Claudio Canuto · Anita Tabacco

Mathematical Analysis II

UNITEXT

Second Edition

Springer

UNITEXT – La Matematica per il 3+2

Volume 85

For further volumes:
http://www.springer.com/series/5418

Claudio Canuto · Anita Tabacco

Mathematical Analysis II

Second Edition

Springer

Claudio Canuto
Department of Mathematical Sciences
Politecnico di Torino
Torino, Italy

Anita Tabacco
Department of Mathematical Sciences
Politecnico di Torino
Torino, Italy

UNITEXT – La Matematica per il 3+2
ISSN 2038-5722 ISSN 2038-5757 (electronic)

ISBN 978-3-319-12756-9 ISBN 978-3-319-12757-6 (eBook)
DOI 10.1007/978-3-319-12757-6
Springer Cham Heidelberg New York Dordrecht London

Library of Congress Control Number: 2014952083

Cover Design: Simona Colombo, Giochi di Grafica, Milano, Italy
Files provided by the Authors

Springer is a part of Springer Science+Business Media (www.springer.com)

To Arianna, Susanna and Chiara, Alessandro, Cecilia

Preface

The purpose of this textbook is to present an array of topics that are found in the syllabus of the typical second lecture course in Calculus, as offered in many universities. Conceptually, it follows our previous book *Mathematical Analysis I*, published by Springer, which will be referred to throughout as Vol. I.

While the subject matter known as ‘Calculus 1’ concerns real functions of real variables, and as such is more or less standard, the choices for a course on ‘Calculus 2’ can vary a lot, and even the way the topics can be taught is not so rigid. Due to this larger flexibility we tried to cover a wide range of subjects, reflected in the fact that the amount of content gathered here may not be comparable to the number of credits conferred to a second Calculus course by the current programme specifications. The reminders disseminated in the text render the sections more independent from one another, allowing the reader to jump back and forth, and thus enhancing the book’s versatility.

The succession of chapters is what we believe to be the most natural. With the first three chapters we conclude the study of one-variable functions, begun in Vol. I, by discussing sequences and series of functions, including power series and Fourier series. Then we pass to examine multivariable and vector-valued functions, investigating continuity properties and developing the corresponding integral and differential calculus (over open measurable sets of $\mathbb{R}^n$ first, then on curves and surfaces). In the final part of the book we apply some of the theory learnt to the study of systems of ordinary differential equations.

Continuing along the same strand of thought of Vol. I, we wanted the presentation to be as clear and comprehensible as possible. Every page of the book concentrates on very few essential notions, most of the time just one, in order to keep the reader focused. For theorems’ statements, we chose the form that hastens an immediate understanding and guarantees readability at the same time. Hence, they are as a rule followed by several examples and pictures; the same is true for the techniques of computation.

The large number of exercises, gathered according to the main topics at the end of each chapter, should help the student test his improvements. We provide the solution to all exercises, and very often the procedure for solving is outlined.

Some graphical conventions are adopted: definitions are displayed over grey backgrounds, while statements appear on blue; examples are marked with a blue vertical bar at the side; exercises with solutions are boxed (e.g., $\boxed{12.}$).

This second edition is enriched by two appendices, devoted to differential and integral calculus, respectively. Therein, the interested reader may find the rigorous explanation of many results that are merely stated without proof in the previous chapters, together with useful additional material. We completely omitted the proofs whose technical aspects prevail over the fundamental notions and ideas. These may be found in other, more detailed, texts, some of which are explicitly suggested to deepen relevant topics.

All figures were created with MATLAB™ and edited using the freely-available package `psfrag`.

This volume originates from a textbook written in Italian, itself an expanded version of the lecture courses on Calculus we have taught over the years at the Politecnico di Torino. We owe much to many authors who wrote books on the subject: A. Bacciotti and F. Ricci, C. Pagani and S. Salsa, G. Gilardi to name a few. We have also found enduring inspiration in the Anglo-Saxon-flavoured books by T. Apostol and J. Stewart.

Special thanks are due to Dr. Simon Chiossi, for the careful and effective work of translation.

Finally, we wish to thank Francesca Bonadei – Executive Editor, Mathematics and Statistics, Springer Italia – for her encouragement and support in the preparation of this textbook.

Torino, August 2014 *Claudio Canuto, Anita Tabacco*

Contents

1

Numerical series

This is the first of three chapters dedicated to series. A series formalises the idea of adding infinitely many terms of a sequence which can involve numbers (numerical series) or functions (series of functions). Using series we can represent an irrational number by the sum of an infinite sequence of increasingly smaller rational numbers, for instance, or a continuous map by a sum of infinitely many piecewise-constant functions defined over intervals of decreasing size. Since the definition itself of series relies on the notion of limit of a sequence, the study of a series' behaviour requires all the instruments used for such limits.

In this chapter we will consider numerical series: beside their unquestionable theoretical importance, they serve as a warm-up for the ensuing study of series of functions. We begin by recalling their main properties. Subsequently we will consider the various types of convergence conditions of a numerical sequence and identify important classes of series, to study which the appropriate tools will be provided.

1.1 Round-up on sequences

We briefly recall here the definition and main properties of sequences, whose full treatise is present in Vol. I.

A real **sequence** is a function from $\mathbb{N}$ to $\mathbb{R}$ whose domain contains a set $\{n \in \mathbb{N} \ : \ n \geq n_0\}$ for some integer $n_0 \geq 0$. If one calls a the sequence, it is common practice to denote the image of n by a_n instead of $a(n)$; in other words we will write $a : n \mapsto a_n$. A standardised way to indicate a sequence is $\{a_n\}_{n \geq n_0}$ (ignoring the terms with $n < n_0$), or even more concisely $\{a_n\}$.

The behaviour of a sequence as n tends to ∞ can be classified as follows. The sequence $\{a_n\}_{n \geq n_0}$ **converges** (**to** ℓ) if the limit $\lim\limits_{n \to \infty} a_n = \ell$ exists and is finite. When the limit exists but is infinite, the sequence is said to **diverge to** $+\infty$ or $-\infty$. If the sequence is neither convergent nor divergent, i.e., if $\lim\limits_{n \to \infty} a_n$ does not exist, the sequence is **indeterminate**.

C. Canuto, A. Tabacco: *Mathematical Analysis II*, 2nd Ed.,
UNITEXT – La Matematica per il 3+2 85, DOI 10.1007/978-3-319-12757-6_1,

The fact that the behaviour of the first terms is completely irrelevant justifies the following definition. A sequence $\{a_n\}_{n \geq n_0}$ satisfies a certain property **eventually**, if there is an integer $N \geq n_0$ such that the sequence $\{a_n\}_{n \geq N}$ satisfies the property.

The main theorems governing the limit behaviour of sequences are recalled below.

Theorems on sequences

1. *Uniqueness of the limit*: if the limit of a sequence exists, it is unique.
2. *Boundedness*: a converging sequence is bounded.
3. *Existence of the limit for monotone sequences*: an eventually-monotone sequence is convergent if bounded, divergent if unbounded (divergent to $+\infty$ if increasing, to $-\infty$ if decreasing).
4. *First Comparison Theorem*: let $\{a_n\}$, $\{b_n\}$ be sequences with finite or infinite limits $\lim\limits_{n\to\infty} a_n = \ell$ and $\lim\limits_{n\to\infty} b_n = m$. If $a_n \leq b_n$ eventually, then $\ell \leq m$.

5a. *Second Comparison Theorem - finite case*: let $\{a_n\}$, $\{b_n\}$ and $\{c_n\}$ be sequences with $\lim\limits_{n\to\infty} a_n = \lim\limits_{n\to\infty} c_n = \ell \in \mathbb{R}$. If $a_n \leq b_n \leq c_n$ eventually, then $\lim\limits_{n\to\infty} b_n = \ell$.

5b. *Second Comparison Theorem - infinite case*: let $\{a_n\}$, $\{b_n\}$ be sequences such that $\lim\limits_{n\to\infty} a_n = +\infty$. If $a_n \leq b_n$ eventually, then $\lim\limits_{n\to\infty} b_n = +\infty$. A similar result holds if the limit is $-\infty$: $\lim\limits_{n\to\infty} b_n = -\infty$ implies $\lim\limits_{n\to\infty} a_n = -\infty$.

6. *Property*: a sequence $\{a_n\}$ is infinitesimal, that is $\lim\limits_{n\to\infty} a_n = 0$, if and only if the sequence of absolute values $\{|a_n|\}$ is infinitesimal.
7. *Theorem*: if $\{a_n\}$ is infinitesimal and $\{b_n\}$ bounded, $\{a_n b_n\}$ is infinitesimal.
8. *Algebra of limits*: let $\{a_n\}$, $\{b_n\}$ be such that $\lim\limits_{n\to\infty} a_n = \ell$ and $\lim\limits_{n\to\infty} b_n = m$ (ℓ, m finite or infinite). Then
$$\lim_{n\to\infty} (a_n \pm b_n) = \ell \pm m\,,$$
$$\lim_{n\to\infty} a_n\, b_n = \ell\, m\,,$$
$$\lim_{n\to\infty} \frac{a_n}{b_n} = \frac{\ell}{m}, \quad \text{if } b_n \neq 0 \text{ eventually,}$$
whenever the right-hand sides are defined.
9. *Substitution Theorem*: let $\{a_n\}$ be a sequence with $\lim\limits_{n\to\infty} a_n = \ell$ and suppose g is a function defined on a neighbourhood of ℓ:
   a) if $\ell \in \mathbb{R}$ and g is continuous at ℓ, then $\lim\limits_{n\to\infty} g(a_n) = g(\ell)$;
   b) if $\ell \notin \mathbb{R}$ and $\lim\limits_{x\to\ell} g(x) = m$ exists, then $\lim\limits_{n\to\infty} g(a_n) = m$.

10. *Ratio Test*: let $\{a_n\}$ be a sequence for which $a_n > 0$ eventually, and suppose the limit
$$\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = q$$
exists, finite or infinite. If $q < 1$ then $\lim\limits_{n\to\infty} a_n = 0$; if $q > 1$ then $\lim\limits_{n\to\infty} a_n = +\infty$.

Let us review some cases of particular relevance.

Examples 1.1

i) Consider the *geometric sequence* $a_n = q^n$, where q is a fixed number in $\mathbb{R}$. In Vol. I, Example 5.18, we proved

$$\boxed{\lim_{n\to\infty} q^n = \begin{cases} 0 & \text{if } |q| < 1, \\ 1 & \text{if } q = 1, \\ +\infty & \text{if } q > 1, \\ \text{does not exist} & \text{if } q \le -1. \end{cases}} \tag{1.1}$$

ii) Let $p > 0$ be a given number and consider the sequence $\sqrt[n]{p}$. Using the Substitution Theorem with $g(x) = p^x$ we have

$$\boxed{\lim_{n\to\infty} \sqrt[n]{p} = \lim_{n\to\infty} p^{1/n} = p^0 = 1\,.}$$

iii) Consider now $\sqrt[n]{n}$; again using the Substitution Theorem,

$$\boxed{\lim_{n\to\infty} \sqrt[n]{n} = \lim_{n\to\infty} \exp \frac{\log n}{n} = \mathrm{e}^0 = 1.}$$

iv) The number e may be defined as the limit of the sequence $a_n = \left(1 + \frac{1}{n}\right)^n$, which converges since it is strictly increasing and bounded from above.

v) At last, look at the sequences, all tending to $+\infty$,

$$\boxed{\log n\,,\ n^\alpha\,,\ q^n\,,\ n!\,,\ n^n \quad (\alpha > 0,\ q > 1)\,.}$$

In Vol. I, Sect. 5.4, we proved that each of these is infinite of order bigger than the preceding one. This means that for $n \to \infty$, with $\alpha > 0$ and $q > 1$, we have

$$\log n = o(n^\alpha)\,, \qquad n^\alpha = o(q^n)\,,$$
$$q^n = o(n!)\,, \qquad n! = o(n^n)\,. \qquad \square$$

1.2 Numerical series

To introduce the notion of numerical series, i.e., of "sum of infinitely many numbers", we examine a simple yet instructive situation borrowed from geometry.

Consider a segment of length $\ell = 2$ (Fig. 1.1). The middle point splits it into two parts of length $a_0 = \ell/2 = 1$. While keeping the left half fixed, we subdivide the right one in two parts of length $a_1 = \ell/4 = 1/2$. Iterating the process indefinitely one can think of the initial segment as the union of infinitely many segments of lengths $1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \ldots$ Correspondingly, the total length of the starting segment can be thought of as sum of the lengths of all sub-segments, in other words

$$2 = 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{8} + \frac{1}{16} + \ldots \tag{1.2}$$

On the right we have a sum of infinitely many terms. This infinite sum can be defined properly using sequences, and leads to the notion of a numerical series.

Given the sequence $\{a_k\}_{k\geq 0}$, one constructs the so-called **sequence of partial sums** $\{s_n\}_{n\geq 0}$ in the following manner:

$$s_0 = a_0\,, \qquad s_1 = a_0 + a_1\,, \qquad s_2 = a_0 + a_1 + a_2,$$

and in general,

$$\boxed{s_n = a_0 + a_1 + \ldots + a_n = \sum_{k=0}^{n} a_k\,.}$$

Note that $s_n = s_{n-1} + a_n$. Then it is natural to study the limit behaviour of such a sequence. Let us (formally) define

$$\boxed{\sum_{k=0}^{\infty} a_k = \lim_{n\to\infty} \sum_{k=0}^{n} a_k = \lim_{n\to\infty} s_n\,.}$$

The symbol $\displaystyle\sum_{k=0}^{\infty} a_k$ is called (numerical) **series**, and a_k is the **general term** of the series.

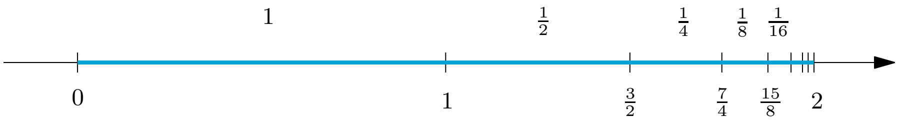


Figure 1.1. Successive splittings of the interval $[0, 2]$. The coordinates of the subdivision points are indicated below the blue line, while the lengths of sub-intervals lie above it

Definition 1.2 *Given the sequence* $\{a_k\}_{k\geq 0}$ *and* $s_n = \sum_{k=0}^{n} a_k$, *consider the limit* $\lim_{n\to\infty} s_n$.

i) If the limit exists (finite or infinite), its value s *is called* **sum of the series** *and one writes*

$$\sum_{k=0}^{\infty} a_k = s = \lim_{n\to\infty} s_n \,.$$

- *If* s *is finite, one says that* **the series** $\sum_{k=0}^{\infty} a_k$ **converges**.
- *If* s *is infinite,* **the series** $\sum_{k=0}^{\infty} a_k$ **diverges**, *to either* $+\infty$ *or* $-\infty$.

ii) If the limit does not exist, **the series** $\sum_{k=0}^{\infty} a_k$ **is indeterminate**.

Examples 1.3

i) Let us go back to the interval split infinitely many times. The length of the shortest segment obtained after $k+1$ subdivisions is $a_k = \frac{1}{2^k}$, $k \geq 0$. Thus, we consider the series $\sum_{k=0}^{\infty} \frac{1}{2^k}$. Its partial sums read

$$s_0 = 1\,, \qquad s_1 = 1 + \frac{1}{2} = \frac{3}{2}\,, \qquad s_2 = 1 + \frac{1}{2} + \frac{1}{4} = \frac{7}{4}\,,$$
$$\vdots$$
$$s_n = 1 + \frac{1}{2} + \ldots + \frac{1}{2^n}\,.$$

Using the fact that $a^{n+1} - b^{n+1} = (a-b)(a^n + a^{n-1}b + \ldots + ab^{n-1} + b^n)$, and choosing $a = 1$ and $b = x$ arbitrary but different from one, we obtain the identity

$$1 + x + \ldots + x^n = \frac{1 - x^{n+1}}{1 - x}\,. \tag{1.3}$$

Therefore

$$s_n = 1 + \frac{1}{2} + \ldots + \frac{1}{2^n} = \frac{1 - \frac{1}{2^{n+1}}}{1 - \frac{1}{2}} = 2\left(1 - \frac{1}{2^{n+1}}\right) = 2 - \frac{1}{2^n}\,,$$

and so

$$\lim_{n\to\infty} s_n = \lim_{n\to\infty}\left(2 - \frac{1}{2^n}\right) = 2\,.$$

The series converges and its sum is 2, which eventually justifies formula (1.2).

ii) The partial sums of the series $\sum_{k=0}^{\infty}(-1)^k$ satisfy

$$\begin{aligned} s_0 &= 1\,, & s_1 &= 1-1=0\,, \\ s_2 &= s_1+1=1\,, & s_3 &= s_2-1=0\,, \\ &\vdots \\ s_{2n} &= 1\,, & s_{2n+1} &= 0\,. \end{aligned}$$

The terms with even index are all equal to 1, whereas the odd ones are 0. Therefore $\lim_{n\to\infty} s_n$ cannot exist and the series is indeterminate.

iii) The two previous examples are special cases of the following series, called **geometric series**,

$$\sum_{k=0}^{\infty} q^k\,,$$

where q is a fixed number in $\mathbb{R}$. The geometric series is particularly important. If $q=1$, then $s_n = a_0+a_1+\ldots+a_n = 1+1+\ldots+1 = n+1$ and $\lim_{n\to\infty} s_n = +\infty$. Hence the series diverges to $+\infty$.

If $q \neq 1$ instead, (1.3) implies

$$s_n = 1+q+q^2+\ldots+q^n = \frac{1-q^{n+1}}{1-q}\,.$$

Recalling Example 1.1 i), we obtain

$$\lim_{n\to\infty} s_n = \lim_{n\to\infty} \frac{1-q^{n+1}}{1-q} = \begin{cases} \dfrac{1}{1-q} & \text{if } |q|<1\,, \\ +\infty & \text{if } q>1\,, \\ \text{does not exist} & \text{if } q \le -1\,. \end{cases}$$

In conclusion

$$\sum_{k=0}^{\infty} q^k \quad \begin{cases} \text{converges to } \dfrac{1}{1-q} & \text{if } |q|<1\,, \\ \text{diverges to } +\infty & \text{if } q \ge 1\,, \\ \text{is indeterminate} & \text{if } q \le -1\,. \end{cases}$$

□

Sometimes the sequence $\{a_k\}$ is only defined for $k \ge k_0$: Definition 1.2 then modifies in the obvious way. Moreover, the following fact holds, whose easy proof is left to the reader.

Property 1.4 *A series' behaviour does not change by adding, modifying or eliminating a finite number of terms.*

Note that this property, in case of convergence, is saying nothing about the sum of the series, which generally changes when the series is altered. For instance,

$$\sum_{k=1}^{\infty} \frac{1}{2^k} = \sum_{k=0}^{\infty} \frac{1}{2^k} - 1 = 2 - 1 = 1 .$$

Examples 1.5

i) The series $\displaystyle\sum_{k=2}^{\infty} \frac{1}{(k-1)k}$ is called **series of Mengoli**. As

$$a_k = \frac{1}{(k-1)k} = \frac{1}{k-1} - \frac{1}{k},$$

it follows that

$$s_2 = a_2 = \frac{1}{1 \cdot 2} = 1 - \frac{1}{2}$$
$$s_3 = a_2 + a_3 = \left(1 - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) = 1 - \frac{1}{3},$$

and in general

$$\begin{aligned} s_n = a_2 + a_3 + \ldots + a_n &= \left(1 - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \ldots + \left(\frac{1}{n-1} - \frac{1}{n}\right) \\ &= 1 - \frac{1}{n} . \end{aligned}$$

Thus

$$\lim_{n\to\infty} s_n = \lim_{n\to\infty} \left(1 - \frac{1}{n}\right) = 1$$

and the series converges to 1.

ii) For the series $\displaystyle\sum_{k=1}^{\infty} \log\left(1 + \frac{1}{k}\right)$ we have

$$a_k = \log\left(1 + \frac{1}{k}\right) = \log \frac{k+1}{k} = \log(k+1) - \log k ,$$

so

$$\begin{aligned} s_1 &= \log 2 , \qquad s_2 = \log 2 + (\log 3 - \log 2) = \log 3 , \\ &\vdots \\ s_n &= \log 2 + (\log 3 - \log 2) + \ldots + \big(\log(n+1) - \log n\big) = \log(n+1) . \end{aligned}$$

Hence

$$\lim_{n\to\infty} s_n = \lim_{n\to\infty} \log(n+1) = +\infty$$

and the series diverges (to $+\infty$). □

The two instances just considered belong to the larger class of **telescopic series**. These are defined by $a_k = b_{k+1} - b_k$ for a suitable sequence $\{b_k\}_{k \geq k_0}$.

Since $s_n = b_{n+1} - b_{k_0}$, the behaviour of a telescopic series is the same as that of the sequence $\{b_k\}$.

There is a simple yet useful necessary condition for a numerical series to converge.

Property 1.6 *Let* $\sum_{k=0}^{\infty} a_k$ *be a converging series. Then*

$$\lim_{k\to\infty} a_k = 0\,. \tag{1.4}$$

Proof. Let $s = \lim_{n\to\infty} s_n$. Since $a_k = s_k - s_{k-1}$,

$$\lim_{k\to\infty} a_k = \lim_{k\to\infty} (s_k - s_{k-1}) = s - s = 0\,.$$ □

Observe that condition (1.4) is not sufficient to guarantee convergence. The general term of a series may tend to 0 without the series having to converge. For example we saw that $\sum_{k=1}^{\infty} \log\left(1+\frac{1}{k}\right)$ diverges (Example 1.5 ii)), while the continuity of the logarithm implies that $\lim_{k\to\infty} \log\left(1+\frac{1}{k}\right) = 0$.

Example 1.7

It is easy to see that $\sum_{k=1}^{\infty} \left(1-\frac{1}{k}\right)^k$ does not converge, because the general term $a_k = \left(1-\frac{1}{k}\right)^k$ tends to $\mathrm{e}^{-1} \neq 0$. □

If a series $\sum_{k=0}^{\infty} a_k$ converges to s, the quantity

$$\boxed{r_n = s - s_n = \sum_{k=n+1}^{\infty} a_k}$$

is called n**th remainder**.

Now comes another necessary condition for convergence.

Property 1.8 *Let* $\sum_{k=0}^{\infty} a_k$ *be a convergent series. Then*

$$\lim_{n\to\infty} r_n = 0\,.$$

Proof. Just note $\lim_{n\to\infty} r_n = \lim_{n\to\infty} (s - s_n) = s - s = 0\,.$ □

It is not always possible to predict the behaviour of a series $\sum_{k=0}^{\infty} a_k$ using merely the definition. It may well happen that the sequence of partial sums cannot be computed explicitly, so it becomes important to have other ways to establish whether the series converges or not. In case of convergence it could also be necessary to determine the actual sum explicitly. This may require using more sophisticated techniques, which go beyond the scopes of this text.

1.3 Series with positive terms

We deal with series $\sum_{k=0}^{\infty} a_k$ for which $a_k \geq 0$ for any $k \in \mathbb{N}$.

Proposition 1.9 *A series $\sum_{k=0}^{\infty} a_k$ with positive terms either converges or diverges to $+\infty$.*

Proof. The sequence s_n is monotonically increasing since

$$s_{n+1} = s_n + a_{n+1} \geq s_n\,, \qquad \forall n \geq 0\,.$$

It is then sufficient to use Theorem 3 on p. 2 to conclude that $\lim_{n\to\infty} s_n$ exists, and is either finite or $+\infty$. □

We list a few tools for studying the convergence of positive-term series.

Theorem 1.10 (**Comparison Test**) *Let $\sum_{k=0}^{\infty} a_k$ and $\sum_{k=0}^{\infty} b_k$ be positive-term series such that $0 \leq a_k \leq b_k$, for any $k \geq 0$.*

i) If $\sum_{k=0}^{\infty} b_k$ converges, also $\sum_{k=0}^{\infty} a_k$ converges and

$$\sum_{k=0}^{\infty} a_k \leq \sum_{k=0}^{\infty} b_k\,.$$

ii) If $\sum_{k=0}^{\infty} a_k$ diverges, then $\sum_{k=0}^{\infty} b_k$ diverges as well.

Proof. *i)* Denote by $\{s_n\}$ and $\{t_n\}$ the sequences of partial sums of $\sum_{k=0}^{\infty} a_k$, $\sum_{k=0}^{\infty} b_k$ respectively. Since $a_k \le b_k$ for all k,

$$s_n \le t_n\,, \qquad \forall n \ge 0\,.$$

By assumption, the series $\sum_{k=0}^{\infty} b_k$ converges, so $\lim_{n\to\infty} t_n = t \in \mathbb{R}$. Proposition 1.9 implies that $\lim_{n\to\infty} s_n = s$ exists, finite or infinite. By the First Comparison Theorem (Theorem 4, p. 2) we have

$$s = \lim_{n\to\infty} s_n \le \lim_{n\to\infty} t_n = t \in \mathbb{R}\,.$$

Therefore $s \in \mathbb{R}$, and the series $\sum_{k=0}^{\infty} a_k$ converges. Furthermore $s \le t$.

ii) By contradiction, if $\sum_{k=0}^{\infty} b_k$ converged, part *i)* would force $\sum_{k=0}^{\infty} a_k$ to converge too. $\square$

Examples 1.11

i) Consider $\displaystyle\sum_{k=1}^{\infty} \frac{1}{k^2}$. As

$$\frac{1}{k^2} < \frac{1}{(k-1)k}\,, \qquad \forall k \ge 2\,,$$

and the series of Mengoli $\displaystyle\sum_{k=2}^{\infty} \frac{1}{(k-1)k}$ converges (Example 1.5 i)), we conclude that the given series converges to a sum ≤ 2. One can prove the sum is precisely $\dfrac{\pi^2}{6}$ (see Example 3.18).

ii) The series $\displaystyle\sum_{k=1}^{\infty} \frac{1}{k}$ is known as **harmonic series**. Since $\log(1+x) \le x\,, \forall x > -1\,,$ (Vol. I, Ch. 6, Exercise 12), it follows

$$\log\Big(1+\frac{1}{k}\Big) \le \frac{1}{k}\,, \qquad \forall k \ge 1\,;$$

but since $\displaystyle\sum_{k=1}^{\infty} \log\Big(1+\frac{1}{k}\Big)$ diverges (Example 1.5 ii)), we conclude that the harmonic series diverges.

iii) Subsuming the previous two examples we have

$$\sum_{k=1}^{\infty} \frac{1}{k^\alpha}\,, \qquad \alpha \quad 0\,,> \tag{1.5}$$

called **generalised harmonic series**. Because

$$\frac{1}{k^\alpha} > \frac{1}{k} \quad \text{for } 0 < \alpha < 1\,, \qquad \frac{1}{k^\alpha} < \frac{1}{k^2} \quad \text{for } \alpha > 2\,,$$

the Comparison Test tells us the generalised harmonic series diverges for $0 < \alpha < 1$ and converges for $\alpha > 2$. The case $1 < \alpha < 2$ will be examined in Example 1.19. □

Here is a useful criterion that generalises the Comparison Test.

Theorem 1.12 (**Asymptotic Comparison Test**) *Let* $\sum_{k=0}^{\infty} a_k$ *and* $\sum_{k=0}^{\infty} b_k$ *be positive-term series and suppose the sequences* $\{a_k\}_{k\geq 0}$ *and* $\{b_k\}_{k\geq 0}$ *have the same order of magnitude for* $k \to \infty$. *Then the series have the same behaviour.*

Proof. Having the same order of magnitude for $k \to \infty$ is equivalent to

$$\lim_{k\to\infty} \frac{a_k}{b_k} = \ell \in \mathbb{R} \setminus \{0\}\,.$$

Therefore the sequences $\left\{\frac{a_k}{b_k}\right\}_{k\geq 0}$ and $\left\{\frac{b_k}{a_k}\right\}_{k\geq 0}$ are both convergent, hence both bounded (Theorem 2, p. 2). There must exist constants $M_1, M_2 > 0$ such that

$$\left|\frac{a_k}{b_k}\right| \leq M_1 \qquad \text{and} \qquad \left|\frac{b_k}{a_k}\right| \leq M_2$$

for any $k > 0$, i.e.,

$$|a_k| \leq M_1 |b_k| \qquad \text{and} \qquad |b_k| \leq M_2 |a_k|\,.$$

Now it suffices to use Theorem 1.10 to finish the proof. □

Examples 1.13

i) Consider $\sum_{k=0}^{\infty} a_k = \sum_{k=0}^{\infty} \frac{k+3}{2k^2+5}$ and let $b_k = \frac{1}{k}$. Then

$$\lim_{k\to\infty} \frac{a_k}{b_k} = \frac{1}{2}$$

and the given series behaves as the harmonic series, hence diverges.

ii) Take the series $\sum_{k=1}^{\infty} a_k = \sum_{k=1}^{\infty} \sin\frac{1}{k^2}$. As $\sin\frac{1}{k^2} \sim \frac{1}{k^2}$ for $k \to \infty$, the series has the same behaviour of $\sum_{k=1}^{\infty} \frac{1}{k^2}$, so it converges. □

Eventually, here are two results – of algebraic flavour and often easy to employ – which prescribe sufficient conditions for a series to converge or diverge.

Theorem 1.14 (**Ratio Test**) *Let the series* $\sum_{k=0}^{\infty} a_k$ *have* $a_k > 0$, $\forall k \geq 0$. *Assume the limit*

$$\lim_{k\to\infty} \frac{a_{k+1}}{a_k} = \ell$$

exists, finite or infinite. If $\ell < 1$ *the series converges; if* $\ell > 1$ *it diverges.*

Proof. First, suppose ℓ finite. By definition of limit we know that for any $\varepsilon > 0$, there is an integer $k_\varepsilon \geq 0$ such that

$$\forall k > k_\varepsilon \quad \Rightarrow \quad \left| \frac{a_{k+1}}{a_k} - \ell \right| < \varepsilon \quad \text{i.e.,} \quad \ell - \varepsilon < \frac{a_{k+1}}{a_k} < \ell + \varepsilon \,.$$

Assume $\ell < 1$. Choose $\varepsilon = \frac{1-\ell}{2}$ and set $q = \frac{1+\ell}{2}$, so

$$0 < \frac{a_{k+1}}{a_k} < \ell + \varepsilon = q \,, \qquad \forall k > k_\varepsilon \,.$$

Repeating the argument we obtain

$$a_{k+1} < q a_k < q^2 a_{k-1} < \ldots < q^{k-k_\varepsilon} a_{k_\varepsilon+1}$$

hence

$$a_{k+1} < \frac{a_{k_\varepsilon+1}}{q^{k_\varepsilon}} q^k \,, \quad \forall k > k_\varepsilon \,.$$

The claim follows by Theorem 1.10 and from the fact that the geometric series, with $q < 1$, converges (Example 1.3).

Now consider $\ell > 1$. Choose $\varepsilon = \ell - 1$, and notice

$$1 = \ell - \varepsilon < \frac{a_{k+1}}{a_k} \,, \quad \forall k > k_\varepsilon \,.$$

Thus $a_{k+1} > a_k > \ldots > a_{k_\varepsilon+1} > 0$, so the necessary condition for convergence fails, for $\lim_{k\to\infty} a_k \neq 0$.

Eventually, if $\ell = +\infty$, we put $A = 1$ in the condition of limit, and there exists $k_A \geq 0$ with $a_k > 1$, for any $k > k_A$. Once again the necessary condition to have convergence does not hold. □

Theorem 1.15 (**Root Test**) *Given a series* $\sum_{k=0}^{\infty} a_k$ *with* $a_k \geq 0$, $\forall k \geq 0$, *suppose*

$$\lim_{k\to\infty} \sqrt[k]{a_k} = \ell$$

exists, finite or infinite. If $\ell < 1$ *the series converges, if* $\ell > 1$ *it diverges.*

Proof. This proof is essentially identical to the previous one, so we leave it to the reader. □

Examples 1.16

i) For $\sum_{k=0}^{\infty} \frac{k}{3^k}$ we have $a_k = \frac{k}{3^k}$ and $a_{k+1} = \frac{k+1}{3^{k+1}}$, therefore

$$\lim_{k\to\infty} \frac{a_{k+1}}{a_k} = \lim_{k\to\infty} \frac{1}{3}\frac{k+1}{k} = \frac{1}{3} < 1\,.$$

The given series converges by the Ratio Test 1.14.

ii) The series $\sum_{k=1}^{\infty} \frac{1}{k^k}$ has

$$\lim_{k\to\infty} \sqrt[k]{a_k} = \lim_{k\to\infty} \frac{1}{k} = 0 < 1\,.$$

The Root Test 1.15 ensures that the series converges. □

We remark that the Ratio and Root Tests do not allow to conclude anything if $\ell = 1$. For example, $\sum_{k=1}^{\infty} \frac{1}{k}$ diverges and $\sum_{k=1}^{\infty} \frac{1}{k^2}$ converges, yet they both satisfy Theorems 1.14 and 1.15 with $\ell = 1$.

In certain situations it may be useful to think of the general term a_k as the value at $x = k$ of a function f defined on the half-line $[k_0, +\infty)$. Under the appropriate assumptions, we can relate the behaviour of the series to that of the integral of f over $[k_0, +\infty)$. In fact,

Theorem 1.17 (**Integral Test**) *Let f be continuous, positive and decreasing on $[k_0, +\infty)$, for $k_0 \in \mathbb{N}$. Then*

$$\sum_{k=k_0+1}^{\infty} f(k) \le \int_{k_0}^{+\infty} f(x)\,\mathrm{d}x \le \sum_{k=k_0}^{\infty} f(k)\,. \tag{1.6}$$

Therefore the integral and the series share the same behaviour:

a) $\int_{k_0}^{+\infty} f(x)\,\mathrm{d}x$ *converges* $\iff$ $\sum_{k=k_0}^{\infty} f(k)$ *converges;*

b) $\int_{k_0}^{+\infty} f(x)\,\mathrm{d}x$ *diverges* $\iff$ $\sum_{k=k_0}^{\infty} f(k)$ *diverges.*

Proof. Since f decreases, for any $k \ge k_0$ we have

$$f(k+1) \le f(x) \le f(k)\,, \qquad \forall x \in [k, k+1]\,,$$

and as the integral is monotone,

$$f(k+1) \le \int_{k}^{k+1} f(x)\,\mathrm{d}x \le f(k)\,.$$

Then for all $n \in \mathbb{N}$ with $n > k_0$ we obtain

$$\sum_{k=k_0+1}^{n+1} f(k) \le \int_{k_0}^{n+1} f(x)\,\mathrm{d}x \le \sum_{k=k_0}^{n} f(k)$$

(after re-indexing the first series). Passing to the limit for $n \to +\infty$ and recalling f is positive and continuous, we conclude. □

From inequalities (1.6) it follows easily that

$$\int_{k_0}^{+\infty} f(x)\,\mathrm{d}x \le \sum_{k=k_0}^{\infty} f(k) \le f(k_0) + \int_{k_0}^{+\infty} f(x)\,\mathrm{d}x\,.$$

Comparing with the improper integral of f allows to estimate, often accurately, the remainder and the sum of the series, and use this to estimate numerically these values:

Property 1.18 *Under the assumptions of Theorem 1.17, if* $\sum_{k=k_0}^{\infty} f(k)$ *converges then for all* $n \ge k_0$

$$\int_{n+1}^{+\infty} f(x)\,\mathrm{d}x \le r_n \le \int_{n}^{+\infty} f(x)\,\mathrm{d}x\,, \tag{1.7}$$

and

$$s_n + \int_{n+1}^{+\infty} f(x)\,\mathrm{d}x \le s \le s_n + \int_{n}^{+\infty} f(x)\,\mathrm{d}x\,. \tag{1.8}$$

Proof. If $\sum_{k=k_0}^{\infty} f(k)$ converges, (1.6) can we re-written substituting k_0 with any integer $n \ge k_0$. Using the first inequality,

$$r_n = s - s_n = \sum_{k=n+1}^{\infty} f(k) \le \int_{n}^{+\infty} f(x)\,\mathrm{d}x\,,$$

while changing k_0 to $n+1$ in the second one yields

$$\int_{n+1}^{+\infty} f(x)\,\mathrm{d}x \le r_n\,.$$

This gives formula (1.7), from which (1.8) follows by adding s_n to each side. □

Examples 1.19

i) The Integral Test is used to study the generalised harmonic series (1.5) for all admissible values of the parameter α. Note in fact that the function $\dfrac{1}{x^\alpha}$, $\alpha > 0$, fulfills the hypotheses and its integral over $[1, +\infty)$ converges if and only if $\alpha > 1$. In conclusion,

$$\boxed{\sum_{k=1}^{\infty} \frac{1}{k^\alpha} \begin{cases} \text{converges if } \alpha > 1\,, \\ \text{diverges if } 0 < \alpha \le 1\,. \end{cases}}$$

ii) In order to study

$$\sum_{k=2}^{\infty} \frac{1}{k \log k}$$

we take the map $f(x) = \dfrac{1}{x \log x}$; its integral over $[2, +\infty)$ diverges, since

$$\int_2^{+\infty} \frac{1}{x \log x}\, \mathrm{d}x = \int_{\log 2}^{+\infty} \frac{1}{t}\, \mathrm{d}t = +\infty.$$

Consequently, the series $\displaystyle\sum_{k=2}^{\infty} \frac{1}{k \log k}$ is divergent.

iii) Suppose we want to estimate the precision of the sum of $\displaystyle\sum_{k=1}^{\infty} \frac{1}{k^3}$ computed up to the first 10 terms.

We need to calculate $\displaystyle\int_n^{+\infty} f(x)\, \mathrm{d}x$ with $f(x) = \dfrac{1}{x^3}$:

$$\int_n^{+\infty} \frac{1}{x^3}\, \mathrm{d}x = \lim_{c \to +\infty} \left[-\frac{1}{2x^2} \right]_n^c = \frac{1}{2n^2}\,.$$

By (1.7) we obtain

$$r_{10} = s - s_{10} \le \int_{10}^{+\infty} \frac{1}{x^3}\, \mathrm{d}x = \frac{1}{2\,(10)^2} = 0.005$$

and

$$r_{10} \ge \int_{11}^{+\infty} \frac{1}{x^3}\, \mathrm{d}x = \frac{1}{2\,(11)^2} = 0.004132\ldots$$

The sum may be estimated with the help of (1.8):

$$s_{10} + \frac{1}{2\,(11)^2} \le s \le s_{10} + \frac{1}{2\,(10)^2}\,.$$

Since

$$s_{10} = 1 + \frac{1}{2^3} + \ldots + \frac{1}{10^3} = 1.197532\ldots,$$

we find $1.201664 \le s \le 1.202532$. The exact value for s is $1.202057\ldots$ □

1.4 Alternating series

These are series of the form

$$\boxed{\sum_{k=0}^{\infty}(-1)^k b_k \qquad \text{with} \qquad b_k > 0\,, \quad \forall k \geq 0.}$$

For them the following result due to Leibniz holds.

Theorem 1.20 (Leibniz's Alternating Series Test) *An alternating series* $\sum_{k=0}^{\infty}(-1)^k b_k$ *converges if the following conditions hold*

i) $\lim_{k\to\infty} b_k = 0\,;$

ii) the sequence $\{b_k\}_{k\geq 0}$ *decreases monotonically.*

Denoting by s *its sum, for all* $n \geq 0$ *one has*

$$|r_n| = |s - s_n| \leq b_{n+1} \qquad \text{and} \qquad s_{2n+1} \leq s \leq s_{2n}\,.$$

Proof. As $\{b_k\}_{k\geq 0}$ is a decreasing sequence,

$$s_{2n} = s_{2n-2} - b_{2n-1} + b_{2n} = s_{2n-2} - (b_{2n-1} - b_{2n}) \leq s_{2n-2}$$

and

$$s_{2n+1} = s_{2n-1} + b_{2n} - b_{2n+1} \geq s_{2n-1}\,.$$

Thus the subsequence of partial sums made by the terms with even index decreases, whereas the subsequence of odd indexes increases. For any $n \geq 0$, moreover,

$$s_{2n} = s_{2n-1} + b_{2n} \geq s_{2n-1} \geq \ldots \geq s_1$$

and

$$s_{2n+1} = s_{2n} - b_{2n+1} \leq s_{2n} \leq \ldots \leq s_0\,.$$

Thus $\{s_{2n}\}_{n\geq 0}$ is bounded from below and $\{s_{2n+1}\}_{n\geq 0}$ from above. By Theorem 3 on p. 2 both sequences converge, so let us put

$$\lim_{n\to\infty} s_{2n} = \inf_{n\geq 0} s_{2n} = s^* \qquad \text{and} \qquad \lim_{n\to\infty} s_{2n+1} = \sup_{n\geq 0} s_{2n+1} = s_*\,.$$

Since

$$s^* - s_* = \lim_{n\to\infty}\big(s_{2n} - s_{2n+1}\big) = \lim_{n\to\infty} b_{2n+1} = 0\,,$$

we conclude that the series $\sum_{k=0}^{\infty}(-1)^k b_k$ has sum $s = s^* = s_*$. In addition,

$$s_{2n+1} \leq s \leq s_{2n}\,, \qquad \forall n \geq 0\,,$$

in other words the sequence $\{s_{2n}\}_{n\geq 0}$ approximates s from above, while $\{s_{2n+1}\}_{n\geq 0}$ approximates s from below. For any $n \geq 0$ we have

$$0 \leq s - s_{2n+1} \leq s_{2n+2} - s_{2n+1} = b_{2n+2}$$

and

$$0 \leq s_{2n} - s \leq s_{2n} - s_{2n+1} = b_{2n+1}\,,$$

so $|r_n| = |s - s_n| \leq b_{n+1}$. □

Examples 1.21

i) Consider the generalised alternating harmonic series $\displaystyle\sum_{k=1}^{\infty}(-1)^k\frac{1}{k^\alpha}$, where $\alpha > 0$. As $\lim\limits_{k\to\infty} b_k = \lim\limits_{k\to\infty} \frac{1}{k^\alpha} = 0$ and the sequence $\left\{\frac{1}{k^\alpha}\right\}_{k\geq 1}$ is strictly decreasing, the series converges.

ii) Condition *i)* in Leibniz's Test is also necessary, whereas *ii)* is only sufficient. In fact, for $k \geq 2$ let

$$b_k = \begin{cases} 1/k & k \text{ even}\,,\\ (k-1)/k^2 & k \text{ odd}\,. \end{cases}$$

It is straightforward that $b_k > 0$ and b_k is infinitesimal. The sequence is not monotone decreasing since $b_k > b_{k+1}$ for k even, $b_k < b_{k+1}$ for k odd. Nevertheless,

$$\sum_{k=2}^{\infty} b_k = \sum_{k=2}^{\infty}(-1)^k\frac{1}{k} + \sum_{\substack{k\geq 3\\ k \text{ odd}}} \frac{1}{k^2}$$

converges, for the two series on the right converge.

iii) We want to approximate the sum of $\displaystyle\sum_{k=0}^{\infty}\frac{(-1)^k}{k!}$ to the third digit, meaning with a margin less than 10^{-3}. The series, alternating for $b_k = \frac{1}{k!}$, converges by Leibniz's Test. From $|s - s_n| \leq b_{n+1}$ we see that for $n = 6$

$$0 < s_6 - s \leq b_7 = \frac{1}{5040} < \frac{1}{5000} = 0.0002 < 10^{-3}\,.$$

As $s_6 = 0.368056\ldots$, the estimate $s \sim 0.368$ is correct up to the third place. □

To study series with arbitrary signs the notion of absolute convergence is useful.

Definition 1.22 *The series* $\displaystyle\sum_{k=0}^{\infty} a_k$ **converges absolutely** *if the positive-term series* $\displaystyle\sum_{k=0}^{\infty} |a_k|$ *converges.*

Example 1.23

The series $\displaystyle\sum_{k=0}^{\infty}(-1)^k\frac{1}{k^2}$ converges absolutely because $\displaystyle\sum_{k=0}^{\infty}\frac{1}{k^2}$ converges. □

The next fact ensures that absolute convergence implies convergence.

Theorem 1.24 (**Absolute Convergence Test**) *If* $\displaystyle\sum_{k=0}^{\infty} a_k$ *converges absolutely then it also converges, and*

$$\left|\sum_{k=0}^{\infty} a_k\right| \le \sum_{k=0}^{\infty} |a_k|\,.$$

Proof. The proof is similar to that of the Absolute Convergence Test for improper integrals.
Define sequences

$$a_k^+ = \begin{cases} a_k & \text{if } a_k \ge 0 \\ 0 & \text{if } a_k < 0 \end{cases} \qquad \text{and} \qquad a_k^- = \begin{cases} 0 & \text{if } a_k \ge 0 \\ -a_k & \text{if } a_k < 0\,. \end{cases}$$

Note $a_k^+, a_k^- \ge 0$ for all $k \ge 0$, and

$$a_k = a_k^+ - a_k^-\,, \qquad\qquad |a_k| = a_k^+ + a_k^-\,.$$

As $0 \le a_k^+, a_k^- \le |a_k|$, for any $k \ge 0$, the Comparison Test (Theorem 1.10) tells us that $\displaystyle\sum_{k=0}^{\infty} a_k^+$ and $\displaystyle\sum_{k=0}^{\infty} a_k^-$ converge. But for any $n \ge 0$,

$$\sum_{k=0}^{n} a_k = \sum_{k=0}^{n} \left(a_k^+ - a_k^-\right) = \sum_{k=0}^{n} a_k^+ - \sum_{k=0}^{n} a_k^-\,,$$

so also the series $\displaystyle\sum_{k=0}^{\infty} a_k = \sum_{k=0}^{\infty} a_k^+ - \sum_{k=0}^{\infty} a_k^-$ converges.
Passing now to the limit for $n \to \infty$ in

$$\left|\sum_{k=0}^{n} a_k\right| \le \sum_{k=0}^{n} |a_k|\,,$$

we obtain the desired inequality. □

Remark 1.25 There are series that converge, but not absolutely. The alternating harmonic series $\displaystyle\sum_{k=1}^{\infty}(-1)^k\frac{1}{k}$ is one such example, for it has a finite sum, but does not converge absolutely, since the harmonic series $\displaystyle\sum_{k=1}^{\infty}\frac{1}{k}$ diverges. In such a situation one speaks about **conditional convergence**. □

The previous criterion allows one to study alternating series by their absolute convergence. As the series of absolute values has positive terms, all criteria seen in Sect 1.3 apply.

1.5 The algebra of series

Two series $\displaystyle\sum_{k=0}^{\infty}a_k$, $\displaystyle\sum_{k=0}^{\infty}b_k$ can be added, multiplied by numbers and multiplied between themselves. The sum is defined in the obvious way as the series whose formal general term reads $c_k = a_k + b_k$:

$$\sum_{k=0}^{\infty}a_k + \sum_{k=0}^{\infty}b_k = \sum_{k=0}^{\infty}(a_k + b_k)\,.$$

Assume the series both converge or diverge, and write $s = \displaystyle\sum_{k=0}^{\infty}a_k$, $t = \displaystyle\sum_{k=0}^{\infty}b_k$ $(s, t \in \mathbb{R} \cup \{\pm\infty\})$. The sum is determinate (convergent or divergent) whenever the expression $s + t$ is defined. If so,

$$\sum_{k=0}^{\infty}(a_k + b_k) = s + t$$

and the sum converges if $s + t \in \mathbb{R}$, diverges if $s + t = \pm\infty$.

If one series converges and the other is indeterminate the sum is necessarily indeterminate.

Apart from these cases, the nature of the sum cannot be deduced directly from the behaviour of the two summands, and must be studied case by case.

Let now $\lambda \in \mathbb{R} \setminus \{0\}$; the series $\lambda\displaystyle\sum_{k=0}^{\infty}a_k$ is by definition the series with general term λa_k. Its behaviour coincides with that of $\displaystyle\sum_{k=0}^{\infty}a_k$. Anyhow, in case of convergence or divergence,

$$\sum_{k=0}^{\infty}\lambda a_k = \lambda s\,.$$

In order to define the product some thinking is needed. If two series converge respectively to $s, t \in \mathbb{R}$ we would like the product series to converge to st. This cannot happen if one defines the general term c_k of the product simply as the product of the corresponding terms, by setting $c_k = a_k b_k$. An example will clarify the issue: consider the two geometric series with terms $a_k = \frac{1}{2^k}$ and $b_k = \frac{1}{3^k}$. Then

$$\sum_{k=0}^{\infty} \frac{1}{2^k} = \frac{1}{1-\frac{1}{2}} = 2\,, \qquad \sum_{k=0}^{\infty} \frac{1}{3^k} = \frac{1}{1-\frac{1}{3}} = \frac{3}{2}$$

while

$$\sum_{k=0}^{\infty} \frac{1}{2^k}\frac{1}{3^k} = \sum_{k=0}^{\infty} \frac{1}{6^k} = \frac{1}{1-\frac{1}{6}} = \frac{6}{5} \neq 2\,\frac{3}{2} = 3\,.$$

One way to multiply series and preserve the above property is the so-called **Cauchy product**, defined by its general term

$$c_k = \sum_{j=0}^{k} a_j b_{k-j} = a_0 b_k + a_1 b_{k-1} + \cdots + a_{k-1} b_1 + a_k b_0\,. \tag{1.9}$$

By arranging the products $a_\ell b_m$ ($\ell, m \geq 0$) in a matrix

$$\begin{array}{c|cccc}
 & b_0 & b_1 & b_2 & \ldots \\
\hline
a_0 & a_0 b_0 & a_0 b_1 & a_0 b_2 & \ldots \\
a_1 & a_1 b_0 & a_1 b_1 & a_1 b_2 & \ldots \\
a_2 & a_2 b_0 & a_2 b_1 & a_2 b_2 & \ldots \\
\vdots & & & &
\end{array}$$

each term c_k becomes the sum of the entries on the kth anti-diagonal.

The reason for this particular definition will become clear when we will discuss power series (Sect. 2.4.1).

It is possible to prove that the absolute convergence of $\sum\limits_{k=0}^{\infty} a_k$ and $\sum\limits_{k=0}^{\infty} b_k$ is sufficient to guarantee the convergence of $\sum\limits_{k=0}^{\infty} c_k$, in which case

$$\sum_{k=0}^{\infty} c_k = \Big(\sum_{k=0}^{\infty} a_k\Big)\Big(\sum_{k=0}^{\infty} b_k\Big) = st\,.$$

1.6 Exercises

1. *Find the general term a_n of the following sequences, and compute $\lim_{n\to\infty} a_n$:*

a) $\frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \frac{4}{5}, \ldots$ b) $-\frac{2}{3}, \frac{3}{9}, -\frac{4}{27}, \frac{5}{81}, \ldots$ c) $0, 1, \sqrt{2}, \sqrt{3}, \sqrt{4}, \ldots$

2. *Study the behaviour of the sequences below and compute the limit if this exists:*

a) $a_n = n(n-1)\,, \quad n \geq 0$

b) $a_n = \dfrac{n+5}{2n-1}\,, \quad n \geq 0$

c) $a_n = \dfrac{2+6n^2}{3n+n^2}\,, \quad n \geq 1$

d) $a_n = \dfrac{\sqrt[3]{n}}{1+\sqrt[3]{n}}\,, \quad n \geq 0$

e) $a_n = \dfrac{5^n}{3^{n+1}}\,, \quad n \geq 0$

f) $a_n = \dfrac{(-1)^{n-1}n^2}{n^2+1}\,, \quad n \geq 0$

g) $a_n = \arctan 5n\,, \quad n \geq 0$

h) $a_n = 3 + \cos n\pi\,, \quad n \geq 0$

i) $a_n = 1 + (-1)^n \sin \dfrac{1}{n}\,, \quad n \geq 1$

ℓ) $a_n = \dfrac{n \cos n}{n^3+1}\,, \quad n \geq 0$

m) $a_n = \sqrt{n+3} - \sqrt{n}\,, \quad n \geq 0$

n) $a_n = \dfrac{\log(2+\mathrm{e}^n)}{4n}\,, \quad n \geq 1$

o) $a_n = -3n + \log(n+1) - \log n\,, \; n \geq 1$

p) $a_n = \dfrac{(-3)^n}{n!}\,, \quad n \geq 1$

3. *Study the behaviour of the following sequences:*

a) $a_n = n - \sqrt{n}$

b) $a_n = (-1)^n \dfrac{n^2+1}{\sqrt{n^2+2}}$

c) $a_n = \dfrac{3^n - 4^n}{1+4^n}$

d) $a_n = \dfrac{(2n)!}{n!}$

e) $a_n = \dfrac{(2n)!}{(n!)^2}$

f) $a_n = \dbinom{n}{3} \dfrac{6}{n^3}$

g) $a_n = \left(\dfrac{n^2-n+1}{n^2+n+2}\right)^{\sqrt{n^2+2}}$

h) $a_n = 2^n \sin(2^{-n}\pi)$

i) $a_n = n \cos \dfrac{n+1}{n} \dfrac{\pi}{2}$

ℓ) $a_n = n! \left(\cos \dfrac{1}{\sqrt{n!}} - 1\right)$

4. *Tell whether the following series converge; if they do, calculate their sum:*

a) $\displaystyle\sum_{k=1}^{\infty} 4 \left(\frac{1}{3}\right)^{k-1}$

b) $\displaystyle\sum_{k=1}^{\infty} \frac{2k}{k+5}$

c) $\displaystyle\sum_{k=0}^{\infty} \tan k$

d) $\displaystyle\sum_{k=1}^{\infty} \left(\sin\frac{1}{k} - \sin\frac{1}{k+1}\right)$

e) $\displaystyle\sum_{k=0}^{\infty} \frac{3^k + 2^k}{6^k}$

f) $\displaystyle\sum_{k=1}^{\infty} \frac{1}{2+3^{-k}}$

5. *Using the geometric series, write the number* $2.3\overline{17} = 2.3171717\dots$ *as a ratio of integers.*

6. *Determine the values of the real number x for which the series below converge. Then compute the sum:*

a) $\displaystyle\sum_{k=2}^{\infty} \frac{x^k}{5^k}$

b) $\displaystyle\sum_{k=1}^{\infty} 3^k (x+2)^k$

c) $\displaystyle\sum_{k=1}^{\infty} \frac{1}{x^k}$

d) $\displaystyle\sum_{k=0}^{\infty} \tan^k x$

7. *Find the real numbers c such that*

$$\sum_{k=2}^{\infty} (1+c)^{-k} = 2\,.$$

8. *Suppose* $\displaystyle\sum_{k=1}^{\infty} a_k$ $(a_k \neq 0)$ *converges. Show that* $\displaystyle\sum_{k=1}^{\infty} \frac{1}{a_k}$ *cannot converge.*

9. *Study the convergence of the following positive-term series:*

a) $\displaystyle\sum_{k=0}^{\infty} \frac{3}{2k^2+1}$

b) $\displaystyle\sum_{k=2}^{\infty} \frac{2^k}{k^5-3}$

c) $\displaystyle\sum_{k=0}^{\infty} \frac{3^k}{k!}$

d) $\displaystyle\sum_{k=1}^{\infty} \frac{k!}{k^k}$

e) $\displaystyle\sum_{k=1}^{\infty} k \arcsin\frac{7}{k^2}$

f) $\displaystyle\sum_{k=1}^{\infty} \log\left(1+\frac{5}{k^2}\right)$

g) $\displaystyle\sum_{k=1}^{\infty} \frac{\log k}{k}$

h) $\displaystyle\sum_{k=1}^{\infty} \frac{1}{2^k-1}$

i) $\displaystyle\sum_{k=1}^{\infty} \sin\frac{1}{k}$

ℓ) $\displaystyle\sum_{k=0}^{\infty} \frac{2+3^k}{2^k}$

m) $\displaystyle\sum_{k=1}^{\infty} \frac{k+3}{\sqrt[3]{k^9+k^2}}$

n) $\displaystyle\sum_{k=1}^{\infty} \frac{\cos^2 k}{k\sqrt{k}}$

10. *Find the real numbers p such that* $\displaystyle\sum_{k=2}^{\infty} \frac{1}{k(\log k)^p}$ *converges.*

11. *Estimate the sum s of the series* $\displaystyle\sum_{k=0}^{\infty} \frac{1}{k^2+4}$ *using the first six terms.*

12. *Study the convergence of the following alternating series:*

a) $\displaystyle\sum_{k=1}^{\infty} (-1)^k \log\left(\frac{1}{k}+1\right)$

b) $\displaystyle\sum_{k=0}^{\infty} (-1)^k \sqrt{\frac{k^3+3}{2k^3-5}}$

c) $\displaystyle\sum_{k=1}^{\infty} \sin\left(k\pi+\frac{1}{k}\right)$

d) $\displaystyle\sum_{k=1}^{\infty} (-1)^k \left(\left(1+\frac{1}{k^2}\right)^{\sqrt{2}}-1\right)$

e) $\displaystyle\sum_{k=1}^{\infty} \frac{(-1)^k 3k}{4k-1}$

f) $\displaystyle\sum_{k=1}^{\infty} (-1)^{k+1} \frac{k^2}{k^3+1}$

13. *Check that the series below converge. Determine the minimum number n of terms necessary for the nth partial sum s_n to approximate the sum with the given margin:*

a) $\displaystyle\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k^4}, \qquad |r_n| < 10^{-3}$

b) $\displaystyle\sum_{k=1}^{\infty} \frac{(-2)^k}{k!}, \qquad |r_n| < 10^{-2}$

c) $\displaystyle\sum_{k=1}^{\infty} \frac{(-1)^k k}{4^k}, \qquad |r_n| < 2\cdot 10^{-3}$

14. *Study the absolute convergence of the following series:*

a) $\displaystyle\sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{\sqrt[3]{k}}$

b) $\displaystyle\sum_{k=1}^{\infty} \frac{(-4)^k}{k^4}$

c) $\displaystyle\sum_{k=1}^{\infty} \frac{(-2)^k}{k!}$

d) $\displaystyle\sum_{k=1}^{\infty} \frac{\cos 3k}{k^3}$

e) $\displaystyle\sum_{k=1}^{\infty} (-1)^k \frac{k}{k^2+3}$

f) $\displaystyle\sum_{k=1}^{\infty} \frac{\sin k\frac{\pi}{6}}{k\sqrt{k}}$

g) $\displaystyle\sum_{k=1}^{\infty} \frac{(-1)^{k+1}5^{k-1}}{(k+1)^2\,4^{k+2}}$

h) $\displaystyle\sum_{k=1}^{\infty} \frac{10^k}{(k+2)\,5^{2k+1}}$

15. *Study the convergence of the series:*

a) $\displaystyle\sum_{k=1}^{\infty}\left(1-\cos\frac{1}{k^3}\right)$

$\boxed{\text{b)}}$ $\displaystyle\sum_{k=1}^{\infty}\frac{\sin k}{k^2}$

c) $\displaystyle\sum_{k=1}^{\infty}\frac{1}{k^3}\binom{k}{2}$

$\boxed{\text{d)}}$ $\displaystyle\sum_{k=1}^{\infty}(-1)^k\left(\sqrt[k]{2}-1\right)$

$\boxed{\text{e)}}$ $\displaystyle\sum_{k=1}^{\infty}\frac{(-1)^{k-1}k!}{1\cdot 3\cdot 5\cdots(2k-1)}$

f) $\displaystyle\sum_{k=1}^{\infty}(-1)^k\frac{3k-1}{2k+1}$

16. *Verify the following series converge and then compute the sum:*

a) $\displaystyle\sum_{k=1}^{\infty}(-1)^k\frac{2^{k-1}}{5^k}$

$\boxed{\text{b)}}$ $\displaystyle\sum_{k=1}^{\infty}\frac{3^k}{2\cdot 4^{2k}}$

$\boxed{\text{c)}}$ $\displaystyle\sum_{k=1}^{\infty}\frac{2k+1}{k^2(k+1)^2}$

d) $\displaystyle\sum_{k=0}^{\infty}\frac{1}{(2k+1)(2k+3)}$

1.6.1 Solutions

1. *General terms and limits:*

a) $a_n=\dfrac{n}{n+1}\,,\ n\geq 1\,,\qquad \lim\limits_{n\to\infty}a_n=1$

b) $a_n=(-1)^n\dfrac{n+1}{3^n}\,,\ n\geq 1\,,\qquad \lim\limits_{n\to\infty}a_n=0$

c) $a_n=\sqrt{n}\,,\ n\geq 0\,,\qquad \lim\limits_{n\to\infty}a_n=+\infty$

2. *Sequences' behaviour and limit:*

a) Diverges to $+\infty$.
b) Converges to $\frac{1}{2}$.
c) Converges to 6.
d) Converges to 1.
e) Diverges to $+\infty$.

f) Since $\lim\limits_{n\to\infty}\dfrac{n^2}{n^2+1}=1$, the sequence is indeterminate because

$$\lim_{n\to\infty}a_{2n}=\lim_{n\to\infty}-\frac{(2n)^2}{(2n)^2+1}=-1\,,\qquad \lim_{n\to\infty}a_{2n+1}=\lim_{n\to\infty}\frac{(2n+1)^2}{(2n+1)^2+1}=1\,.$$

g) Converges to $\frac{\pi}{2}$.

h) Recalling that $\cos n\pi=(-1)^n$, we conclude immediately that the series is indeterminate.

i) Since $\{\sin\frac{1}{n}\}_{n\geq 1}$ is infinitesimal and $\{(-1)^n\}_{n\geq 1}$ bounded, we have

$$\lim_{n\to\infty}(-1)^n\sin\frac{1}{n}=0\,,$$

hence the given sequence converges to 1.

ℓ) Since
$$\left|\frac{n\cos n}{n^3+1}\right| \le \frac{n}{n^3+1} \le \frac{n}{n^3} = \frac{1}{n^2}\,, \qquad \forall n \ge 1\,,$$
by the Comparison Test we have
$$\lim_{n\to\infty} \frac{n\cos n}{n^3+1} = 0\,.$$

m) Converges to 0.

n) We have
$$\frac{\log(2+\mathrm{e}^n)}{4n} = \frac{\log \mathrm{e}^n(1+2\mathrm{e}^{-n})}{4n} = \frac{1}{4} + \frac{\log(1+2\mathrm{e}^{-n})}{4n}$$
so $\lim\limits_{n\to\infty} a_n = \dfrac{1}{4}$.

o) Diverges to $-\infty$.

p) Converges to 0.

3. *Sequences' behaviour:*

a) Diverges to $+\infty$.

b) Indeterminate.

c) Recalling the geometric sequence (Example 1.1 i)), we have
$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} \frac{4^n\left((\frac{3}{4})^n - 1\right)}{4^n(4^{-n}+1)} = -1\,;$$
and the convergence to -1 follows.

d) Diverges to $+\infty$.

e) Let us write
$$a_n = \frac{2n(2n-1)\cdots(n+2)(n+1)}{n(n+1)\cdots 2\cdot 1} = \frac{2n}{n}\cdot\frac{2n-1}{n-1}\cdots\frac{n+2}{2}\cdot\frac{n+1}{1} > n+1\,;$$
as $\lim\limits_{n\to\infty}(n+1) = +\infty$, the Second Comparison Theorem (infinite case), implies the sequence diverges to $+\infty$.

f) Converges to 1.

g) Since
$$a_n = \exp\left(\sqrt{n^2+2}\,\log\frac{n^2-n+1}{n^2+n+2}\right)\,,$$
we consider the sequence
$$b_n = \sqrt{n^2+2}\,\log\frac{n^2-n+1}{n^2+n+2} = \sqrt{n^2+2}\,\log\left(1-\frac{2n+1}{n^2+n+2}\right)\,.$$
Note that
$$\lim_{n\to\infty}\frac{2n+1}{n^2+n+2} = 0\,,$$
so
$$\log\left(1-\frac{2n+1}{n^2+n+2}\right) \sim -\frac{2n+1}{n^2+n+2}\,, \qquad n\to\infty\,.$$

Thus

$$\lim_{n\to\infty} b_n = -\lim_{n\to\infty} \frac{\sqrt{n^2+2}\,(2n+1)}{n^2+n+2} = -\lim_{n\to\infty} \frac{2n^2}{n^2} = -2\,;$$

and the sequence $\{a_n\}$ converges to e^{-2}.

h) Setting $x = 2^{-n}\pi$, we have $x \to 0^+$ for $n \to \infty$, so

$$\lim_{n\to\infty} a_n = \lim_{x\to 0^+} \pi\,\frac{\sin x}{x} = \pi$$

and $\{a_n\}$ tends to π.

i) Observe

$$\cos \frac{n+1}{n}\frac{\pi}{2} = \cos\left(\frac{\pi}{2} + \frac{\pi}{2n}\right) = -\sin\frac{\pi}{2n}\,;$$

therefore, setting $x = \frac{\pi}{2n}$, we have

$$\lim_{n\to\infty} a_n = -\lim_{n\to\infty} n \sin\frac{\pi}{2n} = -\lim_{x\to 0^+} \frac{\pi}{2}\frac{\sin x}{x} = -\frac{\pi}{2}$$

so $\{a_n\}$ converges to $-\pi/2$.

ℓ) Converges to $-1/2$.

4. *Series' convergence and computation of the sum:*

a) Converges with sum 6.

b) Note

$$\lim_{k\to\infty} a_k = \lim_{k\to\infty} \frac{2k}{k+5} = 2 \neq 0\,.$$

Hence the series does not converge, in fact it diverges to $+\infty$.

c) Does not converge.

d) The series is telescopic; we have

$$\begin{aligned} s_n &= (\sin 1 - \sin\frac{1}{2}) + (\sin\frac{1}{2} - \sin\frac{1}{3}) + \cdots + (\sin\frac{1}{n} - \sin\frac{1}{n+1}) \\ &= \sin 1 - \sin\frac{1}{n+1}\,. \end{aligned}$$

As $\lim\limits_{n\to\infty} s_n = \sin 1$, the series converges with sum $\sin 1$.

e) Because

$$\sum_{k=0}^{\infty} \frac{3^k + 2^k}{6^k} = \sum_{k=0}^{\infty}\left(\frac{3}{6}\right)^k + \sum_{k=0}^{\infty}\left(\frac{2}{6}\right)^k = \frac{1}{1-\frac{1}{2}} + \frac{1}{1-\frac{1}{3}} = \frac{7}{2}\,,$$

the series converges to $7/2$.

f) Does not converge.

5. Write

$$2.3\overline{17} = 2.3 + \frac{17}{10^3} + \frac{17}{10^5} + \frac{17}{10^7} + \ldots = 2.3 + \frac{17}{10^3}\left(1 + \frac{1}{10^2} + \frac{1}{10^4} + \ldots\right)$$

$$= 2.3 + \frac{17}{10^3}\sum_{k=0}^{\infty}\frac{1}{10^{2k}} = 2.3 + \frac{17}{10^3}\,\frac{1}{1-\frac{1}{10^2}} = \frac{23}{10} + \frac{17}{1000}\,\frac{100}{99}$$

$$= \frac{23}{10} + \frac{17}{990} = \frac{1147}{495}\,.$$

6. *Series' convergence and computation of the sum:*

a) Converges for $|x| < 5$ and the sum is $s = \frac{x^2}{5(5-x)}$.

b) This geometric series has $q = 3(x+2)$, so it converges if $|3(x+2)| < 1$, i.e., if $x \in (-\frac{7}{3}, -\frac{5}{3})$. For x in this range the sum is

$$s = \frac{1}{1-3(x+2)} - 1 = -\frac{3x+6}{3x+5}\,.$$

c) Converges for $x \in (-\infty, -1) \cup (1, +\infty)$ with sum $s = \frac{1}{x-1}$.

d) This is a geometric series where $q = \tan x$: it converges if $|\tan x| < 1$, that is if $x \in \bigcup_{k\in\mathbb{Z}} \left(-\frac{\pi}{4} + k\pi, \frac{\pi}{4} + k\pi\right)$. For such x, the sum is $s = \frac{1}{1-\tan x}$.

7. This is a geometric series with $q = \frac{1}{1+c}$, which converges for $|1+c| > 1$, i.e., for $c < -2$ or $c > 0$. If so,

$$\sum_{k=2}^{\infty}(1+c)^{-k} = \frac{1}{1-\frac{1}{1-c}} - 1 - \frac{1}{1-c} = \frac{1}{c(1+c)}\,.$$

Imposing $\frac{1}{c(1+c)} = 2$, we obtain $c = \frac{-1\pm\sqrt{3}}{2}$. But as the parameter c varies within $(-\infty, -2) \cup (0, +\infty)$, the only admissible value is $c = \frac{-1+\sqrt{3}}{2}$.

8. As $\displaystyle\sum_{k=1}^{\infty} a_k$ converges, the necessary condition $\displaystyle\lim_{k\to\infty} a_k = 0$ must hold. Therefore $\displaystyle\lim_{k\to\infty}\frac{1}{a_k}$ is not allowed to be 0, so the series $\displaystyle\sum_{k=1}^{\infty}\frac{1}{a_k}$ cannot converge.

9. *Convergence of positive-term series:*

a) Converges.

b) The general term a_k tends to $+\infty$ as $k \to \infty$. By Property 1.6 the series diverges to $+\infty$. Alternatively, one could invoke the Root Test 1.15.

c) By the Ratio Test 1.14:

$$\lim_{k\to\infty} \frac{a_{k+1}}{a_k} = \lim_{k\to\infty} \frac{3^{k+1}}{(k+1)!} \frac{k!}{3^k};$$

writing $(k+1)! = (k+1)k!$ and simplifying, we get

$$\lim_{k\to\infty} \frac{a_{k+1}}{a_k} = \lim_{k\to\infty} \frac{3}{k+1} = 0.$$

The series then converges.

d) Using again the Ratio Test 1.14:

$$\lim_{k\to\infty} \frac{a_{k+1}}{a_k} = \lim_{k\to\infty} \frac{(k+1)!}{(k+1)^{k+1}} \cdot \frac{k^k}{k!} = \lim_{k\to\infty} \left(\frac{k}{k+1}\right)^k = \frac{1}{\mathrm{e}} < 1$$

tells that the series converges.

e) As

$$a_k \sim k\frac{7}{k^2} = \frac{7}{k} \qquad \text{for} \qquad k \to \infty,$$

we conclude that the series diverges, by the Asymptotic Comparison Test 1.12 and the fact that the harmonic series diverges.

f) Converges.

g) Note $\log k > 1$ for $k \geq 3$, so that

$$\frac{\log k}{k} > \frac{1}{k}, \qquad k \geq 3.$$

The Comparison Test 1.10 guarantees divergence.
Alternatively, we may observe that the function $f(x) = \frac{\log x}{x}$ is positive and continuous for $x > 1$. The sign of the first derivative shows f is decreasing when $x > \mathrm{e}$. We can therefore use the Integral Test 1.17:

$$\begin{aligned}\int_3^{+\infty} \frac{\log x}{x}\,\mathrm{d}x &= \lim_{c\to+\infty} \int_3^c \frac{\log x}{x}\,\mathrm{d}x = \lim_{c\to+\infty} \frac{(\log x)^2}{2}\Big|_3^c \\ &= \lim_{c\to+\infty} \frac{(\log c)^2}{2} - \frac{(\log 3)^2}{2} = +\infty,\end{aligned}$$

then conclude that the given series diverges.

h) Converges by the Asymptotic Comparison Test 1.12, because

$$\frac{1}{2^k - 1} \sim \frac{1}{2^k}, \qquad k \to +\infty,$$

and the geometric series $\displaystyle\sum_{k=1}^{\infty} \frac{1}{2^k}$ converges.

i) Diverges by the Asymptotic Comparison Test 1.12, because

$$\sin\frac{1}{k} \sim \frac{1}{k}\,, \qquad k \to +\infty\,,$$

and the harmonic series $\displaystyle\sum_{k=1}^{\infty}\frac{1}{k}$ diverges.

ℓ) Diverges.

m) Converges.

n) Converges by the Comparison Test 1.10, as

$$\frac{\cos^2 k}{k\sqrt{k}} \le \frac{1}{k\sqrt{k}}\,, \qquad k \to +\infty$$

and the generalised harmonic series $\displaystyle\sum_{k=1}^{\infty}\frac{1}{k^{3/2}}$ converges.

10. Converges for $p > 1$.

11. Compute $\displaystyle\int_n^{+\infty} f(x)\,\mathrm{d}x$ where $f(x) = \frac{1}{x^2+4}$ is a positive, decreasing and continuous map on $[0, +\infty)$:

$$\int_n^{+\infty}\frac{1}{x^2+4}\,\mathrm{d}x = \frac{1}{2}\Big[\arctan\frac{x}{2}\Big]_n^{+\infty} = \frac{\pi}{4} - \frac{1}{2}\arctan\frac{n}{2}\,.$$

Since

$$s_6 = \frac{1}{4} + \frac{1}{5} + \ldots + \frac{1}{40} = 0.7614\,,$$

using (1.8)

$$s_6 + \int_7^{+\infty} f(x)\,\mathrm{d}x \le s \le s_6 + \int_6^{+\infty} f(x)\,\mathrm{d}x\,,$$

and we find $0.9005 \le s \le 0.9223$.

12. *Convergence of alternating series:*

a) Converges conditionally.

b) Does not converge.

c) Since

$$\sin\Big(k\pi + \frac{1}{k}\Big) = \cos(k\pi)\,\sin\frac{1}{k} = (-1)^k\sin\frac{1}{k}\,,$$

the series is alternating, with $b_k = \sin\frac{1}{k}$. Then

$$\lim_{k\to\infty} b_k = 0 \qquad \text{and} \qquad b_{k+1} < b_k\,.$$

By Leibniz's Test 1.20, the series converges. It does not converge absolutely since $\sin\frac{1}{k} \sim \frac{1}{k}$ for $k \to \infty$, so the series of absolute values is like a harmonic series, which diverges.

d) Converges absolutely: by the relationship $(1+x)^\alpha - 1 \sim \alpha x$, for $x \to 0$, it follows
$$\left|(-1)^k\Big(\big(1+\frac{1}{k^2}\big)^{\sqrt{2}}-1\Big)\right| \sim \frac{\sqrt{2}}{k^2}\,, \qquad k\to\infty\,.$$
Bearing in mind Example 1.11 i), we may apply the Asymptotic Comparison Test 1.12 to the series of absolute values.

e) Does not converges.

f) This is an alternating series with $b_k = \dfrac{k^2}{k^3+1}$. It is straightforward to check
$$\lim_{k\to\infty} b_k = 0\,.$$
That the sequence b_k is decreasing eventually is, instead, far from obvious. To show this fact, consider the map
$$f(x) = \frac{x^2}{x^3+1}\,,$$
and consider its monotonicity. Since
$$f'(x) = \frac{x(2-x^3)}{(x^3+1)^2}$$
and we are only interested in x positive, we have $f'(x) < 0$ if $2-x^3<0$, i.e., $x > \sqrt[3]{2}$. Therefore, f decreases on the interval $(\sqrt[3]{2},+\infty)$. This means $f(k+1) < f(k)$, so $b_{k+1} < b_k$ for $k \geq 2$. In conclusion, the series converges by Leibniz's Test 1.20.

13. *Series' approximation:*

a) $n = 5$.

b) The series is alternating with $b_k = \frac{2^k}{k!}$. Immediately we see $\lim\limits_{k\to\infty} b_k = 0$, and $b_{k+1} < b_k$ for any $k > 1$ since
$$b_{k+1} = \frac{2^{k+1}}{(k+1)!} < \frac{2^k}{k!} = b_k \qquad\Longleftrightarrow\qquad \frac{2}{k+1} < 1 \qquad\Longleftrightarrow\qquad k>1\,.$$
Imposing $b_{n+1} < 10^{-2} = 0.01$, one may check that
$$b_7 = \frac{8}{315} = 0.02\,, \qquad b_8 = \frac{2}{315} = 0.006 < 0.01\,.$$
The minimum number of terms needed is $n = 7$.

c) $n = 5$.

14. *Absolute convergence:*

a) There is convergence but not absolute convergence. In fact, the alternating series converges by Leibniz's Test 1.20, whereas the series of absolute values is generalised harmonic with exponent $\alpha = \frac{1}{3} < 1$.

b) Does not converge.

c) The series converges absolutely, as one sees easily by the Ratio Test 1.14, for example, since

$$\lim_{k\to\infty} \frac{b_{k+1}}{b_k} = \lim_{k\to\infty} \frac{2^{k+1}}{(k+1)!} \cdot \frac{k!}{2^k} = \lim_{k\to\infty} \frac{2}{k+1} = 0 < 1\,.$$

d) Convergence is absolute, since the series of absolute values converges by the Comparison Test 1.10:

$$\left|\frac{\cos 3k}{k^3}\right| \le \frac{1}{k^3}\,, \qquad \forall k \ge 1\,.$$

e) Converges, but not absolutely.

f) Converges absolutely.

g) Does not converge since its general term does not tend to 0.

h) Converges absolutely.

15. *Convergence of series:*

a) Converges.

b) Observe

$$\left|\frac{\sin k}{k^2}\right| \le \frac{1}{k^2}\,, \qquad \text{for all } k > 0\,;$$

the series $\sum_{k=1}^{\infty} \frac{1}{k^2}$ converges, so the Comparison Test 1.10 forces the series of absolute values to converge, too. Thus the given series converges absolutely.

c) Diverges.

d) This is an alternating series, with $b_k = \sqrt[k]{2}-1$. The sequence $\{b_k\}_{k\ge1}$ decreases, as $\sqrt[k]{2} > \sqrt[k+1]{2}$ for any $k \ge 1$. Thus we can use Leibniz's Test 1.20 to infer convergence. The series does not converge absolutely, for

$$\sqrt[k]{2} - 1 = \mathrm{e}^{\log 2/k} - 1 \sim \frac{\log 2}{k}\,, \qquad k \to \infty\,,$$

just like the harmonic series, which diverges.

e) Note

$$b_k = \frac{k!}{1 \cdot 3 \cdot 5 \cdots (2k-1)} = 1 \cdot \frac{2}{3} \cdot \frac{3}{5} \cdots \frac{k}{2k-1} < \left(\frac{2}{3}\right)^{k-1}$$

since

$$\frac{k}{2k-1} < \frac{2}{3}\,, \qquad \forall k \geq 2\,.$$

The convergence is then absolute, because $\sum_{k=1}^{\infty} b_k$ converges by the Comparison Test 1.10 (it is bounded by a geometric series with $q = \frac{2}{3} < 1$).

f) Does not converge.

16. *Checking series' convergence and computing the sum:*

a) $-1/7$.

b) Up to a factor this is a geometric series; by Example 1.3 iii) we have

$$\sum_{k=1}^{\infty} \frac{3^k}{2 \cdot 4^{2k}} = \frac{1}{2} \sum_{k=1}^{\infty} \left(\frac{3}{16}\right)^k = \frac{1}{2}\left(\frac{1}{1-\frac{3}{16}} - 1\right) = \frac{3}{26}$$

(notice that the sum starts from index 1).

c) It is a telescopic series because

$$\frac{2k+1}{k^2(k+1)^2} = \frac{1}{k^2} - \frac{1}{(k+1)^2}\,;$$

so

$$s_n = 1 - \frac{1}{(n+1)^2}\,,$$

and then $s = \lim_{n\to\infty} s_n = 1$.

d) $1/2$.

2

Series of functions and power series

The idea of approximating a function by a sequence of simple functions, or known ones, lies at the core of several mathematical techniques, both theoretical and practical. For instance, to prove that a differential equation has a solution one can construct recursively a sequence of approximating functions and show they converge to the required solution. At the same time, explicitly finding the values of such a solution may not be possible, not even by analytical methods, so one idea is to adopt numerical methods instead, which can furnish approximating functions with a particularly simple form, like piecewise polynomials. It becomes thus crucial to be able to decide when a sequence of maps generates a limit function, what sort of convergence towards the limit we have, and which features of the functions in the sequence are inherited by the limit. All this will be the content of the first part of this chapter.

The recipe for passing from a sequence of functions to the corresponding series is akin to what we have seen for a sequence of real numbers; the additional complication consists in the fact that now different kinds of convergence can occur. Expanding a function in series represents one of the most important tools of Mathematical Analysis and its applications, again both from the theoretical and practical point of view. Fundamental examples of series of functions are given by power series, discussed in the second half of this chapter, and by Fourier series, to which the whole subsequent chapter is dedicated. Other instances include series of the classical orthogonal functions, like the expansions in Legendre, Chebyshev or Hermite polynomials, Bessel functions, and so on.

In contrast to the latter cases, which provide a global representation of a function over an interval of the real line, power series have a more local nature; in fact, a power series that converges on an interval $(x_0 - R, x_0 + R)$ represents the limit function therein just by using information on its behaviour on an arbitrarily small neighbourhood of x_0. The power series of a functions may actually be thought of as a Taylor expansion of infinite order, centred at x_0. This fact reflects the intuitive picture of a power series as an algebraic polynomial of infinite degree, that is a sum of infinitely many monomials of increasing degree. In the final sections we will address the problem of rigorously determining the convergence set of a

C. Canuto, A. Tabacco: *Mathematical Analysis II,* 2nd Ed.,
UNITEXT – La Matematica per il 3+2 85, DOI 10.1007/978-3-319-12757-6_2,

power series; then we will study the main properties of power series and of series of functions, called analytic functions, that can be expanded in series on a real interval that does not reduce to a point.

2.1 Sequences of functions

Let X be an arbitrary subset of the real line. Suppose that there is a real map defined on X, which we denote $f_n : X \to \mathbb{R}$, for any n larger or equal than a certain $n_0 \geq 0$. The family $\{f_n\}_{n \geq n_0}$ is said **sequence of functions**. Examples are the families $f_n(x) = \sin(x + \frac{1}{n})$, $n \geq 1$ or $f_n(x) = x^n$, $n \geq 0$, on $X = \mathbb{R}$.

As for numerical sequences, we are interested in the study of the behaviour of a sequence of maps as $n \to \infty$. The first step is to analyse the numerical sequence given by the values of the maps f_n at each point of X.

Definition 2.1 *The sequence* $\{f_n\}_{n \geq n_0}$ **converges pointwise** *at* $\overline{x} \in X$ *if the numerical sequence* $\{f_n(\overline{x})\}_{n \geq n_0}$ *converges as* $n \to \infty$. *The subset* $A \subseteq X$ *of such points* $\overline{x}$ *is called* **set of pointwise convergence** *of the sequence* $\{f_n\}_{n \geq n_0}$. *This defines a map* $f : A \to \mathbb{R}$ *by*

$$f(x) = \lim_{n \to \infty} f_n(x) \,, \qquad \forall x \in A \,.$$

We shall write $f_n \to f$ *pointwise on* A, *and speak of the* **limit function** f *of the sequence.*

Note f is the limit function of the sequence if and only if

$$\lim_{n \to \infty} |f_n(x) - f(x)| = 0 \,, \qquad \forall x \in A \,.$$

Examples 2.2

i) Let $f_n(x) = \sin\left(x + \frac{1}{n}\right)$, $n \geq 1$, on $X = \mathbb{R}$. Observing that $x + \frac{1}{n} \to x$ as $n \to \infty$, and using the sine's continuity, we have

$$f(x) = \lim_{n \to \infty} \sin\left(x + \frac{1}{n}\right) = \sin x \,, \qquad \forall x \in \mathbb{R} \,,$$

hence $A = X = \mathbb{R}$.

ii) Consider $f_n(x) = x^n$, $n \geq 0$, on $X = \mathbb{R}$; recalling (1.1) we have

$$f(x) = \lim_{n \to \infty} x^n = \begin{cases} 0 & \text{if } -1 < x < 1 \,, \\ 1 & \text{if } x = 1 \,. \end{cases} \tag{2.1}$$

The sequence converges for no other value x, so $A = (-1, 1]$. □

The notion of pointwise convergence is not sufficient, in many cases, to transfer the properties of the single maps f_n onto the limit function f. Continuity (but also differentiability, or integrability) is one such case. In the above examples the maps $f_n(x)$ are continuous, but the limit function is continuous in case i), not for ii).

A more compelling convergence requirement, that warrants continuity is passed onto the limit, is the so-called uniform convergence. To understand the difference with pointwise convergence, let us make Definition 2.1 explicit. This states that for any point $x \in A$ and any $\varepsilon > 0$ there is an integer $\overline{n}$ such that

$$\forall n \geq n_0\,, \;\; n > \overline{n} \quad \Longrightarrow \quad |f_n(x) - f(x)| < \varepsilon\,.$$

In general $\overline{n}$ depends not only upon ε but also on x, i.e., $\overline{n} = \overline{n}(\varepsilon, x)$. In other terms the index $\overline{n}$, after which the values $f_n(x)$ approximate $f(x)$ with a margin smaller than ε, may vary from point to point. For example consider $f_n(x) = x^n$, with $0 < x < 1$; then the condition

$$|f_n(x) - f(x)| = |x^n - 0| = x^n < \varepsilon$$

holds for any $n > \frac{\log \varepsilon}{\log x}$. Therefore the smallest n for which the condition is valid tends to infinity as x approaches 1. Hence there is no $\overline{n}$ depending on ε and not on x.

The convergence is said uniform whenever the index $\overline{n}$ can be chosen independently of x. This means that, for any $\varepsilon > 0$, there must be an integer $\overline{n}$ such that

$$\forall n \geq n_0\,, n > \overline{n} \quad \Longrightarrow \quad |f_n(x) - f(x)| < \varepsilon\,, \qquad \forall x \in A\,.$$

Using the notion of supremum and infimum, and recalling ε is arbitrary, we can reformulate as follows: for any $\varepsilon > 0$ there is an integer $\overline{n}$ such that

$$\forall n \geq n_0\,, n > \overline{n} \quad \Longrightarrow \quad \sup_{x \in A} |f_n(x) - f(x)| < \varepsilon\,.$$

Definition 2.3 *The sequence* $\{f_n\}_{n \geq n_0}$ **converges uniformly** *on A to the limit function f if*

$$\lim_{n \to \infty} \sup_{x \in A} |f_n(x) - f(x)| = 0\,.$$

Otherwise said, for any $\varepsilon > 0$ there is an $\overline{n} = \overline{n}(\varepsilon)$ such that

$$\forall n \geq n_0\,, \;\; n > \overline{n} \quad \Longrightarrow \quad |f_n(x) - f(x)| < \varepsilon\,, \qquad \forall x \in A\,. \tag{2.2}$$

We will write $f_n \to f$ uniformly on A.

Let us introduce the symbol

$$\|g\|_{\infty,A} = \sup_{x \in A} |g(x)|\,,$$

for a bounded map $g : A \to \mathbb{R}$; this quantity is variously called *infinity norm*, *supremum norm*, or *sup-norm* for short (see Appendix A.2.1, p. 521, for a comprehensive presentation of the concept of norm of a function). An alternative definition of uniform convergence is thus

$$\lim_{n \to \infty} \|f_n - f\|_{\infty,A} = 0\,. \tag{2.3}$$

Clearly, the uniform convergence of a sequence is a stronger condition than pointwise convergence. By definition of sup-norm in fact,

$$\forall x \in A\,, \qquad |f_n(x) - f(x)| \le \|f_n - f\|_{\infty,A}$$

so if the norm on the right tends to 0, so does the absolute value on the left. Therefore

Proposition 2.4 *If the sequence $\{f_n\}_{n\ge n_0}$ converges to f uniformly on A, it converges pointwise on A.*

The converse is false, as some examples show.

Examples 2.5

i) The sequence $f_n(x) = \sin\left(x + \frac{1}{n}\right)$, $n \ge 1$, converges uniformly to $f(x) = \sin x$ on $\mathbb{R}$. In fact, using known trigonometric identities, we have, for any $x \in \mathbb{R}$,

$$\left|\sin\left(x+\frac{1}{n}\right) - \sin x\right| = 2\left|\sin\frac{1}{2n}\right|\left|\cos\left(x+\frac{1}{2n}\right)\right| \le 2\sin\frac{1}{2n};$$

moreover, equality is attained for $x = -\frac{1}{2n}$, for example. Therefore

$$\|f_n - f\|_{\infty,\mathbb{R}} = \sup_{x\in\mathbb{R}}\left|\sin\left(x+\frac{1}{n}\right) - \sin x\right| = 2\sin\frac{1}{2n}$$

and passing to the limit for $n \to \infty$ we obtain the result (see Fig. 2.1, left).

ii) As already seen, $f_n(x) = x^n$, $n \ge 0$, does not converge uniformly on $I = [0,1]$ to f defined in (2.1). For any $n \ge 0$ in fact, $\|f_n - f\|_{\infty,I} = \sup\limits_{0\le x<1} x^n = 1$ (Fig. 2.1, right). Nevertheless, the convergence is uniform on every sub-interval $I_a = [0,a]$, $0 < a < 1$, for

$$\|f_n - f\|_{\infty,I_a} = \sup_{x\in[0,a]} |x^n - 0| = a^n \to 0 \quad \text{as } n \to \infty\,.$$

Therefore the sequence converges to zero uniformly on I_a. More generally one can show the sequence converges uniformly to zero on any interval $[-a,a]$, $0 < a < 1$.
□

The following criterion is immediate to check, and useful for verifying uniform convergence.

Proposition 2.6 *Let the sequence $\{f_n\}_{n\ge n_0}$ converge pointwise on A to a function f. Take a numerical sequence $\{M_n\}_{n\ge n_0}$, infinitesimal for $n \to \infty$, such that*

$$|f_n(x) - f(x)| \le M_n\,, \qquad \forall x \in A\,.$$

Then $f_n \to f$ uniformly on A.

The property has been used in the previous examples, with $M_n = 2\sin\frac{1}{2n}$ for case i), and $M_n = a^n$ for ii).

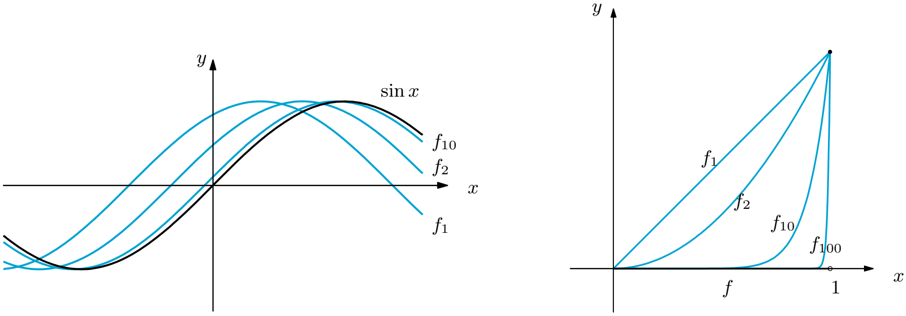


Figure 2.1. Graphs of the functions f_n and their limit f relative to Examples 2.5 i) (left) and ii) (right)

2.2 Properties of uniformly convergent sequences

As announced earlier, under uniform convergence the limit function inherits continuity from the sequence.

Theorem 2.7 *Let the sequence of continuous maps $\{f_n\}_{n\geq n_0}$ converge uniformly to f on the real interval I. Then f is continuous on I.*

Proof. By uniform convergence, given $\varepsilon > 0$, there is an $\overline{n} = \overline{n}(\varepsilon) \geq n_0$ such that for any $n > \overline{n}$ and any $x \in I$

$$|f_n(x) - f(x)| < \frac{\varepsilon}{3}\,.$$

Fix $x_0 \in I$ and take $n > \overline{n}$. As f_n is continuous at x_0, there is a $\delta > 0$ such that, for each $x \in I$ with $|x - x_0| < \delta$,

$$|f_n(x) - f_n(x_0)| < \frac{\varepsilon}{3}\,.$$

Let therefore $x \in I$ with $|x - x_0| < \delta$. Then

$$|f(x) - f(x_0)| \leq |f(x) - f_n(x)| + |f_n(x) - f_n(x_0)| + \\ + |f_n(x_0) - f(x_0)| < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon\,,$$

so f is continuous at x_0. □

This result can be used to say that pointwise convergence is not always uniform. In fact if the limit function is not continuous while the single terms in the sequence are, the convergence cannot be uniform.

2.2.1 Interchanging limits and integrals

Suppose $f_n \to f$ pointwise on $I = [a,b]$. If the maps are integrable, it is not true, in general, that

$$\int_a^b f_n(x)\,\mathrm{d}x \to \int_a^b f(x)\,\mathrm{d}x\,.$$

Example 2.8

Let $f_n(x) = xn^2\mathrm{e}^{-nx}$ on $I = [0,1]$. Then $f_n(x) \to 0 = f(x)$, as $n \to \infty$, pointwise on I. Therefore $\int_0^1 f(x)\,\mathrm{d}x = 0$; on the other hand, setting $\varphi(t) = t\mathrm{e}^{-t}$ we have

$$\int_0^1 f_n(x)\,\mathrm{d}x = \int_0^1 \varphi(nx)n\,\mathrm{d}x = \int_0^n \varphi(t)\,\mathrm{d}t = -n\mathrm{e}^{-n} + \big[-\mathrm{e}^{-t}\big]_0^n$$
$$= -n\mathrm{e}^{-n} - \mathrm{e}^{-n} + 1 \to 1 \quad \text{for } n \to \infty\,. \qquad \square$$

Uniform convergence is a sufficient condition for transferring integrability to the limit.

Theorem 2.9 *Let $I = [a,b]$ be a closed, bounded interval and $\{f_n\}_{n\ge n_0}$ a sequence of integrable functions over I such that $f_n \to f$ uniformly on I. Then f is integrable on I, and*

$$\int_a^b f_n(x)\,\mathrm{d}x \to \int_a^b f(x)\,\mathrm{d}x \qquad \textit{as } n \to \infty\,. \tag{2.4}$$

Proof. The integrability of the limit function is immediate if each f_n is continuous, for in that case f itself is continuous by the previous theorem. In general, one needs to approximate the functions by means of step functions, as prescribed by the definition of an integrable map; the details are left to the reader's good will.
In order to prove (2.4), let us fix $\varepsilon > 0$; then there exists an $\overline{n} = \overline{n}(\varepsilon) \ge n_0$ such that, for any $n > \overline{n}$ and any $x \in I$,

$$|f_n(x) - f(x)| < \frac{\varepsilon}{b-a}\,.$$

Therefore, for all $n > \overline{n}$, we have

$$\left|\int_a^b f_n(x)\,\mathrm{d}x - \int_a^b f(x)\,\mathrm{d}x\right| = \left|\int_a^b \big(f_n(x) - f(x)\big)\,\mathrm{d}x\right|$$
$$\le \int_a^b |f_n(x) - f(x)|\,\mathrm{d}x < \int_a^b \frac{\varepsilon}{b-a}\,\mathrm{d}x = \varepsilon\,. \qquad \square$$

Note that (2.4) can be written as

$$\boxed{\lim_{n\to\infty} \int_a^b f_n(x)\,\mathrm{d}x = \int_a^b \lim_{n\to\infty} f_n(x)\,\mathrm{d}x\,,}$$

showing that uniform convergence allows to exchange the operations of limit and integration.

2.2.2 Interchanging limits and derivatives

When considering limits of differentiable functions, certain assumptions on uniform convergence guarantee the differentiability of the limit.

Theorem 2.10 *Let $\{f_n\}_{n\geq n_0}$ be a sequence of $\mathcal{C}^1$ functions over the interval $I = [a,b]$. Suppose there exist maps f and g on I such that*

i) $f_n \to f$ *pointwise on* I*;*
ii) $f_n' \to g$ *uniformly on* I*.*

Then f is $\mathcal{C}^1$ on I, and $f' = g$. Moreover, $f_n \to f$ uniformly on I (and clearly, $f_n' \to f'$ uniformly on I).

Proof. Fix an arbitrary $x_0 \in I$ and set

$$\tilde{f}(x) = f(x_0) + \int_{x_0}^{x} g(t)\,\mathrm{d}t\,. \tag{2.5}$$

We show first that $f_n \to \tilde{f}$ uniformly on I. For this, let $\varepsilon > 0$ be given. By *i)*, there exists $n_1 = n_1(\varepsilon; x_0) \geq n_0$ such that for any $n > n_1$ we have

$$|f_n(x_0) - f(x_0)| < \frac{\varepsilon}{2}\,.$$

By *ii)*, there is $n_2 = n_2(\varepsilon) \geq n_0$ such that, for any $n > n_2$ and any $t \in [a,b]$,

$$|f_n'(t) - g(t)| < \frac{\varepsilon}{2(b-a)}\,.$$

Note we may write each map f_n as

$$f_n(x) = f_n(x_0) + \int_{x_0}^{x} f_n'(t)\,\mathrm{d}t\,,$$

because of the Fundamental Theorem of Integral Calculus (see Vol. I, Cor. 9.42). Thus, for any $n > \overline{n} = \max(n_1, n_2)$ and any $x \in [a,b]$, we have

$$|f_n(x) - \tilde{f}(x)| = \left| f_n(x_0) - f(x_0) + \int_{x_0}^{x} \big(f_n'(t) - g(t)\big)\,\mathrm{d}t \right|$$

$$\leq |f_n(x_0) - f(x_0)| + \int_{x_0}^{x} |f'_n(t) - g(t)| \, \mathrm{d}t$$

$$< \frac{\varepsilon}{2} + \frac{\varepsilon}{2(b-a)} \int_a^b \mathrm{d}t = \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \, .$$

Therefore $f_n \to \tilde{f}$ uniformly on I, hence also pointwise, by Proposition 2.4. But $f_n \to f$ pointwise on I by assumption; by the uniqueness of the limit then, $\tilde{f}$ coincides with f. From (2.5) and the Fundamental Theorem of Integral Calculus it follows that f is differentiable with first derivative g. □

Under the theorem's assumptions then,

$$\boxed{\lim_{n\to\infty} f'_n(x) = \Big(\lim_{n\to\infty} f_n(x)\Big)' \, , \qquad \forall x \in I \, ,}$$

so the uniform convergence (of the first derivatives) allows to exchange the limit and the derivative.

Remark 2.11 A more general result, with similar proof, states that if a sequence $\{f_n\}_{n \geq n_0}$ of $\mathcal{C}^1$ maps satisfies these two properties:

i) there is an $x_0 \in [a, b]$ such that $\lim\limits_{n\to\infty} f_n(x_0) = \ell \in \mathbb{R}$;
ii) the sequence $\{f'_n\}_{n \geq n_0}$ converges uniformly on I to a map g (necessarily continuous on I),

then, setting

$$f(x) = \ell + \int_{x_0}^{x} g(t) \, \mathrm{d}t \, ,$$

f_n converges uniformly to f on I. Furthermore, $f \in \mathcal{C}^1$ and $f'(x) = g(x)$ for any $x \in [a, b]$. □

Remark 2.12 An example will explain why mere pointwise convergence of the derivatives is not enough to conclude as in the theorem, even in case the sequence of functions converges uniformly. Consider the sequence $f_n(x) = x - x^n/n$; it converges uniformly on $I = [0, 1]$ to $f(x) = x$ because

$$|f_n(x) - f(x)| = \left| \frac{x^n}{n} \right| \leq \frac{1}{n} \, , \qquad \forall x \in I \, ,$$

and hence

$$\|f_n - f\|_{\infty, I} \leq \frac{1}{n} \to 0 \quad \text{as } n \to \infty \, .$$

Yet the derivatives $f'_n(x) = 1 - x^{n-1}$ converge to the discontinuous function

$$g(x) = \begin{cases} 1 & \text{if } x \in [0,1)\,, \\ 0 & \text{if } x = 1\,. \end{cases}$$

So $\{f'_n\}_{n\geq 0}$ converges only pointwise to g on $[0,1]$, not uniformly, and the latter does not coincide on I with the derivative of f. □

2.3 Series of functions

Starting with a sequence of functions $\{f_k\}_{k\geq k_0}$ defined on a set $X \subseteq \mathbb{R}$, we can build a series of functions $\displaystyle\sum_{k=k_0}^{\infty} f_k$ in a similar fashion to numerical sequences and series. Precisely, we consider how the sequence of partial sums

$$s_n(x) = \sum_{k=k_0}^{n} f_k(x)$$

behaves as $n \to \infty$. Different types of convergence can occur.

Definition 2.13 *The series of functions* $\displaystyle\sum_{k=k_0}^{\infty} f_k$ **converges pointwise** *at* $\overline{x}$ *if the sequence of partial sums* $\{s_n\}_{n\geq k_0}$ *converges pointwise at* $\overline{x}$*; equivalently, the numerical series* $\displaystyle\sum_{k=k_0}^{\infty} f_k(\overline{x})$ *converges. Let* $A \subseteq X$ *be the set of such points* $\overline{x}$*, called the* **set of pointwise convergence** *of* $\displaystyle\sum_{k=k_0}^{\infty} f_k$*; we have thus defined the function* $s : A \to \mathbb{R}$*, called* **sum***, by*

$$s(x) = \lim_{n\to\infty} s_n(x) = \sum_{k=k_0}^{\infty} f_k(x)\,, \qquad \forall x \in A\,.$$

The pointwise convergence of a series of functions can be studied using at each point $x \in X$ what we already know about numerical series. In particular, the sequence $\{f_k(x)\}_{k\geq k_0}$ must be infinitesimal, as $k \to \infty$, in order for x to belong to A. What is more, the convergence criteria seen in the previous chapter can be applied, at each point.

Definition 2.14 *The series of functions* $\displaystyle\sum_{k=k_0}^{\infty} f_k$ **converges absolutely** *on* A *if for any* $x \in A$ *the series* $\displaystyle\sum_{k=k_0}^{\infty} |f_k(x)|$ *converges.*

Definition 2.15 *The series of functions* $\sum_{k=k_0}^{\infty} f_k$ **converges uniformly** *to the function* s *on* A *if the sequence of partial sums* $\{s_n\}_{n\geq k_0}$ *converges uniformly to* s *on* A.

Both the absolute convergence (due to Theorem 1.24) and the uniform convergence (Proposition 2.4) imply the pointwise convergence of the series. There are no logical implications between uniform and absolute convergence, instead.

Example 2.16

The series $\sum_{k=0}^{\infty} x^k$ is nothing but the geometric series of Example 1.3 where q is taken as independent variable and re-labelled x. Thus, the series converges pointwise to the sum $s(x) = \frac{1}{1-x}$ on $A = (-1, 1)$; on the same set there is absolute convergence as well. As for uniform convergence, it holds on every closed interval $[-a, a]$ with $0 < a < 1$. In fact,

$$|s_n(x) - s(x)| = \left| \frac{1 - x^{n+1}}{1 - x} - \frac{1}{1 - x} \right| = \frac{|x|^{n+1}}{1 - x} \leq \frac{a^{n+1}}{1 - a},$$

where we have used the fact that $|x| \leq a$ implies $1 - a \leq 1 - x$. Moreover, the sequence $M_n = \frac{a^{n+1}}{1-a}$ tends to 0 as $n \to \infty$, and the result follows from Proposition 2.6. □

It is clear from the definitions just given that Theorems 2.7, 2.9 and 2.10 can be formulated for series of functions, so we re-phrase them for completeness' sake.

Theorem 2.17 *Let* $\{f_k\}_{k\geq k_0}$ *be a sequence of continuous maps on a real interval* I *such that the series* $\sum_{k=k_0}^{\infty} f_k$ *converges uniformly to a function* s *on* I. *Then* s *is continuous on* I.

Theorem 2.18 (Integration by series) *Let* $I = [a, b]$ *be a closed bounded interval and* $\{f_k\}_{k\geq k_0}$ *a sequence of integrable functions on* I *such that* $\sum_{k=k_0}^{\infty} f_k$ *converges uniformly to a function* s *on* I. *Then* s *is integrable on* I, *and*

$$\int_a^b s(x)\,\mathrm{d}x = \int_a^b \sum_{k=k_0}^{\infty} f_k(x)\,\mathrm{d}x = \sum_{k=k_0}^{\infty} \int_a^b f_k(x)\,\mathrm{d}x\,. \tag{2.6}$$

This is worded alternatively by saying that the series is integrable term by term.

Theorem 2.19 (**Differentiation of series**) *Let* $\{f_k\}_{k \geq k_0}$ *be a sequence of* $\mathcal{C}^1$ *maps on* $I = [a, b]$*. Suppose there are maps* s *and* t *on* I *such that*

i) $\displaystyle\sum_{k=k_0}^{\infty} f_k(x) = s(x)\,, \quad \forall x \in I;$

ii) $\displaystyle\sum_{k=k_0}^{\infty} f'_k(x) = t(x)\,, \quad \forall x \in I$ *and the convergence is uniform on* I.

Then $s \in \mathcal{C}^1(I)$ *and* $s' = t$*. Furthermore,* $\displaystyle\sum_{k=k_0}^{\infty} f_k$ *converges uniformly to* s *on* I *(and* $\displaystyle\sum_{k=k_0}^{\infty} f'_k$ *converges uniformly to* s'*).*

That is to say,

$$\boxed{\sum_{k=k_0}^{\infty} f'_k(x) = \left(\sum_{k=k_0}^{\infty} f_k(x)\right)', \qquad \forall x \in I\,,}$$

and the series is differentiable term by term.

The importance of uniform convergence should be clear by now. But checking uniform convergence is another matter, often far from easy. Using the definition requires knowing the sum, and as we have seen with numerical series, the sum is not always computable explicitly. For this reason we will prove a condition sufficient to guarantee the uniform convergence of a series, even without knowing its sum in advance.

Theorem 2.20 (**Weierstrass' M-test**) *Let* $\{f_k\}_{k \geq k_0}$ *be a sequence of maps on* X *and* $\{M_k\}_{k \geq k_0}$ *a sequence of real numbers such that, for any* $k \geq k_0$,

$$|f_k(x)| \leq M_k\,, \qquad \forall x \in X\,.$$

Assume the numerical series $\displaystyle\sum_{k=k_0}^{\infty} M_k$ *converges. Then the series* $\displaystyle\sum_{k=k_0}^{\infty} f_k$ *converges uniformly on* X.

Proof. Fix $x \in X$, so that the numerical series $\sum_{k=k_0}^{\infty} |f_k(x)|$ converges by the Comparison Test, and hence the sum

$$s(x) = \sum_{k=k_0}^{\infty} f_k(x)\,, \qquad \forall x \in X\,,$$

is well defined. It suffices to check whether the partial sums $\{s_n\}_{n \geq k_0}$ converge uniformly to s on X. But for any $x \in X$,

$$|s_n(x) - s(x)| = \left| \sum_{k=n+1}^{\infty} f_k(x) \right| \leq \sum_{k=n+1}^{\infty} |f_k(x)| \leq \sum_{k=n+1}^{\infty} M_k\,,$$

i.e.,

$$\sup_{x \in X} |s_n(x) - s(x)| \leq \sum_{k=n+1}^{\infty} M_k\,.$$

As the series $\sum_{k=k_0}^{\infty} M_k$ converges, the right-hand side is just the nth remainder of a converging series, which goes to 0 as $n \to \infty$.
In conclusion,

$$\lim_{n \to \infty} \sup_{x \in X} |s_n(x) - s(x)| = 0$$

so $\sum_{k=k_0}^{\infty} f_k$ converges uniformly on X. □

Example 2.21

We want to understand the uniform and pointwise convergence of

$$\sum_{k=1}^{\infty} \frac{\sin k^4 x}{k\sqrt{k}}\,, \qquad x \in \mathbb{R}\,.$$

Note

$$\left| \frac{\sin k^4 x}{k\sqrt{k}} \right| \leq \frac{1}{k\sqrt{k}}\,, \qquad \forall x \in \mathbb{R}\,;$$

we may then use the M-test with $M_k = \frac{1}{k\sqrt{k}}$, since the series $\sum_{k=1}^{\infty} \frac{1}{k^{3/2}}$ converges (it is generalised harmonic of exponent $3/2$). Therefore the given series converges uniformly, hence also pointwise, on $\mathbb{R}$. □

2.4 Power series

Power series are very special series in which the maps f_k are polynomials. More precisely,

Definition 2.22 *Fix $x_0 \in \mathbb{R}$ and let $\{a_k\}_{k\geq 0}$ be a numerical sequence. One calls* **power series** *a series of the form*

$$\sum_{k=0}^{\infty} a_k(x-x_0)^k = a_0 + a_1(x-x_0) + a_2(x-x_0)^2 + \cdots \tag{2.7}$$

The point x_0 is said **centre** *of the series and the numbers $\{a_k\}_{k\geq 0}$ are the series'* **coefficients**.

The series converges at its centre, irrespective of the coefficients.

The next three examples exhaust the possible types of convergence set of a series. We will show that such set is always an interval (possibly shrunk to the centre).

Examples 2.23

i) The series

$$\sum_{k=1}^{\infty} k^k x^k = x + 4x^2 + 27x^3 + \cdots$$

converges only at $x = 0$; in fact at any other $x \neq 0$ the general term $k^k x^k$ is not infinitesimal, so the necessary condition for convergence is not fulfilled (Property 1.6).

ii) Consider

$$\sum_{k=0}^{\infty} \frac{x^k}{k!} = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots\,; \tag{2.8}$$

it is known as **exponential series**, because it sums to the function $s(x) = \mathrm{e}^x$. This fact will be proved later in Example 2.46 i).

The exponential series converges for any $x \in \mathbb{R}$. Indeed, with a given $x \neq 0$, the Ratio Test for numerical series (Theorem 1.14) guarantees convergence:

$$\lim_{k\to\infty} \left| \frac{x^{k+1}}{(k+1)!} \cdot \frac{k!}{x^k} \right| = \lim_{k\to\infty} \frac{|x|}{k+1} = 0\,, \qquad \forall x \in \mathbb{R} \setminus \{0\}\,.$$

iii) Another familiar example is the geometric series

$$\sum_{k=0}^{\infty} x^k = 1 + x + x^2 + x^3 + \cdots$$

(recall Example 2.16). We already know it converges, for $x \in (-1, 1)$, to the function $s(x) = \frac{1}{1-x}$. □

In all examples the series converge (absolutely) on a symmetric interval with respect to the centre (the origin, in the specific cases). We will see that the convergence set A of any power series, independently of the coefficients, is either a bounded interval (open, closed or half-open) centered at x_0, or the whole $\mathbb{R}$.

We start by series centered at the origin; this is no real restriction, because the substitution $y = x - x_0$ allows to reduce to that case.

Before that though, we need a technical result, direct consequence of the Comparison Test for numerical series (Theorem 1.10), which will be greatly useful for power series.

Proposition 2.24 *If the series* $\sum_{k=0}^{\infty} a_k \overline{x}^k$, $\overline{x} \neq 0$, *has bounded terms, in particular if it converges, then the power series* $\sum_{k=0}^{\infty} a_k x^k$ *converges absolutely for any* x *such that* $|x| < |\overline{x}|$.

Proof. As $a_k \overline{x}_k$ is bounded, there is a constant $M > 0$ such that

$$|a_k \overline{x}_k| \le M\,, \qquad \forall k \ge 0\,.$$

For any x with $|x| < |\overline{x}|$ then,

$$|a_k x^k| = \left| a_k \overline{x}^k \left(\frac{x}{\overline{x}}\right)^k \right| \le M \left|\frac{x}{\overline{x}}\right|^k, \qquad \forall k \ge 0\,.$$

But $|x| < |\overline{x}|$, so the geometric series $\sum_{k=0}^{\infty} \left(\frac{x}{\overline{x}}\right)^k$ converges absolutely and, by the Comparison Test, $\sum_{k=0}^{\infty} a_k x^k$ converges absolutely. □

Example 2.25

The series $\sum_{k=0}^{\infty} \frac{k-1}{k+1} x^k$ has bounded terms when $x = \pm 1$, since $\left|\frac{k-1}{k+1}\right| \le 1$ for any $k \ge 0$. The above proposition forces convergence when $|x| < 1$. The series does not converge when $|x| \le 1$ because, in case $x = \pm 1$, the general term is not infinitesimal. □

Proposition 2.24 has an immediate, yet crucial, consequence.

Corollary 2.26 *If a power series* $\sum_{k=0}^{\infty} a_k x^k$ *converges at* $x_1 \neq 0$, *it converges absolutely on the open interval* $(-|x_1|, |x_1|)$; *if it does not converge at* $x_2 \neq 0$, *it does not converge anywhere along the half-lines* $(|x_2|, +\infty)$ *and* $(-\infty, -|x_2|)$.

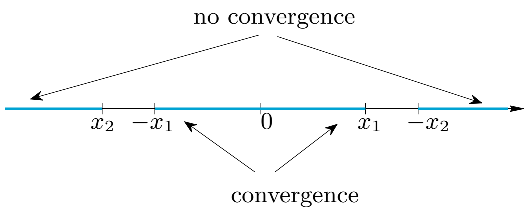


Figure 2.2. Illustration of Corollary 2.26

The statement is depicted in Fig. 2.2, for $x_1 > 0$ and $x_2 < 0$.

Now we are in a position to prove that the convergence set of a power series is a symmetric interval, end-points excluded.

Theorem 2.27 *Given a power series* $\sum_{k=0}^{\infty} a_k x^k$, *only one of the following holds:*

a) the series converges at $x = 0$ *only;*

b) the series converges pointwise and absolutely for any $x \in \mathbb{R}$*; moreover, it converges uniformly on every closed and bounded interval* $[a, b]$*;*

c) there is a unique real number $R > 0$ *such that the series converges pointwise and absolutely for any* $|x| < R$*, and uniformly on all intervals* $[a, b] \subset (-R, R)$*. Furthermore, the series does not converge on* $|x| > R$*.*

Proof. Let A denote the set of convergence of $\sum_{k=0}^{\infty} a_k x^k$.
If $A = \{0\}$, we have case a).
Case b) occurs if $A = \mathbb{R}$. In fact, Corollary 2.26 tells the series converges pointwise and absolutely for any $x \in \mathbb{R}$. As for the uniform convergence on $[a, b]$, set $L = \max(|a|, |b|)$. Then

$$|f_k(x)| = |a_k x^k| \leq |a_k L^k| \,, \qquad \forall x \in [a, b]\,;$$

and we may use Weierstrass' M-test 2.20 with $M_k = |a_k| L^k$.
Now suppose A contains points other than 0 but is smaller that the whole line, so there is an $\overline{x} \notin A$. Corollary 2.26 says A cannot contain any x with $|x| > |\overline{x}|$, meaning that A is bounded. Set $R = \sup A$, so $R > 0$ because A is larger than $\{0\}$. Consider an arbitrary x with $|x| < R$: by definition of supremum there is an x_1 such that $|x| < x_1 < R$ and $\sum_{k=0}^{\infty} a_k x_1^k$ converges.
Hence Corollary 2.26 tells the series converges pointwise and absolutely at x. For uniform convergence we proceed exactly as in case b). At last, by definition of sup the set A cannot contain values $x > R$, but neither values

$x < -R$ (again by Corollary 2.26). Thus if $|x| > R$ the series $\sum_{k=0}^{\infty} a_k x^k$ does not converge. □

Definition 2.28 *One calls* **convergence radius** *of the series* $\sum_{k=0}^{\infty} a_k x^k$ *the number*

$$R = \sup \Big\{ x \in \mathbb{R} : \sum_{k=0}^{\infty} a_k x^k \text{ converges} \Big\}.$$

Going back to Theorem 2.27, we remark that $R = 0$ in case a); in case b), $R = +\infty$, while in case c), R is precisely the strictly-positive real number of the statement.

Examples 2.29

Let us return to Examples 2.23.

i) The series $\sum_{k=1}^{\infty} k^k x^k$ has convergence radius $R = 0$.

ii) For $\sum_{k=0}^{\infty} \frac{x^k}{k!}$ the radius is $R = +\infty$.

iii) The series $\sum_{k=0}^{\infty} x^k$ has radius $R = 1$. □

Beware that the theorem says nothing about the behaviour at $x = \pm R$: the series might converge at both end-points, at one only, or at none, as in the next examples.

Examples 2.30

i) The series

$$\sum_{k=1}^{\infty} \frac{x^k}{k^2}$$

converges at $x = \pm 1$ (generalised harmonic of exponent 2 for $x = 1$, alternating for $x = -1$). It does not converge on $|x| > 1$, as the general term is not infinitesimal. Thus $R = 1$ and $A = [-1, 1]$.

ii) The series

$$\sum_{k=1}^{\infty} \frac{x^k}{k}$$

converges at $x = -1$ (alternating harmonic series) but not at $x = 1$ (harmonic series). Hence $R = 1$ and $A = [-1, 1)$.

iii) The geometric series

$$\sum_{k=1}^{\infty} x^k$$

converges only on $A = (-1, 1)$ with radius $R = 1$. □

Convergence at one end-point ensures the series converges uniformly on closed intervals containing that end-point. Precisely, we have

Theorem 2.31 (**Abel**) *Suppose $R > 0$ is finite. If the series converges at $x = R$, then the convergence is uniform on every interval $[a, R] \subset (-R, R]$. The analogue statement holds if the series converges at $x = -R$.*

If we now center a power series at a generic x_0, the previous results read as follows. The radius R is 0 if and only if $\sum_{k=0}^{\infty} a_k (x - x_0)^k$ converges only at x_0, while $R = +\infty$ if and only if the series converges at any x in $\mathbb{R}$. In the remaining case R is positive and finite, and Theorem 2.27 says the set A of convergence satisfies

$$\{x \in \mathbb{R} : |x - x_0| < R\} \subseteq A \subseteq \{x \in \mathbb{R} : |x - x_0| \leq R\}\,.$$

The importance of determining the radius of convergence is evident. The next two criteria, easy consequences of the analogous Ratio and Root Tests for numerical series, give a rather simple yet useful answer.

Theorem 2.32 (**Ratio Test**) *Given the power series*

$$\sum_{k=0}^{\infty} a_k (x - x_0)^k$$

with $a_k \neq 0$ for all $k \geq 0$, if the limit

$$\lim_{k \to \infty} \left| \frac{a_{k+1}}{a_k} \right| = \ell$$

exists, the radius of convergence R is given by

$$R = \begin{cases} 0 & \text{if } \ell = +\infty\,, \\ +\infty & \text{if } \ell = 0\,, \\ \dfrac{1}{\ell} & \text{if } 0 < \ell < +\infty\,. \end{cases} \tag{2.9}$$

Proof. For simplicity suppose $x_0 = 0$, and let $x \neq 0$. The claim follows by the Ratio Test 1.14 since

$$\lim_{k\to\infty} \left| \frac{a_{k+1}x^{k+1}}{a_k x^k} \right| = \lim_{k\to\infty} \left| \frac{a_{k+1}}{a_k} \right| |x| = \ell|x| \,.$$

When $\ell = +\infty$, we have $\ell|x| > 1$ and the series does not converge for any $x \neq 0$, so $R = 0$; when $\ell = 0$, $\ell|x| = 0 < 1$ and the series converges for any x, so $R = +\infty$. At last, when ℓ is finite and non-zero, the series converges for all x such that $\ell|x| < 1$, so for $|x| < 1/\ell$, and not for $|x| > 1/\ell$; therefore $R = 1/\ell$. □

Theorem 2.33 (**Root Test**) *Given the power series*

$$\sum_{k=0}^{\infty} a_k (x - x_0)^k \,,$$

if the limit

$$\lim_{k\to\infty} \sqrt[k]{|a_k|} = \ell$$

exists, the radius R is given by formula (2.9).

The proof, left to the reader, relies on the Root Test 1.15 and follows the same lines.

Examples 2.34

i) The series $\displaystyle\sum_{k=0}^{\infty} kx^k$ has radius $R = 1$, because $\displaystyle\lim_{k\to\infty} \sqrt[k]{k} = 1$; it does not converge for $x = 1$ nor for $x = -1$.

ii) Consider

$$\sum_{k=0}^{\infty} \frac{k!}{k^k} x^k$$

and use the Ratio Test:

$$\lim_{k\to\infty} \frac{(k+1)!}{(k+1)^{k+1}} \cdot \frac{k^k}{k!} = \lim_{k\to\infty} \left(\frac{k}{k+1} \right)^k = \lim_{k\to\infty} \left(1 + \frac{1}{k} \right)^{-k} = \mathrm{e}^{-1} \,.$$

The radius is thus $R = \mathrm{e}$.

iii) To study

$$\sum_{k=2}^{\infty} \frac{2k+1}{(k-1)(k+2)} (x-2)^{2k} \tag{2.10}$$

set $y = (x-2)^2$ and consider the power series in y

$$\sum_{k=2}^{\infty} \frac{2k+1}{(k-1)(k+2)} y^k \tag{2.11}$$

centred at the origin. Since

$$\lim_{k\to\infty} \sqrt[k]{\frac{2k+1}{(k-1)(k+2)}} = 1$$

the radius is 1. For $y = 1$ the series (2.11) reduces to

$$\sum_{k=2}^{\infty} \frac{2k+1}{(k-1)(k+2)},$$

which diverges like the harmonic series ($\frac{2k+1}{(k-1)(k+2)} \sim \frac{2}{k}$, $k \to \infty$), whereas for $y = -1$ the series (2.11) converges (by Leibniz's Test 1.20). In summary, (2.11) converges for $-1 \le y < 1$.
Going back to the variable x, that means $-1 \le (x-2)^2 < 1$. The left inequality is always true, while the right one holds for $-1 < x-2 < 1$. So, the series (2.10) has radius $R = 1$ and converges on the interval $(1, 3)$ (note the centre is $x_0 = 2$).

iv) The series

$$\sum_{k=0}^{\infty} x^{k!} = x + x + x^2 + x^6 + x^{24} + \cdots$$

is a power series where infinitely many coefficients are 0, and we cannot substitute as we did before; in such cases the aforementioned criteria do not apply, and it is more convenient to use directly the Ratio or Root Test for numerical series. In the case at hand

$$\begin{aligned}\lim_{k\to\infty} \left| \frac{x^{(k+1)!}}{x^{k!}} \right| &= \lim_{k\to\infty} |x|^{(k+1)!-k!} \\ &= \lim_{k\to\infty} |x|^{k!k} = \begin{cases} 0 & \text{if } |x| < 1\,, \\ +\infty & \text{if } |x| > 1\,. \end{cases}\end{aligned}$$

Thus $R = 1$. The series converges neither for $x = 1$, nor for $x = -1$.

v) Consider, for $\alpha \in \mathbb{R}$, the **binomial series**

$$\sum_{k=0}^{\infty} \binom{\alpha}{k} x^k .$$

If $\alpha = n \in \mathbb{N}$ the series is actually a finite sum, and Newton's binomial formula (Vol. I, Eq. (1.13)) tells us that

$$\sum_{k=0}^{n} \binom{n}{k} x^k = (1+x)^n ,$$

hence the name. Let us then study for $\alpha \in \mathbb{R} \setminus \mathbb{N}$ and observe

$$\frac{\left|\binom{\alpha}{k+1}\right|}{\left|\binom{\alpha}{k}\right|} = \frac{|\alpha(\alpha-1)\cdots(\alpha-k)|}{(k+1)!} \cdot \frac{k!}{|\alpha(\alpha-1)\cdots(\alpha-k+1)|} = \frac{|\alpha-k|}{k+1} ;$$

therefore

$$\lim_{k\to\infty} \frac{\left|\binom{\alpha}{k+1}\right|}{\left|\binom{\alpha}{k}\right|} = \lim_{k\to\infty} \frac{|\alpha - k|}{k+1} = 1$$

and the series has radius $R = 1$. The behaviour of the series at the endpoints cannot be studied by one of the criteria presented above; one can prove that the series converges at $x = -1$ only for $\alpha > 0$ and at $x = 1$ only for $\alpha > -1$. □

2.4.1 Algebraic operations

The operations of sum and product of two polynomials extend in a natural manner to power series centred at the same point x_0; the problem remains of determining the radius of convergence of the resulting series. This is addressed in the next theorems, where x_0 will be 0 for simplicity.

Theorem 2.35 *Given* $\Sigma_1 = \sum_{k=0}^{\infty} a_k x^k$ *and* $\Sigma_2 = \sum_{k=0}^{\infty} b_k x^k$, *of respective radii* R_1, R_2, *their sum* $\Sigma = \sum_{k=0}^{\infty} (a_k + b_k) x^k$ *has radius* R *satisfying* $R \geq \min(R_1, R_2)$. *If* $R_1 \neq R_2$, *necessarily* $R = \min(R_1, R_2)$.

Proof. Suppose $R_1 \neq R_2$; we may assume $R_1 < R_2$. Given any point x such that $R_1 < x < R_2$, if the series Σ converged we would have

$$\Sigma_1 = \sum_{k=0}^{\infty} a_k x^k = \sum_{k=0}^{\infty} (a_k + b_k) x^k - \sum_{k=0}^{\infty} b_k x^k = \Sigma - \Sigma_2\,,$$

hence also the series Σ_1 would have to converge, contradicting the fact that $x > R_1$. Therefore $R = R_1 = \min(R_1, R_2)$.
In case $R_1 = R_2$, the radius R is at least equal to such value, since the sum of two convergent series is convergent (see Sect. 1.5). □

In case $R_1 = R_2$, the radius R might be strictly larger than both R_1, R_2 due to possible cancellations of terms in the sum.

Example 2.36

The series

$$\Sigma_1 = \sum_{k=1}^{\infty} \frac{2^k + 1}{4^k - 2^k}\, x^k \qquad \text{and} \qquad \Sigma_2 = \sum_{k=1}^{\infty} \frac{1 - 2^k}{4^k + 2^k}\, x^k$$

have the same radius $R_1 = R_2 = 2$. Their sum

$$\Sigma = \sum_{k=1}^{\infty} \frac{4}{4^k - 1}\, x^k\,,$$

though, has radius $R = 4$. □

The product of two series is defined so that to preserve the distributive property of the sum with respect to the multiplication. In other words,

$$(a_0 + a_1x + a_2x^2 + \ldots)(b_0 + b_1x + b_2x^2 + \ldots) \\ = a_0b_0 + (a_0b_1 + a_1b_0)x + (a_0b_2 + a_1b_1 + a_2b_0)x^2 + \ldots ,$$

i.e.,

$$\Big(\sum_{k=0}^{\infty} a_k x^k\Big)\Big(\sum_{k=0}^{\infty} b_k x^k\Big) = \sum_{k=0}^{\infty} c_k x^k \tag{2.12}$$

where

$$c_k = \sum_{j=0}^{k} a_j b_{k-j} .$$

This multiplication rule is called Cauchy product: putting $x = 1$ returns precisely the Cauchy product (1.9) of numerical series. Then we have the following result.

Theorem 2.37 *Given* $\Sigma_1 = \sum_{k=0}^{\infty} a_k x^k$ *and* $\Sigma_2 = \sum_{k=0}^{\infty} b_k x^k$, *of respective radii* R_1, R_2, *their Cauchy product has convergence radius* $R \geq \min(R_1, R_2)$.

2.4.2 Differentiation and integration

Let us move on to consider the regularity of the sum of a power series. We have already remarked that the functions $f_k(x) = a_k(x - x_0)^k$ are $\mathcal{C}^\infty$ polynomials over all of $\mathbb{R}$. In particular they are continuous and their sum $s(x)$ is continuous, where defined (using Theorem 2.17), because the convergence is uniform on closed intervals in the convergence set (Theorem 2.31). Let us see in detail how term-by-term differentiation and integration fit with power series. For clarity we assume $x_0 = 0$ and begin with a little technical fact.

Lemma 2.38 *The series* $\sum_1 = \sum_{k=0}^{\infty} a_k x^k$ *and* $\sum_2 = \sum_{k=0}^{\infty} k a_k x^k$ *have the same radius of convergence.*

Proof. Call R_1, R_2 the radii of $\sum_1, \sum_2$ respectively. Clearly $R_2 \leq R_1$ (for $|a_k x^k| \leq |k a_k x^k|$). On the other hand if $|x| < R_1$ and $\overline{x}$ satisfies $|x| < \overline{x} < R_1$, the series $\sum_{k=0}^{\infty} a_k \overline{x}^k$ converges and so $|a_k|\overline{x}^k$ is bounded from above by a constant $M > 0$ for any $k \geq 0$. Hence

$$|k a_k x^k| = k|a_k|\overline{x}^k \left|\frac{x}{\overline{x}}\right|^k \leq Mk \left|\frac{x}{\overline{x}}\right|^k .$$

Since $\displaystyle\sum_{k=0}^{\infty} k\left(\frac{x}{\bar{x}}\right)^k$ is convergent (Example 2.34 i)), by the Comparison Test 1.10 also $\displaystyle\sum_{k=0}^{\infty} k a_k x^k$ converges, whence $R_1 \leq R_2$. In conclusion $R_1 = R_2$ and the claim is proved. □

The series $\displaystyle\sum_{k=0}^{\infty} k a_k x^{k-1} = \sum_{k=1}^{\infty} k a_k x^{k-1} = \sum_{k=0}^{\infty}(k+1)a_{k+1}x^k$ is the **derivatives' series** of $\displaystyle\sum_{k=0}^{\infty} a_k x^k$.

Theorem 2.39 *Suppose the radius R of the series $\displaystyle\sum_{k=0}^{\infty} a_k x^k$ is positive, finite or infinite. Then*

a) the sum s is a $\mathcal{C}^\infty$ map over $(-R, R)$. Moreover, the nth derivative of s on $(-R, R)$ can be computed by differentiating $\displaystyle\sum_{k=0}^{\infty} a_k x^k$ term by term n times. In particular, for any $x \in (-R, R)$

$$s'(x) = \sum_{k=1}^{\infty} k a_k x^{k-1} = \sum_{k=0}^{\infty}(k+1)a_{k+1}x^k\,; \tag{2.13}$$

b) for any $x \in (-R, R)$

$$\int_0^x s(t)\,\mathrm{d}t = \sum_{k=0}^{\infty} \frac{a_k}{k+1} x^{k+1}\,. \tag{2.14}$$

Proof. a) By the previous lemma a power series and its derivatives' series have identical radii, because $\displaystyle\sum_{k=1}^{\infty} k a_k x^{k-1} = \frac{1}{x}\sum_{k=1}^{\infty} k a_k x^k$ for any $x \neq 0$. The derivatives' series converges uniformly on every interval $[a, b] \subset (-R, R)$; thus Theorem 2.19 applies, and we conclude that s is $\mathcal{C}^1$ on $(-R, R)$ and that (2.13) holds. Iterating the argument proves the claim.

b) The result follows immediately by noticing $\displaystyle\sum_{k=0}^{\infty} a_k x^k$ is the derivatives' series of $\displaystyle\sum_{k=0}^{\infty} \frac{a_k}{k+1} x^{k+1}$. These two have same radius and we can use Theorem 2.18. □

Example 2.40

Differentiating term by term

$$\sum_{k=0}^{\infty} x^k = \frac{1}{1-x}, \qquad x \in (-1,1), \tag{2.15}$$

we infer that for $x \in (-1,1)$

$$\sum_{k=1}^{\infty} kx^{k-1} = \sum_{k=0}^{\infty} (k+1)x^k = \frac{1}{(1-x)^2}. \tag{2.16}$$

Integrating term by term the series

$$\sum_{k=0}^{\infty} (-1)^k x^k = \frac{1}{1+x}, \qquad x \in (-1,1),$$

obtained from (2.15) by changing x to $-x$, we have for all $x \in (-1,1)$

$$\sum_{k=0}^{\infty} \frac{(-1)^k}{k+1} x^{k+1} = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} x^k = \log(1+x). \tag{2.17}$$

At last from

$$\sum_{k=0}^{\infty} (-1)^k x^{2k} = \frac{1}{1+x^2}, \qquad x \in (-1,1),$$

obtained from (2.15) writing $-x^2$ instead of x, and integrating each term separately, we see that for any $x \in (-1,1)$

$$\sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} x^{2k+1} = \arctan x. \tag{2.18}$$

Proposition 2.41 *Suppose* $\sum_{k=0}^{\infty} a_k (x-x_0)^k$ *has radius* $R > 0$*. The series' coefficients depend on the derivatives of the sum* $s(x)$ *as follows:*

$$a_k = \frac{1}{k!} s^{(k)}(x_0), \qquad \forall k \geq 0.$$

Proof. Write the sum as $s(x) = \sum_{h=0}^{\infty} a_h (x-x_0)^h$; differentiating each term k times gives

$$\begin{aligned} s^{(k)}(x) &= \sum_{h=k}^{\infty} h(h-1)\cdots(h-k+1) a_h (x-x_0)^{h-k} \\ &= \sum_{h=0}^{\infty} (h+k)(h+k-1)\cdots(h+1) a_{h+k} (x-x_0)^h. \end{aligned}$$

For $x = x_0$ only the term indexed by $h = 0$ contributes, and the above expression becomes

$$s^{(k)}(x_0) = k!\, a_k \,, \qquad \forall k \geq 0 \,.$$

□

2.5 Analytic functions

The previous section examined the properties of the sum of a power series, summarised in Theorem 2.39. Now we want to take the opposite viewpoint, and demand that an arbitrary function (necessarily $\mathcal{C}^\infty$) be the sum of some power series. Said better, we take $f \in \mathcal{C}^\infty(X)$, $X \subseteq \mathbb{R}$, $x_0 \in X$ and ask whether, on a suitable interval $(x_0 - \delta, x_0 + \delta) \subseteq X$ with $\delta > 0$, it is possible to represent f as the sum of a power series

$$f(x) = \sum_{k=0}^{\infty} a_k (x - x_0)^k \,; \tag{2.19}$$

by Proposition 2.41 we must necessarily have

$$a_k = \frac{f^{(k)}(x_0)}{k!} \,, \qquad \forall k \geq 0 \,.$$

In particular when $x_0 = 0$, f is given by

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} x^k \,. \tag{2.20}$$

Definition 2.42 *The series* (2.19) *is called the* **Taylor series** *of f centred at x_0. If the radius is positive and the sum coincides with f around x_0 (i.e., on some neighbourhood of x_0), one says the map f has a* **Taylor series expansion**, *or is an* **analytic** *function, at x_0. If $x_0 = 0$, one speaks sometimes of* **Maclaurin series** *of f.*

The definition is motivated by the fact that not all $\mathcal{C}^\infty$ functions admit a power series representation, as in this example.

Example 2.43

Consider

$$f(x) = \begin{cases} e^{-1/x^2} & \text{if } x \neq 0 \,, \\ 0 & \text{if } x = 0 \,. \end{cases}$$

It is not hard to check f is $\mathcal{C}^\infty$ on $\mathbb{R}$ with $f^{(k)}(0) = 0$ for all $k \geq 0$. Therefore the terms of (2.20) all vanish and the sum (the zero function) does not represent f anywhere around the origin. □

The partial sums in (2.19) are precisely the Taylor polynomials of f at x_0:

$$s_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k = Tf_{n,x_0}(x)\,.$$

Therefore f having a Taylor series expansion is equivalent to the convergence to f of the sequence of its own Taylor polynomials:

$$\lim_{n\to\infty} s_n(x) = \lim_{n\to\infty} Tf_{n,x_0}(x) = f(x)\,, \qquad \forall x \in (x_0 - \delta, x_0 + \delta)\,.$$

In such a case the nth remainder of the series $r_n(x) = f(x) - s_n(x)$ is infinitesimal, as $n \to \infty$, for any $x \in (x_0 - \delta, x_0 + \delta)$:

$$\lim_{n\to\infty} r_n(x) = 0\,.$$

There is a sufficient condition for a $\mathcal{C}^\infty$ map to have a Taylor series expansion around a point.

Theorem 2.44 *Take $f \in \mathcal{C}^\infty(x_0 - \delta, x_0 + \delta)$, $\delta > 0$. If there are an index $k_0 \geq 0$ and a constant $M > 0$ such that*

$$|f^{(k)}(x)| \leq M\frac{k!}{\delta^k}\,, \qquad \forall x \in (x_0 - \delta, x_0 + \delta) \tag{2.21}$$

for all $k \geq k_0$, then f has a Taylor series expansion at x_0 whose radius is at least δ.

Proof. Write the Taylor expansion of f at x_0 of order $n \geq k_0$ with Lagrange remainder (see Vol. I, Thm 7.2):

$$f(x) = Tf_{n,x_0}(x) + \frac{1}{(n+1)!}f^{(n+1)}(x_n)(x - x_0)^{n+1}\,,$$

where x_n is a certain point between x_0 and x. By assumption, for any $x \in (x_0 - \delta, x_0 + \delta)$ we have

$$|r_n(x)| = \frac{1}{(n+1)!}|f^{(n+1)}(x_n)|\,|x - x_0|^{n+1} \leq M\left(\frac{|x - x_0|}{\delta}\right)^{n+1}.$$

If we assume $|x - x_0|/\delta < 1$, then

$$\lim_{n\to\infty} r_n(x) = 0$$

and the claim is proved. □

Remark 2.45 Condition (2.21) holds in particular if all derivatives $f^{(k)}(x)$ are uniformly bounded, independently of k: this means there is a constant $M > 0$ for which

$$|f^{(k)}(x)| \leq M\,, \qquad \forall x \in (x_0 - \delta, x_0 + \delta)\,. \tag{2.22}$$

In fact, from Example 1.1 v) we have $\frac{k!}{\delta^k} \to \infty$ as $k \to \infty$, so $\frac{k!}{\delta^k} \geq 1$ for k bigger or equal than a certain k_0.

A similar argument shows that (2.21) is true more generally if

$$|f^{(k)}(x)| \leq M^k\,, \qquad \forall x \in (x_0 - \delta, x_0 + \delta)\,. \tag{2.23}$$

Examples 2.46

i) We can eventually prove the earlier claim on the exponential series, that is to say

$$\mathrm{e}^x = \sum_{k=0}^{\infty} \frac{x^k}{k!} \qquad \forall x \in \mathbb{R}\,. \tag{2.24}$$

We already know the series converges for any $x \in \mathbb{R}$ (Example 2.23 ii)); additionally, the map e^x is $\mathcal{C}^\infty$ on $\mathbb{R}$ with $f^{(k)}(x) = \mathrm{e}^x$, $f^{(k)}(0) = 1$. Fixing an arbitrary $\delta > 0$, inequality (2.22) holds since

$$|f^{(k)}(x)| = \mathrm{e}^x \leq \mathrm{e}^\delta = M\,, \qquad \forall x \in (-\delta, \delta)\,.$$

Hence f has a Maclaurin series and (2.24) is true, as promised.
More generally, e^x has a Taylor series at each $x_0 \in \mathbb{R}$:

$$\mathrm{e}^x = \sum_{k=0}^{\infty} \frac{\mathrm{e}^{x_0}}{k!}(x - x_0)^k\,, \qquad \forall x \in \mathbb{R}\,.$$

ii) Writing $-x^2$ instead of x in (2.24) yields

$$\mathrm{e}^{-x^2} = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{k!}\,, \qquad \forall x \in \mathbb{R}\,.$$

Integrating term by term we obtain a representation in series of the **error function**

$$\operatorname{erf} x = \frac{2}{\sqrt{\pi}} \int_0^x \mathrm{e}^{-t^2}\,\mathrm{d}t = \frac{2}{\sqrt{\pi}} \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \frac{x^{2k+1}}{2k+1}\,,$$

which has a role in Probability and Statistics.

iii) The trigonometric functions $f(x) = \sin x$, $g(x) = \cos x$ are analytic for any $x \in \mathbb{R}$. Indeed, they are $\mathcal{C}^\infty$ on $\mathbb{R}$ and all derivatives satisfy (2.22) with $M = 1$. In the special case $x_0 = 0$,

$$\sin x = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!} x^{2k+1}\,, \qquad \forall x \in \mathbb{R}\,, \tag{2.25}$$

$$\cos x = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!} x^{2k}\,, \qquad \forall x \in \mathbb{R}\,. \tag{2.26}$$

iv) Let us prove that for $\alpha \in \mathbb{R} \setminus \mathbb{N}$

$$(1+x)^\alpha = \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k, \qquad \forall x \in (-1,1). \tag{2.27}$$

In Example 2.34 v) we found the radius of convergence $R = 1$ for the right-hand side. Let $f(x)$ denote the sum of the series:

$$f(x) = \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k, \qquad \forall x \in (-1,1).$$

Differentiating term-wise and multiplying by $(1+x)$ gives

$$(1+x)f'(x) = (1+x)\sum_{k=1}^{\infty} k\binom{\alpha}{k} x^{k-1} = \sum_{k=0}^{\infty}(k+1)\binom{\alpha}{k+1} x^k + \sum_{k=1}^{\infty} k\binom{\alpha}{k} x^k$$

$$= \sum_{k=0}^{\infty}\left[(k+1)\binom{\alpha}{k+1} + k\binom{\alpha}{k}\right] x^k = \alpha \sum_{k=0}^{\infty}\binom{\alpha}{k} x^k = \alpha f(x).$$

Hence $f'(x) = \alpha(1+x)^{-1}f(x)$. Now take the map $g(x) = (1+x)^{-\alpha}f(x)$ and note

$$\begin{aligned} g'(x) &= -\alpha(1+x)^{-\alpha-1}f(x) + (1+x)^{-\alpha}f'(x) \\ &= -\alpha(1+x)^{-\alpha-1}f(x) + \alpha(1+x)^{-\alpha-1}f(x) = 0 \end{aligned}$$

for any $x \in (-1,1)$. Therefore $g(x)$ is constant and we can write

$$f(x) = c(1+x)^\alpha.$$

The value of the constant c is fixed by $f(0) = 1$, so $c = 1$.

v) When $\alpha = -1$, formula (2.27) gives the Taylor series of $f(x) = \frac{1}{1+x}$ at the origin:

$$\frac{1}{1+x} = \sum_{k=0}^{\infty}(-1)^k x^k.$$

Actually, f is analytic at all points $x_0 \neq -1$ and the corresponding Taylor series' radius is $R = |1+x_0|$, i.e., the distance of x_0 from the singular point $x = -1$; indeed, one has

$$f^{(k)}(x) = (-1)^k k!(1+x)^{-(k+1)}$$

and it is not difficult to check estimate (2.21) on a suitable neighbourhood of x_0. Furthermore, the Root Test (Theorem 2.33) gives

$$R = \lim_{k\to\infty} \sqrt[k]{|1+x_0|^{k+1}} = |1+x_0| > 0.$$

vi) One can prove the map $f(x) = \frac{1}{1+x^2}$ is analytic at each point $x_0 \in \mathbb{R}$. This does not mean, though, that the radius of the generic Taylor expansion of f is $+\infty$. For instance, at $x_0 = 0$,

$$f(x) = \sum_{k=0}^{\infty}(-1)^k x^{2k}$$

has radius 1. In general the Taylor series of f at x_0 will have radius $\sqrt{1+x_0^2}$ (see the next section).

vi) The last two instances are, as a matter of fact, rational, and it is known that rational functions – more generally all elementary functions – are analytic at every point lying in the interior of their domain. □

2.6 Power series in $\mathbb{C}$

The definition of power series extends easily to the complex numbers. By a power series in $\mathbb{C}$ we mean an expression like

$$\sum_{k=0}^{\infty} a_k (z - z_0)^k ,$$

where $\{a_k\}_{k\geq 0}$ is a sequence of complex numbers, $z_0 \in \mathbb{C}$ and z is the complex variable. The notions of convergence (pointwise, absolute, uniform) carry over provided we substitute everywhere the absolute value with the modulus.

The convergence interval of a real power series is now replaced by a disc in the complex plane, centred at z_0 and of radius $R \in [0, +\infty]$.

The term analytic map determines a function of one complex variable that is the sum of a power series in $\mathbb{C}$. Examples include rational functions of complex variable

$$f(z) = \frac{P(z)}{Q(z)} ,$$

where P, Q are coprime polynomials over the complex numbers; with $z_0 \in \operatorname{dom} f$ fixed, the convergence radius of the series centred at z_0 whose sum is f coincides with the distance (in the complex plane) between z_0 and the nearest zero of the denominator, i.e., the closest singularity.

The exponential, sine and cosine functions possess a natural extension to $\mathbb{C}$, obtained by substituting the real variable x with the complex z in (2.24), (2.25) and (2.26). These new series converge on the whole complex plane, so the corresponding functions are analytic on $\mathbb{C}$.

2.7 Exercises

1. *Over the given interval I determine the sets of pointwise and uniform convergence, and the limit function, for:*

   $\boxed{\text{a)}}$ $f_n(x) = \dfrac{nx}{1 + n^3x^3} ,\qquad I = [0, +\infty)$

   $\boxed{\text{b)}}$ $f_n(x) = \dfrac{x}{1 + n^2x^2} ,\qquad I = [0, +\infty)$

   $\boxed{\text{c)}}$ $f_n(x) = \mathrm{e}^{nx} ,\qquad I = \mathbb{R}$

d) $f_n(x) = nx\mathrm{e}^{-nx}\,, \qquad I = \mathbb{R}$

e) $f_n(x) = \dfrac{4^{nx}}{3^{nx} + 5^{nx}}\,, \qquad I = \mathbb{R}$

2. *Study uniform and pointwise convergence for the sequence of maps:*

$$f_n(x) = nx(1 - x^2)^n\,, \qquad x \in [-1, 1]\,.$$

Does the following formula hold?

$$\lim_{n\to\infty} \int_0^1 f_n(x)\,\mathrm{d}x = \int_0^1 \lim_{n\to\infty} f_n(x)\,\mathrm{d}x\,.$$

3. *Study uniform and pointwise convergence for the sequence of maps:*

$$f_n(x) = \arctan nx\,, \qquad x \in \mathbb{R}\,.$$

Tell whether the formula

$$\lim_{n\to\infty} \int_a^1 f_n(x)\,\mathrm{d}x = \int_a^1 \lim_{n\to\infty} f_n(x)\,\mathrm{d}x\,,$$

holds, with $a = 0$ or $a = 1/2$.

