Analysis and Topology Course Notes

Patrick Boily
University of Ottawa
pboily@uottawa.ca

December 2023

Preface

Personalities and Stakes

Differential calculus is a collection of algebraic tools that enable the exact resolution of certain geometrical problems posed by the ancients: determining the **length of a curve**, the **area of a geometric figure**, or the **volume of a solid**, for example, or finding the **tangent line to any geometric shape**. By the 17^{th} century, it was already possible to solve these problems.

Isaac Newton and his British contemporaries relied on velocity and rates of change in a **theory of fluxions**, whereas Leibniz and European mathematicians used infinitesimal increments, or **differentials**, mysterious quantities larger than 0, but smaller than any other number.[1] The results obtained were valid in both frameworks, but the methods used in either case were far from satisfactory. It was not necessary to understand *why* the methods were valid for them to work, but the question often recurred, scratching the collective subconscious of mathematicians, philosophers, and theologians of the time: mathematics were considered "divine" or "heavenly", so why was there so much ambiguity?[2]

Both methods used the notion of **infinity** without ever defining the concept; unfortunately, infinity has the troublesome habit of defying intuition when one least expects it.

The French mathematician d'Alembert then attempted to provide a certain formalism by introducing the notion of a **limit**,

> [...] the number which one can approach as closely as one wishes by using a sequence of secant approximations [...],

but his definition hardly proved more precise. What is meant by "approaching"? Do we ever reach this number?

It was in response to this lack of formalism that **mathematical analysis** was established. Its foundations are owed, among others, to Cauchy, Gauss, and Weierstrass; calculus is an "intuitive" version of their analysis.

---

[1] Mathematicians of the era engaged in furious academic battles over the priority of discovery; the British insisted that Newton was the inventor of differential calculus since he had used it to calculate the orbits of planets; but Leibniz published his results on the derivative of a product before Newton. Several collaborations, as well as many friendships, became casualties of the conflict.

[2] Bishop Berkeley, in a (now famous) treatise published in 1734, attacked the approaches of both camps: if velocity is the first derivative (the first fluxion) of a particle's position, what corresponds to the second and third derivative? How can a quantity be smaller than any other quantity? Are infinitesimals the ghosts of departed quantities?

References and Influences

These notes are mostly based on courses I taught at the University of Ottawa between 2004 and 2021, but also on courses that I took as a student between 1994 and 2002. I was blessed with fantastic calculus, analysis, differential systems, and topology instructors and mentors:

- **Richard Blute**, **Luc Demers**, **Marcel Déruaz**, **Benoit Dionne**, **Thierry Giordano**, **Barry Jessup**, **Victor Leblanc**, and **Rémi Vaillancourt** at the University of Ottawa, and
- **Wojciech Jaworski** and **Michael Moore** at Carleton University.

It is no exaggeration to say that I would not be a professional mathematician without their guidance, for which I thank them heartily.

More pragmatically, these notes could not exist without their influence and hard work, in particular that of B. Dionne (chapters 1-6), T. Giordano (chapters 7-14, 21-24), and M. Moore (chapters 15-20). I should also mention Aaron Smith with whom I co-taught MAT 2125 (Elementary Real Analysis) online during the COVID-19 pandemic, who contributed some material and solved problems to chapters 1-6.

I have consulted and borrowed from a whole slew of references over the years, of which the following are the most prominent:

- Bartle, R.G., Sherbert, D.R. [1992], *Introduction to Real Analysis*, 2nd edition, Wiley.
- Brown, J.W., Churchill, R.V. [1996], *Complex Variables and Applications*, 7th edition, McGraw-Hill.
- Gourdon, X. [2000], *Les maths en tête: analyse*, 2e édition, Ellipses.
- Hirsch, M.W., Smale, S. [1974], *Differential Equations, Dynamical Systems, and Linear Algebra*, Academic Press.
- Marsden, J.E., Hoffman, M.J. [1993], *Elementary Classical Analysis*, W.H. Freeman.
- Marsden, J.E., Tromba, A.J. [1988], *Vector Calculus*, W.H. Freeman.
- Munkres, J.R. [1974], *Topology: a First Course*, Prentice-Hall.
- Royden, H.L. [1968], *Real Analysis*, Macmillan.
- Rudin, W.R. [1991], *Functional Analysis*, McGraw-Hill.
- Rudin, W.R. [1987], *Real and Complex Analysis*, McGraw-Hill.
- Savage, A. [2017], *Elementary Real Analysis*, course notes (these also form the basis of Section 6.3).
- Spivak, M. [1965], *Calculus on Manifolds*, Addison-Wesley.

Be sure to give these masterful works the attention they deserve.

Pre-Requisites and Course Notes Overview

Readers are assumed to have taken three semesters of calculus, two semesters of linear algebra, and one course in mathematical reasoning and proofs at the university level (MAT 1320, MAT 1322, MAT1341, MAT 1362, MAT 2122, and MAT 2141 at the University of Ottawa), and more importantly, to have **mastered their contents**.

Each of the first four parts correspond roughly to a course offered (or previously offered) at the University of Ottawa:

- Part I: MAT 2125 (*Elementary Real Analysis*, formerly *Real Analysis I*);
- Part II: MAT 3120 (formerly *Real Analysis III*, currently *Real Analysis*);
- Part III: MAT 2121 (formerly *Real Analysis II*, not in the course catalogue anymore, except as a special topics course), and
- Part IV: MAT 4153 (*General Topology*),

whereas Part V contains tidbits that could easily be found in MAT 3121 (*Complex Analysis I*), MAT 3130 (*Introduction to Dynamical Systems*), MAT 4121 (*Measure and Integration I*), and/or MAT 4124 (*Introduction to Functional Analysis*).

Any analyst and any topologist worth their salt will have to tackle all the topics listed above in their training (and more besides, depending on their individual research interests), but there is no substitute for taking courses and learning from specific instructors.

I make these notes available mainly to help students bridge gaps caused by scheduling issues and to whet their appetites by giving them a chance to look ahead.

No matter how we swing it, there is a **lot of material to cover**, and there is no denying that some of it can be scary the first time it is encountered ... but analysis is mostly **fun** once we get the hang of it.

So roll up your sleeves, and happy learning!

Patrick Boily
Wakefield, Canada
December 2023

Contents

Part I

Elementary Real Analysis

Chapter 1

The Real Numbers

In real analysis, the fundamental object of study is the set of **real numbers**, $\mathbb{R}$. In this chapter, we introduce $\mathbb{R}$ and some of its important properties, discuss the cardinality of sets, and provide a first analytical result, whose proof will serve as an introduction to the discipline.

1.1 Hierarchy of Number Systems

At a basic level, **analysis** is a theory on the real numbers $\mathbb{R}$, that is, the objects with which we work are **real numbers**, **real sets**, and **real functions**. We will see at a later stage that we can conduct analysis on any **metric space** (such as $\mathbb{R}^n$ and $\mathbb{C}$, for instance).

There is a natural hierarchy amongst number sets, which you have no doubt encountered in your courses:

$$\mathbb{N}^\times \subsetneq \mathbb{N} \subsetneq \mathbb{Z} \subsetneq \mathbb{Q} \subsetneq \mathbb{A} \subsetneq \mathbb{R} \subsetneq \mathbb{C}.$$

The **positive integers** $\mathbb{N}^\times$ are the counting numbers; **zero** is added to $\mathbb{N}^\times$ to form $\mathbb{N}$, in which all equations $x + a = b$, $b \geq a \in \mathbb{N}^\times$ have a solution. Similarly, the **integers** $\mathbb{Z}$ are built by adding new numbers to $\mathbb{N}$ in order for all equations of the form $x + a = b$, $a, b \in \mathbb{N}$ to have solutions. For the **rational numbers** $\mathbb{Q}$, the equations in question have the form $ax + b = 0$, $a, b \in \mathbb{Z}$, $b \neq 0$. For the **algebraic numbers** $\mathbb{A}$, we are looking at equations of the form

$$a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 = 0, \quad a_i \in \mathbb{Q},$$

and for **complex numbers** $\mathbb{C}$, equations of the form

$$a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 = 0, \quad a_i \in \mathbb{R}.$$

In other words, number sets are generally easy to construct once we have the right building blocks... except when it comes to the **real numbers** $\mathbb{R}$. In this chapter and the next, we will introduce concepts that will allow us to **"formally" define** $\mathbb{R}$.

In what follows, we will make use of the following axiom about the set $\mathbb{N}$.

> **Axiom** (Well-Ordering Principle)
> *Any non-empty subset of $\mathbb{N}$ has a smallest element.*

We will define the "smallest" element of a set momentarily. We shall also discuss how to measure the "size" of a set in Section 1.2; for the moment, we remark only that while $\mathbb{Q}$ is infinite, **it contains infinitely more holes than it does elements**.

Field and Order Properties of $\mathbb{R}$

A **field** F is a set endowed with two binary operations: an **addition** $+ : F \times F \to F$, defined by $+(a, b) = a + b$, and a **multiplication** $\cdot : F \times F \to F$, defined by $\cdot(a, b) = ab$, which satisfy the 9 **field properties**:

(A1) **commutativity of $+$:** $\forall a, b \in F$, $a + b = b + a$;
(A2) **associativity of $+$:** $\forall a, b, c \in F$, $(a + b) + c = a + (b + c)$;
(A3) **existence of neutral element for $+$:** $\exists 0 \in F$, $\forall a \in F$, $a + 0 = a$;
(A4) **inverse with respect to $+$:** $\forall a \in F$, $\exists! b \in F$, $a + b = 0$;
(M1) **commutativity of $\cdot$:** $\forall a, b \in F$, $ab = ba$
(M2) **associativity of $\cdot$:** $\forall a, b, c \in F$, $(ab)c = a(bc)$
(M3) **existence of neutral element for $\cdot$:** $\exists 1 \in F$, $\forall a \in F$, $1a = a$
(M4) **inverse with respect to $\cdot$:** $\forall a \in F^{\times}$, $\exists! b \in F$, $ab = 1$
(D1) **distributivity of $\cdot$ over $+$:** $\forall a, b, c \in F$, $a(b + c) = ab + ac$

> **Examples:** $\mathbb{Q}$ is a field; $\mathbb{N}$ is not a field since (A4) is not satisfied for $x = 1 \in \mathbb{N}$, say; $\mathbb{Z}$ is not a field since (M4) is not satisfied for $x = 2$, say. □

An **order** on a set F is a binary relation "$<$" satisfying the **order properties**:

(O1) **trichotomy:** $\forall a, b, c \in F$, $a < b$ or $a = b$ or $b < a$;
(O2) **transitivity:** $\forall a, b, c \in F$, if $a < b$ and $b < c$, then $a < c$.
(O3) $\forall a, b, c \in F$, if $a < b$, then $a + c < b + c$.
(O4) **(specific to $\mathbb{R}$):** $\forall a, b, c \in \mathbb{R}$, if $a < b$ and $c > 0$, then $ac < bc$.

> **Examples**
>
> 1. The relation "is born before" is an order relation on the set of human beings (with reasonable assumptions about birth);
> 2. the relation "is smaller than" is an order relation on $\mathbb{N}, \mathbb{Z}, \mathbb{Q}$;
> 3. the relation "is a subset of" is not an order on $\wp(\mathbb{N})$ since we have neither $\{1, 2\} \not\subseteq \{1, 3\}$, $\{1, 2\} \neq \{1, 3\}$, nor $\{1, 3\} \not\subseteq \{1, 2\}$. □

Let $(F,<)$ be an ordered set and $S \subseteq F$. If $a < b$ or $a = b$, we write $a \leq b$. The element $u \in F$ is an **upper bound of** S if $s \leq u$ for all $s \in S$. In that case, we say that S is **bounded above**. If u is the smallest upper bound of S, we say that it is the **supremum** of S, denoted $u = \sup S$.

The element $v \in F$ is a **lower bound of** S if $v \leq s$ for all $s \in S$. In that case, we say that S is **bounded below**. If v is the largest lower bound of S, we say that it is the **infimum** of S, denoted $u = \inf S$. If the set S is bounded both above and below, we say that it is **bounded**.

> **Example:** if $S = \{x \in \mathbb{Q} \mid 2 < x < 3\}$, then $\inf S = 2$.
>
> **Proof:** the rational number $v = 2$ is a lower bound of S since $2 = v < x$ for all $x \in S$ (but so are $v = -1$ and $v = 1.5$). Hence $\inf S \geq 2$.
>
> To show that 2 is indeed the greatest lower bound, we suppose that $u = \inf S > 2$ and derive a contradiction. As we already know that $\inf S \geq 2$, this will only leave one possibility: $\inf S = 2$.
>
> By assumption, there exists $0 < \varepsilon < 1$ in $\mathbb{Q}$ such that $u = 2 + \varepsilon$. Find a rational number $u^* \in (2, u)$. By definition, $u^* \in S$, since $3 > u^* > 2$. But $u > u^*$, and so u cannot be a lower bound of S, which contradicts the hypothesis that $u = \inf S$. Thus $\inf S \not> 2$ and $\inf S = 2$. ■

This "proof" rests on thin ice, however: it assumes that the infimum exists in the first place; that if the infimum exists, it is a rational number, and that a rational number can be found between any two distinct rationals. These assumptions are valid **in this specific case**, but not so in general – more on this later.

> **Example:** show that if $S = \mathbb{N}$, then $\inf S = 1$.
>
> **Proof:** the integer $v = 1$ is a lower bound since $1 = v \leq n$ for all $n \in \mathbb{N}$, so $\inf \mathbb{N} \geq 1$. But any number above 1 cannot be a lower bound of $\mathbb{N}$ since it would not be smaller than 1. Thus, $\inf S = 1$. ■

Completeness of $\mathbb{R}$

A set $(F,<)$ is **complete** if every non-empty bounded subset $S \subseteq F$ has a supremum and an infimum.

> **Example:** show that $\mathbb{Q}$ is not complete.
>
> **Proof:** consider the subset $S = \{x \in \mathbb{Q}^+ \mid 2 < x^2 < 3\}$. Since $1.5 \in \mathbb{Q}^+$, then $1.5^2 = 2.25 \in \mathbb{Q}^+$. We have $2 < 1.5^2 = 2.25 < 3$, so $1.5 \in S$, and thus $S \neq \varnothing$. Furthermore, S is bounded above by 3 since $3^2 > 3$ and bounded below by 1 since $1^2 < 1$, so S is bounded.

We will see shortly that S has no supremum/infimum in $\mathbb{Q}$ (since no rational x is such that $x^2 = 2$ or $x^2 = 3$). Thus $\mathbb{Q}$ is not complete. ■

The set $\mathbb{R}$ of **real numbers** is the smallest complete ordered field containing $\mathbb{N}$, with order $a < b \Longleftrightarrow b - a > 0$.

Archimedean Property

Classically, $\mathbb{R}$ is constructed using **Dedekind cuts** or **Cauchy sequences**: in effect, $\mathbb{R}$ is constructed by "filling the holes" of $\mathbb{Q}$. We will discuss Cauchy sequences in Chapter 2 and provide the outline of $\mathbb{R}$'s construction in Chapter 7. For now, we assume that $\mathbb{R}$ is available and that is satisfies the properties mentioned previously, as well as the next "obvious" result.

Theorem 1 (ARCHIMEDEAN PROPERTY OF $\mathbb{R}$)
Let $x \in \mathbb{R}$. Then $\exists n_x \in \mathbb{N}^\times$ such that $x < n_x$.

Proof: suppose that there is no such integer. Then $x \geq n \, \forall n \in \mathbb{N}$. Consequently, x is an upper bound of $\mathbb{N}^\times$. But $\mathbb{N}^\times$ is a non-empty subset of $\mathbb{R}$. Since $\mathbb{R}$ is complete, $\alpha = \sup \mathbb{N}^\times$ exists.

By definition of the supremum (the smallest upper bound), $\alpha - 1$ is not an upper bound of $\mathbb{N}^\times$ (otherwise α would not be the smallest upper bound, as $\alpha - 1 < \alpha$ would be a smaller upper bound).

Since $\alpha - 1$ is not an upper bound of $\mathbb{N}^\times$, $\exists m \in \mathbb{N}^\times$ such that $\alpha - 1 < m$. Using the properties of $\mathbb{R}$, we must then have $\alpha < m + 1 \in \mathbb{N}^\times$; that is, α is not an upper bound of $\mathbb{N}^\times$.

This contradicts the fact that $\alpha = \sup \mathbb{N}^\times$, and so, since $\mathbb{N}^\times \neq \varnothing$, x cannot be an upper bound of $\mathbb{N}^\times$. Thus $\exists n_x \in \mathbb{N}^\times$ such that $x < n_x$. ■

The Archimedean property of $\mathbb{R}$ is a fundamental construct; it used (often implicitly) in nearly all analytical proofs.

Theorem 2 (VARIANTS OF THE ARCHIMEDEAN PROPERTY)
Let $x, y \in \mathbb{R}^+$. Then $\exists n_1, n_2, n_3 \geq 1$ such that

1. $x < n_1 y$;
2. $0 < \frac{1}{n_2} < y$, *and*
3. $n_3 - 1 \leq x < n_3$.

Proof:

1. Let $z = \frac{x}{y} > 0$. By the Archimedean property, $\exists n_1 \geq 1$ such that $z = \frac{x}{y} < n_1$. Then $x < n_1 y$.
2. If $x = 1$, then part 1 implies $\exists n_2 \geq 1$ such that $0 < 1 < n_2 y$. Then $0 < \frac{1}{n_2} < y$.
3. Let $L = \{m \in \mathbb{N}^\times : x < m\}$. By the Archimedean property, $L \neq \varnothing$. Indeed, there is at least one $n \geq 1$ such that $x < n$. By the well-ordering principle, L has a smallest element, say $m = n_3$. Then $n_3 - 1 \notin L$ (otherwise, $n_3 - 1$ would be the least element of L, which it is not) and so $n_3 - 1 \leq x < n_3$.

There are other variants, but these are the ones we will use the most. ■

Let's look at a basic result which highlights how to use the Archimedean property.

Example: show that $\inf\{\frac{1}{n} \mid n \in \mathbb{N}^\times\} = 0$.

Proof: since $0 \leq \frac{1}{n}$ for all $n \in \mathbb{N}^\times$, 0 is a lower bound of the set. Suppose that $\varepsilon > 0$ is also a lower bound. Then $\varepsilon \leq \frac{1}{n}$ for all $n \in \mathbb{N}^\times$, which means that $n \leq \frac{1}{\varepsilon}$ for all $n \in \mathbb{N}^\times$. This contradicts the Archimedean Property, so 0 is the smallest lower bound of the set. ■

It is thus always possible to find an integer greater than any specified real number. This result is extremely useful – we use it next to show the existence of **irrational numbers**.

Corollary
The positive root of $x^2 = 2$ lies in $\mathbb{R}$ but not in $\mathbb{Q}$.

Proof: we first show that any solution of $x^2 = 2$ cannot be rational. Suppose the equation $x^2 = 2$ has a rational positive root $r = p/q$, with $\gcd(p, q) = 1$. Then $p^2/q^2 = 2$, or $p^2 = 2q^2$. Hence p^2 is even, and so p is also even. Indeed, if $p = 2k + 1$ is odd, then so is $p^2 = 2(2k^2 + 2k) + 1$.

Set $p = 2m$. Then $(2m)^2 = 2q^2$, or $2m^2 = q^2$. Thus q^2 and q are even. Consequently, both p and q are even, which contradicts the hypothesis $\gcd(p, q) = 1$. The equation $r^2 = 2$ cannot then have a solution in $\mathbb{Q}$. But we have not yet shown that the equation has a solution in $\mathbb{R}$.

Consider the set $S = \{s \in \mathbb{R}^+ : s^2 < 2\}$, where $\mathbb{R}^+$ denotes the set of positive real numbers. This set in not empty as $1 \in S$. Furthermore, S is bounded above by 2. Indeed, if $t \geq 2$, then $t^2 \geq 4 > 2$, whence $t \notin S$.

By completeness of $\mathbb{R}$, $u = \sup S \geq 1$ exists. It is enough to show that neither $u^2 < 2$ and $u^2 > 2$ can hold. The only remaining possibility is that $u^2 = 2$.

- If $u^2 < 2$, then $\frac{2u+1}{2-u^2} > 0$. By the Archimedean property, $\exists n > 0$ such that $\frac{2u+1}{2-u^2} < n$. By re-arranging the terms, we get

  $$0 < \frac{1}{n}(2u+1) < 2 - u^2.$$

  Then

  $$\begin{aligned}\left(u + \frac{1}{n}\right)^2 = u^2 + \frac{2u}{n} + \frac{1}{n^2} \quad &\leq \quad u^2 + \frac{2u}{n} + \frac{1}{n} \\ &= \quad u^2 + \frac{1}{n}(2u+1) < u^2 + 2 - u^2 = 2.\end{aligned}$$

  Since $(u + \frac{1}{n})^2 < 2$, $u + \frac{1}{n} \in S$. But $u < u + \frac{1}{n}$; u is then not an upper bound of S, which contradicts the fact that $u = \sup S$. Thus $u^2 \not< 2$.

- If $u^2 > 3$, then $\frac{2u}{u^2-2} > 0$. By the Archimedean property, $\exists n > 0$ such that $\frac{2u}{u^2-3} < n$. By re-arranging the terms, we get

  $$0 > -\frac{2u}{n} > 2 - u^2.$$

  Then

  $$\left(u - \frac{1}{n}\right)^2 = u^2 - \frac{2u}{n} + \frac{1}{n^2} > u^2 - \frac{2u}{n} > u^2 + 2 - u^2 = 2.$$

  Since $(u - \frac{1}{n})^2 > 2$, $u - \frac{1}{n}$ is an upper bound of S. But $u > u - \frac{1}{n}$; u can not then be the supremum of S, which is a contradiction. Thus $u^2 \not> 2$.

That leaves only one alternative (since $u \in \mathbb{R}$): $u^2 = 2$, and $u = \sqrt{2} \in \mathbb{R}$. ■

From this point on, when we mention the Archimedean Property, we mean one of the four variants from Theorems 1 and 2.

Absolute Value and Useful Inequalities

The real numbers enjoy another collection of useful and interesting properties.

Theorem 3 (Bernoulli's Inequality)
Let $x \geq -1$. Then $(1+x)^n \geq 1 + nx$, $\forall n \in \mathbb{N}$.

Proof: we prove the result by induction on n.

- If $n=1$, then $(1+x)^1 = 1+x \geq 1+1x$.
- Suppose that the result is true for $n=k$, that is $(1+x)^k \geq 1+kx$. We have to show that it is also true for $n=k+1$. But

$$(1+x)^{k+1} = (1+x)^k(1+x) \underbrace{\geq (1+kx)(1+x)}_{\boxed{\text{Ind. Hyp.}}} = 1+(k+1)x+kx^2 \geq 1+(k+1)x,$$

which completes the proof. ■

The assumption $x \geq -1$ is essential – if $1+x<0$, the use of the induction hypothesis in the string of inequalities cannot be justified (it would, in fact, be invalid).

Theorem 4 (Cauchy's Inequality)
If $a_1,\ldots,a_n$ and $b_1,\ldots,b_n$ are real numbers, then

$$\left(\sum_{i=1}^n a_ib_i\right)^2 \leq \left(\sum_{i=1}^n a_i^2\right)\left(\sum_{i=1}^n b_i^2\right).$$

Furthermore, if $b_j \neq 0$ for one of $1 \leq j \leq n$, then equality holds if and only if $\exists s \in \mathbb{R}$ such that $a_i = sb_i$ for all $i=1,\ldots,n$.

Proof: for any $t \in \mathbb{R}$,

$$0 \leq \sum_{i=1}^n (a_i+tb_i)^2 = \sum_{i=1}^n a_i^2 + 2t\sum_{i=1}^n a_ib_i + t^2\sum_{i=1}^n b_i^2.$$

The right-hand side of this inequality is a polynomial of degree 2 in t. As it is non-negative, it has at most 1 real root. Thus, its discriminant

$$\left(2\sum_{i=1}^n a_ib_i\right)^2 - 4\left(\sum_{i=1}^n a_i^2\right)\left(\sum_{i=1}^n b_i^2\right) \leq 0,$$

and so

$$\left(\sum_{i=1}^n a_ib_i\right)^2 \leq \left(\sum_{i=1}^n a_i^2\right)\left(\sum_{i=1}^n b_i^2\right).$$

If all the b_i are 0, the equality holds trivially, as both the left and right side of the Cauchy inequality are 0. So suppose $b_i \neq 0$ for at least one of the values j between 1 and n.

If $a_i = sb_i$ for all $i = 1, \ldots, n$ and $s \in \mathbb{R}$ is fixed then

$$\begin{aligned}\left(\sum_{i=1}^n a_i b_i\right)^2 = \left(\sum_{i=1}^n s b_i^2\right)^2 = s^2\left(\sum_{i=1}^n b_i^2\right)^2 = s^2\left(\sum_{i=1}^n b_i^2\right)\left(\sum_{i=1}^n b_i^2\right) \\ = \left(\sum_{i=1}^n s^2 b_i^2\right)\left(\sum_{i=1}^n b_i^2\right) = \left(\sum_{i=1}^n a_i^2\right)\left(\sum_{i=1}^n b_i^2\right).\end{aligned}$$

On the other hand, if

$$\left(\sum_{i=1}^n a_i b_i\right)^2 = \left(\sum_{i=1}^n a_i^2\right)\left(\sum_{i=1}^n b_i^2\right) \quad \text{then} \quad 4\left(\sum_{i=1}^n a_i b_i\right)^2 - 4\left(\sum_{i=1}^n a_i^2\right)\left(\sum_{i=1}^n b_i^2\right) = 0.$$

But the left-hand side of this expression is the discriminant of the following polynomial of degree 2 in t:

$$\sum_{i=1}^n (a_i + tb_i)^2 = \sum_{i=1}^n a_i^2 + 2t\sum_{i=1}^n a_i b_i + t^2 \sum_{i=1}^n b_i^2.$$

Since the discriminant is 0, the polynomial has a unique root, say $t = -s$, therefore

$$\sum_{i=1}^n (a_i - sb_i)^2 = 0.$$

Since $(a_i - sb_i)^2 \geq 0$ for all $i = 1, \ldots, n$, then

$$(a_i - sb_i)^2 = 0 \Longrightarrow a_i - sb_i = 0 \Longrightarrow a_i = sb_i \quad \text{for all } i = 1, \ldots, n,$$

which completes the proof. ■

The next result is used extensively in analytical arguments.

Theorem 5 (TRIANGLE INEQUALITY)
If $a_1, \ldots, a_n, b_1, \ldots, b_n \in \mathbb{R}$,

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

Furthermore, if $b_j \neq 0$ *for one of* $1 \leq j \leq n$*, then equality holds if and only if* $\exists s \in \mathbb{R}$ *such that* $a_i = sb_i$ *for all* $i = 1, \ldots, n$.

Proof: taking the square root on both sides of the inequality below yields the desired result:

$$\begin{aligned}\sum_{i=1}^{n}(a_i+b_i)^2 &= \sum_{i=1}^{n}a_i^2+2\sum_{i=1}^{n}a_ib_i+\sum_{i=1}^{n}b_i^2\\ \boxed{\text{Cauchy Inequality}} &\leq \sum_{i=1}^{n}a_i^2+2\left(\sum_{i=1}^{n}a_i^2\right)^{1/2}\left(\sum_{i=1}^{n}b_i^2\right)^{1/2}+\sum_{i=1}^{n}b_i^2\\ &= \left(\left(\sum_{i=1}^{n}a_i^2\right)^{1/2}+\left(\sum_{i=1}^{n}b_i^2\right)^{1/2}\right)^2.\end{aligned}$$

If all the b_i are 0, the equality holds trivially, as both the left and right side of the Triangle Inequality are $\left(\sum_{i=1}^{n}a_i^2\right)^{1/2}$. So suppose $b_i\neq 0$ for at least one of the values j between 1 and n.

If $a_i=sb_i$ for all $i=1,\ldots,n$ and $s\in\mathbb{R}$ is fixed, then equality holds since

$$\begin{aligned}\left(\sum_{i=1}^{n}(a_i+b_i)^2\right)^{1/2} &= \left(\sum_{i=1}^{n}(sb_i+b_i)^2\right)^{1/2}=\left(\sum_{i=1}^{n}(s+1)^2b_i^2\right)^{1/2}\\ &= \left((s+1)^2\sum_{i=1}^{n}b_i^2\right)^{1/2}=(s+1)\left(\sum_{i=1}^{n}b_i^2\right)^{1/2}\quad\text{and}\end{aligned}$$

$$\begin{aligned}\left(\sum_{i=1}^{n}a_i^2\right)^{1/2}+\left(\sum_{i=1}^{n}b_i^2\right)^{1/2} &= \left(\sum_{i=1}^{n}s^2b_i^2\right)^{1/2}+\left(\sum_{i=1}^{n}b_i^2\right)^{1/2}\\ &= s\left(\sum_{i=1}^{n}b_i^2\right)^{1/2}+\left(\sum_{i=1}^{n}b_i^2\right)^{1/2}=(s+1)\left(\sum_{i=1}^{n}b_i^2\right)^{1/2}.\end{aligned}$$

Conversely, if

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

then

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

Developing both sides of this expression yields

$$\sum_{i=1}^{n}a_i^2+2\sum_{i=1}^{n}a_ib_i+\sum_{i=1}^{n}b_i^2=\sum_{i=1}^{n}a_i^2+2\left(\sum_{i=1}^{n}a_i^2\right)^{1/2}\left(\sum_{i=1}^{n}b_i^2\right)^{1/2}+\sum_{i=1}^{n}b_i^2,$$

or simply

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

Elevating both sides to the second power yields

$$\left(\sum_{i=1}^{n} a_i b_i\right)^2 = \left(\sum_{i=1}^{n} a_i^2\right) \left(\sum_{i=1}^{n} b_i^2\right).$$

By Cauchy's Inequality, $\exists s \in \mathbb{R}$ such that $a_i = sb_i$ for all $i = 1, \ldots, n$. ■

In the triangle inequality, if we set $n = 1$, we obtain the very useful inequality:

$$\sqrt{(a+b)^2} \leq \sqrt{a^2} + \sqrt{b^2},$$

which we usually write as

$$|a+b| \leq |a| + |b|, \quad \text{for all } a, b \in \mathbb{R}.$$

The function $|\cdot| : \mathbb{R} \to \mathbb{R}$ is the **absolute value**, which represents the distance between a real number and the origin. It is defined by

$$|x| = \begin{cases} x, & x \geq 0 \\ x, & x \leq 0 \end{cases}$$

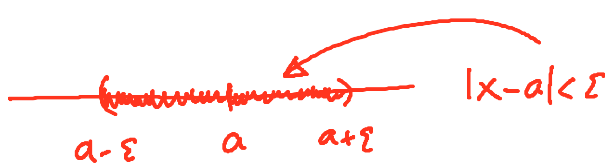
Equipped with this function, $\mathbb{R}$ is an example of a **normed space**. Normed space will be discussed in Chapter 8.

Theorem 6 (Properties of the Absolute Value)
If $x, y \in \mathbb{R}$, then

1. $|x| = \sqrt{x^2}$
2. $-|x| \leq x \leq |x|$
3. $|xy| = |x||y|$
4. $|x+y| \leq |x| + |y|$
5. $|x-y| \leq |x| + |y|$
6. $||x| - |y|| \leq |x-y|$

Remark: the following inequality will play a central role in the chapters to come:

$$|x-a| < \varepsilon \iff a - \varepsilon < x < a + \varepsilon.$$



Density of $\mathbb{Q}$

We finish this section with an intriguing result about the distribution of rationals and irrationals among the reals.

Theorem 7 (DENSITY OF $\mathbb{Q}$)
Let $x, y \in \mathbb{R}$ be such that $x < y$. Then, $\exists r \in \mathbb{Q}$ such that $x < r < y$.

Proof: there are three distinct cases.

1. If $x < 0 < y$, then select $r = 0$.
2. If $0 \leq x < y$, then $y - x > 0$ and $\frac{1}{y-x} > 0$. By the Archimedean property, $\exists n \geq 1$ such that
$$n > \frac{1}{y-x} > 0.$$
By that same property, $\exists m \geq 1$ such that $m - 1 \leq nx < m$. Since $n(y-x) > 1$, then $ny - 1 > nx$ and $nx \geq m - 1$. By the transitivity of the order $<$ on $\mathbb{R}$, we have $ny - 1 > m - 1$, and so $ny > m$. But $m > nx$, so $ny > m > nx$ and $y > \frac{m}{n} > x$. Select $r = \frac{m}{n}$.
3. If $x < y \leq 0$, then $y - x > 0$ and $\frac{1}{y-x} > 0$. By the Archimedean property, $\exists n \geq 1$ such that
$$n > \frac{1}{y-x} > 0.$$
Note that $-nx > 0$. By that same property, $\exists m \geq 0$ such that $m < -nx \leq m+1$ or $-m-1 \leq nx < -m$. Since $n(y-x) > 1$, then $ny - 1 > nx \geq -m - 1$, that is $ny > -m$. But $-m > nx$, so $ny > -m > nx$ and $y > -\frac{m}{n} > x$. Select $r = -\frac{m}{n}$. ■

Theorem 7 has a twin: the set of irrational numbers is also dense in $\mathbb{R}$.

Corollary (DENSITY OF $\mathbb{R} \setminus \mathbb{Q}$)
Let $x, y \in \mathbb{R}$ with $x < y$. Then, $\exists z \notin \mathbb{Q}$ such that $x < z < y$.

Proof: we will prove the case $x, y > 0$, the other cases are left as an exercise. According to Theorem 7, $\exists r \neq 0 \in \mathbb{Q}$ such that $\frac{x}{\sqrt{2}} < r < \frac{y}{\sqrt{2}}$.

Hence $x < r\sqrt{2} < y$. Set $z = r\sqrt{2}$. Then $z \notin \mathbb{Q}$ – indeed, if $z = r\sqrt{2} = \frac{p}{q} \in \mathbb{Q}$, then $\sqrt{2} = \frac{p}{qr} \in \mathbb{Q}$, a contradiction. ■

It is thus possible to find rationals and irrationals between any two real numbers $x < y$. In spite of this, however, $\mathbb{Q}$ is in fact much "smaller" than $\mathbb{R} \setminus \mathbb{Q}$, as we shall presently see.

1.2 Cardinality of Sets

For all $n \in \mathbb{N}^\times$, define $\mathbb{N}_n = \{1, 2, \ldots, n\}$. A set S is **finite** if $S = \varnothing$ or if there exists a bijection $f : \mathbb{N}_n \to S$ for some $n \in \mathbb{N}^\times$. If S is not finite, it is **infinite**. If S is infinite and there exists a bijection $f : \mathbb{N} \to S$, then S is **countable** and we write $|S| = \omega$. Otherwise, it is **uncountable**.[1]

Consider two sets S_n and T_n, both with n distinct elements:

$$S_n = \{s_1, \ldots, s_n\}, \quad T_n = \{t_1, \ldots, t_n\}.$$

These two finite sets have the same size: there is a bijection $f : S_n \to T_n$, $f(s_i) = t_i$ for $1 \le i \le n$ (it is not the only such bijection).

In general, two sets S, T are said to have the same **cardinality**, denoted $|S| = |T|$, if there exists a bijection $f : S \to T$. If S, T are finite, $|S| = |T|$ means that the two sets **have the same number of elements**: $|S| = |T| = |\mathbb{N}_n| = n$ for some $n \in \mathbb{N}$. If S, T are infinite, the "number of elements" is not a well-defined, which can lead to counter-intuitive results.

Examples

1. The set $2\mathbb{N} = \{2, 4, \ldots\}$ is countable because $f : \mathbb{N} \to 2\mathbb{N}$, with $f(n) = 2n$, is a bijection. We would then write $|\mathbb{N}| = |2\mathbb{N}| = \omega$.

2. The set $\mathbb{Z} = \{\ldots, -2, -1, 0, 1, 2, \ldots\}$ is countable since $f : \mathbb{Z} \to \mathbb{N}$ with

$$f(z) = \begin{cases} 2z, & z \ge 0 \\ -2z - 1, & z < 0 \end{cases}$$

is a bijection. Thus $|\mathbb{Z}| = |\mathbb{N}| = \omega$. $\square$

So two sets can have equal cardinality even when one is strictly contained in the other – but this can only happen with infinite sets, however.

Theorem 8
If S is an infinite subset of a countable set A, then S is countable.

Proof: as A is countable, we can list all its elements: $A = \{a_1, a_2, \ldots, \}$. Let $n_1, n_2, \ldots$ be integers obtained by the following algorithm:

- Set $K_1 = \{n \in \mathbb{N} \mid a_n \in S\}$. According to the well-ordering principle, $\exists n_1 \in K_1$ which is minimal. Then $a_{n_1} \in S$ and $a_m \notin S$ for all $m < n_1$.

[1] Finite sets may also be called **finitely countable** sets, and countable sets, **infinitely countable** sets.

- Set $K_2 = K_1 \setminus K_1$. According to the WOP, $\exists n_2 \in K_2$ which is minimal, with $n_1 < n_2$. Then $a_{n_2} \in S$ and $a_m \notin S$ for all $m < n_1$ with $m \neq n_1$; etc.

Repeating this process, we obtain the set $S' = \{a_{n_1}, a_{n_2}, \ldots\}$. But every element of S must be in S' (why?), so $S = S'$. The function $f : \mathbb{N} \to S$ defined by $k \mapsto a_{n_k}$ is thus a bijection, and so S is countable. ■

General Remark: when a proof is difficult to follow, it is never a bad idea to try the reasoning it with specific examples satisfying the hypotheses. If we have to provide a proof, remember that an example only works if we are trying to show that some statement is **false**. A direct proof **never** uses examples.

The contrapositive of Theorem 8 gives a useful way to show that a set is uncountable: if $S \subseteq A$ is uncountable, then A is uncountable.

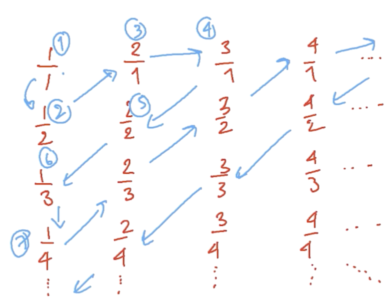
Cardinality of $\mathbb{Q}$

Another way to think of countable sets is that they could be **enumerated**, at least conceptually, in an infinite list.

Theorem 9
The set $\mathbb{Q}$ is countable.

Proof: Write $\mathbb{Q} = \mathbb{Q}^- \cup \{0\} \cup \mathbb{Q}^+$, with the obvious notation. As there is a bijection $f : \mathbb{Q}^+ \to \mathbb{Q}^-$, with $f(r) = -r$, we must have $|\mathbb{Q}^+| = |\mathbb{Q}^-|$. It is then sufficient to show that $|\mathbb{Q}^+| = \omega$.

Indeed, if we can enumerate the elements of $\mathbb{Q}^+$, then we can enumerate the elements of $\mathbb{Q}$ by starting with 0, and alternating from $\mathbb{Q}^-$ to $\mathbb{Q}^+$. But note that every positive rational takes the form $\frac{m}{n}$, with $m, n \in \mathbb{N}^\times$. We can thus arrange all such fractions in an infinite array:



There is a bijection between $\mathbb{N}^\times$ and the set $F = \{\frac{1}{1}, \frac{1}{2}, \frac{2}{1}, \frac{3}{1}, \frac{2}{2}, \ldots\}$, so $|F| = \omega$. But $\mathbb{Q}^+ \subseteq F$, so $\mathbb{Q}^+$ is countable since it is infinite (indeed, $\mathbb{N}^\times \subseteq \mathbb{Q}^+$). According to Theorem 8, we must have $|\mathbb{Q}^+| = \omega$. This completes the proof. ■

Cardinality of $\mathbb{R}$

We now show that a set which would seem to be much smaller than $\mathbb{Q}$ at a first glance is in fact much larger than $\mathbb{Q}$ from a cardinality perspective, using a celebrated argument.

Theorem 10 (Cantor's Diagonal Argument)
The set $I = [0, 1]$ is uncountable.

Proof: every number $x \in I$ has a (not necessarily unique) decimal representation of the form

$$x = 0.a_1a_2a_3\cdots, \quad a_i \in \{0, \dots, 9\}.$$

By convention, we write $1 = .0.99999\overline{9}$ and $0 = 0.00000\overline{0}$. When numbers have two decimal representations, such as $0.4000\overline{0} = 0.3999\overline{9}$, we only consider the representation with a tail of repeating 9s.

Assume that I is countable. Then it is possible to enumerate its elements:

$$I = \{x_1, x_2, \dots\}.$$

Each of the $x_i \in I$ has a unique decimal representation (with the convention given earlier):

$$\begin{aligned} x_1 &= 0.a_{1,1}a_{1,2}a_{1,3}\cdots a_{1,n}\cdots \\ x_2 &= 0.a_{2,1}a_{2,2}a_{2,3}\cdots a_{2,n}\cdots \\ &\vdots \\ x_n &= 0.a_{n,1}a_{n,2}a_{n,3}\cdots a_{n,n}\cdots \\ &\vdots \end{aligned}$$

where $a_{i,j} \in \{0, \dots, 9\}$ for all $i, j \in \mathbb{N}^\times$. Define the real number $y = 0.y_1y_2y_3\cdots$, where

$$y_i = \begin{cases} 2 & \text{if } a_{i,i} \geq 5 \\ 6 & \text{if } a_{i,i} \leq 4 \end{cases} \qquad \text{for } i \in \mathbb{N}^\times.$$

As $0 \leq y \leq 1$, we have $y \in I$. But for all $i \in \mathbb{N}^\times$, we also have $y \neq x_i$ in the list because $y_i \neq a_{i,i}$. Thus $y \notin I$, a contradiction. Consequently, the assumption that I is countable is not valid. ■

Since $[0, 1] \subseteq \mathbb{R}$, then $\mathbb{R}$ is also uncountable. What about $\mathbb{R} \setminus \mathbb{Q}$? In general, is it possible for the union of two countable sets to be uncountable? Is the intersection of two uncountable sets uncountable?

1.3 Nested Intervals Theorem

We end this initial chapter with an important result concerning nested intervals, which we will use shortly. In style and rigour, its proof is representative of analytical reasoning.

Theorem 11 (Nested Intervals)
For every integer $n \geq 1$, let $[a_n, b_n] = I_n$ be such that

$$I_1 \supseteq I_2 \supseteq \cdots I_n \supseteq I_{n+1} \supseteq \cdots$$

Then there exists $\psi, \eta \in \mathbb{R}$ such that $\psi \leq \eta$ and $\bigcap_{n\geq 1} I_n = [\psi, \eta]$. Furthermore, if $\inf\{b_n - a_n \mid n \in \mathbb{N}\} = 0$*, then $\psi = \eta$.*

Proof: since $I_n \subseteq I_1$ for all $n \geq 1$, the set $S = \{a_1, \ldots, a_n\}$ is bounded above by b_1. But $S \neq \varnothing$, so $\psi = \sup S$ exists by completeness of $\mathbb{R}$, and thus

$$a_n \leq \psi, \quad \text{for all } n \geq 1.$$

Fix $n \geq 1$ and let $k \geq 1$ be an integer:

- if $k \geq n$, then $I_n \supseteq I_k$ and $a_k \leq b_k \leq b_n$;
- if $k < n$, then $I_n \subseteq I_k$ and $a_k \leq a_n \leq b_n$.

In both cases, $a_k \leq b_n$ for all $k \geq 1$. Thus b_n is an upper bound of S for all $n \geq 1$. As $\psi = \sup S$, $\psi \leq b_n$ for all $n \geq 1$. Combining these results, we have $a_n \leq \psi \leq b_n$, for all $n \geq 1$.

Since $I_n \subseteq I_1$ for all $n \geq 1$, the set $T = \{b_1, \ldots, b_n\}$ is bounded below by a_1. But $T \neq \varnothing$, so $\eta = \inf T$ exists by completeness of $\mathbb{R}$, and thus

$$b_n \geq \eta, \quad \text{for all } n \geq 1.$$

Fix $n \geq 1$ and let $k \geq 1$ be an integer:

- if $k \geq n$, then $I_n \supseteq I_k$ and $a_n \leq a_k \leq b_k$;
- if $k < n$, then $I_n \subseteq I_k$ and $a_n \leq b_n \leq b_k$.

In both cases, $a_n \leq b_k$ for all $k \geq 1$. Thus a_n is an lower bound of T for all $n \geq 1$. As $\eta = \inf T$, $\eta \geq a_n$ for all $n \geq 1$. Combining these results, we have $a_n \leq \eta \leq b_n$, for all $n \geq 1$.

Since $\psi \leq b_n$ for all $n \geq 1$, ψ is a lower bound of T. As η is the largest such lower bound, $\psi \leq \eta$, which is to say: $a_n \leq \psi \leq \eta \leq b_n$, for all $n \geq 1$, and so $[\psi, \eta] \subseteq I_n$ for all $n \geq 1$.

Consequently,

$$[\psi,\eta]\subseteq\bigcap_{n\geq 1} I_n.$$

Now, suppose that $\gamma\in I_n$ for all $n\geq 1$. Then $a_n\leq\gamma\leq b_n$ for all $n\geq 1$, and so γ is an upper bound of S and a lower bound of T.

But ψ is the smallest upper bound of S, so $\psi=\sup S\leq\gamma$, and η is the largest lower bound of T, so $\gamma\leq\inf T\leq\eta$, and so $\gamma\in[\psi,\eta]$. Thus

$$\bigcap_{n\geq 1} I_n\subseteq[\psi,\eta]\Longrightarrow\bigcap_{n\geq 1} I_n=[\psi,\eta].$$

Finally, suppose that $\inf\{b_n-a_n\mid n\geq 1\}=0$. Let $\varepsilon>0$. By definition, $\exists k\geq 1$ such that $0\leq b_k-a_k<\varepsilon$, otherwise $\varepsilon>0$ would be a lower bound of the set, which would contradict the assumption that 0 is the largest such upper bound.

We have seen that $b_k\geq\eta$ and that $a_k\leq\psi$, so

$$\varepsilon>b_k-a_k\geq\eta-\psi\geq 0.$$

Thus, for all $\varepsilon>0$, we have $0\leq\eta-\psi<\varepsilon$, which is to say $\eta-\psi=0$. ■

Proof note: from this point on, we will avoid repeating nearly identical proof segments, using generic statements like "Similarly, we can show that $a_n\leq\inf\{b_i\mid i\geq 1\}\leq b_n$, for all $n\geq 1$" while leaving the details to be worked out by the reader.

Why can we conclude that $\eta-\psi=0$ if $0\leq\eta-\psi<\varepsilon$ for all $\varepsilon>0$? In general, if $a\leq x<a+\varepsilon$ for all $\varepsilon>0$, then $x=a$. Indeed, if $x\neq a$, $\exists\delta>0$ such that $x=a+\delta$. Thus, if $\varepsilon=\delta$ (which is possible since ε can take on any positive value) we would have $\delta=x-a<\varepsilon=\delta$, a contradiction.

Example: if $I_n=[1-\frac{1}{n},1+\frac{1}{n}]$ for $n\geq 1$, then the conditions of Theorem 11 are satisfied, and so $\bigcap_{n\geq 1} I_n=[\psi,\eta]$. As $\inf\{b_n-a_n\mid n\geq 1\}=\inf\{\frac{2}{n}\mid n\geq 1\}=0$, we have:

$$\psi=\sup\{1-\tfrac{1}{n}\}=1=\inf\{1+\tfrac{1}{n}\}=\eta,\Longrightarrow[\psi,\eta]=\{1\},$$

which concludes the example. □

Warning: we can only use a theorem if the hypotheses are satisfied (even though the conclusion may hold nonetheless). The intervals $I_n=(1-\frac{1}{n},1+\frac{1}{n})$, $n\geq 1$ are such that their intersection is $\{1\}$, but not because of the Theorem 11.

1.4 Solved Problems

1. Let $a, b \in \mathbb{R}$ and suppose that $a \leq b + \varepsilon$ for all $\varepsilon > 0$. Show that $a \leq b$.

   Proof: suppose that $a > b$. Let $\varepsilon_0 = \frac{a-b}{2} > 0$. Then

   $$a > b \Longrightarrow a + a > a + b \text{ (by O3)} \Longrightarrow a = \frac{a+a}{2} > \frac{a+b}{2} = b + \varepsilon_0 \text{ (by O4)}.$$

   Hence, $a > b + \varepsilon_0$, which contradicts the hypothesis that $a \leq b + \varepsilon$ for all $\varepsilon > 0$. Consequently, the assumption $a > b$ is false, that is, $a \not> b$ or $a \leq b$ by trichotomy of the order on $\mathbb{R}$. ■

2. Let $c > 0$ be a real number.

   a) If $c > 1$, show that $c^n \geq c$ for all $n \in \mathbb{N}$ and that $c^n > 1$ if $n > 1$.

   b) If $0 < c < 1$, show that $c^n \leq c$ for all $n \in \mathbb{N}$ and that $c^n < 1$ if $n > 1$.

   Proof: the statements are clearly not true if $n = 0$: as a result, we must interpret $\mathbb{N}$ to stand for the set $\mathbb{N} = \{1, 2, 3, \ldots\}$, without the 0. Generally, we use whatever "version" of $\mathbb{N}$ is appropriate.

   a) If $c > 1$, $\exists x \in \mathbb{R}$ such that $x > 0$ and $c = 1 + x$. Let $n \in \mathbb{N}$. First note that $n - 1 \geq 0$ and so $(n-1)x > 0$.

      Then, by Bernoulli's inequality,

      $$c^n = (1+x)^n \geq 1 + nx = 1 + x + (n-1)x \geq 1 + x = c.$$

      Furthermore, $n - 1 > 0$ and $(n-1)x > 0$ if $n > 1$. Consequently, the last inequality above is strict and so $c^n > c > 1$, which implies $c^n > 1$ (by transitivity of the order $>$).

   b) If $0 < c < 1$, there exists $b > 1$ such that $c = \frac{1}{b}$. Indeed, $\frac{1}{c}$ is such that $c \cdot \frac{1}{c} = 1$. As $c > 0$, then $\frac{1}{c} > 0$ since the product $c \cdot \frac{1}{c} = 1$ is positive.

      But $c < 1$, so that $1 = c \cdot \frac{1}{c} < \frac{1}{c}$.

      In particular, if we let $b = \frac{1}{c}$, then $b > 1$ and so we can apply part (a) of this question to get $b^n \geq b$ for all $n \in \mathbb{N}$ and $b^n > 1$ if $n > 1$.

      Let $n \in \mathbb{N}$. Then

      $$\frac{1}{c^n} = b^n \geq b = \frac{1}{c}$$

      so that $c \geq c^n$ and

      $$\frac{1}{c^n} = b^n > 1$$

      so that $1 > c^n$ if $n > 1$. ■

3. Let $c > 0$ be a real number.

   a) If $c > 1$ and $m, n \in \mathbb{N}$, show that $c^m > c^n$ if and only if $m > n$.
   b) If $0 < c < 1$ and $m, n \in \mathbb{N}$, show that $c^m > c^n$ if and only if $m < n$.

   Proof:

   a) It is sufficient to show that if $m \geq n$, then $c^m \geq c^n$. If $m = n$, the result is clear, so we assume $m > n$. In that case, $\exists k \geq 1$ such that $m = n + k$. An easy induction exercise shows that $c^{n+k} = c^n c^k$ for for all integers n and k.

      In particular, using the previous problem,

      $$c^m = c^{n+k} = c^n c^k \geq c^n \cdot c > c^n \cdot 1 = c^n$$

      and so $c^m > c^n$.
   b) This can be shown from a) using the technique from the previous question. ■

4. Let $S = \{x \in \mathbb{R} \mid x > 0\}$. Does S have lower bounds? Does S have upper bounds? Does $\inf S$ exist? Does $\sup S$ exist? Prove your statements.

   Does S have lower bounds? Yes.
   By definition, any negative real number is a lower bound (so is 0).

   Does S have upper bounds? No.
   Assume that it does. By the completeness of $\mathbb{R}$, $\alpha = \sup \mathbb{R}$ exists. In particular, $\alpha \geq n$ for all $n \in \mathbb{N}$, which contradicts the Archimedean Property of $\mathbb{R}$. Hence S has no upper bound.

   Does $\inf S$ exist? Yes.
   Consider the set $-S = \{x \in \mathbb{R} \mid -x \in S\} = \{x \in \mathbb{R} \mid x < 0\}$. By construction, 0 is an upper bound of $-S$. Note furthermore that neither S nor $-S$ are empty.

   By completeness of $\mathbb{R}$, $\sup(-S)$ exists. Right? The definition of completeness we use is that any non-empty bounded subset of $\mathbb{R}$ has a supremum. But $-S$ is only bounded above, not below. How can we conclude that $\sup(-S)$ exists?

   That definition is one particular version of the Completeness Property of $\mathbb{R}$. An **equivalent** way of stating it is: *The ordered set F is* ***complete*** *if for any $\varnothing \neq S \subset F$, S has a supremum in F whenever S is bounded above and an infimum in F whenever S is bounded below.*

   But $\sup(-S) = -\inf S$. Indeed, let $u = \sup(-S)$. Then $u \geq -x$ for all $-x \in -S$ and if v is another upper bound of $-S$ then $u \leq v$. Note that if v is an upper bound of $-S$, then $v \geq -x$ for all $-x \in -S$, **i.e.** $-v \leq x$ for all $x \in S$: as a result, $-v$ is a lower bound of S.

   Similarly, if $-v$ is a lower bound of S, v is automatically an upper bound of $-S$. Then any lower bound of S is of the form $-v$, where v is an upper bound of $-S$.

Then, $-u \leq x$ for all $x \in S$ and $-v \leq -u$ whenever $-v$ is a lower bound of S. Hence $-u = \inf S$ and so $u = -\inf S$.

As $\sup(-S) = -\inf S$ exists, so does $\inf S$.

Does $\sup S$ **exist?** No.

See second item. ■

5. Let $S = \left\{1 - \frac{(-1)^n}{n} \mid n \in \mathbb{N}\right\}$. Find $\inf S$ and $\sup S$.

Proof: the first few elements of S are:

$$2, \frac{1}{2}, \frac{4}{3}, \frac{3}{4}, \frac{6}{5}, \frac{5}{6}, \cdots .$$

This suggests that S is bounded above by 2 and below by $\frac{1}{2}$. To show that this is indeed the case, note that $(-1)^n$ only takes on the values -1 and 1, whatever the integer n.

Technically, this also has to be shown. One proceeds by induction.

The **base case** is clear: when $n = 1$, $(-1)^1 = -1 \in \{1, -1\}$.

Now, on to the **induction step**: suppose $(-1)^k \in \{1, -1\}$. Then

$$(-1)^{k+1} = (-1)^k(-1) = \begin{cases} 1(-1) = -1 \\ (-1)(-1) = 1 \end{cases} .$$

Hence $(-1)^{k+1} \in \{1, -1\}$.

By induction, $(-1)^n \in \{-1, 1\}$ for all $n \in \mathbb{N}$.

Thus $-1 \leq (-1)^n \leq 1$ for all $n \geq 1$. (In practice, we need only show it once and refer to the result if we need it in the future.)

For any $n \geq 2$, we then have $-n \leq -1 \leq (-1)^n$ and $\frac{n}{2} \geq 1 \geq (-1)^n$, that is

$$-n \leq (-1)^n \leq \frac{n}{2}.$$

A quick check shows the inequalities also hold for $n = 1$. Then, for $n \geq 1$, we have

$$\begin{aligned} -n \leq (-1)^n &\leq \frac{n}{2} \\ \therefore -1 \leq \frac{(-1)^n}{n} &\leq \frac{1}{2} \\ \therefore 1 \geq -\frac{(-1)^n}{n} &\geq -\frac{1}{2} \\ \therefore 2 \geq 1 - \frac{(-1)^n}{n} &\geq \frac{1}{2}. \end{aligned}$$

Hence $2 \geq s \geq \frac{1}{2}$ for all $s \in S$, *i.e.* 2 is an upper bound and $\frac{1}{2}$ is a lower bound of S.

By completeness, $S \subseteq \mathbb{R}$ has a supremum and an infimum in $\mathbb{R}$. If $u = \sup S < 2$, there is a contradiction as $u \not\geq s$ for all $s \in S$ (it "misses" the element 2 in S).

Thus, $\sup S \geq 2$. But 2 is already an upper bound so $\sup S \leq 2$. Consequently $\sup S = 2$. Similarly, $\inf S = \frac{1}{2}$. ■

6. Let $S \subseteq \mathbb{R}$ be non-empty. Show that if $u = \sup S$ exists, then for every number $n \in \mathbb{N}$ the number $u - \frac{1}{n}$ is not an upper bound of S, but the number $u + \frac{1}{n}$ is.

   Proof: let $n \geq 1$. Then $\frac{1}{n} > 0$ and $u < u + \frac{1}{n}$. Since $s \leq u$ for all $s \in S$, $s < u + \frac{1}{n}$ for all $s \in S$ by transitivity of $<$. Consequently, $u + \frac{1}{n}$ is an upper bound of S.

   Furthermore, $u - \frac{1}{n} < u$. Since u is the least upper bound, $u - \frac{1}{n}$ cannot be an upper bound (as it would then be lesser upper bound than u, a contradiction). This completes the proof. Or does it?

   We haven't used the hypothesis $S \neq \varnothing$. Where does it fit? Does it even fit? The definition of an upper bound implies that every real number is an upper bound of the empty set. Indeed, if $v \in \mathbb{R}$, then $v \geq s$ for all $s \in \varnothing$ automatically as there is **no** $s \in \varnothing$.

   The proof rests on the fact that $u = \sup S$. But $\sup \varnothing$ does not exist, as discussed. ■

7. If $S = \left\{ \frac{1}{n} - \frac{1}{m} \mid m, n \in \mathbb{N} \right\}$, find $\inf S$ and $\sup S$.

   Proof: the set $S = \left\{ \frac{1}{n} - \frac{1}{m} \mid n, m \in \mathbb{N} \right\}$ is bounded above by 1 and below by -1 since

   $$\frac{1}{n} \leq 1 \leq 1 + \frac{1}{m} \quad \text{and} \quad \frac{1}{m} \leq 1 \leq 1 + \frac{1}{n} \implies -1 \leq \frac{1}{n} - \frac{1}{m} \leq 1, \quad \forall m, n \in \mathbb{N}.$$

   Note that S is not empty as $0 = \frac{1}{2} - \frac{1}{2}$ is in S, say.

   By completeness, S has a supremum and an infimum. By definition, $s^* = \sup S \leq 1$. Suppose that $s^* < 1$. Then $\exists \varepsilon > 0$ such that $s^* = 1 - \varepsilon$. Furthermore,

   $$\frac{1}{n} - \frac{1}{m} \leq 1 - \varepsilon, \quad \forall m, n \in \mathbb{N}.$$

   In particular, if $n = 1$, then

   $$1 - \frac{1}{m} \leq 1 - \varepsilon, \quad \forall m \in \mathbb{N}.$$

   Equivalently, $\varepsilon \leq \frac{1}{m}$ for all integers m so that $\frac{1}{\varepsilon}$ is an upper bound for $\mathbb{N}$. This contradicts the Archimedean Property of $\mathbb{R}$. Hence $s^* \not< 1$ and so $s^* = 1$.

   To prove that $\inf S = -1$, proceed along the same lines (inf $\leftrightsquigarrow$ sup, etc.). ■

8. Let X be a non-empty set and let $f : X \to \mathbb{R}$ have bounded range in $\mathbb{R}$. If $a \in \mathbb{R}$, show that

$$\sup\{a + f(x) : x \in X\} = a + \sup\{f(x) : x \in X\}$$
$$\inf\{a + f(x) : x \in X\} = a + \inf\{f(x) : x \in X\}.$$

Proof: let $f(X) = \{f(x) \mid x \in X\}$. By hypothesis, $f(X)$ is bounded and not empty and so has a supremum in $\mathbb{R}$, say u^*. We need to show $\sup\{a+f(x); x \in X\} = a+u^*$.

To do so, first note that $a + u^*$ is an upper bound of $\sup\{a + f(x) \mid x \in X\}$ since $u^* \geq f(x)$ for all $x \in X$; as a result $a + u^* \geq a + f(x)$ for all $x \in X$ (we know that $\sup\{a + f(x) \mid x \in X\}$ indeed has a supremum by completeness of $\mathbb{R}$).

Next, we need to show that $a+u^*$ is the smallest upper bound of $\{a+f(x) \mid x \in X\}$. Suppose v is another upper bound of $\{a + f(x) \mid x \in X\}$. Then $v \geq a + f(x)$ for all $x \in X$; in particular, $v - a$ is an upper bound of $f(X)$.

By hypothesis, $v - a \geq u^*$, hence $v \geq a + u^*$. Consequently, $a + u^*$ is the least upper bound of $\{a + f(x) \mid x \in X\}$, *i.e.*

$$\sup\{a + f(x) \mid x \in X\} = a + u^*.$$

The proof for the other equality proceeds in a similar manner. ■

9. Let A and B be bounded non-empty subsets of $\mathbb{R}$, and let

$$A + B = \{a + b \mid a \in A, b \in B\}.$$

Prove that $\sup(A + B) = \sup A + \sup B$ and $\inf(A + B) = \inf A + \inf B$.

Proof: A and B are bounded and non-empty. By completeness, they have infimums (in $\mathbb{R}$), say a_* and b_*, respectively. Then $a_* \leq a$ and $b_* \leq b$ for all $a \in A$, $b \in B$.

The real number $a_* + b_*$ is a lower bound of $A + B$ since $a_* + b_* \leq a + b$ for all $a \in A$, $b \in B$. By completeness of $\mathbb{R}$, $A + B$ has an infimum as it is also not empty. We show that this infimum is indeed $a_* + b_*$.

Let w be a lower bound of $A + B$. Then, $w \leq a + b$ for all $a \in A$ and $b \in B$, or $w - b \leq a$ for all $a \in A$ and $b \in B$.

Thus, $w - b$ is a lower bound of A for all $b \in B$, *i.e.* $w - b \leq a_*$ for all $b \in B \implies w - a_* \leq b$ for all $b \in B$, so $w - a_*$ is a lower bound of B.

Hence $w - a_* \leq b_*$. As a result, $w \leq a_* + b_*$, which concludes the proof. The other equality is shown in the same manner. ■

10. Let X be a non-empty set and let $f, g : X \to \mathbb{R}$ have bounded range in $\mathbb{R}$. Show that

$$\sup\{f(x) + g(x) \mid x \in X\} \leq \sup\{f(x) \mid x \in X\} + \sup\{g(x) \mid x \in X\}$$
$$\inf\{f(x) \mid x \in X\} + \inf\{g(x) \mid x \in X\} \leq \inf\{f(x) + g(x) \mid x \in X\}.$$

Proof: let $f(X) = \{f(x) \mid x \in X\}$ and $g(X) = \{g(x) \mid x \in X\}$. By hypothesis, $f(X)$ and $g(X)$ are both bounded and not empty, so they each have a supremum in $\mathbb{R}$, say u^* and v^* respectively.

Since $f(x) \leq u^*$ and $g(x) \leq v^*$ for all $x \in X$, then $f(x) + g(x) \leq u^* + v^*$ for all $x \in X$. Hence, $\{f(x) + g(x) \mid x \in X\}$ has a supremum in $\mathbb{R}$, as it is a bounded non-empty subset of $\mathbb{R}$. Let w^* be that supremum, *i.e.* the smallest upper bound of $\{f(x) + g(x) \mid x \in X\}$.

Since $u^* + v^*$ is also an upper bound of that set, it's automatically larger than w^*.

Note that we can not say more: it is **not** true, in general, that $w^* = u^* + v^*$. Indeed, take $X = [1, 2]$ and let f and g be defined by

$$f(x) = \frac{1}{x} \quad \text{and} \quad g(x) = -\frac{1}{x}, \quad \forall x \in X.$$

Then $f(X) = \{\frac{1}{x} \mid x \in X\}$, $g(X) = \{-\frac{1}{x} \mid x \in X\}$ and $u^* = 1$, $v^* = -\frac{1}{2}$ and $w^* = 0$ (you should show these results!), and $w^* \leq u^* + v^*$ but $w^* \neq u^* + v^*$.[2]

The other inequality is tackled in a similar manner. ■

11. Let X and Y be non-empty sets and let $h : X \times Y \to \mathbb{R}$ have bounded range in $\mathbb{R}$. Let $F : X \to \mathbb{R}$ and $G : Y \to \mathbb{R}$ be defined by

$$F(x) = \sup\{h(x, y) \mid y \in Y\} \quad \text{and} \quad G(y) = \sup\{h(x, y) \mid x \in X\}.$$

Show that

$$\sup\{h(x, y) \mid (x, y) \in X \times Y\} = \sup\{F(x) \mid x \in X\} = \sup\{G(y) \mid y \in Y\}.$$

Proof: let $h(X, Y) = \{h(x, y) \mid (x, y) \in X \times Y\} \subseteq \mathbb{R}$. By definition, $h(X, Y)$ is bounded and not empty, so it has a supremum in $\mathbb{R}$, and F and G are well-defined.

Let $\alpha = \sup h(X, Y)$. Then $\alpha \geq h(x, y)$ for all $x \in X$ and $y \in Y$. In particular, if $x \in X$ is fixed, $\alpha \geq h(x, y)$ for all $y \in Y$. But $F(x)$ is the smallest upper bound of $\{h(x, y) \mid y \in Y\}$, so $\alpha \geq F(x)$.

But x was arbitrary, so $\alpha \geq F(x)$ for all $x \in X$. Hence α is an upper bound of $\{F(x) \mid x \in X\}$; by completeness, $\{F(x) \mid x \in X\}$ has a supremum in $\mathbb{R}$, say β. Then $\alpha \geq \beta$, by definition of the supremum.

Again, fix $x \in X$. Then $\beta \geq F(x) \geq h(x, y)$ for all $y \in Y$. Hence, for any $x \in X$, $\beta \geq h(x, y)$ for all $y \in Y$. As a result, β is an upper bound of $h(X, Y)$. Then $\beta \geq \alpha$, by definition of the supremum.

Combining these two results yields $\alpha = \beta$ (now do the other). ■

[2] Compare this result with the one from the previous question; what is the difference?

12. Show there exists a positive real number u such that $u^2 = 3$.

    Proof: we first show that u is not rational.[3]

    Suppose the equation $r^2 = 3$ has a positive root r in $\mathbb{Q}$. Let $r = p/q$ with $\gcd(p,q) = 1$ be that solution. Then $p^2/q^2 = 3$, or $p^2 = 3q^2$. Hence p^2 is a multiple of 3, and so p is also a multiple of 3.[4]

    Set $p = 3m$. Then $(3m)^2 = 3q^2$, which is the same as $3m^2 = q^2$. Then q^2 is a multiple of 3, and so q is also a multiple of 3. Consequently, p and q are both divisible by 3, which contradicts the hypothesis $\gcd(p,q) = 1$. The equation $r^2 = 3$ cannot then have a solution in $\mathbb{Q}$.

    But we haven't shown yet that the equation has a solution in $\mathbb{R}$. Consider the set $S = \{s \in \mathbb{R}^+ : s^2 < 3\}$, where $\mathbb{R}^+$ denotes the set of positive real numbers.

    This set in not empty as $1 \in S$. Furthermore, S is bounded above by 3. (Indeed, if $t \geq 3$, then $t^2 \geq 9 > 3$, whence $t \notin S$.) By completeness of $\mathbb{R}$, $x = \sup S \geq 1$ exists. It will be enough to show that neither $x^2 < 3$ and $x^2 > 3$ can hold. The only remaining possibility is that $x = \sqrt{3}$.

    - If $x^2 < 3$, then $\frac{2x+1}{3-x^2} > 0$. By the Archimedean property, $\exists n > 0$ such that $\frac{2x+1}{3-x^2} < n$. By re-arranging the terms, we get

    $$0 < \frac{1}{n}(2x+1) < 3 - x^2.$$

      Then

    $$\begin{aligned}\left(x + \frac{1}{n}\right)^2 = x^2 + \frac{2x}{n} + \frac{1}{n^2} \quad &\leq \quad x^2 + \frac{2x}{n} + \frac{1}{n} \\ &= \quad x^2 + \frac{1}{n}(2x+1) < x^2 + 3 - x^2 = 3.\end{aligned}$$

      Since $(x + \frac{1}{n})^2 < 3$, $x + \frac{1}{n} \in S$. But $x < x + \frac{1}{n}$; x is then not an upper bound of S, which contradicts the fact that $x = \sup S$. Thus $x^2 \not< 3$.
    - If $x^2 > 3$, then $\frac{2x}{x^2-3} > 0$. By the Archimedean property, $\exists n > 0$ such that $\frac{2x}{x^2-3} < n$. By re-arranging the terms, we get

    $$0 > -\frac{2x}{n} > 3 - x^2.$$

      Then

    $$\left(x - \frac{1}{n}\right)^2 = x^2 - \frac{2x}{n} + \frac{1}{n^2} > x^2 - \frac{2x}{n} > x^2 + 3 - x^2 = 3.$$

      Since $(x - \frac{1}{n})^2 > 3$, $x - \frac{1}{n}$ is an upper bound of S. But $x > x - \frac{1}{n}$; then x cannot be the supremum of S, which is a contradiction. Thus $x^2 \not> 3$.

[3] Even though that wasn't part of the question, it will be informative.

[4] Indeed, if p is not a multiple of 3, then neither is p^2. Let $p = 3k+1$ or $p = 3k+2$. Then $p^2 = 3(3k^2+2k)+1$ or $p^2 = 3(3k^2+4k+1)+1$, neither of which is a multiple of 3.

That leaves only one alternative (since we know that $x \in \mathbb{R}$): $x^2 = 3$, whence $x = u = \sqrt{3} > 0$. ■

13. Show there exists a positive real number u such that $u^3 = 2$.

Proof: consider the set $S = \{s \in \mathbb{R}^+ : s^3 < 2\}$, where $\mathbb{R}^+$ denotes the set of positive real numbers.

This set in not empty as $1 \in S$. Furthermore, S is bounded above by 2.[5] By completeness of $\mathbb{R}$, $x = \sup S \geq 1$ exists. It will be enough to show that neither $x^3 < 2$ and $x^3 > 2$ can hold. The only remaining possibility is that $x = \sqrt[3]{2}$.

- If $x^3 < 2$, then $\frac{3x^2+3x+1}{2-x^3} > 0$. By the Archimedean property, $\exists n > 0$ such that $\frac{3x^2+3x+1}{2-x^3} < n$. By re-arranging the terms, we get

$$0 < \frac{1}{n}(3x^2 + 3x + 1) < 2 - x^3.$$

  Then

$$\begin{aligned}\left(x + \frac{1}{n}\right)^3 &= x^3 + \frac{3x^2}{n} + \frac{3x}{n^2} + \frac{1}{n^3} \\ &\leq x^3 + \frac{3x^2}{n} + \frac{3x}{n} + \frac{1}{n} \\ &= x^3 + \frac{1}{n}(3x^2 + 3x + 1) < x^3 + 2 - x^3 = 2.\end{aligned}$$

  Since $(x + \frac{1}{n})^3 < 2$, $x + \frac{1}{n} \in S$. But $x < x + \frac{1}{n}$; x is then not an upper bound of S, which contradicts the fact that $x = \sup S$. Thus $x^3 \not< 2$.
- If $x^3 > 2$, then $\frac{3x^2+1}{x^3-2} > 0$. By the Archimedean property, $\exists n > 0$ such that $\frac{3x^2+1}{x^3-2} < n$. By re-arranging the terms, we get

$$0 > -\frac{(3x^2 + 1)}{n} > 2 - x^3.$$

  Then

$$\begin{aligned}\left(x - \frac{1}{n}\right)^3 &= x^3 - \frac{3x^2}{n} + \frac{3x}{n^2} - \frac{1}{n^3} \\ &\geq x^3 - \frac{3x^2}{n} - \frac{1}{n^3} \geq x^3 - \frac{3x^2}{n} - \frac{1}{n} \\ &= x^3 - \frac{1}{n}(3x^2 + 1) > x^3 + 2 - x^3 = 2.\end{aligned}$$

  Since $(x - \frac{1}{n})^3 > 2$, $x - \frac{1}{n}$ is an upper bound of S. But $x > x - \frac{1}{n}$; x can not then be the supremum of S, which is a contradiction. Thus $x^3 \not> 2$.

That leaves only one alternative (since we know $x \in \mathbb{R}$): $x^3 = 2$ or, equivalently, $x = u = \sqrt[3]{2} > 0$.[6] ■

[5]Indeed, if $t \geq 2$, then $t^3 \geq 8 > 2$, whence $t \notin S$.

[6]We could also show it is irrational, but we'll leave it as an exercise.

14. Let $S \subseteq \mathbb{R}$ and suppose that $s^* = \sup S$ belongs to S. If $u \notin S$, show that $\sup(S \cup \{u\}) = \sup\{s^*, u\}$.

    Proof: in this case, we do not need to verify if s^* exists, as that is one of the hypotheses. Set $v = \sup\{s^*, u\}$. Then, v is an upper bound of $S \cup \{u\}$ since $v \geq u$ and $v \geq s^* \geq s$ for all $s \in S$.

    Furthermore, $v \in S \cup \{u\}$.

    Hence, any upper bound of $S \cup \{u\}$ must be $\geq v$: consequently, v is the smallest upper bound of $\sup(S \cup \{u\})$. ■

15. Show that a non-empty finite set $S \subseteq \mathbb{R}$ contains its supremum.

    Proof: we use induction on the cardinality of S to prove the statement.

    Base case: if $|S| = 1$, then $S = \{s_1\}$ for some $s_1 \in \mathbb{R}$. Clearly, $s_1 = \sup S \in S$.

    Induction step: Suppose that the result holds for any set whose cardinality is $n = k$. Let S be any set with $|S| = k + 1$, say

    $$S = \{s_1, \ldots, s_k, s_{k+1}\}.$$

    Write $S = T \cup \{s_{k+1}\}$, with $T = \{s_1, \ldots, s_k\}$. Note that we can assume that $s_{k+1} \notin T$ (otherwise $|S| = k$).

    Then T is non-empty and bounded since it is finite (exercise: a finite set is bounded); by completeness, $t^* = \sup T$ exists. However, $|T| = k$. By the induction hypothesis, then, $\sup T \in T$, *i.e.* $t^* = s_j$ for some $j \in \{1, \ldots, k\}$.

    According to the preceding problem,

    $$\sup S = \sup(T \cup \{s_{k+1}\}) = \sup\{t^*, s_{k+1}\} \in T \cup \{s_{k+1}\} = S.$$

    By induction, any non-empty finite set then contains its supremum.[7] ■

16. If $S \subseteq \mathbb{R}$ is a non-empty bounded set and $I_S = [\inf S, \sup S]$, show that $S \subseteq I_S$. Moreover, if J is any closed bounded interval of $\mathbb{R}$ such that $S \subseteq J$, show that $I_S \subseteq J$.

    Proof: as S is non-empty and bounded, $\sup S$ and $\inf S$ exist by the completeness of $\mathbb{R}$. Since $\inf S \leq s \leq \sup S$ for all $s \in S$, then $\inf S \leq \sup S$ and so the interval $I_S = [\inf S, \sup S]$ is well-defined. Furthermore, the string of inequalities above also shows that $S \subseteq I_S$.

    Now, let $J = [a, b]$ be a closed interval containing S. Then $a \leq s \leq b$ for all $s \in S$. Thus, a is a lower bound and b is an upper bound of S. By definition,

    $$a \leq \inf S \leq \sup S \leq b,$$

    and so $I_S = [\inf S, \sup S] \subseteq [a, b] = J$. ■

[7] And its infimum too – it's the same idea.

17. Prove that if $K_n = (n, \infty)$ for $n \in \mathbb{N}$, then $\bigcap_{n\in\mathbb{N}} K_n = \varnothing$.

    Proof: suppose $x \in \bigcap K_n$. Then $x \in K_n$ for all n, i.e. $x > n$ for all $n \in \mathbb{N}$. This implies x is an upper bound of $\mathbb{N}$, which contradicts the Archimedean property. Hence, $\bigcap K_n = \varnothing$.[8] ■

18. If S is finite and $s^* \notin S$, show $S \cup \{s^*\}$ is finite.

    Proof: If $S = \varnothing$, then $S \cup \{s^*\} = \{s^*\}$ is finite as the function $f : \mathbb{N}_1 \to \{s^*\}$ defined by $f(1) = s^*$ is a bijection.

    Now, suppose $S \neq \varnothing$. As S is finite, there exist an integer k and a bijection $f : \mathbb{N}_k \to S$.
    Define the associated function $\tilde{f} : \mathbb{N}_{k+1} \to S \cup \{s^*\}$ by

    $$\tilde{f}(i) = \begin{cases} f(i) & \text{if } 1 \leq i \leq k \\ s^* & \text{if } i = k+1 \end{cases}.$$

    As $s^* \notin S$, $\tilde{f}$ is a bijection. Hence $S \cup \{s^*\}$ is finite. ■

19. Show directly that there exists a bijection between $\mathbb{Z}$ and $\mathbb{Q}$.

    Proof: write
    $$\mathbb{Q} = \left\{ \frac{m}{n} \mid m, n \in \mathbb{Z}, n > 0, \text{gcd}(m, n) = 1 \right\},$$
    where gcd(m, n) is the greatest common divisor of m, n. Define the map $f : \mathbb{Q} \to \mathbb{Z}$ by $f(\frac{m}{n}) = m$. To see that f is surjective, note that for all $m \in \mathbb{Z}$, $\frac{m}{1} \in \mathbb{Q}$ and $f(\frac{m}{1}) = m$.

    Next, we define the map $g : \mathbb{Z} \to \mathbb{Q}$ according to three cases: for numbers of the form

    a) $2^a 3^b$ with $a, b \in \{0, 1, 2, \ldots\}$, set $g(2^a 3^b) = \frac{a}{b}$.
    b) $-2^a 3^b$ with $a, b \in \{0, 1, 2, \ldots\}$, set $g(-2^a 3^b) = -\frac{a}{b}$.
    c) every other type n, set $g(n) = 0$.

    We need to check that g is well-defined, and then that it is surjective. To see that it is well-defined, we note that integers have unique prime decompositions, and $2, 3$ are prime.

    This means that every number can have *at most* one decomposition of the form $\pm 2^a 3^b$, so every number is in *at most* one case. But every number n must be in *at least* one case. Thus, every number belongs to *exactly one case*, so it is well-defined.

[8] If you do not like contradiction proofs, here is the same proof, but presented as a direct argument.
Let $x \in \mathbb{R}$. We will show that $x \notin \bigcap K_n$; as x is arbitrary, this implies $\bigcap K_n = \varnothing$. By the Archimedean property, there is a positive integer N such that $N > x$. Hence $x \notin K_n$ for all $n \geq N$. The conclusion follows. ■

To check that g is surjective, we consider some $\frac{m}{n} \in \mathbb{Q}$ and again consider three cases:

a) $\frac{m}{n} > 0$: $g(2^m 3^n) = \frac{m}{n}$.
b) $\frac{m}{n} < 0$: $g(-2^m 3^n) = \frac{m}{n}$.
c) $\frac{m}{n} = 0$: $g(5) = \frac{m}{n}$.

This completes the proof.[9] ■

20. Using only the field axioms of $\mathbb{R}$, show that the multiplicative identity of $\mathbb{R}$ is unique.

Proof: let a, b be two multiplicative identities in a field. Since a is a multiplicative identity, $ab = b$. Since b is a multiplicative identity, $ab = a$. Combining these two equations, we have $b = ab = a$. This completes the proof. ■

21. Using only the field axioms of $\mathbb{R}$, show that $(2x-1)(2x+1) = 4x^2 - 1$.

Proof: each equality is labeled with the field axiom used:

$$\begin{aligned}(2x-1)(2x+1) &\overset{\text{D1}}{=} 2x(2x+1) + (-1)(2x+1) \\ &\overset{\text{D1}}{=} (2x)(2x) + (1)2x + (-1)(2x) + (-1)(1) \\ &\overset{\text{D1}}{=} (2x)(2x) + (1 + (-1))2x + (-1)(1) \\ &\overset{\text{A4}}{=} (2x)(2x) + (-1)(1) \overset{\text{A3}}{=} (2x)(2x) - 1 \\ &\overset{\text{M1}}{=} ((2)(2))(x^2) - 1 = ((1+1)(1+1))(x^2) - 1 \\ &\overset{\text{D1}}{=} (1(1+1) + 1(1+1))x^2 - 1 \\ &\overset{\text{M3}}{=} 4x^2 - 1.\end{aligned}$$

This completes the proof. ■

22. Using only the order axioms, usual arithmetic manipulations, and inequalities between concrete numbers, prove that if $x \in \mathbb{R}$ satisfies $x < \varepsilon$ for all $\varepsilon > 0$, then $x \leq 0$.

Proof: assume first that $x > 0$. By O4 (and the fact that $0 < \frac{1}{2} < 1$), we have

$$\left(\frac{1}{2}\right) x > \left(\frac{1}{2}\right) \cdot 0 = 0$$

as well. By O3, since $\frac{x}{2} > 0$, we have

$$\frac{x}{2} < \frac{x}{2} + \frac{x}{2} = x.$$

Putting together these two sequences of inequalities, we have

$$0 < \frac{x}{2} < x.$$

But then we have found some number $\varepsilon = \frac{x}{2} > 0$ so that $x > \varepsilon$; this contradicts the original assumption. Thus, we conclude that our original assumption $x > 0$ is false; by O1, we conclude $x \leq 0$. ■

[9]Note that other bijections could also be exhibited.

23. Show that there exists some $x \in \mathbb{R}$ satisfying $x^2 + x = 5$.

Proof: consider the interval $I = [0, 10]$, define $S = \{x \in I \mid x^2 + x < 5\}$, and define $A = \sup S$. Note that for $x \in [0, 1]$,

$$x^2 + x - 5 \le 1^2 + 1 - 5 = -3 < 0,$$

so $A \ge 1$. Similarly, for $x \in [9, 10]$,

$$x^2 + x - 5 \ge 9^2 + 9 - 5 > 0,$$

so $A \le 9$.

Claim: $A^2 + A = 5$. This is shown in two parts: first we show that $A^2 + A \le 5$, then we show that $A^2 + A \ge 5$.

We show that $A^2 + A \le 5$ by contradiction. Let us assume $A^2 + A > 5$. Then, by a previous exercise, there exists some $0 < \varepsilon < 1$ so that $A^2 + A > 5 + \varepsilon$. But then for all $0 < \delta < \frac{\varepsilon}{100}$, we have

$$\begin{aligned}(A-\delta)^2 + (A-\delta) &= A^2 - 2A\delta + \delta^2 + A - \delta \\ &\ge A^2 - (2)(10)(\delta) + A - \delta \\ &\ge A^2 + A - 21\delta \\ &> A^2 + A - \varepsilon > 5.\end{aligned}$$

Furthermore, since $A \ge 1$ and $\delta \le 0.01$, we know that $A - \delta \in I$. Thus, in this case $A - \frac{\varepsilon}{100} < A$ is also an upper bound on S, contradicting the fact that A is defined to be the least upper bound on S. We conclude that $A^2 + A \le 5$.

Next, we show that $A^2 + A \ge 5$ by contradiction. Let us assume $A^2 + A < 5$. Then, by a previous exercise, there exists some $0 < \varepsilon < 1$ so that $A^2 + A < 5 - \varepsilon$. But then for all $0 < \delta < \frac{\varepsilon}{100}$, we have

$$\begin{aligned}(A+\delta)^2 + (A+\delta) &= A^2 + A + (2A + 1 + \delta)\delta \\ &\le A^2 + A + 22\delta \\ &< A^2 + A - \varepsilon < 5.\end{aligned}$$

Furthermore, since $A \le 9$ and $\delta \le 0.01$, we know that $A + \delta \in I$. Thus, in this case $A + \frac{\varepsilon}{100} \in S$ and $A + \frac{\varepsilon}{100} > A$, contradicting the fact that A is defined to be an upper bound on S. We conclude that $A^2 + A \le 5$.

Since $A^2 + A \le 5$ and $A^2 + A \ge 5$, we conclude that $A^2 + A = 5$. ■

24. Consider a set S with $0 \le \sup S = A < \infty$ and $A \notin S$. Show that for all $\varepsilon > 0$, $S \cap [A - \varepsilon, A] \neq \varnothing$. Using this fact, conclude that $S \cap [A - \varepsilon, A]$ is infinite.

Proof: we prove the first claim by contradiction.

Assume there is some $\varepsilon > 0$ such that $S \cap [A - \varepsilon, A]$ is empty. Since A is an upper bound for S, we also know that $S \cap (A, \infty)$ is empty. Thus, $S \cap [A - \varepsilon, \infty)$ is empty. But this means that $A - \varepsilon < A$ is an upper bound for s, contradicting the fact that A is the least upper bound for S. We conclude that in fact $S \cap [A - \varepsilon, A]$ is not empty.

We also prove the second part by contradiction. Assume there is some $\varepsilon > 0$ such that $S \cap [A - \varepsilon, A]$ is finite. Then we can enumerate its elements, $\{b_1, \ldots, b_n\}$. Let $B = \max(b_1, \ldots, b_n\}$.

Since $A \notin S$, we know that $b_1, \ldots, b_n < A$. Since B is a maximum of finitely many elements, we must have $B < A$ as well.

But then $A > A - \frac{A-B}{2} > B$, so $[A - \frac{A-B}{2}, A] \cap S$ is empty. This, however, is impossible according to the first part of the question.

This completes the proof. ■

25. Somebody walks up to you with a proof by induction of the statement "*For any integer $N \in \mathbb{N}$, all collections of N sheep are the same colour*," as follows:

   - **Notation:** Let $x_1, x_2, \ldots,$ be the colours of all sheep in the world, in some order.
   - **Base Case:** Obviously the first sheep is a single colour, x_1.
   - **Induction Step:** Assume that the statement is true up to some integer n.

     By the induction hypothesis, the collection of the first n sheep $\{x_1, \ldots, x_n\}$ are one colour (label this "colour 1'), and the collection of the last n sheep $\{x_2, \ldots, x_{n+1}\}$ are also one colour (label this "colour 2" - note that we haven't yet shown it is the same colour as the first collection).

     Since $\{x_2, \ldots, x_n\}$ are in both sets, we must have that "colour 1" and "colour 2" are the same, and so $\{x_1, \ldots, x_{n+1}\}$ are all one colour.

   Explain why this "proof" fails by identifying/explaining a (significant) false statement.

   Solution: the critical error is in the following part of the argument, in the case $n = 1$:

   > "the collection of the first n sheep $\{x_1, \ldots, x_n\}$ are one colour, and the collection of the last n sheep $\{x_2, \ldots, x_{n+1}\}$ are also one (possibly different) colour. Since $\{x_2, \ldots, x_n\}$ are in both sets, both sets must in fact be the same colour, and so $\{x_1, \ldots, x_{n+1}\}$ are all one colour."

   Consider the case $n = 1$. Then the collection $\{x_2, \ldots, x_n\}$ is actually empty, and so we cannot conclude that the two sets $\{x_1\}, \{x_2\}$ share any sheep, and so we cannot conclude that they are the same colour. □

1.5 Exercises

1. Prepare a 2-page summary of this chapter, with important definitions and results.
2. Complete the proof of the corollary on the density of $\mathbb{R} \setminus \mathbb{Q}$.
3. Can the union of two countable sets be uncountable? Is $\mathbb{R} \setminus \mathbb{Q}$ countable or uncountable?
4. Is the intersection of two uncountable sets uncountable or countable?
5. Complete the proof of solved problem 11.

Chapter 2

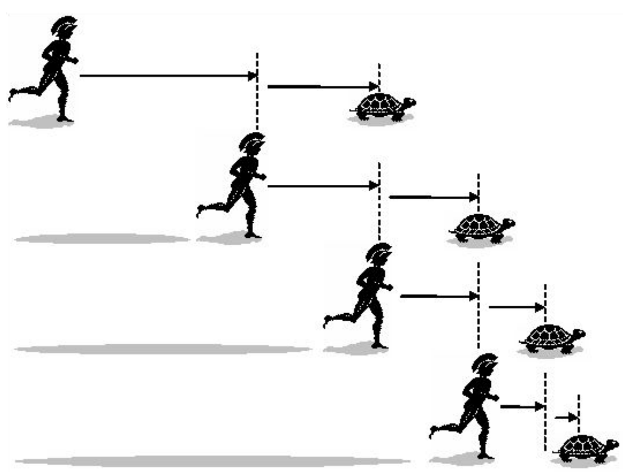
Sequences of Real Numbers

A large chunk of analysis concerns itself with problems of **convergence**. In this chapter, we introduce **sequences** and **limits**, provide results that help to **compute** such limits (when they exist), and identify situations when the limit can be shown to exist without first having to compute it.

2.1 Infinity vs. Intuition

When dealing with infinity, our intuition sometimes falters, as we shall see presently.

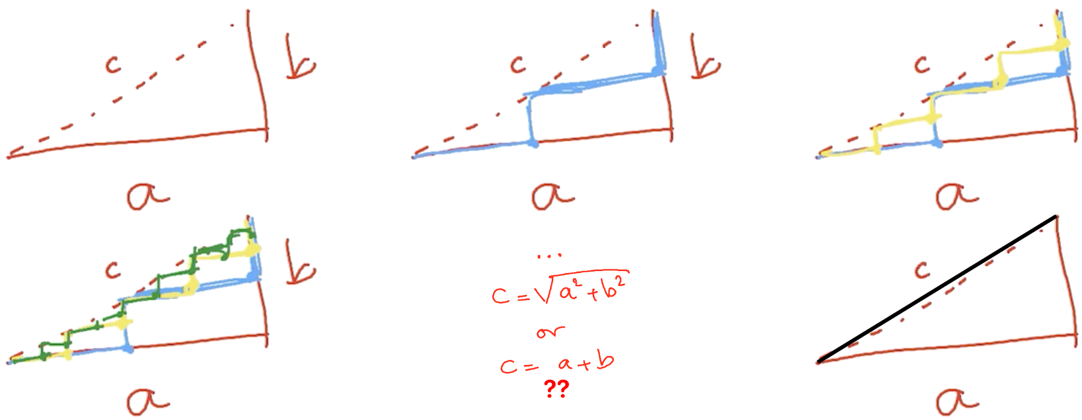
Example (ZENO'S PARADOX)
Achilles pursues a turtle. When he reaches her starting point, she has moved a certain distance. When he crosses that distance, she has moved yet another distance, and so forth. Achilles is always trailing the turtle, so he cannot catch her.



What would happen in reality? □

The next example puts one of the great classical results of planar geometry in doubt.

Example (ANTI-PYTHAGOREAN THEOREM)
Consider a right-angle triangle with base a, height b, and hypotenuse c. We can build staircase structures that each have the same constant length as $a+b$, while increasing the number of stairs (see image below).



This seems to tell us that $c = a + b$. But we know that $c = \sqrt{a^2 + b^2}$ according to Pythagoras' Theorem. Thus, we would expect to have $(a+b)^2 = a^2 + b^2$ for all right-angle triangles, which is to say, that $2ab = 0$, or, equivalently, that each right-angle triangle has at least one side with length 0. But we know this cannot be true, as the $(3, 4, 5)$ right-angle triangle demonstrates. What is going on? $\square$

Finally, we present two baffling "results" about infinite sums.

Examples (INFINITE SUMS)

1. Let $S = 1 + (-1) + 1 + (-1) + \cdots$. Then
   $S = (1 + (-1)) + (1 + (-1)) + \cdots = 0 + 0 + \cdots = 0$
   $S = 1 - (1 + (-1) + 1 + (-1) + \cdots) = 1 + S \implies S = 1/2$
   $S = 1 + ((-1) + 1) + ((-1) + 1) + \cdots = 1 + 0 + 0 + \cdots = 1$
   Therefore $0 = \frac{1}{2} = 1$. Does this make sense?

2. Let $S = 1 + 2 + 4 + 8 + \cdots$. Then
   $$S = 1 + 2(1 + 2 + 4 = 8 + \cdots) = 1 + 2S \implies S = -1.$$
   Can a sum of positive terms yield a negative result? $\square$

2.2 Limit of a Sequence

In each of the examples provided in Section 2.1, the problem arises with a "..." (implicit in Zeno's paradox, explicit in the others): seen individually, each of the steps makes sense. But when we stitch them all together – letting the number of steps increase without bounds – all hell breaks loose.

There are instances where letting $n \to \infty$ leads to **convergent behaviour**, others (as in the preceding examples), where it doesn't.[1] We start by formalizing these notions.

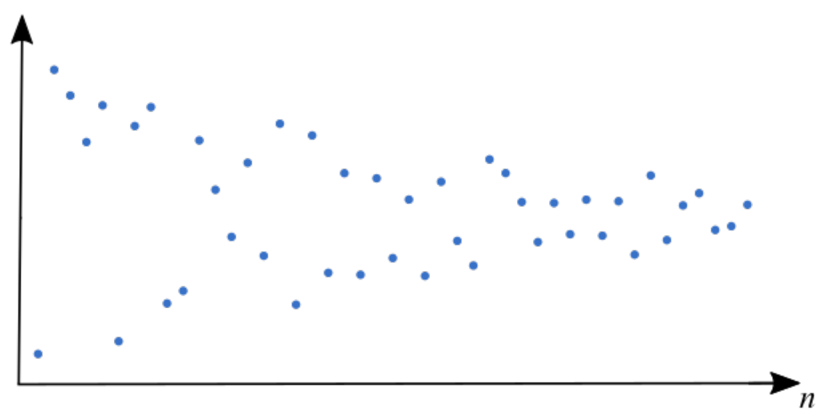
A **sequence** of real numbers is a function $X : \mathbb{N} \to \mathbb{R}$ defined by $X(n) = a_n$, where $a_n \in \mathbb{R}$. We denote the sequence X by $(a_n)_{n\in\mathbb{N}}$ or simply by (a_n).

Examples

1. $X : \mathbb{N} \to \mathbb{R}$, $n \mapsto 2n$ is the sequence with $X(1) = 2$, $X(2) = 4$, etc.; we may also write $X = (x_n) = (2, 4, 6, \ldots)$.[2]

2. $X : \mathbb{N} \to \mathbb{R}$, $n \mapsto \frac{1}{n}$ is the sequence with $X(1) = \frac{1}{2}$, $X(2) = \frac{1}{2}$, etc.; we may also write $X = (x_n) = (1, 1/2, 1/3, \ldots)$. □

In general, we let $\mathbb{N}$ stand for whatever countable subset of $\mathbb{N}$ is required for the definition of the sequence to make sense. Graphically, we can display sequences as a "scatterplot", with the horizontal coordinate being the index n and the vertical axis the value $X(n) = x_n$ of the sequence at n. An example is provided below.



We can also see a sequence as an **ordered set** of **terms** a_n, that is, a set of **indexed values**. The set of all values taken by the sequence (a_n) is called the **range** of (a_n) and we denote it by $\{a_n\}$. Sequences and their ranges are different objects.

[1]It isn't much of a stretch to state that mathematical analysis is about coming to terms with infinity – thankfully, this endeavour has proven to have extremely rich consequences, as we shall see throughout these notes.

Examples

1. The terms of the sequence $(\frac{1}{n^2})$ are $(1, \frac{1}{4}, \frac{1}{9}, \ldots)$, while its range is $\{1, \frac{1}{4}, \frac{1}{9}, \ldots\}$.
2. The terms of the sequence $(\frac{1+(-1)^n}{n})$ are $(0, 1, 0, \frac{1}{2}, 0, \frac{1}{3}, \ldots)$, while its range is $\{0, 1, \frac{1}{2}, \frac{1}{3}, \ldots\}$. $\square$

Certain sequences are defined with the help of a **recurrence relation**: the first few terms are given, and the subsequent terms are computed using the preceding terms and the relation.

Example (Fibonacci Sequence)
The classic sequence $(1, 1, 2, 3, 5, 8, 13, \ldots)$ is a recurrence relation, defined by by $x_1 = 1$, $x_2 = 1$, and $x_n = x_{n-1} + x_{n-2}$ for $n \geq 3$. $\square$

We will now examine in detail a specific sequence,

$$(x_n) = \left(\frac{1}{2n}\right) = \left(\frac{1}{2}, \frac{1}{4}, \frac{1}{6}, \frac{1}{8}, \ldots\right).$$

As n increases, the values of x_n seem to **approach** 0. What does this mean, mathematically? Let $\varepsilon > 0$.[3] Then the real number $\frac{1}{2\varepsilon}$ is positive, *i.e.*,

$$\frac{1}{2\varepsilon} > 0.$$

According to the Archimedean property, there exists a **threshold** $N_\varepsilon \in \mathbb{N}$ such that

$$N_\varepsilon > \frac{1}{2\varepsilon}.$$

Different values of ε lead to different thresholds: for instance, if $\varepsilon = \frac{1}{100}$, then any

$$N_\varepsilon > \frac{1}{2(1/100)} = 50$$

would work; if $\varepsilon = \frac{1}{1000}$, then any $N_\varepsilon > 500$ would work, and so on.

No matter what value $\varepsilon > 0$ takes, however, if we look at indices past the threshold (i.e. when $n > N_\varepsilon$), we have

$$n > N_\varepsilon > \frac{1}{2\varepsilon} \implies n > \frac{1}{2\varepsilon} \iff \varepsilon > \frac{1}{2n}.$$

For all indices n **after** the threshold N_ε (i.e. $\forall n > N_\varepsilon$), we have:

$$|x_n - 0| = |x_n| = \left|\frac{1}{2n}\right| = \frac{1}{2n} < \varepsilon \implies 0 - \varepsilon < x_n < 0 + \varepsilon.$$

[3] In theory, ε could take on any positive value, but in practice we are interested in small values $\varepsilon \ll 1$.

The interval $(-\varepsilon, \varepsilon)$ thus contains **all** the terms of the sequence x_n **after** the N_εth term, which is to say $x_n \in (-\varepsilon, \varepsilon)$ for all $n > N_\varepsilon$.

Another way of saying this is that the interval $(-\varepsilon, \varepsilon)$ contains all the terms of the sequence (x_n), except maybe for a finite number of terms included in $x_1, \ldots, x_{N_\varepsilon}$.

If $\varepsilon = 1/100$, for instance, $\exists N_{1/100} > \frac{1}{2(1/100)} = 50$ ($N_{1/100} = 51$ works) such that

$$n > 51 \Longrightarrow |x_n - 0| = |x_n| = \left|\frac{1}{2n}\right| = \frac{1}{2n} < \frac{1}{2(51)} = \frac{1}{102} < \frac{1}{100} = \varepsilon.$$

In other words, the interval $(-1/100, 1/100)$ contains all the terms of the sequence from $n = 52$ onward.

But the threshold $N_{1/100} = 51$ does not may not necessarily work for ε values smaller than $1/100$, however. If $\varepsilon = 1/1000$, say, then we need $N_{1/1000} > \frac{1}{2(1/1000)} = 500$ to guarantee that all the terms after the threshold fall in the interval $(-1/1000, 1/1000)$.

Obviously, we could find an appropriate threshold N_ε in the same manner using any $\varepsilon > 0$. This leads us to the following definition.

A sequence (x_n) of real numbers **converges** to a **limit** $L \in \mathbb{R}$, which we denote by

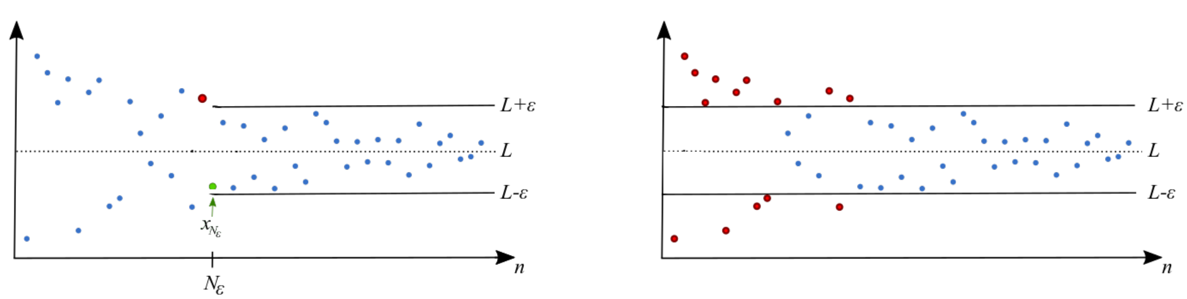
$$x_n \to L \quad \text{or} \quad \lim_{n\to\infty} x_n = L,$$

if

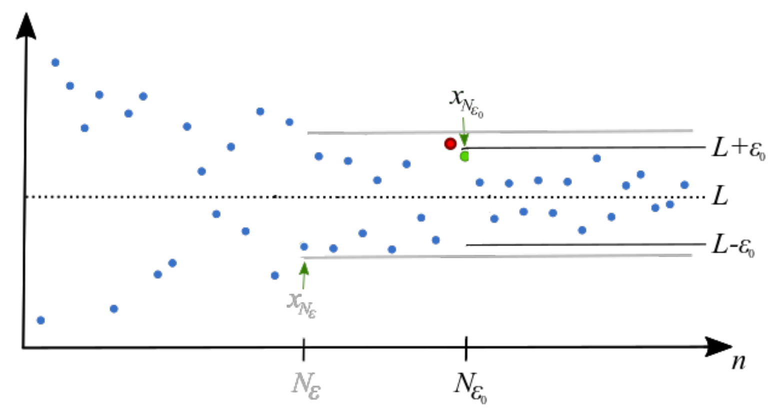
$$\forall \varepsilon > 0,\ \exists N_\varepsilon \in \mathbb{N} \text{ such that } n > N_\varepsilon \Longrightarrow |x_n - L| < \varepsilon.$$

This may look complicated, but it is just the formalized statement of the example above, where $L = 0$: we look for a **systematic** threshold N_ε after which all terms of the sequence x_n lie in $(L - \varepsilon, L + \varepsilon)$.

In the illustration below where $x_n \to L$, we find an acceptable threshold N_ε for ε on the left, and display the finite number of sequence terms falling outside of the interval $(L - \varepsilon, L + \varepsilon)$ on the right.



We also identify a threshold N_{ε_0} for $\varepsilon_0 \leq \varepsilon$ in the illustration below.



A sequence (x_n) which does not converge to a limit is said to be **divergent**:

$$\forall L \in \mathbb{R},\ \exists \varepsilon_L > 0,\ \forall N \in \mathbb{N},\ \exists n_N > N \text{ such that } |x_{n_N} - L| \geq \varepsilon_L;$$

in other words, no real number L can be the limit of (x_n).

There is only one way for a sequence to converge – its values must eventually get closer and closer to the limit; but there is more than one way for a sequence to diverge.

Examples

1. Show that $\frac{1}{n} \to 0$.

   Proof: let $\varepsilon > 0$. By the Archimedean property, $\exists N_\varepsilon > \frac{1}{\varepsilon}$, so $\varepsilon > \frac{1}{N_\varepsilon}$. If $n > N_\varepsilon$, then $\frac{1}{n} < \frac{1}{N_\varepsilon}$ and

   $$\left|\frac{1}{n} - 0\right| = \left|\frac{1}{n}\right| = \frac{1}{n} < \frac{1}{N_\varepsilon} < \varepsilon.$$

   This completes the proof. ■

2. Show that $\frac{n+1}{n^2+1} \to 0$.

   Proof: let $\varepsilon > 0$. By the Archimedean property, $\exists N_\varepsilon > \frac{2}{\varepsilon}$, so $\varepsilon > \frac{2}{N_\varepsilon}$. If $n > N_\varepsilon$, then $\frac{1}{n} < \frac{1}{N_\varepsilon}$ and

   $$\left|\frac{n+1}{n^2+1} - 0\right| = \frac{n+1}{n^2+1} \leq \frac{2n}{n^2+1} < \frac{2n}{n^2} = \frac{2}{n} < \frac{2}{N_\varepsilon} < \varepsilon.$$

   This completes the proof. ■

3. Show that $\frac{4-2n-3n^2}{2n^2+n} \to -\frac{3}{2}$.

   Proof: let $\varepsilon > 0$. By the Archimedean property, $\exists N_\varepsilon > \frac{2}{\varepsilon}$, so $\varepsilon > \frac{2}{N_\varepsilon}$. If $n > N_\varepsilon$, then $\frac{1}{n} < \frac{1}{N_\varepsilon}$ and

   $$\left|\frac{4-2n-3n^2}{2n^2+n} - \left(-\frac{3}{2}\right)\right| = \left|\frac{2(4-2n-3n^2)+3(2n^2+n)}{2(2n^2+n)}\right| = \frac{|8-n|}{4n^2+2n}.$$

   Note that $8-n \leq 8n$ if $1 \leq n \leq 8$, and that $n-8 \leq 8n$ if $n \geq 8$, so that $|8-n| \leq 8n$ for all $n \geq 1$. Thus

   $$\frac{|8-n|}{4n^2+2n} \leq \frac{8n}{4n^2+2n} < \frac{8n}{4n^2} = \frac{2}{n} < \frac{2}{N_\varepsilon} < \varepsilon$$

   when $n > N_\varepsilon$, which completes the proof. ■

4. Show that $(x_n) = (n)$ is divergent.

   Proof: suppose instead that (x_n) converges to $a \in \mathbb{R}$. Let $\varepsilon > 0$. By definition, $\exists N_\varepsilon \in \mathbb{N}$ such that $|x_n - a| = |n - a| < \varepsilon$ whenever $n > N_\varepsilon \implies$ that $n < a + \varepsilon$ for all $n > N_\varepsilon$, $\implies a + \varepsilon$ is a an upper bound for $\mathbb{N}$. This contradicts the Archimedean property, so the sequence (n) must diverge. ■

The main benefit of the formal definition of the limit of a sequence is that it **does not call on infinity**: we write $n \to \infty$, but that is a merely a notation of convenience. On the flip side, the formal definition has 2 major inconveniences:

1. it cannot be used to **determine the limit** of a convergent sequence – it can only be used to verify that a given candidate is (or is not) a limit of a sequence;

2. it can seem artificial to some extent, especially upon a first encounter.

In practice, using the definition is in fact rather simple: in order to **determine a threshold** N_ε that does the trick, we often **backtrack** from the end of the string of inequalities rather than to proceed directly from "Let $\varepsilon > 0$".

We have been careful to refer to "a" limit when the sequence converges, but we should really be talking about "the" limit in such cases.

Theorem 12 (Unique Limit)
A convergent sequence (x_n) of real numbers has exactly one limit.

Proof: suppose that $x_n \to x'$ and $x_n \to x''$. Let $\varepsilon > 0$. Then there exist 2 integers $N'_\varepsilon, N''_\varepsilon \in \mathbb{N}$ such that

$$|x_n - x'| < \varepsilon \text{ whenever } n > N'_\varepsilon \quad \text{and} \quad |x_n - x''| < \varepsilon \text{ whenever } n > N''_\varepsilon.$$

Set $N_\varepsilon = \max\{N'_\varepsilon, N''_\varepsilon\}$. Then whenever $n > N_\varepsilon$, we have

$$0 \leq |x' - x''| = |x' - x_n + x_n - x''| \leq |x_n - x'| + |x_n - x''| < \varepsilon + \varepsilon = 2\varepsilon.$$

Thus $0 \leq \frac{|x'-x''|}{2} < \varepsilon$. As $\varepsilon > 0$ was arbitrary, $\frac{|x'-x''|}{2} = 0$ and $x' = x''$. ■

Sequences have other properties, which we can sometimes use to show that they converge (or diverge). A sequence $(x_n) \subseteq \mathbb{R}$ is **bounded** by $M > 0$ if $|x_n| \leq M$ for all $n \in \mathbb{N}$.

Theorem 13
Any convergent sequence (x_n) of real numbers is bounded.

Proof: let $(x_n) \subseteq \mathbb{R}$ converge to $x \in \mathbb{R}$. Then for $\varepsilon = 1$, say, $\exists N \in \mathbb{N}$ s.t.

$$|x_n - x| < 1 \quad \text{when } n > N.$$

Thanks to the "reverse" triangle inequality (Theorem 6.6), we also have

$$|x_n| - |x| \leq |x_n - x| < 1 \quad \text{when } n > N,$$

so that $|x_n| < |x| + 1$ when $n > N$.

Finally, we set $M = \max\{|x_1|, \ldots, |x_N|, |x| + 1\}$. Then $|x_n| \leq M$ for all n, which means that (x_n) is bounded. ■

About Proofs In general, we may prove results:

- **directly**, as in Theorem 13;
- by **induction**, as in Bernouilli's inequality (Theorem 3), or
- by **contradiction**, as in the Archimedean property (Theorem 1), and so on.

The **contrapositive** of $P \implies Q$ is $\neg Q \implies \neg P$. They are **logically equivalent**, but one may prove easier to demonstrate than the other. On the other hand, the **converse** of $P \implies Q$ is $Q \implies P$. There is no general link between a statement and its converse: sometimes they are both true, sometimes they are both false, sometimes only of them is true.

Example: the contrapositive of Theorem 13 is "*Any unbounded sequence is divergent*", which is valid since Theorem 13 is true. Its converse is "*Any bounded sequence is convergent*" – if we think that the converse is true, then we try to prove it; if we think that it is false, we look for a **counter-example**. Which one is it? □

2.3 Operations on Sequences and Basic Theorems

The following result removes the need to use the formal definition... as long as we have some "ground-level" building blocks to start with.

Theorem 14 (Operations on Convergent Sequences)
Let $(x_n), (y_n)$ be convergent, with $x_n \to x$ and $y_n \to y$. Let $c \in \mathbb{R}$. Then

1. $|x_n| \to |x|$;
2. $(x_n + y_n) \to (x + y)$;
3. $x_n y_n \to xy$ *and* $cx_n \to cx$;
4. $\frac{x_n}{y_n} \to \frac{x}{y}$, *if* $y_n, y \neq 0$ *for all* n.

Proof: we show each part using the definition of the limit of a sequence.

1. Let $\varepsilon > 0$. As $x_n \to x$, $\exists N'_\varepsilon$ such that $|x_n - x| < \varepsilon$ whenever $n > N'_\varepsilon$. But $||x_n| - |x|| \leq |x_n - x|$, according to Theorem 6. Hence, for $\varepsilon > 0$, $\exists N_\varepsilon = N'_\varepsilon$ such that
$$||x_n| - |x|| \leq |x_n - x| < \varepsilon$$
whenever $n > N_\varepsilon$, *i.e.*, $|x_n| \to |x|$.

2. Let $\varepsilon > 0$; then $\frac{\varepsilon}{2} > 0$. As $x_n \to x$ and $y_n \to y$, $\exists N^x_{\frac{\varepsilon}{2}}, N^y_{\frac{\varepsilon}{2}}$ such that
$$|x_n - x| < \frac{\varepsilon}{2} \quad \text{and} \quad |y_n - y| < \frac{\varepsilon}{2} \tag{2.1}$$
whenever $n > N^x_{\frac{\varepsilon}{2}}$ and $n > N^y_{\frac{\varepsilon}{2}}$, respectively. Set $N_\varepsilon = \max\{N^x_{\frac{\varepsilon}{2}}, N^y_{\frac{\varepsilon}{2}}\}$.

   Then, whenever $n > N_\varepsilon$, which is to say, whenever n is strictly larger than both $N^x_{\varepsilon/2}$ and $N^y_{\varepsilon/2}$ simultaneously, we have:
$$\begin{aligned} |(x_n + y_n) - (x + y)| &= |(x_n - x) + (y_n - y)| \leq |x_n - x| + |y_n - y| \\ &\boxed{\text{by (2.1).}} < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon, \end{aligned}$$
   i.e., $(x_n + y_n) \to (x + y)$.

3. According to Theorem 13, (x_n) and (y_n) are bounded since they are convergent sequences. Thus $\exists M_x, M_y \in \mathbb{N}$ such that for all n, we have
$$|x_n| < M_x \quad \text{and} \quad |y_n| < M_y.$$

Let $\varepsilon > 0$; then $\frac{\varepsilon}{2M_x}, \frac{\varepsilon}{2M_y} > 0$. As $x_n \to x$, $y_n \to y$, $\exists N^x_{\frac{\varepsilon}{2M_y}}, N^y_{\frac{\varepsilon}{2M_x}} \in \mathbb{N}$ such that

$$|x_n - x| < \frac{\varepsilon}{2M_y} \quad \text{and} \quad |y_n - y| < \frac{\varepsilon}{2M_x} \tag{2.2}$$

whenever $n > N^x_{\frac{\varepsilon}{2M_y}}$ and $n > N^y_{\frac{\varepsilon}{2M_x}}$ respectively. Moreover, $|y| \leq M_y$ (see Theorem 15).

Set $N_\varepsilon = \max\{N^x_{\frac{\varepsilon}{2M_x}}, N^y_{\frac{\varepsilon}{2M_y}}\}$. Then, whenever $n > N_\varepsilon$, we have:

$$\begin{aligned} |x_n y_n - xy| = |x_n y_n - x_n y + x_n y - xy| &= |x_n(y_n - y) + y(x_n - x)| \\ \leq |x_n||y_n - y| + |y||x_n - x| &< M_x|y_n - y| + M_y|x_n - x| \\ \boxed{\text{by (2.2)}} &< M_x \cdot \frac{\varepsilon}{2M_x} + M_y \cdot \frac{\varepsilon}{2M_y} \\ &= \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon, \end{aligned}$$

i.e., $x_n y_n \to xy$. Furthermore, if the sequence (y_n) is defined by $y_n = c$ for all n, then the preceding result yields $cx_n \to cx$, since $y_n = c \to c$.[4]

4. It is enough to show $1/y_n \to 1/y$ under the Theorem's assumptions; then the result will hold by part 3. Since $y \neq 0$, $\frac{|y|}{2} > 0$. Hence, as $y_n \to y$, $\exists N_{|y|/2} \in \mathbb{N}$ such that $|y_n - y| < |y|/2$, whenever $n > N_{|y|/2}$. According to Theorem 6, we then have

$$|y| - |y_n| < |y - y_n| < \frac{|y|}{2}, \quad \text{and so} \quad \frac{|y|}{2} < |y_n| \quad \text{or} \quad \frac{1}{|y_n|} < \frac{2}{|y|} \tag{2.3}$$

whenever $n > N_{|y|/2}$ – everything is well-defined as neither y_n nor y is 0 for all n.

Let $\varepsilon > 0$. Then $|y|^2\varepsilon/2 > 0$. As $y_n \to y$, $\exists N_{|y|^2\varepsilon/2} \in \mathbb{N}$ such that

$$|y_n - y| < |y|^2 \frac{\varepsilon}{2} \tag{2.4}$$

whenever $n > N_{|y|^2 \cdot \frac{\varepsilon}{2}}$. Set $N_\varepsilon = \max\{N_{\frac{|y|}{2}}, N_{|y|^2\frac{\varepsilon}{2}}\}$. Then, whenever $n > N_\varepsilon$,

$$\begin{aligned} \left|\frac{1}{y_n} - \frac{1}{y}\right| = \left|\frac{y - y_n}{y_n y}\right| &= \frac{|y - y_n|}{|y_n y|} \\ \boxed{\text{by (2.3)}} &< \frac{2|y - y_n|}{|y|^2} \\ \boxed{\text{by (2.4)}} &< \frac{2}{|y|^2} \cdot |y|^2 \frac{\varepsilon}{2} = \varepsilon, \quad \text{i.e.,} \quad \frac{1}{y_n} \to \frac{1}{y}, \end{aligned}$$

which completes the proof. ■

Now that we have some basic tools to work with, we present two results that allow us to compute limits without operating directly on a sequence.

Theorem 15 (Comparison Theorem for Sequences)
Let $(x_n), (y_n)$ be convergent sequences of real numbers with $x_n \to x$, $y_n \to y$, and $x_n \leq y_n$ $\forall n \in \mathbb{N}$. Then $x \leq y$.

Proof: suppose that it is not the case, namely, that $x > y$. Then $x - y > 0$. Set $\varepsilon = \frac{x-y}{2} > 0$. Since $x_n \to x$ and $y_n \to y$, $\exists N_\varepsilon^x, N_\varepsilon^y \in \mathbb{N}$ s.t.

$$|x_n - x| < \varepsilon \quad \text{whenever } n > N_\varepsilon^x \quad \text{and} \quad |y_n - y| < \varepsilon \quad \text{whenever } n > N_\varepsilon^y.$$

Let $N_\varepsilon = \max\{N_\varepsilon^x, N_\varepsilon^y\}$. Then, if $n > N_\varepsilon$, we have

$$y_n < y + \varepsilon = y + \frac{x-y}{2} = \frac{x+y}{2} = x - \frac{x-y}{2} = x - \varepsilon < x_n.$$

But this contradicts the assumption that $x_n \leq y_n$ for all n, and so $x \leq y$. ■

Warning: the "$\leq$"s in the statement of Theorem 15 **cannot** be replaced by "$<$"s throughout. For instance, if $(x_n) = (\frac{1}{n+1})$ and $(y_n) = (\frac{1}{n})$, then $x_n < y_n$ for all $n \in \mathbb{N}$, but $x_n \to x = 0$, $y_n \to y = 0$, and $0 = x \not< y = 0$.

Theorem 16 (Squeeze Theorem for Sequences)
Let $(x_n), (y_n), (z_n) \subseteq \mathbb{R}$ be such that $x_n, z_n \to \alpha$ and $x_n \leq y_n \leq z_n$, $\forall n \in \mathbb{N}$. Then $y_n \to \alpha$.

Proof: let $\varepsilon > 0$. By convergence of $(x_n), (z_n)$ to α, $\exists N_\varepsilon^x, N_\varepsilon^z \in \mathbb{N}$ s.t.

$$|x_n - \alpha| < \varepsilon \text{ whenever } n > N_\varepsilon^x \quad \text{and} \quad |z_n - \alpha| < \varepsilon \text{ whenever } n > N_\varepsilon^z.$$

Let $N_\varepsilon = \max\{N_\varepsilon^x, N_\varepsilon^z\}$. When $n > N_\varepsilon$, $\alpha - \varepsilon < x_n \leq y_n \leq z_n < \alpha + \varepsilon$, which is to say, that $|y_n - \alpha| < \varepsilon$. Consequently, $y_n \to \alpha$. ■

We can use these various results to compute a fair collection of limits.

Examples

1. Compute $\lim\limits_{n\to\infty} \dfrac{3n+1}{n}$, if the limit exists.

   Solution: note that $\frac{3n+1}{n} = 3 + \frac{1}{n}$. According to Theorem 14, **if the limit exists** we must have

$$\lim_{n\to\infty} \frac{3n+1}{n} = \lim_{n\to\infty} \left(3 + \frac{1}{n}\right) = \lim_{n\to\infty} 3 + \lim_{n\to\infty} \frac{1}{n} = 3 + 0 + 3.$$

Reading the string of equations backwards, we see that the original limit must exist and be equal to 3. $\square$

2. Compute $\lim\limits_{n\to\infty} \dfrac{\sin(n^2+212)}{n}$, if the limit exists.

   Solution: we cannot use Theorem 14 since neither the numerator nor the denominator limit exists. This does not necessarily mean that the limit of the quotient does not exist. In order to determine if it does, we need to use another approach.

   By definition of the sin function (which we take for granted for now), we have $-1 \leq \sin x \leq 1, \ \forall x \in \mathbb{R}$. Thus

$$-1 \leq \sin(n^2+212) \leq 1, \ \forall n \Longrightarrow -\frac{1}{n} \leq \frac{\sin(n^2+212)}{n} \leq \frac{1}{n}, \ \forall n.$$

   As $\pm\frac{1}{n} \to 0$, we can use the squeeze theorem to conclude that

$$\lim_{n\to\infty} \frac{\sin(n^2+212)}{n} = 0. \qquad \square$$

3. Compute $\lim\limits_{n\to\infty} \dfrac{2n-1}{n+7}$, if the limit exists.

   Solution: we cannot apply Theorem 14 directly since neither the numerator nor the denominator limits exist. However,

$$\frac{2n-1}{n+7} = \frac{1/n \cdot (2n-1)}{1/n \cdot (n+7)} = \frac{2-1/n}{1+7/n} \quad \text{when } n \neq 0.$$

   Because each of the constituent parts converge (and because the denominator is never equal to 0, either in the limit or in the sequence), repeated applications of Theorem 14 yield

$$\lim_{n\to\infty} \frac{2n-1}{n+7} = \frac{\lim\limits_{n\to\infty}(2-1/n)}{\lim\limits_{n\to\infty}(1+7/n)} = \frac{2-\lim\limits_{n\to\infty} 1/n}{1+7\cdot \lim\limits_{n\to\infty} 1/n} = \frac{2-0}{1+7\cdot 0} = 2.$$

   This is basically a calculus argument. $\square$

4. Let (x_n) be such that $|x_n| \to 0$. Show that $x_n \to 0$.

   Proof: since $-|x_n| \leq x_n \leq |x_n|$ for all $n \in \mathbb{N}$ according to Theorem 6, and since $-|x_n|, |x_n| \to 0$ by assumption, then $x_n \to 0$ according to the squeeze theorem (note, however that if $|x_n| \to \alpha \neq 0$, we cannot necessarily conclude that $x_n \to \alpha$. Consider, for instance, the sequence $(x_n) = (-1)^n$). $\blacksquare$

5. Let $|q| < 1$. Compute $\lim_{n\to\infty} q^n$, if the limit exists.

   Solution: if $q = 0$, then $q^n = 0 \to 0$. If $q \neq 0$, then $\frac{1}{|q|} > 1$. Thus, $\exists t > 0$ such that $\frac{1}{|q|} = 1 + t$.

   From Bernoulli's inequality, we have

$$\left(\frac{1}{|q|}\right)^n = (1+t)^n \geq 1 + nt,\ \forall n \in \mathbb{N},$$

   so that $0 \leq |q^n| \leq |q|^n \leq \frac{1}{1+nt}$. But $\frac{1}{1+nt} = 0$ when $n \to \infty$ (does this need to be proven?); thus $|q^n| \to 0$ according to the squeeze theorem, and so $q^n \to 0$ by the previous example. □

6. Let $|q| < 1$. Compute $\lim_{n\to\infty} nq^n$, if the limit exists.

   Solution: the proof that $nq^n \to 0$ is left as an exercise; it is similar to the proof of part of the previous example, but uses an extension of Bernoulli's inequality:

$$(1+t)^n \geq 1 + nt + \frac{n(n-1)}{2}t^2,\ \text{for } t > 0, n \geq 1,$$

   which can be proven by induction. □

7. Show that $\sqrt[n]{n} \to 1$.

   Solution: let $\varepsilon > 0$. Then $1 + \varepsilon > 1$ and $0 < \frac{1}{1+\varepsilon} < 1$.

   Claim: $n\left(\frac{1}{1+\varepsilon}\right)^n \to 0$ when $n \to \infty$ (use previous example with $q = \frac{1}{1+\varepsilon}$).

   Hence, $\exists M_1 \in \mathbb{N}$ such that

$$\left|\frac{n}{(1+\varepsilon)^n} - 0\right| < 1 \text{ when } n > M_1 \Longrightarrow 1 \leq n < (1+\varepsilon)^n \text{ when } n > M_1.$$

   Set $N_\varepsilon = M_1$. Then $1-\varepsilon < 1 \leq n^{1/n} < 1+\varepsilon$ when $n > N_\varepsilon$. But this is precisely the same as $|n^{1/n} - 1| < \varepsilon$ when $n > N_\varepsilon$; thus $n^{1/n} \to 1$. □

8. Compute $\lim_{n\to\infty} \frac{n!}{n^n}$, if the limit exists.

   Solution: since

$$0 \leq \frac{n!}{n^n} = \frac{n\cdot(n-1)\cdot\cdots\cdot 2\cdot 1}{n\cdot n\cdot\cdots\cdot n\cdot n} \leq \frac{1}{n},\ \forall n \in \mathbb{N},$$

   and $\frac{1}{n} \to 0$, the squeeze theorem implies $\frac{n!}{n^n} \to 0$. □

9. Let $a > 0$. Compute $\lim\limits_{n\to\infty} a^{1/n}$, if the limit exists.

   Solution: since $a > 0$, we have $\frac{1}{a} > 0$. According to the Archimedean property, $\exists N_a \geq \max\{a, \frac{1}{a}\}$. For every $n \geq N_a$, we then have $\frac{1}{n} \leq a \leq n$. Thus $\frac{1}{\sqrt[n]{n}} \leq \sqrt[n]{a} \leq \sqrt[n]{n}$ for all $n \geq N_a$. But $\sqrt[n]{n} \to 1$ by a previous example, so $\sqrt[n]{a} \to 1$ by the squeeze theorem. □

10. Compute $\lim\limits_{n\to\infty} \sqrt[n]{3^n + 5^n}$, if the limit exists.

    Solution: since

    $$5^n \leq 3^n + 5^n \leq 5^n + 5^n = 2 \cdot 5^n \leq n \cdot 5^n, \ \forall n \geq 2,$$

    then

    $$5 \leq \sqrt[n]{3^n + 5^n} \leq \sqrt[n]{n} \cdot 5, \ \forall n \geq 2.$$

    But we have seen previously that $\sqrt[n]{n} \to 1$.

    The squeeze theorem can then be applied to the above chain of inequalities to conclude $\sqrt[n]{3^n + 5^n} \to 5$. □

We can also use the definition and theorems to demonstrate general results (that is, results about general sequences rather than about specific examples).

Theorem 17
Let $y_n \to y$. If $y_n \geq 0 \ \forall n \in \mathbb{N}$, then $\sqrt{y_n} \to \sqrt{y}$.

Proof: according to Theorem 15, we must have $y \geq 0$. There are 2 cases:

- If $y = 0$, let $\varepsilon > 0$. Then $\varepsilon^2 > 0$. Since $y_n \to 0$, $\exists M_{\varepsilon^2} \in \mathbb{N}$ s.t. whenever $n > M_{\varepsilon^2}$, we must have $|y_n - 0| = y_n < \varepsilon^2$. Now, set $N_\varepsilon = M_{\varepsilon^2}$.

  Then whenever $n > N_\varepsilon$, $|\sqrt{y_n} - 0| = \sqrt{y_n} < \sqrt{\varepsilon^2} = \varepsilon$.

- If $y > 0$, let $\varepsilon > 0$. Then $\varepsilon\sqrt{y} > 0$. Since $y_n \to y$, $\exists M_{\varepsilon\sqrt{y}} \in \mathbb{N}$ s.t. whenever $n > M_{\varepsilon\sqrt{y}}$, $|y_n - y| < \varepsilon\sqrt{y}$. Now, set $N_\varepsilon = M_{\varepsilon\sqrt{y}}$.

  Then whenever $n > N_\varepsilon$, $|\sqrt{y_n} - \sqrt{y}| = \frac{|y_n - y|}{\sqrt{y_n} + \sqrt{y}} \leq \frac{|y_n - y|}{\sqrt{y}} < \frac{\varepsilon\sqrt{y}}{\sqrt{y}} = \varepsilon$.

In both cases, we have $\sqrt{y_n} \to \sqrt{y}$. ■

2.4 Bounded Monotone Convergence Theorem

A sequence (x_n) is **increasing** if $x_1 \leq x_2 \leq \cdots x_n \leq x_{n+1} \leq \cdots$, $\forall n \in \mathbb{N}$; it is **decreasing** if $x_1 \geq x_2 \geq \cdots \geq x_n \geq x_{n+1} \cdots$, $\forall n \in \mathbb{N}$. If (x_n) is either increasing or decreasing, we say that it is **monotone**. If it is both increasing and decreasing, it is **constant**.[5]

Monotone sequences play an important role in the theory of convergence, assuming that they satisfy an additional condition.

Theorem 18 (Bounded Monotone Convergence)
Let (x_n) be an increasing sequence bounded above. Then $x_n \to \sup\{x_n \mid n \in \mathbb{N}\}$.

Proof: since the sequence (x_n) is bounded above, so it its range $\{x_n\}$. By completeness of $\mathbb{R}$, $x^* = \sup\{x_n\}$ exists. It remains only to show $x_n \to x^*$.

Let $\varepsilon > 0$. By definition, $x^* - \varepsilon$ is not an upper bound for $\{x_n\}$. Then $\exists N_\varepsilon \in \mathbb{N}$ s.t.

$$x^* - \varepsilon < x_{N_\varepsilon} \leq x^* < x^* + \varepsilon.$$

But (x_n) is increasing; in particular, $x_{N_\varepsilon} \leq x_n$ when $n > N_\varepsilon$. Thus

$$n > N_\varepsilon \Longrightarrow x^* - \varepsilon < x_n < x^* + \varepsilon,$$

so $x_n \to x^*$. ■

A similar result holds for decreasing sequences bounded below.

Examples

- Does the sequence $(x_n) = (1 - \frac{1}{n})$ converge? If so, what is its limit?

  Solution: as $\frac{1}{n} \geq \frac{1}{n+1}$ for all $n \in \mathbb{N}$,

  $$x_n - 1 - \frac{1}{n} \leq 1 - \frac{1}{n+1} \leq x_{n+1},$$

  and so (x_n) is increasing. Furthermore, $x_n \leq 1$ for all $n \in \mathbb{N}$. Then (x_n) converges according to the bounded monotone convergence theorem, and

  $$\lim_{n\to\infty} x_n = \sup_{n\in\mathbb{N}}\{x_n\} = \sup_{n\in\mathbb{N}}\{1 - 1/n\} = 1 + \sup_{n\in\mathbb{N}}\{-1/n\} = 1 - \inf_{n\in\mathbb{N}}\{1/n\} = 1,$$

  which agrees with our intuition. □

[5]When the inequalities are strict, then the sequence is **strictly increasing** or **strictly decreasing**, depending on the specific situation, and is thus **strictly monotone**.

- Let (x_n) be defined by $x_n = \sqrt{2x_{n-1}}$ when $n \geq 2$, with $x_1 = 1$. Does (x_n) converge? If so, to what limit?

  Solution: we first show, by induction, that (x_n) is increasing.

  - **Base Case:** $x_2 = \sqrt{2} \geq 1 = x_1$.
  - **Induction Step:** Suppose $x_k \geq x_{k-1}$. Then

  $$2x_k \geq 2x_{k-1} \Longrightarrow \sqrt{2x_k} \geq \sqrt{2x_{k-1}} \Longrightarrow x_{k+1} \geq x_k.$$

  Thus $x_{n+1} \geq x_n$ for all $n \in \mathbb{N}$.

  Next we show, again by induction, that (x_n) is bounded above by 2.

  - **Base Case:** $1 \leq x_1 = 1 \leq 2$.
  - **Induction Step:** Suppose $1 \leq x_k \leq 2$. Then

  $$2 \leq 2x_k \leq 2 \cdot 2 = 4 \Longrightarrow 1 \leq \sqrt{2} \leq \sqrt{2x_k} \leq \sqrt{4} = 2 \Longrightarrow 1 \leq x_{k+1} \leq 2.$$

  Thus $x_n \leq 2$ for all $n \in \mathbb{N}$ (why did we include the lower bound 1?).

  We then have, according to the bounded monotone convergence theorem,

  $$x_n \to x = \sup\{x_n \mid n \in \mathbb{N}\}.$$

  But

  $$x = \lim_{n\to\infty} x_n = \lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} \sqrt{2x_n} = \sqrt{2 \lim_{n\to\infty} x_n} = \sqrt{2x},$$

  whence $x^2 = 2x$. So either $x = 0$ or $x = 2$. But $x_n \geq 1$ for all $n \in \mathbb{N}$, so $x \geq 1$ according to Theorem 15. Thus $x_n \to 2$. $\square$

2.5 Bolazano-Weierstrass Theorem

The main result of this section, concerning bounded sequences and their subsequences, is a corner stone of analysis.

Let $(x_n) \subseteq \mathbb{R}$ be a sequence and $n_1 < n_2 < \cdots$ be an increasing string of positive integers. The sequence

$$(x_{n_k})_k = (x_{n_1}, x_{n_2}, \ldots)$$

is a **subsequence** of (x_n), denoted by $(x_{n_k}) \subseteq (x_n)$. Note that $n_k \geq k$ for all $k \in \mathbb{N}$.

Examples

- Let $(x_n) = (\frac{1}{n})$. Both $(\frac{1}{2k}) = (\frac{1}{2}, \frac{1}{4}, \ldots)$ and $(1, \frac{1}{3}, \frac{1}{6}, \frac{1}{10}, \frac{1}{15}, \frac{1}{21}, \ldots)$ are subsequences of (x_n) as they sample the original sequence while preserving the order in which the terms appear. But $(1, \frac{1}{3}, \frac{1}{2}, \frac{1}{8}, \ldots)$ is not a subsequence of (x_n) as $\frac{1}{3} = x_3$ appears before $\frac{1}{2} = x_2$.
- The sequence $(x_{3n}) = (x_3, x_6, x_9, \ldots)$ is a subsequence of (x_n) for any sequence (x_n).
- Every sequence (x_n) is a (**non-proper**) subsequence of itself.
- If $(y_k) = (x_{n_k})$ is a subsequence of (x_n) and $(z_j) = (y_{k_j})$ is a subsequence of (y_k), then $(z_j) = (x_{n_{k_j}})$ is a subsequence of (x_n). $\square$

Convergent sequences have well-behaved subsequences, as we see below.

Theorem 19 Let $x_n \to x$. If $(x_{n_k}) \subseteq (x_n)$, then $x_{n_k} \to x$ as well.

Proof: Let $\varepsilon > 0$. Since $x_n \to x$, $\exists N_\varepsilon \in \mathbb{N}$ such that $|x_n - x| < \varepsilon$ whenever $n > N_\varepsilon$. But (x_{n_k}) is a subsequence of (x_n), so $n_k \geq k$ for all $k \in \mathbb{N}$. Then $|x_{n_k} - x| < \varepsilon$ whenever $n_k \geq k > N_\varepsilon$, so $x_{n_k} \to x$ when $k \to \infty$. $\blacksquare$

Note that the converse of Theorem 19 is false (see Exercises).

The next result is surprising (at first glance) and deep, and will prove quite useful.

Theorem 20 (Bolzano-Weierstrass)
If $(x_n) \subseteq \mathbb{R}$ is bounded, it has (at least) one convergent subsequence.

Proof: we build a subsequence as follows: as (x_n) is bounded, there is an interval $I_1 = [a, b]$ s.t. $(x_n) \subseteq I_1$. Let $n_1 = 1$. Then $x_{n_1} = x_1 \in I_1$ and

$$\text{length}(I_1) = b - a = \frac{b-a}{2^0}.$$

Set $I_1' = [a, \frac{a+b}{2}]$ and $I_1'' = [\frac{a+b}{2}, b]$,

$$A_1 = \{n \in \mathbb{N} \mid n > n_1 \text{ and } x_n \in I_1'\}, \quad B_1 = \{n \in \mathbb{N} \mid n > n_1 \text{ and } x_n \in I_1''\}.$$

At least one of A_1, B_1 must be infinite as $A_1 \cup B_1 = \{n \in \mathbb{N} \mid n > n_1\}$:

- If A_1 is infinite, set $I_2 = I_1'$. Since A_1 is an infinite set of integers, it is not empty. By the well-ordering axiom, A_1 contains a smallest element, say n_2.

- If A_1 is finite, set $I_2 = I_1''$. Since B_1 is an infinite set of integers, it is not empty. By the well-ordering axiom, B_1 contains a smallest element, say n_2.

Either way, there is an integer $n_2 > n_1$ such that $x_{n_2} \in I_2$, $I_1 \supseteq I_2$ and

$$\text{length}(I_2) = \frac{b-a}{2^1}.$$

Now, suppose that $I_{k-1} \supseteq I_k$ are intervals with

$$\text{length}(I_{k-1}) = \frac{b-a}{2^{k-2}} \quad \text{and} \quad \text{length}(I_k) = \frac{b-a}{2^{k-1}},$$

that $\exists n_{k-1}, n_k \in \mathbb{N}$ such that $n_{k-1} < n_k$, $x_{n_{j-1}} \in I_{k-1}$, $x_{n_k} \in I_k$, and that at least one of the corresponding sets A_{k-1}, B_{k-1} is infinite.

Write $I_k = [\alpha, \beta]$. Set $I_k' = [\alpha, \frac{\alpha+\beta}{2}]$ and $I_k'' = [\frac{\alpha+\beta}{2}, \beta]$,

$$A_k = \{n \in \mathbb{N} \mid n > n_k \text{ and } x_n \in I_k'\}, \quad B_k = \{n \in \mathbb{N} \mid n > n_k \text{ and } x_n \in I_k''\}.$$

One of A_k, B_k must be infinite as $A_k \cup B_k = \{n \in \mathbb{N} \mid n > n_k \text{ and } x_n \in I_k\}$ is infinite.

- If A_k is infinite, set $I_{k+1} = I_k'$. Since A_k is an infinite set of integers, it is not empty. By the well-ordering axiom, A_k contains a smallest element, say n_{k+1}.
- If A_k is finite, set $I_{k+1} = I_k''$. Since B_k is an infinite set of integers, it is not empty. By the well-ordering axiom, B_k contains a smallest element, say n_{k+1}.

Either way, there is an integer $n_{k+1} > n_k$ s.t. $x_{n_{k+1}} \in I_{k+1}$, $I_k \supseteq I_{k+1}$ and

$$\text{length}(I_{k+1}) = \frac{b-a}{2^k}.$$

By induction, we have

1. $I_1 \supseteq I_2 \supseteq \cdots I_k \supseteq I_{k+1} \supseteq \cdots$;
2. for each $k \in \mathbb{N}$, length$(I_k) = \frac{b-a}{2^{k-1}}$;
3. for each $k \in \mathbb{N}$, $x_{n_k} \in I_k$, and
4. $n_1 < n_2 < \cdots < n_k < n_{k+1} < \cdots$.

Furthermore, $\frac{b-a}{2^k} \to 0$ (since it is a subsequence of $\frac{b-a}{n} \to 0$). According to the nested intervals theorem, then, $\exists \xi \in [a, b]$ such that

$$\bigcap_{k \geq 1} I_k = \{\xi\}.$$

It remains only to show that $x_{n_k} \to \xi$.

Let $\varepsilon > 0$. By the Archimedean property, $\exists K_\varepsilon \in \mathbb{N}$ such that $2^{K_\varepsilon - 1} > \frac{b-a}{\varepsilon}$, and so

$$k > K_\varepsilon \Longrightarrow 2^{K_\varepsilon - 1} < 2^{k-1} \Longrightarrow 0 \leq \frac{b-a}{2^{k-1}} < \frac{b-a}{2^{K_\varepsilon - 1}} < \varepsilon.$$

Since $\xi \in I_k$ for all $k \in \mathbb{N}$, then

$$k > K_\varepsilon \Longrightarrow |x_{n_k} - \xi| \leq \frac{b-a}{2^{k-1}} < \frac{b-a}{2^{K_\varepsilon - 1}} < \varepsilon,$$

which is to say $x_{n_k} \to x$. ∎

We have mentioned that a sequence (x_n) which diverges is one for which

$$\forall L \in \mathbb{R},\ \exists \varepsilon_L > 0,\ \forall N \in \mathbb{N},\ \exists n_N > N \text{ such that } |x_{n_N} - L| \geq \varepsilon_L.$$

If (x_n) **does not converge to** L, it is easy to construct a subsequence (x_{n_k}) which also fails to converge to L:

- let $n_1 \in \mathbb{N}$ be such that $n_1 \geq 1$ and $|x_{n_1} - L| \geq \varepsilon_L$;
- let $n_2 \in \mathbb{N}$ be such that $n_2 \geq n_1$ and $|x_{n_2} - L| \geq \varepsilon_L$;
- etc.

Note that if $x_n \not\to L$, some subsequences of (x_n) **might still converge** to L: for instance, $x_n = (-1)^n \not\to 1$, but $x_{2n} = (-1)^{2n} = 1 \to 1$.

Theorem 21
Let $(x_n) \subseteq \mathbb{R}$ be a bounded sequence such that every one of its proper converging subsequence converges to the same $x \in \mathbb{R}$. Then $x_n \to x$.

Proof: Let $M > 0$ be a bound for (x_n). Then $|x_n| \leq M$ for all $n \in \mathbb{N}$. If (x_n) does not converge to x, then $\exists (x_{n_k}) \subseteq (x_n)$ and an $\varepsilon_0 > 0$ such that

$$|x_{n_k} - x| \geq \varepsilon_0 \quad \text{for all } k \in \mathbb{N}.$$

But (x_{n_k}) is also a bounded sequence, and so, by the Bolzano-Weierstrass theorem, there is convergent subsequence $(x_{n_{k_j}}) \subseteq (x_{n_k}) \subseteq (x_n)$.

But all subsequences of (x_n) converge to x, by assumption, so $x_{n_{k_j}} \to x$. That is to say, for $\varepsilon_0 > 0$, $\exists N_{\varepsilon_0} \in \mathbb{N}$ such that $|x_{n_{k_j}} - x| < \varepsilon_0$ whenever $k_j > j > N_{\varepsilon_0}$, which contradicts the above property. Hence $x_n \to x$. ∎

2.6 Cauchy Sequences

One of the main challenge with the definition of a limit is that we need to know what L is **before** we can show what it is. Thankfully, we can bypass the circularity of the situation.We say that a sequence (x_n) is a **Cauchy sequence** if

$$\forall \varepsilon > 0,\ \exists N_\varepsilon \in \mathbb{N} \text{ such that } m,n > N_\varepsilon \Longrightarrow |x_m - x_n| < \varepsilon.$$

Incidentally, (x_n) is not a Cauchy sequence if

$$\exists \varepsilon_0 > 0,\ \forall N \in \mathbb{N},\ \exists m_N, n_N > N \text{ such that } |x_{m_N} - x_{n_N}| \geq \varepsilon_0.$$

Examples:

1. Is $(x_n) = (\frac{1}{n})$ a Cauchy sequence?

   Solution: let $\varepsilon > 0$. By the Archimedean property, $\exists N_\varepsilon > \frac{2}{\varepsilon}$. Thus

   $$m,n > N_\varepsilon \Longrightarrow \left|\frac{1}{m} - \frac{1}{n}\right| \leq \frac{1}{m} + \frac{1}{n} < \frac{1}{N_\varepsilon} + \frac{1}{N_\varepsilon} = \frac{2}{N_\varepsilon} < \varepsilon.$$

   Thus (x_n) is Cauchy. □

2. Is $(x_n) = (1 + \frac{1}{2} + \cdots + \frac{1}{n})$ a Cauchy sequence?
   Solution: let $m > n$. Then $\frac{1}{n} \geq \frac{1}{n+1} \geq \cdots \geq \frac{1}{m}$ and

   $$|x_m - x_n| = \frac{1}{n+1} + \cdots + \frac{1}{m} \geq \underbrace{\frac{1}{m} + \cdots + \frac{1}{m}}_{m-n \text{ terms}} = \frac{(m-n)}{m} = 1 - \frac{n}{m}.$$

   In particular, if $m = 2n$, then $|x_m - x_n| \geq \frac{1}{2}$ for every $n \in \mathbb{N}$, and so (x_n) is not a Cauchy sequence. □

In essence, a Cauchy sequence is a sequence for which the terms can get as close to one another as one wishes, after a certain index threshold.

The next result shows that Cauchy sequences have at least one of the traits of convergent sequences in $\mathbb{R}$ – we will soon see that the similarity is not pure happenstance.

Theorem 22
If (x_n) is a Cauchy sequence, then it is bounded.

Proof: let $1 > \varepsilon > 0$. If (x_n) is Cauchy, $\exists N_\varepsilon \in \mathbb{N}$ such that $|x_m - x_n| < \varepsilon$ whenever $m,n > N_\varepsilon$. Set $m^* = N_\varepsilon + 1$. If $n > N_\varepsilon$, then

$$|x_n| = |x_{m^*} + (x_n - x_{m^*})| \leq |x_{m^*}| + |x_n - x_{m^*}| < |x_{m^*}| + \varepsilon.$$

Set $M = \max\{|x_1| + 1, \ldots, |x_{N_\varepsilon}| + 1, |x_{m^*}| + 1\}$. Then $|x_n| \leq M$ for all $n \in \mathbb{N}$. ■

We could also show that the sum of two Cauchy sequences is a Cauchy sequence, that every bounded Cauchy sequence admits at least one convergent subsequence, and so on. In fact, any result that applies to convergent sequences in $\mathbb{R}$ also applies to Cauchy sequences in $\mathbb{R}$ (and vice-versa) because of the following result.

Theorem 23
A sequence of real numbers is convergent if and only if it is a Cauchy sequence.

Proof: let (x_n) be the sequence under consideration. Suppose that $x_n \to x$, say. Let $\varepsilon > 0$. Then $\frac{\varepsilon}{2} > 0$ and $\exists M_{\varepsilon/2}$ such that

$$n > M_{\varepsilon/2} \Longrightarrow |x_n - x| < \frac{\varepsilon}{2}.$$

Set $N_\varepsilon = M_{\varepsilon/2}$. When $n, m > N_\varepsilon$, we have

$$|x_m - x_n| \leq |x_m - x + x - x_n| \leq |x_m - x| + |x - x_n| \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

which is to say that (x_n) is Cauchy.

Now suppose that (x_n) is Cauchy. According to Theorem 22, it is a bounded sequence, and so must admit a convergent subsequence $(x_{n_k}) \subseteq (x_n)$ by the Bolzano-Weierstrass theorem, with $x_{n_k} \to x$, say.

Let $\varepsilon > 0$. Since (x_n) is Cauchy, $\exists M_{\varepsilon/2} \in \mathbb{N}$ such that

$$n, m > M_{\varepsilon/2} \Longrightarrow |x_m - x_n| < \frac{\varepsilon}{2}.$$

Since (x_{n_k}) converges to x, $\exists N > M_{\varepsilon/2}$ such that $|x_N - x| < \frac{\varepsilon}{2}$. Set $N_\varepsilon = M_{\varepsilon/2}$. Then

$$n > N_\varepsilon \Longrightarrow |x_n - x| = |x_n - x_N + x_N - x| \leq |x_n - x_N| + |x_N - x| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon,$$

and so $(x_n)$is convergent. ■

This result can help simplify proofs and computations to a considerable extent.

Examples

1. As the sequence $(x_n) = (1 + \frac{1}{2} + \cdots + \frac{1}{n})$ is not a Cauchy sequence, it does not converge.

2. Compute the limit of the sequence defined by $x_n = \frac{1}{2}(x_{n-2} + x_{n-1})$, $n > 2$, with $x_1 = 1$ and $x_2 = 2$.

Solution: we cannot use the bounded monotone convergence theorem as (x_n) is not monotone. However, (x_n) is a Cauchy sequence. Indeed,

$$\begin{aligned}|x_{n+1}-x_n| &= \left|\tfrac{1}{2}(x_{n-1}+x_n)-x_n\right| = \tfrac{1}{2}|x_n-x_{n-1}| = \tfrac{1}{2^2}|x_{n-1}-x_{n-2}| \\ &= \tfrac{1}{2^3}|x_{n-2}-x_{n-3}| = \cdots = \tfrac{1}{2^{n-1}}|x_2-x_1| = \tfrac{1}{2^{n-1}}.\end{aligned}$$

Let $\varepsilon > 0$. By the Archimedean property, $\exists N_\varepsilon \in \mathbb{N}$ such that $\frac{1}{2^{N_\varepsilon-2}} < \varepsilon$. Then, whenever $m \geq n > N_\varepsilon$,

$$\begin{aligned}|x_m-x_n| &\leq |x_m-x_{m-1}| + \cdots + |x_{n+1}-x_n| \\ &= \frac{1}{2^{m-2}} + \cdots + \frac{1}{2^{n-1}} < \frac{1}{2^{n-2}} < \frac{1}{2^{N_\varepsilon-2}} < \varepsilon.\end{aligned}$$

Being a Cauchy sequence, (x_n) is convergent according to Theorem 23. Let $x_n \to x$. From Theorem 19, we must have $x_{2n+1} \to x$ as well.

It is left as an induction exercise to show that

$$x_{2n+1} = 1 + \frac{1}{2} + \frac{1}{2^3} + \cdots + \frac{1}{2^{2n-1}} = 1 + \frac{3}{4}\left(1 - \frac{1}{4^n}\right).$$

Then $x_{2n+1} \to 1 + \frac{2}{3} = \frac{5}{3} = x$. □

Cauchy sequences illustrate the fundamental difference between $\mathbb{R}$ and $\mathbb{Q}$. A sequence is Cauchy if the points of the sequence "**accumulate**" on top of one another. We have seen that in $\mathbb{R}$, every Cauchy sequence is convergent, and *vice-versa*.

In $\mathbb{Q}$, the converging sequences are Cauchy, but there are Cauchy sequences that do not converge: it is possible that the points of such a sequence "accumulate" around one of the (uncountably infinitely) many holes of $\mathbb{Q}$. For instance, the sequence $(1, 1.4, 1.41, 1.414, \ldots)$ is Cauchy in $\mathbb{Q}$, but does not converge in $\mathbb{Q}$.

This remark leads to one of the ways of building $\mathbb{R}$ from $\mathbb{Q}$: we take all Cauchy sequences in $\mathbb{Q}$ and add whatever point the sequences "accumulates" around to $\mathbb{R}$ (there is more to it than that, but that is the main idea – We will revisit this idea in much more detail in Chapter 7). In the example above, the Cauchy sequence would lead us to add $\sqrt{2}$ to $\mathbb{Q}$.

2.7 Solved Problems

1. The first few terms of a sequence (x_n) are given below. Assuming that the "natural pattern" indicated by these terms persists, give a formula for the nth term x_n.

   a) $(5, 7, 9, 11, \ldots)$;

b) $\left(\frac{1}{2}, -\frac{1}{4}, \frac{1}{8}, -\frac{1}{16}, \ldots\right)$;

c) $\left(\frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}, \ldots\right)$;

d) $(1, 4, 9, 16, \ldots)$.

Solution: there is no general method (this question is a wee bit on the easy side...).

a) Odd integers ≥ 5: $x_n = 2n+3$ for all $n \geq 1$;

b) Alternating powers of $\frac{1}{2}$: $x_n = (-1)^{n+1}\frac{1}{2^n}$ for all $n \geq 1$;

c) Fractions where the denominator is one more than the numerator: $x_n = \frac{n}{n+1}$ for all $n \geq 1$;

d) Perfect squares ≥ 1: $x_n = n^2$ for all $n \geq 1$. □

2. Use the definition of the limit of a sequence to establish the following limits.

a) $\lim_{n\to\infty}\left(\frac{1}{n^2+1}\right) = 0$;

b) $\lim_{n\to\infty}\left(\frac{2n}{n+1}\right) = 2$;

c) $\lim_{n\to\infty}\left(\frac{3n+1}{2n+5}\right) = \frac{3}{2}$, and

d) $\lim_{n\to\infty}\left(\frac{n^2-1}{2n^2+3}\right) = \frac{1}{2}$.

Proof:

a) Let $\varepsilon > 0$. By the Archimedean property, there is a positive integer $N_\varepsilon > \frac{1}{\varepsilon}$. Then

$$\left|\frac{1}{n^2+1} - 0\right| = \frac{1}{n^2+1} < \frac{1}{n^2} \leq \frac{1}{n} < \frac{1}{N_\varepsilon} < \varepsilon,$$

whenever $n > N_\varepsilon$.

b) Let $\varepsilon > 0$. By the Archimedean property, there is a positive integer $N_\varepsilon > \frac{2}{\varepsilon}$. Then

$$\left|\frac{2n}{n+1} - 2\right| = \left|-\frac{2}{n+1}\right| = \frac{2}{n+1} < \frac{2}{n} < \frac{2}{N_\varepsilon} < \varepsilon,$$

whenever $n > N_\varepsilon$.

c) Let $\varepsilon > 0$. By the Archimedean property, there is a positive integer $N_\varepsilon > \frac{13}{4} \cdot \frac{1}{\varepsilon}$. Then

$$\left|\frac{3n+1}{2n+5} - \frac{3}{2}\right| = \left|-\frac{13}{2(2n+5)}\right| = \frac{13}{2} \cdot \frac{1}{2n+5} < \frac{13}{2} \cdot \frac{1}{2n} = \frac{13}{4} \cdot \frac{1}{n} < \frac{13}{4} \cdot \frac{1}{N_\varepsilon},$$

which is smaller than ε whenever $n > N_\varepsilon$.

d) Let $\varepsilon > 0$. By the Archimedean property, there is a positive integer $N_\varepsilon > \frac{5}{4} \cdot \frac{1}{\varepsilon}$. Then

$$\left|\frac{n^2-1}{2n^2+3} - \frac{1}{2}\right| = \left|-\frac{5}{2(2n^2+3)}\right| = \frac{5}{2} \cdot \frac{1}{2n^2+3} < \frac{5}{2} \cdot \frac{1}{2n^2} \leq \frac{5}{4} \cdot \frac{1}{n} < \frac{5}{4} \cdot \frac{1}{N_\varepsilon},$$

which is smaller than ε whenever $n > N_\varepsilon$. ■

3. Show that

a) $\displaystyle\lim_{n\to\infty}\left(\frac{1}{\sqrt{n+7}}\right) = 0$;

b) $\displaystyle\lim_{n\to\infty}\left(\frac{2n}{n+2}\right) = 2$;

c) $\displaystyle\lim_{n\to\infty}\left(\frac{\sqrt{n}}{n+1}\right) = 0$, and

d) $\displaystyle\lim_{n\to\infty}\left(\frac{(-1)^n n}{n^2+1}\right) = 0$.

Proof:

a) Let $\varepsilon > 0$. By the Archimedean property, there is a positive integer $N_\varepsilon > \frac{1}{\varepsilon^2}$. Then

$$\left|\frac{1}{\sqrt{n+7}} - 0\right| = \frac{1}{\sqrt{n+7}} < \frac{1}{\sqrt{n}} < \frac{1}{\sqrt{N_\varepsilon}} < \varepsilon,$$

whenever $n > N_\varepsilon$.

b) Let $\varepsilon > 0$. By the Archimedean property, there is a positive integer $N_\varepsilon > \frac{4}{\varepsilon}$. Then

$$\left|\frac{2n}{n+2} - 2\right| = \left|-\frac{4}{n+2}\right| = \frac{4}{n+2} < \frac{4}{n} < \frac{4}{N_\varepsilon} < \varepsilon,$$

whenever $n > N_\varepsilon$.

c) Let $\varepsilon > 0$. By the Archimedean property, there is a positive integer $N_\varepsilon > \frac{1}{\varepsilon^2}$. Then

$$\left|\frac{\sqrt{n}}{n+1} - 0\right| = \frac{\sqrt{n}}{n+1} < \frac{\sqrt{n}}{n} = \frac{1}{\sqrt{n}} < \frac{1}{\sqrt{N_\varepsilon}} < \varepsilon,$$

whenever $n > N_\varepsilon$.

d) Let $\varepsilon > 0$. By the Archimedean property, there is a positive integer $N_\varepsilon > \frac{1}{\varepsilon}$. Then

$$\left|\frac{(-1)^n n}{n^2+1} - 0\right| = \frac{n}{n^2+1} < \frac{n}{n^2} = \frac{1}{n} < \frac{1}{N_\varepsilon} < \varepsilon,$$

whenever $n > N_\varepsilon$. ■

4. Show that $\lim_{n\to\infty}\left(\frac{1}{n}-\frac{1}{n+1}\right)=0.$

Proof: let $\varepsilon>0$. By the Archimedean property, there is a positive integer $N_\varepsilon>\frac{1}{\sqrt{\varepsilon}}$. Then

$$\left|\frac{1}{n}-\frac{1}{n+1}-0\right|=\frac{1}{n(1+n)}<\frac{1}{n^2}<\frac{1}{N_\varepsilon^2}<\varepsilon,$$

whenever $n>N_\varepsilon$. ■

5. Find the limit of the following sequences:

a) $\lim_{n\to\infty}\left(\left(2+\frac{1}{n}\right)^2\right)$;

b) $\lim_{n\to\infty}\left(\frac{(-1)^n}{n+2}\right)$;

c) $\lim_{n\to\infty}\left(\frac{\sqrt{n}-1}{\sqrt{n}+1}\right)$, and

d) $\lim_{n\to\infty}\left(\frac{n+1}{n\sqrt{n}}\right)$.

Solution: we can only use the definition if we have a candidate. Throughout, we will assume that it is known that $\frac{1}{n}\to 0$.

a) Note that $(2+\frac{1}{n})^2=4+\frac{2}{n}+\frac{1}{n^2}$. Then, by Theorem 14 (operations on sequences and limits),

$$\frac{2}{n}=2\cdot\frac{1}{n}\to 2\cdot 0=0 \quad \text{and}\,\frac{1}{n^2}=\frac{1}{n}\cdot\frac{1}{n}\to 0\cdot 0=0,$$

so that $4+\frac{2}{n}+\frac{1}{n^2}\to 4+0+0=4$.

b) Clearly,

$$\frac{-1}{n+2}\leq\frac{(-1)^n}{n+2}\leq\frac{1}{n+2},\quad \forall n\in\mathbb{N}.$$

Note that $n+2\geq n$ for all n so that

$$0\leq\frac{1}{n+2}\leq\frac{1}{n},\quad \forall n\in\mathbb{N};$$

as a result, $\frac{1}{n+2}\to 0$ by the squeeze theorem. Then $-\frac{1}{n+2}\to -0=0$ by Theorem 14, so that $\frac{(-1)^n}{n+2}\to 0$ by the squeeze theorem.

c) Re-write $\frac{\sqrt{n}-1}{\sqrt{n}+1}=1-\frac{2}{\sqrt{n}+1}$. Note that

$$0\leq\frac{1}{\sqrt{n}+1}<\frac{1}{\sqrt{n}},\quad \forall n\in\mathbb{N}.$$

We have seen that $\frac{1}{\sqrt{n}}\to 0$; as a result of the squeeze theorem, $\frac{1}{\sqrt{n}+1}\to 0$. Then $1-\frac{2}{\sqrt{n}+1}\to 1-2\cdot 0=1$, by theorem 14.

d) Note that $n \leq n\sqrt{n} \leq n^2$ for all $n \in \mathbb{N}$ so

$$\frac{1}{n^2} \leq \frac{1}{n\sqrt{n}} \leq \frac{1}{n}, \quad \forall n \in \mathbb{N}.$$

But $\frac{1}{n}, \frac{1}{n^2}, \frac{1}{\sqrt{n}} \to 0$ (see previous problems) so that $\frac{1}{n\sqrt{n}} \to 0$ by the squeeze theorem. Furthermore,

$$\frac{n+1}{n\sqrt{n}} = \frac{1}{\sqrt{n}} + \frac{1}{n\sqrt{n}} \to 0 + 0 = 0,$$

by Theorem 14. □

6. Let $y_n = \sqrt{n+1} - \sqrt{n}$. Show that (y_n) and $(\sqrt{n}y_n)$ converge.

Proof: as

$$0 \leq \sqrt{n+1} - \sqrt{n} = \frac{1}{\sqrt{n+1} + \sqrt{n}} \leq \frac{1}{\sqrt{n}}, \quad \forall n \in \mathbb{N},$$

and $\frac{1}{\sqrt{n}} \to 0$, then $\sqrt{n+1} - \sqrt{n} \to 0$ by the squeeze theorem.
Note that $\sqrt{n}y_n = \sqrt{n(n+1)} - n = \frac{1}{\sqrt{1+\frac{1}{n}+1}}$ for all $n \in \mathbb{N}$. Then, according to theorem 14,

$$\lim_{n\to\infty} \sqrt{n}y_n = \lim_{n\to\infty} \frac{1}{\sqrt{1+\frac{1}{n}}+1} = \frac{1}{\lim_{n\to\infty}\left(\sqrt{1+\dfrac{1}{n}}+1\right)} = \frac{1}{2},$$

since $\sqrt{1+\frac{1}{n}}+1 > 2$ for all $n \in \mathbb{N}$. ■

7. Let $(x_n) \subseteq \mathbb{R}^+$ be such that $x_n^{1/n} \to L < 1$ for all n. Show $\exists r \in (0,1)$ such that $0 < x_n < r^n$ for all sufficiently large $n \in \mathbb{N}$. Use this result to show that $x_n \to 0$.

Proof: since $L < 1$, $\exists \varepsilon_0 > 0$ such that $L < L + \varepsilon_0 < 1$. Then, $\exists N_0 \in \mathbb{N}$ such that

$$|x_n^{1/n} - L| < \varepsilon_0 \quad \text{whenever } n > N_0.$$

Hence $L - \varepsilon_0 < x_n^{1/n} < L + \varepsilon_0$ for all $n > N_0$. Set $r = L + \varepsilon_0$. Then $r \in (0,1)$ and

$$0 < x_n < r^n, \quad \forall n > N_0.$$

Let $\varepsilon > 0$. $r^n \to 0$ (do you know how to show this?), $\exists N_\varepsilon \geq N_0$ such that $r^n < \varepsilon$ whenever $n > N_\varepsilon$, hence

$$|x_n - 0| = x_n < r^n < \varepsilon$$

whenever $n > N_\varepsilon$. ■

8. Give an example of a convergent (resp. divergent) sequence (x_n) of positive real numbers with $x_n^{1/n} \to 1$.

Solution: the sequences $(x_n) = \frac{1}{n}$ and $(x_n) = (n)$ do the trick, among others. □

9. Let $x_1 = 1$, $x_{n+1} = \sqrt{2 + x_n}$ for $n \in \mathbb{N}$. Show that (x_n) converges; find the limit.

   Proof: we show (x_n) is increasing and bounded by induction; according to the bounded monotone convergence theorem, (x_n) must then converge.

   A quick computation shows that $x_2 = \sqrt{3}$.

   Initial case: Clearly, $1 \leq x_1 \leq x_2 \leq 2$.

   Induction hypothesis: Suppose $1 \leq x_k \leq x_{k+1} \leq 2$. Then

   $$3 \leq x_k + 2 \leq x_{k+1} + 2 \leq 4$$

   and so

   $$1 \leq \sqrt{3} \leq \sqrt{x_k + 2} \leq \sqrt{x_{k+1} + 2} \leq \sqrt{4} = 2,$$

   i.e. $1 \leq x_{k+1} \leq x_{k+2} = 2$.

   Hence (x_n) is increasing and bounded above by 2; as such $x_n \to x$ for some $x \in \mathbb{R}$. But

   $$x = \lim_{n\to\infty} x_n = \lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} \sqrt{2 + x_n} = \sqrt{2 + \lim_{n\to\infty} x_n} = \sqrt{2 + x},$$

   that is, $x^2 = 2 + x$. The only solutions are $x = 2$ or $x = -1$, but $x = -1$ must be rejected since $1 \leq x_n$ for all n.

   Thus, $x_n \to 2$.

10. Let $x_n = \displaystyle\sum_{k=1}^{n} \frac{1}{k^2}$ for all $n \in \mathbb{N}$. Show that (x_n) is increasing and bounded above.

   Proof: as $\frac{1}{(n+1)^2} > 0$ for all $n \in \mathbb{N}$, we have

   $$x_n = \frac{1}{1^2} + \cdots + \frac{1}{n^2} \leq \frac{1}{1^2} + \cdots + \frac{1}{n^2} + \frac{1}{(n+1)^2} = x_{n+1}.$$

   Furthermore, for any $k \geq 2 \in \mathbb{N}$, we have $\frac{1}{k^2} < \frac{1}{k-1} - \frac{1}{k}$. Then

   $$\begin{aligned} x_n &= \frac{1}{1^2} + \frac{1}{2^2} + \cdots + \frac{1}{n^2} \\ &\leq 1 + \left(\frac{1}{1} - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \cdots + \left(\frac{1}{n-1} - \frac{1}{n}\right) \\ &= 1 + 1 + 0 + \cdots + 0 - \frac{1}{n} < 2 \end{aligned}$$

   for all $n \in \mathbb{N}$. Hence (x_n) is increasing and bounded above by 2.

11. Show that $c^{1/n} \to 1$ if $0 < c < 1$.

Proof: let $x_n = c^{1/n}$ for all $n \in \mathbb{N}$. Since $x_{n+1} = c^{1/(n+1)} > c^{1/n} = x_n$ for all $n \in \mathbb{N}$ (as $c < 1$), then (x_n) is increasing. Furthermore, $0 < c^{1/n} < 1^{1/n} = 1$ for all $n \in \mathbb{N}$, so (x_n) is bounded above.

Hence (x_n) converges, and $x_n \to x$, for some $x \in \mathbb{R}$. As all subsequences of a convergent sequence converge to the same limit as the convergent sequence, $x_{2n} = c^{1/2n} \to x$. As such,

$$x = \lim_{n\to\infty} c^{1/2n} = \lim_{n\to\infty} \sqrt{c^{1/n}} = \lim_{n\to\infty} \sqrt{x_n} = \sqrt{\lim_{n\to\infty} x_n} = \sqrt{x},$$

and so either $x = 0$ or $x = 1$. But as x_n increases to 1, there comes a point after which all x_n are "far" from 0 (you should mathematicize this statement...), so $x_n \to 1$. ■

12. Let (x_n) be a bounded sequence and let $s_n = \sup\{x_k : k \geq n\}$. If $S = \inf\{s_n\}$, show that there is a subsequence of (x_n) that converges to S.

Proof: as (x_n) is bounded, $\exists M > 0$ such that $-M < x_n < M$ for all $n \in \mathbb{N}$. By definition, $s_1 \geq s_2 \geq \cdots$ and $s_n \geq x_k$ for all $n \in \mathbb{N}$, $k \geq n$.

Hence $s_n > -M$ for all n and (s_n) is bounded below and decreasing, i.e. (s_n) is convergent. Furthermore, for each $n \in \mathbb{N}$, as $s_n = \sup\{x_k : k \geq n\}$, $\exists k_n \in \mathbb{N}$ s.t.

$$s_n - \frac{1}{n} \leq x_{k_n} < s_n$$

(otherwise s_n is not the supremum).

The sequence (x_{k_n}) might not necessarily be a subsequence of (x_n), but by deleting the terms that are out of order, the resulting sequence, which we will also denote by (x_{k_n}) is a subsequence of (x_n).

Then

$$0 \leq |x_{k_n} - s_n| \leq \frac{1}{n}, \quad \forall n \in \mathbb{N}.$$

By the squeeze theorem,

$$0 \leq \lim_{n\to\infty} |x_{k_n} - s_n| \leq 0, \quad \text{so } \lim_{n\to\infty} |x_{k_n} - s_n| = 0.$$

But this means that

$$\lim_{n\to\infty} x_{k_n} = \lim_{n\to\infty} s_n = S, \quad \text{(why?)}$$

where the last equation comes from the theorem on bounded increasing/decreasing sequences. ■

13. Suppose that $x_n \geq 0$ for all $n \in \mathbb{N}$ and that $((-1)^n x_n)$ converges. Show that (x_n) converges.

Proof: Let $(-1)^n x_n \to \alpha$. Consider its subsequences

$$\left((-1)^{2n} x_{2n}\right) = (x_{2n}) \quad \text{and} \quad \left((-1)^{2n+1} x_{2n+1}\right) = (-x_{2n+1}).$$

Then $x_{2n} \to \alpha$ and $(-x_{2n+1}) \to \alpha$. But $x_{2n} \geq 0\ \forall n \in \mathbb{N}$ so $\alpha \geq 0$. Similarly, $-x_{2n+1} \leq 0\ \forall n \in \mathbb{N}$ so $\alpha \leq 0$. Since $0 \leq \alpha \leq 0$, we must then have $\alpha = 0$. By Theorem 14 (operations on limits), we have:

$$\lim_{n\to\infty} |(-1)^n x_n| = |0| = 0.$$

But $|(-1)^n x_n| = x_n\ \forall n$, so $x_n \to 0$. ■

14. Show that if (x_n) is unbounded, there exists a subsequence (x_{n_k}) with $1/x_{n_k} \to 0$.

Proof: as (x_n) is unbounded, $\exists n_1 \in \mathbb{N}$ such that $|x_{n_1}| \geq 1$. Moreover, $\forall k \geq 2$, $\exists n_k \in \mathbb{N}$ such that $|x_{n_k}| \geq k$ and $n_{k+1} > n_k$ (otherwise the sequence would be bounded).

Let $\varepsilon > 0$. According to the Archimedean property, $\exists K_\varepsilon \in \mathbb{N}$ such that $K_\varepsilon > \frac{1}{\varepsilon}$ and

$$\left|\frac{1}{x_{n_k}} - 0\right| = \frac{1}{|x_{n_k}|} \leq \frac{1}{k} < \frac{1}{K_\varepsilon} < \varepsilon$$

whenever $k > K_\varepsilon$. Thus, $1/x_{n_k} \to 0$. ■

15. If $x_n = \frac{(-1)^n}{n}$, find the convergent subsequence in the proof of the Bolzano-Weierstrass theorem, with $I_1 = [-1, 1]$.

Proof: we first note that (x_n) is bounded by -1 and 1, so the question makes sense. Let $n_1 = 1$. Then $x_{n_1} = x_1 = -1$ and length$(I_1) = 2$. Set $I_1' = [-1, 0]$ and $I_1'' = [0, 1]$.

We have

$$A_1 = \{n \in \mathbb{N} \mid n > n_1 \text{ and } x_n \in I_1'\} = \{3, 5, 7, 9, 11, \ldots\}$$

and

$$B_1 = \{n \in \mathbb{N} \mid n > n_1 \text{ and } x_n \in I_1''\} = \{2, 4, 6, 8, 10, \ldots\}.$$

Since A_1 is infinite (why?), set $I_2 = I_1' = [-1, 0]$ and $n_2 = \min A_1 = 3$, so that $x_{n_2} = -1/3$. Note that $n_2 > n_1$, $I_2 \subseteq I_1$, and length$(I_2) = 1$. Set $I_2' = [-1, -1/2]$ and $I_2'' = [-1/2, 0]$.

We have

$$A_2 = \{n \in \mathbb{N} \mid n > n_2 \text{ and } x_n \in I_2'\} = \varnothing$$

and

$$B_2 = \{n \in \mathbb{N} \mid n > n_2 \text{ and } x_n \in I_2''\} = \{5, 7, 9, 11, 13, \ldots\}.$$

Since A_2 is finite, set $I_3 = I_2'' = [-1/2, 0]$ and $n_3 = \min B_2 = 5$, so that $x_{n_3} = -1/5$. Note that $n_3 > n_2 > n_1$, $I_3 \subseteq I_2 \subseteq I_1$, and length$(I_3) = 1/2$.

For $k \geq 3$, we set $I_k' = [-1/2^{k-2}, -1/2^{k-1}]$ and $I_k'' = [-1/2^{k-1}, 0]$. Then

$$A_k = \{n \in \mathbb{N} \mid n > n_k \text{ and } x_n \in I_k'\} = \varnothing$$

and

$$B_k = \{n \in \mathbb{N} \mid n > n_k \text{ and } x_n \in I_k''\} = \{2k+1, 2k+3, 2k+5, \ldots\}.$$

A_k is finite, so set $I_{k+1} = I_k'' = [-1/2^{k-1}, 0]$. Furthermore, $n_{k+1} = \min B_k = 2k+1$ so that $x_{n_k} = \frac{-1}{2k+1}$.

Note that $n_{k+1} > n_k > \cdots > n_2 > n_1$, $I_{k+1} \subseteq I_k \subseteq \cdots \subseteq I_2 \subseteq I_1$ and $\text{length}(I_{k+1}) = 1/2^{k-2}$. The convergent subsequence is thus $-1, -1/3, -1/5, \ldots \to 0$. ∎

16. Show directly that a bounded increasing sequence is a Cauchy sequence.

Proof: let $\varepsilon > 0$. By completeness of $\mathbb{R}$, $x^* = \sup\{x_n \mid n \in \mathbb{N}\}$ exists as $\{x_n \mid n \in \mathbb{N}\}$ is bounded and non-empty. In particular, $\exists M_{\frac{\varepsilon}{2}} \in \mathbb{N}$ such that

$$x^* - \frac{\varepsilon}{2} < x_{M_{\frac{\varepsilon}{2}}} \leq x^*.$$

But $x^* \geq x_n > x_{M_{\frac{\varepsilon}{2}}}$ whenever $n > M_{\frac{\varepsilon}{2}}$.

Let $N_\varepsilon = M_{\frac{\varepsilon}{2}}$. Then

$$|x_m - x_n| = |x_m - x^* + x^* - x_n| \leq |x^* - x_m| + |x^* - x_n| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

whenever $m, n > N_\varepsilon$. ∎

17. If $0 < r < 1$ and $|x_{n+1} - x_n| < r^n$ for all $n \in \mathbb{N}$, show that (x_n) is Cauchy.

Proof: let $\varepsilon > 0$. By the Archimedean property, $\exists N_\varepsilon > \log_r(\varepsilon(1-r)) + 1$, i.e. $r^{N_\varepsilon - 1} < \varepsilon$. Then

$$\begin{aligned} |x_m - x_n| &\leq |x_m - x_{m-1}| + \cdots + |x_{n+1} - x_n| \\ &< r^{m-1} + \cdots + r^n < \frac{r^{n-1}}{1-r} < \frac{r^{N_\varepsilon - 1}}{1-r} < \varepsilon \end{aligned}$$

whenever $m > n > N_\varepsilon$.[6] ∎

18. If $x_1 < x_2$ and $x_n = \frac{1}{2}(x_{n-1} + x_{n-2})$ for all $n \in \mathbb{N}$, show that (x_n) is convergent and compute its limit.

Proof: we start by showing that (x_n) is Cauchy. Let $L = x_2 - x_1$; by induction,

$$|x_n - x_{n-1}| \leq \frac{L}{2^{n-2}}.$$

[6] The third last inequality holds since $r^{m-1} + \cdots + r^n$ is a geometric progression.

Let $\varepsilon > 0$. By the Archimedean property, $\exists N_\varepsilon \in \mathbb{N}$ such that $\frac{L}{2^{N_\varepsilon - 2}} < \varepsilon$. Then

$$\begin{aligned}|x_m - x_n| &\leq |x_m - x_{m-1}| + \cdots + |x_{n+1} - x_n| \\ &\leq \frac{L}{2^{m-2}} + \cdots \frac{L}{2^{n-1}} \leq \frac{L}{2^{n-2}} < \frac{L}{2^{N_\varepsilon - 2}} < \varepsilon\end{aligned}$$

whenever $m > n > N_\varepsilon$. Hence (x_n) is a Cauchy sequence, and so it converges, say to $x_n \to x$. We can show by induction (do it!) that

$$x_{2n+1} = x_1 + \frac{L}{2} + \frac{L}{2^3} + \cdots + \frac{L}{2^{2n-1}}$$

for all $n \in \mathbb{N}$. In particular,

$$\begin{aligned}x = \lim_{n\to\infty} x_{2n+1} &= x_1 + \lim_{n\to\infty}\left(\frac{L}{2} + \frac{L}{2^3} + \cdots + \frac{L}{2^{2n-1}}\right) \\ &= x_1 + \frac{L}{2}\lim_{n\to\infty}\left(1 + \frac{1}{2^2} + \cdots + \frac{1}{2^{2n-2}}\right) \\ &= x_1 + \frac{L}{2}\lim_{n\to\infty}\left(\frac{1-(1/2^2)^n}{1-(1/2^2)}\right) = x_1 + \frac{2}{3}L = \frac{1}{3}(x_1 + 2x_2).\end{aligned}$$

For instance, when $x_1 = 1$ and $x_2 = 2$, $x_n \to 5/3$. ■

19. Suppose that (a_n) is a bounded sequence and $b_n \to 0$. Show that $a_n b_n \to 0$.

Proof: since (a_n) is bounded, there exists some $0 \leq M < \infty$ so that $\sup_n |a_n| \leq M$. Next, we will check that $a_n b_n \to 0$.

Fix some $\varepsilon > 0$. Since $b_n \to 0$, there exists some N_ε so that for all $n > N_\varepsilon$, $|b_n| \leq \frac{\varepsilon}{M}$. Thus, for all $n > N_\varepsilon$,

$$|a_n b_n| \leq M|b_n| \leq M\,\frac{\varepsilon}{M} = \varepsilon.$$

Thus, $a_n b_n \to 0$. ■

20. Let (a_n) be a sequence with no convergent subsequences. Show that $|a_n| \to \infty$.

Proof: we prove this by contradiction. Assume that $|a_n|$ does not diverge to infinity. Then there exists some $M < \infty$ such that the set $\{n \in \mathbb{N} \mid |a_n| < M\}$ is infinite. Let

$$1 \leq m_1 \leq m_2 \leq m_3 \leq \ldots$$

be the indices satisfying $|a_{m_n}| < M$. Set $b_n = a_{m_n}$. Then $\{b_n\}$ is a bounded sequence and so has a convergent subsequence $\{b_{k_n}\}_n$ according to the Bolzano-Weierstrass theorem.

But $\{a_{m_{k_n}}\}_n = \{b_{k_n}\}_n$ is in fact a convergent subsequence of (a_n), contradicting the information given in the question. We conclude that our assumption was false, and so that $|a_n|$ diverges to infinity. ■

21. We define the **limit inferior** and the **limit superior** of a sequence as follows:

$$\liminf_{n\to\infty} a_n = \lim_{n\to\infty} \inf\{a_k \mid k \geq n\}$$
$$\limsup_{n\to\infty} a_n = \lim_{n\to\infty} \sup\{a_k \mid k \geq n\}.$$

Let (a_n) be bounded. Show that $\liminf_{n\to\infty} a_n$ and $\limsup_{n\to\infty} a_n$ exist and are in $\mathbb{R}$.

Proof: define the sequence of sets $B_n = \{a_k \mid k \geq n\}$ and the sequence of numbers $b_n = \sup(B_n)$, so that
$$\limsup_{n\to\infty} a_n = \lim_{n\to\infty} b_n.$$
We note that $B_1 \supset B_2 \supset \ldots$, which implies $\sup(B_1) \leq \sup(B_2) \leq \ldots$, which means that $\{b_n\}$ is monotone decreasing. Furthermore, since (a_n) is bounded, there exists some $-\infty < M < \infty$ so that $a_n \geq M$ for all $n \in \mathbb{N}$.

But this M is a lower bound for (a_n), which means it must be a lower bound for B_n for all $n \in \mathbb{N}$, which means $b_n = \sup(B_n) \geq M$ for all $n \in \mathbb{N}$ as well.

Thus, we have shown that $\{b_n\}$ is a monotone decreasing sequence that is bounded from below. Hence, by the monotone convergence theorem, it has a limit and so

$$\limsup_{n\to\infty} a_n = \lim_{n\to\infty} b_n$$

exists. The proof for the lim inf statement follows a similar path. ■

22. Let (a_n) be unbounded. Show that $\liminf_{n\to\infty} a_n = -\infty$ or $\limsup_{n\to\infty} a_n = \infty$.

Proof: since (a_n) is unbounded, for all $0 < M < \infty$, there exists $n = n(M)$ satisfying $|a_n| > M$.

Define the subsequence $\{b_k\}$ by setting $b_k = a_{n(k)}$, so that $|b_k| > k$ for all $k \in \mathbb{N}$. Since this is an infinite sequence, we have by the Pigeonhole Principle that at least one of the two sets $I_+ = \{k \in \mathbb{N} \mid b_k \geq 0\}$, $I_- = \{k \in \mathbb{N} \mid b_k \leq 0\}$ is infinite.

In the case that I_+ is infinite, write the elements $i_1 < i_2 < i_3 < \ldots$ in order and define the subsequence $\{c_\ell\}$ of $\{b_n\}$ by the formula $c_\ell = b_{i_\ell} = a_{n(i_\ell)}$. But then for all n, we have

$$\begin{aligned}\sup\{a_k \mid k \geq n\} &\geq \sup\{a_{n(i_\ell)} \mid \ell \geq n\} \\ &= \sup\{c_k \mid k \geq n\} \geq \sup\{k \mid k \geq n\} = \infty.\end{aligned}$$

Thus,

$$\limsup_{n\to\infty} a_n = \infty.$$

The case that I_- is infinite is essentially the same, with the conclusion

$$\liminf_{n\to\infty} a_n = -\infty.$$

This completes the proof.[7] ■

23. Let (a_n), (b_n) be two sequences. Show that

$$\liminf_{n\to\infty} a_n + \limsup_{n\to\infty} b_n \leq \limsup_{n\to\infty}(a_n+b_n) \leq \limsup_{n\to\infty} a_n + \limsup_{n\to\infty} b_n.$$

Proof: fix $\varepsilon > 0$. Then there exists some $N_\varepsilon \in \mathbb{N}$ such that, for all $m > N_\varepsilon$, the following inequalities all hold:

$$\frac{\varepsilon}{2} + \limsup_{n\to\infty} a_n \geq a_m \geq -\frac{\varepsilon}{2} + \liminf_{n\to\infty} a_n;$$
$$\frac{\varepsilon}{2} + \limsup_{n\to\infty} b_n \geq b_m \geq -\frac{\varepsilon}{2} + \liminf_{n\to\infty} b_n.$$

Adding the left-hand sided inequalities, we get:

$$a_m + b_m \leq \varepsilon + \limsup_{n\to\infty} a_n + \limsup_{n\to\infty} b_n.$$

We conclude with our first desired inequality,

$$\limsup_{n\to\infty}(a_n+b_n) \leq \limsup_{n\to\infty} a_n + \limsup_{n\to\infty} b_n.$$

To obtain the reverse inequality, again fix $\varepsilon > 0$. Then there exists a sequence $\{k_n\}$ so that

$$b_{k_m} \geq -\frac{\varepsilon}{2} + \limsup_{n\to\infty} b_n \quad \text{for all } m.$$

Chopping off the finitely-many terms in the sequence occurring before the threshold N_ε and applying the above inequalities, we have, for all $m \in \mathbb{N}$:

$$a_{k_m} + b_{k_m} \geq -\frac{\varepsilon}{2} + \liminf_{n\to\infty} a_n - \frac{\varepsilon}{2} + \limsup_{n\to\infty} b_n.$$

We conclude with the desired reverse inequality,

$$\limsup_{n\to\infty}(a_n+b_n) \geq \liminf_{n\to\infty} a_n + \limsup_{n\to\infty} b_n.$$

For the second question, consider the sequences

$$a_n = (-1)^n,\ b_n = (-1)^{n+1}.$$

Thus $a_n + b_n = 0$ for all n, so $\limsup_{n\to\infty}(a_n+b_n) = 0$. However,

$$\limsup_{n\to\infty} a_n = \limsup_{n\to\infty} b_n = 1,$$

which completes the proof. ■

[7] As an aside, if I_-, I_+ are both infinite, then we have

$$\limsup_{n\to\infty} a_n = \infty, \quad \liminf_{n\to\infty} a_n = -\infty,$$

which you can check holds for sequences such as $a_n = (-n)^n$, say.

2.8 Exercises

1. Prepare a 2-page summary of this chapter, with important definitions and results.
2. Is the converse of Theorem 13 true?
3. Let $|q| < 1$. Compute $\lim_{n\to\infty} nq^n$, if the limit exists.
4. Let (x_n) be a decreasing sequence, bounded below. Show that $x_n \to \inf\{x_n \mid n \in \mathbb{N}\}$.
5. Find a divergent sequence with convergent subsequences.
6. Show directly that the sum of two Cauchy sequences is a Cauchy sequence.
7. Show directly that every bounded Cauchy sequence admits at least one convergent subsequence.
8. Complete the induction argument that allows you to compute the limit of the sequence defined by $x_n = \frac{1}{2}(x_{n-2} + x_{n-1})$, $n > 2$, with $x_1 = 1$ and $x_2 = 2$.
9. Show that $(x_n) = \frac{1}{n}$ and $(x_n) = (n)$ are both positive real sequences with $x_n^{1/n} \to 1$, even though one converges and one diverges.
10. Complete the proof of solved problem 21 (do the lim inf case). Consider the sequence given by the recursion $a_{n+1} = \frac{1}{2}(a_n + a_n^{-1})$, with some initial condition $a_1 \in (-\infty, 0) \cup (0, \infty)$. Find and prove the limit, if it exists.

Chapter 3

Limits and Continuity

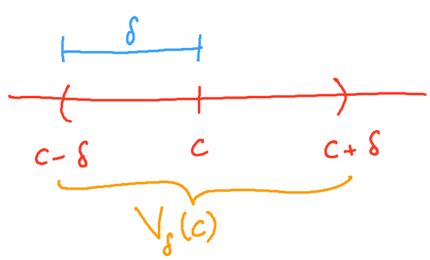
The main objects of study in analysis are **functions**. In this chapter, we introduce the $\varepsilon - \delta$ **definition of the limit** of a function, provide results that help to compute such limits, identify two types of **continuity**, and present some of the theorems that form the basis of analytical endeavours.

3.1 Limit of a Function

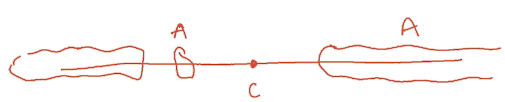
The objects we have studied thus far are functions of $\mathbb{N}$ into $\mathbb{R}$. However, most of calculus deals with functions of $\mathbb{R}$ into $\mathbb{R}$. How do we generalize the concepts and results we have derived for sequences to functions?

Let $A \subseteq \mathbb{R}$ and $c \in \mathbb{R}$. The **neighbourhood** $V_\delta(c)$, where $\delta > 0$, is the interval

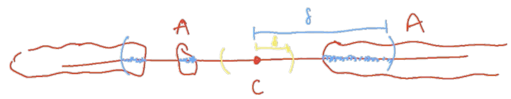
$$V_\delta(c) = \{x \in \mathbb{R} : |x - c| < \delta\} = (c - \delta, c + \delta).$$



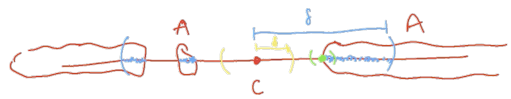
The point $c \in \mathbb{R}$ is a **limit point** (or **cluster point**) of A if every neighbourhood $V_\delta(c)$ contains at least one point $x \in A$ other than c.

> **Example:** consider the set $A \subseteq \mathbb{R}$ drawn below.
>
> 
>

The $V_\delta(c)$−neighbourhood in blue contains points in A other than c, but c is not a limit point of A since the $V_\delta(c)$−neighbourhood in yellow does not contain points of A.



The point at the centre of the green interval is a limit point of A, however.



The **set of all limit points** of A is denoted by $\overline{A}$; a limit point of A does not have to be in A.

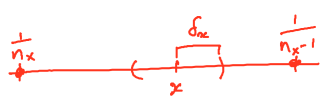
Example: what are the limit points of $A = \{\frac{1}{n} \mid n \in \mathbb{N}\}$?

Solution: let $n \in \mathbb{N}$. The distance between a point $\frac{1}{n}$ and its immediate successor/predecessor $\frac{1}{n\pm1}$ is

$$\frac{1}{n} - \frac{1}{n \pm 1} = \frac{1}{n(n \pm 1)} > \frac{1}{3n^2}.$$

Let $\delta = \frac{1}{3n^2}$. Then $V_\delta(\frac{1}{n}) = (\frac{1}{n} - \frac{1}{3n^2}, \frac{1}{n} + \frac{1}{3n^2}) \subseteq (\frac{1}{n-1}, \frac{1}{n+1})$, so the only point of A in $V_\delta(\frac{1}{n})$ is $\frac{1}{n}$. Thus $\frac{1}{n} \notin \overline{A}$. No negative real number is a limit point of A; indeed, if $x < 0$, set $\delta = \frac{|x|}{2}$. Then $V_\delta(x) \subseteq (-\infty, 0)$ and so contains no point of A. Similarly, no real number strictly greater than 1 is a limit point of A. Hence $\overline{A} \subseteq [0,1] \setminus A$.

Let $x \in (0,1] \setminus A$. By the Archimedean property, $\exists n_x \in \mathbb{N}$ s.t. $n_x > \frac{1}{x} > n_x - 1$, so $\frac{1}{n_x} < x < \frac{1}{n_x-1}$. Set $\delta_x = \frac{1}{2}\min\{|x - \frac{1}{n_x}|, |x - \frac{1}{n_x-1}|\}$. Then $V_{\delta_x}(x)$ contains none of the points of A.



The only remaining possibility is $x = 0$. Let $\delta > 0$. By the Archimedean property, $\exists N_\delta$ such that $\frac{1}{N_\delta} < \delta$. But $0 \neq \frac{1}{N_\delta} \in A$, Thus

$$\varnothing \neq \left\{\frac{1}{N_\delta}\right\} \subseteq V_\delta \cap A = (-\delta, \delta) \cap A,$$

so $x = 0$ is the only limit point of A: $\overline{A} = \{0\}$. $\square$

Directly determining the limit points of a set is a time-intensive endeavour. Thankfully, there is a link between limit points and convergent sequences.

Theorem 24
A point $c \in \mathbb{R}$ is a limit point of A if and only if there is a sequence $(a_n) \subseteq A$, with $a_n \neq c$ for $n \in \mathbb{N}$, such that $a_n \to c$.

Proof: suppose c is a limit point of A. By definition, the neighbourhood $V_{\frac{1}{n}}(c)$ must contain a point $a_n \neq c \in A$, for all $n \in \mathbb{N}$. Let $\varepsilon > 0$. By the Archimedean property, $\exists N_\varepsilon > \frac{1}{\varepsilon}$ s.t. $\frac{1}{N_\varepsilon} < \varepsilon$. Thus

$$n > N_\varepsilon \Longrightarrow 0 < |a_n - c| < \frac{1}{n} < \frac{1}{N_\varepsilon} < \varepsilon, \quad \text{i.e. } a_n \to c.$$

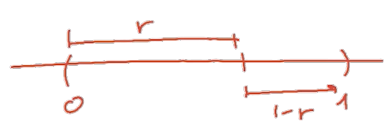
Conversely, suppose that there is a sequence $(a_n) \subseteq A$, with $a_n \neq c$ for all $n \in \mathbb{N}$, such that $a_n \to c$. Let $\delta > 0$. By definition, $\exists N_\delta \in \mathbb{N}$, such that $0 < |a_n - c| < \delta$ for all $n > N_\delta$. Then $a_n \in V_\delta(c)$ and $a_n \neq c$ for all $n > N_\delta$. Thus any neighbourhood of c contains at least one $a_n \neq c$, so $c \in \overline{A}$. ■

Any limit point of A is in fact the limit of a sequence in A, and *vice-versa*.

Example: let $A = [0,1] \cap \mathbb{Q}$. What are the limit points of A?

Solution: any convergent sequence $(a_n) \subseteq A$ is such that $0 \leq a_n \leq 1$ for all $n \in \mathbb{N}$, so its limit must also lie in $[0,1]$, according to Theorem 15. On the other hand, Theorem 24 tells us that any limit point of A is the limit of a sequence of rationals in $[0,1]$. The sequences $(\frac{1}{n})$ and $(1-\frac{1}{n})$ lie in A. Since $\frac{1}{n} \to 0$ and $1 - \frac{1}{n} \to 1$, then $0, 1 \in \overline{A}$.

Now, let $r \in (0,1)$. Set $\eta = \min\{r, 1-r\}$.



Then $\eta > 0$ and $\frac{1}{\eta} > 0$. By the Archimedean property, $\exists M \in \mathbb{N}$ s.t. $M > \frac{1}{\eta}$. Then

$$0 \leq r - \eta < r - \frac{1}{M} > r + \frac{1}{M} < r + \eta \leq 1,$$

since $\eta = r$ if $r \leq 1/2$ and $\eta = 1 - r$ if $r \geq 1/2$. So

$$n > M \Longrightarrow 0 < r - \frac{1}{n} < r + \frac{1}{n} < 1.$$

But the density theorem states that for all $n > M$, $\exists a_n \neq r \in \mathbb{Q}$ such that

$$r - \frac{1}{n} < a_n < r + \frac{1}{n}.$$

The sequence (a_n) thus constructed converges to r. Indeed, let $\varepsilon > 0$. According to the Archimedean property, $\exists N \in \mathbb{N}$ such that $N > \frac{1}{\varepsilon}$.

Set $N_\varepsilon = \max\{M, N\}$. Then

$$n > N_\varepsilon \implies 0 < |a_n - r| < \frac{1}{n} < \frac{1}{N_\varepsilon} < \varepsilon,$$

and so $a_n \to r$ and $r \in \overline{A}$. Consequently, $\overline{A} = [0, 1]$. ■

Intuitively, a limit of a function f at c is a value L towards which $f(x)$ "approaches" as x gets closer to c, if it exists. But what does that actually mean? What would need to happen for the value not to exist?

Let $A \subseteq \mathbb{R}$, $f : A \to \mathbb{R}$, and $c \in \overline{A}$: $L \in \mathbb{R}$ is the **limit of** f **at** c if

$$\forall \varepsilon > 0,\ \exists \delta_\varepsilon > 0 \text{ such that } 0 < |x - c| < \delta_\varepsilon \text{ and } x \in A \implies |f(x) - L| < \varepsilon,$$

which we denote by

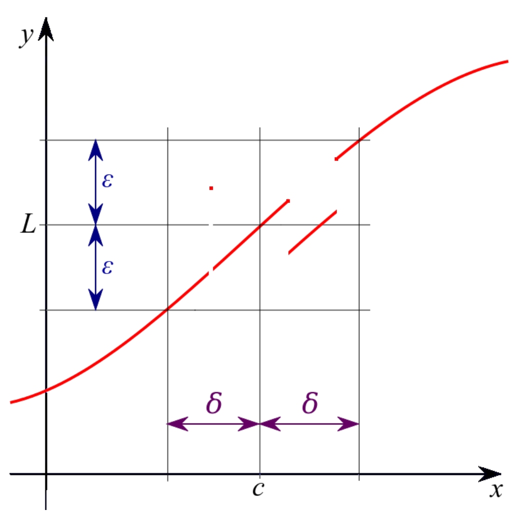
$$\lim_{x \to c} f(x) = L \quad \text{or by} \quad f(x) \to L, \text{ when } x \to c.$$

The limit of f at c is **not** $L \in \mathbb{R}$ if

$$\exists \varepsilon_0 > 0,\ \forall \delta > 0,\ \exists x_\delta \in A \text{ such that } 0 < |x_\delta - c| < \delta_\varepsilon \text{ and } |f(x_\delta) - L| \geq \varepsilon_0,$$

which we denote by

$$\lim_{x \to c} f(x) \neq L \quad \text{or by} \quad f(x) \not\to L, \text{ when } x \to c.$$



The underlying principle is the same as that of the limit of a sequence: given $\varepsilon > 0$, we need to find a $\delta_\varepsilon > 0$ which satisfies the definition. Graphically, this is equivalent to putting a horizontal strip of width 2ε around the line $y = L$, and showing that there is a neighbourhood $V_{\delta_\varepsilon}(c)$ such that $f(x)$ is in the strip for any $x \in V_{\delta_\varepsilon}$.

Examples

1. Let $f : [0,1) \to \mathbb{R}$ be the function defined by

$$f(x) = \begin{cases} 2, & x \in (0,1) \\ 3, & x = 0 \end{cases}$$

Show $\lim\limits_{x\to 0} f(x) = 2$.

Proof: let $\varepsilon > 0$. Set $\delta_\varepsilon = 1$. Then

$$x \in [0,1) \text{ and } 0 < |x - c| < \delta_\varepsilon \Longrightarrow |f(x) = 2| = 0 < 0 \cdot \delta < \varepsilon,$$

which completes the proof. ■

2. Let $f : [0,\infty) \to \mathbb{R}$ be defined by $f(x) = \frac{x^2+2x+2}{x+1}$. Show $\lim\limits_{x\to 2} f(x) = \frac{10}{3}$.

Proof: let $\varepsilon > 0$. Set $\delta_\varepsilon = \varepsilon$. Then

$$\begin{aligned} \left|\frac{x^2+2x+2}{x+1} - \frac{10}{3}\right| &= \left|\frac{3(x^2+2x+2) - 10(x+1)}{x+1}\right| = \left|\frac{3x^2-4x-4}{3x+3}\right| \\ &= \underbrace{\left|\frac{3x+2}{3x+3}\right|}_{<1} |x-2| < |x-2| < \delta_\varepsilon = \varepsilon \end{aligned}$$

when $x \geq 0$ and $0 < |x - 2| < \delta_\varepsilon$. ■

3. Let $f : \mathbb{R} \setminus \{0\} \to \mathbb{R}$, $f(x) = x^2 \cos(1/x)$. Show that $\lim\limits_{x\to 0} f(x) = 0$.

Proof: note that $c \in A = \mathbb{R} \setminus \{0\}$. We can only use the definition of the limit if $c \in \overline{A}$. That it does so is a given, as $(\frac{1}{n}) \subseteq A$ and $\frac{1}{n} \to 0$, with $\frac{1}{n} \neq 0$ for all $n \in \mathbb{N}$, according to Theorem 24.

Let $\varepsilon > 0$ and set $\delta_\varepsilon = \sqrt{\varepsilon}$. Then

$$\left|x^2 \cos(1/x) - 0\right| = |x|^2 \underbrace{|\cos(1/x)|}_{\leq 1} \leq |x|^2 = |x - 0|^2 < \delta_\varepsilon^2 < \varepsilon,$$

whenever $x \in \mathbb{R} \setminus \{0\}$ and $0 < |x - 0| < \delta_\varepsilon$. ■

As is the case with sequences, a function has at most one limit at any of its limit points c.

Theorem 25
Let $A \subseteq \mathbb{R}$, $f : A \to \mathbb{R}$ and c a limit point of A. Then f has at most one limit at c.

Proof: suppose that

$$\lim_{x\to c} f(x) = L' \quad \text{and} \quad \lim_{x\to c} f(x) = L'', \quad \text{where } L' < L''.$$

Let $\varepsilon = \frac{L''-L'}{3} > 0$. By definition, $\exists \delta'_\varepsilon, \delta''_\varepsilon$ s.t. $|f(x) - L'| < \varepsilon$ and $|f(x) - L''| < \varepsilon$ whenever $x \in A$ and $0 < |x - c| < \delta'_\varepsilon$, $0 < |x - c| < \delta''_\varepsilon$.

Set $\delta_\varepsilon = \min\{\delta'_\varepsilon, \delta''_\varepsilon\}$. Then, whenever $x \in A$ and $0 < |x - c| < \delta_\varepsilon$,

$$\begin{aligned} f(x) < L'' + \varepsilon = L' + \frac{L''-L'}{3} &= \frac{2L'+L''}{3} = \frac{L'+L''}{3} + \frac{L'}{3} \\ < \frac{L'+L''}{3} + \frac{L''}{3} < \frac{2L''+L'}{3} &= L'' - \frac{L''-L'}{3} = L'' - \varepsilon < f(x), \end{aligned}$$

which is a contradiction, hence $L' \not< L''$. The proof that $L'' \not< L'$ is identical. ■

As is the case with sequences, the definition is useless if we do not have a candidate for L beforehand. The next result allows us to get such a candidate before using the definition.

Theorem 26 (SEQUENTIAL CRITERION)
Let $A \subseteq \mathbb{R}$, $f : A \to \mathbb{R}$ and c a limit point of A. Then

$$\lim_{x\to c} f(x) = L \quad \textit{if and only if} \quad \lim_{n\to\infty} f(x_n) = L$$

for any sequence $(x_n) \subseteq A$ such that $x_n \to c$, with $x_n \neq c$ for all $n \in \mathbb{N}$.

Proof: assume $\lim\limits_{x\to c} f(x) = L$. Let $\varepsilon > 0$. Then $\exists \delta_\varepsilon > 0$ such that

$$x \in A \text{ and } 0 < |x - c| < \delta_\varepsilon \implies |f(x) - L| < \varepsilon.$$

Suppose $(x_n) \subseteq A$ is such that $x_n \neq c$ for all $n \in \mathbb{N}$ and $x_n \to c$. Then $\exists M_{\delta_\varepsilon} > 0$ such that $0 < |x_n - c| < \delta_\varepsilon$ whenever $n > M_{\delta_\varepsilon}$.

Let $N_\varepsilon = M_{\delta_\varepsilon}$. Then

$$x_n \neq c \in A \text{ and } n > N_\varepsilon \implies 0 < |x_n - c| < \delta_\varepsilon \implies |f(x_n) - L| < \varepsilon,$$

which is to say $f(x_n) \to L$.

Conversely, if $\lim_{x\to c} f(x) \neq L$, then $\exists \varepsilon_0 > 0$ s.t. $\forall \delta > 0$, $\exists x_\delta \in A$ with $0 < |x_\delta - c| < \delta$ but $|f(x) - L| \geq \varepsilon_0$. Thus, for $n \in \mathbb{N}$ and $\delta = \frac{1}{n}$, $\exists x_n = x_\delta$ as above.

