b}}{\lvert b \rvert}$. Notice that

$$\lvert 1/\alpha \rvert = \lvert \alpha \rvert = 1.$$

Therefore the series $\sum_{k=0}^{\infty} c_k (z/\alpha)^k = \sum_{k=0}^{\infty} c_k \alpha^{-k} z^k$ converges to $f(z/\alpha)$ in $B(0,\rho)$. When $z = x$ is real we apply Theorem 11.3.6 at $|b|$ and get a series that converges to $f(z/\alpha)$ on $B(|b|, \rho - |b|)$. That is, there is a convergent series

$$f(z/\alpha) = \sum_{k=0}^{\infty} a_k \big(z - |b|\big)^k.$$

Using $\alpha b = |b|$, we find

$$f(z) = f(\alpha z/\alpha) = \sum_{k=0}^{\infty} a_k (\alpha z - |b|)^k = \sum_{k=0}^{\infty} a_k \alpha^k \big(z - |b|/\alpha\big)^k = \sum_{k=0}^{\infty} a_k \alpha^k (z-b)^k,$$

and this series converges for all z such that $\big|\alpha z - |b|\big| < \rho - |b|$ or $|z - b| < \rho - |b|$. ◻

We proved above that a convergent power series is an analytic function where it converges. We have also shown before that $\frac{1}{1-z}$ is analytic outside of $z = 1$.

Note that just because a real analytic function is analytic on the entire real line it does not necessarily mean that it has a power series representation that converges everywhere. For example, the function

$$f(x) = \frac{1}{1+x^2}$$

happens to be real analytic function on $\mathbb{R}$ (exercise). A power series around the origin converging to f has a radius of convergence of exactly 1. Can you see why? (exercise)

11.3.5 Identity theorem for analytic functions

Lemma 11.3.8. *Suppose $f(z) = \sum_{k=0}^{\infty} a_k z^k$ is a convergent power series and $\{z_n\}_{n=1}^{\infty}$ is a sequence of nonzero complex numbers converging to 0, such that $f(z_n) = 0$ for all n. Then $a_k = 0$ for every k.*

Proof. By continuity we know $f(0) = 0$ so $a_0 = 0$. Suppose there exists some nonzero a_k. Let m be the smallest m such that $a_m \neq 0$. Then

$$f(z) = \sum_{k=m}^{\infty} a_k z^k = z^m \sum_{k=m}^{\infty} a_k z^{k-m} = z^m \sum_{k=0}^{\infty} a_{k+m} z^k.$$

Write $g(z) = \sum_{k=0}^{\infty} a_{k+m} z^k$ (this series converges in on the same set as f). g is continuous and $g(0) = a_m \neq 0$. Thus there exists some $\delta > 0$ such that $g(z) \neq 0$ for all $z \in B(0,\delta)$. As $f(z) = z^m g(z)$, the only point in $B(0,\delta)$ where $f(z) = 0$ is when $z = 0$, but this contradicts the assumption that $f(z_n) = 0$ for all n. ◻

Recall that in a metric space X, a *cluster point* (or sometimes *limit point*) of a set E is a point $p \in X$ such that $B(p,\epsilon) \setminus \{p\}$ contains points of E for all $\epsilon > 0$.

Theorem 11.3.9 (Identity theorem). *Let $U \subset \mathbb{C}$ be open and connected. If $f: U \to \mathbb{C}$ and $g: U \to \mathbb{C}$ are analytic functions that are equal on a set $E \subset U$, and E has a cluster point in U, then $f(z) = g(z)$ for all $z \in U$.*

In most common applications of this theorem E is an open set or perhaps a curve.

Proof. Without loss of generality suppose E is the set of all points $z \in U$ such that $g(z) = f(z)$. Note that E must be closed as f and g are continuous.

Suppose E has a cluster point. Without loss of generality assume that 0 is this cluster point. Near 0, we have the expansions

$$f(z) = \sum_{k=0}^{\infty} a_k z^k \qquad \text{and} \qquad g(z) = \sum_{k=0}^{\infty} b_k z^k,$$

which converge in some ball $B(0, \rho)$. Therefore the series

$$0 = f(z) - g(z) = \sum_{k=0}^{\infty} (a_k - b_k) z^k$$

converges in $B(0, \rho)$. As 0 is a cluster point of E, there is a sequence of nonzero points $\{z_n\}_{n=1}^{\infty}$ such that $f(z_n) - g(z_n) = 0$. Hence, by the lemma above $a_k = b_k$ for all k. Therefore, $B(0, \rho) \subset E$.

Thus the set of cluster points of E is open. The set of cluster points of E is also closed: A limit of cluster points of E is in E as it is closed, and it is clearly a cluster point of E. As U is connected, the set of cluster points of E is equal to U, or in other words $E = U$. □

By restricting our attention to real x, we obtain the same theorem for connected open subsets of $\mathbb{R}$, which are just open intervals.

11.3.6 Exercises

***Exercise* 11.3.1:** *Let*

$$a_{k,m} := \begin{cases} 1 & \text{if } k = m, \\ -2^{k-m} & \text{if } k < m, \\ 0 & \text{if } k > m. \end{cases}$$

Compute (or show the limit doesn't exist):

a) $\displaystyle\sum_{m=1}^{\infty} |a_{k,m}|$ *for all k,* *b)* $\displaystyle\sum_{k=1}^{\infty} |a_{k,m}|$ *for all m,* *c)* $\displaystyle\sum_{k=1}^{\infty}\sum_{m=1}^{\infty} |a_{k,m}|$, *d)* $\displaystyle\sum_{k=1}^{\infty}\sum_{m=1}^{\infty} a_{k,m}$, *e)* $\displaystyle\sum_{m=1}^{\infty}\sum_{k=1}^{\infty} a_{k,m}$.

Hint: Fubini for sums does not apply, in fact, answers to d) and e) are different.

***Exercise* 11.3.2:** *Let $f(x) := \frac{1}{1+x^2}$. Prove that*

a) f is analytic function on all of $\mathbb{R}$ by finding a power series for f at every $a \in \mathbb{R}$,

b) the radius of convergence of the power series for f at the origin is 1.

Exercise* 11.3.3*: *Suppose $f\colon \mathbb{C} \to \mathbb{C}$ is analytic. Show that for each n, there are at most finitely many zeros of f in $B(0,n)$, that is, $f^{-1}(0) \cap B(0,n)$ is finite for each n.*

Exercise* 11.3.4*: *Suppose $U \subset \mathbb{C}$ is open and connected, $0 \in U$, and $f\colon U \to \mathbb{C}$ is analytic. Treating f as a function of a real x at the origin, suppose $f^{(n)}(0) = 0$ for all n. Show that $f(z) = 0$ for all $z \in U$.*

Exercise* 11.3.5*: *Suppose $U \subset \mathbb{C}$ is open and connected, $0 \in U$, and $f\colon U \to \mathbb{C}$ is analytic. For real x and y, let $h(x) := f(x)$ and $g(y) := -i\,f(iy)$. Show that h and g are infinitely differentiable at the origin and $h'(0) = g'(0)$.*

Exercise* 11.3.6*: *Suppose a function f is analytic in some neighborhood of the origin, and that there exists an M such that $|f^{(n)}(0)| \leq M$ for all n. Prove that the series of f at the origin converges for all $z \in \mathbb{C}$.*

Exercise* 11.3.7*: *Suppose $f(z) := \sum_{n=0}^{\infty} c_n z^n$ with a radius of convergence 1. Suppose $f(0) = 0$, but f is not the zero function. Show that there exists a $k \in \mathbb{N}$ and a convergent power series $g(z) := \sum_{n=0}^{\infty} d_n z^n$ with radius of convergence 1 such that $f(z) = z^k g(z)$ for all $z \in B(0,1)$, and $g(0) \neq 0$.*

Exercise* 11.3.8*: *Suppose $U \subset \mathbb{C}$ is open and connected. Suppose that $f\colon U \to \mathbb{C}$ is analytic, $U \cap \mathbb{R} \neq \emptyset$ and $f(x) = 0$ for all $x \in U \cap \mathbb{R}$. Show that $f(z) = 0$ for all $z \in U$.*

Exercise* 11.3.9*: *For $\alpha \in \mathbb{C}$ and $k = 0, 1, 2, 3 \ldots$, define*

$$\binom{\alpha}{k} := \frac{\alpha(\alpha-1)\cdots(\alpha-k)}{k!}.$$

a) Show that the series

$$f(z) := \sum_{k=0}^{\infty} \binom{\alpha}{k} z^k$$

converges whenever $|z| < 1$. In fact, prove that for $\alpha = 0, 1, 2, 3, \ldots$ the radius of convergence is ∞, and for all other α the radius of convergence is 1.

b) Show that for $x \in \mathbb{R}$, $|x| < 1$, we have

$$(1+x)f'(x) = \alpha f(x),$$

meaning that $f(x) = (1+x)^{\alpha}$.

Exercise* 11.3.10*: *Suppose $f\colon \mathbb{C} \to \mathbb{C}$ is analytic and suppose that for some open interval $(a,b) \subset \mathbb{R}$, f is real valued on (a,b). Show that f is real-valued on $\mathbb{R}$.*

Exercise* 11.3.11*: *Let $\mathbb{D} := B(0,1)$ be the unit disc. Suppose $f\colon \mathbb{D} \to \mathbb{C}$ is analytic with power series $\sum_{n=0}^{\infty} c_n z^n$. Suppose $|c_n| \leq 1$ for all n. Prove that for all $z \in \mathbb{D}$, we have $|f(z)| \leq \frac{1}{1-|z|}$.*

11.4 Complex exponential and trigonometric functions

Note: 1 lecture

11.4.1 The complex exponential

Let

$$E(z) := \sum_{k=0}^{\infty} \frac{1}{k!} z^k.$$

This series converges for all $z \in \mathbb{C}$, and so by Corollary 11.3.7, E is analytic on $\mathbb{C}$. We notice that $E(0) = 1$, and that for $z = x \in \mathbb{R}$, $E(x) \in \mathbb{R}$. Keeping x real, direct computation shows

$$\frac{d}{dx}\big(E(x)\big) = E(x).$$

In §5.4 of volume I (or by Picard's theorem), we proved that the unique function satisfying $E' = E$ and $E(0) = 1$ is the exponential. In other words, for $x \in \mathbb{R}$, $e^x = E(x)$.

For complex numbers z, we define

$$e^z := E(z) = \sum_{k=0}^{\infty} \frac{1}{k!} z^k.$$

On the real line this new definition agrees with our previous one. See Figure 11.7. Notice that in the x direction (the real direction) the graph behaves like the real exponential, and in the y direction (the imaginary direction) the graph oscillates.

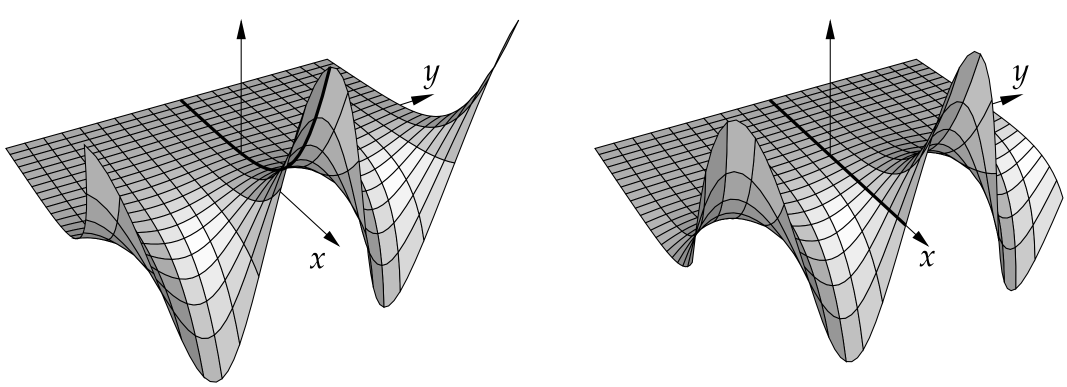

Figure 11.7: Graphs of the real part (left) and imaginary part (right) of the complex exponential $e^z = e^{x+iy}$. The x-axis goes from -4 to 4, the y-axis goes from -6 to 6, and the vertical axis goes from $-e^4 \approx -54.6$ to $e^4 \approx 54.6$. The plot of the real exponential ($y = 0$) is marked in a bold line.

Proposition 11.4.1 (Law of exponents). *Let $z, w \in \mathbb{C}$ be complex numbers. Then*

$$e^{z+w} = e^z e^w.$$

Proof. We already know that the equality $e^{x+y} = e^x e^y$ holds for all real numbers x and y. For every fixed $y \in \mathbb{R}$, consider the expressions as functions of x and apply the identity theorem (Theorem 11.3.9) to get that $e^{z+y} = e^z e^y$ for all $z \in \mathbb{C}$. Fixing an arbitrary $z \in \mathbb{C}$, we get $e^{z+y} = e^z e^y$ for all $y \in \mathbb{R}$. Again by the identity theorem $e^{z+w} = e^z e^w$ for all $w \in \mathbb{C}$. □

A simple consequence of the proposition is that $e^z \neq 0$ for all $z \in \mathbb{C}$, as $e^z e^{-z} = e^{z-z} = 1$. This computation means that $(e^z)^{-1} = e^{-z}$. Combining that fact with the law of exponents gives

$$(e^z)^n = e^{nz} \qquad \text{for all } n \in \mathbb{Z}.$$

A yet more complicated consequence is that we can compute the power series for the exponential at any point $a \in \mathbb{C}$:

$$e^z = e^a e^{z-a} = \sum_{k=0}^{\infty} \frac{e^a}{k!}(z-a)^k.$$

11.4.2 Trigonometric functions and π

We can now finally define *sine* and *cosine* by the equation

$$e^{x+iy} = e^x \big(\cos(y) + i \sin(y)\big).$$

In fact, we define sine and cosine for all complex z:

$$\cos(z) := \frac{e^{iz} + e^{-iz}}{2} \qquad \text{and} \qquad \sin(z) := \frac{e^{iz} - e^{-iz}}{2i}.$$

Let us use our definition to prove common properties of sine and cosine. In the process, we also define the number π.

Proposition 11.4.2. *The sine and cosine functions have the following properties:*

(*i*) *For all $z \in \mathbb{C}$,*

$$e^{iz} = \cos(z) + i \sin(z) \qquad (\textit{Euler's formula}).$$

(*ii*) $\cos(0) = 1$, $\sin(0) = 0$.

(*iii*) *For all $z \in \mathbb{C}$,*

$$\cos(-z) = \cos(z), \qquad \sin(-z) = -\sin(z).$$

(*iv*) *For all $z \in \mathbb{C}$,*

$$\cos(z) = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} z^{2k}, \qquad \sin(z) = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} z^{2k+1}.$$

(*v*) *For all* $x \in \mathbb{R}$
$$\cos(x) = \operatorname{Re}(e^{ix}) \qquad \textit{and} \qquad \sin(x) = \operatorname{Im}(e^{ix}).$$

(*vi*) *For all* $z \in \mathbb{C}$,
$$\big(\cos(z)\big)^2 + \big(\sin(z)\big)^2 = 1.$$

(*vii*) *For all* $x \in \mathbb{R}$,
$$|\sin(x)| \leq 1, \qquad |\cos(x)| \leq 1.$$

(*viii*) *For all* $x \in \mathbb{R}$,
$$\frac{d}{dx}\big[\cos(x)\big] = -\sin(x) \qquad \textit{and} \qquad \frac{d}{dx}\big[\sin(x)\big] = \cos(x).$$

(*ix*) *For all* $x \geq 0$,
$$\sin(x) \leq x.$$

(*x*) *There exists an* $x > 0$ *such that* $\cos(x) = 0$. *We define*
$$\pi := 2 \inf\{x > 0 : \cos(x) = 0\}.$$

(*xi*) *For all* $z \in \mathbb{C}$,
$$e^{2\pi i} = 1 \qquad \textit{and} \qquad e^{z+i2\pi} = e^z.$$

(*xii*) *Sine and cosine are* 2π*-periodic and not periodic with any smaller period. That is,* 2π *is the smallest number such that for all* $z \in \mathbb{C}$,
$$\sin(z + 2\pi) = \sin(z) \qquad \textit{and} \qquad \cos(z + 2\pi) = \cos(z).$$

(*xiii*) *The function* $x \mapsto e^{ix}$ *is a bijective map from* $[0, 2\pi)$ *onto the set of* $z \in \mathbb{C}$ *such that* $|z| = 1$.

The proposition immediately implies that $\sin(x)$ and $\cos(x)$ are real whenever x is real.

Proof. The first three items follow directly from the definition. The computation of the power series for both is left as an exercise. As complex conjugate is a continuous function, the definition of e^z implies $\overline{e^z} = e^{\bar{z}}$. If x is real,
$$\overline{e^{ix}} = e^{-ix}.$$
Thus for real x, $\cos(x) = \frac{e^{ix}-e^{-ix}}{2} = \frac{e^{ix}-\overline{e^{ix}}}{2} = \operatorname{Re}(e^{ix})$ and similarly $\sin(x) = \operatorname{Im}(e^{ix})$.

For real x, we compute
$$1 = e^{ix}e^{-ix} = e^{ix}\,\overline{e^{ix}} = |e^{ix}|^2 = \big|\cos(x) + i\sin(x)\big|^2 = \big(\cos(x)\big)^2 + \big(\sin(x)\big)^2.$$

A slightly more complicated computation shows this fact for complex numbers, see Exercise 11.4.6. In particular, is e^{ix} is unimodular for real x; the values lie on the unit circle. A square of a real number is always nonnegative:
$$\big(\sin(x)\big)^2 = 1 - \big(\cos(x)\big)^2 \leq 1.$$

So $|\sin(x)| \leq 1$ and similarly $|\cos(x)| \leq 1$.