4. *Determine the sets of pointwise convergence of the series:*

a) $\displaystyle\sum_{k=1}^{\infty} \frac{(k+2)^x}{k^3 + \sqrt{k}}\,,$

b) $\displaystyle\sum_{k=1}^{\infty} \left(1 + \frac{x}{k}\right)^{k^2}$

c) $\displaystyle\sum_{k=1}^{\infty} \frac{1}{x^k + x^{-k}}\,, \qquad x > 0$

d) $\displaystyle\sum_{k=1}^{\infty} \frac{x^k}{x^k + 2^k}\,, \qquad x \neq -2$

e) $\displaystyle\sum_{k=1}^{\infty} (-1)^k k^x \sin\frac{1}{k}$

f) $\displaystyle\sum_{k=1}^{\infty} \left(k - \sqrt{k^2 - 1}\right)^x$

5. *Determine the sets of pointwise and uniform convergence of the series* $\displaystyle\sum_{k=2}^{\infty} \mathrm{e}^{kx}$*. Compute its sum, where defined.*

6. *Setting* $f_k(x) = \cos\frac{x}{k}$*, check* $\displaystyle\sum_{k=1}^{\infty} f_k'(x)$ *converges uniformly on* $[-1, 1]$*, while* $\displaystyle\sum_{k=1}^{\infty} f_k(x)$ *converges nowhere.*

7. *Determine the sets of pointwise and uniform convergence of the series:*

$\boxed{\text{a)}}$ $\displaystyle\sum_{k=1}^{\infty} k^{1/x}$ $\boxed{\text{b)}}$ $\displaystyle\sum_{k=1}^{\infty} \frac{(\log k)^x}{k}$ $\boxed{\text{c)}}$ $\displaystyle\sum_{k=1}^{\infty} \Big[(k^2+x^2)^{\frac{1}{k^2+x^2}} - 1\Big]$

$\boxed{8.}$ *Knowing that* $\displaystyle\sum_{k=0}^{\infty} a_k 4^k$ *converges, can one infer the convergence of the following series?*

a) $\displaystyle\sum_{k=0}^{\infty} a_k(-2)^k$ b) $\displaystyle\sum_{k=0}^{\infty} a_k(-4)^k$

9. *Suppose that* $\displaystyle\sum_{k=0}^{\infty} a_k x^k$ *converges for* $x = -4$ *and diverges for* $x = 6$*. What can be said about the convergence or divergence of the following series?*

a) $\displaystyle\sum_{k=0}^{\infty} a_k$ b) $\displaystyle\sum_{k=0}^{\infty} a_k 7^k$

c) $\displaystyle\sum_{k=0}^{\infty} a_k(-3)^k$ d) $\displaystyle\sum_{k=0}^{\infty}(-1)^k a_k 9^k$

$\boxed{10.}$ *Let* p *be a positive integer. Determine, as* p *varies, the radius of convergence of*

$$\sum_{k=0}^{\infty} \frac{(k!)^p}{(pk)!}\, x^k\,.$$

11. *Find radius and set of convergence of the power series:*

a) $\displaystyle\sum_{k=1}^{\infty} \frac{x^k}{\sqrt{k}}$ b) $\displaystyle\sum_{k=0}^{\infty} \frac{(-1)^k x^k}{k+1}$

c) $\displaystyle\sum_{k=0}^{\infty} k x^{2k}$ d) $\displaystyle\sum_{k=2}^{\infty}(-1)^k \frac{x^k}{3^k \log k}$

e) $\displaystyle\sum_{k=0}^{\infty} k^2(x-4)^k$ f) $\displaystyle\sum_{k=0}^{\infty} \frac{k^3(x-1)^k}{10^k}$

g) $\displaystyle\sum_{k=1}^{\infty}(-1)^k \frac{(x+3)^k}{k3^k}$ h) $\displaystyle\sum_{k=1}^{\infty} k!(2x-1)^k$

i) $\displaystyle\sum_{k=1}^{\infty} \frac{kx^k}{1\cdot 3\cdot 5\cdots(2k-1)}$ ℓ) $\displaystyle\sum_{k=1}^{\infty} \frac{(-1)^k x^{2k-1}}{2k-1}$

$\boxed{12.}$ *The function*

$$J_1(x) = \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k+1}}{k!\,(k+1)!\,2^{2k+1}}$$

is called Bessel function of order 1*. Determine its domain.*

$\boxed{13.}$ *Given the function*

$$f(x) = 1 + 2x + x^2 + 2x^3 + \cdots = \sum_{k=0}^{\infty} a_k x^k\,,$$

where $a_{2k} = 1$, $a_{2k+1} = 2$ *for any* $k \geq 0$, *determine the domain of* f *and an explicit formula for it.*

14. *Determine the convergence set of the series:*

$\boxed{\text{a)}}$ $\displaystyle\sum_{k=1}^{\infty} \left(\frac{2}{3}\right)^k (x^2-1)^k$ $\boxed{\text{b)}}$ $\displaystyle\sum_{k=1}^{\infty} \frac{1}{\sqrt[k]{k}} \left(\frac{1+x}{1-x}\right)^k$

$\boxed{\text{c)}}$ $\displaystyle\sum_{k=1}^{\infty} \frac{k+1}{k^2+1}\, 2^{-kx^2}$ $\boxed{\text{d)}}$ $\displaystyle\sum_{k=1}^{\infty} \frac{1}{k} \frac{(3x)^k}{\left(x+\frac{1}{x}\right)^k}$

$\boxed{15.}$ *Determine the radius of convergence of the power series*

$$\sum_{k=0}^{\infty} a^{\sqrt{k}} x^k$$

as the real parameter $a > 0$ *varies.*

16. *Expand in Maclaurin series the following functions, computing the radius of convergence of the series thus obtained:*

$\boxed{\text{a)}}$ $f(x) = \dfrac{x^3}{x+2}$ $\boxed{\text{b)}}$ $f(x) = \dfrac{1+x^2}{1-x^2}$

$\boxed{\text{c)}}$ $f(x) = \log(3-x)$ $\boxed{\text{d)}}$ $f(x) = \dfrac{x^3}{(x-4)^2}$

$\boxed{\text{e)}}$ $f(x) = \log\dfrac{1+x}{1-x}$ f) $f(x) = \sin x^4$

$\boxed{\text{g)}}$ $f(x) = \sin^2 x$ h) $f(x) = 2^x$

17. *Expand the maps below in Taylor series around the point* x_0*, and tell what is the radius of the series:*

$\boxed{\text{a)}}$ $f(x) = \dfrac{1}{x}\,,$ $x_0 = 1$

$\boxed{\text{b)}}$ $f(x) = \sqrt{x}\,,$ $x_0 = 4$

c) $f(x) = \log x\,,$ $x_0 = 2$

$\boxed{18.}$ *Verify that*

$$\sum_{k=1} k^2 x^k = \frac{x^2+x}{(1-x)^3}$$

for $|x| < 1$.

19. *Write the first three terms of the Maclaurin series of:*

$\boxed{\text{a)}}$ $f(x) = \dfrac{\log(1-x)}{e^x}$ b) $f(x) = e^{-x^2}\cos x$ c) $f(x) = \dfrac{\sin x}{1-x}$

20. *Write as Maclaurin series the following indefinite integrals:*

$\boxed{\text{a)}}$ $\displaystyle\int \sin x^2 \, dx$ b) $\displaystyle\int \sqrt{1+x^3}\, dx$

21. *Using series' expansions compute the definite integrals with the accuracy required:*

a) $\displaystyle\int_0^1 \sin x^2 \, dx\,,$ up to the third digit

b) $\displaystyle\int_0^{1/10} \sqrt{1+x^3}\, dx\,,$ with an absolute error $< 10^{-8}$

2.7.1 Solutions

1. *Limits of sequences of functions:*

a) Since $f_n(0) = 0$ for every n, $f(0) = 0$; if $x \neq 0$,

$$f_n(x) \sim \frac{nx}{n^3x^3} = \frac{1}{n^2x^2} \to 0 \qquad \text{for} \qquad n \to +\infty\,.$$

The limit function f is identically zero on I.
For the uniform convergence, we study the maps f_n on I and notice

$$f_n'(x) = \frac{n(1-2n^3x^3)}{(1+n^3x^3)^2}$$

and $f_n'(x) = 0$ for $x_n = \frac{1}{\sqrt[3]{2}n}$ with $f_n(x_n) = \frac{2}{3\sqrt[3]{2}}$ (Fig. 2.3). Hence

$$\sup_{x\in[0,+\infty)} |f_n(x)| = \frac{2}{3\sqrt[3]{2}}$$

and the convergence is not uniform on $[0,+\infty)$. Nonetheless, with $\delta > 0$ fixed and n large enough, f_n is decreasing on $[\delta,+\infty)$, so

$$\sup_{x\in[\delta,+\infty)} |f_n(x)| = f_n(\delta) \to 0 \qquad \text{as} \quad n \to +\infty\,.$$

The convergence is uniform on all intervals $[\delta,+\infty)$ with $\delta > 0$.

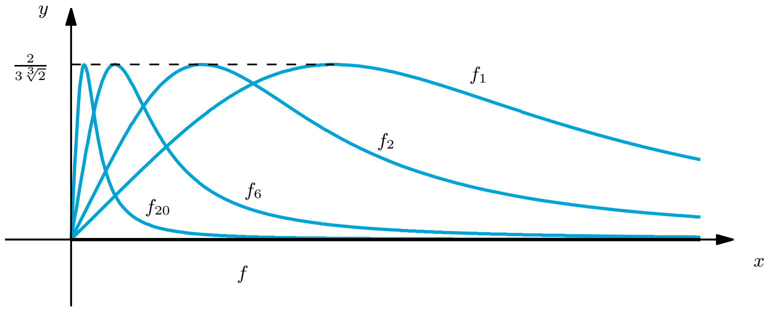


Figure 2.3. Graphs of f_n and f relative to Exercise 1. a)

b) Reasoning as before, the sequence converges pointwise on I to the limit $f(x) = 0$, for any $x \in I$. Moreover

$$f_n'(x) = \frac{1 - n^2x^2}{(1 + n^2x^2)^2}\,,$$

and for any $x \geq 0$

$$f_n'(x) = 0 \iff x = \frac{1}{n} \qquad \text{with} \quad f_n\Big(\frac{1}{n}\Big) = \frac{1}{2n}\,.$$

Thus there is uniform convergence on I, since

$$\lim_{n\to\infty} \sup_{x\in[0,+\infty)} |f_n(x)| = \lim_{n\to\infty} \frac{1}{2n} = 0\,.$$

See Fig. 2.4.

c) The sequence converges pointwise on $(-\infty, 0]$ to

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

We have no uniform convergence on $(-\infty, 0]$ as f is not continuous; but on all half-lines $(-\infty, -\delta]$, for any $\delta > 0$, this is the case, because

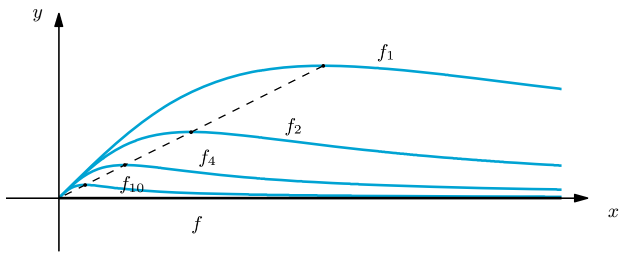


Figure 2.4. Graphs of f_n and f relative to Exercise 1. b)

$$\lim_{n\to\infty} \sup_{x\in(-\infty,-\delta]} \mathrm{e}^{nx} = \lim_{n\to\infty} \mathrm{e}^{-n\delta} = 0\,.$$

d) We have pointwise convergence to $f(x) = 0$ for all $x \in [0, +\infty)$, and uniform convergence on every $[\delta, +\infty)$, $\delta > 0$.

e) Note $f_n(0) = 1/2$. As $n \to \infty$ the maps f_n satisfy

$$f_n(x) \sim \begin{cases} (4/3)^{nx} & \text{if } x < 0\ , \\ (4/5)^{nx} & \text{if } x > 0\,. \end{cases}$$

Anyway for $x \neq 0$

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

Hence the sequence converges pointwise on $\mathbb{R}$ to the limit

$$f(x) = \begin{cases} 0 & \text{if } x \neq 0, \\ 1/2 & \text{if } x = 0. \end{cases}$$

The convergence is not uniform on $\mathbb{R}$ because the limit is not continuous on that domain. But we do have uniform convergence on every set $A_\delta = (-\infty, -\delta] \cup [\delta, +\infty)$, $\delta > 0$, since

$$\lim_{n\to\infty} \sup_{x\in A_\delta} |f_n(x)| = \lim_{n\to\infty} \max\big(f_n(\delta), f_n(-\delta)\big) = 0\,.$$

2. Notice $f_n(1) = f_n(-1) = 0$ for all n. For any $x \in (-1, 1)$ moreover, $1 - x^2 < 1$; hence

$$\lim_{n\to\infty} f_n(x) = 0\,, \qquad \forall x \in (-1, 1)\,.$$

Then the sequence converges to $f(x) = 0\,,\ \forall x \in [-1, 1]$.

What about uniform convergence? For this we consider the odd maps f_n, so it suffices to take $x \in [0, 1]$. Then

$$f_n'(x) = n(1 - x^2)^n - 2nx^2(1 - x^2)^{n-1} = n(1 - x^2)^{n-1}(1 - x^2 - 2nx^2)\,,$$

and $f_n(x)$ has a maximum point $x = 1/\sqrt{1 + 2n}$ (and by symmetry a minimum point $x = -1/\sqrt{1 + 2n}$). Therefore

$$\sup_{x\in[-1,1]} |f_n(x)| = f_n\left(\frac{1}{\sqrt{1 + 2n}}\right) = \frac{n}{\sqrt{1 + 2n}} \left(\frac{2n}{1 + 2n}\right)^n,$$

and the convergence is not uniform on $[-1, 1]$, for

$$\lim_{n\to\infty} \frac{n}{\sqrt{1 + 2n}} \left(\frac{2n}{1 + 2n}\right)^n = \mathrm{e}^{-1/2} \lim_{n\to\infty} \frac{n}{\sqrt{1 + 2n}} = +\infty\,.$$

From this argument the convergence cannot be uniform on the interval $[0, 1]$ either, and we cannot swap the limit with differentiation. Let us check the formula. Putting $t = 1 - x^2$, we have

$$\lim_{n\to\infty} \int_0^1 f_n(x)\,\mathrm{d}x = \lim_{n\to\infty} \frac{n}{2} \int_0^1 t^n\,\mathrm{d}t = \lim_{n\to\infty} \frac{n}{2(n+1)} = \frac{1}{2}\,,$$

while

$$\int_0^1 \lim_{n\to\infty} f_n(x)\,\mathrm{d}x = 0\,.$$

3. Since

$$\lim_{n\to\infty} f_n(x) = \begin{cases} \pi/2 & \text{if } x > 0\ , \\ 0 & \text{if } x = 0\ , \\ -\pi/2 & \text{if } x < 0\,, \end{cases}$$

we have pointwise convergence on $\mathbb{R}$, but not uniform: the limit is not continuous despite the f_n are. Similarly, no uniform convergence on $[0, 1]$. Therefore, if we put $a = 0$, it is not possible to exchange limit and integration automatically. Compute the two sides of the equality independently:

$$\begin{aligned}
\lim_{n\to\infty} \int_0^1 f_n(x)\,\mathrm{d}x &= \lim_{n\to\infty} \left(x \arctan nx \Big|_0^1 - \int_0^1 \frac{nx}{1+n^2x^2}\,\mathrm{d}x \right) \\
&= \lim_{n\to\infty} \left[\arctan n - \frac{\log(1+n^2x^2)}{2n} \Big|_0^1 \right] \\
&= \lim_{n\to\infty} \left(\arctan n - \frac{\log(1+n^2)}{2n} \right) = \frac{\pi}{2}\,.
\end{aligned}$$

Moreover

$$\int_0^1 \lim_{n\to\infty} f_n(x)\,\mathrm{d}x = \int_0^1 \frac{\pi}{2}\,\mathrm{d}x = \frac{\pi}{2}\,.$$

Hence the equality holds even if the sequence does not converge uniformly on $[0, 1]$.

If we take $a = 1/2$, instead, we have uniform convergence on $[1/2, 1]$, so the equality is true by Theorem 2.18.

4. *Convergence set for series of functions:*

a) Fix x, so that

$$f_k(x) = \frac{(k+2)^x}{k^3 + \sqrt{k}} \sim \frac{1}{k^{3-x}}\,, \qquad k \to \infty\,.$$

The series is like the generalised harmonic series of exponent $3 - x$, so it converges if $3 - x > 1$, hence $x < 2$, and diverges if $3 - x \le 1$, so $x \ge 2$.

b) Given x, use the Root Test to see

$$\lim_{k\to\infty} \sqrt[k]{|f_k(x)|} = \lim_{k\to\infty} \left(1+\frac{x}{k}\right)^k = \mathrm{e}^x\,.$$

Then the series converges if $\mathrm{e}^x < 1$, i.e., $x < 0$, and diverges if $\mathrm{e}^x > 1$, so $x > 0$. If $x = 0$, the series diverges because the general term is always 1. The convergence set is the half-line $(-\infty, 0)$.

c) If $x = 1$, the general term does not tend to 0 so the series cannot converge. If $x \neq 1$, as $k \to \infty$ we have

$$f_k(x) \sim \begin{cases} x^{-k} & \text{if } x > 1, \\ x^k & \text{if } x < 1\,. \end{cases}$$

In either case the series converges. Thus the convergence set is $(0, +\infty) \setminus \{1\}$.

d) If $|x| < 2$, the convergence is absolute as $|f_k(x)| \sim \left(\frac{|x|}{2}\right)^k$, $k \to \infty$. If $x \le -2$ or $x \ge 2$ the series does not converge since the general term is not infinitesimal. The set of convergence is $(-2, 2)$.

e) Observe that

$$|f_k(x)| \sim \frac{1}{k^{1-x}}\,, \qquad k \to \infty\,.$$

The series converges absolutely if $1 - x > 1$, so $x < 0$. By Leibnitz's Test, the series converges pointwise if $0 < 1-x \le 1$, i.e., $0 \le x < 1$. It does not converge (it is indeterminate, actually) if $x \ge 1$, since the general term does not tend to 0. The convergence set is thus $(-\infty, 1)$.

f) Given x, the general term is equivalent to that of a generalised harmonic series:

$$f_k(x) = \left(\frac{1}{k+\sqrt{k^2-1}}\right)^x \sim \left(\frac{1}{2k}\right)^x\,, \qquad k \to \infty\,.$$

The convergence set is $(1, +\infty)$.

5. Geometric series with $q = \mathrm{e}^x$, converging pointwise on $(-\infty, 0)$ and uniformly on $(-\infty, -\delta]$, for any $\delta > 0$. Moreover for any $x < 0$

$$\sum_{k=2}^{\infty} (\mathrm{e}^x)^k = \frac{1}{1-\mathrm{e}^x} - 1 - \mathrm{e}^x = \frac{\mathrm{e}^{2x}}{1-e^x}\,.$$

6. For any $x \in \mathbb{R}$, $\displaystyle\lim_{k\to\infty} \cos\frac{x}{k} = 1 \neq 0$. Hence the convergence set of $\displaystyle\sum_{k=1}^{\infty} \cos\frac{x}{k}$ is empty.

As $f_k'(x) = -\frac{1}{k}\sin\frac{x}{k}$, we have

$$|f_k'(x)| \le \frac{|x|}{k^2} \le \frac{1}{k^2}\,, \qquad \forall x \in [-1, 1]\,.$$

Since $\sum_{k=1}^{\infty}\frac{1}{k^2}$ converges, the M-test of Weierstrass tells $\sum_{k=1}^{\infty} f_k'(x)$ converges uniformly on $[-1,1]$.

7. *Sets of pointwise and uniform convergence:*

a) This is harmonic with exponent $-1/x$, so: pointwise convergence on $(-1,0)$, uniform convergence on any sub-interval $[a,b]$ of $(-1,0)$.

b) The Integral Test tells the convergence is pointwise on $(-\infty,-1)$. Uniform convergence happens on every half-line $(-\infty,-\delta]$, for any $\delta>1$.

c) Observing

$$f_k(x)\sim\frac{1}{k^2+x^2}\log(k^2+x^2)\,,\qquad k\to\infty\,,$$

the series converges pointwise on $\mathbb{R}$. Moreover,

$$\sup_{x\in\mathbb{R}}|f_k(x)|=f_k(0)=\exp\left(\frac{1}{k^2}\log k^2\right)-1=M_k\,.$$

The numerical series $\sum_{k=1}^{\infty}M_k$ converges just like $\sum_{k=1}^{\infty}\frac{2\log k}{k^2}$. The M-test implies the convergence is also uniform on $\mathbb{R}$.

8. The assumption ensures the radius of convergence is bigger or equal than 4. Hence the first series converges, while for the second one we cannot say anything.

9. *Convergence of power series:*

a) Converges. b) Diverges. c) Converges. d) Diverges.

10. We have

$$\left|\frac{a_{k+1}}{a_k}\right|=\frac{\big((k+1)!\big)^p\,(pk)!}{\big(p(k+1)\big)!\,(k!)^p}\,\frac{(k+1)^p}{(pk+1)(pk+2)\cdots(pk+p)}\,.$$

Thus

$$\lim_{k\to\infty}\left|\frac{a_{k+1}}{a_k}\right|=\lim_{k\to\infty}\frac{k+1}{pk+1}\,\frac{k+1}{pk+2}\cdots\frac{k+1}{pk+p}=\frac{1}{p^p}\,,$$

and the Ratio Test gives $R=p^p$.

11. *Radius and set of convergence for power series:*

a) $R=1$, $I=[-1,1)$
b) $R=1$, $I=(-1,1]$
c) $R=1$, $I=(-1,1)$
d) $R=3$, $I=(-3,3]$
e) $R=1$, $I=(3,5)$
f) $R=10$, $I=(-9,11)$
g) $R=3$, $I=(-6,0]$
h) $R=0$, $I=\{\frac{1}{2}\}$
i) $R=+\infty$, $I=\mathbb{R}$
ℓ) $R=1$, $I=[-1,1]$

12. Using the Ratio Test:

$$\lim_{k\to\infty}\left|\frac{a_{k+1}}{a_k}\right| = \lim_{k\to\infty}\frac{k!\,(k+1)!\,2^{2k+1}}{(k+1)!\,(k+2)!\,2^{2k+3}} = \lim_{k\to\infty}\frac{1}{4(k+2)(k+1)} = 0\,.$$

Hence $R = +\infty$ and the domain of the function is $\mathbb{R}$.

13. Since $\lim\limits_{k\to\infty}\sqrt[k]{|a_k|} = 1$, the radius is $R = 1$. It is straightforward to see the series does not converge for $x = \pm 1$ because the general term does not tend to 0. Hence $\operatorname{dom} f = (-1,1)$, and for $x \in (-1,1)$,

$$f(x) = \sum_{k=0}^{\infty} x^{2k} + 2\sum_{k=0}^{\infty} x^{2k+1}\frac{1}{1-x^2} + \frac{2x}{1-x^2} = \frac{1+2x}{1-x^2}\,.$$

14. *Set of convergence:*

a) Put $y = x^2 - 1$ and look at the power series

$$\sum_{k=1}^{\infty}\left(\frac{2}{3}\right)^k y^k$$

in the variable y, with radius R_y. Since

$$\lim_{k\to\infty}\sqrt[k]{\left(\frac{2}{3}\right)^k} = \frac{2}{3}\,,$$

$R_y = \frac{3}{2}$. For $y = \pm\frac{3}{2}$ the series does not converge as the general term does not tend to 0. In conclusion, the series converges if $-\frac{3}{2} < x^2 - 1 < \frac{3}{2}$. The first inequality holds for any x, whereas the second one equals $x^2 < \frac{5}{2}$; The series converges on $\left(-\sqrt{\frac{5}{2}}, \sqrt{\frac{5}{2}}\right)$.

b) Let $x \neq 1$; set $y = \frac{1+x}{1-x}$ and consider the series $\sum\limits_{k=1}^{\infty}\frac{1}{\sqrt[k]{k}}y^k$ in y. Since

$$\lim_{k\to\infty}\sqrt[k]{\frac{1}{\sqrt[k]{k}}} = \lim_{k\to\infty} \mathrm{e}^{-\frac{1}{k^2}\log k} = 1\,,$$

we obtain $R_y = 1$. For $y = \pm 1$ there is no convergence as the general term does not tend to 0. Hence, the series converges for

$$-1 < \frac{1+x}{1-x} < 1\,, \qquad \text{i.e.,} \qquad x < 0\,.$$

c) Write $y = 2^{-x^2}$ and consider the power series $\sum\limits_{k=1}^{\infty}\frac{k+1}{k^2+1}y^k$. Immediately we have $R_y = 1$; in addition the series converges if $y = -1$ (like the alternating harmonic series) and diverges if $y = 1$ (harmonic series). Returning to the variable x, the series converges if $-1 \le 2^{-x^2} < 1$. The left inequality is trivial, the right one holds when $x \neq 0$. Overall the set of convergence is $\mathbb{R}\setminus\{0\}$.

d) When $x \neq 0$ we set $y = \dfrac{x}{x+\frac{1}{x}} = \dfrac{x^2}{1+x^2}$ and then study $\displaystyle\sum_{k=1}^{\infty} \frac{3^k}{k} y^k$. Its radius equals $R_y = \frac{1}{3}$, so it converges on $[-\frac{1}{3}, \frac{1}{3})$.
Back to the x, we impose the conditions

$$-\frac{1}{3} \leq \frac{x^2}{1+x^2} < \frac{1}{3}\,.$$

This is equivalent to $2x^2 < 1$, making $\big(-\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2}\big) \setminus \{0\}$ the convergence set.

15. Exploiting the Ratio Test we have

$$\lim_{k\to\infty} \left|\frac{a_{k+1}}{a_k}\right| = \lim_{k\to\infty} \frac{a^{\sqrt{k+1}}}{a^{\sqrt{k}}} = \lim_{k\to\infty} a^{\sqrt{k+1}-\sqrt{k}} = \lim_{k\to\infty} a^{\frac{1}{\sqrt{k+1}+\sqrt{k}}} = 1\,.$$

Hence $R = 1$ for any $a > 0$.

16. *Maclaurin series:*

a) Using the geometric series with $q = -\frac{x}{2}$,

$$f(x) = \frac{x^3}{2}\,\frac{1}{1+\frac{x}{2}} = \frac{x^3}{2}\sum_{k=0}^{\infty}\Big(-\frac{x}{2}\Big)^k = \sum_{k=0}^{\infty}(-1)^k \frac{x^{k+3}}{2^{k+1}};$$

this has radius $R = 2$.

b) With the geometric series where $q = x^2$, we have

$$f(x) = -1 + \frac{2}{1-x^2} = -1 + 2\sum_{k=0}^{\infty} x^{2k}\,,$$

whose radius is $R = 1$.

c) Expanding the function $g(t) = \log(1+t)$, where $t = -\frac{x}{3}$, we obtain

$$f(x) = \log 3 + \log(1-\frac{x}{3}) = \log 3 + \sum_{k=1}^{\infty}\frac{(-1)^{k+1}}{k}\Big(-\frac{x}{3}\Big)^k = \log 3 - \sum_{k=1}^{\infty}\frac{x^k}{k3^k}\,;$$

This has radius $R = 3$.

d) Expanding $g(t) = \dfrac{1}{(1-t)^2}$ (recall (2.16)) with $t = \frac{x}{4}$,

$$f(x) = \frac{x^3}{16}\,\frac{1}{(1-\frac{x}{4})^2} = \frac{x^3}{16}\sum_{k=0}^{\infty}(k+1)\Big(\frac{x}{4}\Big)^k = \sum_{k=0}^{\infty}\frac{(k+1)x^{k+3}}{4^{k+2}}\,;$$

Now the radius is $R = 4$.

e) Recalling the series of $g(t) = \log(1+t)$, with $t = x$ and then $t = -x$, we find

$$f(x) = \log(1+x) - \log(1-x) = \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} x^k - \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} (-x)^k$$
$$= \sum_{k=1}^{\infty} \frac{1}{k}\big((-1)^{k+1} + 1\big) x^k = \sum_{k=0}^{\infty} \frac{2}{2k+1} x^{2k+1} ;$$

thus the radius equals $R = 1$.

f) $f(x) = \sum_{k=0}^{\infty} (-1)^k \frac{x^{8k+4}}{(2k+1)!}$ has radius $R = \infty$.

g) Since $\sin^2 x = \frac{1}{2}(1 - \cos 2x)$, and remembering the expansion of $g(t) = \cos t$ with $t = 2x$, we have

$$f(x) = \frac{1}{2} - \frac{1}{2} \sum_{k=0}^{\infty} (-1)^k \frac{2^{2k} x^{2k}}{(2k)!}, \qquad R = +\infty .$$

h) $f(x) = \sum_{k=0}^{\infty} \frac{(\log 2)^k}{k!} x^k$, with $R = +\infty$.

17. *Taylor series:*

a) One can proceed directly and compute the derivatives of f to obtain

$$f^{(k)}(x) = (-1)^k \frac{k!}{x^{k+1}}, \qquad \text{whence} \qquad f^{(k)}(1) = (-1)^k k!, \quad \forall k \in \mathbb{N} .$$

Therefore

$$f(x) = \sum_{k=0}^{\infty} (-1)^k (x-1)^k, \qquad R = 1 .$$

Alternatively, one could set $t = x - 1$ and take the Maclaurin series of $f(t) = \frac{1}{1+t}$ to arrive at the same result:

$$\frac{1}{1+t} = \sum_{k=0}^{\infty} (-1)^k t^k = \sum_{k=0}^{\infty} (-1)^k (x-1)^k, \qquad R = 1 .$$

b) Here as well we compute directly

$$f'(x) = \frac{1}{2} x^{-1/2}, \qquad f^{(k)}(x) = (-1)^{k+1} \frac{1 \cdot 3 \cdot 5 \cdots (2k-3)}{2^k} x^{-\frac{2k-1}{2}},$$

for all $k \geq 2$; then

$$f(x) = 2 + \frac{1}{4}(x-4) + \sum_{k=1}^{\infty}(-1)^{k+1}\,\frac{1\cdot 3\cdot 5\cdots(2k-3)}{k!2^k}\,2^{-(2k-1)}\,(x-4)^k$$

and the radius is $R = 4$.

Alternatively, put $t = x - 4$, to the effect that

$$\begin{aligned}
\sqrt{x} = \sqrt{4+t} &= 2\sqrt{1+\frac{t}{4}} \\
&= 2 + \frac{1}{4}(x-4) + 2\sum_{k=2}^{\infty}\binom{\frac{1}{2}}{k}\left(\frac{t}{4}\right)^k \\
&= 2 + \frac{1}{4}(x-4) + 2\sum_{k=2}^{\infty}\frac{\frac{1}{2}\left(\frac{1}{2}-1\right)\cdots\left(\frac{1}{2}-k+1\right)}{k!}\,\frac{1}{4^k}\,(x-4)^k \\
&= 2 + \frac{1}{4}(x-4) + 2\sum_{k=2}^{\infty}\frac{1\cdot 3\cdots(2k-3)}{2^k k!}\,\frac{(-1)^{k+1}}{2^{2k}}\,(x-4)^k \\
&= 2 + \frac{1}{4}(x-4) + \sum_{k=2}^{\infty}(-1)^{k+1}\,\frac{1\cdot 3\cdot 5\cdots(2k-3)}{k!2^{3k-1}}\,(x-4)^k\,.
\end{aligned}$$

c) $\log x = \log 2 + \sum_{k=1}^{\infty}\frac{(-1)^{k+1}}{k2^k}(x-2)^k\,, \quad R = 2.$

18. The equality is trivial for $x = 0$. Differentiating term by term, for $|x| < 1$ we have

$$\sum_{k=0}^{\infty} x^k = \frac{1}{1-x}\,, \qquad \sum_{k=1}^{\infty} kx^{k-1} = \frac{1}{(1-x)^2}\,, \qquad \sum_{k=2}^{\infty} k(k-1)x^{k-2} = \frac{2}{(1-x)^3}\,.$$

From the last relationship, when $x \neq 0$, we have

$$\frac{2}{(1-x)^3} = \sum_{k=1}^{\infty} k(k+1)x^{k-1} = \sum_{k=1}^{\infty} k^2x^{k-1} + \sum_{k=1}^{\infty} kx^{k-1} = \frac{1}{x}\sum_{k=1}^{\infty} k^2x^k + \frac{1}{(1-x)^2}\,.$$

Therefore

$$\sum_{k=1}^{\infty} k^2x^k = \frac{2x}{(1-x)^3} - \frac{x}{(1-x)^2} = \frac{x^2+x}{(1-x)^3}\,.$$

19. *Maclaurin series:*

a) Using the well-known series of $g(x) = \log(1-x)$ and $h(x) = \mathrm{e}^{-x}$ yields

$$\begin{aligned}
f(x) = \mathrm{e}^{-x}\log(1-x) &= \left(1 - x + \frac{x^2}{2} - \frac{x^3}{3!} + \cdots\right)\left(-x - \frac{x^2}{2} - \frac{x^3}{3} - \cdots\right) \\
&= -x + x^2 - \frac{1}{2}x^2 - \frac{1}{3}x^3 + \frac{1}{2}x^3 - \frac{1}{2}x^3 + \cdots \\
&= -x + \frac{1}{2}x^2 - \frac{1}{3}x^3 + \cdots
\end{aligned}$$

b) $f(x) = 1 - \frac{3}{2}x^2 + \frac{25}{24}x^4 - \cdots$

c) $f(x) = x + x^2 + \frac{5}{6}x^3 + \cdots$

20. *Indefinite integrals:*

a) Since

$$\sin x^2 = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!}\, x^{2(2k+1)}, \qquad \forall x \in \mathbb{R},$$

a term-by-term integration produces

$$\int \sin x^2 \,\mathrm{d}x = c + \sum_{k=0}^{\infty} \frac{(-1)^k x^{4k+3}}{(4k+3)(2k+1)!}$$

with arbitrary constant c.

b) $\displaystyle\int \sqrt{1+x^3}\,\mathrm{d}x = c + x + \frac{x^4}{8} + \sum_{k=2}^{\infty}(-1)^{k-1}\frac{1\cdot 3\cdot 5\cdots(2k-3)}{2^k k!(3k+1)}\, x^{3k+1}.$

21. *Definite integrals:*

a) $\displaystyle\int_0^1 \sin x^2 \,\mathrm{d}x \sim 0.310.$

b) $\displaystyle\int_0^{1/10} \sqrt{1+x^3}\,\mathrm{d}x \sim 0.10001250.$

3

Fourier series

The sound of a guitar, the picture of a footballer on tv, the trail left by an oil tanker on the ocean, the sudden tremors of an earthquake are all examples of events, either natural or caused by man, that have to do with travelling-wave phenomena. Sound for instance arises from the swift change in air pressure described by pressure waves that move through space. The other examples can be understood similarly using propagating electromagnetic waves, water waves on the ocean's surface, and elastic waves within the ground, respectively.

The language of Mathematics represents waves by one or more functions that model the physical object of concern, like air pressure, or the brightness of an image's basic colours; in general this function depends on time and space, for instance the position of the microphone or the coordinates of the pixel on the screen. If one fixes the observer's position in space, the wave appears as a function of the time variable only, hence as a signal (acoustic, of light, ...). On the contrary, if we fix a moment in time, the wave will look like a collection of values in space of the physical quantity represented (think of a picture still on the screen, a photograph of the trail taken from a plane, and so on).

Propagating waves can have an extremely complicated structure; the desire to understand and be able to control their behaviour in full has stimulated the quest for the appropriate mathematical theories to analyse them, in the last centuries. Generally speaking, this analysis aims at breaking down a complicated wave's structure in the superposition of simpler components that are easy to treat and whose nature is well understood. According to such theories there will be a collection of 'elementary waves', each describing a specific and basic way of propagation. Certain waves are obtained by superposing a finite number of elementary ones, yet the more complex structures are given by an infinite number of elementary waves, and then the tricky problem arises – as always – of making sense of an infinite collection of objects and how to handle them.

The so-called Fourier Analysis is the most acclaimed and widespread framework for describing propagating phenomena (and not only). In its simplest form, Fourier Analysis considers one-dimensional signals with a given periodicity. The elementary waves are sine functions characterised by a certain frequency, phase

C. Canuto, A. Tabacco: *Mathematical Analysis II,* 2nd Ed.,
UNITEXT – La Matematica per il 3+2 85, DOI 10.1007/978-3-319-12757-6_3,

and amplitude of oscillation. The composition of a finite number of elementary waves generates trigonometric polynomials, whereas an infinite number produces a series of functions, called Fourier series. This is an extremely efficient way of representing in series a large class of periodic functions. This chapter introduces the rudiments of Fourier Analysis through the study of Fourier series and the issues of their convergence to a periodic map.

Beside standard Fourier Analysis, other much more sophisticated tools for analysing and representing functions have been developed, especially in the last decades of the XX century, among which the so-called Wavelet Analysis. Some of those theories lie at the core of the recent, striking success of Mathematics in several groundbreaking technological applications such as mobile phones, or digital sound and image processing. Far from obfuscating the classical subject matter, these latter-day developments highlight the importance of the standard theory as foundational for all successive advancements.

3.1 Trigonometric polynomials

We begin by recalling periodic functions.

Definition 3.1 *A map* $f : \mathbb{R} \to \mathbb{R}$ *is* **periodic of period** $T > 0$ *if*

$$f(x+T) = f(x)\,, \qquad \forall x \in \mathbb{R}.$$

If f is periodic of period T it is also periodic of period kT, with $k \in \mathbb{N} \setminus \{0\}$, including that f might be periodic of period T/k, $k \in \mathbb{N} \setminus \{0\}$. The **minimum period** of f is the smallest T (if existent) for which f is periodic. Moreover, f is known when we know it on an interval $[x_0, x_0+T)$ (or $(x_0, x_0+T]$) of length T. Usually one chooses the interval $[0,T)$ or $\big[-\frac{T}{2}, \frac{T}{2}\big)$. Note that if f is constant, it is periodic of period T, for any $T > 0$.

Any map defined on a bounded interval $[a, b)$ can be prolonged to a periodic function of period $T = b - a$, by setting

$$f(x+kT) = f(x)\,, \qquad k \in \mathbb{Z}\,,\ \forall x \in [a,b)\,.$$

Such prolongation is not necessarily continuous at the points $x = a + kT$, even if the original function is.

Examples 3.2

i) The functions $f(x) = \cos x$, $g(x) = \sin x$ are periodic of period $2k\pi$ with $k \in \mathbb{N} \setminus \{0\}$. Both have minimum period 2π.

ii) $f(x) = \cos\frac{x}{4}$ is periodic of period 8π, $16\pi, \ldots$; its minimum period is 8π.

iii) The maps $f(x) = \cos\omega x$, $g(x) = \sin\omega x$, $\omega \neq 0$, are periodic with minimum period $T_0 = \frac{2\pi}{\omega}$.

iv) The mantissa map $f(x) = M(x)$ is periodic of minimum period $T_0 = 1$. □

The following properties are easy to prove.

Proposition 3.3 *Let f be periodic of period $T > 0$; for any $x_0 \in \mathbb{R}$ then,*

$$\int_0^T f(x)\,\mathrm{d}x = \int_{x_0}^{x_0+T} f(x)\,\mathrm{d}x\,.$$

In particular, if $x_0 = -T/2$,

$$\int_0^T f(x)\,\mathrm{d}x = \int_{-T/2}^{T/2} f(x)\,\mathrm{d}x\,.$$

Proof. By the properties of definite integrals,

$$\int_0^T f(x)\,\mathrm{d}x = \int_0^{x_0} f(x)\,\mathrm{d}x + \int_{x_0}^{x_0+T} f(x)\,\mathrm{d}x + \int_{x_0+T}^{T} f(x)\,\mathrm{d}x\,.$$

Putting $x = y + T$ in the last integral, by periodicity

$$\int_{x_0+T}^{T} f(x)\,\mathrm{d}x = \int_{x_0}^{0} f(y+T)\,\mathrm{d}y = \int_{x_0}^{0} f(y)\,\mathrm{d}y - \int_0^{x_0} f(y)\,\mathrm{d}y\,,$$

whence the result. □

Proposition 3.4 *Let f be a periodic map of period $T_1 > 0$ and take $T_2 > 0$. Then $g(x) = f\Big(\frac{T_1}{T_2}x\Big)$ is periodic of period T_2.*

Proof. For any $x \in \mathbb{R}$,

$$g(x+T_2) = f\Big(\frac{T_1}{T_2}(x+T_2)\Big) = f\Big(\frac{T_1}{T_2}x + T_1\Big) = f\Big(\frac{T_1}{T_2}x\Big) = g(x)\,. \qquad \square$$

The periodic function $f(x) = a\sin(\omega x + \phi)$, where a, ω,ϕ are constant, is rather important for the sequel. It describes in Physics a special oscillation of sine type, and goes under the name of **simple harmonic**. Its minimum period equals $T = \frac{2\pi}{\omega}$ and the latter's reciprocal $\frac{\omega}{2\pi}$ is the **frequency**, i.e., the number of wave oscillations on each unit interval (oscillations per unit of time, if x denotes time); ω is said **angular frequency**. The quantities a and ϕ are called **amplitude** and **phase** (offset) of the oscillation. Modifying $a > 0$ has the effect of widening or shrinking the range of f (the oscillation's crests and troughs move apart or get closer, respectively), while a positive or negative variation of ϕ translates the wave left or right (Fig. 3.1). A simple harmonic can also be represented as $a\sin(\omega x+\phi) =$

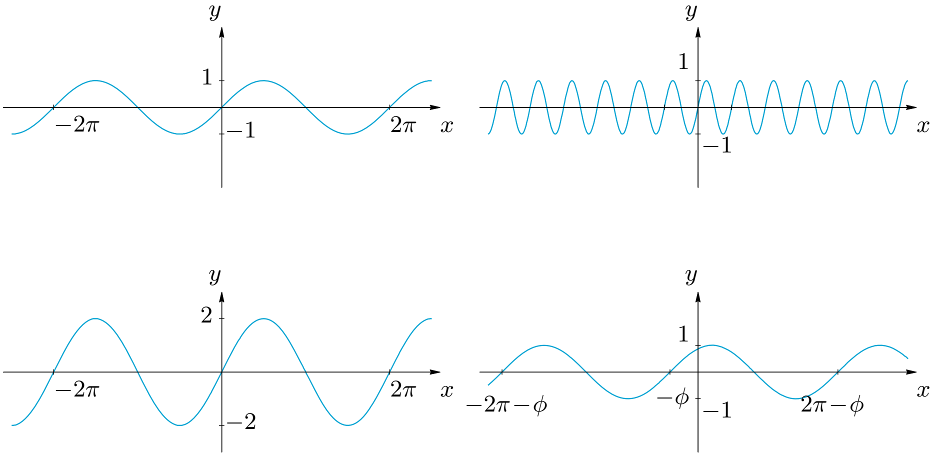


Figure 3.1. Simple harmonics for various values of angular frequency, amplitude and phase: $(\omega, a, \phi) = (1, 1, 0)$, top left; $(\omega, a, \phi) = (5, 1, 0)$, top right; $(\omega, a, \phi) = (1, 2, 0)$, bottom left; $(\omega, a, \phi) = (1, 1, \frac{\pi}{3})$, bottom right

$\alpha \cos \omega x + \beta \sin \omega x$ with $\alpha = a \sin \phi$ and $\beta = a \cos \phi$; the inverse transformation of the parameters is $a = \sqrt{\alpha^2 + \beta^2}$, $\phi = \arctan \frac{\alpha}{\beta}$.

In the sequel we shall concentrate on periodic functions of period 2π because all results can be generalised by a simple variable change, thanks to Proposition 3.4. (More details can be found in Sect. 3.6.)

The superposition of simple harmonics whose frequencies are all multiple of one fundamental frequency, say $1/2\pi$ for simplicity, gives rise to trigonometric polynomials.

Definition 3.5 *A* **trigonometric polynomial** *of order or degree n is a finite linear combination*

$$\begin{aligned} P(x) &= a_0 + a_1 \cos x + b_1 \sin x + \ldots + a_n \cos nx + b_n \sin nx \\ &= a_0 + \sum_{k=1}^{n} (a_k \cos kx + b_k \sin kx) \,, \end{aligned}$$

where a_k, b_k are real constants and at least one of a_n, b_n is non-zero. Representing simple harmonics in terms of their amplitude and phase, a trigonometric polynomial can be written as

$$P(x) = a_0 + \sum_{k=1}^{n} \alpha_k \sin(kx + \varphi_k) \,.$$

The name stems from the observation that each algebraic polynomial of degree n in X and Y generates a trigonometric polynomial of the same degree n, by substituting $X = \cos x$, $Y = \sin x$ and using suitable trigonometric identities. For example, $p(X, Y) = X^3 + 2Y^2$ gives

$$\begin{aligned}
p(\cos x, \sin x) &= \cos^3 x + 2\sin^2 x \\
&= \cos x\,\frac{1+\cos 2x}{2} + 2\,\frac{1-\cos 2x}{2} \\
&= 1 + \frac{1}{2}\cos x - \cos 2x + \frac{1}{2}\cos x \cos 2x \\
&= 1 + \frac{1}{2}\cos x - \cos 2x + \frac{1}{4}(\cos x + \cos 3x) \\
&= 1 + \frac{3}{4}\cos x - \cos 2x + \frac{1}{4}\cos 3x = P(x).
\end{aligned}$$

Obviously, not all periodic maps can be represented as trigonometric polynomials (e.g., $f(x) = \mathrm{e}^{\sin x}$). At the same time though, certain periodic maps (that include the functions appearing in applications) may be approximated, in a sense to be made precise, by trigonometric polynomials: they can actually be expanded in series of trigonometric polynomials. These functions are called **Fourier series** and are the object of concern in this chapter.

3.2 Fourier Coefficients and Fourier series

Although the theory of Fourier series can be developed in a very broad context, we shall restrict to a subclass of all periodic maps (of period 2π), namely those belonging to the space $\tilde{\mathcal{C}}_{2\pi}$, which we will introduce in a moment. First though, a few preliminary definitions are required.

Definition 3.6 *A map f periodic of period 2π is* **piecewise continuous** *if it is continuous on $[0, 2\pi]$ except for at most a finite number of points x_0. At such points there can be a removable singularity or a singularity of the first kind, so the left and right limits*

$$f(x_0^+) = \lim_{x\to x_0^+} f(x) \qquad \text{and} \qquad f(x_0^-) = \lim_{x\to x_0^-} f(x)$$

exist and are finite.
If, in addition,

$$f(x_0) = \frac{1}{2}\big(f(x_0^+) + f(x_0^-)\big), \tag{3.1}$$

at each discontinuity x_0, f is called **regularised**.

Definition 3.7 *We denote by $\tilde{\mathcal{C}}_{2\pi}$ the space of maps defined on $\mathbb{R}$ that are periodic of period 2π, piecewise continuous and regularised.*

The set $\tilde{\mathcal{C}}_{2\pi}$ is an $\mathbb{R}$-vector space (i.e., $\alpha f + \beta g \in \tilde{\mathcal{C}}_{2\pi}$ for any $\alpha,\beta \in \mathbb{R}$ and all $f, g \in \tilde{\mathcal{C}}_{2\pi}$); it is not hard to show that given $f,\, g \in \tilde{\mathcal{C}}_{2\pi}$, the expression

$$(f,\, g) = \int_0^{2\pi} f(x)g(x)\,\mathrm{d}x \tag{3.2}$$

defines a **scalar product** on $\tilde{\mathcal{C}}_{2\pi}$ (we refer to Appendix A.2.1, p. 521, for the general concept of scalar product and of norm of a function). In fact,

i) $(f,\, f) \geq 0$ for any $f \in \tilde{\mathcal{C}}_{2\pi}$, and $(f,\, f) = 0$ if and only if $f = 0$;

ii) $(f,\, g) = (g,\, f)$, for any $f,\, g \in \tilde{\mathcal{C}}_{2\pi}$;

iii) $(\alpha f + \beta g,\, h) = \alpha(f,\, h) + \beta(g,\, h)$, for all $f,\, g,\, h \in \tilde{\mathcal{C}}_{2\pi}$ and any $\alpha,\beta \in \mathbb{R}$.

The only non-trivial fact is that $(f,\, f) = 0$ forces $f = 0$. To see that, let $x_1, \ldots, x_n$ be the discontinuity points of f in $[0, 2\pi]$; then

$$0 = \int_0^{2\pi} f^2(x)\,\mathrm{d}x = \int_0^{x_1} f^2(x)\,\mathrm{d}x + \int_{x_1}^{x_2} f^2(x)\,\mathrm{d}x + \ldots + \int_{x_n}^{2\pi} f^2(x)\,\mathrm{d}x\,.$$

As f is continuous on every sub-interval $(x_i,\, x_{i+1})$, we get $f(x) = 0$ on each of them. At last, $f(x_i) = 0$ at each point of discontinuity, by (3.1).

Associated to the scalar product (3.2) is a norm

$$\|f\|_2 = (f, f)^{1/2} = \Big(\int_0^{2\pi} |f(x)|^2\,\mathrm{d}x\Big)^{1/2}$$

called **quadratic norm**. As any other norm, the quadratic norm enjoys the following characteristic properties:

i) $\|f\|_2 \geq 0$ for any $f \in \tilde{\mathcal{C}}_{2\pi}$ and $\|f\|_2 = 0$ if and only if $f = 0$;

ii) $\|\alpha f\|_2 = |\alpha|\, \|f\|_2$, for all $f \in \tilde{\mathcal{C}}_{2\pi}$ and any $\alpha \in \mathbb{R}$;

iii) $\|f + g\|_2 \leq \|f\|_2 + \|g\|_2$, for any $f,\, g \in \tilde{\mathcal{C}}_{2\pi}$.

The **Cauchy-Schwarz inequality**

$$|(f,\, g)| \leq \|f\|_2\, \|g\|_2\,, \qquad \forall f,\, g \in \tilde{\mathcal{C}}_{2\pi}\,, \tag{3.3}$$

holds.

Let us now recall that a family $\{f_k\}$ of non-zero maps in $\tilde{\mathcal{C}}_{2\pi}$ is said an **orthogonal system** if

$$(f_k,\, f_\ell) = 0 \qquad \text{for } k \neq \ell\,.$$

By *normalising* $\hat{f}_k = \dfrac{f_k}{\|f_k\|_2}$ for any k, an orthogonal system generates an **orthonormal system** $\{\hat{f}_k\}$,

$$(\hat{f}_k,\, \hat{f}_\ell) = \delta_{k\ell} = \begin{cases} 1 & \text{if } k = \ell\,, \\ 0 & \text{if } k \neq \ell\ . \end{cases}$$

The set of functions

$$\begin{aligned} \mathcal{F} &= \Big\{1,\, \cos x,\, \sin x, \ldots, \cos kx,\, \sin kx, \ldots\Big\} \\ &= \Big\{\varphi_k(x) = \cos kx : k \geq 0\Big\} \cup \Big\{\psi_k(x) = \sin kx : k \geq 1\Big\} \end{aligned} \tag{3.4}$$

forms an orthogonal system in $\tilde{\mathcal{C}}_{2\pi}$, whose associated orthonormal system is

$$\hat{\mathcal{F}} = \Big\{\frac{1}{\sqrt{2\pi}},\, \frac{1}{\sqrt{\pi}}\cos x,\, \frac{1}{\sqrt{\pi}}\sin x, \ldots, \frac{1}{\sqrt{\pi}}\cos kx,\, \frac{1}{\sqrt{\pi}}\sin kx, \ldots\Big\}\,. \tag{3.5}$$

This follows from the relationships

$$\boxed{\begin{aligned} &\int_0^{2\pi} \cos^2 kx \,\mathrm{d}x = \int_0^{2\pi} \sin^2 kx \,\mathrm{d}x = \pi\,, \quad \forall k \geq 1 \\ &\int_0^{2\pi} \cos kx\, \cos \ell x \,\mathrm{d}x = \int_0^{2\pi} \sin kx\, \sin \ell x \,\mathrm{d}x = 0\,, \quad \forall k,\ell \geq 0,\ k \neq \ell\,, \\ &\int_0^{2\pi} \cos kx\, \sin \ell x \,\mathrm{d}x = 0\,, \quad \forall k,\ell \geq 0\,. \end{aligned}} \tag{3.6}$$

The above orthonormal system in $\tilde{\mathcal{C}}_{2\pi}$ plays the role of the canonical orthonormal basis $\{\boldsymbol{e}_k\}$ of $\mathbb{R}^n$: each element in the vector space is uniquely writable as linear combination of the elements of the system, with the difference that now the linear combination is an infinite series. Fourier series are precisely the expansions in series of maps in $\tilde{\mathcal{C}}_{2\pi}$, viewed as formal combinations of the orthonormal system (3.5).

A crucial feature of this expansion is the possibility of approximating a map by a linear combination of simple functions. Precisely, for any $n \geq 0$ we consider the $(2n+1)$-dimensional subset $\mathcal{P}_n \subset \tilde{\mathcal{C}}_{2\pi}$ of trigonometric polynomials of degree $\leq n$. This subspace is spanned by maps φ_k and ψ_k of $\mathcal{F}$ with index $k \leq n$, forming a finite orthogonal system $\mathcal{F}_n$. A "natural" approximation in $\mathcal{P}_n$ of a map of $\tilde{\mathcal{C}}_{2\pi}$ is provided by its orthogonal projection on $\mathcal{P}_n$.

Definition 3.8 *We call* **orthogonal projection** *of $f \in \tilde{\mathcal{C}}_{2\pi}$ on $\mathcal{P}_n$ the element $S_{n,f} \in \mathcal{P}_n$ defined by*

$$S_{n,f}(x) = a_0 + \sum_{k=1}^{n}(a_k \cos kx + b_k \sin kx)\,, \tag{3.7}$$

where a_k, b_k are

$$\begin{aligned} a_0 &= \frac{1}{2\pi}\int_0^{2\pi} f(x)\,\mathrm{d}x\,; \\ a_k &= \frac{1}{\pi}\int_0^{2\pi} f(x)\cos kx\,\mathrm{d}x\,, \qquad k \geq 1\,; \\ b_k &= \frac{1}{\pi}\int_0^{2\pi} f(x)\sin kx\,\mathrm{d}x\,, \qquad k \geq 1\,. \end{aligned} \tag{3.8}$$

Note $S_{n,f}$ can be written as

$$S_{n,f}(x) = \sum_{k=0}^{n} a_k\varphi_k(x) + \sum_{k=1}^{n} b_k\psi_k(x)$$

and the coefficients a_k, b_k become

$$a_k = \frac{(f,\varphi_k)}{(\varphi_k,\,\varphi_k)} \quad \text{for } k \geq 0\,, \qquad b_k = \frac{(f,\psi_k)}{(\psi_k,\,\psi_k)} \quad \text{for } k \geq 1\,. \tag{3.9}$$

There is an equivalent representation with respect to the finite orthonormal system $\hat{\mathcal{F}}_n$ made by the elements of $\hat{\mathcal{F}}$ of index $k \leq n$; in fact

$$S_{n,f}(x) = \sum_{k=0}^{n} \hat{a}_k\hat{\varphi}_k(x) + \sum_{k=1}^{n} \hat{b}_k\hat{\psi}_k(x)$$

where

$$\hat{a}_k = (f,\,\hat{\varphi}_k) \quad \text{for } k \geq 0\,, \qquad \hat{b}_k = (f,\,\hat{\psi}_k) \quad \text{for } k \geq 1\,. \tag{3.10}$$

The equivalence follows from

$$a_k = \frac{1}{\|\varphi_k\|_2^2}\,(f,\varphi_k) = \frac{1}{\|\varphi_k\|_2}\Big(f,\,\frac{\varphi_k}{\|\varphi_k\|_2}\Big) = \frac{1}{\|\varphi_k\|_2}\,\hat{a}_k$$

hence

$$a_k\varphi_k(x) = \hat{a}_k\frac{\varphi_k(x)}{\|\varphi_k\|_2} = \hat{a}_k\hat{\varphi}_k(x)\,;$$

similarly one proves $b_k\psi_k(x) = \hat{b}_k\hat{\psi}_k(x)$.

To understand the properties of the orthogonal projection of f on $\mathcal{P}_n$, we stress that the quadratic norm defines a **distance** on $\tilde{\mathcal{C}}_{2\pi}$:

$$d(f,\,g) = \|f-g\|_2\,, \qquad f,\,g \in \tilde{\mathcal{C}}_{2\pi}\,. \tag{3.11}$$

The number $\|f-g\|_2$ measures how "close" f and g are. The expected properties of distances hold:

i) $d(f,\,g) \geq 0$ for any $f,\,g \in \tilde{\mathcal{C}}_{2\pi}$, and $d(f,\,g) = 0$ precisely if $f = g$;

ii) $d(f,\,g) = d(g,\,f)$, for all $f,\,g \in \tilde{\mathcal{C}}_{2\pi}$;

iii) $d(f,\,g) \leq d(f,\,h) + d(h,\,g)$, for any $f,\,g,\,h \in \tilde{\mathcal{C}}_{2\pi}$.

The orthogonal projection of f on $\mathcal{P}_n$ enjoys therefore the following properties, some of which are symbolically represented in Fig. 3.2.

Proposition 3.9 *i) The function $f - S_{n,f}$ is orthogonal to every element of $\mathcal{P}_n$, and $S_{n,f}$ is the unique element of $\mathcal{P}_n$ with this property.*

ii) $S_{n,f}$ is the element in $\mathcal{P}_n$ with minimum distance to f with respect to (3.11), *i.e.,*

$$\|f - S_{n,f}\|_2 = \min_{P\in\mathcal{P}_n} \|f-P\|_2\,.$$

Hence $S_{n,f}$ is the polynomial of $\mathcal{P}_n$ that best approximates f in quadratic norm.

iii) The **minimum square error** *$\|f - S_{n,f}\|_2$ of f satisfies*

$$\|f - S_{n,f}\|_2^2 = \int_0^{2\pi} |f(x)|^2\,\mathrm{d}x - 2\pi a_0^2 - \pi\sum_{k=1}^{n}(a_k^2+b_k^2)\,. \tag{3.12}$$

Proof. This proposition transfers to $\tilde{\mathcal{C}}_{2\pi}$ the general properties of the orthogonal projection of a vector, belonging to a vector space equipped with a dot product, onto the subspace generated by a finite orthogonal system. For the reader not familiar with such abstract notions, we provide an adapted proof.

For *i)*, let

$$P(x) = \sum_{k=0}^{n} \tilde{a}_k\varphi_k(x) + \sum_{k=1}^{n} \tilde{b}_k\psi_k(x)$$

be a generic element of $\mathcal{P}_n$. Then $f - P$ is orthogonal to every element in $\mathcal{P}_n$ if and only if, for any $k \leq n$,

$$(f-P,\varphi_k) = 0 \qquad \text{and} \qquad (f-P,\psi_k) = 0\,.$$

Using the orthogonality of φ_k, ψ_k, that is equivalent to

$$(f,\varphi_k) - \tilde{a}_k(\varphi_k,\,\varphi_k) = 0 \qquad \text{and} \qquad (f,\psi_k) - \tilde{b}_k(\psi_k,\,\psi_k) = 0\,.$$

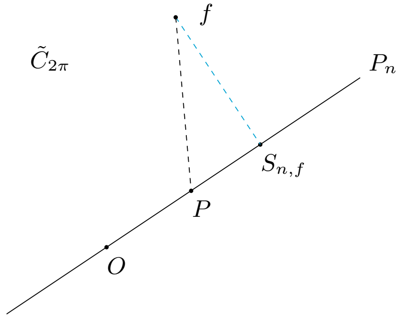


Figure 3.2. Orthogonal projection of $f \in \tilde{\mathcal{C}}_{2\pi}$ on $\mathcal{P}_n$

Hence $\tilde{a}_k = a_k$ and $\tilde{b}_k = b_k$ for any $k \leq n$, recalling (3.9). In other words, $P = S_{n,f}$.
To prove *ii)*, note first that

$$\|f+g\|_2^2 = \|f\|_2^2 + \|g\|_2^2 + 2(f,\, g)\,, \qquad \forall f,\, g \in \tilde{\mathcal{C}}_{2\pi}$$

by definition of norm. Writing $\|f-P\|^2$ as $\|(f-S_{n,f})+(S_{n,f}-P)\|^2$, and using the previous relationship with the fact that $f - S_{n,f}$ is orthogonal to $S_{n,f} - P \in \mathcal{P}_n$ by *i)*, we obtain

$$\|f-P\|_2^2 = \|f-S_{n,f}\|_2^2 + \|S_{n,f}-P\|_2^2\,;$$

the equation is to be considered as a generalisation of Pythagoras' Theorem to spaces with a scalar product (Fig. 3.2). Then

$$\|f-P\|_2^2 \geq \|f-S_{n,f}\|_2^2\,, \qquad \forall P \in \mathcal{P}_n\,,$$

and equality holds if and only if $S_{n,f} - P = 0$ if and only if $P = S_{n,f}$. This proves *ii)*.
Claim *iii)* follows from

$$\begin{aligned}\|f-S_{n,f}\|_2^2 &= \big(f-S_{n,f},\, f-S_{n,f}\big)\\ &= \big(f,\, f-S_{n,f}\big) - \big(S_{n,f},\, f-S_{n,f}\big) = \big(f,\, f-S_{n,f}\big)\\ &= \|f\|_2^2 - \big(f,\, S_{n,f}\big)\end{aligned}$$

and

$$\big(f,\, S_{n,f}\big) = \int_0^{2\pi} f(x)\, S_{n,f}(x)\, \mathrm{d}x = 2\pi a_0^2 + \pi \sum_{k=1}^{n} (a_k^2 + b_k^2)\,. \qquad \square$$

At this juncture it becomes natural to see if, and in which sense, the polynomial sequence $S_{n,f}$ converges to f as $n \to \infty$. These are the partial sums of the series

$$a_0 + \sum_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx)\,,$$

where the coefficients a_k, b_k are given by (3.8). We are thus asking about the convergence of the series.

Definition 3.10 *The* **Fourier series** *of* $f \in \tilde{\mathcal{C}}_{2\pi}$ *is the series of functions*

$$a_0 + \sum_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx)\,, \tag{3.13}$$

where a_0, a_k, b_k $(k \geq 1)$ *are the real numbers* (3.8) *and are called the* **Fourier coefficients** *of* f. *We shall write*

$$f \approx a_0 + \sum_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx)\,. \tag{3.14}$$

The symbol $\approx$ means that the right-hand side of (3.14) represents the Fourier series of f; more explicitly, the coefficients a_k and b_k are prescribed by (3.8). Due to the ample range of behaviours of a series (see Sect. 2.3), one should expect the Fourier series of f not to converge at all, or to converge to a sum other than f. That is why we shall use the equality sign in (3.14) only in case the series pointwise converges to f. We will soon discuss sufficient conditions for the series to converge in some way or another.

It is possible to define the Fourier series of a periodic, piecewise-continuous but not-necessarily-regularised function. Its series coincides with the one of the regularised function built from f.

Example 3.11

Consider the **square wave** (Fig. 3.3)

$$f(x) = \begin{cases} -1 & \text{if } -\pi < x < 0\,, \\ 0 & \text{if } x = 0,\ \pm\pi\,, \\ 1 & \text{if } 0 < x < \pi\ . \end{cases}$$

By Proposition 3.3, for any $k \geq 0$

$$\int_0^{2\pi} f(x) \cos kx \,\mathrm{d}x = \int_{-\pi}^{\pi} f(x) \cos kx \,\mathrm{d}x = 0\,,$$

as $f(x) \cos kx$ is an odd map, whence $a_k = 0$. Moreover

$$\begin{aligned} b_k &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin kx \,\mathrm{d}x = \frac{2}{\pi} \int_0^{\pi} \sin kx \,\mathrm{d}x \\ &= \frac{2}{k\pi}(1 - \cos k\pi) = \frac{2}{k\pi}(1 - (-1)^k) = \begin{cases} 0 & \text{if } k \text{ is even}\,, \\ \frac{4}{k\pi} & \text{if } k \text{ is odd}\,. \end{cases} \end{aligned}$$

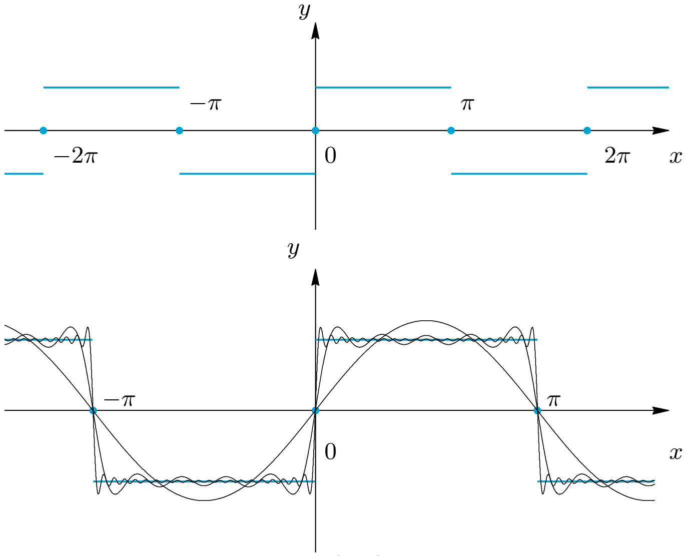


Figure 3.3. Graphs of the square wave (top) and the corresponding polynomials $S_{1,f}(x)$, $S_{9,f}(x)$, $S_{41,f}(x)$ (bottom)

Writing every odd k as $k = 2m+1$, the Fourier series of f reads

$$f \approx \frac{4}{\pi}\sum_{m=0}^{\infty}\frac{1}{2m+1}\sin(2m+1)x\,. \qquad \square$$

The example shows that if the map $f \in \tilde{\mathcal{C}}_{2\pi}$ is symmetric, computing its Fourier coefficients can be simpler. The precise statement goes as follows.

Proposition 3.12 *If the map $f \in \tilde{\mathcal{C}}_{2\pi}$ is odd,*

$$a_k = 0\,, \qquad \forall k \geq 0\,,$$
$$b_k = \frac{2}{\pi}\int_0^{\pi} f(x)\sin kx\,\mathrm{d}x\,, \qquad \forall k \geq 1\,;$$

If f is even,

$$b_k = 0\,, \qquad \forall k \geq 1\,,$$
$$a_0 = \frac{1}{\pi}\int_0^{\pi} f(x)\,\mathrm{d}x\,; \qquad a_k = \frac{2}{\pi}\int_0^{\pi} f(x)\cos kx\,\mathrm{d}x\,, \qquad \forall k \geq 1\,.$$

Proof. Take, for example, f even. Recalling Proposition 3.3, it suffices to note $f(x)\sin kx$ is odd for any $k \geq 1$, and $f(x)\cos kx$ is even for any $k \geq 0$, to obtain that

$$\int_0^{2\pi} f(x)\sin kx\,\mathrm{d}x = \int_{-\pi}^{\pi} f(x)\sin kx\,\mathrm{d}x = 0\,, \qquad \forall k \geq 1\,,$$

and, for any $k \geq 0$,

$$\int_0^{2\pi} f(x)\cos kx\,\mathrm{d}x = \int_{-\pi}^{\pi} f(x)\cos kx\,\mathrm{d}x = 2\int_0^{\pi} f(x)\cos kx\,\mathrm{d}x\,.$$

□

Example 3.13

Let us determine the Fourier series for the **rectified wave** $f(x) = |\sin x|$ (Fig. 3.4). As f is even, the b_k vanish and we just need to compute a_k for $k \geq 0$:

$$\begin{aligned}
a_0 &= \frac{1}{2\pi}\int_{-\pi}^{\pi} |\sin x|\,\mathrm{d}x = \frac{1}{\pi}\int_0^{\pi} \sin x\,\mathrm{d}x = \frac{2}{\pi}\,,\\
a_1 &= \frac{1}{\pi}\int_0^{\pi} \sin 2x\,\mathrm{d}x = 0\,,\\
a_k &= \frac{2}{\pi}\int_0^{\pi} \sin x \cos kx\,\mathrm{d}x = \frac{1}{\pi}\int_0^{\pi}\Big(\sin(k+1)x - \sin(k-1)x\Big)\,\mathrm{d}x\\
&= \frac{1}{\pi}\Big(\frac{1-\cos(k+1)\pi}{k+1} - \frac{1-\cos(k-1)\pi}{k-1}\Big)\\
&= \begin{cases} 0 & \text{if } k \text{ is odd}\,,\\ -\dfrac{4}{\pi(k^2-1)} & \text{if } k \text{ is even}\,,\end{cases} \qquad \forall k > 1\,.
\end{aligned}$$

The Fourier series of the rectified wave thus reads

$$f \approx \frac{2}{\pi} - \frac{4}{\pi}\sum_{m=1}^{\infty} \frac{1}{4m^2-1}\cos 2mx\,.$$

□

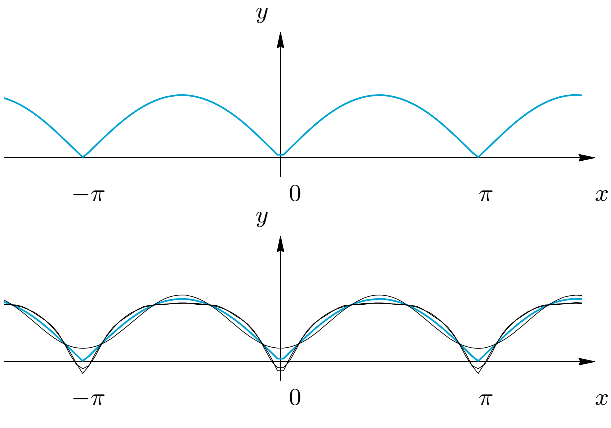


Figure 3.4. Graphs of the rectified wave (top) and the corresponding polynomials $S_{2,f}(x)$, $S_{10,f}(x)$, $S_{30,f}(x)$ (bottom)

3.3 Exponential form

The exponential form is an alternative, but equivalent, way of representing a Fourier series; it is more concise and sometimes easier to use, but the price for this is that it requires complex numbers.

The starting point is the Euler identity (Vol. I, Eq. (8.31))

$$e^{i\theta} = \cos\theta + i\sin\theta\,, \theta \qquad \in \mathbb{R}\,; \tag{3.15}$$

setting $\theta = \pm kx$, $k \geq 1$ integer, we may write $\cos kx$ and $\sin kx$ as linear combinations of the functions e^{ikx}, e^{-ikx}:

$$\cos kx = \frac{1}{2}(e^{ikx} + e^{-ikx})\,, \qquad \sin kx = \frac{1}{2i}(e^{ikx} - e^{-ikx})\,.$$

On the other hand, $1 = e^{i0x}$ trivially. Substituting in (3.14) and rearranging terms yields

$$\boxed{f \approx \sum_{k=-\infty}^{+\infty} c_k e^{ikx}\,,} \tag{3.16}$$

where

$$\boxed{c_0 = a_0\,, \qquad c_k = \frac{a_k - ib_k}{2}\,, \qquad c_{-k} = \frac{a_k + ib_k}{2}\,, \quad \text{for } k \geq 1\,.} \tag{3.17}$$

It is an easy task to check

$$\boxed{c_k = \frac{1}{2\pi}\int_0^{2\pi} f(x)e^{-ikx}\,dx\,, \qquad k \in \mathbb{Z}\,.}$$

Expression (3.16) represents the Fourier series of f in exponential form, and the coefficients c_k are the complex Fourier coefficients of f.

The complex Fourier series embodies the (formal) expansion of a function $f \in \tilde{\mathcal{C}}_{2\pi}$ with respect to the orthogonal system of functions e^{ikx}, $k \in \mathbb{Z}$. In fact,

$$(e^{ikx}, e^{i\ell x}) = 2\pi\delta_{kl} = \begin{cases} 2\pi & \text{if } k = \ell\,, \\ 0 & \text{if } k \neq \ell\,, \end{cases}$$

where

$$(f, g) = \int_0^{2\pi} f(x)\overline{g(x)}\,dx$$

is defined on $\tilde{\mathcal{C}}^*_{2\pi}$, the set of complex-valued maps $f = f_r + if_i : \mathbb{R} \to \mathbb{C}$ whose real part f_r and imaginary part f_i belong to $\tilde{\mathcal{C}}_{2\pi}$.

For this system the complex Fourier coefficients of f become

$$c_k = \frac{(f, \mathrm{e}^{ikx})}{(\mathrm{e}^{ikx}, \mathrm{e}^{ikx})}, \qquad k \in \mathbb{Z}, \tag{3.18}$$

in analogy to (3.9). Since this formula makes sense for all maps in $\tilde{\mathcal{C}}^*_{2\pi}$, equation (3.16) might define Fourier series on $\tilde{\mathcal{C}}^*_{2\pi}$ as well.

It is straighforward that $f \in \tilde{\mathcal{C}}^*_{2\pi}$ is a real map ($f_\mathrm{i} = 0$) if and only if its complex Fourier coefficients satisfy $c_{-k} = \overline{c}_k$, for any $k \in \mathbb{Z}$. If so, the real Fourier coefficients of f are just

$$\boxed{a_0 = c_0\,, \qquad a_k = c_k + c_{-k}\,, \qquad b_k = i(c_k - c_{-k})\,, \quad \text{for } k \geq 1\,.} \tag{3.19}$$

3.4 Differentiation

Consider the real Fourier series (3.13) and let us differentiate it (formally) term by term. This gives

$$\alpha_0 + \sum_{k=1}^{\infty} (\alpha_k \cos kx + \beta_k \sin kx)\,,$$

with

$$\alpha_0 = 0\,, \qquad \alpha_k = kb_k\,, \qquad \beta_k = -ka_k\,, \quad \text{for } k \geq 1\,. \tag{3.20}$$

Supposing $f \in \tilde{\mathcal{C}}_{2\pi}$ is $\mathcal{C}^1$ on $\mathbb{R}$, in which case the derivative f' still belongs to $\tilde{\mathcal{C}}_{2\pi}$, the previous expression coincides with the Fourier series of f'. In fact,

$$\alpha_0 = \frac{1}{2\pi}\int_0^{2\pi} f'(x)\,\mathrm{d}x = \frac{1}{2\pi}\big(f(2\pi) - f(0)\big) = 0$$

by periodicity. Moreover, for $k \geq 1$, integrating by part gives

$$\begin{aligned} \alpha_k &= \frac{1}{\pi}\int_0^{2\pi} f'(x)\cos kx\,\mathrm{d}x \\ &= \frac{1}{\pi}\Big[f(x)\cos kx\Big]_0^{2\pi} + \frac{k}{\pi}\int_0^{2\pi} f(x)\sin kx\,\mathrm{d}x = kb_k\,, \end{aligned}$$

and similarly, $\beta_k = -ka_k$.

In summary,

$$\boxed{f' \approx \sum_{k=1}^{\infty} (k\,b_k \cos kx - k\,a_k \sin kx)\,.} \tag{3.21}$$

A similar reasoning shows that such representation holds under weaker hypotheses on the differentiability of f, e.g., if f is piecewise $\mathcal{C}^1$ (see Definition 3.24).

The derivatives' series becomes all the more explicit if we exploit the complex form (3.16). From (3.15) in fact,

$$\frac{\mathrm{d}}{\mathrm{d}\theta}\,\mathrm{e}^{i\theta} = (-\sin\theta) + i\cos\theta = i(\cos\theta + i\sin\theta) = i\mathrm{e}^{i\theta}\,.$$

Therefore

$$\frac{\mathrm{d}}{\mathrm{d}x}\,\mathrm{e}^{ikx} = ik\mathrm{e}^{ikx}\,,$$

whence

$$\boxed{f' \approx \sum_{k=-\infty}^{+\infty} ikc_k\mathrm{e}^{ikx}\,.}$$

If f is $\mathcal{C}^r$ on $\mathbb{R}$, $r \geq 1$, this fact generalises in the obvious way:

$$f^{(r)} \approx \sum_{k=-\infty}^{+\infty} (ik)^r c_k\mathrm{e}^{ikx}\,;$$

Thus, the kth complex Fourier coefficient γ_k of the rth derivative of f is

$$\boxed{\gamma_k = (ik)^r c_k\,.} \tag{3.22}$$

3.5 Convergence of Fourier series

This section is devoted to the convergence properties of the Fourier series of a piecewise-continuous, periodic map of period 2π (not regularised necessarily). We shall treat three kinds of convergence: quadratic, pointwise and uniform. We omit the proofs of all theorems, due to their prevailingly technical nature[1].

3.5.1 Quadratic convergence

We begin by the definition.

Definition 3.14 *Let f and f_k, $k \geq 0$, be square-integrable functions defined on a closed and bounded interval $[a, b]$. The series $\sum_{k=0}^{\infty} f_k$* **converges in quadratic norm** *to f on $[a, b]$ if*

$$\lim_{n\to\infty} \int_a^b \Big| f(x) - \sum_{k=0}^{n} f_k(x) \Big|^2 \mathrm{d}x = \lim_{n\to\infty} \Big\| f - \sum_{k=0}^{n} f_k \Big\|_2^2 = 0\,.$$

[1] A classical text where the interested reader may find these proofs is Y. Katznelson's, *Introduction to Harmonic Analysis*, Cambridge University Press, 2004.

Directly from this, the uniform convergence of $\sum_{k=0}^{\infty} f_k$ to f implies the quadratic convergence, for

$$\begin{aligned}\int_a^b \Big|f(x) - \sum_{k=0}^{n} f_k(x)\Big|^2 \,\mathrm{d}x &\le \int_a^b \Big(\sup_{x\in[a,b]} \Big|f(x) - \sum_{k=0}^{n} f_k(x)\Big|\Big)^2 \,\mathrm{d}x \\ &\le (b-a) \Big\|f - \sum_{k=0}^{n} f_k\Big\|^2_{\infty,[a,b]};\end{aligned}$$

hence, if the last expression is infinitesimal as $n \to \infty$, so is the first one.

Quadratic convergence of the Fourier series of a map in $\tilde{\mathcal{C}}_{2\pi}$ is guaranteed by the next fundamental result, whose proof we omit.

Theorem 3.15 *The Fourier series of $f \in \tilde{\mathcal{C}}_{2\pi}$ converges to f in quadratic norm:*

$$\lim_{n\to+\infty} \|f - S_{n,f}\|_2 = 0\,.$$

Let us describe some consequences.

Corollary 3.16 *Any $f \in \tilde{\mathcal{C}}_{2\pi}$ satisfies* **Parseval's formula***:*

$$\int_0^{2\pi} |f(x)|^2 \,\mathrm{d}x = 2\pi a_0^2 + \pi \sum_{k=1}^{+\infty} (a_k^2 + b_k^2)\,. \tag{3.23}$$

Proof. This is an easy corollary of the above theorem. By (3.12) in fact,

$$\begin{aligned}0 = \lim_{n\to+\infty} \|f - S_{n,f}\|_2 &= \lim_{n\to+\infty} \Big(\int_0^{2\pi} |f(x)|^2 \,\mathrm{d}x - 2\pi a_0^2 - \pi \sum_{k=1}^{n} (a_k^2 + b_k^2)\Big) \\ &= \int_0^{2\pi} |f(x)|^2 \,\mathrm{d}x - 2\pi a_0^2 - \pi \sum_{k=1}^{\infty} (a_k^2 + b_k^2)\,.\end{aligned}$$

□

Corollary 3.17 (**Riemann-Lebesgue Lemma**) *Given $f \in \tilde{\mathcal{C}}_{2\pi}$,*

$$\lim_{k\to+\infty} a_k = \lim_{k\to+\infty} b_k = 0\,.$$

Proof. From Parseval's identity (3.23) the series $\sum_{k=1}^{\infty} (a_k^2 + b_k^2)$ converges. Therefore its general term $a_k^2 + b_k^2$ goes to zero as $k \to +\infty$, and the result follows. □

If $f \in \tilde{\mathcal{C}}^*_{2\pi}$ is expanded in complex Fourier series (3.16), Parseval's formula becomes

$$\boxed{\int_0^{2\pi} |f(x)|^2 \,\mathrm{d}x = 2\pi \sum_{k=-\infty}^{+\infty} |c_k|^2 \,,}$$

while the Riemann-Lebesgue Lemma says

$$\lim_{k\to\infty} c_k = 0\,.$$

Corollary 3.16 is useful to compute sums of numerical series.

Example 3.18

The map $f(x) = x$, defined on $(-\pi,\pi)$ and prolonged by periodicity to all $\mathbb{R}$ (**sawtooth wave**) (Fig. 3.5), has a simple Fourier series. The map is odd, so $a_k = 0$ for all $k \geq 0$, and $b_k = \dfrac{2}{k}(-1)^{k+1}$, hence

$$f \approx \sum_{k=1}^{\infty} \frac{2}{k}(-1)^{k+1} \sin kx.$$

The series converges in quadratic norm to $f(x)$, and via Parseval's identity we find

$$\int_{-\pi}^{\pi} |f(x)|^2 \,\mathrm{d}x = \pi \sum_{k=1}^{\infty} b_k^2\,.$$

Since

$$\int_{-\pi}^{\pi} x^2 \,\mathrm{d}x = \frac{2\pi^3}{3}\,,$$

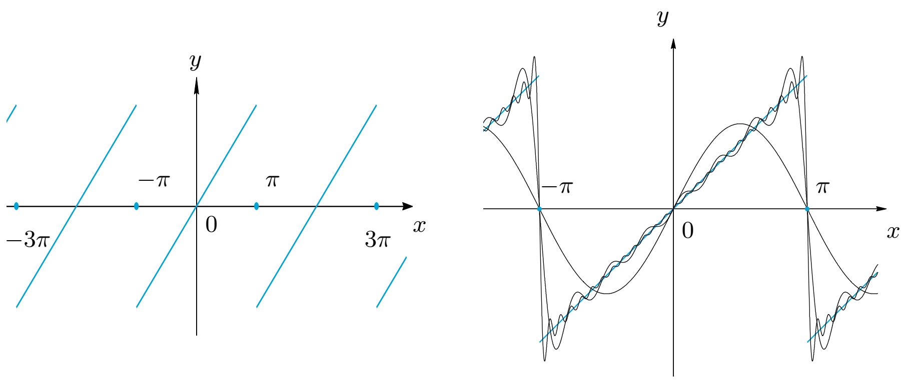


Figure 3.5. Graphs of the sawtooth wave (left) and the corresponding polynomials $S_{1,f}(x)$, $S_{7,f}(x)$, $S_{25,f}(x)$ (right)

we have

$$\frac{2\pi^3}{3} = \pi \sum_{k=1}^{\infty} \frac{4}{k^2},$$

whence the sum of the inverses of all natural numbers squared is

$$\sum_{k=1}^{\infty} \frac{1}{k^2} = \frac{\pi^2}{6};$$

This fact was stated without proof in Example 1.11 i). □

We remark, at last, that additional assumptions of the regularity on f allow to estimate the remainder of the Fourier series in terms of n, and furnish informations on the speed at which the Fourier series tends to f in quadratic norm. For instance, if $f \in \tilde{\mathcal{C}}_{2\pi}$ is $\mathcal{C}^r$ on $\mathbb{R}$, $r \geq 1$, it can be proved that

$$\|f - S_{n,f}\|_2 \leq \frac{1}{n^r} \|f^{(r)}\|_2 .$$

This is coherent with what will happen in Sect. 3.5.4.

3.5.2 Pointwise convergence

We saw that the Fourier series of $f \in \tilde{\mathcal{C}}_{2\pi}$ converges to f in quadratic norm, so in a suitable integral sense; this, though, does not warrant pointwise convergence. We are thus left with the hard task of finding conditions that ensure pointwise convergence: alas, not even assuming f continuous guarantees the Fourier series will converge. On the other hand the uniform convergence of the Fourier series implies pointwise convergence (see the remark after Definition 2.15). As the trigonometric polynomials are continuous, uniform convergence still requires f be continuous (by Theorem 2.17). We shall state, without proving them, some sufficient conditions for the pointwise convergence of the Fourier series of a non-necessarily continuous map. The first ones guarantee convergence on the entire interval $[0, 2\pi]$. But before that, we introduce a piece of notation.

Definition 3.19 *i) A function f is called* **piecewise regular** *on an interval $[a, b] \subset \mathbb{R}$ in case*

a) it is differentiable everywhere on $[a, b]$ except at a finite number of points at most;
b) it is piecewise continuous, together with its derivative f'.

ii) f is **piecewise monotone** *if the interval $[a, b]$ can be divided in a finite number of sub-intervals where f is monotone.*

Theorem 3.20 *Let $f \in \tilde{\mathcal{C}}_{2\pi}$ and suppose one of the following holds:*

a) f is piecewise regular on $[0,\, 2\pi]$;
b) f is piecewise monotone on $[0,\, 2\pi]$.

Then the Fourier series of f converges pointwise to f on $[0,\, 2\pi]$.

The theorem also holds under the assumption that f is not regularised; in such a case (like for the square wave or the sawtooth function),

at a discontinuity point x_0, the Fourier series converges to the regularised value

$$\frac{f(x_0^-) + f(x_0^+)}{2}$$

of f, and not to $f(x_0)$.

Now let us see a local condition for pointwise convergence.

Definition 3.21 *A piecewise-continuous map admits* **left pseudo-derivative** *and* **right pseudo-derivative** *at $x_0 \in \mathbb{R}$ if the following respective limits exist and are finite*

$$f'(x_0^-) = \lim_{x \to x_0^-} \frac{f(x) - f(x_0^-)}{x - x_0}, \qquad f'(x_0^+) = \lim_{x \to x_0^+} \frac{f(x) - f(x_0^+)}{x - x_0}.$$

(Fig. 3.6 explains the geometric meaning of pseudo-derivatives.) If in addition f is continuous at x_0, the pseudo-derivatives are nothing else than the left and right derivatives.

Note that a piecewise-regular function on $[0, 2\pi]$ admits pseudo-derivatives at each point of $[0,\, 2\pi]$; despite this, there are functions that are not piecewise regular yet admit pseudo-derivatives everywhere in $\mathbb{R}$ (an example is $f(x) = x^2 \sin \frac{1}{x}$, for $x \neq 0,\ x \in [-\pi, \pi\,]$ and $f(0) = 0$).

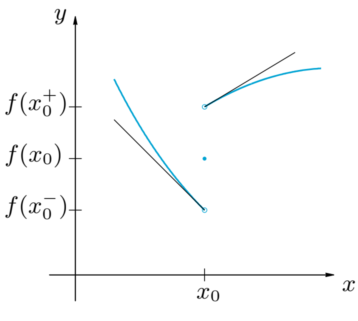


Figure 3.6. Geometric meaning of the left and right pseudo-derivatives of f at x_0

Theorem 3.22 *Let $f \in \tilde{\mathcal{C}}_{2\pi}$. If at $x_0 \in [0, 2\pi]$ the left and right pseudo-derivatives exist, the Fourier series of f at x_0 converges to the (regularised) value $f(x_0)$.*

Example 3.23

Let us go back to the square wave of Example 3.11. Condition *a)* of Theorem 3.20 holds, so we have pointwise convergence on $\mathbb{R}$. Looking at Fig. 3.3 we can see a special behaviour around a discontinuity. If we take a neighbourhood of $x_0 = 0$, the x-coordinates of points closest to the maximum and minimum points of the nth partial sum tend to x_0 as $n \to \infty$, whereas the y-coordinates tend to different limits $\ell_\pm \sim \pm 1.18$; the latter are not the limits $f(0^\pm) = \pm 1$ of f at 0. Such anomaly appears every time one considers a discontinuous map in $\tilde{\mathcal{C}}_{2\pi}$, and goes under the name of **Gibbs phenomenon**. □

3.5.3 Uniform convergence

As already noted, there is no uniform convergence for the Fourier series of a discontinuous function, and we know that continuity is not sufficient (it does not guarantee pointwise convergence either).

Let us then introduce a new class of maps.

Definition 3.24 *A function $f \in \tilde{\mathcal{C}}_{2\pi}$ is* **piecewise $\mathcal{C}^1$** *if it is continuous on $\mathbb{R}$ and piecewise regular on $[0, 2\pi]$.*

The square wave is not piecewise $\mathcal{C}^1$ (since not continuous), in contrast to the rectified wave.

We now state the following important theorem.

Theorem 3.25 *Let $f \in \tilde{\mathcal{C}}_{2\pi}$ be piecewise $\mathcal{C}^1$. Its Fourier series converges uniformly to f everywhere on $\mathbb{R}$.*

More generally, the following localization principle holds.

Theorem 3.26 *Let $f \in \tilde{\mathcal{C}}_{2\pi}$ be piecewise regular on $[0, 2\pi]$. Its Fourier series converges uniformly to f on any closed sub-interval where the map is continuous.*

Example 3.27

i) The rectified wave has uniformly convergent Fourier series on $\mathbb{R}$ (Fig. 3.4).

ii) The Fourier series of the square wave converges uniformly to the function on every interval $[\varepsilon, \pi - \varepsilon]$ or $[\pi + \varepsilon, 2\pi - \varepsilon]$ $(0 < \varepsilon < \pi/2)$, because the square wave is piecewise regular on $[0, 2\pi]$ and continuous on $(0, \pi)$ and $(\pi, 2\pi)$. □

3.5.4 Decay of Fourier coefficients

The equations of Sect. 3.4, relating the Fourier coefficients of a map and its derivatives, help to establish a link between the coefficients' asymptotic behaviour, as $|k| \to \infty$, and the regularity of the function. For simplicity we consider complex coefficients. Let $f \in \tilde{\mathcal{C}}_{2\pi}$ be of class $\mathcal{C}^r$ on $\mathbb{R}$, with $r \geq 1$; from (3.22) we obtain, for any $k \neq 0$,

$$|c_k| = \frac{1}{|k|^r}|\gamma_k|\,.$$

The sequence $|\gamma_k|$ is bounded for $|k| \to \infty$; in fact, using (3.18) on $f^{(r)}$ gives

$$\gamma_k = \frac{(f^{(r)}, \mathrm{e}^{ikx})}{(\mathrm{e}^{ikx}, \mathrm{e}^{ikx})}\,;$$

the inequality of Schwarz tells

$$|\gamma_k| \leq \frac{\|f^{(r)}\|_2 \|\mathrm{e}^{ikx}\|_2}{\|\mathrm{e}^{ikx}\|_2^2} = \frac{1}{\sqrt{2}}\|f^{(r)}\|_2\,.$$

Thus we have proved

$$\boxed{|c_k| = O\left(\frac{1}{|k|^r}\right) \qquad \text{as } |k| \to \infty\,.}$$

The result for real coefficients, i.e.,

$$\boxed{|a_k|,\ |b_k| = O\left(\frac{1}{k^r}\right) \qquad \text{for } k \to +\infty\,,}$$

is similarly found using (3.19) or a direct computation. In any case, if f has period 2π and is $\mathcal{C}^r$ on $\mathbb{R}$, its Fourier coefficients are infinitesimal of order at least r with respect to the test function $1/|k|$.

Vice versa, it can be proved that the speed of decay of Fourier coefficients determines, in a suitable sense, the function's regularity.

3.6 Periodic functions with period $T > 0$

In case a function f belongs to $\tilde{\mathcal{C}}_T$, i.e., is defined on $\mathbb{R}$, piecewise continuous on $[0, T]$, regularised and periodic of period $T > 0$, its Fourier series assumes the form

$$\boxed{f \approx a_0 + \sum_{k=1}^{\infty}\Big(a_k \cos k\frac{2\pi}{T}x + b_k \sin k\frac{2\pi}{T}x\Big)\,,}$$

where

$$\boxed{\begin{aligned} a_0 &= \frac{1}{T}\int_0^T f(x)\,\mathrm{d}x\,, \\ a_k &= \frac{2}{T}\int_0^T f(x)\cos k\frac{2\pi}{T}x\,\mathrm{d}x\,, \qquad k \geq 1\,, \\ b_k &= \frac{2}{T}\int_0^T f(x)\sin k\frac{2\pi}{T}x\,\mathrm{d}x\,, \qquad k \geq 1\,. \end{aligned}}$$

The theorems concerning quadratic, pointwise and uniform convergence of the Fourier series of a map in $\tilde{\mathcal{C}}_{2\pi}$ transfer in the obvious manner to maps in $\tilde{\mathcal{C}}_T$. Parseval's formula reads

$$\boxed{\int_0^T |f(x)|^2\,\mathrm{d}x = Ta_0^2 + \frac{T}{2}\sum_{k=1}^{\infty}(a_k^2 + b_k^2)\,.}$$

As far as the Fourier series' exponential form is concerned, (3.16) must be replaced by

$$\boxed{f \approx \sum_{k=-\infty}^{+\infty} c_k e^{ik\frac{2\pi}{T}x}\,,} \tag{3.24}$$

where

$$\boxed{c_k = \frac{1}{T}\int_0^T f(x)e^{-ik\frac{2\pi}{T}x}\,\mathrm{d}x\,, \qquad k \in \mathbb{Z}\,;}$$

Parseval's identity takes the form

$$\boxed{\int_0^T |f(x)|^2\,\mathrm{d}x = T\sum_{k=-\infty}^{+\infty}|c_k|^2\,.}$$

Example 3.28

Let us write the Fourier expansion for $f(x) = 1 - x^2$ on $I = [-1, 1]$, made periodic of period 2. As f is even, the coefficients b_k are zero. Moreover,

$$a_0 = \frac{1}{2}\int_{-1}^{1}(1-x^2)\,\mathrm{d}x = \int_0^1(1-x^2)\,\mathrm{d}x = \frac{2}{3}\,,$$

$$a_k = \frac{2}{2}\int_{-1}^{1}(1-x^2)\cos k\pi x\,\mathrm{d}x = 2\int_0^1(1-x^2)\cos k\pi x\,\mathrm{d}x = \frac{4}{k^2\pi^2}(-1)^{k+1}$$

for any $k \geq 1$.

Hence

$$f \approx \frac{2}{3} + \frac{4}{\pi^2} \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k^2} \cos k\pi x.$$

Since f is piecewise of class C^1 the convergence is uniform on $\mathbb{R}$, hence also pointwise at every $x \in \mathbb{R}$. We can write

$$f(x) = \frac{2}{3} + \frac{4}{\pi^2} \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k^2} \cos k\pi x\,, \qquad \forall x \in \mathbb{R}.$$

In particular

$$f(0) = 1 = \frac{2}{3} + \frac{4}{\pi^2} \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k^2}\,,$$

whence the sum of the generalised alternating harmonic series can be eventually computed

$$\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k^2} = \frac{\pi^2}{12}. \qquad \square$$

3.7 Exercises

1. *Determine the minimum period of the following maps:*

   a) $f(x) = \cos(3x - 1)$ b) $f(x) = \sin \frac{x}{3} - \cos 4x$

   c) $f(x) = 1 + \cos x + \sin 3x$ d) $f(x) = \sin x \cos x + 5$

   e) $f(x) = 1 + \cos^2 x$ f) $f(x) = |\cos x| + \sin 2x$

2. *Sketch the graph of the functions on $\mathbb{R}$ that on $[0, \pi)$ coincide with $f(x) = \sqrt{x}$ and are:*
   a) *π-periodic;* b) *2π-periodic, even;* c) *2π-periodic, odd.*

3. *Given $f(x) = \cos^3 x + \sin 3x - 4$,*
   a) *determine its minimum period;*
   b) *compute its Fourier series;*
   c) *study quadratic, pointwise, uniform convergence of such expansion.*

4. *Determine the Fourier series expansion of the 2π-periodic maps defined, on $[-\pi, \pi]$, as follows:*

   a) $f(x) = 1 - 2\cos x + |x|$ b) $f(x) = 1 + x + \sin 2x$

   c) $f(x) = 4|\sin^3 x|$ d) $f(x) = \begin{cases} 0 & \textit{if } -\pi \le x < 0\,, \\ -\cos x & \textit{if } 0 \le x < \pi \end{cases}$

5. *Determine the Fourier series of the regularised maps of period $T = 1$ below:*

a) $f(x) = \begin{cases} 1 & \text{if } x \in \left(-\frac{1}{4}, \frac{1}{4}\right), \\ 0 & \text{if } x \in \left(\frac{1}{4}, \frac{3}{4}\right) \end{cases}$ b) $f(x) = |\sin 2\pi x|$

c) $f(x) = \begin{cases} \sin 2\pi x & \text{if } x \in \left[0, \frac{1}{2}\right], \\ 0 & \text{if } x \in \left[\frac{1}{2}, 1\right] \end{cases}$ d) $f(x) = \begin{cases} x & \text{if } x \in \left[0, \frac{1}{2}\right], \\ 1 - x & \text{if } x \in \left[\frac{1}{2}, 1\right] \end{cases}$

6. *Let f be the T-periodic, piecewise-continuous, regularised function whose Fourier series is*

$$f \approx 1 + \sum_{k=1}^{\infty} \frac{k}{(2k+1)^3} \sin 2kx\,.$$

Determine the period T and the symmetries (where present) of f.

7. *Appropriately using the Fourier series of $f(x) = x^2$ and $g(x) = x^3$, calculate the sum of:*

a) $\displaystyle\sum_{k=1}^{\infty} \frac{1}{k^4}$ b) $\displaystyle\sum_{k=1}^{\infty} \frac{1}{k^6}$

8. *Determine the Fourier series of the 2π-periodic map defined on $[-\pi, \pi]$ by*

$$f(x) = \frac{|\varphi(x)| + \varphi(x)}{2}\,,$$

where $\varphi(x) = x^2 - 1$.

9. *Determine the Fourier series for the 2π-periodic, even, regularised function defined on $[0, \pi]$ by*

$$f(x) = \begin{cases} \pi - x & \text{if } 0 \le x < \frac{\pi}{2}\,, \\ 0 & \text{if } \frac{\pi}{2} < x \le \pi\,. \end{cases}$$

Use the expansion to compute the sum of the series

$$\sum_{k=0}^{\infty} \frac{1}{(2k+1)^2}\,.$$

10. *Find the Fourier coefficients for the 2π-periodic, odd map $f(x) = 1 + \sin 2x + \sin 4x$ defined on $[0, \pi]$.*

11. *Consider the 2π-periodic map*

$$f(x) = \begin{cases} \cos 2x & \text{if } |x| \le \frac{\pi}{2}\,, \\ -1 & \text{if } |x| > \frac{\pi}{2} \text{ and } |x| \le \pi \end{cases}$$

defined on $[-\pi,\pi]$. Determine the Fourier series of f and f'; then study the uniform convergence of the two series obtained.

12. *Consider the 2π-periodic function that coincides with $f(x) = x^2$ on $[0, 2\pi]$. Verify its Fourier series is*

$$f \approx \frac{4}{3}\pi^2 + 4\sum_{k=1}^{\infty}\Big(\frac{1}{k^2}\cos kx - \frac{\pi}{k}\sin kx\Big).$$

Study the convergence of the above and use the results to compute the sum of the series $\displaystyle\sum_{k=1}^{\infty}\frac{(-1)^k}{k^2}$.

13. *Consider*

$$f(x) = 2 + \sum_{k=1}^{\infty}\frac{1}{2^k}\sin kx\,, \qquad x \in \mathbb{R}\,.$$

Check that $f \in \mathcal{C}^\infty(\mathbb{R})$. Deduce the Fourier series of f' and the values of $\|f\|_2$ and $\displaystyle\int_0^{2\pi} f(x)\,\mathrm{d}x$.

14. *Consider the 2π-periodic function defined on $[-\pi,\pi)$ as $f(x) = x$. Determine the Fourier series of f and study its quadratic, pointwise and uniform convergence.*

3.7.1 Solutions

1. *Minimum period:*

a) $T = \frac{2}{3}\pi$. b) $T = 6\pi$. c) $T = 2\pi$.

d) $T = \pi$. e) $T = \pi$. f) $T = \pi$.

2. *Graphs of maps:* see Fig. 3.7.

3. a) The minimum period is $T = 2\pi$.

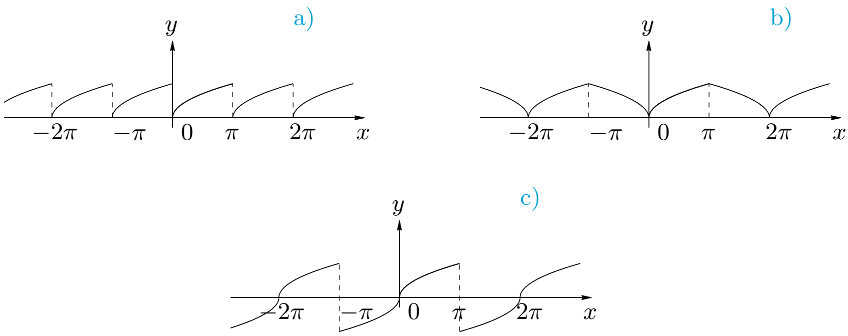


Figure 3.7. The graphs relative to Exercise 2

b) Variously using trigonometric identities one gets

$$\begin{aligned} f(x) &= -4 + \sin 3x + \cos x \cos^2 x \\ &= -4 + \sin 3x + \frac{1}{2}\cos x(1 + \cos 2x) \\ &= -4 + \frac{1}{2}\cos x + \sin 3x + \frac{1}{4}(\cos 3x + \cos x) \\ &= -4 + \frac{3}{4}\cos x + \sin 3x + \frac{1}{4}\cos 3x\,. \end{aligned}$$

c) As f is a trigonometric polynomial, the Fourier series is a finite sum of simple harmonics. Thus all types of convergence hold.

4. *Fourier series' expansions:*

a) The function is the sum of the trigonometric polynomial $1 - 2\cos x$ and the map $g(x) = |x|$. It is sufficient to determine the Fourier series for g. The latter is even, so $b_k = 0$ for all $k \geq 1$. Let us find the coefficients a_k:

$$a_0 = \frac{1}{2\pi}\int_{-\pi}^{\pi} |x|\,\mathrm{d}x = \frac{1}{\pi}\int_0^{\pi} x\,\mathrm{d}x = \frac{\pi}{2}\,,$$

$$\begin{aligned} a_k &= \frac{1}{\pi}\int_{-\pi}^{\pi} |x|\cos kx\,\mathrm{d}x = \frac{2}{\pi}\int_0^{\pi} x\cos kx\,\mathrm{d}x \\ &= \frac{2}{\pi k^2}\big[\cos kx + kx\sin kx\big]_0^{\pi} = \frac{2}{\pi k^2}\big((-1)^k - 1\big) \\ &= \begin{cases} 0 & \text{if } k \text{ even,} \\ -\dfrac{4}{\pi k^2} & \text{if } k \text{ odd}\,, \end{cases} \qquad k \geq 1\,. \end{aligned}$$

In conclusion,

$$g \approx \frac{\pi}{2} - \frac{4}{\pi}\sum_{k=0}^{\infty}\frac{1}{(2k+1)^2}\cos(2k+1)x\,,$$

so

$$f \approx 1 + \frac{\pi}{2} - \Big(2 + \frac{4}{\pi}\Big)\cos x - \frac{4}{\pi}\sum_{k=1}^{\infty}\frac{1}{(2k+1)^2}\cos(2k+1)x\,.$$

b) $f \approx 1 + 2\sin x + 2\sum_{k=3}^{\infty}\frac{(-1)^{k+1}}{k}\sin kx\,.$

c) The map is even, so $b_k = 0$ for all $k \geq 1$. As for the a_k:

$$a_0 = \frac{1}{2\pi}\int_{-\pi}^{\pi}4|\sin^3 x|\,\mathrm{d}x = \frac{4}{\pi}\int_0^{\pi}\sin^3 x\,\mathrm{d}x = \frac{4}{\pi}\left[\frac{\cos^3 x}{3} - \cos x\right]_0^{\pi} = \frac{16}{3\pi}$$

$$\begin{aligned}
a_k &= \frac{1}{\pi}\int_{-\pi}^{\pi}4|\sin^3 x|\cos kx\,\mathrm{d}x = \frac{8}{\pi}\int_0^{\pi}\sin^3 x\cos kx\,\mathrm{d}x\\
&= \frac{2}{\pi}\int_0^{\pi}(3\sin x - \sin 3x)\cos kx\,\mathrm{d}x\\
&= \begin{cases}
\frac{2}{\pi}\big[\sin^4 x\big]_0^{\pi} & \text{if } k=1\ ,\\[1ex]
\frac{6}{\pi}\left[\frac{\cos 2x}{4} - \frac{\cos 4x}{8}\right]_0^{\pi} - \frac{1}{3\pi}\big[\sin^2 3x\big]_0^{\pi} & \text{if } k=3\,,\\[1ex]
\frac{2}{\pi}\Big[-\frac{3}{2}\Big(\frac{\cos(1+k)x}{1+k} + \frac{\cos(1-k)x}{1-k}\Big) + & \\
\qquad\qquad \frac{1}{2}\Big(\frac{\cos(3-k)x}{3-k} + \frac{\cos(3+k)x}{3+k}\Big)\Big]_0^{\pi} & \text{if } k \neq 1,\ 3\,,
\end{cases}\\
&= \begin{cases}
0 & \text{if } k=1,\ k=3,\\[1ex]
\frac{48}{\pi(1-k^2)(9-k^2)}\big((-1)^k+1\big) & \text{if } k \neq 1,3\,,
\end{cases}\\
&= \begin{cases}
0 & \text{if } k \text{ odd},\\[1ex]
\frac{96}{\pi(1-k^2)(9-k^2)} & \text{if } k \text{ even}\,.
\end{cases}
\end{aligned}$$

Overall,

$$f \approx \frac{16}{3\pi} + \frac{96}{\pi}\sum_{k=1}^{\infty}\frac{1}{(1-4k^2)(9-4k^2)}\sin 2kx\,.$$

d) Calculating the coefficients a_k, b_k gives:

$$
\begin{aligned}
a_0 &= \frac{1}{2\pi}\int_0^{\pi}(-\cos x)\,\mathrm{d}x = 0\,,\\
a_k &= -\frac{1}{\pi}\int_0^{\pi}\cos x\cos kx\,\mathrm{d}x\\
&= \begin{cases} -\dfrac{1}{\pi}\left[\dfrac{x}{2}+\dfrac{\sin 2x}{4}\right]_0^{\pi} & \text{for } k=1\,,\\ -\dfrac{1}{\pi}\left[\dfrac{\sin(1-k)x}{2(1-k)}+\dfrac{\sin(1+k)x}{2(1+k)}\right]_0^{\pi} & \text{for } k\neq 1\,,\end{cases}\\
&= \begin{cases} -\dfrac{1}{2} & \text{for } k=1\,,\\ 0 & \text{for } k\neq 1\,;\end{cases}\\
b_k &= -\frac{1}{\pi}\int_0^{\pi}\cos x\sin kx\,\mathrm{d}x\\
&= \begin{cases} -\dfrac{1}{\pi}\left[\dfrac{\sin^2 x}{2}\right]_0^{\pi} & \text{for } k=1\,,\\ \dfrac{1}{\pi}\left[\dfrac{\cos(k-1)x}{2(k-1)}+\dfrac{\cos(k+1)x}{2(k+1)}\right]_0^{\pi} & \text{for } k\neq 1\,,\end{cases}\\
&= \begin{cases} -\dfrac{2k}{\pi(k^2-1)} & \text{for } k \text{ even}\,,\\ 0 & \text{for } k \text{ odd}\,.\end{cases}
\end{aligned}
$$

Therefore

$$f \approx -\frac{1}{2}\cos x - \frac{4}{\pi}\sum_{k=1}^{\infty}\frac{k}{4k^2-1}\sin 2kx\,.$$

5. *Fourier series' expansions:*

a) $f \approx \dfrac{1}{2} + \dfrac{2}{\pi}\displaystyle\sum_{k=1}^{\infty}\frac{(-1)^{k-1}}{2k-1}\cos 2\pi(2k-1)x\,.$

b) $f \approx \dfrac{2}{\pi} - \dfrac{4}{\pi}\displaystyle\sum_{k=1}^{\infty}\frac{1}{4k^2-1}\cos 4\pi kx\,.$

c) $f \approx \dfrac{1}{\pi} + \dfrac{1}{2}\sin 2\pi x - \dfrac{2}{\pi}\displaystyle\sum_{k=1}^{\infty}\frac{1}{4k^2-1}\cos 4\pi kx\,.$

d) $f \approx \dfrac{1}{4} - \dfrac{2}{\pi^2}\displaystyle\sum_{k=1}^{\infty}\frac{1}{(2k-1)^2}\cos 2\pi(2k-1)x\,.$

6. Since

$$f \approx 1 + \frac{1}{27}\sin 2x + \frac{2}{125}\sin 4x + \cdots,$$

we have $T = \pi$, which is the minimum period of the simple harmonic $\sin 2x$.

The function is not symmetric, as sum of the even map $g(x) = 1$ and the odd one $h(x) = \sum_{k=1}^{\infty} \frac{k}{(2k+1)^3}\sin 2kx$.

7. *Sum of series:*

a) We determine the Fourier coefficients for the 2π-periodic map defined as $f(x) = x^2$ on $[-\pi,\pi]$. Being even, it has $b_k = 0$ for all $k \geq 1$. Moreover,

$$\begin{aligned} a_0 &= \frac{1}{\pi}\int_0^{\pi} x^2\,\mathrm{d}x = \frac{\pi^2}{3}, \\ a_k &= \frac{2}{\pi}\int_0^{\pi} x^2\cos kx\,\mathrm{d}x = \frac{2}{\pi}\left[\frac{2x}{k^2}\cos kx + \left(\frac{x^2}{k} - \frac{2}{k^3}\right)\sin kx\right]_0^{\pi} \\ &= \frac{4}{k^2}(-1)^k, \qquad k \geq 1. \end{aligned}$$

Therefore

$$f \approx \frac{\pi^2}{3} + 4\sum_{k=1}^{\infty}\frac{(-1)^k}{k^2}\cos kx;$$

Parseval's identity yields

$$\int_{-\pi}^{\pi} x^4\,\mathrm{d}x = \frac{2}{9}\pi^5 + 16\pi\sum_{k=1}^{\infty}\frac{1}{k^4},$$

so

$$\sum_{k=1}^{\infty}\frac{1}{k^4} = \frac{\pi^4}{90}.$$

b) $\displaystyle\sum_{k=1}^{\infty}\frac{1}{k^6} = \frac{\pi^6}{945}.$

8. Observe

$$f(x) = \begin{cases} x^2 - 1 & \text{if } x \in [-\pi,-1] \cup [1,\pi], \\ 0 & \text{if } x \in (-1,1) \end{cases}$$

(Fig. 3.8). The map is even so $b_k = 0$ for all $k \geq 1$. Moreover

$$\begin{aligned} a_0 &= \frac{1}{\pi}\int_1^{\pi}(x^2-1)\,\mathrm{d}x = \frac{\pi^2}{3} - 1 + \frac{2}{3\pi}, \\ a_k &= \frac{2}{\pi}\int_1^{\pi}(x^2-1)\cos kx\,\mathrm{d}x \end{aligned}$$

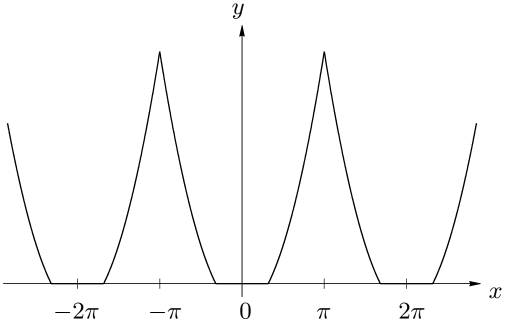


Figure 3.8. The graph of f relative to Exercise 8

$$= \frac{2}{\pi}\left[\frac{1}{k^3}\big(2kx\cos kx + (k^2x^2-2)\sin kx\big) - \frac{\sin kx}{k}\right]_1^{\pi}$$

$$= \frac{4}{\pi k^2}\Big((-1)^k\pi + \frac{\sin k}{k} - \cos k\Big)\,, \qquad k \geq 1\,;$$

then

$$f \approx \frac{\pi^2}{3} - 1 + \frac{2}{3\pi} + \frac{4}{\pi}\sum_{k=1}^{\infty}\frac{1}{k^2}\Big((-1)^k\pi + \frac{\sin k}{k} - \cos k\Big)\cos kx\,.$$

9. For convenience let us draw the graph of f (Fig. 3.9).
As f is even, the coefficients b_k for $k \geq 1$ all vanish. The other ones are:

$$a_0 = \frac{1}{\pi}\int_0^{\pi/2}(\pi - x)\,\mathrm{d}x = \frac{3}{8}\pi\,,$$

$$a_k = \frac{2}{\pi}\int_0^{\pi/2}(\pi - x)\cos kx\,\mathrm{d}x = 2\left[\frac{\sin kx}{x} - \frac{1}{\pi k^2}\big(\cos kx + kx\sin kx\big)\right]_0^{\pi/2}$$

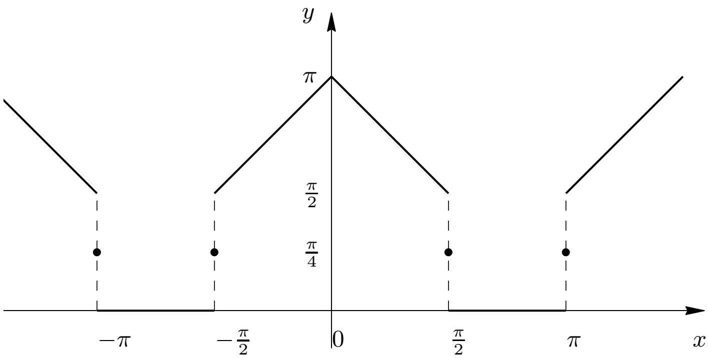


Figure 3.9. The graph of f relative to Exercise 9

$$= \frac{1}{k}\sin\frac{\pi}{2}k - \frac{2}{\pi k^2}\cos\frac{\pi}{2}k + \frac{2}{\pi k^2}\,, \qquad k \geq 1\,.$$

Hence

$$f \approx \frac{3}{8}\pi + \sum_{k=1}^{\infty}\left(\frac{1}{k}\sin\frac{\pi}{2}k - \frac{2}{\pi k^2}\cos\frac{\pi}{2}k + \frac{2}{\pi k^2}\right)\cos kx\,.$$

The series converges pointwise to f, regularised; in particular, for $x = \frac{\pi}{2}$

$$\frac{\pi}{4} = \frac{3}{8}\pi + \sum_{k=1}^{\infty}\left(\frac{1}{k}\sin\frac{\pi}{2}k - \frac{2}{\pi k^2}\cos\frac{\pi}{2}k + \frac{2}{\pi k^2}\right)\cos\frac{\pi}{2}k\,.$$

Now, as

$$\sin\frac{\pi}{2}k = \begin{cases} 0 & \text{if } k = 2m\,, \\ (-1)^m & \text{if } k = 2m+1\,, \end{cases} \qquad \cos\frac{\pi}{2}k = \begin{cases} (-1)^m & \text{if } k = 2m\,, \\ 0 & \text{if } k = 2m+1\,, \end{cases}$$

we have

$$-\frac{\pi}{8} = \sum_{m=1}^{\infty}\frac{1}{2\pi m^2}\big((-1)^m - 1\big) - \sum_{k=0}^{\infty}\frac{1}{\pi(2k+1)^2}\,,$$

whence

$$\sum_{k=0}^{\infty}\frac{1}{(2k+1)^2} = \frac{\pi^2}{8}\,.$$

10. We have

$$a_k = 0\,,\ \forall k \geq 0;$$

$$b_2 = b_4 = 1\,,\ b_{2m} = 0\,,\ \forall m \geq 3\,; \qquad b_{2m+1} = \frac{4}{\pi(2m+1)}\,,\ \forall m \geq 0\,.$$

11. The graph is shown in Fig. 3.10. Let us begin by finding the Fourier coefficients of f. The map is even, implying $b_k = 0$ for all $k \geq 1$. What is more,

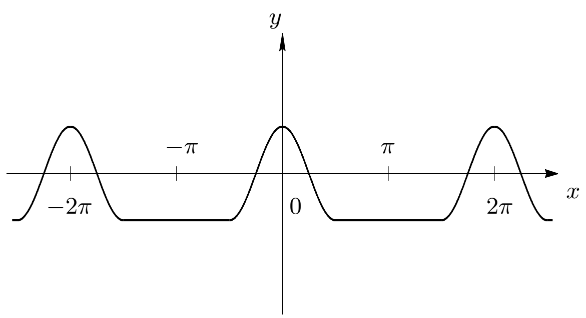


Figure 3.10. The graph of f relative to Exercise 11

$$a_0 = \frac{1}{\pi}\int_0^\pi f(x)\,\mathrm{d}x \frac{1}{\pi}\left(\int_0^{\pi/2}\cos 2x\,\mathrm{d}x - \int_{\pi/2}^{\pi}\mathrm{d}x\right) = -\frac{1}{2}\,,$$

$$a_k = \frac{2}{\pi}\int_0^\pi f(x)\cos kx\,\mathrm{d}x \frac{2}{\pi}\left(\int_0^{\pi/2}\cos 2x\cos kx\,\mathrm{d}x - \int_{\pi/2}^{\pi}\cos kx\,\mathrm{d}x\right)$$

$$= \begin{cases} \dfrac{2}{\pi}\left[\left[\dfrac{x}{2}+\dfrac{\sin 4x}{8}\right]_0^{\pi/2} - \left[\dfrac{1}{2}\sin 2x\right]_{\pi/2}^{\pi}\right] & \text{for } k=2\,, \\ \dfrac{2}{\pi}\left[\left[\dfrac{\sin(2-k)x}{2(2-k)}+\dfrac{\sin(2+k)x}{2(2+k)}\right]_0^{\pi/2} - \left[\dfrac{1}{k}\sin kx\right]_{\pi/2}^{\pi}\right] & \text{for } k\neq 2\,, \end{cases}$$

$$= \begin{cases} \dfrac{1}{2} & \text{for } k=2\,, \\ 0 & \text{for } k=2m\,,\ m>1\,, \\ \dfrac{8(-1)^m}{\pi(2m+1)\big(4-(2m+1)^2\big)} & \text{for } k=2m+1\,,\ m\geq 0\,. \end{cases}$$

Hence

$$f \approx -\frac{1}{2}+\frac{1}{2}\cos 2x + \frac{8}{\pi}\sum_{k=0}^{\infty}\frac{(-1)^k}{(2k+1)(3-4k-4k^2)}\cos(2k+1)x\,.$$

Note the series converges uniformly on $\mathbb{R}$ by Weierstrass' M-test: for $k \geq 0$,

$$\left|\frac{(-1)^k}{(2k+1)(3-4k-4k^2)}\cos(2k+1)x\right| \leq \frac{1}{(2k+1)(4k^2+4k-3)} = M_k\,,$$

and $\sum_{k=0}^{\infty} M_k$ converges (like the generalised harmonic series of exponent 3, because $M_k \sim \frac{1}{8k^3}$, as $k\to\infty$). In particular, the Fourier series converges for any $x\in\mathbb{R}$ pointwise. Alternatively, one could invoke Theorem 3.20.

Instead of computing directly the Fourier coefficients of f' with the definition, we shall check if the convergence is uniform for the derivatives' Fourier series; thus, we will be able to use Theorem 2.19. Actually one sees rather immediately f is $\mathcal{C}^1(\mathbb{R})$ with

$$f'(x) = \begin{cases} -2\sin 2x & \text{if } |x|<\frac{\pi}{2}\,, \\ 0 & \text{if } \frac{\pi}{2}\leq |x| \leq \pi\,, \end{cases}$$

while f' is piecewise $\mathcal{C}^1$ on $\mathbb{R}$ (f'' has a jump discontinuity at $\pm\frac{\pi}{2}$). Therefore the Fourier series of f' converges uniformly (hence, pointwise) on $\mathbb{R}$, and so

$$f'(x) = -\sin 2x + \frac{8}{\pi}\sum_{k=0}^{\infty}\frac{(-1)^k}{4k^2+4k-3}\sin(2k+1)x\,.$$

12. We have

$$
\begin{aligned}
a_0 &= \frac{1}{2\pi}\int_0^{2\pi} x^2\,\mathrm{d}x = \frac{4}{3}\pi^2\,,\\
a_k &= \frac{1}{\pi}\int_0^{2\pi} x^2\cos kx\,\mathrm{d}x\\
&= \frac{1}{\pi}\left[\frac{1}{k}x^2\sin kx + \frac{2}{k^2}x\cos kx - \frac{2}{k^3}\sin kx\right]_0^{2\pi} = \frac{4}{k^2}\,,\qquad k\ge 1\,,\\
b_k &= \frac{1}{\pi}\int_0^{2\pi} x^2\sin kx\,\mathrm{d}x\\
&= \frac{1}{\pi}\left[-\frac{1}{k}x^2\cos kx + \frac{2}{k^2}x\sin kx + \frac{2}{k^3}\cos kx\right]_0^{2\pi} = -\frac{4\pi}{k}\,,\qquad k\ge 1\,;
\end{aligned}
$$

so the Fourier series of f is the given one.

The function f is continuous and piecewise monotone on $\mathbb{R}$, so its Fourier series converges to the regularised f pointwise ($\tilde f(2k\pi) = 2\pi^2$, $\forall k \in \mathbb{Z}$). Furthermore, the series converges to f uniformly on all closed sub-intervals not containing the points $2k\pi$, $\forall k \in \mathbb{Z}$.

In particular,

$$
f(\pi) = \pi^2 = \frac{4}{3}\pi^2 + 4\sum_{k=1}^{\infty}\frac{1}{k^2}\cos k\pi = \frac{4}{3}\pi^2 + 4\sum_{k=1}^{\infty}\frac{(-1)^k}{k^2}
$$

whence

$$
\sum_{k=1}^{\infty}\frac{(-1)^k}{k^2} = \frac{1}{4}\left(\pi^2 - \frac{4}{3}\pi^2\right) = -\frac{\pi^2}{12}\,.
$$

13. The series $\displaystyle\sum_{k=1}^{\infty}\frac{1}{2^k}\sin kx$ converges to $\mathbb{R}$ uniformly because Weierstrass' M-test applies with $M_k = \frac{1}{2^k}$; this is due to

$$
|f_k(x)| = \left|\frac{1}{2^k}\sin kx\right| \le \frac{1}{2^k}\,,\qquad \forall x\in\mathbb{R}\,.
$$

Analogous results hold for the series of derivatives:

$$
\begin{aligned}
|f_k'(x)| &= \left|\frac{k}{2^k}\cos kx\right| \le \frac{k}{2^k}\,,\qquad \forall x\in\mathbb{R}\,,\\
|f_k''(x)| &= \left|\frac{k^2}{2^k}\sin kx\right| \le \frac{k^2}{2^k}\,,\qquad \forall x\in\mathbb{R}\,,\\
&\ \ \vdots\\
|f_k^{(n)}(x)| &\le \frac{k^n}{2^k}\,,\qquad \forall x\in\mathbb{R}\,.
\end{aligned}
$$

Consequently, for all $n \geq 0$ the series $\displaystyle\sum_{k=0}^{\infty} f_k^{(n)}(x)$ converges on $\mathbb{R}$ uniformly; by Theorem 2.19 the map f is differentiable infinitely many times: $f \in \mathcal{C}^{\infty}(\mathbb{R})$. In particular, the Fourier series of f' is

$$f'(x) = \sum_{k=1}^{\infty} \frac{k}{2^k} \cos kx\,, \qquad \forall x \in \mathbb{R}\,.$$

To compute $\|f\|_2$ we use Parseval's formula

$$\begin{aligned}\|f\|_2^2 &= \int_0^{2\pi} |f(x)|^2 \,\mathrm{d}x = 2\pi a_0 + \pi \sum_{k=1}^{\infty} (a_k^2 + b_k^2) = 4\pi + \pi \sum_{k=1}^{\infty} \frac{1}{4^k} \\ &= 4\pi + \pi \left(\frac{1}{1-\frac{1}{4}} - 1 \right) = 4\pi + \frac{\pi}{3} = \frac{13}{3}\pi\,,\end{aligned}$$

from which $\|f\|_2 = 5\sqrt{\frac{13}{3}\pi}$. At last,

$$\int_0^{2\pi} f(x)\,\mathrm{d}x = 2\pi a_0 = 4\pi\,.$$

14. We have

$$f \approx 2 \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} \sin kx\,.$$

This converges quadratically; it converges pointwise to the regularised map coinciding with f for $x \neq \pi + 2k\pi$ and equal to 0 for $x = \pi + 2k\pi$, $k \in \mathbb{Z}$; it converges uniformly on every closed interval not containing $\pi + 2k\pi$, $k \in \mathbb{Z}$.

4

Functions between Euclidean spaces

This chapter sees the dawn of the study of multivariable and vector-valued functions, that is, maps between the Euclidean spaces $\mathbb{R}^n$ and $\mathbb{R}^m$ or subsets thereof, with one of n and m bigger than 1. Subsequent chapters treat the relative differential and integral calculus and constitute a large part of the course.

To warm up we briefly recall the main notions related to vectors and matrices, which students should already be familiar with. Then we review the indispensable topological foundations of Euclidean spaces, especially neighbourhood systems of a point, open and closed sets, and the boundary of a set. We discuss the properties of subsets of $\mathbb{R}^n$, which naturally generalise those of real intervals, highlighting the features of this richer, higher-dimensional landscape.

We then deal with the continuity features of functions and their limits; despite these extend the ones seen in dimension one, they require particular care, because of the subtleties and snags specific to the multivariable setting.

At last, we start exploring a remarkable class of functions describing one- and two-dimensional geometrical objects – present in our everyday life – called curves and surfaces. The careful study of curves and surfaces will continue in the second part of Chapter 6, at which point the differential calculus apparatus will be available. The aspects connected to integral calculus will be postponed to Chapter 9.

4.1 Vectors in $\mathbb{R}^n$

Recall $\mathbb{R}^n$ is the vector space of ordered n-tuples $\boldsymbol{x} = (x_i)_{i=1,\dots,n}$, called **vectors**. The components x_i of $\boldsymbol{x}$ may be written either horizontally, to give a row vector

$$\boldsymbol{x} = (x_1, \dots, x_n)\,,$$

or vertically, producing a column vector

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = (x_1, \dots, x_n)^T\,.$$

C. Canuto, A. Tabacco: *Mathematical Analysis II,* 2nd Ed.,
UNITEXT – La Matematica per il 3+2 85, DOI 10.1007/978-3-319-12757-6_4,

These two expressions will be equivalent in practice, except for some cases where writing vertically or horizontally will make a difference. For typesetting reasons the horizontal notation is preferable.

Any vector of $\mathbb{R}^n$ can be represented using the **canonical basis** $\{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$ whose vectors have components all zero except for one that equals 1

$$\boldsymbol{e}_i = (\delta_{ij})_{1\le j\le n} \qquad \text{where } \delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases} . \tag{4.1}$$

Then

$$\boldsymbol{x} = \sum_{i=1}^{n} x_i \boldsymbol{e}_i \,. \tag{4.2}$$

The vectors of the canonical basis are usually denoted by $\boldsymbol{i}$, $\boldsymbol{j}$ in $\mathbb{R}^2$ and $\boldsymbol{i}$, $\boldsymbol{j}$, $\boldsymbol{k}$ in $\mathbb{R}^3$. It can be useful to identify a vector $(x_1, x_2) = x_1\boldsymbol{i} + x_2\boldsymbol{j} \in \mathbb{R}^2$ with the vector $(x_1, x_2, 0) = x_1\boldsymbol{i} + x_2\boldsymbol{j} + 0\,\boldsymbol{k} \in \mathbb{R}^3$: the expression $x_1\boldsymbol{i} + x_2\boldsymbol{j}$ will thus indicate a vector of $\mathbb{R}^2$ or $\mathbb{R}^3$, according to the context.

In $\mathbb{R}^n$ the **dot product** of two vectors is defined as

$$\boldsymbol{x} \cdot \boldsymbol{y} = x_1 y_1 + \ldots + x_n y_n = \sum_{i=1}^{n} x_i y_i \,.$$

This in turn defines the **Euclidean norm**

$$\|\boldsymbol{x}\| = \sqrt{\boldsymbol{x} \cdot \boldsymbol{x}} = \sqrt{\sum_{i=1}^{n} x_i^2} \,,$$

for which the **Cauchy-Schwarz inequality**

$$|\boldsymbol{x} \cdot \boldsymbol{y}| \le \|\boldsymbol{x}\| \, \|\boldsymbol{y}\| \tag{4.3}$$

and the **triangle inequality**

$$\|\boldsymbol{x} + \boldsymbol{y}\| \le \|\boldsymbol{x}\| + \|\boldsymbol{y}\| \tag{4.4}$$

hold. Two vectors $\boldsymbol{x}$ and $\boldsymbol{y}$ satisfying $\boldsymbol{x} \cdot \boldsymbol{y} = 0$ are called **orthogonal**, and a vector $\boldsymbol{x}$ such that $\|\boldsymbol{x}\| = 1$ is said a **unit vector**, or **of length** 1. The canonical basis of $\mathbb{R}^n$ is an example of an **orthonormal system** of vectors, i.e., a set of n normalised and pairwise orthogonal vectors; its elements satisfy in fact

$$\boldsymbol{e}_i \cdot \boldsymbol{e}_j = \delta_{ij} \,, \qquad 1 \le i, j \le n \,.$$

From (4.2) we have

$$x_i = \boldsymbol{x} \cdot \boldsymbol{e}_i$$

for the ith component of a vector $\boldsymbol{x}$.

It makes sense to associate $\boldsymbol{x} \in \mathbb{R}^n$ to the unique point P in Euclidean n-space whose coordinates in an orthogonal Cartesian frame are the components of

$\boldsymbol{x}$; this fact extends what we already know for the plane and space. Under this identification $\|\boldsymbol{x}\|$ is the Euclidean distance between the point P, of coordinates $\boldsymbol{x}$, and the origin O. The quantity $\|\boldsymbol{x}-\boldsymbol{y}\| = \sqrt{(x_1-y_1)^2+\ldots+(x_n-y_n)^2}$ is the distance between the points P and Q of respective coordinates $\boldsymbol{x}$ and $\boldsymbol{y}$.

In $\mathbb{R}^3$ the **cross** or **wedge product** $\boldsymbol{x}\wedge\boldsymbol{y}$ of two vectors $\boldsymbol{x} = x_1\boldsymbol{i}+x_2\boldsymbol{j}+x_3\boldsymbol{k}$ and $\boldsymbol{y} = y_1\boldsymbol{i}+y_2\boldsymbol{j}+y_3\boldsymbol{k}$ is the vector of $\mathbb{R}^3$ defined by

$$\boldsymbol{x}\wedge\boldsymbol{y} = (x_2y_3 - x_3y_2)\boldsymbol{i} + (x_3y_1 - x_1y_3)\boldsymbol{j} + (x_1y_2 - x_2y_1)\boldsymbol{k}\,. \tag{4.5}$$

It can also be computed by the formula

$$\boldsymbol{x}\wedge\boldsymbol{y} = \det\begin{pmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{pmatrix} \tag{4.6}$$

by expanding the determinant formally along the first row (see definition (4.11)). The cross product of two vectors is orthogonal to both (Fig. 4.1, left):

$$(\boldsymbol{x}\wedge\boldsymbol{y})\cdot\boldsymbol{x} = 0\,, \qquad (\boldsymbol{x}\wedge\boldsymbol{y})\cdot\boldsymbol{y} = 0\,. \tag{4.7}$$

The number $\|\boldsymbol{x}\wedge\boldsymbol{y}\|$ is the area of the parallelogram with the vectors $\boldsymbol{x}$, $\boldsymbol{y}$ as sides, while $\|(\boldsymbol{x}\wedge\boldsymbol{y})\cdot\boldsymbol{z}\|$ represents the volume of the prism of sides $\boldsymbol{x}$, $\boldsymbol{y}$, $\boldsymbol{z}$ (Fig. 4.1).

We have some properties:

$$\begin{aligned} &\boldsymbol{y}\wedge\boldsymbol{x} = -(\boldsymbol{x}\wedge\boldsymbol{y})\,, \\ &\boldsymbol{x}\wedge\boldsymbol{y} = \boldsymbol{0} \quad\Longleftrightarrow\quad \boldsymbol{x} = \lambda\boldsymbol{y} \text{ for some } \lambda\in\mathbb{R}\,, \\ &(\boldsymbol{x}+\boldsymbol{y})\wedge\boldsymbol{z} = \boldsymbol{x}\wedge\boldsymbol{z} + \boldsymbol{y}\wedge\boldsymbol{z}\,, \end{aligned} \tag{4.8}$$

the first of which implies $\boldsymbol{x}\wedge\boldsymbol{x} = \boldsymbol{0}$.

Furthermore,

$$\boldsymbol{i}\wedge\boldsymbol{j} = \boldsymbol{k}\,, \qquad \boldsymbol{j}\wedge\boldsymbol{k} = \boldsymbol{i}\,, \qquad \boldsymbol{k}\wedge\boldsymbol{i} = \boldsymbol{j}\,;$$

these and previous ones prove for instance that if $\boldsymbol{x} = x_1\boldsymbol{i}+x_2\boldsymbol{j}$, $\boldsymbol{y} = y_1\boldsymbol{i}+y_2\boldsymbol{j}$ then $\boldsymbol{x}\wedge\boldsymbol{y} = (x_1y_2 - x_2y_1)\boldsymbol{k}$.

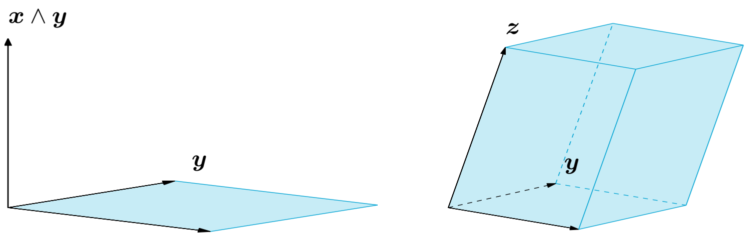


Figure 4.1. Wedge products $\boldsymbol{x}\wedge\boldsymbol{y}$ (left) and $(\boldsymbol{x}\wedge\boldsymbol{y})\cdot\boldsymbol{z}$ (right)

A triple $(\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3)$ of pairwise non-aligned, unit vectors is said **positively oriented** (or **right-handed**) if

$$(\boldsymbol{v}_1 \wedge \boldsymbol{v}_2) \cdot \boldsymbol{v}_3 > 0\,,$$

i.e., if $\boldsymbol{v}_3$ and $\boldsymbol{v}_1 \wedge \boldsymbol{v}_2$ lie on the same side of the plane spanned by $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$. A practical way to decide whether a triple is positively oriented is to use the *right-hand rule*: the pointing finger is $\boldsymbol{v}_1$, the long finger $\boldsymbol{v}_2$, the thumb $\boldsymbol{v}_3$. (The left hand works, too, as long as we take the long finger for $\boldsymbol{v}_1$, the pointing finger for $\boldsymbol{v}_2$ and the thumb for $\boldsymbol{v}_3$.)

The triple $(\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3)$ is **negatively oriented** (or **left-handed**) if $(\boldsymbol{v}_1, \boldsymbol{v}_2, -\boldsymbol{v}_3)$ is oriented positively

$$(\boldsymbol{v}_1 \wedge \boldsymbol{v}_2) \cdot \boldsymbol{v}_3 < 0\,.$$

4.2 Matrices

A real matrix $\boldsymbol{A}$ with m rows and n columns (an $m \times n$ matrix) is a collection of $m \times n$ real numbers arranged in a table

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & & & \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

or, more concisely,

$$\boldsymbol{A} = (a_{ij})_{\substack{1 \le i \le m \\ 1 \le j \le n}} \in \mathbb{R}^{mn}\,.$$

The vectors formed by the entries of one row (resp. column) of $\boldsymbol{A}$ are the row vectors (column vectors) of $\boldsymbol{A}$. Thus $m \times 1$ matrices are vectors of $\mathbb{R}^m$ written as column vectors, while $1 \times n$ matrices are vectors of $\mathbb{R}^n$ seen as row vectors. When $m = n$ the matrix is called **square of order** n. The set of $m \times n$ matrices is a vector space: usually indicated by $\mathbb{R}^{m,n}$, it is isomorphic to the Euclidean space $\mathbb{R}^{mn}$. The matrix $\boldsymbol{C} = \lambda\boldsymbol{A} + \mu\boldsymbol{B}$, λ, $\mu \in \mathbb{R}$, has entries

$$c_{ij} = \lambda a_{ij} + \mu b_{ij}\,, \qquad 1 \le i \le m\,,\ 1 \le j \le n\,.$$

If $\boldsymbol{A}$ is $m \times n$ and $\boldsymbol{B}$ is $n \times p$, the product $\boldsymbol{C} = \boldsymbol{AB}$ has by definition m rows and p columns; its generic entry is the dot product of a row vector of $\boldsymbol{A}$ with a column vector of $\boldsymbol{B}$:

$$c_{ij} = \sum_{k=1}^{n} a_{ik} b_{kj}\,, \qquad 1 \le i \le m\,,\ 1 \le j \le p\,.$$

In particular, if $\boldsymbol{B} = \boldsymbol{x}$ is $n \times 1$, hence a column vector in $\mathbb{R}^n$, the matrix product with the vector $\boldsymbol{Ax}$ is well defined: it is a column vector in $\mathbb{R}^m$. If $p = m$, we have square matrices $\boldsymbol{AB}$ of order m and $\boldsymbol{BA}$ of order n. If $n = m$, $\boldsymbol{AB} \neq \boldsymbol{BA}$ in general, because the product of matrices is not commutative.

The **first minor** $\boldsymbol{A}_{ij}$ $(1 \le i \le m,\ 1 \le j \le n)$ of an $m \times n$ matrix $\boldsymbol{A}$ is the $(m-1) \times (n-1)$ matrix obtained erasing from $\boldsymbol{A}$ the ith row and jth column.

The **rank** $r \le \min(m, n)$ of $\boldsymbol{A}$ is the maximum number of linearly independent rows thought of as vectors of $\mathbb{R}^n$ (or linearly independent columns, seen as vectors of $\mathbb{R}^m$).

Given an $m \times n$ matrix $\boldsymbol{A}$, its **transpose** is the $n \times m$ matrix $\boldsymbol{A}^T$ with entries

$$a_{ij}^T = a_{ji}\,, \qquad 1 \le i \le n\,,\ 1 \le j \le m\,;$$

otherwise put, $\boldsymbol{A}^T$ is obtained from $\boldsymbol{A}$ by orderly swapping rows and columns; clearly, $(\boldsymbol{A}^T)^T = \boldsymbol{A}$. Whenever defined, the product of matrices satisfies $(\boldsymbol{AB})^T = \boldsymbol{B}^T\boldsymbol{A}^T$. The dot product of vectors in $\mathbb{R}^n$ is a special case of matrix product $\boldsymbol{x} \cdot \boldsymbol{y} = \boldsymbol{x}^T\boldsymbol{y} = \boldsymbol{y}^T\boldsymbol{x}$, provided we think $\boldsymbol{x}$ and $\boldsymbol{y}$ as column vectors.

In $\mathbb{R}^{m,n}$ there is a **norm**, associated to the Euclidean norm,

$$\|\boldsymbol{A}\| = \max\{\|\boldsymbol{Ax}\| : \boldsymbol{x} \in \mathbb{R}^n,\ \|\boldsymbol{x}\| = 1\}\,, \tag{4.9}$$

that satisfies the inequalities

$$\|\boldsymbol{Ax}\| \le \|\boldsymbol{A}\|\,\|\boldsymbol{x}\| \qquad \text{and} \qquad \|\boldsymbol{AB}\| \le \|\boldsymbol{A}\|\,\|\boldsymbol{B}\| \tag{4.10}$$

for any $\boldsymbol{x} \in \mathbb{R}^n$ and $\boldsymbol{B} \in \mathbb{R}^{n,p}$.

Square matrices

From now on we will consider square matrices of order n. Among them a particularly important one is the **identity** matrix $\boldsymbol{I} = (\delta_{ij})_{1 \le i,j \le n}$, which satisfies $\boldsymbol{AI} = \boldsymbol{IA} = \boldsymbol{A}$ for any square matrix $\boldsymbol{A}$ of order n. A matrix $\boldsymbol{A}$ is called **symmetric** if it coincides with its transpose

$$a_{ij} = a_{ji}\,, \qquad 1 \le i, j \le n\,;$$

it is **normal** if $\boldsymbol{AA}^T = \boldsymbol{A}^T\boldsymbol{A}$, and in particular **orthogonal** if $\boldsymbol{AA}^T = \boldsymbol{A}^T\boldsymbol{A} = \boldsymbol{I}$.

To each square matrix $\boldsymbol{A}$ one can associate a number $\det \boldsymbol{A}$, the **determinant** of $\boldsymbol{A}$, which may be computed recursively using Laplace's rule: $\det \boldsymbol{A} = a$ if $n = 1$ and $\boldsymbol{A} = (a)$, whereas for $n > 1$

$$\det \boldsymbol{A} = \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det \boldsymbol{A}_{ij}\,, \tag{4.11}$$

where $i \in \{1, \ldots, n\}$ is arbitrary, but fixed. For instance

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc\,.$$

The following properties hold:

$$\det(\boldsymbol{AB}) = \det \boldsymbol{A} \, \det \boldsymbol{B} \,, \qquad \det(\lambda \boldsymbol{A}) = \lambda^n \det \boldsymbol{A} \,,$$
$$\det \boldsymbol{A}^T = \det \boldsymbol{A} \,, \qquad \det \boldsymbol{A}' = -\det \boldsymbol{A}$$

if $\boldsymbol{A}'$ is $\boldsymbol{A}$ with two rows (or columns) exchanged. Immediate consequences are that $\det \boldsymbol{I} = 1$ and $|\det \boldsymbol{A}| = 1$ for $\boldsymbol{A}$ orthogonal.

The matrix $\boldsymbol{A}$ is said **non-singular** if $\det \boldsymbol{A} \neq 0$. This is equivalent to the *invertibility* of $\boldsymbol{A}$, i.e., the existence of a matrix $\boldsymbol{A}^{-1}$ of order n, called the **inverse** of $\boldsymbol{A}$, such that
$$\boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{A}^{-1}\boldsymbol{A} = \boldsymbol{I} \,.$$

If $\boldsymbol{A}$ and $\boldsymbol{B}$ are invertible,

$$\det(\boldsymbol{A}^{-1}) = (\det \boldsymbol{A})^{-1} \,, \qquad (\boldsymbol{AB})^{-1} = \boldsymbol{B}^{-1}\boldsymbol{A}^{-1} \,, \qquad (\boldsymbol{A}^T)^{-1} = (\boldsymbol{A}^{-1})^T \,;$$

from the last equation it follows, as special case, that the inverse of a symmetric matrix is still symmetric. Every orthogonal matrix $\boldsymbol{A}$ is invertible, for $\boldsymbol{A}^{-1} = \boldsymbol{A}^T$.

There are several equivalent conditions to invertibility, among which:

- the rank of $\boldsymbol{A}$ equals n;
- the linear system $\boldsymbol{Ax} = \boldsymbol{b}$ has a solution $\boldsymbol{x} \in \mathbb{R}^n$ for any $\boldsymbol{b} \in \mathbb{R}^n$;
- the homogeneous linear system $\boldsymbol{Ax} = \boldsymbol{0}$ has only the trivial solution $\boldsymbol{x} = \boldsymbol{0}$.

The last one amounts to saying that 0 is no eigenvalue of $\boldsymbol{A}$.

Eigenvalues and eigenvectors

The **eigenvalues** of $\boldsymbol{A}$ are the zeroes (in $\mathbb{C}$) of the **characteristic polynomial** of degree n
$$\chi(\lambda) = \det(\boldsymbol{A} - \lambda \boldsymbol{I}) \,.$$

In other words, the eigenvalues are complex numbers λ for which there exists a vector $\boldsymbol{v} \neq \boldsymbol{0}$, called (right) **eigenvector of $\boldsymbol{A}$ associated to** λ, such that

$$\boldsymbol{Av} = \lambda \boldsymbol{v} \tag{4.12}$$

(here and henceforth all vectors should be thought of as column vectors). The Fundamental Theorem of Algebra predicts the existence of p distinct eigenvalues $\lambda^{(1)}, \ldots, \lambda^{(p)}$ with $1 \leq p \leq n$, each having *algebraic multiplicity* (as root of the polynomial χ) $\mu^{(i)}$, such that $\mu^{(1)} + \ldots + \mu^{(p)} = n$. The maximum number of linearly independent eigenvectors associated to $\lambda^{(i)}$ is the *geometric multiplicity* of $\lambda^{(i)}$; we denote it by $m^{(i)}$, and observe $m^{(i)} \leq \mu^{(i)}$. Eigenvectors associated to distinct eigenvalues are linearly independent. Therefore when the algebraic and geometric multiplicities of each eigenvalue coincide, there are in total n linearly independent eigenvectors, and thus a basis made of eigenvectors. In such a case the matrix is **diagonalisable**.