The sequence $(x_n) \subseteq A$ is such that $0 < |x_n - c| < \frac{1}{n}$ and $|f(x_n) - L| \geq \varepsilon_0$. According to the squeeze theorem, $x_n \to c$, with $|f(x_n) - L| \geq \varepsilon_0$ for all $n \in \mathbb{N}$. Thus $f(x_n) \not\to L$. ■

Let us take a look at a few examples.

Examples

1. Let $f : \mathbb{R} \to \mathbb{R}$, $f(x) = 3x^3 + x + 1$. Compute $\lim_{x\to 7} f(x)$.

   Solution: let $(x_n) \subseteq \mathbb{R} \setminus \{7\}$ with $x_n \to 7$. Then

   $$\begin{aligned}\lim_{n\to\infty} f(x_n) &= \lim_{n\to\infty} (3x_n^2 + x_n + 1) = 3\left(\lim_{n\to\infty} x_n\right)^2 + \lim_{n\to\infty} x_n + 1 \\ &= 3 \cdot 7^3 + 7 + 1 = 1037.\end{aligned}$$

   Thus $f(x) \to 1037$ when $x \to 7$, according to Theorem 26. □

2. Let $f : (2,\infty) \to \mathbb{R}$, $f(x) = \frac{(x-1)(x-2)}{(x-2)}$. Compute $\lim_{x\to 2} f(x)$.

   Solution: let $(x_n) \subseteq \mathbb{R} \setminus \{2\}$ with $x_n \to 2$. Then

   $$\begin{aligned}\lim_{n\to\infty} f(x_n) &= \lim_{n\to\infty} \frac{(x_n - 1)(x_n - 2)}{(x_n - 2)} = \lim_{n\to\infty} (x_n - 1) = \lim_{n\to\infty} x_n - 1 \\ &= 2 - 1 = 1.\end{aligned}$$

   Since (x_n) was arbitrary, $f(x) \to 1$ when $x \to 2$, according to Theorem 26. □

3. Let $f : \mathbb{R} \setminus \{0\} \to \mathbb{R}$, $f(x) = x^2 \cos(1/x)$. Show that $\lim_{x\to 0} f(x) = 0$.

   Proof: let $(x_n) \subseteq \mathbb{R} \setminus \{0\}$ be any sequence converging to 0. Then

   $$0 \leq |x_n^2 \cos(1/x_n)| \leq |x_n^2| = |x_n|^2.$$

   However, since $x_n \to 0$, then both $|x_n| \to 0$ and $|x_n|^2 \to 0$, which is to say that

   $$\lim_{n\to 0} |x_n^2 \cos(1/x_n)| = 0$$

   according to the squeeze theorem. Thus $x_n^2 \cos(1/x_n) \to 0$. Since (x_n) was arbitrary, $f(x) \to 0$ when $x \to 0$, according to the sequential criterion. ■

4. Let $f : \mathbb{R} \to \mathbb{R}$ be the function defined by

$$f(x) = \begin{cases} 0, & x \in \mathbb{Q} \\ 1, & x \notin \mathbb{Q} \end{cases}$$

Show that $\lim_{x \to 0} f(x)$ does not exist.

Proof: define $(x_n), (y_n)$ by $x_n = \frac{1}{n}$, $y_n = \frac{\sqrt{2}}{n}$ for all $n \in \mathbb{N}$. Then $(x_n) \subseteq \mathbb{Q}$ and $(y_n) \subseteq \mathbb{R} \setminus \mathbb{Q}$. Furthermore, $x_n, y_n \to 0$, with $x_n, y_n \neq 0$ for all $n \in \mathbb{N}$. But $f(x_n) = 0$ and $f(y_n) = 1$ for all $n \in \mathbb{N}$, so

$$\lim_{n \to \infty} f(x_n) = 0 \neq 1 = \lim_{n \to \infty} f(y_n),$$

thus $\lim_{x \to 0} f(x)$ does not exist. ■

5. Let sgn $: \mathbb{R} \to \mathbb{R}$ be the function defined by

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

Show that $\lim_{x \to 0}(x + \text{sgn}(x))$ does not exist.

Proof: define $(x_n), (y_n)$ by $x_n = \frac{1}{n}$, $y_n = -\frac{1}{n}$ for all $n \in \mathbb{N}$. Then $x_n, y_n \to 0$, with $x_n, y_n \neq 0$ for all $n \in \mathbb{N}$.

But $f(x_n) = \frac{1}{n} + \text{sgn}\left(\frac{1}{n}\right) = \frac{1}{n} + 1$, and $f(y_n) = -\frac{1}{n} + \text{sgn}\left(-\frac{1}{n}\right) = -\frac{1}{n} - 1$ for all $n \in \mathbb{N}$, so

$$\lim_{n \to \infty} f(x_n) = \lim_{n \to \infty} \left(\frac{1}{n} + 1\right) \neq -1 = \lim_{n \to \infty} \left(\frac{1}{n} + 1\right) = \lim_{n \to \infty} f(x_n),$$

thus $\lim_{x \to 0} f(x)$ does not exist. ■

To show that the limit does not exist, it is enough to find two specific sequences $(x_n), (y_n) \subseteq A$, with $x_n, y_n \neq c$ for all $n \in \mathbb{N}$ and $x_n, y_n \to c$, such that $f(x_n) \to L_1$, $f(y_n) \to L_2$, $L_1 \neq L_2$.

But we cannot show that the limit L exists by finding two sequences $(x_n), (y_n) \subseteq A$ with $x_n, y_n \neq c$ for all $n \in \mathbb{N}$, $x_n, y_n \to c$, and $f(x_n), f(y_n) \to L$.

Note that at no point have we needed to use the graph of a function to compute a limit or prove its existence.

3.2 Properties of Limits

Limits behave quite nicely with respect to the usual operations.

Theorem 27 (Operations on Limits)
Let $A \subseteq \mathbb{R}$, $f, g : A \to \mathbb{R}$, and c a limit point of A. Suppose $f(x) \to L$ and $g(x) \to M$ when $x \to c$. Then

1. $\lim_{x\to c} |f(x)| = |L|$;
2. $\lim_{x\to c}(f(x) + g(x)) = L + M$;
3. $\lim_{x\to c} f(x)g(x) = LM$;
4. $\lim_{x\to c} \frac{f(x)}{g(x)} = \frac{L}{M}$, *if* $g(x) \neq 0$ *for all* $x \in A$ *and if* $M \neq 0$.

Proof: this result is an easy consequence of Theorems 14 and 26. Let $(x_n) \subseteq A$ with $x_n \neq c$ and $x_n \to c$ for all $n \in \mathbb{N}$. Then $f(x_n) \to L$ and $g(x_n) \to M$.

1. $\lim_{x\to c} |f(x)| = \lim_{n\to\infty} |f(x_n)| = \left|\lim_{n\to\infty} f(x_n)\right| = L$.
2. $\lim_{x\to c}\big[f(x) + g(x)\big] = \lim_{n\to\infty} \big[f(x_n) + g(x_n)\big] = \lim_{n\to\infty} f(x_n) + \lim_{n\to\infty} g(x_n) = L + M$.
3. $\lim_{x\to c}\big[f(x)g(x)\big] = \lim_{n\to\infty} \big[f(x_n)g(x_n)\big] = \lim_{n\to\infty} f(x_n) \cdot \lim_{n\to\infty} g(x_n) = LM$.
4. $\lim_{x\to c} \left[\frac{f(x)}{g(x)}\right] = \lim_{n\to\infty} \left[\frac{f(x_n)}{g(x_n)}\right] = \frac{\lim_{n\to\infty} f(x_n)}{\lim_{n\to\infty} g(x_n)} = \frac{L}{M}$, if $g(x) \neq 0$ for $x \in A$ and if $M = 0$. ■

There is also a squeeze theorem for functions, but it is not nearly as useful as the corresponding result for sequences.

Theorem 28 (squeeze theorem for Functions)
Let $A \subseteq \mathbb{R}$, $f, g, h : A \to \mathbb{R}$, and c a limit point of A. If $f(x) \leq g(x) \leq h(x)$ for all $x \in A$ and if $f(x), h(x) \to L$ when $x \to c$, then $g(x) \to L$ when $x \to c$.

Proof: let $(x_n) \subseteq A$, with $x_n \neq c$ for all $n \in \mathbb{N}$ and $x_n \to c$. According to the sequential criterion,

$$\lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} h(x_n) = L.$$

Since $f(x_n) \leq g(x_n) \leq h(x_n)$ for all $n \in \mathbb{N}$, then $\lim_{n\to\infty} g(x_n) = L$, by the squeeze theorem (for sequences). Since (x_n) was arbitrary, we conclude that $g(x) \to L$, again by the sequential criterion. ■

Let's take a look at some examples.

Examples

1. Let $f:\mathbb{R}\to\mathbb{R}$, $f(x)=k$, $k\in\mathbb{R}$. Show that $\lim\limits_{x\to c} f(x)=k$ for all $c\in\mathbb{R}$.

   Proof: let $\varepsilon>0$. Set $\delta_\varepsilon=\varepsilon$. Then $|f(x)-k|=|k-k|=0<\varepsilon$, when $0<|x-c|<\delta_\varepsilon=\varepsilon$. ■

2. Let $f:\mathbb{R}\to\mathbb{R}$, $f(x)=x$. Show that $\lim\limits_{x\to c} f(x)=f(c)$ for all $c\in\mathbb{R}$.

   Proof: let $\varepsilon>0$. Set $\delta_\varepsilon=\varepsilon$. Then $|f(x)-c|=|x-c|<\delta_\varepsilon=\varepsilon$, when $0<|x-c|<\delta_\varepsilon=\varepsilon$. ■

3. Let $f:\mathbb{R}\to\mathbb{R}$, $f(x)=\frac{x^3+2x-4}{x^2+1}$. Compute $\lim\limits_{x\to 3} f(x)$.

   Solution: according to Theorem 27, and the preceding examples,

   $$\begin{aligned}\lim_{x\to 3}(x^3+2x+4)&=\left(\lim_{x\to 3}x\right)^3+2\left(\lim_{x\to 3}x\right)+\lim_{x\to 3}4=3^2+2(3)+3=37\\ \lim_{x\to 3}(x^2+1)&=\left(\lim_{x\to 3}x\right)^2+1=3^2+1=10,\end{aligned}$$

   and so $\lim\limits_{x\to 3}\dfrac{x^3+2x-4}{x^2+1}=\frac{10}{3}$, **because** $x^2+1\neq 0$ for all $x\in\mathbb{R}$. □

4. Let $f:\mathbb{R}\setminus\{0\}\to\mathbb{R}$, $f(x)=x^2\cos(1/x)$. Show that $\lim\limits_{x\to 0} f(x)=0$.

   Proof: we cannot use the multiplication component of Theorem 27 to compute the limit since $\lim\limits_{x\to 0}\cos(1/x)$ does not exist.

   Indeed, let $(x_n),(y_n)\subseteq\mathbb{R}\setminus\{0\}$ be such that $x_n=\frac{1}{(2n-1)\pi}$, and $y_n=\frac{1}{2n\pi}$ for all $n\in\mathbb{N}$. Then $x_n,y_n\to 0$. But

   $$\cos\left(\frac{1}{x_n}\right)=\cos((2n-1)\pi)=-1\quad\text{and}\quad\cos\left(\frac{1}{y_n}\right)=\cos(2n\pi)=1$$

   for all $n\in\mathbb{N}$. Then

   $$\cos(1/x_n)\to -1\neq 1\leftarrow\cos(1/y_n).$$

   This does not mean that

   $$\lim_{x\to 0}x^2\cos\left(\frac{1}{x}\right)$$

   does not exist, only that we cannot use Theorem 27 to compute it.

   In fact, the squeeze theorem for functions does the trick, with $-x^2\leq f(x)\leq x^2$. ■

Other sequence concepts have analogous definitions in the world of functions. Let $A \subseteq \mathbb{R}$, $f : A \to \mathbb{R}$ and $c \in \overline{A}$. The function f is **bounded on some neighbourhood of** c if $\exists \delta > 0$ and $M > 0$ are such that $|f(x)| \leq M$ for all $x \in A \cap V_\delta(c)$.

Theorem 29
If $A \subseteq \mathbb{R}$, $f : A \to \mathbb{R}$, $c \in \overline{A}$, and $\lim\limits_{x \to c} f(x) = L$ for some $L \in \mathbb{R}$, then f is bounded on some neighbourhood of c.

Proof: Let $\varepsilon = 1$. By definition, $\exists \delta_1 > 0$ such that $|f(x) - L| < 1$ whenever $x \in A$ and $0 < |x - c| < \delta_1$. Since

$$|f(x)| - |L| < |f(x) - L|,$$

then $|f(x)| - |L| \leq 1$ whenever $x \in A$ and $0 < |x - c| < \delta_1$.

If $c \notin A$, set $M = |L| + 1$. If $c \in A$, set $M = \max\{|f(c)|, |L| + 1\}$. In either case, $|f(x)| \leq M$ whenever $x \in A$ and $0 < |x - c| < \delta_1$. ■

3.3 Continuous Functions

Functions like polynomials, or trigonometric functions, are continuous, which is a fundamental notion of calculus.

Intuitively, a function is continuous at a point if the graph of the function at that point can be traced without lifting the pen. The notion of "continuity" is fundamental is calculus.

But we emphasized earlier that limits could be computed/shown to exist without referring to the graph of a function. What does that mean for continuity?

Let $A \subseteq \mathbb{R}$, $f : A \to \mathbb{R}$, and $c \in A$; f **is continuous at** c if

$$\forall \varepsilon > 0, \exists \delta_\varepsilon > 0 \text{ such that } |x - c| < \delta_\varepsilon \text{ and } x \in A \Longrightarrow |f(x) - f(c)| < \varepsilon.$$

When computing the limit of f at c, we are interested in the behaviour of the function near c, but not at c. When we are dealing with continuity, we also **include the behaviour at** c. When c is a **limit point of** A, this definition actually means that

$$\lim_{x \to c} f(x) = f(c).$$

If $c \notin \overline{A}$, the expression $\lim\limits_{x \to c} f(x)$ is meaningless.[1] In that case, f is automatically continuous at c. Indeed, there will then be a $\delta > 0$ such that $V_\delta(c)$ contains no point of A but c. Then for $\varepsilon > 0$, whenever $x \in A$ and $|x - c| < \delta$ (i.e., whenever $x = c$), we have

$$|f(x) - f(c)| = |f(c) - f(c)| = 0 < \varepsilon.$$

[1]Since there are no sequence $(x_n) \subseteq A$ with $x_n \neq c$ for all $n \in \mathbb{N}$ and $x_n \to c$.

The definition contains 3 statements: a function f is continuous at c if

1. $f(c)$ is defined;
2. $\lim\limits_{x\to c} f(x)$ exists, and
3. $\lim\limits_{x\to c} f(x) = f(c)$.

Let $B \subseteq A$. If f is continuous for all $c \in B$, then we say that f **is continuous on** B.

Examples

- Let $f : [0,\infty) \to \mathbb{R}$, $f(x) = \frac{x^2+2x+2}{x+1}$. Is f continuous at $c = 2$?

  Solution: since 2 is a limit point of $[0,\infty)$, we need only verify if $\lim\limits_{x\to 2} f(x) = f(2)$. But we have already seen that $f(x) \to \frac{10}{3} = f(2)$ when $x \to 2$, so f is continuous at $c = 2$. □

- Let $f : [0,1) \to \mathbb{R}$,

  $$f(x) = \begin{cases} 2, & x \in (0,1) \\ 3, & x = 0 \end{cases}$$

  Is f continuous at $c = 0$?

  Solution: since 0 is a limit point of $[0,1)$, we need only verify if $\lim\limits_{x\to 0} f(x) = f(0)$. But we have already seen that $f(x) \to 2 \neq 3 = f(0)$ when $x \to 0$, so f is not continuous at $c = 0$. □

- Let $f : \mathbb{R} \to \mathbb{R}$, $f(x) = 3x^3 + x + 1$. Is f continuous at $c = 7$?

  Solution: since 7 is a limit point of $\mathbb{R}$, we need only verify if $\lim\limits_{x\to 7} f(x) = f(7)$. But we have already seen that $f(x) \to 1037 = f(7)$ when $x \to 7$, so f is continuous at $c = 7$. □

- Let $f : \mathbb{R} \to \mathbb{R}$,

  $$f(x) = \begin{cases} 0, & x \in \mathbb{Q} \\ 1, & x \notin \mathbb{Q} \end{cases}$$

  Is f continuous at $c = 0$?

  Solution: as $f(0) = 0$, we only need to verify if $\lim\limits_{x\to 0} f(x) = f(0)$. But we have already seen that $\lim\limits_{x\to 0} f(x)$ does not exist, so f is not continuous at $c = 0$. ■

- Let $f:(2,\infty)\to\mathbb{R}$, $f(x)=\frac{(x-1)(x-2)}{(x-2)}$. Is f continuous at $c=2$?

  Solution: since f is not defined at $c=2$ and since $2\notin A$, f is not continuous at $c=2$. $\square$

- Let $f:\mathbb{R}\to\mathbb{R}$, $f(x)=k$, $k\in\mathbb{R}$. Is f continuous on $\mathbb{R}$?

  Solution: since all $c\in\mathbb{R}$ are limit points of $\mathbb{R}$, we need only verify if $\lim_{x\to c} f(x)=f(c)$. But we have already seen that $f(x)\to k=f(c)$ for all $c\in\mathbb{R}$, so f is continuous on $\mathbb{R}$. $\square$

- Let $f:[0,\infty)\to\mathbb{R}$, $f(x)=\sqrt{x}$. Is f continuous on $[0,\infty)$?

  Solution: let $\varepsilon>0$. If $c=0$, set $\delta_\varepsilon=\varepsilon$. Then

  $$x\geq 0 \text{ and } |x-0|<\delta_\varepsilon \Longrightarrow f(x)-f(0)|=\sqrt{x}=\sqrt{|x-0|}<\sqrt{\delta_\varepsilon=\varepsilon},$$

  so f is continuous at $c=0$. If $c>0$, set $\delta_\varepsilon=\sqrt{c}\varepsilon$. Then

  $$|f(x)-f(c)|=|\sqrt{x}-\sqrt{c}|=\frac{|x-c|}{\sqrt{x}+\sqrt{c}}<\frac{|x-c|}{\sqrt{c}}<\frac{\delta_\varepsilon}{\sqrt{c}}=\varepsilon$$

  whenever $x\geq 0$ and $|x-c|<\delta_\varepsilon$. Hence f is continuous at any $c>0$. $\square$

- Let $f:\mathbb{R}\to\mathbb{R}$,

  $$f(x)=\begin{cases} x, & x\in\mathbb{Q} \\ 0, & x\notin\mathbb{Q}\end{cases}$$

  Is f continuous at $c=0$? At $c\neq 0$?

  Solution: since $f(0)=0$, we need to see if $\lim_{x\to 0} f(x)=0$. Let $\varepsilon>0$ and set $\delta_\varepsilon>0$. Then $|x-0|<\delta_\varepsilon \Longrightarrow |f(x)-f(0)|=|f(x)|\leq |x|=|x-0|<\delta_\varepsilon=\varepsilon$, so f is continuous at $c=0$. Now let $n\in\mathbb{N}$. According to the density theorem, $\exists x_n\in\mathbb{Q}$, $y_n\notin\mathbb{Q}$ such that

  $$c<x_n+c+\frac{1}{n} \quad \text{and} \quad c<y_n<c+\frac{1}{n}.$$

  According to the sequence squeeze theorem, $x_n,y_n\to c$. But $f(x_n)=x_n$ and $f(y_n)=0$ for all $n\in\mathbb{N}$, so

  $$\lim_{n\to\infty} f(x_n)=\lim_{n\to\infty} x_n=c \quad \text{and} \quad \lim_{n\to\infty} f(y_n)=\lim_{n\to\infty} 0=0.$$

  Since $c\neq 0$, these limits are different, and so $\lim_{x\to c} f(x)$ does not exist, according to the sequential criterion. $\square$

- Let $A = \{x \in \mathbb{R} \mid x > 0\}$. Consider the function $f : A \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} 0 & \text{if } x \notin \mathbb{Q} \\ \frac{1}{n} & \text{if } x = \frac{m}{n} \in \mathbb{Q}, \text{ with } \gcd(m,n) = 1 \end{cases}$$

Where is f is continuous?

Solution: we consider two types of limit points of A: $a \in \mathbb{Q}$ and $b \notin \mathbb{Q}$. If $0 < a \in \mathbb{Q}$, let $(x_n) \subseteq A \cap \mathbb{Q}^\complement$ be such that $x_n \to a$. Then $f(x_n) \to 0$. But $f(a) > 0$, so $f(x) \not\to f(a)$ when $x \to a$, according to the sequential criterion.

If $0 < b \notin \mathbb{Q}$, let $\varepsilon > 0$. By the Archimedean property, there exists an integer $N_0 > \frac{1}{\varepsilon}$. There can only be a finite set of rationals with denominator $< N_0$ in the interval $(b-1, b+1)$. Indeed, if $n < N_0$ and $\frac{m}{n} \in (b-1, b+1)$ then whenever $|k| > 2n$, we have:

$$\left| \frac{m+k}{n} - \frac{m}{n} \right| = \frac{|k|}{n} > 2 \Longrightarrow \frac{m+k}{n} \notin (b-1, b+1).$$

Consequently, $\exists \delta > 0$ such that there are no rational number $\frac{m}{n}$ with denominator $< N_0$ in $(b-\delta, b+\delta)$, which is to say that for all $x \in (b-\delta, b+\delta)$, either $f(x) = 0$ (when x is irrational) or $f(x) = \frac{1}{n} \leq \frac{1}{N_0}$ (when x is rational).

Thus, if $|x - b| < \delta$ and $x \in A$, we have

$$|f(x) - f(b)| = |f(x) - 0| = |f(x)| \leq \frac{1}{N_0} < \varepsilon,$$

so $f(x) \to f(b)$ when $x \to b$, i.e., f is only continuous on $A \cap (\mathbb{R} \setminus \mathbb{Q})$. □

Continuity behaves very nicely with respect to elementary operations on functions.

Theorem 30 (Operations on Continuous Functions)
Let $A \subseteq \mathbb{R}$, $f, g : A \to \mathbb{R}$, and $c \in A$. If f, g are continuous at c, then

1. *$|f|$ is continuous at c;*
2. *$f + g$ is continuous at c;*
3. *fg is continuous at c;*
4. *$\frac{f}{g}$ is continuous at c if $g \neq 0$ on A.*

Proof: if $c \notin \overline{A}$, there is nothing to prove. If $c \in \overline{A}$, then

$$\lim_{x\to c} f(x) = f(c) \quad \text{and} \quad \lim_{x\to c} g(x) = g(c).$$

We can then apply Theorem 27 directly with $L = f(c)$ and $M = g(c)$. ■

Since constants and the identity function are continuous on $\mathbb{R}$ (as we saw in the preceding examples), so are polynomial functions. Furthermore, rational functions are continuous on their domain.

The **composition** of the functions $f : A \to B$ and $g : B \to C$ is the function $g \circ f : A \to C$, with $(g \circ f)(x) = g(f(x))$ for all $x \in A$.

Theorem 31 (Composition of Continuous Functions)
Let $A, B \subseteq \mathbb{R}$, $f : A \to \mathbb{R}$, $g : B \to \mathbb{R}$, $c \in A$. If f is continuous at c, g is continuous at $f(c)$, and $f(A) \subseteq B$, then $g \circ f : A \to B$ is continuous at c.

Proof: let $\varepsilon > 0$. As g is continuous at $f(c)$, $\exists \delta_\varepsilon > 0$ such that

$$y \in B \text{ and } |y - f(c)| < \delta_\varepsilon \Longrightarrow |g(y) - g(f(c))| < \varepsilon.$$

Since f is continuous at c, $\exists \eta_{\delta_\varepsilon} = \eta_\varepsilon > 0$ such that

$$x \in A \text{ and } |x - c| < \eta_{\delta_\varepsilon} \Longrightarrow |f(x) - f(c)| < \delta_\varepsilon \Longrightarrow$$

$$x \in A \text{ and } |x - c| < \eta_\varepsilon \Longrightarrow |(g \circ f)(x) - (g \circ f)(c)| = |g(f(x)) - g(f(c))| < \varepsilon,$$

which completes the proof. ■

It is not too difficult to see that Theorems 30 and 31 remain valid if we replace "continuous at c" with "continuous at A".

Example: let $f : [0, \infty) \to \mathbb{R}$, defined by $f(x) = \sqrt{3x^3 + x + 1}$. Show that f is continuous on $[0, \infty)$.

Proof: we can write $f = g \circ h$, where $g : [0, \infty) \to \mathbb{R}$, $g(y) = \sqrt{y}$ and $h : \mathbb{R} \to \mathbb{R}$, $h(x) = 3x^2 + x + 1$. Since g and h are both continuous on their domains and $h(\mathbb{R}) \subseteq [0, \infty)$, g is continuous on $[0, \infty)$, according to Theorem 31. ■

An **algebraic** function is a function obtained via the (possibly repeated) composition of rational functions and root functions. The class of algebraic functions is continuous on its domain. The same goes for trigonometric, exponential, and logarithmic functions, *via* their power series definition.

3.4 Max/Min Theorem

We begin our study of the classical theorems of calculus. Let $A \subseteq \mathbb{R}$, $f : A \to \mathbb{R}$. The function $f : A \to \mathbb{R}$ is **bounded** on A if $\exists M > 0$ such that $|f(x)| < M$ for all $x \in A$.

Examples

1. $f : [0,1] \to \mathbb{R}$, $f(x) = x^2$, is bounded on $[0,1]$ as $|f(x)| < 2, \forall x \in [0,1]$.
2. $g : \mathbb{R} \to \mathbb{R}$, $g(x) = x^2$, is not bounded on $\mathbb{R}$ Indeed, suppose $\exists M > 0$ such that $|f(x)| < M$ for all $x \in \mathbb{R}$. Then $|x^2| = |x|^2 < M$ for all $x \in \mathbb{R}$, i.e. $|x| < \sqrt{M}$ for all $x \in \mathbb{R} \implies M$ is an upper bound of $\mathbb{R}$. But there is no such bound, $\therefore g$ is not bounded on $\mathbb{R}$.
3. $f : (0,1) \to \mathbb{R}$, $f(x) = \frac{1}{x}$, is not bounded on $(0,1]$, but it is bounded on $[a,1]$ for all $a \in (0,1]$. $\square$

There is a link between continuity and boundedness.

Theorem 32
If $f : [a,b] \to \mathbb{R}$ is continuous on $[a,b]$, then f is bounded on $[a,b]$.

Proof: suppose f is not bounded on $[a,b]$. Hence, for all $n \in \mathbb{N}$, $\exists x_n \in [a,b]$ such that $|f(x_n)| > n$. However, $(x_n) \subseteq [a,b]$ so that (x_n) is bounded.

According to Bolzano-Weierstrass, $\exists (x_{n_k}) \subseteq (x_n)$ such that $x_{n_k} \to \hat{x} \in [a,b]$, since

$$a \leq x_{n_k} \leq b \quad \text{for all } k.$$

Since f is continuous, we have

$$f(\hat{x}) = \lim_{x \to \hat{x}} f(x) = \lim_{k \to \infty} f(x_{n_k}),$$

so $(f(x_{n_k}))$ is bounded, being a convergent sequence. But this contradicts the assumption that $|f(x_{n_k})| > n_k \geq k$ for all k. Hence f is bounded on $[a,b]$. $\blacksquare$

Continuous functions on closed, bounded sets have a useful property. Let $A \subseteq \mathbb{R}$, $f : A \to \mathbb{R}$. We say that f **reaches a global maximum on** A if $\exists x^* \in A$ such that $f(x^*) \geq f(x)$ for all $x \in A$. Similarly, f **reaches a global minimum on** A if $\exists x_* \in A$ such that $f(x_*) \leq f(x)$ for all $x \in A$.

Theorem 33 (Max/Min Theorem)
If $f : [a,b] \to \mathbb{R}$ is continuous, then f reaches a global maximum and a global minimum of $[a,b]$.

Proof: let $f([a,b]) = \{f(x) \mid x \in [a,b]\}$. According to Theorem 32, $f([a,b])$ is bounded as f is continous, and so, by completeness of $\mathbb{R}$,

$$s^* = \sup\{f(x) \mid x \in [a,b]\} \quad \text{and} \quad s_* = \inf\{f(x) \mid x \in [a,b]\}$$

both exist. We need only show $\exists x^*, x_* \in [a,b]$ such that $f(x^*) = s^*$ and $f(x_*) = s_*$.

Since $s^* - \frac{1}{n}$ is not an upper bound of $f([a,b])$ for every $n \in \mathbb{N}$, $\exists x_n \in [a,b]$ with

$$s^* - \frac{1}{n} < f(x_n) \leq s^*, \quad \text{for all } n \in \mathbb{N}.$$

According to the squeeze theorem, we must have $f(x_n) \to s^*$ (this says nothing about whether x_n converges or not, however).

But $(x_n) \subseteq [a,b]$ is bounded, so applying the Bolzano-Weierstrass theorem, we find that $\exists (x_{n_k}) \subseteq (x_n)$ such that $x_{n_k} \to x^* \in [a,b]$. As f is continuous,

$$s^* = \lim_{k\to\infty} f(x_{n_k}) = f\Big(\lim_{k\to\infty} x_{n_k}\Big) = f(x^*).$$

The existence of $x_* \in [a,b]$ such that $f(x_*) = s_*$ is shown similarly. ■

Let's take a look at some examples.

Examples

1. The function $f : [0,1] \to \mathbb{R}$, $f(x) = x^2$, reaches its maximum and minimum on $[0,1]$ since f is continuous, being a polynomial.

2. Let $f : [0,1) \to \mathbb{R}$ be the function defined by

$$f(x) = \begin{cases} 2, & x \in (0,1) \\ 3, & x = 0 \end{cases}$$

   The function f is not continuous on $[0,1)$, and $[0,1)$ is not closed and bounded, so we cannot use the max/min theorem to conclude that f reaches its global max/min on $[0,1)$... even though it does: 3 at $x^* = 0$ and 2 at any $x_* \in (0,1)$.[2]

3. The function $f : [a,1] \to \mathbb{R}$, $a \in (0,1]$, defined by $f(x) = \frac{1}{x}$ reaches its global max/global min on $[a,1]$ as f is continuous on $[a,1]$, being rational there.

4. The function $f : (0,1] \to \mathbb{R}$ defined by $f(x) = \frac{1}{x}$ is continuous on $(0,1]$, but we cannot use the max/min theorem as $(0,1]$ is not closed. In this case, f has no global maximum, but it does have a global minimum at $x = 1$. □

3.5 Intermediate Value Theorem

The following result has many applications; notably it can help locate the roots of a function.

Theorem 34
Let $f : [a, b] \to \mathbb{R}$ be continuous. If $\exists \alpha, \beta \in [a, b]$ such that $f(\alpha)f(\beta) < 0$, then $\exists \gamma \in (a, b)$ such that $f(\gamma) = 0$.

Proof: we prove the result for $f(\alpha) < 0 < f(\beta)$; the other case is similar. Write $\alpha_1 = \alpha$, $\beta_1 = \beta$, $I_1 = [\alpha_1, \beta_1]$, and $\gamma_1 = \frac{\alpha_1+\beta_1}{2}$. There are 3 possibilities:

i. if $f(\gamma_1) = 0$, set $\gamma = \gamma_1$; then $\gamma \in (\alpha_1, \beta_1)$ and the theorem is proven;
ii. if $f(\gamma_1) > 0$, set $\alpha_2 = \alpha_1$, $\beta_2 = \gamma_1$;
iii. if $f(\gamma_1) < 0$, set $\alpha_2 = \gamma_1$, $\beta_2 = \beta_1$.

In the last two cases, set $I_2 = [\alpha_2, \beta_2]$. Then $I_1 \supseteq I_2$, length$(I_1) = \frac{\beta_1-\alpha_1}{2^0}$ and

$$f(\alpha_2) < 0 < f(\beta_2).$$

This is the base case $n = 1$ of an induction process, which can be extended for all $n \in \mathbb{N}$. Either one of two things can occur:

1. $\exists n \in \mathbb{N}$ such that $f(\gamma_n) = 0$, with $\gamma_n \in (\alpha_n, \beta_n) \subseteq (\alpha, \beta)$, in which case the theorem is proven, or

2. there is a chain of nested intervals

$$I_1 \supseteq I_2 \supseteq \cdots I_k \supseteq I_{k+1} \supseteq \cdots$$

where $I_n = [\alpha_n, \beta_n]$, length$(I_n) = \frac{\beta_n-\alpha_n}{2^{n-1}}$, $f(\alpha_n) < 0 < f(\beta_n)$ $\forall n \in \mathbb{N}$.

According to the nested intervals theorem, since

$$\inf_{n\in\mathbb{N}}\{\text{length}(I_n)\} = \lim_{n\to\infty} \frac{\beta_n - \alpha_n}{2^{n-1}} = 0,$$

$\exists c \in [\alpha, \beta] \subseteq [a, b]$ such that $\bigcap_{n\in\mathbb{N}} I_n = \{c\}$.

It remains to show that $f(c) = 0$. Note that the sequences $(\alpha_n), (\beta_n)$ both converge to c. Indeed, let $\varepsilon > 0$. By the Archimedean property, $\exists N_\varepsilon \in \mathbb{N}$ such that $N_\varepsilon > \log_2(\frac{\beta-\alpha}{\varepsilon}) + 1$.

Since $c \in I_n$ for all $n \in \mathbb{N}$, then $|\alpha_n - c| <$ length$(I_n) = \frac{\beta-\alpha}{2^{n-1}} < \varepsilon$ whenever $n > N_\varepsilon$. The proof that $\beta_n \to c$ is identical.

Since f is continuous on $[a,b]$, it is also continuous at c. Thus,

$$\lim_{n\to\infty} f(\alpha_n) = \lim_{n\to\infty} f(\beta_n) = f(c).$$

But $f(\alpha_n) < 0$ for all n, so, Theorem 15:

$$f(c) = \lim_{n\to\infty} f(\alpha_n) \leq 0.$$

Using the same Theorem, we have $f(c) \geq 0$. Then $f(c) = 0$. Lastly, note that $c \neq \alpha, \beta$; otherwise, $f(\alpha)f(\beta) = 0$.

This concludes the proof, with $\gamma = c$. ■

We can use the result to revisit a corollary from Chapter 1.

Example: Show that $\exists x \in \mathbb{R}^+$ such that $x^2 = 2$.

Proof: the function $f : [0,2] \to \mathbb{R}$ defined by $f(x) = x^2 - 2$ is continuous on $[0,2]$. As $f(0) = 0^2 - 2 = -2 < 0$ and $f(2) = 2^2 - 2 = 2 > 0$, $\exists \gamma \in (0,2)$ such that $\gamma^2 - 2 = 0$, so $\gamma^2 = 2$, according to Theorem 34. ■

This result easily generalizes to the following.

Theorem 35 (INTERMEDIATE VALUE THEOREM)
Let $f : [a,b] \to \mathbb{R}$ be continuous. If $\exists \alpha < \beta \in [a,b]$ s.t. $f(\alpha) < k < f(\beta)$ or $f(\alpha) > k > f(\beta)$, then $\exists \gamma \in (a,b)$ such that $f(\gamma) = k$.

Proof: assume that $f(\alpha) < k < f(\beta)$; the proof for the other case is similar. Consider the function $g : [a,b] \to \mathbb{R}$ defined by $g(x) = f(x) - k$. Theorem 30 shows that g is continuous on $[a,b]$. Furthermore,

$$g(\alpha) = f(\alpha) - k < k - k = 0 < f(\beta) - k = g(\beta).$$

By Theorem 34, $\exists \gamma \in (\alpha, \beta)$ such that $g(\gamma) = f(\gamma) - k = 0$. Thus $f(\gamma) = k$. ■

The following result combines the max/min and the intermediate value theorems.

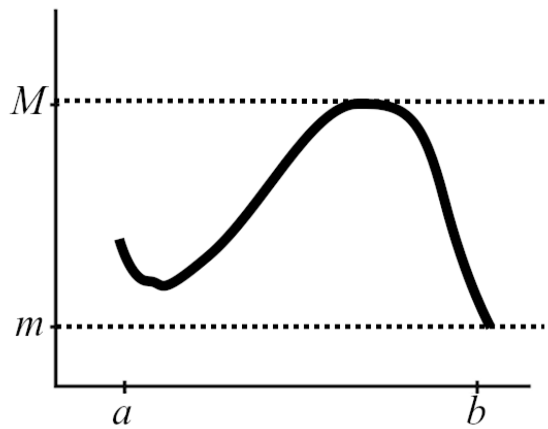
Theorem 36
If $f : [a,b] \to \mathbb{R}$ is continuous, then $f([a,b])$ is a closed and bounded interval.

Proof: Let $m = \inf\{f[a,b]\}$ and $M = \sup\{f[a,b]\}$. According to the max/min theorem, $\exists \alpha, \beta \in [a,b]$ such that $f(\alpha) = m$ and $f(\beta) = M$.

If $m = M$, then f is constant and $f([a,b]) = [m,m] = [M,M]$. If $m < M$, then $\alpha \neq \beta$. Furthermore, $m \leq f(x) \leq M$ for all $x \in [a,b]$, so that $f([a,b]) \subseteq [m,M]$.

Now, let $k \in [m,M]$. According to the intermediate value theorem, $\exists \gamma$ between α and β such that $f(\gamma) = k$. Hence $k \in f([a,b])$ and so $[m,M] \subseteq f([a,b])$. Consequently, $f([a,b]) = [m,M]$. ■

The image of any interval by a continuous function is always an interval, but the only time that we know for a fact that image is of the same type as the original is when the original is closed and bounded.



Examples

1. Let $f : [0,1] \to \mathbb{R}$, $f(x) = 2x - 1$. Then $f([0,1])$ is closed and bounded (in fact, $f([0,1]) = [-1,1]$, but the endpoints of $f([-1,1])$ are not provided by Theorem 36).

2. The function $f : (0, 2\pi) \to \mathbb{R}$ defined by $f(x) = \sin x$ is continuous and $f((0,2\pi)) = [-1,1]$, but Theorem 36 does not apply.

3.6 Uniform Continuity

If $f : A \to \mathbb{R}$ is continuous (on A), then for $\varepsilon > 0$ and $c \in A$, the $\delta_\varepsilon > 0$ that is used to show continuity of f at c generally depends on ε **and** on c. But there might be instances when δ_ε depends only on ε.

The function f is **uniformly continuous** on A if

$$x, y \in A \text{ and } |x - y| < \delta_\varepsilon \Longrightarrow |f(x) - f(y)| < \varepsilon.$$

The notion of uniform continuity is more restrictive than that of (simple) continuity.

Theorem 37
If $f : A \to \mathbb{R}$ is uniformly continuous on A, then f is continuous on A.

Proof: let $c \in A$ and $\varepsilon > 0$. As f is uniformly continuous on A, $\exists \delta_\varepsilon > 0$ such that
$$|f(x) - f(y)| < \varepsilon \quad \text{whenever } |x - y| < \delta_\varepsilon \text{ and } x, y \in A.$$
In particular, if $y = c$ then
$$|f(x) - f(c)| < \varepsilon \quad \text{whenever } |x - c| < \delta_\varepsilon \text{ and } x \in A.$$
As c is arbitrary, f is continuous on A. ■

The converse of Theorem 37 is false, as the following example shows.

Example: show that $f : (0, \infty) \to \mathbb{R}$ defined by $f(x) = \frac{1}{x}$ is continuous on $(0, \infty)$ but not uniformly continuous on $(0, \infty)$.

Proof: that f is continuous on $(0, \infty)$ is immediate, as it is a rational function. Let $(x_n) = (\frac{1}{n}) \subseteq (0, \infty)$. Clearly, (x_n) is a Cauchy sequence as it is a convergent sequence. But $f(x_n) = \frac{1}{1/n} = n$ for all $n \in \mathbb{N}$, so $(f(x_n))$ is not a Cauchy sequence in $\mathbb{R}$ (as it is not bounded, and thus divergent).

According to a lemma that we will prove next, f cannot be uniformly continuous on $(0, \infty)$. ■

In a sense, continuity only requires that there be no "holes" in the function; uniform continuity requires that the combination of domain and rule plays "nicely".

Lemma: if f is uniformly continuous on A and $(x_n) \subseteq A$ is a Cauchy sequence, then $f(x_n)$ is a Cauchy sequence.

Proof: if $(x_n) \subseteq A$ is a Cauchy sequence and $\delta > 0$, $\exists N_\delta \in \mathbb{N}$ such that $|x_m - x_n| < \delta$ whenever $m, n > N_\delta$.

But f is uniformly continuous on A, so that $\forall \varepsilon > 0$, $\exists \delta_\varepsilon > 0$ such that
$$x, y \in A \text{ and } |x - y| < \delta_\varepsilon \Longrightarrow |f(x) - f(y)| < \varepsilon.$$
Combining these two statements, with $N_\varepsilon = M_{\delta_\varepsilon}$, yields
$$m, n > N_\varepsilon \Longrightarrow |x_m - x_n| < \delta_\varepsilon \Longrightarrow |f(x_m) - f(x_n)| < \varepsilon,$$
and so $(f(x_n))$ is a Cauchy sequence. ■

While continuous functions are not generally uniformly continuous, there is a specific class of functions for which continuity is equivalent to uniform continuity.

Theorem 38
Let $f:[a,b]\to\mathbb{R}$. Then f is uniformly continuous on $[a,b]$ if it is continuous on $[a,b]$.

Proof: this is the converse of Theorem 37. Assume f is continuous on $[a,b]$. If f is not uniformly continuous, then $\exists\varepsilon_0>0$ such that $\forall\delta>0$, $\exists x_\delta,y_\delta\in[a,b]$ with

$$|f(x_\delta)-f(y_\delta)|\geq\varepsilon_0 \text{ and } |x_\delta-y_\delta|<\delta.$$

For $n\in\mathbb{N}$, let $\delta_n=\frac{1}{n}$. The corresponding sequences $(x_{\delta_n}),(y_{\delta_n})$ lie in $[a,b]$, with

$$|x_{\delta_n}-y_{\delta_n}|<\delta_n=\frac{1}{n}\quad\text{and}\quad|f(x_{\delta_n})-f(y_{\delta_n})|\geq\varepsilon_0,\quad\forall n\in\mathbb{N}.$$

As (x_{δ_n}) is bounded, $\exists(x_{\delta_{n_k}})\subseteq(x_{\delta_n})$ such that $x_{\delta_{n_k}}\to z$ with $k\to\infty$, according to the Bolazano-Weierstrass theorem.

Furthermore, $z\in[a,b]$ according to Theorem 15. The corresponding sequence $(y_{\delta_{n_k}})$ also converges to z since

$$0\leq|y_{\delta_{n_k}}-z|\leq|y_{\delta_{n_k}}-x_{\delta_{n_k}}|+|x_{\delta_{n_k}}-z|<\frac{1}{n_k}+|x_{\delta_{n_k}}-z|$$

according to the squeeze theorem, as both $\frac{1}{n_k},|x_{\delta_{n_k}}-z|\to 0$ with $k\to\infty$. But f is continuous, so both $(f(x_{\delta_{n_k}})),(f(y_{\delta_{n_k}}))\to f(z)$, which is impossible as we have $|f(x_{\delta_n})-f(y_{\delta_n})|\geq\varepsilon_0,\forall n\in\mathbb{N}$. Thus f must be uniformly continuous. ■

There is something "special" about the interval $[a,b]$ that allows for all sorts of interesting results when combined with continuous functions; as we shall see in Chapters 8, 9, 16-17.

Example: show $f:\mathbb{R}\to\mathbb{R}$, $f(x)=\frac{1}{1+x^2}$ is uniformly continuous on $(0,1)$.

Proof: let $\varepsilon>0$. Set $\delta_\varepsilon=\varepsilon$. Note that $\forall z\in\mathbb{R}$, $0\leq(|z|-1)^2=z^2-2|z|+1\implies 2|z|\leq 1+z^2\implies|\frac{z}{1+z^2}|\leq 1/2$. Then whenever $|x-y|<\delta_\varepsilon$, we have:

$$\begin{aligned}|f(x)-f(y)|&=\left|\tfrac{1}{1+x^2}-\tfrac{1}{1+y^2}\right|=\left|\tfrac{y^2-x^2}{(1+x^2)(1+y^2)}\right|=\left|\tfrac{x+y}{(1+x^2)(1+y^2)}\right||x-y|\\&\leq\left(\left|\tfrac{y}{1+y^2}\right|\cdot\underbrace{\tfrac{1}{1+x^2}}_{\leq 1}+\left|\tfrac{x}{1+x^2}\right|\cdot\underbrace{\tfrac{1}{1+y^2}}_{\leq 1}\right)|x-y|\\&\leq\left(\underbrace{\left|\tfrac{y}{1+y^2}\right|}_{\leq 1/2}+\underbrace{\left|\tfrac{x}{1+x^2}\right|}_{\leq 1/2}\right)|x-y|\leq|x-y|<\delta_\varepsilon=\varepsilon,\end{aligned}$$ ■

3.7 Solved Problems

1. Show $\lim\limits_{x\to c} x^3 = c^3$ for any $c \in \mathbb{R}$.

   Proof: if $|x-c|<1$, then $|x|<|c|+1$. Let $\varepsilon>0$ and set $\delta_\varepsilon = \min\{1, \frac{\varepsilon}{3|c|^2+3|c|+1}\}$. Then

$$\begin{aligned}
|x^3-c^3| &= |x-c||x^2+cx+c^2| \leq |x-c|\left(|x|^2+|c||x|+|c|^2\right) \\
&< |x-c|\left((|c|+1)^2+|c|(|c|+1)+|c|^2\right) \\
&= |x-c|\left(3|c|^2+3|c|+1\right) \\
&< \delta_\varepsilon \cdot \left(3|c|^2+3|c|+1\right) \leq \frac{\varepsilon}{3|c|^2+3|c|+1}\cdot\left(3|c|^2+3|c|+1\right) = \varepsilon,
\end{aligned}$$

   whenever $0<|x-c|<\delta_\varepsilon$ and $x\in\mathbb{R}$. ■

2. Let $f:\mathbb{R}\to\mathbb{R}$ and let $c\in\mathbb{R}$. Show that $\lim\limits_{x\to c} f(x) = L$ if and only if $\lim\limits_{x\to 0} f(x+c) = L$.

   Proof: we have

$$\begin{gathered}
\lim_{x\to c} f(x) = L \\
\Updownarrow \\
\forall\varepsilon>0, \exists\delta_\varepsilon>0 \text{ s.t. } |f(x)-L|<\varepsilon \text{ when } 0<|x-c|<\delta_\varepsilon \\
\Updownarrow \\
\text{Set } x=y+c: \ \forall\varepsilon>0, \exists\delta_\varepsilon>0 \text{ s.t. } |f(y+c)-L|<\varepsilon \text{ when } 0<|y|<\delta_\varepsilon \\
\Updownarrow \\
\forall\varepsilon>0, \exists\delta_\varepsilon>0 \text{ s.t. } |f(y+c)-L|<\varepsilon \text{ when } 0<|y-0|<\delta_\varepsilon \\
\Updownarrow \\
\lim_{y\to 0} f(y+c) = L,
\end{gathered}$$

   which completes the proof. ■

3. Use either the ε–δ definition of the limit or the sequential criterion for limits to establish the following limits:

   a) $\lim\limits_{x\to 2} \dfrac{1}{1-x} = -1$;

   b) $\lim\limits_{x\to 1} \dfrac{x}{1+x} = \dfrac{1}{2}$;

   c) $\lim\limits_{x\to 0} \dfrac{x^2}{|x|} = 0$, and

   d) $\lim\limits_{x\to 1} \dfrac{x^2-x+1}{x+1} = \dfrac{1}{2}$

Proof:

a) Let $\varepsilon > 0$ and set $\delta_\varepsilon = \min\{\frac{1}{2}, \frac{\varepsilon}{2}\}$. Then

$$0 < |x-2| < \delta_\varepsilon \implies |x-2| < \frac{1}{2} \iff \frac{3}{2} < x < \frac{5}{2}$$
$$\iff \frac{1}{2} < x-1 < \frac{3}{2} \iff \frac{1}{x-1} < 2.$$

Thus

$$\left|\frac{1}{1-x} - (-1)\right| = \frac{1}{|x-1|}|x-2| = \frac{1}{x-1}|x-2| < 2\delta_\varepsilon < \varepsilon$$

whenever $0 < |x-2| < \delta_\varepsilon$ and $x \in \mathbb{R}$. (Note that if $0 < |x-2| < \delta_\varepsilon$, we've seen that $x > \frac{3}{2}$ and so that $|x-1| = x-1$. This explains why we have gotten rid of the absolute values above.)

b) Let $\varepsilon > 0$ and set $\delta_\varepsilon = \min\{\frac{1}{2}, 3\varepsilon\}$. Then

$$0 < |x-1| < \delta_\varepsilon \implies |x-1| < \frac{1}{2} \iff \frac{1}{2} < x < \frac{3}{2}$$
$$\iff 3 < 2(x+1) < 5 \iff \frac{1}{2(x+1)} < \frac{1}{3}.$$

Thus

$$\left|\frac{x}{1+x} - \frac{1}{2}\right| = \frac{1}{2|x+1|}|x-1| = \frac{1}{2(x+1)}|x-1| < \frac{1}{3}\delta_\varepsilon < \varepsilon$$

whenever $0 < |x-1| < \delta_\varepsilon$ and $x \in \mathbb{R}$. (Note that if $0 < |x-1| < \delta_\varepsilon$, we've seen that $2(x+1) > 3$ and so that $2|x+1| = 2(x+1)$. This explains why we have gotten rid of the absolute values above.)

c) Let $(x_n) \subseteq \mathbb{R}$ be a sequence s.t. $x_n \to 0$ and $x_n \neq 0$ for all n. Then

$$\frac{x_n^2}{|x_n|} = \frac{|x_n|^2}{|x_n|} = |x_n| \to 0,$$

by theorem 14. By the sequence squeeze theorem, the limit must be thus 0.

d) Let $\varepsilon > 0$ and set $\delta_\varepsilon = \min\{\frac{1}{2}, \frac{3}{2}\varepsilon\}$. Then

$$0 < |x-1| < \delta_\varepsilon \implies |2x-1| < 2 \text{ and } \left|\frac{1}{2(x+1)}\right| < \frac{1}{3}.$$

Thus ,whenever $0 < |x-1| < \delta_\varepsilon$ and $x \in \mathbb{R}$, we have

$$\left|\frac{x^2-x+1}{x+1} - \frac{1}{2}\right| = \left|\frac{2x-1}{2(x+1)}\right| |x-1| < \frac{2}{3}|x-1| < \frac{2}{3}\delta_\varepsilon < \varepsilon.$$

This completes the exercise. ∎

4. Show that the following limits do not exist:

a) $\lim_{x \to 0} \frac{1}{x^2}$, with $x > 0$;

b) $\lim\limits_{x\to 0} \dfrac{1}{\sqrt{x}}$, with $x > 0$;

c) $\lim\limits_{x\to 0}(x + \text{sgn}(x))$, and

d) $\lim\limits_{x\to 0} \sin(1/x^2)$, with $x > 0$.

Solution: in each instance, we only give some sequence(s) for which Theorem 26 shows the limit does not exist.

a) $x_n = \frac{1}{n} \to 0$, but $f(x_n) = \frac{1}{1/n^2} = n^2 \to \infty$.

b) $x_n = \frac{1}{n} \to 0$, but $f(x_n) = \frac{1}{1/\sqrt{n}} = \sqrt{n} \to \infty$.

c) $x_n = \frac{1}{n}, y_n = -\frac{1}{n} \to 0$, but $f(x_n) = \frac{1}{n} + 1 \to 1, f(y_n) = -\frac{1}{n} - 1 \to -1$.

d) $x_n = \sqrt{\frac{2}{(4n+1)\pi}}, y_n = \sqrt{\frac{2}{(4n+3)\pi}} \to 0$ but

$$f(x_n) = \sin\left(\frac{4n+1}{2}\pi\right) \to 1,\ f(y_n) = \sin\left(\frac{4n+3}{2}\pi\right) \to -1.$$

This completes the exercise. □

5. Let $c \in \mathbb{R}$ and let $f : \mathbb{R} \to \mathbb{R}$ be such that $\lim\limits_{x\to c}(f(x))^2 = L$. Show that if $L = 0$, then $\lim\limits_{x\to c} f(x) = 0$. Show that if $L \neq 0$, then f may not have a limit at c.

Proof: if $\lim\limits_{x\to c}(f(x))^2 = 0$ then $\forall \eta > 0, \exists \delta_\eta > 0$ such that

$$|f(x)|^2 = \left|(f(x))^2 - 0\right| < \eta$$

whenever $0 < |x - c| < \delta_\eta$. Let $\varepsilon > 0$.

By definition of the real numbers, $\exists \eta_\varepsilon > 0$ such that $\varepsilon = \sqrt{\eta_\varepsilon}$. Set $\delta_\varepsilon = \delta_{\eta_\varepsilon}$. Then

$$|f(x) - 0| = |f(x)| = \sqrt{|f(x)|^2} < \sqrt{\eta_\varepsilon} = \varepsilon$$

whenever $0 < |x - c| < \delta_\varepsilon$.

Now, consider the function $f : \mathbb{R} \to \mathbb{R}$ defined by

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

Then $(f(x))^2 \equiv 1$ and

$$\lim_{x\to 0}(f(x))^2 = \lim_{x\to 0} 1 = 1.$$

But $\lim\limits_{x\to 0} f(x)$ does not exist since $(x_n) = (\frac{1}{n}), (y_n) = (-\frac{1}{n})$ are sequences such that $x_n, y_n \to 0$, $x_n, y_n \neq 0$ for all n and

$$f(x_n) = -1 \to -1 \neq 1 \leftarrow 1 = f(y_n).$$

This completes the proof. ■

6. Let $f : \mathbb{R} \to \mathbb{R}$, let J be a closed interval in $\mathbb{R}$ and let $c \in J$. If f_2 is the restriction of f to J, show that if f has a limit at c then f_2 has a limit at c. Show the converse is not necessarily true.

   Proof: suppose $\lim_{x\to c} f(x) = L$ exists. Then, $\forall \varepsilon > 0$, $\exists \delta_\varepsilon > 0$ s.t. $|f(x) - L| < \varepsilon$ whenever $0 < |x - c| < \delta_\varepsilon$. But $f_2(x) = f(x)$ for all $x \in J \subseteq \mathbb{R}$, so $\forall \varepsilon > 0$, $\exists \delta_\varepsilon > 0$ (exactly as above) s.t. $|f_2(x) - L| = |f(x) - L| < \varepsilon$ whenever $0 < |x - c| < \delta_\varepsilon$ and $x \in J$, and so $\lim_{x\to c} f_2(x) = L$.

   Now consider $f : \mathbb{R} \to \mathbb{R}$ defined by

   $$f(x) = \begin{cases} 0 & \text{if } x \in (-\infty, 0) \cup (1, \infty) \\ 1 & \text{if } x \in [0, 1] \end{cases},$$

   with $J = [0, 1]$ and $f_2 = f|_J$. Then $\lim_{x\to 1} f_2(x) = 1$ but $\lim_{x\to 1} f(x)$ does not exist. ■

7. Determine the following limits and state which theorems are used in each case.

   a) $\lim_{x\to 2} \sqrt{\dfrac{2x+1}{x+3}}$, $(x > 0)$;

   b) $\lim_{x\to 2} \dfrac{x^2 - 4}{x - 2}$, $(x > 0)$;

   c) $\lim_{x\to 0} \sqrt{\dfrac{(x+1)^2 - 1}{x}}$, $(x > 0)$, and

   d) $\lim_{x\to 1} \dfrac{\sqrt{x} - 1}{x - 1}$, $(x > 0)$.

   Solution: We will do c) in its entirety and only give the answers to the others.

   Consider the sequence $(x_n) = (\frac{1}{n})$. Then $x_n \to 0$, $x_n \neq 0\ \forall n \in \mathbb{N}$, and

   $$\frac{(x_n+1)^2 - 1}{x_n} = \frac{\left(\frac{1}{n} + 1\right)^2 - 1}{\frac{1}{n}} = \frac{1}{n} + 2 \to 2.$$

   Hence, if $\lim_{x\to 0} \dfrac{(x+1)^2 - 1}{x}$ exists, its value must be 2, by Theorem 26.

   Let $\varepsilon > 0$. Set $\delta_\varepsilon = \varepsilon$. Then when $0 < |x - 0| < \delta_\varepsilon$ and $x > 0$, we have

   $$\left|\frac{(x+1)^2 - 1}{x} - 2\right| = \left|\frac{x^2 + 2x + 1 - 1 - 2x}{x}\right| = \left|\frac{x^2}{x}\right| = |x| = |x - 0| < \delta_\varepsilon = \varepsilon.$$

   a) 1 b) 4 d) $\frac{1}{2}$ □

8. Give examples of functions f and g such that f and g do not have limits at point c, but both $f+g$ and fg have limits at c.

Solution: Let $f, g : \mathbb{R} \to \mathbb{R}$ be defined by

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

and $g(x) = -f(x)$ for all $x \in \mathbb{R}$. Then $f(x) + g(x) \equiv 0$ and $f(x)g(x) \equiv -1$. As a result,

$$\lim_{x \to 0} (f+g)(x) = 0 \quad \text{and} \lim_{x \to 0} (fg)(x) = -1,$$

but the limits of f and g don't exist at 0 (see solved problem 5). □

9. Determine whether the following limits exist in $\mathbb{R}$:

a) $\displaystyle\lim_{x \to 0} \sin\left(\frac{1}{x^2}\right), \quad$ with $x \neq 0$;

b) $\displaystyle\lim_{x \to 0} x \sin\left(\frac{1}{x^2}\right), \quad$ with $x \neq 0$;

c) $\displaystyle\lim_{x \to 0} \operatorname{sgn} \sin\left(\frac{1}{x}\right), \quad$ with $x \neq 0$, and

d) $\displaystyle\lim_{x \to 0} \sqrt{x} \sin\left(\frac{1}{x^2}\right), \quad$ with $x > 0$.

Solution:

a) Let $(x_n) = (\frac{1}{\sqrt{n\pi}})$ and $(y_n) = (\sqrt{\frac{2}{(4n+1)\pi}})$ for all $n \in \mathbb{N}$. Then $x_n, y_n \to 0$ and $x_n, y_n \neq 0$ for all $n \in \mathbb{N}$. But

$$\sin\left(\frac{1}{x_n^2}\right) = \sin(n\pi) = 0 \quad \text{and} \quad \sin\left(\frac{1}{y_n^2}\right) = \sin\left(\frac{(4n+1)\pi}{2}\right) = 1$$

for all $n \in \mathbb{N}$.

Then $\sin(1/x_n^2) \to 0$ and $\sin(1/y_n^2) \to 1$. As $0 \neq 1$, $\displaystyle\lim_{x \to 0} \sin\left(\frac{1}{x^2}\right)$ doesn't exist.

b) Consider the sequence $(x_n) = (\frac{1}{\sqrt{n\pi}})$. Then $x_n \to 0$ and $x_n \neq 0$ for all $n \in \mathbb{N}$. Furthermore,

$$x_n \sin\left(\frac{1}{x_n^2}\right) = \frac{1}{\sqrt{n\pi}} \sin(n\pi) = \frac{1}{\sqrt{n\pi}} \cdot 0 \to 0.$$

As a result, if $\displaystyle\lim_{x \to 0} x \sin\left(\frac{1}{x^2}\right)$ exists, it must take the value 0. Let $\varepsilon > 0$. Set $\delta_\varepsilon = \varepsilon$. Then

$$\left| x \sin\left(\frac{1}{x^2}\right) - 0 \right| = |x| \left| \sin\left(\frac{1}{x^2}\right) \right| \leq |x| = |x - 0| < \delta_\varepsilon = \varepsilon$$

whenever $0 < |x - 0| < \delta_\varepsilon$ and $x > 0$. Hence $\displaystyle\lim_{x \to 0} x \sin\left(\frac{1}{x^2}\right) = 0$.

c) Let $(x_n) = \left(\frac{2}{(2n+1)\pi}\right)$. Then $x_n \to 0$, $x_n \neq 0$ for all $n \in \mathbb{N}$ and

$$\operatorname{sgn}\left(\sin\left(\frac{1}{x_n}\right)\right) = \operatorname{sgn}\left((-1)^n\right) = (-1)^n,$$

which does not converge. Hence $\lim_{x\to 0} \operatorname{sgn}\left(\sin\left(\frac{1}{x}\right)\right)$ does not exist.

d) $\lim_{x\to 0} \sqrt{x}\sin\left(\frac{1}{x^2}\right) = 0$, with the same proof as b), save for $\delta_\varepsilon = \varepsilon^2$. □

10. Let $f : \mathbb{R} \to \mathbb{R}$ be s.t. $f(x+y) = f(x) + f(y)$ for all $x, y \in \mathbb{R}$. Assume $\lim_{x\to 0} f(x) = L$ exists. Prove that $L = 0$ and that f has a limit at every point $c \in \mathbb{R}$.

Proof: as f is additive, we have $f(2x) = f(x+x) = f(x) + f(x) = 2f(x)$, so that

$$L = \lim_{y\to 0} f(y) = \lim_{2x\to 0} f(2x) = \lim_{x\to 0} f(2x) = \lim_{x\to 0} 2f(x) = 2\lim_{x\to 0} f(x) = 2L;$$

hence $L = 2L$ and $L = 0$, i.e., $\lim_{x\to 0} f(x) = 0$.

Now, let $c \in \mathbb{R}$. Then

$$\begin{aligned}\lim_{x\to c} f(x) &= \lim_{x\to c} (f(x-c) + f(c)) = \lim_{x\to c} f(x-c) + \lim_{x\to c} f(c)\\ &= \lim_{y\to 0} f(y) + f(c) = 0 + f(c) = f(c).\end{aligned}$$

As f is defined on all of $\mathbb{R}$, $f(c)$ exists for all $c \in \mathbb{R}$, and so $\lim_{x\to c} f(x) = f(c)$ exists for all $c \in \mathbb{R}$. ■

11. Let $K > 0$ and let $f : \mathbb{R} \to \mathbb{R}$ satisfy the condition

$$|f(x) - f(y)| \leq K|x-y|$$

for all $x, y \in \mathbb{R}$. Show that f is continuous on $\mathbb{R}$.

Proof: let $c \in \mathbb{R}$ and $\varepsilon > 0$. Set $\delta_\varepsilon = \frac{\varepsilon}{K}$. Then

$$|f(x) - f(c)| \leq K|x-c| < K\delta_\varepsilon < K\frac{\varepsilon}{K} = \varepsilon$$

whenever $|x-c| < \delta_\varepsilon$. ■

12. Let $f : (0,1) \to \mathbb{R}$ be bounded and s.t. $\lim_{x\to 0} f(x)$ does not exist. Show that there are two convergent sequences $(x_n), (y_n) \subseteq (0,1)$ with $x_n, y_n \to 0$ and $f(x_n) \to \xi$, $f(y_n) \to \zeta$, but $\xi \neq \zeta$.

Proof: for $n \in \mathbb{N}$, let $I_n = (0, 1/n)$ and set

$$s_n = \sup f(I_n) \quad \text{and} \quad t_n = \inf f(I_n).$$

These are well-defined as $f(I_n)$ is bounded.

By construction, (s_n) is decreasing and (t_n) is increasing. Since

$$s_1 \geq s_n = \sup f(I_n) \geq \inf f(I_n) = t_n \geq t_1,$$

(s_n) is bounded below by t_1 and (t_n) is bounded above by s_1. Hence $s_n \to s$ and $t_n \to t$ exist, by the bounded monotone convergence theorem.

For $n \in \mathbb{N}$, let $x_n, y_n \in I_n$ be s.t.

$$|f(x_n) - s_n| < \frac{1}{n} \quad \text{and} \quad |f(y_n) - t_n| < \frac{1}{n}.$$

This can always be done as $s_n - \frac{1}{n}$ and $t_n + \frac{1}{n}$ are not the supremum and the infimum, respectively, of $f(I_n)$. Then, $x_n, y_n \to 0$ and $x_n, y_n \neq 0$ for all $n \in \mathbb{N}$.

Furthermore, $f(x_n) \to s$ and $f(y_n) \to t$ according to the sequence squeeze theorem; indeed, $s_n - \frac{1}{n} < f(x_n) \leq s_n$, $t_n \leq f(y_n) < t_n + \frac{1}{n}$, $s_n \to s$, and $t_n \to t$, and the statement follows.

Now, suppose that $s = t = \ell$. Then $s_n, t_n \to \ell$. Let $\varepsilon > 0$. $\exists N_1, N_2 \in \mathbb{N}$ s.t. $|s_n - \ell| < \varepsilon$ whenever $n > N_1$ and $|t_n - \ell| < \varepsilon$ whenever $n > N_2$. Set $N_\varepsilon = \max\{N_1, N_2\}$. Then

$$\ell - \varepsilon < t_n \leq s_n < \ell - \varepsilon$$

whenever $n > N_\varepsilon$. Set $\delta_\varepsilon = \frac{1}{N_\varepsilon}$. Then

$$\ell - \varepsilon < t_{N_\varepsilon} = \inf f(I_{N_\varepsilon}) \leq f(x) \leq \sup f(I_{N_\varepsilon}) \leq s_{N_\varepsilon} < \ell + \varepsilon,$$

that is, $|f(x) - \ell| < \varepsilon$ whenever $0 < |x - 0| < \frac{1}{N_\varepsilon} = \delta_\varepsilon$. Hence $\lim\limits_{x\to 0} f(x) = \ell$, which contradicts the hypothesis that the limit does not exist. Consequently, $s \neq t$, which completes the proof. ■

13. Let $f : \mathbb{R} \to \mathbb{R}$ be continuous on $\mathbb{R}$ and let $P = \{x \in \mathbb{R} : f(x) > 0\}$. If $c \in P$, show that there exists a neighbourhood $V_\delta(c) \subseteq P$.

    Proof: let $c \in P$. Then $f(c) > 0$ and $\exists \varepsilon_0 > 0$ s.t. $f(c) - \varepsilon_0 > 0$. By the continuity of f, $\exists \delta_{\varepsilon_0}$ s.t. $|f(x) - f(c)| < \varepsilon_0$ whenever $|x - c| < \delta_{\varepsilon_0}$.

    Thus, $0 < f(c) - \varepsilon_0 < f(x)$ for all $x \in V_{\delta_{\varepsilon_0}}$, i.e. $V_{\delta_{\varepsilon_0}} \subseteq P$. ■

14. Prove that if an additive function is continuous at some point $c \in \mathbb{R}$, it is continuous on $\mathbb{R}$.

    Proof: in the light of a previous question on the topic, it is sufficient to show that if $\lim\limits_{x\to c} f(x) = f(c)$ for some $c \in \mathbb{R}$, then $\lim\limits_{x\to 0} f(x) = 0$. Let f be continuous at c. Then

$$\begin{aligned} f(c) &= \lim_{x\to c} f(x) = \lim_{x\to c} (f(x-c) + f(c)) \\ &= \lim_{x\to c} f(x-c) + \lim_{x\to c} f(c) = \lim_{y\to 0} f(y) + f(c), \end{aligned}$$

    hence $\lim\limits_{y\to 0} f(y) = 0$, which completes the proof. ■

15. If f is a continuous additive function on $\mathbb{R}$, show that $f(x) = cx$ for all $x \in \mathbb{R}$, where $c = f(1)$.

Proof: let $n \in \mathbb{N}$. Then

$$f(1) = f\left(\frac{n}{n}\right) = f\left(\frac{1}{n} + \cdots + \frac{1}{n}\right) = f\left(\frac{1}{n}\right) + \cdots + f\left(\frac{1}{n}\right) = nf\left(\frac{1}{n}\right),$$

hence $\frac{1}{n}f(1) = f\left(\frac{1}{n}\right)$.

Set $c = f(1)$. Let $y \in \mathbb{Q}$. Then $y = \frac{m}{n}$, where $m \in \mathbb{Z}$ and $n \in \mathbb{N}^{\times}$, and

$$f(y) = f\left(\frac{m}{n}\right) = mf\left(\frac{1}{n}\right) = m\frac{1}{n}f(1) = yc.$$

Let $x \in \mathbb{R}$. Since x is a limit point of $\mathbb{Q}$, $\exists (x_n) \subseteq \mathbb{Q}$ s.t. $x_n \to x$, with $x_n \neq x$ for all $n \in \mathbb{N}$. But $f(x_n) \to f(x)$, by continuity, so $f(x_n) = cx_n \to cx$, by the above discussion. Hence, $f(x) = cx$. ■

16. Let $I = [a, b]$ and $f : I \to \mathbb{R}$ be a continuous function on I s.t. $\forall x \in I$, $\exists y \in I$ s.t. $|f(y)| \leq \frac{1}{2}|f(x)|$. Show $\exists c \in I$ s.t. $f(c) = 0$.

Proof: let $x_1 \in I$. By hypothesis, $\exists x_2 \in I$ s.t.

$$|f(x_2)| \leq \frac{1}{2}|f(x_1)|.$$

Since $x_2 \in I$, $\exists x_3 \in I$ s.t.

$$|f(x_3)| \leq \frac{1}{2}|f(x_2)| \leq \frac{1}{2}\left(\frac{1}{2}|f(x_1)|\right) = \frac{1}{2^2}|f(x_1)|,$$

and so on. The sequence $(x_n) \subseteq I$ thusly built satistfies

$$0 \leq |f(x_n)| \leq \frac{1}{2^{n-1}}|f(x_1)|,$$

by induction (can you show this?).

Then $\lim\limits_{n\to\infty} |f(x_n)| = 0$, by the squeeze theorem, and so $f(x_n) \to 0$. As (x_n) is bounded, it has a convergent subsequence (x_{n_k}) (according to the Bolzano-Weierstrass theorem) whose limit c is in I (because $a \leq x_n \leq b$ for all n).

Since $(f(x_{n_k}))$ is a subsequence of $(f(x_n))$, then

$$\lim_{k\to\infty} f(x_{n_k}) = 0.$$

However,

$$\lim_{k\to\infty} f(x_{n_k}) = f\left(\lim_{k\to\infty} x_{n_k}\right) = f(c),$$

as f is continuous. Hence $f(c) = 0$. ■

17. Show that every polynomial with odd degree has at least one real root.

Proof: let
$$f(x) = a_{2n+1}x^{2n+1} + a_{2n}x^{2n} + \cdots + a_1x + a_0,$$
where $a_i \in \mathbb{R}$ for $i = 0, \ldots, 2n+1$. Assume that $a_{2n} \neq 0$.[3] Let
$$M = \max\left\{(2n+1)\frac{|a_{2n}|}{|a_{2n+1}|}, \left(\frac{|a_{2n-k}|}{|a_{2n}|}\right)^{1/k}; k = 1, \ldots, 2n\right\}.$$
Then, whenever $|x| \geq M$,
- $|a_{2n}||x^{2n}| \geq |a_{2n}||x^{2n}|$;
- $|a_{2n}||x^{2n}| \geq |a_{2n-1}||x^{2n-1}|$;
- $\cdots$;
- $|a_{2n}||x^{2n}| \geq |a_1||x|$, and
- $|a_{2n}||x^{2n}| \geq |a_0|$,

and so
$$\begin{aligned}|a_{2n}x^{2n} + \cdots a_0| &\leq |a_{2n}||x^{2n}| + \cdots + |a_0| \leq |a_{2n}||x^{2n}| + \cdots + |a_{2n}||x^{2n}| \\ &= (2n+1)|a_{2n}||x^{2n}| \leq |a_{2n+1}||x^{2n+1}| = |a_{2n+1}x^{2n+1}|,\end{aligned}$$
from which we concude that $f(M+1)f(-M-1) < 0$.

As f is continuous on $[-M-1, M+1]$, $\exists c \in [-M-1, M+1]$ s.t. $f(c) = 0$, by the intermediate value theorem. ■

18. Let $f : [0,1] \to \mathbb{R}$ be continuous, with $f(0) = f(1)$. Show $\exists c \in [0, \frac{1}{2}]$ s.t. $f(c) = f(c + \frac{1}{2})$.

Proof: let $g : [0, \frac{1}{2}] \to \mathbb{R}$ be defined by $g(x) = f(x) - f(x + \frac{1}{2})$. By construction, g is continuous on $[0, \frac{1}{2}]$. If $g(0) = g(1/2) = 0$, there is nothing else to show. Otherwise,
$$g(0) = f(0) - f(1/2) \quad \text{and} \quad g(1/2) = f(1/2) - f(1) = f(1/2) - f(0);$$
hence $g(0)g(\frac{1}{2}) < 0$. By the intermediate value theorem, $\exists c \in [0, \frac{1}{2}]$ s.t. $g(c) = 0$, that is $f(c) - f(c + \frac{1}{2}) = 0$. This completes the proof. ■

19. Show that $f(x) = \frac{1}{x^2}$ is uniformly continuous on $A = [1, \infty)$, but not on $B = (0, \infty)$.

Proof: if $x, y \in A$, then $x, y \geq 1$. In particular, $|x| = x$ and $|y| = y$, and $\frac{1}{x^2y}, \frac{1}{xy^2} \leq 1$. Let $\varepsilon > 0$ and set $\delta_\varepsilon = \frac{\varepsilon}{2}$. Then
$$\begin{aligned}|f(x) - f(y)| &= \left|\frac{1}{x^2} - \frac{1}{y^2}\right| = \left|\frac{y^2 - x^2}{x^2y^2}\right| = \frac{|y+x||y-x|}{x^2y^2} \\ &= |y-x|\left(\frac{y}{x^2y^2} + \frac{x}{x^2y^2}\right) = |x-y|\left(\frac{1}{x^2y} + \frac{1}{xy^2}\right) \leq 2|x-y| < 2\delta_\varepsilon = \varepsilon\end{aligned}$$
whenever $|x - y| < \delta_\varepsilon$ and $x, y \in A$.

[3] If that is not the case, the proof will proceed in a similar fashion, but a_{2n} will be replaced by the first a_i that is non-zero, starting with a_{2n-1}; if all coefficients are 0, then the real root is 0.

We show that the negation of the definition of uniform continuity holds on B. Let $\varepsilon = 1$ and $\delta > 0$. Then, $\exists N \in \mathbb{N}$ s.t. $\frac{1}{N^2} < \delta$. Set $x_N = \frac{1}{N}$ and $y_N = \frac{1}{N+1}$. Clearly, $x_N, y_N \in B$ and

$$|x_N - y_N| = \left|\frac{1}{N} - \frac{1}{N+1}\right| = \frac{1}{N(N+1)} \leq \frac{1}{N^2} < \delta.$$

However,

$$|f(x_N) - f(y_N)| = |N^2 - (N+1)^2| = 2N + 1 > \varepsilon,$$

that is, f is not uniformly continuous on B. ■

20. If $f(x) = x$ and $g(x) = \sin x$, show that f and g are both uniformly continuous on $\mathbb{R}$ but that their product is not uniformly continuous on $\mathbb{R}$.

Proof: let $\varepsilon > 0$ and set $\delta_\varepsilon = \varepsilon$. Then

$$|f(x) - f(y)| = |x - y| < \delta_\varepsilon = \varepsilon$$

and

$$\begin{aligned} |g(x) - g(y)| = |\sin x - \sin y| &= 2\left|\sin\left(\frac{1}{2}(x-y)\right)\cos\left(\frac{1}{2}(x+y)\right)\right| \\ &\leq 2\frac{1}{2}|x-y| \cdot 1 = |x-y| < \delta_\varepsilon = \varepsilon \end{aligned}$$

(the second-last inequality can be obtained using Taylor's theorem on the sin function, see Chapter 4), whenever $|x - y| < \delta_\varepsilon$ and $x, y \in \mathbb{R}$. Hence f and g are both uniformly continuous.

Set $h(x) = x \sin x$. Let $\varepsilon = 1$ and $\delta > 0$. Them $\exists N \in \mathbb{N}$ s.t. $\frac{1}{N} < \delta$ and $K \in \mathbb{N}$ s.t.

$$K > \frac{1}{4}\left(1 - \cos\frac{1}{N}\right)^{-1} + 3.$$

Define

$$x_K = \frac{4K-3}{2}\pi \quad \text{and} \quad y_K = \frac{4K-3}{2}\pi - \frac{1}{N}.$$

Then $|x_K - y_K| = \frac{1}{N} < \delta$ and

$$|h(x_K) - h(y_K)| \geq \frac{4K-3}{2}\pi\left(1 - \cos\frac{1}{N}\right) > \frac{\pi}{2} > 1 = \varepsilon,$$

and so h is not uniformly continuous. ■

21. Let $A \subseteq \mathbb{R}$ and suppose that f has the following property: $\forall \varepsilon > 0$, $\exists g_\varepsilon : A \to \mathbb{R}$ s.t. g_ε is uniformly continuous on A with $|f(x) - g_\varepsilon(x)| < \varepsilon$ for all $x \in A$. Show f is uniformly continuous on A.

Proof: let $\varepsilon > 0$. Then $\frac{\varepsilon}{3} > 0$ and there exists $g_{\varepsilon/3}$ as in the hypothesis: hence $\exists \eta_{\varepsilon/3} > 0$ s.t. $|g_{\varepsilon/3}(x) - g_{\varepsilon/3}(y)| < \frac{\varepsilon}{3}$ whenever $|x-y| < \eta_{\varepsilon/3}$ and $x, y \in A$. Set $\delta_\varepsilon = \eta_{\varepsilon/3}$. Then

$$\begin{aligned} |f(x) - f(y)| &= |f(x) - g_{\varepsilon/3}(x) + g_{\varepsilon/3}(x) - g_{\varepsilon/3}(y) + g_{\varepsilon/3}(y) - f(y)| \\ &\leq |f(x) - g_{\varepsilon/3}(x)| + |g_{\varepsilon/3}(x) - g_{\varepsilon/3}(y)| + |g_{\varepsilon/3}(y) - f(y)| \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon \end{aligned}$$

whenever $|x-y| < \delta_\varepsilon$ and $x, y \in A$. Hence, f is uniformly continuous on A. ■

22. Is a continuous $p-$periodic fonction on $\mathbb{R}$ bounded and uniformly continuous on $\mathbb{R}$?

Proof: since f is continuous, then $|f|$ is also continuous, being the composition of two continuous functions. As f is $p-$periodic, $\exists c \in [0, p]$ s.t.

$$\sup_{x \in \mathbb{R}} |f(x)| = \sup_{x \in [0,p]} |f(x)| = |f(c)|,$$

by the max/min theorem. Hence f is bounded by $|f(c)|$ on $\mathbb{R}$.

Let $\varepsilon > 0$. By hypothesis, f is continuous on the closed interval $[-1, p+1]$, which implies that that f is uniformly continuous on $[-1, p+1]$ (according to Theorem 38). Then, $\exists \delta_\varepsilon > 0$ s.t. $|f(x) - f(y)| < \varepsilon$ whenever $|x-y| < \delta_\varepsilon$ and $x, y \in [-1, p+1]$.

Without loss of generality, we can assume that $\delta_\varepsilon < 1$. Let $x, y \in \mathbb{R}$ s.t. $|x-y| < \delta_\varepsilon$. Then $\exists k \in \mathbb{Z}$ and $\alpha, \beta \in [-1, p+1]$ s.t. $x = \alpha + kp$ and $y = \beta + kp$.

Thus $|\alpha - \beta| = |x-y| < \delta_\varepsilon$ and $|f(x) - f(y)| = |f(\alpha) - f(\beta)| < \varepsilon$, since f is uniformly continuous on $[-1, p+1]$; consequently, f is uniformly continuous. ■

23. Define $g : \mathbb{R} \to \mathbb{R}$ by

$$g(x) = \begin{cases} \frac{(-1)^n}{n} & \text{if } x = 1/n \text{ for some } n \in \mathbb{N}, \\ 0 & \text{otherwise} \end{cases}$$

Prove that g is continuous at 0.

Proof: let $\varepsilon > 0$. Set $\delta_\varepsilon = \varepsilon$. Then,

$$\left|\frac{1}{n} - 0\right| < \delta \Longrightarrow \left|g\left(\frac{1}{n}\right) - g(0)\right| = \left|\frac{1}{n}\right| = \left|\frac{1}{n} - 0\right| < \delta_\varepsilon = \varepsilon$$

whenever $|1/n - 0| < \delta_\varepsilon$, so g is continuous at 0. ■

3.8 Exercises

1. Prepare a 2-page summary of this chapter, with important definitions and results.
2. Let $A = \{\frac{1}{n} \mid n \in \mathbb{N}\}$. Show that no real number strictly greater than 1 can be a limit point of A.
3. Prove the "min" part of Theorem 33.
4. Complete the solution of solved problem 7.
5. Let $f : \mathbb{R} \to \mathbb{R}$. The **pre-image** of a subset $B \subseteq \mathbb{R}$ under f is

$$f^{-1}(B) = \{a \in A \mid f(a) \in B\}.$$

   Prove that f is continuous if and only if the pre-image of every open subset of $\mathbb{R}$ is an open subset of $\mathbb{R}$.
6. A function $f : A \to \mathbb{R}$ is said to be **Lipschitz** if there is a positive number M such that

$$|f(x) - f(y)| \leq M|x - y| \quad \forall\, x, y \in A.$$

   Show that a Lipschitz function must be uniformly continuous, but that uniformly continuous functions do not have to be Lipschitz.

Chapter 4

Differential and Integral Calculus

We have spent a fair amount of time and energy on concepts like the limit, continuity, and uniform continuity, with the goal of making differential and integral calculus sound. In this chapter, we introduce the concepts of **differentiability** and **Riemann-integrability** for functions, and prove a number of useful calculus results.

4.1 Differentiation

Let $I \subseteq \mathbb{R}$ be an interval, $f : I \to \mathbb{R}$, and $c \in I$. The real number L is the **derivative of** f **at** c, denoted by $f'(c) = L$, if

$$\forall \varepsilon > 0, \exists \delta_\varepsilon > 0 \text{ s.t. } x \in I \text{ and } 0 < |x - c| < \delta_\varepsilon \implies \left| \frac{f(x) - f(c)}{x - c} - L \right| < \varepsilon.$$

In that case, we say that f is **differentiable** at c.[1] While $f'(c)$ (if it exists) is a real number, , $f' : I \to \mathbb{R}$ is a function – the **derivative function**.

> **Example:** let $f : I \to \mathbb{R}$ be defined by $f(x) = x^3$. Set $c \in I$. Then
>
> $$f'(c) = \lim_{x \to c} \frac{f(x) - f(c)}{x - c} = \lim_{x \to c} \frac{x^3 - c^3}{x - c} = \lim_{x \to c} (x^2 + cx + c^2) = 3c^2.$$
>
> The corresponding derivative function is $f' : I \to \mathbb{R}$, $f'(x) = 3x^2$. □

As we learn in calculus courses, there is a link between differentiability and continuity.

[1] This definition simply states that $f'(c)$ exists if $\lim_{x \to c} \frac{f(x)-f(c)}{x-c}$ exists, and that, in that case,

$$f'(c) = \lim_{x \to c} \frac{f(x) - f(c)}{x - c}.$$

Theorem 39
If $f : I \to \mathbb{R}$ has a derivative at c, then f is continuous at c.

Proof: let $x, c \in I$, $x \neq c$. Then $f(x) - f(c) = \left(\frac{f(x)-f(c)}{x-c}\right)(x-c)$ and so

$$\lim_{x\to c}(f(x) - f(c)) = \lim_{x\to c} \frac{f(x) - f(c)}{x - c} \cdot \lim_{x\to c}(x - c),$$

if all the limits exist. But $x - c \to 0$ when $x \to c$, and

$$\lim_{x\to c} \frac{f(x) - f(c)}{x - c} = f'(c)$$

by hypothesis, so

$$\lim_{x\to c}(f(x) - f(c)) = f'(c) = 0 = 0 \Longrightarrow \lim_{x\to c} f(x) = f(c),$$

which means that f is continuous at c. ■

The converse of Theorem 39 does not always hold, however. The function $|\cdot| : \mathbb{R} \to \mathbb{R}$, for instance, is continuous at $x = 0$, but it has no derivative there as $|x|/x$ has no limit when $x \to 0$. Continuity is a **necessary condition** for differentiability, but it is **not sufficient**.

Example (Weierstrass' Monster) Weierstrass provided the first example of such a function in 1872: $f : \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = \sum_{n\in\mathbb{N}} \frac{\cos(3^n x)}{2^n}.$$

That it took so long to find an example is mostly due to the fact that the definition of a function has evolved a fair amount over the last 200 years. □

The usual rules of differentiability are easily demonstrated.

Theorem 40
Let $c \in I$, I an interval, $\alpha \in \mathbb{R}$, $f, g : I \to \mathbb{R}$ be differentiable at c, with $g(c) \neq 0$. Then

1. *αf is differentiable at c and $(\alpha f)'(c) = \alpha f'(c)$;*
2. *$f + g$ is differentiable at c and $(f + g)'(c) = f'(c) + g'(c)$;*
3. *fg is differentiable at c and $(fg)'(c) = f'(c)g(c) + f(c)g'(c)$;*
4. *f/g is differentiable at c and $(f/g)'(c) = \dfrac{f'(c)g(c) - f(c)g'(c)}{[g(c)]^2}$.*

Proof: in all instances, we compute the limit of the differential quotient, taking into account the fact that f and g are differentiable at c.

1. If αf is differentiable at c, then

$$(\alpha f)'(c)=\lim_{x\to c}\frac{(\alpha f)(x)-(\alpha f)(c)}{x-c}=\lim_{x\to c}\frac{\alpha(f(x)-f(c))}{x-c}=\alpha\lim_{x\to c}\frac{f(x)-f(c)}{x-c}.$$

   But f is differentiable at c, so the last limit exists, validating the string of equations, and is equal to $f'(c)$, and so $(\alpha f)'(c)=\alpha f'(c)$.

2. If $f+g$ is differentiable at c, then

$$\begin{aligned}(f+g)'(c)&=\lim_{x\to c}\frac{(f+g)(x)-(f+g)(c)}{x-c}=\lim_{x\to c}\frac{f(x)+g(x)-f(c)-g(c)}{x-c}\\&=\lim_{x\to c}\frac{f(x)-f(c)}{x-c}+\lim_{x\to c}\frac{g(x)-g(c)}{x-c}.\end{aligned}$$

   But both f,g is differentiable at c, so the sum of limits exists, validating the string of equations, and is equal to $f'(c)+g'(c)$, and so $(f+g)'(c)=f'(c)+g'(c)$.

3. If fg is differentiable at c, then

$$\begin{aligned}(fg)'(c)&=\lim_{x\to c}\frac{(fg)(x)-(fg)(c)}{x-c}=\lim_{x\to c}\frac{f(x)g(x)-f(c)g(c)}{x-c}\\&=\lim_{x\to c}\frac{f(x)g(x)-f(c)g(x)+f(c)g(x)-f(c)g(c)}{x-c}\\&=\lim_{x\to c}\frac{f(x)-f(c)}{x-c}g(x)+\lim_{x\to c}f(c)\frac{g(x)-g(c)}{x-c}\\&=\lim_{x\to c}\frac{f(x)-f(c)}{x-c}\cdot\lim_{x\to c}g(x)+f(c)\lim_{x\to c}\frac{g(x)-g(c)}{x-c}\end{aligned}$$

   But both f,g is differentiable at c, so the differential quotient limits exist, validating the string of equations. Furthermore, g is continuous at c, being differentiable at c (acccording to Theorem 39). Hence

$$(fg)'(c)=f'(c)\cdot\lim_{x\to c}g(x)+f(c)g'(c)=f'(c)g(c)+f(c)g'(c).$$

4. Set $h=f/g$; then $f(c)=g(c)h(c)$ and $f'(c)=g'(c)h(c)+g(c)h'(c)$ by the previous rule. Thus

$$(f/g)'(c)=h'(c)=\frac{f'(c)-g'(c)h(c)}{g(c)}=\frac{f'(c)-g'(c)f(c)/g(c)}{g(c)}=\frac{f'(c)g(c)-g'(c)f(c)}{[g(c)]^2},$$

   which completes the proof. ∎

Be careful! Although what we wrote in the proof for the fourth rule is undeniably true, we still need to show that h is differentiable at c under the given conditions before we can use the product rule. The proof could instead look like the following (reprise).

4. Since g is continuous at c (being differentiable at c) and $g(c) \neq 0$, $\exists$ an interval $J \subseteq I$ such that $c \in J$ an $g \neq 0$ on J. If f/g is differentiable at c, then

$$\begin{aligned}(f/g)'(c) &= \lim_{x\to c} \frac{(f/g)(x) - (f/g)(c)}{x-c} = \lim_{x\to c} \frac{f(x)/g(x) - f(c)/g(c)}{x-c} \\ &= \lim_{x\to c} \frac{f(x)g(c) - f(c)g(x)}{g(x)g(c)(x-c)} = \lim_{x\to c} \frac{f(x)g(c) - f(c)g(c) + f(c)g(c) - f(c)g(c)}{g(x)g(c)(x-c)},\end{aligned}$$

so that

$$\begin{aligned}(f/g)'(c) &= \lim_{x\to c} \frac{1}{g(x)g(c)} \left[\frac{f(x)-f(c)}{x-c} g(c) - f(c) \frac{g(x)-g(c)}{x-c} \right] \\ &= \lim_{x\to c} \frac{1}{g(x)g(c)} \cdot \left[\lim_{x\to c} \frac{f(x)-f(c)}{x-c} g(c) - f(c) \lim_{x\to c} \frac{g(x)-g(c)}{x-c} \right].\end{aligned}$$

But both f, g is differentiable at c, so the differential quotient limits exist, validating the string of equations.

Furthermore, g is continuous at c, being differentiable at c (cf. Theorem 39), and $g \neq 0$ on J, so that $\frac{1}{g(x)} \to \frac{1}{g(c)}$ when $x \to c$. Thus

$$(f/g)'(c) = \frac{f'(c)g(c) - f(c)g'(c)}{[g(c)]^2}. \qquad \blacksquare$$

Using mathematical induction, we can easily show that

$$\left[\sum_{i=1}^n f_i \right]'(c) = \sum_{i=1}^n f_i'(c) \quad \text{and} \quad \left[\prod_{i=1}^n f_i \right]'(c) = \sum_{i=1}^n \Big(\prod_{j\neq i} f_j(c) \Big) f_i'(c),$$

if $f_1, \ldots, f_n$ are all differentiable at c. In particular, if $f_1 = \cdots = f_n$, then

$$(f^n)'(c) = n f^{n-1}(c) \cdot f'(c).$$

If we consider the identity function f, then for $c \in \mathbb{R}$, we have

$$f'(c) = \lim_{x\to c} \frac{x-c}{x-c} = 1 \Longrightarrow (f^n)'(x) = n f^{n-1}(x) \cdot f'(x) = n x^{n-1}$$

for all $x \in \mathbb{R}, n \in \mathbb{N}$; this can be extended to $n \in \mathbb{Z}$ using Theorem 40.4.

Theorem 41 (CARATHÉODORY)
Let $I = [a,b]$ and $f : I \to \mathbb{R}$. Then f is differentiable at $c \in I$ if and only if $\exists \varphi_c : I \to \mathbb{R}$, continuous at c such that $f(x) - f(c) = \varphi_c(x)(x-c)$, for all $x \in I$. In that case, $\varphi_c(c) = f'(c)$.

Proof: let $c \in I$. Assume that $f'(c)$ exists. Define $\varphi_c : I \to \mathbb{R}$ by

$$\varphi_c(x) = \begin{cases} \frac{f(x)-f(c)}{x-c}, & x \neq c \\ f'(c), & x = c \end{cases}$$

Then φ_c is continuous at c since

$$\lim_{x \to c} \varphi_c(x) = \lim_{x \to c} \frac{f(x)-f(c)}{x-c} = f'(c) = \varphi_c(c).$$

If $x = c$, then $f(x) = f(c)$ and

$$f(x) - f(c) = f(c) - f(c) = 0 = \varphi_c(c)(c-c) = \varphi_c(x)(x-c).$$

If $x \neq c$ and $x \in I$, then, by definition, $f(x) - f(c) = \varphi_c(x)(x-c)$. Assume now that $\exists \varphi_c : I \to \mathbb{R}$, continuous at c, and such that $f(x) - f(c) = \varphi_c(x)(x-c)$, for all $x \in I$.

If $x \neq c$, then

$$\varphi_c(x) = \frac{f(x)-f(c)}{x-c}$$

and, since φ_c is continuous at c,

$$\varphi_c(c) = \lim_{x \to c} \varphi_c(x) = \lim_{x \to c} \frac{f(x)-f(c)}{x-c}$$

exists. Then $\varphi_c(c) = f'(c)$ and f is differentiable at c. ■

It is important to recognize that φ_c is not, as a function, the same as f', in general **– it is only at c that they can be guaranteed to coincide**, although in certain cases (such as when f is a linear function), $f'(x) = \varphi_c(x)$ for all c in I. Carethéodory's Theorem can be used to prove an important rule of calculus.

Theorem 42 (CHAIN RULE)
Let I, J be closed bounded intervals, $g : I \to \mathbb{R}$ and $f : I \to \mathbb{R}$ be functions such that $f(J) \subseteq I$, and let $c \in J$, with $d = f(c)$. If f is differentiable at c and g is differentiable at d, then the composition $g \circ f : J \to \mathbb{R}$ is differentiable at c and

$$(g \circ f)'(c) = g'(f(c))f'(c) = g'(d)f'(c)$$

.

Proof: since $f'(c)$ exists, Carathéodory's Theorem implies that $\varphi_c : J \to \mathbb{R}$ such that φ_c is continuous at $c \in J$ with

$$\varphi_c(c) = f'(c), \quad \text{and } f(x) - f(c) = \varphi_c(x)(x-c), \quad \text{for all } x \in J.$$

Since $g'(d)$ exists, $\exists \psi_d : I \to \mathbb{R}$ such that ψ_d is continuous at $d \in I$, with

$$\psi_d(d) = g'(d), \quad \text{and } g(y) - g(d) = \psi_d(y)(y-d), \quad \text{for all } y \in I.$$

Thus, if $y = f(x)$ and $d = f(c)$, we have

$$\begin{aligned}(g \circ f)(x) - (g \circ f)(c) = g(f(x)) - g(f(c)) &= \psi_d(f(x))(f(x) - f(c)) \\ &= \psi_d(f(x))\varphi_c(x)(x-c) = \big[(\psi_d \circ f)(x) \cdot \varphi_c(x)\big](x-c),\end{aligned}$$

for all $x \in J$ such that $f(x) \in I$.

However $(\psi_d \circ f) \cdot \varphi_c$ is continuous at c, being the product of two functions which are continuous at c. According to Carathéodory, $(\psi_d \circ f)(c) \cdot \varphi_c(c) = (g \circ f)'(c)$. But

$$(\psi_d \circ f)(c) \cdot \varphi_c(c) = \psi_d(f(c))\varphi_c(c) = g'(f(c))f'(c) = g'(d)f'(c),$$

which completes the proof. ■

The chain rule can be used to determine some of the other classical rules of differentiation.

Examples

- Suppose that $f : I \to \mathbb{R}$ is differentiable at c and that $f, f' \neq 0$ on I. If h is defined by $h(y) = \frac{1}{y}$, $y \neq 0$, then $h'(y) = -\frac{1}{y^2}$. Thus

$$(1/f)'(x) = (h \circ f)'(x) = h'(f(x)) \cdot f'(x) = -\frac{f'(x)}{(f(x))^2}, \quad \text{for all } x \in I.$$

- Let $g = |\cdot|$. Then $g'(c) = \text{sgn}(c)$ for all $c \neq 0$. Indeed,

$$\lim_{x\to c} \frac{|x| - |c|}{x - c} = \begin{cases} \lim_{x\to c} \dfrac{x-c}{x-c}, & c > 0 \\ -\lim_{x\to c} \dfrac{x-c}{x-c}, & c < 0 \end{cases} = \begin{cases} 1, & c > 0 \\ -1, & c < 0 \end{cases} = \text{sgn}(c),$$

  but $g'(0)$ does not exist (even though $\text{sgn}(0) = 0$). If $f : [a,b] \to \mathbb{R}$ is differentiable, the chain Rule states that $|f|'(x) = \text{sgn}(f(x)) \cdot f'(x)$. What happens if $f(c) = 0$? Is $|f|$ differentiable at c?

4.1.1 Mean Value Theorem

With basic calculus in the bag, we can now tackle some of the heavy analysis hitters. Let I be an interval; a function $f : I \to \mathbb{R}$ has a **relative maximum** at $c \in I$ if $\exists \delta > 0$ s.t.

$$f(x) \leq f(c), \quad \forall x \in V_\delta(c) = (c-\delta, c+\delta);$$

it has a **relative minimum** at $c \in I$ if $\exists \delta > 0$ such that

$$f(x) \geq f(c), \quad \forall x \in V_\delta(c) = (c-\delta, c+\delta).$$

If f has either a relative maximum or a relative minimum at c, we say that it has a **relative extremum** at c.[2]

Theorem 43
Let $f : [a,b] \to \mathbb{R}$, $c \in (a,b)$. If f has a relative extremum at c and if f is differentiable at c, then $f'(c) = 0$.

Proof: without loss of generality, assume that f has a relative maximum at c; the proof for a relative minimum follows the same lines. Let $\tilde{\delta}$ be the quantity whose existence is guaranteed by the definition:

$$f(x) \leq f(c), \quad \forall x \in V_{\tilde{\delta}}.$$

If $f'(c) > 0$, then $\exists \delta > 0$ such that $\frac{f(x)-f(c)}{x-c} > 0$ whenever $0 < |x-c| < \delta$. Indeed, according to the definition of the derivative, if $\varepsilon = \frac{1}{2}f'(c) > 0$, $\exists \delta_\varepsilon > 0$ such that

$$\left| \frac{f(x)-f(c)}{x-c} - f'(c) \right| < \varepsilon = \frac{1}{2}f'(c)$$

whenever $0 < |x-c| < \delta_\varepsilon$. Set $\delta = \min\{\delta_\varepsilon, \tilde{\delta}\}$. Then

$$-\frac{1}{2}f'(c) < \frac{f(x)-f(c)}{x-c} - f'(c) < \frac{1}{2}f'(c), \quad \text{whenever } 0 < |x-c| < \delta,$$

and so

$$0 < \frac{1}{2}f'(c) < \frac{f(x)-f(c)}{x-c}, \quad \text{whenever } 0 < |x-c| < \delta.$$

But if $x \in V_\delta(c)$ with $x > c$, then

$$f(x) - f(c) = \underbrace{(x-c)}_{>0} \cdot \underbrace{\frac{f(x)-f(c)}{x-c}}_{>0} > 0,$$

and so $f(x) > f(c)$, which contradicts the fact that f has a relative maximum at c. Thus, $f'(c) \not> 0$. We can prove that $f'(c) \not< 0$ using a similar argument. As neither $f'(c) > 0$ nor $f'(c) < 0$, we must have $f'(c) = 0$. ■

[2]Note that the definition of relative extremum makes no mention of continuity or differentiability.

This result justifies the common practice of looking for relative extrema at **roots of the derivative**. Since c is not an endpoint of I, we must also include a and b in the search for extrema.[3] The next theorem has far-reaching consequences.

Theorem 44 (Rolle)
Let $f : [a,b] \to \mathbb{R}$ be continuous on $[a,b]$ and differentiable on (a,b). If $f(a) = 0$ and $f(b) = 0$, $\exists c \in (a,b)$ such that $f'(c) = 0$.

Proof: if $f \equiv 0$ on $[a,b]$, then the conclusion holds for any $c \in (a,b)$. If $\exists x^*$ such that $f(x^*) \neq 0$, we may suppose, without loss of generality, that $f(x^*) > 0$. According to the max/min theorem, f reaches its maximum

$$\sup\{f(x) \mid x \in [a,b]\} > 0$$

somewhere in $[a,b]$. But since $f(a) = f(b) = 0$, the maximum must be reached in (a,b). Denote that point by c. Then $f'(c)$ exists and since f has a relative maximum at c, Theorem 43 implies that $f'(c) = 0$. ■

This subsection's main result is an easy corollary of Rolle's Theorem.

Theorem 45 (Mean Value Theorem)
Let $f : [a,b] \to \mathbb{R}$ be continuous on $[a,b]$. If f is differentiable on (a,b), $\exists c \in (a,b)$ such that $f(b) - f(a) = f'(c)(b-a)$.

Proof: let $\varphi : [a,b] \to \mathbb{R}$ be defined by

$$\varphi(x) = f(x) - (a) - \frac{f(b)-f(a)}{b-a}(x-a).$$

Then

$$\varphi(a) = f(a) - f(a) - \frac{f(b)-f(a)}{b-a}(a-a) = 0, \quad \text{and}$$
$$\varphi(b) = f(b) - f(a) - \frac{f(b)-f(a)}{b-a}(b-a) = f(b) - f(a) - (f(b) - f(a)) = 0.$$

But φ is continuous on $[a,b]$ as f and $x \mapsto x - a$ are continuous on $[a,b]$. According to Rolle's Theorem, $\exists c \in (a,b)$ such that $\varphi'(c) = 0$. But

$$\varphi'(x) = f'(x) - \frac{f(b)-f(a)}{b-a},$$

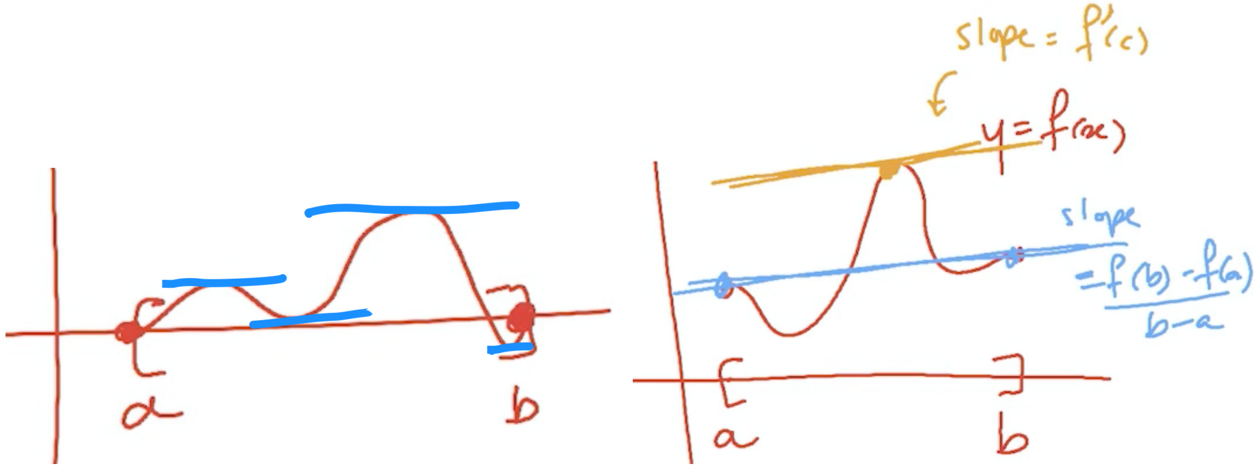
so that $f'(c) - \frac{f(b)-f(a)}{b-c} = 0$, which completes the proof. ■

[3]What happens if f is not differentiable at c in Theorem 43?

Among other things, this tells us something about functions whose derivatives is identically zero on $[a, b]$.

> **Theorem 46**
> *Let $f : [a, b] \to \mathbb{R}$ be continuous on $[a, b]$ and differentiable on (a, b). If $f' \equiv 0$ on (a, b), then f is constant on $[a, b]$.*
>
> **Proof:** let $x \in (a, b]$. According to the mean value theorem, $\exists c \in (a, x)$ such that $f(x) - f(a) = f'(c)(x - a)$. But $f'(c) = 0$, so that $f(x) - f(a) = 0$ for all $x \in [a, b]$. ■

Illustrations of Rolle's theorem (left) and the mean value theorem (right) are shown below.



4.1.2 Taylor Theorem

This subsection's main result is used extensively in applications. It is, in a way, an extension of the mean value theorem to higher order derivatives. We can naturally obtain the **higher-order derivatives** of a function f by formally applying the differentiation rules repeatedly. Hence, $f^{(2)} = f'' = (f')'$, $f^{(3)} = f''' = (f'')' = ((f')')'$, etc. Suppose $f = f^{(0)}$ can be differentiated n times at $x = x_0$. The n**th Taylor polynomial of** f **at** $x = x_0$ is

$$P_n(x; f, x_0) = \sum_{i=0}^{n} \frac{f^{(i)}(x_0)}{i!}(x - x_0)^i.$$

> **Theorem 47** (Taylor)
> *Let $n \in \mathbb{N}$ and $f : [a.b] \to \mathbb{R}$ be such that f and its derivatives $f', f'', \ldots, f^{(n)}$ are continuous on $[a, b]$, and $f^{(n+1)}$ exists on (a, b). If $x_0 \in [a, b]$, then for all $x \neq x_0 \in [a, b]$, $\exists c$ between x and x_0 such that*
>
> $$f(x) = P_n(x; f, x_0) + \frac{f^{(n+1)}(c)}{(n+1)!}(x - x_0)^{n+1}.$$

Proof: let $x \in [a,b]$. If $x_0 < x$, set $J = [x_0, x]$. Otherwise, set $J = [x, x_0]$. Let $F : J \to \mathbb{R}$ be defined by

$$F(t) = f(x) - P_n(t; f, x) = f(x) - f(t) - f'(t)(x-t) - \cdots - \frac{f^{(n)}(t)}{n!}(x-t)^n.$$

Note that F is continuous on J as f and its n higher-order derivatives are continuous on J, and that

$$\begin{aligned} F'(t) = -f'(t) &- \Big[f''(t)(x-t) - f'(t)\Big] - \left[\frac{f'''(t)}{2!}(x-t)^2 - f''(t)(x-t)\right] \\ &- \cdots - \\ &- \left[\frac{f^{(n+1)}(t)}{n!}(x-t)^n - f^{(n)}(t)(x-t)^{n-1}\right]. \end{aligned}$$

Thus $F'(t) = -\frac{f^{(n+1)}(t)}{n!}(x-t)^n$. Let $G : J \to \mathbb{R}$ be defined by

$$G(t) = F(t) - \left(\frac{x-t}{x-x_0}\right)^{n+1} F(x_0).$$

Then

$$\begin{aligned} G(x_0) &= F(x_0) - \left(\frac{x-x_0}{x-x_0}\right)^{n+1} F(x_0) = 0 \\ G(x) &= F(x) - \left(\frac{x-x}{x-x_0}\right)^{n+1} F(x_0) = F(x). \end{aligned}$$

But

$$F(x) = f(x) - f(x) - f'(x)(x-x) - \cdots - \frac{f^{(n)(x)}}{n!}(x-x)^n = 0.$$

Thus $G(x) = 0$. Note that G is continuous on J. Furthermore, G is differentiable on J since

$$G'(t) = F'(t) + \frac{(n+1)}{x-x_0}\left(\frac{x-t}{x-x_0}\right)^n F(x_0) = -\frac{f^{(n+1)}(t)}{n!}(x-t)^n + \frac{(n+1)}{x-x_0}\left(\frac{x-t}{x-x_0}\right)^n F(x_0).$$

As G satisfies the hypotheses of Rolle's theorem, $\exists c$ between x and x_0 such that $G'(c) = 0$. Thus

$$\begin{aligned} \frac{f^{(n+1)}(c)}{n!}(x-c)^n &= (n+1)\frac{(x-c)^n}{(x-x_0)^{n+1}} F(x_0) \quad \Longrightarrow \\ F(x_0) = \frac{f^{(n+1)}(c)}{n!(n+1)}\frac{(x-c)^n}{(x-c)^n}(x-x_0)^{n+1} &= \frac{f^{(n+1)}(c)}{(n+1)!}(x-x_0)^{n+1}. \end{aligned}$$

But

$$F(x_0) = f(x) - P_n(x_0; f, x) \Longrightarrow f(x) = P_n(f; x_0) + F(x_0),$$

which completes the proof. ■

One of the obvious uses of Taylor's theorem is for approximations.

Example: use Taylor's Theorem with $n = 2$ to approximate $\sqrt[4]{1+x}$ near $x_0 = 0$ (for $x > -1$).

Solution: let $f(x) = (1+x)^{1/4}$. Then

$$f'(x) = \frac{1}{4}(1+x)^{-3/4}, \quad f''(x) = -\frac{3}{16}(1+x)^{-7/4}, \quad f'''(x) = \frac{21}{64}(1+x)^{-11/4}$$

are all continuous in closed intervals $[-a, a]$, $1 > a > 0$, so Taylor's theorem can be brought to bear on the situation. Note that $f(0) = 1$, $f'(0) = \frac{1}{4}$ and $f''(0) = -\frac{3}{16}$. According to Taylor's Theorem, for every $x \in [-a, a]$, $1 > a > 0$, $\exists c$ between x and 0 such that

$$f(x) = P_2(x; f, 0) + \frac{f'''(c)}{3!}x^3 = 1 + \frac{1}{4}x - \frac{3}{32}x^2 + \frac{7}{128(1+c)^{11/4}}x^3.$$

For instance, $\sqrt[4]{1.4}$ can be approximated by

$$f(0.4) \approx P_2(0.4) = 1 + \tfrac{1}{4}(0.4) - \tfrac{3}{32}(0.4)^2 \approx 1.085.$$

Moreover, since $c \in (0, 0.4)$,

$$\tfrac{f'''(c)}{6}(0.4)^3 = \tfrac{7}{128}(1+c)^{-11/4}(0.4)^3 \leq \tfrac{7}{128}(0.4)^3 = 0.0035,$$

so $|\sqrt[4]{1.4} - 1.085| \leq 0.0035$, which is to say that the approximation is correct to 2 decimal places. $\square$

4.1.3 Relative Extrema

We end the section on differentiability by giving a characterization of **relative extrema** using the derivative.

A function $f : I \to \mathbb{R}$ is **increasing** (resp. **decreasing**) if

$$f(x_1) \leq f(x_2), \quad \text{(resp. } f(x_1) \geq f(x_2)\text{)} \quad \forall x_1 \leq x_2 \in I.$$

If the inequalities are strict, then the function is **strictly increasing** (resp. **strictly decreasing**). A function that is either increasing or decreasing (exclusively) is **monotone**. If the function is also differentiable, then a link exists.

Theorem 48
Let $f : [a, b] \to \mathbb{R}$ be continuous on $[a, b]$, differentiable on (a, b). Then f is increasing on $[a, b]$ if and only if $f' \geq 0$ on (a, b).

Proof: suppose f is increasing and let $c \in (a,b)$. For all $x < c$ in (a,b), we have $f(x) \leq f(c)$; for all $x > c$ in (a,b), we have $f(x) \geq f(c)$. Thus

$$\frac{f(x)-f(c)}{x-c} \geq 0, \quad \text{for all } x \neq c \in (a,b).$$

Since f is differentiable at c, we must have

$$f'(c) = \lim_{x \to c} \frac{f(x)-f(c)}{x-c} \geq 0.$$

As c is arbitrary, we have $f'(x) \geq 0$ for all $x \in (a,b)$. If, conversely, $f'(x) \geq 0$ for all $x \in (a,b)$, let $x_1 < x_2 \in [a,b]$. By the Mean Value Theorem, $\exists c \in (x_1,x_2)$ such that

$$f(x_2) - f(x_1) = f'(c)(x_2 - x_1).$$

Since $f'(c) \geq 0$ an $x_2 > x_1$, then

$$\frac{f(x_2)-f(x_1)}{x_2-x_1} = f'(c) \geq 0 \Longrightarrow f(x_2) - f(x_1) \geq 0 \Longrightarrow f(x_2) \geq f(x_1),$$

which is to say, f is increasing on $[a,b]$. ■

Theorem 48 holds for decreasing functions as well (after having made the obvious changes to the statement).[4]

The next theorem is a celebrated result from calculus.

Theorem 49 (First Derivative Test)
Let f be continuous on $[a,b]$ and let $c \in (a,b)$. Suppose f is differentiable on (a,c) and on (b,c), but not necessarily at c. Then

1. *if $\exists V_\delta(c) \subseteq [a,b]$ such that $f'(x) \geq 0$ for $c-\delta < x < c$ and $f'(x) \leq 0$ for $c < x < c+\delta$, then f has a relative maximum at c;*
2. *if $\exists V_\delta(c) \subseteq [a,b]$ such that $f'(x) \leq 0$ for $c-\delta < x < c$ and $f'(x) \geq 0$ for $c < x < c+\delta$, then f has a relative minimum at c.*

Proof: we only prove 1.; the proof for 2. follows the same lines. If $x \in (c-\delta, c)$, the mean value theorem states that $\exists c_x \in (x,c)$ such that

$$f(c) - f(x) = \underbrace{f'(c_x)}_{\geq 0}\underbrace{(c-x)}_{\geq 0} \geq 0,$$

so that $f(x) \leq f(c)$ for all $x \in (c-\delta, c)$.

[4] If we switch to strictly monotone functions, only one direction holds in all cases – which one?

If $x \in (c, c+\delta)$, the mean value theorem states that $\exists c_x \in (c, x)$ such that

$$f(c) - f(x) = \underbrace{f'(c_x)}_{\leq 0}\underbrace{(c - x)}_{\leq 0} \geq 0,$$

so that $f(x) \leq f(c)$ for all $x \in (c, c+\delta)$.

Combining these statements with the fact that $f(c) \leq f(c)$, we obtain $f(x) \leq f(c)$ for all $x \in V_\delta(c)$, so f has a relative maximum at c. ■

The converse of the first derivative test is not necessarily true. For instance, the function defined by

$$f(x) = \begin{cases} 2x^4 + x^4 \sin(1/x), & x \neq 0 \\ 0 & x = 0 \end{cases}$$

has an absolute minimum at $x = 0$, but it has derivatives of either sign on either side of any neighbourhood of $x = 0$.

We end this section with a rather surprising result.

Theorem 50 (DARBOUX)
Let $f : [a, b] \to \mathbb{R}$ be differentiable, continuous on $[a, b]$ and let k be strictly confined between $f'(a)$ and $f'(b)$. Then $\exists c \in (a, b)$ with $f'(c) = k$.

Proof: without loss of generality, assume $f'(a) < k < f'(b)$. Define $g : [a, b] \to \mathbb{R}$ by $g(x) = kx - f(x)$; g is then continuous and differentiable on $[a, b]$ given that both f and $x \mapsto kx$ also are.

By the max/min theorem, g reaches its maximum value at some $c \in [a, b]$. However, $g'(a) = k - f'(a) > 0$, so that $c \neq a$, and $g'(b) = k - f'(b) < 0$, so that $c \neq b$. Hence $g'(c) = 0$ for some $c \in (a, b)$, according to Theorem 43, and so $f'(c) = k$, which completes the proof. ■

Darboux's theorem states that the derivative of a continuous function, which needs not be continuous, nevertheless satisfies the intermediate value property.[5]

There are a number of other results which could be shown about differentiable functions, but they are left as exercises (see question 4).

[5]That seems like witchcraft, right? It shouldn't be possible, but the argument is sound. One of the lessons from this result is that analytical reasoning *can* be informed by intuition and geometry, but ultimately, the validity of results rests on proofs.

4.2 Riemann Integral

Calculus as a discipline only took flight after Newton announced his **theory of fluxions**. With Leibniz' independent discovery that the reversal of the process for fining tangents lead to areas under curves, integration was born. Riemann was the first to discuss integration as a process **separate** from differentiation.

We start by studying the integration of a functions $\mathbb{R} \to \mathbb{R}$. Later on, we will tackle integration of functions $\mathbb{R}^n \to \mathbb{R}$ (see Chapter 21) and of functions $\mathbb{R}^n \to \mathbb{R}^n$ (see Chapter 14).

Let $I = [a, b]$. A **partition** $P \in \mathcal{P}([a, b])$ is a subset $P = \{x_0, \dots, x_n\} \subseteq I$ such that

$$a = x_0 < x_1 < \cdots < x_{n-1} < x_n = b.$$

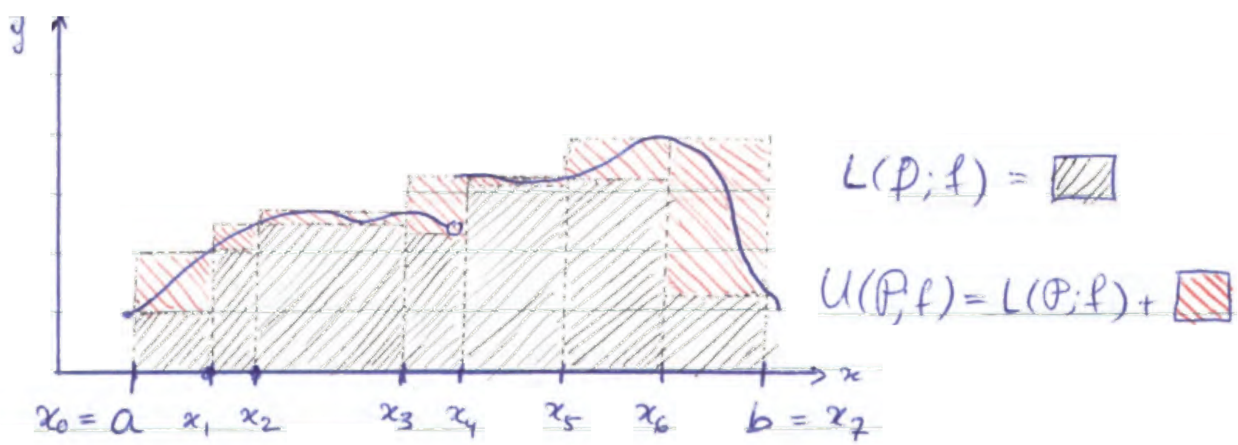
If $f : I \to \mathbb{R}$ is bounded and P is a partition of I, the sums

$$L(P; f) = \sum_{i=1}^{n} m_i(x_i - x_{i-1}) < \infty, \quad U(P; f) = \sum_{i=1}^{n} M_i(x_i - x_{i-1}) < \infty,$$

where

$$m_i = \inf\{f(x) \mid x \in [x_{i-1}, x_i]\}, \quad M_i = \sup\{f(x) \mid x \in [x_{i-1}, x_i]\}, \quad 1 \leq i \leq n$$

are the **lower** and the **upper sum of f corresponding to** P, respectively. If $f : I \to \mathbb{R}_0^+$, we can give a graphical representation of these sums; $L(P; f)$ is the area of the union of the rectangles with base $[x_{k-1}, x_k]$ and height m_k, and $U(P; f)$ is the area of the union of the rectangles with base $[x_{k-1}, x_k]$ and height M_k.



A partition Q of I is a **refinement** of a partition P of I if $P \subseteq Q$.

> **Example:** both $P = \{0, 1, 4, 10\}$ and $Q = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 10\}$ are partitions of $[0, 10]$; since $Q \supseteq P$, Q is a refinement of P. ☐

We will use the following lemma repeatedly in this section.

Lemma: let $I = [a, b]$ and $f : I \to \mathbb{R}$ be bounded. Then

1. $L(P; f) \leq U(P; f)$ for any partition P of I;
2. $L(P; f) \leq L(Q; f)$ and $U(Q; f) \leq L(Q; f)$ for any refinement $Q \supseteq P$ of I, and
3. $L(P_1; f) \leq U(P_2; f)$ for any pair of partitions P_1, P_2 of I.

Proof:

1. Let $P = \{x_0, \ldots, x_n\}$ be a partition of I. Since

$$m_i = \inf\{f(x) \mid x \in [x_{i-1}, x_i]\} \leq \sup\{f(x) \mid x \in [x_{i-1}, x_i]\} = M_i$$

for all $1 \leq i \leq n$, then

$$L(P; f) = \sum_{i=1}^{n} m_i(\underbrace{x_i - x_{i-1}}_{>0}) \leq \sum_{i=1}^{n} M_i(\underbrace{x_i - x_{i-1}}_{>0}) = U(P; f).$$

2. Let $Q = \{y_0, \ldots, y_m\}$ be a refinement of $P = \{x_0, \ldots, x_n\}$. Set $I_i = [x_{i-1}, x_i]$ and $\tilde{I}_j = [y_{j-1}, y_j]$, for $1 \leq i \leq n, 1 \leq j \leq m$. Write $m_i = \inf\{f(x) \mid x \in I_i\}$ and $\tilde{m}_j = \inf\{f(x) \mid x \in \tilde{I}_j\}$ and fix $1 \leq i \leq n$. Then $\exists j, k$ such that

$$I_i = \tilde{I}_{j+1} \cup \cdots \cup \tilde{I}_{j+k} = \bigcup_{\ell=1}^{k} \tilde{I}_{j+\ell}.$$

Then

$$\begin{aligned} m_i(x_i - x_i - 1) &= m_i(y_j + k - y_j) = m_i(y_{j+1} - y_j + \cdots + y_{j+k} - y_{j+k-1}) \\ &= m_i(y_{j+1} - y_j) + \cdots + m_i(y_{j+k} - y_{j+k-1}) \\ &= \sum_{\ell=1}^{k} m_i(y_{j+\ell} - y_{j+\ell-1}) \leq \sum_{\ell=1}^{k} \tilde{m}_{j+\ell}(y_{j+\ell} - y_{j+\ell-1}) \end{aligned}$$

since $\tilde{I}_{j+\ell} \subseteq I_i$ for all $\ell = 1, \ldots, k$. Hence

$$L(P; f) = \sum_{i=1}^{n} m_i(x_i - x_{i-1}) \leq \sum_{j=1}^{m} \tilde{m}_j(y_j - y_{j-1}) = L(Q; f).$$

The proof for $U(P; f) \geq U(Q; f)$ follows a similar argument.

3. Let P_1, P_2 be partitions of I. Set $Q = P_1 \cup P_2$. Then Q is a refinement of both P_1 and P_2. By the results proven in the previous parts of this lemma, we have

$$L(P_1; f) \leq L(Q; f) \leq U(Q; f) \leq U(P_2; f),$$

which completes the proof. ■

Let $I = [a, b]$ and $f : I \to \mathbb{R}$ be bounded. The **lower integral of f on I** is the number

$$L(f) = \sup\{L(P; f) \mid P \text{ a partition of } I\}.$$

The **upper integral of f on I** is the number

$$U(f) = \inf\{U(P; f) \mid P \text{ a partition of } I\}.$$

Since f is bounded on I, $\exists m, M$ such that $m \leq f(x) \leq M$ for all $x \in I$. Consider the **trivial partition** $P_0 = \{a, b\}$. Since any partition P of I is a refinement of P_0, we thus have

$$L(P; f) \leq U(P_0; f) \leq M(b-a) \quad \text{and} \quad U(P; f) \geq L(P_0; f) \geq m(b-a).$$

Thus $L(f), U(f)$ exist, by completeness. But we can say more.

Theorem 51
Let $f : [a, b] \to \mathbb{R}$ be bounded. Then $L(f) \leq U(f)$.

Proof: let P_1, P_2 be partitions of $[a, b]$. Then $L(P_1; f) \leq U(P_2; f)$. If we fix P_2, $U(P_2; f)$ is an upper bound for

$$A = \{L(P_1; f) \mid P_1 \text{ is a partition of } [a, b]\}.$$

Since $L(f) = \sup(A)$ and since P_2 was arbitrary, $L(f)$ is a lower bound for

$$B = \{U(P_2; f) \mid P_2 \text{ is a partition of } [a, b]\}.$$

Thus $L(f) \leq \inf(B) = U(f)$. ■

When $L(f) = U(f)$, we say that f is **Riemann-integrable** on $[a, b]$; the **integral of f on $[a, b]$** is the real number

$$L(f) = U(f) = \int_a^b f = \int_a^b f(x)\,\mathrm{d}x.$$

By convention, we define $\int_a^b f = -\int_b^a f$ when $b < a$. Note that $\int_a^a f = 0$ for all bounded functions f.

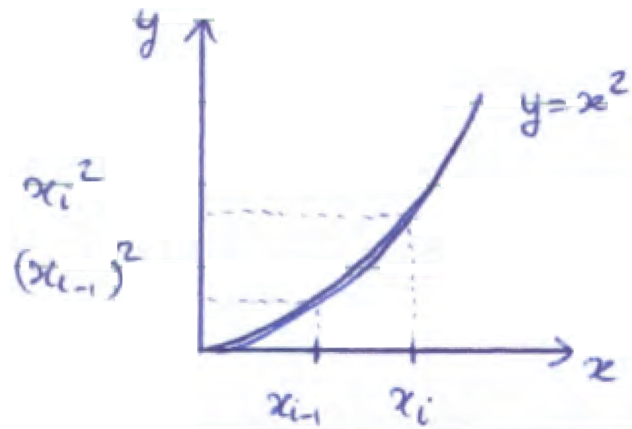
Example: show directly that the function defined by $h(x) = x^2$ is Riemann-integrable on $[a, b]$, $b > a \geq 0$. Furthermore show that $\int_a^b h = \frac{b^3 - a^3}{3}$.

Proof: let $P_n = \left\{x_i = a + \frac{b-a}{n} \cdot i \mid i = 0, \ldots, n\right\} \in \mathcal{P}([a, b])$. For $i = 1, \ldots, n$ set $m_i = \inf\{h(x) \mid x \in [x_{i-1}, x_i]\}$. With this notation, we have

$$L(P_n; h) = \sum_{i=1}^{n} m_i(x_i - x_{i-1}) = \frac{b-a}{n} \sum_{i=1}^{n} m_i.$$

But $h'(x) = 2x \geq 0$ when $x \geq 0$, and so h is increasing on $[a,b]$. Consequently, for $i = 1, \ldots, n$, we have

$$m_i = x_{i-1}^2 = \big(a + \frac{b-a}{n}(i-1)\big)^2 = a^2 + 2\frac{a(b-a)}{n}(i-1) + \frac{(b-a)^2}{n^2}(i-1)^2.$$



The lower sum of h associated to P_n is thus

$$\begin{aligned}
L(P_n;h) &= \frac{b-a}{n}\sum_{i=1}^{n}\big(a^2 + 2\frac{a(b-a)}{n}(i-1) + \frac{(b-a)^2}{n^2}(i-1)^2\big) \\
&= \frac{na^2(b-a)}{n} + \frac{2a(b-a)^2}{n^2}\sum_{i=1}^{n}(i-1) + \frac{(b-a)^3}{n^3}\sum_{i=1}^{n}(i-1)^2 \\
&= a^2(b-a) + \frac{2a(b-a)^2}{n^2}\cdot\frac{n(n-1)}{2} + \frac{(b-a)^3}{n^3}\cdot\frac{n(n-1)(2n-1)}{6} \\
&= a^2(b-a) + a(b-a)^2\big(1-\tfrac{1}{n}\big) + \tfrac{(b-a)^3}{6}\big(1-\tfrac{1}{n}\big)\big(2-\tfrac{1}{n}\big).
\end{aligned}$$

For the lower sum of h on $[a,b]$, we have

$$\begin{aligned}
L(h) &= \sup\{L(P;h) \mid P \in \mathcal{P}([a,b])\} \geq \sup_{n\in\mathbb{N}}\{L(P_n;h)\} \\
&= \sup_{n\in\mathbb{N}}\left\{a^2(b-a) + a(b-a)^2\left(1-\tfrac{1}{n}\right) + \tfrac{(b-a)^3}{6}\left(1-\tfrac{1}{n}\right)\left(2-\tfrac{1}{n}\right)\right\} \\
&= \lim_{n\to\infty}\left[a^2(b-a) + a(b-a)^2\left(1-\tfrac{1}{n}\right) + \tfrac{(b-a)^3}{6}\left(1-\tfrac{1}{n}\right)\left(2-\tfrac{1}{n}\right)\right] \\
&= a^2(b-a) + a(b-a)^2 + \tfrac{(b-a)^3}{6}\cdot 2 = \tfrac{b^3-a^3}{3}.
\end{aligned}$$

Similarly, we can show that

$$U(P_n;h) = a^2(b-a) + a(b-a)^2\left(1+\tfrac{1}{n}\right) + \tfrac{(b-a)^3}{6}\left(1+\tfrac{1}{n}\right)\left(2+\tfrac{1}{n}\right).$$

For the upper sum of h on $[a,b]$, we have

$$\begin{aligned} U(h) &= \inf\{U(P;h) \mid P \in \mathcal{P}([a,b])\} \leq \inf_{n\in\mathbb{N}}\{U(P_n;h)\} \\ &= \inf_{n\in\mathbb{N}}\left\{a^2(b-a)+a(b-a)^2\left(1+\tfrac{1}{n}\right)+\tfrac{(b-a)^3}{6}\left(1+\tfrac{1}{n}\right)\left(2+\tfrac{1}{n}\right)\right\} \\ &= \lim_{n\to\infty}\left[a^2(b-a)+a(b-a)^2\left(1+\tfrac{1}{n}\right)+\tfrac{(b-a)^3}{6}\left(1+\tfrac{1}{n}\right)\left(2+\tfrac{1}{n}\right)\right] \\ &= a^2(b-a)+a(b-a)^2+\tfrac{(b-a)^3}{6}\cdot 2 = \tfrac{b^3-a^3}{3}. \end{aligned}$$

Thus $\frac{b^3-a^3}{3} \leq L(h) \leq U(h) \leq \frac{b^3-a^3}{3}$ and so $L(h) = U(h) = \int_a^b h = \frac{b^3-a^3}{3}$, which completes the proof. ■

It is clearly not the most efficient process in practice, but it works!

Example: show directly that the **Dirichlet function** defined by

$$f(x) = \begin{cases} 1, & x \in \mathbb{Q} \\ 0, & x \notin \mathbb{Q} \end{cases}$$

is not Riemann-integrable on $[0,1]$.

Proof: let $P = \{x_0,\ldots,x_n\} \in \mathcal{P}([0,1])$. Since both $\mathbb{Q}$ and $\mathbb{R}\setminus\mathbb{Q}$ are dense in $\mathbb{R}$, for each $1 \leq i \leq n$, $\exists q_i \in \mathbb{Q}, t_i \notin \mathbb{Q}$ such that $q_i, t_i \in [x_{i-1}, x_i]$.
But $f(q_i) = 0$ and $f(t_i) = 1$, so that $m_i = 0$, $M_i = 1$ for all $1 \leq i \leq n$. This implies that $L(P;f) = 0$ and $U(P;f) = 1$ for any partition P. Thus $L(f) = 0 \neq 1 = U(f)$, and so f is not Riemann-integrable. ■

This last example underlines some of the shortcomings of the Riemann integral – by any account the integral of Dirichlet's function should really be 0 on $[0,1]$: the set $\mathbb{R}\setminus\mathbb{Q}$ is so much larger than $\mathbb{Q}$ that whatever happens on $\mathbb{Q}$ should largely be irrelevant (see Section 1.2). There are various theories of integration – as we shall see in Chapter 21, the Lebesgue-Borel integral of f on $[0,1]$ is indeed 0.

Other issues arise with the Riemann integral, which we will discuss in the coming sections.

4.2.1 Riemann's Criterion

We focus on two **fundamental questions** associated with the Riemann integral of a function over an interval $[a,b]$: **does it exist? If so, what value does it take?**

The direct approach is cumbersome, even for the simplest of functions. The following result allows us to bypass the need to compute $L(f)$ and $U(f)$ to determine if a function is Riemann-integrable or not.

Theorem 52 (Riemann's Criterion)
Let $I = [a, b]$ and $f : I \to \mathbb{R}$ be a bounded function. Then f is Riemann-integrable if and only if $\forall \varepsilon > 0$, $\exists P_\varepsilon$ a partition of I such that the lower sum and the upper sum of f corresponding to P_ε satisfy $U(P_\varepsilon; f) - L(P_\varepsilon; f) < \varepsilon$.

Proof: if f is Riemann-integrable, then $L(f) = U(f) = \int_a^b f$. Let $\varepsilon > 0$.

Since $\int_a^b f - \frac{\varepsilon}{2}$ is not an upper bound of $\{L(P; f) \mid P \text{ a partition of } [a, b]\}$, there exists a partition P_1 such that

$$\int_a^b f - \frac{\varepsilon}{2} < L(P_1; f) \leq \int_a^b f.$$

Using a similar argument, there exists a partition P_2 such that

$$\int_a^b f + \frac{\varepsilon}{2} \geq U(P_2; f) > \int_a^b f.$$

Set $P_\varepsilon = P_1 \cup P_2$. Then P_ε is a refinement of P_1 and P_2. Consequently,

$$\int_a^b f - \frac{\varepsilon}{2} < L(P_1; f) \leq L(P_\varepsilon; f) \leq U(P_\varepsilon; f) \leq U(P_2; f) < \int_a^b f + \frac{\varepsilon}{2}$$

which implies that

$$U(P_\varepsilon; f) - L(P_\varepsilon; f) < \varepsilon.$$

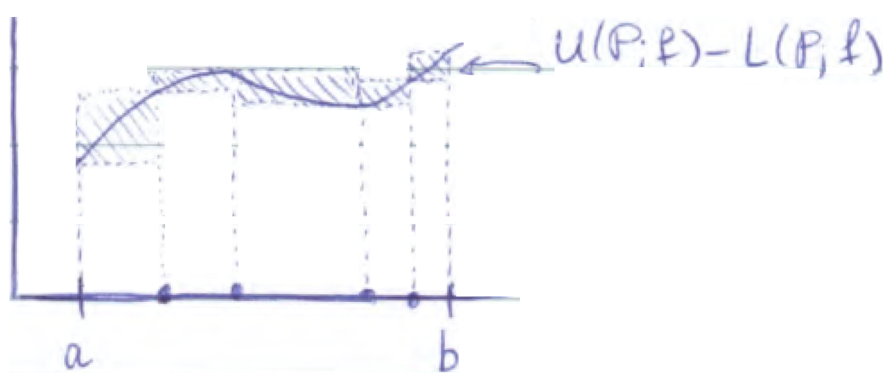
Conversely, let $\varepsilon > 0$ and P_ε be such that $U(P_\varepsilon; f) - L(P_\varepsilon; f) < \varepsilon$. Since

$$U(f) \leq U(P_\varepsilon; f) \quad \text{and} \quad L(f) \geq L(P_\varepsilon; f),$$

then

$$0 \leq U(f) - L(f) \leq U(P_\varepsilon; f) - L(P_\varepsilon; f) < \varepsilon.$$

But $\varepsilon > 0$ was arbitrary, so $U(f) - L(f) = 0$, which implies that $U(f) = L(f)$ and that f is Riemann-integrable on $[a, b]$. ∎



In the illustration on the previous page (for a continuous function), the smaller the shaded area is, the closer $U(P;f)$ and $L(P;f)$ are to $\int_a^b f$.

There are 2 instances where the Riemann-integrability of a function f on $[a,b]$ is guaranteed: when f is **monotone**, and when it is **continuous**.

Theorem 53
Let $I=[a,b]$ and $f:I\to\mathbb{R}$ be a monotone function on I. Then f is Riemann-integrable on I.

Proof: we show that the result holds for increasing functions. The proof for decreasing functions is similar. Let

$$P_n=\{x_i=a+i\left(\tfrac{b-a}{n}\right)\mid i=0,\ldots,n\}$$

be the partition of I into n equal sub-intervals. Since f is increasing on I, we have, for $1\leq i\leq n$,

$$\begin{aligned} m_i &= \inf\{f(x)\mid x\in[x_{i-1},x_i]\}=f(x_{i-1}),\\ M_i &= \sup\{f(x)\mid x\in[x_{i-1},x_i]\}=f(x_i). \end{aligned}$$

Hence,

$$\begin{aligned} U(P_n;f)-L(P_n;f) &= \sum_{i=1}^n M_i(x_i-x_{i-1})-\sum_{i=1}^n m_i(x_i-x_{i-1})\\ &= \sum_{i=1}^n (M_i-m_i)(x_i-x_{i-1})\\ &= \frac{b-a}{n}\sum_{i=1}^n (f(x_i)-f(x_{i-1}))\\ &= \frac{b-a}{n}\Big[f(x_1)-f(x_0)+\cdots+f(x_n)-f(x_{n-1})\Big]\\ &= \frac{b-a}{n}(f(b)-f(a))\geq 0. \end{aligned}$$

Let $\varepsilon>0$. By the Archimedean property, $\exists N_\varepsilon\in\mathbb{N}$ such that

$$\frac{(b-a)(f(b)-f(a))}{\varepsilon}<n.$$

Set $P_\varepsilon=P_n$. Then

$$U(P_\varepsilon;f)-L(P_\varepsilon;f)<\frac{b-a}{N_\varepsilon}(f(b)-f(a))<\varepsilon,$$

and f is Riemann-integrable on $[a,b]$, according to Riemann's criterion. ■

Theorem 54 Let $I = [a, b]$ and $f : I \to \mathbb{R}$ be continuous, with $a < b$. Then f is Riemann-integrable on I.

Proof: let $\varepsilon > 0$. According to Theorem 38, f is uniformly continuous on I. Hence $\exists \delta_\varepsilon > 0$ s.t. $|f(x) - f(y)| < \frac{\varepsilon}{b-a}$ whenever $|x - y| < \delta_\varepsilon$ and $x, y \in [a, b]$.

Pick $n \in \mathbb{N}$ such that $\frac{b-a}{n} < \delta_\varepsilon$ and let

$$P_\varepsilon = \{x_i = a + i\left(\tfrac{b-a}{n}\right) \mid i = 0, \ldots, n\}$$

be the partition of $[a, b]$ into n equal sub-intervals.

As f is continuous on $[x_{i-1}, x_i]$, $\exists u_i, v_i \in [x_{i-1}, x_i]$ such that

$$m_i = \inf\{f(x) \mid x \in [x_{i-1}, x_i]\} = f(u_i), \quad M_i = \sup\{f(x) \mid x \in [x_{i-1}, x_i]\} = f(v_i),$$

for all $1 \leq i \leq n$, according to the max/min Theorem. Since $|u_i - v_i| \leq \frac{b-a}{n} < \delta_\varepsilon$ for all i, we have:

$$\begin{aligned} U(P_\varepsilon; f) - L(P_\varepsilon; f) &= \sum_{i=1}^{n}(M_i - m_i)(x_i - x_{i-1}) = \frac{b-a}{n}\sum_{i=1}^{n}(f(v_i) - f(u_i)) \\ &< \frac{b-a}{n}\sum_{i=1}^{n}\frac{\varepsilon}{b-a} = \varepsilon, \end{aligned}$$

by uniform continuity of f. According to Theorem 52, f is Riemann-integrable. ■

4.2.2 Properties of the Riemann Integral

The Riemann integral has a whole slew of interesting properties.

Theorem 55 (Properties of the Riemann Integral)
Let $I = [a, b]$ and $f, g : I \to \mathbb{R}$ be Riemann-integrable on I. Then

1. *$f + g$ is Riemann-integrable on I, with $\int_a^b (f + g) = \int_a^b f + \int_a^b g$;*
2. *if $k \in \mathbb{R}$, $k \cdot f$ is Riemann-integrable on I, with $\int_a^b k \cdot f = k \int_a^b f$;*
3. *if $f(x) \leq g(x)\ \forall x \in I$, then $\int_a^b f \leq \int_a^b g$, and*
4. *if $|f(x)| \leq K\ \forall x \in I$, then $\left|\int_a^b f\right| \leq K(b - a)$.*

Proof: we use a variety of pre-existing results.

1. Let $\varepsilon > 0$. Since f, g are Riemann-integrable, $\exists P_1, P_2$ partitions of I such that $U(P_1; f) - L(P_1; f) < \frac{\varepsilon}{2}$ and $U(P_2; g) - L(P_2; g) < \frac{\varepsilon}{2}$.

   Set $P = P_1 \cup P_2$. Then P is a refinement of P_1 and P_2, and

   $$\begin{aligned} U(P; f+g) &\leq U(P; f) + U(P; g) \\ &< L(P; f) + L(P; g) + \varepsilon \leq L(P; f+g) + \varepsilon, \end{aligned}$$

   since, over non-empty subsets of I, we have

   $$\begin{aligned} \inf\{f(x) + g(x)\} &\geq \inf\{f(x)\} + \inf\{g(x)\} \\ \sup\{f(x) + g(x)\} &\leq \sup\{f(x)\} + \sup\{g(x)\}. \end{aligned}$$

   Hence $f + g$ is Riemann-integrable according to Riemann's criterion. Furthermore, we see from above that

   $$\int_a^b (f+g) \leq U(P; f+g) < L(P; f) + L(P; g) + \varepsilon \leq \int_a^b f + \int_a^b g + \varepsilon$$

   and

   $$\int_a^b f + \int_a^b g \leq U(P; f) + U(P; g) < L(P; f+g) + \varepsilon \leq \int_a^b (f+g) + \varepsilon.$$

   Since $\varepsilon > 0$ is arbitrary, $\int_a^b f + \int_a^b g \leq \int_a^b (f+g) \leq \int_a^b f + \int_a^b g$, from which we conclude that $\int_a^b (f+g) = \int_a^b f + \int_a^b g$.

2. The proof for $k = 0$ is trivial. We show that the result holds for $k < 0$ (the proof for $k > 0$ is similar). Let $P = \{x_0, \ldots, x_n\}$ be a partition of I.

   Since $k < 0$, we have $\inf\{kf(x)\} = k \sup\{f(x)\}$ over non-empty subsets of I, and so we have $L(P; kf) = kU(P; f)$. In particular,

   $$\begin{aligned} L(kf) &= \sup\{L(P; kf) \mid P \text{ a partition of } I\} \\ &= \sup\{kU(P; f) \mid P \text{ a partition of } I\} \\ &= k \inf\{U(P; f) \mid P \text{ a partition of } I\} = kU(f) \end{aligned}$$

   Similarly, $U(P; kf) = kL(P; f)$ and $U(fk) = kL(f)$, so

   $$L(fk) = \underbrace{kU(f) = kL(f)}_{\text{since } f \text{ is R-int.}} = U(kf).$$

   Thus kf is Riemann-integrable on I and $\int_a^b kf = L(k) = kU(f) = \int_a^b f$.

3. We start by showing that if $h : I \to \mathbb{R}$ is integrable on I and $h(x) \geq 0$ for all $x \in I$, then $\int_a^b h(x) \geq 0$. Let $P_0 = \{a, b\} = \{x_0, x_1\}$ and $m_1 = \inf\{h(x) \mid x \in [a, b]\} \geq 0$. Then,

$$0 \leq m_1(b-a) = L(P_0; h) \leq L(P; h)$$

for any partition P of I, as $P \supseteq P_0$. But h is Riemann-integrable by assumption, thus

$$\int_a^b h = \sup\{L(P; h) \mid P \text{ a partition of } I\} \geq L(P_0; h) \geq 0.$$

Then, set $h = g - f$. By hypothesis, $h(x) = g(x) - f(x) \geq 0$. Then

$$\int_a^b h = \int_a^b (g - f) = \int_a^b g - \int_a^b f \geq 0,$$

which implies that $\int_a^b g \geq \int_a^b f$.

4. Let $P_0 = \{a, b\} = \{x_0, x_1\}$. As always, set $m_1 = \inf\{f(x) \mid x \in [a, b]\}$, and $M_1 = \sup\{f(x) \mid x \in [a, b]\}$. Then for any partition P of I, we have

$$\begin{aligned} m_1(b-a) &= L(P_0; f) \leq L(P; f) \leq L(f) = \int_a^b f \\ &= U(f) \leq U(P; f) \leq U(P_0; f) = M_1(b-a). \end{aligned}$$

In particular,

$$m_1(b-a) \leq \int_a^b f \leq M_1(b-a).$$

Now, if $|f(x)| \leq K$ for all $x \in I$, then $-K \leq m_1$ and $M_1 \leq K$ so that

$$-K(b-a) \leq m_1(b-a) \leq \int_a^b f \leq M_1(b-a) \leq K(b-a),$$

so that $|\int_a^b f| \leq K(b-a)$. ■

When all the functions involved are non-negative, these results and the next one are compatible with the calculus interpretation of the Riemann integral as the **area under the curve**.

Theorem 56 (Additivity of the Riemann Integral)
Let $I = [a, b]$, $c \in (a, b)$, and $f : I \to \mathbb{R}$ be bounded on I. Then f is Riemann-integrable on I if and only if it is Riemann-integrable on $I_1 = [a, c]$ and on $I_2 = [c, b]$. When that is the case, $\int_a^b f = \int_a^c f + \int_c^b f$.

Proof: we start by assuming that f is Riemann-integrable on I. Let $\varepsilon > 0$. According to the Riemann criterion, $\exists P_\varepsilon$ a partition of I such that $U(P_\varepsilon; f) - L(P_\varepsilon; f) < \varepsilon$. Now, set $P = P_\varepsilon \cup \{c\}$. Then P is a refinement of P_ε so that

$$U(P; f) - L(P; f) \leq U(P_\varepsilon; f) - L(P_\varepsilon; f) < \varepsilon.$$

Set $P_1 = P \cap I_1$ and $P_2 = P \cap I_2$. Then P_i is a partition of I_i, and

$$\begin{aligned} \varepsilon > U(P; f) - L(P; f) &\geq U(P_1; f) + U(P_2; f) - L(P_1; f) - L(P_2; f) \\ &= \big[U(P_1; f) - L(P_1; f)\big] + \big[U(P_2; f) - L(P_2; f)\big] \end{aligned}$$

Consequently, $U(P_i; f) - L(P_i; f) < \varepsilon$ for $i = 1, 2$ and f is Riemann-integrable on I_1 and I_2, according to the Riemann criterion.

Now assume that f is Riemann-integrable on I_1 and I_2. Let $\varepsilon > 0$. According to the Riemann criterion, for $i = 1, 2$, $\exists P_i$ a partition of I_i such that

$$U(P_i; f) + L(P_i; f) < \frac{\varepsilon}{2}.$$

Set $P = P_1 \cup P_2$. Then P is a partition of I. Furthermore,

$$\begin{aligned} U(P; f) - L(P; f) &= U(P_1; f) + U(P_2; f) - L(P_1; f) - L(P_2; f) \\ &= U(P_1; f) - L(P_1; f) + U(P_2; f) - L(P_2; f) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon, \end{aligned}$$

thus f is Riemann-integrable on I according the Riemann criterion.

Finally, let's assume that f is Riemann-integrable on I (and so on I_1, I_2), or vice-versa. Let P_1, P_2 be partitions of I_1, I_2, respectively, such that

$$U(P_i; f) - L(P_i; f) < \frac{\varepsilon}{2}, \quad i = 1, 2.$$

Set $P = P_1 \cup P_2$. Then P is a partition of I and

$$\begin{aligned} \int_a^b f &\leq U(P; f) = U(P_1; f) + U(P_2; f) \\ &< L(P_1; f) + L(P_2; f) + \varepsilon = \int_a^c f + \int_c^b f + \varepsilon. \end{aligned}$$

Similarly,

$$\int_a^b f \geq L(P_1; f) + L(P_2; f) > U(P_1; f) + U(P_2; f) - \varepsilon \geq \int_a^c f + \int_c^b f - \varepsilon$$

Since $\varepsilon > 0$ is arbitrary, $\int_a^b f = \int_a^c f + \int_c^b f$. ■

The next theorem is the crowning achievement of what has come before, combining results from previous chapters and sections. Its proof constitutes the first "real" example of what we might as well refer to as **analytical reasoning**.

Theorem 57 (Composition Theorem for Integrals)
Let $I = [a,b]$ and $J = [\alpha,\beta]$, $f : I \to \mathbb{R}$ Riemann-integrable on I, $\varphi : J \to \mathbb{R}$ continuous on J and $f(I) \subseteq J$. Then $\varphi \circ f : I \to \mathbb{R}$ is Riemann-integrable on I.

Proof: let $\varepsilon > 0$, $K = \sup\{|\varphi(x)| \mid x \in J\}$ (wich is guaranteed to exist according to the max/min theorem) and $\varepsilon' = \frac{\varepsilon}{b-a+2K}$.

Since φ is uniformly continuous on J (being continuous on a closed, bounded interval), $\exists \delta_\varepsilon > 0$ s.t.

$$|x-y| < \delta_\varepsilon,\ x,y,\in J \Longrightarrow |\varphi(x)-\varphi(y)| < \varepsilon'.$$

Without loss of generality, pick $\delta_\varepsilon < \varepsilon'$.

Since f is Riemann-integrable on I, there is a partition $P = \{x_0,\ldots,x_n\}$ of $I = [a,b]$ such that

$$U(P;f) - L(P;f) < \delta_\varepsilon^2$$

(according to Riemann's criterion).

We show that $U(P;\varphi\circ f) - L(P;\varphi\circ f) < \varepsilon$, and so that $\varphi \circ f$ is Riemann-integrable according to Riemann's criterion. On $[x_{i-1},x_i]$ for $i = 1,\ldots,n$, set

$$m_i = \inf\{f(x)\},\ M_i = \sup\{f(x)\},\ \tilde{m}_i = \inf\{\varphi(f(x))\},\ \tilde{M}_i = \sup\{\varphi(f(x))\}.$$

With those, set $A = \{i \mid M_i - m_i < \delta_\varepsilon\}$, $B = \{i \mid M_i - m_i \geq \delta_\varepsilon\}$.

- If $i \in A$, then

$$x,y \in [x_{i-1},x_i] \Longrightarrow |f(x)-f(y)| \leq M_i - m_i < \delta_\varepsilon,$$

  so $|\varphi(f(x)) - \varphi(f(y)| < \varepsilon'\ \forall x,y \in [x_{i-1},x_i]$. In particular, $\tilde{M}_i - \tilde{m}_i \leq \varepsilon'$.

- If $i \in B$, then

$$x,y \in [x_{i-1},x_i] \Longrightarrow |\varphi(f(x)) - \varphi(f(y))| \leq |\varphi(f(x))| + |\varphi(f(y))| \leq 2K.$$

  In particular, $\tilde{M}_i - \tilde{m}_i \leq 2K$, since $-K \leq \tilde{m}_i \leq \varphi(z) \leq \tilde{M}_i \leq K$ for all $z \in [x_{i-1},x_i]$.

Then

$$\begin{aligned}U(P;\varphi\circ f)-L(P;\varphi\circ f) &= \sum_{i=1}^{n}(\tilde{M}_i-\tilde{m}_i)(x_i-x_{i-1})\\ &= \sum_{i\in A}(\tilde{M}_i-\tilde{m}_i)(x_i-x_{i-1})+\sum_{i\in B}(\tilde{M}_i-\tilde{m}_i)(x_i-x_{i-1})\\ &\le \varepsilon'\sum_{i\in A}(x_i-x_{i-1})+2K\sum_{i\in B}(x_i-x_{i-1})\\ &\le \varepsilon'(b-a)+2K\sum_{i\in B}\frac{(M_i-m_i)}{\delta_\varepsilon}(x_i-x_{i-1})\\ &= \varepsilon'(b-a)+\frac{2K}{\delta_\varepsilon}\sum_{i=1}^{n}(M_i-m_i)(x_i-x_{i-1}).\end{aligned}$$

By earlier work in the proof, we have

$$\sum_{i=1}^{n}(M_i-m_i)(x_i-x_{i-1})\le U(P;f)-L(P;f)<\delta_\varepsilon^2,$$

so that

$$\begin{aligned}U(P;\varphi\circ f)-L(P;\varphi\circ f) &< \varepsilon'(b-a)+\frac{2K}{\delta_\varepsilon}\cdot\delta_\varepsilon^2\\ &= \varepsilon'(b-a)+2K\delta_\varepsilon<\varepsilon'(b-a)+2K\varepsilon'\\ &= \varepsilon'(b-a+2K)=\varepsilon,\end{aligned}$$

which completes the proof. ■

The proof of the composition theorem requires the intervals I and J to be closed, as the following example shows.

Example: let $f,\varphi:(0,1)\to\mathbb{R}$ be defined by $f(x)=x$ and $\varphi(x)=\frac{1}{x}$. Then f is Riemann-integrable on $(0,1)$, φ is continuous on $(0,1)$, but $\varphi\circ f:(0,1)\to\mathbb{R}$, $(\varphi\circ f)(x)=1/x$, is not Riemann-integrable on $(0,1)$. □

Note, however, that there are examples of functions defined on open intervals for which the conclusion of the composition theorem still holds.

Example: let $f,\varphi:(0,1)\to\mathbb{R}$ be defined by $f(x)=x$ and $\varphi(x)=x$. Then f is Riemann-integrable on $(0,1)$, φ is continuous on $(0,1)$, and $\varphi\circ f:(0,1)\to\mathbb{R}$, $(\varphi\circ f)(x)=x$, is Riemann-integrable on $(0,1)$. □

Theorem 57 is rather technical, but it can be used to show a variety of results.

Theorem 58
Let $I = [a,b]$ and $f, g : I \to \mathbb{R}$ be Riemann-integrable on I. Then fg and $|f|$ are Riemann-integrable on I, and $\left|\int_a^b f\right| \leq \int_a^b |f|$.

Proof: the function defined by $\varphi(t) = t^2$ is continuous. by the Composition theorem, $\varphi \circ (f+g) = (f+g)^2$ and $\varphi \circ (f-g) = (f-g)^2$ are both Riemann-integrable on I. But the product fg can be re-written as

$$fg = \frac{1}{4}\left[(f+g)^2 - (f-g)^2\right].$$

According to Theorem 55, fg is Riemann-integrable on I.

Now, consider the function defined by $\varphi(t) = |t|$. It is continuous, so $\varphi \circ f = |f|$ is Riemann-integrable on I according to the composition theorem.

Pick $c \in \{\pm 1\}$ such that $c\int_a^b f \geq 0$. Hence

$$\left|\int_a^b f\right| = c\int_a^b f = \int_a^b cf \leq \int_a^b |f|,$$

since $cf(x) \leq |f(x)|$ for all $x \in I$. ■

Note that even if the product of Riemann-integrable functions is itself Riemann-integrable there is no simple way to express $\int_a^b fg$ in terms of $\int_a^b f$ and $\int_a^b g$.

Given all that has come so far, we might suspect that the composition of Riemann-integrable functions is also Riemann-integrable. The following counter-example shows that this **need not be the case**.

Example: let $I = [0,1]$ and let $f : I \to \mathbb{R}$ be **Thomae's function**:

$$f(x) = \begin{cases} 1, & x = 0 \\ 1/n, & x = m/n \in \mathbb{Q},\ \text{gcd}(m,n) = 1 \\ 0, & x \notin \mathbb{Q} \end{cases}$$

It can be shown that f is Riemann-integrable on $[0,1]$ and that $\int_0^1 f = 0$. Consider the function $g : [0,1] \to \mathbb{R}$ defined by $g(x) \equiv 1$ on $(0,1]$ and $g(0) = 0$. Then g is Riemann-integrable on $[0,1]$, with $\int_0^1 g = 1$, but $g \circ f : [0,1] \to \mathbb{R}$ is the **Dirichlet function**, and is therefore not Riemann-integrable on $[0,1]$. □

4.2.3 Fundamental Theorem of Calculus

With Descartes' creation of analytical geometry, it became possible to find the **tangents** to curves that are **algebraically** described.[6] Fermat then showed the connection between that problem and the problem of finding the **maximum/minimum** of a (continuous) function. In the 1680s, Newton and Leibniz eventually discovered that computing the **area underneath a curve** is exactly the opposite of finding the tangent. **Calculus** provided a general framework to solve problems that had hitherto been very difficult to solve.[7] In this section, we study the connection between these concepts.

Theorem 59 (Fundamental Theorem of Calculus, 1st version)
Let $I = [a, b]$, $f : I \to \mathbb{R}$ be Riemann-integrable on I, and $F : I \to \mathbb{R}$ be such that F is continuous on I and differentiable on (a, b). If $F'(x) = f(x)$ for all $x \in (a, b)$, then $\int_a^b f = F(b) - F(a)$.

Proof: let $\varepsilon > 0$. Since f is Riemann-integrable on I, $\exists P_\varepsilon \in \mathcal{P}(I)$ such that

$$U(P_\varepsilon; f) - L(P_\varepsilon; f) < \varepsilon.$$

Applying the mean value theorem to F on $[x_{i-1}, x_i]$ for each $1 \leq i \leq n$, we conclude that $\exists t_i \in (x_{i-1}, x_i)$ such that

$$\frac{F(x_i) - F(x_{i-1})}{x_i - x_{i-1}} = F'(t_i) = f(t_i), \quad 1 \leq i \leq n.$$

Let $\tilde{m}_i = \inf\{f(x) \mid x \in [x_{i-1}, x_i]\}$, $\tilde{M}_i = \sup\{f(x) \mid x \in [x_{i-1}, x_i]\}$ for $1 \leq i \leq n$. Then

$$L(P_\varepsilon; f) = \sum_{i=1}^n \tilde{m}_i(x_i - x_{i-1}) \leq \sum_{i=1}^n f(t_i)(x_i - x_{i-1}) = \sum_{i=1}^n (F(x_i) - F(x_{i-1})) = F(b) - F(a),$$

and, similarly, $U(P_\varepsilon; f) \geq F(b) - F(a)$. Then $L(P_\varepsilon; f) \leq F(b) - F(a) \leq U(P_\varepsilon; f)$ for all $\varepsilon > 0$. Since we have

$$L(P_\varepsilon; f) \leq \int_a^b f \leq U(P_\varepsilon; f)$$

and $U(P_\varepsilon; f) - L(P_\varepsilon; f) < \varepsilon$, for all $\varepsilon > 0$, we must also have

$$\left| \int_a^b f - (F(b) - F(a)) \right| < \varepsilon, \quad \text{for all } \varepsilon > 0,$$

so that $\int_a^b f = F(b) - F(a)$. ∎

[6] That is, curves who can be expressed in $\mathbb{R}^2$ as $f(x, y) = 0$ for algebraic functions f.
[7] And even then, only in specific circumstances.

This classical calculus result is quite useful in applications,[8] as is its cousin.

Theorem 60 (Fundamental Theorem of Calculus, 2nd version)
Let $I=[a,b]$, $f:I\to\mathbb{R}$ be Riemann-integrable on I. Define a function $F:I\to\mathbb{R}$ by $F(x)=\int_a^x f$. Then F is continuous on I. Furthermore, if f is continuous at $c\in(a,b)$, then F is differentiable at c and $F'(c)=f(c)$.

Proof: since f is Riemann-integrable on I, then f is bounded on I. Let $K>0$ be such that $|f(x)|<K$ for all $x\in I$. Let $x\in I$ and $\varepsilon>0$. Set $\delta_\varepsilon=\frac{\varepsilon}{K}$. Then whenever $|x-y|<\delta_\varepsilon=\frac{\varepsilon}{K}$ and $y\in I$, we have

$$|F(y)-F(x)|=\left|\int_a^y f-\int_a^x f\right|=\left|\int_x^y f\right|\leq K|x-y|<\varepsilon.$$

Then F is uniformly continuous on I, and so is continuous on I. Now assume that f is continuous at c and let $\varepsilon>0$. Then $\exists\delta_\varepsilon>0$ such that $|f(x)-f(c)|<\varepsilon$ whenever $|x-c|<\delta_\varepsilon$ and $x\in I$.

Thus, if $0\leq|h|=|x-c|<\delta_\varepsilon$ and $x\in I$, we have

$$\begin{aligned}\left|\frac{F(c+h)-F(c)}{h}-f(c)\right|&=\left|\frac{1}{h}\int_a^{c+h} f-\frac{1}{h}\int_a^c f-f(c)\right|\\&=\left|\frac{1}{h}\int_c^{c+h} f-\frac{1}{h}\int_c^{c+h} f(c)\right|=\left|\frac{1}{h}\int_c^{c+h}(f-f(c))\right|\\&\leq\frac{1}{|h|}\left|\int_c^{c+h}(f-f(c))\right|<\frac{1}{|h|}\cdot\varepsilon\left|\int_c^{c+h}1\right|=\frac{1}{|h|}\cdot\varepsilon|h|=\varepsilon,\end{aligned}$$

which is to say, $F'(c)=f(c)$. ■

The first version of the fundamental theorem of calculus provides a justification of the **method used to evaluate definite integrals** in calculus; the second version, which allows the upper bound of the Riemann integral to vary, provides a basis for finding antiderivatives.

Let $I=[a,b]$ an $f:I\to\mathbb{R}$. An **antiderivative** of f on I is a differentiable function $F:I\to\mathbb{R}$ such that $F'(x)=f(x)$ for all $x\in I$. If f is Riemann-integrable on I, the function $F:I\to\mathbb{R}$ defined by $F(x)=\int_a^x f$ for $x\in I$ is the **indefinite integral of f on I.** If f is Riemann-integrable on I and if F is an antiderivative of f on I, then

$$\int_a^b f=F(b)-F(a).$$

However, Riemann-integrable functions on I may not have antiderivatives on I (such as the signum and Thomae's functions), and functions with antiderivatives may not be Riemann-integrable on I (such as the reciprocal of the square root function on $[0,1]$).

[8]We will see in Chapter 14 that the Theorem 59 (1st version) is a special case of a more general result.

If f is Riemann-integrable on I, then $F(x) = \int_a^x f$ exists. Moreover, if f is continuous on I, than F is an antiderivative of f on I, since $F'(x) = f(x)$ for all $x \in I$. Continuous functions thus **always have antiderivatives**.[9]

But if f is not continuous on I, the indefinite integral F may not be an antiderivative of f on I – it may fail to be differentiable at certain points of I, or F' may exists but be different from f at various points of I.

4.2.4 Evaluation of Integrals

We complete this chapter by presenting some common methods used to evaluate integrals, and the proof for two of them.

Theorem 61 (INTEGRATION BY PARTS)
Let $f, g : [a,b] \to \mathbb{R}$ both be Riemann-integrable on $[a,b]$, with antiderivatives $F, G : [a,b] \to \mathbb{R}$, respectively. Then

$$\int_a^b Fg = F(b)G(b) - F(a)G(a) - \int_a^b fG.$$

Proof: Let $H : [a,b] \to \mathbb{R}$ be defined by $H = FG$. As F and G are both differentiable, so is H: $H' = F'G + FG' = fG + Fg$.

Then $\int_a^b H' = H(b) - H(a)$, so

$$\int_a^b (fG + Fg) = F(b)G(b) - F(a)G(a) \Longrightarrow \int_a^b Fg = H(b) - H(a) - \int_a^b fG.$$

This completes the proof. ■

Theorem 62 (FIRST SUBSTITUTION THEOREM)
Let $J = [\alpha, \beta]$, and $\varphi \to \mathbb{R}$ be a function with a continuous derivative on J. If $f : I \to \mathbb{R}$ is continuous on $I = [a,b] \supseteq \varphi(J)$, then

$$\int_\alpha^\beta (f \circ \varphi)\varphi' = \int_{\varphi(\alpha)}^{\varphi(\beta)} f.$$

Proof: Since f is continuous on I, it is Riemann-integrable on I and so we can define a function $F : I \to \mathbb{R}$ through

$$F(x) = \int_{\varphi(\alpha)}^x f, \quad x \in I.$$

By construction F is continuous and differentiable on I. Furthermore, $F' = f$ on I, according to the second version of the fundamental theorem of calculus.

[9] Even if they can't be expressed using elementary functions.

Define $H : J \to \mathbb{R}$ by $H = F \circ \varphi$. Then H is differentiable on I, being the composition of two differentiable functions on I, and $H' = (F' \circ \varphi)\varphi' = (f' \circ \varphi)\varphi'$ is Riemann-integrable since $\varphi, f \circ \varphi$ are Riemann-integrable (being continuous) on I, according to Theorem 58. The first version of the Fundamental Theorem of Calculus then yields

$$\int_\alpha^\beta (f \circ \varphi)\varphi' = \int_\alpha^\beta H' = H(\beta) - H(\alpha) = F(\varphi(\beta)) - F(\varphi(\alpha)) = \int_{\varphi(\alpha)}^{\varphi(\beta)} f,$$

which completes the proof. ∎

The proofs of the last three theorems are left as an exercise.

Theorem 63 (Second Substitution Theorem)
Let $J = [\alpha, \beta]$, and $\varphi \to \mathbb{R}$ be a function with a continuous derivative on J and such that $\varphi' \neq 0$ on J. Let $I = [a, b] \supseteq \varphi(J)$, and $\psi : I \to \mathbb{R}$ be the inverse of φ (which exists as φ is montoone). If $f : I \to \mathbb{R}$ is continuous on I, then

$$\int_\alpha^\beta f \circ \varphi = \int_{\varphi(\alpha)}^{\varphi(\beta)} f\psi'.$$

Theorem 64 (Mean Value Theorem for Integrals)
Let $I = [a, b]$, $f : I \to \mathbb{R}$ be continuous on I, and $p : I \to \mathbb{R}$ be Riemann-integrable on I, with $p \geq 0$ on I. Then $\exists c \in (a, b)$ such that

$$\int_a^b fp = f(c) \int_a^b p.$$

Theorem 65 (Squeeze Theorem for Integrals)
Let $I = [a, b]$ and $f \leq g \leq h : I \to \mathbb{R}$ be bounded on I. If f, h are Riemann-integrable on I with $\int_a^b f = \int_a^b h$, then g is Reimann-integrable on I and $\int_a^b g = \int_a^b f = \int_a^b h$.

4.3 Solved Problems

1. Use the definition to find the derivative of the function defined by $g(x) = \frac{1}{x}$, $x \in \mathbb{R}$, $x \neq 0$.

   Solution: from calculus, we "know" that $g'(x) = -\frac{1}{x^2}$. Let $c \in \mathbb{R}$ s.t. $c \neq 0$. Set $a_c = \frac{c}{2}$ and $b_c = \frac{3c}{2}$. Clearly, if $c > 0$, $0 < a_c < c < b_c$, whereas $b_c < c < a_c < 0$ if $c < 0$. In both cases, $\frac{1}{|x|} \leq \frac{1}{|a_c|}$ whenever x lies between a_c and b_c. We restrict g on the interval between a_c and b_c (denote this interval by A).

   Let $\varepsilon > 0$ and set $\delta_\varepsilon = |a_c|c^2\varepsilon$. Then whenever $0 < |x - c| < \delta_\varepsilon$ and $x \in A$, we have

   $$\left| \frac{\frac{1}{x} - \frac{1}{c}}{x - c} + \frac{1}{c^2} \right| = \left| \frac{c - x}{xc(x - c)} + \frac{1}{c^2} \right| = \left| \frac{1}{c^2} - \frac{1}{xc} \right| = \frac{|x - c|}{|x|c^2} \leq \frac{|x - c|}{|a_c|c^2} < \frac{\delta_\varepsilon}{|a_c|c^2} = \varepsilon,$$

   which validates our calculus guess. □

2. Prove that the derivative of an even differentiable function is odd, and vice-versa.

   Proof: if f is even, then $f(x) = f(-x)$ for all $x \in \mathbb{R}$. Let $g(x) = f(-x)$. Then g is differentiable by the chain rule and $f(x) = g(x)$ for all $x \in \mathbb{R}$. Furthermore,

   $$f'(x) = g'(x) = (f(-x))' = f'(-x) \cdot -1,$$

   that is, $-f'(-x) = f'(-x)$, or f' is odd. The other statement is proved similarly. ■

3. Let $a > b > 0$ and $n \in \mathbb{N}$ with $n \geq 2$. Show that $a^{1/n} - b^{1/n} < (a - b)^{1/n}$.

   Proof: consider the continuous function $f : [1, \infty) \to \mathbb{R}$ defined by $f(x) = x^{1/n} - (x - 1)^{1/n}$, whose derivative is

   $$f'(x) = \frac{1}{n} \left(x^{\frac{1-n}{n}} - (x - 1)^{\frac{1-n}{n}} \right).$$

   Now,

   $$0 \leq x - 1 < x, \quad \forall x \geq 1 \Longrightarrow 0 \leq (x - 1)^n < x^n, \quad \forall x \geq 1, n \geq 2$$
   $$\therefore 0 \leq (x - 1)^{\frac{n}{n-1}} < x^{\frac{n}{n-1}}, \quad \forall x \geq 1, n \geq 2$$

   and so

   $$\frac{1}{x^{\frac{n}{n-1}}} < \frac{1}{(x - 1)^{\frac{n}{n-1}}},$$

   or $x^{\frac{1-n}{n}} < (x - 1)^{\frac{1-n}{n}}$ for all $x \geq 1$, $n \geq 2$.

   Hence $f'(x) < 0$ for all $x \in [1, \infty)$, that is f is strictly decreasing over $[1, \infty)$. But $f(\frac{a}{b}) < f(1)$, as $\frac{a}{b} > 1$. But

   $$f\left(\frac{a}{b}\right) = \left(\frac{a}{b}\right)^{\frac{1}{n}} - \left(\frac{a}{b} - 1\right)^{\frac{1}{n}} = \frac{1}{b^{\frac{1}{n}}} \left(a^{\frac{1}{n}} - (a - b)^{\frac{1}{n}} \right)$$

and $f(1) = 1$, so

$$\frac{1}{b^{\frac{1}{n}}}\left(a^{\frac{1}{n}} - (a-b)^{\frac{1}{n}}\right) < 1,$$

that is $a^{\frac{1}{n}} - (a-b)^{\frac{1}{n}} < b^{\frac{1}{n}}$, which completes the proof. ∎

4. Let $f : [a,b] \to \mathbb{R}$ be continuous on $[a,b]$ and differentiable on (a,b). Show that if $\lim_{x\to a} f'(x) = A$, then $f'(a)$ exists and equals A.

   Proof: let $x \in (a,b)$. By the Mean Value Theorem, $\exists c_x \in (a,x)$ s.t.

$$\frac{f(x)-f(a)}{x-a} = f'(c_x).$$

   When $x \to a$, $c_x \to a$ (indeed, let $\varepsilon > 0$ and set $\delta_\varepsilon = \varepsilon$; then $|c_x - a| < |x-a| < \delta_\varepsilon = \varepsilon$ whenever $0 < |x-a| < \delta_\varepsilon$). Then

$$\lim_{x\to a} f'(c_x) = \lim_{c_x\to a} f'(c_x) = A$$

   by hypothesis. Hence $\lim_{x\to a} f'(x)$ exists and so

$$f'(a) = \lim_{x\to a}\frac{f(x)-f(a)}{x-a} = \lim_{x\to a} f'(x) = A$$

   exists. ∎

5. If $x > 0$, show $1 + \frac{1}{2}x - \frac{1}{8}x^2 \leq \sqrt{1+x} \leq 1 + \frac{1}{2}x$.

   Proof: let $x_0 = 0$ and $f(x) = \sqrt{1+x}$. According to Taylor's theorem, since f is C^3 when $x > 0$, $f(x) = P_1(x) + R_1(x)$ and $f(x) = P_2(x) + R_2(x)$, where

$$\begin{aligned}
P_1(x) &= f(x_0) + f'(x_0)(x-x_0) = \sqrt{1+0} + \frac{1}{2\sqrt{1+0}}x = 1 + \frac{1}{2}x \\
P_2(x) &= P_1(x) + \frac{f''(x_0)}{2}(x-x_0)^2 = 1 + \frac{1}{2}x - \frac{1}{8\sqrt[3]{1+0}}x^2 = 1 + \frac{1}{2}x - \frac{1}{8}x^2 \\
R_1(x) &= \frac{f''(c_1)}{2}(x-x_0)^2 = -\frac{1}{8\sqrt[3]{1+c_1}}x^2, \quad \text{for some } c_1 \in [0,x] \\
R_2(x) &= \frac{f'''(c_2)}{6}(x-x_0)^3 = \frac{3}{48\sqrt[5]{1+c_2}}x^3, \quad \text{for some } c_2 \in [0,x].
\end{aligned}$$

   When $x > 0$, $R_1(x) \leq 0$ and $R_2(x) \geq 0$, so $P_2(x) \leq f(x) \leq P_1(x)$. ∎

6. Let $a \in \mathbb{R}$ and $f : \mathbb{R} \to \mathbb{R}$ be defined by

$$f(x) = \begin{cases} x^2 & \text{if } x \geq 0, \\ ax & \text{if } x < 0. \end{cases}$$

   For which values of a is f differentiable at $x = 0$? For which values of a is f continuous at $x = 0$?

Solution: We have

$$f'_+(0) = \lim_{x\to 0^+} \frac{f(x)-f(0)}{x-0} = \lim_{x\to 0^+} \frac{x^2}{x} = \lim_{x\to 0^+} x = 0$$

and

$$f'_-(0) = \lim_{x\to 0^-} \frac{f(x)-f(0)}{x-0} = \lim_{x\to 0^-} \frac{ax}{x} = \lim_{x\to 0^+} a = a.$$

Thus, f is differentiable at $x=0$ if and only if $a=0$.

Since both x^2 and ax are continuous functions, we have

$$\lim_{x\to 0^+} f(x) = \lim_{x\to 0^+} x^2 = 0 = f(0) = 0 = \lim_{x\to 0^-} ax = \lim_{x\to 0^-} f(x)$$

and the the function f is continuous at $x=0$ for all values of a. □

7. Let $f : [a,b] \to \mathbb{R}$ be continuous on $[a,b]$ and differentiable on (a,b). Show that f is Lipschitz if and only if f' is bounded on (a,b).

Proof: Suppose that f satisfies the Lipschitz condition on $[a,b]$ with constant M. Then, for all $x_0 \in (a,b)$, we have

$$\left|\frac{f(x)-f(x_0)}{x-x_0}\right| \le M \qquad \forall\, x \in (a,b)\setminus\{x_0\}.$$

Thus

$$|f'(x_0)| = \left|\lim_{x\to x_0} \frac{f(x)-f(x_0)}{x-x_0}\right| = \lim_{x\to x_0} \left|\frac{f(x)-f(x_0)}{x-x_0}\right| \le M,$$

where we used the fact that the absolute value function is continuous to pull the limit out of the absolute value. So the derivative of f is bounded on (a,b).

Now assume that $|f'(x)| \le M$ for all $x \in (a,b)$. Let $x,y \in [a,b]$, $x<y$. Applying the Mean Value Theorem to f on the interval $[x,y]$ yields the existence of $c \in (x,y)$ such that

$$\frac{f(y)-f(x)}{y-x} = f'(c).$$

Thus

$$\left|\frac{f(x)-f(y)}{x-y}\right| \le M \Longrightarrow |f(x)-f(y)| \le M|x-y|.$$

This completes the proof. ■

8. Prove that $\int_0^1 g = \frac{1}{2}$ if

$$g(x) = \begin{cases} 1 & x \in (\frac{1}{2}, 1] \\ 0 & x \in [0, \frac{1}{2}] \end{cases}.$$

Is that still true if $g(\frac{1}{2}) = 7$ instead?

Proof: let $\varepsilon > 0$ and define the partition $P_\varepsilon = \{0, \frac{1}{2} - \varepsilon, \frac{1}{2} + \varepsilon, 1\}$. Since g is bounded on $[0,1]$, $L(g) \leq U(g)$ exist and

$$L(g) \geq L(P_\varepsilon; g) = \frac{1}{2} - \varepsilon \quad \text{and} \quad U(g) \leq U(P_\varepsilon; g) = \frac{1}{2} + \varepsilon.$$

Hence

$$\frac{1}{2} - \varepsilon \leq L(g) \leq U(g) \leq \frac{1}{2} + \varepsilon, \quad \text{for all } \varepsilon > 0.$$

Since $\varepsilon > 0$ is arbitrary, then $\frac{1}{2} \leq L(g) \leq U(g) \leq \frac{1}{2}$; by definition, g is Riemann-integrable on $[0,1]$ and $L(g) = U(g) = \int_0^1 g = \frac{1}{2}$.

If instead $g(1/2) = 7$, the exact same work as above yields

$$\frac{1}{2} - \varepsilon \leq L(g) \leq U(g) \leq \frac{1}{2} + 13\varepsilon, \quad \text{for all } \varepsilon > 0.$$

Since $\varepsilon > 0$ is arbitrary, then $\frac{1}{2} \leq L(g) \leq U(g) \leq \frac{1}{2}$; by definition, g is also Riemann-integrable on $[0,1]$ and $L(g) = U(g) = \int_a^b f = \frac{1}{2}$. ■

9. Let $f : [a,b] \to \mathbb{R}$ be bounded and such that $f(x) \geq 0, \forall x \in [a,b]$. Show $L(f) \geq 0$.

Proof: as f is bounded on $[a,b]$, $L(f)$ exists and the set

$$\{f(x) \mid x \in [a,b]\} \neq \varnothing$$

is bounded below. By completeness of $\mathbb{R}$, $m_1 = \inf\{f(x) \mid x \in [a,b]\}$ exists. Furthermore, $m_1 \geq 0$ since $f(x) \geq 0$ for all $x \in [a,b]$.

Let $P = \{x_0, x_1\} = \{a, b\}$ be the trivial partition of $[a,b]$. Then

$$L(f) \geq L(P; f) = m_1(b-a) \geq 0,$$

which completes the proof. ■

10. Let $f : [a,b] \to \mathbb{R}$ be increasing on $[a,b]$. If P_n partitions $[a,b]$ into n equal parts, show that

$$0 \leq U(P_n; f) - \int_a^b f \leq \frac{f(b) - f(a)}{n}(b-a).$$

Proof: as f is increasing, it is monotone and thus Riemann-integrable by Theorem 53. Then $L(f) = U(f) = \int_a^b f$. Let

$$P_n = \{x_i = a + i\tfrac{b-a}{n} \mid i = 0, \ldots, n\}$$

be the partition of $[a,b]$ into n equal sub-intervals. By definition, $L(P_n; f) \leq \int_a^b f$ and $U(P_n; f) \geq \int_a^b f$. Then

$$U(P_n; f) - L(P_n; f) \geq U(P_n; f) - \int_a^b f \geq \int_a^b f - \int_a^b f = 0.$$

In particular, $U(P_n;f) - \int_a^b f \geq 0$. As f is increasing on $[a,b]$,

$$M_i = \sup_{[x_{i-1},x_i]}\{f(x)\} = f(x_i), \quad m_i = \inf_{[x_{i-1},x_i]}\{f(x)\} = f(x_{i-1}), \quad \text{and}$$

$$\begin{aligned} U(P_n;f) - L(P_n;f) &= \sum_{i=1}^{n}(M_i - m_i)(x_i - x_{i-1}) \\ &= \frac{b-a}{n}\sum_{i=1}^{n}(f(x_i) - f(x_{i-1})) = \frac{b-a}{n}(f(b) - f(a)). \end{aligned}$$

Since $L(P_n;f) \leq \int_a^b f$, then

$$\frac{b-a}{n}(f(b) - f(a)) = U(P_n;f) - L(P_n;f) \geq U(P_n;f) - \int_a^b f \geq 0,$$

which completes the proof. ■

11. Let $f : [a,b] \to \mathbb{R}$ be an integrable function and let $\varepsilon > 0$. If P_ε is the partition whose existence is asserted by the Riemann Criterion, show that $U(P;f) - L(P;f) < \varepsilon$ for all refinement P of P_ε.

    Proof: let P be a refinement of P_ε. Then $U(P_{\varepsilon;f}) \geq U(P;f)$ and $L(P_\varepsilon;f) \leq L(P;f)$, and so

    $$U(P_\varepsilon;f) \geq U(P;f) \geq L(P;f) \geq L(P_\varepsilon;f).$$

    By the Riemann Criterion, $U(P_\varepsilon;f) < \varepsilon + L(P_\varepsilon;f)$. Then

    $$\varepsilon + L(P;f) \geq \varepsilon + L(P_\varepsilon;f) > U(P_\varepsilon;f) \geq U(P;f),$$

    i.e. $\varepsilon + L(P;f) > U(P;f)$, which completes the proof. ■

12. Let $a > 0$ and $J = [-a,a]$. Let $f : J \to \mathbb{R}$ be bounded and let $\mathcal{P}^*$ be the set of symmetric partitions of J that contain 0. Show $L(f) = \sup\{L(P;f) \mid P \in \mathcal{P}^*\}$.

    Proof: let $\alpha = \sup\{L(P;f) \mid P \in \mathcal{P}^*\}$. By definition,

    $$\alpha \leq L(f) = \sup\{L(P;f) \mid P \text{ is a partition of } [-a,a]\}.$$

    Let $\varepsilon > 0$ and $P_\varepsilon = \{x_0, x_1, \ldots, x_n\}$ be a partition of $[-a,a]$ such that

    $$L(f) - \varepsilon < L(P_\varepsilon;f) \leq L(f).$$

    Such a partition exists as $L(f) - \varepsilon$ is not the supremum of the aforementioned set.

    Consider the set $\{0, \pm x_0, \ldots, \pm x_n\}$. Eliminate all the repetitions from this set and re-order its elements. Denote the new set by Q_ε.

    Then Q_ε is a refinement of P_ε and $Q_\varepsilon \in \mathcal{P}^*$; so $\alpha \geq L(Q_\varepsilon;f)$, and

    $$L(f) - \varepsilon < L(P_\varepsilon;f) \leq L(Q_\varepsilon;f) \leq \alpha \leq L(f),$$

    as $\varepsilon > 0$ is arbitrary, $L(f) = \alpha$. ■

13. Let $a > 0$ and $J = [-a, a]$. Let f be integrable on J. If f is even (i.e. $f(-x) = f(x)$ for all x), show that
$$\int_{-a}^{a} f = 2\int_{0}^{a} f.$$
If f is odd (i.e. $f(-x) = -f(x)$ for all x), show that $\int_{-a}^{a} f = 0.$

Proof: as f is integrable over $[-a, a]$, Theorem 56 implies that f is integrable over $[0, a]$. If f is even, let $P \in \mathcal{P}^*$. There is a partition $\tilde{P}$ of $[0, a]$ s.t. $L(P; f) = 2L(\tilde{P}; f)$ and vice-versa. Indeed, let
$$P = \{x_{-n}, \ldots, x_{-1}, x_0, x_1, \ldots, x_n\},$$
where $x_0 = 0$ and $x_{-i} = -x_i$ for all $i = 1, \ldots, n$. Then $P \in \mathcal{P}^*$.

Let $m_i = \inf\{f(x) \mid x \in [x_{i-1}, x_i]\}$, for $i = -n-1, \ldots, 0, \ldots, n$. Since f is even, $m_i = m_{-i+1}$ for $i = -n-1, \ldots, 0, \ldots, n$. Then
$$L(P; f) = \sum_{i=-n-1}^{0} m_i(x_i - x_{i-1}) + \sum_{i=1}^{n} m_i(x_i - x_{i-1}) = 2\sum_{i=1}^{n} m_i(x_i - x_{i-1}) = L(\tilde{P}; f),$$
where $\tilde{P}$ is a partition of $[0, a]$.

This, combined with the previous solved problem, yields
$$\begin{aligned}\int_{-a}^{a} f &= \sup\{L(P; f) \mid P \in \mathcal{P}^*\} = \sup\{2L(\tilde{P}; f) \mid \tilde{P} \text{ is a partition of } [0, a]\} \\ &= 2\sup\{L(\tilde{P}; f) \mid \tilde{P} \text{ is a partition of } [0, a]\} = 2\int_{0}^{a} f.\end{aligned}$$

If f is odd, consider the function $h : \mathbb{R} \to \mathbb{R}$ given by
$$h(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ -1 & \text{if } x < 0 \end{cases}.$$

The product fh is an even function, so
$$2\int_{0}^{a} f = 2\int_{0}^{a} hf = \int_{-a}^{a} hf = \int_{-a}^{0} hf + \int_{0}^{a} hf = \int_{-a}^{0} -f + \int_{0}^{a} f,$$
and so $\int_0^a f = \int_{-a}^0 -f = -\int_{-a}^0 f$. Then
$$\int_{-a}^{a} f = \int_{-a}^{0} f + \int_{0}^{a} f = -\int_{0}^{a} f + \int_{0}^{a} f = 0,$$
which completes the proof. ■

14. Give an example of a function $f : [0, 1] \to \mathbb{R}$ that is not integrable on $[0, 1]$, but such that $|f|$ is integrable on $[0, 1]$.

Solution: here is one example – $f : [0,1] \to \mathbb{R}$, defined by $f(x) = -1$ if $x \notin \mathbb{Q}$ and $f(x) = 1$ if $x \in \mathbb{Q}$. The proof that f is not Riemann-integrable on $[0,1]$ is similar to the proof that the Dirichlet function is not Rimeann-integrable on $[0,1]$. □

15. Let $f : [a,b] \to \mathbb{R}$ be integrable on $[a,b]$. Show $|f|$ is integrable on $[a,b]$ directly.

Proof: let $\varepsilon > 0$. By the Riemann criterion, there exists a partition $P_\varepsilon = \{x_0, \ldots, x_n\}$ of $[a,b]$ such that $U(P_\varepsilon; f) - L(P_\varepsilon; f) < \varepsilon$.

For all $i = 1, \ldots, n$, let

$$M_i = \text{sup}\{f(x) \mid x \in [x_{i-1}, x_i]\} \quad \text{and} \quad m_i = \text{inf}\{f(x) \mid x \in [x_{i-1}, x_i]\}.$$

For all $i = 1, \ldots, n$, we then have $|f(x) - f(y)| \leq M_i - m_i$ on $[x_{i-1}, x_i]$. As

$$||f(x)| - |f(y)|| \leq |f(x) - f(y)| \leq M_i - m_i \quad \text{for all } x, y \in [x_{i-1}, x_i],$$

we have $\tilde{M}_i - \tilde{m}_i \leq M_i - m_i$, where

$$\tilde{M}_i = \text{sup}\{|f(x)| \mid x \in [x_{i-1}, x_i]\} \quad \text{and} \quad \tilde{m}_i = \text{inf}\{|f(x)| \mid x \in [x_{i-1}, x_i]\}$$

for all $i = 1, \ldots, n$. Then

$$\begin{aligned} U(P_\varepsilon; |f|) - L(P_\varepsilon; |f|) &= \sum_{i=1}^{n} \left(\tilde{M}_i - \tilde{m}_i\right)(x_i = x_{i-1}) \\ &\leq \sum_{i=1}^{n} (M_i - m_i)(x_i = x_{i-1}) = U(P_\varepsilon; |f|) - L(P_\varepsilon; |f|) < \varepsilon. \end{aligned}$$

According to the Riemann criterion, $|f|$ is thus integrable on $[a,b]$. ■

16. If f is integrable on $[a,b]$ and $0 \leq m \leq f(x) \leq M$ for all $x \in [a,b]$, show that

$$m \leq \left[\frac{1}{b-a} \int_a^b f^2\right]^{1/2} \leq M.$$

Proof: by hypothesis, $m^2 \leq f^2(x) \leq M^2$ for all $x \in [a,b]$. As f is integrable on $[a,b]$, so is f^2, by Theorem 58.

Then

$$\int_a^b m^2 \leq \int_a^b f^2 \leq \int_a^b M^2$$

by the squeeze theorem for integrals and so

$$m^2(b-a) \leq \int_a^b f^2 \leq M^2(b-a).$$

We obtain the result by re-arranging the terms and extracting square roots. ■

17. If f is continuous on $[a,b]$ and $f(x) \geq 0$ for all $x \in [a,b]$, show there exists $c \in [a,b]$ such that
$$f(c) = \left[\frac{1}{b-a}\int_a^b f^2\right]^{1/2}.$$

Proof: by the max/min theorem, $\exists x_0, x_1 \in [a,b]$ such that
$$m = \inf_{[a,b]}\{f(x)\} = f(x_0),\ M = \sup_{[a,b]}\{f(x)\} = f(x_1).$$
By the preceding solved problem, we then have
$$f(x_0) \leq \left[\frac{1}{b-a}\int_a^b f^2\right]^{1/2} \leq f(x_1).$$
As f is continuous on $[x_0, x_1]$ (or $[x_1, x_0]$), the intermediate value theorem states $\exists c \in [a,b]$ such that
$$f(c) = \left[\frac{1}{b-a}\int_a^b f^2\right]^{1/2},$$
which completes the proof. ■

18. If f is continuous on $[a,b]$ and $f(x) > 0$ for all $x \in [a,b]$, show that $\frac{1}{f}$ is integrable on $[a,b]$.

Proof: since f is continuous on $[a,b]$ it is integrable on $[a,b]$; by Theorem 36, since f is continuous and $[a,b]$ is a closed bounded interval, then $f([a,b]) = [m,M]$ is also closed bounded interval. Furthermore, $0 < m \leq M$ since $f(x) > 0$ for all $x \in [a,b]$.

Let $\varphi : [m,M] \to \mathbb{R}$ be defined by $\varphi(t) = \frac{1}{t}$. Then φ is continuous and bounded on $[m,M]$ and so $\varphi \circ f : [a,b] \to \mathbb{R}$, defined by $\varphi(f(x)) = \frac{1}{f(x)}$ is integrable on $[a,b]$, by Theorem 57. ■

19. Let f be continuous on $[a,b]$. Define $H : [a,b] \to \mathbb{R}$ by
$$H(x) = \int_x^b f \quad \text{for all } x \in [a,b].$$
Find $H'(x)$ for all $x \in [a,b]$.

Proof: define $F(x) = \int_a^x f$. Since f is continuous, F is differentiable and the fundamental theorem of calculus (2nd version) yields $F'(x) = f(x)$ for all $x \in [a,b]$. Then, by the additivity theorem, we have:
$$F(x) + H(x) = \int_a^x f + \int_x^b f = \int_a^b f.$$
In particular,
$$H(x) = \int_a^b f - F(x).$$
As F is differentiable, $\int_a^b f - F(x)$ is also differentiable; so is H since $H'(x) = 0 - F'(x) = -f(x)$. ■

20. Suppose $f:[0,\infty)\to\mathbb{R}$ is continuous and $f(x)\neq 0$ for all $x>0$. If

$$(f(x))^2 = 2\int_0^x f \quad \text{for all } x>0,$$

show that $f(x)=x$ for all $x\geq 0$.

Proof: as f is continuous, $F(x)=\int_0^x f$ is continuous; the fundamental theorem of calculus (2nd version) then yields $F'(x)=f(x)$ for all $x\in[0,\infty)$.

Now, either $f(x)>0$ for all $x>0$ or $f(x)<0$ for all $x>0$ – otherwise f admits a root $c>0$ by the intermediate value theorem, which would contradict $f(x)\neq 0$ $\forall x>0$.

But

$$F(x)=\int_0^x f = \frac{(f(x))^2}{2} > 0 \quad \text{for all } x>0,$$

so $\int_0^x f>0$ for all $x>0$, which is to say that $f>0$ for all $x>0$ – otherwise, $\int_0^x f\leq\int_0^x 0=0$, which contradicts one of the above inequalities.

By construction,

$$\frac{(f(0))^2}{2} = F(0) = \int_0^0 f = 0,$$

that is, $f(0)=0$. Now, let $c>0$. By hypothesis, $F'(c)=f(c)>0$. Furthermore, $F(c)=\frac{(f(c))^2}{2}$. As f is continuous at c,

$$\lim_{x\to c}\frac{1}{2}\left(f(x)+f(c)\right)=f(c).$$

Thus we have:

$$1=\frac{F'(c)}{f(c)}=\frac{\displaystyle\lim_{x\to c}\frac{F(x)-F(c)}{x-c}}{\displaystyle\lim_{x\to c}\frac{1}{2}\left(f(x)+f(c)\right)}=\lim_{x\to c}\frac{(f(x))^2-(f(c))^2}{(x-c)\left(f(x)+f(c)\right)}$$
$$=\lim_{x\to c}\frac{\left(f(x)-f(c)\right)\left(f(x)+f(c)\right)}{(x-c)\left(f(x)+f(c)\right)}=\lim_{x\to c}\frac{f(x)-f(c)}{x-c}=f'(c).$$

Then, the function f is differentiable and $f'(c)=1$ for all $c>0$. By the fundamental theorem of calculus (1st version),

$$\int_0^x f' = f(x)-f(0)=f(x)-0=f(x)$$

for all $x\geq 0$. As $\int_0^x f'=\int_0^x 1 = x-0=x$, this completes the proof (which, incidentally, is one of my favourite analysis proofs). ■

21. Let $f,g:[a,b]\to\mathbb{R}$ be continuous and such that

$$\int_a^b f=\int_a^b g.$$

Show that there exists $c\in[a,b]$ such that $f(c)=g(c)$.

Proof: as f and g are continuous, the functions

$$F(x)=\int_a^x f \quad \text{and} \quad G(x)=\int_a^x g$$

are continuous and differentiable on $[a,b]$, with $F'(x)=f(x)$ and $G'(x)=g(x)$, according to the fundamental theorem of calculus (2nd version). Then $H(x)=F(x)-G(x)$ is continuous.

But by hypothesis, we have

$$\begin{aligned} H(a)&=F(a)-G(a)=\int_a^a f-\int_a^a g=0-0=0 \\ H(b)&=F(b)-G(b)=\int_a^b f-\int_a^b g=0. \end{aligned}$$

Since H is also differentiable, $\exists c\in(a,b)$ such that $H'(c)=0$, by Rolle's theorem. As

$$H'(c)=F'(c)-G'(c)=f(c)-g(c)=0,$$

this completes the proof. ■

22. Let $f:[0,3]\to\mathbb{R}$ be defined by

$$f(x)=\begin{cases} x & x\in[0,1) \\ 1 & x\in[1,2) \\ x & x\in[2,3] \end{cases}.$$

Find $F:[0,3]\to\mathbb{R}$, where
$$F(x)=\int_0^x f.$$
Where is F differentiable? What is F' there?

Solution: the function f is increasing on $[0,3]$ so it is Riemann-integrable there. The function F is given by
$$F(x)=\begin{cases} \frac{x^2}{2}, & x\in[0,1) \\ x-\frac{1}{2}, & x\in[1,2) \\ \frac{x^2-1}{2}, & x\in[2,3] \end{cases}$$
By the fundamental theorem of calculus, F is differentiable wherever f is continuous, that is, on $[0,2)\cup(2,3]$, and $F'=f$ there. □

23. If $f:[0,1]\to\mathbb{R}$ is continuous and $\int_0^x f=\int_x^1 f$ for all $x\in[0,1]$, show that $f\equiv 0$.

Proof: as f is continuous, then $F(x)=\int_0^x f$ is continuous and differentiable on $[0,1]$, with $F'(x)=f(x)$, by the fundamental theorem of calculus. By the additivity theorem,
$$\int_0^x f+\int_x^1 f=\int_0^1 f.$$

But $\int_0^x f = \int_x^1 f$ so $2\int_0^x f = \int_0^1 f$. In particular,

$$F(x) = \frac{1}{2}\int_0^1 f = \text{constant}.$$

Then $f(x) = F'(x) = 0$ for all $x \in [0,1]$. ■

24. Let $f : [a,b] \to \mathbb{R}$ be continuous, $f \geq 0$ on $[a,b]$, and $\int_a^b f = 0$. Show that $f \equiv 0$ on $[a,b]$.

Proof: We show the contrapositive. Suppose that there exists $z \in [a,b]$ such that $f(z) > 0$. Since f is continuous, we may assume $z \in (a,b)$, as if $f(z) = 0$ for all $z \in (a,b)$, then $f(a) = f(b) = 0$.

Then, taking $\varepsilon = f(z)/2$ in the definition of continuity, there exists a $\delta > 0$ such that

$$|x - z| < \delta \Longrightarrow |f(x) - f(z)| < f(z)/2 \Longrightarrow f(x) > f(z)/2.$$

Reducing δ if necessary, we may assume $\delta \leq \min\{z-a, b-a\}$. Therefore,

$$[z - \delta/2, z + \delta/2] \subseteq (z - \delta, z + \delta) \subseteq [a,b].$$

Thus

$$\int_a^b f = \int_a^{z-\delta/2} f + \int_{z-\delta/2}^{z+\delta/2} f + \int_{z+\delta/2}^b f \geq 0 + \delta f(z)/2 + 0 > 0.$$

This completes the proof. ■

25. Let $f : [a,b] \to \mathbb{R}$ be continuous and let $\int_a^b f = 0$. Show $\exists c \in [a,b]$ such that $f(c) = 0$.

Proof: we show the contrapositive. Suppose $f(c) \neq 0$ for all $c \in [a,b]$. Then, by the intermediate value theorem, either $f(x) > 0$ for all $x \in [a,b]$ or $f(x) < 0$ for all $x \in [a,b]$.

If $f(x) > 0$ for all $x \in [a,b]$, then $\int_a^b f > 0$ by the preceding solved problem. Similarly, if $f(x) < 0$ for all $x \in [a,b]$, then $\int_a^b (-f) > 0$, which implies that $-\int_a^b f > 0$. In both cases, $\int_a^b f \neq 0$. ■

26. Compute $\dfrac{\mathrm{d}}{\mathrm{d}x}\displaystyle\int_{-x}^x e^{t^2}\,dt$.

Solution: according to the additivity property of the Riemann integral and the fundamental theorem of calculus, we have

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}x}\int_{-x}^x e^{t^2}\,dt &= \frac{\mathrm{d}}{\mathrm{d}x}\left(\int_{-x}^0 e^{t^2}\,dt + \int_0^x e^{t^2}\,dt\right)\\
&= \frac{\mathrm{d}}{\mathrm{d}x}\left(-\int_0^{-x} e^{t^2}\,dt + \int_0^x e^{t^2}\,dt\right)\\
&= -\frac{\mathrm{d}}{\mathrm{d}x}\int_0^{-x} e^{t^2}\,dt + \frac{\mathrm{d}}{\mathrm{d}x}\int_0^x e^{t^2}\,dt\\
&= -e^{x^2}\cdot(-1) + e^{x^2} = 2e^{x^2},
\end{aligned}$$

where we used the chain rule in the second-to-last equation. □

27. Let $f : [a, b] \to \mathbb{R}$ be Riemann-integrable on $[a + \delta, b]$ and unbounded in the interval $(a, a + \delta)$ for every $0 < \delta < b - a$. Define
$$\int_a^b f = \lim_{\delta \to 0^+} \int_{a+\delta}^b f,$$
where $\delta \to 0^+$ means that $\delta \to 0$ and $\delta > 0$. A similar construction allows us to define
$$\int_a^b g = \lim_{\delta \to 0^+} \int_a^{b-\delta} g.$$
Such integrals are said to be **improper**; when the limits exist, they are further said to be **convergent**. How can the expression
$$\int_0^1 \frac{1}{\sqrt{|x|}}\, \mathrm{d}x$$
be interpreted as an improper integral? Is it convergent? If so, what is its value?

Solution: by definition,
$$\int_0^1 \frac{1}{\sqrt{|x|}}\, \mathrm{d}x = \lim_{a \to 0^+} \int_a^1 \frac{1}{\sqrt{x}}\, \mathrm{d}x = \lim_{a \to 0^+} \left(2\sqrt{1} - 2\sqrt{a}\right) = 2.$$
Thus the improper integral converges to 2. □

28. Let $G : \mathbb{R} \to \mathbb{R}$ be defined according to
$$G(x) = \begin{cases} x^2 \sin\left(\frac{\pi}{x^2}\right) & x \neq 0 \\ 0 & x = 0 \end{cases}$$
Show that G is the antiderivative of some function $g : [0, 1] \to \mathbb{R}$, but that g is not Riemann-integrable on $[0, 1]$.

Proof: the derivative of G is
$$G'(x) = g(x) = \begin{cases} 2x \sin\left(\frac{\pi}{x^2}\right) - \frac{2\pi}{x} \cos\left(\frac{\pi}{x^2}\right), & x \neq 0 \\ 0, & x = 0 \end{cases}.$$
But g is not bounded on $[0, 1]$, so it cannot be Riemann-integrable on $[0, 1]$. ■

29. Let $f : \mathbb{R} \to \mathbb{R}$ be Thomae's function. Show that the indefinite integral of f on $[1, 2]$ is not an antiderivative of f on $[1, 2]$.

Proof: for any $x \in \mathbb{Q} \cap [1, 2]$, the indefinite integral F is such that $F'(x) \neq f(x)$; F cannot then be an antiderivative of f on $[1, 2]$.[10] ■

30. Without evaluating the integrals, show that $\displaystyle\int_1^4 e^{-8t}\, dt = \frac{1}{8} \int_4^8 te^{-t^2/2}\, dt.$

Proof: we can use the 2nd substitution theorem with $f(x) = e^{-x^2/2}$, $\varphi(t) = 4\sqrt{t}$, $\psi(t) = \frac{t^2}{16}$, $J = [1, 4]$. ■

[10] Of course, this will only make sense if you've managed to find F...