We leave the computation of the derivatives to the reader as exercises. Let us prove that $\sin(x) \leq x$ for $x \geq 0$. Consider $f(x) := x - \sin(x)$ and differentiate:

$$f'(x) = \frac{d}{dx}\big[x - \sin(x)\big] = 1 - \cos(x) \geq 0,$$

for all $x \in \mathbb{R}$ as $|\cos(x)| \leq 1$. In other words, f is increasing and $f(0) = 0$. So f must be nonnegative when $x \geq 0$ and hence, $\sin(x) \geq x$.

Next, we claim there exists a positive x such that $\cos(x) = 0$. As $\cos(0) = 1 > 0$, $\cos(x) > 0$ for x near 0. Namely, there is some $y > 0$, such that $\cos(x) > 0$ on $[0, y)$. Then $\sin(x)$ is strictly increasing on $[0, y)$. As $\sin(0) = 0$, then $\sin(x) > 0$ for $x \in (0, y)$. Take $a \in (0, y)$. By the mean value theorem, there is a $c \in (a, y)$ such that

$$2 \geq \cos(a) - \cos(y) = \sin(c)(y - a) \geq \sin(a)(y - a).$$

As $a \in (0, y)$, then $\sin(a) > 0$ and so

$$y \leq \frac{2}{\sin(a)} + a.$$

Hence there is some largest y such that $\cos(x) > 0$ in $[0, y)$, and let y be the largest such number. By continuity, $\cos(y) = 0$. In fact, y is the smallest positive y such that $\cos(y) = 0$. As mentioned, π is defined to be $2y$.

As $\cos(\pi/2) = 0$, then $\big(\sin(\pi/2)\big)^2 = 1$. As sin is positive on $(0, \pi/2)$, we have $\sin(\pi/2) = 1$. Hence,

$$e^{i\pi/2} = i,$$

and by the law of exponents,

$$e^{i\pi} = -1, \qquad e^{i2\pi} = 1.$$

So $e^{i2\pi} = 1 = e^0$. The law of exponents also says

$$e^{z+i2\pi} = e^z e^{i2\pi} = e^z$$

for all $z \in \mathbb{C}$. Immediately, we also obtain $\cos(z + 2\pi) = \cos(z)$ and $\sin(z + 2\pi) = \sin(z)$. So sin and cos are 2π-periodic.

We claim that sin and cos are not periodic with a smaller period. It would suffice to show that if $e^{ix} = 1$ for the smallest positive x, then $x = 2\pi$. Let x be the smallest positive x such that $e^{ix} = 1$. Of course, $x \leq 2\pi$. By the law of exponents,

$$\big(e^{ix/4}\big)^4 = 1.$$

If $e^{ix/4} = a + ib$, then

$$(a + ib)^4 = a^4 - 6a^2b^2 + b^4 + i\big(4ab(a^2 - b^2)\big) = 1.$$

Then either $a = 0$ or $a^2 = b^2$. As $x/4 \leq \pi/2$, then $a = \cos(x/4) \geq 0$ and $b = \sin(x/4) > 0$. If $a^2 = b^2$, then $a^4 - 6a^2b^2 + b^4 = -4a^4 < 0$ and in particular not equal to 1. Therefore $a = 0$, in which case $x/4 = \pi/2$. Hence 2π is the smallest period we could choose for e^{ix} and so also for cos and sin.

Finally, we wish to show that e^{ix} is one-to-one and onto from the set $[0, 2\pi)$ to the set of $z \in \mathbb{C}$ such that $|z| = 1$. Suppose $e^{ix} = e^{iy}$ and $x > y$. Then $e^{i(x-y)} = 1$, meaning $x - y$ is a multiple of 2π and hence only one of them can live in $[0, 2\pi)$. To show onto, pick $(a, b) \in \mathbb{R}^2$ such that $a^2 + b^2 = 1$. Suppose first that $a, b \geq 0$. By the intermediate value theorem, there must exist an $x \in [0, \pi/2]$ such that $\cos(x) = a$, and hence $b^2 = \bigl(\sin(x)\bigr)^2$. As b and $\sin(x)$ are nonnegative, $b = \sin(x)$. Since $-\sin(x)$ is the derivative of $\cos(x)$ and $\cos(-x) = \cos(x)$, then $\sin(x) < 0$ for $x \in [-\pi/2, 0)$. Using the same reasoning, we obtain that if $a > 0$ and $b \leq 0$, we can find an x in $[-\pi/2, 0)$, and by periodicity, $x \in [3\pi/2, 2\pi)$ such that $\cos(x) = a$ and $\sin(x) = b$. Multiplying by -1 is the same as multiplying by $e^{i\pi}$ or $e^{-i\pi}$. So we can always assume that $a \geq 0$ (details are left as exercise). □

11.4.3 The unit circle and polar coordinates

The arclength of a curve parametrized by $\gamma \colon [a, b] \to \mathbb{C}$ is given by

$$\int_a^b |\gamma'(t)| \, dt.$$

We have that e^{it} parametrizes the circle for t in $[0, 2\pi)$. As $\frac{d}{dt}\bigl(e^{it}\bigr) = ie^{it}$, the circumference of the circle (the arclength) is

$$\int_0^{2\pi} |ie^{it}| \, dt = \int_0^{2\pi} 1 \, dt = 2\pi.$$

More generally, e^{it} parametrizes the circle by arclength. That is, t measures the arclength on a circle of radius 1 by the angle in radians. So the definitions of sin and cos given above agree with the standard geometric definitions.

All the points on the unit circle can be achieved by e^{it} for some t. Therefore, we can write a complex number $z \in \mathbb{C}$ (in so-called *polar coordinates*) as

$$z = re^{i\theta}$$

for some $r \geq 0$ and $\theta \in \mathbb{R}$. The θ is, of course, not unique as θ or $\theta + 2\pi$ gives the same number. The law of exponents $e^{a+b} = e^a e^b$ leads to a useful formula for powers and products of complex numbers in polar coordinates:

$$\bigl(re^{i\theta}\bigr)^n = r^n e^{in\theta}, \qquad (re^{i\theta})(se^{i\gamma}) = rse^{i(\theta+\gamma)}.$$

11.4.4 Exercises

***Exercise* 11.4.1:** *Derive the power series for* $\sin(z)$ *and* $\cos(z)$ *at the origin.*

***Exercise* 11.4.2:** *Using the power series, show that for real* x, *we have* $\frac{d}{dx}\bigl[\sin(x)\bigr] = \cos(x)$ *and* $\frac{d}{dx}\bigl[\cos(x)\bigr] = -\sin(x)$.

***Exercise* 11.4.3:** *Finish the proof of the argument that* $x \mapsto e^{ix}$ *from* $[0, 2\pi)$ *is onto the unit circle. In particular, assume that we get all points of the form* (a, b) *where* $a^2 + b^2 = 1$ *and* $a \geq 0$. *By multiplying by* $e^{i\pi}$ *or* $e^{-i\pi}$, *show that we get everything, that is, even points where* $a < 0$.

***Exercise* 11.4.4:** *Show that the exponential is onto* $\mathbb{C} \setminus \{0\}$, *and in fact, that for every nonzero* w, *there are infinitely many* $z \in \mathbb{C}$ *such that* $e^z = w$.

***Exercise* 11.4.5:** *Prove that for every* $w \neq 0$ *and every* $\epsilon > 0$, *there exists a* $z \in \mathbb{C}$ *with* $|z| < \epsilon$ *such that* $e^{1/z} = w$.

***Exercise* 11.4.6:** *We showed* $\bigl(\cos(x)\bigr)^2 + \bigl(\sin(x)\bigr)^2 = 1$ *for all* $x \in \mathbb{R}$. *Prove that* $\bigl(\cos(z)\bigr)^2 + \bigl(\sin(z)\bigr)^2 = 1$ *for all* $z \in \mathbb{C}$.

***Exercise* 11.4.7:** *Prove the trigonometric identities* $\sin(z + w) = \sin(z)\cos(w) + \cos(z)\sin(w)$ *and* $\cos(z + w) = \cos(z)\cos(w) - \sin(z)\sin(w)$ *for all* $z, w \in \mathbb{C}$.

***Exercise* 11.4.8:** *Define* $\operatorname{sinc}(z) := \frac{\sin(z)}{z}$ *for* $z \neq 0$ *and* $\operatorname{sinc}(0) := 1$. *Show that sinc is analytic and compute its power series at zero.*

Define the *hyperbolic sine* and *hyperbolic cosine* by

$$\sinh(z) := \frac{e^z - e^{-z}}{2}, \qquad \cosh(z) := \frac{e^z + e^{-z}}{2}.$$

***Exercise* 11.4.9:** *Derive the power series at the origin for the hyperbolic sine and cosine.*

***Exercise* 11.4.10:** *Show*

a) $\sinh(0) = 0$, $\cosh(0) = 1$.

b) For $x \in \mathbb{R}$, $\frac{d}{dx}\bigl[\sinh(x)\bigr] = \cosh(x)$ *and* $\frac{d}{dx}\bigl[\cosh(x)\bigr] = \sinh(x)$.

c) $\cosh(x) > 0$ *for all* $x \in \mathbb{R}$ *and show that* $\sinh(x)$ *is strictly increasing and bijective from* $\mathbb{R}$ *to* $\mathbb{R}$.

d) $\bigl(\cosh(x)\bigr)^2 = 1 + \bigl(\sinh(x)\bigr)^2$ *for all* x.

***Exercise* 11.4.11:** *Define* $\tan(x) := \frac{\sin(x)}{\cos(x)}$ *as usual.*

a) Show that for $x \in (-\pi/2, \pi/2)$ *both* $\sin$ *and* $\tan$ *are strictly increasing, and hence* $\sin^{-1}$ *and* $\tan^{-1}$ *exist when we restrict to that interval.*

b) Show that $\sin^{-1}$ *and* $\tan^{-1}$ *are differentiable and that* $\frac{d}{dx}\sin^{-1}(x) = \frac{1}{\sqrt{1-x^2}}$ *and* $\frac{d}{dx}\tan^{-1}(x) = \frac{1}{1+x^2}$.

c) Using the finite geometric sum formula show

$$\tan^{-1}(x) = \int_0^x \frac{1}{1+t^s}\,dt = \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} x^{2k+1}$$

converges for all $-1 \leq x \leq 1$ *(including the end points). Hint: Integrate the finite sum, not the series.*

d) Use this to show that

$$1 - \frac{1}{3} + \frac{1}{5} - \cdots = \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} = \frac{\pi}{4}.$$

11.5 Maximum principle and the fundamental theorem of algebra

Note: half a lecture, optional

In this section we study the local behavior of polynomials, and analytic functions in general, and the growth of polynomials as z goes to infinity. As an application we prove the fundamental theorem of algebra: Any nonconstant polynomial has a complex root.

Lemma 11.5.1. *Let $\epsilon > 0$, let $p(z)$ be a nonconstant complex polynomial, or more generally a nonconstant power series converging in $B(z_0, \epsilon)$, and suppose $p(z_0) \neq 0$. Then there exists a $w \in B(z_0, \epsilon)$ such that $|p(w)| < |p(z_0)|$.*

Proof. We prove this lemma for a polynomial and leave the general case as Exercise 11.5.1. Without loss of generality assume that $z_0 = 0$ and $p(0) = 1$. Write

$$p(z) = 1 + a_k z^k + a_{k+1} z^{k+1} + \cdots + a_d z^d,$$

where $a_k \neq 0$. Pick t such that $a_k e^{ikt} = -|a_k|$, which we can do by the discussion on trigonometric functions. Suppose $r > 0$ is small enough such that $1 - r^k|a_k| > 0$. We have

$$p(re^{it}) = 1 - r^k|a_k| + r^{k+1} a_{k+1} e^{i(k+1)t} + \cdots + r^d a_d e^{idt}.$$

So

$$\begin{aligned} \left|p(re^{it})\right| - \left|r^{k+1} a_{k+1} e^{i(k+1)t} + \cdots + r^d a_d e^{idt}\right| &\leq \left|p(re^{it}) - r^{k+1} a_{k+1} e^{i(k+1)t} - \cdots - r^d a_d e^{idt}\right| \\ &= \left|1 - r^k|a_k|\right| = 1 - r^k|a_k|. \end{aligned}$$

In other words,

$$\left|p(re^{it})\right| \leq 1 - r^k \left(|a_k| - r \left| a_{k+1} e^{i(k+1)t} + \cdots + r^{d-k-1} a_d e^{idt} \right| \right).$$

For small enough r, the expression in the parentheses is positive as $|a_k| > 0$. Hence, $\left|p(re^{it})\right| < 1 = p(0)$. □

What the lemma says is that the only minima the modulus of analytic functions has are precisely at the zeros. It is sometimes called the *minimum modulus principle*. If f is analytic and nonzero at a point, then $1/f$ is analytic near that point. Applying the lemma and the identity theorem, one obtains the *maximum modulus principle*, or sometimes just the *maximum principle*.

Theorem 11.5.2 (Maximum modulus principle)**.** *If $U \subset \mathbb{C}$ is open and connected, $f : U \to \mathbb{C}$ is analytic, and $|f(z)|$ attains a relative maximum at $z_0 \in U$, then f is constant.*

The details of the proof is left as Exercise 11.5.2.

Remark 11.5.3. The lemma (and the maximum principle) does not hold if we restrict to the real numbers. For example, $x^2 + 1$ has a minimum at $x = 0$, but no zero there. There is a w arbitrarily close to 0 such that $|w^2 + 1| < 1$, but this w is necessarily not real. Letting $w = i\epsilon$ for small $\epsilon > 0$ works.

The moral of the story is that if $p(0) = 1$, then very close to 0, the series (or polynomial) looks like $1 + az^k$, and $1 + az^k$ has no minimum at the origin. All the higher powers of z are too small to make a difference. For polynomials, we find similar behavior at infinity.

Lemma 11.5.4. *Let $p(z)$ be a nonconstant complex polynomial. Then for an $M > 0$, there exists an $R > 0$ such that $|p(z)| \geq M$ whenever $|z| \geq R$.*

Proof. Write $p(z) = a_0 + a_1 z + \cdots + a_d z^d$ and suppose that $d \geq 1$ and $a_d \neq 0$. Suppose $|z| \geq R$ (so also $|z|^{-1} \leq R^{-1}$). We estimate:

$$\begin{aligned} |p(z)| &\geq |a_d z^d| - |a_0| - |a_1 z| - \cdots - |a_{d-1} z^{d-1}| \\ &= |z|^d \bigl(|a_d| - |a_0|\,|z|^{-d} - |a_1|\,|z|^{-d+1} - \cdots - |a_{d-1}|\,|z|^{-1}\bigr) \\ &\geq R^d \bigl(|a_d| - |a_0| R^{-d} - |a_1| R^{1-d} - \cdots - |a_{d-1}| R^{-1}\bigr). \end{aligned}$$

Then the expression in parentheses is eventually positive for large enough R. In particular, for large enough R we get that this expression is greater than $\frac{|a_d|}{2}$, and so

$$|p(z)| \geq R^d \frac{|a_d|}{2}.$$

Therefore, we can pick R large enough to be bigger than a given M. □

This second lemma does *not* generalize to analytic functions, even those defined on the entire plane $\mathbb{C}$. The function $\cos(z)$ is a counterexample. We had to look at the term with the largest degree, and we only have such a term for a polynomial. In fact, something that we will not prove is that an analytic function defined on all of $\mathbb{C}$ satisfying the conclusion of the lemma must be a polynomial.

The moral of the story here is that for very large $|z|$ (far away from the origin) a polynomial of degree d really looks like a constant multiple of z^d.

Theorem 11.5.5 (Fundamental theorem of algebra)**.** *Let $p(z)$ be a nonconstant complex polynomial, then there exists a $z_0 \in \mathbb{C}$ such that $p(z_0) = 0$.*

Proof. Let $\mu := \inf\bigl\{|p(z)| : z \in \mathbb{C}\bigr\}$. Find an R such that for all z with $|z| \geq R$, we have $|p(z)| \geq \mu + 1$. Therefore, every z with $|p(z)|$ close to μ must be in the closed ball $C(0, R) = \bigl\{z \in \mathbb{C} : |z| \leq R\bigr\}$. As $|p(z)|$ is a continuous real-valued function, it achieves its minimum on the compact set $C(0, R)$ (closed and bounded) and this minimum must be μ. So there is a $z_0 \in C(0, R)$ such that $|p(z_0)| = \mu$. As that is a minimum of $|p(z)|$ on $\mathbb{C}$, then by the first lemma above, we have $|p(z_0)| = 0$. □

The fundamental theorem also does not generalize to analytic functions. The exponential e^z is an analytic function on $\mathbb{C}$ with no zeros.