The name means that the matrix can be made diagonal. To see this we number the eigenvalues by λ_k, $1 \leq k \leq n$ for convenience, repeating each one as many times as its multiplicity; let $\boldsymbol{v}_k$ be an eigenvector associated to λ_k, chosen so that the set $\{\boldsymbol{v}_k\}_{1\le k\le n}$ is made of linearly independent vectors. The equations

$$\boldsymbol{A}\boldsymbol{v}_k = \lambda_k \boldsymbol{v}_k\,, \qquad 1 \le k \le n\,,$$

can be rewritten in matrix form

$$\boldsymbol{A}\boldsymbol{P} = \boldsymbol{P}\boldsymbol{\Lambda}\,, \tag{4.13}$$

where $\boldsymbol{\Lambda} = \mathbf{diag}\,(\lambda_1, \ldots, \lambda_n)$ is the diagonal matrix with the eigenvalues as entries, and $\boldsymbol{P} = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ is the square matrix of order n whose columns are the eigenvectors of $\boldsymbol{A}$. The linear independence of the eigenvectors is equivalent to the invertibility of $\boldsymbol{P}$, so that (4.13) becomes

$$\boldsymbol{A} = \boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^{-1}\,, \tag{4.14}$$

and consequently $\boldsymbol{A}$ is **similar** to a diagonal matrix.

Returning to the general setting, it is relevant to notice that as $\boldsymbol{A}$ is a real matrix (making the characteristic polynomial a real polynomial), its eigenvalues are either real or complex conjugate; the same is true for the corresponding eigenvectors. The determinant of $\boldsymbol{A}$ coincides with the product of the eigenvalues

$$\det \boldsymbol{A} = \lambda_1 \lambda_2 \cdots \lambda_{n-1}\lambda_n\,.$$

The eigenvalues of $\boldsymbol{A}^2 = \boldsymbol{A}\boldsymbol{A}$ are the squares of those of $\boldsymbol{A}$, with the same eigenvectors, and the analogue fact will hold for the generic power of $\boldsymbol{A}$. At the same time, if $\boldsymbol{A}$ is invertible, the eigenvalues of the inverse matrix are the *inverses* of the eigenvalues of $\boldsymbol{A}$, while the eigenvectors stay the same; by assumption in fact, (4.12) is equivalent to

$$\boldsymbol{v} = \lambda \boldsymbol{A}^{-1}\boldsymbol{v}\,, \qquad \text{so} \qquad \boldsymbol{A}^{-1}\boldsymbol{v} = \frac{1}{\lambda}\boldsymbol{v}\,.$$

The **spectral radius** of $\boldsymbol{A}$ is by definition the maximum modulus of the eigenvalues

$$\rho(\boldsymbol{A}) = \max_{1\le k\le n} |\lambda_k|\,.$$

The (Euclidean) norm of $\boldsymbol{A}$ satisfies

$$\|\boldsymbol{A}\| = \sqrt{\rho(\boldsymbol{A}^T\boldsymbol{A})}\,;$$

a special case is that of symmetric matrices $\boldsymbol{A}$, for which $\|\boldsymbol{A}\| = \rho(\boldsymbol{A})$; if $\boldsymbol{A}$ is additionally orthogonal, then $\|\boldsymbol{A}\| = \sqrt{\rho(\boldsymbol{I})} = 1$.

Symmetric matrices

Symmetric matrices $\boldsymbol{A}$ have pivotal properties concerning eigenvalues and eigenvectors. Each eigenvalue is real and its algebraic and geometric multiplicities coincide, rendering $\boldsymbol{A}$ always diagonalisable. The eigenvectors, all real, may be chosen to form an orthonormal basis (in fact, eigenvectors relative to distinct eigenvalues are orthogonal, those with same eigenvalue can be made orthonormal); the matrix $\boldsymbol{P}$ associated to such a basis is thus orthogonal, so (4.14) reads

$$\boldsymbol{A} = \boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^T . \tag{4.15}$$

Hence the transformation $\boldsymbol{x} \mapsto \boldsymbol{P}^T\boldsymbol{x} = \boldsymbol{y}$ defines an orthogonal change of basis in $\mathbb{R}^n$, from the canonical basis $\{\boldsymbol{e}_i\}_{1\le i\le n}$ to the basis of eigenvectors $\{\boldsymbol{v}_k\}_{1\le k\le n}$; in this latter basis $\boldsymbol{A}$ is diagonal. The inverse transformation is $\boldsymbol{y} \mapsto \boldsymbol{P}\boldsymbol{y} = \boldsymbol{x}$.

Every real symmetric matrix $\boldsymbol{A}$ is associated to a **quadratic form** Q, which is a function $Q : \mathbb{R}^n \to \mathbb{R}$ satisfying $Q(\lambda\boldsymbol{x}) = \lambda^2 Q(\boldsymbol{x})$ for any $\boldsymbol{x} \in \mathbb{R}^n$, $\lambda \in \mathbb{R}$; to be precise,

$$Q(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}\cdot\boldsymbol{A}\boldsymbol{x} = \frac{1}{2}\boldsymbol{x}^T\boldsymbol{A}\boldsymbol{x} . \tag{4.16}$$

The eigenvalues of $\boldsymbol{A}$ determine the quadratic form Q: substituting to $\boldsymbol{A}$ the expression (4.15) and setting $\boldsymbol{y} = \boldsymbol{P}^T\boldsymbol{x}$, we get

$$Q(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x}^T\boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^T\boldsymbol{x} = \frac{1}{2}(\boldsymbol{P}^T\boldsymbol{x})^T\boldsymbol{\Lambda}(\boldsymbol{P}^T\boldsymbol{x}) = \frac{1}{2}\boldsymbol{y}^T\boldsymbol{\Lambda}\boldsymbol{y} .$$

Since $\boldsymbol{\Lambda}\boldsymbol{y} = (\lambda_k y_k)_{1\le k\le n}$ if $\boldsymbol{y} = (y_k)_{1\le k\le n}$, we conclude

$$Q(\boldsymbol{x}) = \frac{1}{2}\sum_{k=1}^{n}\lambda_k y_k^2 . \tag{4.17}$$

Consequently, one can classify $\boldsymbol{A}$ according to the sign of Q:

- $\boldsymbol{A}$ is **positive definite** if $Q(\boldsymbol{x}) > 0$ for any $\boldsymbol{x} \in \mathbb{R}^n$, $\boldsymbol{x} \ne \boldsymbol{0}$; equivalently, all eigenvalues of $\boldsymbol{A}$ are strictly positive.
- $\boldsymbol{A}$ is **positive semi-definite** if $Q(\boldsymbol{x}) \ge 0$ for any $\boldsymbol{x} \in \mathbb{R}^n$; equivalently, all eigenvalues of $\boldsymbol{A}$ are non-negative.
- $\boldsymbol{A}$ is **indefinite** if Q assumes on $\mathbb{R}^n$ both positive and negative values; this is to say $\boldsymbol{A}$ has positive and negative eigenvalues.

The notion of negative-definite and negative semi-definite matrices are clear.

Positive-definite symmetric matrices may be characterised in many other equivalent ways. For example, all first minors of $\boldsymbol{A}$ (those obtained by erasing the same rows and columns) have positive determinant. In particular, the diagonal entries a_{ii} are positive.

A crucial geometrical characterisation is the following: if $\boldsymbol{A}$ is positive definite and symmetric, the level sets

$$\{\boldsymbol{x} \in \mathbb{R}^n : Q(\boldsymbol{x}) = c > 0\}$$

of Q are generalised ellipses (e.g., ellipses in dimension 2, ellipsoids in dimension 3), with axes collinear to the eigenvectors of $\boldsymbol{A}$.

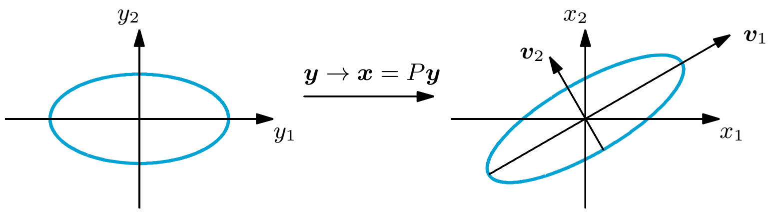


Figure 4.2. The conic associated to a positive-definite symmetric matrix

In fact, (restricting to dimension two for simplicity) the equation

$$\frac{1}{2}(\lambda_1 y_1^2 + \lambda_2 y_2^2) = c\,, \qquad \text{i.e.,} \qquad \frac{y_1^2}{(2c/\lambda_1)} + \frac{y_2^2}{(2c/\lambda_2)} = 1$$

defines an ellipse in canonical form with semi-axes of length $\sqrt{2c/\lambda_1}$ and $\sqrt{2c/\lambda_2}$ in the coordinates (y_1, y_2) associated to the eigenvectors $\boldsymbol{v}_1$, $\boldsymbol{v}_2$. In the original coordinates (x_1, x_2) the ellipse is rotated in the directions of $\boldsymbol{v}_1$, $\boldsymbol{v}_2$ (Fig. 4.2).

Equation (4.17) implies easily

$$Q(\boldsymbol{x}) \geq \frac{\lambda_*}{2}\|\boldsymbol{x}\|^2\,, \qquad \forall \boldsymbol{x} \in \mathbb{R}^n\,, \tag{4.18}$$

where $\lambda_* = \min\limits_{1 \leq k \leq n} \lambda_k$. In fact,

$$Q(\boldsymbol{x}) \geq \frac{\lambda_*}{2}\sum_{k=1}^{n} y_k^2 = \frac{\lambda_*}{2}\|\boldsymbol{y}\|^2$$

and $\|\boldsymbol{y}\|^2 = \|\boldsymbol{P}^T\boldsymbol{x}\|^2 = \boldsymbol{x}^T\boldsymbol{P}\boldsymbol{P}^T\boldsymbol{x} = \boldsymbol{x}^T\boldsymbol{x} = \|\boldsymbol{x}\|^2$ as $\boldsymbol{P}$ is orthogonal.

Example 4.1

Take the symmetric matrix of order 2

$$\boldsymbol{A} = \begin{pmatrix} 4 & \alpha \\ \alpha & 2 \end{pmatrix}$$

with α a real parameter. Solving the characteristic equation $\det(\boldsymbol{A} - \lambda \boldsymbol{I}) = (4-\lambda)(2-\lambda) - \alpha^2 = 0$, we find the eigenvalues:

$$\lambda_1 = 3 - \sqrt{1+\alpha^2}\,, \qquad \lambda_2 = 3 + \sqrt{1+\alpha^2} > 0\,.$$

Then

$$|\alpha| < 2\sqrt{2} \quad \Rightarrow \quad \lambda_1 > 0 \quad \Rightarrow \quad \boldsymbol{A} \text{ positive definite}\,,$$

$$|\alpha| = 2\sqrt{2} \quad \Rightarrow \quad \lambda_1 = 0 \quad \Rightarrow \quad \boldsymbol{A} \text{ positive semi-definite}\,,$$

$$|\alpha| > 2\sqrt{2} \quad \Rightarrow \quad \lambda_1 < 0 \quad \Rightarrow \quad \boldsymbol{A} \text{ indefinite}\,.$$

□

4.3 Sets in $\mathbb{R}^n$ and their properties

The study of limits and the continuity of functions of several variables require that we introduce some notions on vectors and subsets of $\mathbb{R}^n$.

Using distances we define neighbourhoods in $\mathbb{R}^n$.

Definition 4.2 *Take $\boldsymbol{x}_0 \in \mathbb{R}^n$ and let $r > 0$ be a real number. One calls* **neighbourhood** *of $\boldsymbol{x}_0$ of radius r the set*

$$B_r(\boldsymbol{x}_0) = \{\boldsymbol{x} \in \mathbb{R}^n : \|\boldsymbol{x} - \boldsymbol{x}_0\| < r\}$$

consisting of all points in $\mathbb{R}^n$ with distance from $\boldsymbol{x}_0$ smaller than r. The set

$$\overline{B}_r(\boldsymbol{x}_0) = \{\boldsymbol{x} \in \mathbb{R}^n : \|\boldsymbol{x} - \boldsymbol{x}_0\| \le r\}$$

is called **closed neighbourhood** *of $\boldsymbol{x}_0$ of radius r.*

Therefore $\overline{B}_r(\boldsymbol{x}_0)$ is the disc $(n = 2)$ or the ball $(n = 3)$ centred at $\boldsymbol{x}_0$ of radius r, while $B_r(\boldsymbol{x}_0)$ is a disc or ball without boundary.

If X is a subset of $\mathbb{R}^n$, by $\mathscr{C}X = \mathbb{R}^n \setminus X$ we denote the complement of X.

Definition 4.3 *A point $\boldsymbol{x} \in \mathbb{R}^n$ is called*

i) an **interior point** *of X if there is a neighbourhood $B_r(\boldsymbol{x})$ contained in X;*

ii) an **exterior point** *of X if it belongs to the interior of $\mathscr{C}X$;*

iii) a **boundary point** *of X if it is neither interior nor exterior for X.*

Fig. 4.3 depicts the various possibilities.

Boundary points can also be defined as the points whose every neighbourhood contains points of X and $\mathscr{C}X$ alike. It follows that X and its complement have the same boundary set.

Definition 4.4 *The set of interior points of X forms the* **interior** *of X, denoted by $\mathring{X}$ or $\operatorname{int} X$. Similarly, boundary points form the* **boundary** *of X, written ∂X. Exterior points form the* **exterior** *of X. Eventually, the set $X \cup \partial X$ is the* **closure** *of X, written $\overline{X}$.*

To be absolutely accurate, given a topological space X the set ∂X defined above should be called 'frontier', because the term 'boundary' is used in a different sense for topological manifolds (defining the topology of which is a rather delicate mat-

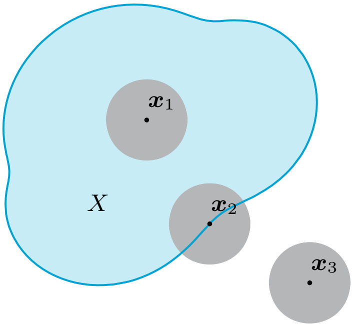


Figure 4.3. An interior point $\boldsymbol{x}_1$, a boundary point $\boldsymbol{x}_2$, an exterior point $\boldsymbol{x}_3$ of a set X

ter). Section 6.7.3 will discuss explicit examples where X is a surface. We shall not be this subtle and always speak about the boundary ∂X of X.

From the definition,

$$\overset{\circ}{X} \subseteq X \subseteq \overline{X}\,.$$

When one of the above inclusions is an equality the space X deserves one of the following names:

Definition 4.5 *The set X is* **open** *if all its points are interior points, $\overset{\circ}{X} = X$. It is* **closed** *if it contains its boundary, $X = \overline{X}$.*

A set is then open if and only if it contains a neighbourhood of each of its points, i.e., when it does not contain boundary points. Consequently, X is closed precisely if its complement is open, since a set and its complement share the boundary.

Examples 4.6

i) The sets $\mathbb{R}^n$ and $\emptyset$ are simultaneously open and closed (and are the only such subsets of $\mathbb{R}^n$, by the way). Their boundary is empty.

ii) The half-plane $X = \{(x, y) \in \mathbb{R}^2 : x > 1\}$ is open. In fact, given $(x_0, y_0) \in X$, any neighbourhood of radius $r \leq x_0 - 1$ is contained in X (Fig. 4.4, left). The boundary ∂X is the line $x = 1$ parallel to the y-axis (Fig. 4.4, right), hence the complementary set $\mathscr{C}X = \{(x, y) \in \mathbb{R}^2 : x \leq 1\}$ is closed.

Any half-plane $X \subset \mathbb{R}^2$ defined by an inequality like

$$ax + by > c \qquad \text{or} \qquad ax + by < c\,,$$

with one of a, b non-zero, is open; inequalities like

$$ax + by \geq c \qquad \text{or} \qquad ax + by \leq c$$

define closed half-planes. In either case the boundary ∂X is the line $ax + by = c$.

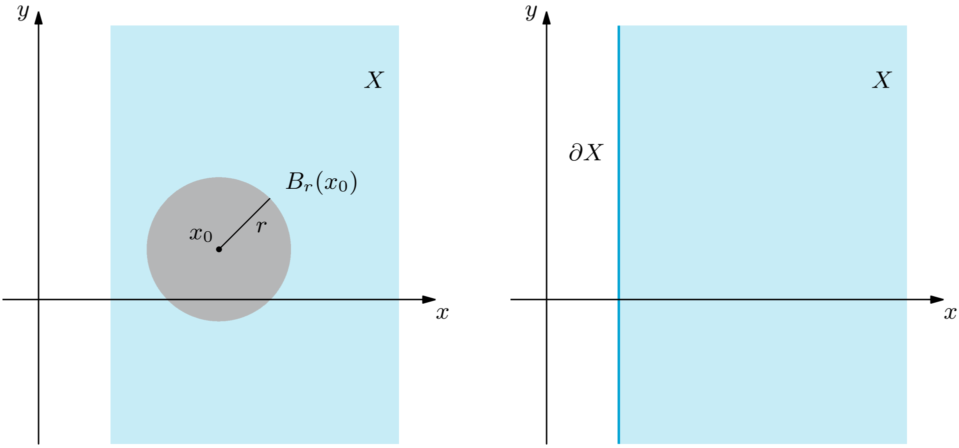


Figure 4.4. The boundary (right) of the half-plane $x > 1$ and the neighbourhood of a point (left)

iii) Let $X = B_1(\mathbf{0}) = \{\boldsymbol{x} \in \mathbb{R}^n : \|\boldsymbol{x}\| < 1\}$ be the n-dimensional ball with centre the origin and radius 1, without boundary. It is open, because for any $\boldsymbol{x}_0 \in X$, $B_r(\boldsymbol{x}_0) \subseteq X$ if we set $r \leq 1 - \|\boldsymbol{x}_0\|$: for $\|\boldsymbol{x} - \boldsymbol{x}_0\| < r$ in fact, we have

$$\|\boldsymbol{x}\| = \|\boldsymbol{x} - \boldsymbol{x}_0 + \boldsymbol{x}_0\| \leq \|\boldsymbol{x} - \boldsymbol{x}_0\| + \|\boldsymbol{x}_0\| < 1 - \|\boldsymbol{x}_0\| + \|\boldsymbol{x}_0\| = 1\,.$$

The boundary $\partial X = \{\boldsymbol{x} \in \mathbb{R}^n : \|\boldsymbol{x}\| = 1\}$ is the 'surface' of the ball (the n-dimensional sphere), while the closure of X is the ball itself, i.e., the closed neighbourhood $\overline{B}_1(\mathbf{0})$.

In general, for arbitrary $\boldsymbol{x}_0 \in \mathbb{R}^n$ and $r > 0$, the neighbourhood $B_r(\boldsymbol{x}_0)$ is open, it has boundary $\{\boldsymbol{x} \in \mathbb{R}^n : \|\boldsymbol{x} - \boldsymbol{x}_0\| = r\}$, and the closed neighbourhood $\overline{B}_r(\boldsymbol{x}_0)$ is the closure.

iv) The set $X = \{\boldsymbol{x} \in \mathbb{R}^n : 2 \leq \|\boldsymbol{x} - \boldsymbol{x}_0\| < 3\}$ (an annulus for $n = 2$, a spherical shell for $n = 3$) is neither open nor closed, for its interior is

$$\overset{\circ}{X} = \{\boldsymbol{x} \in \mathbb{R}^n : 2 < \|\boldsymbol{x} - \boldsymbol{x}_0\| < 3\}$$

and the closure

$$\overline{X} = \{\boldsymbol{x} \in \mathbb{R}^n : 2 \leq \|\boldsymbol{x} - \boldsymbol{x}_0\| \leq 3\}\,.$$

The boundary,

$$\partial X = \{\boldsymbol{x} \in \mathbb{R}^n : \|\boldsymbol{x} - \boldsymbol{x}_0\| = 2\} \cup \{\boldsymbol{x} \in \mathbb{R}^n : \|\boldsymbol{x} - \boldsymbol{x}_0\| = 3\}\,,$$

is the union of the two spheres delimiting X.

v) The set $X = [0,1]^2 \cap \mathbb{Q}^2$ of points with rational coordinates inside the unit square has empty interior and the whole square $[0,1]^2$ as boundary. Since rational numbers are dense in $\mathbb{R}$, any neighbourhood of a point in $[0,1]^2$ contains infinitely many points of X and infinitely many of the complement. □

Definition 4.7 *A point* $\boldsymbol{x} \in \mathbb{R}^n$ *is called a* **limit point** *of* X *if each of its neighbourhoods contains points of* X *different from* $\boldsymbol{x}$*:*

$$\forall r > 0\,, \qquad \big(B_r(\boldsymbol{x}) \setminus \{\boldsymbol{x}\}\big) \cap X \neq \emptyset\,.$$

A point $\boldsymbol{x} \in X$ *is* **isolated** *if there exists a neighbourhood* $B_r(\boldsymbol{x})$ *not containing points of* X *other than* $\boldsymbol{x}$*:*

$$\exists r > 0\,, \qquad \big(B_r(\boldsymbol{x}) \setminus \{\boldsymbol{x}\}\big) \cap X = \emptyset\,,$$

or equivalently,

$$\exists r > 0\,, \qquad B_r(\boldsymbol{x}) \cap X = \{\boldsymbol{x}\}\,.$$

Interior points of X are certainly limit points in X, whereas no exterior point can be a limit point. A boundary point of X must necessarily be either a limit point of X (belonging to X or not) or an isolated point. Conversely, a limit point can be interior or a non-isolated boundary point, and an isolated point is forced to lie on the boundary.

Example 4.8

Consider X the set $X = \{(x, y) \in \mathbb{R}^2 : \frac{y^2}{x^2} \leq 1 \text{ or } x^2 + (y-1)^2 \leq 0\}$.
Requiring $\frac{y^2}{x^2} \leq 1$, hence $\frac{|y|}{|x|} \leq 1$, defines the region lying between the lines $y = \pm x$ (included) and containing the x-axis except the origin (due to the denominator). To this we have to add the point $(0, 1)$, the unique solution to $x^2 + (y-1)^2 \leq 0$. See Fig. 4.5.
The limit points of X are those satisfying $y^2 \leq x^2$. They are either interior, when $y^2 < x^2$, or boundary points, when $y = \pm x$ (origin included). An additional boundary point is $(0, 1)$, also the only isolated point. □

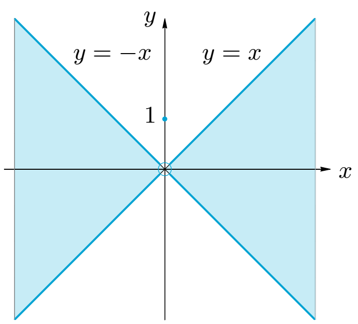


Figure 4.5. The set $X = \{(x, y) \in \mathbb{R}^2 : \frac{y^2}{x^2} \leq 1 \text{ or } x^2 + (y-1)^2 \leq 0\}$

Let us define a few more notions that will be useful in the sequel.

Definition 4.9 *A set X is* **bounded** *if there exists a real number $R > 0$ such that*
$$\|\boldsymbol{x}\| \le R\,, \qquad \forall \boldsymbol{x} \in X\,,$$
i.e., if X is contained in a closed neighbourhood $\overline{B}_R(\mathbf{0})$ of the origin.

Definition 4.10 *A set X is said* **compact** *if it is closed and bounded.*

Examples 4.11

i) The unit square $X = [0,1]^2 \subset \mathbb{R}^2$ is compact; it manifestly contains its boundary, and if we take $\boldsymbol{x} \in X$,
$$\|\boldsymbol{x}\| = \sqrt{x_1^2 + x_2^2} \le \sqrt{1+1} = \sqrt{2}\,.$$
In general, the n-dimensional cube $X = [0,1]^n \subset \mathbb{R}^n$ is compact.

ii) The elementary plane figures (such as rectangles, polygons, discs, ovals) and solids (tetrahedra, prisms, pyramids, cones and spheres) are all compact.

iii) The closure of a bounded set is compact. □

Let $\boldsymbol{a}$, $\boldsymbol{b}$ be distinct points of $\mathbb{R}^n$, and call $S[\boldsymbol{a},\boldsymbol{b}]$ the (closed) **segment** with end points $\boldsymbol{a}$, $\boldsymbol{b}$, i.e., the set of points on the line through $\boldsymbol{a}$ and $\boldsymbol{b}$ that lie between the two points:
$$\begin{aligned} S[\boldsymbol{a},\boldsymbol{b}] &= \{\boldsymbol{x} = \boldsymbol{a} + t(\boldsymbol{b}-\boldsymbol{a}) : 0 \le t \le 1\} \\ &= \{\boldsymbol{x} = (1-t)\boldsymbol{a} + t\boldsymbol{b} : 0 \le t \le 1\}\,. \end{aligned} \tag{4.19}$$

Definition 4.12 *A set X is called* **convex** *if the segment between any two points of X is all contained in X.*

Given $r+1$ points $\boldsymbol{a}_0$, $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_r$ in $\mathbb{R}^n$, all distinct (except possibly for $\boldsymbol{a}_0 = \boldsymbol{a}_r$), one calls **polygonal path** of vertices $\boldsymbol{a}_0$, $\boldsymbol{a}_1, \ldots, \boldsymbol{a}_r$ the union of the r segments $S[\boldsymbol{a}_{i-1}, \boldsymbol{a}_i]$, $1 \le i \le r$, joint at the end points:
$$P[\boldsymbol{a}_0, \ldots, \boldsymbol{a}_r] = \bigcup_{i=1}^{r} S[\boldsymbol{a}_{i-1}, \boldsymbol{a}_i]\,.$$

Definition 4.13 *An open set $A \subseteq \mathbb{R}^n$ is* (**path-**)**connected** *if, given two arbitrary points $\boldsymbol{x}$, $\boldsymbol{y}$ in A, there is a polygonal path joining them that is entirely contained in A.*

Figure 4.6 shows an example.

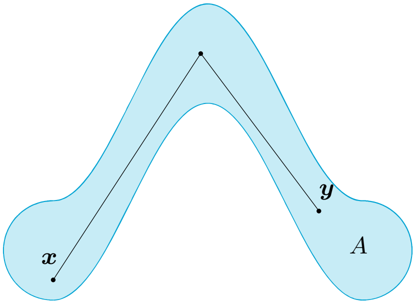


Figure 4.6. A connected (but not convex) set A

Examples 4.14

i) The only open connected subsets of $\mathbb{R}$ are the open intervals.

ii) An open and convex set A in $\mathbb{R}^n$ is obviously connected.

iii) Any annulus $A = \{\boldsymbol{x} \in \mathbb{R}^2 : r_1 < \|\boldsymbol{x} - \boldsymbol{x}_0\| < r_2\}$, with $\boldsymbol{x}_0 \in \mathbb{R}^2$, $r_2 > r_1 \geq 0$, is connected but not convex (Fig. 4.7, left). Finding a polygonal path between any two points $\boldsymbol{x}, \boldsymbol{y} \in A$ is intuitively quite easy.

iv) The open set $A = \{\boldsymbol{x} = (x, y) \in \mathbb{R}^2 : xy > 1\}$ is not connected (Fig. 4.7, right). □

It is a basic fact that an open non-empty set A of $\mathbb{R}^n$ is the union of a family $\{A_i\}_{i\in\mathcal{I}}$ of non-empty, connected and pairwise-disjoint open sets:

$$A = \bigcup_{i\in\mathcal{I}} A_i \qquad \text{with} \qquad A_i \cap A_j = \emptyset \text{ for } i \neq j\,.$$

Each A_i is called a **connected component** of A.

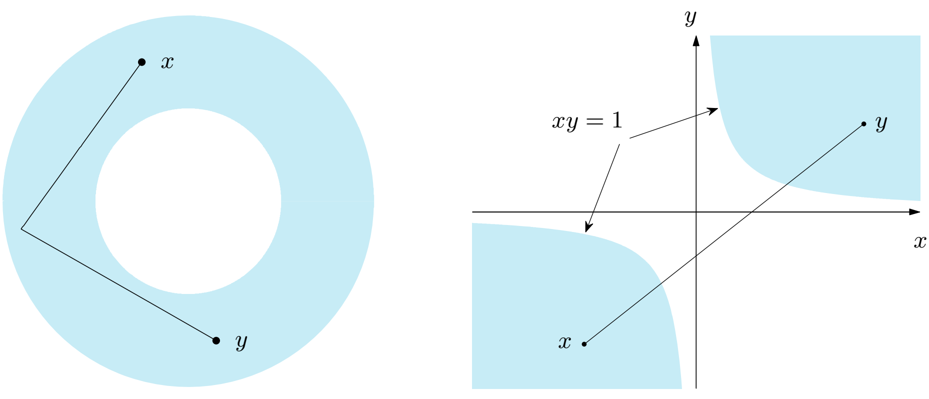


Figure 4.7. The set A of Examples 4.14 iii) (left) and iv) (right)

Every open connected set has only one connected component, namely itself. The set A of Example 4.14 iv) has two connected components

$$A_1 = \{\boldsymbol{x} = (x, y) \in A : x > 0\} \qquad \text{and} \qquad A_2 = \{\boldsymbol{x} = (x, y) \in A : x < 0\}.$$

Definition 4.15 *We call* **region** *any subset $\mathcal{R}$ of $\mathbb{R}^n$ made by the union of a non-empty, open, connected set A and part of the boundary ∂A*

$$\mathcal{R} = A \cup Z \qquad \textit{with} \qquad \emptyset \subseteq Z \subseteq \partial A.$$

When $Z = \emptyset$ the region is open, when $Z = \partial A$ it is closed.

A region may be defined equivalently as a non-empty set $\mathcal{R}$ of $\mathbb{R}^n$ whose interior $A = \overset{\circ}{\mathcal{R}}$ is connected and such that $A \subseteq \mathcal{R} \subseteq \overline{A}$.

Example 4.16

The set $\mathcal{R} = \{\boldsymbol{x} \in \mathbb{R}^2 : \sqrt{4 - x^2 - y^2} < 1\}$ is a region in the plane, for

$$\mathcal{R} = \{\boldsymbol{x} \in \mathbb{R}^2 : \sqrt{3} < \|\boldsymbol{x}\| \leq 2\}.$$

Therefore $A = \overset{\circ}{\mathcal{R}}$ is the open annulus $\{\boldsymbol{x} \in \mathbb{R}^2 : \sqrt{3} < \|\boldsymbol{x}\| < 2\}$, while $Z = \{\boldsymbol{x} \in \mathbb{R}^2 : \|\boldsymbol{x}\| = 2\} \subset \partial A$. □

4.4 Functions: definitions and first examples

We begin discussing real-valued maps, sometimes called (real) **scalar functions**. Let n be a given integer ≥ 1. A function f defined on $\mathbb{R}^n$ with values in $\mathbb{R}$ is a **real function of n real variables**; if $\operatorname{dom} f$ denotes its domain, we write

$$f : \operatorname{dom} f \subseteq \mathbb{R}^n \to \mathbb{R}.$$

The graph $\Gamma(f) = \{(\boldsymbol{x}, f(\boldsymbol{x})) \in \mathbb{R}^{n+1} : \boldsymbol{x} \in \operatorname{dom} f\}$ is a subset of $\mathbb{R}^{n+1}$.

The case $n = 1$ (one real variable) was dealt with in Vol. I exhaustively, so in the sequel we will consider scalar functions of two or more variables. If $n = 2$ or 3 the variable $\boldsymbol{x}$ will also be written as (x, y) or (x, y, z), respectively.

Examples 4.17

i) The map $z = f(x, y) = 2x - 3y$, defined on $\mathbb{R}^2$, has the plane $2x - 3y - z = 0$ as graph.

ii) The function $z = f(x, y) = \dfrac{y^2 + x^2}{y^2 - x^2}$ is defined on $\mathbb{R}^2$ without the lines $y = \pm x$.

iii) The function $w = f(x, y, z) = \sqrt{z - x^2 - y^2}$ has domain $\operatorname{dom} f = \{(x, y, z) \in \mathbb{R}^3 : z \geq x^2 + y^2\}$, which is made of the elliptic paraboloid $z = x^2 + y^2$ and the region inside of it.

iv) The function $f(\boldsymbol{x}) = f(x_1, \ldots, x_n) = \log(1 - x_1^2 - \ldots - x_n^2)$ is defined inside the n-dimensional sphere $x_1^2 + \ldots + x_n^2 = 1$, since $\operatorname{dom} f = \left\{\boldsymbol{x} \in \mathbb{R}^n : \sum_{i=1}^{n} x_i^2 < 1\right\}$. □

In principle, a scalar function can be drawn only when $n \leq 2$. For example, $f(x,y) = \sqrt{9 - x^2 - y^2}$ has graph in $\mathbb{R}^3$ given by $z = \sqrt{9 - x^2 - y^2}$. Squaring the equation yields

$$z = 9 - x^2 - y^2 \qquad \text{hence} \qquad x^2 + y^2 + z^2 = 9\,.$$

We recognise the equation of a sphere with centre the origin and radius 3. As $z \geq 0$, the graph of f is the hemisphere of Fig. 4.8.

Another way to visualise a function's behaviour, in two or three variables, is by finding its level sets. Given a real number c, the **level set**

$$\boxed{L(f,c) = \{\boldsymbol{x} \in \operatorname{dom} f : f(\boldsymbol{x}) = c\}} \tag{4.20}$$

is the subset of $\mathbb{R}^n$ where the function is constant, equal to c. Figure 4.9 shows some level sets for the function $z = f(x,y)$ of Example 4.9 ii).

Geometrically, in dimension $n = 2$, a level set is the projection on the xy-plane of the intersection between the graph of f and the plane $z = c$. Clearly, $L(f,c)$ is not empty if and only if $c \in \operatorname{im} f$. A level set may have an extremely complicated shape. That said, we shall see in Sect. 7.2 certain assumptions on f that guarantee $L(f,c)$ consists of curves (dimension two) or surfaces (dimension three).

Consider now the more general situation of a map between Euclidean spaces, and precisely: given integers $n, m \geq 1$, we denote by $\boldsymbol{f}$ an arbitrary function on $\mathbb{R}^n$ with values in $\mathbb{R}^m$

$$\boldsymbol{f} : \operatorname{dom} \boldsymbol{f} \subseteq \mathbb{R}^n \to \mathbb{R}^m\,.$$

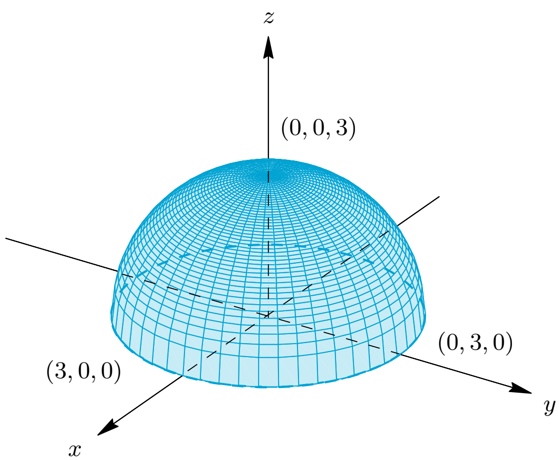


Figure 4.8. The function $f(x,y) = \sqrt{9 - x^2 - y^2}$

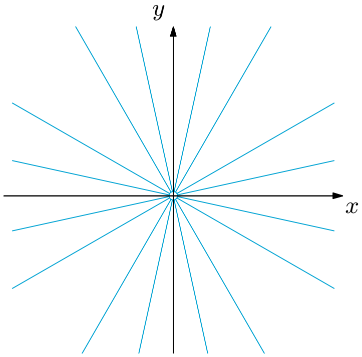


Figure 4.9. Level sets for $z = f(x, y) = \dfrac{y^2 + x^2}{y^2 - x^2}$

If $m = 1$, we have the scalar functions of above. If $m \geq 2$ we shall say $\boldsymbol{f}$ is a real **vector-valued** function.

Let us see some interesting cases. Curves, seen in Vol. I for $m = 2$ (plane curves) and $m = 3$ (curves in space), are special vector-valued maps where $n = 1$; surfaces in space are vector-valued functions with $n = 2$ and $m = 3$. The study of curves and surfaces is postponed to Sects. 4.6, 4.7 and Chapter 6 in particular. In the case $n = m$, $\boldsymbol{f}$ is called a **vector field**. An example with $n = m = 3$ is the Earth's gravitational field.

Let f_i, $1 \leq i \leq m$, be the components of $\boldsymbol{f}$ with respect to the canonical basis $\{\boldsymbol{e}_i\}_{1 \leq i \leq m}$ of $\mathbb{R}^m$:

$$\boldsymbol{f}(\boldsymbol{x}) = \big(f_i(\boldsymbol{x})\big)_{1 \leq i \leq m} = \sum_{i=1}^{m} f_i(\boldsymbol{x})\boldsymbol{e}_i \,.$$

Each f_i is a real scalar function of one or more real variables, defined on dom $\boldsymbol{f}$ at least; actually, dom $\boldsymbol{f}$ is the intersection of the domains of the components of $\boldsymbol{f}$.

Examples 4.18

i) Consider the vector field on $\mathbb{R}^2$

$$\boldsymbol{f}(x, y) = (-y, x)\,.$$

The best way to visualise a two-dimensional field is to draw the vector corresponding to $\boldsymbol{f}(x, y)$ as position vector at the point (x, y). This is clearly not possible everywhere on the plane, but a sufficient number of points might still give a reasonable idea of the behaviour of $\boldsymbol{f}$. Since $\boldsymbol{f}(1, 0) = (0, 1)$, we draw the vector $(0, 1)$ at the point $(1, 0)$; similarly, we plot the vector $(-1, 0)$ at $(0, 1)$ because $\boldsymbol{f}(0, 1) = (-1, 0)$ (see Fig. 4.10, left).
Notice that each vector is tangent to a circle centred at the origin. In fact, the dot product of the position vector $\boldsymbol{x} = (x, y)$ with $\boldsymbol{f}(\boldsymbol{x})$ is zero:

$$\boldsymbol{x} \cdot \boldsymbol{f}(\boldsymbol{x}) = (x, y) \cdot (-y, x) = -xy + xy = 0\,,$$

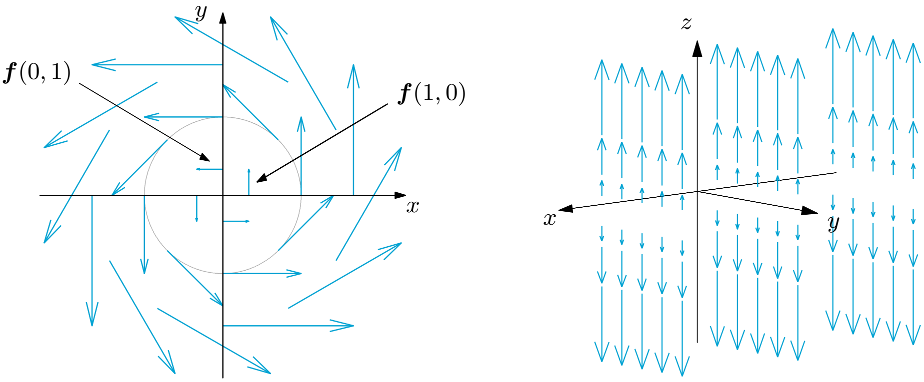


Figure 4.10. The fields $\boldsymbol{f}(x, y) = (-y, x)$ (left) and $\boldsymbol{f}(x, y, z) = (0, 0, z)$ (right)

making $\boldsymbol{x}$ and $\boldsymbol{f}(\boldsymbol{x})$ orthogonal vectors. Additionally, $\|\boldsymbol{f}(\boldsymbol{x})\| = \|\boldsymbol{x}\|$, so the length of $\boldsymbol{f}(x, y)$ coincides with the circle's radius.
This vector field represents the velocity of a wheel spinning counter-clockwise.

ii) Vector fields in $\mathbb{R}^3$ can be understood in a similar way. Figure 4.10, right, shows a picture of the vector field

$$\boldsymbol{f}(x, y, z) = (0, 0, z)\,.$$

All vectors are vertical, and point upwards if they lie above the xy-plane, downwards if below the plane $z = 0$. The magnitude increases as we move away from the xy-plane.

iii) Imagine a fluid running through a pipe with velocity $\boldsymbol{f}(x, y, z)$ at the point (x, y, z). The function $\boldsymbol{f}$ assignes a vector to each pont (x, y, z) in a certain domain Ω (the region inside the pipe) and so is a vector field of $\mathbb{R}^3$, called the velocity vector field. A concrete example is in Fig. 4.11.

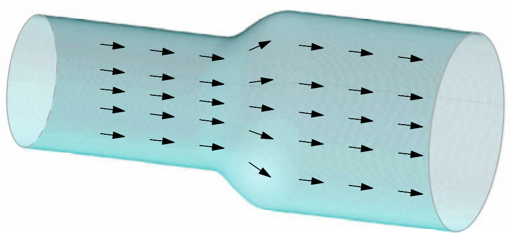


Figure 4.11. The velocity vector field of a fluid moving in a pipe

iv) The $\mathbb{R}^3$-valued map

$$\boldsymbol{f}(x,y,z) = \left(\frac{x}{(x^2+y^2+z^2)^{3/2}}, \frac{y}{(x^2+y^2+z^2)^{3/2}}, \frac{z}{(x^2+y^2+z^2)^{3/2}}\right) = \frac{\boldsymbol{x}}{\|\boldsymbol{x}\|^3},$$

on $\mathbb{R}^3 \setminus \{\mathbf{0}\}$ represents the electrostatic force field generated by a charged particle placed in the origin.

v) Let $\boldsymbol{A} = (a_{ij})_{\substack{1 \le i \le m \\ 1 \le j \le n}}$ be a real $m \times n$ matrix. The function

$$\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x}\,,$$

is a linear map from $\mathbb{R}^n$ to $\mathbb{R}^m$. □

4.5 Continuity and limits

The notion of continuity for functions between Euclidean spaces is essentially the same as what we have seen for one-variable functions (Vol. I, Ch. 3), with the proviso that the absolute value of $\mathbb{R}$ must be replaced by an arbitrary norm in $\mathbb{R}^n$ or $\mathbb{R}^m$ (which we shall indicate with $\|\cdot\|$ for simplicity).

Definition 4.19 *A function $\boldsymbol{f} : \mathrm{dom}\,\boldsymbol{f} \subseteq \mathbb{R}^n \to \mathbb{R}^m$ is said* **continuous at** $\boldsymbol{x}_0 \in \mathrm{dom}\,\boldsymbol{f}$ *if for any $\varepsilon > 0$ there exists a $\delta > 0$ such that*

$$\forall \boldsymbol{x} \in \mathrm{dom}\,\boldsymbol{f}, \qquad \|\boldsymbol{x} - \boldsymbol{x}_0\| < \delta \quad \Rightarrow \quad \|\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{x}_0)\| < \varepsilon\,;$$

that is to say,

$$\forall \boldsymbol{x} \in \mathrm{dom}\,\boldsymbol{f}, \qquad \boldsymbol{x} \in B_\delta(\boldsymbol{x}_0) \quad \Rightarrow \quad f(\boldsymbol{x}) \in B_\varepsilon(\boldsymbol{f}(\boldsymbol{x}_0)).$$

A map $\boldsymbol{f}$ is continuous on a set $\Omega \subseteq \mathrm{dom}\,\boldsymbol{f}$ if it is continuous at each point $\boldsymbol{x} \in \Omega$.

The following result is used a lot to study the continuity of vector-valued functions. Its proof is left to the reader.

Proposition 4.20 *The map $\boldsymbol{f} = (f_i)_{1 \le i \le m}$ is continuous at $\boldsymbol{x}_0 \in \mathrm{dom}\,\boldsymbol{f}$ if and only if all its components f_i are continuous.*

Due to this result we shall merely provide some examples of scalar functions.

Examples 4.21

i) Let us verify $f : \mathbb{R}^2 \to \mathbb{R}$, $f(\boldsymbol{x}) = 2x_1 + 5x_2$ is continuous at $\boldsymbol{x}_0 = (3,1)$. Using the fact that $|y_i| \le \|\boldsymbol{y}\|$ for all i (mentioned earlier), we have

$$|f(\boldsymbol{x}) - f(\boldsymbol{x}_0)| = |2(x_1 - 3) + 5(x_2 - 1)| \le 2|x_1 - 3| + 5|x_2 - 1| \le 7\|\boldsymbol{x} - \boldsymbol{x}_0\|\,.$$

Given $\varepsilon > 0$ then, it is enough to choose $\delta = \varepsilon/7$ to conclude. The same argument shows f is continuous at each $\boldsymbol{x}_0 \in \mathbb{R}^2$.

ii) The above function is an affine map $f : \mathbb{R}^n \to \mathbb{R}$, i.e., a map of type $f(\boldsymbol{x}) = \boldsymbol{a} \cdot \boldsymbol{x} + b$ ($\boldsymbol{a} \in \mathbb{R}^n,\ b \in \mathbb{R}$). Affine maps are continuous at each $\boldsymbol{x}_0 \in \mathbb{R}^n$ because

$$|f(\boldsymbol{x}) - f(\boldsymbol{x}_0)| = |\boldsymbol{a} \cdot \boldsymbol{x} - \boldsymbol{a} \cdot \boldsymbol{x}_0| = |\boldsymbol{a} \cdot (\boldsymbol{x} - \boldsymbol{x}_0)| \le \|\boldsymbol{a}\| \, \|\boldsymbol{x} - \boldsymbol{x}_0\| \qquad (4.21)$$

by the Cauchy-Schwarz inequality (4.3).

If $\boldsymbol{a} = \boldsymbol{0}$, the result is trivial. If $\boldsymbol{a} \neq \boldsymbol{0}$, continuity holds if one chooses $\delta = \varepsilon/\|\boldsymbol{a}\|$ for any given $\varepsilon > 0$. □

The map $\boldsymbol{f}$ is *uniformly continuous* on Ω if we may choose δ independently of $\boldsymbol{x}_0$ in the above definition; this is made precise as follows.

Definition 4.22 *A function is said* **uniformly continuous** *on* $\Omega \subseteq \operatorname{dom} \boldsymbol{f}$ *if for any* $\varepsilon > 0$ *there exists a* $\delta > 0$ *such that*

$$\forall \boldsymbol{x}', \boldsymbol{x}'' \in \Omega, \qquad \|\boldsymbol{x}' - \boldsymbol{x}''\| < \delta \quad \Rightarrow \quad \|\boldsymbol{f}(\boldsymbol{x}') - \boldsymbol{f}(\boldsymbol{x}'')\| < \varepsilon\,. \qquad (4.22)$$

For instance, the above affine function f is uniformly continuous on $\mathbb{R}^n$.

A continuous function on a closed and bounded set (i.e., a compact set) Ω is uniformly continuous therein (Theorem of Heine-Cantor, given in Appendix A.1.3, p. 515).

Often one can study the continuity of a function of several variables without turning to the definition. To this end the next three criteria are rather practical.

i) If the map φ is defined and continuous on a set $I \subseteq \mathbb{R}$, then

$$f(\boldsymbol{x}) = \varphi(x_1)$$

is defined and continuous on $\Omega = I \times \mathbb{R}^{n-1} \subseteq \mathbb{R}^n$.
In general, any continuous map of m variables is continuous if we think of it as a map of $n > m$ variables.

For example, the following functions are continuous:

$$\begin{aligned} f_1(x,y) &= \mathrm{e}^x && \text{on } \operatorname{dom} f_1 = \mathbb{R}^2\,, \\ f_2(x,y,z) &= \sqrt{1-y^2} && \text{on } \operatorname{dom} f_2 = \mathbb{R} \times [-1,1] \times \mathbb{R}\,. \end{aligned}$$

ii) If f and g are continuous on $\Omega \subseteq \mathbb{R}^n$, then also $f+g$, $f-g$ and fg are continuous on Ω, while f/g is continuous on the subset of Ω where $g \neq 0$.

Examples of continuous maps:

$$h_1(x,y) = \mathrm{e}^x + \sin y \qquad \text{on } \operatorname{dom} h_1 = \mathbb{R}^2\,,$$

$$h_2(x,y) = y\log x \qquad \text{on } \operatorname{dom} h_2 = (0,+\infty)\times\mathbb{R}\,,$$
$$h_3(x,y,z) = \frac{\arctan y}{x^2+z^2} \quad \text{on } \operatorname{dom} h_3 = \mathbb{R}^3 \setminus \big(\{0\}\times\mathbb{R}\times\{0\}\big)\,.$$

iii) If f is continuous on $\Omega \subseteq \mathbb{R}^n$ and g is continuous on $I \subseteq \mathbb{R}$, the composite map $g\circ f$ is continuous on $\operatorname{dom} g\circ f = \{\boldsymbol{x} \in \Omega : f(\boldsymbol{x}) \in I\}$.

For instance,

$$h_1(x,y) = \log(1+xy) \qquad \text{on } \operatorname{dom} h_1 = \{(x,y)\in\mathbb{R}^2 : xy > -1\}\,,$$
$$h_2(x,y,z) = \left|\frac{x^3+y^2}{z-1}\right| \qquad \text{on } \operatorname{dom} h_2 = \mathbb{R}^2 \times \big(\mathbb{R}\setminus\{1\}\big)\,,$$
$$h_3(\boldsymbol{x}) = \sqrt[4]{x_1^4+\ldots+x_n^4} \qquad \text{on } \operatorname{dom} h_3 = \mathbb{R}^n\,.$$

In particular, Proposition 4.20 and criterion *iii)* imply the next result, which is about the continuity of a composite map in the most general setting.

Proposition 4.23 *Let*

$$\boldsymbol{f} : \operatorname{dom}\boldsymbol{f} \subseteq \mathbb{R}^n \to \mathbb{R}^m\,, \qquad \boldsymbol{g} : \operatorname{dom}\boldsymbol{g} \subseteq \mathbb{R}^m \to \mathbb{R}^p$$

be functions and $\boldsymbol{x}_0 \in \operatorname{dom}\boldsymbol{f}$ a point such that $\boldsymbol{y}_0 = \boldsymbol{f}(\boldsymbol{x}_0) \in \operatorname{dom}\boldsymbol{g}$. Consider the composite map

$$\boldsymbol{h} = \boldsymbol{g}\circ\boldsymbol{f} : \operatorname{dom}\boldsymbol{h} \subseteq \mathbb{R}^n \to \mathbb{R}^p\,,$$

where $\boldsymbol{x}_0 \in \operatorname{dom}\boldsymbol{h}$. Then if $\boldsymbol{f}$ is continuous at $\boldsymbol{x}_0$ and $\boldsymbol{g}$ is continuous at $\boldsymbol{y}_0$, $\boldsymbol{h}$ is continuous at $\boldsymbol{x}_0$.

The definition of finite limit of a vector-valued map, for $\boldsymbol{x} \to \boldsymbol{x}_0 \in \mathbb{R}^n$, is completely analogous to the one-variable case. From now on we will suppose $\boldsymbol{f}$ is defined on $\operatorname{dom}\boldsymbol{f} \subseteq \mathbb{R}^n$ and $\boldsymbol{x}_0 \in \mathbb{R}^n$ is a limit point of $\operatorname{dom}\boldsymbol{f}$.

Definition 4.24 *One says that* $\boldsymbol{f}$ **has limit** $\boldsymbol{\ell} \in \mathbb{R}^m$ (*or* **tends to** $\boldsymbol{\ell}$) **as** $\boldsymbol{x}$ **tends to** $\boldsymbol{x}_0$, *in symbols*

$$\lim_{\boldsymbol{x}\to\boldsymbol{x}_0} \boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{\ell},$$

if for any $\varepsilon > 0$ there is a $\delta > 0$ such that

$$\forall \boldsymbol{x} \in \operatorname{dom}\boldsymbol{f}, \qquad 0 < \|\boldsymbol{x}-\boldsymbol{x}_0\| < \delta \quad \Rightarrow \quad \|\boldsymbol{f}(\boldsymbol{x})-\boldsymbol{\ell}\| < \varepsilon\,, \tag{4.23}$$

i.e.,

$$\forall \boldsymbol{x} \in \operatorname{dom}\boldsymbol{f}, \qquad \boldsymbol{x} \in B_\delta(\boldsymbol{x}_0)\setminus\{\boldsymbol{x}_0\} \quad \Rightarrow \quad \boldsymbol{f}(\boldsymbol{x}) \in B_\varepsilon(\boldsymbol{\ell})\,.$$

As for one real variable, if $\boldsymbol{f}$ is defined on $\boldsymbol{x}_0$, then

$$\boxed{\boldsymbol{f} \text{ continuous at } \boldsymbol{x}_0 \quad \iff \quad \lim_{\boldsymbol{x}\to\boldsymbol{x}_0} \boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{x}_0)\,.}$$

The analogue of Proposition 4.20 holds for limits, justifying the component-by-component approach.

Examples 4.25

i) The map

$$f(x,y) = \frac{x^4+y^4}{x^2+y^2}$$

is defined on $\mathbb{R}^2 \setminus \{(0,0)\}$ and continuous on its domain, as is clear by using criteria *i)* and *ii)* on p. 131. What is more,

$$\lim_{(x,y)\to(0,0)} f(x,y) = 0\,,$$

because

$$x^4+y^4 \le x^4+2x^2y^2+y^4 = (x^2+y^2)^2$$

and by definition of limit

$$\left|\frac{x^4+y^4}{x^2+y^2}\right| \le x^2+y^2 < \varepsilon \qquad \text{if} \qquad 0 < \|\boldsymbol{x}\| < \sqrt{\varepsilon}\,;$$

therefore the condition is true with $\delta = \sqrt{\varepsilon}$.

ii) Let us check that

$$\lim_{(x,y)\to(0,0)} \frac{|x|+|y|}{x^2+y^2} = +\infty\,.$$

Since

$$x^2+y^2 \le x^2+2|x|\,|y|+y^2 = (|x|+|y|)^2,$$

so that $\|\boldsymbol{x}\| \le |x|+|y|$, we have

$$f(x,y) = \frac{|x|+|y|}{\|\boldsymbol{x}\|^2} = \frac{|x|+|y|}{\|\boldsymbol{x}\|}\,\frac{1}{\|\boldsymbol{x}\|} \ge \frac{1}{\|\boldsymbol{x}\|}\,,$$

and $f(x,y) > A$ if $\|\boldsymbol{x}\| < 1/A$; the condition for the limit is true by taking $\delta = 1/A$. □

A *necessary condition* for the limit of $f(\boldsymbol{x})$ to exist (finite or not) as $\boldsymbol{x} \to \boldsymbol{x}_0$, is that the restriction of f to any line through $\boldsymbol{x}_0$ has the same limit. This observation is often used to show that a certain limit does not exist.

Example 4.26

The map

$$f(x,y) = \frac{x^3+y^2}{x^2+y^2}$$

does not admit limit for $(x,y) \to (0,0)$. Suppose the contrary, and let

$$L = \lim_{(x,y)\to(0,0)} f(x,y)$$

be finite or infinite. Necessarily then,

$$L = \lim_{x\to 0} f(x,0) = \lim_{y\to 0} f(0,y)\,.$$

But $f(x,0) = x$, so $\lim_{x\to 0} f(x,0) = 0$, and $f(0,y) = 1$ for any $y \neq 0$, whence $\lim_{y\to 0} f(0,y) = 1\,.$ □

One should not be led to believe, though, that the behaviour of the function restricted to lines through $\boldsymbol{x}_0$ is sufficient to compute the limit. Lines represent but one way in which we can approach the point $\boldsymbol{x}_0$. Different relationships among the variables may give rise to completely different behaviours of the limit.

Example 4.27

Let us determine the limit of

$$f(x,y) = \begin{cases} x\mathrm{e}^{x/y} & \text{if } y \neq 0\,, \\ 0 & \text{if } y = 0\,, \end{cases}$$

at the origin. The map tends to 0 along each straight line passing through $(0,0)$:

$$\lim_{x\to 0} f(x,0) = \lim_{y\to 0} f(0,y) = 0$$

because the function is identically zero on the coordinate axes, and along the other lines $y = kx$, $k \neq 0$,

$$\lim_{x\to 0} f(x,kx) = \lim_{x\to 0} x\mathrm{e}^{k} = 0\,.$$

Yet along the parabolic arc $y = x^2$, $x > 0$, the map tends to infinity, for

$$\lim_{x\to 0^+} f(x,x^2) = \lim_{x\to 0^+} x\mathrm{e}^{1/x} = +\infty\,.$$

Therefore the function does not admit limit as $(x,y) \to (0,0)$. □

A function of several variables can be proved to admit limit for $\boldsymbol{x}$ tending to $\boldsymbol{x}_0$ (see Remark 4.35 for more details) if and only if the limit behaviour is independent of the path through $\boldsymbol{x}_0$ chosen (see Fig. 4.12 for some examples). The previous cases show that if that is not true, the limit does not exist.

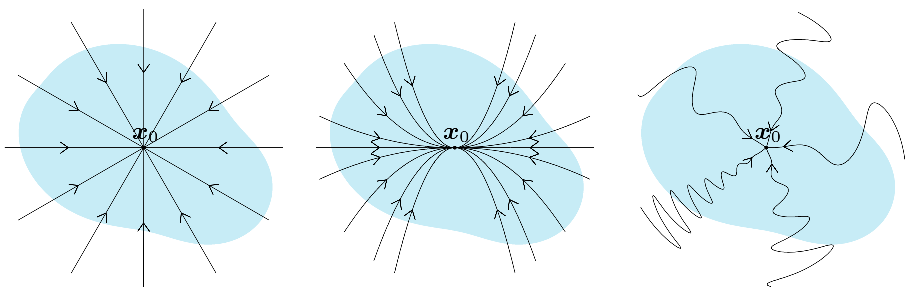


Figure 4.12. Different ways of approaching the point $\boldsymbol{x}_0$

For two variables, a useful *sufficient condition* to study the limit's existence relies on polar coordinates

$$x = x_0 + r\cos\theta\,, \qquad y = y_0 + r\sin\theta\,.$$

Proposition 4.28 *Suppose there exist $\ell \in \mathbb{R}$ and a map g depending on the variable r such that, on a neighbourhood of (x_0, y_0),*

$$|f(x_0 + r\cos\theta\,, y_0 + r\sin\theta) - \ell| \le g(r) \qquad \textit{with} \qquad \lim_{r\to 0^+} g(r) = 0\,.$$

Then

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

Examples 4.29

i) The above result simplifies the study of the limit of Example 4.25 i):

$$\begin{aligned} f(r\cos\theta, r\sin\theta) &= \frac{r^4\sin^4\theta + r^4\cos^4\theta}{r^2} = r^2(\sin^4\theta + \cos^4\theta) \\ &\le r^2(\sin^2\theta + \cos^2\theta) = r^2\,, \end{aligned}$$

recalling $|\sin\theta| \le 1$ and $|\cos\theta| \le 1$. Thus

$$|f(r\cos\theta, r\sin\theta)| \le r^2$$

and the criterion applies with $g(r) = r^2$.

ii) Consider

$$\lim_{(x,y)\to(0,1)} \frac{x\log y}{\sqrt{x^2 + (y-1)^2}}\,.$$

Then

$$f(r\cos\theta, 1 + r\sin\theta) = \frac{r\cos\theta\log(1 + r\sin\theta)}{r} = \cos\theta\log(1 + r\sin\theta)\,.$$

Remembering that

$$\lim_{t\to 0} \frac{\log(1+t)}{t} = 1\,,$$

for t small enough we have

$$\left|\frac{\log(1+t)}{t}\right| \le 2\,,$$

i.e., $|\log(1+t)| \le 2|t|$. Then

$$|f(r\cos\theta, 1 + r\sin\theta)| = |\cos\theta\log(1 + r\sin\theta)| \le 2r|\sin\theta\cos\theta| \le 2r\,.$$

The criterion can be used with $g(r) = 2r$, to the effect that

$$\lim_{(x,y)\to(0,1)} \frac{x\log y}{\sqrt{x^2 + (y-1)^2}} = 0\,.$$

□

We would also like to understand what happens when the norm of the independent variable tends to infinity. As there is no natural ordering on $\mathbb{R}^n$ for $n \geq 2$, one cannot discern, in general, how the argument of the map moves away from the origin. In contrast to dimension one, where it is possible to distinguish the limits for $x \to +\infty$ and $x \to -\infty$, in higher dimensions there is only one "point" at infinity ∞. Neighbourhoods of the point at infinity are by definition

$$\boxed{B_R(\infty) = \{\boldsymbol{x} \in \mathbb{R}^n : \|\boldsymbol{x}\| > R\} \qquad \text{with} \quad R > 0\,.}$$

Each $B_R(\infty)$ is the complement of the closed neighbourhood $\overline{B}_R(\mathbf{0})$ of radius R centred at the origin.

With this, the definition of limit (finite or infinite) assumes the usual form. For example a function $\boldsymbol{f}$ with unbounded domain in $\mathbb{R}^n$ **has limit $\boldsymbol{\ell} \in \mathbb{R}^m$ as $\boldsymbol{x}$ tends to $\boldsymbol{\infty}$**, written

$$\boxed{\lim_{\boldsymbol{x}\to\boldsymbol{\infty}} \boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{\ell} \in \mathbb{R}^m\,,}$$

if for any $\varepsilon > 0$ there is an $R > 0$ such that

$$\boxed{\forall \boldsymbol{x} \in \operatorname{dom} \boldsymbol{f}\,, \qquad \|\boldsymbol{x}\| > R \quad \Rightarrow \quad \|\boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{\ell}\| < \varepsilon\,,}$$

i.e.,

$$\forall \boldsymbol{x} \in \operatorname{dom} \boldsymbol{f}\,, \qquad \boldsymbol{x} \in B_R(\infty) \quad \Rightarrow \quad \boldsymbol{f}(\boldsymbol{x}) \in B_\varepsilon(\boldsymbol{\ell})\,.$$

Eventually, we discuss infinite limits. A scalar function f **has limit** $+\infty$ (or **tends to** $+\infty$) **as $\boldsymbol{x}$ tends to $\boldsymbol{x}_0$**, written

$$\boxed{\lim_{\boldsymbol{x}\to\boldsymbol{x}_0} f(\boldsymbol{x}) = +\infty,}$$

if for any $R > 0$ there is a $\delta > 0$ such that

$$\boxed{\forall \boldsymbol{x} \in \operatorname{dom} f, \qquad 0 < \|\boldsymbol{x} - \boldsymbol{x}_0\| < \delta \quad \Rightarrow \quad f(\boldsymbol{x}) > R\,,} \tag{4.24}$$

i.e.,

$$\forall \boldsymbol{x} \in \operatorname{dom} f, \qquad \boldsymbol{x} \in B_\delta(\boldsymbol{x}_0) \setminus \{\boldsymbol{x}_0\} \quad \Rightarrow \quad f(\boldsymbol{x}) \in B_R(+\infty).$$

The definitions of

$$\lim_{\boldsymbol{x}\to\boldsymbol{x}_0} f(\boldsymbol{x}) = -\infty \qquad \text{and} \qquad \lim_{\boldsymbol{x}\to\boldsymbol{\infty}} f(\boldsymbol{x}) = -\infty$$

descend from the previous ones, substituting $f(\boldsymbol{x}) > R$ with $f(\boldsymbol{x}) < -R$.

In the vectorial case, the limit is infinite when at least one component of $\boldsymbol{f}$ tends to ∞. Here as well we cannot distinguish how $\boldsymbol{f}$ grows. Precisely, $\boldsymbol{f}$ **has limit $\boldsymbol{\infty}$** (or **tends to $\boldsymbol{\infty}$**) **as $\boldsymbol{x}$ tends to $\boldsymbol{x}_0$**, which one indicates by

$$\boxed{\lim_{\boldsymbol{x}\to\boldsymbol{x}_0} \boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{\infty},}$$

if

$$\lim_{\boldsymbol{x}\to\boldsymbol{x}_0} \|\boldsymbol{f}(\boldsymbol{x})\| = +\infty\,.$$

Similarly we define

$$\lim_{\boldsymbol{x}\to\boldsymbol{\infty}} \boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{\infty}\,.$$

4.5.1 Properties of limits and continuity

The main theorems on limits of one-variable maps, discussed in Vol. 1, Ch. 4, carry over to several variables. Precisely, we still have uniqueness of the limit, local invariance of tha function's sign, the various Comparison Theorems plus corollaries, the algebra of limits with the relative indeterminate forms. Clearly, $\boldsymbol{x} \to \boldsymbol{c}$ should be understood by thinking $\boldsymbol{c}$ as either a point $\boldsymbol{x}_0 \in \mathbb{R}^n$ or $\boldsymbol{\infty}$.

We can also use the Landau formalism for a local study in several variables. The definition and properties of the symbols O, o, $\asymp$, $\sim$ depend on limits, and thus extend. The expression $f(\boldsymbol{x}) = o(\|\boldsymbol{x}\|)$ as $\boldsymbol{x} \to \boldsymbol{0}$, for instance, means

$$\lim_{\boldsymbol{x}\to\boldsymbol{0}} \frac{f(\boldsymbol{x})}{\|\boldsymbol{x}\|} = 0\,.$$

An example satisfying this property is $f(x,y) = 2x^2 - 5y^3$.

Some continuity theorems of global nature, see Vol. I, Sect. 4.3, have a counterpart for scalar maps of several variables. For example, the theorem on the existence of zeroes goes as follows.

Theorem 4.30 *Let f be a continuous map on a region $\mathcal{R} \subseteq \mathbb{R}^n$. If f assumes on $\mathcal{R}$ both positive and negative values, it necessarily has a zero on $\mathcal{R}$.*

Proof. If $\boldsymbol{a}, \boldsymbol{b} \in \mathcal{R}$ satisfy $f(\boldsymbol{a}) < 0$ and $f(\boldsymbol{b}) > 0$, and if $P[\boldsymbol{a}, \ldots, \boldsymbol{b}]$ is an arbitrary polygonal path in $\mathcal{R}$ joining $\boldsymbol{a}$ and $\boldsymbol{b}$, the map f restricted to $P[\boldsymbol{a}, \ldots, \boldsymbol{b}]$ is a function of one variable that satisfies the ordinary Theorem of Existence of Zeroes. Therefore an $\boldsymbol{x}_0 \in P[\boldsymbol{a}, \ldots, \boldsymbol{b}]$ exists with $f(\boldsymbol{x}_0) = 0$. □

From this follows, as in the one-dimensional case, the Mean Value Theorem. We also have Weierstrass's Theorem, which we will see in Sect. 5.6, Theorem 5.24.

For maps valued in $\mathbb{R}^m$, $m \geq 1$, we have the results on limits that make sense for vectorial quantities (e.g., the uniqueness of the limit and the limit of a sum of functions, but not the Comparison Theorems).

The Substitution Theorem holds, just as in dimension one (Vol. I, Thm. 4.15); for continuous maps this guarantees the continuity of the composite map, as mentioned in Proposition 4.23.

Examples 4.31

i) We prove that

$$\lim_{(x,y)\to\infty} \mathrm{e}^{-\frac{1}{|x|+|y|}} = 1\,.$$

As $(x,y) \to \infty$, we have $|x|+|y| \to +\infty$, hence $t = -\frac{1}{|x|+|y|} \to 0$; but the exponential map $t \to \mathrm{e}^t$ is continuous at the origin, so the result follows.

ii) Let us show

$$\lim_{(x,y)\to\infty} \frac{x^4+y^4}{x^2+y^2} = +\infty\,.$$

This descends from

$$\frac{x^4+y^4}{x^2+y^2} \geq \frac{1}{4}\|\boldsymbol{x}\|^2\,, \qquad \forall \boldsymbol{x} \in \mathbb{R}^2\,,\ \boldsymbol{x} \neq \boldsymbol{0} \tag{4.25}$$

and the Comparison Theorem, for $(x,y) \to \infty$ is clearly the same as $\|\boldsymbol{x}\| \to +\infty$. To get (4.25), note that if $x^2 \geq y^2$, we have

$$\|\boldsymbol{x}\|^2 = x^2+y^2 \leq 2x^2\,,$$

so

$$\frac{x^4+y^4}{x^2+y^2} \geq \frac{x^4}{2x^2} = \frac{1}{2}x^2 \geq \frac{1}{4}\|\boldsymbol{x}\|^2\,;$$

at the same result we arrive if $y^2 \geq x^2$. □

4.6 Curves in $\mathbb{R}^m$

A curve can describe the way the boundary of a plane region encloses the region itself – think of a polygon or an ellipse, or the trajectory of a point-particle moving in time under the effect of a force. Chapter 9 will provide us with a means of integrating along a curve, hence allowing us to formulate mathematically the physical notion of the work of a force.

Let us start with the definition of curve. Given a real interval I and a map $\boldsymbol{\gamma} : I \to \mathbb{R}^m$, we denote by $\boldsymbol{\gamma}(t) = \big(x_i(t)\big)_{1\leq i\leq m} = \sum_{i=1}^m x_i(t)\boldsymbol{e}_i \in \mathbb{R}^m$ the image of $t \in I$ under $\boldsymbol{\gamma}$.

Definition 4.32 *A continuous function* $\boldsymbol{\gamma} : I \subseteq \mathbb{R} \to \mathbb{R}^m$ *is called a* **curve**. *The set* $\Gamma = \boldsymbol{\gamma}(I) \subseteq \mathbb{R}^m$ *is said* **trace** *of the curve.*

If the trace of the curve lies on a plane, one speaks about a **plane curve**.

The most common curves are those in the plane ($m = 2$)

$$\boxed{\boldsymbol{\gamma}(t) = \big(x(t), y(t)\big) = x(t)\,\boldsymbol{i} + y(t)\,\boldsymbol{j}\,,}$$

and curves in space ($m = 3$)

$$\boxed{\boldsymbol{\gamma}(t) = \big(x(t), y(t), z(t)\big) = x(t)\,\boldsymbol{i} + y(t)\,\boldsymbol{j} + z(t)\,\boldsymbol{k}\,.}$$

We wish to stress the difference occurring between a curve, which is a function of one real variable, and its trace, a set in Euclidean space $\mathbb{R}^m$. A curve provides a way to parametrise its trace by associating to the parameter $t \in I$ one (and only one) point of the trace. Still, Γ may be the trace of many curves, because it may be parametrised in distinct ways. For instance the plane curves $\boldsymbol{\gamma}(t) = (t,t)$, $t \in [0,1]$, and $\boldsymbol{\delta}(t) = (t^2, t^2)$, $t \in [0,1]$ have the segment AB of end points $A = (0,0)$ and $B = (1,1)$ as common trace; $\boldsymbol{\gamma}$ and $\boldsymbol{\delta}$ are two parametrisations of the segment. The mid-point of AB, for example, corresponds to the value $t = \frac{1}{2}$ for $\boldsymbol{\gamma}$, $t = \frac{\sqrt{2}}{2}$ for $\boldsymbol{\delta}$.

A **curve** $\boldsymbol{\gamma}$ is said **simple** if $\boldsymbol{\gamma}$ is a one-to-one map, i.e., if distinct parameter's values determine distinct points of the trace.

If the interval $I = [a,b]$ is closed and bounded, as in the previous examples, the curve $\boldsymbol{\gamma}$ will be named **arc**. We call **end points** of the arc the points $P_0 = \boldsymbol{\gamma}(a)$, $P_1 = \boldsymbol{\gamma}(b)$; precisely, P_0 is the **initial point** and P_1 the **end point**, and $\boldsymbol{\gamma}$ **joins** P_0 and P_1. An **arc** is **closed** if its end points coincide: $\boldsymbol{\gamma}(a) = \boldsymbol{\gamma}(b)$; although a closed arc cannot be simple, one speaks anyway of a **simple closed arc** (or **Jordan arc**) when the point $\boldsymbol{\gamma}(a) = \boldsymbol{\gamma}(b)$ is the unique point on the trace where injectivity fails. Figure 4.13 illustrates a few arcs; in particular, the trace of the Jordan arc (bottom left) shows a core property of Jordan arcs, known as Jordan Curve Theorem.

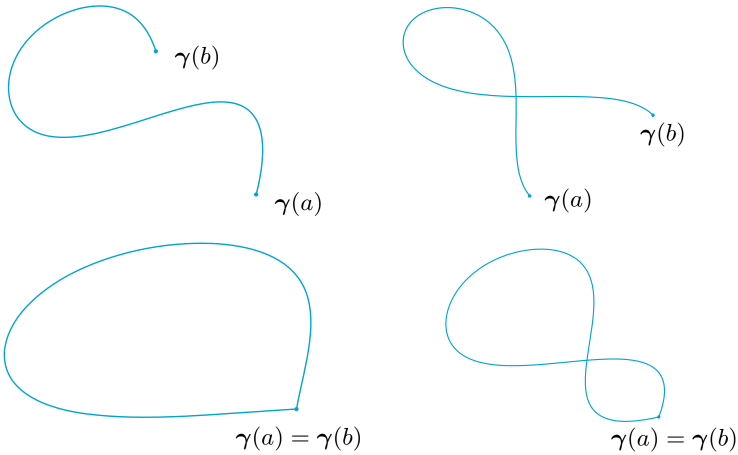


Figure 4.13. The trace $\Gamma = \boldsymbol{\gamma}([a,b])$ of a simple arc (top left), a non-simple arc (top right), a simple closed arc or Jordan arc (bottom left) and a closed non-simple arc (bottom right)

Theorem 4.33 *The trace Γ of a Jordan arc in the plane divides the plane in two connected components Σ_i and Σ_e with common boundary Γ, where Σ_i is bounded, Σ_e is unbounded.*
Conventionally, Σ_i is called the region "inside" Γ, while Σ_e is the region "outside" Γ.

Like for curves, there is a structural difference between an arc and its trace. It must be said that very often one uses the term 'arc' for a subset of Euclidean space $\mathbb{R}^m$ (as in: 'circular arc'), understanding the object as implicitly parametrised somehow, usually in the simplest and most natural way.

Examples 4.34

i) The simple plane curve

$$\boldsymbol{\gamma}(t) = (at+b, ct+d)\,, \qquad t \in \mathbb{R}\,,\ a \neq 0$$

has the line $y = \dfrac{c}{a}x + \dfrac{ad-bc}{a}$ as trace. Indeed, putting $x = x(t) = at+b$ and $y = y(t) = ct+d$ gives $t = \dfrac{x-b}{a}$, so

$$y = \frac{c}{a}(x-b) + d = \frac{c}{a}x + \frac{ad-bc}{a}\,.$$

ii) Let $\varphi : I \to \mathbb{R}$ be a continuous function on the interval I; the curve

$$\boldsymbol{\gamma}(t) = t\boldsymbol{i} + \varphi(t)\boldsymbol{j}\,, \qquad t \in I$$

has the graph of φ as trace.

iii) The trace of

$$\boldsymbol{\gamma}(t) = \big(x(t), y(t)\big) = (1+\cos t, 3+\sin t)\,, \qquad t \in [0, 2\pi]$$

is the circle with centre $(1,3)$ and radius 1, for in fact $\big(x(t)-1\big)^2 + \big(y(t)-3\big)^2 = \cos^2 t + \sin^2 t = 1$. The arc is simple and closed, and is the standard parametrisation of the circle starting at the point $(2,3)$ and running counter-clockwise.

More generally, the closed, simple arc

$$\boldsymbol{\gamma}(t) = \big(x(t), y(t)\big) = (x_0 + r\cos t, y_0 + r\sin t)\,, \qquad t \in [0, 2\pi]$$

has trace the circle centred at (x_0, y_0) with radius r.

If t varies in an interval $[0, 2k\pi]$, with $k \geq 2$ a positive integer, the curve has the same trace seen as a set; but because we wind around the centre k times, the curve is not simple.

Instead, if t varies in $[0, \pi]$, the curve is a circular arc, simple but not closed.

iv) Given $a, b > 0$, the map

$$\boldsymbol{\gamma}(t) = \big(x(t), y(t)\big) = (a\cos t, b\sin t)\,, \qquad t \in [0, 2\pi]$$

is a simple closed curve parametrising the ellipse with centre in the origin and semi-axes a and b:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1\,.$$

v) The trace of

$$\boldsymbol{\gamma}(t) = \big(x(t), y(t)\big) = (t\cos t, t\sin t) = t\cos t\,\boldsymbol{i} + t\sin t\,\boldsymbol{j}\,, \qquad t \in [0, +\infty]$$

is a spiral coiling counter-clockwise around the origin, see Fig. 4.14, left. Since the point $\boldsymbol{\gamma}(t)$ has distance $\sqrt{x^2(t)+y^2(t)} = t$ from the origin, so it moves farther afield as t grows, the spiral a simple curve.

vi) The simple curve

$$\boldsymbol{\gamma}(t) = \big(x(t), y(t), z(t)\big) = (\cos t, \sin t, t) = \cos t\,\boldsymbol{i} + \sin t\,\boldsymbol{j} + t\,\boldsymbol{k}\,, \qquad t \in \mathbb{R}$$

has the circular helix of Fig. 4.14, right, as trace. It rests on the infinite cylinder $\{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 = 1\}$ along the z-axis with radius 1.

vii) Let P and Q be distinct points of $\mathbb{R}^m$, with $m \geq 2$ arbitrary. The trace of the simple curve

$$\boldsymbol{\gamma}(t) = P + (Q - P)t\,, \qquad t \in \mathbb{R}$$

is the line through P and Q, because $\boldsymbol{\gamma}(0) = P$, $\boldsymbol{\gamma}(1) = Q$ and the vector $\boldsymbol{\gamma}(t) - P$ has constant direction, being parallel to $Q - P$.

There is a more general parametrisation of the same line

$$\boldsymbol{\gamma}(t) = P + (Q - P)\frac{t - t_0}{t_1 - t_0}\,, \qquad t \in \mathbb{R}\,, \tag{4.26}$$

with $t_0 \neq t_1$; in this case $\boldsymbol{\gamma}(t_0) = P$, $\boldsymbol{\gamma}(t_1) = Q$. □

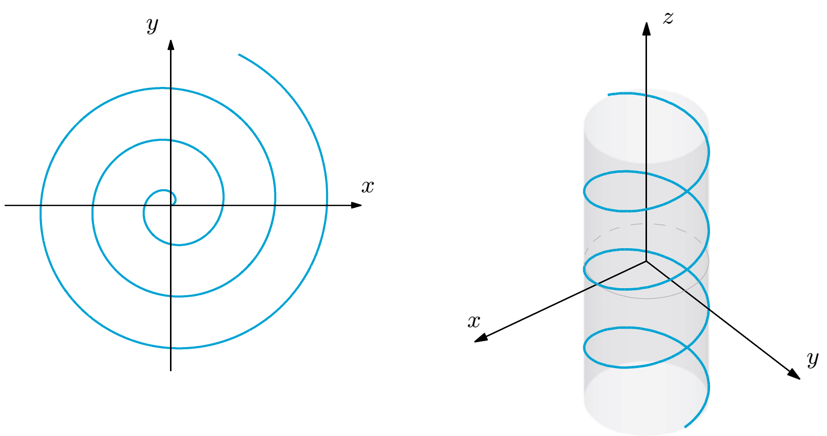


Figure 4.14. Spiral and circular helix, see Examples 4.34 v) and vi)

Some of the above curves have a special trace, given as the locus of points in the plane satisfying an equation of type

$$f(x, y) = c\,; \tag{4.27}$$

otherwise said, they parametrise level sets of f. With suitable assumptions on f, one can start from an implicit equation like (4.27) and obtain a curve $\boldsymbol{\gamma}(t) = \big(x(t), y(t)\big)$ of solutions. The details may be found in Sect. 7.2.

Remark 4.35 We posited in Sect. 4.5 that the existence of the limit of f, as $\boldsymbol{x}$ tends to $\boldsymbol{x}_0 \in \mathbb{R}^n$, is equivalent to the fact that the restrictions of f to the trace of any curve through $\boldsymbol{x}_0$ admit the same limit; namely, one can prove that

$$\boxed{\lim_{\boldsymbol{x}\to\boldsymbol{x}_0} f(\boldsymbol{x}) = \ell \qquad \Longleftrightarrow \qquad \lim_{t\to t_0} f\big(\boldsymbol{\gamma}(t)\big) = \ell}$$

for any curve $\boldsymbol{\gamma} : I \to \operatorname{dom} f$ such that $\boldsymbol{\gamma}(t_0) = \boldsymbol{x}_0$ for some $t_0 \in I$. □

Curves in polar, cylindrical and spherical coordinates

Representing a curve using Cartesian coordinates, as we have done so far, is but one possibility. Sometimes it may be better to use polar coordinates in dimension 2, and cylindrical or spherical coordinates in dimension 3 (these were introduced in Vol. I, Sect. 8.1, and will be discusses anew in Sect. 6.6.1).

A plane curve $\mathbb{R}^2$ can be defined by a continuous map $\boldsymbol{\gamma}_p : I \subseteq \mathbb{R} \to [0, +\infty)\times\mathbb{R}$, where $\boldsymbol{\gamma}_p(t) = \big(r(t), \theta(t)\big)$ are the polar coordinates of the point image of the value t of the parameter. It corresponds to the curve $\boldsymbol{\gamma}(t) = r(t)\cos\theta(t)\,\boldsymbol{i} + r(t)\sin\theta(t)\,\boldsymbol{j}$ in Cartesian coordinates. The spiral of Example 4.34 v) can be parametrised, in polar coordinates, by $\boldsymbol{\gamma}_p(t) = (t, t)$, $t \in [0, +\infty)$.

Similarly, we can represent a curve in space using cylindrical coordinates, $\boldsymbol{\gamma}_c(t) = \big(r(t), \theta(t), z(t)\big)$, with $\boldsymbol{\gamma}_c : I \subseteq \mathbb{R} \to [0, +\infty) \times \mathbb{R}^2$ continuous, or using spherical coordinates, $\boldsymbol{\gamma}_s(t) = \big(r(t), \varphi(t), \theta(t)\big)$, with $\boldsymbol{\gamma}_s : I \subseteq \mathbb{R} \to [0, +\infty) \times \mathbb{R}^2$ continuous. The circular helix of Example 4.34 vi) is $\boldsymbol{\gamma}_c(t) = (1, t, t)$, $t \in \mathbb{R}$, while a meridian arc of the unit sphere, joining the North and the South poles, is $\boldsymbol{\gamma}_s(t) = (1, t, \theta_0)$, for $t \in [0, \pi]$ and with given $\theta_0 \in [0, 2\pi]$.

4.7 Surfaces in $\mathbb{R}^3$

Surfaces are continuous functions defined on special subsets of $\mathbb{R}^2$, namely plane regions (see Definition 4.15); these regions play the role intervals had for curves.

Definition 4.36 *Let $\mathcal{R} \subseteq \mathbb{R}^2$ be a region. A continuous map $\boldsymbol{\sigma} : \mathcal{R} \to \mathbb{R}^3$ is said* **surface**, *and $\Sigma = \boldsymbol{\sigma}(\mathcal{R}) \subseteq \mathbb{R}^3$ is the* **trace** *of the surface.*

The independent variables in $\mathcal{R}$ are usually denoted (u, v), and

$$\boxed{\boldsymbol{\sigma}(u,v) = \big(x(u,v),\, y(u,v),\, z(u,v)\big) = x(u,v)\boldsymbol{i} + y(u,v)\boldsymbol{j} + z(u,v)\boldsymbol{k}}$$

is the Cartesian representation of $\boldsymbol{\sigma}$.

A surface is **simple** if the restriction of $\boldsymbol{\sigma}$ to the interior of $\mathcal{R}$ is one-to-one.

A surface is **compact** if the region $\mathcal{R}$ is compact. It is the 2-dimensional analogue of an arc.

Examples 4.37

i) Consider vectors $\boldsymbol{a}$, $\boldsymbol{b} \in \mathbb{R}^3$ such that $\boldsymbol{a} \wedge \boldsymbol{b} \neq \boldsymbol{0}$, and $\boldsymbol{c} \in \mathbb{R}^3$. The surface $\boldsymbol{\sigma} : \mathbb{R}^2 \to \mathbb{R}^3$

$$\begin{aligned}\boldsymbol{\sigma}(u,v) &= \boldsymbol{a}u + \boldsymbol{b}v + \boldsymbol{c}\\ &= (a_1u + b_1v + c_1)\boldsymbol{i} + (a_2u + b_2v + c_2)\boldsymbol{j} + (a_3u + b_3v + c_3)\boldsymbol{k}\end{aligned}$$

parametrises the plane Π passing through the point $\boldsymbol{c}$ and parallel to $\boldsymbol{a}$ and $\boldsymbol{b}$ (Fig. 4.15, left).

The Cartesian equation is found by setting $\boldsymbol{x} = (x, y, z) = \boldsymbol{\sigma}(u, v)$ and observing

$$\boldsymbol{x} - \boldsymbol{c} = \boldsymbol{a}u + \boldsymbol{b}v\,,$$

i.e., $\boldsymbol{x} - \boldsymbol{c}$ is a linear combination of $\boldsymbol{a}$ and $\boldsymbol{b}$. Hence by (4.7)

$$(\boldsymbol{a} \wedge \boldsymbol{b}) \cdot (\boldsymbol{x} - \boldsymbol{c}) = (\boldsymbol{a} \wedge \boldsymbol{b}) \cdot \boldsymbol{a}\, u + (\boldsymbol{a} \wedge \boldsymbol{b}) \cdot \boldsymbol{b}\, v = 0\,,$$

so

$$(\boldsymbol{a} \wedge \boldsymbol{b}) \cdot \boldsymbol{x} = (\boldsymbol{a} \wedge \boldsymbol{b}) \cdot \boldsymbol{c}\,.$$

The plane has thus equation

$$\alpha x + \beta y + \gamma z = \delta\,,$$

where α,β and γ are the components of $\boldsymbol{a} \wedge \boldsymbol{b}$, and $\delta = (\boldsymbol{a} \wedge \boldsymbol{b}) \cdot \boldsymbol{c}$.

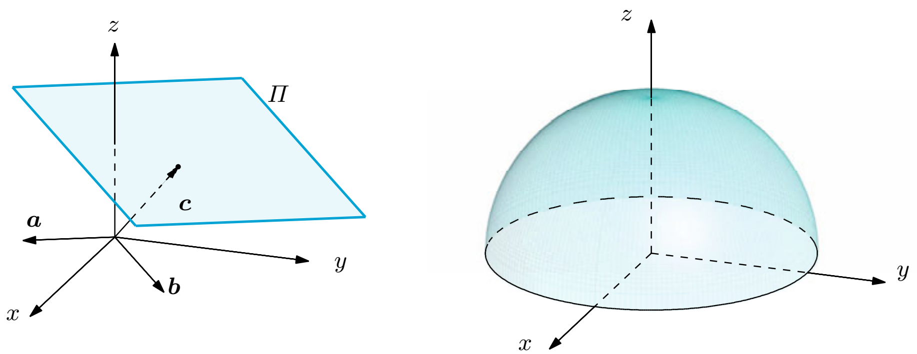


Figure 4.15. Representation of the plane Π (left) and the hemisphere (right) relative to Examples 4.37 i) and ii)

ii) Any continuous scalar function $\varphi : \mathcal{R} \to \mathbb{R}$ defined on a plane region produces a surface $\boldsymbol{\sigma} : \mathcal{R} \to \mathbb{R}^3$

$$\boxed{\boldsymbol{\sigma}(u,v) = u\boldsymbol{i} + v\boldsymbol{j} + \varphi(u,v)\boldsymbol{k}\,,}$$

whose trace is the graph of φ. Such a surface is sometimes referred to as **topographic** (with respect to the z-axis).
The surface $\boldsymbol{\sigma} : \mathcal{R} = \{(u,v) \in \mathbb{R}^2 : u^2 + v^2 \le 1\} \to \mathbb{R}^3$,

$$\boldsymbol{\sigma}(u,v) = u\boldsymbol{i} + v\boldsymbol{j} + \sqrt{1 - u^2 - v^2}\,\boldsymbol{k}$$

produces as trace the upper hemisphere centred at the origin and with unit radius (Fig. 4.15, right).
Permuting the components u, v, $\varphi(u,v)$ of $\boldsymbol{\sigma}$ clearly gives rise to surfaces whose equation is solved for x or y instead of z.

iii) The plane curve $\boldsymbol{\gamma} : I \subseteq \mathbb{R} \to \mathbb{R}^3$ given by $\boldsymbol{\gamma}(t) = \big(\gamma_1(t), 0, \gamma_3(t)\big)$, $\gamma_1(t) \ge 0$ for any $t \in I$ has trace Γ in the half-plane xz where $x \ge 0$.
Rotating Γ around the z-axis gives the trace Σ of the surface $\boldsymbol{\sigma} : I \times [0, 2\pi] \to \mathbb{R}^3$ defined by

$$\boxed{\boldsymbol{\sigma}(u,v) = \gamma_1(u)\cos v\,\boldsymbol{i} + \gamma_1(u)\sin v\,\boldsymbol{j} + \gamma_3(u)\boldsymbol{k}\,.}$$

One calls **surface of revolution** (around the z-axis) a surface thus obtained. The generating arc is said **meridian** (arc). For example, revolving around the z-axis the parabolic arc Γ parametrised by $\boldsymbol{\gamma} : [-1,1] \to \mathbb{R}^3$, $\boldsymbol{\gamma}(t) = (4 - t^2, 0, t)$ yields the lateral surface of a plinth, see Fig. 4.16, left.
Akin surfaces can be defined by revolution around the other coordinate axes.

iv) The vector-valued map

$$\boldsymbol{\sigma}(u,v) = (x_0 + r\sin u\cos v)\boldsymbol{i} + (y_0 + r\sin u\sin v)\boldsymbol{j} + (z_0 + r\cos u)\boldsymbol{k}$$

on $\mathcal{R} = [0,\pi] \times [0, 2\pi]$ is a compact surface with trace the spherical surface of centre $\boldsymbol{x}_0 = (x_0, y_0, z_0)$, radius $r > 0$. Slightly more generally, consider the (surface of the) ellipsoid, centred at $\boldsymbol{x}_0$ with semi-axes $a, b, c > 0$, defined by

$$\frac{(x-x_0)^2}{a^2} + \frac{(y-y_0)^2}{b^2} + \frac{(z-z_0)^2}{c^2} = 1\,;$$

This surface is parametrised by

$$\boldsymbol{\sigma}(u,v) = (x_0 + a\sin u\cos v)\boldsymbol{i} + (y_0 + b\sin u\sin v)\boldsymbol{j} + (z_0 + c\cos u)\boldsymbol{k}\,.$$

v) The function $\boldsymbol{\sigma} : \mathcal{R} \to \mathbb{R}^3$,

$$\boldsymbol{\sigma}(u,v) = u\cos v\,\boldsymbol{i} + u\sin v\boldsymbol{j} + v\boldsymbol{k}$$

with $\mathcal{R} = [0,1] \times [0, 4\pi]$, defines a compact surface; its trace is the ideal parking-lot ramp, see Fig. 4.16, right. The name **helicoid** is commonly used for this surface defined on $\mathcal{R} = [0,1] \times \mathbb{R}$.

All surfaces considered so far are simple. $\square$

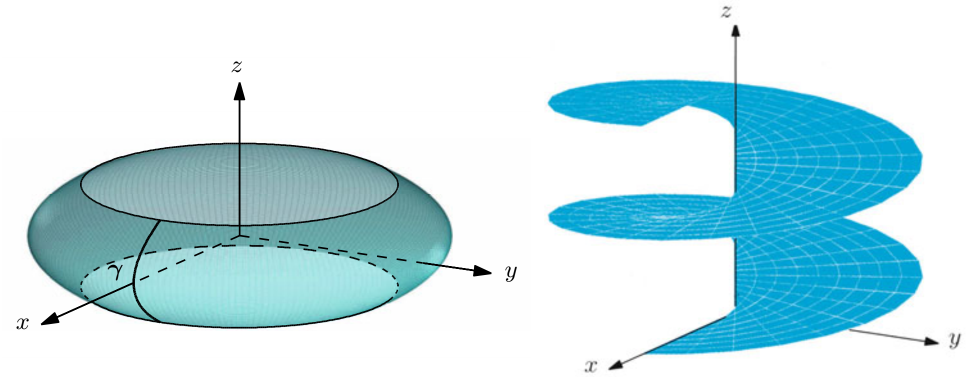


Figure 4.16. Plinth (left) and helicoid (right) relative to Examples 4.37 iii) and v)

We will see in Sect. 7.2 sufficient conditions for an implicit equation

$$f(x, y, z) = c$$

to be solved in one of the variables, i.e., to determine the trace of a surface which is locally a graph.

For surfaces, too, there is an alternative representation making use of cylindrical or spherical coordinates. The lateral surface of an infinite cylinder with axis z and radius 1, for instance, has cylindrical parametrisation $\boldsymbol{\sigma}_c : [0, 2\pi] \times \mathbb{R} \to \mathbb{R}^3$, $\boldsymbol{\sigma}_c(u, v) = \big(r(u, v), \theta(u, v), z(u, v)\big) = (1, u, v)$. Similarly, the unit sphere at the origin has a spherical parametrisation $\boldsymbol{\sigma}_s : [0, \pi] \times [0, 2\pi] \to \mathbb{R}^3$, $\boldsymbol{\sigma}_s(u, v) = \big(r(u, v), \varphi(u, v), \theta(u, v)\big) = (1, u, v)$. The surfaces of revolution of Example 4.37 iii) are $\boldsymbol{\sigma}_c : I \times [0, 2\pi] \to \mathbb{R}^3$, with $\boldsymbol{\sigma}_c(u, v) = \big(\gamma_1(u), v, \gamma_3(u)\big)$ in cylindrical coordinates.

4.8 Exercises

1. *Determine the interior, the closure and the boundary of the following sets. Say if the sets are open, closed, connected, convex or bounded:*

   a) $A = \{(x, y) \in \mathbb{R}^2 : 0 \le x \le 1,\ 0 < y < 1\}$

   b) $B = \big(B_2(\mathbf{0}) \setminus ([-1, 1] \times \{0\})\big) \cup \big((-1, 1) \times \{3\}\big)$

   c) $C = \{(x, y) \in \mathbb{R}^2 : |y| > 2\}$

2. *Determine the domain of the functions:*

   a) $f(x, y) = \dfrac{x - 3y + 7}{x - y^2}$

$\boxed{\text{b)}}$ $f(x,y) = \sqrt{1-3xy}$

$\boxed{\text{c)}}$ $f(x,y) = \sqrt{3x+y+1} - \dfrac{1}{\sqrt{2y-x}}$

$\boxed{\text{d)}}$ $f(x,y,z) = \log(x^2+y^2+z^2-9)$

e) $\boldsymbol{f}(x,y) = \left(\dfrac{y}{\sqrt{x^2+y^2}}, -\dfrac{x}{\sqrt{x^2+y^2}}\right)$

$\boxed{\text{f)}}$ $\boldsymbol{f}(x,y,z) = \left(\arctan y, \dfrac{\log z}{x}\right)$

g) $\gamma(t) = (t^3, \sqrt{t-1}, \sqrt{5-t})$

$\boxed{\text{h)}}$ $\gamma(t) = \left(\dfrac{t-2}{t+2}, \log(9-t^2)\right)$

$\boxed{\text{i)}}$ $\boldsymbol{\sigma}(u,v) = \left(\log(1-u^2-v^2), u, \dfrac{1}{u^2+v^2}\right)$

$\boxed{\ell)}$ $\boldsymbol{\sigma}(u,v) = \left(\sqrt{u+v}, \dfrac{2}{4-u^2-v^2}, v\right)$

3. *Say whether the following limits exist, and compute them:*

$\boxed{\text{a)}}$ $\displaystyle\lim_{(x,y)\to(0,0)} \frac{xy}{\sqrt{x^2+y^2}}$

$\boxed{\text{b)}}$ $\displaystyle\lim_{(x,y)\to(0,0)} \frac{xy}{x^2+y^2}$

$\boxed{\text{c)}}$ $\displaystyle\lim_{(x,y)\to(0,0)} \frac{x}{y}\log(1+x)$

d) $\displaystyle\lim_{(x,y)\to(0,0)} \frac{3x^2y}{x^2+y^2}$

e) $\displaystyle\lim_{(x,y)\to(0,0)} \frac{x^2y}{x^4+y^2}$

f) $\displaystyle\lim_{(x,y,z)\to(0,0,0)} \frac{xy+yz^2+xz^2}{x^2+y^2+z^4}$

$\boxed{\text{g)}}$ $\displaystyle\lim_{(x,y)\to(0,0)} \frac{x^2}{\sqrt{x^2+y^2}}$

h) $\displaystyle\lim_{(x,y)\to(0,0)} \frac{x^2-x^3+y^2+y^3}{x^2+y^2}$

$\boxed{\text{i)}}$ $\displaystyle\lim_{(x,y)\to(0,0)} \frac{x-y}{x+y}$

$\ell)$ $\displaystyle\lim_{(x,y)\to(0,0)} (x^2+y^2)\log(x^2+y^2)+5$

$\boxed{\text{m)}}$ $\displaystyle\lim_{(x,y)\to\infty} \frac{x^2+y+1}{x^2+y^4}$

$\boxed{\text{n)}}$ $\displaystyle\lim_{(x,y)\to\infty} \frac{\sqrt{1+3x^2+5y^2}}{x^2+y^2}$

4. *Determine the set on which the following maps are continuous:*

$\boxed{\text{a)}}$ $f(x,y) = \arcsin(xy-x-2y)$

b) $f(x,y,z) = \dfrac{xyz}{x^2+y^2-z}$

$\boxed{5.}$ *As the real number $\alpha \geq 0$ varies, study the continuity of:*

$$f(x,y) = \begin{cases} |x|^\alpha \dfrac{\sin y}{x^2+y^2} & \text{if } (x,y) \neq (0,0)\,, \\ 0 & \text{if } (x,y) = (0,0)\,. \end{cases}$$

$\boxed{6.}$ *The cylinders $z = x^2$ and $z = 4y^2$ intersect along the traces of two curves. Find a parametrisation of the one containing the point $(2, -1, 4)$.*

$\boxed{7.}$ *Parametrise (with a curve) the intersection of the plane $x + y + z = 1$ and the cylinder $z = x^2$.*

8. *Parametrise (with a curve) the intersection of $x^2 + y^2 = 16$ and $z = x + y$.*

9. *Find a Cartesian parametrisation for the curve $\boldsymbol{\gamma}(t) = \big(r(t), \theta(t)\big)$, $t \in [0, 2\pi]$, given in polar coordinates, when:*

$\boxed{\text{a)}}$ $\boldsymbol{\gamma}(t) = (\sin^2 t, t)$ $\qquad$ $\boxed{\text{b)}}$ $\boldsymbol{\gamma}(t) = \big(\sin \frac{t}{2}, t\big)$

10. *Eliminate the parameters u, v to obtain a Cartesian equation in x, y, z representing the trace of:*

$\boxed{\text{a)}}$ $\boldsymbol{\sigma}(u, v) = au \cos v\boldsymbol{i} + bu \sin v\boldsymbol{j} + u^2\boldsymbol{k}\,, \qquad a, b \in \mathbb{R}$
elliptic paraboloid

b) $\boldsymbol{\sigma}(u, v) = u\boldsymbol{i} + a \sin v\boldsymbol{j} + a \cos v\boldsymbol{k}\,, \qquad a \in \mathbb{R}$
cylinder

$\boxed{\text{c)}}$ $\boldsymbol{\sigma}(u, v) = (a + b \cos u) \sin v\boldsymbol{i} + (a + b \cos u) \cos v\boldsymbol{j} + b \sin u\boldsymbol{k}\,, \quad 0 < b < a$
torus

4.8.1 Solutions

1. *Properties of sets:*

a) The interior of A is

$$\mathring{A} = \{(x, y) \in \mathbb{R}^2 : 0 < x < 1,\ 0 < y < 1\}\,,$$

which is the open square $(0, 1)^2$. Its closure is

$$\overline{A} = \{(x, y) \in \mathbb{R}^2 : 0 \le x \le 1,\ 0 \le y \le 1\}\,,$$

i.e., the closed square $[0, 1]^2$. The boundary is

$$\begin{aligned}\partial A = \{(x, y) \in \mathbb{R}^2 : x = 0 \text{ or } x = 1,\ 0 \le y \le 1\} \cup \\ \cup\{(x, y) \in \mathbb{R}^2 : y = 0 \text{ or } y = 1,\ 0 \le x \le 1\}\,,\end{aligned}$$

which is the union of the four sides (perimeter) of $[0, 1]^2$.
Then A is neither open nor closed, but connected, convex and bounded, see Fig. 4.17, left.

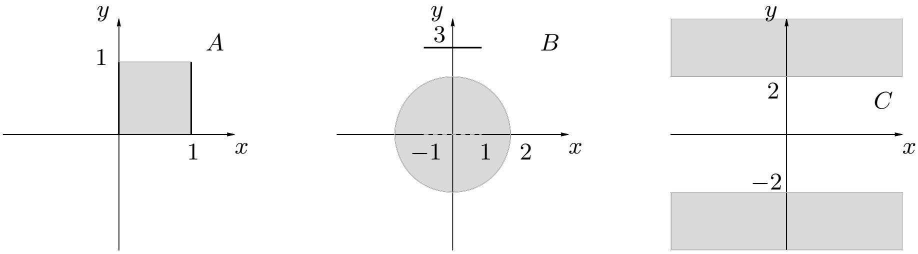


Figure 4.17. The sets A, B, C of Exercise 1

b) The interior of B is

$$\overset{\circ}{B} = B_2(\mathbf{0}) \setminus ([-1,1] \times \{0\}),$$

the closure

$$\overline{B} = \overline{B_2(\mathbf{0})} \cup \big([-1,1] \times \{3\}\big),$$

the boundary

$$\partial B = \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 = 2\} \cup ([-1,1] \times \{0\}) \cup ([-1,1] \times \{3\}).$$

Hence B is not open, closed, connected, nor convex; but it is bounded, see Fig. 4.17, middle.

c) The interior of C is C itself; its closure is

$$\overline{C} = \{(x,y) \in \mathbb{R}^2 : |y| \geq 2\},$$

while the boundary is

$$\partial C = \{(x,y) \in \mathbb{R}^2 : y = \pm 2\},$$

making C open, not connected, nor convex, nor bounded (Fig. 4.17, right).

2. *Domain of functions:*

a) The domain is $\{(x,y) \in \mathbb{R}^2 : x \neq y^2\}$, the set of all points in the plane except those on the parabola $x = y^2$.

b) The map is defined where the radicand is ≥ 0, so the domain is

$$\{(x,y) \in \mathbb{R}^2 : y \leq \frac{1}{3x} \text{ if } x > 0,\ y \geq \frac{1}{3x} \text{ if } x < 0,\ y \in \mathbb{R} \text{ if } x = 0\},$$

which describes the points lying between the two branches of the hyperbola $y = \frac{1}{3x}$.

c) The map is defined for $3x + y + 1 \geq 0$ and $2y - x > 0$, implying the domain is

$$\{(x,y) \in \mathbb{R}^2 : y \geq -3x - 1\} \cap \{(x,y) \in \mathbb{R}^2 : y > \frac{x}{2}\}.$$

See Fig. 4.18.

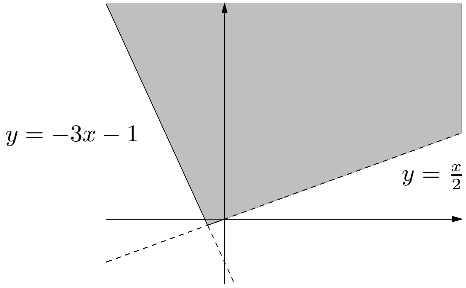


Figure 4.18. The domain of the function f of Exercise 2. c)

d) The function's domain is defined by the positivity of the log's argument, whence

$$\{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + x^2 > 9\} .$$

These are the points of the plane outside the sphere at the origin of radius 3.

e) $\operatorname{dom} \boldsymbol{f} = \mathbb{R}^2 \setminus \{\mathbf{0}\}$.

f) The map is defined for $z > 0$ and $x \neq 0$; there are no constraints on y. The domain is thus the subset of $\mathbb{R}^3$ given by the half-space $z > 0$ without the half-plane $x = 0$.

g) $\operatorname{dom} \boldsymbol{\gamma} = I = [1, 5]$.

h) The components $x(t) = \dfrac{t-2}{t+2}$ and $y(t) = \log(9 - t^2)$ of the curve are well defined for $t \neq -2$ and $t \in (-3, 3)$ respectively. Therefore $\boldsymbol{\gamma}$ is defined on the intervals $I_1 = (-3, -2)$ and $I_2 = (-2, 3)$.

i) The surface is defined for $0 < u^2 + v^2 < 1$, i.e., for points of the punctured open unit disc on the plane uv.

ℓ) The domain of $\boldsymbol{\sigma}$ is $\operatorname{dom} \boldsymbol{\sigma} = \{(u, v) \in \mathbb{R}^2 : u + v \geq 0, u^2 + v^2 \neq 4\}$. See Fig. 4.19 for a picture.

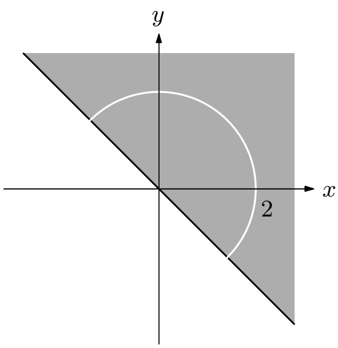


Figure 4.19. The domain of the surface $\boldsymbol{\sigma}$ of Exercise 2. ℓ)

3. *Limits:*

a) From
$$|x| = \sqrt{x^2} \le \sqrt{x^2+y^2}$$
follows
$$\frac{|x|}{\sqrt{x^2+y^2}} \le 1\,.$$
Hence
$$|f(x,y)| = \frac{|x|\,|y|}{\sqrt{x^2+y^2}} \le |y|\,,$$
and by the Squeeze Rule we conclude the limit is 0.

b) If $y=0$ or $x=0$ the map $f(x,y)=\dfrac{xy}{x^2+y^2}$ is identically zero. When computing limits along the axes, then, the result is 0:
$$\lim_{x\to 0} f(x,0) = \lim_{y\to 0} f(0,y) = 0\,.$$
But along the line $y=x$ the function equals $\frac{1}{2}$, so
$$\lim_{x\to 0} f(x,x) = \lim_{x\to 0} \frac{x^2}{x^2+x^2} = \frac{1}{2}\,.$$
In conclusion, the limit does not exist.

c) Let us compute the function $f(x,y)=\dfrac{x}{y}\log(1+x)$ along the lines $y=kx$ with $k\neq 0$; from
$$f(x,kx) = \frac{1}{k}\log(1+x)$$
follows
$$\lim_{x\to 0} f(x,kx) = 0\,.$$
Along the parabola $y=x^2$ though, $f(x,x^2)=\frac{1}{x}\log(1+x)$, i.e.,
$$\lim_{x\to 0} f(x,x^2) = 1\,.$$
The limit does not exist.

d) $0\,.$

e) Does not exist.

f) Does not exist.

g) Since
$$|f(r\cos\theta, r\sin\theta)| = \frac{r^2\cos^2\theta}{r} = r\cos^2\theta \le r\,,$$
Proposition 4.28, with $g(r)=r$, implies the limit is 0.

h) 1.

i) From
$$f(r\cos\theta, r\sin\theta) = \frac{\cos\theta - \sin\theta}{\cos\theta + \sin\theta}$$
follows
$$\lim_{r\to 0} f(r\cos\theta, r\sin\theta) = \frac{\cos\theta - \sin\theta}{\cos\theta + \sin\theta};$$
thus the limit does not exist.

ℓ) 5.

m) Set $f(x,y) = \dfrac{x^2+y+1}{x^2+y^4}$, so that
$$f(x,0) = \frac{x^2+1}{x^2} \qquad \text{and} \qquad f(0,y) = \frac{y+1}{y^4}.$$
Hence
$$\lim_{x\to\pm\infty} f(x,0) = 1 \qquad \text{and} \qquad \lim_{y\to\pm\infty} f(0,y) = 0$$
and we conclude the limit does not exist.

n) Note
$$0 \le \frac{\sqrt{1+3x^2+5y^2}}{x^2+y^2} \le \frac{\sqrt{1+5(x^2+y^2)}}{x^2+y^2}, \qquad \forall (x,y) \ne (0,0).$$
Set $t = x^2+y^2$, so that
$$\lim_{(x,y)\to\boldsymbol{\infty}} \frac{\sqrt{1+5(x^2+y^2)}}{x^2+y^2} = \lim_{t\to+\infty} \frac{\sqrt{1+5t}}{t} = 0$$
and
$$0 \le \lim_{(x,y)\to\boldsymbol{\infty}} \frac{\sqrt{1+3x^2+5y^2}}{x^2+y^2} \le \lim_{t\to+\infty} \frac{\sqrt{1+5t}}{t} = 0.$$
The required limit is 0.

4. *Continuity sets:*

a) The function is continuous on its domain as composite map of continuous functions. For the domain, recall that arcsin is defined when the argument lies between -1 and 1, so
$$\operatorname{dom} f = \{(x,y) \in \mathbb{R}^2 : -1 \le xy - x - 2y \le 1\}.$$
Let us draw such set. The points of the line $x = 2$ do not belong to $\operatorname{dom} f$. If $x > 2$, the condition $-1 + x \le (x-2)y \le 1 + x$ is equivalent to
$$1 + \frac{1}{x-2} = \frac{x-1}{x-2} \le y \le \frac{x+1}{x-2} = 1 + \frac{3}{x-2};$$

this means dom f contains all points lying between the graphs of the two hyperbolas $y = 1 + \dfrac{1}{x-2}$ and $y = 1 + \dfrac{3}{x-2}$. Similarly, if $x < 2$, the domain is

$$1 + \frac{3}{x-2} \le y \le 1 + \frac{1}{x-2}\,.$$

Overall, dom f is as in Fig. 4.20.

b) The continuity set is $C = \{(x, y, z) \in \mathbb{R}^3 : z \neq x^2 + y^2\}$.

5. Let us compute $\lim\limits_{(x,y)\to(0,0)} f(x, y)$. Using polar coordinates,

$$f(r\cos\theta, r\sin\theta) = r^{\alpha-2}|\cos\theta|^{\alpha}\,\sin(r\sin\theta)\,;$$

but $|\sin t| \le |t|$, valid for any $t \in \mathbb{R}$, implies

$$|f(r\cos\theta, r\sin\theta)| \le r^{\alpha-1}|\cos\theta|^{\alpha}|\sin\theta|\,,$$

so the limit is zero if $\alpha > 1$, and does not exist if $\alpha \le 1$. Therefore if $\alpha > 1$ the map is continuous on $\mathbb{R}^2$, if $0 \le \alpha \le 1$ it is continuos only on $\mathbb{R}^2 \setminus \{\mathbf{0}\}$.

6. The system

$$\begin{cases} z = x^2 \\ z = 4y^2 \end{cases}$$

gives $x^2 = 4y^2$. The cylinders' intersection projects onto the plane xy as the two lines $x = \pm 2y$ (Fig. 4.21). As we are looking for the branch containing $(2, -1, 4)$, we choose $x = -2y$. One possible parametrisation is given by $t = y$, hence

$$\boldsymbol{\gamma}(t) = (-2t, t, 4t^2)\,, \qquad t \in \mathbb{R}\,.$$

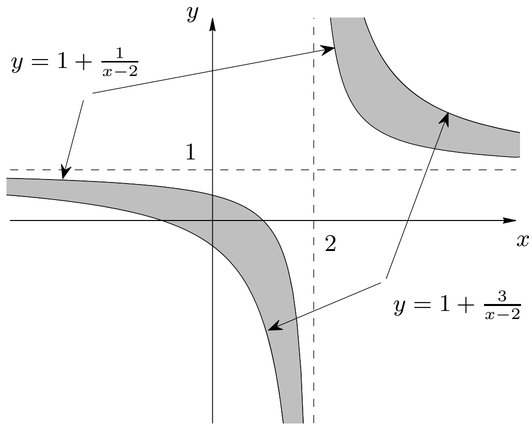


Figure 4.20. The continuity set of $f(x, y) = \arcsin(xy - x - 2y)$

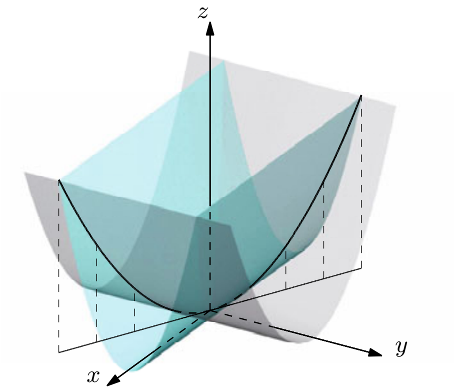


Figure 4.21. The intersection of the cylinders $z = x^2$ and $z = 4y^2$

7. We have $x + y + x^2 = 1$, and choosing $t = x$ as parameter,

$$\boldsymbol{\gamma}(t) = (t, 1 - t - t^2, t^2)\,, \qquad t \in \mathbb{R}\,.$$

8. Use cylindrical coordinates with $x = 4\cos t$, $y = 4\sin t$. Then

$$\boldsymbol{\gamma}(t) = (4\cos t, 4\sin t, 4(\cos t + \sin t))\,, \qquad t \in [0, 2\pi]\,.$$

9. *Curves in Cartesian coordinates:*

a) We have $r(t) = \sin^2 t$ and $\theta(t) = t$; recalling that $x = r\cos\theta$ and $y = r\sin\theta$, gives

$$x = \sin^2 t \cos t \qquad \text{and} \qquad y = \sin^3 t\,.$$

In Cartesian coordinates then, $\boldsymbol{\gamma} : [0, 2\pi] \to \mathbb{R}^2$ can be written $\boldsymbol{\gamma}(t) = \big(x(t), y(t)\big) = \big(\sin^2 t \cos t, \sin^3 t\big)$, see Fig. 4.22, left.

b) We have $\boldsymbol{\gamma}(t) = \big(\sin\frac{t}{2}\cos t, \sin\frac{t}{2}\sin t\big)$. The curve is called cardioid, see Fig. 4.22, right.

10. Graphs:

a) It is straightforward to see

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = u^2\cos^2 v + u^2\sin^2 v = u^2 = z,$$

giving the equation of an elliptic paraboloid.

b) $y^2 + z^2 = a^2\,.$

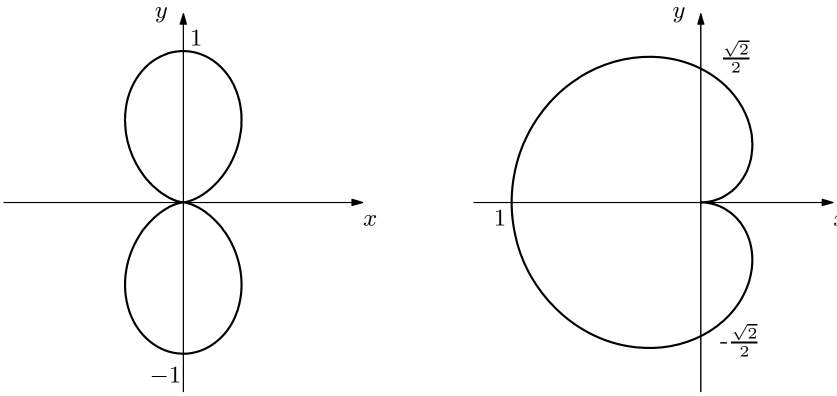


Figure 4.22. Traces of the curves of Exercise 9. a) (left) and 9. b) (right)

c) Note that $x^2+y^2 = (a+b\cos u)^2$ and $a+b\cos u > 0$; therefore $\sqrt{x^2+y^2}-a = b\cos u$. Moreover,

$$\big(\sqrt{x^2+y^2}-a\big)^2 + z^2 = b^2\cos^2 u + b^2\sin^2 u = b^2\,;$$

in conclusion

$$\big(\sqrt{x^2+y^2}-a\big)^2 + z^2 = b^2\,.$$

5

Differential calculus for scalar functions

While the previous chapter dealt with the continuity of multivariable functions and their limits, the next three are dedicated to differential calculus. We start in this chapter by discussing scalar functions.

The power of differential tools should be familiar to students from the first Calculus course; knowing the derivatives of a function of one variable allows to capture the function's global behaviour, on intervals in the domain, as well as the local one, on arbitrarily small neighbourhoods. In passing to higher dimension, there are more instruments available that adapt to a more varied range of possibilities. If on one hand certain facts attract less interest (drawing graphs for instance, or understanding monotonicity, which is tightly related to the ordering of the reals, not present any more), on the other new aspects come into play (from Linear Algebra in particular) and become central. The first derivative in one variable is replaced by the gradient vector field, and the Hessian matrix takes the place of the second derivative. Due to the presence of more variables, some notions require special attention (differentiability at a point is a more delicate issue now), whereas others (convexity, Taylor expansions) translate directly from one to several variables. The study of so-called unconstrained extrema of a function of several real variables carries over effortlessly, thus generalising the known Fermat and Weierstrass' Theorems, and bringing to the fore a new kind of stationary points, like saddle points, at the same time.

5.1 First partial derivatives and gradient

The simplest case where partial derivatives at a point can be seen is on the plane, i.e., in dimension two, which we start from.

Let $f : \operatorname{dom} f \subseteq \mathbb{R}^2 \to \mathbb{R}$ be a function of two variables defined in a neighbourhood of the point $\boldsymbol{x}_0 = (x_0, y_0)$. The map $x \mapsto f(x, y_0)$, obtained by fixing the second variable to a constant, is a real-valued map of one real variable, defined

C. Canuto, A. Tabacco: *Mathematical Analysis II*, 2nd Ed.,
UNITEXT – La Matematica per il 3+2 85, DOI 10.1007/978-3-319-12757-6_5,

around the point $x_0 \in \mathbb{R}$. If this is differentiable at x_0, one says f admits **partial derivative with respect to** x at $\boldsymbol{x}_0$ and sets

$$\boxed{\frac{\partial f}{\partial x}(\boldsymbol{x}_0) = \frac{\mathrm{d}}{\mathrm{d}x} f(x, y_0)\Big|_{x=x_0} = \lim_{x \to x_0} \frac{f(x, y_0) - f(x_0, y_0)}{x - x_0}\,.}$$

Similarly, if $y \mapsto f(x_0, y)$ is differentiable at y_0, f is said to admit **partial derivative with respect to** y at $\boldsymbol{x}_0$ and one defines

$$\boxed{\frac{\partial f}{\partial y}(\boldsymbol{x}_0) = \frac{\mathrm{d}}{\mathrm{d}y} f(x_0, y)\Big|_{y=y_0} = \lim_{y \to y_0} \frac{f(x_0, y) - f(x_0, y_0)}{y - y_0}\,.}$$

The geometric meaning of partial derivatives is explained in Fig. 5.1.

This can be generalised to functions of n variables, $n \geq 3$, in the most obvious way. Precisely, let a map in n variables $f : \operatorname{dom} f \subseteq \mathbb{R}^n \to \mathbb{R}$ be defined on the neighbourhood of $\boldsymbol{x}_0 = (x_{01}, \ldots, x_{0n}) = \sum_{i=1}^{n} x_{0i}\boldsymbol{e}_i$, where $\boldsymbol{e}_i$ is the ith unit vector of the canonical basis of $\mathbb{R}^n$ seen in (4.1). One says f admits **partial derivative with respect to** x_i at $\boldsymbol{x}_0$ if the function of one real variable

$$x \mapsto f(x_{01}, \ldots, x_{0,i-1}, x, x_{0,i+1}, \ldots, x_{0n})\,,$$

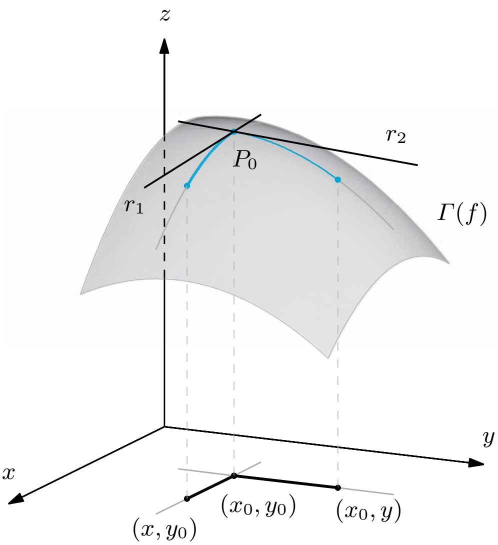


Figure 5.1. The partial derivatives of f at $\boldsymbol{x}_0$ are the slopes of the lines r_1, r_2, tangent to the graph of f at P_0

obtained by making all independent variables but the ith one constant, is differentiable at $x = x_{0i}$. If so, one defines the symbol

$$\begin{aligned}\frac{\partial f}{\partial x_i}(\boldsymbol{x}_0) &= \frac{\mathrm{d}}{\mathrm{d}x} f(x_{01}, \ldots, x_{0,i-1}, x, x_{0,i+1}, \ldots, x_{0n})\Big|_{x=x_{0i}} \\ &= \lim_{\Delta x \to 0} \frac{f(\boldsymbol{x}_0 + \Delta x\, \boldsymbol{e}_i) - f(\boldsymbol{x}_0)}{\Delta x}.\end{aligned} \tag{5.1}$$

The partial derivative of f at $\boldsymbol{x}_0$ with respect to the variable x_i is also denoted as follows

$$\mathrm{D}_{x_i} f(\boldsymbol{x}_0), \qquad \mathrm{D}_i f(\boldsymbol{x}_0), \qquad f_{x_i}(\boldsymbol{x}_0)$$

(or simply $f_i(\boldsymbol{x}_0)$, if no confusion arises).

Definition 5.1 *Assume f admits partial derivatives at $\boldsymbol{x}_0$ with respect to all variables. The* **gradient** $\nabla f(\boldsymbol{x}_0)$ *of f at $\boldsymbol{x}_0$ is the vector defined by*

$$\nabla f(\boldsymbol{x}_0) = \Big(\frac{\partial f}{\partial x_i}(\boldsymbol{x}_0)\Big)_{1 \le i \le n} = \frac{\partial f}{\partial x_1}(\boldsymbol{x}_0)\, \boldsymbol{e}_1 + \ldots + \frac{\partial f}{\partial x_n}(\boldsymbol{x}_0)\, \boldsymbol{e}_n \in \mathbb{R}^n.$$

Another notation for it is **grad** $f(\boldsymbol{x}_0)$.

We remind that, as any vector in $\mathbb{R}^n$, $\nabla f(\boldsymbol{x}_0)$ can be written both as row or column vector, according to need.

Examples 5.2

i) Let $f(x,y) = \sqrt{x^2+y^2}$ be the function 'distance from the origin'. At the point $\boldsymbol{x}_0 = (2,-1)$,

$$\frac{\partial f}{\partial x}(\boldsymbol{x}_0) = \frac{\mathrm{d}}{\mathrm{d}x}\sqrt{x^2+1}\Big|_{x=2} = \frac{x}{\sqrt{x^2+1}}\Big|_{x=2} = \frac{2}{\sqrt{5}},$$

$$\frac{\partial f}{\partial y}(\boldsymbol{x}_0) = \frac{\mathrm{d}}{\mathrm{d}y}\sqrt{4+y^2}\Big|_{y=-1} = \frac{y}{\sqrt{4+y^2}}\Big|_{y=-1} = -\frac{1}{\sqrt{5}}.$$

Therefore

$$\nabla f(\boldsymbol{x}_0) = \Big(\frac{2}{\sqrt{5}}, -\frac{1}{\sqrt{5}}\Big) = \frac{1}{\sqrt{5}}(2,-1).$$

ii) For $f(x,y,z) = y \log(2x - 3z)$, at $\boldsymbol{x}_0 = (2,3,1)$ we have

$$\frac{\partial f}{\partial x}(\boldsymbol{x}_0) = \frac{\mathrm{d}}{\mathrm{d}x} 3\log(2x-3)\Big|_{x=2} = 3\,\frac{2}{2x-3}\Big|_{x=2} = 6,$$

$$\frac{\partial f}{\partial y}(\boldsymbol{x}_0) = \frac{\mathrm{d}}{\mathrm{d}y}\, y \log 1\Big|_{y=3} = 0,$$

$$\frac{\partial f}{\partial z}(\boldsymbol{x}_0) = \frac{\mathrm{d}}{\mathrm{d}z} 3\log(4-3z)\Big|_{z=1} = 3\,\frac{-3}{4-3z}\Big|_{z=1} = -9\,,$$

whence

$$\nabla f(\boldsymbol{x}_0) = (6, 0, -9)\,.$$

iii) Let $f : \mathbb{R}^n \to \mathbb{R}$ be the affine map $f(\boldsymbol{x}) = \boldsymbol{a}\cdot\boldsymbol{x} + b$, $\boldsymbol{a} \in \mathbb{R}^n$, $b \in \mathbb{R}$. At any point $\boldsymbol{x}_0 \in \mathbb{R}^n$,

$$\nabla f(\boldsymbol{x}_0) = \boldsymbol{a}\,.$$

iv) Take a function f not depending on the variable x_i on a neighbourhood of $\boldsymbol{x}_0 \in \operatorname{dom} f$. Then $\dfrac{\partial f}{\partial x_i}(\boldsymbol{x}_0) = 0$.
In particular, f constant around $\boldsymbol{x}_0$ implies $\nabla f(\boldsymbol{x}_0) = \boldsymbol{0}$. □

The function

$$\boxed{\frac{\partial f}{\partial x_i} : \boldsymbol{x} \mapsto \frac{\partial f}{\partial x_i}(\boldsymbol{x})\,,}$$

defined on a suitable subset $\operatorname{dom}\dfrac{\partial f}{\partial x_i} \subseteq \operatorname{dom} f \subseteq \mathbb{R}^n$ and with values in $\mathbb{R}$, is called (first) **partial derivative** of f with respect to x_i. The **gradient function** of f,

$$\boxed{\nabla f : \boldsymbol{x} \mapsto \nabla f(\boldsymbol{x}),}$$

whose domain $\operatorname{dom}\nabla f$ is the intersection of the domains of the single first partial derivatives, is an example of a *vector field*, being a function defined on a subset of $\mathbb{R}^n$ with values in $\mathbb{R}^n$.

In practice, each partial derivative $\dfrac{\partial f}{\partial x_i}$ is computed by freezing all variables of f different from x_i (taking them as constants), and differentiating in the only one left x_i. We can then use on this function everything we know from Calculus 1.

Examples 5.3

We shall use the previous examples.

i) For $f(x,y) = \sqrt{x^2+y^2}$ we have

$$\nabla f(\boldsymbol{x}) = \left(\frac{x}{\sqrt{x^2+y^2}}, \frac{y}{\sqrt{x^2+y^2}}\right) = \frac{\boldsymbol{x}}{\|\boldsymbol{x}\|}$$

with $\operatorname{dom}\nabla f = \mathbb{R}^2 \setminus \{\boldsymbol{0}\}$. The formula holds in any dimension n if $f(\boldsymbol{x}) = \|\boldsymbol{x}\|$ is the norm function on $\mathbb{R}^n$.

ii) For $f(x,y,z) = y\log(2x-3z)$ we obtain

$$\nabla f(\boldsymbol{x}) = \left(\frac{2y}{2x-3z}, \log(2x-3z), \frac{-3y}{2x-3z}\right)\,,$$

with $\operatorname{dom}\nabla f = \operatorname{dom} f = \{(x,y,z) \in \mathbb{R}^3 : 2x-3z > 0\}$.

iii) The total resistance R produced by three conductors with resistances R_1, R_2, R_3 in a parallel circuit is given by the formula:

$$\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3}\,.$$

We want to compute the partial derivative of R with respect to one of the R_i, say R_1. As

$$R(R_1, R_2, R_3) = \frac{1}{\frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3}} = \frac{R_1 R_2 R_3}{R_2 R_3 + R_1 R_3 + R_1 R_2}\,,$$

we have

$$\frac{\partial R}{\partial R_1}(R_1, R_2, R_3) = \frac{R_2^2 R_3^2}{(R_2 R_3 + R_1 R_3 + R_1 R_2)^2}\,. \qquad \square$$

Partial derivatives with respect to the x_i, $i = 1, \ldots, n$, are special cases of the directional derivative along a vector. Let f be a map defined around a point $\boldsymbol{x}_0 \in \mathbb{R}^n$ and let $\boldsymbol{v} \in \mathbb{R}^n$ be a given non-zero vector. Then f admits **partial derivative**, or **directional derivative**, **along $\boldsymbol{v}$** at $\boldsymbol{x}_0$ if

$$\boxed{\frac{\partial f}{\partial \boldsymbol{v}}(\boldsymbol{x}_0) = \lim_{t \to 0} \frac{f(\boldsymbol{x}_0 + t\boldsymbol{v}) - f(\boldsymbol{x}_0)}{t}} \qquad (5.2)$$

exists and is finite.

Another notation for this is $\mathrm{D}_{\boldsymbol{v}} f(\boldsymbol{x}_0)$. The condition spells out the fact that the map $t \mapsto f(\boldsymbol{x}_0 + t\boldsymbol{v})$ is differentiable at $t_0 = 0$ (the latter is defined in a neighbourhood of $t_0 = 0$, since $\boldsymbol{x}_0 + t\boldsymbol{v}$ belongs to the neighbourhood of $\boldsymbol{x}_0$ where f is defined, for t small enough). See Fig. 5.2 to interpret geometrically the directional derivative.

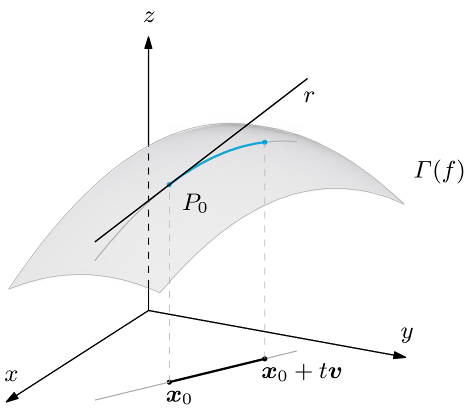


Figure 5.2. The directional derivative is the slope of the line r tangent to the graph of f at P_0

The directional derivative of f at $\boldsymbol{x}_0$ with respect to x_i is obtained by choosing $\boldsymbol{v} = \boldsymbol{e}_i$; thus

$$\frac{\partial f}{\partial \boldsymbol{e}_i}(\boldsymbol{x}_0) = \frac{\partial f}{\partial x_i}(\boldsymbol{x}_0), \qquad i = 1, \ldots, n,$$

as is immediate by comparing (5.2) with (5.1), using $\Delta x = t$.

In the next section we discuss the relationship between the gradient and directional derivatives of differentiable functions.

5.2 Differentiability and differentials

Recall from Vol. I, Sect. 6.6 that a real map of one real variable f, differentiable at $x_0 \in \mathbb{R}$, satisfies the first formula of the finite increment

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + o(x - x_0), \qquad x \to x_0. \tag{5.3}$$

This is actually equivalent to differentiability at x_0, because if there is a number $a \in \mathbb{R}$ such that

$$f(x) = f(x_0) + a\,(x - x_0) + o(x - x_0), \qquad x \to x_0,$$

necessarily f is differentiable at x_0, and $a = f'(x_0)$. From the geometric viewpoint, furthermore, differentiability at x_0 amounts to the existence of the tangent line to the graph of f at the point $P_0 = \big(x_0, f(x_0)\big)$:

$$y = t(x) = f(x_0) + f'(x_0)(x - x_0).$$

In presence of several variables, the picture is more involved and the existence of the gradient of f at $\boldsymbol{x}_0$ does not guarantee the validity of a formula like (5.3), e.g.,

$$f(\boldsymbol{x}) = f(\boldsymbol{x}_0) + \nabla f(\boldsymbol{x}_0) \cdot (\boldsymbol{x} - \boldsymbol{x}_0) + o(\|\boldsymbol{x} - \boldsymbol{x}_0\|), \qquad \boldsymbol{x} \to \boldsymbol{x}_0, \tag{5.4}$$

nor the existence of the tangent plane (or hyperplane, if $n > 2$) to the graph of f at $P_0 = \big(\boldsymbol{x}_0, f(\boldsymbol{x}_0)\big) \in \mathbb{R}^{n+1}$. Consider for example the function

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

The map is identically zero on the coordinate axes ($f(x,0) = f(0,y) = 0$ for any $x, y \in \mathbb{R}$), so

$$\frac{\partial f}{\partial x}(0,0) = \frac{\partial f}{\partial y}(0,0) = 0, \qquad \text{i.e.,} \qquad \nabla f(0,0) = (0,0).$$

If we had (5.4), at $\boldsymbol{x}_0 = (0,0)$ we would find

$$\lim_{(x,y)\to(0,0)} \frac{x^2 y}{(x^2+y^2)\sqrt{x^2+y^2}} = 0\,;$$

but moving along the line $y = x$, for instance, we have

$$\lim_{x\to 0^\pm} \frac{x^3}{2\sqrt{2}x^2|x|} = \pm\frac{1}{2\sqrt{2}} \neq 0\,.$$

Not even the existence of *every* partial derivative at $\boldsymbol{x}_0$ warrants (5.4) will hold. It is easy to see that the above f has directional derivatives along any given vector $\boldsymbol{v}$.

Thus it makes sense to introduce the following definition.

Definition 5.4 *A function f is* **differentiable** *at an interior point $\boldsymbol{x}_0$ of the domain, if $\nabla f(\boldsymbol{x}_0)$ exists and the following formula holds:*

$$f(\boldsymbol{x}) = f(\boldsymbol{x}_0) + \nabla f(\boldsymbol{x}_0)\cdot(\boldsymbol{x}-\boldsymbol{x}_0) + o(\|\boldsymbol{x}-\boldsymbol{x}_0\|)\,, \qquad \boldsymbol{x}\to\boldsymbol{x}_0\,. \tag{5.5}$$

In the case $n = 2$,

$$z = f(\boldsymbol{x}_0) + \nabla f(\boldsymbol{x}_0)\cdot(\boldsymbol{x}-\boldsymbol{x}_0)\,, \tag{5.6}$$

i.e.,

$$\boxed{z = f(x_0,y_0) + \frac{\partial f}{\partial x}(x_0,y_0)(x-x_0) + \frac{\partial f}{\partial y}(x_0,y_0)(y-y_0)\,,} \tag{5.7}$$

defines a plane, called the **tangent plane** to the graph of the function f at $P_0 = \big(x_0, y_0, f(x_0,y_0)\big)$. This is the plane that best approximates the graph of f on a neighbourhood of P_0, see Fig. 5.3. The differentiability at $\boldsymbol{x}_0$ is equivalent to the existence of the tangent plane at P_0. In higher dimensions $n > 2$, equation (5.6) defines the hyperplane (affine subspace of codimension 1, i.e., of dimension $n-1$) tangent to the graph of f at the point $P_0 = \big(\boldsymbol{x}_0, f(\boldsymbol{x}_0)\big)$.

Equation (5.5) suggests a natural way to approximate the map f around $\boldsymbol{x}_0$ by means of a polynomial of degree one in $\boldsymbol{x}$. Neglecting the higher-order terms, we have in fact

$$\boxed{f(\boldsymbol{x}) \sim f(\boldsymbol{x}_0) + \nabla f(\boldsymbol{x}_0)\cdot(\boldsymbol{x}-\boldsymbol{x}_0)}$$

on a neighbourhood of $\boldsymbol{x}_0$. This approximation, called **linearisation** of f at $\boldsymbol{x}_0$, often allows to simplify in a constructive and efficient way the mathematical description of a physical phenomenon.

Here is yet another interpretation of (5.5): put $\Delta\boldsymbol{x} = \boldsymbol{x}-\boldsymbol{x}_0$ in the formula, so that

$$f(\boldsymbol{x}_0+\Delta\boldsymbol{x}) = f(\boldsymbol{x}_0) + \nabla f(\boldsymbol{x}_0)\cdot\Delta\boldsymbol{x} + o(\|\Delta\boldsymbol{x}\|)\,, \qquad \Delta\boldsymbol{x}\to 0\,,$$

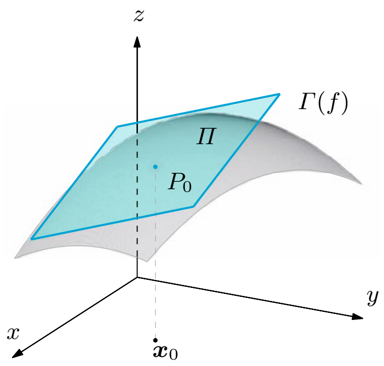


Figure 5.3. Tangent plane at P_0

in other words

$$\Delta f = f(\boldsymbol{x}_0 + \Delta \boldsymbol{x}) - f(\boldsymbol{x}_0) = \nabla f(\boldsymbol{x}_0) \cdot \Delta \boldsymbol{x} + o(\|\Delta \boldsymbol{x}\|) , \qquad \Delta \boldsymbol{x} \to 0 .$$

Then the increment Δf of the dependent variable is, up to an infinitesimal of order bigger than one, proportional to the increment $\Delta \boldsymbol{x}$ of the independent variable. This means the linear map $\Delta \boldsymbol{x} \mapsto \nabla f(\boldsymbol{x}_0) \cdot \Delta \boldsymbol{x}$ is an approximation, often sufficiently accurate, of the variation of f in a neighbourhood of $\boldsymbol{x}_0$. This fact justifies the next definition.

Definition 5.5 *The linear map* $\mathrm{d}f_{\boldsymbol{x}_0} : \mathbb{R}^n \to \mathbb{R}$

$$\mathrm{d}f_{\boldsymbol{x}_0}(\Delta \boldsymbol{x}) = \nabla f(\boldsymbol{x}_0) \cdot \Delta \boldsymbol{x}$$

is called **differential** *of* f *at* $\boldsymbol{x}_0$.

Example 5.6

Consider $f(x, y) = \sqrt{1 + x + y}$ and set $\boldsymbol{x}_0 = (1, 2)$. Then $\nabla f(\boldsymbol{x}_0) = \left(\frac{1}{4}, \frac{1}{4}\right)$ and the differential at $\boldsymbol{x}_0$ is the function

$$\mathrm{d}f_{\boldsymbol{x}_0}(\Delta x, \Delta y) = \frac{1}{4}\Delta x + \frac{1}{4}\Delta y .$$

Choosing for instance $(\Delta x, \Delta y) = \left(\frac{1}{100}, \frac{1}{20}\right)$, we will have

$$\Delta f = \sqrt{\frac{203}{50}} - 2 = 0.014944\ldots \qquad \text{while} \qquad \mathrm{d}f_{\boldsymbol{x}_0}\Big(\frac{1}{100}, \frac{1}{20}\Big) = 0.015 . \qquad \square$$

Just like in dimension one, differentiability implies continuity also for several variables.

Proposition 5.7 *A differentiable map f at $\boldsymbol{x}_0$ is continuous at $\boldsymbol{x}_0$.*

Proof. From (5.4),

$$\lim_{\boldsymbol{x}\to\boldsymbol{x}_0} f(\boldsymbol{x}) = \lim_{\boldsymbol{x}\to\boldsymbol{x}_0} \big(f(\boldsymbol{x}_0) + \nabla f(\boldsymbol{x}_0)\cdot(\boldsymbol{x}-\boldsymbol{x}_0) + o(\|\boldsymbol{x}-\boldsymbol{x}_0\|)\big) = f(\boldsymbol{x}_0)\,. \qquad \square$$

This property shows that the definition of a differentiable function of several variables is the correct analogue of what we have in dimension one.

In order to check whether a map is differentiable at a point of its domain, the following sufficient condition if usually employed. Its proof may be found in Appendix A.1.1, p. 511.

Proposition 5.8 *Assume f admits continuous partial derivatives on a neighbourhood of $\boldsymbol{x}_0$. Then f is differentiable at $\boldsymbol{x}_0$.*

Here is another feature of differentiable maps.

Proposition 5.9 *If a function f is differentiable at $\boldsymbol{x}_0$, it admits at $\boldsymbol{x}_0$ directional derivatives along any vector $\boldsymbol{v} \neq \boldsymbol{0}$, and moreover*

$$\frac{\partial f}{\partial \boldsymbol{v}}(\boldsymbol{x}_0) = \nabla f(\boldsymbol{x}_0)\cdot\boldsymbol{v} = \frac{\partial f}{\partial x_1}(\boldsymbol{x}_0)\,v_1 + \cdots + \frac{\partial f}{\partial x_n}(\boldsymbol{x}_0)\,v_n\,. \tag{5.8}$$

Proof. Using (5.4),

$$f(\boldsymbol{x}_0 + t\boldsymbol{v}) = f(\boldsymbol{x}_0) + t\nabla f(\boldsymbol{x}_0)\cdot\boldsymbol{v} + o(\|t\boldsymbol{v}\|)\,, \qquad \|t\boldsymbol{v}\| \to 0\,.$$

Since $\|t\boldsymbol{v}\| = |t|\|\boldsymbol{v}\|$, we have $o(\|t\boldsymbol{v}\|) = o(t)\,,\ t\to 0$, and hence

$$\begin{aligned}\frac{\partial f}{\partial \boldsymbol{v}}(\boldsymbol{x}_0) &= \lim_{t\to 0}\frac{f(\boldsymbol{x}_0+t\boldsymbol{v}) - f(\boldsymbol{x}_0)}{t}\\ &= \lim_{t\to 0}\frac{t\nabla f(\boldsymbol{x}_0)\cdot\boldsymbol{v} + o(t)}{t} = \nabla f(\boldsymbol{x}_0)\cdot\boldsymbol{v}\,. \qquad \square\end{aligned}$$

Note that (5.8) furnishes the expressions

$$\frac{\partial f}{\partial x_i}(\boldsymbol{x}_0) = \boldsymbol{e}_i\cdot\nabla f(\boldsymbol{x}_0), \qquad i = 1,\ldots,n\,,$$

which might turn out to be useful.

Formula (5.8) establishes a simple-yet-crucial result concerning the behaviour of a map around $\boldsymbol{x}_0$ in case the gradient is non-zero at that point. How does the directional derivative of f at $\boldsymbol{x}_0$ vary, when we change the direction along which we differentiate? To answer this we first of all establish bounds for $\dfrac{\partial f}{\partial \boldsymbol{v}}(\boldsymbol{x}_0)$ when $\boldsymbol{v}$ is a unit vector ($\|\boldsymbol{v}\| = 1$) in $\mathbb{R}^n$. By (5.8), recalling the Cauchy-Schwarz inequality (4.3), we have

$$\left|\frac{\partial f}{\partial \boldsymbol{v}}(\boldsymbol{x}_0)\right| \leq \|\nabla f(\boldsymbol{x}_0)\|$$

i.e.,

$$\boxed{-\|\nabla f(\boldsymbol{x}_0)\| \leq \frac{\partial f}{\partial \boldsymbol{v}}(\boldsymbol{x}_0) \leq \|\nabla f(\boldsymbol{x}_0)\|\ .}$$

Now, equality is attained for $\boldsymbol{v} = \pm\dfrac{\nabla f(\boldsymbol{x}_0)}{\|\nabla f(\boldsymbol{x}_0)\|}$, and then $\dfrac{\partial f}{\partial \boldsymbol{v}}(\boldsymbol{x}_0)$ reaches its (positive) maximum or (negative) minimum according to whether $\boldsymbol{v}$ is plus or minus the unit vector parallel to $\nabla f(\boldsymbol{x}_0)$. In summary we have proved the following result, shown in Fig. 5.4 (see Sect. 7.2.1 for more details).

Proposition 5.10 *At points $\boldsymbol{x}_0$ where the gradient of f does not vanish, f has the greatest rate of increase, starting from $\boldsymbol{x}_0$, in the direction of the gradient, and the greatest decrease in the opposite direction.*

The next property will be useful in the sequel. The proof of Proposition 5.9 showed that the map $\varphi(t) = f(\boldsymbol{x}_0 + t\boldsymbol{v})$ is differentiable at $t = 0$, and

$$\varphi'(0) = \nabla f(\boldsymbol{x}_0) \cdot \boldsymbol{v}\,. \tag{5.9}$$

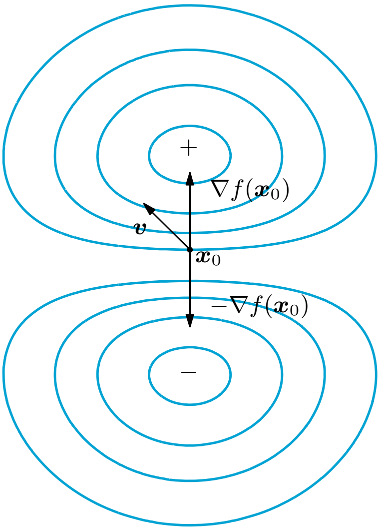


Figure 5.4. Level curves and direction of steepest slope

More generally,

Property 5.11 *Let* $\boldsymbol{a}$, $\boldsymbol{v} \in \mathbb{R}^n$ *be given points and suppose* f *is differentiable at* $\boldsymbol{a}+t_0\boldsymbol{v}$, $t_0 \in \mathbb{R}$. *Then the map* $\varphi(t) = f(\boldsymbol{a}+t\boldsymbol{v})$ *is differentiable at* t_0, *and*

$$\varphi'(t_0) = \nabla f(\boldsymbol{a}+t_0\boldsymbol{v}) \cdot \boldsymbol{v}\,. \tag{5.10}$$

Proof. Set $\Delta t = t - t_0$, so that $\boldsymbol{a} + t\boldsymbol{v} = (\boldsymbol{a} + t_0\boldsymbol{v}) + \Delta t\,\boldsymbol{v}$ and then $\varphi(t) = f(\boldsymbol{a}+t_0\boldsymbol{v}+\Delta t\boldsymbol{v}) = \psi(\Delta t)$. Now apply (5.9) to the map ψ. □

Formula (5.10) is nothing else but a special case of the rule for differentiating a composite map of several variables, for which see Sect. 6.4.