4.4 Exercises

1. Prepare a 2-page summary of this chapter, with important definitions and results.
2. With the assumptions of Theorem 48, show that f is decreasing on $[a, b]$ if and only if $f' \leq 0$ on (a, b).
3. Prove part 2. of the first derivative test.
4. Let $f : (a, b) \to \mathbb{R}$ be differentiable on (a, b), with $f'(x) \neq 0$. Prove the following statements.
   a) f is monotone on (a, b) and $f((a, b))$ is an open interval (α, β);
   b) f has an **inverse** $f^{-1} : (\alpha, \beta) \to \mathbb{R}$ such that
   $$f^{-1}(f(x)) = x, \quad f(f^{-1}(y)) = y, \quad \forall x \in (a, b), y \in (\alpha, \beta),$$
   c) f^{-1} is differentiable on (α, β), with
   $$(f^{-1})'(y) = \frac{1}{f'(f^{-1}(y))}, \quad \forall y \in (\alpha, \beta).$$
5. Let $I = [a, b]$ and $f : I \to \mathbb{R}$ be bounded. Then $U(Q; f) \leq L(Q; f)$ for any refinement $Q \supseteq P$ of I.
6. Prove that $f \equiv 1$ is Riemann-integrable on $[0, 1]$.
7. Show that Theorem 53 holds for decreasing functions.
8. Show that Thomae's function f is Riemann-integrable over $[0, 1]$ and that $\int_0^1 f = 0$.
9. Show that the signum function and Thomae's function do not have antiderivatives on any closed, bounded interval $I \subseteq \mathbb{R}$.
10. Show that the reciprocal of the square root function has an anti-derivative on $[0, 1]$, but that it is not Riemann-integrable on $[0, 1]$.
11. Find a function $f : [a, b] \to \mathbb{R}$ such that the indefinite integral $F : [a, b] \to \mathbb{R}$ defined by $F(x) = \int_a^x f$ is not an antiderivative of f.
12. Prove Theorems 63, 64, and 65.
13. For which values of s does the integral $\int_0^1 x^s \, \mathrm{d}x$ converge?
14. Show that the indefinite integral of sgn is not an antiderivative of sgn on $[-1, 1]$.

Chapter 5

Sequences of Functions

We now look at **sequences of functions**, which arise naturally in analysis and its applications. In particular, we discuss two **types of convergence** (**pointwise** and **uniform**and prove the **limit interchange theorems**.

5.1 Pointwise and Uniform Convergence

Let $A \subseteq \mathbb{R}$ and $(f_n)_n$ be a **sequence of functions** $f_n : A \to \mathbb{R}$. The sequence $(f_n(x))_n$ may converge for some $x \in A$ and diverge for others. Let $A_0 = \{x \in A \mid (f_n(x))_n \text{ converges}\} \subseteq A$. For each $x \in A_0$, $(f_n(x))$ converges to a unique limit

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

the **pointwise limit** of (f_n); we denote the situation by $f_n \to f$ on A_0.

Examples

1. Let $f_n : \mathbb{R} \to \mathbb{R}$ be defined by $f_n(x) = \frac{x}{n}$ for all $n \in \mathbb{N}, x \in \mathbb{R}$, and let f be the zero function on $\mathbb{R}$. Show that $f_n \to f$ on $\mathbb{R}$.

   Proof: let $\varepsilon > 0$ and $x \in \mathbb{R}$. By the Archimedean property, $\exists N_{\varepsilon,x} > \frac{|x|}{\varepsilon}$ so that

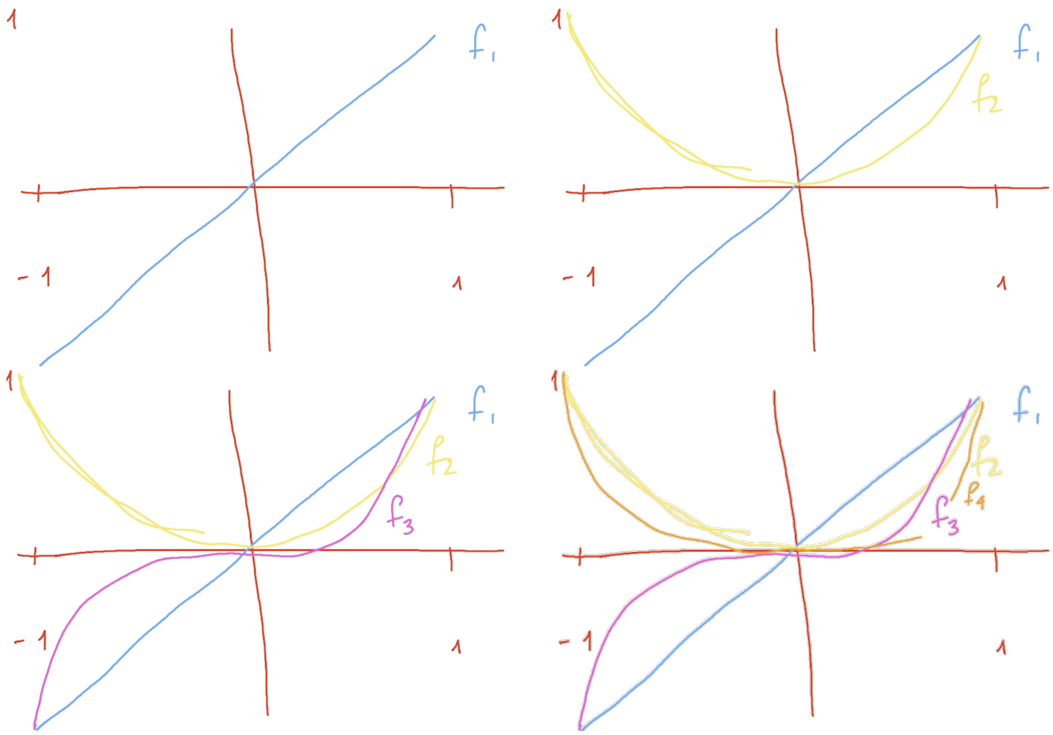
   $$n > N_{\varepsilon,x} \Longrightarrow \left|\tfrac{x}{n} - 0\right| < \tfrac{|x|}{n} < \tfrac{|x|}{N_{\varepsilon,x}} < \varepsilon,$$

   thus $f_n \to 0$ on $\mathbb{R}$. ■

2. Let $f_n : \mathbb{R} \to \mathbb{R}$ be defined by $f_n(x) = x^n$ for all $n \in \mathbb{N}, x \in \mathbb{R}$, and let f be the zero function on $\mathbb{R}$, except at $x = 1$ where $f(1) = 1$. Show that $f_n \to f$ on $(-1, 1]$.

   Proof: using various results seen in Chapters 2 and 3 (and in the solved problems and exercises), we know that

$$\lim_{n\to\infty} x^n = \begin{cases} 0 & x \in (-1,1) \\ 1 & x = 1 \\ \text{does not exist} & \text{otherwise} \end{cases}$$



Thus $f_n \to f$ on $(-1,1]$. Note that all f_n are continuous on $(1,1]$, but that f is not. ■

3. Let $f_n : \mathbb{R} \to \mathbb{R}$ be defined by $f_n(x) = \frac{x^2+nx}{n}$ for all $n \in \mathbb{N}, x \in \mathbb{R}$, and let f be the identity function on $\mathbb{R}$. Show that $f_n \to f$ on $\mathbb{R}$.

   Proof: as $f_n(x) = \frac{x^2}{n} + x \to f(x) = x, \forall x \in \mathbb{R}$, we have $f_n \to f$ on $\mathbb{R}$. ■

The last example show that there is something "incomplete" about pointwise convergence – why is continuity not preserved by the process? As it happens, we can define a different type of convergence which will preserve this important property.

A sequence of functions $f_n : A \to \mathbb{R}$ **converges uniformly on** $A_0 \subseteq A$ **to** $f : A_0 \to \mathbb{R}$, denoted by $f_n \rightrightarrows f$ on A_0, if the threshold $N_{\varepsilon,x} \in \mathbb{N}$ in the pointwise definition is in fact **independent** of $x \in A_0$:

$$\forall \varepsilon > 0,\ \exists N_\varepsilon \in \mathbb{N} \text{ such that } n > N_\varepsilon \text{ and } x \in A_0 \Longrightarrow |f_n(x) - f(x)| < \varepsilon.$$

The distinction between pointwise and uniform convergence is not unlike that between continuity and uniform continuity: convergence is uniform if the threshold is the same for all $x \in A_0$.

Clearly, if $f_n \rightrightarrows f$ on A_0, then $f_n \to f$ on A_0, but the converse is not necessarily true.

Examples

1. Show that the sequence $f_n : [1,2] \to \mathbb{R}$ defined by $f_n(x) = \frac{\sin x}{nx}$ for $n \in \mathbb{N}$ converges uniformly to the zero function on $[1,2]$.

   Proof: let $\varepsilon > 0$. According to the Archimedean property, $\exists N_\varepsilon > \frac{1}{\varepsilon}$ so that

   $$n > N_\varepsilon \text{ and } x \in [1,2] \Longrightarrow \left|\frac{\sin x}{nx} - 0\right| = \left|\frac{\sin x}{nx}\right| \leq \frac{1}{nx} \leq \frac{1}{n} < \frac{1}{N_\varepsilon} < \varepsilon,$$

   thus $f_n \rightrightarrows 0$ on $[1,2]$. ■

2. Let $f_n : \mathbb{R} \to \mathbb{R}$ be defined by $f_n(x) = x^n$ for all $n \in \mathbb{N}, x \in \mathbb{R}$, and let f be the zero function on $\mathbb{R}$, except at $x = 1$ where $f(1) = 1$. Show that $f_n \not\rightrightarrows f$ on $(-1,1]$.

   Proof: we use the negation of the definition. Let $\varepsilon_0 = \frac{1}{4}$, and set $x_k = \frac{1}{2^{1/k}}$ and $(n_k) = (k)$. Then

   $$|f_{n_k}(x_k) - f(x_k)| = \left|\tfrac{1}{2} - 0\right| = \frac{1}{2} \geq \varepsilon_0,$$

   which completes the proof. ■

A sequence of functions f_n does not converge uniformly to f on A_0 if

$$\exists \varepsilon_0 > 0 \text{ with } (f_{n_k}) \subseteq (f_n) \text{ and } (x_k) \subseteq A_0 \text{ s.t. } |f_{n_k}(x_k) - f(x_k)| \geq \varepsilon_0,\ \forall k \in \mathbb{N}.$$

Example: let $f_n : [0,1] \to \mathbb{R}$ be the sequence of functions defined by

$$f_n(x) = \begin{cases} nx, & x \in [0,1/n] \\ 2 - nx, & x \in [1/n, 2/n] \\ 0 & x \in [2/n, 1] \end{cases}$$

for all $n \in \mathbb{N}$. Let $f : [0,1] \to \mathbb{R}$ be the zero function on $[0,1]$. Show that $f_n \to f$ on $[0,1]$ but $f_n \not\rightrightarrows f$ on $[0,1]$.

Proof: if $x = 0$, $f_n(0) = 0$ for all n so $(f_n(0))$ converges to 0. If $x \in (0, 1]$, $\exists N_x > 2/x$ by the Archimedean property. Thus, for $n > N_x$, $f_n(x) = 0$ since $x > \frac{2}{N} > \frac{2}{n}$, so $f_n(x) \to 0$ on $(0, 1]$. Combining these results, $f_n \to f$ on $[0, 1]$.

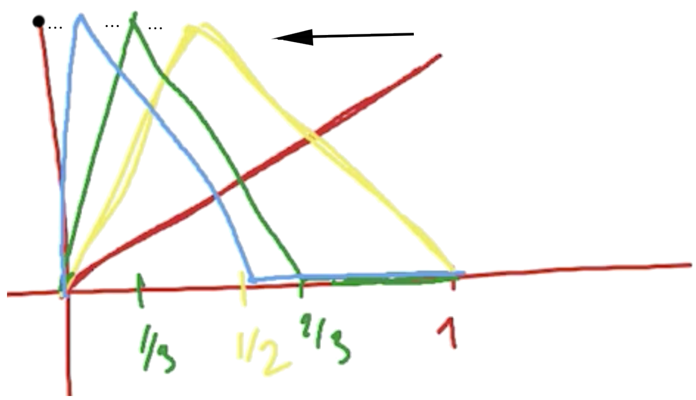
Now, let $\varepsilon_0 = \frac{1}{2}$. Note that since $|f_n(\frac{1}{n}) - f(\frac{1}{n})| = 1$ for all $n \in \mathbb{N}$, we can never obtain

$$|f_n(x) - f(x)| < \varepsilon$$

for all $x \in [0, 1]$, and so $f_n \not\rightrightarrows f$ on $[0, 1]$. ■

The fact that we have to separate the proof for pointwise convergence into distinct arguments depending on the value of x is a strong indication that the convergence cannot be uniform.[1]

Intuitively, we can think of the convergence process in the last example as being a flattening process: what happens to the tents' peak as $n \to \infty$? That we have to "break" the tents to get to the pointwise limit is another indication that the convergence cannot be uniform.



The definition of uniform convergence is only ever useful if a candidate for a uniform limit is available, a situation that we have encountered before. As was the case for number sequences, the completeness of $\mathbb{R}$ comes to the rescue.

Theorem 66 (Cauchy's Criterion for Sequences of Functions)
Let $f_n : A \to \mathbb{R}$, $\forall n \in \mathbb{N}$. Then, $f_n \rightrightarrows f$ on $A_0 \subseteq A$ if and only if $\forall \varepsilon > 0$, $\exists N_\varepsilon \in \mathbb{N}$ (indep. of $x \in A_0$) such that $|f_m(x) - f_n(x)| < \varepsilon$ whenever $m \geq n > N_\varepsilon \in \mathbb{N}$ and $x \in A_0$.

Proof: let $\varepsilon > 0$. If $f_n \rightrightarrows f$ on A_0, $\exists N_\varepsilon \in \mathbb{N}$ such that $|f_n(x) - f(x)| < \varepsilon$ when $x \in A_0$ and $n > N_\varepsilon$. Hence,

$$\begin{aligned}|f_m(x) - f_n(x)| &= |f_m(x) - f(x) + f(x) - f_n(x)| \\ &\leq |f_m(x) - f(x)| + |f_n(x) - f(x)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon\end{aligned}$$

whenever $x \in A_0$ and $m \geq n > N_\varepsilon$.

[1]Although it could be that it was possible to do a one-pass proof and that the insight escaped us.

Conversely, let $\varepsilon > 0$ and assume that $\exists N_{\varepsilon/2} \in \mathbb{N}$ (independent of $x \in A_0$) such that

$$m \geq n > N_{\varepsilon/2} \text{ and } x \in A_0 \Longrightarrow -\frac{\varepsilon}{2} < f_m(x) - f_n(x) < \frac{\varepsilon}{2}.$$

Since $x \in A_0$, we know that $f_m(x) \to f$ on A_0 when $m \to \infty$. Thus,

$$m \geq n > N_{\varepsilon/2} \text{ and } x \in A_0 \Longrightarrow \lim_{m\to\infty} -\frac{\varepsilon}{2} \leq \lim_{m\to\infty} (f_m(x) - f_n(x)) \leq \lim_{m\to\infty} \frac{\varepsilon}{2},$$

or

$$m \geq n > N_{\varepsilon/2} \text{ and } x \in A_0 \Longrightarrow -\varepsilon < -\frac{\varepsilon}{2} \leq f(x) - f_n(x) \leq \frac{\varepsilon}{2} < \varepsilon,$$

and so $f_n \rightrightarrows f$ on A_0. ■

5.2 Limit Interchange Theorems

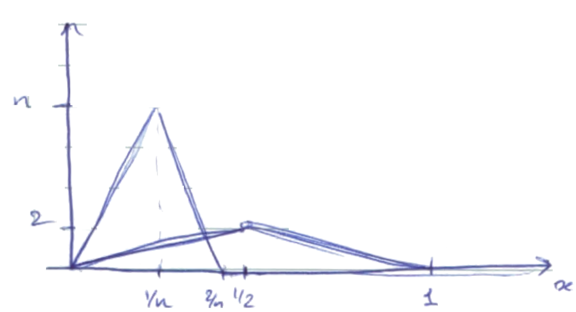
It is often necessary to know if the limit f of a sequence of functions (f_n) is continuous, differentiable, or Riemann-integrable. Unfortunately, we cannot guarantee that this will be the case, even when the f_n are continuous, differentiable, or Riemann-integrable, respectively.

Examples

1. Consider the sequence of functions $f_n : [0,1] \to \mathbb{R}$ defined by $f_n(x) = x^n$ for $n \in \mathbb{N}$ and $f : [0,1] \to \mathbb{R}$ be the zero function except at $x = 1$ where $f(1) = 1$. Then f_n is continuous on $[0,1]$ for all $n \in \mathbb{N}$, but f is not. □
2. The same functions f_n are differentiable on $[0,1]$ for all $n \in \mathbb{N}$, but f is not (as it is not continuous at $x = 1$). □
3. Consider the functions $f_n : [0,1] \to \mathbb{R}$ defined by

$$f_n(x) = \begin{cases} n^2 x, & x \in [0, 1/n] \\ -n^2(x - 2/n), & x \in [1/n, 2/n] \\ 0 & x \in [2/n, 1] \end{cases}$$

for $n \geq 2$.



Since f_n is continuous on $[0,1]$ for all $n \geq 2$, f_n is Riemann-integrable on $[0,1]$ for all $n \geq 2$, with

$$\int_0^1 f_n = \frac{1}{2} \cdot \frac{2}{n} \cdot n = 1, \quad \text{for all } n \geq 2.$$

If $x = 0$, $f_n(0) = 0$ for all n so $(f_n(0))$ converges to 0.

If $x \in (0,1]$, $\exists N_x > 2/x$ by the Archimedean property. Thus, for $n > N_x$, $f_n(x) = 0$ since $x > \frac{2}{N} > \frac{2}{n}$, so $f_n(x) \to 0$ on $(0,1]$. So $f_n \to f$ on $[0,1]$, but

$$\int_0^1 f = 0 \neq 1 = \lim_{n\to\infty} \int_0^1 f,$$

which is to say we cannot interchange the limit and the integral here. $\square$

Note that none of the "convergences" in the previous example are uniform on $[0,1]$. When the convergence $f_n \rightrightarrows f$ on A is **uniform**, then if the f_n are

- continuous on A, so is f;
- differentiable on A, so is f, with

$$f' = \frac{\mathrm{d}}{\mathrm{d}x}\Big[\lim_{n\to\infty} f_n\Big] = \lim_{n\to\infty}\Big[\frac{\mathrm{d}}{\mathrm{d}x} f_n\Big] = \lim_{n\to\infty} f_n';$$

- Riemann-integrable on A, then so is f, with

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

We finish this chapter by proving three **limit interchange theorems**, with applications in analysis, engineering, and mathematical physics.[2]

Theorem 67
Let $f_n : A \to \mathbb{R}$ be continuous on A for all $n \in \mathbb{N}$. If $f_n \rightrightarrows f$ on A, then f is continuous on A.

Proof: let $\varepsilon > 0$. By definition, $\exists H_{\varepsilon/3} \in \mathbb{N}$ such that

$$n > H_{\varepsilon/3} \text{ and } x \in A \Longrightarrow |f_n(x) - f(x)| < \tfrac{\varepsilon}{3}.$$

[2]Although their conclusions are often used without verifying that the convergence is indeed uniform.

Let $c \in A$. According to the triangle inequality,

$$\begin{aligned}|f(x)-f(c)| &\leq |f(x)-f_{H_{\varepsilon/3}}(x)| + |f_{H_{\varepsilon/3}}(x)-f_{H_{\varepsilon/3}}(c)| + |f_{H_{\varepsilon/3}}(c)-f(c)| \\ &< \tfrac{\varepsilon}{3} + |f_{H_{\varepsilon/3}}(x)-f_{H_{\varepsilon/3}}(c)| + \tfrac{\varepsilon}{3}\end{aligned}$$

whenever $n > H_{\varepsilon/3}$.

But $f_{H_{\varepsilon/3}}$ is continuous at c, so $\exists \delta_{\varepsilon/3} > 0$ such that $|f_{H_{\varepsilon/3}}(x) - f_{H_{\varepsilon/3}}(c)| < \frac{\varepsilon}{3}$ when $x \in A$ and $|x-c| < \delta_{\varepsilon/3}$. Thus $|f(x)-f(c)| < \varepsilon$ whenever $x \in A$ and $|x-c| < \delta_{\varepsilon/3}$, so f is continuous at c. As $c \in A$ is arbitrary, f is continuous on A. ■

The next two results are slightly more complicated to prove.

Theorem 68
Let $f_n : [a,b] \to \mathbb{R}$ be a sequence of differentiable functions on $[a,b]$ such that $\exists x_0 \in [a,b]$ with $f_n(x_0) \to z_0$, and $f_n'' \rightrightarrows g$ on $[a,b]$. Then $f_n \rightrightarrows f$ on $[a,b]$ for some function $f : [a,b] \to \mathbb{R}$ such that $f' = g$.

Proof: let $\varepsilon > 0$ and $x \in [a,b]$. Since $f_n' \rightrightarrows g$ on $[a,b]$, the sequence f_n' satisfies Cauchy's criterion, and so $\exists N_1 \in \mathbb{N}$ such that

$$m \geq n > N_1 \text{ and } y \in [a,b] \Longrightarrow |f_m'(y) - f_n'(y)| < \frac{\varepsilon}{2(b-a)}.$$

As $(f_n(x_0))$ converges it is also a Cauchy sequence, so $\exists N_2 \in \mathbb{N}$ such that

$$m \geq n > N_2 \Longrightarrow |f_m(x_0) - f_n(x_0)| < \frac{\varepsilon}{2}.$$

According to the mean value theorem, $\exists y$ between x and x_0 such that

$$(f_m(x) - f_n(x)) - (f_m(x_0) - f_n(x_0)) = (f_m'(y) - f_n'(y))(x - x_0).$$

Hence,

$$\begin{aligned}|f_m(x) - f_n(x)| &\leq |f_m(x_0) - f_n(x_0)| + |f_m'(y) - f_n'(y)| \cdot |x - x_0| \\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2(b-a)}(b-a) = \varepsilon\end{aligned}$$

for all $m \geq n > \max\{N_1, N_2\}$.

Both N_1 and N_2 are independent of x, so $N_\varepsilon = \max\{N_1, N_2\}$ also is, and thus $(f_n)_n$ satisfies Cauchy's criterion, which yields $f_n \rightrightarrows f$ on $[a,b]$.

It remains only to show that $f' = g$ on $[a,b]$. Let $\varepsilon > 0$ and $c \in [a,b]$. Since (f_n') satisfies Cauchy's criterion (as $f_n' \rightrightarrows g$), $\exists K_1 \in \mathbb{N}$ (independent of x) such that

$$m \geq n > K_1 \text{ and } y \in [a,b] \Longrightarrow |f_m'(y) - f_n'(y)| < \frac{\varepsilon}{3}.$$

But $f' \rightrightarrows g'$, so $\exists K_2 \in \mathbb{N}$ (independent of c) such that

$$n \geq K_2 \text{ and } c \in [a,b] \Longrightarrow |f_n'(c) - g(c)| < \frac{\varepsilon}{3}.$$

Set $K_\varepsilon > \max\{K_1, K_2\}$.

As $f_{K_\varepsilon}'(c)$ exists, $\exists \delta_\varepsilon > 0$ such that

$$0 < |x-c| < \delta_\varepsilon \text{ and } x \in [a,b] \Longrightarrow \left| \frac{f_{K_\varepsilon}(x) - f_{K_\varepsilon}(c)}{x-c} - f_{K_\varepsilon}'(c) \right| < \frac{\varepsilon}{3}.$$

According to the mean value theorem, $\exists y$ between x and c such that

$$(f_m(x) - f_n(x)) - (f_m(c) - f_n(c)) = (f_m'(y) - f_n'(y))(x-c).$$

If $x \neq c$, then $m \geq n > K_\varepsilon$ and $x \in [a,b] \Longrightarrow$

$$\left| \frac{f_m(x) - f_m(c)}{x-c} - \frac{f_n(x) - f_n(c)}{x-c} \right| = |f_m'(y) - f_n'(y)| < \frac{\varepsilon}{3}.$$

Letting $m \to \infty$ (i.e. $f_m \to f$ on A), we get

$$n > K_\varepsilon \text{ and } x \in [a,b] \Longrightarrow \left| \frac{f(x) - f(c)}{x-c} - \frac{f_m(c) - f_n(c)}{x-c} \right| \leq \frac{\varepsilon}{3}.$$

Combining all of these inequalities, for $0 < |x-c| < \delta_\varepsilon$, $x \in [a,b]$, and $k > K_\varepsilon$, we have

$$\begin{aligned} \left| \frac{f(x) - f(c)}{x-c} - g(c) \right| &= \left| \frac{f(x) - f(c)}{x-c} - \frac{f_k(x) - f_k(c)}{x-c} \right. \\ &\qquad \left. + \frac{f_k(x) - f_k(c)}{x-c} - f_k'(c) + f_k'(c) - g(c) \right| \\ &\leq \left| \frac{f(x) - f(c)}{x-c} - \frac{f_k(x) - f_k(c)}{x-c} \right| + \left| \frac{f_k(x) - f_k(c)}{x-c} - f_k'(c) \right| \\ &\qquad + |f_k'(c) - g(c)| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon, \end{aligned}$$

which is to say that $f'(c) = g(c)$. ■

Theorem 69
Let $f_n : [a, b] \to \mathbb{R}$ be Riemann-integrable on $[a, b]$ for all $n \in \mathbb{N}$. If $f_n \rightrightarrows f$ on $[a, b]$, then f is Riemann-integrable on $[a, b]$ and

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

Proof: let $\varepsilon > 0$. Since $f_n \rightrightarrows f$ on $[a, b]$, $\exists K_\varepsilon \in \mathbb{N}$ (independent of x) such that

$$n \geq K_\varepsilon \Longrightarrow |f_n(x) - f(x)| < \frac{\varepsilon}{4(b-a)}.$$

Since f_{K_ε} is Riemann-integrable, $\exists P_\varepsilon = \{x_0, \ldots, x_n\}$ a partition of $[a, b]$ such that

$$U(P_\varepsilon; f_{K_\varepsilon}) - L(P_\varepsilon; f_{K_\varepsilon}) < \frac{\varepsilon}{2},$$

according to the Riemann criterion.

For all $1 \leq i \leq n$, set

$$\begin{aligned} m_i(f) &= \inf\{f(x) \mid x \in [x_{i-1}, x_i]\},\ m_i(f_{K_\varepsilon}) = \inf\{f_{K_\varepsilon}(x) \mid x \in [x_{i-1}, x_i]\}, \\ M_i(f) &= \sup\{f(x) \mid x \in [x_{i-1}, x_i]\},\ M_i(f_{K_\varepsilon}) = \sup\{f_{K_\varepsilon}(x) \mid x \in [x_{i-1}, x_i]\}. \end{aligned}$$

Then according to the reverse triangle inequality, we have

$$\begin{aligned} |f(x)| < |f_{K_\varepsilon}(x)| + \frac{\varepsilon}{4(b-a)} &\Longrightarrow |f(x)| < M_i(f_{K_\varepsilon}) + \frac{\varepsilon}{4(b-a)} \text{ on } [x_{i-1}, x_i] \\ &\Longrightarrow M_i(f) < M_i(f_{K_\varepsilon}) + \frac{\varepsilon}{4(b-a)} \text{ on } [x_{i-1}, x_i]. \end{aligned}$$

Similarly, $m_i(f) \geq m_i(f_{K_\varepsilon}) - \frac{\varepsilon}{4(b-a)}$ on $[x_{i-1}, x_i]$. Thus,

$$\begin{aligned} U(P_\varepsilon; f) &= \sum_{i=1}^n M_i(f)(x_i - x_{i-1}) \\ &\leq \sum_{i=1}^n M_i(f_{K_\varepsilon})(x_i - x_{i-1}) + \frac{\varepsilon}{4(b-a)} \sum_{i=1}^n (x_i - x_{i-1}) = U(P_\varepsilon; f_{K_\varepsilon}) + \frac{\varepsilon}{4}. \end{aligned}$$

Similarly, $L(P_\varepsilon; f) \geq L(P_\varepsilon; f_{K_\varepsilon}) - \frac{\varepsilon}{4}$. Hence

$$U(P_\varepsilon; f) - L(P_\varepsilon; f) \leq U(P_\varepsilon; f_{K_\varepsilon}) - L(P_\varepsilon; f_{K_\varepsilon}) + \frac{\varepsilon}{2} < \varepsilon.$$

Thus, according to the Riemann criterion, f is Riemann-integrable.

Finally, let $\varepsilon > 0$. As $f_n \rightrightarrows f$ on $[a,b]$, $\exists \hat{K}_\varepsilon$ (independent of x) such that

$$n > \hat{K}_\varepsilon \text{ and } x \in [a,b] \implies |f_n(x) - f(x)| < \frac{\varepsilon}{2(b-a)}.$$

Consequently, $\int_a^b f_n \to \int_a^b f$, since $n > \hat{K}_\varepsilon \implies$

$$\left|\int_a^b f_n - \int_a^b f\right| = \left|\int_a^b (f_n - f)\right| \leq \int_a^b |f_n - f| \leq \int_a^b \frac{\varepsilon}{2(b-a)} = \frac{\varepsilon}{2} < \varepsilon,$$

which completes the proof. ■

5.3 Solved Problems

1. Show that $\lim\limits_{n\to\infty} \dfrac{nx}{1+n^2x^2} = 0$ for all $x \in \mathbb{R}$.

   Proof: if $x = 0$, then $\frac{nx}{1+n^2x^2} = 0 \to 0$. If $x \neq 0$, let $\varepsilon > 0$. By the Archimedean property, $\exists N_\varepsilon > \frac{1}{\varepsilon|x|}$ (depending on x) s.t.

   $$\left|\frac{nx}{1+n^2x^2} - 0\right| = \frac{n|x|}{1+n^2x^2} < \frac{n|x|}{n^2x^2} = \frac{1}{n|x|} < \frac{1}{N_\varepsilon|x|} < \varepsilon$$

   whenever $n > N_\varepsilon$, i.e. $\frac{nx}{1+n^2x^2} \to 0$ on $\mathbb{R}$. ■

2. Show that if $f_n(x) = x + \frac{1}{n}$ and $f(x) = x$ for all $x \in \mathbb{R}$, $n \in \mathbb{N}$, then $f_n \rightrightarrows f$ on $\mathbb{R}$ but $f_n^2 \not\rightrightarrows g$ on $\mathbb{R}$ for any function g.

   Proof: let $\varepsilon > 0$. By the Archimedean property, $\exists N_\varepsilon > \frac{1}{\varepsilon}$ (independent of x) s.t.

   $$|f_n(x) - f(x)| = \left|x + \frac{1}{n} - x\right| = \frac{1}{n} < \frac{1}{N_\varepsilon} < \varepsilon$$

   whenever $n > N_\varepsilon$, i.e. $f_n \rightrightarrows 0$ on $\mathbb{R}$.

   Now, $(f_n(x))^2 = x^2 + \frac{2x}{n} + \frac{1}{n^2} \to x^2$ for all $x \in \mathbb{R}$. Hence, $f_n^2 \to g$ on $\mathbb{R}$, where $g(x) = x^2$. If f_n^2 converges uniformly to any function, it will have to do so to g. But let $\varepsilon_0 = 2$ and $x_n = n$. Then

   $$\left|(f_n(x_n))^2 - g(x_n)\right| = \left|\frac{2x_n}{n} + \frac{1}{n^2}\right| = 2 + \frac{1}{n^2} \geq 2 = \varepsilon_0$$

   for all $n \in \mathbb{N}$ (this is the negation of the definition of uniform convergence). Hence f_n^2 does not converge uniformly on $\mathbb{R}$. ■

3. Let $f_n(x) = \frac{1}{(1+x)^n}$ for $x \in [0,1]$. Denote by f the pointwise limit of f_n on $[0,1]$. Does $f_n \rightrightarrows f$ on $[0,1]$?

Proof: first note that $1 \leq 1+x$ on $[0,1]$. In particular, $\frac{1}{1+x} \leq 1$ on $[0,1]$. If $x \in (0,1]$, then $\frac{1}{(1+x)^n} \to 0$, according to one of the chapter's examples.

If $x = 0$,

$$\frac{1}{(1+x)^n} = \frac{1}{1^n} = 1 \to 1;$$

i.e. $f_n \to f$ on $[0,1]$, where

$$f(x) = \begin{cases} 0, & x \in (0,1] \\ 1, & x = 0 \end{cases}.$$

However, $f_n \not\rightrightarrows f$ by theorem 67, since f_n is continuous on $[0,1]$ for all $n \in \mathbb{N}$, but f is not. ■

4. Let (f_n) be the sequence of functions defined by $f_n(x) = \frac{x^n}{n}$, for $x \in [0,1]$ and $n \in \mathbb{N}$. Show that (f_n) converges uniformly to a differentiable function $f : [0,1] \to \mathbb{R}$, and that the sequence (f_n') converges pointwise to a function $g : [0,1] \to \mathbb{R}$, but that $g(1) \neq f'(1)$.

   Proof: the sequence $f_n(x) = \frac{x^n}{n} \to f(x) \equiv 0$ on $[0,1]$. Indeed, let $\varepsilon > 0$. By the Archimedean property, $\exists N_\varepsilon > \frac{1}{\varepsilon}$ s.t.

$$\left| \frac{x^n}{n} - 0 \right| \leq \frac{|x|^n}{n} \leq \frac{1}{n} < \frac{1}{N_\varepsilon} < \varepsilon$$

   whenever $n > N_\varepsilon$. Note that f is differentiable and $f'(x) = 0$ for all $x \in [0,1]$. Furthermore, $f_n'(x) = \frac{nx^{n-1}}{n} = x^{n-1} \to g(x)$ on $[0,1]$, where

$$g(x) = \begin{cases} 0, & x \in [0,1) \\ 1, & x = 1 \end{cases},$$

   by one of the examples I did in class. Then $g(1) = 1 \neq 0 = f'(1)$. ■

5. Show that $\lim\limits_{n\to\infty} \int_1^2 e^{-nx^2}\, \mathrm{d}x = 0$.

   Proof: as $\left(e^{-nx^2}\right)' = -2nxe^{-nx^2} < 0$ on $[1,2]$ for all $n \in \mathbb{N}$, e^{-nx^2} is decreasing on $[1,2]$ for all n, that is

$$e^{-nx^2} \leq e^{-n(1)^2} = e^{-n} \quad \text{for all } n \in \mathbb{N}.$$

   Now,

$$f_n(x) = e^{-nx^2} \rightrightarrows f(x) \equiv 0 \quad \text{on } [1,2].$$

   Indeed, let $\varepsilon > 0$. By the Archimedean property, $\exists N_\varepsilon > \ln \frac{1}{\varepsilon}$ (independent of x) s.t.

$$\left| e^{-nx^2} - 0 \right| = e^{-nx^2} < e^{-Nx^2} \leq e^{-N} < \varepsilon$$

   whenever $n > N_\varepsilon$. Then (and only because of this uniform convergence),

$$\lim_{n\to\infty} \int_1^2 e^{-nx^2}\, \mathrm{d}x = \int_1^2 \lim_{n\to\infty} e^{-nx^2}\, \mathrm{d}x = \int_1^2 0\, \mathrm{d}x = 0,$$

   by the limit interchange theorem for integrals. ■

6. Show that $\lim_{n\to\infty} \int_{\pi/2}^{\pi} \frac{\sin(nx)}{nx}\,\mathrm{d}x = 0$.

   Proof: for $n \in \mathbb{N}$, define $f_n : [\pi/2, \pi] \to \mathbb{R}$ by

   $$f_n(x) = \frac{\sin(nx)}{nx}.$$

   Then each f_n is continuous. For all $n \in \mathbb{N}$, we have

   $$\sup_{x\in[\pi/2,\pi]} \left\{ \left| \frac{\sin(nx)}{nx} \right| \right\} \le \frac{2}{n\pi}.$$

   Since $2/n\pi \to 0$ as $n \to \infty$, we have $f_n \rightrightarrows 0$ (why?). Then the limit interchange theorem for integrals applies, and we have

   $$\lim_{n\to\infty} \int_{\pi/2}^{\pi} \frac{\sin(nx)}{nx}\,\mathrm{d}x = \int_{\pi/2}^{\pi} 0\,\mathrm{d}x = 0.$$

   This completes the proof. ■

7. Show that if $f_n \rightrightarrows f$ on $[a, b]$, and each f_n is continuous, then the sequence of functions $(F_n)_n$ defined by

   $$F_n(x) = \int_a^x f_n(t)\,dt$$

   also converges uniformly on $[a, b]$.

   Proof: define $F(x) = \int_a^x f(t)\,dt$. Let $\varepsilon > 0$. Since $f_n \rightrightarrows f$, $\exists N \in \mathbb{N}$ such that, for all $n \ge N$, we have

   $$|f_n(x) - f(x)| < \frac{\varepsilon}{b-a} \quad \forall\, x \in [a, b].$$

   Then, for all $n \ge N$ and $x \in [a, b]$, we have

   $$\begin{aligned} |F_n(x) - F(x)| &= \left| \int_a^x f_n(t)\,dt - \int_a^x f(t)\,dt \right| \le \int_a^x |f_n(t) - f(t)|\,dt \\ &\le (x-a) \cdot \frac{\varepsilon}{b-a} \le \varepsilon. \end{aligned}$$

   Thus $F_n \rightrightarrows F$ on $[a, b]$. ■

5.4 Exercises

1. Prepare a 2-page summary of this chapter, with important definitions and results.

2. Are all the hypotheses of Theorem 68 necessary?

Chapter 6

Series of Functions

In the final chapter of this part, we discuss a specific type of sequence: the **series** (series of **numbers**, series of **functions**, and **power** series). Note that the latter is more naturally expressed using a complex analysis framework (see Chapter 22), but we present it here, as well as important theorems for regular series, in the real analysis framework.

6.1 Series of Numbers

Let $(x_n) \subseteq \mathbb{R}$. The **series** associated with (x_n), denoted by

$$S_{(x_n)} = \sum_{n=1}^{\infty} x_n,$$

is the sequence (s_n), where

$$s_1 = x_1, \quad s_2 = x_1 + x_2, \quad s_3 = x_1 + x_2 + x_3, \quad \ldots$$

If the sequence of partial sums s_n converges to S, we say the series $S_{(x_n)}$ **converges to the sum** S. When the context is clear, we may also write $\sum x_n (= S)$.

We start by producing a necessary condition for convergence.

Theorem 70
If $\sum_{n=1}^{\infty} x_n$ *converges, then* $x_n \to 0$*.*

Proof: let S be the limit of the partial sums. Then

$$\lim_{n\to\infty}(s_n - s_{n-1}) = \lim_{n\to\infty} s_n - \lim_{n\to\infty} s_{n-1} = S - S = 0,$$

with the second equation being guaranteed by Theorem 14 and the convergence of the series. ■

We can bypass the need to know the limit in order to prove convergence.

Theorem 71 (Cauchy Criterion for Series)
The series $\sum_{n=1}^{\infty} x_n$ *converges if and only if* $\forall \varepsilon > 0$, $\exists N_\varepsilon \in \mathbb{N}$ *such that*

$$m > n > N_\varepsilon \implies |x_{n+1} + \cdots + x_m| < \varepsilon.$$

Proof: let (s_n) be the series of partial sums. If (s_n) converges, it is a Cauchy sequence, so that $\exists N_\varepsilon \in \mathbb{N}$ such that $m > n > N_\varepsilon \implies |s_m - s_n| < \varepsilon$. But $|s_m - s_n| = |x_m + \cdots + x_{n+1}|$, so Cauchy's criterion holds.

Conversely, if Cauchy's criterion holds, the sequence of partial terms is a Cauchy sequence, and so the series converges by completeness of $\mathbb{R}$. ■

Other tests can be used to show the convergence of a series without knowing the limit.

Theorem 72 (Comparison Test)
Let $\sum_{n=1}^{\infty} x_n$, $\sum_{n=1}^{\infty} y_n$ *be series whose terms are all non-negative. If* $\exists K \in \mathbb{N}$ *such that* $0 \leq x_n \leq y_n$ *when* $n > K$, *then*

1. $\sum_{n=1}^{\infty} y_n$ *converges* $\implies \sum_{n=1}^{\infty} x_n$ *converges.*
2. $\sum_{n=1}^{\infty} x_n$ *diverges* $\implies \sum_{n=1}^{\infty} y_n$ *diverges.*

Proof: we prove 1.; the proof for the other part is simply the contrapositive. Let $\varepsilon > 0$. As $\sum y_n$ converges, $\exists N_\varepsilon \in \mathbb{N}$ such that $0 \leq y_{n+1} + \cdots + y_m < \varepsilon$ according to Cauchy's criterion for series.

Hence, whenever $m \geq n > M_\varepsilon = \max\{N_\varepsilon, K\}$, then

$$0 \leq \sum_{i=n+1}^{m} x_i \leq \sum_{i=n+1}^{m} y_i < \varepsilon.$$

As such, $\sum x_n$ converges as it satisfies Cauchy's criterion for series. ■

Typical problems may look like the following.

Example: discuss the convergence of $\sum_{n=1}^{\infty} \frac{1}{n(n+1)}$ and $\sum_{n=1}^{\infty} \frac{1}{n^2}$.

Solution: the limit of the partial sums of the first series converges to 1 as

$$\lim_{k\to\infty} s_k = \lim_{k\to\infty} \sum_{n=1}^{k} \left(\frac{1}{n} - \frac{1}{n+1}\right) = \lim_{k\to\infty} \left(1 - \frac{1}{k+1}\right) = 1 - 0 = 1.$$

For the second series, since $n^2 \geq \frac{1}{2}(n^2+n) \geq 0$ for all $n \in \mathbb{N}$, then $\frac{2}{n(n+1)} \geq \frac{1}{n^2} \geq 0$ for all $n \in \mathbb{N}$, and

$$\infty > 2\sum_{n=1}^{\infty} \frac{1}{n(n+1)} \geq \sum_{n=1}^{\infty} \frac{1}{n^2},$$

thus the series converges, according to the comparison theorem. □

When the sign of the underlying sequence terms alternates, convergence is particularly easy to establish.

Theorem 73 (Alternating Series Test)
Let (a_n) be a sequence of non-negative numbers such that $a_n \searrow 0$ (i.e., $a_n \to 0$ and $a_{n+1} \leq a_n$). Then $\sum_{n=0}^{\infty} (-1)^n a_n$ converges.

Proof: let (s_k) be the series of partial sums

$$s_k = \sum_{n=0}^{k} (-1)^n a_n.$$

The subsequence of even terms is $s_{2k} = s_{2k-2} - (a_{2k-1} - a_{2k})$; that of the odd terms is $s_{2k+1} = s_{2k-1} - (a_{2k} - a_{2k+1})$. Since $a_n \searrow 0$, $a_{n+1} \leq a_n$ for all n. Thus $s_{2k} \leq s_{2k-2}$ and $s_{2k+1} \geq s_{2k-1}$ for all $k \in \mathbb{N}$. But $s_{2k} \geq s_{2m+1}$ for all $k, m \in \mathbb{N}$ (left as an exercise), and so

$$a_0 = s_0 \geq s_2 \geq s_4 \geq \cdots \geq s_5 \geq s_3 \geq s_1 = a_0 - a_1.$$

Thus (s_{2k}) is a bounded decreasing sequence and (s_{2k-1}) is a bounded increasing sequence, and so $\lim_{k\to\infty} s_{2k}$ and $\lim_{k\to\infty} s_{2k-1}$ exist. According to Theorem 14, then, we have

$$\lim_{k\to\infty} (s_{2k} - s_{2k-1}) = \lim_{k\to\infty} a_{2k} = 0$$

since $a_n \searrow 0$, which implies that the alternating series converges:

$$\lim_{k\to\infty} \sum_{n=0}^{2k} (-1)^n a_n = \lim_{k\to\infty} s_{2k} = \lim_{k\to\infty} s_{2k+1} = \lim_{k\to\infty} \sum_{n=0}^{2k+1} (-1)^n a_n,$$

which completes the proof. ■

Even though it was not part of the statement, the proof of Theorem 73 allows us to conclude that the value of a convergent alternating series lies between a_{2k} and a_{2m+1} for all $k, m \in \mathbb{N}$.

Example: the alternating harmonic series $-1 + 1/2 - 1/3 + \cdots$ converges. Indeed, consider the sequence $(a_n) = (\frac{1}{n}) = (1, \frac{1}{2}, \frac{1}{3}, \ldots)$. As $\frac{1}{n} \to 0$ and $\frac{1}{n+1} \leq \frac{1}{n}$ for all n, then the corresponding alternating series converges. Its value lies between $s_0 = 1$ and $s_1 = 1 - \frac{1}{2} = \frac{1}{2}$, $s_1 = \frac{1}{2}$ and $s_2 = \frac{1}{2} + \frac{1}{3} = \frac{5}{6}$, $s_2 = \frac{5}{6}$ and $s_3 = \frac{5}{6} - \frac{1}{4} = \frac{7}{12}$, etc. $\square$

Two other convergence tests are often used in practice: the ratio test and the root test.

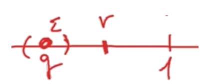
Theorem 74 (Ratio Test)
Let (a_n) be a sequence of positive real numbers.

1. *If* $\displaystyle\lim_{n\to\infty} \frac{a_{n+1}}{a_n} < 1$, *then* $\displaystyle\sum_{n=1}^{\infty} a_n$ *converges.*

2. *If* $\displaystyle\lim_{n\to\infty} \frac{a_{n+1}}{a_n} > 1$, *then* $\displaystyle\sum_{n=1}^{\infty} a_n$ *diverges.*

Proof:

1. Assume $0 \leq \frac{a_{n+1}}{a_n} \to q < 1$. Let $r = \frac{q+1}{2}$. Thus $q < r < 1$ and there are only finitely many indices n for which $\frac{a_{n+1}}{a_n} > r$. Indeed, let $\varepsilon \in (0, \frac{1-q}{2})$.



Then, $\exists N_\varepsilon \in \mathbb{N}$ such that

$$n > N_\varepsilon \Longrightarrow \frac{a_{n+1}}{a_n} - q < \varepsilon < \frac{1-q}{2} \Longrightarrow \frac{a_{n+1}}{a_n} \leq \frac{q+1}{2} = r.$$

Then

$$n > N_\varepsilon \Longrightarrow a_n = \frac{a_n}{a_{n-1}} \cdot \cdots \cdot \frac{a_{N+1}}{a_N} \cdot a_N \leq r^{n-N} a_N.$$

The tail of the original series converges, as

$$\sum_{n=N+1}^{\infty} a_n \leq \sum_{n=N+1}^{\infty} a_N r^{n-N} = \frac{a_N}{r^N} \sum_{n=N+1}^{\infty} r^n = \frac{a_N}{r^N}\left(\frac{r^{N+1}}{1-r}\right) < \infty,$$

where the last equation is left as an exercise. As $a_0 + \cdots + a_N$ is also finite, the full series converges.

2. Assume $\frac{a_{n+1}}{a_n} \to q > 1$. Using a similar argument as in part 1., we can show that $\exists r > 1$ and $N \in \mathbb{N}$ such that $\frac{a_{n+1}}{a_n} \geq r > 1$ for all $n \in \mathbb{N}$, so that $a_{n+1} > a_n$ for all $n \geq 1$.

   Thus $a_n \not\to 0$, and so $\sum_{n=0}^{\infty} a_n$ diverges, according to Theorem 70. ∎

If $\frac{a_{n+1}}{a_n} \to 1$, then the series may converge or may diverge, depending on the specific nature of a_n. The key parts of the proof (namely, the convergence of the tail in the first case and the condition $a_n \not\to 0$ in the second) are also valid if the statement is relaxed to some extent.

Theorem 74 (Ratio Test Reprise)
Let (a_n) be a sequence of real numbers with $a_n \neq 0$ for all n.

1. *If* $\limsup_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| < 1$, *then* $\sum_{n=1}^{\infty} a_n$ *converges.*
2. *If* $\liminf_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| > 1$, *then* $\sum_{n=1}^{\infty} a_n$ *diverges.*

The root test is similar (we will not prove it).

Theorem 75 (Root Test)
Let (a_n) be a sequence of positive real numbers.

1. *If* $\limsup_{n\to\infty} \sqrt[n]{a_n} < 1$, *then* $\sum_{n=1}^{\infty} a_n$ *converges.*
2. *If* $\liminf_{n\to\infty} \sqrt[n]{a_n} > 1$, *then* $\sum_{n=1}^{\infty} a_n$ *diverges.*

This general result also has a stricter version, replacing lim sup and lim inf by lim. In either version, if the limit is 1, then the series may converge or diverge, depending on the specific nature of the terms a_n.

Examples: discuss the convergence of

$$\sum_{n=1}^{\infty} \frac{(-1)^n}{n2^n}, \quad \sum_{n=1}^{\infty} \frac{3^n}{n2^n}, \quad \text{and} \quad \sum_{n=1}^{\infty} \frac{1}{n^p},\ p > 0.$$

1. The terms are all non-zero. We compute

$$\lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| = \lim_{n\to\infty}\left|\frac{(-1)^{n+1}}{(n+1)2^{n+1}}\cdot\frac{n2^n}{(-1)^n}\right| = \frac{1}{2}\lim_{n\to\infty}\left|\frac{n}{n+1}\right| = \frac{1}{2} < 1,$$

so the series converges according to the ratio test. □

2. The terms are all positive. We compute

$$\lim_{n\to\infty}\sqrt[n]{a_n} = \lim_{n\to\infty}\sqrt[n]{\frac{3^n}{n2^n}} = \frac{3}{2}\lim_{n\to\infty}\frac{1}{n^{1/n}} = \frac{3}{2} > 1,$$

so the series diverges according to the root test. □

3. The terms are all positive. For all $p > 0$, we compute

$$\lim_{n\to\infty}\left|\frac{1}{(n+1)^p}\cdot\frac{n^p}{1}\right| = \lim_{n\to\infty}\left(\frac{n}{n+1}\right)^p \to 1^p = 1.$$

Thus we cannot use the ratio test to determine if the series converges. If $p = 1$, the harmonic series is bounded below by a divergent series

$$\begin{aligned}\sum_{n=1}^{\infty}\frac{1}{n} &= 1+\frac{1}{2}+\frac{1}{3}+\frac{1}{4}+\frac{1}{5}+\frac{1}{6}+\frac{1}{7}+\frac{1}{8}+\cdots\\ &\geq 1+\frac{1}{2}+\underbrace{\frac{1}{4}+\frac{1}{4}}_{=1/2}+\underbrace{\frac{1}{8}+\frac{1}{8}+\frac{1}{8}+\frac{1}{8}}_{=1/2}+\cdots = 1+\frac{1}{2}+\frac{1}{2}+\frac{1}{2}+\cdots=\infty,\end{aligned}$$

and so must itself be divergent. As $\frac{1}{n^p} > \frac{1}{n}$ for all n when $p < 1$, then the series diverges for all $0 < p \leq 1$ according to the comparison theorem. If $p > 1$, the $p-$series is bounded above by a convergent series

$$\begin{aligned}\sum_{n=1}^{\infty}\frac{1}{n^p} &= 1+\frac{1}{2^p}+\frac{1}{3^p}+\frac{1}{4^p}+\frac{1}{5^p}+\frac{1}{6^p}+\frac{1}{7^p}+\frac{1}{8^p}+\cdots\\ &\leq 1+\underbrace{\frac{1}{2^p}+\frac{1}{2^p}}_{\text{2 times}}+\underbrace{\frac{1}{4^p}+\frac{1}{4^p}+\frac{1}{4^p}+\frac{1}{4^p}}_{\text{4 times}}+\frac{1}{8^p}+\cdots\\ &= 1+2^1\cdot\frac{1}{(2^1)^p}+2^2\cdot\frac{1}{(2^2)^p}+\cdots = \sum_{k=0}^{\infty}2^{k(1-p)} = \sum_{k=0}^{\infty}\frac{1}{(2^{p-1})^k}.\end{aligned}$$

But this series converges according to the root test. Indeed, all the terms are positive, and, because $p > 1$,

$$\lim_{k\to\infty}\sqrt[k]{\frac{1}{(2^{p-1})^k}} = \lim_{k\to\infty}\frac{1}{2^{p-1}} < 1.$$

Thus the $p-$series diverges for $0 < 1 \leq p$ and converges for $p > 1$. □

The next result (provided without proof) shows that the series of the absolute values may play an important role in the convergence of the "raw" series.

Theorem 76 (Absolute Convergence)
If the series $\sum_{n=0}^{\infty} |a_n|$ *converges, so does* $\sum_{n=0}^{\infty} a_n$ *(note that this is not an "if and only if" statement).*

The final result explains when the terms of a series can be re-arranged without affecting the convergence of the original series.

Theorem 77 (Series Re-arrangement)
If the series $\sum_{n=0}^{\infty} |a_n|$ *converges, so does* $\sum_{n=0}^{\infty} a_{\varphi(n)}$*,* $\varphi : \mathbb{N} \to \mathbb{N}$ *a bijection.*

6.2 Series of Functions

Series of functions play the same role for sequences of functions that series played for sequences of numbers. Let $I \subseteq \mathbb{R}$ and $f_n : I \to \mathbb{R}$, $\forall n \in \mathbb{N}$. If the sequence of partial sums

$$s_1(x) = f_1(x), \quad s_2(x) = f_1(x) + f_2(x), \quad s_3(x) = f_1(x) + f_2(x) + f_3(x), \quad \ldots$$

converges to some function $f : I \to \mathbb{R}$ for all $x \in I$, we say that the series of functions $\sum f_n$ **converges pointwise to** f **on** I.

Example: consider the sequence of functions $f_n : \mathbb{R} \to \mathbb{R}$, with $f_n(x) = x^n$ for each $n \in \mathbb{N}$. Does the sequence of partial sums $s_k(x)$ converge to some pointwise limit over some $A \subseteq \mathbb{R}$?

Solution: formally, we have

$$(1 - x^{k+1}) = (1 - x)(1 + x + x^2 + \cdots + x^k) = (1 - x)s_k(x).$$

Thus

$$x \neq -1 \Longrightarrow s_k(x) = \sum_{n=0}^{k} x^n = \frac{1 - x^{k+1}}{1 - x},$$

and so

$$\sum_{n=0}^{\infty} x^n = \lim_{k\to\infty} s_k(x) = \frac{1}{1 - x}$$

when $x \in (-1, 1)$. $\square$

If the sequence of partial sums (s_n) converges uniformly to f on I, we say that the series of functions $\sum f_n$ **converges uniformly to** f **on** I. If the convergence of the series of functions is uniform, the limit interchange theorems can be applied.

Theorem 78 (Cauchy Criterion for Series of Functions)
Let $f_n : I \to \mathbb{R}$ for all $n \in \mathbb{N}$. The series of functions with term f_n converges uniformly to some function $f : I \to \mathbb{R}$ if and only if $\forall \varepsilon > 0$, $\exists N_\varepsilon \in \mathbb{N}$ (independent of $x \in I$) such that

$$m > n > N_\varepsilon \Longrightarrow \left| \sum_{i=n+1}^{m} f_i(x) \right| < \varepsilon.$$

Proof: the proof follows directly from Theorem 66 applied to the sequence of partial sums $s_m : I \to \mathbb{R}$. ■

The next result is a powerful tool to prove uniform convergence (and as a pre-requisite to the use of the limit interchange theorems). The simplicity of its proof belies its importance.

Theorem 79 (Weierstrass $M-$Test)
Let $f_n : I \to \mathbb{R}$ and $M_n \geq 0$ for all $n \in \mathbb{N}$. Assume that $|f_n(x)| \leq M_n$ for all $x \in I$, $n \in \mathbb{N}$. Then

$$\sum_{n=1}^{\infty} M_n \text{ converges} \Longrightarrow \sum_{n=1}^{\infty} f_n \text{ converges uniformly on } I.$$

Proof: let $\varepsilon > 0$. Since $\sum M_n$ converges, its sequences of partial sums (s_k) is Cauchy and $\exists K_\varepsilon \in \mathbb{N}$ such that

$$m > n > K_\varepsilon \Longrightarrow \sum_{i=n+1}^{m} M_i < \varepsilon.$$

But

$$m > n > K_\varepsilon \Longrightarrow \left| \sum_{i=n+1}^{m} f_i(x) \right| \leq \sum_{i=n+1}^{m} |f_i(x)| \leq \sum_{i=n+1}^{m} M_i < \varepsilon;$$

since K_ε is independent of $x \in I$, $\displaystyle\sum_{n=1}^{\infty} f_n$ converges uniformly on I. ■

The following example showcases its usefulness.

Example: let $\varepsilon \in (0,1)$. Consider the sequence of functions $g_n : \mathbb{R} \to \mathbb{R}$ defined by $g_n(x) = nx^{n-1}$ for each $n \in \mathbb{N}$. Does $\sigma_k(x) \rightrightarrows \sigma(x)$ on $I_\varepsilon = (-1+\varepsilon, 1-\varepsilon)$ for some σ? If so, find σ.

Solution: consider the sequence of functions $f_n : \mathbb{R} \to \mathbb{R}$ defined by $f_n(x) = x^n$ for each $n \in \mathbb{N}$, and the corresponding sequence of partial sums $s_k(x)$ defined by $s_k(x) = 1 + x + \cdots + x^k$.

We have already shown that $s_k(x) \to \frac{1}{1-x}$ pointwise on $(-1+\varepsilon, 1-\varepsilon)$. The partials sums s_k are differentiable on I_ε since

$$\sigma_k(x) = s_k'(x) = 1 + 2x + 3x^2 + \cdots + kx^{k-1}$$

are polynomials (in fact, σ_k is also continuous on I_ε). Furthermore, note that the sequence of derivatives of partial sums $\sigma_k(x)$ converge uniformly on I_ε. To show this, note that

$$|g_n(x)| = |nx^{n-1}| \leq n|1-\varepsilon|^{n-1} = M_n \quad \forall x \in I_\varepsilon,\ \forall n \in \mathbb{N}.$$

But

$$\sum_{n=0}^{\infty} M_n = \sum_{n=0}^{\infty} n(1-\varepsilon)^{n-1}.$$

Since

$$\lim_{n\to\infty} \frac{(n+1)(1-\varepsilon)^n}{n(1-\varepsilon)^{n-1}} = (1-\varepsilon) \lim_{n\to\infty} \frac{n+1}{n} = (1-\varepsilon) < 1,$$

then $\sum M_n$ converges according to the ratio test.

According to the Weierstrass $M-$test, then, $\sigma_k(x) \rightrightarrows \sigma(x)$ on I_ε for some function $\sigma : I_\varepsilon \to \mathbb{R}$. We can use the limit interchange theorem 68 to identify σ:

$$\sigma(x) = \lim_{k\to\infty} \sigma_k(x) = \lim_{k\to\infty} \frac{\mathrm{d}}{\mathrm{d}x}[s_k(x)] = \frac{\mathrm{d}}{\mathrm{d}x}\Big[\lim_{k\to\infty} s_k(x)\Big] = \frac{\mathrm{d}}{\mathrm{d}x}\Big[\frac{1}{1-x}\Big],$$

which is to say $\sigma(x) = \frac{1}{(1-x)^2}$. $\square$

Incidentally, Theorem 68 also tells us that $s_k(x) \rightrightarrows \frac{1}{1-x}$ on I_ε, for all $0 < \varepsilon < 1$, and that for all $k \in \mathbb{N}$ and $x \in I_\varepsilon$, $\varepsilon \in (0,1)$, we have

$$\sum_{n=0}^{\infty} \frac{\mathrm{d}^k}{\mathrm{d}x^k}[x^n] = \frac{\mathrm{d}^k}{\mathrm{d}x^k} \sum_{n=0}^{\infty} x^n = \frac{\mathrm{d}^k}{\mathrm{d}x^k}\Big(\frac{1}{1-x}\Big).$$

6.3 Power Series

A **power series** around its **center** $x = x_0$ is a formal expression of the form

$$\sum_{n=0}^{\infty} a_n(x-x_0)^n.$$

We have already seen an example of such a series, which converged uniformly on intervals containing $x_0 = 0$:

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x} \quad \text{on } I_\varepsilon = (-1+\varepsilon, 1-\varepsilon),\ \forall \varepsilon \in (0,1);$$

note, however, that the convergence is only **pointwise** on $(-1,1)$. The function $f : A \to \mathbb{R}$, $f(x) = \frac{1}{1-x}$ is defined for all $x \neq 1$, however, yet the power series $1 + x + x^2 + \cdots$ does not converge to f outside of $(-1,1)$.[1]

Examples: where do the following power series converge:

$$\sum_{n=0}^{\infty} x^n, \quad \sum_{n=1}^{\infty} (nx)^n, \quad \sum_{n=1}^{\infty} \left(\frac{x}{n}\right)^n ?$$

Solution: we have seen that the first power series converges only on $(-1,1)$.

The second power series obviously converges when $x = 0$. To show that it fails to converge on $\mathbb{R} \setminus \{0\}$, note that if $|x| > 0$, then $\exists N \in \mathbb{N}$ such that $N > \frac{2}{|x|}$ by the Archimedean property. Thus,

$$n > N \Longrightarrow |(nx)^n| = n^n |x|^n > 2^n$$

and the sequence $(nx)^n$ is unbounded, which means that the terms do not go to 0, and so the series diverges.

For the third power series, let $x \in \mathbb{R}$. By the Archimedean property, $\exists N \in \mathbb{N}$ such that $N > 2|x|$. Thus,

$$n > N \Longrightarrow \left| \left(\frac{x}{n}\right)^n \right| = \frac{|x|^n}{n^n} < \frac{1}{2^n}.$$

According to the Weierstrass $M-$test and Theorem 76, the series thus converges uniformly on $\mathbb{R}$. $\square$

[1] Power series are commonly used as a formal **guessing procedure** to solve differential equations, but this is not a topic we will tackle at the moment. It is also natural to try to determine for which functions $f : A \to \mathbb{R}$ (and which A) we can find a **sequence of coefficients** (a_n) such that

$$f(x) = \sum_{n=0}^{\infty} a_n (x - x_0)^n, \quad \forall x \in A;$$

questions of this ilk are more naturally answered in $\mathbb{C}$; a more complete treatment would be provided in a complex analysis course.

The **radius of convergence** of a power series $\sum_{n=0}^{\infty} a_n(x-x_0)^n$ is

$$R = \frac{1}{\limsup_{n\to\infty} |a_n|^{1/n}}.$$

If the limit exists, we can replace lim sup by lim. Intuitively, for all large enough n, we have:

$$-R^{-n} \leq -|a_n| \leq a_n \leq |a_n| \leq R^{-n},$$

so that

$$-\sum_{n>N} \left(\frac{x-x_0}{R}\right)^n \leq \sum_{n>N} a_n(x-x_0)^n \leq \sum_{n>N} \left(\frac{x-x_0}{R}\right)^n.$$

The bounds are geometric series, and they converge when $|x-x_0| < R$. We would then expect the original power series to converge on the **interval of convergence** $|x-x_0| < R$.

> **Theorem 80**
> *Let R be the radius of convergence of the power series*
>
> $$\sum_{n=0}^{\infty} a_n(x-x_0)^n.$$
>
> *Then, if*
>
> - $R = 0$*, the power series converges for* $x = x_0$ *and diverges for* $x \neq x_0$*;*
> - $R = \infty$*, the power series converges absolutely on* $\mathbb{R}$*, and*
> - $0 < R < \infty$*, the power series converges absolutely on* $|x-x_0| < R$*, diverges on* $|x-x_0| > R$*; the extremities must be analyzed separately.*
>
> **Proof:** follows immediately from the root test. ■

But we can provide a stronger convergence statement.

> **Theorem 81**
> *The power series of Theorem 80 converges uniformly on any compact sub-interval*
>
> $$[a,b] \subseteq (x_0 - R, x_0 + R).$$
>
> **Proof:** let $\ell = \max\{|a-x_0|, |b-x_0|\} < R$. For every $n \in \mathbb{N}$, set $M_n = \ell^n|a_n| \geq 0$ and $\varepsilon = \frac{1}{4}(R-\ell)$.

Since $\frac{1}{R} = \limsup_{n\to\infty} |a_n|^{1/n}$, $\exists N_\varepsilon \in \mathbb{N}$ such that $n > N_\varepsilon \implies |a_n| \leq (\frac{1}{R-\varepsilon})^n$. Thus, for all $n > N_\varepsilon$, we have

$$0 \leq M_n = \ell^n |a_n| = (R-4\varepsilon)^n |a_n| \leq \left(\frac{R-4\varepsilon}{R-\varepsilon}\right)^n = \Big(1 - \underbrace{\frac{3\varepsilon}{R-\varepsilon}}_{>0}\Big)^n,$$

so that

$$\begin{aligned}\sum_{n=0}^{\infty} M_n &= \sum_{n=0}^{N_\varepsilon} M_n + \sum_{n>N_\varepsilon} M_n \leq \sum_{n=0}^{N_\varepsilon} M_n + \sum_{n>N_\varepsilon} \left(1 - \frac{3\varepsilon}{R-\varepsilon}\right)^n \\ &\leq \sum_{n=0}^{N_\varepsilon} M_n + \sum_{n=0}^{\infty} \left(1 - \frac{3\varepsilon}{R-\varepsilon}\right)^n = \underbrace{\sum_{n=0}^{N_\varepsilon} M_n}_{\text{finite}} + \frac{R-\varepsilon}{3\varepsilon} < \infty.\end{aligned}$$

But for all $x \in [a,b]$, we have

$$|a_n(x-x_0)^n| \leq |a_n|\ell^n = M_n, \quad \text{for all } n \in \mathbb{N}.$$

According to Theorem 79, the power series converges uniformly on $[a,b]$. ■

In what follows, we let $f : (x_0 - R, x_0 + R) \to \mathbb{R}$ be the function defined by

$$f(x) = \sum_{n=0}^{\infty} a_n(x-x_0)^n, \quad \text{and} \quad s_N(x) = \sum_{n=0}^{N} a_n(x-x_0)^n;$$

these have multiple nice properties, courtesy of the limit interchange theorems.

Theorem 82
The function f is continuous on any closed bounded interval $[a,b] \subseteq (x_0 - R, x_0 + R)$.

Proof: the functions $a_n(x-x_0)^n$ are continuous on $[a,b]$ for all n, and

$$s_N(x) = \sum_{n=0}^{N} a_n(x-x_0)^n \rightrightarrows f(x) \text{ on } [a,b] \quad \text{when } N \to \infty.$$

According to Theorem 67, f is continuous on $[a,b]$. ■

We get more than continuity, however.

Theorem 83
Let $x \in (x_0 - R, x_0 + R)$. Then f is Riemann-integrable between x_0 and x and

$$\int_{x_0}^{x} f(t)\,dt = \sum_{n=0}^{\infty} \frac{a_n}{n+1}(x - x_0)^{n+1}.$$

Proof: without loss of generality, assume $x > x_0$. As in the proof of Theorem 82, $s_N(x) \rightrightarrows f(x)$ on $[x_0, x]$ when $N \to \infty$. Thus, according to Limit Interchange Theorem 69, we have

$$\begin{aligned}\int_{x_0}^{x} f(t)\,dt &= \lim_{N\to\infty} \int_{x_0}^{x} s_N(t)\,dt = \lim_{N\to\infty} \int_{x_0}^{x} \sum_{n=0}^{N} a_n(t - x_0)^n\,dt \\ &= \lim_{N\to\infty} \sum_{n=0}^{N} \int_{x_0}^{x} a_n(t - x_0)^n\,dt = \sum_{n=0}^{\infty} \frac{a_n}{n+1}(x - x_0)^{n+1},\end{aligned}$$

which completes the proof. ■

The last result shows that power series really do behave nicely on their convergence interval.

Theorem 84
The function f is differentiable on $(x_0 - R, x_0 + R)$ and

$$f'(x) = \sum_{n=1}^{\infty} n a_n (x - x_0)^{n-1}.$$

Proof: as $n^{1/n} \to 1$,

$$\limsup_{n\to\infty} (n|a_n|)^{1/n} = \limsup_{n\to\infty} n^{1/n} \cdot \limsup_{n\to\infty} |a_n|^{1/n} = \frac{1}{R},$$

so the radius of convergence of both power series is identical, and so, in particular, $s'_N(x)$ converges uniformly on any closed bounded interval $[a, b] \subseteq (x_0 - R, x_0 + R)$.

Thus, according to limit interchange theorem 68, we have

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}x}\big[f(x)\big] &= \lim_{N\to\infty} \frac{\mathrm{d}}{\mathrm{d}x}\big[s_N(x)\big] = \lim_{N\to\infty} \frac{\mathrm{d}}{\mathrm{d}x} \sum_{n=0}^{N} \big[a_n(x - x_0)^n\big] \\ &= \lim_{N\to\infty} \sum_{n=0}^{N} \frac{\mathrm{d}}{\mathrm{d}x}\big[a_n(x - x_0)^n\big] = \sum_{n=1}^{\infty} n a_n (x - x_0)^{n-1},\end{aligned}$$

which completes the proof. ■

How do we compute the power series coefficients a_n? Combining Theorems 82 and 84, we see that f is **smooth** in its interval of convergence (i.e., all of its derivatives are continuous).

> **Theorem 85**
> *If $R > 0$, then*
> $$a_n = \frac{f^{(n)}(x_0)}{n!}.$$
> **Proof:** if $x = x_0$, then $f(x_0) = a_0$, which corresponds to the case $n = 0$. When $n = k > 0$, then repeated application of Theorem 84 yields
> $$f^{(k)}(x) = \sum_{n=k}^{\infty} \frac{n!}{(n-k)!} a_n (x - x_0)^{n-k} \quad \text{on } (x_0 - R, x_0 - R).$$
> If we evaluate at $x = x_0$, we get $f^{(k)}(x_0) = k! a_k$, thus $a_k = \frac{f^{(k)}(x_0)}{k!}$. ■

As a corollary, if $\exists r > 0$ such that

$$f(x) = \sum_{n=0}^{\infty} a_n (x - x_0)^n \quad \text{and} \quad g(x) = \sum_{n=0}^{\infty} b_n (x - x_0)^n$$

and $f(x) = g(x)$ for all $x \in (x_0 - r, x_0 + r)$, then $a_n = b_n$ for all $n \in \mathbb{N}$.[2]

> **Example:** consider the function $f : \mathbb{R} \to \mathbb{R}$ defined by
> $$f(x) = \begin{cases} \exp(-1/x^2), & x \neq 0 \\ 0, & x = 0 \end{cases}$$
> Show that f does not have a power series expansion.
>
> **Proof:** for all $n \in \mathbb{N}$, it can be shown that
> $$f^{(n)}(x) = \begin{cases} \frac{\mathrm{d}^n}{\mathrm{d}x^n}\left[\exp(-1/x^2)\right], & x \neq 0 \\ 0, & x = 0 \end{cases}$$
> is continuous and that $f^{(n)}(0) = 0$. According to the corollary to Theorem 85, if f is equal to its power series on some interval $(-r, r)$, then all the a_n would be 0, and so $f \equiv 0$, but $f \not\equiv 0$, so f cannot be equal to its power series expansion. ■

Thus, we cannot always assume that a function is equal to its power series. There are other ways to expand a function as an infinite series, most notably *via* **Laurent Series** and **Fourier Series**. These topics are covered in courses in complex analysis and partial differential equations, respectively, although we briefly discuss the latter in Chapter 11.

[2] Attempts to strengthen this **uniqueness** result must necessarily fail.

6.4 Solved Problems

1. Answer the following questions about series.

   a) If $\sum_{k=1}^{\infty}(a_k+b_k)$ converges, what about $\sum_{k=1}^{\infty}a_k$ and $\sum_{k=1}^{\infty}b_k$?

   b) If $\sum_{k=1}^{\infty}(a_k+b_k)$ diverges, what about $\sum_{k=1}^{\infty}a_k$ and $\sum_{k=1}^{\infty}b_k$?

   c) If $\sum_{k=1}^{\infty}(a_{2k}+a_{2k-1})$ converges, what about $\sum_{k=1}^{\infty}a_k$?

   d) If $\sum_{k=1}^{\infty}a_k$ converges, what about $\sum_{k=1}^{\infty}(a_{2k}+a_{2k-1})$?

   Solution:

   a) They might both diverge. Consider $a_k=-k$ and $b_k=k$. However, if one converges, then so does the other, by the arithmetic of limits/series.

   b) At least one of them diverges because if they both converged, then the series of sums would converge as well (according to a proposition seen in the notes).

   c) Nothing. Consider $a_{2k}=k$, $a_{2k+1}=-k$, for which $\sum_{k=1}^{\infty}a_k$ diverges, but $a_{2k}=\frac{1}{k^2}$, $a_{2k+1}=0$, for which $\sum_{k=1}^{\infty}a_k$ converges.

   d) It also converges. The sequence of partial sums of the second series is
   $$(a_1+a_2, a_1+a_2+a_3+a_4, , a_1+a_2+a_3+a_4+a_5+a_6,\ldots)$$
   is a subsequence of the sequence of partial sums of the first series
   $$(a_1, a_1+a_2, a_1+a_2+a_3, a_1+a_2+a_3+a_4,\ldots).$$
   If the first series sequence of partial sums converges, so does the subsequence's series. □

2. For all $r>1$, show that
   $$\frac{1}{r-1}=\frac{1}{r+1}+\frac{2}{r^2+1}+\frac{4}{r^4+1}+\frac{8}{r^8+1}+\cdots$$
   Solution: we see that
   $$\frac{1}{\ell+1}=\frac{1}{\ell-1}-\frac{2}{\ell^2-1}.$$
   Thus, for all $k\in\mathbb{N}$, if $\ell=2^k$, we have
   $$\begin{aligned}\frac{1}{r^{2^k}+1}&=\frac{1}{r^{2^k}-1}-\frac{2}{r^{2^{k+1}}-1}\\ \Longrightarrow\frac{2^k}{r^{2^k}+1}&=\frac{2^k}{r^{2^k}-1}-\frac{2^{k+1}}{r^{2^{k+1}}-1}.\end{aligned}$$

Therefore, we have a telescoping sum

$$\sum_{k=1}^{\infty} \frac{2^k}{r^{2^k}+1} = \lim_{n\to\infty} \sum_{k=1}^{n} \frac{2^k}{r^{2^k}+1} = \lim_{n\to\infty} \left(\frac{1}{r-1} - \frac{2^n}{r^{2^n}-1} \right) = \frac{1}{r-1},$$

where the last equation follows from the fact that, for $r > 1$, we have

$$\lim_{m\to\infty} \frac{m}{r^m} = 0.$$

This completes the proof. ■

3. Using Riemann integration, find the values of p for which the series $\displaystyle\sum_{n=1}^{\infty} \frac{1}{n^p}$ converges.

Solution: if $p \leq 0$, then $\frac{1}{n^p} \not\to 0$ so the series diverges. In what follows, then, let $p > 0$. For $k \in \mathbb{N}$, consider the function $f_{k;p} : [1,k] \to \mathbb{R}$ defined by $f_{k;p}(x) = \frac{1}{x^p}$. Since $f'_{k;p}(x) = -\frac{p}{x^{p+1}} < 0$ for all $x \geq 1$, $f_{k;p}$ is strictly decreasing on $[1,k]$. Thus $f_{k;p}$ is Riemann-integrable on $[1,k]$. Consider the partition $P_k = \{1, 2, \ldots, k, k+1\}$ of $[1, k+1]$. Since $f_{k;p}$ is Riemann-integrable,

$$L(f_{k;p}; P_k) \leq \int_1^{k+1} f_{k;p} \leq U(f_{k;p}; P_k).$$

As $f_{k;p}$ is decreasing on the sub-interval $[\mu, \nu]$, $f_{k;p}$ reaches its maximum at μ and its minimum at ν; Hence

$$\begin{aligned} U(f_{k;p}; P_k) &= \sum_{n=1}^{k} f_{k;p}(n)(n+1-n) = \sum_{n=1}^{k} \frac{1}{n^p}, \quad \text{and} \\ L(f_{k;p}; P_k) &= \sum_{n=2}^{k+1} f_{k;p}(n+1)(n+1-n) = \sum_{n=2}^{k+1} \frac{1}{n^p}. \end{aligned}$$

But

$$\sum_{n=2}^{k+1} \frac{1}{n^p} = \frac{1}{(k+1)^p} - 1 + \sum_{n=1}^{k} \frac{1}{n^p}.$$

Thus

$$\frac{1}{(k+1)^p} - 1 + \sum_{n=1}^{k} \frac{1}{n^p} \leq \int_1^{k+1} f_{k;p} \leq \sum_{n=1}^{k} \frac{1}{n^p}.$$

Write $s_{k;p}$ for the partial sum and note that

$$\int_1^{k+1} f_{k;p} = \int_1^{k+1} \frac{dx}{x^p} = \begin{cases} \ln(k+1), & \text{when } p = 1 \\ \frac{1}{1-p}(k^{1-p} - 1), & \text{when } p \neq 1 \end{cases}$$

If $p = 1$, then $\ln(k+1) \leq s_{k;1}$ for all k. Since the sequence $\{\ln(k+1)\}_k$ is unbounded, so must $\{s_{k;1}\}_k$ be unbounded, which means that the corresponding series cannot converge. If $p > 1$, then

$$\lim_{k\to\infty} \left(\frac{1}{1-p}(k^{1-p} - 1) + 1 - \frac{1}{(k+1)^p} \right) = \frac{p}{p-1}.$$

Since $s_{k;p}$ is monotone (as every additional $\frac{1}{n^p}$ added to the partial sum is positive) and since $s_{k;p}$ is bounded above by the convergent sequence

$$\left\{\frac{1}{1-p}(k^{1-p}-1)+1-\frac{1}{(k+1)^p}\right\}_k,$$

$s_{k;p}$ is a convergent sequence. If $p<1$, then

$$\left\{\frac{1}{1-p}(k^{1-p}-1)\right\}_k$$

is unbounded. As $s_{k;p}\geq \frac{1}{1-p}(k^{1-p}-1)$ for all k, $\{s_{k;p}\}$ is also unbounded, which means that the corresponding series cannot converge. Thus, the series converges if and only if $p>1$. $\square$

4. Which of the following series converge?

   a) $\displaystyle\sum_{n=1}^{\infty}\frac{n(n+1)}{(n+2)^2}$

   b) $\displaystyle\sum_{n=1}^{\infty}\frac{2+\sin^3(n+1)}{2^n+n^2}$

   c) $\displaystyle\sum_{n=1}^{\infty}\frac{1}{2^n-1+\cos^2 n^3}$

   d) $\displaystyle\sum_{n=1}^{\infty}\frac{n+1}{n^2+1}$

   e) $\displaystyle\sum_{n=1}^{\infty}\frac{n+1}{n^3+1}$

   f) $\displaystyle\sum_{n=1}^{\infty}\frac{n!}{n^n}$

   g) $\displaystyle\sum_{n=1}^{\infty}\frac{n!}{5^n}$

   h) $\displaystyle\sum_{n=1}^{\infty}\frac{n^n}{3^{1+2n}}$

   i) $\displaystyle\sum_{n=1}^{\infty}\left(\frac{5n+3n^3}{7n^3+2}\right)^n$

   Solution: we use the various tests at our disposal.

   a) Since

   $$\lim_{n\to\infty}\frac{n(n+1)}{(n+2)^2}=1\neq 0,$$

   the series diverges .

b) Since $-1 \le \sin^3(n+1) \le 1$, we have

$$0 \le \frac{2+\sin^3(n+1)}{2^n+n^2} \le \frac{1}{2^n+n^2} \le \frac{1}{2^n}.$$

Thus the given series converges by comparison with the geometric series $\displaystyle\sum_{n=1}^{\infty} \frac{1}{2^n}$.

c) If a_n denotes the n-th term of the series, we have

$$\frac{a_{n+1}}{a_n} = \frac{2^n - 1 + \cos^2 n^3}{2^{n+1} - 1 + \cos^2(n+1)^3} \to \frac{1}{2} < 1.$$

Thus the series converges by the ratio test.

d) We have

$$\frac{n+1}{n^2+1} \ge \frac{n}{2n^2} = \frac{1}{2n}.$$

Thus the series diverges by comparison with the harmonic series.

e) We have

$$0 \le \frac{n+1}{n^3+1} \le \frac{2n}{n^3} = \frac{2}{n^2}.$$

Thus the series converges by comparison with $\displaystyle\sum_{n=1}^{\infty} \frac{2}{n^2}$.

f) For $n \ge 2$, we have

$$0 \le \frac{n!}{n^n} = \frac{1}{n} \cdot \frac{2}{n} \cdot \frac{3 \cdot 4 \cdots n}{n^{n-2}} \le \frac{2}{n^2}.$$

Thus the series converges by comparison with $\displaystyle\sum_{n=1}^{\infty} \frac{2}{n^2}$.

g) If a_n denotes the n-th term in the series, we have

$$\frac{a_{n+1}}{a_n} = \frac{(n+1)!}{5^{n+1}} \frac{5^n}{n!} = \frac{n+1}{5} \to \infty.$$

Thus the series diverges by the ratio test.

h) We have

$$\left(\frac{n^n}{3^{1+2n}}\right)^{1/n} = \frac{n}{3^{2+1/n}} \to \infty.$$

Thus the series diverges by the root test.

i) We have

$$\left(\left(\frac{5n+3n^3}{7n^3+2}\right)^n\right)^{1/n} = \frac{5n+3n^3}{7n^3+2} \to \frac{3}{7} < 1.$$

Thus the series converges by the root test. $\square$

5. Give an example of a power series $\sum_{k=0}^{\infty} a_k x^k$ with interval of convergence $[-\sqrt{2}, \sqrt{2})$.

Proof: consider the series
$$\sum_{k=1}^{\infty} \frac{x^k}{k}.$$
We have
$$\limsup_{k\to\infty} \sqrt[k]{\frac{|x|^k}{k}} = \limsup_{k\to\infty} \frac{|x|}{\sqrt[k]{k}} = |x|.$$
Therefore, by the root test, the series converges when $|x| < 1$ and diverges for $|x| > 1$. For $x = 1$, the series is the harmonic series, which diverges. For $x = -1$, it is the alternating harmonic series, which converges. Thus, the series converges precisely on the interval $[-1, 1)$.

Now, replace x by $x/\sqrt{2}$. The corresponding power series is thus
$$\sum_{k=0}^{\infty} \frac{1}{\sqrt{2}^k k} x^k.$$
We have
$$\limsup_{k\to\infty} \sqrt[k]{\frac{|x|^k}{\sqrt{2}^k k}} = \limsup_{k\to\infty} \frac{|x|}{\sqrt{2}\sqrt[k]{k}} = \frac{|x|}{\sqrt{2}}.$$
The series converges on $\frac{|x|}{\sqrt{2}} < 1$ and diverges on $\frac{|x|}{\sqrt{2}} > 1$. For $x = \sqrt{2}$, the series is the harmonic series, which diverges. For $x = -\sqrt{2}$, it is the alternating harmonic series, which converges.

Thus, the series converges precisely on the interval $[-\sqrt{2}, \sqrt{2})$. ■

6. Find the values of x for which the following series converge:

a) $\sum_{n=1}^{\infty} (nx)^n$;

b) $\sum_{n=1}^{\infty} x^n$;

c) $\sum_{n=1}^{\infty} \frac{x^n}{n^2}$;

d) $\sum_{n=1}^{\infty} \frac{x^n}{n!}$.

Solution:

a) The series diverges whenever $x \neq 0$ since the terms $(nx)^n$ do not tend to zero when $n \to \infty$. (For large enough n, we have $n|x| \geq 1$.) Thus, this power series converges *only* at its centre.

b) The geometric series converges precisely on the interval $(-1, 1)$, and the series takes on the value $\frac{1}{1-x}$ there.

c) For $|x| \leq 1$, we have

$$\left|\frac{x^n}{n^2}\right| \leq \frac{1}{n^2},$$

and thus the series converges for these values of x. If $|x| > 1$, the terms $|x^n/n^2| \to \infty$, and so the series diverges. Hence the series converges precisely on the interval $[-1, 1]$.

d) Let $x \in \mathbb{R}$. Using the ratio test we have

$$\frac{x^{n+1}}{(n+1)!} \cdot \frac{n!}{x^n} = \frac{x}{n+1} \to 0.$$

Thus the series converges for all $x \in \mathbb{R}$ (and takes on the value e^x). ■

7. If the power series $\sum a_k x^k$ has radius of convergence R, what is the radius of convergence of the series $\sum a_k x^{2k}$?

Solution: the new series can be written as $\sum_{k=0}^{\infty} b_k x^k$, where $b_k = a_{k/2}$ if k is even and $b_k = 0$ if k is odd. Thus

$$\begin{aligned}\limsup_{k\to\infty} \sqrt[k]{|b_k|} &= \lim_{k\to\infty} \sqrt[k]{|a_{k/2}|} = \lim_{k\to\infty} \sqrt[2k]{|a_k|} = \lim_{k\to\infty} \left(\sqrt[k]{|a_k|}\right)^{1/2} \\ &= \left(\lim_{k\to\infty} \sqrt[k]{|a_k|}\right)^{1/2} = R^{1/2}.\end{aligned}$$

Therefore, the radius of convergence of the new series is $\sqrt{R}$. □

8. Obtain power series expansions for the following functions.

a) $\dfrac{x}{1+x^2}$;

b) $\dfrac{x}{(1+x^2)^2}$;

c) $\dfrac{x}{1+x^3}$;

d) $\dfrac{x^2}{1+x^3}$;

e) $f(x) = \displaystyle\int_0^1 \frac{1-e^{-sx}}{s}\, ds$, about $x = 0$.

Solution:

a) Since

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

we have

$$\frac{x}{1+x^2}=x\sum_{k=0}^{\infty}(-x^2)^k=\sum_{k=0}^{\infty}(-1)^k x^{2k+1}.$$

b) We know that, for $x\in(-1,1)$, $\dfrac{1}{1-x}=\displaystyle\sum_{k=1}^{\infty}x^k$. For any $-1<a<b<1$, the series $\displaystyle\sum_{k=1}^{\infty}kx^{k-1}$ converges uniformly on $[a,b]$. Indeed, let $c=\max\{|a|,|b|\}<1$. Then, for all $x\in[a,b]$, we have

$$|kx^{k-1}|\leq kc^{k-1}.$$

Now,

$$\frac{(k+1)c^k}{kc^{k-1}}=\frac{k+1}{k}c\to c\quad\text{as } k\to\infty.$$

Since $c<1$, the ratio test tells us that $\displaystyle\sum_{k=1}^{\infty}kc^{k-1}$ converges. Thus, $\displaystyle\sum_{k=1}^{\infty}kx^{k-1}$ converges uniformly by the Weierstrass M-test. Consequently, we have

$$\sum_{k=1}^{\infty}kx^{k-1}=\frac{\mathrm{d}}{\mathrm{d}x}\left(\frac{1}{1-x}\right)=\frac{1}{(1-x)^2},$$

and so for any $x\in[a,b]\subseteq(-1,1)$:

$$\frac{x}{(1+x^2)^2}=x\sum_{k=1}^{\infty}k(-x^2)^{k-1}=\sum_{k=1}^{\infty}(-1)^{k-1}kx^{2k-1}.$$

c) Using the geometric series, we have

$$\frac{x}{1+x^3}=x\sum_{k=0}^{\infty}(-x^3)^k=\sum_{k=0}^{\infty}(-1)^k x^{3k+1}.$$

d) Using the geometric series, we have

$$\frac{x^2}{1+x^3}=x^2\sum_{k=0}^{\infty}(-x^3)^k=\sum_{k=0}^{\infty}(-1)^k x^{3k+2}.$$

e) Since

$$e^x=\sum_{k=0}^{\infty}\frac{x^k}{k!},$$

we have

$$\frac{1-e^{-sx}}{s} = -\frac{1}{s}\sum_{k=1}^{\infty}\frac{(-sx)^k}{k!} = \sum_{k=1}^{\infty}(-1)^{k+1}\frac{s^{k-1}x^k}{k!}.$$

This series converges absolutely for all s and all x (use the ratio test or compare it to the series for e^x). Therefore, viewing it as a power series in s (for some fixed x), its interval of convergence is ∞, and its centre is 0. Thus the series can be integrated term by term:

$$\begin{aligned}\int_0^1 \frac{1-e^{-sx}}{s}\,ds &= \int_0^1 \sum_{k=1}^{\infty}(-1)^{k+1}\frac{s^{k-1}x^k}{k!}\,ds \\ &= \sum_{k=1}^{\infty}(-1)^{k+1}\left(\int_0^1 s^{k-1}\,ds\right)\frac{x^k}{k!} \\ &= \sum_{k=1}^{\infty}(-1)^{k+1}\left[\frac{s^k}{k}\right]_{s=0}^{s=1}\frac{x^k}{(k!)} = \sum_{k=1}^{\infty}(-1)^{k+1}\frac{x^k}{k(k!)}.\end{aligned}$$

□

6.5 Exercises

1. Prepare a 2-page summary of this chapter, with important definitions and results.
2. Prove the relaxed version of Theorem 74.
3. Prove Theorem 75, as well as its relaxed version.
4. Prove Theorem 76.
5. Prove Theorem 77.
6. Explain the infinite sums paradoxes of Chapter 2 in light of Theorems 76 and 77.

Part II

Real Analysis and Metric Spaces

Chapter 7

The Real Numbers (Reprise)

In a course on real analysis, the fundamental object of study is the set of real numbers. In Chapter 1, we introduced $\mathbb{R}$ in an intuitive and informal way. In this chapter, we show how $\mathbb{R}$ can be built using Cauchy sequences.

7.1 Cauchy Sequences in $\mathbb{Q}$

$(\mathbb{R}, |\cdot|)$ and $(\mathbb{Q}, |\cdot|)$ are both ordered fields. There is a fundamental difference between them, however: in $(\mathbb{R}, |\cdot|)$, every Cauchy sequence converges; in $(\mathbb{Q}, |\cdot|)$, some do not.

Lemma
If $(x_n) \subseteq \mathbb{Q}$ converges to $x \in \mathbb{Q}$, then (x_n^2) converges to $x^2 \in \mathbb{Q}$.

Proof: first, note that if $x \in \mathbb{Q}$, then $x^2 \in \mathbb{Q}$, since $\mathbb{Q}$ is a field. Now, let $\varepsilon > 0$. By hypothesis, $\exists N \in \mathbb{N}$ such that $n > N \implies |x_n - x| < \varepsilon$. Hence, for all $n > N$,

$$\begin{aligned}|x_n^2 - x^2| &= |x_n - x||x_n + x| < \varepsilon|x_n + x| \leq \varepsilon(|x_n| + |x|) \\ &= \varepsilon(|x_n - x + x| + |x|) \leq \varepsilon(|x_n - x| + 2|x|) < \varepsilon(\varepsilon + 2|x|).\end{aligned}$$

As ε can be made arbitrarily small, this completes the proof. ■

The following result sets the stage to show that $\mathbb{Q}$ is **incomplete** (see proof on pages 7-8).

Lemma
There is no rational number a for which $a^2 = 2$.

We build a sequence of rational numbers a_n for which $a_n^2 \to 2$:

$$a_1 = \frac{1}{1}, \quad a_2 = \frac{14}{10}, \quad a_3 = \frac{141}{100}, \quad a_4 = \frac{1414}{1000}, \quad \ldots$$

We can show by induction that

$$0 < a_1 < a_2 < \cdots < a_{n-1} < a_n < \cdots < 2 \quad \text{and} \quad 0 < a_1^2 < a_2^2 < \cdots < a_{n-1}^2 < a_n^2 < \cdots < 2.$$

For $n \in \mathbb{N}$, write $b_n = a_n + \frac{1}{10^{n-1}}$. Then $b_n^2 > 2 > a_n^2$ for all n.

Consequently, $a_n^2 \to 2$ since

$$|a_n^2 - 2| \leq |b_n^2 - a_n^2| = |b_n - a_n||b_n + a_n| \leq \frac{1}{10^{n-1}}\left(2a_n + \frac{1}{10^{n-1}}\right) \to 0.$$

It is easy to see that (a_n) is a Cauchy sequence in $\mathbb{Q}$; indeed, $|a_n - a_m| < 10^{-n}$ whenever $m \geq n$.

However, (a_n) cannot be a convergent sequence in $\mathbb{Q}$: were it to converge to a number $a \in \mathbb{Q}$, we would have $a_n^2 \to a^2 = 2 \in \mathbb{Q}$ according to the first Lemma, but $a \notin \mathbb{Q}$ according to the second Lemma.

A metric space (E, d) in which every Cauchy sequence also converges in (E, d) is termed **complete.**[1] The previous discussion shows that $(\mathbb{Q}, |\cdot|)$ is not complete.

7.2 Building $\mathbb{R}$ by Completing $\mathbb{Q}$

Is the fact that $\mathbb{Q}$ incomplete problematic? Not in the sense that arithmetic in $\mathbb{Q}$ is compromised. But it is still fairly inconvenient.

If we take a closer look at the formal definition, we notice that we can only claim a sequence to be convergent **once we know what its limit is**. But if we already know that the sequence has a limit, then it automatically converges.

At this stage, the main advantage a complete metric space holds over a non-complete one is simply that it allows one to talk about the convergence of a sequence without knowing a thing about its limit, save that it exists. But this does not change the fact that $\mathbb{Q}$ is not complete. What can we do about that?

The sequence (a_n) described previously does not converge in $\mathbb{Q}$, but its values get closer and closer to one of the "holes" of $\mathbb{Q}$.

If we fill up that hole (in effect starting the process of "completing" $\mathbb{Q}$), we may expect that the sequence would now converge in the bigger set. This leads to the following definition of the **real numbers** $\mathbb{R}$:

1. any Cauchy sequence in $\mathbb{Q}$ corresponds to a real number;

2. two Cauchy sequences (x_n) and (y_n) in $\mathbb{Q}$ define the same real number if $(x_n) \sim (y_n)$:

$$\forall \varepsilon > 0, \exists N \in \mathbb{N} \text{ such that } n > N \implies |x_n - y_n| < \varepsilon.$$

[1]We will discuss metric spaces in the coming chapters – for now, we simply think of it as a space in which we can compute the "distance" between points.

It is not too difficult to show that $\sim$ is an equivalence relation on the set of Cauchy sequences in $\mathbb{Q}$ (see exercises), and so we define $\mathbb{R}$ as the quotient set

$$\mathbb{R} = \{(x_n) \mid (x_n) \text{ is a Cauchy sequence in } \mathbb{Q}\}/\sim .$$

How does this definition of $\mathbb{R}$ compare with our usual intuition?

For starters, there should be an **addition** and a **multiplication** in $\mathbb{R}$ that behave as we expect them to (**commutative**, **associative**, **invertible, and so on**). We achieve this by endowing our definition of $\mathbb{R}$ with the following operations: if $\alpha = [(a_n)], \beta = [(b_n)] \in \mathbb{R}$, define

$$\alpha + \beta = [(a_n + b_n)] \quad \text{and} \quad \alpha\beta = [(a_n b_n)].$$

In order for this definition to make sense, we need to verify that if (a_n) and (b_n) are Cauchy sequences, then so are $(a_n + b_n)$ and $(a_n b_n)$, and that the choice or representative in the equivalence classes are irrelevant:

$$(a_n) \sim (a_n') \text{ and } (b_n) \sim (b_n') \Longrightarrow (a_n + b_n) \sim (a_n' + b_n') \text{ and } (a_n b_n) \sim (a_n' b_n').$$

The proof is left as an exercise, and relies on the following inequalities:

$$|(a_n + b_n) - (a_n' + b_n')| \leq |a_n - a_n'| + |b_n - b_n'|$$

and

$$|a_n b_n - a_n' b_n'| \leq |a_n||b_n - b_n'| + |b_n'||a_n - a_n'|$$

and on Cauchy sequences being bounded in $\mathbb{Q}$.

Finally, in order for $\mathbb{Q}$ to be a subset of $\mathbb{R}$, we complete its definition as follows: if $\alpha \in \mathbb{R}$ is such that

$$\alpha = [(a, a, a, \ldots)], \quad a \in \mathbb{Q},$$

we identify α with $a \in \mathbb{Q}$. Consequently, if a Cauchy sequence (b_n) converges to $b \in \mathbb{Q}$, the real number $\beta = [(b_n)]$ **is** the rational number b.

7.3 An Order Relation on $\mathbb{R}$

To show that $\mathbb{R}$ is indeed complete, we next need to introduce an order on $\mathbb{R}$. If (a_n) and (b_n) are Cauchy sequences in $\mathbb{Q}$, there are three possibilities:

1. $\exists N \in \mathbb{N}$ such that $(n > N \Longrightarrow a_n \geq b_n)$;
2. $\exists N \in \mathbb{N}$ such that $(n > N \Longrightarrow a_n \leq b_n)$, or
3. (a_n) and (b_n) "overlap" infinitely often, in which case we must have $(a_n) \sim (b_n)$.

Write $\alpha = [(a_n)]$ and $\beta = [(b_n)]$. We define an **order** $<$ on $\mathbb{R}$ as follows:

1. $\alpha \geq \beta$ if cases 1 or 3 hold;
2. $\alpha \leq \beta$ if cases 2 or 3 hold.

But it is not enough to write $\leq$ or $\geq$; we still need to show that the relation is indeed an **order** (this is left as an exercise).

Lemma
Let $\varepsilon \in \mathbb{Q}$ and $N \in \mathbb{N}$. If (a_n) is a Cauchy sequence in $\mathbb{Q}$ for which $a_n \leq \varepsilon$ for all $n > N$, then $\alpha = [(a_n)] \leq \varepsilon$.

Proof: it suffices to identify $\varepsilon \in \mathbb{Q}$ with the equivalence class of the constant sequence

$$[(\varepsilon, \varepsilon, \ldots)].$$

Then the above definition of $\leq$ in $\mathbb{R}$ yields the desired conclusion. ■

We see now why we define $\mathbb{R}$ using Cauchy sequence in $\mathbb{Q}$.

Theorem 86
Let (a_n) be a Cauchy sequence in $\mathbb{Q}$ and set $\alpha = [(a_n)] \in \mathbb{R}$. Then (a_n) converges to α in $\mathbb{R}$.

Proof: We want to show that given any (real) $\varepsilon > 0$, we can find an integer $N \in \mathbb{N}$ such that $|a_n - \alpha| < \varepsilon$ whenever $n > N$.

For all $n \in \mathbb{N}$, the sequence $(a_n, a_n, \ldots)$ defines the real number a_n; similarly, the sequence $(a_1, a_2, \ldots)$ defines the real number α. Consequently, the sequences

$$(a_n - a_1, a_n - a_2, \ldots, a_n - a_m, \ldots) \quad \text{and} \quad (|a_n - a_1|, |a_n - a_2|, \ldots, |a_n - a_m|, \ldots)$$

define respectively the real numbers $a_n - \alpha$ and $|a_n - \alpha|$.

Let $\varepsilon > 0$. Since (a_n) is a Cauchy sequence, there is an integer $N \in \mathbb{N}$ such that $|a_n - a_m| < \varepsilon$ (as rational numbers) for each $n, m > N$. Fix $n > N$. Then we have $|a_n - a_m| < \varepsilon$ (as rational numbers) whenever $m > N$; consequently, $|a_n - \alpha| < \varepsilon$. Since this holds whenever $n > N$, we have $a_n \to \alpha$ in $\mathbb{R}$. ■

As a corollary, every real number is the limit of a Cauchy sequence of rational numbers.

Theorem 87 (Completeness of $\mathbb{R}$)
$\mathbb{R}$ is complete.

Proof: let (α_n) be a Cauchy sequence in $\mathbb{R}$. We show that it converges in $\mathbb{R}$ as follows:

1. construct a sequence (a_n) in $\mathbb{Q}$ for which $|a_n - \alpha_n| < \frac{1}{10^n}$ (where a_n is viewed as the constant sequence);
2. verify that (a_n) is a Cauchy sequence in $\mathbb{Q}$ and denote the associated real number by α;
3. show that $\alpha_n \to \alpha$.

That is, once more, left as an exercise. ■

We have not put emphasis on the fact that there are multiple ways of completing sets, but the **completion** of $\mathbb{Q}$ is entirely dependent on the notion of closeness that is being used: traditionally, the metric we use is that two rational numbers are considered close to one another if their respective decimal expansions start to differ far to the **right** of the decimal point.

For instance, the distance between 23410.0001 and 23410.0008 is smaller than 10^{-3} because the decimal expansions start to differ at the 4th digit to the right of the decimal point. In base 10, if $q, r \in \mathbb{Q}$, then we can write

$$q = \sum_{i\in\mathbb{Z}} q_i 10^i, \quad r = \sum_{i\in\mathbb{Z}} r_i 10^i$$

Under the usual metric $d_{10}(q,r) = \left|\sum_{i\in\mathbb{Z}} (q_i - r_i)10^i\right|$, we have

$$\begin{aligned} d_{10}(23410.0001, 23410.0008) = \big| &\cdots + (0-0)10^n + \cdots + (0-0)10^5 \\ &+ (2-2)10^4 + (3-3)10^3 + (4-4)10^2 + (1-1)10^1 \\ &+ (0-0)10^0 + (0-0)10^{-1} + (0-0)10^{-2} \\ &+ (0-0)10^{-3} + (1-8)10^{-4} + (0-0)10^{-5} + \cdots \\ &+ (0-0)10^{-n} + \cdots \big| = 7\cdot 10^{-4}. \end{aligned}$$

But that is an artificial convention. What would happen if we defined a metric the other way? Two rational numbers would be considered close to one another if their respective decimal expansions start to differ far to the **left** of the decimal point, say.

Under this new metric $\tilde{d}_{10}(q,r) = \left|\sum_{i\in\mathbb{Z}} (q_i - r_i)10^{-i}\right|$, we have

$$\begin{aligned} \tilde{d}_{10}(23410.0001, 23410.0008) = \big| &\cdots + (0-0)10^{-n} + \cdots + (0-0)10^{-5} \\ &+ (2-2)10^{-4} + (3-3)10^{-3} + (4-4)10^{-2} + (1-1)10^{-1} \\ &+ (0-0)10^0 + (0-0)10^1 + (0-0)10^2 + (0-0)10^3 \\ &+ (1-8)10^4 + (0-0)10^5 + \cdots + (0-0)10^n + \cdots \big| = 7\cdot 10^4, \end{aligned}$$

so that 23410.0001 and 23410.0008 are actually far apart, whereas 20000000012 and 12 are very close to one another since $\tilde{d}_{10}(20000000012, 12) = 2 \cdot 10^{-10}$.

If $\tilde{d}_{10}$ is indeed a metric on $\mathbb{Q}$ (see exercise 10), then Cauchy sequences in $(\mathbb{Q}, d)$ will not have a lot in common with Cauchy sequences in $(\mathbb{Q}, \tilde{d})$. There is no reason to expect that the completion of $\mathbb{Q}$ will be the same in both instances, and in fact, it is not.

When we complete $\mathbb{Q}$ using the metric $\tilde{d}_p$, where p is a prime integer, the resulting set we obtain is called the **field of $p-$adic numbers**, and it is distinct from $\mathbb{R}$. Just about everything we will do in these course notes could also apply to these new sets.

The moral of the story is that different metrics lead to different completions of $\mathbb{Q}$, and that neither of those is intrinsically superior to the others.

7.4 Exercises

1. Prepare a 2-page summary of this chapter, with important definitions and results.
2. Show that the relation $(x_n) \sim (y_n)$ is an equivalence relation on the space of Cauchy sequences in $\mathbb{Q}$ (i.e., show that it is reflexive, symmetric, and transitive).
3. If (a_n) and (b_n) are Cauchy sequences in $\mathbb{Q}$, show that so are $(a_n + b_n)$ and $(a_n b_n)$.
4. If (a_n), (b_n), (a_n') and (b_n') are Cauchy sequences in $\mathbb{Q}$ such that $(a_n) \sim (a_n')$ and $(b_n) \sim (b_n')$, show that $(a_n + b_n) \sim (a_n' + b_n')$ and $(a_n b_n) \sim (a_n' b_n')$.
5. Show that $\mathbb{R}$ is a field.
6. If (a_n) and (b_n) are Cauchy sequences which "overlap" infinitely often, show that $(a_n) \sim (b_n)$.
7. Let $\alpha, \beta, \gamma \in \mathbb{R}$. If $\alpha \leq \beta$ and $\beta \leq \gamma$, show that $\alpha \leq \gamma$.
8. Let $\alpha, \beta \in \mathbb{R}$. If $\alpha \leq \beta$ and $\beta \leq \alpha$, show that $\alpha = \beta$.
9. Fill the details in the proof of Theorem 7.3.
10. Show that $\tilde{d}_{10}$ is a metric on $\mathbb{Q}$ (use the definition in Section 8.1.1).
11. Let p be a prime integer. What can you say about the field of $p-$adic numbers?

Chapter 8

Metric Spaces and Sequences

> One of the natural ways we can extend the concepts we have discussed in the previous chapters is by moving from $\mathbb{R}$ to $\mathbb{R}^m$. Some of the notions that generalize nicely to vectors and functions on vectors include **norms** and **distances**, **sequences**, and **continuity**.

The symbol $\mathbb{K}$ is sometimes used to denote either $\mathbb{R}$ or $\mathbb{C}$; $\mathcal{C}_{\mathbb{R}}([0,1])$ represents the $\mathbb{R}-$vector space of continuous functions $[0,1] \to \mathbb{R}$, and $\mathcal{F}_{\mathbb{R}}([0,1])$ represents the $\mathbb{R}-$vector space of functions $[0,1] \to \mathbb{R}$.

8.1 Preliminaries

Most of the results of the previous chapters rely heavily on the properties of the absolute value. Its fundamental role in $\mathbb{R}$ is as a measure of the **magnitude** of a real number: $|x|$ is the distance from the real number x to the origin.

We can generalize the concept of the absolute value to higher-dimensional spaces in various ways. In this chapter, we discuss **norms** and **metrics**, and the **topologies** they induce.

8.1.1 Norms, Metrics, and Topology

Let E be a $\mathbb{K}-$vector space, such as $\mathbb{R}$, $\mathbb{C}^n$ or $\mathcal{C}_{\mathbb{R}}([0,1])$, say. A **norm** over E is a mapping $\|\cdot\| : E \to \mathbb{R}$ for which the following properties hold:

1. $\forall \mathbf{x} \in E$, $\|\mathbf{x}\| \geq 0$;
2. $\|\mathbf{x}\| = 0 \iff \mathbf{x} = \mathbf{0}$;
3. $\forall \mathbf{x} \in E, \forall \lambda \in \mathbb{K}$, $\|\lambda \mathbf{x}\| = |\lambda| \|\mathbf{x}\|$, and
4. $\forall \mathbf{x}, \mathbf{y} \in E$, $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$.

If the 4 properties hold, we say that $(E, \|\cdot\|)$ is a **normed space**.

Examples

1. $\mathbb{R}$ is a normed space together with the absolute value $|\cdot|$.

2. $\mathbb{C}$ is a normed space together with the modulus $|\cdot|$.

3. $\mathbb{R}^n$ is a normed space together with the **Euclidean norm**

$$\|\mathbf{x}\|_2 = \|(x_1, \ldots, x_n)\|_2 = \sqrt{x_1^2 + \cdots + x_n^2}.$$

   The Euclidean norm over $\mathbb{R}^n$ will play a special role in our explorations: note that it is intimately linked to the **inner product**

$$(\cdot \mid \cdot) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}, \quad \text{defined by} \quad (\mathbf{x} \mid \mathbf{y}) = \sum x_i y_i \Longrightarrow \|\mathbf{x}\| = (\mathbf{x} \mid \mathbf{x})^{1/2}.$$

4. $E = \mathcal{C}_{\mathbb{R}}([0,1])$ together with the **sup norm** $\|f\|_\infty = \sup_{x\in[0,1]} |f(x)|$ is another important normed space.

5. For $p \geq 1$, the $p-$**norm** over $\mathbb{R}^n$ is defined as follows:

$$\|\mathbf{x}\| = \left(\sum_{i=1}^{n} |x_i|^p \right)^{1/p}.$$

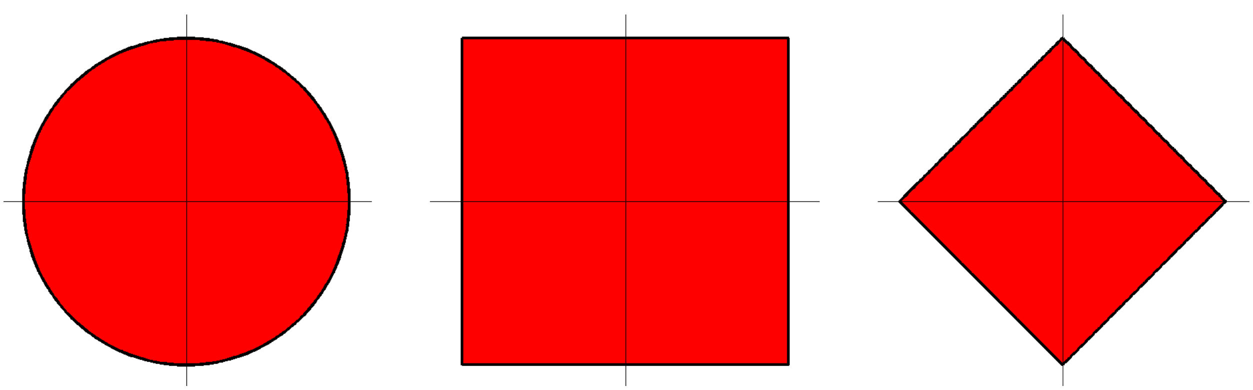
   Special cases of the $p-$norm over $\mathbb{R}^n$ include the Euclidean norm ($p = 2$), the sup norm ($p = \infty$) and the $1-$norm:

$$\|\mathbf{x}\|_\infty = \max_{1\leq i\leq n} |x_i|, \quad \|\mathbf{x}\|_\infty = \sum_{i=1}^{n} |x_i|. \qquad \square$$

The **open ball of radius** 1 **induced by the** $p-$**norm around the origin in** $\mathbb{R}^n$ is the set

$$B^p(\mathbf{0}, 1) = \{\mathbf{x} \in \mathbb{R}^n \mid \|\mathbf{x}\|_p < 1\};$$

different values of p leading to different geometrical sets $B^p(\mathbf{0}, 1)$: $p = 2, \infty, 1$ (left to right).[1]



[1]We can also talk of **closed balls**, or of **general balls** of radius r centered at some point $\mathbf{a} \in \mathbb{R}^n$.

The open balls have different shapes (only the regions in red, not the boundaries), but we will see that they are all equivalent, in the sense that they all induce the same topologies.

Since there are similarities between summation and integration (the Riemann-integral of a function over an interval is, essentially, the limit of a sum), it could tempting to conclude that there are equivalent $p-$norms over $\mathcal{F}_{\mathbb{R}}([0,1])$: something along the lines of

$$\|f\|_p = \Big(\int_{[0,1]} |f|^p \, dm\Big)^{1/p} \tag{8.1}$$

where m is the Lebesgue measure (see Chapters 21 and 26), but these mappings are not in fact norms on $\mathcal{F}_{\mathbb{R}}([0,1])$.

Indeed, consider the Dirichlet function $\chi_{\mathbb{Q}} \in \mathcal{F}_{\mathbb{R}}([0,1])$, say. It can be shown that $\|f\|_1 = 0$. However, $\chi_{\mathbb{Q}} \neq 0$ which contradicts the second property of norms (in fact, $\|\cdot\|_p$ is a **seminorm** on $\mathcal{F}_{\mathbb{R}}([0,1])$).

If we instead restrict the function space to $\mathcal{C}_{\mathbb{R}}([0,1])$, $\|\cdot\|_p$ is indeed a norm for all $p \geq 1$, but unfortunately, $(\mathcal{C}_{\mathbb{R}}([0,1]), \|\cdot\|_p)$ is not complete (more on this later).

Let E be any set. A **metric** over E is a mapping $d : E \times E \to \mathbb{R}$ for which the following properties hold:

1. $\forall \mathbf{x}, \mathbf{y} \in E$, $d(\mathbf{x}, \mathbf{y}) \geq 0$;
2. $\forall \mathbf{x} \in E$, $d(\mathbf{x}, \mathbf{x}) = 0$;
3. $d(\mathbf{x}, \mathbf{y}) = 0 \Longleftrightarrow \mathbf{x} = \mathbf{y}$;
4. $\forall \mathbf{x}, \mathbf{y} \in E$, $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$, and
5. $\forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in E$, $d(\mathbf{x}, \mathbf{y}) \leq d(\mathbf{x}, \mathbf{z}) + d(\mathbf{z}, \mathbf{y})$.

If the 5 properties hold, we say that (E, d) is a **metric space**.

An important property of such spaces is that every normed space gives rise to a metric space.

Theorem 88
Let $(E, \|\cdot\|)$ be a normed space, and define $d : E \times E \to \mathbb{R}$ by

$$d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|.$$

Then (E, d) is a metric space.

Proof: we show that all the metric space properties hold. Property 1, for instance, is a direct consequence of norm property 1:

$$\forall \mathbf{x}, \mathbf{y} \in E,\ d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\| \geq 0.$$

Properties 2 and 3 are a direct consequence of norm property 2:

$$\forall \mathbf{x} \in E, d(\mathbf{x}, \mathbf{x}) = \|\mathbf{x} - \mathbf{x}\| = \|\mathbf{0}\| = 0;$$
$$\forall \mathbf{x}, \mathbf{y} \in E, d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\| = 0 \iff \mathbf{x} - \mathbf{y} = \mathbf{0} \iff \mathbf{x} = \mathbf{y}.$$

Property 4 is a direct consequence of norm property 3:

$$\forall \mathbf{x}, \mathbf{y} \in E,\ d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\| = |-1| \cdot \|\mathbf{x} - \mathbf{y}\| = \|\mathbf{y} - \mathbf{x}\| = d(\mathbf{y}, \mathbf{x}).$$

Property 5 is a direct consequence of norm property 5:

$$\begin{aligned}\forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in E,\ d(\mathbf{x}, \mathbf{y}) &= \|\mathbf{x} - \mathbf{y}\| = \|\mathbf{x} - \mathbf{z} + \mathbf{z} - \mathbf{y}\| \\ &\leq \|\mathbf{x} - \mathbf{z}\| + \|\mathbf{z} - \mathbf{y}\| = d(\mathbf{x}, \mathbf{z}) + d(\mathbf{z}, \mathbf{y}).\end{aligned}$$

Thus (E, d) is a metric space. ■

Not every metric space arises from a norm, however.

Examples

1. Let E be any set and define $d : E \times E \to \mathbb{R}$ by

$$d(\mathbf{x}, \mathbf{y}) = \begin{cases} 0 & \text{if } \mathbf{x} = \mathbf{y} \\ 1 & \text{otherwise} \end{cases} \tag{8.2}$$

   Then (E, d) is a metric space in which every point is considered to be *far* from every other distinct point. We call such metric spaces **discrete**.

2. Let $E = \mathbb{R}^n$ and define $d : E \times E \to \mathbb{R}$ by $d_2(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|_2$. Then (E, d_2) is a metric space, which we usually refer to has having the **standard topology.** □

Let (E, d) be a metric space. The **open ball centered at a** $\in E$ **with radius** $r > 0$ is the set

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

the **closed ball centered at a** $\in E$ **with radius** $r > 0$ is the set

$$D(\mathbf{a}, r) = D_d(\mathbf{a}, r) = \{\mathbf{x} \in E \mid d(\mathbf{a}, \mathbf{x}) \leq r\},$$

and the **sphere centered at a** $\in E$ **with radius** $r > 0$ is the set

$$S(\mathbf{a}, r) = S_d(\mathbf{a}, r) = D(\mathbf{a}, r) \setminus B(\mathbf{a}, r) = \{\mathbf{x} \in E \mid d(\mathbf{a}, \mathbf{x}) = r\}.$$

Examples

1. Let $a \in E = \mathbb{R}$ and define $d(x,y) = |x-y|$ for all $x,y \in E$. Then, for $r > 0$, the balls reduce to intervals:

$$B(a,r) = (a-r, a+r), \quad D(a,r) = [a-r, a+r],$$

and the sphere to a discrete set $S(a,r) = \{a-r, a+r\}$.

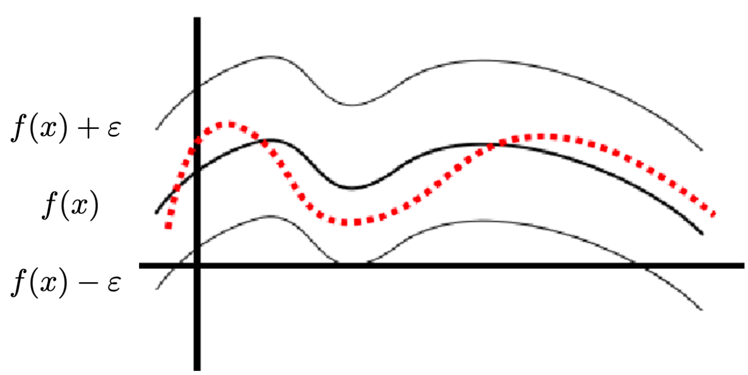
2. Let (E,d) be a discrete metric space and $\mathbf{a} \in E$. Then

$$B(\mathbf{a}, r) = \begin{cases} \{\mathbf{a}\}, & \text{if } r < 1 \\ E, & \text{otherwise} \end{cases}$$

3. Let $E = \mathcal{C}_{\mathbb{R}}([0,1])$, $d_\infty(f,g) = \|f-g\|_\infty$. Then, for $\varepsilon > 0$,

$$\begin{aligned} B(f,\varepsilon) &= \{g \in E \mid \|f-g\|_\infty < \varepsilon\} = \left\{g \in E \mid \sup_{x\in[0,1]} |f(x)-g(x)| < \varepsilon\right\} \\ &= \{g \in E \mid |f(x)-g(x)| < \varepsilon \; \forall x \in [0,1]\} \end{aligned}$$

We see $B(f,\varepsilon)$ in the image below; f is the solid curve in the middle, the two bounding curves are ε away from f, and the red dashes show a function g in $B(f,\varepsilon)$.



4. Let $A, B \neq \varnothing$ be subsets of a metric space (E,d). The **distance** between A and B is defined by

$$d(A,B) = \inf_{\mathbf{x}\in A, \mathbf{y}\in B} \{d(\mathbf{x},\mathbf{y})\}. \tag{8.3}$$

Unfortunately, d does not define a metric on $\wp(E) \setminus \varnothing$ (see exercise 10). When $A = \{\mathbf{x}\}$, we write $d(A,B) = d(\mathbf{x}, B)$. $\square$

Lemma 89
Let (E,d) be a metric space, $\mathbf{x},\mathbf{a} \in E$, $r > 0$ and $\mathbf{x} \notin B(\mathbf{a},r)$. Show that $d(\mathbf{x},B(\mathbf{a},r)) \geq d(\mathbf{x},\mathbf{a}) - r$.

Proof: for all $\mathbf{y} \in B(\mathbf{a},r)$, we have $d(\mathbf{x},\mathbf{y}) + d(\mathbf{y},\mathbf{a}) \geq d(\mathbf{x},\mathbf{a})$, whence

$$d(\mathbf{x},\mathbf{y}) \geq d(\mathbf{x},\mathbf{a}) - d(\mathbf{y},\mathbf{a}) \geq d(\mathbf{x},\mathbf{a}) - r.$$

Consequently,

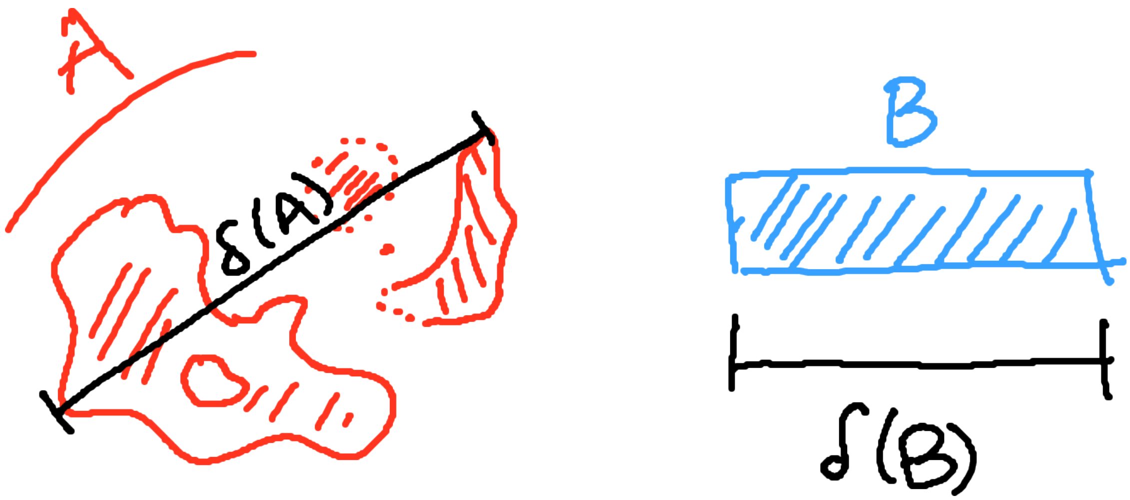
$$d(\mathbf{x},B(\mathbf{a},r)) = \inf_{\mathbf{y}\in B(\mathbf{a},r)} \{d(\mathbf{x},\mathbf{y})\} \geq d(\mathbf{x},\mathbf{a}) - r$$

whenever $\mathbf{x} \notin B(\mathbf{a},r)$. ■

Let (E,d) be a metric space and let $\varnothing \neq A \subseteq E$. The **diameter** of A under d is defined by

$$\delta_d(A) = \sup_{\mathbf{x},\mathbf{y}\in A} \{d(\mathbf{x},\mathbf{y})\}.$$

For instance, in $(\mathbb{R}^n, d_2)$, we have $\delta_{d_2}(B(\mathbf{a},r)) = 2r$; the diameter of two subsets $A, B \subseteq \mathbb{R}^2$ is illustrated below.

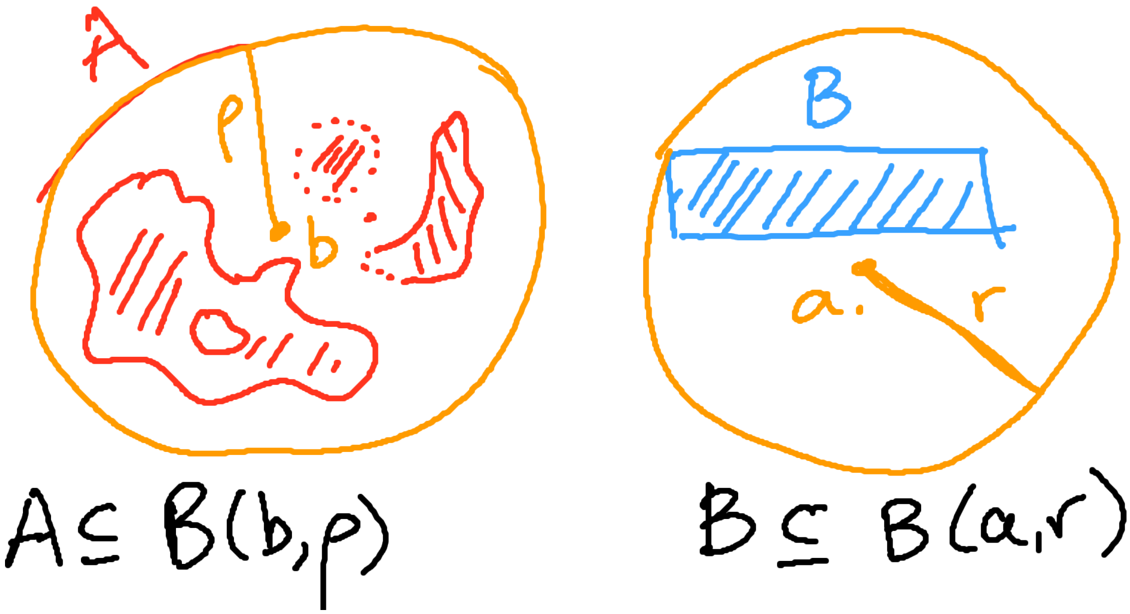


We say that A is **bounded in** (E,d) if $\delta_d(A) < \infty$.

Proposition 90
Let (E,d) be a metric space and let $\varnothing \neq A \subseteq E$. Then, A is bounded in (E,d) if and only if $\exists \mathbf{x} \in E$, $\exists r > 0$ such that $A \subseteq B(\mathbf{x},r)$.

Proof: one direction is immediate: if $\exists \mathbf{x} \in E$, $\exists r > 0$ such that $A \subseteq B(\mathbf{x},r)$, then $d(\mathbf{y},\mathbf{z}) < r$ for all $\mathbf{y},\mathbf{z} \in A \subseteq B(\mathbf{x},r)$, so that $\delta_d(A) \leq r$.

Conversely, if $\delta_d(A) \leq M$, say, then $d(\mathbf{y},\mathbf{z}) < r = M + 1$ for all $\mathbf{y},\mathbf{z} \in A$. Pick any $\mathbf{x} \in A$. Then for any other $\mathbf{y}$ in A, $d(\mathbf{x},\mathbf{y}) < r$, so that $\mathbf{y} \in B(\mathbf{x},r)$. Thus $A \subseteq B(\mathbf{x},r)$. ■

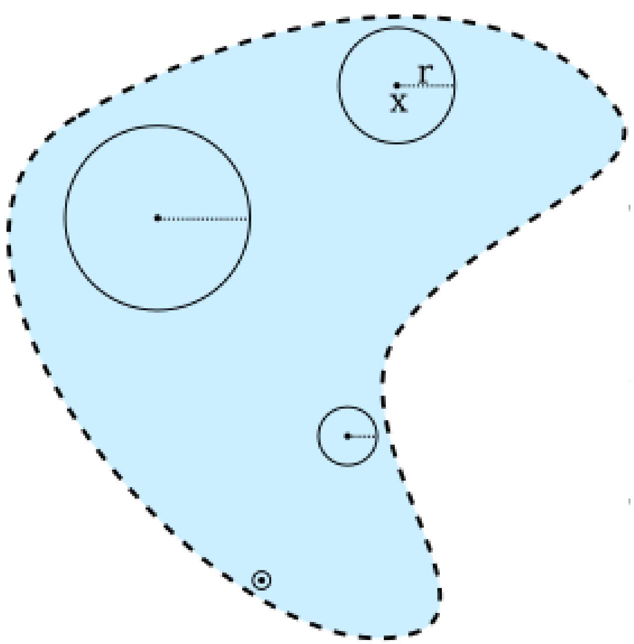


In this subsection, (E,d) is always a metric space, so we drop the d to lighten the text.

A subset $A \subseteq E$ is an **open subset of E under d** (or simply "open" if the context is clear) if either

- $A = \varnothing$, or
- $\forall \mathbf{x} \in E, \exists r > 0$ such that $B(\mathbf{x}, r) \subseteq A$.

We denote this relationship by $A \subseteq_O E$; an open subset of $\mathbb{R}^2$ in the Euclidean topology is shown below (D.J. Eck).

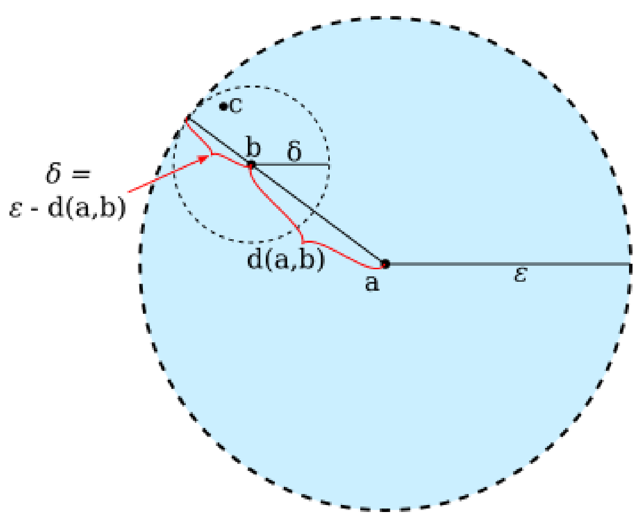


Proposition 91
Open sets in E have the following properties:

1. $E \subseteq_O E$;
2. $\forall \mathbf{a} \in E, r > 0$, *then* $B(\mathbf{a}, r) \subseteq_O E$;
3. *the union of an arbitrary family* $\{A_i\}_{i\in I}$ *of open subsets of* E *is an open subset of* E, *and*
4. *the intersection of a finite family* $\{A_i\}_{i=1}^{\ell}$ *of open subsets of* E *is an open subset of* E.

Proof:

1. Let $\mathbf{x} \in E$. Since $B(\mathbf{x}, r) \subseteq E$ for all $r > 0$, then $E \subseteq_O E$.
2. Let $B(\mathbf{a}, R)$ be an open ball in E, and let $\mathbf{x} \in B(\mathbf{a}, R)$. By definition, $d(\mathbf{a}, \mathbf{x}) < R$ implies $\exists \rho > 0$ with $\rho = \frac{R - d(\mathbf{a},\mathbf{x})}{2}$. It is not hard to show that with such a ρ, we have $B(\mathbf{x}, \rho) \subseteq B(\mathbf{a}, R)$.
3. Let $A = \bigcup A_i$. If $A = \varnothing$ then $A \subseteq_O E$. If $A \neq \varnothing$, let $\mathbf{x} \in A$. By definition, $\exists i \in I$ such that $\mathbf{x} \in A_i$. But $A_i \subseteq_O E$ and, as such, $\exists \rho > 0$ for which $B(\mathbf{x}, \rho) \subseteq A_i \subseteq \bigcup A_i = A$. Consequently, $A \subseteq_O E$.
4. It suffices to prove the result for $\ell = 2$ (why?). Let $A = A_1 \cap A_2$. If $A = \varnothing$ then $A \subseteq_O E$. If $A \neq \varnothing$, let $\mathbf{x} \in A$. Then $\mathbf{x} \in A_1$. But $A_1 \subseteq_O E$ and, as such, $\exists r_1 > 0$ for which $B(\mathbf{x}, r_1) \subseteq A_1 \subseteq A$. As well, $\mathbf{x} \in A_2$. But $A_2 \subseteq_O E$ and, as such, $\exists r_2 > 0$ for which $B(\mathbf{x}, r_2) \subseteq A_2 \subseteq A$. Set $\rho = \min\{r_1, r_2\}$. Then $B(\mathbf{x}, r) \subseteq A_1 \cap A_2$, and, consequently, $A \subseteq_O E$. ■



We have seen plenty of examples in Part I.

Examples

1. Let $a \in \mathbb{R}$. Then $(-\infty, a)$ and (a, ∞) are both open in $E = \mathbb{R}$ since
$$(-\infty, a) = \bigcup_{x<a}(x, a) \quad \text{and} \quad (a, \infty) = \bigcup_{x>a}(a, x).$$

2. The intersection of an arbitrary family of open subsets of E could be open, but need not be:
$$\bigcap_{n\in\mathbb{N}}(-n, n) = (-1, 1) \subseteq_O \mathbb{R},$$
but
$$\bigcap_{n\in\mathbb{N}}(-\tfrac{1}{n}, \tfrac{1}{n}) = \{0\} \text{ is not open in } \mathbb{R};$$
we will have more to say on the topic of arbitrary intersection of open sets in Part IV and Chapter 21. □

The collection of a metric space (E, d)'s open subsets forms a **topology** τ on E:

1. $\varnothing, E \in \tau$;
2. if $U_i \in \tau$ for all $i \in I$, then $\bigcup_I U_i \in \tau$, and
3. if $U_1, U_2 \in \tau$, then $U_1 \cap U_2 \in \tau$.

Examples

1. Let (E, d) be a metric space. The collection of all open subsets of E under d forms a topology on E, the **metric space topology**.
2. Let E be any set. The collection $\tau = \{\varnothing, E\}$ forms a topology on E, the **indiscrete topology**.
3. Let E be any set. The collection $\tau = \wp(E)$ forms a topology on E, the **discrete topology**. □

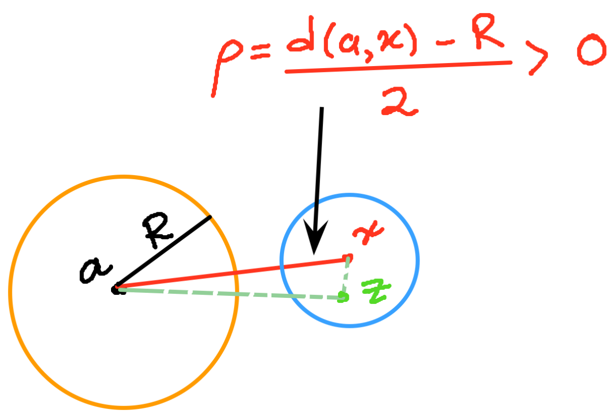
A subset $A \subseteq E$ is a **closed subset of** E **under** d if $E \setminus A \subseteq_O E$. We denote this relationship by $A \subseteq_C E$.

As a consequence of the definition of closed sets in opposition to open sets, we get a whole slew of properties of closed subsets, for free, such as $\varnothing, E \subseteq_C E$. But there are more substantial ones as well.

Examples

1. Every closed ball in (E, d) is closed.

   Proof: let $A = D(\mathbf{a}, R)$ be a closed ball in E and set
   $$E \setminus A = \{\mathbf{x} \in E \mid d(\mathbf{a}, \mathbf{x}) > R\}.$$
   We need to show that $E \setminus A$ is open. Let $\mathbf{x} \in E \setminus A$; by definition, $d(\mathbf{a}, \mathbf{x}) > R$ and $\rho = \frac{d(\mathbf{a},\mathbf{x})-R}{2} > 0$.

   It remains only to show that $B(\mathbf{x}, \rho) \subseteq E \setminus A$. Let $\mathbf{z} \in B(\mathbf{x}, \rho)$. Then
   $$d(\mathbf{x}, \mathbf{z}) < \rho \quad \text{and} \quad -d(\mathbf{x}, \mathbf{z}) > -\rho.$$
   Thus, according to the triangle inequality we have
   $$d(\mathbf{a}, \mathbf{z}) \geq d(\mathbf{a}, \mathbf{x}) - d(\mathbf{x}, \mathbf{z}) \geq 2\rho + R - d(\mathbf{x}, \mathbf{z}) \geq R + \rho > R;$$
   as such, $z \in E \setminus A$. This completes the proof. ■

2. Every sphere in (E, d) is closed.

   Proof: Let $S = S(\mathbf{a}, R)$. Note that
   $$E \setminus S = B(\mathbf{a}, R) \cup [E \setminus D(\mathbf{a}, R)] \subseteq_O E$$
   since it is a union of open sets. Consequently, $S \subseteq_C E$. ■

3. The intersection of an arbitrary family $\{A_i\}_{i \in I}$ of closed subsets of E is a closed subset of E. □

4. The union of a finite family $\{A_i\}_{i=1}^{\ell}$ of closed subsets of E is a closed subset of E. Note however that the union of an arbitrary family of closed subsets of E need not be closed (see exercise 18) in E. □

The **closure** of a subset $A \subseteq E$ with respect to a metric d is the smallest closed subset $\overline{A}$ of E (again, with respect to d) containing A (with possible equality).

The closure has a number of interesting properties, one of which being that $\overline{A}$ is the intersection of all closed sets containing A, and that $A \subseteq \overline{A}$ (see exercises 19 and 20).

Examples

1. In the Euclidean topology, $\overline{(0,1)} = [0,1]$.
2. In the discrete topology, $\overline{(0,1)} = (0,1)$.
3. In the Euclidean topology, $\overline{S(\mathbf{a},R)} = S(\mathbf{a},R)$. □

The closure provides us with a clear way to characterize closed subsets.

Lemma 92
Let A be a subset of E. Then $A \subseteq_C E \iff A = \overline{A}$.

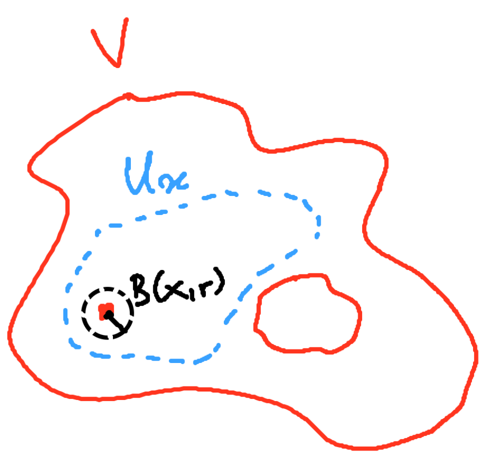
Proof: one direction is immediate. Let $A \subseteq_C E$. The smallest closed subset of E containing A is thus A itself, so $A = \overline{A}$.

Conversely, assume $A = \overline{A}$. As $\overline{A}$ is the smallest closed subset of A containing A, then $A = \overline{A}$ is closed in E. ■

A **neighbourhood** of $\mathbf{x} \in E$ is a subset $V \subseteq E$ containing an open subset $U_{\mathbf{x}} \subseteq_O E$ with $\mathbf{x} \in U_{\mathbf{x}}$. In other words, V is a neighbourhood of $\mathbf{x}$ if $\exists r > 0$ such that $B(\mathbf{x}, r) \subseteq V$ (but V is not necessarily open). The set of all neighbourhoods of $\mathbf{x}$ is denoted by

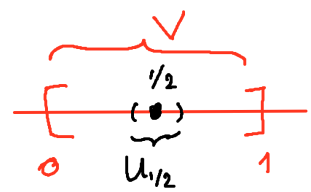
$$\mathcal{V}(\mathbf{x}) = \{V \subseteq E \mid V \text{ is a neighbourhood of } \mathbf{x}\}.$$

The image below shows a neighbourhood V of $\mathbf{x}$, with an open set $U_{\mathbf{x}}$.

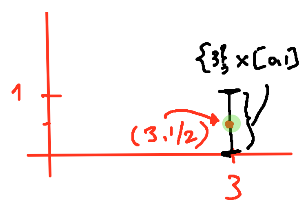


Examples

1. In $\mathbb{R}$ with the standard topology, $[0,1]$ and $(0,1]$ are neighbourhoods of $\frac{1}{2}$.



2. In $\mathbb{R}^2$ with the standard topology, $\{3\} \times [0,1]$ is not a neighbourhood of $(3, \frac{1}{2})$.



The various definitions give us an easy lemma.

Lemma 93
Let (E,d) be a metric space with $U \subseteq E$. Then U is a neighbourhood of each of its points if and only if $U \subseteq_O E$.

Proof: one direction holds as a consequence of the definition of open sets; the other as a consequence of the definition of neighbourhoods. ■

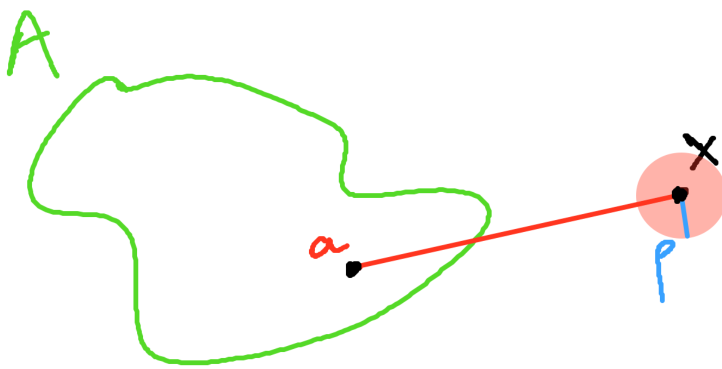
Points in $\overline{A}$ have useful (equivalent) properties.

Proposition 94
Let $A \subseteq E$. The following conditions are equivalent:

1. $\mathbf{x} \in \overline{A}$
2. $\forall \varepsilon > 0, \exists \mathbf{a} \in A$ *such that* $d(\mathbf{a}, \mathbf{x}) < \varepsilon$
3. $\forall V \in \mathcal{V}(\mathbf{x}), V \cap A \neq \varnothing$
4. $d(\{\mathbf{x}\}, A) = d(\mathbf{x}, A) = 0$

Proof: we will only prove that $1. \iff 2.$ The proof that $2. \iff 3 \iff 4.$ is left as an exercise.

Assume $\mathbf{x} \notin \overline{A}$. Then $\mathbf{x} \in E \setminus \overline{A} \subseteq_O E$. Thus $\exists \rho > 0$ such that $B(\mathbf{x}, \rho) \subseteq E \setminus \overline{A}$. Consequently, $d(\mathbf{a}, \mathbf{x}) \geq \rho, \forall \mathbf{a} \in A$.



Conversely, let $\mathbf{x} \in E$ and assume $\exists \varepsilon > 0$ such that

$$A \subseteq \underbrace{E \setminus B(\mathbf{x}, \varepsilon)}_{\text{closed}}.$$

Since $\overline{A}$ is the smallest closed set containing A, we must have

$$A \subseteq \overline{A} \subseteq E \setminus B(\mathbf{x}, \varepsilon)$$

and so $\mathbf{x} \notin \overline{A}$. ■

A subset A of E is **dense in** (E, d) if $\overline{A} = E$. A metric space (E, d) is **separable** if it has at least one dense subset.

Examples

1. $\mathbb{Q}$ and $\mathbb{R} \setminus \mathbb{Q}$ are both dense in $\mathbb{R}$ in the usual topology. □
2. Neither of these sets are dense in $\mathbb{R}$ in the discrete topology. □
3. Every non-empty subset of E is dense in E in the indiscrete topology. □
4. **Weierstrass' Theorem:** *let P be the set of polynomial functions $[0, 1] \to \mathbb{R}$. Then P is dense in $(\mathcal{C}_{\mathbb{R}}([0, 1]), d_\infty)$.*

   Thus real continuous functions on $[0, 1]$ (which need not even be C^1) can be approximated as closely as desired/needed by smooth (polynomial) functions (we will discuss this further in Chapter 23).
5. $\mathbb{R}$ and $\mathbb{R}^n$ are separable in the Euclidean topology. □

A family $\mathcal{G} = \{G_\lambda\}_{\lambda \in L}$, $\varnothing \neq G_\lambda \subseteq_O E$ forms a **basis for the open subsets of** E if every non-empty open subset of E can be written as a union of members of $\mathcal{G}$.

Examples

1. $\{B(x,r) \mid x \in \mathbb{Q}, r \in \mathbb{Q}^*_+\}$ and $\{B(x,r) \mid x \in \mathbb{R}, r \in \mathbb{R}^*_+\}$ both form a basis for the open subsets of $\mathbb{R}$. □

2. $\{B(\mathbf{x},r) \mid \mathbf{x} \in \mathbb{Q}^n, r \in \mathbb{Q}^*_+\}$ forms a basis for the open subsets of $\mathbb{R}^n$. □

There is a nice way to characterize such bases.

Proposition 95
A family $\mathcal{G} = \{G_\lambda\}_{\lambda \in L}$ is a basis for the open subsets of E if and only if $\forall \mathbf{x} \in E$, $\forall V \in \mathcal{V}(\mathbf{x})$, $\exists \lambda \in L$ such that $\mathbf{x} \in G_\lambda \subseteq V$.

Proof: the direction $\Longrightarrow$ holds as a consequence of the definition of neighbourhood and of a base.

Conversely, let $\varnothing \neq U \subseteq_O E$. Note that, being open, U is a neighbourhood of all its points. Then, by hypothesis, $\forall \mathbf{x} \in U$ $\exists \lambda(\mathbf{x}) \in L$ such that $\mathbf{x} \in G_{\lambda(\mathbf{x})} \subseteq U$. However,

$$U = \bigcup_{\mathbf{x} \in U} \{\mathbf{x}\} \subseteq \bigcup_{\mathbf{x} \in U} G_{\lambda(\mathbf{x})} \subseteq U,$$

so that U is the union of elements of $\mathcal{G}$. ■

By analogy with the closure, the **interior** of a subset $A \subseteq E$ is the largest open subset of E contained in A; we denote that subset by int(A) (or sometimes A°). It is not difficult to show that int(A) is the union of all the open subsets of E contained in A, and that $A \subseteq_O E$ if and only if int$(A) = A$ (see exercises).

Examples

1. In the discrete topology, int$([0,1]) = [0,1]$; while in the Euclidean topology, int$([0,1]) = (0,1)$. □

2. In the Euclidean topology, int$(S(\mathbf{a},R)) = \varnothing$ and int$(D(\mathbf{a},R)) = B(\mathbf{a},R)$. □

3. While int$(\overline{(a,b)}) = (a,b)$ and $\overline{\text{int}([a,b])} = [a,b]$ in $(\mathbb{R}, d_2)$, int$(\overline{W}) \neq W$, in general, as we can see with $W = (0, \frac{1}{2}) \cup (\frac{1}{2}, 1) \subseteq (\mathbb{R}, d_2)$. □

The next concepts are not crucial to our study, but still nice to have: $U \subseteq E$ is a **regular open subset of** E if int$(\overline{U}) = U$; $B \subseteq E$ is a **regular closed subset of** E if $\overline{\text{int}(B)} = B$.

Not all metrics are derived from a norm (the discrete metric fails in that regard, for instance), but **normed vector spaces** have a very nice property when it comes to closure and balls.

Lemma 96
If (E,d) is a normed vector space, then $D(\mathbf{0},1)=\overline{B(\mathbf{0},1)}$.

Proof: since $B(\mathbf{0},1)\subseteq D(\mathbf{0},1)\subseteq_C E$, we have $\overline{B(\mathbf{0},1)}\subseteq D(\mathbf{0},1)$ as $\overline{B(\mathbf{0},1)}$ since the **smallest** closed subset of E containing $B(\mathbf{0},1)$.

As $D(\mathbf{0},1)=B(\mathbf{0},1)\cup S(\mathbf{0},1)$, we only need to show that $S(\mathbf{0},1)\subseteq\overline{B(\mathbf{0},1)}$ as $B(\mathbf{0},1)\subseteq\overline{B(\mathbf{0},1)}$. Let $\mathbf{x}\in S(\mathbf{0},1)$; then $\|x\|=1$. Let $1>\varepsilon>0$ and set $\mathbf{z}=(1-\frac{\varepsilon}{2})\mathbf{x}$.

Then $\mathbf{z}\in B(\mathbf{0},1)$, since $\|\mathbf{z}\|=|1-\frac{\varepsilon}{2}|\cdot\|\mathbf{x}\|<1$; we note further that $d(\mathbf{z},\mathbf{x})=\|\mathbf{z}-\mathbf{x}\|=\frac{\varepsilon}{2}\|\mathbf{x}\|=\frac{\varepsilon}{2}<\varepsilon$ and so, according to Proposition 94 with $\mathbf{a}=\mathbf{z}$ and $A=B(\mathbf{0},1)$, we indeed have $\mathbf{x}\in\overline{B(\mathbf{0},1)}$. ■

We can use this lemma to show that the discrete metric is not derived from a norm: were it so, we would have $D(\mathbf{0},1)=\overline{B(\mathbf{0},1)}$. However, in $\mathbb{R}^n$ we have

$$B(\mathbf{0},1)=\{\mathbf{0}\}\subseteq_C\mathbb{R}\text{ and }D(\mathbf{0},1)=\mathbb{R}\Longrightarrow\overline{B(\mathbf{0},1)}=\{\mathbf{0}\}\neq\mathbb{R}=D(\mathbf{0},1).$$

Proposition 97
Let $A\subseteq E$. The following conditions are equivalent:

1. $\mathbf{x}\in\text{int}(A)$
2. $A\in\mathcal{V}(\mathbf{x})$
3. $\exists\varepsilon>0$ *such that* $B(\mathbf{x},\varepsilon)\subseteq A$.

Proof: by definition, we have $2.\iff 3.$ It remains only to show that $1.\iff 3.$

$3.\implies 1.$: Let $\varepsilon>0$ and $B(\mathbf{x},\varepsilon)\subseteq A$. Since $\text{int}(A)$ is the largest open subset of E contained in A and since $B(\mathbf{x},\varepsilon)$ is an open subset of E contained in A, we must have $B(\mathbf{x},\varepsilon)\subseteq\text{int}(A)$, whence $\mathbf{x}\in\text{int}(A)$.

$1.\implies 3.$: Let $\mathbf{x}\in\text{int}(A)\subseteq_O E$. By definition, there must exist some $\varepsilon>0$ such that $B(\mathbf{x},\varepsilon)\subseteq\text{int}(A)\subseteq A$. ■

As an example of the usefulness of this result, note that by the density of $\mathbb{Q}$ and its complement $\mathbb{R}\setminus\mathbb{Q}$ in $\mathbb{R}$, we automatically get $\text{int}(\mathbb{Q})=\text{int}(\mathbb{R}\setminus\mathbb{Q})=\varnothing$ with the usual topology on $\mathbb{R}$.

We end this section with a few other topological concepts:

- the **boundary** of a subset $A \subseteq E$ is simply defined by $\partial A = \overline{A} \setminus \text{int}(A)$ and the **exterior** of A is given by $\text{int}(E \setminus A)$;[2]
- we say that $\mathbf{x} \in E$ is a **cluster point** of A if

$$\forall \varepsilon > 0,\ \exists \mathbf{y}_\varepsilon \in B(\mathbf{x}, \varepsilon) \cap A \text{ such that } \mathbf{y}_\varepsilon \neq \mathbf{x};$$

- we also say that $\mathbf{x} \in E$ is an **isolated point** of A if $\exists \varepsilon > 0$ for which $B(\mathbf{x}, \varepsilon) \cap A = \{\mathbf{x}\}$.

Examples Let $A = \{\frac{1}{n} : n \geq 1\}$.

1. 0 is a cluster point of A since $B(0, \varepsilon) \cap A$ contains all $\frac{1}{n}$, where $n > \frac{1}{\varepsilon}$.
2. For all $n \geq 1$, $\frac{1}{n}$ is an isolated point of A, as $B(\frac{1}{n}, \frac{1}{2n(n+1)}) \cap A = \{\frac{1}{n}\}$.

There is a link between cluster points of a set and its closure.

Lemma 98
If $\mathbf{x}$ is a cluster point of A, then $\mathbf{x} \in \overline{A}$ and every neighbourhood of $\mathbf{x}$ contains an infinite set of points in A.

Proof: that $\mathbf{x} \in \overline{A}$ is a direct consequence of Propostion 94. The rest of the proof can be done by showing that if a neighbourhood of $\mathbf{x}$ exists which contain only a finite number of points of A, then $\mathbf{x}$ cannot be a cluster point of A. ■

Finally, if (E, d) is a metric space and $F \subseteq E$, then (F, d) is also a metric space, called a **metric subspace** of E. The topology on F is completely determined by the topology on E.

Proposition 99
Let (E, d) be a metric space and $F \subseteq E$. Then

$$B \subseteq_O F \iff \exists A \subseteq_O E \textit{ such that } B = A \cap F$$

and

$$B \subseteq_C F \iff \exists A \subseteq_C E \textit{ such that } B = A \cap F.$$

Proof: left as an exercise. ■

[2] In a nutshell, the exterior is the largest open subset of E which excludes A in its entirety.

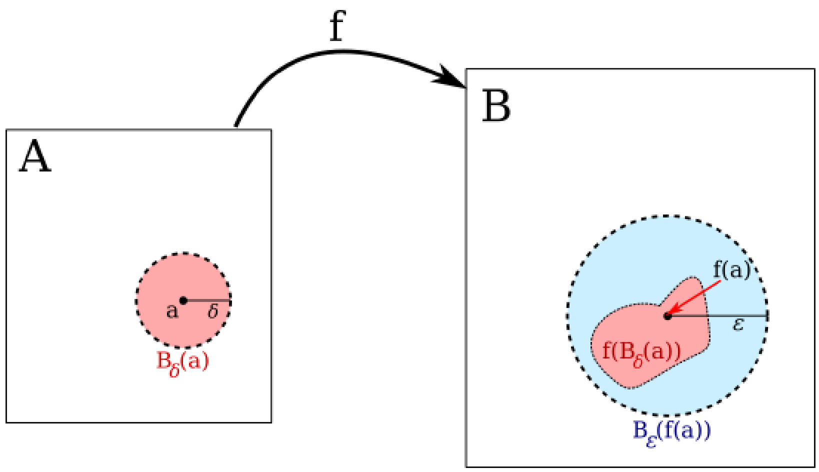
8.1.2 Continuity

The concept of **continuity** is fundamental in all aspects of analysis. Let $(A, d_A), (B, d_B)$ be metric spaces. Since we view $d_A(\mathbf{a}, \mathbf{x})$ and $d_B(f(\mathbf{a}), f(\mathbf{x}))$ as generalizations of $|a - x|$ and $|f(a) - f(x)|$, respectively, and we say that a map $f : A \to B$ is **continuous at a** $\in A$ if

$$\forall \varepsilon > 0,\ \exists \delta > 0, (\mathbf{x} \in A \text{ and } d_A(\mathbf{a}, \mathbf{x}) < \delta) \Longrightarrow d_B(f(\mathbf{a}), f(\mathbf{x})) < \varepsilon;$$

or, equivalently, if for any open $\varepsilon-$ball W centered at $f(\mathbf{a})$, there is an open $\delta-$ball V centered at **a** such that $f(V) \subseteq W$; or yet again equivalently, if for any neighbourhood $W \subseteq_O B$ of $f(\mathbf{a})$, there is a neighbourhood $V \subseteq_O A$ of **a** such that $f(V) \subseteq W$.[3]

The continuity of $f : (\mathbb{R}^2, d_2) \to (\mathbb{R}^2, d_2)$ at $\mathbf{a} \in \mathbb{R}^2$ is illustrated below (D.J. Eck).



We further say that the map f is **continuous on** A if it is continuous at each $\mathbf{a} \in A$.

Proposition 100
Let $(E, d), (\tilde{E}, \tilde{d})$ be metric spaces, and let $f : E \to \tilde{E}$. The following conditions are equivalent:

1. *f is continuous on E;*
2. *for any $W \subseteq_O \tilde{E}$, $f^{-1}(W) = \{\mathbf{x} \in E | f(\mathbf{x}) \in W\} \subseteq_O E$, and*
3. *for any $Y \subseteq_C \tilde{E}$, $f^{-1}(Y) \subseteq_C E$.*

Proof: that $2. \Longleftrightarrow 3.$ follows directly from the fact that

$$f^{-1}(\tilde{E} \setminus Y) = E \setminus f^{-1}(Y).$$

[3]That these definitions are equivalent is left as an exercise.

1. $\implies$ 2.: Let $W \subseteq_O \tilde{E}$ and $\mathbf{x} \in f^{-1}(W)$. Since W is open in $\tilde{E}$, $\exists \varepsilon > 0$ such that $B(f(\mathbf{x}), \varepsilon) \subseteq W$. By continuity, $\exists \delta > 0$ such that $f(B(\mathbf{x}, \delta)) \subseteq B(f(\mathbf{x}), \varepsilon) \subseteq W$. But this means that

$$B(\mathbf{x}, \delta) = f^{-1}(f(B(\mathbf{x}, \delta)) \subseteq f^{-1}(W)$$

(see exercises) and so $f^{-1}(W) \subseteq_O E$.

2. $\implies$ 1.: Let $f(\mathbf{x}) \in W \subseteq_O \tilde{E}$. Set $V = f^{-1}(W) \subseteq_O E$. Then $\mathbf{x} \in V$ and $f(V) \subseteq W$; consequently, f is continuous. ■

Consider a map $f : E \to \tilde{E}$ as above. If $f(W) \subseteq_O \tilde{E}$ for all $W \subseteq_O E$, then we say that f is an **open mapping**; by analogy, if $f(Y) \subseteq_C \tilde{E}$ for all $Y \subseteq_C E$, then we say that f is a **closed mapping**.

Generally speaking, continuous maps are neither open nor closed; the constant function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = a$ provides an example of a continuous function which is not open in the standard topology, as $(0, 1) \subseteq_O \mathbb{R}$, but $f((0, 1)) = \{a\} \subseteq_C \mathbb{R}$, for instance.

Proposition 101
Let $f : (E, d) \to (\tilde{E}, \tilde{d})$ and $g : (\tilde{E}, \tilde{d}) \to (\hat{E}, \hat{d})$ be continuous. Then the composition $g \circ f : (E, d) \to (\hat{E}, \hat{d})$ is continuous.

Proof: let $\mathbf{a} \in E$ and $\varepsilon > 0$. As g is continuous at $f(\mathbf{a}) \in \tilde{E}$, $\exists \delta_\varepsilon > 0$ such that

$$\mathbf{y} \in \tilde{E} \text{ and } \mathbf{y} \in B_{\tilde{d}}(f(\mathbf{a}), \delta_\varepsilon) \implies g(\mathbf{y}) \in B_{\hat{d}}(g(f(\mathbf{a})), \varepsilon).$$

Since f is continuous at $\mathbf{a}$, $\exists \eta_{\delta_\varepsilon} = \eta_\varepsilon > 0$ such that

$$\mathbf{x} \in E \text{ and } \mathbf{x} \in B_d(\mathbf{a}, \eta_{\delta_\varepsilon}) \implies f(\mathbf{x}) \in B_{\tilde{d}}(f(\mathbf{a}), \delta_\varepsilon).$$

Combining these results together, we get

$$\mathbf{x} \in E \text{ and } \mathbf{x} \in B_d(\mathbf{a}, \eta_{\delta_\varepsilon}) \implies g(f(\mathbf{x})) \in B_{\hat{d}}(g(f(\mathbf{a})), \varepsilon),$$

which completes the proof. ■

As we can see, in many instances, the broad strokes of proofs in the multi-dimensional cases follow those of the corresponding one-dimensional proofs.

Corollary 102 *Let $f : (E, d) \to (\tilde{E}, \tilde{d})$ be a continuous function. If $F \subseteq E$, then the restriction $f|_F : (F, d|_F) \to (\tilde{E}, \tilde{d})$ is continuous.*

Proof: it suffices to show that the inclusion $F \hookrightarrow E_1$ is continuous, which is left as an exercise, and then to apply Proposition 101. ■

Some standard examples are shown below.

Examples

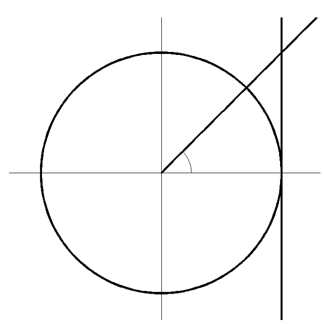
1. The functions $f : (\mathbb{R}, d_2) \to (\mathbb{R}, d_2)$ defined by $f(x) = x^3$ is continuous. $\square$
2. The identity function $\text{id} : (\mathbb{R}, d_{\text{discrete}}) \to (\mathbb{R}, d_2)$ is continuous, since $\text{id}^{-1}(V) = V \subseteq_O (\mathbb{R}, d_{\text{discrete}})$ for all $V \subseteq_O (\mathbb{R}, d_2)$. $\square$
3. The identity function $\text{id}^{\text{inv}} : (\mathbb{R}, d_2) \to (\mathbb{R}, d_{\text{discrete}})$ is not continuous, since, for instance,
$$\left(\text{id}^{\text{inv}}\right)^{-1}(\{a\}) = \{a\}$$
is not open in $(\mathbb{R}, d_2)$ even though $\{a\} \subseteq_O (\mathbb{R}, d_{\text{discrete}})$. $\square$
4. Consider the characteristic function $\chi_{\mathbb{R}\setminus\mathbb{Q}} : \mathbb{R} \to \mathbb{R}$. Then $\chi_{\mathbb{R}\setminus\mathbb{Q}}$ is continuous when restricted to $\mathbb{Q}$ (being a constant function), but $\chi_{\mathbb{R}\setminus\mathbb{Q}}$ is nowhere continuous on $\mathbb{R}$. $\square$

A metric d on E gives rise to a topology by **defining the open sets** of E. A natural question to ask is: can two different metrics give rise to the same topology? In order to answer that question, we need to introduce a new concept.

Let $(E, d), (\tilde{E}, \tilde{d})$ be metric spaces. A function $f : E \to \tilde{E}$ is a **homeomorphism** if f is bijective and both f and f^{inv} are continuous.[4]

Examples

1. $f : (\mathbb{R}, d_2) \to (\mathbb{R}, d_2)$, $f(x) = x^3$, is a homeomorphism. $\square$
2. $\text{id} : (\mathbb{R}, d_{\text{discrete}}) \to (\mathbb{R}, d_2)$, $\text{id}(x) = x$, is not a homeomorphism. $\square$
3. The function $g : (\mathbb{R}, d_2) \to \left((-\frac{\pi}{2}, \frac{\pi}{2}), d_2\right)$ defined by $g(x) = \arctan(x)$ is a homeomorphism. $\square$



[4]Alternatively, f is a homeomorphism if it is bijective, continuous and open.

These examples illustrate that the notion of **boundedness** is not necessarily preserved by homeomorphisms: for instance, $\mathbb{R}$ is unbounded while $(-\frac{\pi}{2}, \frac{\pi}{2})$ is bounded, but both spaces are homemorphic to one another *via* arctan.

Furthermore, neither is the notion of **distance** necessarily preserved by homeomorphisms: in general,

$$d(x_1, x_2) \neq \tilde{d}(f(x_1), f(x_2)).$$

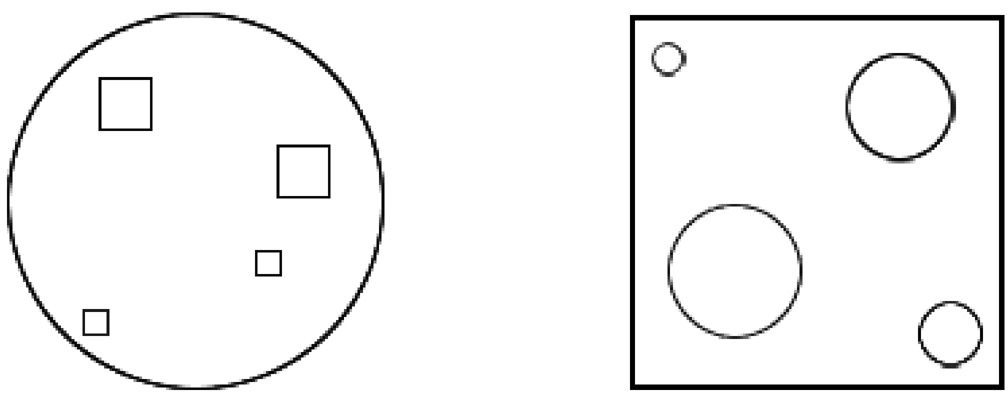
For instance, in the first example,

$$d(0,2) = |0-2| = 2 \neq \tilde{d}(0^3, 2^3) = |0^3 - 2^3| = 9.$$

However, homeomorphisms $f : E \to \tilde{E}$ **preserve the topologies of** E **and** $\tilde{E}$**:**

$$W \subseteq_O E \Longleftrightarrow f(W) \subseteq_O \tilde{E} = f(E)$$
$$Y \subseteq_C E \Longleftrightarrow f(Y) \subseteq_C \tilde{E} = f(E).$$

Two metrics $d, \tilde{d}$ on E are **topologically equivalent** if $\text{id} : (E,d) \to (E,\tilde{d})$ is a homeomorphism. In that case, d and $\tilde{d}$ give rise to the same topologies on E.

> **Example:** if $p, q \geq 1$, d_p and d_q induce the same topologies on $\mathbb{R}^n$.
>
> For instance, to show that d_2 and d_∞ are topologically equivalent in $\mathbb{R}^2$, it suffices to show that any point of a $2-$ball has an $\infty-$neighbourhood contained in the $2-$ball, and, conversely, that any point of an $\infty-$ball has a $2-$neighbourhood contained in the $\infty-$ball (see exercises). In the illustration below, we see a $2-$ball filled with $\infty-$balls (left) and an $\infty-$ball filled with with $2-$balls (right). $\square$
>
> 

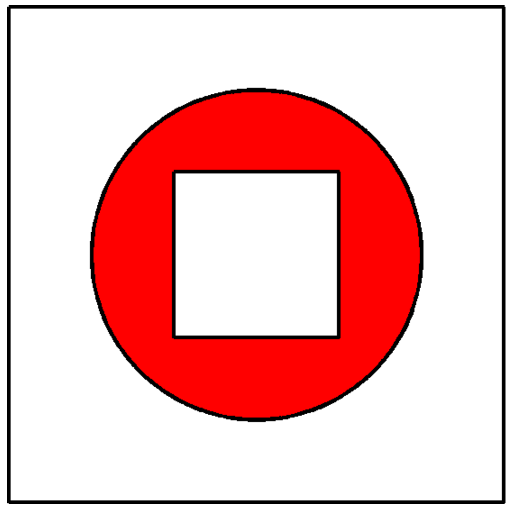
There is an associated notion: two metrics $d, \tilde{d}$ on E are **(strongly) equivalent** if $\exists A, B > 0$ such that

$$Ad(\mathbf{x},\mathbf{y}) \leq \tilde{d}(\mathbf{x},\mathbf{y}) \leq Bd(\mathbf{x},\mathbf{y}) \quad \forall \mathbf{x},\mathbf{y} \in E.$$

Intuitively, two metrics are equivalent if it is always possible to fit a $\tilde{d}-$ball between two $d-$balls, while maintaining the ratios of the balls' radii. Topological equivalence is not an equivalent notion, as we see in exercise 36.

Example: if $p, q \geq 1$, d_p and d_q are equivalent on $\mathbb{R}^n$.

For instance, to show that d_2 and d_∞ are equivalent in $\mathbb{R}^2$, it suffices to show that $\exists A, B > 0$ such that any $2-$ball of radius $R > 0$ contains an $\infty-$ball of radius $\frac{R}{A}$, and is contained in an $\infty-$ball of radius $\frac{R}{B}$.



Given the geometry of squares and circles, what values can A and B take? $\square$

There is also a similar notion for norms. Two norms $\|\cdot\|^*$, $\|\cdot\|^\circ$ on E are **equivalent** if $\exists a, b > 0$ such that

$$a\|\mathbf{x}\|^* \leq \|\mathbf{x}\|^\circ \leq b\|\mathbf{x}\|^*, \quad \forall \mathbf{x} \in E.$$

Clearly, two equivalent norms on E give rise to two equivalent metrics on E. But there is an important difference: over a **finite**$-$dimensional vector space, **any two norms are equivalent**, which we can show using the following proof outline:

1. without loss of generality, assume $\|\cdot\|^* = \|\cdot\|_1$;
2. only the vectors $\mathbf{x} \in S_1(\mathbf{0}, 1)$ need to be considered (why?);
3. show that $\|\cdot\|^\circ$ is continuous with respect to $\|\cdot\|_1$, and
4. use the max/min theorem over $S_1(\mathbf{0}, 1)$ to bound $a \leq \|\mathbf{x}\|^\circ \leq b$.

We end this section on preliminaries with two definitions that generalize the notion of a continuous function.

Let $f : (E, d) \to (\tilde{E}, \tilde{d})$. We say that f is

1. **uniformly continuous** if $\forall \varepsilon > 0$, $\exists \delta = \delta(\varepsilon) > 0$ such that $\forall \mathbf{x}, \mathbf{y} \in E$, $d(\mathbf{x}, \mathbf{y}) < \delta \implies \tilde{d}(f(\mathbf{x}), f(\mathbf{y})) < \varepsilon$;
2. **Lipschitz continuous** if $\exists K > 0$ such that $\tilde{d}(f(\mathbf{x}), f(\mathbf{y})) \leq K d(\mathbf{x}, \mathbf{y})$ $\forall \mathbf{x}, \mathbf{y} \in E$.

The conceptual difference between continuity and uniform continuity is that δ may depend on **x** and **y** as well as ε in the former case, but it can only depend on ε in the latter case.

Examples

1. Any polynomial $p : \mathbb{R} \to \mathbb{R}$ is uniformly continuous over a closed, bounded interval. □
2. Any uniformly continuous function is automatically continuous. □
3. Any Lipschitz continuous function is automatically uniformly continuous, hence continuous. □
4. The function $f : (0,1] \to \mathbb{R}$ defined by $f(x) = \frac{1}{x}$ is continuous but not uniformly continuous. □

This allows us to define another type of equivalence between metrics: two metrics $d, \tilde{d}$ on E are **uniformly equivalent** if $\text{id} : (E,d) \to (E,\tilde{d})$ is uniformly continuous, and so is its inverse.

Uniformly equivalent metrics are topologically equivalent, as uniform continuity also implies continuity, but there are topologically equivalent metrics that are not uniformly equivalent. However, uniform equivalence and strong equivalence of metrics are ... well, equivalent.

Lastly, note that uniform continuity, unlike continuity, is not a **topological notion**: given a function $f : E \to \tilde{E}$, the knowledge of the topologies on E and $\tilde{E}$, respectively, is sufficient to determine if f is continuous. But more must be known in order to determine if f is uniformly continuous. There is something fundamental at play here; we will return to it at a later stage.

8.2 Sequence in a Metric Space

Consider the sequence $(\mathbf{x}_n) \subseteq (E,d)$. The sequence **converges** to $\mathbf{x} \in (E,d)$, which we denote by $\mathbf{x}_n \to \mathbf{x}$, if

$$\forall \varepsilon > 0,\ \exists N \in \mathbb{N} \text{ such that } n > N \implies d(\mathbf{x}_n, \mathbf{x}) < \varepsilon.$$

In light of the notions presented in the previous section, this is equivalent to the following definition: $\mathbf{x}_n \to \mathbf{x} \in E$ if

$$\forall V \in \mathcal{V}(\mathbf{x}),\ \exists N \in \mathbb{N} \text{ such that } n > N \implies \mathbf{x}_n \in V.$$

Thus a sequence converges to **x** if any neighbourhood of **x** contains infinitely many terms in the sequence.

A **subsequence of** $(\mathbf{x}_n)$ is a sequence $(\mathbf{y}_n)$ such that $\mathbf{y}_n = \mathbf{x}_{\varphi(n)}$ for some strictly increasing function $\varphi : \mathbb{N} \to \mathbb{N}$. It is easy to show that if $\mathbf{x}_n \to \mathbf{x}$, then any subsequence of $(\mathbf{x}_n)$ also converges to **x** (see the exercises).

Let $(\mathbf{x}_n)$ be a sequence in a metric space (E,d). We say that $\mathbf{a} \in E$ is a **limit point of** $(\mathbf{x}_n)$ if $\forall \varepsilon > 0, \forall \rho \in \mathbb{N}, \exists n \geq \rho$ such that $d(\mathbf{x}_n, \mathbf{a}) < \varepsilon$.[5]

Proposition 103
Let $(\mathbf{x}_n) \subseteq (E,d)$, $\mathbf{a} \in E$. The following are equivalent:

1. $\mathbf{a}$ *is a limit point of* $(\mathbf{x}_n)$*;*
2. *there is a subsequence of* $(\mathbf{x}_n)$ *which converges to* $\mathbf{a}$*;*
3. $\forall \rho \in \mathbb{N}$*, we have* $\mathbf{a} \in \overline{A_\rho}$*, where* $A_\rho = \{\mathbf{x}_n | n \geq \rho\}$*, and*
4. *either* $\mathbf{a}$ *is a cluster point of* A_1 *or* $\{\mathbf{x}_n \mid \mathbf{x}_n = \mathbf{a}\}$ *is infinite (in the latter case, we say that* $\mathbf{a}$ *is a* ***replicating point of*** $(\mathbf{x}_n)$*).*

Proof: we prove $1. \implies 2. \implies 3. \implies 4. \implies 1.$

$1. \implies 2.$: Set $\varepsilon_n = \frac{1}{n}$. Since $\mathbf{a}$ is a limit point of the sequence $(\mathbf{x}_n)$, there is a smallest integer n for which $d(\mathbf{y}_n, \mathbf{a}) < \frac{1}{n}$, where $\mathbf{y}_n$ is a member of the sequence $(\mathbf{x}_m)_{m \geq n}$. By construction, $(\mathbf{y}_n)$ is a subsequence of $(\mathbf{x}_n)$ and $\mathbf{y}_n \to \mathbf{a}$.

$2. \implies 3.$: If there is a subsequence $(\mathbf{y}_n) \subseteq (\mathbf{x}_n)$ which converges to $\mathbf{a}$, then $\forall \varepsilon > 0, \forall \rho \in \mathbb{N}, \exists N \in \mathbb{N}$ such that $\mathbf{y}_n \in A_\rho \cap B(\mathbf{a}, \varepsilon)$ whenever $n > N$. But according to Proposition 94, $\mathbf{a} \in \overline{A_\rho}$ if and only if $\forall \varepsilon > 0,\ A_\rho \cap B(\mathbf{a}, \varepsilon) \neq \varnothing$. Consequently, $\forall \rho \in \mathbb{N}$, $\mathbf{a} \in \overline{A_\rho}$.

$3. \implies 4.$: If $\forall \rho \in \mathbb{N}$, $\mathbf{a} \in \overline{A_\rho}$, then $\forall \rho \in \mathbb{N}$, $\forall \varepsilon > 0$, $\exists$ a smallest $n_\rho \geq \rho$ such that $d(\mathbf{x}_{n_\rho}, a) < \varepsilon$. As such, $\mathbf{x}_{n_\rho}$ is a subsequence of $(\mathbf{x}_n)$ and

$$\lim_{\rho \to \infty} \mathbf{x}_{n_\rho} = \mathbf{a}.$$

- If $(\mathbf{x}_n)$ converges, it must do so to $\mathbf{a}$, according to exercise 40. Consequently, $\forall \eta > 0$, $A_1 \cap B(\mathbf{a}, \varepsilon)$ is infinite and so must contain at least one point distinct from $\mathbf{a}$. Consequently, $\mathbf{a}$ is a cluster point of A_1.
- If $(\mathbf{x}_n)$ diverges and $\mathbf{a}$ is not a replicating point of $(\mathbf{x}_n)$, then $\mathbf{x}_{n_\rho} \not\to \mathbf{a}$ (why?), which is a contradiction. Consequently, if $(\mathbf{x}_n)$ diverges then $\mathbf{a}$ is a replicating point of $(\mathbf{x}_n)$.

$4. \implies 1.$: Left as an exercise. ■

8.2.1 Closure, Closed Subsets, and Continuity

We can conclude from Proposition 103 that the set $\bigcap_{\rho \in \mathbb{N}} \overline{A_\rho}$ of limit points of $(\mathbf{x}_n)$ is closed and that if $\mathbf{x}_n \to \mathbf{x}$, then $\mathbf{x}$ is the unique limit point of $(\mathbf{x}_n)$.

[5] Compare with the notion of a cluster point.

There is a nice way to characterize closure, closed subsets and continuity using sequences and convergence, provided by the next three results.

Proposition 104
Let (E,d) be a metric space, $A \subseteq E$ and $\mathbf{x} \in E$. Then,

$$\mathbf{x} \in \overline{A} \Longleftrightarrow \exists (\mathbf{x}_n) \subseteq A \text{ such that } \mathbf{x}_n \to \mathbf{x}.$$

Proof: the direction $\Longleftarrow$ is a clear consequence of the remark at the start of this subsection. For $\Longrightarrow$, consider the following argument. Let $n \in \mathbb{N}$. Since $\mathbf{x} \in \overline{A}$, $\exists \mathbf{x}_n (\neq \mathbf{x}) \in A$ with $d(\mathbf{x}_n, \mathbf{x}) < \frac{1}{n}$. Clearly, $\mathbf{x}_n \to \mathbf{x}$. ■

Proposition 105
Let (E,d) be a metric space, with $F \subseteq E$. Then, $F \subseteq_C E$ if and only if any sequence $(\mathbf{x}_n) \subseteq F$ which converges in E converges to a point in F.

Proof: if $F \subseteq_C E$, then $\overline{F} = F$. Assume that $\mathbf{x}_n \in F$ and $\mathbf{x}_n \to \mathbf{x}$. We must show that $\mathbf{x} \in F = \overline{F}$. If (x_n) is eventually constant, then $\mathbf{x}_n = \mathbf{x} \in F$ for all n greater than some index. Otherwise $\forall \varepsilon > 0$, $B(\mathbf{x}, \varepsilon) \cap F$ contains an infinite subset of $\{\mathbf{x}_n \mid n \geq 1\}$; consequently, $\mathbf{x} \in \overline{F}$.

Conversely, let $\mathbf{x} \in \overline{F}$. According to Proposition 104, there is a subsequence $(\mathbf{x}_n) \subseteq F$ such that $\mathbf{x}_n \to \mathbf{x}$. By hypothesis, any such sequence must converge *in* F. Hence, $\mathbf{x} \in F$. Consequently, $F = \overline{F}$ and $F \subseteq_C E$. ■

Proposition 106
Let $(E,d), (\tilde{E}, \tilde{d})$ be a metric spaces. Then $f : E \to \tilde{E}$ is continuous if and only $f(\mathbf{x}_n) \to f(\mathbf{x})$ whenever $\mathbf{x}_n \to \mathbf{x}$.

Proof: the direction $\Longleftarrow$ is a clear consequence of the definition of a continuous function.

Conversely, let $F \subseteq_C \tilde{E}$. We want to show that $f^{-1}(F) \subseteq_C E$. Let $(\mathbf{x}_n) \subseteq f^{-1}(F)$ with $\mathbf{x}_n \to \mathbf{x}$. By hypothesis, $f(\mathbf{x}_n) \to f(\mathbf{x})$. But $F \subseteq_C \tilde{E}$ so that $f(\mathbf{x}) \in F$, according to Proposition 105.

Consequently, $\mathbf{x} \in f^{-1}(F)$. According to Proposition 105, we must then have $f^{-1}(F) \subseteq_F E$; in other words, f is continuous. ■

We will see in Part IV that these characterizations do not always apply to general (as in, non-metric) topological spaces.

8.2.2 Complete Spaces and Cauchy Sequences

The sequence $(\mathbf{x}_n) \subseteq (E, d)$ is a **Cauchy sequence** if

$$\forall \varepsilon > 0, \exists N \in \mathbb{N} \text{ such that } n, m > N \implies d(\mathbf{x}_n, \mathbf{x}_m) < \varepsilon.$$

Some properties of Cauchy sequences in $\mathbb{R}$ carry over to metric spaces.

> **Proposition 107**
> *Convergent sequences in (E, d) are Cauchy.*
>
> **Proof:** let $\mathbf{x}_n \to \mathbf{x}$ and $\varepsilon > 0$; thus $\exists N \in \mathbb{N}$ such that $d(\mathbf{x}_n, \mathbf{x}) < \frac{\varepsilon}{2}$ whenever $n > N$. Now, let $m > N$. According to the triangle inequality,
>
> $$d(\mathbf{x}_n, \mathbf{x}_m) \leq d(\mathbf{x}_n, \mathbf{x}) + d(\mathbf{x}, \mathbf{x}_m) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$
>
> Consequently, $(\mathbf{x}_n)$ is a Cauchy sequence. ■

In a normed space $(E, \|\cdot\|)$, a sequence $(\mathbf{x}_n)$ is **bounded** if $\exists M \in \mathbb{N}$ such that $\|\mathbf{x}_n\| < M$ for all $n \in \mathbb{N}$.

But a metric space (E, d) is not necessarily a normed vector space, so there might not be a norm available to determine boundedness.

In a general metric space (E, d), a sequence $(\mathbf{x}_n)$ is **bounded** if $\exists M > 0$ s.t. $\mathbf{x}_n \in B(\mathbf{0}, M)$ for all $n \in \mathbb{N}$. Similarly, $A \subseteq E$ is **bounded** if $\delta(A) < \infty$ (using the definition from p. 192).

> **Proposition 108**
> *Every Cauchy sequence in (E, d) is bounded.*
>
> **Proof:** let $(\mathbf{x}_n)$ be a Cauchy sequence. If $1 > \varepsilon > 0$, then $\exists N \in \mathbb{N}$ such that $d(\mathbf{x}_n, \mathbf{x}_m) < \varepsilon$ whenever $n, m > N$. Now, let
>
> $$M = \max\{d(\mathbf{0}, \mathbf{x}_1), d(\mathbf{0}, \mathbf{x}_2), \ldots, d(\mathbf{0}, \mathbf{x}_N), d(\mathbf{0}, \mathbf{x}_{N+1})\} + 2.$$
>
> Then, for any $n > N$, the triangle inequality yields
>
> $$d(\mathbf{0}, \mathbf{x}_n) \leq d(\mathbf{0}, \mathbf{x}_{N+1}) + d(\mathbf{x}_{N+1}, \mathbf{x}_n) \leq M - 2 + 1,$$
>
> *i.e.* for any $n > N$, $\mathbf{x}_n \in B(\mathbf{0}, M)$. Since $\mathbf{x}_n \in B(\mathbf{0}, M - 2)$ for all $1 \leq n \leq N$, then $\mathbf{x}_n \in B(\mathbf{0}, M)$ for all $n \in \mathbb{N}$. ■

Interestingly, given its link to convergence in the case of complete spaces, the notion of a Cauchy sequence is not topological.

Example: let $A = (0, \infty)$. Consider the following metrics on A:

$$d_1(x, y) = |x - y| \quad \text{and} \quad d_2(x, y) = |\ln x - \ln y|.$$

Show that both metrics induce the same topology on A, but that Cauchy sequences under one are not necessarily Cauchy sequences under the other.

Proof: the mapping id $: (A, d_1) \to (A, d_2)$ is homeomorphic. Indeed, for $x, z \in A$ and $\varepsilon, \eta > 0$, we have

$$B_{d_1}(x, \varepsilon) = \{y \in A \mid |x - y| < \varepsilon\} = (x - \varepsilon, x + \varepsilon) \cap A,$$

and

$$B_{d_2}(z, \eta) = \{y \in A \mid |\ln z - \ln y| < \eta\} = \{y \in A \mid e^{-\eta} < \frac{y}{z} < e^{\eta}\} = (ze^{-\eta}, ze^{\eta}).$$

It is left as an exercise to show that

$$B_{d_1}(z, \tfrac{1}{2}z(1 - e^{-\eta})) \subseteq B_{d_2}(z, \eta) \quad \text{and} \quad B_{d_2}(x, \ln(\tfrac{2x+\varepsilon}{2x})) \subseteq B_{d_1}(x, \varepsilon)$$

for all $x, z \in A$, $\varepsilon, \eta > 0$. Thus $W \subseteq_O (A, d_1) \iff W \subseteq_O (A, d_2)$. We already know that the sequence $(\frac{1}{n})$ is Cauchy in (A, d_1). But if $m = 2n$, then

$$d_2(\tfrac{1}{m}, \tfrac{1}{n}) = \left|\ln \tfrac{1}{m} - \ln \tfrac{1}{n}\right| = \left|\ln \tfrac{n}{m}\right| = \left|\ln \tfrac{n}{2n}\right| = \ln 2 \geq 1/2$$

for every $n \in \mathbb{N}$, and so $(\frac{1}{n})$ is not a Cauchy sequence in (A, d_2). ∎

This could not happen, however, if the metrics are strongly equivalent, which further illustrates the distinctness of the notions of strong equivalence and topological equivalence.

Proposition 109
Let d and $\tilde{d}$ be two equivalent metrics on E. Then, $(\mathbf{x}_n)$ is a Cauchy sequence in (E, d) if and only if $(\mathbf{x}_n)$ is a Cauchy sequence in $(E, \tilde{d})$.

Proof: since d and $\tilde{d}$ are equivalent, $\exists a, b > 0$ such that

$$ad(\mathbf{x}, \mathbf{y}) \leq \tilde{d}(\mathbf{x}, \mathbf{y}) \leq bd(\mathbf{x}, \mathbf{y}) \quad \forall \mathbf{x}, \mathbf{y} \in E.$$

If $(\mathbf{x}_n)$ is a Cauchy sequence in $(E, \tilde{d})$, then, $\forall \varepsilon > 0$, $\exists N \in \mathbb{N}$ such that $m, n > N \implies \tilde{d}(\mathbf{x}_n, \mathbf{x}_m) < \varepsilon$. Thus, it is the case that

$$ad(\mathbf{x}_n, \mathbf{x}_m) \leq \tilde{d}(\mathbf{x}_n, \mathbf{x}_m) < \varepsilon \quad \forall m, n > N \implies d(\mathbf{x}_n, \mathbf{x}_m) < \frac{\varepsilon}{a} \quad \forall m, n > N.$$

Consequently, $(\mathbf{x}_n)$ is also a Cauchy sequence in (E, d). By symmetry, the reverse implication is clearly true. ∎

A metric space (E, d) is **complete** if every single one of its Cauchy sequences is convergent. If a complete metric space is also a normed vector space, then it is a **Banach space**. If a Banach space is also an inner product space, then it is a **Hilbert space**.

Examples (COMPLETE, BANACH, AND HILBERT SPACES)

1. We have already seen that $(\mathbb{R}, d_2)$ is a complete space. Since it is a normed space, it is also a Banach space. The inner product $(x \mid y) = xy$ makes it a Hilbert space.
2. The same applies to $(\mathbb{K}^n, d_2)$, with the inner product $(\mathbf{x} \mid \mathbf{y}) = \sum x_i \overline{y_i}$.
3. The space $\mathcal{C} = (C_{\mathbb{K}}([0,1]), \|\cdot\|_\infty)$ is a Hilbert space with the inner product
$$(f \mid g) = \int_{[0,1]} f\overline{g}\, dm, \quad f \sim g \Longleftrightarrow f = g \text{ a.e.}$$
4. It is a bit less obvious that the space
$$\ell^2(\mathbb{N}) = \{\mathcal{X} \mid \mathcal{X} = (x_n)_{n\in\mathbb{N}};\ x_n \in \mathbb{C},\ \sum |x_n|^2 < \infty\}$$
is a Hilbert space, together with
$$(\mathcal{X} \mid \mathcal{Y}) = \sum x_n \overline{y_n} \quad \text{and} \quad \|\mathcal{X}\|_2 = (\mathcal{X} \mid \mathcal{X})^{1/2} = \left(\sum |x_n|^2\right)^{1/2},$$
but it is a classical result (see Chapter 27). □

Closed subsets of complete spaces are especially well-behaved, as we see in the next two results.

Proposition 110
Every closed subset of a complete metric space is complete.

Proof: let $A \subseteq_C E$ and $(\mathbf{x}_n) \subseteq A$ be a Cauchy sequence. Since E is complete, $\mathbf{x}_n \to \mathbf{x}$ converges in E. But A is closed, so $\mathbf{x} \in A$, according to Proposition 105. ■

Proposition 111
Every complete subspace of a metric space is closed.

Proof: let $A \subseteq (E, d)$ be complete. Let $\mathbf{x} \in \overline{A}$. According to Proposition 104, $\exists (\mathbf{x}_n) \subseteq A$ such that $\mathbf{x}_n \to \mathbf{x}$. Therefore, $(\mathbf{x}_n)$ is a convergent sequence in E. In particular, it is a Cauchy sequence of points in A, according to Proposition 107. But A is complete so that $\mathbf{x} \in A$. Hence $\overline{A} \subseteq A$ and so $\overline{A} = A$, which means that $A \subseteq_C E$. ■

The **product** of two metric spaces (E', d') and (E^*, d^*) is the metric space

$$(E, d) = (E' \times E^*, \sup\{d', d^*\});$$

it is easy to see how this definition can be extended to a product of n metric spaces. At any rate, the product of metric spaces is also a metric space.[6]

Proposition 112
Let (E_i, d_i) be metric spaces for $i = 1, \ldots, n$. The product metric space $(E, d) = (E_1 \times \cdots \times E_n, \sup_{i=1,\ldots,n}\{d_i\})$ is complete if and only if (E_i, d_i) for all $i = 1, \ldots, n$.

Proof: left as an exercise. ∎

The following result is a generalization of the nested intervals theorem of Chapter 1.

Proposition 113
Let (E, d) be a complete metric space. If (F_n) is a decreasing sequence of non-empty closed subsets of E

$$E \supseteq F_1 \supseteq F_2 \supseteq \cdots \supseteq F_n \supseteq \cdots$$

such that $\lim_{n\to\infty} \delta(F_n) = 0$, then $\bigcap_{n\geq 1} F_n = \{\mathbf{x}\}$ for some $\mathbf{x} \in E$.

Proof: let $\Gamma = \bigcap F_n$. For each $n \in \mathbb{N}$, pick $\mathbf{x}_n \in F_n$.

Let $\varepsilon > 0$. Since $\delta(F_n) \to 0$, $\exists N_\varepsilon \in \mathbb{N}$ such that

$$n > N_\varepsilon \Longrightarrow \delta(F_n) < \sup\{d(\mathbf{w}, \mathbf{z}) \mid \mathbf{w}, \mathbf{z} \in F_n\} < \tfrac{\varepsilon}{2}.$$

Let $m > n > N_\varepsilon$ and pick $\mathbf{y} \in F_m \subseteq F_n$. Then

$$m > n > N_\varepsilon \Longrightarrow d(\mathbf{x}_n, \mathbf{x}_m) \leq d(\mathbf{x}_n, y) + d(\mathbf{y}, \mathbf{x}_m) < \tfrac{\varepsilon}{2} + \tfrac{\varepsilon}{2} = \varepsilon.$$

As $(\mathbf{x}_n) \subseteq E$ is Cauchy and E is complete, $\exists \mathbf{x} \in E$ such that $\mathbf{x}_n \to \mathbf{x}$. For all $p \geq 1$, $(\mathbf{x}_n)_{n\geq p} \subseteq F_p$. As $F_p \subseteq_C E$, $(\mathbf{x}_n)_{n\geq p}$ converges in F_p, according to Proposition 105. Hence $\mathbf{x} \in F_p$ for all $p \geq 1$. Consequently, $\mathbf{x} \in \Gamma$.

But if $\mathbf{y} \in \Gamma$, then $\mathbf{y} \in F_n$ for all n, so that $0 \leq d(\mathbf{x}, \mathbf{y}) \leq \delta(F_n) \to 0$ for all n. Thus $d(\mathbf{x}, \mathbf{y}) = 0$, so that $\mathbf{y} = \mathbf{x}$ and $\Gamma = \{\mathbf{x}\}$. ∎

If $r \in (0, 1)$, for instance, and if we have $F_n = \overline{B(\mathbf{0}, r^n)} \subseteq (\mathbf{R}^m, d_2)$ for some $m \geq 1$, then $\bigcap F_n = \{\mathbf{0}\}$.

[6] In fact, the definition can be generalized to arbitrary collections $\{E_\alpha\}_{\alpha\in J}$, but we will see in Part IV that there are complications.

The following contraction result is representative of a family of extremely useful theorems.

Theorem 114 (Fixed Point Theorem)
Let (E, d) be a a complete metric space and let $f : E \to E$ be a ***contraction*** *on E, that is,*

$$\exists k \in (0,1) \text{ such that } d(f(\mathbf{x}), f(\mathbf{y})) \leq kd(\mathbf{x}, \mathbf{y}) \text{ for all } \mathbf{x}, \mathbf{y} \in E.$$

Then $\exists! \mathbf{x}^* \in E$ such that $f(\mathbf{x}^*) = \mathbf{x}^*$; $\mathbf{x}^*$ is a ***fixed point of*** *f.*

Proof: let $\mathbf{x}_0 \in E$. If $f(\mathbf{x}_0) = \mathbf{x}_0$, we are done. Otherwise, consider the sequence $(f^n(\mathbf{x}_0))_n$, where f^n represents n successive compositions of f:

$$\begin{aligned} d(f^n(\mathbf{x}_0), f^{n+1}(\mathbf{x}_0)) &= d(f(f^{n-1}(\mathbf{x}_0)), f(f^n(\mathbf{x}_0))) \leq kd(f^{n-1}(\mathbf{x}_0), f^n(\mathbf{x}_0)) \\ &= kd(f(f^{n-2})(\mathbf{x}_0), f(f^{n-1})(\mathbf{x}_0)) \leq \cdots \leq k^n d(\mathbf{x}_0, f(\mathbf{x}_0)). \end{aligned}$$

Then, for any $m > n$,

$$\begin{aligned} d(f^m(\mathbf{x}_0), f^n(\mathbf{x}_0)) &\leq d(f^m(\mathbf{x}_0), f^{m-1}(\mathbf{x}_0)) + \cdots + d(f^{n+1}(\mathbf{x}_0), f^n(\mathbf{x}_0)) \\ &\leq (k^n + \cdots + k^{m-1}) d(\mathbf{x}_0, f(\mathbf{x}_0)) \leq \frac{k^n}{1-k} d(\mathbf{x}_0, f(\mathbf{x}_0)) \end{aligned}$$

For any ε, let $M_\varepsilon = \left\lceil \ln\left(\frac{\varepsilon}{d(\mathbf{x}_0, f(\mathbf{x}_0))}(1-k)\right) - \ln k \right\rceil$. Then, whenever $m > n > M_\varepsilon$, we have

$$d(f^m(\mathbf{x}_0), f^n(\mathbf{x}_0)) \leq \frac{k^n}{1-k} d(\mathbf{x}_0, f(\mathbf{x}_0)) \leq \frac{k^{M_\varepsilon}}{1-k} d(\mathbf{x}_0, f(\mathbf{x}_0)) < \varepsilon.$$

Consequently, $(f^n(\mathbf{x}_0))$ is a Cauchy sequence in E. But E is complete so that $f^n(\mathbf{x}_0) \to \mathbf{x}$ for some $\mathbf{x} \in E$.

By definition, contraction mappings are Lipschitz continuous, and thus also continuous, and so

$$f(\mathbf{x}) = f\big(\lim_{n\to\infty} f^n(\mathbf{x}_0)\big) = \lim_{n\to\infty} f(f^n(\mathbf{x}_0)) = \lim_{n\to\infty} f^{n+1}(\mathbf{x}_0) = \mathbf{x}.$$

Now, suppose that $\mathbf{x}$ and $\mathbf{y}$ are two fixed points of f. Then,

$$d(\mathbf{x}, \mathbf{y}) = d(f(\mathbf{x}), f(\mathbf{y})) \leq kd(\mathbf{x}, \mathbf{y}).$$

Since $k < 1$, the only way for the inequality to be valid is if $d(\mathbf{x}, \mathbf{y}) = 0$, which implies that $\mathbf{x} = \mathbf{y}$. The fixed point of f is thus unique. Call it $\mathbf{x}^*$ to match with the statement of the theorem. ■

The choice of $\mathbf{x}_0 \in E$ in the proof of Theorem 114 is arbitrary; if f is a contraction, the sequence $(f^n(\mathbf{x}))$ converges to the unique fixed point $\mathbf{x}^*$ for all $\mathbf{x} \in E$. Note that the restriction $k \in (0,1)$ is necessary, as the following example demonstrates.

> **Example:** let $f : \mathbb{R} \to \mathbb{R}$ be defined by
>
> $$f(x) = \begin{cases} 1, & x < 0 \\ x + \frac{1}{x+1}, & x \geq 0 \end{cases}.$$
>
> It is not hard to see that f has no fixed point (see exercise 45), yet
>
> $$d(f(x), f(y)) \leq d(x,y) \quad \text{for all } x, y \in \mathbb{R}. \qquad \square$$

8.3 Solved Problems

1. Let A, B be subsets of a metric space (E, d). Show that

   a) $B \subseteq A \Longrightarrow \text{int}(B) \subseteq \text{int}(A)$
   b) $B \subseteq A \Longrightarrow \overline{B} \subseteq \overline{A}$
   c) $\text{int}(A \cap B) = \text{int}(A) \cap \text{int}(B)$
   d) $\overline{A \cup B} = \overline{A} \cup \overline{B}$
   e) $\text{int}(A) \cup \text{int}(B) \subseteq \text{int}(A \cup B)$
   f) $\overline{A \cap B} \subseteq \overline{A} \cap \overline{B}$

   Proof:

   a) By definition, $\text{int}(B) \subseteq B \subseteq A$, i.e. $\text{int}(B)$ is an open set contained in A. Consequently, $\text{int}(B)$ is contained in the largest open set contained in A, namely $\text{int}(A)$.
   b) By definition, $B \subseteq A \subseteq \overline{A}$, i.e. $\overline{A}$ is a closed set containing B. Consequently, $\overline{A}$ contains the smallest closed set containing B, i.e. $\overline{B}$.
   c) Since $\text{int}(A) \cap \text{int}(B) \subseteq_O E$ and since $\text{int}(A) \subseteq A$ and $\text{int}(B) \subseteq B$, we must have $\text{int}(A) \cap \text{int}(B) \subseteq A \cap B$. As such, $\text{int}(A) \cap \text{int}(B)$ must be contained in the largest open set contained in $A \cap B$, so that $\text{int}(A) \cap \text{int}(B) \subseteq \text{int}(A \cap B)$. On the other hand, since $A \cap B \subseteq A, B$, then we must have $\text{int}(A \cap B) \subseteq \text{int}(A), \text{int}(B)$ and so
   $$\text{int}(A \cap B) \subseteq \text{int}(A) \cap \text{int}(B).$$
   d) Basically the same proof with $\cap \leftrightsquigarrow \cup$, $\subseteq \leftrightsquigarrow \supseteq$, $\text{int}(\cdot) \leftrightsquigarrow \overline{(\cdot)}$.
   e) Since $A, B \subseteq A \cup B$, then $\text{int}(A), \text{int}(B) \subseteq \text{int}(A \cup B)$. Hence $\text{int}(A) \cup \text{int}(B) \subseteq \text{int}(A \cup B)$.
   f) Basically the same proof with $\cap \leftrightsquigarrow \cup$, $\subseteq \leftrightsquigarrow \supseteq$, $\text{int}(\cdot) \leftrightsquigarrow \overline{(\cdot)}$. $\blacksquare$

2. In each instance, give an example showing that, in general,

   a) $\text{int}(A) \cup \text{int}(B) \neq \text{int}(A \cup B)$
   b) $\overline{A \cap B} \neq \overline{A} \cap \overline{B}$

   Solution:

   a) Let $E = \mathbb{R}$ with the Euclidean metric, and let $A = [a, b]$ and $B = [b, c]$ with $c > b > a$, for instance. Then

   $$\begin{aligned} \text{int}(A) &= (a, b), \quad \text{int}(B) = (b, c), \quad A \cup B = [a, c], \\ \text{int}(A \cup B) &= (a, c), \quad \text{int}(A) \cup \text{int}(B) = (a, b) \cup (b, c) = (a, c) \setminus \{b\}. \end{aligned}$$

   b) Let $E = \mathbb{R}$ with the Euclidean metric, and $A = (a, b)$ and $B = (b, c)$ with $c > b > a$, for instance. Then

   $$\overline{A} = [a, b], \quad \overline{B} = [b, c], \quad A \cap B = \varnothing, \quad \overline{A \cap B} = \varnothing, \quad \overline{A} \cap \overline{B} = \{b\}. \qquad \square$$

3. Let A be subset of a metric space (E, d). Show that

   a) $E \setminus \text{int}(A) = \overline{E \setminus A}$
   b) $E \setminus \overline{A} = \text{int}(E \setminus A)$
   c) $\partial(\text{int}(A)) \subseteq \partial A$
   d) $\partial \overline{A} \subseteq \partial A$

   Proof:

   a) We have

   $$\begin{aligned} \text{int}(A) &\subseteq A, \quad \text{by definition} \\ E \setminus A &\subseteq E \setminus \text{int}(A), \quad \text{again by definition} \\ \overline{E \setminus A} &\subseteq \overline{E \setminus \text{int}(A)} = E \setminus \text{int}(A), \quad \text{as } E \setminus \text{int}(A) \subseteq_C E \end{aligned}$$

   On the other hand, we have

   $$\begin{aligned} E \setminus A &\subseteq \overline{E \setminus A}, \quad \text{by definition} \\ E \setminus \overline{E \setminus A} &\subseteq E \setminus (E \setminus A) = A, \quad \text{again by definition} \\ E \setminus \overline{E \setminus A} &= \text{int}(E \setminus \overline{E \setminus A}) \subseteq \text{int}(A) = E \setminus \text{int}(A), \quad \text{as } E \setminus \overline{E \setminus A} \subseteq_O E \\ E \setminus \text{int}(A) &\subseteq \overline{E \setminus A} \end{aligned}$$

   b) We have

   $$\begin{aligned} A &\subseteq \overline{A}, \quad \text{by definition} \\ E \setminus \overline{A} &\subseteq E \setminus A, \quad \text{again by definition} \\ E \setminus \overline{A} &= \text{int}(E \setminus \overline{A}) \subseteq \text{int}(E \setminus A), \quad \text{as } E \setminus \overline{A} \subseteq_O E \end{aligned}$$

On the other hand, we have

$$\begin{aligned}
\text{int}(E \setminus A) &\subseteq E \setminus A, \quad \text{by definition} \\
A = E \setminus (E \setminus A) &\subseteq E \setminus \text{int}(E \setminus A), \quad \text{again by definition} \\
\overline{A} \subseteq \overline{E \setminus \text{int}(E \setminus A)} &= E \setminus \text{int}(E \setminus A) \quad \text{as } E \setminus \text{int}(E \setminus A) \subseteq_C E \\
\text{int}(E \setminus A) &\subseteq E \setminus \overline{A}
\end{aligned}$$

c) Since $\text{int}(A) \subseteq A$, we have $\overline{\text{int}(A)} \subseteq \overline{A}$ and so

$$\partial\, \text{int}(A) = \overline{\text{int}(A)} \setminus \text{int}(A) \subseteq \overline{A} \setminus \text{int}(A) = \partial A.$$

d) Basically the same idea, as above, but with $X \setminus \text{int}(\overline{A}) \subseteq X \setminus \text{int}(A)$. ■

4. Find an example of a subset A of a metric space (E, d) for which $\partial(\text{int}(A))$, ∂A and $\partial \overline{A}$ are all different.