11.5.1 Exercises

***Exercise* 11.5.1:** *Prove Lemma 11.5.1 for an analytic function. That is, suppose that $p(z)$ is a nonconstant power series converging in $B(z_0, \epsilon)$.*

***Exercise* 11.5.2:** *Use Lemma 11.5.1 for analytic functions to prove Theorem 11.5.2.*

***Exercise* 11.5.3:** *Let $U \subset \mathbb{C}$ be open and $z_0 \in U$. Suppose $f : U \to \mathbb{C}$ is analytic and $f(z_0) = 0$. Show that there exists an $\epsilon > 0$ such that either $f(z) \neq 0$ for all z with $0 < |z| < \epsilon$ or $f(z) = 0$ for all $z \in B(z_0, \epsilon)$. In other words, zeros of analytic functions are isolated. Of course, same holds for polynomials.*

A *rational function* is a function $f(z) := \frac{p(z)}{q(z)}$ where p and q are polynomials and q is not identically zero. A point $z_0 \in \mathbb{C}$ where $f(z_0) = 0$ (and therefore $p(z_0) = 0$) is called a *zero*. A point $z_0 \in \mathbb{C}$ is called an *singularity* of f if $q(z_0) = 0$. As all zeros are isolated and so all singularities of rational functions are isolated and so are called an *isolated singularity*. An isolated singularity is called *removable* if $\lim_{z \mapsto z_0} f(z)$ exists. An isolated singularity is called a *pole* if $\lim_{z \mapsto z_0} |f(z)| = \infty$. We say f has pole at ∞ if

$$\lim_{z \to \infty} |f(z)| = \infty,$$

that is, if for every $M > 0$ there exists an $R > 0$ such that $|f(z)| > M$ for all z with $|z| > R$.

***Exercise* 11.5.4:** *Show that a rational function which is not identically zero has at most finitely many zeros and singularities. In fact, show that if p is a polynomial of degree $n > 0$ it has at most n zeros.*
Hint: If z_0 is a zero of p, then without loss of generality assume $z_0 = 0$. Then use induction.

***Exercise* 11.5.5:** *Prove that if z_0 is a removable singularity of a rational function $f(z) := \frac{p(z)}{q(z)}$, then there exist polynomials $\widetilde{p}$ and $\widetilde{q}$ such that $\widetilde{q}(z_0) \neq 0$ and $f(z) = \frac{\widetilde{p}(z)}{\widetilde{q}(z)}$.*
Hint: Without loss of generality assume $z_0 = 0$.

***Exercise* 11.5.6:** *Given a rational function f with an isolated singularity at z_0, show that z_0 is either removable or a pole.*
Hint: See the previous exercise.

***Exercise* 11.5.7:** *Let f be a rational function and $S \subset \mathbb{C}$ is the set of the singularities of f. Prove that f is equal to a polynomial on $\mathbb{C} \setminus S$ if and only if f has a pole at infinity and all the singularities are removable.*
Hint: See previous exercises.

11.6 Equicontinuity and the Arzelà–Ascoli theorem

Note: 2 lectures

We would like an analogue of Bolzano–Weierstrass. Something to the tune of "every bounded sequence of functions (with some property) has a convergent subsequence." Matters are not as simple even for continuous functions. Not every bounded sequence in the metric space $C([0,1],\mathbb{R})$ has a convergent subsequence.

Definition 11.6.1. Let X be a set. Let $f_n \colon X \to \mathbb{C}$ be functions in a sequence. We say that $\{f_n\}_{n=1}^\infty$ is *pointwise bounded* if for every $x \in X$, there is an $M_x \in \mathbb{R}$ such that

$$|f_n(x)| \leq M_x \qquad \text{for all } n \in \mathbb{N}.$$

We say that $\{f_n\}_{n=1}^\infty$ is *uniformly bounded* if there is an $M \in \mathbb{R}$ such that

$$|f_n(x)| \leq M \qquad \text{for all } n \in \mathbb{N} \text{ and all } x \in X.$$

If X is a compact metric space, then a sequence in $C(X,\mathbb{C})$ is uniformly bounded if it is bounded as a set in the metric space $C(X,\mathbb{C})$ using the uniform norm.

Example 11.6.2: There exist sequences of continuous functions on $[0,1]$ that are uniformly bounded but contain no subsequence converging even pointwise. Let us state without proof that $f_n(x) := \sin(2\pi n x)$ is one such sequence. Below we will show that there must always exist a subsequence converging at countably many points, but $[0,1]$ is uncountable.

Example 11.6.3: The sequence $f_n(x) := x^n$ of continuous functions on $[0,1]$ is uniformly bounded, but contains no subsequence that converges uniformly, although the sequence converges pointwise (to a discontinuous function).

Example 11.6.4: The sequence $\{f_n\}_{n=1}^\infty$ of functions in $C([0,1],\mathbb{R})$ given by $f_n(x) := \frac{n^3 x}{1+n^4x^2}$ converges pointwise to the zero function (obvious at $x = 0$, and for $x > 0$, we have $\frac{n^3x}{1+n^4x^2} \leq \frac{1}{nx}$). As for each x, $\{f_n(x)\}_{n=1}^\infty$ converges to 0, it is bounded so $\{f_n\}_{n=1}^\infty$ is pointwise bounded.

Via calculus, we find that the maximum of f_n on $[0,1]$ occurs at the critical point $x = 1/n^2$:

$$\|f_n\|_{[0,1]} = f_n\left(1/n^2\right) = n/2.$$

So $\lim_{n\to\infty} \|f_n\|_{[0,1]} = \infty$, and this sequence is not uniformly bounded.

When the domain is countable, we can locate a subsequence converging at least pointwise. The proof uses a very common and useful diagonal argument.

Proposition 11.6.5. *Let X be a countable set and $f_n \colon X \to \mathbb{C}$ give a pointwise bounded sequence of functions. Then $\{f_n\}_{n=1}^\infty$ has a subsequence that converges pointwise.*

Proof. Let $x_1, x_2, x_3, \ldots$ be an enumeration of the elements of X. The sequence $\{f_n(x_1)\}_{n=1}^{\infty}$ is bounded and hence we have a subsequence of $\{f_n\}_{n=1}^{\infty}$, which we denote by $\{f_{1,k}\}_{k=1}^{\infty}$, such that $\{f_{1,k}(x_1)\}_{k=1}^{\infty}$ converges. Next $\{f_{1,k}(x_2)\}_{k=1}^{\infty}$ is bounded and so $\{f_{1,k}\}_{k=1}^{\infty}$ has a subsequence $\{f_{2,k}\}_{k=1}^{\infty}$ such that $\{f_{2,k}(x_2)\}_{k=1}^{\infty}$ converges. Note that $\{f_{2,k}(x_1)\}_{k=1}^{\infty}$ is still convergent.

In general, we have a sequence $\{f_{m,k}\}_{k=1}^{\infty}$, which is a subsequence of $\{f_{m-1,k}\}_{k=1}^{\infty}$, such that $\{f_{m,k}(x_j)\}_{k=1}^{\infty}$ converges for $j = 1, 2, \ldots, m$. We let $\{f_{m+1,k}\}_{k=1}^{\infty}$ be a subsequence of $\{f_{m,k}\}_{k=1}^{\infty}$ such that $\{f_{m+1,k}(x_{m+1})\}_{k=1}^{\infty}$ converges (and hence it converges for all x_j for $j = 1, 2, \ldots, m+1$). Rinse and repeat.

If X is finite, we are done as the process stops at some point. If X is countably infinite, we pick the sequence $\{f_{k,k}\}_{k=1}^{\infty}$. This is a subsequence of the original sequence $\{f_n\}_{n=1}^{\infty}$. For every m, the tail $\{f_{k,k}\}_{k=m}^{\infty}$ is a subsequence of $\{f_{m,k}\}_{k=1}^{\infty}$ and hence for any m the sequence $\{f_{k,k}(x_m)\}_{k=1}^{\infty}$ converges. □

For larger than countable sets, we need the functions of the sequence to be related. When we look at continuous functions, the concept we need is equicontinuity.

Definition 11.6.6. Let (X, d) be a metric space. A set S of functions $f \colon X \to \mathbb{C}$ is *uniformly equicontinuous* if for every $\epsilon > 0$, there is a $\delta > 0$ such that if $x, y \in X$ with $d(x, y) < \delta$, we have

$$|f(x) - f(y)| < \epsilon \qquad \text{for all } f \in S.$$

Notice that functions in a uniformly equicontinuous sequence are all uniformly continuous. It is not hard to show that a finite set of uniformly continuous functions is uniformly equicontinuous. The definition is really interesting if S is infinite.

Just as for continuity, one can define equicontinuity at a point. That is, S is *equicontinuous* at $x \in X$ if for every $\epsilon > 0$, there is a $\delta > 0$ such that for $y \in X$ with $d(x, y) < \delta$, we have $|f(x) - f(y)| < \epsilon$ for all $f \in S$. We will only deal with compact X here, and one can prove (exercise) that for a compact metric space X, if S is equicontinuous at every $x \in X$, then it is uniformly equicontinuous. For simplicity we stick to uniform equicontinuity.

Proposition 11.6.7. *Suppose (X, d) is a compact metric space, $f_n \in C(X, \mathbb{C})$, and $\{f_n\}_{n=1}^{\infty}$ converges uniformly, then $\{f_n\}_{n=1}^{\infty}$ is uniformly equicontinuous.*

Proof. Let $\epsilon > 0$ be given. As $\{f_n\}_{n=1}^{\infty}$ converges uniformly, there is an $N \in \mathbb{N}$ such that for all $n \geq N$

$$|f_n(x) - f_N(x)| < \epsilon/3 \qquad \text{for all } x \in X.$$

As X is compact, every continuous function is uniformly continuous. So $\{f_1, f_2, \ldots, f_N\}$ is a finite set of uniformly continuous functions. And so, as we mentioned above, the set is uniformly equicontinuous. Hence there is a $\delta > 0$ such that

$$|f_j(x) - f_j(y)| < \epsilon/3 < \epsilon$$

whenever $d(x, y) < \delta$ and $1 \leq j \leq N$.

Take $n > N$. For $d(x, y) < \delta$, we have

$$|f_n(x) - f_n(y)| \leq |f_n(x) - f_N(x)| + |f_N(x) - f_N(y)| + |f_N(y) - f_n(y)| < \epsilon/3 + \epsilon/3 + \epsilon/3 = \epsilon. \qquad \square$$

Proposition 11.6.8. *A compact metric space (X,d) contains a countable dense subset, that is, there exists a countable $D \subset X$ such that $\overline{D} = X$.*

Proof. For each $n \in \mathbb{N}$ there are finitely many balls of radius $1/n$ that cover X (as X is compact). That is, for every n, there exists a finite set of points $x_{n,1}, x_{n,2}, \ldots, x_{n,k_n}$ such that

$$X = \bigcup_{j=1}^{k_n} B(x_{n,j}, 1/n).$$

Let $D := \bigcup_{n=1}^{\infty} \{x_{n,1}, x_{n,2}, \ldots, x_{n,k_n}\}$. The set D is countable as it is a countable union of finite sets. For every $x \in X$ and every $\epsilon > 0$, there exists an n such that $1/n < \epsilon$ and an $x_{n,j} \in D$ such that

$$x \in B(x_{n,j}, 1/n) \subset B(x_{n,j}, \epsilon).$$

Hence $x \in \overline{D}$, so $\overline{D} = X$, and D is dense. ◻

We are now ready for the main result of this section, the Arzelà–Ascoli theorem* about existence of convergent subsequences.

Theorem 11.6.9 (Arzelà–Ascoli)**.** *Let (X,d) be a compact metric space, and let $\{f_n\}_{n=1}^{\infty}$ be pointwise bounded and uniformly equicontinuous sequence of functions $f_n \in C(X,\mathbb{C})$. Then $\{f_n\}_{n=1}^{\infty}$ is uniformly bounded and $\{f_n\}_{n=1}^{\infty}$ contains a uniformly convergent subsequence.*

Basically, a uniformly equicontinuous sequence in the metric space $C(X,\mathbb{C})$ that is pointwise bounded is bounded (in $C(X,\mathbb{C})$) and furthermore contains a convergent subsequence in $C(X,\mathbb{C})$.

As we mentioned before, as X is compact, it is enough to just assume that $\{f_n\}_{n=1}^{\infty}$ is equicontinuous as uniform equicontinuity is automatic via an exercise.

Proof. We first show that the sequence is uniformly bounded. By uniform equicontinuity, there is a $\delta > 0$ such that for all $x \in X$ and all $n \in \mathbb{N}$,

$$B(x,\delta) \subset f_n^{-1}\bigl(B(f_n(x), 1)\bigr).$$

The space X is compact, so there exist $x_1, x_2, \ldots, x_k$ such that

$$X = \bigcup_{j=1}^{k} B(x_j, \delta).$$

As $\{f_n\}_{n=1}^{\infty}$ is pointwise bounded there exist $M_1, M_2, \ldots, M_k$ such that for $j = 1, 2, \ldots, k$,

$$|f_n(x_j)| \leq M_j \qquad \text{for all } n.$$

*Named after the Italian mathematicians Cesare Arzelà (1847–1912), and Giulio Ascoli (1843–1896).

Let $M := 1 + \max\{M_1, M_2, \ldots, M_k\}$. Given any $x \in X$, there is a j such that $x \in B(x_j, \delta)$. Therefore, for all n, we have $x \in f_n^{-1}\bigl(B(f_n(x_j), 1)\bigr)$, or in other words

$$\lvert f_n(x) - f_n(x_j)\rvert < 1.$$

By the reverse triangle inequality,

$$\lvert f_n(x)\rvert < 1 + \lvert f_n(x_j)\rvert \le 1 + M_j \le M.$$

As x was arbitrary, $\{f_n\}_{n=1}^{\infty}$ is uniformly bounded.

Next, pick a countable dense subset $D \subset X$. By Proposition 11.6.5, we find a subsequence $\{f_{n_j}\}_{j=1}^{\infty}$ that converges pointwise on D. Write $g_j := f_{n_j}$ for simplicity. The sequence $\{g_n\}_{n=1}^{\infty}$ is uniformly equicontinuous. Let $\epsilon > 0$ be given, then there exists a $\delta > 0$ such that for all $x \in X$ and all $n \in \mathbb{N}$,

$$B(x, \delta) \subset g_n^{-1}\bigl(B(g_n(x), \epsilon/3)\bigr).$$

By density of D and because δ is fixed, every $x \in X$ is in $B(y, \delta)$ for some $y \in D$. By compactness of X, there is a finite subset $\{x_1, x_2, \ldots, x_k\} \subset D$ such that

$$X = \bigcup_{j=1}^{k} B(x_j, \delta).$$

As $\{x_1, x_2, \ldots, x_k\}$ is a finite set and $\{g_n\}_{n=1}^{\infty}$ converges pointwise on D, there exists a single N such that for all $n, m \ge N$,

$$\lvert g_n(x_j) - g_m(x_j)\rvert < \epsilon/3 \qquad \text{for all } j = 1, 2, \ldots, k.$$

Let $x \in X$ be arbitrary. There is some j such that $x \in B(x_j, \delta)$ and so for all $\ell \in \mathbb{N}$,

$$\lvert g_\ell(x) - g_\ell(x_j)\rvert < \epsilon/3.$$

So for $n, m \ge N$,

$$\begin{aligned} \lvert g_n(x) - g_m(x)\rvert &\le \lvert g_n(x) - g_n(x_j)\rvert + \lvert g_n(x_j) - g_m(x_j)\rvert + \lvert g_m(x_j) - g_m(x)\rvert \\ &< \epsilon/3 + \epsilon/3 + \epsilon/3 = \epsilon. \end{aligned}$$

Hence, $\{g_n\}_{n=1}^{\infty}$ is uniformly Cauchy. By completeness of $\mathbb{C}$, it is uniformly convergent. □

Corollary 11.6.10. *Let (X, d) be a compact metric space. Let $S \subset C(X, \mathbb{C})$ be a closed, bounded and uniformly equicontinuous set. Then S is compact.*

The theorem says that S is sequentially compact and that means compact in a metric space. Recall that the closed unit ball in $C\bigl([0,1], \mathbb{R}\bigr)$, and therefore also in $C\bigl([0,1], \mathbb{C}\bigr)$, is not compact. Hence it cannot be a uniformly equicontinuous set.

Corollary 11.6.11. *Suppose $\{f_n\}_{n=1}^{\infty}$ is a sequence of differentiable functions on $[a, b]$, $\{f_n'\}_{n=1}^{\infty}$ is uniformly bounded, and there is an $x_0 \in [a, b]$ such that $\{f_n(x_0)\}_{n=1}^{\infty}$ is bounded. Then there exists a uniformly convergent subsequence $\{f_{n_j}\}_{j=1}^{\infty}$.*

Proof. The trick is to use the mean value theorem. If M is the uniform bound on $\{f_n'\}_{n=1}^{\infty}$, then by the mean value theorem for every n

$$|f_n(x) - f_n(y)| \leq M|x-y| \qquad \text{for all } x, y \in X.$$

All the f_n are Lipschitz with the same constant and hence the sequence is uniformly equicontinuous.

Suppose $|f_n(x_0)| \leq M_0$ for all n. For all $x \in [a,b]$,

$$|f_n(x)| \leq |f_n(x_0)| + |f_n(x) - f_n(x_0)| \leq M_0 + M|x - x_0| \leq M_0 + M(b-a).$$

So $\{f_n\}_{n=1}^{\infty}$ is uniformly bounded. We apply Arzelà–Ascoli to find the subsequence. □

A classic application of the corollary above to Arzelà–Ascoli in the theory of differential equations is to prove the Peano existence theorem, that is, the existence of solutions to ordinary differential equations. See Exercise 11.6.11 below.

Another application of Arzelà–Ascoli using the same idea as the corollary above is the following. Take a continuous $k \colon [0,1] \times [0,1] \to \mathbb{C}$. For every $f \in C\big([0,1],\mathbb{C}\big)$ define

$$T\big(f\big)(x) := \int_0^1 f(t)\, k(x,t)\, dt.$$

In exercises to earlier sections you have shown that T is a linear operator on $C\big([0,1],\mathbb{C}\big)$. Via Arzelà–Ascoli, we also find (exercise) that the image of the unit ball of functions

$$T\big(B(0,1)\big) = \Big\{ Tf \in C\big([0,1],\mathbb{C}\big) : \|f\|_{[0,1]} < 1 \Big\}$$

has compact closure, usually called *relatively compact*. Such an operator is called a *compact operator*. And they are very useful. Generally operators defined by integration tend to be compact.