5.2.1 Mean Value Theorem and Lipschitz functions

Let $\boldsymbol{a}$, $\boldsymbol{b}$ be distinct points in $\mathbb{R}^n$ and

$$S[\boldsymbol{a},\boldsymbol{b}] = \{\boldsymbol{x}(t) = (1-t)\boldsymbol{a} + t\boldsymbol{b} : 0 \le t \le 1\}$$

be the segment between $\boldsymbol{a}$ and $\boldsymbol{b}$, as in (4.19). The result below is the n-dimensional version of the famous result for one-variable functions due to Lagrange.

Theorem 5.12 (**Mean Value Theorem** or **Lagrange Theorem**) *Let* $f : \operatorname{dom} f \subseteq \mathbb{R}^n \to \mathbb{R}$ *be defined and continuous at any point of* $S[\boldsymbol{a},\boldsymbol{b}]$, *and differentiable at any point of* $S[\boldsymbol{a},\boldsymbol{b}]$ *with the (possible) exception of the end-points* $\boldsymbol{a}$ *and* $\boldsymbol{b}$. *Then there exists an* $\overline{\boldsymbol{x}} \in S[\boldsymbol{a},\boldsymbol{b}]$ *different from* $\boldsymbol{a}$, $\boldsymbol{b}$ *such that*

$$f(\boldsymbol{b}) - f(\boldsymbol{a}) = \nabla f(\overline{\boldsymbol{x}}) \cdot (\boldsymbol{b}-\boldsymbol{a})\,. \tag{5.11}$$

Proof. Consider the auxiliary map $\varphi(t) = f\big(\boldsymbol{x}(t)\big)$, defined - and continuous - on $[0,1] \subset \mathbb{R}$, as composite of the continuous maps $t \mapsto \boldsymbol{x}(t)$ and f, for any $\boldsymbol{x}(t)$. Because $\boldsymbol{x}(t) = \boldsymbol{a}+t(\boldsymbol{b}-\boldsymbol{a})$, and using Property 5.11 with $\boldsymbol{v} = \boldsymbol{b}-\boldsymbol{a}$, we obtain that φ is differentiable (at least) on $(0,1)$, plus

$$\varphi'(t) = \nabla f\big(\boldsymbol{x}(t)\big) \cdot (\boldsymbol{b}-\boldsymbol{a})\,, \qquad 0 < t < 1\,.$$

Then φ satisfies the one-dimensional Mean Value Theorem on $[0,1]$, and there must be a $\bar{t} \in (0,1)$ such that

$$\varphi(1) - \varphi(0) = \varphi'(\bar{t})\,.$$

Putting $\overline{\boldsymbol{x}} = \boldsymbol{x}(\bar{t})$ in the above gives precisely (5.11). □

As an application, we will find a result known to the reader at least in dimension one.

Proposition 5.13 *Let $\mathcal{R}$ be a region in $\mathbb{R}^n$ and $f : \mathcal{R} \subseteq \mathbb{R}^n \to \mathbb{R}$ a continuous map on $\mathcal{R}$, differentiable everywhere on $A = \overset{\circ}{\mathcal{R}}$. Then*

$$\nabla f = 0 \quad \text{on } A \qquad \Longleftrightarrow \qquad f \text{ is constant on } \mathcal{R}\,.$$

Proof. We have seen that f constant on $\mathcal{R}$, hence on A *a fortiori*, implies $\nabla f = 0$ at each point of A (Example 5.2 iv)).

Conversely, let us fix an arbitrary $\boldsymbol{a}$ in A, and choose any $\boldsymbol{b}$ in $\mathcal{R}$. There is a polygonal path $P[\boldsymbol{a}_0, \ldots, \boldsymbol{a}_r]$ with $\boldsymbol{a}_0 = \boldsymbol{a}$, $\boldsymbol{a}_r = \boldsymbol{b}$, going from one to the other (directly from the definition of open connected set if $\boldsymbol{b} \in A$, while if $\boldsymbol{b} \in \partial A$ we can join it to a point of A through a segment). On each segment $S[\boldsymbol{a}_{j-1}, \boldsymbol{a}_j]$, $1 \le j \le r$ forming the path, the hypotheses of Theorem 5.12 hold, so from (5.11) and the fact that $\nabla f = 0$ identically, follows

$$f(\boldsymbol{a}_j) - f(\boldsymbol{a}_{j-1}) = 0\,,$$

whence $f(\boldsymbol{b}) = f(\boldsymbol{a})$ for any $\boldsymbol{b} \in \mathcal{R}$. That means f is constant. □

Consequently, if f is differentiable everywhere on an open set A and its gradient is zero on A, then f is constant on each connected component of A.

The property we are about to introduce is relevant for the applications. For this, let $\mathcal{R}$ be a region of $\operatorname{dom} f$.

Definition 5.14 *The map f is* **Lipschitz** *on $\mathcal{R}$ if there is a constant $L \ge 0$ such that*

$$|f(\boldsymbol{x}_1) - f(\boldsymbol{x}_2)| \le L\|\boldsymbol{x}_1 - \boldsymbol{x}_2\|\,, \qquad \forall \boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathcal{R}\,. \tag{5.12}$$

The smallest number L verifying the condition is said **Lipschitz constant** *of f on $\mathcal{R}$.*

By definition of least upper bound we easily see that the Lipschitz constant of f on $\mathcal{R}$ is

$$\sup_{\substack{\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathcal{R} \\ \boldsymbol{x}_1 \ne \boldsymbol{x}_2}} \frac{|f(\boldsymbol{x}_1) - f(\boldsymbol{x}_2)|}{\|\boldsymbol{x}_1 - \boldsymbol{x}_2\|}\,.$$

To say that f is Lipschitz on $\mathcal{R}$ tantamounts to the assertion that the supremum is finite.

Examples 5.15

i) The affine map $f(\boldsymbol{x}) = \boldsymbol{a} \cdot \boldsymbol{x} + b$ ($\boldsymbol{a} \in \mathbb{R}^n,\ b \in \mathbb{R}$) is Lipschitz on $\mathbb{R}^n$, as shown by (4.21). It can be proved the Lipschitz constant equals $\|\boldsymbol{a}\|$.

ii) The function $f(\boldsymbol{x}) = \|\boldsymbol{x}\|$, mapping $\boldsymbol{x} \in \mathbb{R}^n$ to its Euclidean norm, is Lipschitz on $\mathbb{R}^n$ with Lipschitz constant 1, as

$$\big|\|\boldsymbol{x}_1\| - \|\boldsymbol{x}_2\|\big| \le \|\boldsymbol{x}_1 - \boldsymbol{x}_2\|, \qquad \forall \boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathbb{R}^n .$$

This is a consequence of $\boldsymbol{x}_1 = \boldsymbol{x}_2 + (\boldsymbol{x}_1 - \boldsymbol{x}_2)$ and the triangle inequality (4.4)

$$\|\boldsymbol{x}_1\| \le \|\boldsymbol{x}_2\| + \|\boldsymbol{x}_1 - \boldsymbol{x}_2\|,$$

i.e.,

$$\|\boldsymbol{x}_1\| - \|\boldsymbol{x}_2\| \le \|\boldsymbol{x}_1 - \boldsymbol{x}_2\|;$$

swapping $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ yields the result.

iii) The map $f(\boldsymbol{x}) = \|\boldsymbol{x}\|^2$ is not Lipschitz on all $\mathbb{R}^n$. In fact, choosing $\boldsymbol{x}_2 = \boldsymbol{0}$, equation (5.12) becomes

$$\|\boldsymbol{x}_1\|^2 \le L\|\boldsymbol{x}_1\|,$$

which is true if and only if $\|\boldsymbol{x}_1\| \le L$. It becomes Lipschitz on any *bounded* region $\mathcal{R}$, for

$$\begin{aligned}\big|\|\boldsymbol{x}_1\|^2 - \|\boldsymbol{x}_2\|^2\big| &= \big(\|\boldsymbol{x}_1\| + \|\boldsymbol{x}_2\|\big)\,\big|\|\boldsymbol{x}_1\| - \|\boldsymbol{x}_2\|\big| \\ &\le 2M\big|\|\boldsymbol{x}_1\| - \|\boldsymbol{x}_2\|\big|, \quad \forall \boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathcal{R},\end{aligned}$$

where $M = \sup\{\|\boldsymbol{x}\| : \boldsymbol{x} \in \mathcal{R}\}$ and using the previous example. □

Note if $\mathcal{R}$ is open, the property of being Lipschitz implies the (uniform) continuity on $\mathcal{R}$.

There is a sufficient condition for being Lipschitz, namely:

Proposition 5.16 *Let f be differentiable over a convex region $\mathcal{R}$ inside* $\operatorname{dom} f$, *with bounded (first) partial derivatives. Then f is Lipschitz on $\mathcal{R}$. Precisely, for any $M \ge 0$ such that*

$$\left|\frac{\partial f}{\partial x_i}(\boldsymbol{x})\right| \le M, \qquad \forall \boldsymbol{x} \in \mathcal{R},\ i = 1, \ldots, n,$$

one has (5.12) *with $L = \sqrt{n}\,M$.*

Proof. Pick $\boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathcal{R}$. By assumption the segment $S[\boldsymbol{x}_1, \boldsymbol{x}_2]$ is contained in $\mathcal{R}$ and f is differentiable (hence, continuous) at any point of $S[\boldsymbol{x}_1, \boldsymbol{x}_2]$. Thus by the Mean Value Theorem 5.12 there is an $\overline{\boldsymbol{x}} \in \mathcal{R}$ with

$$f(\boldsymbol{x}_1) - f(\boldsymbol{x}_2) = \nabla f(\overline{\boldsymbol{x}}) \cdot (\boldsymbol{x}_1 - \boldsymbol{x}_2).$$

The Cauchy-Schwarz inequality (4.3) tells us

$$|f(\boldsymbol{x}_1) - f(\boldsymbol{x}_2)| \le \|\nabla f(\overline{\boldsymbol{x}})\|\,\|\boldsymbol{x}_1 - \boldsymbol{x}_2\|.$$

To conclude, observe

$$\|\nabla f(\overline{\boldsymbol{x}})\| = \left(\sum_{i=1}^n \left|\frac{\partial f}{\partial x_i}(\overline{\boldsymbol{x}})\right|^2\right)^{1/2} \le \left(\sum_{i=1}^n M^2\right)^{1/2} = \sqrt{n}\,M. \qquad \square$$

The assertion that the existence and boundedness of the partial derivatives implies being Lipschitz is a fact that holds under more general assumptions: for instance if $\mathcal{R}$ is compact, or if the boundary of $\mathcal{R}$ is sufficiently regular.

5.3 Second partial derivatives and Hessian matrix

Let f admit partial derivative in x_i on a whole neighbourhood of $\boldsymbol{x}_0$. If $\dfrac{\partial f}{\partial x_i}$ admits first derivative at $\boldsymbol{x}_0$ with respect to x_j, we say f admits at $\boldsymbol{x}_0$ **second partial derivative** with respect to x_i and x_j, and set

$$\boxed{\frac{\partial^2 f}{\partial x_j \partial x_i}(\boldsymbol{x}_0) = \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right)(\boldsymbol{x}_0)\,.}$$

If $i \neq j$ one speaks of **mixed** second partial derivatives, while for $i = j$ second partial derivatives are said of pure type and denoted by the symbol $\dfrac{\partial^2 f}{\partial x_i^2}$. Other ways to denote second partial derivatives are

$$\mathrm{D}^2_{x_j x_i} f(\boldsymbol{x}_0)\,, \qquad \mathrm{D}^2_{ji} f(\boldsymbol{x}_0)\,, \qquad f_{x_j x_i}(\boldsymbol{x}_0)\,, \qquad f_{ji}(\boldsymbol{x}_0)\,.$$

In case i is different from j, and assuming f admits at $\boldsymbol{x}_0$ mixed derivatives $\dfrac{\partial^2 f}{\partial x_j \partial x_i}(\boldsymbol{x}_0)$ and $\dfrac{\partial^2 f}{\partial x_i \partial x_j}(\boldsymbol{x}_0)$, these might differ. Take for example

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

a function such that $\dfrac{\partial^2 f}{\partial y \partial x}(0,0) = -1$ while $\dfrac{\partial^2 f}{\partial x \partial y}(0,0) = 1$.

We have arrived at an important sufficient condition, very often fulfilled, for the mixed derivatives to coincide.

Theorem 5.17 (**Schwarz**) *If the mixed partial derivatives* $\dfrac{\partial^2 f}{\partial x_j \partial x_i}$ *and* $\dfrac{\partial^2 f}{\partial x_i \partial x_j}$ $(j \neq i)$ *exist on a neighbourhood of* $\boldsymbol{x}_0$ *and are continuous at* $\boldsymbol{x}_0$*, they coincide at* $\boldsymbol{x}_0$.

Proof. See Appendix A.1.1, p. 512. □

Definition 5.18 *If f possesses all second partial derivatives at $\boldsymbol{x}_0$, the matrix*

$$\boldsymbol{H}f(\boldsymbol{x}_0) = (h_{ij})_{1 \le i,j \le n} \qquad \textit{where} \qquad h_{ij} = \frac{\partial^2 f}{\partial x_j \partial x_i}(\boldsymbol{x}_0)\,, \tag{5.13}$$

or

$$\boldsymbol{H}f(\boldsymbol{x}_0) = \begin{pmatrix} \dfrac{\partial^2 f}{\partial x_1^2}(\boldsymbol{x}_0) & \dfrac{\partial^2 f}{\partial x_2 \partial x_1}(\boldsymbol{x}_0) & \dots & \dfrac{\partial^2 f}{\partial x_n \partial x_1}(\boldsymbol{x}_0) \\ \dfrac{\partial^2 f}{\partial x_1 \partial x_2}(\boldsymbol{x}_0) & \dfrac{\partial^2 f}{\partial x_2^2}(\boldsymbol{x}_0) & \dots & \dfrac{\partial^2 f}{\partial x_n \partial x_2}(\boldsymbol{x}_0) \\ \vdots & & & \vdots \\ \dfrac{\partial^2 f}{\partial x_1 \partial x_n}(\boldsymbol{x}_0) & \dots & \dots & \dfrac{\partial^2 f}{\partial x_n^2}(\boldsymbol{x}_0) \end{pmatrix},$$

is called **Hessian matrix** *of f at $\boldsymbol{x}_0$* (**Hessian** *for short*).

The Hessian matrix is also commonly denoted by $\mathbf{Hess}_f(\boldsymbol{x}_0)$, or $\boldsymbol{H}_f(\boldsymbol{x}_0)$.

If the mixed derivatives are continuous at $\boldsymbol{x}_0$, the matrix $\boldsymbol{H}f(\boldsymbol{x}_0)$ is symmetric by Schwarz's Theorem. Then the order of differentiation is irrelevant, and $f_{x_i x_j}(\boldsymbol{x}_0)$, $f_{ij}(\boldsymbol{x}_0)$ is the same as $f_{x_j x_i}(\boldsymbol{x}_0)$, $f_{ji}(\boldsymbol{x}_0)$.

The Hessian makes its appearance when studying the local behaviour of f at $\boldsymbol{x}_0$, as explained in Sect. 5.6.

Examples 5.19

i) For the map $f(x,y) = x\sin(x+2y)$ we have

$$\frac{\partial f}{\partial x}(x,y) = \sin(x+2y) + x\cos(x+2y)\,, \qquad \frac{\partial f}{\partial y}(x,y) = 2x\cos(x+2y)\,,$$

so that

$$\begin{aligned} \frac{\partial^2 f}{\partial x^2}(\boldsymbol{x}) &= 2\cos(x+2y) - x\sin(x+2y)\,, \\ \frac{\partial^2 f}{\partial x \partial y}(\boldsymbol{x}) = \frac{\partial^2 f}{\partial y \partial x}(\boldsymbol{x}) &= 2\cos(x+2y) - 2x\sin(x+2y)\,, \\ \frac{\partial^2 f}{\partial y^2}(\boldsymbol{x}) &= -4x\sin(x+2y)\,. \end{aligned}$$

At the origin $\boldsymbol{x}_0 = \mathbf{0}$ the Hessian of f is

$$\boldsymbol{H}f(\mathbf{0}) = \begin{pmatrix} 2 & 2 \\ 2 & 0 \end{pmatrix}.$$

ii) Given the symmetric matrix $\boldsymbol{A} \in \mathbb{R}^{n\times n}$, a vector $\boldsymbol{b} \in \mathbb{R}^n$ and a constant $c \in \mathbb{R}$, we define the map $f : \mathbb{R}^n \to \mathbb{R}$ by

$$f(\boldsymbol{x}) = \boldsymbol{x} \cdot \boldsymbol{A}\boldsymbol{x} + \boldsymbol{b} \cdot \boldsymbol{x} + c = \sum_{p=1}^{n} x_p \left(\sum_{q=1}^{n} a_{pq} x_q \right) + \sum_{p=1}^{n} b_p x_p + c\,.$$

For the product $\boldsymbol{Ax}$ to make sense, we are forced to write $\boldsymbol{x}$ as a column vector, as will happen for all vectors considered in the example. By the chain rule

$$\frac{\partial f}{\partial x_i}(\boldsymbol{x}) = \sum_{p=1}^{n} \frac{\partial x_p}{\partial x_i}\left(\sum_{q=1}^{n} a_{pq}x_q\right) + \sum_{p=1}^{n} x_p\left(\sum_{q=1}^{n} a_{pq}\frac{\partial x_q}{\partial x_i}\right) + \sum_{p=1}^{n} b_p\frac{\partial x_p}{\partial x_i}\,.$$

For each pair of indices p, i between 1 and n,

$$\frac{\partial x_p}{\partial x_i} = \delta_{pi} = \begin{cases} 1 & \text{if } p = i\,, \\ 0 & \text{if } p \neq i\,, \end{cases}$$

so

$$\frac{\partial f}{\partial x_i}(\boldsymbol{x}) = \sum_{q=1}^{n} a_{iq}x_q + \sum_{p=1}^{n} x_p a_{pi} + b_i\,.$$

On the other hand $\boldsymbol{A}$ is symmetric and the summation is index arbitrary, so

$$\sum_{p=1}^{n} x_p a_{pi} = \sum_{p=1}^{n} a_{ip}x_p = \sum_{q=1}^{n} a_{iq}x_q\,.$$

Then

$$\frac{\partial f}{\partial x_i}(\boldsymbol{x}) = 2\sum_{q=1}^{n} a_{iq}x_q + b_i\,,$$

i.e.,

$$\nabla f(\boldsymbol{x}) = 2\boldsymbol{Ax} + \boldsymbol{b}\,.$$

Differentiating further,

$$h_{ij} = \frac{\partial^2 f}{\partial x_j \partial x_i}(\boldsymbol{x}) = 2a_{ij}\,, \qquad 1 \leq i, j \leq n\,,$$

$$\boldsymbol{H}f(\boldsymbol{x}) = 2\boldsymbol{A}\,.$$

Note the Hessian of f is independent of $\boldsymbol{x}$.

iii) The kinetic energy of a body with mass m moving at velocity v is $K = \frac{1}{2}mv^2$. Then

$$\nabla K(m, v) = \left(\frac{1}{2}v^2, mv\right) \qquad \text{and} \qquad \boldsymbol{H}K(m, v) = \begin{pmatrix} 0 & v \\ v & m \end{pmatrix}.$$

Note that

$$\frac{\partial K}{\partial m}\frac{\partial^2 K}{\partial v^2} = K\,.$$

□

5.4 Higher-order partial derivatives

The second partial derivatives $\dfrac{\partial^2 f}{\partial x_j \partial x_i}$ of f have been defined as first partial derivatives of the functions $\dfrac{\partial f}{\partial x_i}$; under Schwarz's Theorem, the order of differentiation is inessential. Similarly one defines the partial derivatives of order three as first derivatives of the maps $\dfrac{\partial^2 f}{\partial x_j \partial x_i}$ (assuming this is possible, of course). In general, by successive differentiation one defines kth **partial derivatives** (or **partial derivatives of order** k) of f, for any integer $k \geq 1$. Supposing that the mixed partial derivatives of all orders are continuous, and thus satisfy Schwarz's Theorem, one indicates by

$$\boxed{\frac{\partial^k f}{\partial x_1^{\alpha_1} \partial x_2^{\alpha_2} \dots \partial x_n^{\alpha_n}}(\boldsymbol{x}_0)}$$

the kth partial derivative of f at $\boldsymbol{x}_0$, obtained differentiating α_1 times with respect to x_1, α_2 times with respect to x_2, ..., α_n times with respect to x_n. The exponents α_i are integers between 0 and k such that $\alpha_1 + \alpha_2 + \dots + \alpha_n = k$. Frequent symbols include $\mathrm{D}^{\boldsymbol{\alpha}} f(\boldsymbol{x}_0)$, that involves the multi-index $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \dots, \alpha_n) \in \mathbb{N}^n$, but also, e.g., f_{xxy} to denote differentiation in x twice, in y once.

One last piece of notation concerns functions f, called **of class** $\mathcal{C}^k$ $(k \geq 1)$ over an open set $\Omega \subseteq \operatorname{dom} f$ if the partial derivatives of any order $\leq k$ exist and are continuous everywhere on Ω. The set of such maps is indicated by $\mathcal{C}^k(\Omega)$. By extension, $\mathcal{C}^0(\Omega)$ will be the set of continuous maps on Ω, and $\mathcal{C}^\infty(\Omega)$ the set of maps belonging to $\mathcal{C}^k(\Omega)$ for any k, hence the functions admitting partial derivatives of any order on Ω. Notice the inclusions

$$\mathcal{C}^\infty(\Omega) \subset \dots \subset \mathcal{C}^k(\Omega) \subset \mathcal{C}^{k-1}(\Omega) \subset \dots \subset \mathcal{C}^0(\Omega)\,.$$

Thanks to Propostion 5.8, a function in $\mathcal{C}^1(\Omega)$ is differentiable everywhere on Ω.

Instead of the open set Ω, we may take the closure $\overline{\Omega}$, and assume $\overline{\Omega}$ is contained in $\operatorname{dom} f$. If so, we write $f \in \mathcal{C}^0(\overline{\Omega})$ if f is continuous at any point of $\overline{\Omega}$; for $1 \leq k \leq \infty$, we write $f \in \mathcal{C}^k(\overline{\Omega})$, or say f is $\mathcal{C}^k$ on $\overline{\Omega}$, if there is an open set Ω' with $\overline{\Omega} \subset \Omega' \subseteq \operatorname{dom} f$ and $f \in \mathcal{C}^k(\Omega')$.

5.5 Taylor expansions; convexity

Taylor expansions allow to approximate, locally, a function using a polynomial in the independent variables by the knowledge of certain partial derivatives, just as in dimension one. We already encountered examples; for a differentiable map formula (5.5) holds, which is the Taylor expansion at first order with Peano's remainder. Note

$$\boxed{Tf_{1,\boldsymbol{x}_0}(\boldsymbol{x}) = f(\boldsymbol{x}_0) + \nabla f(\boldsymbol{x}_0) \cdot (\boldsymbol{x} - \boldsymbol{x}_0)}$$

is a polynomial of degree less or equal than 1 in the x_i, called **Taylor polynomial of f at $\boldsymbol{x}_0$ of order** 1. Besides, if f is $\mathcal{C}^1$ on a neighbourhood of $\boldsymbol{x}_0$, then (5.11) holds with $\boldsymbol{a} = \boldsymbol{x}_0$ and $\boldsymbol{b} = \boldsymbol{x}$ arbitrarily chosen in the neighbourhood, so

$$\boxed{f(\boldsymbol{x}) = f(\boldsymbol{x}_0) + \nabla f(\overline{\boldsymbol{x}}) \cdot (\boldsymbol{x} - \boldsymbol{x}_0)\,, \qquad \overline{\boldsymbol{x}} \in S[\boldsymbol{x}, \boldsymbol{x}_0]\,.} \tag{5.14}$$

Viewing the constant $f(\boldsymbol{x}_0)$ as a 0-degree polynomial, the formula gives the Taylor expansion of f at $\boldsymbol{x}_0$ of order 0, with Lagrange's remainder.

Increasing the map's regularity we can find Taylor formulas that are more and more precise. For $\mathcal{C}^2$ maps we have the following results, whose proofs are given in Appendix A.1.2, p. 513 and p. 514.

Theorem 5.20 *A function f of class $\mathcal{C}^2$ around $\boldsymbol{x}_0$ admits at $\boldsymbol{x}_0$ the* **Taylor expansion of order one with Lagrange's remainder**:

$$f(\boldsymbol{x}) = f(\boldsymbol{x}_0) + \nabla f(\boldsymbol{x}_0) \cdot (\boldsymbol{x} - \boldsymbol{x}_0) + \frac{1}{2}(\boldsymbol{x} - \boldsymbol{x}_0) \cdot \boldsymbol{H}f(\overline{\boldsymbol{x}})(\boldsymbol{x} - \boldsymbol{x}_0)\,, \tag{5.15}$$

where $\overline{\boldsymbol{x}}$ is interior to the segment $S[\boldsymbol{x}, \boldsymbol{x}_0]$.

Formulas (5.4) and (5.15) are two different ways of writing the remainder of f in the degree-one Taylor polynomial $Tf_{1,\boldsymbol{x}_0}(\boldsymbol{x})$.

The expansion of order two is given by

Theorem 5.21 *A function f of class $\mathcal{C}^2$ around $\boldsymbol{x}_0$ admits at $\boldsymbol{x}_0$ the following* **Taylor expansion of order two with Peano's remainder**:

$$\begin{aligned} f(\boldsymbol{x}) = f(\boldsymbol{x}_0) + \nabla f(\boldsymbol{x}_0) \cdot (\boldsymbol{x} - \boldsymbol{x}_0) + \frac{1}{2}(\boldsymbol{x} - \boldsymbol{x}_0) \cdot \boldsymbol{H}f(\boldsymbol{x}_0)(\boldsymbol{x} - \boldsymbol{x}_0) \\ +o(\|\boldsymbol{x} - \boldsymbol{x}_0\|^2)\,, \qquad \boldsymbol{x} \to \boldsymbol{x}_0\,. \end{aligned} \tag{5.16}$$

The expression

$$\boxed{Tf_{2,\boldsymbol{x}_0}(\boldsymbol{x}) = f(\boldsymbol{x}_0) + \nabla f(\boldsymbol{x}_0) \cdot (\boldsymbol{x} - \boldsymbol{x}_0) + \frac{1}{2}(\boldsymbol{x} - \boldsymbol{x}_0) \cdot \boldsymbol{H}f(\boldsymbol{x}_0)(\boldsymbol{x} - \boldsymbol{x}_0)}$$

is a polynomial of degree ≤ 2 in the x_i, called **Taylor polynomial of f at $\boldsymbol{x}_0$ of order** 2. It gives the best quadratic approximation of the map on the neighbourhood of $\boldsymbol{x}_0$ (see Fig. 5.5 for an example).

For clarity's sake, let us render (5.16) explicit for an $f(x, y)$ of two variables:

$$\begin{aligned} f(x,y) &= f(x_0,y_0) + f_x(x_0,y_0)(x - x_0) + f_y(x_0,y_0)(y - y_0) \\ &\quad + \frac{1}{2} f_{xx}(x_0,y_0)(x - x_0)^2 + f_{xy}(x_0,y_0)(x - x_0)(y - y_0) + \frac{1}{2} f_{yy}(x_0,y_0)(y - y_0)^2 \\ &\quad + o\big((x - x_0)^2 + (y - y_0)^2\big)\,, \qquad (x,y) \to (x_0,y_0)\,. \end{aligned}$$

For the quadratic term we used the fact that the Hessian is symmetric by Schwarz's Theorem.

Taylor formulas of arbitrary order n can be written assuming f is $\mathcal{C}^n$ around $\boldsymbol{x}_0$. These, though, go beyond the scope of the present text.

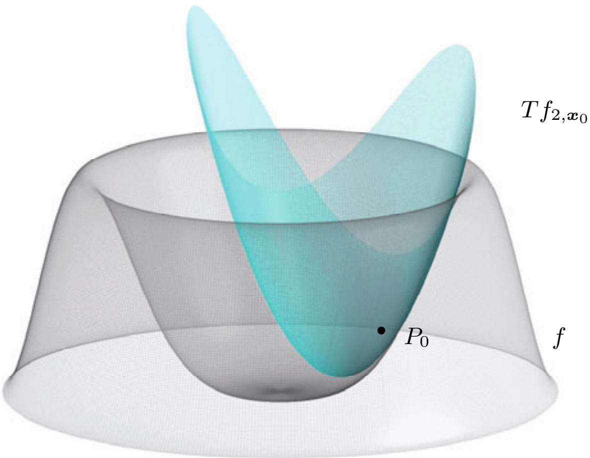


Figure 5.5. Graph of the order-two Taylor polynomial of f at $\boldsymbol{x}_0$ (osculating paraboloid at P_0)

5.5.1 Convexity

The idea of convexity, both global and local, for one-variable functions (Vol. I, Sect. 6.9) can be generalised to several variables.

Global convexity goes through the convexity of the set of points above the graph, and namely: take $f : C \subseteq \mathbb{R}^n \to \mathbb{R}$ with C a convex set, and define

$$E_f = \big\{(\boldsymbol{x}, y) \in \mathbb{R}^{n+1} : \boldsymbol{x} \in C,\ y \geq f(\boldsymbol{x})\big\}.$$

Then f is **convex** on C if the set E_f is convex.

Local convexity depends upon the mutual position of f's graph and the tangent plane. Precisely, f differentiable at $\boldsymbol{x}_0 \in \operatorname{dom} f$ is said **convex at $\boldsymbol{x}_0$** if there is a neighbourhood $B_r(\boldsymbol{x}_0)$ such that

$$\boxed{f(\boldsymbol{x}) \geq f(\boldsymbol{x}_0) + \nabla f(\boldsymbol{x}_0) \cdot (\boldsymbol{x} - \boldsymbol{x}_0)\,, \qquad \forall \boldsymbol{x} \in B_r(\boldsymbol{x}_0)\,;}$$

f is **strictly convex at $\boldsymbol{x}_0$** if the inequality is strict for any $\boldsymbol{x} \neq \boldsymbol{x}_0$.

It can be proved that the local convexity of a differentiable map f at any point in a convex subset $C \subseteq \operatorname{dom} f$ is equivalent to the global convexity of f on C.

Take a $\mathcal{C}^2$ map f around a point $\boldsymbol{x}_0$ in $\operatorname{dom} f$. Using Theorem 5.21 and the properties of the symmetric matrix $\boldsymbol{H}f(\boldsymbol{x}_0)$ (see Sect. 4.2), we can say that

$$\boxed{\begin{array}{lcl} \boldsymbol{H}f(\boldsymbol{x}_0) \text{ is positive semi-definite} & \Longleftrightarrow & f \text{ is convex at } \boldsymbol{x}_0 \\[1ex] \boldsymbol{H}f(\boldsymbol{x}_0) \text{ is positive definite} & \Longrightarrow & f \text{ is strictly convex at } \boldsymbol{x}_0\,. \end{array}}$$

5.6 Extremal points of a function; stationary points

Extremum values and extremum points (local or global) in several variables are defined in analogy to dimension one.

Definition 5.22 *A point* $\boldsymbol{x}_0 \in \operatorname{dom} f$ *is a* **relative** (or **local**) **maximum point for** f *if, on a neighbourhood* $B_r(\boldsymbol{x}_0)$ *of* $\boldsymbol{x}_0$*,*

$$\forall \boldsymbol{x} \in B_r(\boldsymbol{x}_0) \cap \operatorname{dom} f, \qquad f(\boldsymbol{x}) \leq f(\boldsymbol{x}_0).$$

The value $f(\boldsymbol{x}_0)$ *is a* **relative maximum of** f*.*
Moreover, $\boldsymbol{x}_0$ *is said an* **absolute***, or* **global**, **maximum point for** f *if*

$$\forall \boldsymbol{x} \in \operatorname{dom} f, \qquad f(\boldsymbol{x}) \leq f(\boldsymbol{x}_0),$$

and correspondingly $f(\boldsymbol{x}_0)$ *is the* **absolute maximum of** f*. Either way, the maximum is* **strict** *if* $f(\boldsymbol{x}) < f(\boldsymbol{x}_0)$ *when* $\boldsymbol{x} \neq \boldsymbol{x}_0$*.*

Inverting the inequalities defines **relative** and **absolute minimum points**. Minimum and maximum points alike are called **extrema**, or **extremum points**, for f.

Examples 5.23

i) The map $f(\boldsymbol{x}) = \|\boldsymbol{x}\|$ has a strict global minimum at the origin, for $f(\boldsymbol{0}) = \boldsymbol{0}$ and $f(\boldsymbol{x}) > 0$ for any $\boldsymbol{x} \neq \boldsymbol{0}$. Clearly f has no maximum points (neither relative, nor absolute) on $\mathbb{R}^n$.

ii) The function $f(x, y) = x^2(\mathrm{e}^{-y^2} - 1)$ is always ≤ 0, since $x^2 \geq 0$ and $\mathrm{e}^{-y^2} \leq 1$ for all $(x, y) \in \mathbb{R}^2$. Moreover, it vanishes if $x = 0$ or $y = 0$, i.e., $f(0, y) = 0$ for any $y \in \mathbb{R}$ and $f(x, 0) = 0$ for all $x \in \mathbb{R}$. Hence all points on the coordinate axes are global maxima (not strict). □

Extremum points as of Definition 5.22 are commonly called *unconstrained*, because the independent variable is "free" to roam the whole domain of the function. Later (Sect. 7.3) we will see the notion of "constrained" extremum points, for which the independent variable is restricted to a subset of the domain, like a curve or a surface.

A sufficient condition for having extrema in several variables is Weierstrass's Theorem (seen in Vol. I, Thm. 4.31); the proof is completely analogous.

Theorem 5.24 (**Weierstrass**) *Let* f *be continuous on a compact set* $K \subseteq \operatorname{dom} f$*. Its image* $f(K)$ *is a closed and bounded subset of* $\mathbb{R}$*, and in particular,* $f(K)$ *is a closed interval if* K *is connected.*
Consequently, f *has on* K *a maximum and a minimum value.*

For example, $f(\boldsymbol{x}) = \|\boldsymbol{x}\|$ admits absolute maximum on every compact set $K \subset \mathbb{R}^n$.

There is also a notion of critical point for functions of several variables.

Definition 5.25 *A point $\boldsymbol{x}_0$ at which f is differentiable is* **critical** *or* **stationary** *for f if $\nabla f(\boldsymbol{x}_0) = \mathbf{0}$. If, instead, $\nabla f(\boldsymbol{x}_0) \neq \mathbf{0}$, $\boldsymbol{x}_0$ is said a* **regular point** *for f.*

By (5.8) a stationary point annihilates all directional derivatives of f.

In two variables stationary points have a neat geometric interpretation. Recalling (5.7), a point is stationary if the tangent plane to the graph is horizontal (Fig. 5.6).

It is Fermat's Theorem that justifies the interest in finding stationary points; we state it below for the several-variable case.

Theorem 5.26 (**Fermat**) *Let f be differentiable at the extremum point $\boldsymbol{x}_0$. Then $\boldsymbol{x}_0$ is stationary for f.*

Proof. By assumption the map of one variable

$$x \mapsto f(x_{01}, \ldots, x_{0,i-1}, x, x_{0,i+1}, \ldots, x_{0n})$$

is, for any i, defined and differentiable on a neighbourhood of x_{0i}; the latter is an extremum point. Thus, Fermat's Theorem for one-variable functions gives $\dfrac{\partial f}{\partial x_i}(\boldsymbol{x}_0) = 0$. □

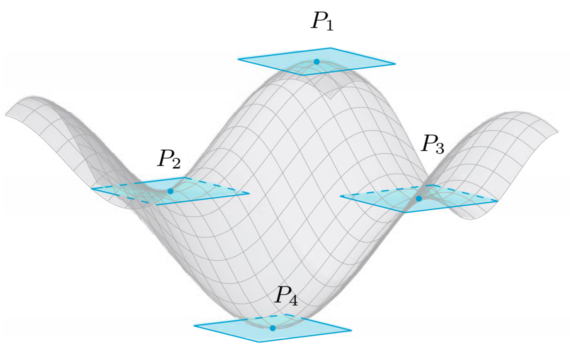


Figure 5.6. Tangent planes at stationary points in two variables

In the light of this result, we make a few observations on the relationship between extrema and stationary points.

i) Having an extremum point for f does not mean f is differentiable at it, nor that the point is stationary. This is exactly what happens to $f(\boldsymbol{x}) = \|\boldsymbol{x}\|$ of Example 5.23 i): the origin is an absolute minimum, but f does not admit partial derivatives there. In fact f behaves, along each direction, as the absolute value, for $f(0, \ldots, 0, x, 0, \ldots, 0) = |x|$.

ii) For maps that are differentiable on the whole domain, Fermat's Theorem provides a *necessary* condition for an extremum point; this means extremum points are to be found among stationary points.

iii) That said, not all stationary points are extrema. Consider $f(x, y) = x^3y^3$: the origin is stationary, yet neither a maximum nor a minimum point. This map is zero along the axes in fact, positive in the first and third quadrants, negative in the others.

Taking all this into consideration, it makes sense to search for *sufficient* conditions ensuring a stationary point $\boldsymbol{x}_0$ is an extremum point. Apropos which, the Hessian matrix $\boldsymbol{H}f(\boldsymbol{x}_0)$ is useful when the function f is at least $\mathcal{C}^2$ around $\boldsymbol{x}_0$. With these assumptions and the fact that $\boldsymbol{x}_0$ is stationary, we have Taylor's expansion (5.16)

$$f(\boldsymbol{x}) - f(\boldsymbol{x}_0) = Q(\boldsymbol{x} - \boldsymbol{x}_0) + o(\|\boldsymbol{x} - \boldsymbol{x}_0\|^2)\,, \quad \boldsymbol{x} \to \boldsymbol{x}_0\,, \tag{5.17}$$

where $Q(\boldsymbol{v}) = \frac{1}{2}\boldsymbol{v} \cdot \boldsymbol{H}f(\boldsymbol{x}_0)\boldsymbol{v}$ is the quadratic form associated to the symmetric matrix $\boldsymbol{H}f(\boldsymbol{x}_0)$ (see Sect. 4.2).

Now we do have a sufficient condition for a stationary point to be extremal.

Theorem 5.27 *Let f be $\mathcal{C}^2$ on some neighbourhood of $\boldsymbol{x}_0$, a stationary point for f. Then*

i) if $\boldsymbol{x}_0$ is a minimum (respectively, maximum) point for f, $\boldsymbol{H}f(\boldsymbol{x}_0)$ is positive (negative) semi-definite;
ii) if $\boldsymbol{H}f(\boldsymbol{x}_0)$ is positive (negative) definite, the point $\boldsymbol{x}_0$ is a local strict minimum (maximum) for f.

Proof. i) To fix ideas let us suppose $\boldsymbol{x}_0$ is a local minimum for f, and $B_r(\boldsymbol{x}_0)$ is a neighbourhood of $\boldsymbol{x}_0$ where $f(\boldsymbol{x}) \geq f(\boldsymbol{x}_0)$. Choosing an arbitrary $\boldsymbol{v} \in \mathbb{R}^n$, let $\boldsymbol{x} = \boldsymbol{x}_0 + \varepsilon\boldsymbol{v}$ with $\varepsilon > 0$ small enough so that $\boldsymbol{x} \in B_r(\boldsymbol{x}_0)$. From (5.17),

$$Q(\boldsymbol{x} - \boldsymbol{x}_0) + o(\|\boldsymbol{x} - \boldsymbol{x}_0\|^2) \geq 0\,, \qquad \boldsymbol{x} \to \boldsymbol{x}_0\,.$$

But $Q(\boldsymbol{x} - \boldsymbol{x}_0) = Q(\varepsilon\boldsymbol{v}) = \varepsilon^2 Q(\boldsymbol{v})$, and $o(\|\boldsymbol{x} - \boldsymbol{x}_0\|^2) = o(\varepsilon^2\|\boldsymbol{v}\|^2) = \varepsilon^2 o(1)$ as $\varepsilon \to 0^+$. Therefore

$$\varepsilon^2 Q(\boldsymbol{v}) + \varepsilon^2 o(1) \geq 0\,, \qquad \varepsilon \to 0^+\,,$$

i.e.,

$$Q(\boldsymbol{v}) + o(1) \geq 0\,, \qquad \varepsilon \to 0^+\,.$$

Taking the limit $\varepsilon \to 0^+$ and noting $Q(\boldsymbol{v})$ does not depend on ε, we get $Q(\boldsymbol{v}) \geq 0$. But as $\boldsymbol{v}$ is arbitrary, $\boldsymbol{H}f(\boldsymbol{x}_0)$ is positive semi-definite.

ii) Let $\boldsymbol{H}f(\boldsymbol{x}_0)$ be positive definite. Then $Q(\boldsymbol{v}) \geq \alpha\|\boldsymbol{v}\|^2$ for any $\boldsymbol{v} \in \mathbb{R}^n$, where $\alpha = \lambda_*/2$ and $\lambda_* > 0$ denotes the smallest eigenvalue of $\boldsymbol{H}f(\boldsymbol{x}_0)$ (see (4.18)). By (5.17),

$$\begin{aligned} f(\boldsymbol{x}) - f(\boldsymbol{x}_0) &\geq \alpha\|\boldsymbol{x}-\boldsymbol{x}_0\|^2 + \|\boldsymbol{x}-\boldsymbol{x}_0\|^2 o(1) \\ &= \big(\alpha + o(1)\big)\|\boldsymbol{x}-\boldsymbol{x}_0\|^2\,, \qquad \boldsymbol{x} \to \boldsymbol{x}_0\,. \end{aligned}$$

On a neighbourhood $B_r(\boldsymbol{x}_0)$ of sufficiently small radius, $\alpha + o(1) > 0$ hence $f(\boldsymbol{x}) \geq f(\boldsymbol{x}_0)$. □

As corollary of the theorem, on a neighbourhood of a minimum point $\boldsymbol{x}_0$ for f, where $\boldsymbol{H}f(\boldsymbol{x}_0)$ is positive definite, the graph of f is well approximated by the quadratic map $g(\boldsymbol{x}) = f(\boldsymbol{x}_0) + Q(\boldsymbol{x}-\boldsymbol{x}_0)$, an elliptic paraboloid in dimension 2. Furthermore, the level sets are approximated by the level sets of $Q(\boldsymbol{x}-\boldsymbol{x}_0)$; as we recalled in Sect. 4.2, these are ellipses (in dimension 2) or ellipsoids (in dimension 3) centred at $\boldsymbol{x}_0$.

Remark 5.28 One could prove that if f is $\mathcal{C}^2$ on its domain and $\boldsymbol{H}f(\boldsymbol{x})$ everywhere positive (or negative) definite, then f admits at most one stationary point $\boldsymbol{x}_0$, which is also a global minimum (maximum) for f. □

Examples 5.29

i) Consider

$$f(x,y) = 2x\mathrm{e}^{-(x^2+y^2)}$$

on $\mathbb{R}^2$ and compute

$$\frac{\partial f}{\partial x}(x,y) = 2(1-2x^2)\mathrm{e}^{-(x^2+y^2)}\,, \qquad \frac{\partial f}{\partial y}(x,y) = -4xy\mathrm{e}^{-(x^2+y^2)}\,.$$

The zeroes of these expressions are the stationary points $\boldsymbol{x}_1 = \big(\frac{\sqrt{2}}{2},0\big)$ and $\boldsymbol{x}_2 = \big(-\frac{\sqrt{2}}{2},0\big)$. Moreover,

$$\begin{aligned} \frac{\partial^2 f}{\partial x^2}(x,y) &= 4x(2x^2-3)\mathrm{e}^{-(x^2+y^2)}\,, \\ \frac{\partial^2 f}{\partial x\partial y}(x,y) &= \frac{\partial^2 f}{\partial y\partial x}(x,y) = 4y(2x^2-1)\mathrm{e}^{-(x^2+y^2)}\,, \\ \frac{\partial^2 f}{\partial y^2}(x,y) &= 4x(2y^2-1)\mathrm{e}^{-(x^2+y^2)}\,, \end{aligned}$$

so

$$\boldsymbol{H}f\Big(\frac{\sqrt{2}}{2},0\Big) = \begin{pmatrix} -4\sqrt{2}\mathrm{e}^{-1/2} & 0 \\ 0 & -2\sqrt{2}\mathrm{e}^{-1/2} \end{pmatrix}$$

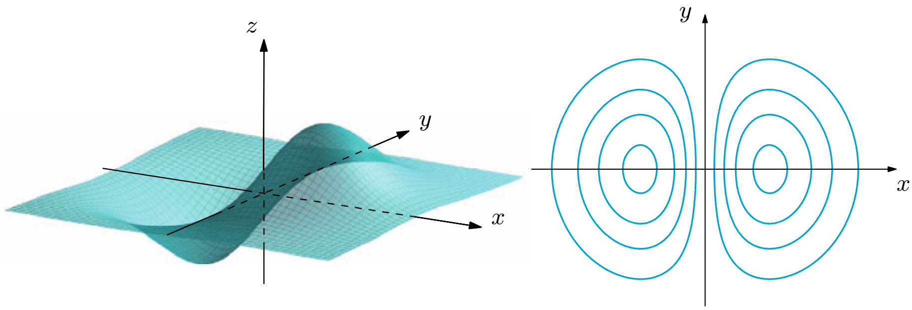


Figure 5.7. Graph and level curves of $f(x,y) = 2x\mathrm{e}^{-(x^2+y^2)}$

and

$$\boldsymbol{H}f\big(-\frac{\sqrt{2}}{2},0\big) = \begin{pmatrix} 4\sqrt{2}\mathrm{e}^{-1/2} & 0 \\ 0 & 2\sqrt{2}\mathrm{e}^{-1/2} \end{pmatrix}.$$

The Hessian matrices are diagonal, whence $\boldsymbol{H}f(\boldsymbol{x}_1)$ is negative definite while $\boldsymbol{H}f(\boldsymbol{x}_2)$ positive definite. In summary, $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ are local extrema (a local maximum and a local minimum, respectively). Fig. 5.7 shows the graph and the level curves of f.

ii) The function

$$f(x,y,z) = \frac{1}{x} + y^2 + \frac{1}{z} + xz$$

is defined on $\mathrm{dom}\, f = \{(x,y,z) \in \mathbb{R}^3 : x \neq 0 \text{ and } z \neq 0\}$. As

$$\nabla f(x,y,z) = \Big(-\frac{1}{x^2} + z, 2y, -\frac{1}{z^2} + x\Big),$$

imposing $\nabla f(x,y,z) = \boldsymbol{0}$ produces only one stationary point $\boldsymbol{x}_0 = (1,0,1)$. Then

$$\boldsymbol{H}f(x,y,z) = \begin{pmatrix} 2/x^3 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 2/z^3 \end{pmatrix}, \qquad \text{whence} \qquad \boldsymbol{H}f(1,0,1) = \begin{pmatrix} 2 & 0 & 1 \\ 0 & 2 & 0 \\ 1 & 0 & 2 \end{pmatrix}.$$

The characteristic equation of $\boldsymbol{A} = \boldsymbol{H}f(1,0,1)$ reads

$$\det(\boldsymbol{A} - \lambda\boldsymbol{I}) = (2-\lambda)\big((2-\lambda)^2 - 1\big) = 0,$$

solved by $\lambda_1 = 1$, $\lambda_2 = 2$, $\lambda_3 = 3$. Therefore the Hessian at $\boldsymbol{x}_0$ is positive definite, making $\boldsymbol{x}_0$ a local minimum point. □

5.6.1 Saddle points

Recalling what an indefinite matrix is (see Sect. 4.2), statement *i)* of Theorem 5.27 may be formulated in the following equivalent way.

Proposition 5.30 *Let f be of class $\mathcal{C}^2$ around a stationary point $\boldsymbol{x}_0$. If $\boldsymbol{H}f(\boldsymbol{x}_0)$ is indefinite, $\boldsymbol{x}_0$ is not an extremum point.*

A stationary point $\boldsymbol{x}_0$ for f such that $\boldsymbol{H}f(\boldsymbol{x}_0)$ is indefinite is called a **saddle point**. The name stems from the shape of the graph in the following example around $\boldsymbol{x}_0$.

Example 5.31

Consider $f(x,y) = x^2 - y^2$: as $\nabla f(x,y) = (2x, -2y)$, there is only one stationary point, the origin. The Hessian matrix

$$\boldsymbol{H}f(x,y) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$$

is independent of the point, hence indefinite. Therefore the origin is a saddle point.

It is convenient to consider in more detail the behaviour of f around such a point. Moving along the x-axis, the map $f(x,0) = x^2$ has a minimum at the origin. Along the y-axis, by contrast, the function $f(0,y) = -y^2$ has a maximum at the origin:

$$f(0,0) = \min_{x\in\mathbb{R}} f(x,0) = \max_{x\in\mathbb{R}} f(0,y)\,.$$

Level curves and the graph (from two viewpoints) of the function f are shown in Fig. 5.8 and 5.9. □

The kind of behaviour just described is typical of stationary points at which the Hessian matrix is indefinite and non-singular (the eigenvalues are non-zero and have different signs). Let us see more examples.

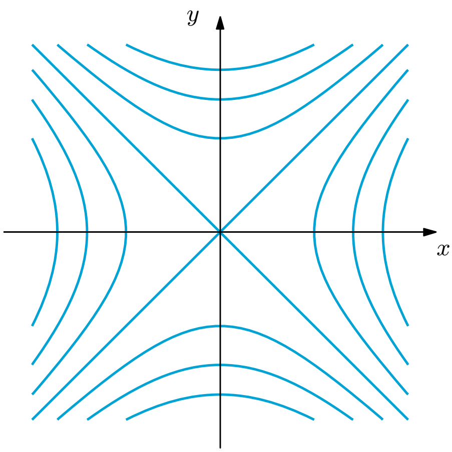


Figure 5.8. Level curves of $f(x,y) = x^2 - y^2$

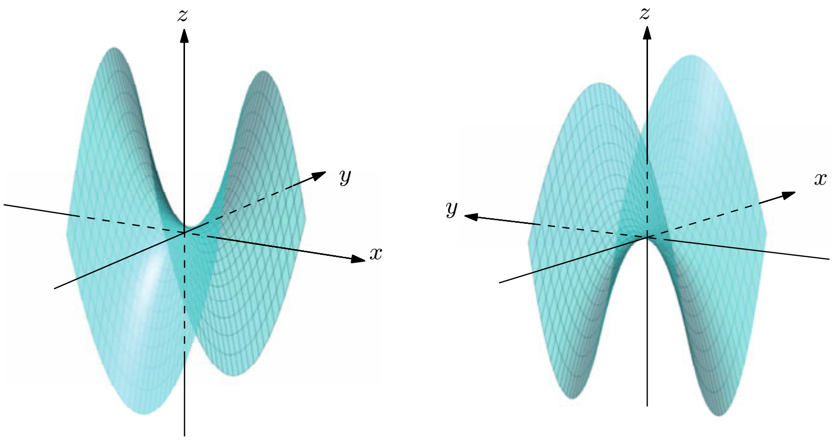


Figure 5.9. Graphical view of the function $f(x, y) = x^2 - y^2$ from different angles

Examples 5.32

i) For the map $f(x, y) = xy$ we have

$$\nabla f(x, y) = (y, x) \qquad \text{and} \qquad \boldsymbol{H}f(x, y) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

As before, $\boldsymbol{x}_0 = (0, 0)$ is a saddle point, because the eigenvalues of the Hessian are $\lambda_1 = 1$ and $\lambda_2 = -1$. Moving along the bisectrix of the first and third quadrant, f has a minimum at $\boldsymbol{x}_0$, while along the orthogonal bisectrix f has a maximum at $\boldsymbol{x}_0$:

$$f(0, 0) = \min_{x \in \mathbb{R}} f(x, x) = \max_{x \in \mathbb{R}} f(x, -x).$$

The directions of these lines are those of the eigenvectors $\boldsymbol{w}_1 = (1, 1)$, $\boldsymbol{w}_2 = (-1, 1)$ associated to the eigenvalues λ_1, λ_2.

Changing variables $x = u - v$, $y = u + v$ (corresponding to a rotation of $\pi/4$ in the plane, see Sect. 6.6 and Example 6.31 in particular), f becomes

$$f(x, y) = (u - v)(u + v) = u^2 - v^2 = \tilde{f}(u, v),$$

the same as in the previous example in the new variables u, v.

ii) The function $f(x, y, z) = x^2 + y^2 - z^2$ has gradient $\nabla f(x, y, z) = (2x, 2y, -2z)$, with unique stationary point the origin. The matrix

$$\boldsymbol{H}f(x, y, z) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -2 \end{pmatrix}$$

is indefinite, and the origin is a saddle point.

A closer look uncovers the hidden structure of the saddle point. Moving on the xy-plane, we see that the origin is a minimum point for $f(x, y, 0) = x^2 + y^2$. At the same time, along the z-axis the origin is a maximum for $f(0, 0, z) = -z^2$. Thus

$$f(0,0,0) = \min_{(x,y)\in\mathbb{R}^2} f(x,y,0) = \max_{z\in\mathbb{R}} f(0,0,z)\,.$$

iii) A slightly more elaborate situation is provided by the function $f(x, y, z) = x^2 + y^3 - z^2$. Since $\nabla f(x, y, z) = (2x, 3y^2, -2z)$ the origin is again stationary. The Hessian

$$\boldsymbol{H}f(0,0,0) = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -2 \end{pmatrix}$$

is indefinite, making the origin a saddle point. More precisely, $(0, 0, 0)$ is a minimum if we move along the x-axis, a maximum along the z-axis, but on the y-axis we have an inflection point. □

The notion of **saddle points** can be generalised to subsume stationary points where the Hessian matrix is (positive or negative) semi-definite. In such a case, the mere knowledge of $\boldsymbol{H}f(\boldsymbol{x}_0)$ is not sufficient to determine the nature of the point. In order to have a genuine saddle point $\boldsymbol{x}_0$ one must additionally require that *there exist a direction along which f has a maximum point at $\boldsymbol{x}_0$ and a direction along which f has a minimum at $\boldsymbol{x}_0$*. Precisely, there should be vectors $\boldsymbol{v}_1$, $\boldsymbol{v}_2$ such that the maps $t \mapsto f(\boldsymbol{x}_0 + t\boldsymbol{v}_i)$, $i = 1, 2$, have a strict minimum and a strict maximum point respectively, for $t = 0$.

As an example take $f(x, y) = x^2 - y^4$. The origin is stationary, and the Hessian at that point is $\boldsymbol{H}f(0,0) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$, hence positive semi-definite. Since $f(x, 0) = x^2$ has a minimum at $x = 0$ and $f(0, y) = -y^4$ has a maximum at $y = 0$, the above requirement holds by taking $\boldsymbol{v}_1 = \boldsymbol{i} = (1, 0)$, $\boldsymbol{v}_2 = \boldsymbol{j} = (0, 1)$. We still call $\boldsymbol{x}_0$ a saddle point.

Consider now $f(x, y) = x^2 - y^3$, for which the origin is stationary the Hessian is the same as in the previous case. Despite this, for any $m \in \mathbb{R}$ the map $f(x, mx) = x^2 - m^3x^3$ has a minimum at $x = 0$, and $f(0, y) = -y^3$ has an inflection point at $y = 0$. Therefore no vector $\boldsymbol{v}_2$ exists that maximises $t \mapsto f(t\boldsymbol{v}_2)$ at $t = 0$. For this reason $\boldsymbol{x}_0 = \boldsymbol{0}$ will not be called a saddle point.

We note that if $\boldsymbol{x}_0$ is stationary and $\boldsymbol{H}f(\boldsymbol{x}_0)$ is positive semi-definite, then every eigenvector $\boldsymbol{w}$ associated to an eigenvalue $\lambda > 0$ ensures that $t \mapsto f(\boldsymbol{x}_0 + t\boldsymbol{w})$ has a strict minimum point at $t = 0$ (whence $\boldsymbol{w}$ can be chosen as vector $\boldsymbol{v}_1$); in fact, (5.17) gives

$$f(\boldsymbol{x}_0 + t\boldsymbol{w}) = f(\boldsymbol{x}_0) + \frac{1}{2}\lambda\|\boldsymbol{w}\|^2t^2 + o(t^2)\,, \qquad t \to 0\,.$$

Then $\boldsymbol{x}_0$ is a saddle point if and only if we can find a vector $\boldsymbol{v}_2$ in the kernel of the matrix $\boldsymbol{H}f(\boldsymbol{x}_0)$ such that $t \mapsto f(\boldsymbol{x}_0 + t\boldsymbol{v}_2)$ has a strict maximum at $t = 0$. Similarly when $\boldsymbol{H}f(\boldsymbol{x}_0)$ is negative semi-definite.

To conclude the chapter, we wish to provide the reader with a simple procedure for classifying stationary points in two variables. The whole point is to determine the eigenvalues' sign in the Hessian matrix, which can be done without effort for 2×2 matrices. If $\boldsymbol{A}$ is a symmetric matrix of order 2, the determinant is the product of the two eigenvalues. Therefore if $\det \boldsymbol{A} > 0$ the eigenvalues have the same sign and $\boldsymbol{A}$ is definite, positive or negative according to the sign of either diagonal term a_{11} or a_{22}. If $\det \boldsymbol{A} < 0$, the eigenvalues have different sign, making $\boldsymbol{A}$ indefinite. If $\det \boldsymbol{A} = 0$, the matrix is semi-definite, not definite.

Recycling this discussion with the Hessian $\boldsymbol{H}f(\boldsymbol{x}_0)$ of a stationary point, and bearing in mind Theorem 5.27, gives

$$\det \boldsymbol{H}f(\boldsymbol{x}_0) > 0 \ \Rightarrow \begin{cases} \boldsymbol{x}_0 \text{ strict local minimum point, if } \ f_{xx}(\boldsymbol{x}_0) > 0 \\ \boldsymbol{x}_0 \text{ strict local maximum point, if } \ f_{xx}(\boldsymbol{x}_0) < 0 \end{cases}$$

$$\det \boldsymbol{H}f(\boldsymbol{x}_0) < 0 \ \Rightarrow \ \boldsymbol{x}_0 \text{ saddle point}$$

$$\det \boldsymbol{H}f(\boldsymbol{x}_0) = 0 \ \Rightarrow \ \text{the nature of } \boldsymbol{x}_0 \text{ cannot be determined by } \boldsymbol{H}f(\boldsymbol{x}_0) \text{ only}\,.$$

In the first case $\det \boldsymbol{H}f(\boldsymbol{x}_0) > 0$, the dichotomy maximum vs. minimum may be sorted out by the sign of $f_{yy}(\boldsymbol{x}_0)$, which is also the sign of $f_{xx}(\boldsymbol{x}_0)$.

Example 5.33

The first derivatives of $f(x,y) = 2xy + \mathrm{e}^{-(x+y)^2}$ are

$$f_x(x,y) = 2\big(y - (x+y)\,\mathrm{e}^{-(x+y)^2}\big)\,, \qquad f_y(x,y) = 2\big(x - (x+y)\,\mathrm{e}^{-(x+y)^2}\big)\,,$$

and the second derivatives read

$$f_{xx}(x,y) = f_{yy}(x,y) = -2\mathrm{e}^{-(x+y)^2}\big(1 - 2(x+y)^2\big)\,,$$
$$f_{xy}(x,y) = f_{xy}(x,y) = 2 - 2\mathrm{e}^{-(x+y)^2}\big(1 - 2(x+y)^2\big)\,.$$

There are three stationary points

$$\boldsymbol{x}_0 = (0,0)\,, \qquad \boldsymbol{x}_1 = \big(\frac{1}{2}\sqrt{\log 2}, \frac{1}{2}\sqrt{\log 2}\big)\,, \qquad \boldsymbol{x}_2 = -\boldsymbol{x}_1\,,$$

and correspondingly,

$$\boldsymbol{H}f(\boldsymbol{x}_0) = \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix}, \quad \boldsymbol{H}f(\boldsymbol{x}_1) = \boldsymbol{H}f(\boldsymbol{x}_2) = \begin{pmatrix} 1 - 2\log 2 & 1 + 2\log 2 \\ 1 + 2\log 2 & 1 - 2\log 2 \end{pmatrix}.$$

Therefore $\boldsymbol{x}_0$ is a maximum point, for $\boldsymbol{H}f(\boldsymbol{x}_0)$ is negative definite, while $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ are saddle points since $\det \boldsymbol{H}f(\boldsymbol{x}_1) = \det \boldsymbol{H}f(\boldsymbol{x}_2) = -8\log 2 < 0$. □

5.7 Exercises

1. *Compute the first partial derivatives at the points indicated:*

   a) $f(x,y) = \sqrt{3x+y^2}$ at $(x_0,y_0) = (1,2)$

   b) $f(x,y,z) = y\mathrm{e}^{x+yz}$ at $(x_0,y_0,z_0) = (0,1,-1)$

   $\boxed{\text{c)}}$ $f(x,y) = 8x^2 + \displaystyle\int_1^y \mathrm{e}^{-t^2}\,\mathrm{d}t$ at $(x_0,y_0) = (3,1)$

2. *Compute the first partial derivatives of:*

   $\boxed{\text{a)}}$ $f(x,y) = \log(x+\sqrt{x^2+y^2})$

   $\boxed{\text{b)}}$ $f(x,y) = \displaystyle\int_x^y \cos t^2\,dt$

   $\boxed{\text{c)}}$ $f(x,y,z,t) = \dfrac{x-y}{z-t}$

   $\boxed{\text{d)}}$ $f(x_1,\ldots,x_n) = \sin(x_1+2x_2+\ldots+nx_n)$

3. *Compute the partial derivative indicated:*

   a) $f(x,y) = x^3y^2 - 3xy^4\,,$ $\quad f_{yyy}$

   b) $f(x,y) = x\sin y\,,$ $\quad \dfrac{\partial^3 f}{\partial x\partial y^2}$

   c) $f(x,y,z) = \mathrm{e}^{xyz}\,,$ $\quad f_{xyx}$

   d) $f(x,y,z) = x^a y^b z^c\,,$ $\quad \dfrac{\partial^6 f}{\partial x\partial y^2\partial z^3}$

4. *Determine which maps f satisfy $f_{xx}+f_{yy}=0$, known as Laplace equation:*

   a) $f(x,y) = x^2+y^2$ b) $f(x,y) = x^3+3xy^2$

   $\boxed{\text{c)}}$ $f(x,y) = \log\sqrt{x^2+y^2}$ $\boxed{\text{d)}}$ $f(x,y) = \mathrm{e}^{-x}\cos y - \mathrm{e}^{-y}\cos x$

$\boxed{5.}$ *Check that $f(x,t) = \mathrm{e}^{-t}\sin kx$ satisfies the so-called heat equation $f_t = \dfrac{1}{k^2}f_{xx}$.*

6. *Check that the following maps solve $f_{tt} = f_{xx}$, known as wave equation:*

   $\boxed{\text{a)}}$ $f(x,t) = \sin x\,\sin t$ $\boxed{\text{b)}}$ $f(x,t) = \sin(x-t)+\log(x+t)$

$\boxed{7.}$ *Given*

$$f(x,y) = \begin{cases} \dfrac{x^3y - xy^3}{x^2+y^2} & \text{if } (x,y) \neq (0,0)\,, \\ 0 & \text{if } (x,y) = (0,0)\,, \end{cases}$$

a) compute $f_x(x,y)$ and $f_y(x,y)$ for any $(x,y) \neq (0,0)$;
b) calculate $f_x(0,0)$, $f_y(0,0)$ using the definition of second partial derivative;
c) discuss the results obtained in the light of Schwarz's Theorem 5.17.

8. *Determine the gradient map of:*

a) $f(x,y) = \arctan \dfrac{x+y}{x-y}$ b) $f(x,y) = (x+y)\log(2x-y)$

c) $f(x,y,z) = \sin(x+y)\cos(y-z)$ d) $f(x,y,z) = (x+y)^z$

9. *Compute the directional derivatives along $\boldsymbol{v}$, at the indicated points:*

a) $f(x,y) = x\sqrt{y-3}$ $\boldsymbol{v} = (-1,6)$ $\boldsymbol{x}_0 = (2,12)$

b) $f(x,y,z) = \dfrac{1}{x+2y-3z}$ $\boldsymbol{v} = (12,-9,-4)$ $\boldsymbol{x}_0 = (1,1,-1)$

10. *Determine the tangent plane to the graph of $f(x,y)$ at the point $P_0 = (x_0, y_0, f(x_0,y_0))$:*

a) $f(x,y) = 3x^2 - y^2 + 3y$ *at* $P_0 = (-1,2,f(-1,2))$

$\boxed{\text{b)}}$ $f(x,y) = e^{y^2-x^2}$ *at* $P_0 = (-1,1,f(-1,1))$

c) $f(x,y) = x\log y$ *at* $P_0 = (4,1,f(4,1))$

11. *Relying on the definition, check the maps below are differentiable at the given point:*

$\boxed{\text{a)}}$ $f(x,y) = y\sqrt{x}$ *at* $(x_0,y_0) = (4,1)$

$\boxed{\text{b)}}$ $f(x,y) = |y|\log(1+x)$ *at* $(x_0,y_0) = (0,0)$

$\boxed{\text{c)}}$ $f(x,y) = xy - 3x^2$ *at* $(x_0,y_0) = (1,2)$

$\boxed{12.}$ *Given*

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

compute $f_x(0,0)$ and $f_y(0,0)$. Is f differentiable at the origin?

13. *Study the differentiability at* $(0,0)$ *of*

$$f(x,y) = \begin{cases} \dfrac{x^2y^3}{x^4+y^4} & \text{if } (x,y) \neq (0,0)\,, \\ 0 & \text{if } (x,y) = (0,0)\,. \end{cases}$$

$\boxed{14.}$ *Discuss the differentiability of*

$$f(x,y) = |x| \sin(x^2+y^2)$$

at any point of the plane.

15. *Study continuity and differentiability at the origin of*

$$f(x,y) = \begin{cases} |y|^\alpha \sin x & \text{if } y \neq 0\,, \\ 0 & \text{if } y = 0 \end{cases}$$

as α *varies in the reals.*

$\boxed{16.}$ *Study the differentiability at* $(0,0,0)$ *of*

$$f(x,y,z) = \begin{cases} (x^2+y^2+z^2)\sin \dfrac{1}{\sqrt{x^2+y^2+z^2}} & \text{if } (x,y,z) \neq (0,0,0)\,, \\ 0 & \text{if } (x,y,z) = (0,0,0)\,. \end{cases}$$

$\boxed{17.}$ *Given* $f(x,y) = x^2 + 3xy - y^2$*, find its differential at* $(x_0, y_0) = (2,3)$*. If* x *varies between* 2 *and* 2.05*, and* y *between* 3 *and* 2.96*, compare the increment* Δf *with the corresponding differential* $\mathrm{d}f_{(x_0,y_0)}$.

18. *Determine the differential at a generic point* (x_0, y_0) *of the functions:*

$\boxed{\text{a)}}$ $f(x,y) = \mathrm{e}^x \cos y$

b) $f(x,y) = x \sin xy$

c) $f(x,y,z) = \log(x^2+y^2+z^2)$

d) $f(x,y,z) = \dfrac{x}{y+2z}$

$\boxed{19.}$ *Find the differential of* $f(x,y,z) = x^3\sqrt{y^2+z^2}$ *at* $(2,3,4)$ *and then use it to approximate the number* $1.98^3\sqrt{3.01^2+3.97^2}$.

$\boxed{20.}$ *Tell whether*

$$f(x,y) = \frac{1}{x+y+1}$$

is Lipschitz over the rectangle $\mathcal{R} = [0,2] \times [0,1]$*; if yes compute the Lipschitz constant.*

21. *Determine if, how and where in* $\mathbb{R}^3$ *is*

$$f(x,y,z) = \mathrm{e}^{-(3x^2+2y^4+z^6)}$$

a Lipschitz map.

22. *Write the Taylor polynomial of order two for the following functions at the point given:*

$\boxed{\text{a)}}$ $f(x,y) = \cos x \, \cos y$ *at* $(x_0, y_0) = (0,0)$

b) $f(x,y,z) = \mathrm{e}^{x+y+z}$ *at* $(x_0, y_0, z_0) = (1,1,1)$

c) $f(x,y,z) = \cos(x+2y-3z)$ *at* $(x_0, y_0, z_0) = (0,0,1)$

23. *Determine the stationary points (if existent), specifying their nature:*

$\boxed{\text{a)}}$ $f(x,y) = x^2y + x^2 - 2y$

$\boxed{\text{b)}}$ $f(x,y) = y\log(x+y)$

$\boxed{\text{c)}}$ $f(x,y) = x + \frac{1}{6}x^6 + y^2(y^2-1)$

$\boxed{\text{d)}}$ $f(x,y) = xy\mathrm{e}^{-\frac{x}{5}-\frac{y}{6}}$

$\boxed{\text{e)}}$ $f(x,y) = \dfrac{x}{y} + \dfrac{8}{x} - y$

$\boxed{\text{f)}}$ $f(x,y) = 2y\log(2-x^2) + y^2$

$\boxed{\text{g)}}$ $f(x,y) = \mathrm{e}^{3x^2-6xy+2y^3}$

$\boxed{\text{h)}}$ $f(x,y) = y\log x$

$\boxed{\text{i)}}$ $f(x,y) = \log(x^2+y^2-1)$

$\boxed{\ell)}$ $f(x,y,z) = xyz + \dfrac{1}{x} + \dfrac{1}{y}\dfrac{1}{z}$

$\boxed{\text{24.}}$ *Determine and draw the domain of*

$$f(x,y) = \sqrt{y^2 - x^2}\,.$$

Find stationary points and extrema.

$\boxed{\text{25.}}$ *Determine domain and stationary points of*

$$f(x,y) = x^2\log(1+y) + x^2y^2\,.$$

What is their nature?

5.7.1 Solutions

1. *Partial derivatives of maps:*

a) $\dfrac{\partial f}{\partial x}(1,2) = \dfrac{3}{2\sqrt{7}}\,, \quad \dfrac{\partial f}{\partial y}(1,2) = \dfrac{2}{\sqrt{7}}\,.$

b) $\dfrac{\partial f}{\partial x}(0,1,-1) = \mathrm{e}^{-1}\,, \quad \dfrac{\partial f}{\partial y}(0,1,-1) = 0\,, \quad \dfrac{\partial f}{\partial z}(0,1,-1) = \mathrm{e}^{-1}\,.$

c) We have
$$\frac{\partial f}{\partial x}(x,y) = 16x \qquad \text{and} \qquad \frac{\partial f}{\partial y}(x,y) = \mathrm{e}^{-y^2}\,,$$
the latter computed by means of the Fundamental Theorem of Integral Calculus. Thus
$$\frac{\partial f}{\partial x}(3,1) = 48 \qquad \text{and} \qquad \frac{\partial f}{\partial y}(3,1) = \mathrm{e}^{-1}\,.$$

2. *Partial derivatives:*

a) We have
$$f_x(x,y) = \frac{1}{\sqrt{x^2+y^2}}\,, \qquad f_x(x,y) = \frac{1}{x+\sqrt{x^2+y^2}} \cdot \frac{y}{\sqrt{x^2+y^2}}\,.$$

b) Using the Fundamental Theorem of Integral Calculus we have
$$f_x(x,y) = -\frac{\partial}{\partial x}\int_y^x \cos t^2\,dt = -\cos x^2\,,$$
$$f_y(x,y) = \frac{\partial}{\partial y}\int_x^y \cos t^2\,dt = \cos y^2\,.$$

c) We have
$$f_x(x,y,z,t) = \frac{1}{z-t}\,, \qquad f_y(x,y,z,t) = \frac{1}{t-z}\,,$$
$$f_z(x,y,z,t) = \frac{y-x}{(z-t)^2}\,, \qquad f_t(x,y,z,t) = \frac{x-y}{(z-t)^2}\,.$$

d) We have
$$\frac{\partial f}{\partial x_1}(x_1,\ldots,x_n) = \cos(x_1+2x_2+\ldots+nx_n)\,,$$
$$\frac{\partial f}{\partial x_2}(x_1,\ldots,x_n) = 2\cos(x_1+2x_2+\ldots+nx_n)\,,$$
$$\vdots$$
$$\frac{\partial f}{\partial x_n}(x_1,\ldots,x_n) = n\cos(x_1+2x_2+\ldots+nx_n)\,, \ .$$

3. *Partial derivatives:*

a) $f_{yyy} = -72xy\,.$ b) $f_{yyx} = -\sin y\,.$ c) $f_{xyx} = yz^2\mathrm{e}^{xyz}(2+xyz)\,.$

d) $f_{zzzyyx} = abc(b-1)(c-1)(c-2)x^{a-1}y^{b-2}z^{c-3}\,.$

4. *Solutions of the Laplace equation:*

a) No. b) No.

c) Since $f(x,y) = \frac{1}{2}\log(x^2+y^2)$,

$$f_x = \frac{x}{x^2+y^2}\,, \qquad f_{xx} = \frac{y^2-x^2}{(x^2+y^2)^2}\,,$$

$$f_y = \frac{y}{x^2+y^2}\,, \qquad f_{yy} = \frac{x^2-y^2}{(x^2+y^2)^2}\,,$$

hence $f_{xx} + f_{yy} = 0\,,\ \forall x,y \neq 0$. Therefore the function is a solution of the Laplace equation on $\mathbb{R}^2 \setminus \{\mathbf{0}\}$.

d) Since

$$f_x = -\mathrm{e}^{-x}\cos y + \mathrm{e}^{-y}\sin x\,, \qquad f_{xx} = \mathrm{e}^{-x}\cos y + \mathrm{e}^{-y}\cos x\,,$$

$$f_y = -\mathrm{e}^{-x}\sin y + \mathrm{e}^{-y}\cos x\,, \qquad f_{yy} = -\mathrm{e}^{-x}\cos y - \mathrm{e}^{-y}\cos x\,,$$

we have $f_{xx} + f_{yy} = 0\,,\ \forall (x,y) \in \mathbb{R}^2$, and the function satisfies Laplace's equation on $\mathbb{R}^2$.

5. The assertion follows from

$$f_t = -\mathrm{e}^{-t}\sin kx\,, \qquad f_x = k\mathrm{e}^{-t}\cos kx\,, \qquad f_{xx} = -k^2\mathrm{e}^{-t}\sin kx\,.$$

6. *Solutions of the wave equation:*

a) From

$$f_x = \cos x\,\sin t\,, \qquad f_{xx} = -\sin x\,\sin t\,,$$

$$f_t = \sin x\,\cos t\,, \qquad f_{tt} = -\sin x\,\sin t$$

follows $f_{xx} = f_{tt}\,,\ \forall (x,t) \in \mathbb{R}^2$.

b) As

$$f_x = \cos(x-t) + \frac{1}{x+t}\,, \qquad f_{xx} = -\sin(x-t) - \frac{1}{(x+t)^2}\,,$$

$$f_t = -\cos(x-t) + \frac{1}{x+t}\,, \qquad f_{tt} = -\sin(x-t) - \frac{1}{(x+t)^2}\,,$$

we have $f_{xx} = f_{tt}\,,\ \forall (x,t) \in \mathbb{R}^2$ such that $x+t>0$.

7. a) Using the usual rules, for $(x,y) \neq (0,0)$ we have

$$f_x(x,y) = \frac{(3x^2y-y^3)(x^2+y^2) - 2x(x^3y-xy^3)}{(x^2+y^2)^2} = \frac{y(x^4+4x^2y^2-y^4)}{(x^2+y^2)^2}\,,$$

$$f_y(x,y) = \frac{(x^3-3xy^2)(x^2+y^2) - 2y(x^3y-xy^3)}{(x^2+y^2)^2} = \frac{x(x^4-4x^2y^2-y^4)}{(x^2+y^2)^2}\,.$$

b) To compute $f_x(0,0)$ and $f_y(0,0)$ let us resort to the definition:

$$f_x(0,0) = \lim_{x\to 0} \frac{f(x,0)-f(0,0)}{x} = \lim_{x\to 0} 0 = 0\,,$$
$$f_y(0,0) = \lim_{y\to 0} \frac{f(0,y)-f(0,0)}{y} = \lim_{y\to 0} 0 = 0\,.$$

Moreover,

$$f_{xy}(0,0) = \frac{\partial^2 f}{\partial y \partial x}(0,0) = \lim_{y\to 0} \frac{f_x(0,y)-f_x(0,0)}{y} = \lim_{y\to 0} -\frac{y^5}{y^5} = -1\,,$$
$$f_{yx}(0,0) = \frac{\partial^2 f}{\partial x \partial y}(0,0) = \lim_{x\to 0} \frac{f_y(x,0)-f_y(0,0)}{x} = \lim_{x\to 0} \frac{x^5}{x^5} = 1\,.$$

c) Schwarz's Theorem 5.17 does not apply, as the partial derivatives f_{xy}, f_{yx} are not continuous at $(0,0)$ (polar coordinates show that the limits $\lim\limits_{(x,y)\to(0,0)} f_{xy}(x,y)$ and $\lim\limits_{(x,y)\to(0,0)} f_{yx}(x,y)$ do not exist).

8. *Gradients:*

a) $\nabla f(x,y) = \left(-\dfrac{y}{x^2+y^2}\,,\ \dfrac{x}{x^2+y^2}\right)\,.$

b) $\nabla f(x,y) = \left(\log(2x-y) + \dfrac{2(x+y)}{2x-y}\,,\ \log(2x-y) - \dfrac{x+y}{2x-y}\right)\,.$

c) $\nabla f(x,y,z) = \big(\cos(x+y)\cos(y-z)\,,\ \cos(x+2y-z)\,,\ \sin(x+y)\sin(y-z)\big)\,.$

d) $\nabla f(x,y,z) = \big(z(x+y)^{z-1}\,,\ z(x+y)^{z-1}\,,\ (x+y)^z\log(x+y)\big)\,.$

9. *Directional derivatives:*

a) $\dfrac{\partial f}{\partial \boldsymbol{v}}(\boldsymbol{x}_0) = -1\,.$ b) $\dfrac{\partial f}{\partial \boldsymbol{v}}(\boldsymbol{x}_0) = -\dfrac{1}{6}\,.$

10. *Tangent planes:*

a) $z = -6x - y + 1\,.$

b) By (5.7), we compute

$$f_x(x,y) = -2x\mathrm{e}^{y^2-x^2}\,, \qquad f_y(x,y) = 2y\mathrm{e}^{y^2-x^2}\,,$$
$$f(-1,1) = 1\,, \qquad f_x(-1,1) = 2\,, \qquad f_y(-1,1) = 2\,.$$

The equation is thus

$$z = f(-1,1) + f_x(-1,1)(x+1) + f_y(-1,1)(y-1) = 2x + 2y + 1\,.$$

c) $z = 4y - 4\,.$

11. *Checking differentiability:*

a) From $f_x(x,y) = \frac{y}{2\sqrt{x}}$ and $f_y(x,y) = \sqrt{x}$ follows

$$f(4,1) = 2\,, \qquad f_x(4,1) = \frac{1}{4}\,, \qquad f_y(4,1) = 2\,.$$

Then f is differentiable at $(4,1)$ if and only if

$$\lim_{(x,y)\to(4,1)} \frac{f(x,y) - f(4,1) - f_x(4,1)(x-4) - f_y(4,1)(y-1)}{\sqrt{(x-4)^2+(y-1)^2}} = 0$$

i.e.,

$$L = \lim_{(x,y)\to(4,1)} \frac{y\sqrt{x} - 2 - \frac{1}{4}(x-4) - 2(y-1)}{\sqrt{(x-4)^2+(y-1)^2}} = 0\,.$$

Put $x = 4 + r\cos\theta$ e $y = 1 + r\sin\theta$ and observe that for $r \to 0$,

$$\sqrt{4 + r\cos\theta} = 2\sqrt{1 + \frac{r}{4}\cos\theta} = 2\left(1 + \frac{r}{8}\cos\theta + o(r)\right)\,,$$

so that

$$\begin{aligned} y\sqrt{x} - 2 - \frac{1}{4}(x-4) - 2(y-1) &= 2(1 + r\sin\theta)\left(1 + \frac{r}{8}\cos\theta + o(r)\right) + \\ &\quad -2 - \frac{1}{4}r\cos\theta - 2r\sin\theta = o(r)\,, \qquad r \to 0\,. \end{aligned}$$

Hence

$$L = \lim_{r\to 0} \frac{o(r)}{r} = \lim_{r\to 0} o(1) = 0\,.$$

b) Note $f(x,0) = f(0,y) = 0$, so $f_x(0,0) = f_y(0,0) = 0$. Differentiability at the origin is the same as proving

$$\lim_{(x,y)\to(0,0)} \frac{f(x,y) - f(0,0) - \nabla f(0,0)\cdot(x,y)}{\sqrt{x^2+y^2}} = 0\,,$$

i.e., given that $f(0,0) = 0$,

$$\lim_{(x,y)\to(0,0)} \frac{|y|\log(1+x)}{\sqrt{x^2+y^2}} = 0\,.$$

Polar coordinates come to the rescue:

$$\left|\frac{y\log(1+x)}{\sqrt{x^2+y^2}}\right| = \frac{r|\sin\theta\log(1+r\cos\theta)|}{r} \le 2r|\sin\theta\cos\theta| \le 2r\,;$$

Proposition 4.28 with $g(r) = 2r$ allows to conclude.

c) We have $f(1,2) = -1$, $f_x(1,2) = -4$ and $f_y(1,2) = 1$; f is differentiable at $(1,2)$ precisely when

$$L = \lim_{(x,y)\to(1,2)} \frac{xy - 3x^2 + 1 + 4(x-1) - (y-2)}{\sqrt{(x-1)^2 + (y-2)^2}} = 0\,.$$

Setting $x = 1 + r\cos\theta\,,\; y = 2 + r\sin\theta$ and computing,

$$L = \lim_{r\to 0} r\cos\theta(\sin\theta - 3\cos\theta).$$

The limit is zero by Proposition 4.28 with $g(r) = 4r$.

12. Observe $f(x,0) = f(0,y) = 0$, so

$$f_x(0,0) = \lim_{x\to 0} \frac{f(x,0) - f(0,0)}{x} = \lim_{x\to 0} 0 = 0\,,$$
$$f_y(0,0) = \lim_{y\to 0} \frac{f(0,y) - f(0,0)}{y} = \lim_{y\to 0} 0 = 0\,.$$

The map is certainly not differentiable at the origin because it is not even continuous; in fact $\lim\limits_{(x,y)\to(0,0)} f(x,y)$ does not exist, as one sees taking the limit along the coordinate axes,

$$\lim_{x\to 0} f(x,0) = \lim_{y\to 0} f(0,y) = 0\,,$$

and then along the line $y = x$,

$$\lim_{x\to 0} f(x,x) = \frac{1}{2}\,.$$

13. The function is not differentiable.

14. The map is certainly differentiable at all points $(x,y) \in \mathbb{R}^2$ with $x \neq 0$, by Proposition 5.8. To study the points on the axis, let us fix $(0, y_0)$ and compute the derivatives. No problems arise with f_y since

$$f_y(x,y) = 2|x|y\sin(x^2+y^2)\,, \qquad \forall (x,y) \in \mathbb{R}^2\,,$$

and $f_y(0,y_0) = 0$. As far as f_x is concerned,

$$\lim_{x\to 0} \frac{f(x,y_0) - f(0,y_0)}{x} = \lim_{x\to 0} \frac{|x|\sin(x^2+y_0^2)}{x}\,.$$

If $y_0 = \pm\sqrt{n\pi}$, with $n \in \mathbb{N}$, we have

$$f_x(0, \pm\sqrt{n\pi}) = \lim_{x\to 0} \frac{|x|}{x}(-1)^n \sin x^2 = 0\,,$$

otherwise
$$\lim_{x\to 0^+} \frac{|x|}{x} \sin(x^2+y_0^2) = \sin y_0^2\,,$$
while
$$\lim_{x\to 0^-} \frac{|x|}{x} \sin(x^2+y_0^2) = -\sin y_0^2\,.$$
Thus $f_x(0,y_0)$ exists only for $y_0 = \pm\sqrt{n\pi}$, where f is continuous and therefore differentiable.

15. The map is continuous if $\alpha \ge 0$; it is differentiable if $\alpha > 0$.

16. Observe $f(x,0,0) = x^2 \sin \frac{1}{|x|}$, hence $f_x(0,0,0) = 0$, since
$$f_x(0,0,0) = \lim_{x\to 0} \frac{f(x,0,0)-f(0,0,0)}{x} = \lim_{x\to 0} x \sin \frac{1}{|x|} = 0\,.$$
Similarly, $f_y(0,0,0) = f_z(0,0,0) = 0$. For differentiability at the origin we consider
$$\lim_{(x,y,z)\to(0,0,0)} \frac{f(x,y,z)}{\sqrt{x^2+y^2+z^2}} = \lim_{(x,y,z)\to(0,0,0)} \sqrt{x^2+y^2+z^2}\sin\frac{1}{\sqrt{x^2+y^2+z^2}}\,.$$
The limit is zero by the Squeeze rule:
$$0 \le \left|\sqrt{x^2+y^2+z^2}\sin\frac{1}{\sqrt{x^2+y^2+z^2}}\right| \le \sqrt{x^2+y^2+z^2}\,,$$
for all $(x,y,z) \neq (0,0,0)$.

17. Since $f_x(x,y) = 2x+3y\,,\ f_y(x,y) = 3x-2y$ we have
$$\mathrm{d}f_{(2,3)}(\Delta x,\Delta y) = \nabla f(2,3)\cdot(x-2,y-3) = 13(x-2)\,.$$
Let now $\boldsymbol{\Delta x} = (2.05-x_0, 2.96-y_0) = \left(\frac{5}{100}, -\frac{4}{100}\right)$, so
$$\Delta f = f(2.05, 2.96) - f(2,3) = 0.6449$$
and
$$\mathrm{d}f_{(2,3)}\left(\frac{5}{100}, -\frac{4}{100}\right) = 0.65\,.$$

18. *Differentials:*

a) As $f_x(x,y) = \mathrm{e}^x\cos y$ and $f_y(x,y) = -\mathrm{e}^x \sin y$, the differential is
$$\mathrm{d}f_{(x_0,y_0)}(\Delta x,\Delta y) = \nabla f(x_0,y_0)\cdot(\Delta x,\Delta y) = \mathrm{e}^{x_0}\cos y_0\,\Delta x - \mathrm{e}^{x_0}\sin y_0\,\Delta y\,.$$

b) $\mathrm{d}f_{(x_0,y_0)}(\Delta x,\Delta y) = (\sin x_0y_0 + x_0y_0\cos x_0y_0)\,\Delta x + x_0^2\cos x_0y_0\,\Delta y\,.$

c) $$\mathrm{d}f_{(x_0,y_0,z_0)}(\Delta x,\Delta y,\Delta z) = \frac{2x_0}{x_0^2+y_0^2+z_0^2}\,\Delta x + \frac{2y_0}{x_0^2+y_0^2+z_0^2}\,\Delta y + \\ + \frac{2z_0}{x_0^2+y_0^2+z_0^2}\,\Delta z\,.$$

d) $$\mathrm{d}f_{(x_0,y_0,z_0)}(\Delta x,\Delta y,\Delta z) = \frac{1}{y_0+2z_0}\,\Delta x - \frac{x_0}{(y_0+2z_0)^2}\,\Delta y - \frac{2x_0}{(y_0+2z_0)^2}\,\Delta z\,.$$

19. From

$$f_x(x,y,z) = 3x^2\sqrt{y^2+z^2}\,, \quad f_y(x,y,z) = \frac{x^3y}{\sqrt{y^2+z^2}}\,, \quad f_z(x,y,z) = \frac{x^3z}{\sqrt{y^2+z^2}}$$

follows

$$\mathrm{d}f_{(2,3,4)}(\Delta x,\Delta y,\Delta z) = 60\Delta x + \frac{24}{5}\Delta y + \frac{32}{5}\Delta z\,.$$

Set $\Delta \boldsymbol{x} = \big(-\frac{2}{100}, \frac{1}{100}, -\frac{3}{100}\big)$, so that

$$\mathrm{d}f_{(2,3,4)}\big(-\frac{2}{100}, \frac{1}{100}, -\frac{3}{100}\big) = -1.344\,.$$

By linearising, we may approximate $1.98^3\sqrt{3.01^2+3.97^2}$ by

$$f(2,3,4) + \mathrm{d}f_{(2,3,4)}\big(-\frac{2}{100}, \frac{1}{100}, -\frac{3}{100}\big) = 40 - 1.344 = 38.656\,.$$

20. First of all

$$\frac{\partial f}{\partial x}(x,y) = \frac{\partial f}{\partial y}(x,y) = -\frac{1}{(x+y+1)^2}\,,$$

and secondly

$$\sup_{(x,y)\in\mathcal{R}}\Big|\frac{\partial f}{\partial x}(x,y)\Big| = \sup_{(x,y)\in\mathcal{R}}\Big|\frac{\partial f}{\partial x}(x,y)\Big| = 1\,;$$

Proposition 5.16 then tells us f is Lipschitz on $\mathcal{R}$, with $L=\sqrt{2}$.

21. The map is Lipschitz on the entire $\mathbb{R}^3$.

22. *Taylor polynomials:*

a) We have to find the Taylor polynomial for f at $\boldsymbol{x}_0$ of order 2

$$Tf_{2,\boldsymbol{x}_0}(\boldsymbol{x}) = f(\boldsymbol{x}_0) + \nabla f(\boldsymbol{x}_0)\cdot(\boldsymbol{x}-\boldsymbol{x}_0) + \frac{1}{2}(\boldsymbol{x}-\boldsymbol{x}_0)\cdot \boldsymbol{H}f(\boldsymbol{x}_0)(\boldsymbol{x}-\boldsymbol{x}_0)\,.$$

Let us begin by computing the partial derivatives involved:

$$\begin{aligned}
&f_x(x,y) = -\sin x\cos y && \text{so } f_x(0,0)=0\\
&f_y(x,y) = -\cos x\sin y && \text{so } f_y(0,0)=0\\
&f_{xx}(x,y) = -\cos x\cos y && \text{so } f_{xx}(0,0)=-1\\
&f_{yy}(x,y) = -\cos x\cos y && \text{so } f_{yy}(0,0)=-1\\
&f_{xy}(x,y) = f_{yx}(x,y) = \sin x\sin y && \text{so } f_{xy}(0,0)=f_{yx}(0,0)=0
\end{aligned}$$

As $f(0,0)=1$,

$$\begin{aligned} Tf_{2,(0,0)}(x,y) &= 1+\frac{1}{2}(x,y)\cdot\begin{pmatrix}-1 & 0\\ 0 & -1\end{pmatrix}\cdot\begin{pmatrix}x\\ y\end{pmatrix}\\ &= 1+\frac{1}{2}(x,y)\cdot(-x,-y) = 1-\frac{1}{2}x^2-\frac{1}{2}y^2\,. \end{aligned}$$

Alternatively, we might want to recall that, for $x\to 0$ and $y\to 0$,

$$\cos x = 1-\frac{1}{2}x^2+o(x^2)\,,\qquad \cos y = 1-\frac{1}{2}y^2+o(y^2)\,;$$

which we have to multiply. As the Taylor polynomial is unique, we find immediately $Tf_{2,(0,0)}(x,y)=1-\frac{1}{2}x^2-\frac{1}{2}y^2$.

b) Computing directly,

$$\begin{aligned} Tf_{2,(1,1,1)}(x,y,z) = \mathrm{e}^3\Big(1+(x-1)+(y-1)+(z-1)&+\\ +\frac{1}{2}\big((x-1)^2+(y-1)^2+(z-1)^2\big)&+\\ +(x-1)(y-1)+(x-1)(z-1)+(y-1)(z-1)&\Big)\,. \end{aligned}$$

c) We have

$$\begin{aligned} Tf_{2,(0,0,1)}(x,y,z) = \cos 3+\sin 3\big(x+2y-3(z-1)\big)&+\\ +\frac{1}{2}\cos 3\big(-x^2-4y^2-9(z-1)^2\big)&+\\ +\cos 3\big(-2xy+6x(z-1)+6y(z-1)\big)&\,. \end{aligned}$$

23. *Stationary points and type:*

a) From

$$\frac{\partial f}{\partial x}(x,y)=2x(y+1)\,,\qquad \frac{\partial f}{\partial y}(x,y)=x^2-2$$

and the condition $\nabla f(x,y)=\mathbf{0}$ we obtain the stationary points $P_1=(\sqrt{2},-1)$, $P_2=(-\sqrt{2},-1)$. Since

$$\frac{\partial^2 f}{\partial x^2}(x,y)=2(y+1)\,,\qquad \frac{\partial^2 f}{\partial y^2}(x,y)=0\,,$$
$$\frac{\partial^2 f}{\partial x\partial y}(x,y)=\frac{\partial^2 f}{\partial y\partial x}(x,y)=2x\,,$$

the Hessians at those points read

$$\boldsymbol{H}f(P_1)=\begin{pmatrix}0 & 2\sqrt{2}\\ 2\sqrt{2} & 0\end{pmatrix}\,,\qquad \boldsymbol{H}f(P_2)=\begin{pmatrix}0 & -2\sqrt{2}\\ -2\sqrt{2} & 0\end{pmatrix}\,.$$

In either case the determinant is negative, so P_1,P_2 are saddle points.

b) The function is defined for $x+y>0$, i.e., on the half-plane $\operatorname{dom} f = \{(x,y) \in \mathbb{R}^2 : y > -x\}$. As

$$\frac{\partial f}{\partial x}(x,y) = \frac{y}{x+y}\,, \qquad \frac{\partial f}{\partial y}(x,y) = \log(x+y) + \frac{y}{x+y}\,,$$

imposing $\nabla f(x,y) = \mathbf{0}$ yields the system

$$\begin{cases} \dfrac{y}{x+y} = 0 \\ \log(x+y) + \dfrac{y}{x+y} = 0 \end{cases}$$

whose unique solution is the stationary point $P = (1,0) \in \operatorname{dom} f$. We have then

$$\frac{\partial^2 f}{\partial x^2}(x,y) = -\frac{y}{(x+y)^2}\,, \qquad \frac{\partial^2 f}{\partial y^2}(x,y) = \frac{1}{x+y} + \frac{x}{(x+y)^2}\,,$$
$$\frac{\partial^2 f}{\partial x \partial y}(x,y) = \frac{\partial^2 f}{\partial y \partial x}(x,y) = \frac{x}{(x+y)^2}\,,$$

so

$$\boldsymbol{H}f(P) = \begin{pmatrix} 0 & 1 \\ 1 & 2 \end{pmatrix}.$$

The Hessian determinant is $-1 < 0$, making P a saddle point.

c) From

$$\frac{\partial f}{\partial x}(x,y) = 1 + x^5\,, \qquad \frac{\partial f}{\partial y}(x,y) = 4y^3 - 2y$$

and $\nabla f(x,y) = \mathbf{0}$ we find three stationary points

$$P_1 = (-1,0)\,, \qquad P_2 = \Big(-1, \frac{\sqrt{2}}{2}\Big)\,, \qquad P_3 = \Big(-1, -\frac{\sqrt{2}}{2}\Big)\,.$$

Then

$$\frac{\partial^2 f}{\partial x^2}(x,y) = 5x^4\,, \quad \frac{\partial^2 f}{\partial x \partial y}(x,y) = \frac{\partial^2 f}{\partial y \partial x}(x,y) = 0\,, \quad \frac{\partial^2 f}{\partial y^2}(x,y) = 12y^2 - 2\,.$$

Consequently,

$$\boldsymbol{H}f(-1,0) = \begin{pmatrix} 5 & 0 \\ 0 & -2 \end{pmatrix}, \quad \boldsymbol{H}f\Big(-1, \frac{\sqrt{2}}{2}\Big) = \boldsymbol{H}f\Big(-1, -\frac{\sqrt{2}}{2}\Big) = \begin{pmatrix} 5 & 0 \\ 0 & 4 \end{pmatrix}.$$

The Hessians are diagonal; the one of P_1 is indefinite, so P_1 is a saddle point; the other two are positive definite, and P_2, P_3 are local minimum points.

d) The partial derivatives read

$$\frac{\partial f}{\partial x}(x,y) = y\left(1-\frac{x}{5}\right)\mathrm{e}^{-\frac{x}{5}-\frac{y}{6}}, \qquad \frac{\partial f}{\partial y}(x,y) = x\left(1-\frac{y}{6}\right)\mathrm{e}^{-\frac{x}{5}-\frac{y}{6}}.$$

The equation $\nabla f(x,y) = \mathbf{0}$ implies

$$\begin{cases} y\left(1-\frac{x}{5}\right) = 0 \\ x\left(1-\frac{y}{6}\right) = 0\,, \end{cases}$$

hence $P_1 = (0,0)$ and $P_2 = (5,6)$ are the stationary points. Now,

$$\begin{aligned}
&\frac{\partial^2 f}{\partial x^2}(x,y) = \frac{1}{5}y\left(\frac{x}{5}-2\right)\mathrm{e}^{-\frac{x}{5}-\frac{y}{6}}, \\
&\frac{\partial^2 f}{\partial x \partial y}(x,y) = \frac{\partial^2 f}{\partial y \partial x}(x,y) = \left(1-\frac{x}{5}\right)\left(1-\frac{y}{6}\right)\mathrm{e}^{-\frac{x}{5}-\frac{y}{6}}, \\
&\frac{\partial^2 f}{\partial y^2}(x,y) = \frac{1}{6}x\left(\frac{y}{6}-2\right)\mathrm{e}^{-\frac{x}{5}-\frac{y}{6}},
\end{aligned}$$

so that

$$\boldsymbol{H}f(0,0) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \boldsymbol{H}f(5,6) = \begin{pmatrix} -\frac{6}{5}\mathrm{e}^{-2} & 0 \\ 0 & -\frac{5}{6}\mathrm{e}^{-2} \end{pmatrix}.$$

Since $\det \boldsymbol{H}f(0,0) = -1 < 0$, P_1 is a saddle point for f, while $\det \boldsymbol{H}f(5,6) = \mathrm{e}^{-4} > 0$ and $\dfrac{\partial^2 f}{\partial x^2}(5,6) < 0$ mean P_2 is a relative maximum point. This last fact could also have been established by noticing both eigenvectors are negative, so the matrix is negative definite.

e) The map is defined on the plane minus the coordinate axes $x = 0$, $y = 0$. As

$$\frac{\partial f}{\partial x}(x,y) = \frac{1}{y} - \frac{8}{x^2}, \qquad \frac{\partial f}{\partial y}(x,y) = -\frac{x}{y^2} - 1,$$

$\nabla f(x,y) = \mathbf{0}$ has one solution $P = (-4,2)$ only. What this point is is readily said, for

$$\frac{\partial^2 f}{\partial x^2}(x,y) = \frac{16}{x^3}, \quad \frac{\partial^2 f}{\partial x \partial y}(x,y) = \frac{\partial^2 f}{\partial y \partial x}(x,y) = -\frac{1}{y^2}, \quad \frac{\partial^2 f}{\partial y^2}(x,y) = \frac{2x}{y^3},$$

and

$$\boldsymbol{H}f(-4,2) = \begin{pmatrix} -\frac{1}{4} & -\frac{1}{4} \\ -\frac{1}{4} & -1 \end{pmatrix}.$$

Since $\det \boldsymbol{H}f(-4,2) = \dfrac{3}{16} > 0$ and $\dfrac{\partial^2 f}{\partial x^2}(-4,2) = -\dfrac{1}{4} < 0$, P is a relative maximum.

f) The function is defined on

$$\operatorname{dom} f = \{(x, y) \in \mathbb{R}^2 : 2 - x^2 > 0\},$$

which is the horizontal strip between the lines $y = \pm\sqrt{2}$. From

$$\frac{\partial f}{\partial x}(x, y) = -\frac{4xy}{2 - x^2}, \qquad \frac{\partial f}{\partial y}(x, y) = 2\log(2 - x^2) + 2y,$$

equation $\nabla f(x, y) = \mathbf{0}$ gives points $P_1 = (1, 0)$, $P_2 = (-1, 0)$, $P_3 = (0, -\log 2)$. The second derivatives read:

$$\frac{\partial^2 f}{\partial x^2}(x, y) = -\frac{4y(x^2 + 2)}{(2 - x^2)^2}, \qquad \frac{\partial^2 f}{\partial y^2}(x, y) = 2,$$
$$\frac{\partial^2 f}{\partial x \partial y}(x, y) = \frac{\partial^2 f}{\partial y \partial x}(x, y) = -\frac{4x}{2 - x^2},$$

so

$$\boldsymbol{H}f(1, 0) = \begin{pmatrix} 0 & -4 \\ -4 & 2 \end{pmatrix}, \quad \boldsymbol{H}f(-1, 0) = \begin{pmatrix} 0 & 4 \\ 4 & 2 \end{pmatrix},$$

$$\boldsymbol{H}f(0, -\log 2) = \begin{pmatrix} 2\log 2 & 0 \\ 0 & 2 \end{pmatrix}.$$

As $\det \boldsymbol{H}f(1, 0) = \det \boldsymbol{H}f(-1, 0) = -16 < 0$, P_1 and P_2 are saddle points; P_3 is a relative minimum because the Hessian $\boldsymbol{H}f(P_3)$ is positive definite.

g) Using

$$\frac{\partial f}{\partial x}(x, y) = 6(x - y)\mathrm{e}^{3x^2 - 6xy + 2y^3}, \qquad \frac{\partial f}{\partial y}(x, y) = 6(y^2 - x)\mathrm{e}^{3x^2 - 6xy + 2y^3},$$

$\nabla f(x, y) = \mathbf{0}$ produces two stationary points $P_1 = (0, 0)$, $P_2 = (1, 1)$. As for second-order derivatives,

$$\frac{\partial^2 f}{\partial x^2}(x, y) = 6\big(1 + 6(x - y)^2\big)\mathrm{e}^{3x^2 - 6xy + 2y^3},$$
$$\frac{\partial^2 f}{\partial x \partial y}(x, y) = \frac{\partial^2 f}{\partial y \partial x}(x, y) = 6\big(-1 + 6(y^2 - x)(x - y)\big)\mathrm{e}^{3x^2 - 6xy + 2y^3},$$
$$\frac{\partial^2 f}{\partial y^2}(x, y) = 6\big(2y + 6(y^2 - x)^2\big)\mathrm{e}^{3x^2 - 6xy + 2y^3},$$

so

$$\boldsymbol{H}f(0, 0) = \begin{pmatrix} 6 & -6 \\ -6 & 0 \end{pmatrix}, \qquad \boldsymbol{H}f(1, 1) = \begin{pmatrix} 6\mathrm{e}^{-1} & -6\mathrm{e}^{-1} \\ -6\mathrm{e}^{-1} & 12\mathrm{e}^{-1} \end{pmatrix}.$$

The first is a saddle point, for $\det \boldsymbol{H}f(0, 0) = -36 < 0$; as for the other point, $\det \boldsymbol{H}f(1, 1) = 36\mathrm{e}^{-1} > 0$ and $\dfrac{\partial^2 f}{\partial x^2}(1, 1) = 6\mathrm{e}^{-1} > 0$ imply P_2 is a relative minimum.

h) The function is defined on $x > 0$. The derivatives

$$\frac{\partial f}{\partial x}(x,y) = \frac{y}{x}\,, \qquad \frac{\partial f}{\partial y}(x,y) = \log x$$

are zero at the point $P = (1,0)$. Since

$$\frac{\partial^2 f}{\partial x^2}(x,y) = -\frac{y}{x^2}\,, \qquad \frac{\partial^2 f}{\partial x \partial y}(x,y) = \frac{\partial^2 f}{\partial y \partial x}(x,y) = \frac{1}{x}\,, \qquad \frac{\partial^2 f}{\partial y^2}(x,y) = 0\,,$$

we have

$$\boldsymbol{H}f(1,0) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

and $\det \boldsymbol{H}f(1,0) = -1 < 0$, telling that P is a saddle point.

i) There are no stationary points because

$$\nabla f(x,y) = \Big(\frac{2x}{x^2+y^2-1}, \frac{2y}{x^2+y^2-1}\Big)$$

vanishes only at $(0,0)$, which does not belong to the domain of f; in fact $\operatorname{dom} f = \{(x,y) \in \mathbb{R}^2 : x^2+y^2 > 1\}$ consists of the points lying outside the unit circle centred in the origin.

ℓ) Putting

$$\frac{\partial f}{\partial x}(x,y,z) = yz - \frac{1}{x^2}\,, \quad \frac{\partial f}{\partial y}(x,y,z) = xz - \frac{1}{y^2}\,, \quad \frac{\partial f}{\partial z}(x,y,z) = xy - \frac{1}{z^2}$$

all equal 0 gives $P_1 = (1,1,1)$ and $P_2 = -P_1$. Moreover,

$$\begin{aligned} &f_{xx}(x,y) = \frac{2}{x^3}\,, && f_{yy}(x,y) = \frac{2}{y^3}\,, && f_{zz}(x,y) = \frac{2}{z^3} \\ &f_{xy}(x,y) = f_{yx} = z\,, && f_{xz}(x,y) = f_{zx} = y\,, && f_{yz}(x,y) = f_{zy} = x\,, \end{aligned}$$

so the Hessians $\boldsymbol{H}f(P_1)$, $\boldsymbol{H}f(P_2)$ are positive definite (the eigenvalues are $\lambda_1 = 1$, with multiplicity 2, and $\lambda_2 = 4$ in both cases). Therefore P_1 and P_2 are local minima for f.

24. The domain reads

$$\operatorname{dom} f = \{(x,y) \in \mathbb{R}^2 : y^2 - x^2 \geq 0\}\,.$$

The inequality $y^2 - x^2 \geq 0$ is $(y-x)(y+x) \geq 0$, satisfied if the factors $(y-x)$, $(y+x)$ have the same sign. Thus

$$\operatorname{dom} f = \{(x,y) \in \mathbb{R}^2 : y \geq x \text{ and } y \geq -x\} \cup \{(x,y) \in \mathbb{R}^2 : y \leq x \text{ and } y \leq -x\}\,,$$

see Fig. 5.10.

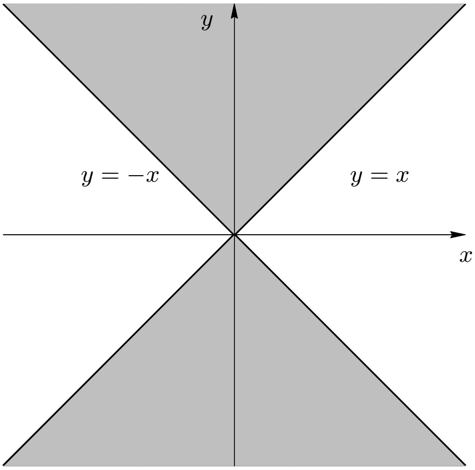


Figure 5.10. Domain of $f(x,y) = \sqrt{y^2 - x^2}$

The first derivatives are

$$\frac{\partial f}{\partial x}(x,y) = -\frac{x}{\sqrt{y^2 - x^2}}\,, \qquad \frac{\partial f}{\partial y}(x,y) = \frac{y}{\sqrt{y^2 - x^2}}\,.$$

The lines $y = x$, $y = -x$ are not contained in the domain of the partial derivatives, preventing the possibility of having stationary points.

On the other hand it is easy to see that

$$f(x,x) = f(x,-x) = 0 \qquad \text{and} \qquad f(x,y) \ge 0\,,\ \forall (x,y) \in \operatorname{dom} f\,.$$

All points on $y = x$ and $y = -x$, i.e., those of coordinates (x,x) and $(x,-x)$, are therefore absolute minima for f.

25. First of all the function is defined on

$$\operatorname{dom} f = \{(x,y) \in \mathbb{R}^2 : 1 + y > 0\}\,,$$

which is the open half-plane determined by the line $y = -1$.

Secondly, the points annihilating the gradient function

$$\nabla f(x,y) = \left(\frac{\partial f}{\partial x}(x,y), \frac{\partial f}{\partial y}(x,y)\right) = \left(2x\log(1+y) + 2xy^2, \frac{x^2}{1+y} + 2x^2 y\right)$$

are the solutions of

$$\begin{cases} 2x\left(\log(1+y) + y^2\right) = 0 \\ x^2\left(\dfrac{1}{1+y} + 2y\right) = 0\,. \end{cases}$$

We find $(0, y)$, with $y > -1$ arbitrary. Thirdly, we compute the second derivatives

$$\begin{aligned}
&\frac{\partial^2 f}{\partial x^2}(x, y) = 2\left(\log(1+y) + y^2\right) ,\\
&\frac{\partial^2 f}{\partial x \partial y}(x, y) = \frac{\partial^2 f}{\partial y \partial x}(x, y) = 2x\left(\frac{1}{1+y} + 2y\right) ,\\
&\frac{\partial^2 f}{\partial y^2}(x, y) = x^2\left(2 - \frac{1}{(1+y)^2}\right) .
\end{aligned}$$

These tell that the Hessian at the stationary points is

$$\boldsymbol{H}f(0, y) = \begin{pmatrix} 2\left(\log(1+y) + y^2\right) & 0 \\ 0 & 0 \end{pmatrix} ,$$

which unfortunately does not help to understand the points' nature.

This can be accomplished by direct inspection of the function. Write $f(x, y) = \alpha(x)\beta(y)$ with $\alpha(x) = x^2$ and $\beta(y) = \log(1+y)+y^2$. Note also that $f(0, y) = 0$, for any $y > -1$. It is not difficult to see that $\beta(y) > 0$ when $y > 0$, and $\beta(y) < 0$ when $y < 0$ (just compare the graphs of the elementary functions $\varphi(y) = \log(1 + y)$ and $\psi(y) = -y^2$). For any (x, y) in a suitable neighbourhood of $(0, y)$ then,

$$f(x, y) \geq 0 \qquad \text{if} \quad y > 0 \qquad \text{and} \qquad f(x, y) \leq 0 \qquad \text{if} \quad y < 0 .$$

In conclusion, for $y > 0$ the points $(0, y)$ are relative minima, whereas for $y < 0$ they are relative maxima. The origin is neither.

6

Differential calculus for vector-valued functions

In resuming the study of vector-valued functions started in Chapter 4, we begin by the various definitions concerning differentiability, and introduce the Jacobian matrix, which gathers the gradients of the function's components, and the basic differential operators of order one and two. Then we will present the tools of differential calculus; among them, the so-called chain rule for differentiating composite maps has a prominent role, for it lies at the core of the idea of coordinate-system changes. After discussing the general theory, we examine in detail the special, but of the foremost importance, frame systems of polar, cylindrical, and spherical coordinates.

The second half of the chapter devotes itself to regular, or piecewise-regular, curves and surfaces, from a differential point of view. The analytical approach, that focuses on the functions, gradually gives way to the geometrically-intrinsic aspects of curves and surfaces as objects in the plane or in space. The fundamental vectors of a curve (the tangent, normal, binormal vectors and the curvature) are defined, and we show how to choose one of the two possible orientations of a curve. For surfaces we introduce the normal vector and the tangent plane, then discuss the possibility of fixing a way to cross the surface, which leads to the dichotomy between orientable and non-orientable surfaces, plus the notions of boundary and closed surface. All this will be the basis upon which to build, in Chapter 9, an integral calculus on curves and surfaces, and to establish the paramount Theorems of Gauss, Green and Stokes.

6.1 Partial derivatives and Jacobian matrix

Given $\boldsymbol{x}_0 \in \operatorname{dom} \boldsymbol{f}$, suppose every component f_i of $\boldsymbol{f}$ admits at $\boldsymbol{x}_0$ all first partial derivatives $\dfrac{\partial f_i}{\partial x_j}$, $j = 1, \ldots, n$, so that to have the gradient vector

$$\nabla f_i(\boldsymbol{x}_0) = \Big(\frac{\partial f_i}{\partial x_j}(\boldsymbol{x}_0)\Big)_{1 \le j \le n} = \Big(\frac{\partial f_i}{\partial x_1}(\boldsymbol{x}_0), \ldots, \frac{\partial f_i}{\partial x_n}(\boldsymbol{x}_0)\Big),$$

here written as row vector.

C. Canuto, A. Tabacco: *Mathematical Analysis II,* 2nd Ed.,
UNITEXT – La Matematica per il 3+2 85, DOI 10.1007/978-3-319-12757-6_6,

Definition 6.1 *The matrix with m rows and n columns*

$$\boldsymbol{J f}(\boldsymbol{x}_0) = \Big(\frac{\partial f_i}{\partial x_j}(\boldsymbol{x}_0)\Big)_{\substack{1 \le i \le m \\ 1 \le j \le n}} = \begin{pmatrix} \nabla f_1(\boldsymbol{x}_0) \\ \vdots \\ \nabla f_m(\boldsymbol{x}_0) \end{pmatrix}$$

is called **Jacobian matrix** *of* $\boldsymbol{f}$ *at the point* $\boldsymbol{x}_0$.

The Jacobian (matrix) is also indicated by $\boldsymbol{J}_{\boldsymbol{f}}(\boldsymbol{x}_0)$ or $D\boldsymbol{f}(\boldsymbol{x}_0)$.
In particular if $\boldsymbol{f} = f$ is a scalar map ($m = 1$), then

$$\boldsymbol{J f}(\boldsymbol{x}_0) = \nabla f(\boldsymbol{x}_0)\,.$$

Examples 6.2

i) The function $\boldsymbol{f} : \mathbb{R}^3 \to \mathbb{R}^2$, $\boldsymbol{f}(x,y,z) = xyz\,\boldsymbol{i} + (x^2+y^2+z^2)\,\boldsymbol{j}$ has components $f_1(x,y,z) = xyz$ and $f_2(x,y,z) = x^2+y^2+z^2$, whose partial derivatives we write as entries of

$$\boldsymbol{J f}(x,y,z) = \begin{pmatrix} yz & xz & xy \\ 2x & 2y & 2z \end{pmatrix}.$$

ii) Consider

$$\boldsymbol{f} : \mathbb{R}^n \to \mathbb{R}^n\,, \qquad \boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}\,,$$

where $\boldsymbol{A} = (a_{ij})_{\substack{1\le i\le m \\ 1\le j\le n}} \in \mathbb{R}^{m\times n}$ is an $m \times n$ matrix and $\boldsymbol{b} = (b_i)_{1\le i\le m} \in \mathbb{R}^m$.
The ith component of $\boldsymbol{f}$ is

$$f_i(\boldsymbol{x}) = \sum_{j=1}^{n} a_{ij}x_j + b_i\,,$$

whence $\dfrac{\partial f_i}{\partial x_j}(\boldsymbol{x}) = a_{ij}$ for all $j = 1,\ldots,n$ $\boldsymbol{x} \in \mathbb{R}^n$. Therefore $\boldsymbol{J f}(\boldsymbol{x}) = \boldsymbol{A}$. □

6.2 Differentiability and Lipschitz functions

Now we shall see if and how the previous chapter's results extend to vector-valued functions. Starting from differentiability, let us suppose each component of $\boldsymbol{f}$ is differentiable at $\boldsymbol{x}_0 \in \operatorname{dom} \boldsymbol{f}$ (see Definition 5.5),

$$f_i(\boldsymbol{x}) = f_i(\boldsymbol{x}_0) + \nabla f_i(\boldsymbol{x}_0)\cdot(\boldsymbol{x}-\boldsymbol{x}_0) + o(\|\boldsymbol{x}-\boldsymbol{x}_0\|)\,, \qquad \boldsymbol{x} \to \boldsymbol{x}_0$$

for any $i = 1,\ldots,n$. The dot product $\nabla f_i(\boldsymbol{x}_0)\cdot(\boldsymbol{x}-\boldsymbol{x}_0)$ is to be thought of as a matrix product between the row vector $\nabla f_i(\boldsymbol{x}_0)$ and the column vector $\boldsymbol{x}-\boldsymbol{x}_0$.

By definition of Jacobian matrix, we may write, vectorially,

$$\boxed{\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{x}_0) + \boldsymbol{J}\boldsymbol{f}(\boldsymbol{x}_0)(\boldsymbol{x}-\boldsymbol{x}_0) + o(\|\boldsymbol{x}-\boldsymbol{x}_0\|)\,, \qquad \boldsymbol{x} \to \boldsymbol{x}_0\,.} \tag{6.1}$$

One says then $\boldsymbol{f}$ is **differentiable at** $\boldsymbol{x}_0$.

Let us consider a special case: putting $\Delta\boldsymbol{x} = \boldsymbol{x} - \boldsymbol{x}_0$, we rewrite the above as

$$\boldsymbol{f}(\boldsymbol{x}_0 + \Delta\boldsymbol{x}) = \boldsymbol{f}(\boldsymbol{x}_0) + \boldsymbol{J}\boldsymbol{f}(\boldsymbol{x}_0)\Delta\boldsymbol{x} + o(\|\Delta\boldsymbol{x}\|)\,, \qquad \Delta\boldsymbol{x} \to 0\,;$$

the linear map $\mathbf{d}\boldsymbol{f}_{\boldsymbol{x}_0}$ from $\mathbb{R}^n$ to $\mathbb{R}^m$ defined by

$$\boxed{\mathbf{d}\boldsymbol{f}_{\boldsymbol{x}_0} : \Delta\boldsymbol{x} \mapsto \boldsymbol{J}\boldsymbol{f}(\boldsymbol{x}_0)\Delta\boldsymbol{x}}$$

is called **differential** of $\boldsymbol{f}$ at $\boldsymbol{x}_0$. Up to infinitesimals of order greater than one, the formula says the increment $\Delta\boldsymbol{f} = \boldsymbol{f}(\boldsymbol{x}_0+\Delta\boldsymbol{x}) - \boldsymbol{f}(\boldsymbol{x}_0)$ is approximated by the value of the differential $\mathbf{d}\boldsymbol{f}_{\boldsymbol{x}_0} = \boldsymbol{J}\boldsymbol{f}(\boldsymbol{x}_0)\Delta\boldsymbol{x}$.

As for scalar functions, equation (6.1) **linearises** $\boldsymbol{f}$ at $\boldsymbol{x}_0$, written

$$\boxed{\boldsymbol{f}(\boldsymbol{x}) \sim \boldsymbol{f}(\boldsymbol{x}_0) + \boldsymbol{J}\boldsymbol{f}(\boldsymbol{x}_0)(\boldsymbol{x}-\boldsymbol{x}_0)}$$

on a neighbourhood of $\boldsymbol{x}_0$; in other words it approximates $\boldsymbol{f}$ by means of a degree-one polynomial in $\boldsymbol{x}$ (the Taylor polynomial of order 1 at $\boldsymbol{x}_0$).

Propositions 5.7 and 5.8 carry over, as one sees by taking one component at a time.

Also vectorial functions can be Lipschitz. The next statements generalise Definition 5.14 and Proposition 5.16. Let $\mathcal{R}$ be a region inside $\operatorname{dom}\boldsymbol{f}$.

Definition 6.3 *The map* $\boldsymbol{f}$ *is* **Lipschitz** *on* $\mathcal{R}$ *if there is a constant* $L \geq 0$ *such that*

$$\|\boldsymbol{f}(\boldsymbol{x}_1) - \boldsymbol{f}(\boldsymbol{x}_2)\| \leq L\|\boldsymbol{x}_1 - \boldsymbol{x}_2\|\,, \qquad \forall \boldsymbol{x}_1, \boldsymbol{x}_2 \in \mathcal{R}\,. \tag{6.2}$$

The smallest such L *is the* **Lipschitz constant** *of* $\boldsymbol{f}$ *on* $\mathcal{R}$*.*

Clearly $\boldsymbol{f}$ is Lipschitz on $\mathcal{R}$ if and only if all its components are.

Proposition 6.4 *Let* $\mathcal{R}$ *be a connected region in* $\operatorname{dom}\boldsymbol{f}$*. Suppose* $\boldsymbol{f}$ *is differentiable on such region and assume there is an* $M \geq 0$ *such that*

$$\left|\frac{\partial f_i}{\partial x_j}(\boldsymbol{x})\right| \leq M\,, \qquad \forall \boldsymbol{x} \in \mathcal{R},\ i = 1,\ldots,m\,,\ j = 1,\ldots,n\,.$$

Then $\boldsymbol{f}$ *is Lipschitz on* $\mathcal{R}$ *with* $L = \sqrt{nm}\,M$*.*

The proof is an easy consequence of Proposition 5.16 applied to each component of $\boldsymbol{f}$.

Vector-valued functions do not have an analogue of the Mean Value Theorem 5.12, since there might not be an $\overline{\boldsymbol{x}} \in S[\boldsymbol{a}, \boldsymbol{b}]$ such that $\boldsymbol{f}(\boldsymbol{b}) - \boldsymbol{f}(\boldsymbol{a}) = \boldsymbol{J}\boldsymbol{f}(\overline{\boldsymbol{x}})(\boldsymbol{b} - \boldsymbol{a})$ (whereas for each component f_i there clearly is a point $\overline{\boldsymbol{x}}_i \in S[\boldsymbol{a}, \boldsymbol{b}]$ satisfying (5.11)). A simple counterexample is the curve $\boldsymbol{f} : \mathbb{R} \to \mathbb{R}^2$, $\boldsymbol{f}(t) = (t^2, t^3)$ with $a = 0$, $b = 1$, for which $\boldsymbol{f}(1) - \boldsymbol{f}(0) = (1, 1)$ but $\boldsymbol{J}\boldsymbol{f}(t) = (2t, 3t^2)$. We cannot find any $\bar{t}$ such that $\boldsymbol{J}\boldsymbol{f}(\bar{t})(b - a) = (2\bar{t}, 3\bar{t}^2) = (1, 1)$.

Nevertheless, one could prove, under hypotheses similar to Lagrange's statement 5.12, that

$$\|\boldsymbol{f}(\boldsymbol{b}) - \boldsymbol{f}(\boldsymbol{a})\| \leq \sup_{\boldsymbol{x} \in S[\boldsymbol{a}, \boldsymbol{b}]} \|\boldsymbol{J}\boldsymbol{f}(\boldsymbol{x})\| \, \|\boldsymbol{b} - \boldsymbol{a}\| \,,$$

where the Jacobian's norm is defined as in Sect. 4.2.

At last, we extend the notion of functions of class $\mathcal{C}^k$, see Sect. 5.4. A map $\boldsymbol{f}$ is of class $\mathcal{C}^k$ ($0 \leq k \leq \infty$) on the open set $\Omega \subseteq \operatorname{dom} \boldsymbol{f}$ if all components are of class $\mathcal{C}^k$ on Ω; we shall write $\boldsymbol{f} \in \big(\mathcal{C}^k(\Omega)\big)^n$. A similar definition is valid if we take $\overline{\Omega}$ instead of Ω.

6.3 Basic differential operators

Given a real function φ, defined on an open set Ω in $\mathbb{R}^n$ and differentiable on Ω, we saw in Sect. 5.2 how to associate to such a scalar field on Ω the (first) partial derivatives $\dfrac{\partial \varphi}{\partial x_j}$ with respect to the coordinates x_j, $j = 1, \ldots, n$; these are still scalar fields on Ω. Each mapping $\varphi \mapsto \dfrac{\partial \varphi}{\partial x_j}$ is a linear operator, because

$$\frac{\partial}{\partial x_j}(\lambda \varphi + \mu \psi) = \lambda \frac{\partial \varphi}{\partial x_j} + \mu \frac{\partial \psi}{\partial x_j}$$

for any pair of functions φ, ψ differentiable on Ω and any pair of numbers λ, $\mu \in \mathbb{R}$. The operator maps $\mathcal{C}^1(\Omega)$ to $\mathcal{C}^0(\Omega)$: each partial derivative of a $\mathcal{C}^1$ function on Ω is of class $\mathcal{C}^0$ on Ω; in general, each operator $\dfrac{\partial}{\partial x_j}$ maps $\mathcal{C}^k(\Omega)$ to $\mathcal{C}^{k-1}(\Omega)$, for any $k \geq 1$.

6.3.1 First-order operators

Using operators involving (first) partial derivatives we can introduce a host of linear differential operators of order one that act on (scalar or vector) fields defined and differentiable on Ω, and return (scalar or vector) fields on Ω. The first we wish to describe is the **gradient** operator; as we know, it associates to a differentiable scalar field the vector field of first derivatives:

$$\mathbf{grad}\,\varphi = \nabla\varphi = \left(\frac{\partial\varphi}{\partial x_j}\right)_{1\le j\le n} = \frac{\partial\varphi}{\partial x_1}\,\boldsymbol{e}_1 + \cdots + \frac{\partial\varphi}{\partial x_n}\,\boldsymbol{e}_n\,.$$

Thus if $\varphi \in \mathcal{C}^1(\Omega)$, then $\mathbf{grad}\,\varphi \in \big(\mathcal{C}^0(\Omega)\big)^n$, meaning the gradient is a linear operator from $\mathcal{C}^1(\Omega)$ to $\big(\mathcal{C}^0(\Omega)\big)^n$.

Let us see two other fundamental linear differential operators.

Definition 6.5 *The* **divergence** *of a vector field* $\boldsymbol{f} = f_1\boldsymbol{e}_1 + \cdots + f_n\boldsymbol{e}_n$*, differentiable on* $\Omega \subseteq \mathbb{R}^n$*, is the scalar field*

$$\operatorname{div}\boldsymbol{f} = \frac{\partial f_1}{\partial x_1} + \cdots + \frac{\partial f_n}{\partial x_n} = \sum_{j=1}^{n}\frac{\partial f_j}{\partial x_j}\,. \tag{6.3}$$

The divergence operator maps $\big(\mathcal{C}^1(\Omega)\big)^n$ *to* $\mathcal{C}^0(\Omega)$*.*

Definition 6.6 *The* **curl** *of a vector field* $\boldsymbol{f} = f_1\boldsymbol{i} + f_2\boldsymbol{j} + f_3\boldsymbol{k}$*, differentiable on* $\Omega \subseteq \mathbb{R}^3$*, is the vector field*

$$\begin{aligned}\mathbf{curl}\,\boldsymbol{f} &= \left(\frac{\partial f_3}{\partial x_2} - \frac{\partial f_2}{\partial x_3}\right)\boldsymbol{i} + \left(\frac{\partial f_1}{\partial x_3} - \frac{\partial f_3}{\partial x_1}\right)\boldsymbol{j} + \left(\frac{\partial f_2}{\partial x_1} - \frac{\partial f_1}{\partial x_2}\right)\boldsymbol{k} \\ &= \det\begin{pmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ \dfrac{\partial}{\partial x_1} & \dfrac{\partial}{\partial x_2} & \dfrac{\partial}{\partial x_3} \\ f_1 & f_2 & f_3 \end{pmatrix}\end{aligned} \tag{6.4}$$

(*the determinant is computed along the first row*)*. The curl operator maps* $\big(\mathcal{C}^1(\Omega)\big)^3$ *to* $\big(\mathcal{C}^0(\Omega)\big)^3$*. Another symbol used in many European countries is* **rot***, standing for* **rotor***.*

We remark that the curl, as above defined, acts only on three-dimensional vector fields. In dimension 2, one defines the curl of a vector field $\boldsymbol{f}$, differentiable on an open set $\Omega \subseteq \mathbb{R}^2$, as the scalar field

$$\boxed{\operatorname{curl}\boldsymbol{f} = \frac{\partial f_2}{\partial x_1} - \frac{\partial f_1}{\partial x_2}\,.} \tag{6.5}$$

Note that this is the only non-zero component (the third one) of the curl of $\mathbf{\Phi}(x_1, x_2, x_3) = f_1(x_1, x_2)\boldsymbol{i} + f_2(x_1, x_2)\boldsymbol{j} + 0\boldsymbol{k}$, associated to $\boldsymbol{f}$; in other words

$$\mathbf{curl}\,\mathbf{\Phi} = 0\boldsymbol{i} + 0\boldsymbol{j} + (\operatorname{curl} \boldsymbol{f})\,\boldsymbol{k}\,. \tag{6.6}$$

Sometimes, in dimension 2, the curl of a differentiable scalar field φ on an open set $\Omega \subseteq \mathbb{R}^2$ is also defined as the (two-dimensional) vector field

$$\boxed{\mathbf{curl}\,\varphi = \frac{\partial \varphi}{\partial x_2}\boldsymbol{i} - \frac{\partial \varphi}{\partial x_1}\boldsymbol{j}\,.} \tag{6.7}$$

Here, too, the definition is suggested by a suitable three-dimensional curl: setting $\mathbf{\Phi}(x_1, x_2, x_3) = 0\boldsymbol{i} + 0\boldsymbol{j} + \varphi(x_1, x_2)\boldsymbol{k}$, we see immediately

$$\mathbf{curl}\,\mathbf{\Phi} = \mathbf{curl}\,\varphi + 0\boldsymbol{k}\,.$$

Higher-dimensional curl operators exist, but go beyond the purposes of this book.

Occasionally one finds useful to have a single formalism to represent the three operators gradient, divergence and curl. To this end, denote by ∇ the symbolic vector whose components are the partial differential operators $\dfrac{\partial}{\partial x_1}, \ldots, \dfrac{\partial}{\partial x_n}$:

$$\boxed{\nabla = \left(\frac{\partial}{\partial x_j}\right)_{1 \le j \le n} = \frac{\partial}{\partial x_1}\,\boldsymbol{e}_1 + \cdots + \frac{\partial}{\partial x_n}\,\boldsymbol{e}_n\,.}$$

In this way the gradient of a scalar field φ, denoted $\nabla\varphi$, may be thought of as obtained from the multiplication (on the right) of the vector ∇ by the scalar φ. Similarly, (6.3) shows the divergence of a vector field $\boldsymbol{f}$ is the dot product of the two vectors ∇ and $\boldsymbol{f}$, whence one writes

$$\boxed{\operatorname{div} \boldsymbol{f} = \nabla \cdot \boldsymbol{f}\,.}$$

In dimension 3 at last, the curl of a vector field $\boldsymbol{f}$ can be obtained, as (6.4) suggests, computing the cross product of the vectors ∇ and $\boldsymbol{f}$, allowing one to write

$$\boxed{\mathbf{curl}\,\boldsymbol{f} = \nabla \wedge \boldsymbol{f}\,.}$$

Let us illustrate the geometric meaning of the divergence and the curl of a three-dimensional vector field, and show that the former is related to the change in volume of a portion of mass moving under the effect of the vector field, while the latter has to do with the rotation of a solid around a point. So let $\boldsymbol{f} : \mathbb{R}^3 \to \mathbb{R}^3$ be a $\mathcal{C}^1$ vector field, which we shall assume to have bounded first derivatives on the whole $\mathbb{R}^3$. For any $\boldsymbol{x} \in \mathbb{R}^3$ let $\mathbf{\Phi}(t, \boldsymbol{x})$ be the trajectory of $\boldsymbol{f}$ passing through $\boldsymbol{x}$ at time $t = 0$ or, equivalently, the solution of the Cauchy problem for the autonomous differential system

$$\begin{cases} \boldsymbol{\Phi}' = \boldsymbol{f}(\boldsymbol{\Phi}), & t > 0, \\ \boldsymbol{\Phi}(0, \boldsymbol{x}) = \boldsymbol{x}. \end{cases}$$

We may also suppose that the solution exists at any time $t \geq 0$ and is differentiable with continuity with respect to both t and $\boldsymbol{x}$ (that this is true will be proved in Ch. 10). Fix a bounded open set Ω_0 and follow its evolution in time by looking at its images under $\boldsymbol{\Phi}$

$$\Omega_t = \boldsymbol{\Phi}(t,\Omega_0) = \{\boldsymbol{z} = \boldsymbol{\Phi}(t, \boldsymbol{x}) : \boldsymbol{x} \in \Omega_0\}.$$

We will introduce in Chapter 8 the triple integral of a map g defined on Ω_t,

$$\int_{\Omega_t} g(x, y, z)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z,$$

and show that, when g is the constant function 1, the integral represents the volume of the set Ω_t. Well, one can prove that

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\Omega_t} \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \int_{\Omega_t} \mathrm{div}\,\boldsymbol{f}\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z,$$

which shows it is precisely the divergence of $\boldsymbol{f}$ that governs the volume variation along the field's trajectories. In particular, if $\boldsymbol{f}$ is such that $\mathrm{div}\,\boldsymbol{f} = 0$ on $\mathbb{R}^3$, the volume of the image of any open set Ω_0 is constant with time (see Fig. 6.1 for a picture in dimension two).

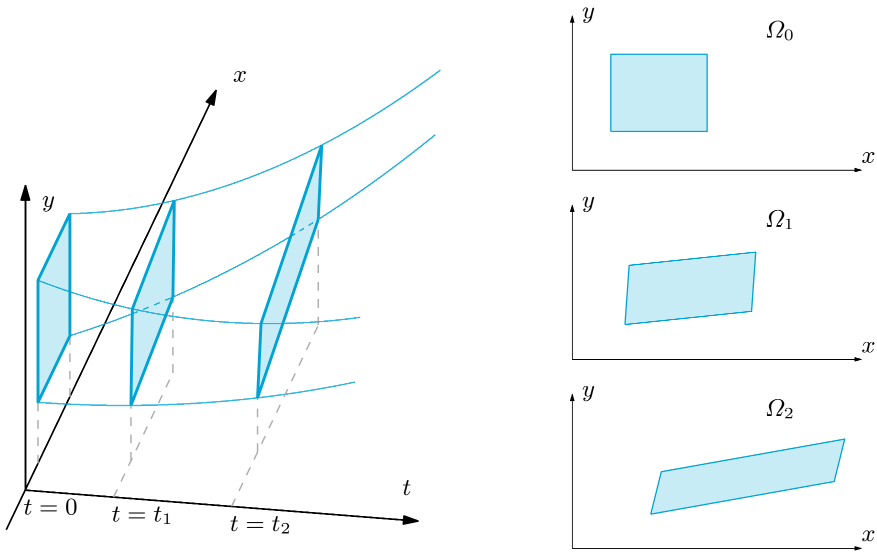


Figure 6.1. The area of a surface evolving in time under the effect of a two-dimensional field with zero divergence does not change

Let now $\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x}$ be a particular rigid motion of a three-dimensional solid S, namely a (clockwise) rotation around the z-axis by an angle θ. Then $\boldsymbol{A}$ is orthogonal (viz. it does not affect distances); precisely, it takes the form

$$\boldsymbol{A} = \begin{pmatrix} \cos\theta & \sin\theta & 0 \\ -\sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

It is easy to see that

$$\mathbf{curl}\, \boldsymbol{f} = 0\boldsymbol{i} + 0\boldsymbol{j} - 2\sin\theta \boldsymbol{k}\,;$$

so the curl of $\boldsymbol{f}$ has only one non-zero component, in the direction of rotation, that depends on the angle.

After this digression we return to the general properties of the operators gradient, divergence and curl, we observe that also the latter two are linear, like the gradient

$$\boxed{\begin{aligned} &\operatorname{div}(\lambda \boldsymbol{f} + \mu \boldsymbol{g}) = \lambda \operatorname{div} \boldsymbol{f} + \mu \operatorname{div} \boldsymbol{g}\,, \\ &\mathbf{curl}\,(\lambda \boldsymbol{f} + \mu \boldsymbol{g}) = \lambda\, \mathbf{curl}\, \boldsymbol{f} + \mu\, \mathbf{curl}\, \boldsymbol{g} \end{aligned}}$$

for any pair of fields $\boldsymbol{f}$, $\boldsymbol{g}$ and scalars λ, μ. Moreover, we have a list of properties expressing how the operators interact with various products between (scalar and vector) fields of class $\mathcal{C}^1$:

$$\boxed{\begin{aligned} &\mathbf{grad}\,(\varphi\psi) = \psi\, \mathbf{grad}\, \varphi + \varphi\, \mathbf{grad}\, \psi\,, \\ &\mathbf{grad}\,(\boldsymbol{f} \cdot \boldsymbol{g}) = \boldsymbol{g}\, \boldsymbol{J}\boldsymbol{f} + \boldsymbol{f}\, \boldsymbol{J}\boldsymbol{g}\,, \\ &\operatorname{div}(\varphi \boldsymbol{f}) = \mathbf{grad}\, \varphi \cdot \boldsymbol{f} + \varphi \operatorname{div} \boldsymbol{f}\,, \\ &\operatorname{div}(\boldsymbol{f} \wedge \boldsymbol{g}) = \boldsymbol{g} \cdot \mathbf{curl}\, \boldsymbol{f} - \boldsymbol{f} \cdot \mathbf{curl}\, \boldsymbol{g}\,, \\ &\mathbf{curl}\,(\varphi \boldsymbol{f}) = \mathbf{grad}\, \varphi \wedge \boldsymbol{f} + \varphi\, \mathbf{curl}\, \boldsymbol{f}\,, \\ &\mathbf{curl}\,(\boldsymbol{f} \wedge \boldsymbol{g}) = \boldsymbol{f} \operatorname{div} \boldsymbol{g} - \boldsymbol{g} \operatorname{div} \boldsymbol{f} + \boldsymbol{g}\, \boldsymbol{J}\boldsymbol{f} - \boldsymbol{f}\, \boldsymbol{J}\boldsymbol{g}\,. \end{aligned}}$$

Their proof is straightforward from the definitions and the rule for differentiating a product.

In two special cases, the successive action of two of $\mathbf{grad}$, div, $\mathbf{curl}$ on a sufficiently regular field gives the null vector field. The following results ensue from the definitions by using Schwarz's Theorem 5.17.

Proposition 6.7 *i) Let φ be a scalar field of class $\mathcal{C}^2$ on an open set Ω of $\mathbb{R}^3$. Then*

$$\mathbf{curl}\,\mathbf{grad}\,\varphi = \nabla \wedge (\nabla\varphi) = \mathbf{0} \qquad \text{on } \Omega\,.$$

ii) Let $\mathbf{\Phi}$ be a $\mathcal{C}^2$ vector field on an open set Ω of $\mathbb{R}^3$. Then

$$\operatorname{div}\mathbf{curl}\,\mathbf{\Phi} = \nabla \cdot (\nabla \wedge \mathbf{\Phi}) = 0 \qquad \text{on } \Omega\,.$$

The two-dimensional version of the above results reads

Proposition 6.8 *Let φ be a scalar field of class $\mathcal{C}^2$ on an open set Ω in $\mathbb{R}^2$. Then*

$$\operatorname{curl}\mathbf{grad}\,\varphi = 0 \qquad \text{and} \qquad \operatorname{div}\mathbf{curl}\,\varphi = 0 \qquad \text{on } \Omega\,.$$

These propositions steer us to the examination of fields with null gradient, null curl or zero divergence on a set Ω, a study with relevant applications. It is known (Proposition 5.13) that a scalar field φ has null gradient on Ω if and only if φ is constant on connected components of Ω. For the other two operators we preliminarily need some terminology.

Definition 6.9 *i) A vector field $\boldsymbol{f}$, differentiable on an open set Ω in $\mathbb{R}^3$ and such that* $\mathbf{curl}\,\boldsymbol{f} = \mathbf{0}$, *is said* **irrotational** (*or* **curl-free**) *on Ω.*
ii) A vector field $\boldsymbol{f}$, differentiable on an open set Ω of $\mathbb{R}^n$ and such that $\operatorname{div}\boldsymbol{f} = 0$ *is said* **divergence-free** *on Ω.*

Definition 6.10 *i) A vector field $\boldsymbol{f}$ on an open set Ω of $\mathbb{R}^n$ is* **conservative** *in Ω if there exists a scalar field φ such that* $\boldsymbol{f} = \mathbf{grad}\,\varphi$ *on Ω. The function φ is called a (scalar)* **potential** *of $\boldsymbol{f}$.*
ii) A vector field $\boldsymbol{f}$ on an open set Ω of $\mathbb{R}^3$ is **of curl type** *if there exists a vector field $\mathbf{\Phi}$ such that* $\boldsymbol{f} = \mathbf{curl}\,\mathbf{\Phi}$ *on Ω. The function $\mathbf{\Phi}$ is called a (vector)* **potential** *for $\boldsymbol{f}$.*

Taking these into consideration, Proposition 6.7 modifies as follows: i) if a $\mathcal{C}^1$ vector field $\boldsymbol{f}$ on an open set Ω of $\mathbb{R}^3$ is conservative (and so admits a scalar potential of class $\mathcal{C}^2$), then it is necessarily irrotational. ii) If a $\mathcal{C}^1$ vector field $\boldsymbol{f}$ on an open set Ω of $\mathbb{R}^3$ admits a $\mathcal{C}^2$ vector potential, it is divergence-free. Equivalently, we may concisely say that: i) $\boldsymbol{f}$ conservative implies $\boldsymbol{f}$ irrotational; ii) $\boldsymbol{f}$ of curl type implies $\boldsymbol{f}$ divergence-free.

The natural question is whether the above necessary conditions are also sufficient to guarantee the existence of a (scalar or vectorial) potential for $\boldsymbol{f}$. The answer will be given in Sect. 9.6, at least for fields with no curl. But we can say that in the absence of additional assumptions on the open set Ω the answer is negative. In fact, open subsets of $\mathbb{R}^3$ exist, on which there are irrotational vector fields of class $\mathcal{C}^1$ that are not conservative. Notwithstanding these counterexamples, we will provide conditions on Ω turning the necessary condition into an equivalence. In particular, each $\mathcal{C}^1$ vector field on an open, **convex** subset of $\mathbb{R}^3$ (like the interior of a cube, of a sphere, of an ellipsoid) is irrotational if and only if it is conservative.

Comparable results hold for divergence-free fields, in relationship to the existence of vector potentials.

Examples 6.11

i) Let $\boldsymbol{f}$ be the affine vector field

$$\boldsymbol{f} : \mathbb{R}^3 \to \mathbb{R}^3 , \qquad \boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x} + \boldsymbol{b}$$

defined by the 3×3 matrix $\boldsymbol{A}$ and the vector $\boldsymbol{b}$ of $\mathbb{R}^3$. Immediately we have

$$\operatorname{div} \boldsymbol{f} = a_{11} + a_{22} + a_{33} ,$$

$$\mathbf{curl}\, \boldsymbol{f} = (a_{32} - a_{23})\boldsymbol{i} + (a_{31} - a_{13})\boldsymbol{j} + (a_{21} - a_{12})\boldsymbol{k} .$$

Therefore $\boldsymbol{f}$ has no divergence on $\mathbb{R}^3$ if and only if the trace of $\boldsymbol{A}$, $\operatorname{tr} \boldsymbol{A} = a_{11} + a_{22} + a_{33}$, is zero. The field is, instead, irrotational on $\mathbb{R}^3$ if and only if $\boldsymbol{A}$ is symmetric.

Since $\mathbb{R}^3$ is clearly convex, to say $\boldsymbol{f}$ has no curl is the same as asserting $\boldsymbol{f}$ is conservative. In fact if $\boldsymbol{A}$ is symmetric, a (scalar) potential for $\boldsymbol{f}$ is

$$\varphi(\boldsymbol{x}) = \frac{1}{2}\boldsymbol{x} \cdot \boldsymbol{A}\boldsymbol{x} - \boldsymbol{b} \cdot \boldsymbol{x} .$$

Similarly, $\boldsymbol{f}$ is divergence-free if and only if it is of curl type, for if $\operatorname{tr} \boldsymbol{A}$ is zero, a vector potential for $\boldsymbol{f}$ is

$$\boldsymbol{\Phi}(\boldsymbol{x}) = \frac{1}{3}(\boldsymbol{A}\boldsymbol{x}) \wedge \boldsymbol{x} + \frac{1}{2}\boldsymbol{b} \wedge \boldsymbol{x} .$$

Note at last the field

$$\boldsymbol{f}(\boldsymbol{x}) = (y + z)\,\boldsymbol{i} + (x - z)\,\boldsymbol{j} + (x - y)\,\boldsymbol{k} ,$$

corresponding to

$$\boldsymbol{A} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & -1 \\ 1 & -1 & 0 \end{pmatrix} \quad \text{and} \quad \boldsymbol{b} = \boldsymbol{0} ,$$

is an example of a simultaneously irrotational and divergence-free field.

ii) Given $\boldsymbol{f} \in \big(\mathcal{C}^1(\Omega)\big)^3$, and $\boldsymbol{x}_0 \in \Omega$, consider the Jacobian $\boldsymbol{J}\boldsymbol{f}(\boldsymbol{x}_0)$ of $\boldsymbol{f}$ at $\boldsymbol{x}_0$. Then $(\operatorname{div} \boldsymbol{f})(\boldsymbol{x}_0) = 0$ if and only if $\operatorname{tr}\big(\boldsymbol{J}\boldsymbol{f}(\boldsymbol{x}_0)\big) = 0$, while $\big(\mathbf{curl}\, \boldsymbol{f}\big)(\boldsymbol{x}_0) = \boldsymbol{0}$ if and only if $\boldsymbol{J}\boldsymbol{f}(\boldsymbol{x}_0)$ is symmetric. In particular, the curl of $\boldsymbol{f}$ measures the failure of the Jacobian to be symmetric. □

Finally, we state without proof a theorem that casts some light on Definitions 6.5 and 6.6.