Solution: let $E = \mathbb{R}$ with the Euclidean metric, and let $A = \mathbb{Q} \cup (0,1)$, for instance. Then

$$\begin{aligned}
\overline{A} &= \overline{\mathbb{Q} \cup (0,1)} = \overline{\mathbb{Q}} \cup \overline{(0,1)} = \mathbb{R} \\
\text{int}(A) &= \{x \in \mathbb{R} \mid \exists r > 0 \text{ s.t. } B(x,r) \subseteq A\} = (0,1) \\
\partial(\text{int}(A)) &= \overline{\text{int}(A)} \setminus \text{int}(A) = \overline{(0,1)} \setminus (0,1) = [0,1] \setminus (0,1) = \{0,1\} \\
\partial A &= \overline{A} \setminus \text{int}(A) = \mathbb{R} \setminus (0,1) \\
\partial \overline{A} &= \overline{A} \setminus \text{int}(\overline{A}) = \mathbb{R} \setminus \text{int}(\mathbb{R}) = \mathbb{R} \setminus \mathbb{R} = \varnothing
\end{aligned}$$

which are all distinct. □

5. Find two subsets $A, B \subseteq (R, d_2)$ for which $A \cup B$, $\text{int}(A) \cup B$, $A \cup \text{int}(B)$, $\text{int}(A) \cup \text{int}(B)$, and $\text{int}(A \cup B)$ are all distinct.

Solution: let $E = \mathbb{R}$ with the Euclidean metric, and let

$$A = [\sqrt{2}, \varphi) \cup (\varphi, e) \cup \{\pi\} \cup (\mathbb{Q} \cap (8,9)), \quad \varphi = \frac{1+\sqrt{5}}{2}$$

for instance. Then

$$\begin{aligned}
\text{int}(A) &= (\sqrt{2}, \varphi) \cup (\varphi, e), \quad \overline{A} = [\sqrt{2}, e] \cup \{\pi\} \cup [8,9] \\
\text{int}(\overline{A}) &= (\sqrt{2}, e) \cup (8,9) \\
\overline{\text{int}(A)} &= [\sqrt{2}, e] \\
\overline{\text{int}(\overline{A})} &= [\sqrt{2}, e] \cup [8,9] \\
\text{int}\left(\overline{\text{int}(A)}\right) &= (\sqrt{2}, e)
\end{aligned}$$

are all distinct. □

6. Find a subset $A \subseteq (R, d_2)$ for which A, $\text{int}(A)$, $\overline{A}$, $\text{int}(\overline{A})$, $\overline{\text{int}(A)}$, $\overline{\text{int}(\overline{A})}$ and $\text{int}\left(\overline{\text{int}(A)}\right)$ are all distinct.

Solution: let $E = \mathbb{R}$ with the Euclidean metric, and let $A = [\sqrt{2}, e]$ and $B = [e, \pi]$, for instance. Then

$$\begin{aligned} A \cup B &= [\sqrt{2}, \pi] \\ \text{int}(A) \cup B &= (\sqrt{2}, \pi] \\ A \cup \text{int}(B) &= [\sqrt{2}, \pi) \\ \text{int}(A) \cup \text{int}(B) &= (\sqrt{2}, \pi) \setminus \{e\} \\ \text{int}(A \cup B) &= (\sqrt{2}, \pi) \end{aligned}$$

which are all distinct. □

7. For any subset $A \subseteq (R, d_2)$, show that $\text{int}\left(\overline{\text{int}(\overline{A})}\right) = \text{int}(\overline{A})$.

Proof: By definition,

$$\text{int}(\overline{A}) \subseteq \overline{A} \Longrightarrow \overline{\text{int}(\overline{A})} \subseteq \overline{\overline{A}} = \overline{A} \Longrightarrow \text{int}\left(\overline{\text{int}(\overline{A})}\right) \subseteq \text{int}(\overline{A}).$$

On the other hand, whenever B is open we have

$$B \subseteq \overline{B} \Longrightarrow B = \text{int}\, B \subseteq \text{int}(\overline{B}).$$

Set $B = \text{int}(\overline{A})$. Then B is open and

$$\text{int}(\overline{A}) \subseteq \text{int}(\overline{B}) = \text{int}\left(\overline{\text{int}(\overline{A})}\right),$$

which completes the proof. ■

(Could we replace $(\mathbb{R}, d_2)$ by any metric space? Any topological space?)

8. We say that $A \subseteq E$ is **meagre** (or nowhere dense) if and only if $\text{int}(\overline{A}) = \varnothing$. Show that
   a) A is meagre if and only if $\text{int}(E \setminus A)$ is dense in E (A is **dense** in B if $A \subseteq B \subseteq \overline{A}$);
   b) A is meagre if and only if A is contained in a closed subset of E whose interior is empty;
   c) A is closed and meagre if and only if $A = \partial A$, and
   d) A is meagre $\Longrightarrow \overline{A} = \partial A$.

Proof:

a) $\boxed{\Longrightarrow}$ If $\text{int}(\overline{A}) = \varnothing$, then

$$E = E \setminus \varnothing = E \setminus \text{int}(\overline{A}) = \overline{E \setminus \overline{A}} = \overline{\text{int}\, E \setminus A}.$$

Hence $\text{int}(E \setminus A)$ is dense in E.

$\boxed{\Longleftarrow}$ It's pretty much the same thing: if $\overline{\text{int}(E \setminus A)} = E$, then

$$E = \overline{\text{int}\, E \setminus A} = \overline{E \setminus \overline{A}} = E \setminus \text{int}(\overline{A}).$$

Hence $\text{int}(\overline{A}) = \varnothing$.

b) $\boxed{\Longrightarrow}$ If int$(\overline{A}) = \varnothing$, then $\overline{A}$ does not have interior points. Since $\overline{A} \subseteq_C E$ and since $A \subseteq \overline{A}$, then A is contained in a closed set whose interior is empty.

$\boxed{\Longleftarrow}$ Let $A \subseteq B$, where $B \subseteq_C E$ and int$(B) = \varnothing$. By definition, $\overline{A} \subseteq B$ and so int$(\overline{A}) \subseteq$ int$(B) = \varnothing$.

c) $\boxed{\Longrightarrow}$ If $\overline{A} = A$ and int$(\overline{A}) = \varnothing$, then int$(A) =$ int$(\overline{A}) = \varnothing$. Then

$$\partial A = \overline{A} \setminus \text{int}(A) = \overline{A} \setminus \varnothing = \overline{A} = A.$$

$\boxed{\Longleftarrow}$ We have $A = \partial A \iff A = \overline{A} \setminus A \Longrightarrow A \subseteq \overline{A} \setminus \text{int}(A)$. However int$(A) \subseteq A$ so that int$(A) \neq \varnothing \Longrightarrow A \not\subseteq \overline{A} \setminus \text{int}(A)$. Consequently, int$(A) = \varnothing$, which means that $A = \partial A = \overline{A}$ and so $A \subseteq_C E$. Then int$(\overline{A}) =$ int$(A) = \varnothing$.

d) If int$(\overline{A}) = \varnothing$, we have $A \subseteq \overline{A} \Longrightarrow$ int$(A) \subseteq$ int$(\overline{A}) = \varnothing$. Hence

$$\partial A = \overline{A} \setminus \text{int}(A) = \overline{A} \setminus \varnothing = \overline{A}.$$

(What condition must hold for the converse to be satisfied?) ■

9. Show that d_∞, d_1 and d_2 are equivalent on $\mathbb{R}^2$.

Proof: we could do it directly, but notice that these metrics are all derived from norms on $\mathbb{R}^2$. Since $\mathbb{R}^2$ is a finite-dimensional vector space, all norms on $\mathbb{R}^2$ are equivalent. Hence the three metrics are equivalent. That is all there is to it. ■

10. For $i = 1, \ldots, n$, let (E_i, d_i) be metric spaces and $U_i \subseteq_O E_i$. Show that $U_1 \times \cdots \times U_n$ is an open subset of

$$(E, d) = (E_1 \times \cdots \times E_n, \sup\{d_i \mid i = 1, \ldots, n\}).$$

Proof: consider the subset $U = U_1 \times \cdots \times U_n \subseteq E$, where $U_i \subseteq_O E_i$ for all i. Let $\mathbf{x} \in U$. Then $\pi_i(\mathbf{x}) = \mathbf{x}_i \in U_i$ for all i. But $U_i \subseteq_O E_i$ so that $\exists \eta_i > 0$ with $B_{d_i}(\mathbf{x}_i, \eta_i) \subseteq U_i$. Set $\eta = \min\{\eta_i\}_{i=1}^n > 0$. Then

$$\begin{aligned} B(\mathbf{x}, \eta) &= \{\mathbf{y} | d(\mathbf{x}, \mathbf{y}) < \eta\} = \{\mathbf{y} | \sup\{d_i(\mathbf{x}_i, \mathbf{y}_i)\}_{i=1}^n < \eta\} \\ &= \{\mathbf{y} | d_i(\mathbf{x}_i, \mathbf{y}_i) < \eta \, \forall i = 1, \ldots, n\} = \prod_{i=1}^n B_{d_i}(\mathbf{x}_i, \eta) \subseteq \prod_{i=1}^n U_i = U \end{aligned}$$

Consequently, $U \subseteq_O E$. ■

11. For $i = 1, \ldots, n$, let (E_i, d_i) be metric spaces and let $\pi_i : E_1 \times \cdots \times E_n \to E_i$ be defined by $\pi_i(\mathbf{x}_1, \ldots, \mathbf{x}_n) = \mathbf{x}_i$. Show that π_i is open and continuous.

Proof: let $i \in \{1, \ldots, n\}$ and $U \subseteq_O E_i$. Since

$$\pi_i^{-1}(U) = E_1 \times \cdots E_{i-1} \times U \times E_{i+1} \times \cdots E_n,$$

then $\pi_i^{-1}(U) \subseteq_0 E_1 \times \cdots \times E_n$ according to the previous problem, and so π_i is continuous.

Now, suppose that $V \subseteq_O E_1 \times \cdots \times E_n$. We need to show that

$$\pi_i(V) = \{\mathbf{x} \in E_i | \mathbf{x} = \pi_i(\mathbf{y}), \mathbf{y} \in V\} \subseteq_O E_i.$$

Let $\mathbf{u} \in \pi_i(V)$ and consider $\mathbf{x} \in \pi_i^{-1}(\mathbf{u})$. Since $V \subseteq_O E_1 \times \cdots \times E_n$, $\exists r_{\mathbf{x}} > 0$ such that $B_d(\mathbf{x}, r_{\mathbf{x}}) \subseteq V$. We will show that $B_{d_i}(\mathbf{u}, r_{\mathbf{x}}) \subseteq \pi_i(V)$. Let $\mathbf{z} \in B_{d_i}(\mathbf{u}, r_{\mathbf{x}})$. Then $d_i(\mathbf{u}, \mathbf{z}) < r_{\mathbf{x}}$. Set $\mathbf{w} = \mathbf{x}$, except in the ith position, where $\mathbf{w}_i = \mathbf{z}$. Then $\pi_i(\mathbf{w}) = \mathbf{z}$ and

$$d(\mathbf{w}, \mathbf{x}) = \sup\{d_i(\mathbf{w}_i, \mathbf{x}_i)\} = \sup\{0, \ldots, d_i(\mathbf{z}, \mathbf{u}), \ldots, 0\} = d_i(\mathbf{z}, \mathbf{u}) < r_{\mathbf{x}},$$

that is, $\mathbf{w} \in B_d(\mathbf{x}, r_{\mathbf{x}}) \subseteq V$. Thus $\mathbf{z} = \pi_i(\mathbf{w}) \in \pi_i(V)$, and so π_i is open. ■

12. Show that a map $f : (F, \delta) \to (E_1, d_1) \times \cdots \times (E_n, d_n)$ is continuous at $\mathbf{a} \in F$ if and only if $\pi_i \circ f$ is continuous at $\mathbf{a} \in F$ for all i.

Proof: if f is continuous at $\mathbf{a}$, then $\pi \circ f$ is continuous at $\mathbf{a}$ for all i, since π_i is continuous and the composition of continuous functions is continuous.

Now, if $\pi_i \circ f$ is continuous at $\mathbf{a} \in F$ for all i, then, for all $\varepsilon > 0$, $\exists \eta_1, \ldots, \eta_n > 0$ such that $d_i(\pi_i(f(\mathbf{x})), \pi_i(f(\mathbf{a}))) < \varepsilon$ whenever $\delta(\mathbf{x}, \mathbf{a}) < \eta_i$ for all $i = 1, \ldots, n$.

Set $\eta = \sup\{\eta_i\} > 0$. Then, for all $\varepsilon > 0$,

$$d(f(\mathbf{x}), f(\mathbf{a})) = \sup\{d_i(\pi_i(f(\mathbf{x})), \pi_i(f(\mathbf{a})))\} < \varepsilon$$

whenever $\delta(\mathbf{x}, \mathbf{a}) < \eta$; as such, f is continuous at $\mathbf{a}$. ■

13. Let $f : (E_1, d_1) \times \cdots \times (E_n, d_n) \to (F, \delta)$ and $\mathbf{a} = (\mathbf{a}_1, \ldots, \mathbf{a}_n) \in E$. For all i, define $f_i : (E_i, d_i) \to (F, \delta)$ by $f_i(\mathbf{x}) = f(\mathbf{a}_1, \ldots, \mathbf{a}_{i-1}, \mathbf{x}, \mathbf{a}_{i+1}, \ldots, \mathbf{a}_n)$. Show that if f is continuous at $\mathbf{a}$, then f_i is continuous at $\mathbf{a}$ for all i.

Proof: by continuity of f, for all $\varepsilon > 0$, $\exists \eta > 0$ such that

$$d(\mathbf{x}, \mathbf{a}) < \eta \Longrightarrow \delta(f(\mathbf{x}), f(\mathbf{a})) < \varepsilon.$$

For any $\mathbf{x} \in E_i$, write $\tilde{\mathbf{x}} = (\mathbf{a}_1, \ldots, \mathbf{a}_{i-1}, \mathbf{x}, \mathbf{a}_{i+1}, \ldots, \mathbf{a}_n)$. Then, if $d(\tilde{\mathbf{x}}, \mathbf{a}) < \eta$, we have

$$\delta(f_i(\mathbf{x}), f_i(\mathbf{a})) = \delta(f(\tilde{\mathbf{x}}), f(\mathbf{a})) < \varepsilon.$$

Since $d_i(\mathbf{x}, \mathbf{a}_i) \leq d(\tilde{\mathbf{x}}, \mathbf{a}) < \eta$, f_i is continuous at $\mathbf{a}$. ■

14. Show that $d = \sup\{d_i \mid i = 1, \ldots, n\}$ defines a metric on $E = \prod_{i=1}^n (E_i, d_i)$.

Proof: the only property which is not immediately obvious is the triangle inequality (and even at that, it is pretty obvious). Let $\mathbf{x}, \mathbf{y}, \mathbf{z} \in E$. Then

$$\begin{aligned} d(\mathbf{x}, \mathbf{y}) &= \sup\{d_i(\mathbf{x}_i, \mathbf{y}_i)\} \leq \sup\{d_i(\mathbf{x}_i, \mathbf{z}_i) + d_i(\mathbf{z}_i, \mathbf{y}_i)\} \\ &\leq \sup\{d_i(\mathbf{x}_i, \mathbf{z}_i)\} + \sup\{d_i(\mathbf{z}_i, \mathbf{y}_i)\} = d(\mathbf{x}, \mathbf{z}) + d(\mathbf{z}, \mathbf{y}) \end{aligned}$$

So we've got that going for us, which is nice. ■

15. Let (E_i, d_i) be metric spaces for $i = 1, \ldots, n$. Show that the metric product space $(E, d) = (\prod E_i, \sup\{d_i\})$ is complete if and only if (E_i, d_i) is complete for each i.

Proof: Assume (E, d) is complete, and let $(\mathbf{x}_n)$ be a Cauchy sequence in (E_i, d_i) for some i. Then for all $\varepsilon > 0$, $\exists M \in \mathbb{N}$ such that $d_i(\mathbf{x}_n, \mathbf{x}_m) < \varepsilon$ whenever $n, m > M$.

For each $j \neq i$, pick $\mathbf{a}_j \in E_j$.

Write $\mathbf{w}_n = (\mathbf{a}_1, \ldots, \mathbf{a}_{i-1}, \mathbf{x}_n, \mathbf{a}_{i+1}, \ldots, \mathbf{a}_n)$. Then $(\mathbf{w}_n)$ is a Cauchy sequence in E: indeed for all $\varepsilon > 0$, we have

$$\begin{aligned} d(\mathbf{w}_n, \mathbf{w}_m) &= \sup\{d_i(\pi_i(\mathbf{w}_n), \pi_i(\mathbf{w}_m))\} \\ &= \sup\{d_1(\mathbf{a}_1, \mathbf{a}_1), \ldots, d_i(\mathbf{x}_n, \mathbf{x}_m), \ldots, d_n(\mathbf{a}_n, \mathbf{a}_n)\} \\ &= \sup\{0, \ldots, 0, d_i(\mathbf{x}_n, \mathbf{x}_m), 0, \ldots, 0\} = d_i(\mathbf{x}_n, \mathbf{x}_m) < \varepsilon \end{aligned}$$

whenever $n, m > M$.

Since (E, d) is complete, $\exists \mathbf{w} \in E$ for which $\mathbf{w}_n \to \mathbf{w}$. Furthermore, π_i is continuous, so that $\mathbf{x}_n = \pi_i(\mathbf{w}_n) \to \pi_i(\mathbf{w}) \in E_i$, and so $(\mathbf{x}_n)$ converges in (E_i, d_i). Consequently, (E_i, d_i) is complete for all i.

On the other hand, suppose that (E_i, d_i) is complete for all i, and let $(\mathbf{w}_n)$ be a Cauchy sequence in (E, d).

Since $d_i(\pi_i(\mathbf{w}_n), \pi_i(\mathbf{w}_m)) \leq d(\mathbf{w}_n, \mathbf{w}_m)$ for all i, $(\pi_i(\mathbf{w}_n))$ is a Cauchy sequence in (E_i, d_i) for all i. As all (E_i, d_i) are complete, $\exists \mathbf{x}_1, \ldots, \mathbf{x}_n$, $\mathbf{x}_i \in E_i$, such that $\pi_i(\mathbf{w}_n) \to \mathbf{x}_i$ for all i, i.e. for all $\varepsilon > 0$, $\exists M_1, \ldots, M_n \in \mathbb{N}$ such that

$$\forall i,\ d_i(\pi_i(\mathbf{w}_n), \mathbf{x}_i) < \varepsilon \quad \text{whenever } n > M_i.$$

Set $M = \max\{M_i | i = 1, \ldots, n\} < \infty$ and $\mathbf{w} = (\mathbf{x}_1, \ldots, \mathbf{x}_n)$. Let $\varepsilon > 0$.

Then

$$d(\mathbf{w}_n, \mathbf{w}) = \sup\{d_i(\pi_i(\mathbf{w}_n), \pi_i(\mathbf{w}))\} = \sup\{d_i(\pi_i(\mathbf{w}_n), \mathbf{x}_i)\} < \varepsilon$$

whenever $n > M$.

As we have shown that $\mathbf{w}_n \to \mathbf{w} \in E$, we conclude that (E, d) is complete. ■

16. Show that the converse of the previous result does not hold in general, for instance for $f : \mathbb{R}^2 \to \mathbb{R}$ defined by

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

Solution: the problem is that $f(x, 0)$ is continuous at $x = 0$, $f(0, y)$ is continuous at $y = 0$, but $f(x, y)$ is not continuous at $(x, y) = (0, 0)$ since, among other things, $\lim\limits_{z \to 0} f(z, z) = \dfrac{1}{2} \neq 0$. ■

17. Let $d_1, d_2 : \mathbb{N} \times \mathbb{N} \to \mathbb{R}$ be defined according to

$$d_1(m,n) = \begin{cases} 0, & \text{if } m = n \\ 1 + \frac{1}{m+n}, & \text{otherwise} \end{cases} \qquad d_2(m,n) = \frac{|m-n|}{mn}.$$

a) Show that d_1 and d_2 are metrics on $\mathbb{N}$.
b) Show that the topologies of $(\mathbb{N}, d_1)$ and $(\mathbb{N}, d_2)$ are both discrete.
c) Show that $(\mathbb{N}, d_1)$ is complete but that $(\mathbb{N}, d_2)$ is not.
d) What does this say about completeness as a topological property of a space?

Proof:

a) The only property which is not immediately obvious is the triangle inequality.
- If $d_1(m,n) = 0$, then $0 = d_1(m,n) \leq d_1(m,k) + d_1(k,n)$ for all k.

  If $d_1(m,n) \neq 0$ and $d_1(m,k) = 0$, then $d_1(m,n) \leq d_1(m,k) + d_1(k,n)$.

  If $d_1(m,n), d_1(m,k), d_1(k,n) \neq 0$, then

$$d_1(m,n) = 1 + \frac{1}{m+n} \leq 2 + \frac{1}{m+k} + \frac{1}{k+n} = d_1(m,k) + d_1(k,n)$$

  since $\frac{1}{m+n} < 1$.
- For d_2, notice that

$$\begin{aligned} d_2(m,k) + d_2(k,n) &= \frac{|m-k|}{mk} + \frac{|k-n|}{kn} = \frac{n|m-k| + m|k-n|}{mkn} \\ &= \frac{|nm-nk| + |mk-mn|}{mkn} \\ &\geq \frac{|mk-nk|}{mkn} = \frac{|m-n|k}{mkn} = \frac{|m-n|}{mn} = d_2(m,n) \end{aligned}$$

b) For all $n \in \mathbb{N}$, we need to show that $\{n\}$ is open in both $(\mathbb{N}, d_1)$ and $(\mathbb{N}, d_2)$, that is, we must show $\exists r_1, r_2 > 0$ such that $B_{d_i}(n, r_i) \subseteq \{n\}$.
- Pick any $r_1 < 1$. Then

$$B_{d_1}(n, r_1) = \{y \in \mathbb{N} \mid_1 (y,n) < r_1\} = \left\{y \in \mathbb{N} \mid y = n \text{ or } \tfrac{1}{n+y} < 1\right\} = \{n\}.$$

- Simple algebraic manipulations show that $d_2(n,m) \geq \frac{1}{n(n+1)}$ whenever $n \neq m \in \mathbb{N}$. Set $r_2 = \frac{1}{n(n+1)} > 0$. Then

$$B_{d_2}(n, r_2) = \left\{y \in \mathbb{N} \mid_2 (n,y) < \frac{1}{n(n+1)}\right\} = \{n\}$$

c) For completeness:
   - Let (k_n) be a Cauchy sequence in $(\mathbb{N}, d_1)$. Then, for all $1 > \varepsilon > 0$, $\exists M \in \mathbb{N}$ such that $d_1(k_n, k_m) < \varepsilon$ whenever $n, m > M$.

     Since $d_1(x, y) > 1$ for all $x \neq y$, we must have $k_n = k_m$ for all $n, m > M$. Then (k_n) is constant for all $n > M$, and as such, it is a convergent sequence in $(\mathbb{N}, d_1)$.
   - Consider the sequence (n) in $(\mathbb{N}, d_2)$. To show that (n) is a Cauchy sequence, let $\varepsilon > 0$ and $M > \frac{2}{\varepsilon}$. Then

$$d_2(m, n) = \frac{|m - n|}{mn} \leq \frac{m + n}{mn} = \frac{1}{m} + \frac{1}{n} \leq \frac{2}{\min\{m, n\}} < \frac{2}{M} < \varepsilon$$

     whenever $m, n > M$.

     Now, if $n \to K$ in $(\mathbb{N}, d_2)$, then, for $\varepsilon = \frac{1}{K(K+1)}$, $\exists M \in \mathbb{N}$ such that $d(K, n) < \frac{1}{K(K+1)}$ whenever $n > M$ (except for possibly $K = n$).

     But this contradicts the fact that $d(K, n) \geq \frac{1}{K(K+1)}$ whenever $K \neq n$. Hence (n) cannot converge in $(\mathbb{N}, d_2)$.

d) This is yet another example that completeness is not a topological property... ■

18. Let (E, d) be a metric space. Define $d_1, d_2 : E \times E \to \mathbb{R}$ by $d_1(\mathbf{x}, \mathbf{y}) = \frac{d(\mathbf{x},\mathbf{y})}{1+d(\mathbf{x},\mathbf{y})}$ and $d_2(\mathbf{x}, \mathbf{y}) = \min\{d(\mathbf{x}, \mathbf{y}), 1\}$.

   a) Show that d_1 and d_2 are metrics on E.
   b) Show that d is topologically equivalent to d_2.
   c) Show that d_1 is topologically equivalent to d_2.

   Proof:

   a) The only property which is not immediately obvious is the triangle inequality.
      - Let $\mathbf{x}, \mathbf{y}, \mathbf{z} \in E$.

        Write $t = d(\mathbf{x}, \mathbf{y}) \geq 0$, $k = d(\mathbf{x}, \mathbf{z}) \geq 0$, $\ell = d(\mathbf{z}, \mathbf{y}) \geq 0$. Since d is a metric, $t \leq k + \ell$. Since the function $f(w) = \frac{w}{1+w}$ is increasing over $[0, \infty)$,

$$\begin{aligned} d_1(\mathbf{x}, \mathbf{y}) = \frac{t}{1+t} &\leq \frac{k+\ell}{1+k+\ell} = \frac{k}{1+k+\ell} + \frac{\ell}{1+k+\ell} \\ &\leq \frac{k}{1+k} + \frac{\ell}{1+\ell} = d_1(\mathbf{x}, \mathbf{z}) + d_1(\mathbf{z}, \mathbf{w}). \end{aligned}$$

      - Let $\mathbf{x}, \mathbf{y}, \mathbf{z} \in E$. If $d_2(\mathbf{x}, \mathbf{z}) \geq 1$ or $d_2(\mathbf{z}, \mathbf{y}) \geq 1$, then

$$d_2(\mathbf{x}, \mathbf{z}) + d_2(\mathbf{z}, \mathbf{y}) \geq 1 \geq d_2(\mathbf{x}, \mathbf{y}).$$

        If $d_2(\mathbf{x}, \mathbf{z}) < 1$ and $d_2(\mathbf{z}, \mathbf{y}) < 1$, then

$$d_2(\mathbf{x}, \mathbf{z}) \leq d(\mathbf{x}, \mathbf{y}) \leq d(\mathbf{x}, \mathbf{z}) + d(\mathbf{z}, \mathbf{y}) = d_2(\mathbf{x}, \mathbf{z}) + d_2(\mathbf{z}, \mathbf{y}).$$

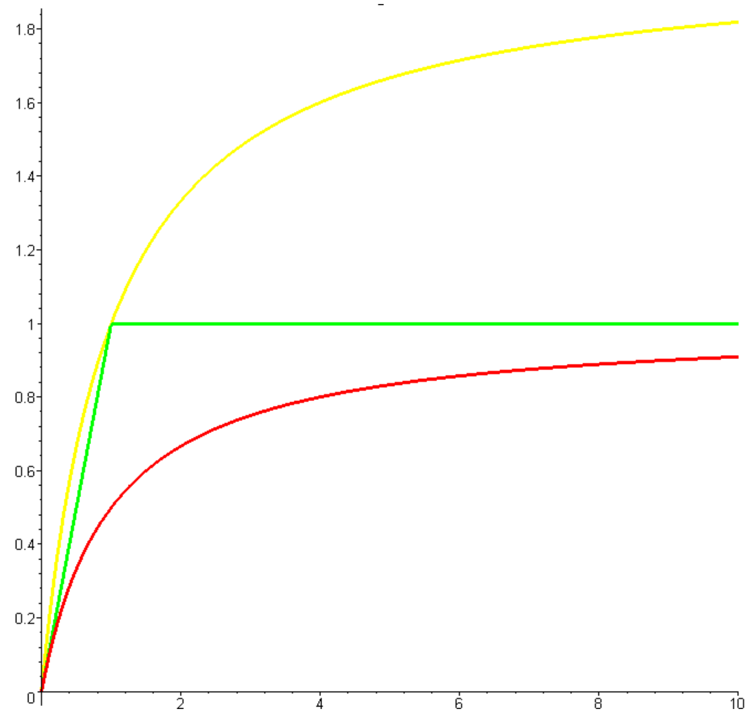
b) Since $d_2 \leq d$, $B_d(\mathbf{x}, r) \subseteq B_{d_2}(\mathbf{x}, r)$ for all $\mathbf{x} \in E$ and $r > 0$. That is, $B_{d_2}(\mathbf{x}, r)$ is open in the $d-$topology.

Similarly, $B_{d_2}(\mathbf{x}, \min\{r, 1\}) \subseteq B_d(\mathbf{x}, r)$ for all $\mathbf{x} \in E$. That is, $B_d(\mathbf{x}, r)$ is open in the d_2-topology. Hence d and d_2 are equivalent.

c) Lengthy but simple manipulations show that

$$\underbrace{d_1}_{\text{red}} \leq \underbrace{d_2}_{\text{green}} \leq \underbrace{2d_1}_{\text{yellow}}$$

and so the metrics are equivalent. □



19. Let (E, d) and $(F, \hat{d})$ be two metric spaces, and let $A \subseteq E$ be dense in E.

a) If $f : (A, d) \to (F, \hat{d})$ is continuous and if $\lim_{\mathbf{y}\to\mathbf{x}, \mathbf{y}\in A} f(\mathbf{y})$ exists for all $\mathbf{x} \in E \setminus A$, show that there exists a unique continuous function $g : E \to F$ with $g|_A = f$.

b) Assume further that $(F, \hat{d})$ is complete. If $f : (A, d) \to (F, \hat{d})$ is uniformly continuous, show that there exists a unique function $g : E \to F$, uniformly continuous, with $g|_A = f$.

Proof:

a) The function $g : E \to F$ that does the trick is given by

$$g(\mathbf{x}) = \begin{cases} f(\mathbf{x}), & \mathbf{x} \in A \\ \lim_{\mathbf{y}\to\mathbf{x}, \mathbf{y}\in A} f(\mathbf{y}), & \mathbf{x} \in E \setminus A \end{cases} \tag{8.4}$$

In order to show that g is continuous, let $\mathbf{x} \in E$ and $(\mathbf{x}_n) \subseteq E$ be such that $\mathbf{x}_n \to \mathbf{x}$. For all $n \in \mathbb{N}$, $g(\mathbf{x}_n) = \lim_{\mathbf{y}\to\mathbf{x}_n, \mathbf{y}\in A} f(\mathbf{y})$. Consequently, for any $n \in \mathbb{N}$, $\exists \mathbf{y}_n \in A$ such that

$$d(\mathbf{x}_n, \mathbf{y}_n) \leq \frac{1}{n} \quad \text{and} \quad \hat{d}(g(\mathbf{x}_n), f(\mathbf{y}_n)) < \frac{1}{n}.$$

From the triangle inequality

$$d(\mathbf{x}, \mathbf{y}_n) \leq d(\mathbf{x}, \mathbf{x}_n) + d(\mathbf{x}_n, \mathbf{y}_n) \leq \frac{1}{n} + d(\mathbf{x}, \mathbf{x}_n)$$

we conclude that $\mathbf{y}_n \to \mathbf{x}$ and so that $f(\mathbf{y}_n) \to g(\mathbf{x})$. Combining this result with

$$\hat{d}(g(\mathbf{x}_n), g(\mathbf{x})) \leq \hat{d}(g(\mathbf{x}_n), f(\mathbf{y}_n)) + \hat{d}(f(\mathbf{y}_n), g(\mathbf{x})) \leq \frac{1}{n} + \hat{d}(f(\mathbf{y}_n), g(\mathbf{x})),$$

we conclude that $g(\mathbf{x}_n) \to g(\mathbf{x})$. By the Sequential Criterion, g is thus continuous at $\mathbf{x}$ for all $\mathbf{x} \in E$, and so it is continuous on E.

It remains only to show that g is the unique function satisfying the conditions outlined in the statement of the problem.

Let $g, h : E \to F$ be two continuous functions with $g|_A = h|_A = f|_A$. Then $g(\mathbf{x}) = h(\mathbf{x})$ for all $\mathbf{x} \in A$.

Now, let $\mathbf{x} \in E \setminus A$. Since A is dense in E, there is a sequence $(\mathbf{x}_n) \subseteq A$ such that $\mathbf{x}_n \to \mathbf{x}$. Since g and h are continuous,

$$g(\mathbf{x}) = \lim_{n\to\infty} g(\mathbf{x}_n) = \lim_{n\to\infty} f(\mathbf{x}_n) = \lim_{n\to\infty} h(\mathbf{x}_n) = h(\mathbf{x}).$$

Hence $g(\mathbf{x}) = h(\mathbf{x})$ for all $\mathbf{x} \in E$. Consequently, $g = h$ on E.

b) Let $\mathbf{x}_0 \in E \setminus A$ and $\varepsilon > 0$. Since f is uniformly continuous on A, $\exists \alpha > 0$ such that $\hat{d}(f(\mathbf{x}), f(\mathbf{y})) < \varepsilon$ whenever $\mathbf{x}, \mathbf{y} \in A$ and $d(\mathbf{x}, \mathbf{y}) < \alpha$.

In particular, if $\mathbf{x}, \mathbf{y} \in A$ are such that $d(\mathbf{x}, \mathbf{x}_0), d(\mathbf{y}, \mathbf{x}_0) < \frac{\alpha}{2}$, then $d(\mathbf{x}, \mathbf{y}) < \alpha$ and $\hat{d}(f(\mathbf{x}), f(\mathbf{y})) < \varepsilon$.

Since $(F, \hat{d})$ is complete, the Cauchy Criterion for Functions (we will discuss this one later) applies and we conclude that $\lim_{\mathbf{y}\to\mathbf{x}_0, \mathbf{y}\in A} f(\mathbf{y})$ exists. According to the result of part (a), the function $g : E \to F$ defined by (8.4) is continuous on E.

It remains only to show that g is uniformly continuous on E.

Let $\varepsilon > 0$. By hypothesis, f is uniformly continuous on A. As a result, $\exists \alpha > 0$ such that $\hat{d}(f(\mathbf{x}), f(\mathbf{y})) < \varepsilon$ whenever $\mathbf{x}, \mathbf{y} \in A$ and $d(\mathbf{x}, \mathbf{y}) < \alpha$.

Let $\mathbf{x}, \mathbf{y} \in E$ satisfy $d(\mathbf{x}, \mathbf{y}) < \alpha$. Since A is dense in E, two sequences $(\mathbf{x}_n), (\mathbf{y}_n) \subseteq$

A can be found such that $\mathbf{x}_n \to \mathbf{x}$ and $\mathbf{y}_n \to \mathbf{y}$. Since d is a continuous mapping, $d(\mathbf{x}_n, \mathbf{y}_n) \to d(\mathbf{x}, \mathbf{y}) < \alpha$ which shows the existence of an index $N \in \mathbb{N}$ such that $d(\mathbf{x}_n, \mathbf{y}_n) < \alpha$ for all $n > N$.

Hence, $\hat{d}(f(\mathbf{x}_n), f(\mathbf{y}_n)) < \varepsilon$ for all $n > N$. By continuity,

$$\hat{d}(f(\mathbf{x}_n), f(\mathbf{y}_n)) \to \hat{d}(g(\mathbf{x}), g(\mathbf{y})) \leq \varepsilon,$$

which shows that g is uniformly continuous on E. ■

20. Let (E, d) be a metric space. Let $\mathcal{C}$ denote the set of Cauchy sequences in E.

a) i. Let $U = (\mathbf{u}_n), V = (\mathbf{v}_n) \in \mathcal{C}$. Show that $(d(\mathbf{u}_n, \mathbf{v}_n))$ converges, and denote its limit by $\delta(U, V)$.

ii. Show that δ is symmetric and satisfies the triangle inequality.

b) Consider the equivalence relation $\sim$ on $\mathcal{C}$ defined by

$$U \sim V \iff \delta(U, V) = 0.$$

Write $\hat{E} = \mathcal{C}/\sim$ and denote the equivalence class of $U \in \mathcal{C}$ in $\hat{E}$ by $\hat{U}$.

i. What is the equivalence class of a sequence which converges in E?

ii. If $U \sim U'$ and $V \sim V'$, show that $\delta(U, V) = \delta(U', V')$. Thus, for $\hat{U}, \hat{V} \in \hat{E}$, the real number $\delta(\hat{U}, \hat{V}) = \delta(U, V)$ is well-defined, not being dependent on the choice of class representatives.

iii. Show that δ is a metric on $\hat{E}$.

iv. Let $\iota : E \to \hat{E}$ be defined by $\iota(\alpha) = \widehat{(\alpha)}$, where (α) is the constant sequence. Show that ι is an isometry (and so also $1-1$). Furthermore, show that $\iota(E)$ is dense in $\hat{E}$.

c) Show that $(\hat{E}, \delta)$ is complete.

d) Let (E_1, d_1) and (E_2, d_2) be complete metric spaces, and suppose that there are isometries $\iota_k : E \to E_k$ with $\iota_k(E)$ dense in E_k, for $k = 1, 2$. Show that there is a unique bijective isometry $\varphi : E_1 \to E_2$ such that $\varphi(\iota_1(\mathbf{x})) = \iota_2(\mathbf{x})$ for all $\mathbf{x} \in E$.

Proof:

a) i. Since $\mathbb{R}$ is complete, it will suffice to show that $(d(\mathbf{u}_n, \mathbf{v}_n))$ is a Cauchy sequence. For all $p, q \in \mathbb{N}$,

$$d(\mathbf{u}_p, \mathbf{v}_p) \leq d(\mathbf{u}_p, \mathbf{u}_q) + d(\mathbf{u}_q, \mathbf{v}_q) + d(\mathbf{v}_p, \mathbf{v}_q)$$
$$d(\mathbf{u}_q, \mathbf{v}_q) \leq d(\mathbf{u}_p, \mathbf{u}_q) + d(\mathbf{u}_p, \mathbf{v}_p) + d(\mathbf{v}_p, \mathbf{v}_q)$$

whence

$$d(\mathbf{u}_p, \mathbf{v}_p) - d(\mathbf{u}_q, \mathbf{v}_q) \leq d(\mathbf{u}_p, \mathbf{u}_q) + d(\mathbf{v}_p, \mathbf{v}_q)$$
$$d(\mathbf{u}_q, \mathbf{v}_q) - d(\mathbf{u}_p, \mathbf{v}_p) \leq d(\mathbf{u}_p, \mathbf{u}_q) + d(\mathbf{v}_p, \mathbf{v}_q)$$

and so $|d(\mathbf{u}_p, \mathbf{v}_p) - d(\mathbf{u}_q, \mathbf{v}_q)| \leq d(\mathbf{u}_p, \mathbf{u}_q) + d(\mathbf{v}_p, \mathbf{v}_q) \to 0$, since both U and V are Cauchy sequences. Consequently, $(d(\mathbf{u}_n, \mathbf{v}_n))$ is a Cauchy sequence.

ii. Symmetry is clear, since the limit of a convergent sequence is unique in a metric space and

$$\delta(V,U) \leftarrow d(\mathbf{v}_n,\mathbf{u}_n) = d(\mathbf{u}_n,\mathbf{v}_n) \to \delta(U,V).$$

The triangle inequality is also obvious since

$$\delta(U,V) \leftarrow d(\mathbf{u}_n,\mathbf{v}_n) \leq d(\mathbf{u}_n,\mathbf{w}_n) + d(\mathbf{w}_n,\mathbf{v}_n) \to \delta(U,W) + \delta(W,V)$$

implies that $\delta(U,V) \leq \delta(U,W) + \delta(W,V)$.

b) i. Let $U = (\mathbf{u}_n)$ be a convergent sequence in E which converges to $\alpha \in E$. Since any convergent sequence is a Cauchy sequence, $U \in \mathcal{C}$. Let $V = (\mathbf{v}_n) \in \mathcal{C}$. Then

$$U \sim V \iff \delta(U,V) = 0 \iff d(\mathbf{u}_n,\mathbf{v}_n) \to 0.$$

Thanks to the inequalities

$$d(\alpha,\mathbf{v}_n) \leq d(\alpha,\mathbf{u}_n) + d(\mathbf{u}_n,\mathbf{v}_n) \quad \text{and} \quad d(\mathbf{u}_n,\mathbf{v}_n) \leq d(\alpha,\mathbf{u}_n) + d(\alpha,\mathbf{v}_n),$$

we see that $U \sim V$ if and only if $d(\alpha,\mathbf{v}_n) \to 0$ (since we already have $d(\alpha,\mathbf{u}_n) \to 0$). Then, $\hat{U} = \{V = (\mathbf{v}_n) \in \mathcal{C} \mid \mathbf{v}_n \to \alpha\}$.

ii. If $U \sim U'$ and $V \sim V'$, then, according to the triangle inequality, we have

$$\delta(U,V) \leq \delta(U,U') + \delta(U',V') + \delta(V,V') = \delta(U',V').$$

Similarly, $\delta(U',V') \leq \delta(U,V)$ so that $\delta(U,V) = \delta(U',V')$.

iii. It remains only to show that $\delta(\hat{U},\hat{V}) = 0$ if and only if $\hat{U} = \hat{V}$. But that is exactly how the equivalence relation was built in the first place.

iv. For any $\alpha \in E$, let $(\alpha) \in \mathcal{C}$ be the constant sequence. Then

$$\delta(\iota(\alpha),\iota(\beta)) = \delta((\alpha),(\beta)) = d(\alpha,\beta)$$

and so ι is an isometry.

Let $\hat{U} \in \hat{E}$, with $U = (\mathbf{u}_n) \in \mathcal{C}$, and $\varepsilon > 0$. Since U is a Cauchy sequence, $\exists N \in \mathbb{N}$ such that for all $p,q > N$ we have $d(\mathbf{u}_p,\mathbf{u}_q) < \varepsilon$. Now, fix $p > N$. Then

$$\delta(\hat{U},\iota(\mathbf{u}_p)) = \delta(U,(\mathbf{u}_p)) = \lim_{n\to\infty} d(\mathbf{u}_n,\mathbf{u}_p) \leq \varepsilon.$$

Since this holds for all $p > N$, we conclude that $\iota(\mathbf{u}_n) \to \hat{U}$. Hence any element of $\hat{E}$ is the limit of a sequence of elements of $\iota(E)$, i.e. $\iota(E)$ is dense in $\hat{E}$.

c) Let (α_n) be a Cauchy sequence in $\hat{E}$. Since $\iota(E)$ is dense in $\hat{E}$, $\forall n \in \mathbb{N}$, $\exists \mathbf{x}_n \in E$ with $\delta(\alpha_n, \iota(\mathbf{x}_n)) < \frac{1}{n}$. Then

$$\begin{aligned} d(\mathbf{x}_p, \mathbf{x}_q) = \delta(\iota(\mathbf{x}_p), \iota(\mathbf{x}_q)) &\leq \delta(\iota(\mathbf{x}_p), \alpha_p) + \delta(\alpha_p, \alpha_q) + \delta(\alpha_q, \iota(\mathbf{x}_q)) \\ &\leq \delta(\alpha_p, \alpha_q) + \frac{1}{p} + \frac{1}{q} \end{aligned}$$

so that $d(\mathbf{x}_p, \mathbf{x}_q) \to 0$ as $p, q \to \infty$, which is to say that $(\mathbf{x}_n) \in \mathcal{C}$. Denote $\alpha = (\hat{\mathbf{x}}_n) \in \hat{E}$.

We will show that $\alpha_n \to \alpha$. Since

$$\delta(\alpha_n, \alpha) \leq \delta(\alpha_n, \iota(\mathbf{x}_n)) + \delta(\iota(\mathbf{x}_n), \alpha) < \frac{1}{n} + \delta(\iota(\mathbf{x}_n), \alpha),$$

it suffices to show that $\delta(\iota(\mathbf{x}_n), \alpha) \to 0$.

Let $\varepsilon > 0$. The sequence $(\mathbf{x}_n)$ being Cauchy in E, $\exists N \in \mathbb{N}$ such that $d(\mathbf{x}_p, \mathbf{x}_q) < \varepsilon$ whenever $p, q \geq N$. Thus, fixing n and letting $p \to \infty$, we have

$$\delta(\iota(\mathbf{x}_n), \alpha) = \lim_{p \to \infty} d(\mathbf{x}_n, \mathbf{x}_p) \leq \varepsilon$$

for all $n > N$, whence we have the desired result.

d) Define φ on $\iota_1(E)$ by setting $\varphi(\iota_1(\mathbf{x})) = \iota_2(\mathbf{x})$ for all $\mathbf{x} \in E$. Restricted to $\iota_1(E)$, the mapping φ is an isometry since

$$d_2(\varphi(\iota_1(\mathbf{x})), \varphi(\iota_1(\mathbf{y}))) = d_2(\iota_2(\mathbf{x}), \iota_2(\mathbf{y})) = d(\mathbf{x}, \mathbf{y}) = d_1(\iota_1(\mathbf{x}), \iota_1(\mathbf{y}))$$

for all $\mathbf{x}, \mathbf{y} \in E$. Thus, φ is uniformly continuous on $\iota_1(E)$. Since $\iota_1(E)$ is dense in E_1 and since E_2 is complete, we can apply the result of a previous problem to show that φ can be extended to a unique uniformly continuous function on E_1.

Furthermore, φ is an isometry on $\iota_1(E)$; since $\iota_1(E)$ is dense in E_1 and since φ is continuous on E_1, φ is an isometry on E_1 in its entirety. In particular φ is $1-1$.

It remains only to show that φ is onto. Let $\beta \in E_2$. As $\iota_2(E)$ is dense in E_2, $\exists (\beta_n) = (\iota_2(\mathbf{x}_n)) \subseteq \iota_2(E)$ such that $\beta_n \to \beta$. Since

$$d_1(\iota_1(\mathbf{x}_p), \iota_1(\mathbf{x}_q)) = d(\mathbf{x}_p, \mathbf{x}_q) = d_2(\iota_2(\mathbf{x}_p), \iota_2(\mathbf{x}_q)) = d_2(\beta_p, \beta_q)$$

for all $p, q \in \mathbb{N}$, the sequence $(\iota_1(\mathbf{x}_n))$ is a Cauchy sequence in E_1. But E_1 is complete so that $\iota_1(\mathbf{x}_n) \to \alpha \in E_1$. Since φ is continuous, we have

$$\varphi(\alpha) = \lim_{n \to \infty} \varphi(\iota_1(\mathbf{x}_n)) = \lim_{n \to \infty} \iota_2(\mathbf{x}_n) = \lim_{n \to \infty} \beta_n = \beta,$$

that is, φ is onto. ■

21. Let $A, B \subseteq E$, where E is endowed with any metric you care to imagine. Show that

   a) $A \subseteq \overline{A}$
   b) $\overline{(\overline{A})} = \overline{A}$
   c) $\overline{A \cup B} = \overline{A} \cup \overline{B}$
   d) $\overline{\varnothing} = \varnothing$
   e) in general, $\overline{A \cap B} \neq \overline{A} \cap \overline{B}$

   Proof:

   a) This one is clear by definition.
   b) By part (a), $\overline{A} \subseteq \overline{(\overline{A})}$. Conversely, since $\overline{(\overline{A})}$ is the smallest closed set containing $\overline{A}$ and since $\overline{A}$ is also a closed set containing $\overline{A}$, then $\overline{(\overline{A})} \subseteq \overline{A}$. Hence, $\overline{A} = \overline{(\overline{A})}$.
   c) Since the union of two closed sets is closed, $\overline{A} \cup \overline{B}$ is a closed set containing $A \cup B$ and so $\overline{A \cup B} \subseteq \overline{A} \cup \overline{B}$. Conversely, $\overline{A \cup B}$ is a closed set containing both A and B, so both $\overline{A}, \overline{B} \subseteq \overline{A \cup B}$; therefore $\overline{A} \cup \overline{B} \subseteq \overline{A \cup B}$. Thus $\overline{A \cup B} = \overline{A} \cup \overline{B}$.
   d) Since $\varnothing$ is always a closed set, $\overline{\varnothing} = \varnothing$.
   e) Consider the following example in $(\mathbb{R}, d_2)$: let $A = (-1, 0)$ and $B = (0, 1)$. Then $\overline{A} = [-1, 0]$, $\overline{B} = [0, 1]$, $A \cap B = \varnothing$, $\overline{A \cap B} = \varnothing$ while $\overline{A} \cap \overline{B} = \{0\}$. ■

22. Let A be a subset of (E, d). Show that $\overline{A} = \text{int}(A) \cup \partial A$.

   Proof: suppose that $\mathbf{x} \in \text{int}(A)$. Then $\mathbf{x} \in A \subseteq \overline{A}$. Now suppose that $\mathbf{x} \in \partial A$. We proceed by contradiction. If $\mathbf{x} \notin \overline{A}$ then, since $E \setminus \overline{A} \subseteq_O E$, $\exists r > 0$ such that $B(\mathbf{x}, r) \subseteq E \setminus \overline{A} \subseteq E \setminus A$. This contradicts the fact that $\mathbf{x} \in \partial A$ (how?) and so we must have $\mathbf{x} \in \overline{A}$. Thus $\text{int}(A) \cup \partial A \subseteq \overline{A}$.

   Conversely, suppose that $\mathbf{x} \in \overline{A}$. There are only three possibilities: $\mathbf{x} \in \text{int}(A)$, $\mathbf{x} \in \partial A$ or $\mathbf{x} \in \text{int}(E \setminus A)$ (why?). If $\mathbf{x} \in \text{int}(E \setminus A)$, then $\exists r > 0$ such that $B(\mathbf{x}, r) \subseteq E \setminus A$. This implies that $A \subseteq E \setminus B(\mathbf{x}, r)$. Therefore $\overline{A} \subseteq E \setminus B(\mathbf{x}, r)$, since $E \setminus B(\mathbf{x}, r) \subseteq_C E$, which in turns implies that $\mathbf{x} \notin \overline{A}$, a contradiction.

   Thus $\mathbf{x} \in \text{int}(A) \cup \partial A$ and so $\overline{A} \subseteq \text{int}(A) \cup \partial A$. ■

23. Let $A = \{\frac{1}{n} \mid n \in \mathbb{N}^\times\}$. Under the usual topology on $\mathbb{R}$, show that every point of A is a boundary point and that the only cluster point of A is 0.

   Proof: To show that every point of $x \in A$ is a boundary point, note that any neighbourhood V of x contains an open interval $I_r = (x - r, x + r)$, for some $r > 0$. But $x \in I_r \cap A$ and since any open interval contains an irrational number $I_r \cap (\mathbb{R} \setminus A) \neq \varnothing$. Consequently, any neighbourhood of x contains both points in A and points not in A, which is another definition of $x \in \partial A$.

   To show that 0 is a cluster point of A, note that any neighbourhood of 0 in $(\mathbb{R}, d_2)$ contains an interval of the form $(-\varepsilon, \varepsilon)$ for some $\varepsilon > 0$. By the Archimedean property, $\exists N \in \mathbb{N}$ such that $\frac{1}{N} < \varepsilon$. Hence $0 \neq \frac{1}{N} \in B(0, \varepsilon)$ and so 0 is a cluster point of A.

In order to show that there are no other cluster points, first observe that any $x < 0$ cannot be a cluster point of A since the neighbourhood $(-2x, 0)$ contains no points in A. Likewise, any $x > 1$ cannot be a cluster point of A since the neighbourhood $(1, 2x)$ contains no point of A.

If $x \in (0, 1]$, then either $x \in A$ or $x \notin A$. If $x = \frac{1}{n} \in A$, then the open neighbourhood $(x - r, x + r)$ contains no other point of A as long as $r < \frac{1}{n(n-1)}$, and so x is not a cluster point of A. If $x \notin A$, choose $k \in \mathbb{N}$ such that $x \in (\frac{1}{k}, \frac{1}{k-1})$. Then the open neighbourhood $(x - r, x + r)$ contains no other point of A if $r < \min\{x - \frac{1}{k}, \frac{1}{k-1} - x\}$ and so x cannot be a cluster point of A. ■

24. Let $\tau_1 = \{U \subseteq \mathbb{R} \mid \mathbb{R} \setminus U \text{ is finite or } U = \varnothing\}$, $\tau_2 = \{U \subseteq \mathbb{R} \mid \mathbb{R} \setminus U \text{ is countable or } U = \varnothing\}$.

   a) Show that τ_1 and τ_2 define topologies on $\mathbb{R}$ (the **co-finite topology** and **countable complement** topology, respectively).
   b) What is the boundary of the set $A = \{\frac{1}{n} \mid n \in \mathbb{N}^\times\}$ under these two topologies?

Proof:

a) It suffices to verify that the three properties hold for τ_1:

   i. $\varnothing \in \tau_1$ by definition; $\mathbb{R} \in \tau_1$ since $\mathbb{R} \setminus \mathbb{R} = \varnothing$ is finite.

   ii. Let $\{X_\alpha\} \subseteq \tau_1$. Then $\mathbb{R} \setminus X_\alpha$ is finite for all α. According to the de Morgan's Laws, the set
   $$\mathbb{R} \setminus \bigcup_\alpha X_\alpha = \bigcap_\alpha (\mathbb{R} \setminus X_\alpha)$$
   is a finite set as it is the intersection of an arbitrary collection of finite sets. Hence, $\bigcup X_\alpha \in \tau_1$.

   iii. Let $\{X_i\}_{i=1}^n \subseteq \tau_1$. Then $\mathbb{R} \setminus X_i$ is finite for all $i = 1, \ldots, n$.

   According to the de Morgan's Laws, the set
   $$\mathbb{R} \setminus \bigcap_{i=1}^n X_i = \bigcup_{i=1}^n (\mathbb{R} \setminus X_i)$$
   is a finite set as it is the union of a finite collection of finite sets. Hence, $\bigcap_{i=1}^n X_i \in \tau_1$.

Now for τ_2:

   i. $\varnothing \in \tau_2$ by definition; $\mathbb{R} \in \tau_2$ since $\mathbb{R} \setminus \mathbb{R} = \varnothing$ is countable.

   ii. Let $\{X_\alpha\} \subseteq \tau_2$. Then $\mathbb{R} \setminus X_\alpha$ is countable for all α.

   According to the de Morgan's Laws, the set
   $$\mathbb{R} \setminus \bigcup_\alpha X_\alpha = \bigcap_\alpha (\mathbb{R} \setminus X_\alpha)$$
   is a countable set as it is the intersection of an arbitrary collection of countable sets. Hence, $\bigcup X_\alpha \in \tau_2$.

iii. Let $\{X_i\}_{i=1}^n \subseteq \tau_2$. Then $\mathbb{R} \setminus X_i$ is countable for all $i = 1, \ldots, n$. According to the de Morgan's Laws, the set

$$\mathbb{R} \setminus \bigcap_{i=1}^n X_i = \bigcup_{i=1}^n (\mathbb{R} \setminus X_i)$$

is a countable set as it is the union of a finite collection of countable sets. Hence, $\bigcap_{i=1}^n X_i \in \tau_2$.

b) In the countable complement topology, $A \subseteq_C \mathbb{R}$, because $\mathbb{R} \setminus (\mathbb{R} \setminus A) = A$ is countable and so $\mathbb{R} \setminus A \subseteq_O \mathbb{R}$. Consequently, $\overline{A} = A$. Furthermore, the only open set of $\mathbb{R}$ contained in A is the empty set, as any other open set is uncountable. Hence $\text{int}(A) = \varnothing$ and $\partial A = \overline{A} \setminus \text{int}(A) = A$.

In the co-finite topology, the only closed set containing A is $\mathbb{R}$, as any other closed set is finite. Consequently, $\overline{A} = \mathbb{R}$. Furthermore, the only open set of $\mathbb{R}$ contained in A is the empty set, as any other open set is infinite. Hence $\text{int}(A) = \varnothing$ and $\partial A = \mathbb{R}$. ■

25. Let $A, B \subseteq (E, d)$. If $\mathbf{x} \in E$ is a cluster point of $A \cap B$, show that $\mathbf{x}$ is a cluster point of both A and B.

Proof: let $\mathbf{x}$ be a cluster point of $A \cap B$. Then any neighbourhood V of $\mathbf{x}$ contains a point $\mathbf{y} \in A \cap B \subseteq A$ such that $\mathbf{y} \neq \mathbf{x}$. Thus $\mathbf{y}$ is a cluster point of A. The argument for B is identical. ■

26. Show that $B \subseteq (\mathbb{R}^p, d_2)$ is closed if and only if every convergent sequence in B converges to a point in B.

Proof: first, assume that B is closed. Let $\mathbf{x} = \lim \mathbf{x}_n$. Then, for any $\varepsilon > 0$, $\exists n_\varepsilon > 0$ such that $\mathbf{x}_n \in B(\mathbf{x}, \varepsilon)$ for all $n \geq n_\varepsilon$. Consequently, $B \cap B(\mathbf{x}, \varepsilon) \neq \varnothing$ for all $\varepsilon > 0$. Since $\mathbb{R}^p \setminus B \subseteq_O \mathbb{R}^p$, it follows that $\mathbf{x} \in B$ (why?).

Conversely, assume that for every convergent sequence $(\mathbf{x}_k) \subseteq \mathbb{R}^p$, we have $\mathbf{x} = \lim \mathbf{x}_k \in B$. If $\mathbb{R}^p \setminus B$ is not open in $\mathbb{R}^p$, $\exists \mathbf{x} \in \mathbb{R}^p \setminus B$ such that $B(\mathbf{x}, \frac{1}{n}) \cap B \neq \varnothing$ for all $n \in \mathbb{N}$. Then $\exists \mathbf{x}_n \in B(\mathbf{x}, \frac{1}{n}) \cap B$; the sequence $(\mathbf{x}_n) \subseteq B$ converges to $\mathbf{x} \notin B$, which contradicts the hypothesis. Hence $\mathbb{R}^p \setminus B \subseteq_O \mathbb{R}^p$. ■

27. Let $(\mathbf{x}_n) \subseteq (\mathbb{R}^p, \|\cdot\|)$ such that

$$\|\mathbf{x}_{n+1} - \mathbf{x}_n\| \leq r\|\mathbf{x}_n - \mathbf{x}_{n-1}\|$$

where $r < 1$. Show that $(\mathbf{x}_n)$ converges.

Proof: we have $\|\mathbf{x}_3 - \mathbf{x}_2\| \leq r\|\mathbf{x}_2 - \mathbf{x}_1\|$ and it is easily seen by induction that if

$$\|\mathbf{x}_{n+1} - \mathbf{x}_n\| \leq r^{n-1}\|\mathbf{x}_2 - \mathbf{x}_1\|$$

then

$$\|\mathbf{x}_{n+2} - \mathbf{x}_{n+1}\| \leq r\|\mathbf{x}_{n+1} - \mathbf{x}_n\| \leq r^n\|\mathbf{x}_2 - \mathbf{x}_1\|.$$

Therefore, if $m > n$,

$$\|\mathbf{x}_m - \mathbf{x}_n\| = \left\| \sum_{k=n}^{m-1} (\mathbf{x}_{k+1} - \mathbf{x}_k) \right\| \leq \sum_{k=n}^{m-1} \|\mathbf{x}_{k+1} - \mathbf{x}_k\|$$
$$\leq \sum_{k=n}^{\infty} \|\mathbf{x}_{k+1} - \mathbf{x}_k\| \leq \sum_{k=n}^{\infty} r^{k-1} \|\mathbf{x}_2 - \mathbf{x}_1\| \leq \frac{r^{n-1}}{1-r} \|\mathbf{x}_2 - \mathbf{x}_1\|.$$

Let $\varepsilon > 0$. Since $r < 1$, $\exists N_\varepsilon$ so that

$$r^{n-1} < \varepsilon \frac{1-r}{\|\mathbf{x}_2 - \mathbf{x}_1\|} \quad \text{for all } n \geq N,$$

and so $\|\mathbf{x}_m - \mathbf{x}_n\| < \varepsilon$ for all $m \geq n \geq N_\varepsilon$. It follows that $(\mathbf{x}_n)$ is Cauchy and that it is convergent, since $(\mathbb{R}^p, \|\cdot\|)$ is a Banach space. ■

8.4 Exercises

1. Prepare a 2-page summary of this chapter, with important definitions and results.
2. Show that the absolute value defines a norm on $\mathbb{R}$.
3. Show that the modulus defines a norm on $\mathbb{C}$.
4. Show that the sup norm $\|\cdot\|_\infty$ is indeed a norm on $\mathcal{C}_{\mathbb{R}}([0,1])$.
5. Let $\infty \geq p \geq 1$. Show that the $p-$norm $\|\cdot\|_\infty$ is indeed a norm on $\mathbb{R}^n$.
6. Let $p \geq 1$. Show that (8.1, p. 189), defines a norm on $\mathcal{L}^p([0,1])$.
7. Prove Lemma 8.1.1, p. 189.
8. Let E be any set. Show that (8.2, p. 190) defines a metric on E.
9. Let $E = \mathbb{R}^n$. Show that d_2 is a metric on E.
10. Let $E = \mathbb{R}$, $d(x,y) = |x-y|$, $A = \mathbb{N}$ and $B = \{\frac{n-1}{n} \mid n \in \mathbb{N}\}$. Compute $d(A,B)$, where d is as in (8.3, p. 191)). Can you use this result to show that (8.3, p. 191) does not define a metric on $\wp(E) \setminus \varnothing$?
11. In a metric space, show that $\delta(A) \in [0,\infty]$. Also, show that $\delta(A) = 0 \iff A$ is a singleton.
12. Prove or disprove: In any metric space (E,d), $\delta_d(B(\mathbf{a},r)) = 2r$.
13. Prove or disprove: Let d, d' be metrics on E. Then, A is bounded in (E,d) if and only if A is bounded in (E,d').
14. Where does the proof that a finite intersection of open subsets is open fail for arbitrary intersections?

15. Show that the metric space topology on a discrete metric space is the discrete topology.

16. Show that the intersection of an arbitrary family $\{A_i\}_{i\in I}$ of closed subsets of E is a closed subset of E.

17. Show that the union of a finite family $\{A_i\}_{i=1}^{\ell}$ of closed subsets of E is a closed subset of E.

18. Show that the union of an arbitrary family of closed subsets of E need not be closed in E.

19. Let A be a subset of a metric space (E,d). Show that $\overline{A}$ is the intersection of all closed subsets of E containing A.

20. Let A be a subset of a metric space (E,d). Show that $A \subseteq \overline{A}$.

21. Prove Lemma 92, p. 197.

22. In Proposition 94, p. 198, show that $2. \iff 3 \iff 4.$

23. Let A be a subset of a metric space (E,d). Show that int(A) is the union of all open subsets of E contained in A.

24. Let A be a subset of a metric space (E,d). Show that int$(A) \subseteq A$.

25. Let A be a subset of a metric space (E,d). Show that $A \subseteq_O E \iff A = \text{int}(A)$.

26. Complete the proof of Lemma 98, p. 202.

27. Prove Proposition 99, p. 202.

28. Show that the three definitions of continuity are equivalent.

29. Let $f: C \to D$, $A \subseteq C$ and $B \subseteq D$. Show that $f^{-1}(f(A)) = A$ and that in general, the best we can say is that $f(f^{-1}(B)) \subseteq B$.

30. Can you find a function $f: E \to \tilde{E}$ which is continuous but not closed?

31. Can you find a function $f: E \to \tilde{E}$ which is open and closed but not continuous?

32. Can you find a function $f: E \to \tilde{E}$ which is open and continuous but not closed?

33. Complete the proof of Proposition 101, p. 204.

34. Complete the proof of Corollary 102, p. 204.

35. Provide the details showing that d_2 and d_∞ are topologically equivalent on $\mathbb{R}^2$.

36. Consider the metric space $(\mathbb{R}, d_2)$. Define a new function $\tilde{d} : \mathbb{R} \times \mathbb{R} \to \mathbb{R} \to \mathbb{R}$ by
$$\tilde{d}(x,y) = \frac{d(x,y)}{1+d(x,y)}.$$
Show that $\tilde{d}$ defines a metric on $\mathbb{R}$, that d and $\tilde{d}$ are topologically equivalent but that they are not equivalent.

37. Let (E,d) be a metric space. Show that $d : E \times E \to \mathbb{R}$ is Lipschitz continuous (with $k = 2$) and so that it is a continuous map.

38. Find a function which is uniformly continuous but not Lipschitz continuous.

39. Show that the two definitions of convergence of a sequence are equivalent.

40. Show that if $\mathbf{x}_n \to \mathbf{x}$, then any subsequence of $(\mathbf{x}_n)$ also converges to $\mathbf{x}$.

41. Show that the set of limit points of a sequence is closed.

42. Complete the proof of Proposition 103, p. 209.

43. Prove Proposition 8.2.2, p. 214.

44. Show that the space $\ell^2(\mathbb{N})$ is a Hilbert space as follows.
    a) Show that $\ell^2(\mathbb{N})$ is a vector space over $\mathbb{C}$.
    b) Show that $(\cdot|\cdot)$ defined in the text is indeed an inner product over $\ell^2(\mathbb{N})$.
    c) Show that $(\cdot|\cdot)$ defines a norm $\|\cdot\|$ over $\ell^2(\mathbb{N})$.
    d) Show that $\ell^2(\mathbb{N})$ is complete under $\|\cdot\|$.

45. Let $f : \mathbb{R} \to \mathbb{R}$ be defined by
$$f(x) = \begin{cases} 1, & x < 0 \\ x + \frac{1}{x+1}, & x \geq 0 \end{cases}.$$
Show that f has no fixed point but that $d(f(x), f(y)) \leq d(x,y)$ for all $x, y \in \mathbb{R}$.

46. Let X be a compact metric space. Define
$$C_{\mathbb{R}}(X) = \{f | f : X \to \mathbb{R}, f \text{ continuous}\}.$$
Show that $(C_{\mathbb{R}}(X), \|\cdot\|_\infty)$ is a Banach space, but that neither $(C_{\mathbb{R}}(X), \|\cdot\|_1)$ nor $(C_{\mathbb{R}}(X), \|\cdot\|_2)$ is complete.

47. Let $E = \{f \in C_B(\mathbb{R},\mathbb{R}) | f \text{ uniformly continuous}\}$. Show that E is a complete subalgebra of $C_B(\mathbb{R},\mathbb{R})$.

48. Let (E, d) be a complete metric space and $f : E \to E$. If there exists a positive integer r and $k \in (0, 1)$ such that
$$f^r = \underbrace{f \circ f \circ \cdots \circ f}_{r \text{ times}}$$
and $d(f^r(x), f^r(y)) \leq kd(x, y)$ for all $x, y \in E$, show that f has a unique fixed point.

49. Let $X = (0, \infty)$. Consider the function $\tilde{d} : X \times X \to \mathbb{R}_0^+$ defined by
$$\tilde{d}(x, y) = \left| \frac{1}{x} - \frac{1}{y} \right|.$$
   a) Prepare a 2-page summary of this chapter; identify the important definitions and results.
   b) Show that $\tilde{d}$ is a metric on X.
   c) Show that $\tilde{d}$ and d_2 induce the same topology on X (*i.e.* the open sets of X are exactly the same under both metrics).
   d) Show that $(X, \tilde{d})$ is not a complete metric space.
   e) Show that $((0, 1], \tilde{d})$ is a complete metric space.

50. Let $\mathcal{B}(X, \mathbb{R})$ denote the set of bounded functions from X to $\mathbb{R}$. It is easy to see that $\mathcal{B}(X, \mathbb{R})$ is a vector space over $\mathbb{R}$. The norm of $f \in \mathcal{B}(X, \mathbb{R})$ is defined by
$$\|f\| = \sup_{x \in X} |f(x)|.$$
Show that $\mathcal{B}(X, \mathbb{R})$ is a Banach space with this norm.

51. Are the co-finite topologies and the countable complement topologies derived from a metric?

Chapter 9

Metric Spaces and Topology

One of the natural ways we can extend the concepts we have discussed in the previous chapters is by moving from $\mathbb{R}$ to $\mathbb{R}^m$. Some of the notions that generalize nicely to vectors and functions on vectors include **compactness** and **connectedness**.

The symbol $\mathbb{K}$ is sometimes used to denote either $\mathbb{R}$ or $\mathbb{C}$.

9.1 Compact Spaces

Let A be a finite set. A function $f : A \to \mathbb{K}$ is necessarily **bounded** (in the sense that $\exists M \in \mathbb{K}$ such that $|f(a)| \leq M$ for all $a \in A$).

Might this be due to the **finiteness** of A? While finiteness is sufficient, it is not a necessary condition for boundedness: the Dirichlet function $\chi_{\mathbb{Q}} : [0,1] \to \mathbb{R}$ is bounded, even though its domain is the **uncountable** set $[0,1]$.

Perhaps it is the **boundedness of the function's domain** that does the trick? Unfortunately, that condition is neither sufficient nor necessary, as can be seen from the functions

$$f : [0,1] \to \mathbb{R}, \quad f(x) = \frac{1}{x} \quad \text{for } x > 0, \quad \text{and} \quad f(0) = 0,$$

and $g : \mathbb{R} \to \mathbb{R}$ defined by $g(x) = \exp(-x^2)$.

Could the culprit instead be the **continuous nature** of the function? Not as such, no, as we have examples of continuous functions being bounded, others being unbounded; and non-continuous functions being bounded, others being unbounded.

A condition on the domain of the function alone cannot guarantee boundedness; and neither can one on the nature of the function. However, a **combination** of two conditions, one each on the domain and on the function, can provide such a guarantee.

In this section, we study the appropriate property on the domain, that of **compactness**, which generalizes the property of finiteness. Its definition, which in all honesty is not super intuitive, is due to Borel and Lebesgue, is applicable to metric and general topological spaces alike.

9.1.1 The Borel-Lebesgue Property

A space E is **compact** if any family of open subsets covering E contains a finite sub-family which also covers E. In other words, E is compact if, for any collection $\mathcal{U} = \{U_i\}_{i\in I}$ of open subsets $U_i \subseteq_O E$ with $E \subseteq \bigcup_{i\in I} U_i$, $\exists$ a finite $J \subseteq I$ such that $E \subseteq \bigcup_{j\in J} U_j$.

Examples

1. Every finite metric space (E, d) is compact.

   Proof: let $\mathcal{U}$ be an open cover of $E = \{\mathbf{x}_1, \dots, \mathbf{x}_n\}$. Thus, for each $1 \leq i \leq n$, $\exists U_i \in \mathcal{U}$ such that $\mathbf{x}_i \in U_i$. Then $\{U_1, \dots, U_n\}$ is a finite subcover of E. ■

2. In the standard topology, $\mathbb{R}$ is not compact.

   Proof: consider the open cover $\mathbb{R} = \bigcup_{n\in\mathbb{N}} (-n, n)$.

   Any finite subcollection $\{(-n_1, n_1), \dots, (-n_m, n_m)\}$ is bounded by $M = \max\{n_j \mid 1 \leq j \leq m\}$, and thus cannot be a cover of $\mathbb{R}$ according to the Archimedean Property. Consequently, no such finite subcover exists and $\mathbb{R}$ is not compact. ■

3. Show that $\mathbb{R}$ is compact in the indiscrete topology.

   Proof: the only open cover of $\mathbb{R}$ in the indiscrete topology is $\{\mathbb{R}\}$, which is already a finite sub-cover of $\mathbb{R}$ (the only other open subset of $\mathbb{R}$ in the indiscrete topology is $\varnothing$). ■

4. Show that any compact metric (E, d) space is bounded.

   Proof: consider the open cover $\mathcal{U} = \{B(\mathbf{x}, 1) \mid \mathbf{x} \in E\}$. Since E is compact, $\exists \mathbf{x}_1, \dots, \mathbf{x}_n \in E$ such that $E = B(\mathbf{x}_1, 1) \cup \cdots B(\mathbf{x}_n, 1)$. Consequently, E has a finite diameter $\leq n$ and is thus bounded. ■

By abuse of notation, we often write: "let $\bigcup U_i$ be an open cover of E" rather than "let $\{U_i\}$ be an open cover of E," as in the second example above.

Incidentally, does the fourth example contradict the third one? It doesn't actually, but what does that imply about the indiscrete topology?

The duality open $\leftrightsquigarrow$ closed/ union $\leftrightsquigarrow$ intersection yields an equivalent definition: a space E is **compact** if any family of closed subsets of E with an empty intersection contains a finite sub-family whose intersection is also empty.

In other words, E is compact if, for any collection $\mathcal{W} = \{V_i\}_{i\in I}$ of closed subsets $V_i \subseteq_C E$ with $\bigcap_{i\in I} V_i = \varnothing$, $\exists$ a finite $J \subseteq I$ such that $\bigcap_{j\in J} V_j = \varnothing$.

Proposition 115
Let $(F_n)_{n\geq 1}$ be a decreasing sequence of non-empty closed subsets of a compact space E. Then $\bigcap_{n\geq 1} F_n \neq \varnothing$.

Proof: if $\bigcap_{n\geq 1} F_n = \varnothing$, then $E = \bigcup_{n\geq 1} E \setminus F_n$, where $E \setminus F \subseteq_O E$. Since E is compact, $\exists$ a finite subsequence of indices $n_1 < \cdots < n_k$ such that

$$E = \bigcup_{i=1}^{k} E \setminus F_{n_i}.$$

Consequently, $\bigcap_{i=1}^{k} F_{n_i} = \varnothing$. But the original sequence is decreasing, so that

$$\bigcap_{i=1}^{k} F_{n_i} = F_{n_k} = \varnothing,$$

which contradicts the hypothesis that all F_n are non-empty. As a result, we conclude that $\bigcap_{n\geq 1} F_n \neq \varnothing$. ■

Continuous functions on compact domains have quite useful properties.

Proposition 116
Let $f : (E, d) \to (F, \delta)$ be any continuous function over a compact metric space. Then f is uniformly continuous.

Proof: let $\mathbf{x} \in E$. Since f is continuous at $\mathbf{x} \in E$, $\forall \varepsilon > 0$, $\exists M_{\mathbf{x}}(\varepsilon) > 0$ such that

$$f(B(\mathbf{x}, M_{\mathbf{x}})) \subseteq B(f(\mathbf{x}), \varepsilon).$$

Furthermore, $E = \bigcup_{\mathbf{x}\in E} B(\mathbf{x}, M_{\mathbf{x}})$ is an open cover of E, which is compact. Consequently, $\exists \mathbf{x}_1, \ldots, \mathbf{x}_n \in E$ such that $E = \bigcup_{i=1}^{n} B(\mathbf{x}_i, M_{\mathbf{x}_i})$. Set

$$M = M(\varepsilon) = \tfrac{1}{2} \cdot \min\{M_{\mathbf{x}_1}, \ldots, M_{\mathbf{x}_n}\} > 0.$$

Then, $\forall_{\varepsilon>0}$, $\exists M(\varepsilon) > 0$ such that $f(B(\mathbf{x}, M)) \subseteq B(f(\mathbf{x}), \varepsilon)$ for all $\mathbf{x} \in E$. As M does not depend on $\mathbf{x}$, f is uniformly continuous. ■

A subset $A \subseteq E$ is deemed to be a **compact subset of** E, which we denote by $A \subseteq_K E$, if any family of open subsets of E covering A contains a finite sub-family which also covers A.

Proposition 117
A finite union of compact subsets of E is itself compact.

Proof: let $A_1, \ldots, A_n \subseteq_K E$ and write $A = \bigcup_{k=1}^n A_k$. Let $\{U_i\}_{i\in I} \subseteq \wp(E)$ be an open cover of A. Then $\{U_i\}_{i\in I}$ is also an open cover of A_k for each k.

Since all A_k are compact, $\exists$ finite $J_1, \ldots, J_k \subseteq I$ such that $A_k \subseteq \bigcup_{j\in J_k} U_j$ for each k. Thus, $A \subseteq \bigcup_{k=1}^n \bigcup_{j\in J_k} U_j$. But $\bigcup_{k=1}^n \{U_j\}_{j\in J_k}$ is a finite sub-family of $\{U_i\}_{i\in I}$, from which we conclude that $A \subseteq_K E$. ■

The infinite union of compact subsets could be compact or not, however.

Examples

1. Both $[0,1], [2,3] \subseteq_K (\mathbb{R}, d_1)$, so $[0,1] \cup [2,3] \subseteq_K (\mathbb{R}, d_1)$. □
2. For any $x \geq 1$, $[0, \frac{1}{x}] \subseteq_K (\mathbb{R}, d_1)$. The union $\bigcup_{x\geq 1}[0, \frac{1}{x}] = [0,1]$ is also a compact subset of $(\mathbb{R}, d_1)$. □
3. For any $n \in \mathbb{N}$, $[-n, n] \subseteq_K (\mathbb{R}, d_1)$, but the union $\bigcup_{n\in\mathbb{N}}[-n, n] = \mathbb{R}$ is not a compact subset of $(\mathbb{R}, d_1)$. □

9.1.2 The Bolzano-Weierstrass Property

For metric spaces, compactness can also be established via a property of **sequences** which is often easier to ascertain than the Borel-Lebesgue property, but it comes with a warning: **the two properties are not equivalent in general for non-metric spaces**.

Let (E, d) be a metric space. We say that E is **precompact** if $\forall \varepsilon > 0$, $\exists \mathbf{x}_1, \ldots, \mathbf{x}_n \in E$ such that $E = \bigcup_{i=1}^n B(\mathbf{x}_i, \varepsilon)$.

Proposition 118
A compact space is precompact.

Proof: left as an exercise. ■

We now present the section's main result, a "special case" of which we saw in Theorem 20.

Theorem 119 (Bolzano-Weierstrass Compactness)
Let (E,d) be a metric space. Then E is compact if and only if any sequence in E has a convergent sub-sequence in E.

Proof: assume E is compact and let $(\mathbf{x}_n) \subseteq E$. If the range of $(\mathbf{x}_n)$ is finite, there is a constant subsequence which would then automatically be convergent. We thus consider sequences with infinite range $A = \{\mathbf{x}_n \mid n \in \mathbb{N}\}$.

We show that such an A has at least one cluster point. Suppose, instead, that there A has no cluster point. Thus for any $\mathbf{x} \in E$, $\exists r_{\mathbf{x}} > 0$ with $B(\mathbf{x}, r_{\mathbf{x}}) \cap A$ is finite. Since E is compact, there exists a finite $J \subseteq E$ such that $E = \bigcup_{\mathbf{x} \in J} B(\mathbf{x}, r_{\mathbf{x}})$.

Then
$$A = \bigcup_{\mathbf{x} \in J} (B(\mathbf{x}, r_{\mathbf{x}}) \cap A)$$
is a finite union of finite sets, hence A is itself finite.

But this contradicts the fact that A is infinite. Hence, A has at least one cluster point $\mathbf{x} \in E$. Such a cluster point is a limit point of $(\mathbf{x}_n)$: consequently, there is a subsequence of $(\mathbf{x}_n)$ which converges to $\mathbf{x} \in E$ (in which case we say that E satisfies the **Bolzano-Weierstrass property**).

Conversely, assume all sequences in E have convergent subsequence in E. First, note that any metric space (E,d) satisfying the Bolzano-Weierstrass property is precompact. Indeed, suppose that $\exists \varepsilon > 0$ such that E can not be covered with a finite number of $\varepsilon-$balls. Let $\mathbf{x}_0 \in E$. By assumption, $B(\mathbf{x}_0, \varepsilon) \neq E$. Thus $\exists \mathbf{x}_1 \in E$ such that $d(\mathbf{x}_0, \mathbf{x}_1) \geq \varepsilon$.

Since $B(\mathbf{x}_0, \varepsilon) \cup B(\mathbf{x}_1, \varepsilon) \neq E$, $\exists \mathbf{x}_2 \in E$ such that $d(\mathbf{x}_0, \mathbf{x}_1), d(\mathbf{x}_0, \mathbf{x}_2) \geq \varepsilon$. Continuing this process, we build a list $\mathbf{x}_0, \mathbf{x}_1, \ldots, \mathbf{x}_n$ for which $d(\mathbf{x}_i, \mathbf{x}_j) \geq \varepsilon$ for all $i < j \leq n$.

Since $\bigcup_{i=0}^n B(\mathbf{x}_i, \varepsilon) \neq E$, $\exists \mathbf{x}_{n+1} \in E$ such that $d(\mathbf{x}_i, \mathbf{x}_{n+1}) \geq \varepsilon$ for all $0 \leq i \leq n$. By induction, there is a sequence $(\mathbf{x}_n) \subseteq E$ such that $d(\mathbf{x}_i, \mathbf{x}_j) \geq \varepsilon$ whenever $i \neq j$. Consequently, this sequence has no convergent subsequence, since no subsequence is a Cauchy sequence. This contradicts the hypothesis that E satisfies the Bolzano-Weierstrass property, thus E is precompact.

Next, we show that if the metric space (E,d) satisfies the Bolzano-Weierstrass property and if $\{U_i\}_{i \in I}$ is an open cover of E, then
$$\exists \alpha > 0, \forall \mathbf{x} \in E, \exists i \in I \Longrightarrow B(\mathbf{x}, \alpha) \subseteq U_i. \tag{9.1}$$

Indeed, suppose that

$$\forall \alpha > 0, \exists \mathbf{x} \in E, \forall i \in I \implies B(\mathbf{x}, \alpha) \not\subseteq U_i. \tag{9.2}$$

In particular,

$$\forall n \in \mathbb{N}^\times, \exists \mathbf{x}_n \in E, \forall i \in I \implies B(\mathbf{x}, \tfrac{1}{n}) \not\subseteq U_i.$$

Let $(\mathbf{x}_{\varphi(n)})$ be a convergent subsequence of $(\mathbf{x}_n)$ (such a sequence exists since E satisfies the Bolzano-Weierstrass property).

Write $\mathbf{x}_{\varphi(n)} \to \mathbf{x}$. Since $\{U_i\}_{i\in I}$ covers E, $\exists i \in I$ such that $\mathbf{x} \in U_i$. But $U_i \subseteq_O E$, so $\exists r > 0$ such that $B(\mathbf{x}, 2r) \subseteq U_i$.

Accordingly, $\exists N \in \mathbb{N}$ such that $d(\mathbf{x}_{\varphi(n)}, \mathbf{x}) < r$ and $\varphi(n) > \frac{1}{r}$ for all $n > N$. Consequently, $\forall n > N$ and $\forall \mathbf{y} \in B(\mathbf{x}_{\varphi(n)}, \frac{1}{\varphi(n)})$, we have

$$d(\mathbf{x}, \mathbf{y}) \leq d(\mathbf{x}, \mathbf{x}_{\varphi(n)}) + d(\mathbf{x}_{\varphi(n)}, \mathbf{y}) < r + r = 2r.$$

Thus $\forall n > N$, $B(\mathbf{x}_{\varphi(n)}, \frac{1}{\varphi(n)}) \subseteq U_i$, which contradicts (9.2), and so (9.1) holds.

To show E is compact, let $\{U_i\}_{i\in I}$ be an open cover of E. We know from (9.1) that

$$\exists \alpha > 0, \forall \mathbf{x} \in E, \exists i \in I \implies B(\mathbf{x}, \alpha) \subseteq U_i.$$

But E is precompact, so $\exists \mathbf{x}_1, \ldots, \mathbf{x}_n \in E$ such that $E = \bigcup_{j=1}^n B(\mathbf{x}_j, \alpha)$.

Let $i_1, \ldots, i_n$ be the indices for which $B(\mathbf{x}_j, \alpha) \subseteq U_{i_j}$, $1 \leq j \leq n$. Then $E = \bigcup_{j=1}^n U_{i_j}$ is a finite subcover of E; E is indeed compact. ■

The following result has a similar flavour.

Theorem 120
Let (E, d) be a metric space. Then E is compact if and only if any sequence in E has a limit point if and only if every infinite subset of E has a cluster point.

Proof: left as an exercise. ■

It is usually easier to show that the Bolzano-Weierstrass is violated than to show that it holds.

Example: Show that the set $(0, 1)$ is not a compact subset of $(\mathbb{R}, d_1)$.

Proof: Consider the sequence $(1/n) \subseteq (0, 1)$. Every subsequence of $(1/n)$ converges to $0 \notin (0, 1)$. According to Theorem 119, $(0, 1)$ is not a compact subset of $(\mathbb{R}, d_1)$. ■

Compact sets really have quite useful properties.

Proposition 121
Let (E,d) be a metric space.

1. *If E is compact and $A \subseteq_C E$, then $A \subseteq_K E$.*
2. *If $A \subseteq_K E$, then $A \subseteq_C E$ and A is bounded.*

Proof:

1. Since E is compact, it is precompact (see the proof of Theorem 119) and so is A. The set E is also complete (see exercise 2). Thus A is a closed subset of the complete set E: A is then complete (see Proposition 110). But A is precompact and complete, and so $A \subseteq_K E$ (see exercise 3).

2. Since $A \subseteq_K E$, it is precompact. Hence for $\varepsilon > 0$, $\exists \mathbf{x}_1, \ldots, \mathbf{x}_n \in A$ such that

$$A \subseteq \bigcup_{j=1}^{n} B(\mathbf{x}_j, \varepsilon).$$

Thus, $\delta(A) \leq n\varepsilon < \infty$ and A is bounded.

To show that $A \subseteq_C E$, it suffices to show that any sequence in A which converges does so in A, according to Proposition 105. So let $(\mathbf{x}_n) \subseteq A$ be such that $\mathbf{x}_n \to \mathbf{x} \in E$. But A is compact, so that $\exists$ a convergent subsequence $(\mathbf{x}_{\varphi(n)})$ which converges in A. Since any subsequence of a sequence converging to $\mathbf{x}$ also converges to $\mathbf{x}$, $\mathbf{x}_{\varphi(n)} \to \mathbf{x} \in A$ and so $A \subseteq_C E$. ■

Unlike completeness, compactness is a **topological notion**.

Proposition 122
Let (E,d) and (F,δ) be metric spaces, together with a continuous function $f : (E,d) \to (F,\delta)$. If $A \subseteq_K E$ then $f(A) \subseteq_K F$.

Proof: let $\{U_\lambda\}_{\lambda\in\Lambda}$ be an open cover of $f(A)$. Since f is continuous, we have that $A \cap f^{-1}(U_\lambda) \subseteq_O A$ for all $\lambda \in \Lambda$. Thus $\{A \cap f^{-1}(U_\lambda)\}_{\lambda\in\Lambda}$ is an open cover of A. But $A \subseteq_K E$ so that $\exists$ a finite $H \subseteq \Lambda$ such that

$$\bigcup_{\lambda\in H} \left(A \cap f^{-1}(U_\lambda)\right) = A.$$

As such, $\{f(U_\lambda)\}_{\lambda\in H}$ is a finite sub-cover of $f(A)$, and so $f(A) \subseteq_K F$. ■

There is also a link with homeomorphisms.

Proposition 123
Let $f : (E, d) \to (F, \delta)$ be a continuous bijection. If (E, d) is compact, then f is a homeomorphism.

Proof: let $Y \subseteq_C E$. We need to show that $f(Y) \subseteq_C F$. According to Proposition 122, $f(Y) \subseteq_K F$. But, according to Proposition 121, part 2, $f(Y) \subseteq_C F$. So f is closed, meaning that f^{inv} is continuous. ■

Perhaps the most famous theorem linking continuous functions and compact spaces is the result to which we were alluding to at the start of this section (we proved a restricted case in Theorem 33).

Proposition 124 (MAX/MIN THEOREM (REPRISE))
Let $f : (E, d) \to \mathbb{R}$ be continuous. If (E, d) is compact, then f is bounded and $\exists \mathbf{a}, \mathbf{b} \in E$ such that $f(\mathbf{a}) = \inf_{\mathbf{x}\in E} f(\mathbf{x})$ and $f(\mathbf{b}) = \sup_{\mathbf{x}\in E} f(\mathbf{x})$.

Proof: since E is compact and f is continuous, then $f(E)$ is compact according to Proposition 122. As such, $f(E)$ is both closed and bounded in $\mathbb{R}$, according to Proposition 121.

Now, set $A = \inf_{\mathbf{x}\in E} f(\mathbf{x})$. By definition, for each $n \geq 1$, $\exists \mathbf{a}_n \in E$ such that $A \leq f(\mathbf{a}_n) < A + \frac{1}{n}$ (otherwise $\inf_{\mathbf{x}\in E} f(\mathbf{x}) \geq A + \frac{1}{n} > A$).

But $(\mathbf{a}_n)$ is a subsequence of the compact space E (hence a subsequence of a closed space) so $\exists$ a subsequence $(\mathbf{a}_{\varphi(n)})$ which converges to some $\mathbf{a} \in A$ according to Proposition 105.

As f is continuous, $f(\mathbf{a}_{\varphi(n)}) \to f(\mathbf{a})$. But $f(\mathbf{a}_{\varphi(n)}) \to A$, since

$$A \leq f(\mathbf{a}_{\varphi(n)}) < A + \frac{1}{\varphi(n)} \to A.$$

The limit of a convergent sequence is unique in a metric space, so $f(\mathbf{a}) = A$.

A similar argument shows $\exists \mathbf{b} \in E$ such that $f(\mathbf{b}) = \sup_{\mathbf{x}\in E} f(\mathbf{x})$. ■

The next result is often used as the definition of a compact set, but it cannot be generalized to infinite dimensional spaces (such as $\ell^2(\mathbb{N})$ or other infinite dimensional Banach spaces).

Proposition 125 (HEINE-BOREL)
Any closed bounded subset of $\mathbb{K}^n$ is compact in the usual topology.

Proof: since $\mathbb{C}^m \simeq \mathbb{R}^{2m}$, we only need to verify that this is the case for $\mathbb{R}^n$. Furthermore, the proposition will be established if we can show it to be valid for any $A = [a_1, b_1] \times \cdots \times [a_n, b_n] \subseteq_C \mathbb{R}^n$ (why is that the case?).

Since $\mathbb{R}^n$ is complete and $A \subseteq \mathbb{R}^n$, then A is a complete subset of $\mathbb{R}^n$, according to Proposition 110. It will then be sufficient to show that A is precompact, according to the proof of Theorem 119.

But that is obvious (see exercise 5). ■

9.2 Connected Spaces

Let $f : A \subseteq \mathbb{R} \to \mathbb{R}$ be a continuous function such that $\exists a, b \in A$ with $f(a)f(b) < 0$. What condition do we need on A in order to guarantee the existence of a solution to $f(x) = 0$ on A?

Whether A is compact or not is irrelevant: for instance, in the standard topology, the function $f : A = [0, 1] \cup [2, 3] \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} -1 & x \in [0, 1] \\ 1 & x \in [2, 3] \end{cases}$$

is continuous over the compact set A, there are points $a, b \in A$ such that $f(a)f(b) < 0$, yet $f(x) \neq 0$ for all $x \in A$. On the other hand, $f : A = [-1, 1] \to \mathbb{R}$ defined by $f(x) = x$ is such that $f(-1)f(1) < 0$ and $\exists x \in A$ such that $f(x) = 0$ (namely, $x = 0$).

The key notion is that of **connectedness**. Let (E, d) be a metric space. A **partition** of E is a collection of two disjoint non-empty subsets $U, V \subseteq E$ such that $E = U \cup V$.[1] An **open partition** of E is a partition where $U, V \subseteq_O E$; a **closed partition** of E is a partition where $U, V \subseteq_C E$.

Examples

1. There are many partitions of $\mathbb{R}$ in the usual topology, such as

$$(-\infty, 0] \sqcup (0, \infty) \quad \text{or} \quad [(-\infty, -3] \cup \{0\}] \sqcup [(-3, 0) \cup (0, \infty)],$$

   but no such partition can be an open partition or a closed partition. □

2. The metric space $A = [0, 1] \cup [2, 3]$ is partitioned by $[0, 1]$ and $[2, 3]$. This is both an open partition and a closed partition in the usual **subspace** topology (note that this is not the case in $\mathbb{R}$, but we are only interested in the set A, not the space in which it is embedded). □

3. The singleton set $E = \{*\}$ cannot be partitioned. □

[1]We denote the disjoint union by $E = U \sqcup V$.

The next result establishes an "easy" way to determine if a space has such partitions.

Proposition 126
Let (E,d) be a metric space. The following conditions are equivalent:

1. *E has no open partition;*
2. *E has no closed partition;*
3. *The only subsets of E that are both open and closed are $\varnothing$ and E (such sets are rather unfortunately known as **clopen sets**).*

Proof: we show $1. \implies 2. \implies 3. \implies 1.$

$1. \implies 2.$: Suppose that $\{F_1, F_2\}$ forms a closed partition of E. Then $F_i = E \setminus F_{i-1} \subseteq_O E$ for $i = 1, 2$. Hence $\{F_1, F_2\}$ also forms an open partition of E, which contradicts the hypothesis that no such partition of E exists. Thus E has no closed partition.

$2. \implies 3.$: Let $A \subseteq E$ be such that $A \subseteq_C E$ and $A \subseteq_O E$. Then $\{A, E \setminus A\}$ is a closed partition of E. By hypothesis, there can be no such partition of E. Hence $A = \varnothing$ or $E \setminus A = \varnothing$.

$3. \implies 1.$: This is clear once one realizes that any open partition is automatically also a closed partition. ■

A metric space (E,d) is said to be **connected** if it satisfies any of the conditions listed in Proposition 126. Similarly, a subset $A \subseteq E$ is **connected** if its only clopen partition is trivial, that is: whenever $A = X \sqcup Y$, $X, Y \subseteq_O E$, either $X = \varnothing$ or $Y = \varnothing$. We will denote such a situation with $A \subseteq_{©} E$ (this is emphatically not a notation you will find anywhere else).

Examples

1. In the usual topology, $\mathbb{R}$ is connected. □
2. In the same topology, $A = [0,1] \cup [2,3]$ is not a connected subspace of $\mathbb{R}$. □
3. The singleton set $E = \{*\}$ is vacuously connected. □
4. Is $A = \{\frac{1}{n} \mid n \in \mathbb{N}\}$ a connected subset of $\mathbb{R}$ in the usual topology?

   Solution: since $A = \{1\} \sqcup \{\frac{1}{n} \mid n \geq 2\}$ is a non-trivial open partition of A, A is not a connected subset of $\mathbb{R}$ in the usual topology. Indeed, $\{1\} \subseteq_O A$ since $\{1\} = (\frac{1}{2}, \infty) \cap A$, $\{\frac{1}{n} \mid n \geq 2\} \subseteq_O A$ since $\{\frac{1}{n} \mid n \geq 2\} = (0,1) \cap A$. □

As was the case with compactness, connectedness is a **topological notion**.

Proposition 127
Let $f:(E,d)\to(F,\delta)$ be continuous. If $A\subseteq_{©} E$, then $f(A)\subseteq_{©} F$.

Proof: Let $B\subseteq_{O,C} f(A)$. We will show that $B=\varnothing$ or $B=f(A)$.

Since $B\subseteq_O f(A)$, then $\exists U\subseteq_O F$ such that $B=f(A)\cap U$. Similarly, since $B\subseteq_C f(A)$, then $\exists W\subseteq_C F$ such that $B=f(A)\cap W$. But f is continuous so $f^{-1}(U)\subseteq_O E$ and $f^{-1}(W)\subseteq_C E$. Therefore,

$$f^{-1}(B)=A\cap f^{-1}(U)\subseteq_O A \quad\text{and}\quad f^{-1}(B)=A\cap f^{-1}(W)\subseteq_C A.$$

Thus $f^{-1}(B)\subseteq_{O,C} A$. However A is a connected subset of E, so either $f^{-1}(B)=\varnothing$ or $f^{-1}(B)=A$. Since $B\subseteq f(A)$, that leaves only two possibilities: $B=\varnothing$ or $B=f(A)$, which means $f(A)\subseteq_{©} B$. ■

9.2.1 Characterization of Connected Spaces

We now give a simple necessary and sufficient condition for connectedness. Throughout, we endow the set $\{0,1\}$ with the discrete metric.

Proposition 128
A metric space (E,d) is connected if and only if every continuous function $f:E\to\{0,1\}$ is constant.

Proof: assume (E,d) is connected. If $f:E\to\{0,1\}$ is continuous and not constant, then $f^{-1}(0),f^{-1}(1)\subseteq_{O,C} E$ and $E=f^{-1}(0)\sqcup f^{-1}(1)$.

Since f is not constant, neither $f^{-1}(0)$ nor $f^{-1}(1)$ is $\varnothing$ or all of E. Hence E is not connected, as it contains non-trivial clopens, which contradicts our starting assumption. Thus f is constant.

Conversely, if E is not connected, $\exists$ non-trivial clopens X,Y such that $E=X\sqcup Y$. Consider the characteristic function $\chi_X:E\to\{0,1\}$: we have $f^{-1}(0)=Y\subseteq_O E$ and $f^{-1}(1)=X\subseteq_O E$. Consequently, f is continuous and clearly not constant. ■

In practice, Proposition 128 is typically easier to use to show that a space is **not connected**.

Proposition 129
Let (E,d) be a metric space and $A\subseteq_{©} E$. If $B\subseteq E$ is such that $A\subseteq B\subseteq\overline{A}$, then $B\subseteq_{©} E$.

Proof: if such a B is not connected, then $\exists$ a non-trivial open partition $\{X, Y\}$ of B. In particular, $\{A \cap X, A \cap Y\}$ is an open (in A) partition of A. But A is dense in B: if $\mathbf{x} \in B$, every neighbourhood around $\mathbf{x}$ contains at least a point of A.

In particular, if $\mathbf{x} \in B \cap X$, then any neighbourhood around $\mathbf{x}$ must contain at least a point of $A \cap X$. Consequently, $A \cap X \neq \varnothing$. Similarly, $A \cap Y \neq \varnothing$.

Thus, $\{A \cap X, A \cap Y\}$ is a non-trivial open partition of A, which contradicts the fact that A is connected. So B must be connected. ■

There is a series of other useful propositions about connected spaces.

Proposition 130
If $(B_i)_{i\in I}$ is a family of connected subsets of a metric space (E, d) such that $\bigcap_{i\in I} B_i \neq \varnothing$, then $B = \bigcup_{i\in I} B_i \subseteq_© E$.

Proof: if $\{X, Y\}$ is a non-trivial open partition of B and if $\mathbf{b} \in \bigcap_{i\in I} B_i$, we may assume $\mathbf{b} \in X$ without loss of generality. But $B = \bigcup_{i\in I} = X \sqcup Y$ and $Y \neq \varnothing$; hence $\exists i_0 \in I$ such that $Y \cap B_{i_0} \neq \varnothing$.

Since $\mathbf{b} \in \bigcap_{i\in I} B_i$, then $\mathbf{b} \in X \cap B_{i_0} \neq \varnothing$ and so $\{X \cap B_{i_0}, Y \cap B_{i_0}\}$ is a non-trivial open partition of B_{i_0}, which contradicts the hypothesis that $B_{i_0} \subseteq_© E$. Consequently, $B \subseteq_© E$. ■

Proposition 131
If $(C_n)_{n\in\mathbb{N}}$ is a sequence of connected subsets of a metric space (E, d) such that $C_{n-1} \cap C_n \neq \varnothing$, then $C = \bigcup_{n\in\mathbb{N}} C_n \subseteq_© E$.

Proof: left as an exercise. ■

Proposition 132
Let $(E_1, d_1), \dots, (E_n, d_n)$ be metric spaces. Then

$$(E, d) = (E_1 \times \cdots \times E_n, \sup\{d_i \mid 1 \leq i \leq n\})$$

is connected if and only if (E_i, d_i) is connected for all i.

Proof: left as an exercise. ■

Let (E, d) be a metric space once more. We define an equivalence relation on E as follows:

$$\mathbf{x} R \mathbf{y} \iff \exists C \subseteq_© E \text{ such that } \mathbf{x}, \mathbf{y} \in C. \tag{9.3}$$

The equivalence class

$$[\mathbf{x}] = \{\mathbf{y} \in E \mid \mathbf{y}R\mathbf{x}\} = \bigcup_{\substack{C \subseteq_{©} E \\ \mathbf{x} \in C}} C$$

is a connected subset of E, which we call the **connected component** of **x**. It is not difficult to show that $[\mathbf{x}] \subseteq_C E$ and that if a metric space only has a finite number of connected components, then each of those components is a clopen subset of E (see exercises 10 and 11).

Proposition 133
Consider $\mathbb{R}$ with the usual topology. Then, $A \subseteq_{©} \mathbb{R}$ if and only if A is an interval.

Proof: let $A \subseteq_{©} \mathbb{R}$. If A is not an interval, $\exists a, b \in A$ for which $\exists c \in (a, b)$ with $c \notin A$. Thus, $A \subseteq (-\infty, c) \cup (c, \infty)$.

Hence $\{A \cap (-\infty, c), A \cap (c, \infty)\}$ is a non-trivial open partition of A, which implies that A is not a connected subset of $\mathbb{R}$, a contradiction as $A \subseteq_{©} E$, and so A is an interval.

Conversely, if $A = \{*\}$, we have already shown that $A \subseteq_{©} \mathbb{R}$. According to Proposition 129, it is sufficient to verify that $A = (a, b) \subseteq_{©} \mathbb{R}$ for any $a < b$. We will show that any continuous map $f : (a, b) \to \{0, 1\}$ is constant.

Suppose otherwise that $\exists x, y \in (a, b)$ such that $x < y$ and $f(x) \neq f(y)$. Without loss of generality, let $f(x) = 0$ and $f(y) = 1$. Set

$$\Gamma = \{z \mid z \geq x \text{ and } f(t) = 0 \, \forall t \in [x, z]\}.$$

Clearly, $\Gamma \neq \varnothing$ since $x \in \Gamma$. Furthermore Γ is bounded above by y. Thus, since $\mathbb{R}$ is complete, $\exists c \in [x, y] \subseteq (a, b)$ such that $c = \sup \Gamma$.

By continuity of f at c, $f(c) = 0$ and $\exists \delta > 0$ such that

$$s \in (c - \delta, c + \delta) \implies |f(s)| = |f(s) - f(c)| < \tfrac{1}{2}.$$

As such, $f(s) < \frac{1}{2}$ for all $s \in (c - \delta, c + \delta)$. But f can only take two values: 0 or 1. Consequently, $f(s) = 0$ for all $s \in (c - \delta, c + \delta)$.

This in turn implies that $c + \frac{\delta}{2} \in \Gamma$, which contradicts the fact that $c = \sup \Gamma$. Thus, f is constant, and $(a, b) \subseteq_{©} \mathbb{R}$. ■

We can now give a proof of the remark made after Theorem 36.

Corollary 134 (Bolzano's Theorem)
Consider $\mathbb{R}$ with the usual topology and a continuous function $f : \mathbb{R} \to \mathbb{R}$. The image of any interval by f is an interval.

Proof: let $A \subseteq_{©} \mathbb{R}$. By the preceding proposition, A is an interval. Since f is continuous, $f(A) \subseteq_{©} \mathbb{R}$. But the only connected subsets of $\mathbb{R}$ are the intervals. Consequently, $f(A)$ is an interval. ■

9.2.2 Path-Connected Spaces

We can also define other types of connectedness.

Let (E, d) be a metric space. We say that E is **path-connected** if for any two points $\mathbf{x}, \mathbf{y} \in E$, there is a continuous function $\gamma : [0, 1] \to E$ such that $\gamma(0) = \mathbf{x}$ and $\gamma(1) = \mathbf{y}$. The **segment between x and y** is

$$[\mathbf{x}, \mathbf{y}] = \{t\mathbf{x} + (1 - t)\mathbf{y} \mid t \in [0, 1]\}.$$

The continuous function associated to this segment is the function

$$f_{\mathbf{x},\mathbf{y}} : [0, 1] \to E, \quad \text{defined by} \quad f_{\mathbf{x},\mathbf{y}}(t) = t\mathbf{x} + (1 - t)\mathbf{y}.$$

If $[\mathbf{x}, \mathbf{y}]$ and $[\mathbf{z}, \mathbf{w}]$ are two segments, define their **sum** (concatenation) to be

$$[\mathbf{x}, \mathbf{y}] + [\mathbf{z}, \mathbf{w}] = \{2t\mathbf{x} + (1 - 2t)\mathbf{y} \mid t \in [0, \tfrac{1}{2}]\} \cup \{(2t - 1)\mathbf{z} + (2 - 2t)\mathbf{w} \mid t \in [\tfrac{1}{2}, 1]\}.$$

If $\mathbf{y} = \mathbf{z}$, the continuous function associated to this sum is the function

$$g_{\mathbf{x},\mathbf{y},\mathbf{w}} : [0, 1] \to E, \quad \text{defined by} \quad g_{\mathbf{x},\mathbf{y},\mathbf{w}}(t) = \begin{cases} 2t\mathbf{x} + (1 - 2t)\mathbf{y} & \text{if } t \in [0, \frac{1}{2}] \\ (2t - 1)\mathbf{y} + (2 - 2t)\mathbf{w} & \text{if } t \in [\frac{1}{2}, 1] \end{cases}$$

Examples

1. Show that $B(\mathbf{0}, 1)$ is path-connected in $(\mathbb{R}^2, d_2)$.

   Proof: Let $\mathbf{a} \neq \mathbf{b} \in B(\mathbf{0}, 1)$. Then $[\mathbf{a}, \mathbf{0}], [\mathbf{0}, \mathbf{b}] \subseteq B(\mathbf{0}, 1)$. Indeed, if $\mathbf{x} \in [\mathbf{a}, \mathbf{0}]$, then $\mathbf{x} = t\mathbf{a}$ for $t \in [0, 1]$. But $\|\mathbf{x}\| = |t|\|\mathbf{a}\| \leq \|\mathbf{a}\| < 1$, so that $\mathbf{x} \in B(\mathbf{0}, 1)$. Then $g_{\mathbf{a},\mathbf{0},\mathbf{b}} \in C_{B(\mathbf{0},1)}([0, 1])$ is such that $g_{\mathbf{a},\mathbf{0},\mathbf{b}}(0) = \mathbf{a}$ and $g_{\mathbf{a},\mathbf{0},\mathbf{b}}(1) = \mathbf{b}$. ■

2. In any normed vector space $(E, \|\cdot\|)$ over $\mathbb{K}$, any open ball $B(\mathbf{x}, \rho)$ is path-connected (see exercise 13). □

There is clearly a link between the two connectedness definitions.

Proposition 135
If (E, d) is path-connected, then it is also connected.

Proof: let $f : E \to \{0,1\}$ be a continuous function and $\mathbf{a}, \mathbf{b} \in E$. Since E is path-connected, $\exists$ a continuous path $\gamma : [0,1] \to \mathbb{R}$ such that $\gamma(0) = \mathbf{a}$ and $\gamma(1) = \mathbf{b}$.

Since the composition $f \circ \gamma : [0,1] \to \{0,1\}$ is continuous and since $[0,1] \subseteq_{©} \mathbb{R}$, then $f \circ \gamma$ is constant: in particular,

$$f(\mathbf{a}) = f(\gamma(0)) = f(\gamma(1)) = f(\mathbf{b}),$$

so that f itself is constant. Consequently, E is connected. ■

If $E = (\mathbb{K}^n, d_{\text{Euclidean}})$, the converse is also true.

Proposition 136
If $A \subseteq_{©} \mathbb{K}^n$ in the usual topology, then A is path-connected.

Proof: left as an exercise. ■

But connected spaces are not path-connected, in general (see exercise 22, for instance). The following result will allow us to segue gently into Chapter 10.

Theorem 137
Let $(E, \|\cdot\|)$ be a normed vector space over $\mathbb{K}$. Then any $A \subseteq_{O,©} E$ is path-connected.

Proof: Let $\mathbf{x}_0 \in A$ and set

$$F_{\mathbf{x}_0} = \{\mathbf{x} \in A \mid \exists \gamma \in C_E([0,1]) \text{ such that } \gamma(0) = \mathbf{x}_0, \gamma(1) = \mathbf{x}\}.$$

We need to show that $F_{\mathbf{x}_0} = A$. In order to do so, note that $F_{\mathbf{x}_0} \neq \varnothing$ as $\mathbf{x}_0 \in F_{\mathbf{x}_0}$. If we can show that $F_{\mathbf{x}_0} \subseteq_{O,C} A$, then we are done as $A \subseteq_{©} E$.

- Let $\mathbf{x} \in F_{\mathbf{x}_0} \subseteq A$. Since $A \subseteq_O E$, $\exists \rho > 0$ such that $B(\mathbf{x}, \rho) \subseteq A$. For any $\mathbf{y} \in B(\mathbf{x}, \rho)$, $[\mathbf{y}, \mathbf{x}] \in B(x, \rho)$ (modify the proof of exercise 13). Since $\mathbf{x}_0 \in F_{\mathbf{x}_0}$, $B(\mathbf{x}, \rho) \subseteq F_{\mathbf{x}_0}$. Consequently, $F_{\mathbf{x}_0} \subseteq_O A$.

- If $\mathbf{x} \in \overline{F_{\mathbf{x}_0}} \cap A$, then for any $\rho > 0$ we have $B(\mathbf{x}, \rho) \cap F_{\mathbf{x}_0} \neq \varnothing$. Since $A \subseteq_O E$, $\exists \rho_0 > 0$ such that $B(\mathbf{x}, \rho_0) \subseteq A$; in particular $\varnothing \neq B(\mathbf{x}, \rho_0) \cap F_{\mathbf{x}_0} \subseteq A$. Now, let $\mathbf{y} \in B(\mathbf{x}, \rho_0) \cap F_{\mathbf{x}_0}$. Since $[\mathbf{y}, \mathbf{x}] \subseteq B(\mathbf{x}, \rho_0)$, there is a continuous path in A from $\mathbf{y}$ to $\mathbf{x}$. Since $\mathbf{y} \in F_{\mathbf{x}_0}$, there is a continuous path in A from $\mathbf{x}_0$ to $\mathbf{y}$. Combining these paths, there is a continuous path in A from $\mathbf{x}_0$ to $\mathbf{x}$. Hence, $\mathbf{x} \in F_{\mathbf{x}_0}$. Consequently, $F_{\mathbf{x}_0} \subseteq_C A$.

This concludes the proof. ■

Finally, we note that path-connectedness is a **topological notion**.

> **Proposition 138** Let $f : (E, d) \to (F, \delta)$ be a continuous map. If E is path-connected, then $f(E)$ is path-connected. **Proof:** left as an exercise. ■

9.3 Solved Problems

1. Let (E, d) be a metric space.

   a) If $W_1, W_2 \subseteq_K E$, show that $\exists \mathbf{x}_i \in W_i$ such that $d(\mathbf{x}_1, \mathbf{x}_2) = d(W_1, W_2)$.

   b) If $W \subseteq_K E$ and $F \subseteq_C E$ are such that $W \subseteq F = \varnothing$, show that $d(W, F) \neq 0$. Is the conclusion still valid when $W \subseteq_C E$ is not necessarily compact?

   Proof:

   a) The mapping $\varphi : K_1 \to \mathbb{R}$ defined by $\varphi(\mathbf{x}) = d(\mathbf{x}, K_2)$ is continuous. Since K_1 is compact, the Max/Min Theorem applies: $\exists \mathbf{x}_1 \in K_1$ such that

   $$\varphi(\mathbf{x}_1) = d(\mathbf{x}_1, K_2) = \inf_{\mathbf{x} \in K_1} \{d(\mathbf{x}, K_2)\} = d(K_1, K_2).$$

   Similarly, the mapping $\eta : K_2 \to \mathbb{R}$ defined by $\eta(\mathbf{y}) = d(\mathbf{x}_1, \mathbf{y})$ is continuous on a compact set: as such, $\exists x_2 \in K_2$ such that

   $$\eta(\mathbf{x}_2) = d(\mathbf{x}_1, \mathbf{x}_2) = \inf_{\mathbf{y} \in K_2} \{d(\mathbf{x}_1, K_2)\} = d(K_1, K_2).$$

   b) The mapping $\theta : K \to \mathbb{R}$ defined by $\theta(\mathbf{x}) = d(\mathbf{x}, F)$ is continuous on the compact K so that $\exists \mathbf{x}_0 \in K$ such that

   $$\theta(\mathbf{x}_0) = d(\mathbf{x}_0, F) = \inf_{\mathbf{x} \in K} \{d(\mathbf{x}, F)\} = d(K, F).$$

   If $d(\mathbf{x}_0, F) = 0$ then $\mathbf{x}_0 \in F$ since F is closed. But that is impossible as $K \cap F = \varnothing$ and so $d(\mathbf{x}_0, F) \neq 0$.

   If K is only assumed closed, the conclusion may not hold. For instance in $\mathbb{R}^2$, the sets $K = \{(x, y) \mid y \leq 0\}$ and $F = \{(x, y) \mid y \geq e^x\}$ are closed and disjoints, yet $d(K, F) = 0$. ■

2. Let $(E, d) = (\mathbb{R}^n, d_2)$.

   a) If $F \subseteq_C E$ is unbounded and $f : F \to \mathbb{R}$ is a continuous map such that

   $$\lim_{\|\mathbf{x}\| \to \infty} f(\mathbf{x}) = +\infty, \qquad \mathbf{x} \in F,$$

   show $\exists \mathbf{x} \in F$ such that $f(\mathbf{x}) = \inf_{\mathbf{y} \in F} f(\mathbf{y})$.

   b) If $W \subseteq_K E$ and $F \subseteq_C E$, show $\exists \mathbf{x} \in W, \mathbf{y} \in F$ such that $d(\mathbf{x}, \mathbf{y}) = d(W, F)$. Is the conclusion still valid when E is an infinite-dimensional vector space over $\mathbb{R}$?

Proof:

a) Fix $\mathbf{a} \in F$ and consider the set $\Gamma = \{\mathbf{x} \in F \mid f(\mathbf{x}) \leq f(\mathbf{a})\}$. Since f is continuous, $\Gamma = f^{-1}((-\infty, f(a)]) \subseteq_C F$ and so $\Gamma \subseteq_C E$. It is also bounded since

$$\lim_{\|\mathbf{x}\| \to \infty} f(\mathbf{x}) = +\infty, \qquad \mathbf{x} \in F.$$

Thus $\Gamma \subseteq_K \mathbb{R}^n$ by the Heine-Borel Theorem. Furthermore, $\Gamma \neq \varnothing$ since $\mathbf{a} \in \Gamma$. According to the Max/Min Theorem, $\exists \mathbf{x} \in \Gamma$ such that $f(\mathbf{x}) = \inf_{\mathbf{y} \in \Gamma}\{f(\mathbf{y})\}$. By construction,

$$\inf_{\mathbf{y} \in \Gamma}\{f(\mathbf{y})\} = \inf_{\mathbf{y} \in F}\{f(\mathbf{y})\},$$

whence $f(\mathbf{x}) = \inf_{\mathbf{y} \in F}\{f(\mathbf{y})\}$ for some $\mathbf{x} \in F$.

b) Since the mapping $\varphi : K \to \mathbb{R}$ defined by $\varphi(\mathbf{x}) = d(\mathbf{x}, F)$ is continuous, $\exists \mathbf{x} \in K$ such that

$$d(\mathbf{x}, F) = \inf_{\mathbf{y} \in K}\{d(\mathbf{y}, F)\} = d(K, F).$$

Note that the mapping $\psi_{\mathbf{x}} : F \to \mathbb{R}$ defined by $\psi_{\mathbf{x}}(\mathbf{y}) = d(\mathbf{x}, \mathbf{y})$ is also continuous. If F is bounded, then $F \subseteq_K \mathbb{R}^n$ and the desired result is derived from the result in (a).

Otherwise, if F is unbounded we have

$$\lim_{\|\mathbf{y}\| \to \infty} \psi_{\mathbf{x}}(\mathbf{y}) = \infty, \quad \mathbf{y} \in F$$

so that $\exists \mathbf{y} \in F$ such that

$$\psi_{\mathbf{x}}(\mathbf{y}) = \inf_{\mathbf{z} \in F}\{\psi_{\mathbf{x}}(\mathbf{z})\} = d(\mathbf{x}, F) = d(K, F),$$

which proves the desired result.

The result is false in general if E is infinite-dimensional: consider for instance the vector space of bounded sequences in $\mathbb{R}$, with the norm $\|(u_n)\| = \sup_{n \in \mathbb{N}}\{|u_n|\}$.

For any $n \in \mathbb{N}$, let $\mathcal{X}_n$ be the sequence where the n^{th} term is $1 + 2^{-n}$ and all the other terms are 0. The set $F = \{\mathcal{X}_n \mid n \in \mathbb{N}\}$ is closed in E since all its points are isolated points. If $K = \{\mathbf{0}\}$, it is obvious that $d(K, F) = 1$, yet $d(K, \mathcal{X}_n) = 1 + 2^{-n} > 1$ for all $n \in \mathbb{N}$. ∎

3. Let (E, d) be a compact metric space with a map $f : E \to E$ such that $\forall \mathbf{x} \neq \mathbf{y} \in E$, $d(f(\mathbf{x}), f(\mathbf{y})) < d(\mathbf{x}, \mathbf{y})$.

   a) Show that f admits a unique fixed point $\alpha \in E$.

   b) Let $\mathbf{x}_0 \in E$. For each $n \in \mathbb{N}$, set $\mathbf{x}_{n+1} = f(\mathbf{x}_n)$. Show that $\mathbf{x}_n \to \alpha$.

   c) Are these results still valid if E is complete but not compact?

Proof:

a) First note that, being Lipschitz, f is continuous. Then, the mapping $\varphi_f : E \to \mathbb{R}$ defined by $\varphi_f(\mathbf{x}) = d(\mathbf{x}, f(\mathbf{x}))$ is continuous as it is a composition of continuous functions. But E is compact so that $\exists \alpha \in E$ such that $d(\alpha, f(\alpha)) = \inf_{\mathbf{x}\in E}\{d(\mathbf{x}, f(\mathbf{x}))\}$. If $\alpha \neq f(\alpha) = \beta$, then

$$d(\beta, f(\beta)) = d(f(\alpha), f(\beta)) < d(\alpha, \beta) = d(\alpha, f(\alpha))$$

by hypothesis, which contradicts the definition of α. Thus $\alpha = f(\alpha)$.

Now, suppose $\beta = f(\beta)$ with $\beta \neq \alpha$. Then we have

$$d(f(\alpha), f(\beta)) = d(\alpha, \beta),$$

which contradicts the hypothesis. Thus $\alpha = \beta$.

b) Write $\mathbf{u}_n = d(\alpha, \mathbf{x}_n)$. If $\exists n_0 \in \mathbb{N}$ such that $\mathbf{u}_{n_0} = 0$, then $\mathbf{u}_n = \mathbf{u}_{n_0} = 0$ for all $n \geq n_0$ and the result follows. Otherwise, for all $n \in \mathbb{N}$ we have

$$\mathbf{u}_{n+1} = d(f(\alpha), f(\mathbf{x}_n)) < d(\alpha, \mathbf{x}_n) = \mathbf{u}_n,$$

i.e. $(\mathbf{u}_n)$ is a strictly decreasing sequence. As it is bounded below by 0, it is necessarily convergent. Let $\mathbf{u}_n \to \ell \geq 0$. We need to show $\ell = 0$.

Assume that $\ell > 0$. Since $(\mathbf{u}_n)$ is decreasing, $\mathbf{u}_n \geq \ell$ for all n. Since $(\mathbf{x}_n)$ is a sequence in the compact set E, there is a convergent subsequence $(\mathbf{x}_{\varphi(n)})$, with $\varphi : \mathbb{N} \to \mathbb{N}$ strictly increasing. Let $\beta = \lim \mathbf{x}_{\varphi(n)}$. Then

$$\ell = \lim_{n\to\infty} d(\alpha, \mathbf{x}_{\varphi(n)}) = d(\alpha, \beta).$$

Since f is continuous, we have

$$\lim_{n\to\infty} d(\alpha, f(\mathbf{x}_{\varphi(n)})) = d(\alpha, f(\beta)).$$

But that is impossible since

$$d(\alpha, f(\beta)) = d(f(\alpha), f(\beta)) < d(\alpha, \beta) = \ell$$

and

$$d(\alpha, f(\mathbf{x}_{\varphi(n)})) = d(\alpha, \mathbf{x}_{\varphi(n)+1}) \geq \ell \quad \forall n.$$

The only remaining possibility is thus that $\ell = 0$.

c) Completeness of E is not sufficient. For instance, the function $f : \mathbb{R} \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} 1 & \text{if } x < 0 \\ x + \frac{1}{1+x} & \text{if } x \geq 0 \end{cases}$$

satisfies the hypothesis, but it admits no fixed point. ■

4. Let (E, d) and (F, δ) be two metric spaces, together with a injective map $f : E \to F$. Show that f is continuous if and only if $f(W) \subseteq_K F$ for all $W \subseteq_K E$.

Proof: we already know that if f is continuous and $W \subseteq_K E$, then $f(W) \subseteq_K F$.

Now assume that $f(W) \subseteq_K F$ for all $W \subseteq_K E$. Let $\mathbf{x} \in E$ and $(\mathbf{x}_n) \subseteq E$ be such that $\mathbf{x}_n \to \mathbf{x}$. The set $V = \{\mathbf{x}_n \mid n \in \mathbb{N}\} \cup \{\mathbf{x}\}$ is compact in E, according to the Borel-Lebesgue property. Thus, we have $V' = f(V) \subseteq_K F$.

Let $g : V \to F$ be such that $g = f|_V$. Since f is injective, g is a bijection from V to V'. The map $g^{-1} : V' \to V$ is continuous since any closed subset $W \subseteq_C V$ is automatically compact in V.

As such $(g^{-1})^{-1}(W) = g(W) \subseteq_K V'$ is automatically closed in V'. Since V' is compact, $(g^{-1})^{-1} = g$ is continuous. Thus

$$f(\mathbf{x}_n) = g(\mathbf{x}_n) \to g(\mathbf{x}) = f(\mathbf{x}) \implies f \text{ is continuous.}$$

ote that if f is not injective, the result does not hold in general. For instance, the Heaviside function $f : \mathbb{R} \to \mathbb{R}$ defined by $f(x) = 0$ if $x < 0$ and $f(x) = 1$ if $x \geq 0$ sends any compact set to a compact set, but it is not continuous. ■

5. Let (E, d) be a metric space. If $\varepsilon > 0$, we say that E is $\varepsilon-$**chained** if for all $\mathbf{a}, \mathbf{b} \in E$, $\exists n \in \mathbb{N}^\times$ and $\mathbf{x}_0, \ldots, \mathbf{x}_n \in E$ such that $\mathbf{x}_0 = \mathbf{a}$, $\mathbf{x}_n = \mathbf{b}$ and $d(\mathbf{x}_i, \mathbf{x}_{i-1}) < \varepsilon$ for all $i = 1, \ldots, n$. We say that E is **well-chained** if it is $\varepsilon-$chained for all $\varepsilon > 0$.

   a) If E is connected, show that E is well-chained.

   b) If E is compact and well-chained, show that E is connected. Is the result still true if E is not necessarily compact?

   Proof:

   a) Let $\varepsilon > 0$. We define an equivalence relation $\mathcal{R}_\varepsilon$ on E according to the following: $\mathbf{x}\mathcal{R}_\varepsilon\mathbf{y}$ if and only if $\exists n \in \mathbb{N}^\times$ and $\mathbf{x}_0, \ldots, \mathbf{x}_n \in E$ such that $\mathbf{x}_0 = \mathbf{x}$, $\mathbf{x}_n = \mathbf{y}$ and $d(\mathbf{x}_i, \mathbf{x}_{i-1}) < \varepsilon$ for all $i = 1, \ldots, n$.

      Let $\mathbf{x} \in E$ and $\mathbf{y} \in [\mathbf{x}]$. Then, for all $\mathbf{z} \in B(\mathbf{y}, \varepsilon)$ we have $\mathbf{z} \in [\mathbf{y}] = [\mathbf{x}]$. Thus $B(\mathbf{y}, \varepsilon) \subseteq [\mathbf{x}]$ and so $[\mathbf{x}] \subseteq_O E$.

      Since
      $$[\mathbf{x}] = E \setminus \bigcup_{\mathbf{y} \notin [\mathbf{x}]} [\mathbf{y}]$$
      is the complement of an open set, $[\mathbf{x}] \subseteq_C E$. Consequently, $[\mathbf{x}]$ is a clopen subset of E. But E is connected; we must then have $[\mathbf{x}] = E$ since $[\mathbf{x}] \neq \varnothing$. Hence, every pair of point of E can be joined by an $\varepsilon-$ chain. As ε is arbitrary, E is **well-chained**.

   b) Suppose that E is not connected. Then we can write $E = F_1 \sqcup F_2$, where $\varnothing \neq F_1, F_2 \subseteq_C E$. Since E is compact, $F_1, F_2 \subseteq_K E$.It is left as an exercise to show that $\exists \mathbf{a}_1 \in F_1$ and $\mathbf{a}_2 \in F_2$ such that $d(\mathbf{a}_1, \mathbf{a}_2) = d(F_1, F_2)$.

Since $F_1 \cap F_2 \neq \varnothing$, $\mathbf{a}_1 \neq \mathbf{a}_2$ and so $\varepsilon = d(\mathbf{a}_1, \mathbf{a}_2) > 0$; as such, $d(\mathbf{x}, \mathbf{y}) \geq \varepsilon$ for all $(\mathbf{x}, \mathbf{y}) \in F_1 \times F_2$.

Let $(\mathbf{x}, \mathbf{y})$ be such a point. Since E is well-chained, $\exists$ an $\varepsilon-$chain $(\mathbf{x}_0, \ldots, \mathbf{x}_n) \in E^{n+1}$ such that

$$\mathbf{x}_0 = \mathbf{x},\ \mathbf{x}_n = \mathbf{y} \quad \text{and} \quad d(\mathbf{x}_i, \mathbf{x}_{i-1}) < \varepsilon \quad \text{for all } i = 1, \ldots, n.$$

Since $\mathbf{x}_0 \in F_1$ and $\mathbf{x}_n \in F_2$, $\exists i$ such that $\mathbf{x}_{i-1} \in F_1$ and $\mathbf{x}_i \in F_2$.

But this would imply that $\varepsilon > d(\mathbf{x}_{i-1}, \mathbf{x}_i) \geq d(F_1, F_2) = \varepsilon$, which is a contradiction. Consequently, E is connected.

If E is not compact, the result is not valid in general: Q is well-chained when endowed with the usual metric because it is dense in $\mathbb{R}$, but it is not connected. ■

6. Let (E, d) be a metric space, with two disjoint sets $A, B \subseteq_C E$. Show that there exists a continuous function $f : E \to [0, 1]$ such that $A = f^{-1}(\{0\})$ and $B = f^{-1}(\{1\})$, as well as two disjoint sets $U, V \subseteq_O E$ such that $A \subseteq U$ and $B \subseteq V$.

Proof: Let $F \subseteq_C E$. Define $g_F : (E, d) \to (\mathbb{R}, |\cdot|)$ by

$$g_F(\mathbf{x}) = d(\mathbf{x}, F) = \inf_{\mathbf{y} \in F}\{d(\mathbf{x}, \mathbf{y})\}$$

According to the Triangle Inequality, for all $\mathbf{y} \in F$ we have

$$g_F(\mathbf{x}) = d(\mathbf{x}, F) \leq d(\mathbf{x}, \mathbf{y}) \leq d(\mathbf{x}, \mathbf{z}) + d(\mathbf{z}, \mathbf{y}) \quad \forall \mathbf{x}, \mathbf{z} \in E,$$

thus we must have $g_F(\mathbf{x}) \leq d(\mathbf{x}, \mathbf{z}) + g_F(\mathbf{z})$ for all $\mathbf{x}, \mathbf{z} \in E$, that is, $g_F(\mathbf{x}) - g_F(\mathbf{z}) \leq d(\mathbf{x}, \mathbf{z})$ for all $\mathbf{x}, \mathbf{z} \in E$. In a similar fashion, $g_F(\mathbf{z}) - g_F(\mathbf{x}) \leq d(\mathbf{x}, \mathbf{z})$ for all $\mathbf{x}, \mathbf{z} \in E$. Thus,

$$|g_F(\mathbf{x}) - g_F(\mathbf{z})| \leq d(\mathbf{x}, \mathbf{z}) \quad \text{for all } \mathbf{x}, \mathbf{z} \in E,$$

i.e. g_F is Lipschitz (and so continuous).

Since $F \subseteq_C E$, $g_F(\mathbf{x}) = 0$ if and only if $\mathbf{x} \in F$. Let $f : (E, d) \to (\mathbb{R}, |\cdot|)$ be defined by

$$f(\mathbf{x}) = \frac{g_A(\mathbf{x})}{g_A(\mathbf{x}) + g_B(\mathbf{x})} = \frac{d(\mathbf{x}, A)}{d(\mathbf{x}, A) + d(\mathbf{x}, B)};$$

it is well-defined since whenever $d(\mathbf{x}, A) + d(\mathbf{x}, B) = 0$, we must have $d(\mathbf{x}, A) = d(\mathbf{x}, B) = 0$, i.e. $\mathbf{x} \in A$ and $\mathbf{x} \in B$. But $A \cap B = \varnothing$ and so for all $\mathbf{x} \in E$, we have $d(\mathbf{x}, A) + d(\mathbf{x}, B) \neq 0$.

Furthermore, $f(\mathbf{x}) = 0$ if and only if $d(\mathbf{x}, A) = 0$, i.e. $\mathbf{x} \in A$; $f(\mathbf{x}) = 1$ if and only if $d(\mathbf{x}, B) = 0$, i.e. $\mathbf{x} \in B$.

The function f is continuous since it is the composition of continuous functions. It is clear that $0 \leq f(\mathbf{x}) \leq 1$, so that $f : E \to [0, 1]$.

Finally, let

$$A \subseteq U = f^{-1}([0,1/2)) \subseteq_O [0,1] \quad \text{and} \quad B \subseteq V = f^{-1}((1/2,1]) \subseteq_O [0,1].$$

Then $U \cap V = \varnothing$ by construction and we are done. ■

9.4 Exercises

1. Prepare a 2-page summary of this chapter, with important definitions and results.
2. Show that any compact metric space is precompact and complete.
3. Show that any complete precompact metric space is compact.
4. Prove Theorem 120.
5. With the usual metric, show that $A \subseteq \mathbb{R}^n$ is precompact if and only if $\overline{A} \subseteq_K \mathbb{R}^n$.
6. Prove Proposition 131.
7. Prove Proposition 132.
8. Let $(E_1, d_1), \ldots, (E_n, d_n)$ be metric spaces. Show that

   $$(E, d) = (E_1 \times \cdots \times E_n, \sup\{d_i \mid 1 \leq i \leq n\})$$

   is compact if and only if (E_i, d_i) is compact for all $i = 1, \ldots, n$.[2]
9. Show that (9.3) defines an equivalence relation on a metric space (E, d).
10. Let (E, d) be a metric space and let $\mathbf{x} \in E$. Show that $[\mathbf{x}] \subseteq_C E$.
11. Let (E, d) be a metric space with finitely many connected components. Show that each of those components is a clopen subset of E.
12. Prove Proposition 136.
13. Show that if $(E, \|\cdot\|)$ is a normed vector space over $\mathbb{K}$, then any open ball $B(\mathbf{x}, \rho)$ is path-connected.
14. Let (E, d) be a metric space, $B \subseteq_{©} E$ and $A \subseteq E$ such that

    $$B \cap \text{int}(A) \neq \varnothing \quad \text{and} \quad B \cap \text{int}(E \setminus A) \neq \varnothing.$$

    Show that $B \cap \partial A \neq \varnothing$.

[2]This result cannot be generalized to infinite products (**Tychonoff's Theorem**) without calling upon the **Axiom of Choice**, a.k.a **Zorn's Lemma**, a.k.a. the **Existence of Non-Measurable Sets**, a.k.a. the **Banach-Tarksi Paradox**.

15. Let (A, d_1) and (B, d_2) be two metric spaces. Let $X \subsetneq A$ and $Y \subsetneq B$. Show that

$$(A \times B) \setminus (X \times Y) \subseteq_{©} A \times B.$$

16. Prove Proposition 9.2.2.

17. In the usual topology, give an example of a subset $A \subseteq_{©} \mathbb{R}^2$ for which $\mathrm{int}(A)$ is not connected.

18. In the usual topology, give an example of a subset $A \subseteq \mathbb{R}^2$ for which $\overline{A} \subseteq_{©} \mathbb{R}^2$ but A is not connected.

19. Show that if the connected components of a compact set are open, then there are finitely many of them.

20. Let (E, d) and (F, δ) be metric spaces, together with a continuous map $f : E \to F$ such that $f_{-1}(W) \subseteq_K E$ for all $W \subseteq_K F$. Show that f is a closed map.

21. Let (E, d) be a connected metric space and let $F \subseteq_C E$, with $\partial F \subseteq_{©} E$. Show that $F \subseteq_{©} E$. Is the result still true if F is not necessarily closed?

22. Let $\Gamma = \left[\bigcup_{x \in \mathbb{Q}} (\{x\} \times (0, \infty))\right] \cup \left[\bigcup_{x \in \mathbb{R} \setminus \mathbb{Q}} (\{x\} \times (-\infty, 0))\right] \subseteq \mathbb{R}^2$.

    a) Show that $\Gamma \subseteq_{©} \mathbb{R}^2$.
    b) Show that Γ is not path-connected.

Chapter 10

Normed Vector Spaces

> The main objective of this chapter is to show that linear transformations of **finite-dimensional** normed vector spaces over $\mathbb{K}$ are **continuous**.

Norms were introduced in chapter 8; we provided a family of examples, the $p-$norms on $\mathbb{K}^n$. Let $p \geq 1$ and $A \in \mathbb{M}_{m,n}(\mathbb{K})$, the set $\mathbb{M}_{m,n}(\mathbb{K})$ of matrices of size $m \times n$ with entries in $\mathbb{K}$. The **induced** $p-$**norm on** $\mathbb{M}_{m,n}(\mathbb{K})$ is given by

$$\|A\|_p = \sup_{\|\mathbf{x}\|_p \leq 1} \{\|A\mathbf{x}\|_p\} .$$

It is easy to show that:

$$\|A\|_\infty = \max_{1\leq i\leq m} \left\{ \sum_{j=1}^{n} |a_{ij}| \right\}, \quad \|A\|_1 = \max_{1\leq j\leq n} \left\{ \sum_{i=1}^{m} |a_{ij}| \right\}, \quad \|A\|_2 = \text{ largest singular value of } A.$$

A **normed vector space** $(E, \|\cdot\|_E)$ is a vector space $(E, +, \cdot, \mathbf{0}_E)$ over $\mathbb{K}$ endowed with a norm $\|\cdot\|_E$; with matrix addition and multiplication by a scalar, the set $\mathbb{M}_{m,n}(\mathbb{K})$ is such a space for any of the induced $p-$norms. A normed vector space's operations behave as well as they could be hoped to, under the circumstances.

Proposition 139
Let E be a normed vector space over $\mathbb{K}$. The maps $+ : E \times E \to E$ and $\cdot : \mathbb{K} \times E \to E$ are continuous.

Proof: left as an exercise. ■

In what follows, let $(E, \|\cdot\|_E)$ and $(F, \|\cdot\|_F)$ be normed vector spaces over $\mathbb{K}$. A map $T : E \to F$ is **linear** if

$$T(\mathbf{0}_E) = \mathbf{0}_F \quad \text{and} \quad T(a\mathbf{x} + b\mathbf{y}) = aT(\mathbf{x}) + bT(\mathbf{y}), \quad \forall a, b \in \mathbb{K}, \mathbf{x}, \mathbf{y} \in E.$$

The set of **all linear maps from** E **to** F is denoted by $L(E, F)$. For instance, if $E = \mathbb{K}^n$ and $F = \mathbb{K}^m$, then $L(E, F) \simeq \mathbb{M}_{m,n}(\mathbb{K})$.

Theorem 140
Let $(E, \|\cdot\|_E)$ and $(F, \|\cdot\|_F)$ be two normed vector spaces over $\mathbb{K}$ and let $f \in L(E, F)$. The following conditions are equivalent:

1. *f is continuous over E*
2. *f is continuous at $\mathbf{0} \in E$*
3. *f is bounded over $\overline{B(\mathbf{0}, 1)}$*
4. *f is bounded over $S(\mathbf{0}, 1)$*
5. *$\exists M > 0$ such that $\|f(\mathbf{x})\|_F \leq M\|\mathbf{x}\|_E$ for all $\mathbf{x} \in E$.*
6. *f is Lipschitz continuous*
7. *f is uniformly continuous*

Proof: the implications $1. \implies 2.$, $3. \implies 4.$, $5. \implies 6. \implies 7. \implies 1.$ are clear.

$2. \implies 3.$: Let $\varepsilon = 1$. By continuity at $\mathbf{0}$, $\exists \delta > 0$ such that

$$\|f(\mathbf{x}) - f(\mathbf{0})\|_F = \|f(\mathbf{x})\|_F \leq 1$$

whenever $\|\mathbf{x} - \mathbf{0}\|_E = \|\mathbf{x}\|_E \leq \delta$. Now, let $\mathbf{y} \in \overline{B(\mathbf{0}, 1)}$. Since f is linear, we have

$$\|f(\mathbf{y})\|_F = \|f(\tfrac{1}{\delta}\delta\mathbf{y})\|_F = \tfrac{1}{\delta}\|f(\delta\mathbf{y})\|_F.$$

Since $\|\delta\mathbf{y}\|_E \leq \delta\|\mathbf{y}\|_E \leq \delta$. Consequently, $\|f(\delta\mathbf{y})\|_F \leq 1$ and

$$\|f(\mathbf{y})\|_F = \tfrac{1}{\delta}\|f(\delta\mathbf{y})\|_F \leq \tfrac{1}{\delta}.$$

But $\mathbf{y} \in \overline{B(\mathbf{0}, 1)}$ is arbitrary, so that f is bounded by $\frac{1}{\delta}$ over $\overline{B(\mathbf{0}, 1)}$.

$4. \implies 5.$: Since f is bounded over $S(\mathbf{0}, 1)$, $\exists N > 0$ such that $\|f(\mathbf{x})\|_F \leq N$ whenever $\|\mathbf{x}\|_E = 1$. Suppose $\mathbf{y} \neq 0_E \in E$. Then, since f is linear we have

$$\|f(\mathbf{y})\|_F = \left\| f\left(\|\mathbf{y}\|_E \tfrac{\mathbf{y}}{\|\mathbf{y}\|_E}\right)\right\|_F = \|\mathbf{y}\|_E \left\| f\left(\tfrac{\mathbf{y}}{\|\mathbf{y}\|_E}\right)\right\|_F. \tag{10.1}$$

However, $\left\|\frac{\mathbf{y}}{\|\mathbf{y}\|_E}\right\|_E = 1$ so that $\left\| f\left(\frac{\mathbf{y}}{\|\mathbf{y}\|_E}\right)\right\|_F \leq N$.

Substituting this last result in (10.1), we get that $\|f(\mathbf{y})\|_F \leq N\|\mathbf{y}\|_E$ for all $\mathbf{0} \neq \mathbf{y} \in E$. When $\mathbf{y} = 0$, the inequality remains valid since $f(\mathbf{0}_E) = \mathbf{0}_F$ and $0 = \|\mathbf{0}_F\|_F \leq N\|\mathbf{0}_E\|_E = 0$. This completes the proof. ■

If $f \in L(E,F)$ is also **continuous** (that is, if $f \in L_c(E,F)$), it then makes sense to define

$$\|f\| = \sup_{\|\mathbf{x}\|_E=1} \{\|f(\mathbf{x})\|_F\} = \sup_{\|\mathbf{x}\|_E\leq 1} \{\|f(\mathbf{x})\|_F\}.$$

With this definition, $(L_c(E,F), \|\cdot\|)$ is a normed vector space. Furthermore, if $f \in L_c(E,F)$ and $g \in L_c(F,G)$ then $g \circ f \in L_c(E,G)$ and we have

$$\|(g \circ f)(\mathbf{x})\| = \|g(f(\mathbf{x}))\| \leq \|g\|\|f(\mathbf{x})\| \leq \|g\|\|f\|\|\mathbf{x}\| \leq M\|\mathbf{x}\|$$

for some $M > 0$ and for all $\mathbf{x} \in E$. In particular, $\|f \circ g\| \leq \|f\|\|g\|$. The composition thus defines a kind of multiplication on $L_c(E,E)$; together with this multiplication, $L_c(E,E)$ is a **normed algebra**.

Theorem 141
If F is a Banach space over $\mathbb{K}$, then so is $L_c(E,F)$.

Proof: let $(f_n)_{n\in\mathbb{N}}$ be a Cauchy sequence in $L_c(E,F)$. For all $\mathbf{x} \in E$, $(f_n(\mathbf{x}))_{n\in\mathbb{N}}$ is a sequence in F. Fix such an $\mathbf{x}$. Thus, for all $p,q \in \mathbb{N}$,

$$\|f_p(\mathbf{x}) - f_q(\mathbf{x})\|_F = \|(f_p - f_q)(\mathbf{x})\|_F \leq \|f_p - f_q\|\|\mathbf{x}\|_E.$$

Let $\varepsilon > 0$. Since (f_n) is a Cauchy sequence in $L_c(E,F)$, $\exists M \in \mathbb{N}$ such that $\|f_p - f_q\|_F \leq \varepsilon$ whenever $p,q > M$. As a result, $\|f_p(\mathbf{x}) - f_q(\mathbf{x})\|_F < \varepsilon\|\mathbf{x}\|_E$ whenever $p,q > M$, so that $(f_n(\mathbf{x}))_{n\in\mathbb{N}}$ is a Cauchy sequence in F.

But F is complete so that $f_n(\mathbf{x}) \to f(\mathbf{x}) \in F$ for all $\mathbf{x} \in E$, which defines a map $f : E \to F$. It remains only to show that $f \in L_c(E,F)$ and that $f_n \to f$ in $(L_c(E,F), \|\cdot\|)$. The map f is linear as

$$f(a\mathbf{x} + b\mathbf{y}) = \lim_{n\to\infty} f_n(a\mathbf{x} + b\mathbf{y}) = \lim_{n\to\infty} [af_n(\mathbf{x}) + bf_n(\mathbf{y})] = af(\mathbf{x}) + bf(\mathbf{y})$$

for all $\mathbf{x}, \mathbf{y} \in E$, $a,b \in \mathbb{K}$. Furthermore, f is continuous since, as the Cauchy sequence (f_n) is necessarily bounded, $\exists N > 0$ such that $\|f_n\| \leq N$. Fix $\mathbf{x} \in E$ to get $\|f_n(\mathbf{x})\|_F \leq N\|\mathbf{x}\|_E$ for all n. As $n \to \infty$, we see that $\|f(\mathbf{x})\|_F \leq N\|\mathbf{x}\|_E$.

Finally, we need to show that $f_n \to f$ in $L_c(E,F)$. Let $\varepsilon > 0$. Since (f_n) is a Cauchy sequence in $L_c(E,F)$, $\exists K > 0$ such that $\|f_p - f_q\| < \varepsilon$ whenever $p,q > K$. Now, fix $\mathbf{x} \in E$. Then,

$$\|f_p(\mathbf{x}) - f_q(\mathbf{x})\|_F \leq \|f_p - f_q\|\|\mathbf{x}\|_E < \varepsilon\|\mathbf{x}\|_E$$

whenever $p,q > N$. If we fix p and let $q \to \infty$, then we have

$$\|f_p(\mathbf{x}) - f(\mathbf{x})\|_F < \varepsilon\|\mathbf{x}\|_E$$

whenever $p > N$. Since this holds for all $\mathbf{x} \in E$, we have $\|f_p - f\| \leq \varepsilon$ for all $p > N$, i.e. $f_n \to f$ in $L_c(E,F)$. ■

We have seen that the metrics d_p are equivalent in $\mathbb{K}^n$, for $p \geq 1$. Can the same be said about the norms? In fact, we can say even more: not only are the $p-$norms equivalent, but *all* norms on $\mathbb{K}^n$ are equivalent.

Proposition 142
Let E be a finite dimensional vector space over $\mathbb{K}$. All norms on E are equivalent.

Proof: suppose $\dim_{\mathbb{K}}(E) = n < \infty$. If $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is a basis of E, any $\mathbf{x} \in E$ can be written uniquely as a linear combination $\mathbf{x} = \sum_{i=1}^n x_i \mathbf{e}_i$. It is easy to see that the function $N_0 : E \to \mathbb{R}$, where

$$N_0(\mathbf{x}) = \|\varphi(\mathbf{x})\|_\infty = \|(x_1, \ldots, x_n)\|_\infty = \sup\{|x_i| \mid i = 1, \ldots, n\},$$

defines a norm on E. Let $N : E \to \mathbb{R}$ be any norm on E and set $a = \sum_{i=1}^n N(\mathbf{e}_i)$. If $\mathbf{x} \in E$, we have:

$$N(\mathbf{x}) = N\left(\sum_{i=1}^n x_i \mathbf{e}_i\right) \leq \sum_{i=1}^n N(x_i \mathbf{e}_i) \leq \sum_{i=1}^n |x_i| N(\mathbf{e}_i) \leq \sup_{i=1,\ldots,n}\{|x_i|\} \sum_{i=1}^n N(\mathbf{e}_i) = N_0(\mathbf{x}) \cdot a$$

so that $N(\mathbf{x}) \leq a N_0(\mathbf{x})$ for all $\mathbf{x} \in E$.

But the map $\varphi : (E, N_0) \to (\mathbb{K}^n, \|\cdot\|_\infty)$ is an **isometry** since $N_0(\mathbf{x}) = \|\mathbf{x}\|_\infty$ for all $\mathbf{x} \in E$, which means that it must be continuous (why?). Since

$$\tilde{S} = \{(x_1, \ldots, x_n) \in \mathbb{K}^n \mid \|(x_1, \ldots, x_n)\|_\infty = 1\} \subseteq_K \mathbb{K}^n,$$

then $S = \varphi^{-1}(\tilde{S}) = \{\mathbf{x} \in E | N_0(\mathbf{x}) = 1\} \subseteq_K E$; the norm $N : (E, N_0) \to (\mathbb{R}, |\cdot|)$ is also a continuous function – according to the max/min theorem, $\exists \mathbf{x}^* \in S$ such that $N(\mathbf{x}^*) = \inf_{\mathbf{x} \in S}\{N(\mathbf{x})\}$. Clearly, $N(\mathbf{x}^*) \neq 0$; otherwise we have $\mathbf{x}^* = \mathbf{0}$, which contradicts the fact that $\mathbf{x} \in S$ as $N_0(\mathbf{x}^*) = N_0(\mathbf{0}) = 0 \neq 1$. Hence $\inf_{\mathbf{x} \in S}\{N(\mathbf{x})\} > 0$.

Write $\inf_{\mathbf{x} \in S}\{N(\mathbf{x})\} = 1/b$ for the appropriate $b > 0$. If $\mathbf{x} = \mathbf{0} \in E$, then

$$N(\mathbf{x}) = N(\mathbf{0}) = 0 \geq 0 = \frac{1}{b} N_0(\mathbf{0}) = \frac{1}{b} N_0(\mathbf{x}).$$

If $\mathbf{x} \neq \mathbf{0} \in E$, then $\frac{\mathbf{x}}{N_0(\mathbf{x})} \in S$ and

$$N(\mathbf{x}) = N\left(N_0(\mathbf{x}) \frac{\mathbf{x}}{N_0(\mathbf{x})}\right) = N_0(\mathbf{x}) N\left(\frac{\mathbf{x}}{N_0(\mathbf{x})}\right) \geq N_0(\mathbf{x}) \cdot \frac{1}{b}.$$

In both cases, $N_0(\mathbf{x}) \leq b N(\mathbf{x})$ for all $\mathbf{x} \in E$, and so any norm N on E is equivalent to the norm N_0. By transitivity, any such norms are then equivalent to one another. ■

In general, this result is not valid if E is infinite-dimensional.

Corollary 143
Let E be a finite-dimensional vector space over $\mathbb{K}$ and let $(F, \|\cdot\|_F)$ be any normed vector space over $\mathbb{K}$. If $f : E \to F$ is a linear mapping, then f is continuous.

Proof: Let $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ be a basis of E. For any $\mathbf{x} \in E$, we have

$$\begin{aligned}\|f(\mathbf{x})\|_F &= \|f\left(\sum x_i \mathbf{e}_i\right)\|_F = \|\sum x_i f(\mathbf{e}_i)\|_F \\ &\leq \sum |x_i| \|f(\mathbf{e}_i)\|_F \leq N_0(\mathbf{x}) \cdot \sum \|f(\mathbf{e}_i)\|_F := aN_0(\mathbf{x}).\end{aligned}$$

Then for any $\varepsilon > 0$, $\exists \delta = \frac{\varepsilon}{a}$ such that

$$\|f(\mathbf{x}) - f(\mathbf{y})\|_F = \|f(\mathbf{x} - \mathbf{y})\|_F \leq aN_0(\mathbf{x} - \mathbf{y}) < a\delta = \varepsilon$$

whenever $N_0(\mathbf{x} - \mathbf{y}) < \delta$, and so f is continuous. ■

This leads to a series of useful results.

Corollary 144
Any finite-dimensional vector space over $\mathbb{K}$ is a Banach space.

Proof: this is an easy consequence of the facts that the map

$$\varphi : (E, N_0) \to (\mathbb{K}^n, \|\cdot\|_\infty)$$

is an isometry and that $(\mathbb{K}^n, \|\cdot\|_\infty)$ is a Banach space. ■

Corollary 145
Any finite-dimensional subspace of a normed vector space over $\mathbb{K}$ is closed.

Corollary 146
The compact subsets of a finite-dimensional normed vector are the subsets that are both closed and bounded under the norm.

10.1 Solved Problems

1. Let E be a normed vector space over $\mathbb{R}$ and $A, B \subseteq E$. Denote

$$A + B = \{\mathbf{a} + \mathbf{b} \mid (\mathbf{a}, \mathbf{b}) \in A \times B\}.$$

   a) If $A \subseteq_O E$, show that $A + B \subseteq_O E$.
   b) If $A \subseteq_K E$ and $B \subseteq_C E$, show that $A + B \subseteq_C E$. Is the result still true if A is only assumed to be closed in E?

Proof:

a) We can write

$$A + B = \bigcup_{\mathbf{b}\in B} (A + \{\mathbf{b}\}).$$

If $A \subseteq_O E$, we obviously have $A + \{\mathbf{b}\} \subseteq_O E$ for any $\mathbf{b} \in B$.

Indeed, if $B(\mathbf{x}, \rho) \subseteq A$, then $B(\mathbf{x} + \mathbf{b}, \rho) \subseteq A + \{\mathbf{b}\}$. Thus $A + B$ is a union of open sets: as a result, $A + B \subseteq_O E$.

b) Let $(\mathbf{z}_n) = (\mathbf{x}_n + \mathbf{y}_n) \subseteq A + B$ be such that $\mathbf{z}_n \to \mathbf{z}$ where $(\mathbf{x}_n) \subseteq A$ and $(\mathbf{y}_n) \subseteq B$. Since $A \subseteq_K E$, there is a convergent subsequence $(\mathbf{x}_{\varphi(n)})$ with $\mathbf{x}_{\varphi(n)} \to \mathbf{x} \in A$.

Since $(\mathbf{z}_{\varphi(n)})$ converges to $\mathbf{z}$, the sequence $(\mathbf{y}_{\varphi(n)}) \subseteq B$ converges to $\mathbf{y} = \mathbf{z} - \mathbf{x}$. But $B \subseteq_C E$ so that $\mathbf{y} \in B$. Thus, $\mathbf{z} = \mathbf{x} + \mathbf{y} \in A + B$, which proves the desired result. If A is only closed (and not compact), the result is false in general. Let $E = \mathbb{R}^2$, $A = \{(x, e^x) \mid x \in \mathbb{R}\}$ and $B = \mathbb{R} \times \{0\}$. Both $A, B \subseteq_C \mathbb{R}^2$ but $A + B = \mathbb{R} \times (0, \infty)$ is not closed in $\mathbb{R}^2$. ■ ■

2. Let E be a normed vector space over $\mathbb{R}$ and $\varphi : E \to \mathbb{R}$ be a linear functional on E.

   a) Show directly that φ is continuous on E if and only if $\ker \varphi \subseteq_C E$.

   b) i. Let F be a subspace of E. Show that the map $N : E/F \to \mathbb{R}$ defined by

   $$N([\mathbf{x}]) = \inf_{\mathbf{y}\in[\mathbf{x}]} \{\|\mathbf{y}\|\}$$

   is a **semi-norm** on the quotient space E/F. What can you say if $F \subseteq_C E$?

   ii. Show part a) again, this time using part b)i.

Proof:

a) If φ is continuous, then $\ker \varphi = \varphi^{-1}(\{0\}) \subseteq_C E$ since $\{0\} \subseteq_C \mathbb{R}$.

Conversely, suppose that $\ker \varphi \subseteq_C E$. If φ is not continuous, φ is unbounded on the unit sphere $S(\mathbf{0}, 1)$. Thus, $\exists (\mathbf{x}_n) \subseteq E$ such that $\|\mathbf{x}_n\| = 1$ for all $n \in \mathbb{N}$ and for which $|\varphi(\mathbf{x}_n)| \to \infty$. Let $\mathbf{u} \in E$ be such that $\varphi(\mathbf{u}) = 1$: such a $\mathbf{u} \in E$ necessarily exists because φ is linear. Indeed, if $0 \neq \varphi(\mathbf{w}) = r \in \mathbb{R}$, then $\mathbf{w} \neq \mathbf{0}$. Set $\mathbf{u} = \frac{\mathbf{w}}{\varphi(\mathbf{w})}$. Then

$$\varphi(\mathbf{u}) = \varphi\left(\frac{\mathbf{w}}{\varphi(\mathbf{w})}\right) = \frac{1}{\varphi(\mathbf{w})}\varphi(\mathbf{w}) = 1.$$

For any $n \in \mathbb{N}$, set $\mathbf{u}_n = \mathbf{u} - \frac{\mathbf{x}_n}{\varphi(\mathbf{x}_n)}$. Then

$$\varphi(\mathbf{u}_n) = \varphi(\mathbf{u}) - \varphi\left(\frac{\mathbf{x}_n}{\varphi(\mathbf{x}_n)}\right) = \varphi(\mathbf{u}) - \frac{\varphi(\mathbf{x}_n)}{\varphi(\mathbf{x}_n)} = \varphi(\mathbf{u}_n) - 1 = 0,$$

whence $\mathbf{u}_n \in \ker\varphi$ for all $n \in \mathbb{N}$. Note that $\mathbf{u}_n = \mathbf{u} - \frac{\mathbf{x}_n}{\varphi(\mathbf{x}_n)} \to \mathbf{u}$ since $|\varphi(\mathbf{x}_n)| \to \infty$ and $\|\mathbf{x}_n\| = 1$ for all n. Since $\ker\varphi$, the limit $\mathbf{u} \in \ker\varphi$, *i.e.* $\varphi(\mathbf{u}) = 0$. But this contradicts the fact that $\varphi(\mathbf{u}) = 1$. Hence φ is continuous.

b) i. Let $\mathbf{x} \in E$ and $\lambda \in \mathbb{R}$. Recall that $[\mathbf{x}] = \mathbf{x} + F$. Since $[\lambda\mathbf{x}] = \lambda[\mathbf{x}]$, we have

$$N(\lambda[\mathbf{x}]) = |\lambda|N([\mathbf{x}]).$$

It remains only to show that N satisfies the triangle inequality.
Let $\mathbf{x}, \mathbf{y} \in E$. For any $\mathbf{u}, \mathbf{v} \in F$, we have

$$N([\mathbf{x}+\mathbf{y}]) \leq \|(\mathbf{x}+\mathbf{y}) + (\mathbf{u}+\mathbf{v})\| \leq \|\mathbf{x}+\mathbf{u}\| + \|\mathbf{y}+\mathbf{v}\|.$$

Thus

$$\begin{aligned} N([\mathbf{x}+\mathbf{y}]) &\leq \inf_{\mathbf{u},\mathbf{v}\in F}\{\|\mathbf{x}+\mathbf{u}\| + \|\mathbf{y}+\mathbf{v}\|\} \\ &\leq \inf_{\mathbf{u}\in F}\{\|\mathbf{x}+\mathbf{u}\|\} + \inf_{\mathbf{v}\in F}\{\|\mathbf{y}+\mathbf{v}\|\} = N([\mathbf{x}]) + N([\mathbf{y}]). \end{aligned}$$

As such, N is a semi-norm on E/F. Since $[\mathbf{x}] = \mathbf{x}+F$ for all $\mathbf{x} \in E$, $N([\mathbf{x}]) = \inf_{\mathbf{y}\in F}\{\|\mathbf{x}-\mathbf{y}\|\} = d(\mathbf{x}, F)$. As a result, if $F \subseteq_C E$, $N([\mathbf{x}]) = 0$ if and only if $\mathbf{x} \in F$, i.e. $[\mathbf{x}] = \mathbf{0}$. Consequently, if $F \subseteq_C E$, N is a norm on E/F.

ii. Let $\varphi : E \to \mathbb{R}$ be a linear functional for which $\ker\varphi \subseteq_C E$. If $\varphi \equiv 0$, φ is clearly continuous. Otherwise, $\varphi(E) = \mathbb{R}$. Indeed, let $x \in \mathbb{R}$. If $\varphi(\mathbf{u}) = 1$ for some $\mathbf{u} \in E$, then $x\mathbf{u} \in E$, $\varphi(x\mathbf{u}) = x$ and φ is onto. Let $\eta : E \to E/\ker\varphi$ be the canonical surjection $\eta(\mathbf{u}) = \mathbf{u}+\ker\varphi$. By the Isomorphism Theorem for vector spaces, it is possible to factor $\varphi = \psi \circ \eta$, where $\psi : E/\ker\varphi \to \mathbb{R}$ is linear.

According to Corollary 143, ψ is thus continuous, being linear. If N is the norm defined in (b)i. with $F = \ker\varphi$, we have

$$N([\mathbf{x}] - [\mathbf{y}]) = N([\mathbf{x}-\mathbf{y}]) \leq \|\mathbf{x}-\mathbf{y}\| \quad \forall \mathbf{x}, \mathbf{y} \in E$$

and so η is continuous Thus, ϕ is continuous being the composition of two continuous functions. ■

3. If $\mathbf{x} = (x_1, \ldots, x_n) \in \mathbb{R}^n$, define $\|\mathbf{x}\|_\infty = \sup\{|x_1|, \ldots, |x_n|\}$. Show that $\mathbf{x} \mapsto \|\mathbf{x}\|_\infty$ defines a norm on $\mathbb{R}^n$.

Proof: There are 4 conditions to verify:

a) $\|\mathbf{x}\|_\infty = \sup\{|x_1|, \ldots, |x_n|\} \geq 0$ is clear since $|x_i| \geq 0$ for all i.
b) $\|\mathbf{x}\|_\infty = 0 \iff \sup\{|x_1|, \ldots, |x_n|\} = 0 \iff |x_i| = 0,\ \forall i \iff$
$x_i = 0,\ \forall i \iff \mathbf{x} = \mathbf{0}$.
c) If $a \in \mathbb{R}$, then

$$\|a\mathbf{x}\|_\infty = \sup\{|ax_1|, \ldots, |ax_n|\} = |a|\sup\{|x_1|, \ldots, |x_n|\} = |a|\|\mathbf{x}\|_\infty.$$

d) Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Then

$$\begin{aligned}\|\mathbf{x}+\mathbf{y}\|_\infty &= \sup\{|x_1+y_1|, \ldots, |x_n+y_n|\} \le \sup\{|x_1|+|y_1|, \ldots, |x_n|+|y_n|\} \\ &\le \sup\{|x_1|, \ldots, |x_n|\} + \sup\{|y_1|, \ldots, |y_n|\} = \|\mathbf{x}\|_\infty + \|\mathbf{y}\|_\infty.\end{aligned}$$

Thus, $\mathbf{x} \to \|\mathbf{x}\|_\infty$ defines a norm on $\mathbb{R}^n$. ■

4. Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and define the inner product $(\mathbf{x} \mid \mathbf{y}) = x_1y_1 + \cdots + x_ny_n$. As seen in the notes, this inner product defines a norm $\|\mathbf{x}\| = \sqrt{(\mathbf{x} \mid \mathbf{x})}$. Show the **Parallelogram Identity**: $\|\mathbf{x}+\mathbf{y}\|^2 + \|\mathbf{x}-\mathbf{y}\|^2 = 2(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2)$, $\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n$.

   Proof: We have

$$\begin{aligned}\|\mathbf{x}+\mathbf{y}\|^2 + \|\mathbf{x}-\mathbf{y}\|^2 &= (\mathbf{x}+\mathbf{y} \mid \mathbf{x}+\mathbf{y}) + (\mathbf{x}-\mathbf{y} \mid \mathbf{x}-\mathbf{y}) \\ &= (\mathbf{x} \mid \mathbf{x}) + 2(\mathbf{x} \mid \mathbf{y}) + (\mathbf{y} \mid \mathbf{y}) + (\mathbf{x} \mid \mathbf{x}) - 2(\mathbf{x} \mid \mathbf{y}) + (\mathbf{y} \mid \mathbf{y}) \\ &= 2(\mathbf{x} \mid \mathbf{x}) + 2(\mathbf{y} \mid \mathbf{y}) = 2(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2)\end{aligned}$$

   Now, consider a parallelogram with vertices $\mathbf{0}, \mathbf{x}, \mathbf{y}, \mathbf{x}+\mathbf{y}$. Then the sum of squares of the lengths of the four sides is $2(\|\mathbf{x}\|^2 + \|\mathbf{y}\|^2)$, while the sum of squares of the diagonals is $\|\mathbf{x}+\mathbf{y}\|^2 + \|\mathbf{x}-\mathbf{y}\|^2$. ■

5. Let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Is it true that $\|\mathbf{x}+\mathbf{y}\|_\infty = \|\mathbf{x}\|_\infty + \|\mathbf{y}\|_\infty$ if and only if $\mathbf{x} = c\mathbf{y}$ or $\mathbf{y} = c\mathbf{x}$ for some $c \ge 0$?

   Proof: No. Consider the following example in $\mathbb{R}^2$: let $\mathbf{x} = (1,0)$ and $\mathbf{y} = (1,1)$. Then $\mathbf{x}+\mathbf{y} = (2,1)$ and $\|\mathbf{x}\|_\infty + \|\mathbf{y}\|_\infty = \|\mathbf{x}+\mathbf{y}\|_\infty = 2$, but $\mathbf{x} \ne c\mathbf{y}$ for any $c \in \mathbb{R}$. ■

10.2 Exercises

1. Prepare a 2-page summary of this chapter, with important definitions and results.
2. Show that $\|A\|_\infty$, $\|A\|_1$, and $\|A\|_2$ (from the first page of this chapter) define norms over $\mathbb{M}_{m,n}(\mathbb{K})$.
3. Show that the induced $p-$norm is a norm on $\mathbb{M}_{m,n}(\mathbb{K})$ for all $p \ge 1$.
4. Prove Proposition 139.
5. Show that all isometries are continuous.
6. Prove Corollary 145.
7. Prove Corollary 146.
8. Let E be a normed vector space with a countably infinite basis. Show that E cannot be complete.
9. Let E be an infinite-dimensional normed vector space over $\mathbb{R}$. Show that $D(\mathbf{0}, 1)$ is not compact in E by showing that it is not pre-compact in E (by what name is this result usually known?).

Chapter 11

Sequences of Functions in Metric Spaces

> In this chapter, we study properties of sequences and series of functions, extending Chapters 5 and 6 to **general metric spaces** and provide important **Fourier analysis** results.

The symbol $\mathbb{K}$ is used to denote either $\mathbb{R}$ or $\mathbb{C}$; $\mathcal{C}^\ell(X,\mathbb{K})$ represents the $\mathbb{K}-$vector space of ℓ times continuously differentiable functions $X \to \mathbb{K}$; $\mathcal{F}(X,\mathbb{K})$, the $\mathbb{K}-$vector space of functions $X \to \mathbb{K}$; $\mathcal{R}(X,\mathbb{K})$, the $\mathbb{K}-$vector space of Riemann-integrable functions $X \to \mathbb{K}$, $\mathcal{C}_c(\mathbb{R},\mathbb{C})$ is the set of continuous functions $\mathbb{R} \to \mathbb{C}$ with **compact support**,[1] and $\mathcal{B}(X,\mathbb{K})$, the $\mathbb{K}-$vector space of bounded functions $X \to \mathbb{K}$.

11.1 Uniform Convergence

Let X be a set and let (E,d) be a metric space. A sequence $(f_n)_{n\in\mathbb{N}}$ of functions $f_n : X \to E$ is said to **converge pointwise** to a function $f : X \to E$ (denoted by $f_n \to f$ on X) if $f_n(\mathbf{x}) \to f(\mathbf{x})$ for all $\mathbf{x} \in X$.

Symbolically, $f_n \to f$ on X if

$$\forall \varepsilon > 0, \forall \mathbf{x} \in X, \exists N = N_{\varepsilon,\mathbf{x}} \text{ such that } n > N \implies d(f_n(\mathbf{x}), f(\mathbf{x})) < \varepsilon$$

(note the **explicit dependence** of N on $\mathbf{x}$).

As we have discussed in Chapters 5 and 6, pointwise convergence is quite often **not strong enough** of a property for our needs. Consequently, we introduce a second kind of convergence: the sequence (f_n) is said to **converge uniformly** to a function $f : X \to E$ (denoted by $f_n \rightrightarrows f$ on X) if we can remove the explicit dependence of N on $\mathbf{x}$.

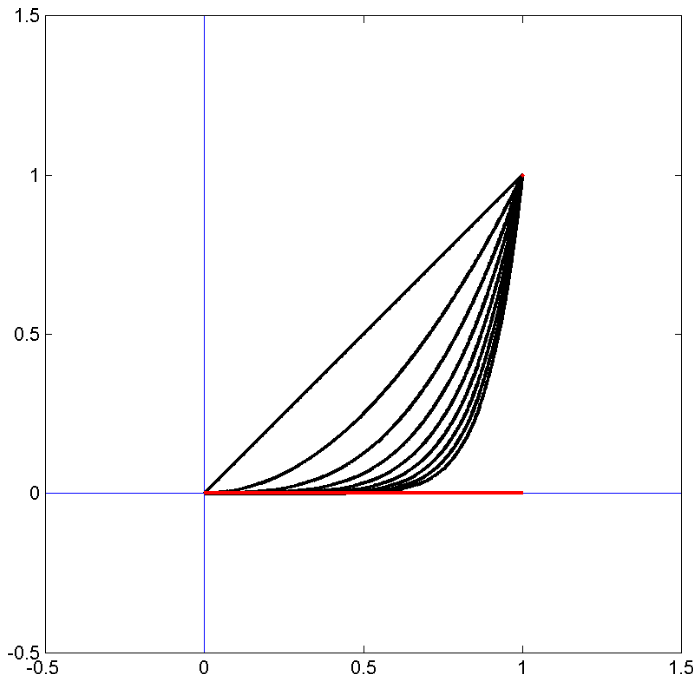
Symbolically, $f_n \rightrightarrows f$ on X if

$$\forall \varepsilon > 0, \exists N = N_\varepsilon \text{ such that } n > N \implies \sup_{\mathbf{x}\in X}\{d(f_n(\mathbf{x}), f(\mathbf{x}))\} < \varepsilon.$$

[1] That is, functions taking on the value 0 outside of some compact subset $K \subseteq \mathbb{R}$.

Examples

1. Let $(E,d)=(\mathbb{R},|\cdot|)$, $X=[0,1]$ and $f_n:X\to E$ be defined by $f_n(x)=x^n$. Then $f_n\to f$ on X, where $f:X\to E$ is given by $f(x)=0$ if $x\neq 1$ and $f(1)=1$. Note that f is not continuous on X, even though each of the f_n is continuous.



The sequence (f_n) in black, the limit f in red.

2. With the definitions as in the last example, $f_n \not\rightrightarrows f$ on X. Indeed,

$$\sup_{x\in[0,1]}\{d(f_n(x),f(x))\}=\sup_{x\in[0,1]}\{|x^n|\}=1^n=1,$$

which can never be smaller than any $1>\varepsilon>0$.

However, $f_n\rightrightarrows f$ on $[0,a]$ for all $a\in[0,1)$ (see Chapter 5).

Theorem 66 generalizes to metric spaces as one would expect.

Proposition 147 (Cauchy's Criterion for Sequences of Functions)
Let (E,d) be a complete metric space and (f_n) be a sequence of functions $f_n:X\to E$. Then, $f_n\rightrightarrows f$ on X if and only if

$$\forall\varepsilon>0,\exists N=N_\varepsilon>0\textit{ s.t. } n,m>N \implies \sup_{\mathbf{x}\in X}\{d(f_n(\mathbf{x}),f_m(\mathbf{x}))\}<\varepsilon.$$

Proof: suppose that $f_n \rightrightarrows f$ on X and let $\varepsilon > 0$. By hypothesis, $\exists N_1, N_2$ such that

$$\sup_{\mathbf{x}\in X}\{d(f_n(\mathbf{x}), f(\mathbf{x}))\} < \frac{\varepsilon}{2}, \quad \sup_{\mathbf{x}\in X}\{d(f_m(\mathbf{x}), f(\mathbf{x}))\} < \frac{\varepsilon}{2}$$

whenever $n > N_1$ and $n > N_2$. Set $N = \max\{N_1, N_2\}$.

Then, whenever $n, m > N$, we have

$$\begin{aligned}\sup_{\mathbf{x}\in X}\{d(f_n(\mathbf{x}), f_m(\mathbf{x}))\} &\leq \sup_{\mathbf{x}\in X}\{d(f_n(\mathbf{x}), f(\mathbf{x})) + d(f_m(\mathbf{x}), f(\mathbf{x}))\} \\ &\leq \sup_{\mathbf{x}\in X}\{d(f_n(\mathbf{x}), f(\mathbf{x}))\} + \sup_{\mathbf{x}\in X}\{d(f_m(\mathbf{x}), f(\mathbf{x}))\} < \varepsilon.\end{aligned}$$

Conversely, suppose that the $\varepsilon-$statement holds. Then, for any $\mathbf{x} \in X$, $(f_n(\mathbf{x}))$ is a Cauchy sequence in E and thus converges to a $f(\mathbf{x}) \in E$, as E is complete. As a result, $f_n \to f$ on X. It remains to show that $f_n \rightrightarrows f$ on X.

Let $\varepsilon > 0$. By hypothesis, $\exists N > 0$ such that $\sup_{\mathbf{x}\in X}\{d(f_n(\mathbf{x}), f_m(\mathbf{x}))\} < \frac{\varepsilon}{2}$ whenever $n, m > N$. Now, fix $n > N$ and let

$$a_m(\mathbf{x}) = d(f_n(\mathbf{x}), f_m(\mathbf{x})) \quad \text{and} \quad a(\mathbf{x}) = d(f_n(\mathbf{x}), f(\mathbf{x})).$$

Then $a_m(\mathbf{x}) \to a(\mathbf{x})$ Since $a_m(\mathbf{x}) < \frac{\varepsilon}{2}$ for all $\mathbf{x} \in X$, then $a(\mathbf{x}) \leq \frac{\varepsilon}{2}$ for all $\mathbf{x} \in X$. Hence,

$$\sup_{\mathbf{x}\in X}\{d(f_n(\mathbf{x}), f(\mathbf{x}))\} \leq \sup_{\mathbf{x}\in X}\{a(\mathbf{x})\} \leq \frac{\varepsilon}{2} < \varepsilon.$$

As such, $f_n \rightrightarrows f$ on X. ■

Series of Functions

Similar notions exist for **series** of functions. Let (E, d) be a metric space and let (u_n) be a sequence of functions $u_n : X \to E$. For any $m \in \mathbb{N}$, define the **partial sum** $f_m : X \to E$ by

$$f_m(\mathbf{x}) = u_1(\mathbf{x}) + \cdots + u_m(\mathbf{x}) = \sum_{n=1}^{m} u_n(\mathbf{x}).$$

The sequence (f_m) is the **series generated** by (u_n), and it is usually denoted by $\sum_{n\in\mathbb{N}} u_n$.

If $f_m \to f$ on X, we say that the series **converges (pointwise)** on X; if $f_m \rightrightarrows f$ on X, we say that the series **converges uniformly** on X. In both cases, f is said to be the **sum** of the series. If (f_m) does not converge, we say that the series **diverges**.

Finally, let E be a Banach space and let (g_n) be a sequence of functions $g_n \in \mathcal{B}(X, E)$. The series $\sum g_n$ **converges absolutely** on X if $\sum \|g_n\|_\infty$ converges.[2]

[2] There is no need to stipulate the type of convergence in the latter case, since that is a numerical series.

Proposition 148
If $\sum g_n$ converges absolutely on X, then $\sum g_n$ converges uniformly on X.

Proof: according to the Cauchy criterion, it suffices to show that $\forall \varepsilon > 0$, $\exists N \in \mathbb{N}$ such that

$$\Big\| \sum_{k=n}^{m} g_k \Big\|_\infty < \varepsilon.$$

But according to the triangle inequality,

$$\Big\| \sum_{k=n}^{m} g_k \Big\|_\infty \leq \sum_{k=n}^{m} \|g_k\|_\infty.$$

Since $\sum g_k$ converges absolutely, $\forall \varepsilon > 0$, $\exists N > 0$ such that

$$\sum_{k=n}^{m} \|g_k\|_\infty < \varepsilon$$

whenever $n > N$. ■

11.1.1 Properties

The two main types of convergence are not created equal, however. We establish the superiority of uniform convergence over pointwise convergence in a series of well-known theorems (which all have counterparts in Chapter 5).

Theorem 149
Let (E, d) and $(F, \tilde{d})$ be metric spaces. If $(f_n) \subseteq \mathcal{C}(E, F)$ is such that $f_n \rightrightarrows f$ on E, then $f \in \mathcal{C}(E, F)$.

Proof: let $\varepsilon > 0$ and $\mathbf{x}_0 \in E$.

Since $f_n \rightrightarrows f$ on E, then $\exists n > N$ for which $\sup_{\mathbf{x} \in E}\{d(f_n(\mathbf{x}), f(\mathbf{x}))\} < \frac{\varepsilon}{3}$. Furthermore, since f_n is continuous at $\mathbf{x}_0$, $\exists \delta > 0$ such that

$$\tilde{d}(f_n(\mathbf{x}), f_n(\mathbf{x}_0)) < \frac{\varepsilon}{3} \quad \text{whenever } d(\mathbf{x}, \mathbf{x}_0) < \delta.$$

Then

$$\begin{aligned} \tilde{d}(f(\mathbf{x}), f(\mathbf{x}_0)) &= \tilde{d}(f(\mathbf{x}), f_n(\mathbf{x})) + \tilde{d}(f_n(\mathbf{x}), f_n(\mathbf{x}_0)) + \tilde{d}(f_n(\mathbf{x}_0), f(\mathbf{x})) \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon \end{aligned}$$

whenever $d(\mathbf{x}, \mathbf{x}_0) < \delta$, hence f is continuous at $\mathbf{x}_0$. ■

We have already seen an example showing that this may not hold for pointwise convergence.

Theorem 150 (LIMIT INTERCHANGE; RIEMANN-INTEGRABLE FUNCTIONS)
Let $(E, \|\cdot\|)$ be a Banach space. If $(f_n) \subseteq \mathcal{F}([a,b],E)$ is such that $f_n \rightrightarrows f$ on $[a,b]$, and if f_n is Riemann-integrable over $[a,b]$ for all n, then f is Riemann-integrable and $\int_a^b f_n(x)\,dx \to \int_a^b f(x)\,dx$.

Proof: left as an exercise (see Chapter 5). ■

Although, the fact that the limit interchange is not necessarily valid if $f_n \to f$ instead of $f_n \rightrightarrows f$ on $[a,b]$ could be seen as an indictment of the **Riemann integral** rather than as an indictment of pointwise convergence. In chapter 21, we will take the former position and introduce the **Lebesgue (Borel) integral** to circumvent this difficulty.

The next result is a companion to Theorem 150.

Theorem 151 (LIMIT INTERCHANGE; FUNDAMENTAL THEOREM)
Let $(E, \|\cdot\|)$ be a Banach space. If $(f_n) \subseteq \mathcal{F}([a,b],E)$ is such that $f_n \rightrightarrows f$ on $[a,b]$, and if f_n is Riemann-integrable over $[a,b]$ for all n, then f is Riemann-integrable according to Theorem 150. Define $F_n, F : [a,b] \to E$ by $F_n(x) = \int_a^x f_n(t)\,dt$ and $F(x) = \int_a^x f(t)\,dt$. Then $F_n \rightrightarrows F$ on $[a,b]$.

Proof: let $\varepsilon > 0$.

Since $f_n \rightrightarrows f$ on $[a,b]$, $\exists N \in \mathbb{N}$ such that $\|f - f_n\|_\infty < \frac{\varepsilon}{2(b-a)}$ whenever $n > N$. Now,

$$\begin{aligned}\|F_n(\mathbf{x}) - F(\mathbf{x})\| &= \left\| \int_a^x (f_n(t) - f(t))\,dt \right\| \leq \int_a^x \|f_n(t) - f(t)\|\,dt \\ &\leq \int_a^x \|f_n - f\|_\infty\,dt < \frac{\varepsilon}{2(b-a)}(x-a) \leq \frac{\varepsilon}{2(b-a)}(b-a) = \frac{\varepsilon}{2}.\end{aligned}$$

Since this is true for all $x \in [a,b]$, then $\|F_n - F\|_\infty \leq \frac{\varepsilon}{2} < \varepsilon$. By the Cauchy criterion, $F_n \rightrightarrows F$ on $[a,b]$. ■

Theorem 151 has an interesting corollary when applied to series, which is often assumed to hold (without proof) when solving differential equations.

Theorem 152 Let $(E, \|\cdot\|)$ be a Banach space and let $\sum g_n$ be a series of functions in $\mathcal{R}([a,b],E)$. If $\sum g_n$ is uniformly convergent, then

$$\int_a^b \left(\sum_{n\in\mathbb{N}} g_n(t) \right)\,dt = \sum_{n\in\mathbb{N}} \left(\int_a^b g_n(t)\,dt \right).$$

Proof: this is a direct consequence of Theorem 151. ■

We have not defined differentiability of functions $\mathbb{R} \to E$ in a general normed vector space E, but we can use functions $\mathbb{R} \to \mathbb{K}^n$ as a template: a function $f : \mathbb{R} \to \mathbb{K}^n$ is **differentiable at** t if

$$f'(t) = \lim_{h \to 0} \frac{f(t+h) - f(t)}{h}$$

exists; it is **differentiable over** $\mathbb{R}$ if it is differentiable at all $t \in \mathbb{R}$. Differentiability is also the subject of a limit interchange theorem.

Theorem 153 (Limit Interchange; Differentiable Functions)
Let $(E, \|\cdot\|)$ be a Banach space. If $(f_n) \subseteq \mathcal{C}^1([a,b], E)$ is such that $f_n(x_0) \to f(x_0)$ for some $x_0 \in [a,b]$ and if $\exists g \in \mathcal{C}([a,b], E)$ such that $f_n' \rightrightarrows g$ on $[a,b]$, then $\exists f \in \mathcal{C}^1([a,b], E)$ such that $f_n \rightrightarrows f$ on $[a,b]$ and $f' = g$.

Proof: according to the fundamental theorem of calculus, for all $n \in \mathbb{N}$ we have $f_n(x) - f_n(a) = \int_a^x f_n'(t)\,dt$. Since $f_n' \rightrightarrows g$, then

$$f_n(x) - f_n(a) = \int_a^x f_n'(t)\,dt \rightrightarrows \int_a^x g(t)\,dt \quad \text{on } [a,b],$$

according to Theorem 150. In particular, the sequence $(f_n(x_0) - f(a))_n$ converges, which implies that $(f_n(a))_n$ converges to some $\ell \in E$. It is easy to show that $f_n \rightrightarrows f$, where $f : [a,b] \to E$ is defined by

$$f(x) = \ell + \int_a^x g(t)\,dt.$$

Since all the f_n are continuous and the convergence is uniform, then f is continuous. It is also differentiable, and its derivative is continuous as $f' = g \in \mathcal{C}([a,b], E)$ (again, according to the fundamental theorem of calculus). ■

We can use these theorems to compute various quantities that would be difficult to compute directly.

Examples

1. Compute $\int_0^\infty f(x)\,dx$, where $f(x) = \frac{x^2}{\exp(x) - 1}$.

   Solution: consider $(g_n) \subseteq \mathcal{C}(\mathbb{R}^+, \mathbb{R}^+)$ defined by $g_n(x) = \exp(-nx)x^2$ for all $n \in \mathbb{N}^\times$. Then $\sum g_n$ converges pointwise to $f : \mathbb{R}^+ \to \mathbb{R}^+$.

Indeed,

$$\begin{aligned}\sum_{n=1}^m g_n(x) = x^2\left(\sum_{n=1}^m \exp(-nx)\right) &= x^2\left(\sum_{n=1}^m (\exp(-x))^n\right)\\ &= x^2\left(\frac{\exp(-x)-\exp(-(m+1)x)}{1-\exp(-x)}\right) \leq f(x),\end{aligned}$$

since $\exp(-x) < 1$ for all $x \in \mathbb{R}^+$.

Then,

$$\begin{aligned}\sum_{n\in\mathbb{N}^\times} g_n(x) = \lim_{m\to\infty}\sum_{n=1}^m g_n(x) &= \lim_{m\to\infty} x^2\left(\frac{\exp(-x)-\exp(-(m+1)x)}{1-\exp(-x)}\right)\\ &= \frac{x^2}{\exp(x)-1}.\end{aligned}$$

Furthermore, $\sum g_n$ converges absolutely to f on $[a,b] \subseteq (0,\infty)$.

Indeed, for all $x \in [a,b]$, we have $|g_n(x)| \leq \exp(-na)b^2$. Note that

$$\sum_{n\in\mathbb{N}^\times} \exp(-na)b^2 = b^2\sum_{n\in\mathbb{N}^\times}(\exp(-a))^n = \frac{b^2}{\exp(a)-1}, \quad \text{since } a > 0.$$

Hence $\sum_{n\in\mathbb{N}^\times}\exp(-na)b^2$ converges and so, according to Exercise 1, $\sum g_n$ is absolutely convergent.

Since $\int_0^\infty f(t)\,dt$ converges (use the Comparison Theorem with $\exp(-\sqrt{x})$, for instance), then, according to Theorem 152,

$$\int_0^\infty f(t)\,dt = \int_0^\infty\left(\sum_{n\in\mathbb{N}^\times} g_n(t)\right)dt = \sum_{n\in\mathbb{N}^\times}\left(\int_0^\infty g_n(t)\,dt\right)$$

Repeated integration by parts shows that $\int_0^\infty g_n(t)\,dt = \frac{2}{n^3}$, so that

$$\int_0^\infty \frac{x^2}{\exp(x)-1}\,dx = 2\sum_{n\in\mathbb{N}^\times}\frac{1}{n^3} = 2\zeta(3). \qquad \blacksquare$$

2. Show that uniform convergence is not equivalent to absolute convergence.

Proof: it will be sufficient to exhibit a series which is uniformly convergent but not absolutely convergent. Consider (u_k) a series of constant functions from an interval I to $\mathbb{R}$ defined by $u_k(x) = \frac{(-1)^k}{k}$ for all $x \in I$.

Since $\|u_k\|_\infty = \frac{1}{k}$, and since $\sum \frac{1}{k}$ diverges (it is the **harmonic series**, after all), then $\sum u_k$ is not absolutely convergent. However,

$$\left\| \sum_{k=n}^{m} u_k \right\|_\infty = \left| \sum_{k=n}^{m} \frac{(-1)^k}{k} \right| \leq \frac{1}{n} \to 0 \quad \text{as } n, m \to \infty,$$

so that $\sum u_k$ is uniformly convergent. ■

11.1.2 Abel's Criterion

In calculus courses and in Chapters 5 and 6, we have seen a number of tests can be used to gauge the convergence of series (whether numerical series or series of functions):

- $p-$test;
- comparison test;
- alternating series test;
- integral test;
- d'Alembert test (also known as the ratio test), or
- Cauchy test (also known as the root test).

In this section, we present a new test.

Proposition 154 (Abel's Criterion)
Let $(\mathbf{a}_n) \subseteq E$, where E is a Banach space over $\mathbb{R}$. If we can write $\mathbf{a}_n = \varepsilon_n \mathbf{b}_n$ with

1. *$\varepsilon_n \searrow 0$ a sequence in $\mathbb{R}$, and*
2. *$\exists \sigma \in \mathbb{R}$ such that $\| \sum_{n \leq N} \mathbf{b}_n \| \leq \sigma$ for all $N \in \mathbb{N}$.*

Then $\sum \mathbf{a}_n$ is pointwise convergent and $\| \sum_{n \geq N} \mathbf{a}_n \| \leq 2\sigma\varepsilon_N$ for all $N \in \mathbb{N}$.

Proof: for any $q > p$, let $S_p^q = \mathbf{b}_{p+1} + \cdots + \mathbf{b}_q$. Since $S_p^q = \sum_{n \leq q} \mathbf{b}_n - \sum_{n \leq p} \mathbf{b}_n$, we have $\left\| S_p^q \right\| \leq 2\sigma$. If we write

$$\mathbf{b}_{p+1} = S_p^{p+1},\ \mathbf{b}_{p+2} = S_p^{p+2} - S_p^{p+1},\ \cdots,\ \mathbf{b}_q = S_p^q - S_p^{q-1},$$

then

$$\begin{aligned}\varepsilon_{p+1}\mathbf{b}_{p+1}+\cdots+\varepsilon_q\mathbf{b}_q &= \varepsilon_{p+1}S_p^{p+1}+\varepsilon_{p+2}\left(S_p^{p+2}-S_p^{p+1}\right)+\cdots+\varepsilon_q\left(S_p^q-S_p^{q-1}\right)\\ &= S_p^{p+1}\left(\varepsilon_{p+1}-\varepsilon_{p+2}\right)+\cdots+S_p^{q-1}\left(\varepsilon_{q-1}-\varepsilon_q\right)+\varepsilon_q S_p^q,\end{aligned}$$

whence

$$\begin{aligned}\Big\|\sum_{k=p+1}^{q}\mathbf{a}_k\Big\| &= \|\varepsilon_{p+1}\mathbf{b}_{p+1}+\cdots+\varepsilon_q\mathbf{b}_q\|\\ &\leq \left\|S_p^{p+1}\right\|\left|\varepsilon_{p+1}-\varepsilon_{p+2}\right|+\cdots+\left\|S_p^{q-1}\right\|\left|\varepsilon_{q-1}-\varepsilon_q\right|+\left|\varepsilon_q\right|\left\|S_p^q\right\|\\ &\leq 2\sigma\left(\varepsilon_{p+1}-\varepsilon_{p+2}\right)+\cdots+2\sigma\left(\varepsilon_{q-1}-\varepsilon_q\right)+2\sigma\varepsilon_q\\ &= 2\sigma\varepsilon_{p+1}\to 0 \quad \text{as } p,q\to\infty\end{aligned}$$

Hence, $\sum \mathbf{a}_k$ converges by the Cauchy Criterion. ■

We can easily generalize this result to sequences of functions.

Proposition 155 (ABEL'S CRITERION (REPRISE))
Let $\sum f_n$ be such that $f_n=\varepsilon_n g_n \in \mathcal{F}([a,b],E)$, where E is a Banach space over $\mathbb{R}$. If

1. $\varepsilon_n(x)\searrow 0$ *for all* $x\in[a,b]$;
2. $\exists\sigma\in\mathbb{R}$ *such that* $\|\sum_{n\leq N} g_n(x)\|\leq\sigma$ *for all* $N\in\mathbb{N}$ *and all* $x\in[a,b]$, *and*
3. $\|\varepsilon_n\|_\infty\to 0$.

Then $\sum f_n$ is uniformly convergent on $[a,b]$.

Proof: left as an exercise. ■

The three conditions are in fact independent (see Exercise 7). For the next example (and the rest of the chapter), we assume some familiarity with complex numbers (see Chapter 22 if necessary).

Example: consider the series $\sum_{k\in\mathbb{N}} c_k b_k(x)$, where $b_k(x)=e^{ikx}$, $x\in\mathbb{R}$ and $c_k\searrow 0$. Show that the series converges (pointwise) for any $x\in(0,2\pi)$ and that it converges uniformly on $[\delta,2\pi-\delta]$ for any $\delta\in(0,\pi)$.

Proof: since $|e^{ikx}|=1$, then $\sum_{k\in\mathbb{N}} c_k e^{ikx}$ is absolutely convergent whenever $\sum_{k\in\mathbb{N}}|c_k|<\infty$. If $x\neq 2k\pi$, $k\in\mathbb{N}$, then

$$1+e^{ix}+\cdots+e^{inx}=\frac{1-e^{i(n+1)x}}{1-e^{ix}},$$

whence
$$\left|\sum_{k=1}^{n} b_k(x)\right| = |1+e^{ix}+\cdots+e^{inx}| \leq \frac{2}{|1-e^{ix}|} := \sigma_x.$$
According to Abel's criterion for numerical series, $\sum_{k\in\mathbb{N}} c_k e^{ikx}$ thus converges pointwise for any $x\in(0,2\pi)$.

Now, let $\pi>\delta>0$ and $x\in[\delta,2\pi-\delta]$. Then
$$|1-e^{ix}| = \left|e^{ix/2}(e^{-ix/2}-e^{ix/2})\right| = 2\left|\frac{e^{ix/2}-e^{-ix/2}}{2i}\right| = 2|\sin(x/2)| > \sin\delta,$$
from which we can conclude that
$$\left|\sum_{k=1}^{n} b_k(x)\right| \leq \frac{2}{\sin\delta} := \sigma.$$
Consequently, again according to Abel's criterion applied to series of functions, $\sum_{k\in\mathbb{N}} c_k e^{ikx}$ converges uniformly for any on $[\delta,2\pi-\delta]$ for any $\pi>\delta>0$. ■

11.2 Fourier Series

The series $\sum_{k\in\mathbb{N}} c_k e^{ikx}$ in the previous example is continuous on $(0,2\pi)$ even though it fails to converge uniformly on $(0,2\pi)$. It is an example of a **Fourier Series**, a monumental idea in the development of modern mathematics. They were first proposed as solutions to the **heat equation**, in which we seek functions $u: U\subseteq_O \mathbb{R}^2\times(a,b)\to\mathbb{R}$ satisfying the partial differential equation
$$\frac{\partial u}{\partial t} = \Delta u = \frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}+\frac{\partial^2 u}{\partial z^2}.$$
Since the Fourier series approach gave rise to **already-known solutions** of the heat equation, the process with which they were formed was accepted as valid, even though a number of mathematicians had objections concerning the use of infinity and (possibly divergent) series.

The importance of rigour in mathematics was just starting to be understood by some of the best mathematical minds; while these objections may sound a bit odd nowadays, it is important to remember that the current definitions of the concepts that made some of our predecessors queasy have been distilled of all offending material after years of polishing, which was driven by the very objections that they brought up.

It is no exaggeration to say that analysis would not be what it is today without this particular episode; while it remains in fashion amongst some mathematicians to deride engineers and physicists for "playing with tools beyond their understanding", let us keep in mind that analytical advances mostly arise from the application of mathematics to so-called 'real-world' problems, in the grand tradition of Archimedes and Newton.

In this section, we introduce and discuss the convergence of Fourier Series.

11.2.1 Trigonometric Series and Periodic Functions

A **trigonometric polynomial** is any (finite) linear combination of positive powers of sines and cosines:

$$p(t) = a_0 + \sum_{k=1}^{n} \left(a_k \cos(kt) + b_k \sin(kt)\right), \quad \text{where } a_k, b_k \in \mathbb{C}.$$

Since

$$\cos t = \frac{e^{it} + e^{-it}}{2}, \quad \sin t = \frac{e^{it} - e^{-it}}{2i},$$

we can write

$$p(t) = a_0 + \sum_{k=1}^{n} \left(a_k \cos(kt) + b_k \sin(kt)\right) = \sum_{k=-n}^{n} c_k e^{ikt},$$

with

$$a_0 = c_0, \quad a_k = c_k + c_{-k}, \quad \text{and} \quad b_k = i(c_k - c_{-k}),$$

or

$$c_0 = a_0, \quad c_k = \frac{a_k - ib_k}{2}, \quad \text{and} \quad c_{-k} = \frac{a_k + ib_k}{2},$$

for all $1 \leq k \leq n$.

A **trigonometric series** is a formal expression of the form

$$\sum_{k\in\mathbb{Z}} c_k e^{ikt} = a_0 + \sum_{k\in\mathbb{N}} \left(a_k \cos(kt) + b_k \sin(kt)\right).$$

We say that a series indexed by $\mathbb{Z}$ **converges** if both the series indexed by the positive integers **and** the series indexed by the negative integers converges.

> **Proposition 156**
> *If $\sum_{k\in\mathbb{Z}} c_k e^{ikt}$ converges absolutely for some t, then $\sum_{k\in\mathbb{Z}} |c_k| < \infty$. Furthermore, if $\sum_{k\in\mathbb{Z}} |c_k| < \infty$, then $\exists f \in \mathcal{C}(\mathbb{R}, \mathbb{C})$ such that $\sum_{k\in\mathbb{Z}} c_k e^{ikt} \rightrightarrows f$ on $\mathbb{R}$.*
>
> **Proof:** left as an exercise. ■

These ideas will become more clear with a concrete example.

> **Example:** Let $b \in (-1, 1)$. Consider the trigonometric series $\sum_{k\in\mathbb{N}} b^k \sin(kt)$. What is its complex form? Does it converge anywhere? If so, what to?
>
> **Solution:** according to the previous formulas, we formally have
>
> $$c_0 = 0, \quad c_k = \frac{0 - ib^k}{2} = \frac{b^k}{2i} \quad \text{and} \quad c_{-k} = \frac{0 + ib^k}{2} = -\frac{b^k}{2i},$$
>
> for $k \geq 1$.

We also have

$$\sum_{k=1}^{n} b^k \sin(kt) = -\frac{1}{2i}\sum_{k=-n}^{-1} b^{-k}e^{ikt} + \frac{1}{2i}\sum_{k=1}^{n} b^k e^{ikt},$$

so that, formally,

$$\sum_{k=1}^{\infty} b^k \sin(kt) = -\frac{1}{2i}\sum_{k=-\infty}^{-1} b^{-k}e^{ikt} + \frac{1}{2i}\sum_{k=1}^{\infty} b^k e^{ikt}.$$

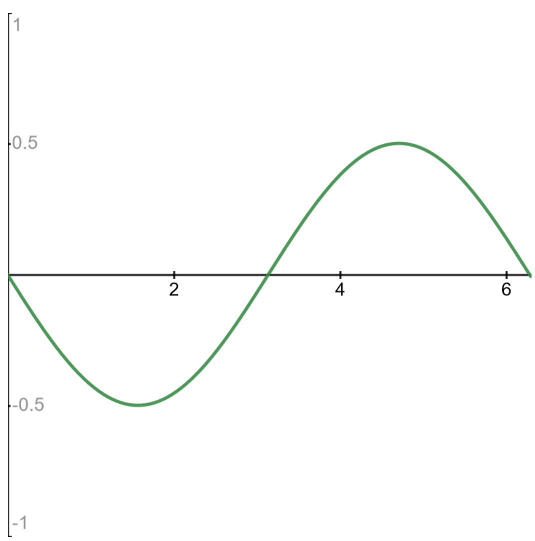
The series converges absolutely (and thus at least pointwise), as

$$\sum_{k\geq 1} \|b^k \sin(kt)\|_\infty = \sum_{k\geq 1} |b|^k = \frac{|b|}{1-|b|} < \infty, \quad \text{since } |b| < 1.$$

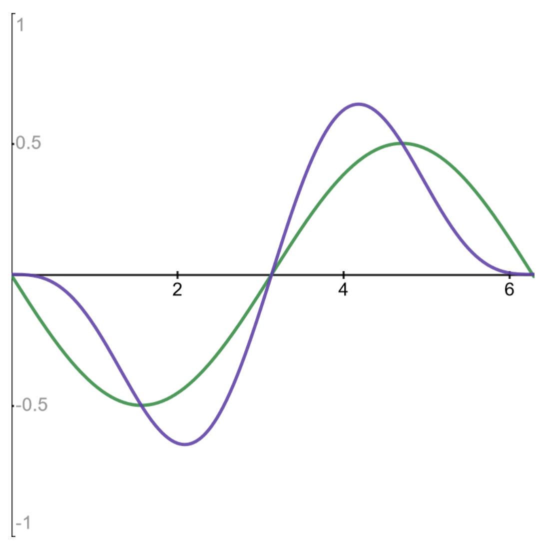
According to Proposition 148, $\exists f \in \mathcal{C}(\mathbb{R},\mathbb{C})$ to which the series converges uniformly on $\mathbb{R}$. We can re-write the convergent series as

$$\begin{aligned}\sum_{k=1}^{\infty} b^k \sin(kt) &= \frac{1}{2i}\left[\sum_{k=1}^{\infty}\left(be^{it}\right)^k - \sum_{k=1}^{\infty}\left(be^{-it}\right)^k\right] = \frac{1}{2i}\left(\frac{be^{it}}{1-be^{it}} - \frac{be^{-it}}{1-be^{-it}}\right)\\ &= \frac{b}{2i}\cdot\frac{e^{it}-e^{-it}}{1-b(e^{it}+e^{-it})+b^2} = b\cdot\underbrace{\frac{e^{it}-e^{-it}}{2i}}_{=\sin t}\cdot\frac{1}{1-2b\underbrace{\frac{e^{it}+e^{-it}}{2}}_{=\cos t}+b^2}.\end{aligned}$$

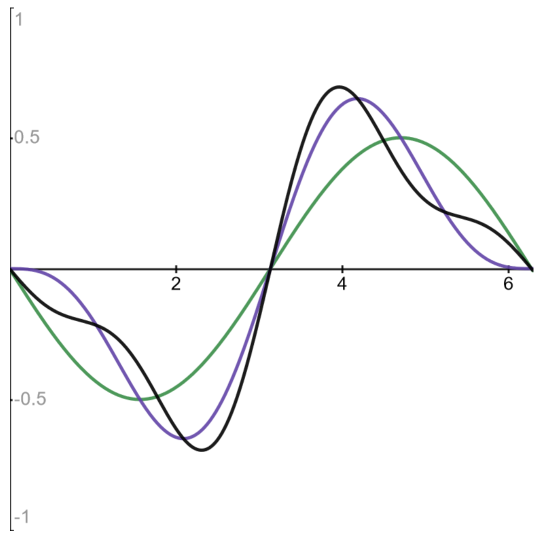
Thus the series converges uniformly to $f : t \mapsto \frac{b\sin t}{1-2b\sin t+b^2}$ on $\mathbb{R}$. ■



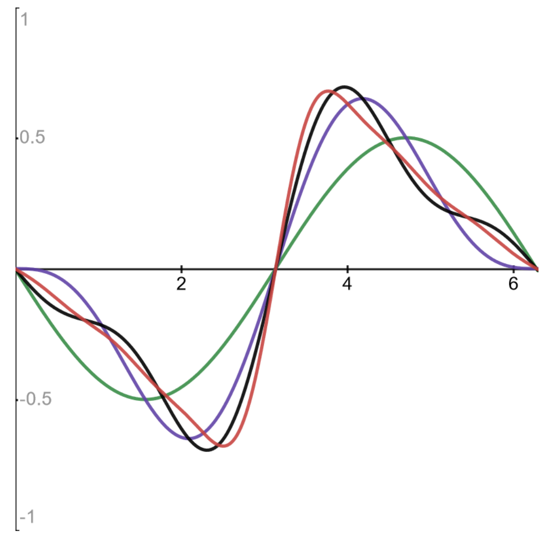
$b = -1/2, k = 1$



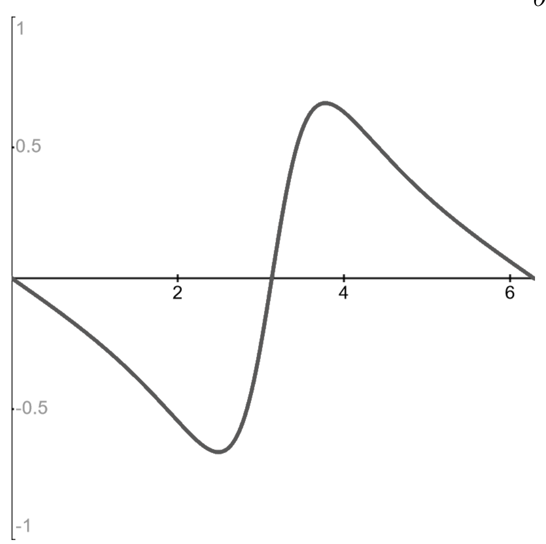
$b = -1/2, k = 1, 2$



$b = -1/2, k = 1, 2, 3$



$b = -1/2, k = 1, \ldots, 6$



$b = -1/2, f(t) = \frac{b \sin t}{1 - 2b \sin t + b^2}$

11.2.2 Again, Abel's Criterion

Proposition 157
Let $\sum_{k\in\mathbb{Z}} c_k e^{ikt}$ be such that $c_k \geq 0$ and $c_k \searrow 0$ both as $k \to \infty$ and as $k \to -\infty$. Then $\sum_{k\in\mathbb{Z}} c_k e^{ikt}$ converges uniformly on $[\delta, 2\pi - \delta]$ for any $\delta \in (0, \pi)$. Consequently, the sum $f(t) = \sum_{k\in\mathbb{Z}} c_k e^{ikt}$ is continuous on $(0, 2\pi)$.