11.6.1 Exercises

***Exercise* 11.6.1:** *Let $f_n \colon [-1,1] \to \mathbb{R}$ be given by $f_n(x) := \frac{nx}{1+(nx)^2}$. Prove that the sequence is uniformly bounded, converges pointwise to 0, yet there is no subsequence that converges uniformly. Which hypothesis of Arzelà–Ascoli is not satisfied? Prove your assertion.*

***Exercise* 11.6.2:** *Define $f_n \colon \mathbb{R} \to \mathbb{R}$ by $f_n(x) := \frac{1}{(x-n)^2+1}$. Prove that this sequence is uniformly bounded, uniformly equicontinuous, the sequence converges pointwise to zero, yet there is no subsequence that converges uniformly. Which hypothesis of Arzelà–Ascoli is not satisfied? Prove your assertion.*

***Exercise* 11.6.3:** *Let (X,d) be a compact metric space, $C > 0$, $0 < \alpha \leq 1$, and suppose $f_n \colon X \to \mathbb{C}$ are functions such as $\big|f_n(x) - f_n(y)\big| \leq C d(x,y)^{\alpha}$ for all $x, y \in X$ and $n \in \mathbb{N}$. Suppose also that there is a point $p \in X$ such that $f_n(p) = 0$ for all n. Show that there exists a uniformly convergent subsequence converging to an $f \colon X \to \mathbb{C}$ that also satisfies $f(p) = 0$ and $\big|f(x) - f(y)\big| \leq C d(x,y)^{\alpha}$.*

***Exercise* 11.6.4:** *Let $T: C\big([0,1],\mathbb{C}\big) \to C\big([0,1],\mathbb{C}\big)$ be the operator given by*

$$T\big(f\big)(x) := \int_0^x f(t)\,dt.$$

(That T is linear and that Tf is continuous follows from linearity of the integral and the fundamental theorem of calculus.)

a) *Show that T takes the unit ball centered at 0 in $C\big([0,1],\mathbb{C}\big)$ into a relatively compact set (a set with compact closure). That is, T is a compact operator.*
Hint: See Exercise 7.4.20 in volume I.

b) *Let $C \subset C\big([0,1],\mathbb{C}\big)$ the closed unit ball, prove that the image $T(C)$ is not closed (though it is relatively compact).*

***Exercise* 11.6.5:** *Given $k \in C\big([0,1]\times[0,1],\mathbb{C}\big)$, define the operator $T: C\big([0,1],\mathbb{C}\big) \to C\big([0,1],\mathbb{C}\big)$ by*

$$T\big(f\big)(x) := \int_0^1 f(t)\,k(x,t)\,dt.$$

Show that T takes the unit ball centered at 0 in $C\big([0,1],\mathbb{C}\big)$ into a relatively compact set (a set with compact closure). That is, T is a compact operator.
Hint: See Exercise 7.4.20 in volume I.
Note: That T is a well-defined linear operator was proved in Exercise 8.1.6.

***Exercise* 11.6.6:** *Suppose $S^1 \subset \mathbb{C}$ is the unit circle, that is, the set where $|z| = 1$. Suppose the continuous functions $f_n: S^1 \to \mathbb{C}$ are uniformly bounded. Let $\gamma: [0,1] \to S^1$ be a parametrization of S^1, and $g(z,w)$ a continuous function on $C(0,1)\times S^1$ (here $C(0,1) \subset \mathbb{C}$ is the closed unit ball). Define the functions $F_n: C(0,1) \to \mathbb{C}$ by the path integral (see §9.2)*

$$F_n(z) := \int_\gamma f_n(w)\,g(z,w)\,ds(w).$$

Show that $\{F_n\}_{n=1}^\infty$ has a uniformly convergent subsequence.

***Exercise* 11.6.7:** *Suppose (X,d) is a compact metric space, $\{f_n\}_{n=1}^\infty$ a uniformly equicontinuous sequence of functions in $C(X,\mathbb{C})$. Suppose $\{f_n\}_{n=1}^\infty$ converges pointwise. Show that it converges uniformly.*

***Exercise* 11.6.8:** *Suppose that $\{f_n\}_{n=1}^\infty$ is a uniformly equicontinuous uniformly bounded sequence of 2π-periodic functions $f_n: \mathbb{R} \to \mathbb{R}$. Show that there is a uniformly convergent subsequence.*

***Exercise* 11.6.9:** *Show that for a compact metric space X, a sequence $\{f_n\}_{n=1}^\infty$ that is equicontinuous at every $x \in X$ is uniformly equicontinuous.*

***Exercise* 11.6.10:** *Define $f_n: [0,1] \to \mathbb{C}$ by $f_n(t) := e^{i(2\pi t+n)}$, which gives a uniformly equicontinuous uniformly bounded sequence. Prove a stronger conclusion than that of Arzelà–Ascoli for this sequence. Let $\gamma \in \mathbb{R}$ be given, and define $g(t) := e^{i(2\pi t+\gamma)}$. Show that there exists a subsequence of $\{f_n\}_{n=1}^\infty$ converging uniformly to g.*
Hint: Feel free to use the Kronecker density theorem*: *The sequence $\{e^{in}\}_{n=1}^\infty$ is dense in the unit circle.*

*Named after the German mathematician Leopold Kronecker (1823–1891).

Exercise* 11.6.11*: *Prove the* Peano existence theorem *(note the weaker hypotheses than Picard, but also the lack of uniqueness in this theorem):*

Theorem: Suppose $F: I \times J \to \mathbb{R}$ is a continuous function where $I, J \subset \mathbb{R}$ are closed bounded intervals, let I° and J° be their interiors, and let $(x_0, y_0) \in I^\circ \times J^\circ$. Then there exists an $h > 0$ and a differentiable function $f: [x_0 - h, x_0 + h] \to J \subset \mathbb{R}$, such that

$$f'(x) = F\bigl(x, f(x)\bigr) \qquad \textit{and} \qquad f(x_0) = y_0.$$

Use the following outline:

a) *We wish to define the Picard iterates, that is, set $f_0(x) := y_0$, and*

$$f_{n+1}(x) := y_0 + \int_{x_0}^{x} F\bigl(t, f_n(t)\bigr)\, dt.$$

Prove that there exists an $h > 0$ such that $f_n: [x_0 - h, x_0 + h] \to \mathbb{C}$ is well-defined for all n. Hint: F is bounded (why?).

b) *Show that $\{f_n\}_{n=1}^\infty$ is equicontinuous and bounded, in fact it is Lipschitz with a uniform Lipschitz constant. Arzelà–Ascoli then says that there exists a uniformly convergent subsequence $\{f_{n_k}\}_{k=1}^\infty$.*

c) *Prove $\bigl\{F\bigl(x, f_{n_k}(x)\bigr)\bigr\}_{k=1}^\infty$ converges uniformly on $[x_0 - h, x_0 + h]$. Hint: F is uniformly continuous (why?).*

d) *Finish the proof of the theorem by taking the limit under the integral and applying the fundamental theorem of calculus.*

11.7 The Stone–Weierstrass theorem

Note: 3 lectures

11.7.1 Weierstrass approximation

Perhaps surprisingly, even a very badly behaved continuous function is a uniform limit of polynomials. We cannot really get any "nicer" functions than polynomials. The idea of the proof is a very common approximation or "smoothing" idea (convolution with an approximate delta function) that has applications far beyond pure mathematics.

Theorem 11.7.1 (Weierstrass approximation theorem)**.** *If $f \colon [a,b] \to \mathbb{C}$ is continuous, then there exists a sequence $\{p_n\}_{n=1}^{\infty}$ of polynomials converging to f uniformly on $[a,b]$. Furthermore, if f is real-valued, we can find p_n with real coefficients.*

Proof. For $x \in [0,1]$, define

$$g(x) \coloneqq f\bigl((b-a)x+a\bigr) - f(a) - x\bigl(f(b)-f(a)\bigr).$$

If we prove the theorem for g and find the sequence $\{p_n\}_{n=1}^{\infty}$ for g, it is proved for f as we simply composed with an invertible affine function and added an affine function to f: We reverse the process and apply that to our p_n, to obtain polynomials approximating f. The function g is defined on $[0,1]$ and $g(0) = g(1) = 0$. For simplicity, assume that g is defined on $\mathbb{R}$ by letting $g(x) \coloneqq 0$ if $x < 0$ or $x > 1$. This extended g is continuous.

Define

$$c_n \coloneqq \left(\int_{-1}^{1} (1-x^2)^n \, dx \right)^{-1}, \qquad q_n(x) \coloneqq c_n (1-x^2)^n .$$

The choice of c_n is so that $\int_{-1}^{1} q_n(x)\,dx = 1$. See Figure 11.8.

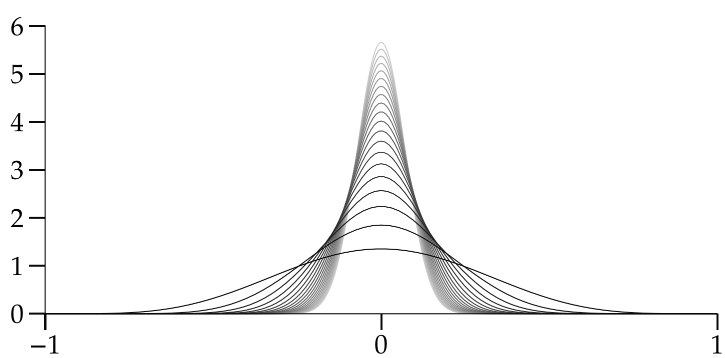

Figure 11.8: Plot of the approximate delta functions q_n on $[-1,1]$ for $n = 5, 10, 15, 20, \ldots, 100$ with higher n in lighter shade.

The functions q_n are peaks around 0 (ignoring what happens outside of $[-1,1]$) that get narrower and taller as n increases, while the area underneath is always 1. A classic approximation idea is to do a *convolution* integral with peaks like this: For for $x \in [0,1]$, let

$$p_n(x) := \int_0^1 g(t) q_n(x-t)\, dt \qquad \left(= \int_{-\infty}^{\infty} g(t) q_n(x-t)\, dt \right).$$

The idea of this convolution is that we do a "weighted average" of the function g around the point x using q_n as the weight. See Figure 11.9.

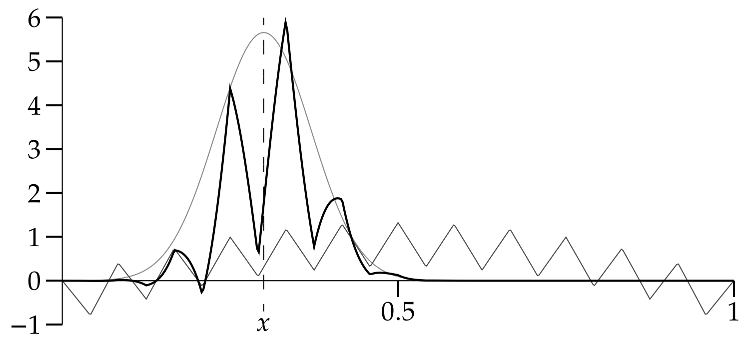

Figure 11.9: For $x = 0.3$, the plot of $q_{100}(x-t)$ (light gray peak centered at x), some continuous function $g(t)$ (the jagged line) and the product $g(t)q_{100}(x-t)$ (the bold line).

As q_n is a narrow peak, the integral mostly sees the values of g that are close to x and it does the weighted average of them. When the peak gets narrower, we compute this average closer to x and we expect the result to get closer to the value of $g(x)$. Really, we are approximating what is called a delta function* (don't worry if you have not heard of this concept), and functions like q_n are often called *approximate delta functions*. We could do this with any set of polynomials that look like narrower and narrower peaks near zero. These just happen to be the simplest ones. We only need this behavior on $[-1,1]$ as the convolution sees nothing further than this as g is zero outside $[0,1]$.

Because q_n is a polynomial, we write

$$q_n(x-t) = a_0(t) + a_1(t)\, x + \cdots + a_{2n}(t)\, x^{2n},$$

where $a_k(t)$ are polynomials in t, and hence integrable functions. So

$$\begin{aligned} p_n(x) &= \int_0^1 g(t) q_n(x-t)\, dt \\ &= \left(\int_0^1 g(t) a_0(t)\, dt \right) + \left(\int_0^1 g(t) a_1(t)\, dt \right) x + \cdots + \left(\int_0^1 g(t) a_{2n}(t)\, dt \right) x^{2n}. \end{aligned}$$

*The delta function is not actually a function, it is a "thing" that should give "$\int_{-\infty}^{\infty} g(t)\delta(x-t)\, dt = g(x)$."

In other words, p_n is a polynomial* in x. If $g(t)$ is real-valued, then the functions $g(t)a_j(t)$ are real-valued and p_n has real coefficients, proving the "furthermore" part of the theorem.

We still need to prove that $\{p_n\}_{n=1}^{\infty}$ converges to g. We start with estimating the size of c_n. For $x \in [0,1]$, we have that $1 - x \leq 1 - x^2$. We estimate

$$\begin{aligned} c_n^{-1} = \int_{-1}^{1} (1-x^2)^n \, dx &= 2 \int_0^1 (1-x^2)^n \, dx \\ &\geq 2 \int_0^1 (1-x)^n \, dx = \frac{2}{n+1}. \end{aligned}$$

So $c_n \leq \frac{n+1}{2} \leq n$. Let us see how small q_n is if we ignore some small interval around the origin, where the peak is. Given any $\delta > 0$, $\delta < 1$, we have that for all x such that $\delta \leq |x| \leq 1$,

$$q_n(x) \leq c_n (1-\delta^2)^n \leq n(1-\delta^2)^n,$$

because q_n is increasing on $[-1,0]$ and decreasing on $[0,1]$. By the ratio test, $n(1-\delta^2)^n$ goes to 0 as n goes to infinity.

The function q_n is even, $q_n(t) = q_n(-t)$, and g is zero outside of $[0,1]$. So for $x \in [0,1]$,

$$p_n(x) = \int_0^1 g(t) q_n(x-t) \, dt = \int_{-x}^{1-x} g(x+t) q_n(-t) \, dt = \int_{-1}^{1} g(x+t) q_n(t) \, dt.$$

Let $\epsilon > 0$ be given. As $[-1,2]$ is compact and g is continuous on $[-1,2]$, we have that g is uniformly continuous. Pick $0 < \delta < 1$ such that if $|x-y| < \delta$ (and $x, y \in [-1,2]$), then

$$|g(x) - g(y)| < \frac{\epsilon}{2}.$$

Let M be such that $|g(x)| \leq M$ for all x. Let N be such that for all $n \geq N$,

$$4Mn(1-\delta^2)^n < \frac{\epsilon}{2}.$$

Note that $\int_{-1}^{1} q_n(t) \, dt = 1$ and $q_n(t) \geq 0$ on $[-1,1]$. So for $n \geq N$ and every $x \in [0,1]$,

$$\begin{aligned} |p_n(x) - g(x)| &= \left| \int_{-1}^{1} g(x+t) q_n(t) \, dt - g(x) \int_{-1}^{1} q_n(t) \, dt \right| \\ &= \left| \int_{-1}^{1} \big(g(x+t) - g(x)\big) q_n(t) \, dt \right| \\ &\leq \int_{-1}^{1} |g(x+t) - g(x)| q_n(t) \, dt \\ &= \int_{-1}^{-\delta} |g(x+t) - g(x)| q_n(t) \, dt \quad + \int_{-\delta}^{\delta} |g(x+t) - g(x)| q_n(t) \, dt \\ &\quad + \int_{\delta}^{1} |g(x+t) - g(x)| q_n(t) \, dt \end{aligned}$$

*Do note that the functions a_j depend on n, so the coefficients of p_n change as n changes.

$$
\begin{aligned}
&\le 2M \int_{-1}^{-\delta} q_n(t)\,dt \quad + \quad \frac{\epsilon}{2}\int_{-\delta}^{\delta} q_n(t)\,dt \quad + \quad 2M\int_{\delta}^{1} q_n(t)\,dt \\
&\le 2Mn{(1-\delta^2)}^n(1-\delta) \quad + \quad \frac{\epsilon}{2} \quad + \quad 2Mn{(1-\delta^2)}^n(1-\delta) \\
&< 4Mn{(1-\delta^2)}^n + \frac{\epsilon}{2} < \epsilon.
\end{aligned}
$$
□

A convolution often inherits some property of the functions we are convolving. In our case the convolution p_n inherited the property of being a polynomial from q_n. The same idea of the proof is often used to get other properties. If q_n or g is infinitely differentiable, so is p_n. If q_n or g is a solution to a linear differential equation, so is p_n. Etc.

Let us note an immediate application of the Weierstrass theorem. We have already seen that countable dense subsets can be very useful.

Corollary 11.7.2. *The metric spaces $C\big([a,b],\mathbb{R}\big)$ and $C\big([a,b],\mathbb{C}\big)$ each contain a countable dense subset.*

Proof. Without loss of generality, consider only $C\big([a,b],\mathbb{R}\big)$ (why?). Real polynomials are dense in $C\big([a,b],\mathbb{R}\big)$ by Weierstrass. If we show that every real polynomial can be approximated by polynomials with rational coefficients, we are done. Indeed, there are only countably many rational numbers and so there are only countably many polynomials with rational coefficients (a countable union of countable sets is countable).