Theorem 6.12 *Let Ω be an open convex subset of $\mathbb{R}^3$. Every $\mathcal{C}^1$ vector field on Ω decomposes (not uniquely) into the sum of an irrotational field and a field with no divergence. In other terms, for any $\boldsymbol{f} \in \big(\mathcal{C}^1(\Omega)\big)^3$ there exist $\boldsymbol{f}^{(irr)} \in \big(\mathcal{C}^1(\Omega)\big)^3$ with $\mathbf{curl}\,\boldsymbol{f}^{(irr)} = \mathbf{0}$ on Ω, and $\boldsymbol{f}^{(divfree)} \in \big(\mathcal{C}^1(\Omega)\big)^3$ with $\operatorname{div} \boldsymbol{f}^{(divfree)} = 0$ on Ω, such that*

$$\boldsymbol{f} = \boldsymbol{f}^{(irr)} + \boldsymbol{f}^{(divfree)} .$$

Such a representation is known as **Helmholtz decomposition** *of $\boldsymbol{f}$.*

Example 6.13

We return to the affine vector field on $\mathbb{R}^3$ (Example 6.11 i)). Let us decompose the matrix $\boldsymbol{A}$ in the sum of its symmetric part $\boldsymbol{A}^{(sym)} = \frac{1}{2}(\boldsymbol{A} + \boldsymbol{A}^T)$ and skew-symmetric part $\boldsymbol{A}^{(skew)} = \frac{1}{2}(\boldsymbol{A} - \boldsymbol{A}^T)$,

$$\boldsymbol{A} = \boldsymbol{A}^{(sym)} + \boldsymbol{A}^{(skew)} .$$

Setting $\boldsymbol{f}^{(irr)}(\boldsymbol{x}) = \boldsymbol{A}^{(sym)}\boldsymbol{x} + \boldsymbol{b}$ and $\boldsymbol{f}^{(divfree)}(\boldsymbol{x}) = \boldsymbol{A}^{(skew)}\boldsymbol{x}$ realises the Helmholtz decomposition of $\boldsymbol{f}$: $\boldsymbol{f}^{(irr)}$ is irrotational as $\boldsymbol{A}^{(sym)}$ is symmetric, $\boldsymbol{f}^{(divfree)}$ is divergence-free as the diagonal of a skew-symmetric matrix is zero. Adding to $\boldsymbol{A}^{(sym)}$ an arbitrary traceless diagonal matrix $\boldsymbol{D}$ and subtracting the same from $\boldsymbol{A}^{(skew)}$ gives new fields $\boldsymbol{f}^{(irr)}$ and $\boldsymbol{f}^{(divfree)}$ for a different Helmholtz decomposition of $\boldsymbol{f}$. □

6.3.2 Second-order operators

The consecutive action of two linear, first-order differential operators typically produces a linear differential operator of order two, obviously defined on a sufficiently-regular (scalar or vector) field. We have already remarked (Proposition 6.7) how letting the curl act on the gradient, or computing the divergence of the curl, of a $\mathcal{C}^2$ field produces the null operator. We list a few second-order operators with crucial applications.

i) The operator $\operatorname{div} \mathbf{grad}$ maps a scalar field $\varphi \in \mathcal{C}^2(\Omega)$ to the $\mathcal{C}^0(\Omega)$ scalar field

$$\boxed{\operatorname{div} \mathbf{grad}\, \varphi = \nabla \cdot \nabla \varphi = \sum_{j=1}^{n} \frac{\partial^2 \varphi}{\partial x_j^2}\,,} \tag{6.8}$$

sum of the second partial derivatives of φ of pure type. The operator $\operatorname{div} \mathbf{grad}$ is known as **Laplace operator**, or simply **Laplacian**, and often denoted by Δ.

Therefore

$$\boxed{\Delta = \frac{\partial^2}{\partial x_1^2} + \cdots + \frac{\partial^2}{\partial x_n^2}\,,}$$

and (6.8) reads $\Delta\varphi$. Another possibility to denote Laplace's operator is ∇^2, which intuitively reminds the second equation of (6.8). Note $\nabla^2\varphi$ cannot be confused with $\nabla(\nabla\varphi)$, a meaningless expression as ∇ acts on scalars, not on vectors; $\nabla^2\varphi$ is purely meant as a shorthand symbol for $\nabla\cdot(\nabla\varphi)$.
A map φ such that $\Delta\varphi = 0$ on an open set Ω is called **harmonic** on Ω. Harmonic maps enjoy key mathematical properties and intervene in the description of several physical phenomena. For example, the electrostatic potential generated in vacuum by an electric charge at the point $\boldsymbol{x}_0 \in \mathbb{R}^3$ is a harmonic function defined on the open set $\mathbb{R}^3 \setminus \{\boldsymbol{x}_0\}$.

ii) The Laplace operator acts (component-wise) on vector fields as well, and one sets

$$\boxed{\Delta\boldsymbol{f} = \Delta f_1\boldsymbol{e}_1 + \cdots + \Delta f_n\boldsymbol{e}_n}$$

where $\boldsymbol{f} = f_1\boldsymbol{e}_1 + \cdots f_n\boldsymbol{e}_n$. Thus the vector Laplacian maps $\big(\mathcal{C}^2(\Omega)\big)^n$ to $\big(\mathcal{C}^0(\Omega)\big)^n$.

iii) The operator $\mathbf{grad}\,\mathrm{div}$ transforms vector fields into vector fields, and precisely it maps $\big(\mathcal{C}^2(\Omega)\big)^n$ to $\big(\mathcal{C}^0(\Omega)\big)^n$.

iv) Similarly the operator $\mathbf{curl}\,\mathbf{curl}$ goes from $\big(\mathcal{C}^2(\Omega)\big)^3$ to $\big(\mathcal{C}^0(\Omega)\big)^3$.
The latter three operators are related by the formula

$$\boxed{\Delta\boldsymbol{f} - \mathbf{grad}\,\mathrm{div}\,\boldsymbol{f} + \mathbf{curl}\,\mathbf{curl}\,\boldsymbol{f} = \mathbf{0}\,.}$$

6.4 Differentiating composite functions

Let

$$\boldsymbol{f} : \mathrm{dom}\,\boldsymbol{f} \subseteq \mathbb{R}^n \to \mathbb{R}^m \qquad \text{and} \qquad \boldsymbol{g} : \mathrm{dom}\,\boldsymbol{g} \subseteq \mathbb{R}^m \to \mathbb{R}^p$$

be two maps and $\boldsymbol{x}_0 \in \mathrm{dom}\,\boldsymbol{f}$ a point such that $\boldsymbol{y}_0 = \boldsymbol{f}(\boldsymbol{x}_0) \in \mathrm{dom}\,\boldsymbol{g}$, so that the composite

$$\boldsymbol{h} = \boldsymbol{g} \circ \boldsymbol{f} : \mathrm{dom}\,\boldsymbol{h} \subseteq \mathbb{R}^n \to \mathbb{R}^p\,,$$

for $\boldsymbol{x}_0 \in \mathrm{dom}\,\boldsymbol{h}$, is well defined. We know the composition of continuous functions is continuous (Proposition 4.23).

As far as differentiability is concerned, we have a result whose proof is similar to the case $n = m = p = 1$.

Theorem 6.14 *Let $\boldsymbol{f}$ be differentiable at $\boldsymbol{x}_0 \in \mathrm{dom}\,\boldsymbol{h}$ and $\boldsymbol{g}$ differentiable at $\boldsymbol{y}_0 = \boldsymbol{f}(\boldsymbol{x}_0)$. Then $\boldsymbol{h} = \boldsymbol{g} \circ \boldsymbol{f}$ is differentiable at $\boldsymbol{x}_0$ and its Jacobian matrix is*

$$\boldsymbol{J}(\boldsymbol{g} \circ \boldsymbol{f})(\boldsymbol{x}_0) = \boldsymbol{J}\boldsymbol{g}(\boldsymbol{y}_0)\,\boldsymbol{J}\boldsymbol{f}(\boldsymbol{x}_0) \tag{6.9}$$

Note $\boldsymbol{Jf}(\boldsymbol{x}_0)$ is $m \times n$, $\boldsymbol{Jg}(\boldsymbol{y}_0)$ is $p \times m$, hence the product of matrices on the right-hand side is well defined and produces a $p \times n$ matrix.

We shall make (6.9) explicit by writing the derivatives of the components of $\boldsymbol{h}$ in terms of the derivatives of $\boldsymbol{f}$ and $\boldsymbol{g}$. For this, set $\boldsymbol{x} = (x_j)_{1 \le j \le n}$ and

$$\boldsymbol{y} = \boldsymbol{f}(\boldsymbol{x}) = \big(f_k(\boldsymbol{x})\big)_{1 \le k \le m}, \qquad \boldsymbol{z} = \boldsymbol{g}(\boldsymbol{y}) = \big(g_i(\boldsymbol{y})\big)_{1 \le i \le p},$$

$$\boldsymbol{z} = \boldsymbol{h}(\boldsymbol{x}) = \big(h_i(\boldsymbol{x})\big)_{1 \le i \le p}.$$

The entry on row i and column j of $\boldsymbol{Jh}(\boldsymbol{x}_0)$ is

$$\frac{\partial h_i}{\partial x_j}(\boldsymbol{x}_0) = \sum_{k=1}^{m} \frac{\partial g_i}{\partial y_k}(\boldsymbol{y}_0)\, \frac{\partial f_k}{\partial x_j}(\boldsymbol{x}_0). \tag{6.10}$$

To remember these equations, we can write them as

$$\boxed{\frac{\partial z_i}{\partial x_j}(\boldsymbol{x}_0) = \sum_{k=1}^{m} \frac{\partial z_i}{\partial y_k}(\boldsymbol{y}_0)\, \frac{\partial y_k}{\partial x_j}(\boldsymbol{x}_0), \qquad 1 \le i \le p,\ 1 \le j \le n,} \tag{6.11}$$

also known as the **chain rule** for differentiating composite maps.

Examples 6.15

i) Let $\boldsymbol{f} = (f_1, f_2) : \mathbb{R}^2 \to \mathbb{R}^2$ and $g : \mathbb{R}^2 \to \mathbb{R}$ be differentiable. Call $h = g \circ \boldsymbol{f} : \mathbb{R}^2 \to \mathbb{R}$ the composition, $h(x, y) = g\big(f_1(x, y), f_2(x, y)\big)$. Then

$$\boxed{\nabla h(\boldsymbol{x}) = \nabla g\big(\boldsymbol{f}(\boldsymbol{x})\big)\, \boldsymbol{Jf}(\boldsymbol{x});} \tag{6.12}$$

putting $u = f_1(x, y)$, $v = f_2(x, y)$, this becomes

$$\frac{\partial h}{\partial x}(x, y) = \frac{\partial g}{\partial u}(u, v)\, \frac{\partial f_1}{\partial x}(x, y) + \frac{\partial g}{\partial v}(u, v)\, \frac{\partial f_2}{\partial x}(x, y)$$

$$\frac{\partial h}{\partial y}(x, y) = \frac{\partial g}{\partial u}(u, v)\, \frac{\partial f_1}{\partial y}(x, y) + \frac{\partial g}{\partial v}(u, v)\, \frac{\partial f_2}{\partial y}(x, y).$$

ii) Let $\varphi : I \subseteq \mathbb{R} \to \mathbb{R}$ be a differentiable map and $f : \mathbb{R}^2 \to \mathbb{R}$ a scalar differentiable function. The composite $h(x) = f\big(x, \varphi(x)\big)$ is differentiable on I, and

$$\boxed{h'(x) = \frac{\mathrm{d}h}{\mathrm{d}x}(x) = \frac{\partial f}{\partial x}\big(x, \varphi(x)\big) + \frac{\partial f}{\partial y}\big(x, \varphi(x)\big)\varphi'(x),}$$

as follows from Theorem 6.14, since $h = f \circ \boldsymbol{\Phi}$ with $\boldsymbol{\Phi} : I \to \mathbb{R}^2$, $\boldsymbol{\Phi}(x) = \big(x, \varphi(x)\big)$.

When a function depends on one variable only, the partial derivative symbol in (6.9) should be replaced by the more precise ordinary derivative.

iii) Consider a curve $\boldsymbol{\gamma} : I \subseteq \mathbb{R} \to \mathbb{R}^3$ of differentiable components γ_i, together with a differentiable map $f : \mathbb{R}^3 \to \mathbb{R}$. Let $h = f \circ \boldsymbol{\gamma} : I \to \mathbb{R}$ be the composition. Then

$$h'(t) = \nabla f\big(\boldsymbol{\gamma}(t)\big) \cdot \boldsymbol{\gamma}'(t)\,, \tag{6.13}$$

or, putting $(x, y, z) = \boldsymbol{\gamma}(t)$,

$$\boxed{\frac{\mathrm{d}h}{\mathrm{d}t}(t) = \frac{\partial f}{\partial x}(x,y,z)\,\frac{\mathrm{d}\gamma_1}{\mathrm{d}t}(t) + \frac{\partial f}{\partial y}(x,y,z)\,\frac{\mathrm{d}\gamma_2}{\mathrm{d}t}(t) + \frac{\partial f}{\partial z}(x,y,z)\,\frac{\mathrm{d}\gamma_3}{\mathrm{d}t}(t)\,.}$$

□

Theorem 6.14 can be successfully applied to extend the one-variable result, known to the student, about the derivative of the inverse function (Vol. I, Thm. 6.9). We show that under suitable hypotheses the Jacobian of the inverse function is roughly speaking the inverse Jacobian matrix. Sect. 7.1.1 will give us a sufficient condition for the following corollary to hold.

Corollary 6.16 *Let $\boldsymbol{f} : \mathrm{dom}\,\boldsymbol{f} \subseteq \mathbb{R}^n \to \mathbb{R}^n$ be differentiable at $\boldsymbol{x}_0$ with non-singular $\boldsymbol{J}\boldsymbol{f}(\boldsymbol{x}_0)$. Assume further that $\boldsymbol{f}$ is invertible on a neighbourhood of $\boldsymbol{x}_0$, and that the inverse map $\boldsymbol{f}^{-1}$ is differentiable at $\boldsymbol{y}_0 = \boldsymbol{f}(\boldsymbol{x}_0)$. Then*

$$\boldsymbol{J}(\boldsymbol{f}^{-1})(\boldsymbol{y}_0) = \big(\boldsymbol{J}\boldsymbol{f}(\boldsymbol{x}_0)\big)^{-1}\,.$$

Proof. Applying the theorem with $\boldsymbol{g} = \boldsymbol{f}^{-1}$ will meet our needs, because $\boldsymbol{h} = \boldsymbol{f}^{-1} \circ \boldsymbol{f}$ is the identity map ($\boldsymbol{h}(\boldsymbol{x}) = \boldsymbol{x}$ for any $\boldsymbol{x}$ around $\boldsymbol{x}_0$), hence $\boldsymbol{J}\boldsymbol{h}(\boldsymbol{x}_0) = \boldsymbol{I}$. Therefore

$$\boldsymbol{J}(\boldsymbol{f}^{-1})(\boldsymbol{y}_0)\,\boldsymbol{J}\boldsymbol{f}(\boldsymbol{x}_0) = \boldsymbol{I}\,,$$

whence the claim, as $\boldsymbol{J}\boldsymbol{f}(\boldsymbol{x}_0)$ is invertible. □

6.4.1 Functions defined by integrals

We encounter a novel way to define a function, a way that takes a given map of two (scalar or vectorial) variables and integrates it with respect to one of them. This kind of map has many manifestations, for instance in the description of electromagnetic fields. In the sequel we will restrict to the case of two scalar variables, although the more general treatise is not, conceptually, that more difficult.

Let then g be a real map defined on the set $\mathcal{R} = I \times J \subset \mathbb{R}^2$, where I is an arbitrary real interval and $J = [a, b]$ is closed and bounded. Suppose g is continuous on $\mathcal{R}$ and define

$$\boxed{f(x) = \int_a^b g(x,y)\,\mathrm{d}y\,;} \tag{6.14}$$

this is a one-variable map defined on I, because for any $x \in I$ the function $y \mapsto g(x,y)$ is continuous hence integrable on J.

Our function f satisfies the following properties, whose proof can be found in Appendix A.1.3, p. 516.

Proposition 6.17 *The function f of* (6.14) *is continuous on I. Moreover, if g admits continuous partial derivative $\dfrac{\partial g}{\partial x}$ on $\mathcal{R}$, then f is of class $\mathcal{C}^1$ on I and*

$$f'(x) = \int_a^b \frac{\partial g}{\partial x}(x,y)\,\mathrm{d}y\,.$$

This proposition spells out a *rule for differentiating integrals*: differentiating in one variable and integrating in the other are commuting operations, namely

$$\frac{\mathrm{d}}{\mathrm{d}x}\int_a^b g(x,y)\,\mathrm{d}y = \int_a^b \frac{\partial g}{\partial x}(x,y)\,\mathrm{d}y\,.$$

The above formula extends to higher derivatives:

$$f^{(k)}(x) = \int_a^b \frac{\partial^k g}{\partial x^k}(x,y)\,\mathrm{d}y\,, \qquad k \geq 1\,,$$

provided the integrand exists and is continuous on $\mathcal{R}$.

A more general form of (6.14) is

$$\boxed{f(x) = \int_{\alpha(x)}^{\beta(x)} g(x,y)\,\mathrm{d}y\,,} \tag{6.15}$$

where α and β are defined on I with values in $[a,b]$.

Notice that the integral function

$$f(x) = \int_a^x g(y)\,\mathrm{d}y\,,$$

considered in Vol. I, Sect. 9.8, is a special case.

Proposition 6.17 generalises in the following manner.

Proposition 6.18 *If α and β are continuous on I, the map f defined by* (6.15) *is continuous on I. If moreover g admits continuous partial derivative $\dfrac{\partial g}{\partial x}$ on $\mathcal{R}$ and α, β are $\mathcal{C}^1$ on I, then f is $\mathcal{C}^1$ on I, and*

$$f'(x) = \int_{\alpha(x)}^{\beta(x)} \frac{\partial g}{\partial x}(x,y)\,\mathrm{d}y + \beta'(x) g\big(x,\beta(x)\big) - \alpha'(x) g\big(x,\alpha(x)\big)\,. \tag{6.16}$$

Proof. We shall only prove the formula, referring to Appendix A.1.3, p. 517, for the rest of the proof.

Define on $I \times J^2$ the map

$$F(x,p,q) = \int_p^q g(x,y)\,\mathrm{d}y\,;$$

it admits continuous first derivatives everywhere on its domain

$$\frac{\partial F}{\partial x}(x,p,q) = \int_p^q \frac{\partial g}{\partial x}(x,y)\,\mathrm{d}y\,,$$

$$\frac{\partial F}{\partial p}(x,p,q) = -g(x,p)\,, \qquad \frac{\partial F}{\partial q}(x,p,q) = g(x,q)\,.$$

The last two descend from the Fundamental Theorem of Integral Calculus. The assertion now follows from the fact that $f(x) = F\big(x,\alpha(x),\beta(x)\big)$ by applying the chain rule

$$\frac{\mathrm{d}f}{\mathrm{d}x}(x) = \frac{\partial F}{\partial x}\big(x,\alpha(x),\beta(x)\big) + \\ + \frac{\partial F}{\partial p}\big(x,\alpha(x),\beta(x)\big)\alpha'(x) + \frac{\partial F}{\partial q}\big(x,\alpha(x),\beta(x)\big)\beta'(x)\,.$$

□

Example 6.19

The map

$$f(x) = \int_x^{x^2} \frac{\mathrm{e}^{-xy^2}}{y}\,\mathrm{d}y$$

is of the form (6.15) if we set $g(x,y) = \frac{\mathrm{e}^{-xy^2}}{y}$, $\alpha(x) = x$, $\beta(x) = x^2$. As g is not integrable in elementary functions with respect to y, we cannot compute $f(x)$ explicitly. But g is $\mathcal{C}^1$ on any closed region contained in the first quadrant $x > 0$, $y > 0$, while α and β are regular everywhere. Invoking Proposition 6.18 we deduce f is $\mathcal{C}^1$ on $(0,+\infty)$, with derivative

$$f'(x) = -\int_x^{x^2} ye^{-xy^2}\,\mathrm{d}y + g(x,x^2)\,2x - g(x,x)$$

$$= \frac{\mathrm{e}^{-xy^2}}{y}\Big|_{y=x}^{y=x^2} + \frac{2}{x}\,\mathrm{e}^{-x^5} - \frac{1}{x}\,\mathrm{e}^{-x^3} = \frac{5}{2x}\mathrm{e}^{-x^5} - \frac{3}{2x}\mathrm{e}^{-x^3}\,.$$

Using this we can deduce the behaviour of f around the point $x_0 = 1$. Firstly, $f(1) = 0$ and $f'(1) = \mathrm{e}^{-1}$; secondly, differentiating $f'(x)$ once more gives $f''(1) = -7\mathrm{e}^{-1}$. Therefore

$$f(x) = \frac{1}{\mathrm{e}}(x-1) - \frac{7}{2\mathrm{e}}(x-1)^2 + o\big((x-1)^2\big)\,, \qquad x \to 1\,,$$

implying f is positive, increasing and concave around x_0. □

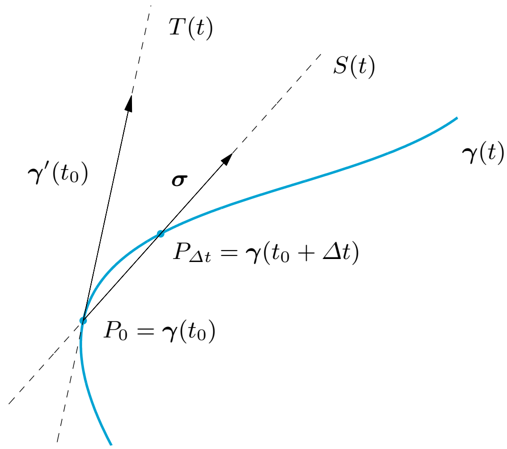


Figure 6.2. Tangent and secant vectors to a curve at P_0

6.5 Regular curves

Curves were introduced in Sect. 4.6. A **curve** $\gamma : I \to \mathbb{R}^m$ is said **differentiable** if its components $x_i : I \to \mathbb{R}$, $i \leq i \leq m$, are differentiable on I (recall a map is differentiable on an interval I if differentiable at all interior points of I, and at the end-points of I where present). We denote by $\gamma' : I \to \mathbb{R}^m$ the derivative $\gamma'(t) = \big(x_i'(t)\big)_{1\leq i\leq m} = \sum_{i=1}^{m} x_i'(t)\boldsymbol{e}_i$.

Definition 6.20 *A* **curve** $\gamma : I \to \mathbb{R}^m$ *is* **regular** *if it is differentiable on I with continuous derivative (the components are $\mathcal{C}^1$ on I) and if $\gamma'(t) \neq \mathbf{0}$, for any $t \in I$.*

A **curve** $\gamma : I \to \mathbb{R}^m$ *is* **piecewise regular** *if I is the finite union of intervals on which γ is regular.*

If γ is a regular curve and $t_0 \in I$, we can interpret the vector $\gamma'(t_0)$ geometrically (Fig. 6.2). Calling $P_0 = \gamma(t_0)$ and taking $t_0 + \Delta t \in I$ such that the point $P_{\Delta t} = \gamma(t_0 + \Delta t)$ is distinct from P_0, we consider the lines through P_0 and $P_{\Delta t}$; by (4.26), a secant line can be parametrised as

$$S(t) = P_0 + \big(P_{\Delta t} - P_0\big)\frac{t-t_0}{\Delta t} = \gamma(t_0) + \frac{\gamma(t_0+\Delta t) - \gamma(t_0)}{\Delta t}(t - t_0)\,. \qquad (6.17)$$

Letting now Δt tend to 0, the point $P_{\Delta t}$ moves towards P_0 (in the sense that each component of $P_{\Delta t}$ tends to the corresponding component of P_0). Meanwhile, the vector $\boldsymbol{\sigma} = \boldsymbol{\sigma}(t_0, \Delta t) = \dfrac{\gamma(t_0+\Delta t) - \gamma(t_0)}{\Delta t}$ tends to $\gamma'(t_0)$ by regularity.

The limit position of (6.17) is thus

$$\boxed{T(t) = \boldsymbol{\gamma}(t_0) + \boldsymbol{\gamma}'(t_0)(t - t_0)\,, \qquad t \in \mathbb{R}\,,}$$

the tangent line to the trace of the curve at P_0. For this reason we introduce

Definition 6.21 *Let $\boldsymbol{\gamma} : I \to \mathbb{R}^m$ be a regular curve, and $t_0 \in I$. The vector $\boldsymbol{\gamma}'(t_0)$ is called* **tangent vector** *to the trace of the curve at $P_0 = \boldsymbol{\gamma}(t_0)$.*

To be truly rigorous, the tangent vector at P_0 is the position vector $\big(P_0, \boldsymbol{\gamma}'(t_0)\big)$, but it is common practice to denote it by $\boldsymbol{\gamma}'(t_0)$. Later on we shall see the tangent line at a point is an intrinsic object, and does not depend upon the chosen parametrisation; its length and orientation, instead, do depend on the parametrisation.

In kinematics, a curve in $\mathbb{R}^3$ models the trajectory of a point-particle occupying the position $\boldsymbol{\gamma}(t)$ at time t. When the curve is regular, $\boldsymbol{\gamma}'(t)$ is the particle's velocity at the instant t.

Examples 6.22

i) All curves considered in Examples 4.34 are regular.

ii) Let $\varphi : I \to \mathbb{R}$ denote a continuously-differentiable map on I; the curve

$$\boldsymbol{\gamma}(t) = \big(t, \varphi(t)\big)\,, \qquad t \in I\,,$$

is regular, and has the graph of φ as trace. In fact,

$$\boldsymbol{\gamma}'(t) = \big(1, \varphi'(t)\big) \neq (0,0)\,, \qquad \text{for any } t \in I\,.$$

iii) The arc $\boldsymbol{\gamma} : [0,2] \to \mathbb{R}^2$ defined by

$$\boldsymbol{\gamma}(t) = \begin{cases} (t,1)\,, & t \in [0,1)\,, \\ (t,t)\,, & t \in [1,2]\,, \end{cases}$$

parametrises the polygonal path ABC (see Fig. 6.3, left); the arc

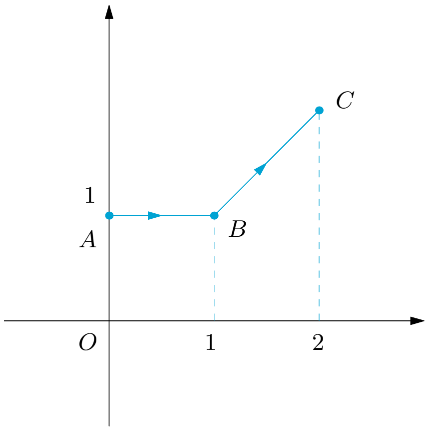
$$\boldsymbol{\gamma}(t) = \begin{cases} (t,1)\,, & t \in [0,1)\,, \\ (t,t)\,, & t \in [1,2)\,, \\ \big(4-t, 2-\frac{1}{2}(t-2)\big)\,, & t \in [2,4]\,, \end{cases}$$

parametrises the path $ABCA$ (Fig. 6.3, right). Both curves are piecewise regular.

iv) The curves

$$\boldsymbol{\gamma}(t) = \big(1 + \sqrt{2}\cos t, \sqrt{2}\sin t\big)\,, \qquad t \in [0, 2\pi]\,,$$
$$\widetilde{\boldsymbol{\gamma}}(t) = \big(1 + \sqrt{2}\cos 2t, -\sqrt{2}\sin 2t\big)\,, \qquad t \in [0, \pi]\,,$$

are different parametrisations (counter-clockwise and clockwise, respectively) of the same circle C, whose centre is $(1,0)$ and radius $\sqrt{2}$.



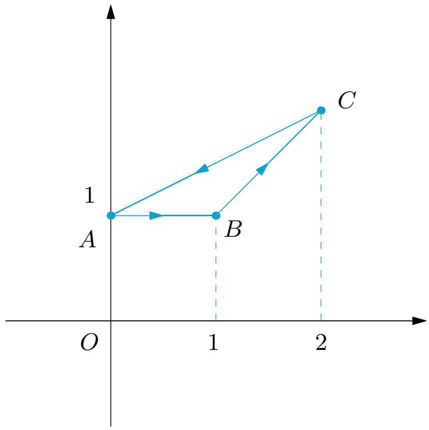


Figure 6.3. The polygonal paths ABC (left) and $ABCA$ (right) of Example 6.22 iii)

They are regular with derivatives

$$\boldsymbol{\gamma}'(t) = \sqrt{2}\big(-\sin t, \cos t\big)\,, \qquad \widetilde{\boldsymbol{\gamma}}'(t) = 2\sqrt{2}\big(-\sin 2t, -\cos 2t\big)\,.$$

The point $P_0 = (0,1) \in C$ is the image under $\boldsymbol{\gamma}$ of $t_0 = \frac{3}{4}\pi$, and of $\widetilde{t}_0 = \frac{5}{8}\pi$ under $\widetilde{\boldsymbol{\gamma}}$: $P_0 = \boldsymbol{\gamma}(t_0) = \widetilde{\boldsymbol{\gamma}}(\widetilde{t}_0)$. In the first case the tangent vector is $\boldsymbol{\gamma}'(t_0) = (-1,-1)$ and the tangent line at P_0

$$T(t) = (0,1) - (1,1)\big(t - \frac{3}{4}\pi\big) = \big(-t + \frac{3}{4}\pi, 1 - t + \frac{3}{4}\pi\big)\,, \qquad t \in \mathbb{R}\,;$$

in the second case $\widetilde{\boldsymbol{\gamma}}'(\widetilde{t}_0) = (2,2)$ and

$$\widetilde{T}(t) = (0,1) + (2,2)\big(t - \frac{5}{8}\pi\big) = \big(2(t - \frac{5}{8}\pi), 1 + 2(t - \frac{5}{8}\pi)\big)\,, \qquad t \in \mathbb{R}\,.$$

The tangent vectors at P_0 have different length and orientation, but the same direction. Recalling Example 4.34 i), in both cases $y = 1 + x$ is the tangent line.

v) The curve $\boldsymbol{\gamma}_s : \mathbb{R} \to [0,+\infty) \times \mathbb{R}^2$, defined in spherical coordinates by

$$\boldsymbol{\gamma}_s(t) = \big(r(t), \varphi(t), \theta(t)\big) = \big(1, \frac{\pi}{2}(1 + \frac{1}{2}\sin 8t), t\big)\,,$$

describes the periodic motion, on the unit sphere, of the point $P(t)$ that revolves around the z-axis and simultaneously oscillates between the parallels of colatitude $\varphi_{\min} = \frac{\pi}{4}$ and $\varphi_{\max} = \frac{3}{4}\pi$ (Fig. 6.4). The curve is regular, for

$$\boldsymbol{\gamma}_s'(t) = (0, 2\pi\cos 8t, 1)\,.$$

□

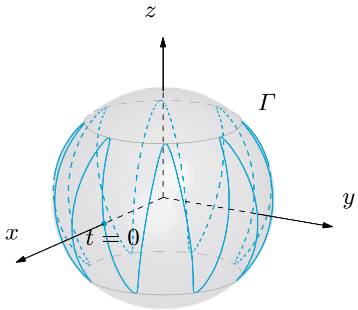


Figure 6.4. A point moving on a sphere (Example 6.22 v))

6.5.1 Congruence of curves; orientation

Now we discuss some useful relationships between curves parametrising the same trace.

Definition 6.23 *Let* $\boldsymbol{\gamma} : I \to \mathbb{R}^m$ *and* $\boldsymbol{\delta} : J \to \mathbb{R}^m$ *be regular curves. They are called* **congruent** *if there is a bijection* $\varphi : J \to I$*, differentiable with non-zero continuous derivative, such that*

$$\boldsymbol{\delta} = \boldsymbol{\gamma} \circ \varphi$$

(hence $\boldsymbol{\delta}(\tau) = \boldsymbol{\gamma}\big(\varphi(\tau)\big)$ *for any* $\tau \in J$*).*

For the sequel we remark that congruent curves have the same trace, for $\boldsymbol{x} = \boldsymbol{\delta}(\tau)$ if and only if $\boldsymbol{x} = \boldsymbol{\gamma}(t)$ with $t = \varphi(\tau)$. In addition, the tangent vectors at the point $P_0 = \boldsymbol{\gamma}(t_0) = \boldsymbol{\delta}(\tau_0)$ are collinear, because differentiating $\boldsymbol{\delta}(\tau) = \boldsymbol{\gamma}\big(\varphi(\tau)\big)$ at τ_0 gives

$$\boldsymbol{\delta}'(\tau_0) = \boldsymbol{\gamma}'\big(\varphi(\tau_0)\big)\varphi'(\tau_0) = \boldsymbol{\gamma}'(t_0)\varphi'(\tau_0)\,, \tag{6.18}$$

with $\varphi'(\tau_0) \neq 0$. Consequently, the tangent line at P_0 is the same for the two curves.

The map φ, relating the congruent curves $\boldsymbol{\gamma}$, $\boldsymbol{\delta}$ as in Definition 6.23, has φ' always > 0 or always < 0 on J; in fact, by assumption φ' is continuous and never zero on J, hence has constant sign by the Theorem of Existence of Zeroes. This entails we can divide congruent curves in two classes.

Definition 6.24 *The congruent curves* $\boldsymbol{\gamma} : I \to \mathbb{R}^m$, $\boldsymbol{\delta} : J \to \mathbb{R}^m$ *are* **equivalent** *if the bijection* $\varphi : J \to I$ *has strictly positive derivative, while they are* **anti-equivalent** *if* φ' *is strictly negative.*

Here is an example of two anti-equivalent curves.

Definition 6.25 *Let* $\boldsymbol{\gamma} : I \to \mathbb{R}^m$ *be a regular curve. Denoting by* $-I$ *the interval* $\{t \in \mathbb{R} : -t \in I\}$*, the curve* $-\boldsymbol{\gamma} : -I \to \mathbb{R}^m$*,* $(-\boldsymbol{\gamma})(t) = \boldsymbol{\gamma}(-t)$ *is said* **opposite** *to* $\boldsymbol{\gamma}$*, and is anti-equivalent to* $\boldsymbol{\gamma}$*.*

We may write $(-\boldsymbol{\gamma}) = \boldsymbol{\gamma} \circ \varphi$, where $\varphi : -I \to I$ is the bijection $\varphi(t) = -t$. If $\boldsymbol{\gamma} : [a, b] \to \mathbb{R}^m$ is a regular arc then $-\boldsymbol{\gamma}$ is still a regular arc defined on the interval $[-b, -a]$.

If we take anti-equivalent curves $\boldsymbol{\gamma}$ and $\boldsymbol{\delta}$ we may also write

$$\boldsymbol{\delta}(\tau) = \boldsymbol{\gamma}\big(\varphi(\tau)\big) = \boldsymbol{\gamma}\big(-(-\varphi(\tau))\big) = (-\boldsymbol{\gamma})\big(\psi(\tau)\big)$$

where $\psi : J \to -I$, $\psi(\tau) = -\varphi(\tau)$. Since $\psi'(\tau) = -\varphi'(\tau) > 0$, the curves $\boldsymbol{\delta}$ and $(-\boldsymbol{\gamma})$ will be equivalent. In conclusion,

Property 6.26 *Congruent curves are either equivalent or one is equivalent to the opposite of the other.*

Due to this observation we shall adopt the notation $\boldsymbol{\delta} \sim \boldsymbol{\gamma}$ to denote two equivalent curves, and $\boldsymbol{\delta} \sim -\boldsymbol{\gamma}$ for anti-equivalent ones.

By (6.18) now, equivalent curves have tangent vectors pointing in the same direction, whereas anti-equivalent curves have opposite tangent vectors.

Assume from now on that curves are simple. It is immediate to verify that all curves congruent to a simple curve are themselves simple. Moreover, one can prove the following property.

Proposition 6.27 *If* Γ *denotes the trace of a simple, regular curve* $\boldsymbol{\gamma}$*, any other simple regular curve* $\boldsymbol{\delta}$ *having* Γ *as trace is congruent to* $\boldsymbol{\gamma}$*.*

Thus all parametrisations of Γ by simple regular curves are grouped into two classes: two curves belong to the same class if they are equivalent, and live in different classes if anti-equivalent. To each class we associate an **orientation** of Γ. Given any parametrisation of Γ in fact, we say the point $P_2 = \boldsymbol{\gamma}(t_2)$ follows $P_1 = \boldsymbol{\gamma}(t_1)$ on $\boldsymbol{\gamma}$ if $t_2 > t_1$ (Fig. 6.5). Well, it is easy to see that if $\boldsymbol{\delta}$ is equivalent to $\boldsymbol{\gamma}$, P_2 follows P_1 also in this parametrisation, in other words $P_2 = \boldsymbol{\delta}(\tau_2)$ and $P_1 = \boldsymbol{\delta}(\tau_1)$ with $\tau_2 > \tau_1$. Conversely, if $\boldsymbol{\delta}$ is anti-equivalent to $\boldsymbol{\gamma}$, then $P_2 = \boldsymbol{\delta}(\tau_2)$ and $P_1 = \boldsymbol{\delta}(\tau_1)$ with $\tau_2 < \tau_1$, so P_1 will follow P_2.

As observed earlier, the orientation of Γ can be also determined from the orientation of its tangent vectors.

The above discussion explains why simple regular curves are commonly thought of as geometrical objects, rather than as parametrisations, and often the two notions are confused, on purpose. This motivates the following definition.

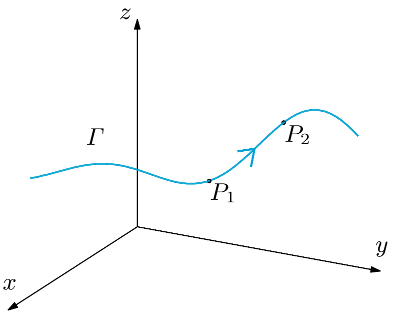


Figure 6.5. An arc Γ in $\mathbb{R}^3$ and an orientation on it

Definition 6.28 *A subset Γ of $\mathbb{R}^m$ is called a* **simple regular curve** *if it can be described as the trace of a curve with the same properties.*

If necessary one associates to Γ one of the two possible orientations. The choice of Γ and of an orientation on it determine a class of simple regular curves all equivalent to each other; within this class, one might select the most suitable parametrisation for the specific needs.

Every definition and property stated for regular curves adapts easily to piecewise-regular curves.

6.5.2 Length and arc length

We define the **length** of the regular arc $\boldsymbol{\gamma} : [a,b] \to \mathbb{R}^m$ as the number

$$\ell(\boldsymbol{\gamma}) = \int_a^b \|\boldsymbol{\gamma}'(t)\| \, \mathrm{d}t = \int_a^b \sqrt{\sum_{i=1}^m \big(x_i'(t)\big)^2} \, \mathrm{d}t \,. \tag{6.19}$$

The reason is geometrical (see Fig. 6.6). We subdivide $[a,b]$ using $a = t_0 < t_1 < \ldots, t_{K-1} < t_K = b$ and consider the points $P_k = \boldsymbol{\gamma}(t_k) \in \Gamma$, $k = 0, \ldots, K$. They determine a polygonal path in $\mathbb{R}^m$ (possibly degenerate) whose length is

$$\ell(t_0, t_1, \ldots, t_K) = \sum_{k=1}^K \operatorname{dist}(P_{k-1}, P_k) \,,$$

where $\operatorname{dist}(P_{k-1}, P_k) = \|P_k - P_{k-1}\|$ is the Euclidean distance of two points. Note

$$\|P_k - P_{k-1}\| = \sqrt{\sum_{i=1}^m \big(x_i(t_k) - x_i(t_{k-1})\big)^2} = \sqrt{\sum_{i=1}^m \left(\frac{\Delta x_i}{\Delta t}\right)_k^2} \, \Delta t_k$$

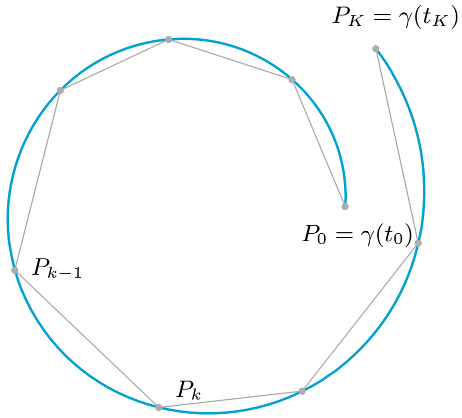


Figure 6.6. Approximation of the trace of an arc by a polygonal path

where $\Delta t_k = t_k - t_{k-1}$ and

$$\left(\frac{\Delta x_i}{\Delta t}\right)_k = \left(\frac{x_i(t_k) - x_i(t_{k-1})}{t_k - t_{k-1}}\right).$$

Therefore

$$\ell(t_0, t_1, \ldots, t_K) = \sum_{k=1}^{K} \sqrt{\sum_{i=1}^{m} \left(\frac{\Delta x_i}{\Delta t}\right)_k^2} \Delta t_k\,;$$

notice the analogy with the last integral in (6.19), of which the above is an approximation. One could prove that if the curve is piecewise regular, the least upper bound of $\ell(t_0, t_1, \ldots, t_K)$ over all possible partitions of $[a, b]$ is finite, and equals $\ell(\boldsymbol{\gamma})$.

The length (6.19) of an arc depends not only on the trace Γ, but also on the chosen parametrisation. For example, parametrising the circle $x^2 + y^2 = r^2$ by $\boldsymbol{\gamma}_1(t) = (r\cos t, r\sin t)$, $t \in [0, 2\pi]$, we have

$$\ell(\boldsymbol{\gamma}_1) = \int_0^{2\pi} r\,\mathrm{d}t = 2\pi r\,,$$

as is well known from elementary geometry. But if we take $\boldsymbol{\gamma}_2(t) = (r\cos 2t, r\sin 2t)$, $t \in [0, 2\pi]$, then

$$\ell(\boldsymbol{\gamma}_2) = \int_0^{2\pi} 2r\,\mathrm{d}t = 4\pi r\,.$$

In the latter case we went around the origin twice. Recalling (6.18), it is easy to see that two congruent arcs have the same length (see also Proposition 9.3 below, where $f = 1$). We shall prove in Sect. 9.1 that the length of a *simple* (or *Jordan*) arc depends only upon the trace Γ; it is called **length** of Γ, and denoted by $\ell(\Gamma)$.

In the previous example γ_1 is simple while γ_2 is not; the length $\ell(\Gamma)$ of the circle is $\ell(\gamma_1)$.

Let γ be a regular curve defined on the interval I, on which we fix an arbitrary point $t_0 \in I$, and introduce the function $s : I \to \mathbb{R}$

$$\boxed{s(t) = \int_{t_0}^{t} \|\gamma'(\tau)\| \,\mathrm{d}\tau\,.} \tag{6.20}$$

Recalling expression (6.19) for the length of an arc, we have

$$s(t) = \begin{cases} \ell(\gamma_{|[t_0,t]})\,, & t > t_0\,, \\ 0\,, & t = t_0\,, \\ -\ell(\gamma_{|[t,t_0]})\,, & t < t_0\,. \end{cases}$$

The function s allows to define an equivalent curve that gives a new parametrisation of the trace of γ. By the Fundamental Theorem of Integral Calculus, in fact,

$$s'(t) = \|\gamma'(t)\| > 0\,, \qquad \forall t \in I\,,$$

so s is a strictly increasing map, and hence invertible on I. Call $J = s(I)$ the image interval under s, and let $t : J \to I \subseteq \mathbb{R}$ be the inverse map of s. In other terms we write t as a function of another parameter s, $t = t(s)$. The curve $\widetilde{\gamma} : J \to \mathbb{R}^m$, $\widetilde{\gamma}(s) = \gamma(t(s))$ is equivalent to γ (in particular it has the same trace Γ). If $P_1 = \gamma(t_1)$ is an arbitrary point on Γ, then $P_1 = \widetilde{\gamma}(s_1)$ with t_1 and s_1 related by $t_1 = t(s_1)$. The number s_1 is the **arc length** of P_1.

Recalling the rule for differentiating an inverse function,

$$\widetilde{\gamma}'(s) = \frac{d\widetilde{\gamma}}{ds}(s) = \frac{d\gamma}{dt}\big(t(s)\big)\,\frac{dt}{ds}(s) = \frac{\gamma'(t)}{\|\gamma'(t)\|}\,,$$

whence

$$\|\widetilde{\gamma}'(s)\| = 1\,, \qquad \forall s \in J\,. \tag{6.21}$$

This means that the arc length parametrises the curve with constant "speed" 1.

The definitions of length of an arc and arc length extend to piecewise-regular curves.

Example 6.29

The curve $\gamma : \mathbb{R} \to \mathbb{R}^3$, $\gamma(t) = (\cos t, \sin t, t)$ has trace the circular helix (see Example 4.34 vi)). Then

$$\|\gamma'(t)\| = \|(-\sin t, \cos t, 1)\| = (\sin^2 t + \cos^2 t + 1)^{1/2} = \sqrt{2}\,.$$

Choosing $t_0 = 0$,

$$s(t) = \int_0^t \|\gamma'(\tau)\|\, d\tau = \sqrt{2}\int_0^t d\tau = \sqrt{2}t\,.$$

Therefore $t = t(s) = \frac{\sqrt{2}}{2}s$, $s \in \mathbb{R}$, and the helix can be parametrised anew by arc length

$$\widetilde{\gamma}(s) = \left(\cos \frac{\sqrt{2}}{2}s, \sin \frac{\sqrt{2}}{2}s, \frac{\sqrt{2}}{2}s\right) . \qquad \square$$

6.5.3 Elements of differential geometry for curves

This section is dedicated to the intrinsic geometry of curves in $\mathbb{R}^3$, and is not strictly essential for the sequel. As such it may be skipped at first reading.

We consider a regular, simple curve Γ in $\mathbb{R}^3$ parametrised by arc length s (defined from an origin point P^*). Call $\boldsymbol{\gamma} = \boldsymbol{\gamma}(s)$ such parametrisation, defined on an interval $J \subseteq \mathbb{R}$, and suppose $\boldsymbol{\gamma}$ is of class $\mathcal{C}^2$ on J.

If $\boldsymbol{t}(s) = \boldsymbol{\gamma}'(s)$ is the tangent vector to Γ at $P = \boldsymbol{\gamma}(s)$, by (6.21) we have

$$\boxed{\|\boldsymbol{t}(s)\| = 1 \, , \qquad \forall s \in J \, ,}$$

making $\boldsymbol{t}(s)$ a unit vector. Differentiating once more in s gives $\boldsymbol{t}'(s) = \boldsymbol{\gamma}''(s)$, a vector orthogonal to $\boldsymbol{t}(s)$; in fact differentiating

$$\|\boldsymbol{t}(s)\|^2 = \sum_{i=1}^{3} t_i^2(s) = 1 \, , \qquad \forall s \in J \, ,$$

we find

$$2\sum_{i=1}^{3} t_i(s)t_i'(s) = 0 \, ,$$

i.e., $\boldsymbol{t}(s) \cdot \boldsymbol{t}'(s) = 0$. Recall now that if $\boldsymbol{\gamma}(s)$ is the trajectory of a point-particle in time, its velocity $\boldsymbol{t}(s)$ has constant speed $= 1$. Therefore $\boldsymbol{t}'(s)$ represents the acceleration, and depends exclusively on the change of direction of the velocity vector. Thus the acceleration is perpendicular to the direction of motion.

If at $P_0 = \boldsymbol{\gamma}(s_0)$ the vector $\boldsymbol{t}'(s_0)$ is not zero, we may define

$$\boxed{\boldsymbol{n}(s_0) = \frac{\boldsymbol{t}'(s_0)}{\|\boldsymbol{t}'(s_0)\|} \, ,} \qquad (6.22)$$

called **principal normal** (vector) to Γ at P_0. The orthonormal vectors $\boldsymbol{t}(s_0)$ and $\boldsymbol{n}(s_0)$ lie on a plane passing through P_0, the **osculating plane** to the curve Γ at P_0. Among all planes passing through P_0, the osculating plane is the one that best adapts to the curve; to be precise, the distance of a point $\boldsymbol{\gamma}(s)$ on the curve from the osculating plane is infinitesimal of order bigger than $s - s_0$, as $s \to s_0$. The osculating plane of a plane curve is, at each point, the plane containing the curve.

The orientation of the principal normal has to do with the curve's convexity (Fig. 6.7). If the curve is plane, in the frame system of $\boldsymbol{t}$ and $\boldsymbol{n}$ the curve can be represented around P_0 as the graph of a convex function.

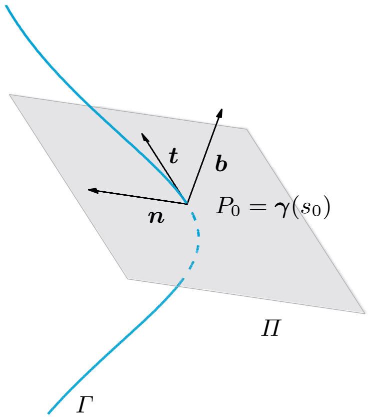


Figure 6.7. Osculating plane and tangent, normal, binormal vectors of Γ at P_0

The number $\mathcal{K}(s_0) = \|\boldsymbol{t}'(s_0)\|$ is the **curvature** of Γ at P_0, and its inverse $\mathcal{R}(s_0)$ is called **curvature radius**. These names arise from the following considerations. For simplicity suppose the curve is plane, and let us advance by $\mathcal{R}(s_0)$ in the direction of $\boldsymbol{n}$, starting from $P_0(s_0)$; the point $C(s_0)$ thus reached is the **centre of curvature** (Fig. 6.8). The circle with centre $C(s_0)$ and radius $\mathcal{R}(s_0)$ is tangent to Γ at P_0, and among all tangent circles we are considering the one that best approximates the curve around P_0 (osculating circle).

The orthogonal vector to the osculating plane,

$$\boxed{\boldsymbol{b}(s_0) = \boldsymbol{t}(s_0) \wedge \boldsymbol{n}(s_0)\,,} \tag{6.23}$$

is known as the **binormal** (vector) to Γ at P_0. This completes a positively-oriented triple $(\boldsymbol{t}, \boldsymbol{n}, \boldsymbol{b})$ of orthonormal vectors, that defines a moving frame along a curve.

The binormal unit vector, being orthogonal to the osculating plane, is constant along the curve if the latter is plane. Therefore, its variation measures how far the curve is from being plane. If the curve is $\mathcal{C}^3$ it makes sense to consider the vector $\boldsymbol{b}'(s_0)$, the **torsion vector** of the curve at P_0. Differentiating definition (6.23) gives

$$\boldsymbol{b}'(s_0) = \boldsymbol{t}'(s_0) \wedge \boldsymbol{n}(s_0) + \boldsymbol{t}(s_0) \wedge \boldsymbol{n}'(s_0) = \boldsymbol{t}(s_0) \wedge \boldsymbol{n}'(s_0)$$

as $\boldsymbol{n}(s_0)$ is parallel to $\boldsymbol{t}'(s_0)$. This explains why $\boldsymbol{n}'(s_0)$ is orthogonal to $\boldsymbol{t}(s_0)$. At the same time, differentiating $\|\boldsymbol{b}(s_0)\|^2 = 1$ gives $\boldsymbol{b}'(s_0) \cdot \boldsymbol{b}(s_0) = 0$, making $\boldsymbol{b}'(s_0)$ orthogonal to $\boldsymbol{b}(s_0)$ as well. Therefore $\boldsymbol{b}'(s_0)$ is parallel to $\boldsymbol{n}(s_0)$, and there must exist a scalar $\tau(s_0)$, called **torsion** of the curve, such that $\boldsymbol{b}'(s_0) = \tau(s_0)\boldsymbol{n}(s_0)$. It turns out that a curve is plane if and only if the torsion vanishes identically.

Differentiating the equation $\boldsymbol{n}(s) = \boldsymbol{b}(s) \wedge \boldsymbol{t}(s)$ gives

$$\begin{aligned}
\boldsymbol{n}'(s_0) &= \boldsymbol{b}(s_0) \wedge \boldsymbol{t}'(s_0) + \boldsymbol{b}'(s_0) \wedge \boldsymbol{t}(s_0) \\
&= \boldsymbol{b}(s_0) \wedge \mathcal{K}(s_0)\boldsymbol{n}(s_0) + \tau(s_0)\boldsymbol{n}(s_0) \wedge \boldsymbol{t}(s_0) \\
&= -\mathcal{K}(s_0)\boldsymbol{t}(s_0) - \tau(s_0)\boldsymbol{b}(s_0)\,.
\end{aligned}$$

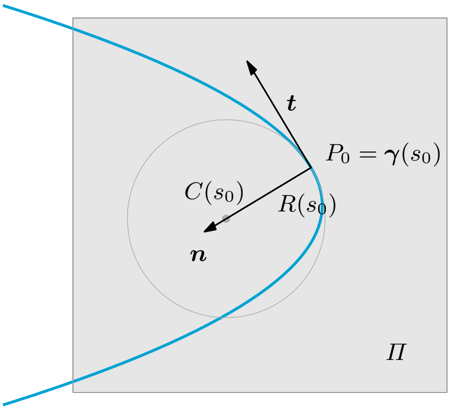


Figure 6.8. Osculating circle, centre and radius of curvature at P_0

To summarise, the unit vectors $\boldsymbol{t}, \boldsymbol{n}, \boldsymbol{b}$ of a $\mathcal{C}^3$ curve Γ satisfy the **Frenet formulas**

$$\boxed{\boldsymbol{t}' = \mathcal{K}\boldsymbol{n}\,, \qquad \boldsymbol{n}' = -\mathcal{K}\boldsymbol{t} - \tau\boldsymbol{b}\,, \qquad \boldsymbol{b}' = \tau\boldsymbol{n}\,.} \tag{6.24}$$

We conclude by pointing out that the vectors $\boldsymbol{t}, \boldsymbol{n}, \boldsymbol{b}$ admit a representation in terms of an arbitrary parametrisation of Γ.

6.6 Variable changes

Let $\mathcal{R}$ be a region of $\mathbb{R}^n$. The generic point $P \in \mathcal{R}$ is completely determined by its Cartesian coordinates $x_1, x_2, \ldots, x_n$ which, as components of a vector $\boldsymbol{x}$, allow to identify $P = \boldsymbol{x}$. For $i = 1, \ldots, n$ the line

$$R_i = \boldsymbol{x} + \mathbb{R}\boldsymbol{e}_i = \{\boldsymbol{x}_t = (x_1, \ldots, x_{i-1}, x_i + t, x_{i+1}, \ldots, x_n) : t \in \mathbb{R}\}$$

contains P and is parallel to the ith canonical unit vector $\boldsymbol{e}_i$. Thus the lines R_i, called *coordinate lines* through P, are mutually orthogonal.

Definition 6.30 *Let $\mathcal{R}'$ be another region of $\mathbb{R}^n$, with interior A'. A vector field $\boldsymbol{\Phi} : \mathcal{R}' \to \mathcal{R}$ defines a* **change of variables***, or* **change of coordinates***, on $\mathcal{R}$ if $\boldsymbol{\Phi}$ is:*

i) a bijective map between $\mathcal{R}'$ and $\mathcal{R}$;
ii) of class $\mathcal{C}^1$ on A';
iii) regular on A', or equivalently, the Jacobian $\boldsymbol{J\Phi}$ is non-singular everywhere on A'.

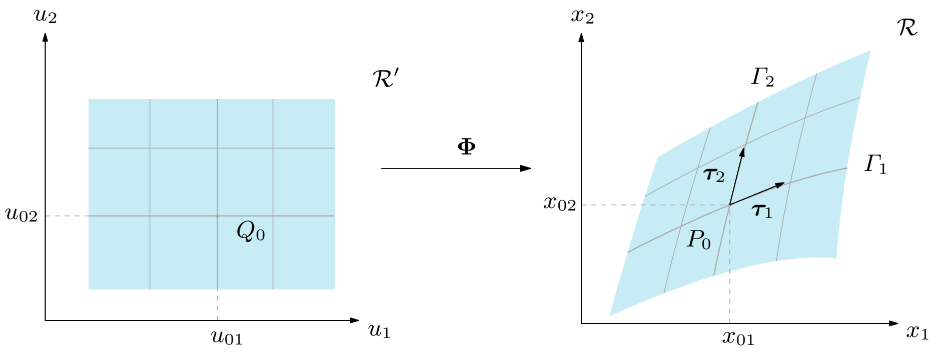


Figure 6.9. Change of variables in $\mathcal{R}$

Let A indicate the image $\boldsymbol{\Phi}(A')$: then it is possible to prove that A is open, and consequently $A \subseteq \overset{\circ}{\mathcal{R}}$; furthermore, $\boldsymbol{\Phi}(\partial\mathcal{R}') = \boldsymbol{\Phi}(\partial A')$ is contained in $\mathcal{R} \setminus A$.

Denote by $Q = \boldsymbol{u}$ the generic point of $\mathcal{R}'$, with Cartesian coordinates $u_1, \ldots, u_n$. Given $P_0 = \boldsymbol{x}_0 \in A$, part i) implies there exists a unique point $Q_0 = \boldsymbol{u}_0 \in A'$ such that $P_0 = \boldsymbol{\Phi}(Q_0)$, or $\boldsymbol{x}_0 = \boldsymbol{\Phi}(\boldsymbol{u}_0)$. Thus we may determine P_0, besides by its Cartesian coordinates $x_{01}, \ldots, x_{0n}$, also by the Cartesian coordinates $u_{01}, \ldots, u_{0n}$ of Q_0; the latter are called **curvilinear coordinates** of P_0 (relative to the transformation $\boldsymbol{\Phi}$). The coordinate lines through Q_0 produce curves in $\mathcal{R}$ passing through P_0. Precisely, the set

$$\Gamma_i = \{\boldsymbol{x} = \boldsymbol{\Phi}(\boldsymbol{u}_0 + t\boldsymbol{e}_i) : \boldsymbol{u}_0 + t\boldsymbol{e}_i \in \mathcal{R}' \text{ with } t \in I_i\},$$

where $i = 1, \ldots, n$ and I_i is an interval containing the origin, is the trace of the curve (simple, by i))

$$t \mapsto \boldsymbol{\gamma}_i(t) = \boldsymbol{\Phi}(\boldsymbol{u}_0 + t\boldsymbol{e}_i)$$

defined on I_i. These sets are the **coordinate lines** through the point P_0 (relative to $\boldsymbol{\Phi}$); see Fig. 6.9.

If $P_0 \in A$ these are regular curves (by iii)), so tangent vectors at P_0 exist

$$\boxed{\boldsymbol{\tau}_i = \boldsymbol{\gamma}_i'(0) = \boldsymbol{J\Phi}(\boldsymbol{u}_0) \cdot \boldsymbol{e}_i = \frac{\partial \boldsymbol{\Phi}}{\partial u_i}(\boldsymbol{u}_0), \qquad 1 \leq i \leq n\,.} \tag{6.25}$$

They are the columns of the Jacobian matrix of $\boldsymbol{\Phi}$ at $\boldsymbol{u}_0$. (Warning: the reader should pay attention not to confuse these with the row vectors $\nabla\varphi_i(\boldsymbol{u}_0)$, which are the gradients of the components of $\boldsymbol{\Phi} = (\varphi_i)_{1 \leq i \leq n}$.) Therefore by iii), the vectors $\boldsymbol{\tau}_i$, $1 \leq i \leq n$, are linearly independent (hence, non-zero), and so they form a basis of $\mathbb{R}^n$. We shall denote by $\boldsymbol{t}_i = \boldsymbol{\tau}_i / \|\boldsymbol{\tau}_i\|$ the corresponding unit vectors. If at every $P_0 \in A$ the tangent vectors are orthogonal, i.e., if the matrix $\boldsymbol{J\Phi}(\boldsymbol{u}_0)^T \boldsymbol{J\Phi}(\boldsymbol{u}_0)$ is diagonal, the change of variables will be called **orthogonal**. If so, $\{\boldsymbol{t}_1, \ldots, \boldsymbol{t}_n\}$ will be an orthonormal basis of $\mathbb{R}^n$ relative to the point P_0.

Properties ii) and iii) have yet another important consequence. The map $\boldsymbol{u} \mapsto \det \boldsymbol{J\Phi}(\boldsymbol{u})$ is continuous on A' because the determinant depends continuously upon the matrix' entries; moreover, $\det \boldsymbol{J\Phi}(\boldsymbol{u}) \neq 0$ for any $\boldsymbol{u} \in A'$. We therefore deduce

$$\det \boldsymbol{J\Phi}(\boldsymbol{u}) > 0\,, \quad \forall \boldsymbol{u} \in A'\,, \qquad \text{or} \qquad \det \boldsymbol{J\Phi}(\boldsymbol{u}) < 0\,, \quad \forall \boldsymbol{u} \in A'\,. \tag{6.26}$$

In fact, if $\det \boldsymbol{J\Phi}$ were both positive and negative on A', which is open and connected, then Theorem 4.30 would necessarily force the determinant to vanish at some point of A', contradicting iii).

Now we focus on low dimensions. In dimension 2, the first (second, respectively) of (6.26) says that for any $\boldsymbol{u}_0 \in A'$ the vector $\boldsymbol{\tau}_1$ can be aligned to $\boldsymbol{\tau}_2$ by a *counter-clockwise* (*clockwise*) rotation of $\theta \in (0, \pi]$. Identifying in fact each $\boldsymbol{\tau}_i$ with $\tilde{\boldsymbol{\tau}}_i = \boldsymbol{\tau}_i + 0\boldsymbol{k} \in \mathbb{R}^3$, we have

$$\tilde{\boldsymbol{\tau}}_1 \wedge \tilde{\boldsymbol{\tau}}_2 = \big(\det \boldsymbol{J\Phi}(\boldsymbol{u}_0)\big)\boldsymbol{k}\,, \tag{6.27}$$

orienting the triple $(\tilde{\boldsymbol{\tau}}_1, \tilde{\boldsymbol{\tau}}_2, \boldsymbol{k})$ positively (negatively).

In dimension 3, the triple $(\boldsymbol{\tau}_1, \boldsymbol{\tau}_2, \boldsymbol{\tau}_3)$ is positively-oriented (negatively-oriented) if the first (second) of (6.26) holds; in fact, $(\boldsymbol{\tau}_1 \wedge \boldsymbol{\tau}_2) \cdot \boldsymbol{\tau}_3 = \det \boldsymbol{J\Phi}(\boldsymbol{u}_0)$.

Changes of variable on a region in $\mathbb{R}^n$ are of two types, depending on the sign of the Jacobian determinant. If the sign is plus, the variable change is said **orientation-preserving**, if the sign is minus, **orientation-reversing**.

Example 6.31

The map

$$\boldsymbol{x} = \boldsymbol{\Phi}(\boldsymbol{u}) = \boldsymbol{A}\boldsymbol{u} + \boldsymbol{b}\,,$$

with $\boldsymbol{A}$ a non-singular matrix of order n, defines an **affine** change of variables in $\mathbb{R}^n$. For example, the change in the plane ($\boldsymbol{x} = (x, y)$, $\boldsymbol{u} = (u, v)$)

$$x = \frac{1}{\sqrt{2}}(v + u) + 2\,, \qquad y = \frac{1}{\sqrt{2}}(v - u) + 1\,,$$

is a translation of the origin to the point $(2, 1)$, followed by a counter-clockwise rotation of the axes by $\pi/4$ (Fig. 6.10). The coordinate lines $u =$ constant ($v =$ constant) are parallel to the bisectrix of the first and third quadrant (resp. second and fourth).

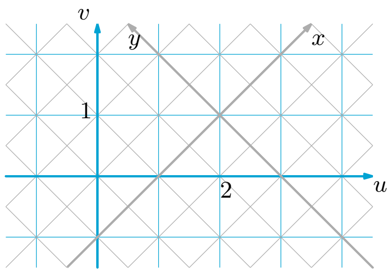


Figure 6.10. Affine change of variables in the plane

The new frame system is still orthogonal, confirmed by the matrix

$$\boldsymbol{A} = \begin{pmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix},$$

which is orthogonal since $\boldsymbol{A}^T\boldsymbol{A} = \boldsymbol{I}$. (In general, an affine change of variables is orthogonal if the associated matrix is orthogonal.) The change preserves the orientation, for $\det \boldsymbol{A} = 1 > 0$. □

6.6.1 Special frame systems

We examine the changes of variables associated to relevant transformations of the plane and of space.

i) **Polar coordinates.** Denote

$$\boldsymbol{\Phi} : [0,+\infty) \times \mathbb{R} \to \mathbb{R}^2\,, \qquad (r,\theta) \mapsto (x,y) = (r\cos\theta, r\sin\theta)$$

the map transforming polar coordinates into Cartesian ones (Fig. 6.11, left). It is differentiable, and its Jacobian

$$\boxed{J\boldsymbol{\Phi}(r,\theta) = \begin{pmatrix} \frac{\partial x}{\partial r} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial r} & \frac{\partial y}{\partial \theta} \end{pmatrix} = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}} \tag{6.28}$$

has determinant

$$\det J\boldsymbol{\Phi}(r,\theta) = r\cos^2\theta + r\sin^2\theta = r\,. \tag{6.29}$$

Its positivity on $(0,+\infty)\times\mathbb{R}$ makes the Jacobian invertible; in terms of r and θ, we have

$$\boxed{J\boldsymbol{\Phi}(r,\theta)^{-1} = \begin{pmatrix} \frac{\partial r}{\partial x} & \frac{\partial r}{\partial y} \\ \frac{\partial \theta}{\partial x} & \frac{\partial \theta}{\partial y} \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\dfrac{\sin\theta}{r} & \dfrac{\cos\theta}{r} \end{pmatrix}}\,. \tag{6.30}$$

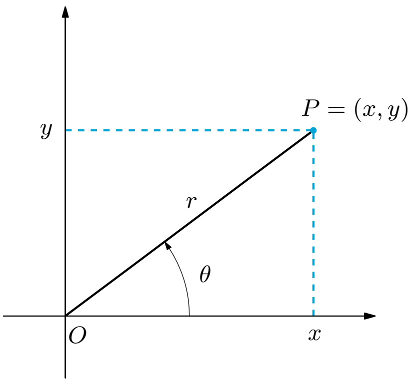
Therefore, $\boldsymbol{\Phi}$ is a change of variables in the plane if we choose, for example, $\mathcal{R} = \mathbb{R}^2$ and $\mathcal{R}' = (0,+\infty)\times(-\pi,\pi] \cup \{(0,0)\}$. The interior $A' = (0,+\infty)\times(-\pi,\pi)$ has open image under $\boldsymbol{\Phi}$ given by $A = \mathbb{R}^2 \setminus \{(x,0)\in\mathbb{R}^2 : x \le 0\}$, the plane minus the negative x-axis. The change is orthogonal, as $J\boldsymbol{\Phi}^T J\boldsymbol{\Phi} = \mathbf{diag}\,(1,r^2)$, and preserves the orientation, because $\det J\boldsymbol{\Phi} > 0$ on A'.

Rays emanating from the origin (for θ constant) and circles centred at the origin (r constant) are the coordinate lines. The tangent vectors $\boldsymbol{\tau}_i$, $i = 1,2$, columns of $J\boldsymbol{\Phi}$, will henceforth be indicated by $\boldsymbol{\tau}_r$, $\boldsymbol{\tau}_\theta$ and written as row vectors

$$\boldsymbol{\tau}_r = (\cos\theta, \sin\theta)\,, \qquad \boldsymbol{\tau}_\theta = r(-\sin\theta, \cos\theta)\,.$$

Normalising the second one we obtain an orthonormal basis of $\mathbb{R}^2$

$$\boxed{\boldsymbol{t}_r = (\cos\theta, \sin\theta)\,, \qquad \boldsymbol{t}_\theta = (-\sin\theta, \cos\theta)} \tag{6.31}$$



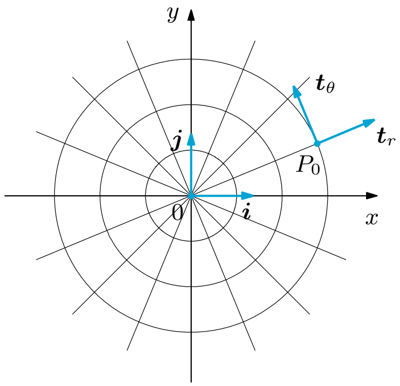


Figure 6.11. Polar coordinates in the plane (left); coordinate lines and unit tangent vectors (right)

formed by the unit tangent vectors to the coordinate lines, at each point $P \in \mathbb{R}^2 \setminus \{(0,0)\}$ (Fig. 6.11, right).

Take a scalar map $f(x,y)$ defined on a subset of $\mathbb{R}^2$ not containing the origin. In polar coordinates it will be

$$\tilde{f}(r,\theta) = f(r\cos\theta, r\sin\theta),$$

or $\tilde{f} = f \circ \boldsymbol{\Phi}$. Supposing f differentiable on its domain, the chain rule (Theorem 6.14 or, more precisely, formula (6.12)) gives

$$\nabla_{(r,\theta)}\tilde{f} = \nabla f_{(x,y)}\, \boldsymbol{J\Phi}\,,$$

whose inverse is $\nabla f_{(x,y)} = \nabla_{(r,\theta)}\tilde{f}\, \boldsymbol{J\Phi}^{-1}$. Dropping the symbol $\sim$ for simplicity, those identities become, by (6.28), (6.30),

$$\boxed{\frac{\partial f}{\partial r} = \frac{\partial f}{\partial x}\cos\theta + \frac{\partial f}{\partial y}\sin\theta\,, \qquad \frac{\partial f}{\partial \theta} = -\frac{\partial f}{\partial x} r\sin\theta + \frac{\partial f}{\partial y} r\cos\theta} \tag{6.32}$$

and

$$\boxed{\frac{\partial f}{\partial x} = \frac{\partial f}{\partial r}\cos\theta - \frac{\partial f}{\partial \theta}\frac{\sin\theta}{r}\,, \qquad \frac{\partial f}{\partial y} = \frac{\partial f}{\partial r}\sin\theta + \frac{\partial f}{\partial \theta}\frac{\cos\theta}{r}\,.} \tag{6.33}$$

Any vector field $\boldsymbol{g}(x,y) = g_1(x,y)\boldsymbol{i} + g_2(x,y)\boldsymbol{j}$ on a subset of $\mathbb{R}^2$ without the origin can be written, at each point $(x,y) = (r\cos\theta, r\sin\theta)$ of the domain, using the basis $\{\boldsymbol{t}_r, \boldsymbol{t}_\theta\}$:

$$\boldsymbol{g} = g_r\boldsymbol{t}_r + g_\theta\boldsymbol{t}_\theta \tag{6.34}$$

(here and henceforth the subscripts r, θ do not denote partial derivatives, rather components).

As the basis is orthonormal,

$$\begin{aligned} g_r &= \boldsymbol{g}\cdot\boldsymbol{t}_r = (g_1\boldsymbol{i}+g_2\boldsymbol{j})\cdot\boldsymbol{t}_r = g_1\cos\theta+g_2\sin\theta\,, \\ g_\theta &= \boldsymbol{g}\cdot\boldsymbol{t}_\theta = (g_1\boldsymbol{i}+g_2\boldsymbol{j})\cdot\boldsymbol{t}_\theta = -g_1\sin\theta+g_2\cos\theta\,. \end{aligned} \tag{6.35}$$

In particular, if $\boldsymbol{g}$ is the gradient of a differentiable function, $\mathbf{grad}\, f = \dfrac{\partial f}{\partial x}\boldsymbol{i}+\dfrac{\partial f}{\partial y}\boldsymbol{j}$, by (6.32) we have

$$g_r = \frac{\partial f}{\partial x}\cos\theta+\frac{\partial f}{\partial y}\sin\theta = \frac{\partial f}{\partial r}\,, \qquad g_\theta = -\frac{\partial f}{\partial x}\sin\theta+\frac{\partial f}{\partial y}\cos\theta = \frac{1}{r}\frac{\partial f}{\partial \theta}\,;$$

therefore the gradient in polar coordinates reads

$$\boxed{\mathbf{grad}\, f = \frac{\partial f}{\partial r}\boldsymbol{t}_r + \frac{1}{r}\frac{\partial f}{\partial \theta}\boldsymbol{t}_\theta\,.} \tag{6.36}$$

The divergence of a differentiable vector field $\boldsymbol{g}$ in polar coordinates is, similarly,

$$\begin{aligned} \operatorname{div}\boldsymbol{g} &= \frac{\partial g_1}{\partial x}+\frac{\partial g_2}{\partial y} = \frac{\partial g_1}{\partial r}\cos\theta - \frac{\partial g_1}{\partial \theta}\frac{\sin\theta}{r} + \frac{\partial g_2}{\partial r}\sin\theta + \frac{\partial g_2}{\partial \theta}\frac{\cos\theta}{r} \\ &= \frac{\partial}{\partial r}(g_1\cos\theta+g_2\sin\theta)+\frac{1}{r}\Big[\frac{\partial}{\partial\theta}(-g_1\sin\theta+g_2\cos\theta)+(g_1\cos\theta+g_2\sin\theta)\Big]\,; \end{aligned}$$

using (6.34) and (6.35), we find

$$\boxed{\operatorname{div}\boldsymbol{g} = \frac{\partial g_r}{\partial r}+\frac{1}{r}g_r+\frac{1}{r}\frac{\partial g_\theta}{\partial\theta}\,.} \tag{6.37}$$

Combining this with (6.36) yields the polar representation of the Laplacian Δf of a $\mathcal{C}^2$ map on a plane region without the origin. Setting $\boldsymbol{g} = \mathbf{grad}\, f$ in (6.37),

$$\boxed{\Delta f = \frac{\partial^2 f}{\partial r^2}+\frac{1}{r}\frac{\partial f}{\partial r}+\frac{1}{r^2}\frac{\partial^2 f}{\partial\theta^2}\,.} \tag{6.38}$$

Let us return to (6.34), and notice that – in contrast to the canonical unit vectors $\boldsymbol{i}$, $\boldsymbol{j}$ – the unit vectors $\boldsymbol{t}_r$, $\boldsymbol{t}_\theta$ vary from point to point; from (6.31), in particular, follow

$$\boxed{\frac{\partial \boldsymbol{t}_r}{\partial r} = 0\,, \qquad \frac{\partial \boldsymbol{t}_r}{\partial \theta} = \boldsymbol{t}_\theta\,; \qquad\quad \frac{\partial \boldsymbol{t}_\theta}{\partial r} = 0\,, \qquad \frac{\partial \boldsymbol{t}_\theta}{\partial \theta} = -\boldsymbol{t}_r\,.} \tag{6.39}$$

To differentiate $\boldsymbol{g}$ with respect to r and θ then, we will use the usual product rule

$$\begin{aligned} \frac{\partial \boldsymbol{g}}{\partial r} &= \frac{\partial g_r}{\partial r}\boldsymbol{t}_r + \frac{\partial g_\theta}{\partial r}\boldsymbol{t}_\theta\,, \\ \frac{\partial \boldsymbol{g}}{\partial \theta} &= \Big(\frac{\partial g_r}{\partial \theta}-g_\theta\Big)\boldsymbol{t}_r + \Big(\frac{\partial g_\theta}{\partial \theta}+g_r\Big)\boldsymbol{t}_\theta\,. \end{aligned}$$

ii) **Cylindrical coordinates.** Consider the $\mathcal{C}^\infty$ transformation

$$\mathbf{\Phi} : [0, +\infty) \times \mathbb{R}^2 \to \mathbb{R}^3 , \qquad (r, \theta, t) \mapsto (x, y, z) = (r \cos \theta, r \sin \theta, t)$$

describing the passage from cylindrical to Cartesian coordinates (Fig. 6.12, left). The Jacobian

$$\boxed{\boldsymbol{J\Phi}(r, \theta, t) = \begin{pmatrix} \cos\theta & -r\sin\theta & 0 \\ \sin\theta & r\cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}} \tag{6.40}$$

has determinant

$$\det \boldsymbol{J\Phi}(r, \theta, t) = r , \tag{6.41}$$

strictly positive on $(0, +\infty) \times \mathbb{R}^2$, so the matrix is invertible. Moreover, $\boldsymbol{J\Phi}^T \boldsymbol{J\Phi} = \mathbf{diag}\,(1, r^2, 1)$.

Thus $\mathbf{\Phi}$ is an orthogonal, orientation-preserving change of coordinates from $\mathcal{R}' = (0, +\infty) \times (-\pi, \pi] \times \mathbb{R} \cup \{(0, 0, t) : t \in \mathbb{R}\}$ to $\mathcal{R} = \mathbb{R}^3$. The interior of $\mathcal{R}'$ is $A' = (0, +\infty) \times (-\pi, \pi) \times \mathbb{R}$, whose image under $\mathbf{\Phi}$ is the open set $A = \mathbb{R}^3 \setminus \{(x, 0, z) \in \mathbb{R}^3 : x \leq 0,\ z \in \mathbb{R}\}$, the whole space minus half a plane.

The coordinate lines are: horizontal rays emanating from the z-axis, horizontal circles centred along the z-axis, and vertical lines. The corresponding unit tangent vectors at $P \in \mathbb{R}^3 \setminus \{(0, 0, z) : z \in \mathbb{R}\}$ are, with the obvious notations,

$$\boxed{\boldsymbol{t}_r = (\cos\theta, \sin\theta, 0) , \qquad \boldsymbol{t}_\theta = (-\sin\theta, \cos\theta, 0) , \qquad \boldsymbol{t}_t = (0, 0, 1) .} \tag{6.42}$$

They form an orthonormal frame at P (Fig. 6.12, right).

A scalar map $f(x, y, z)$, differentiable on a subset of $\mathbb{R}^3$ not containing the axis z, is

$$\tilde{f}(r, \theta, t) = f(r \cos\theta, r \sin\theta, t) ;$$

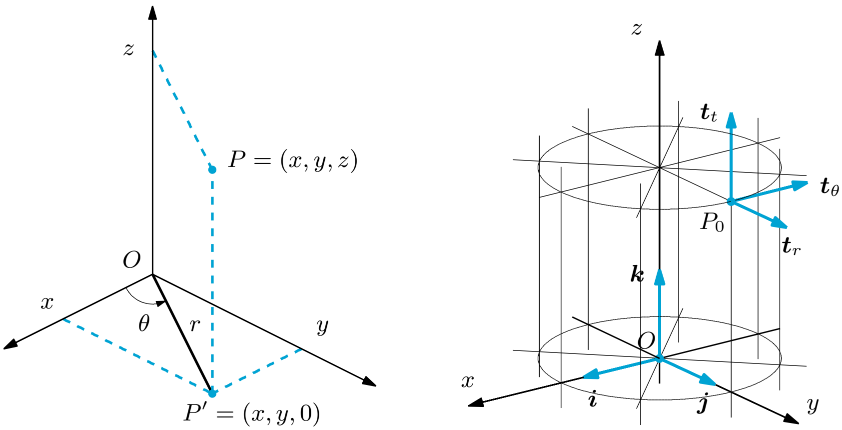


Figure 6.12. Cylindrical coordinates in space (left); coordinate lines and unit tangent vectors (right)

spelling out $\nabla_{(r,\theta,t)}\tilde{f} = \nabla f_{(x,y,z)} \boldsymbol{J\Phi}$ gives

$$\boxed{\frac{\partial f}{\partial r} = \frac{\partial f}{\partial x}\cos\theta + \frac{\partial f}{\partial y}\sin\theta\,, \quad \frac{\partial f}{\partial \theta} = -\frac{\partial f}{\partial x} r\sin\theta + \frac{\partial f}{\partial y} r\cos\theta\,, \quad \frac{\partial f}{\partial t} = \frac{\partial f}{\partial z}\,.}$$

The inverse formulas are

$$\boxed{\frac{\partial f}{\partial x} = \frac{\partial f}{\partial r}\cos\theta - \frac{\partial f}{\partial \theta}\frac{\sin\theta}{r}\,, \qquad \frac{\partial f}{\partial y} = \frac{\partial f}{\partial r}\sin\theta + \frac{\partial f}{\partial \theta}\frac{\cos\theta}{r}\,, \qquad \frac{\partial f}{\partial z} = \frac{\partial f}{\partial t}\,.}$$

At last, this is how the basic differential operators look like in cylindrical coordinates

$$\boxed{\begin{aligned}
&\mathbf{grad}\, f = \frac{\partial f}{\partial r}\boldsymbol{t}_r + \frac{1}{r}\frac{\partial f}{\partial \theta}\boldsymbol{t}_\theta + \frac{\partial f}{\partial t}\boldsymbol{t}_t\,,\\
&\operatorname{div} \boldsymbol{f} = \frac{\partial f_r}{\partial r} + \frac{1}{r} f_r + \frac{1}{r}\frac{\partial f_\theta}{\partial \theta} + \frac{\partial f_t}{\partial t}\,,\\
&\mathbf{curl}\, \boldsymbol{f} = \Big(\frac{1}{r}\frac{\partial f_t}{\partial \theta} - \frac{\partial f_\theta}{\partial t}\Big)\boldsymbol{t}_r + \Big(\frac{\partial f_r}{\partial t} - \frac{\partial f_t}{\partial r}\Big)\boldsymbol{t}_\theta + \Big(\frac{\partial f_\theta}{\partial r} - \frac{1}{r}\frac{\partial f_r}{\partial \theta} + \frac{1}{r} f_\theta\Big)\boldsymbol{t}_t\,,\\
&\Delta f = \frac{\partial^2 f}{\partial r^2} + \frac{1}{r}\frac{\partial f}{\partial r} + \frac{1}{r^2}\frac{\partial^2 f}{\partial \theta^2} + \frac{\partial^2 f}{\partial t^2}\,.
\end{aligned}}$$

iii) **Spherical coordinates.** The function

$$\boldsymbol{\Phi} : [0,+\infty) \times \mathbb{R}^2 \to \mathbb{R}^3\,, \quad (r,\varphi,\theta) \mapsto (x,y,z) = (r\sin\varphi\cos\theta, r\sin\varphi\sin\theta, r\cos\varphi)$$

is differentiable infinitely many times, and describes the passage from spherical coordinates to Cartesian coordinates (Fig. 6.13, left). As

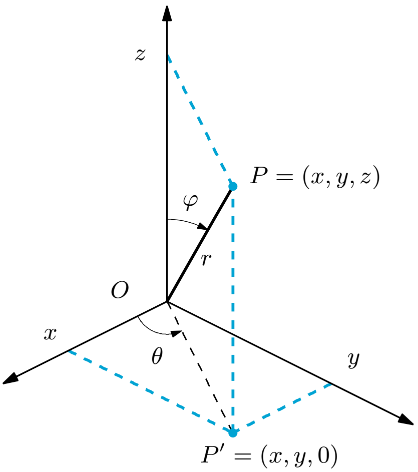
$$\boxed{\boldsymbol{J\Phi}(r,\varphi,\theta) = \begin{pmatrix} \sin\varphi\cos\theta & r\cos\varphi\cos\theta & -r\sin\varphi\sin\theta \\ \sin\varphi\sin\theta & r\cos\varphi\sin\theta & r\sin\varphi\cos\theta \\ \cos\varphi & -r\sin\varphi & 0 \end{pmatrix},} \tag{6.43}$$

we compute the determinant by expanding along the last row; recalling $\sin^2\theta + \cos^2\theta = 1$ and $\sin^2\varphi + \cos^2\varphi = 1$, we find

$$\det \boldsymbol{J\Phi}(r,\varphi,\theta) = r^2\sin\varphi\,, \tag{6.44}$$

strictly positive on $(0,+\infty)\times(0,\pi)\times\mathbb{R}$. The Jacobian is invertible on that domain. Furthermore, $\boldsymbol{J\Phi}^T\boldsymbol{J\Phi} = \mathbf{diag}\,(1, r(\cos^2\varphi + r\sin^2\varphi), r^2\sin\varphi)$.

Then $\boldsymbol{\Phi}$ is an orthogonal, orientation-preserving change of variables mapping $\mathcal{R}' = (0,+\infty)\times[0,\pi]\times(-\pi,\pi] \cup \{(0,0,0)\}$ onto $\mathcal{R} = \mathbb{R}^3$. The interior of $\mathcal{R}'$



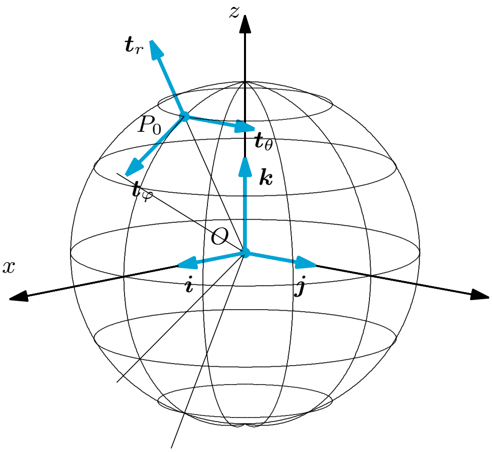


Figure 6.13. Spherical coordinates in space (left); coordinate lines and unit tangent vectors (right)

is $A' = (0,+\infty) \times (0,\pi) \times (-\pi,\pi)$, whose image is in turn the open set $A = \mathbb{R}^3 \setminus \{(x,0,z) \in \mathbb{R}^3 : x \le 0,\ z \in \mathbb{R}\}$, as for cylindrical coordinates.

There are three types of coordinate lines, namely rays from the origin, vertical half-circles centred at the origin (the Earth's meridians), and horizontal circles centred on the z-axis (the parallels). The unit tangent vectors at $P \in \mathbb{R}^3 \setminus \{(0,0,0)\}$ result from normalising the columns of $\boldsymbol{J\Phi}$

$$\boxed{\begin{aligned} \boldsymbol{t}_r &= (\sin\varphi\cos\theta, \sin\varphi\sin\theta, \cos\varphi)\,, \\ \boldsymbol{t}_\varphi &= (\cos\varphi\cos\theta, \cos\varphi\sin\theta, -\sin\varphi)\,, \\ \boldsymbol{t}_\theta &= (-\sin\theta, \cos\theta, 0)\,. \end{aligned}} \tag{6.45}$$

They are an orthonormal basis of $\mathbb{R}^3$ at the point P (Fig. 6.13, right).

Given a scalar map $f(x,y,z)$, differentiable away from the origin, in spherical coordinates

$$\tilde{f}(r,\varphi,\theta) = f(r\sin\varphi\cos\theta, r\sin\varphi\sin\theta, r\cos\varphi)\,,$$

we express $\nabla_{(r,\varphi,\theta)}\tilde{f} = \nabla f_{(x,y,z)}\boldsymbol{J\Phi}$ (dropping $\sim$) as

$$\boxed{\begin{aligned} \frac{\partial f}{\partial r} &= \frac{\partial f}{\partial x}\sin\varphi\cos\theta + \frac{\partial f}{\partial y}\sin\varphi\sin\theta + \frac{\partial f}{\partial z}\cos\varphi \\ \frac{\partial f}{\partial \varphi} &= \frac{\partial f}{\partial x}r\cos\varphi\cos\theta + \frac{\partial f}{\partial y}r\cos\varphi\sin\theta - \frac{\partial f}{\partial z}r\sin\varphi \\ \frac{\partial f}{\partial \theta} &= -\frac{\partial f}{\partial x}r\sin\varphi\sin\theta + \frac{\partial f}{\partial y}r\sin\varphi\cos\theta\,. \end{aligned}}$$

The inverse relationships read

$$\begin{aligned}
\frac{\partial f}{\partial x} &= \frac{\partial f}{\partial r}\sin\varphi\cos\theta + \frac{\partial f}{\partial\varphi}\frac{\cos\varphi\cos\theta}{r} - \frac{\partial f}{\partial\theta}\frac{\sin\theta}{r\sin\varphi}\\
\frac{\partial f}{\partial y} &= \frac{\partial f}{\partial r}\sin\varphi\sin\theta + \frac{\partial f}{\partial\varphi}\frac{\cos\varphi\sin\theta}{r} + \frac{\partial f}{\partial\theta}\frac{\cos\theta}{r\sin\varphi}\\
\frac{\partial f}{\partial z} &= \frac{\partial f}{\partial r}\cos\varphi - \frac{\partial f}{\partial\varphi}\frac{\sin\varphi}{r}\,.
\end{aligned}$$

The usual differential operators in spherical coordinates read

$$\begin{aligned}
&\mathbf{grad}\, f = \frac{\partial f}{\partial r}\boldsymbol{t}_r + \frac{1}{r}\frac{\partial f}{\partial\varphi}\boldsymbol{t}_\varphi + \frac{1}{r\sin\varphi}\frac{\partial f}{\partial\theta}\boldsymbol{t}_\theta\,,\\
&\mathrm{div}\,\boldsymbol{f} = \frac{\partial f_r}{\partial r} + \frac{2}{r}f_r + \frac{1}{r}\frac{\partial f_\varphi}{\partial\varphi} + \frac{\tan\varphi}{r}f_\varphi + \frac{1}{r\sin\varphi}\frac{\partial f_\theta}{\partial\theta}\,,\\
&\mathbf{curl}\,\boldsymbol{f} = \Big(\frac{1}{r}\frac{\partial f_\theta}{\partial\varphi} + \frac{\tan\varphi}{r}f_\theta - \frac{\partial f_\varphi}{\partial\theta}\Big)\boldsymbol{t}_r + \Big(\frac{1}{r\sin\varphi}\frac{\partial f_r}{\partial\theta} - \frac{\partial f_\theta}{\partial r} - \frac{1}{r}f_\theta\Big)\boldsymbol{t}_\varphi\\
&\qquad + \Big(\frac{\partial f_\varphi}{\partial r} + \frac{1}{r}f_\varphi - \frac{\partial f_r}{\partial\varphi}\Big)\boldsymbol{t}_\theta\,,\\
&\Delta f = \frac{\partial^2 f}{\partial r^2} + \frac{2}{r}\frac{\partial f}{\partial r} + \frac{1}{r^2}\frac{\partial^2 f}{\partial\varphi^2} + \frac{\tan\varphi}{r^2}\frac{\partial f}{\partial\varphi} + \frac{1}{r^2\sin^2\varphi}\frac{\partial^2 f}{\partial\theta^2}\,.
\end{aligned}$$

6.7 Regular surfaces

Surfaces, and in particular compact ones (defined by a compact region $\mathcal{R}$), were introduced in Sect. 4.7.

Definition 6.32 *A* **surface** $\boldsymbol{\sigma} : \mathcal{R} \to \mathbb{R}^3$ *is* **regular** *if* $\boldsymbol{\sigma}$ *is* $\mathcal{C}^1$ *on* $A = \mathring{\mathcal{R}}$ *and the Jacobian matrix* $\boldsymbol{J\sigma}$ *has maximal rank* $(= 2)$ *at every point of* A.
A compact surface is **regular** *if it is the restriction to* $\mathcal{R}$ *of a regular surface defined on an open set containing* $\mathcal{R}$.

The condition on $\boldsymbol{J\sigma}$ is equivalent to the fact that the vectors $\dfrac{\partial\boldsymbol{\sigma}}{\partial u}(u_0, v_0)$ and $\dfrac{\partial\boldsymbol{\sigma}}{\partial v}(u_0, v_0)$ are linearly independent for any $(u_0, v_0) \in A$. By Definition 6.1, such vectors form the columns of the Jacobian $\boldsymbol{J\sigma}$

$$\boldsymbol{J\sigma} = \Big(\frac{\partial\boldsymbol{\sigma}}{\partial u} \quad \frac{\partial\boldsymbol{\sigma}}{\partial v}\Big)\,. \tag{6.46}$$

Examples 6.33

i) The surfaces seen in Examples 4.37 i), iii), iv) are regular, as simple calculations will show.

ii) The surface of Example 4.37 ii) is regular if (and only if) φ is $\mathcal{C}^1$ on A. If so,

$$\boldsymbol{J\sigma} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ \dfrac{\partial \varphi}{\partial u} & \dfrac{\partial \varphi}{\partial v} \end{pmatrix},$$

whose first two rows grant the matrix rank 2.

iii) The surface $\boldsymbol{\sigma} : \mathbb{R} \times [0, 2\pi] \to \mathbb{R}^3$,

$$\boldsymbol{\sigma}(u, v) = u \cos v\, \boldsymbol{i} + u \sin v \boldsymbol{j} + u\, \boldsymbol{k}$$

parametrises the cone

$$x^2 + y^2 - z^2 = 0\,.$$

Its Jacobian is

$$\boldsymbol{J\sigma} = \begin{pmatrix} \cos v & -u \sin v \\ \sin v & u \cos v \\ 1 & 0 \end{pmatrix}.$$

As the determinants of the first minors are u, $u \sin v$, $-u \cos v$, the surface is not regular: at points $(u_0, v_0) = (0, v_0)$, mapped to the cone's apex, all minors are singular. □

Before we continue, let us point out the use of terminology. Although Definition 4.36 privileges the analytical aspects, the prevailing (by standard practice) geometrical viewpoint uses the term surface to mean the trace $\Sigma \subset \mathbb{R}^3$ as well, retaining for the function $\boldsymbol{\sigma}$ the role of **parametrisation** of the surface. We follow the mainstream and adopt this language, with the additional assumption that all parametrisations $\boldsymbol{\sigma}$ be simple. In such a way, many subsequent notions will have a more immediate, and intuitive, geometrical representation. With this matter settled, we may now see a definition.

Definition 6.34 *A subset Σ of $\mathbb{R}^3$ is a* **regular and simple surface** *if Σ admits a regular and simple parametrisation $\sigma : \mathcal{R} \to \Sigma$.*

Let then $\Sigma \subset \mathbb{R}^3$ be a regular simple surface parametrised by $\boldsymbol{\sigma} : \mathcal{R} \to \Sigma$, and $P_0 = \boldsymbol{\sigma}(u_0, v_0)$ the point on Σ image of $(u_0, v_0) \in A = \overset{\circ}{\mathcal{R}}$. Since the vectors $\dfrac{\partial \boldsymbol{\sigma}}{\partial u}(u_0, v_0)$, $\dfrac{\partial \boldsymbol{\sigma}}{\partial v}(u_0, v_0)$ are by hypothesis linearly independent, we introduce the map $\Pi : \mathbb{R}^2 \to \mathbb{R}^3$,

$$\boxed{\Pi(u, v) = \boldsymbol{\sigma}(u_0, v_0) + \frac{\partial \boldsymbol{\sigma}}{\partial u}(u_0, v_0)\,(u - u_0) + \frac{\partial \boldsymbol{\sigma}}{\partial v}(u_0, v_0)\,(v - v_0)\,,} \qquad (6.47)$$

that parametrises a plane through P_0 (recall Example 4.37 i)). It is called the **tangent plane** to the surface at P_0. Justifying the notation, notice that the functions

$$u \mapsto \boldsymbol{\sigma}(u, v_0) \qquad \text{and} \qquad v \mapsto \boldsymbol{\sigma}(u_0, v)$$

define two regular curves lying on Σ and passing through P_0. Their tangent vectors at P_0 are $\dfrac{\partial \boldsymbol{\sigma}}{\partial u}(u_0, v_0)$ and $\dfrac{\partial \boldsymbol{\sigma}}{\partial v}(u_0, v_0)$; therefore the tangent lines to such curves at P_0 lie on the tangent plane, and actually they span it by linear combination. In general, the tangent plane contains the tangent vector to any regular curve passing through P_0 and lying on the surface Σ.

Definition 6.35 *The vector*

$$\boldsymbol{\nu}(u_0, v_0) = \frac{\partial \boldsymbol{\sigma}}{\partial u}(u_0, v_0) \wedge \frac{\partial \boldsymbol{\sigma}}{\partial v}(u_0, v_0) \tag{6.48}$$

is the surface's **normal** (*vector*) *at* P_0. *The corresponding* **unit normal** *will be indicated by* $\boldsymbol{n}(u_0, v_0) = \dfrac{\boldsymbol{\nu}(u_0, v_0)}{\|\boldsymbol{\nu}(u_0, v_0)\|}$.

Due to the surface's regularity, $\boldsymbol{\nu}(u_0, v_0)$ is non-zero. It is also orthogonal to $\dfrac{\partial \boldsymbol{\sigma}}{\partial u}(u_0, v_0)$ and $\dfrac{\partial \boldsymbol{\sigma}}{\partial v}(u_0, v_0)$, so if we compute it at P_0 it will be orthogonal to the tangent plane at P_0; this explains the name (Fig. 6.14).

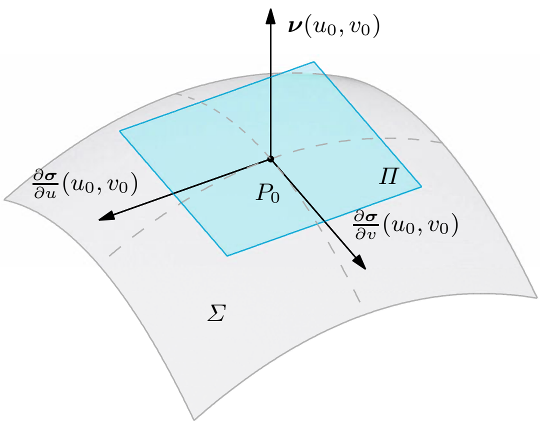


Figure 6.14. Tangent plane Π and normal vector $\boldsymbol{\nu}$ at P_0

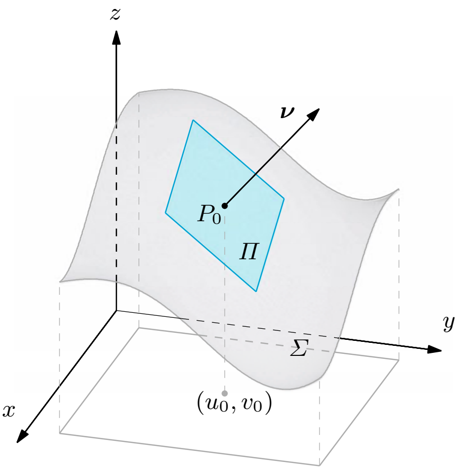


Figure 6.15. Normal vector of a surface

Example 6.36

i) The regular surface $\boldsymbol{\sigma} : \mathcal{R} \to \mathbb{R}^3$, $\boldsymbol{\sigma}(u,v) = u\,\boldsymbol{i} + v\,\boldsymbol{j} + \varphi(u,v)\,\boldsymbol{k}$, with $\varphi \in \mathcal{C}^1(\mathcal{R})$ (see Example 6.33 ii)), has

$$\boxed{\boldsymbol{\nu}(u,v) = -\frac{\partial \varphi}{\partial u}(u,v)\,\boldsymbol{i} - \frac{\partial \varphi}{\partial v}(u,v)\,\boldsymbol{j} + \boldsymbol{k}\,.} \tag{6.49}$$

Note that, at any point, the normal vector points upwards, because $\boldsymbol{\nu} \cdot \boldsymbol{k} > 0$ (Fig. 6.15).

ii) Let $\boldsymbol{\sigma} : [0,\pi] \times [0, 2\pi] \to \mathbb{R}^3$ be the parametrisation of the ellipsoid with centre the origin and semi-axes $a, b, c > 0$ (see Example 4.37 iv)). Then

$$\frac{\partial \boldsymbol{\sigma}}{\partial u}(u,v) = a\cos u \cos v\,\boldsymbol{i} + b\cos u \sin v\,\boldsymbol{j} - c\sin u\,\boldsymbol{k}\,,$$
$$\frac{\partial \boldsymbol{\sigma}}{\partial v}(u,v) = -a\sin u \sin v\,\boldsymbol{i} + b\sin u \cos v\,\boldsymbol{j} + 0\boldsymbol{k}\,,$$

whence

$$\boldsymbol{\nu}(u,v) = bc\sin^2 u \cos v\,\boldsymbol{i} + ac\sin^2 u \sin v\,\boldsymbol{j} + ab\sin u \cos u\,\boldsymbol{k}\,.$$

If the surface is a sphere of radius r (when $a = b = c = r$),

$$\boldsymbol{\nu}(u,v) = r\sin u(r\sin u \cos v\,\boldsymbol{i} + r\sin u \sin v\,\boldsymbol{j} + r\cos u\,\boldsymbol{k})$$

so the normal $\boldsymbol{\nu}$ at $\boldsymbol{x}$ is proportional to $\boldsymbol{x}$, and thus aligned with the radial vector. □

Just like with the tangent to a curve, one could prove that the tangent plane is intrinsic to the surface, in other words it does not depend on the parametrisation. As a result the normal's direction is intrinsic as well, while length and orientation vary with the chosen parametrisation.

6.7.1 Changing parametrisation

Suppose Σ is regular, with two parametrisations $\boldsymbol{\sigma} : \mathcal{R} \to \Sigma$ and $\widetilde{\boldsymbol{\sigma}} : \widetilde{\mathcal{R}} \to \Sigma$.

Definition 6.37 *The* **parametrisations** $\widetilde{\boldsymbol{\sigma}}$ *and* $\boldsymbol{\sigma}$ *are* **congruent** *if there is a change of variables* $\Phi : \widetilde{\mathcal{R}} \to \mathcal{R}$ *such that* $\widetilde{\boldsymbol{\sigma}} = \boldsymbol{\sigma} \circ \Phi$. *If* Σ *is a compact surface,* $\boldsymbol{\Phi}$ *is required to be the restriction of a change of variables between open sets containing* $\widetilde{\mathcal{R}}$ *and* $\mathcal{R}$.

Although the definition does not require the parametrisations to be simple, we shall assume they are throughout. The property of being regular and simple is preserved by congruence.

Bearing in mind the discussion of Sect. 6.6, if A and $\widetilde{A}$ are the interiors of $\mathcal{R}$ and $\widetilde{\mathcal{R}}$, then either

$$\det \boldsymbol{J\Phi} > 0 \quad \text{on} \quad \widetilde{A} \qquad \text{or} \qquad \det \boldsymbol{J\Phi} < 0 \quad \text{on} \quad \widetilde{A}\,.$$

In the former case $\widetilde{\boldsymbol{\sigma}}$ is **equivalent** to $\boldsymbol{\sigma}$, while in the latter $\widetilde{\boldsymbol{\sigma}}$ is **anti-equivalent** to $\boldsymbol{\sigma}$. In other terms, an orientation-preserving (orientation-reversing) change of variables induces a parametrisation equivalent (anti-equivalent) to the given one.

The tangent vectors $\dfrac{\partial \widetilde{\boldsymbol{\sigma}}}{\partial \tilde{u}}$ and $\dfrac{\partial \widetilde{\boldsymbol{\sigma}}}{\partial \tilde{v}}$ can be easily expressed as linear combinations of $\dfrac{\partial \boldsymbol{\sigma}}{\partial u}$ and $\dfrac{\partial \boldsymbol{\sigma}}{\partial v}$. As it happens, recalling (6.46) and the chain rule (6.9), and omitting to write the point $(u_0, v_0) = \Phi(\tilde{u}_0, \tilde{v}_0)$ of differentiation, we have

$$\begin{pmatrix} \dfrac{\partial \widetilde{\boldsymbol{\sigma}}}{\partial \tilde{u}} & \dfrac{\partial \widetilde{\boldsymbol{\sigma}}}{\partial \tilde{v}} \end{pmatrix} = \begin{pmatrix} \dfrac{\partial \boldsymbol{\sigma}}{\partial u} & \dfrac{\partial \boldsymbol{\sigma}}{\partial v} \end{pmatrix} \boldsymbol{J\Phi}\,;$$

setting $\boldsymbol{J\Phi} = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix}$, these become

$$\frac{\partial \widetilde{\boldsymbol{\sigma}}}{\partial \tilde{u}} = m_{11} \frac{\partial \boldsymbol{\sigma}}{\partial u} + m_{21} \frac{\partial \boldsymbol{\sigma}}{\partial v}\,, \qquad \frac{\partial \widetilde{\boldsymbol{\sigma}}}{\partial \tilde{v}} = m_{12} \frac{\partial \boldsymbol{\sigma}}{\partial u} + m_{22} \frac{\partial \boldsymbol{\sigma}}{\partial v}\,.$$

They show, on one hand, that $\dfrac{\partial \widetilde{\boldsymbol{\sigma}}}{\partial \tilde{u}}$ and $\dfrac{\partial \widetilde{\boldsymbol{\sigma}}}{\partial \tilde{v}}$ generate the same plane of $\dfrac{\partial \boldsymbol{\sigma}}{\partial u}$ and $\dfrac{\partial \boldsymbol{\sigma}}{\partial v}$, as prescribed by (6.47); therefore

the tangent plane to a surface Σ is invariant for congruent parametrisations.

On the other hand, the previous expressions and property (4.8) imply

$$\frac{\partial \widetilde{\boldsymbol{\sigma}}}{\partial \tilde{u}} \wedge \frac{\partial \widetilde{\boldsymbol{\sigma}}}{\partial \tilde{v}} = m_{11} m_{12} \frac{\partial \boldsymbol{\sigma}}{\partial u} \wedge \frac{\partial \boldsymbol{\sigma}}{\partial u} + m_{11} m_{22} \frac{\partial \boldsymbol{\sigma}}{\partial u} \wedge \frac{\partial \boldsymbol{\sigma}}{\partial v} +$$

$$+m_{21}m_{12}\frac{\partial\boldsymbol{\sigma}}{\partial v}\wedge\frac{\partial\boldsymbol{\sigma}}{\partial u}+m_{21}m_{22}\frac{\partial\boldsymbol{\sigma}}{\partial v}\wedge\frac{\partial\boldsymbol{\sigma}}{\partial v}$$

$$=(m_{11}m_{22}-m_{12}m_{21})\frac{\partial\boldsymbol{\sigma}}{\partial u}\wedge\frac{\partial\boldsymbol{\sigma}}{\partial v}\,.$$

From these follows that the normal vectors at the point P_0 satisfy

$$\boxed{\widetilde{\boldsymbol{\nu}}=(\det \boldsymbol{J}\varPhi)\,\boldsymbol{\nu}\,,} \tag{6.50}$$

which in turn corroborates that

two congruent parametrisations generate parallel normal vectors; the two orientations are the same for equivalent parametrisations (orientation-preserving change of variables), otherwise they are opposite (orientation-reversing change).

6.7.2 Orientable surfaces

Proposition 6.27 guarantees two parametrisations of the same regular, simple curve $\varGamma$ of $\mathbb{R}^m$ are congruent, hence there are always two orientations to choose from. The analogue result for surfaces cannot hold, and we will see counterexamples (the Möbius strip, the Klein bottle). Thus the following definition makes sense.

Definition 6.38 *A regular simple surface $\varSigma\subset\mathbb{R}^3$ is said* **orientable** *if any two regular and simple parametrisations are congruent.*

The name comes from the fact that the surface may be endowed with two **orientations** (naïvely, the crossing directions) depending on the normal vector. All equivalent parametrisations will have the same orientation, whilst anti-equivalent ones will have opposite orientations. It can be proved that a regular and simple surface parametrised by an *open* plane region $\mathcal{R}$ is always orientable.

When the regular, simple surface is parametrised by a non-open region, the picture becomes more complicated and the surface may, or not, be orientable. With the help of the next two examples we shed some light on the matter. Consider first the cylindrical strip of radius 1 and height 2, see Fig. 6.16, left.

We parametrise this compact surface by the map

$$\boldsymbol{\sigma}:[0,2\pi]\times[-1,1]\to\mathbb{R}^3\,,\qquad \boldsymbol{\sigma}(u,v)=\cos u\,\boldsymbol{i}+\sin u\,\boldsymbol{j}+v\boldsymbol{k}\,, \tag{6.51}$$

regular and simple because, for instance, injective on $[0,2\pi)\times[-1,1]$. The associated normal

$$\boldsymbol{\nu}(u,v)=\cos u\,\boldsymbol{i}+\sin u\,\boldsymbol{j}+0\boldsymbol{k}$$

constantly points outside the cylinder. To nurture visual intuition, we might say that a person walking on the surface (staying on the plane xy, for instance) can return to the starting point always keeping the normal vector on the same side. The surface is orientable, it has two sides – an inside and an outside – and if the person wanted to go to the opposite side it would have to cross the top rim $v = 1$ or the bottom one $v = -1$. (The rims in question are the surface's boundaries, as we shall see below.)

The second example is yet another strip, called **Möbius strip**, and shown in Fig. 6.16, right. To construct it, start from the previous cylinder, cut it along a vertical segment, and then glue the ends back together after twisting one by 180°. The precise parametrisation $\boldsymbol{\sigma} : [0, 2\pi] \times [-1, 1] \to \mathbb{R}^3$ is

$$\boldsymbol{\sigma}(u, v) = \big(1 - \frac{v}{2}\cos\frac{u}{2}\big)\cos u\,\boldsymbol{i} + \big(1 - \frac{v}{2}\cos\frac{u}{2}\big)\sin u\,\boldsymbol{j} - \frac{v}{2}\sin\frac{u}{2}\,\boldsymbol{k}\,.$$

For $u = 0$, i.e., at the point $P_0 = (1, 0, 0) = \boldsymbol{\sigma}(0, 0)$, the normal vector is $\boldsymbol{\nu} = \frac{1}{2}\boldsymbol{k}$. As u increases, the normal varies with continuity, except that as we go back to $P_0 = \boldsymbol{\sigma}(2\pi, 0)$ after a complete turn, we have $\boldsymbol{\nu} = -\frac{1}{2}\boldsymbol{k}$, opposite to the one we started with. This means that our imaginary friend starts from P_0 and after one loop returns to the same point (without crossing the rim) upside down! The Möbius strip is a **non-orientable** surface; it has one side only and one boundary: starting at any point of the boundary and going around the origin twice, for instance, we reach all points of the boundary, in contrast to what happens for the cylinder.

At any rate the Möbius strip embodies a somehow "pathological" situation. Surfaces (compact or not) delimiting elementary solids – spheres, ellipsoid, cylinders, cones and the like – are all orientable. More generally,

Proposition 6.39 *Every regular simple surface Σ contained in the boundary $\partial\Omega$ of an open, connected and bounded set $\Omega \subset \mathbb{R}^3$ is orientable.*

We can thus choose for Σ either the orientation from inside towards outside Ω, or the converse. In the first case the unit normal $\boldsymbol{n}$ **points outwards**, in the second it **points inwards**.

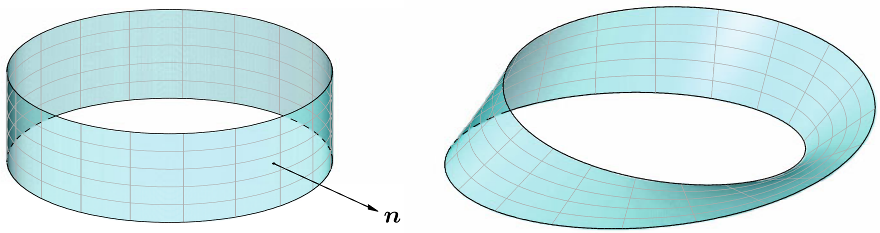


Figure 6.16. Cylinder (left) and Möbius strip (right)

6.7.3 Boundary of a surface; closed surfaces

The other important object concerning a surface is the boundary, mentioned above. This is a rather delicate notion if one wishes to discuss it in full generality, but we shall present the matter in an elementary way.

Let $\Sigma \subset \mathbb{R}^3$ be a regular and simple surface, which we assume to be a closed subset $\Sigma = \overline{\Sigma}$ of $\mathbb{R}^3$ (in the sense of Definition 4.5). Then let $\boldsymbol{\sigma} : \mathcal{R} \subset \mathbb{R}^2 \to \Sigma$ be a parametrisation over the closed region $\mathcal{R}$. Calling $A = \overset{\circ}{\mathcal{R}}$ the interior of $\mathcal{R}$, the image $\Sigma^\circ_{\boldsymbol{\sigma}} = \boldsymbol{\sigma}(A)$ is a subset of $\Sigma = \boldsymbol{\sigma}(\mathcal{R})$. The difference set

$$\partial\Sigma_{\boldsymbol{\sigma}} = \Sigma \setminus \Sigma^\circ_{\boldsymbol{\sigma}}$$

will be called **boundary of Σ (relative to $\boldsymbol{\sigma}$)**. Subsequent examples will show that $\partial\Sigma_{\boldsymbol{\sigma}}$ contains all points that are obvious boundary points of Σ in purely geometrical terms, but might also contain points that depend on the chosen parametrisation; this justifies the subscript $\boldsymbol{\sigma}$. It is thus logical to define the **boundary of Σ** (viewed as an intrinsic object to the surface) as the set

$$\boxed{\partial\Sigma = \bigcap_{\boldsymbol{\sigma}} \partial\Sigma_{\boldsymbol{\sigma}}\,,}$$

where the intersection is taken over all possible parametrisations (non-regular as well) of Σ. We point out that the symbol $\partial\Sigma$ denoted in Sect. 4.3 the frontier of Σ, seen as subset of $\mathbb{R}^3$; for regular surfaces this set coincides with $\overline{\Sigma} = \Sigma$, since no interior points are present. Our treatise of surfaces will only involve the notion of boundary, not of frontier. Here are the promised examples.

Examples 6.40

i) The upper hemisphere with centre $(0, 0, 1)$ and radius 1

$$\Sigma = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + (z-1)^2 = 1, z \ge 1\}$$

is a closed set $\Sigma = \overline{\Sigma}$ inside $\mathbb{R}^3$. Geometrically, it is quite intuitive (Fig. 6.17, left) that its boundary is the circle

$$\partial\Sigma = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + (z-1)^2 = 1, z = 1\}\,.$$

If one parametrises Σ by

$$\boldsymbol{\sigma} : \mathcal{R} = \{(u, v) \in \mathbb{R}^2 : u^2+v^2 \le 1\} \to \mathbb{R}^3\,, \quad \boldsymbol{\sigma}(u, v) = u\boldsymbol{i}+v\boldsymbol{j}+(1+\sqrt{u^2+v^2})\boldsymbol{k}\,,$$

it follows that

$$\Sigma^\circ_{\boldsymbol{\sigma}} = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + (z-1)^2 = 1, z > 1\}\,,$$

so $\Sigma \setminus \Sigma^\circ_{\boldsymbol{\sigma}}$ is precisely the boundary we claimed. Parametrising Σ with spherical coordinates, instead,

$$\widetilde{\boldsymbol{\sigma}} : \widetilde{\mathcal{R}} = [0, \frac{\pi}{2}] \times [0, 2\pi] \to \mathbb{R}^3\,, \quad \widetilde{\boldsymbol{\sigma}}(u, v) = \sin u \cos v\, \boldsymbol{i} + \sin u \sin v\, \boldsymbol{j} + (1+\cos u)\boldsymbol{k}\,,$$

so $\widetilde{A} = (0, \frac{\pi}{2}) \times (0, 2\pi)$, we obtain $\Sigma^\circ_{\widetilde{\boldsymbol{\sigma}}}$ as the portion of surface lying above the plane $z = 1$, without the equatorial arc joining the North pole $(0, 0, 2)$ to $(1, 0, 1)$ on the plane xz (Fig. 6.17, right); analytically,

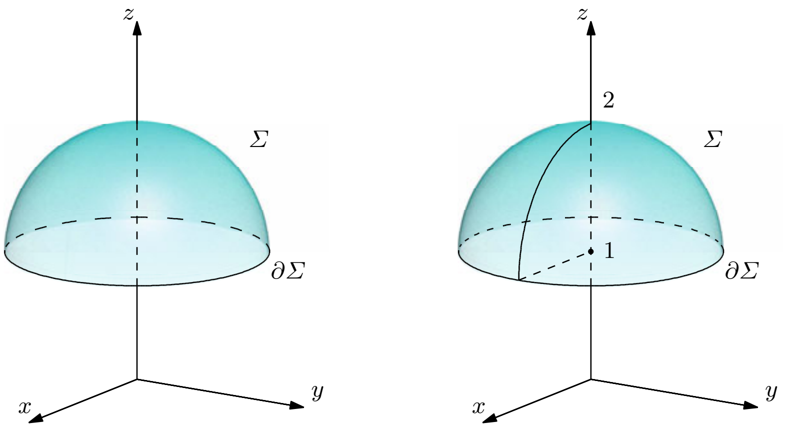


Figure 6.17. The hemisphere of Example 6.40 i)

$$\Sigma^{\circ}_{\widetilde{\boldsymbol{\sigma}}} = \{(x,y,z) \in \mathbb{R}^3 : x^2+y^2+(z-1)^2 = 1, z > 1\}$$
$$\setminus \{(x,y,z) \in \mathbb{R}^3 : x > 0,\ y = 0,\ x^2+(z-1)^2 = 1\}\,.$$

Besides the points on the geometric boundary, $\Sigma^{\circ}_{\widetilde{\boldsymbol{\sigma}}}$ consists of the aforementioned arc, which is image of part of the boundary of $\widetilde{\mathcal{R}}$; these points depend on the particular parametrisation we have chosen.

ii) Let

$$\Sigma = \{(x,y,z) \in \mathbb{R}^3 : x^2+y^2 = 1, |z| \le 1\}$$

be the (closed) cylinder of Fig. 6.18, whose boundary $\partial\Sigma$ consists of the circles $x^2+y^2 = 1, z = \pm 1$. Given $u_0 \in [0, 2\pi]$, we may parametrise Σ by

$$\boldsymbol{\sigma}_{u_0} : [0,2\pi] \times [-1,1] \to \mathbb{R}^3\,, \qquad \boldsymbol{\sigma}_{u_0}(u,v) = \cos(u-u_0)\,\boldsymbol{i} + \sin(u-u_0)\,\boldsymbol{j} + v\boldsymbol{k}\,,$$

generalising the parametrisation $\boldsymbol{\sigma} = \boldsymbol{\sigma}_0$ seen in (6.51). Then

$$\Sigma^{\circ}_{\boldsymbol{\sigma}_{u_0}} = \{(x,y,z) \in \mathbb{R}^3 : x^2+y^2 = 1, |z| < 1\}$$
$$\setminus \{(\cos u_0, \sin u_0, z) \in \mathbb{R}^3 : |z| < 1\}\,,$$

and $\partial\Sigma_{\boldsymbol{\sigma}_{u_0}}$ contains, apart from the two circles giving $\partial\Sigma$, also the vertical segment $\{(\cos u_0, \sin u_0, z) \in \mathbb{R}^3 : |z| < 1\}$. Intersecting all sets $\partial\Sigma_{\boldsymbol{\sigma}_{u_0}}$ – any two of them is enough – gives the geometric boundary of Σ.

iii) Consider at last the surface

$$\Sigma = \big\{(x,y,z) \in \mathbb{R}^3 : (x,y) \in \Gamma,\ 0 \le z \le 1\big\}\,,$$

where Γ is the Jordan arc in the plane xy defined by

$$\boldsymbol{\gamma} : [0,1] \to \mathbb{R}^2\,, \qquad \gamma_1(t) = \varphi(t) = t(1-t)^2 \quad \text{and} \quad \gamma_2(t) = \psi(t) = (1-t)t^2$$

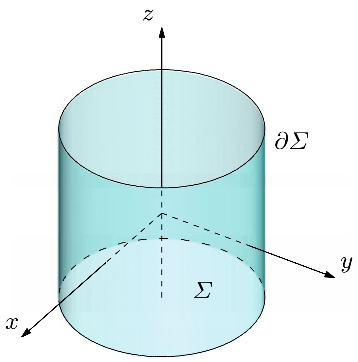


Figure 6.18. The cylinder relative to Example 6.40 ii)

(see Fig. 6.19, left). Let us first parametrise Σ with $\boldsymbol{\sigma} : \mathcal{R} \to \mathbb{R}^3$, where $\mathcal{R} = [0,1]^2$ and $\boldsymbol{\sigma}(u,v) = \big(\varphi(u), \psi(u), v\big)$. It is easy to convince ourselves that the boundary $\partial\Sigma_{\boldsymbol{\sigma}}$ of $\boldsymbol{\sigma}$ is $\partial\Sigma_{\boldsymbol{\sigma}} = \Gamma_0 \cup \Gamma_1 \cup S_0$, where $\Gamma_0 = \Gamma \times \{0\}$, $\Gamma_1 = \Gamma \times \{1\}$ and $S_0 = \{(0,0)\} \times [0,1]$ (Fig. 6.19, right). But if we extend the maps φ and ψ periodically to the intervals $[k, k+1]$, we may consider, for any $u_0 \in \mathbb{R}$, the parametrisations $\boldsymbol{\sigma}_{u_0} : \mathcal{R} \to \mathbb{R}^3$ given by $\boldsymbol{\sigma}_{u_0}(u,v) = \big(\varphi(u-u_0), \psi(u-u_0), z\big)$, for which $\Sigma^{\circ}_{\boldsymbol{\sigma}_{u_0}} = \Gamma_0 \cup \Gamma_1 \cup S_{u_0}$, $S_{u_0} = \{\boldsymbol{\gamma}(u_0)\} \times [0,1]$. We conclude that the vertical segment S_{u_0} does depend upon the parametrisation, whereas the top and bottom loops Γ_0, Γ_1 are common to all parametrisations. In summary, the boundary of Σ is the union of the loops, $\partial\Sigma = \Gamma_0 \cup \Gamma_1$. □

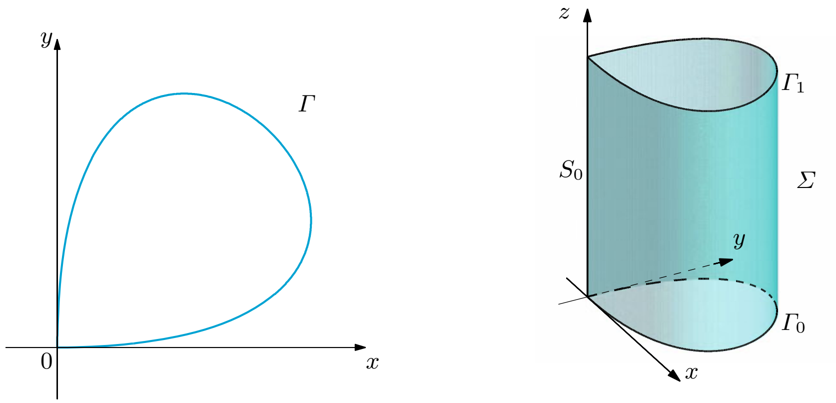


Figure 6.19. The arc Γ and surface Σ of Example 6.40 iii)

Among surfaces a relevant role is played by those that are closed. A regular, simple surface $\Sigma \subset \mathbb{R}^3$ is **closed** if it is bounded (as a subset of $\mathbb{R}^3$) and has no boundary ($\partial\Sigma = \emptyset$). This notion, too, is heavily dependent on the surface's geometry, and differs from the topological closure of Definition 4.5 (namely, each closed surface is necessarily a closed subset of $\mathbb{R}^3$, but there are topologically-closed surfaces with non-empty boundary). Examples of closed surfaces include the sphere and the torus, which we introduce now.

Examples 6.41

i) Arguing as in Example 6.40 i), we see that the parametrisation of the unit sphere $\Sigma = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = 1\}$ by spherical coordinates (Example 4.37 iii)) has a boundary, the semi-circle from the North pole to the South pole lying in the half-plane $x \geq 0$, $y = 0$. Any rotation of the coordinate system will produce another parametrisation whose boundary is a semi-circle joining antipodal points. It is straightforward to see that the boundaries' intersection is empty.

ii) A **torus** is the surface Σ built by identifying the top and bottom boundaries of the cylinder of Example 6.40 ii). It can also be obtained by a 2π-revolution around the z-axis of a circle of radius r lying on a plane containing the axis, having centre at a point of the plane xy at a distance $R + r$, $R \geq 0$, from the axis (see Fig. 6.20). A parametrisation is $\boldsymbol{\sigma} : [0, 2\pi] \times [0, 2\pi] \to \mathbb{R}^3$ with

$$\boldsymbol{\sigma}(u, v) = (R + r\cos u)\cos v\,\boldsymbol{i} + (R + r\cos u)\sin v\,\boldsymbol{j} + r\sin u\,\boldsymbol{k}\,.$$

The boundary $\partial\Sigma_{\boldsymbol{\sigma}}$ is the union of the circle on $z = 0$ with centre the origin and radius $R + r$, and the circle on $y = 0$ centred at $(R, 0, 0)$ with radius r.
Here, as well, changing the parametrisation shows the geometric boundary $\partial\Sigma$ is empty, making the torus a closed surface. □

The next theorem is the closest kin to the Jordan Curve Theorem 4.33 for plane curves.

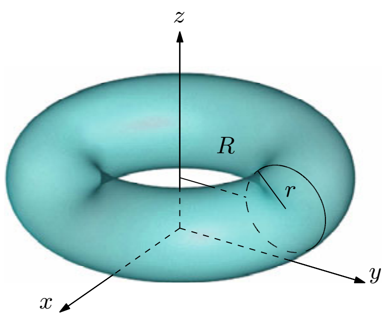


Figure 6.20. The torus of Example 6.41 ii)

Theorem 6.42 *A closed orientable surface Σ divides the space $\mathbb{R}^3$ in two open regions A_i, A_e whose common frontier is Σ. The region A_i (the* **interior***) is bounded, while the other region A_e (the* **exterior***) is unbounded.*

Nonetheless, there do exist closed non-orientable surfaces, which do not separate space in an inside and an outside. A traditional example is the **Klein bottle**, built from the Möbius strip in the same fashion the torus is constructed by glueing the cylinder's two rims.

6.7.4 Piecewise-regular surfaces

This generalisation of regularity allows the surface to contain curves where differentiability fails, and grants us a means of dealing with surfaces delimiting polyhedra such as cubes, pyramids, and truncated pyramids, or solids of revolution like truncated cones or cylinders. As in Sect. refsec:bordo, we shall assume all surfaces are *closed* subsets of $\mathbb{R}^3$.

Definition 6.43 *A subset $\Sigma \subset \mathbb{R}^3$ is a* **piecewise-regular, simple** *surface if it is the union of finitely many regular, simple surfaces $\Sigma_1, \ldots, \Sigma_K$ whose pairwise intersections $\Sigma_h \cap \Sigma_k$ (if not empty or a point) are piecewise-regular curves Γ_{hk} contained in the boundary of both Σ_h and Σ_k.*
The analogue definition holds for piecewise-regular compact surfaces.

Each surface Σ_k is called a **face** (or **component**) of Σ, and the intersection curve between any two faces is an **edge** of Σ .

A piecewise regular simple surface Σ is **orientable** if every component is orientable in such a way that adjacent components have compatible orientations. Proposition 6.39 then extends to (compact) piecewise-regular, simple surfaces.

The **boundary** $\partial\Sigma$ of a piecewise-regular simple surface Σ is the closure of the set of points belonging to the boundary of one, and one only, component. Equivalently, consider the union of the single components' boundaries, without the points common to two or more components, and then take the closure of this set: this is $\partial\Sigma$. A piecewise-regular simple surface is **closed** if bounded and with empty boundary.

Examples 6.44

i) The frontier of a cube is piecewise regular, and has the cube's six faces as components; it is obviously orientable and closed (Fig. 6.21, left).

ii) The piecewise-regular surface obtained from the previous one by taking away two opposite faces (Fig. 6.21, right) is still orientable, but no longer closed. Its boundary is the union of the boundaries of the faces removed. □

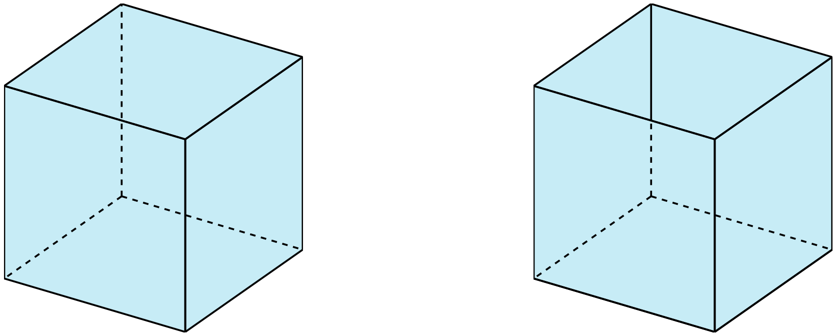


Figure 6.21. The cubes of Examples 6.44

Remark 6.45 Example ii) provides us with the excuse for explaining why one should insist on the closure of the boundary, instead of taking the mere boundary. Consider a cube aligned with the Cartesian axes, from which we have removed the top and bottom faces. The union of the boundaries of the 4 faces left is given by the 12 edges forming the skeleton. Getting rid of the points common to two faces leaves us with the 4 horizontal top edges and the 4 bottom ones, without the 8 vertices. Closing what is left recovers the 8 missing vertices, so the boundary is precisely the union of the 8 horizontal edges. □

6.8 Exercises

1. *Find the Jacobian matrix of the following functions:*

   a) $\boldsymbol{f}(x,y) = \mathrm{e}^{2x+y}\,\boldsymbol{i} + \cos(x+2y)\,\boldsymbol{j}$

   b) $\boldsymbol{f}(x,y,z) = (x+2y^2+3z^3)\,\boldsymbol{i} + (x+\sin 3y+\mathrm{e}^z)\,\boldsymbol{j}$

2. *Determine the divergence of the following vector fields:*

   a) $\boldsymbol{f}(x,y) = \cos(x+2y)\,\boldsymbol{i} + \mathrm{e}^{2x+y}\,\boldsymbol{j}$

   b) $\boldsymbol{f}(x,y,z) = (x+y+z)\,\boldsymbol{i} + (x^2+y^2+z^2)\,\boldsymbol{j} + (x^3+y^3+z^3)\,\boldsymbol{k}$

3. *Compute the curl of the vector fields:*

   a) $\boldsymbol{f}(x,y,z) = x\,\boldsymbol{i} + y\,\boldsymbol{j} + z\,\boldsymbol{k}$

   b) $\boldsymbol{f}(x,y,z) = xyz\,\boldsymbol{i} + z\sin y\,\boldsymbol{j} + x\mathrm{e}^y\,\boldsymbol{k}$

   c) $\boldsymbol{f}(x,y) = \left(\dfrac{y}{\sqrt{x^2+y^2}}\right)\boldsymbol{i} + \left(\dfrac{x}{\sqrt{x^2+y^2}}\right)\boldsymbol{j}$

   d) $\boldsymbol{f}(x,y) = \mathbf{grad}\,(\log_x y)^2$

4. *Given* $f(x,y) = 3x + 2y$ *and* $\boldsymbol{g}(u,v) = (u+v)\,\boldsymbol{i} + uv\,\boldsymbol{j}$*, compute the composite map* $f \circ \boldsymbol{g}$ *explicitly and find its gradient.*

5. *Compute the composite of the following pairs of maps and the composite's gradient:*

   a) $f(s,t) = \sqrt{s+t}\,, \qquad \boldsymbol{g}(x,y) = xy\,\boldsymbol{i} + \dfrac{x}{y}\,\boldsymbol{j}$

   b) $f(x,y,z) = xyz\,, \qquad \boldsymbol{g}(r,s,t) = (r+s)\,\boldsymbol{i} + (r+3t)\,\boldsymbol{j} + (s-t)\,\boldsymbol{k}$

6. *Let* $\boldsymbol{f} : \mathbb{R}^2 \to \mathbb{R}^2$ *and* $\boldsymbol{g} : \mathbb{R}^3 \to \mathbb{R}^2$ *be defined by*

$$\boldsymbol{f}(x,y) = \sin(2x+y)\,\boldsymbol{i} + \mathrm{e}^{x+2y}\,\boldsymbol{j}\,, \quad \boldsymbol{g}(u,v,z) = (u + 2v^2 + 3z^3)\,\boldsymbol{i} + (u^2 - 2v)\,\boldsymbol{j}\,.$$

   a) *Compute their Jacobians* $J\boldsymbol{f}$ *and* $J\boldsymbol{g}$.

   b) *Determine the composite* $\boldsymbol{h} = \boldsymbol{f} \circ \boldsymbol{g}$ *and its Jacobian* $J\boldsymbol{h}$ *at the point* $(1,-1,0)$.

7. *Given the following maps, determine the first derivative and the monotonicity on the interval indicated:*

   a) $f(x) = \displaystyle\int_0^x \frac{\arctan x^2 y}{y}\,\mathrm{d}y\,, \qquad t \in [1, +\infty)$

   b) $f(x) = \displaystyle\int_0^{\sqrt{1-x}} y\,\sqrt[3]{8 + y^4 - \frac{x}{2}}\,\mathrm{d}y\,, \qquad t \in [0,1]$

8. *Compute the length of the arcs:*

   a) $\gamma(t) = (t, 3t^2)\,, \qquad t \in [0,1]$

   b) $\gamma(t) = (t\cos t, t\sin t, t)\,, \qquad t \in [0, 2\pi]$

   c) $\gamma(t) = (t^2, t^2, t^3)\,, \qquad t \in [0,1]$

9. *Determine the values of* $\alpha \in \mathbb{R}$ *for which the length of* $\gamma(t) = (t, \alpha t^2, t^3)$*,* $t \in [0,T]$*, equals* $\ell(\gamma) = T + T^3$.

10. *Consider the closed arc* γ *whose trace is the union of the segment between* $A = (-\log 2, 1/2)$ *and* $B = (1,0)$*, the circular arc* $x^2 + y^2 = 1$ *joining* B *to* $C = (0,1)$*, and the arc* $\gamma_3(t) = (t, \mathrm{e}^t)$ *from* C *to* A*. Compute its length.*

11. *Tell whether the following arcs are closed, regular or piecewise regular:*

   a) $\gamma(t) = \big(t(t-\pi)^2(t-2\pi t), \cos t\big)\,, \qquad t \in [0, 2\pi]$

   b) $\gamma(t) = (\sin^2 2\pi t, t)\,, \qquad t \in [0,1]$

12. *What are the values of the parameter $\alpha \in \mathbb{R}$ for which the curve*

$$\gamma(t) = \begin{cases} (\alpha t, t^3, 0) & \text{if } t < 0 \\ (\sin \alpha^2 t, 0, t^3) & \text{if } t \geq 0 \end{cases},$$

is regular?

13. *Calculate the principal normal vector $\boldsymbol{n}(s)$, the binormal $\boldsymbol{b}(s)$ and the torsion $\boldsymbol{b}'(s)$ for the circular helix $\gamma(t) = (\cos t, \sin t, t)$, $t \in \mathbb{R}$.*

14. *If a curve $\gamma(t) = x(t)\boldsymbol{i} + y(t)\boldsymbol{j}$ is $(r(t), \theta(t))$ in polar coordinates, verify that*

$$\|\gamma'(t)\|^2 = \Big(\frac{\mathrm{d}r}{\mathrm{d}t}\Big)^2 + \Big(r(t)\frac{\mathrm{d}\theta}{\mathrm{d}t}\Big)^2 .$$

15. *Check that if a curve $\gamma(t) = x(t)\boldsymbol{i} + y(t)\boldsymbol{j} + z(t)\boldsymbol{k}$ is given by $\big(r(t), \theta(t), z(t)\big)$ in cylindrical coordinates, then*

$$\|\gamma'(t)\|^2 = \Big(\frac{\mathrm{d}r}{\mathrm{d}t}\Big)^2 + \Big(r(t)\frac{\mathrm{d}\theta}{\mathrm{d}t}\Big)^2 + \Big(\frac{\mathrm{d}z}{\mathrm{d}t}\Big)^2 .$$

16. *Check that the curve $\gamma(t) = x(t)\boldsymbol{i}+y(t)\boldsymbol{j}+z(t)\boldsymbol{k}$, given in spherical coordinates by $\big(r(t), \varphi(t), \theta(t)\big)$, satisfies*

$$\|\gamma'(t)\|^2 = \Big(\frac{\mathrm{d}r}{\mathrm{d}t}\Big)^2 + \big(r(t)\big)^2\Big(\frac{\mathrm{d}\varphi}{\mathrm{d}t}\Big)^2 + \big(r(t)\sin\varphi(t)\big)^2\Big(\frac{\mathrm{d}\theta}{\mathrm{d}t}\Big)^2 .$$

17. *Using Exercise 15, compute the length of the arc $\gamma(t)$ given in cylindrical coordinates by*

$$\gamma(t) = \big(r(t), \theta(t), z(t)\big) = \Big(\sqrt{2}\cos t, \frac{1}{\sqrt{2}}t, \sin t\Big), \qquad t \in \big[0, \frac{\pi}{2}\big].$$

18. *Consider the parametric surface*

$$\boldsymbol{\sigma}(u, v) = uv\,\boldsymbol{i} + (1 + 3u)\,\boldsymbol{j} + (v^3 + 2u)\,\boldsymbol{k}.$$

a) *Tell whether it is simple.*
b) *Determine the set $\mathcal{R}$ on which $\boldsymbol{\sigma}$ is regular.*
c) *Determine the normal to the surface at every point of $\mathcal{R}$.*
d) *Write the equation of the plane tangent to the surface at $P_0 = \boldsymbol{\sigma}(u_0, v_0) = (1, 4, 3)$.*

19. *The surfaces*

$$\boldsymbol{\sigma}_1(u,v) = \cos(2-u)\,\boldsymbol{i} + \sin(2-u)\,\boldsymbol{j} + v^2\,\boldsymbol{k}\,, \qquad (u,v) \in [0,2\pi]\times[0,1]$$

and

$$\boldsymbol{\sigma}_2(u,v) = \sin(3+2u)\,\boldsymbol{i} + \cos(3+2u)\,\boldsymbol{j} + (1-v)\,\boldsymbol{k}\,, \qquad (u,v) \in [0,\pi]\times[0,1]$$

parametrise the same set Σ in $\mathbb{R}^3$.

a) Determine Σ.

b) Say whether the orientations defined by $\boldsymbol{\sigma}_1$ and $\boldsymbol{\sigma}_2$ coincide or not.

c) Compute the unit normals to Σ, at $P_0 = (0,1,\frac{1}{4})$, relative to $\boldsymbol{\sigma}_1$ and $\boldsymbol{\sigma}_2$.

20. *Consider the surface $\boldsymbol{\sigma} : \mathbb{R}^2 \to \mathbb{R}^3$,*

$$\boldsymbol{\sigma}(u,v) = u\,\boldsymbol{i} + v\,\boldsymbol{j} + (u^2 + 3uv + v^2)\,\boldsymbol{k}\,.$$

a) What is the unit normal $\boldsymbol{n}(u,v)$?

b) Determine the points on the image Σ at which the normal is orthogonal to the plane $8x + 7y - 2z = 4$.

6.8.1 Solutions

1. *Jacobian matrices:*

a) $J\boldsymbol{f}(x,y) = \begin{pmatrix} 2e^{2x+y} & e^{2x+y} \\ -\sin(x+2y) & -2\sin(x+2y) \end{pmatrix}$

b) $J\boldsymbol{f}(x,y,z) = \begin{pmatrix} 1 & 4y & 9z^2 \\ 1 & 3\cos 3y & e^z \end{pmatrix}$

2. *Vector fields' divergence:*

a) $\operatorname{div} \boldsymbol{f}(x,y) = -\sin(x+2y) + e^{2x+y}$

b) $\operatorname{div} \boldsymbol{f}(x,y,z) = 1 + 2y + 3z^2$

3. *Curl of vector fields:*

a) $\mathbf{curl}\, \boldsymbol{f}(x,y,z) = \mathbf{0}$

b) $\mathbf{curl}\, \boldsymbol{f}(x,y,z) = (xe^y - \sin y)\,\boldsymbol{i} - (e^y - xy)\,\boldsymbol{j} - xz\,\boldsymbol{k}$

c) $\mathbf{curl}\, \boldsymbol{f}(x,y) = \dfrac{2xy}{(x^2+y^2)^{3/2}}$

d) Since $\boldsymbol{f}$ is a gradient, Proposition 6.8 ensures its curl is null.

4. We have

$$f \circ \boldsymbol{g}(u,v) = f\big(\boldsymbol{g}(u,v)\big) = f(u+v, uv) = 3(u+v) + 2uv$$

and

$$\nabla f \circ \boldsymbol{g}(u,v) = (3+2v, 3+2u)\,.$$

5. *Composite maps and gradients:*

a) We have

$$f \circ \boldsymbol{g}(x,y) = f\big(\boldsymbol{g}(x,y)\big) = f\big(xy, \frac{x}{y}\big) = \sqrt{xy + \frac{x}{y}}$$

and

$$\nabla f \circ \boldsymbol{g}(x,y) = \left(\frac{1}{2}\sqrt{\frac{y}{xy^2+x}}\,\Big(y + \frac{1}{y}\Big), \frac{1}{2}\sqrt{\frac{y}{xy^2+x}}\,\Big(x - \frac{x}{y^2}\Big)\right)\,.$$

b) Since

$$f\big(\boldsymbol{g}(r,s,t)\big) = f(r+s, r+3t, s-t) = (r+s)\,(r+3t)\,(s-t)\,,$$

setting $h = f \circ \boldsymbol{g}$ gives the gradient's components

$$\begin{aligned}
\frac{\partial h}{\partial r}(r,s,t) &= (r+3t)\,(s-t) + (r+s)\,(s-t) = 2rs + 2st - 2rt + s^2 - 3t^2 \\
\frac{\partial h}{\partial s}(r,s,t) &= (r+3t)\,(s-t) + (r+s)\,(r+3t) = 2rs + 6st + 2rt - 3t^2 + r^2 \\
\frac{\partial h}{\partial t}(r,s,t) &= 3(r+s)\,(s-t) - (r+s)\,(r+3t) = 2rs - 6rt - 6st + 3s^2 - r^2\,.
\end{aligned}$$

6. a) We have

$$\boldsymbol{Jf}(x,y) = \begin{pmatrix} 2\cos(2x+y) & \cos(2x+y) \\ \mathrm{e}^{x+2y} & 2\mathrm{e}^{x+2y} \end{pmatrix}, \qquad \boldsymbol{Jg}(u,v,z) = \begin{pmatrix} 1 & 4v & 9z^2 \\ 2u & -2 & 0 \end{pmatrix}.$$

b) Since

$$\begin{aligned}
\boldsymbol{h}(u,v,z) &= \boldsymbol{f}(u + 2v^2 + 3z^3, u^2 - 2v) \\
&= \sin(u^2 + 4v^2 + 6z^3 + 2u - 2v)\,\boldsymbol{i} + \mathrm{e}^{2u^2+2v^2+3z^3+u-4v}\,\boldsymbol{j}
\end{aligned}$$

and $\boldsymbol{g}(1,-1,0) = (3,3)$, it follows that

$$\begin{aligned}
\boldsymbol{Jh}(1,-1,0) &= \boldsymbol{Jf}\big(\boldsymbol{g}(1,-1,0)\big)\,\boldsymbol{Jg}(1,-1,0) = \boldsymbol{Jf}(3,3)\,\boldsymbol{Jg}(1,-1,0) \\
&= \begin{pmatrix} 2\cos 9 & \cos 9 \\ \mathrm{e}^9 & 2\mathrm{e}^9 \end{pmatrix} \begin{pmatrix} 1 & -4 & 0 \\ 2 & -2 & 0 \end{pmatrix} = \begin{pmatrix} 4\cos 9 & -10\cos 9 & 0 \\ 5\mathrm{e}^9 & -8\mathrm{e}^9 & 0 \end{pmatrix}.
\end{aligned}$$

7. *Derivatives of integral functions:*

a) As
$$f'(x) = \frac{\arctan x^3}{x} > 0\,, \qquad \forall x \in \mathbb{R}\,,$$
the map is (monotone) increasing on $[1, +\infty)$.

b) Since
$$\begin{aligned} f'(x) &= -\frac{1}{6}\int_0^{\sqrt{1-x}} y\left(8+y^4-\frac{x}{2}\right)^{-3/2} \mathrm{d}y - \frac{1}{2}\sqrt[3]{8+(1-x)^2-\frac{x}{2}} \\ &= -\frac{1}{2}\left(\frac{1}{3}\int_0^{\sqrt{1-x}} y\left(8+y^4-\frac{x}{2}\right)^{-3/2} \mathrm{d}y + \sqrt[3]{x^2-\frac{5}{2}x+9}\right), \end{aligned}$$
we obtain $f'(x) \le 0$ for any $x \in [0,1]$; hence f is decreasing on $[0,1]$.