Proof: it suffices to show that

$$\sum_{k\geq 0} c_k e^{ikt} \quad \text{and} \sum_{k\leq -1} c_k e^{ikt}$$

both converge uniformly on $[\delta, 2\pi-\delta]$ for all $0<\delta<\pi$, and to apply Abel's criterion for each of the series.

Let $\delta \in (0,\pi)$. Since

$$\begin{aligned}\left|\sum_{k=0}^{n} e^{ikt}\right| &= \left|1+\cdots+e^{int}\right| = \left|\frac{1-e^{i(n+1)t}}{1-e^{it}}\right| \leq \frac{2}{|1-e^{it}|} \leq \frac{2}{\sin\delta} \\ \left|\sum_{k=-n}^{-1} e^{ikt}\right| &= \left|e^{-int}+\cdots+e^{-it}\right| = \left|e^{-int}\right|\left|1+\cdots+e^{i(n-1)t}\right| \\ &= \left|1+\cdots+e^{i(n-1)t}\right| = \left|\frac{1-e^{int}}{1-e^{it}}\right| \leq \frac{2}{|1-e^{it}|} \leq \frac{2}{\sin\delta}\end{aligned}$$

for all $t\in[\delta,2\pi-\delta]$, the series converge uniformly on $[\delta,2\pi-\delta]$. ■

Abel's criterion could also be used even in circumstances where c_k is not always positive. For instance, let $\sum_{k\in\mathbb{Z}}(-1)^k c_k e^{ikt}$ where the coefficient c_k are as in the statement of Proposition 157. What does the fact that

$$\left|\sum_{k\in\mathbb{Z}}(-1)^k e^{ikt}\right| = \left|\frac{1+(-1)^{n+1}e^{i(n+1)t}}{1-e^{it}}\right| \leq \frac{2}{|1+e^{it}|}$$

tell you? These results also apply to the real part and the imaginary part of $\sum_{k\in\mathbb{Z}} c_k e^{ikt}$, i.e. to the series

$$a_0 + \sum_{k\geq 1} a_k \cos(kt) \quad \text{and} \quad \sum_{k\geq 1} b_k \sin(kt).$$

For instance, $\sum_{k\geq 1}\frac{\sin(kt)}{k}$ converges uniformly on $[\delta,2\pi-\delta]$ for any $\delta>0$. As a result, the sum is continuous on $(0,2\pi)$. However, even though $\sum_{k\geq 1}\frac{\sin(kt)}{k}$ converges for $t=0$ and $t=2\pi$, the function is not continuous on $[0,2\pi]$ (see Exercise 9).

Let $T>0$. A function $f:\mathbb{R}\to\mathbb{C}$ is $T-$**periodic** if $f(t+T)=f(t)$ for all $t\in\mathbb{R}$. The smallest positive T for which this holds is the period of the function. Periodic functions play an important role in Fourier analysis.

Examples

1. The functions cos and sin are $2\pi-$periodic. □

2. The function tan is $\pi-$periodic. □

3. The function defined by e^{ikt} is $\frac{2\pi}{k}-$periodic for any $k \in \mathbb{Z}$. □

4. The function defined by e^{ikwt}, where $w = \frac{2\pi}{T}$ and $k \in \mathbb{Z}$, is $T-$periodic. □

5. Let $f \in \mathcal{C}_c(\mathbb{R}, \mathbb{C})$, with **compact support** K (that is, $f(t) = 0$ when $t \notin K$). Show that $\varphi_f : t \mapsto \sum_{k\in\mathbb{Z}} f(t-k)$ is $1-$periodic.

   Proof: this series converges for all t since there is only a finite set of integers k for which $t - k \in K$ (because K is compact). Then

$$\varphi(t+1) = \sum_{k\in\mathbb{Z}} f(t+1-k) = \sum_{k\in\mathbb{Z}} f(t-k) = \varphi_f(t),$$

   so φ_f is $1-$periodic. ■

If $f \in \mathcal{C}(\mathbb{R}, \mathbb{C})$ is a $T-$periodic function, then f is bounded on the interval $[0, T]$, with

$$c_0(f) = \frac{1}{T}\int_0^T f(t)\,dt < \infty.$$

The complex number c_0 is the **mean value of** f, also given by

$$c_0(f) = \frac{1}{T}\int_0^T f(t)\,dt.$$

If $w = \frac{2\pi}{T}$ and $k \neq 0$, the function $g : t \mapsto e^{ikwt}$ is $T-$periodic. Then

$$c_0(g) = \frac{1}{T}\int_0^T e^{ikwt}\,dt = \frac{1}{T}\left[\frac{e^{ikwt}}{ikw}\right]_0^T = 0.$$

Hence, if $f(t) = \sum_{k\in\mathbb{Z}} c_k e^{ikwt}$ is uniformly convergent on $[0, T]$ and $T-$periodic, then

$$c_0(f) = \frac{1}{T}\int_0^T f(t)\,dt = \frac{1}{T}\int_0^T \left(\sum_{k\in\mathbb{Z}} c_k e^{ikwt}\right) dt = \sum_{k\in\mathbb{Z}} \frac{c_k}{T}\int_0^T e^{ikwt}\,dt = c_0$$

The sum and the integral can be interchanged because the series converges uniformly on $[0, T]$. If $f \in \mathcal{C}(\mathbb{R}, \mathbb{C})$ is $T-$periodic, the sequence $(c_k(f))$, where

$$c_k(f) = c_0\left(fe^{-ikwt}\right) = \frac{1}{T}\int_0^T f(t)e^{-ikwt}\,dt, \quad k \in \mathbb{Z},$$

is the sequence of **Fourier coefficients of** f. Clearly, if $w = \frac{2\pi}{T}$ and $f(t) = \sum_{k\in\mathbb{Z}} c_k e^{ikwt}$ is uniformly convergent on $[0, T]$, then $c_k(f) = c_k$.

Proposition 158
The mapping $f \mapsto (c_k(f))_{k\in\mathbb{Z}}$ is a linear map from the vector space of continuous $T-$periodic functions to the space of bounded sequences indexed by $\mathbb{Z}$. More precisely,

$$\sup_{k\in\mathbb{Z}}\{|c_k(f)|\} \leq \|f\|_1 \leq \|f\|_\infty < \infty,$$

where $\|f\|_1 = \frac{1}{T}\int_0^T |f(t)|\,dt$.

Proof: left as an exercise. ■

We can improve on Proposition 158 once we show that

$$\|f\|_2 = \left(\sum_{k\in\mathbb{Z}} |c_k(f)|^2\right)^{1/2}.$$

Proposition 159 Let f be a $2\pi-$periodic function such that $f \in C^n$, $n > 0$. Then

$$c_k(f) = \frac{1}{(ik)^n} c_k\left(f^{(n)}\right), \quad k \neq 0.$$

In particular,

$$|c_k(f)| \leq \frac{\|f^{(n)}\|_\infty}{|k|^n}$$

and so $|c_k(f)| \to 0$ as $|k| \to \infty$.

Proof: this is easily shown by induction on n. If $n = 1$, we have

$$\begin{aligned} c_k(f) &= \frac{1}{2\pi}\int_0^{2\pi} f(t)e^{-ikt}\,dt = \frac{1}{2\pi}\left[\left.\frac{f(t)e^{-ikt}}{-ik}\right|_0^{2\pi} + \frac{1}{ik}\int_0^{2\pi} f'(t)e^{-ikt}\,dt\right] \\ &= \frac{1}{ik}c_k(f'). \end{aligned}$$

A sequence of integrations by parts yields the result for general n. ■

As a corollary, if $f \in C^2$ is $2\pi-$periodic, then $\sum_{k\in\mathbb{Z}} c_k(f)e^{ikt}$ converges absolutely (and so uniformly) on $\mathbb{R}$.

All that precedes leads us to the crucial definition: the **Fourier series** of a $2\pi-$periodic function f is the series $\sum_{k\in\mathbb{Z}} c_k(f)e^{ikt}$; in that case, we write $f(t) \sim \sum_{k\in\mathbb{Z}} c_k(f)e^{ikt}$. Note that it is possible for f **not to equal** its Fourier series.

11.2.3 Convergence of Fourier Series

The next results discuss the **convergence** of Fourier series.

Theorem 160
Let f be $2\pi-$periodic. If $f \in C^2$, then the Fourier series $\sum_{k\in\mathbb{Z}} c_k(f)e^{ikt}$ converges absolutely (and so uniformly) to f on $\mathbb{R}$.

Proof: according to the corollary to Proposition 159, the Fourier series $g(t) = \sum_{k\in\mathbb{Z}} c_k(f)e^{ikt}$ converges absolutely on $\mathbb{R}$, and thus g is continuous and $2\pi-$periodic. We want to show that $g = f$.

Let $h = f - g$. Then h is continuous and $2\pi-$periodic. We also have

$$c_k(h) = c_k(f) - c_k(g) = 0,$$

so that $c_k(f) = c_k(g)$ for all $k \in \mathbb{Z}$.

It remains only to show that when h is continuous, $2\pi-$periodic, and $c_k(h) = 0$ for all $k \in \mathbb{Z}$, then $h \equiv 0$. According to a corollary of the Stone-Weierstrass theorem (see Chapter 23), $\exists (p_n)_{n\in\mathbb{N}}$ such that $p_n(t) = \sum_{k\in\mathbb{Z}} a_k(n)e^{ikt}$ and $p_n \rightrightarrows \overline{h}$. Note that for a fixed k, we must have $a_k(n) \to 0$ when $n \to \infty$.

Then

$$\begin{aligned}\frac{1}{2\pi}\int_0^{2\pi} |h(t)|^2\,dt &= \frac{1}{2\pi}\int_0^{2\pi} h(t)\overline{h(t)}\,dt \overset{\text{thm 150}}{=} \lim_{n\to\infty}\frac{1}{2\pi}\int_0^{2\pi} h(t)p_n(t)\,dt \\ &\overset{\text{thm 152}}{=} \sum_{k\in\mathbb{Z}}\left(\lim_{n\to\infty} a_k(n)\frac{1}{2\pi}\int_0^{2\pi} h(t)e^{ikt}\,dt\right) = \sum_{k\in\mathbb{Z}}\left(\lim_{n\to\infty} a_k(n)c_{-k}(h)\right) = 0.\end{aligned}$$

Since $|h(t)|^2$ is continuous, $|h(t)|^2 = 0$ for all $t \in [0, 2\pi]$, so that $h(t) = 0$ for all $t \in [0, 2\pi]$. ■

The next result provides a sufficient condition for a function to be equal to its Fourier series.

Theorem 161
Let f be a continuous $2\pi-$periodic function such that

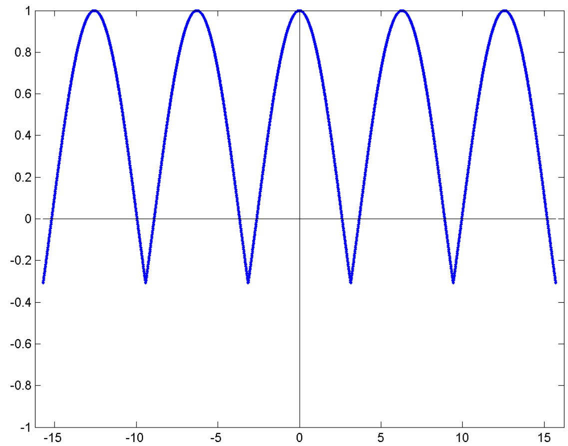
$$\sum_{k\in\mathbb{Z}} |c_k(f)| = M < \infty.$$

Then the Fourier series of f converges absolutely to f on $\mathbb{R}$ and is equal to f on $\mathbb{R}$.

Proof: left as an exercise. ■

Let us take a look at an example.

Example: fix $a \in \mathbb{R}$ and let $f_a(t) = \cos(at)$, $|t| \leq \pi$. Extend f_a to $\mathbb{R}$ by periodicity. What is the Fourier series of f_a? Is it equal to f_a on $\mathbb{R}$? **Solution:** if $a \notin \mathbb{Z}$, f_a is not differentiable (see example below).



If $a \in \mathbb{Z}$ then $\cos(at)$ is already a trigonometric polynomial so the Fourier series of f_a is simply $\cos(at)$. So assume that $a \notin \mathbb{Z}$.

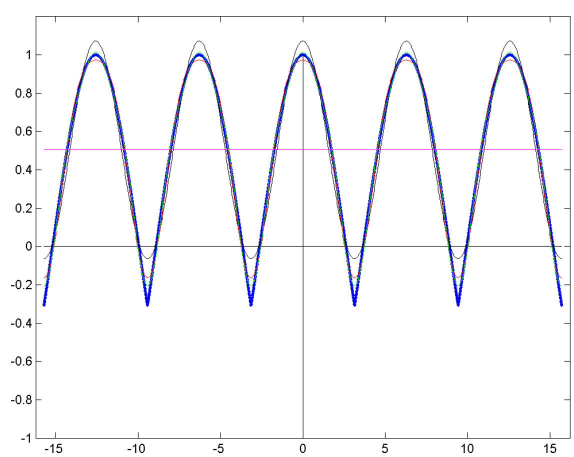
Let $k \in \mathbb{Z}$. Then

$$c_k(f_a) = \frac{1}{2\pi}\int_{-\pi}^{\pi} \cos(at)e^{-ikt}\,dt = \frac{1}{2\pi}\int_{-\pi}^{\pi} \frac{e^{iat} - e^{-iat}}{2}e^{-ikt}\,dt = \frac{a(-1)^k \sin(\pi a)}{\pi(a^2 - k^2)}$$

Using the comparison test with $|c_k(f)| \sim \frac{1}{k^2}$, we see that $\sum_{k\in\mathbb{Z}} |c_k(f)| < \infty$. According to Theorem 161,

$$f_a(t) = \sum_{k\in\mathbb{Z}} \frac{a(-1)^k \sin(\pi a)}{\pi(a^2 - k^2)} e^{ikt}$$

converges absolutely on $\mathbb{R}$. ■



11.2.4 Dirichlet's Convergence Theorem

Let $f:\mathbb{R}\to\mathbb{C}$ be a $2\pi-$periodic Riemann-integrable function. For $k\in\mathbb{Z}$, set

$$e_k(t)=e^{ikt}=\left(e^{it}\right)^k=(e_1(t))^k.$$

Let $N\in\mathbb{N}$. Define

$$S_N(f)(t):=\sum_{k=-N}^{N}c_k(f)e_k(t);$$

$S_N(f)$ is the **partial sum of degree** N for the Fourier series of f.[3] We can write these partial sums as **convolutions**: indeed, we have

$$\begin{aligned}
S_N(f)(t)&:=\sum_{k=-N}^{N}c_k(f)e_k(t)=\sum_{k=-N}^{N}e_k(t)\int f(y)e_k(-y)\,dy\\
&=\int f(y)\left\{\sum_{k=-N}^{N}e_k(t)e_k(-y)\right\}dy\\
&=\int f(y)\left\{\sum_{k=-N}^{N}e_k(t-y)\right\}dy\\
&=\int f(y)K_N(t-y)\,dy:=(\hat{D}_N*f)(t),
\end{aligned}$$

where the **Dirichlet kernel** of order N is, formally,

$$\begin{aligned}
K_N(t)&=\sum_{k=-N}^{N}e_k(t)=\sum_{k=-N}^{N}e^{ikt}=\frac{e^{-iNt}-e^{i(N+1)t}}{1-e^{it}}\\
&=\frac{1}{e^{iNt}}\left(\frac{1-e^{i(2N+1)t}}{1-e^{it}}\right)=\frac{\sin((N+1/2)t)}{\sin(t/2)},\quad\text{when } t\notin 2\pi\mathbb{Z}.
\end{aligned}$$

Proposition 162
The Dirichlet kernel is even, $2\pi-$periodic, $c_0(K_N)=1$, $\int_0^\pi K_N(t)\,dt=\pi$, and

$$K_N(0)=\lim_{t\to}K_N(t)=2N+1.$$

Proof: left as an exercise. ■

The next result is substantially more difficult to prove.

[3]In what follows, we will write $\int:=\frac{1}{2\pi}\int_0^{2\pi}=\frac{1}{2\pi}\int_a^{a+2\pi}$ for any $a\in\mathbb{R}$.

Lemma 163 (Riemann-Lebesgue Lemma)
Let $f:[a,b]\to\mathbb{C}$ *be integrable over* $[a,b]$. *Then* $\lim_{n\to\infty}\int_a^b f(t)e^{int}\,dt=0$.

Proof: left as a (difficult) exercise. ■

We can now state and prove this section's main result.

Theorem 164 (Dirichlet's Convergence Theorem)
Let $f:\mathbb{R}\to\mathbb{C}$ *be piecewise (with a finite number of discontinuities) and* $2\pi-$*periodic. If the following one-sided limits exist* $\forall x\in\mathbb{R}$*:*

$$f(x^{\pm})=\lim_{h\searrow 0}f(x\pm h),\quad f'(x^{\pm})=\lim_{h\searrow 0}\frac{f(x\pm h)-f(x)}{h},$$

then

$$S_N(f)(x)=\sum_{k=-N}^{N}c_k(f)e_k(x)\to\frac{f(x^+)+f(x^-)}{2},\quad \textit{as } N\to\infty.$$

Proof: without loss of generality, we can assume that $x=0$ by translating the variable x to the origin as needed. Consider the sequence of partial sums

$$s_N:=S_N(f)(0)=\sum_{k=-N}^{N}c_k(f)e_k(0)=\sum_{k=-N}^{N}c_k(f).$$

For $N\in\mathbb{N}$, we have

$$s_N=\sum_{|k|\leq N}\int f(t)e^{-ikt}\,dt=\int f(t)K_N(t)\,dt.$$

Since $K_N(t)$ is even, then

$$\int_{-\pi}^{0}f(t)K_N(t)\,dt=\int_0^{\pi}f(-t)K_N(t)\,dt,$$

whence (remember the notation convention for integrals)

$$s_N=\frac{1}{2\pi}\int_0^{\pi}\{f(t)+f(-t)\}K_N(t)\,dt.$$

Write

$$u_N=s_N-\tfrac{f(0^+)+f(0^-)}{2}.$$

Then

$$\begin{aligned} u_N &= \frac{1}{2\pi}\int_0^{\pi}\{f(t)+f(-t)\}K_N(t)\,dt - \tfrac{f(0^+)+f(0^-)}{2}\cdot\frac{1}{\pi}\int_0^{\pi}K_N(t)\,dt \\ &= \frac{1}{2\pi}\int_0^{\pi}\{f(t)+f(-t)-f(0^+)-f(0^-)\}K_N(t)\,dt \\ &= \frac{1}{2\pi}\int_0^{\pi} g(t)\sin((N+1/2)t)\,dt, \end{aligned}$$

where

$$g(t)=\begin{cases} \dfrac{f(t)-f(0^+)+f(-t)-f(0^-)}{\sin(t/2)}, & \text{if } t\in(0,\pi] \\ 0, & \text{otherwise} \end{cases}$$

By construction, g is clearly piecewise continuous on $(0,\pi]$. It is necessarily bounded in a neighbourhood of $t=0$ according to de l'Hôpital's Rule:

$$\lim_{t\searrow 0} g(t)=\lim_{t\searrow 0}\frac{2(f'(t)-f'(-t))}{\cos(t/2)}=2(f'(0^+)+f'(0^-))<\infty.$$

The function g is thus nicely-behaved: it is bounded and piecewise continuous (with at most a finite number of discontinuities) over $[0,\pi]$ and so is integrable on every continuous piece of $[0,\pi]$, using an easy generalization of Theorem 54 (see Chapter 4).

According to the Riemann-Lebesgue lemma 155,

$$\lim_{n\to\infty}\int_0^{\pi} g(t)e^{int}\,dt=0.$$

The relation still holds with the change of variable $n=N+1/2$.

Since $2\pi u_N$ is the imaginary part of $\int_0^{\pi} g(t)e^{i(N+1/2)t}\,dt$, then $2\pi u_N \to 0$ and $s_N\to \frac{f(0^+)+f(0^-)}{2}$ when $N\to\infty$. ■

In other words, if a periodic function f is "**nice enough**" (piecewise C^1), then it is **equal to its Fourier series wherever f is continuous**. At discontinuities of f, the Fourier series converges to the **mean of the one-sided limits**.[4]

Example: let $f:[0,2\pi]\to\mathbb{R}$ be defined by $f(t)=t^2$. Extend f to $\mathbb{R}$ by periodicity. What is the Fourier series of f. Is it equal to f on $\mathbb{R}$?

[4]Be careful: some piecewise C^0 periodic functions have **divergent** Fourier series.

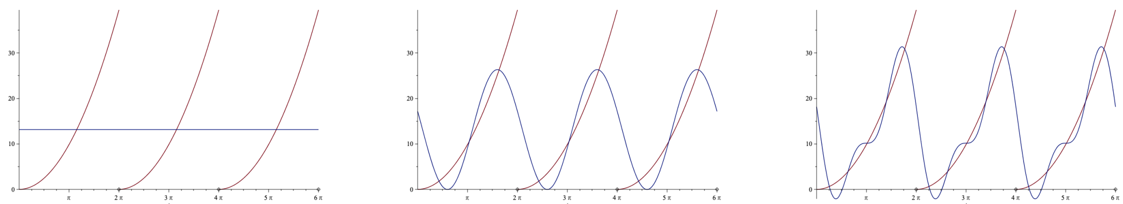
Solution: the Fourier coefficients of f are

$$c_k(f_a) = \frac{1}{2\pi}\int_0^{2\pi} t^2 e^{-ikt}\,dt = \begin{cases} [c]\frac{2}{n^2}(i\pi k+1), & k \neq 0 \\ \frac{4\pi^2}{3}, & k = 0 \end{cases}$$

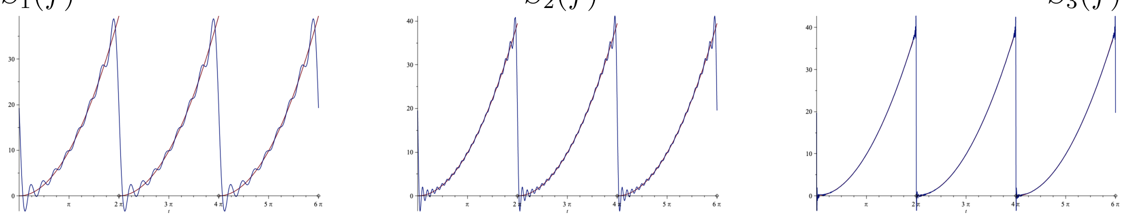
According to Dirichlet's convergence theorem,

$$\sum_{k\in\mathbb{Z}} c_k(f)e^{ikt} = \frac{4\pi^2}{3} + \sum_{k\in\mathbb{Z}^\times} \frac{2}{k^2}(i\pi k+1)e^{ikt}$$

converges (at least pointwise) to t^2 for $t \notin 2\pi\mathbb{Z}$, and to $\frac{f(2\pi)+f(0)}{2} = 2\pi^2$ for $t \in 2\pi\mathbb{Z}$, since f is piecewise C^1.



$S_1(f)$ $S_2(f)$ $S_3(f)$



$S_8(f)$ $S_{20}(f)$ $S_{200}(f)$

The convergence turns out to be uniform on $[2\pi\ell+\delta, 2\pi(\ell+1)-\delta]$, for all $\delta \in (0,\pi)$, $\ell \in \mathbb{Z}$ (more on this in the next section), but only pointwise over $\mathbb{R}$ as a whole, in keeping with Theorem 164. ■

Notice the overshooting of the partial sums as $t \to 2\pi\ell$, $\ell \in \mathbb{Z}$, which does not seem to dampen when $N \to \infty$. This "universal" behaviour at discontinuities is termed **Gibbs' Phenomenon** (contrast the behaviour of the Fourier series of t^2 with that of $\cos(at)$ discussed earlier).

The explanation of the problem is linked with the $\limsup$ and $\liminf$ of the partial sums $S_n(f)(x_N)$ at points x_N that approach a discontinuity at x_0, but we will not discuss this any further.

11.2.5 Quadratic Mean Convergence

The set of $2\pi-$periodic piecewise continuous functions from $\mathbb{R}$ to $\mathbb{C}$ is an **inner product space** together with

$$(f \mid g) = \frac{1}{2\pi}\int_0^{2\pi} f(t)\overline{g(t)}\,dt,$$

with **associated norm** $\|f\|_2 = \sqrt{(f \mid f)}$.

Note that for $\mu,\nu \in \mathbb{Z}$, we have

$$(e_\mu \mid e_\nu) = \frac{1}{2\pi}\int_0^{2\pi} e^{i\mu t}e^{-i\nu t}\,dt = \frac{1}{2\pi}\int_0^{2\pi} e^{i(\mu-\nu)t}\,dt = \delta_{\mu,\nu} = \begin{cases} 0, & \mu \neq \nu \\ 1, & \mu = \nu \end{cases}$$

For a given $N \in \mathbb{N}$ and a function f in the inner product space of the previous page, consider the partial sum

$$S_N(f) = \sum_{|k|\leq N} c_k(f)e_k(t).$$

For any $|k| \leq N$, we must have

$$c_k(f) = \frac{1}{2\pi}\int_0^{2\pi} f(t)e^{-ikt}\,dt = (f \mid e_k).$$

But

$$(S_N(f) \mid e_k) = \sum_{|\ell|\leq N} c_\ell(f)(e_\ell \mid e_k) = \sum_{|\ell|\leq N} c_\ell(f)\delta_{\ell,k} = c_k(f).$$

Thus, $(f - S_N(f) \mid e_k) = 0$ for all $|k| \leq N$ and we can write

$$f = S_N(f) + (f - S_N(f)),$$

with $S_N(f) \in \mathcal{P}_N = \text{Span}\{e_k \mid -N \leq k \leq N\}$ and $f - S_N(f) \in \mathcal{P}_N^\perp$.

Note furthermore that since $(S_N \mid f - S_N(f)) = 0$, then

$$\begin{aligned}\|f\|_2^2 &= (f \mid f) = (S_N(f) + (f - S_N(f)) \mid S_N(f) + (f - S_N(f))) \\ &= (S_N(f) \mid S_N(f)) + 2\text{Re}\,\underbrace{(S_N(f) \mid f - S_N(f))}_{=0} + (f - S_N(f) \mid f - S_N(f)) \\ &= \|S_N(f)\|_2^2 + \|f - S_N(f)\|_2^2.\end{aligned}$$

For any other function $g \in \mathcal{P}_N$, we see that

$$\begin{aligned}\|f-g\|^2 &= \|\underbrace{f - S_N(f)}_{\in\mathcal{P}_N^\perp} + \underbrace{S_N(f) - g}_{\in\mathcal{P}_N}\|_2^2 \\ &= \|f - S_N(f)\|_2^2 + \|S_N(f) - g\|_2^2 \geq \|f - S_N(f)\|_2^2.\end{aligned}$$

Since g was arbitrary,

$$\inf_{g\in\mathcal{P}_N}\|f-g\|_2^2=\|f-S_N(f)\|_2^2=\|f\|_2^2-\|S_N(f)\|_2^2. \tag{11.1}$$

The partial sum $S_N(f)$ is thus the nearest trigonometric polynomial to f in $\mathcal{P}_N$, in the sense of the **quadratic mean**.

Theorem 165 (Parseval's Identity)
Let f be a $2\pi-$periodic piecewise continuous function from $\mathbb{R}$ to $\mathbb{C}$. Then

$$\frac{1}{2\pi}\int_0^{2\pi}|f(t)|^2\,dt=\sum_{k=-\infty}^{\infty}|c_k(f)|^2.$$

Proof: as $|f|^2$ is Riemann-integrable on $[0,2\pi]$, the convergence of the series will be assured once the equality is established. By construction,

$$\begin{aligned}\|S_N(f)\|_2^2&=\left(\sum_{|k|\leq N}c_k(f)e^{ikt}\,\middle|\,\sum_{|\ell|\leq N}c_\ell(f)e^{i\ell t}\right)=\sum_{k,\ell=-N}^{N}c_k(f)\overline{c_\ell(f)}(e_k\mid e_\ell)\\&=\sum_{k,\ell=-N}^{N}c_k(f)\overline{c_\ell(f)}\delta_{k,\ell}=\sum_{k=-N}^{N}|c_k(f)|^2.\end{aligned}$$

The sequence of infimums given in (11.1) by

$$(x_N)=\left(\inf_{g\in\mathcal{P}_N}\{\|f-g\|_2^2\}\right)$$

is bounded below by 0.

Let $N\in\mathbb{N}$. Clearly, $\|S_N(f)\|_2^2\leq\|S_{N+1}(f)\|_2^2$, and so

$$x_N=\|f-S_N(f)\|_2^2=\|f\|_2^2-\|S_N(f)\|_2^2\geq\|f\|_2^2-\|S_{N+1}(f)\|_2^2=x_{N+1}.$$

Thus (x_N) is a decreasing and bounded sequence; as such, it converges to $0\leq x_*=\inf\{x_N\mid N\in\mathbb{N}\}$ by the bounded monotone convergence theorem.

In particular, this means that

$$x_*=\lim_{N\to\infty}x_N=\|f\|_2^2-\lim_{N\to\infty}\|S_N(f)\|_2^2=\frac{1}{2\pi}\int_0^{2\pi}|f(t)|^2\,dt\;-\sum_{k=-\infty}^{\infty}|c_k(f)|^2,$$

which guarantees the convergence of the series, as $|f|^2$ is Riemann-integrable over $[0,2\pi]$ (being continuous).

Write $\mathcal{P} = \bigcup_{N\in\mathbb{N}} \mathcal{P}_N$. Since $\mathcal{P}_N \subseteq \mathcal{P}$ for all $N \in \mathbb{N}$, we have

$$\inf_{g\in\mathcal{P}} \|f-g\|_2^2 \leq \inf_{g\in\mathcal{P}_N} \|f-g\|_2^2 = x_N, \quad \text{for all } N \in \mathbb{N},$$

which implies that

$$0 \leq \inf_{g\in\mathcal{P}} \|f-g\|_2^2 \leq x_*.$$

Conversely, $x_* \leq \|f-g\|_2^2$ for all $g \in \mathcal{P}_N$, $N \in \mathbb{N}$. Thus $x_* \leq \|f-g\|_2^2$ for all $g \in \mathcal{P}$, so that

$$x_* \leq \inf_{g\in\mathcal{P}} \|f-g\|_2^2.$$

Combining these, we obtain

$$\inf_{g\in\mathcal{P}} \|f-g\|_2^2 = \frac{1}{2\pi}\int_0^{2\pi} |f(t)|^2\,dt - \sum_{k=-\infty}^{\infty} |c_k(f)|^2.$$

Let $\varepsilon > 0$. As f is a $2\pi-$periodic piecewise continuous function, we can find a $2\pi-$periodic continuous function f_c such that

$$\|f-f_c\|_2 < K\varepsilon, \quad \text{for some } K > 0.$$

If f is constant, simply set $f_c = f$; we do the same if f is continuous.

Otherwise, assume that f admits m discontinuities at

$$x_1 < \ldots < x_m \in (\delta, 2\pi+\delta), \quad \text{for some } \delta > 0,$$

and denote the closed ε^2-neighbourhood around x_α by

$$B_{\alpha,\varepsilon^2} = [y_{\alpha,\varepsilon^2}, y_{\alpha,\varepsilon^2} + 2\varepsilon^2],$$

for $\alpha = 1,\ldots,m$, and their union by B_{ε^2} (restrict ε as needed to ensure that the $B_{\alpha,\varepsilon^2} = [y_{\alpha,\varepsilon^2}, y_{\alpha,\varepsilon^2} + 2\varepsilon^2]$ do not overlap).

Outside of B_{ε^2} but in $[\delta, 2\pi+\delta]$, define $f_c \equiv f$. In each of the $B_{\alpha,\varepsilon^2} \cap [\delta, 2\pi+\delta]$, let f_c be the linear function joining the points

$$(y_{\alpha,\varepsilon^2}, f(y_{\alpha,\varepsilon^2})) \quad \text{and} \quad (y_{\alpha,\varepsilon^2} + 2\varepsilon^2, f(y_{\alpha,\varepsilon^2} + 2\varepsilon^2)).$$

The function $f_c : [\delta, 2\pi+\delta] \to \mathbb{C}$ is "clearly" continuous, and can be extended to a $2\pi-$periodic continuous function over $\mathbb{R}$.

In particular, $|f-f_c|^2$ is real-valued and continuous over $[\delta, 2\pi+\delta]$. Consequently, the latter reaches its maximum $M > 0$ somewhere on $[\delta, 2\pi+\delta]$, by the max/min theorem.

Thus, for any $\delta > 0$,

$$\begin{aligned}\|f - f_c\|_2^2 &= \frac{1}{2\pi}\int_{\delta}^{2\pi+\delta} |f(t) - f_c(t)|^2\,dt = \frac{1}{2\pi}\sum_{\alpha=1}^{m}\int_{B_{\alpha,\varepsilon^2}} |f(t) - f_c(t)|^2\,dt \\ &\leq \frac{1}{2\pi}\sum_{\alpha=1}^{m}\int_{B_{\alpha,\varepsilon^2}} M\,dt = \frac{1}{2\pi}\sum_{\alpha=1}^{m} 2\varepsilon^2 \cdot M = \underbrace{\frac{mM}{\pi}}_{>0}\varepsilon^2 := K^2\varepsilon^2\end{aligned}$$

According to the Stone-Weierstrass theorem (see Chapter 23), the set of $2\pi-$periodic trigonometric polynomials $\mathcal{P}$ is dense in the set of $2\pi-$periodic continuous functions w.r.t. to $\|\cdot\|_2$, and so $\exists g \in \mathcal{P}$ with $\|f_c - g\|_2 < \varepsilon$.

Putting this together, we see that

$$\|f - g\|_2 \leq \|f - f_c\|_2 + \|f_c - g\|_2 < K\varepsilon + \varepsilon = (K+1)\varepsilon.$$

Thus

$$\inf_{g\in\mathcal{P}} \|f - g\|_2 < (K+1)\varepsilon \quad \text{for all } \varepsilon \implies \inf_{g\in\mathcal{P}} \|f - g\|_2 = 0. \qquad \blacksquare$$

Parseval's identity remains valid for locally Riemann-integrable functions ($\int_K |f|\,dt < \infty$ for all $K \subseteq_K [0, 2\pi]$), instead of piecewise continuous, with multiple consequences: the series

$$\sum_{k\in\mathbb{Z}} |c_k(f)|^2$$

converges, which shows that $|c_k(f)|^2 \to 0$, and thus $c_k(f) \to 0$ as $k \to \pm\infty$ (by the Riemann-Lebesgue lemma). It can also be used to show that any $2\pi-$periodic continuous function $f : \mathbb{R} \to \mathbb{C}$ whose Fourier series converges uniformly on $\mathbb{R}$ must be equal to said series (compare with Dirichlet's convergence theorem).

11.3 Exercises

1. Let (g_n) be a sequence of functions. Show that $\sum g_n$ converges absolutely if and only if $\exists (a_n) \subseteq \mathbb{R}^+$ such that $\sum a_n$ converges and $\|g_n\|_\infty \leq a_n$ for all n. Use that result to show that the series of functions $\sum g_n$, where $g_n : [0,1] \to \mathbb{R}$ is defined by $g_n(x) = \frac{x^n}{n^2}$, is absolutely convergent on $[0,1]$.
2. For each of the theorems of Section 11.1.1 (save for Theorem 152), find an example showing that the result does not hold if uniform convergence is replaced by pointwise convergence.
3. Prove Theorems 152, 153, and 161, as well as Propositions 156 and 158.

4. Find some examples showing that the result of Theorem 152 does not hold in general if absolute convergence is replaced by a weaker type of convergence.

5. Let $g_n : \mathbb{R} \to \mathbb{R}$ be defined by $g_n(x) = \frac{x^n}{n!}$ for each $n \in \mathbb{N}$. Show that each of the following series of functions converges absolutely on $\mathbb{R}$.

   a) $S = \sum(-1)^{n+1} g_{2n+1}$
   b) $C = \sum(-1)^n g_{2n}$
   c) $E = \sum g_n$

6. Let S, C, E be as in the previous question. Using the appropriate theorems, show that for any $x \in \mathbb{R}$ show that $S'(x) = C(x)$, $C'(x) = -S(x)$, and $E'(x) = E(x)$.

7. Find examples showing that the three conditions in the statement of Proposition 155 are independent from one another.

8. Improve the bound in Proposition 158 by showing that

$$\|f\|_2 = \left(\sum_{k\in\mathbb{Z}} |c_k(f)|^2\right)^{1/2}.$$

9. Show that the function $f : [0, 2\pi] \to \mathbb{R}$ defined by $f(t) = \sum_{k\geq 1} \frac{\sin(kt)}{k}$ is not continuous on $[0, 2\pi]$.

10. Using the Fourier series of the cosine, show that $\pi \cot(a\pi) = \sum_{k\in\mathbb{Z}} \frac{a}{a^2-k^2}$ for all $a \notin \mathbb{Z}$ (also known as **Euler's Formula**).

11. Prove the properties of the Dirichlet kernel (Proposition 11.2.4).

12. Show that $(f \mid g)$ (see page 289) defines an inner product on the set of $2\pi-$periodic piecewise continuous functions from $\mathbb{R}$ to $\mathbb{C}$.

13. Prove the Riemann-Lebesgue lemma without using Parseval's identity.

14. Show that any $2\pi-$periodic continuous function $f : \mathbb{R} \to \mathbb{C}$ whose Fourier coefficients are all 0 must be the zero function.

15. Let $(a_n) \subseteq \mathbb{C}$ be such that $a_n \to \ell$ and let $(\varepsilon_n) \subseteq \mathbb{R}^+$ be a divergent sequence. Define a sequence $(b_n) \subseteq \mathbb{C}$ by

$$b_n = \frac{\sum_{i=1}^n a_i\varepsilon_i}{\sum_{i=1}^n \varepsilon_i}.$$

    Show that $b_n \to \ell$.

16. a) Let (f_n) be the sequence of functions defined by

$$f_n : \mathbb{R}_0^+ \to \mathbb{R}, \quad f_n(x) = \begin{cases} \left(1 - \frac{x}{n}\right)^n & x \in [0, n] \\ 0 & x > n \end{cases}$$

    Show that $f_n \rightrightarrows f$ on $\mathbb{R}_0^+$, where $f : \mathbb{R}_0^+ \to \mathbb{R}$ is defined by $f(x) = e^{-x}$.

b) Let $U \subseteq_K \mathbb{C}$ and let (f_n) be the sequence of functions defined by

$$f : \mathbb{C} \to \mathbb{C}, \quad f(z) = \left(1 + \frac{z}{n}\right)^n.$$

Show that $f_n \rightrightarrows f$ on K, where $f : \mathbb{C} \to \mathbb{C}$ is defined by $f(z) = e^z$.

17. For any $n \in \mathbb{N}^\times$, let $u_n : \mathbb{R}_0^+ \to \mathbb{R}$ be defined by $u(x) = \frac{x}{n^2+x^2}$.

   a) Show that $\sum u_n \to f$ for some $f \in \mathcal{C}(\mathbb{R}_0^+, \mathbb{R})$, but that $\sum u_n \not\rightrightarrows f$ on $\mathbb{R}_0^+$.
   b) Show that $\sum(-1)^n u_n \rightrightarrows g$ on $\mathbb{R}_0^+$ for some $g \in \mathcal{C}(\mathbb{R}_0^+, \mathbb{R})$, but that $\sum(-1)^n u_n$ is not absolutely convergent on $\mathbb{R}_0^+$.

18. What can you say about a function $f : \mathbb{R} \to \mathbb{R}$ which is the uniform limit of a sequence of polynomials (P_n)?

19. Consider the sequence of functions $(f_n) \subseteq \mathcal{C}([0, \pi/2], \mathbb{R})$ defined by $f_n(x) = \cos^n x \sin x$ for all $n \in \mathbb{N}$.

   a) Let $\mathcal{O} : [0, \pi/2] \to \mathbb{R}$ be the zero function. Show that $f_n \rightrightarrows \mathcal{O}$ on $[0, \pi/2]$.
   b) Consider the sequence of functions (g_n) defined by $g_n = (n+1)f_n$. Let $\delta > 0$. Show that $g_n \rightrightarrows \mathcal{O}$ on $[\delta, \pi/2]$ but that

$$\int_0^{\pi/2} g_n(t)\, dt \not\to 0.$$

20. Theses results are due to Dini.

   a) Let $(f_n) \in \mathcal{C}([a,b], \mathbb{R})$ be an increasing sequence of functions (i.e. for all $x \in [a,b]$ and for all $n \in \mathbb{N}$, we have $f_n(x) \leq f_{n+1}(x)$). If $f_n \to f$ on $[a,b]$ where $f \in \mathcal{C}([a,b], \mathbb{R})$, show that $f_n \rightrightarrows f$ on $[a,b]$.
   b) Let $(f_n) \in \mathcal{C}([a,b], \mathbb{R})$ be a sequence of increasing functions (i.e. for all $x \geq y \in [a,b]$ and for all $n \in \mathbb{N}$, we have $f_n(x) \geq f_n(y)$). If $f_n \to f$ on $[a,b]$ where $f \in \mathcal{C}([a,b], \mathbb{R})$, show that $f_n \rightrightarrows f$ on $[a,b]$.

21. Determine whether $\sum \mathbf{x}_n$ converges in $(\mathbb{R}^2, \|\cdot\|_2)$, where

$$\mathbf{x}_n = \left(\frac{(\sin n)^n}{n^2}, \frac{1}{n^2}\right).$$

   If so, does $\sum \mathbf{x}_n$ converge absolutely?

22. Compute the values of the following convergent series

$$\sum_{n=1}^{\infty} \frac{1}{n^2}, \quad \sum_{n=1}^{\infty} \frac{1}{(2n-1)^2}, \quad \sum_{n=1}^{\infty} \frac{1}{n^4},$$

   using the $2\pi-$periodic function defined by $f(x) = 1 - x^2/\pi^2$ over the interval $[-\pi, \pi]$.

23. Prepare a 2-page summary of this chapter, with important definitions and results.

Part III

Vector Analysis and Differential Forms

Chapter 12

Alternating Multilinear Forms

In order to define the notion of **differential forms** (and to learn how to integrate them), we need concepts from linear algebra. In this chapter, E is a finite dimensional vector space over $\mathbb{R}$ (i.e., $\dim(E) = n \Longrightarrow E \simeq \mathbb{R}^n$).

12.1 Linear Algebra Notions

A **(linear)** $1-$**form over** E is a linear map $f : E \to \mathbb{R}$; a **(linear)** $p-$**form over** E is a linear map $f : E^p = E \times \cdots \times E \to \mathbb{R}$ which is linear in each of its arguments.

Examples

1. The projection map $f_1 : \mathbb{R}^n \to \mathbb{R}$, defined by $f_1(\mathbf{x}) = f_1(x_1, \ldots, x_n) = x_1$ is a $1-$form over $\mathbb{R}^n$. Generally, the projection $f_i : \mathbb{R}^n \to \mathbb{R}$ defined by $f_i(\mathbf{x}) = x_i$ is a $1-$form over $\mathbb{R}^n$ for all $i = 1, \ldots, n$.

   If $\mathcal{B} = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is a basis of E, then for any $\mathbf{x} \in E$ we can write

   $$\mathbf{x} = x_1\mathbf{e}_1 + \cdots + x_1\mathbf{e}_1$$

   and the projection $f_i^B : E \to \mathbb{R}$ defined by $f_i^B(\mathbf{x}) = x_i$ is a $1-$form over E. $\square$

2. The inner product $(\cdot \mid \cdot) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ defined by

   $$(\mathbf{x} \mid \mathbf{y}) = ((x_1, \ldots, x_n) \mid (y_1, \ldots, y_n)) = \sum_{i=1}^{n} x_i y_i$$

   is a **(bilinear)** $2-$form over $\mathbb{R}^n$.

   If $(\mathbf{x} \mid \mathbf{y}) = (\mathbf{y} \mid \mathbf{x})$ for all $\mathbf{x}, \mathbf{y} \in E$, the $2-$form is **symmetric**. $\square$

3. The $2-$determinant $\det : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ defined by

$$\det(\mathbf{x}, \mathbf{y}) = \det \begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \end{pmatrix} = x_1 y_2 - x_2 y_1$$

is a bilinear form over $\mathbb{R}^2$, but it is not symmetric since $\det(\mathbf{x}, \mathbf{y}) = -\det(\mathbf{y}, \mathbf{x})$ for all $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$. Note as well that $\det(\mathbf{x}, \mathbf{x}) = 0$ for all $\mathbf{x} \in \mathbb{R}^2$. □

A $p-$form f over E is **alternating** if $f(\mathbf{x}_1, \ldots, \mathbf{x}_p) = 0$ whenever $\mathbf{x}_i = \mathbf{x}_j$ for some $i < j$.

Example: $\det : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ is an alternating bilinear form. More generally,

$$\det : \underbrace{\mathbb{R}^n \times \cdots \times \mathbb{R}^n}_{n \text{ times}} \to \mathbb{R}$$

is an alternating linear $n-$form. □

Let $f : \mathbb{R}^2 \times \mathbb{R}^2 \to \mathbb{R}$ be an alternating bilinear form on $\mathbb{R}^2$. If $\{\mathbf{e}_1, \mathbf{e}_2\}$ is a basis of $\mathbb{R}^2$, then f is **completely determined** by the value taken by $f(\mathbf{e}_1, \mathbf{e}_2)$. Indeed, let $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$, Then

$$\begin{aligned} f(\mathbf{x}, \mathbf{y}) &= f(x_1\mathbf{e}_1 + x_2\mathbf{e}_2, y_1\mathbf{e}_1 + y_2\mathbf{e}_2) = x_1 f(\mathbf{e}_1, y_1\mathbf{e}_1 + y_2\mathbf{e}_2) + x_2 f(\mathbf{e}_2, y_1\mathbf{e}_1 + y_2\mathbf{e}_2) \\ &= x_1y_1 \underbrace{f(\mathbf{e}_1, \mathbf{e}_1)}_{=0} + x_1y_2 f(\mathbf{e}_1, \mathbf{e}_2) + x_2y_1 \underbrace{f(\mathbf{e}_2, \mathbf{e}_1)}_{=-f(\mathbf{e}_1, \mathbf{e}_2)} + x_2y_2 \underbrace{f(\mathbf{e}_2, \mathbf{e}_2)}_{=0} \\ &= (x_1y_2 - x_2y_1) f(\mathbf{e}_1, \mathbf{e}_2) = \det \begin{pmatrix} x_1 & x_2 \\ y_1 & y_2 \end{pmatrix} f(\mathbf{e}_1, \mathbf{e}_2). \end{aligned}$$

Let $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ be a basis of $E = \mathbb{R}^n$ and let $\{\mathbf{x}_1, \ldots, \mathbf{x}_n\} \subseteq E = \mathbb{R}^n$. For $1 \leq i \leq n$, Write

$$\mathbf{x}_i = \sum_{j=1}^{n} s_{i,j} \mathbf{e}_j.$$

If $f : E^n \to \mathbb{R}$ is an alternating (linear) $n-$form, then

$$f(\mathbf{x}_1, \ldots, \mathbf{x}_n) = \det \begin{pmatrix} x_{1,1} & \cdots & x_{1,n} \\ \vdots & \ddots & \vdots \\ x_{n,1} & \cdots & x_{n,n} \end{pmatrix} f(\mathbf{e}_1, \ldots, \mathbf{e}_n) = \det \begin{pmatrix} \mathbf{x}_1 & \cdots & \mathbf{x}_n \end{pmatrix}^\top f(\mathbf{e}_1, \ldots, \mathbf{e}_n).$$

Let $f_1, \ldots, f_p$ be p linear $1-$forms over E.[1] Define $f : E^p \to \mathbb{R}$ by

$$f(\mathbf{x}_1, \ldots, \mathbf{x}_p) = f_1(\mathbf{x}_1) \cdots f_p(\mathbf{x}_p).$$

Then f is the **tensor product** of the f_i; it is a linear $p-$form over E, which we usually denote by $f = f_1 \otimes \cdots \otimes f_p$.

[1]We can also write this as $f_1, \ldots, f_p \in E^*$, where $E^* = \{f : E \to \mathbb{R} \mid f \text{ linear}\}$ is the **dual space** of E.

If $\mathcal{B} = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is a basis of E, then for $1 \leq i \leq n$, we define the **linear functionals** $\mathbf{e}_i^* \in E^*$ by

$$\mathbf{e}_i^*(\mathbf{x}) = \mathbf{e}_i^*(x_1\mathbf{e}_1 + \cdots + x_n\mathbf{e}_n) = x_1\mathbf{e}_i^*(\mathbf{e}_1) + \cdots + x_n\mathbf{e}_i^*(\mathbf{e}_n) = x_i\mathbf{e}_i^*(\mathbf{e}_i) = x_i$$

for all $\mathbf{x} = (x_1, \ldots, x_n) \in E$. In that case, the set

$$\left\{\mathbf{e}_{i_1}^* \otimes \cdots \otimes \mathbf{e}_{i_p}^* \mid i_j \in \{1, \ldots, n\}\right\}$$

forms a basis of the vector space of $p-$forms over E, and $\dim(\{p - \text{forms over } E\}) = n^p$.

12.2 Anti-Symmetric Forms

In introductory linear algebra and group theory courses, we learn that if $A = (a_{i,j}) \subseteq \mathbb{M}_n(\mathbb{R})$, then we can write the determinant of A using the **Laplace expansion**:

$$\det(A) = \sum_{\sigma \in S_n} \epsilon(\sigma) a_{1,\sigma(1)} \cdots a_{n,\sigma(n)},$$

where S_n is the **permutation group** on $\{1, \ldots, n\}$ (whence $|S_n| = n!$) and $\epsilon : S_n \to \{\pm 1\}$ is the **signature** of a permutation σ (more on this in the first footnote of Section 12.3).

Proposition 166
Let f be a linear $p-$form over E. If $g : E^p \to \mathbb{R}$ is defined by

$$g(\mathbf{x}_1, \ldots, \mathbf{x}_p) = \sum_{\sigma \in S_p} \epsilon(\sigma) f(\mathbf{x}_{\sigma(1)}, \ldots, \mathbf{x}_{\sigma(p)}),$$

then g is an alternating $p-$form.

Proof: we only prove the statement for $p = 2$. The proof for $p \geq 3$ is left as an exercise.

Let $p = 2$. Then $S_2 = \{\text{id}, \sigma = \begin{pmatrix} 1 & 2 \end{pmatrix}\}$ and we have $\epsilon(\text{id}) = 1$ and $\epsilon(\sigma) = -1$. Therefore,

$$g(\mathbf{x}_1, \mathbf{x}_2) = f(\mathbf{x}_1, \mathbf{x}_2) - f(\mathbf{x}_2, \mathbf{x}_1).$$

Clearly $g(\mathbf{x}, \mathbf{x}) = 0$, and so g is alternating. ■

The alternating $p-$form g in Proposition 166 is the **anti-symmetric form built from** f.

Let $f_1, \ldots, f_p$ be linear $1-$forms over E. The anti-symmetric form built from the tensor product $f_1 \otimes \cdots \otimes f_p$ is the **wedge product of** $f_1, \ldots, f_p$, denoted by $(f_1 \wedge \cdots \wedge f_p)$.[2]

[2]Formally, we should be using the brackets around the wedge product (and the tensor product) of linear forms, but we will often omit them to simplify the notation.

By definition, then, we have

$$\begin{aligned}(f_1\wedge\cdots\wedge f_p)(\mathbf{x}_1,\ldots,\mathbf{x}_p) &= \sum_{\sigma\in S_p}\epsilon(\sigma)(f_1\otimes\cdots\otimes f_p)(\mathbf{x}_{\sigma(1)},\ldots,\mathbf{x}_{\sigma(p)})\\ &= \sum_{\sigma\in S_p}\epsilon(\sigma)f_1(\mathbf{x}_{\sigma(1)})\cdots f_p(\mathbf{x}_{\sigma(p)}) = \det\begin{pmatrix} f_1(\mathbf{x}_1) & \cdots & f_1(\mathbf{x}_p)\\ \vdots & \ddots & \vdots\\ f_p(\mathbf{x}_1) & \cdots & f_p(\mathbf{x}_p)\end{pmatrix}.\end{aligned}$$

A few examples will help to illustrate the concept.

Examples: consider the case $p = 2$; let f_1, f_2 be linear $1-$form over $E = \mathbb{R}^2$ and $\mathbf{x}_1, \mathbf{x}_2 \in E$. Then:

1. $f_1 \wedge f_2(\mathbf{x}_1, \mathbf{x}_2) = f_1(\mathbf{x}_1)f_2(\mathbf{x}_2) - f_1(\mathbf{x}_2)f_2(\mathbf{x}_1)$.
2. $f_2 \wedge f_1(\mathbf{x}_1, \mathbf{x}_2) = f_2(\mathbf{x}_1)f_1(\mathbf{x}_2) - f_2(\mathbf{x}_2)f_1(\mathbf{x}_1) = -f_1 \wedge f_2(\mathbf{x}_1, \mathbf{x}_2)$.
3. $f_1 \wedge f_1(\mathbf{x}_1, \mathbf{x}_2) = f_2 \wedge f_2(\mathbf{x}_1, \mathbf{x}_2) = 0$. □

Generally, if $f_i = f_j$ for some $i =\neq j$, then $f_1 \wedge \cdots \wedge f_p = 0$. Furthermore, if $\sigma \in S_p$, then

$$f_{\sigma(1)} \wedge \cdots \wedge f_{\sigma(p)} = \epsilon(\sigma) f_1 \wedge \cdots \wedge f_p.$$

Example: let $\mathcal{B} = \{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ be a basis of E (i.e., $n = \dim(E) = 3$) and let $g : E \times E \to \mathbb{R}$ be a bilinear alternating form (i.e., $p = 2$). Then

$$g(\mathbf{x}, \mathbf{y}) = g\left(\sum_{i=1}^{3} x_i\mathbf{e}_i, \sum_{j=1}^{3} y_j\mathbf{e}_j\right) = \sum_{i,j=1}^{3} x_iy_jg(\mathbf{e}_i, \mathbf{e}_j).$$

Since g is alternating, we must have:

$$g(\mathbf{e}_i, \mathbf{e}_j) = -g(\mathbf{e}_j, \mathbf{e}_i), \quad g(\mathbf{e}_i, \mathbf{e}_i) = 0, \quad \text{for all } i, j = 1, \ldots, 3.$$

Thus,

$$\begin{aligned} g(\mathbf{x}, \mathbf{y}) &= x_1y_2g(\mathbf{e}_1, \mathbf{e}_2) + x_1y_3g(\mathbf{e}_1, \mathbf{e}_3) + x_2y_3g(\mathbf{e}_2, \mathbf{e}_3)\\ &\quad - x_2y_1g(\mathbf{e}_1, \mathbf{e}_2) - x_3y_1g(\mathbf{e}_1, \mathbf{e}_3) - x_3y_2g(\mathbf{e}_2, \mathbf{e}_3)\\ &= (x_1y_2 - x_2y_1)g(\mathbf{e}_1, \mathbf{e}_2) + (x_1y_3 - x_3y_1)g(\mathbf{e}_1, \mathbf{e}_3) + (x_2y_3 - x_3y_2)g(\mathbf{e}_2, \mathbf{e}_3). \end{aligned}$$

But note that for $i < j$, we have

$$\mathbf{e}_i^* \wedge \mathbf{e}_j^*(\mathbf{x}, \mathbf{y}) = \mathbf{e}_i^*(\mathbf{x})\mathbf{e}_j^*(\mathbf{y}) - \mathbf{e}_i^*(\mathbf{y})\mathbf{e}_j^*(\mathbf{x}) = x_iy_j - x_jy_i.$$

Combining the last two results, we have

$$g(\mathbf{x},\mathbf{y}) = g(\mathbf{e}_1,\mathbf{e}_2)\mathbf{e}_1^*\wedge\mathbf{e}_2^* + g(\mathbf{e}_1,\mathbf{e}_3)\mathbf{e}_1^*\wedge\mathbf{e}_3^* + g(\mathbf{e}_2,\mathbf{e}_3)\mathbf{e}_2^*\wedge\mathbf{e}_3^*.$$

Consequently, g is a linear combination of the wedge products $\{\mathbf{e}_2^*\wedge\mathbf{e}_3^* \mid i<j\}$. Furthermore, $\{\mathbf{e}_1^*\wedge\mathbf{e}_2^*, \mathbf{e}_1^*\wedge\mathbf{e}_3^*, \mathbf{e}_2^*\wedge\mathbf{e}_3^*\}$ are linearly independent.

Indeed, suppose that

$$(d_{1,2}\mathbf{e}_1^*\wedge\mathbf{e}_2^* + d_{1,3}\mathbf{e}_1^*\wedge\mathbf{e}_3^* + d_{2,3}\mathbf{e}_2^*\wedge\mathbf{e}_3^*)(\mathbf{x},\mathbf{y}) = 0 \quad \text{for all } \mathbf{x},\mathbf{y}.$$

In particular, this would hold for $(\mathbf{x},\mathbf{y}) = (\mathbf{e}_1,\mathbf{e}_2)$, and so

$$\begin{aligned}
0 &= d_{1,2}\mathbf{e}_1^*\wedge\mathbf{e}_2^*(\mathbf{e}_1,\mathbf{e}_2) + d_{1,3}\mathbf{e}_1^*\wedge\mathbf{e}_3^*(\mathbf{e}_1,\mathbf{e}_3) + d_{2,3}\mathbf{e}_2^*\wedge\mathbf{e}_3^*(\mathbf{e}_2,\mathbf{e}_3)\\
&= d_{1,2}(\mathbf{e}_1^*(\mathbf{e}_1)\mathbf{e}_2^*(\mathbf{e}_2) - \mathbf{e}_1^*(\mathbf{e}_2)\mathbf{e}_2^*(\mathbf{e}_1)) + d_{1,3}((\mathbf{e}_1^*(\mathbf{e}_1)\mathbf{e}_3^*(\mathbf{e}_2) - \mathbf{e}_1^*(\mathbf{e}_2)\mathbf{e}_3^*(\mathbf{e}_1))\\
&\qquad + d_{2,3}((\mathbf{e}_2^*(\mathbf{e}_1)\mathbf{e}_3^*(\mathbf{e}_2) - \mathbf{e}_2^*(\mathbf{e}_2)\mathbf{e}_3^*(\mathbf{e}_1))\\
&= d_{1,2}(1\cdot 1 - 0\cdot 0) + d_{1,3}(1\cdot 0 - 0\cdot 0) + d_{2,3}(0\cdot 0 - 0\cdot 0) = d_{1,2} \implies d_{1,2} = 0.
\end{aligned}$$

Similarly, using $(\mathbf{x},\mathbf{y}) = (\mathbf{e}_1,\mathbf{e}_3)$ and $(\mathbf{x},\mathbf{y}) = (\mathbf{e}_2,\mathbf{e}_3)$ yields $d_{1,3} = d_{2,3} = 0$.

Thus $\{\mathbf{e}_1^*\wedge\mathbf{e}_2^*, \mathbf{e}_1^*\wedge\mathbf{e}_3^*, \mathbf{e}_2^*\wedge\mathbf{e}_3^*\}$ forms a basis for the space of alternating bilinear $(2-)$forms over E. ■

The **space of alternating $p-$forms over** $E\simeq\mathbb{R}^n$ will constantly be appearing in what follows; to lighten the text, we denote it by $\Lambda^p(E)$.

Theorem 167
Let $\{\mathbf{e}_1,\ldots,\mathbf{e}_n\}$ be a basis of $E\simeq\mathbb{R}^n$ and $\{\mathbf{e}_1^*,\ldots,\mathbf{e}_n^*\}$ be the dual basis of E^*. Then

$$\{\mathbf{e}_{i_1}^*\wedge\cdots\wedge\mathbf{e}_{i_p}^* \mid i_1<\cdots<i_p\}$$

is a basis of $\Lambda^p(E)$.

Proof: left as an exercise. ■

Corollary 168
Let $E\simeq\mathbb{R}^n$. If $1\leq p\leq n$, then

$$\dim(\Lambda^p(E)) = \binom{n}{p} = \frac{n!}{p!(n-p)!};$$

if $p>n$, then $\dim(\Lambda^p(E)) = 0$.

Proof: left as an exercise. ■

12.3 Wedge Product of Alternating Forms

If $f \in \Lambda^p(E)$ and $g \in \Lambda^q(E)$, is there a **natural way** to build a form $f \wedge g \in \Lambda^{p+q}(E)$? It turns out that it can be done, with a small group theory detour.

Let S_{p+q} be the permutation group on $\{1, \ldots, p+q\}$,[3] and set

$$A = \{\sigma \in S_{p+q} \mid \sigma(1) < \cdots < \sigma(p) \quad \text{and} \quad \sigma(p+1) < \cdots < \sigma(p+q)\}.$$

Examples

1. If $p = 1$ and $q = 2$, then $A = \{\sigma \in S_3 \mid \sigma(2) < \sigma(3)\}$. But

   $$S_3 = \{\text{id}, \begin{pmatrix}1 & 2\end{pmatrix}, \begin{pmatrix}1 & 3\end{pmatrix}, \begin{pmatrix}2 & 3\end{pmatrix}, \begin{pmatrix}1 & 2 & 3\end{pmatrix}, \begin{pmatrix}1 & 3 & 2\end{pmatrix}\},$$

   so that
   $$A = \{\text{id}, \begin{pmatrix}1 & 2\end{pmatrix}, \begin{pmatrix}1 & 3 & 2\end{pmatrix}\}.$$

2. If $p = 2$ and $q = 2$, then $A = \{\sigma \in S_4 \mid \sigma(1) < \sigma(2) \quad \text{and} \quad \sigma(3) < \sigma(4)\}$; S_4 has $4! = 24$ permutations, and we can show that

   $$A = \{\text{id}, \begin{pmatrix}2 & 3\end{pmatrix}, \begin{pmatrix}2 & 4 & 3\end{pmatrix}, \begin{pmatrix}1 & 2 & 3\end{pmatrix}, \begin{pmatrix}1 & 2 & 4 & 3\end{pmatrix}, \begin{pmatrix}1 & 3\end{pmatrix}\begin{pmatrix}2 & 4\end{pmatrix}\}.$$

   Permutation calculations can quickly become cumbersome! $\square$

If $f \in \Lambda^p(E)$ and $g \in \Lambda^q(E)$, the **wedge product of f and g** is given by

$$f \wedge g(\mathbf{x}_1, \ldots, \mathbf{x}_p, \mathbf{x}_{p+1}, \ldots, \mathbf{x}_{p+q}) = \sum_{\sigma \in A} \epsilon(\sigma) f(\mathbf{x}_{\sigma(1)}, \ldots, \mathbf{x}_{\sigma(p)}) g(\mathbf{x}_{\sigma(p+1)}, \ldots, \mathbf{x}_{\sigma(p+q)}).$$

As $f \wedge g$ depends linearly on each of $\mathbf{x}_1, \ldots, \mathbf{x}_{p+q}$, then it is a linear $(p+q)-$form on E. Is it alternating?

Example: if $p = 1$ and $q = 3$, then

$$A = \{\sigma \in S_4 \mid \sigma(2) < \sigma(3) < \sigma(4)\} = \{\text{id}, \begin{pmatrix}1 & 2\end{pmatrix}, \begin{pmatrix}1 & 3 & 2\end{pmatrix}, \begin{pmatrix}1 & 4 & 3 & 2\end{pmatrix}\};$$

the corresponding signatures are $1, -1, 1, -1$. If all we know of f, g is that $f \in \Lambda^1(E)$ and $g \in \Lambda^q(E)$, then we must have:

[3] A permutation $\sigma \in S_n$ is a bijection $\sigma : \{1, \ldots, n\} \to \{1, \ldots, n\}$. We can also write σ in **cycle notation**, as illustrated as follows: suppose that σ acts on $\{1, 2, 3, 4, 5\}$ according to $\sigma(1) = 2$, $\sigma(2) = 5$, $\sigma(5) = 1$, $\sigma(3) = 3$, and $\sigma(4) = 4$. Then we write σ as $\begin{pmatrix}1 & 2 & 5\end{pmatrix}\begin{pmatrix}3\end{pmatrix}\begin{pmatrix}4\end{pmatrix}$, or usually as $\begin{pmatrix}1 & 2 & 5\end{pmatrix}$ since $3, 4$ are left unchanged by σ. The signature $\epsilon(\sigma)$ of a permutation σ is determined as follows. We write σ as a product of disjoint cycles (as above); the signature is -1 if and only if the factorization contains an **odd** number of **even-length** cycles. As $\sigma = \begin{pmatrix}1 & 2 & 5\end{pmatrix}$ contains no even-length cycle, $\epsilon(\sigma) = 1$.

$$\begin{aligned} f \wedge g(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4) &= f(\mathbf{x}_1) g(\mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4) - f(\mathbf{x}_2) g(\mathbf{x}_1, \mathbf{x}_3, \mathbf{x}_4) \\ &\quad + f(\mathbf{x}_3) g(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_4) - f(\mathbf{x}_4) g(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3). \end{aligned}$$

If $\mathbf{x}_1 = \mathbf{x}_2$, $\mathbf{x}_1 = \mathbf{x}_3$, $\mathbf{x}_1 = \mathbf{x}_4$, $\mathbf{x}_2 = \mathbf{x}_3$, $\mathbf{x}_2 = \mathbf{x}_4$, or $\mathbf{x}_3 = \mathbf{x}_4$, the g components of $f \wedge g$ are either 0 because they are alternating and contain a repeated argument, or they cancel one another out (try it!); thus $f \wedge g$ is alternating. □

The wedge product has the right kinds of properties: if $f, f_1, f_2 \in \Lambda^p(E)$, $g, g_1, g_2 \in \Lambda^q(E)$, and $\alpha \in \mathbb{R}$, then

$$\begin{aligned} (f_1 + f_2) \wedge g &= f_1 \wedge g + f_2 \wedge g, \\ f \wedge (g_1 + g_2) &= f \wedge g_1 + f \wedge g_2, \\ (\alpha f) \wedge g &= \alpha (f \wedge g) = f \wedge (\alpha g). \end{aligned}$$

This leads us to the following crucial result.

Lemma 169
Let $f_i \in E^*$, $1 \leq i \leq p+q$. Then $f = f_1 \wedge \cdots \wedge f_p \in \Lambda^p(E)$, $g = g_{p+1} \wedge \cdots \wedge g_{p+q} \in \Lambda^q(E)$ and

$$f \wedge g = f_1 \wedge \cdots \wedge f_{p+q}.$$

Proof: by definition,

$$\begin{aligned} f_1 \wedge \cdots \wedge f_p(\mathbf{x}_1, \ldots, \mathbf{x}_p) &= \sum_{\sigma \in S_p} \epsilon(\sigma) f_1(\mathbf{x}_{\sigma(1)}) \cdots f_p(\mathbf{x}_{\sigma(p)}), \\ f_{p+1} \wedge \cdots \wedge f_{p+q}(\mathbf{x}_{p+1}, \ldots, \mathbf{x}_{p+q}) &= \sum_{\tau \in S_q} \epsilon(\tau) f_{p+1}(\mathbf{x}_{\tau(p+1)}) \cdots f_{p+q}(\mathbf{x}_{\tau(p+q)}). \end{aligned}$$

It is easy to see that

$$S_p \simeq \{\sigma \in S_{p+q} \mid \sigma(j) = j, p+1 \leq j \leq p+q\} \quad \text{and} \quad S_q \simeq \{\tau \in S_{p+q} \mid \tau(j) = j, 1 \leq j \leq p\}.$$

In Lemma 171, we will see that every $\tilde{\sigma} \in S_{p+q}$ can be written uniquely as $\tilde{\sigma} = \sigma\sigma'\sigma''$, with $\sigma \in A$, $\sigma' \in S_p$, and $\sigma'' \in S_q$. Then

$$\begin{aligned} f_1 \wedge \cdots \wedge f_{p+q}(\mathbf{x}_1, \ldots, \mathbf{x}_{p+q}) &= \sum_{\tilde{\sigma} \in S_{p+q}} \epsilon(\tilde{\sigma}) f_1(\mathbf{x}_{\tilde{\sigma}(1)}) \cdots f_{p+q}(\mathbf{x}_{\tilde{\sigma}(p+q)}) \\ &= \sum_{\sigma, \sigma', \sigma''} \epsilon(\sigma\sigma'\sigma'') f_1(\mathbf{x}_{\sigma\sigma'\sigma''(1)}) \cdots f_{p+q}(\mathbf{x}_{\sigma\sigma'\sigma''(p+q)}) \\ &= \sum_{\sigma, \sigma', \sigma''} \epsilon(\sigma)\epsilon(\sigma')\epsilon(\sigma'') f_1(\mathbf{x}_{\sigma\sigma'(1)}) \cdots f_p(\mathbf{x}_{\sigma\sigma'(p)}) f_{p+1}(\mathbf{x}_{\sigma\sigma''(p+1)}) \cdots f_{p+q}(\mathbf{x}_{\sigma\sigma''(p+q)}) \end{aligned}$$

so that

$$\begin{aligned}
&f_1 \wedge \cdots \wedge f_{p+q}(\mathbf{x}_1, \ldots, \mathbf{x}_{p+q}) \\
&= \sum_{\sigma \in A} \epsilon(\sigma) \Big(\sum_{\sigma' \in S_p} \epsilon(\sigma'') f_1(\mathbf{x}_{\sigma\sigma'(1)}) \cdots f_p(\mathbf{x}_{\sigma\sigma'(p)}) \Big) \Big(\sum_{\sigma'' \in S_q} \epsilon(\sigma'') f_{p+1}(\mathbf{x}_{\sigma\sigma''(p+1)}) \cdots f_{p+q}(\mathbf{x}_{\sigma\sigma''(p+q)}) \Big) \\
&= \sum_{\sigma \in A} f(\mathbf{x}_{\sigma(1)}, \ldots, \mathbf{x}_{\sigma(p)}) g(\mathbf{x}_{\sigma(p+1)}, \ldots, \mathbf{x}_{\sigma(p+q)}) = f \wedge g(\mathbf{x}_{\sigma(p)}, \ldots, \mathbf{x}_{\sigma(p+q)}).
\end{aligned}$$

That this $(p+q)-$form is alternating is left as an exercise. ■

This leads us to the main result of this section.

Theorem 170
Let $f \in \Lambda^p(E)$ and $g \in \Lambda^q(E)$. Then $f \wedge g \in \Lambda^{p+q}(E)$.

Proof: according to Theorem 167, f is a linear combination of wedge products of $p-$forms over E of the form $\mathbf{e}^*_{i_1} \wedge \cdots \wedge \mathbf{e}^*_{i_p}$; similarly, g is a linear combination of wedge products of $q-$forms over E of the form $\mathbf{e}^*_{j_1} \wedge \cdots \wedge \mathbf{e}^*_{j_q}$.

According to Lemma 169, expressions of the form

$$(\mathbf{e}^*_{i_1} \wedge \cdots \wedge \mathbf{e}^*_{i_p}) \wedge (\mathbf{e}^*_{j_1} \wedge \cdots \wedge \mathbf{e}^*_{j_q}) \tag{12.1}$$

are alternating $(p+q)-$forms.

Thus $f \wedge g$ is a linear combination of alternating $(p+q)-$forms as in (12.1); since $\Lambda^{p+q}(E)$ is a vector space over E (see Corollary 168), $f \wedge g$ is alternating. ■

The wedge product of alternating forms is thus well-defined, and it has a set of useful properties. Let $f \in \Lambda^p(E), g \in \Lambda^q(E), h \in \Lambda^r(E)$. Then:

1. $f \wedge (g \wedge h) = (f \wedge g) \wedge h$ (the wedge product is **associative**);

2. $f \wedge g = (-1)^{pq} g \wedge f$ (it is **not commutative**), and

3. if $u : E \to F$ is a linear transformation, $f \in \Lambda^p(E)$, and $g \in \Lambda^q(F)$, then $u(f) \in \Lambda^p(E)$, where
$$u(f)(\mathbf{x}_1, \ldots, \mathbf{x}_p) = f(u(\mathbf{x}_1), \ldots, u(\mathbf{x}_p));$$
$u(g) \in \Lambda^q(E)$, where
$$u(g)(\mathbf{x}_1, \ldots, \mathbf{x}_p) = g(u(\mathbf{x}_1), \ldots, u(\mathbf{x}_q)),$$
and $u(f \wedge g) = u(f) \wedge u(g) \in \Lambda^{p+q}(E)$ (the proof is left as an exercise).

We finish this section with the promised lemma.

Lemma 171
If $\tilde{\sigma} \in S_{p+q}$, there is a unique triplet $\sigma \in A$, $\sigma' \in S_p$, and $\sigma'' \in S_q$ such that $\tilde{\sigma} = \sigma\sigma'\sigma''$.

Proof: let $A' = \{\tilde{\sigma}(1), \ldots, \tilde{\sigma}(p)\} \subseteq \{1, \ldots, p+q\}$. List the integers in A' in increasing order, and define σ' by $\sigma'(j) =$rank of $\tilde{\sigma}(j)$ in A', for $1 \leq j \leq p$.

Similarly, define σ'' by $\sigma''(j) =$rank of $\tilde{\sigma}(i)$ in $A'' =$ ordered$\{\tilde{\sigma}(p+1), \ldots, \tilde{\sigma}(p+q)\}$, for $p+1 \leq i \leq p+q$.

If we write $A' = \{i_1 < \cdots < i_p\}$ and $A'' = \{i_{p+1} < \cdots < i_{p+q}\}$, we can then define σ by $\sigma(j) = i_j$, $1 \leq j \leq p+q$. Then $\tilde{\sigma} = \sigma\sigma'\sigma''$. ■

12.4 Solved Problems

1. Let E be a finite-dimensional vector space over $\mathbb{R}$, with $\dim(E) = 3$. If $\mathbf{x}, \mathbf{y}, \mathbf{z} \in E$ are linearly dependent, show that $f(\mathbf{x}, \mathbf{y}, \mathbf{z}) = 0$ for any alternating linear $3-$form f.

   Proof: let $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ be the canonical basis of E. Since $\mathbf{x}, \mathbf{y}, \mathbf{z} \in E$ are linearly dependent, (at least) one of them may be expressed as a linear combination of the other two. Without loss of generality, say $\mathbf{x} = a\mathbf{y} + b\mathbf{z}$, with $a, b \in \mathbb{R}$. Then

   $$f(\mathbf{x}, \mathbf{y}, \mathbf{z}) = f(a\mathbf{y} + b\mathbf{z}, \mathbf{y}, \mathbf{z}) = af(\mathbf{y}, \mathbf{y}, \mathbf{z}) + bf(\mathbf{z}, \mathbf{y}, \mathbf{z}) = a \cdot 0 + b \cdot 0 = 0,$$

   since f is alternating. ■

2. Let E be a finite-dimensional vector space over $\mathbb{R}$, with $\dim(E) = 3$. If $\mathbf{x}, \mathbf{y}, \mathbf{z} \in E$ are linearly independent, show that $f(\mathbf{x}, \mathbf{y}, \mathbf{z}) \neq 0$ for any alternating linear $3-$form $f \neq 0$.

   Proof: let $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ be the canonical basis of E. Since $f \neq 0$, $f(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) \neq 0$. Write

   $$\begin{aligned} \mathbf{x} &= x_1\mathbf{e}_1 + x_2\mathbf{e}_2 + x_3\mathbf{e}_3 \\ \mathbf{y} &= y_1\mathbf{e}_1 + y_2\mathbf{e}_2 + y_3\mathbf{e}_3 \\ \mathbf{z} &= z_1\mathbf{e}_1 + z_2\mathbf{e}_2 + z_3\mathbf{e}_3 \end{aligned}$$

   Since $\{\mathbf{x}, \mathbf{y}, \mathbf{z}\}$ are linearly independent,

   $$\det \begin{pmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{pmatrix} \neq 0.$$

   Then

   $$f(\mathbf{x}, \mathbf{y}, \mathbf{z}) = \sum_{\substack{i \neq j \\ i \neq k \\ j \neq k}} x_i y_j z_k f(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) = \det \begin{pmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{pmatrix} \cdot f(\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3) \neq 0. \quad ■$$

3. Show that the inner product $(\cdot \mid |\cdot) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is a bilinear form.

   Proof: the inner product $(\cdot \mid \cdot) : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ is defined by

   $$(\mathbf{x} \mid \mathbf{y}) = \sum_{i=1}^{n} x_i y_i.$$

   In order to show it is bilinear, we need to show that for all $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n$, $a, b \in \mathbb{R}$, we have

   - $(a\mathbf{x} + b\mathbf{y} \mid \mathbf{z}) = a(\mathbf{x}|\mathbf{z}) + b(\mathbf{y} \mid \mathbf{z})$
   - $(\mathbf{x} \mid a\mathbf{y} + b\mathbf{z}) = a(\mathbf{x} \mid \mathbf{y}) + b(\mathbf{x} \mid \mathbf{z})$

   But

   $$(a\mathbf{x} + b\mathbf{y} \mid \mathbf{z}) = \sum_{i=1}^{n}(ax_i + by_i)z_i = a\sum_{i=1}^{n} x_i z_i + b\sum_{i=1}^{n} y_i z_i = a(\mathbf{x} \mid \mathbf{z}) + b(\mathbf{y} \mid \mathbf{z})$$

   and

   $$(\mathbf{x} \mid a\mathbf{y} + b\mathbf{z}) = \sum_{i=1}^{n} x_i(ay_i + bz_i) = a\sum_{i=1}^{n} x_i y_i + b\sum_{i=1}^{n} x_i z_i = a(\mathbf{x} \mid \mathbf{y}) + b(\mathbf{x} \mid \mathbf{z})$$

   so that the inner product is indeed bilinear. It is not alternating, however, since we would need $(\mathbf{x} \mid \mathbf{x}) = 0$ for all $\mathbf{x} \in \mathbb{R}^n$ but $(\mathbf{e}_1 \mid \mathbf{e}_1) = 1$. ■

4. Show that det : $\mathbb{R}^n \times \cdots \times \mathbb{R}^n \to \mathbb{R}$ is a bilinear form.

   Proof: that this form is both multilinear and alternating is immediate due to the properties of the determinant that you have seen/will see in your linear algebra courses:

   - Firstly, det $(\mathbf{x}_1, \ldots, a_1\mathbf{y}_1 + a_2\mathbf{y}_2, \ldots, \mathbf{x}_n) = \sum_{j=1}^{2} a_j$ det $(\mathbf{x}_1, \ldots, \mathbf{y}_j, \ldots, \mathbf{x}_n)$
   - Secondly, det $(\mathbf{x}_1, \ldots, \mathbf{x}_n) = 0$ if $x_i = x_j$ for some $i \neq j$.

   ■

5. Show that $\{\mathbf{e}^*_{i_1} \otimes \mathbf{e}^*_{i_2} \otimes \cdots \mathbf{e}^*_{i_p} : 1 \leq i_j \leq n\}$ forms a basis of the vector space of linear $p-$forms over E. What is the dimension of that vector space?

   Proof: recall that $\mathbf{e}^*_i : E \to \mathbb{R}$ is the linear functional such that $\mathbf{e}^*_i(\mathbf{e}_j) = \delta_{i,j}$.

   Let us first assume that the set in question is indeed a basis of the space of all linear (but not necessarily alternating) $p-$forms. There are n possible choices for each $1-$form $\mathbf{e}^*_{i_j}$ appearing in the tensor product. Since there are p such forms, there is a total of n^p tensor products. Hence, dim$(\{$space of $p-$linear forms over $E\}) = n^p$.

We now show that the set is such a basis. First, note that for any choice of indices i_j, $1 \leq j \leq p$, $\mathbf{e}^*_{i_1} \otimes \cdots \mathbf{e}^*_{i_p}$ is a $p-$linear form over E; indeed,

$$\begin{aligned}
&\mathbf{e}^*_{i_1} \otimes \cdots \otimes \mathbf{e}^*_{i_j} \otimes \cdots \otimes \mathbf{e}^*_{i_p}(\mathbf{x}_1, \ldots, a\mathbf{y}_1 + b\mathbf{y}_2, \ldots, \mathbf{x}_p) \\
&\quad = \mathbf{e}^*_{i_1}(\mathbf{x}_1) \cdots \mathbf{e}^*_{i_j}(a\mathbf{y}_1 + b\mathbf{y}_2) \cdots \mathbf{e}^*_{i_p}(\mathbf{x}_p) \\
&\quad = a\mathbf{e}^*_{i_1}(\mathbf{x}_1) \cdots \mathbf{e}^*_{i_j}(\mathbf{y}_1) \cdots \mathbf{e}^*_{i_p}(\mathbf{x}_p) + b\mathbf{e}^*_{i_1}(\mathbf{x}_1) \cdots \mathbf{e}^*_{i_j}(\mathbf{y}_2) \cdots \mathbf{e}^*_{i_p}(\mathbf{x}_p) \\
&\quad = a\mathbf{e}^*_{i_1} \otimes \cdots \otimes \mathbf{e}^*_{i_j} \otimes \cdots \otimes \mathbf{e}^*_{i_p}(\mathbf{x}_1, \ldots, \mathbf{y}_1, \ldots, \mathbf{x}_p) + b\mathbf{e}^*_{i_1} \otimes \cdots \otimes \mathbf{e}^*_{i_j} \otimes \cdots \otimes \mathbf{e}^*_{i_p}(\mathbf{x}_1, \ldots, \mathbf{y}_2, \ldots, \mathbf{x}_p)
\end{aligned}$$

since $\mathbf{e}^*_{i_j}$ is linear. Hence,

$$\text{Span}\{\mathbf{e}^*_{i_1} \otimes \mathbf{e}^*_{i_2} \otimes \cdots \mathbf{e}^*_{i_p} : 1 \leq i_j \leq n\} \subseteq \{\text{space of } p-\text{linear forms over } E\}$$

Now, let f be a $p-$linear form, and suppose $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ be the canonical basis of E. For $1 \leq j \leq p$, write

$$\mathbf{x}_j = \sum_{i=1}^{n} x_{j,i}\mathbf{e}_i.$$

Then

$$\begin{aligned}
f(\mathbf{x}_1, \ldots, \mathbf{x}_p) &= \sum_{j_1,\ldots,j_p=1}^{n} x_{j_1,1} \cdots x_{j_p,1} f(\mathbf{e}_{j_1}, \ldots, \mathbf{e}_{j_p}) \\
&= \sum_{j_1,\ldots,j_p=1}^{n} \mathbf{e}^*_{j_1}(\mathbf{x}_1) \cdots \mathbf{e}^*_{j_p}(\mathbf{x}_p) f(\mathbf{e}_{j_1}, \ldots, \mathbf{e}_{j_p}) \\
&= \sum_{j_1,\ldots,j_p=1}^{n} f(\mathbf{e}_{j_1}, \ldots, \mathbf{e}_{j_p}) \mathbf{e}^*_{j_1} \otimes \cdots \otimes \mathbf{e}^*_{j_p}(\mathbf{x}_1, \ldots, \mathbf{x}_p)
\end{aligned}$$

and so $f \in \text{Span}\{\mathbf{e}^*_{i_1} \otimes \mathbf{e}^*_{i_2} \otimes \cdots \mathbf{e}^*_{i_p} : 1 \leq i_j \leq n\}$. Consequently,

$$\text{Span}\{\mathbf{e}^*_{i_1} \otimes \mathbf{e}^*_{i_2} \otimes \cdots \mathbf{e}^*_{i_p} : 1 \leq i_j \leq n\} = \{\text{space of } p-\text{linear forms over } E\}$$

It remains only to show that the tensor products are linearly independent. To do so, suppose that

$$\sum_{j_1,\ldots,j_p=1}^{n} a_{j_1,\ldots,j_p} \mathbf{e}^*_{j_1} \otimes \cdots \otimes \mathbf{e}^*_{j_p} = 0$$

Then

$$\sum_{j_1,\ldots,j_p=1}^{n} a_{j_1,\ldots,j_p} \mathbf{e}^*_{j_1} \otimes \cdots \otimes \mathbf{e}^*_{j_p}(\mathbf{x}_1, \ldots, \mathbf{x}_p) = 0$$

for all $(\mathbf{x}_1, \ldots, \mathbf{x}_p) \in E^p$. Fix $j_1^*, \ldots, j_p^*$. Then $(\mathbf{e}_{j_1^*}, \ldots, \mathbf{e}_{j_p^*}) \in E^p$ and so

$$\sum_{j_1,\ldots,j_p=1}^{n} a_{j_1,\ldots,j_p} \mathbf{e}^*_{j_1} \otimes \cdots \otimes \mathbf{e}^*_{j_p}(\mathbf{e}_{j_1^*}, \ldots, \mathbf{e}_{j_p^*}) = 0$$

But

$$\sum_{j_1,\ldots,j_p=1}^{n} a_{j_1,\ldots,j_p}\mathbf{e}^*_{j_1}\otimes\cdots\otimes\mathbf{e}^*_{j_p}(\mathbf{e}_{j_1^*},\ldots,\mathbf{e}_{j_p^*}) = a_{j_1^*,\ldots,j_p^*}$$

so that $a_{j_1^*,\ldots,j_p^*}=0$. But $j_1^*,\ldots,j_p^*$ were arbitrary, so that we indeed have $a_{j_1,\ldots,j_p}=0$ for all $1\leq j_1,\ldots,j_p\leq n$, and the tensor products are linearly independent. ■

6. Let $f_1,f_2,\ldots,f_p$ be linear $1-$forms over E and $\sigma\in S_p$. Show that

$$f_{\sigma(1)}\wedge\cdots\wedge f_{\sigma(p)} = \varepsilon(\sigma)f_1\wedge\cdots\wedge f_p.$$

Proof: by definition, we have

$$\begin{aligned}
f_{\sigma(1)}\wedge\cdots\wedge f_{\sigma(p)}(\mathbf{x}_1,\ldots,\mathbf{x}_p) &= \det\begin{pmatrix} f_{\sigma(1)}(\mathbf{x}_1) & \cdots & f_{\sigma(1)}(\mathbf{x}_p)\\ \vdots & & \vdots\\ f_{\sigma(p)}(\mathbf{x}_1) & \cdots & f_{\sigma(p)}(\mathbf{x}_p)\end{pmatrix}\\
&= \epsilon(\sigma)\det\begin{pmatrix} f_1(\mathbf{x}_1) & \cdots & f_1(\mathbf{x}_p)\\ \vdots & & \vdots\\ f_p(\mathbf{x}_1) & \cdots & f_p(\mathbf{x}_p)\end{pmatrix}\\
&= \epsilon(\sigma)f_1\wedge\cdots\wedge f_p(\mathbf{x}_1,\ldots,\mathbf{x}_p)
\end{aligned}$$

■

7. Let $f_1,f_2,\ldots,f_p$ be linear $1-$forms over E such that $f_i=f_j$ for some $i\neq j$. Show that $f_1\wedge\cdots\wedge f_p=0$.

Proof: by definition, we have

$$f_1\wedge\cdots\wedge f_p(\mathbf{x}_1,\ldots,\mathbf{x}_p) = \det\begin{pmatrix} f_1(\mathbf{x}_1) & \cdots & f_1(\mathbf{x}_p)\\ \vdots & & \vdots\\ f_p(\mathbf{x}_1) & \cdots & f_p(\mathbf{x}_p)\end{pmatrix}$$

If $f_i=f_j$ for $i\neq j$, two of the rows in the above matrix are identical; as a result, the determinant is 0. ■

8. Provide a proof of Corollary 168.

Proof: you should be able to make an informal argument for this one. In essence, the proof runs as follows:

a) $\Lambda^p(E)$ is a subspace of the space of linear $p-$forms over E.

b) $\mathbf{e}^*_{i_1}\wedge\cdots\wedge\mathbf{e}^*_{i_p}\in\Lambda^p(E)$ for all $1\leq i_1,\ldots,i_p\leq n$, so that

$$\text{Span}\{\mathbf{e}^*_{i_1}\wedge\cdots\wedge\mathbf{e}^*_{i_p} : 1\leq i_1,\ldots,i_p\leq n\}\subseteq\Lambda^p(E).$$

c) Any $f\in\Lambda^p(E)$ can be written as

$$f = \sum_{i_1<\cdots<i_p} f(\mathbf{e}_{i_1},\ldots,\mathbf{e}_{i_p})\mathbf{e}^*_{i_1}\wedge\cdots\wedge\mathbf{e}^*_{i_p}$$

so that
$$\Lambda^p(E) \subseteq \text{Span}\{\mathbf{e}_{i_1}^* \wedge \cdots \wedge \mathbf{e}_{i_p}^* : 1 \leq i_1, \ldots, i_p \leq n\}.$$

Consequently,

$$\Lambda^p(E) = \text{Span}\{\mathbf{e}_{i_1}^* \wedge \cdots \wedge \mathbf{e}_{i_p}^* : 1 \leq i_1, \ldots, i_p \leq n\}.$$

d) For each fixed choice of $i_1, \ldots, i_p$ there are two possibilities:
   i. if the indices are all distinct, let

   $$A_{i_1,\ldots,i_p} = \big\{\sigma \in S_p : \sigma(\{i_1, \ldots, i_p\}) \subseteq \{i_1, \ldots, i_p\}\big\}.$$

   Then $\mathbf{e}_{i_1}^* \wedge \cdots \wedge \mathbf{e}_{i_p}^* = \epsilon(\sigma)\mathbf{e}_{\sigma(i_1)}^* \wedge \cdots \wedge \mathbf{e}_{\sigma(i_p)}^*$ for each $\sigma \in A_{i_1,\ldots,i_p}$. Consequently, all wedge products containing $\mathbf{e}_{i_1}^*, \ldots, \mathbf{e}_{i_p}^*$ are linearly dependent. Remove all of them except the canonical one, i.e. the one for which $i_1 < \ldots < i_p$ (this can be done since all indices are distinct);
   ii. if some of the indices repeat, then $\mathbf{e}_{i_1}^* \wedge \cdots \wedge \mathbf{e}_{i_p}^* = 0$ (see exercise 8). Consequently, all such wedge products are linearly dependent. Remove all of them.

e) The remaining wedge products $\{\mathbf{e}_{i_1}^* \wedge \cdots \wedge \mathbf{e}_{i_p}^* : i_1 < \cdots < i_p\}$ span $\Lambda^p(E)$. One can show that they are linearly independent just as was done at the end of exercise 6. Thus
$$\{\mathbf{e}_{i_1}^* \wedge \cdots \wedge \mathbf{e}_{i_p}^* : i_1 < \cdots < i_p\}$$
is a basis of $\Lambda^p(E)$.

f) If $n \leq p$, there are $\binom{n}{p}$ ways of selecting p distinct indices from a set of n indices, and so $\dim(\Lambda^p(E)) = \binom{n}{p}$.

g) In the event where $p > n$, there is no way of selecting p distinct indices from a set of n indices, and so $\Lambda^p(E) = \{0\}$. ■

9. Let $f = f_1 \wedge f_2$ and $g = g_1 \wedge g_2$ be alternating $p-$forms over E. Work out the details and express $f \wedge g$ in terms of f and g, and show that $f \wedge g$ is alternating.

Proof: we have

$$\begin{aligned} A &= \{\sigma \in S_4 | \sigma(1) < \sigma(2) \ \underline{\text{and}} \ \sigma(3) < \sigma(4)\} \\ &= \{\text{id}, (2\,3), (2\,4\,3), (1\,2\,3), (1\,2\,4\,3), (1\,3)(2\,4)\} \end{aligned}$$

Consequently,

$$\begin{aligned} f \wedge g(\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4) &= \sum_{\sigma \in A} \epsilon(\sigma) f(\mathbf{x}_{\sigma(1)}, \mathbf{x}_{\sigma(2)}) g(\mathbf{x}_{\sigma(3)}, \mathbf{x}_{\sigma(4)}) \\ &= f(\mathbf{x}_1, \mathbf{x}_2) g(\mathbf{x}_3, \mathbf{x}_4) - f(\mathbf{x}_1, \mathbf{x}_3) g(\mathbf{x}_2, \mathbf{x}_4) + f(\mathbf{x}_1, \mathbf{x}_2) g(\mathbf{x}_3, \mathbf{x}_4) \\ &\quad + f(\mathbf{x}_2, \mathbf{x}_3) g(\mathbf{x}_1, \mathbf{x}_4) - f(\mathbf{x}_2, \mathbf{x}_4) g(\mathbf{x}_1, \mathbf{x}_3) + f(\mathbf{x}_3, \mathbf{x}_4) g(\mathbf{x}_1, \mathbf{x}_2) \end{aligned}$$

We can easily verify that $f \wedge g$ is alternating, using the fact that both f and g are alternating. ■

12.5 Exercises

1. Prepare a 2-page summary of this chapter, with important definitions and results.
2. Prove Proposition 166 for $p \geq 3$.
3. Let $f_1, \dots, f_p \in E^*$. Show that if $f_i = f_j$ for some $i \neq j$, then $f_1 \wedge \cdots \wedge f_p = 0$.
4. Prove Theorem 167.
5. Show that if $p = q = 2$, the set $A \subseteq S_4$ contains only 6 permutations.
6. Let f, f_1, f_2 be alternating $p-$forms over E, g, g_1, g_2 be alternating $q-$forms over E, and $\alpha \in \mathbb{R}$. Show that
   a) $(f_1 + f_2) \wedge g = f_1 \wedge g + f_2 \wedge g$
   b) $f \wedge (g_1 + g_2) = f \wedge g_1 + f \wedge g_2$
   c) $(\alpha f) \wedge g = \alpha(f \wedge g) = f \wedge (\alpha g)$
7. Complete the proof of Lemma 169.
8. Let $f \in \Lambda^p(E)$, $g \in \Lambda^q(E)$, and $h \in \Lambda^r(E)$. Show that $f \wedge (g \wedge h) = (f \wedge g) \wedge h \in \Lambda^{p+q+r}(E)$ and that $f \wedge g = (-1)^{pq} g \wedge f$.
9. Prove property 3 on p. 304.
10. Let k be odd and $\omega \in \Lambda^k(\mathbb{R}^n)$. Show that $\omega \wedge \omega = 0$. Is the condition on k necessary, sufficient, or both?

Chapter 13

Differential Forms

In this chapter, we introduce the notion of **differential** $p-$**forms over** $\mathbb{R}^n$, which are derivatives of alternating linear $p-$forms over $\mathbb{R}^n$. This new notion is a generalization of the **differential** of a function and admits a number of applications in mathematical physics (Grand Unified Theories, Yang-Mills theory, superstring theory, etc.)

13.1 Differential $p-$Forms

We start by discussing the situation for $n=3$. Let $U \subseteq_O \mathbb{R}^3$. A **differential** $1-$**form over** U is a function $U \to (\mathbb{R}^3)^*$; the set of all such differential forms is denoted $\Omega^1(U)$.

If $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ is the **canonical basis** of $\mathbb{R}^3$, then for any $\mathbf{w} \in \mathbb{R}^3$ we have

$$\mathbf{w} = w_1\mathbf{e}_1 + w_2\mathbf{e}_2 + w_3\mathbf{e}_3.$$

We denote the **dual basis** of $(\mathbb{R}^3)^*$ by $\{\,\mathrm{d}x,\ \mathrm{d}y,\ \mathrm{d}z\}$, which is to say that

$$\mathrm{d}x,\ \mathrm{d}y,\ \mathrm{d}z : \mathbb{R}^3 \to \mathbb{R} \quad \text{and} \quad \mathrm{d}x(\mathbf{w}) = w_1,\ \ \mathrm{d}y(\mathbf{w}) = w_2,\ \ \mathrm{d}z(\mathbf{w}) = w_3 \quad \text{for all } \mathbf{w} \in \mathbb{R}^3.$$

Then, if $\alpha \in (\mathbb{R}^3)^*$, there are unique $P, Q, R \in R$ such that

$$\alpha = P\,\mathrm{d}x + Q\,\mathrm{d}y + R\,\mathrm{d}z.$$

In general, if $\omega \in \Omega^1(U)$, $\exists! P, Q, R : U \to \mathbb{R}$ such that

$$\omega(\mathbf{u}) = P(\mathbf{u})\,\mathrm{d}x + Q(\mathbf{u})\,\mathrm{d}y + R(\mathbf{u})\,\mathrm{d}z, \quad \text{for all } \mathbf{u} \in U.$$

Let $f : U \to \mathbb{R}$ be differentiable on U; the **differential of** f is $\mathrm{d}f \in \Omega^1(U)$, where

$$\mathrm{d}f(\mathbf{u}) = \frac{\partial f}{\partial x}(\mathbf{u})\,\mathrm{d}x + \frac{\partial f}{\partial y}(\mathbf{u})\,\mathrm{d}y + \frac{\partial f}{\partial z}(\mathbf{u})\,\mathrm{d}z, \quad \text{for all } \mathbf{u} \in U.$$

Let $\omega \in \Omega^1(U)$. If the constituents $P, Q, R : U \to \mathbb{R}$ are continuous on U (respectively $\mathcal{C}^1$ or $\mathcal{C}^\infty$), then ω is continuous U (respectively $\mathcal{C}^1$ or $\mathcal{C}^\infty$).[1]

[1] These restrictions on P, Q, R make $\Omega^1(U)$ a $\mathcal{C}^0(U,\mathbb{R})-$ **module** (respectively, $\mathcal{C}^1(U,\mathbb{R})$ or $\mathcal{C}^\infty(U,\mathbb{R})$).

Proposition 172
$\Omega^1(U)$ is an infinite-dimensional vector space over $\mathbb{R}$.

Proof: left as an exercise. ■

If $U \subseteq_O \mathbb{R}^n$, $f : U \to \mathbb{R}$ is $\mathcal{C}^0$ (respectively $\mathcal{C}^1$ or $\mathcal{C}^\infty$) and ω is a $\mathcal{C}^0$ (respectively $\mathcal{C}^1$ or $\mathcal{C}^\infty$) differential $1-$form over U, then $f\omega \in \Omega^1(U)$, where

$$f\omega(\mathbf{u}) = f(\mathbf{u})P(\mathbf{u})\,\mathrm{d}x + f(\mathbf{u})Q(\mathbf{u})\,\mathrm{d}y + f(\mathbf{u})R(\mathbf{u})\,\mathrm{d}z, \quad \forall \mathbf{u} \in U.$$

A **differential** $p-$**form** ω **over** U is a map $\omega : U \to \Lambda^p(\mathbb{R}^n)$; the set of all such differential forms is denoted by $\Omega^p(U)$. If $p = 0$, $\Omega^0(U) = \mathbf{C}^k(U, \mathbb{R})$, where $k \in \{0, 1, \infty\}$; Corollary 168 shows that $\Omega^p(U) = \{0\}$ when $p > n$.

Proposition 173
$\Omega^p(U)$ is an infinite-dimensional vector space over $\mathbb{R}$ and a $\mathbf{C}^k(U)-$module (i.e., if $f \in \mathbf{C}^k(U, \mathbb{R})$ and $\omega \in \Omega^p(U)$, then $f\omega \in \Omega^p(U)$ for $k \in \{0, 1, \infty\}$.

Proof: left as an exercise. ■

Let $\omega_1 \in \Omega^{p_1}(U)$ and $\omega_2 \in \Omega^{p_2}(U)$. By definition, $\omega_i(\mathbf{u}) \in \Lambda^{p_i}(U)$ for all $\mathbf{u} \in U$, for $i = 1, 2$; according to Theorem 170, we must have

$$\omega_1(\mathbf{u}) \wedge \omega_2(\mathbf{u}) \in \Lambda^{p_1+p_2}(U),$$

and so the function $\omega_1 \wedge \omega_2 : U \to \Lambda^{p_1+p_2}(U)$ defined by

$$(\omega_1 \wedge \omega_2)(\mathbf{u}) = \omega_1(\mathbf{u}) \wedge \omega_2(\mathbf{u}), \quad \text{for all } \mathbf{u} \in U$$

is a differential $(p_1 + p_2)-$form over U, which is to say that $\omega_1 \wedge \omega_2 \in \Omega^{p_1+p_2}(U)$. This differential form is called the **we dge (or exterior) product of** ω_1 **and** ω_2.[2]

Example: if $n = 3$, we have

- $\Omega^0(U) = \{\omega = f \mid f \in \mathbf{C}^k(U, \mathbb{R})\}$;
- $\Omega^1(U) = \{\omega = f\,\mathrm{d}x + g\,\mathrm{d}y + h\,\mathrm{d}z \mid f, g, h \in \mathbf{C}^k(U, \mathbb{R})\}$;
- $\Omega^2(U) = \{\omega = f\,\mathrm{d}x \wedge \mathrm{d}y + g\,\mathrm{d}x \wedge \mathrm{d}z + h\,\mathrm{d}y \wedge \mathrm{d}z \mid f, g, h \in \mathbf{C}^k(U, \mathbb{R})\}$;
- $\Omega^3(U) = \{\omega = f\,\mathrm{d}x \wedge \mathrm{d}y \wedge \mathrm{d}z \mid f \in \mathbf{C}^k(U, \mathbb{R})\}$, and
- $\Omega^p(U) = \{0\}$, when $p > 3$. □

[2] It is also sometimes denoted by $\omega_1\omega_2$.

Theorem 174

1. *For* $i = 1, 2$*, let* $\omega_i, \omega_i' \in \Omega^{p_i}(U)$ *and* $f : U \to \mathbb{R}$*. Then:*
   - $(\omega_1 + \omega_1') \wedge \omega_2 = \omega_1 \wedge \omega_2 + \omega_1' \wedge \omega_2$*;*
   - $\omega_1 \wedge (\omega_2 + \omega_2') = \omega_1 \wedge \omega_2 + \omega_1 \wedge \omega_2'$*, and*
   - $(f\omega_1) \wedge \omega_2 = f(\omega_1 \wedge \omega_2) = \omega_1 \wedge (f\omega_2)$*.*
2. *If* $\omega_1, \ldots, \omega_q \in \Omega^1(U)$*, then*
   - *when* $\omega_i = \omega_j$ *for some* $i \neq j$*, we have* $\omega_1 \wedge \cdots \wedge \omega_q = 0$;
   - *for* $\sigma \in S_q$*,* $\omega_{\sigma(1)} \wedge \cdots \wedge \omega_{\sigma(q)} = \epsilon(\sigma)\omega_1 \wedge \cdots \wedge \omega_q$.
3. *For* $i = 1, 2, 3$*, let* $\omega_i \in \Omega^{p_i}(U)$*. Then:*
   - $\omega_1 \wedge (\omega_2 \wedge \omega_3) = (\omega_1 \wedge \omega_2) \wedge \omega_3$*, and*
   - $\omega_1 \wedge \omega_2 = (-1)^{p_1 p_2} \omega_2 \wedge \omega_1$*.*

Proof: left as an exercise. ■

A few examples will help illustrate the main principles.

Examples: let $n = 3$, $f : U \to \mathbb{R}$, and set

$$\omega_1 = \mathrm{d}x_1 = \mathbf{e}_1^*, \quad \omega_2 = \mathrm{d}x_2 = \mathbf{e}_3^*, \quad \omega_3 = \mathrm{d}x_3 = \mathbf{e}_3^* \in \Omega^1(U).$$

- $\mathrm{d}x_1 \wedge \mathrm{d}x_2 = (-1)^{1\cdot 1}\, \mathrm{d}x_2 \wedge \mathrm{d}x_1$;
- $\mathrm{d}x_1 \wedge \mathrm{d}x_2 \wedge \mathrm{d}x_3 = \mathrm{d}x_3 \wedge \mathrm{d}x_1 \wedge \mathrm{d}x_2 = -\,\mathrm{d}x_1 \wedge \mathrm{d}x_3 \wedge \mathrm{d}x_2$;
- $\mathrm{d}x_1 \wedge \mathrm{d}x_1 = \mathrm{d}x_2 \wedge \mathrm{d}x_2 = \mathrm{d}x_3 \wedge \mathrm{d}x_3 = 0$, and
- $(f\, \mathrm{d}x_1 \wedge \mathrm{d}x_2) \wedge \mathrm{d}x_3 = (-1)^{2\cdot 1}\, \mathrm{d}x_3 \wedge (f\, \mathrm{d}x_1 \wedge \mathrm{d}x_2)$. □

This section's final result will set the stage for the rest of the chapter and the next one.

Theorem 175
Let $\omega \in \Omega^p(U)$*. We can uniquely write*

$$\omega = \sum P_{i_1,\cdots,i_p}\, dx_{i_1} \wedge \cdots \wedge dx_{i_p},$$

where $P_{i_1,\cdots,i_p} : U \to \mathbb{R}$ *for* $i_1 < \cdot < i_p$*.*

Proof: left as an exercise. ■

13.2 Exterior Derivative

In what follows, we fix $k = \infty$ so that $\Omega^p(U)$ represents the vector space of $\mathcal{C}^\infty$ (smooth) differential $p-$forms over $U \subseteq_O \mathbb{R}^n$.

The exterior derivative (or differential) of $\omega \in \Omega^p(U)$ is defined recursively.

1. If $f \in \Omega^0(U)$ (that is, $f : U \to \mathbb{R}$ is smooth), then its exterior derivative is
$$\mathrm{d}f = \sum_{i=1}^n \frac{\partial f}{\partial x_i}\,\mathrm{d}x_i \in \Omega^1(U).$$

2. If $\omega = \sum_{i=1}^n P_i\,\mathrm{d}x_i \in \Omega^1(U)$, $P_i \in \mathcal{C}^\infty(U,\mathbb{R})$ for $1 \leq i \leq n$, then its exterior derivative is
$$\mathrm{d}\omega = \sum_{i=1}^n \mathrm{d}P_i \wedge \mathrm{d}x_i = \sum_{i=1}^n \left(\sum_{j=1}^n \frac{\partial P_i}{\partial x_j}\,\mathrm{d}x_j\right) \wedge \mathrm{d}x_i = \sum_{i<j}\left(\frac{\partial P_j}{\partial x_i} - \frac{\partial P_i}{\partial x_j}\right)\mathrm{d}x_i \wedge \mathrm{d}x_j \in \Omega^2(U).$$

...

p. In general, if
$$\omega = \sum_{i_1<\cdots<i_p} P_{i_1,\cdots,i_p}\,\mathrm{d}x_{i_1} \wedge \cdots \wedge \mathrm{d}x_{i_p} \in \Omega^p(U),$$
then its exterior derivative is
$$\mathrm{d}\omega = \sum_{i_1<\cdots<i_p} \mathrm{d}P_{i_1,\cdots,i_p} \wedge \mathrm{d}x_{i_1} \wedge \cdots \wedge \mathrm{d}x_{i_p} \in \Omega^{p+1}(U).$$

As we shall see after the next examples, the exterior derivative behaves as a regular derivative with respect to the sum of differential forms and to the product of functions, but there is a twist for a general product of differential forms.

> **Examples:** throughout, let $f, g, h \in \mathcal{C}^\infty(\mathbb{R}^n, \mathbb{R})$ for an appropriate n.
>
> 1. In $\mathbb{R}^2$, let $\omega = f\,\mathrm{d}x + g\,\mathrm{d}y \in \Omega^1(\mathbb{R}^2)$. Then
> $$\begin{aligned}\mathrm{d}\omega &= \mathrm{d}f \wedge \mathrm{d}x + \mathrm{d}g \wedge \mathrm{d}y = \left(\frac{\partial f}{\partial x}\,\mathrm{d}x + \frac{\partial f}{\partial y}\,\mathrm{d}y\right) \wedge \mathrm{d}x + \left(\frac{\partial g}{\partial x}\,\mathrm{d}x + \frac{\partial g}{\partial y}\,\mathrm{d}y\right) \wedge \mathrm{d}y \\ &= \frac{\partial f}{\partial x}\,\mathrm{d}x \wedge \mathrm{d}x + \frac{\partial f}{\partial y}\,\mathrm{d}y \wedge \mathrm{d}x + \frac{\partial g}{\partial x}\,\mathrm{d}x \wedge \mathrm{d}y + \frac{\partial g}{\partial y}\,\mathrm{d}y \wedge \mathrm{d}y \\ &= \frac{\partial f}{\partial x} \cdot 0 - \frac{\partial f}{\partial y}\,\mathrm{d}x \wedge \mathrm{d}y + \frac{\partial g}{\partial x}\,\mathrm{d}x \wedge \mathrm{d}y + \frac{\partial g}{\partial y} \cdot 0 = \left(\frac{\partial g}{\partial x} - \frac{\partial f}{\partial y}\right)\mathrm{d}x \wedge \mathrm{d}y \in \Omega^2(\mathbb{R}^2).\end{aligned}$$

2. In $\mathbb{R}^3$, let $\omega = f\,\mathrm{d}x + g\,\mathrm{d}y + h\,\mathrm{d}z \in \Omega^1(\mathbb{R}^3)$. Then

$$\begin{aligned}
\mathrm{d}\omega &= \mathrm{d}f \wedge \mathrm{d}x + \mathrm{d}g \wedge \mathrm{d}y + \mathrm{d}h \wedge \mathrm{d}z \\
&= \left(\frac{\partial f}{\partial x}\,\mathrm{d}x + \frac{\partial f}{\partial y}\,\mathrm{d}y + \frac{\partial f}{\partial z}\,\mathrm{d}z\right) \wedge \mathrm{d}x + \left(\frac{\partial g}{\partial x}\,\mathrm{d}x + \frac{\partial g}{\partial y}\,\mathrm{d}y + \frac{\partial g}{\partial z}\,\mathrm{d}z\right) \wedge \mathrm{d}y + \\
&\qquad = \left(\frac{\partial h}{\partial x}\,\mathrm{d}x + \frac{\partial h}{\partial y}\,\mathrm{d}y + \frac{\partial h}{\partial z}\,\mathrm{d}z\right) \wedge \mathrm{d}z \\
&= \left(\frac{\partial g}{\partial x} - \frac{\partial f}{\partial y}\right)\mathrm{d}x \wedge \mathrm{d}y - \left(\frac{\partial f}{\partial z} - \frac{\partial h}{\partial x}\right)\mathrm{d}x \wedge \mathrm{d}z + \left(\frac{\partial h}{\partial y} - \frac{\partial g}{\partial z}\right)\mathrm{d}y \wedge \mathrm{d}z \in \Omega^2(\mathbb{R}^3).
\end{aligned}$$

3. In $\mathbb{R}^3$, let $\omega = f\,\mathrm{d}x \wedge \mathrm{d}y + g\,\mathrm{d}x \wedge \mathrm{d}z + h\,\mathrm{d}y \wedge \mathrm{d}z \in \Omega^2(\mathbb{R}^3)$. Then

$$\begin{aligned}
\mathrm{d}\omega &= \mathrm{d}f \wedge \mathrm{d}x \wedge d_y + \mathrm{d}g \wedge \mathrm{d}x \wedge \mathrm{d}z + \mathrm{d}h \wedge \mathrm{d}y \wedge \mathrm{d}z \\
&= \left(\frac{\partial f}{\partial x}\,\mathrm{d}x + \frac{\partial f}{\partial y}\,\mathrm{d}y + \frac{\partial f}{\partial z}\,\mathrm{d}z\right) \wedge \mathrm{d}x \wedge \mathrm{d}y + \left(\frac{\partial g}{\partial x}\,\mathrm{d}x + \frac{\partial g}{\partial y}\,\mathrm{d}y + \frac{\partial g}{\partial z}\,\mathrm{d}z\right) \wedge \mathrm{d}x \wedge \mathrm{d}z + \\
&\qquad = \left(\frac{\partial h}{\partial x}\,\mathrm{d}x + \frac{\partial h}{\partial y}\,\mathrm{d}y + \frac{\partial h}{\partial z}\,\mathrm{d}z\right) \wedge \mathrm{d}y \wedge \mathrm{d}z \\
&= \frac{\partial f}{\partial z}\,\mathrm{d}z \wedge \mathrm{d}x \wedge \mathrm{d}y + \frac{\partial g}{\partial y}\,\mathrm{d}y \wedge \mathrm{d}x \wedge \mathrm{d}z + \frac{\partial h}{\partial x}\,\mathrm{d}x \wedge \mathrm{d}y \wedge \mathrm{d}z \\
&= \left(\frac{\partial f}{\partial z} - \frac{\partial g}{\partial y} + \frac{\partial h}{\partial x}\right)\mathrm{d}x \wedge \mathrm{d}y \wedge \mathrm{d}z \in \Omega^3(\mathbb{R}^3). \qquad \square
\end{aligned}$$

Theorem 176
Let $\omega_1, \omega_2 \in \Omega^p(U)$. Then $d(\omega_1 + \omega_2) = d\omega_1 + d\omega_2$.

Proof: left as an exercise. ■

Lemma 177
If $f, g \in \Omega^0(\mathbb{R}^n)$, then $d(fg) = (df)g + f(dg)$.

Proof: the product $fg \in \Omega^0(\mathbb{R}^n)$ is itself a function $\mathbb{R}^n \to \mathbb{R}$. By definition,

$$\begin{aligned}
\mathrm{d}(fg) &= \sum_{i=1}^n \frac{\partial(fg)}{\partial x_i}\,\mathrm{d}x_i = \sum_{i=1}^n \left(\frac{\partial f}{\partial x_i} g + f\frac{\partial g}{\partial x_i}\right)\mathrm{d}x_i \\
&= \left(\sum_{i=1}^n \frac{\partial f}{\partial x_i}\,\mathrm{d}x_i\right) g + f\left(\sum_{i=1}^n \frac{\partial g}{\partial x_i}\,\mathrm{d}x_i\right) = (\mathrm{d}f)g + f(\mathrm{d}g). \qquad ■
\end{aligned}$$

Lemma 177 is a special case (with $p = 0$) of the more general rule for the derivative of the product of differential forms.

Theorem 178
Let $\omega \in \Omega^p(U), \omega' \in \Omega^q(U)$. *Then* $d(\omega \wedge \omega') = d\omega \wedge \omega' + (-1)^p \omega \wedge d\omega'$.

Proof: if $\{i_1, \dots, i_\ell\} \subseteq \{1, \dots, n\}$ (in increasing order) and $f \in \mathcal{C}^\infty(U, \mathbb{R})$, then

$$\mathrm{d}(f\,\mathrm{d}x_{i_1} \wedge \cdots \wedge \mathrm{d}x_{i_\ell}) = \mathrm{d}f \wedge \mathrm{d}x_{i_1} \wedge \cdots \wedge \mathrm{d}x_{i_\ell}.$$

Since $\mathrm{d}(\omega_1 + \omega_2) = \mathrm{d}\omega_1 + \mathrm{d}\omega_2$, we only need to verify the conclusion for

$$\begin{aligned} \omega &= f\,\mathrm{d}x_{i_1} \wedge \cdots \wedge \mathrm{d}x_{i_p}, \quad i_1 < \cdots < i_p \\ \omega' &= g\,\mathrm{d}x_{j_1} \wedge \cdots \wedge \mathrm{d}x_{j_q}, \quad j_1 < \cdots < j_q. \end{aligned}$$

Then

$$\begin{aligned} \mathrm{d}(\omega \wedge \omega') &= \mathrm{d}(f\,\mathrm{d}x_{i_1} \wedge \cdots \wedge \mathrm{d}x_{i_p} \wedge g\,\mathrm{d}x_{j_1} \wedge \cdots \wedge \mathrm{d}x_{j_q}) \\ \boxed{\text{thm 174.1}} &= \mathrm{d}(fg\,\mathrm{d}x_{i_1} \wedge \cdots \wedge \mathrm{d}x_{i_p} \wedge \mathrm{d}x_{j_1} \wedge \cdots \wedge \mathrm{d}x_{j_q}) \\ &= \mathrm{d}(fg) \wedge \mathrm{d}x_{i_1} \wedge \cdots \wedge \mathrm{d}x_{i_p} \wedge \mathrm{d}x_{j_1} \wedge \cdots \wedge \mathrm{d}x_{j_q} \\ \boxed{\text{lemma 177}} &= [(\mathrm{d}f)g + f(\mathrm{d}g)] \wedge \mathrm{d}x_{i_1} \wedge \cdots \wedge \mathrm{d}x_{i_p} \wedge \mathrm{d}x_{j_1} \wedge \cdots \wedge \mathrm{d}x_{j_q} \\ &= (\mathrm{d}f)g \wedge \mathrm{d}x_{i_1} \wedge \cdots \wedge \mathrm{d}x_{i_p} \wedge \mathrm{d}x_{j_1} \wedge \cdots \wedge \mathrm{d}x_{j_q} \\ &\quad + f(\mathrm{d}g)\,\mathrm{d}x_{i_1} \wedge \cdots \wedge \mathrm{d}x_{i_p} \wedge \mathrm{d}x_{j_1} \wedge \cdots \wedge \mathrm{d}x_{j_q} \\ &= \underbrace{\mathrm{d}f \wedge \mathrm{d}x_{i_1} \wedge \cdots \wedge \mathrm{d}x_{i_p}}_{=\mathrm{d}\omega} \wedge \underbrace{g\,\mathrm{d}x_{j_1} \wedge \cdots \wedge \mathrm{d}x_{j_q}}_{=\omega'} \\ &\quad + (-1)^p \underbrace{f\,\mathrm{d}x_{i_1} \wedge \cdots \wedge \mathrm{d}x_{i_p}}_{=\omega} \wedge \underbrace{\mathrm{d}g \wedge \mathrm{d}x_{j_1} \wedge \cdots \wedge \mathrm{d}x_{j_q}}_{=\mathrm{d}\omega'} \\ &= \mathrm{d}\omega \wedge \omega' + (-1)^p \omega \wedge \mathrm{d}\omega'. \qquad \blacksquare \end{aligned}$$

We illustrate this in the case where $\omega = \sum_{i=1}^n f_i\,\mathrm{d}x_i \in \Omega^1(\mathbb{R}^n)$ and $\omega' = h \in \Omega^0(\mathbb{R}^n)$. Then

$$\omega \wedge \omega' = \sum_{i=1}^n f_i h\,\mathrm{d}x_i \quad \text{and} \quad \mathrm{d}(\omega \wedge \omega') = \mathrm{d}\left(\sum_{i=1}^n f_i h\,\mathrm{d}x_i\right) = \sum_{i=1}^n \mathrm{d}(f_i h\,\mathrm{d}x_i) = \sum_{i=1}^n \mathrm{d}(f_i h) \wedge \mathrm{d}x_i$$

$$\begin{aligned} &= \sum_{i=1}^n [(\mathrm{d}f_i)h + f_i(\mathrm{d}h)] \wedge \mathrm{d}x_i = \sum_{i=1}^n (\mathrm{d}f_i \wedge \mathrm{d}x_i)h + \sum_{i=1}^n f_i\,\mathrm{d}h \wedge \mathrm{d}x_i \\ &= \mathrm{d}\omega \wedge \omega' + \sum_{i=1}^n f_i(-\,\mathrm{d}x_i \wedge \mathrm{d}h) = \mathrm{d}\omega \wedge \omega' - \omega \wedge \mathrm{d}\omega' \\ &= \mathrm{d}\omega \wedge \omega' + (-1)^1 \omega \wedge \mathrm{d}\omega'. \end{aligned}$$

The next result showcases a crucial property of exterior derivatives.

Theorem 179
Let $\omega \in \Omega^p(U)$. Then $d(d\omega) = 0$.

Proof: if $f \in \mathcal{C}^\infty(U,\mathbb{R}) = \Omega^0(U)$, then $\mathrm{d}f \in \Omega^1(U)$ and

$$\mathrm{d}(\mathrm{d}f) = \mathrm{d}\left(\sum_{i=1}^n \frac{\partial f}{\partial x_i}\,\mathrm{d}x_i\right) = \sum_{i=1}^n \mathrm{d}\left(\frac{\partial f}{\partial x_i}\right)\wedge \mathrm{d}x_i = \sum_{i=1}^n \left(\sum_{j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}\,\mathrm{d}x_j\right)\wedge \mathrm{d}x_i.$$

When $i=j$, $\mathrm{d}x_i \wedge \mathrm{d}x_j = 0$; when $i>j$, $\mathrm{d}x_i\wedge \mathrm{d}x_j = -\,\mathrm{d}x_j \wedge \mathrm{d}x_i$, so that

$$\mathrm{d}^2 f = \sum_{i<j} \underbrace{\left(\frac{\partial^2 f}{\partial x_i \partial x_j} - \frac{\partial^2 f}{\partial x_j \partial x_i}\right)}_{=0 \text{ since } f\in\mathcal{C}^\infty(U,\mathbb{R})} \mathrm{d}x_i \wedge \mathrm{d}x_j = 0.$$

Furthermore,

$$\mathrm{d}(\mathrm{d}x_i) = \mathrm{d}(1\cdot\,\mathrm{d}x_i) = d(1)\wedge \mathrm{d}x_i = 0 \wedge \mathrm{d}x_i = 0.$$

Since $\mathrm{d}(\omega+\omega') = \mathrm{d}\omega + \mathrm{d}\omega'$, it is sufficient to show that $\mathrm{d}^2(f\,\mathrm{d}x_{i_1}\wedge\cdots\wedge \mathrm{d}x_{i_p}) = 0$, where $\{i_1 < \ldots < i_p\} \subseteq \{1,\ldots,n\}$ and f is as above. As

$$\begin{aligned}\mathrm{d}(\mathrm{d}(f\,\mathrm{d}x_{i_1}\wedge\cdots\wedge \mathrm{d}x_{i_p})) &= \mathrm{d}(\mathrm{d}f \wedge \mathrm{d}x_{i_1}\wedge\cdots\wedge \mathrm{d}x_{i_p}) \\ &= \ \mathrm{d}(\mathrm{d}f)\wedge \mathrm{d}x_{i_1}\wedge\cdots\wedge \mathrm{d}x_{i_p} + (-1)^{0+1}\mathrm{d}f\wedge \mathrm{d}(\mathrm{d}x_{i_1}\wedge\cdots\wedge \mathrm{d}x_{i_p}) \\ &= 0\wedge \mathrm{d}x_{i_1}\wedge\cdots\wedge \mathrm{d}x_{i_p} - \mathrm{d}f\wedge 0 = 0. \qquad \blacksquare\end{aligned}$$

A differential form $\omega \in \Omega^p(U)$ is **closed** if $\mathrm{d}\omega = 0$.

Example: let $n=1$ and $\omega\in\Omega^1(\mathbb{R}^1)$. Then $\mathrm{d}\omega \in \Omega^2(\mathbb{R}^1)$; since $\Omega^2(\mathbb{R}^1) = \{0\}$, ω is automatically closed. $\square$

13.3 Antiderivative

Let $p>1$, $U \subseteq_O \mathbb{R}^n$ and $\omega\in\Omega^p(U)$; ω is **exact** if $\exists \eta \in \Omega^{p-1}(U)$ such that $\mathrm{d}\eta = \omega$. The differential form η is an **antiderivative of** ω. If ω is exact, then $\mathrm{d}\omega = \mathrm{d}^2\eta = 0$ and so every exact form is also closed.

If $n=1$, let $f\in\Omega^0(\mathbb{R})$. Then $\Omega^1(\mathbb{R}) = \{g\,\mathrm{d}x \mid g\in\Omega^0(\mathbb{R})\}$. If $F:\mathbb{R}\to\mathbb{R}$ is such that $F'(x) = f(x)$ for all $x\in\mathbb{R}$, then $F\in\Omega^0(\mathbb{R})$ and

$$dF = \frac{\partial F}{\partial x}\,\mathrm{d}x = f\,\mathrm{d}x.$$

Such an F exists by Theorem 60 since f is continuous on $\mathbb{R}$. Hence, every $\omega\in\Omega^1(\mathbb{R})$ is exact.

Examples

1. Let $\omega = P_1(x,y)\,\mathrm{d}x + P_2(x,y)\,\mathrm{d}y = y\,\mathrm{d}x - x\,\mathrm{d}y \in \Omega^1(\mathbb{R}^2)$. Since

$$\mathrm{d}\omega = \left(\frac{\partial P_2}{\partial x} - \frac{\partial P_1}{\partial y}\right)\,\mathrm{d}x \wedge \mathrm{d}y = (-1-1)\,\mathrm{d}x \wedge \mathrm{d}y = -2\,\mathrm{d}x \wedge \mathrm{d}y \neq 0;$$

since ω is not closed, it cannot be exact. □

2. Let $\omega = f(x,y)\,\mathrm{d}x + g(x,y)\,\mathrm{d}y = (3x^2+2xy+y^2)\,\mathrm{d}x + (x^2+2xy+3y^2)\,\mathrm{d}y \in \Omega^1(\mathbb{R}^2)$. Since

$$\begin{aligned}
\mathrm{d}\omega &= \mathrm{d}f \wedge \mathrm{d}x + \mathrm{d}g \wedge \mathrm{d}y \\
&= \left(\frac{\partial f}{\partial x}\,\mathrm{d}x + \frac{\partial f}{\partial y}\,\mathrm{d}y\right) \wedge \mathrm{d}x + \left(\frac{\partial g}{\partial x}\,\mathrm{d}x + \frac{\partial g}{\partial y}\,\mathrm{d}y\right) \wedge \mathrm{d}y \\
&= \frac{\partial f}{\partial y}\,\mathrm{d}y \wedge \mathrm{d}x + \frac{\partial g}{\partial x}\,\mathrm{d}x \wedge \mathrm{d}y = \left(\frac{\partial g}{\partial x} - \frac{\partial f}{\partial y}\right)\,\mathrm{d}x \wedge \mathrm{d}y.
\end{aligned}$$

But

$$\frac{\partial g}{\partial x} = 2x + 2y = \frac{\partial f}{\partial y}$$

in this specific case, so $\mathrm{d}\omega = 0$, which means that ω is closed.

We can show that this particular closed form is also exact, which is to say that $\exists F \in \Omega^0(\mathbb{R}^2) = \mathcal{C}^\infty(\mathbb{R}^2,\mathbb{R})$ such that $\mathrm{d}F = \omega$. If such a F exists,

$$\mathrm{d}F = \frac{\partial F}{\partial x}\,\mathrm{d}x + \frac{\partial F}{\partial y}\,\mathrm{d}y = f(x,y)\,\mathrm{d}x + g(x,y)\,\mathrm{d}y,$$

and we must have

$$\frac{\partial F}{\partial x} = f(x,y) = 3x^2 + 2xy + y^2 \quad \text{and} \quad \frac{\partial F}{\partial y} = g(x,y) = x^2 + 2xy + 3y^2.$$

Integrating the first of these with respect to x yields

$$F(x,y) = x^3 + x^2y + y^2x + \varphi(y).$$

Differentiating with respect to y yields

$$\frac{\partial F}{\partial y} = x^3 + 2xy + \varphi'(y) = x^2 + 2xy + 3y^2,$$

so that $\varphi'(y) = 3y^2$, and so $\varphi(y) = y^3 + C$. Thus the antiderivatives of ω take the form

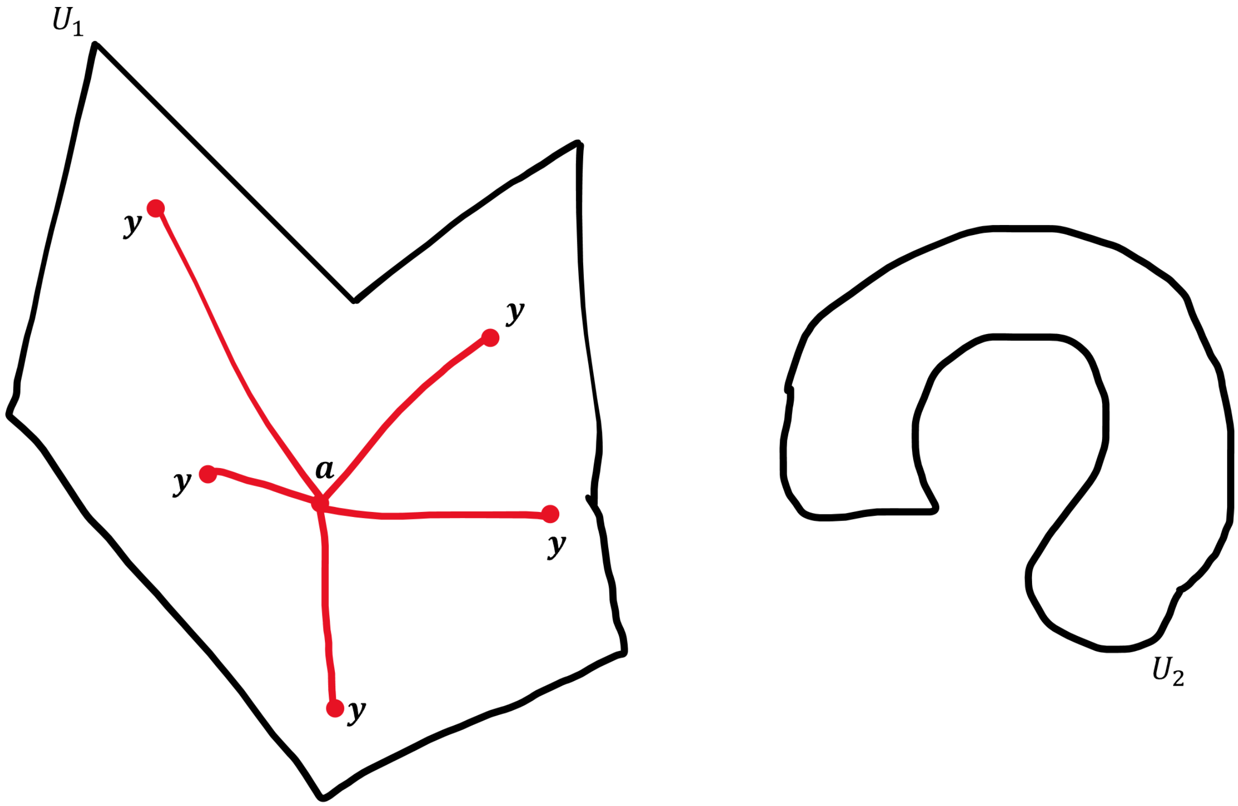
$$F(x,y) = x^3 + x^2y + xy^2 + y^3 + C,$$

where $C \in \mathbb{R}$. □

Exact forms are necessarily closed; the converse is valid when $U \subseteq_O \mathbb{R}^n$ has an additional property. A set $U \subseteq \mathbb{R}$ is **star-shaped** if $\exists \mathbf{a} \in U$ such that $\forall \mathbf{y} \in U$ we have

$$[\mathbf{a}, \mathbf{y}] = \{(1-t)\mathbf{a} + t\mathbf{y} \mid 0 \leq t \leq 1\} = \{\mathbf{a} + t(\mathbf{y} - \mathbf{a}) \mid 0 \leq t \leq 1\} \subseteq U.$$

In $\mathbb{R}^2$, for instance, U_1 (on the left) is star-shaped, whereas U_2 (on the right) is not.



We now present a highly technical lemma that will allow us to prove the desired result.

Theorem 180
Let $U \subseteq_O \mathbb{R}^n$, $I = [0,1]$, and $\varphi : U \times I \to \mathbb{R}$ a continuous function in the Euclidean metric. Then the function $\psi : U \to \mathbb{R}$ defined by

$$\psi(\mathbf{x} = \int_0^1 \varphi(\mathbf{x}, t)dt$$

is continuous.

Furthermore, if $D_{\mathbf{x}}\varphi : U \times I \to End(\mathbb{R}^n, \mathbb{R}) \simeq (\mathbb{R}^n)^*$ exists and is continuous, then ψ is C^1 and

$$D_{\mathbf{x}}\psi(\mathbf{x}) = \int_0^1 D_{\mathbf{x}}\varphi(\mathbf{x}, t)dt.$$

Proof: we start by proving the continuity of ψ. We want to show that $\forall \varepsilon > 0, \exists \delta_\varepsilon > 0$ such that

$$\|\mathbf{x} - \mathbf{x}'\| < \delta_\varepsilon \Longrightarrow |\psi(\mathbf{x}) - \psi(\mathbf{x}')| < \varepsilon.$$

For $\mathbf{x}, \mathbf{x}' \in U$, we have

$$|\psi(\mathbf{x}) - \psi(\mathbf{x}')| = \left| \int_0^1 (\varphi(\mathbf{x}, t) - \varphi(\mathbf{x}', t)) dt \right| \leq \int_0^1 |\varphi(\mathbf{x}, t) - \varphi(\mathbf{x}', t)| dt.$$

Let $\varepsilon > 0$ and $(\mathbf{x}, t) \in U \times I$. Since φ is continuous, $\exists \delta_\varepsilon = \delta_\varepsilon(\mathbf{x}, t)$ such that

$$\|\mathbf{x} - \mathbf{x}'\|, |t - t'| < \delta_\varepsilon \Longrightarrow |\varphi(\mathbf{x}, t) - \varphi(\mathbf{x}', t')| < \varepsilon/12.$$

In particular,

$$\|\mathbf{x} - \mathbf{x}'\| < \delta_\varepsilon \Longrightarrow |\varphi(\mathbf{x}, t') - \varphi(\mathbf{x}', t')| < \varepsilon/6.$$

For a $\mathbf{x}$ fixed, define $V_t = \{t' \in \mathbb{R} \mid |t - t'| < \delta_\varepsilon(\mathbf{x}, t)\} \cap I$; then $\{V_t\}_{t \in I}$ is an open cover of the subspace $I \subseteq \mathbb{R}$. But I is a compact subspace of $\mathbb{R}$ in the Euclidean topology, and so there is a finite subcover $\{V_{t_1}, \ldots, V_{t_K}\}$ of I with

$$\bigcup_{i=1}^{K} V_{t_i} = I.$$

Let $\delta_\varepsilon(\mathbf{x}) = \min\{\delta(\mathbf{x}, t_i) \mid i = 1, \ldots, K\}$. Thus for any $t' \in I$, we can find a $t_i \in I$ such that $|t_i - t'| < \delta_\varepsilon(\mathbf{x}, t_i)$. If we also have $\|\mathbf{x} - \mathbf{x}'\| < \delta_\varepsilon(\mathbf{x})$, then

$$\begin{aligned} |\varphi(\mathbf{x}, t') - \varphi(\mathbf{x}', t')| &\leq |\varphi(\mathbf{x}, t') - \varphi(\mathbf{x}, t_i)| + |\varphi(\mathbf{x}, t_i) - \varphi(\mathbf{x}', t_i)| + |\varphi(\mathbf{x}', t_i) - \varphi(\mathbf{x}', t')| \\ &< \varepsilon/6 + \varepsilon/6 + \varepsilon/6 = \varepsilon/2. \end{aligned}$$

Set $\delta_\varepsilon = \delta_\varepsilon(\mathbf{x})$. Then for all $\mathbf{x}, \mathbf{x}' \in U$ we have

$$|\psi(\mathbf{x}) - \psi(\mathbf{x}')| \leq \int_0^1 \frac{\varepsilon}{2} dt = \frac{\varepsilon}{2} < \varepsilon.$$

We now tackle the differentiability of ψ. Since $D_\mathbf{x}\varphi$ is continuous by assumption, the same argument as above shows that the function

$$\mathbf{x} \in U \mapsto \lambda(\mathbf{x}) = \int_0^1 D_\mathbf{x}\varphi(\mathbf{x}, t) dt$$

is continuous. It remains only to show that $\lambda(x) = D_\mathbf{x}\psi(\mathbf{x})$, that is, $\forall \varepsilon > 0, \exists \delta_\varepsilon > 0$ such that

$$\|\mathbf{h}\| < \delta_\varepsilon \Longrightarrow |\psi(\mathbf{x} + \mathbf{h}) - \psi(\mathbf{x}) - \lambda(\mathbf{x})\mathbf{h}| < \varepsilon \cdot \|\mathbf{h}\|.$$

But

$$|\psi(\mathbf{x} + \mathbf{h}) - \psi(\mathbf{x}) - \lambda(\mathbf{x})\mathbf{h}| = \left| \int_0^1 (\varphi(\mathbf{x} + \mathbf{h}, t) - \varphi(\mathbf{x}, t)) dt - \int_0^1 D_\mathbf{x}\varphi(\mathbf{x}, t)\mathbf{h} dt \right|$$

$$\leq \int_0^1 |\varphi(\mathbf{x}+\mathbf{h},t)-\varphi(\mathbf{x},t))-D_{\mathbf{x}}\varphi(\mathbf{x},t)\mathbf{h}|dt$$

$$\boxed{\text{Taylor's thm}} = \int_0^1 |D_{\mathbf{x}}\varphi(\mathbf{x}+\boldsymbol{\theta},t)-D_{\mathbf{x}}\varphi(\mathbf{x},t)|dt,$$

for $\boldsymbol{\theta} \in [\mathbf{0},\mathbf{h}]$. But $D_{\mathbf{x}}\varphi$ is continuous so $\forall\varepsilon>0$, $\exists\delta_\varepsilon>0$ such that

$$\|\boldsymbol{\theta}\| \leq \|\mathbf{h}\| < \delta_\varepsilon \Longrightarrow |D_{\mathbf{x}}\varphi(\mathbf{x}+\boldsymbol{\theta},t)-D_{\mathbf{x}}\varphi(\mathbf{x},t)| < \varepsilon.$$

Hence

$$|\psi(\mathbf{x}+\mathbf{h})-\psi(\mathbf{x})-\lambda(\mathbf{x})\mathbf{h}| < \int_0^1 \varepsilon\|\mathbf{h}\|dt = \varepsilon\|\mathbf{h}\|,$$

which completes the proof. ■.

And now, the *pièce de résistance*.

Theorem 181 (Poincaré's Lemma)
Let $U \subseteq \mathbb{R}^n$ be star-shaped and containing **0**. *If $\omega \in \Omega^p(U)$ is closed, then it is exact.*

Proof: we start by proving the result for $n=1$, $p=1$. Let $\omega \in \Omega^1(U)$. Then $\omega = f\,\mathrm{d}x$, with $f \in \mathcal{C}^\infty(U,\mathbb{R})$. Since $\Omega^2(U)=\{0\}$, we have $\mathrm{d}\omega = 0 \in \Omega^2(U)$. We show that $\exists F \in \Omega^0(U)$ such that $\mathrm{d}F=\omega$.

Recall that

$$F(x) = \int_0^x f(t)dt = \int_0^1 f(xs)xds = \int_0^1 g(x,s)ds.$$

According to Lemma 180,

$$F'(x) = \int_0^1 \frac{\partial g}{\partial x}(x,s)ds = \int_0^1 (f(xs)+sf'(xs))ds$$
$$= \int_0^1 \frac{d}{ds}[sf(x,s)]ds = 1\cdot f(x,1) - 0\cdot f(x,0) = f(x).$$

Hence $\mathrm{d}F = \frac{\partial F}{\partial x}\,\mathrm{d}x = F'(x)\,\mathrm{d}x = f(x)\,\mathrm{d}x = \omega$.

Now suppose that $n>1$, $p=1$. Let $\omega \in \Omega^1(U)$ with $\mathrm{d}\omega=0$. We want to show $\exists\eta = F \in \Omega^0(U) = \mathcal{C}^\infty(U,\mathbb{R})$ such that $\mathrm{d}\eta=\omega$. By hypothesis,

$$\omega = \sum_{i=1}^n f_i\,\mathrm{d}x_i, \quad \text{with } f_i \in \mathcal{C}^\infty(U,\mathbb{R})$$

and

$$\mathrm{d}\omega = \sum_{i=1}^{n} \mathrm{d}f_i \wedge \mathrm{d}x_i = \sum_{i,j=1}^{n} \frac{\partial f_i}{\partial x_j} \,\mathrm{d}x_i \wedge \mathrm{d}x_j = \sum_{i<j} \left(\frac{\partial f_i}{\partial x_j} - \frac{\partial f_j}{\partial x_i} \right) \mathrm{d}x_i \wedge \mathrm{d}x_j = 0,$$

and so

$$\frac{\partial f_i}{\partial x_j} = \frac{\partial f_j}{\partial x_i}, \quad \text{for all } 1 \le i < j \le n.$$

Let

$$F(\mathbf{x}) = F(x_1, \ldots, x_n) = \sum_{i=1}^{n} f_i(\mathbf{x}s)x_i ds = \sum_{i=1}^{n} \underbrace{f_i(x_1 s, \ldots, x_n s)x_i}_{=g_i(\mathbf{x},s)} ds.$$

We show that $\mathrm{d}F = \omega$:

$$\begin{aligned}
\frac{\partial F}{\partial x_1}(\mathbf{x}) &= \sum_{i=1}^{n} \frac{\partial}{\partial x_1} \int_0^1 g_i(\mathbf{x}, s)ds = \sum_{i=1}^{n} \int_0^1 \frac{\partial}{\partial x_1} g_i(\mathbf{x}, s)ds \\
&= \int_0^1 \left[f_1(\mathbf{x}s) + x_1 s \frac{\partial}{\partial x_1} f(\mathbf{x}s) \right] ds + \sum_{j=2}^{n} \int_0^1 x_j s \frac{\partial}{\partial x_1} f_j(\mathbf{x}s)ds \\
&= \int_0^1 \left[f_1(\mathbf{x}s) + \sum_{j=1}^{n} x_j s \frac{\partial}{\partial x_j} f_1(\mathbf{x}s) \right] ds,
\end{aligned}$$

by the equality of partial derivatives above. Set $k_1(s) = sf_1(\mathbf{x}s)$. Then

$$k_1'(s) = f_1(\mathbf{x}s) + \sum_{j=1}^{n} x_j s \frac{\partial}{\partial x_j} f_1(\mathbf{x}s),$$

so that

$$\frac{\partial F}{\partial x_1}(\mathbf{x}) = \int_0^1 k'(s)ds = k(1) - k(0) = f_1(\mathbf{x}).$$

In a similar fashion, we can see that

$$\frac{\partial F}{\partial x_i}(\mathbf{x}) = f_i(\mathbf{x}), \quad \text{for all } 1 \le j \le n,$$

and so

$$\mathrm{d}F = \sum_{i=1}^{n} \frac{\partial F}{\partial x_i} \,\mathrm{d}x_i = \sum_{i=1}^{n} f_i \,\mathrm{d}x_i = \omega.$$

We will not be providing the proof for $p > 1$. ∎

Where exactly was the hypothesis that U is star-shaped used?[3]

[3]Hint: look at the definition of $F(x)$ (in the case $n = 1$) and $F(\mathbf{x})$ (in the case $n > 1$).

In a nutshell, we have shown the following result.

> **Proposition 182**
> *Let $U \subseteq_O \mathbb{R}$ and $\omega = \sum_{i=1}^n f_i\, dx_i \in \Omega^1(U)$. Consider the following conditions:*
>
> 1. *ω is exact in U;*
> 2. *ω is closed in U;*
> 3. *$\frac{\partial f_i}{\partial x_j} = \frac{\partial f_j}{\partial x_i}$ for all i, j.*
>
> Then $1. \implies 2. \iff 3.$ Furthermore, if U is star-shaped, then the three conditions are equivalent.

13.4 Pullback of a Differential Form

Let $U \subseteq_O \mathbb{R}^m$, $V \subseteq_O \mathbb{R}^n$, $\mathbf{g} \in \mathcal{C}^\infty(U, V)$.[4] The **pullback** function $\mathbf{g}^* : \Omega^k(V) \to \Omega^k(U)$ satisfies

$$\mathbf{g}^*(\bigwedge_i \omega_i) = \bigwedge_i \mathbf{g}^*(\omega_i).$$

We define it as follows.

Case $k = 0$: if $f \in \mathcal{C}^\infty(V, \mathbb{R}) = \Omega^0(V)$, the pullback is

$$\mathbf{g}^*(f) = f \circ \mathbf{g} : U \to \mathbb{R} \in \mathcal{C}^\infty(U, V) = \Omega^0(U).$$

Case $k = 1$: if a smooth $\mathbf{g} : U \subseteq_o \mathbb{R}^m \to V \subseteq_O \mathbb{R}^n$ maps

$$\mathbf{u} = (u_1, \ldots, u_m) \in U \mapsto \mathbf{v} = \mathbf{g}(\mathbf{u}) = (g_1(\mathbf{u}), \ldots, g_n(\mathbf{u})) \in V,$$

and $\omega \in \Omega^1(V)$, then $\omega = \sum_{i=1}^n f_i\, \mathrm{d}x_i$ and the pullback is

$$\mathbf{g}^*(\omega) = \sum_{i=1}^n \mathbf{g}^*(f_i)\mathbf{g}^*(\mathrm{d}x_i) = \sum_{i=1}^n (f_i \circ \mathbf{g})\, \mathrm{d}g_i = \sum_{i=1}^n (f_i \circ \mathbf{g}) \left(\sum_{j=1}^m \frac{\partial g_i}{\partial u_j}\, \mathrm{d}u_j \right).$$

Let us take a look at some examples.

> **Examples**
>
> 1. Let $\mathbf{g} : U = \mathbb{R} \to V = \mathbb{R}$ and consider $\omega = f\, \mathrm{d}x \in \Omega^1(V)$. Then the pullback $\mathbf{g}^*(\omega) \in \Omega^1(U)$ is given by
>
> $$\mathbf{g}^*(\omega)(u) = (f \circ g)\mathbf{g}^*(\mathrm{d}x)(u) = f(\mathbf{g}(u)) \cdot \mathbf{g}'(u)\, \mathrm{d}u. \qquad \square$$

[4] We will encounter such functions when we discuss vector fields in Section 13.5.

2. Let $\mathbf{g} : U = \mathbb{R} \to V = \mathbb{R}^2$ be defined by

$$\mathbf{g}(t) = (\cos t, \sin t)$$

and $\omega = -y\,\mathrm{d}x + x\,\mathrm{d}y \in \Omega^1(V)$. Then

$$\mathbf{g}^*(\mathrm{d}x)(t) = (\mathrm{d}g_1)(t) = -\sin t\,\mathrm{d}t, \quad \mathbf{g}^*(\mathrm{d}y)(t) = (\mathrm{d}g_2)(t) = \cos t\,\mathrm{d}t,$$

and the pullback $\mathbf{g}^*(\omega) \in \Omega^1(U)$ is given by

$$\begin{aligned}\mathbf{g}^*(\omega)(t) &= f_1(\mathbf{g}(t))(\mathrm{d}g_1)(t) + f_2(\mathbf{g}(t))(\mathrm{d}g_2)(t)\\ &= (-\sin t)(-\sin t\,\mathrm{d}t) + (\cos t)(\cos t\,\mathrm{d}t) = (\sin^2 t + \cos^2 t)\,\mathrm{d}t = \mathrm{d}t. \qquad \square\end{aligned}$$

3. Let $\mathbf{g} : U = \mathbb{R}^2 \to V = \mathbb{R}^2$ be defined by

$$\mathbf{g}(\mathbf{u}) = (g_1(u_1, u_2), g_2(u_1, u_2)) = (u_1 \cos u_2, u_1 \sin u_2)$$

and $\omega = f_1(x_1, x_2)\,\mathrm{d}x_1 + f_2(x_1, x_2)\,\mathrm{d}x_2 = x_1\,\mathrm{d}x_1 + x_2\,\mathrm{d}x_2 \in \Omega^1(V)$. Then

$$\begin{aligned}\mathbf{g}^*(\mathrm{d}x_1)(u_1, u_2) &= (\mathrm{d}g_1)(u_1, u_2) = \frac{\partial g_1(u_1, u_2)}{\partial u_1}\,\mathrm{d}u_1 + \frac{\partial g_1(u_1, u_2)}{\partial u_2}\,\mathrm{d}u_2\\ &= \cos u_2\,\mathrm{d}u_1 - u_1 \sin u_2\,\mathrm{d}u_2\\ \mathbf{g}^*(\mathrm{d}x_2)(u_1, u_2) &= (\mathrm{d}g_2)(u_1, u_2) = \frac{\partial g_2(u_1, u_2)}{\partial u_1}\,\mathrm{d}u_1 + \frac{\partial g_2(u_1, u_2)}{\partial u_2}\,\mathrm{d}u_2\\ &= \sin u_2\,\mathrm{d}u_1 + u_1 \cos u_2\,\mathrm{d}u_2,\end{aligned}$$

and the pullback $\mathbf{g}^*(\omega) \in \Omega^1(U)$ is given by

$$\begin{aligned}\mathbf{g}^*(\omega)(u_1, u_2) &= f_1(\mathbf{g}(u_1, u_2))(\mathrm{d}g_1)(u_1, u_2) + f_2(\mathbf{g}(u_1, u_2))(\mathrm{d}g_2)(u_1, u_2)\\ &= u_1 \cos u_2(\cos u_2\,\mathrm{d}u_1 - u_1 \sin u_2\,\mathrm{d}u_2) + u_1 \sin u_2(\sin u_2\,\mathrm{d}u_1 + u_1 \cos u_2\,\mathrm{d}u_2)\\ &= u_1(\cos^2 u_2 + \sin^2 u_2), \mathrm{d}u = u_1\,\mathrm{d}u_1. \qquad \square\end{aligned}$$

Case $k > 1$: if $\mathbf{g} : U \subseteq_O \mathbb{R}^m \to V \subseteq_O \mathbb{R}^n$ is smooth and $\omega = \mathrm{d}x_{i_1} \wedge \cdots \wedge \mathrm{d}x_{i_k} \in \Omega^k(V)$, we define the pullback

$$\mathbf{g}^*(\omega) = \mathbf{g}^*(\mathrm{d}x_{i_1} \wedge \cdots \wedge \mathrm{d}x_{i_k}) = \mathrm{d}g_{i_1} \wedge \cdots \wedge \mathrm{d}g_{i_k} \in \Omega^k(U).$$

If

$$\omega = \sum_{i_1 < \cdots < i_k} P_{i_1, \cdots, i_k} \mathrm{d}x_{i_1} \wedge \cdots \wedge \mathrm{d}x_{i_k} \in \Omega^k(V),$$

then the pullback is

$$\mathbf{g}^*(\omega) = \sum_{i_1 < \cdots < i_k} \mathbf{g}^*(P_{i_1, \cdots, i_k})\mathbf{g}^*(\mathrm{d}x_{i_1} \wedge \cdots \wedge \mathrm{d}x_{i_k}) = \sum_{i_1 < \cdots < i_k} (P_{i_1, \cdots, i_k} \circ \mathbf{g})\mathrm{d}g_{i_1} \wedge \cdots \wedge \mathrm{d}g_{i_k} \in \Omega^k(U).$$

Example: let $\mathbf{g} : U = \mathbb{R}^2 \to V = \mathbb{R}^2$ be defined by

$$\mathbf{g}(\mathbf{u}) = (g_1(u_1, u_2), g_2(u_1, u_2)) = (u_1 \cos u_2, u_1 \sin u_2)$$

and $\omega = \mathrm{d}x_1 \wedge \mathrm{d}x_2 \in \Omega^2(V)$. Then

$$(\mathrm{d}g_1)(u_1, u_2) = \cos u_2 \,\mathrm{d}u_1 - u_1 \sin u_2 \,\mathrm{d}u_2, \quad (\mathrm{d}g_2)(u_1, u_2) = \sin u_2 \,\mathrm{d}u_1 + u_1 \cos u_2 \,\mathrm{d}u_2,$$

and the pullback $\mathbf{g}^*(\omega) \in \Omega^2(U)$ is given by

$$\begin{aligned} \mathbf{g}^*(\omega)(u_1, u_2) &= \mathbf{g}^*(\mathrm{d}x_1 \wedge \mathrm{d}x_2)(u_1, u_2) = (\mathrm{d}g_1)(u_1, u_2) \wedge (\mathrm{d}g_2)(u_1, u_2) \\ &= (\cos u_2 \,\mathrm{d}u_1 - u_1 \sin u_2 \,\mathrm{d}u_2) \wedge (\sin u_2 \,\mathrm{d}u_1 + u_1 \cos u_2 \,\mathrm{d}u_2) \\ &= u_1 \cos^2 u_2 \,\mathrm{d}u_1 \wedge \mathrm{d}u_2 - u_1 \sin^2 u_2 \,\mathrm{d}u_2 \wedge \mathrm{d}u_1 \\ &= u_1(\cos^2 u_2 + \sin^2 u_2) \,\mathrm{d}u_1 \wedge \mathrm{d}u_2 = u_1 \,\mathrm{d}u_1 \wedge \mathrm{d}u_2. \qquad \square \end{aligned}$$

While none of the computations are particularly difficult to perform (although they can be tedious), there is a simpler way to express pullbacks, as the following discussion illustrates.

If $\mathbf{g} : U = \mathbb{R}^2 \to V = \mathbb{R}^2$ is smooth, then the pullback of $\mathrm{d}x_1 \wedge \mathrm{d}x_2 \in \Omega^2(V)$ by $\mathbf{g}$ is

$$\begin{aligned} \mathbf{g}^*(\mathrm{d}x_1 \wedge \mathrm{d}x_2) = \mathrm{d}g_1 \wedge \mathrm{d}g_2 &= \left(\frac{\partial g_1}{\partial u_1} \,\mathrm{d}u_1 + \frac{\partial g_1}{\partial u_2} \,\mathrm{d}u_2 \right) \wedge \left(\frac{\partial g_2}{\partial u_1} \,\mathrm{d}u_1 + \frac{\partial g_2}{\partial u_2} \,\mathrm{d}u_2 \right) \\ &= \left(\frac{\partial g_1}{\partial u_1}\frac{\partial g_2}{\partial u_2} - \frac{\partial g_1}{\partial u_2}\frac{\partial g_2}{\partial u_1} \right) \mathrm{d}u_1 \wedge \mathrm{d}u_2 = \det(D\mathbf{g}) \,\mathrm{d}u_1 \wedge \mathrm{d}u_2 \in \Omega^2(U), \end{aligned}$$

where $D\mathbf{g}$ is the **Jacobian matrix of g** (see Section 21.7).

Generally, if $\mathbf{g} : U \subseteq_O \mathbb{R}^m \to V \subseteq_O \mathbb{R}^m$ is smooth, then the pullback of $\mathrm{d}x_{i_1} \wedge \cdots \wedge \mathrm{d}x_{i_k} \in \Omega^k(V)$ by $\mathbf{g}$ is

$$\begin{aligned} \mathbf{g}^*(\mathrm{d}x_{i_1} \wedge \cdots \wedge \mathrm{d}x_{i_k}) = \mathrm{d}g_{i_1} \wedge \cdots \wedge \mathrm{d}g_{i_k} &= \left(\sum_{j=1}^m \frac{\partial g_{i_1}}{\partial u_j} \,\mathrm{d}u_j \right) \wedge \cdots \wedge \left(\sum_{j=1}^m \frac{\partial g_{i_k}}{\partial u_j} \,\mathrm{d}u_j \right) \\ &= \sum_{j_1 < \cdots < j_k} \det \begin{pmatrix} \frac{\partial g_{i_1}}{\partial u_{j_1}} & \cdots & \frac{\partial g_{i_1}}{\partial u_{j_k}} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_{i_k}}{\partial u_{j_1}} & \cdots & \frac{\partial g_{i_k}}{\partial u_{j_k}} \end{pmatrix} \mathrm{d}u_{j_1} \wedge \cdots \wedge \mathrm{d}u_{j_k} \in \Omega^k(U). \end{aligned}$$

If $U, V \subseteq_O \mathbb{R}^n$, $g : U \to V$ smooth, $f \in \mathcal{C}^\infty(V, \mathbb{R})$, and $\omega = f \,\mathrm{d}x_1 \wedge \cdots \wedge \mathrm{d}x_n \in \Omega^n(V)$, then the pullback of ω by $\mathbf{g}$ is

$$\mathbf{g}^*(\omega) = (f \circ \mathbf{g}) \,\mathrm{d}g_1 \wedge \cdots \wedge \mathrm{d}g_n = \mathbf{g}^*(f) \det(D\mathbf{g}) \,\mathrm{d}u_1 \wedge \cdots \wedge \mathrm{d}u_n \in \Omega^n(U).$$

The pullback commutes with the exterior derivative for $0-$differential forms.

Lemma 183
With the usual assumptions of this section, if $f \in \Omega^0(V)$, then $d(\mathbf{g}^*(f)) = \mathbf{g}^*(df)$.

Proof: we use the definition and see that

$$\begin{aligned}\mathrm{d}(\mathbf{g}^*(f)) = \mathrm{d}(f\circ\mathbf{g}) &= \sum_{j=1}^m \frac{\partial(f\circ\mathbf{g})}{\partial u_j}\,\mathrm{d}u_j = \sum_{j=1}^m \left(\sum_{i=1}^n \left(\frac{\partial f}{\partial x_i}\circ\mathbf{g}\right)\frac{\partial g_i}{\partial u_j}\right)\mathrm{d}u_j \\ &= \sum_{i=1}^n \left(\frac{\partial f}{\partial x_i}\circ\mathbf{g}\right)\left(\sum_{j=1}^m \frac{\partial g_i}{\partial u_j}\,\mathrm{d}u_j\right) = \mathbf{g}^*\left(\sum_{i=1}^n \frac{\partial f}{\partial x_i}\,\mathrm{d}x_i\right) = \mathbf{g}^*(\mathrm{d}f),\end{aligned}$$

which completes the proof. ■

But this result does not apply solely to $\Omega^0(V)$.

Proposition 184
Let $\mathbf{g} : U \subseteq_O \mathbb{R}^m \to V \subseteq_O \mathbb{R}^m$ be smooth. If $\omega \in \Omega^0(V)$, then $d(\mathbf{g}^*(\omega)) = \mathbf{g}^*(d\omega)$.

Proof: the case $k = 0$ was proven in Lemma 183. For $k > 0$, since $\mathrm{d}(\omega_1+\omega_2) = \mathrm{d}\omega_1 + \mathrm{d}\omega_2$ and

$$\omega = \sum_{i_1<\cdots<i_k} f_{i_1,\ldots,i_k}\,\mathrm{d}x_{i_1}\wedge\cdots\wedge\mathrm{d}x_{i_l}, \quad f_{i_1,\ldots,i_k}\in\Omega^0(V),$$

it is sufficient to show that

$$\mathbf{g}^*(\mathrm{d}(f\,\mathrm{d}x_{i_1}\wedge\cdots\wedge\mathrm{d}x_{i_k})) = \mathrm{d}\left(\mathbf{g}^*(f\,\mathrm{d}x_{i_1}\wedge\cdots\wedge\mathrm{d}x_{i_k})\right).$$

But the left side of this equation reduces to

$$\begin{aligned}\mathbf{g}^*(\mathrm{d}(f\,\mathrm{d}x_{i_1}\wedge\cdots\wedge\mathrm{d}x_{i_k})) &= \mathbf{g}^*(\mathrm{d}f)\wedge\mathbf{g}^*(\mathrm{d}x_{i_1}\wedge\cdots\wedge\mathrm{d}x_{i_k}) \\ \boxed{\text{lemma 183}} &= \mathrm{d}(\mathbf{g}^*(f))\wedge\mathbf{g}^*(\mathrm{d}x_{i_1}\wedge\cdots\wedge\mathrm{d}x_{i_k}) \\ &= \mathrm{d}(\mathbf{g}^*(f))\wedge(\mathrm{d}g_{i_1}\wedge\cdots\wedge\mathrm{d}g_{i_k}).\end{aligned}$$

Thanks to repeated use of Theorem 177, the right side, on the other hand, reduces to

$$\begin{aligned}\mathrm{d}(f\circ\mathbf{g}\,\mathrm{d}g_{i_1}\wedge\cdots\wedge\mathrm{d}g_{i_k}) &= \mathrm{d}(f\circ\mathbf{g})\wedge\mathrm{d}g_{i_1}\wedge\cdots\wedge\mathrm{d}g_{i_k} + (-1)^0(f\circ\mathbf{g})\underbrace{\mathrm{d}(\mathrm{d}g_{i_1}\wedge\cdots\wedge\mathrm{d}g_{i_k})}_{=0} \\ &= \mathrm{d}(f\circ\mathbf{g})\wedge\mathrm{d}g_{i_1}\wedge\cdots\wedge\mathrm{d}g_{i_k}. \quad ■\end{aligned}$$

The machinery we have developed up to now may seem hopelessly formal and mechanical; its practical value comes through once we identify differential forms with vector fields.

13.5 Vector Fields

Let $U \subseteq_O \mathbb{R}^n$. A **vector field** is a function $\mathbf{F} : U \to \mathbb{R}^n$; it is of class $\mathbf{C}^k$ if $\mathbf{F} \in \mathcal{C}^k(U, \mathbb{R}^n)$. A function $f : U \to \mathbb{R}$ is called a **scalar field.**

> **Example:** let $f : U \to \mathbb{R}$ be continuously differentiable and consider $\nabla f : U \to \mathbb{R}^n$ defined by
> $$\nabla f(\mathbf{u}) = \left(\frac{\partial f(\mathbf{u})}{\partial x_1}, \ldots, \frac{\partial f(\mathbf{u})}{\partial x_n}\right).$$
> Then f is a scalar field and ∇f is a vector field.

We can associate to any vector field $\mathbf{F} : U \to \mathbb{R}^n$, defined by $\mathbf{F}(\mathbf{x}) = (F_1(\mathbf{x}), \ldots, F_n(\mathbf{x}))$ a unique differential form $\omega_{\mathbf{F}} \in \Omega^1(U)$ defined by

$$\omega_{\mathbf{F}} = F_1 \, \mathrm{d}x_1 + \cdots + F_n \, \mathrm{d}x_n.$$

In particular, if $f : U \to \mathbb{R}$ is smooth, the differential form associated to ∇f is

$$\omega_{\nabla f} = \frac{\partial f}{\partial x_1} \, \mathrm{d}x_1 + \cdots + \frac{\partial f}{\partial x_n} \, \mathrm{d}x_n = \mathrm{d}f \in \Omega^1(U).$$

> **Theorem 185**
> *Let $U \subseteq_O \mathbb{R}^n$ and $\mathbf{F} : U \to \mathbb{R}^n$ be smooth. Consider the following conditions:*
>
> 1. $\mathbf{F} = \nabla f$ *for some* $f : U \to \mathbb{R}$ *smooth;*
> 2. $\frac{\partial F_i}{\partial x_j} = \frac{\partial F_j}{\partial x_i}$ *for all* i, j.
>
> *Then* $1. \implies 2.$ *If U is star-shaped then, the conditions are equivalent.*
>
> **Proof:** if $\mathbf{F} = \nabla f$, then $\omega_{\mathbf{F}} = \omega_{\nabla f} = \mathrm{d}f \in \Omega^1(U)$ is exact and so condition 2. holds according to Proposition 182.
>
> If U is star-shaped and $\frac{\partial F_i}{\partial x_j} = \frac{\partial F_j}{\partial x_i}$ for all i, j, then $\omega_{\mathbf{F}} = F_1 \, \mathrm{d}x_1 + \cdots + \mathrm{d}x_n$ is exact (again, by Theorem 182), so that
> $$\omega_{\mathbf{F}} = \mathrm{d}f = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i} \, \mathrm{d}x_i$$
> for some $f : U \to \mathbb{R} \in \Omega^0(U)$. By unicity of $\omega_{\mathbf{F}}$, we must have $F_i = \frac{\partial f}{\partial x_i}$ for all i, which is to say that $\mathbf{F} = \nabla f$. ■

When $\mathbf{F} = \nabla f$, we say that $\mathbf{F}$ is a **conservative vector field** (or a **gradient field**) and that f is a **scalar potential** for $\mathbf{F}$.

Until the end of the chapter, we work with vector fields $\mathbf{F}: U \subseteq_O \mathbb{R}^3 \to \mathbb{R}^3$. Recall that, seen as a vector field over $\mathbb{R}$,

$$\dim\left(\Lambda^p(\mathbb{R}^n)\right) = \binom{n}{p},$$

according to Corollary 168; in that case, we have

$$\dim(\Lambda^1(\mathbb{R}^3)) = \dim(\Lambda^2(\mathbb{R}^3)), \quad \dim(\Lambda^0(\mathbb{R}^3)) = \dim(\Lambda^3(\mathbb{R}^3)) = 1.$$

Consider the vector space isomorphism $\Phi_1 : \mathbb{R}^3 \to \Lambda^1(\mathbb{R}^3)$ defined by

$$\Phi_1(\mathbf{a}) = \Phi_1(a_1, a_2, a_3) = a_1\,\mathrm{d}x_1 + a_2\,\mathrm{d}x_2 + a_3\,\mathrm{d}x_3.$$

If we "multiply" two vectors in $\mathbb{R}^3$, we should get the same "result" as if we "multiply" two $1-$forms over $\mathbb{R}^3$; the problem is that we while the wedge product can play the role of a multiplication, the wedge product of two $1-$forms over $\mathbb{R}^3$ is a $2-$form over $\mathbb{R}^3$.

Over other spaces this would be a deal-breaker, but over $\mathbb{R}^3$ the problem evaporates once we introduce a second vector space isomorphism $\Phi_2 : \mathbb{R}^3 \to \Lambda^2(\mathbb{R}^3)$, defined by

$$\Phi_2(\mathbf{a}) = \Phi_2(a_1, a_2, a_3) = a_1\,\mathrm{d}x_2 \wedge \mathrm{d}x_3 + a_2\,\mathrm{d}x_3 \wedge \mathrm{d}x_1 + a_3\,\mathrm{d}x_1 \wedge \mathrm{d}x_2,$$

and define the **cross-product** over $\mathbb{R}^3$ by

$$\begin{aligned}
\mathbf{a} \times \mathbf{b} &= (a_1, a_2, a_3) \times (b_1, b_2, b_3) \\
&\simeq \Phi_1(a_1, a_2, a_3) \wedge \Phi_1(b_1, b_2, b_3) \\
&= (a_1\,\mathrm{d}x_1 + a_2\,\mathrm{d}x_2 + a_3\,\mathrm{d}x_3) \wedge (b_1\,\mathrm{d}x_1 + b_2\,\mathrm{d}x_2 + b_3\,\mathrm{d}x_3) \\
&= (a_2b_3 - a_3b_2)\,\mathrm{d}x_2 \wedge \mathrm{d}x_3 + (a_3b_1 - a_1b_3)\,\mathrm{d}x_1 \wedge \mathrm{d}x_2 + (a_1b_2 - a_2b_1)\,\mathrm{d}x_1 \wedge \mathrm{d}x_2 \\
&\simeq \Phi_2^{-1}((a_2b_3 - a_3b_2)\,\mathrm{d}x_2 \wedge \mathrm{d}x_3 + (a_3b_1 - a_1b_3)\,\mathrm{d}x_1 \wedge \mathrm{d}x_2 + (a_1b_2 - a_2b_1)\,\mathrm{d}x_1 \wedge \mathrm{d}x_2) \\
&= (a_2b_3 - a_3b_2, a_3b_1 - a_1b_3, a_1b_2 - a_2b_1),
\end{aligned}$$

which should go some way towards elucidating the mystery of where the apparently random definition of the cross-product come from when it is first introduced in linear algebra courses.

In applications, it is typical to use $x = x_1$, $y = x_2$, and $z = x_3$. In that case, we could also write the vector field $\mathbf{F} : U \to \mathbb{R}^3$ as

$$\mathbf{F}(x, y, z) = (P(x, y, z), Q(x, y, z), R(x, y, z));$$

the composition

$$\Phi_1 \circ \mathbf{F} = \omega_{\mathbf{F}} = P\,\mathrm{d}x + Q\,\mathrm{d}y + R\,\mathrm{d}z \in \Omega^1(U)$$

is the corresponding differential $1-$form over U.

Then, we have:

$$\begin{aligned}
\mathrm{d}\omega_{\mathbf{F}} &= \mathrm{d}P \wedge \mathrm{d}x + \mathrm{d}Q \wedge \mathrm{d}y + \mathrm{d}R \wedge \mathrm{d}z \\
&= \left(\frac{\partial P}{\partial x}\,\mathrm{d}x + \frac{\partial P}{\partial y}\,\mathrm{d}y + \frac{\partial P}{\partial z}\,\mathrm{d}z\right) \wedge \mathrm{d}x + \left(\frac{\partial Q}{\partial x}\,\mathrm{d}x + \frac{\partial Q}{\partial y}\,\mathrm{d}y + \frac{\partial Q}{\partial z}\,\mathrm{d}z\right) \wedge \mathrm{d}y \\
&\quad + \left(\frac{\partial R}{\partial x}\,\mathrm{d}x + \frac{\partial R}{\partial y}\,\mathrm{d}y + \frac{\partial R}{\partial z}\,\mathrm{d}z\right) \wedge \mathrm{d}z \\
&= \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right) \mathrm{d}y \wedge \mathrm{d}z + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right) \mathrm{d}z \wedge \mathrm{d}x + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) \mathrm{d}x \wedge \mathrm{d}y \in \Omega^2(U).
\end{aligned}$$

The vector field $\Phi_2^{-1}(\mathrm{d}\omega_{\mathbf{F}}) = \Phi_2^{-1}(\Phi_1(\mathbf{F}))$ associated with $\mathrm{d}\omega_{\mathbf{F}}$ is the **curl of F** and is denoted by $\mathrm{curl}(\mathbf{F}) = \nabla \times \mathbf{F} : U \to \mathbb{R}^3$ and

$$\mathrm{curl}(\mathbf{F}) = \nabla \times \mathbf{F} = \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}, \frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}, \frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right).$$

Theorem 186
Let $U =\subseteq_O \mathbb{R}^3$ and $\mathbf{F} : U \to \mathbb{R}^3$ be smooth. Consider the following conditions:

1. $\mathbf{F} = \nabla f$ *for some smooth* $f : U \to \mathbb{R}$;
2. $\nabla \times \mathbf{F} = \mathbf{0}$.

Then $1. \implies 2.$ *If U is star-shaped then, the conditions are equivalent.*

Proof: direct application of Theorem 185. ■

If instead we consider the composition

$$\Phi_2 \circ \mathbf{F} = \varphi_{\mathbf{F}} = P\,\mathrm{d}y \wedge \mathrm{d}x + Q\,\mathrm{d}z \wedge \mathrm{d}x + R\,\mathrm{d}x \wedge \mathrm{d}z \in \Omega^2(U),$$

then we have

$$\begin{aligned}
\mathrm{d}\varphi_{\mathbf{F}} &= \mathrm{d}P \wedge \mathrm{d}y \wedge \mathrm{d}z + \mathrm{d}Q \wedge \mathrm{d}z \wedge \mathrm{d}x + \mathrm{d}R \wedge \mathrm{d}x \wedge \mathrm{d}y \\
&= \left(\frac{\partial P}{\partial x}\,\mathrm{d}x + \frac{\partial P}{\partial y}\,\mathrm{d}y + \frac{\partial P}{\partial z}\,\mathrm{d}z\right) \wedge \mathrm{d}y \wedge \mathrm{d}z + \left(\frac{\partial Q}{\partial x}\,\mathrm{d}x + \frac{\partial Q}{\partial y}\,\mathrm{d}y + \frac{\partial Q}{\partial z}\,\mathrm{d}z\right) \wedge \mathrm{d}z \wedge \mathrm{d}x \\
&\quad + \left(\frac{\partial R}{\partial x}\,\mathrm{d}x + \frac{\partial R}{\partial y}\,\mathrm{d}y + \frac{\partial R}{\partial z}\,\mathrm{d}z\right) \wedge \mathrm{d}x \wedge \mathrm{d}y \\
&= \frac{\partial P}{\partial x}\,\mathrm{d}x \wedge \mathrm{d}y \wedge \mathrm{d}z + \frac{\partial Q}{\partial y}\,\mathrm{d}y \wedge \mathrm{d}z \wedge \mathrm{d}x + \frac{\partial R}{\partial z}\,\mathrm{d}z \wedge \mathrm{d}x \wedge \mathrm{d}y \\
&= \left(\frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}\right) \mathrm{d}x \wedge \mathrm{d}y \wedge \mathrm{d}z \in \Omega^3(U).
\end{aligned}$$

The scalar field associated with $\mathrm{d}\varphi_{\mathbf{F}}$ is the **divergence of F** and is denoted by $\mathrm{div}(\mathbf{F}) = \nabla \cdot \mathbf{F} : U \to \mathbb{R}$ and

$$\mathrm{div}(\mathbf{F}) = \nabla \cdot \mathbf{F} = \frac{\partial P}{\partial x} + \frac{\partial Q}{\partial y} + \frac{\partial R}{\partial z}.$$

As a consequence of Poincaré's lemma, we obtain the following result.

Theorem 187
Let $U =\subseteq_O \mathbb{R}^3$ *and* $\mathbf{F} : U \to \mathbb{R}^3$ *be smooth. If there is a vector field* $\mathbf{G} : U \to \mathbb{R}^3$ *such that curl*$(\mathbf{G}) = \nabla \times \mathbf{G} = \mathbf{F}$*, then div*$(\mathbf{F}) = \nabla \cdot \mathbf{F} = 0$*. If* U *is star-shaped and div*$(\mathbf{F}) = \nabla \cdot \mathbf{F} = 0$*, then there is a* $\mathbf{G} : U \to \mathbb{R}^3$ *such that curl*$(\mathbf{G}) = \nabla \times \mathbf{G} = \mathbf{F}$*.*

Proof: let $\omega_{\mathbf{G}} \in \Omega^1(U)$ and $\varphi_{\mathbf{F}} \in \Omega^2(U)$ be the associated differential forms. If $\mathrm{curl}(\mathbf{G}) = \mathbf{F}$, then $\mathrm{d}\omega_{\mathbf{G}} = \varphi_{\mathbf{F}}$, so that $\mathrm{d}\varphi_{\mathbf{F}} = \mathrm{d}(\mathrm{d}\omega_{\mathbf{G}}) = 0$, and thus $\mathrm{div}(\mathbf{F}) = 0$.

If U is star-shaped and $\mathrm{div}(\mathbf{F}) = 0$, then $\mathrm{d}\varphi_{\mathbf{F}} = 0$, and so $\varphi_{\mathbf{F}}$ is closed. According to Poincaré's lemma, $\varphi_{\mathbf{F}}$ is exact, which is to say that $\exists \omega \in \Omega^1(U)$ such that $\mathrm{d}\omega = \varphi_{\mathbf{F}}$. If $\mathbf{G}$ is the vector field corresponding to ω, then we have $\mathrm{curl}(\mathbf{G}) = \mathbf{F}$. ■

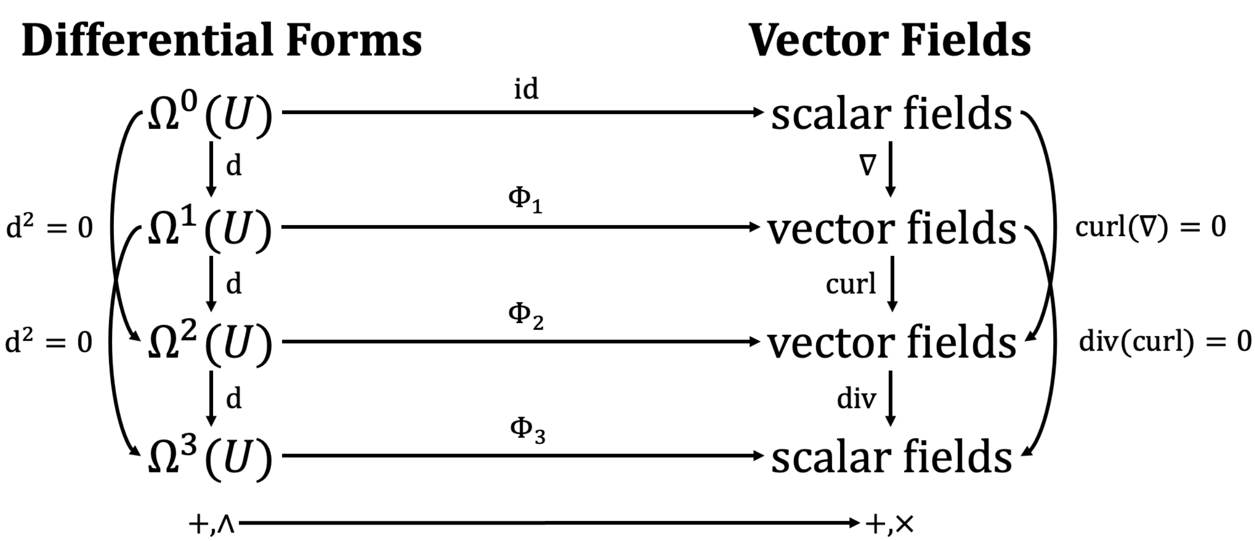
When $\mathbf{F} = \mathrm{curl}(\mathbf{G})$ for some $\mathbf{G} : U \to GR^3$, the vector field $\mathbf{G}$ is a **vector potential** for $\mathbf{F}$. Such a vector potential is not unique; indeed if $f : U \to \mathbb{R}$ is smooth, then $\mathrm{curl}(\mathbf{G} + \nabla f) = \mathrm{curl}(\mathbf{G})$, as we can see below: if

$$\mathbf{G} \leftrightsquigarrow \omega_{\mathbf{G}} \in \Omega^1(U), \quad \mathrm{curl}(\mathbf{G}) \leftrightsquigarrow \mathrm{d}\omega_{\mathbf{G}} \in \Omega^2(U), \quad \nabla f \leftrightsquigarrow \mathrm{d}f \in \Omega^1(U),$$

then

$$\mathrm{curl}(\mathbf{G} + \nabla f) \leftrightsquigarrow \mathrm{d}(\omega_{\mathbf{G}} + \mathrm{d}f) = \mathrm{d}\omega_{\mathbf{G}} \leftrightsquigarrow \mathrm{curl}(\mathbf{G}).$$

In short, differential forms provide a tool to work with vector fields, which are the objects of interests in applications; the correspondence is diagrammed below.



13.6 Solved Problems

1. Are the following $1-$forms exact?

   a) $\omega = 2xy\,\mathrm{d}x + x^2\,\mathrm{d}y$ b) $\omega = (x^2+yz)\,\mathrm{d}x + (xz+\cos y)\,\mathrm{d}y + (z+xy)\,\mathrm{d}z$
   c) $\omega = y\,\mathrm{d}x + z\,\mathrm{d}y + x\,\mathrm{d}z$ d) $\omega = \frac{x}{x^2+y^2}\,\mathrm{d}x + \frac{y}{x^2+y^2}\,\mathrm{d}y$

   Solution:

   a) We have $\omega = 2xy\,\mathrm{d}x + x^2\,\mathrm{d}y \in \Omega^1(\mathbb{R}^2)$, where $\mathbb{R}^2$ is star-shaped. Since

   $$\mathrm{d}\omega = 2\big[(\mathrm{d}x)y + x(\mathrm{d}y)\big] \wedge \mathrm{d}x + (2x\,\mathrm{d}x) \wedge \mathrm{d}y = 2x\big[\,\mathrm{d}y \wedge \mathrm{d}x + \mathrm{d}x \wedge \mathrm{d}y\big] = 0,$$

   ω is closed. According to Poincaré's lemma, ω is also exact. In fact, $\eta = x^2y$ is an anti-derivative of ω (i.e. $\mathrm{d}\eta = \omega$).

   b) We have $\omega = (x^2+yz)\,\mathrm{d}x + (xz+\cos y)\,\mathrm{d}y + (z+xy)\,\mathrm{d}z \in \Omega^1(\mathbb{R}^3)$, where $\mathbb{R}^3$ is star-shaped. Since

   $$\mathrm{d}\omega = z\,\mathrm{d}y \wedge \mathrm{d}x + y\,\mathrm{d}z \wedge \mathrm{d}x + x\,\mathrm{d}z \wedge \mathrm{d}y + z\,\mathrm{d}x \wedge \mathrm{d}y + x\,\mathrm{d}y \wedge \mathrm{d}z + y\,\mathrm{d}x \wedge \mathrm{d}z = 0,$$

   ω is closed. According to Poincaré's lemma, ω is also exact. In fact,

   $$\eta = \frac{x^3}{3} + xyz + \sin y + \frac{z^2}{2}$$

   is an anti-derivative of ω (i.e. $\mathrm{d}\eta = \omega$).

   c) Since $\mathrm{d}\omega = \mathrm{d}y \wedge \mathrm{d}x + \mathrm{d}z \wedge \mathrm{d}y + \mathrm{d}x \wedge \mathrm{d}z \neq 0$, ω is not closed. Consequently, ω is not exact (remember, this has nothing to do with Poincaré's lemma).

   d) We have $\omega = \frac{x}{x^2+y^2}\,\mathrm{d}x + \frac{y}{x^2+y^2}\,\mathrm{d}y \in \Omega^1(\mathbb{R}^2 - \{(0,0)\})$. Note that $U = \mathbb{R}^2 - \{(0,0)\}$ is NOT star-shaped, and so we cannot use Poincaré's lemma to determine whether ω is exact or not. If ω is not closed, then it will necessarily not be exact, by contraposition. However,

   $$\mathrm{d}\omega = \frac{-2xy}{(x^2+y^2)^2}\,\mathrm{d}y \wedge \mathrm{d}x - \frac{2xy}{(x^2+y^2)^2}\,\mathrm{d}x \wedge \mathrm{d}y = 0,$$

   and so ω is closed and we cannot use this approach. We are left with no other option than to try to find an anti-derivative. The brute force method yields $\eta = \ln(\sqrt{x^2+y^2})$ as an anti-derivative of ω (i.e. $\mathrm{d}\eta = \omega$). $\square$

2. Are the following $2-$forms exact?

   a) $\omega = \mathrm{d}x \wedge \mathrm{d}y$
   b) $\omega = z\,\mathrm{d}x \wedge \mathrm{d}y + y\,\mathrm{d}x \wedge \mathrm{d}z + z\,\mathrm{d}y \wedge \mathrm{d}z$

   Solution:

   a) We have $\omega = \mathrm{d}x \wedge \mathrm{d}y \in \Omega^2(\mathbb{R}^2)$, where $\mathbb{R}^2$ is star-shaped. Since

   $$\mathrm{d}\omega = d(\mathrm{d}x \wedge \mathrm{d}y) = d^2x \wedge \mathrm{d}y - \mathrm{d}x \wedge \mathrm{d}y^2 = 0 - 0 = 0,$$

   ω is closed. According to Poincaré's lemma, ω is also exact.

b) We have

$$\omega = z\,\mathrm{d}x \wedge \mathrm{d}y + y\,\mathrm{d}x \wedge \mathrm{d}z + z\,\mathrm{d}y \wedge \mathrm{d}z \in \Omega^2(\mathbb{R}^3),$$

where $\mathbb{R}^3$ is star-shaped. Since

$$\begin{aligned}\mathrm{d}\omega &= \mathrm{d}z \wedge \mathrm{d}x \wedge \mathrm{d}y + \mathrm{d}y \wedge \mathrm{d}x \wedge \mathrm{d}z + \mathrm{d}z \wedge \mathrm{d}y \wedge \mathrm{d}z \\ &= \mathrm{d}z \wedge \mathrm{d}x \wedge \mathrm{d}y - \mathrm{d}z \wedge \mathrm{d}x \wedge \mathrm{d}y + 0 = 0,\end{aligned}$$

ω is closed. According to Poincaré's lemma, ω is also exact. In fact, $\eta = xz\,\mathrm{d}y + xy\,\mathrm{d}z$ is an anti-derivative of ω (i.e. $\mathrm{d}\eta = \omega$). $\square$

13.7 Exercises

1. Prepare a 2-page summary of this chapter, with important definitions and results.
2. Prove results 172, 173, 174, 175, 176, 180 (try, at least), and 184.
3. If $f \in \Omega^0(U)$ and $\omega \in \Omega^p(U)$, show that $f \wedge \omega = f\omega$.
4. Show that if ω and φ are two closed differential forms, then so is $\omega \wedge \varphi$. Show that if ω is also exact, then $\omega \wedge \varphi$ is exact.
5. What is the pullback $\mathbf{g}^*(\omega) \in \Omega^1(\mathbb{R})$ if $\mathbf{g} : \mathbb{R} \to \mathbb{R}^2$ is defined by $\mathbf{g}(v) = (3\cos 2v, 3\sin 2v)$ and $\omega = -y\,\mathrm{d}x + x\,\mathrm{d}y \in \Omega^1(\mathbb{R}^2)$? Simplify your answer as much as possible.
6. What is the pullback $\mathbf{g}^*(\omega) \in \Omega^1(\mathbb{R}^2)$ if $g : \mathbb{R}^2 \to \mathbb{R}^3$ is defined by $\mathbf{g}(u, v) = (\cos u, \sin u, v)$ and $\omega = z\,\mathrm{d}x + x\,\mathrm{d}y + y\,\mathrm{d}z \in \Omega^1(\mathbb{R}^3)$? Simplify your answer as much as possible.
7. What is the pullback $\mathbf{g}^*(\omega) \in \Omega^2(\mathbb{R}^2)$ if $g : \mathbb{R}^2 \to \mathbb{R}^3$ is defined by $\mathbf{g}(u, v) = (\cos u, \sin u, v)$ and $\omega = z\,\mathrm{d}x \wedge \mathrm{d}y + y\,\mathrm{d}z \wedge \mathrm{d}x \in \Omega^2(\mathbb{R}^3)$? Simplify your answer as much as possible.
8. For each of the three previous exercises, compute $\mathbf{g}^*(\mathrm{d}\omega)$ and $d(\mathbf{g}^*\omega)$.
9. Let $\mathbf{g} : (0, \infty) \times (0, \pi) \times (0, 2\pi) \to \mathbb{R}^3$ the map defining the spherical coordinates in $\mathbb{R}^3$. Compute $g^*(\mathrm{d}x \wedge \mathrm{d}y \wedge \mathrm{d}z)$.
10. Let $\mathbf{F}, \mathbf{G} : \mathbb{R}^3 \to \mathbb{R}^3$, $f : \mathbb{R}^3 \to \mathbb{R}$ be smooth mappings and $\cdot$ and $\times$ represent the inner product and cross product in $\mathbb{R}^3$, respectively. Show that
    a) $\mathrm{div}(\mathbf{F} + \mathbf{G}) = \mathrm{div}(\mathbf{F}) + \mathrm{div}(\mathbf{G})$
    b) $\mathrm{div}(f\mathbf{F}) = f\mathrm{div}(\mathbf{F}) + \mathbf{F} \cdot \nabla f$
    c) $\mathrm{div}(\mathbf{F} \times \mathbf{G}) = \mathbf{G} \cdot \mathrm{curl}\,\mathbf{F} - \mathbf{F} \cdot \mathrm{curl}\,\mathbf{G}$
    d) $\mathrm{curl}(f\mathbf{F}) = f\,\mathrm{curl}(\mathbf{F}) + (\nabla f) \times \mathbf{F}$
    e) $\mathrm{div}(f\nabla f) = |\nabla f|^2$
11. Let $U \subseteq_O \mathbb{R}^n$ and $p \geq 0$. Show that $\Omega^p(U)$ is a vector space over $\mathbb{R}$.
12. Let $U \subseteq_O \mathbb{R}^n$, $p \geq 0$ and $\omega_1, \omega_2 \in \Omega^p(U)$. Show that $d(\omega_1 + \omega_2) = \mathrm{d}\omega_1 + \mathrm{d}\omega_2$.

Chapter 14

Integrating Differential Forms

The integral of a differential form generalizes the concept of the integral of a function of a single variable (see Chapter 21 for another). In this chapter, we formalize the concepts of the **line**, **surface**, and **flux integral**, and present **Stokes' Theorem**, a deep unifying result of vector analysis.

14.1 Line Integral of a Differential $1-$Form

Let $U \subseteq_O \mathbb{R}^n$. Assume that $\boldsymbol{\gamma}$ is a differentiable path in U and that $\omega \in \Omega^1(U)$. This section's objective is to define $\int_{\boldsymbol{\gamma}} \omega$ **meaningfully**. A **path** in U is a continuous function $\boldsymbol{\gamma} : [a, b] \to U$; $\boldsymbol{\gamma}(a)$ is the **starting point** while $\boldsymbol{\gamma}(b)$ is the path's **finishing point**.

Examples

1. Let $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$. The path $\boldsymbol{\gamma} : [0, 1] \to \mathbb{R}^n$ defined by $\boldsymbol{\gamma}(t) = t\mathbf{v} + (1 - t)\mathbf{u}$ is the (oriented) line segment joining $\mathbf{u}$ and $\mathbf{v}$. $\square$

2. Let $\boldsymbol{\gamma} : [0, 2\pi] \to \mathbb{R}^2$ be defined by $\boldsymbol{\gamma}(t) = (\cos t, \sin t)$. Then $\boldsymbol{\gamma}([0, 2\pi])$ is the **unit circle** in $\mathbb{R}^2$, starting at $\boldsymbol{\gamma}(0) = (1, 0)$ and ending at $\boldsymbol{\gamma}(2\pi) = (1, 0)$, travelling counter-clockwise. $\square$

In that last example, $\boldsymbol{\gamma}$ is a **closed**, **simple** curve, which is to say that

$$\boldsymbol{\gamma}(0) = \boldsymbol{\gamma}(2\pi) \quad \text{and} \quad \boldsymbol{\gamma}(t) \neq \boldsymbol{\gamma}(s) \text{ for all } t \neq s \in (0, 2\pi).$$

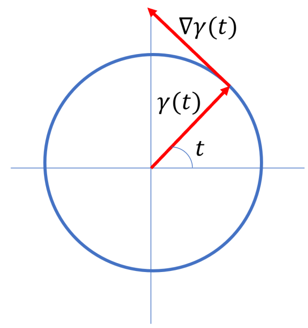
A path $\boldsymbol{\gamma}$ is **continuously differentiable** (denoted $\mathcal{C}^1$) if its derivative $\boldsymbol{\gamma}' : [a, b] \to \text{End}(\mathbb{R}, \mathbb{R}^n)$ varies continuously with t; the derivative is one-sided at the endpoints a and b. In that case,

$$\boldsymbol{\gamma}'(t) : \mathbb{R} \to \mathbb{R}^n, \quad x \mapsto \boldsymbol{\gamma}'(t)x = \nabla\boldsymbol{\gamma}(t)x = (\gamma_1'(t), \dots, \gamma_n'(t))x.$$

Examples

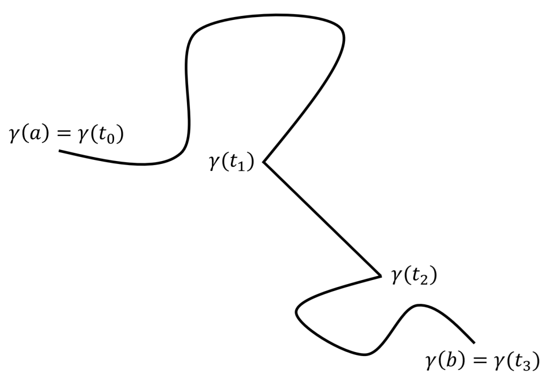
1. In the line segment example above, $\boldsymbol{\gamma}'(t) = \mathbf{v} - \mathbf{u} \in \mathbb{R}^n$. $\square$

2. In the circle example above, $\nabla\boldsymbol{\gamma}(t) = (-\sin t, \cos t)$. Note that $\boldsymbol{\gamma}(t) \perp \nabla\boldsymbol{\gamma}(t)$ for all t. $\square$



If $\boldsymbol{\gamma} : [a, b] \to U \subseteq_O \mathbb{R}^n$ represents the **position** of a particle at time t, then $\boldsymbol{\gamma}'(t)$ represents the **velocity vector** of the particle at time t; $\boldsymbol{\gamma}'(t_0)$ is necessarily tangent to the path $\boldsymbol{\gamma}$ at $t = t_0$.

A path $\boldsymbol{\gamma}$ is **piecewise differentiable** if $a = t_0 < t_1 < \cdots < t_n = b$ and $\boldsymbol{\gamma}|_{[t_i, t_{i+1}]}$ is $\mathcal{C}^1$ for all i.



Now we come to the section's important definition. Let $\boldsymbol{\gamma}$ be a $\mathcal{C}^1$ path in $U \subseteq_O \mathbb{R}^n$ and

$$\omega = \sum_{i=1}^{n} P_i(\mathbf{x})\, \mathrm{d}x_i \in \Omega^1(U).$$

The **line integral of** ω **along** $\boldsymbol{\gamma}$ is given by

$$\int_{\boldsymbol{\gamma}} \omega = \int_{\boldsymbol{\gamma}} \sum_{i=1}^{n} P_i(\mathbf{x})\, \mathrm{d}x_i := \int_{[a,b]} \sum_{i=1}^{n} P_i(\boldsymbol{\gamma}(t))\gamma_i(t)\, \mathrm{d}t,$$

where $\boldsymbol{\gamma}(t) = (\gamma_1(t), \ldots, \gamma_n(t))$, $\boldsymbol{\gamma}'(t) = (\gamma_1'(t), \ldots, \gamma_n'(t))$, and $\boldsymbol{\gamma} : [a, b] \to U \subseteq_O \mathbb{R}^n$.

Example: if $\boldsymbol{\gamma} : [-1,1] \to \mathbb{R}^2$ is $\boldsymbol{\gamma}(t) = (t, t^2)$ and $\omega = -y\,\mathrm{d}x + x\,\mathrm{d}y \in \Omega^1(\mathbb{R}^2)$, then $\boldsymbol{\gamma}'(t) = (1, 2t)$, $P_1(x,y) = -y$, $P_2(x,y) = x$, and

$$\begin{aligned}\int_{\boldsymbol{\gamma}} \omega &= \int_{-1}^{1} \left(P_1(\boldsymbol{\gamma}(t))\gamma_1'(t) + P_2(\boldsymbol{\gamma}(t))\gamma_2'(t)\right)\,\mathrm{d}t \\ &= \int_{-1}^{1} (P_1(t,t^2)(1) + P_2(t,t^2)(2t))\,\mathrm{d}t \\ &= \int_{-1}^{1} (-t^2 + t(2t))\,\mathrm{d}t = \int_{-1}^{1} t^2\,\mathrm{d}t = \left[\frac{t^3}{3}\right]_{-1}^{1} = \frac{2}{3},\end{aligned}$$

using the regular rules of integration. □

But we could also approach the problem from a different (but ultimately equivalent) angle: the pullback of ω by $\boldsymbol{\gamma}$ is

$$\boldsymbol{\gamma}^*(\omega) = \boldsymbol{\gamma}^*(-y\,\mathrm{d}x + x\,\mathrm{d}y) = P_1(\boldsymbol{\gamma}(t))\frac{\partial \gamma_1}{\partial t}\,\mathrm{d}t + P_2(\boldsymbol{\gamma}(t))\frac{\partial \gamma_1}{\partial t}\,\mathrm{d}t = (-\gamma_2(t)\gamma_1'(t) + \gamma_1(t)\gamma_2'(t))\,\mathrm{d}t \in \Omega^1(\mathbb{R}),$$

so that $\int_{\boldsymbol{\gamma}} \omega = \int_{-1}^{1} \boldsymbol{\gamma}^*(\omega)$.

In general, if $\gamma : [a,b] \to U \subseteq_O \mathbb{R}^n$ and $\omega = \sum_{i=1}^n P_i\,\mathrm{d}x_i \in \Omega^1(U)$, then

$$\int_{\boldsymbol{\gamma}} \omega = \int_{[a,b]} \boldsymbol{\gamma}^*(\omega) = \int_{[a,b]} \sum_{i=1}^n P_i(\boldsymbol{\gamma}(t))\,\mathrm{d}\gamma_i = \int_{[a,b]} \sum_{i=1}^n P_i(\boldsymbol{\gamma}(t))\gamma_i(t)\,\mathrm{d}t.$$

Example: consider $\omega = -y\,\mathrm{d}x + x\,\mathrm{d}y \in \Omega^1(\mathbb{R}^2)$ and two paths from $(1,0)$ to $(0,1)$, $\boldsymbol{\gamma} : [0,\pi/2] \to \mathbb{R}^2$ (a circle arc) and $\boldsymbol{\eta} : [0,1] \to \mathbb{R}^2$ (a line segment), defined by $\boldsymbol{\gamma}(t) = (\cos t, \sin t)$ and $\boldsymbol{\eta}(t) = (1-t, t)$. Then

$$\begin{aligned}\int_{\boldsymbol{\gamma}} \omega &= \int_0^{\pi/2} \boldsymbol{\gamma}^*(\omega) = \int_0^{\pi/2} \left[(-\sin t)(\sin t) + (\cos t)(\cos t)\right]\,\mathrm{d}t = \int_0^{\pi/2} 1\,\mathrm{d}t = [t]_0^{\pi/2} = \frac{\pi}{2}, \\ \int_{\boldsymbol{\eta}} \omega &= \int_0^1 \boldsymbol{\eta}^*(\omega) = \int_0^1 \left[(-t)(-1) + (1-t)(1)\right]\,\mathrm{d}t = \int_0^1 1\,\mathrm{d}t = [t]_0^1 = 1.\end{aligned}$$

Evidently, the value of the line integral depends on the path and the endpoints. □

If $\mathbf{P} : U \to \mathbb{R}^n$ is the vector field corresponding to $\omega \in \Omega^1(U)$, then

$$\sum_{i=1}^n P_i(\boldsymbol{\gamma}(t))\gamma_i'(t) = \mathbf{P}(\boldsymbol{\gamma}(t)) \cdot \boldsymbol{\gamma}'(t) = \left(\mathbf{P}(\boldsymbol{\gamma}(t)) \mid \boldsymbol{\gamma}'(t)\right),$$

we sometimes write

$$\int_{\boldsymbol{\gamma}} \omega = \int_{[a,b]} \mathbf{P}(\boldsymbol{\gamma}(t)) \cdot \boldsymbol{\gamma}'(t)\,\mathrm{d}t = \int_{\boldsymbol{\gamma}} \mathbf{P} \cdot \mathrm{d}\mathbf{r},$$

where $\mathbf{r}$ is a parameterization of γ (i.e., $\mathrm{d}\mathbf{r}(t) = \boldsymbol{\gamma}'(t)\mathrm{d}t$).

Let $\varphi : [a', b'] \to [a, b]$ be a $\mathcal{C}^1$ **diffeomorphism**;[1] this entails that $\varphi'(t) \neq 0$ for all $t \in [a', b']$. Since φ' is continuous, there are 2 possibilities:

1. $\varphi'(t) > 0 \Longrightarrow \varphi(a') = a$ and $\varphi(b') = b$, in which case φ **preserves the orientation**;

2. $\varphi'(t) < 0 \Longrightarrow \varphi(a') = b$ and $\varphi(b') = a$, in which case φ **reverses the orientation**.

> **Examples:** $\varphi : [1, 2] \to [1, 4]$ defined by $\varphi(t) = t^2$ preserves the orientation as $\varphi'(t) = 2t > 0$ on $[1, 2]$; but $\varphi : [-2, -1] \to [1, 4]$ defined by $\varphi(t) = t^2$ reverses the orientation as $\varphi'(t) = 2t < 0$ on $[-2, -1]$. □

The distinction comes in at the following level.