Further without loss of generality, suppose $[a,b] = [0,1]$. Let

$$p(x) := \sum_{k=0}^{n} a_k\, x^k$$

be a polynomial of degree n where $a_k \in \mathbb{R}$. Given $\epsilon > 0$, pick $b_k \in \mathbb{Q}$ such that $|a_k - b_k| < \frac{\epsilon}{n+1}$. Then if we let

$$q(x) := \sum_{k=0}^{n} b_k\, x^k,$$

we have

$$\lvert p(x) - q(x)\rvert = \left\lvert \sum_{k=0}^{n} (a_k - b_k)x^k \right\rvert \le \sum_{k=0}^{n} \lvert a_k - b_k\rvert x^k \le \sum_{k=0}^{n} \lvert a_k - b_k\rvert < \sum_{k=0}^{n} \frac{\epsilon}{n+1} = \epsilon.$$
□

Remark 11.7.3. While we will not prove so, the corollary above implies that $C\big([a,b],\mathbb{C}\big)$ has the same cardinality as $\mathbb{R}$, which may be a bit surprising. The set of all functions $[a,b] \to \mathbb{C}$ has cardinality strictly greater than the cardinality of $\mathbb{R}$, it has the cardinality of the power set of $\mathbb{R}$. So the set of continuous functions is a very tiny subset of the set of all functions.

Warning! The fact that every continuous function $f\colon [-1,1] \to \mathbb{C}$ (or any interval $[a,b]$) can be uniformly approximated by polynomials

$$\sum_{k=0}^{n} a_k\, x^k$$

does not mean that every continuous f is analytic, that is, equal to a power series

$$\sum_{k=0}^{\infty} c_k \, x^k .$$

An analytic function is infinitely differentiable, however, the function $|x|$ is continuous and near the origin approximable by polynomials, and so provides a counterexample.

The key distinction is that the polynomials coming from the Weierstrass theorem are not the partial sums of a power series. For each one, the coefficients a_k above can be completely different—they do not need to come from a single sequence $\{c_k\}_{k=1}^{\infty}$.

Interestingly, to generalize Weierstrass, we will only need to use it to approximate the absolute value function by polynomials without a constant term.

Corollary 11.7.4. *Let $[-a, a]$ be an interval. Then there is a sequence of real polynomials $\{p_n\}_{n=1}^{\infty}$ that converges uniformly to $|x|$ on $[-a, a]$ and such that $p_n(0) = 0$ for all n.*

Proof. As $f(x) := |x|$ is continuous and real-valued on $[-a, a]$, the Weierstrass theorem gives a sequence of real polynomials $\{\widetilde{p}_n\}_{n=1}^{\infty}$ that converges to f uniformly on $[-a, a]$. Let

$$p_n(x) := \widetilde{p}_n(x) - \widetilde{p}_n(0).$$

Obviously $p_n(0) = 0$.

Given $\epsilon > 0$, let N be such that for $n \geq N$, we have $\left|\widetilde{p}_n(x) - |x|\right| < \epsilon/2$ for all $x \in [-a, a]$. In particular, $|\widetilde{p}_n(0)| < \epsilon/2$. Then for $n \geq N$,

$$\left|p_n(x) - |x|\right| = \left|\widetilde{p}_n(x) - \widetilde{p}_n(0) - |x|\right| \leq \left|\widetilde{p}_n(x) - |x|\right| + |\widetilde{p}_n(0)| < \epsilon/2 + \epsilon/2 = \epsilon. \qquad \square$$

Generalizing the corollary, we can make the polynomials from the Weierstrass theorem be equal to our target function at one point, not just for $|x|$, but that's the one we will need. It is also possible (see Exercise 11.7.14) to make the polynomials equal at finitely many points by subtracting not a constant but a properly crafted polynomial.

11.7.2 Stone–Weierstrass approximation

We want to abstract away what is not really necessary and prove a general version of the Weierstrass theorem, the Stone–Weierstrass theorem*. Polynomials are dense in the space of continuous functions on a compact interval. What other kind of families of functions are also dense? And if the domain is an arbitrary metric space, then we no longer have polynomials to begin with.

*Named after the American mathematician Marshall Harvey Stone (1903–1989), and the German mathematician Karl Theodor Wilhelm Weierstrass (1815–1897).

Definition 11.7.5. A set $\mathcal{A}$ of complex-valued functions $f: X \to \mathbb{C}$ is said to be an *algebra* (sometimes *complex algebra* or *algebra over* $\mathbb{C}$) if for all $f, g \in \mathcal{A}$ and $c \in \mathbb{C}$, we have

(i) $f + g \in \mathcal{A}$.

(ii) $fg \in \mathcal{A}$.

(iii) $cg \in \mathcal{A}$.

A *real algebra* or an *algebra over* $\mathbb{R}$ is a set of real-valued functions that satisfies the three properties above for $c \in \mathbb{R}$.

We are interested in the case when X is a compact metric space. Then $C(X, \mathbb{C})$ and $C(X, \mathbb{R})$ are metric spaces. Given a set $\mathcal{A} \subset C(X, \mathbb{C})$, the set of all uniform limits is the metric space closure $\overline{\mathcal{A}}$. When we talk about closure of an algebra from now on we mean the closure in $C(X, \mathbb{C})$ as a metric space. Same for $C(X, \mathbb{R})$.

The set $\mathcal{P}$ of all polynomials is an algebra in $C\big([a, b], \mathbb{C}\big)$, and we have shown that its closure $\overline{\mathcal{P}} = C\big([a, b], \mathbb{C}\big)$. That is, it is dense. That is the sort of result that we wish to prove.

We leave the following proposition as an exercise.

Proposition 11.7.6. *Suppose X is a compact metric space. If $\mathcal{A} \subset C(X, \mathbb{C})$ is an algebra, then the closure $\overline{\mathcal{A}}$ is also an algebra. Similarly for a real algebra in $C(X, \mathbb{R})$.*

We distill the properties of polynomials that are sufficient for an approximation theorem.

Definition 11.7.7. Let $\mathcal{A}$ be a set of complex-valued functions defined on a set X.

(i) $\mathcal{A}$ *separates points* if for every $x, y \in X$ with $x \neq y$, there is an $f \in \mathcal{A}$ such that $f(x) \neq f(y)$.

(ii) $\mathcal{A}$ *vanishes at no point* if for every $x \in X$ there is an $f \in \mathcal{A}$ such that $f(x) \neq 0$.

Example 11.7.8: Given any $X \subset \mathbb{R}$ (or $X \subset \mathbb{C}$), the set $\mathcal{P}$ of polynomials in one variable separates points and vanishes at no point on X. That is, $1 \in \mathcal{P}$, so it vanishes at no point. And for $x, y \in X$, $x \neq y$, take $f(t) := t$. Then $f(x) = x \neq y = f(y)$. So $\mathcal{P}$ separates points.

Example 11.7.9: The set of functions of the form

$$f(t) = a_0 + \sum_{n=1}^{k} a_n \cos(nt)$$

is an algebra, which follows by the identity $\cos(mt)\cos(nt) = \frac{\cos((n+m)t)}{2} + \frac{\cos((n-m)t)}{2}$. The algebra vanishes at no point as it contains a constant function. It does not separate points if the domain is an interval $[-a, a]$, as $f(-t) = f(t)$ for all t. It does separate points if the domain is $[0, \pi]$; $\cos(t)$ is one-to-one on $[0, \pi]$.

Example 11.7.10: The set $\mathcal{P}$ of real polynomials with no constant term is an algebra that vanishes at the origin. Clearly, any function in the closure of $\mathcal{P}$ also vanishes at the origin, so the closure of $\mathcal{P}$ cannot be $C\big([0, 1], \mathbb{R}\big)$.

Similarly, the set of constant functions is an algebra that does not separate points. Uniform limits of constants are constants, so we also do not obtain all continuous functions.

It is interesting that these two properties, "vanishes at no point" and "separates points," are sufficient to obtain approximation of any real-valued continuous function. Before we prove this theorem, we note that such an algebra can interpolate a finite number of values exactly. We will state this result only for two points as that is all that we will require.

Proposition 11.7.11. *Suppose $\mathscr{A}$ is an algebra of complex-valued functions on a set X that separates points and vanishes at no point. Suppose x, y are distinct points of X, and $c, d \in \mathbb{C}$. Then there is an $f \in \mathscr{A}$ such that*

$$f(x) = c, \qquad f(y) = d.$$

If $\mathscr{A}$ is a real algebra, the conclusion holds for $c, d \in \mathbb{R}$.

Proof. There must exist an $g, h, k \in \mathscr{A}$ such that $g(x) \neq g(y)$, $h(x) \neq 0$, $k(y) \neq 0$. Let

$$\begin{aligned} f &:= c \frac{\big(g - g(y)\big)h}{\big(g(x) - g(y)\big)h(x)} + d \frac{\big(g - g(x)\big)k}{\big(g(y) - g(x)\big)k(y)} \\ &= c \frac{gh - g(y)h}{g(x)h(x) - g(y)h(x)} + d \frac{gk - g(x)k}{g(y)k(y) - g(x)k(y)}. \end{aligned}$$

We are not dividing by zero (clear from the first formula). Also by the first formula, $f(x) = c$ and $f(y) = d$. By the second formula, $f \in \mathscr{A}$ (as $\mathscr{A}$ is an algebra). □

Theorem 11.7.12 (Stone–Weierstrass, real version)**.** *Let X be a compact metric space and $\mathscr{A}$ a real algebra of real-valued continuous functions on X, such that $\mathscr{A}$ separates points and vanishes at no point. Then the closure $\overline{\mathscr{A}} = C(X, \mathbb{R})$.*

The proof is divided into several claims.

Claim 1: *If $f \in \overline{\mathscr{A}}$, then $|f| \in \overline{\mathscr{A}}$.*

Proof. The function f is bounded (continuous on a compact set), so there is an M such that $|f(x)| \leq M$ for all $x \in X$. Let $\epsilon > 0$ be given. By the corollary to the Weierstrass theorem, there exists a real polynomial $c_1 y + c_2 y^2 + \cdots + c_n y^n$ (vanishing at $y = 0$) such that

$$\left| |y| - \sum_{k=1}^{n} c_k y^k \right| < \epsilon$$

for all $y \in [-M, M]$. Because $\overline{\mathscr{A}}$ is an algebra and because there is no constant term in the polynomial,

$$\sum_{k=1}^{n} c_k f^k \in \overline{\mathscr{A}}.$$

As $|f(x)| \leq M$, then for all $x \in X$

$$\left| |f(x)| - \sum_{k=1}^{n} c_k \big(f(x)\big)^k \right| < \epsilon.$$

So $|f|$ is in the closure of $\overline{\mathscr{A}}$, which is itself closed. In other words, $|f| \in \overline{\mathscr{A}}$. □

Claim 2: *If* $f \in \overline{\mathcal{A}}$ *and* $g \in \overline{\mathcal{A}}$*, then* $\max(f, g) \in \overline{\mathcal{A}}$ *and* $\min(f, g) \in \overline{\mathcal{A}}$*, where*

$$\big(\max(f,g)\big)(x) := \max\big\{f(x), g(x)\big\}, \qquad \text{and} \qquad \big(\min(f,g)\big)(x) := \min\big\{f(x), g(x)\big\}.$$

Proof. Write:

$$\max(f,g) = \frac{f+g}{2} + \frac{|f-g|}{2}, \qquad \text{and} \qquad \min(f,g) = \frac{f+g}{2} - \frac{|f-g|}{2}.$$

As $\overline{\mathcal{A}}$ is an algebra we are done. □

By induction, the claim is also true for the minimum or maximum of a finite collection of functions.

Claim 3: *Given* $f \in C(X, \mathbb{R})$*,* $x \in X$*, and* $\epsilon > 0$*, there exists a* $g_x \in \overline{\mathcal{A}}$ *with* $g_x(x) = f(x)$ *and*

$$g_x(t) > f(t) - \epsilon \qquad \text{for all } t \in X.$$

Proof. Fix f, x, and ϵ. By Proposition 11.7.11, for every $y \in X$, find an $h_y \in \mathcal{A}$ such that

$$h_y(x) = f(x), \qquad h_y(y) = f(y).$$

As h_y and f are continuous, the function $h_y - f$ is continuous, and the set

$$U_y := \big\{t \in X : h_y(t) > f(t) - \epsilon\big\} = (h_y - f)^{-1}\big((-\epsilon, \infty)\big)$$

is open (it is the inverse image of an open set by a continuous function). Furthermore $y \in U_y$. So the sets U_y cover X. The space X is compact, so there exist finitely many points $y_1, y_2, \dots, y_n$ in X such that

$$X = \bigcup_{k=1}^{n} U_{y_k}.$$

Let

$$g_x := \max(h_{y_1}, h_{y_2}, \dots, h_{y_n}).$$

By Claim 2, $g_x \in \overline{\mathcal{A}}$. See Figure 11.10. Moreover,

$$g_x(t) > f(t) - \epsilon$$

for all $t \in X$, since for every t, there is a y_k such that $t \in U_{y_k}$, and so $h_{y_k}(t) > f(t) - \epsilon$. Finally, $h_y(x) = f(x)$ for all $y \in X$, so $g_x(x) = f(x)$. □

What we have now is for each x a function $g_x \in \overline{\mathcal{A}}$ that is within ϵ of f near x (being continuous), but also g_x is within ϵ of f from at least one side at all points. If we cover X with neighborhoods where g_x is a good approximation, we can repeat the idea of the argument with a minimum to get a function that is within ϵ from both sides.

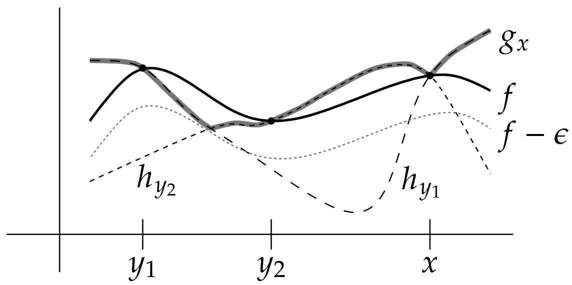

Figure 11.10: Construction of g_x out of two h_{y_1} (longer dashes) and h_{y_2} (shorter dashes).

Claim 4: *If $f \in C(X, \mathbb{R})$ and $\epsilon > 0$ is given, then there exists an $\varphi \in \overline{\mathscr{A}}$ such that*

$$|f(x) - \varphi(x)| < \epsilon.$$

Proof. For every $x \in X$, find the function g_x as in Claim 3. Let

$$V_x := \{t \in X : g_x(t) < f(t) + \epsilon\}.$$

The sets V_x are open as g_x and f are continuous. As $g_x(x) = f(x)$, then $x \in V_x$. So the sets V_x cover X. By compactness of X, there are finitely many points $x_1, x_2, \ldots, x_n$ such that

$$X = \bigcup_{k=1}^{n} V_{x_k}.$$

Let

$$\varphi := \min(g_{x_1}, g_{x_2}, \ldots, g_{x_n}).$$

By Claim 2, $\varphi \in \overline{\mathscr{A}}$. Similarly as before (same argument as in Claim 3), for all $t \in X$,

$$\varphi(t) < f(t) + \epsilon.$$

Since all the g_x satisfy $g_x(t) > f(t) - \epsilon$ for all $t \in X$, $\varphi(t) > f(t) - \epsilon$ as well. Hence, for all t,

$$-\epsilon < \varphi(t) - f(t) < \epsilon,$$

which is the desired conclusion. □

The proof of the theorem follows from Claim 4. The claim states that an arbitrary continuous function is in the closure of $\overline{\mathscr{A}}$, which is already closed. The theorem is proved.

Example 11.7.13: The functions of the form

$$f(t) = \sum_{k=1}^{n} c_k\, e^{kt},$$

for $c_k \in \mathbb{R}$, are dense in $C\big([a,b], \mathbb{R}\big)$. Such functions are a real algebra, which follows from $e^{kt}e^{\ell t} = e^{(k+\ell)t}$. They separate points as e^t is one-to-one. As $e^t > 0$ for all t, the algebra does not vanish at any point.

In general, given a set of functions that separates points and does not vanish at any point, we let these functions *generate* an algebra by considering all the linear combinations of arbitrary multiples of such functions. That is, we consider all real polynomials without constant term of such functions. In the example above, the algebra is generated by e^t. We consider polynomials in e^t without constant term.

Example 11.7.14: We mentioned that the set of all functions of the form

$$a_0 + \sum_{n=1}^{N} a_n \cos(nt)$$

is an algebra. When considered on $[0,\pi]$, it separates points and vanishes nowhere so Stone–Weierstrass applies. As for polynomials, you *do not* want to conclude that every continuous function on $[0,\pi]$ has a uniformly convergent Fourier cosine series, that is, that every continuous function can be written as

$$a_0 + \sum_{n=1}^{\infty} a_n \cos(nt).$$

That is *not true*! There exist continuous functions whose Fourier series does not converge even pointwise let alone uniformly. See §11.8.

To obtain Stone–Weierstrass for complex algebras, we must make an extra assumption.

Definition 11.7.15. An algebra $\mathscr{A}$ is *self-adjoint* if for all $f \in \mathscr{A}$, the function $\bar{f}$ defined by $\bar{f}(x) := \overline{f(x)}$ is in $\mathscr{A}$, where by the bar we mean the complex conjugate.

Theorem 11.7.16 (Stone–Weierstrass, complex version)**.** *Let X be a compact metric space and $\mathscr{A}$ an algebra of complex-valued continuous functions on X, such that $\mathscr{A}$ separates points, vanishes at no point, and is self-adjoint. Then the closure $\overline{\mathscr{A}} = C(X,\mathbb{C})$.*

Proof. Suppose $\mathscr{A}_{\mathbb{R}} \subset \mathscr{A}$ is the set of the real-valued elements of $\mathscr{A}$. For $f \in \mathscr{A}$, write $f = u + iv$ where u and v are real-valued. Then

$$u = \frac{f+\bar{f}}{2}, \qquad v = \frac{f-\bar{f}}{2i}.$$

So $u,v \in \mathscr{A}$ as $\mathscr{A}$ is a self-adjoint algebra, and since they are real-valued $u,v \in \mathscr{A}_{\mathbb{R}}$.