8. *Length of arcs:*

We shall make use of the following indefinite integral (Vol. I, Example 9.13 v)):
$$\int \sqrt{1+x^2}\,\mathrm{d}x = \frac{1}{2}x\sqrt{1+x^2} + \frac{1}{2}\log(\sqrt{1+x^2}+x) + c\,.$$

a) From
$$\gamma'(t) = (1, 6t) \qquad \text{and} \qquad \|\gamma'(t)\| = \sqrt{1+36t^2}$$
follows
$$\begin{aligned} \ell(\gamma) &= \int_0^1 \sqrt{1+36t^2}\,\mathrm{d}t = \frac{1}{6}\int_0^6 \sqrt{1+x^2}\,\mathrm{d}x \\ &= \frac{1}{6}\left[\frac{1}{2}x\sqrt{1+x^2} + \frac{1}{2}\log(\sqrt{1+x^2}+x)\right]_0^6 \\ &= \frac{1}{2}\sqrt{37} + \frac{1}{12}\log(\sqrt{37}+6)\,. \end{aligned}$$

b) Since
$$\gamma'(t) = (\cos t - t\sin t, \sin t + t\cos t, 1) \qquad \text{and} \qquad \|\gamma'(t)\| = \sqrt{2+t^2}\,,$$
we have
$$\begin{aligned} \ell(\gamma) &= \int_0^{2\pi} \sqrt{2+t^2}\,\mathrm{d}t = 2\int_0^{\sqrt{2}\pi} \sqrt{1+x^2}\,\mathrm{d}x \\ &= 2\left[\frac{1}{2}x\sqrt{1+x^2} + \frac{1}{2}\log(\sqrt{1+x^2}+x)\right]_0^{\sqrt{2}\pi} \\ &= \sqrt{2}\pi\sqrt{1+2\pi^2} + \log\left(\sqrt{1+2\pi^2}+\sqrt{2}\pi\right). \end{aligned}$$

c) $\ell(\boldsymbol{\gamma}) = \frac{1}{27}\big(17\sqrt{17} - 16\sqrt{2}\big)$.

9. As $\boldsymbol{\gamma}'(t) = (1, 2\alpha t, 3t^2)$, we must have

$$\ell(\boldsymbol{\gamma}) = T + T^3 = \int_0^T \sqrt{1 + 4\alpha^2 t^2 + 9t^4}\, \mathrm{d}t\,.$$

Set $g(T) = T + T^3$: the Fundamental Theorem of Integral Calculus yields

$$g'(T) = 1 + 3T^2 = \sqrt{1 + 4\alpha^2 T^2 + 9T^4}\,,$$

i.e.,

$$(1 + 3T^2)^2 = 1 + 4\alpha^2 T^2 + 9T^4\,.$$

From this, $4\alpha^2 = 6$, so $\alpha = \pm\sqrt{3/2}$.

10. We have

$$\ell(\boldsymbol{\gamma}) = \ell(\boldsymbol{\gamma}_1) + \ell(\boldsymbol{\gamma}_2) + \ell(\boldsymbol{\gamma}_3)$$

where

$$\begin{aligned}
\boldsymbol{\gamma}_1(t) &= \Big(t, \frac{1-t}{2(1+\log 2)}\Big)\,, \qquad t \in [-\log 2, 1]\,,\\
\boldsymbol{\gamma}_2(t) &= (\cos t, \sin t)\,, \qquad t \in \big[0, \frac{\pi}{2}\big]\\
\boldsymbol{\gamma}_3(t) &= (t, \mathrm{e}^t)\,, \qquad t \in [-\log 2, 0]\,.
\end{aligned}$$

From elementary geometry we know that

$$\begin{aligned}
\ell(\boldsymbol{\gamma}_1) &= d(A, B) = \sqrt{(1+\log 2)^2 + \frac{1}{4}} = \sqrt{\frac{5}{4} + 2\log 2 + \log^2 2}\,,\\
\ell(\boldsymbol{\gamma}_2) &= \frac{2\pi}{4} = \frac{\pi}{2}\,.
\end{aligned}$$

For $\ell(\boldsymbol{\gamma}_3)$, observe that $\boldsymbol{\gamma}_3'(t) = (1, \mathrm{e}^t)$, so

$$\ell(\boldsymbol{\gamma}_3) = \int_{-\log 2}^0 \sqrt{1 + \mathrm{e}^{2t}}\, \mathrm{d}t\,.$$

Now, setting $u = \sqrt{1 + \mathrm{e}^{2t}}$, we obtain $\mathrm{e}^{2t} = u^2 - 1$ and $\mathrm{d}u = \frac{\mathrm{e}^{2t}}{\sqrt{1+\mathrm{e}^{2t}}}\, \mathrm{d}t = \frac{u^2-1}{u}\, \mathrm{d}t$. Therefore

$$\begin{aligned}
\ell(\boldsymbol{\gamma}_3) &= \int_{\sqrt{5}/2}^{\sqrt{2}} \frac{u^2}{u^2-1}\, \mathrm{d}u = \int_{\sqrt{5}/2}^{\sqrt{2}} \Big(1 + \frac{1/2}{u-1} - \frac{1/2}{u+1}\Big)\, \mathrm{d}u\\
&= \Big[u + \frac{1}{2}\log\Big|\frac{u-1}{u+1}\Big|\Big]_{\sqrt{5}/2}^{\sqrt{2}} = \sqrt{2} - \frac{\sqrt{5}}{2} + \log(\sqrt{2}-1)(\sqrt{5}+2)\,.
\end{aligned}$$

To sum up,

$$\ell(\gamma) = \frac{\pi}{2} + \sqrt{\frac{5}{4} + 2\log 2 + \log^2 2} + \sqrt{2} - \frac{\sqrt{5}}{2} + \log(\sqrt{2}-1)(\sqrt{5}+2)\,.$$

11. *Closed, regular, piecewise-regular arcs:*

a) The arc is closed because $\gamma(0) = (0,1) = \gamma(2\pi)$.
Moreover, $\gamma'(t) = \big(2(t-\pi)(2t^2 - 4\pi t + \pi^2), -\sin t\big)$ implies that $\gamma'(t)$ is $\mathcal{C}^1$ on $[0,2\pi]$. Next, notice $\sin t = 0$ only when $t = 0, \pi, 2\pi$, and $\gamma'(0) = (-2\pi^3, 0) \neq \mathbf{0}$, $\gamma'(2\pi) = (-10\pi^3, 0) \neq \mathbf{0}$, $\gamma'(\pi) = \mathbf{0}$. The only non-regular point is thus the one corresponding to $t = \pi$. Altogether, the arc is piecewise regular.

b) The fact that $\gamma(0) = (0,0)$ and $\gamma(1) = (0,1)$ implies the arc is not closed. What is more,

$$\gamma'(t) = (4\pi \sin 2\pi t \cos 2\pi t, 1) \neq (0,0)\,, \qquad \forall t \in [0,1]\,,$$

making the arc regular.

12. Note γ is continuous for any $t \in \mathbb{R}$ with

$$\gamma'(t) = \begin{cases} (\alpha, 3t^2, 0) & \text{if } t < 0\,, \\ (\alpha^2 \cos \alpha^2 t, 0, 3t^2) & \text{if } t \geq 0\,. \end{cases}$$

Certainly $\gamma'(t) \neq \mathbf{0}$ for $t \neq 0$ and any $\alpha \in \mathbb{R}$. To study the point $t = 0$, observe that

$$\gamma'(0^-) = \lim_{t\to 0^-} \gamma'(t) = (\alpha, 0, 0)\,, \qquad \gamma'(0^+) = \lim_{t\to 0^+} \gamma'(t) = (\alpha^2, 0, 0)\,,$$

so $\gamma'(0)$ exists if $\alpha = \alpha^2$, i.e., if $\alpha = 0$ or $\alpha = 1$. If $\alpha = 0$ then $\gamma'(0) = \mathbf{0}$, while for $\alpha = 1$ we have $\gamma'(0) = (1,0,0) \neq \mathbf{0}$. In conclusion, the only value of regularity is $\alpha = 1$.

13. Bearing in mind Example 6.29, we may re-parametrise the helix by arc length:

$$\widetilde{\gamma}(s) = \Big(\cos\frac{\sqrt{2}}{2}s, \sin\frac{\sqrt{2}}{2}s, \frac{\sqrt{2}}{2}s\Big)\,, \quad s \in \mathbb{R}\,.$$

For any $s \in \mathbb{R}$ then,

$$\boldsymbol{t}(s) = \widetilde{\gamma}'(s) = \Big(-\frac{\sqrt{2}}{2}\sin\frac{\sqrt{2}}{2}s, \frac{\sqrt{2}}{2}\cos\frac{\sqrt{2}}{2}s, \frac{\sqrt{2}}{2}\Big)\,,$$

and

$$\boldsymbol{t}'(s) = \Big(-\frac{1}{2}\cos\frac{\sqrt{2}}{2}s, -\frac{1}{2}\sin\frac{\sqrt{2}}{2}s, 0\Big)$$

with curvature $\mathcal{K}(s) = \|\boldsymbol{t}'(s)\| = 1/2$. Therefore

$$\boldsymbol{n}(s) = \Big(-\cos\frac{\sqrt{2}}{2}s, -\sin\frac{\sqrt{2}}{2}s, 0\Big)$$

and

$$\boldsymbol{b}(s) = \boldsymbol{t}(s) \wedge \boldsymbol{n}(s) = \Big(\frac{\sqrt{2}}{2}\sin\frac{\sqrt{2}}{2}s, -\frac{\sqrt{2}}{2}\cos\frac{\sqrt{2}}{2}s, \frac{\sqrt{2}}{2}\Big).$$

At last,

$$\boldsymbol{b}'(t) = \Big(\frac{1}{2}\cos\frac{\sqrt{2}}{2}s, \frac{1}{2}\sin\frac{\sqrt{2}}{2}s, 0\Big)$$

and the scalar torsion is $\tau(s) = -\frac{1}{2}$.

14. Remembering that

$$x = r\cos\theta\,, \qquad y = r\sin\theta\,,$$

and using the chain rule,

$$x'(t) = \frac{\mathrm{d}x}{\mathrm{d}t} = \frac{\partial x}{\partial r}\frac{\mathrm{d}r}{\mathrm{d}t} + \frac{\partial x}{\partial \theta}\frac{\mathrm{d}\theta}{\mathrm{d}t} = \cos\theta\frac{\mathrm{d}r}{\mathrm{d}t} - r\sin\theta\frac{\mathrm{d}\theta}{\mathrm{d}t}$$

$$y'(t) = \frac{\mathrm{d}y}{\mathrm{d}t} = \frac{\partial y}{\partial r}\frac{\mathrm{d}r}{\mathrm{d}t} + \frac{\partial y}{\partial \theta}\frac{\mathrm{d}\theta}{\mathrm{d}t} = \sin\theta\frac{\mathrm{d}r}{\mathrm{d}t} + r\cos\theta\frac{\mathrm{d}\theta}{\mathrm{d}t}\,.$$

The required equation follows by using $\sin^2\theta + \cos^2\theta = 1$, for any $\theta \in \mathbb{R}$, and $\|\boldsymbol{\gamma}'(t)\|^2 = \big(x'(t)\big)^2 + \big(y'(t)\big)^2$.

16. Recalling

$$x = r\sin\varphi\cos\theta\,, \qquad y = r\sin\varphi\sin\theta\,, \qquad x = r\cos\varphi\,,$$

and the chain rule, we have

$$\begin{aligned} x'(t) &= \frac{\mathrm{d}x}{\mathrm{d}t} = \frac{\partial x}{\partial r}\frac{\mathrm{d}r}{\mathrm{d}t} + \frac{\partial x}{\partial \varphi}\frac{\mathrm{d}\varphi}{\mathrm{d}t} + \frac{\partial x}{\partial \theta}\frac{\mathrm{d}\theta}{\mathrm{d}t} \\ &= \sin\varphi\cos\theta\frac{\mathrm{d}r}{\mathrm{d}t} + r\cos\varphi\cos\theta\frac{\mathrm{d}\varphi}{\mathrm{d}t} - r\sin\varphi\sin\theta\frac{\mathrm{d}\theta}{\mathrm{d}t} \\ y'(t) &= \frac{\mathrm{d}y}{\mathrm{d}t} = \frac{\partial y}{\partial r}\frac{\mathrm{d}r}{\mathrm{d}t} + \frac{\partial y}{\partial \varphi}\frac{\mathrm{d}\varphi}{\mathrm{d}t} + \frac{\partial y}{\partial \theta}\frac{\mathrm{d}\theta}{\mathrm{d}t} \\ &= \sin\varphi\sin\theta\frac{\mathrm{d}r}{\mathrm{d}t} + r\cos\varphi\sin\theta\frac{\mathrm{d}\varphi}{\mathrm{d}t} + r\sin\varphi\cos\theta\frac{\mathrm{d}\theta}{\mathrm{d}t} \\ z'(t) &= \frac{\mathrm{d}z}{\mathrm{d}t} = \frac{\partial z}{\partial r}\frac{\mathrm{d}r}{\mathrm{d}t} + \frac{\partial z}{\partial \varphi}\frac{\mathrm{d}\varphi}{\mathrm{d}t} + \frac{\partial z}{\partial \theta}\frac{\mathrm{d}\theta}{\mathrm{d}t} \\ &= \cos\varphi\frac{\mathrm{d}r}{\mathrm{d}t} - r\sin\varphi\frac{\mathrm{d}\varphi}{\mathrm{d}t}\,. \end{aligned}$$

Now using $\sin^2\xi + \cos^2\xi = 1$, for any $\xi \in \mathbb{R}$, and $\|\boldsymbol{\gamma}'(t)\|^2 = \big(x'(t)\big)^2 + \big(y'(t)\big)^2 + \big(z'(t)\big)^2$, a little computation gives the result.

17. Since

$$r'(t) = -\sqrt{2}\sin t\,,\theta \qquad '(t) = \frac{1}{\sqrt{2}}\,, \qquad z'(t) = \cos t\,,$$

we have

$$\ell(\gamma) = \int_0^{\pi/2} \sqrt{2\sin^2 t + \big(2\cos^2 t\big)\,\frac{1}{2} + \cos^2 t}\;\mathrm{d}t = \sqrt{2}\int_0^{\pi/2} \mathrm{d}t = \frac{\sqrt{2}}{2}\pi\,.$$

Notice the arc's trace lies on the surface of the ellipsoid $x^2 + y^2 + 2z^2 = 2$. The initial point is $(\sqrt{2}, 0, 0)$, the end point $(0, 0, 1)$ (see Fig. 6.22).

18. a) The surface is simple if, for $(u_1, v_1), (u_2, v_2) \in \mathbb{R}^2$, we have

$$\boldsymbol{\sigma}(u_1, v_1) = \boldsymbol{\sigma}(u_2, v_2) \qquad \Rightarrow \qquad (u_1, v_1) = (u_2, v_2)\,.$$

The left equation means

$$\begin{cases} u_1 v_1 = u_2 v_2 \\ 1 + 3u_1 = 1 + 3u_2 \\ v_1^3 + 2u_1 = v_2^3 + 2u_2\,. \end{cases}$$

From the middle line we obtain $u_1 = u_2$, and then $v_1 = v_2$ by substitution.

b) Consider the Jacobian

$$\boldsymbol{J\sigma} = \begin{pmatrix} v & u \\ 3 & 0 \\ 2 & 3v^2 \end{pmatrix}\,.$$

The three minors' determinants are $-3u$, $9v^2$, $3v^3 - 2u$. The only point at which they all vanish is the origin. Thus $\boldsymbol{\sigma}$ is regular on $\mathcal{R} = \mathbb{R}^2 \setminus \{(0,0)\}$.

c) For $(u, v) \neq 0$,

$$\nu(u, v) = \det\begin{pmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ v & 3 & 2 \\ u & 0 & 3v^2 \end{pmatrix} = 9v^2\,\boldsymbol{i} + (2u - 3v^3)\,\boldsymbol{j} - 3u\,\boldsymbol{k}\,.$$

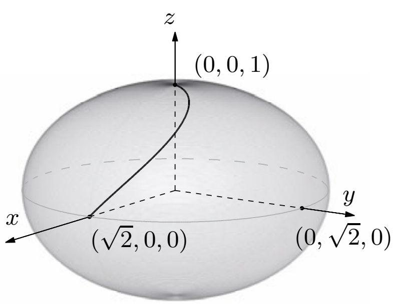


Figure 6.22. The arc of Exercise 17

d) Imposing

$$uv = 1\,, \qquad 1 + 3u = 4\,, \qquad v^2 + 2u = 3$$

shows the point $P_0 = (1, 4, 3)$ is the image under $\boldsymbol{\sigma}$ of $(u_0, v_0) = (1, 1)$. Therefore the tangent plane in question is

$$\begin{aligned} \Pi(u, v) &= \boldsymbol{\sigma}(1, 1) + \frac{\partial \boldsymbol{\sigma}}{\partial u}(1, 1)\,(u - 1) + \frac{\partial \boldsymbol{\sigma}}{\partial v}(1, 1)\,(v - 1) \\ &= (1, 4, 3) + (1, 3, 2)(u - 1) + (1, 0, 3)(v - 1) \\ &= (u + v - 1)\,\boldsymbol{i} + (3u + 1)\,\boldsymbol{j} + (2u + 3v - 2)\,\boldsymbol{k}\,. \end{aligned}$$

Using Cartesian coordinates $x = u + v - 1$, $y = 3u + 1$, $z = 2u + 3v - 2$, the plane reads $9x - y - 3z + 4 = 0$.

19. a) As u varies in $[0, 2\pi]$, the vector $\cos(2 - u)\,\boldsymbol{i} + \sin(2 - u)\,\boldsymbol{j}$ runs along the unit circle in the plane $z = 0$, while for v in $[0, 1]$, the vector $v^2\,\boldsymbol{k}$ describes the segment $[0, 1]$ on the z-axis. Consequently Σ is a cylinder of height 1 with axis on the z's.

b) Using formula (6.48) the normal vector defined by $\boldsymbol{\sigma}_1$ is

$$\boldsymbol{\nu}_1(u, v) = -2v\cos(2 - u)\,\boldsymbol{i} - 2v\sin(2 - u)\,\boldsymbol{j}\,,$$

while the normal of $\boldsymbol{\sigma}_2$ is

$$\boldsymbol{\nu}_2(u, v) = 2\sin(3 + 2u)\,\boldsymbol{i} + 2\cos(3 + 2u)\,\boldsymbol{j}\,.$$

If $P = \boldsymbol{\sigma}(u_1, v_1) = \boldsymbol{\sigma}_2(u_2, v_2)$ is an arbitrary point on Σ, then

$$\cos(2 - u_1) = \sin(3 + 2u_2) \quad \text{and} \quad \sin(2 - u_1) = \cos(3 + 2u_2)\,,$$

and so

$$\boldsymbol{\nu}_1(u_1, v_1) = -v_1\boldsymbol{\nu}_2(u_2, v_2)\,.$$

Since v_1 is non-negative, the orientations are opposite.

c) We have $P_0 = \boldsymbol{\sigma}\big(2 - \frac{\pi}{2}, \frac{1}{2}\big) = \boldsymbol{\sigma}\big(\pi - \frac{2}{3}, \frac{3}{4}\big)$. Hence $\boldsymbol{\nu}_1(P_0) = -\boldsymbol{j}$, while $\boldsymbol{\nu}_2(P_0) = 2\boldsymbol{j}$. The unit normals then are $\boldsymbol{n}_1(P_0) = -\boldsymbol{j}$ and $\boldsymbol{n}_2(P_0) = \boldsymbol{j}$.

20. a) Expression (6.49) yields

$$\boldsymbol{\nu}(u, v) = -(2u + 3v)\,\boldsymbol{i} - (3u + 2v)\,\boldsymbol{j} + \boldsymbol{k}\,,$$

normalising which produces

$$\boldsymbol{n}(u, v) = -\frac{(2u + 3v)}{\|\boldsymbol{\nu}\|}\,\boldsymbol{i} - \frac{(3u + 2v)}{\|\boldsymbol{\nu}\|}\,\boldsymbol{j} + \frac{1}{\|\boldsymbol{\nu}\|}\,\boldsymbol{k}\,,$$

with $\|\boldsymbol{\nu}\|^2 = 13(u^2 + v^2) + 24uv + 1$.

b) The orthogonality follows by imposing $\boldsymbol{\nu}$ be parallel to the vector $8\boldsymbol{i} + 7\boldsymbol{j} - 2\boldsymbol{k}$. Thus we impose

$$\begin{cases} -(2u + 3v) = 8\lambda \\ -(3u + 2v) = 7\lambda \\ 1 = -2\lambda\,. \end{cases}$$

Solving the system gives $\lambda = -1/2$ and $u = 1/2$, $v = 1$. The required condition is valid for the point $P_0 = \boldsymbol{\sigma}\big(\frac{1}{2}, 1\big) = \big(\frac{1}{2}, 1, \frac{11}{4}\big)$.

7

Applying differential calculus

We conclude with this chapter the treatise of differential calculus for multivariate and vector-values functions. Two are the themes of concern: the Implicit Function Theorem with its applications, and the study of constrained extrema.

Given an equation in two or more independent variables, the Implicit Function Theorem provides sufficient conditions to express one variable in terms of the others. It helps to examine the nature of the level sets of a function, which, under regularity assumptions, are curves and surfaces in space. Finally, it furnishes the tools for studying the locus defined by a system of equations.

Constrained extrema, in other words extremum points of a map restricted to a given subset of the domain, can be approached in two ways. The first method, called parametric, reduces the problem to understading unconstrained extrema in lower dimension; the other geometrically-rooted method, relying on Lagrange multipliers, studies the stationary points of a new function, the Lagrangian of the problem.

7.1 Implicit Function Theorem

An equation of type

$$\boxed{f(x,y) = 0} \tag{7.1}$$

sets up an implicit relationship between the variables x and y; in geometrical terms it defines the locus of points $P = (x, y)$ in the plane whose coordinates satisfy the equation. Very often it is possible to write one variable as a function of the other at least locally (in the neighbourhood of one solution), as $y = \varphi(x)$ or $x = \psi(y)$. For example, the equation

$$f(x,y) = ax + by + c = 0\,, \qquad \text{with} \quad a^2 + b^2 \neq 0\,,$$

defines a line on the plane, and is equivalent to $y = \varphi(x) = -\frac{1}{b}(ax + c)$, if $b \neq 0$, or $x = \psi(y) = -\frac{1}{a}(by + c)$ if $a \neq 0$.

C. Canuto, A. Tabacco: *Mathematical Analysis II,* 2nd Ed.,
UNITEXT – La Matematica per il 3+2 85, DOI 10.1007/978-3-319-12757-6_7,

Instead,

$$f(x,y) = x^2 + y^2 - r^2 = 0\,, \qquad \text{with} \quad r > 0$$

defines a circle centred at the origin with radius r; if (x_0, y_0) belongs to the circle and $y_0 > 0$, we may solve the equation for y on a neighbourhood of x_0, and write $y = \varphi(x) = \sqrt{r^2 - x^2}$; if $x_0 < 0$, we can write $x = \psi(y) = -\sqrt{r^2 - y^2}$ on a neighbourhood of y_0. It is not possible to write y in terms of x on a neighbourhood of $x_0 = r$ or $-r$, nor x as a function of y around $y_0 = r$ or $-r$, unless we violate the definition of function itself. But in any of the cases where solving explicitly for one variable is possible, one partial derivative of f is non-zero ($\frac{\partial f}{\partial y} = b \neq 0$, $\frac{\partial f}{\partial x} = a \neq 0$ for the line, $\frac{\partial f}{\partial y} = 2y_0 > 0$, $\frac{\partial f}{\partial x} = 2x_0 < 0$ for the circle); conversely, where one variable cannot be made a function of the other the partial derivative is zero.

Moving on to more elaborate situations, we must distinguish between the *existence* of a map between the independent variables expressing (7.1) in an explicit way, and the *possibility of representing* such map in terms of known elementary functions. For example, in

$$y\,\mathrm{e}^{2x} + \cos y - 2 = 0$$

we can solve for x explicitly,

$$x = \frac{1}{2}\log(2 - \cos y) - \frac{1}{2}\log y\,,$$

but we are not able to do the same with

$$x^5 + 3x^2 y - 2y^4 - 1 = 0\,.$$

Nonetheless, analysing the graph of the function, e.g., around the solution $(x_0, y_0) = (1, 0)$, shows that the equation defines y in terms of x, and the study we are about to embark on will confirm rigorously the claim (see Example 7.2). Thus there is a huge interest in finding criteria that ensure a function $y = y(x)$, say, at least exists, even if the variable y cannot be written in an analytically-explicit way in terms of x. This lays a solid groundwork for the employ of numerical methods. We discuss below a milestone result in Mathematics, the Implicit Function Theorem, also known as Dini's Theorem in Italy, which yields a sufficient condition for the function $y(x)$ to exist; the condition is the non-vanishing of the partial derivative of f in the variable we are solving for.

All these considerations evidently apply to equations with three or more variables, like

$$\boxed{f(x,y,z) = 0\,.} \tag{7.2}$$

Under special assumptions, we are allowed to define $z = \varphi(x, y)$ as a function of x and y. A second generalisation concerns systems of equations.

Numerous are the applications of the Implicit Function Theorem. To name a few, many laws of Physics (think of Thermodynamics) tie two or more quantities through implicit relationships which, depending on the system's specific state, can be made explicit. In geometry the Theorem allows to describe as a regular simple surface, if locally, the locus of points whose coordinates satisfy (7.2); it lies at the heart of an 'intrinsic' viewpoint that regards a surface as a set of points defined by algebraic and transcendental equations.

So let us begin to see the two-dimensional version. Its proof may be found in Appendix A.1.4, p. 518.

Theorem 7.1 *Let Ω be a non-empty open set in $\mathbb{R}^2$ and $f:\Omega\to\mathbb{R}$ a $\mathcal{C}^1$ map. Assume at the point $(x_0,y_0)\in\Omega$ we have $f(x_0,y_0)=0$. If $\dfrac{\partial f}{\partial y}(x_0,y_0)\neq 0$, there exists a neighbourhood I of x_0 and a function $\varphi: I\to\mathbb{R}$ such that:*

i) $\big(x,\varphi(x)\big)\in\Omega$ *for any* $x\in I$;

ii) $y_0=\varphi(x_0)$;

iii) $f\big(x,\varphi(x)\big)=0$ *for any* $x\in I$;

iv) φ *is a* $\mathcal{C}^1$ *map on* I *with derivative*

$$\varphi'(x) = -\frac{\dfrac{\partial f}{\partial x}\big(x,\varphi(x)\big)}{\dfrac{\partial f}{\partial y}\big(x,\varphi(x)\big)}\,. \tag{7.3}$$

On a neighbourhood of (x_0,y_0) moreover, the zero set of f coincides with the graph of φ.

We remark that $x\mapsto\boldsymbol{\gamma}(x)=\big(x,\varphi(x)\big)$ is a curve in the plane whose trace is precisely the graph of φ.

Example 7.2

Consider the equation

$$x^5+3x^2y-2y^4=1\,,$$

already considered in the introduction. It is solved by $(x_0,y_0)=(1,0)$. To study the existence of other solutions in the neighbourhood of that point we define the function

$$f(x,y)=x^5+3x^2y-2y^4-1\,.$$

We have $f_y(x,y)=3x^2-8y^3$, hence $f_y(1,0)=3>0$. The above theorem guarantees there is a function $y=\varphi(x)$, differentiable on a neighbourhood I of $x_0=1$, such that $\varphi(1)=0$ and

$$x^5+3x^2\varphi(x)-2\big(\varphi(x)\big)^4=1\,,\qquad x\in I\,.$$

From $f_x(x,y)=5x^4+6xy$ we get $f_x(1,0)=5$, so $\varphi'(1)=-\frac{5}{3}$. The map φ is decreasing at $x_0=1$, where the tangent to its graph is $y=-\frac{5}{3}(x-1)$. □

In case $\dfrac{\partial f}{\partial x}(x_0, y_0) \neq 0$, we arrive at a similar result to Theorem 7.1 in which x and y are swapped. To summarise (see Fig. 7.1),

Corollary 7.3 *Let Ω be a non-empty open set in $\mathbb{R}^2$ and $f : \Omega \to \mathbb{R}$ a $\mathcal{C}^1$ map. The equation $f(x,y) = 0$ can be solved for one variable, $y = \varphi(x)$ or $x = \psi(y)$, around any zero (x_0, y_0) of f at which $\nabla f(x_0, y_0) \neq 0$ (i.e., around each regular point).*

There remains to understand what is the true structure of the zero set in the neighbourhood of a stationary point. Assuming f is $\mathcal{C}^2$, the previous chapter's study of the Hessian matrix $\boldsymbol{H}f(x_0, y_0)$ provides useful information.

i) If $\boldsymbol{H}f(x_0, y_0)$ is definite (positive or negative), (x_0, y_0) is a local strict minimum or maximum; therefore f will be strictly positive or negative on a whole punctured neighbourhood of (x_0, y_0), making (x_0, y_0) an *isolated zero* for f. An example is the origin for the map $f(x,y) = x^2 + y^2$.

ii) If $\boldsymbol{H}f(x_0, y_0)$ is indefinite, i.e., it has eigenvalues $\lambda_1 > 0$ and $\lambda_2 < 0$, one can prove that the zero set of f around (x_0, y_0) *is not* the graph of a function, because it consists of two distinct curves that meet at (x_0, y_0) and form an angle that depends upon the eigenvalues' ratio. For instance, the zeroes of $f(x,y) = 4x^2 - 25y^2$ belong to the lines $y = \pm\frac{2}{5}x$, crossing at the origin (Fig. 7.2, left).

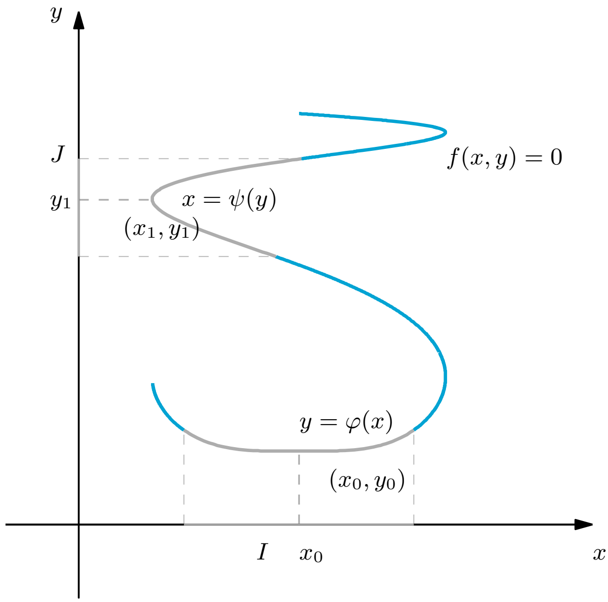


Figure 7.1. Some intervals where $f(x,y) = 0$ can be solved for x or y

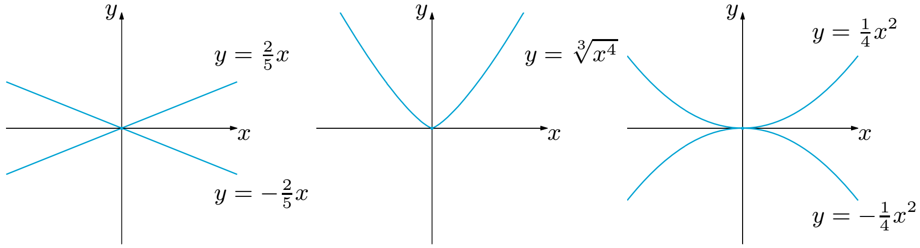


Figure 7.2. The zero sets of $f(x,y) = 4x^2 - 25y^2$ (left), $f(x,y) = x^4 - y^3$ (centre) and $f(x,y) = x^4 - 16y^2$ (right)

iii) If $\boldsymbol{H}f(x_0, y_0)$ has one or both eigenvalues equal zero, nothing can be said in general. For example, assuming the origin as point (x_0, y_0), the zeroes of $f(x,y) = x^4 - y^3$ coincide with the graph of $y = \sqrt[3]{x^4}$ (Fig. 7.2, centre). The map $f(x,y) = x^4 - 16y^2$, instead, vanishes along the parabolas $y = \pm\frac{1}{4}x^2$ that meet at the origin (Fig. 7.2, right).

We state the Implicit Function Theorem for functions in three variables. Its proof is easily obtained by adapting the one given in two dimensions.

Theorem 7.4 *Let Ω be a non-empty open subset of $\mathbb{R}^3$ and $f : \Omega \to \mathbb{R}$ a map of class $\mathcal{C}^1$ with a zero at $(x_0, y_0, z_0) \in \Omega$. If $\dfrac{\partial f}{\partial z}(x_0, y_0, z_0) \neq 0$, there exist a neighbourhood A of (x_0, y_0) and a function $\varphi : A \to \mathbb{R}$ such that:*

i) $\big(x, y, \varphi(x,y)\big) \in \Omega$ *for any* $(x,y) \in A$*;*

ii) $z_0 = \varphi(x_0, y_0)$*;*

iii) $f\big(x, y, \varphi(x,y)\big) = 0$ *for any* $(x,y) \in A$*;*

iv) φ *is* $\mathcal{C}^1$ *on* A*, with partial derivatives*

$$\frac{\partial \varphi}{\partial x}(x,y) = -\frac{\dfrac{\partial f}{\partial x}\big(x, y, \varphi(x,y)\big)}{\dfrac{\partial f}{\partial z}\big(x, y, \varphi(x,y)\big)}, \qquad \frac{\partial \varphi}{\partial y}(x,y) = -\frac{\dfrac{\partial f}{\partial y}\big(x, y, \varphi(x,y)\big)}{\dfrac{\partial f}{\partial z}\big(x, y, \varphi(x,y)\big)}. \qquad (7.4)$$

Moreover, on a neighbourhood of (x_0, y_0, z_0) the zero set of f coincides with the graph of φ.

The function $(x,y) \mapsto \boldsymbol{\sigma}(x,y,z) = \big(x, y, \varphi(x,y)\big)$ defines a regular simple surface in space, whose properties will be examined in Sect. 7.2.

The result may be applied, as happened above, under the assumption the point (x_0, y_0, z_0) be regular, possibly after swapping the independent variables' roles.

These two theorems are in fact subsumed by a general statement on vector-valued maps.

Let n, m be integers with $n \geq 2$ and $1 \leq m \leq n-1$. Take an open, non-empty set Ω in $\mathbb{R}^n$ and let $\boldsymbol{F} : \Omega \to \mathbb{R}^m$ be a $\mathcal{C}^1$ map on it. Select m of the n independent variables $x_1, x_2, \ldots, x_n$; without loss of generality we can take the last m variables (always possible by permuting the variables). Then write $\boldsymbol{x}$ as two sets of coordinates $(\boldsymbol{\xi}, \boldsymbol{\mu})$, where $\boldsymbol{\xi} = (x_1, \ldots, x_{n-m}) \in \mathbb{R}^{n-m}$ and $\boldsymbol{\mu} = (x_{n-m+1}, \ldots, x_n) \in \mathbb{R}^m$. Correspondingly, decompose the Jacobian of $\boldsymbol{F}$ at $\boldsymbol{x} \in \Omega$ into the matrix $\boldsymbol{J_\xi F}(\boldsymbol{x})$ with m rows and $n-m$ columns containing the first $n-m$ columns of $\boldsymbol{JF}(\boldsymbol{x})$, and the $m \times m$ matrix $\boldsymbol{J_\mu F}(\boldsymbol{x})$ formed by the remaining m columns of $\boldsymbol{JF}(\boldsymbol{x})$.

We are ready to solve locally the implicit equation $\boldsymbol{F}(\boldsymbol{x}) = \boldsymbol{0}$, now reading $\boldsymbol{F}(\boldsymbol{\xi}, \boldsymbol{\mu}) = \boldsymbol{0}$, and obtain $\boldsymbol{\mu} = \boldsymbol{\Phi}(\boldsymbol{\xi})$.

Theorem 7.5 (**Implicit Function Theorem**, or **Dini's Theorem**) *With the above conventions, let $\boldsymbol{x}_0 = (\boldsymbol{\xi}_0, \boldsymbol{\mu}_0) \in \Omega$ be a point such that $\boldsymbol{F}(\boldsymbol{x}_0) = \boldsymbol{0}$. If the matrix $\boldsymbol{J_\mu F}(\boldsymbol{x}_0)$ is non-singular there exist neighbourhoods $A \subseteq \Omega$ of $\boldsymbol{x}_0$, and I of $\boldsymbol{\xi}_0$ in $\mathbb{R}^{n-m}$, such that the zero set*

$$\{\boldsymbol{x} \in A : \boldsymbol{F}(\boldsymbol{x}) = \boldsymbol{0}\}$$

of $\boldsymbol{F}$ coincides with the graph

$$\{\boldsymbol{x} = \big(\boldsymbol{\xi}, \boldsymbol{\Phi}(\boldsymbol{\xi})\big) : \boldsymbol{\xi} \in I\}$$

of a $\mathcal{C}^1$ map $\boldsymbol{\Phi} : I \to \mathbb{R}^m$ satisfying $\boldsymbol{\Phi}(\boldsymbol{\xi}_0) = \boldsymbol{\mu}_0$. On I, the Jacobian of $\boldsymbol{\Phi}$ (with m rows and $n-m$ columns) is the solution of the linear system

$$\boldsymbol{J_\mu F}\big(\boldsymbol{\xi}, \boldsymbol{\Phi}(\boldsymbol{\xi})\big)\, \boldsymbol{J\Phi}(\boldsymbol{\xi}) = -\boldsymbol{J_\xi F}\big(\boldsymbol{\xi}, \boldsymbol{\Phi}(\boldsymbol{\xi})\big)\,. \tag{7.5}$$

Theorems 7.1 and 7.4 are special instances of the above, as can be seen taking $n = 2$, $m = 1$, and $n = 3$, $m = 1$. The next example elucidates yet another case of interest.

Example 7.6

Consider the system of two equations in three variables

$$\begin{cases} f_1(x, y, z) = 0 \\ f_2(x, y, z) = 0 \end{cases}$$

where f_i are $\mathcal{C}^1$ on an open set Ω in $\mathbb{R}^3$, and call $\boldsymbol{x}_0 = (x_0, y_0, z_0) \in \Omega$ a solution. Specialising the theorem with $n = 3$, $m = 2$ and $\boldsymbol{F} = (f_1, f_2)$ ensures that if

$$\begin{pmatrix} \dfrac{\partial f_1}{\partial y}(\boldsymbol{x}_0) & \dfrac{\partial f_1}{\partial z}(\boldsymbol{x}_0) \\[2ex] \dfrac{\partial f_2}{\partial y}(\boldsymbol{x}_0) & \dfrac{\partial f_2}{\partial z}(\boldsymbol{x}_0) \end{pmatrix}$$

is not singular, the system admits infinitely many solutions around $\boldsymbol{x}_0$; these are of the form $\big(x, \varphi_1(x), \varphi_2(x)\big)$, where $\boldsymbol{\Phi} = (\varphi_1, \varphi_2) : I \subset \mathbb{R} \to \mathbb{R}^2$ is $\mathcal{C}^1$ on a neighbourhood I of x_0.

The components of the first derivative $\boldsymbol{\Phi}'$ at $x \in I$ solve the system

$$\begin{pmatrix} \frac{\partial f_1}{\partial y}\big(x,\varphi_1(x),\varphi_2(x)\big) & \frac{\partial f_1}{\partial z}\big(x,\varphi_1(x),\varphi_2(x)\big) \\ \frac{\partial f_2}{\partial y}\big(x,\varphi_1(x),\varphi_2(x)\big) & \frac{\partial f_2}{\partial z}\big(x,\varphi_1(x),\varphi_2(x)\big) \end{pmatrix} \begin{pmatrix} \varphi_1'(x) \\ \\ \varphi_2'(x) \end{pmatrix} =$$

$$= - \begin{pmatrix} \frac{\partial f_1}{\partial x}\big(x,\varphi_1(x),\varphi_2(x)\big) \\ \frac{\partial f_2}{\partial x}\big(x,\varphi_1(x),\varphi_2(x)\big) \end{pmatrix}.$$

The function $x \mapsto \boldsymbol{\gamma}(x) = \big(x,\varphi_1(x),\varphi_2(x)\big)$ represents a (regular) curve in space. □

7.1.1 Local invertibility of a function

One application is the following: if the Jacobian of a regular map of $\mathbb{R}^n$ is non-singular, then the function is invertible around the point.

This generalises to multivariable calculus a property of functions of one real variable, namely: if f is $\mathcal{C}^1$ around $x_0 \in \operatorname{dom} f \subseteq \mathbb{R}$ and if $f'(x_0) \neq 0$, then f' has constant sign around x_0; as such it is strictly monotone, hence invertible; furthermore, f^{-1} is differentiable and $(f^{-1})'(y_0) = 1/f'(x_0)$.

Proposition 7.7 *Let $\boldsymbol{f} : \operatorname{dom} \boldsymbol{f} \subseteq \mathbb{R}^n \to \mathbb{R}^n$ be $\mathcal{C}^1$ around a point $\boldsymbol{x}_0 \in \operatorname{dom} \boldsymbol{f}$. If $J\boldsymbol{f}(\boldsymbol{x}_0)$ is non-singular, the inverse function $\boldsymbol{x} = \boldsymbol{f}^{-1}(\boldsymbol{y})$ is well defined on a neighbourhood of $\boldsymbol{y}_0 = \boldsymbol{f}(\boldsymbol{x}_0)$, of class $\mathcal{C}^1$, and satisfies*

$$J(\boldsymbol{f}^{-1})(\boldsymbol{y}_0) = \big(J\boldsymbol{f}(\boldsymbol{x}_0)\big)^{-1}. \tag{7.6}$$

Proof. Define the auxiliary map $\boldsymbol{F} : \operatorname{dom} \boldsymbol{f} \times \mathbb{R}^n \subseteq \mathbb{R}^{2n} \to \mathbb{R}^n$ by $\boldsymbol{F}(\boldsymbol{x},\boldsymbol{y}) = \boldsymbol{f}(\boldsymbol{x}) - \boldsymbol{y}$. Then $\boldsymbol{F}(\boldsymbol{x}_0,\boldsymbol{y}_0) = \boldsymbol{0}$ and $J\boldsymbol{F}(\boldsymbol{x}_0,\boldsymbol{y}_0) = \big(\, J\boldsymbol{f}(\boldsymbol{x}_0)\,, \boldsymbol{I}\,\big)$ (an $n \times 2n$ matrix). Since $J_{\boldsymbol{x}}\boldsymbol{F}(\boldsymbol{x}_0,\boldsymbol{y}_0) = J\boldsymbol{f}(\boldsymbol{x}_0)$, Theorem 7.5 guarantees there is a neighbourhood $B_r(\boldsymbol{y}_0)$ and a $\mathcal{C}^1$ map $\boldsymbol{g} : B_r(\boldsymbol{y}_0) \to \operatorname{dom} \boldsymbol{f}$ such that $\boldsymbol{F}\big(\boldsymbol{g}(\boldsymbol{y}),\boldsymbol{y}\big) = \boldsymbol{0}$, i.e., $\boldsymbol{f}\big(\boldsymbol{g}(\boldsymbol{y})\big) = \boldsymbol{y}$, $\forall \boldsymbol{y} \in B_r(\boldsymbol{y}_0)$. Therefore $\boldsymbol{g}(\boldsymbol{y}) = \boldsymbol{f}^{-1}(\boldsymbol{y})$, so (7.6) follows from (7.5). □

The formula for the Jacobian of the inverse function was obtained in Corollary 6.16 as a consequence of the chain rule; that corollary's assumptions are eventually substantiated by the previous proposition.

What we have seen can be read in terms of a system of equations. Let us interpret

$$\boldsymbol{f}(\boldsymbol{x}) = \boldsymbol{y} \tag{7.7}$$

as a (non-linear) system of n equations in n unknowns, where the right-hand side $\boldsymbol{y}$ is given and $\boldsymbol{x}$ is the solution. Suppose, for a given datum $\boldsymbol{y}_0$, that we know a solution $\boldsymbol{x}_0$. If the Jacobian $\boldsymbol{J}\boldsymbol{f}(\boldsymbol{x}_0)$ is non-singular, the equation (7.7) admits one, and only one solution $\boldsymbol{x}$ in the proximity of $\boldsymbol{x}_0$, for any $\boldsymbol{y}$ sufficiently close to $\boldsymbol{y}_0$. In turn, the Jacobian's invertibility at $\boldsymbol{x}_0$ is equivalent to the fact that the linear system

$$\boldsymbol{f}(\boldsymbol{x}_0) + \boldsymbol{J}\boldsymbol{f}(\boldsymbol{x}_0)(\boldsymbol{x} - \boldsymbol{x}_0) = \boldsymbol{y}\,, \tag{7.8}$$

obtained by *linearising* the left-hand side of (7.7) around $\boldsymbol{x}_0$, admits a solution whichever $\boldsymbol{y} \in \mathbb{R}^n$ is taken (it suffices to write the system as $\boldsymbol{J}\boldsymbol{f}(\boldsymbol{x}_0)\boldsymbol{x} = \boldsymbol{y} - \boldsymbol{f}(\boldsymbol{x}_0) + \boldsymbol{J}\boldsymbol{f}(\boldsymbol{x}_0)\boldsymbol{x}_0$ and notice the right-hand side assumes any value of $\mathbb{R}^n$ as $\boldsymbol{y}$ varies in $\mathbb{R}^n$).

In conclusion, the knowledge of one solution $\boldsymbol{x}_0$ of a non-linear equation for a given datum $\boldsymbol{y}_0$, provided the linearised equation around that point can be solved, guarantees the solvability of the non-linear equation for any value of the datum that is sufficiently close to $\boldsymbol{y}_0$.

Example 7.8

The system of non-linear equations

$$\begin{cases} x^3 + y^3 - 2xy = a \\ 3x^2 + xy^2 = b \end{cases}$$

admits $(x_0, y_0) = (1, 2)$ as solution for $(a, b) = (5, 7)$. Calling $\boldsymbol{f}(x, y) = (x^3 + y^3 - 2xy, 3x^2 + xy^2)$, we have

$$\boldsymbol{J}\boldsymbol{f}(x, y) = \begin{pmatrix} 3x^2 - 2y & 3y^2 - 2x \\ 6x + y^2 & 2xy \end{pmatrix}.$$

As

$$\det \boldsymbol{J}\boldsymbol{f}(1, 2) = \det \begin{pmatrix} 10 & 4 \\ -1 & 10 \end{pmatrix} = 104 > 0\,,$$

we can say the system admits a unique solution (x, y) for any choice of a, b close enough to $5, 7$ respectively. □

7.2 Level curves and level surfaces

At this juncture we can resume level sets $L(f, c)$ of a given function f, seen in (4.20). The differential calculus developed so far allows us to describe such sets at least in the case of suitably regular functions. The study of level sets offers useful informations on the function's behaviour, as the constant c varies. At the same time, solving an equation like

$$f(x, y) = 0\,, \qquad \text{or} \qquad f(x, y, z) = 0\,,$$

is, as a matter of fact, analogous to determining the level set $L(f, 0)$ of f. Under suitable assumptions, as we know, implicit relations of this type among the independent variables generate regular and simple curves, or surfaces, in space.

In the sequel we shall consider only two and three dimensions, beginning from the former situation.

7.2.1 Level curves

Let then f be a real map of two real variables; given $c \in \operatorname{im} f$, denote by $L(f,c)$ the corresponding level set and let $\boldsymbol{x}_0 \in L(f,c)$. Suppose f is $\mathcal{C}^1$ on a neighbourhood of the regular point $\boldsymbol{x}_0$, so that $\nabla f(\boldsymbol{x}_0) \neq \mathbf{0}$. Corollary 7.3 applies to the map $f(\boldsymbol{x}) - c$ to say that on a certain neighbourhood $B(\boldsymbol{x}_0)$ of $\boldsymbol{x}_0$ the points of $L(f,c)$ lie on a regular simple curve $\boldsymbol{\gamma} : I \to B(\boldsymbol{x}_0)$ (a graph in one of the variables); moreover, the point $t_0 \in I$ such that $\boldsymbol{\gamma}(t_0) = \boldsymbol{x}_0$ is in the interior of I. In other terms,

$$L(f,c) \cap B(\boldsymbol{x}_0) = \{\boldsymbol{x} \in \mathbb{R}^2 : \boldsymbol{x} = \boldsymbol{\gamma}(t), \text{ with } t \in I\},$$

whence

$$f\big(\boldsymbol{\gamma}(t)\big) = c \qquad \forall t \in I\,. \tag{7.9}$$

If the argument is valid for any point of $L(f,c)$, hence if f is $\mathcal{C}^1$ on an open set containing $L(f,c)$ and all points of $L(f,c)$ are regular, the level set is made of traces of regular, simple curves. If so, one speaks of *level curves* of f. The next property is particularly important (see Fig. 7.3).

Proposition 7.9 *The gradient of a map is orthogonal to level curves at any of their regular points.*

Proof. Differentiating (7.9) at $t_0 \in I$ (recall (6.13)) gives

$$\nabla f(\boldsymbol{x}_0) \cdot \boldsymbol{\gamma}'(t) = 0\,,$$

and the result follows. □

Moving along a level curve from $\boldsymbol{x}_0$ maintains the function constant, whereas in the perpendicular direction the function undergoes the maximum variation (by Proposition 5.10). It can turn out useful to remark that $\mathbf{curl} f$, as of (6.7), is always orthogonal to the gradient, so it is tangent to the level curves of f.

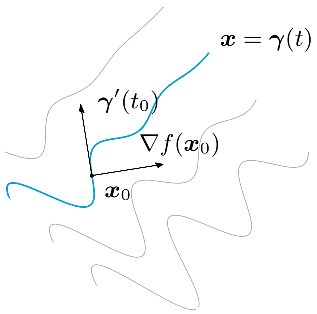


Figure 7.3. Level curves and gradient of a map

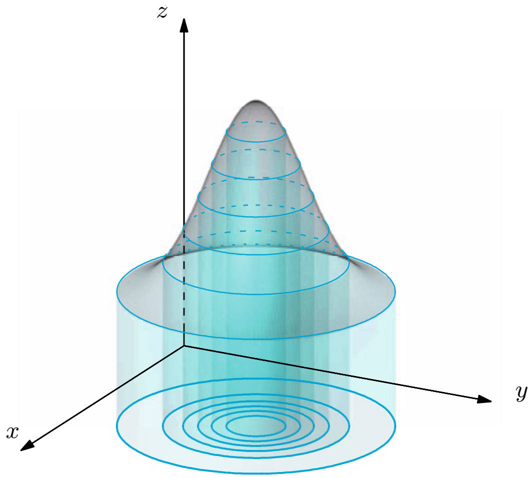


Figure 7.4. The relationship between graph and level sets

Fig. 7.4 shows the relationship between level curves and the graph of a map. Recall that the level curves defined by $f(x, y) = c$ are the projections on the xy-plane of the intersection between the graph of f and the planes $z = c$. In this way, if we vary c by a fixed increment and plot the corresponding level curves, the latter will be closer to each other where the graph is "steeper", and sparser where it "flattens out".

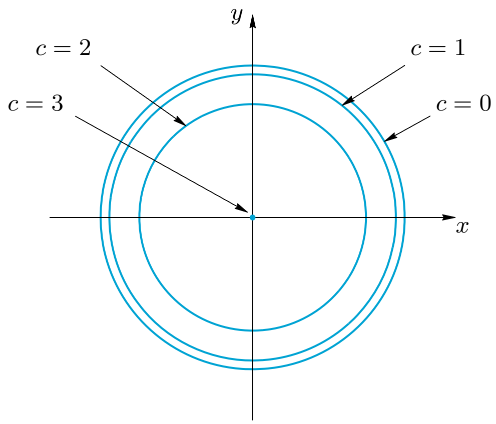


Figure 7.5. Level curves of $f(x, y) = \sqrt{9 - x^2 - y^2}$

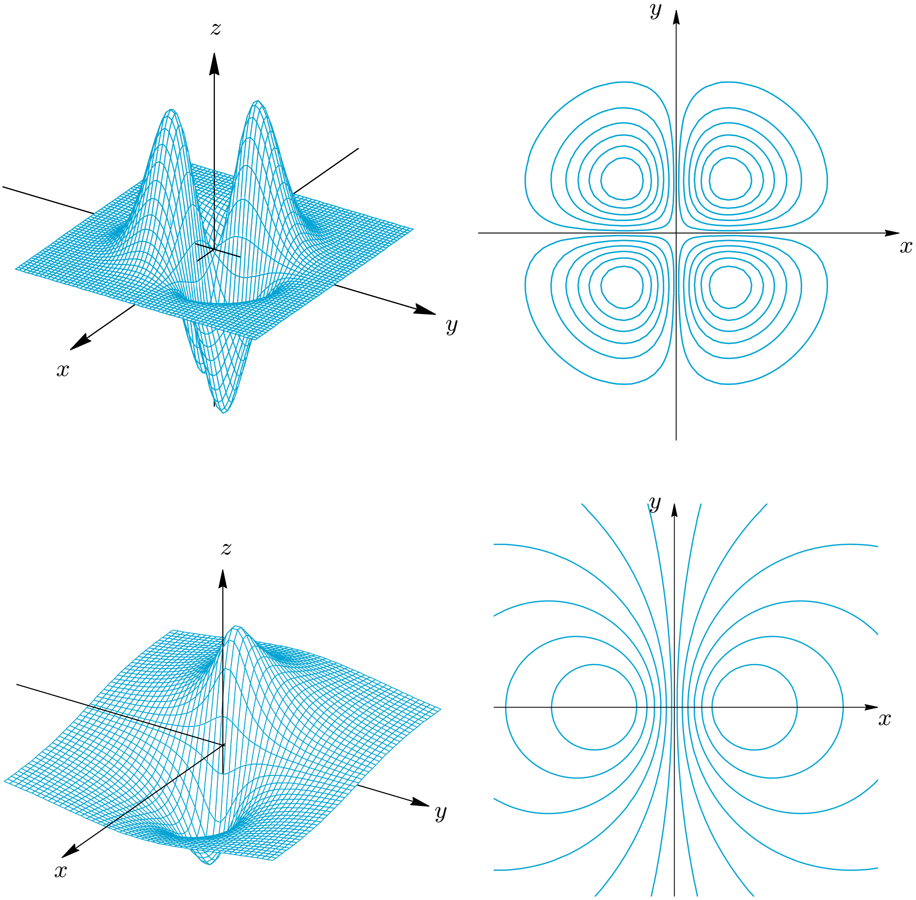


Figure 7.6. Graph and level curves of $f(x,y) = -xye^{-x^2-y^2}$ (top), and $f(x,y) = -\dfrac{x}{x^2+y^2+1}$ (bottom)

Examples 7.10

i) The graph of $f(x,y) = \sqrt{9-x^2-y^2}$ was shown in Fig. 4.8. The level curves are

$$\sqrt{9-x^2-y^2} = c \qquad \text{i.e.,} \qquad x^2+y^2 = 9-c^2\,.$$

This is a family of circles with centre the origin and radii $\sqrt{9-c^2}$, see Fig. 7.5.

ii) Fig. 7.6 shows some level curves and the corresponding graph. □

Patently, a level set may contain non-regular points of f, hence stationary points. In such a case the level set may not be representable, around the point, as trace of a curve.

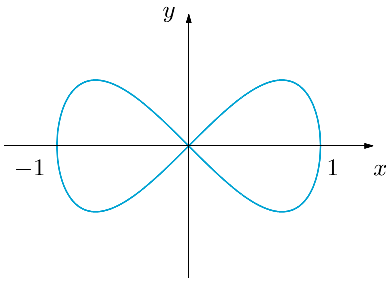


Figure 7.7. Level set $L(f,0)$ for $f(x,y) = x^4 - x^2 + y^2$

Example 7.11

The level set $L(f,0)$ of

$$f(x,y) = x^4 - x^2 + y^2$$

is shown in Fig. 7.7. The origin is a saddle-like stationary point. On every neighbourhood of the origin the level set consists of two branches that intersect orthogonally; as such, it cannot be the graph of a function in one of the variables. □

Archetypal examples of how level curves might be employed are the maps used in topography, as in Fig. 7.8. On pictures representing a piece of land level curves join points on the Earth's surface at the same height above sea level: such paths do not go uphill nor downhill, and in fact they are called *isoclines* (lit. 'of equal inclination').

Another frequent instance is the representation of the temperature of a certain region at a given time. The level curves are called *isotherms* and connect points with the same temperature. Similarly for *isobars*, the level curves on a map joining points having identical atmospheric pressure.

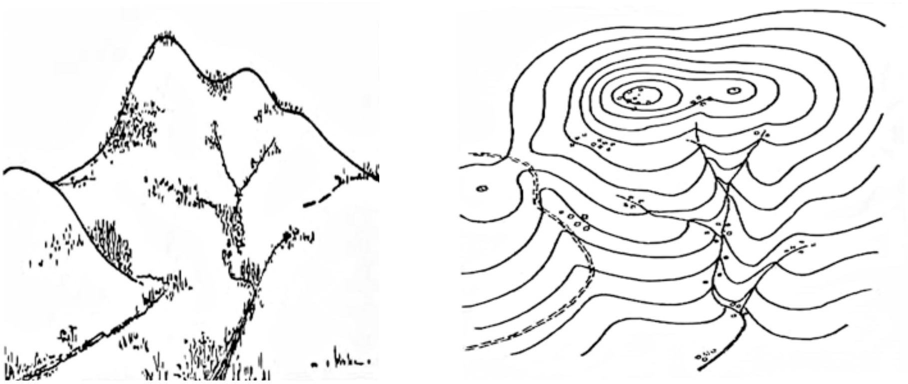


Figure 7.8. Three-dimensional representation of a territory and the corresponding isoclines

7.2.2 Level surfaces

In presence of three independent variables, level sets can be parametrised around regular points. To be precise, if f is $\mathcal{C}^1$ around $\boldsymbol{x}_0 \in L(f,c)$ and $\nabla f(\boldsymbol{x}_0) \neq \boldsymbol{0}$, the 3-dimensional analogue of Corollary 7.3 (that easily descends from Theorem 7.4) guarantees the existence of a neighbourhood $B(\boldsymbol{x}_0)$ of $\boldsymbol{x}_0$ and of a regular, simple surface $\boldsymbol{\sigma} : \mathcal{R} \to B(\boldsymbol{x}_0)$ such that

$$L(f,c) \cap B(\boldsymbol{x}_0) = \{\boldsymbol{x} \in \mathbb{R}^3 : \boldsymbol{x} = \boldsymbol{\sigma}(u,v), \text{ with } (u,v) \in \mathcal{R}\}\,,$$

hence

$$f\big(\boldsymbol{\sigma}(u,v)\big) = c \qquad \forall (u,v) \in \mathcal{R}\,. \tag{7.10}$$

In other terms $L(f,c)$ is locally the image of a level surface of f, and Proposition 7.9 has an analogue

Proposition 7.12 *At any regular point, the gradient of a function is parallel to the normal of the level surface.*

Proof. The partial derivatives of $f \circ \boldsymbol{\sigma}$ at $(u_0,v_0) \in \mathcal{R}$ such that $\boldsymbol{\sigma}(u_0,v_0) = \boldsymbol{x}_0$ are

$$\nabla f(\boldsymbol{x}_0) \cdot \frac{\partial \boldsymbol{\sigma}}{\partial u}(u_0,v_0) = 0\,, \qquad \nabla f(\boldsymbol{x}_0) \cdot \frac{\partial \boldsymbol{\sigma}}{\partial v}(u_0,v_0) = 0\,.$$

The claim follows now from (6.48) □

Examples 7.13

i) The level surfaces of $f(x,y,z) = x^2 + y^2 + z^2$ form a family of concentric spheres with radii $\sqrt{c}$. The normal vectors are aligned with the gradient of f, see Fig. 7.9.

ii) Let us explain how the Implicit Function Theorem, and its corollaries, may be used for studying a surface defined by an algebraic equation. Let $\Sigma \subset \mathbb{R}^3$ be the set of points satisfying

$$f(x,y,z) = 2x^4 + 2y^4 + 2z^4 + x - y - 6 = 0\,.$$

The gradient $\nabla f(x,y,z) = (8x^3+1, 8y^3-1, 8z^3)$ vanishes only at $P_0 = (-\frac{1}{2}, \frac{1}{2}, 0)$, but the latter does not satisfy the equation, i.e., $P_0 \notin \Sigma$. Therefore all points P of Σ are regular, and Σ is a regular simple surface; around any point P the surface can be represented as the graph of a function expressing one variable in terms of the other two.

Notice also

$$\lim_{\boldsymbol{x} \to \boldsymbol{\infty}} f(\boldsymbol{x}) = +\infty\,,$$

so the open set

$$\Omega = \{(x,y,z) \in \mathbb{R}^3 : f(x,y,z) < 0\}\,,$$

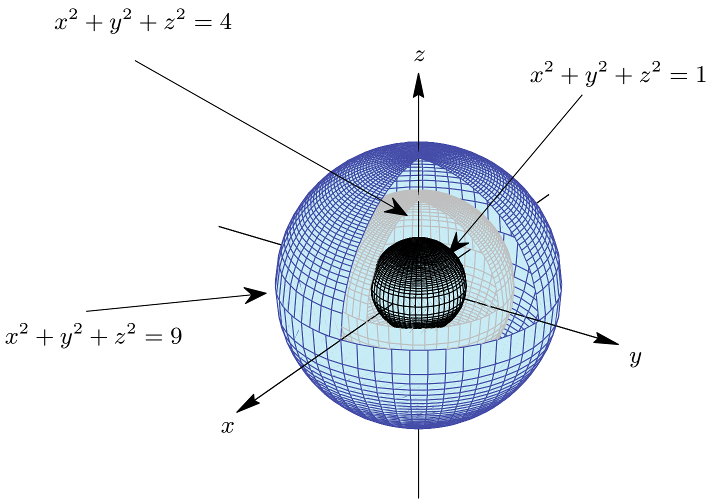


Figure 7.9. Level surfaces of $f(x, y, z) = \sqrt{x^2 + y^2 + z^2}$

containing the origin and of which Σ is the boundary, is bounded; consequently Σ is a compact set. One could prove Σ has no boundary, making it a closed surface.

Proposition 7.9 proves, for example, that the normal vector (up to sign) to Σ at $P_1 = (1, 1, 1)$ is proportional to $\nabla f(1, 1, 1) = (9, 7, 8)$. The unit normal $\boldsymbol{\nu}$ at P_1, chosen to point away from the origin ($\boldsymbol{\nu} \cdot \boldsymbol{x} > 0$), is then

$$\boldsymbol{\nu} = \frac{1}{\sqrt{194}}(9\boldsymbol{i} + 7\boldsymbol{j} + 8\boldsymbol{k}) . \qquad \square$$

7.3 Constrained extrema

With Sect. 5.6 we have learnt how to find extremum points lying inside the domain of a sufficiently regular function. That does not exhaust all possible extrema though, because there might be some lying on the domain's boundary (without mentioning some extrema could be non-regular points). At the same time in many applications it is required we search for minimum and maximum points on a particular subset of the domain; for instance, if the subset is defined through equations or inequalities of the independent variable, one speaks about **constrained extrema**.

The present section develops two methods to find extrema of the kind just described, and for this we start with an example.

Suppose we want to find the minimum of the linear map $f(\boldsymbol{x}) = \boldsymbol{a} \cdot \boldsymbol{x}$, $\boldsymbol{a} = (1, \sqrt{3}) \in \mathbb{R}^2$, among the unit vectors $\boldsymbol{x}$ of the plane, hence under the constraint

$\|\boldsymbol{x}\| = 1$. Geometrically, the point $P = \boldsymbol{x} = (x, y)$ can only move on the unit circle $x^2 + y^2 = 1$. If $g(\boldsymbol{x}) = x^2 + y^2 - 1 = 0$ is the equation of the circle and $G = \{\boldsymbol{x} \in \mathbb{R}^2 : g(\boldsymbol{x}) = 0\}$ the set of constrained points, we are looking for $\boldsymbol{x}_0$ satisfying

$$\boxed{\boldsymbol{x}_0 \in G \qquad \text{and} \qquad f(\boldsymbol{x}_0) = \min_{\boldsymbol{x}\in G} f(\boldsymbol{x})\,.}$$

The problem can be tackled in two different ways, one privileging the analytical point of view, the other the geometrical aspects. The first consists in reducing the number of variables from two to one, by observing that the constraint is a simple closed arc and as such can be written parametrically. Precisely, set $\boldsymbol{\gamma} : [0, 2\pi] \to \mathbb{R}^2$, $\boldsymbol{\gamma}(t) = (\cos t, \sin t)$, so that G coincides with the trace of $\boldsymbol{\gamma}$. Then f restricted to G becomes $f \circ \boldsymbol{\gamma}$, a function of the variable t; moreover,

$$\min_{\boldsymbol{x}\in G} f(\boldsymbol{x}) = \min_{t\in[0,2\pi]} f\big(\boldsymbol{\gamma}(t)\big)\,.$$

Now we find the extrema of $\varphi(t) = f\big(\boldsymbol{\gamma}(t)\big) = \cos t + \sqrt{3}\sin t$; the map is periodic of period 2π, so we can think of it as a map on $\mathbb{R}$ and ignore extrema that fall outside $[0, 2\pi]$. The first derivative $\varphi'(t) = -\sin t + \sqrt{3}\cos t$ vanishes at $t = \frac{\pi}{3}$ and $\frac{4}{3}\pi$. As $\varphi''(\frac{\pi}{3}) = -2 < 0$ and $\varphi''(\frac{4}{3}\pi) = 2 > 0$, we have a maximum at $t = \frac{\pi}{3}$ and a minimum at $t = \frac{4}{3}\pi$. Therefore there is only one solution to the original constrained problem, namely $\boldsymbol{x}_0 = \big(\cos\frac{4}{3}\pi, \sin\frac{4}{3}\pi\big) = -\big(\frac{1}{2}, \frac{\sqrt{3}}{2}\big)$.

The same problem can be understood from a geometrical perspective relying on Fig. 7.10. The gradient of f is $\nabla f(\boldsymbol{x}) = \boldsymbol{a}$, and we recall f is increasing along

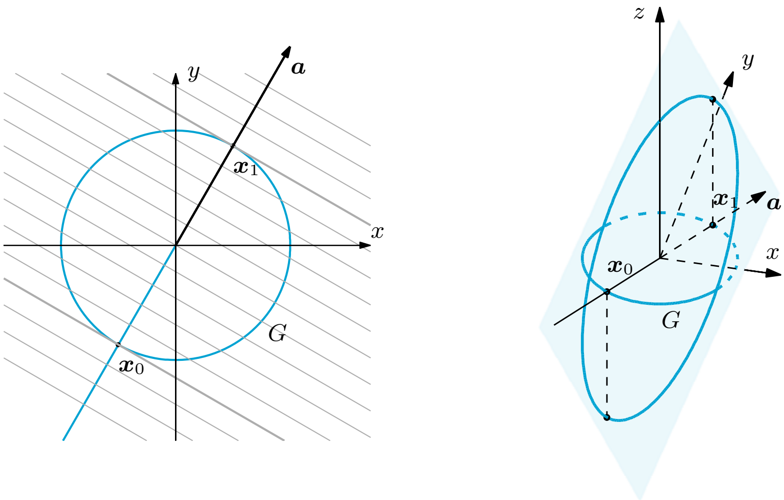


Figure 7.10. Level curves of $f(\boldsymbol{x}) = \boldsymbol{a} \cdot \boldsymbol{x}$ (left) and restriction to the constraint G (right)

$\boldsymbol{a}$ (with the greatest rate, actually, see Proposition 5.10); the level curves of f are the perpendicular lines to $\boldsymbol{a}$. Obviously, f restricted to the unit circle will reach its minimum and maximum at the points where the level curves touch the circle itself. These points are, respectively, $\boldsymbol{x}_0 = -\left(\frac{1}{2}, \frac{\sqrt{3}}{2}\right)$, which we already know of, and its antipodal point $\boldsymbol{x}_1 = \left(\frac{1}{2}, \frac{\sqrt{3}}{2}\right)$. They may also be characterised as follows. The gradient is orthogonal to level curves, and the circle is indeed a level curve for $g(\boldsymbol{x})$; to say therefore that the level curves of f and g are tangent at $\boldsymbol{x}_i$, $i = 0, 1$ tantamounts to requiring that the gradients are parallel; otherwise said at each point $\boldsymbol{x}_i$ there is a constant λ such that

$$\nabla f(\boldsymbol{x}) = \lambda \nabla g(\boldsymbol{x}) .$$

These, together with the constraining equation $g(\boldsymbol{x}) = 0$, allow to find the constrained extrema of f. In fact, we have

$$\begin{cases} \lambda x = 1 \\ \lambda y = \sqrt{3} \\ x^2 + y^2 = 1 ; \end{cases}$$

substituting the first and second in the third equation gives

$$\begin{cases} x = \dfrac{1}{\lambda} \\ y = \dfrac{\sqrt{3}}{\lambda} \\ \lambda^2 = 4 , \quad \text{so} \quad \lambda = \pm 2 , \end{cases}$$

and the solutions are $\boldsymbol{x}_0$ and $\boldsymbol{x}_1$. Since $f(\boldsymbol{x}_0) < f(\boldsymbol{x}_1)$, $\boldsymbol{x}_0$ will be the minimum point and $\boldsymbol{x}_1$ the maximum point for f.

In the general case, let $f : \operatorname{dom} f \subseteq \mathbb{R}^n \to \mathbb{R}$ be a map in n real variables, and $G \in \operatorname{dom} f$ a proper subset of the domain of f, called in the sequel **admissible set**. First of all we introduce the notion of constrained extremum.

Definition 7.14 *A point $\boldsymbol{x}_0 \in G$ is said a* **relative extremum point** *of f* **constrained to** *G if $\boldsymbol{x}_0$ is a relative extremum for the restriction $f_{|G}$ of f to G. In other words, there exists a neighbourhood $B_r(\boldsymbol{x}_0)$ such that*

$$\forall \boldsymbol{x} \in B_r(\boldsymbol{x}_0) \cap G \qquad f(\boldsymbol{x}) \leq f(\boldsymbol{x}_0)$$

for a constrained maximum, or

$$\forall \boldsymbol{x} \in B_r(\boldsymbol{x}_0) \cap G \qquad f(\boldsymbol{x}) \geq f(\boldsymbol{x}_0)$$

in case of a constrained minimum.

Constrained absolute maxima and minima are respectively defined by

$$f(\boldsymbol{x}_0) = \max_{\boldsymbol{x}\in G} f(\boldsymbol{x}) \qquad \textit{and} \qquad f(\boldsymbol{x}_0) = \min_{\boldsymbol{x}\in G} f(\boldsymbol{x})\,.$$

A constrained extremum f is not necessarily an extremum as well. The function $f(x,y) = xy$, for example, has a saddle at the origin, but when we restrict to the bisectrix of the first and third quadrant (or second and fourth), the origin becomes a constrained absolute minimum (maximum).

A recurring situation is that in which the set G is a subset defined by equations or inequalities, called **constraints**. We begin by examining one constraint and one equation, and consider a map $g : \operatorname{dom} g \subseteq \mathbb{R}^n \to \mathbb{R}$, with $G \subset \operatorname{dom} g$, such that

$$G = \{\boldsymbol{x} \in \operatorname{dom} g : g(\boldsymbol{x}) = 0\}\,;$$

in other terms G is the zero set of g, or the level set $L(g,0)$. We discuss both approaches mentioned in the foreword for this admissible set.

Assume henceforth f and g are $\mathcal{C}^1$ on an open set containing G. The aim is to find the constrained extrema of f by looking at the stationary points of suitable maps.

7.3.1 The method of parameters

This method originates from the possibility of writing the admissible set G in parametric form, i.e., as image of a map defined on a subset of $\mathbb{R}^{n-1}$. To be precise, suppose we know a $\mathcal{C}^1$ map $\boldsymbol{\psi} : A \subseteq \mathbb{R}^{n-1} \to \mathbb{R}^n$ such that

$$\boxed{G = \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{x} = \boldsymbol{\psi}(\boldsymbol{u}) \text{ with } \boldsymbol{u} \in A\}\,.}$$

In dimensions 2 and 3, this forces G to be the trace of a curve $\boldsymbol{\psi} = \boldsymbol{\gamma} : I \to \mathbb{R}^2$ or a surface $\boldsymbol{\psi} = \boldsymbol{\sigma} : \mathcal{R} \to \mathbb{R}^3$. The function $\boldsymbol{\psi}$ can be typically recovered from the equation $g(\boldsymbol{x}) = 0$; the Implicit Function Theorem 7.5 guarantees this is possible provided G consists of regular points for g. Studying f restricted to G is equivalent to studying the composite $f \circ \boldsymbol{\psi}$, so we will detect the latter's extrema on A, which has one dimension less than the domain of f. In particular, when $n = 2$ we will consider the map $t \mapsto f\big(\boldsymbol{\gamma}(t)\big)$ of one variable, when $n = 3$ we will have the two-variable map $(u,v) \mapsto f\big(\boldsymbol{\sigma}(u,v)\big)$.

Let us then examine the interior points of A first, because if one such, say $\boldsymbol{u}_0$, is an extremum for $f \circ \boldsymbol{\psi}$, then it is stationary; putting $\boldsymbol{x}_0 = \boldsymbol{\psi}(\boldsymbol{u}_0)$, we necessarily have

$$\boxed{\nabla_{\boldsymbol{u}}(f \circ \boldsymbol{\psi})(\boldsymbol{u}_0) = \nabla f\big(\boldsymbol{\psi}(\boldsymbol{u}_0)\big)\boldsymbol{J}\boldsymbol{\psi}(\boldsymbol{u}_0) = \boldsymbol{0}\,.}$$

Stationary points solving the above equation have to be examined carefully to tell whether they are extrema or not. After that, we inspect the boundary ∂A to find other possible extrema.

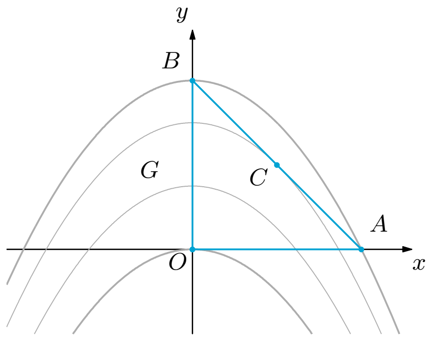


Figure 7.11. The level curves and the admissible set of Example 7.15

Example 7.15

Consider $f(x, y) = x^2 + y - 1$ and let G be the perimeter of the triangle with vertices $O = (0, 0)$, $A = (1, 0)$, $B = (0, 1)$, see Fig. 7.11. We want to find the minima and maxima of f on G, which exist by compactness. Parametrise the three sides OA, OB, AB by $\boldsymbol{\gamma}_1(t) = (t, 0)$, $\boldsymbol{\gamma}_2(t) = (0, t)$, $\boldsymbol{\gamma}_3(t) = (t, 1 - t)$, where $0 \le t \le 1$ for all three. The map $f_{|OA}(t) = t^2 - 1$ has minimum at O and maximum at A; $f_{|OB}(t) = t - 1$ has minimum at O and maximum at B, and $f_{|AB}(t) = t^2 - t$ has minimum at $C = \left(\frac{1}{2}, \frac{1}{2}\right)$ and maximum at A and B. Since $f(O) = -1$, $f(A) = f(B) = 0$ and $f(C) = -\frac{1}{4}$, the function reaches its minimum value at O and its maximum at A and B. □

7.3.2 Lagrange multipliers

The idea behind Lagrange multipliers is a particular feature of any regular constrained extremum point, shown in Fig. 7.12.

Proposition 7.16 *Let $\boldsymbol{x}_0 \in G$ be a regular point for g. If $\boldsymbol{x}_0$ is an extremum for f constrained to G, there exists a unique constant $\lambda_0 \in \mathbb{R}$, called* **Lagrange multiplier***, such that*

$$\nabla f(\boldsymbol{x}_0) = \lambda_0 \nabla g(\boldsymbol{x}_0)\,. \tag{7.11}$$

Proof. For simplicity we assume $n = 2$ or 3. In the former case what we have seen in Sect. 7.2.1 applies to g, since $\boldsymbol{x}_0 \in G = L(g, 0)$. Thus a regular simple curve $\boldsymbol{\gamma} : I \to \mathbb{R}^2$ exists, with $\boldsymbol{x}_0 = \boldsymbol{\gamma}(t_0)$ for some t_0 interior to I, such that $g(\boldsymbol{x}) = g\big(\boldsymbol{\gamma}(t)\big) = 0$ on a neighbourhood of $\boldsymbol{x}_0$; additionally,

$$\nabla g(\boldsymbol{x}_0) \cdot \boldsymbol{\gamma}'(t_0) = 0\,.$$

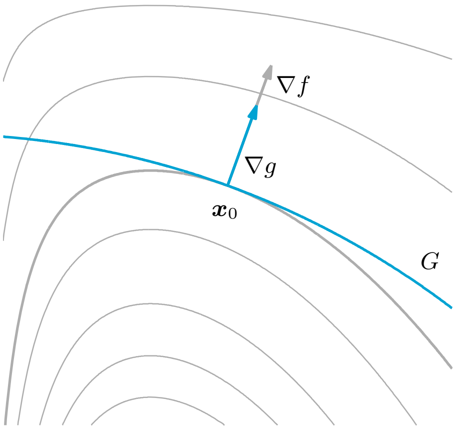


Figure 7.12. At a constrained extremum the gradients of f and g are parallel

But by assumption $t \mapsto f\big(\boldsymbol{\gamma}(t)\big)$ has an extremum at t_0, which is stationary by Fermat's Theorem. Consequently

$$\frac{\mathrm{d}}{\mathrm{d}t} f\big(\boldsymbol{\gamma}(t)\big)_{|t=t_0} = \nabla f(\boldsymbol{x}_0) \cdot \boldsymbol{\gamma}'(t_0) = 0 \,.$$

As $\nabla f(\boldsymbol{x}_0)$ and $\nabla g(\boldsymbol{x}_0)$ are both orthogonal to the same vector $\boldsymbol{\gamma}'(t_0) \neq 0$ ($\boldsymbol{\gamma}$ being regular), they are parallel, i.e., (7.11) holds. The uniqueness of λ_0 follows from $\nabla g(\boldsymbol{x}_0) \neq 0$.
The proof in three dimensions is completely similar, because on the one hand $g\big(\boldsymbol{\sigma}(u,v)\big) = 0$ around $\boldsymbol{x}_0$ for a suitable regular and simple surface $\boldsymbol{\sigma} : A \to \mathbb{R}^3$ (Sect. 7.2.2); on the other the map $f\big(\boldsymbol{\sigma}(u,v)\big)$ has a relative extremum at (u_0, v_0), interior to A and such that $\boldsymbol{\sigma}(u_0,v_0) = \boldsymbol{x}_0$. At the same time then

$$\nabla g(\boldsymbol{x}_0) \boldsymbol{J\sigma}(u_0,v_0) = 0 \,, \qquad \text{and} \qquad \nabla f(\boldsymbol{x}_0) \boldsymbol{J\sigma}(u_0,v_0) = 0 \,.$$

As $\boldsymbol{\sigma}$ is regular, the column vectors of $\boldsymbol{J\sigma}(u_0,v_0)$ are linearly independent and span the tangent plane to $\boldsymbol{\sigma}$ at $\boldsymbol{x}_0$. The vectors $\nabla f(\boldsymbol{x}_0)$ and $\nabla g(\boldsymbol{x}_0)$ are both perpendicular to the plane, and so parallel. □

This is the right place to stress that (7.11) can be fulfilled, with $g(\boldsymbol{x}) = 0$, also by a point $\boldsymbol{x}_0$ which is not a constrained extremum. That is because the two conditions are necessary yet *not sufficient* for the existence of a constrained extremum. For example, $f(x,y) = x - y^5$ and $g(x,y) = x - y^3$ satisfy $g(0,0) = 0$, $\nabla f(0,0) = \nabla g(0,0) = (1,0)$; nevertheless, f restricted to $G = \{(x,y) \in \mathbb{R}^2 : x = y^3\}$ has neither a minimum nor a maximum at the origin, for it is given by $y \mapsto f(y^3, y) = y^3 - y^5$ ($y_0 = 0$ is a horizontal inflection point).

There is an equivalent formulation for the previous proposition, that associates to $\boldsymbol{x}_0$ an unconstrained stationary point relative to a new function depending on f and g.

Definition 7.17 *Set* $\Omega = \operatorname{dom} f \cap \operatorname{dom} g \subseteq \mathbb{R}^n$. *The function* $\mathcal{L} : \Omega \times \mathbb{R} \to \mathbb{R}$ *defined by*

$$\mathcal{L}(\boldsymbol{x}, \lambda) = f(\boldsymbol{x}) - \lambda g(\boldsymbol{x})$$

is said **Lagrangian** (**function**) **of** f **constrained to** g.

The gradient of $\mathcal{L}$ looks as follows

$$\nabla_{(\boldsymbol{x},\lambda)}\mathcal{L}(\boldsymbol{x}, \lambda) = \Big(\nabla_{\boldsymbol{x}}\mathcal{L}(\boldsymbol{x}, \lambda), \frac{\partial \mathcal{L}}{\partial \lambda}(\boldsymbol{x}, \lambda)\Big) = \big(\nabla f(\boldsymbol{x}) - \lambda \nabla g(\boldsymbol{x}), g(\boldsymbol{x})\big) .$$

Hence the condition $\nabla_{(\boldsymbol{x},\lambda)}\mathcal{L}(\boldsymbol{x}_0, \lambda_0) = \mathbf{0}$, expressing that $(\boldsymbol{x}_0, \lambda_0)$ is stationary for $\mathcal{L}$, is equivalent to the system

$$\boxed{\begin{cases} \nabla f(\boldsymbol{x}_0) = \lambda_0 \nabla g(\boldsymbol{x}_0) , \\ g(\boldsymbol{x}_0) = 0 . \end{cases}}$$

Proposition 7.16 ensures that each regular extremum point of f constrained by g determines a unique stationary point for the Lagrangian $\mathcal{L}$.

Under this new light, the procedure breaks into the following steps.

i) Write the system of $n+1$ equations in $n+1$ unknowns $\boldsymbol{x} = (x_1, \ldots, x_n)$ and λ

$$\begin{cases} \nabla f(\boldsymbol{x}) = \lambda \nabla g(\boldsymbol{x}) \\ g(\boldsymbol{x}) = 0 \end{cases} \tag{7.12}$$

and solve it; as the system is not, in general, linear, there may be a number of distinct solutions.

ii) For each solution $(\boldsymbol{x}_0, \lambda_0)$ found, decide whether $\boldsymbol{x}_0$ is a constrained extremum for f, often with ad hoc arguments.
In general, after (7.12) has been solved, the multiplier λ_0 stops being useful. If G is compact for instance, Weierstrass' Theorem 5.24 guarantees the existence of an absolute minimum and maximum of $f_{|G}$ (distinct if f is not constant on G); therefore, assuming G only consists of regular points for g, these two must be among the ones found previously; comparing values will permit us to pin down minima and maxima.

iii) In presence of non-regular points (stationary) for g in G (or points where f and/or g are not differentiable) we are forced to proceed case by case, lacking a general procedure.

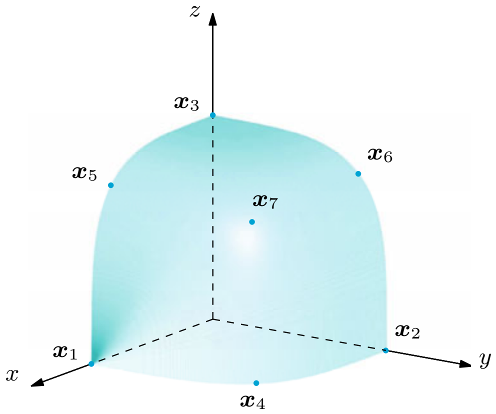


Figure 7.13. The graph of the admissible set G (first octant only) and the stationary points of f on G (Example 7.18)

Example 7.18

Let us find the points of $\mathbb{R}^3$ lying on the manifold defined by $x^4 + y^4 + z^4 = 1$ with smallest and largest distance from the origin (see Fig. 7.13).

The problem consists in extremising the function $f(x, y, z) = \|\boldsymbol{x}\|^2 = x^2+y^2+z^2$ on the set G defined by $g(x, y, z) = x^4 + y^4 + z^4 - 1 = 0$. We consider the square distance, rather than the distance $\|\boldsymbol{x}\| = \sqrt{x^2 + y^2 + z^2}$, because the two have the same extrema, but the former has the advantage of simplifying computations. We use Lagrange multipliers, and consider the system (7.12):

$$\begin{cases} 2x = \lambda 4x^3 \\ 2y = \lambda 4y^3 \\ 2z = \lambda 4z^3 \\ x^4 + y^4 + z^4 - 1 = 0\,. \end{cases}$$

As f is invariant under sign change in its arguments, $f(\pm x, \pm y, \pm z) = f(x, y, z)$, and similarly for g, we can just look for solutions belonging in the first octant ($x \geq 0$, $y \geq 0$, $z \geq 0$). The first three equations are solved by

$$x = 0 \text{ or } x = \frac{1}{\sqrt{2\lambda}}\,, \qquad y = 0 \text{ or } y = \frac{1}{\sqrt{2\lambda}}\,, \qquad z = 0 \text{ or } z = \frac{1}{\sqrt{2\lambda}}\,,$$

combined in all possible ways. The point $(x, y, z) = (0, 0, 0)$ is to be excluded because it fails to satisfy the last equation. The choices $\big(\frac{1}{\sqrt{2\lambda}}, 0, 0\big)$, $\big(0, \frac{1}{\sqrt{2\lambda}}, 0\big)$ or $\big(0, 0, \frac{1}{\sqrt{2\lambda}}\big)$ force $\frac{1}{4\lambda^2} = 1$, hence $\lambda = \frac{1}{2}$ (since $\lambda > 0$). Similarly, $(x, y, z) = \big(\frac{1}{\sqrt{2\lambda}}, \frac{1}{\sqrt{2\lambda}}, 0\big)$, $\big(\frac{1}{\sqrt{2\lambda}}, 0, \frac{1}{\sqrt{2\lambda}}\big)$ or $\big(0, \frac{1}{\sqrt{2\lambda}}, \frac{1}{\sqrt{2\lambda}}\big)$ satisfy the fourth equation if $\lambda = \frac{\sqrt{2}}{2}$, while $\big(\frac{1}{\sqrt{2\lambda}}, \frac{1}{\sqrt{2\lambda}}, \frac{1}{\sqrt{2\lambda}}\big)$ fulfills it if $\lambda = \frac{\sqrt{3}}{2}$. The solutions then are:

$$\boldsymbol{x}_1 = (1, 0, 0)\,, \;\; f(\boldsymbol{x}_1) = 1\,; \qquad \boldsymbol{x}_4 = \big(\tfrac{1}{\sqrt[4]{2}}, \tfrac{1}{\sqrt[4]{2}}, 0\big)\,, \;\; f(\boldsymbol{x}_4) = \sqrt{2}\,;$$

$$\boldsymbol{x}_2 = (0, 1, 0)\,, \;\; f(\boldsymbol{x}_2) = 1\,; \qquad \boldsymbol{x}_5 = \big(\tfrac{1}{\sqrt[4]{2}}, 0, \tfrac{1}{\sqrt[4]{2}}\big)\,, \;\; f(\boldsymbol{x}_5) = \sqrt{2}\,;$$

$\boldsymbol{x}_3 = (0,0,1)\,, \qquad f(\boldsymbol{x}_3) = 1\,; \qquad \boldsymbol{x}_6 = \big(0, \frac{1}{\sqrt[4]{2}}, \frac{1}{\sqrt[4]{2}}\big)\,, \ f(\boldsymbol{x}_6) = \sqrt{2}\,;$

$\boldsymbol{x}_7 = \big(\frac{1}{\sqrt[4]{3}}, \frac{1}{\sqrt[4]{3}}, \frac{1}{\sqrt[4]{3}}\big)\,, \ f(\boldsymbol{x}_7) = \sqrt{3}\,.$

In conclusion, the first octant contains 3 points of G, $\boldsymbol{x}_1$, $\boldsymbol{x}_2$, $\boldsymbol{x}_3$, with the shortest distance to the origin, and 1 farthest point $\boldsymbol{x}_7$. The distance function on G has stationary points $\boldsymbol{x}_4$, $\boldsymbol{x}_5$, $\boldsymbol{x}_6$, but no minimum nor maximum. □

Now we pass to briefly consider the case of an admissible set G defined by $m < n$ equalities, of the type $g_i(\boldsymbol{x}) = 0$, $1 \le i \le m$. If $\boldsymbol{x}_0 \in G$ is a constrained extremum for f on G and regular for each g_i, the analogue to Proposition 7.16 tells us there exist m constants λ_{0i}, $1 \le i \le m$ (the Lagrange multipliers), such that

$$\nabla f(\boldsymbol{x}_0) = \sum_{i=1}^{m} \lambda_{0i} \nabla g_i(\boldsymbol{x}_0)\,.$$

Equivalently, set $\boldsymbol{\lambda} = (\lambda_i)_{i=1,\ldots,m}) \in \mathbb{R}^m$ and $\boldsymbol{g}(\boldsymbol{x}) = \big(g_1(\boldsymbol{x})\big)_{i=1,\ldots,m}$: then the Lagrangian

$$\boxed{\mathcal{L}(\boldsymbol{x}, \boldsymbol{\lambda}) = f(\boldsymbol{x}) - \boldsymbol{\lambda} \cdot \boldsymbol{g}(\boldsymbol{x})} \tag{7.13}$$

admits a stationary point $(\boldsymbol{x}_0, \boldsymbol{\lambda}_0)$. To find such points we have to write the system of $n+m$ equations in the $n+m$ unknowns given by the components of $\boldsymbol{x}$ and $\boldsymbol{\lambda}$,

$$\boxed{\begin{cases} \nabla f(\boldsymbol{x}) = \displaystyle\sum_{i=1}^{m} \lambda_i \nabla g_i(\boldsymbol{x})\,, & \\ g_i(\boldsymbol{x}) = 0\,, & 1 \le i \le m\,, \end{cases}}$$

which generalises (7.12).

Example 7.19

We want to find the extrema of $f(x,y,z) = 3x + 3y + 8z$ constrained to the intersection of two cylinders, $x^2 + z^2 = 1$ and $y^2 + z^2 = 1$. Define

$$g_1(x,y,z) = x^2 + z^2 - 1 \quad \text{and} \quad g_2(x,y,z) = y^2 + z^2 - 1$$

so that the admissible set is $G = G_1 \cap G_2$, $G_i = L(g_i, 0)$. Each point of G_i is regular for g_i, so the same is true for G. Moreover G is compact, so f certainly has minimum and maximum on it.

As $\nabla f = (3,3,8)$, $\nabla g_1 = (2x, 0, 2z)$, $\nabla g_2 = (0, 2y, 2z)$, system (7.13) reads

$$\begin{cases} 3 = \lambda_1 2x \\ 3 = \lambda_2 2y \\ 8 = (\lambda_1 + \lambda_2) 2z \\ x^2 + z^2 - 1 = 0 \\ y^2 + z^2 - 1 = 0\,; \end{cases}$$

the first three equations tell $x = \frac{3}{2\lambda_1}$, $y = \frac{3}{2\lambda_2}$, $z = \frac{4}{\lambda_1+\lambda_2}$, (with $\lambda_1 \neq 0$, $\lambda_2 \neq 0$, $\lambda_1 + \lambda_2 \neq 0$), so that the remaining two give

$$\lambda_1 = \lambda_2 = \pm\frac{5}{2}\,.$$

Then, setting $\boldsymbol{x}_0 = \left(\frac{3}{5}, \frac{3}{5}, \frac{4}{5}\right)$ and $\boldsymbol{x}_1 = -\boldsymbol{x}_0$, we have $f(\boldsymbol{x}_0) > 0$ and $f(\boldsymbol{x}_1) < 0$. We conclude that $\boldsymbol{x}_0$ is an absolute constrained maximum point, while $\boldsymbol{x}_1$ is an absolute constrained minimum. □

Finally, let us assess the case in which the set G is defined by m inequalities; without loss of generality we may assume the constraints are of type $g_i(\boldsymbol{x}) \geq 0$, $i = 1, \ldots, m$. So first we examine interior points of G, with the techniques of Sect. 5.6. Secondly, we look at the boundary ∂G of G, indeed at points where at least one inequality is actually an equality. This generates a constrained extremum problem on ∂G, which we know how to handle.

Due to the profusion of situations, we just describe a few possibilities using examples.

Examples 7.20

i) Let us return to Example 7.15, and suppose we want to find the extrema of f on the whole triangle G of vertices O, A, B.
Set $g_1(x, y) = x$, $g_2(x, y) = y$, $g_3(x, y) = 1 - x - y$, so that $G = \{(x, y) \in \mathbb{R}^2 : g_i(x, y) \geq 0,\ i = 1, 2, 3\}$. There are no extrema on the interior of G, since $\nabla f(\boldsymbol{x}) = (2x, 1) \neq 0$ precludes the existence of stationary points. The extrema of f on G, which have to be present by the compactness of G, must then belong to the boundary, and are those we already know of from Example 7.15.

ii) We look for the extrema of f of Example 7.18 that belong to G defined by $x^4 + y^4 + z^4 \leq 1$. As $\nabla f(\boldsymbol{x}) = 2\boldsymbol{x}$, the only extremum interior to G is the origin, where f reaches the absolute minimum. But G is compact, so there must also be an absolute maximum somewhere on the boundary ∂G. The extrema constrained to the latter are exactly those of the example, so we conclude that f is maximised on G by $\boldsymbol{x}_7 = \left(\frac{1}{\sqrt[4]{3}}, \frac{1}{\sqrt[4]{3}}, \frac{1}{\sqrt[4]{3}}\right)$, and at the other seven points obtained from this by flipping any sign.

iii) We determine the extrema of $f(x, y) = (x + y)\mathrm{e}^{-(x^2+y^2)}$ subject to $g(x, y) = 2x + y \geq 0$. Note immediately that $f(x, y) \to 0$ as $\|\boldsymbol{x}\| \to \infty$, so f must admit absolute maximum and minimum on G. The interior points where the constraint holds form the (open) half-space $y > -2x$. Since

$$\nabla f(\boldsymbol{x}) = \mathrm{e}^{-(x^2+y^2)}\big(1 - 2x(x + y), 1 - 2y(x + y)\big)\,,$$

f has stationary points $\boldsymbol{x} = \pm\left(\frac{1}{2}, \frac{1}{2}\right)$; the only one of these inside G is $\boldsymbol{x}_0 = \left(\frac{1}{2}, \frac{1}{2}\right)$, for which $\boldsymbol{H}f(\boldsymbol{x}_0) = -\mathrm{e}^{-1/2}\begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$. The Hessian matrix tells us that $\boldsymbol{x}_0$ is a relative maximum for f. On the boundary ∂G, where $y = -2x$, the composite map

$$\varphi(x) = f(x, -2x) = -x\mathrm{e}^{-5x^2}$$

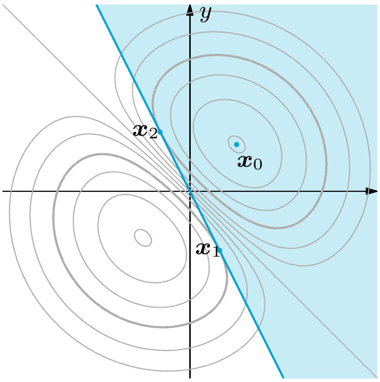


Figure 7.14. The level curves of f and the admissible set G of Example 7.20 iii)

admits absolute maximum at $x = -\frac{1}{\sqrt{10}}$ and minimum at $x = \frac{1}{\sqrt{10}}$. Setting $\boldsymbol{x}_1 = \big(\frac{1}{\sqrt{10}}, -\frac{2}{\sqrt{10}}\big)$ and $\boldsymbol{x}_2 = \big(-\frac{1}{\sqrt{10}}, \frac{2}{\sqrt{10}}\big)$, we can without doubt say $\boldsymbol{x}_1$ is the unique absolute minimum on G. For the absolute maxima, we need to compare the values attained at $\boldsymbol{x}_0$ and $\boldsymbol{x}_2$. But since $f(\boldsymbol{x}_0) = \mathrm{e}^{-1/2}$ and $f(\boldsymbol{x}_2) = \frac{1}{\sqrt{10}}\mathrm{e}^{-1/2}$, $\boldsymbol{x}_0$ is the only absolute maximum point on G (Fig. 7.14).

iv) Consider $f(x,y) = x + 2y$ on the set G defined by

$$x + 2y + 8 \geq 0\,, \qquad 5x + y + 13 \geq 0\,, \qquad x - 4y + 11 \geq 0\,,$$
$$2x + y - 5 \leq 0\,, \qquad 5x - 2y - 8 \leq 0\,.$$

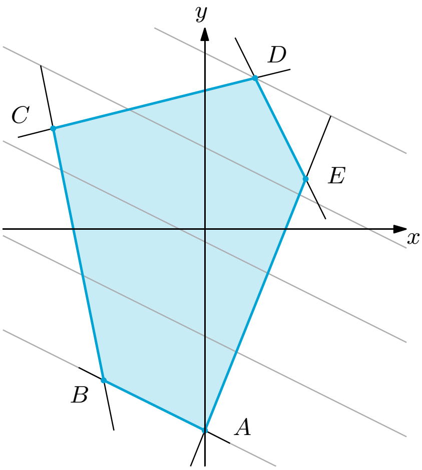


Figure 7.15. Level curves of f and admissible set G of Example 7.20 iv)

The set G is the (irregular) pentagon of Fig. 7.15, having vertices $A=(0,-4)$, $B=(-2,-3)$, $C=(-3,2)$, $D=(1,3)$, $E=(2,1)$ and obtained as intersection of the five half-planes defined by the above inequalities. As $\nabla f(\boldsymbol{x})=(1,2)\neq \mathbf{0}$, the function f attains minimum and maximum on the boundary of G; and since f is linear on the perimeter, any extremum point must be a vertex. Thus it is enough to compare the values at the corners

$$f(A)=f(B)=-8\,,\qquad f(C)=1\,,\qquad f(D)=7\,,\quad f(E)=4\,;$$

f restricted to G is smallest at each point of AB and largest at D. This simple example is the typical problem dealt with by Linear Programming, a series of methods to find extrema of linear maps subject to constraints given by linear inequalities. Linear Programming is relevant in many branches of Mathematics, like Optimization and Operations Research. The reader should refer to the specific literature for further information on the matter. □

7.4 Exercises

1. *Supposing you are able to write $x^3y+xy^4=2$ in the form $y=\varphi(x)$, compute $\varphi'(x)$. Determine a point x_0 around which this is feasible.*

2. *The equation $x^2+y^3z=\frac{xz}{y}$ admits the solution $(2,1,4)$. Verify it can be written as $y=\varphi(x,z)$, with φ defined on a neighbourhood of $(2,4)$. Compute the partial derivatives of φ at $(2,4)$.*

3. *Check that*

$$e^{x-y}+x^2-y^2-e(x+1)+1=0$$

defines a function $y=\varphi(x)$ on a neighbourhood of $x_0=0$. What is the nature of the point x_0 for φ?

4. *Verify that the equation*

$$x^2+2x+e^y+y-2z^3=0$$

gives a map $y=\varphi(x,z)$ defined around $P=(x_0,z_0)=(-1,0)$. Such map defines a surface, of which the tangent plane at the point P should be determined.

5. a) *Verify that around $P_0=(0,2,-1)$ the equation*

$$x\log(y+z)+3(y-2)z+\sin x=0$$

defines a regular simple surface Σ.
b) *Determine the tangent plane to Σ at P_0, and the unit normal forming with $\boldsymbol{v}=4\boldsymbol{i}+2\boldsymbol{j}-5\boldsymbol{k}$ an acute angle.*

$\boxed{6.}$ *Check that the system*

$$\begin{cases} 3x - \cos y + y + e^z = 0 \\ x - e^x - y + z + 1 = 0 \end{cases}$$

yields, around the origin, a curve in space of equations

$$\gamma(t) = \big(t, \varphi_1(t), \varphi_2(t)\big)\,, \qquad t \in I(0)\,.$$

Write the equation of the tangent to the curve at the origin.

$\boxed{7.}$ *Verify that*

$$y^7 + 3y - 2xe^{3x} = 0$$

defines a function $y = \varphi(x)$ *for any* $x \in \mathbb{R}$. *Study* φ *and sketch its graph.*

8. *Represent the following maps' level curves:*

$\boxed{\text{a)}}$ $f(x,y) = 6 - 3x - 2y$ $\boxed{\text{b)}}$ $f(x,y) = 4x^2 + y^2$

c) $f(x,y) = 2xy$ d) $f(x,y) = 2y - 3\log x$

e) $f(x,y) = \sqrt{4x + 3y}$ f) $f(x,y) = 4x - 2y^2$

$\boxed{9.}$ *Match the functions below to the graphs A–F of Fig. 7.16 and to the level curves I–VI of Fig. 7.17:*

a) $z = \cos\sqrt{x^2 + 2y^2}$ b) $z = (x^2 - y^2)e^{-x^2-y^2}$

c) $z = \dfrac{15}{9x^2 + y^2 + 1}$ d) $z = x^3 - 3xy^2$

e) $z = \cos x \,\sin 2y$ f) $z = 6\cos^2 x - \dfrac{1}{10}x^2$

$\boxed{10.}$ *Describe the level surfaces of the following maps:*

$\boxed{\text{a)}}$ $f(x,y,z) = x + 3y + 5z$ $\boxed{\text{b)}}$ $f(x,y,z) = x^2 - y^2 + z^2$

$\boxed{11.}$ *Determine the maximum and minimum points of*

$$f(x,y) = x^2 + y^2 + \frac{3}{2}x + 1$$

on the set $G = \{(x,y) \in \mathbb{R}^2 : 4x^2 + y^2 - 1 = 0\}$.

$\boxed{12.}$ *Find maximum and minimum points for* $f(x,y) = x + 3y + 2$ *on the compact set* $G = \{(x,y) \in \mathbb{R}^2 : x, y \geq 0,\; x^2 + y^2 \leq 1\}$.

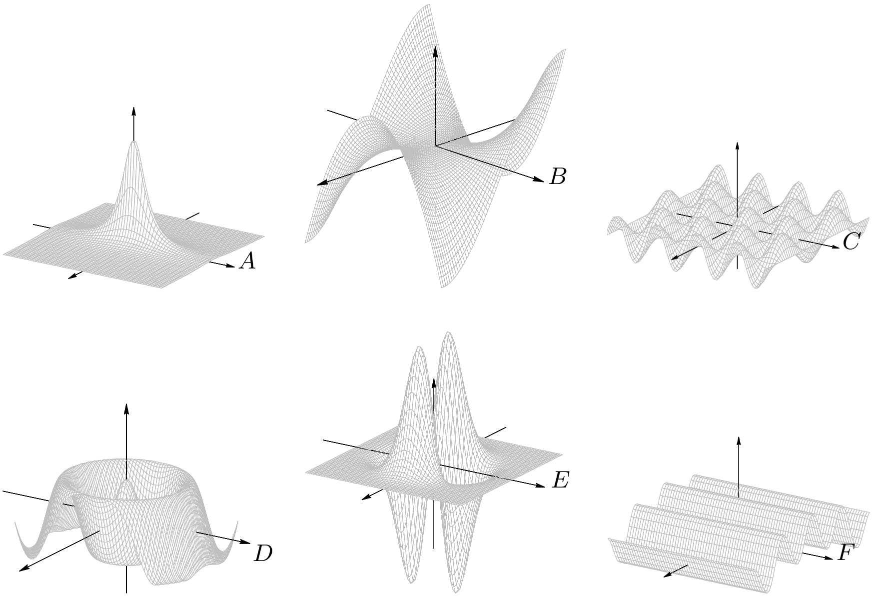


Figure 7.16. The graphs of Exercise 9

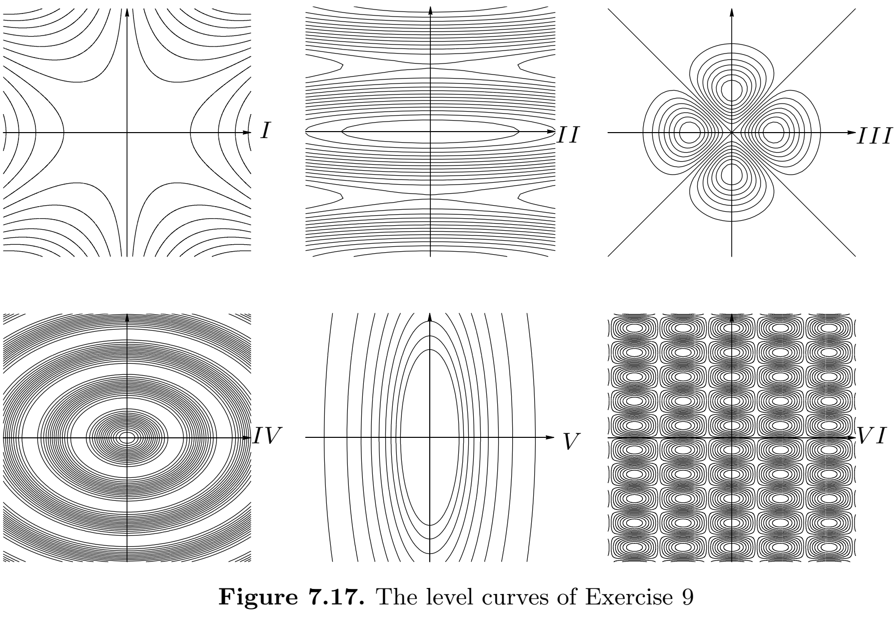


Figure 7.17. The level curves of Exercise 9

13. *Consider the function*

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

a) Find its stationary points and describe their type.
b) Compute the map's absolute minimum and maximum on the set

$$G = \{(x,y) \in \mathbb{R}^2 : \sqrt{1+x^2} \le y \le 2\}\,.$$

14. *Determine the extrema of $f(x,y) = 2x^2 + y^2$ constrained to*

$$G = \{(x,y) \in \mathbb{R}^2 : x^4 - x^2 + y^2 - 5 = 0\}\,.$$

15. *What are the absolute minimum and maximum of $f(x,y) = 4x^2 + y^2 - 2x - 4y + 1$ on*

$$G = \{(x,y) \in \mathbb{R}^2 : 4x^2 + y^2 - 1 = 0\}\,?$$

7.4.1 Solutions

1. Set $f(x,y) = x^3y + xy^4 - 2$; then

$$f_x(x,y) = 3x^2y + y^4 \qquad \text{and} \qquad f_y(x,y) = x^3 + 4xy^3\,,$$

so

$$\varphi'(x) = -\frac{3x^2\varphi(x) + \big(\varphi(x)\big)^4}{x^3 + 4x\big(\varphi(x)\big)^3}$$

for any $x \neq 0$ with $x^2 \neq -4y^3$.

For instance, the point $(x_0, y_0) = (1,1)$ solves the equation, thus there is a map $y = \varphi(x)$ defined around $x_0 = 1$ because $f_y(1,1) = 5 \neq 0$. Moreover, $\varphi'(1) = -4/5$.

2. Setting $f(x,y,z) = x^2 + y^3z - \frac{xz}{y}$, we have

$$\begin{aligned}
f_x(x,y,z) &= 2x - \frac{z}{y} && \text{with } f_x(2,1,4) = 0\,,\\
f_y(x,y,z) &= 3y^2z + \frac{xz}{y^2} && \text{with } f_y(2,1,4) = 20 \neq 0\,,\\
f_z(x,y,z) &= y^3 - \frac{x}{y} && \text{with } f_z(2,1,4) = -1 \neq 0\,.
\end{aligned}$$

Due to Theorem 7.4 we can solve for y, hence express y as a function of x and z, around $(2,4)$; otherwise said, there exists $y = \varphi(x,z)$ such that

$$\frac{\partial\varphi}{\partial x}(2,4) = 0\,, \qquad \frac{\partial\varphi}{\partial z}(2,4) = \frac{1}{20}\,.$$

3. Call $f(x,y) = \mathrm{e}^{x-y} + x^2 - y^2 - \mathrm{e}(x+1) + 1$ and notice $(0,-1)$ is a solution. Then

$$f_x(x,y) = \mathrm{e}^{x-y} + 2x - \mathrm{e}\,, \qquad f_y(x,y) = -\mathrm{e}^{x-y} - 2y$$

with $f_x(0,-1) = 0$, $f_y(0,-1) = 2 - \mathrm{e} \neq 0$. Theorem 7.1 guarantees the existence of a map $y = \varphi(x)$, defined around the origin, such that

$$\varphi'(0) = -\frac{f_x(0,-1)}{f_y(0,-1)} = 0\,.$$

Hence $x_0 = 0$ is a critical point of φ.

4. If we set

$$f(x,y,z) = x^2 + 2x + \mathrm{e}^y + y - 2z^3$$

then $f(-1,0,0) = 0$. We have

$$\begin{aligned} f_x(x,y,z) &= 2x + 2 &&\text{with } f_x(-1,0,0) = 0\,,\\ f_y(x,y,z) &= \mathrm{e}^y + 1 &&\text{with } f_y(-1,0,0) = 2 \neq 0\,,\\ f_z(x,y,z) &= -6z^2 &&\text{with } f_z(-1,0,0) = 0\,. \end{aligned}$$

By Theorem 7.4 there is a map $y = \varphi(x,z)$ around $(-1,0)$ satisfying

$$\frac{\partial\varphi}{\partial x}(-1,0) = \frac{\partial\varphi}{\partial z}(-1,0) = 0\,.$$

The tangent plane is therefore

$$y = \varphi(-1,0) + \varphi_x(-1,0)(x+1) + \varphi_z(-1,0)(z-0) = 0\,.$$

5. a) The gradient of $f(x,y,z) = x\log(y+z) + 3(y-2)z + \sin x$ is

$$\nabla f(x,y,z) = \Big(\log(y+z) + \cos x, \frac{x}{y+z} + 3z, \frac{x}{y+z} + 3(y-2)\Big)\,,$$

so $\nabla f(P_0) = (1,-3,0) \neq \mathbf{0}$. By Theorem 7.4 then, we can express x via y and z, or y in terms of x, z. Therefore, around P_0 the surface Σ is locally a regular simple graph.

b) The tangent plane at $P_0 = \boldsymbol{x}_0$ is, recalling Proposition 7.9,

$$\nabla f(\boldsymbol{x}_0)\cdot(\boldsymbol{x} - \boldsymbol{x}_0) = x - 3(y-2) = 0\,,$$

so $x - 3y = -6$. The unit normal at P_0 will be $\boldsymbol{\nu} = \frac{\sigma}{\sqrt{10}}(\boldsymbol{i} - 3\boldsymbol{j})$, with $\sigma \in \{\pm 1\}$ defined by $\boldsymbol{\nu}\cdot\boldsymbol{v} = -\frac{2\sigma}{\sqrt{10}} > 0$, whence $\sigma = -1$.

6. Referring to Theorem 7.5 and Example 7.6, let us set

$$\begin{cases} f_1(x,y,z) = 3x - \cos y + y + \mathrm{e}^z \\ f_2(x,y,z) = x - \mathrm{e}^x - y + z + 1\,. \end{cases}$$

Then

$$\frac{\partial f_1}{\partial y}(x,y,z) = \sin y + 1\,, \qquad \frac{\partial f_1}{\partial z}(x,y,z) = \mathrm{e}^z\,,$$
$$\frac{\partial f_2}{\partial y}(x,y,z) = -1\,, \qquad \frac{\partial f_2}{\partial z}(x,y,z) = 1\,.$$

Now consider the matrix

$$\begin{pmatrix} \frac{\partial f_1}{\partial y}(\mathbf{0}) & \frac{\partial f_1}{\partial z}(\mathbf{0}) \\ \frac{\partial f_2}{\partial y}(\mathbf{0}) & \frac{\partial f_2}{\partial z}(\mathbf{0}) \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix};$$

this being non-singular, it defines around $t = 0$ a curve $\boldsymbol{\gamma}(t) = \big(t, \varphi_1(t), \varphi_2(t)\big)$.

For the tangent line we need to compute $\boldsymbol{\gamma}'(t) = \big(1, \varphi_1'(t), \varphi_2'(t)\big)$, and in particular $\boldsymbol{\gamma}'(0) = \big(1, \varphi_1'(0), \varphi_2'(0)\big)$. But

$$\begin{pmatrix} \frac{\partial f_1}{\partial y}(\mathbf{0}) & \frac{\partial f_1}{\partial z}(\mathbf{0}) \\ \frac{\partial f_2}{\partial y}(\mathbf{0}) & \frac{\partial f_2}{\partial z}(\mathbf{0}) \end{pmatrix} \begin{pmatrix} \varphi_1'(0) \\ \varphi_2'(0) \end{pmatrix} = - \begin{pmatrix} \frac{\partial f_1}{\partial x}(\mathbf{0}) \\ \frac{\partial f_2}{\partial x}(\mathbf{0}) \end{pmatrix}$$

is

$$\begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} \varphi_1'(0) \\ \varphi_2'(0) \end{pmatrix} = - \begin{pmatrix} 3 \\ 0 \end{pmatrix},$$

so

$$\begin{cases} \varphi_1'(0) + \varphi_2'(0) = -3 \\ \varphi_1'(0) - \varphi_2'(0) = 0 \end{cases},$$

solved by $\varphi_1'(0) = \varphi_2'(0) = -3/2$. In conclusion, the tangent line is

$$T(t) = \boldsymbol{\gamma}(0) + \boldsymbol{\gamma}'(0)t = \big(1, -\frac{3}{2}, -\frac{3}{2}\big)t = \big(t, -\frac{3}{2}t, -\frac{3}{2}t\big)\,.$$

7. The map $f(x,y) = y^7 + 3y - 2x\mathrm{e}^{3x}$ has derivatives

$$f_x(x,y) = -2\mathrm{e}^{3x}(1+3x)\,, \qquad f_y(x,y) = 7y^6 + 3 > 0\,.$$

Theorem 7.1 says we can have y as a function of x around every point of $\mathbb{R}^2$, so let $y = \varphi(x)$ be such a map. Then

$$\varphi'(x) = -\frac{f_x(x,y)}{f_y(x,y)} = \frac{2\mathrm{e}^{3x}(1+3x)}{7y^6+3}\,,$$

and $\varphi'(x) = 0$ for $x = -1/3$, $\varphi'(x) > 0$ for $x > -1/3$. Observe that $f(0,0) = 0$, so $\varphi(0) = 0$. The function passes through the origin, is increasing when $x > -1/3$, decreasing when $x < -1/3$. The point $x = -1/3$ is an (absolute) minimum for φ, with $\varphi(-1/3) < 0$.

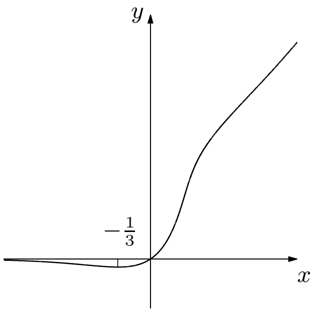


Figure 7.18. The graph of the implicit function of Exercise 7

To compute the limits as $x \to \pm\infty$, note

$$\big(\varphi(x)\big)^7 + 3\big(\varphi(x)\big) = 2x\mathrm{e}^{3x}$$

and

$$\lim_{x\to-\infty} 2x\mathrm{e}^{3x} = 0\,, \qquad \lim_{x\to+\infty} 2x\mathrm{e}^{3x} = +\infty\,.$$

Let $\ell = \lim_{x\to-\infty} \varphi(x)$, so that

$$\ell^7 + 3\ell = \ell(\ell^6+3) = 0 \qquad \text{and hence} \qquad \ell = 0\,.$$

Then call $m = \lim_{x\to+\infty} \varphi(x)$; necessarily $m = +\infty$, for otherwise we would have $m^7 + 3m = +\infty$, a contradiction.

In summary,

$$\lim_{x\to-\infty} \varphi(x) = 0\,, \qquad \lim_{x\to+\infty} \varphi(x) = +\infty\,;$$

the point $x = -1/3$ is an absolute minimum and the graph of f can be seen in Fig. 7.18.

8. *Level curves:*

a) These are the level curves:

$$6 - 3x - 2y = k \qquad \text{i.e.,} \qquad 3x + 2y + k - 6 = 0\,.$$

They form a family of parallel lines with slope $-3/2$, see Fig. 7.19, left.

b) The level curves

$$4x^2 + y^2 = k \qquad \text{i.e.,} \qquad \frac{x^2}{k/4} + \frac{y^2}{k} = 1\,,$$

are, for $k > 0$, a family of ellipses centred at the origin and semi-axes $\sqrt{k}/2$, $\sqrt{k}$. See Fig. 7.19, right.

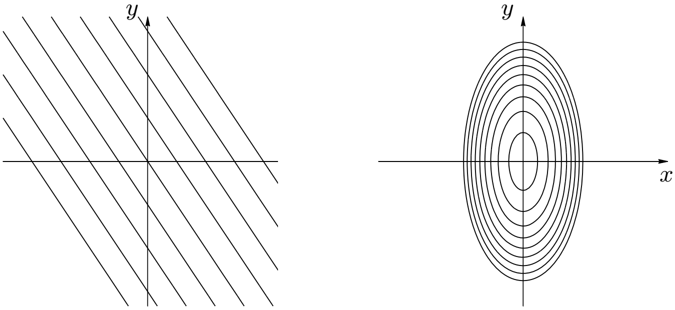


Figure 7.19. Level curves of $f(x, y) = 6 - 3x - 2y$ (left), and of $f(x, y) = 4x^2 + y^2$ (right)

c) See Fig. 7.20, left.
d) See Fig. 7.20, right.
e) See Fig. 7.21, left.
f) See Fig. 7.21, right.

9. The correct matches are : a-D-IV; b-E-III; c-A-V; d-B-I; e-C-VI; f-F-II.

10. *Level surfaces:*

a) We are considering a family of parallel planes of equation

$$x + 3y + 5z - k = 0\,.$$

b) These are hyperboloids (with one or two sheets) with axis on the y-axis.

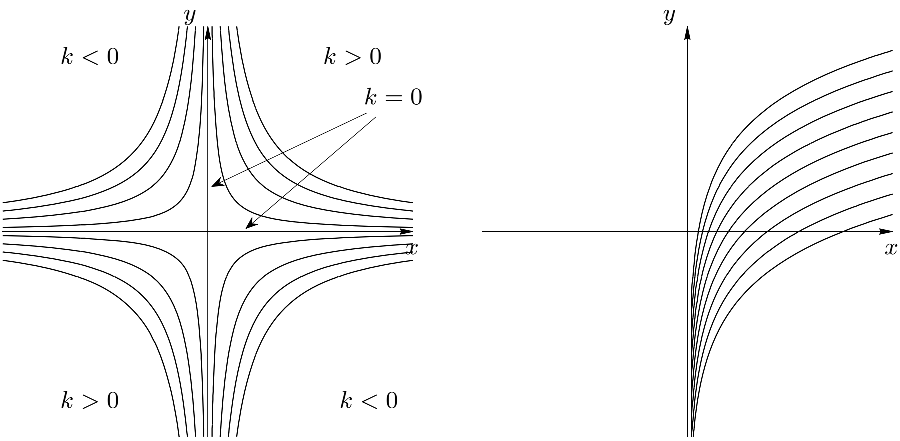


Figure 7.20. Level curves of $f(x, y) = 2xy$ (left), and of $f(x, y) = 2y - 3\log x$ (right)

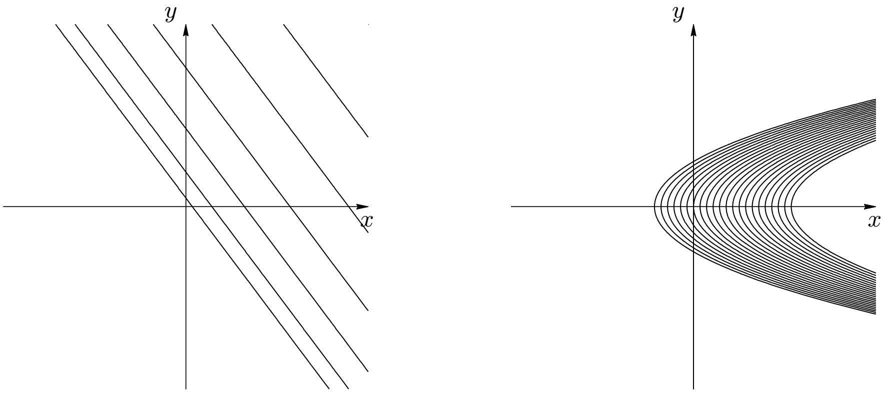


Figure 7.21. Level curves of $f(x,y)=\sqrt{4x+3y}$ (left), and of $f(x,y)=4x-2y^2$ (right)

11. The set G is an ellipse that we can parametrise as $\boldsymbol{\gamma}(t)=\left(\frac{1}{2}\cos t,\sin t\right)$, $t\in[0,2\pi)$. Thus

$$\varphi(t)=f\circ\boldsymbol{\gamma}(t)=\frac{1}{4}\cos^2 t+\sin^2 t+\frac{3}{4}\cos t+1=-\frac{3}{4}\cos^2 t+\frac{3}{4}\cos t+2\,,$$

with

$$\varphi'(t)=\frac{3}{2}\sin t\cos t-\frac{3}{4}\sin t=\frac{3}{2}\sin t\big(\cos t-\frac{1}{2}\big)\,.$$

Note $\varphi'(t)=0$ for $\sin t=0$ or $\cos t=\frac{1}{2}$, so for $t_1=0$, $t_2=\pi$, $t_3=\frac{\pi}{3}$, $t_4=\frac{5}{3}\pi$. Moreover, $\varphi'(t)>0$ for $t\in\left(0,\frac{\pi}{3}\right)\cup\left(\pi,\frac{5}{3}\pi\right)$, hence φ increases on $\left(0,\frac{\pi}{3}\right)$ and $\left(\pi,\frac{5}{3}\pi\right)$, while it decreases on $\left(\frac{\pi}{3},\pi\right)$ and $\left(\frac{5}{3}\pi,2\pi\right)$. Therefore $P_1=\boldsymbol{\gamma}(t_1)=(\frac{1}{2},0)$ and $P_2=\boldsymbol{\gamma}(t_2)=(-\frac{1}{2},0)$ are local minima with values $f(P_1)=2$, $f(P_2)=\frac{1}{2}$; the points $P_3=\boldsymbol{\gamma}(t_3)=(\frac{1}{4},\frac{\sqrt{3}}{2})$, $P_4=\boldsymbol{\gamma}(t_4)=(\frac{1}{4},-\frac{\sqrt{3}}{2})$ are local maxima with $f(P_3)=f(P_4)=\frac{35}{16}$. In particular, the maximum value of f on G is $\frac{35}{16}$, reached at both P_3 and P_4, while the minimum value $1/2$ is attained at P_2.

An alternative way to solve the exercise is to use Lagrange multipliers. If we set $g(x,y)=4x^2+y^2-1$, the Lagrangian is

$$\mathcal{L}(x,y,\lambda)=f(x,y)-\lambda g(x,y)\,.$$

Since $\nabla f(x,y)=(2x+\frac{3}{2},2y)$, $\nabla g(x,y)=(8x,2y)$, we have to solve the system

$$\begin{cases}2x+\frac{3}{2}=8\lambda x\\ 2y=2\lambda y\\ 4x^2+y^2=1\,.\end{cases}$$

From the second equation we get $\lambda=1$ or $y=0$. With $\lambda=1$ we find $x=\frac{1}{4}$, $y=\pm\frac{\sqrt{3}}{2}$ (the points P_3, P_4); taking $y=0$ gives $x=\pm\frac{1}{2}$ (and we obtain P_1 and P_2). Computing f at these points clearly gives the same result of the parameters' method.

12. Since $\nabla f(x,y) = (1,3) \neq (0,0)$ for all (x,y), there are no critical points on the interior of G. We look for extrema on the boundary, which must exist by Weierstrass' Theorem. Now, ∂G decomposes in three pieces: two segments, on the x- and the y-axis, that join the origin to the points $A = (1,0)$ and $B = (0,1)$, and the arc of unit circle connecting A to B. Restricted to the segment OA, the function is $f(x,0) = x+2$, $x \in [0,1]$, so $x = 0$ is a (local) minimum with $f(0,0) = 2$, and $x = 1$ is a (local) maximum with $f(1,0) = 3$. Similarly, on OB the map is $f(0,y) = 3y+2$, $y \in [0,1]$, so $y = 0$ is a (local) minimum, $f(0,0) = 2$, and $y = 1$ a (local) maximum with $f(0,1) = 5$. At last, parametrising the arc by $\boldsymbol{\gamma}(t) = (\cos t, \sin t)$, $t \in [0, \pi/2]$ gives

$$\varphi(t) = f \circ \boldsymbol{\gamma}(t) = \cos t + 3 \sin t + 2\,.$$

Since $\varphi'(t) = -\sin t + 3\cos t$ is zero at $t_0 = \arctan 3$, and $\varphi'(t) > 0$ for $t \in [0, t_0]$, the function f has a (local) maximum at $\boldsymbol{\gamma}(t_0) = (x_0, y_0)$. To compute x_0, y_0 explicitly, observe that

$$\begin{cases} \sin t_0 = 3\cos t_0 \\ \sin^2 t_0 + \cos^2 t_0 = 1\,, \end{cases}$$

whence $9\cos^2 t_0 + \cos^2 t_0 = 9x_0^2 + x_0^2 = 10x_0^2 = 1$, $x_0 = 1/\sqrt{10}$, and so $y_0 = 3/\sqrt{10}$ (remember $x, y \geq 0$ on G). Furthermore, $f(x_0, y_0) = 2 + \sqrt{10}$. Overall, the origin is the absolute minimum and (x_0, y_0) the absolute maximum for f.

13. a) Given that $\nabla f(x,y) = \big(2x(y+1), x^2 - 2\big)$, the stationary points are $P_1 = (\sqrt{2}, -1)$ and $P_2 = (-\sqrt{2}, -1)$. The Hessian

$$\boldsymbol{H}f(x,y) = \begin{pmatrix} 2(y+1) & 2x \\ 2x & 0 \end{pmatrix}$$

at those points reads

$$\boldsymbol{H}f(P_1) = \begin{pmatrix} 0 & 2\sqrt{2} \\ 2\sqrt{2} & 0 \end{pmatrix}, \qquad \boldsymbol{H}f(P_2) = \begin{pmatrix} 0 & -2\sqrt{2} \\ -2\sqrt{2} & 0 \end{pmatrix},$$

so P_1, P_2 are both saddle points.

b) By part a), there are no extrema on the interior of G. But since G is compact, Weierstrass' Theorem ensures there are extremum points, and they belong to ∂G. The boundary of G is the union of the horizontal segment between $A = (-\sqrt{3}, 2)$ and $B = (\sqrt{3}, 2)$ and of the arc joining A to B with equation $y = \sqrt{1+x^2}$ (Fig. 7.22).

On AB we have

$$f(x,2) = 3x^2 - 4\,, \quad x \in [-\sqrt{3}, \sqrt{3}]$$

whence f has local minimum at $x = 0$ and local maximum at $x = \pm\sqrt{3}$, with $f(0,2) = -4$ and $f(\pm\sqrt{3}, 2) = 5$ respectively.

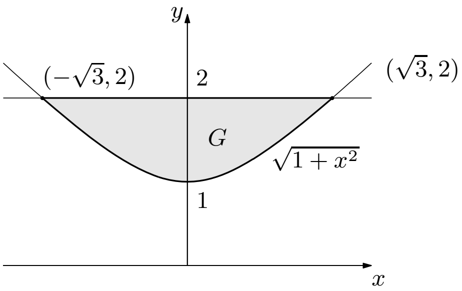


Figure 7.22. The admissible set G of Exercise 13

On the arc connecting A and B along $y = \sqrt{1+x^2}$, we have

$$\varphi(x) = f(x, \sqrt{1+x^2}) = x^2(\sqrt{1+x^2}+1) - 2\sqrt{1+x^2}\,, \quad x \in [-\sqrt{3}, \sqrt{3}]\,.$$

As

$$\varphi'(x) = x\,\frac{3x^2 + 2\sqrt{1+x^2}}{\sqrt{1+x^2}}$$

vanishes at $x = 0$ only, and is positive for $x > 0$, the point $x = 0$ is a local minimum with $f(0,1) = -2$, while $x = \pm\sqrt{3}$ are local maxima.

Therefore, $P_1 = (0,2)$ is the absolute minimum, $P_2 = (\sqrt{3},0)$ and $P_3 = (-\sqrt{3},0)$ are absolute maxima on G.

14. Define $g(x,y) = x^4 - x^2 + y^2 - 5$ and use Lagrange multipliers. As

$$\nabla f(x,y) = (4x, 2y)\,, \qquad \nabla g(x,y) = (4x^3 - 2x, 2y)\,,$$

we consider the system

$$\begin{cases} 4x = \lambda 2x(2x^2-1) \\ 2y = \lambda 2y \\ x^4 - x^2 + y^2 - 5 = 0\,. \end{cases}$$

The second equation gives $\lambda = 1$ or $y = 0$. In the former case we have $x = 0$ or $x = \pm\sqrt{3/2}$, and correspondingly $y = \pm\sqrt{5}$ or $y = \pm\sqrt{17}/2$; in the latter case $x^2 = \frac{1\pm\sqrt{21}}{2}$, so $x = \pm\sqrt{\frac{1+\sqrt{21}}{2}}$ (note $1-\sqrt{21} < 0$ gives a non-valid x^2). Therefore,

$$P_{1,2} = (0, \pm\sqrt{5})\,, \qquad P_{3,4,5,6} = \Big(\pm\sqrt{\frac{3}{2}}, \pm\frac{\sqrt{17}}{2}\Big)\,, \qquad P_{7,8} = \Big(\pm\sqrt{\frac{1+\sqrt{21}}{2}}, 0\Big)$$

are extrema, constrained to G. From

$$f(P_{1,2}) = 5\,, \qquad f(P_{3,4,5,6}) = \frac{29}{4}\,, \qquad f(P_{7,8}) = 1 + \sqrt{21}$$

the maximum value of f is $29/4$, the minimum 5.

15. Let us use Lagrange's multipliers putting $g(x, y) = 4x^2 + y^2 - 1$ and observing

$$\nabla f(x, y) = (4x + 2, 2y - 4)\,, \qquad \nabla g(x, y) = (8x, 2y)\,.$$

We have to solve

$$\begin{cases} 8x + 2 = 8\lambda x \\ 2y - 4 = 2\lambda y \\ 4x^2 + y^2 - 1 = 0 \end{cases} \qquad \Longleftrightarrow \qquad \begin{cases} x = \dfrac{1}{4(\lambda - 1)} \\ y = -\dfrac{2}{\lambda - 1} \\ \dfrac{1}{4}\dfrac{1}{(\lambda - 1)^2} + \dfrac{4}{(\lambda - 1)^2} = 1\,. \end{cases}$$

Note $\lambda \neq 1$, for otherwise the first two equations would be inconsistent. The third equation on the right gives $\lambda - 1 = \pm\frac{\sqrt{17}}{2}$, hence $x = \pm\frac{1}{2\sqrt{17}}$ and $y = \mp\frac{4}{\sqrt{17}}$. The extrema are therefore

$$P_1 = \Big(\frac{1}{2\sqrt{17}}, -\frac{4}{\sqrt{17}}\Big) \qquad \text{and} \qquad P_2 = \Big(-\frac{1}{2\sqrt{17}}, \frac{4}{\sqrt{17}}\Big)\,.$$

From $f(P_1) = 2 + \sqrt{17}$, $f(P_2) = 2 - \sqrt{17}$, we see the maximum is $2 + \sqrt{17}$, the minimum $2 - \sqrt{17}$.

8

Integral calculus in several variables

The definite integral of a function of one real variable allowed us, in Vol. I, to define and calculate the area of a sufficiently regular region in the plane. The present chapter extends this notion to multivariable maps by discussing multiple integrals; in particular, we introduce double integrals for dimension 2 and triple integrals for dimension 3. These new tools rely on the notions of a measurable subset of $\mathbb{R}^n$ and the corresponding n-dimensional measure; the latter extends the idea of the area of a plane region ($n = 2$), and the volume of a solid ($n = 3$) to more general situations.

We continue by introducing methods for computing multiple integrals. Among them, dimensional-reduction techniques transform multiple integrals into one-dimensional integrals, which can be tackled using the rules the reader is already familiar with. On the other hand a variable change in the integration domain can produce an expression, as for integrals by substitution, that is computable with more ease than the multiple integral.

In the sequel we shall explain the part played by multiple integrals in the correct formulation of physical quantities, like mass, centre of gravity, and moments of inertia of a body with given density.

The material cannot hope to exhaust multivariable integral calculus. Integrating a map that depends on n variables over a lower-dimensional manifold gives rise to other kinds of integrals, with great applicative importance, such as curvilinear integrals, or flux integrals through a surface. These in particular will be carefully dealt with in the subsequent chapter, where we will also show how to recover a map from its gradient, i.e., find a primitive of sorts, essentially.

The technical nature of many proofs, that are often adaptations of one-dimensional arguments, has induced us to skip them[1]. The wealth of examples we present will in any case illustrate the statements thoroughly.

[1] The interested reader may find the proofs in, e.g., the classical textbook by R. Courant and F. John, *Introduction to Calculus and Analysis, Vol. II*, Springer, 1999.

C. Canuto, A. Tabacco: *Mathematical Analysis II*, 2nd Ed.,
UNITEXT – La Matematica per il 3+2 85, DOI 10.1007/978-3-319-12757-6_7,

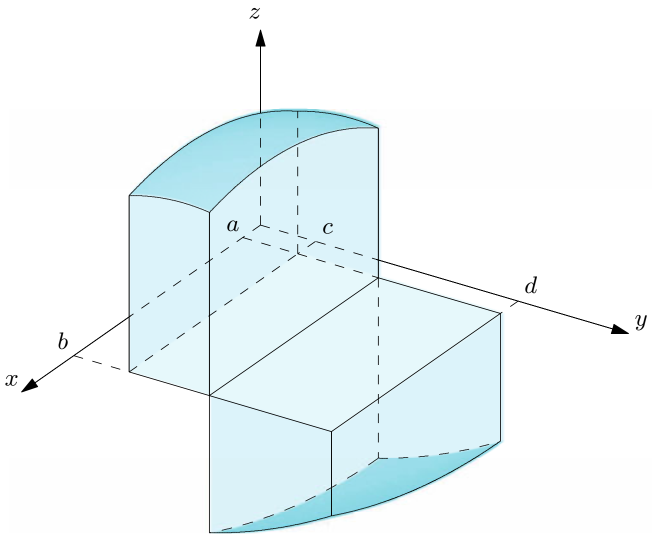


Figure 8.1. The cylindroid of a map

8.1 Double integral over rectangles

Consider a real function $f : B \to \mathbb{R}$, defined on a closed rectangle $B = [a,b] \times [c,d] \subset \mathbb{R}^2$ and bounded over it. We call **cylindroid of** f the three-dimensional region $\mathcal{C}(f;B)$ between B and the graph of f

$$\boxed{\mathcal{C}(f;B) = \{(x,y,z) \in \mathbb{R}^3 : (x,y) \in B,\, 0 \le z \le f(x,y) \text{ or } f(x,y) \le z \le 0\}\,,}$$

see Fig. 8.1. (The choice of bounds for z depends on the sign of $f(x,y)$.) If f satisfies certain requirements, we may associate to the cylindroid of f a number called the *double integral of* f *over* B. In case f is positive, this number represents the region's volume. In particular, when the cylindroid is a simple solid (e.g., a parallelepiped, a prism, and so on) it gives the usual expression for the volume.

Many are the ways to construct the double integral of a function; we will explain the method due to Riemann, which generalises what we saw in Vol. I, Sects. 9.4 and 9.5, for dimension 1.

Let us thus consider arbitrary partitions of $[a,b]$ and $[c,d]$ associated to the ordered points $\{x_0, x_1, \ldots, x_p\}$ and $\{y_0, y_1, \ldots, y_q\}$

$$a = x_0 < x_1 < \cdots < x_{p-1} < x_p = b\,, \qquad c = y_0 < y_1 < \cdots < y_{q-1} < y_q = d\,,$$

with

$$I = [a,b] = \bigcup_{h=1}^{p} I_h = \bigcup_{h=1}^{p} [x_{h-1}, x_h]\,, \qquad J = [c,d] = \bigcup_{k=1}^{q} J_k = \bigcup_{k=1}^{q} [y_{k-1}, y_k]\,.$$

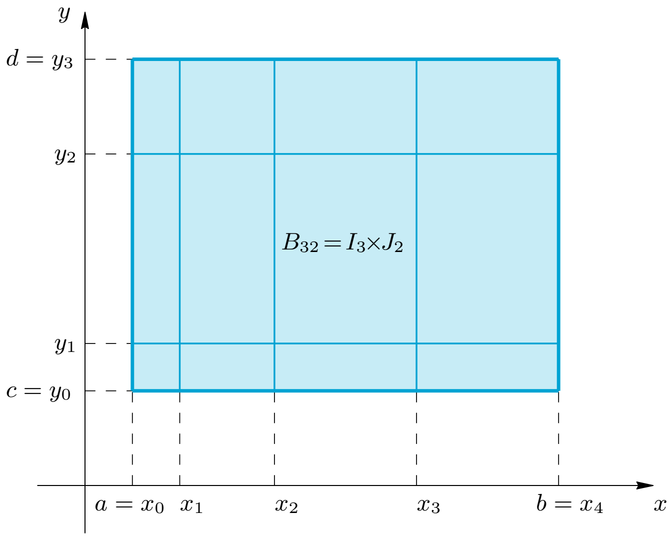


Figure 8.2. Subdivision of the rectangle B

The rectangle B is made of $p \cdot q$ products $B_{hk} = I_h \times J_k$; we have thus built a *partition* or *subdivision* of B, say $\mathcal{D} = \{B_{hk} : h = 1, \dots p,\ k = 1, \dots, q\}$, which is the product of the partitions of $[a, b]$ and $[c, d]$ (Fig. 8.2). Set

$$m_{hk} = \inf_{(x,y)\in B_{hk}} f(x, y) \qquad \text{and} \qquad M_{hk} = \sup_{(x,y)\in B_{hk}} f(x, y)$$

and define the **lower** and **upper sum** of f on B relative to the subdivision $\mathcal{D}$ by

$$\boxed{\begin{aligned} s = s(\mathcal{D}, f) &= \sum_{k=1}^{q}\sum_{h=1}^{p} m_{hk}(x_h - x_{h-1})(y_k - y_{k-1})\,, \\ S = S(\mathcal{D}, f) &= \sum_{k=1}^{q}\sum_{h=1}^{p} M_{hk}(x_h - x_{h-1})(y_k - y_{k-1})\,. \end{aligned}} \tag{8.1}$$

Since f is bounded on B, there exist constants m and M such that, for each subdivision $\mathcal{D}$,

$$m(b - a)(d - c) \le s(\mathcal{D}, f) \le S(\mathcal{D}, f) \le M(b - a)(d - c)\,.$$

The following quantities are thus well defined

$$\boxed{\overline{\int_B} f = \inf_{\mathcal{D}} S(\mathcal{D}, f) \qquad \text{and} \qquad \underline{\int_B} f = \sup_{\mathcal{D}} s(\mathcal{D}, f)\,,} \tag{8.2}$$

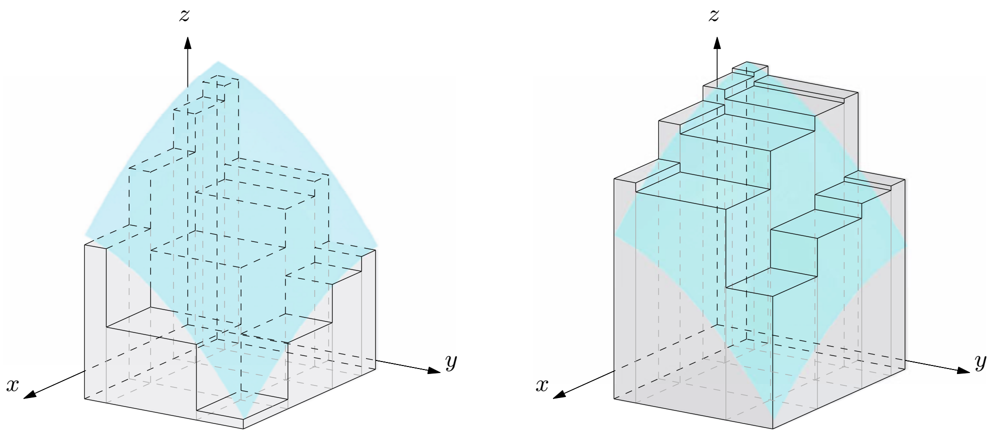


Figure 8.3. Lower (left) and upper sum (right)

respectively called **upper integral** and **lower integral** of f over B. As in the one-dimensional case, it is not hard to check that

$$\underline{\int_B} f \leq \overline{\int_B} f\,.$$

Definition 8.1 *A map f bounded on $B = [a,b] \times [c,d] = I \times J$ is* **Riemann integrable on** B *if*

$$\underline{\int_B} f = \overline{\int_B} f\,.$$

This value is called the **double integral of** f **over** B*, and denoted by one of the symbols*

$$\int_B f\,,\quad \iint_B f\,,\quad \int_B f(x,y)\,\mathrm{d}x\,\mathrm{d}y\,,\quad \int_J \int_I f(x,y)\,\mathrm{d}x\,\mathrm{d}y\,,\quad \int_c^d \int_a^b f(x,y)\,\mathrm{d}x\,\mathrm{d}y\,.$$

The geometrical meaning is clear when f is positive on B. Given a partition $\mathcal{D}$ of B, the cylindroid of f is contained inside the solid formed by the union of the parallelepipeds with base B_{hk} and height M_{hk}, and it contains the solid made by the parallelepipeds with the same base and m_{hk} as height (see Fig. 8.3).

The upper integral is an *over-estimate* of the region, while the lower integral an *under-estimate*. Thus f is integrable when these two concide, i.e., when the region defines a number representing its volume.

Examples 8.2

i) Suppose f is constant on B, say equal K. Then for any partition $\mathcal{D}$, we have $m_{hk} = M_{hk} = K$, so

$$s(\mathcal{D}, f) = S(\mathcal{D}, f) = \sum_{k=1}^{q}\sum_{h=1}^{p} K\,(x_h - x_{h-1})(y_k - y_{k-1}) = K\,(b-a)(d-c)\,.$$

Therefore

$$\int_B f = K\,(b-a)(d-c) = K \cdot \text{area}(B)\,.$$

ii) Let f be the two-dimensional Dirichlet function on $B = [0,1] \times [0,1]$

$$f(x,y) = \begin{cases} 1 & \text{if } x, y \in \mathbb{Q},\ 0 \le x, y \le 1, \\ 0 & \text{otherwise}\,. \end{cases}$$

For any given partiton $\mathcal{D}$ then,

$$s(\mathcal{D}, f) = \sum_{k=1}^{q}\sum_{h=1}^{p} 0 \cdot (x_h - x_{h-1})(y_k - y_{k-1}) = 0\,,$$

$$S(\mathcal{D}, f) = \sum_{k=1}^{q}\sum_{h=1}^{p} 1 \cdot (x_h - x_{h-1})(y_k - y_{k-1}) = 1\,.$$

This shows f is not integrable on B. □

Remark 8.3 Let us examine, in detail, similarities and differences with the definition, as of Vol. I, of a Riemann integrable map on an interval $[a,b]$. We may subdivide the rectangle $B = [a,b] \times [c,d]$ without having to use Cartesian products of partitions of $[a,b]$ and $[c,d]$. For instance, we could subdivide $B = \cup_{i=1}^{N} B_i$ by taking coordinate rectangles B_i as in Fig. 8.4. Clearly, the kind of partitions we consider are a special case of these.