> **Proposition 188**
> *Let $\omega = \sum_{i=1}^n P_i(\mathbf{x})\, dx_i \in \Omega^1(U)$, $\boldsymbol{\gamma} : [a, b] \to U$, $\boldsymbol{\gamma} \in \mathcal{C}^1$. If $\varphi : [a', b'] \to [a, b]$ is a $\mathcal{C}^1$ diffeomorphism, then*
>
> 1. $\int_{\boldsymbol{\gamma}\circ\varphi} \omega = \int_{\boldsymbol{\gamma}} \omega$ *if φ is orientation-preserving;*
> 2. $\int_{\boldsymbol{\gamma}\circ\varphi} \omega = -\int_{\boldsymbol{\gamma}} \omega$ *if φ is orientation-reversing.*
>
> **Proof:**
>
> 1. By construction, $\boldsymbol{\gamma} \circ \varphi : [a', b'] \to U$ is a $\mathcal{C}^1$ path and $\boldsymbol{\gamma}'(\varphi(t))\varphi'(t)$ exists for all $t \in [a', b']$. If we write $t = \varphi(s)$, then $\mathrm{d}t = \varphi'(s)\,\mathrm{d}s$, $a = \varphi(a')$, and $b = \varphi(b')$, and so
> $$\begin{aligned}\int_{\boldsymbol{\gamma}} \omega &= \int_{t=a}^{t=b} \sum_{i=1}^{n} (P_i \circ \boldsymbol{\gamma}(t))\gamma_i'(t)\,\mathrm{d}t = \int_{s=a'}^{s=b'} \sum_{i=1}^{n} (P_i \circ \boldsymbol{\gamma}(\varphi(s)))\gamma_i'(\varphi(s))\varphi'(s)\,\mathrm{d}s \\ &= \int_{a'}^{b'} \sum_{i=1}^{n} \left[P_i \circ (\boldsymbol{\gamma} \circ \varphi)(s)\right](\boldsymbol{\gamma} \circ \varphi)_i'(s)\,\mathrm{d}s = \int_{\boldsymbol{\gamma}\circ\varphi} \omega.\end{aligned}$$
> 2. The proof is similar, except that the change of variable is $t = \varphi(s)$, then $\mathrm{d}t = \varphi'(s)\,\mathrm{d}s$, $a = \varphi(b')$, and $b = \varphi(a')$, and so
> $$\int_{\boldsymbol{\gamma}} \omega = \int_{s=b'}^{s=a'} \sum_{i=1}^{n} (P_i \circ \boldsymbol{\gamma}(\varphi(s)))\gamma_i'(\varphi(s))\varphi'(s)\,\mathrm{d}s = -\int_{s=a'}^{s=b'} \cdots = -\int_{\boldsymbol{\gamma}\circ\varphi} \omega. \quad \blacksquare$$

The line integral has two properties that are the counterparts of Theorems 55.1 and 56.

[1] That is, both φ and φ^{-1} are $\mathcal{C}^1$.

Proposition 189
Let $U \subseteq_O \mathbb{R}^n$, $\omega, \omega_1, \omega_2 \in \Omega^1(U)$, and $\boldsymbol{\gamma}, \boldsymbol{\eta}$ be $\mathcal{C}^1$ paths in U such that the finishing point of $\boldsymbol{\gamma}$ is the starting point of $\boldsymbol{\eta}$. The concatenation $\boldsymbol{\gamma}+\boldsymbol{\eta}$ is piecewise $\mathcal{C}^1$. Then:

1. *the line integral is linear in the sum (concatenation) of paths:*
$$\int_{\boldsymbol{\gamma}+\boldsymbol{\eta}} \omega = \int_{\boldsymbol{\gamma}} \omega + \int_{\boldsymbol{\eta}} \omega;$$
2. *the line integral is linear in the sum of differential forms:*
$$\int_{\boldsymbol{\gamma}} (\omega_1+\omega_2) = \int_{\boldsymbol{\gamma}} \omega_1 + \int_{\boldsymbol{\gamma}} \omega_2.$$

Proof: left as an exercise. ■

Proposition 189, together with the next property, justifies the naming of the line integral: if it looks like an integral and it behaves like an integral...

Theorem 190 (Fundamental Theorem of Line Integrals)
Let $\boldsymbol{\gamma} : [a,b] \to \mathbb{R}$ be a piecewise $\mathcal{C}^1$ path and $\omega = df \in \Omega^1(U)$ for some vector field $f \in \mathcal{C}^\infty(U,\mathbb{R})$. Then
$$\int_{\boldsymbol{\gamma}} \omega = \int_{\boldsymbol{\gamma}} df = f(\boldsymbol{\gamma}(b)) - f(\boldsymbol{\gamma}(a)).$$

Proof: according to Proposition 189.1, it is sufficient to show the result for $\mathcal{C}^1$ paths $\boldsymbol{\gamma}$; according to Proposition 184, we know that $\mathrm{d}(\boldsymbol{\gamma}^*(f)) = \boldsymbol{\gamma}^*(\mathrm{d}f)$. Then
$$\begin{aligned}\int_{\boldsymbol{\gamma}} \omega &= \int_{\boldsymbol{\gamma}} \mathrm{d}f = \int_{[a,b]} \gamma^*(f)(\mathrm{d}f) = \int_a^b \mathrm{d}(\boldsymbol{\gamma}^*(f)) = \int_a^b \mathrm{d}(f\circ\boldsymbol{\gamma}) = \int_a^b (f\circ\boldsymbol{\gamma})'(t)\,\mathrm{d}t \\ &= [f\circ\boldsymbol{\gamma}(t)]_a^b = f(\boldsymbol{\gamma}(b)) - f(\boldsymbol{\gamma}(a)),\end{aligned}$$
which completes the proof. ■

In the example on page 335, we have $\int_{\boldsymbol{\gamma}} -y\,\mathrm{d}x + x\,\mathrm{d}x \neq \int_{\boldsymbol{\eta}} -y\,\mathrm{d}x + x\,\mathrm{d}x$, even though $\boldsymbol{\gamma}$ and $\boldsymbol{\eta}$ have the same starting points and finishing points, and so Theorem 190 does not apply. What is the problem?

Corollary 191
If $\omega = dg \in \Omega^1(U)$ and $\boldsymbol{\gamma}$ is a $\mathcal{C}^1$ path in U, then $\int_{\boldsymbol{\gamma}} \omega = \int_{\boldsymbol{\gamma}} dg$ depends only on the endpoints of $\boldsymbol{\gamma}$. **Proof:** immediately follows from Theorem 190. ■

An open subset $U \subseteq_O \mathbb{R}^n$ is **path-connected** if for all $\mathbf{u}, \mathbf{v} \in U$, there is a path $\boldsymbol{\gamma} : [a,b] \to U$ such that $\boldsymbol{\gamma}(a) = \mathbf{u}$ and $\boldsymbol{\gamma}(b) = \mathbf{v}$; open balls and open annulii/torii are path-connected in $\mathbb{R}^2/\mathbb{R}^3$, but a set made up of disjoint open balls isn't.

A **loop** $\boldsymbol{\gamma}$ is a path $\boldsymbol{\gamma} : [a,b] \to U$ for which $\boldsymbol{\gamma}(a) = \boldsymbol{\gamma}(b)$; the path $\boldsymbol{\gamma} : [0, 2\pi] \to \mathbb{R}^2 \simeq \mathbb{C}$ defined by $\boldsymbol{\gamma}(t) = (\cos t, \sin t) \simeq e^{it}$ is a loop.

Theorem 192
Let $U \subseteq_O \mathbb{R}^n$ be path-connected. For a continuous differential form $\omega \in \Omega^1(U)$, the following are equivalent:

1. *ω is exact in U;*
2. *$\int_{\boldsymbol{\gamma}} \omega = 0$ for any loop $\boldsymbol{\gamma} : [a,b] \to U$;*
3. *if $\boldsymbol{\gamma}$ is any path in U, $\int_{\boldsymbol{\gamma}} \omega$ only depends on the endpoints of $\boldsymbol{\gamma}$.*

Proof: $\boxed{1. \Longrightarrow 2.}$ follows from Theorem 190 since $\boldsymbol{\gamma}(a) = \boldsymbol{\gamma}(b)$ for any loop $\boldsymbol{\gamma} : [a,b] \to U$.

For $\boxed{2. \Longrightarrow 3.}$, let $\boldsymbol{\gamma}, \boldsymbol{\eta}$ be two paths in U with the same endopoints. Then $\boldsymbol{\gamma} - \boldsymbol{\eta}$ is a loop in U, and

$$0 = \int_{\boldsymbol{\gamma}-\boldsymbol{\eta}} \omega = \int_{\boldsymbol{\gamma}} \omega + \int_{-\boldsymbol{\eta}} \omega = \int_{\boldsymbol{\gamma}} \omega - \int_{\boldsymbol{\eta}} \omega \Longrightarrow \int_{\boldsymbol{\gamma}} \omega = \int_{\boldsymbol{\eta}} \omega.$$

For $\boxed{3. \Longrightarrow 1.}$, let $\mathbf{x}_0 \in U$ be fixed. For any $\mathbf{x} \in U$, let $\boldsymbol{\gamma}_{\mathbf{x}}$ be a path in U from $\mathbf{x}_0$ to $\mathbf{x}$. Define $f : U \to \mathbb{R}$ by $f(\mathbf{x}) = \int_{\boldsymbol{\gamma}_{\mathbf{x}}} \omega$. By assumption, if $\tilde{\boldsymbol{\gamma}}_{\mathbf{x}}$ is any other path in U from $\mathbf{x}_0$ to $\mathbf{x}$, then $\boldsymbol{\gamma}_{\mathbf{x}} - \tilde{\boldsymbol{\gamma}}_x$ is a loop in U and

$$0 = \int_{\boldsymbol{\gamma}_{\mathbf{x}} - \tilde{\boldsymbol{\gamma}}_{\mathbf{x}}} \omega = \int_{\boldsymbol{\gamma}_{\mathbf{x}}} \omega - \int_{\tilde{\boldsymbol{\gamma}}_{\mathbf{x}}} \omega \Longrightarrow f(\mathbf{x}) = \int_{\boldsymbol{\gamma}_{\mathbf{x}}} \omega = \int_{\tilde{\boldsymbol{\gamma}}_{\mathbf{x}}} \omega,$$

no matter which path $\boldsymbol{\gamma}_{\mathbf{x}}$ we use. Hence, f is well-defined.

It remains to see that $\mathrm{d}f = \omega$. Since

$$\mathrm{d}f = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}\, \mathrm{d}x_i \quad \text{and} \quad \omega = \sum_{i=1}^{n} P_i\, \mathrm{d}x_i,$$

we need to show that $\frac{\partial f}{\partial x_i} = P_i$, $1 \leq i \leq n$. We know that

$$\frac{\partial f}{\partial x_i} = \lim_{t \to 0} \frac{f(\mathbf{x} + t\mathbf{e}_i) - f(\mathbf{x})}{t},$$

for $1 \leq i \leq n$ if the limit exists. Since U is open, $\mathbf{x} + t\mathbf{e}_i \in U$ for all i if t is small enough.

For each i, we have

$$\frac{1}{t}\left(f(\mathbf{x}+t\mathbf{e}_i)-f(\mathbf{x})\right)=\frac{1}{t}\left[\int_{\boldsymbol{\gamma}_{\mathbf{x}+t\mathbf{e}_i}}\omega-\int_{\boldsymbol{\gamma}_{\mathbf{x}}}\omega\right]=\frac{1}{t}\int_{\boldsymbol{\gamma}_i^t}\omega,$$

where $\boldsymbol{\gamma}_i^t$ is the straight line path from $\mathbf{x}$ to $\mathbf{x}+t\mathbf{e}_i$ (which is possible, again, if t is small enough), that is $\boldsymbol{\gamma}_i^t:[0,1]\to U$ defined by

$$\boldsymbol{\gamma}_i^t(s)=s(\mathbf{x}+t\mathbf{e}_i)+(1-s)\mathbf{x}=\mathbf{x}+st\mathbf{e}_i;$$

then $(\boldsymbol{\gamma}_i^t)'(s)=t\mathbf{e}_i$. In particular, for $1\leq j\leq n$ we have

$$\mathrm{d}(\boldsymbol{\gamma}_i^t)_j=\sum_{j=1}^n\frac{\partial(\boldsymbol{\gamma}_i^t)_j}{\partial s}\,\mathrm{d}s=\begin{cases}0 & \text{if } i\neq j\\ t\,\mathrm{d}s & \text{if } i=j\end{cases}$$

so that the pullback of ω by $\boldsymbol{\gamma}_i^t$ is

$$(\boldsymbol{\gamma}_i^t)^*(\omega)=\sum_{j=1}^n\left(P_j\circ\boldsymbol{\gamma}_i^t\right)\,\mathrm{d}(\boldsymbol{\gamma}_i^t)_j$$

and so

$$\begin{aligned}\frac{1}{t}\int_{\boldsymbol{\gamma}_i^t}\omega&=\frac{1}{t}\int_0^1(\boldsymbol{\gamma}_i^t)^*(\omega)=\frac{1}{t}\int_0^1\sum_{j=1}^nP_j\circ\boldsymbol{\gamma}_i^t(s)\,\mathrm{d}(\boldsymbol{\gamma}_i^t)_j=\frac{1}{t}\int_0^1P_i(\boldsymbol{\gamma}_i^t(s))t\,\mathrm{d}s\\&=\int_0^1P_i(\mathbf{x}+st\mathbf{e}_i)\,\mathrm{d}s=\int_0^1\left(P_i(\mathbf{x})+P_i(\mathbf{x}+st\mathbf{e}_i)-P_i(\mathbf{x})\right)\mathrm{d}s\\&=P_i(\mathbf{x})+\int_0^1\left(P_i(\mathbf{x}+st\mathbf{e}_i)-P_i(\mathbf{x})\right)\mathrm{d}s.\end{aligned}$$

Hence,

$$\begin{aligned}\frac{\partial f}{\partial x_i}&=\lim_{t\to 0}\left[P_i(\mathbf{x}+\int_0^1\left(P_i(\mathbf{x}+st\mathbf{e}_i)-P_i(\mathbf{x})\right)\mathrm{d}s\right]\\&=P_i(\mathbf{x})+\int_0^1\underbrace{\lim_{t\to 0}\Big(P_i(\mathbf{x}+st\mathbf{e}_i)-P_i(\mathbf{x})\Big)}_{=0\text{ since }\omega\text{ is }\mathcal{C}^0}\,\mathrm{d}s=P_i(\mathbf{x}),\end{aligned}$$

which completes the proof. ■

We extract a specific implication from this result, for future ease of access.

Corollary 193
With the same hypotheses as in Theorem 192, if $\int_{\gamma}\omega=0$ for any loop γ in U, then ω is exact.

Finally, we show how to build an antiderivative for $\omega \in \Omega^1(U)$.

Example: consider the differential form

$$\omega = P_1(x,y)\,\mathrm{d}x + P_2(x,y)\,\mathrm{d}y = (e^x + 2xy)\,\mathrm{d}x + (x^2 + \cos y)\,\mathrm{d}y \in \Omega^1(\mathbb{R}^2).$$

Since

$$\mathrm{d}\omega = \left(\frac{\partial P_2}{\partial x} - \frac{\partial P_1}{\partial y}\right)\,\mathrm{d}x \wedge \mathrm{d}y = (2x - 2x)\,\mathrm{d}x \wedge \mathrm{d}y = 0,$$

then ω is closed. According to Poincaré's lemma, since $\mathbb{R}^2$ is star-shaped (and thus path-connected), then ω is exact, so it has an antiderivative $f : \mathbb{R}^2 \to \mathbb{R}$. We will compute f in two ways, exploiting Theorem 192.

1. Let $\mathbf{z}_0 = (x_0, y_0) \in U$ be fixed and consider the path $\boldsymbol{\gamma} : [0,1] \to \mathbb{R}^2$ given by $\boldsymbol{\gamma}(t) = t\mathbf{z}_0$ ($\boldsymbol{\gamma}$ is the line segment joining the origin to $\mathbf{z}_0$). Then $\boldsymbol{\gamma}'(t) = \mathbf{z}_0$. Set

$$\begin{aligned} f(\mathbf{z}_0) &= \int_{\boldsymbol{\gamma}} \omega = \int_0^1 \boldsymbol{\gamma}^*(\omega) = \int_0^1 P_1(\boldsymbol{\gamma}(t))\gamma_1'(t)\,\mathrm{d}t + P_2(\boldsymbol{\gamma}(t))\gamma_2'(t)\,\mathrm{d}t \\ &= \int_0^1 P_1(tx_0, ty_0)x_0\,\mathrm{d}t + P_2(tx_0, ty_0)y_0\,\mathrm{d}t \\ &= \int_0^1 \left(e^{tx_0} + 2(tx_0)(ty_0)x_0\right)\mathrm{d}t + \int_0^1 \left((tx_0)^2 + \cos(ty_0)y_0\right)\mathrm{d}t \\ &= \left[e^{tx_0} + \frac{2}{3}t^3x_0^2y_0 + \frac{1}{3}t^3x_0^2y_0 + \sin(ty_0)\right]_0^1 = e^{x_0} + x_0^2y_0 + \sin y_0 - 1. \end{aligned}$$

2. If instead we join the origin to $\mathbf{z}_0 = (x_0, y_0)$ by first travelling horizontally to $(x_0, 0)$ along $\boldsymbol{\gamma}_1$, then travelling vertically to (x_0, y_0) along $\boldsymbol{\gamma}_2$, we have

$$\boldsymbol{\gamma}_1 : [0, x_0] \to \mathbb{R}^2, t \mapsto (t, 0), \quad \boldsymbol{\gamma}_2 : [0, y_0] \to \mathbb{R}^2, t \mapsto (x_0, t),$$

and $\boldsymbol{\gamma}_1'(t) = (1,0)$, $\boldsymbol{\gamma}_2'(t) = (0,1)$, so that

$$\begin{aligned} f(\mathbf{z}_0) &= \int_{\boldsymbol{\gamma}_1} \omega + \int_{\boldsymbol{\gamma}_2} \omega = \int_{\boldsymbol{\gamma}_1} \omega + \int_{\boldsymbol{\gamma}_2} \omega \\ &= \int_0^{x_0} e^t\,\mathrm{d}t + \int_0^{y_0} (x_0^2 + \cos t)\,\mathrm{d}t = e^{x_0} - 1 + [x_0^2 t + \sin t]_0^{y_0} \\ &= e^{x_0} - 1 + x_0^2 y_0 + \sin y_0. \end{aligned}$$

No surprise there: they're the same! □

Interpretation of the Line Integral Suppose a point particle proceeds along the path $\boldsymbol{\gamma}$ and is subjected to the effects of a vector field $\mathbf{F}$. Then the **work** done by the particle on its journey is given by $\int_{\boldsymbol{\gamma}} \Phi_1 \circ \mathbf{F} = \int_{\boldsymbol{\gamma}} \omega_{\mathbf{F}}$.

14.2 Integral of a Differential $p-$Form

Let $U \subseteq_O \mathbb{R}^n$. Given a differential form $\omega \in \Omega^1(U)$ and a $\mathcal{C}^1$ function $\boldsymbol{\gamma} : V = [a,b] \subseteq \mathbb{R}^1 \to U$, we have seen how we could define a quantity, the line integral $\int_{\boldsymbol{\gamma}} \omega$, that behaves in many ways like the Riemann integral.

If we remember that dim$(\Lambda^1(\mathbb{R}^1)) = 1$, we can define an vector space isomorphism

$$\tilde{\Phi}_1 : \mathbb{R}^1 \to \Lambda$$

by $\tilde{\Phi}_1(a) = a\,\mathrm{d}t$ and thus re-write the line integral formulation as

$$\int_{\boldsymbol{\gamma}} \omega = \int_V \boldsymbol{\gamma}^*(\omega) := \int_V \tilde{\Phi}_1^{-1}(\boldsymbol{\gamma}^*(\omega))\,\mathrm{d}m = \int_{[a,b]} \tilde{\Phi}_1^{-1}(\boldsymbol{\gamma}^*(\omega))\,\mathrm{d}m,$$

where m is the Borel-Lebesgue measure on $\mathbb{R}$ (see Chapter 21).[2]

We can generalize this definition to **differential** $p-$**forms**. Let $V \subseteq \mathbb{R}^p$ and consider a $\mathcal{C}^1$ function $\boldsymbol{\sigma} : V \to U$ and a differential form $\varphi \in \Omega^p(U) \subseteq \Omega^p(\mathbb{R}^n)$. The pullback of φ by $\boldsymbol{\sigma}$ is itself a differential form $\boldsymbol{\sigma}^*(\varphi) \in \Omega^1(V) \subseteq \Omega^p(\mathbb{R}^p)$. Since dim $\Lambda^p(\mathbb{R}^p) = 1$, we there is a vector space isomorphism

$$\tilde{\Phi}_p : \mathbb{R}^1 \to \Lambda^p(\mathbb{R}^p)$$

given by $\tilde{\Phi}_p(a) = a\,\mathrm{d}t_1 \wedge \cdots \wedge \mathrm{d}t_p$. Suppose that $\boldsymbol{\sigma}$ is **orientable** (more on this later), then we define the "surface" integral of φ on V by

$$\int_{\boldsymbol{\sigma}} \varphi = \int_V \boldsymbol{\sigma}^*(\varphi) := \int_V \tilde{\Phi}_p^{-1}(\boldsymbol{\gamma}^*(\varphi))\,\mathrm{d}m.$$

Example: consider $\boldsymbol{\sigma} : [0,1]^2 \to \mathbb{R}^3$, which is defined by $\boldsymbol{\sigma}(s,t) = (s,t,s^2+t^2)$, and $\varphi = \mathrm{d}x \wedge \mathrm{d}z - \mathrm{d}x \wedge \mathrm{d}y \in \Omega^2(\mathbb{R}^3)$. Then

$$\begin{aligned}\boldsymbol{\sigma}^*(\varphi) &= \boldsymbol{\sigma}(\mathrm{d}x \wedge \mathrm{d}z) - \boldsymbol{\sigma}^*(\mathrm{d}x \wedge \mathrm{d}y) = \mathrm{d}\sigma_1 \wedge \mathrm{d}\sigma_3 - \mathrm{d}\sigma_1 \wedge \mathrm{d}\sigma_2 \\ &= \left(\frac{\partial \sigma_1}{\partial s}\,\mathrm{d}s + \frac{\partial \sigma_1}{\partial t}\,\mathrm{d}t\right) \wedge \left(\frac{\partial \sigma_3}{\partial s}\,\mathrm{d}s + \frac{\partial \sigma_3}{\partial t}\,\mathrm{d}t\right) - \left(\frac{\partial \sigma_1}{\partial s}\,\mathrm{d}s + \frac{\partial \sigma_1}{\partial t}\,\mathrm{d}t\right) \wedge \left(\frac{\partial \sigma_2}{\partial s}\,\mathrm{d}s + \frac{\partial \sigma_2}{\partial t}\,\mathrm{d}t\right) \\ &= (1 \cdot\, \mathrm{d}s + 0 \cdot\, \mathrm{d}t) \wedge (2s\,\mathrm{d}s + 2t\,\mathrm{d}t) - (1 \cdot\, \mathrm{d}s + 0 \cdot\, \mathrm{d}t) \wedge (0 \cdot \mathrm{d}s + 1 \cdot \mathrm{d}t) = (2t-1)\,\mathrm{d}s \wedge \mathrm{d}t.\end{aligned}$$

Hence $\tilde{\Phi}_2^{-1}(\boldsymbol{\sigma}^*(\varphi)) = 2t-1$ and

$$\int_{\boldsymbol{\sigma}} \varphi = \int_{[0,1]^2} = \int_0^1 \int_0^1 (2t-1)\,\mathrm{d}s\,\mathrm{d}t = \int_0^1 (2t-1)\,\mathrm{d}t = 0,$$

assuming that the reader knows how to compute multivariate integrals. □

[2] Note that $\tilde{\Phi}_1$ and Φ_1 defined in the previous section are different functions.

We have seen in Chapter 13 that

$$\varphi = \varphi_{\mathbf{F}} = P(x,y,z)\,\mathrm{d}y \wedge \mathrm{d}z + Q(x,y,z)\,\mathrm{d}z \wedge \mathrm{d}x + R(x,y,z)\,\mathrm{d}x \wedge \mathrm{d}y \in \Omega^2(\mathbb{R}^3)$$

corresponds to the vector field $\mathbf{F} : \mathbb{R}^3 \to \mathbb{R}^3$ defined by

$$\mathbf{F}(x,y,z) = (P(x,y,z), Q(x,y,z), R(x,y,z)).$$

If we set $\mathrm{d}\mathbf{A} = (\mathrm{d}y \wedge \mathrm{d}z, \mathrm{d}z \wedge \mathrm{d}x, \mathrm{d}x \wedge \mathrm{d}y)$, then we often write

$$\int_{\sigma} \varphi = \int_{\boldsymbol{\sigma}} \mathbf{F} \cdot \mathrm{d}\mathbf{A} = \iint_S \mathbf{F} \cdot \mathrm{d}\mathbf{A},$$

where $S = \boldsymbol{\sigma}(V) = \{\boldsymbol{\sigma}(s,t) \mid (s,t) \in V\}$ is orientable. In that case, the **surface integral** (also known as the **flux integral**) of φ over $\boldsymbol{\sigma}$ is

$$\int_{\boldsymbol{\sigma}} \varphi = \int_V \boldsymbol{\sigma}^*(\varphi) = \pm \int_V \mathbf{F}(\boldsymbol{\sigma}) \cdot \left[\frac{\partial \boldsymbol{\sigma}}{\partial s} \times \frac{\partial \boldsymbol{\sigma}}{\partial t}\right] \mathrm{d}m$$

(the $\pm$ comes from the surface orientation).

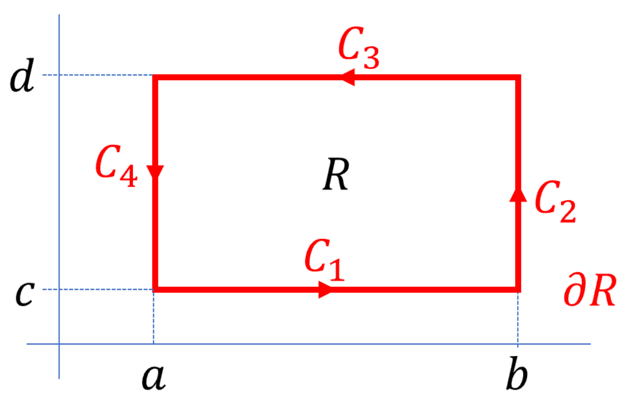
Interpretation of the Surface Integral Suppose a surface S parameterized by $\boldsymbol{\sigma}$ is "dropped" into a fluid whose flow is governed by the vector field $\mathbf{F}$. Then the **flux** of the fluid through S is given by $\int_{\boldsymbol{\sigma}} \Phi_2 \circ \mathbf{F} = \int_{\boldsymbol{\sigma}} \varphi_{\mathbf{F}}$.

14.3 Green's Theorem

Consider a **rectangle** $R = [a,b] \times [c,d] \subseteq \mathbb{R}^2$ and let ∂R be its **boundary**:

$$\partial R = ([a,b] \times \{c\}) \cup (\{b\} \times [c,d]) \cup ([a,b] \times \{d\}) \cup (\{a\} \times [c,d]),$$

together with the **induced orientation**, chosen so that as we travel ∂R, along the direction given by the orientation, the surface R falls to the left, as shown below.



Theorem 194 (Green's Theorem for a Rectangle)
Let $R=[a,b]\times[c,d]\subseteq\mathbb{R}^2$ (with the induced orientation) and $\omega\in\Omega^1(U)$, where $R\subseteq U\subseteq_O\mathbb{R}^2$. Then

$$\int_{\mathbf{R}} d\omega=\int_{\partial\mathbf{R}}\omega,$$

where $\mathbf{R}:R\to U$ and $\partial\mathbf{R}:\partial R\to U$ are the identity functions.

Proof: write $\omega=P(x,y)\,\mathrm{d}x+Q(x,y)\,\mathrm{d}y\in\Omega^1(U)$. We have seen that

$$\mathrm{d}\omega=\left(\frac{\partial Q(x,y)}{\partial y}-\frac{\partial P(x,y)}{\partial x}\right)\,\mathrm{d}x\wedge\mathrm{d}y$$

and

$$\begin{aligned}\int_{\mathbf{R}}\mathrm{d}\omega&=\int_R\mathbf{R}^*(\mathrm{d}\omega)=\int_R\tilde{\Phi}_2^{-1}(\mathbf{R}^*(\mathrm{d}\omega))\,\mathrm{d}m=\int_R\left(\frac{\partial Q(x,y)}{\partial y}-\frac{\partial P(x,y)}{\partial x}\right)\,\mathrm{d}m\\
&=\int_a^b\int_c^d\left(\frac{\partial Q(x,y)}{\partial y}-\frac{\partial P(x,y)}{\partial x}\right)\,\mathrm{d}y\,\mathrm{d}x=\int_a^b\int_c^d\frac{\partial Q(x,y)}{\partial y}\,\mathrm{d}y\,\mathrm{d}x-\int_a^b\int_c^d\frac{\partial P(x,y)}{\partial x}\,\mathrm{d}y\,\mathrm{d}x\\
&=\int_c^d\int_a^b\frac{\partial Q(x,y)}{\partial y}\,\mathrm{d}x\,\mathrm{d}y-\int_a^b\int_c^d\frac{\partial P(x,y)}{\partial x}\,\mathrm{d}y\,\mathrm{d}x,\quad\text{by Fubini's theorem (see Chapter 21)}\\
&=\int_c^d(Q(b,y)-Q(a,y))\,\mathrm{d}y-\int_a^b(P(x,d)-P(x,c))\,\mathrm{d}x\\
&=\int_a^b P(x,c)\,\mathrm{d}x+\int_c^d Q(b,y)\,\mathrm{d}y+\int_b^a P(x,d)\,\mathrm{d}x+\int_d^c Q(a,y)\,\mathrm{d}y\\
&=\int_{[a,b]}P(t,c)\,\mathrm{d}m+\int_{[c,d]}Q(b,t)\,\mathrm{d}m-\int_{[a,b]}P(t,d)\,\mathrm{d}m-\int_{[c,d]}Q(a,t)\,\mathrm{d}m.\end{aligned}$$

Now write $\partial\mathbf{R}=\mathbf{C}_1+\mathbf{C}_2+\mathbf{C}_3+\mathbf{C}_4$, where

$$\begin{aligned}&\mathbf{C}_1:[a,b]\to\mathbb{R}^2,\ \mathbf{C}_1(t)=(t,c);\quad \mathbf{C}_3:[a,b]\to\mathbb{R}^2,\ \mathbf{C}_3(t)=(b+a-t,d);\\
&\mathbf{C}_2:[c,d]\to\mathbb{R}^2,\ \mathbf{C}_2(t)=(b,t);\quad \mathbf{C}_4:[c,d]\to\mathbb{R}^2,\ \mathbf{C}_4(t)=(a,d+c-t).\end{aligned}$$

According to Proposition 189,

$$\begin{aligned}&\int_{\partial\mathbf{R}}\omega=\int_{\mathbf{C}_1}\omega+\int_{\mathbf{C}_2}\omega+\int_{\mathbf{C}_3}\omega+\int_{\mathbf{C}_4}\omega\\
&=\int_{[a,b]}\tilde{\Phi}_1^{-1}(\mathbf{C}_1^*(\omega))+\int_{[c,d]}\tilde{\Phi}_1^{-1}(\mathbf{C}_2^*(\omega))+\int_{[a,b]}\tilde{\Phi}_1^{-1}(\mathbf{C}_3^*(\omega))+\int_{[c,d]}\tilde{\Phi}_1^{-1}(\mathbf{C}_4^*(\omega)),\\
&=\int_{[a,b]}\Big[P(t,c)\cdot 1+Q(t,c)\cdot 0\Big]\,\mathrm{d}m+\int_{[a,b]}\Big[P(b+a-t,d)\cdot(-1)+Q(b+a-t,d)\cdot 0\Big]\,\mathrm{d}m\\
&+\int_{[c,d]}\Big[P(b,t)\cdot 0+Q(b,t)\cdot 1\Big]\,\mathrm{d}m+\int_{[c,d]}\Big[P(a,d+c-t)\cdot 0+Q(a,d+c-t)\cdot(-1)\Big]\,\mathrm{d}m\end{aligned}$$

so

$$\begin{aligned}\int_{\partial\mathbf{R}}\omega &= \int_{[a,b]} P(t,c)\,\mathrm{d}m + \int_{[c,d]} Q(b,t)\,\mathrm{d}m - \int_{[a,b]} P(b+a-t,d)\,\mathrm{d}m - \int_{[c,d]} Q(a,d+c-t)\,\mathrm{d}m \\ &= \int_{[a,b]} P(t,c)\,\mathrm{d}m + \int_{[c,d]} Q(b,t)\,\mathrm{d}m + \int_{[b,a]} P(s,d)\,\mathrm{d}m + \int_{[d,c]} Q(a,s)\,\mathrm{d}m \\ &= \int_{[a,b]} P(t,c)\,\mathrm{d}m + \int_{[c,d]} Q(b,t)\,\mathrm{d}m - \int_{[a,b]} P(t,d)\,\mathrm{d}m - \int_{[c,d]} Q(a,t)\,\mathrm{d}m\end{aligned},$$

which completes the proof. ■

This is a remarkable result: integrating a derivative on a rectangle is equivalent to integrating the antiderivative on the rectangle's boundary. As it happens, it is not specific to rectangles.[3]

Theorem 195 (Green's Theorem)
Let $K \subseteq_K \mathbb{R}^2$, and assume that ∂K can be given the induced orientation. If

$$\omega = P(x,y)\,dx + Q(x,y)\,dy \in \Omega^1(U)$$

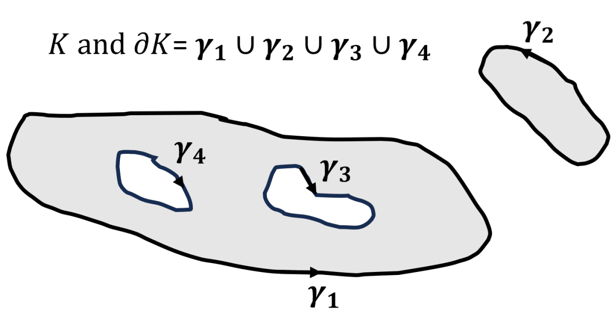
for $K \subseteq U \subseteq_O \mathbb{R}^2$, then

$$\int_{\mathbf{K}} d\omega = \int_{\partial\mathbf{K}} \omega,$$

where $\mathbf{K} : K \to \mathbb{R}^2$ and $\partial\mathbf{K} : \partial K \to \mathbb{R}^2$ are identity functions.

Proof: we only provide a sketch. Green's theorem for a rectangle can be shown to apply to unions of rectangles where each pair shares at most an edge: if the rectangles do not share edges, then the result is obvious – if they do share edges, then the induced orientation ensures that the shared edges are traversed one way for one rectangle, and the other way for another, meaning that their contribution to the integral will cancel out and only the outside boundary counts.

We can write any compact set K as a (potentially infinite) union of such rectangles $\{R_n\}$; Green's theorem holds in the limit. ■



[3] It's not even specific to $\mathbb{R}^2$, as we shall see shortly.

The classical version of Green's theorem is

$$\iint_K \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right) \mathrm{d}x\,\mathrm{d}y = \oint_{\partial K} P\,\mathrm{d}x + Q\,\mathrm{d}y.$$

Let $K \subseteq_K \mathbb{R}^2$ have a boundary with the induced orientation. By definition, we have

$$\mathrm{d}(x\,\mathrm{d}y) = \mathrm{d}x \wedge \mathrm{d}y = \mathrm{d}(-y\,\mathrm{d}x) \Longrightarrow \mathrm{d}x \wedge \mathrm{d}y = \mathrm{d}\left(\frac{1}{2}(-y\,\mathrm{d}x + \mathrm{d}y)\right) := \mathrm{d}\omega.$$

Thus, according to Green's theorem,

$$\text{Area}(K) = \iint_K \mathrm{d}A = \int_K 1 \cdot \mathrm{d}m = \int_{\mathbf{K}} \mathrm{d}\omega = \int_{\partial \mathbf{K}} \omega = \frac{1}{2}\int_{\partial \mathbf{K}} -y\,\mathrm{d}x + x\,\mathrm{d}y.$$

Example: what is the area of the ellipse

$$K = \left\{(x,y) \in \mathbb{R}^2 \;\middle|\; \frac{x^2}{a^2} + \frac{y^2}{b^2} \leq 1\right\}, \quad a, b > 0?$$

Solution: let $\boldsymbol{\gamma} : [0, 2\pi] \to \mathbb{R}^2$ be defined by $\boldsymbol{\gamma}(t) = (a\cos t, b\sin t)$; then $\boldsymbol{\gamma}$ is a parameterization of $\partial K = \boldsymbol{\gamma}([0,2\pi])$, and so

$$\begin{aligned}
\text{Area}(K) &= \frac{1}{2}\int_{\partial \mathbf{K}} -y\,\mathrm{d}x + x\,\mathrm{d}y = \frac{1}{2}\int_{[0,2\pi]} \tilde{\Phi}_1^{-1}(\boldsymbol{\gamma}^*(\omega)) \\
&= \frac{1}{2}\int_{[0,2\pi]} P(\boldsymbol{\gamma}(t))\gamma_1'(t)\,\mathrm{d}t + Q(\boldsymbol{\gamma}(t))\gamma_2'(t)\,\mathrm{d}t \\
&= \frac{1}{2}\int_{[0,2\pi]} P(a\cos t, b\sin t)(-a\sin t)\,\mathrm{d}t + Q(a\cos t, b\sin t)(b\cos t)\,\mathrm{d}t \\
&= \frac{1}{2}\int_0^{2\pi} \Big[(-b\sin t)(a\sin t) + (a\cos t)(b\cos t)\Big]\,\mathrm{d}t = \frac{1}{2}\int_0^{2\pi} ab\,\mathrm{d}t = \pi ab,
\end{aligned}$$

which we could have derived by viewing ellipses as generalized circles, but it's nice to be able to do it analytically. □

A subset $X \subseteq \mathbb{R}^n$ is **simply connected**, denoted $\pi_1(X) \simeq 1$, if X is connected and if each loop in X is **homotopic** to a single point, which is to say that each loop in X can be **deformed continuously** to a single point (see Chapter 20 for more on this topic).[4]

Example: the connected component bounded by $\boldsymbol{\gamma}_2$ in the image on the previous page is simply connected; the connected component bounded by $\boldsymbol{\gamma}_1 \cup \boldsymbol{\gamma}_3 \cup \boldsymbol{\gamma}_4$ isn't.

[4]Roughly speaking, X is simply connected if its interior contains no "hole".

Corollary 196
Let $U \subseteq_O \mathbb{R}^2$ be simply connected. If $\omega \in \Omega^1(U)$ is closed, then ω is exact.

Proof: according to Theorem 194, for any rectangle $R \subseteq U$, we have

$$\int_{\partial \mathbf{R}} \omega = \int_{\mathbf{R}} \mathrm{d}\omega;$$

since ω is closed, then $\mathrm{d}\omega = 0$, so that $\int_{\partial \mathbf{R}} \omega = 0$.

For a fixed $\mathbf{x}_0 \in U$ and for all $\mathbf{x} \in U$, there is a piecewise $\mathcal{C}^1$ path $\boldsymbol{\gamma}_{\mathbf{x}}$ connecting $\mathbf{x}_0$ to $\mathbf{x}$ that is made up of horizontal and vertical segments in U.

We would like to define $f(\mathbf{x}) = \int_{\boldsymbol{\gamma}_{\mathbf{x}}} \omega$, so that $\mathrm{d}f = \omega$ (as in the proof of Theorem 192). But this is only possible if f is well-defined, meaning that $f(\mathbf{x})$ takes on the same value independently of the piecewise $\mathcal{C}^1$ path $\boldsymbol{\gamma}_{\mathbf{x}}$ taken from $\mathbf{x}_0$ to $\mathbf{x}$, as long as it is a path of horizontal and vertical segments.

If $\boldsymbol{\gamma}_1$ and $\boldsymbol{\gamma}_2$ are two such paths, then $\boldsymbol{\gamma}_1 - \boldsymbol{\gamma}_2$ enclose a region made up of contiguous rectangles, say $R_1 \cup \cdots \cup R_k$. According to Green's theorem for rectangles,

$$\int_{\mathbf{R}_1 \cup \cdots \cup \mathbf{R}_k} \mathrm{d}\omega = \int_{\mathbf{R}_1} \mathrm{d}\omega + \cdots + \int_{\mathbf{R}_k} \mathrm{d}\omega = \int_{\partial \mathbf{R}_1} \omega + \cdots + \int_{\partial \mathbf{R}_k} \omega = \int_{\boldsymbol{\gamma}_1 - \boldsymbol{\gamma}_2} \omega = \int_{\boldsymbol{\gamma}_1} \omega - \int_{\boldsymbol{\gamma}_2} \omega.$$

Since ω is closed in U, the left hand-side of that string of equations is 0, so that $\int_{\boldsymbol{\gamma}_1} \omega = \int_{\boldsymbol{\gamma}_2} \omega$. Thus f is well-defined and the proof is complete. ■

The condition that U be simply connected is necessary: if

$$\omega = \frac{-y}{x^2 + y^2}\,\mathrm{d}x + \frac{x}{x^2 + y^2}\,\mathrm{d}y \in \Omega^1(U = \mathbb{R}^2 \setminus \{\mathbf{0}\}),$$

then we have

$$\mathrm{d}\omega = \left(\frac{\partial}{\partial x}\left(\frac{x}{x^2 + y^2}\right) - \frac{\partial}{\partial y}\left(\frac{-y}{x^2 + y^2}\right)\right) \mathrm{d}x \wedge \mathrm{d}y = \left(\frac{-x^2 + y^2}{(x^2 + y^2)^2} - \frac{y^2 - x^2}{(x^2 + y^2)^2}\right) \mathrm{d}x \wedge \mathrm{d}y = 0.$$

If $\boldsymbol{\gamma} : [0, 2\pi] \to \mathbb{R}^2$ defined by $\boldsymbol{\gamma}(t) = (\cos t, \sin t) \in U$ is a parameterization of the unit circle, we have

$$\int_{\boldsymbol{\gamma}} \omega = \int_{[0,2\pi]} \tilde{\Phi}_1^{-1}(\boldsymbol{\gamma}^*(\omega)) = \int_0^{2\pi} \mathrm{d}t = 2\pi \neq 0 = \int_{\mathbf{B}^1} \mathrm{d}\omega,$$

and so ω cannot be exact in U since the 3rd statement in Theorem 192.3 does not hold. The only fly in the ointment is that $U = \mathbb{R}^2 \setminus \{\mathbf{0}\}$ is not simply connected.

14.4 Surfaces and Orientable Surfaces in $\mathbb{R}^3$

It is fairly easy (?) to parameterize areas in $\mathbb{R}^2$, but the addition of a 3rd dimension can complicate matters to some extent (especially when it comes to their boundaries).

There are 3 classical ways to describe a plane $S \subseteq \mathbb{R}^3$.

- The **implicit** approach requires a **normal vector n** to S and a point $P_0 \in S$:

$$S = \{\mathbf{v} \in \mathbb{R}^3 \mid (\mathbf{v} - P_0) \cdot \mathbf{n} = 0\} = \{(x, y, z) \mid \underbrace{ax + by + cz}_{=F(x,y,z)} - \underbrace{(ax_0 + by_0 + cz_0)}_{=d} = 0\}.$$

- The **explicit** approach views the plane as the **graph** of a function: as $\mathbf{n} = (a, b, c) \neq \mathbf{0}$, we may assume that $c \neq 0$.[5] Then we have $cz = d - ax - by$, so that

$$z = \frac{d - ax - by}{c} = f(x, y), \quad f : \mathbb{R}^2 \to \mathbb{R}^2,$$

  and we have $F(x, y, f(x, y)) = 0$ and $S = \{(x, y, f(x, y)) \mid (x, y) \in \mathbb{R}^2\}$.

- Finally, in the **parametric** approach, let $\mathbf{v}_1, \mathbf{v}_2 \in S_0$ be linearly independent, where

$$S_0 = \{(x, y, z) \mid F(x, y, z) = ax + by + cz = 0\};$$

  hence $S_0 = \text{Span}\{\mathbf{v}_1, \mathbf{v}_2\}$. If $\mathbf{v}_0 \in S$, we have $S = \mathbf{v}_0 + S_0$. Let $\mathbf{g} : \mathbb{R}^2 \to \mathbb{R}^3$, defined by $g(s, t) = \mathbf{v}_0 + s\mathbf{v}_1 + t\mathbf{v}_2$; then $\mathbf{g}(\mathbb{R}^2) = S$ and so g is a **parameterization** of S.

These approaches generalize to non-planar surfaces. A subset $S \subseteq \mathbb{R}$ is a **surface** in $\mathbb{R}^3$ if one of the three following equivalent conditions hold.[6]

- **Explicit description:** $\forall \mathbf{p} \in S$, $\exists W_\mathbf{p} \subseteq_0 \mathbb{R}^3$ and $f : \pi_{x,y}(W_\mathbf{p}) \subseteq \mathbb{R}^2 \to \mathbb{R}$ smooth such that $S \cap W_\mathbf{p} = \text{Graph}(f)$.

- **Implicit description:** $\forall \mathbf{p} \in S$, $\exists W_\mathbf{p} \subseteq_0 \mathbb{R}^3$ and $\mathbf{F} : W_\mathbf{p} \to \mathbb{R}^3$ smooth such that

$$S \cap W_\mathbf{p} = \mathbf{F}^{-1}(\mathbf{0}) = \{\mathbf{w} \in W_\mathbf{p} \mid \mathbf{F}(\mathbf{w}) = \mathbf{0}\}$$

  and $\det(D\mathbf{F}) \neq 0$ on $S \cap W_\mathbf{p}$.

- **Parametric description:** $\forall \mathbf{p} \in S$, $\exists W_\mathbf{p} \subseteq_0 \mathbb{R}^3$ and a smooth injection $\mathbf{g} : U \subseteq \mathbb{R}^2 \to \mathbb{R}^3$ such that $\text{rank}(D\mathbf{g}(\mathbf{x})) = 2$ for all $\mathbf{x} \in U$ and such that $\mathbf{g}^{-1} : S \cap W_\mathbf{p} \to U$ is continuous. In that case, we say that $\mathbf{g}$ is a local parameterization of S.

In the latter case, the challenge is usually to find the "right" **g**.

[5] Change the variable representation, if necessary.
[6] The equivalence of the conditions is a consequence of the **implicit function theorem**.

Examples

1. Consider the unit sphere $S \subseteq \mathbb{R}^3$.

   - **Implicit descriptions:** $S = \{(x,y,z) \mid x^2+y^2+z^2=1\}$
   - **Explicit description:**
     a) If $W_1^+ = \{(x,y,z) \mid z>0\}$, $V_1 = \pi_{x,y}(W_1^+) = \{(x,y) \mid x^2+y^2<1\}$, and $f_1^+ : V_1 \to \mathbb{R}$ is given by $f_1^+(x,y) = \sqrt{1-x^2-y^2} = z$, then $S \cap W_1^+$ is the northern hemisphere.
     b) If $W_1^- = \{(x,y,z) \mid z<0\}$, $V_1 = \pi_{x,y}(W_1^-) = \{(x,y) \mid x^2+y^2<1\}$, and $f_1^- : V_1 \to \mathbb{R}$ is given by $f_1^-(x,y) = -\sqrt{1-x^2-y^2} = z$, then $S \cap W_1^-$ is the southern hemisphere.
     c) If $W_2^+ = \{(x,y,z) \mid y>0\}$, $V_2 = \pi_{x,z}(W_2^+) = \{(x,z) \mid x^2+z^2<1\}$, and $f_2^+ : V_2 \to \mathbb{R}$ is given by $f_2^+(x,z) = \sqrt{1-x^2-z^2} = y$, and so on.
   - **Parameteric description:** consider $\mathbf{g} : (0,\pi)\times(-\pi,\pi) \to \mathbb{R}^3$ defined by
     $$\mathbf{g}(\theta,\varphi) = (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta) = (x,y,z).$$
     Then
     $$D\mathbf{g}(\theta,\varphi) = \begin{pmatrix} \cos\theta\cos\varphi & -\sin\theta\sin\varphi \\ \cos\theta\sin\varphi & \sin\theta\cos\varphi \\ -\sin\theta & 0 \end{pmatrix}.$$
     It is an exercise to show that $\text{rank}(D\mathbf{g}(\theta,\varphi)) = 2$ for all (θ,φ). Furthermore, $\mathbf{g}$ is injective over $U = (0,\pi)\times(-\pi,\pi)$. Indeed, if $(\theta,\varphi), (\theta',\varphi') \in U$ and $\mathbf{g}(\theta,\varphi) = \mathbf{g}(\theta',\varphi')$, then:
     – $\cos\theta = \cos\theta' \implies \theta = \theta'$;
     – $\sin\theta\cos\varphi = \sin\theta\cos\varphi \implies \cos\varphi = \cos\varphi'$;
     – $\sin\theta\sin\varphi = \sin\theta\sin\varphi \implies \sin\varphi = \sin\varphi'$.
     – the last two equations yield $\varphi = \varphi'$ over $(-\pi,\pi)$.

     Finally, we show that that $\mathbf{g}^{-1} : \mathbf{g}(U) \to U$ defined by $\mathbf{g}(x,y,z) = (\theta,\varphi)$ is continuous. Since $z = \cos\theta$< then $\theta = \arccos z$, which is continuous. Since $-\pi/2 < \varphi/2 < \pi/2$, we have $\cos(\varphi/2) \neq 0$, and we can write
     $$\tan\frac{\varphi}{2} = \frac{\sin\theta\sin\varphi}{\sin\theta+\sin\theta\cos\varphi} = \frac{y}{\sqrt{1-z^2}+x},$$
     whence
     $$\varphi = 2\arctan\left(\frac{y}{\sqrt{1-z^2}+x}\right),$$
     which is also continuous.

   But $C = \{(x,0,z) \mid x^2+z^2=1, x\leq 0\} \subseteq S$, so we have $\mathbf{g}(U) = S \setminus C$, and so $\mathbf{g}$ is a **local** parametrization of S – it is impossible to get all of S with $\mathbf{g}$.

2. Consider the infinite cone $S : z^2 = x^2 + y^2$, $z \geq 0$.

   - **Implicit description:** $S = \{(x, y, z) \mid |x^2 + y^2 - z^2 = 0\}$
   - **Explicit description:** If $f : \mathbb{R}^2 \to \mathbb{R}$ is given by $f(x, y) = x^2 + y^2$, then $S = \{(x, y, f(x, y)) \mid (x, y) \in \mathbb{R}^2\}$
   - **Parameteric description:** consider $\mathbf{g} : U = (0, 2\pi) \times (0, a) \to \mathbb{R}^3$ defined by
   $$\mathbf{g}(\varphi, r) = (r \cos\varphi, r \sin\varphi, r).$$
   We can show that $D\mathbf{g}$ is of full rank on U, that $\mathbf{g}$ is injective on U, and that $\mathbf{g}^{-1}$ is continuous on U (see exercises).

   Finally, if $C_0 = \{(x, 0, z) \mid a > x - z \geq 0\}$, then
   $$\mathbf{g}(U) = \{(x, y, z) \mid x^2 + y^2 = z^2 < a^2\} \setminus C_0;$$
   the parameterization is local. $\square$

In both examples, the local parameterization covers the surface entirely, except for a set of **measure** (area) zero (see Chapter 21) – the missing pieces do not contribute to the integrals.

A subset $S \subseteq \mathbb{R}^3$ is a **surface with a boundary** in $\mathbb{R}^3$ if for at least some point $\mathbf{p} \in S$, there is a $W_{\mathbf{p}} \subseteq_O \mathbb{R}^3$ and a parameterization $\mathbf{g} : U \to \mathbb{R}^3$ such that $\mathbf{g}(U) = V = W_{\mathbf{p}} \cap S$ and $U \subseteq_0 \mathbb{R}^2_+$. We write $\mathbf{p} \in \partial S$ if $\mathbf{p} = \mathbf{g}(\mathbf{u})$ for some $\mathbf{u} \in \partial\mathbb{R}^2_+ = \{(x, y) \mid y = 0\}$.

Examples

1. Consider the surface S which is the northern hemisphere of the unit sphere in $\mathbb{R}^3$. Let $\mathbf{p}$ be a point of S which is not on the equator: $\exists \mathbf{0} \in U \subseteq_O \mathbb{R}^2$ and a local parameterization $\mathbf{g} : U \to \mathbb{R}^3$ such that $\mathbf{g}(\mathbf{0}) = \mathbf{p}$ and $\mathbf{g}(U) \subseteq S$. For a point $\mathbf{p}$ on the equator, we can find $\mathbf{0} \in U' \subseteq_O \mathbb{R}^2_+$ and a local parameterization $\mathbf{g}' : U' \to \mathbb{R}^3$ such that $\mathbf{g}'(\mathbf{0}) = \mathbf{p}$ and $\mathbf{g}'(U') \subseteq S$. Thus ∂S is the equator.

2. A pair of trousers S is a “surface” in $\mathbb{R}^3$; the boundary ∂S consists of the top of the waistband and the bottom of the two leg openings.

3. The ellipsoid
$$S = \left\{ (x, y, z) \in \mathbb{R}^3 \;\middle|\; \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1 \right\}$$
is a surface without a boundary. $\square$

In the last example, there is a sense in which the **volume**

$$V = \left\{ (x,y,z) \in \mathbb{R}^3 \;\middle|\; \frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} \leq 1 \right\}$$

(which is not the same as the surface S) DOES have a "boundary", namely $\partial V = S$. In general, if S is a $m-$**dimensional object**, its boundary should be a $m-1-$**dimensional object**.

14.5 Integral of a Form on an Orientable Surface

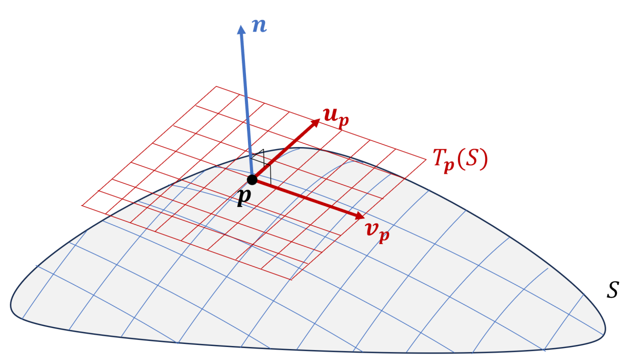
We have seen that we can induce an orientation on the boundary of planar regions; can we orient surfaces as well? Let $\mathcal{E} = \{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ and $\mathcal{E} = \{\mathbf{f}_1, \ldots, \mathbf{f}_n\}$ be two bases of $\mathbb{R}^n$, and let P be the **change of basis matrix** from $\mathcal{E}$ to $\mathcal{F}$. We say that $\mathcal{E}$ and $\mathcal{F}$ have the **same orientation** if $\det(P) > 0$ and that they have **opposite orientation** if $\det(P) < 0$.

Examples

1. In $\mathbb{R}^2$, if $\mathcal{E} = \{(1,0),(0,1)\}$ and $\mathcal{F}_\alpha = \{(\cos\alpha, \sin\alpha), (-\sin\alpha, \cos\alpha)\}$, the change of basis matrix is $P = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$ and $\det P = 1$, so $\mathcal{E}$ and $\mathcal{F}_\alpha$ have the same orientation. $\square$
2. In $\mathbb{R}^2$, if $\mathcal{E} = \{(1,0),(0,1)\}$ and $\mathcal{F} = \{(1,0),(0,-1)\}$, then $P = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ and $\det P = -1$, so $\mathcal{E}$ and $\mathcal{F}$ have opposite orientations. $\square$

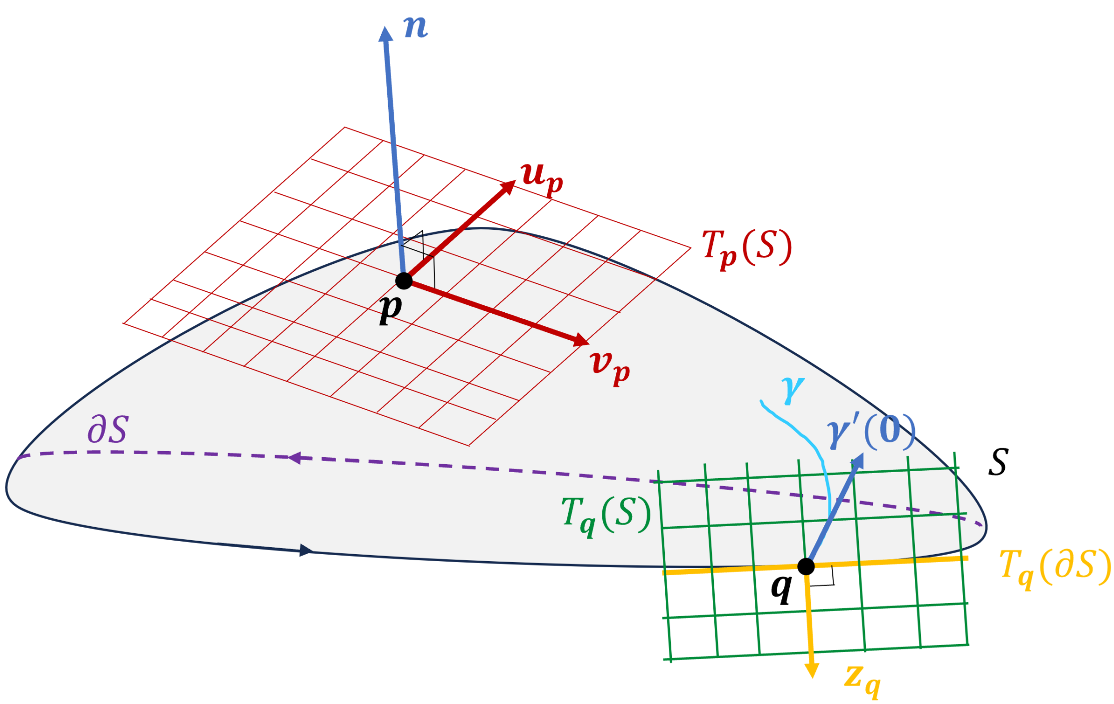
By convention, the orientation of the **canonical basis** of $\mathbb{R}^n$ is taken to be positive.

Let $S \subseteq \mathbb{R}^3$ be a surface. For all $\mathbf{p}$, let $T_{\mathbf{p}}(S) \subseteq \mathbb{R}^3$ denote the **tangent plane to** S **at p**. By definition, $T_{\mathbf{p}}(S) \simeq \mathbb{R}^2 = \text{Span}(\mathbf{u_p}, \mathbf{v_p})$, $\mathbf{n} \perp T_{\mathbf{p}}(S)$, as below. We say that S is **orientable** if it is possible to **continuously** select a basis $\{\mathbf{u_p}, \mathbf{v_p}\}$ of $T_{\mathbf{p}}(S)$ as $\mathbf{p} \in S$ varies **continuously**.[7]

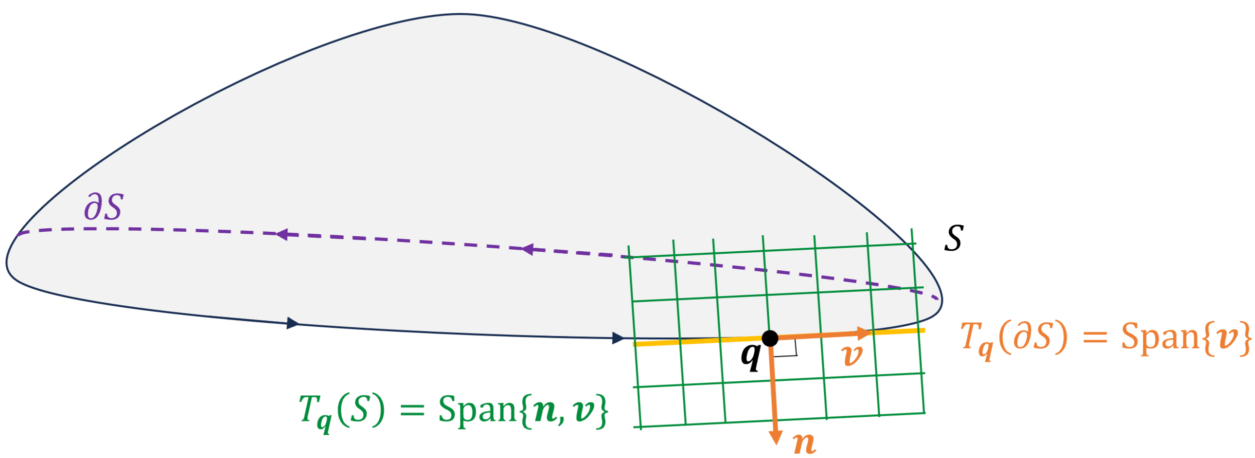


[7] Importantly, not every surface is orientable (such as a Möbius strip or a Klein bottle, for instance).

Let $S \subseteq \mathbb{R}^3$ be a compact surface with boundary ∂S. Let $\mathbf{q} \in \partial S$ and define $T_{\mathbf{q}}(\partial S) \subseteq T_{\mathbf{q}}(S)$ to be the $1-$dimensional line tangent to ∂S at $\mathbf{p}$. Pick $\alpha > 0$ and let $\boldsymbol{\gamma} : [0, t) \to S$ be a $\mathcal{C}^1$ path on S with $\boldsymbol{\gamma}(0) = \mathbf{q}$. Pick a $\mathbf{z_q} \in T_{\mathbf{q}}(S)$ such that $\mathbf{z_q} \perp T_{\mathbf{q}}(\partial S)$ and the angle between $\mathbf{z_q}$ and $\boldsymbol{\gamma}'(0) \in T_{\mathbf{q}}(S)$ is greater than a right angle. We say that $\mathbf{z_q}$ points to the **exterior of** S, whereas $-\mathbf{z_q}$ points to the **interior of** S.



The boundary ∂S is **orientable** when for all $\mathbf{q} \in \partial S$, the orientation of $T_{\mathbf{q}}(\partial S)$ is given by a vector $\mathbf{v}$ such that the orientation of $T_{\mathbf{q}}(S)$ is given by the basis $\{\mathbf{n}, \mathbf{v}\}$, where $\mathbf{n}$ is normal to $T_{\mathbf{q}}(\partial S)$ and points towards the exterior of S.



At any point of the boundary, the cross-product $\mathbf{n} \times \mathbf{v}$ (in that order) points towards the **positive orientation of the surface** S (the direction given by the **right-hand rule**).

Recall that if $U \subseteq_0 \mathbb{R}^2$ and $\omega = P(x,y)\,\mathrm{d}x \wedge \mathrm{d}y \in \Omega^2(\mathbb{R}^2)$ where P is integrable over U (see Chapter 21 for details), then

$$\int_{\mathbf{U}} \omega = \int_U P\,\mathrm{d}m, \quad \text{where} \quad \mathbf{U} : U \to \mathbb{R}^2 \equiv \text{identity on } U.$$

Let $W \subseteq_O \mathbb{R}^3$, U a Borel[8] subset of $\mathbb{R}^2$, $U \subseteq_O U$, $U \subseteq_O \mathbb{R}^2$ with $\text{Area}(U - U_0) = 0$ and let $\boldsymbol{\varphi} : U \to W$ be such that $\boldsymbol{\varphi}|_{U_0} = \boldsymbol{\varphi}_0 : U_0 \to W$ is $\mathcal{C}^1$. If $\omega \in \Omega^2(W)$, then

$$\int_{\boldsymbol{\varphi}} \omega = \int_{\boldsymbol{\varphi}_0} \omega.$$

This is well-defined, as we can see below. Let $U_0', \boldsymbol{\varphi}_0'$ be objects that satisfy the same properties as $U_0, \boldsymbol{\varphi}_0'$. Denote $\boldsymbol{\varphi}_0^*(\omega) = P_0(x,y)\,\mathrm{d}x \wedge \mathrm{d}y$ and $\boldsymbol{\varphi}'^*_0(\omega) = P_0'(x,y)\,\mathrm{d}x \wedge \mathrm{d}y$. We must show that

$$\int_U P\,\mathrm{d}m = \int_{U'} P'\,\mathrm{d}m.$$

Write $U_0'' = U_0 \cap U_0'$; we have $P_0 = P_0'$ on U_0'' and

$$U_0 \setminus U_0'' = U_0 \cap (U_0')^c \subseteq U \cap (U_0')^c = U \setminus U_0'.$$

Thus,

$$\text{Area}(U_0 \setminus U_0'') \leq \text{Area}(U_0 \setminus U_0') = 0.$$

Similarly, $\text{Area}(U_0' \setminus U_0'') = 0$, and so

$$\int_{U_0} P_0\,\mathrm{d}m = \int_{U_0''} P_0\,\mathrm{d}m = \int_{U_0''} P_0'\,\mathrm{d}m = \int_{U_0'} P_0'\,\mathrm{d}m.$$

> **Example:** let $\omega = xz^2\,\mathrm{d}y \wedge \mathrm{d}z + yx^2\,\mathrm{d}x \wedge \mathrm{d}y + zy^2\,\mathrm{d}x \wedge \mathrm{d}y \in \Omega^2(\mathbb{R}^3)$ and set $a > 0$. We consider the function $\mathbf{\Phi} : [0,\pi] \times [0, 2\pi) \to \mathbb{R}^3$ defined by
>
> $$(\theta, \varphi) \mapsto a(\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta);$$
>
> $\mathbf{\Phi}$ is a parameterization in spherical coordinates of the surface
>
> $$S_a = \{(x,y,z) \mid x^2 + y^2 + z^2 = a^2\}.$$
>
> Let $U = [0,\pi] \times [0,2\pi)$ and $U_0 = (0,\pi) \times (0,2\pi)$; then $\mathbf{\Phi}_0 = \mathbf{\Phi}|_{U_0}$ is $\mathcal{C}^1$. Since $\text{Area}(U \setminus U_0) = 0$, we have
>
> $$\int_{\mathbf{\Phi}} \omega = \int_{U_0} \mathbf{\Phi}^*(\omega).$$

[8] For all intents and purposes, U is sufficiently "nice" (see Chapter 21).

We can show that

$$\mathbf{\Phi}^*(\omega) = a^5(\sin^3\theta\cos^2\theta + \sin^5\theta\cos^2\varphi\sin^2\varphi)\,\mathrm{d}\theta\wedge\mathrm{d}\varphi,$$

and so

$$\int_{\mathbf{\Phi}}\omega = \int_0^{\pi}\int_0^{2\pi} a^5(\sin^3\theta\cos^2\theta + \sin^5\theta\cos^2\varphi\sin^2\varphi)\,\mathrm{d}\theta\,\mathrm{d}\varphi = \frac{4}{5}\pi a^5.$$

For any (θ,φ), the basis $\{\frac{\partial\mathbf{\Phi}}{\partial\theta},\frac{\partial\mathbf{\Phi}}{\partial\varphi}\}$ defines the positive orientation on S_a *via* the right-hand rule; $\mathbf{\Phi}_0$ then defines a local parameterization of S_a up to a set of area 0. □

If S is orientable in $\mathbb{R}^3$ and $\mathbf{\Phi}: U\to\mathbb{R}^3$, $\mathbf{\Psi}: V\to\mathbb{R}^3$ are two orientation-preserving parameterizations of S, let $\boldsymbol{\eta}: U\to V$ be the unique bijection such that $\mathbf{\Phi}=\mathbf{\Psi}\circ\boldsymbol{\eta}$. Then $\boldsymbol{\eta}$ is a diffeomorphism and $\forall\mathbf{u}\in U$,

$$D\mathbf{\Phi}(\mathbf{u}) = D\mathbf{\Psi}(\boldsymbol{\eta}(\mathbf{u}))D\boldsymbol{\eta}(\mathbf{u}).$$

Since $\{\frac{\partial\mathbf{\Phi}(\mathbf{u})}{\partial u_1},\frac{\partial\mathbf{\Phi}(\mathbf{u})}{\partial u_2}\}$ is a positive basis of $T_{\mathbf{\Phi}(\mathbf{u})}(S)$ and since $\{\frac{\partial\mathbf{\Psi}(\boldsymbol{\eta}(\mathbf{u}))}{\partial v_1},\frac{\partial\mathbf{\Psi}(\boldsymbol{\eta}(\mathbf{u}))}{\partial v_2}\}$ is a positive basis of $T_{\mathbf{\Phi}(\boldsymbol{\eta}(\mathbf{u}))}(S)$, both $D\mathbf{\Phi}(\mathbf{u})$ and $D\mathbf{\Psi}(\boldsymbol{\eta}(\mathbf{u}))$ transform the canonical basis of $\mathbb{R}^2$ into positive-orientation bases of $T_{\mathbf{\Phi}(\mathbf{u})}(S)$.

In that case, $D\boldsymbol{\eta}(\mathbf{u})$ preserves the orientation of $\mathbb{R}^2$ and $\det(D\boldsymbol{\eta}(\mathbf{u}))>0$ for all $\mathbf{u}$.

If $\omega\in\Omega^2(\mathbb{R}^3)$, we have $\mathbf{\Phi}^*(\omega)=a(u_1,u_2)\,\mathrm{d}u_1\wedge\mathrm{d}u_2$, $\mathbf{\Psi}^*(\omega)=b(v_1,v_2)\,\mathrm{d}v_1\wedge\mathrm{d}v_2$ for $a\in\Omega^0(U)$ and $b\in\Omega^0(V)$. Since $\mathbf{\Phi}=\mathbf{\Psi}\circ\boldsymbol{\eta}$, we have

$$\mathbf{\Phi}^*(\omega) = a\,\mathrm{d}u_1\wedge\mathrm{d}u_2 = \boldsymbol{\eta}^*(\mathbf{\Psi}^*(\omega)) = \boldsymbol{\eta}^*(b\,\mathrm{d}v_1\wedge\mathrm{d}v_2) = (b\circ\boldsymbol{\eta})\det(D\boldsymbol{\eta})\,\mathrm{d}u_1\wedge\mathrm{d}u_2.$$

Thus, according to the change of variable theorem (see Chapter 21), we have

$$\begin{aligned}\int_U\mathbf{\Phi}^*(\omega) = \int_{\mathbf{\Phi}}\omega = \int_U a\,\mathrm{d}u_1\,\mathrm{d}u_2 = \int_U(b\circ\boldsymbol{\eta})\det(D\boldsymbol{\eta})\,\mathrm{d}u_1\,\mathrm{d}u_2 = \int_U(b\circ\boldsymbol{\eta})|\det(D\boldsymbol{\eta})|\,\mathrm{d}u_1\,\mathrm{d}u_2 \\ = \int_V b\,\mathrm{d}v_1\,\mathrm{d}v_2 = \int_V\mathbf{\Psi}^*(\omega) = \int_{\mathbf{\Psi}}\omega.\end{aligned}$$

We have then proven the following result.

Theorem 197
Under the hypotheses outlined above, the integrability of ω with respect to $\mathbf{\Phi}$ and the value of $\int_{\mathbf{\Phi}}\omega$ depend only on ω and the surface $S=\mathbf{\Phi}(U)$.

We say that $\omega\in\Omega^2(\mathbb{R}^3)$ is integrable over $S\subseteq\mathbb{R}^3$ if ω is integrable with respect to a parameterization $\mathbf{\Phi}$ of S and we write $\int_S\omega=\int_{\mathbf{\Phi}}\omega$.

14.6 Area of a Surface and Flux Integral

In an exercise from the previous chapter, we saw that if $\mathbf{u}, \mathbf{v}, \mathbf{n} \in \mathbb{R}^3$ are such that **u** and **v** are not parallel, $\mathbf{n} \perp \mathbf{u}, \mathbf{v}$ with $\|\mathbf{n}\| = 1$ and

$$\varphi = n_1\,\mathrm{d}y \wedge \mathrm{d}x + n_2\,\mathrm{d}z \wedge \mathrm{d}x + n_3\,\mathrm{d}x \wedge \mathrm{d}y \in \Lambda^2(\mathbb{R}^3),$$

then $\varphi(\mathbf{u}, \mathbf{v})$ represents the **signed area** of the parallelogram bound by **u** and **v.** Thus:

- if $\mathbf{n} = \frac{\mathbf{u}\times\mathbf{v}}{\|\mathbf{u}\times\mathbf{v}\|}$, then Area $= \varphi(\mathbf{u}, \mathbf{v})$;
- if $\mathbf{n} = -\frac{\mathbf{u}\times\mathbf{v}}{\|\mathbf{u}\times\mathbf{v}\|}$, then Area $= -\varphi(\mathbf{u}, \mathbf{v})$.

Let $S \subseteq \mathbb{R}^3$ be an orientable surface, and let $\mathbf{n} : S \to \mathbb{R}^3$ be the **vector field of unit vectors normal to S, pointing towards the exterior of S.**[9]

> **Example:** consider the sphere of radius $a > 0$ centered at the origin:
>
> $$S_a = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 - a^2 = F(x, y, z) = 0\}.$$
>
> Then $\nabla F(x, y, z) = (2x, 2y, 2z) \perp S_a$ and points towards the exterior of S_a for all $(x, y, z) \in S_a$, so we could pick
>
> $$\mathbf{n}(x, y, z) = \frac{\nabla F(x, y, z)}{\|\nabla F(x, y, z)\|}. \qquad \square$$

The **area differential** $\boldsymbol{\sigma} = n_1\,\mathrm{d}y \wedge \mathrm{d}x + n_2\,\mathrm{d}z \wedge \mathrm{d}x + n_3\,\mathrm{d}x \wedge \mathrm{d}y \in \Omega^2(\mathbb{R}^3)$ is such that $\boldsymbol{\sigma} : \mathbb{R}^3 \to \Lambda^2(\mathbb{R}^3)$. According to the preceding discussion, for all $\mathbf{s} \in S \subseteq \mathbb{R}^3$, and for all $\mathbf{u}, \mathbf{v} \in T_{\mathbf{s}}(S)$, we have

$$\boldsymbol{\sigma}(\mathbf{s})(\mathbf{u}, \mathbf{v}) = \text{signed area of parallelogram bound by } \mathbf{u} \text{ and } \mathbf{v}.$$

Using the above notation, we then have the following result.

> **Proposition 198**
> *For an orientable surface $S \subseteq \mathbb{R}^3$, let $\boldsymbol{\sigma} \in \Omega^2(\mathbb{R}^3)$ be the area differential of S. Then the signed area of S is given by $\int_S \omega$.*

We sometimes used the following formulation:

$$\text{Signed Area}(S) = \iint_{U_0} \left\| \frac{\partial \boldsymbol{\sigma}}{\partial s} \times \frac{\partial \boldsymbol{\sigma}}{\partial t} \right\| \mathrm{d}s\,\mathrm{d}t,$$

where $\boldsymbol{\Phi} : U_0 \to \mathbb{R}^3$ is a parameterization of S.

[9] In other words, we can find a continuous mapping $\mathbf{s} \in S \mapsto \{\mathbf{u}(\mathbf{s}), \mathbf{v}(\mathbf{s})\}$, where $\{\mathbf{u}(\mathbf{s}), \mathbf{v}(\mathbf{s})\} \in T_{\mathbf{s}}(S)$ defines the orientation of S, so that $\{\mathbf{n}(\mathbf{s}), \mathbf{u}(\mathbf{s}), \mathbf{v}(\mathbf{s})\}$ forms a basis of $\mathbb{R}^3$ with **positive orientation.**

Example: consider the unit sphere

$$S = \{(x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 - 1 = F(x, y, z) = 0\}.$$

The outward normal vector field $\mathbf{n} : S \to \mathbb{R}^3$ is given by

$$\mathbf{n}(x, y, z) = \frac{\nabla F(x, y, z)}{\|\nabla F(x, y, z)\|} = (x, y, z) \perp S.$$

The area differential of S is thus $\boldsymbol{\sigma} = x\,\mathrm{d}y \wedge \mathrm{d}x + y\,\mathrm{d}z \wedge \mathrm{d}x + z\,\mathrm{d}x \wedge \mathrm{d}y \in \Omega^2(\mathbb{R}^3)$. In order to calculate $\int_S \boldsymbol{\sigma}$, we use the following parameterization of S:

$$\mathbf{\Phi} : U_0 = [0, \pi] \times [0, 2\pi) \to \mathbb{R}^3, \quad \text{where} \quad \mathbf{\Phi}(\theta, \varphi) = (\sin\theta\cos\varphi, \sin\theta\sin\varphi, \cos\theta),$$

and

$$\int_S \boldsymbol{\sigma} = \int_{\mathbf{\Phi}} \boldsymbol{\sigma} = \int_{U_0} \mathbf{\Phi}^*(\boldsymbol{\sigma}).$$

But $\mathbf{\Phi}^*(\boldsymbol{\sigma}) = (\sin^3\theta + \cos^2\theta\sin\theta)\,\mathrm{d}\theta \wedge \mathrm{d}\varphi$, so that

$$\int_{U_0} \mathbf{\Phi}^*(\boldsymbol{\sigma}) = \int_0^\pi \int_0^{2\pi} (\sin^3\theta + \cos^2\theta\sin\theta)\,\mathrm{d}\theta\,\mathrm{d}\varphi = 4\pi. \qquad \square$$

14.7 Stokes' Theorem

We finish this chapter (and this part of the course notes) with a generalization of Green's theorem, which we unfortunately present without proof.

Theorem 199 (STOKES' THEOREM)
Let $M \subseteq W \subseteq_O \mathbb{R}^n$ be a compact orientable manifold with orientable boundary ∂M such that $\dim(M) = p$. If $\omega \in \Omega^{p-1}(W)$, then $\int_{\partial M} \omega = \int_M d\omega$.

When $M = S \subseteq \mathbb{R}^3$ and $p = \dim(M) = 2$, then we usually write Stokes' theorem as

$$\int_S (\nabla \times \mathbf{F}) \cdot \mathrm{d}\mathbf{A} = \oint_{\partial S} \mathbf{F} \cdot \mathrm{d}\mathbf{r}.$$

Corollary 200
Let $\partial M = \varnothing$ in Theorem 199. If $\phi \in \Omega^p(W)$ is exact, then $\int_M \phi = 0$.

Proof: since φ is exact, $\exists \eta \in \Omega^{p-1}(W)$ such that $\mathrm{d}\eta = \varphi$, so that

$$\int_M \varphi = \int_M \mathrm{d}\eta = \int_{\partial M} \eta = 0. \qquad \blacksquare$$

14.8 Solved Problems

Let's do some vector calculus!

1. Let $\mathbf{F}(x,y) = (xy, x-y)$ and C be the boundary of the triangle with vertices $(1,0)$, $(-1,0)$ and $(0,1)$. Compute the line integral $\int_C \mathbf{F}\cdot d\mathbf{r}$.

   Solution: the triangle is parameterized by
   $$C_1 : (t,0), \quad -1\le t\le 1, \quad C_2 : (1-t,t), \quad 0\le t\le 1, \quad C_3 : (-t,1-t), \quad 0\le t\le 1.$$
   Thus, the line integral of interest is
   $$\begin{aligned} I &= \oint_C \mathbf{F}\cdot d\mathbf{r} = \int_{C_1}\mathbf{F}\cdot d\mathbf{r} + \int_{C_2}\mathbf{F}\cdot d\mathbf{r} + \int_{C_3}\mathbf{F}\cdot \mathrm{d}\mathbf{r} \\ &= \int_{-1}^{1}(t^2,t)\cdot(1,0)\,\mathrm{d}t + \int_0^1 (t-t^2, 1-2t)\cdot(-1,1)\,\mathrm{d}t + \int_0^1 (t^2-t,-1)\cdot(-1,-1)\,\mathrm{d}t = 1. \end{aligned}$$
   Under the other orientation, the answer is -1. $\square$

2. Let $\mathbf{F}(x,y) = (2xe^{x^2}\sin y, e^{x^2}\cos y)$ and C be the path defined by $x(t)=t$, $y(t)=\frac{\pi}{2}t$, $0\le t\le 1$.

   a) Compute $\int_C \mathbf{F}\cdot d\mathbf{r}$ directly.
   b) Compute $\int_C \mathbf{F}\cdot d\mathbf{r}$ using the fundamental theorem of line integrals.

   Solution:

   a) We have
   $$\begin{aligned} I &= \int_C \mathbf{F}\cdot \mathrm{d}\mathbf{r} = \int_0^1 (2te^{t^2}\sin(\pi t/2), e^{t^2}\cos(\pi t/2))\cdot(1,\pi/2)\,\mathrm{d}t \\ &= \int_0^1 e^{t^2}(2t\sin(\pi t/2) + \pi/2\cos(\pi t/2))\,\mathrm{d}t = \left[e^{t^2}\sin(\pi t/2)\right]_0^1 = e. \end{aligned}$$

   b) Let $f(x,y) = e^{x^2}\sin y$. Then $\mathbf{F} = \nabla f$ and
   $$\int_C \mathbf{F}\cdot\mathrm{d}\mathbf{r} = f(1,\pi/2) - f(0,0) = e - 0 = e,$$
   according to the fundamental theorem of line integrals. $\square$

3. Compute $\int_C \mathbf{F}\cdot \mathrm{d}\mathbf{r}$, if $\mathbf{F}(x,y) = (x^2y, -xy)$ and $C = \{\mathbf{r}(t) = (t^3,t^4) \mid 0\le t\le 1\}$.

   Solution: we have $\mathbf{r}'(t) = (3t^2, 4t^3)$. Thus,
   $$\begin{aligned} \int_C \mathbf{F}\cdot\mathrm{d}\mathbf{r} &= \int_0^1 \mathbf{F}(\mathbf{r}(t))\cdot\mathbf{r}'(t)\,\mathrm{d}t = \int_0^1 \mathbf{F}\left(t^3,t^4\right)\cdot(3t^2,4t^3)\,\mathrm{d}t \\ &= \int_0^1 \left(t^{10},-t^7\right)\cdot(3t^2,4t^3)\,\mathrm{d}t = \int_0^1 \left(3t^{12}-4t^{10}\right)\,\mathrm{d}t \\ &= \left[\frac{3t^{13}}{13} - \frac{4t^{11}}{11}\right]_0^1 = -\frac{19}{143}. \qquad \square \end{aligned}$$

4. Compute $\int_C \mathbf{F}\cdot d\mathbf{r}$, where $\mathbf{F}(x,y,z)=(y+z,-x^2,-4y^2)$ and
$$C=\{\mathbf{r}(t)=(t,t^2,t^4)\mid 0\leq t\leq 1\}.$$

Solution: in this case, we have $\mathbf{r}'(t)=(1,2t,4t^3)$. Thus,
$$\begin{aligned}\int_C \mathbf{F}\cdot d\mathbf{r}&=\int_0^1 \mathbf{F}(\mathbf{r}(t))\cdot\mathbf{r}'(t)\,dt=\int_0^1 \mathbf{F}\left(t,t^2,t^4\right)\cdot(1,2t,4t^3)\,dt\\&=\int_0^1\left(t^2+t^4,-t^2,-4t^4\right)\cdot(1,2t,4t^3)\,dt=\int_0^1\left(t^2-2t^3+t^4-16t^7\right)\,dt\\&=\left[\frac{t^3}{3}-\frac{t^4}{2}+\frac{t^5}{5}-2t^8\right]_0^1=-\frac{59}{30}.\qquad\square\end{aligned}$$

5. Compute $\int_C \mathbf{F}\cdot d\mathbf{r}$ if $\mathbf{F}(x,y,z)=(\sin x,\cos y,xz)$ and
$$C=\{\mathbf{r}(t)=(t^3,-t^2,t)\mid 0\leq t\leq 1\}.$$

Solution: in this case, we have $\mathbf{r}'(t)=(3t^2,-2t,1)$. Thus,
$$\begin{aligned}\int_C \mathbf{F}\cdot d\mathbf{r}&=\int_0^1 \mathbf{F}(\mathbf{r}(t))\cdot\mathbf{r}'(t)\,dt=\int_0^1 \mathbf{F}\left(t^3,-t^2,t\right)\cdot(3t^2,-2t,1)\,dt\\&=\int_0^1\left(\sin(t^3),\cos(-t^2),t^4\right)\cdot(3t^2,-2t,1)\,dt=\int_0^1\left(3t^2\sin(t^3)-2t\cos(-t^2)+t^4\right)\,dt\\&=\left[-\cos(t^3)-\sin(t^2)+\frac{t^5}{5}\right]_0^1=\frac{6}{5}-\cos(1)-\sin(1).\qquad\square\end{aligned}$$

6. Are $\mathbf{F}(x,y)=(ye^x+\sin y,e^x+x\cos y)$ and $\mathbf{F}(x,y)=(ye^{xy}+4x^3y,xe^{xy}+x^4)$ a conservative vector fields? If so, find their potential.

Solution: the vector field $\mathbf{F}$ is conservative if and only if
$$\frac{\partial F_1}{\partial y}=\frac{\partial F_2}{\partial x}.$$

Since
$$\begin{aligned}\frac{\partial F_1}{\partial y}&=\frac{\partial}{\partial y}(ye^x+\sin y)=e^x+\cos y\\\frac{\partial F_2}{\partial x}&=\frac{\partial}{\partial x}(e^x+x\cos y)=e^x+\cos y\end{aligned}$$

the field is conservative. In this case, the potential f satisfies $\nabla f=\mathbf{F}$, that is
$$\begin{aligned}f_x(x,y)&=F_1(x,y)=ye^x+\sin y\\f_y(x,y)&=F_2(x,y)=e^x+x\cos y\end{aligned}$$

whence
$$f(x,y)=\int f_x(x,y)dx=\int(ye^x+\sin y)\,dx=ye^x+x\sin y+k(y),$$

where $k(y)$ is a function of y. Substituting this function f in the equation for f_y, we have

$$f_y(x,y) = e^x + x\cos y + k'(y) = e^x + x\cos y;$$

the function $k(y)$ is a constant since the derivative in y is zero. Thus, the family of potential for **F** is $f(x,y) = ye^x + x\sin y + k$, $k \in \mathbb{R}$.

Since

$$\begin{aligned}\frac{\partial F_1}{\partial y} &= \frac{\partial}{\partial y}\left(ye^{xy} + 4x^3y\right) = e^{xy} + xye^{xy} + 4x^3\\ \frac{\partial F_2}{\partial x} &= \frac{\partial}{\partial x}\left(xe^{xy} + x^4\right) = e^{xy} + xye^{xy} + 4x^3\end{aligned}$$

the second field is conservative. In this case, the potential f satisfies $\nabla f = \mathbf{F}$, that is

$$\begin{aligned}f_x(x,y) &= F_1(x,y) = ye^{xy} + 4x^3y\\ f_y(x,y) &= F_2(x,y) = xe^{xy} + x^4\end{aligned}$$

whence

$$f(x,y) = \int f_x(x,y)dx = \int \left(ye^{xy} + 4x^3y\right)dx = e^{xy} + x^4y + k(y),$$

where $k(y)$ is a function of y. Substituting this function f in the equation for f_y, we have

$$f_y(x,y) = xe^{xy} + x^4 + k'(y) = xe^{xy} + x^4;$$

the function $k(y)$ is a constant since the derivative in y is zero. Thus, the family of potential for **F** is $f(x,y) = e^{xy} + x^4y + k$, $k \in \mathbb{R}$. □

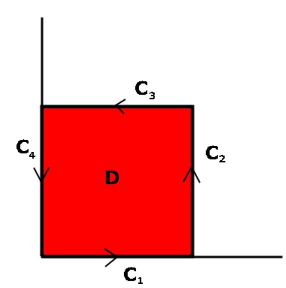
7. Find a potential for these vector fields, if one exists.

   a) $\mathbf{F}(x,y) = (2xy^3, 3x^2y + x)$;
   b) $\mathbf{F}(x,y) = (2xy^3 + y, 3x^2y + x)$;
   c) $\mathbf{F}(x,y) = (2xy, x^2 + 8y)$.

   Solution: a) and b) do not have potential functions, but $f(x,y) = x^2y + 4y^2$ is a potential function for c). □

8. Using the direct approach and Green's theorem, compute $\int_C \mathbf{F}\cdot d\mathbf{r}$, where C is the square with vertices $(0,0)$, $(1,0)$, $(1,1)$, $(0,1)$, and $\mathbf{F}(x,y) = (x^2y, xy^3)$.

   Solution: the region is shown below.



Let C_1 be the segment from $(0,0)$ to $(1,0)$; C_2 the segment from $(1,0)$ to $(1,1)$; C_3 the segment from $(1,1)$ to $(0,1)$, and C_4 the segment from $(0,1)$ to $(0,0)$. Thus

$$\begin{aligned}
C_1 &= \{\mathbf{r}(t) = (t,0) \mid 0 \le t \le 1\} \\
C_2 &= \{\mathbf{r}(t) = (1,t) \mid 0 \le t \le 1\} \\
C_3 &= \{\mathbf{r}(t) = (1-t,1) \mid 0 \le t \le 1\} \\
C_4 &= \{\mathbf{r}(t) = (0,1-t) \mid 0 \le t \le 1\}
\end{aligned}$$

and

$$\int_C \mathbf{F}\cdot \mathrm{d}\mathbf{r} = \int_{C_1} \mathbf{F}\cdot \mathrm{d}\mathbf{r} + \int_{C_2} \mathbf{F}\cdot \mathrm{d}\mathbf{r} + \int_{C_3} \mathbf{F}\cdot \mathrm{d}\mathbf{r} + \int_{C_4} \mathbf{F}\cdot \mathrm{d}\mathbf{r}.$$

We can show with ease that

$$\begin{aligned}
\int_{C_1} \mathbf{F}\cdot \mathrm{d}\mathbf{r} &= \int_0^1 (t^2(0), t(0)^3)\cdot(1,0)\,\mathrm{d}t = 0 \\
\int_{C_2} \mathbf{F}\cdot \mathrm{d}\mathbf{r} &= \int_0^1 (1^2(t), 1t^3)\cdot(0,1)\,\mathrm{d}t = \int_0^1 t^3\,\mathrm{d}t = \frac{1}{4} \\
\int_{C_3} \mathbf{F}\cdot \mathrm{d}\mathbf{r} &= \int_0^1 ((1-t)^2(1), (1-t)(1)^3)\cdot(-1,0)\,\mathrm{d}t = \int_0^1 -(1-t)^2\,\mathrm{d}t = -\frac{1}{3} \\
\int_{C_4} \mathbf{F}\cdot \mathrm{d}\mathbf{r} &= \int_0^1 (0^2(1-t), 0(1-t)^3)\cdot(0,-1)\,\mathrm{d}t = 0
\end{aligned}$$

so that

$$\int_C \mathbf{F}\cdot \mathrm{d}\mathbf{r} = 0 + \frac{1}{4} - \frac{1}{3} + 0 = -\frac{1}{12}.$$

Using Green's theorem instead, we have

$$\int_C \mathbf{F}\cdot \mathrm{d}\mathbf{r} = \iint_D \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}\right)\,\mathrm{d}A = \iint_D \left(y^3 - x^2\right)\,\mathrm{d}A,$$

where the region of integration D (in red) is bounded by the curve C, with the positive orientation. Since $D = \{(x,y) \mid 0 \le x \le 1, 0 \le y \le 1\}$, we have

$$\begin{aligned}
\iint_D \left(y^3 - x^2\right)\,\mathrm{d}A &= \int_0^1 \int_0^1 (y^3 - x^2)\,dy\,dx = \int_0^1 \left[\frac{y^4}{4} - x^2 y\right]_{y=0}^{y=1} dx \\
&= \int_0^1 \left(\frac{1}{4} - x^2\right) dx = \left[\frac{x}{4} - \frac{x^3}{3}\right]_0^1 = \frac{1}{4} - \frac{1}{3} = -\frac{1}{12}.
\end{aligned}$$

This completes the computations. □

9. Compute $\int_C \mathbf{F}\cdot \mathrm{d}\mathbf{r}$, where C is the circle $x^2 + y^2 = 4$ from $(2,0)$ to $(\sqrt{2},\sqrt{2})$, then along the segment from $(\sqrt{2},\sqrt{2})$ to the origin and finally along the segment from the origin to $(2,0)$ (with the positive orientation), for $\mathbf{F}(x,y) = (y^2 - x^2 y, xy^2)$.

   Solution: according to Green's theorem,

$$\int_C \mathbf{F}\cdot \mathrm{d}\mathbf{r} = \iint_D \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}\right)\,\mathrm{d}A = \iint_D \left(y^2 - 2y + x^2\right)\,\mathrm{d}A,$$

where the region D is bounded by the curve C, oriented positively. In polar coordinates,

$$D_{(r,\theta)} = \{(r,\theta) \mid 0 \leq r \leq 2, 0 \leq \theta \leq \frac{\pi}{4}\},$$

and $y^2 - 2y + x^2 = r^2 - 2r\sin\theta$, whence

$$\iint_D \left(y^2 - 2y + x^2\right)\, \mathrm{d}A = \int_0^{\pi/4}\int_0^2 (r^2 - 2r\sin\theta) r\, \mathrm{d}r\, \mathrm{d}\theta = \int_0^{\pi/4}\left[4 - \frac{16}{3}\sin\theta\right]\, \mathrm{d}\theta$$

$$= \left[4\theta + \frac{16}{3}\cos\theta\right]_0^{\pi/4} = \pi + \frac{8}{3}(\sqrt{2} - 2).$$

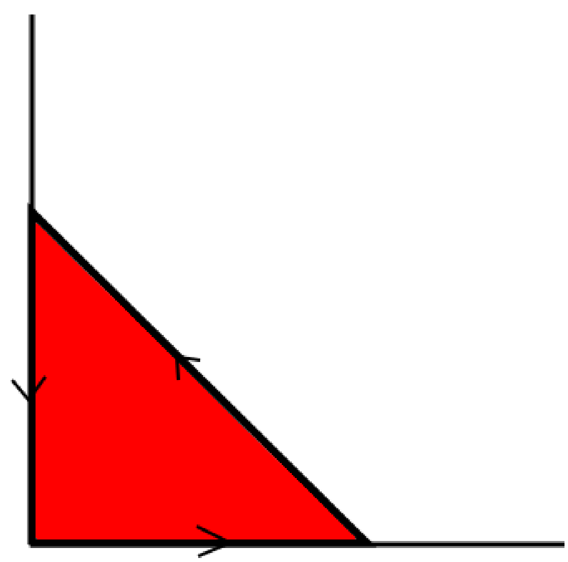
□

10. What is the work accomplished by the vector field $\mathbf{F}(x,y) = (x(x+y), xy^2)$ on a particle traveling along the $x-$axis from the origin to $(1,0)$, then from $(1,0)$ to $(0,1)$ along a straight line, and finally back to the origin along the $y-$axis?

Solution: the work in question is given by

$$W = \int_C \mathbf{F}\cdot \mathrm{d}\mathbf{r} = \iint_D \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}\right)\, \mathrm{d}A = \iint_D \left(y^2 - x\right)\, \mathrm{d}A,$$

where the region D (in red) is bounded by the curve C, oriented positively.



Since $D = \{(x,y) \mid 0 \leq x \leq 1, 0 \leq y \leq 1 - x\}$,

$$W = \iint_D \left(y^2 - x\right)\, \mathrm{d}A = \int_0^1\int_0^{1-x} (y^2 - x)\, dy\, dx = \int_0^1 \left[\frac{y^3}{3} - xy\right]_{y=0}^{y=1-x} dx$$

$$= \int_0^1 \left(\frac{(x-1)^3}{3} - x(1-x)\right) dx = \left[-\frac{1}{12}(1-x)^4 - \frac{x^2}{2} + \frac{x^3}{3}\right]_0^1 = -\frac{1}{12}. \quad \square$$

11. Let $\mathbf{F}(x,y,z) = (\frac{z}{2}, y, 2x)$ and S be the rectangle with vertices $(2,0,4)$, $(2,3,4)$, $(0,0,4)$ et $(0,3,4)$. Compute the surface integral $\iint_S \mathbf{F}\cdot d\mathbf{A}$.

Solution: the region S is parameterized by

$$x(s,t) = s,\ y(s,t) = t,\ z(s,t) = 4,\ (s,t) \in D : 0 \leq s \leq 2,\ 0 \leq t \leq 3$$

Thus, $\mathbf{v}_s = (1,0,0)$, $\mathbf{v}_t = (0,1,0)$ and

$$\mathbf{v}_s \times \mathbf{v}_t = (0,0,1).$$

Restricted to S, the vector field takes the form

$$\mathbf{F}(x(s,t), y(s,t), z(s,t)) = (2, t, 2s).$$

The positive orientation of the surface S was not specified, so we select the upwards orientation as the positive orientation. Since $\mathbf{v}_s \times \mathbf{v}_t = (0,0,1)$ points upwards,

$$I = \iint_S \mathbf{F}\cdot \mathrm{d}\mathbf{A} = \iint_D (2,t,2s)\cdot(0,0,1)\,ds\,\mathrm{d}t = \int_0^3 \int_0^2 2s\,ds\,\mathrm{d}t = 12. \qquad \square$$

12. Let $\mathbf{F}(x,y,z) = (x,y,z)$ and S be the surface defined by $z = -2x - 4y + 1$ in the first octant. Compute the surface integral $\iint_S \mathbf{F}\cdot d\mathbf{A}$.

Solution: the region S is parameterized by

$$x(s,t) = s,\ y(s,t) = t,\ z(s,t) = -2s - 4t + 1,\ (s,t) \in D : 0 \leq t \leq 1/4,\ 0 \leq s \leq 1/2 - 2t$$

Thus, $\mathbf{v}_s = (1,0,-2)$, $\mathbf{v}_t = (0,1,-4)$ and

$$\mathbf{v}_s \times \mathbf{v}_t = (2,4,1).$$

Restricted to S, the vector field becomes

$$\mathbf{F}(x(s,t), y(s,t), z(s,t)) = (s, t, -2s - 4t + 1).$$

The positive orientation of S is still not specified, so we select the upwards orientation. Since $\mathbf{v}_s \times \mathbf{v}_t = (2,4,1)$ points upwards, we have

$$\begin{aligned} I &= \iint_S \mathbf{F}\cdot \mathrm{d}\mathbf{A} = \iint_D (s,t,-2s-4t+1)\cdot(2,4,1)\,ds\,\mathrm{d}t \\ &= \int_0^{1/4} \int_0^{1/2-2t} 1\,ds\ \mathrm{d}t = \int_0^{1/4} (1/2 - 2t)\,\mathrm{d}t = \frac{1}{16}. \qquad \square \end{aligned}$$

13. Let $\mathbf{F}(x,y,z) = (-xz, -yz, z^2)$ and S be the surface $z^2 = x^2 + y^2$ lying above the plane $z = 0$ and below the plane $z = 1$. Compute $\iint_S \mathbf{F}\cdot d\mathbf{A}$.

Solution: the region S is parameterized by

$$x(s,t) = s,\ y(s,t) = t,\ z(s,t) = \sqrt{s^2 + t^2},\ (s,t) \in D : s^2 + t^2 \leq 1.$$

Thus, $\mathbf{v}_s = (1, 0, \frac{s}{\sqrt{s^2+t^2}})$, $\mathbf{v}_t = (0, t, \frac{t}{\sqrt{s^2+t^2}})$ and

$$\mathbf{v}_s \times \mathbf{v}_t = (-s(s^2+t^2)^{-1/2}, -t(s^2+t^2)^{-1/2}, 1).$$

Restricted to S, the vector field becomes

$$\mathbf{F}(x(s,t), y(s,t), z(s,t)) = (-s\sqrt{s^2+t^2}, -t\sqrt{s^2+t^2}, s^2+t^2).$$

The positive orientation of S is once again not specified, we again select the upwards orientation as the positive one. Since $\mathbf{v}_s \times \mathbf{v}_t$ points upwards, we have

$$\begin{aligned} I &= \iint_S \mathbf{F} \cdot \mathrm{d}\mathbf{A} = \iint_D (-s\sqrt{s^2+t^2}, -t\sqrt{s^2+t^2}, s^2+t^2) \cdot (\frac{-s}{\sqrt{s^2+t^2}}, \frac{-t}{\sqrt{s^2+t^2}}, 1)\, ds\, \mathrm{d}t \\ &= 2\iint_D (s^2+t^2)\, ds\, \mathrm{d}t. \end{aligned}$$

In polar coordinates, this last integral is easy to evaluate: $s = r\cos\theta$, $t = r\sin\theta$, $0 \leq r \leq 1$, $0 \leq \theta \leq 2\pi$:

$$I = 2\iint_D (s^2+t^2)\, ds\ \mathrm{d}t = 2\int_0^1 \int_0^{2\pi} (r^2\cos^2\theta + r^2\sin^2\theta) r\, \mathrm{d}\theta\, \mathrm{d}r = 2\int_0^1 \int_0^{2\pi} r^3\, \mathrm{d}\theta\, dr = \pi. \qquad \square$$

14. Let $\mathbf{F}(x,y,z) = (y,x,0)$ and S be the surface defined by $x^2+y^2 = 9$, $0 \leq x \leq 3$, $-3 \leq y \leq 3$, $1 \leq z \leq 2$. Compute the surface integral $\iint_S \mathbf{F} \cdot d\mathbf{A}$.

    Solution: without a single computation, it is possible to determine that the flux must be zero. Why is that? $\square$

15. Let $\mathbf{F}(x,y,z) = (x,0,0)$ and let S be the surface parameterized by $x = e^p$, $y = \cos 3q$, $z = 6p$, $0 \leq p \leq 4$, $0 \leq q \leq \frac{\pi}{6}$. Compute the surface integral $\iint_S \mathbf{F} \cdot d\mathbf{A}$.

    Solution: the region S is parameterized by

$$x(p,q) = e^p,\ y(p,q) = \cos(3q),\ z(p,q) = 6p,\ (s,t) \in D : 0 \leq p \leq 4,\ 0 \leq q \leq \frac{\pi}{6}.$$

Thus, $\mathbf{v}_p = (e^p, 0, 6)$, $\mathbf{v}_q = (0, -3\sin 3q, 0)$ and

$$\mathbf{v}_p \times \mathbf{v}_q = (18\sin(3q), 0, -3e^p\sin(3q)).$$

Restricted to S, the vector field becomes

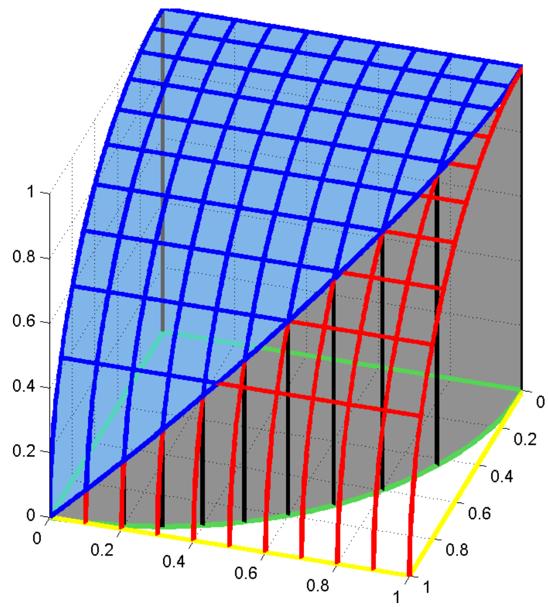
$$\mathbf{F}(x(p,q), y(p,q), z(p,q)) = (e^p, 0, 0).$$

Guess what, the surface orientation has not been specified, so we select the positive $x-$axis as a positive orientation. Since the first component of $\mathbf{v}_p \times \mathbf{v}_q$ is positive when $0 \leq q \leq \frac{\pi}{6}$, we have

$$\begin{aligned} I &= \iint_S \mathbf{F} \cdot \mathrm{d}\mathbf{A} = \iint_D (e^p, 0, 0) \cdot (18\sin(3q), 0, -3e^p\sin(3q))\, dp\, dq \\ &= 18\int_0^{\pi/6} \int_0^4 e^p \sin(3q)\, dp\, dq = 6(e^4 - 1). \qquad \square \end{aligned}$$

16. What is the area of the piece S of the cylinder $x^2 + z^2 = a^2$ bounded by the surface of the cylinder $x^2 + y^2 = a^2$, where $a > 0$?

Solution: in the image below, the situation is illustrated in the first octant, for $a = 1$: the cylinder $x^2 + z^2 = a^2$ appears in grey, the cylinder $x^2 + y^2 = a^2$ in red. The part of S in the first octant shows up in blue.



The surface S is parameterized by

$$x = p, \quad y = q, \quad z = \sqrt{a^2 - p^2}, \quad (p, q) \in \Omega$$

where Ω is the region of the $xy-$plane bounded by the green curve. Accordingly,

$$A(S) = 8 \iint_\Omega \|\mathbf{v}_p \times \mathbf{v}_q\| \, dq \, dp,$$

where

$$\mathbf{v}_p = \left(1, 0, -\frac{p}{\sqrt{a^2 - p^2}}\right), \quad \mathbf{v}_q = (0, 1, 0)$$

and

$$\mathbf{v}_p \times \mathbf{v}_q = \left(\frac{p}{\sqrt{a^2 - p^2}}, 0, 1\right),$$

whence

$$\|\mathbf{v}_p \times \mathbf{v}_q\| = \frac{a}{\sqrt{a^2 - p^2}}.$$

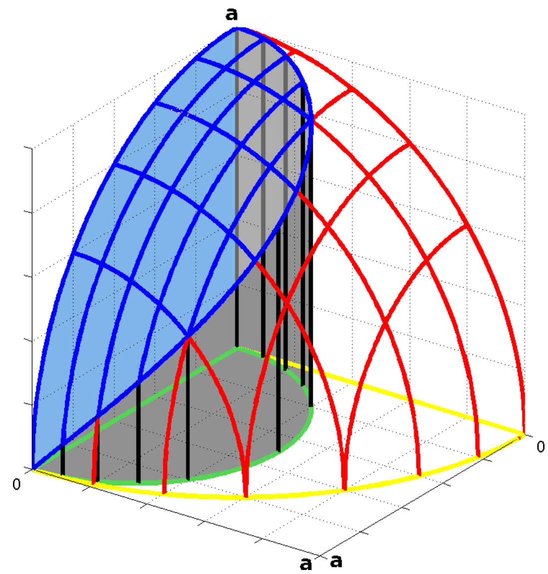
Thus,

$$A(S) = 8 \iint_\Omega \|\mathbf{v}_p \times \mathbf{v}_q\| \, dp \, dq = 8 \iint_\Omega \frac{a}{\sqrt{a^2 - p^2}} \, dq \, dp = 8 \int_0^a \int_0^{\sqrt{a^2-p^2}} \frac{a}{\sqrt{a^2 - p^2}} \, dq \, dp$$

$$= 8 \int_0^a \left[\frac{a}{\sqrt{a^2 - p^2}} q\right]_0^{\sqrt{a^2-p^2}} dp = 8a \int_0^a dp = 8a^2. \qquad \square$$

17. What is the area of the piece S of the sphere $x^2+y^2+z^2=a^2$ bounded by the surface of the cylinder $x^2+y^2=ax$, where $a>0$?

Solution: in the image below, the situation is illustrated in the first octant, for $a=1$: the cylinder $x^2+y^2=ax$ appears in grey, the sphere $x^2+y^2+z^2=a^2$ in red. The part of S in the first octant shows up in blue.



The surface S is parameterized by

$$x=p,\quad y=q,\quad z=\sqrt{a^2-p^2-q^2},\quad (p,q)\in\Omega$$

where Ω is the region of the $xy-$plane bounded by the green curve. Accordingly,

$$A(S)=4\iint_\Omega \|\mathbf{v}_p\times\mathbf{v}_q\|\,dq\,dp,$$

where

$$\mathbf{v}_p=\left(1,0,-\frac{p}{\sqrt{a^2-p^2-q^2}}\right),\quad \mathbf{v}_q=\left(0,1,-\frac{q}{\sqrt{a^2-p^2-q^2}}\right)$$

and

$$\mathbf{v}_p\times\mathbf{v}_q=\left(\frac{p}{\sqrt{a^2-p^2-q^2}},\frac{q}{\sqrt{a^2-p^2-q^2}},1\right),\quad \text{whence } \|\mathbf{v}_p\times\mathbf{v}_q\|=\frac{a}{\sqrt{a^2-p^2-q^2}}.$$

Thus,

$$A(S)=4\iint_\Omega \|\mathbf{v}_p\times\mathbf{v}_q\|\,dp\,dq=4\iint_\Omega \frac{a}{\sqrt{a^2-p^2-q^2}}\,dq\,dp=4\int_0^a\int_0^{\sqrt{ap-p^2}}\frac{a}{\sqrt{a^2-p^2-q^2}}\,dq\,dp$$

$$=4\int_0^a\left[a\arctan\left(\frac{q}{\sqrt{a^2-p^2-q^2}}\right)\right]_0^{\sqrt{ap-p^2}}dp=4a\int_0^a\arctan\left(\sqrt{\frac{p}{a}}\right)dp$$

$$=4a\left[(p+a)\arctan\left(\sqrt{\frac{p}{a}}\right)-\sqrt{ap}\right]_0^a=2a^2(\pi-2).\qquad\square$$

18. Let $\mathbf{F}(x, y, z) = (2x - y, x + 4y, 0)$. Compute the line integral $\int_C \mathbf{F} \cdot d\mathbf{r}$ using Stokes' theorem, when C is a circle of radius 10 centered at the origin

   a) in the plane $z = 0$;

   b) in the plane $x = 0$.

Solution: Since curl $\mathbf{F}(x, y, z) = (0, 0, 2)$, if C is oriented positively, we have

$$\int_C \mathbf{F} \cdot \mathrm{d}\mathbf{r} = \iint_S \mathrm{curl}\, \mathbf{F} \cdot \mathrm{d}\mathbf{A}$$

according to Stokes' Theorem.

a) We select the $xy-$plane region S parameterized by

$$x = r\cos\theta,\ y = r\sin\theta,\ z = 0,$$
$$(s, t) \in D = \{0 \leq r \leq 10,\ 0 \leq \theta \leq 2\pi\}.$$

Thus, $\mathbf{v}_r = (\cos\theta, \sin\theta, 0)$,

$$\mathbf{v}_\theta = (-r\sin\theta, r\cos\theta, 0)$$

and $\mathbf{v}_r \times \mathbf{v}_\theta = (0, 0, r)$. The positive orientation has to be the upwards orientation. Since $\mathbf{v}_r \times \mathbf{v}_\theta$ points upwards when $r \geq 0$,

$$\begin{aligned} I &= \iint_S \mathrm{curl}\, \mathbf{F} \cdot \mathrm{d}\mathbf{A} \\ &= \iint_S (0, 0, 2) \cdot (0, 0, r)\, \mathrm{d}r\, \mathrm{d}\theta \\ &= \iint_D 2r\, \mathrm{d}r\, \mathrm{d}\theta = 2\int_0^{2\pi} \int_0^{10} r\, \mathrm{d}r\, \mathrm{d}\theta = 200\pi. \end{aligned}$$

b) We select the $yz-$plane region S parameterized by

$$x = 0,\ y = r\cos\theta,\ z = r\sin\theta,\ (r, \theta) \in D : 0 \leq r \leq 10,\ 0 \leq \theta \leq 2\pi.$$

Thus, $\mathbf{v}_r = (0, \cos\theta, \sin\theta)$,

$$\mathbf{v}_\theta = (0, -r\sin\theta, r\cos\theta)$$

and $\mathbf{v}_r \times \mathbf{v}_\theta = (r, 0, 0)$. Independently of the orientation of S, we have

$$\begin{aligned} I &= \iint_S \mathrm{curl}\, \mathbf{F} \cdot \mathrm{d}\mathbf{A} \\ &= \iint_D (0, 0, 2) \cdot (r, 0, 0)\, \mathrm{d}r\, \mathrm{d}\theta = \iint_S 0\, \mathrm{d}r\, \mathrm{d}\theta = 0.\square \end{aligned}$$

14.9 Exercises

1. Prepare a 2-page summary of this chapter, with important definitions and results.
2. Translate all of the solved problems of this section (and their solutions) into the language of differential forms.
3. If $\varphi : [a,b] \to [c,d]$ is a $\mathcal{C}^1$ diffeomorphism, show that $\varphi'(t) \neq 0$ for all $t \in [a,b]$.
4. Prove Proposition 189.
5. Flesh out the details in the proof of Green's theorem.
6. For the parametric description of the unit sphere $S \subseteq \mathbb{R}^3$, show that $\text{rank}(D\mathbf{g}(\theta,\varphi)) = 2$ for all (θ,φ).
7. For the parametric description of the cone $S \subseteq \mathbb{R}^3$, show that $\text{rank}(D\mathbf{g}(\varphi,r)) = 2$ for all (φ,r), that $\mathbf{g}$ is injective, and that $\mathbf{g}^{-1}$ is continuous.
8. Complete the calculations of the example on pp. 352-352.
9. Complete the calculations of the example on p. 355.
10. Consider the following classical mathematical results.

    Fundamental Theorem of Calculus: Let $f : [a,b] \to \mathbb{R}$ be R-int and $F : [a,b] \to \mathbb{R}$ be such that F is continuous on $[a,b]$, differentiable on (a,b) and $F'(x) = f(x)$ for all $x \in (a,b)$. Then $\int_a^b f = F(b) - F(a)$.

    Fundamental Theorem of Line Integrals: Let $U \subseteq_O \mathbb{R}^n$, $\phi : U \to \mathbb{R}$ be C^1 and L be a piecewise-C^1 path from A to B in U. Then $\int_L \nabla\phi(\mathbf{r}) \cdot d\mathbf{r} = \phi(B) - \phi(A)$.

    Green's Theorem: Let C be a positively oriented, piecewise smooth, simple closed curve in $\mathbb{R}^2$ and let D be the region bounded by C. If L and M are C^1 on an open region containing D, then

    $$\oint_C (L\,\mathrm{d}x + M\,\mathrm{d}y) = \iint_D \left(\frac{\partial M}{\partial x} - \frac{\partial L}{\partial y}\right)\,\mathrm{d}A.$$

    Classical Stokes' Theorem: Let $S \subseteq \mathbb{R}^3$ be a compact surface with a piecewise-smooth boundary C. If $\mathbf{F} : S \to \mathbb{R}^3$ is C^1, then

    $$\int_S \text{curl}\,\mathbf{F} \cdot \mathrm{d}\mathbf{A} = \oint_C \mathbf{F} \cdot \mathrm{d}\mathbf{r}.$$

    Divergence Theorem: Let $W \subseteq \mathbb{R}^3$ be a compact solid with a piecewise-smooth boundary ∂W. If $\mathbf{F} : W \to \mathbb{R}^3$ is C^1, then

    $$\iiint_W \text{div}\,\mathbf{F}\,\mathrm{d}V = \int_{\partial W} \mathbf{F} \cdot \mathrm{d}\mathbf{A}.$$

    Using the language of differential forms, explain why these five results are special instances of the same result.

Part IV

Topology

Chapter 15

General Topology Concepts

In this chapter, we begin the study of general topology, which extends the concepts of analysis to **general spaces** (on which metrics may not necessarily be definable). We start by presenting the basic concepts and definitions of topology: **open sets**, **bases**, **separation axioms**, **continuity**, and **homeomorphisms**, and we present a few examples of frequently-encountered **topologies**: order, box, subspace, product, and quotient.

15.1 Basic Definitions

Let X be a set. A **topology $\mathfrak{T}$ on X** is a collection of subsets of X.[1] such that

1. $\varnothing, X \in \mathfrak{T}$;
2. if $U_1, \ldots, U_n \in \mathfrak{T}$, then $\bigcap_{i=1}^n U_i \in \mathfrak{T}$;
3. if $\{U_\alpha\}_{\alpha \in \mathcal{A}} \in \mathfrak{T}$, then $\bigcup_{\alpha \in \mathcal{A}} U_\alpha \in \mathfrak{T}$.

The ordered pair $(X, \mathfrak{T})$ is a **topological space**. The sets $U \in \mathfrak{T}$ are called the **open sets of** X. If U is an open set in X containing x, we say that U is a **neighbourhood of x in X**.

Examples: The following collections are topologies on X.

1. $\mathfrak{T} = \wp(X)$ is the **discrete topology** on X.
2. $\mathfrak{T} = \{\varnothing, X\}$ is the **indiscrete topology** on X.
3. If $X = \mathbb{R}$, $\mathfrak{T} = \{A \mid A =$ union of open intervals in $\mathbb{R}\}$ is the **standard topology on $\mathbb{R}$.**

[1] Or a subset $\mathfrak{T}$ of the power set $\wp(X)$.

4. If X is a metric space, $\mathfrak{T} = \{A \mid A \text{ is open in } X \text{ under the metric}\}$ is the **metric topology** on X.

5. $\mathfrak{T} = \{A \mid X \setminus A \text{ is finite}\} \cup \{\varnothing\}$ is the **finite complement topology** on X.

6. $\mathfrak{T} = \{A \mid X \setminus A \text{ is countable}\} \cup \{\varnothing\}$ is the **countable complement topology** on X. $\square$

Let $\mathfrak{T}_1$ and $\mathfrak{T}_2$ be two topologies on a set X. If $\mathfrak{T}_1 \subseteq \mathfrak{T}_2$, then $\mathfrak{T}_2$ is **finer** than $\mathfrak{T}_1$ and $\mathfrak{T}_1$ is **coarser** than $\mathfrak{T}_2$. Obviously, the discrete topology is finer than all other topologies on X.

If $\mathfrak{T}_1 \subsetneq \mathfrak{T}_2$, then $\mathfrak{T}_2$ is **strictly finer** than $\mathfrak{T}_1$ and $\mathfrak{T}_1$ is **strictly coarser** than $\mathfrak{T}_2$. The collection of all topologies on a set X and the inclusion relation form a **poset**, but that will not be that important for us.

A **basis** $\mathfrak{B}$ for a topology is a family of subsets of X such that

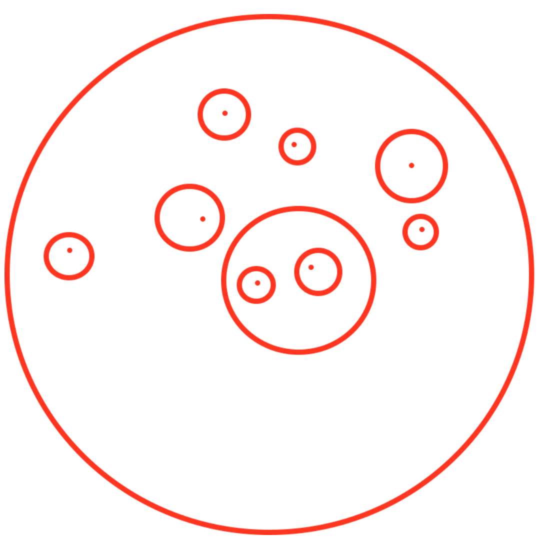
1. if $x \in X$, then there exists $B \in \mathfrak{B}$ such that $x \in B$;[2]

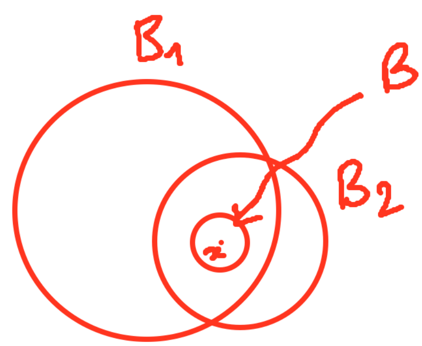
2. if $B_1, B_2 \in \mathfrak{B}$ and $x \in B_1 \cap B_2$, then there exists $B \in \mathfrak{B}$ such that $x \in B \subseteq B_1 \cap B_2$.

The **topology generated by the basis** $\mathfrak{B}$ is

$$\mathfrak{T}(\mathfrak{B}) = \left\{ \bigcup_{B \in \mathfrak{B}'} B \;\middle|\; \mathfrak{B}' \subseteq \mathfrak{B} \right\}.$$

We illustrate conditions 1 (left), 2 (right) for the standard topology on $\mathbb{R}^2$ below.



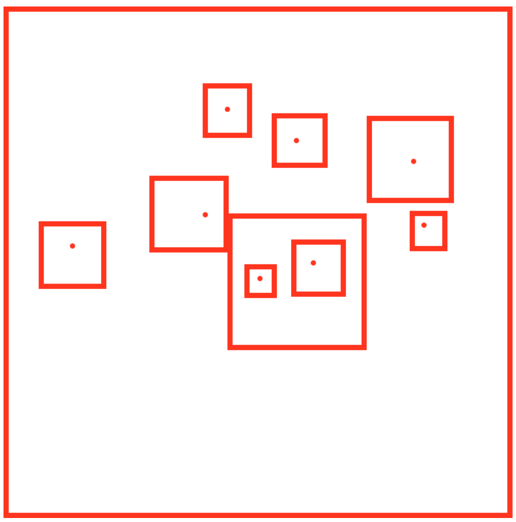


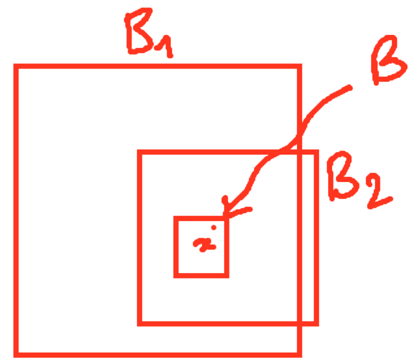
[2] Note that for a given x, the set B need not be unique.

Examples

1. The standard topology on $\mathbb{R}$ has the **open intervals** as a basis.
2. Let $X = \mathbb{R}^2$, $\mathfrak{B}_1$ be the set of all open discs in X, and $\mathfrak{B}_2$ the set of all open squares. Then $\mathfrak{B}_1$ and $\mathfrak{B}_2$ are bases. $\square$

We illustrate conditions 1 (left), 2 (right) for the ℓ_1 topology on $\mathbb{R}^2$.





Theorem 201
Suppose that $\mathfrak{B}_1$ and $\mathfrak{B}_2$ are bases for topologies $\mathfrak{T}_1$ and $\mathfrak{T}_2$, respectively. Then $\mathfrak{T}_1$ is finer than $\mathfrak{T}_2$ if and only if for each $B_2 \in \mathfrak{B}_2$ and any $x \in B_2$, there exists $B_1 \in \mathfrak{B}_1$ such that $x \in B_1 \subseteq B_2$.

Proof: suppose $\mathfrak{T}_1$ is finer than $\mathfrak{T}_2$. Then $B_2 \in \mathfrak{T}_1$ exists $B \in \mathfrak{T}_1$ such that $x \in B \subseteq B_2$. Then, since $\mathfrak{B}_1$ is a basis for $\mathfrak{T}_1$, there exists $B_1 \in \mathfrak{B}_1$ such that $x \in B_1 \subseteq B \subseteq B_2$.

Conversely, let $B \in \mathfrak{B}_2$ and $x \in B$. Then there exists $B_x \in \mathfrak{B}_1$ such that $x \in B_x \subseteq B$, so

$$B = \bigcup_{x \in B} B_x,$$

and $B \in \mathfrak{T}_1$. But any $B_2 \in \mathfrak{T}_2$ is a union of open sets B, so $\mathfrak{T}_2 \subseteq \mathfrak{T}_1$. $\blacksquare$

In the preceding example (second item), it is possible to fit a square inside any circle and vice-versa, and so $\mathfrak{T}(\mathfrak{B}_1) = \mathfrak{T}(\mathfrak{B}_2)$.

A **sub-basis** for a topology on a set X is a collection $\mathfrak{S}$ of subsets of X such that for each $x \in X$, there exists $S \in \mathfrak{S}$ with $x \in \mathfrak{S}$ (note that this means that $X = \bigcup_{S \in \mathfrak{S}} S$).

Examples

1. Let X be a set. Then $\mathfrak{S} = \{x \mid x \in X\}$ is a sub-basis for the discrete topology and $\mathfrak{S}' = \{\varnothing, X\}$ is a sub-basis for the indiscrete topology.

2. Either of the following sets of semi-finite intervals form a sub-basis for the standard topology on $\mathbb{R}$:

$$\mathfrak{S} = \{(a, +\infty) \mid a \in \mathbb{R}\} \cup \{(-\infty, b) \mid b \in \mathbb{R}\}$$
$$\mathfrak{S}' = \{(a, +\infty) \mid a \in \mathbb{R}\}.$$

A basis $\mathfrak{B}$ can be built from a sub-basis $\mathfrak{S}$ by adding to it **all finite intersections** of its elements. Indeed, $B_1, B_2 \in \mathfrak{B} \implies B_1 \cap B_2 \in \mathfrak{B}$ if

$$\mathfrak{B} = \mathfrak{S} \cup \left\{ \bigcap_{i=1}^{n} S_i \,\middle|\, S_i \in \mathfrak{S} \right\}.$$

Example: consider $X = \mathbb{R}$ and $\mathfrak{B} = \{[a, b) \mid a, b \in \mathbb{R}\}$. Then,

$$[a, b) \cap [c, d) = \begin{cases} \varnothing & \text{if } b \leq c \\ [a, b) & \text{if } b \geq c, a \geq c, b \leq d \\ [c, d) & \text{if } b \geq c, a \leq c, b \geq d \\ [c, b) & \text{if } b \geq c, a \leq c, b \leq d \\ [a, d) & \text{if } b \geq c, a \geq c, b \geq d \end{cases}$$

The set $\mathfrak{B}$ is a basis for some topology $\mathfrak{T}'$ on $\mathbb{R}$. We compare $\mathfrak{T}'$ with the standard topology $\mathfrak{T}$ on $\mathbb{R}$ and show that the two topologies are not equal. Suppose $(a, b) \in \mathfrak{T}$. Then, for any $x \in (a, b)$, we get $[x, b) \in \mathfrak{B}$ and $[x, b) \subset (a, b)$. Hence $(a, b) \in \mathfrak{T}'$, and $\mathfrak{T} \subseteq \mathfrak{T}'$, i.e. $\mathfrak{T}'$ is finer than $\mathfrak{T}$.

However, the inclusion is not reversed, which is to say, $[a, b[\notin \mathfrak{T}$. If it were, since $a \in [a, b[$, there would exist (c, d) such that $a \in (c, d) \subseteq [a, b)$, but this is impossible. Thus $\mathfrak{T} \subsetneq \mathfrak{T}'$, i.e. $\mathfrak{T}'$ is strictly finer than $\mathfrak{T}$.

The topology $\mathfrak{T}'$ on $\mathbb{R}$ is the **lower limit topology**, denoted by $\mathbb{R}_l$. □

Let X be a set with a **total order** $\mathcal{R}$. By definition,

1. for every $x, y \in X$, if $x \neq y$, then $x\mathcal{R}y$ or $y\mathcal{R}x$;
2. there is no $x \in X$ such that $x\mathcal{R}x$, and
3. for every $x, y, z \in X$, if $x\mathcal{R}y$ and $y\mathcal{R}z$, then $x\mathcal{R}z$.

We usually write $x < y$ instead of $x\mathcal{R}y$.

It is possible to generalize the concept of an **interval** by writing

$$(a,b) = \{x \in X \mid a < x < b\}, \quad [a,b] = \{x \in X \mid a \leq x \leq b\},$$

and so on.

The **order topology** on X is generated by the basis $\mathfrak{B}$ having as elements intervals of the following forms:

1. (a,b), for $a < b$;
2. $[\perp, b)$, if $\perp$ is a **smallest element** of X ($\perp \leq a$ for all $a \in X$), and
3. $(a, \top]$, if $\top$ is a **greatest element** of X ($\top \geq b$ for all $b \in X$).

Examples

1. The order topology on $\mathbb{R}$ is the standard topology on $\mathbb{R}$, as $\mathbb{R}$ has no lowest or greatest element (all basis elements are of the form (a,b), for $a < b$).
2. In the order topology on $\mathbb{N}$, every point is open as

   $$\{1\} = [1,2) \qquad \text{and } \{n\} = (n-1, n+1) \qquad \text{for } n > 1.$$

   Hence the order topology on $\mathbb{N}$ is the discrete topology on $\mathbb{N}$.
3. Let $X = \{0\} \cup \left\{\frac{1}{n} \mid n \in \mathbb{N}\right\}$. Then

   $$\{1\} = (1/2, 1] \quad \text{and} \quad \left\{\frac{1}{n}\right\} = \left(\frac{1}{n+1}, \frac{1}{n-1}\right) \quad \text{for } n > 1.$$

   But any open set containing 0 will contain a basic set of the form $[0, \frac{1}{N})$, with $\frac{1}{N+1} \in [0, \frac{1}{N})$. Hence $\{0\}$ is not open, and the order topology on X is not discrete. $\square$

15.2 Box and Subspace Topologies

Suppose X and Y are topological spaces. Consider the family of subsets of $X \times Y$ given by

$$\mathfrak{B} = \{U \times V \mid U \subseteq_O X, V \subseteq_O Y\},$$

where $A \subseteq_O X$ stands for $S \in \mathfrak{T}$ ("*A is an open subset of X in the topology on X*").

As $X \subseteq_O X$ and $Y \subseteq_O Y$, we have $X \times Y \in \mathfrak{B}$, and so every element of $X \times Y$ lies in (at least) one element of $\mathfrak{B}$.

Now suppose $U_1 \times V_1, U_2 \times V_2 \in \mathfrak{B}$. As $U_1 \cap U_2 \subseteq_O X$ and $V_1 \cap V_2 \subseteq_O Y$, we have

$$(U_1 \times V_1) \cap (U_2 \times V_2) = (U_1 \cap U_2) \times (V_1 \cap V_2) \in \mathfrak{B}.$$

This means that $\mathfrak{B}$ is a basis for a topology on $X \times Y$, which we call the **box product topology on** $X \times Y$.

Two mappings come with this topology:

$$\pi_1 \mid X \times Y \to X \quad \text{and} \quad \pi_2 \mid X \times Y \to Y,$$

defined by $\pi_1(x, y) = x$ and $\pi_2(x, y) = y$. These mappings are called the **projections onto the first and second coordinates**; we have

$$U \times V = (U \times Y) \cap (X \times V) = \pi_1^{-1}(U) \cap \pi_2^{-1}(V),$$

where

$$\pi_1^{-1}(U) = \{(x, y) : \pi_1(x, y) \in U\} \quad \text{and} \quad \pi_2^{-1}(V) = \{(x, y) : \pi_2(x, y) \in V\}.$$

The set $\mathfrak{S} = \{\pi_1^{-1}(U) \mid U \subseteq_O X\} \cup \{\pi_2^{-1}(V) \mid V \subseteq_O Y\}$ is thus a sub-basis of the box product topology on $X \times Y$.

> **Example:** if $X = Y = \mathbb{R}$, the box product topology on $\mathbb{R}^2$ is the standard topology on $\mathbb{R}^2$ (and is also the same as the ℓ_1 and ℓ_2 topologies on $\mathbb{R}^2$). $\square$

Suppose $Y \subseteq X$, where X is a topological space. For each $V \subseteq_O X$, we define $U = V \cap Y$ to be an open set in Y. This creates a topology on Y.

1. $\varnothing, Y \subseteq_O Y$ since $\varnothing = \varnothing \cap Y$ and $Y = X \cap Y$, and $\varnothing, X \subseteq_O X$.
2. Suppose $U_\alpha \subseteq_O Y$. Then $\exists V_\alpha \subseteq_O X$ such that $U_\alpha = V_\alpha \cap Y$. But

$$\bigcup_\alpha V_\alpha \subseteq_O X \quad \text{and} \quad \bigcup_\alpha U_\alpha = \left(\bigcup_\alpha V_\alpha\right) \cap Y \Longrightarrow \bigcup_\alpha U_\alpha \subseteq_O Y.$$

3. Suppose $U_i \subseteq_O Y$, for $1 \leq i \leq n$. Then $\exists V_i \subseteq_O X$ such that $U_i = V_i \cap Y$, for $1 \leq i \leq n$. But

$$\bigcap_{i=1}^n V_i \subseteq_O X \quad \text{and} \quad \bigcap_{i=1}^n U_i = \left(\bigcap_{i=1}^n V_i\right) \cap Y \Longrightarrow \bigcap_{i=1}^n U_i \subseteq_O Y.$$

This topology on Y is called the **subspace topology on** Y **relative to** X. The open sets in Y are called **relatively open**; they are not always open in X.

Theorem 202
Suppose Y is a subspace of X and $\mathfrak{B}$ is a basis for the topology on X. Then $\mathfrak{B}_Y = \{U \cap Y \mid U \in \mathfrak{B}\}$ is a basis for the subspace topology.

Proof: let $V = U \cap Y$ and suppose $y \in V$ and $U \subseteq_O X$. Let $B \in \mathfrak{B}$ such that $y \in B \subseteq U$. Hence $y \in B_Y = B \cap Y \subseteq U \cap Y$, and so $\mathfrak{B}_Y$ is a basis for the subspace topology on Y. ■

Some examples will help to solidify the concepts.

Examples

1. Let $X = \mathbb{R}$ and $Y = \mathbb{Q}$. A basic open set of Y is a set of the form $B = (a, b) \cap \mathbb{Q}$, where $a, b \in \mathbb{R}$. Note that B contains no interval of real numbers. Hence, no open set of $\mathbb{Q}$ can be open in $\mathbb{R}$.

2. Let $X = \mathbb{R}$ and $Y = [0, 1]$. A basic open set of Y is a set of the form $B = (a, b) \cap [0, 1]$, where $a, b \in \mathbb{R}$. If $0 \leq a < b \leq 1$, the relatively open sets of Y will be open in $\mathbb{R}$. The basic sets in Y are the sets of the form $[0, b)$, $(a, 1]$, and (a, b), and the subspace topology on Y is the order topology.

3. Let $X = \mathbb{R}$ and $Y = \{-1\} \cup \left\{\frac{1}{n}\right\}_{n\in\mathbb{N}}$. In this case, the subspace topology is discrete. Indeed,
$$\{-1\}=(-3/2, -1/2) \cap Y, \qquad \left\{\frac{1}{n}\right\}=\left(\frac{1}{n+1/2}, \frac{1}{n-1/2}\right) \cap Y.$$

4. Let $X = \mathbb{R}$ and $Y = \{0\} \cup \left\{\frac{1}{n}\right\}_{n\in\mathbb{N}}$. In this case, the subspace topology is not discrete. Indeed, while
$$\left\{\frac{1}{n}\right\} = \left(\frac{1}{n+1/2}, \frac{1}{n-1/2}\right) \cap Y,$$
we have $\{0\} \neq (a, b) \cap Y$ for all $a < b \in X$. □

15.3 Dual Definitions and Separation Axioms

It is possible to define all the notions of topology in terms of **closed** sets, instead of open sets. Let X be a set. A **topology** $\mathfrak{T}$ **on** X is a collection of subsets of X such that

1. $\varnothing, X \in \mathfrak{T}$;

2. if $C_1, \ldots, C_n \in \mathfrak{T}$, then $\bigcup_{i=1}^n C_i \in \mathfrak{T}$;

3. if $\{C_\alpha\}_{\alpha\in\mathcal{A}} \in \mathfrak{T}$, then $\bigcap_{\alpha\in\mathcal{A}} C_\alpha \in \mathfrak{T}$.

The ordered pair $(X, \mathfrak{T})$ is a **topological space.** The sets $C \in \mathfrak{T}$ are called the **closed sets of** X. In general, a set V is closed in X, denoted by $V \subseteq_C X$, if and only if its **complement is open in** X.

Using this definition, it is easy to prove the following propositions.

Proposition 203
Let Y be a subspace of X. A set A is closed in Y if and only if it is the intersection of a closed set in X with Y.

Proof: left as an exercise. □

Proposition 204
Let Y be a subspace of X. If A is closed in Y and Y is closed in X, then A is closed in X.

Proof: left as an exercise. □

Again, let's take a look at some examples.

Examples

1. Let $X = \mathbb{R}$. Then $[a, b]$ is **closed** in $\mathbb{R}$ for all $a < b$.
2. Let $X = \mathbb{R}$. The set $[0, 1[$ is **neither open nor closed** in $\mathbb{R}$ with the standard topology.
3. If X has the discrete topology, then every set is **both open and closed**, since every set is the union of open singletons, and the complement of every set is also the union of open singletons.
4. Let $X = \{a, b, c, d\}$ be a set with 4 distinct elements. Define a topology on X by

$$\mathfrak{T} = \{\varnothing, \{a, b\}, \{c, d\}, X\}.$$

   All sets which are open **are also closed, and vice-versa**; the topology is not discrete as $\{b, c\}$ **is neither open nor closed**. □

The **closure** of a set A in X is the smallest closed set containing A, usually denoted by $\overline{A}$. Obviously, $A \subseteq \overline{A}$. By definition, we have

$$\overline{A} = \bigcap_{A \subseteq C \subseteq_C X} C.$$

If $A \subseteq_C X$, then $A = \overline{A}$, as $\overline{A} \subseteq A$. Thus, A is closed if and only if $A = \overline{A}$.

Similarly, the **interior** of a set A in X is the largest open set contained in A, usually denoted by A°. Obviously, $A^\circ \subseteq A$. By definition, we also have

$$A^\circ = \bigcup_{V \subseteq A, V \subseteq_O X} V.$$

If $A \subseteq_O X$, then $A = A^\circ$, as $A^\circ \subseteq A$. Thus A is open if and only if $A = A^\circ$.

Examples

1. The closure of $(0,1)$ in $\mathbb{R}$ is $[0,1]$.
2. Let $X = \mathbb{R}$ and $A = \mathbb{Q}$. Then $A^\circ = \varnothing$ and $\overline{A} = \mathbb{R}$. □

The result from the last example follows from Theorem 206.

Theorem 205
Let A be a subset of X. Then $x \in \overline{A}$ if and only if every neighbourhood V of x has a non-empty intersection with A.

Proof: we show that $x \notin \overline{A}$ if and only if there is a neighbourhood V of x such that $A \cap V = \varnothing$. Suppose $x \notin \overline{A}$. Then there is a closed set C containing A with $x \notin C$. Let $V = X \backslash C \subseteq_O X$. Then $x \in V$ and $A \cap V \subseteq C \cap V = \varnothing$, so $A \cap V = \varnothing$.

Conversely, suppose there is a neighbourhood V of x such that $A \cap V = \varnothing$. Let $C = X \setminus V \subseteq_C X$. Then $A \subseteq C$ and $\overline{A} \subseteq C$, as C is closed. But $V \cap C = \varnothing$, so $x \notin C$ and thus $x \notin \overline{A}$. ■

Let A be a subset of X. A point $a \in X$ is a **limit point** of A if every neighbourhood of a contains a point of A different from a, i.e. $a \in \overline{A \setminus \{a\}}$.

Examples

1. Let $X = \mathbb{R}$ and $A = \{\frac{n}{n+1} \mid n \in \mathbb{N}\}$. Then $\{1\}$ is a limit point of A, and $\overline{A} = A \cup \{1\}$, according to Theorem 206.
2. Let X be a set with the indiscrete topology. For any non-empty subset A of X and any point $a \in X$, a is a limit point of A as long as $A \neq \{a\}$. For instance, Let $X = \{a, b\}$ with topology $\mathfrak{T} = \{\varnothing, X\}$. If $A = \{b\}$, then a is a limit point of A. Indeed, the only neighbourhood of a is X, and $A \cap X = \{b\} \neq \varnothing$.

We've alluded to it a few times already, so now it's time for Theorem 206.

> **Theorem 206**
> If A' is the set of all limit points of A, then $\overline{A} = A \cup A'$.
>
> **Proof:** if $x \in A \cup A'$, then $x \in A$ or $x \in A'$. In the first case, $x \in A \subseteq \overline{A}$. In the other, every neighbourhood of x contains a point of A. Thus $x \in \overline{A}$.
>
> Conversely, suppose $x \in \overline{A}$. Either $x \in A$ or $x \notin A$. It is sufficient to show that if $x \notin A$, then $x \in A'$. If $x \notin A$, every neighbourhood of x meets A in at least one point other than x. But $x \notin A$, so $x \in A'$. ■

We have the following corollary.

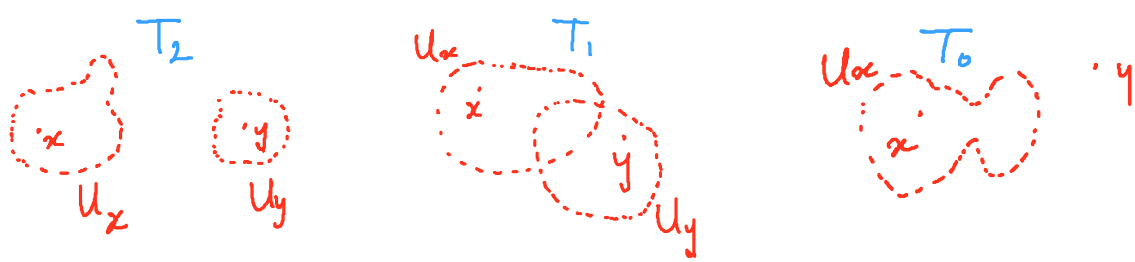
> **Corollary 207**
> *A is closed in X if and only if $A' \subseteq A$.*
>
> **Proof:** left as an exercise.

To avoid degenerate situations like the one found in the preceding example (which is to say, that any point could be the limit point of all non-singleton subsets in the indiscrete topology), we introduce the notion of **separation axioms**.

A space X is:

1. T_2 or **Hausdorff** if for every pair $x \neq y \in X$, there exist disjoint neighbourhoods U_x of x and U_y of y;
2. T_1 if for every pair $x \neq y \in X$, there exist neighbourhoods U_x of x and U_y of y such that $y \notin U_x$ and $x \notin U_y$;
3. T_0 if for every pair $x \neq y \in X$, there exist a neighbourhood U of either x or y that misses the other.[3]

Note that every T_2 space is T_1, and every T_1 space is T_0, but that there are T_0 spaces that are not T_1, and T_1 spaces that are not T_2; the conditions are illustrated below.



[3]Other separation axioms will be discussed at a later stage. In their studies, many topologists are only interested in spaces that are at least Hausdorff.

Theorem 208
If X is Hausdorff and $x \in X$ is a limit point of $A \subseteq X$, then every neighbourhood of x contains infinitely many points of A.

Proof: let x be a limit point of A and V be a neighbourhood of x. Since X is a T_2 space, its singletons are closed sets. Indeed, let $x \in X$. For all $y \neq x \in X$, there exist neighbourhoods U_x of x and U_y of y such that $x \notin U_y$ and $y \notin U_x$ (the T_1 condition holds for T_2 spaces). Then

$$X \setminus \{x\} = \bigcup_{y \in Y} U_y$$

is open in X and $\{x\}$ is closed; if x has a neighbourhood V such that $A \cap V$ is finite,

$$A \cap V = \{a_1, \ldots, a_n\}$$

must be closed, being the finite union of closed sets.

Let $W = V \setminus (A \cap V)$. If $x \in W$, then W is a neighbourhood of x such that $W \cap A = \varnothing$, which contradicts x being a limit point of A. Hence $x \in A \cap V$. After reordering if necessary, suppose $x = a_1$. Then

$$W_1 = V \setminus \{a_2, \ldots, a_n\}$$

is a neighbourhood of x such that $W_1 \cap A = \{a_1\} = \{x\}$, so that x cannot be a limit point of A. By *reductio ad absurdum*, $A \cap V$ is infinite. ■

Hausdorff spaces are particularly well-behaved with respect to toplogies.

Theorem 209
Every simply ordered set is T_2 in the order topology. The product of two T_2 spaces is T_2. A subspace of a T_2 space is T_2.

Proof: left as an exercise. ■

15.4 Continuity and Homeomorphisms

Suppose that X and Y are topological spaces. A function $f : X \to Y$ is **continuous** if $f^{-1}(V)$ is open in X whenever V is open in Y.[4]

Theorem 210
Let $f : X \to Y$. If $\mathfrak{B}$ is a basis for the topology of Y, then f is continuous if and only if $f^{-1}(B) \subseteq_O X$ for every $B \in \mathfrak{B}$.

[4]Similarly, if $\mathfrak{S}$ is a sub-basis for Y, then f is continuous if and only if $f^{-1}(S) \subseteq_O X$ for all $S \in \mathfrak{S}$.

Proof: if f is continuous, $f^{-1}(B) \subseteq_O X$ for all $B \in \mathfrak{B}$ since such $B \subseteq_O Y$. Conversely, suppose $f^{-1}(B)$ is open for all $B \in \mathfrak{B}$. Let $V = \bigcup_{i \in I} B_i$ be an open subset of Y. Then

$$f^{-1}(V) = f^{-1}\left(\bigcup_{i \in I} B_i\right) = \bigcup_{i \in I} f^{-1}(B_i)$$

is open in X since all $f^{-1}(B_i)$ is open in X for all $i \in I$. ■

Continuous functions are to topology what linear maps are to linear algebra.

Examples

1. If X and Y are metric spaces and $f : X \to Y$ is continuous with respect to the metrics in the usual sense, it is continuous in the topological sense.

2. For a product space $X \times Y$, the projections π_1, π_2 are continuous. Indeed, $\pi_1^{-1}(U) = U \times Y, \pi_2^{-1}(V) = X \times V \subseteq_O X \times Y$ when $U \subseteq_0 X, V \subseteq_O Y$.

3. For each $b \in Y$, the **inclusion map** $i_b : X \to X \times Y$ defined by $i_b(x) = (x, b)$ is continuous. Indeed, let $U \times V$ be a basic neighbourhood in $X \times Y$. Then

$$i_b^{-1}(U \times V) = \begin{cases} \varnothing, & b \notin V, \\ U, & b \in V, \end{cases}$$

   which is open in X. Thus the inclusion map is continuous.

4. For any X, the identity map id $: X \to X$ is continuous when X has the same topology as a domain as it has as a range.

5. The function id $: \mathbb{R} \to \mathbb{R}_l$ is not continuous. Indeed, let $[a, b)$ be an open set in $\mathbb{R}_l$. Then $\text{id}^{-1}([a, b)) = [a, b)$ is not open in $\mathbb{R}$, so id is not continuous. The function id $: \mathbb{R}_l \to \mathbb{R}$ is continuous, however. Let (a, b) be a basic open set in $\mathbb{R}$. Then $\text{id}^{-1}(a, b) = (a, b) = \bigcup_{n \in \mathbb{N}} [a + 1/n, b)$ is open in $\mathbb{R}_l$, so id is continuous.

6. Let $f : X \to Y$ and $g : Y \to Z$ be continuous functions. Then $g \circ f : X \to Z$ is a continuous function. Indeed, let $U \subseteq_O Z$. Then $V = g^{-1}(U) \subseteq_O Y$ since g is continuous, and $f^{-1}(V) \subseteq_O X$ as f is continuous. Then

$$(g \circ f)^{-1}(U) = f^{-1}(g^{-1}(U)) = f^{-1}(V)$$

   is open in X and $g \circ f$ is continuous. □

There are other ways to verify if a function is continuous.

Theorem 211
Let $f : X \to Y$. The following statements are equivalent:

1. *f is continuous;*
2. *for any $A \subseteq X$, $f(\overline{A}) \subseteq \overline{f(A)}$;*
3. *if C is closed in Y, then $f^{-1}(C)$ is closed in X.*

Proof:

1. $\implies 2.$: If $x \in \overline{A}$, then every neighbourhood of x contains a point of A. If V is a neighbourhood of $f(x)$ then $f^{-1}(V)$ is open in X and $x \in f^{-1}(V)$. As x is a limit point of A, there exists $a \in A$ with $a \in f^{-1}(V)$ and $f(a) \in V$, so $f(a) \in V \cap f(A)$. But this just means that $f(x)$ is a limit point of $f(A)$, so $f(x) \in \overline{f(A)}$, that is $f(\overline{A}) \subseteq \overline{f(A)}$.

2. $\implies 3.$: If C is closed in Y, then $C = \overline{C}$. Let $A = f^{-1}(C)$ then $A \subseteq \overline{A}$ and

$$f(\overline{A}) = \overline{f(A)} = \overline{f(f^{-1}(C))} \subseteq \overline{C} = C.$$

Then $\overline{A} \subseteq f^{-1}(C)$ so $f^{-1}(C)$ is closed.

3. $\implies 1.$: If $f^{-1}(C)$ is closed whenever C is closed, then if V is open in Y, $Y \setminus V$ is closed in Y, so $f^{-1}(Y \setminus V)$ is closed in X. But

$$f^{-1}(Y \setminus V) = f^{-1}(Y) \setminus f^{-1}(V) = X \setminus f^{-1}(V),$$

so $f^{-1}(V)$ is open. Hence f is continuous. ■

A **homeomorphism** $f : X \to Y$ is a bijection for which both f and the inverse function $g : Y \to X$ are continuous. We say that X and Y are **homeomorphic** when there is a homeomorphism $f : X \to Y$. [5]

Examples

1. Let $X = \mathbb{R}$, $Y = (0, \infty)$. The function $f : X \to Y$, defined by $f(x) = e^x$ is continuous. The inverse function $g : Y \to X$ defined by $g(y) = \ln y$ is also continuous. Both these functions are bijections, so $\mathbb{R}$ and $(0, \infty)$ are homeomorphic in the standard topology.

[5] Homeomorphisms play the same role for topological spaces as isomorphisms play for groups. Consequently, homeomorphism of spaces is an equivalence relation on the 'set' of topological spaces. Homeomorphic spaces are **identical** from the point of view of topology.

2. The bijections $\tan : (-\pi/2, \pi/2) \to \mathbb{R}$ and $\arctan : \mathbb{R} \to (-\pi/2, \pi/2)$ are both continuous, so $\mathbb{R}$ is homeomorphic to $(-\pi/2, \pi/2)$.

3. The continuous bijections $f : (a, b) \to (c, d)$ and $g : (c, d) \to (a, b)$,
$$f(x) = c + \frac{d-c}{b-a}(x-a) \quad \text{and} \quad g(y) = a + \frac{b-a}{d-c}(y-c),$$
are inverses of one another, so (a, b) is homeomorphic to (c, d). $\square$

The continuous function $f : X \to Y$ is an **embedding** of X into Y if the map $g : X \to f(X)$ defined by $g(x) = f(x)$ is a homeomorphism when $f(X)$ has the subspace topology.

Examples

1. For $b \in Y$, the **inclusion map** $i_b : X \to X \times Y$, $x \mapsto (x, b)$, is an embedding.

2. Let $A \subseteq X$. The inclusion map $\iota : A \to X$, $a \mapsto a$, is an embedding. $\square$

Continuous functions enjoy a whole slew of attractive properties.

Theorem 212
Let X, Y, Z be top. spaces, and $V_\alpha \subseteq_O X$, $A_i \subseteq_C X$.

1. *Constant functions are continuous.*

2. *The inclusion function $\iota : A \subseteq X \to X$ is continuous.*

3. *If $f : X \to Y$ is continuous, then the restriction function $f|_A$ for all subsets $A \subseteq X$ is continuous.*

4. *If $f : X \to Y$ is continuous, then $f : X \to Z$ is continuous, assuming that $f(X) \subseteq Z$ and either $Z \subseteq Y$ or $Y \subseteq Z$.*

5. *If $X = \bigcup V_\alpha$ and the restriction $f|_{V_\alpha} : V_\alpha \to Y$ is continuous for each α, then $f : X \to Y$ is continuous.*

6. *If $X = \bigcup_{i=1}^n A_i$ and the restriction $f|_{A_i} : A_i \to Y$ is continuous for each $1 \leq i \leq n$, then $f : X \to Y$ is continuous.*

Proof: left as an exercise. $\blacksquare$

As a special case of Theorem 212, we get the following result.

Lemma 213 (Pasting Lemma)
Suppose $X = A \cup B$ where A and B are closed sets. If $f : A \to Y$ and $g : B \to Y$ are such that $f(x) = g(x)$ for all $x \in A \cap B$, then the function $h : X \to Y$ defined by

$$h(x) = \begin{cases} f(x), & \text{if } x \in A\,, \\ g(x), & \text{if } x \in B \end{cases}$$

is continuous. The same holds if A and B are both open.

Proof: left as an exercise. ■

Lemma 213 is extremely useful.

Examples

1. If $X = Y = \mathbb{R}$, let $A = [0, \infty)$ and $B = (-\infty, 0]$, and define $f : A \to Y$ by $f(x) = x$ and $g : B \to Y$ by $g(x) = -x$. Then $h(x) = |x|$ is continuous by Lemma 213.
2. Instead, take $B = (-\infty, 0)$ and define $f : A \to Y$ by $f(x) = x + 1$ and $g : B \to Y$ by $g(x) = x$. The function h obtained by Lemma 213 construction is not continuous as $h^{-1}(1/2, 3/2) = [0, 1/2)$. □

This last example shows that Lemma 213 does not hold if A and B are not both closed, or open.

Theorem 214
Let $f : X \to Y \times Z$. Then f is continuous if and only if the functions $\pi_1 f$ and $\pi_2 f$ are continuous.

Proof: if f is continuous then $\pi_1 f$ and $\pi_2 f$ are continuous since the projections are continuous. Conversely, suppose $\pi_1 f$ and $\pi_2 f$ are continuous. If $U \times V$ is a basic open set in $Y \times Z$, then

$$f^{-1}(U \times V) = (\pi_1 f)^{-1}(U) \cap (\pi_2 f)^{-1}(V),$$

which is open as $\pi_1 f$ and $\pi_2 f$ are continuous. Hence f is continuous. ■

The following local formulation of continuity is sometimes useful in applications. A function $f : X \to Y$ is **locally continuous** at $x \in X$ if for any open set V with $f(x) \in V$, there is a neighbourhood U of x such that $f(U) \subseteq V$. A function $f : X \to Y$ is thus continuous if and only if it is locally continuous at every point of X, as can easily be verified.

15.5 Product Topology

Suppose $\{X_\alpha\}_{\alpha\in A}$ is a family of topological spaces, where A is an arbitrary **indexing set**.[6] Then

$$X = \prod_{\alpha\in A} X_\alpha$$

is the set of all maps $x : A \to \bigcup_{\alpha\in A} X_\alpha$ such that $x(\alpha) \in X_\alpha, \forall\alpha \in A$. We write x_α for $x(\alpha)$ and $x = (x_\alpha)_{\alpha\in A}$. This set X comes equipped with **projection mappings** π_α for each $\alpha \in A$, defined by $\pi_\alpha(x) = x_\alpha$ for all $x \in X$.

We can endow X with a topology by extending the box product topology to **arbitrary products**. A basic open set in this **box topology** is a set of the form $\prod_\alpha U_\alpha$, where $U_\alpha \subseteq_O X_\alpha$ for each $\alpha \in A$.

Alternatively, extending the topology obtained by the sub-basis

$$\mathfrak{S} = \bigcup_{\alpha\in A} \{\pi_\alpha^{-1}(V_\alpha) \mid V_\alpha \subseteq_O X_\alpha\}$$

to arbitrary products yields a topology called the **product topology** on X. The basic open sets in the product topology have the form $\prod_\alpha U_\alpha$, where $U_\alpha = X_\alpha$ **except in a finite number of cases** $U_{\alpha_i} \subseteq_O X_{\alpha_i}$, for $1 \leq i \leq n$.

Note that when A is finite, the box and product topologies coincide. Furthermore, the basic open sets in the product topology are open in the box topology, and so **the box topology is finer than the product topology**. But this inclusion is strict. For instance, $(-1,1)^\omega$ is open in $\mathbb{R}^\omega = \prod_{n\in\mathbb{N}} \mathbb{R}$ with the box topology, but it is not open in $\mathbb{R}^\omega$ with the product topology as this would imply that $\mathbb{R} \subseteq (-1,1)$.

Theorem 215
If $\mathfrak{B}_\alpha$ is a basis for the topology on X_α, then

$$\mathfrak{B} = \left\{ \prod_{\alpha\in A} B_\alpha \,\middle|\, B_\alpha \in \mathfrak{B}_\alpha \right\}$$

is a basis for $\prod_\alpha X_\alpha$ in the box topology.

Proof: left as an exercise. ■

Theorem 216
In both the box and product topologies, the product of subspaces is a subspace and the product of Hausdorff spaces is Hausdorff.

Proof: left as an exercise. ■

[6] In particular, A may not be countable.

While the definition of the box topology might seem the more natural of the two generalizations to infinite products, there is at least one way in which the product topology is superior (and hence, preferable).

Theorem 217
Let $f : Y \to X = \prod_\alpha X_\alpha$ and and $f_\alpha = \pi_\alpha f$ for all α. When X is endowed with the product topology, f is continuous if and only if f_α is continuous for all α.

Proof: suppose f is continuous. The projections π_α are continuous. Indeed, pick α. Let V_α be a basic open of X_α. Then $W = \pi_\alpha^{-1}(V_\alpha) = \prod_\beta W_\beta$, where $W_\alpha = V_\alpha$ and $W_\beta = X_\beta$. But W is open in the product topology, so π_α is continuous. Thus, $f_\alpha = \pi_\alpha f$ is continuous for each α, being the composition of two continuous functions.

Conversely, suppose that f_α is continuous for all α. Let $\pi_\alpha^{-1}(U_\alpha)$ be a sub-basic subset of X. As $f_\alpha = \pi_\alpha f$ is continuous,

$$f^{-1}\left(\pi_\alpha^{-1}(U_\alpha)\right) = f_\alpha^{-1}(U_\alpha)$$

is open in Y, which is to say that f is continuous. ■

This result need not be true in the box topology.

Example: consider the function $f : \mathbb{R} \to \mathbb{R}^\omega$, defined by $f_n(x) = nx$ for all $x \in \mathbb{R}$. Each f_n is continuous on $\mathbb{R}$, and $f(x) = (nx)_{n\in\mathbb{N}}$. In the box topology, $(-1,1)^\omega \subseteq_O \mathbb{R}^\omega$. But $f_n^{-1}(-1,1) = (-1/n, 1/n)$ and $f^{-1}\left((-1,1)^\omega\right) = \{0\}$, which is not open in $\mathbb{R}$. Hence f is not continuous in the box topology. □

15.6 Quotient Topology

Let X be a topological space and $f : X \to Y$ be a **surjective** mapping. We make f continuous by defining a topology on Y through

$$V \subseteq_O Y \iff f^{-1}(V) \subseteq_O X.$$

That this defines a topology is clear:

1. $\varnothing \subseteq_O Y$ as $f^{-1}(\varnothing) = \varnothing \subseteq_O X$; $Y \subseteq_O Y$ as $f^{-1}(Y) = X \subseteq_O X$ since f is surjective.
2. If $U, V \subseteq_O Y$, then $f^{-1}(U), f^{-1}(V) \subseteq_O X$. But
$$f^{-1}(U \cap V) = f^{-1}(U) \cap f^{-1}(V) \subseteq_O X \Longrightarrow \text{so } U \cap V \subseteq_O Y.$$
3. If $U_\alpha \subseteq_O Y$ for all α, then $f^{-1}(U_\alpha) \subseteq_O X$ for all α. But
$$f^{-1}\left(\bigcup_\alpha U_\alpha\right) = \bigcup_\alpha f^{-1}(U_\alpha) \subseteq_O X \Longrightarrow \text{so } \bigcup U_\alpha \subseteq_O Y.$$

This is the **quotient topology on** Y, and $f : X \to Y$ is a **quotient map**. Thus, any continuous map $f : X \to Y$ is a quotient map whenever it is a surjective. Note that a quotient map need not be open.

> **Example:** let $X = [0,2]$ have the subspace topology from $\mathbb{R}$, and set $Y = \{a,b\}$ where $\mathfrak{T}_Y = \{\varnothing, \{a\}, Y\}$, and define $f : X \to Y$ by
>
> $$f(x) = \begin{cases} a, & \text{if } 0 \leq x < 1, \\ b, & \text{if } 1 \leq x \leq 2. \end{cases}$$
>
> As $f^{-1}(\{a\}) = [0,1) \subseteq_O X$, f is continuous and a quotient map (as it is also surjective). However, $f(1,2) = \{b\}$ is not open in Y, so f is not open. □

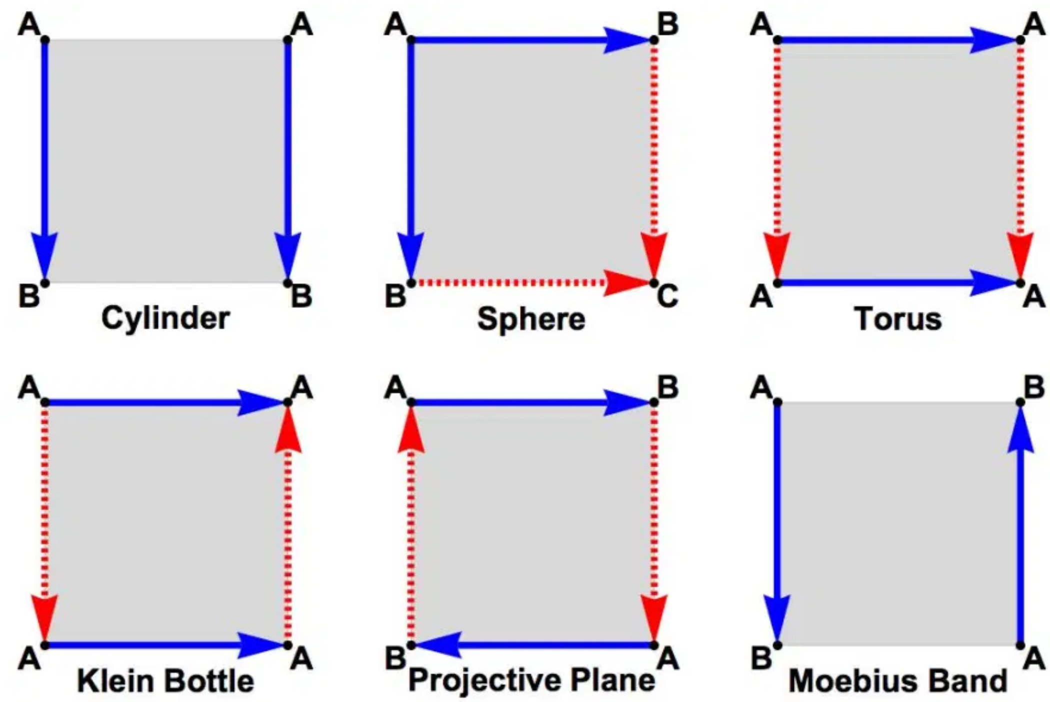
If $f : X \to Y$ is a quotient map, we define an **equivalence relation on** X by

$$x_1 \sim x_2 \Longleftrightarrow f(x_1) = f(x_2).$$

Equivalence classes of $X/\sim$ are in $1-$to-1 correspondence with elements of Y; $X/\sim$ and Y are homeomorphic under the **identification topology**.

> **Examples:** in what follows, we set $X = I \times I$, where $I = [0,1]$, with the usual subspace topology from $\mathbb{R}^2$.
>
> 1. The **cylinder** is defined via the following equivalence relation on X:
>
> $$(x,y) \sim (x,y') \Longleftrightarrow y - y' \in \mathbb{Z}^2.$$
>
> 2. The **torus** is defined via the following equivalence relation on X:
>
> $$(x,y) \sim (x',y') \Longleftrightarrow (x - x', y - y') \in \mathbb{Z}^2.$$
>
> 3. The **Möbius band** is defined via the following equivalence relation on X:
>
> $$(x,y) \sim (x',y') \Longleftrightarrow x - x' \in \mathbb{Z} \text{ and } y + y' = 1.$$
>
> 4. The **Klein bottle** is defined via the following equivalence relation on X:
>
> $$(x,y) \sim (x',y') \Longleftrightarrow (x - x' \in \mathbb{Z} \text{ and } y + y' = 1) \text{ or } (x = x' \text{ and } y - y' \in \mathbb{Z}).$$
>
> 5. The **projective plane** is defined via the following equivalence relation on X:
>
> $$(x,y) \sim (x',y') \Longleftrightarrow (x - x' \in \mathbb{Z} \text{ and } y + y' = 1) \text{ or } (x + x' = 1 \text{ and } y - y' \in \mathbb{Z}).$$

The identification topologies of those spaces on X are shown below (from `thatsmaths.com`).



15.7 Solved Problems

1. Show that if $\mathfrak{B}$ is a basis for a topology on X, then the topology generated by $\mathfrak{B}$ is the intersection of all topologies on X that contain $\mathfrak{B}$. Prove the same if $\mathfrak{B}$ is a sub-basis.

Proof: let $\mathfrak{B}$ be a basis, and suppose $\mathfrak{T}(\mathfrak{B})$ is the topology on X generated by $\mathfrak{B}$. We first show that $\mathfrak{T}(\mathfrak{B}) \subseteq \bigcap_{\mathfrak{B}\subseteq\mathfrak{T}} \mathfrak{T}$.

Let $U \in \mathfrak{T}(\mathfrak{B})$. Then $U = \bigcup_{B\in\mathfrak{B}_U} B$, for some $\mathfrak{B}_U \subseteq \mathfrak{B}$. Let $\mathfrak{T}$ be any topology on X containing $\mathfrak{B}$. In particular, it also contains $\mathfrak{B}_U$, and

$$\bigcup_{B\in\mathfrak{B}_U} B = U \in \mathfrak{T},$$

since arbitrary unions of open sets in $\mathfrak{T}$ are open in $\mathfrak{T}$. But $\mathfrak{T}$ was arbitrary, so $U \in \bigcap_{\mathfrak{B}\subseteq\mathfrak{T}} \mathfrak{T}$, and $\mathfrak{T}(\mathfrak{B}) \subseteq \bigcap_{\mathfrak{B}\subseteq\mathfrak{T}} \mathfrak{T}$. Conversely, since $\mathfrak{T}(\mathfrak{B})$ is a topology on X containing $\mathfrak{B}$, then

$$\bigcap_{\mathfrak{B}\subseteq\mathfrak{T}} \mathfrak{T} \subseteq \mathfrak{T}(\mathfrak{B}).$$

Hence $\bigcap_{\mathfrak{B}\subseteq\mathfrak{T}} \mathfrak{T} = \mathfrak{T}(\mathfrak{B})$.

Now suppose $\mathfrak{B}$ is a sub-basis. The proof follows the same lines. The sole difference is that the topology on X generated by $\mathfrak{B}$ is

$$\mathfrak{T}(\mathfrak{B}) = \left\{ \bigcup_{\text{arbitrary}} \left(\bigcap_{\text{finite}} B_i \right) \middle| B_i \in \mathfrak{B} \right\}.$$

So we need only to verify that if $U \in \mathfrak{T}(\mathfrak{B})$, then $U \in \bigcap_{\mathfrak{B} \subseteq \mathfrak{T}} \mathfrak{T}$. Let

$$U = \bigcup_{\text{arb.}} \left(\bigcap_{\text{fin.}} B_i \right)$$

and $\mathfrak{T}$ be any topology on X containing $\mathfrak{B}$. Then $U \in \mathfrak{T}$ since arbitrary unions and finite intersections of open sets in $\mathfrak{T}$ are open in $\mathfrak{T}$.

The rest of the proof is identical to the above proof for when $\mathfrak{B}$ is a basis. ■

2. Show that the collection

$$\mathfrak{B} = \{[a,b) \mid a < b, a, b \in \mathbb{Q}\}$$

is a basis that generates a topology different from that of $\mathbb{R}_l$.

Proof: to show that $\mathfrak{B}$ is a basis, it suffices to show the second property, since $\mathbb{R} = \bigcup_{n \in \mathbb{N}} [-n, n)$. Let $[a,b)$ and $[c,d)$ belong to $\mathfrak{B}_2$. Then

$$[a,b) \cap [c,d) = \begin{cases} \varnothing & \text{if } b \leq c \\ [a,b) & \text{if } b \geq c, a \geq c, b \leq d \\ [c,d) & \text{if } b \geq c, a \leq c, b \geq d \\ [c,b) & \text{if } b \geq c, a \leq c, b \leq d \\ [a,d) & \text{if } b \geq c, a \geq c, b \geq d \end{cases},$$

where $a, b, c, d \in \mathbb{Q}$.

Thus, whenever $x \in [a,b) \cap [c,d)$, there exists an interval $I \in \mathfrak{B}$ such that $x \in I \subseteq [a,b) \cap [c,d)$. Hence B is a basis. Denote the topology on $\mathbb{R}$ generated by $\mathfrak{B}$ by $\mathfrak{T}$, and that of the lower limit topology on $\mathbb{R}$ by $\mathfrak{T}_l$. Clerly, $[\pi, 4) \in \mathfrak{T}_l$. Does it also belong to $\mathfrak{T}$?

If it does, we can then write

$$[\pi, 4) = \bigcup_{\alpha \in \mathcal{A}} [a_\alpha, b_\alpha),$$

for $a_\alpha, b_\alpha \in \mathbb{Q}$. But notice that each of the a_α must be greater than π. In particular, since $\pi \notin \mathbb{Q}$, each of the a_α must be strictly greater than π, since they are all rational. Hence, we can at best obtain

$$(\pi, 4) = \bigcup_{\alpha \in \mathcal{A}} [a_\alpha, b_\alpha),$$

if $a_\alpha, b_\alpha \in \mathbb{Q}$. Hence $[\pi, 4) \notin \mathfrak{T}$ and $\mathfrak{T}_l \neq \mathfrak{T}$. ■

3. Show that if Y is a subspace of X, and A is a subset of Y, then the subspace topology on A as a subspace of Y is the same as the subspace topology on A as a subspace of X.

Proof: let U be open in the subspace topology on A as a subspace of X, and V be open in the subspace topology on A as a subspace of Y.

Then, there exists $W \subseteq_O X$ and $Z \subseteq_O Y$ such that $U = A \cap W$ and $V = A \cap Z$. But if $Z \subseteq_O Y$, there exist $Z' \subseteq_O X$ such that $Z = Y \cap Z'$, and so $V = A \cap Y \cap Z'$.

Since $A \subseteq Y$,

$$U = A \cap W = A \cap Y \cap W, \quad V = A \cap Y \cap Z' = A \cap Z',$$

where W and Z' are open in X.

Hence U is open in the subspace topology on A as a subspace of Y, and V is open in the subspace topology on A as a subspace of X, and so the two topologies are equal. ■

4. If $\mathfrak{T}$ and $\mathfrak{T}'$ are topologies on X and $\mathfrak{T}'$ is strictly finer than $\mathfrak{T}$, what can you say about the corresponding subspace topologies on the subset Y of X?

Solution: let $\mathfrak{T}_Y$ and $\mathfrak{T}'_Y$ be the subspaces topologies on a subset Y of X corresponding to $\mathfrak{T}$ and $\mathfrak{T}'$ respectively. It should be clear that $\mathfrak{T}'_Y$ is finer than $\mathfrak{T}_Y$. Indeed let $B = V \cap Y$ for some $V \in \mathfrak{T} \subsetneq \mathfrak{T}'$. Hence $B = V \cap Y$ for some $V \in \mathfrak{T}'$.

Can we necessarily say that $\mathfrak{T}'_Y$ is strictly finer than $\mathfrak{T}$? Well, suppose all $U \in \mathfrak{T}'$ where $U \notin \mathfrak{T}$ are such that $U \cap Y = \varnothing$.[7] Then

$$A = Y \cap U = Y \cap \varnothing \in \mathfrak{T}_Y$$

since $\varnothing$ is open in $\mathfrak{T}$.

For all other $V \in \mathfrak{T}'$, we have $V \in \mathfrak{T}$, and so we have $A = V \cap Y \in \mathfrak{T}$. Hence, in this case $\mathfrak{T}_Y = \mathfrak{T}'_Y$. The following example shows that $\mathfrak{T}'_Y$ could be strictly finer than $\mathfrak{T}_Y$.

Let $X = \mathbb{R}$ (as a set), $Y = (0,1)$ and suppose $\mathfrak{T}$ and $\mathfrak{T}'$ are the usual topology on $\mathbb{R}$ and the lower limit topology on $\mathbb{R}$, respectively.

Then $[0.5,1) \in \mathfrak{T}'_Y$, but it is not open in the usual subspace topology on Y since there is no interval (a,b) such that

$$[0.5,1) = (0,1) \cap (a,b).$$

In this case, $\mathfrak{T}'_Y$ is strictly finer than $\mathfrak{T}_Y$. Thus, the most we can say without more information is that $\mathfrak{T}'_Y$ is finer than $\mathfrak{T}_Y$. □

[7]For instance, let $X = \{a,b,c\}$, $Y = \{c\}$, $\mathfrak{T} = \{\varnothing, X\}$ and $\mathfrak{T}' = \{\varnothing, \{a,b\}, X\}$. Then $\mathfrak{T} \subsetneq \mathfrak{T}'$, and the only $U \in \mathfrak{T}'$ where $U \notin \mathfrak{T}$ is $U = \{a,b\}$, so $Y \cap U = \varnothing$.

5. Show that the projections $\pi_1 : X \times Y \to X$ and $\pi_2 : X \times Y \to Y$ are open maps.

   Proof: we show that π_1 is open, the proof that π_2 is open is similar. Let B be a basic open set in $X \times Y$. Hence $B = U \times V$, where U is open in X and V is open in Y. Then $\pi_1(B) = U$ is open in X. Now, any open W in $X \times Y$ is written

   $$W = \bigcup_{\alpha \in \mathcal{A}} (U_\alpha \times V_\alpha),$$

   where $U_\alpha \times V_\alpha$ is a basic open set for all $\alpha \in \mathcal{A}$. Now

   $$\begin{aligned} \pi_1(W) &= \pi_1(\{(u,v) \in X \times Y | (u,v) \in U_\alpha \times V_\alpha \text{ for some } \alpha \in A\} \\ &= \{u \in X | u \in U_\alpha \text{ for some } \alpha \in A\} = \cup_{\alpha \in \mathcal{A}} U_\alpha, \end{aligned}$$

   which is open in X, since it is an arbitrary union of open sets in X, so π_1 is open. ■

6. Show that X is Hausdorff if and only if the diagonal

   $$\Delta = \{(x,x) | x \in X\}$$

   is closed in $X \times X$.

   Proof: since $\varnothing$ and any one point set are vacuously Hausdorff, and since their respective Δ are $\varnothing$ and X, which are closed sets in X, the result holds when $X = \varnothing$ and $X = \{*\}$. We can thus restrict ourselves to spaces X with at least two elements. For any such X, $X \times X \setminus \Delta \neq \varnothing$.

   Suppose X is Hausdorff. We show that $X \times X \setminus \Delta$ is open in $X \times X$, and so that Δ is closed in $X \times X$.

   Let $(x,y) \in X \times X \setminus \Delta$. Then $x \neq y$. So there exists two sets U_x, V_y (open in X) such that $x \in U_x$, $y \in V_y$ and $U_x \cap V_y = \varnothing$. Now $(x,y) \in U_x \times V_y$, which is open in $X \times X$. We show that $(U_x \times V_y) \cap \Delta = \varnothing$. Suppose

   $$(z,z) \in (U_x \times V_y) \cap \Delta \neq \varnothing.$$

   Then $z \in U_x$ and $z \in V_y$, so $z \in U_x \cap V_y$. But $U_x \cap V_y = \varnothing$, so there is no such (z,z). Hence, we can fit an open set around each $(x,y) \in X \times X \setminus \Delta$, and so $X \times X \setminus \Delta$ is open in $X \times X$.

   Conversely, suppose Δ is closed in $X \times X$, and let $x, y \in X$ such that $x \neq y$. Then $(x,y) \in X \times X \setminus \Delta$, an open set of $X \times X$. Hence there exists a basic open set $U \times V$ of $X \times X$ such that

   $$(x,y) \in U \times V \subseteq X \times X \setminus \Delta.$$

   But $U \cap V = \varnothing$, otherwise there would exist $z \in X$ such that

   $$(z,z) \in U \times V \nsubseteq X \times X \setminus \Delta.$$

   Thus U, V are open subsets of X with $x \in U$, $y \in V$, and $U \cap V = \varnothing$, and so X is Hausdorff. ■

7. Let $A \subseteq X$, and let $f : A \to Y$ be continuous; let Y be Hausdorff. Show that if f may be extended to a continuous function $g : \overline{A} \to Y$, then g is uniquely determined.

 Proof: suppose f can be extended to g and h, as in the statement of the problem. Suppose $g \neq h$. Then, there exists $x_0 \in \overline{A} \setminus A = \partial A$ such that $g(x_0) \neq h(x_0)$, since $f = g|_A = h|_A$.

 But Y is Hausdorff, so $\exists U, V \subseteq_O Y$ such that $g(x_0) \in U$, $h(x_0) \in V$, and $U \cap V = \varnothing$.

 Since g and h are continuous, $g^{-1}(U), h^{-1}(V) \subseteq_O X$. Furthermore,

 $$x_0 \in g^{-1}(U) \cap h^{-1}(V) \subseteq_O X.$$

 As $x_0 \in \overline{A}$, there exists $a \neq x_0$ in A such that $a \in g^{-1}(U) \cap h^{-1}(V)$, and so $g(a) \in U$ and $h(a) \in V$. But $g(a) = h(a) = f(a)$ since $a \in A$, which yields $f(a) \in U \cap V$, a contradiction, as this set is supposed empty. Thus when f can be extended, it can be done uniquely. ■

8. If $f_1 : X_1 \to Y_1$, $f_2 : X_2 \to Y_2$ are continuous, show that $F : X_1 \times X_2 \to Y_1 \times Y_2$ is continous, where $F(x_1, x_2) = (f_1(x_1), f_2(x_2))$.

 Proof: the set $\mathfrak{B} = \{U \times V \mid U \subseteq_O Y_1, V \subseteq_O Y_2\}$ is a basis for the product topology on $Y_1 \times Y_2$. Then, it is enough to show that $F^{-1}(U \times V) \subseteq_O X_1 \times X_2$ for all $U \times V \in \mathfrak{B}$. But

 $$\begin{aligned} F^{-1}(U \times V) &= \{(x_1, x_2) \in X_1 \times X_2 \mid F(x_1, x_2) \in U \times V\} \\ &= \{(x_1, x_2) \in X_1 \times X_2 \mid f_1(x_1) \in U, f_2(x_2) \in V\} \\ &= \{(x_1, x_2) \in X_1 \times X_2 \mid x_1 \in f_1^{-1}(U), x_2 \in f_2^{-1}(V)\} = f_1^{-1}(U) \times f_2^{-1}(V). \end{aligned}$$

 But $f_1^{-1}(U) \subseteq_O X_1$ and $f_2^{-1}(V) \subseteq_O X_2$ since f_1 and f_2 are continuous, and so

 $$F^{-1}(U \times V) = f_1^{-1}(U) \times f_2^{-1}(V) \subseteq_O X_1 \times X_2$$

 in the product topology, which means that F is continuous. ■

9. Let $f : X \to Y$ be an onto mapping. For each of the properties T_1 and T_2, prove or disprove that if one of X, Y has the property, then so must the other when

 a) f is continuous;
 b) f is open;
 c) f is both open and continuous.

 Solution: throughout, we assume that both X and Y have at least two elements – otherwise, all the statements are vacuously or trivially true. Recall that a space W is T_1 when, for each pair of distinct points $x, y \in W$, there exists U_x, V_y open sets in W such that $x \in U_x \not\ni y$ and $y \in V_y \not\ni x$.

a) f is continuous: $f^{-1}(U)$ is open in X whenever U is open in Y.

   i. $\underline{X \text{ is } T_1}$. Let $X = \mathbb{R}$, $Y = \{a, b\}$ with the indiscrete topology and define the surjection $f : \mathbb{R} \to Y$ by $f(0) = a$ and $f(x) = b$ for all $x \neq 0$. Then $\mathbb{R}$ is T_1, since it is T_2, and f is continuous, since $f^{-1}(Y) = \mathbb{R}$ is open in $\mathbb{R}$, but Y is not T_1 since every neighbourhood of a contains b. So X is $T_1 \not\Rightarrow Y$ is T_1.

   ii. $\underline{Y \text{ is } T_1}$. Let $X = \{a, b, c, d\}$ with $\mathfrak{T}_X = \{\varnothing, \{a, c\}, \{b, d\}, X\}$, $Y = \{a, b\}$ with the discrete topology and define the surjection $f : X \to Y$ by $f(a) = a$, $f(b) = b$, $f(c) = a$ and $f(d) = b$. Then f is continuous, since both $f^{-1}(\{a\}) = \{a, c\}$, $f^{-1}(\{b\}) = \{b, d\}$ lie in $\mathfrak{T}_X$, but X is not T_1 since every neighbourhood of a contains c. So Y is $T_1 \not\Rightarrow X$ is T_1.

   iii. $\underline{X \text{ is } T_2}$. In the counter-example a)i., X is also T_2, but Y is not T_1, so it is certainly not T_2. Hence X is $T_2 \not\Rightarrow Y$ is T_2.

   iv. $\underline{Y \text{ is } T_2}$. In the counter-example a)ii., Y is also T_2, but X is not T_1, so it is certainly not T_2. Hence Y is $T_2 \not\Rightarrow X$ is T_2.

b) f is open: $f(V)$ is open in Y whenever V is open in X.

   i. $\underline{X \text{ is } T_1}$. See b)iii. X is $T_1 \not\Rightarrow Y$ is T_1.

   ii. $\underline{Y \text{ is } T_1}$. In the counter-example a)ii., f is surjective, it is open since Y has the discrete topology, and Y is T_1. But X is not T_1. So Y is $T_1 \not\Rightarrow X$ is T_1.

   iii. $\underline{X \text{ is } T_2}$. Let $X = \mathbb{R}$, $Y = \{a, b\}$ with the indiscrete topology, and define the surjection $f : \mathbb{R} \to Y$ by $f(x) = a$ whenever $x \in \mathbb{Q}$ and $f(x) = b$ whenever $x \notin \mathbb{Q}$. Then f is open. Indeed, any basic open set (a, b) contains both rational and irrational numbers, and so $f(a, b) = Y \subseteq_O Y$. Note that $\mathbb{R}$ is T_2, but Y is not T_2, as it is not even T_1. Thus, X is $T_2 \not\Rightarrow Y$ is T_2.

   iv. $\underline{Y \text{ is } T_2}$. In the counter-example a)ii., f is surjective, it is open since Y has the discrete topology, and Y is T_2. But X is not T_2, as it is not T_1. Thus, Y is $T_2 \not\Rightarrow X$ is T_2.

c) f is both open and continuous: $f^{-1}(U)$ is open in X whenever U is open in Y and $f(V)$ is open in Y whenever V is open in X.

   i. $\underline{X \text{ is } T_1}$. See b)iii. X is $T_1 \not\Rightarrow Y$ is T_1.

   ii. $\underline{Y \text{ is } T_1}$. In the counter-example a)ii., f is surjective, it is open since Y has the discrete topology, it is continuous by definition and Y is T_1. But X is not T_1. Hence Y is $T_1 \not\Rightarrow X$ is T_1.

   iii. $\underline{X \text{ is } T_2}$. In the counter-example b)iii., the function f is also continuous since Y has the indiscrete topology and $X = \mathfrak{T}_2$. But Y is not T_2 as it is not even T_1. Hence X is $T_2 \not\Rightarrow Y$ is T_2.

   iv. $\underline{Y \text{ is } T_2}$. In the counter-example a)ii., f is surjective, it is open since Y has the discrete topology, it is continuous by definition and Y is T_2. But X is not T_2. As it is not even T_1. Hence Y is $T_2 \not\Rightarrow X$ is T_2.

And that's it, folks: T_1 and T_2 do not behave nicely with respect to continuous functions. □

10. Show that the set A of all bounded sequences is both open and closed in the box topology on $\mathbb{R}^\omega$.

Proof: let A be the set

$$A = \{(x_n)_{n\in\mathbb{N}} \mid \exists M \in \mathbb{R} \text{ with } |x_n| < M \ \forall n \in \mathbb{N}\}.$$

We start by showing that $A \subseteq_O \mathbb{R}^\omega$.

Let $(x_n)_{n\in\mathbb{N}} \in A$. Then $\exists M \in \mathbb{R}$ such that $|x_n| < M$ for all $n \in \mathbb{N}$. Set

$$U_n = (x_n - 1, x_n + 1), \text{ for all } n \in \mathbb{N}.$$

Then

$$U = \prod_{n\in\mathbb{N}} (x_n - 1, x_n + 1)$$

is open in the box topology on $\mathbb{R}^\omega$, since $U_n \subseteq_O \mathbb{R}$ for all $n \in \mathbb{N}$. Clearly, $(x_n)_{n\in\mathbb{N}} \in U$, since $x_n \in U_n$ for all $n \in \mathbb{N}$. But $U \subseteq A$.

Indeed, suppose $(w_n)_{n\in\mathbb{N}} \in U$. Then $w_n \in U_n$ for all $n \in \mathbb{N}$ and so $x_n - 1 < w_n < x_n + 1$ for all $n \in \mathbb{N}$. But this means that

$$-M - 1 < x_n - 1 < w_n < x_n + 1 < M + 1$$

and $|w_n| < M + 1$ for all $n \in \mathbb{N}$. Hence $(w_n)_{n\in\mathbb{N}} \in A$ and $U \subseteq A$ so we conclude that $A \subseteq_O \mathbb{R}^\omega$.

We now show that $A \subseteq_C \mathbb{R}^\omega$. Suppose $(x_n)_{n\in\mathbb{N}} \in \overline{A}$, and let

$$V = \prod_{n\in\mathbb{N}} \left(x_n - \frac{1}{n}, x_n + \frac{1}{n}\right).$$

Then $(x_n)_{n\in\mathbb{N}} \in V \subseteq_O \mathbb{R}^\omega$ and there exists $(a_n)_{n\in\mathbb{N}} \in A$ such that $(a_n)_{n\in\mathbb{N}} \in V$. In that case, there exists $M \in \mathbb{R}$ such that $-M < a_n < M$ for all $n \in \mathbb{N}$. However, $a_n \in (x_n - \frac{1}{n}, x_n + \frac{1}{n})$ so that

$$a_n - \frac{1}{n} < x_n < a_n + \frac{1}{n}$$

for all $n \in \mathbb{N}$, and so

$$-M - 1 < a_n - \frac{1}{n} < x_n < a_n + \frac{1}{n} < M + 1$$

and $|x_n| < M + 1$ for all $n \in \mathbb{N}$.

Thus $(x_n)_{n\in\mathbb{N}} \in A$, and $\overline{A} \subseteq A$, which yields $A \subseteq_C \mathbb{R}^\omega$. ■

15.8 Exercises

1. Prepare a 2-page summary of this chapter, with important definitions and results.
2. Let Z be a subspace of Y. Are the subspace topologies on Z relative to X and Y the same?
3. Show directly that the box product topology on $\mathbb{R}^2$ is identical to the ℓ_1 and ℓ_2 topologies on $\mathbb{R}^2$.
4. Provide a proof of Results 203, 204, 207, 209, 212, 213, 215, and 216.
5. Show that a function which is locally continuous at every point is continuous, and *vice-versa*.
6. Provide the details for the homeomorphism examples of pp. 381-382.
7. Provide the details for the embedding examples of p. 382.
8. Provide the equivalence relation for the identification topology of the cylinder, the sphere, and the projective plane.
9. Show that the map $f : X \to Y$ is continuous if and only if $f(\overline{A}) \subseteq \overline{f(A)}$ for any subset A of X.
10. Let $f, g : X \to Y$ be continuous maps from a space X to a Hausdorff space Y. Prove that the set $C = \{x \mid f(x) = g(x)\}$ is closed in X.
11. Suppose that $f : X \to Y$ is a bijection. If $\mathfrak{B}$ is a basis for the topology on X, prove that f is a homeomorphism if and only if the collection $\{f(B) \mid B \in \mathfrak{B}\}$ is a basis for the topology on Y.
12. Show that the map $f : X \to Y$ is continuous if and only if $f(\overline{A}) \subseteq \overline{f(A)}$ for any subset A of X.
13. Let $f, g : X \to Y$ be continuous maps from a space X to a Hausdorff space Y. Prove that the set $C = \{x \mid f(x) = g(x)\}$ is closed in X.
14. Suppose that $f : X \to Y$ is a bijection. If $\mathfrak{B}$ is a basis for the topology on X, prove that f is a homeomorphism if and only if the collection $\{f(B) \mid B \in \mathfrak{B}\}$ is a basis for the topology on Y.

Chapter 16

Connected Spaces

In Chapter 9, we discussed **connectedness** and **path-connectedness** in the context of metric spaces. In this chapter, we discuss how these notions extend to **general topological spaces**.

16.1 Connected Sets

A **separation** of a space X is a pair of **disjoint** non-empty open sets U and V such that $X = U \cup V$. Note that both U and V are open and closed. When no separation of X exists, we say that X is **connected**. Alternatively, X is connected if the only sets that are closed and open in X are $\varnothing$ and X.

> **Example:** let $X = [1,2] \cup [3,4]$ be a subspace of $\mathbb{R}$. $U = [1,2]$ is closed in X as $U = X \cap [1,2]$ and $[1,2]$ is closed in $\mathbb{R}$. But $U = X \cap (0.5, 3.5)$, so $U \subseteq_O X$. Consequently, X is not connected. $\square$

In general, a subspace $Y \subseteq X$ is connected if it is connected in the **subspace topology**.

> **Theorem 218**
> *A separation of a subset Y is a pair of non-empty subsets A and B whose union is Y and such that $\overline{A} \cap B = \varnothing$ and $A \cap \overline{B} = \varnothing$.*
>
> **Proof:** Suppose A and B satisfy the conditions of the theorem. Then
>
> $$\overline{A} \cap Y = \overline{A} \cap (A \cup B) = (\overline{A} \cap A) \cup (\overline{A} \cap B) = A \cup \varnothing = A,$$
>
> and A is closed in the subspace topology on Y (i.e., relatively closed). Similarly, B is relatively closed, so A and B are relatively open in Y. Consequently, A and B form a separation of Y.

Conversely, suppose A and B are a separation of Y. Then A is relatively closed and so $A = \overline{A} \cap Y$. Then

$$\overline{A} \cap B = \overline{A} \cap Y \cap B = A \cap B = \varnothing.$$

Similarly $A \cap \overline{B} = \varnothing$. ■

If $Y \subseteq X$ is a connected set, and U and V is a separation of X, then $Y \subseteq U$ or $Y \subseteq V$.[1]

Theorem 219
If $\{C_\alpha\}_{\alpha \in A}$ is a family of connected sets such that $\bigcap_\alpha C_\alpha \neq \varnothing$, then $\bigcup_\alpha C_\alpha$ is connected.

Proof: suppose $x \in \bigcap_\alpha C_\alpha$. If U and V is a separation of $\bigcup_\alpha C_\alpha$, then either $x \in U$ or $x \in V$. Without loss of generality, let $x \in U$. Let $\alpha \in A$. Since C_α is connected, either $C_\alpha \subseteq U$ or $C_\alpha \subseteq V$. But $x \in C_\alpha$, so $C_\alpha \subseteq U$. Then $\bigcup_\alpha C_\alpha \subseteq U$. Hence

$$\left(\bigcup_\alpha C_\alpha\right) \cap V \subseteq U \cap V = \varnothing.$$

As $\bigcup_\alpha C_\alpha = U \cup V$, this means that $V = \varnothing$, which is a contradiction since U and V form a separation. Consequently, there could be no such separation to start with, and $\bigcup_\alpha C_\alpha$ is connected. ■

Connectedness behaves well with respect to the closure of a set, as we can see below.

Theorem 220
If A is connected, and $A \subseteq B \subseteq \overline{A}$, then B is connected.

Proof: if U and V forms a separation of B, then $A \subseteq U$ or $A \subseteq V$. Without loss of generality, suppose $A \subseteq U$. Then $V = B \cap V \subseteq \overline{A} \cap V \subseteq \overline{U} \cap V = \varnothing$, by Theorem 218. But $V \neq \varnothing$ as U and V form a separation of B. Thus there cannot be a separation of B and B is connected. ■

As mentioned in Chapter 9, connectedness is a **topological property**.

Theorem 221
Let $f : X \to Y$ be a continuous function. If X is connected, $f(X)$ is connected.

Proof: suppose that $f(X)$ is not connected. Let U and V form a separation of $f(X)$. Then $f^{-1}(U)$ and $f^{-1}(V)$ form a separation of X and X is not connected. ■

[1] Otherwise $U \cap Y$ and $V \cap Y$ would form a separation of Y.

Theorem 222 If X and Y are connected spaces, so is $X \times Y$.

Proof: if $x \in X$, the function $i_x : Y \to X \times Y$ defined by $i_x(y) = (x, y)$ is continuous. Then $i_x(Y) = \{x\} \times Y$ is connected. Similarly, $i_y(X) = X \times \{y\}$ is connected for all $y \in Y$. Then

$$i_x(Y) \cap i_y(X) = \{(x, y)\} \neq \varnothing$$

for all $y \in Y$. Then $C_y = i_y(X) \cup i_x(Y) = (X \times \{y\}) \cup (\{x\} \times Y)$ is connected for all $y \in Y$. Now

$$\bigcap_{y \in Y} C_y = \{x\} \times Y = i_x(Y) \neq \varnothing,$$

so $\bigcup_{y \in Y} C_y = X \times Y$ is connected. ■

As a result, any finite product of connected sets is connected. What about an infinite product of connected sets?

Theorem 223
Let $\{X_\alpha\}_{\alpha \in A}$ be a collection of connected sets. Then $\prod_\alpha X_\alpha$ is connected in the product topology.

Proof: if $\prod_\alpha X_\alpha = \varnothing$, then the theorem is trivially true, so let $b = (b_\alpha)_\alpha \in \prod_\alpha X_\alpha$. For each finite set $\{\alpha_1, \ldots, \alpha_n\}$ of A, consider the space

$$X(\alpha_1, \ldots, \alpha_n) = \left\{ (x_\alpha)_{\alpha \in A} \,\middle|\, \begin{matrix} x_\alpha = b_\alpha \text{ if } \alpha \notin \{\alpha_1, \ldots, \alpha_n\} \\ x_\alpha \in X_\alpha \text{ if } \alpha \in \{\alpha_1, \ldots, \alpha_n\} \end{matrix} \right\},$$

which is homeomorphic to $X_{\alpha_1} \times \cdots \times X_{\alpha_n}$, and so connected. Let $\mathfrak{B}$ be the collection of all finite subsets of A. Note that $b \in X(\alpha_1, \ldots, \alpha_n)$ for all $\{\alpha_1, \ldots, \alpha_n\} \in \mathfrak{B}$, hence

$$b \in \bigcap_{\{\alpha_1, \ldots, \alpha_n\} \in \mathfrak{B}} X(\alpha_1, \ldots, \alpha_n) \neq \varnothing.$$

Thus $Y = \bigcup_{\mathfrak{B}} X(\alpha_1, \ldots, \alpha_n)$ is connected. We show that $\overline{Y} = \prod_\alpha X_\alpha$. Since Y is connected, $\overline{Y}$ is connected and the theorem is proven.

Let $x = (x_\alpha)_\alpha \in \prod_\alpha X_\alpha \neq b$ and let V be a basic neighbourhood of x. Then $V = \prod_\alpha V_\alpha$, where $V_\alpha = X_\alpha$ for all but a finite number of open sets V_{α_i}, $1 \leq i \leq n$.

Define $y = (y_\alpha)_\alpha$ by

$$y_\alpha = \begin{cases} b_\alpha, & \text{if } \alpha \neq \alpha_i \text{ for all } 1 \leq i \leq n, \\ x_\alpha, & \text{if } \alpha = \alpha_i \text{ for some } 1 \leq i \leq n. \end{cases}$$

Then, $y_\alpha = b_\alpha \in V_\alpha = X_\alpha$ for $\alpha \notin \{\alpha_1, \ldots, \alpha_n\}$, and $y_\alpha = x_\alpha \in V_\alpha$ for $\alpha \in \{\alpha_1, \ldots, \alpha_n\}$. Hence $y_\alpha \in V_\alpha$ for all α and $y \in V$.

But, by construction, $y \in X(\alpha_1, \ldots, \alpha_n) \subseteq Y$, so $y \in V \cap Y \neq \varnothing$. As $b \neq x$, we get $y \neq x$, and x is a limit point of Y, that is $x \in \overline{Y}$. Consequently, $\overline{Y} = \prod_\alpha X_\alpha$. ■

In the usual topology, $\mathbb{R}$ has some useful properties, some of which can be extended to general spaces. A **linear continuum**. for isntance, is an ordered set X in which the following hold:

i. if $x < y \in X$, there exists $z \in X$ such that $x < z < y$;

ii. any non-empty set $A \subset X$ with an upper bound has a least upper bound.

A rather tedious, but not very difficult, argument ([Munkres, , p.153] shows that linear continua are connected, and that rays and intervals are connected subsets in a linear continuum. As $\mathbb{R}$ is a linear continuum, it is connected. The next result is a generalization of a very important theorem from analysis (see Theorem 35, Chapter 3).

Theorem 224 (Intermediate Value Theorem)
Suppose $f : X \to Y$ is continuous and Y has the order topology for some ordering $<$. If X is connected and $a, b \in X$ are such that $f(a) < f(b)$, then for any $y \in Y$ such that $f(a) < y < f(b)$, there exists $x \in X$ such that $f(x) = y$.

Proof: let $A = \{z \in Y : z > y\}$ and $B = \{z \in Y : z < y\}$. Then $A, B \subseteq_O Y$, and, as f is continuous, $f^{-1}(A), f^{-1}(B) \subseteq_O X$. Furthermore, $f^{-1}(A) \cap f^{-1}(B) = \varnothing$, $a \in f^{-1}(B)$ and $b \in f^{-1}(A)$. Since X is connected, $X \neq f^{-1}(A) \cup f^{-1}(B)$ (otherwise, $f^{-1}(A)$ and $f^{-1}(B)$ would form a separation of X).

Hence, there exists $x \in X \setminus (f^{-1}(A) \cup f^{-1}(B))$. As $f(x) \notin A$ and $f(x) \notin B$, $f(x) = y$. ■

If $x \in X$, the (connected) **component** of x in X, denoted C_x is the union of all connected sets containing x. It is connected as the intersection of all these sets contain x. As C_x is connected, $\overline{C_x}$ is connected and so $\overline{C_x} \subseteq C_x$. Then the component C_x is closed in X; if X has a finite number of components, each component is also open.

We can define an equivalence relation on X as follows: xRy if and only if there is a connected set containing both x and y.

Then:

1. for all $x \in X$, xRx;
2. if xRy, then yRx, and
3. if xRy and yRx, then xRz.

The **equivalence class** of x is simply the (connected) component of x in X.

> **Examples** (Components)
>
> 1. Let $X = [1,2) \cup (3,4)$ be a subspace of X. Then X has two components, $[1,2)$ and $(3,4)$.
> 2. Let $x \in \mathbb{Q}$. Then the component of x is $\{x\}$ as the only connected subsets of $\mathbb{Q}$ are one-point sets. When all the components of X are singletons, we say that the space X is **totally disconnected**. □

16.2 Path-Connectedness

A **path** in a space X is a continuous map $p : [0,1] \to X$. Throughout, we denote $[0,1]$ by I. If $p(0) = a$ and $p(1) = b$, we say that p is a **path from** a **to** b, a is the **initial point** of p, while b is the **terminal point** of p. A space X is **path-connected** if for any pair of points $a, b \in X$, there is a path p from a to b.

> **Proposition 225**
> *A path-connected space X is connected.*
>
> **Proof:** Suppose A, B were a separation of X. Let $a \in A$ and $b \in B$. As X is path-connected, there is a path p from a to b. But $p(I)$ is connected in X as I is connected, so $p(I) \subseteq A$ or $p(I) \subseteq B$. But $p(0) \in A$ and $p(1) \in B$, a contradiction. Hence X is connected. ■

We have already discussed paths in Chapter 14.

> **Examples** (Paths and path-connectedness)
>
> 1. Let $a \in X$. The map $p_a : I \to X$ defined by $p_a(t) = a$ is a path, the **constant path at** a.
> 2. For $n > 1$, $\mathbb{R}^n \setminus \{0\}$ is path-connected. Let $a, b \in \mathbb{R}^n \setminus \{0\}$. Define $S_{a,b}$ to be the circle with diameter $\overline{ab}$. If $0 \notin S_{a,b}$, then either of the semi-circles form a path from a to b in $\mathbb{R}^n \setminus \{0\}$. If $0 \in S_{a,b}$, it can only lie on one of the semi-circles. Then the other semi-circle gives the desired path.

3. Any **convex** subset C of $\mathbb{R}^n$ is connected. Indeed, let $a, b \in C$ and define a path $p : I \to X$ by

$$p(t) = (1-t)a + tb = t(b-a) + a.$$

Then p is continuous, $p(0) = a$ and $p(1) = b - a + a = b$. Hence C is path-connected, so connected.

4. $\mathbb{R} \setminus \{0\}$ is not connected, as $(-\infty, 0)$, $(0, \infty)$ is a separation. Let $n > 1$. Then $\mathbb{R}^n \setminus \{0\}$ and $\mathbb{R} \setminus \{0\}$ are not homeomorphic. But this actually means that $\mathbb{R}^n$ is not homeomorphic to $\mathbb{R}$. Suppose $f : \mathbb{R}^n \to \mathbb{R}$ was a homeomorphism. Then $f(\mathbb{R}^n \setminus \{0\}) = \mathbb{R} \setminus \{f(0)\}$ would be the continuous image of a connected set, so should be connected. But it clearly isn't, so there can be no homeomorphism.