If $x \neq y$, then find an $f \in \mathscr{A}$ such that $f(x) \neq f(y)$. If $f = u+iv$, then it is obvious that either $u(x) \neq u(y)$ or $v(x) \neq v(y)$. So $\mathscr{A}_{\mathbb{R}}$ separates points. Similarly, for every x find $f \in \mathscr{A}$ such that $f(x) \neq 0$. If $f = u + iv$, then either $u(x) \neq 0$ or $v(x) \neq 0$. So $\mathscr{A}_{\mathbb{R}}$ vanishes at no point. The set $\mathscr{A}_{\mathbb{R}}$ is a real algebra, and satisfies the hypotheses of the real Stone–Weierstrass theorem. Given any $f = u + iv \in C(X,\mathbb{C})$, we find $g,h \in \mathscr{A}_{\mathbb{R}}$ such that $|u(t)-g(t)| < \nicefrac{\epsilon}{2}$ and $|v(t)-h(t)| < \nicefrac{\epsilon}{2}$ for all $t \in X$. Next, $g + ih \in \mathscr{A}$, and

$$\begin{aligned} \bigl|f(t) - \bigl(g(t)+ih(t)\bigr)\bigr| &= \bigl|u(t)+iv(t) - \bigl(g(t)+ih(t)\bigr)\bigr| \\ &\leq |u(t)-g(t)| + |v(t)-h(t)| < \nicefrac{\epsilon}{2} + \nicefrac{\epsilon}{2} = \epsilon \end{aligned}$$

for all $t \in X$. So $\overline{\mathscr{A}} = C(X,\mathbb{C})$. $\square$

The self-adjoint requirement is necessary, although it is not so obvious to see it. For an example, see Exercise 11.7.9.

We give an interesting application. When working with functions of two variables, it may be useful to work with functions of the form $f(x)g(y)$ rather than $F(x,y)$. For example, they are easier to integrate. We have the following.

Example 11.7.17: Any continuous $F\colon [0,1]\times[0,1] \to \mathbb{C}$ can be approximated uniformly by functions of the form

$$\sum_{j=1}^{n} f_j(x)g_j(y),$$

where $f_j\colon [0,1] \to \mathbb{C}$ and $g_j\colon [0,1] \to \mathbb{C}$ are continuous.

Proof: It is not hard to see that the functions of the above form are a complex algebra. It is equally easy to show that they vanish nowhere, separate points, and the algebra is self-adjoint. As $[0,1]\times[0,1]$ is compact, Stone–Weierstrass obtains the result.

11.7.3 Exercises

Exercise* 11.7.1*: *Prove Proposition 11.7.6. Hint: If $\{f_n\}_{n=1}^{\infty}$ is a sequence in $C(X,\mathbb{R})$ converging to f, then as f is bounded, show that f_n is uniformly bounded, that is, there exists a single bound for all f_n (and f).*

Exercise* 11.7.2*: *Suppose $X := \mathbb{R}$ (not compact in particular). Show that $f(t) := e^t$ is not possible to uniformly approximate by polynomials on X. Hint: Consider $\left|\frac{e^t}{t^n}\right|$ as $t \to \infty$.*

Exercise* 11.7.3*: *Suppose $f\colon [0,1] \to \mathbb{C}$ is a uniform limit of a sequence of polynomials of degree at most d, then the limit is a polynomial of degree at most d. Conclude that to approximate a function which is not a polynomial, we need the degree of the approximations to go to infinity.*
Hint: First prove that if a sequence of polynomials of degree d converges uniformly to the zero function, then the coefficients converge to zero. One way to do this is linear algebra: Consider a polynomial p evaluated at $d+1$ points to be a linear operator taking the coefficients of p to the values of p (an operator in $L(\mathbb{R}^{d+1})$).

Exercise* 11.7.4*: *Suppose $f\colon [0,1] \to \mathbb{R}$ is continuous and $\int_0^1 f(x)x^n\,dx = 0$ for all $n = 0,1,2,\ldots$. Show that $f(x) = 0$ for all $x \in [0,1]$. Hint: Approximate by polynomials to show that $\int_0^1 \bigl(f(x)\bigr)^2\,dx = 0$.*

Exercise* 11.7.5*: *Suppose $I\colon C\bigl([0,1],\mathbb{R}\bigr) \to \mathbb{R}$ is a linear continuous function such that $I(x^n) = \frac{1}{n+1}$ for all $n = 0,1,2,3,\ldots$. Prove that $I(f) = \int_0^1 f$ for all $f \in C\bigl([0,1],\mathbb{R}\bigr)$.*

Exercise* 11.7.6*: *Let $\mathcal{A}$ be the collection of real polynomials in x^2, that is, polynomials of the form $c_0 + c_1x^2 + c_2x^4 + \cdots + c_dx^{2d}$.*

a) Show that every $f \in C\bigl([0,1],\mathbb{R}\bigr)$ is a uniform limit of polynomials from $\mathcal{A}$.

b) Find an $f \in C\bigl([-1,1],\mathbb{R}\bigr)$ that is not a uniform limit of polynomials from $\mathcal{A}$.

c) Which hypothesis of the real Stone–Weierstrass is not satisfied for the domain $[-1,1]$?

***Exercise* 11.7.7:** *Let $|z| = 1$ define the unit circle $S^1 \subset \mathbb{C}$.*

a) *Show that functions of the form*

$$\sum_{k=-n}^{n} c_k \, z^k$$

are dense in $C(S^1, \mathbb{C})$. Notice the negative powers.

b) *Show that functions of the form*

$$c_0 + \sum_{k=1}^{n} c_k \, z^k + \sum_{k=1}^{n} c_{-k} \, \bar{z}^k$$

are dense in $C(S^1, \mathbb{C})$. These are so-called harmonic polynomials, and this approximation leads to, for example, the solution of the steady state heat problem.

Hint: A good way to write the equation for S^1 is $z\bar{z} = 1$.

***Exercise* 11.7.8:** *Show that for complex numbers c_j, the set of functions of x on $[-\pi, \pi]$ of the form*

$$\sum_{k=-n}^{n} c_k \, e^{ikx}$$

satisfies the hypotheses of the complex Stone–Weierstrass theorem and therefore such functions are dense in the $C\big([-\pi, \pi], \mathbb{C}\big)$.

***Exercise* 11.7.9:** *Let $S^1 \subset \mathbb{C}$ be the unit circle, that is, the set where $|z| = 1$. Orient this set counterclockwise. Let $\gamma(t) := e^{it}$. For the one-form $f(z)\,dz$ we write**

$$\int_{S^1} f(z)\,dz := \int_0^{2\pi} f(e^{it})\, ie^{it}\, dt.$$

a) *Prove that for all nonnegative integers $k = 0, 1, 2, 3, \ldots$, we have $\int_{S^1} z^k \, dz = 0$.*

b) *Prove that if $P(z) = \sum_{k=0}^{n} c_k z^k$ is a polynomial in z, then $\int_{S^1} P(z)\,dz = 0$.*

c) *Prove $\int_{S^1} \bar{z}\, dz \neq 0$.*

d) *Conclude that polynomials in z (this algebra of functions is not self-adjoint) are not dense in $C(S^1, \mathbb{C})$.*

***Exercise* 11.7.10:** *Let (X, d) be a compact metric space and suppose $\mathcal{A} \subset C(X, \mathbb{R})$ is a real algebra that separates points, but vanishes at exactly one point $x_0 \in X$. That is, $f(x_0) = 0$ for all $f \in \mathcal{A}$, but for every $y \in X \setminus \{x_0\}$ there is a $\varphi \in \mathcal{A}$ such that $\varphi(y) \neq 0$. Prove that every function $g \in C(X, \mathbb{R})$ such that $g(x_0) = 0$ is a uniform limit of functions from $\mathcal{A}$.*

***Exercise* 11.7.11:** *Let (X, d) be a compact metric space and suppose $\mathcal{A} \subset C(X, \mathbb{R})$ is a real algebra. Suppose that for each $y \in X$ the closure $\overline{\mathcal{A}}$ contains the function $\varphi_y(x) := d(y, x)$. Then $\overline{\mathcal{A}} = C(X, \mathbb{R})$.*

---

*Alternatively, one could define $dz := dx + i\,dy$ and extend the path integral from chapter 9 to complex-valued one-forms.

Exercise* 11.7.12*:

a) Suppose $f: [a,b] \to \mathbb{C}$ is continuously differentiable. Show that there exists a sequence of polynomials $\{p_n\}_{n=1}^{\infty}$ that converges in the C^1 norm to f, that is, $\|f - p_n\|_{[a,b]} + \|f' - p_n'\|_{[a,b]} \to 0$ as $n \to \infty$.

b) Suppose $f: [a,b] \to \mathbb{C}$ is k times continuously differentiable. Show that there exists a sequence of polynomials $\{p_n\}_{n=1}^{\infty}$ that converges in the C^k norm to f, that is,

$$\sum_{j=0}^{k} \left\| f^{(j)} - p_n^{(j)} \right\|_{[a,b]} \to 0 \qquad \text{as} \qquad n \to \infty.$$

Exercise* 11.7.13*:

a) Show that an even function $f: [-1,1] \to \mathbb{R}$ is a uniform limit of polynomials with even powers only, that is, polynomials of the form $a_0 + a_1 x^2 + a_2 x^4 + \cdots + a_k x^{2k}$.

b) Show that an odd function $f: [-1,1] \to \mathbb{R}$ is a uniform limit of polynomials with odd powers only, that is, polynomials of the form $b_1 x + b_2 x^3 + b_3 x^5 + \cdots + b_k x^{2k-1}$.

Exercise* 11.7.14*: *Let $f: [a,b] \to \mathbb{R}$ be continuous.*

a) Given two points $x_1, x_2 \in [a,b]$, show that there exists a sequence of real polynomials $\{p_n\}_{n=1}^{\infty}$ so that $p_n(x_1) = f(x_1)$ and $p_n(x_2) = f(x_2)$ for all n.

b) Generalize the previous part to k points: Given the points $x_1, x_2, \ldots, x_k \in [a,b]$, show that there exists a sequence of real polynomials $\{p_n\}_{n=1}^{\infty}$ so that for all n, $p_n(x_j) = f(x_j)$ for $j = 1, 2, \ldots, k$.
Hint: The polynomial $(x - x_1)(x - x_2) \cdots (x - x_{\ell-1})(x - x_{\ell+1}) \cdots (x - x_k)$ is zero at x_j for $j \neq \ell$ but nonzero at x_ℓ. Use it to construct a polynomial that takes prescribed values at $x_1, x_2, \ldots, x_k$.

11.8 Fourier series

Note: 3–4 lectures

Fourier series* is perhaps the most important (and the most difficult) of the series that we cover in this book. We saw a few examples already, but let us start at the beginning.

11.8.1 Trigonometric polynomials

A *trigonometric polynomial* is an expression of the form

$$a_0 + \sum_{n=1}^{N} \bigl(a_n \cos(nx) + b_n \sin(nx)\bigr),$$

or equivalently, thanks to Euler's formula ($e^{i\theta} = \cos(\theta) + i\sin(\theta)$):

$$\sum_{n=-N}^{N} c_n e^{inx}.$$

The second form is usually more convenient. If $z \in \mathbb{C}$ with $|z| = 1$, we write $z = e^{ix}$, and so

$$\sum_{n=-N}^{N} c_n e^{inx} = \sum_{n=-N}^{N} c_n z^n.$$

So a trigonometric polynomial is really a rational function of the complex variable z (we are allowing negative powers) evaluated on the unit circle. There is a wonderful connection between power series (actually Laurent series because of the negative powers) and Fourier series because of this observation, but we will not investigate this further.

Another reason why Fourier series is important and comes up in so many applications is that the functions e^{inx} are eigenfunctions† of various differential operators. For example,

$$\frac{d}{dx}\bigl[e^{inx}\bigr] = (in)e^{inx}, \qquad \frac{d^2}{dx^2}\bigl[e^{inx}\bigr] = (-n^2)e^{inx}.$$

That is, they are the functions whose derivative is a scalar (the eigenvalue) times itself. Just as eigenvalues and eigenvectors are important in studying matrices, eigenvalues and eigenfunctions are important when studying linear differential equations.

The functions $\cos(nx)$, $\sin(nx)$, and e^{inx} are 2π-periodic and hence trigonometric polynomials are also 2π-periodic. We could rescale x to make the period different, but the theory is the same, so we stick with the period 2π. The antiderivative of e^{inx} is $\frac{e^{inx}}{in}$ and so

$$\int_{-\pi}^{\pi} e^{inx}\,dx = \begin{cases} 2\pi & \text{if } n = 0, \\ 0 & \text{otherwise.} \end{cases}$$

*Named after the French mathematician Jean-Baptiste Joseph Fourier (1768–1830).

†Eigenfunction is like an eigenvector for a matrix, but for a linear operator on a vector space of functions.

Consider

$$f(x) := \sum_{n=-N}^{N} c_n e^{inx},$$

and for $m = -N, \dots, N$ compute

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-imx}\,dx = \frac{1}{2\pi}\int_{-\pi}^{\pi}\left(\sum_{n=-N}^{N} c_n e^{i(n-m)x}\right)dx = \sum_{n=-N}^{N} c_n \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{i(n-m)x}\,dx = c_m.$$

We just found a way of computing the coefficients c_m using an integral of f. If $|m| > N$, the integral is 0, so we might as well have included enough zero coefficients to make $|m| \le N$.

Proposition 11.8.1. *A trigonometric polynomial $f(x) = \sum_{n=-N}^{N} c_n\, e^{inx}$ is real-valued for real x if and only if $c_{-m} = \overline{c_m}$ for all $m = -N, \dots, N$.*

Proof. If $f(x)$ is real-valued, that is, $\overline{f(x)} = f(x)$, then

$$\overline{c_m} = \overline{\frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-imx}\,dx} = \frac{1}{2\pi}\int_{-\pi}^{\pi} \overline{f(x)e^{-imx}}\,dx = \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{imx}\,dx = c_{-m}.$$

The complex conjugate goes inside the integral because the integral is done on real and imaginary parts separately.

On the other hand, if $c_{-m} = \overline{c_m}$, then

$$\overline{c_{-m}\, e^{-imx} + c_m\, e^{imx}} = \overline{c_{-m}}\, e^{imx} + \overline{c_m}\, e^{-imx} = c_m\, e^{imx} + c_{-m}\, e^{-imx},$$

which is real valued. Also $c_0 = \overline{c_0}$, so c_0 is real. By pairing up the terms, we obtain that f has to be real-valued. □

The functions e^{inx} are also linearly independent.

Proposition 11.8.2. *If*

$$\sum_{n=-N}^{N} c_n\, e^{inx} = 0$$

for all $x \in [-\pi, \pi]$, then $c_n = 0$ for all n.

Proof. The result follows immediately from the integral formula for c_n. □

11.8.2 Fourier series

We now take limits. The series

$$\sum_{n=-\infty}^{\infty} c_n\, e^{inx}$$

is called the *Fourier series* and the numbers c_n the *Fourier coefficients*. Using Euler's formula $e^{i\theta} = \cos(\theta) + i\sin(\theta)$, we could also develop everything with sines and cosines, that is, as the series $a_0 + \sum_{n=1}^{\infty} a_n \cos(nx) + b_n \sin(nx)$. It is equivalent, but slightly more messy.

Several questions arise. What functions are expressible as Fourier series? Obviously, they have to be 2π-periodic, but not every periodic function is expressible with the series. Furthermore, if we do have a Fourier series, where does it converge (where and if at all)? Does it converge absolutely? Uniformly? Also note that the series has two limits. When talking about Fourier series convergence, we often talk about the following limit:

$$\lim_{N\to\infty} \sum_{n=-N}^{N} c_n e^{inx}.$$

There are other ways we can sum the series to get convergence in more situations, but we refrain from discussing those. In light of this, define the *symmetric partial sums*

$$s_N(f;x) := \sum_{n=-N}^{N} c_n\, e^{inx}.$$

Conversely, for an integrable function $f\colon [-\pi,\pi] \to \mathbb{C}$, call the numbers

$$c_n := \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx}\, dx$$

its *Fourier coefficients*. To emphasize the function the coefficients belong to, we write $\hat{f}(n)$.*
We then formally write down a Fourier series:

$$f(x) \sim \sum_{n=-\infty}^{\infty} c_n\, e^{inx}.$$

As you might imagine such a series might not even converge. The $\sim$ doesn't imply anything about the two sides being equal in any way. It is simply that we created a formal series using the formula for the coefficients. We will see that when the functions are "nice enough," we do get convergence.

Example 11.8.3: Consider the step function $h(x)$ so that $h(x) := 1$ on $[0,\pi]$ and $h(x) := -1$ on $(-\pi,0)$, extended periodically to a 2π-periodic function. With a little bit of calculus, we compute the coefficients:

$$\hat{h}(0) = \frac{1}{2\pi} \int_{-\pi}^{\pi} h(x)\, dx = 0, \qquad \hat{h}(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} h(x) e^{-inx}\, dx = \frac{i\big((-1)^n - 1\big)}{\pi n} \quad \text{for } n \geq 1.$$

A little bit of simplification leads to

$$s_N(h;x) = \sum_{n=-N}^{N} \hat{h}(n)\, e^{inx} = \sum_{n=1}^{N} \frac{2\big(1-(-1)^n\big)}{\pi n} \sin(nx).$$

See the left hand graph in Figure 11.11 for a graph of h and several symmetric partial sums.