Furthermore, if we are given a generic two-dimensional step function φ associated to such a rectangular partition, it is possible to define its double integral in an elementary way, namely: let

$$\varphi(x,y) = c_i\,, \qquad \forall (x,y) \in B_i\,, \quad i = 1, \ldots, N\,;$$

then

$$\int_B \varphi = \sum_{i=1}^{N} c_i \text{area}(B_i)\,.$$

Notice the lower and upper sums of a bounded map $f : B \to \mathbb{R}$, defined in (8.1), are precisely the integrals of two step functions, one smaller and one larger than f. Their constant values on each sub-rectangle coincide respectively with the greatest lower bound and least upper bound of f on the sub-rectangle. We could have considered the set $\mathcal{S}_f^-$ (resp. $\mathcal{S}_f^+$) of step functions that are smaller (larger) than f. The lower and upper integrals thus obtained,

$$\underline{\int_B} f = \sup\{\int_B g : g \in \mathcal{S}_f^-\} \qquad \text{and} \qquad \overline{\int_B} f = \inf\{\int_B g : g \in \mathcal{S}_f^+\},$$

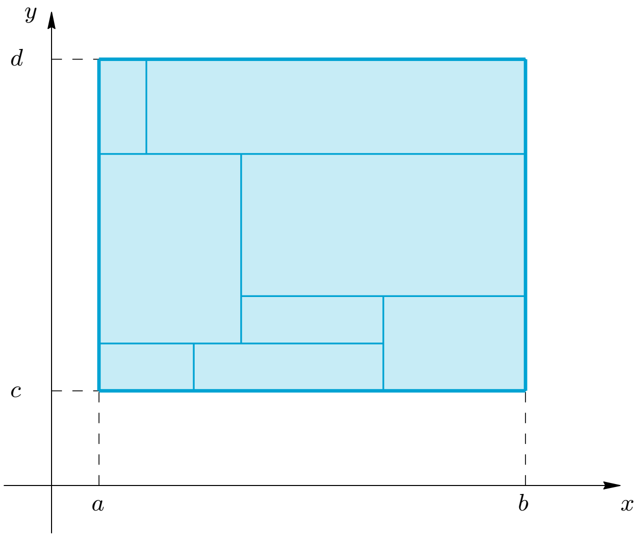


Figure 8.4. Generic partition of the rectangle B

coincide with those introduced in (8.2). In other terms, the two recipes for the definite integral give the same result. □

As for one-variable maps, it is imperative to find classes of integrable functions and be able to compute their integrals directly, without turning to the definition. A partial answer to the first problem is provided by the following theorem. Further classes of integrable functions will be considered in subsequent sections.

Theorem 8.4 *If f is continuous on the rectangle B, then it is integrable on B.*

Our next result allows us to reduce a double integral, under suitable assumptions, to the computation of two integrals over real intervals.

Theorem 8.5 *Let f be integrable over $B = [a,b] \times [c,d]$.*

a) If, for any $y \in [c,d]$, the integral $g(y) = \int_a^b f(x,y)\,\mathrm{d}x$ exists, the map $g : [c,d] \to \mathbb{R}$ is integrable on $[c,d]$ and

$$\int_B f = \int_c^d g(y)\,\mathrm{d}y = \int_c^d \left(\int_a^b f(x,y)\,\mathrm{d}x \right) \mathrm{d}y\,. \tag{8.3}$$

b) If, for any $x \in [a,b]$*, the integral* $h(x) = \int_c^d f(x,y)\,\mathrm{d}y$ *exists, the map* $h : [a,b] \to \mathbb{R}$ *is integrable on* $[a,b]$ *and*

$$\int_B f = \int_a^b h(x)\,\mathrm{d}x = \int_a^b \left(\int_c^d f(x,y)\,\mathrm{d}y \right) \mathrm{d}x\,. \tag{8.4}$$

In particular, if f *is continuous on* B*,*

$$\int_B f = \int_a^b \left(\int_c^d f(x,y)\,\mathrm{d}y \right) \mathrm{d}x = \int_c^d \left(\int_a^b f(x,y)\,\mathrm{d}x \right) \mathrm{d}y\,. \tag{8.5}$$

Formulas (8.3) and (8.4) are said, generically, *reduction formulas for iterated integrals*. When they hold simultaneously, as happens for continuous maps, we say that *the order of integration can be swapped* in the double integral.

Examples 8.6

i) A special case occurs when f has the form $f(x,y) = h(x)g(y)$, with h integrable on $[a,b]$ and g integrable on $[c,d]$. Then it can be proved that f is integrable on $B = [a,b] \times [c,d]$, and formulas (8.3), (8.4) read

$$\int_B f = \left(\int_a^b h(x)\,\mathrm{d}x \right) \cdot \left(\int_c^d g(y)\,\mathrm{d}y \right)\,.$$

This means the double integral coincides with the product of the two one-dimensional integrals of h and g.

ii) Let us determine the double integral of $f(x,y) = \cos(x+y)$ over $B = [0, \frac{\pi}{4}] \times [0, \frac{\pi}{2}]$. The map is continuous, so we may indifferently use (8.3) or (8.4). For example,

$$\begin{aligned}\int_B f &= \int_0^{\pi/4} \left(\int_0^{\pi/2} \cos(x+y)\,\mathrm{d}y \right) \mathrm{d}x = \int_0^{\pi/4} \Big[\sin(x+y) \Big]_{y=0}^{y=\pi/2} \mathrm{d}x \\ &= \int_0^{\pi/4} \left(\sin(x + \frac{\pi}{2}) - \sin x \right) \mathrm{d}x = \sqrt{2} - 1\,.\end{aligned}$$

iii) Let us compute the double integral of $f(x,y) = x \cos xy$ over $B = [1,2] \times [0,\pi]$. Although (8.3) and (8.4) are both valid, the latter is more convenient here. In fact,

$$\begin{aligned}\int_B x \cos xy\,\mathrm{d}x\,\mathrm{d}y &= \int_1^2 \left(\int_0^\pi x \cos xy\,\mathrm{d}y \right) \mathrm{d}x \\ &= \int_1^2 \Big[\sin xy \Big]_{y=0}^{y=\pi} \mathrm{d}x = \int_1^2 \sin \pi x\,\mathrm{d}x = -\frac{2}{\pi}\,.\end{aligned}$$

Formula (8.3) involves more elaborate computations which the reader might want to perform. □

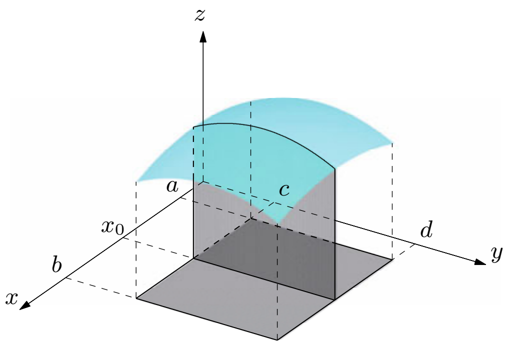


Figure 8.5. The geometrical meaning of integrating on cross-sections

Remark 8.7 To interpret an iterated integral geometrically, let us assume for simplicity f is positive and continuous on B, and consider (8.4). Given an $x_0 \in [a, b]$, $h(x_0) = \int_c^d f(x_0, y)\,\mathrm{d}y$ represents the area of the region obtained as intersection between the cylindroid of f and the plane $x = x_0$. The region's volume is the integral from a to b of such area (Fig. 8.5). □

8.2 Double integrals over measurable sets

For one-dimensional integrals the region over which to integrate is always an interval, or a finite union of intervals. Since double-integrating only over rectangles (or finite unions thereof) is too restrictive, we need to introduce the kind of sets over which we will discuss integrability. To this end, let Ω be an arbitrary bounded subset of $\mathbb{R}^2$, and denote by $\chi_\Omega : \mathbb{R}^2 \to \mathbb{R}$ its characteristic function

$$\chi_\Omega(\boldsymbol{x}) = \begin{cases} 1 & \text{if } \boldsymbol{x} \in \Omega\,, \\ 0 & \text{if } \boldsymbol{x} \notin \Omega \end{cases}$$

(see Fig. 8.6).

We fix an arbitrary rectangle B containing Ω, and ask ourselves whether χ_Ω is integrable on B. It is easy to check that if yes, and if B' is any other rectangle containing Ω, the function χ_Ω is still integrable on B'. Moreover the two integrals $\int_B \chi_\Omega$ and $\int_{B'} \chi_\Omega$ coincide, so the common value can be denoted by $\int_\Omega \chi_\Omega$. That said, let us introduce the notion of a measurable set à la Peano-Jordan.

Definition 8.8 *A bounded subset $\Omega \subset \mathbb{R}^2$ is* **measurable** *if, for an arbitrary rectangle B containing Ω, the function χ_Ω is integrable on B. If so, the non-negative number*

$$|\Omega| = \int_\Omega \chi_\Omega$$

is the **measure** (*or area*) *of Ω.*

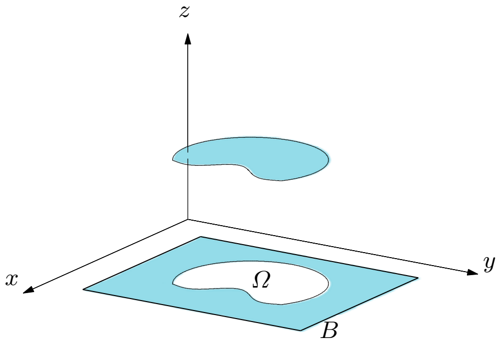


Figure 8.6. The characteristic function of a set Ω

Examples of measurable sets include polygons and discs, and the above notion is precisely what we already know as their surface area. At times we shall write area(Ω) instead of $|\Omega|$. In particular, for any rectangle $B = [a, b] \times [c, d]$ we see immediately that

$$|B| = \text{area}(B) = \int_B \mathrm{d}x\,\mathrm{d}y = \int_c^d\!\!\int_a^b \mathrm{d}x\,\mathrm{d}y = (b-a)(d-c).$$

Not all bounded sets in the plane are measurable. Think of the points in the square $[0, 1] \times [0, 1]$ with rational coordinates. The characteristic function of this set is the Dirichlet function relative to Example 8.2 ii), which is not integrable. Therefore the set is not measurable.

Another definition will turn out useful in the sequel.

Definition 8.9 *A set Ω has* **zero measure** *if it is measurable and $|\Omega| = 0$.*

With this we can characterise measurable sets in the plane. In fact, the next result is often used the tell whether a given set is measurable or not.

Theorem 8.10 *A bounded set $\Omega \subset \mathbb{R}^2$ is measurable if and only if its boundary $\partial\Omega$ has measure zero.*

Among zero-measure sets are:

i) subsets of sets with zero measure;

ii) finite unions of sets with zero measure;

iii) sets consisting of finitely many points;

iv) segments;

v) graphs $\Gamma = \{(x, f(x))\}$ or $\Gamma = \{(f(y), y)\}$ of integrable maps $f : [a, b] \to \mathbb{R}$;

vi) traces of (piecewise-)regular plane curves.

As for the last, we should not confuse the measure of the trace of a plane curve with its length, which is in essence a one-dimensional measure for sets that are not contained in a line.

A case of zero-measure set of particular relevance is the graph of a continuous map defined on a closed and bounded interval. Therefore, measurable sets include bounded sets whose boundary is a finite union of such graphs, or more generally, a finite union of regular Jordan arcs. At last, bounded and convex sets are measurable.

Measurable sets and their measures enjoy some properties.

Property 8.11 *If Ω_1 and Ω_2 are measurable,*

i) $\Omega_1 \subseteq \Omega_2$ *implies* $|\Omega_1| \leq |\Omega_2|$;

ii) $\Omega_1 \cup \Omega_2$ *and* $\Omega_1 \cap \Omega_2$ *are measurable with*

$$|\Omega_1 \cup \Omega_2| = |\Omega_1| + |\Omega_2| - |\Omega_1 \cap \Omega_2| \,, \quad \textit{hence} \quad |\Omega_1 \cup \Omega_2| \leq |\Omega_1| + |\Omega_2| \,.$$

Here is a simple, yet useful result for the sequel.

Property 8.12 *If Ω is a measurable set, so are the interior $\mathring{\Omega}$, the closure $\overline{\Omega}$, and in general all sets $\tilde{\Omega}$ such that $\mathring{\Omega} \subseteq \tilde{\Omega} \subseteq \overline{\Omega}$. Any of these has measure $|\Omega|$.*

Proof. All sets $\tilde{\Omega}$ have the same boundary $\partial\tilde{\Omega} = \partial\Omega$. Therefore Theorem 8.10 tells they are measurable. But they differ from one another by subsets of $\partial\Omega$, whose measure is zero. □

Now we introduce the notion of integrability for bounded maps on a given measurable set Ω. The method is completely similar to the way we selected measurable sets. Let $f : \Omega \to \mathbb{R}$ be a bounded map, and consider $\tilde{f} : \mathbb{R}^2 \to \mathbb{R}$ (called the trivial extension of f to $\mathbb{R}^2$), defined by

$$\tilde{f}(\boldsymbol{x}) = \begin{cases} f(\boldsymbol{x}) & \text{if } \boldsymbol{x} \in \Omega, \\ 0 & \text{if } \boldsymbol{x} \notin \Omega \end{cases} \tag{8.6}$$

(see Fig. 8.7).

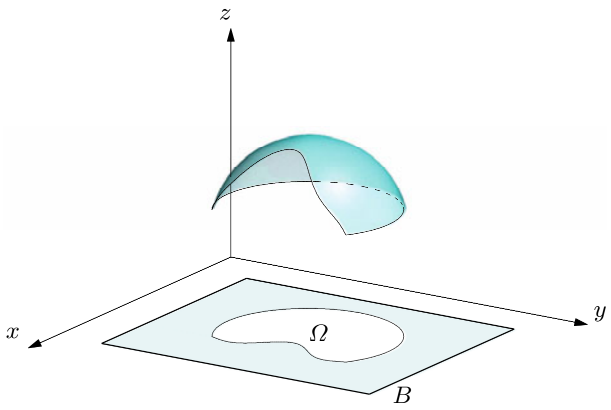


Figure 8.7. Trivial extension of a map to a rectangle B

Definition 8.13 *The map f is (Riemann)* **integrable** *on Ω if $\tilde{f}$ is integrable on any rectangle B containing Ω. In that case, the integral $\displaystyle\int_B \tilde{f}$ is independent of the choice of B, and we set*

$$\int_\Omega f = \int_B \tilde{f} .$$

This value is called the **double integral of** f **over** Ω. *Other symbols to denote it are*

$$\iint_\Omega f , \qquad \int_\Omega f(x,y)\,\mathrm{d}x\,\mathrm{d}y , \qquad \int_\Omega f\,\mathrm{d}\Omega .$$

Observe that if f coincides with the characteristic function of Ω, what we have just defined is exactly the measure of Ω.

The first large class of integrable maps we wish to describe is a sort of generalisation of piecewise-continuous functions in one variable.

Definition 8.14 *A map $f : \Omega \to \mathbb{R}$, bounded on a measurable set Ω, is said* **generically continuous** *on Ω if the discontinuity set has zero measure.*

A bounded, continuous map on Ω is clearly generically continuous, as no discontinuity points are present. An example of a generically continuous, but not continuous, function is given by the map $f(x, y) = \operatorname{sign}(x - y)$ on the square $\Omega = (0, 1)^2$.

Then we have the following result.

Theorem 8.15 *Let f be generically continuous on a measurable set Ω. Then f is integrable on Ω.*

There is a way to compute the double integral on special regions in the plane, which requires a definition.

Definition 8.16 *A set $\Omega \subset \mathbb{R}^2$ is* **normal with respect to the y-axis** *if it is of the form*

$$\Omega = \{(x, y) \in \mathbb{R}^2 : a \leq x \leq b,\ g_1(x) \leq y \leq g_2(x)\}$$

with g_1, $g_2 : [a, b] \to \mathbb{R}$ continuous.
Analogously, Ω is **normal with respect to the x-axis** *if*

$$\Omega = \{(x, y) \in \mathbb{R}^2 : c \leq y \leq d,\ h_1(y) \leq y \leq h_2(y)\}$$

with h_1, $h_2 : [c, d] \to \mathbb{R}$ continuous. We will shorten this by the terms **normal for y** *and* **normal for x** *respectively.*

This notion can be understood geometrically. The set Ω is normal with respect to the y-axis if any vertical line $x = x_0$ either does not meet Ω, or it intersects it in the segment (possibly reduced to a point) with end points $\big(x_0, g_1(x_0)\big)$, $\big(x_0, g_2(x_0)\big)$. Similarly, Ω is normal for x if a horizontal line $y = y_0$ has intersection either empty or the segment between $\big(h_1(y_0), y_0\big)$, $\big(h_2(y_0), y_0\big)$ with Ω. Examples are shown in Fig. 8.8 and Fig. 8.9.

Normal sets for either variable are clearly measurable, because their boundary is a finite union of zero-measure sets (graphs of continuous maps and segments). The next proposition allows us to compute a double integral by iteration, thus generalising Theorem 8.5 for rectangles.

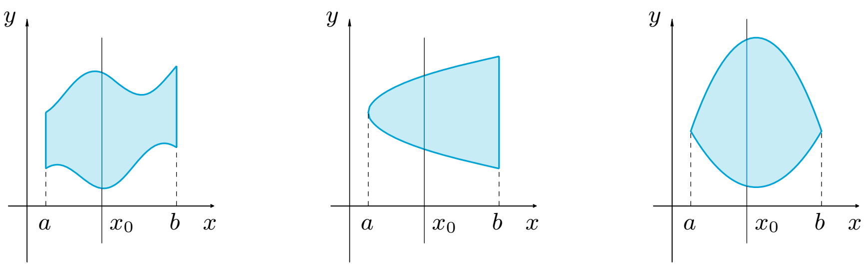


Figure 8.8. Normal sets for y

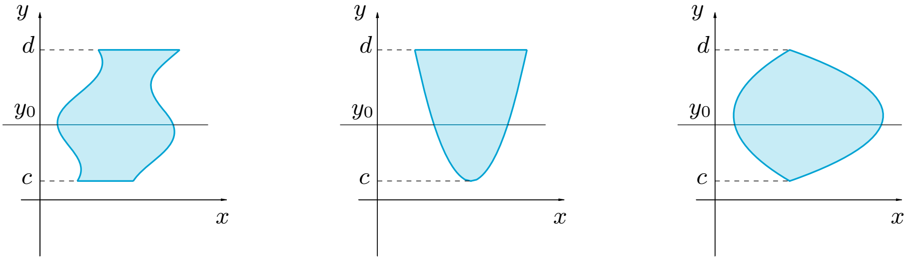


Figure 8.9. Normal sets for x

Theorem 8.17 *Let $f : \Omega \to \mathbb{R}$ be continuous on Ω. If Ω is normal for y, then*

$$\int_\Omega f = \int_a^b \left(\int_{g_1(x)}^{g_2(x)} f(x,y)\,\mathrm{d}y \right) \mathrm{d}x\,. \tag{8.7}$$

If Ω is normal for x,

$$\int_\Omega f = \int_c^d \left(\int_{h_1(y)}^{h_2(y)} f(x,y)\,\mathrm{d}x \right) \mathrm{d}y\,. \tag{8.8}$$

Proof. Suppose Ω is normal for x. From the definition, Ω is a closed and bounded, hence compact, subset of $\mathbb{R}^2$. By Weierstrass' Theorem 5.24, f is bounded on Ω. Then Theorem 8.15 implies f is integrable on Ω. As for formula (8.8), let $B = [a,b] \times [c,d]$ be a rectangle containing Ω, and consider the trivial extension $\tilde{f}$ of f to B. The idea is to use Theorem 8.5 a); for this, we observe that for any $y \in [c,d]$, the integral

$$g(y) = \int_a^b \tilde{f}(x,y)\,\mathrm{d}x$$

exists because $x \mapsto \tilde{f}(x,y)$ has not more than two discontinuity points (see Fig. 8.10). Moreover

$$\int_a^b \tilde{f}(x,y)\,\mathrm{d}x = \int_{h_1(y)}^{h_2(y)} f(x,y)\,\mathrm{d}x\,,$$

so

$$\int_\Omega f = \int_B \tilde{f} = \int_c^d \left(\int_a^b \tilde{f}(x,y)\,\mathrm{d}x \right) \mathrm{d}y = \int_c^d \left(\int_{h_1(y)}^{h_2(y)} f(x,y)\,\mathrm{d}x \right) \mathrm{d}y\,,$$

proving (8.8). Formula (8.7) is similar. $\square$

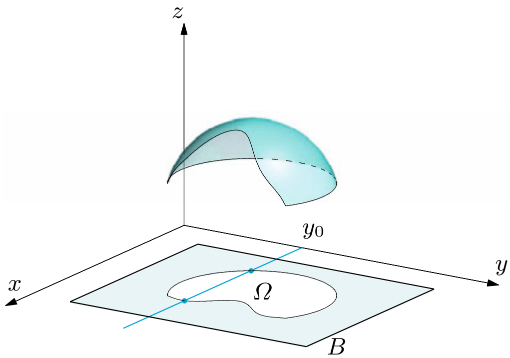


Figure 8.10. Reduction formula for an iterated integral

Also (8.7) and (8.8) are *iterated integrals*, respectively referred to as *formula of vertical* and *of horizontal integration.*

Examples 8.18

i) Compute

$$\int_\Omega (x+2y)\,\mathrm{d}x\,\mathrm{d}y\,,$$

where Ω is the region in the first quadrant bounded by the curves $y = 2x$, $y = 3 - x^2$, $x = 0$ (as in Fig. 8.11, left). The first line and the parabola meet at two points with $x = -3$ and $x = 1$, the latter of which delimits Ω on the right. The set Ω is normal for both x and y, but given its shape we prefer to integrate in y first. In fact,

$$\Omega = \big\{(x,y) \in \mathbb{R}^2 : 0 \le x \le 1,\, 2x \le y \le 3 - x^2\big\}.$$

Hence,

$$\begin{aligned}\int_\Omega (x+2y)\,\mathrm{d}x\,\mathrm{d}y &= \int_0^1 \left(\int_{2x}^{3-x^2} (x+2y)\,\mathrm{d}y\right)\mathrm{d}x = \int_0^1 \Big[xy + y^2\Big]_{y=2x}^{y=3-x^2}\,\mathrm{d}x \\ &= \int_0^1 \big(x^4 - x^3 - 12x^2 + 3x + 9\big)\,\mathrm{d}x \\ &= \left[\frac{1}{5}x^5 - \frac{1}{4}x^4 - 4x^3 + \frac{3}{2}x^2 + 9x\right]_{x=0}^{x=1} = \frac{129}{20}.\end{aligned}$$

Had we chosen to integrate in x first, we would have had to write the domain as

$$\Omega = \big\{(x,y) \in \mathbb{R}^2 : 0 \le y \le 3,\, 0 \le x \le h_2(y)\big\},$$

where $h_2(y)$ is

$$h_2(y) = \begin{cases} \dfrac{1}{2}y & \text{if } 0 \le y \le 2\,, \\ \sqrt{3-y} & \text{if } 2 < y \le 3\,. \end{cases}$$

The form of $h_2(y)$ clearly suggests why the former technique is to be preferred.

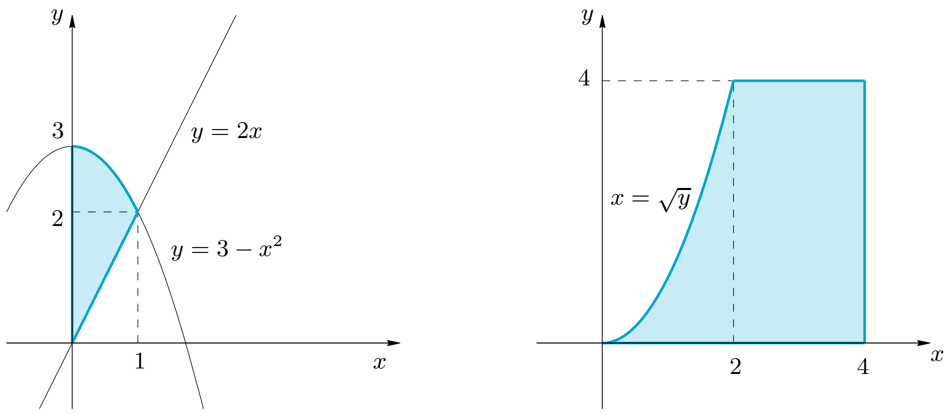


Figure 8.11. The set $\Omega = \{(x,y) \in \mathbb{R}^2 : 0 \le x \le 1,\ 2x \le y \le 3 - x^2\}$ (left) and $\Omega = \{(x,y) \in \mathbb{R}^2 : 0 \le y \le 4,\ \sqrt{y} \le x \le 4\}$ (right)

ii) Consider

$$\int_\Omega (5y + 2x)\,\mathrm{d}x\,\mathrm{d}y$$

where Ω is the plane region bounded by $y = 0$, $y = 4$, $x = 4$ and by the graph of $x = \sqrt{y}$ (see Fig. 8.11, right). Again, Ω is normal for both x and y, but horizontal integration will turn out to be better. In fact, we write

$$\Omega = \big\{(x,y) \in \mathbb{R}^2 : 0 \le y \le 4,\ \sqrt{y} \le x \le 4\big\};$$

therefore

$$\begin{aligned}\int_\Omega (5y + 2x)\,\mathrm{d}x\,\mathrm{d}y &= \int_0^4 \left(\int_{\sqrt{y}}^4 (5y + 2x)\,\mathrm{d}x \right) \mathrm{d}y = \int_0^4 \Big[5xy + x^2 \Big]_{x=\sqrt{y}}^{x=4} \mathrm{d}y \\ &= \int_0^4 \Big(19y + 16 - 5y^{3/2} \Big)\,\mathrm{d}y \\ &= \left[\frac{19}{2} y^2 + 16y - 2y^{5/2} \right]_{y=0}^{y=4} = 152. \end{aligned}$$

□

For positive functions that are integrable on generic measurable sets Ω, e.g., rectangles, the double integral makes geometrical sense, namely it represents the volume $\mathrm{vol}(\mathcal{C})$ of the cylindroid of f

$$\boxed{\mathcal{C} = \mathcal{C}(f;\Omega) = \{(x,y,z) \in \mathbb{R}^3 : (x,y) \in \Omega,\ 0 \le z \le f(x,y)\}\,.}$$

See Fig. 8.12.

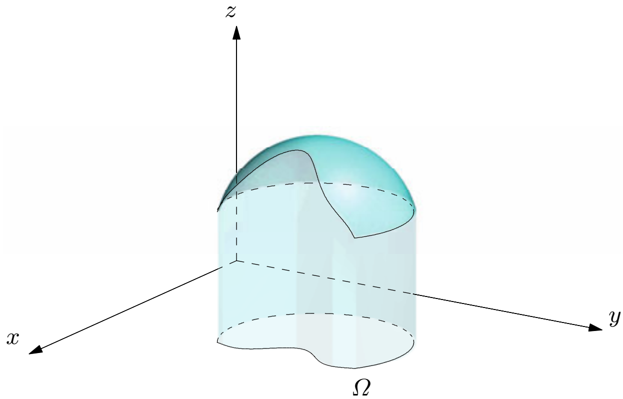


Figure 8.12. The cylindroid of a positive map

Example 8.19

Compute the volume of

$$\mathcal{C} = \left\{(x,y,z) \in \mathbb{R}^3 : 0 \le y \le \frac{3}{4}x,\ x^2+y^2 \le 25,\ z \le xy\right\}.$$

The solid is the cylindroid of $f(x,y) = xy$ with base

$$\Omega = \left\{(x,y) \in \mathbb{R}^2 : 0 \le y \le \frac{3}{4}x,\ x^2+y^2 \le 25\right\}.$$

The map is integrable (as continuous) on Ω, which is normal for both x and y (Fig. 8.13). Therefore $\mathrm{vol}(\mathcal{C}) = \int_\Omega xy\,\mathrm{d}x\,\mathrm{d}y$.

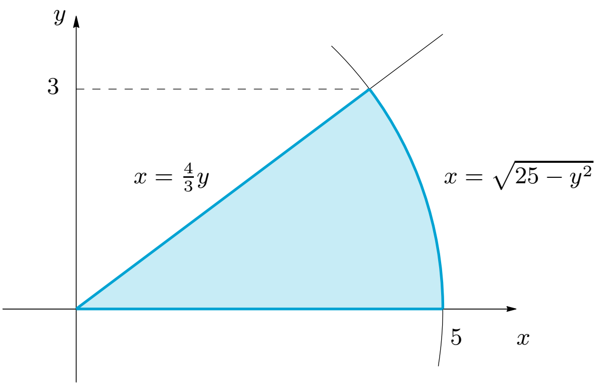


Figure 8.13. The set $\Omega = \left\{(x,y) \in \mathbb{R}^2 : 0 \le y \le 3,\ \frac{4}{3}y \le x \le \sqrt{25-y^2}\right\}$

The region Ω lies in the first quadrant, is bounded by the line $y = \frac{3}{4}x$ and the circle $x^2 + y^2 = 25$ centred at the origin with radius 5. For simplicity let us integrate horizontally, since

$$\Omega = \Big\{(x, y) \in \mathbb{R}^2 : 0 \le y \le 3,\ \frac{4}{3}y \le x \le \sqrt{25 - y^2}\Big\}.$$

Thus

$$\begin{aligned}\int_\Omega xy \, dx \, dy &= \int_0^3 \left(\int_{\frac{4}{3}y}^{\sqrt{25-y^2}} xy \, dx \right) dy \\ &= \int_0^3 \left[\frac{1}{2} yx^2 \right]_{x=\frac{4}{3}y}^{x=\sqrt{25-y^2}} dy \\ &= \frac{1}{2} \int_0^3 \Big(y(25 - y^2) - \frac{16}{9} y^3 \Big) dy = \frac{225}{8}.\end{aligned}$$

□

8.2.1 Properties of double integrals

In this section we state a bunch of useful properties of double integrals.

Theorem 8.20 *Let f, g be integrable maps on a measurable set $\Omega \subset \mathbb{R}^2$.*

i) (**Linearity**) *For any $\alpha, \beta \in \mathbb{R}$, the map $\alpha f + \beta g$ is integrable on Ω and*

$$\int_\Omega \alpha f + \beta g = \alpha \int_\Omega f + \beta \int_\Omega g.$$

ii) (**Positivity**) *If $f \ge 0$ on Ω, then*

$$\int_\Omega f \ge 0.$$

In addition, if f is continuous and Ω measurable with $|\Omega| > 0$, we have equality above if and only if f is identically zero.

iii) (**Comparison/Monotonicity**) *If $f \le g$ on Ω, then*

$$\int_\Omega f \le \int_\Omega g.$$

iv) (**Boundedness**) *The map $|f|$ is integrable on Ω and*

$$\left| \int_\Omega f \right| \le \int_\Omega |f|.$$

v) (**Mean Value Theorem**) *If Ω is measurable and*

$$m = \inf_{(x,y)\in\Omega} f(x,y)\,, \qquad M = \sup_{(x,y)\in\Omega} f(x,y)\,,$$

then

$$m \le \frac{1}{|\Omega|}\int_\Omega f \le M\,.$$

The number

$$\frac{1}{|\Omega|}\int_\Omega f \tag{8.9}$$

is called (**integral**) **mean value** *of f on Ω.*

vi) (**Additivity of domains**) *Let $\Omega = \Omega_1 \cup \Omega_2$, with $\Omega_1 \cap \Omega_2$ of zero measure. If f is integrable on Ω_1 and on Ω_2, then f is integrable on Ω, and*

$$\int_\Omega f = \int_{\Omega_1} f + \int_{\Omega_2} f\,.$$

vii) If $f = g$ except that on a zero-measure subset of Ω, then

$$\int_\Omega f = \int_\Omega g\,.$$

Property *vi)* is extremely useful when integrals are defined over unions of finitely many normal sets for one variable.

There is a counterpart to Property 8.12 for integrable functions.

Property 8.21 *Let f be integrable on a measurable set Ω, and suppose it is defined on the closure $\overline{\Omega}$ of Ω. Then f is integrable over any subset $\tilde{\Omega}$ such that $\mathring{\Omega} \subseteq \tilde{\Omega} \subseteq \overline{\Omega}$, and*

$$\int_{\tilde{\Omega}} f = \int_\Omega f\,.$$

Otherwise said, the double integral of an integrable map does not depend on whether bits of the boundary belong to the integration domain.

Examples 8.22

i) Consider

$$\int_\Omega (1+x)\,\mathrm{d}x\,\mathrm{d}y\,,$$

where $\Omega = \{(x,y) \in \mathbb{R}^2 : y > |x|,\ y < \frac{1}{2}x + 2\}$.

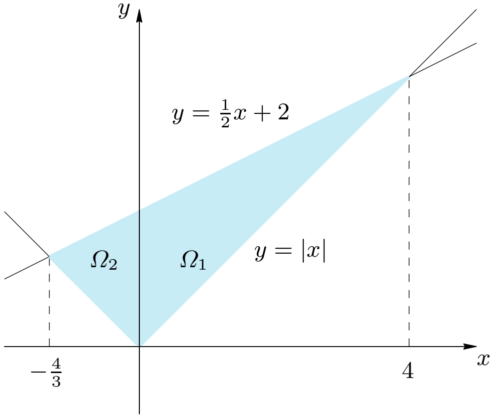


Figure 8.14. The set Ω relative to Example 8.22 i)

The domain Ω, depicted in Fig. 8.14, is made of the points lying between the graphs of $y = |x|$ and $y = \frac{1}{2}x+2$. The set Ω is normal for both x and y; it is more convenient to start integrating in y. Due to the presence of $y = |x|$, moreover, it is better to compute the integral on Ω as sum of integrals over the subsets Ω_1 and Ω_2 of the picture. The graphs of $y = |x|$ and $y = \frac{1}{2}x + 2$ meet for $x = 4$ and $x = -\frac{4}{3}$. Then Ω_1 and Ω_2 are, respectively,

$$\Omega_1 = \{(x, y) \in \mathbb{R}^2 : 0 \le x < 4,\ x < y < \frac{1}{2}x + 2\}$$

and

$$\Omega_2 = \{(x, y) \in \mathbb{R}^2 : -\frac{4}{3} < x < 0,\ -x < y < \frac{1}{2}x + 2\}.$$

Therefore

$$\begin{aligned}\int_{\Omega_1} (1 + x)\ \mathrm{d}x\,\mathrm{d}y &= \int_0^4 \left(\int_x^{\frac{1}{2}x+2} (1 + x)\,\mathrm{d}y \right) \mathrm{d}x = \int_0^4 [y + xy]_x^{\frac{1}{2}x+2}\ \mathrm{d}x \\ &= \int_0^4 \left(-\frac{1}{2}x^2 + \frac{3}{2}x + 2 \right) \mathrm{d}x = \left[-\frac{1}{6}x^3 + \frac{3}{4}x^2 + 2x \right]_0^4 = \frac{28}{3}\end{aligned}$$

and

$$\begin{aligned}\int_{\Omega_2} (1 + x)\ \mathrm{d}x\,\mathrm{d}y &= \int_{-\frac{4}{3}}^0 \left(\int_{-x}^{\frac{1}{2}x+2} (1 + x)\,\mathrm{d}y \right) \mathrm{d}x = \int_{-\frac{4}{3}}^0 [y + xy]_{-x}^{\frac{1}{2}x+2}\ \mathrm{d}x \\ &= \int_{-\frac{4}{3}}^0 \left(\frac{3}{2}x^2 + \frac{7}{2}x + 2 \right) \mathrm{d}x = \left[\frac{1}{2}x^3 + \frac{7}{4}x^2 + 2x \right]_{-\frac{4}{3}}^0 = \frac{20}{27}.\end{aligned}$$

In conclusion,

$$\int_{\Omega} (1 + x)\ \mathrm{d}x\,\mathrm{d}y = \int_{\Omega_1} (1 + x)\ \mathrm{d}x\,\mathrm{d}y + \int_{\Omega_2} (1 + x)\ \mathrm{d}x\,\mathrm{d}y = \frac{28}{3} + \frac{20}{27} = \frac{272}{27}.$$

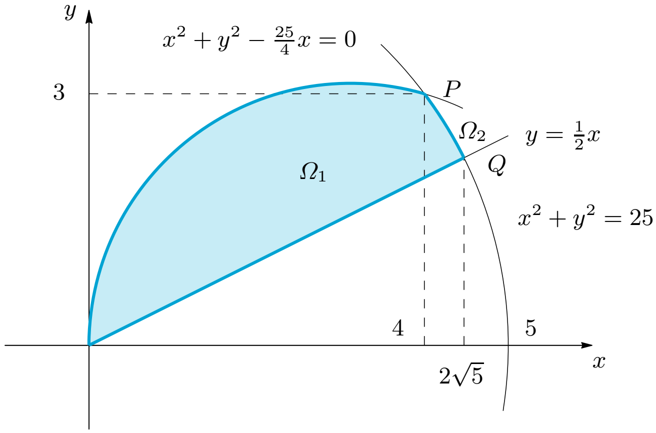


Figure 8.15. The set Ω relative to Example 8.22 ii)

ii) Consider

$$\int_\Omega \frac{y}{x+1}\,\mathrm{d}x\,\mathrm{d}y\,,$$

with Ω bounded by the circles $x^2+y^2=25$, $x^2+y^2-\frac{25}{4}x=0$ and the line $y=\frac{1}{2}x$ (Fig. 8.15). The first curve has centre in the origin and radius 5, the second in $(\frac{25}{8},0)$ and radius $\frac{25}{8}$. They meet at $P=(4,3)$ in the first quadrant; moreover, the line $y=\frac{1}{2}x$ intersects $x^2+y^2=25$ in $Q=(2\sqrt{5},\sqrt{5})$. The region is normal with respect to both axes; we therefore integrate by dividing Ω in two parts Ω_1, Ω_2, whose horizontal projections on the x-axis are $[0,4]$ and $[4,2\sqrt{5}]$:

$$\begin{aligned}
\int_\Omega \frac{y}{x+1}\,\mathrm{d}x\,\mathrm{d}y &= \int_{\Omega_1} \frac{y}{x+1}\,\mathrm{d}x\,\mathrm{d}y + \int_{\Omega_2} \frac{y}{x+1}\,\mathrm{d}x\,\mathrm{d}y \\
&= \int_0^4 \left(\int_{x/2}^{\sqrt{\frac{25}{4}x-x^2}} \frac{y}{x+1}\,\mathrm{d}y\right)\mathrm{d}x + \int_4^{2\sqrt{5}} \left(\int_{x/2}^{\sqrt{25-x^2}} \frac{y}{x+1}\,\mathrm{d}y\right)\mathrm{d}x \\
&= \frac{1}{2}\int_0^4 \frac{25x-5x^2}{x+1}\,\mathrm{d}x + \frac{1}{8}\int_4^{2\sqrt{5}} \frac{100-5x^2}{x+1}\,\mathrm{d}x \\
&= \frac{1}{8}\int_0^4 \Big(-5x+30-\frac{30}{x+1}\Big)\,\mathrm{d}x + \frac{1}{8}\int_4^{2\sqrt{5}} \Big(-5x+5+\frac{95}{x+1}\Big)\,\mathrm{d}x \\
&= \frac{5}{8}\big(10+2\sqrt{5}-25\log 5+19\log(1+2\sqrt{5})\big)\,. \qquad \square
\end{aligned}$$

8.3 Changing variables in double integrals

This section deals with the analogue of the substitution method for one-dimensional integrals. This generalisation represents an important computational tool, besides furnishing an alternative way to understand the Jacobian determinant of plane transformations.

Consider a measurable region $\Omega \subset \mathbb{R}^2$ and a continuous map f that is bounded on it. As f is integrable on Ω, we may ask how $\int_\Omega f$ varies if we change variables in the plane.

Precisely, retaining the notation of Sect. 6.6, let Ω' be a measurable region and $\mathbf{\Phi} : \Omega' \to \Omega$, $(x, y) = \mathbf{\Phi}(u, v)$ a variable change, like in Definition 6.30. We want to write the integral of $f(x, y)$ over Ω as integral over Ω' by means of the composite map $\tilde{f} = f \circ \mathbf{\Phi}$, i.e., $\tilde{f}(u, v) = f\big(\mathbf{\Phi}(u, v)\big)$, defined on Ω'. To do that though, we ought to recall the integral is defined using the areas of elementary sets, such as rectangles, into which the domain is divided. It becomes thus relevant to understand how areas change when passing from Ω' to Ω via $\mathbf{\Phi}$.

So let us start with a rectangle B' in Ω', whose sides are parallel to the axes u and v and whose vertices are $\boldsymbol{u}_0$, $\boldsymbol{u}_1 = \boldsymbol{u}_0 + \Delta u\, \boldsymbol{e}_1$, $\boldsymbol{u}_2 = \boldsymbol{u}_0 + \Delta v\, \boldsymbol{e}_2$ and $\boldsymbol{u}_3 = \boldsymbol{u}_0 + \Delta u\, \boldsymbol{e}_1 + \Delta v\, \boldsymbol{e}_2$ (Fig. 8.16, left); here Δu,Δv denote positive, and small, increments. The area of B' is $|B'| = \Delta u \Delta v$, clearly.

Denote by B the image of B' under $\mathbf{\Phi}$ (Fig. 8.16, right); it has sides along the coordinate lines of $\mathbf{\Phi}$ and vertices given by $\boldsymbol{x}_i = \mathbf{\Phi}(\boldsymbol{u}_i)$, $i = 0, \ldots, 3$. In order to express the area of B in terms of B' we need to make some approximations, justified by the choice of sufficiently small Δu and Δv. First of all, let us approximate B with the parallelogram B^p with three vertices at $\boldsymbol{x}_0$, $\boldsymbol{x}_1$, $\boldsymbol{x}_2$. It is known that its area is the absolute value of the cross product of any two adjacent sides:

$$|B| \sim |B^p| = \|(\boldsymbol{x}_1 - \boldsymbol{x}_0) \wedge (\boldsymbol{x}_2 - \boldsymbol{x}_0)\| \,.$$

On the other hand, the Taylor expansion of first order at $\boldsymbol{u}_0$ gives, also using (6.25),

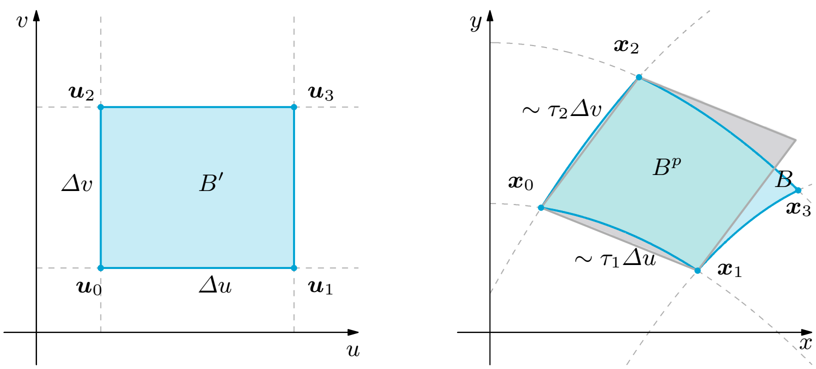


Figure 8.16. How a variable change transforms a rectangle

$$\boldsymbol{x}_1 - \boldsymbol{x}_0 = \boldsymbol{\Phi}(\boldsymbol{u}_1) - \boldsymbol{\Phi}(\boldsymbol{u}_0) \sim \frac{\partial \boldsymbol{\Phi}}{\partial u}(\boldsymbol{u}_0)\Delta u = \boldsymbol{\tau}_1 \Delta u\,,$$

and similarly

$$\boldsymbol{x}_2 - \boldsymbol{x}_0 \sim \boldsymbol{\tau}_2 \Delta v\,, \qquad \text{with} \quad \boldsymbol{\tau}_2 = \frac{\partial \boldsymbol{\Phi}}{\partial v}(\boldsymbol{u}_0)\,.$$

Therefore

$$|B| \sim \|\boldsymbol{\tau}_1 \wedge \boldsymbol{\tau}_2\| \Delta u \Delta v = \|\boldsymbol{\tau}_1 \wedge \boldsymbol{\tau}_2\|\, |B'|\,.$$

Now (6.27) yields

$$\|\boldsymbol{\tau}_1 \wedge \boldsymbol{\tau}_2\| = |\det \boldsymbol{J\Phi}(\boldsymbol{u}_0)|\,,$$

so that

$$|B| \sim |\det \boldsymbol{J\Phi}(\boldsymbol{u}_0)|\, |B'|\,. \tag{8.10}$$

Up to infinitesimals of order greater than Δu and Δv then, the quantity $|\det \boldsymbol{J\Phi}|$ represents the ratio between the areas of two (small) surface elements $B' \subset \Omega'$, $B \subset \Omega$, which correspond under $\boldsymbol{\Phi}$. This relationship is usually written symbolically as

$$\boxed{\mathrm{d}x\,\mathrm{d}y = |\det \boldsymbol{J\Phi}(u,v)|\,\mathrm{d}u\,\mathrm{d}v\,,} \tag{8.11}$$

where $\mathrm{d}u\,\mathrm{d}v$ is the area of the 'infinitesimal' surface element in Ω' and $\mathrm{d}x\,\mathrm{d}y$ the area of the image in Ω.

Examples 8.23

i) Call $\boldsymbol{\Phi}$ the affine transformation $\boldsymbol{\Phi}(\boldsymbol{u}) = \boldsymbol{A}\boldsymbol{u} + \boldsymbol{b}$, where $\boldsymbol{A}$ is a non-singular matrix. Then B is a parallelogram of vertices $\boldsymbol{x}_0$, $\boldsymbol{x}_1$, $\boldsymbol{x}_2$, $\boldsymbol{x}_3$, and concides with B^p, so all approximations are actually exact. In particular, the position vectors along the sides at $\boldsymbol{x}_0$ read

$$\boldsymbol{x}_1 - \boldsymbol{x}_0 = \boldsymbol{a}_1 \Delta u\,, \qquad \boldsymbol{x}_2 - \boldsymbol{x}_0 = \boldsymbol{a}_2 \Delta v\,,$$

where $\boldsymbol{a}_1$ and $\boldsymbol{a}_2$ are the column vectors of $\boldsymbol{A}$. Hence

$$|B| = \|\boldsymbol{a}_1 \wedge \boldsymbol{a}_2\| \Delta u \Delta v = |\det \boldsymbol{A}| \Delta u \Delta v = |\det \boldsymbol{J\Phi}|\, |B'|\,,$$

and (8.10) is an equality.

ii) Consider the transformation $\boldsymbol{\Phi} : (r,\theta) \mapsto (r\cos\theta, r\sin\theta)$ relative to polar coordinates in the plane. The image under $\boldsymbol{\Phi}$ of a rectangle $B' = [r_0, r_0 + \Delta r] \times [\theta_0, \theta_0 + \Delta\theta]$ in the $r\theta$-plane is the region of Fig. 8.17. It has height Δr in the radial direction and base $\sim r\Delta\theta$ along the angle direction; but since polar coordinates are orthogonal,

$$|B| \sim r\Delta r \Delta\theta = r|B'|\,.$$

This is the form that (8.10) takes in the case at hand, because indeed $r = \det \boldsymbol{J\Phi}(r,\theta) > 0$ (recall (6.29)). □

After this detour we are ready to integrate. If $\mathcal{D}' = \{B'_i\}_{i\in I}$ is a partition of Ω' into rectangular elements, $\mathcal{D} = \{B_i = \boldsymbol{\Phi}(B'_i)\}_{i\in I}$ will be a partition of Ω in rectangular-like regions; denote by $|\det \boldsymbol{J\Phi}_i|$ the value appearing in (8.10) for the

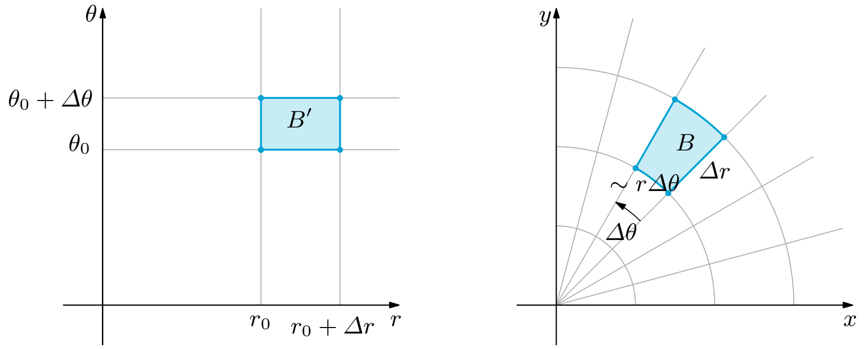


Figure 8.17. Area element in polar coordinates

element B_i. If f_i is an approximation of f on B_i (consequently, an approximation of the transform $\tilde{f} = f \circ \mathbf{\Phi}$ on B_i' as well), then by using (8.10) on each element we obtain

$$\sum_{i\in I} f_i |B_i| \sim \sum_{i\in I} f_i |\det \boldsymbol{J\Phi}_i|\,|B_i'|\,.$$

Now refining the partition of Ω' further and further, by taking rectangles of sides Δu, Δv going to 0, one can prove that the sums on the right converge to $\int_{\Omega'} f\big(\mathbf{\Phi}(u,v)\big)|\det \boldsymbol{J\Phi}(u,v)|\,\mathrm{d}u\,\mathrm{d}v$, while those on the left to $\int_{\Omega} f(x,y)\,\mathrm{d}x\,\mathrm{d}y$, provided f is continuous and bounded on Ω.

All the above discussion should justify, though heuristically, the following key result (clearly, everything can be made rigorous).

Theorem 8.24 *Let $\mathbf{\Phi} : \Omega' \to \Omega$ be a change of variables between measurable regions Ω', Ω in $\mathbb{R}^2$, as in Definition 6.30, with components*

$$x = \varphi(u,v) \qquad \textit{and} \qquad y = \psi(u,v).$$

If f is a continuous and bounded map on Ω,

$$\int_{\Omega} f(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_{\Omega'} f\big(\varphi(u,v), \psi(u,v)\big)|\det \boldsymbol{J\Phi}(u,v)|\,\mathrm{d}u\,\mathrm{d}v\,. \tag{8.12}$$

In the applications, it may be more convenient to use the inverse transformation $(u,v) = \mathbf{\Psi}(x,y)$, such that $\Omega' = \mathbf{\Psi}(\Omega)$ and $\det \boldsymbol{J\Phi}(u,v) = \big(\det \boldsymbol{J\Psi}(x,y)\big)^{-1}$.

Let us make this formula explicit by considering two special changes of variables, namely an affine map and the passage to polar coordinates.

For *affine transformations*, set $\boldsymbol{x} = \mathbf{\Phi}(\boldsymbol{u}) = \boldsymbol{A}\boldsymbol{u} + \boldsymbol{b}$ where $\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ and $\boldsymbol{b} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$; hence $x = \varphi(u,v) = a_{11}u + a_{12}v + b_1$ and $y = \psi(u,v) = a_{21}u +$

$a_{22}v + b_2$. If Ω is measurable and $\Omega' = \boldsymbol{\Phi}^{-1}(\Omega)$, we have

$$\boxed{\int_{\Omega} f(x,y)\,\mathrm{d}x\,\mathrm{d}y = |\det \boldsymbol{A}| \int_{\Omega'} f(a_{11}u + a_{12}v + b_1, a_{21}u + a_{22}v + b_2)\,\mathrm{d}u\,\mathrm{d}v\,.}$$

As for *polar coordinates*, $x = \varphi(r,\theta) = r\cos\theta$, $y = \psi(r,\theta) = r\sin\theta$, and $\det \boldsymbol{J\Phi}(r,\theta) = r$. Therefore, setting again $\Omega' = \boldsymbol{\Phi}^{-1}(\Omega)$, we obtain

$$\boxed{\int_{\Omega} f(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_{\Omega'} f(r\cos\theta, r\sin\theta)\, r\,\mathrm{d}r\,\mathrm{d}\theta\,.}$$

Examples 8.25

i) The map

$$f(x,y) = (x^2 - y^2)\log\big(1 + (x+y)^4\big)$$

is defined on $\Omega = \{(x,y) \in \mathbb{R}^2 : x > 0,\ 0 < y < 2 - x\}$, see Fig. 8.18, right. The map is continuous and bounded on Ω, so $\int_{\Omega} f(x,y)\,\mathrm{d}x\,\mathrm{d}y$ exists. The expression of f suggests defining

$$u = x + y \qquad \text{and} \qquad v = x - y\,,$$

which entails we are considering the linear map

$$x = \varphi(u,v) = \frac{u+v}{2}\,, \qquad y = \psi(u,v) = \frac{u-v}{2}$$

defined by $\boldsymbol{A} = \begin{pmatrix} 1/2 & 1/2 \\ 1/2 & -1/2 \end{pmatrix}$ with $|\det \boldsymbol{A}| = 1/2$. The pre-image of Ω under $\boldsymbol{\Phi}$ is the set

$$\Omega' = \{(u,v) \in \mathbb{R}^2 : 0 < u < 2,\ -u < v < u\}$$

of Fig. 8.18, left.

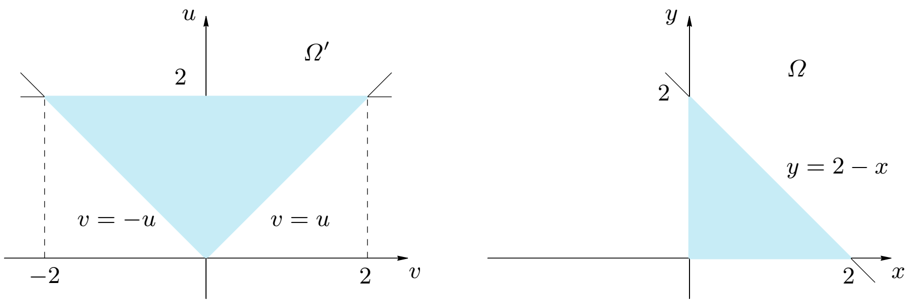


Figure 8.18. The sets $\Omega = \{(x,y) \in \mathbb{R}^2 : x > 0,\ 0 < y < 2 - x\}$ (right) and $\Omega' = \{(u,v) \in \mathbb{R}^2 : 0 < u < 2,\ -u < v < u\}$ (left)

Then

$$\int_{\Omega}(x^2-y^2)\log\big(1+(x+y)^4\big)\,\mathrm{d}x\,\mathrm{d}y = \frac{1}{2}\int_{\Omega'} uv\log(1+u^4)\,\mathrm{d}u\,\mathrm{d}v$$

$$= \frac{1}{2}\int_0^2\left(\int_{-u}^{u} v\,\mathrm{d}v\right)u\log(1+u^4)\,\mathrm{d}u$$

$$= \frac{1}{4}\int_0^2\Big[v^2\Big]_{v=-u}^{v=u} u\log(1+u^4)\,\mathrm{d}u = 0\,.$$

ii) Consider

$$f(x,y) = \frac{1}{1+x^2+y^2}$$

over

$$\Omega = \{(x,y)\in\mathbb{R}^2 : 0<y<\sqrt{3}x,\ 1<x^2+y^2<4\}.$$

Passing to polar coordinates gives

$$\Omega' = \{(r,\theta) : 1<r<2,\ 0<\theta<\frac{\pi}{3}\}$$

(Fig. 8.19), so

$$\int_{\Omega} f(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_{\Omega'}\frac{1}{1+r^2}r\,\mathrm{d}r\,\mathrm{d}\theta = \int_0^{\pi/3}\left(\int_1^2\frac{r}{1+r^2}\,\mathrm{d}r\right)\mathrm{d}\theta$$

$$= \left(\int_0^{\pi/3}\mathrm{d}\theta\right)\left(\int_1^2\frac{r}{1+r^2}\,\mathrm{d}r\right) = \frac{\pi}{6}\log\frac{5}{2}. \qquad \square$$

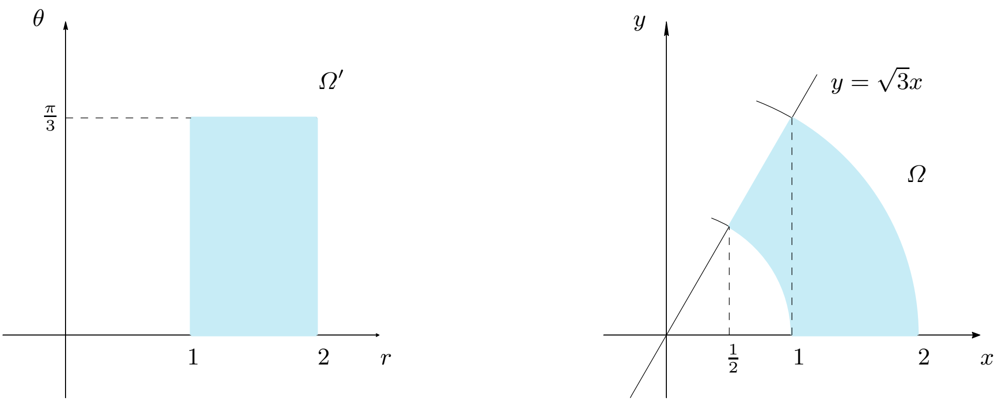


Figure 8.19. The sets $\Omega = \{(x,y)\in\mathbb{R}^2 : 0<y<\sqrt{3}x,\ 1<x^2+y^2<4\}$ (right) and $\Omega' = \{(r,\theta) : 1<r<2,\ 0<\theta<\frac{\pi}{3}\}$ (left)

8.4 Multiple integrals

Since the definition of multiple integral – or n-dimensional integral – for a map of $n \geq 3$ real variables closely resembles the one just seen for 2 variables, we merely point out the differences with the previous situations due to the increased dimension in the case $n = 3$.

The role of plane rectangles is now taken by parallelepipeds in space. The generic set $B = [a_1, b_1] \times [a_2, b_2] \times [a_3, b_3]$ can be broken up into the product of partitions of the intervals $[a_1, b_1]$, $[a_2, b_2]$, $[a_3, b_3]$, using the points $\{x_0, x_1, \dots, x_p\}$, $\{y_0, y_1, \dots, y_q\}$ and $\{z_0, z_1, \dots, z_r\}$ respectively. The solid B is thus the union of parallelepipeds

$$B_{hk\ell} = [x_{h-1}, x_h] \times [y_{k-1}, y_k] \times [z_{\ell-1}, z_\ell].$$

Take $f : B \to \mathbb{R}$ bounded; the lower and upper sums of f over B relative to the above partition $\mathcal{D}$ are, by definition,

$$s = s(\mathcal{D}, f) = \sum_{\ell=1}^{r} \sum_{k=1}^{q} \sum_{h=1}^{p} m_{hk\ell}(x_h - x_{h-1})(y_k - y_{k-1})(z_\ell - z_{\ell-1}) ,$$

$$S = S(\mathcal{D}, f) = \sum_{\ell=1}^{r} \sum_{k=1}^{q} \sum_{h=1}^{p} M_{hk\ell}(x_h - x_{h-1})(y_k - y_{k-1})(z_\ell - z_{\ell-1}) ,$$

where

$$m_{hk\ell} = \inf_{(x,y,z)\in B_{hk\ell}} f(x, y, z) , \qquad M_{hk\ell} = \sup_{(x,y,z)\in B_{hk\ell}} f(x, y, z) .$$

A map f is said **integrable on** B if

$$\inf_{\mathcal{D}} S(\mathcal{D}, f) = \sup_{\mathcal{D}} s(\mathcal{D}, f);$$

such value, called the (**triple**) **integral of** f **on** B, is denoted by one of the symbols

$$\int_B f , \qquad \iiint_B f , \qquad \int_B f(x, y, z)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z .$$

Here, as well, continuity guarantees integrability. One also has the analogue of Theorem 8.5.

Example 8.26

Compute $\int_B xyz\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z$ over $B = [0, 1] \times [-1, 2] \times [0, 2]$. We have

$$\begin{aligned}\int_B xyz\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z &= \int_0^2 \left(\int_{-1}^2 \left(\int_0^1 xyz\,\mathrm{d}x\right) \mathrm{d}y\right) \mathrm{d}z \\ &= \frac{1}{2}\int_0^2 \left(\int_{-1}^2 yz\,\mathrm{d}y\right) \mathrm{d}z = \frac{3}{4}\int_0^2 z\,\mathrm{d}z = \frac{3}{2}.\end{aligned}$$

□

To be able to integrate over bounded sets $\Omega \subset \mathbb{R}^3$, it is necessary to define what 'measurable' means in space. In analogy to Definition 8.8, we substitute the rectangle B with a parallelepiped containing Ω, and put

$$|\Omega| = \int_\Omega \chi_\Omega = \int_B \chi_\Omega \,.$$

This is independent of the choice of B. The familiar sets of Euclidean geometry, like spheres, cylinders, polyhedra and so on, turn out to be all measurable. This notion of measure is indeed the volume, so we shall sometimes write $|\Omega| = \mathrm{vol}(\Omega)$. Definition 8.9, Theorem 8.10 and Property 8.12 hold also in higher dimensions. For instance, if $D \subset \mathbb{R}^2$ is measurable and $g : D \to \mathbb{R}$ integrable on D, the graph has 3-dimensional measure equal zero. In particular, the traces of regular surfaces have, as subsets of $\mathbb{R}^3$, zero measure.

Take $f : \Omega \to \mathbb{R}$ bounded, with Ω measurable. Given an arbitrary parallelepiped B containing Ω, we extend f by setting it to zero on $B \setminus \Omega$, and call the extension $\tilde{f}$. One says f is **integrable on** Ω if $\tilde{f}$ is integrable on B, in which case

$$\int_\Omega f = \int_B \tilde{f}\,;$$

this number, not depending on B, is the (triple) **integral of** f **on** Ω, also denoted by

$$\iiint_\Omega f\,, \qquad \int_\Omega f(x,y,z)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\,, \qquad \int_\Omega f\,\mathrm{d}\Omega\,.$$

Again, Definition 8.14 and Theorem 8.15 adapt to the present situation, and similarly happens for Theorem 8.20.

A completely analogous construction leads to multiple integrals in dimension $n > 3$.

We would like to find examples of regions in space where integrals are easy to compute explicitly. These should work as normal sets did in the plane, i.e., reducing triple integrals to lower-dimensional ones. We consider in detail normal regions for the z-axis, the other cases being completely similar.

Definition 8.27 *We call $\Omega \subset \mathbb{R}^3$ a* **normal** *set* **for** z *(normal with respect to the z-axis, to be precise) if*

$$\Omega = \{(x,y,z) \in \mathbb{R}^3 : (x,y) \in D,\ g_1(x,y) \le z \le g_2(x,y)\}\,, \tag{8.13}$$

for a closed measurable region D in $\mathbb{R}^2$ and continuous maps g_1, $g_2 : D \to \mathbb{R}$.

Fig. 8.20 shows a few normal regions.

The boundary of a normal set has thus zero measure, hence a normal set is measurable. Triple integrals over normal domains can be reduced to iterated double

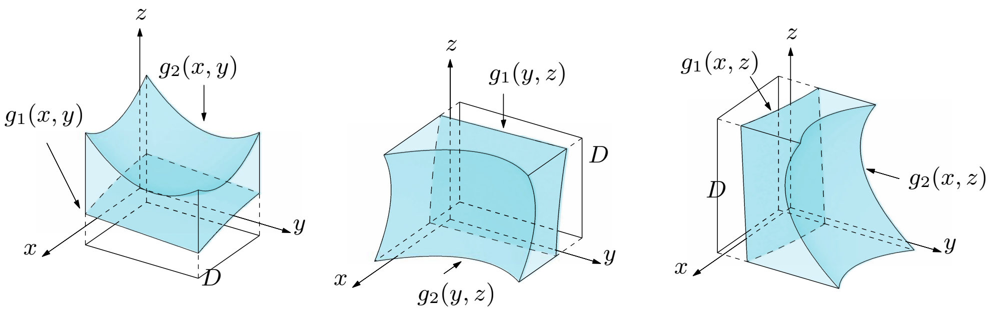


Figure 8.20. Normal sets

and/or simple integrals. For clarity, we consider normal domains for the vertical axis, and let Ω be defined as in (8.13); then we have an *iterated integral*

$$\boxed{\int_\Omega f = \int_D \left(\int_{g_1(x,y)}^{g_2(x,y)} f(x,y,z)\,\mathrm{d}z \right) \mathrm{d}x\,\mathrm{d}y\,,} \tag{8.14}$$

see Fig. 8.21, where f is integrated first along vertical segments in the domain.

Dimensional reduction is possible for other kinds of measurable sets. Suppose the measurable set $\Omega \subset \mathbb{R}^3$ is such that the z-coordinate of its points varies within an interval $[\alpha,\beta] \subset \mathbb{R}$. For any z_0 in this interval, the plane $z = z_0$ cuts Ω along the cross-section Ω_{z_0} (Fig. 8.22); so let us denote by

$$\boxed{A_{z_0} = \{(x,y) \in \mathbb{R}^2 : (x,y,z_0) \in \Omega\}}$$

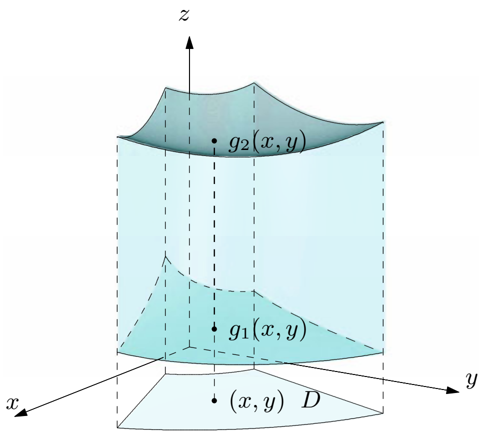


Figure 8.21. Integration along segments

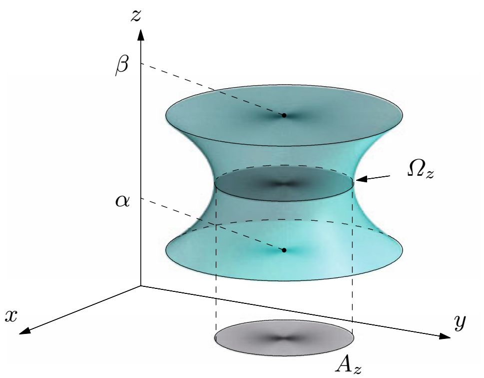


Figure 8.22. Integration over cross-sections

the xy-projection of the slice Ω_{z_0}. Thus Ω reads

$$\Omega = \{(x, y, z) \in \mathbb{R}^3 : z \in [\alpha,\beta],\ (x, y) \in A_z\} .$$

Now assume every A_z is measurable in $\mathbb{R}^2$. This is the case, for example, if the boundary of Ω is a finite union of graphs, or if Ω is convex.

Taking a continuous and bounded function $f : \Omega \to \mathbb{R}$ guarantees all integrals $\int_{A_z} f(x, y, z)\, \mathrm{d}x\, \mathrm{d}y$, $z \in [\alpha,\beta]$, exist, and one could prove that

$$\boxed{\int_\Omega f = \int_\alpha^\beta \left(\int_{A_z} f(x, y, z)\, \mathrm{d}x\, \mathrm{d}y \right) \mathrm{d}z\,;} \tag{8.15}$$

this is yet another *iterated integral*, where now f is integrated first over two-dimensional cross-sections of the domain.

Examples 8.28

i) Compute

$$\int_\Omega x\, \mathrm{d}x\, \mathrm{d}y\, \mathrm{d}z$$

where Ω denotes the tetrahedron enclosed by the planes $x = 0$, $y = 0$, $z = 0$ and $x + y + z = 1$ (Fig. 8.23). The solid is a normal region for z, because

$$\Omega = \{(x, y, z) \in \mathbb{R}^3 : (x, y) \in D,\ 0 \le z \le 1 - x - y\}$$

where

$$D = \{(x, y) \in \mathbb{R}^2 : 0 \le x \le 1,\ 0 \le y \le 1 - x\} .$$

At the same time we can describe Ω as

$$\Omega = \{(x, y, z) \in \mathbb{R}^3 : 0 \le z \le 1,\ (x, y) \in A_z\}$$

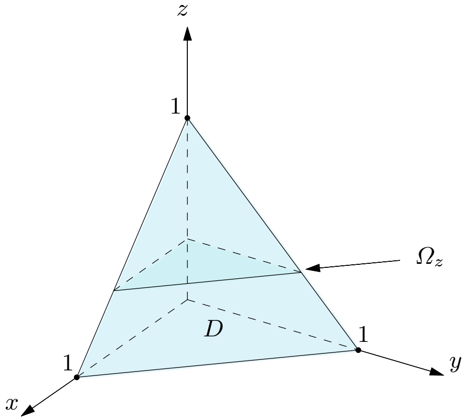


Figure 8.23. The tetrahedron relative to Example 8.28 i)

with

$$A_z = \{(x, y) \in \mathbb{R}^2 : 0 \le x \le 1 - z,\, 0 \le y \le 1 - z - x\}.$$

See Fig. 8.24 for a picture of D and A_z. Then we can integrate in z first, so that

$$\begin{aligned}\int_\Omega x \,\mathrm{d}x \,\mathrm{d}y \,\mathrm{d}z &= \int_D \left(\int_0^{1-x-y} x \,\mathrm{d}z \right) \mathrm{d}x \,\mathrm{d}y \int_D (1 - x - y)x \,\mathrm{d}x \,\mathrm{d}y \\ &= \int_0^1 \left(\int_0^{1-x} (1 - x - y)x \,\mathrm{d}y \right) \mathrm{d}x = -\frac{1}{2} \int_0^1 \Big[(1 - x - y)^2 \Big]_{y=0}^{y=1-x} x \,\mathrm{d}x \\ &= \frac{1}{2} \int_0^1 (1 - x)^2 x \,\mathrm{d}x = \frac{1}{24}.\end{aligned}$$

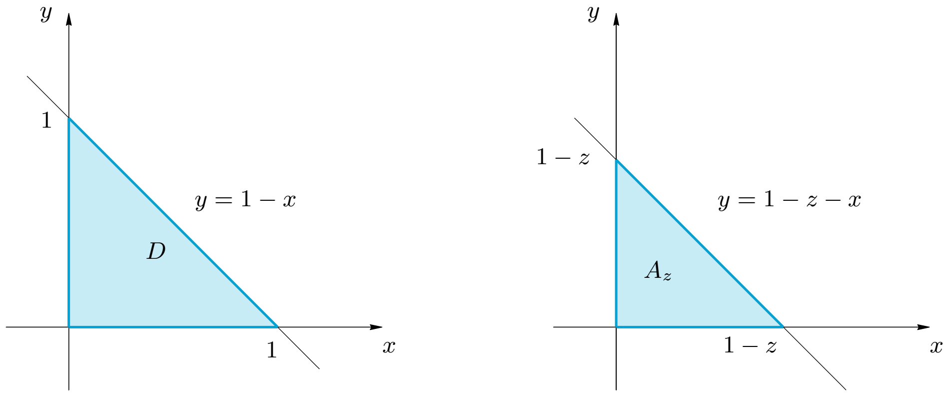


Figure 8.24. The sets D (left) and A_z (right) relative to Example 8.28 i)

But we can also compute the double integral in $dx\,dy$ first,

$$\int_\Omega x\,dx\,dy\,dz = \int_0^1 \Big(\int_{A_z} x\,dx\,dy\Big)\,dz = \int_0^1 \Big(\int_0^{1-z}\Big(\int_0^{1-x-z} x\,dy\Big)\,dx\Big)\,dz$$
$$= \int_0^1 \Big(\int_0^{1-z} x(1-x-z)\,dx\Big)\,dz = \int_0^1 \Big[\frac{1}{2}x^2(1-z) - \frac{1}{3}x^3\Big]_{x=0}^{x=1-z}\,dz$$
$$= \frac{1}{6}\int_0^1 (1-z)^3\,dz = \frac{1}{24}.$$

ii) Consider

$$\int_\Omega \sqrt{y^2+z^2}\,dx\,dy\,dz\,,$$

where Ω is bounded by the paraboloid $x = y^2+z^2$ and the plane $x = 2$ (Fig. 8.25, left). We may write

$$\Omega = \{(x,y,z) \in \mathbb{R}^3 : 0 \le x \le 2,\ (y,z) \in A_x\}$$

with slices

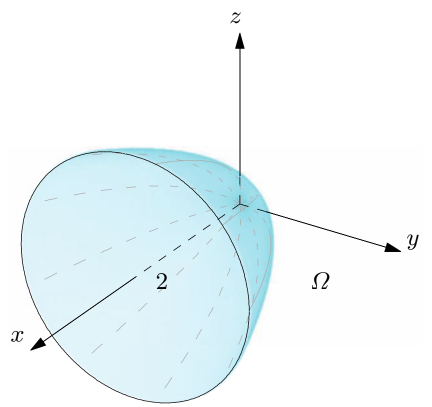
$$A_x = \{(y,z) \in \mathbb{R}^2 : y^2+z^2 \le x\}\,.$$

See Fig. 8.25, right, for the latter. We can thus integrate over A_x first, and the best option is to use polar coordinates in the plane yz, given the shape of the region and the function involved. Then

$$\int_{A_x} \sqrt{y^2+z^2}\,dy\,dz = \int_0^{2\pi}\Big(\int_0^{\sqrt{x}} r^2\,dr\Big)\,d\theta = \frac{2}{3}\pi x\sqrt{x}\,,$$

and consequently

$$\int_\Omega \sqrt{y^2+z^2}\,dx\,dy\,dz = \frac{2}{3}\pi\int_0^2 x\sqrt{x}\,dx = \frac{16}{15}\sqrt{2}\pi\,.$$



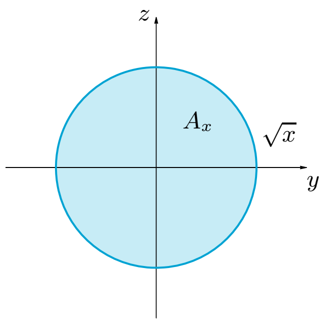


Figure 8.25. Example 8.28 ii): paraboloid (left) and the set A_x (right)

Notice the region Ω is normal for any coordinate, so the integral may be computed by reducing in different ways; the most natural choice is to look at Ω as being normal for x:

$$\Omega = \{(x,y,z) \in \mathbb{R}^3 : (y,z) \in D,\ y^2+z^2 \le x \le 2\}$$

with $D = \{(y,z) \in \mathbb{R}^2 : 0 \le x \le y^2+z^2\}$. The details are left to the reader. □

8.4.1 Changing variables in triple integrals

Theorem 8.24, governing variable changes, now reads as follows.

Theorem 8.29 *Let $\mathbf{\Phi} : \Omega' \to \Omega$, with Ω', Ω measurable in $\mathbb{R}^3$, be a change of variables on Ω, and set $\boldsymbol{x} = \mathbf{\Phi}(\boldsymbol{u})$. If f is a continuous and bounded map on Ω, we have*

$$\int_\Omega f(\boldsymbol{x})\,\mathrm{d}\Omega = \int_{\Omega'} f\big(\mathbf{\Phi}(\boldsymbol{u})\big)\,|\det \boldsymbol{J}\mathbf{\Phi}(\boldsymbol{u})|\,\mathrm{d}\Omega' . \tag{8.16}$$

The same formula holds in any dimension $n > 3$.

Let us see how a triple integral transforms if we use cylindrical or spherical coordinates.

Let $\mathbf{\Phi}$ define *cylindrical coordinates* in $\mathbb{R}^3$, see Sect. 6.6; Fig. 8.26, left, shows the corresponding volume element. If Ω is a measurable set and $\Omega' = \mathbf{\Phi}^{-1}(\Omega)$, from formula (6.41) we have

$$\boxed{\int_\Omega f(x,y,z)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \int_{\Omega'} f(r\cos\theta, r\sin\theta, t)\,r\,\mathrm{d}r\,\mathrm{d}\theta\,\mathrm{d}t\,.}$$

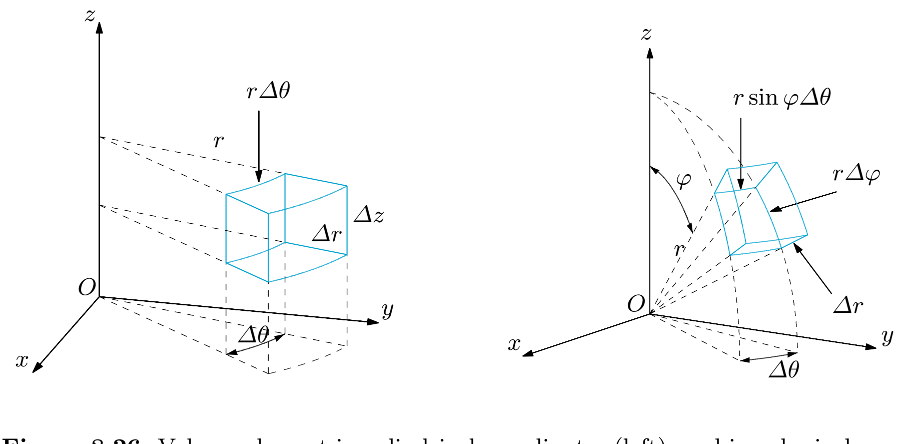


Figure 8.26. Volume element in cylindrical coordinates (left), and in spherical coordinates (right)

The variables x, y and z can be exchanged according to the need. Example 8.28 ii) indeed used the cylindrical transformation $\mathbf{\Phi}(r,\theta,t) = (t, r\cos\theta, r\sin\theta)$.

Consider the transformation $\mathbf{\Phi}$ defining *spherical coordinates* in $\mathbb{R}^3$; the volume element is shown in Fig. 8.26, right. If Ω' and Ω are related by $\Omega' = \mathbf{\Phi}^{-1}(\Omega)$, equation (6.44) gives

$$\int_{\Omega} f(x,y,z)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \int_{\Omega'} f(r\sin\varphi\cos\theta, r\sin\varphi\sin\theta, r\cos\varphi)\, r^2\sin\varphi\,\mathrm{d}r\,\mathrm{d}\varphi\,\mathrm{d}\theta\,.$$

Here, too, the roles of the variables can be exchanged.

Examples 8.30

i) Let us compute

$$\int_{\Omega} (x^2+y^2)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\,,$$

Ω being the region inside the cylinder $x^2+y^2=1$, below the plane $z=3$ and above the paraboloid $x^2+y^2+z=1$ (Fig. 8.27, left).
In cylindrical coordinates the cylinder has equation $r=1$, the paraboloid $t = 1-r^2$. Hence,

$$0 \le r \le 1\,, \qquad 0 \le \theta \le 2\pi\,, \qquad 1-r^2 \le t \le 3$$

and

$$\int_{\Omega} (x^2+y^2)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \int_0^{2\pi}\int_0^1\int_{1-r^2}^3 r^3\,\mathrm{d}t\,\mathrm{d}r\,\mathrm{d}\theta = 2\pi\int_0^1 \Big[r^3 t\Big]_{t=1-r^2}^{t=3}\mathrm{d}r$$
$$= 2\pi\int_0^1 (2r^3+r^5)\,\mathrm{d}r = \frac{4}{3}\pi\,.$$

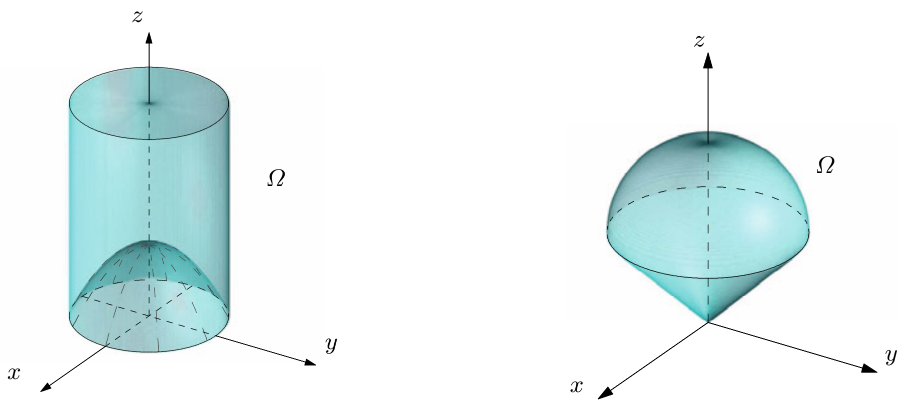


Figure 8.27. The regions relative to Examples 8.30 i) e ii)

ii) Find the volume of the solid defined by $z \geq \sqrt{x^2+y^2}$ and $x^2+y^2+z^2-z \leq 0$. This region Ω lies above the cone $z = \sqrt{x^2+y^2}$ and inside the sphere $x^2+y^2+\left(z-\frac{1}{2}\right)^2 = \frac{1}{4}$ with centre $(0,0,\frac{1}{2})$ and radius $\frac{1}{2}$ (Fig. 8.27, right).
The volume is given by $\int_\Omega \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z$, for which we use spherical coordinates. Then

$$0 \leq \theta \leq 2\pi\,, \qquad 0 \leq \varphi \leq \frac{\pi}{4}\,.$$

The bounds on r are found by noting that the sphere has equation $r^2 = r\cos\varphi$, i.e., $r = \cos\varphi$. Therefore $0 \leq r \leq \cos\varphi$. In conclusion,

$$\begin{aligned}\mathrm{vol}(\Omega) &= \int_\Omega \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \int_0^{2\pi}\int_0^{\pi/4}\int_0^{\cos\varphi} r^2\sin\varphi\,\mathrm{d}r\,\mathrm{d}\varphi\,\mathrm{d}\theta \\ &= \frac{2}{3}\pi\int_0^{\pi/4}\cos^3\varphi\sin\varphi\,\mathrm{d}\varphi = -\frac{1}{6}\pi\Big[\cos^4\varphi\Big]_{\varphi=0}^{\varphi=\pi/4} = \frac{\pi}{8}\,.\end{aligned}$$ □

8.5 Applications and generalisations

This last section is devoted to some applications of multiple integrals. We will consider solids of revolution and show how to compute their volume, and eventually extend the notion of integral to cover vector-values functions and integrals over unbounded sets in $\mathbb{R}^n$.

8.5.1 Mass, centre of mass and moments of a solid body

Double and triple integrals are used to determine the mass, the centre of gravity and the moments of inertia of plane regions or solid figures. Let us consider, for a start, a physical body whose width in the z-direction is negligible with respect to the other dimensions x and y, such as a thin plate. Suppose its mean section, in the z direction, is given by a region Ω in the plane. Call $\mu(x,y)$ the surface's density of mass (the mass per unit of area); then the body's **total mass** is

$$\boxed{m = \int_\Omega \mu(x,y)\,\mathrm{d}x\,\mathrm{d}y\,.} \tag{8.17}$$

The **centre of mass** (also known as **centre of gravity**, or **centroid**), of Ω is the point $G = (x_G, y_G)$ with coordinates

$$\boxed{x_G = \frac{1}{m}\int_\Omega x\mu(x,y)\,\mathrm{d}x\,\mathrm{d}y\,, \qquad y_G = \frac{1}{m}\int_\Omega y\mu(x,y)\,\mathrm{d}x\,\mathrm{d}y\,.} \tag{8.18}$$

Assuming the body has constant density (a homogeneous body), we have

$$x_G = \frac{1}{\text{area}(\Omega)} \int_\Omega x \,\mathrm{d}x\,\mathrm{d}y\,, \qquad y_G = \frac{1}{\text{area}(\Omega)} \int_\Omega y \,\mathrm{d}x\,\mathrm{d}y\,;$$

the coordinates of the centre of mass of Ω are the mean values (see (8.9)) of the coordinates of its generic point.

The **moment (of inertia) of Ω about** a given line r (the axis) is

$$I_r = \int_\Omega d_r^2(x,y)\mu(x,y)\,\mathrm{d}x\,\mathrm{d}y\,,$$

where $d_r(x,y)$ denotes the distance of (x,y) from r. In particular, the moments about the coordinate axes are

$$I_x = \int_\Omega y^2\mu(x,y)\,\mathrm{d}x\,\mathrm{d}y\,, \qquad I_y = \int_\Omega x^2\mu(x,y)\,\mathrm{d}x\,\mathrm{d}y\,.$$

Their sum is called **(polar) moment (of inertia) about the origin**

$$I_0 = I_x + I_y = \int_\Omega (x^2+y^2)\mu(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_\Omega d_0^2(x,y)\mu(x,y)\,\mathrm{d}x\,\mathrm{d}y\,,$$

with $d_0(x,y)$ denoting the distance between (x,y) and the origin.

Similarly, for a solid body Ω in $\mathbb{R}^3$ with mass density $\mu(x,y,z)$:

$$m = \int_\Omega \mu(x,y,z)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\,, \qquad x_G = \frac{1}{m}\int_\Omega x\mu(x,y,z)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\,,$$
$$y_G = \frac{1}{m}\int_\Omega y\mu(x,y,z)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\,, \qquad z_G = \frac{1}{m}\int_\Omega z\mu(x,y,z)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\,.$$

Its moments about the axes are

$$I_x = \int_\Omega (y^2+z^2)\mu(x,y,z)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\,,$$
$$I_y = \int_\Omega (x^2+z^2)\mu(x,y,z)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\,,$$
$$I_z = \int_\Omega (x^2+y^2)\mu(x,y,z)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\,,$$

and the moment about the origin is

$$I_0 = I_x + I_y + I_z = \int_\Omega (x^2+y^2+z^2)\mu(x,y)\,\mathrm{d}x\,\mathrm{d}y\,.$$

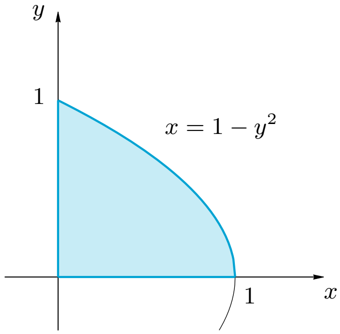


Figure 8.28. The set Ω relative to Example 8.31

Example 8.31

Consider a thin plate Ω in the first quadrant bounded by the parabola $x = 1-y^2$ and the axes. Knowing its density $\mu(x,y) = y$, we want to compute the above quantities.
The region Ω is represented in Fig. 8.28. We have

$$m = \int_\Omega y \, \mathrm{d}x \, \mathrm{d}y = \int_0^1 \left(\int_0^{1-y^2} \mathrm{d}x \right) y \, \mathrm{d}y = \int_0^1 y(1-y^2) \, \mathrm{d}y = \frac{1}{4},$$

$$x_G = 4 \int_\Omega xy \, \mathrm{d}y \, \mathrm{d}x = 4 \int_0^1 \left(\int_0^{1-y^2} x \, \mathrm{d}x \right) y \, \mathrm{d}y = 2 \int_0^1 y(1-y^2)^2 \, \mathrm{d}y = \frac{1}{3},$$

$$y_G = 4 \int_\Omega y^2 \, \mathrm{d}y \, \mathrm{d}x = 4 \int_0^1 \left(\int_0^{1-y^2} \mathrm{d}x \right) y^2 \, \mathrm{d}y = 4 \int_0^1 y^2(1-y^2) \, \mathrm{d}y = \frac{8}{15},$$

$$I_x = \int_\Omega y^3 \, \mathrm{d}y \, \mathrm{d}x = \int_0^1 \left(\int_0^{1-y^2} \mathrm{d}x \right) y^3 \, \mathrm{d}y = \int_0^1 y^3(1-y^2) \, \mathrm{d}y = \frac{1}{12},$$

$$I_y = \int_\Omega x^2 y \, \mathrm{d}y \, \mathrm{d}x = \int_0^1 \left(\int_0^{1-y^2} x^2 \, \mathrm{d}x \right) y \, \mathrm{d}y = \frac{1}{3} \int_0^1 y(1-y^2)^3 \, \mathrm{d}y = \frac{1}{24}.$$

□

8.5.2 Volume of solids of revolution

Let Ω be obtained by rotating around the z-axis the trapezoid T determined by the function $f : [a,b] \to \mathbb{R}$, $y = f(z)$ in the yz-plane (Fig. 8.29). The region T is called the **meridian section** of Ω. The volume of Ω equals the integral $\int_\Omega \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z$. The region Ω_{z_0}, intersection of Ω with the plane $z = z_0$, is a circle of radius $f(z_0)$. Integrating over the slices Ω_z and recalling the notation of p. 325, then

$$\mathrm{vol}(\Omega) = \int_a^b \left(\int_{A_z} \mathrm{d}x \, \mathrm{d}y \right) \mathrm{d}z = \pi \int_a^b \big(f(z)\big)^2 \, \mathrm{d}z, \tag{8.19}$$

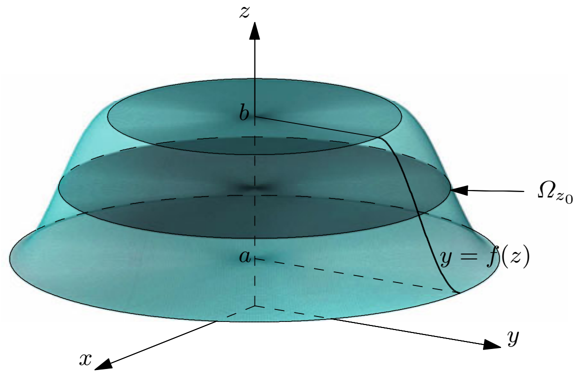


Figure 8.29. A solid of revolution

since the double integral over A_z coincides with the area of Ω_z, that is $\pi\big(f(z)\big)^2$. Formula (8.19) can be understood geometrically using the centre of mass of the section T. In fact,

$$y_G = \frac{\int_T y\,\mathrm{d}y\,\mathrm{d}z}{\int_T \mathrm{d}y\,\mathrm{d}z} = \frac{1}{\mathrm{area}(T)} \int_a^b \int_0^{f(z)} y\,\mathrm{d}y\,\mathrm{d}z = \frac{1}{2\mathrm{area}(T)} \int_a^b \big(f(z)\big)^2\,\mathrm{d}z\,.$$

Using (8.19) then,

$$y_G = \frac{\mathrm{vol}(\Omega)}{2\pi\,\mathrm{area}(T)}$$

or

$$\boxed{\mathrm{vol}(\Omega) = 2\pi y_G\,\mathrm{area}(T)\,.} \tag{8.20}$$

This proves the so-called Centroid Theorem of Pappus (variously known also as Guldin's Theorem, or Pappus–Guldin Theorem).

Theorem 8.32 *The volume of a solid of revolution is the product of the area of the meridian section times the length of the circle described by the section's centre of mass.*

This result extends to the solids of revolution whose meridian section T is not the trapezoid of a function f, but instead a measurable set in the plane. The examples that follow are exactly of this kind.

Examples 8.33

i) Compute the volume of the solid Ω, obtained revolving the triangle T of vertices $A = (1,1)$, $B = (2,1)$, $C = (1,3)$ in the plane yz around the z-axis. Fig. 8.30 shows T, whose area is clearly 1. As the line through B and C is $z = -2y + 5$, we have

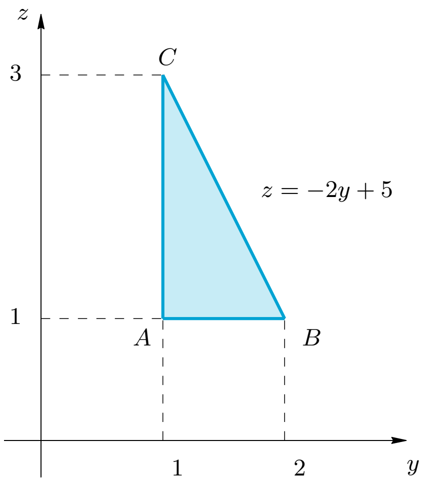


Figure 8.30. The triangle T in Example 8.33 i)

$$\begin{aligned}\mathrm{vol}(\Omega) &= 2\pi y_G = 2\pi \int_T y \,\mathrm{d}y\,\mathrm{d}z = 2\pi \int_1^2 \Big(\int_1^{-2y+5} \mathrm{d}z \Big) y\,\mathrm{d}y \\ &= 4\pi \int_1^2 (2-y)y\,\mathrm{d}y = \frac{8}{3}\pi\,.\end{aligned}$$

ii) Rotate the circle T given by $(y-y_0)^2+(z-z_0)^2 = r^2$, $0<r<y_0$, in the plane yz. The solid Ω obtained is a **torus**, see Fig. 8.31. Since area$(T) = \pi r^2$ and the centre of mass of a circle coincides with its geometric centre, we have

$$\mathrm{vol}(\Omega) = 2\pi^2 r^2 y_0\,. \qquad \square$$

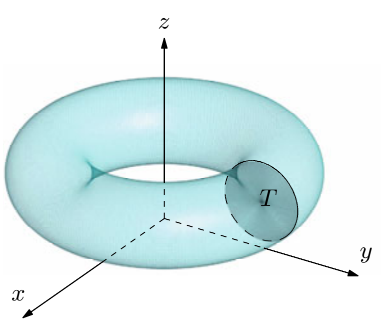


Figure 8.31. The torus relative to Example 8.33 ii)

8.5.3 Integrals of vector-valued functions

Let $\boldsymbol{f} : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$ be a vector-valued map defined on a measurable set Ω (with n, $m \geq 2$ arbitrary); let $\boldsymbol{f} = \sum_{i=1}^{m} f_i \boldsymbol{e}_i$ be its representation in the canonical basis of $\mathbb{R}^m$.

Then $\boldsymbol{f}$ is said **integrable on** Ω if all its coordinates f_i are integrable, in which case one defines the **integral of $\boldsymbol{f}$ over** Ω as the vector of $\mathbb{R}^m$

$$\boxed{\int_\Omega \boldsymbol{f}\, \mathrm{d}\Omega = \sum_{i=1}^{m} \Big(\int_\Omega f_i \, \mathrm{d}\Omega \Big) \boldsymbol{e}_i \, .}$$

For instance, if $\boldsymbol{v}$ is the velocity field of a fluid inside a measurable domain $\Omega \subset \mathbb{R}^3$, then

$$\boldsymbol{v}_\Omega = \frac{1}{|\Omega|} \int_\Omega \boldsymbol{v} \, \mathrm{d}\Omega$$

is the *mean velocity* vector in Ω.

8.5.4 Improper multiple integrals

We saw in Vol. I how to define rigorously integrals of unbounded maps and integrals over unbounded intervals. As far as multivariable functions are concerned, we shall only discuss unbounded domains of integrations. For simplicity, let $f(x, y)$ be a map of two real variables defined on the entire $\mathbb{R}^2$, positive and integrable on every disc $B_R(0)$ centred at the origin with radius R. As R grows to $+\infty$, the discs clearly become bigger and tend to cover the plane $\mathbb{R}^2$. So it is natural to expect that f will be integrable on $\mathbb{R}^2$ if the limit of the integrals of f over $B_R(0)$ exists and is finite, as $R \to +\infty$. More precisely,

Definition 8.34 *Let $f : \mathbb{R}^2 \to \mathbb{R}$ be a positive and integrable map over any disc $B_R(0)$. Then f is said* **integrable on** $\mathbb{R}^2$ *if the limit* $\lim_{R\to+\infty} \int_{B_R(0)} f(x, y) \, \mathrm{d}x \, \mathrm{d}y$ *exists and is finite. The latter, called* **improper integral of** f *on* $\mathbb{R}^2$*, is denoted by*

$$\int_{\mathbb{R}^2} f(x, y) \, \mathrm{d}x \, \mathrm{d}y = \lim_{R\to+\infty} \int_{B_R(0)} f(x, y) \, \mathrm{d}x \, \mathrm{d}y \, .$$

Remark 8.35 i) The definition holds for maps of any sign by requiring $|f(x, y)|$ to be integrable. Recall that each map can be written as the difference of two positive functions, its positive part $f_+(x, y) = \max\big(f(x, y), 0\big)$ and its negative part $f_-(x, y) = \max\big(-f(x, y), 0\big)$, i.e., $f(x, y) = f_+(x, y) - f_-(x, y)$; then

$$\int_{\mathbb{R}^2} f(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_{\mathbb{R}^2} f_+(x,y)\,\mathrm{d}x\,\mathrm{d}y - \int_{\mathbb{R}^2} f_-(x,y)\,\mathrm{d}x\,\mathrm{d}y\,.$$

ii) Other types of sets may be used to cover $\mathbb{R}^2$. Consider for example squares $Q_R(0) = [-R,R] \times [-R,R]$ of base $2R$, all centred at the origin. One can prove that f integrable implies

$$\lim_{R\to+\infty} \int_{B_R(0)} f(x,y)\,\mathrm{d}x\,\mathrm{d}y = \lim_{R\to+\infty} \int_{Q_R(0)} f(x,y)\,\mathrm{d}x\,\mathrm{d}y\,.$$

iii) If the map is not defined on all $\mathbb{R}^2$, but rather on an unbounded subset Ω, one considers the limit of $\displaystyle\int_{B_R(0)\cap\Omega} f(x,y)\,\mathrm{d}x\,\mathrm{d}y$ as $R \to +\infty$. If this is finite, again by definition we set

$$\int_{\Omega} f(x,y)\,\mathrm{d}x\,\mathrm{d}y = \lim_{R\to+\infty} \int_{B_R(0)\cap\Omega} f(x,y)\,\mathrm{d}x\,\mathrm{d}y\,.$$

□

Examples 8.36

i) Consider $f(x,y) = \dfrac{1}{\sqrt{(3+x^2+y^2)^3}}$, a map defined on $\mathbb{R}^2$, positive and continuous. The integral on the disc $B_R(0)$ can be computed in polar coordinates:

$$\begin{aligned}\int_{B_R(0)} f(x,y)\,\mathrm{d}x\,\mathrm{d}y &= \int_0^{2\pi}\int_0^R \frac{r}{\sqrt{(3+r^2)^3}}\,\mathrm{d}r\,\mathrm{d}\theta \\ &= -2\pi\Big[(3+r^2)^{-1/2}\Big]_0^R = 2\pi\Big(\frac{1}{\sqrt{3}} - \frac{1}{\sqrt{3+R^2}}\Big)\,.\end{aligned}$$

Therefore

$$\int_{\mathbb{R}^2} f(x,y)\,\mathrm{d}x\,\mathrm{d}y = \lim_{R\to+\infty} 2\pi\Big(\frac{1}{\sqrt{3}} - \frac{1}{\sqrt{3+R^2}}\Big) = \frac{2\pi}{\sqrt{3}}\,,$$

making f integrable on $\mathbb{R}^2$.

ii) Take $f(x,y) = \dfrac{1}{1+x^2+y^2}$. Proceeding as before, we have

$$\begin{aligned}\int_{B_R(0)} f(x,y)\,\mathrm{d}x\,\mathrm{d}y &= \int_0^{2\pi}\int_0^R \frac{r}{1+r^2}\,\mathrm{d}r\,\mathrm{d}\theta \\ &= \pi\Big[\log(1+r^2)\Big]_0^R = \pi\log(1+R^2)\,.\end{aligned}$$

Thus

$$\lim_{R\to+\infty} \int_{B_R(0)} f(x,y)\,\mathrm{d}x\,\mathrm{d}y = \lim_{R\to+\infty} \pi\log(1+R^2) = +\infty$$

so the map is not integrable on $\mathbb{R}^2$.

iii) With what we have learned we are now able to compute an integral, of paramount importance in Probability, that has to do with Gaussian density:

$$S = \int_{-\infty}^{+\infty} \mathrm{e}^{-x^2}\,\mathrm{d}x\,.$$

Let $f(x,y) = \mathrm{e}^{-x^2-y^2}$ and observe

$$\int_{\mathbb{R}^2} \mathrm{e}^{-x^2-y^2}\,\mathrm{d}x\,\mathrm{d}y = \Big(\int_{\mathbb{R}} \mathrm{e}^{-x^2}\,\mathrm{d}x\Big)\Big(\int_{\mathbb{R}} \mathrm{e}^{-y^2}\,\mathrm{d}y\Big) = S^2\,.$$

On the other hand

$$\int_{\mathbb{R}^2} \mathrm{e}^{-x^2-y^2}\,\mathrm{d}x\,\mathrm{d}y = \lim_{R\to+\infty} \int_{B_R(0)} \mathrm{e}^{-x^2-y^2}\,\mathrm{d}x\,\mathrm{d}y = \lim_{R\to+\infty} \int_0^{2\pi}\int_0^R \mathrm{e}^{-r^2} r\,\mathrm{d}r\,\mathrm{d}\theta$$

$$= \lim_{R\to+\infty} \pi\Big[-\mathrm{e}^{-r^2}\Big]_0^R = \lim_{R\to+\infty} \pi(1-\mathrm{e}^{-R^2}) = \pi\,.$$

Therefore

$$\boxed{S = \int_{-\infty}^{+\infty} \mathrm{e}^{-x^2}\,\mathrm{d}x = \sqrt{\pi}\,.}$$

□

8.6 Exercises

1. *Draw a rough picture of the region A and then compute the double integrals below:*

   a) $\displaystyle\int_A \frac{x}{x^2+y^2}\,\mathrm{d}x\,\mathrm{d}y$ where $A = [1,2]\times[1,2]$

   b) $\displaystyle\int_B xy\,\mathrm{d}x\,\mathrm{d}y$ where $B = \{(x,y)\in\mathbb{R}^2 : x\in[0,1],\ x^2\le y\le 1+x\}$

   c) $\displaystyle\int_C (x+y)\,\mathrm{d}x\,\mathrm{d}y$ where $C = \{(x,y)\in\mathbb{R}^2 : 2x^3\le y\le 2\sqrt{x}\}$

   d) $\displaystyle\int_D \sqrt{x}\,\mathrm{d}x\,\mathrm{d}y$ where $D = \{(x,y)\in\mathbb{R}^2 : y\in[0,1],\ y\le x\le \mathrm{e}^y\}$

   $\boxed{\text{e)}}$ $\displaystyle\int_E y^3\,\mathrm{d}x\,\mathrm{d}y$ where E is the triangle of vertices $(0,2)$, $(1,1)$, $(3,2)$

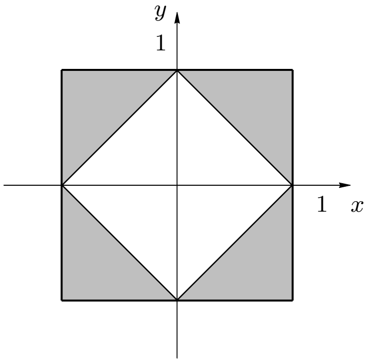
   $\boxed{\text{f)}}$ $\displaystyle\int_F \mathrm{e}^{y^2}\,\mathrm{d}x\,\mathrm{d}y$ where $F = \{(x,y)\in\mathbb{R}^2 : 0\le y\le 1,\ 0\le x\le y\}$

   $\boxed{\text{g)}}$ $\displaystyle\int_G x\cos y\,\mathrm{d}x\,\mathrm{d}y$ where G is bounded by $y=0$, $y=x^2$, $x=1$

   $\boxed{\text{h)}}$ $\displaystyle\int_H y\mathrm{e}^x\,\mathrm{d}x\,\mathrm{d}y$ where H is the triangle of vertices $(0,0)$, $(2,4)$, $(6,0)$

2. *Let $f(x,y)$ be a generic continuous map on $\mathbb{R}^2$. Find the domain of integration in the plane and change the order of integration in the following integrals:*

   $\boxed{\text{a)}}$ $\displaystyle\int_0^1\int_0^x f(x,y)\,\mathrm{d}y\,\mathrm{d}x$ $\boxed{\text{b)}}$ $\displaystyle\int_0^{\pi/2}\int_0^{\sin x} f(x,y)\,\mathrm{d}y\,\mathrm{d}x$



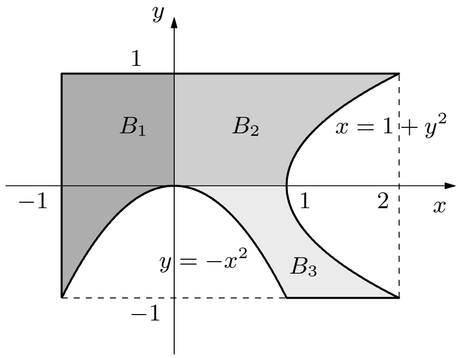


Figure 8.32. The sets A and B relative to Exercise 4

$\boxed{\text{c)}}$ $\displaystyle\int_1^2 \int_0^{\log x} f(x,y)\,\mathrm{d}y\,\mathrm{d}x$ $\boxed{\text{d)}}$ $\displaystyle\int_0^1 \int_{y^2}^{2-y} f(x,y)\,\mathrm{d}x\,\mathrm{d}y$

$\boxed{\text{e)}}$ $\displaystyle\int_0^4 \int_{y/2}^{2} f(x,y)\,\mathrm{d}x\,\mathrm{d}y$ $\boxed{\text{f)}}$ $\displaystyle\int_0^1 \int_{\arctan x}^{\pi/4} f(x,y)\,\mathrm{d}y\,\mathrm{d}x$

3. *Draw a picture of the integration domain and compute the integrals:*

$\boxed{\text{a)}}$ $\displaystyle\int_0^1 \int_{\pi y}^{\pi} \frac{\sin x}{x}\,\mathrm{d}x\,\mathrm{d}y$ $\boxed{\text{b)}}$ $\displaystyle\int_0^1 \int_y^1 \mathrm{e}^{-x^2}\,\mathrm{d}x\,\mathrm{d}y$

$\boxed{\text{c)}}$ $\displaystyle\int_0^1 \int_x^1 \frac{\sqrt{y}}{x^2+y^2}\,\mathrm{d}y\,\mathrm{d}x$ d) $\displaystyle\int_0^1 \int_{\sqrt{x}}^1 \mathrm{e}^{y^3}\,\mathrm{d}y\,\mathrm{d}x$

e) $\displaystyle\int_0^3 \int_{y^2}^9 y\cos x^2\,\mathrm{d}x\,\mathrm{d}y$ $\boxed{\text{f)}}$ $\displaystyle\int_0^1 \int_{\arcsin y}^{\pi/2} \cos x\sqrt{1+\cos^2 x}\,\mathrm{d}x\,\mathrm{d}y$

4. *Referring to Fig. 8.32, write the domains of integration as unions of normal domains, then compute the integrals:*

a) $\displaystyle\int_A x^2\,\mathrm{d}x\,\mathrm{d}y$ $\boxed{\text{b)}}$ $\displaystyle\int_B xy\,\mathrm{d}x\,\mathrm{d}y$

$\boxed{5.}$ *Compute* $\displaystyle\int_A y\,\mathrm{d}x\,\mathrm{d}y$ *over the domain A defined by*

$$R^2 \le x^2 + y^2\,, \qquad 0 \le x \le 2\,, \qquad 0 \le y \le x$$

as $R \ge 0$ varies.

6. *As $R \ge 0$ varies, calculate the double integral* $\displaystyle\int_A x\,\mathrm{d}x\,\mathrm{d}y$ *over A defined by*

$$R^2 \ge x^2 + y^2\,, \qquad 0 \le x \le 1\,, \qquad -x \le y \le 0\,.$$

$\boxed{7.}$ *Compute*

$$\int_A x \sin|x^2 - y| \, \mathrm{d}x \, \mathrm{d}y$$

where A is the unit square $[0,1] \times [0,1]$.

8. *Determine*

$$\int_A |\sin x - y| \, \mathrm{d}x \, \mathrm{d}y$$

where A is the rectangle $[0,\pi] \times [0,1]$.

9. *For any real α calculate*

$$\int_A y \mathrm{e}^{-\alpha|x-y^2|} \, \mathrm{d}x \, \mathrm{d}y$$

where A is the square $[0,1] \times [0,1]$.

10. *Draw a picture of the given region and, using suitable coordinate changes, compute the integrals:*

$\boxed{\text{a)}}$ $\displaystyle\int_A \frac{1}{x^2y^2} \, \mathrm{d}x \, \mathrm{d}y$ where A is bounded by $y = x^2$, $y = 2x^2$, and by $x = y^2$, $x = 3y^2$

$\boxed{\text{b)}}$ $\displaystyle\int_B \frac{x^5y^5}{x^3y^3+1} \, \mathrm{d}x \, \mathrm{d}y$ where B, in the first quadrant, is bounded by $y = x$, $y = 3x$, $xy = 2$, $xy = 6$

$\boxed{\text{c)}}$ $\displaystyle\int_C (3x + 4y^2) \, \mathrm{d}x \, \mathrm{d}y$ where $C = \{(x,y) \in \mathbb{R}^2 : y \geq 0,\ 1 \leq x^2 + y^2 \leq 4\}$

d) $\displaystyle\int_D xy \, \mathrm{d}x \, \mathrm{d}y$ where $D = \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 \leq 9\}$

e) $\displaystyle\int_E \mathrm{e}^{-x^2-y^2} \, \mathrm{d}x \, \mathrm{d}y$ where E is bounded by $x = \sqrt{4-y^2}$ and $x = 0$

f) $\displaystyle\int_F \arctan\frac{y}{x} \, \mathrm{d}x \, \mathrm{d}y$ where $F = \{(x,y) \in \mathbb{R}^2 : 1 \leq x^2 + y^2 \leq 4,\ |y| \leq |x|\}$

g) $\displaystyle\int_G \frac{x}{\sqrt{x^2+y^2}} \, \mathrm{d}x \, \mathrm{d}y$ where $G = \{(x,y) \in \mathbb{R}^2 : y \geq 0,\ x + y \geq 0,\ 3 \leq x^2 + y^2 \leq 9\}$

$\boxed{\text{h)}}$ $\displaystyle\int_H x \, \mathrm{d}x \, \mathrm{d}y$ where $H = \{(x,y) \in \mathbb{R}^2 : x, y \geq 0,\ x^2 + y^2 \leq 4,\ x^2 + y^2 - 2y \geq 0\}$

11. *By transforming into Cartesian coordinates, compute the following integral in polar coordinates*

$$\int_{A'} \frac{1}{\sqrt{\cos^2\theta + \sin\theta/r + 1/r^2}}\,\mathrm{d}r\,\mathrm{d}\theta$$

where $A' = \left\{(r,\theta) : 0 \le \theta \le \frac{\pi}{2},\ r \le \frac{1}{\cos\theta},\ r \le \frac{1}{\sin\theta}\right\}$.

12. *By transforming into Cartesian coordinates, compute the following integral in polar coordinates*

$$\int_{A'} \frac{\log r(\cos\theta + \sin\theta)}{\cos\theta}\,\mathrm{d}r\,\mathrm{d}\theta$$

where $A' = \left\{(r,\theta) : 0 \le \theta \le \frac{\pi}{4},\ 1 \le r\cos\theta \le 2\right\}$.

13. *Use polar coordinates to write the sum of the integrals*

$$\int_{1/\sqrt{2}}^{1}\int_{\sqrt{1-x^2}}^{x} xy\,\mathrm{d}y\,\mathrm{d}x + \int_{1}^{\sqrt{2}}\int_{0}^{x} xy\,\mathrm{d}y\,\mathrm{d}x + \int_{\sqrt{2}}^{2}\int_{0}^{\sqrt{4-x^2}} xy\,\mathrm{d}y\,\mathrm{d}x$$

as a single double integral, then compute it.

14. *Determine mass and centre of gravity of a thin triangular plate of vertices* $(0,0)$, $(1,0)$, $(0,2)$ *if its density is* $\mu(x,y) = 1 + 3x + y$.

15. *A thin plate, of the shape of a semi-disc of radius* a, *has density proportional to the distance of the centre from the origin. Find the centre of mass.*

16. *Determine the moments* I_x, I_y, I_0 *of a disc with constant density* $\mu(x,y) = \mu$, *centre at the origin and radius* a.

17. *Let* D *be a disc with unit density, centre at* $C = (a,0)$ *and radius* a. *Verify the equation*

$$I_0 = I_C + a^2\mathcal{A}$$

holds, where I_0 *and* I_C *are the moments about the origin and the centre* C, *and* $\mathcal{A}$ *is the disc's area.*

18. *Compute the moment about the origin of the thin plate* C *defined by*

$$C = \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 \le 4,\ x^2 + y^2 - x \ge 0\}$$

knowing that its density $\mu(x,y)$ *equals the distance of* (x,y) *from the origin.*

$\boxed{19.}$ *A thin plate C occupies the quarter of disc $x^2+y^2 \leq 1$ contained in the first quadrant. Find its centre of mass knowing the density at each point equals the distance of that point from the x-axis.*

20. *A thin plate has the shape of a parallelogram with vertices $(3,0)$, $(0,6)$, $(-1,2)$, $(2,-4)$. Assuming it has unit density, compute I_y, the moment about the axis y.*

21. *Determine the following multiple integrals:*

a) $\displaystyle\int_A (xy - z^3)\,dx\,dy\,dz$ where $A = [-1,1] \times [0,1] \times [0,2]$

b) $\displaystyle\int_B 2y\,dx\,dy\,dz$ where $B = \{(x,y,z) \in \mathbb{R}^3 : 0 \leq x \leq z \leq 2,\ 0 \leq y \leq \sqrt{4-z^2}\}$

c) $\displaystyle\int_C xz \sin y^5\,dx\,dy\,dz$ where $C = \{(x,y,z) \in \mathbb{R}^3 : 0 \leq x \leq y \leq 1,\ y \leq z \leq 2y\}$

$\boxed{\text{d)}}$ $\displaystyle\int_D y\,dx\,dy\,dz$ where D, in the first octant, is bounded by the planes $x+y=1$, $y+z=1$

$\boxed{\text{e)}}$ $\displaystyle\int_E y\,dx\,dy\,dz$ where E is bounded by the paraboloid $y = 4x^2+4z^2$ and the plane $y=4$

$\boxed{\text{f)}}$ $\displaystyle\int_F z\,dx\,dy\,dz$ where F, in the first octant, is bounded by the plane $y=3x$ and the cylinder $y^2+z^2=9$

22. *Write the triple integral* $\displaystyle\int_A f(x,y,z)\,dx\,dy\,dz$ *as an iterated integral in at least three different ways, with A being the solid with lateral surface:*

a) $x=0$, $x=z$, $y^2=1-z$ b) $x^2+y^2=9$, $z=0$, $z=6$

$\boxed{23.}$ *Consider the region A in the first octant bounded by the planes $x+y-z+1=0$ and $x+y=a$. Determine the real number $a>0$ so that the volume equals vol$(A) = \frac{5}{6}$.*

$\boxed{24.}$ *Compute the volume of the region A common to the cylinders $x^2+y^2 \leq 1$ and $x^2+z^2 \leq 1$.*

$\boxed{25.}$ *Compute*

$$\int_A \frac{1}{3-z}\,dx\,dy\,dz$$

over the region A defined by

$$9z \le 1 + y^2 + 9x^2\,, \qquad\qquad 0 \le z \le \sqrt{9-(y^2+9x^2)}\,.$$

$\boxed{26.}$ *The plane $x = K$ divides the tetrahedron of vertices $(1,0,0)$, $(0,2,0)$, $(0,0,3)$, $(0,0,0)$ in two regions C_1, C_2. Determine K so that the two solids obtained have the same volume.*

$\boxed{27.}$ *By suitable variable changes, compute the following integrals:*

$\boxed{\text{a)}}$ $\displaystyle\int_A \sqrt{x^2+y^2}\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z$

with A bounded by the cylinder $x^2+y^2=25$ and the planes $z=-1$, $z=2$

$\boxed{\text{b)}}$ $\displaystyle\int_B x\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z$

with B bounded by the cylinders $x^2+z^2=1$, $x^2+z^2=4$ and the planes $y=0$ and $y=z+2$

c) $\displaystyle\int_C (x+z)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z$

with $C = \{(x,y,z)\in\mathbb{R}^3 : 1\le x\le 2,\ x^2+y^2+z^2\le 4\}$

$\boxed{\text{d)}}$ $\displaystyle\int_D x\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z$

with D in the first octant and bounded by the spheres $x^2+y^2+z^2=1$ and $x^2+y^2+z^2=4$

e) $\displaystyle\int_E \sqrt{x^2+y^2+z^2}\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z$

with E bounded below by the cone $\varphi=\frac{\pi}{6}$ and above by the sphere $r=2$

$\boxed{28.}$ *Compute the following integrals with the help of a variable change:*

$\boxed{\text{a)}}$ $\displaystyle\int_{-1}^{1}\int_{-\sqrt{1-z^2}}^{\sqrt{1-z^2}}\int_{x^2+z^2}^{2-x^2-z^2}(x^2+z^2)^{3/2}\,\mathrm{d}y\,\mathrm{d}x\,\mathrm{d}z$

b) $\displaystyle\int_0^3\int_0^{\sqrt{9-y^2}}\int_{\sqrt{x^2+y^2}}^{\sqrt{18-x^2-y^2}}(x^2+y^2+z^2)\,\mathrm{d}z\,\mathrm{d}x\,\mathrm{d}y$

$\boxed{29.}$ *Calculate the integral*

$$\int_\Omega \Big(4x^2+\frac{16}{9}y^2+z^2\Big)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z$$

where Ω is the solid bounded by the ellipsoid $\dfrac{x^2}{4}+\dfrac{y^2}{9}+\dfrac{z^2}{16}=1.$

$\boxed{30.}$ *Find mass and centre of gravity of Ω, the part of cylinder $x^2+y^2 \le 1$ in the first octant bounded by the plane $z = 1$, assuming the density is $\mu(x, y, z) = x + y$.*

$\boxed{31.}$ *What is the volume of the solid Ω obtained by revolution around the z-axis of*

$$T = \{(y, z) \in \mathbb{R}^2 : 2 \le y \le 3,\ 0 \le z \le y,\ y^2 - z^2 \le 4\}\,?$$

32. *The region T, lying on the plane xy, is bounded by the curves $y = x^2$, $y = 4$, $x = 0$ with $0 \le x \le 2$. Compute the volume of Ω, obtained by revolution of T around the axis y.*

$\boxed{33.}$ *Rotate around the z-axis the region*

$$T = \{(x, z) \in \mathbb{R}^2 : \sin z < x < \pi - z,\ 0 < z < \pi\}$$

to obtain the solid Ω. Find its volume and centre of mass.

8.6.1 Solutions

1. *Double integrals:*

a) The region A is a square, and the integral equals

$$I = \frac{1}{2}\Big(7\log 2 - 3\log 5 - 2\arctan 2 - 4\arctan\frac{1}{2} + \frac{3}{2}\pi\Big)\,.$$

b) The domain B is represented in Fig. 8.33, left. The integral equals $I = \frac{5}{8}$.

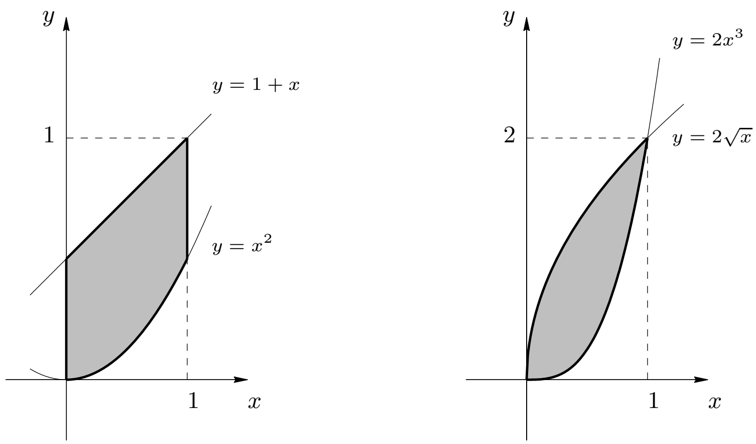


Figure 8.33. The sets B relative to Exercise 1. b) (left) and C to Exercise 1. c) (right)

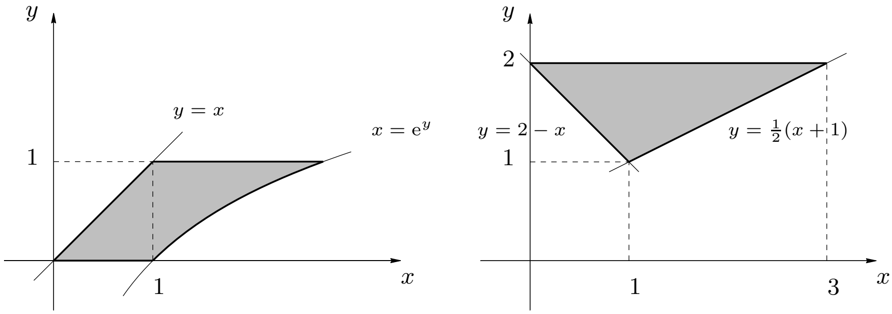


Figure 8.34. The sets D relative to Exercise 1. d) (left) and E to Exercise 1. e) (right)

c) See Fig. 8.33, right for the region C. The integral is $I = \frac{39}{35}$.

d) The region D is represented in Fig. 8.34, left. The integral's value is $I = \frac{4}{9}e^{3/2} - \frac{32}{45}$.

e) The region E can be seen in Fig. 8.34, right. The lines through the triangle's vertices are $y = 2$, $y = \frac{1}{2}(x+1)$ and $y = -x+2$. Integrating horizontally with the bounds $1 \le y \le 2$ and $2 - y \le x \le 2y - 1$ gives

$$\int_E y^3 \, dx \, dy = \int_1^2 \left(\int_{2-y}^{2y-1} y^3 \, dx \right) dy = \frac{147}{20} \, .$$

f) The region F is shown in Fig. 8.35, left. Let us integrate between $0 \le y \le 1$ and $0 \le x \le y$ to obtain

$$\int_F e^{y^2} \, dx \, dy = \int_0^1 \left(\int_0^y e^{y^2} \, dx \right) dy = \int_0^1 e^{y^2} \, [x]_0^y \, dy$$

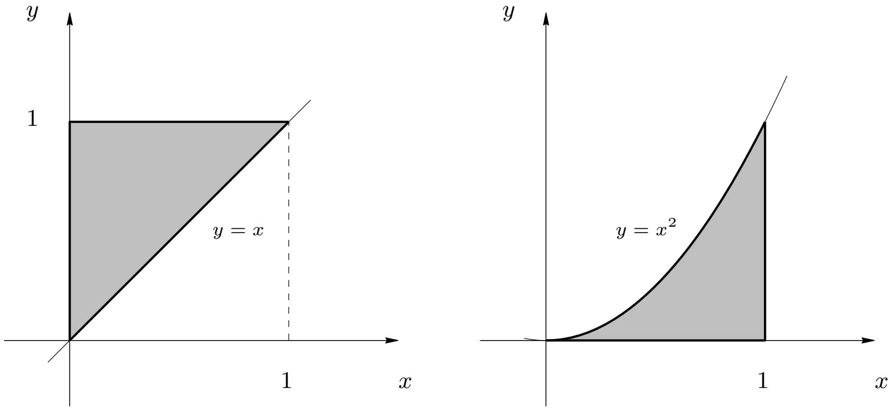


Figure 8.35. The sets F relative to Exercise 1. f) (left), and G to Exercise 1. g) (right)

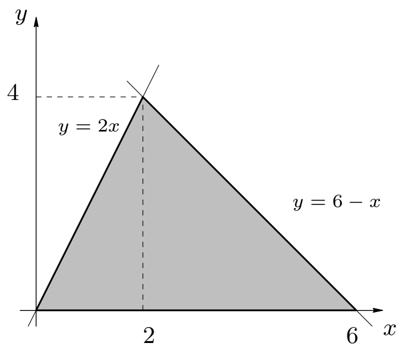


Figure 8.36. The set H relative to Exercise 1. h)

$$= \int_0^1 y e^{y^2}\,\mathrm{d}y = \left[\frac{1}{2}e^{y^2}\right]_0^1 = \frac{1}{2}(e-1)\,.$$

g) The set G is represented in Fig. 8.35, right. We integrate vertically with $0 \le x \le 1$, $0 \le y \le x^2$:

$$\int_G x\cos y\,\mathrm{d}x\,\mathrm{d}y = \int_0^1 \left(\int_0^{x^2} x\cos y\,\mathrm{d}y\right)\mathrm{d}x = \int_0^1 x\,[\sin y]_0^{x^2}\,\mathrm{d}x$$

$$= \int_0^1 x\sin x^2\,\mathrm{d}x = \left[-\frac{1}{2}\cos x^2\right]_0^1 = \frac{1}{2}(1-\cos 1)\,.$$

h) The region H is shown in Fig. 8.36. The lines passing through the vertices of the triangle are $y = 0$, $y = -x + 6$, $y = 2x$. It is convenient to integrate horizontally with $0 \le y \le 4$ and $y/2 \le x \le 6 - y$. Integrating by parts then,

$$\int_H y e^x\,\mathrm{d}x\,\mathrm{d}y = \int_0^4 \left(\int_{y/2}^{6-y} y e^x\,\mathrm{d}x\right)\mathrm{d}y = e^6 - 9e^2 - 4\,.$$

2. *Order of integration:*

a) The domain A is represented in Fig. 8.37, left. Exchanging the integration order gives

$$\int_0^1\int_0^x f(x,y)\,\mathrm{d}y\,\mathrm{d}x = \int_0^1\int_y^1 f(x,y)\,\mathrm{d}x\,\mathrm{d}y\,.$$

b) For the domain B see Fig. 8.37, right. Exchanging the integration order gives

$$\int_0^{\pi/2}\int_0^{\sin x} f(x,y)\,\mathrm{d}y\,\mathrm{d}x = \int_0^1\int_{\arcsin y}^{\pi/2} f(x,y)\,\mathrm{d}x\,\mathrm{d}y\,.$$

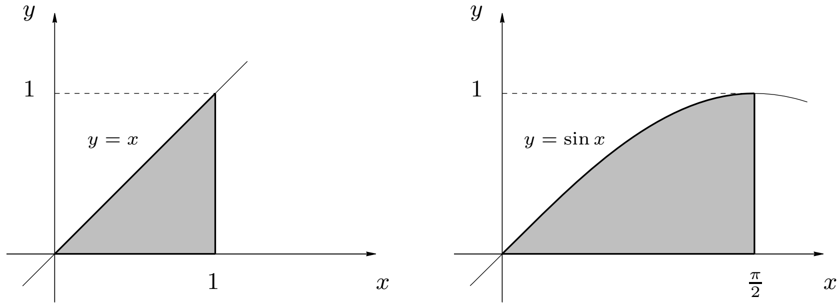


Figure 8.37. The sets A relative to Exercise 2. a) (left) and B to Exercise 2. b) (right)

c) The domain C is represented in Fig. 8.38, left. Exchanging the integration order gives

$$\int_1^2 \int_0^{\log x} f(x,y)\,\mathrm{d}y\,\mathrm{d}x = \int_0^{\log 2} \int_{\mathrm{e}^y}^2 f(x,y)\,\mathrm{d}x\,\mathrm{d}y\,.$$

d) The region D is drawn in Fig. 8.38, right. In order to integrate vertically, we must first divide D in two parts, to the effect that

$$\int_0^1 \int_{y^2}^{2-y} f(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_0^1 \int_0^{\sqrt{x}} f(x,y)\,\mathrm{d}y\,\mathrm{d}x + \int_1^2 \int_0^{2-x} f(x,y)\,\mathrm{d}y\,\mathrm{d}x\,.$$

e) The domain E is given in Fig. 8.39, left. Exchanging the integration order gives

$$\int_0^4 \int_{y/2}^2 f(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_0^2 \int_0^{2x} f(x,y)\,\mathrm{d}y\,\mathrm{d}x\,.$$

f) F is represented in Fig. 8.39, right. Exchanging the integration order gives

$$\int_0^1 \int_{\arctan x}^{\pi/4} f(x,y)\,\mathrm{d}y\,\mathrm{d}x = \int_0^{\pi/4} \int_0^{\tan y} f(x,y)\,\mathrm{d}x\,\mathrm{d}y\,.$$

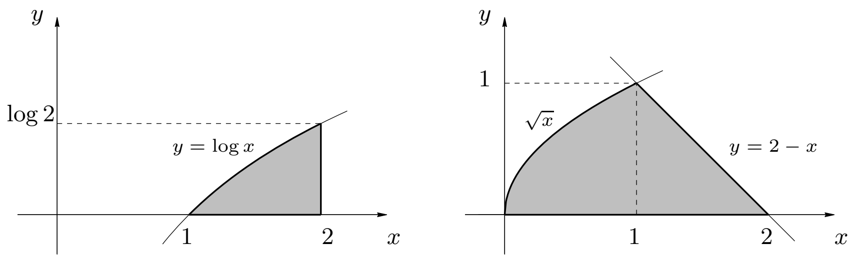


Figure 8.38. The sets C relative to Exercise 2. c) (left) and D to Exercise 2. d) (right)

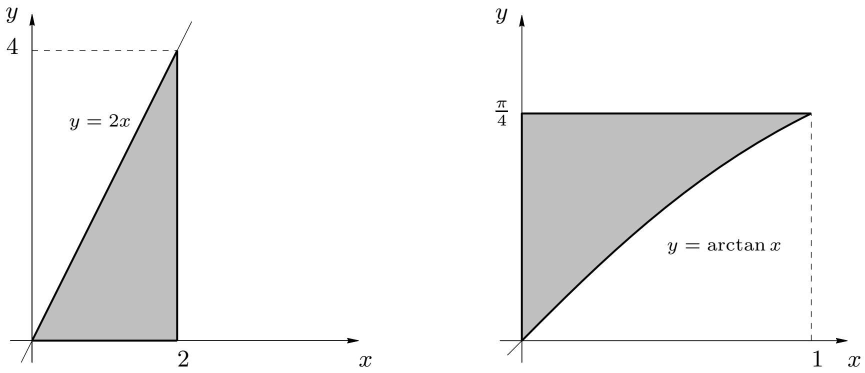


Figure 8.39. The sets E relative to Exercise 2. e) (left) and F to Exercise 2. f) (right)

3. *Double integrals:*

a) The domain A is normal for x and y, so we may write

$$\begin{aligned} A &= \{(x,y) \in \mathbb{R}^2 : 0 \le y \le 1,\ \pi\, y \le x \le \pi\} \\ &= \{(x,y) \in \mathbb{R}^2 : 0 \le x \le \pi,\ 0 \le y \le \frac{x}{\pi}\}\,. \end{aligned}$$

See Fig. 8.40, left.
Given that the integrand map is not integrable in elementary functions in x, it is necessary to exchange the order of integration. Then

$$\begin{aligned} \int_0^1 \int_{\pi y}^{\pi} \frac{\sin x}{x}\, \mathrm{d}x\, \mathrm{d}y &= \int_0^{\pi} \left(\int_0^{x/\pi} \frac{\sin x}{x}\, \mathrm{d}y \right) \mathrm{d}x = \int_0^{\pi} \frac{\sin x}{x} [y]_0^{x/\pi}\, \mathrm{d}x \\ &= \frac{1}{\pi} \int_0^{\pi} \sin x\, \mathrm{d}x = \frac{2}{\pi}\,. \end{aligned}$$

b) The domain is normal for x and y, so

$$\begin{aligned} B &= \{(x,y) \in \mathbb{R}^2 : 0 \le y \le 1,\ y \le x \le 1\} \\ &= \{(x,y) \in \mathbb{R}^2 : 0 \le x \le 1,\ 0 \le y \le x\}\,. \end{aligned}$$

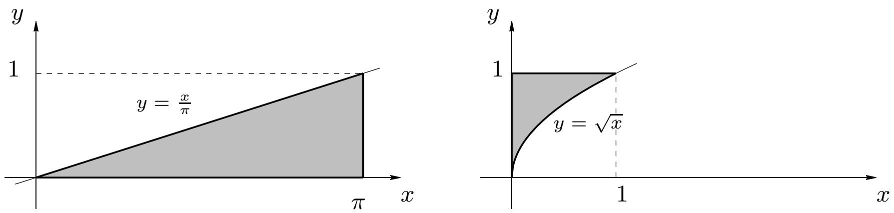


Figure 8.40. The sets A relative to Exercise 3. a) (left) and D to Exercise 3. d) (right)

It coincides with the set A of Exercise 2. a), see Fig. 8.37, left.
As in the previous exercise, the map is not integrable in elementary functions in x, so we exchange the order

$$\int_0^1 \int_y^1 \mathrm{e}^{-x^2}\,\mathrm{d}x\,\mathrm{d}y = \int_0^1 \left(\int_0^x \mathrm{e}^{-x^2}\,\mathrm{d}y\right)\mathrm{d}x = \int_0^1 x\mathrm{e}^{-x^2}\,\mathrm{d}x = \frac{1}{2}(1-\mathrm{e}^{-1})\,.$$

c) The integration domain is normal for x and y, so we write

$$\begin{aligned} C &= \{(x,y)\in\mathbb{R}^2 : 0\le x\le 1,\ x\le y\le 1\} \\ &= \{(x,y)\in\mathbb{R}^2 : 0\le y\le 1,\ 0\le x\le y\}\,. \end{aligned}$$

It is the same as the set F of Exercise 1. f), see Fig. 8.35, left.
As in the previous exercise, the map is not integrable in elementary functions in y, but we can exchange integration order

$$\begin{aligned} \int_0^1 \int_x^1 \frac{\sqrt{y}}{x^2+y^2}\,\mathrm{d}y\,\mathrm{d}x &= \int_0^1 \left(\int_0^y \frac{\sqrt{y}}{x^2+y^2}\,\mathrm{d}x\right)\mathrm{d}y \\ &= \int_0^1 \left[\sqrt{y}\,\frac{1}{y}\arctan\frac{x}{y}\right]_0^y \mathrm{d}y = \frac{\pi}{4}\int_0^1 \frac{1}{\sqrt{y}}\,\mathrm{d}y = \frac{\pi}{2}\,. \end{aligned}$$

d) The domain D is as in Fig. 8.40, right, and the integral equals $I = (\mathrm{e}-1)/3$.
e) The domain E is represented by Fig. 8.41, the integral is $I = \frac{1}{4}\sin 81$.
f) The domain is normal for x and y, so

$$\begin{aligned} F &= \{(x,y)\in\mathbb{R}^2 : 0\le y\le 1,\ \arcsin y\le x\le \frac{\pi}{2}\} \\ &= \{(x,y)\in\mathbb{R}^2 : 0\le x\le \frac{\pi}{2},\ 0\le y\le \sin x\}\,. \end{aligned}$$

It is precisely the set B from Exercise 2. b), see Fig. 8.37, right.
As in the previous exercise the map is not integrable in elementary functions in x, so we exchange orders

$$\int_0^1 \int_{\arcsin y}^{\pi/2} \cos x\sqrt{1+\cos^2 x}\,\mathrm{d}x\,\mathrm{d}y = \int_0^{\pi/2}\left(\int_0^{\sin x} \cos x\sqrt{1+\cos^2 x}\,\mathrm{d}y\right)\mathrm{d}x$$

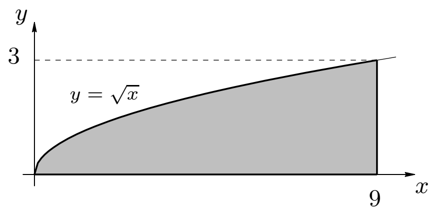


Figure 8.41. The set E relative to Exercise 3. e)

$$\int_0^{\pi/2} \sin x \cos x \sqrt{1+\cos^2 x}\, \mathrm{d}x = \frac{1}{2}\int_1^2 \sqrt{t}\, \mathrm{d}t = \frac{2\sqrt{2}-1}{3}.$$

4. *Double integrals:*

a) The integral is 1.

b) Referring to Fig. 8.32 we divide B into three subsets:

$$B_1 = \{(x,y) \in \mathbb{R}^2 : -1 \le x \le 0,\ -x^2 \le y \le 1\}$$
$$B_2 = \{(x,y) \in \mathbb{R}^2 : 0 \le y \le 1,\ 0 \le x \le 1+y^2\}$$
$$B_3 = \{(x,y) \in \mathbb{R}^2 : -1 \le y \le 0,\ \sqrt{-y} \le x \le 1+y^2\}.$$

After a little computation,

$$\begin{aligned}\int_B xy\, \mathrm{d}x\, \mathrm{d}y &= \int_{B_1} xy\, \mathrm{d}x\, \mathrm{d}y + \int_{B_2} xy\, \mathrm{d}x\, \mathrm{d}y + \int_{B_3} xy\, \mathrm{d}x\, \mathrm{d}y \\ &= \int_{-1}^0 \int_{-x^2}^1 xy\, \mathrm{d}y\, \mathrm{d}x + \int_0^1 \int_0^{1+y^2} xy\, \mathrm{d}x\, \mathrm{d}y + \int_{-1}^0 \int_{\sqrt{-y}}^{1+y^2} xy\, \mathrm{d}x\, \mathrm{d}y \\ &= 0.\end{aligned}$$

5. When $0 \le x \le 2$ the circle $x^2+y^2=R^2$ and the line $y=x$ meet at $P = \left(\frac{R}{\sqrt{2}}, \frac{R}{\sqrt{2}}\right)$. Hence for $\frac{R}{\sqrt{2}} > 2$, i.e., $R > 2\sqrt{2}$, A is empty and the integral is 0.

So let $0 \le R \le 2\sqrt{2}$, and we have to distinguish two cases: $0 \le R \le 2$ and $2 < R \le 2\sqrt{2}$. See Fig. 8.42 for A.

Take $0 \le R \le 2$: using Fig. 8.42 we can write

$$\int_A y\, \mathrm{d}x\, \mathrm{d}y = \int_{A_1} y\, \mathrm{d}x\, \mathrm{d}y + \int_{A_2} y\, \mathrm{d}x\, \mathrm{d}y$$

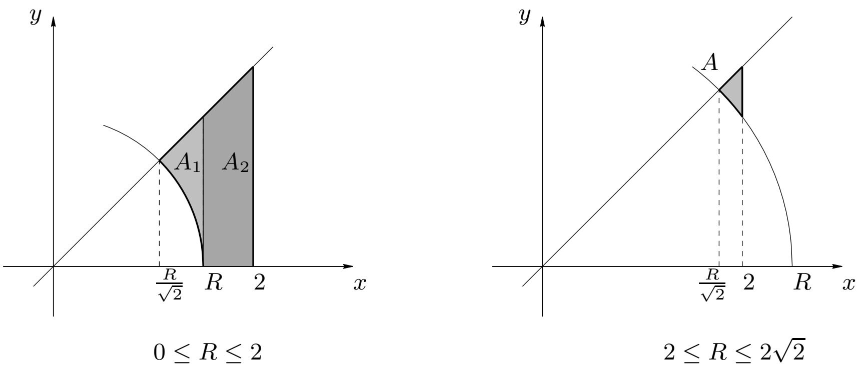


Figure 8.42. The set A relative to Exercise 5

$$= \int_{R/\sqrt{2}}^{R} \int_{\sqrt{R^2-x^2}}^{x} y \,\mathrm{d}y \,\mathrm{d}x + \int_{R}^{2} \int_{0}^{x} y \,\mathrm{d}y \,\mathrm{d}x$$

$$= \frac{1}{2} \int_{R/\sqrt{2}}^{R} (2x^2 - R^2) \,\mathrm{d}x + \frac{1}{2} \int_{R}^{2} x^2 \,\mathrm{d}x = \frac{4}{3} + \frac{1}{6} R^3(\sqrt{2} - 2) \,.$$

Now suppose $2 < R \le 2\sqrt{2}$: then

$$\int_A y \,\mathrm{d}x \,\mathrm{d}y = \int_{R/\sqrt{2}}^{2} \int_{\sqrt{R^2-x^2}}^{x} y \,\mathrm{d}y \,\mathrm{d}x$$

$$= \frac{1}{2} \int_{R/\sqrt{2}}^{2} (2x^2 - R^2) \,\mathrm{d}x = \frac{8}{3} - R^2 + \frac{1}{3\sqrt{2}} R^3 \,.$$

6. We have

$$\int_A x \,\mathrm{d}x \,\mathrm{d}y = \begin{cases} \dfrac{\sqrt{2}}{6} R^3 & \text{if } 0 \le R < 1 \,, \\ \dfrac{\sqrt{2}}{6} R^3 - \dfrac{1}{3}(R^2 - 1)^{3/2} & \text{if } 1 \le R < \sqrt{2} \,, \\ \dfrac{1}{3} & \text{if } R \ge \sqrt{2} \,. \end{cases}$$

7. The integrand changes sign according to whether $(x, y) \in A$ is above or below the parabola $y = x^2$ (see Fig. 8.43). Precisely, for any $(x, y) \in A$,

$$x \sin|x^2 - y| = \begin{cases} x \sin(x^2 - y) & \text{if } y \le x^2 \,, \\ -x \sin(x^2 - y) & \text{if } y > x^2 \,. \end{cases}$$

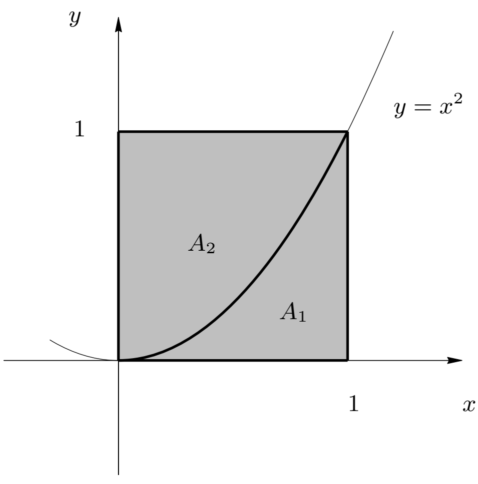


Figure 8.43. The set A relative to Exercise 7

Looking at Fig. 8.43 then, we have

$$\begin{aligned}\int_A x\sin|x^2-y|\,\mathrm{d}x\,\mathrm{d}y &= \int_{A_1} x\sin(x^2-y)\,\mathrm{d}x\,\mathrm{d}y - \int_{A_2} x\sin(x^2-y)\,\mathrm{d}x\,\mathrm{d}y\\ &= \int_0^1\int_0^{x^2} x\sin(x^2-y)\,\mathrm{d}y\,\mathrm{d}x - \int_0^1\int_{x^2}^1 x\sin(x^2-y)\,\mathrm{d}y\,\mathrm{d}x\\ &= \int_0^1 \big[x\cos(x^2-y)\big]_0^{x^2}\,\mathrm{d}x - \int_0^1 \big[x\cos(x^2-y)\big]_{x^2}^1\,\mathrm{d}x\\ &= \int_0^1 x(1-\cos x^2)\,\mathrm{d}x - \int_0^1 x\big(\cos(x^2-1)-1\big)\,\mathrm{d}x = 1-\sin 1\,.\end{aligned}$$

8. The result is $I = \pi - 2$.

9. We have

$$\int_A y\mathrm{e}^{-\alpha|x-y^2|}\,\mathrm{d}x\,\mathrm{d}y = \begin{cases} \dfrac{1}{\alpha^2}(\mathrm{e}^{-\alpha}-1+\alpha) & \text{if } \alpha\neq 0\,,\\ \dfrac{1}{2} & \text{if } \alpha = 0\,.\end{cases}$$

10. *Double integrals:*

a) The region A is bounded by four parabolas, see Fig. 8.44, left. Set $u = \dfrac{y}{x^2}$, $v = \dfrac{x}{y^2}$, which define the transformation $(u,v) = \boldsymbol{\Psi}(x,y)$; with that change, A becomes the rectangle $A' = [1,2]\times[1,3]$. The Jacobian of $\boldsymbol{\Psi}$ reads

$$\boldsymbol{J\Psi}(x,y) = \begin{pmatrix} -2y/x^3 & 1/x^2\\ 1/y^2 & -2x/y^3\end{pmatrix}$$

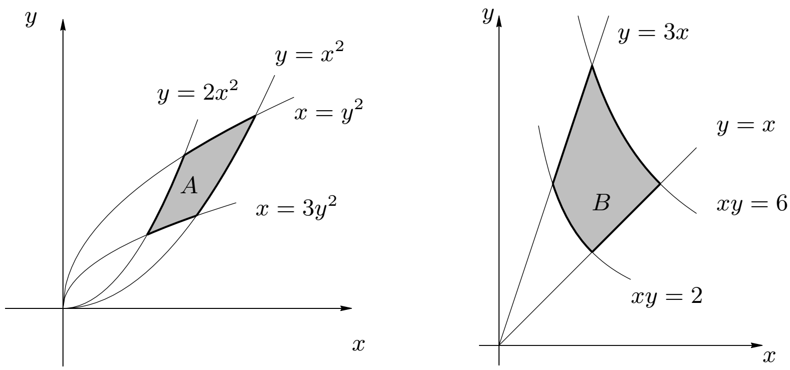


Figure 8.44. The sets A relative to Exercise 10. a) (left) and B to Exercise 10. b) (right)

with

$$\det \boldsymbol{J\Psi}(x,y) = \frac{4}{x^2y^2} - \frac{1}{x^2y^2} = \frac{3}{x^2y^2} = 3u^2v^2\,;$$

therefore, if $\boldsymbol{\Phi} = \boldsymbol{\Psi}^{-1}$, we have $\det \boldsymbol{J\Phi}(u,v) = \frac{1}{3u^2v^2}$. Thus

$$\int_A \frac{1}{x^2y^2}\,\mathrm{d}x\,\mathrm{d}y = \int_1^2 \int_1^3 u^2v^2 \frac{1}{3u^2v^2}\,\mathrm{d}v\,\mathrm{d}u = \frac{2}{3}\,.$$

b) B is bounded by 2 lines and 2 hyperbolas, see Fig. 8.44, right.
Define $(u,v) = \boldsymbol{\Psi}(x,y)$ by $u = xy$, $v = \dfrac{y}{x}$. Then B becomes the rectangle $B' = [2,6] \times [1,3]$, and

$$\boldsymbol{J\Psi}(x,y) = \begin{pmatrix} y & x \\ -y/x^2 & 1/x \end{pmatrix}, \qquad \det \boldsymbol{J\Psi}(x,y) = 2\frac{y}{x} = 2v\,.$$

Calling $\boldsymbol{\Phi} = \boldsymbol{\Psi}^{-1}$, we have $\det \boldsymbol{J\Phi}(u,v) = 1/2v$, so

$$\int_B \frac{x^5y^5}{x^3y^3+1}\,\mathrm{d}x\,\mathrm{d}y = \int_1^3\int_2^6 \frac{u^5}{u^3+1}\,\frac{1}{2v}\,\mathrm{d}u\,\mathrm{d}v = \frac{1}{2}\big[\log v\big]_1^3 \int_2^6 \Big(u^2 - \frac{u^2}{u^3+1}\Big)\,\mathrm{d}u$$

$$= \frac{1}{2}\log 3\left[\frac{1}{3}u^3 - \frac{1}{3}\log(u^3+1)\right]_2^6 = \frac{1}{6}\log 3\big(208 + \log\frac{9}{217}\big)\,.$$

c) The region C is shown in Fig. 8.45, left.
Passing to polar coordinates (r,θ), C transforms into $C' = [1,2]\times[0,\pi]$, so

$$\int_C (3x+4y^2)\,\mathrm{d}x\,\mathrm{d}y = \int_0^\pi \int_1^2 (3r\cos\theta + 4r^2\sin^2\theta)r\,\mathrm{d}r\,\mathrm{d}\theta$$

$$= \int_0^\pi \big[r^3\cos\theta + r^4\sin^2\theta\big]_1^2\,\mathrm{d}\theta = \int_0^\pi \big(7\cos\theta + 15\sin^2\theta\big)\,\mathrm{d}\theta$$

$$= \int_0^\pi \Big(7\cos\theta + \frac{15}{2}(1-\cos 2\theta)\Big)\,\mathrm{d}\theta = \frac{15}{2}\pi\,.$$

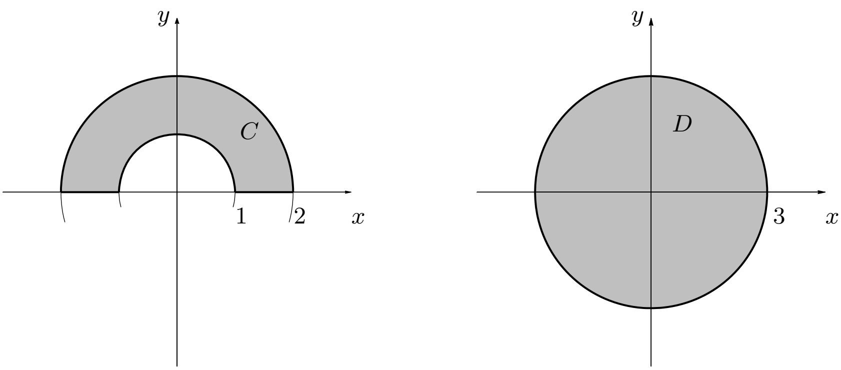


Figure 8.45. The sets C relative to Exercise 10. c) (left) and D to Exercise 10. d) (right)

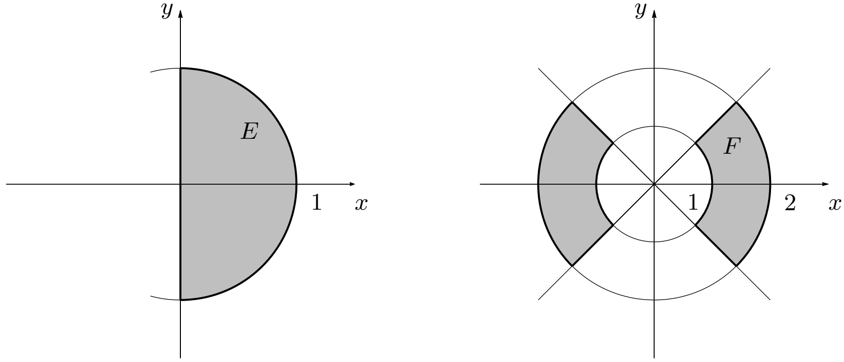


Figure 8.46. The sets E relative to Exercise 10. e) (left) and F to Exercise 10. f) (right)

d) For D see Fig. 8.45, right. The integral is $I = 0$.

e) Fig. 8.46, left, shows the region E, and the integral is $I = \frac{\pi}{2}(1 - e^{-4})$.

f) F is represented in Fig. 8.46, right, the integral is $I = 0$.

g) G is shown in Fig. 8.47, left. The integral equals $I = \frac{3\sqrt{2}}{2}$.

h) The region H lies in the first quadrant, and consists of points inside the circle centred at the origin of radius 2 but outside the circle with centre $(0,1)$ and unit radius (Fig. 8.47, right). In polar coordinates H becomes H' in the (r,θ)-plane defined by $0 \le \theta \le \frac{\pi}{2}$ and $2\sin\theta \le r \le 2$; that is because the circle $x^2 + y^2 - 2y \ge 0$ reads $r - 2\sin\theta \ge 0$ in polar coordinates. Therefore

$$\begin{aligned}\int_H x\,dx\,dy &= \int_0^{\pi/2}\int_{2\sin\theta}^{2} r^2\cos\theta\,dr\,d\theta = \frac{1}{3}\int_0^{\pi/2}\cos\theta\big[r^3\big]_{2\sin\theta}^{2}\,d\theta \\ &= \frac{8}{3}\int_0^{\pi/2}(\cos\theta - \cos\theta\sin^3\theta)\,d\theta = \frac{8}{3}\left[\sin\theta - \frac{1}{4}\sin^4\theta\right]_0^{\pi/2} = 2\,.\end{aligned}$$

11. If we pass to Cartesian coordinates, the condition $0 \le \theta \le \frac{\pi}{2}$ means we only have to consider points in the first quadrant, while $r = \frac{1}{\cos\theta}$ and $r = \frac{1}{\sin\theta}$

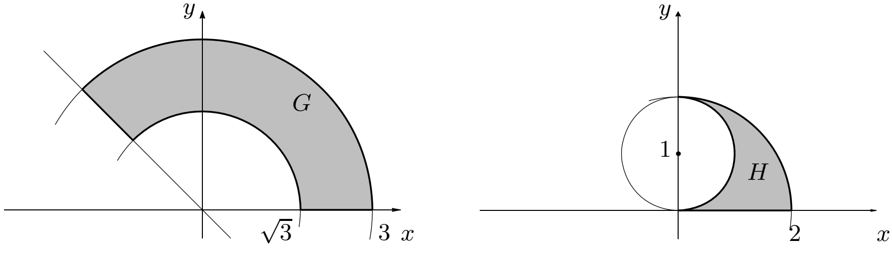


Figure 8.47. The sets G relative to Exercise 10. g) (left) and H to Exercise 10. h) (right)

correspond to the lines $x = 1$, $y = 1$. In the plane (x, y) then, A' becomes the square $A = [0,1] \times [0,1]$, so

$$\begin{aligned}
\int_{A'} \frac{1}{\sqrt{\cos^2\theta + \sin\theta/r + 1/r^2}}\, dr\, d\theta &= \int_{A'} \frac{1}{\sqrt{r^2\cos^2\theta + r\sin\theta + 1}}\, r\, dr\, d\theta \\
&= \int_A \frac{1}{\sqrt{x^2+y+1}}\, dx\, dy = \int_0^1 \int_0^1 \frac{1}{\sqrt{x^2+y+1}}\, dx\, dy \\
&= 2\int_0^1 \Big[\sqrt{x^2+y+1}\Big]_0^1\, dx = 2\int_0^1 \big(\sqrt{x^2+2} - \sqrt{x^2+1}\big)\, dx \\
&= 2\left[\frac{x\sqrt{x^2+2}}{2} + \log(x+\sqrt{x^2+2}) - \frac{x\sqrt{x^2+1}}{2} - \frac{1}{2}\log(x+\sqrt{x^2+1})\right]_0^1 \\
&= \sqrt{3} - \sqrt{2} + \log\frac{2+\sqrt{3}}{1+\sqrt{2}}\,.
\end{aligned}$$

12. $I = 4\log 2 - 2$.