5. Let $A = \{(x, y) \mid x = ny, n \in \mathbb{N}, 0 \leq x \leq 1\}$. Graphically, A represents the union of lines through the origin of slopes 1, $\frac{1}{2}$, $\frac{1}{3}$, $\ldots$, restricted to $I \times I$. A is connected, as it is clearly path-connected. Let $X = A \cup \{(1, 0)\}$. Then X is connected since $A \subseteq X \subseteq \overline{A}$. We show that X is not path-connected by showing that there is no path in X from $b = (1, 0)$ to any point of A. As a result, connected spaces need not be path connected.

   Suppose $p : I \to X$ is a path with $p(0) = b$ and let V be a neighbourhood of b, excluding $(0, 0)$. Let $t_0 \in p^{-1}(b)$. As p is continuous, there exists a basic (hence connected) neighbourhood U of t_0 such that $p(U) \subseteq V$. If $t_1 \in U$ and $p(t_1) \neq b$, then $p(t_1)$ lies on $x = ny$ for some $n \in \mathbb{N}$. Write

$$W_1 = \left\{(x, y) : x < \left(n + \frac{1}{2}\right) y\right\} \cap V$$

   and

$$W_2 = \left\{(x, y) : x > \left(n + \frac{1}{2}\right) y\right\} \cap V.$$

   Then W_1 and W_2 forms a separation of V. Thus $p(U) \subseteq W_1$ or $p(U) \subseteq W_2$. But $t_0 \in U$, so $b = p(t_0) \in p(U)$ and $b = (1, 0) \in W_2$. Then $p(U) \subseteq W_2$. However $p(t_1) \in W_1$. So there can be no such t_1 and $p(U) = \{b\}$. Consequently, $p^{-1}(b) = I$, as it is non-empty and both open and closed in I. So p is the constant path p_b, and no point in A can be reached from b. $\square$

It is possible to define another relation on X: xPy if there is a path in X from x to y.

1. For all $x \in X$, xPx as there is a path $p : I \to X$ defined by $p(t) = x$ for all $t \in I$;
2. if xPy there is a path $p : I \to X$ such that $p(0) = x$ and $p(1) = y$. Then, yPx as there is a path $q : I \to X$ defined by $q(t) = p(1-t)$.

3. if xPy and yPx there are paths $p, q : I \to X$ such that $p(1) = q(0) = y$, $p(0) = x$ and $q(1) = z$. Then xPz as there is a path $r = p.q : I \to X$ defined by

$$r(t) = (p.q)(t) = \begin{cases} p(2t) & \text{if } t \in [0, 1/2], \\ q(2t-1) & \text{if } t \in [1/2, 1]. \end{cases}$$

So P is an equivalence relation. The equivalence class of x is the **path component** of x in X. A path component need not be closed. Consider the space X from example 5 on p. 399. The subset A is a path component of X, but A is not closed in X since $(1, 0) \in \overline{A}$ but $(1, 0) \notin A$.

16.3 Local (Path) Connectedness

A space X is **locally (path) connected** if for each $x \in X$, every neighbourhood V_x of x contains a (path) connected neighbourhood of x. The following examples show that local (path) connectedness and (path) connectedness are independent properties.

Examples (Local (Path) Connectedness)

1. The space X from example 5 on p. 399 is connected but not locally connected, since the only connected neighbourhood of $(1, 0)$ is X.
2. The space $X = (0, 1) \cup (2, 3)$ is locally connected and locally path-connected, but it is clearly neither connected nor path connected.
3. Let $Y = X \cup S$, where X is the space from example 5 on p. 399 and S is an arc joining $(1, 0)$ to $(1, 1)$ without meeting any other point of X. Then X is path connected, but it is not locally path-connected. Indeed, the neighbourhood $V = B((1, 0), 1/2) \cap Y$ contains no path-connected neighbourhood.

There is a simple characterization of locally connected spaces.

Theorem 226
A space X is locally connected if and only if the components of each open subset V of X are open.

Proof: if X is locally path-connected and $V \subseteq_O X$, let C be a component of V. If $x \in V$, there is a connected neighbourhood U of x where $U \subseteq V$. As C is a maximal connected set, $U \subseteq C$ and C is open.

Conversely, suppose the components of open subsets are open. If V is a neighbourhood of x, let U be the component of x in V. Then U is a connected neighbourhood of x lying in V, so X is locally connected. ■

A similar theorem holds for locally path-connected spaces. We finish this section with the following result.

> **Theorem 227**
> *If X is a locally path-connected space, then the components and path components of X coincide.*
>
> **Proof:** If $x \in X$, there is a component C and a path component D of x. Since D is connected, $D \subseteq C$. By the previous theorem, $D \subseteq_O C$. If $y \in C \setminus D$, then there exists a path-connected neighbourhood V of y such that $V \subseteq C$. Then $V \cap D = \varnothing$. Otherwise $y \in D$ since there would be a path from x to y. Hence $y \in V \subseteq C \setminus D$ and $C \setminus D \subseteq_O C$. Then D is closed and open in C. Since C is connected, either $D = \varnothing$ or $D = C$. But $x \in D$, so $D = C$. ■

16.4 Solved Problems

1. Let A and B be connected subsets of a space X. For each of the following condition, either prove it to be sufficient to ensure that $A \cup B$ be connected or provide a counter-example to show that $A \cup B$ need not be connected:

   a) $\overline{A} \cap \overline{B} \neq \varnothing$;
   b) $\overline{A} \cap B \neq \varnothing$ and $A \cap \overline{B} \neq \varnothing$;
   c) $\overline{A} \cap B \neq \varnothing$ or $A \cap \overline{B} \neq \varnothing$.

   Solution:

   a) Let $X = \mathbb{R}$, $a \in \mathbb{R}$, $A = (-\infty, a)$ and $B = (a, +\infty)$. Then $\overline{A} = (-\infty, a]$, $\overline{B} = [a, +\infty)$ and $\overline{A} \cap \overline{B} = \{a\} \neq \varnothing$, but $A \cap B = \varnothing$, so $A \cup B$ is not connected. The condition is not sufficient.

   (b and c) Let $Y = A \cup B$. By a theorem seen in class, a separation of Y is a pair of non-empty subsets W and Z of Y such that $\overline{W} \cap Z = \varnothing$, $W \cap \overline{Z} = \varnothing$ and $Y = W \cup Z$. By hypothesis (in both cases), A and B can not form a separation of Y. Now suppose W and Z formed a separation of Y. Since A and B are connected, each of W and Z must contain exactly one of A and B, say $A \subseteq W$ and $B \subseteq Z$.[2] Since W and Z are disjoint, and $W \cup Z \subseteq A \cup B$, we get $W \subseteq A$ and $Z \subseteq B$, and so W and Z can not form a separation of Y, which is a contradiction. Hence, in both cases, $A \cup B$ is connected. □

2. Let X be locally path-connected. Show that every connected open set in X is path-connected.

   Proof: If $U = \varnothing$, the statement is vacuously true. So suppose $U \neq \varnothing$ is an open connected set in X. Since $U \subseteq_O X$, and X is locally path-connected, then, for every

[2]The only other possibility is that Y lies in one of W xor Z, which would make the other subset empty, and so W and Z could not form a separation of Y.

$x \in U$, there exists $V_x \subseteq_O X$ such that $x \in V_x \subseteq U$ and V_x is path-connected. Now, pick $z \in U$, define V to be the path component of U containing z and let $Y = U - V$. Since X is locally path-connected, V is open in X. Note that

$$\left(\bigcup_{y \in Y} V_y\right) \cap V = \varnothing;$$

otherwise, there would be a $y \in Y \cap V$, a contradiction. Hence we have $Y = \bigcup_{y \in Y} V_y$ and $Y \subseteq_O X$ since $V_y \subseteq_O X$ for all $y \in Y$.

But U is connected, so either $V = \varnothing$ or $Y = \varnothing$. Since $z \in V$, we must have $Y = \varnothing$ and $U = V$. Hence U is path-connected. ■

3. Let X be an ordered set (with at least two elements) in the order topology. Show that if X is connected, then X is a linear continuum.

Proof: a linear continuum is an ordered set in which

i. if $x < y$, there exists z such that $x < z < y$;
ii. any non-empty set A with an upper bound has a least upper bound.

Define the upper open ray and the lower open ray at x by

$$\begin{aligned} \text{UR}(x) &= \{y \in X | y < x\} \\ \text{LR}(x) &= \{y \in X | x < y\} \end{aligned}$$

for all $x \in X$. In the order topology, $\text{UR}(x), \text{LR}(x) \subseteq_O X$ for all $x \in X$. Now let $x, y \in X$ be such that $x < y$, and suppose that there does not exist $z \in X$ such that $x < z < y$. Then $\text{UR}(y) \cap \text{LR}(x) = \emptyset$, and

$$\text{UR}(y) \cup \text{LR}(x) = X.$$

Hence $\text{UR}(y), \text{LR}(x)$ is a separation of X, a contradiction since X is connected, so there must exist a $z \in X$ such that $x < z < y$.
Now, let A be a subset of X with at least one upper bound. Define the sets

$$\begin{aligned} U &= \bigcup_{a \in A} \text{UR}(a) \\ V &= \bigcup_{\substack{w > a \\ \forall a \in A}} \text{LR}(w). \end{aligned}$$

By construction, both U and V are open, and $U \cap V = \emptyset$. Since X is connected, $U \cup V \neq X$, otherwise U and V would be a separation of X. Suppose $b, c \in X - (U \cup V)$. Then, either $b < c$, $c < b$ or $b = c$. If $b < c$, then $c > a$ for all $a \in A$. By i., there exists $w \in X$ such that $b < w < c$, and $c \in \text{LR}(w) \subseteq V$. Similarly, if $c < b$, $b \in V$. This leaves only the possibility that $b = c$, that is $X - (U \cup V) = \{b\}$. By construction, b is smaller than any upper bound of A, and it is greater (or equal) than any element of A, so it is the least upper bound of A. Hence, X is a linear continuum. ■

16.5 Exercises

1. Prepare a 2-page summary of this chapter, with important definitions and results.

2. Is the product of an arbitrary collection of connected spaces connected in the box topology?

3. Show that a space X is locally path-connected if and only if the path-connected components of each open subset V of X are open.

4. Let A be a connected subset of a space X. If $A \subseteq B \subseteq \overline{A}$, show that B is connected. Are the interior and the boundary of A necessarily connected? If either of these is connected, must A be connected? What if both of them are connected?

5. Let A be a subset of a locally connected space. Prove or disprove:

   a) If A is path-connected and $A \subseteq B \subseteq \overline{A}$, then B is path-connected.

   b) If A is open and connected, then A is path-connected.

   c) If A is open, the path components are open.

6. Let X be the subspace
$$X = \left\{ \frac{t}{1+t} e^{it} \,\middle|\, t \geq 0 \right\} \cup \{e^{i\pi}\}.$$
   Give detailed answers to the following:

   a) Is X connected?

   b) Is X locally connected?

   c) Is X path-connected?

   d) Is X locally path-connected?

7. Let $\mathfrak{T}$ and $\mathfrak{T}'$ be two topologies on a space X. If $\mathfrak{T}'$ is finer than $\mathfrak{T}$, does connectedness of X in one topology imply anything about its connectedness in the other?

8. If $|X|$ is infinite, show that X is connected in the finite complement topology.

9. If X_α is path-connected for each α, show that $\prod_\alpha X_\alpha$ is path-connected. If each X_α is also locally path-connected, show that $\prod_\alpha X_\alpha$ is also locally path-connected. Investigate what happens when each X_α is locally path connected, but not necessarily path-connected.

Chapter 17

Compact Spaces

In Chapter 9, we discussed **compactness** in the context of metric spaces. In this chapter, we discuss the notion from a **topological perspective**.

17.1 Compactness

A **covering of a space** X is a family $\mathfrak{F}$ of subsets of X such that

$$\bigcup_{F \in \mathfrak{F}} F = X.$$

A subset Y of X is **covered by a family** $\mathfrak{F}$ if

$$Y \subseteq \bigcup_{F \in \mathfrak{F}} F.$$

We say that $\mathfrak{F}$ is an **open covering** when every $F \in \mathfrak{F}$ is open. A sub-collection $\mathcal{A} \subseteq \mathfrak{F}$ that still covers X is called a **sub-covering** of X.

> **Examples** (Coverings and Sub-coverings)
>
> 1. Consider the sets $\mathfrak{F} = \{(a,b) \mid a < b \in \mathbb{R}\}$, $\mathcal{A} = \{(a,b) \mid a < b \in \mathbb{Q}\}$ and $\xi = \{(n-1, n+1) \mid n \in \mathbb{Z}\}$. Then $\mathfrak{F}$ is an open covering of $\mathbb{R}$, $\mathcal{A}$ and ξ are sub-coverings, but ξ has no proper sub-covering.
>
> 2. The collection $\mathfrak{F} = \{[a,b) \mid a < b \in \mathbb{R}\}$ is an open covering of $\mathbb{R}_l$. $\square$

A space X is **compact** if every open covering of X contains a finite sub-covering. A subspace C of X is **compact in** X if every open covering of C contains a finite sub-covering.[1]

[1] This definition seems rather straightforward, on the face of it, but it is the culmination of a rather long and arduous process, with dead ends and wrong turns – we will look into some of these in the coming pages.

Examples (Compact Spaces)

1. $\mathbb{R}$ is not compact, since the covering ξ from the previous example contains no proper sub-covering, hence no finite sub-covering.

2. Let $X = \{0\} \cup \left\{\frac{1}{n} \,\middle|\, n \in \mathbb{N}\right\}$. Any open covering of X will contain a neighbourhood of 0, say V_0. For some N, we have $\frac{1}{n} \in V_0$ for all $n > N$. For each n such that $1 \leq n \leq N$, pick V_n from the open covering such that $\frac{1}{n} \in V_n$. Then $\{V_0, V_1, \ldots, V_n\}$ is a sub-covering, so X is compact. □

It takes some practice to get the hang of the definition.

Theorem 228
Every closed subset C of a compact set X is compact.

Proof: suppose $\{U_\alpha\}_\alpha$ is an open covering of C. As C is closed, $X \setminus C$ is open and $\{U_\alpha\}_\alpha \cup \{X \setminus C\}$ is an open covering of X. As X is compact, there exists a finite sub-covering of X, say $\{U_{\alpha_i}\}_{i=1}^n$. If $X \setminus C = U_{\alpha_j}$ for some j, discard U_{α_j}. The remaining $\{U_{\alpha_i}\}_{i(\neq j)=1}^n$ is a finite sub-covering of C. In the other case, the finite sub-covering of X is clearly a finite sub-covering of C. Hence C is compact. ■

In general, the converse is not true (see example 1 on p. 407). However, it holds for a broad class of spaces.

Theorem 229
If X is Hausdorff, every compact subset of X is closed.

Proof: let Y be a compact subset of X. As X is Hausdorff, if $x \notin Y$, for each $y \in Y$, there is two disjoint neighbourhoods U_y of y, V_y of x. Then $\{U_y\}_{y \in Y}$ is an open covering of Y. But Y is compact so there is a finite sub-covering, say $\{U_{y_i}\}_{i=1}^n$.

Now, write

$$V = \bigcap_{i=1}^n V_{y_i}, \quad \text{and} \quad U = \bigcup_{i=1}^n U_{y_i}.$$

Then V is a neighbourhood of x such that

$$V \cap Y \subseteq V \cap U = \bigcup_{i=1}^n (V \cap U_{y_i}) = \varnothing.$$

Hence we can fit an open set V around every $x \notin Y$, which means $X \setminus Y$ is open and Y is closed. ■

Note that we have in fact proven the following result.

Corollary
If X is Hausdorff, and C is a compact subset of X, then for $x \notin C$, there exists disjoint open sets U, V such that $x \in V$ and $C \subseteq U$.

What can we say when the spaces are not Hausdorff? Depends on the situation, actually.

Examples

1. If $X = \{a, b\}$ has the indiscrete topology, then every subset of X is compact. In particular, $\{a\}$ is compact. However, it is not closed since $\{b\}$ is not open.
2. In $\mathbb{R}$ with the finite complement topology, every subset is compact. Indeed, let C be a subset of $\mathbb{R}$, with open covering $\mathfrak{F}$. For $F \in \mathfrak{F}$, F covers C for at most a finite number of points, say $\{c_i\}_{i=1}^n$. Pick $F_i \in \mathfrak{F}$ such that $c_i \in F_i$ for all i. Then $\{F, F_1, \ldots, F_n\}$ covers C, and so C is compact.

In the topology of the last example, even the open sets are compact. This does not contradict Theorem 229 since $\mathbb{R}$ is not Hausdorff in the finite complement topology. As it happens, compactness is a **topological notion**.

Theorem 230
The continuous image of a compact set $C \subseteq X$ by $f : X \to Y$ is compact.

Proof: let $\mathfrak{F}$ be an open covering of $f(C)$. By continuity, $\{f^{-1}(F)\}_{F \in \mathfrak{F}}$ is an open covering of C. So there is a finite sub-covering, say $\{f^{-1}(F_1), \ldots, f^{-1}(F_n)\}$, as C is compact, and

$$f(C) \subseteq f\left(\bigcup_{i=1}^n f^{-1}(F_i)\right) = \bigcup_{i=1}^n f\left(f^{-1}(F_i)\right) \subseteq \bigcup_{i=1}^n F_i.$$

Then $\{F_1, \ldots, F_n\}$ covers $f(C)$, and so $f(C)$ is compact. ■

There are all sorts of results about compact spaces and continuous functions.

Theorem 231
If X is compact, Y is Hausdorff, and $f : X \to Y$ is a continuous bijection, then f is a homeomorphism.

Proof: let C be a closed subset of X. As X is compact, C is compact. Since $f : X \to Y$ is continuous, $f(C)$ is compact in Y and thus closed in Y, as Y is Hausdorff. So f is closed. As f is a continuous bijection, f is a homeomorphism. ■

As we had done with connectedness, we would like to show that finite products of compact spaces are compact.[2] To do this we will need the following lemma.

Lemma 232 (TUBE LEMMA)
If Y is compact and N is an open set in $X \times Y$ which contains the slice $\{x_0\} \times Y$, then there exists a neighbourhood W of x_0 such that $W \times Y \subseteq N$.

Proof: if $y \in Y$, then $(x_0, y) \in N$. As N is open, there exists two neighbourhoods U_y of x_0 and V_y of y such that $U_y \times V_y \subseteq N$. Repeating this process for all $y \in Y$ yields an open covering $\{V_y\}_{y \in Y}$ of Y. As Y is compact, there is a finite sub-covering, say $\{V_{y_1}, \ldots, V_{y_n}\}$, with $U_{y_i} \times V_{y_i} \subseteq N$ for all $1 \leq i \leq n$. Let

$$W = \bigcap_{i=1}^{n} U_i.$$

then W is open in X as it is a finite intersection of open sets. Furthermore, $x_0 \in W$ as $x_0 \in U_{y_i}$ for all $1 \leq i \leq n$. Now, let $(x, y) \in W \times Y$. There is a j such that $y \in V_{y_j}$. As $x \in W$, $x \in U_{y_j}$. Then $(x, y) \in U_j \times V_j \subseteq N$, so $W \times Y \subseteq N$. ■

We now have all the machinery to prove the following result.

Theorem 233
If X and Y are compact, then $X \times Y$ is compact.

Proof: let $\mathfrak{F}$ be an open covering for $X \times Y$. For each $x \in X$ we get a finite sub-covering of $\{x\} \times Y$ from $\mathfrak{F}$, say $F(x)_1, \ldots F(x)_n$. Let N be the open set $N = \bigcup_{i=1}^{n} F(x)_i$. By the Tube Lemma, there is a neighbourhood W_x of x in X such that $W_x \times Y \subseteq N$. Repeating this procedure for all $x \in X$, we get that $\{W_x\}_{x \in X}$ is an open covering of X. But X is compact, so there is a finite sub-covering $\{W_{x_1}, \ldots, W_{x_m}\}$. For each of these W_{x_i}, there were n_i corresponding sets $F(x_i)_j$ in $\mathfrak{F}$. Define

$$\mathfrak{F}' = \{F(x_i)_j \mid 1 \leq i \leq m, 1 \leq j \leq n_i\}.$$

$\mathfrak{F}'$ is a finite open collection, with $\sum_{i=1}^{m} n_i$ elements. For any $(x, y) \in X \times Y$, $x \in W_{x_i}$ for some i. Then $(x, y) \in W_{x_i} \times Y$ and $(x, y) \in F(x_i)_j$ for some j, so

$$X \times Y \subseteq \bigcup_{i=1}^{m} \bigcup_{j=1}^{n_i} F(x_i)_j.$$

Thus $\mathfrak{F}'$ is a finite sub-covering of $X \times Y$ from $\mathfrak{F}$, and so $X \times Y$ is compact. ■

[2] It is not a simple matter to generalize to arbitrary products of compact spaces. This will be the content of chapter 19.

Du to the complementary nature of open and closed sets, it is also possible to express compactness in term of closed sets. A family $\mathfrak{F}$ of sets has the **finite intersection property** whenever

$$\bigcap_{i=1}^{n} F_i \neq \varnothing$$

for any selection $F_i \in \mathfrak{F}$, $1 \leq i \leq n$.

Theorem 234
A space X is compact if and only if every family $\{F_\alpha\}_\alpha$ of closed subsets of X having the finite intersection property has a non-void intersection, that is, $\bigcap_\alpha F_\alpha \neq \varnothing$.

Proof: we make the following three remarks: $\{X \setminus F_\alpha\}_\alpha$ is an open family if and only if $\{F_\alpha\}_\alpha$ is a closed family;

$$\bigcup_{\alpha}(X \setminus F_\alpha) = X \iff \bigcap_{\alpha} F_\alpha = \varnothing$$

and

$$\bigcup_{i=1}^{n}(X \setminus F_{\alpha_i}) = X \iff \bigcap_{i=1}^{n} F_{\alpha_i} = \varnothing$$

for any selection $F_i \in \mathfrak{F}$, $1 \leq i \leq n$. The theorem is easily proved using the contrapositive statement and the three remarks. ■

There is another version of this theorem:

Theorem 234 (Reprise)
A space X is compact if and only if for every family $\mathcal{A}$ of subsets of X satisfying the finite intersection property, the intersection $\bigcap_{A\in\mathcal{A}} \overline{A}$ is not empty.

Proof: left as an exercise. ■

As an easy corollary we get the following result.

Corollary 235 Let X be a compact space, and suppose

$$C_1 \supseteq C_2 \supseteq \cdots \supseteq C_n \supseteq \cdots$$

is a nested sequence of closed sets. Then

$$\bigcap_{n\in\mathbb{N}} C_n \neq \varnothing.$$

Interest in compact spaces arose when we realized that there was something special about intervals in the usual topology on $\mathbb{R}$ that made the max/min theorem come out as it did.

Theorem 236 If X has the order topology, where the ordering has the least upper bound property, then each interval

$$[a,b] = \{x \mid a \leq x \leq b\}$$

is compact.

Proof: the proof is similar to that of the Heine-Borel theorem (see Proposition 125 in Chapter 9, and Theorem 237). This means that the key point of the proof is the least upper bound property, and not the metric. ■

The first step in the process was a generalization of intervals to $\mathbb{R}^n$.

Theorem 237 (Heine-Borel Theorem – Reprise)
In the usual topology, the compact sets of $\mathbb{R}^n$ are exactly the closed and bounded sets.

Proof: since $\mathbb{R}^n$ is Hausdorff, any of its compact subset is closed. If C is compact in $\mathbb{R}^n$, it can be covered by

$$\{(-m,m)^n \mid m \in \mathbb{N}\}.$$

But C is compact, so it has a finite sub-covering and there exists $M \in \mathbb{N}$ such that $C \subseteq (-M,M)^n$. Thus C is bounded.

Conversely, suppose that C is a closed bounded set. Then, there exists $M \in \mathbb{N}$ such that $C \subseteq [-M,M]^n$. But $[-M,M]^n$ and C is a finite product of the compact spaces $[-M,M]$, and so is itself compact. C is then compact since it is a closed subset of a compact set. ■

Note that this result need not hold for a general metric space (where boundedness may not be defined, for instance), as we shall see shortly.[3]

Theorem 238 (Maximum and Minimum Value Theorem)
Let C be a compact subset of X, and suppose $f : X \to Y$ is continuous, where Y has a (total) order topology. Then f is bounded on C and actually attains its bounds there.

Proof: as C is compact and f is continuous, $f(C)$ is compact. If f does not have a largest value on C, then, for each $a \in C$, there exists $a' \in C$ such that $f(a) < f(a')$. For any $y \in Y$, denote $(-\infty, y) = \{z \in Y \mid z < y\}$.

[3]This is the reason for the less-than-intuitive definition of compactness currently in use.

Then

$$\{(-\infty, f(a))\}_{a\in C}$$

is an open covering of $f(C)$. But $f(C)$ is compact, so there exists a finite subcovering, say

$$\{(-\infty, f(a_i))\}_{i=1}^n.$$

Let $a_0 \in C$ be the a_i that maximizes $f(a_i)$. Since $f(a_0) \in f(C)$, $f(a_0) \in (-\infty, f(a_j))$ for some j, which means that $f(a_0) < f(a_j) \leq f(a_0)$, a contradiction, since $x \not< x$ in Y. Hence f has a largest value on C. The proof that f has a smallest value on C is similar. ■

This result is the generalization to topological spaces of one of the fundamental results of analysis (see Theorem 33 in Chapter 3.)

Metric Spaces (Reprise)

Let us revisit metric spaces from the vantage point of topology. If d is a metric on a space X, the basic open sets in X are the open balls

$$B_d(a,r) = \{x \in X \mid d(a,x) < r\}.$$

The topology generated by these basic sets is called the **metric topology on** X.[4]

Let us suppose that the metrics d and d' generate the topologies $\mathfrak{T}$ and $\mathfrak{T}'$ on X. $\mathfrak{T}$ is finer than $\mathfrak{T}'$ whenever $B_{d'}(a,r')$ is open in $\mathfrak{T}$ for all $a \in X$, $r' \in \mathbb{R}^+$, and so whenever there exists $r \in \mathbb{R}^+$ such that

$$B_d(a,r) \subseteq B_{d'}(a,r').$$

Example: let d be the Euclidean metric on $\mathbb{R}^2$ and d' be defined on $\mathbb{R}^2$ by $d'((x_1,y_1),(x_2,y_2)) = \max\{|x_1-x_2|, |y_1-y_2|\}$. Then

$$B_d(0,1) = \{(x,y) \mid x^2+y^2 < 1\},\ B_{d'}(0,1) = \{(x,y) \mid -1 < x < 1 \text{ and } -1 < y < 1\},$$

$B_d(0,1) \subseteq B_{d'}(0,1)$ and $B_{d'}(0, \frac{1}{\sqrt{2}}) \subseteq B_d(0,1)$. Generalizing to all open balls, one gets $\mathfrak{T} = \mathfrak{T}'$. □

[4]The collection of all open balls is a basis. Indeed, $x \in B_d(x,1)$ for all $x \in X$. The empty set is a ball of radius 0. Suppose that $y \in B_d(x_1,r_1) \cap B_d(x_2,r_2) \neq \varnothing$. Then $y \in B_d(y,r) \subseteq B_d(x_1,r_1) \cap B_d(x_2,r_2)$, where

$$r = \frac{\min\{d(x_1,y)-r_1, d(x_2,y)-r_2\}}{2}.$$

Let X be a metric space with metric d. The **standard bounded metric** $\overline{d}$ on X is the metric defined by

$$\overline{d}(x, y) = \min\{d(x, y), 1\}$$

(the only property that is not trivially true is the triangle inequality). For any ball $B_d(0, \varepsilon)$, put $\delta = \min\{\varepsilon, 1\}$. Then

$$B_{\overline{d}}(a, \delta) \subseteq B_d(a, \varepsilon).$$

For a ball $B_{\overline{d}}(a, \delta)$, if $\delta \leq 1$, then $B_{\overline{d}}(a, \delta) = B_d(a, \delta)$. If $\delta > 1$, $B_{\overline{d}}(a, \delta) = X$. This means that the topology generated by the bounded standard metric $\overline{d}$ on X is the same as that generated by the metric d. Consequently, we may assume that the metric d is bounded.

A space X is **metrizable** if there is a metric d on X where the metric topology on X coincides with the topology on X. This leads us to one of the fundamental differences between metric spaces and general topological spaces, a result which is simple to state, but whose proof is surprisingly sophisticated.[5]

Theorem 239
Any countable product of metrizable spaces is metrizable.

Proof: Suppose (X_n, d_n) is a metric space and d_n is the standard bounded metric on X_n for all $n \in \mathbb{N}$. Let $x, y \in X = \prod X_n$ and define

$$d(x, y) = \text{l.u.b.} \left\{ \frac{d_n(x_n, y_n)}{n} \right\}_{n \in \mathbb{N}}$$

It is not hard to see that this defines a metric on X. We need to verify that the topology generated on X by d is that given by the product topology.

Suppose $U \subseteq X$ is open in the product topology. If $x = (x_n) \in U$, there is a basic set $\prod V_n$, where $V_n \subseteq_O X_n$, and $V_n = X_n$ for all but a finite number of n's, i.e for all $n > N$ for some N. Then there exists $\varepsilon_n > 0$ such that $B(x_n, \varepsilon_n) \subseteq V_n$ for all $n \in \mathbb{N}$. Let

$$\varepsilon = \min \left\{ \frac{\varepsilon_n}{n} \right\}_{n=1}^{N}.$$

If $y = (y_n) \in B(x, \varepsilon)$, then $d(x, y) < \varepsilon$, so

$$\frac{d_n(x_n, y_n)}{n} \leq d(x, y) < \varepsilon \leq \frac{\varepsilon_n}{n}$$

for $1 \leq n \leq N$.

[5] As a reminder, the notations $A \subseteq_O X$, $A \subseteq_C X$, and $A \subseteq_K X$ are used respectively for A *is open in* X, A *is closed in* X, and A *is compact in* X.

Hence $d(x_n, y_n) < \varepsilon_n$, and so

$$y_n \in B(x_n, \varepsilon_n) \subseteq V_n$$

for $1 \leq n \leq N$ and $y_n \in X_n = V_n$ for all $n > N$, so $y \in \prod V_n$ and $B(x, \varepsilon) \subseteq \prod X_n$. Then $\prod X_n$ is open in the metric topology. As a result, around each point of U, we can fit an open set in the metric topology, i.e. $U \subseteq X$ is open in the metric topology.

Conversely, suppose $U \subseteq X$ is open in the metric topology. Then, if $x \in U$, there exists $\varepsilon > 0$ such that $B(x, \varepsilon) \subseteq U$. Choose N such that $\frac{1}{N} < \varepsilon$. Put $V_n = B(x_n, n\varepsilon)$ for all n, so that $V_n = X_n$ whenever $n > N$ (remember, the metrics d_n are standard bounded metrics). If $y = (y_n) \in \prod V_n$, then $d_n(x_n, y_n) < n\varepsilon$ for all $n \in \mathbb{N}$. In particular,

$$\frac{d_n(x_n, y_n)}{n} < \varepsilon$$

whenever $1 \leq n \leq N$ and

$$\frac{d_n(x_n, y_n)}{n} \leq \frac{1}{n} < \frac{1}{N} < \varepsilon$$

for all $n > N$. By construction,

$$d(x, y) = \text{l.u.b.} \left\{ \frac{d_n(x_n, y_n)}{n} \right\}_{n \in \mathbb{N}} < \varepsilon.$$

Then $y \in B(x, \varepsilon)$ and $\prod V_n \subseteq B(x, \varepsilon)$. As a result, around each point of U, we can fit an open set in the product topology, i.e. $U \subseteq X$ is open in the product topology. ■

Let (X_α, d_α) be a collection (not necessarily countable) of metric spaces, where d_α is a standard bounded metric on X_α. Define a metric d on $\prod_\alpha X_\alpha$ by

$$d(x, y) = \text{l.u.b.}\{d_\alpha(x_\alpha, y_\alpha)\}$$

for all $x, y \in X$.[6] This metric is called the **uniform metric**, and the topology it generates on $\prod_\alpha X_\alpha$ is called the **uniform topology** on X. We will in the solved problems that the uniform topology is finer than the product topology and coarser than the box topology, and that for infinite products, the inclusions are strict.

We now introduce another concept that allows us to tell if a space is metrizable. A sequence $\{x_n\}_n \in \mathbb{N}$ in a space X (not necessarily metric) **converges to** $x \in X$ (denoted $x_n \to x$) if for every neighbourhood V of x, there exists $N \in \mathbb{N}$ such that $x_n \in V$ for every $n > N$.

In a general topological space, the **limit** of a sequence is not necessarily unique!

[6] The only non-trivial component here is again the triangle inequality.

Examples (LIMITS)

1. Let $X = [0,1]$, where the basic open sets in X are of the form (a,b) and $[0,a) \cup (b,1]$ for $0 < a < b < 1$. In the topology generated by this basis, every neighbourhood of 0 is a neighbourhood of 1, and vice-versa. Thus $\frac{1}{n} \to 0$ as usual, but $\frac{1}{n} \to 1$ as well.

2. Let X be a space with the indiscrete topology. Then every sequence in X converges to every element of X. □

Suppose $\{a_n\}_{n\in\mathbb{N}}$ is a sequence in a set $A \subseteq X$, and let $a_n \to a \notin A$. Then a is a **limit point** of A. Indeed, for any neighbourhood V of a, there is some index N for which $a_n \in V$ when $n > N$. Consequently, $a_n \in V \cap A$ and $a \neq a_n$ for all $n > N$ (as $a \notin A$), so $a \in \overline{A}$.

In general, if a sequence in A converges to a point not in A, the limit is a limit point. However, the converse statement is false: if $a \in \overline{A}$, there might not be a sequence in A converging to a, as can be seen in the next example.[7]

Example: let Ω be the first uncountable ordinal; let X be the set $\Omega^+ = \Omega \cup \{\Omega\}$, with the order topology. Consider $A = \Omega = [0,\Omega)$. Suppose the sequence $\{\alpha_n\}_{n\in\mathbb{N}}$, where $\alpha_n \in A$, has the limit α. As

$$\alpha_n \leq \bigcup_{m\in\mathbb{N}} \alpha_m = \beta,$$

then $\alpha \leq \beta$. But β is a countable union of countable sets, hence it is countable. Therefore, $\beta < \Omega$, so $\alpha < \Omega$ and $\alpha_n \notin (\beta,\Omega)$ for all $n \in \mathbb{N}$. Now $\overline{A} = [0,\Omega]$, and so $\Omega \in \overline{A}$, but no sequence in A converges to Ω. □

This example may seem a bit far-fetched, but that is the nature of the discipline – in general topology, exotic counter-examples are entirely legitimate. In metric spaces, however, things tend to be substantially better behaved.

Lemma 240 (SEQUENCE LEMMA)
Let X be a metrizable space. For any subset A of X, if $a \in \overline{A}$, then there is a sequence $\{a_n\}_{n\in\mathbb{N}} \subseteq A$ with $a_n \to a$.

Proof: let d be the metric generating the topology on X. For each $n \in \mathbb{N}$, construct the neighbourhood $B(a,\frac{1}{n})$. As $a \in \overline{A}$, we have $A \cap B(a,\frac{1}{n}) \neq \varnothing$ for all $n \in \mathbb{N}$. Let $a_n \in A \cap B(a,\frac{1}{n})$ for all $n \in \mathbb{N}$. Then $a_n \to a$. Indeed, let V be a neighbourhood of a. Then there is a basic neighbourhood $B(a,\varepsilon) \subseteq V$. Let $N \geq \frac{1}{\varepsilon}$. Then, whenever $n > N$, we get $d(a,a_n) < \frac{1}{n} < \frac{1}{N} \leq \varepsilon$, hence $a_n \in V$ and $a_n \to a$. ■

[7]The example requires some familiarity with the **first uncountable ordinal**, see `https://en.wikipedia.org/wiki/First_uncountable_ordinal` for details.

Next, we see that one of the sacred cows of analysis see Proposition 106 in Chapter 8) may not necessarily hold in general topological spaces.

Theorem 241
The function $f : X \to Y$ is continuous if, whenever $a_n \to a$ in X, then $f(a_n) \to f(a)$ in Y. If X is metrizable, the converse holds.

Proof: suppose that f is continuous and $a_n \to a$ in X. Let V be a neighbourhood of $f(a)$. Then $f^{-1}(V)$ is a neighbourhood of a, and so there exists N such that $a_n \in f^{-1}(V)$ whenever $n > N$. Then $f(a_n) \in V$ whenever $n > N$ and $f(a_n) \to f(a)$ in Y.

Conversely, suppose X is metrizable and that the sequence condition holds. Let $A \subseteq X$. By the sequence lemma, if $a \in \overline{A}$, there is a sequence $\{a_n\}_{n\in\mathbb{N}} \subseteq A$ such that $a_n \to a$. By hypothesis, $f(a_n) \to f(a)$, so $f(a) \in \overline{f(A)}$, as $f(a_n) \in f(A)$ for all $n \in \mathbb{N}$. Hence $f(\overline{A}) \subseteq \overline{f(A)}$ which is equivalent to f being continuous. ■

17.2 Limit Point and Sequential Compactness

Throughout the history of topology, many definitions of compactness have been formulated. At the time, each were thought to have isolated the crucial property of a set like $[0,1]$ that made the maximum/minimum theorem possible, amongst others.

As our understanding of topology increased, these different notions were discarded, to be replaced by the modern concept. But the failed candidates are interesting in their own rights, as they coincide with compactness in the case of metric spaces, as we shall see.

A subset A in a space X is said to be **sequentially compact** if every sequence $\{a_n\}_{n\in\mathbb{N}} \subseteq A$ contains a convergent subsequence. A subset A in a space X is said to be **limit point compact** if every infinite subset of A has a limit point.The next few results show how the various compactness notions are related.

Proposition 242
If X is compact, then X is limit point compact.

Proof: suppose X is compact and let A be a subset of X with no limit point. Then $A = \overline{A}$, and A is closed, so compact. Also, for any $a \in A$, there is a neighbourhood V_a such that $V_a \cap A = \{a\}$. Thus $\{V_a\}_{a\in A}$ is an open covering of A. Since A is compact, there is a finite sub-covering $\{V_{a_i}\}_{i=1}^n$. But

$$A = A \cap \left(\bigcup_{i=1}^n V_{a_i}\right) = \bigcup_{i=1}^n (A \cap V_{a_i}) = \bigcup_{i=1}^n \{a_i\}.$$

Hence A is finite. By contraposition, X is limit point compact. ■

Proposition 243
If X is a limit point compact metric space, then X is sequentially compact.

Proof: let $\{a_n\}_{n\in\mathbb{N}} \subseteq X$, and write $A = \{a_1, a_2, \ldots\}$. If A is finite, there has to exist a constant (hence convergent) subsequence $\{a_{n_m}\}_{m\in\mathbb{N}}$. Otherwise, A is infinite. As X is limit point compact, A has a limit point, say a and every neighbourhood of a contains a point in A different from a. In particular, since X is a metric space, for each $m \in \mathbb{N}$, $B(a, \frac{1}{m})$ is a neighbourhood of a and there exists $a_{n_m} \in B(a, \frac{1}{m}) \cap A$ such that $a_{n_m} \neq a$. By construction, $a_{n_m} \to a$, so X is sequentially compact. ■

Proposition 244
If X is sequentially compact, then X is limit point compact.

Proof: let A be an infinite subset of the X. Then A contains a countable subset $\{a_1, a_2, \ldots\}$. As X is sequentially compact, there is a convergent subsequence $a_{n_m} \to a$. By construction, a is a limit point of A and X is limit point compact. ■

The following result to show that the notions of compactness are equivalent for metric spaces.

Theorem 245
Let X be a compact metric space. For any open covering $\mathfrak{F}$ of X, there is a number $\delta > 0$ satisfying the following property: if $A \subseteq X$ is such that diam$(A) < \delta$, then there exists $F \in \mathfrak{F}$ such that $A \subseteq F$.

Proof: we prove the theorem by contradiction. Let $\mathfrak{F}$ be an open covering of X, and suppose that no δ satisfying the property exists. Then, for each $n \in \mathbb{N}$, we can find a set A_n such that diam$(A_n) < \frac{1}{n}$ where $A_n \nsubseteq F$ for all $F \in \mathfrak{F}$. As $A_n \neq \varnothing$ for all $n \in \mathbb{N}$, we can select $a_n \in A_n$ for all $n \in \mathbb{N}$, and get the sequence $\{a_n\}_{n\in\mathbb{N}}$.

In a metric space, compactness implies sequential compactness, so there is a convergent subsequence $\{a_{n_m}\}_{m\in\mathbb{N}}$, with $a_{n_m} \to a \in X$. Pick $F \in \mathfrak{F}$ such that $a \in F$. As F is open, there exists $r > 0$ such that $B(a, 2r) \subseteq F$.

Since the subsequence is convergent, there is a number $N \in \mathbb{N}$ such that $a_{n_m} \in B(a, r)$ for all $n_m > N$. Pick $n_m >$ such that $\frac{1}{n_m} < r$. If $x \in A_{n_m}$, then

$$d(x, a) \leq d(x, a_{n_m}) + d(a_{n_m}, a) < \frac{1}{n_m} + r < 2r$$

since diam$(A_{n_m}) < \frac{1}{n_m}$ and $A_{n_k} \subseteq B(a, 2r) \subseteq F$, a contradiction. So there must be a number $\delta > 0$ satisfying the property. ■

The number δ in the proof of Theorem 245 is called a **Lebesgue number** of the covering $\mathfrak{F}$.

We need one more definition before we are ready to prove our big result. A metric space X is **totally bounded** if, for every $\varepsilon > 0$, X can be covered by a finite number of ε-balls.

Theorem 246
In a metric space, compactness, sequential compactness, and limit point compactness are equivalent.

Proof: according to Propositions 242, 243, and 244, it only remains to show that a sequentially compact set X is compact. Let $\mathfrak{F}$ be an open covering of X, and suppose X is not totally bounded. Then there exists $\varepsilon > 0$ such that there is no finite covering of X by ε-balls.

Let $x_1 \in X$. As $B(x_1, \varepsilon) \neq X$, select $x_2 \in X \setminus B(x_1, \varepsilon)$. It is possible to select

$$x_{n+1} \in X \setminus \bigcup_{i=1}^{n} B(x_i, \varepsilon)$$

since $\bigcup_{i=1}^{n} B(x_i, \varepsilon)$ does not cover X. By recursion, $\{x_n\}_{n\in\mathbb{N}}$ is a sequence, and it contains a converging subsequence $\{x_{n_m}\}_{m\in\mathbb{N}}$, where $x_{n_m} \to x$, since X is sequentially compact. Then, there exists M such that $x_{n_m} \in B(x, \frac{\varepsilon}{2})$ and

$$d(x_{n_{m+1}}, x_{n_m}) \leq d(x, x_{n_{m+1}}) + d(x, x_{n_m}) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

whenever $m > M$. But this yields $x_{n_{m+1}} \in B(x_{n_m}, \varepsilon)$, which is a contradiction by construction of the sequence $\{x_n\}_{n\in\mathbb{N}}$. Hence X must be totally bounded.

Let 3ε be a Lebesgue number of $\mathfrak{F}$. Then there exists a finite collection $\mathfrak{B} = \{B(y_i, \varepsilon)\}_{i=1}^{n}$ covering X. As diam$(B(y_i, \varepsilon)) \leq 2\varepsilon < 3\varepsilon$, $\exists F_i \in F$ such that $B(y_i, \varepsilon) \subseteq F_i$ for all $1 \leq i \leq n$. If $x \in X$, then $x \in B(y_i, \varepsilon) \subseteq F_i$ for some i. Since $\mathfrak{B}$ is a finite covering of X, $\{F_i\}_{i=1}^{n}$ is a finite sub-covering of X, and so X is compact. ■

The converse of Proposition 242 is not in general true, as can be seen in the following example (which again uses the smallest uncountable ordinal).

Example: let Ω be the first uncountable ordinal, and let X be the set $\Omega^+ = \Omega \cup \{\Omega\}$, with the order topology. Now, Ω is limit point compact. Indeed, suppose C is an infinite (countable) subset of Ω. Then C is bounded above by $\bigcup_{\gamma\in C} \gamma = \beta$, and so $C \subseteq [0, \beta]$. It is clear that Ω has the l.u.b. property, so, by Theorem 236, $[0, \beta]$ is compact. By Proposition 242, $[0, \beta]$ is limit point compact, and so C contains a limit point. Thus Ω is limit point compact. But Ω isn't closed in the Hausdorff space X, so Ω is not compact. □

17.3 Local Compactness and One-Point Compactification

By analogy with local connectedness, we can also define a notion of local compactness: a space X is **locally compact at** $x \in X$ if there exists a compact set C which contains a neighbourhood V of x. We say that X is **locally compact** if it is locally compact at each $x \in X$.

There is an equivalent definition if X is Hausdorff space. For each $x \in X$, if there exists a neighbourhood V and a compact set C such that $x \in V \subseteq C$, then, as X is Hausdorff, C is closed, so $\overline{V} \subseteq C$, and x has a neighbourhood with compact closure.

Examples (Local Compactness)

1. Every compact space is locally compact.
2. $\mathbb{R}$ is locally compact, since, for any basic open set $]a, b[$, the closure $[a, b]$ is compact. Similarly, $\mathbb{R}^n$ is locally compact for all $n \in \mathbb{N}$. However $\mathbb{R}^\omega$ is not locally compact in the product topology. Indeed, let

$$V = (a_1, b_1) \times \cdots (a_n, b_n) \times \mathbb{R} \times \cdots$$

be a basic neighbourhood in the product topology. Then

$$\overline{V} = [a_1, b_1] \times \cdots [a_n, b_n] \times \mathbb{R} \times \cdots ,$$

which is not compact in the product topology. □

Let X be a locally compact Hausdorff space, and suppose that ∞ is a point not in X. Construct a new set $Y = X \cup \{\infty\}$, with the following topology: $V \subseteq_O Y$ if either

- $V = U \subseteq_O X$ whenever $\infty \notin V$, or;
- $V = Y \setminus C$, where C is a compact subset of X whenever $\infty \in V$.

This is indeed a topology on Y, as we see presently.

1. $\varnothing$ is an open set of type 1, Y is an open set of type 2.
2. Let $V_1, V_2 \subseteq_O Y$. Then
   a) $V_1, V_2 \subseteq_O X$, so $V_1 \cap V_2 \subseteq_O X$, hence $V_1 \cap V_2 \subseteq_O Y$; or
   b) $V_1 \subseteq_O X$ and $V_2 = Y \setminus C$, where $C \subseteq_K X$. Then

$$V_1 \cap V_2 = V_1 \cap (Y \setminus C) = V_1 \cap (X \setminus C) \subseteq_O Y,$$

   as C is closed in X, since X is Hausdorff; or
   c) $V_1 = Y \setminus C_1$, $V_2 = Y \setminus C_2$ where $C_1, C_2 \subseteq_K X$. Then

$$V_1 \cap V_2 = (Y \setminus C_1) \cap (Y \setminus C_2) = Y \setminus (C_1 \cup C_2) \subseteq_O Y,$$

   since $C_1 \cup C_2 \subseteq_K X$ whenever $C_1, C_2 \subseteq_K X$.

3. a) $V_\beta \subseteq_O X$, so $\bigcup_\beta V_\beta \subseteq_O X$, hence $\bigcup_\beta V_\beta \subseteq_O Y$; or
   b) $V_\alpha \subseteq_O X$ (i.e $\bigcup_\alpha V_\alpha \subseteq_O X$) and $V_\beta = Y \setminus C_\beta$, where $C_\beta \subseteq_K X$. Then

$$\left(\bigcup_\alpha V_\alpha\right) \cup \left(\bigcup_\beta V_\beta\right) = \left(\bigcup_\alpha V_\alpha\right) \cup \left(\bigcup_\beta (Y \setminus C_\beta)\right) = \left(\bigcup_\alpha V_\alpha\right) \cup \left(Y \setminus \bigcap_\beta C_\beta\right)$$
$$= Y \setminus \left(\bigcap_\beta C_\beta - \bigcup_\alpha V_\alpha\right) \subseteq_O Y,$$

   as $\bigcap_\beta C_\beta - \bigcup_\alpha V_\alpha$ is compact since it is a closed subset of a compact set; or
   c) $V_\beta = Y \setminus C_\beta$, where $C_\beta \subseteq_K X$. Then

$$\bigcup_\beta V_\beta = \bigcup_\beta (Y \setminus C_\beta) = Y \setminus (\bigcap_\beta C_\beta) \subseteq_O Y,$$

   since $\bigcap_\beta C_\beta \subseteq_K X$ whenever $C_\beta \subseteq_K X$.

The subspace topology on X agrees with the original topology on X. Indeed, in the subspace topology, open sets look like $V \cap X$, where $V \subseteq_O Y$. If $\infty \notin V$, then $V \cap X = V \subseteq_O X$ in the original topology.

On the other hand, if $\infty \in V$, $V = Y \setminus C$ for some compact C, and $V = Y \setminus C = X \setminus C$. But X is Hausdorff, so C is closed, and $V \subseteq_O X$ in the original topology. Conversely, every open set in the original topology is an open set of type 1 in the subspace topology.

Theorem 247

Let X be a non-compact locally compact Hausdorff space and $\infty \notin X$. Then $Y = X \cup \{\infty\}$ is compact Hausdorff with the topology defined above and $\overline{X} = Y$.

Proof: let $\mathfrak{F}$ be an open covering of Y. Then, there exists $F_0 \in \mathfrak{F}$ with $\infty \in F_0$. By definition, $C = Y \setminus F_0$ is a compact subset of X and $\mathfrak{F}' = \{F \cap X\}_{F \in \mathfrak{F}'}$ is an open covering of C in X. As C is compact, there is a finite sub-covering

$$\{F_1 \cap X, \ldots, F_n \cap X\}$$

of C. Hence $\{F_0, F_1, \ldots, F_n\}$ is a finite sub-covering of Y and Y is compact.

As X is not compact, $\{\infty\} = Y \setminus X$ is not open in Y. So every neighbourhood of ∞ looks like $Y \setminus C$, where $C \subsetneq_K X$, and so meets X. By definition, ∞ is a limit point of X in Y, so $\overline{X} = Y$.

We show now that Y is Hausdorff. If $x \neq y \in X$, there are open neighbourhoods in X satisfying the T_2 condition as X is Hausdorff. So suppose $x \in X$ and $y = \infty$. As X is locally compact, there is a compact set C and a neighbourhood V of x such that $x \in V \subseteq C$. Then $U = Y \setminus C$ is a neighbourhood of ∞ and $U \cap V = \varnothing$, which proves that Y is Hausdorff. ■

The space Y is the **one-point compactification of** X.

Examples (ONE-POINT COMPACTIFICATION)

1. Let $X = \mathbb{R}$. Then X is a non-compact locally compact Hausdorff space. By Theorem 247, there is a one-point compactification $Y = X \cup \{\infty\}$ of X. Y is in fact homeomorphic to
$$S^1 = \{(x, y) \mid x^2 + y^2 = 1\},$$
through the homeomorphism $f : S^1 \to Y$ defined by $f(x, y) = \frac{x}{1-y}$ whenever $y \neq 1$ and $f(0, 1) = \infty$.

2. Let $X = \mathbb{R}^2$. Then X is a non-compact locally compact Hausdorff space. By Theorem 247, there exists a one-point compactification Y, or $X \cup \{\infty\}$ of X. Y is in fact homeomorphic to
$$S^2 = \{(x, y, z) \mid x^2 + y^2 + z^2 = 1\},$$
through the homeomorphism $f : S^2 \to Y$ defined by
$$f(x, y, z) = \left(\frac{x}{1-z}, \frac{y}{1-z}\right)$$
whenever $z \neq 1$ and $f(0, 0, 1) = \infty$.

17.4 Solved Problems

1. Let $\mathfrak{B}$ be a basis for a topology on a space X. Show that a subset A of X is compact if and only if every covering of A by sets from $\mathfrak{B}$ has a finite subcovering.

Proof: if $A \subseteq_K X$, then every open covering of A contains a finite subcovering of A. But every covering of A by sets from $\mathfrak{B}$ is an open covering of A as all sets in $\mathfrak{B}$ are open, and so contains a finite subcovering of A.

Conversely, suppose that every covering of A by sets from $\mathfrak{B}$ contains a finite subcovering of A, and let $\mathfrak{U} = \{U_\gamma\}_{\gamma \in \Gamma}$ be an open covering of A. Since $U_\gamma \subseteq_O X$, and since $\mathfrak{B}$ is a basis for the topology on X, there exists, for each $\gamma \in \Gamma$, a subset $\mathfrak{B}_\gamma \subseteq \mathfrak{B}$ such that
$$U_\gamma = \bigcup_{B \in \mathfrak{B}_\gamma} B.$$
Thus, the collection $\{B | B \in \mathfrak{B}_\gamma \text{ for some } \gamma \in \Gamma\}$ is a covering of A by sets from $\mathfrak{B}$, and by hypothesis, it contains a finite subcovering of A, say $\{B_1, \ldots, B_n\}$. Now, for $1 \leq i \leq n$, choose $U_i \in \mathfrak{U}$ such that $B_i \subseteq U_i$. Then $\{U_1, \ldots, U_n\}$ is a finite subcovering of A, and $A \subseteq_K X$. ■

2. Let A and B be disjoint compact subsets of the Hausdorff space X. Show that there exist disjoint open sets U and V containing A and B, respectively.

Proof: assume that $A, B \neq \varnothing$, otherwise the statement is vacuously true. Let $b \in B$. Since X is Hausdorff and $A \cap B = \varnothing$, for every $a \in A$, there exists $U_{b,a}, V_{b,a} \subseteq_O X$, such that $a \in U_{b,a}$, $b \in V_{b,a}$ and $U_{b,a} \cap V_{b,a} = \varnothing$. The collection $\{U_{b,a}\}_{a\in A}$ is an open covering of $A \subseteq_K X$, and so we can extract from it a finite subcovering, say $\{U_{b,a_1}, \ldots, U_{b,a_n}\}$. Now, put

$$U(b) = \bigcup_{i=1}^{n} U_{b,a_i} \quad \text{and} \quad V(b) = \bigcap_{i=1}^{n} V_{b,a_i}.$$

Then $A \subseteq U(b), U(b), V(b) \subseteq_O X$ and $U(b) \cap V(b) \neq \varnothing$. This process can be repeated for every $b \in B$ so that $\{V(b)\}_{b\in B}$ covers B. Since $B \subseteq_K X$, we can extract a finite subcovering of B, say $\{V(b_1), \ldots, V(b_m)\}$. Let

$$V = \bigcup_{i=1}^{m} V(b_i) \quad \text{and} \quad U = \bigcap_{i=1}^{m} U(b_i).$$

Then $U, V \subseteq_O X$, $A \subseteq U$, $B \subseteq V$ and $U \cap V = \varnothing$. Indeed

$$\begin{aligned} U \cap V &= U \cap \left(\textstyle\bigcup_{i=1}^{m} V(b_i)\right) = \textstyle\bigcup_{i=1}^{m} (U \cap V(b_i)) \\ &\subseteq \qquad \bigcup_{i=1}^{n} (V(b_i) \cap U(b_i)) = \bigcup_{i=1}^{m} \varnothing = \varnothing, \end{aligned}$$

and the statement is proven. ■

3. Show that $[0,1]$ is not compact in $\mathbb{R}_l$. Is it compact in the countable complement topology on $\mathbb{R}$?

Proof: we prove the statement by exhibiting an open covering of $[0,1]$ from which it is impossible to extract a finite subcovering. Let

$$\mathfrak{U} = \{[1,2)\} \cup \{[0, 1-1/n)\}_{n\geq 2}.$$

Then $\mathfrak{U}$ is an open covering of $[0,1]$ since

$$[0,1) \subseteq \bigcup_{U\in\mathfrak{U}} U = [0,2),$$

and since $[a,b)$ is open in $\mathbb{R}_l$ for all $a < b$ in $\mathbb{R}$. Any subcovering of $[0,1)$ must contain $[1,2)$ as $1 \notin [0, 1-1/n)$ for all $n \geq 2$. Any finite subcovering must then look like

$$\mathfrak{V} = \{[1,2), [0, 1-1/n_1), \ldots, [0, 1-1/n_m)\},$$

where the n_i's are ordered such that $n_1 > n_2 > \ldots > n_m \geq 2$. With this ordering,

$$\bigcup_{i=1}^{m-1} [0, 1-1/n_i) \subseteq [0, 1-1/n_m).$$

However, $1 - 1/(n_m - 1) \notin [0, 1 - 1/n_m)$ and $1 - 1/(n_m - 1) \notin [1, 2)$. Any finite subcollection $\mathfrak{V}$ taken from $\mathfrak{U}$ cannot cover all of $[0, 1]$, so $[0, 1]$ is not compact in $\mathbb{R}_l$.

We show now that $[0, 1]$ is not compact in $\mathbb{R}$ with the finite complement topology. First, recall that a space is compact if and only if every family $\mathfrak{C} = \{C_\alpha\}$ of closed subsets having the finite intersection property, that is $\bigcap_{i=1}^n C_{\alpha_i} \neq \varnothing$ for all $C_{\alpha_i} \in \mathfrak{C}$, $1 \leq i \leq n$, has a non-empty intersection:

$$\bigcap_\alpha C_\alpha \neq \varnothing.$$

We construct a family of closed subsets having the finite intersection property, while their full intersection is empty. The closed subsets of $[0, 1]$ in this topology are the countable subsets of $[0, 1]$, as well as $[0, 1]$ itself. Now let

$$A_n = \left\{ \frac{1}{m} \right\}_{m \geq n} \subseteq [0, 1]$$

for all $n \in \mathbb{N}$. Each of the $A_n \neq \varnothing$ is countable and so closed in $[0, 1]$. Now, take $A_{n_1}, \ldots, A_{n_k}$, where $n_k > n_{k-1} > \ldots > n_1$. By construction,

$$\bigcap_{i=1}^k A_{n_i} = A_{n_k} \neq \varnothing.$$

But $\bigcap_{n \in \mathbb{N}} A_n = \varnothing$, since, otherwise, there would exist $m \in \mathbb{N}$ such that $\frac{1}{m} \leq \frac{1}{n}$ for all $n \in \mathbb{N}$, a contradiction. We can thus conclude that $[0, 1]$ is not a compact subspace in the finite complement topology. ■

4. Let X be a locally compact space. If $f : X \to Y$ is continuous, is the space $f(X)$ necessarily locally compact? What if f is both continuous and open?

Proof: let $X = \{-1\} \cup (0, 1)$ be a subspace of $\mathbb{R}$ and

$$T = \{(x, \sin(1/x)) \mid 0 < x < 1\} \cup \{(0, 0)\}$$

be a subspace of $\mathbb{R}^2$. This is the *topologist's sine curve*. Let $f : X \to T$ be the map sending -1 to $(0, 0)$ and x to $(x, \sin(1/x))$ for $0 < x < 1$. This map is continuous, since the pre-image of open subsets of T in the subspace topology are unions of open intervals in X, possibly with $\{-1\}$. Furthermore, $f(X) = T$. X is clearly locally compact at x for $0 < x < 1$. And $\{-1\}$ is a compact neighbourhood of $\{-1\}$, so that X is locally compact at -1. But T is not locally compact at $(0, 0)$. Indeed any candidate for a compact subset around $(0, 0)$ must contain an infinity of points who are as close as desired from the slice $\{0\} \times [-1, 1]$. Hence, no such sets are closed in $\mathbb{R}^2$, and so they can not be compact.

Suppose f is continuous and open. Then for any $y \in f(X)$ there exists $x \in X$ such that $f(x) = y$. The space X is locally compact so there is a compact set C_x and an open set U_x such that $x \in U_x \subseteq C_x$. Now, applying f yields

$$y = f(x) \in f(U_x) \subseteq f(C_x).$$

Since f is continuous and open $f(C_x)$ is compact and $f(U_x)$ is open. So $f(X)$ is locally compact at $f(x)$ for all $f(x) \in f(X)$, and $f(X)$ is locally compact. ■

5. Show that $[0,1]^\omega$ is not locally compact in the uniform topology.

Proof: Throughout, we assume that $[0,1]^\omega$ has the uniform topology. Let $\delta_{m,n}$ be the Kronecker δ, and let $d_m = (\delta_{m,n})_{n\in\mathbb{N}}$. Hence $d_m \in [0,1]^\omega$ for all $m \in \mathbb{N}$ and $d_U(d_m, d_k) = 1$ when $m \neq k$. Consequently, the sequence $d_1, d_2, \ldots$ has no convergent subsequence. Now, consider the open ball $B(x,r)$ in $[0,1]^\omega$. It contains a sequence $(x_n \pm \frac{rd_n}{2})_{n\in\mathbb{N}}$,[8] with no convergent subsequence. Hence $B(x,r)$ can not be contained in a compact set as compact set are sequentially compact in $[0,1]^\omega$. Thus $[0,1]^\omega$ is not locally compact in the uniform topology. ■

6. Let $\mathfrak{T}_P$, $\mathfrak{T}_U$, $\mathfrak{T}_B$ denote the product, uniform and box topologies respectively on $\mathbb{R}^\omega$.

   a) Show that $\mathfrak{T}_B$ is strictly finer than $\mathfrak{T}_U$.

   b) In which of the topologies are the following functions from $\mathbb{R}$ to $\mathbb{R}^\omega$ continuous?

      i. $f(t) = (t, 2t+1, 3t+2, 4t+3, \ldots)$

      ii. $g(t) = (t/2, t/3, t/4, t/5, \ldots)$

   c) In which of the topologies do the following sequences converge?

$$\begin{aligned}
x_1 &= (1,1,1,1,\ldots) & y_1 &= (1,0,0,0,\ldots)\\
x_2 &= (0,2^2,2^2,2^2,\ldots) & y_2 &= ((1/2)^2,(1/2)^2,0,0,\ldots)\\
x_3 &= (0,0,3^3,3^3,\ldots) & y_3 &= ((1/3)^3,(1/3)^3,(1/3)^3,0,\ldots)\\
x_4 &= (0,0,0,4^4,\ldots) & y_4 &= ((1/4)^4,(1/4)^4,(1/4)^4,(1/4)^4,\ldots)\\
&\vdots
\end{aligned}$$

$$\begin{aligned}
z_1 &= (1,1,0,0,\ldots)\\
z_2 &= ((1/2)^2,(1/2)^2,0,0,\ldots)\\
z_3 &= ((1/3)^2,(1/3)^2,0,0,\ldots)\\
z_4 &= ((1/4)^4,(1/4)^4,0,0,\ldots)\\
&\vdots
\end{aligned}$$

Solution:

a) Let $B(x,\varepsilon_x)$ be an open ball in the uniform topology. The set

$$B_x = \prod_{n\in\mathbb{N}} \left(x_n - \frac{\varepsilon_x}{4}, x_n + \frac{\varepsilon_x}{4}\right)$$

is open in the box topology, and $x \in B_x \subseteq B(x,\varepsilon_x)$. Indeed, let $z \in B_x$. Then $d_n(x_n,z_n) < \frac{\varepsilon_x}{2}$ for all $n \in \mathbb{N}$, so

$$d(x,z) = \text{l.u.b.}\{d_n(x_n,z_n)\} \le \frac{\varepsilon_x}{2} < \varepsilon_x,$$

[8] Select either one of $+$ or $-$ so that x_n stays in $[0,1]$.

and thus $z \in B(x, \varepsilon_x)$. Now suppose $x \neq y \in B(x, \varepsilon_x)$. As $B(x, \varepsilon_x)$ is open in the uniform (metric) topology, there exists $\varepsilon_y > 0$ such that $B(y, \varepsilon_y) \subseteq B(x, \varepsilon_x)$. Using the same reasoning as above yields

$$y \in B_y \subseteq B(y, \varepsilon_y) \subseteq B(x, \varepsilon_x),$$

where B_y is open in the box topology for all $y \in B(x, \varepsilon_x)$. Hence, around each point of $B(x, \varepsilon_x)$, we can fit an open set in the box topology, i.e. $B(x, \varepsilon_x)$ is open in the box topology and $\mathfrak{T}_U \subseteq \mathfrak{T}_B$.

We show that $\mathfrak{T}_U \subsetneq \mathfrak{T}_B$ by showing that $\mathbb{R}^\omega$ is not metrizable in the box topology. Since $\mathbb{R}^\omega$ has a metric in the uniform topology, $\mathfrak{T}_B$ is strictly finer than $\mathfrak{T}_U$. Let $X = \mathbb{R}^\omega$ and $A = (0,1)^\omega$. Clearly $0 = (0,0,0,\ldots) \in \overline{A}$ since, in the box topology every neighbourhood of 0 contains positive sequences. However, there is no sequence $x_n \in A$ such that $x_n \to 0$. Suppose x_n is a sequence in A. Then,

$$\begin{aligned} x_1 &= (x_{1,1}, x_{1,2}, x_{1,3}, \ldots) \\ x_2 &= (x_{2,1}, x_{2,2}, x_{2,3}, \ldots) \\ x_3 &= (x_{3,1}, x_{3,2}, x_{3,3}, \ldots) \\ &\vdots \end{aligned}$$

Let $\varepsilon < 0$, and construct the open set (in the box topology)

$$U_\varepsilon = \prod_{m \in \mathbb{N}} (\varepsilon, y_m),$$

where $0 < y_m < x_{m,m}$ for all $m \in \mathbb{N}$. By construction, U_ε is a neighbourhood of 0 in the box topology, and $x_n \notin U_\varepsilon$ for all $n \in \mathbb{N}$. Hence, x_n can not converge to 0. By the Sequence Lemma, $\mathbb{R}^\omega$ (in the box topology) is not metrizable.

b) Both of the functions are continuous in the product topology as each of the components are continuous. In the uniform topology, f is not continuous. Indeed, let $\varepsilon = 1/2$. Then, for every $\delta > 0$,

$$d_U(f(x), f(x+\delta)) = \text{l.u.b.}\{\min\{n\delta, 1\}\} = 1 > \varepsilon.$$

In the box topology f is not continuous. Indeed, let

$$U = \prod_{n \in \mathbb{N}} \left(\frac{n^2 - n - 1}{n}, \frac{n^2 - n + 1}{n} \right).$$

U is open, but $f^{-1}(U) = \{0\}$[9] which is closed in $\mathbb{R}$. Similarly, g isn't continuous in the box topology. Let

$$V = \prod_{n \in \mathbb{N}} \left(-\frac{1}{(n+1)^2}, \frac{1}{(n+1)^2} \right).$$

[9] $t \in f^{-1}(U) \iff f(t) \in U \iff \frac{n^2-n-1}{n} < nt + (n-1) < \frac{n^2-n+1}{n}\ \forall n \iff -\frac{1}{n} < nt < \frac{1}{n}\ \forall n \iff t = 0.$

V is open, but $g^{-1}(V) = \{0\}$[10] which is closed in $\mathbb{R}$. But it is continuous in the uniform topology. Indeed, let $\varepsilon > 0$ and put $\delta = 2\varepsilon$. If $|x-y| < \delta$, then

$$d_Y(g(x), g(y)) < \text{l.u.b.}\{\min\{\delta/(n+1), 1\}\} = \frac{1}{2}\delta = \varepsilon.$$

c) All three sequences have to converge to $0 = (0, 0, 0...)$ if they converge at all. They all converge in the product topology. Indeed, suppose U is a basic neighbourhood of 0 in the product topology. Then

$$U = U_1 \times U_2 \times \cdots \times U_m \times \mathbb{R} \times \mathbb{R} \times \cdots$$

for some $m \in \mathbb{N}$, and where each of the $U_i =]a_i, b_i[$ are basic neighbourhood of 0 in $\mathbb{R}$. The sequence x_n lies in U for $n > m$, y_n lies in U for all n such that $\left(\frac{1}{n}\right)^n < \min_{1\leq i\leq m}\{b_i\}$,[11] and z_n lies in U for all

$$n > \frac{1}{(\min_{1\leq i\leq m}\{b_i\})^2}.$$

Let's look at what happens in the uniform topology. The sequence x_n does not converge to 0. Indeed, let $B(0,\varepsilon)$ be a ε-neighbourhood of 0, so

$$B(0,\varepsilon) \quad = \{\xi \mid \text{l.u.b.}|\xi_i| < \varepsilon\}.$$

For the sequence x_n,

$$\text{l.u.b.}\{\min\{|x_{nj}|, 1\}\} = 1,$$

which is bigger than every $\varepsilon < 1$. Hence, there does not exist a N for which $x_n \in B(0,\varepsilon)$ when $n > N$, and $\varepsilon < 1$. At the same time, (x_n) does not converge in the box topology. For y_n, z_n, all elements of the sequence are less than 1 for large enough n, so we can forget about the metric being bounded, and

$$\begin{aligned} \text{l.u.b.}\{|y_{nj}|\} &= (1/n)^n \\ \text{l.u.b.}\{|z_{nj}|\} &= (1/n)^2. \end{aligned}$$

For these least upper bounds, there exists N such that $y_n, z_n \in B(0,\varepsilon)$ whenever $n > N$, so the sequences converge to 0 in the uniform topology.[12] In the box topology, (y_n) doesn't converge, but (z_n) does. $\square$

[10] $t \in g^{-1}(V) \iff g(t) \in V \iff -\frac{1}{(n+1)^2} < \frac{t}{n+1} < \frac{1}{(n+1)^2}\ \forall n \iff -\frac{1}{n+1} < t < \frac{1}{n+1}\ \forall n \iff t = 0.$

[11] As $f(x) = x^{-x}$ is eventually decreasing, y_n is eventually in U for all $n > N$.

[12] Wait, you say. For this sequence to converge to 0, every neighbourhhod of 0 must contain all y_n when $n > N$ for some N. Ah, but every neighbourhood of 0 contains $B(0,\varepsilon)$ for some $\varepsilon > 0$, and this ball contains all y_n when $n > N$, so the original neighbourhood did as well...

17.5 Exercises

1. Prepare a 2-page summary of this chapter, with important definitions and results.
2. Let X be the subspace
$$X = \left\{ \frac{t}{1+t} e^{it} \,\middle|\, t \geq 0 \right\} \cup \{e^{i\pi}\}.$$
Give detailed answers to the following:
   a) Is X compact?
   b) Is X locally compact?
3. Prove that if Y is compact and N is an open set in $X \times Y$ containing $\{x_0\} \times Y$, then there is a neighbourhood W of x_0 such that $W \times Y \subseteq N$.
4. If Y is compact, show that the projection $\pi_1 : X \times Y \to X$ is closed.
5. Prove Theorem 234 (Reprise) and Corollary 235.
6. Show that the standard bounded metric $\overline{d}$ and the uniform metric are indeed metrics on (X, d).

Chapter 18

Countability and Separation

In Chapter 15, we introduced a few simple separation definitions (T_0, T_1, and T_2/Hausdorff); in this chapter, we extend the discussion to more sophisticated **separation** axioms, and introduce the notions of **first** and **second countable** spaces.

18.1 Countability Axioms

A **basis at** $x \in X$ is a collection $\mathfrak{B}$ of open sets containing x and such that, for each neighbourhood V of x, there exists $B \in \mathfrak{B}$ with $x \in B \subseteq V$.

We say that a space X is **first countable at** $x \in X$ if there is a **countable** basis at x; X is simply **first countable** if it is first countable at **every** $x \in X$. It is **second countable** if its topology has a countable basis.

Examples (First and Second Countability)

1. If X is second countable, then it has a countable basis $\mathfrak{B}$. Let $x \in X$. If U is an (open) neighbourhood of x, then

$$U = \bigcup_{n \in \mathbb{N}} B_n,$$

where $B_n \in \mathfrak{B}$ for all $n \in \mathbb{N}$. As $x \in U$, then $x \in B_m$ for some m. Hence

$$\mathfrak{B}_x = \{B \in \mathfrak{B} \mid x \in B\}$$

is a countable basis at x since $\mathfrak{B}_x \subseteq \mathfrak{B}$, and so X is first countable at x. But x is arbitrary, so X is first countable.

2. Let $X = \mathbb{R}$ in the usual topology; it is second countable since
$$\mathfrak{B} = \{(a,b) \mid a < b \in \mathbb{Q}\}$$
is a countable basis of X. In light of the previous example, X is also first countable.

3. Let $X = \mathbb{R}$ in the discrete topology. $\mathbb{R}$ is not second countable as every set is open and $\mathbb{R}$ is uncountable. However, it is first countable, since $\{x\}$ is a basis at x for each $x \in X$. □

A space X is **Lindelöf** if every open covering of X contains a **countable** (not necessarily finite) sub-covering. A subset A of X is **Lindelöf** if it is Lindelöf in the subspace topology.

Theorem 248
If X is second countable, then it is Lindelöf.

Proof: let $\mathfrak{F}$ be an open covering and $\mathfrak{B} = \{B_n\}_n$ be a countable basis of X. For each $n \in \mathbb{N}$, whenever it is possible to do so, let $F_n \in \mathfrak{F}$ be such that $B_n \subseteq F_n$. Otherwise let $F_n = \varnothing$. Then
$$X = \bigcup_{n\in\mathbb{N}} B_n \subseteq \bigcup_{n\in\mathbb{N}} F_n$$
and $\{F_n\}_n$ is a countable sub-collection; it is also an open cover. Indeed, let $x \in X$. Then there exists $F \in \mathfrak{F}$ such that $x \in F$. As F is open, there exists a basic set $B_n \in \mathfrak{B}$ such that $x \in B_n \subseteq F$. By construction, $F \in \{F_n\}_n$. Hence X is Lindelöf. ■

Let us take a look at some examples.

Examples (LINDELÖF SPACES)

1. The space $\mathbb{R}$ is second countable (hence Lindelöf), since
$$\mathfrak{B} = \{(a,b) \mid a < b \in \mathbb{Q}\}$$
is a countable basis.

2. The space $\mathbb{R}_l$ is Lindelöf but not second countable. Indeed, let $\mathfrak{B}$ be any basis for the lower limit topology on $\mathbb{R}_l$. Then, for any $x \in \mathbb{R}$ and $\varepsilon > 0$, we have $[x, x+\varepsilon) \subseteq_O \mathbb{R}_l$, that is, there is a basic set $B_{x,\varepsilon}$ such that
$$x \in B_{x,\varepsilon} \subseteq [x, x+\varepsilon).$$

If $x < y$, then, for $\varepsilon = y - x$, $y \notin B_{x,\varepsilon}$. So $\mathfrak{B}$ must contain an uncountable sub-collection and $\mathbb{R}_l$ is not second countable.

We show that $\mathbb{R}_l$ is Lindelöf by showing that every open covering by basic sets contains a countable sub-covering. Let $\mathfrak{F} = \{[\alpha_a, \beta_a)\}_{a \in J}$ be an open covering of $\mathbb{R}_l$ and

$$C = \bigcup_{a \in J} (\alpha_a, \beta_a)$$

be a subspace of $\mathbb{R}$. As $\mathbb{R}$ is second countable, so is C; it is thus also Lindelöf, as of Theorem 248. The collection $\{(\alpha_a, \beta_a)\}_{a \in J}$ is an open covering of C, so there exists a sub-covering $\{(\alpha_{a_n}, \beta_{a_n})\}_{n \in \mathbb{N}}$ of C. Then

$$\mathfrak{F}' = \{[\alpha_{a_n}, \beta_{a_n})\}_{n \in \mathbb{N}}$$

also covers C and $\mathfrak{F}' \cup (\mathbb{R} \setminus C)$ is a covering of $\mathbb{R}$.

Let $x \in \mathbb{R} \setminus C$. Then $x = \alpha_a$ for some $a \in J$. Let $q_x \in (\alpha_a, \beta_a) \cap \mathbb{Q}$. Then

$$(x, q_x) \subseteq (\alpha_a, \beta_a) \subseteq C.$$

Now suppose $x < y \in \mathbb{R} \setminus C$. Necessarily, $q_x < q_y$ since, otherwise,

$$y \in (x, q_y) \subseteq (x, q_x) \subseteq C,$$

a contradiction as $y \notin C$. Thus the map $x \mapsto q_x$ is an injection of $\mathbb{R} \setminus C$ into $\mathbb{Q}$, which means that $\mathbb{R} \setminus C$ is countable. Write $\mathbb{R} \setminus C = \{z_n\}_{n \in \mathbb{N}}$, and find $[\alpha_m, \beta_m) \in \mathfrak{F}$ with $z_m \in [\alpha_m, \beta_m)$ for all $m \in \mathbb{N}$ – this can be done as $\mathfrak{F}$ is an open covering of $\mathbb{R}_l$. Then $\mathfrak{F}' \cup \{[\alpha_m, \beta_m)\}_{m \in \mathbb{N}}$ is a countable sub-cover of $\mathbb{R}_l$ extracted from $\mathfrak{F}$.

3. The space $\mathbb{R}_l^2$ is not Lindelöf. To show this, let $L = \{(x, -x)\}_{x \in \mathbb{R}}$. Then $\mathbb{R}_l^2 \setminus L$ is open in $\mathbb{R}_l^2$. Indeed, let $(x, y) \in \mathbb{R}_l^2 \setminus L$ and put $\varepsilon = \frac{x+y}{2}$. Then

$$(x, y) \in [x, x + \varepsilon) \times [y, y + \varepsilon)$$

and $([x, x + \varepsilon) \times [y, y + \varepsilon)) \cap L = \varnothing$. Now $\mathfrak{F} = \{\mathbb{R}_l^2 \setminus L\} \cup \{F_a\}_{a \in \mathbb{R}}$, where

$$F_a = [a, a + 1) \times [-a, -a + 1),$$

is an open covering of $\mathbb{R}_l^2$. But F_a is the only set in $\mathfrak{F}$ containing $(a, -a)$, so any sub-covering will contain F_a for all $a \in \mathbb{R}$. As $\mathbb{R}$ is uncountable, $\mathfrak{F}$ does not contain a countable sub-covering. Hence $\mathbb{R}_l^2$ is not Lindelöf. This demonstrates that the product of two Lindelöf spaces need not be Lindelöf.

4. Let Ω be the first uncountable ordinal. Then $\Omega = [0, \Omega)$ is first countable but not Lindelöf, so it is not second countable. Indeed, suppose $a \in \Omega$. Then

$$\mathfrak{B}_a = \{(c, a+1)\}_{c<a}$$

is countable as $a < \Omega$. Let U be a neighbourhood of a. Then $a \in (c, a+1) \subseteq U$ for $c < a$. This makes $\mathfrak{B}_a$ a countable basis at $a \in \Omega$. Then Ω is first countable.

To show that Ω is not second countable, consider the open covering $\mathfrak{F} = \{[0, b)\}_{b \in \Omega}$ of Ω, and let $\mathfrak{F}'$ be any countable sub-collection from $\mathfrak{F}$. Let

$$\beta = \bigcup_{[0,b) \in \mathfrak{F}'} b.$$

As it is a countable union of countable sets, β is countable, that is $\beta \in \Omega$. But $\beta \notin [0, b)$ for all $[0, b)$ in $\mathfrak{F}'$, and so $\mathfrak{F}'$ cannot be a sub-covering from $\mathfrak{F}$. Hence Ω is not Lindelöf, nor is it second countable. □

We can show fairly easily that countability behaves as expected for subspaces and products.

Theorem 249
If X is first (resp. second) countable, then any subspace of X is first (resp. second) countable. If X_n is first (resp. second) countable for all $n \in \mathbb{N}$, then

$$\prod_{n \in \mathbb{N}} X_n$$

is first (resp. second) countable.

Proof: the statement about subspaces is clearly true. We show that the countable product of second countable spaces is second countable. The proof for first countable spaces is similar, and is left as an exercise.

Let $X = \prod X_n$, $\mathfrak{B}_n$ be a countable basis for X_n, and define

$$\mathfrak{C}_m = \left\{ \prod_{n \in \mathbb{N}} V_n \,\middle|\, V_n \in \mathfrak{B}_n \text{ for } 0 \leq n \leq m, V_n = X_n \text{ for } m < n \right\}$$

for all $m \in \mathbb{N}$. Then $\mathfrak{C} = \bigcup_{m \in \mathbb{N}} \mathfrak{C}_m$ is countable. Furthermore, it is a basis for the product topology on X. So X is second countable. ■

We shall see in the next section that there is a link between countability and separation.

18.2 Separation Axioms

Let X be a space. In Chapter 15, we introduced a number of separation axioms:

0. X is T_0 if for every pair $x \neq y \in X$, there exist a neighbourhood U of either x or y that misses the other;
1. X is T_1 if for every pair $x \neq y \in X$, there exist neighbourhoods U_x of x and U_y of y such that $y \notin U_x$ and $x \notin U_y$;
2. X is T_2 or **Hausdorff** if for every pair $x \neq y \in X$, there exist disjoint neighbourhoods U_x of x and U_y of y.

We have also seen that if a space X is T_1, then every singleton is closed in X. Note that the condition T_2 is strictly stronger than the condition T_1: there are T_1 spaces that fail to be T_2.

We introduce two new **separation axioms.**[1] We say that a space X is:

3. T_3 or **regular** if X is T_1 and if for every pair consisting of a $x \in X$ and a closed set B disjoint from x, there exist disjoint neighbourhoods U_x of x and U_B containing B;
4. T_4 or **normal** if X is T_1 and if for every pair consisting of disjoint closed sets A and B, there exist disjoint neighbourhoods U_A containing A and U_B containing B.

Some of the conditions imply some of the others: a regular space is Hausdorff, for instance, since singletons are closed. Indeed let $x \neq y$. Then x and the closed set $\{y\}$ are disjoint and there exist U_x and $U_{\{y\}}$ such that $x \in U_x$, $\{y\} \subseteq U_{\{y\}}$ and $U_x \cap U_{\{y\}} = \varnothing$. For the same reasons, a normal space is regular. The following examples (without proof) show that none of the implications

$$T_4 \Longrightarrow T_3 \Longrightarrow T_2 \Longrightarrow T_1 \Longrightarrow T_0$$

can be reversed and that normal spaces are not as well behaved as we might expect.

Example: (Regularity and Normality)

1. Let $K = \left\{\frac{1}{n} : n \in \mathbb{N}\right\}$ be a subset of $\mathbb{R}$, with basic open sets of the form (a,b) and $(a,b) \setminus K$ for all $a,b \in \mathbb{R}$. With this topology $\mathbb{R}$ is Hausdorff. But it is not regular, since it is possible to separate the point 0 and the closed set K. Hence a Hausdorff space need not be regular.
2. Let Ω be the least uncountable ordinal. The spaces Ω and Ω^+ are normal in the order topology. But their product is not normal. The product $\Omega^+ \times \Omega^+$ is normal however, so a subspace of a normal space need not be normal. And, as we will see later, $\Omega \times \Omega^+$ is regular, being the product of two regular spaces, so a regular space need not be normal.
3. If A is uncountable, the product space $\mathbb{R}^A$ is not normal.

[1]There are other such axioms; see `https://en.wikipedia.org/wiki/Separation_axiom` for more.

We can also formulate the conditions of regularity and normality differently, as that could be more useful in different contexts (as in the next section).

Lemma 250
Let X be T_1. Then

1. *X is regular if and only if given a point of $x \in X$ and a neighbourhood U of x, there is a neighbourhood V of x such that $\overline{V} \subseteq U$.*

2. *X is normal if and only if given a closed set $A \subseteq X$ and an open set U containing A, there is an open set V containing A such that $\overline{V} \subseteq U$.*

Proof:

1. Suppose X is regular and let $x \in X$ have a neighbourhood U. Then

$$x \notin X \setminus U \subseteq_C X.$$

By regularity of X, there exist open subsets V and W such that $x \in V$, $X \setminus U \subseteq W$ and $V \cap W = \varnothing$. Suppose $y \in W$. Then W is a neighbourhood of y that does not meet V, and so $\overline{V} \cap W = \varnothing$. Hence $\overline{V} \subseteq X \setminus W = U$.

Conversely, suppose $B \subseteq_C X$ and $x \notin B$. then

$$x \in X \setminus B \subseteq_O X.$$

By hypothesis, there exists a neighbourhood V of x such that

$$x \in \overline{V} \subseteq X \setminus B.$$

Then by construction, $B \subseteq X \setminus \overline{V} \subseteq_O X$, $x \in V$ and $X \setminus \overline{V} \cap V = \varnothing$. In other words, X is regular.

2. The proof of the second statement uses sensibly the same argument. ■

T_2 and T_3 spaces behave particularly well with respect to subspaces and products.

Theorem 251 Let W, $\{W_\alpha\}$ be Hausdorff, X, $\{X_\beta\}$ be regular.

1. Each subspace Y of W is Hausdorff, and the product $\prod W_\alpha$ is Hausdorff.

2. Each subspace Y of X is regular, and the product $\prod X_\beta$ is regular.

Proof:

1. Let Y be a subspace of W. If $x \neq y \in Y$, then there exist disjoint $U, V \subseteq_O X$ such that $x \in U$ and $y \in V$. But $U \cap Y, V \cap Y \subseteq_O Y$ are disjoint and $x \in U \cap Y$, $y \in V \cap Y$, so Y is Hausdorff.

   Let $W = \prod W_\alpha$. If $x = (x_\alpha) \neq y = (y_\alpha)$, then there is a coordinate γ such that $x_\gamma \neq y_\gamma$. As W_γ is Hausdorff, there exist disjoint $U, V \subseteq_O X_\gamma$ such that $x_\gamma \in U$, $y_\gamma \in V$. Then $\pi_\gamma^{-1}(U), \pi_\gamma^{-1}(V) \subseteq_O W$ are disjoint and $x \in \pi_\gamma^{-1}(U)$, $y \in \pi_\gamma^{-1}(V)$, so W is Hausdorff.

2. Let Y be a subspace of X. Since Y is Hausdorff, one point sets are closed in Y. If $x \in Y$, and B is a closed subset of Y disjoint from x, then

   $$\overline{B} \cap Y \subseteq \overline{B \cap Y} = B \cap Y = B.$$

   So $x \notin \overline{B}$ (in X). By regularity of X, there exist disjoint $U, V \subseteq_O X$ such that $x \in U$ and $\overline{B} \subseteq V$. Then $U \cap Y, V \cap Y \subseteq_O Y$ are disjoint, $x \in U \cap Y$ and $B \subseteq V \cap Y$. Hence Y is regular.

   Let $X = \prod X_\beta$. Since X is Hausdorff, one point sets are closed in X. Let $x = (x_\beta) \in X$ and suppose U is a neighbourhood of x. Choose a basis $\prod U_\beta$ such that

   $$x \in \prod U_\beta \subseteq U.$$

   For each β, X_β is regular. Then there exists a neighbourhood V_β such that $x_\beta \in \overline{V_\beta} \subseteq U_\beta$.[2] Then, $V = \prod V_\beta$ is a neighbourhood of $x \in X$. But $\overline{V} = \prod \overline{V_\beta}$, so

   $$\overline{V} \subseteq \prod U_\beta \subseteq U$$

   and so X is regular according to Lemma 250. ■

The following three theorems give sets of hypotheses under which **normality** is assured.

Theorem 252
Let X be metrizable. Then X is normal.

Proof: let d be the metric on X, and A and B be disjoint closed subsets of X. For each $a \in A$, choose ε_a such that $B(a, \varepsilon_a) \cap B = \varnothing$ – this can always be done as B is closed in X so $X \setminus B \subseteq_O X$.

Similarly, for each $b \in B$, choose ε_b such that $B(b, \varepsilon_b) \cap A = \varnothing$. Then

$$U = \bigcup_{a \in A} B(a, \varepsilon_a/2) \quad \text{and} \quad V = \bigcup_{b \in B} B(b, \varepsilon_b/2)$$

are open subsets of X containing A and B respectively. They are also disjoint. Otherwise, $B(a, \varepsilon_a/2) \cap B(b, \varepsilon_b/2) \neq \varnothing$ for some $a \in A$, $b \in B$. Suppose z lies in that intersection. Then

$$d(a,b) \leq d(a,z) + d(z,b) < \frac{\varepsilon_a + \varepsilon_b}{2}.$$

If $\varepsilon_a \leq \varepsilon_b$, then $d(a,b) < \varepsilon_b$ and $a \in B(b, \varepsilon_b)$. If $\varepsilon_a \geq \varepsilon_b$, then $d(a,b) < \varepsilon_a$ and $b \in B(a, \varepsilon_a)$. But both these statements are false, so $U \cap V = \varnothing$ and X is normal. ■

As usual, compact Hausdorff space behave nicely.

Theorem 253
Let X be a compact Hausdorff space. Then X is normal.

Proof: see the solved problems. ■

We establish a link with second countability below.

Theorem 254
Let X be a second countable regular space. Then X is normal.

Proof: let $\mathfrak{B}$ be a countable basis for X. Suppose A and B are disjoint closed subsets of X. As B is closed, each $x \in A$ has a neighbourhood U_x not meeting B. By regularity, there is a neighbourhood V_x of x such that

$$x \in \overline{V_x} \subseteq U.$$

As $V_x \subseteq_O X$, there exists $W_x \in \mathfrak{B}$ such that $x \in W_x \subseteq V_x$, and

$$\overline{W_x} \subseteq U_x \subseteq X \setminus B,$$

so $\overline{W_x} \cap B = \varnothing$. Then $\{W_x\}_{x \in A}$ is a countable open covering of A since it is contained in $\mathfrak{B}$. Let us re-index it, and write $\{W_n\}_{n \in \mathbb{N}}$. Similarly, it is possible to construct a countable open covering $\{Z_n\}_{n \in \mathbb{N}}$ of B such that $\overline{Z_n} \cap A = \varnothing$ for all $n \in \mathbb{N}$. Given $n \in \mathbb{N}$, define

$$W_n' = W_n \setminus \bigcup_{i=1}^{n} \overline{Z_i} \quad \text{and} \quad Z_n' = Z_n \setminus \bigcup_{i=1}^{n} \overline{W_i}.$$

Then $W'_n, Z'_n \subseteq_O X$ as $W_n, Z_n \subseteq_O X$ and $\bigcup_{i=1}^n \overline{W_i}, \bigcup_{i=1}^n \overline{Z_i} \subseteq_C X$. Let

$$W' = \bigcup_{n\in\mathbb{N}} W'_n \quad \text{and} \quad Z' = \bigcup_{n\in\mathbb{N}} Z'_n.$$

Then $W', Z' \subseteq_O X$ and $A \subseteq W'$ and $B \subseteq Z'$. Indeed if $x \in A$, then $x \in W_n$ for some n. But, by construction, $x \notin \overline{Z_i}$ for all $i \in \mathbb{N}$. Then $x \in W'_n$. Similarly, if $y \in B$, then $y \in Z'_n$ for some $n \in \mathbb{N}$. It remains only to show that $W' \cap Z' = \varnothing$.

Suppose $\xi \in W' \cap Z'$. Then $\xi \in W'_n \cap Z'_m$ for some $m, n \in \mathbb{N}$. If $m \geq n$, then

$$\xi \in W'_n \Longrightarrow \xi \in W_n, \quad \text{and} \quad \xi \in Z'_m \Longrightarrow \xi \notin \overline{W_n},$$

which is a contradiction. If $m \leq n$, then

$$\xi \in W'_n \Longrightarrow \xi \notin \overline{Z_m}, \quad \text{and} \quad \xi \in Z'_m \Longrightarrow \xi \in Z_m,$$

another contradiction. Then $W' \cap Z' = \varnothing$ and X is normal. ■

18.3 Results of Urysohn and Tietze

Let X be a normal space, and A, B be disjoint closed subsets of X. Put $U_1 = X \setminus B \subseteq_O X$, so $A \subseteq U_1$. As X is normal, there exists $U_0 \subseteq_O X$ such that $A \subseteq U_0$ and $\overline{U_0} \subseteq U_1$. For each dyadic rational $r = \frac{m}{2^n}$ in $[0,1]$, we can associate an open set U_r such that

$$r < s \Longrightarrow \overline{U_r} \subseteq U_s. \tag{18.1}$$

To do so, we start with any $U_{\frac{1}{2}} \subseteq_O X$ such that

$$\overline{U_0} \subseteq U_{\frac{1}{2}} \subseteq \overline{U_{\frac{1}{2}}} \subseteq U_1$$

(this can be done as X is normal, $U_1 \subseteq_O X$, $\overline{U_0} \subseteq_C X$, and $\overline{U_0} \subseteq U_1$). Then, by the same process, it is possible to obtain $U_{\frac{1}{4}}, U_{\frac{3}{4}} \subseteq_O X$ satisfying

$$\overline{U_0} \subseteq U_{\frac{1}{4}} \subseteq \overline{U_{\frac{1}{4}}} \subseteq U_{\frac{1}{2}} \quad \text{and} \quad \overline{U_{\frac{1}{2}}} \subseteq U_{\frac{3}{4}} \subseteq \overline{U_{\frac{3}{4}}} \subseteq U_1.$$

Recursively, suppose we have sets $U_{\frac{m}{2^n}}$ satisfying (18.1), for $m = 0, 1, \ldots, 2^n$.

Then $\overline{U_{\frac{m}{2^n}}} \subseteq U_{\frac{m+1}{2^n}}$ for all $m = 0, 1, \ldots, 2^n - 1$. By normality of X, for $m = 0, 1, \ldots, 2^n - 1$, there is an set $U_{\frac{2m+1}{2^{n+1}}} \subseteq_O X$ such that

$$\overline{U_{\frac{m}{2^n}}} \subseteq U_{\frac{2m+1}{2^{n+1}}} \subseteq \overline{U_{\frac{2m+1}{2^{n+1}}}} \subseteq U_{\frac{m+1}{2^n}}.$$

Let r be a **dyadic rational** not in $[0,1]$.[3] If $r > 1$, take $U_r = X$. If $r < 0$, take $U_r = \varnothing$. Then (18.1) holds for all dyadic rational.

[3]A dyadic rational is a real number that can be written as a fraction with denominator 2^q for some non-negative integer q.

Now, let $x \in X$, and define

$$Q(x) = \{p \mid x \in U_p\}.$$

For all $x \in X$, $p \notin Q(x)$ whenever $p < 0$ since $x \notin U_p = \varnothing$, and $q \in Q(x)$ whenever $q > 1$ since $x \in U_p = X$. Hence $Q(x)$ is bounded below and its **greatest lower bound** lies in $[0, 1]$. Define $f : X \to [0, 1]$ by

$$f(x) = \text{g.l.b.}\{Q(x)\}.$$

Then, $f(a) = 0$ for $a \in A$ since $Q(a)$ is the set of dyadic rational in $[0, \infty)$, and $f(b) = 1$ for $b \in B$ since $Q(b)$ is the set of dyadic rational in $(1, \infty)$. By construction,

1. $x \in \overline{U_p} \implies f(x) \leq p$;
2. $x \notin U_p \implies f(x) \geq p$.

Indeed, if $x \in \overline{U_p}$, then $x \in U_q$ for all $q > p$. Then $q \in Q(x)$ for all $q > p$, and so that $f(x) \leq p$. If $x \notin U_p$, then p is a lower bound for $Q(x)$ so that $f(x) \geq p$.

Theorem 255 (URYSOHN LEMMA)
The function f defined above is continuous.

Proof: suppose $x_0 \in X$ and (a, b) is a neighbourhood of $f(x_0)$. We find a set $U \subseteq_O X$ such that $x_0 \in U \subseteq f^{-1}((a, b))$. Choose two dyadic rationals $p < q$ such that $a < p < f(x_0) < q < b$. Let $U = U_q \setminus \overline{U_p}$. Then $U \subseteq_O X$ as U_q is open and $\overline{U_p}$ is closed. Since $q > f(x_0) > p$, we have $x_0 \notin \overline{U_p}$ and $x_0 \in U_q$, so $x_0 \in U$.

If $x \in U$, then $x \in U_q \subseteq \overline{U_q}$ and $f(x) \leq q < b$; but $x \notin \overline{U_p}$ so $x \notin U_p$ and $a < p \leq f(x)$. Thus $f(U) \subseteq (a, b)$ and $U \subseteq f^{-1}((a, b))$, so $f^{-1}((a, b)) \subseteq_O X$ and f is continuous. ■

We have shown that in any normal space X, it is possible to separate any two **disjoint closed sets** A and B by a continuous function $f : X \to [0, 1]$, where $f(a) = 0$ for all $a \in A$ and $f(b) = 1$ for all $b \in B$. Note that this does not necessarily mean that $f^{-1}(\{0\}) = A$ and $f^{-1}(\{1\}) = B$. This prompts the following definition.

A T_1-space is $T_{3\frac{1}{2}}$ or **completely regular** if, given a point x_0 and a closed subset A with $x_0 \notin A$, there is a continuous function $f : X \to [0, 1]$ such that $f(x_0) = 0$ and $f(a) = 1$ for all $a \in A$. Suppose X is completely regular. Let x_0 and A be a closed subset of X such that $x_0 \notin A$. Then there is a continuous function $f : X \to [0, 1]$ with $f(x_0) = 0$ and $f(a) = 1$ for all $a \in A$. Define

$$U = f^{-1}([0, 1/3)) \quad \text{and} \quad V = f^{-1}((2/3, 1]).$$

Then $U, V \subseteq_O X$, $x_0 \in U$, $A \subseteq V$, and $U \cap V = \varnothing$, and so X is regular.

One of the most important corollaries of the Urysohn lemma is the Tietze extension theorem. Before stating and proving it, we first prove the following useful lemma.

Lemma 256
Let X be a normal space and $A \subseteq_C X$. If $h : A \to [-r, r]$ is continuous, then there is a continuous function $g : X \to \left[-\frac{r}{3}, \frac{r}{3}\right]$ such that $|h(x) - g(x)| \leq \frac{2}{3}r$ for all $x \in A$.

Proof: let $B = h^{-1}([r/3, r])$ and $C = h^{-1}([-r, -r/3])$. Then $B, C \subseteq_C A$ as h is continuous, so $B, C \subseteq_C X$ as $A \subseteq_C X$, and $B \cap C = \varnothing$. Since X is normal, we can use the Urysohn lemma to construct a continuous function $g : X \to \left[-\frac{r}{3}, \frac{r}{3}\right]$ such that $g(b) = \frac{r}{3}$ for all $b \in B$ and $g(c) = -\frac{r}{3}$ for all $c \in C$. Now, let $x \in A$.

Then there are three cases:

1. If $x \in B$, then $r \geq h(x) \geq \frac{r}{3} = g(x)$, so $\frac{2}{3}r \geq h(x) - g(x) \geq 0$.
2. If $x \in C$, then $-r \leq h(x) \leq -\frac{r}{3} = g(x)$, so $\frac{2}{3}r \geq g(x) - h(x) \geq 0$.
3. If $x \in A \setminus (B \cup C)$ then $|h(x)| < r$ and $|g(x)| \leq \frac{r}{3}$, so
$$|h(x) - g(x)| \leq |h(x)| + |g(x)| \leq 2r/3.$$

Hence $|h(x) - g(x)| \leq 2r/3$ whenever $x \in A$. ■

We are now ready to prove the extension result.

Theorem 257 (Tietze Extension Theorem)
Let X be a normal space and A a closed subset of X.

1. *If $f : A \to [a, b]$ is continuous, there is a continuous function $g : X \to [a, b]$ such that $g|_A = f$.*
2. *If $f : A \to \mathbb{R}$ is continuous, there is a continuous function $g : X \to \mathbb{R}$ such that $g|_A = f$.*

Proof:

1. It is sufficient to prove the theorem for $a = -1$, $b = +1$, as $[-1, 1]$ is homeomorphic to $[a, b]$ for all $a < b \in \mathbb{R}$. Let $r = 1$, $h = f$ and apply Lemma 256 to get a a continuous function g_1 on X such that
$$|g_1(x)| \leq 1/3 \quad \text{and} \quad |f(a) - g_1(a)| \leq 2/3$$
for all $x \in X$, $a \in A$.

Then, for $r = \frac{2}{3}$, $h = f - g_1$, we repeat the process to get a continuous function g_2 on X such that, for all $x \in X$, $a \in A$, we have:

$$|g_2(x)| \leq 1/3 \cdot (2/3) \quad \text{and} \quad |f(a) - g_1(a) - g_2(a)| \leq (2/3)^2 .$$

By recursion, suppose that $s_n = \sum_{k=1}^{n} g_k$, where $|f(a) - s_n(a)| \leq \left(\frac{2}{3}\right)^n$ for all $a \in A$. Take $r = \left(\frac{2}{3}\right)^n$, and $h = f - s_n$. Then by Lemma 256, there is a continuous function g_{n+1} on X such that

$$|g_{n+1}(x)| \leq 1/3 \cdot (2/3)^n \quad \text{and} \quad |f(a) - s_n(a) - g_{n+1}(a)| \leq (2/3)^{n+1}$$

for all $x \in X$, $a \in A$. By induction, the continuous functions g_n are defined for all $n \in \mathbb{N}$, and $|g_n(x)| \leq 1/3 \cdot (2/3)^{n-1} = M_n$ for all $x \in X$. By the Weierstrass M-test (Theorem 79),

$$g = \sum_{n \in \mathbb{N}} g_n$$

is uniformly convergent hence continuous. By construction

$$|g(x)| \leq \sum_{n \in \mathbb{N}} |g_n(x)| \leq \sum_{n \in \mathbb{N}} \frac{1}{3} \left(\frac{2}{3}\right)^{n-1} = \frac{1/3}{1 - 2/3} = 1$$

for all $x \in X$. For $a \in A$, $|f(a) - s_n(a)| \leq \left(\frac{2}{3}\right)^n$. Then, as $n \to \infty$, $s_n(a) \to f(a)$ and $s_n(a) \to g(a)$. As X is Hausdorff, limits are unique, so $g|_A = f$.

2. It is sufficient to prove the theorem for continuous $f : A \to (-1, 1)$, as $(-1, 1)$ is homeomorphic to $\mathbb{R}$. If $f : A \to (-1, 1) \subseteq [-1, 1]$ is a continuous function, using part 1 of the theorem, there is a continuous extension $h : X \to [-1, 1]$. Define

$$D = h^{-1}(\{-1\}) \cup h^{-1}(\{1\}) \subseteq X.$$

As h is continuous and X is Hausdorff, $D \subseteq_C X$. Since

$$h(A) = f(A) \subseteq (-1, 1),$$

then $A \cap D = \varnothing$. Using the Urysohn lemma, there is a continuous function $\phi : X \to [0, 1]$ such that $\phi(D) = \{0\}$ and $\phi(A) = \{1\}$. Let $g(x) = \phi(x)h(x)$ for all $x \in X$.

Then g is continuous and $g|_A = f$ since $g(a) = \phi(a)h(a) = 1 \cdot h(a) = f(a)$ for all $a \in A$. Finally, $g : X \to (-1, 1)$. Indeed, if $x \in D$, then $g(x) = \phi(x)h(x) = 0 \cdot h(x) = 0 \in (-1, 1)$. If $x \notin D$, then $|h(x)| < 1$, so $|g(x)| \leq |\phi(x)||h(x)| < 1$. ■

In the remaining part of this chapter, we prove a result that provides conditions under which a topological space is **metrizable**.

Theorem 258 (Urysohn Metrization Theorem)
Every regular second countable space X is metrizable.

Proof: we show X is metrizable by showing it is homeomorphic to a subspace of $\mathbb{R}^\omega$ in the product topology. Let $\mathfrak{B} = \{B_n\}_{n\in\mathbb{N}}$ be a basis of X. Then, using appendix A, section 4, question 3, there is a countable collection of continuous function $f_n : X \to [0,1]$, where $f_n(x) > 0$ for $x \in B_n$ and $f(X \setminus B_n) = \{0\}$ for all $n \in \mathbb{N}$. Given $x_0 \in X$, and a neighbourhood U of x_0, there is an index $n \in \mathbb{N}$ such that $f_n(x_0) > 0$ and $f(X \setminus U) = \{0\}$. Indeed, choose a basis element B_n such that $x_0 \in B_n \subseteq U$. Then the index n satisfies the property. Now, define a function $F : X \to \mathbb{R}^\omega$ (in the product topology) by

$$F(x) = (f_1(x), f_2(x), \ldots).$$

We show that F is an embedding. Clearly, F is continuous, since f_n is continuous for all $n \in \mathbb{N}$. Furthermore, it is injective. Indeed, let $x \neq y$. As X is Hausdorff, there is a neighbourhood U of x disjoint from y. Using the property above, there is an index $n \in \mathbb{N}$ such that $f_n(x) > 0$ and $f_n(y) = 0$. Hence $F(x) \neq F(y)$. It remains only to show that F is an homeomorphism from X to $F(X)$. As F is already continuous, it will be sufficient to show that F is open.

Let $U \subseteq_O X$ and $z_0 \in F(X)$. Then there exists $W \subseteq_O F(X)$ such that $z_0 \in W \subseteq F(U)$. Indeed, let $x_0 \in U$ such that $F(x_0) = z_0$. As above, there is an index $N \in \mathbb{N}$ such that $f_N(x_0) > 0$ and $f_N(X \setminus U) = \{0\}$. Let

$$V = \pi_N^{-1}((0,\infty)) \subseteq_O \mathbb{R}^\omega,$$

and set $W = V \cap F(X)$. Then $W \subseteq_O F(X)$, and

$$\pi_N(z_0) = \pi_N(F(x_0)) = f_N(x_0) > 0,$$

so that $z_0 \in W$. We show now that $W \subseteq F(U)$. If $z \in W$, there exists an $x \in X$ such that $z = F(x)$ and $\pi_N(z) > 0$. But $0 < \pi_N(z) = f_N(x)$, so $x \in U$. Then $F(x) \in F(U)$ and $W \subseteq F(U)$. Hence F is open. ■

In the proof, we have called upon a special countable collection of continuous functions. The following theorem shows how to generalize to arbitrary collections.

Theorem 259 (Embedding Theorem)
Suppose X is Hausdorff and $\{f_\alpha\}$ is a family of real-valued continuous functions (indexed by A) such that if U is a neighbourhood of $x_0 \in X$, there is an $\alpha \in A$ such that $f_\alpha(x_0) > 0$ and $f_\alpha(x) = 0$ if $x \notin U$. Then X is homeomorphic to a subspace of $\mathbb{R}^A$.

Proof: the proof is similar to that of the Urysohn metrization theorem, just replace $n \in \omega$ throughout by $\alpha \in A$. ∎

Let's take a look at another embedding result.

Theorem 260
Let X be a completely regular space. Then X can be embedded in $\mathbb{R}^A$ for some A.

Proof: we first define an index set. Let $A = \{(C, x) \mid C \subseteq_C X, x \notin C\}$. For $x_0 \in X$, if U is a neighbourhood of x_0, then $C = X \setminus U \subseteq_C X$ and $x_0 \notin C$, so $\alpha = (C, x_0) \in A$.

Since X is completely regular, there is a continuous function $f_\alpha : X \to [0,1]$ such that $f(x_0) = 1$ and $f(x) = 0$ for all $x \in C$, so for all $x \notin U$. Hence there is a family of continuous functions $\{f_\alpha\}_\alpha$ satisfying the hypotheses of the embedding theorem. As X is Hausdorff, we apply the embedding theorem to obtain the desired result. ∎

The next result shows that $T_{3\frac{1}{2}}$ spaces behave in a nice fashion, not unlike their T_3 cousins.

Theorem 261
Subspaces and product of completely regular spaces are completely regular.

Proof: suppose Y is a subspace of the completely regular space X. If $y \in Y$ and $y \notin A \subseteq_C Y$, then $A = Y \cap \overline{A}$ (closure in X) and $y \notin \overline{A}$. Since X is completely regular, there is a continuous function $f : X \to [0,1]$ such that $f(\overline{A}) = \{0\}$ and $f(y) = 1$. Then the restriction of f to Y is continuous and $f|_Y(A) = \{0\}$ and $f|_Y(y) = 1$, so that Y is completely regular.

Now suppose that X_α is completely regular for every α. Let $X = \prod_\alpha X_\alpha$. If $C \subseteq_C X$ and $x_0 = (x_\alpha)_\alpha \notin C$, then there is a basic neighbourhood $\prod_\alpha U_\alpha$ of x_0 disjoint from C. By definition, $U_\alpha = X_\alpha$, except when $\alpha = \alpha_i$, $1 \leq i \leq n$ for some n. Let $i \in \{1, \ldots, n\}$. Then $x_{\alpha_i} \in U_{\alpha_i} \subsetneq_O X_{\alpha_i}$, so $X_{\alpha_i} \setminus U_{\alpha_i} \subseteq_C X_{\alpha_i}$. By complete regularity of X_{α_i}, there is a continuous function $f_{\alpha_i} : X_{\alpha_i} \to [0,1]$ such that $f_{\alpha_i}(x_{\alpha_i}) = 1$ and $f_{\alpha_i}(X_{\alpha_i} \setminus U_{\alpha_i}) = \{0\}$. This can be done for all $1 \leq i \leq n$. Now define a function $f : X \to [0,1]$ by

$$f(x) = f_{\alpha_1}(\pi_{\alpha_1}(x)) \cdots f_{\alpha_n}(\pi_{\alpha_n}(x)).$$

Then, f is continuous, being the product of continuous functions. Furthermore, $f(x_0) = f_{\alpha_1}(\pi_{\alpha_1}(x_0)) \cdots f_{\alpha_n}(\pi_{\alpha_n}(x_0)) = 1$. Now suppose $y \notin \prod_\alpha U_\alpha$. Then, there exists α_i such that $\pi_{\alpha_i}(y) \notin U_{\alpha_i}$ and $f_{\alpha_i}(\pi_{\alpha_i}(y)) = 0$. Hence $f(y) = 0$, and X is completely regular. ∎

18.4 Solved Problems

1. Show that if X is Lindelöf and Y is compact, then $X \times Y$ is Lindelöf.

   Proof: the proof is nearly identical to that showing $X \times Y$ is compact whenever X and Y are compact. Let $\mathfrak{F}$ be an open covering for $X \times Y$. For each $x \in X$ we get a finite subcovering of $\{x\} \times Y$ from $\mathfrak{F}$, say $F(x)_1, \ldots F(x)_n$. Let N be the open set

   $$N = \bigcup_{i=1}^{n} F(x)_i.$$

   By the Tube Lemma, there is a neighbourhood W_x of x in X such that $W_x \times Y \subseteq N$. Repeating this procedure for all $x \in X$, we get that $\{W_x\}_{x \in X}$ is an open covering of X. But X is Lindelöf, so there is a countable subcovering

   $$W_{x_1}, W_{x_2}, \ldots$$

   For each of these W_{x_i}, there were n_i corresponding sets $F(x_i)_j$ in $\mathfrak{F}$. Define

   $$\mathfrak{F}' = \{F(x_i)_j \mid i \in \mathbb{N}, 1 \leq j \leq n_i\}.$$

   $\mathfrak{F}'$ is an open countable collection. For any $(x, y) \in X \times Y$, $x \in W_{x_i}$ for some i. Then $(x, y) \in W_{x_i} \times Y$ and $(x, y) \in F(x_i)_j$ for some j, so

   $$X \times Y \subseteq \bigcup_{i \in \mathbb{N}} \bigcup_{j=1}^{n_i} F(x_i)_j.$$

   So $\mathfrak{F}'$ is a countable subcovering of $X \times Y$ extracted from $\mathfrak{F}$, so $X \times Y$ is Lindelöf. ■

2. Let X be a space with the order topology. Show that X is regular.

   Proof: let $A \subseteq_C X$, with $b \notin A$. As A is closed, there exists an open interval (c, d) such that

   $$b \in (c, d) \subseteq X \setminus A.$$

   There are now four possibilities.

   a) If there exists $e, f \in X$ such that $c < e < b < f < d$, put $U = (e, f)$ and $V = (-\infty, e) \cup (f, +\infty)$.
   b) If there is an f, but no such e, that is $(c, b) = \varnothing$, put $U = (c, f)$ and $V = (-\infty, b) \cup (f, +\infty)$.
   c) If there is an e, but no such f, that is $(b, d) = \varnothing$, put $U = (e, d)$ and $V = (-\infty, e) \cup (b, +\infty)$.
   d) If there are no such e, f, that is $(c, b) = (b, d) = \varnothing$, put $U = (c, d) = \{b\}$ and $V = (-\infty, b) \cup (b, +\infty)$.

   In all cases, $b \in U \subseteq_O X$, $A \subseteq V \subseteq_O X$ and $U \cap V = \varnothing$, so X is regular. ■

3. a) If X is a Lindelöf space, show that every closed subset of X is Lindelöf.
   b) If A, B are disjoint closed subsets of a regular space, show that there are open coverings $\mathfrak{E}$, $\mathfrak{F}$ of A, B respectively such that $\overline{U} \cap B = \varnothing$ and $\overline{V} \cap A = \varnothing$ for all $U \in \mathfrak{E}, V \in \mathfrak{F}$.
   c) If X is a regular Lindelöf space, show that X is normal.

Proof:

a) Let A be a closed subset of X, and suppose that $\mathfrak{F}$ is an open covering of A. Then $X \setminus A$ is open and $\mathfrak{F} \cup \{X \setminus A\}$ is an open covering of X. But X is Lindelöf, so there is a countable sub-covering of X, say

$$\{X \setminus A, F_1, F_2, \ldots\},$$

where $F_n \in \mathfrak{F}$ for all $n \in \mathbb{N}$. Consequently,

$$\{F_1, F_2, \ldots\} \subseteq \mathfrak{F}$$

is a countable sub-covering of A, and A is Lindelöf.

b) Let $a \in A$. Since $X \setminus B$ is open, there exists an open set W_a such that $a \in W_a$ and $W_a \cap B = \varnothing$. By regularity of X, there exists an open set U_a such that

$$a \in U_a \subseteq \overline{U_a} \subseteq W_a.$$

Then $\overline{U_a} \cap B \subseteq W_a \cap B = \varnothing$. The collection $\{U_a\}_{a \in A}$ is an open covering of A satisfying the requisite property. Similarly, we can construct an open covering of B satisfying the property.

c) Let A and B be disjoint closed subsets of the regular Lindelöf space X. Then A and B are Lindelöf, by part (a), and there are open coverings $\mathfrak{E}$ and $\mathfrak{F}$ of A and B respectively such that

$$\overline{U} \cap B = \varnothing \quad \text{and} \quad \overline{V} \cap A = \varnothing$$

for all $U \in \mathfrak{E}$, $V \in \mathfrak{F}$. Since A and B are Lindelöf, it is possible to extract countable sub-coverings

$$\{U_1, U_2, \ldots\} \subseteq \mathfrak{E} \quad \text{and} \quad \{V_1, V_2, \ldots\} \subseteq \mathfrak{F}$$

of A and B respectively. Now define

$$U_n' = U_n \setminus \bigcup_{i=1}^{n} \overline{V_i} \quad \text{and} \quad V_n' = V_n \setminus \bigcup_{i=1}^{n} \overline{U_i}.$$

Then U_n', V_n' are open in X as U_n, V_n are open in X and $\bigcup_{i=1}^n \overline{U_i}, \bigcup_{i=1}^n \overline{V_i}$ are closed in X. Let

$$U' = \bigcup_{n \in \mathbb{N}} U_n' \quad \text{and} \quad V' = \bigcup_{n \in \mathbb{N}} V_n'.$$

Then U', V' are open in X with $A \subseteq U'$ and $B \subseteq V'$. Indeed if $x \in A$, then $x \in U_n$ for some n. But, by construction, $x \notin \overline{V_i}$ for all $i \in \mathbb{N}$. Then $x \in U_n'$. Similarly, if $y \in B$, then $y \in V_n'$ for some $n \in \mathbb{N}$. It remains only to show that $U' \cap V' = \varnothing$. Suppose $\xi \in U' \cap V'$. Then $\xi \in U_n' \cap V_m'$ for some $m, n \in \mathbb{N}$. If $m \geq n$, then

$$\begin{aligned} \xi \in U_n' &\implies \xi \in U_n \\ \xi \in V_m' &\implies \xi \notin \overline{U_n}, \end{aligned}$$

a contradiction. If $m \leq n$, then

$$\begin{aligned} \xi \in U_n' &\implies \xi \notin \overline{V_m} \\ \xi \in V_m' &\implies \xi \in V_m, \end{aligned}$$

another contradiction. Then $U' \cap V' = \varnothing$ and X is normal. ■

4. Let X be a second countable regular space and let U be an open set.

a) Let $\{f_n\}_{n\in\mathbb{N}}$ be a sequence of real-valued functions on a space X. If there exists $M \in \mathbb{R}$ such that $|f_n(x)| \leq M$ for all $x \in X$, $n \in \mathbb{N}$ show that

$$\sum_{n\in\mathbb{N}} \frac{1}{2^n} f_n$$

converges uniformly on X.

b) Show that U is a countable union of closed sets in X.

c) Show that there is a continuous function $f : X \to [0, 1]$ such that $f(x) > 0$ for all $x \in U$ and $f(x) = 0$ for all $x \notin U$.

Proof:

a) Let $\varepsilon > 0$, and choose $N_\varepsilon \in \mathbb{N}$ such that

$$N_\varepsilon > \frac{\log M - \log \varepsilon}{\log 2}.$$

Then, for all $x \in X$ and $n > N_\varepsilon$,

$$\left| \sum_{i=n+1}^{\infty} \frac{1}{2^i} f_i(x) \right| \leq \sum_{i=n+1}^{\infty} \frac{1}{2^i} |f_i(x)| \leq M \left(\sum_{i=n+1}^{\infty} \frac{1}{2^i} \right) = \frac{M}{2^n} < \frac{M}{2^{N_\varepsilon}} < \varepsilon,$$

and so $\sum 2^{-n} f_n$ converges uniformly on X.

b) Suppose $\mathfrak{B} = \{B_n\}_{n\in\mathbb{N}}$ is a basis for X and U is open in X. Then $X \setminus U$ is closed in X. Since X is regular, if $x \in U$, there exist $B_x \in \mathfrak{B}$ and an open set V_x such that $x \in B_x$, $X \setminus U \subseteq V_x$ and $B_x \cap V_x = \varnothing$. But

$$\bigcup_{x\in U} B_x = U$$

since $B_x \cap (X \setminus U) \subseteq B_x \cap V_x = \varnothing$ for all $x \in U$. As $\mathfrak{B}$ is countable, there exists a sequence $\{x_n\}_{n\in\mathbb{N}}$ in U such that

$$U = \bigcup_{x\in U} B_x = \bigcup_{n\in\mathbb{N}} B_{x_n}.$$

By construction, $X \setminus U \subseteq V_{x_n}$, so $X \setminus V_{x_n} \subseteq U$ for all $n \in \mathbb{N}$, and

$$\bigcup_{n\in\mathbb{N}} (X \setminus V_{x_n}) \subseteq U.$$

Now, suppose $x \in U$. Then $x \in B_{x_n}$ for some $n \in \mathbb{N}$, so $x \notin V_{x_n}$ and $x \in X \setminus V_{x_n}$ for that n. Hence

$$\bigcup_{n\in\mathbb{N}} (X \setminus V_{x_n}) = U.$$

But $X \setminus V_{x_n}$ is closed in X for all $n \in \mathbb{N}$ so U is a countable union of closed sets.

c) By hypothesis, $U = \bigcup_{n\in\mathbb{N}} C_n$, where C_n is closed in X and $X \setminus U$ is closed. But X is normal, as it is regular and second countable, so, by the Urysohn lemma, there exists a family of continuous functions $\{f_n\}_{n\in\mathbb{N}}$, where $f_n : X \to [0,1]$, such that $f_n(X \setminus U) = \{0\}$ and $f_n(C_n) = \{1\}$. Define the function f on X by

$$f(x) = \sum_{n\in\mathbb{N}} \frac{1}{2^n} f_n(x).$$

Since each f_n is bounded by 1 above and 0 below, we can apply the result obtained in q. 2 to show that f is defined for all $x \in X$ and that f is continuous, since the series is uniformly convergent. Now we show that $f : X \to [0,1]$. Let $x \in X$. Then $f_n(x) \in [0,1]$ for all $n \in \mathbb{N}$, so

$$0 \le \underbrace{\sum_{n\in\mathbb{N}} 2^{-n} f_n(x)}_{=f(x)} \le \sum_{n\in\mathbb{N}} 2^{-n} = 1.$$

It remains to show that f satisfies the requisite property. Suppose $x \notin U$. Then $f(x) = \sum 2^{-n} f_n(x) = \sum 2^{-n} \cdot 0 = 0$. Now suppose $x \in U$. Then $x \in C_n$ for some n, and $f(x) \ge 2^{-n} > 0$. ■

5. For disjoint closed sets A, B in a completely regular space, if A is compact show that there is a continuous function $f : X \to [0,1]$ with $f(A) = \{0\}$ and $f(B) = \{1\}$.

Proof: let $a \in A$. Then, by the previous question, there exists a continuous function $f_a : X \to [0,1]$ such that $f_a(B) = \{1\}$ and $f_a(U_a) = \{0\}$ for some neighbourhood U_a of a, disjoint from B. The collection $\mathfrak{F} = \{U_a\}_{a\in A}$ is then an open covering of A, disjoint from B. But A is compact, so there is a finite sub-covering

$$\{U_{a_1}, \ldots, U_{a_n}\} \subseteq \mathfrak{F}$$

of A. Pick the associated functions f_{a_i}, $1 \leq i \leq n$, and construct the function $f : X \to [0,1]$ defined by

$$f(x) = f_{a_1}(x) f_{a_2}(x) \cdots f_{a_n}(x).$$

Then f is continuous, since the finite product of continuous functions is continuous. Furthermore, $f(A) = \{0\}$ and $f(B) = \{1\}$. Indeed, suppose $x \in A$. Then $x \in U_{a_i}$ for some i and $f_{a_i}(x) = 0$, so $f(x) = 0$. If $x \in B$, $f_{a_i}(x) = 1$ for all $1 \leq i \leq n$, so $f(x) = 1$. ■

6. a) Show that a connected normal space X having more than one point is uncountable.
   b) Show that a connected regular space X having more than one point is uncountable.

   Proof:

   a) By hypothesis, there exists $x \neq y \in X$. Since X is normal, singletons are closed in X and X is completely regular. Then there exists a continuous function $f : X \to [0,1]$ such that $f(x) = 0$ and $f(y) = 1$. But X is connected. By the intermediate value theorem, for every $0 = f(x) \leq r \leq f(y) = 1$, there exists $z_r \in X$ such that $f(z_r) = r$. Then f is a surjection of X onto $[0,1]$. Hence X is uncountable.
   b) Suppose X was a countable connected regular space with at least two points. Then X is clearly Lindelöf, so it is normal by a previous solved problem. But this would make X uncountable by this problem's first part, which is a contradiction. Hence X has to be uncountable. ■

7. Show that every locally compact Hausdorff space is completely regular.

   Proof: as X is a locally compact Hausdorff space, it has a one-point compactification $Y = X \cup \{\infty\}$, where $Y = \overline{X}$ and X is a subspace of Y. But Y is compact Hausdorff, so X is homeomorphic to a subspace of a compact Hausdorff space, hence X is completely regular. ■

18.5 Exercises

1. Prepare a 2-page summary of this chapter, with important definitions and results.
2. If A is a subspace of a first countable space X, show that $x \in \overline{A}$ if and only if there is a sequence of points in A converging to x.
3. If X is a first countable space, show that $f : X \to Y$ is continuous if and only if for any convergent sequence $x_n \to x$, the sequence $f(x_n)$ converges to $f(x)$.
4. If X is second countable, show that every collection of disjoint open sets in X is countable.
5. If Y is compact and X is Lindelöf, show that $X \times Y$ is Lindelöf.

6. Let X be a regular, second countable space. Show that every open set U in X is a countable union of closed sets.

7. Use the fact that X is completely regular to show that there is a continuous function $f : X \to [0, 1]$ such that $f(x) > 0$ for all $x \in U$ and $f(x) = 0$ for $x \notin U$.

8. Show that subspaces and products of completely regular spaces are completely regular.

9. If X_n is first countable for all $n \in \mathbb{N}$, show that $\prod_{n \in \mathbb{N}} X_n$ is first countable.

10. Provide proofs for the examples of p. 431.

11. Complete the proof of Lemma 250.

12. Illustrate the separation axioms as in Chapter 15 (see p. 321, and footnote for a list).

Chapter 19

Advanced Topics

In Chapter 17, we showed that the **finite** product of compact spaces is compact in the **box**, **uniform**, and **product** topologies. **Arbitrary** products of compact spaces, on the other hand, are surprisingly more complicated to handle.

19.1 Tychonoff's Theorem

Our formulation of **compactness** in terms of closed sets uses the **finite intersection property** (f.i.p.).[1] In this section, we will use the following notation:

- a is an element of X;
- A is a subset of X;
- $\mathfrak{A}$ is a collection of subsets of X;
- $\mathbb{A}$ is a family of collections of subsets of X;

as well as a slightly altered re-formulation of that statement (see Theorem 234 (Reprise) in Chapter 17):

> **Theorem 234 (Reprise, Reprise)**
> *X is **compact** if and only if for every family $\mathfrak{F}$ of subsets of X satisfying the f.i.p., we have*
>
> $$\bigcap_{F\in\mathfrak{F}} \overline{F} \neq \varnothing.$$
>
> **Proof:** left as an exercise. ■

[1]We note that the projection mappings are not closed in general.

Our goal is to show that arbitrary products of compact spaces are compact; the following lemmas will bring us to the promise land.

Lemma 262
For any set X and any collection $\mathfrak{F}$ of subsets of X satisfying the f.i.p., there exists a maximal collection $\mathfrak{G}$ with respect to the f.i.p., that is, $\mathfrak{F} \subseteq \mathfrak{G}$, and $\mathfrak{G} \subsetneq \mathfrak{G}' \implies \mathfrak{G}'$ does not satisfy the f.i.p.

Proof: consider all possible collections of subsets of X satisfying the f.i.p., and define a partial order on them by strict inclusion. Then $\{\mathfrak{F}\}$ is a totally ordered family, so by the **maximum principle** of set theory, there is a maximal totally ordered family $\mathbb{A}$ containing it. Define

$$\mathfrak{G} = \bigcup_{\mathfrak{F}' \in A} \mathfrak{F}'.$$

Then $\mathfrak{G}$ satisfies the f.i.p. Indeed, if $G_1, \ldots, G_n \in \mathfrak{G}$, then, for each i, there exists $\mathfrak{F}_i \in \mathbb{A}$ such that $G_i \in \mathfrak{F}_i$. But $\mathbb{A}$ is totally ordered, so one of the $\mathfrak{F}_i$, say $\mathfrak{F}_k$, contains all the others. Then $G_1, \ldots, G_n \in \mathfrak{F}_k$. But $\mathfrak{F}_k$ satisfies the f.i.p., so

$$\bigcap_{i=1}^{n} G_i \neq \varnothing.$$

As $\mathfrak{F} \in \mathbb{A}$, we have $\mathfrak{F} \subseteq \mathfrak{G}$.

Now, suppose $\mathfrak{G} \subseteq \mathfrak{G}'$, where $\mathfrak{G}'$ also satisfies the f.i.p. Then $\mathfrak{F} \subseteq G'$. Furthermore, if $\mathfrak{F}' \in \mathbb{A}$, then $\mathfrak{F}' \subseteq \mathfrak{G}'$. So $\mathfrak{G}'$ is comparable with every collection in $\mathbb{A}$. Thus $\mathbb{A} \cup \{\mathfrak{G}'\}$ is totally ordered and each of its constituent collection satisfies the f.i.p. But $\mathbb{A}$ was maximal with respect to the f.i.p., so $\mathfrak{G}' \in \mathbb{A}$, hence $\mathfrak{G}' \subseteq \mathfrak{G}$ and $\mathfrak{G}' \subseteq \mathfrak{G}$. ■

The **Haussdorf maximum principle** states that in any partially ordered set, every totally ordered subset is contained in a maximal totally ordered subset. This benign looking statement is in fact equivalent to the infamous **axiom of choice**; as it is a a fundamental part of the proof of Lemma 262, it is also a fundamental constituent of its descendents.

Lemma 263
If $\mathfrak{F}$ is maximal with respect to the f.i.p. and $F_1, \ldots, F_n \in \mathfrak{F}$, then $\bigcap_{i=1}^n F_i \in \mathfrak{F}$.

Proof: let $G = \bigcap_{i=1}^n F_i$ and $\mathfrak{G} = \mathfrak{F} \cup \{G\}$. Suppose $G_1, \ldots, G_m \in \mathfrak{G}$ are all distinct.

1. If G is not one of the G_j's, then $G_j \in \mathfrak{F}$ for $1 \leq j \leq m$. Then

$$\bigcap_{j=1}^{m} G_j \neq \varnothing,$$

since $\mathfrak{F}$ satisfies the f.i.p.

2. If $G = G_m$, then then $G_j \in \mathfrak{F}$ for $1 \leq j \leq m-1$. Then

$$\bigcap_{j=1}^{m} G_j = \left(\bigcap_{j=1}^{m-1} G_j\right) \cap \left(\bigcap_{i=1}^{n} F_i\right) \neq \varnothing,$$

since $\mathfrak{F}$ satisfies the f.i.p.

Thus $\mathfrak{G}$ satisfies the f.i.p. By maximality of $\mathfrak{F}$, $\mathfrak{G} = \mathfrak{F}$, hence $G \in \mathfrak{F}$. ■

We will need one more lemma.

Lemma 264
If $\mathfrak{F}$ is a maximal collection with respect to the f.i.p. and $A \subseteq X$ is such that $A \cap F \neq \varnothing$ for any $F \in \mathfrak{F}$, then $A \in \mathfrak{F}$.

Proof: let $\mathfrak{G} = \{A\} \cup \mathfrak{F}$. We show $\mathfrak{G}$ satisfies the f.i.p. Let $G_1, \ldots, G_n \in \mathfrak{G}$.

1. If $G_i \neq A$ for $1 \leq i \leq n$, then $\bigcap_{i=1}^{n} G_i \neq \varnothing$, since $\mathfrak{F}$ satisfies the f.i.p.
2. If $G_n = A$, let $F = \bigcap_{i=1}^{n-1} G_i$, where $G_i \in F$ for $1 \leq i \leq n-1$. By Lemma 263, $F \in \mathfrak{F}$. But by hypothesis,

$$\bigcap_{i=1}^{n} G_i = A \cap F \neq \varnothing.$$

Hence $\mathfrak{G} = \mathfrak{F}$ and $A \in \mathfrak{F}$. ■

We are now ready to state and prove this section's main result.

Theorem 265 (Tychonoff Theorem)
Let $\{X_\alpha\}_\alpha$ be a family of compact sets. Then $\prod_\alpha X_\alpha$ is compact.

Proof: we show that any collection of subsets of $X = \prod_\alpha X_\alpha$ satisfying the f.i.p. has a non-trivial intersection. If $\mathfrak{A}$ is such a collection, then let $\mathfrak{F}$ be the corresponding maximal collection with respect to the f.i.p., given by Lemma 262.

Then

$$\bigcap_{F\in\mathfrak{F}}\overline{F}\subseteq\bigcap_{F\in\mathfrak{A}}\overline{F},$$

and it will be sufficient to show

$$\bigcap_{F\in\mathfrak{F}}\overline{F}\neq\varnothing.$$

For each α, let $\mathfrak{F}_\alpha=\{\pi_\alpha(F)\}_{F\in\mathfrak{F}}$. Then, since $\mathfrak{F}$ satisfies the f.i.p.,

$$\pi_\alpha\underbrace{\left(\bigcap_{i=1}^n F_i\right)}_{\neq\varnothing}\subseteq\bigcup_{i=1}^n\pi_\alpha(F_i)$$

for any $F_1,\ldots,F_n\in\mathfrak{F}$. Hence $\mathfrak{F}_\alpha$ satisfies the f.i.p. But X_α is compact, so

$$P(\alpha)=\bigcap_{F\in\mathfrak{F}}\overline{\pi_\alpha(F)}\neq\varnothing.$$

Let $x_\alpha\in P(\alpha)\subseteq X_\alpha$ and set $x=(x_\alpha)_\alpha$. Then $x\in X$. If U_β is a neighbourhood of x_β in X_β, then $\pi_\beta^{-1}(U_\beta)$ is a sub-basic open set in X, and $U_\beta\cap\pi_\beta(F)\neq\varnothing$ for every $F\in\mathfrak{F}$, since $x_\beta\in\overline{\pi_\beta(F)}$ for all $F\in\mathfrak{F}$.

Consequently, $\pi_\beta^{-1}(U_\beta)\cap F\neq\varnothing$ for all $F\in\mathfrak{F}$. Then, by Lemma 264, $\pi_\beta^{-1}(U_\beta)\in\mathfrak{F}$. If V is a neighbourhood of x in X, then V contains a basic neighbourhood $U=\prod_\alpha U_\alpha$ around x, where $U_\alpha=X_\alpha$ for all but finitely many α.

But

$$U=\bigcap_{i=1}^n\pi_{\beta_i}^{-1}(U_{\beta_i}).$$

By Lemma 19.1, $U\in\mathfrak{F}$. Then $U\cap F\neq\varnothing$ for all $F\in\mathfrak{F}$ since $\mathfrak{F}$ satisfies the f.i.p., so $V\cap F\neq\varnothing$ for all $F\in\mathfrak{F}$. But V was arbitrary, so $x\in\overline{F}$ for all $F\in\mathfrak{F}$ and $x\in\bigcap_{F\in\mathfrak{F}}\overline{F}$. Hence X is compact. ■

Note that as $[0,1]$ is compact, $[0,1]^A$ is compact in the product topology. As a result, any completely regular space can be embedded in $[0,1]^A$ for some index set A, according to the embedding theorem (Theorem 259). Hence, any completely regular space is homeomorphic to a subspace of a compact Hausdorff space, which is to say, a **normal space**. This opens the door for us to continue the discussion on **compactification**.

19.2 Stone-Čech Compactification

A **compactification** of a space X is a compact Hausdorff space Y which contains X as a subspace and such that $\overline{X} = Y$. For X to have a compactification, it must be **completely regular**.

As Y is compact Hausdorff, it is necessarily normal, and so completely regular, and its subspaces are also completely regular. We now show that this condition is sufficient.

Theorem 266
If X is completetly regular, then X has a compactification Y.

Proof: since X is completely regular, it is possible to embed X into a space $Z = [0,1]^A$. If $f : X \to Z$ is the embedding, let $X_0 = f(X)$ and take $Y = \overline{X}$. Then Y_0 is compact, since it is closed in the compact space Z. Let X_1 be a set disjoint from X, in one-to-one correspondence with $Y_0 \setminus X_0$. Then, put $Y = X \cup X_1$. If $g : X_1 \to Y_0 \setminus X_0$ is the bijection, then define $h : Y \to Y_0$ by

$$h(x) = \begin{cases} f(x) & \text{if } x \in X, \\ g(x) & \text{if } x \in X_1. \end{cases}$$

Then h is a bijection. Topologize Y by setting

$$V \subseteq_O Y \iff h(V) \subseteq_O Y_0.$$

This clearly makes $h : Y \to Y_0$ a homeomorphism, and so Y is compact, Hausdorff. But the restriction of h on X is a homeomorphism of X onto X_0, so X is a subspace of Y and $\overline{X_0} = Y_0$ implies $\overline{X} = Y$. ■

The compactification clearly depends on the embedding $f : X \to Z$.

Examples: let $X = (0,1)$ in the usual topology and $f : X \to Z$.

1. If $Z = [0,1]^2$ in the usual topology and $f(x) = e^{2\pi i x}$, then the resulting compactification is the one-point compactification.
2. If $Z = [0,1]$ in the usual topology and $f(x) = x$, then the resulting compactification is a two-point compactification.
3. If $Z = [0,1]^2$ in the usual topology and $f(x) = (x, \sin(1/x))$, then the resulting compactification is given by adding the sets $\{0\} \times [-1,1]$ and $\{(1, \sin 1)\}$ to the topologist's sine curve.

Now, suppose X is a completely regular space. Let $\{f_\alpha\}_{\alpha\in A}$ be the set of all continuous bounded real-valued functions on X. For each $\alpha \in A$, let

$$I_\alpha = \left[\inf_{x\in X}\{f_\alpha(x)\}, \sup_{x\in X}\{f_\alpha(x)\}\right].$$

Then I_α is a closed bounded interval in $\mathbb{R}$, so I_α is compact and $\prod_\alpha I_\alpha$ is compact by Tychonoff's theorem. Define $\hat{F} : X \to \prod_\alpha I_\alpha$ by

$$\hat{F}(x) = (f_\alpha(x))_\alpha,$$

and so $\hat{F}$ is continuous as $f_\alpha(x)$ is continuous for all α. Since X is completely regular, the set $\{f_\alpha\}_{\alpha\in A}$ satisfies the conditions of the embedding theorem.

Consequently, X is homeomorphic to a subspace of $Z = \prod_\alpha I_\alpha$, and we obtain a compactification of X that is homeomorphic to the closure of $\hat{F}(X)$ in Z. This compactification is called the **Stone-Čech compactification** of X, and is denoted $\beta(X)$.[2]

If Y and Z are compactifications of X for which there exists an homeomorphism $f : Y \to Z$, we say that Y and Z are **equivalent** if $f(x) = x$ for all $x \in X$.

Theorem 267
If X is completely regular, then every continuous bounded real-valued function on X can be uniquely extended to a continuous function on $\beta(X)$.

Proof: let f_γ be a continuous bounded real-valued function on X. Then

$$f_\gamma = \pi_\gamma \circ F|_X,$$

where $F : \beta(X) \to \prod I_\alpha$ is the embedding given in footnote 2. Define g on $\beta(X)$ by

$$g(x) = \pi_\gamma F(x).$$

Then $g|_X = f_\gamma$; according to a previous solved problem, the extension is unique as $\beta(X) = \overline{X}$. ■

This leads to the following useful result.

Theorem 268
Suppose that $g : X \to Z$ is continuous, where Z is compact Hausdorff. Suppose Y is a compactification of X such that every continuous real-valued function on X can be extended to Y. Then g can be extended to Y.

[2]Note that we have just uniquely extended the continuous function $\hat{F}$ on X to a continuous function F on $\beta(X) = \overline{X}$ using one of the solved problems from a previous section.

Proof: since Z is a compact Hausdorff it is normal, and so completely regular. Then Z can be embedded into $[0,1]^A$ for some A. Without loss of generality, we may take Z as a subspace of $[0,1]^A$. Note that Z is closed in $[0,1]^A$, since it is a compact subset of $[0,1]^A$. Then $g : X \to [0,1]^A$ is continuous and $g_\alpha = \pi_\alpha \circ g : X \to [0,1]$ is continuous for all $\alpha \in A$. By hypothesis, g_α can be extended to a continuous function $f_\alpha : Y \to \mathbb{R}$. Define $f : Y \to \mathbb{R}^A$ by

$$f(y) = (f_\alpha(y))_{\alpha \in A}.$$

As each coordinate function is continuous, f is continuous. Furthermore, $f|_X = g$. It remains only to show that f maps Y into Z. But

$$f(Y) = f(\overline{X}) \subseteq \overline{f(X)} = \overline{g(X)}.$$

But $g(X) \subseteq Z$ and Z is closed, so $\overline{g(X)} \subseteq Z$. Consequently, $f(Y) \subseteq Z$. Thus $f : Y \to [0,1]^A$ is the desired extension. ■

In a certain sense, the Stone-Čech compactification is unique.

Theorem 269
Suppose Y_1 and Y_2 are compactifications of X satisfying the conditions of Theorem 268. If every continuous function $g : X \to Z$ can be extended, Y_1 and Y_2 are equivalent.

Proof: let $i_1 : X \to Y_1$ be the injection of X into the compact normal space Y_1. Then, i_1 can be extended to $f_1 : Y_2 \to Y_1$. Similarly, we can extend $i_2 : X \to Y_2$ to $f_2 : Y_1 \to Y_2$. Then $f_1 f_2 : Y_1 \to Y_1$, and

$$f_1 f_2(x) = f_1 i_2(x) = f_1(x) = i_1(x) = x$$

for $x \in X$. Hence $f_1 f_2$ extends id $: X \to Y_1$ to $Y_1 = \overline{X}$. Since id_{Y_1} is also such a continuous extension, $f_1 f_2 = \mathrm{id}_{Y_1}$ and, similarly, $f_2 f_1 = \mathrm{id}_{Y_2}$. Hence f_1 and f_2 are homeomorphisms and Y_1 and Y_2 are equivalent.

$$\begin{tikzcd} X \arrow[r, "i_1"] \arrow[dr, "i_2"'] & Y_1 \arrow[d, shift left, "f_2"] \\ & Y_2 \arrow[u, shift left, "f_1"] \end{tikzcd}$$

■

19.3 Solved Problems

1. Let $\{X_\alpha\}$ be a family of non-empty topological spaces. Prove that the product space is locally compact if and only if each X_α is locally compact and all but a finite number of the X_α are compact.

   Proof: let $X = \prod X_\alpha$ and assume the axiom of choice holds. Suppose $x = (x_\alpha)_\alpha \in X$. If X is locally compact, then it is locally compact at x and there exist a compact set C and a basic neighbourhood U such that $x \in U \subseteq C \subseteq X$. But U takes the form

   $$U = U_{\alpha_1} \times \cdots \times U_{\alpha_n} \times \prod_{\alpha \neq \alpha_1, \ldots, \alpha_n} X_\alpha,$$

   where U_{α_i} is open in X_{α_i} for all $1 \leq i \leq n$. Since $U \subseteq C$, then

   $$C = C_{\alpha_1} \times \cdots \times C_{\alpha_n} \times \prod_{\alpha \neq \alpha_1, \ldots, \alpha_n} X_\alpha,$$

   where $U_{\alpha_i} \subseteq C_{\alpha_i}$. But C is compact, so X_α is compact for all $\alpha \neq \alpha_i$, and so is C_{α_i}, for $1 \leq i \leq n$. Now, consider X_{α_i} for $1 \leq i \leq n$. By construction, C_{α_i} is compact, U_{α_i} is open and

   $$x_{\alpha_i} \in U_{\alpha_i} \subseteq C_{\alpha_i} \subseteq X_{\alpha_i}$$

   for $1 \leq i \leq n$. But this means that X_{α_i} is locally compact at x_{α_i}, so X_{α_i} is locally compact for $1 \leq i \leq n$.

   Conversely, suppose X_{α_i} is locally compact for $1 \leq i \leq n$ and X_α is compact for $\alpha \neq \alpha_i$, $1 \leq i \leq n$. Write

   $$W = \prod_{\alpha \neq \alpha_1, \ldots, \alpha_n} X_\alpha.$$

   By Tychonoff's theorem, W is compact, and so locally compact. Then

   $$X = X_{\alpha_1} \times \cdots \times X_{\alpha_n} \times W$$

   is a finite product of locally compact spaces, and so is locally compact. ■

2. Show that if X is completely regular and B is a closed set with $a \notin B$, then there is a continuous function $f : X \to [0,1]$ such that $f(x) = 1$ for all $x \in B$ and $f(x) = 0$ in some neighbourhood of a.

   Proof: since X is completely regular, it is homeomorphic to a subspace of a normal space Y (we identify X with its homeomorphic copy in Y). Since B is closed in X, there exists B_Y closed in Y such that $B = B_Y \cap X$. As $a \in X$, $a \in Y \setminus B_Y$. By normality of Y, there is an open set U_Y in Y such that

   $$a \in U_Y \subseteq \overline{U_Y} \subseteq Y \setminus B_Y.$$

Then $\overline{U_Y} \cap B_Y = \varnothing$, and we can apply the Urysohn lemma to find a continuous function $f : Y \to [0,1]$ such that $f(\overline{U_Y}) = \{0\}$ and $f(B_Y) = \{1\}$. The restriction of a continuous function to a subspace is continuous, so the restriction

$$f|_X : X \to [0,1]$$

is continuous. Put $U = U_Y \cap X$ and $\overline{U} = \overline{U_Y} \cap X$, so that U is open and $\overline{U}$ is closed in X and $U \subseteq \overline{U}$. Then

$$f|_X(B) = f(B_Y \cap X) = \{1\} \quad \text{and} \quad f|_X(\overline{U}) = f(\overline{U_Y} \cap X) = \{0\},$$

so that $f|_X(U) = \{0\}$. But by construction, $a \in U$, so $f|_X$ is the desired function. ■

3. Let X be completely regular. Show that X is connected if and only if $\beta(X)$ is connected.

Proof: if X is connected, $\beta(X) \simeq \overline{X}$ is connected. Now suppose X is not connected, and let A, B be a separation of X. Note that

$$\beta(X) \simeq \overline{X} = \overline{A \cup B} = \overline{A} \cup \overline{B}.$$

Hence $\beta(X)$ is disconnected if $\overline{A}$, $\overline{B}$ is a separation of $\beta(X)$. It will be sufficient to show that $\overline{A} \cap \overline{B} = \varnothing$. Define $f : X \to [0,1]$ by $f(A) = \{0\}$ and $f(B) = \{1\}$. Then, f is continuous. Indeed,

$$\begin{aligned} f^{-1}([0,1]) &= X \\ f^{-1}((a,b)) &= \varnothing \text{ for } 0 \leq a < b \leq 1 \\ f^{-1}([0,b)) &= A \text{ for } 0 < b \leq 1 \\ f^{-1}((a,1]) &= B \text{ for } 0 \leq a < 1, \end{aligned}$$

and X, $\varnothing$, A and B are all open in X. Then f can be extended to a continuous function $\widehat{f} : \beta(X) \to Y$ where $\widehat{f}|_X = f$. As $\widehat{f}$ is continuous,

$$\{0\} \subseteq \widehat{f}(A) \subseteq \widehat{f}(\overline{A}) \subseteq \overline{\widehat{f}(A)} = \overline{f(A)} = \overline{\{0\}} = \{0\}$$

and

$$\{1\} \subseteq \widehat{f}(B) \subseteq \widehat{f}(\overline{B}) \subseteq \overline{\widehat{f}(B)} = \overline{f(B)} = \overline{\{1\}} = \{1\}.$$

Then $\widehat{f}(\overline{A}) = \{0\}$ and $\widehat{f}(\overline{B}) = \{1\}$. Hence $\overline{A} \cap \overline{B} = \varnothing$, since otherwise there would be a $x \in \beta(X)$ such that $\widehat{f}(x) = 0$ and $\widehat{f}(x) = 1$, a contradiction as $\widehat{f}$ is a function. ■

4. Let Y be an arbitrary compactification of X. Show there is a continuous surjective closed map $g : \beta(X) \to Y$ such that $g|_X = \text{id}_X$.

Proof: if Y is a compactification of X, there is an embedding $f : X \to Y$ with $\overline{f(X)} = Y$. Hence, by the properties of the Stone-Čech compactification, and since Y is compact Hausdorff, f can be extended continuously to $g : \beta(X) \to Y$, where $g|_X = f$. As $\beta(X)$ is compact and Y is Hausdorff, the map g is closed. Indeed, let C be a closed subset of $\beta(X)$. As $\beta(X)$ is compact, C is compact, so $g(C)$ is compact in Y. But Y is Hausdorff, so $g(C)$ is closed.

It remains only to show that g is surjective. To do this, we show that $Y \subseteq g(\beta(X))$. As g is an extension of f on X, $f(X) \subseteq g(\beta(X))$. But g is closed, so $g(\beta(X))$ is closed in Y. Thus $Y = \overline{f(X)} \subseteq g(\beta(X))$ and g is surjective. ■

19.4 Exercises

1. Prepare a 2-page summary of this chapter, with important definitions and results.
2. Provide a proof for Theorem 234 (Reprise, Reprise).
3. If $X \neq \beta(X)$, show that $\beta(X)$ is not metrizable.
4. Let X be a discrete.
   a) If $A \subseteq X$, show that $\overline{A}, \overline{X \setminus A} \subseteq_C \beta(X)$ are disjoint.
   b) If $U \subseteq_O \beta(X)$, show that $\overline{U} \subseteq_O \beta(X)$.
   c) Is $\beta(X)$ totally disconnected?

Chapter 20

Introduction to Algebraic Topology

While there are tons of other interesting results and counter-examples in **point set topology**, we have touched upon most of the important ideas of the discipline in Chapters 15-19. In this chapter, we introduce the basic concepts of **algebraic topology**, which is both a precursor and an application of **category theory**, and which provides a stepping stone to **homology theory**, a fascinating (but out-of-scope) offshoot of general topology.

20.1 Fundamental Groups

A **path** in a space X from x to y is a continuous function $p : I = [0, 1] \to X$ where $p(0) = x$ and $p(1) = y$. A **path homotopy** between 2 paths p_0 and p_1 from x_0 to x_1 is a continuous function $F : I \times I \to X$, where

$$F(t, 0) = p_0(t), \quad F(t, 1) = p_1(t), \quad F(0, s) = x_0, \quad F(1, s) = x_1.$$

If such an F exists, we say that p_0 is **(path) homotopic to** p_1 **under** F, which we denote by $p_0 \sim_F p_1$, or $p_0 \sim p_1$ if the dependence on F does not need to be emphasized. Path homotpy is an **equivalence relation** on the set of paths.

Reflexivity: if p is a path from x_0 to x_1 in X, set $F(t, s) = p(t)$ for all s, t. Then $p \sim_F p$.

Symmetry: if p_0, p_1 are homotopic paths from x_0 to x_1 with $p_0 \sim_F p_1$, set $G(t, s) = F(t, 1-s)$ for all s, t. Then $p_1 \sim_G p_0$.

Transitivity: let p_0 and p_1 be paths from x_0 to x_1 with $p_0 \sim_F p_1$, and let p_1 and p_2 be paths from x_1 to x_2 with $p_1 \sim_G p_2$. Then $p_0 \sim_H p_2$, where

$$H(s, t) = \begin{cases} F(t, 2s) & s \in [0, 1/2] \\ G(t, 2s - 1) & s \in [1/2, 1] \end{cases}$$

for all $t, s \in I$. By the pasting lemma (Lemma 213), H is continuous since $F(t, 1) = G(t, 0)$ for all $t \in I$.

Examples (Path Homotopies)

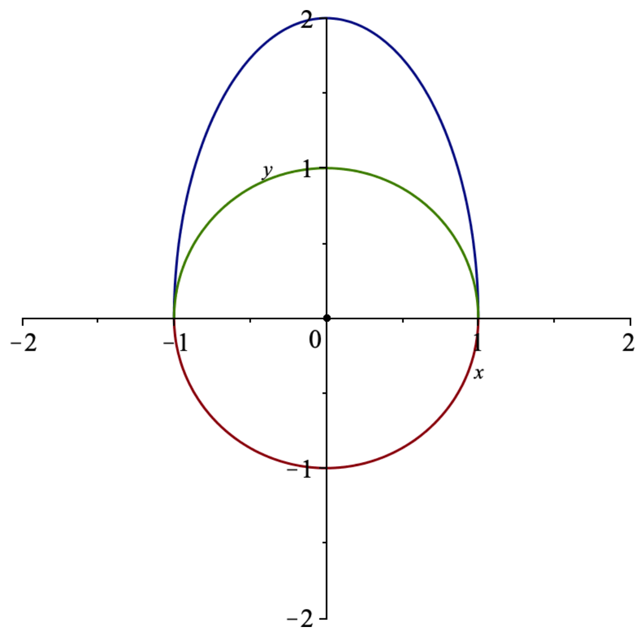
1. Let p and q be any paths with the same endpoints in $\mathbb{R}^n$. Then $p \sim_F q$ where

$$F(t,s) = (1-s)p(t) + sq(t).$$

This path homotopy is called the **straight-line homotopy**.

2. Let p, q, and r be paths from $x_0 = (1,0)$ to $x_1 = (-1,0)$ in the punctured plane $\mathbb{R}^2 \setminus \{0\}$, defined by:

$$\begin{array}{rcll} p(s) & = & (\cos \pi s, \sin \pi s) & \text{(in green)}, \\ q(s) & = & (\cos \pi s, 2\sin \pi s) & \text{(in blue)}, \\ r(s) & = & (\cos \pi s, -\sin \pi s) & \text{(in red)}. \end{array}$$



Then p and q are path homotopic (through the straight-line homotopy, say). But p and r are not path homotopic – we will prove this at a later point. □

The **equivalence class of a path** p is denoted by $[p]$. We show that the equivalence classes of paths behave very much like the elements of a **group**. Let X be a topological space.

Composition If p, q are paths in X from x_0 to x_1 and from x_1 to x_2, respectively, then pq is a path from x_0 to x_2, and we have:

$$pq(t) = \begin{cases} p(2t) & t \in [0, 1/2], \\ q(2t-1) & t \in [1/2, 1]. \end{cases}$$

If $p_0 \sim_F p_1$ from x_0 to x_1 and $q_0 \sim_G q_1$ from x_1 to x_2, define $H : I \times I \to X$ by

$$H(t,s) = \begin{cases} F(2t,s) & t \in [0,1/2] \\ G(2t-1,s) & t \in [1/2,1]. \end{cases}$$

By the pasting lemma, H is continuous since $F(1,s) = G(0,s) = x_1$. Hence $p_0q_0 \sim_H p_1q_1$. Whenever the composition pq is defined, we can define the product of the path classes by $[p][q] = [pq]$.

Associativity If p,q,r are paths in X from x_0 to x_1, x_1 to x_2 and x_2 to x_3 respectively, then $(pq)r$ and $p(qr)$ are paths from x_0 to x_3, and we have:

$$\begin{aligned} (pq)r(t) &= \begin{cases} p(4t) & t \in [0,1/4], \\ q(4t-1) & t \in [1/4,1/2], \\ r(2t-1) & t \in [1/2,1]. \end{cases} \\ p(qr)(t) &= \begin{cases} p(2t) & t \in [0,1/2], \\ q(4t-2) & t \in [1/2,3/4], \\ r(4t-3) & t \in [3/4,1]. \end{cases} \end{aligned}$$

Clearly, $(pq)r \neq p(qr)$. But $(pq)r \sim_F p(qr)$, where

$$F(t,s) = \begin{cases} p\left(\frac{4t}{s+1}\right) & 0 \leq t \leq \frac{1}{4}(s+1), \\ q(4t-1-s) & \frac{1}{4}(s+1) \leq t \leq \frac{1}{4}(s+2), \\ r\left(\frac{4t-s-2}{2-s}\right) & \frac{1}{4}(s+2) \leq t \leq 1. \end{cases}$$

Hence $([p][q])[r] = [p]([q][r])$ whenever these multiplications are defined.

Identities The **constant path** c_x at x is defined by $c_x(t) = x$ for all $t \in I$. If p is a path from x to y, then $c_x \sim_F p \sim_G pc_y$. One gets

$$\begin{aligned} c_xp(t) &= \begin{cases} x & t \in [0,1/2], \\ q(2t-1) & t \in [1/2,1]. \end{cases} \\ pc_y(t) &= \begin{cases} p(2t) & t \in [0,1/2], \\ y & t \in [1/2,1]. \end{cases} \end{aligned}$$

Then

$$\begin{aligned} F(t,s) &= \begin{cases} x & t \in [0,(1-s)/2], \\ p\left(\frac{2t+s-1}{s+1}\right) & t \in [0,(1-s)/2], \end{cases} \\ G(t,s) &= \begin{cases} p\left(\frac{2t}{2-s}\right) & t \in [0,1-s/2], \\ y & t \in [1-s/2,1]. \end{cases} \end{aligned}$$

Then F and G are the required homotopies. Hence for any path p from x to y, $[c_x][p] = [p]$ and $[p] = [p][c_y]$.

Inverses If p is a path in X from x to y, then $\overline{p}$ is a path from y to x defined by $\overline{p}(t) = p(1-t)$ with $p\overline{p} \sim_F c_x$ and $\overline{p}p \sim_G c_y$, where

$$F(t,s) = \begin{cases} p\left(2t\right) & 0 \leq t \leq \frac{s}{2}, \\ p(s) & \frac{s}{2} \leq t \leq 1 - \frac{s}{2}, \\ p(2-2t) & 1 - \frac{s}{2} \leq t \leq 1. \end{cases}$$

Note that $\overline{\overline{p}} = p$, so we get

$$\begin{aligned} G(t,s) &= \begin{cases} \overline{p}\left(2t\right) & 0 \leq t \leq \frac{s}{2}, \\ \overline{p}(s) & \frac{s}{2} \leq t \leq 1 - \frac{s}{2}, \\ \overline{p}(2-2t) & 1 - \frac{s}{2} \leq t \leq 1. \end{cases} \\ &= \begin{cases} p\left(1-2t\right) & 0 \leq t \leq \frac{s}{2}, \\ p(1-s) & \frac{s}{2} \leq t \leq 1 - \frac{s}{2}, \\ p(2t-1) & 1 - \frac{s}{2} \leq t \leq 1. \end{cases} \end{aligned}$$

Hence $[p][\overline{p}] = [c_x]$ and $[\overline{p}][p] = [c_y]$, which means that $[\overline{p}] = [p]^{-1}$.

But it is not always possible to multiply path classes, as two paths may not have matching endpoints, so the group idea is **not complete**. To remedy the situation, we introduce a new concept. A path in X from x to x is a **loop** in X based at x. When p is a loop at x we call the path class $[p]$ a loop at x.

For a fixed $x_0 \in X$, if we consider only loops based at x_0, then pq is always defined. This means that the composition of path classes is always defined and so, for any path classes α, β, γ, with ε the path class of the constant path c_{x_0}, we have

$$(\alpha\beta)\gamma = \alpha(\beta\gamma), \quad \alpha\varepsilon = \varepsilon\alpha = \alpha, \quad \alpha\alpha^{-1} = \alpha^{-1}\alpha = \varepsilon;$$

the path classes of loops in X at x_0 thus form a group, the **fundamental group of X based at** x_0, denoted by $\pi(X, x_0)$. It is also sometimes known as the **first homotopy group of X at** x_0, denoted by $\pi_1(X, x_0)$. The fundamental group does depend on the chosen base point.

Examples (Fundamental Group)

1. If $X = \mathbb{R}^n$ and $x_0 = 0$, then $\pi(\mathbb{R}^n, 0) = \{\varepsilon\}$, as every loop at 0 is path homotopic to the constant loop c_0.

2. If X is any convex subset of $\mathbb{R}^n$ and $x_0 \in X$, then $\pi(X, x_0) = \{\varepsilon\}$, as every loop at x_0 is path homotopic to the constant loop c_{x_0} through the straight-line homotopy.

3. If $X = \mathbb{R}^n \setminus \{0\}$ and p, q and r are defined as in the 2nd example on p. 458, then $p\overline{q}$ and $p\overline{r}$ are two loops based at $(-1, 0)$. But these loops are not path homotopic and so their path classes differ, which means that $\pi(X, (-1, 0))$ is not the trivial group. The fundamental group of the punctured plane will be computed in Section 20.3.

If X is a path-connected space for which there exists $x_0 \in X$ such that $\pi(X, x_0) = \{\varepsilon\}$, we say that X is **simply connected**. The reason why we only need one $x_0 \in X$ is that the fundamental groups of path-connected spaces are independent of the chosen base point.

Theorem 270
If X is path-connected, then $\pi(X, x) \cong \pi(X, y)$ for $x, y \in X$.

Proof: As X is path-connected, there is a path class γ from x to y. Define $\hat{\gamma} : \pi(X, x) \to \pi(X, y)$ by $\hat{\gamma}(\alpha) = \gamma^{-1}\alpha\gamma$. We show that $\hat{\gamma}$ is the desired isomorphism. First, let $\alpha, \beta \in \pi(X, x)$. Then

$$\hat{\gamma}(\alpha)\hat{\gamma}(\beta) = \gamma^{-1}\alpha\gamma\gamma^{-1}\beta\gamma = \gamma^{-1}\alpha\beta\gamma = \hat{\gamma}(\alpha\beta),$$

so $\hat{\gamma}$ is a homomorphism. The reverse class $\overline{\gamma}$ also provides a fundamental group homomorphism $\widehat{\overline{\gamma}} : \pi(X, y) \to \pi(X, x)$ defined by $\widehat{\overline{\gamma}}(\xi) = \gamma\xi\gamma^{-1}$. Then $\hat{\gamma}^{-1} = \widehat{\overline{\gamma}}$, which implies that γ is an isomorphism. ■

In the proof of Theorem 270, if we use a different path class δ from x to y, we get a different isomorphism $\widehat{\delta} : \pi(X, x) \to \pi(X, y)$. But

$$\widehat{\delta}^{-1}\widehat{\gamma}(\alpha) = \delta\gamma^{-1}\alpha\gamma\delta^{-1} = (\delta\gamma^{-1})\alpha(\delta\gamma^{-1})^{-1}$$

for all $\alpha \in \pi(X, x)$. Hence $\widehat{\delta}$ and $\widehat{\gamma}$ differ by an **inner automorphism**.

Suppose $\varphi : X \to Y$ is a continuous function and $p : I \to X$ is a path, then $\varphi \circ p : I \to Y$ is a path, denoted φp. If the composition pq is defined, then $\varphi(pq) = (\varphi p)(\varphi q)$. Thus, if $p \sim_F q$, then $\varphi p \sim_{\varphi F} \varphi q$, and φ induces a homomorphism of path classes

$$\varphi^* : \pi(X, x) \to \pi(Y, \varphi(x)),$$

defined by $\varphi^*([p]) = [\varphi p]$ for all $[p] \in \pi(X, x)$. If furthermore $\psi : Y \to Z$ is a continuous function, then $(\psi\varphi)^* = \psi^*\varphi^*$. From this, if φ is a homeomorphism, $(\varphi^{-1})^* = (\varphi^*)^{-1}$ and φ^* is an isomorphism. As a result, if X is homeomorphic to Y, then $\pi(X, x)$ is isomorphic to $\pi(Y, \varphi(x))$, where φ is the homeomorphism between X and Y.

Corollary 271
If $\pi(X, x) \ncong \pi(Y, y)$, then X and Y are not homeomorphic.

Note that φ^* need not be surjective (injective) when φ is surjective (injective).

1. Let $X = \mathbb{R}$, $Y = S^1$ and define $\varphi : \mathbb{R} \to S^1$ by $\varphi(x) = e^{2\pi i x}$. Then φ is continuous and surjective, and $\varphi(0) = 1$. But $\pi(\mathbb{R}, 0) = \{\varepsilon_0\}$, so $\varphi^*(\pi(\mathbb{R}, 0)) = \{\varepsilon_1\}$. As we shall see in Section 20.3, $\pi(S^1, 1) = \mathbb{Z}$. Hence φ^* is not surjective.

2. Let $X = S^1$, $Y = \mathbb{C}$, and $\varphi : S^1 \to \mathbb{C}$ with $\varphi(z) = z$. Then φ is continuous and injective, and $\varphi(1) = 1$. But $\pi(S^1, 1) = \mathbb{Z}$ and $\pi(\mathbb{C}, 1) = \{\varepsilon_1\}$, so $\ker \varphi^* = \pi(S^1, 1) \neq \{\varepsilon_1\}$ and φ^* is not injective.

20.2 Covering Spaces

Suppose $p : \widetilde{X} \to X$ is a continuous map. Let V be a neighbourhood of $x \in X$. We say that V **is evenly covered by** p **at** x if $p^{-1}(V)$ can be written as a disjoint union of sets $\widetilde{V}$ (the **slices** of $p^{-1}(V)$) such that the restriction $p|_{\widetilde{V}} : \widetilde{V} \to V$ is a homeomorphism. If for every $x \in X$, there is some neighbourhood V of x that is evenly covered by p, then p is a **covering map** and $(\widetilde{X}, p)$ is a **covering space of** X. Note that a covering map is automatically surjective.

Example (Covering Spaces)

1. Let $\widetilde{X} = \mathbb{R}$, $X = S^1$ and define $p : \mathbb{R} \to S^1$ by $p(\widetilde{x}) = e^{2\pi i\widetilde{x}}$. Let $z \in S^1$. Then there exists $\theta_z \in \mathbb{R}$ such that $z = e^{2\pi i\theta_z}$ and $p^{-1}(z) = \{\theta_z + n \mid n \in \mathbb{Z}\}$. Let $V_z = \left\{e^{2\pi i\phi} \mid |\phi - \theta_z| < \frac{1}{2}\right\}$. We show that V_z is evenly covered by p and so that $(\mathbb{R}, p)$ is a covering space of S^1. Note that $p^{-1}(V_z) = \bigsqcup_{n\in\mathbb{Z}} \widetilde{V}_n$, where $\widetilde{V}_n = (\theta_z + n - \frac{1}{2}, \theta_z + n - \frac{1}{2})$ for all $n \in \mathbb{Z}$. But, for all $n \in \mathbb{Z}$,

   $$p(\widetilde{V}_n) \quad = \{e^{2\pi i\phi} \mid \phi \in \widetilde{V}_n\} = \{e^{2\pi i\phi} \mid |\phi - \theta_z| < 1/2\} = V_z,$$

   and $p|_{\widetilde{V}_n}^{-1}(V_z) = \widetilde{V}_n$, so $p|_{\widetilde{V}_n}$ is an homeomorphism and V_z is evenly covered.

2. Let $p : \widetilde{X} \to X$ be a homeomorphism. Then every open set $U \subseteq X$ is evenly covered by p since $p^{-1}(U) \simeq U$. Hence $(\widetilde{X}, p)$ is a covering space of X.

3. Let $\widetilde{X} = S^1$, $X = S^1$ and define $p : S^1 \to S^1$ by $p(z) = z^n$, for all $z \in S^1$ and for some $n \in \mathbb{Z}$. Let $z \in S^1$. Then there exists $\theta_z \in \mathbb{R}$ such that $z = e^{2\pi i\theta_z}$. By definition, $p^{-1}(z) = \left\{e^{\frac{2\pi i m}{n}\theta_z} \mid 0 \le m \le n-1\right\}$. Let $U_z = \left\{e^{2\pi i\phi} \mid |\phi - \theta_z| < \frac{1}{4n}\right\}$. We show that U_z is evenly covered by p and so that (S^1, p) is a covering space of S^1. But $p^{-1}(U_z) = \bigsqcup_{m=0}^{n-1} \widetilde{U}_m$, where $\widetilde{U}_m = \left\{e^{2\pi i\phi} \mid |\phi + m - \theta_z| < \frac{1}{4n}\right\}$ for all $0 \le m \le n-1$, and so

   $$p(\widetilde{U}_m) = \{e^{2\pi i\phi} \mid |\phi - \theta_z| < 1/(4n)\} = U_z$$

   for all $0 \le m \le n-1$, hence $p|_{\widetilde{U}_m}^{-1}(U_z) = \widetilde{U}_n$, so $p|_{\widetilde{U}_n}$ is an homeomorphism and U_z is evenly covered.

4. Let $\widetilde{X} = S^2$ and $X = \mathbb{R}P^2$ be the real projective plane. Then the quotient map $p : S^2 \to \mathbb{R}P^2$ where $p(v) = p(-v)$ for all $v \in S^2$ is a covering map.

A continuous function $f : X \to Y$ is a **local homeomorphism** if for each $x \in X$, there is a neighbourhood V of x such that $f|_V : V \to f(V)$ is a homeomorphism. Consequently, every covering map is a local homeomorphism. But the converse is not necessarily true.

Example let $X = \mathbb{R}^+$, $Y = S^1$ and define $p : \mathbb{R}^+ \to S^1$ by $p(x) = e^{2\pi i x}$ for $x \in \mathbb{R}^+$. Then p is continuous and surjective. Let $x \in \mathbb{R}^+$. Any basic neighbourhood $(x - \varepsilon, x + \eta)$ in $\mathbb{R}^+$, where $\varepsilon + \eta < 1/2$ is mapped homeomorphically to $\left\{e^{2\pi i \phi} \mid -\varepsilon < \phi - x < \eta\right\}$ by p. This makes p a local homeomorphism.

But p is not a covering map. Indeed, if U is an evenly covered neighbourhood of $e^{2\pi i}$ *via* p, then $p^{-1}(U) = \bigsqcup_{n=0}^{\infty} V_n$, where V_n is a small neighbourhood around n when $n > 0$ and $V_0 = (0, \varepsilon)$ for some small ε. But $p(V_0)$ is not homeomorphic to U. So there is no neighbourhood of $e^{2\pi i}$ which is evenly covered by p. $\square$

Suppose $p : \widetilde{X} \to X$ is a covering map and $f : Y \to X$ is a continuous function. A **lift** of f is a map $\widetilde{f} : Y \to \widetilde{X}$ such that $p\widetilde{f} = f$. The following theorems show that paths and path homotopies can be lifted.

Theorem 272 (Path Lifting Property)
Suppose $p : \widetilde{X} \to X$ is a covering map and $f : I \to X$ is a path with $f(0) = x_0$. For each $\widetilde{x}_0 \in p^{-1}(x_0)$, there is a unique path $\widetilde{f} : I \to \widetilde{X}$ such that $\widetilde{f}(0) = \widetilde{x}_0$ and $p\widetilde{f} = f$.

$$\begin{tikzcd} & \widetilde{X} \arrow[d, "p"] \\ I \arrow[ur, "\widetilde{f}"] \arrow[r, "f"'] & X \end{tikzcd}$$

Proof: the sets $f^{-1}(V)$ where V is a **canonical** (which is to say, evenly covered) neighbourhood of a point in $f(I)$ give an open covering $\mathfrak{F}$ of I. As I is a compact metric space, Theorem 245 guarantees the existence of a Lebesgue number ε of $\mathfrak{F}$. Let $n \in \mathbb{N}$ be such that $\frac{1}{n} < \varepsilon$. Let $t_m = \frac{m}{n}$ for $1 \leq m \leq n$ and set $t_0 = 0$. Then $I_m = [t_{m-1}, t_m]$ has diameter less than ε, so it lies in $f^{-1}(V_m)$ for some canonical V_m and $f(I_m) \subseteq V_m$ for $1 \leq m \leq n$.

But V_1 is a canonical neighbourhood of x_0. Let $\widetilde{V}_1$ be the slice of $p^{-1}(V_1)$ containing $\widetilde{x}_0$. Define $\widetilde{f}$ on I_1 by
$$\widetilde{f}(t) = p_1^{-1} f(t),$$
where $p_1 = p|_{\widetilde{V}_1}$. As f is continuous and p_1 is a homeomorphism, $\widetilde{f}$ is continous on I_1.

Now suppose $\widetilde{f}$ is defined on $[0, t_m]$ and let $x_m = f(t_m)$ and $\widetilde{x}_m = \widetilde{f}(t_m)$. Take $\widetilde{V}_{m+1}$ to be the slice of $p^{-1}(V_m)$ containing $\widetilde{x}_m$ and let $p_{m+1} = p|_{\widetilde{V}_{m+1}}$. Define $\widetilde{f}$ on I_{m+1} by
$$\widetilde{f}(t) = p_{m+1}^{-1} f(t).$$

Since $\widetilde{f}$ is defined at t_m, the pasting lemma guarantees that $\widetilde{f}$ is continuous on $[0, t_{m+1}]$. After n steps, the continuous function $\widetilde{f}$ is defined on I and, by construction, $p\widetilde{f} = f$.

Now suppose $g : I \to \widetilde{X}$ is another path such that $g(0) = \widetilde{x}_0$ and $pg = f$. By construction $p_1 g = p_1 \widetilde{f}$ on I_1. Since p_1 is a homeomorphism, $g = \widetilde{f}$ on I_1. Using an argument identical to that used in the construction of $\widetilde{f}$, if $g = \widetilde{f}$ on $[0, t_m]$, then $g = \widetilde{f}$ on $[t_m, t_{m+1}]$. Recursively, $g = \widetilde{f}$ on I. ■

Theorem 273 (Square Lifting Property)
Suppose $p : \widetilde{X} \to X$ is a covering map and $F : I \times I \to X$ is a continuous function with $F(0,0) = x_{0,0}$. For each $\widetilde{x}_{0,0} \in p^{-1}(x_{0,0})$, there is a unique lift of F to $\widetilde{F} : I \times I \to \widetilde{X}$ where $\widetilde{F}(0,0) = \widetilde{x}_{0,0}$.

$$\begin{tikzcd} & \widetilde{X} \arrow[d, "p"] \\ I \times I \arrow[ur, "\widetilde{F}"] \arrow[r, "F"'] & X \end{tikzcd}$$

Proof: the sets $F^{-1}(V)$ where V is a canonical neighbourhood of a point in $F(I \times I)$ form an open covering $\mathfrak{F}$ of $I \times I$ with Lebesgue number ε. Subdivide $I \times I$ into n^2 small squares of diameter less than ε. Using arguments similar to that of the previous proof, lift F to $\widetilde{F}$ on $I_1 \times I_1$, then across the base of $I \times I$ on $I \times I_1$. Next, fill the square one layer at a time. Special care has to be taken to extend $\widetilde{F}$ to $I_k \times I_{l+1}$ from the previous rectangles. This hinges on the fact that the union of the bottom and leftmost edges is connected. Then $\widetilde{F} : I \times I \to \widetilde{X}$ is uniquely defined. ■

Theorem 274
If $f_0, f_1 : I \to X$ are paths with initial point x_0, $p : \widetilde{X} \to X$ is a covering map and $p(\widetilde{x}_0) = x_0$, then the lifts $\widetilde{f}_0, \widetilde{f}_1 : I \to \widetilde{X}$ with initial point $\widetilde{x}_0$ are path homotopic under $\widetilde{F}$ if and only if f_0, f_1 are path homotopic under F, where $\widetilde{F}$ is the unique lift of F based at $\widetilde{x}_0$.

Proof: suppose $\widetilde{f}_0 \sim_{\widetilde{F}} \widetilde{f}_1$, then let $F = p\widetilde{F}$, so $f_0 \sim_F f_1$. Conversely, suppose $f_0 \sim_F f_1$ and let $\widetilde{F}$ be the lift of F obtained by the previous theorem. Then

$$p\widetilde{F}(t,0) = F(t,0) = f_0(t),$$

so $\widetilde{F}(t,0)$ is a lift of f_0 at $\widetilde{F}(0,0) = \widetilde{x}_0$. By uniqueness of lifts, $\widetilde{F}(t,0) = \widetilde{f}_0(t)$.

Similarly, $\widetilde{F}(t,1) = \widetilde{f}_1(t)$. Now

$$p\widetilde{F}(0,s) = f(0,s) = x_0$$

and $\widetilde{F}(0,s)$ is a lift of the constant path $e_{x_0}(s)$. But the constant path $e_{\widetilde{x}_0}(s) = \widetilde{x}_0$ is a lift of e_{x_0}. By uniqueness of lifts,

$$\widetilde{F}(0,s) = e_{\widetilde{x}_0}(s) = \widetilde{x}_0.$$

Similarly $\widetilde{F}(1,s)$ is a constant path and $\widetilde{F}$ is a path homotopy. ■

Corollary 275
If X and $\widetilde{X}$ are path-connected, then $p^{-1}(x)$ has the same cardinality at every point $x \in X$.

Proof: for any path f in X from x to y, if $\widetilde{x} \in p^{-1}(x)$, then the lift of f to $\widetilde{f}$ with initial point $\widetilde{x}$ gives a path in $\widetilde{X}$ from $\widetilde{x}$ to $\widetilde{f}(1) = \widetilde{y}$. Define $\varphi : p^{-1}(x) \to p^{-1}(y)$ by $\varphi(\widetilde{x}) = \widetilde{y}$.

For $\overline{f}$ the reverse path of f from y to x, we get a unique lift from $\widetilde{y}$ to some terminal point. But that terminal point has to be $\widetilde{x}$, since $\widetilde{\overline{f}} = \overline{\widetilde{f}}$. Thus $\overline{\varphi} : p^{-1}(y) \to p^{-1}(x)$ and $\overline{\varphi} = \varphi^{-1}$. ■

The cardinality of $p^{-1}(x)$ is the number of **sheets** of the covering.

Examples (SHEETS)

1. The map $p : S^1 \to S^1$ defined by $p(z) = z^n$ is an n-sheeted covering.
2. The map $p : \mathbb{R} \to S^1$ defined by $p(x) = e^{2\pi i x}$ is an ω-sheeted covering. □

20.3 Fundamental Groups of S^1 and $\mathbb{R}^2 \setminus \{\mathbf{0}\}$

In this section we show how to compute the fundamental group of the circle and of the punctured plane, using techniques introduced in the previous section.

Theorem 276
The fundamental group of S^1 is infinite cyclic, that is it is isomorphic to the additive group $\mathbb{Z}$.

Proof: since S^1 is path-connected, the fundamental group can be based at any point of S^1. For convenience, take $z = e^{2\pi i} = 1$. The map $p : \mathbb{R} \to S^1$ defined by $p(x) = e^{2\pi i x}$ is a covering map.

Let $\alpha \in \pi(S^1, 1)$. Then $\alpha = [f]$, where f is a loop in S^1 based at 1. Then, by the path lifting property, there exists a unique $\widetilde{f}$ with initial point $0 \in p^{-1}(1)$ such that the following diagram commutes.

$$\begin{tikzcd}
 & \mathbb{R} \arrow[d, "p"] \\
I \arrow[ur, "\widetilde{f}"] \arrow[r, "f"'] & S^1
\end{tikzcd}$$

Then

$$p\widetilde{f}(0) = p(0) = 1 \quad \text{and} \quad p\widetilde{f}(1) = f(1) = 1.$$

Hence $\widetilde{f}(1) \in \mathbb{Z}$, say $\widetilde{f}(1) = n$. This integer is independent of the choice of the representative f, by Theorem 274. Define a map $\varphi : \pi(S^1, 1) \to \mathbb{Z}$ by $\varphi(\alpha) = \widetilde{f}(1)$. We show that φ is an isomorphism, which yields the desired result.

φ **is a homomorphism:** Let $\alpha = [f], \beta = [g] \in \pi(S^1, 1)$. By construction, $\varphi(\alpha) = \widetilde{f}(1) = n$ and $\varphi(\beta) = \widetilde{g}(1) = m$ for some $m, n \in \mathbb{Z}$. Define $\widetilde{h}$ by $\widetilde{h}(t) = n + \widetilde{g}(t)$. Then $\widetilde{f}\widetilde{h}$ is a path from 0 to $n + m$. Then

$$p(\widetilde{h}(t)) = e^{2\pi i(n + \widetilde{g}(t))} = p(\widetilde{g}(t)) = g(t)$$

and $p(\widetilde{f}.\widetilde{h}) = p(\widetilde{f}).p(\widetilde{h}) = fg$, so $\widetilde{f}.\widetilde{h}$ is a lift of fg starting at 0. Consequently,

$$\varphi(\alpha\beta) = \widetilde{f}.\widetilde{h}(1) = n + m = \varphi(\alpha) + \varphi(\beta).$$

φ **is injective:** Suppose $\varphi(\alpha) = 0$ for $\alpha = [f]$. Then, if $\widetilde{f}$ is a lift of f starting at 0, $\widetilde{f}(1) = 0$ and so $\widetilde{f}$ is a loop in $\mathbb{R}$ based at 0. But $\mathbb{R}$ is simply connected, so $\widetilde{f} \sim e_0$. By Theorem 274, $f \sim e_1$, or $\alpha = \varepsilon_1$. Then $\ker \varphi = \{\varepsilon_1\}$.

φ **is surjective:** For any $n \in \mathbb{Z}$, let $\widetilde{f}(t) = nt$. Then $\widetilde{f}$ is a path from 0 to n and $f = p\widetilde{f}$ is a loop in S^1. Let $\alpha = [f]$. Then $\varphi(\alpha) = \widetilde{f}(1) = n$. ■

Interestingly, the punctured plane has the same fundamental group as the circle.

Theorem 277
The fundamental group of $\mathbb{R}^2 \setminus \{0\}$ *is infinite cyclic.*

Proof: the point $b = (1, 0)$ belongs to both S^1 and $X = \mathbb{R}^2 \setminus \{0\}$. Let $i : S^1 \to X$ be the inclusion map and $r : X \to S^1$ be the **radial map** defined by $r(z) = z/|z|$ on X.

Both i and r are continuous, and these maps induce the homomorphisms

$$i^* : \pi(S^1, b) \to \pi(X, b) \quad \text{and} \quad r^* : \pi(X, b) \to \pi(S^1, b).$$

Note that $ri = \mathrm{id}_{S^1}$ and so that $r^*i^* = \mathrm{id}_{\pi(S^1,b)}$. Then i^* is injective and r^* is surjective. It remains only to show that $i^*r^* = \mathrm{id}_{\pi(X,b)}$.

Let $\alpha = [f] \in \pi(X, b)$ and define $F : I \times I \to X$ by

$$F(t, s) = (1 - s)f(t) + s\frac{f(t)}{|f(t)|}.$$

Then F is continuous and defined everywhere since $|f(t)| \neq 0$ in X. Furthermore $F(t, s) \neq 0$, as can be easily verified.

$$F(0, s) = F(1, s) = b \quad \text{and} \quad F(t, 0) = f(t),\ F(t, 1) = \frac{f(t)}{|f(t)|}.$$

Then if $g = f/|f|$, F is a path homotopy between f and g. Hence $\alpha = [g]$. But g is a loop in S^1 based at b, so $r(g(t)) = g(t)$ and

$$i^*r^*(\alpha) = i^*([r(g)]) = i^*([g]) = \alpha.$$

Then $i^*r^* = \mathrm{id}_{\pi(X,b)}$ and i^* and r^* are isomorphisms. Consequently, $\pi(X, b)$ is isomorphic to the additive group $\mathbb{Z}$. ■

This last result tells us that puncturing the plane changes the topological nature of $\mathbb{R}^2$.

Corollary 278
$\mathbb{R}^2 \setminus \{0\}$ *and* $\mathbb{R}^2$ *are not homeomorphic.*

Note that $\mathbb{R}^n \setminus \{0\}$ and $\mathbb{R}^n$ **are** homeomorphic when $n > 2$, however.

A subspace A of X is a **retract of** X if there is a continuous function $r : X \to A$ such that $r(a) = a$ for all $a \in A$. Such a function is called a **retraction**. If $r : X \to A$ is a retraction, $ri = \mathrm{id}_A$ where $i : A \to X$ is the inclusion mapping. If $a \in A$, this induces $r^*i^* = \mathrm{id}_{\pi(A,a)}$, so that r^* is surjective and i^* is injective.

Examples (RETRACTS)

1. S^1 is a retract of $\mathbb{R}^2 \setminus \{0\}$ with the radial map $r : \mathbb{R}^2 \setminus \{0\} \to S^1$.
2. Since $\pi(\mathbb{R}^2, 0) = \{\varepsilon_0\}$ and $\pi(S^1, 1) \cong \mathbb{Z}$, there is no surjective homomorphism $r^* : \pi(\mathbb{R}^2, 0) \to \pi(S^1, 1)$. Hence there cannot be a retraction $r : \mathbb{R}^2 \to S^1$, so S^1 is not a retract of $\mathbb{R}^2$.

3. The disc $D = \{z \mid |z| \leq 1\}$ is a retract of $\mathbb{C}$ with the continuous map $r : \mathbb{C} \to D$ defined by

$$r(z) = \begin{cases} z & \text{if } |z| \leq 1, \\ z/|z| & \text{if } |z| > 1. \end{cases}$$

□

Two continuous maps $f, g : X \to Y$ are **homotopic** if $\exists$ a continuous map $F : X \times I \to Y$ such that $F(x, 0) = f(x)$ and $F(x, 1) = g(x)$ for all $x \in X$. A subset A of X is a **strong deformation retract** if there is a retraction $r : X \to A$ and a homotopy $F : X \times I \to X$ such that $F(x, 0) = x$ and $F(x, 1) = r(x)$ for all $x \in X$ and $F(a, t) = a$ for all $a \in A$, that is if $ir \sim_F \mathrm{id}_X$. The importance of strong deformation retracts is explained by the following theorem.

Theorem 279
If A is a strong deformation retract of X, then $\pi(X, a) \simeq \pi(A, a)$ for $a \in A$.

Proof: suppose $r : X \to A$ is a retraction. Then the induced homomorphisms

$$r^* : \pi(X, a) \to \pi(A, a) \quad \text{and} \quad i^* : \pi(A, a) \to \pi(X, a)$$

are respectively surjective and injective. It will be sufficient to show that i^* is also surjective. Let f be a loop in X based at $a \in A$. Then $rf = g$ is a loop in A based at a. Let F be a homotopy between ir and id_X. Then, setting $F_1(t, s) = F(f(t), s)$ yields $f \sim_{F_1} g$, since

$$F_1(t, 0) = F(f(t), 0) = f(t) \quad \text{and} \quad F_1(t, 1) = F(f(t), 1) = rf(t) = g(t).$$

Therefore $[g] = [f]$ and $i^*([g]) = [f]$. Hence i^* is surjective. ■

Suppose that $f : X \to Y$ and $g : Y \to X$ are continuous functions such that $fg \sim \mathrm{id}_Y$ and $gf \sim \mathrm{id}_X$, then X and Y are said to be **homotopy equivalent**, denoted $X \equiv Y$, and f and g are said to be **homotopy inverses**. The relation $\equiv$ is an equivalence relation. Reflexivity and symmetry are trivially shown. To show that $\equiv$ is transitive, let $X \equiv Y$ and $Y \equiv Z$. Then there exist continuous functions

$$f : X \to Y,\ g : Y \to Z,\ h : Y \to Z \text{ and } k : Z \to Y$$

such that $fg \sim \mathrm{id}_X,\ gf \sim \mathrm{id}_Y,\ hk \sim \mathrm{id}_Z$ and $kh \sim \mathrm{id}_Y$. Then

$$(hf)(gk) \sim h(fg)k \sim h\,\mathrm{id}_Y\,k \sim hk \sim \mathrm{id}_Z \quad \text{and} \quad (gk)(hf) \sim g(kh)f \sim g\,\mathrm{id}_Y\,f \sim gk \sim \mathrm{id}_X,$$

so $X \equiv Z$ through the homotopy inverses hf and gk.

Examples (STRONG DEFORMATION RETRACTS)

1. The figure 8 is a strong deformation retract of the doubly-punctured plane. Intuitively, this is done by representing the figure 8 as two petals teaching the axes at the origin. Puncture each petal once. Points interior to the petal slide radially away from the puncture. Points outside the petals slide radially towards the origin until they reach a petal. Timing it so that each point takes exactly one unit of time to reach the appropriate petal yields the desired homotopy.

2. If A is a strong deformation retract of X, then $A \equiv X$. Indeed, let $r : X \to A$ be a retraction. Then $ri = \text{id}_A$ and $ir \sim \text{id}_X$.

From this point on, the spaces we consider are all path-connected.

Theorem 280
Suppose $f, g : X \to Y$ are continuous functions, $x_0 \in X$ and $f(x_0) = y_0$, $g(x_0) = y_1$. If f and g are homotopic, then there is a path class α from y_0 to y_1 such that $g^* = \widehat{\alpha} f^*$, where $f^* : \pi(X, x_0) \to \pi(Y, y_0)$, $g^* : \pi(X, x_0) \to \pi(Y, y_1)$ and $\widehat{\alpha} : \pi(Y, y_0) \to \pi(Y, y_1)$.

Proof: suppose $F : X \times I \to Y$ is a homotopy between f and g, that is, suppose $F(x, 0) = f(x)$ and $F(x, 1) = g(x)$. Let $q : I \to Y$ be such that $q(s) = F(x_0, s)$. As F is continuous, q is a path from y_0 to y_1 since

$$\begin{aligned} q(0) &= F(x_0, 0) = f(x_0) = y_0 \\ q(1) &= F(x_0, 1) = g(x_0) = y_1. \end{aligned}$$

Let $\alpha = [q]$. For any loop h in X based at x_0, we show that

$$g^*([h]) = \widehat{\alpha} f^*([h]),$$

that is $[g \circ h] = \widehat{\alpha}([f \circ h]) = [\overline{q}][f \circ h][q]$, or $g \circ h \sim (\overline{q}(f \circ h))q$.

Let $e = e_{y_1}$. Then $g \circ h \sim e(g \circ h) \sim (e(g \circ h))e$. We next show that

$$(e(g \circ h))e \sim_G (\overline{q}(f \circ h))q$$

for an appropriate path homotopy G. Define $G : I \times I \to Y$ by

$$G(t, s) = \begin{cases} q(1 - 4t(1 - s)) & t \in [0, 1/4], \\ F(h(4t - 1), s) & t \in [1/4, 1/2], \\ q(2t - 1 + 2(1 - t)s) & t \in [1/2, 1]. \end{cases}$$

At $t = \frac{1}{4}$, $q(s) = F(x_0, s)$ and at $t = \frac{1}{2}$, $F(x_0, s) = q(s)$ so, by the pasting lemma, G is continuous on $I \times I$. Now $G(0, s) = G(1, s) = q(1) = y_1$ and

$$\begin{aligned} G(t,0) &= \begin{cases} q(1-4t) & t \in [0,1/4], \\ F(h(4t-1),0) & t \in [1/4,1/2], \\ q(2t-1) & t \in [1/2,1], \end{cases} \\ G(t,1) &= \begin{cases} q(1) & t \in [0,1/4], \\ F(h(4t-1),1) & t \in [1/4,1/2], \\ q(1) & t \in [1/2,1]. \end{cases} \end{aligned}$$

Then $G(t, 0) = (\overline{q}(f \circ h))q(t)$, $G(t, 1) = (e(g \circ h))e(t)$ and $g^* = \widehat{\alpha} f^*$. ■

The existence of homotopy inverses between X and Y imply that the corresponding fundamental groups are isomorphic.

Corollary 281
If $f : X \to Y, g : Y \to X$ are homotopy inverses, then $f^* : \pi(X, x_0) \to \pi(Y, f(x_0))$ is an isomorphism.

Proof: let $y_0 = f(x_0)$ and $x_1 = g(y_0)$. As f and g are homotopy inverses, $g \circ f \sim \mathrm{id}_X$ and the preceding theorem yields $(g \circ f)^* = \widehat{\alpha}\, \mathrm{id}^*_{\pi(X,x_0)} = \widehat{\alpha}$ for some path class α from x_0 to x_1. Then $g^* f^* = \widehat{\alpha}$. As $\widehat{\alpha}$ is an isomorphism, g^* is surjective and f^* is injective. It is then sufficient to show that g^* is injective.

Let $y_1 = f(x_1)$ and denote by f_1^* the homomorphism induced by f from $\pi(X, x_1)$ to $\pi(Y, y_1)$. As before, $fg \sim \mathrm{id}_Y$ and the preceding theorem yields $(f \circ g)^* = \widehat{\beta}\, \mathrm{id}^*_{\pi(Y,f(x_0))} = \widehat{\beta}$ for some path class β from y_0 to y_1. But this means that g^* is injective as $\widehat{\beta}$ is an isomorphism. Hence g^* is an isomorphism and $f^* = (g^*)^{-1}\widehat{\alpha}$ is an isomorphism. ■

Note that X and Y may have isomorphic fundamental groups yet fail to be homeomorphic and/or homotopy equivalent (compare with Corollary 271).

Examples

1. Consider $\mathbb{R}$ in the usual topology and the singleton set $\{*\}$. We have seen that $\pi(\mathbb{R}) = \pi(\{*\}) = \{\varepsilon\}$, but $\mathbb{R}$ and $\{*\}$ are not homeomorphic since $\{*\}$ is compact but $\mathbb{R}$ isn't.
2. Consider S^2 in the usual topology and the singleton set $\{0\} \subseteq \mathbb{R}^3$. We can show (see next section) that $\pi(S^2) = \pi(\{0\}) = \{\varepsilon\}$, but S^2 and $\{0\}$ are not homotopy equivalent (this is harder to prove). □

20.4 Special Seifert-Van Kampen Theorem

The special Seifert-Van Kampen theorem allows us to determine when the fundamental group of a space is . The following lemma will be helpful.

Lemma 282
Suppose $f : I \to X$ is a path and $0 = a_0 < a_1 < \ldots < a_n = 1$. Define $f_i : I \to X$ by $f_i(t) = f((1-t)a_{i-1} + ta_i)$ for $1 \leq i \leq n$. Then

$$f \sim f_1(f_2(\cdots f_n)\cdots).$$

Proof: left as an exercise.

The main result is stated and proven below.

Theorem 283 (Special Seifert-Van Kampen Theorem)
Let U, V, and $U \cap V$ be non-empty, open, path-connected subsets of $X = U \cup V$. Let $x_0 \in U \cap V$. If the inclusions $i : U \to X$ and $j : V \to X$ induce respectively the trivial homomorphisms

$$\begin{aligned} i^* &: \pi(U, x_0) \to \pi(X, x_0), \\ j^* &: \pi(V, x_0) \to \pi(X, x_0), \end{aligned}$$

then $\pi(X, x_0)$ is trivial.

Proof: suppose $f : I \to X$ is a loop based at x_0. The sets $f^{-1}(U)$ and $f^{-1}(V)$ form an open covering of the compact metric space I, so the covering has a Lebesgue number. It is then possible to subdivide I into n intervals of the form $I_i = [a_{i-1}, a_i]$ such that $f(I)$ lies entirely in U or entirely in V for $1 \leq i \leq n$.

Should the image of consecutive intervals I_i and I_{i+1} lie in the same set U or V, amalgamate them to form a single interval. After having done this whenever it was possible to do so, we get a new collection of intervals with images lying entirely either in U or in V, and such that the images of their endpoints lie in $U \cap V$ for all such endpoints. Rename these intervals $I_i = [a_{i-1}, a_i]$ for $1 \leq i \leq m$. Then $f(I_i) \subseteq U$ or $f(I_i) \subseteq V$ and $f(a_i) \in U \cap V$ for $1 \leq i \leq m$.

Let f_i be the image of I_i under f. Then f is a path in U or in V from $f(a_{i-1})$ to $f(a_i)$. Let g_{i-1} be a path in $U \cap V$ from x_0 to $f(a_{i-1})$ and g_i be a path in $U \cap V$ from x_0 to $f(a_i)$. As $U \cap V$ is path connected, the paths g_{i-1} and g_i exist. For consistency, define g_0 and g_m to be the constant paths x_0.

If f_i is a path in V, set $f_i' = (g_{i-1}f_i)\overline{g_i}$. Then f_i' is a loop in V based at x_0.

By hypothesis, $j^*([f_i]) = [\varepsilon]$ in X, so $(g_{i-1}f_i)\overline{g_i} \sim e_{x_0}$ and $f_i \sim \overline{g}_{i-1}g_i$. Define

$$h_i = \begin{cases} \overline{g}_{i-1}g_i & \text{when } f_i \text{ lies in } V, \\ f_i & \text{when } f_i \text{ lies in } U. \end{cases}$$

Then $f_i \sim h_i$ for all i. By the preceding lemma, $f \sim h_1(h_2(\cdots h_m)\cdots)$, which is a loop in U. But loops in U are homotopic to the constant loop e_{x_0} in X, so $f \sim e_{x_0}$ in X and $\pi(X, x_0)$ is trivial as f was arbitrary. ■

We have an easy corollary.

Corollary 284 *If $X = U \cup V$, where U and V are open and simply connected and $U \cap V$ is path-connected, then X is simply connected.*

Using the special Seifert-Van Kampen theorem, we can easily compute the fundamental group of S^n, for $n \geq 2$.

Example: if $n \geq 2$, $\pi(S^n) \simeq \{\varepsilon\}$. Indeed, consider S^n as the unit sphere in $\mathbb{R}^{n+1}$, and let N and S be the north and south pole of S^n, respectively. Let $U = S^n \setminus \{N\}$ and $V = S^n \setminus \{S\}$.

Then U and V are both homeomorphic to $\mathbb{R}^n$ under **stereographic projection**, so U and V are simply connected as $\mathbb{R}^n$ is simply connected for $n \geq 2$. Clearly $S^n = U \cup V$, where U and V are open. But $U \cap V$ is path connected, as it is homeomorphic to $S^{n-1} \times (-1, 1)$, which is path-connected when $n \geq 2$. By the preceding corollary, S^n is simply connected for $n \geq 2$. □

As $\mathbb{R}^n \setminus \{0\}$ and S^n have the same fundamental group (the proof is similar to that of Theorem 277), then $\pi(\mathbb{R}^{n+1} \setminus \{0\})$ is trivial for $n \geq 2$.

Corollary 285
$\mathbb{R}^2$ is not homeomorphic to $\mathbb{R}^n$ when $n \geq 3$.

20.5 Solved Problems

1. Suppose that $f, g : I \to X$ are paths in a space X such that $f(t) = g(t)$ for $t \in [a, 1]$. If the paths $f_a, g_a : I \to X$ defined by $f_a(t) = f(at)$ and $g_a(t) = g(at)$ are path homotopic, show that f and g are path homotopic.

Proof: first, note that if $a = 0$, the result is trivially true. So suppose $a \neq 0$. Let $x_0 = f(0)$, $x_a = f(a)$ and $x_1 = f(1)$. If f_a and g_a are path homotopic, they both start at $f_a(0) = f(0) = x_0$, and they both end at $f_a(1) = f(a) = x_a$. Then, there is a continuous function $H_1 : I \times I \to X$ such that

$$\begin{aligned} H_1(t,0) &= f_a(t) = f(at) \\ H_1(t,1) &= g_a(t) = g(at) \\ H_1(0,s) &= x_0 \\ H_1(1,s) &= x_a. \end{aligned}$$

Let $H_2 : I \times I \to X$ be defined by $H_2(t,s) = f(a + t(1-a))$. Then H_2 is continuous since f is a path, and

$$\begin{aligned} H_2(t,0) &= f(a + t(1-a)) \\ H_2(t,1) &= g(a + t(1-a)) \\ H_2(0,s) &= x_a \\ H_2(1,s) &= x_1. \end{aligned}$$

This makes H_2 into a path homotopy between f and g from x_a to x_1. Now define the map $H : I \times I \to X$ by

$$H(t,s) = \begin{cases} H_1\left(\frac{t}{a}, s\right) & \text{for } t \in [0,a], \\ H_2\left(\frac{t-a}{1-a}, s\right) & \text{for } t \in [a,1]. \end{cases}$$

Then H is continuous by the pasting lemma, as H_1 and H_2 are continuous and at $t = 1$, $H_1(1,s) = H_2(0,s) = x_a$. Furthermore

$$\begin{aligned} H(t,0) &= \begin{cases} H_1\left(\frac{t}{a}, 0\right) & \text{for } t \in [0,a], \\ H_2\left(\frac{t-a}{1-a}, 0\right) & \text{for } t \in [a,1] \end{cases} \\ &= \begin{cases} f_a(t/a) & \text{for } t \in [0,a], \\ f(a + \frac{t-a}{1-a}(1-a)) & \text{for } t \in [a,1] \end{cases} \\ &= \begin{cases} f(t) & \text{for } t \in [0,a], \\ f(t) & \text{for } t \in [a,1] \end{cases} \\ &= f(t), \\ H(t,1) &= \begin{cases} H_1\left(\frac{t}{a}, 1\right) & \text{for } t \in [0,a], \\ H_2\left(\frac{t-a}{1-a}, 1\right) & \text{for } t \in [a,1] \end{cases} \\ &= \begin{cases} g_a(t/a) & \text{for } t \in [0,a], \\ g(a + \frac{t-a}{1-a}(1-a)) & \text{for } t \in [a,1] \end{cases} \\ &= \begin{cases} g(t) & \text{for } t \in [0,a], \\ g(t) & \text{for } t \in [a,1] \end{cases} \\ &= g(t), \\ H(0,s) &= H_1(0,s) = x_0, \\ H(1,s) &= H_2(1,s) = x_1. \end{aligned}$$

Hence H is a path homotopy from f to g between x_0 and x_1. ■

2. Let x_0 and x_1 be two given points of the path-connected space X. Show that $\pi_1(X, x_0)$ is abelian if and only if for every pair α and β of paths from x_0 to x_1, the induced isomorphisms $\widehat{\alpha}$ and $\widehat{\beta}$ are equal.

   Proof: suppose $\pi_1(X, x_0)$ is abelian, and let α and β be two paths from x_0 to x_1. Then $\beta\overline{\alpha}$ is a loop at x_0, so $[\beta\overline{\alpha}] \in \pi_1(X, x_0)$ and

$$[\beta\overline{\alpha}][f] = [f][\beta\overline{\alpha}]$$

   for all $[f] \in \pi_1(X, x_0)$. Then

$$\begin{aligned} [f] &= [\alpha\overline{\beta}][f][\beta\overline{\alpha}] \\ &= [\alpha]\widehat{\beta}([f])[\overline{\alpha}] \\ &= \widehat{\overline{\alpha}}(\widehat{\beta}([f])). \end{aligned}$$

   Hence $\widehat{\alpha}([f]) = \widehat{\beta}([f])$ for all $[f] \in \pi_1(X, x_0)$, so $\widehat{\alpha} = \widehat{\beta}$.