*The notation should seem similar to Fourier transform to those readers that have seen it. The similarity is not just coincidental, we are taking a type of Fourier transform here.

For a second example, consider the function $g(x) := |x|$ on $[-\pi,\pi]$ and then extended to a 2π-periodic function. Computing the coefficients, we find

$$\hat{g}(0) = \frac{1}{2\pi}\int_{-\pi}^{\pi} g(x)\,dx = \frac{\pi}{2}, \qquad \hat{g}(n) = \frac{1}{2\pi}\int_{-\pi}^{\pi} g(x)e^{-inx}\,dx = \frac{(-1)^n-1}{\pi n^2} \quad \text{for } n \geq 1.$$

A little simplification yields

$$s_N(g;x) = \sum_{n=-N}^{N} \hat{g}(n)\,e^{inx} = \frac{\pi}{2} + \sum_{n=1}^{N} \frac{2\big((-1)^n-1\big)}{\pi n^2}\cos(nx).$$

See the right hand graph in Figure 11.11.

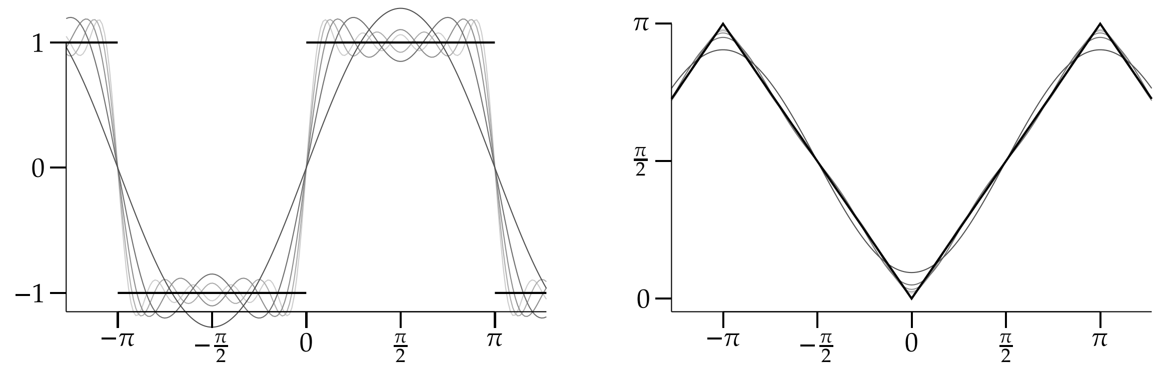

Figure 11.11: The functions h and g in bold, with several symmetric partial sums in gray.

Note that for both f and g, the even coefficients (except $\hat{g}(0)$) happen to vanish, but that is not really important. What is important is convergence. First, at the discontinuity at $x = 0$, we find $s_N(h;0) = 0$ for all N, so $s_N(h;0)$ converges to a different number from $h(0)$ (at a nice enough jump discontinuity, the limit is the average of the two-sided limits, see the exercises). That should not be surprising; the coefficients are computed by an integral, and integration does not notice if the value of a function changes at a single point. We should remark, however, that we are not guaranteed that in general the Fourier series converges to the function even at a point where the function is continuous. We will prove convergence if the function is at least Lipschitz.

What is really important is how fast the coefficients go to zero. For the discontinuous h, the coefficients $\hat{h}(n)$ go to zero approximately like $1/n$. On the other hand, for the continuous g, the coefficients $\hat{g}(n)$ go to zero approximately like $1/n^2$. The Fourier coefficients "see" the discontinuity in some sense.

Do note that continuity in this setting is the continuity of the periodic extension, that is, we include the endpoints $\pm\pi$. So the function $f(x) = x$ defined on $(-\pi,\pi]$ and extended periodically would be discontinuous at the endpoints $\pm\pi$.

In general, the relationship between regularity of the function and the rate of decay of the coefficients is somewhat more complicated than the example above might make it seem, but there are some quick conclusions we can make. We forget about finding a series for a function for a moment, and we consider simply the limit of some given series. A few sections ago, we proved that the Fourier series

$$\sum_{n=1}^{\infty} \frac{\sin(nx)}{n^2}$$

converges uniformly and hence converges to a continuous function. This example and its proof can be extended to a more general criterion.

Proposition 11.8.4. *Let $\sum_{n=-\infty}^{\infty} c_n \, e^{inx}$ be a Fourier series, and C, $\alpha > 1$ constants such that*

$$|c_n| \leq \frac{C}{|n|^{\alpha}} \qquad \textit{for all } n \in \mathbb{Z} \setminus \{0\}.$$

Then the series converges (absolutely and uniformly) to a continuous function on $\mathbb{R}$.

The proof is to apply the Weierstrass M-test (Theorem 11.2.4) and the p-series test to find that the series converges uniformly and hence to a continuous function (Corollary 11.2.8). We can also take derivatives.

Proposition 11.8.5. *Let $\sum_{n=-\infty}^{\infty} c_n \, e^{inx}$ be a Fourier series, and C, $\alpha > 2$ constants such that*

$$|c_n| \leq \frac{C}{|n|^{\alpha}} \qquad \textit{for all } n \in \mathbb{Z} \setminus \{0\}.$$

Then the series converges to a continuously differentiable function on $\mathbb{R}$.

The proof is to note that the series converges to a continuous function by the previous proposition. In particular, it converges at some point. Then differentiate the partial sums

$$\sum_{n=-N}^{N} in c_n \, e^{inx}$$

and notice that for all nonzero n

$$|in c_n| \leq \frac{C}{|n|^{\alpha-1}}.$$

The differentiated series converges uniformly by the M-test again. Since the differentiated series converges uniformly, we find that the original series $\sum_{n=-\infty}^{\infty} c_n \, e^{inx}$ converges to a continuously differentiable function, whose derivative is the differentiated series (see Theorem 11.2.14).

We can iterate this reasoning. Suppose there is some C and $\alpha > k+1$ ($k \in \mathbb{N}$) such that for all nonzero integers n,

$$|c_n| \leq \frac{C}{|n|^{\alpha}}.$$

Then the Fourier series converges to a k-times continuously differentiable function. Therefore, the faster the coefficients go to zero, the more regular the limit is.

11.8.3 Orthonormal systems

Let us abstract away the exponentials, and study a more general series for a function. One fundamental property of the exponentials that makes Fourier series work is that the exponentials are a so-called *orthonormal system*. Fix an interval $[a, b]$. We define an *inner product* for the space of functions. We restrict our attention to Riemann integrable functions as we do not have the Lebesgue integral, which would be the natural choice. Let f and g be complex-valued Riemann integrable functions on $[a, b]$ and define the inner product

$$\langle f, g \rangle := \int_a^b f(x)\overline{g(x)}\,dx.$$

If you have seen Hermitian inner products in linear algebra, this is precisely such a product. We must include the conjugate as we are working with complex numbers. We then have the "size" of f, that is, the L^2 norm $\|f\|_2$, by (defining the square)

$$\|f\|_2^2 := \langle f, f \rangle = \int_a^b |f(x)|^2\,dx.$$

Remark 11.8.6. Note the similarity to finite dimensions. For $z = (z_1, z_2, \ldots, z_d) \in \mathbb{C}^d$, one defines

$$\langle z, w \rangle := \sum_{n=1}^{d} z_n \overline{w_n}.$$

Then the norm is (usually denoted simply by $\|z\|$ in $\mathbb{C}^d$ rather than by $\|z\|_2$)

$$\|z\|^2 = \langle z, z \rangle = \sum_{n=1}^{d} |z_n|^2.$$

This is just the euclidean distance to the origin in $\mathbb{C}^d$ (same as $\mathbb{R}^{2d}$).

In what follows, we will assume all functions are Riemann integrable.

Definition 11.8.7. Let $\{\varphi_n\}_{n=1}^{\infty}$ be a sequence of integrable complex-valued functions on $[a, b]$. We say that this is an *orthonormal system* if

$$\langle \varphi_n, \varphi_m \rangle = \int_a^b \varphi_n(x)\,\overline{\varphi_m(x)}\,dx = \begin{cases} 1 & \text{if } n = m, \\ 0 & \text{otherwise.} \end{cases}$$

In particular, $\|\varphi_n\|_2 = 1$ for all n. If we only require that $\langle \varphi_n, \varphi_m \rangle = 0$ for $m \neq n$, then the system would be called an *orthogonal system*.

We noticed above that

$$\left\{ \frac{1}{\sqrt{2\pi}}\, e^{inx} \right\}_{n=1}^{\infty}$$

is an orthonormal system on $[-\pi, \pi]$. The factor out in front is to make the norm be 1.

Having an orthonormal system $\{\varphi_n\}_{n=1}^{\infty}$ on $[a,b]$ and an integrable function f on $[a,b]$, we can write a Fourier series relative to $\{\varphi_n\}_{n=1}^{\infty}$. Let

$$c_n := \langle f, \varphi_n \rangle = \int_a^b f(x)\overline{\varphi_n(x)}\, dx,$$

and write

$$f(x) \sim \sum_{n=1}^{\infty} c_n \varphi_n.$$

In other words, the series is

$$\sum_{n=1}^{\infty} \langle f, \varphi_n \rangle \varphi_n(x).$$

Notice the similarity to the expression for the orthogonal projection of a vector onto a subspace from linear algebra. We are in fact doing just that, but in a space of functions.

Theorem 11.8.8. *Suppose f is a Riemann integrable function on $[a,b]$. Let $\{\varphi_n\}_{n=1}^{\infty}$ be an orthonormal system on $[a,b]$ and suppose*

$$f(x) \sim \sum_{n=1}^{\infty} c_n \varphi_n(x).$$

If

$$s_k(x) := \sum_{n=1}^{k} c_n \varphi_n(x) \quad \text{and} \quad p_k(x) := \sum_{n=1}^{k} d_n \varphi_n(x)$$

for some other sequence $\{d_n\}_{n=1}^{\infty}$, then

$$\int_a^b |f(x) - s_k(x)|^2\, dx = \|f - s_k\|_2^2 \leq \|f - p_k\|_2^2 = \int_a^b |f(x) - p_k(x)|^2\, dx$$

with equality only if $d_n = c_n$ for all $n = 1, 2, \ldots, k$.

In other words, the partial sums of the Fourier series are the best approximation with respect to the L^2 norm.

Proof. Let us write

$$\int_a^b |f - p_k|^2 = \int_a^b |f|^2 - \int_a^b f\overline{p_k} - \int_a^b \overline{f} p_k + \int_a^b |p_k|^2.$$

Now

$$\int_a^b f\overline{p_k} = \int_a^b f \sum_{n=1}^{k} \overline{d_n}\,\overline{\varphi_n} = \sum_{n=1}^{k} \overline{d_n} \int_a^b f\,\overline{\varphi_n} = \sum_{n=1}^{k} \overline{d_n} c_n,$$

and
$$\int_a^b |p_k|^2 = \int_a^b \sum_{n=1}^k d_n \varphi_n \sum_{m=1}^k \overline{d_m} \overline{\varphi_m} = \sum_{n=1}^k \sum_{m=1}^k d_n \overline{d_m} \int_a^b \varphi_n \overline{\varphi_m} = \sum_{n=1}^k |d_n|^2.$$
So
$$\begin{aligned} \int_a^b |f - p_k|^2 &= \int_a^b |f|^2 - \sum_{n=1}^k \overline{d_n} c_n - \sum_{n=1}^k d_n \overline{c_n} + \sum_{n=1}^k |d_n|^2 \\ &= \int_a^b |f|^2 - \sum_{n=1}^k |c_n|^2 + \sum_{n=1}^k |d_n - c_n|^2. \end{aligned}$$
This is minimized precisely when $d_n = c_n$. □

When we do plug in $d_n = c_n$, then
$$\int_a^b |f - s_k|^2 = \int_a^b |f|^2 - \sum_{n=1}^k |c_n|^2,$$
and so for all k,
$$\sum_{n=1}^k |c_n|^2 \leq \int_a^b |f|^2.$$
Note that
$$\sum_{n=1}^k |c_n|^2 = \|s_k\|_2^2$$
by the calculation above. We take a limit to obtain the so-called *Bessel's inequality*.

Theorem 11.8.9 (Bessel's inequality*)**.** *Suppose f is a Riemann integrable function on $[a, b]$. Let $\{\varphi_n\}_{n=1}^\infty$ be an orthonormal system on $[a, b]$ and suppose*
$$f(x) \sim \sum_{n=1}^\infty c_n \varphi_n(x).$$
Then
$$\sum_{n=1}^\infty |c_n|^2 \leq \int_a^b |f|^2 = \|f\|_2^2.$$

In particular, $\int_a^b |f|^2 < \infty$ implies the series converges and hence
$$\lim_{k\to\infty} c_k = 0.$$

*Named after the German astronomer, mathematician, physicist, and geodesist Friedrich Wilhelm Bessel (1784–1846).

11.8.4 The Dirichlet kernel and approximate delta functions

We return to the trigonometric Fourier series. The system $\{e^{inx}\}_{n=1}^{\infty}$ is orthogonal, but not orthonormal if we simply integrate over $[-\pi,\pi]$. We can rescale the integral and hence the inner product to make $\{e^{inx}\}_{n=1}^{\infty}$ orthonormal. That is, if we replace

$$\int_a^b \qquad \text{with} \qquad \frac{1}{2\pi}\int_{-\pi}^{\pi},$$

(we are just rescaling the dx really)*, then everything works and we obtain that the system $\{e^{inx}\}_{n=1}^{\infty}$ is orthonormal with respect to the inner product

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

Suppose $f\colon \mathbb{R}\to\mathbb{C}$ is 2π-periodic and integrable on $[-\pi,\pi]$. Write

$$f(x) \sim \sum_{n=-\infty}^{\infty} c_n\,e^{inx}, \qquad \text{where} \quad c_n \coloneqq \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x)e^{-inx}\,dx.$$

Recall the notation for the symmetric partial sums, $s_N(f;x) \coloneqq \sum_{n=-N}^{N} c_n\,e^{inx}$. The inequality leading up to Bessel now reads:

$$\frac{1}{2\pi}\int_{-\pi}^{\pi}\lvert s_N(f;x)\rvert^2\,dx = \sum_{n=-N}^{N}\lvert c_n\rvert^2 \leq \frac{1}{2\pi}\int_{-\pi}^{\pi}\lvert f(x)\rvert^2\,dx.$$

Let the *Dirichlet kernel* be

$$D_N(x) \coloneqq \sum_{n=-N}^{N} e^{inx}.$$

We claim that

$$D_N(x) = \frac{\sin\bigl((N+1/2)x\bigr)}{\sin(x/2)},$$

for x such that $\sin(x/2)\neq 0$. The left-hand side is continuous on $\mathbb{R}$, and hence the right-hand side extends continuously to all of $\mathbb{R}$. To show the claim, we use a familiar trick:

$$(e^{ix}-1)D_N(x) = e^{i(N+1)x} - e^{-iNx}.$$

Multiply by $e^{-ix/2}$

$$(e^{ix/2}-e^{-ix/2})D_N(x) = e^{i(N+1/2)x} - e^{-i(N+1/2)x}.$$

The claim follows.

*Mathematicians in this field sometimes simplify matters with a tongue-in-cheek definition that $1 = 2\pi$.

Expand the definition of s_N

$$s_N(f;x) = \sum_{n=-N}^{N} \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) e^{-int}\, dt\, e^{inx}$$

$$= \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \sum_{n=-N}^{N} e^{in(x-t)}\, dt = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) D_N(x-t)\, dt.$$

Convolution strikes again! As D_N and f are 2π-periodic, we may also change variables and write

$$s_N(f;x) = \frac{1}{2\pi} \int_{x-\pi}^{x+\pi} f(x-t) D_N(t)\, dt = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x-t) D_N(t)\, dt.$$

See Figure 11.12 for a plot of D_N for $N = 5$ and $N = 20$.

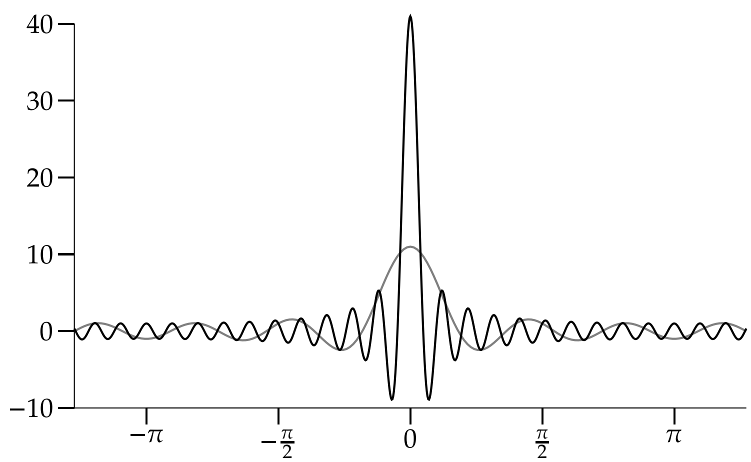

Figure 11.12: Plot of $D_N(x)$ for $N = 5$ (gray) and $N = 20$ (black).

The central peak gets taller and taller as N gets larger, and the side peaks stay small. We are convolving (again) with *approximate delta functions*, although these functions have all these oscillations away from zero. The oscillations on the side do not go away but they are eventually so fast that we expect the integral to just sort of cancel itself out there. Overall, we expect that $s_N(f)$ goes to f. Things are not always simple, but under some conditions on f, such a conclusion holds. For this reason people write

$$2\pi\, \delta(x) \sim \sum_{n=\infty}^{\infty} e^{inx},$$

where δ is the "delta function" (not really a function), which is an object that will give something like "$\int_{-\pi}^{\pi} f(x-t)\delta(t)\, dt = f(x)$." We can think of $D_N(x)$ converging in some sense to $2\pi\, \delta(x)$. However, we have not defined (and will not define) what the delta function is, nor what does it mean for it to be a limit of D_N or have a Fourier series.

11.8.5 Localization

If f satisfies a Lipschitz condition at a point, then the Fourier series converges at that point.

Theorem 11.8.10. *Let x be fixed and let f be a 2π-periodic function Riemann integrable on $[-\pi,\pi]$. Suppose there exist $\delta > 0$ and M such that*

$$|f(x+t) - f(x)| \leq M|t|$$

for all $t \in (-\delta,\delta)$, then

$$\lim_{N\to\infty} s_N(f;x) = f(x).$$

In particular, if f is continuously differentiable at x, then we obtain convergence at x (exercise). A function $f\colon [a,b] \to \mathbb{C}$ is *continuous piecewise smooth* if it is continuous and there exist points $x_0 = a < x_1 < x_2 < \cdots < x_k = b$ such that for every j, f restricted to $[x_j, x_{j+1}]$ is continuously differentiable (up to the endpoints).

Corollary 11.8.11. *Let f be a 2π-periodic function Riemann integrable on $[-\pi,\pi]$. Suppose there exist $x \in \mathbb{R}$ and $\delta > 0$ such that f is continuous piecewise smooth on $[x-\delta, x+\delta]$, then*

$$\lim_{N\to\infty} s_N(f;x) = f(x).$$

The proof of the corollary is left as an exercise. Let us prove the theorem.

Proof of Theorem 11.8.10. For all N,

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} D_N = 1.$$

Write

$$\begin{aligned} s_N(f;x) - f(x) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} f(x-t)D_N(t)\,dt - f(x)\frac{1}{2\pi}\int_{-\pi}^{\pi} D_N(t)\,dt \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \bigl(f(x-t)-f(x)\bigr)D_N(t)\,dt \\ &= \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{f(x-t)-f(x)}{\sin(t/2)}\sin\bigl((N+1/2)t\bigr)\,dt. \end{aligned}$$

By the hypotheses, for small nonzero t,

$$\left|\frac{f(x-t)-f(x)}{\sin(t/2)}\right| \leq \frac{M|t|}{|\sin(t/2)|}.$$

As $\sin(\theta) = \theta + h(\theta)$ where $\frac{h(\theta)}{\theta} \to 0$ as $\theta \to 0$, we notice that $\frac{M|t|}{|\sin(t/2)|}$ is continuous at the origin. Hence, $\frac{f(x-t)-f(x)}{\sin(t/2)}$, as a function of t, is bounded near the origin. As $t = 0$ is the only place on $[-\pi,\pi]$ where the denominator vanishes, it is the only place where there could be a problem. So, the function is bounded near $t = 0$ and clearly Riemann integrable on any interval not including 0, and thus it is Riemann integrable on $[-\pi, \pi]$. We use the trigonometric identity

$$\sin\bigl((N + {}^1\!/\!_2)t\bigr) = \cos({}^t\!/\!_2)\sin(Nt) + \sin({}^t\!/\!_2)\cos(Nt),$$

to compute

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{f(x-t)-f(x)}{\sin({}^t\!/\!_2)} \sin\bigl((N + {}^1\!/\!_2)t\bigr)\,dt = \\ \frac{1}{2\pi}\int_{-\pi}^{\pi} \left(\frac{f(x-t)-f(x)}{\sin({}^t\!/\!_2)}\cos({}^t\!/\!_2)\right)\sin(Nt)\,dt + \frac{1}{2\pi}\int_{-\pi}^{\pi}\bigl(f(x-t)-f(x)\bigr)\cos(Nt)\,dt.$$

As functions of t, $\frac{f(x-t)-f(x)}{\sin({}^t\!/\!_2)}\cos({}^t\!/\!_2)$ and $\bigl(f(x-t)-f(x)\bigr)$ are bounded Riemann integrable functions and so their Fourier coefficients go to zero by Theorem 11.8.9. So the two integrals on the right-hand side, which compute the Fourier coefficients for the real version of the Fourier series go to 0 as N goes to infinity. This is because $\sin(Nt)$ and $\cos(Nt)$ are also orthonormal systems with respect to the same inner product. Hence $s_N(f;x) - f(x)$ goes to 0, that is, $s_N(f;x)$ goes to $f(x)$. □

The theorem also says that convergence depends only on local behavior. That is, to understand convergence of $s_N(f;x)$ we only need to know f in some neighborhood of x.

Corollary 11.8.12. *Suppose f is a 2π-periodic function, Riemann integrable on $[-\pi, \pi]$. If J is an open interval and $f(x) = 0$ for all $x \in J$, then $\lim\limits_{N\to\infty} s_N(f;x) = 0$ for all $x \in J$.*

In particular, if f and g are 2π-periodic functions, Riemann integrable on $[-\pi, \pi]$, J an open interval, and $f(x) = g(x)$ for all $x \in J$, then for all $x \in J$, the sequence $\{s_N(f;x)\}_{N=1}^{\infty}$ converges if and only if $\{s_N(g;x)\}_{N=1}^{\infty}$ converges.

The first claim follows by taking $M = 0$ in the theorem. The "In particular" follows by considering $f - g$, which is zero on J and $s_N(f-g) = s_N(f) - s_N(g)$. So convergence at x depends only on the values of the function near x. However, we saw that the rate of convergence, that is, how fast does $s_N(f)$ converge to f, depends on global behavior of f.

Note a subtle difference between the results above and what Stone–Weierstrass theorem gives. Any continuous function on $[-\pi, \pi]$ can be uniformly approximated by trigonometric polynomials, but these trigonometric polynomials may not be the partial sums s_N.

11.8.6 Parseval's theorem

Finally, convergence always happens in the L^2 sense and operations on the (infinite) vectors of Fourier coefficients are the same as the operations using the integral inner product.

Theorem 11.8.13 (Parseval*). *Let f and g be 2π-periodic functions, Riemann integrable on $[-\pi, \pi]$ with*

$$f(x) \sim \sum_{n=-\infty}^{\infty} c_n\, e^{inx} \qquad \text{and} \qquad g(x) \sim \sum_{n=-\infty}^{\infty} d_n\, e^{inx}.$$

Then

$$\lim_{N\to\infty} \|f - s_N(f)\|_2^2 = \lim_{N\to\infty} \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(x) - s_N(f;x)|^2\, dx = 0.$$

Also

$$\langle f, g\rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x)\overline{g(x)}\, dx = \sum_{n=-\infty}^{\infty} c_n \overline{d_n},$$

and

$$\|f\|_2^2 = \frac{1}{2\pi} \int_{-\pi}^{\pi} |f(x)|^2\, dx = \sum_{n=-\infty}^{\infty} |c_n|^2.$$

Proof. There exists (exercise) a continuous 2π-periodic function h such that

$$\|f - h\|_2 < \epsilon.$$

Via Stone–Weierstrass, approximate h with a trigonometric polynomial uniformly. That is, there is a trigonometric polynomial $P(x)$ such that $|h(x) - P(x)| < \epsilon$ for all x. Hence

$$\|h - P\|_2 = \sqrt{\frac{1}{2\pi} \int_{-\pi}^{\pi} |h(x) - P(x)|^2\, dx} \le \epsilon.$$

If P is of degree N_0, then for all $N \ge N_0$,

$$\|h - s_N(h)\|_2 \le \|h - P\|_2 \le \epsilon,$$

as $s_N(h)$ is the best approximation for h in L^2 (Theorem 11.8.8). By the inequality leading up to Bessel,

$$\|s_N(h) - s_N(f)\|_2 = \|s_N(h - f)\|_2 \le \|h - f\|_2 \le \epsilon.$$

The L^2 norm satisfies the triangle inequality (exercise). Thus, for all $N \ge N_0$,

$$\|f - s_N(f)\|_2 \le \|f - h\|_2 + \|h - s_N(h)\|_2 + \|s_N(h) - s_N(f)\|_2 \le 3\epsilon.$$

Hence, the first claim follows.

Next,

$$\langle s_N(f), g\rangle = \frac{1}{2\pi} \int_{-\pi}^{\pi} s_N(f;x)\overline{g(x)}\, dx = \sum_{n=-N}^{N} c_n \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{inx}\overline{g(x)}\, dx = \sum_{n=-N}^{N} c_n \overline{d_n}.$$

*Named after the French mathematician Marc-Antoine Parseval (1755–1836).

We need the Schwarz (or Cauchy–Schwarz or Cauchy–Bunyakovsky–Schwarz) inequality for L^2, that is,

$$\left| \int_a^b f\bar{g} \right|^2 \leq \left(\int_a^b |f|^2 \right) \left(\int_a^b |g|^2 \right).$$

Its proof is left as an exercise; it is not much different from the finite-dimensional version. So

$$\begin{aligned} \left| \int_{-\pi}^{\pi} f\bar{g} - \int_{-\pi}^{\pi} s_N(f)\bar{g} \right| &= \left| \int_{-\pi}^{\pi} (f - s_N(f))\bar{g} \right| \\ &\leq \left(\int_{-\pi}^{\pi} |f - s_N(f)|^2 \right)^{1/2} \left(\int_{-\pi}^{\pi} |g|^2 \right)^{1/2}. \end{aligned}$$

The right-hand side goes to 0 as N goes to infinity by the first claim of the theorem. That is, as N goes to infinity, $\langle s_N(f), g \rangle$ goes to $\langle f, g \rangle$, and the second claim is proved. The last claim in the theorem follows by using $g = f$. □

11.8.7 Exercises

***Exercise* 11.8.1:** *Consider the Fourier series*

$$\sum_{k=1}^{\infty} \frac{1}{2^k} \sin(2^k x).$$

Show that the series converges uniformly and absolutely to a continuous function. Remark: This is another example of a nowhere differentiable function (you do not have to prove that). See Figure 11.13.*

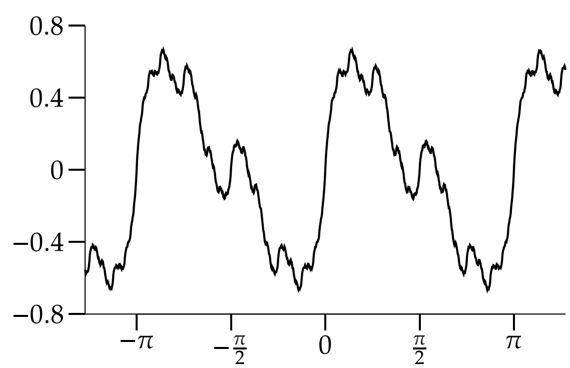

Figure 11.13: Plot of $\sum_{n=1}^{\infty} \frac{1}{2^n} \sin(2^n x)$.

*See G. H. Hardy, *Weierstrass's Non-Differentiable Function*, Transactions of the American Mathematical Society, **17**, No. 3 (Jul., 1916), pp. 301–325. A thing to notice here is the nth Fourier coefficient is $1/n$ if $n = 2^k$ and zero otherwise, so the coefficients go to zero like $1/n$.

***Exercise* 11.8.2:** *Suppose that a 2π-periodic function that is Riemann integrable on $[-\pi,\pi]$, and such that f is continuously differentiable on some open interval (a,b). Prove that for every $x \in (a,b)$, we have $\lim_{N\to\infty} s_N(f;x) = f(x)$.*

***Exercise* 11.8.3:** *Prove Corollary 11.8.11, that is, suppose a 2π-periodic function is continuous piecewise smooth near a point x, then $\lim_{N\to\infty} s_N(f;x) = f(x)$. Hint: See the previous exercise.*

***Exercise* 11.8.4:** *Given a 2π-periodic function $f\colon \mathbb{R} \to \mathbb{C}$, Riemann integrable on $[-\pi,\pi]$, and $\epsilon > 0$, show that there exists a continuous 2π-periodic function $g\colon \mathbb{R} \to \mathbb{C}$ such that $\lVert f - g\rVert_2 < \epsilon$.*

***Exercise* 11.8.5:** *Prove the Cauchy–Bunyakovsky–Schwarz inequality for Riemann integrable functions:*

$$\left| \int_a^b f\bar{g} \right|^2 \leq \left(\int_a^b |f|^2 \right) \left(\int_a^b |g|^2 \right).$$

***Exercise* 11.8.6:** *Prove the L^2 triangle inequality for Riemann integrable functions on $[-\pi,\pi]$:*

$$\lVert f+g\rVert_2 \leq \lVert f\rVert_2 + \lVert g\rVert_2.$$

***Exercise* 11.8.7:** *Suppose for some C and $\alpha > 1$, we have a real sequence $\{a_n\}_{n=1}^\infty$ with $|a_n| \leq \frac{C}{n^\alpha}$ for all n. Let*

$$g(x) := \sum_{n=1}^\infty a_n \sin(nx).$$

a) Show that g is continuous.

b) Formally (that is, suppose you can differentiate under the sum) find a solution (formal solution, that is, do not yet worry about convergence) to the differential equation

$$y'' + 2y = g(x)$$

of the form

$$y(x) = \sum_{n=1}^\infty b_n \sin(nx).$$

c) Then show that this solution y is twice continuously differentiable, and in fact solves the equation.

***Exercise* 11.8.8:** *Let f be a 2π-periodic function such that $f(x) = x$ for $0 < x < 2\pi$. Use Parseval's theorem to find*

$$\sum_{n=1}^\infty \frac{1}{n^2} = \frac{\pi^2}{6}.$$

***Exercise* 11.8.9:** *Suppose that $c_n = 0$ for all $n < 0$ and $\sum_{n=0}^\infty |c_n|$ converges. Let $\mathbb{D} := B(0,1) \subset \mathbb{C}$ be the unit disc, and $\overline{\mathbb{D}} = C(0,1)$ be the closed unit disc. Show that there exists a continuous function $f\colon \overline{\mathbb{D}} \to \mathbb{C}$ that is analytic on $\mathbb{D}$ and such that on the boundary of $\mathbb{D}$ we have $f(e^{i\theta}) = \sum_{n=0}^\infty c_n e^{in\theta}$.*
Hint: If $z = re^{i\theta}$, then $z^n = r^n e^{in\theta}$.

***Exercise* 11.8.10:** *Show that*

$$\sum_{n=1}^{\infty} e^{-1/n} \sin(nx)$$

converges to an infinitely differentiable function.

***Exercise* 11.8.11:** *Let f be a 2π-periodic function such that $f(x) = f(0) + \int_0^x g$ for a function g that is Riemann integrable on every interval. Suppose*

$$f(x) \sim \sum_{n=-\infty}^{\infty} c_n \, e^{inx}.$$

Show that there exists a $C > 0$ such that $|c_n| \leq \frac{C}{|n|}$ for all nonzero n.

***Exercise* 11.8.12:**

a) Let φ be the 2π-periodic function defined by $\varphi(x) := 0$ if $x \in (-\pi, 0)$, and $\varphi(x) := 1$ if $x \in (0, \pi)$, letting $\varphi(0)$ and $\varphi(\pi)$ be arbitrary. Show that $\lim\limits_{N\to\infty} s_N(\varphi; 0) = 1/2$.

b) Let f be a 2π-periodic function Riemann integrable on $[-\pi, \pi]$, $x \in \mathbb{R}$, $\delta > 0$, and there are continuously differentiable $g \colon [x-\delta, x] \to \mathbb{C}$ and $h \colon [x, x+\delta] \to \mathbb{C}$ where $f(t) = g(t)$ for all $t \in [x-\delta, x)$ and where $f(t) = h(t)$ for all $t \in (x, x+\delta]$. Then $\lim\limits_{N\to\infty} s_N(f;x) = \frac{g(x)+h(x)}{2}$, or in other words,

$$\lim_{N\to\infty} s_N(f;x) = \frac{1}{2}\left(\lim_{t\to x^-} f(t) + \lim_{t\to x^+} f(t)\right).$$

***Exercise* 11.8.13:** *Let $\{a_n\}_{n=1}^{\infty}$ be such that $\lim_{n\to\infty} a_n = 0$. Show that there is a continuous 2π-periodic function f whose Fourier coefficients c_n satisfy that for each N there is a $k \geq N$ where $|c_k| \geq a_k$.*
Remark: The exercise says that if f is only continuous, there is no "minimum rate of decay" of the coefficients. Compare with Exercise 11.8.11.
Hint: Look at Exercise 11.8.1 for inspiration.

Further Reading

[R1] Maxwell Rosenlicht, *Introduction to Analysis*, Dover Publications Inc., New York, 1986. Reprint of the 1968 edition.

[R2] Walter Rudin, *Principles of Mathematical Analysis*, 3rd ed., McGraw-Hill Book Co., New York, 1976. International Series in Pure and Applied Mathematics.

[T] William F. Trench, *Introduction to Real Analysis*, Pearson Education, 2003. http://ramanujan.math.trinity.edu/wtrench/texts/TRENCH_REAL_ANALYSIS.PDF.

Index

List of Notation

Notation	Description	Page
e^z	complex exponential function	163
$\sin(z)$	sine function	164
$\cos(z)$	cosine function	164
π	the number π	165
$s_N(f;x)$	symmetric partial sum of a Fourier series	194
$f(x) \sim \sum_{n=-\infty}^{\infty} c_n e^{inx}$	Fourier series for f	194
$\langle f, g \rangle$	inner product of functions	197
$\Vert f \Vert_2$	L^2 norm of f	197