   Conversely, suppose the induced isomorphisms of any two paths in X from x_0 to x_1 are equal. Let α be such a path, and let f be a loop at x_0. Then $f\alpha$ is a path from x_0 to x_1 and $\widehat{\alpha} = \widehat{f\alpha}$. Let $[g] \in \pi_1(X, x_0)$. Then

$$[\overline{\alpha}][g][\alpha] = \widehat{\alpha}([g]) = \widehat{f\alpha}([g]) = [\overline{f\alpha}][g][f\alpha] = [\overline{\alpha}][\overline{f}][g][f][\alpha],$$

   thus $[g] = [\overline{f}][g][f]$ for all loops f and g at x_0, and $\pi_1(X, x_0)$ is abelian. ■

3. Suppose that $\widetilde{X}$ is a two-sheeted covering space of X, that is for each $x \in X$, there are two values $\widetilde{x}_1$ and $\widetilde{x}_2$ with $p^{-1}(x) = \{\widetilde{x}_1, \widetilde{x}_2\}$. Prove that the map $\phi : \widetilde{X} \to \widetilde{X}$, which interchanges the values $\widetilde{x}_1$ and $\widetilde{x}_2$ is a homeomorphism.

   Proof: the map ϕ is clearly a bijection and $\phi^2 = \mathrm{id}$, so ϕ is its own inverse. Furthermore, $\phi(z) \neq z$ for all $z \in \widetilde{X}$. To show ϕ is a homeomorphism, it is then sufficient to show that ϕ is a continuous map. To do so, we find a collection $\{Z_\alpha\}$ of open sets in $\widetilde{X}$ such that $\bigcup_{\alpha \in A} Z_\alpha = \widetilde{X}$ and such that $\phi|_{Z_\alpha} \colon Z_\alpha \to \widetilde{X}$ is continuous for all $\alpha \in A$. Then ϕ will be a continuous map.

   First note that

$$\begin{tikzcd} \widetilde{X} \arrow[r, "p"] & X \\ \widetilde{X} \arrow[u, "\phi"] \arrow[ur, "p"'] & \end{tikzcd}$$

   is a commutative diagram, since for every $x \in X$, there exists $\tilde{x} \in \widetilde{X}$ such that $p^{-1}(x) = \{\tilde{x}, \phi(\tilde{x})\}$. Thus $p\phi = p$. As $(\widetilde{X}, p)$ is a two-sheeted covering of X there exists, for every $x \in X$, a neighbourhood V_x of x in X and two disjoint open sets U_x and W_x in $\widetilde{X}$ such that $p^{-1}(V_x) = U_x \cup W_x$ and such that the mappings

$$p|_{U_x} \colon U_x \to V_x \quad \text{and} \quad p|_{W_x} \colon W_x \to V_x$$

are homeomorphisms.[1] Then U_x is homeomorphic to W_x. We show that $\phi(U_x) = W_x$, and so that $\phi(U_x)$ is homeomorphic to U_x. Suppose however that $\phi(U_x) \neq W_x$, that is, suppose there is $y \in U_x$ such that $\phi(y) \notin W_x$. Then

$$p(y) = p(\phi(y)) \in V_x,$$

and so $\phi(y) \in p^{-1}(V_x) \cup U_x$, since $\phi(x) \neq W_x$. But this would mean that $p|_{U_x} \colon U_x \to V_x$ is not injective as $y \neq \phi(y)$ and $p(y) = p(\phi(y))$. Then $\phi(U_x) \subseteq W_x$, and so $\phi(U_x) = W_x$ since U_x and W_x have the same cardinality and since ϕ is a bijection. Thus $\phi|_{U_x} \colon U_x \to W_x$ is a homeomorphism and $\phi|_{U_x} \colon U_x \to \widetilde{X}$ is continuous. Similarly, $\phi(W_x) = U_x$ and $\phi|_{W_x} \colon W_x \to \widetilde{X}$ is continuous. But

$$\widetilde{X} = p^{-1}(X) = p^{-1}\left(\bigcup_{x\in X} V_x\right) = \bigcup_{x\in X} p^{-1}(V_x) = \bigcup_{x\in X}(U_x \cup W_x),$$

where U_x and W_x are open in $\widetilde{X}$. By the argument in the first paragraph, ϕ is a homeomorphism. ■

4. If $(\widetilde{X}, p)$ and $(\widetilde{Y}, q)$ are covering spaces of X and Y respectively, show that $(\widetilde{X}\times\widetilde{Y}, (p,q))$ is a covering space of $X \times Y$.

Proof: let $h = (p,q)$. We need to show that h is a continuous surjective map and that for every $(x,y) \in X \times Y$, there exists a neighbourhood V of (x,y) such that $h^{-1}(V)$ is a disjoint union of open sets in $\widetilde{X} \times \widetilde{Y}$ and that each of these open sets is homeomorphic to V *via* h.

h **is continuous** Let $U_1 \times U_2$ be a basic neighbourhood of $X \times Y$. Then

$$\begin{aligned} h^{-1}(U_1 \times U_2) &= \{(\tilde{x},\tilde{y}) \in \widetilde{X} \times \widetilde{Y} : (p(\tilde{x}), q(\tilde{y})) \in U_1 \times U_2\} \\ &= p^{-1}(U_1) \times q^{-1}(U_2). \end{aligned}$$

But p and q are continuous, so $p^{-1}(U_1) \times q^{-1}(U_2)$ is a basic neighbourhood of $\widetilde{X} \times \widetilde{Y}$, so h is continuous.

h **is surjective** Let $(x,y) \in X \times Y$. As p and q are surjective, there exist $\tilde{x} \in \widetilde{X}$ and $\tilde{y} \in \widetilde{Y}$ such that $p(\tilde{x}) = x$ and $q(\tilde{y}) = y$. Then we have $h(\tilde{x},\tilde{y}) = (x,y)$ and h is surjective.

h **is a covering map** If $(x,y) \in X \times Y$, as p and q are covering maps, there exist neighbourhoods V_x of x in X and V_y of y in Y that are evenly covered by p and q respectively. That is $p^{-1}(V_x)$ is a disjoint union of open sets $\widetilde{V}_x$ in $\widetilde{X}$, each homeomorphic to V_x *via* p, and $q^{-1}(V_y)$ is a disjoint union of open sets $\widetilde{V}_y$ in $\widetilde{Y}$, each homeomorphic to V_y *via* q. Set $V = V_x \times V_y$. Then $(x,y) \in V$ and

$$h^{-1}(V) = p^{-1}(V_x) \times q^{-1}(V_y) = \left(\bigcup \widetilde{V}_x\right) \times \left(\bigcup \widetilde{V}_y\right) = \bigcup (\widetilde{V}_x \times \widetilde{V}_y),$$

[1]Strictly speaking, $p^{-1}(V)$ should be the disjoint union of an arbitrary collection of homeomorphic open sets in $\widetilde{X}$. But there cannot be more than two of them, since this would violate the condition that $(\widetilde{X}, p)$ be a two-sheet covering of X. Similarly, there cannot be less than two of them, since p has to be a homeomorphism when restricted to $p^{-1}(V)$.

that is $h^{-1}(V)$ is a disjoint union of open sets $\widetilde{V}_x \times \widetilde{V}_y$. But

$$h(\widetilde{V}_x \times \widetilde{V}_y) = p(\widetilde{V}_x) \times q(\widetilde{V}_y) \simeq V_x \times V_y,$$

so $\widetilde{V}_x \times \widetilde{V}_y$ is homeomorphic to $V_x \times V_y$ *via* h.

Then $(\widetilde{X} \times \widetilde{Y}, h)$ is a covering space of $X \times Y$. ■

5. a) For X as in the previous problem, if (X, p') is an n-sheeted covering space of X_1, show that $(\widetilde{X}, p'p)$ is a covering space of X_1.

   b) If X is either i. Hausdorff or ii. completely regular, show that $\widetilde{X}$ has the same property.

Proof:

a) That $p'p$ is a continuous surjective mapping is clear, as it is the composition of two such mappings. It remains only to show that it is a covering map of X.

Let $x \in X_1$. We show that we can find an open neighbourhood V of x in X_1 evenly covered by p'. We then show that the disjoint open sets in X making up $(p')^{-1}(V)$, each of which is homeomorphic to V *via* p', are themselves evenly covered by p. Then there is a disjoint union of open sets in $\widetilde{X}$ making up

$$p^{-1}\left((p')^{-1}(V)\right) = (p'p)^{-1}(V),$$

each of which is homeomorphic to V *via* $p'p$. It is going to be messy, so let's get down to it methodically.

Let $x \in X_1$. Then $(p')^{-1}(x) = \{y_1, \ldots, y_n\}$ in X, as (X, p') is an n-sheeted covering of X_1. First, the *dramatis personæ*.

- V_x is a neighbourhood of x in X_1 evenly covered by p';
- $(p')^{-1}(V_x) = \bigsqcup_{j=1}^{n} W_j$, where $\sqcup$ denotes a disjoint union, W_j is open in X and homeomorphic to V_x *via* p' and $y_j \in W_j$ for all $1 \le j \le n$.
- For $1 \le i \le n$, U_i is a neighbourhood of y_i in X evenly covered by p;
- For $1 \le i \le n$, $p^{-1}(U_i) = \bigsqcup_{\alpha} Z(i)_\alpha$, where $Z(i)_\alpha$ is open in $\widetilde{X}$ and homeomorphic to U_i *via* p for all α;
- $V = (\bigcap_{i=1}^{n} p'(U_i) \cap V_x)$;
- For $1 \le j \le n$, $K_j = \left(p'|_{W_j}\right)^{-1}(V) \subseteq W_j$ and $y_j \in K_j$;
- For $1 \le j \le n$, $M_j = K_j \cap U_j$ and $y_j \in M_j$;
- For $1 \le j \le n$ and for α, $N(j)_\alpha = \left(p|_{Z(j)_\alpha}\right)^{-1}(M_j) \subseteq Z(j)_\alpha$.

Since p' is a covering map, it is an open mapping. Then V is an open subset of X contained in V_x, since it is a finite intersection of open sets in X. As

$$p'|_{W_j} : W_j \to V_x$$

is a homeomorphism, K_j is homeomorphic to V *via* p' for $1 \le j \le n$. Note that K_j is open for $1 \le j \le n$ since V is open and that the K_j are disjoint since the

W_j are disjoint. Then M_j is open, $M_j \subseteq U_j$ for $1 \leq j \leq n$. Note further that the M_j are disjoint since the K_j are disjoint. As

$$p|_{Z(j)_\alpha} \colon Z(j)_\alpha \to U_j$$

is a homeomorphism, M_j is homeomorphic to $N(j)_\alpha$ *via* p for α and $1 \leq j \leq n$. Note that $N(j)_\alpha$ is open for $1 \leq j \leq n$ and α since M_j is open for $1 \leq j \leq n$ and that the $N(j)_\alpha$ are disjoint since the $Z(j)_\alpha$ are disjoint.

Then $N(j)_\alpha$ is homeomorphic to, say, the open subset $p'(M_1) \subseteq V_x$ *via* $p'p$ for $1 \leq j \leq n$. But $p'(M_1)$ is a neighbourhood of x in X_1 so that $p'p$ evenly covers $p'(M_1)$ at x. Hence $(\widetilde{X}, p'p)$ is a covering space of X_1.

b) i. Let $\tilde{x} \neq \tilde{y} \in \widetilde{X}$ and set $x = p(\tilde{x})$, $y = p(\tilde{y})$. Suppose V_x and V_y are neighbourhoods of x and y respectively, who are evenly covered by p. Let W_x and W_y be the (open) slices of $p^{-1}(V_x)$ and $p^{-1}(V_y)$ containing $\tilde{x}$ and $\tilde{y}$ respectively.

A. If $x = y$, W_x and W_y meet $p^{-1}(x) = p^{-1}(y)$ in exactly one point respectively, namely $\tilde{x}$ and $\tilde{y}$. Hence $\tilde{y} \notin W_x$ and $\tilde{x} \notin W_y$.

B. If $x \neq y$, let U_x and U_y be the Hausdorff neighbourhoods of x and y in X. Then $U_x \cap V_x$ is a neighbourhood of x in X disjoint from the neighbourhood $U_y \cap V_y$ of y in X. Furthermore, $O_x = (p|_{W_x})^{-1}(U_x \cap V_x)$ and $O_y = (p|_{W_y})^{-1}(U_y \cap V_y)$ are open in $\widetilde{X}$, as p is a covering map. Then $\tilde{x} \in O_x$, $\tilde{y} \in O_y$ and $O_x \cap O_y = \varnothing$ as

$$U_x \cap V_x \cap U_y \cap V_y = \varnothing.$$

In both cases, $\widetilde{X}$ is Hausdorff.

ii. Suppose $\widetilde{X}$ is non-empty and completely regular. If $\widetilde{W}$ is a neighbourhood of $\tilde{x} \in \widetilde{X}$, let U be a neighbourhood of $p(\tilde{x})$ evenly covered by p such that at least one of the slices, say M, of $p^{-1}(U)$ lies in $\widetilde{W}$.

As X is completely regular, there is a neighbourhood V of $p(\tilde{x})$ such that $\overline{V} \subseteq U$. Take $Z = p^{-1}(V) \cap M$. Then Z is homeomorphic to the slice of $p^{-1}(V)$ in M. By complete regularity of X, there is a continuous function $f : X \to [0,1]$ such that $f(p(\tilde{x})) = 1$ and $f(X - V) = \{0\}$.

Define $g_1 : \overline{Z} \to [0,1]$ by $g_1 = fp$ and $g_2 : \widetilde{X} - Z \to [0,1]$ to be the constant zero-function. As $p(\widetilde{X} - Z) \subseteq X - V$, g_1 and g_2 are both constantly zero on $\overline{Z} \cap \widetilde{X} - Z$. Then, by the pasting lemma, g_1 and g_2 define a continuous function g from $\overline{Z} \cup (\widetilde{X} - Z) = \widetilde{X}$ to $[0,1]$.

By construction, $g(\tilde{x}) = f(p(\tilde{x})) = 1$ and $g(\widetilde{X} - \widetilde{W}) = \{0\}$ since $y \notin \widetilde{W}$ implies $y \notin Z$. Hence $\widetilde{X}$ is completely regular. ■

20.6 Exercises

1. Prepare a 2-page summary of this chapter, with important definitions and results.
2. Let $h : I \to I$ be a continuous such that $h(0) = 0$ and $h(1) = 1$. For any path $f : I \to X$, prove that f and fh are path homotopic.
3. For a product space $X = \prod X_\alpha$, let $f : I \to X$ be a path, and define $f_\alpha = \pi_\alpha f$ for all α. Prove that
    a) two paths f and g in X are path homotopic if and only if $f_\alpha \sim g_\alpha$ for every α;
    b) if $f(1) = g(0)$ for paths $f, g : I \to X$, and $h = fg$, then $h_\alpha = f_\alpha g_\alpha$ for every α.
4. Prove that the fundamental group of X is isomorphic to the direct product of the fundamental groups $\pi(X_\alpha, x_\alpha)$.
5. Let $T = S^1 \times S^1$ denote the torus. For $z_0 \in S^1$, show that $S^1 \times \{z_0\}$ is a retract of T but not a strong deformation retract.
6. Show that $\varphi : X \to Y$ induces a homomorphism of path classes φ^*, as in the discussion on p. 461.
7. Show that $\mathbb{R}^3$ and $\mathbb{R}^3 \setminus \{0\}$ are homeomorphic.
8. Prove Lemma 282.
9. Show that $\mathbb{R}^n \setminus \{0\}$ and S^n have the same fundamental group.

Part V

Special Topics in Analysis and Topology

Chapter 21

Borel-Lebesgue Integration

In this chapter, we present an extension of the theory of integration that overcomes some of the issues associated with Riemann integration, and show how to integrate multi-variate functions in this new framework.

One of the problems associated with Riemann integration (see Chapters 4 and 5) is that some functions that should be integrable in any reasonable theory of integration fail to be so, for a variety of reasons.

Examples

1. Consider the Dirichlet function $\chi_{\mathbb{Q}} : \mathbb{R} \to \mathbb{R}$ defined by

$$\chi_{\mathbb{Q}}(x) = \begin{cases} 0 & x \in \mathbb{R} \setminus \mathbb{Q} \\ 1 & x \in \mathbb{Q} \end{cases}$$

   We have seen in Chapter 4 that this function is not Riemann-integrable over any interval $[a, b]$, but ...it should be, right? $\mathbb{R} \backslash \mathbb{Q}$ is so much "bigger" than $\mathbb{Q}$ that the first branch should dominate and give us an integral of 0. Unfortunately, it doesn't.

2. Consider the function $f : [0, 1] \to \mathbb{R}$ defined by

$$f(x) = \begin{cases} \frac{1}{\sqrt{x}} & x \in (0, 1] \\ 1 & x = 0 \end{cases}$$

   It is not Riemann-integrable on $[0, 1]$ as it is not bounded on $[0, 1]$, but it is Riemann-integrable of $[a, 1]$ for all $1 \geq a > 0$ since it is continuous on $[a, 1]$ for all $1 \geq a > 0$.

Furthermore
$$\int_a^1 f\,\mathrm{d}x = \left[2\sqrt{x}\right]_a^1 = 2(1-\sqrt{a}).$$
As $a \to 0^+$, we see that
$$\int_a^1 f\,\mathrm{d}x \to 2(1-\sqrt{0}) = 2,$$
and we would at the very least consider an extension of Riemann integration for which $\int_0^1 f\,\mathrm{d}x = 2$.

3. The function $g : [0,\infty) \to \mathbb{R}$ defined by $g(x) = e^{-x}$ is not Riemann-integrable on $[0,\infty)$ since the domain of integration cannot even be partitioned. But it is clearly Riemann-integrable on $[0,n]$, $n > 0$, since it is continuous on $[0,n]$; in fact,
$$\int_0^n e^{-x}\,\mathrm{d}x = [-e^{-x}]_0^n = 1 - e^{-n}.$$
Since
$$\lim_{n\to\infty} \int_0^n e^{-x}\,\mathrm{d}x = \lim_{n\to\infty}(1-e^{-n}) = 1 - 0 = 1;$$
any extension of Riemann integration should at least give us $\int_0^\infty g = 1$. $\square$

In this chapter, we will introduce an **extension** of the Riemann integral in which all of these examples will work out as we think they should. The **Lebesgue-Borel** approach to integration views the problem from a different example:[1] fundamentally, instead of building **vertical** boxes under the graph of f, we stack **horizontal** boxes under it. This conceptual shift has far-ranging consequences.[2]

We will also extend our definition of the integral to **multivariate domains** (which is to say, the functions we consider will be functions of $\mathbb{R}^n$ to $\mathbb{R}$). To help illustrate the concepts, we will often work with functions $f : \mathcal{A} \subseteq \mathbb{R}^2 \to \mathbb{R}_+$, where f is **bounded** (as a function), as is A (as a set). By analogy to the 1-dimensional case, we will want to define
$$I = \iint_A f(x,y)\,\mathrm{d}x\,\mathrm{d}y$$
so that
$$I = \mathrm{Vol}\Big(\{(x,y,t) \mid (x,y) \in A,\ 0 \le t \le f(x,y)\}\Big).$$

[1] There are other approaches: **improper Riemann integration** and **generalized Riemann integration**, say, but we will not be touching on those.

[2] It does not resolve all difficulties, however: there are differentiable functions $F : [a,b] \to \mathbb{R}$ for which F' is not Lebesgue-integrable and some important improper integrals do not exist, for instance.

21.1 Borel Sets and Borel Functions

Generally speaking, the **Borel subsets** of $\mathbb{R}^n$ are the $\sigma-$algebra of subsets for which we know how to compute the **length**, and/or the **surface area**, and/or the **volume**, and so on.[3]

Formally, a $\sigma-$**algebra** $\mathfrak{S}$ of $\mathbb{R}^n$ is a collection of subsets of $\mathbb{R}^n$ such that

1. $A_1, A_2, \dots, A_n, \dots \in \mathfrak{S} \implies \bigcup_{n\geq 1} A_n \in \mathfrak{S}$, and
2. $A \in \mathfrak{S} \implies A^c = \mathbb{R}^n \backslash A \in \mathfrak{S}$.

Consequently (see exercises),

1. $A_1, A_2, \dots, A_n, \dots \in \mathfrak{S} \implies \bigcap_{n\geq 1} A_n \in \mathfrak{S}$;
2. $A, B \in \mathfrak{S} \implies A \cap B^c \in \mathfrak{S}$, and
3. $\varnothing, \mathbb{R}^n \in \mathfrak{S}$.

Examples

1. The **power set** $\wp(\mathbb{R}^n)$ is the **largest** $\sigma-$algebra of $\mathbb{R}^n$, since the union of any collection subsets of $\mathbb{R}^n$ is itself a subset of $\mathbb{R}^n$, and since the complement of any subset of $\mathbb{R}^n$ is also a subset of $\mathbb{R}^n$.
2. The **standard topology** $\tau = \{U \subseteq \mathbb{R}^n \mid U \subseteq_O \mathbb{R}^n\}$ is **not** a $\sigma-$algebra of $\mathbb{R}^n$ since the complement of the open ball of radius 1 centered at the origin, say, is not open in $\mathbb{R}^n$ (see Part IV).
3. $\mathfrak{S}_0(\mathbb{R}^n) = \{\mathbb{R}^n, \varnothing\}$ is the smallest $\sigma-$algebra of $\mathbb{R}^n$. □

Note that $\mathfrak{S}$ of $\mathbb{R}^n$ is a subset of $\wp(\mathbb{R}^n)$.

Proposition 286
If $(\mathfrak{S}_i)_{i\geq 1}$ is a collection of $\sigma-$algebras of $\mathbb{R}^n$ then $\mathfrak{S} = \bigcap_{i\geq 1} \mathfrak{S}_i$ is a $\sigma-$algebra of $\mathbb{R}^n$.

Proof:

1. Suppose $A_1, \dots, A_n \dots \in \mathfrak{S}$.Then, $A_1, \dots, A_n, \dots \in \mathfrak{S}_i\ \forall i$. But, $\mathfrak{S}_i$ is a $\sigma-$algebra for all i so that $\bigcup_{n\geq 1} A_n \in \mathfrak{S}_i\ \forall i$. Then, $\bigcup_{n\geq 1} A_n \in \bigcap_{i\geq 1} \mathfrak{S}_i = \mathfrak{S}$.
2. Suppose $A \in \mathfrak{S}$ then we have that $A \in \mathfrak{S}_i\ \forall\, i$. But $\mathfrak{S}_i$ is a $\sigma-$ algebra so that $A^c \in \mathfrak{S}_i\ \forall\, i \implies A^c \in \bigcap_{i\geq 1} \mathfrak{S}_i = \mathfrak{S}$. ■

[3] We only present a restricted version of the Borel-Lebesgue theory of integration; the full version is built on **measurable subsets** of $\mathbb{R}^n$, where the **measure** generalizes the notions of of length, surface area, volume, etc. to "not-as-nice" geometric subsets of $\mathbb{R}^n$ (a feature of the theory is that not every subset of $\mathbb{R}^n$ is measurable).

The standard topology is not a $\sigma-$algebra of $\mathbb{R}^n$, but since $\tau \in \wp(\mathbb{R}^n)$, there is at least one σ-algebra containing the open sets of $\mathbb{R}^n$. The **Borel $\sigma-$algebra of $\mathbb{R}^n$** is the intersection of all $\sigma-$algebras containing the open sets of $\mathbb{R}^n$, we denote it by:

$$\mathcal{B} = \mathcal{B}(\mathbb{R}^n) = \bigcap_{\tau \subseteq \mathfrak{S} \in \wp(\mathbb{R}^n)} \mathfrak{S}.$$

An element of $\mathcal{B}$ is called a **Borel set of** $\mathbb{R}^n$.

Just about every subset of $\mathbb{R}^n$ that we encounter **in practice** is a Borel set:

- every open subset of $\mathbb{R}^n$ is a Borel set of $\mathbb{R}^n$;
- every closed subset of $\mathbb{R}^n$ is a Borel set of $\mathbb{R}^n$;
- any set built *via* unions, intersections, and complements with open sets and/or closed sets is a Borel set of $\mathbb{R}^n$.

Theorem 287
Let $\mathcal{B} = \mathcal{B}(\mathbb{R}^2)$. There exists a unique function Area $: \mathcal{B} \to [0, \infty]$ such that:

1. *Area$(A) \geq 0,\ \forall A \in \mathcal{B}$*
2. *if $A_1, \ldots, A_n, \ldots \in \mathcal{B}$ are pairwise disjoint then:*

$$\textit{Area}\left(\bigcup_{n\geq 1} A_n\right) = \sum_{n\geq 1} \textit{Area}(A_n)$$

3. *Area$([a, a'] \times [b, b']) = (a' - a)(b' - b)$.*

The area function whose existence is guaranteed by theorem 287 corresponds to our intuition of area in $\mathbb{R}^2$, but such a function cannot be defined on the entirety of $\wp(\mathbb{R}^2)$ (see the Banach-Tarski paradox).[4]

Theorem 288
Let $A, B \in \mathcal{B}(\mathbb{R}^2)$ such that $A \subseteq B$, then Area$(A) \leq$ Area(B).

Proof: by definition $B = (A \cap B) \cup (A^c \cap B) = A \cup (B \backslash A^c)$ where $B \backslash A^c \in \mathcal{B}(\mathbb{R}^n)$. Hence, we have

$$\text{Area}(B) = \text{Area}(A) + \text{Area}(B \backslash A^c) \geq \text{Area}(A),$$

which completes the proof. ■

We can extend Theorem 287.2 to not necessarily pairwise disjoint Borel sets.

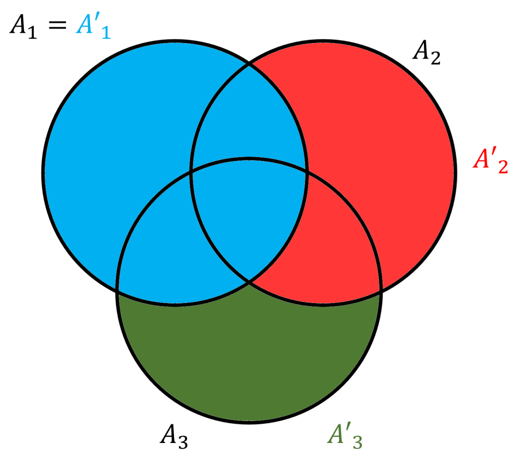
[4]Proving the existence of the function and of a set whose area cannot be measured is rather difficult and is properly tackled in advanced measure theory courses.

Theorem 289
Let $A_1, A_2, \ldots, A_n, \ldots \in \mathcal{B}(\mathbb{R}^2)$. *Then Area*$(\bigcup_{n\geq 1} A_n) \leq \sum_{n\geq 1}$ *Area*(A_n).

Proof: construct the sequence $A_n' \in \mathcal{B}(\mathbb{R}^2)$ as follows:

1. $A_1' = A_1$;
2. $A_2' = A_2 \cap A_1^c$;
3. $A_3' = A_3 \cap (A_1 \cup A_2)^c$, etc.

The process is illustrated below on A_1, A_2, A_3.



Then $A_1', \ldots, A_n', \ldots \in \mathcal{B}(\mathbb{R}^2)$ are pairwise disjoint and

$$A_1 \cup A_2 \cup \ldots \cup A_n = A_1' \cup A_2' \cup \ldots \cup A_n'$$

for all $n \geq 1$. Since $A_n' \subseteq A_n \, \forall n \geq 1$. Then

$$\text{Area}\left(\bigcup_{n\geq 1} A_n\right) = \text{Area}\left(\bigcup_{n\geq 1} A_n'\right) = \sum_{n\geq 1} \text{Area}(A_n') \leq \sum_{n\geq 1} \text{Area}(A_n),$$

which completes the proof. ■

We say that $B \subseteq \mathbb{R}^2$ has a $(2D)$ **measure** 0 if $\forall \varepsilon > 0$, there is a **cover**

$$\{R_1, R_2, \ldots, R_n, \ldots\}$$

of B by rectangles $R_n = [a_n, a_n'] \times [b_n, b_n']$ with $a_n < a_n'$ and $b_n < b_n'$ for all $n \geq 1$, such that

$$\sum_{n\geq 1} \text{Area}(R_n) < \varepsilon.$$

Examples

1. Show that $B = \mathbb{R} \times \{b\}$ has a $2D$ measure 0 for any choice of $b \in \mathbb{R}$.

   Proof: let $\varepsilon > 0$ and set
   $$R_n = [-n, n] \times \left[b - \frac{\varepsilon}{2n2^{n+2}}, b + \frac{\varepsilon}{2n2^{n+2}}\right].$$
   Then Area$(R_n) = 2n \cdot \frac{\varepsilon}{n2^{n+2}} = \frac{\varepsilon}{2^{n+1}}$ for all $n \in \mathbb{N}$, and $B \subseteq \bigcup_{n\geq 1} R_n$, so that
   $$0 \leq \text{Area}(B) \leq \sum_{n\geq 1}(R_n) = \varepsilon \sum_{n\geq 1} \frac{1}{2^{n+1}} < \varepsilon.$$
   As $\varepsilon > 0$ is arbitrary, Area$(B) = 0$. ■

2. $S^1 = \{(x, y) \mid x^2 + y^2 = 1\}$ has $2D$ measure 0.

3. Show that Area$((a, a') \times (b, b'))$ = Area$([a, a'] \times [b, b'])$.

   Proof: write
   $$[a, a'] \times [b, b'] = (a, a') \times (b, b') \sqcup \{a\} \times [b, b'] \sqcup \{a'\} \times [b, b'] \sqcup [a, a'] \times \{b\} \sqcup [a, a'] \times \{b'\}.$$
   Each of the components $[*, *] \times \{*\}$ are subsets of $\mathbb{R} \times \{*\}$, so that they have 2D area 0 (and similarly for the components $\{*\} \times [*, *]$). Thus
   $$\text{Area}([a, a'] \times [b, b']) \leq \text{Area}((a, a') \times (b, b')) + 0 + 0 + 0 + 0 = \text{Area}((a, a') \times (b, b')).$$
   But Area$((a, a') \times (b, b')) \leq$ Area$([a, a'] \times [b, b'])$ since $(a, a') \times (b, b') \subseteq [a, a'] \times [b, b']$, so Area$((a, a') \times (b, b'))$ = Area$([a, a'] \times [b, b'])$. ■

4. Show that every finite subset $B \subseteq \mathbb{R}^2$ has $2D$ measure 0.

   Proof: let $B = \{(x_1, y_1), \ldots (x_n, y_n)\}$ and $\varepsilon > 0$. Pick:
   - a closed rectangle R_1 with Area$(R_1) = \frac{\varepsilon}{2}$ and $(x_1, y_1) \in R_1$;
   - a closed rectangle R_2 with Area$(R_2) = \frac{\varepsilon}{2^2}$ and $(x_2, y_2) \in R_2$;

     $\ldots$
   - a closed rectangle R_n with Area$(R_n) = \frac{\varepsilon}{2^n}$ and $(x_n, y_n) \in R_n$;
   - for $m > n$, any closed rectangle with Area$(R_m) = \frac{\varepsilon}{2^{m+1}}$ will do.

   Then $B \subseteq \bigcup_{m\geq 1} R_m$ and
   $$\sum_{m\geq 1} \text{Area}(R_m) = \sum_{m\geq 1} \frac{\varepsilon}{2^{m+1}} < \varepsilon,$$
   which completes the proof. ■

5. Every countable subset of $\mathbb{R}^2$ has $2D$ measure 0.
6. Let $\varphi : [0,1] \to \mathbb{R}^2$ be continuous and such that there exists $M > 0$ with

$$\|\varphi(s) - \varphi(t)\|_\infty \leq M|s-t| \quad \forall s,t \in [0,1].$$

Then $\varphi([0,1])$ has $2D$ measure 0.

Proof: recall that $\|(x_1,x_2)\|_\infty = \max\{|x_1|,|x_2|\}$. For all $N \geq 1$, let

$$0 = t_0 < t_1 < \cdots < t_N = 1,$$

with $t_i = \frac{i}{N}$. Let $s_i, s_i' \in [t_{i-1}, t_i]$. By hypothesis,

$$\|\varphi(s_i) - \varphi(s_i')\|_\infty \leq M|s_i - s_i'| \leq M|t_{i-1} - t_i| \leq M\left|\frac{i-1}{N} - \frac{i}{N}\right| \leq \frac{M}{N}.$$

Thus, there exists a square $I_i \subseteq \mathbb{R}^2$ whose sides have length $\frac{2M}{N}$ such that $\varphi([t_{i-1},t_i]) \subseteq I_i$. By construction, for all $1 \leq i \leq N$ we have

$$\text{Area}(I_i) = \frac{4M^2}{N^2} \quad \text{and} \quad \sum_{i=1}^{N} \text{Area}(I_i) = \frac{4M^2}{N}.$$

Let $\varepsilon > 0$ and select $N > \frac{4M^2}{\varepsilon}$. Going through the above procedure yields a sequence of rectangles $R_i = I_i$ for $1 \leq i \leq N$; for $n > N$, set $R_n = \{*\} \subseteq \mathbb{R}^2$, a singleton square of area 0. Then

$$\varphi([0,1]) \subseteq \bigcup_{i=1} R_i \implies \sum_{i\geq 1} \text{Area}(R_i) = \frac{4M^2}{N} < \varepsilon,$$

which completes the proof. ■

In the rest of this section, we introduce the class of functions $f : \mathbb{R}^2 \to \overline{\mathbb{R}}$ for which we may expect that

$$\iint_{\mathbb{R}^2} f(x,y)\, \mathrm{d}x\, \mathrm{d}y \in \overline{\mathbb{R}}$$

exists.[5] As we see below, we cannot untangle the function rule from its domain. If $A \in \mathcal{B}(\mathbb{R}^2)$, let the **characteristic function** $\chi_A : \mathbb{R}^2 \to \mathbb{R}$ be defined by

$$\chi_A(x,y) = \begin{cases} 0 & \text{if } (x,y) \notin A \\ 1 & \text{if } (x,y) \in A \end{cases}$$

[5]The set $\overline{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$ is the one-point compactification of $\mathbb{R}$ (see Section 17.4).

Characteristic functions are the building blocks of **Borel-Lebesgue integrable functions**; their integral is easy to obtain. Let $k \in \mathbb{R}$; if $f : \mathbb{R}^2 \to \overline{\mathbb{R}}$ is defined by $f(x,y) = k \cdot \chi_A(x,y)$, then the **Borel-Lebesgue integral of f over $\mathbb{R}^2$ is**

$$\iint_{\mathbb{R}^2} f \, dx \, dy = k \cdot \text{Area}(A) \in \overline{\mathbb{R}}.$$

A function $f : \mathbb{R}^2 \to \mathbb{R}$ is **simple** if $\exists A_1, \ldots, A_n \in \mathcal{B}(\mathbb{R}^2)$ and $a_1, \ldots, a_n \in \mathbb{R}$ such that

$$\mathbb{R}^2 = A_1 \sqcup A_2 \sqcup \cdots \sqcup A_n \quad \text{and} \quad f|_{A_i} \equiv a_i;$$

in that case, $f = \sum_{i=1}^n a_i \chi_{A_i}$.

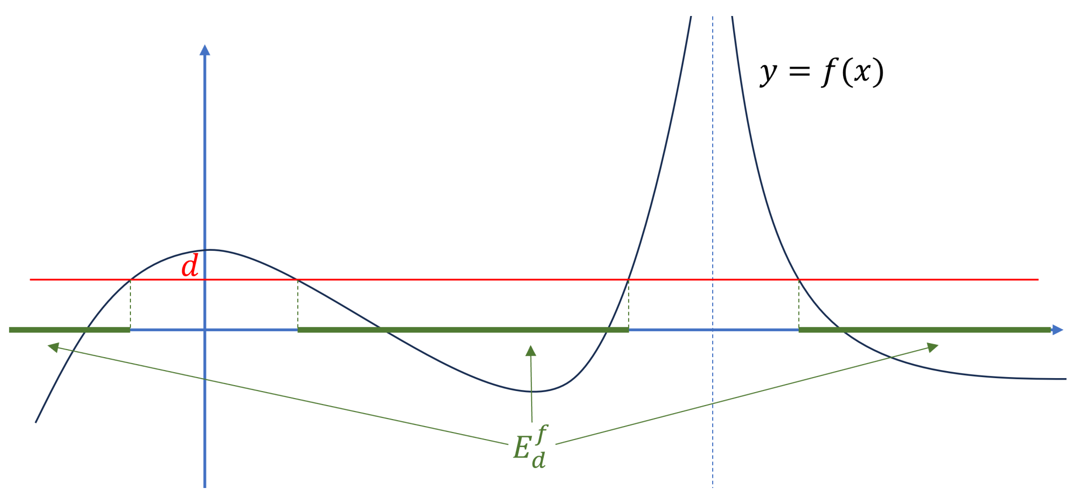
Examples (SIMPLE FUNCTIONS)

1. If $f(x,y) = k$ for all $(x,y) \in \mathbb{R}^2$, then f is a simple function.
2. If $f = \sum_{i=1}^n a_i \chi_{A_i}$, then $|f| = \sum_{i=1}^n |a_i| \chi_{A_i}$ is a simple function.
3. If $f = \sum_{i=1}^n a_i \chi_{A_i}$ and $g = \sum_{j=1}^m b_j \chi_{B_j}$ are simple functions, then
   a) $\mathbb{R} = \bigsqcup_{i=1}^n \bigsqcup_{j=1}^m A_i \cap B_j$;
   b) $f + g = \sum_{i=1}^n \sum_{j=1}^m (a_i + b_j) \chi_{A_i \cap B_j}$ is a simple function, and
   c) $fg = \sum_{i=1}^n \sum_{j=1}^m a_i \, b_j \, \chi_{A_i \cap B_j}$ is a simple function. □

A **Borel function** is a function $f : \mathbb{R}^2 \to \overline{\mathbb{R}}$ for which

$$E_d^f = \{(x,y) \mid f(x,y) \leq d\} \in \mathcal{B}(\mathbb{R}^2), \quad \forall d \in \mathbb{R}.$$

We illustrate the concept below, for a function over $\mathbb{R}$.



Since every subset of $\mathbb{R}^2$ we encounter in practice is a Borel set, every function $f : \mathbb{R}^2 \to \overline{\mathbb{R}}$ we encounter in practice is a Borel function.[6]

[6]It is in fact rather difficult to construct a non-Borel function, although they do exist.

Proposition 289
Let $f, g : \mathbb{R}^2 \to \overline{\mathbb{R}}$ be Borel functions. Then, $|f|, f+g, fg$ are also Borel functions.

Proof: we prove the result only for $|f|$; the proof for the other two functions is left as an exercise. Write $z = (x, y) \in \mathbb{R}^2$. Then, we want to show

$$E_d^{|f|} = \{z \in \mathbb{R}^2 \mid |f(z)| \leq d\} \in \mathcal{B}(\mathbb{R}^2 \quad \forall d \in \mathbb{R}$$

1. if $d < 0$, then $E_d^{|f|} = \varnothing \in \mathcal{B}(\mathbb{R}^2)$;
2. if $d \geq 0$, then

$$\begin{aligned} E_d^{|f|} &= \{z \mid -d \leq f(z) \leq d\} = \{z \mid -d \leq f(z)\} \cap \{z \mid f(z) \leq d\} \\ &= \{z \mid -d \leq f(z)\} \cap E_d^f = \{z \mid f(z) < -d\}^c \cap E_d^f \\ &= \left(\bigcup_{n \geq 1} E^f_{-d-\frac{1}{n}} \right)^c \cap E_d^f. \end{aligned}$$

But f is a Borel function, so $E_d^f, E^f_{-d-\frac{1}{n}} \in \mathcal{B}(\mathbb{R}^2)$ for all $n \geq 1$. This implies that

$$\bigcup_{n \geq 1} E^f_{-d-\frac{1}{n}} \in \mathcal{B}(\mathbb{R}^2),$$

as $\mathcal{B}(\mathbb{R}^2)$ is a $\sigma-$algebra, and so that

$$\mathbb{R}^2 \setminus \left(\bigcup_{n \geq 1} E^f_{-d-\frac{1}{n}} \right) \in \mathcal{B}(\mathbb{R}^2),$$

for the same reason; hence $E_d^{|f|} \in \mathcal{B}(\mathbb{R}^2)$. ■

We can approximate positive-valued Borel functions with simple functions.

Theorem 290
Let $f : \mathbb{R}^2 \to [0, \infty]$ be a Borel function; then there is a sequence (f_n) of simple functions such that:

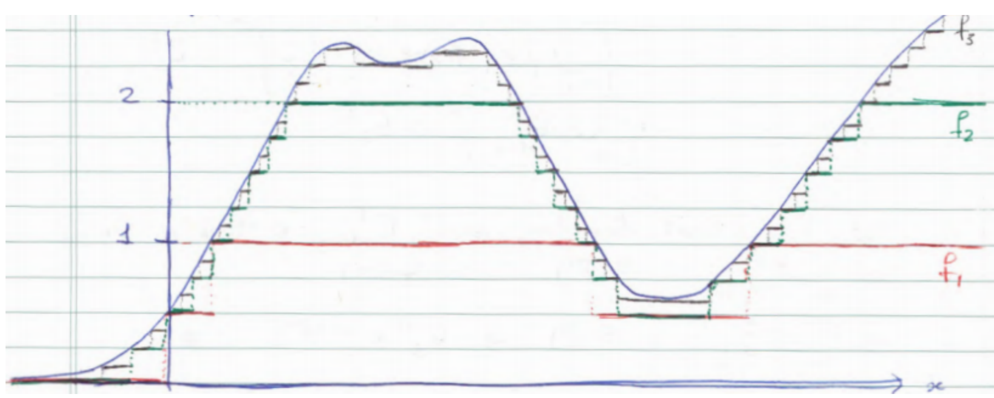
1. *$\forall z \in \mathbb{R}^2$, $f_n(z) \to f(z)$, and*
2. *$0 \leq f_n \leq f$, for all $n \geq 1$.*

Proof: we provide a proof for $f : \mathbb{R} \to [0,\infty]$; the proof for functions on $\mathbb{R}^k$ is identical, but the simpler case is easier to illustrate.

We build the sequence (f_n) as follows.

1. For f_1, write
$$\mathbb{R} = \underbrace{\left\{x \,\middle|\, 0 \le f(x) < \frac{1}{2^1}\right\}}_{A_1^1} \sqcup \underbrace{\left\{x \,\middle|\, \frac{1}{2} \le f(x) < 1\right\}}_{A_2^1} \sqcup \underbrace{\{x \mid f(x) \ge 1\}}_{A^1},$$
and set
$$f_1 = 0 \cdot \chi_{A_1^1} + \frac{1}{2}\chi_{A_2^1} + 1 \cdot \chi_{A^1}, \quad A_1^1, A_2^1, A^1 \in \mathcal{B}(\mathbb{R}^2).$$

2. For f_2, write
$$\mathbb{R} = \left(\bigsqcup_{i=1}^{8} \underbrace{\left\{x \,\middle|\, \frac{i-1}{2^2} \le f(x) < \frac{i}{2^2}\right\}}_{A_i^2}\right) \sqcup \underbrace{\{x \mid f(x) \ge 2\}}_{A^2} = \left(\bigsqcup_{i=1}^{8} A_i^2\right) \cup A^2,$$
and set
$$f_2 = \sum_{i=1}^{8} \frac{i-1}{2^2}\chi_{A_i^2} + 2\chi_{A^2}.$$



...

n. For f_n, write $A^n = \{x \mid f(x) \ge n\}$ and
$$A_i^n = \left\{x \,\middle|\, \frac{i-1}{2^n} \le f(x) < \frac{i}{2^n}\right\}, \quad \text{for } 1 \le i \le n \cdot 2^n.$$

We then have $\mathbb{R} = \left(\bigsqcup_{i=1}^{n\cdot 2^n} A_i^n\right) \sqcup A^n$. Set $f_n = \sum_{i=1}^{n\cdot 2^n} \frac{i-1}{2^n}\chi_{A_i^n} + n \cdot \chi_{A^n}$.

By construction, each f_n is simple and

$$0 \leq f_1(x) \leq f_2(x) \leq \cdots \leq f_n(x) \leq \cdots \leq f(x) \quad \forall x \in \mathbb{R}.$$

1. If $f(x) = \infty$, then $x \in A^n$ for all $n \geq 1$, whence $f_n(x) = n \to \infty = f(x)$
2. If $f(x) < \infty$, then for $n > f(x)$, there exists $1 \leq i \leq u \leq n \cdot 2^n$ such that

$$\frac{i-1}{2^n} \leq f(x) < \frac{i}{2^n}.$$

In that case $x \in A_i^n$ and

$$|f(x) - f_n(x)| = \left| f(x) - \frac{i-1}{2^n} \right| < \frac{1}{2^n} \to 0,$$

which completes the proof. ■

21.2 Integral of Simple Functions

Let $f = \sum_{i=1}^k \alpha_i \chi_{A_i}$ be a simple function $\mathbb{R}^2 \to [0,\infty]$, that is, $\alpha_i \in [0,\infty]$ for $1 \leq i \leq k$ and $\mathbb{R}^2 = A_1 \sqcup \cdots \sqcup A_k$. Since simple functions are finite linear combinations of characteristic functions, we define the **integral of a simple function** as

$$\iint_{\mathbb{R}^2} f(x,y)\,\mathrm{d}x\,\mathrm{d}y = \sum_{i=1}^k \alpha_i \cdot \text{Area}(A_i) \in [0,\infty]$$

(in the Borel-Lebesgue theory of integration, we have $0 \cdot (+\infty) = 0$, by convention). But there might be multiple ways to write a simple function as a sum of characteristic functions: if

$$f = \sum_{i=1}^k \alpha_i \chi_{A_i} = \sum_{j=1}^m \beta_j \chi_{B_j},$$

is the integral the same in both cases? For each $1 \leq i \leq k$, let $J_i = \{j \mid \beta_j = \alpha_i\}$. Then

$$\begin{aligned}\sum_{j=1}^m \beta_j \cdot \text{Area}(B_j) &= \sum_{i=1}^k \sum_{j \in J_i} \beta_j \cdot \text{Area}(B_j) = \sum_{i=1}^k \alpha_i \sum_{j \in J_i} \text{Area}(B_j) \\ &= \sum_{i=1}^k \alpha_i \cdot \text{Area}\left(\bigsqcup_{j \in J_i} B_j \right) = \sum_{i=1}^k \alpha_i \cdot \text{Area}(A_i).\end{aligned}$$

In what follows, we denote the **set of simple functions on** $\mathbb{R}^n$ by $\zeta^{(n)}$ and the **set of positive simple functions on** $\mathbb{R}^n$ by $\zeta_+^{(n)}$.

Lemma 291
Let $f, g \in \zeta_+^{(2)}, \alpha \geq 0$. *Then:*

1. $\iint_{\mathbb{R}^2} \alpha f \, \mathrm{d}x \, \mathrm{d}y = \alpha \iint_{\mathbb{R}^2} f \, \mathrm{d}x \, \mathrm{d}y$;
2. $\iint_{\mathbb{R}^2} (f+g) \, \mathrm{d}x \, \mathrm{d}y = \iint_{\mathbb{R}^2} f \, \mathrm{d}x \, \mathrm{d}y + \iint_{\mathbb{R}^2} g \, \mathrm{d}x \, \mathrm{d}y$, *and*
3. if $f \leq g$ on $\mathbb{R}^2$, then $\iint_{\mathbb{R}^2} f \, \mathrm{d}x \, \mathrm{d}y \leq \iint_{\mathbb{R}^2} g \, \mathrm{d}x \, \mathrm{d}y$.

Proof: note that the results hold over general multi-dimensional spaces, but we restrict the demonstration to $\mathbb{R}^2$.

1. The first statement is clear; its proof is left as an exercise.
2. If $f = \sum_{i=1}^n \alpha_i \chi_{A_i}$ and $g = \sum_{j=1}^m \beta_j \chi_{B_j}$ then $f + g = \displaystyle\sum_{i,j} (\alpha_i + \beta_j) \chi_{A_i \cap B_j}$ and

$$\begin{aligned}
\iint_{\mathbb{R}^2} (f+g) \, \mathrm{d}x \, \mathrm{d}y &= \sum_{i,j} (\alpha_i + \beta_j) \cdot \text{Area}(A_i \cap B_j) \\
&= \sum_{i,j} \alpha_i \cdot \text{Area}(A_i \cap B_j) + \sum_{i,j} \beta_j \cdot \text{Area}(A_i \cap B_j) \\
&= \sum_{i=1}^n \alpha_i \sum_{j=1}^m \text{Area}(A_i \cap B_j) + \sum_{j=1}^m \beta_j \sum_{i=1}^n \text{Area}(A_i \cap B_j) \\
&= \sum_{i=1}^n \alpha_i \cdot \text{Area}\left[A_i \cap \left(\bigsqcup_{j=1}^m B_j \right) \right] + \sum_{j=1}^m \beta_j \cdot \text{Area}\left[B_j \cap \left(\bigsqcup_{i=1}^n A_i \right) \right] \\
&= \sum_{i=1}^n \alpha_i \cdot \text{Area}(A_i) + \sum_{j=1}^m \beta_j \cdot \text{Area}(B_j) = \iint_{\mathbb{R}^2} f \, \mathrm{d}x \, \mathrm{d}y + \iint_{\mathbb{R}^2} g \, \mathrm{d}x \, \mathrm{d}y
\end{aligned}$$

3. If $f \leq g$ on $\mathbb{R}^2$, then $g - f \in \zeta_+^{(2)}$ and

$$\iint_{\mathbb{R}^2} = \iint_{\mathbb{R}^2} [f + (g-f)] \, \mathrm{d}x \, \mathrm{d}y = \iint_{\mathbb{R}^2} f \, \mathrm{d}x \, \mathrm{d}y + \underbrace{\iint_{\mathbb{R}^2} (g-f) \, \mathrm{d}x \, \mathrm{d}y}_{\geq 0} \geq \iint_{\mathbb{R}^2} f \, \mathrm{d}x \, \mathrm{d}y,$$

since $g - f \geq 0$. ■

The first two properties of Lemma 291 indicate that the integral of a simple function behaves as a **linear operator** on the set of positive simple functions on $\mathbb{R}^n$.[7]

[7]We cannot say "over the vector space of positive simple functions" since $\zeta_+^{(n)}$ is not a vector space over $\mathbb{R}$... but $\zeta^{(n)}$ is, however.

Furthermore, if $f = \chi_A$, $A \in \mathcal{B}(\mathbb{R}^2)$, then $\iint f \, \mathrm{d}x \, \mathrm{d}y = \text{Area}(A)$.[8]

As mentioned in the proof of Lemma 291, we can generalize the notion of the integral of positive simple functions directly to higher dimensions. For instance, if $f : \mathbb{R}^3 \to \mathbb{R} \in \zeta_+^{(3)}$, then

$$\iiint_{\mathbb{R}^3} f(x,y,z) \, \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z = \iiint_{\mathbb{R}^3} \sum_{k=1}^{\ell} \gamma_k \chi_{A_k} \, \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z = \sum_{k=1}^{\ell} \gamma_k \cdot \text{Vol}(A_k),$$

and so on with $n \geq 3$:

$$\int \cdots \int f(x_1, \ldots, x_n) \, \mathrm{d}x_1 \ldots \mathrm{d}x_n$$

if $f : \mathbb{R}^n \to \mathbb{R}$ is in $\zeta_+^{(n)}$.

21.3 Integral of Positive Borel Functions

Of course, the overwhelming majority of functions on $\mathbb{R}^n$ are not simple positive functions; but large classes of non-negative functions can be approximated by simple functions (as we have seen Theorem 290). If f is a **positive Borel function** of $\mathbb{R}^2$ to $[0, \infty]$, its Borel-Lebesgue integral

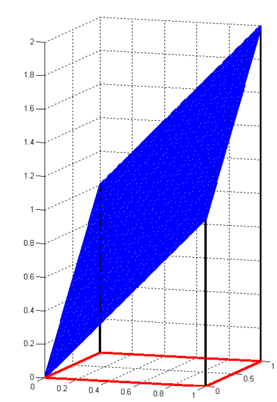
$$\iint f(x,y) \, \mathrm{d}x \, \mathrm{d}y = \sup_{s \in \zeta_+^{(2)}} \left\{ \iint s \, \mathrm{d}x \, \mathrm{d}y \;\middle|\; s \leq f \right\};$$

this definition can be extended to higher-dimensional domains in the obvious way. We illustrate how it applies in practice with a deceptively complicated example.

Example: using the definition, find $\iint f \, \mathrm{d}x \, \mathrm{d}y$, where

$$f(x,y) = \begin{cases} x + y & \text{if } (x,y) \in [0,1]^2 \\ 0 & \text{otherwise} \end{cases}$$

Solution: the function is shown below.



[8]When the context is clear, we may omit the domain of integration.

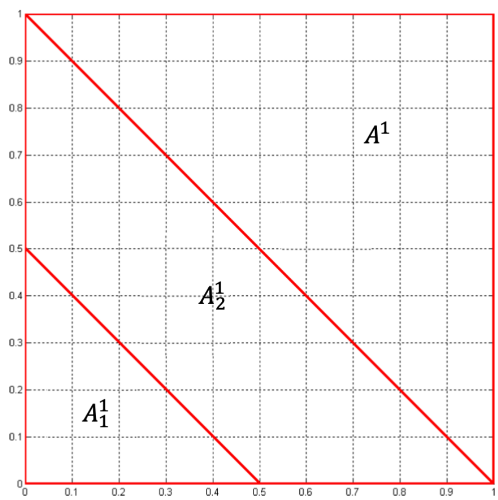
We start by building the sequence of positive simple functions

$$s_1 \leq \ldots \leq s_n \leq \ldots \leq f$$

from Theorem 290.

For $n = 1$, we have:

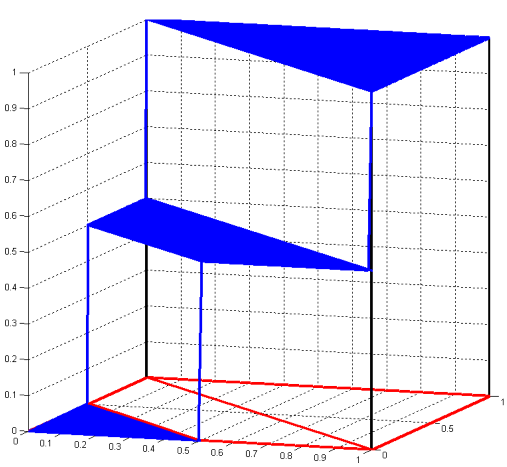
- $A_1^1 = \{(x,y) \mid 0 \leq f(x,y) < \frac{1}{2}\} = (\{(x,y) \mid 0 \leq x+y < \frac{1}{2}\} \cap [0,1]^2) \cup (\mathbb{R}^2 \setminus [0,1]^2)$,
- $A_2^1 = \{(x,y) \mid \frac{1}{2} \leq f(x,y) < 1\} = \{(x,y) \mid \frac{1}{2} \leq x+y < 1\} \cap [0,1]^2$, and
- $A^1 = \{(x,y) \mid f(x,y) \geq 1\} = \{(x,y) \mid x+y \geq 1\} \cap [0,1]^2$ (see below).



The first simple approximation is thus

$$s_1 = 0 \cdot \chi_{A^1} + \frac{1}{2} \cdot \chi_{A_2^1} + 1 \cdot \chi_{A^1},$$

whose graph is shown below:



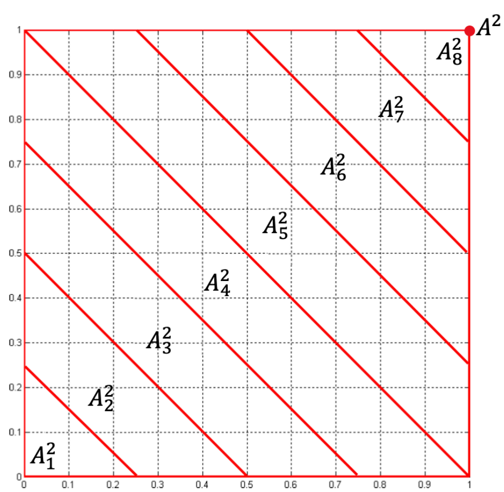
We then have

$$\iint s_1(x,y)\,\mathrm{d}x\,\mathrm{d}y = 0\cdot \text{Area}(A_1^1) + \frac{1}{2}\cdot \text{Area}(A_2^1) + 1\cdot \text{Area}(A^1),$$

whose value we leave un-evaluated.

For $n=2$, we have

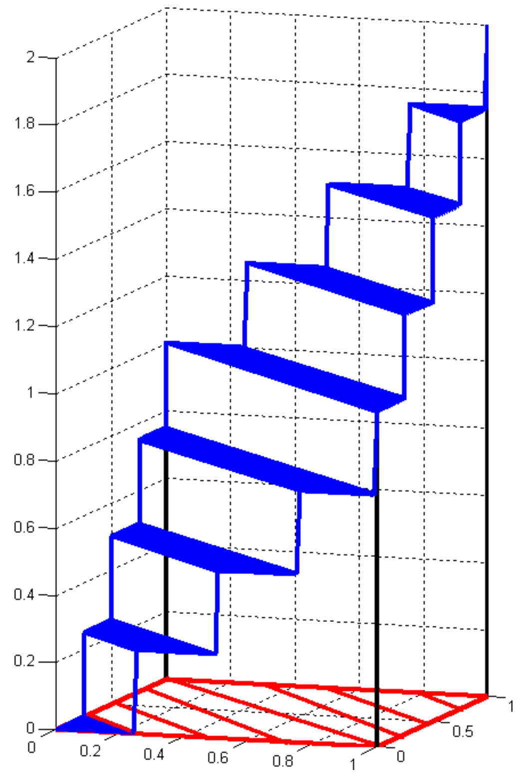
- $A_1^2 = \{(x,y) \mid 0 \leq f(x,y) < \frac{1}{4}\} = (\{(x,y) \mid 0 \leq x+y < \frac{1}{4}\} \cap [0,1]^2) \cup (\mathbb{R}^2 \setminus [0,1]^2)$,
- $A_2^2 = \{(x,y) \mid \frac{1}{4} \leq f(x,y) < \frac{2}{4}\} = \{(x,y) \mid \frac{1}{4} \leq x+y < \frac{1}{2}\} \cap [0,1]^2$,
- $A_3^2 = \{(x,y) \mid \frac{2}{4} \leq f(x,y) < \frac{3}{4}\} = \{(x,y) \mid \frac{1}{2} \leq x+y < \frac{3}{4}\} \cap [0,1]^2$,
- $A_4^2 = \{(x,y) \mid \frac{3}{4} \leq f(x,y) < \frac{4}{4}\} = \{(x,y) \mid \frac{3}{4} \leq x+y < 1\} \cap [0,1]^2$,
- $A_5^2 = \{(x,y) \mid \frac{4}{4} \leq f(x,y) < \frac{5}{4}\} = \{(x,y) \mid 1 \leq x+y < \frac{5}{4}\} \cap [0,1]^2$,
- $A_6^2 = \{(x,y) \mid \frac{5}{4} \leq f(x,y) < \frac{6}{4}\} = \{(x,y) \mid \frac{5}{4} \leq x+y < \frac{3}{2}\} \cap [0,1]^2$,
- $A_7^2 = \{(x,y) \mid \frac{6}{4} \leq f(x,y) < \frac{7}{4}\} = \{(x,y) \mid \frac{3}{2} \leq x+y < \frac{7}{4}\} \cap [0,1]^2$,
- $A_8^2 = \{(x,y) \mid \frac{7}{4} \leq f(x,y) < \frac{8}{4}\} = \{(x,y) \mid \frac{7}{4} \leq x+y < 8\} \cap [0,1]^2$, and
- $A^2 = \{(1,1)\}$ (see below).



The second simple approximation is thus

$$s_2 = \sum_{i=1}^{2(2^2)} \frac{i-1}{2^2}\cdot \chi_{A_i^2} + 2\cdot \chi_{A^2} = \sum_{i=1}^{8} \frac{i-1}{4}\cdot \chi_{A_i^2} + 2\cdot \chi_{A^2},$$

whose graph is shown on the next page:

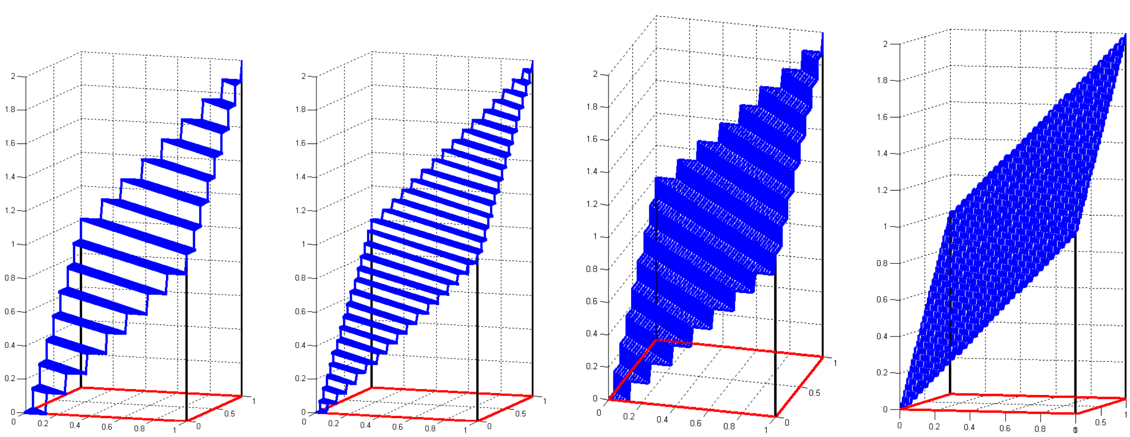


We then have

$$\iint s_2(x,y)\,\mathrm{d}x\,\mathrm{d}y = \sum_{i=1}^{8} \frac{i-1}{4} \cdot \mathrm{Area}(A_i^2) + 2 \cdot \mathrm{Area}(A^2),$$

whose value we again leave un-evaluated.

The process continues in the same way for all n, yielding a sequence of positive simple functions.



At step n, we have:

- $A_1^n = \left(\{(x,y) \mid 0 \leq x+y < \frac{1}{2^n}\} \cap [0,1]^2\right) \cup (\mathbb{R}^2 \setminus [0,1]^2)$,
- $A_i^n = \{(x,y) \mid \frac{i-1}{2^n} \leq x+y < \frac{i}{2^n}\} \cap [0,1]^2$ for $2 \leq i \leq 2^{n+1}$,
- $A_{2^{n+1}+1}^n = \{(1,1)\}$ and $A^n = A_j^n = \varnothing$ for $j > 2^{n+1}+1$.

Then the nth simple approximation is

$$s_n = \sum_{i=1}^{n \cdot 2^n} \frac{i-1}{2^n} \cdot \chi_{A_i} + n \cdot \chi_{A^n} = \sum_{i=1}^{2^{n+1}} \frac{i-1}{2^n} \cdot \chi_{A_i^n} + 2 \cdot \chi_{A_{2^{n+1}+1}^n},$$

so that

$$\begin{aligned}\iint s_n(x,y)\,\mathrm{d}x\,\mathrm{d}y &= \sum_{i=1}^{2^{n+1}} \frac{i-1}{2^n} \cdot \text{Area}(A_i^n) + 2 \cdot \underbrace{\text{Area}(A_{2^{n+1}+1}^n)}_{=0} \\ &= \sum_{i=1}^{2^n} \frac{i-1}{2^n} \cdot \text{Area}(A_i^n) + \sum_{i=2^n+1}^{2^{n+1}} \frac{i-1}{2^n} \cdot \text{Area}(A_i^n).\end{aligned}$$

We can show (see Exercises) that

$$\text{Area}(A_i^n) = \begin{cases} \frac{1}{4^n}\left(i - \frac{1}{2}\right) & \text{for } 1 \leq i \leq 2^n \\ \frac{1}{4^n}\left(2^{n+1} - i - \frac{1}{2}\right) & \text{for } 2^n+1 \leq i \leq 2^{n+1} \end{cases}$$

In general, then, we have:

$$\begin{aligned}\iint s_n(x,y)\,\mathrm{d}x\,\mathrm{d}y &= \sum_{i=1}^{2^n} \frac{i-1}{2^n} \cdot \frac{1}{4^n}\left(i - \frac{1}{2}\right) + \sum_{i=2^n+1}^{2^{n+1}} \frac{i-1}{2^n} \cdot \frac{1}{4^n}\left(2^{n+1} - i - \frac{1}{2}\right) \\ &= \frac{1}{2^n r^n}\left[\sum_{i=1}^{2^n}(i-1)(i-1/2) + \sum_{i=1}^{2^{n+1}}(i-1)(2^{n+1}-i-1/2)\right] \\ &= 1 - \frac{1}{2^{n-1}} + \frac{1}{2 \cdot 4^n}.\end{aligned}$$

Write $B_n = \iint s_n\,\mathrm{d}x\,\mathrm{d}y$; we clearly have $B_n < 1$ for all n, and $B_n \to 1$. Then

$$\iint f(x,y)\,\mathrm{d}x\,\mathrm{d}y = \sup\left\{\iint s(x,y)\,\mathrm{d}x,\ \mathrm{d}y \;\middle|\; s \in \zeta_+^{(2)}, s \leq f\right\} \geq 1 = \lim_{n\to\infty} B_n.$$

For $s \in \zeta_+^{(2)}$, we have seen that

$$\iint s(x,y)\,\mathrm{d}x\,\mathrm{d}y = \sum_{j=1}^{m} \alpha_j \cdot \text{Area}(A_i),$$

and so the integral represents the volume of a collection of m prisms with base area A_j and height α_j. By construction,

$$\iint s(x,y)\,\mathrm{d}x\,\mathrm{d}y \leq \text{Volume}(\text{solid bounded by } 0\leq x,y\leq 1 \text{ and } 0\leq z\leq x+y).$$

We cannot compute the volume using integrals as we have not yet established that the integral of a general positive Borel function over a domain A is the volume of the solid bounded by f over A, but we see easily that the solid in question is exactly the bottom half of the prism defined by $0\leq x,y\leq 1$ and $0\leq z\leq 2$, whose volume we know to be 2, from geometry (see the bottom image on p. 493).

By definition, we must then have

$$\sup\left\{\iint s(x,y)\,\mathrm{d}x,\ \mathrm{d}y \;\middle|\; s\in\zeta_+^{(2)}, s\leq f\right\} \leq \frac{1}{2}(2)=1,$$

which, combined with the previous inequality, shows that

$$\iint f(x,y)\,\mathrm{d}x\,\mathrm{d}y = 1.$$

Phew! □

If $f\in\zeta_+^{(2)}$, both definitions **coincide**: i.e, if $f=\sum\alpha_i\chi_{A_i}$, with $\alpha_i\in\overline{\mathbb{R}}$, $A_i\in\mathcal{B}(\mathbb{R}^2)$, and $A_1\sqcup\cdots\sqcup A_n=\mathbb{R}^2$, then

$$\iint f(x,y)\,\mathrm{d}x\,\mathrm{d}y = \sum_{i=1}^n \alpha_i\cdot\text{Area}(A_i) = I(f) = \sup\left\{\iint s(x,y)\,\mathrm{d}x\,\mathrm{d}y \;\middle|\; s\in\zeta_+^{(2)}, s\leq f\right\}.$$

Indeed, if $f\in\zeta_+^{(2)}$, we have $\iint f(x,y)\,\mathrm{d}x\,\mathrm{d}y\leq I(f)$. On the other hand, if $s\in\zeta_+^{(2)}$, with $s\leq f$, then

$$\iint s(x,y)\,\mathrm{d}x\,\mathrm{d}y \leq \iint f(x,y)\,\mathrm{d}x\,\mathrm{d}y,$$

according to Lemma 291.3, from which we conclude that

$$I(f)=\sup\left\{\iint s(x,y)\,\mathrm{d}x\,\mathrm{d}y \;\middle|\; s\in\zeta_+^{(2)}, s\leq f\right\} \leq \iint f(x,y)\,\mathrm{d}x\,\mathrm{d}y \leq I(f).$$

The next result shows that Lemma 291.3 also applies to positive Borel functions.

Proposition 292
If f,g are positive Borel functions and if $f\leq g$, then

$$\iint f\,\mathrm{d}x\,\mathrm{d}y \leq \iint g\,\mathrm{d}x\,\mathrm{d}y.$$

Proof: if $f \leq g$, then $\{s \in \zeta_+^{(2)} \mid s \leq f\} \subseteq \{s \in \zeta_+^{(2)} \mid s \leq g\}$ whence

$$\left\{ \iint s(x,y)\,\mathrm{d}x\,\mathrm{d}y \,\middle|\, s \in \zeta_+^{(2)}, s \leq f \right\} \subseteq \left\{ \iint s(x,y)\,\mathrm{d}x\,\mathrm{d}y \,\middle|\, s \in \zeta_+^{(2)}, s \leq g \right\}$$

So that

$$\sup \left\{ \iint s(x,y)\,\mathrm{d}x\,\mathrm{d}y \,\middle|\, s \in \zeta_+^{(2)}, s \leq f \right\} \subseteq \sup \left\{ \iint s(x,y)\,\mathrm{d}x\,\mathrm{d}y \,\middle|\, s \in \zeta_+^{(2)}, s \leq g \right\}$$

■

One might wonder why exactly we bothered to introduce the Borel-Lebesgue integral – while going from Riemann sums to simple functions does change our viewpoint of integration, are the corresponding integrals equivalent, or is one "preferable" over the other?

Theorem 293 (Lebesgue Monotone Convergence Theorem)
Let $(f_n)_{n\geq 1}$ be a sequence of Borel functions on $\mathbb{R}^2$ such that

1. $0 \leq f_1(x,y) \leq f_2(x,y) \leq \cdots \leq f_n(x,y) \leq \cdots \quad \forall (x,y) \in \mathbb{R}^2$*, and*
2. $f_n(x,y) \to f(x,y) \quad \forall (x,y) \in \mathbb{R}^2$.

Then f is a Borel function on $\mathbb{R}^2$ and $\iint f_n(x,y)\,\mathrm{d}x\,\mathrm{d}y \to \iint f(x,y)\,\mathrm{d}x\,\mathrm{d}y$. In particular, $\iint f\,\mathrm{d}x\,\mathrm{d}y = \lim_{n\to\infty} \iint s_n\,\mathrm{d}x\,\mathrm{d}y$, whenever (s_n) is a monotonically increasing sequence of positive simple functions bounded above by f, with $s_n \to f$ (pointwise).

Proof: left as a (difficult) exercise. ■.

Theorem 293 suggests that the new definition has a clear advantage: what additional constraint does the equivalent limit interchange theorem 69 of Riemann integration require?

Corollary 294
Let $f, g : \mathbb{R}^2 \to [0,\infty]$ be Borel functions and $\alpha \geq 0$. Then

1. $\iint (f+g)\,\mathrm{d}x\,\mathrm{d}y = \iint f\,\mathrm{d}x\,\mathrm{d}y + \iint g\,\mathrm{d}x\,\mathrm{d}y$,
2. $\iint \alpha f\,\mathrm{d}x\,\mathrm{d}y = \alpha \iint f\,\mathrm{d}x\,\mathrm{d}y$.

Proof: left as an exercise. ■.

From this point on, in order to not have to rely on the notation of **iterated integrals**, we write

$$\int f\,\mathrm{d}m = \int \cdots \int f(x_1,\ldots,x_n)\,\mathrm{d}x_1 \cdots \mathrm{d}x_n$$

and $m(B)$ for the **measure** of $B \subseteq \mathbb{R}^n$ (a generalization of the length, area, volume).

Theorem 295
Let f be a positive Borel function, taking on the value 0 outside of a Borel set A with Area$(A) = 0$. Then $\iint f \, \mathrm{d}x \, \mathrm{d}y = 0$.

Proof: let

$$k(x, y) = \begin{cases} \infty & (x, y) \in A \\ 0 & (x, y) \notin A \end{cases}$$

THen $k \in \zeta_+^{(2)}$ and

$$\int k \, \mathrm{d}m == 0 \cdot \text{Area}(\mathbb{R}^2 \setminus A) + \infty \cdot \text{Area}(A) = 0 \cdot \infty + \infty \cdot 0 = 0,$$

by convention. Since $f \leq k$, then

$$0 \leq \int f \, \mathrm{d}m \leq \int k \, \mathrm{d}m = 0,$$

which completes the proof. ∎

We say that a positive Borel function f is **(Borel-Lebesgue) integrable** if $\int f \, \mathrm{d}m < \infty$. If $f \geq 0$ is integrable and $g \leq f$ is a Borel function, then

$$\infty > \int f \, \mathrm{d}m = \int g \, \mathrm{d}m + \int (f - g) \, \mathrm{d}m \geq \int g \, \mathrm{d}m,$$

and so g is also integrable. This result definitely does not hold in general for Riemann integration.[9]

Theorem 296
Let g be a bounded positive Borel function, taking on the value 0 outside a bounded Borel set A. Then g is integrable.

Proof: let M be such that $g(z) \leq M$. By definition, $\exists B = [a_1, a_1'] \times [a_2, a_2']$ such that $A \subseteq B$ and $g(z) = 0$ if $z \notin B$. Then $g \leq M\chi_B$ and

$$\int g \, \mathrm{d}m \leq \int M\chi_B \, \mathrm{d}m = M \cdot \text{Area}(\chi_B) < \infty,$$

which completes the proof. ∎

We can extend the idea to general Borel functions using the positive and negative parts.

Note that the Riemann and Borel-Lebesgue integral **coincide** when the former exists.

[9] Can you think of a counterexample?

21.4 Integral of Borel Functions

For a general function $f : \mathbb{R}^n \to \mathbb{R}$, define the **positive part of** f by

$$f_+(x) = \begin{cases} f(x) & \text{when } f(x) \geq 0 \\ 0 & \text{when } f(x) < 0, \end{cases}$$

and the **negative part of** f by

$$f_-(x) = \begin{cases} -f(x) & \text{when } f(x) \leq 0 \\ 0 & \text{when } f(x) > 0. \end{cases}$$

Then $f = f_+ - f_-$ and $|f| = f_+ + f_-$.

If $f : \mathbb{R}^n \to \mathbb{R}$ is a finite Borel function, then f_+, f_- are positive Borel functions, by definition. A Borel function $f : \mathbb{R}^n \to \overline{\mathbb{R}}$ is **integrable** if both f_+ and f_- are integrable. In this case, we define

$$\int f \,\mathrm{d}m = \int f_+ \,\mathrm{d}m - \int f_- \,\mathrm{d}m.$$

We see now that Lemma 291 has a counterpart for Borel functions.

Theorem 297
Let f, g be integrable functions and $\lambda \in \mathbb{R}$. Then

1. $\int \lambda f \,\mathrm{d}m = \lambda \int f \,\mathrm{d}m$,
2. $\int (f + g) \,\mathrm{d}m = \int f \,\mathrm{d}m + \int g \,\mathrm{d}m$, and
3. If $f \leq g$ then $\int f \,\mathrm{d}m \leq \int g \,\mathrm{d}m$.

Proof: since f, g are integrable, we have

$$\int f \,\mathrm{d}m = \int f_+ \,\mathrm{d}m - \int f_- \,\mathrm{d}m < \infty, \quad \text{and} \quad \int g \,\mathrm{d}m = \int g_+ \,\mathrm{d}m - \int g_- \,\mathrm{d}m < \infty.$$

1. Assume $\lambda \geq 0$. Then

$$\begin{aligned} \infty > \lambda \int f \,\mathrm{d}m &= \lambda \left(\int f_+ \,\mathrm{d}m - \int f_- \,\mathrm{d}m \right) = \lambda \int f_+ \,\mathrm{d}m - \lambda \int f_- \,\mathrm{d}m \\ \boxed{\text{Corollary 294}} &= \int \lambda f_+ \,\mathrm{d}m - \int \lambda f_- \,\mathrm{d}m = \int (\lambda f)_+ \,\mathrm{d}m - \int (\lambda f)_+ \,\mathrm{d}m \\ &= \int \lambda f \,\mathrm{d}m, \end{aligned}$$

which simultaneously shows that λf is integrable.

The only thing left to do is to show that the property holds for $\lambda - 1$. Note that $(-f)_+ = f_-$ and that $(-f_-) = f_+$, so that $-f$ is itself integrable. Then

$$\begin{aligned} -\int f \, \mathrm{d}m = -\int f_+ \, \mathrm{d}m + \int f_- \, \mathrm{d}m = \int f_- \, \mathrm{d}m - \int f_+ \, \mathrm{d}m \\ = \int (-f)_+ \, \mathrm{d}m - \int (-f)_- \, \mathrm{d}m = \int (-f) \, \mathrm{d}m, \end{aligned}$$

because $-f$ is integrable.

2. By definition, we have

$$f + g = (f_+ - f_-) + (g_+ - g_-) = (f_+ + g_+) - (f_- + g_-).$$

According to the second solved problem (see p. 512), $f + g$ is thus integrable and

$$\begin{aligned} \int (f+g) \, \mathrm{d}m &= \int [(f_+ + g_+) - (f_- + g_-)] \, \mathrm{d}m \\ &= \int (f_+ + g_+) \, \mathrm{d}m - \int (f_- + g_-) \, \mathrm{d}m \\ \boxed{\text{Corollary 294}} &= \int f_+ \, \mathrm{d}m + \int g_+ \, \mathrm{d}m - \int f_- \, \mathrm{d}m - \int g_- \, \mathrm{d}m = \int f \, \mathrm{d}m + \int g \, \mathrm{d}m. \end{aligned}$$

3. Since $g - f \geq 0$ and $g = f + (g - f)$, we have

$$\int g \, \mathrm{d}m = \int f \, \mathrm{d}m + \int g - f \, \mathrm{d}m \geq \int f \, \mathrm{d}m,$$

according to Corollary 294 and Proposition 292. ■

The set $\mathcal{V}_n = \{f : \mathbb{R}^n \to \mathbb{R} \mid f \text{ finite, Borel, integrable}\}$ is a **vector space over** $\mathbb{R}$; the integral of f over $\mathbb{R}^n$ is a **linear functional**, which is to say that

$$\int_{\mathbb{R}^n} \_\_ \, \mathrm{d}m : \mathcal{V}_n \to \mathbb{R}$$

is a **linear functional**.

Theorem 298
Let $B \in \mathcal{B}(\mathbb{R}^n)$, with $m(B) = 0$. If f, g are Borel functions such that $f = g$ on $\mathbb{R}^n \setminus B$ and if f is integrable, then g is integrable and $\int f \, \mathrm{d}m = \int g \, \mathrm{d}m$.

Proof: the functions $f-g$ is a Borel function with $f-g\equiv 0$ on $\mathbb{R}^n\setminus B$. Since $f=g+(f-g)$, we have

$$\int f\,\mathrm{d}m=\int g\,\mathrm{d}m+\int (f-g)\,\mathrm{d}m.$$

Write $h=f-g$; then $\int h\,\mathrm{d}m=0$. Since $h_+,h_-=0$ on $\mathbb{R}^n\setminus B$, we must have

$$\int h_+\,\mathrm{d}m=\int h_-\,\mathrm{d}m=0,$$

according to Theorem 295. Then

$$\int h\,\mathrm{d}m=\int h_+\,\mathrm{d}m-\int h_-\,\mathrm{d}m=0\quad\text{and}\quad\int f\,\mathrm{d}m-\int g\,\mathrm{d}m=0\Longrightarrow\int f\,\mathrm{d}m=\int g\,\mathrm{d}m,$$

which completes the proof. ■

21.5 Integration Over a Subset

To this point, we have studied integration over $\mathbb{R}^n$ in its entirety:

$$\int f\,\mathrm{d}m=\int f\,\mathrm{d}mA.$$

But we can also integrate functions over substes of $\mathbb{R}^n$. Let $A\in\mathcal{B}(\mathbb{R}^n)$ and $f:A\subseteq\mathbb{R}^n\to\mathbb{R}$. If the function $f\chi_A:\mathbb{R}^n\to\mathbb{R}$ defined by

$$(f\chi_A)(x)=\begin{cases}f(x) & x\in A\\ 0 & x\notin A\end{cases}$$

is a Borel function and if $f\chi_A\geq 0$ or $f\chi_A$ is integrable, we define

$$\int_A f\,\mathrm{d}m=\int f\chi_A\,\mathrm{d}m.$$

We can show (see Exercises and Theorem 296) that if f is bounded on A and $f\chi_A$ is a Borel function, then $f\chi_A$ is **integrable**. When $\int_A f\,\mathrm{d}m<\infty$, we say that f is **integrable on** A.

Theorem 299
Let $A,B\in\mathcal{B}(\mathbb{R}^n)$, $A\cap B=\varnothing$. If f is a Borel function on $A\cup B$, then

1. *if $f\geq 0$, $\int_{A\cup B} f\,dm=\int_A f\,dm+\int_B f\,dm$, and*
2. *f is integrable over $A\cup B$ if and only if f is integrable over A and B.*

Proof: left as an exercise. ■

If $m(B)=0$, then $\int_B f\,\mathrm{d}m=0$. In that case $\displaystyle\int_{A\cup B} f\,\mathrm{d}m=\int_A f\,\mathrm{d}m.$

21.6 Multiple Integrals

The example of Section 21.4 shows that while we can compute the (Borel-Lebesgue) integral of a relatively straightforward integrand f, the process can leave a lot to be desired.[10] Let $f : \mathbb{R}^2 \to \mathbb{R}$ be a bounded Borel function, that is 0 outside of a bounded region. For all $y \in \mathbb{R}$, $x \mapsto f(x,y)$ is a Borel bounded function that is 0 outside of a bounded subset of $\mathbb{R}$, hence $x \mapsto f(x,y)$ is integrable.

Theorem 300 (Fubini's Theorem)
Let $f : \mathbb{R}^2 \to [0,\infty]$ be a Borel function. For every y, let $F(y) = \int_{\mathbb{R}} f(x,y)\,dx$. Then F is a Borel function and

$$\int_{\mathbb{R}^2} f\,\mathrm{d}m = \iint_{\mathbb{R}^2} f(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_{\mathbb{R}} F(y)\,\mathrm{d}y = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(x,y)\,\mathrm{d}x \right)\,\mathrm{d}y.$$

Proof: left as an exercise. ■

Similarly, if $G(x) = \int_{\mathbb{R}} f(x,y)\,\mathrm{d}x$, we have

$$\int_{\mathbb{R}^2} f\,\mathrm{d}m = \int_{\mathbb{R}} G(x)\,\mathrm{d}x = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} f(x,y)\,\mathrm{d}y \right)\,\mathrm{d}x,$$

Example: let $f : \mathbb{R}^2 \to [0,\infty]$ be defined by $f(x,y) = (x+y)^{-4}$, where $A \subseteq \mathbb{R}^2$ is the triangle bounded by $x=1$, $y=1$, and $x+y=4$. Compute $\int_A f\,\mathrm{d}m$.

Solution: the triangle's three vertices are located at $(1,1)$, $(1,3)$, and $(3,1)$. For a fixed $x \in \mathbb{R}$, we have

$$F(x) = \int_{\mathbb{R}} f(x,y)\,\mathrm{d}y = \begin{cases} 0 & \text{if } x \not\in [1,3] \\ \int_{[1,4-x]} (x+y)^{-4}\,\mathrm{d}y & \text{otherwise} \end{cases}$$

But

$$\int_{[1,4-x]} \frac{\mathrm{d}y}{(x+y)^4} = \int_1^{4-x} (x+y)^{-4}\,\mathrm{d}y = \left[\frac{(x+y)^{-3}}{-3} \right]_{y=1}^{y=4-x} = \frac{(x+1)^{-3}}{3} - \frac{1}{192},$$

from which we have

$$\int_A f\,\mathrm{d}m = \int_{[1,3]} F(x)\,\mathrm{d}x = \int_1^3 \left[\frac{(x+1)^{-3}}{3} - \frac{1}{192} \right]\mathrm{d}x = \left[\frac{(x+1)^{-2}}{3(-2)} - \frac{x}{192} \right]_{x=1}^{3} = \frac{1}{48}. \qquad \square$$

[10] As in the previous sections, we will provide the important details for functions $\mathbb{R}^2 \to \mathbb{R}$; the process is easy to generalize to $\mathbb{R}^n$.

If f is a positive Borel function, we can interchange the order of integration (as in Theorem 300); for general functions, there are complications. One way out of the quagmire is to decompose $f = f_+ - f_-$ and to integrate f_+ and f_- separately, but that can quickly get cumbersome.

Theorem 301 (SPECIAL FUBINI THEOREM)
Let $f : \mathbb{R}^2 \to \mathbb{R}$ be a bounded Borel function taking on the value 0 outside of a bounded region. For all y, $x \mapsto f(x,y)$ is a bounded Borel function taking on the value 0 outside of a bounded subset of $\mathbb{R}$. Set $F(y) = \int_{\mathbb{R}} f(x,y)\,dx$. Then F is a bounded Borel function and

$$\int_{\mathbb{R}^2} f\,\mathrm{d}m = \iint_{\mathbb{R}^2} f(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_{\mathbb{R}} F(y)\,\mathrm{d}y = \int_{\mathbb{R}} G(x)\,\mathrm{d}x.$$

Proof: by hypothesis, $\exists M, N > 0$ such that $|f(x,y)| \leq M$ for all $(x,y) \in \mathbb{R}^2$ and $f(x,y) = 0$ for all $(x,y) \notin [-N,N]^2$.

For a fixed $y = y_0$, $x \mapsto f(x,y_0)$ is a Borel function, with $|f(x,y_0)| \leq M$ for all x (and y_0) and $f(x,y_0) = 0$ when $|x| > N$. If $|y_0| > N$, $F(y_0) = 0$; more generally,

$$|F(y_0)| \leq \int_{-N}^{N} M\,\mathrm{d}x = 2MN,$$

so it is bounded.

It remains to see that F is a Borel function and that conclusion of the theorem holds. Using the decomposition $f = f_+ - f_-$, we reduce the problem to the case $f \geq 0$; it then suffices to apply Theorem 300 to each of the positive and negative parts of f, completing the proof. ■

The result generalizes to $\mathbb{R}^n$ in the natural way.

Example: Let $f : A \subseteq \mathbb{R}^3 \to \mathbb{R}$ be defined by $f(x,y,z) = 2xyz \cdot \chi_A(x,y,z)$, where

$$A = \{(x,y,z) \mid x \geq 0, y \geq 0, z \geq 0, x^2 + y^2 + z^2 \leq 1\}.$$

Compute

$$I = \int f\,\mathrm{d}m = \iiint_{\mathbb{R}^3} f(x,y,z)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z.$$

Solution: let $B = \{(x,y,z) \mid x^2 + y^2 \leq 1, x \geq 0, y \geq 0, z = 0\}$. For fixed $x, y \in \mathbb{R}^2$, we have

$$F(x,y) = \int_{\mathbb{R}} 2xyz \cdot \chi_A(x,y,z)\,\mathrm{d}z = \begin{cases} 0 & \text{if } (x,y,0) \notin B \\ \int_{[0,\sqrt{1-x^2-y^2}]} 2xyz\mathrm{d}z & \text{if } (x,y,0) \in B \end{cases}$$

Since

$$\int_0^{\sqrt{1-x^2-y^2}} 2xyz \,\mathrm{d}z = 2xy \left[\frac{z^2}{2}\right]_{z=0}^{z=\sqrt{1-x^2-y^2}} = xy(1-x^2-y^2),$$

the desired integral is

$$I = \iint_{\mathbb{R}^2} F(x,y)\,\mathrm{d}x\,\mathrm{d}y = \iint_B xy(1-x^2-y^2)\,\mathrm{d}x\,\mathrm{d}y.$$

We can decompose this double integral as follows: for $0 \le x \le 1$, set

$$G(x) = \int_0^{\sqrt{1-x^2}} xy(1-x^2-y^2)\,\mathrm{d}y = \frac{x}{4}(1-x^2)^2;$$

otherwise, set $G(x) = 0$. Then

$$I = \int_{\mathbb{R}} G(x)\,\mathrm{d}x = \frac{1}{4}\int_{[0,1]} x(1-x^2)^2\,\mathrm{d}x = \frac{1}{24}. \qquad \square$$

In general, if $D \subseteq \mathbb{R}^n$ is a Borel set, then

$$m(D) = \int \chi_D \,\mathrm{d}m.$$

If $n = 2$, this takes the form

$$\text{Area}(D) = \iint_{\mathbb{R}^2} \chi_D(x,y)\,\mathrm{d}x\,\mathrm{d}y;$$

if $n = 3$, we have

$$\text{Vol}(D) = \iiint_{\mathbb{R}^3} \chi_D(x,y,z)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z.$$

Examples

1. Let $a, b > 0$. Find the area of the ellipse $A = \{(x,y) \in \mathbb{R}^2 \mid x^2/a^2 + y^2/b^2 \le 1\}$.

   Solution: rewrite

   $$A = \left\{ (x,y) \in \mathbb{R}^2 \,\middle|\, -a \le x \le a, -\frac{b}{a}\sqrt{a^2-x^2} \le y \le \frac{b}{a}\sqrt{a^2-x^2} \right\}.$$

   Then

   $$\text{Area}(A) = \iint_{\mathbb{R}^2} \chi_A(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_{-a}^{a} \left(\int_{\mathbb{R}} \chi_A(x,y)\,\mathrm{d}y \right) \mathrm{d}x$$

But

$$\int_{\mathbb{R}} \chi_A(x,y)\,\mathrm{d}y = \begin{cases} 0 & \text{if } x \notin [-a,a] \\ \int_{-b/a\sqrt{a^2-x^2}}^{b/a\sqrt{a^2-x^2}} \mathrm{d}y = \frac{2b}{a}\sqrt{a^2-x^2} & \text{if } x \in [-a,a] \end{cases}$$

Then

$$\begin{aligned} \text{Area}(A) &= \frac{2b}{a}\int_{x=-a}^{x=a} \sqrt{a^2-x^2}\,\mathrm{d}x \\ \boxed{x = a\cos\varphi,\ \mathrm{d}x = -a\sin\varphi\,\mathrm{d}\varphi} &= \frac{2b}{a}\int_{\varphi=\pi}^{\varphi=0} \sqrt{a^2(1-\cos^2\varphi)}(-a\sin\varphi)\,\mathrm{d}\varphi \\ &= -\frac{2b}{a}\int_{\pi}^{0} a^2\sin^2\varphi\,\mathrm{d}\varphi = 2ab\int_0^{\pi}\sin^2\varphi\,\mathrm{d}\varphi \\ &= 2ab\int_0^{\pi}\left(\frac{1-\cos 2\varphi}{2}\right)\mathrm{d}\varphi = ab\left[\varphi - \frac{\sin 2\varphi}{2}\right]_0^{\pi} = \pi ab. \end{aligned}$$

2. Let $a,b,c>0$ and $E = \{(x,y,z) \mid x^2/a^2+y^2/b^2+z^2/c^2 \leq 1\}$. Find Vol$(E)$.

Solution: we have

$$\text{Vol}(E) = \iiint_{\mathbb{R}^3} \chi_E(x,y,z)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \int_{-c}^{c} \underbrace{\left(\iint_{\mathbb{R}^2} \chi_E(x,y,z)\,\mathrm{d}x\,\mathrm{d}y\right)}_{=\text{Area}(E_z)} \mathrm{d}z,$$

where

$$E_z = \left\{(x,y) \;\middle|\; \frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2} \leq 1\right\} = \left\{(x,y) \;\middle|\; \frac{x^2}{(ah)^2}+\frac{y^2}{(bh)^2} \leq 1\right\},$$

where $h = \sqrt{1-z^2/c^2} > 0$.

According to the preceding example, we know that

$$\text{Area}(E_z) = \pi(ah)(bh) = \pi abh^2 = \pi ab(1-z^2/c^2)$$

when $|z| \leq c$, so that

$$\text{Vol}(E) = \int_{-c}^{c} \pi ab\left(1-\frac{z^2}{c^2}\right)\mathrm{d}z = \pi ab\left[z - \frac{z^3}{3c^2}\right]_{z=-c}^{z=c} = \frac{4\pi}{3}abc. \qquad \square$$

We finish the chapter with some detail regarding one of the most commonly-used integration shortcuts: **changes of variables**.

21.7 Change of Variables and/or Coordinates

In the preceding section's example where we compute the area of an ellipse, we encounter an integral in x which we cannot compute directly; instead we introduce a new variable φ and a relation between x and φ that we leverage to easily compute the integral. We formalize the process in this section.

Let $\Psi : U \subseteq_O \mathbb{R}^n \to V \subseteq_O \mathbb{R}^n$ be a **diffeomorphism**; thus, Ψ and Ψ^{-1} are C^1, $\Psi \circ \Psi^{-1}(v) = v$, $\Psi^{-1} \circ \Psi(u) = u$, the **Jacobians** $\mathrm{d}\Psi(u), \mathrm{d}\Psi^{-1}(v) : \mathbb{R}^n \to \mathbb{R}^n$ are linear maps and

$$\mathrm{d}(\Psi \circ \Psi^{-1})(v) = \mathrm{d}\Psi(\Psi^{-1}(v))\mathrm{d}\Psi^{-1}(v) = I_n,$$

for all $u \in U, v \in V$, which means that $\mathrm{d}\Psi(u)$ and $\mathrm{d}\Psi^{-1}(v)$ are invertible for all $u \in U, v \in V$.

Examples

1. For $n = 1$, define $\Psi : U = (0, \pi) \to V = (-1, 1)$ by $\Psi(u) = \cos u$. Then $\mathrm{d}\Psi(u) = -\sin u < 0$ for all $u \in (0, \pi)$, i.e., Ψ is decreasing on $(0, \pi)$, with $\Psi(0) = 1$ and $\Psi(\pi) = -1$.

2. For $n = 1$, let $U = V = (0, 1)$ and define $\Psi : U \to V$ by $\Psi(u) = u^2$. Then $\mathrm{d}\Psi(u) = 2u > 0$ for all $u \in U$, i.e., Ψ is increasing on U, with $\Psi(0) = 0$ and $\Psi(1) = 1$.

3. For $n = 2$, let $U = \{(r, \theta) \mid r > 0 \text{ and } -\pi < \theta < \pi\}$, $V = \mathbb{R}^2 \setminus \{(x, 0) \mid x \leq 0\}$, and define $\Psi(r, \theta) = (r \cos \theta, r \sin \theta)$. Then

$$\mathrm{d}\Psi(r, \theta) = \begin{pmatrix} \cos \theta & \sin \theta \\ -r \sin \theta & r \cos \theta \end{pmatrix}.$$

   Note that $J_\Psi(r, \theta) = \det(\mathrm{d}\Psi) = r \cos^2 \theta + r \sin^2 \theta = r > 0$ and that Ψ is:

   - **injective** since if $\Psi(r_1, \theta_1) = \Psi(r_2, \theta_2)$, then

$$r_1 = \|\Psi(r_1, \theta_1)\|_2 = \|\Psi(r_2, \theta_2)\|_2 = r_2$$

     and $\cos \theta_1 = \cos \theta_2$ and $\sin \theta_1 = \sin \theta_2$ yields $\theta_1 = \theta_2 \in (-\pi, \pi)$;
   - **surjective** since if $(x, y) \in V$, set $r = \sqrt{x^2 + y^2} > 0$; then

$$1 = \frac{x^2 + y^2}{r^2} = \frac{x^2}{r^2} + \frac{y^2}{r^2} \implies x = r \cos \theta, y = r \sin \theta \quad \text{for some } \theta \in (-\pi, \pi].$$

     But if $\theta = \pi$, then $x = -r$ and $y = 0$, so that $(x, y) \notin V$, a contradiction; thus $\theta \in (-\pi, \pi)$.

Thus $\Psi : U \to V$ is a bijection; its inverse is $\Psi^{-1} : V \to U$ is defined by $\Psi^{-1}(x,y) = (r,\theta)$, as given on the previous page. It is easy to verify that $\Psi \circ \Psi^{-1} : V \to V$ is the identity, as

$$\Psi(\Psi^{-1}(x,y)) = \Psi(\sqrt{x^2+y^2},\theta) = \Psi(r,\theta) = (r\cos\theta, r\sin\theta) = (x,y).$$

Both Ψ and Ψ^{-1} are C^1 and the Jacobians $\mathrm{d}\Psi(r,\theta)$ and $\mathrm{d}\Psi^{-1}(x,y)$ are invertible (see Exercises); as such, Ψ is a diffeomorphism between U and V. In this particular case, we can express θ explicitly in terms of (x,y):

$$\theta \in (-\pi,\pi) \Longrightarrow \frac{\theta}{2} \in \left(-\frac{\pi}{2},\frac{\pi}{2}\right) \Longrightarrow \cos(\theta/2) \neq 0;$$

then

$$\tan(\theta/2) = \frac{\sin(\theta/2)}{\cos(\theta/2)} = \frac{\sin\theta}{1+\cos\theta} = \frac{r\sin\theta}{r(1+\cos\theta)} = \frac{y}{\sqrt{x^2+y^2}+x}$$
$$\Longrightarrow \theta = 2\text{Arctan}\left(\frac{y}{\sqrt{x^2+y^2}+x}\right). \qquad \square$$

If $f : V \to \overline{\mathbb{R}}$ is a Borel function, let $J_\Psi(z) = \det(\mathrm{d}\Psi(z))$; then $J_\Psi(z) \neq 0$ since Ψ is a diffeomorphism, and the composition $f \circ \Psi : U \to \overline{\mathbb{R}}$ is also a Borel function. In $\mathbb{R}^2$, for instance, if $\Psi(s,t) = (x,y) = (x(s,t), y(s,t))$, then

$$J_\Psi(s,t) = \det\begin{pmatrix} \frac{\partial x(s,t)}{\partial s} & \frac{\partial x(s,t)}{\partial t} \\ \frac{\partial y(s,t)}{\partial s} & \frac{\partial y(s,t)}{\partial t} \end{pmatrix} = \frac{\partial x(s,t)}{\partial s}\cdot\frac{\partial y(s,t)}{\partial t} - \frac{\partial x(s,t)}{\partial t}\cdot\frac{\partial y(s,t)}{\partial s} \neq 0.$$

Theorem 301 (CHANGE OF VARIABLES)

1. *Let $f : V \to [0,\infty]$ be a positive Borel function. Then*

$$\iint_V f(x,y)\,\mathrm{d}x\,\mathrm{d}y = \iint_U f(x(s,t),y(s,t))\,|J_\Psi(s,t)|\,\mathrm{d}s\,\mathrm{d}t.$$

2. If $f : V \to \overline{\mathbb{R}}$ is an integrable Borel function, then $f \circ \Psi |J_\Psi|$ is Borel and integrable on U and

$$\iint_V f(x,y)\,\mathrm{d}x\,\mathrm{d}y = \iint_U f \circ \Psi(s,t)\,|J_\varphi(s,t)|\,\mathrm{d}s\,\mathrm{d}t.$$

Proof: left as an exercise. ■

As usual, this result easily generalizes to $\mathbb{R}^n$.

Examples

1. For $n = 1$, if $\Psi : [\alpha, \beta] \to [a, b]$ is a bijection with $\Psi(\alpha) = a$, $\Psi(\beta) = b$, Ψ is C^1, and $\Psi' > 0$ on (α, β), then Ψ is an increasing diffeomorphism between $[\alpha, \beta]$ and $[a, b]$. Let $f : [a, b] \to \mathbb{R}$ be a continuous function. Then

$$\int_a^b f(u)\,\mathrm{d}u = \int_{[a,b]} f(u)\,\mathrm{d}u = \int_{(a,b)} f(u)\,\mathrm{d}u = \int_{[\alpha,\beta]} f(\Psi(t))|\Psi'(t)|\,\mathrm{d}t = \int_\alpha^\beta f(\Psi(t))\Psi'(t)\,\mathrm{d}t.$$

2. If Ψ is as in the previous example, but with $\Psi' < 0$ on (α, β), then

$$\int_a^b f(u)\,\mathrm{d}u = -\int_\beta^\alpha f(\Psi(t))\Psi'(t)\,\mathrm{d}t. \qquad \square$$

21.7.1 Polar Coordinates

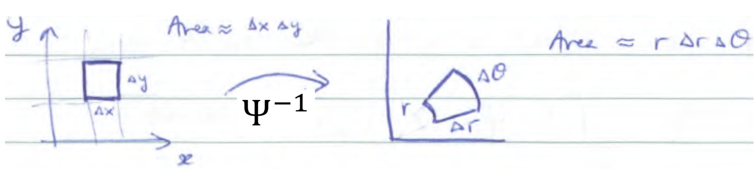
Let U, V, Ψ be as in the example on pp. 508-509. Then $J_\Psi(r, \theta) = r$. If $I = \{(x, 0) \mid x \leq 0\}$, then Area$(I) = 0$. Then, if $f : \mathbb{R}^2 \to [0, \infty]$ is a positive Borel function, we have

$$\iint_{\mathbb{R}^2} f(x, y)\,\mathrm{d}x\,\mathrm{d}y = \iint_V f(x, y)\,\mathrm{d}x\,\mathrm{d}y = \iint_U f(r\cos\theta, r\sin\theta)r\,\mathrm{d}r\,\mathrm{d}\theta.$$

If f is Borel and integrable over $\mathbb{R}^2$, then $(r, \theta) \mapsto f(r\cos\theta, r\sin\theta)r$ is integrable over U and

$$\iint_{\mathbb{R}^2} f(x, y)\,\mathrm{d}x\,\mathrm{d}y = \iint_U f(r\cos\theta, r\sin\theta)r\,\mathrm{d}r\,\mathrm{d}\theta.$$

This transformation yields **polar coordinates**, as illustrated below.



Example: for the Borel function $f : \mathbb{R}^2 \to \mathbb{R}$ defined by $f(x, y) = \exp(-x^2 - y^2)$, we have

$$\begin{aligned} I = \iint_{\mathbb{R}^2} \exp(-x^2 - y^2)\,\mathrm{d}x\,\mathrm{d}y &= \iint_U \exp(-r^2)r\,\mathrm{d}r\,\mathrm{d}\theta = \int_0^\infty \int_{-\pi}^{\pi} \exp(-r^2)r\,\mathrm{d}r\,\mathrm{d}\theta \\ &= \pi \int_0^\infty 2r\exp(-r^2)\,\mathrm{d}r = \pi \int_{u=0}^{u=\infty} \exp(-u)\,\mathrm{d}u = \pi. \end{aligned}$$

Since

$$I = \left(\int_{\mathbb{R}} \exp(-x^2)\,\mathrm{d}x\right)\left(\int_{\mathbb{R}} \exp(-y^2)\,\mathrm{d}y\right) = \left(\int_{\mathbb{R}} \exp(-x^2)\,\mathrm{d}x\right)^2 = \pi,$$

then

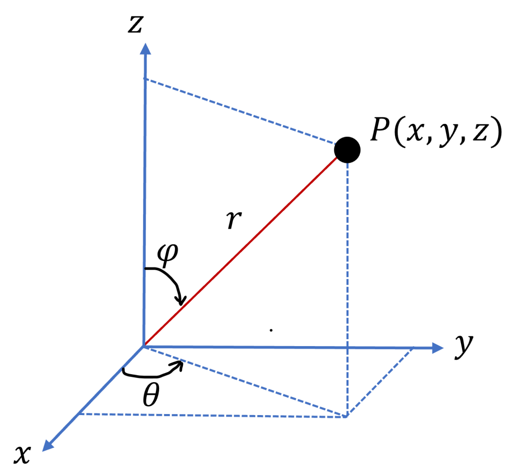
$$\int_{\mathbb{R}} \exp(-x^2)\,\mathrm{d}x = \sqrt{\pi};$$

we can compute the integral even though $\exp(-x^2)$ does not have an elementary anti-derivative. $\square$

21.7.2 Spherical Coordinates

In **spherical coordinates**, we represent the point $P(x,y,z) \in \mathbb{R}^3$ using the coordinates (r,φ,θ):

$$x = r\sin\varphi\cos\theta, \quad y = r\sin\varphi\sin\theta, \quad z = r\cos\varphi.$$



Let $U = \{(r,\varphi,\theta) \mid r > 0, 0 < \varphi < \pi, 0 < \theta < 2\pi\}$ and $V = \mathbb{R}^2 \setminus I_x = \mathbb{R}^3 \setminus \{(x,0,z) \mid x \geq 0\}$. Set $\Psi : U \to V$, with

$$\Psi(r,\varphi,\theta) = (r\sin\varphi\cos\theta, r\sin\varphi\sin\theta, r\cos\varphi).$$

Then

$$\mathrm{d}\Psi(r,\varphi,\theta) = \begin{pmatrix} \sin\varphi\cos\theta & \sin\varphi\sin\theta & \cos\varphi \\ r\cos\varphi & r\cos\varphi\sin\theta & -r\sin\varphi \\ -r\sin\varphi\sin\theta & r\sin\varphi\cos\theta & 0 \end{pmatrix},$$

so that $|J_\Psi(r,\varphi,\theta)| = r^2\sin\varphi$, because of the restrictions in the definition of U. Furthermore, $\mathrm{Vol}(I_x) = 0$; if $f : \mathbb{R}^3 \to [0,\infty]$ is a positive Borel function, we then have

$$\begin{aligned}\iiint_{\mathbb{R}^3} f(x,y,z)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z &= \iiint_V f(x,y,z)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z \\ &= \iiint_U f(r\sin\varphi\cos\theta, r\sin\varphi\sin\theta, r\cos\varphi)\,r^2\sin\varphi\,dr\,\mathrm{d}\varphi\,\mathrm{d}\theta.\end{aligned}$$

More generally, that relationship also holds if $f : \mathbb{R}^3 \to \overline{\mathbb{R}}$ is Borel and integrable.

> **Example:** compute the volume of the ball $B_R = \{(x, y, z) \mid x^2 + y^2 + z^2 \leq R^2\}$, for $R \geq 0$.
>
> **Solution:** according to the definition,
>
> $$\begin{aligned}\text{Vol}(B_R) &= \iiint_{B_R} \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \iiint_{\mathbb{R}^3} \chi_{B_R}(x, y, z)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z \\ &= \int_0^R \left(\int_0^\pi \left(\int_0^{2\pi} r^2 \sin\varphi\,\mathrm{d}\theta\,\mathrm{d}\varphi\,dr\right)\right) = 2\pi \int_0^R r^2 \left(\int_0^\pi \sin\varphi\,\mathrm{d}\varphi\right) dr \\ &= 2\pi \int_0^R r^2[-\cos\varphi]_0^\pi\,dr = 4\pi \int_0^R r^2\,dr = 4\pi \left[\frac{r^3}{3}\right]_0^R = \frac{4}{3}\pi R^3. \qquad \square\end{aligned}$$

21.8 Solved Problems

21.8.1 Borel-Lebesgue Integral on $\mathbb{R}^n$

1. Show that a bounded Borel function which is identically zero outside of a bounded set is integrable.

   Proof: by hypothesis, $\exists M \in \mathbb{R}^+$ such that $|g(z)| < M$ for all $z \in \mathbb{R}^n$. Furthermore, there is a bounded set A such that $g(z) = 0$ for all $z \notin A$. Since A is bounded, there exist $a_i, a_i' \in \mathbb{R}$ such that

   $$A \subseteq B = \prod_{i=1}^n [a_i, a_i']$$

   and $g(z) = 0$ for all $z \notin B$. Finally, $|g| \leq M\chi_B$ and

   $$\left|\int g\right| \leq \int |g| \leq \int M\chi_B = M \int \chi_B = M \cdot m(B) = M \prod_{i=1}^n (a_i' - a_i) < \infty,$$

   that is, g is integrable. ■

2. Let u, v be positive, integrable Borel functions. Show that $u - v$ is integrable and that

   $$\int (u - v)\,\mathrm{d}m = \int u\,\mathrm{d}m - \int v\,\mathrm{d}m.$$

   Proof: by hypothesis, $0 \leq \int u, \int v < \infty$, and so we also have $-\infty \leq \int u, \int v < \infty$. Then,

   $$\infty > \int u = \int (u - v + v) = \int (u - v) + \int v > -\infty$$

so that

$$\infty - \int v > \int (u-v) > -\infty - \int v$$

Since $-\infty < \int v < \infty$, $\infty - \int v = \infty$ and $-\infty - \int v = -\infty$. Finally, this yields

$$\infty > \int (u-v) > -\infty$$

and $u - v$ is integrable. We proved the other required result in the first inequality.■

3. If f is bounded on $A \in \mathcal{B}(\mathbb{R}^2)$, $f\chi_A$ is a Borel function, and Area$(A) < \infty$, show that $f\chi_A$ is integrable.

Proof: let $M > 0$ be such that $|f(x)| < M$ for all $x \in A$. Then, under they hypotheses,

$$\left| \int f\chi_A \right| \leq \int |f\chi_A| = \int |f|\chi_A < \int M\chi_A = M \int \chi_A = M \cdot \text{Area}(A) < \infty,$$

which completes the proof. ■

4. Let $A, B \in \mathcal{B}(\mathbb{R}^n)$, $A \cap B = \varnothing$, and f be a Borel function on $A \cup B$.

   a) If $f \geq 0$, show that
   $$\int_{A\cup B} f \,\mathrm{d}m = \int_A f \,\mathrm{d}m + \int_B f \,\mathrm{d}m.$$
   b) In general, show that f is integrable over $A \cup B$ if and only if f is integrable over A and integrable over B.
   c) If f is integrable over $A \cup B$, show that the equation of part a) holds.

Proof:

a) Let s_n be the sequence of positive simple functions guaranteed by one of the theorems. Then we have
   i. $s_n(z) \to f(z)$ for all z
   ii. $0 \leq s_n(z) \leq f(z)$ for all z
   iii. $s_n(z) \leq s_{n+1}(z)$ for all z

   Let $C \in \mathcal{B}$. Consider the function $f\chi_C$. Then,
   i. $(s_n\chi_C)(z) \to (f\chi_C)(z)$ for all z
   ii. $0 \leq (s_n\chi_C)(z) \leq (f\chi_C)(z)$ for all z
   iii. $(s_n\chi_C)(z) \leq (s_{n+1}\chi_C)(z)$ for all z

   According to the Lebesgue convergence theorem,
   $$\int_C s_n = \int s_n\chi_C \to \int f\chi_C = \int_C f. \tag{21.1}$$
   For any $n \in \mathbb{N}$, we have
   $$s\chi_{A\cup B} = s\chi_A + s\chi_B$$

since $A \cap B = \varnothing$. Then

$$\begin{aligned}\int_{A\cup B} s_n = \int s_n \chi_{A\cup B} \geq \int s_n \chi_{A\cup B} = \int (s_n\chi_A + s_n\chi_B) \\ = \int s_n\chi_A + \int s_n \chi_B = \int_A s_n + \int_B s_n.\end{aligned}$$

If we let $C = A \cup B$ in (21.1), we have

$$\int_{A\cup B} s_n \to \int_{A\cup B} f.$$

If we let $C = A$ in (21.1), we have

$$\int_A s_n \to \int_A f.$$

Finally, if we let $C = B$ in (21.1), we have

$$\int_B s_n \to \int_B f.$$

Combining all these results yields

$$\begin{array}{ccc} \displaystyle\int_A s_n + \int_B s_n & = & \displaystyle\int_{A\cup B} s_n \\ \Big\downarrow & & \Big\downarrow \\ \displaystyle\int_A f + \int_B f & & \displaystyle\int_{A\cup B} f \end{array}$$

so that we can conclude that

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

as limits are unique.

b) Suppose that f is a general (not necessarily positive) function, integrable over A and B, i.e.

$$\left|\int_A f\right|, \left|\int_B f\right| < \infty.$$

By a remark made in class, this also means that

$$0 \leq \int_A f_+, \int_A f_-, \int_B f_+, \int_B f_- < \infty.$$

Since f_- and f_+ are positive integrable Borel functions, we can apply part a) to obtain

$$\begin{aligned} 0 \leq \int_{A\cup B} f_+ = \int_A f_+ + \int_B f_+ < \infty \\ 0 \leq \int_{A\cup B} f_- = \int_A f_- + \int_B f_- < \infty \end{aligned}$$

so that f_+ and f_- are both integrable over $A \cup B$. Consequently, f is integrable over $A \cup B$.

Conversely, suppose that f is a general (not necessarily positive) function, integrable over $A \cup B$, i.e.

$$\left| \int_{A \cup B} f \right| < \infty.$$

By a remark made in class, this also means that

$$0 \leq \int_{A \cup B} f_+, \int_{A \cup B} f_- < \infty.$$

Since f_- and f_+ are positive integrable Borel functions, we can apply part a) to obtain

$$\begin{aligned} 0 &\leq \int_A f_+ + \int_B f_+ = \int_{A \cup B} f_+ < \infty \\ 0 &\leq \int_A f_- + \int_B f_- = \int_{A \cup B} f_- < \infty \end{aligned}$$

This implies that

$$0 \leq \int_A f_+, \int_A f_-, \int_B f_+, \int_B f_- < \infty$$

and so that f_+ and f_- are both integrable over A and over B. Consequently, f is integrable over A and over B.

c) Let us assume that f is a general (not necessarily positive) function, integrable over $A \cup B$ (and so also over A and over B, see part b). By construction,

$$\begin{aligned} \int_{A \cup B} f &= \int_{A \cup B} f_+ - \int_{A \cup B} f_- \\ &= \int_A f_+ + \int_B f_+ - \int_A f_- - \int_B f_- \\ &= \int_A f_+ - \int_A f_- + \int_B f_+ - \int_B f_- \\ &= \int_A f + \int_B f \end{aligned}$$

■

5. Show that the area of the circle $S^1 = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$ is zero.

Proof: we use the following intermediary result.

LEMMA: let $\varphi : [0, T] \to \mathbb{R}^2$ be continuous, with $T > 0$. If $\exists M > 0$ such that

$$\|\varphi(s) - \varphi(t)\|_\infty \leq M|s - t| \tag{21.2}$$

for all $s, t \in [0,T]$, then $\varphi([0,1])$ has $2D$ measure 0.

Proof: for all $N \geq 1$, let

$$0 = t_0 < t_1 < \cdots < t_N = 1, \quad t_i = \frac{i}{N}.$$

Recall that $\|\vec{x}\|_\infty = \max\{|x_1|, |x_2|\}$. Then, according to (21.2), $\varphi([t_{i-1}, t_i]) \subseteq I_i$ for some square I_i of length $\frac{2M}{N}$ (think about this for a second). Then, $\text{Area}(I_i) = \frac{4M^2}{N^2}$ and

$$\sum_{i=1}^{N} \text{Area}(I_i) = \frac{4M^2}{N}.$$

Now, let $\varepsilon > 0$ and select $N > \frac{4M^2}{\varepsilon}$. QED

Let $\varphi : [0, 2\pi] \to \mathbb{R}^2$ be defined by $\varphi(t) = (\cos t, \sin t)$. Then φ is continuous and $\varphi([0, 2\pi]) = S^1$. According to the mean value theorem,

$$\begin{aligned}\|\varphi(s) - \varphi(t)\|_\infty &\leq \max\{\sup_\eta |D\varphi_1(\eta)|, \sup_\eta |D\varphi_2(\eta)\}|s-t| \\ &\leq \max\{\sup_\eta |\sin\eta|, \sup_\eta |\cos\eta\}|s-t| \\ &\leq |s-t|\end{aligned}$$

We can then apply the preceding Lemma to obtain $\text{Area}(S^1) = 0$. ■

6. Show that if $f, g : \mathbb{R}^2 \to \mathbb{R}$ are Borel functions, then so is $f + g$.

Proof: let $d \in \mathbb{R}$. For any $r, s \in \mathbb{Q}$ such that $r + s < d$, we have

$$\{z \mid f(z) < r\} \cap \{z \mid g(z) < s\} \subseteq \{z \mid f(z) + g(z) < d\},$$

or

$$E_r^f \cap E_s^g \subseteq E_d^{f+g}.$$

Then

$$\bigcup_{\substack{r,s\in\mathbb{Q} \\ r+s<d}} \left(E_r^f \cap E_s^g\right) \subseteq E_d^{f+g}.$$

If $z_0 \in E_d^{f+g}$, i.e. if $f(z_0) + g(z_0) < d$, then $\exists r, s \in \mathbb{Q}$ such that $f(z_0) < r$, $g(z_0) < s$ and $r + s < d$ (because $\mathbb{Q}$ is dense in $\mathbb{R}$), so that $z_0 \in E_r^f \cap E_s^g$. Then

$$\bigcup_{\substack{r,s\in\mathbb{Q} \\ r+s<d}} \left(E_r^f \cap E_s^g\right) = E_d^{f+g}.$$

But f, g are Borel functions; as a result,$E_r^f, E_s^g \in \mathcal{B}$ for all $r, s \in \mathbb{Q}$. Since $\mathcal{B}$ is a $\sigma-$algebra,

$$E_d^{f+g} = \bigcup_{\substack{r,s\in\mathbb{Q} \\ r+s<d}} \left(E_r^f \cap E_s^g\right) \in \mathcal{B}$$

and $f + g$ is a Borel function. ■

7. Show that every countable subset of $\mathbb{R}^2$ has $2D$ measure zero.

   Proof: let $\varepsilon > 0$. List the elements of the countable subset as $A = \{a_1, a_2, \ldots, a_n, \ldots\}$. Let R_n be a square centered at a_n with Area$(R_n) = \frac{\varepsilon}{2^{n+1}}$. Then

$$\sum_{n\in\mathbb{N}} \text{Area}(R_n) = \sum_{n\in\mathbb{N}} \frac{\varepsilon}{2^{n+1}} = \frac{\varepsilon}{2}\sum_{n\in\mathbb{N}} \frac{1}{2^n} = \frac{\varepsilon}{2} < \varepsilon.$$

   Thus, Area$(A) = 0$. ■

8. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be defined by $f(x,y) = \sin(x)$ and set $A = [0, 2\pi] \times [0, 1]$. Compute $\int_A f \,\mathrm{d}m$.

   Solution: we have

$$\int_A f = \int f\chi_A = \int (f\chi_A)_+ - \int (f\chi_A)_- = \int f_+\chi_A - \int f_-\chi_A$$

   where

$$f_+(x,y)\chi_A(x,y) = \begin{cases} \sin x & \text{if } x \in [0,\pi] \\ 0 & \text{otherwise} \end{cases}$$
$$f_-(x,y)\chi_A(x,y) = \begin{cases} -\sin x & \text{if } x \in [\pi,2\pi] \\ 0 & \text{otherwise} \end{cases}$$

   Clearly, $\int f_+\chi_A = \int f_-\chi_A$, so that $\int_A f = 0$. □

9. Show that the set

$$\mathcal{I} = \{f : \mathbb{R}^n \to \mathbb{R} : f \text{ finite, Borel, integrable}\}$$

   is a vector space over $\mathbb{R}$.

   Proof: since $\mathcal{I}$ is a subset of the vector space of all functions from $\mathbb{R}^n$ to $\mathbb{R}$ over the scalar field $\mathbb{R}$, it suffices to verify that the three subspace conditions hold:

   a) $\mathcal{O} \in \mathcal{I}$: this is the case since the function defined by $\mathcal{O}(\mathbf{x}) = 0$ for all $\mathbf{x} \in \mathbb{R}^n$ is Borel as I was able to write it down, finite since $|\mathcal{O}(\mathbf{x})| = 0 < \infty$ for all $\mathbf{x} \in \mathbb{R}^n$, and integrable as $\int \mathcal{O} = 0 < \infty$.

   b) $f, g \in \mathcal{I} \implies f + g \in \mathcal{I}$: if f, g are Borel, finite and integrable, then $f + g$ is clearly Borel and finite. It is also clearly integrable, albeit I have to use Theorem 25 (in disguise) to show this:

$$-\infty < -\left|\int f\right| - \left|\int g\right| \leq \left|\int (f+g)\right| \leq \left|\int f\right| + \left|\int g\right| < \infty.$$

   Thus, $f + g \in \mathcal{I}$.

c) $f \in \mathcal{I}, \alpha \in \mathbb{R} \implies \alpha f \in \mathcal{I}$: if f is Borel, finite and integrable, and $\alpha \in \mathbb{R}$, then αf is clearly Borel and finite (since $|\alpha| \neq \infty$). It is also clearly integrable, albeit I have to use Theorem 25 (once again in disguise) to show this:

$$-\infty < -\alpha \left| \int f \right| \leq \left| \int \alpha f \right| \leq \alpha \left| \int f \right| < \infty.$$

Thus, $\alpha f \in \mathcal{I}$.

Consequently, $\mathcal{I}$ is a vector space. ■

10. Show that $I : \mathcal{I} \to \mathbb{R}$ defined by $I(f) = \int f \,\mathrm{d}m$ is a linear functional.

Proof: now that we know that $\mathcal{I}$ is a vector space over $\mathbb{R}$, it suffices to show that $I : \mathcal{I} \to \mathbb{R}$ acts linearly on $\mathcal{I}$, i.e. that

$$I(\alpha f + \beta g) = \alpha I(f) + \beta I(g)$$

for all $f, g \in \mathcal{I}$, $\alpha, \beta \in \mathbb{R}$.

But that is the content of Theorem 25 (since f, g are integrable):

$$\begin{aligned} I(\alpha f + \beta g) &= \int (\alpha f + \beta g) = \int (\alpha f) + \int (\beta g) \\ &= \alpha \int f + \beta \int g = \alpha I(f) + \beta I(g), \end{aligned}$$

which completes the proof. ■

11. Let $f : \mathbb{R} \to \mathbb{R}$ be defined by

$$f(x) = \begin{cases} 1 & \text{if } x \in \mathbb{Q} \\ 0 & \text{otherwise} \end{cases}$$

Is f integrable? If so, what value does $\int f \,\mathrm{d}m$ take? If not, where does the problem lie?

Proof: note that $f(x) \geq 0$ for all $x \in \mathbb{R}$ and $\mathbb{Q} \in \mathcal{B}(\mathbb{R})$ with $\text{Length}(\mathbb{Q}) = 0$. Thus,

$$\int_{\mathbb{R}} f = \int_{\mathbb{R}-\mathbb{Q}} f = \int_{\mathbb{R}-\mathbb{Q}} 0 = 0 < \infty$$

and f is integrable. ■

12. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be defined by $f(x, y) = x + y$ Is f integrable? If so, what value does $\int f \,\mathrm{d}m$ take? If not, where does the problem lie?

Proof: a function is integrable if and only if both its positive part and negative part are integrable. Here, $f_+, f_- : \mathbb{R}^2 \to \mathbb{R}$ are defined by

$$f_+(x, y) = \begin{cases} x + y & \text{if } y \geq -x \\ 0 & \text{else} \end{cases}$$

$$f_-(x, y) = \begin{cases} -x - y & \text{if } y \leq -x \\ 0 & \text{else} \end{cases}$$

Consider the positive simple functions

$$s_1(x,y) = \begin{cases} 1 & \text{if } x,y \geq 1 \\ 0 & \text{else} \end{cases}$$
$$s_2(x,y) = \begin{cases} 1 & \text{if } x,y \leq -1 \\ 0 & \text{else} \end{cases}$$

Then

$$0 \leq s_1(x,y) \leq f_+(x,y)$$
$$0 \leq s_2(x,y) \leq f_-(x,y)$$

for all $(x,y) \in \mathbb{R}^2$. Consequently,

$$0 \leq \int s_1 \leq \int f_+$$
$$0 \leq \int s_2 \leq \int f_-$$

But $\int s_1, \int s_2 = \infty$, so $\int f_+, \int f_- = \infty$ and f is not integrable, as neither its positive part nor its negative part is integrable. ■

13. Suppose that f is R-integrable over $[a,b]$. Is f integrable over $[a,b]$? What relation is there between $\int_{[a,b]} f \, \mathrm{d}m$ and $\int_a^b f(x)\,\mathrm{d}x$, if any?

Proof: if f is R-integrable over $[a,b]$, then on the one hand we have $\int_a^b f(x)\,\mathrm{d}x = \int_{[a,b]} f\,\mathrm{d}m$ and on the other hand we have $\infty > \left|\int_a^b f(x)\,\mathrm{d}x\right|$. Consequently, $\left|\int_{[a,b]} f\,\mathrm{d}m\right| < \infty$ and f is integrable over $[a,b]$. ■

14. Suppose that f is integrable over $[a,b]$. Is f R-integrable over $[a,b]$? What relation is there between $\int_{[a,b]} f\,\mathrm{d}m$ and $\int_a^b f(x)\,\mathrm{d}x$, if any?

Proof: there is no relation in this case. There are instances of integrable functions which are also R-integrable, such as $f : [0,1] \to \mathbb{R}$ defined by $f(x) = x^2$. Then

$$\int_{[0,1]} f\,\mathrm{d}m = \int_0^1 f(x)\,\mathrm{d}x = \frac{1}{3} < \infty.$$

But there are also instances of integrable functions which are not R-integrable.

Consider the function $f : [0,1] \to [0,\infty]$ defined by

$$f(x) = \begin{cases} \infty & x \in \mathbb{Q} \cap [0,1] \\ 0 & \text{else} \end{cases}.$$

We have seen that $\int_{[0,1]} f\,\mathrm{d}m = 0 < \infty$ so that f is integrable. We have also seen that $\int_0^1 f(x)\,\mathrm{d}x$ does not exist, so that it is not R-integrable.

The moral of the story: Lebesgue integration is more general than Riemann integration. But you already knew that. ■

21.8.2 Multivariate Calculus

1. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be independent of y, that is, there exists a function $g : \mathbb{R} \to \mathbb{R}$ such that $f(x, y) \equiv g(x)$ for all $(x, y) \in \mathbb{R}^2$.

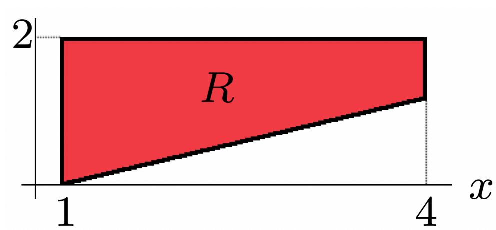
   a) What general property does the surface $z = f(x, y)$ possess?

   b) Let $R = \{(x, y) \mid a \leq x \leq b, c \leq y \leq d\}$. By interpreting the integral as a volume and by using the answer from part a), write $\int_R f \, \mathrm{d}A$ using a function of one variable.

   Solution: if f is independent of y, the surface $z = f(x, y)$ is constant in the $y-$direction, that is, for any $x \in \mathbb{R}$, $f(x, y_1) = f(x, y_2)$ for all y_1, y_2. As such,

$$\int_R f \, \mathrm{d}A = \left(\int_a^b g(x) \, \mathrm{d}x \right) (d - c). \qquad \square$$

2. Let $f : R \subset \mathbb{R}^2 \to \mathbb{R}$ be an integrable function and R be as below.



   Write $\int_R f \, \mathrm{d}A$ as an iterated integral.

   Solution: the vertices of R are: $(1, 0)$, $(2, 1)$, $(4, 2)$ and $(4, a)$, where $1 < a < 2$. The line from $(1, 2)$ to $(4, a)$ is $y = \frac{a}{3}(x - 1)$. Thus, R is the region defined by

$$\frac{a}{3}(x - 1) \leq y \leq 2, \quad 1 \leq x \leq 4,$$

   and $\int_1^4 \int_{\frac{a}{3}(x-1)}^2 f(x, y) \, \mathrm{d}y \, \mathrm{d}x$ is one way to write the iterated integral. $\square$

3. Compute the integral $\int_0^2 \int_0^x e^{x^2} \, \mathrm{d}y \, \mathrm{d}x$.

   Solution: the region of integration is given by

$$0 \leq y \leq x, \quad 0 \leq x \leq 2.$$

   As such, it is the triangle with vertices $(0, 0)$, $(2, 2)$ and $(2, 0)$ (we're not drawing it but you probably should). Thus,

$$\int_0^2 \int_0^x e^{x^2} \, \mathrm{d}y \, \mathrm{d}x = \int_0^2 \left[y e^{x^2} \right]_0^x \mathrm{d}x = \int_0^2 x e^{x^2} \, \mathrm{d}x = \left[\frac{1}{2} e^{x^2} \right]_0^2 = \frac{1}{2}(e^4 - 1). \qquad \square$$

4. Compute $\int_0^3 \int_{y^2}^9 y \sin(x^2) \, \mathrm{d}x \, \mathrm{d}y$.

Solution: the region of integration is

$$y^2 \leq x \leq 9, \quad 0 \leq y \leq 3.$$

Since it is difficult (read: impossible) to find an anti-derivative of $\sin(x^2)$ with respect to x, we change the order of integration. To do so cleanly, it suffices to notice that the region can be written as

$$0 \leq y \leq \sqrt{x}, \quad 0 \leq x \leq 9.$$

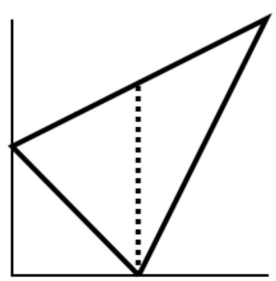
Thus,

$$\int_0^3 \int_{y^2}^9 y \sin(x^2)\, dx\, dy = \int_0^9 \int_0^{\sqrt{x}} y \sin(x^2)\, dy\, dx = \int_0^9 \left[\frac{y^2}{2}\sin(x^2)\right]_0^{\sqrt{x}} dx = \int_0^9 \frac{x}{2}\sin(x^2)\, dx$$
$$= \left[-\frac{1}{4}\cos(x^2)\right]_0^9 = \frac{1}{4}(1-\cos 81). \qquad \square$$

5. What is the volume of the solid bounded by the planes $z = x + 2y + 4$ and $z = 2x + y$, above the triangle in the xy plane with vertices $A(1,0,0)$, $B(2,1,0)$ and $C(0,1,0)$?

Solution: in the $xy-$plane, the equations of the boundary of ΔABC are

$$\begin{aligned} AC: &\quad y = -x + 1 \leftrightsquigarrow x = -y + 1 \\ BC: &\quad y = 1 \\ AB: &\quad y = x - 1 \leftrightsquigarrow x = y + 1 \end{aligned}$$



The region of integration R can be written as

$$0 \leq y \leq 1, \ -y + 1 \leq x \leq y + 1,$$

and the volume of interest is

$$V = \int_R |(x + 2y + 4) - (2x + y)|\, dA = \int_0^1 \int_{-y+1}^{y+1} (y - x + 4)\, dx\, dy = \int_0^1 \left[yx - \frac{x^2}{2} + 4x\right]_{-y+1}^{y+1} dy$$
$$= \int_0^1 \left[\left(y(y+1) - \frac{(y+1)^2}{2} + 4(y+1)\right) - \left(y(-y+1) - \frac{(-y+1)^2}{2} + 4(-y+1)\right)\right] dy$$
$$= \int_0^1 (2y^2 + 6y)\, dy = \left[\frac{2y^3}{3} + 3y\right]_0^1 = \frac{11}{3}. \qquad \square$$

6. Compute $\int_W h \; \mathrm{d}V$, where $h(x, y, z) = ax + by + cz$ and

$$W = \{(x, y, z) \mid 0 \leq x \leq 1, 0 \leq y \leq 1, 0 \leq z \leq 2\}.$$

Solution: the region of integration is rectangular, so there are no hardships:

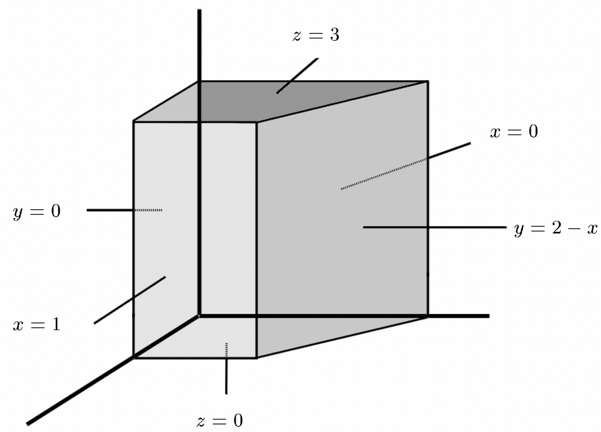
$$\begin{aligned} \int_W h \, \mathrm{d}V &= \int_0^1 \int_0^1 \int_0^2 (ax + by + cz) \, \mathrm{d}z \, \mathrm{d}x \, \mathrm{d}y = \int_0^1 \int_0^1 \left[axz + byz + c\frac{z^2}{2} \right]_0^2 \mathrm{d}x \, \mathrm{d}y \\ &= \int_0^1 \int_0^1 (2ax + 2by + 2c) \, \mathrm{d}x \, \mathrm{d}y = \int_0^1 \left[ax^2 + 2bxy + 2cx \right]_0^1 \mathrm{d}y \\ &= \int_0^1 (a + 2by + 2c) \, \mathrm{d}y = \left[ay + by^2 + 2cy \right]_0^1 = a + b + 2c. \qquad \square \end{aligned}$$

7. Sketch the region of integration W of the triple integral $\int_0^1 \int_0^{2-x} \int_0^3 f(x, y, z) \, \mathrm{d}z \, \mathrm{d}y \; \mathrm{d}x$.

Solution: the region is defined by

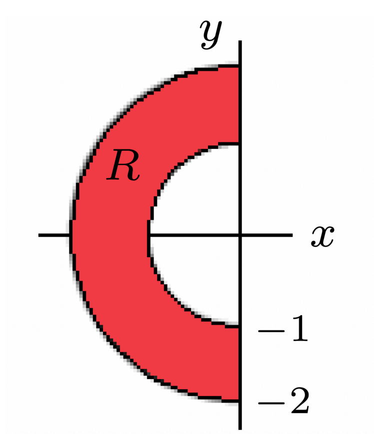
$$0 \leq z \leq 3, \quad 0 \leq y \leq 2 - x, \quad 0 \leq x \leq 1.$$

Thus, it is a box bounded by 6 planes: $z = 0$, $z = 3$, $y = 0$, $y = 2 - x$, $x = 0$, $x = 1$.



$\square$

8. Let $f : R \to \mathbb{R}$ be defined as below. Write $\int_R f \; \mathrm{d}A$ as an iterated integral.



Solution: in polar coordinates, the region becomes

$$1 \leq r \leq 2, \quad \frac{\pi}{2} \leq \theta \leq \frac{3\pi}{2}.$$

Thus,

$$\int_R f(x,y)\,\mathrm{d}A = \int_1^2 \int_{\pi/2}^{3\pi/2} f(r\cos\theta, r\sin\theta) r\,\mathrm{d}\theta\,dr. \qquad \square$$

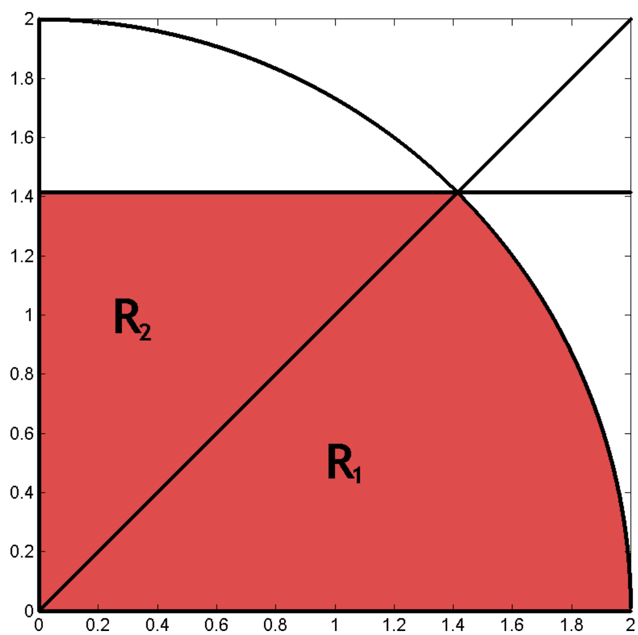
9. Compute $\int_0^{\sqrt{2}} \int_0^{\sqrt{4-y^2}} xy\,\mathrm{d}x\,\mathrm{d}y$.

Solution: The region of integration R is defined by

$$0 \leq x \leq \sqrt{4-y^2}, \quad 0 \leq y \leq \sqrt{2}.$$

We separate this region into two subregions R_1 and R_2 with the line $y = x$. Thus,

$$\int_R xy\,\mathrm{d}A = \int_{R_1} xy\,\mathrm{d}A + \int_{R_2} xy\,\mathrm{d}A.$$



The regions' geometry indicates that polar coordinates have to be used in the first region, while cartesian coordinates will be appropriate in the second region.
In polar coordinates, R_1 is

$$0 \leq r \leq 2, \quad 0 \leq \theta \leq \frac{\pi}{4},$$

whence

$$\int_{R_1} xy\,\mathrm{d}A = \int_0^2 \int_0^{\pi/4} (r\cos\theta)(r\sin\theta) r\,\mathrm{d}\theta\,dr = \int_0^2 \int_0^{\pi/4} r^3 \cos\theta \sin\theta\,\mathrm{d}\theta\,dr$$

$$= \int_0^2 \int_0^{\pi/4} \frac{r^3}{2} \sin 2\theta\,\mathrm{d}\theta\,dr = \int_0^2 \left[-\frac{r^3}{4}\cos 2\theta\right]_0^{\pi/4} \mathrm{d}\theta\,dr = \int_0^2 \frac{r^3}{4}\,dr = \left[\frac{r^4}{16}\right]_0^2 = 1.$$

In cartesian coordinates, R_2 is

$$0 \leq x \leq \sqrt{2}, \quad x \leq y \leq \sqrt{2},$$

whence

$$\begin{aligned}\int_{R_2} xy \,\mathrm{d}A &= \int_0^{\sqrt{2}} \int_x^{\sqrt{2}} xy \,\mathrm{d}y \,\mathrm{d}x = \int_0^{\sqrt{2}} \left[\frac{xy^2}{2}\right]_0^{\sqrt{2}} \mathrm{d}x = \int_0^{\sqrt{2}} \frac{x(2-x^2)}{2} \,\mathrm{d}x \\ &= \left[\frac{x^2}{2} - \frac{x^4}{8}\right]_0^{\sqrt{2}} = \frac{1}{2}.\end{aligned}$$

Thus, $\int_R xy \,\mathrm{d}A = \int_{R_1} xy \,\mathrm{d}A + \int_{R_2} xy \,\mathrm{d}A = 1 + \frac{1}{2} = \frac{3}{2}$. □

10. Compute $\int_W \sin(x^2 + y^2) \,\mathrm{d}V$, where W is the cylinder centered about the z axis from $z = -1$ to $z = 3$ and with radius 1.

    Solution: in cylindrical coordinates, W is

    $$0 \leq r \leq 1, \quad 0 \leq \theta \leq 2\pi, \quad -3 \leq z \leq 1.$$

    Thus,

    $$\int_W \sin(x^2 + y^2) \,\mathrm{d}V = \int_{-3}^{1} \int_0^{2\pi} \int_0^1 \sin(r^2) r \,dr \,\mathrm{d}\theta \,\mathrm{d}z = 4\pi(1 - \cos 1).$$ □

11. Using spherical coordinates, compute the triple integral of $f(\rho, \theta, \varphi) = \sin\varphi$ on the region defined by $0 \leq \theta \leq 2\pi$, $0 \leq \varphi \leq \frac{\pi}{4}$, $1 \leq \rho \leq 2$.

    Solution: in spherical coordinates, the region is

    $$0 \leq \theta \leq 2\pi, \quad 0 \leq \varphi \leq \frac{\pi}{4}, \quad 1 \leq \rho \leq 2.$$

    Thus, the integral is

    $$\begin{aligned} I &= \int_0^{2\pi} \int_0^{\pi/4} \int_1^2 \sin\varphi \rho^2 \sin\varphi \,\mathrm{d}\rho \,\mathrm{d}\varphi \,\mathrm{d}\theta = \int_0^{2\pi} \int_0^{\pi/4} \int_1^2 \rho^2 \sin^2\varphi \,\mathrm{d}\rho \,\mathrm{d}\varphi \,\mathrm{d}\theta \\ &= \int_0^{2\pi} \int_0^{\pi/4} \left[\frac{\rho^3}{3} \sin^2\varphi\right]_1^2 \mathrm{d}\varphi \,\mathrm{d}\theta = \int_0^{2\pi} \int_0^{\pi/4} \frac{7}{3} \sin^2\varphi \,\mathrm{d}\varphi \,\mathrm{d}\theta \\ &= \int_0^{2\pi} \frac{7}{6} \left[\varphi - \sin\varphi\cos\varphi\right]_0^{\pi/4} \,\mathrm{d}\theta = \int_0^{2\pi} \frac{7}{12}\left(\frac{\pi}{2} - 1\right) \mathrm{d}\theta = \frac{14\pi}{12}\left(\frac{\pi}{2} - 1\right) = \frac{7\pi}{6}(\pi - 1). \end{aligned}$$ □

12. Compute

    $$\int_0^1 \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \int_{-\sqrt{1-x^2-z^2}}^{\sqrt{1-x^2-z^2}} (x^2 + y^2 + z^2)^{-1/2} \,\mathrm{d}y \,\mathrm{d}z \,\mathrm{d}x.$$

    Solution: in spherical coordinates, the region of integration is

    $$-\frac{\pi}{2} \leq \theta \leq \frac{\pi}{2}, \quad 0 \leq \varphi \leq \pi, \quad 0 \leq \rho \leq 1.$$

Thus, the integral is

$$\begin{aligned} I &= \int_0^1 \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \int_{-\sqrt{1-x^2-z^2}}^{\sqrt{1-x^2-z^2}} (x^2+y^2+z^2)^{-1/2}\, \mathrm{d}y\, \mathrm{d}z\, \mathrm{d}x = \int_{-\pi/2}^{\pi/2} \int_0^{\pi} \int_0^1 \frac{1}{\sqrt{\rho^2}} \rho^2 \sin\varphi\, \mathrm{d}\rho\, \mathrm{d}\varphi\, \mathrm{d}\theta \\ &= \int_{-\pi/2}^{\pi/2} \int_0^{\pi} \int_0^1 \rho \sin\varphi\, \mathrm{d}\rho\, \mathrm{d}\varphi\, \mathrm{d}\theta = \int_{-\pi/2}^{\pi/2} \int_0^{\pi} \left[\frac{\rho^2}{2}\sin\varphi\right]_0^1 \mathrm{d}\varphi\, \mathrm{d}\theta = \int_{-\pi/2}^{\pi/2} \int_0^{\pi} \frac{\sin\varphi}{2}\, \mathrm{d}\varphi\, \mathrm{d}\theta \\ &= \int_{-\pi/2}^{\pi/2} \left[-\frac{\cos\varphi}{2}\right]_0^{\pi} \mathrm{d}\theta = \int_{-\pi/2}^{\pi/2} \mathrm{d}\theta = \pi. \qquad \square \end{aligned}$$

13. Compute

$$\int_0^1 \int_{-1}^1 \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} (x^2+y^2)^{-1/2}\, \mathrm{d}y\, \mathrm{d}x\, \mathrm{d}z.$$

Solution: in cylindrical coordinates, the region of integration is

$$0 \leq r \leq 1, \quad 0 \leq \theta \leq 2\pi, \quad 0 \leq z \leq 1.$$

In that case, the integral of interest is

$$\begin{aligned} I &= \int_0^1 \int_{-1}^1 \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} (x^2+y^2)^{-1/2}\, \mathrm{d}y\, \mathrm{d}x\, \mathrm{d}z = \int_0^1 \int_0^{2\pi} \int_0^1 \frac{1}{\sqrt{r^2}} r\, dr\, \mathrm{d}\theta\, \mathrm{d}z \\ &= \int_0^1 \int_0^{2\pi} \int_0^1 dr\, \mathrm{d}\theta\, \mathrm{d}z = 2\pi. \qquad \square \end{aligned}$$

14. Compute $\int_0^1 \int_{\sqrt{x}}^1 e^{y^3}\, \mathrm{d}y\, \mathrm{d}x$.

Solution: the region of integration is given by

$$0 \leq x \leq y^2, \quad 0 \leq y \leq 1.$$

Thus, the integral of interest is

$$\int_0^1 \int_{\sqrt{x}}^1 e^{y^3}\, \mathrm{d}y\, \mathrm{d}x = \int_0^1 \int_0^{y^2} e^{y^3}\, \mathrm{d}x\, \mathrm{d}y = \int_0^1 \left[x e^{y^3}\right]_{x=0}^{x=y^2} \mathrm{d}y = \int_0^1 y^2 e^{y^3}\, \mathrm{d}y = \left[\frac{e^{y^3}}{3}\right]_0^1 = \frac{e-1}{3}. \qquad \square$$

15. Sketch the solid bounded by the the surfaces $z=0$, $y=0$, $z=a-x+y$ and $y=a-\frac{1}{a}x^2$, where a is a positive constant. What is the volume of that solid?

Solution: the solid's base is the parabolic region in the $xy-$plane bounded by the line $y=0$ and the parabola $y=a-\frac{1}{a}x^2$. The volume of this solid is thus

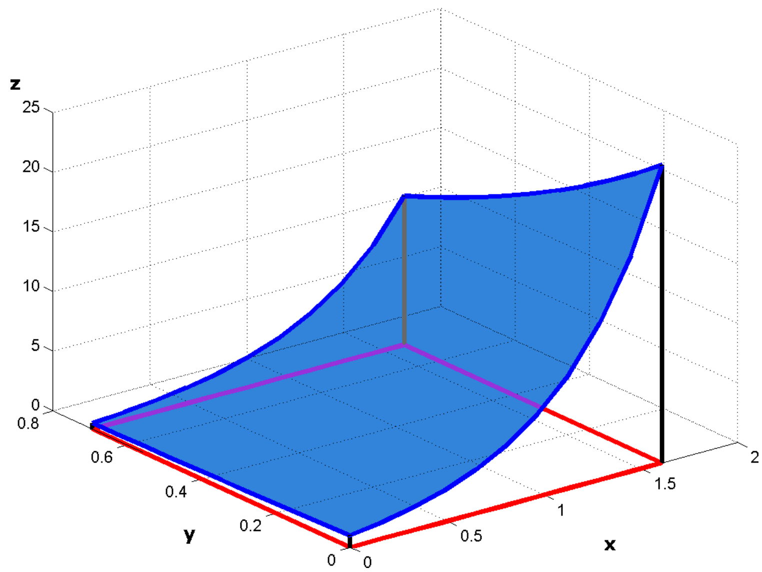
$$V = \iint_D (a-x+y)\, \mathrm{d}A = \iint_D (a+y)\, \mathrm{d}A,$$

(why can we eliminate the x in the integral?) so that

$$\begin{aligned} V &= \int_{-a}^{a} \int_0^{a-\frac{1}{a}x^2} (a+y)\, \mathrm{d}y\, \mathrm{d}x = \int_{-a}^{a} \left[ay + \frac{y^2}{2}\right]_{y=0}^{y=a-\frac{1}{a}x^2} \mathrm{d}x \\ &= 2\int_0^a \left(\frac{3}{2}a^2 - 2x^2 + \frac{x^4}{2a^2}\right) \mathrm{d}x = \left[3a^2x - \frac{4}{3}x^3 + \frac{1}{5a^2}x^5\right]_0^a = 3a^3 - \frac{4}{3}a^3 + \frac{1}{5}a^3 = \frac{28}{15}. \qquad \square \end{aligned}$$

16. Evaluate $\int_0^{\ln 2}\int_0^{\ln 5} e^{2x-y}\,\mathrm{d}x\,\mathrm{d}y$.

Solution: the region of integration appears in red R, while the surface $z = e^{2x-y}$ shows up in blue.

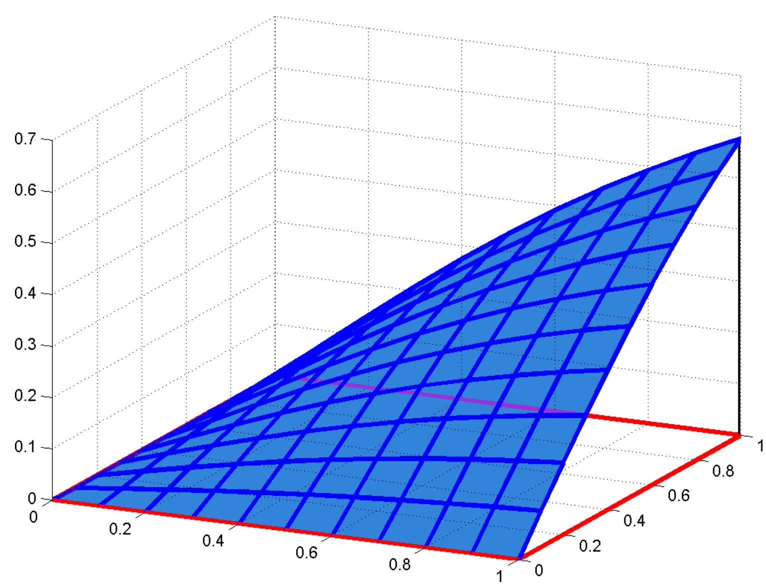


Since R is a rectangle, we can proceed directly:

$$\int_0^{\ln 2}\int_0^{\ln 5} e^{2x-y}\,\mathrm{d}x\,\mathrm{d}y = \int_0^{\ln 2}\left[\frac{1}{2}e^{2x-y}\right]_{x=0}^{x=\ln 5}\mathrm{d}y = \int_0^{\ln 2} 12e^{-y}\,\mathrm{d}y = \left[-12e^{-y}\right]_{y=0}^{y=\ln 2} = 6. \qquad \square$$

17. Evaluate $\int_0^1\int_0^1 \frac{xy}{\sqrt{x^2+y^2+1}}\,\mathrm{d}x\,\mathrm{d}y$.

Solution: the region of integration appears in red R, while the surface $z = \frac{xy}{\sqrt{x^2+y^2+1}}$ shows up in blue.

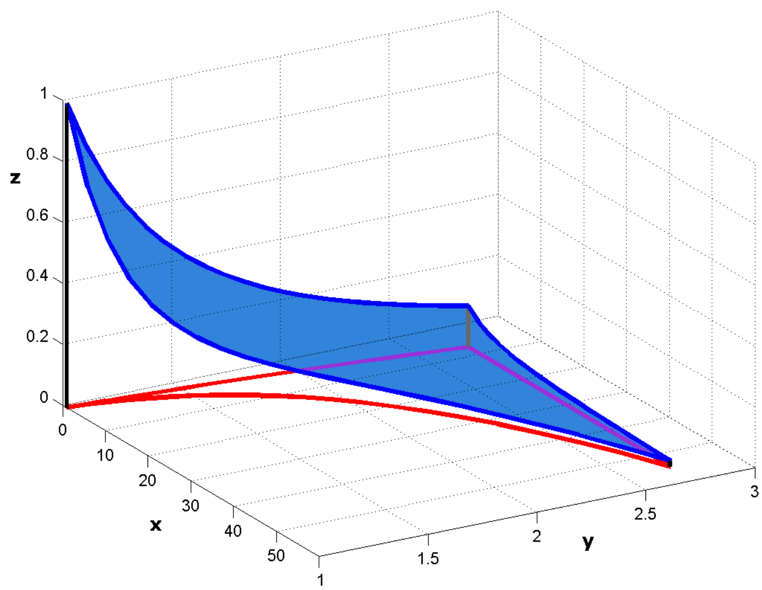


Since R is a rectangle, we can proceed directly:

$$\begin{aligned}
I &= \int_0^1 \int_0^1 \frac{xy}{\sqrt{x^2+y^2+1}}\,\mathrm{d}x\,\mathrm{d}y \\
&= \int_0^1 \left[y\sqrt{x^2+y^2+1}\right]_{x=0}^{x=1}\mathrm{d}y = \int_0^1 y\left[\sqrt{y^2+2}-\sqrt{y^2+1}\right]\mathrm{d}y \\
&= \int_0^1 y\sqrt{y^2+2}\,\mathrm{d}y - \int_0^1 y\sqrt{y^2+1}\,\mathrm{d}y = \left[\frac{1}{3}(y^2+2)^{3/2}\right]_{y=0}^{y=1} - \left[\frac{1}{3}(y^2+1)^{3/2}\right]_{y=0}^{y=1} \\
&= \sqrt{3} - \frac{4}{3}\sqrt{2} + \frac{1}{3}. \qquad \square
\end{aligned}$$

18. Let $D = \{(x,y) \mid 1 \leq y \leq e, y^2 \leq x \leq y^4\}$. Compute $\iint_D \frac{1}{x}\,\mathrm{d}A$.

Solution: the region of integration appears in red R, while the surface $z = \frac{1}{x}$ shows up in blue.

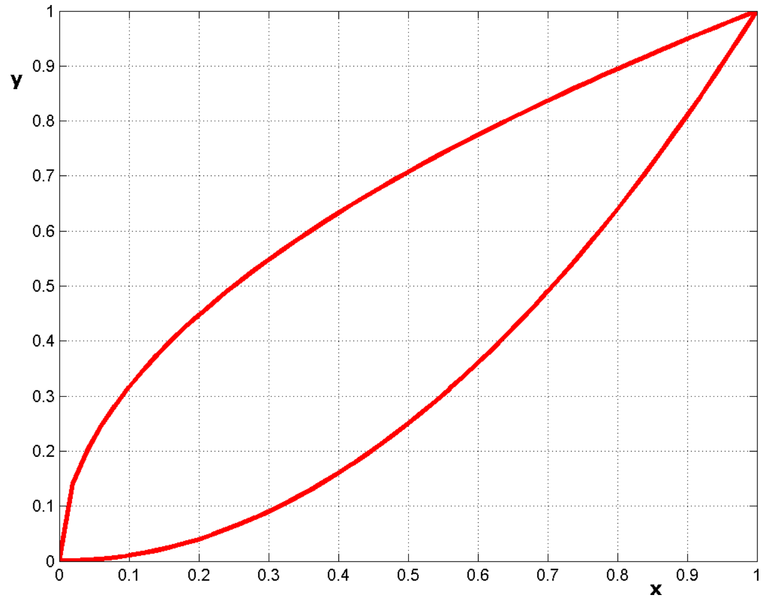


The double integral can be expressed as an iterated integral:

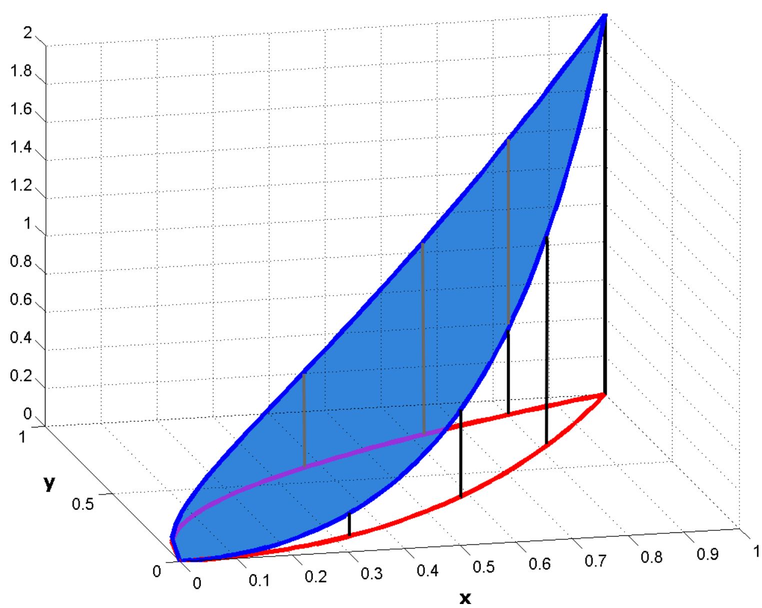
$$\begin{aligned}
\iint_D \frac{1}{x}\,\mathrm{d}A &= \int_1^e \int_{y^2}^{y^4} \frac{1}{x}\,\mathrm{d}x\,\mathrm{d}y = \int_1^e [\ln|x|]_{y^2}^{y^4}\,\mathrm{d}y = \int_1^e \left[\ln|y^4| - \ln|y^2|\right]\,\mathrm{d}y \\
&= \int_1^e \left[\ln|y^2|\right]\,\mathrm{d}y = \int_1^e \left[\ln y^2\right]\,\mathrm{d}y = 2\int_1^e \ln y\,\mathrm{d}y = 2\left[y\ln y - y\right]_1^e = 2. \qquad \square
\end{aligned}$$

19. What is the volume of the solid lying under the paraboloid $z = x^2 + y^2$ and above the domain bounded by $y = x^2$ and $x = y^2$?

Solution: the domain D is shown below:



Thus, $D = \{(x,y) \mid 0 \le x \le 1, x^2 \le y \le \sqrt{x}\}$ and the solid of interest is shown in the following figure:



Its volume is thus

$$V = \iint_D (x^2+y^2)\,\mathrm{d}A = \int_0^1 \int_{x^2}^{\sqrt{x}} (x^2+y^2)\,\mathrm{d}y\,\mathrm{d}x = \int_0^1 \left[x^2y + \frac{y^3}{3}\right]_{x^2}^{\sqrt{x}} \mathrm{d}x$$

$$= \int_0^1 \left[x^{5/2} - x^4 + \frac{x^{3/2}}{3} - \frac{x^6}{3}\right] \mathrm{d}x = \left[\frac{2}{7}x^{7/2} - \frac{x^5}{5} + \frac{2}{15}x^{5/2} - \frac{x^7}{21}\right]_0^1 = \frac{6}{35}. \qquad \square$$

20. Let R be the disk of radius 5, centered at the origin. Evaluate $\iint_R x\,\mathrm{d}A$.

Solution: in polar coordinates, R rewrites as

$$R_{(r,\theta)} = \{(r,\theta) \mid 0 \le r \le 5, 0 \le \theta \le 2\pi\}.$$

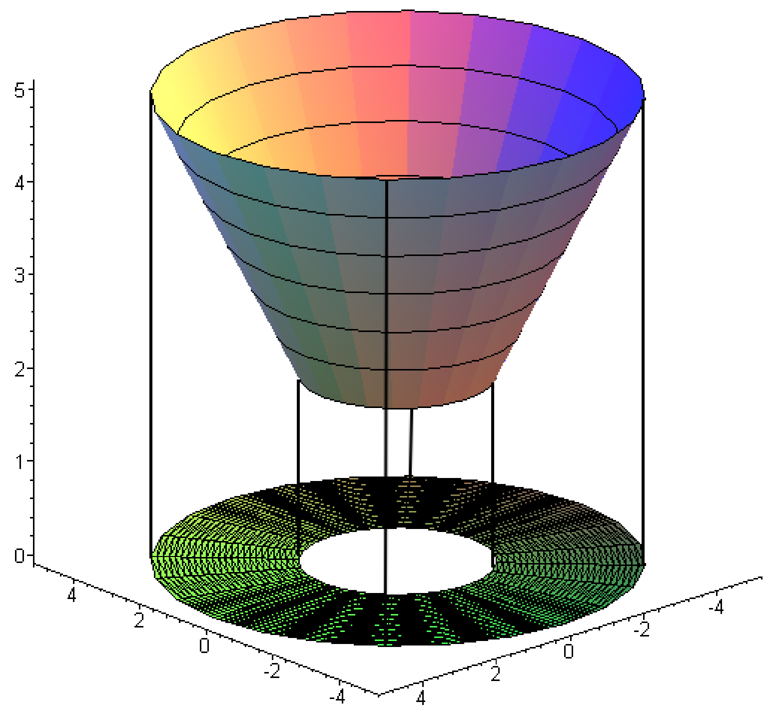
Since $x = r\cos\theta$, the change of variables formula yields

$$\iint_R x\,\mathrm{d}A = \int_0^5 \int_0^{2\pi} r\cos\theta \cdot r\,\mathrm{d}\theta\,dr = \int_0^5 \int_0^{2\pi} r^2\cos\theta\,\mathrm{d}\theta\,dr = \int_0^5 \left[r^2\sin\theta\right]_0^{2\pi}\,dr = 0.$$

Are you suprised by this result? You should not be. $\square$

21. What is the volume of the solid lying under the cone $z = \sqrt{x^2+y^2}$ and above the ring $4 \le x^2+y^2 \le 25$ located in the $xy-$plane?

Solution: the solid of interest is shown here:



If $R = \{(x,y) \mid 4 \le x^2+y^2 \le 25\}$, we wish to evaluate $\iint_R \sqrt{x^2+y^2}\,\mathrm{d}A$. In polar coordinates, we have

$$R_{(r,\theta)} = \{(r,\theta) \mid 2 \le r \le 5, 0 \le \theta \le 2\pi\}$$

and $\sqrt{x^2+y^2} = \sqrt{r^2} = r$, whence

$$\iint_R \sqrt{x^2+y^2}\,\mathrm{d}A = \int_2^5 \int_0^{2\pi} r \cdot r\,\mathrm{d}\theta\,dr = \int_2^5 \int_0^{2\pi} r^2\,\mathrm{d}\theta\,dr = \int_2^5 2\pi r^2\,dr = 78\pi. \quad \square$$

22. Compute $\int_0^2 \int_0^{\sqrt{2x-x^2}} \sqrt{x^2+y^2}\,\mathrm{d}y\,\mathrm{d}x$.

Solution: the region of integration is shown below:

In polar coordinates, this regions rewrites as

$$R_{(r,\theta)} = \{(r,\theta) : 0 \le \theta \le \pi/2, 0 \le r \le 2\cos\theta\},$$

whence the integral of interest is

$$I = \int_0^2 \int_0^{\sqrt{2x-x^2}} \sqrt{x^2+y^2}\, \mathrm{d}y\, \mathrm{d}x = \int_0^{\pi/2} \int_0^{2\cos\theta} \sqrt{r^2} \cdot r\, dr\, \mathrm{d}\theta = \int_0^{\pi/2} \left[\frac{r^3}{3}\right]_{r=0}^{r=2\cos\theta} \mathrm{d}\theta$$
$$= \int_0^{\pi/2} \left(\frac{8}{3}\cos^3\theta\right) \mathrm{d}\theta = \left[\frac{8}{9}\cos^2\theta\sin\theta + \frac{16}{9}\sin\theta\right]_0^{\pi/2} = \frac{16}{9}. \qquad \square$$

23. Find the mass and the centre of mass of the metal plate occupying the domain

$$D = \{(x,y) \mid 0 \le x \le 2,\ 0 \le y \le 3\},$$

if the density function of the plate is $\rho(x,y) = y$.

Solution: the total mass of the plate is $m = \iint_D \rho(x,y)\, \mathrm{d}A$, while the coordinates of the centre of mass $(\overline{x}, \overline{y})$ are given by

$$\overline{x} = \frac{1}{m}\iint_D x\rho(x,y)\, \mathrm{d}A \quad \text{and} \quad \overline{y} = \frac{1}{m}\iint_D y\rho(x,y)\, \mathrm{d}A.$$

Thus,

$$m = \int_0^2 \int_0^3 y\, \mathrm{d}y\, \mathrm{d}x = \int_0^2 \left[\frac{y^2}{2}\right]_0^3 \mathrm{d}x = \int_0^2 \frac{9}{2}\, \mathrm{d}x = 9$$
$$\overline{x} = \frac{1}{9}\int_0^2 \int_0^3 xy\, \mathrm{d}y\, \mathrm{d}x = \frac{1}{9}\int_0^2 \left[x\frac{y^2}{2}\right]_0^3 \mathrm{d}x = \frac{1}{9}\int_0^2 \frac{9}{2}x\, \mathrm{d}x = 1$$
$$\overline{y} = \frac{1}{9}\int_0^2 \int_0^3 y^2\, \mathrm{d}y\, \mathrm{d}x = \frac{1}{9}\int_0^2 \left[\frac{y^3}{3}\right]_0^3 \mathrm{d}x = \frac{1}{9}\int_0^2 9\, \mathrm{d}x = 2. \qquad \square$$

24. Evaluate $\int_0^3 \int_0^{\sqrt{9-x^2}} \int_0^x yz \, \mathrm{d}y \, \mathrm{d}z \, \mathrm{d}x$.

Solution: this can be done directly:

$$I = \int_0^3 \int_0^{\sqrt{9-x^2}} \int_0^x yz \, \mathrm{d}y \, \mathrm{d}z \, \mathrm{d}x = \int_0^3 \int_0^{\sqrt{9-x^2}} \left[\frac{y^2 z}{2}\right]_0^x \mathrm{d}z \, \mathrm{d}x = \int_0^3 \int_0^{\sqrt{9-x^2}} \frac{x^2 z}{2} \, \mathrm{d}z \, \mathrm{d}x$$
$$= \int_0^3 \left[\frac{x^2 z^2}{4}\right]_0^{\sqrt{9-x^2}} \mathrm{d}x = \int_0^3 \frac{x^2(9-x^2)}{4} \, \mathrm{d}x = \left[-\frac{x^5}{20} + \frac{3}{4}x^3\right]_0^3 = \frac{81}{10}. \qquad \square$$

25. Compute $\iiint_E e^x \, \mathrm{d}V$, where

$$E = \{(x,y,z) : 0 \leq y \leq 1, 0 \leq x \leq y, 0 \leq z \leq x+y\}.$$

Solution: again, this can be done directly, with the help of an iterated integral.

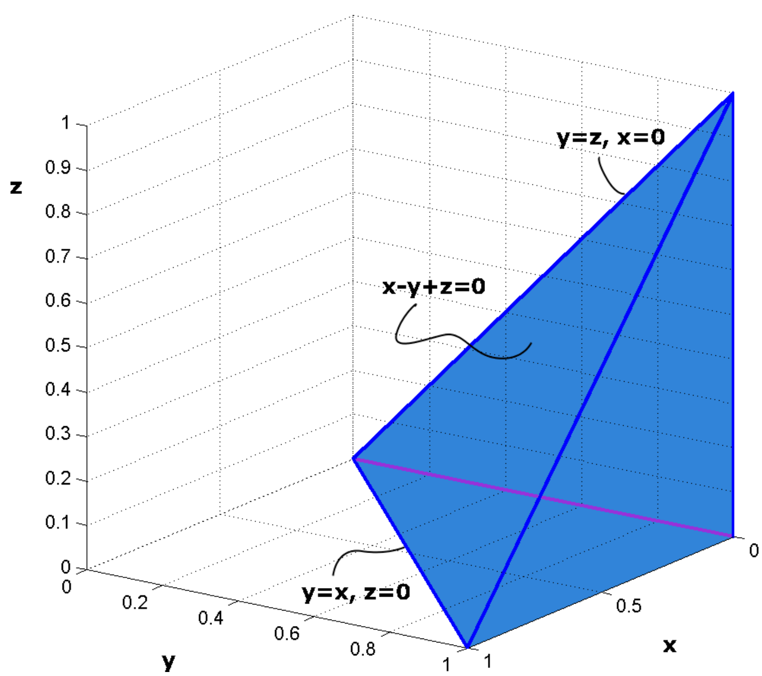
$$I = \int_0^1 \int_0^y \int_0^{x+y} e^x \, \mathrm{d}z \, \mathrm{d}x \, \mathrm{d}y = \int_0^1 \int_0^y [e^x z]_{z=0}^{z=x+y} \, \mathrm{d}x \, \mathrm{d}y = \int_0^1 \int_0^y e^x(x+y) \, \mathrm{d}x \, \mathrm{d}y$$
$$= \int_0^1 [e^x(x+y-1)]_{x=0}^{x=y} \, \mathrm{d}y = \int_0^1 (e^y-1)(y-1) \, \mathrm{d}y = \left[2ye^y - 3e^y + y - \frac{y^2}{2}\right]_0^1 = \frac{7}{2} - e. \qquad \square$$

26. Compute $\iiint_E xz \, \mathrm{d}V$, where E is the pyramid with vertices $(0,0,0)$, $(0,1,0)$, $(1,1,0)$ and $(0,1,1)$.

Solution: we can define E by

$$E = \{(x,y,z) \mid 0 \leq y \leq 1, 0 \leq z \leq y, 0 \leq x \leq y - z\},$$

as can be seen on the figure below.



Thus,

$$\iiint_E xz\,\mathrm{d}V = \int_0^1\int_0^y\int_0^{y-z} xz\,\mathrm{d}x\,\mathrm{d}z\,\mathrm{d}y = \int_0^1\int_0^y \frac{1}{2}(y-z)^2 z\,\mathrm{d}z\,\mathrm{d}y$$
$$= \frac{1}{2}\int_0^1 \left[\frac{1}{2}y^2z^2 - \frac{2}{3}yz^3 + \frac{1}{4}z^4\right]_{z=0}^{z=y} \mathrm{d}y = \frac{1}{24}\int_0^1 y^4\,\mathrm{d}y = \frac{1}{24}\left[\frac{1}{25}\right]_0^1 = \frac{1}{120}. \qquad \square$$

27. Let W be a three-dimensional solid. Its volume can be computed by the following iterated integral:

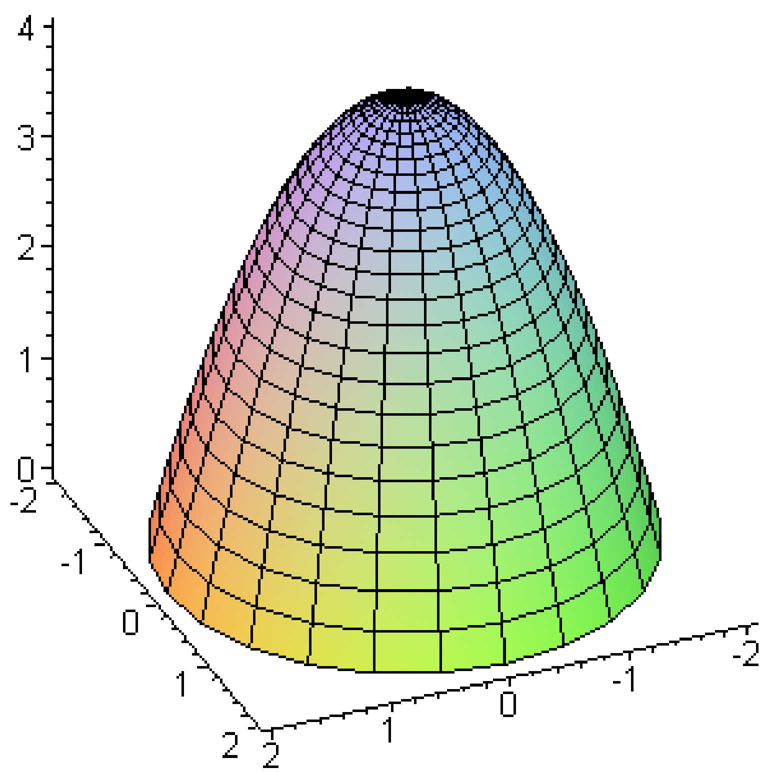
$$V(W) = \int_0^{2\pi}\int_0^2\int_0^{4-r^2} r\,\mathrm{d}z\,dr\,\mathrm{d}\theta.$$

Find W and $V(W)$.

Solution: in cartesian coordinates, $V(W) = \iiint_W \mathrm{d}V$. The volume integral is given in cylindrical coordinates, from which we can conclude that

$$W_{(r,\theta,z)} = \{(r,\theta,z) \mid 0 \le \theta \le 2\pi, 0 \le r \le 2, 0 \le z \le 4-r^2\}.$$

In cartesian coordinates, the solid of interest lies under the paraboloid $z = 4-x^2-y^2$ and above the disk in the $xy-$plane of radius 2 centered at the origin.



Thus,

$$\begin{aligned}
V(W) &= \int_{-2}^{-2}\int_{-\sqrt{4-x^2}}^{\sqrt{4-x^2}}\int_0^{4-x^2-y^2} \mathrm{d}z\,\mathrm{d}y\,\mathrm{d}x = \int_0^{2\pi}\int_0^2\int_0^{4-r^2} r\,\mathrm{d}z\,dr\,\mathrm{d}\theta \\
&= \int_0^{2\pi}\int_0^2 [rz]_{z=0}^{z=4-r^2}\,dr\,\mathrm{d}\theta = \int_0^{2\pi}\int_0^2 r(4-r^2)\,dr\,\mathrm{d}\theta \\
&= \int_0^{2\pi}\left[-\frac{r^4}{4}+2r^2\right]_0^2 \mathrm{d}\theta = 4\int_0^{2\pi}\mathrm{d}\theta = 8\pi. \qquad \square
\end{aligned}$$

28. Let W be a three-dimensional solid. Its volume can be computed by the following iterated integral:

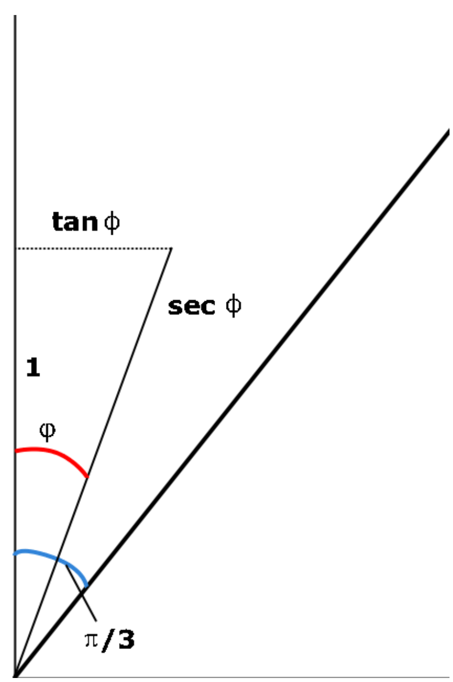
$$\int_0^{\pi/3}\int_0^{2\pi}\int_0^{\sec\varphi} \rho^2\sin\varphi\,\mathrm{d}\rho\,\mathrm{d}\theta\,\mathrm{d}\varphi.$$

Find W and $V(W)$.

Solution: in cartesian coordinates, $V(W)=\iiint_W \mathrm{d}V$. The volume integral is given in spherical coordinates, from which we can conclude that

$$W_{(\rho,\theta,\varphi)} = \{(\rho,\theta,\varphi)\mid 0\le\rho\le\pi/3, 0\le\theta\le 2\pi, 0\le\rho\le\sec\varphi\}.$$

Using the first two sets of inequalities, we see that the solid is part of the cone whose surface is $z=\frac{1}{\sqrt{3}}\sqrt{x^2+y^2}$ (in cartesian coordinates): when the radius is $\rho=\sec\varphi$, the height of the of the point in cartesian coordinates is automatically 1, as can be seen when we provide a transverse slice of the cone:



Thus, the volume of the cone is

$$\begin{aligned} V(W) &= \int_{-\sqrt{3}}^{\sqrt{3}} \int_{-\sqrt{3-x^2}}^{\sqrt{3-x^2}} \int_{\frac{1}{\sqrt{3}}\sqrt{x^2+y^2}}^{1} \mathrm{d}z\,\mathrm{d}y\,\mathrm{d}x = \int_0^{\pi/3} \int_0^{2\pi} \int_0^{\sec\varphi} \rho^2 \sin\varphi \,\mathrm{d}\rho\,\mathrm{d}\theta\,\mathrm{d}\varphi \\ &= \int_0^{\pi/3} \int_0^{2\pi} \left[\frac{1}{3}\rho^3 \sin\varphi\right]_{\rho=0}^{\rho=\sec(\varphi)} \mathrm{d}\theta\,\mathrm{d}\varphi = \int_0^{\pi/3} \left[\frac{1}{3}\sec^3\varphi \sin\varphi\theta\right]_{\theta=0}^{\theta=2\pi} \mathrm{d}\varphi \\ &= \int_0^{\pi/3} \frac{2\pi}{3} \sec^3\varphi \sin\varphi \,\mathrm{d}\varphi = \left[\frac{\pi}{3}\sec^2\varphi\right]_0^{\pi/3} = \pi, \end{aligned}$$

However, you do know how to compute the volume of a cone when the height and the radius are known: $V = \frac{1}{3}\pi r^2 h$. How does that compare to your answer? $\square$

29. Compute $\iiint_B (x^2+y^2+z^2)\,\mathrm{d}V$, where B is the unit ball $x^2+y^2+z^2 \leq 1$.

Solution: in spherical coordinates, the region can be written as

$$B_{(\rho,\theta,\varphi)} = \{(\rho,\theta,\varphi) \mid 0 \leq \rho \leq 1, 0 \leq \theta \leq 2\pi, 0 \leq \varphi \leq \pi\},$$

with $\rho^2 = x^2+y^2+z^2$, whence

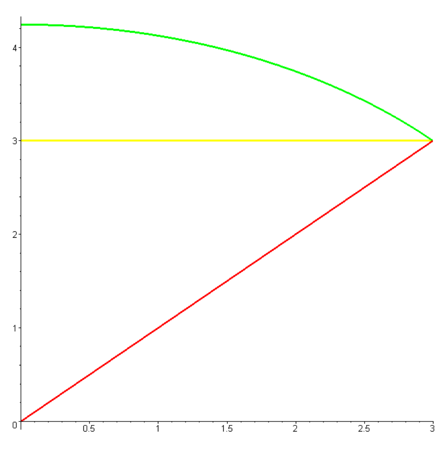
$$\begin{aligned} I &= \iiint_B (x^2+y^2+z^2)\,\mathrm{d}V = \int_0^1 \int_0^{2\pi} \int_0^{\pi} \rho^2 \cdot \rho^2 \sin\varphi \,\mathrm{d}\varphi\,\mathrm{d}\theta\,\mathrm{d}\rho \\ &= \int_0^1 \int_0^{2\pi} \int_0^{\pi} \rho^4 \sin\varphi \,\mathrm{d}\varphi\,\mathrm{d}\theta\,\mathrm{d}\rho = \int_0^1 \int_0^{2\pi} \left[-\rho^4\cos\varphi\right]_{\varphi=0}^{\varphi=\pi/3} \mathrm{d}\theta\,\mathrm{d}\rho \\ &= \int_0^1 \int_0^{2\pi} \frac{\rho^4}{2} \mathrm{d}\theta\,\mathrm{d}\rho = \int_0^1 \pi\rho^4 \,\mathrm{d}\rho = \pi\left[\frac{\rho^5}{5}\right]_0^1 = \frac{\pi}{5}. \qquad \square \end{aligned}$$

30. Evaluate

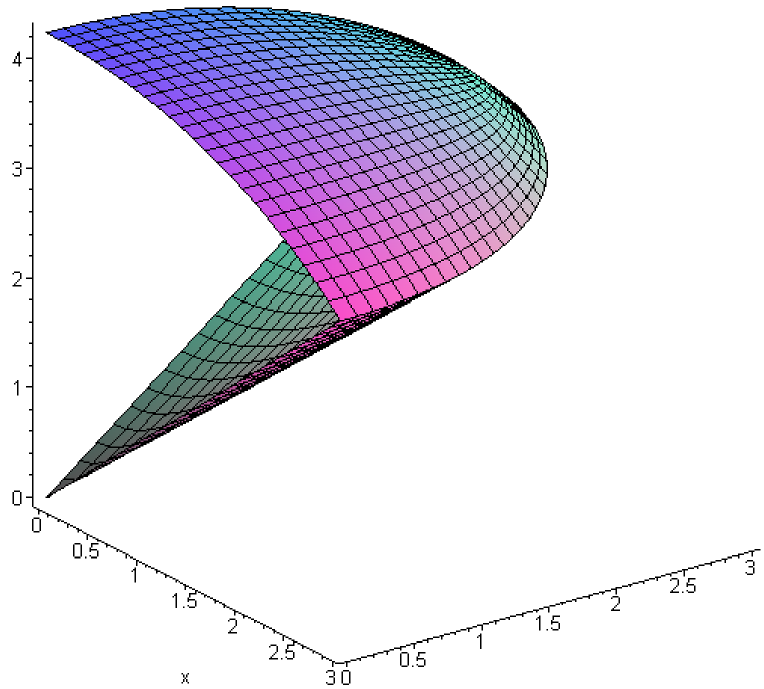
$$\int_0^3 \int_0^{\sqrt{9-y^2}} \int_{\sqrt{x^2+y^2}}^{\sqrt{18-x^2-y^2}} (x^2+y^2+z^2)\,\mathrm{d}z\,\mathrm{d}x\,\mathrm{d}y.$$

Solution: $\square$

31. **Solution:** the volume of integration is defined by the solid lying above the the disk of radius 3 in the first quadrant of the $xy-$plane and bounded by the cone $z^2 = x^2+y^2$ and the sphere $x^2+y^2+z^2 = 18$; as such, it is the solid of revolution of the following curve



around the z axis, under a rotation of $\frac{\pi}{2}$ radians:



In spherical coordinates, the region becomes

$$\left\{(\rho,\theta,\varphi) \mid 0 \leq \rho \leq \sqrt{18}, 0 \leq \theta \leq \frac{\pi}{2}, 0 \leq \varphi \leq \frac{\pi}{4}\right\}$$

with $\rho^2 = x^2+y^2+z^2$, whence

$$\begin{aligned}
I &= \int_0^3 \int_0^{\sqrt{9-y^2}} \int_{\sqrt{x^2+y^2}}^{\sqrt{18-x^2-y^2}} (x^2+y^2+z^2)\,\mathrm{d}z\,\mathrm{d}x\,\mathrm{d}y = \int_0^{\sqrt{18}} \int_0^{\pi/2} \int_0^{\pi/4} \rho^2 \cdot \rho^2 \sin\varphi\,\mathrm{d}\varphi\,\mathrm{d}\theta\,\mathrm{d}\rho \\
&= \int_0^{\sqrt{18}} \int_0^{\pi/2} \int_0^{\pi/4} \rho^4 \sin\varphi\,\mathrm{d}\varphi\,\mathrm{d}\theta\,\mathrm{d}\rho = \int_0^{\sqrt{18}} \int_0^{\pi/2} \left[-\rho^4\cos\varphi\right]_{\varphi=0}^{\varphi=\pi/4}\,\mathrm{d}\theta\,\mathrm{d}\rho \\
&= \int_0^{\sqrt{18}} \int_0^{\pi/2} [1-\cos(\pi/4)]\,\rho^4\,\mathrm{d}\theta\,\mathrm{d}\rho = \int_0^{\sqrt{18}} \frac{\pi}{2}(1-\cos(\pi/4))\rho^4\,\mathrm{d}\rho \\
&= \left[\frac{\pi}{2}(1-\cos(\pi/4))\frac{\rho^5}{5}\right]_0^{\sqrt{18}} = \frac{\pi}{2}(1-\cos(\pi/4))\frac{\sqrt{18}^5}{5}. \qquad \square
\end{aligned}$$

32. Compute the volume of the solid bounded by the cone $z=\sqrt{x^2+y^2}$ and the sphere of radius $a>0$ whose center is located at the origin.

Solution: let

$$A = \overline{B}(0,a) \cap \text{Cone} = \{(x,y,z) \mid x^2+y^2+z^2 \leq a^2 \text{ and } z \geq \sqrt{x^2+y^2}\}$$

If $(x,y,z) \in A$, then

$$x^2+y^2 \leq z^2 \leq a^2-(x^2+y^2),$$

whence $x^2 + y^2 \leq \frac{a^2}{2}$. Denote

$$C = \left\{(x, y) \mid x^2 + y^2 \leq \frac{a^2}{2}\right\}.$$

We then have

$$A = \{(x, y, z) : (x, y) \in C, \sqrt{x^2 + y^2} \leq z \leq \sqrt{a^2 - (x^2 + y^2)}\}$$

and so

$$\begin{aligned}\text{Vol}(A) &= \iiint_A \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \iint_C \left(\sqrt{a^2 - (x^2 + y^2)} - \sqrt{x^2 + y^2}\right)\,\mathrm{d}x\,\mathrm{d}y \\ &= \int_{[0, a/\sqrt{2}]} \int_{[-\pi, \pi]} \left(\sqrt{a^2 - r^2} - r\right) r\,\mathrm{d}\theta\,dr = \cdots = \frac{2\pi a^3}{3}\left(1 - \frac{1}{\sqrt{2}}\right). \qquad \square\end{aligned}$$

33. Compute the volume of the solid bounded by the paraboloïds $z = 10 - x^2 - y^2$ and $z = 2(x^2 + y^2 - 1)$.

Solution: let

$$A = \{(x, y, z) \mid 2(x^2 + y^2 - 1) \leq z \leq 10 - x^2 - y^2\}$$

If $(x, y, z) \in A$, then $x^2 + y^2 \leq 4$ (why?). Denote

$$B = \{(x, y) : x^2 + y^2 \leq 4\}.$$

We then have

$$A = \{(x, y, z) \mid (x, y) \in B, 2(x^2 + y^2 - 1) \leq z \leq 10 - x^2 - y^2\}$$

and so

$$\begin{aligned}\text{Vol}(A) &= \iiint_A \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \iint_B \left((10 - x^2 - y^2) - 2(x^2 + y^2 - 1)\right)\,\mathrm{d}x\,\mathrm{d}y \\ &= 3\iint_B \left(4 - (x^2 + y^2)\right)\,\mathrm{d}x\,\mathrm{d}y = 3\int_{[0,2]}\int_{[-\pi,\pi]} (4 - r^2) r\,\mathrm{d}\theta\,dr = \cdots = 24\pi. \qquad \square\end{aligned}$$

34. Let T be the triangle with vertices $(0, 0)$, $(0, 1)$ and $(1, 0)$. Compute $\iint_T \exp\left(\frac{y-x}{y+x}\right)\,\mathrm{d}x\,\mathrm{d}y$ using

a) polar coordinates;

b) the change of variables $u = y - x$, $v = y + x$.

Solution:

a) Let $x = r\cos\theta$, $y = r\sin\theta$. Then

$$\begin{aligned} I &= \iint_T \exp\left(\frac{y-x}{y+x}\right)\,\mathrm{d}x\,\mathrm{d}y \\ &= \int_{[0,\pi/2]}\int_{[0,(\sin\theta+\cos\theta)^{-1}]} \exp\left(\frac{\sin\theta-\cos\theta}{\sin\theta+\cos\theta}\right) r\,dr\,\mathrm{d}\theta \\ &= \int_{[0,\pi/2]} \exp\left(\frac{\sin\theta-\cos\theta}{\sin\theta+\cos\theta}\right)\left(\int_{[0,(\sin\theta+\cos\theta)^{-1}]} r\,dr\right)\mathrm{d}\theta \\ &= \frac{1}{2}\int_{[0,\pi/2]} \exp\left(\frac{\sin\theta-\cos\theta}{\sin\theta+\cos\theta}\right)\left(\frac{1}{\sin\theta+\cos\theta}\right)^2 \mathrm{d}\theta \end{aligned}$$

Set $t = \frac{\sin\theta-\cos\theta}{\sin\theta+\cos\theta}$. Then $dt = \frac{2}{(\sin\theta+\cos\theta)^2}\,\mathrm{d}\theta$ so that

$$I = \frac{1}{4}\int_{[-1,1]} \exp(t)\,dt = \frac{e-e^{-1}}{4}.$$

b) Let $y = \frac{1}{2}(u+v)$, $x = \frac{1}{2}(v-u)$. Then

$$I = \frac{1}{2}\iint_{T'} \exp\left(\frac{u}{v}\right)\,\mathrm{d}u\,\mathrm{d}V$$

where T' is the triangle in the $uv-$plane bounded by the points $(0,0)$, $(-1,1)$ and $(1,1)$. Then

$$I = \frac{1}{2}\int_{[0,1]}\int_{[-v,v]} \exp\left(\frac{u}{v}\right)\,\mathrm{d}u\,\mathrm{d}V = \cdots = \frac{e-e^{-1}}{4}. \qquad \square$$

35. Compute the area of the planar region bounded by $y = x^2$, $y = 2x^2$, $x = y^2$ and $x = 3y^2$.

Solution: denote the region in question by D and set $u = \frac{y}{x^2}$ and $v = \frac{x}{y^2}$. Then $(x,y) \in D$ if and only if $(u,v) \in R$, where R is the rectangle defined by $1 \leq u \leq 2$ and $1 \leq v \leq 3$. Let $\varphi : D \to R$ be defined by $\varphi(x,y) = (u,v) = (\frac{y}{x^2}, \frac{x}{y^2})$. Then we have

$$J_\varphi(x,y) = \det D\varphi(x,y) = \frac{3}{x^2y^2} = 3u^2v^2$$

and

$$\left|J_{\varphi^{-1}}(u,v)\right| = \frac{1}{|J_\varphi(x,y)|} = \frac{1}{3u^2v^2}.$$

Consequently,

$$\text{Area}(D) = \iint_D \mathrm{d}x\,\mathrm{d}y = \iint R\frac{1}{3u^2v^2}\,\mathrm{d}u\,\mathrm{d}V = \frac{1}{3}\int_{[1,2]}\int_{[1,3]} \frac{1}{v^2u^2}\,\mathrm{d}V\,\mathrm{d}u = \cdots = \frac{1}{9}. \qquad \square$$

36. For what values of $k \in \mathbb{R}$ does the integral

$$\iint_{x^2+y^2\leq 1} \frac{\mathrm{d}x\,\mathrm{d}y}{(x^2+y^2)^k}$$

converge? For each such k, find the value to which it converges.

Solution: first, note that

$$\iint_{x^2+y^2\leq 1} \frac{\mathrm{d}x\,\mathrm{d}y}{(x^2+y^2)^k} = \lim_{\varepsilon\to 0} \iint_{\varepsilon^2\leq x^2+y^2\leq 1} \frac{\mathrm{d}x\,\mathrm{d}y}{(x^2+y^2)^k}.$$

In polar coordinates, we have

$$\iint_{\varepsilon^2\leq x^2+y^2\leq 1} \frac{\mathrm{d}x\,\mathrm{d}y}{(x^2+y^2)^k} = \int_{[\varepsilon,1]} \int_{[0,2\pi]} \frac{1}{r^{2k-1}}\,\mathrm{d}\theta\,dr = 2\pi \int_{[\varepsilon,1]} \frac{dr}{r^{2k-1}}.$$

Then,

$$\lim_{\varepsilon\to 0} \int_{[\varepsilon,1]} \frac{dr}{r^{2k-1}}$$

if and only if $2k-1<1$, i.e. $k<1$. Furthermore,

$$\int_{[\varepsilon,1]} \frac{dr}{r^{2k-1}} = \frac{1}{2(1-k)} - \frac{\varepsilon^{2(1-k)}}{2(1-k)},$$

and so

$$\iint_{x^2+y^2\leq 1} \frac{\mathrm{d}x\,\mathrm{d}y}{(x^2+y^2)^k} = \frac{\pi}{1-k}$$

when $k<1$. $\square$

37. Find the volume of the solid bounded by the interior of the sphere $x^2+y^2+z^2=a^2$ and the interior of the cylinder $x^2+y^2=a^2$, $a>0$.

Solution: let V be the volume sought. Set

$$B = \{(x,y) \mid x^2+y^2 \leq a^2\}.$$

We have

$$\begin{aligned} V &= 2\iint_B \sqrt{2a^2-(x^2+y^2)}\,\mathrm{d}x\,\mathrm{d}y = 2\int_{[0,a]}\int_{[0,2\pi]} \sqrt{2a^2-r^2}\,\mathrm{d}\theta\,dr \\ &= 4\pi \int_{[0,a]} \sqrt{2a^2-r^2}\,r\,dr = \cdots = \frac{4\pi}{3}\left(2^{3/2}-1\right)a^3. \qquad \square \end{aligned}$$

38. Find the volume of the solid bounded by the interior of the cone $z^2=x^2+y^2$ lying above the paraboloïd $z=6-x^2-y^2$.

Solution: let V be the volume sought. Set

$$B = \{(x,y) \mid x^2+y^2 \leq 4\}.$$

We have

$$\begin{aligned} V &= 2\iint_B \left(6-(x^2+y^2)-\sqrt{x^2+y^2}\right)\,\mathrm{d}x\,\mathrm{d}y = \int_{[0,2]}\int_{[0,2\pi]} (6-r^2-r)r\,\mathrm{d}\theta\,dr \\ &= 2\pi \int_{[0,2]} (6-r^2-r)r\,\mathrm{d}\theta\,dr = \cdots = \frac{32\pi}{3}. \qquad \square \end{aligned}$$

39. Find the volume of the solid bounded by the plane $z = 3x+4y$ lying below the paraboloïd $z = x^2 + y^2$.

Solution: the intersection of the paraboloïd and the plane is $\{(x, y, z) \mid 3x + 4y = z = x^2 + y^2\}$. The set

$$D = \{(x, y) \mid 3x + 4y = x^2 + y^2\}$$

is the circle of radius $\frac{5}{2}$ centered at $(\frac{3}{2}, 2)$. For every $(x, y) \in D$, $x^2 + y^2 \leq 3x + 4y$. Let V be the volume sought. Set

$$B = \{(x, y) \mid x^2 + y^2 \leq 4\}.$$

We have

$$V = \iint_B \left(3x + 4y - (x^2 + y^2)\right) \, \mathrm{d}x \, \mathrm{d}y.$$

Using the change of variable

$$x = \frac{3}{2} + r \cos\theta, \quad y = 2 + r \sin\theta,$$

we obtain $V = \frac{1875\pi}{64}$. $\square$

21.9 Exercises

1. Prepare a 2-page summary of this chapter, with important definitions and results.
2. Let $\mathfrak{S}$ be a $\sigma-$algebra. Show that
   a) $A_1, A_2, \ldots, A_n, \ldots \in \mathfrak{S} \implies \bigcap_{n\geq 1} A_n \in \mathfrak{S}$;
   b) $A, B \in \mathfrak{S} \implies A \cap B^c \in \mathfrak{S}$, and
   c) $\varnothing, \mathbb{R}^n \in \mathfrak{S}$.
3. Complete the proof of Lemma 291.1.
4. Compute $\iint s_1(x, y) \, \mathrm{d}x \, \mathrm{d}y$ and $\iint s_2(x, y) \, \mathrm{d}x \, \mathrm{d}y$ in the example of Section 21.3.
5. In the example of Section 21.3, show that:
   a) for $1 \leq i \leq 2^n$, we have Area$(A_i^n) = \frac{1}{4^n}\left(i - \frac{1}{2}\right)$;
   b) for $2^n + 1 \leq i \leq 2^{n+1}$, we have Area$(A_i^n) = \frac{1}{4^n}\left(2^{n+1} - i - \frac{1}{2}\right)$.
6. Complete the proof of Corollary 294.
7. Is the converse of the third solved problem (Borel-Lebesgue integration on $\mathbb{R}^n$) true?
8. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be a Borel function and $d \in \mathbb{R}$. Show that $\{z \in \mathbb{R}^2 \mid f(z) < d\} \in \mathcal{B}(\mathbb{R}^2)$.
9. Complete the proof of Proposition 289 for $f + g$ and fg.

10. Show that if $g : \mathbb{R}^2 \to \mathbb{R}$ and

$$\{z \in \mathbb{R}^2 \mid g(z) < d\} \in \mathcal{B}(\mathbb{R}^2)$$

for all $d \in \mathbb{R}$, then g is a Borel function.

11. Show that $\mathbb{Q}^2$ is dense in $\mathbb{R}^2$ but that Area$(\mathbb{Q}^2) = 0$.

12. Show that $\mathcal{V}_n = \{f : \mathbb{R}^n \to \mathbb{R} \mid f$ finite, Borel, integrable$\}$ is a vector space and that the Borel-Lebesgue integral is a linear functional over $\mathcal{V}_n$.

13. Complete the proof of Theorem 301.

14. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be defined by $f(z) = \exp(-\|z\|^2)$. Find a sequence of simple functions

$$0 \leq s_1 \leq s_2 \leq \ldots \leq s_n \leq f$$

for which $s_n(z) \to f(z)$ for all $z \in \mathbb{R}^2$. Can you use the sequence to compute $\int f \, \mathrm{d}m$? If so, do so.

15. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be defined by

$$f(z) = \begin{cases} x^2 + y^2 & \text{if } (x, y) \in [0, 1] \times [0, 1] \\ 0 & \text{otherwise} \end{cases}$$

Find a sequence of simple functions

$$0 \leq s_1 \leq s_2 \leq \ldots \leq s_n \leq f$$

for which $s_n(z) \to f(z)$ for all $z \in \mathbb{R}^2$. Can you use the sequence to compute $\int f \, \mathrm{d}m$? If so, do so.

16. Let $f : \mathbb{R}^2 \to \mathbb{R}$ be defined by

$$f(z) = \begin{cases} x^2 + y^2 & \text{if } x^2 + y^2 \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

Find a sequence of simple functions

$$0 \leq s_1 \leq s_2 \leq \ldots \leq s_n \leq f$$

for which $s_n(z) \to f(z)$ for all $z \in \mathbb{R}^2$. Can you use the sequence to compute $\int f \, \mathrm{d}m$? If so, do so.

17. Let $f : \mathbb{R}^3 \to \mathbb{R}$ be defined by

$$f(z) = \begin{cases} x + y + z & \text{if } (x, y, z) \in [0, 1] \times [0, 1] \times [0, 1] \\ 0 & \text{otherwise} \end{cases}$$

Find a sequence of simple functions

$$0 \leq s_1 \leq s_2 \leq \ldots \leq s_n \leq f$$

for which $s_n(z) \to f(z)$ for all $z \in \mathbb{R}^3$. Can you use the sequence to compute $\int f \, \mathrm{d}m$? If so, do so.

18. Give a proof of the Lebesgue monotone convergence theorem.

19. Prove Theorem 300.

20. Show that $\Psi(r, \theta) = (x, y)$ is a diffeomorphism between U and V for polar coordinates.

21. Show that $|J_\Psi(r, \varphi, \theta)| = r^2 \sin \varphi$ for spherical coordinates.

22. What is the volume of the solid defined by the intersection of the two cylinders $x^2 + z^2 = 1$ and $y^2 + z^2 = 1$?

23. What is the volume of the solid Q directly above the region bounded by $0 \leq x \leq 1$, $1 \leq y \leq 2$ in the $xy-$plane and below the plane $z = 4 - x - y$?

24. Evaluate the integral $\iint_D x^2 y \, \mathrm{d}x \, \mathrm{d}y$ where D is the region bounded by the curves $y = x^2$ and $x = y^2$ in the first quadrant.

25. Let $f, f_1 : I \to \mathbb{R}$ be two continuous functions for which $f_1 \leq f$. If

$$A = \{(x, y) \in \mathbb{R}^2 \mid f_1(x) \leq y \leq f(x)\},$$

show that

$$\iint \chi_A(x, y) \, \mathrm{d}x \, \mathrm{d}y = \int_I (f_1(x) - f(x)) \, \mathrm{d}x.$$

Can you use this result to show that

$$\text{Graph}(f) = \{(x, f(x)) \mid x \in I\}$$

has $2D$ measure 0?

26. The Gamma and Beta functions are defined by

$$\Gamma(x) = \int_0^\infty t^{x-1} e^{-t} \, dt, \quad \text{for } x > 0$$
$$B(x, y) = \int_0^1 t^{x-1} (1 - t)^{y-1} \, dt, \quad \text{for } x > 0, y > 0$$

Show that the following properties hold:

a) $\Gamma(x + 1) = x\Gamma(x), \quad (x > 0);$
b) $\Gamma(n + 1) = n!, \quad (n = 0, 1, 2, \ldots);$

c) $\Gamma(x) = 2\int_0^\infty s^{2x-1}e^{-s^2}\,ds, \quad (x>0)$;

d) $\Gamma(\frac{1}{2}) = \sqrt{\pi}$, $\quad \Gamma(\frac{3}{2}) = \frac{\sqrt{\pi}}{2}$;

e) $B(x,y) = 2\int_0^{\pi/2} \cos^{2x-1}\theta \sin^{2y-1}\theta \,\mathrm{d}\theta$, $\quad (x>0, y>0)$;

f) $B(x,y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}$, $\quad (x>0, y>0)$.

27. Find the volume of the solid bounded by the interior of each of the cylinders $x^2+y^2=a^2$, $x^2+z^2=a^2$ and $y^2+z^2=a^2$, $a>0$.

28. Let S be the sphere of radius $a>0$ centered at $(0,0,a)$. Show that $\iiint_S z^2\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \frac{8}{5}\pi a^5$.

29. Compute $\iiint e^{-(x^2+y^2+z^2)}\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z$.

30. Show that $S^1 = \{(x,y) \mid x^2+y^2=1\}$ has $2D$ measure 0.

31. Show that every countable subset of $\mathbb{R}^2$ has $2D$ measure 0.

Chapter 22

Complex Analysis Fundamentals

Coming soon.

Chapter 23

Stone-Weierstrass Theorem

Coming soon.

Chapter 24

Baire's Theorem

Coming soon.

Chapter 25

Hale's Theorem

Coming soon.

Chapter 26

Basics of Functional Analysis

Coming soon.

Chapter 27

A Classical Hilbert Space Example

Coming soon.