13. Note $x \in \left[\frac{1}{\sqrt{2}}, 2\right]$ and the curves $y = \sqrt{1-x^2}$ and $y = \sqrt{4-x^2}$ are semi-circles at the origin with radius 1 and 2. The region A in the coordinates (x, y) is made of the following three sets:

$$\begin{aligned}
A_1 &= \{(x,y) \in \mathbb{R}^2 : \frac{1}{\sqrt{2}} \le x \le 1,\ \sqrt{1-x^2} \le y \le x\} \\
A_2 &= \{(x,y) \in \mathbb{R}^2 : 1 \le x \le \sqrt{2},\ 0 \le y \le x\} \\
A_3 &= \{(x,y) \in \mathbb{R}^2 : \sqrt{2} \le x \le 2,\ 0 \le y \le \sqrt{4-x^2}\}\,;
\end{aligned}$$

see Fig. 8.48.

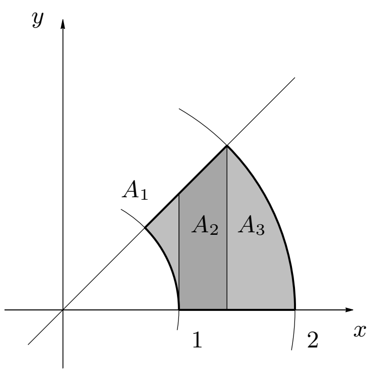


Figure 8.48. The set A relative to Exercise 13

In polar coordinates A becomes $A' = \{(r,\theta) : 0 \le \theta \le \frac{\pi}{4},\ 1 \le r \le 2\}$, so

$$I = \int_{A'} r^3 \sin\theta \cos\theta \,\mathrm{d}r\,\mathrm{d}\theta = \int_0^{\pi/4} \int_1^2 r^3 \sin\theta\cos\theta \,\mathrm{d}r\,\mathrm{d}\theta = \frac{15}{16}\,.$$

14. The thin plate is shown in Fig. 8.49, left, and we can write

$$A = \{(x,y) \in \mathbb{R}^2 : 0 \le x \le 1,\ 0 \le y \le 2-2x\}\,.$$

Hence

$$m(A) = \int_A \mu(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_0^1 \int_0^{2-2x} (1+3x+y)\,\mathrm{d}y\,\mathrm{d}x = \frac{8}{3}\,,$$

$$x_B(A) = \frac{1}{m(A)} \int_A x\mu(x,y)\,\mathrm{d}x\,\mathrm{d}y = \frac{3}{8}\int_0^1 \int_0^{2-2x} x(1+3x+y)\,\mathrm{d}y\,\mathrm{d}x = \frac{3}{8}\,,$$

$$y_B(A) = \frac{1}{m(A)} \int_A y\mu(x,y)\,\mathrm{d}x\,\mathrm{d}y = \frac{3}{8}\int_0^1 \int_0^{2-2x} y(1+3x+y)\,\mathrm{d}y\,\mathrm{d}x = \frac{11}{16}\,,$$

so the centre of mass has coordinates $(\frac{3}{8}, \frac{11}{16})$.

15. The thin plate A is the upper semi-circle of $x^2+y^2=a^2$ (Fig. 8.49, right).

The distance of (x,y) from the centre (the origin) is $\sqrt{x^2+y^2}$, whence the density becomes $\mu(x,y) = K\sqrt{x^2+y^2}$, with given $K>0$. In polar coordinates,

$$m(A) = \int_A \mu(x,y)\,\mathrm{d}x\,\mathrm{d}y = \int_0^\pi \int_0^a Kr^2\,\mathrm{d}r\,\mathrm{d}\theta = \frac{K\pi a^3}{3}\,.$$

Since the thin plate and its density are symmetric with respect to y, the centre of mass has to lie on the axis, so $x_B(A)=0$. As for the other coordinate:

$$y_B(A) = \frac{1}{m(A)} \int_A y\mu(x,y)\,\mathrm{d}x\,\mathrm{d}y = \frac{3}{K\pi a^3}\int_0^\pi \int_0^a Kr^3 \sin\theta\,\mathrm{d}r\,\mathrm{d}\theta = \frac{3a}{2\pi}\,.$$

Therefore the centroid's coordinates are $(0, \frac{3a}{2\pi})$.

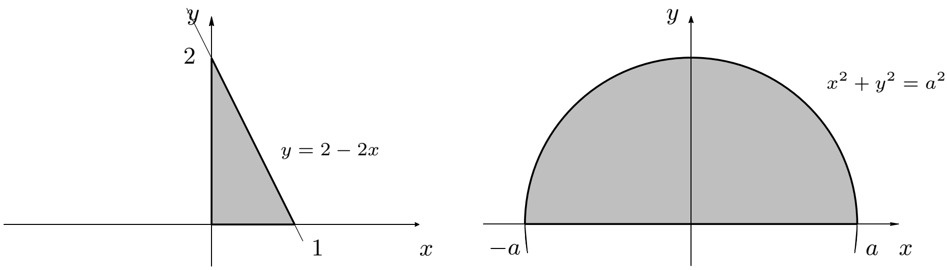


Figure 8.49. The sets A relative to Exercises 14 (left) and 15 (right)

16. In polar coordinates

$$I_0 = \int_D (x^2+y^2)\mu \,\mathrm{d}x\,\mathrm{d}y = \mu \int_0^{2\pi}\int_0^a r^3\,\mathrm{d}r\,\mathrm{d}\theta = \frac{\pi}{2}\mu a^4 .$$

By symmetry, $I_x = I_y$, so the equation $I_0 = I_x + I_y$ gives immediately

$$I_x = I_y = \frac{\pi}{4}\mu a^4 .$$

17. Fig. 8.50 shows the disc D, whose boundary has equation $(x-a)^2+y^2=a^2$, or polar $r = 2a\cos\theta$. Thus

$$\begin{aligned} I_0 &= \int_D (x^2+y^2)\,\mathrm{d}x\,\mathrm{d}y = \int_{-\pi/2}^{\pi/2}\int_0^{2a\cos\theta} r^3\,\mathrm{d}r\,\mathrm{d}\theta \\ &= 4\int_{-\pi/2}^{\pi/2} a^4\cos^4\theta\,\mathrm{d}\theta = 8a^4\int_0^{\pi/2}\cos^4\theta\,\mathrm{d}\theta = \frac{3}{2}\pi a^4 , \end{aligned}$$

$$\begin{aligned} I_C &= \int_D [(x-a)^2+y^2]\,\mathrm{d}x\,\mathrm{d}y = \int_{D'} (t^2+y^2)\,\mathrm{d}t\,\mathrm{d}y \\ &= \int_0^{2\pi}\int_0^a r^3\,\mathrm{d}r\,\mathrm{d}\theta = \frac{1}{2}\pi a^4 . \end{aligned}$$

But as $\mathcal{A} = \pi a^2$, we immediately find $I_0 = I_C + a^2\mathcal{A}$.

18. $I_0 = \frac{64}{5}\pi - \frac{16}{75}$.

19. Since $\mu(x,y) = y$ we have

$$m(A) = \int_C y\,\mathrm{d}x\,\mathrm{d}y = \int_0^{\pi/2}\int_0^1 r^2\sin\theta\,\mathrm{d}r\,\mathrm{d}\theta = \frac{1}{3} ,$$

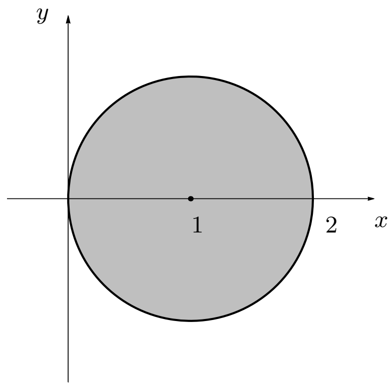


Figure 8.50. The disc D relative to Exercise 17

$$x_B(A) = 3\int_C xy\,\mathrm{d}x\,\mathrm{d}y = 3\int_0^{\pi/2}\int_0^1 r^3\sin\theta\cos\theta\,\mathrm{d}r\,\mathrm{d}\theta = \frac{3}{8}\,,$$

$$y_B(A) = 3\int_C y^2\,\mathrm{d}x\,\mathrm{d}y = 3\int_0^{\pi/2}\int_0^1 r^3\sin^2\theta\,\mathrm{d}r\,\mathrm{d}\theta = \frac{3}{16}\pi\,.$$

20. $I_y = 33$.

21. *Triple integrals:*

a) $I = -8$. b) $I = 4$. c) $I = \frac{3}{20}(1-\cos 1)$.

d) We have

$$\int_D y\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \int_0^1 y\left(\int_0^{1-y}\mathrm{d}x\int_0^{1-y}\mathrm{d}z\right)\mathrm{d}y$$
$$= \int_0^1 y(1-y)^2\,\mathrm{d}y = \frac{1}{12}\,.$$

See Fig. 8.51, left.

e) Since

$$\int_E y\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \int_0^4\left(\int_{E_y}\mathrm{d}x\,\mathrm{d}z\right)y\,\mathrm{d}y$$

with $E_y = \{(x,z)\in\mathbb{R}^2 : x^2+z^2 = \frac{y}{4}\}$, the integral $\int_{E_y}\mathrm{d}x\,\mathrm{d}z$ is the area of E_y, hence $\pi\frac{y}{4}$. Therefore

$$\int_E y\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \frac{\pi}{4}\int_0^4 y^2\,\mathrm{d}y = \frac{16}{3}\pi\,,$$

see Fig. 8.51, right.

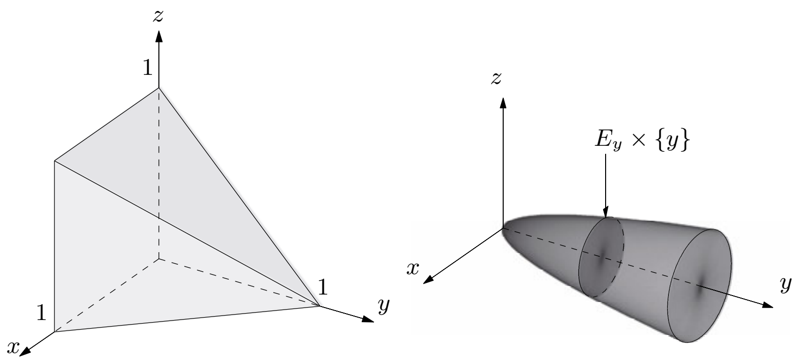


Figure 8.51. The regions relative to Exercise 21. d) (left) and Exercise 21. e) (right)

f) We have

$$\int_F z \,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \int_0^3 \left(\int_0^{\frac{1}{3}\sqrt{9-z^2}} \mathrm{d}x\right)\left(\int_0^{\sqrt{9-z^2}} \mathrm{d}y\right) z\,\mathrm{d}z = \frac{27}{4}\,.$$

See Fig. 8.52, left.

22. *Triple integrals:*

a) We have

$$I = \int_{-1}^1 \int_0^{1-y^2} \int_0^z f(x,y,z)\,\mathrm{d}x\,\mathrm{d}z\,\mathrm{d}y = \int_0^1 \int_{-\sqrt{1-z}}^{\sqrt{1-z}} \int_0^z f(x,y,z)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z$$

$$= \int_0^1 \int_x^1 \int_{-\sqrt{1-z}}^{\sqrt{1-z}} f(x,y,z)\,\mathrm{d}y\,\mathrm{d}z\,\mathrm{d}x = \int_0^1 \int_0^z \int_{-\sqrt{1-z}}^{\sqrt{1-z}} f(x,y,z)\,\mathrm{d}y\,\mathrm{d}x\,\mathrm{d}z$$

$$= \int_{-1}^1 \int_0^{1-y^2} \int_x^{1-y^2} f(x,y,z)\,\mathrm{d}z\,\mathrm{d}x\,\mathrm{d}y = \int_0^1 \int_{-\sqrt{1-x}}^{\sqrt{1-x}} \int_x^{1-y^2} f(x,y,z)\,\mathrm{d}z\,\mathrm{d}y\,\mathrm{d}x\,.$$

b) Simply computing,

$$I = \int_{-3}^3 \int_0^6 \int_{-\sqrt{9-y^2}}^{\sqrt{9-y^2}} f(x,y,z)\,\mathrm{d}x\,\mathrm{d}z\,\mathrm{d}y = \int_0^6 \int_{-3}^3 \int_{-\sqrt{9-y^2}}^{\sqrt{9-y^2}} f(x,y,z)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z$$

$$= \int_{-3}^3 \int_0^6 \int_{-\sqrt{9-x^2}}^{\sqrt{9-x^2}} f(x,y,z)\,\mathrm{d}y\,\mathrm{d}z\,\mathrm{d}x = \int_0^6 \int_{-3}^3 \int_{-\sqrt{9-x^2}}^{\sqrt{9-x^2}} f(x,y,z)\,\mathrm{d}y\,\mathrm{d}x\,\mathrm{d}z$$

$$= \int_{-3}^3 \int_{-\sqrt{9-y^2}}^{\sqrt{9-y^2}} \int_0^6 f(x,y,z)\,\mathrm{d}z\,\mathrm{d}x\,\mathrm{d}y = \int_{-3}^3 \int_{-\sqrt{9-x^2}}^{\sqrt{9-x^2}} \int_0^6 f(x,y,z)\,\mathrm{d}z\,\mathrm{d}y\,\mathrm{d}x\,.$$

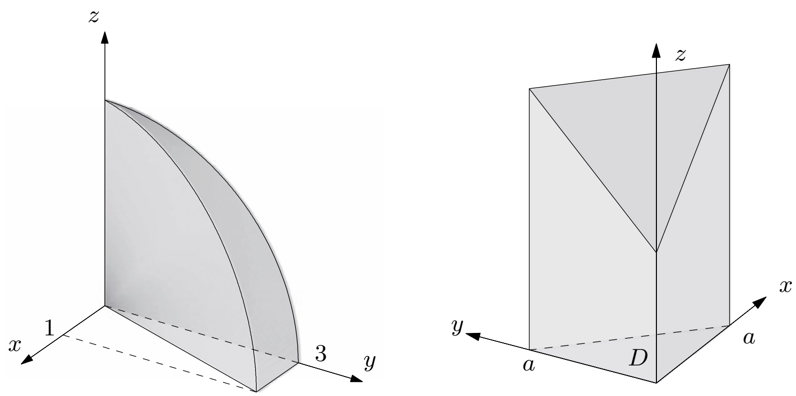


Figure 8.52. The regions relative to Exercise 21. f) (left) and to Exercise 23 (right)

23. Integrating in z first, with $0 \le z \le x+y+1$, gives

$$\mathrm{vol}(A) = \int_D \int_0^{x+y+1} \mathrm{d}z\,\mathrm{d}x\,\mathrm{d}y$$

where $D = \{(x,y) \in \mathbb{R}^2 : 0 \le x \le a,\, 0 \le y \le a-x\}$, see Fig. 8.52, right. The computations yield

$$\mathrm{vol}(A) = \int_0^a \int_0^{a-x} (x+y+1)\,\mathrm{d}y\,\mathrm{d}x = \frac{1}{2}a^2 + \frac{1}{3}a^3 .$$

Imposing $\mathrm{vol}(A) = \frac{5}{6}$ we obtain

$$\frac{1}{2}a^2 + \frac{1}{3}a^3 = \frac{5}{6} \qquad \text{i.e.,} \qquad (a-1)(2a^2+5a+5) = 0 .$$

In conclusion, $a = 1$.

24. By symmetry it suffices to compute the volume of the region restricted to the first octant, then multiply it by 8. Reducing the integral,

$$\mathrm{vol}(A) = 8\int_0^1 \int_0^{\sqrt{1-x^2}} \int_0^{\sqrt{1-x^2}} \mathrm{d}z\,\mathrm{d}x\,\mathrm{d}y = \frac{16}{3} .$$

See Fig. 8.53.

25. The surface $9z = 1 + y^2 + 9x^2$ is a paraboloid, whereas $z = \sqrt{9-(y^2+9x^2)}$ is an ellipsoid centred at the origin with semi-axes $a = 1$, $b = 3$, $c = 3$. Then

$$\int_A \frac{1}{3-z}\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \int_0^1 \frac{1}{3-z}\left(\int_{A_z} \mathrm{d}x\,\mathrm{d}y\right)\mathrm{d}z .$$

For the integral inbetween brackets we distinguish $0 \le z \le \frac{1}{9}$ and $\frac{1}{9} \le z \le 1$. In the first case A_z is an ellipse in (x,y) with semi-axes $a = \frac{1}{3}\sqrt{9-z^2}$ and $b = \sqrt{9-z^2}$

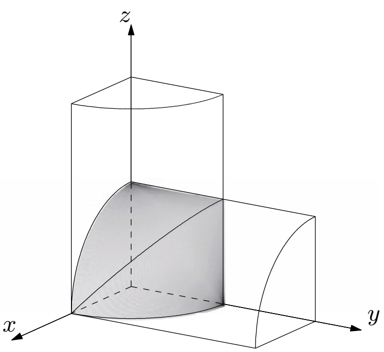


Figure 8.53. The region relative to Exercise 24

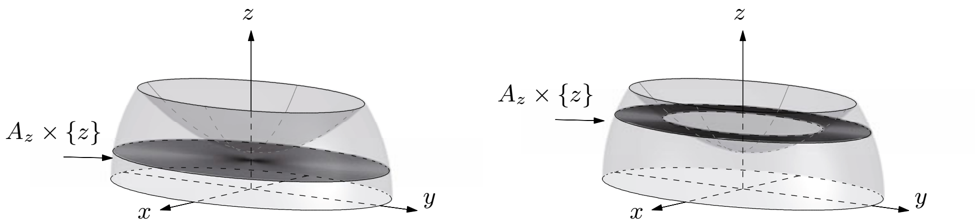


Figure 8.54. The regions relative to Exercise 25

(Fig. 8.54, left); the integral is the area of A_z, whence $\pi ab = \frac{\pi}{3}(9-z^2)$. In case $\frac{1}{9} \le z \le 1$, A_z is the region bounded by the ellipses $9x^2+y^2 = 9z-1$ and $9x^2+y^2 = 9-z^2$ (Fig. 8.54, right) with

$$a_1 = \frac{\sqrt{9z-1}}{3}, \qquad b_1 = \sqrt{9z-1}\,, \qquad a_2 = \frac{\sqrt{9-z^2}}{3}, \qquad b_2 = \sqrt{9-z^2}\,.$$

Therefore

$$\int_{A_z} \mathrm{d}x\,\mathrm{d}y = \pi a_2 b_2 - \pi a_1 b_1 = \frac{\pi}{3}(9-z^2) - \frac{\pi}{3}(9z-1)\,.$$

Returning to the starting integral,

$$\begin{aligned}\int_A \frac{1}{3-z}\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z &= \frac{\pi}{3}\int_0^{1/9} \frac{9-z^2}{3-z}\,\mathrm{d}z + \frac{\pi}{3}\int_{1/9}^1 \left(\frac{9-z^2}{3-z} - \frac{9z-1}{3-z}\right)\mathrm{d}z \\ &= \frac{\pi}{3}\int_0^1 (3+z)\,\mathrm{d}z + \frac{\pi}{3}\int_{1/9}^1 \left(9 + \frac{26}{z-3}\right)\mathrm{d}z \\ &= \frac{23}{6}\pi + \frac{26}{3}\pi \log\frac{9}{13}\,.\end{aligned}$$

26. Both volumes can be found by integrating in x the cross-sections:

$$\mathrm{vol}(C_1) = \int_0^K \left(\int_{A_x} \mathrm{d}y\,\mathrm{d}z\right)\mathrm{d}x = \int_K^1 \left(\int_{A_x} \mathrm{d}y\,\mathrm{d}z\right)\mathrm{d}x = \mathrm{vol}(C_2)\,.$$

The set A_x, contained in (y,z), is

$$A_x = \left\{(y,z) \in \mathbb{R}^2 : 0 \le y \le 2(1-x),\ 0 \le z \le 3-3x-\frac{3}{2}y\right\};$$

Fig. 8.55 shows the product $A_x \times \{x\}$. We must thus have

$$\int_0^K \int_0^{2(1-x)} \int_0^{3-3x-\frac{3}{2}y} \mathrm{d}z\,\mathrm{d}y\,\mathrm{d}x = \int_K^1 \int_0^{2(1-x)} \int_0^{3-3x-\frac{3}{2}y} \mathrm{d}z\,\mathrm{d}y\,\mathrm{d}x\,,$$

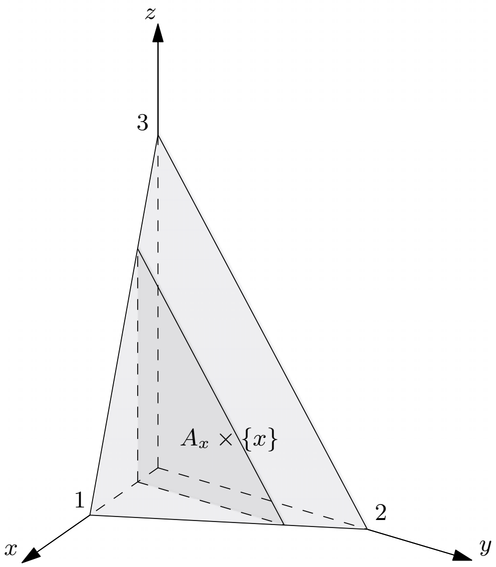


Figure 8.55. The region relative to Exercise 26

hence

$$1 - (1-K)^3 = (1-K)^3\,,$$

from which $K = 1 - 1/\sqrt[3]{2}$.

27. *Triple integrals:*

a) In cylindrical coordinates A becomes

$$A' = \{(r,\theta,t) : 0 \le r \le 5,\ 0 \le \theta \le 2\pi,\ -1 \le t \le 2\}\,.$$

Therefore

$$\int_A \sqrt{x^2+y^2}\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \int_{-1}^{2}\int_0^{2\pi}\int_0^5 r^2\,\mathrm{d}r\,\mathrm{d}\theta\,\mathrm{d}t = 250\pi\,.$$

b) We shall use cylindrical coordinates with axis y, i.e., $x = r\cos\theta$, $y = t$, $z = r\sin\theta$. Then (see Fig. 8.56)

$$B = \{(x,y,z) \in \mathbb{R}^3 : 1 \le x^2+z^2 \le 4,\ 0 \le y \le z+2\}$$

becomes

$$B' = \{(r,\theta,t) : 1 \le r \le 2,\ 0 \le \theta \le 2\pi,\ 0 \le t \le r\sin\theta + 2\}\,;$$

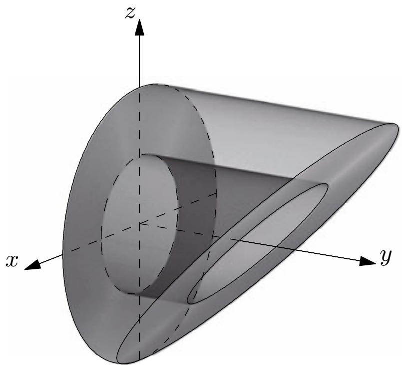


Figure 8.56. The region relative to Exercise 27. b)

therefore

$$\begin{aligned}\int_B x\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z &= \int_0^{2\pi}\int_1^2\int_0^{r\sin\theta+2} r^2\cos\theta\,\mathrm{d}t\,\mathrm{d}r\,\mathrm{d}\theta\\ &= \int_0^{2\pi}\int_1^2 r^2\cos\theta(r\sin\theta+2)\,\mathrm{d}r\,\mathrm{d}\theta\\ &= \Big[\frac{1}{4}r^4\Big]_1^2\Big[\frac{1}{2}\sin^2\theta\Big]_0^{2\pi}+\Big[\frac{2}{3}r^3\Big]_1^2\Big[\sin\theta\Big]_0^{2\pi}=0\,.\end{aligned}$$

c) $\frac{9}{4}\pi$.

d) The region D is, in spherical coordinates,

$$D' = \{(r,\varphi,\theta) : 1\le r\le 2,\ 0\le\varphi,\theta\ \le\frac{\pi}{2}\}\,,$$

so

$$\begin{aligned}\int_D x\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z &= \int_1^2\int_0^{\pi/2}\int_0^{\pi/2} r^3\sin^2\varphi\cos\theta\,\mathrm{d}\theta\,\mathrm{d}\varphi\,\mathrm{d}r\\ &= \Big(\int_1^2 r^3\,\mathrm{d}r\Big)\Big(\int_0^{\pi/2}\cos\theta\,\mathrm{d}\theta\Big)\Big(\int_0^{\pi/2}\sin^2\varphi\,\mathrm{d}\varphi\Big)\\ &= \frac{15}{4}\cdot 1\cdot\frac{1}{2}\int_0^{\pi/2}(1-\cos 2\varphi)\,\mathrm{d}\varphi = \frac{15}{16}\pi\,.\end{aligned}$$

e) $4(2-\sqrt{3})\pi$.

28. *Triple integrals:*

a) In spherical coordinates Ω is defined by

$$-\sqrt{1-z^2} \le x \le \sqrt{1-z^2}\,, \quad x^2+z^2 \le y \le 2-x^2-z^2\,, \quad -1 \le z \le 1\,.$$

The first and last constraints determine the inside of the cylinder $x^2+z^2=1$, so cylindrical coordinates (along y) are convenient:

$$x = r\cos\theta\,, \qquad y = t\,, \qquad z = r\sin\theta\,.$$

Thus Ω becomes

$$\Omega' = \{(r,\theta,t) : 0 \le r \le 1,\, 0 \le \theta \le 2\pi,\, r^2 \le t \le 2-r^2\}$$

and the integral reads

$$\int_\Omega (x^2+z^2)^{3/2}\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \int_0^1\int_0^{2\pi}\int_{r^2}^{2-r^2} r^4\,\mathrm{d}t\,\mathrm{d}\theta\,\mathrm{d}r = \frac{8}{35}\pi\,.$$

b) $\frac{486}{5}(\sqrt{2}-1)\pi$.

29. Let us use generalised spherical coordinates

$$x = 2r\sin\varphi\cos\theta\,, \qquad y = 3r\sin\varphi\sin\theta\,, \qquad z = 4r\cos\varphi$$

where $\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = 24r^2\sin\varphi\,\mathrm{d}r\,\mathrm{d}\varphi\,\mathrm{d}\theta$; now Ω' is described by $0 \le r \le 1\,,\ 0 \le \theta \le 2\pi\,,\ 0 \le \varphi \le \pi\,.$ Thus,

$$\int_\Omega \Big(4x^2 + \frac{16}{9}y^2 + z^2\Big)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = 16\cdot 24\int_0^1\int_0^{2\pi}\int_0^{\pi} r^4\sin\varphi\,\mathrm{d}\varphi\,\mathrm{d}\theta\,\mathrm{d}r = \frac{1536}{5}\pi\,.$$

30. Let us first find the mass using cylindrical coordinates. As Ω is

$$\Omega' = \{(r,\theta,t) : 0 \le r \le 1,\, 0 \le \theta \le \frac{\pi}{2},\, 0 \le t \le 1\}\,,$$

we have

$$m = \int_\Omega (x+y)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \int_0^1\int_0^1\int_0^{\pi/2} r^2(\cos\theta+\sin\theta)\,\mathrm{d}\theta\,\mathrm{d}r\,\mathrm{d}t = \frac{2}{3}\,.$$

By symmetry, $x_G = y_G$, and, similarly to the previous calculation,

$$\begin{aligned}
x_G &= \frac{3}{2}\int_\Omega x(x+y)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \frac{3}{2}\int_0^1\int_0^1\int_0^{\pi/2} r^3\cos\theta(\cos\theta+\sin\theta)\,\mathrm{d}\theta\,\mathrm{d}r\,\mathrm{d}t \\
&= \frac{3}{8}\int_0^{\pi/2}(\cos^2\theta+\cos\theta\sin\theta)\,\mathrm{d}\theta\,\mathrm{d}r\,\mathrm{d}t = \frac{3}{16}\Big[\theta+\frac{1}{2}\sin 2\theta+\sin^2\theta\Big]_0^{\pi/2} \\
&= \frac{3}{16}\big(1+\frac{\pi}{2}\big) = y_G\,,
\end{aligned}$$

$$z_G = \frac{3}{2}\int_\Omega z(x+y)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \frac{3}{2}\int_0^1\int_0^1\int_0^{\pi/2} t\,r^2(\cos\theta+\sin\theta)\,\mathrm{d}\theta\,\mathrm{d}r\,\mathrm{d}t = \frac{1}{2}\,.$$

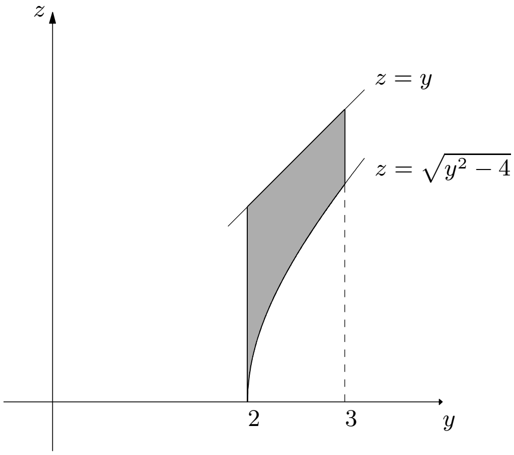


Figure 8.57. The meridian section relative to Exercise 31

31. The meridian section T is shown in Fig. 8.57. Theorem 8.32 tells us

$$\mathrm{vol}(\Omega) = 2\pi \int_T y \, \mathrm{d}y \, \mathrm{d}z \, .$$

Then

$$\begin{aligned} \mathrm{vol}(\Omega) &= 2\pi \int_2^3 \left(\int_{\sqrt{y^2-4}}^{y} y \, \mathrm{d}z \right) \mathrm{d}y = 2\pi \int_2^3 \left(y^2 - y\sqrt{y^2-4} \right) \mathrm{d}y \\ &= 2\pi \Big[\frac{1}{3} y^3 - \frac{1}{3}(y^2-4)^{3/2} \Big]_2^3 = \frac{2}{3}(19 - 5\sqrt{5})\pi \, . \end{aligned}$$

32. 8π.

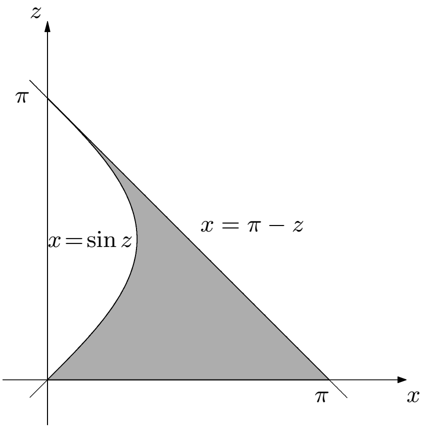
33. The section T is drawn in Fig. 8.58, left. Using Theorem 8.32 the volume is

$$\begin{aligned} \mathrm{vol}(\Omega) &= 2\pi \int_T x \, \mathrm{d}x \, \mathrm{d}z = 2\pi \int_0^\pi \Big(\int_{\sin z}^{\pi - z} x \, \mathrm{d}x \Big) \mathrm{d}z = \pi \int_0^\pi \big((\pi - z)^2 - \sin^2 z \big) \, \mathrm{d}z \\ &= \pi \Big[-\frac{1}{3}(\pi - z)^3 - \frac{1}{2}\big(z - \frac{1}{2}\sin 2z\big) \Big]_0^\pi = \frac{1}{3}\pi^4 - \frac{1}{2}\pi^2 \, . \end{aligned}$$

But as Ω is a solid of revolution around z, the centre of mass is on that axis, so $x_G = y_G = 0$. To find z_G, we integrate cross-sections:

$$z_G = \frac{1}{\mathrm{vol}(\Omega)} \int_\Omega z \, \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}z = \frac{1}{\mathrm{vol}(\Omega)} \int_0^\pi z \left(\int_{A_z} \mathrm{d}x \, \mathrm{d}y \right) \mathrm{d}z$$

where A_z is the projection on xy of the annulus Ω_z of Fig. 8.58, right.



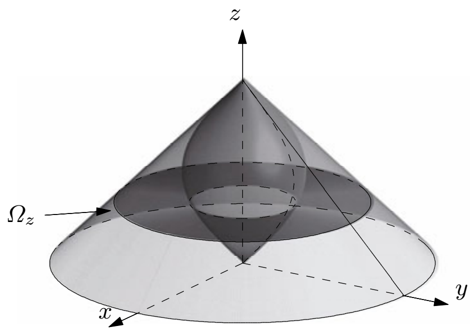


Figure 8.58. Meridian section (left) and region Ω_z (right) relative to Exercise 33

The integral $\displaystyle\int_{A_z} \mathrm{d}x\,\mathrm{d}y$ is the area of A_z, which is known to be $\pi\big((\pi - z)^2 - \sin^2 z\big)$. In conclusion,

$$z_G = \frac{\pi}{\mathrm{vol}(\Omega)} \int_0^{\pi} z\big((\pi - z)^2 - \sin^2 z\big)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \frac{1}{\mathrm{vol}(\Omega)}\Big(\frac{\pi^5}{12} - \frac{\pi^3}{4}\Big) = \frac{\pi^3 - 3\pi}{4\pi^2 - 6}\,.$$

9

Integral calculus on curves and surfaces

With this chapter we conclude the study of multivariable integral calculus. In the first part we define integrals along curves in $\mathbb{R}^m$ and over surfaces in space, by considering first real-valued maps, then vector-valued functions. Integrating a vector field's tangential component along a curve, or its normal component on a surface, defines line and flux integrals respectively; these are interpreted in Physics as the work done by a force along a path, or the flow across a membrane immersed in a fluid. Curvilinear integrals rely, de facto, on integrals over real intervals, in the same way as surface integrals are computed by integrating over domains in the plane. A certain attention is devoted to how integrals depend upon the parametrisations and orientations of the manifolds involved.

Path and flux integrals crop up in a series of results, among which three are pivotal: the Divergence Theorem (also known as Gauss' Theorem), Green's Theorem and Stokes' Theorem. They transfer to a multivariable framework the idea at the heart of one-dimensional integration by parts, and namely: changing the integral over a given domain into an integral over its boundary by modifying the integrand function. The aforementioned theorems have a large number of applications, in particular differential equations governing physical laws; for instance, the Divergence Theorem allows to compute the variation of the flow of a unit of matter by means of an integral over the volume, thus giving rise to the so-called conservation laws.

The last part deals with conservative fields, which are gradients of vector fields, and with the problem of computing the potential of a given field. That is the multivariable version of indefinite integration, which seeks primitive maps on the real line. The applicative importance of conservative fields and their potentials is well known: in many cases a gravitational field or an electric force field is conservative, a fact allowing to compute the force's work simply by difference of two values of the potential.

C. Canuto, A. Tabacco: *Mathematical Analysis II,* 2nd Ed.,
UNITEXT – La Matematica per il 3+2 85, DOI 10.1007/978-3-319-12757-6_9,

9.1 Integrating along curves

Curvilinear integrals in several variables are the natural generalisation of the definite integral of a real map over a real interval. Curves were introduced in Sect. 4.6, while their differential aspects were discussed in Sect. 6.5.

Let $\boldsymbol{\gamma} : I = [a, b] \to \mathbb{R}^m$ ($m \geq 2$) be a regular arc, and $\Gamma = \boldsymbol{\gamma}(I)$ its trace. Suppose $f : \operatorname{dom} f \subseteq \mathbb{R}^m \to \mathbb{R}$ is a map defined on Γ, at least, meaning $\Gamma \subseteq \operatorname{dom} f$. Suppose further that the composite $f \circ \boldsymbol{\gamma} : [a, b] \to \mathbb{R}$, $(f \circ \boldsymbol{\gamma})(t) = f(\boldsymbol{\gamma}(t))$, is (piecewise) continuous on $[a, b]$.

Definition 9.1 *The* **integral** *of f* **along** $\boldsymbol{\gamma}$ *is the number*

$$\int_{\boldsymbol{\gamma}} f = \int_a^b f\big(\boldsymbol{\gamma}(t)\big)\, \|\boldsymbol{\gamma}'(t)\|\, \mathrm{d}t\,. \tag{9.1}$$

Sometimes the name 'curvilinear integral' can be met.

The right-hand side of (9.1) is well defined, because the integrand function $f\big(\boldsymbol{\gamma}(t)\big)\, \|\boldsymbol{\gamma}'(t)\|$ is (piecewise) continuous on $[a, b]$. As $\boldsymbol{\gamma}$ is regular, in fact, its components' first derivatives are continuous and so is the norm $\|\boldsymbol{\gamma}'(t)\|$, by composition; moreover, $f\big(\boldsymbol{\gamma}(t)\big)$ is (piecewise) continuous by hypothesis.

The geometrical meaning is the following. Let $\boldsymbol{\gamma}$ be a simple plane arc and f non-negative along Γ; call

$$G(f) = \{(x, y, z) \in \mathbb{R}^3 : (x, y) \in \operatorname{dom} f,\ z = f(x, y)\}$$

the graph of f. If

$$\Sigma = \{(x, y, z) \in \mathbb{R}^3 : (x, y) \in \Gamma,\ 0 \leq z \leq f(x, y)\}$$

denotes the upright surface fencing Γ up to the graph of f (see Fig. 9.1), it can be proved that the area of Σ is given by the integral of f along $\boldsymbol{\gamma}$. Say, for instance, f is constant equal to h on Γ: then the area of Σ is the product of h times the length of Γ; by Sect. 6.5.2 the length is $\ell(\Gamma) = \int_a^b \|\boldsymbol{\gamma}'(t)\|\, \mathrm{d}t$, whence

$$\operatorname{area}(\Sigma) = h\, \ell(\Gamma) = \int_a^b f\big(\boldsymbol{\gamma}(t)\big) \|\boldsymbol{\gamma}'(t)\|\, \mathrm{d}t = \int_{\boldsymbol{\gamma}} f\,.$$

If f is not constant, instead, we can divide the interval I into K sub-intervals $I_k = [t_{k-1}, t_k]$ of width $\Delta t_k = t_k - t_{k-1}$ sufficiently small. Call $\Gamma_k = \boldsymbol{\gamma}(I_k)$ the restriction to I_k, whose length is $\ell(\Gamma_k) = \int_{t_{k-1}}^{t_k} \|\boldsymbol{\gamma}'(t)\|\, \mathrm{d}t$; then let $\Sigma_k = \{(x, y, z) \in \mathbb{R}^3 : (x, y) \in \Gamma_k,\ 0 \leq z \leq f(x, y)\}$ be the part of Σ over Γ_k (see again Fig. 9.1). Given any $t_k^* \in I_k$, with $P_k^* = \boldsymbol{\gamma}(t_k^*) \in \Gamma_k$, we will have

$$\operatorname{area}(\Sigma_k) \simeq f(P_k^*)\ell(\Gamma_k) = \int_{t_{k-1}}^{t_k} f\big(\boldsymbol{\gamma}(t_k^*)\big) \|\boldsymbol{\gamma}'(t)\|\, \mathrm{d}t\,.$$

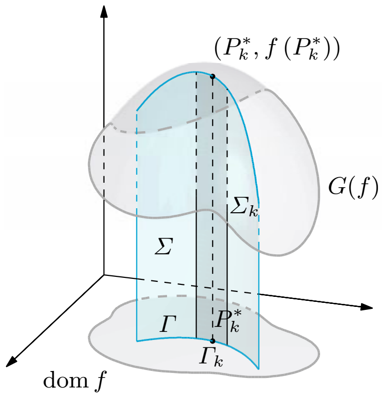


Figure 9.1. The geometrical meaning of integrating along a curve

Summing over k and letting the intervals' width shrink to zero yields precisely formula (9.1).

Recalling Definition 9.1, we can observe that the arc length $s = s(t)$ of (6.20) satisfies

$$\frac{\mathrm{d}s}{\mathrm{d}t} = \|\boldsymbol{\gamma}'(t)\|$$

which, in Leibniz's notation, we may re-write as

$$\boxed{\mathrm{d}s = \|\boldsymbol{\gamma}'(t)\|\,\mathrm{d}t\,.}$$

The differential $\mathrm{d}s$ is the 'infinitesimal' line element along the curve, corresponding to an 'infinitesimal' increment $\mathrm{d}t$ of the parameter t. Such considerations justify indicating the integral of (9.1) by

$$\int_{\boldsymbol{\gamma}} f\,\mathrm{d}s\,, \quad \text{or} \quad \int_{\boldsymbol{\gamma}} f\,\mathrm{d}\gamma \tag{9.2}$$

(in the latter case, the map $s = s(t)$ is written as $\gamma = \gamma(t)$, not to be confused with the vector notation of an arc $\boldsymbol{\gamma} = \boldsymbol{\gamma}(t)$).

Examples 9.2

i) Let $\boldsymbol{\gamma} : [0,1] \to \mathbb{R}^2$ be the regular arc $\boldsymbol{\gamma}(t) = (t, t^2)$ parametrising the parabola $y = x^2$ between $O = (0,0)$ and $A = (1,1)$. Since $\boldsymbol{\gamma}'(t) = (1, 2t)$, we have $\|\boldsymbol{\gamma}'(t)\| = \sqrt{1+4t^2}$. Suppose $f : \mathbb{R} \times [0, +\infty) \to \mathbb{R}$ is the map $f(x,y) = 3x + \sqrt{y}$. The composite $f \circ \boldsymbol{\gamma}$ is $f\big(\boldsymbol{\gamma}(t)\big) = 3t + \sqrt{t^2} = 4t$, so

$$\int_{\boldsymbol{\gamma}} f = \int_0^1 4t\sqrt{1+4t^2}\,\mathrm{d}t\,;$$

this is computed by substituting $r = 1 + 4t^2$,

$$\int_{\gamma} f = \frac{1}{2}\int_1^5 \sqrt{r}\,\mathrm{d}r = \frac{1}{2}\Big[\frac{2}{3}r^{3/2}\Big]_1^5 = \frac{1}{3}(5\sqrt{5}-1)\,.$$

ii) The regular arc $\boldsymbol{\gamma} : [0,1] \to \mathbb{R}^3$, $\boldsymbol{\gamma}(t) = (\frac{1}{3}t^3, \frac{\sqrt{2}}{2}t^2, t)$, gives

$$\boldsymbol{\gamma}'(t) = (t^2, \sqrt{2}t, 1) \quad \text{whence} \quad \|\boldsymbol{\gamma}'(t)\| = \sqrt{t^4 + 2t^2 + 1} = (1+t^2)\,.$$

Take now $f(x,y,z) = 3\sqrt{2}xyz^2$; then

$$\int_{\gamma} f = \int_0^1 t^7(1+t^2)\,\mathrm{d}t = \Big[\frac{1}{8}t^8 + \frac{1}{10}t^{10}\Big]_0^1 = \frac{9}{40}\,.$$

iii) Parametrise the circle centred in $(2,1)$ with radius 2 by $\boldsymbol{\gamma} : [0,2\pi] \to \mathbb{R}^2$, $\boldsymbol{\gamma}(t) = (2+2\cos t, 1+2\sin t)$. Then

$$\|\boldsymbol{\gamma}'(t)\| = \sqrt{4\sin^2 t + 4\cos^2 t} = 2\,, \qquad \forall t\,.$$

For the map $f(x,y) = (x-2)(y-1)+1$ we have $f\big(\boldsymbol{\gamma}(t)\big) = 4\sin t\cos t + 1$, hence

$$\int_{\gamma} f = 2\int_0^{2\pi}(4\sin t\cos t + 1)\,\mathrm{d}t = 2\big[-\cos 2t + t\big]_0^{2\pi} = 4\pi\,.$$

Using the curve $\boldsymbol{\gamma}_*$ with the same components of $\boldsymbol{\gamma}$ but t ranging in $[0, 2k\pi]$ (i.e., going around k times, instead of only once), we have

$$\int_{\gamma_*} f = 2\int_0^{2k\pi}(4\sin t\cos t + 1)\,\mathrm{d}t = 4k\pi. \qquad \square$$

This last example explains that curvilinear integrals depend not only on the trace of the curve, but also on the chosen parametrisation. Nonetheless, congruent parametrisations give rise to equal integrals, as we now show.

Let f be defined on the trace of a regular arc $\boldsymbol{\gamma} : [a,b] \to \mathbb{R}^m$ so that $f \circ \boldsymbol{\gamma}$ is (piecewise) continuous and the integral along $\boldsymbol{\gamma}$ exists. Then $f \circ \boldsymbol{\delta}$, too, where $\boldsymbol{\delta}$ is an arc congruent to $\boldsymbol{\gamma}$, will be (piecewise) continuous because composition of a continuous map between real intervals and the (piecewise-)continuous function $f \circ \boldsymbol{\gamma}$.

Proposition 9.3 *Let $\boldsymbol{\gamma} : [a,b] \to \mathbb{R}^m$ be a regular arc of trace Γ, and f a map on Γ such that $f \circ \boldsymbol{\gamma}$ is (piecewise) continuous. Then*

$$\int_{\delta} f = \int_{\gamma} f\,,$$

for any curve $\boldsymbol{\delta}$ congruent to $\boldsymbol{\gamma}$.

Proof. Let $\boldsymbol{\delta} = \boldsymbol{\gamma} \circ \varphi$, with $\varphi : [c,d] \to [a,b]$, be any congruent arc to $\boldsymbol{\gamma}$, so $\boldsymbol{\delta}'(\tau) = \boldsymbol{\gamma}'\big(\varphi(\tau)\big)\varphi'(\tau)$. Then

$$\int_{\boldsymbol{\delta}} f = \int_c^d f\big(\boldsymbol{\delta}(\tau)\big)\,\|\boldsymbol{\delta}'(\tau)\|\,\mathrm{d}\tau = \int_c^d f\big(\boldsymbol{\gamma}(\varphi(\tau))\big)\,\|\boldsymbol{\gamma}'\big(\varphi(\tau)\big)\varphi'(\tau)\|\,\mathrm{d}\tau$$
$$= \int_c^d f\big(\boldsymbol{\gamma}(\varphi(\tau))\big)\,\|\boldsymbol{\gamma}'\big(\varphi(\tau)\big)\|\,|\varphi'(\tau)|\,\mathrm{d}\tau\,.$$

Substituting $t = \varphi(\tau)$ gives $dt = \varphi'(\tau)\,d\tau$. Note that $\varphi(c) = a$, $\varphi(d) = b$ if $\varphi' > 0$ ($\boldsymbol{\delta}$ and $\boldsymbol{\gamma}$ are equivalent), or $\varphi(c) = b$, $\varphi(d) = a$ if $\varphi' < 0$ ($\boldsymbol{\delta}$ and $\boldsymbol{\gamma}$ are anti-equivalent). In the first case we obtain

$$\int_{\boldsymbol{\delta}} f = \int_a^b f\big(\boldsymbol{\gamma}(t)\big)\,\|\boldsymbol{\gamma}'(t)\|\,\mathrm{d}t = \int_{\boldsymbol{\gamma}} f$$

whereas in the second case

$$\int_{\boldsymbol{\delta}} f = -\int_b^a f\big(\boldsymbol{\gamma}(t)\big)\,\|\boldsymbol{\gamma}'(t)\|\,\mathrm{d}t = \int_a^b f\big(\boldsymbol{\gamma}(t)\big)\,\|\boldsymbol{\gamma}'(t)\|\,\mathrm{d}t = \int_{\boldsymbol{\gamma}} f\,.$$ □

By the previous proposition the following result is straightforward.

Corollary 9.4 *The curvilinear integral of a function does not change by taking opposite arcs.*

Given an arbitrary point c in (a,b) and setting $\boldsymbol{\gamma}_1 = \boldsymbol{\gamma}_{|[a,c]}$, $\boldsymbol{\gamma}_2 = \boldsymbol{\gamma}_{|[c,b]}$, by additivity we have

$$\int_{\boldsymbol{\gamma}} f = \int_{\boldsymbol{\gamma}_1} f + \int_{\boldsymbol{\gamma}_2} f\,. \tag{9.3}$$

This suggests how to extend the notion of integral to include piecewise-regular arcs. More precisely, let $\boldsymbol{\gamma} : [a,b] \to \mathbb{R}^m$ be a piecewise-regular arc and $a = a_0 < a_1 < \ldots < a_n = b$ points in $[a,b]$ such that the arcs $\boldsymbol{\gamma}_i = \boldsymbol{\gamma}_{|[a_{i-1},a_i]}$, $i = 1,\ldots,n$, are regular. Take f, as above, defined at least on Γ such that $f \circ \boldsymbol{\gamma}$ is (piecewise) continuous on $[a,b]$. By definition, then, we set

$$\boxed{\int_{\boldsymbol{\gamma}} f = \sum_{i=1}^{n} \int_{\boldsymbol{\gamma}_i} f\,.} \tag{9.4}$$

Remark 9.5 Computing integrals along curves is made easier by Proposition 9.3. In fact,

$$\boxed{\int_{\boldsymbol{\gamma}} f = \sum_{i=1}^{n} \int_{\boldsymbol{\delta}_i} f\,,} \tag{9.5}$$

where each $\boldsymbol{\delta}_i$ is an arc congruent to $\boldsymbol{\gamma}_i$, $i = 1,\ldots,n$, chosen in order to simplify the corresponding integral on the right. □

Example 9.6

Compute $\int_{\boldsymbol{\gamma}} x^2$, where $\boldsymbol{\gamma} : [0,4] \to \mathbb{R}^2$ is the following parametrisation of the square $[0,1] \times [0,1]$

$$\boldsymbol{\gamma}(t) = \begin{cases} \boldsymbol{\gamma}_1(t) = (t,0) & \text{if } 0 \le t < 1\,, \\ \boldsymbol{\gamma}_2(t) = (1,t-1) & \text{if } 1 \le t < 2\,, \\ \boldsymbol{\gamma}_3(t) = (3-t,1) & \text{if } 2 \le t < 3\,, \\ \boldsymbol{\gamma}_4(t) = (0,4-t) & \text{if } 3 \le t \le 4 \end{cases}$$

(see Fig. 9.2, left). Consider the parametrisations

$$\begin{aligned} &\boldsymbol{\delta}_1(t) = \boldsymbol{\gamma}_1(t) && 0 \le t \le 1\,, && \boldsymbol{\delta}_1 = \boldsymbol{\gamma}_1\,, \\ &\boldsymbol{\delta}_2(t) = (1,t) && 0 \le t \le 1\,, && \boldsymbol{\delta}_2 \sim \boldsymbol{\gamma}_2\,, \\ &\boldsymbol{\delta}_3(t) = (t,1) && 0 \le t \le 1\,, && \boldsymbol{\delta}_3 \sim -\boldsymbol{\gamma}_3\,, \\ &\boldsymbol{\delta}_4(t) = (0,t) && 0 \le t \le 1\,, && \boldsymbol{\delta}_4 \sim -\boldsymbol{\gamma}_4 \end{aligned}$$

(see Fig. 9.2, right). Then

$$\int_{\boldsymbol{\gamma}} x^2 = \int_0^1 t^2\,\mathrm{d}t + \int_0^1 1\,\mathrm{d}t + \int_0^1 t^2\,\mathrm{d}t + \int_0^1 0\,\mathrm{d}t = \frac{5}{3}\,. \qquad \square$$

Curvilinear integrals bear the same properties of linearity, positivity, monotonicity and so on, seen for ordinary integrals.

Remark 9.7 Let $\boldsymbol{\gamma} : [a,b] \to \mathbb{R}$ be a regular arc. Its length $\ell(\boldsymbol{\gamma})$ can be written as

$$\ell(\boldsymbol{\gamma}) = \int_{\boldsymbol{\gamma}} 1\,. \tag{9.6}$$

The arc length $s = s(t)$, see (6.20) with $t_0 = a$, satisfies $s(a) = 0$ and $s(b) = \int_a^b \|\boldsymbol{\gamma}'(\tau)\|\,\mathrm{d}\tau = \ell(\boldsymbol{\gamma})$. We want to use it to integrate a map f by means of the

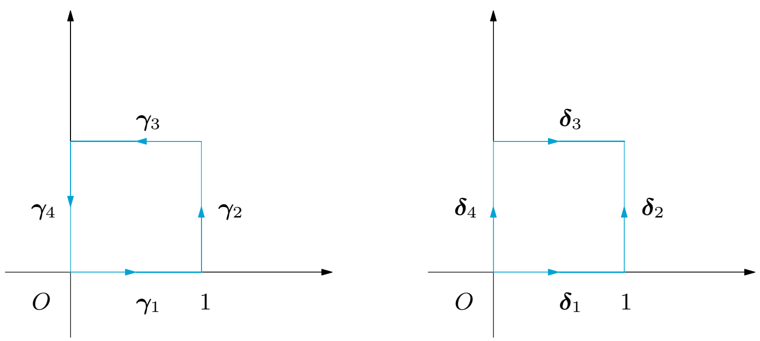


Figure 9.2. Parametrisations of the unit square relative to Example 9.6

parametrisation $\widetilde{\gamma}(s) = \gamma(t(s))$, equivalent to γ, already seen on p. 224. Then

$$\int_{\gamma} f = \int_{\widetilde{\gamma}} f = \int_0^{\ell(\gamma)} f\big(\widetilde{\gamma}(s)\big)\,\mathrm{d}s\,,$$

useful to simplify integrals if we know the arc length's analytical expression. □

In view of subsequent applications, it makes sense to consider integrals along a simple and regular arc Γ, thought of as a subset of $\mathbb{R}^m$, see Definition 6.28.

We already know that all parametrisations of Γ by simple, regular arcs are congruent (Proposition 6.27), and hence give rise to the same integral of a (piecewise-) continuous map f along Γ (Proposition 9.3). For this reason one calls **integral of** f **along** Γ the number

$$\boxed{\int_{\Gamma} f = \int_{\gamma} f\,,} \tag{9.7}$$

where γ is *any* regular and simple parametrisation of Γ. Equivalent symbols are

$$\int_{\Gamma} f\,\mathrm{d}s\,, \qquad \int_{\Gamma} f\,\mathrm{d}\gamma\,, \qquad \int_{\Gamma} f\,\mathrm{d}\ell\,.$$

The generalisation to piecewise-regular, simple arcs should be clear. In case Γ is closed, the integral might be denoted using the symbols

$$\oint_{\Gamma} f\,, \qquad \oint_{\Gamma} f\,\mathrm{d}s\,, \qquad \oint_{\Gamma} f\,\mathrm{d}\gamma\,, \qquad \oint_{\Gamma} f\,\mathrm{d}\ell\,.$$

Remark 9.8 The integral of a (piecewise-)continuous f can be defined along a regular curve $\gamma : I \to \mathbb{R}^m$, where I is a bounded, but *not closed*, interval. For this, $t \mapsto f\big(\gamma(t)\big)\|\gamma'(t)\|$ must be integrable (improperly as well) on I, with integral

$$\int_{\gamma} f = \int_I f\big(\gamma(t)\big)\|\gamma'(t)\|\,\mathrm{d}t\,.$$

For example, parametrise the unit semi-circle centred at the origin and lying on $y \geq 0$ by $\gamma : [-1,1] \to \mathbb{R}^2$, $\gamma(t) = (t, \sqrt{1-t^2})$; the curve is regular only on $(-1,1)$, for $\gamma'(t) = \big(1, -\frac{t}{\sqrt{1-t^2}}\big)$. If $f(x,y) = 1$,

$$\int_{\gamma} f = \int_{-1}^{1} \frac{1}{\sqrt{1-t^2}}\,\mathrm{d}t = \big[\arcsin t\big]_{-1}^{1} = \pi\,,$$

the right-hand-side integral being improper. □

9.1.1 Centre of mass and moments of a curve

As with thin plates and solid bodies, see Sect. 8.5, we can use curvilinear integrals to define certain physical quantities related to a thin wire resting along a (piecewise-)regular, simple arc $\Gamma \subset \mathbb{R}^3$.

Let $\mu = \mu(P)$ be the wire's linear density (mass per unit of length) at a generic $P = \boldsymbol{x} = (x, y, z) \in \Gamma$. The total mass will be

$$\boxed{m = \int_\Gamma \mu\,,}$$

and its centre of mass $G = \boldsymbol{x}_G = (x_G, y_G, z_G)$ is

$$\boxed{\boldsymbol{x}_G = \frac{1}{m}\int_\Gamma \boldsymbol{x}\mu\,,}$$

or

$$x_G = \frac{1}{m}\int_\Gamma x\mu\,, \qquad y_G = \frac{1}{m}\int_\Gamma y\mu\,, \qquad z_G = \frac{1}{m}\int_\Gamma z\mu\,.$$

The moment (of inertia) about a line or a point is

$$\boxed{I = \int_\Gamma d^2\mu\,,}$$

where $d = d(P)$ is the distance of $P \in \Gamma$ from the line or point considered. The axial moments

$$\boxed{I_x = \int_\Gamma (y^2+z^2)\mu\,, \qquad I_y = \int_\Gamma (x^2+z^2)\mu\,, \qquad I_z = \int_\Gamma (x^2+y^2)\mu\,,}$$

are special cases of the above, and their sum $I_0 = I_x + I_y + I_z$ represents the wire's moment about the origin.

Example 9.9

Let us determine the moment about the origin of the arc Γ

$$x^2+y^2+z^2 = 1\,, \qquad y = z\,, \qquad y \geq 0\,,$$

joining $A = (1,0,0)$ to $B = (0, \frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2})$. We can parametrise the arc by $\boldsymbol{\gamma}(t) = (\sqrt{1-2t^2}, t, t)$, $0 \leq t \leq \frac{\sqrt{2}}{2}$, so that

$$\|\boldsymbol{\gamma}'(t)\| = \sqrt{\frac{4t^2}{1-2t^2}+1+1} = \frac{\sqrt{2}}{\sqrt{1-2t^2}}\,,$$

and

$$I_0 = \int_0^{\sqrt{2}/2} (1-2t^2+t^2+t^2)\frac{\sqrt{2}}{\sqrt{1-2t^2}}\,dt = \big[\arcsin\sqrt{2}t\big]_0^{\sqrt{2}/2} = \frac{\pi}{2}\,. \qquad \square$$

9.2 Path integrals

Integrating vector fields along curves gives rise to the notion of path integral, presented below.

Let $I = [a,b]$ and $\boldsymbol{\gamma} : I \to \mathbb{R}^m$ be a regular arc with trace $\varGamma = \boldsymbol{\gamma}(I)$. Take a vector field $\boldsymbol{f}$ in $\mathbb{R}^m$, defined on $\varGamma$ at least. The composite map $\boldsymbol{f} \circ \boldsymbol{\gamma} : t \mapsto \boldsymbol{f}\big(\boldsymbol{\gamma}(t)\big)$ from I to $\mathbb{R}^m$ is thus defined. We shall assume the latter (piecewise) continuous, so that all $f_i\big(\boldsymbol{\gamma}(t)\big)$ are (piecewise) continuous from I to $\mathbb{R}$. For any $t \in I$, recall that

$$\boldsymbol{\tau}(t) = \boldsymbol{\tau}_{\boldsymbol{\gamma}}(t) = \frac{\boldsymbol{\gamma}'(t)}{\|\boldsymbol{\gamma}'(t)\|}$$

is the unit tangent vector at $P(t) = \boldsymbol{\gamma}(t)$. The map $f_\tau = \boldsymbol{f} \cdot \boldsymbol{\tau}$,

$$f_\tau(t) = \big(\boldsymbol{f} \cdot \boldsymbol{\tau}\big)(t) = \boldsymbol{f}\big(\boldsymbol{\gamma}(t)\big) \cdot \boldsymbol{\tau}(t)\,,$$

is the component of $\boldsymbol{f}$ along the tangent direction to $\boldsymbol{\gamma}$ at $P(t)$.

Definition 9.10 *The* **path integral of $\boldsymbol{f}$ along $\boldsymbol{\gamma}$** *is the integral along $\boldsymbol{\gamma}$ of the map f_τ:*

$$\int_{\boldsymbol{\gamma}} \boldsymbol{f} \cdot \boldsymbol{\tau} = \int_{\boldsymbol{\gamma}} f_\tau\,.$$

Another name for it is **line integral of $\boldsymbol{f}$ along $\boldsymbol{\gamma}$**.

By definition of $\boldsymbol{\tau}$ we have

$$\int_{\boldsymbol{\gamma}} \boldsymbol{f} \cdot \boldsymbol{\tau} = \int_a^b \boldsymbol{f}\big(\boldsymbol{\gamma}(t)\big) \cdot \boldsymbol{\tau}(t)\, \|\boldsymbol{\gamma}'(t)\|\, \mathrm{d}t = \int_a^b \boldsymbol{f}\big(\boldsymbol{\gamma}(t)\big) \cdot \boldsymbol{\gamma}'(t)\, \mathrm{d}t\,.$$

Thus, the integral of $\boldsymbol{f}$ along the path $\boldsymbol{\gamma}$ can be computed using

$$\boxed{\int_{\boldsymbol{\gamma}} \boldsymbol{f} \cdot \boldsymbol{\tau} = \int_a^b \boldsymbol{f}\big(\boldsymbol{\gamma}(t)\big) \cdot \boldsymbol{\gamma}'(t)\, \mathrm{d}t\,.} \tag{9.8}$$

Writing $\dfrac{\mathrm{d}\boldsymbol{\gamma}}{\mathrm{d}t} = \boldsymbol{\gamma}'(t)$ as Leibniz does, or $\mathrm{d}\boldsymbol{\gamma} = \boldsymbol{\gamma}'(t)\,\mathrm{d}t$, we may also use the notation

$$\int_{\boldsymbol{\gamma}} \boldsymbol{f} \cdot \mathrm{d}\boldsymbol{\gamma}\,.$$

There is a difference between the symbol $\mathrm{d}\boldsymbol{\gamma}$ (a vector) and the differential $\mathrm{d}\gamma$ of (9.2) (a scalar).

The physical meaning of path integrals is paramount. If $\boldsymbol{f}$ models a field of forces applied to the trace of the curve, the path integral is the work (the work integral) of the force during the motion along $\boldsymbol{\gamma}$. The counterpart to Proposition 9.3 is the following.

Proposition 9.11 *Let $\boldsymbol{\gamma} : [a,b] \to \mathbb{R}^m$ be a regular arc of trace Γ, $\boldsymbol{f}$ a vector field on Γ such that $\boldsymbol{f} \circ \boldsymbol{\gamma}$ is (piecewise) continuous. Then*

$$\int_{\boldsymbol{\delta}} \boldsymbol{f} \cdot \boldsymbol{\tau}_{\boldsymbol{\delta}} = \int_{\boldsymbol{\gamma}} \boldsymbol{f} \cdot \boldsymbol{\tau}_{\boldsymbol{\gamma}}\,, \qquad \textit{for any arc } \boldsymbol{\delta} \textit{ equivalent to } \boldsymbol{\gamma},$$

and

$$\int_{\boldsymbol{\delta}} \boldsymbol{f} \cdot \boldsymbol{\tau}_{\boldsymbol{\delta}} = -\int_{\boldsymbol{\gamma}} \boldsymbol{f} \cdot \boldsymbol{\tau}_{\boldsymbol{\gamma}}\,, \qquad \textit{for any arc } \boldsymbol{\delta} \textit{ anti-equivalent to } \boldsymbol{\gamma}.$$

Proof. This is a consequence of Proposition 9.3, because unit tangent vectors satisfy $\boldsymbol{\tau}_{\boldsymbol{\delta}} = \boldsymbol{\tau}_{\boldsymbol{\gamma}}$ when $\boldsymbol{\delta} \sim \boldsymbol{\gamma}$, and $\boldsymbol{\tau}_{\boldsymbol{\delta}} = -\boldsymbol{\tau}_{\boldsymbol{\gamma}}$ when $\boldsymbol{\delta} \sim -\boldsymbol{\gamma}$. □

Corollary 9.12 *Swapping an arc with its opposite changes the sign of the path integral:*

$$\int_{-\boldsymbol{\gamma}} \boldsymbol{f} \cdot \boldsymbol{\tau}_{(-\boldsymbol{\gamma})} = -\int_{\boldsymbol{\gamma}} \boldsymbol{f} \cdot \boldsymbol{\tau}_{\boldsymbol{\gamma}}\,.$$

In Physics it means that the work done by a force changes sign if one reverses the orientation of the arc; once that is fixed, the work depends only on the path and not on the way one moves along it.

Here is another recurring definition in the applications.

Definition 9.13 *If $\boldsymbol{\gamma}$ is a regular, closed arc, the path integral of $\boldsymbol{f}$ is said* **circulation of $\boldsymbol{f}$ along $\boldsymbol{\gamma}$** *and denoted*

$$\oint_{\boldsymbol{\gamma}} \boldsymbol{f} \cdot \boldsymbol{\tau}\,.$$

Examples 9.14

i) The vector field $\boldsymbol{f} : \mathbb{R}^3 \to \mathbb{R}^3$, $\boldsymbol{f}(x,y,z) = (\mathrm{e}^x, x+y, y+z)$ along $\boldsymbol{\gamma} : [0,1] \to \mathbb{R}^3$, $\boldsymbol{\gamma}(t) = (t, t^2, t^3)$, is given by

$$\boldsymbol{f}\big(\boldsymbol{\gamma}(t)\big) = (\mathrm{e}^t, t+t^2, t^2+t^3) \quad \text{with} \quad \boldsymbol{\gamma}'(t) = (1, 2t, 3t^2)\,.$$

Hence, the path integral of $\boldsymbol{f}$ on $\boldsymbol{\gamma}$ is

$$\begin{aligned}\int_{\boldsymbol{\gamma}} \boldsymbol{f} \cdot \boldsymbol{\tau} &= \int_0^1 (\mathrm{e}^t, t+t^2, t^2+t^3) \cdot (1, 2t, 3t^2)\,\mathrm{d}t \\ &= \int_0^1 \big[\mathrm{e}^t + 2(t^2+t^3) + 3(t^4+t^5)\big]\,\mathrm{d}t = \mathrm{e} + \frac{19}{15}\,.\end{aligned}$$

ii) Take the vector field $\boldsymbol{f} : \mathbb{R}^2 \to \mathbb{R}^2$ given by $\boldsymbol{f}(x,y) = (y,x)$. Parametrise the ellipse $\frac{x^2}{9} + \frac{y^2}{4} = 1$ by $\boldsymbol{\gamma} : [0, 2\pi] \to \mathbb{R}^2$, $\boldsymbol{\gamma}(t) = (3\cos t, 2\sin t)$. Then $\boldsymbol{f}\big(\boldsymbol{\gamma}(t)\big) = (2\sin t, 3\cos t)$ and $\boldsymbol{\gamma}'(t) = (-3\sin t, 2\cos t)$. The integral of $\boldsymbol{f}$ along the ellipse is zero:

$$\begin{aligned}\oint_{\boldsymbol{\gamma}} \boldsymbol{f}\cdot\boldsymbol{\tau} &= \int_0^{2\pi} (2\sin t, 3\cos t)\cdot(-3\sin t, 2\cos t)\,\mathrm{d}t = 6\int_0^{2\pi}(-\sin^2 t + \cos^2 t)\,\mathrm{d}t \\ &= 6\int_0^{2\pi}(2\cos^2 t - 1)\,\mathrm{d}t = 12\int_0^{2\pi}\cos^2 t\,\mathrm{d}t - 12\pi = 0\,,\end{aligned}$$

where we have used

$$\int_0^{2\pi}\cos^2 t\,\mathrm{d}t = \Big[\frac{1}{2}t + \frac{1}{4}\sin 2t\Big]_0^{2\pi} = \pi\,. \qquad \square$$

Path integrals exist on piecewise-regular arcs, too: it is enough to define, as in formula (9.4),

$$\boxed{\int_{\boldsymbol{\gamma}} \boldsymbol{f}\cdot\boldsymbol{\tau} = \sum_{i=1}^{n}\int_{\boldsymbol{\gamma}_i} \boldsymbol{f}\cdot\boldsymbol{\tau}_{\boldsymbol{\gamma}_i}\,,} \tag{9.9}$$

where $\boldsymbol{\gamma}_i$ are regular arcs constituting $\boldsymbol{\gamma}$.

At last, if we think geometrically (see Definition 6.28) and take a piecewise-regular simple arc $\Gamma \subset \mathbb{R}^m$, the path integral along Γ can be defined properly only after one orientation on Γ has been chosen, and the orientation depends on the tangent vector $\boldsymbol{\tau}$. Then we can define the **path integral of $\boldsymbol{f}$ along Γ** by

$$\boxed{\int_{\Gamma} \boldsymbol{f}\cdot\boldsymbol{\tau} = \int_{\boldsymbol{\gamma}} \boldsymbol{f}\cdot\boldsymbol{\tau}_{\boldsymbol{\gamma}}\,,} \tag{9.10}$$

where $\boldsymbol{\gamma}$ is *any* simple, (piecewise-)regular parametrisation of Γ with the chosen orientation. Clearly, reversing the orientation has the effect of changing the integral's sign. The circulation of $\boldsymbol{f}$ along a piecewise-regular Jordan arc (closed and simple) will be indicated by

$$\oint_{\Gamma} \boldsymbol{f}\cdot\boldsymbol{\tau}\,.$$

A different notation for the path integral, based on the language of *differential forms*, is provided in Appendix A.2.3, p. 529.

9.3 Integrals over surfaces

In perfect analogy to curves, the integral on a surface of a map in three variables is a natural way to extend double integrals over flat regions. This section is dedicated to integrating over surfaces. To review surfaces, see Sect. 4.7, whilst for their differential calculus see Sect. 6.7.

Let $\mathcal{R}$ be a compact, measurable region in $\mathbb{R}^2$, $\boldsymbol{\sigma} : \mathcal{R} \to \mathbb{R}^3$ a compact regular surface with trace Σ. The normal vector $\boldsymbol{\nu} = \boldsymbol{\nu}(u, v)$ at $P = \boldsymbol{\sigma}(u, v)$ was defined in (6.48). Take a map $f : \operatorname{dom} f \subseteq \mathbb{R}^3 \to \mathbb{R}$ defined on Σ at least, and assume $f \circ \boldsymbol{\sigma}$ is generically continuous on $\mathcal{R}$.

Definition 9.15 *The* **integral** *of* f **over the surface** $\boldsymbol{\sigma}$ *is the number*

$$\int_{\boldsymbol{\sigma}} f = \int_{\mathcal{R}} f\big(\boldsymbol{\sigma}(u, v)\big) \|\boldsymbol{\nu}(u, v)\| \, du \, dv \,. \tag{9.11}$$

As the surface is regular, the function $(u, v) \mapsto \|\boldsymbol{\nu}(u, v)\|$ is continuous on $\mathcal{R}$, so the right-hand-side integrand is generically continuous, hence integrable.

The definition is inspired by the following considerations. Suppose $\mathcal{R}$ is a rectangle with sides parallel to the (u, v)-axes, and let us divide it into rectangles $\mathcal{R}_{hk} = [u_h, u_h + \Delta u] \times [v_k, v_k + \Delta v]$ with lengths Δu, Δv, such that the interiors $\overset{\circ}{\mathcal{R}}_{hk}$ are pairwise disjoint (as in Fig. 9.3, left). The image of $\mathcal{R}_{hk}$ is the subset Σ_{hk} of Σ bounded by the coordinate lines through $\boldsymbol{\sigma}(u_h, v_k)$, $\boldsymbol{\sigma}(u_h + \Delta u, v_k)$, $\boldsymbol{\sigma}(u_h, v_k + \Delta v)$ and $\boldsymbol{\sigma}(u_h + \Delta u, v_k + \Delta v)$ (Fig. 9.3, right). If Δu and Δv are small enough, the surface is well approximated by the parallelogram Π_{hk} lying on the surface's tangent plane at $\boldsymbol{\sigma}(u_h, v_k)$, which is spanned by the tangent vectors $\dfrac{\partial \boldsymbol{\sigma}}{\partial u}(u_h, v_k)\Delta u$ and $\dfrac{\partial \boldsymbol{\sigma}}{\partial v}(u_h, v_k)\Delta v$ (Fig. 9.4). The area $\Delta\sigma$ of Π_{hk} is

$$\Delta\sigma = \left\| \frac{\partial \boldsymbol{\sigma}}{\partial u}(u_h, v_k) \wedge \frac{\partial \boldsymbol{\sigma}}{\partial v}(u_h, v_k) \right\| \Delta u \Delta v = \|\boldsymbol{\nu}(u_h, v_k)\| \Delta u \Delta v \,,$$

a number that approximates the area of Σ_{hk}. Therefore, we may consider the term

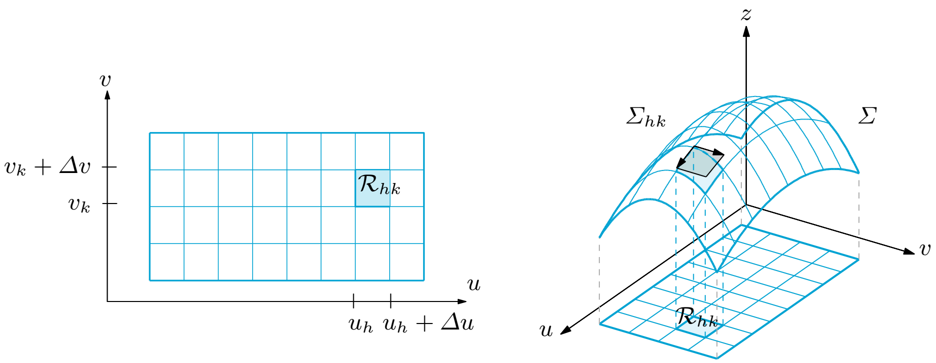


Figure 9.3. Partitions of the domain (left) and of the trace (right) of a compact surface

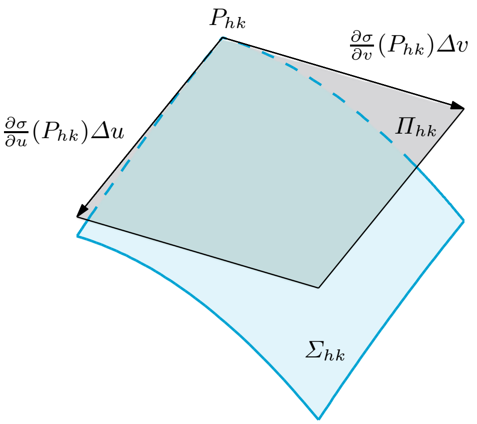


Figure 9.4. Element Σ_{hk} and corresponding tangent plane Π_{hk}

$$\boxed{\mathrm{d}\sigma = \|\boldsymbol{\nu}(u,v)\|\,\mathrm{d}u\,\mathrm{d}v}$$

in the right-hand side of (9.11) as an 'infinitesimal' surface element on Σ. In particular, if the surface is simple, we shall see that the area of Σ is precisely

$$\int_{\boldsymbol{\sigma}} 1 = \int_{\mathcal{R}} \|\boldsymbol{\nu}(u,v)\|\,\mathrm{d}u\,\mathrm{d}v\,.$$

Example 9.16

Let us suppose the regular compact surface $\boldsymbol{\sigma} : \mathcal{R} \to \mathbb{R}^3$ is given by $\boldsymbol{\sigma}(u,v) = u\boldsymbol{i} + v\boldsymbol{j} + \varphi(u,v)\boldsymbol{k}$. Recalling (6.49), we have

$$\|\boldsymbol{\nu}(u,v)\| = \sqrt{1 + \left(\frac{\partial\varphi}{\partial u}\right)^2 + \left(\frac{\partial\varphi}{\partial v}\right)^2}\,,$$

and so the integral of $f = f(x,y,z)$ on $\boldsymbol{\sigma}$ is

$$\int_{\boldsymbol{\sigma}} f = \int_{\mathcal{R}} f\big(u,v,\varphi(u,v)\big)\sqrt{1 + \left(\frac{\partial\varphi}{\partial u}\right)^2 + \left(\frac{\partial\varphi}{\partial v}\right)^2}\,\mathrm{d}u\,\mathrm{d}v\,.$$

For example, if $\boldsymbol{\sigma}$ is defined by $\varphi(u,v) = uv$ over the unit square $\mathcal{R} = [0,1]^2$, and $f(x,y,z) = z/\sqrt{1+x^2+y^2}$, then

$$\int_{\boldsymbol{\sigma}} f = \int_0^1\int_0^1 \frac{uv}{\sqrt{1+u^2+v^2}}\sqrt{1+v^2+u^2}\,\mathrm{d}u\,\mathrm{d}v$$
$$= \int_0^1\int_0^1 uv\,\mathrm{d}u\,\mathrm{d}v = \frac{1}{4}\,.$$
□

Surface integrals are invariant under congruent parametrisations, as stated by the next proposition.

Proposition 9.17 *Let $\boldsymbol{\sigma} : \mathcal{R} \to \mathbb{R}^3$ be a compact surface with trace Σ and f a map on Σ such that $f \circ \boldsymbol{\sigma}$ is generically continuous on $\mathcal{R}$. Then for any parametrisation $\widetilde{\boldsymbol{\sigma}} : \widetilde{\mathcal{R}} \to \Sigma$ congruent to $\boldsymbol{\sigma}$,*

$$\int_{\widetilde{\boldsymbol{\sigma}}} f = \int_{\boldsymbol{\sigma}} f .$$

Proof. Call $\boldsymbol{\Phi} : \widetilde{\mathcal{R}} \to \mathcal{R}$ the change of variables such that $\widetilde{\boldsymbol{\sigma}} = \boldsymbol{\sigma} \circ \boldsymbol{\Phi}$. By (6.50) and Theorem 8.24,

$$\begin{aligned}
\int_{\widetilde{\boldsymbol{\sigma}}} f &= \int_{\widetilde{\mathcal{R}}} f\big(\widetilde{\boldsymbol{\sigma}}(\widetilde{u},\widetilde{v})\big)\, \|\widetilde{\boldsymbol{\nu}}(\widetilde{u},\widetilde{v})\| \, \mathrm{d}\widetilde{u}\, \mathrm{d}\widetilde{v} \\
&= \int_{\widetilde{\mathcal{R}}} f\big(\boldsymbol{\sigma}(\boldsymbol{\Phi}(\widetilde{u},\widetilde{v}))\big)\, \|\det \boldsymbol{J\Phi}(\widetilde{u},\widetilde{v})\boldsymbol{\nu}\big(\boldsymbol{\Phi}(\widetilde{u},\widetilde{v})\big)\| \, \mathrm{d}\widetilde{u}\, \mathrm{d}\widetilde{v} \\
&= \int_{\widetilde{\mathcal{R}}} f\big(\boldsymbol{\sigma}(\boldsymbol{\Phi}(\widetilde{u},\widetilde{v}))\big)\, \|\boldsymbol{\nu}\big(\boldsymbol{\Phi}(\widetilde{u},\widetilde{v})\big)\| \, |\det \boldsymbol{J\Phi}(\widetilde{u},\widetilde{v})| \, \mathrm{d}\widetilde{u}\, \mathrm{d}\widetilde{v} \\
&= \int_{\mathcal{R}} f\big(\boldsymbol{\sigma}(u,v)\big)\, \|\boldsymbol{\nu}(u,v)\| \, \mathrm{d}u\, \mathrm{d}v = \int_{\boldsymbol{\sigma}} f .
\end{aligned}$$

□

The above result allows us to define integrals over regular, simple, compact surfaces $\Sigma \subset \mathbb{R}^2$ thought of as geometrical surfaces (see Definition 6.34). To be precise, one calls **integral** of f **on the surface** Σ the quantity

$$\boxed{\int_{\Sigma} f = \int_{\boldsymbol{\sigma}} f ,} \tag{9.12}$$

where $\boldsymbol{\sigma}$ is any regular, simple parametrisation of Σ. Since all such $\boldsymbol{\sigma}$ are congruent, the definition makes sense. Alternatively, one may also write

$$\int_{\Sigma} f(\boldsymbol{\sigma})\, \mathrm{d}\sigma \qquad \text{or} \qquad \int_{\Sigma} f \, \mathrm{d}\sigma .$$

Integrals can be defined over piecewise-regular compact surfaces Σ, meaning the union of n regular, simple compact surfaces $\Sigma_1, \ldots, \Sigma_n$ as of Definition 6.43. In such a case, one declares

$$\boxed{\int_{\Sigma} f = \sum_{i=1}^{n} \int_{\Sigma_i} f .}$$

Remark 9.18 The definition extends to cover non-compact surfaces, as we did for curvilinear integrals (Remark 9.8), with the proviso that the right-hand-side map of (9.11) be integrable on $\mathcal{R}$.

For instance, the unit hemisphere Σ defined by

$$x^2 + y^2 + z^2 = 1\,, \qquad z \geq 0\,,$$

is not compact in Cartesian coordinates, for

$$\boldsymbol{\sigma}(u,v) = \big(u, v, \sqrt{1-u^2-v^2}\big)$$

cannot be prolonged differentiably to an arbitrary open set containing the unit disc D. Nonetheless, we may still use (9.9) to find the surface's area, because the map

$$\|\boldsymbol{\nu}(u,v)\| = \frac{1}{\sqrt{1-u^2-v^2}}$$

is integrable on D. In fact,

$$\int_\Sigma 1 = \int_D \|\boldsymbol{\nu}(u,v)\|\,\mathrm{d}u\,\mathrm{d}v = \int_0^{2\pi}\int_0^1 \frac{r}{\sqrt{1-r^2}}\,\mathrm{d}r\,\mathrm{d}\theta = 2\pi\,.$$

□

9.3.1 Area of a surface

Via surface integrals we can define the **area of a compact surface** Σ (piecewise regular and simple) thoroughly, by

$$\boxed{\operatorname{area}(\Sigma) = \int_\Sigma 1\,.}$$

Example 9.19

i) The example of Remark 9.18 adapts to show that the area of the hemisphere Σ of radius r is

$$\operatorname{area}(\Sigma) = \int_\Sigma 1 = 2\pi r^2\,,$$

as elementary geometry tells us.

ii) Let us compute the lateral surface Σ of a cylinder of radius r and height L. Supposing the cylinder's axis is the segment $[0, L]$ along the z-axis, the surface Σ is parametrised by $\boldsymbol{\sigma}(u,v) = r\cos u\,\boldsymbol{i} + r\sin u\,\boldsymbol{j} + v\,\boldsymbol{k}$, $(u,v) \in \mathcal{R} = [0,2\pi]\times[0,L]$. Easily then,

$$\boldsymbol{\nu}(u,v) = r\cos u\,\boldsymbol{i} + r\sin u\,\boldsymbol{j} + 0\,\boldsymbol{k}\,,$$

whence $\|\boldsymbol{\nu}(u,v)\| = r$. In conclusion,

$$\operatorname{area}(\Sigma) = \int_0^{2\pi}\int_0^L r\,\mathrm{d}u\,\mathrm{d}v = 2\pi r L\,,$$

another old acquaintance of elementary geometry's fame. □

It is no coincidence that the area is found by integrating along a meridian arc. The cylinder is in fact a surface of revolution (Example 4.37 iii)):

Proposition 9.20 *Let Σ define the surface of revolution generated by revolving the arc Γ, on the plane xz, around the z-axis. Then*

$$\text{area}(\Sigma) = 2\pi \int_\Gamma x\,.$$

Proof. Retaining the notation of Example 4.37 iii) we easily obtain

$$\begin{aligned}\boldsymbol{\nu}(u,v) &= -\gamma_1(u)\gamma_3'(u)\cos v\,\boldsymbol{i} + \gamma_1(u)\gamma_3'(u)\sin v\,\boldsymbol{j}\\ &\qquad + \big(\gamma_1(u)\gamma_1'(u)\cos^2 v + \gamma_1(u)\gamma_1'(u)\sin^2 v\big)\,\boldsymbol{k}\,,\end{aligned}$$

whence

$$\|\boldsymbol{\nu}(u,v)\| = \sqrt{\gamma_1^2(u)\big((\gamma_1'(u))^2 + (\gamma_3'(u))^2\big)} = \gamma_1(u)\|\boldsymbol{\gamma}'(u)\|\,.$$

Above we assumed $x = \gamma_1(u)$ non-negative along the curve. Therefore

$$\begin{aligned}\text{area}(\Sigma) &= \int_\Sigma 1 = \int_0^{2\pi}\int_I \gamma_1(u)\|\boldsymbol{\gamma}'(u)\|\,du dv\\ &= 2\pi\int_I \gamma_1(u)\|\boldsymbol{\gamma}'(u)\|\,du dv = 2\pi\int_\Gamma x\,.\end{aligned}$$

□

As $2\pi x$ is the length of the circle described by $P = (x,0,z) \in \Gamma$ during the revolution around the axis, and

$$x_G = \frac{\int_\Gamma x}{\int_\Gamma 1} = \frac{\int_\Gamma x}{\ell(\Gamma)}$$

is the coordinate of the centre of mass G of Γ, corresponding to unit density along the curve, we can state the formula as

$$\boxed{\text{area}(\Sigma) = 2\pi\, x_G\, \ell(\Gamma)\,.}$$

This is known as Guldin's Theorem.

Theorem 9.21 *The area of a surface of revolution Σ is the product of the length of the meridian section times the length of the circle described by the arc's centre of mass.*

9.3.2 Centre of mass and moments of a surface

Imagine a thin shell covering a (piecewise-)regular, simple, compact surface $\Sigma \subset \mathbb{R}^3$. Denoting by $\mu = \mu(P)$ the shell's density of mass (mass per unit of area) at $P \in \Sigma$, we can compute the total mass

$$\boxed{m = \int_\Sigma \mu}$$

and the centre of mass $G = (x_G, y_G, z_G)$

$$\boxed{\boldsymbol{x}_G = \frac{1}{m}\int_\Sigma \boldsymbol{x}\mu\,,}$$

so

$$x_G = \frac{1}{m}\int_\Sigma x\mu\,, \qquad y_G = \frac{1}{m}\int_\Sigma y\mu\,, \qquad z_G = \frac{1}{m}\int_\Sigma z\mu\,.$$

The moment (of inertia) about an axis or a point is

$$\boxed{I = \int_\Sigma d^2\mu\,,}$$

where $d = d(P)$ is the distance of the generic $P \in \Sigma$ from the line or point given. The moments about the coordinate axes

$$\boxed{I_x = \int_\Sigma (y^2+z^2)\mu\,, \qquad I_y = \int_\Sigma (x^2+z^2)\mu\,, \qquad I_z = \int_\Sigma (x^2+y^2)\mu\,,}$$

are special cases; their sum $I_0 = I_x + I_y + I_z$ represents the moment about the origin.

9.4 Flux integrals

After learning how to integrate scalar fields on surfaces, we now turn to vector-valued maps and define the fundamental notion of flux integral, which will occupy this section.

Let $\boldsymbol{f}$ be a vector field on $\mathbb{R}^3$ that is defined (at least) on the trace Σ of a regular, compact surface $\boldsymbol{\sigma} : \mathcal{R} \to \Sigma$, with $\mathcal{R}$ compact and measurable in $\mathbb{R}^2$. We assume the composite $\boldsymbol{f} \circ \boldsymbol{\sigma}$ is generically continuous on $\mathcal{R}$. If $\boldsymbol{n} = \boldsymbol{n}(u, v)$ is the unit normal to Σ, one calls *normal component* of $\boldsymbol{f}$ the component $f_n = \boldsymbol{f} \cdot \boldsymbol{n}$ along $\boldsymbol{n}$.

Definition 9.22 *The* **surface integral** *of* $\boldsymbol{f}$ *on* $\boldsymbol{\sigma}$ *is the integral on the surface* $\boldsymbol{\sigma}$ *of the map* f_n*:*

$$\int_{\boldsymbol{\sigma}} \boldsymbol{f}\cdot\boldsymbol{n} = \int_{\boldsymbol{\sigma}} f_n\,.$$

Because of Definition 6.35 we can compute surface integrals by

$$\boxed{\int_{\boldsymbol{\sigma}} \boldsymbol{f}\cdot\boldsymbol{n} = \int_{\mathcal{R}} \boldsymbol{f}\big(\boldsymbol{\sigma}(u,v)\big)\cdot\boldsymbol{\nu}(u,v)\,\mathrm{d}u\,\mathrm{d}v\,.} \tag{9.13}$$

As already seen for path integrals, congruent parametrisations in surface integrals possibly entail a sign ambiguity.

Proposition 9.23 *Let* $\boldsymbol{\sigma}:\mathcal{R}\to\mathbb{R}^3$ *be a regular compact surface with trace* Σ*,* $\boldsymbol{f}$ *a vector field on* Σ *such that* $\boldsymbol{f}\circ\boldsymbol{\sigma}$ *is generically continuous. Then*

$$\int_{\widetilde{\boldsymbol{\sigma}}} \boldsymbol{f}\cdot\widetilde{\boldsymbol{n}} = \int_{\boldsymbol{\sigma}} \boldsymbol{f}\cdot\boldsymbol{n}\,, \qquad \textit{for any compact surface } \widetilde{\boldsymbol{\sigma}} \textit{ equivalent to } \boldsymbol{\sigma},$$

and

$$\int_{\widetilde{\boldsymbol{\sigma}}} \boldsymbol{f}\cdot\widetilde{\boldsymbol{n}} = -\int_{\boldsymbol{\sigma}} \boldsymbol{f}\cdot\boldsymbol{n}\,, \qquad \textit{for any compact surface } \widetilde{\boldsymbol{\sigma}} \textit{ anti-equivalent to } \boldsymbol{\sigma}\,.$$

Proof. The proof descends from Proposition 9.17 by observing that unit normals are equal, $\widetilde{\boldsymbol{n}}=\boldsymbol{n}$, if $\widetilde{\boldsymbol{\sigma}}$ is equivalent to $\boldsymbol{\sigma}$, and opposite, $\widetilde{\boldsymbol{n}}=-\boldsymbol{n}$, if $\widetilde{\boldsymbol{\sigma}}$ is anti-equivalent to $\boldsymbol{\sigma}$. □

The proposition allows us to define the surface integral of a vector field $\boldsymbol{f}$ over a (piecewise-)regular, simple, compact surface Σ seen as embedded in $\mathbb{R}^3$. In the light of Sect. 6.7.2, it is though necessary to consider orientable surfaces only (Definition 6.38).

Let us thus assume that Σ is orientable, and that we have fixed an orientation on it, corresponding to one of the unit normals, henceforth denoted $\boldsymbol{n}$. Now we are in the position to define the **flux integral of $\boldsymbol{f}$ on** Σ as

$$\boxed{\int_{\Sigma} \boldsymbol{f}\cdot\boldsymbol{n} = \int_{\boldsymbol{\sigma}} \boldsymbol{f}\cdot\boldsymbol{n}\,,} \tag{9.14}$$

where $\boldsymbol{\sigma}$ is any simple and (piecewise-)regular parametrisation of Σ inducing the chosen orientation. This integral is often called simply **flux of $\boldsymbol{f}$ across**, or **through**, Σ. The terminology stems from Physics; suppose the surface is immersed in a fluid of density $\mu=\mu(\boldsymbol{x})$, and $\boldsymbol{v}=\boldsymbol{v}(\boldsymbol{x})$ is the velocity of the point-particle at $P = \boldsymbol{x}$. Set $\boldsymbol{f} = \mu\boldsymbol{v}$ and denote by $\Delta\Sigma$ the element of area $\Delta\sigma$ and normal $\boldsymbol{n}$. Then $(\boldsymbol{f}\cdot\boldsymbol{n})\,\Delta\sigma$ is the volume rate of fluid flow through $\Delta\Sigma$ per unit of time, i.e., the discharge. Summing, and passing to the limit, the flux integral of $\boldsymbol{f}$ across Σ is the difference between the overall outflow and inflow through the surface Σ.

When the surface Σ is bounded and encloses an open domain Ω, one speaks about **outgoing flux** or **ingoing flux** according to whether the normal $\boldsymbol{n}$ leaves Ω or enters Ω respectively.

Example 9.24

Let us determine the outgoing flux of $\boldsymbol{f}(\boldsymbol{x}) = y\boldsymbol{i} - x\boldsymbol{j} + z\boldsymbol{k}$ through the sphere centred in the origin and of radius r. We opt for spherical coordinates and parametrise by $\boldsymbol{\sigma} : [0,\pi]\times[0,2\pi] \to \mathbb{R}^3$, $\boldsymbol{\sigma}(u,v) = r\sin u\cos v\,\boldsymbol{i} + r\sin u\sin v\,\boldsymbol{j} + r\cos u\,\boldsymbol{k}$ (see Example 4.37 iv)). The outgoing normal is

$$\boldsymbol{\nu}(\boldsymbol{x}) = r\sin u\,\boldsymbol{x} = r\sin u(x\boldsymbol{i} + y\boldsymbol{j} + z\boldsymbol{k})\,,$$

see Example 6.36 ii). Therefore $(\boldsymbol{f}\cdot\boldsymbol{\nu})(\boldsymbol{x}) = r\sin u z^2 = r^3\sin u\cos^2 u$, and recalling (9.13), we have

$$\int_\Sigma \boldsymbol{f}\cdot\boldsymbol{n} = \int_0^{2\pi}\int_0^{\pi} r^3\sin u\cos^2 u\,\mathrm{d}u\mathrm{d}v = \frac{4}{3}\pi r^3\,. \qquad \square$$

For the flux integral, too, we may use an alternative notation based on the language of differential forms; see Appendix A.2.3, p. 529.

9.5 The Theorems of Gauss, Green, and Stokes

The three theorems of the title should be considered multi-dimensional versions of the formula of integration by parts. Each one of them allows to transform the integral of an expression involving the differential operators of Sect. 6.3.1 on a two- or three-dimensional domain, or a surface, into the integral over the boundary of an expression without derivatives.

The importance of such results is paramount, both from the theoretical point of view and in relationship to applications. They typically manifest themselves, for example, when one formulates a law of Physics (e.g., the conservation of mass or energy) in mathematical language (a PDE); but they may also play a role in determining the conditions that guarantee the solvability of said equations (existence and uniqueness of solutions); at last, several numerical techniques for solving equations (such as the finite-volume method and the finite-element method) are implemented by using one of the theorems.

At a more immediate level, these results enable to simplify an integral, and crop up when examining special vector fields, like conservatives fields.

We start by discussing a class of open sets and surfaces, that we will call admissible, for which the theorems hold. A first, basic study of Sects. 9.5.2 - 9.5.4 does not require such level of detail.

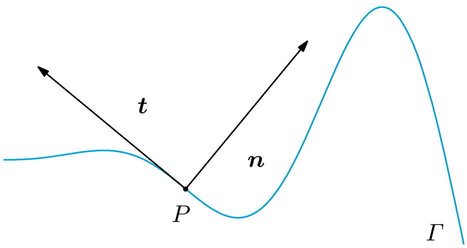


Figure 9.5. Unit tangent to a portion of the Jordan arc Γ, and unit normal rotated by $\pi/2$

9.5.1 Open sets, admissible surfaces and boundaries

Open sets in the plane. Let $\Gamma \subset \mathbb{R}^2$ be a piecewise-regular Jordan arc with a given orientation. The unit tangent vector $\boldsymbol{t} = t_1\boldsymbol{i} + t_2\boldsymbol{j}$ exists at all points P of Γ, with the exception of a finite number. At P the normal direction $\boldsymbol{v}$ to Γ (orthogonal to $\boldsymbol{t}$) is thus well defined. In particular, the unit vector $\boldsymbol{n} = n_1\boldsymbol{i}+n_2\boldsymbol{j} = t_2\boldsymbol{i}-t_1\boldsymbol{j}$ to Γ at P is obtained rotating $\boldsymbol{t}$ clockwise by $\pi/2$ (Fig. 9.5); otherwise said, identifying $\boldsymbol{n}$ and $\boldsymbol{t}$ with $\boldsymbol{n} + 0\boldsymbol{k}$ and $\boldsymbol{t} + 0\boldsymbol{k}$ in $\mathbb{R}^3$ makes the triple $(\boldsymbol{n}, \boldsymbol{t}, \boldsymbol{k})$ right-handed, for $(\boldsymbol{n} \wedge \boldsymbol{t}) \cdot \boldsymbol{k} = 1$. (Notice that the unit vector $\boldsymbol{n}$ might not coincide with the principal normal of (6.22), whose orientation varies with the curve's convexity; at any rate the two vectors clearly differ by a sign, at most.) Furthermore, Γ separates the region inside Γ from the external region, by virtue of Jordan's Curve Theorem 4.33.

> **Definition 9.25** *We call a bounded open set $\Omega \subset \mathbb{R}^2$* ***G-admissible*** *if the following hold:*
>
> *i) the boundary $\partial\Omega$ is a finite union of piecewise-regular, pairwise-disjoint Jordan arcs $\Gamma_1, \ldots, \Gamma_K$;*
>
> *ii) Ω is entirely contained either inside, or outside, each Γ_k.*

Each point $P \in \partial\Omega$ will belong to one, and one only, Jordan arc Γ_k, so there will be (save for a finite number of points) a unit normal $\boldsymbol{n}$ to Γ_k at P, that is chosen to point **outwards** Ω (precisely, all points $Q = P + \varepsilon\boldsymbol{n}$, with $\varepsilon > 0$ sufficiently small, lie outside Ω). We will say $\boldsymbol{n}$ is the **outgoing unit normal to** $\partial\Omega$, or for short, the **outgoing normal** of $\partial\Omega$.

The choice of the outward-pointing orientation induces an orientation on the boundary $\partial\Omega$. In fact, on every arc Γ_k we will fix the orientation so that, if $\boldsymbol{t}$ denotes the tangent vector, the frame $(\boldsymbol{n}, \boldsymbol{t}, \boldsymbol{k})$ is oriented positively. Intuitively, one could say that **a three-dimensional observer standing as $\boldsymbol{k}$ and walking along Γ_k will see Ω on his left** (Fig. 9.6). We shall call this the **positive** orientation of $\partial\Omega$ (and the opposite one negative).

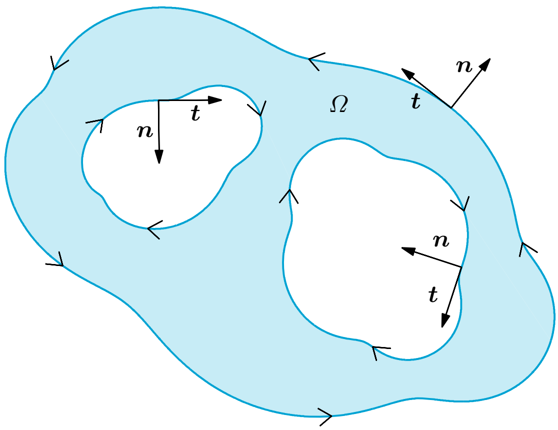


Figure 9.6. A G-admissible open set in the plane

Open sets in space. The whole discussion on two-dimensional G-admissible open sets extends easily to space. For this, we recall that every closed and orientable surface divides space in a region enclosed by the surface and one outside it (Theorem 6.42).

Definition 9.26 *We call a bounded open set $\Omega \subset \mathbb{R}^3$* ***G-admissible*** *if:*

i) its boundary $\partial\Omega$ is the union of a finite number of pairwise-disjoint surfaces $\Sigma_1, \ldots, \Sigma_K$;

ii) each Σ_k is piecewise regular, simple, orientable and closed;

iii) Ω lies entirely inside or outside every surface Σ_k.

For a given G-admissible open set there is a well-defined **outgoing** normal $\boldsymbol{n}$ to $\partial\Omega$, which will coincide with the normal to the surface Σ_k oriented from the inside towards the outside of Ω (Fig. 9.7).

Here are some examples of G-admissible open sets.

Examples 9.27

i) The inside of the elementary solids (e.g., parallelepipeds, polyhedra, cylinders, cones, spheres), and of any regular deformation of these, are G-admissible.

ii) We say an open bounded set $\Omega \subset \mathbb{R}^3$ is **regular and normal** for z in case Ω is normal for z as in Definition 8.27,

$$\Omega = \big\{(x, y, z) \in \mathbb{R}^3 : (x, y) \in D, \alpha\,(x, y) < z < \beta(x, y)\big\}\,,$$

where D is open in $\mathbb{R}^2$ with boundary ∂D a piecewise-regular Jordan arc, and α, β are $\mathcal{C}^1$ maps on D. Such an open set is G-admissible.

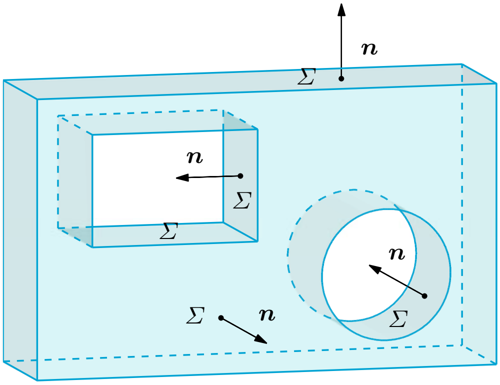


Figure 9.7. A G-admissible open set in $\mathbb{R}^3$ (with $K = 1$)

The boundary of Ω consists of a single surface Σ_1, piecewise regular, normal, orientable and closed; in fact, we can decompose $\Sigma_1 = \Sigma_\beta \cup \Sigma_\alpha \cup \Sigma_\ell$, where

$$\Sigma_\beta = \{(x, y, \beta(x, y)) \in \mathbb{R}^3 : (x, y) \in D\},$$

$$\Sigma_\alpha = \{(x, y, \alpha(x, y)) \in \mathbb{R}^3 : (x, y) \in D\},$$

$$\Sigma_\ell = \{(x, y, z) \in \mathbb{R}^3 : (x, y) \in \partial D, \alpha(x, y) \le z \le \beta(x, y)\}.$$

The outgoing unit normal $\boldsymbol{n}_\beta$ to Σ_β is obtained by normalising

$$\boldsymbol{\nu}_\beta(x, y) = -\frac{\partial \beta}{\partial x}(x, y)\,\boldsymbol{i} - \frac{\partial \beta}{\partial y}(x, y)\,\boldsymbol{j} + \boldsymbol{k},$$

Σ_β being a local graph (see Example 6.36, i)).
Similarly, the outgoing unit normal $\boldsymbol{n}_\alpha$ to Σ_α is the unit vector corresponding to

$$\boldsymbol{\nu}_\alpha(x, y) = \frac{\partial \alpha}{\partial x}(x, y)\,\boldsymbol{i} + \frac{\partial \alpha}{\partial y}(x, y)\,\boldsymbol{j} - \boldsymbol{k}.$$

The outgoing normal $\boldsymbol{n}_\ell$ to Σ_ℓ is

$$\boldsymbol{n} = \boldsymbol{n}_{\partial D} + 0\boldsymbol{k},$$

where $\boldsymbol{n}_{\partial D}$ is the outgoing unit normal (in two dimensions) of ∂D.
Regular and normal sets for x or y are defined in the same way.

iii) Let us see how to generalise the above situation. The open sets $\Omega_1, \ldots, \Omega_K$ form a **partition** of an open set Ω if the Ω_k are pairwise disjoint and the union of their closures coincides with the closure of Ω:

$$\overline{\Omega} = \bigcup_{k=1}^{K} \overline{\Omega}_k \qquad \text{with} \qquad \Omega_h \cap \Omega_k = \emptyset \quad \text{if } h \ne k.$$

We then say an open bounded set Ω of $\mathbb{R}^3$ is **piecewise regular and normal** for x_i ($i = 1, 2, 3$) if it admits a partition into open, regular, normal sets for x_i (see Fig. 9.8 for the two-dimensional picture). Such Ω is G-admissible. □

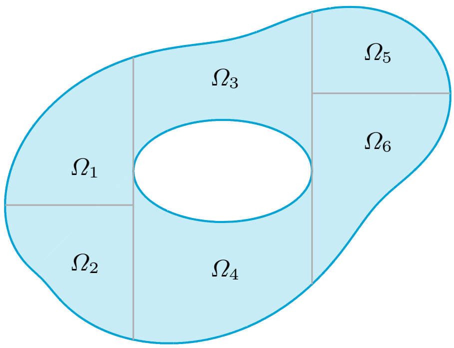


Figure 9.8. Partition of $\Omega \subset \mathbb{R}^2$ into the union of normal sets for y

Compact surfaces. Henceforth Σ will denote a piecewise-regular, normal and orientable compact surface as of Sect. 6.7.4; let Σ_k, $k = 1, \ldots, K$, be its faces, each of which is normal, regular and compact.

The notion of S-admissible compact surfaces is relevant in view of Stokes' Theorem.

Definition 9.28 *We call S-**admissible** a compact, piecewise-regular, normal, orientable surface Σ whose faces Σ_k, $k = 1, \ldots, K$, can be parametrised by maps $\boldsymbol{\sigma}_k : \mathcal{R}_k \to \Sigma_k$ where $\mathcal{R}_k = \overline{\Omega}_k$ is the closure of a G-admissible open set Ω_k in the plane.*

Given such an S-admissible, compact surface Σ, we will assume to have fixed one orientation by choosing a normal $\boldsymbol{n}$ to Σ. On each face Σ_k, $\boldsymbol{n}$ coincides with one of the unit normals to Σ_k, say $\boldsymbol{n}_k$. Without loss of generality we may suppose $\boldsymbol{n}_k$ is the unit normal associated to the parametrisation $\boldsymbol{\sigma}_k$, see Definition 6.35. (If not, it is enough to swap $\boldsymbol{\sigma}_k$ with $\widetilde{\boldsymbol{\sigma}}_k$ on $\widetilde{\mathcal{R}}_k = \{(u, v) \in \mathbb{R}^2 : (-u, v) \in \mathcal{R}_k\}$ given by $\widetilde{\boldsymbol{\sigma}}_k(u, v) = (-u, v)$, whose unit normal is opposite to that of $\boldsymbol{\sigma}_k$.)

By compactness, the unit normal $\boldsymbol{n}$ is defined right up to the boundary of Σ (with the exception of finitely many points, at most). For this reason we can choose an orientation on $\partial\Sigma$. Roughly speaking, **the positive orientation is given by the walking direction of an ideal observer standing as $\boldsymbol{n}$ that proceeds along the boundary and keeps the surface at his left**.

A more accurate definition requires a little extra work. Every point P belonging to $\partial\Sigma$, except for a finite number, lies on the boundary of exactly one face Σ_k (Fig. 9.9); moreover, around P the boundary $\partial\Sigma_k$ is a regular arc Γ_k in $\mathbb{R}^3$, given by the image under $\boldsymbol{\sigma}_k$ of a regular arc Δ_k in $\mathbb{R}^2$ contained in the boundary of a region $\mathcal{R}_k$. In other terms, $\Delta_k = \{\boldsymbol{\gamma}_k(t) : t \in I_k\}$, where $\boldsymbol{\gamma}_k$ is the parametrisation of the Jordan arc containing Δ_k, and correspondingly $\Gamma_k = \{\boldsymbol{\eta}_k(t) = \boldsymbol{\sigma}_k\big(\boldsymbol{\gamma}_k(t)\big) : t \in I_k\}$; the point $P \in \Gamma_k$ will be image under $\boldsymbol{\sigma}_k$ of a point $\boldsymbol{p}_0 \in \Delta_k$ identified by $t = t_0$, hence $P = \boldsymbol{\sigma}_k(\boldsymbol{p}_0) = \boldsymbol{\sigma}_k\big(\boldsymbol{\gamma}_k(t_0)\big) = \boldsymbol{\eta}_k(t_0)$. To the (column) vector $\boldsymbol{\tau} = \boldsymbol{\gamma}_k'(t_0)$,

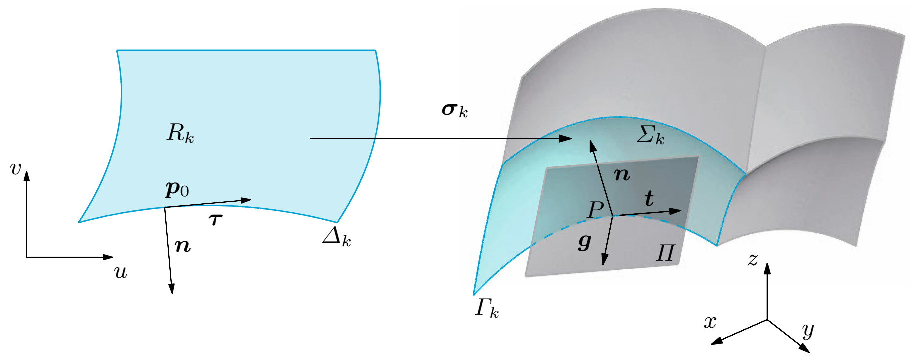


Figure 9.9. The positive orientation of a compact surface's boundary

tangent to Δ_k at $\boldsymbol{p}_0$, corresponds the (column) vector

$$\frac{\mathrm{d}\boldsymbol{\eta}_k}{\mathrm{d}t}(t_0) = \boldsymbol{J\sigma}_k(\boldsymbol{p}_0)\gamma_k'(t_0)$$

(recall the chain rule), tangent to Γ_k at P. Let $\boldsymbol{t}$ be the corresponding unit vector. Then, if we assume that $\boldsymbol{\gamma}_k$ induces the positive orientation on the Jordan arc where Δ_k lies, we will say **the orientation of the arc $\Gamma_k \subset \partial\Sigma$ induced by the unit vector $\boldsymbol{t}$ is positive**. We can say the same in the following way. Let Π be the tangent plane to Σ at P that contains $\boldsymbol{t}$, and denote by $\boldsymbol{g}$ the unit vector orthogonal to $\boldsymbol{t}$, lying on Π and pointing outside Σ; then the positive orientation of $\partial\Sigma$ is the one rendering $(\boldsymbol{g}, \boldsymbol{t}, \boldsymbol{n})$ a positively-oriented triple (see again Fig. 9.9).

Example 9.29

Suppose Σ is given by

$$\Sigma = \{(x,y,z) \in \mathbb{R}^3 : (x,y) \in \mathcal{R}, z = \varphi(x,y)\}\,, \tag{9.15}$$

where $\mathcal{R}$ is a closed, bounded region of the plane, the boundary $\partial\mathcal{R}$ is a regular Jordan arc Γ parametrised by $\boldsymbol{\gamma} : I \to \Gamma$, and φ is a $\mathcal{C}^1$ map on the open set A containing $\mathcal{R}$; let Ω indicate the region inside Γ.
The surface Σ is thus compact, S-admissible and parametrised by $\boldsymbol{\sigma} : \mathcal{R} \to \Sigma$, $\boldsymbol{\sigma}(x,y) = \big(x, y, \varphi\,(x,y)\big)$. The corresponding unit normal is

$$\boldsymbol{n} = \frac{\boldsymbol{\nu}}{\|\boldsymbol{\nu}\|}\,, \qquad \text{where} \qquad \boldsymbol{\nu} = -\varphi_x \boldsymbol{i} - \varphi_y \boldsymbol{j} + \boldsymbol{k}\,. \tag{9.16}$$

If $\boldsymbol{\gamma}$ parametrises Γ with positive orientation (counter-clockwise), also the boundary

$$\partial\Sigma = \big\{\boldsymbol{\eta}(t) = \big(\gamma_1(t), \gamma_2(t), \varphi(\gamma_1(t), \gamma_2(t)\big) : t \in I\big\}$$

will be positively oriented, that is to say, Σ lies constantly on its left.

The corresponding unit tangent to $\partial\Sigma$ is $\boldsymbol{t} = \frac{\boldsymbol{\eta}'}{\|\boldsymbol{\eta}'\|}$, where

$$\boldsymbol{\eta}'(t) = \big(\gamma_1'(t), \gamma_2'(t), \varphi_x(\boldsymbol{\gamma}(t))\gamma_1'(t) + \varphi_y(\boldsymbol{\gamma}(t))\gamma_2'(t)\big)\,. \tag{9.17}$$

□

9.5.2 Divergence Theorem

The theorem in question asserts that under suitable hypotheses the integral of a field's divergence over an open bounded set of $\mathbb{R}^n$ is the flux integral of the field across the boundary: letting Ω denote the open set and $\boldsymbol{n}$ the outward normal to the boundary $\partial\Omega$, we have

$$\int_\Omega \operatorname{div} \boldsymbol{f} = \int_{\partial\Omega} \boldsymbol{f}\cdot\boldsymbol{n}\,. \tag{9.18}$$

We held back on purpose from making precise hypotheses on Ω and $\boldsymbol{f}$, as multiple possibilities exist. The theorem's proofs may be sensibly simplified by sufficiently restrictive assumptions, at the cost of diminishing its far-reaching impact. Finding 'minimal' hypotheses for its validity is a task beyond the scope of our study.

In the sequel the Divergence Theorem will be stated under the assumption that the domain Ω be G-admissible, as discussed in Sect. 9.5.1; far from being the most general, our statement will hold nonetheless in the majority of cases of interest. The reader that wishes to skip the details might think of a G-admissible set as an open bounded set whose boundary is made of finitely many graphs of regular maps, locally viewing the open set on the same side. The outward normal of the open set will be the outward normal of each graph.

Let us begin in dimension three, by the following preparatory, but relevant irrespectively, result. The proof is available in Appendix A.2.2, p. 524, under more stringent, but still significant, assumptions on the domain. Hereafter, Cartesian coordinates will be equivalently denoted by x_1, x_2, x_3 or by x, y, z.

Proposition 9.30 *Let the open set $\Omega \subset \mathbb{R}^3$ be G-admissible, and assume $f \in \mathcal{C}^0(\overline{\Omega})$ with $\dfrac{\partial f}{\partial x_i} \in \mathcal{C}^0(\overline{\Omega})$, $i \in \{1,2,3\}$. Then*

$$\int_\Omega \frac{\partial f}{\partial x_i}\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \int_{\partial\Omega} f n_i\,\mathrm{d}\sigma\,,$$

where n_i is the ith component of the outward normal to $\partial\Omega$.

Proposition 9.30 is the most straightforward multi-dimensional generalisation of the recipe for integrating by parts on a bounded real interval, as we mentioned in the chapter's introduction. Its importance is cardinal, because the Divergence Theorems and Green's Theorem descend easily from it. Let us see the first of these consequences.

Theorem 9.31 (**Divergence Theorem of Gauss**) *Let $\Omega \subset \mathbb{R}^3$ be a G-admissible open set, $\boldsymbol{n}$ the outward normal to $\partial\Omega$. For any vector field $\boldsymbol{f} \in \big(\mathcal{C}^1(\overline{\Omega})\big)^3$,*

$$\int_\Omega \operatorname{div} \boldsymbol{f} \, dx \, dy \, dz = \int_{\partial\Omega} \boldsymbol{f} \cdot \boldsymbol{n} \, d\sigma \,. \tag{9.19}$$

Proof. Each component f_i of $\boldsymbol{f}$ fulfills the assumptions of the previous proposition, so

$$\int_\Omega \frac{\partial f_i}{\partial x_i} \, dx \, dy \, dz = \int_{\partial\Omega} f_i n_i \, d\sigma \qquad \text{for } i = 1, 2, 3\,.$$

Summing over i proves the claim. □

There is actually a version of the Divergence Theorem in every dimension $n \geq 2$. We will only show how the two-dimensional form below is a consequence of the three-dimensional one; for this we shall use a trick.

Theorem 9.32 *Let the open set $\Omega \subset \mathbb{R}^2$ be G-admissible and the normal $\boldsymbol{n}$ to $\partial\Omega$ point outwards. Then for any vector field $\boldsymbol{f} \in \big(\mathcal{C}^1(\overline{\Omega})\big)^2$*

$$\int_\Omega \operatorname{div} \boldsymbol{f} \, dx \, dy = \int_{\partial\Omega} \boldsymbol{f} \cdot \boldsymbol{n} \, d\gamma \,. \tag{9.20}$$

Proof. Define the open set $Q = \Omega \times (0, 1) \subset \mathbb{R}^3$, which is G-admissible (see Fig. 9.10); let $\boldsymbol{\Phi} = \boldsymbol{f} + 0\boldsymbol{k} \in \big(\mathcal{C}^1(\overline{\Omega})\big)^3$, a vector field constant with respect to z for which $\operatorname{div} \boldsymbol{\Phi} = \operatorname{div} \boldsymbol{f}$. Then

$$\int_\Omega \operatorname{div} \boldsymbol{f} \, dx \, dy = \int_0^1 \int_\Omega \operatorname{div} \boldsymbol{f} \, dx \, dy \, dz = \int_Q \operatorname{div} \boldsymbol{\Phi} \, dx \, dy \, dz \,.$$

On the other hand, if $\boldsymbol{N}$ is the outward normal to ∂Q, it is easily seen that $\boldsymbol{f} \cdot \boldsymbol{n} = \boldsymbol{\Phi} \cdot \boldsymbol{N}$ on $\partial\Omega \times (0, 1)$, whereas $\boldsymbol{\Phi} \cdot \boldsymbol{N} = 0$ on $\Omega \times \{0\}$ and $\Omega \times \{1\}$.

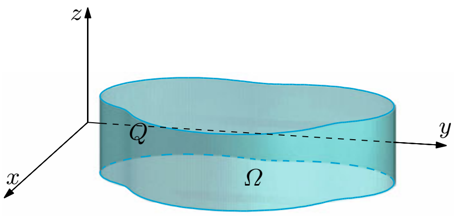


Figure 9.10. From the two-dimensional Ω to the three-dimensional Q

Therefore

$$\int_{\partial\Omega} \boldsymbol{f}\cdot\boldsymbol{n}\,\mathrm{d}\gamma = \int_0^1\int_{\partial\Omega} \boldsymbol{f}\cdot\boldsymbol{n}\,\mathrm{d}\gamma = \int_{\partial\Omega\times(0,1)} \boldsymbol{\Phi}\cdot\boldsymbol{N}\,\mathrm{d}\sigma = \int_{\partial Q} \boldsymbol{\Phi}\cdot\boldsymbol{N}\,\mathrm{d}\sigma\,.$$

The result now follows by applying the Divergence Theorem to the field $\boldsymbol{\Phi}$ on Q. □

Example 9.33

Let us compute the flux of $\boldsymbol{f}(x,y,z) = 2\boldsymbol{i} - 5\boldsymbol{j} + 3\boldsymbol{k}$ through the lateral surface Σ of the solid Ω defined by

$$x^2 + y^2 < 9 - z\,, \qquad 0 < z < 8\,.$$

First, $\partial\Omega = \Sigma \cup B_0 \cup B_1$, where B_0 is the circle of centre the origin and radius 3 on the plane $z = 0$, while B_1 is the circle with radius 1 and centre in the origin of the plane $z = 8$. Since $\operatorname{div}\boldsymbol{f} = 0$, the Divergence Theorem implies

$$0 = \int_\Omega \operatorname{div}\boldsymbol{f}\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \int_{\partial\Omega}\boldsymbol{f}\cdot\boldsymbol{n} = \int_\Sigma \boldsymbol{f}\cdot\boldsymbol{n} + \int_{B_0}\boldsymbol{f}\cdot\boldsymbol{n} + \int_{B_1}\boldsymbol{f}\cdot\boldsymbol{n}\,.$$

But as

$$\int_{B_0}\boldsymbol{f}\cdot\boldsymbol{n} = \int_{B_0}(-3) = -27\pi \qquad \text{and} \qquad \int_{B_1}\boldsymbol{f}\cdot\boldsymbol{n} = \int_{B_1} 3 = 3\pi\,,$$

we conclude that

$$\int_\Sigma \boldsymbol{f}\cdot\boldsymbol{n} = 24\pi\,.$$

□

9.5.3 Green's Theorem

Other important facts ensue from Proposition 9.30. Let $\Omega \subset \mathbb{R}^3$ be a G-admissible open set and $\boldsymbol{f}$ a $\mathcal{C}^1$ vector field on $\overline{\Omega}$. As we know, the first component of the curl of $\boldsymbol{f}$ is

$$(\mathbf{curl}\,\boldsymbol{f})_1 = \frac{\partial f_3}{\partial y} - \frac{\partial f_2}{\partial z}\,.$$

Therefore integrating over Ω and using the proposition repeatedly, we obtain

$$\int_\Omega (\mathbf{curl}\,\boldsymbol{f})_1\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \int_\Omega \Big(\frac{\partial f_3}{\partial y} - \frac{\partial f_2}{\partial z}\Big)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \int_{\partial\Omega}(\boldsymbol{n}\wedge\boldsymbol{f})_1\,\mathrm{d}\sigma\,.$$

Identities of this kind hold for the other components of the curl of $\boldsymbol{f}$. Altogether we then have the following result, that we might call *Curl Theorem*.

Theorem 9.34 *Let $\Omega \subset \mathbb{R}^3$ be open, G-admissible, and $\boldsymbol{n}$ the outward normal to $\partial\Omega$. Then, for any vector field $\boldsymbol{f} \in \big(\mathcal{C}^1(\overline{\Omega})\big)^3$, we have*

$$\int_\Omega \mathbf{curl}\,\boldsymbol{f}\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \int_{\partial\Omega}\boldsymbol{n}\wedge\boldsymbol{f}\,\mathrm{d}\sigma\,. \tag{9.21}$$

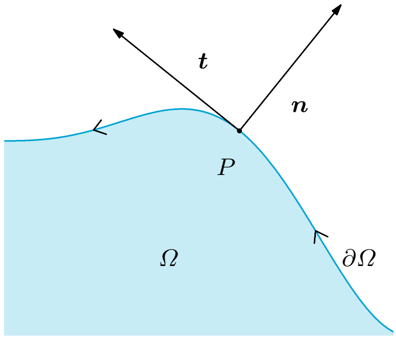


Figure 9.11. Outward normal vector and positive orientation of $\partial\Omega$

If we write the curl as $\nabla \wedge \boldsymbol{f}$, the formula can be remembered as follows: the integral over Ω becomes an integral on $\partial\Omega$ and the vector $\boldsymbol{n}$ takes the place of ∇.

As we show in the online material, Green's Theorem should be considered a two-dimensional version of the above, and can be easily deduced with the trick used for Theorem 9.32.

One last remark is in order in the run up to Green's Theorem. Let $\Omega \subset \mathbb{R}^2$ be open and G-admissible. If $\boldsymbol{n} = n_1\boldsymbol{i} + n_2\boldsymbol{j}$ is the outward normal to $\partial\Omega$, the unit vector $\boldsymbol{t} = -n_2\boldsymbol{i} + n_1\boldsymbol{j}$ is tangent to $\partial\Omega$ and oriented along the positive direction of the boundary: a three-dimensional observer standing as $\boldsymbol{k}$ and walking along $\partial\Omega$ will see Ω constantly on his left (Fig. 9.11). Now recall that given a field $\boldsymbol{f} = f_1\boldsymbol{i} + f_2\boldsymbol{j} \in \big(\mathcal{C}^1(\overline{\Omega})\big)^2$, we defined in (6.5) the function $\operatorname{curl} \boldsymbol{f} = \dfrac{\partial f_2}{\partial x} - \dfrac{\partial f_1}{\partial y}$.

Finally everything is in place for the statement, whose proof may be found in Appendix A.2.2, p. 525.

Theorem 9.35 (**Green**) *Let $\Omega \subset \mathbb{R}^2$ be a G-admissible open set whose boundary $\partial\Omega$ is positively oriented. Take a vector field $\boldsymbol{f} = f_1\boldsymbol{i} + f_2\boldsymbol{j}$ in $\big(\mathcal{C}^1(\overline{\Omega})\big)^2$. Then*

$$\int_\Omega \left(\frac{\partial f_2}{\partial x} - \frac{\partial f_1}{\partial y}\right) \mathrm{d}x\,\mathrm{d}y = \oint_{\partial\Omega} \boldsymbol{f}\cdot\boldsymbol{\tau}\,. \tag{9.22}$$

The theorem can be successfully employed to reduce the computation of the area of a domain in the plane to a path integral. Fix $\Omega \subset \mathbb{R}^2$ open and G-admissible as in the theorem. Then

$$\boxed{\operatorname{area}(\Omega) = \oint_{\partial\Omega} \frac{1}{2}(-y\boldsymbol{i} + x\boldsymbol{j})\cdot\boldsymbol{\tau}\,.} \tag{9.23}$$

In fact, $\boldsymbol{f}(x,y) = -y\boldsymbol{i} + x\boldsymbol{j}$ has constant $\mathbf{curl}\,\boldsymbol{f} = 2$, so Green's Theorem gives

$$\oint_{\partial\Omega} \boldsymbol{f}\cdot\boldsymbol{\tau} = 2\int_{\Omega} \mathrm{d}x\,\mathrm{d}y\,,$$

whence (9.23).

But notice that any vector field with constant curl on Ω may be used to obtain other expressions for the area of Ω, like

$$\mathrm{area}(\Omega) = \oint_{\partial\Omega} (-y)\boldsymbol{i}\cdot\boldsymbol{\tau} = \oint_{\partial\Omega} x\boldsymbol{j}\cdot\boldsymbol{\tau}\,.$$

Example 9.36

Let us determine the area of the elliptical region $E = \{(x,y)\in\mathbb{R}^2 : \frac{x^2}{a^2}+\frac{y^2}{b^2}\le 1\}$. We may parametrise the boundary by $\boldsymbol{\gamma}(t) = a\cos t\,\boldsymbol{i} + b\sin t\,\boldsymbol{j}$, $t\in[0,2\pi]$. Then

$$\mathrm{area}(E) = \frac{1}{2}\int_0^{2\pi} (ab\sin^2 t + ab\cos^2 t)\,\mathrm{d}t = \frac{1}{2}ab\int_0^{2\pi}\mathrm{d}t = \pi ab\,. \qquad \square$$

9.5.4 Stokes' Theorem

We discuss Stokes' Theorem for a rather large class of surfaces, that is S-admissible compact surfaces, introduced with Definition 9.28. Eschewing the formal definition, the reader may think of an S-admissible compact surface as the union of finitely many regular local graphs forming an orientable and simple compact surface. With a given crossing direction fixed, we shall say the boundary is oriented positively if an observer standing as the normal and advancing along the boundary, views the surface on the left.

First of all we re-phrase Green's Theorem in an equivalent way, the advantage being to understand it now as a special case of Stokes' Theorem. We can identify the closure of Ω, in $\mathbb{R}^2$, with the compact surface $\Sigma = \overline{\Omega}\times\{0\}$ in $\mathbb{R}^3$ (see Fig. 9.12); the latter admits the trivial parametrisation $\boldsymbol{\sigma}:\Omega\to\Sigma$, $\boldsymbol{\sigma}(u,v) = (u,v,0)$, and is obviously regular, simple and orientable, hence S-admissible. The boundary is $\partial\Sigma = \partial\Omega\times\{0\}$. Fix as crossing direction of Σ the one given by the z-axis: by calling $\boldsymbol{n}$ the unit normal, we have $\boldsymbol{n} = \boldsymbol{k}$. Furthermore, the positive orientation on $\partial\Omega$ coincides patently with the positive orientation of $\partial\Sigma$. The last piece of notation is the vector field $\boldsymbol{\Phi} = \boldsymbol{f} + 0\boldsymbol{k}$ (constant in z), for which $\mathrm{curl}\,\boldsymbol{f} = (\mathbf{curl}\,\boldsymbol{\Phi})_3 = (\mathbf{curl}\,\boldsymbol{\Phi})\cdot\boldsymbol{n}$; Equation (9.22) then becomes

$$\int_{\Sigma} (\mathbf{curl}\,\boldsymbol{\Phi})\cdot\boldsymbol{n} = \oint_{\partial\Sigma} \boldsymbol{\Phi}\cdot\boldsymbol{\tau}\,,$$

which – as we shall see – is precisely what Stokes' Theorem claims.

We are then ready to state Stokes' Theorem in full generality; the proof is available in Appendix A.2.2, p. 526, in the case the faces are sufficiently regular local graphs.

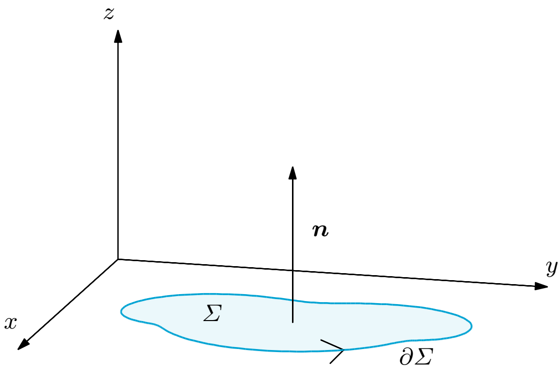


Figure 9.12. Green's Theorem as a special case of Stokes' Theorem

Theorem 9.37 (**Stokes**) *Let $\Sigma \subset \mathbb{R}^3$ be an S-admissible compact surface oriented by the unit normal $\boldsymbol{n}$; correspondingly, let the boundary $\partial\Sigma$ be oriented positively. Suppose the vector field $\boldsymbol{f}$, defined on an open set $A \subseteq \mathbb{R}^3$ containing Σ, is such that $\boldsymbol{f} \in \big(\mathcal{C}^1(A)\big)^3$. Then*

$$\int_\Sigma (\mathbf{curl}\, \boldsymbol{f}) \cdot \boldsymbol{n} = \oint_{\partial\Sigma} \boldsymbol{f} \cdot \boldsymbol{\tau}\,. \tag{9.24}$$

In other words, the flux of the curl of $\boldsymbol{f}$ across the surface equals the path integral of $\boldsymbol{f}$ along the surface's (closed) boundary.

Example 9.38

We use Stokes' Theorem to tackle Example 9.33 in an alternative way. The idea is to write $\boldsymbol{f} = 2\boldsymbol{i} - 5\boldsymbol{j} + 3\boldsymbol{k}$ as $\boldsymbol{f} = \mathbf{curl}\boldsymbol{\Phi}$. Since the components of $\boldsymbol{f}$ are constant, it is natural to look for a $\boldsymbol{\Phi}$ of the form

$$\boldsymbol{\Phi}(x,y,z) = (\alpha_1 x + \alpha_2 y + \alpha_3 z)\boldsymbol{i} + (\beta_1 x + \beta_2 y + \beta_3 z)\boldsymbol{j} + (\gamma_1 x + \gamma_2 y + \gamma_3 z)\boldsymbol{k}\,,$$

so

$$\mathbf{curl}\boldsymbol{\Phi} = (\gamma_2 - \beta_3)\boldsymbol{i} + (\alpha_3 - \gamma_1)\boldsymbol{j} + (\beta_1 - \alpha_2)\boldsymbol{k}\,.$$

Then $\gamma_2 - \beta_3 = 2$, $\alpha_3 - \gamma_1 = -5$, $\beta_1 - \alpha_2 = 3$. A solution is then $\boldsymbol{\Phi}(x,y,z) = -5z\boldsymbol{i} + 3x\boldsymbol{j} + 2y\boldsymbol{k}$. (Notice that the existence of a field $\boldsymbol{\Phi}$ such that $\boldsymbol{f} = \mathbf{curl}\boldsymbol{\Phi}$ is warranted by Sect. 6.3.1, for $\operatorname{div} \boldsymbol{f} = 0$ and Ω is convex.) By Stokes' Theorem,

$$\int_\Sigma \boldsymbol{f} \cdot \boldsymbol{n} = \int_\Sigma \mathbf{curl}\boldsymbol{\Phi} \cdot \boldsymbol{n} = \int_{\partial\Sigma} \boldsymbol{\Phi} \cdot \boldsymbol{\tau}\,.$$

We know $\partial\Sigma = \partial B_0 \cup \partial B_1$, see Example 9.33, where the circle ∂B_0 is oriented clockwise, while ∂B_1 counter-clockwise.

Then

$$\begin{aligned}\int_{B_0} \boldsymbol{\Phi} \cdot \boldsymbol{\tau} &= \int_0^{2\pi} (0\boldsymbol{i} + 9\cos t\,\boldsymbol{j} + 6\sin t\,\boldsymbol{k}) \cdot (3\cos t\,\boldsymbol{i} + 3\sin t\,\boldsymbol{j} + 0\boldsymbol{k})\,\mathrm{d}t \\ &= 27\int_0^{2\pi} \cos^2 t\,\mathrm{d}t = 27\pi\,, \\ \int_{B_1} \boldsymbol{\Phi} \cdot \boldsymbol{\tau} &= -\int_0^{2\pi} (-40\boldsymbol{i} + 3\cos t\,\boldsymbol{j} + 2\sin t\,\boldsymbol{k}) \cdot (\cos t\,\boldsymbol{i} + \sin t\,\boldsymbol{j} + 0\boldsymbol{k})\,\mathrm{d}t \\ &= 40\int_0^{2\pi} \cos t\,\mathrm{d}t - 3\int_0^{2\pi} \cos^2 t\,\mathrm{d}t = -3\pi\,,\end{aligned}$$

so eventually

$$\int_{\partial\Sigma} \boldsymbol{f} \cdot \boldsymbol{n} = 24\pi\,. \qquad \square$$

9.6 Conservative fields and potentials

In Sect. 6.3.1, Definition 6.10, we introduced the notion of a *conservative field* on an open set Ω of $\mathbb{R}^n$ as a field $\boldsymbol{f}$ that is the gradient of a map φ, called the *potential* of $\boldsymbol{f}$

$$\boxed{\boldsymbol{f} = \mathbf{grad}\,\varphi\,, \qquad \text{on } \Omega\,.}$$

Path integrals of conservative fields enjoy very special properties. The first one we encounter is in a certain sense the generalisation to curves of the Fundamental Theorem of Integral Calculus (see in particular Vol. I, Cor. 9.39).

Proposition 9.39 *If $\boldsymbol{f} = \mathbf{grad}\,\varphi$ is a conservative and continuous field on $\Omega \subseteq \mathbb{R}^n$, then*

$$\int_{\boldsymbol{\gamma}} \boldsymbol{f} \cdot \boldsymbol{\tau} = \varphi\big(\boldsymbol{\gamma}(b)\big) - \varphi\big(\boldsymbol{\gamma}(a)\big)$$

for any (piecewise-)regular arc $\boldsymbol{\gamma} : [a, b] \to \Omega$.

Proof. It suffices to consider a regular curve. Recalling formula (9.8) for path integrals, and the chain rule (esp. (6.13)), we have

$$(\mathbf{grad}\,\varphi)\big(\boldsymbol{\gamma}(t)\big) \cdot \boldsymbol{\gamma}'(t) = \frac{\mathrm{d}}{\mathrm{d}t}\varphi\big(\boldsymbol{\gamma}(t)\big)\,,$$

so

$$\int_{\boldsymbol{\gamma}} \boldsymbol{f} \cdot \boldsymbol{\tau} = \int_a^b \frac{\mathrm{d}}{\mathrm{d}t}\varphi\big(\boldsymbol{\gamma}(t)\big)\,\mathrm{d}t = \Big[\varphi\big(\boldsymbol{\gamma}(t)\big)\Big]_a^b = \varphi\big(\boldsymbol{\gamma}(b)\big) - \varphi\big(\boldsymbol{\gamma}(a)\big)\,. \qquad \square$$

Corollary 9.40 *Under the hypotheses of the previous proposition, let $\Gamma \subset \mathbb{R}^n$ be a (piecewise-)regular, simple arc oriented by the tangent vector $\boldsymbol{\tau}$. Then*

$$\int_\Gamma \boldsymbol{f} \cdot \boldsymbol{\tau} = \varphi(P_1) - \varphi(P_0)\,,$$

where P_0 and P_1 are the initial and end points of Γ.

An elementary but remarkable use of this corollary is that the potential of a conservative field is defined up to a constant on each connected component of Ω. Hence potentials behave somehow similarly to primitives in $\mathbb{R}$, namely:

Proposition 9.41 *Two scalar fields φ, ψ are potentials of the same continuous vector field $\boldsymbol{f}$ on Ω if and only if on every connected component Ω_i of Ω there is a constant c_i such that $\varphi - \psi = c_i$.*

Proof. It is clear that if φ and ψ differ by a constant on Ω_i, then $\nabla\varphi = \nabla\psi$. For the converse, fix P_0 arbitrarily in Ω_i; given any $P \in \Omega_i$, let Γ be a polygonal path starting at P_0 and ending at P (such will exist because Ω_i is connected, see Definition 4.13). Then the corollary guarantees

$$\int_\Gamma \boldsymbol{f} \cdot \boldsymbol{\tau} = \varphi(P) - \varphi(P_0) = \psi(P) - \psi(P_0)\,,$$

from which $\varphi(P) - \psi(P) = \varphi(P_0) - \psi(P_0) = c_i$. □

Proposition 9.39 and Corollary 9.40 tell us the path integral of a conservative field depends only on the end points and not on the path itself. Equivalently, arcs joining the same two points give rise to equal path integrals. In particular, the integral along a closed arc is zero. That each of these two facts characterise conservative fields is of primary importance. To establish this, we need some notation, also useful for later. If $\boldsymbol{\gamma} : [a, b] \to \mathbb{R}^n$ is an arc between $P_0 = \boldsymbol{\gamma}(a)$ and $P_1 = \boldsymbol{\gamma}(b)$, we shall write $\boldsymbol{\gamma}[P_0, P_1]$ to mean that $\boldsymbol{\gamma}$ goes from P_0 to P_1. The opposite arc $-\boldsymbol{\gamma}$ (see Definition 6.25) joins P_1 to P_0, i.e., $-\boldsymbol{\gamma}[P_1, P_0] = \boldsymbol{\gamma}[P_0, P_1]$. Given $\boldsymbol{\gamma}_1 = \boldsymbol{\gamma}_1[P_0, P_1]$ and $\boldsymbol{\gamma}_2 = \boldsymbol{\gamma}_2[P_1, P_2]$, by $\boldsymbol{\gamma} \sim \boldsymbol{\gamma}_1 + \boldsymbol{\gamma}_2$ we will denote any arc $\boldsymbol{\gamma}$ with the following property: if $\boldsymbol{\gamma} : [a, b] \to \mathbb{R}^n$, there is a $c \in (a, b)$ with $\boldsymbol{\gamma}_{|[a,c]} \sim \boldsymbol{\gamma}_1$ and $\boldsymbol{\gamma}_{|[c,b]} \sim \boldsymbol{\gamma}_2$. (An example can be easily found using increasing linear maps from $[a, c]$, $[c, b]$ to the domains of $\boldsymbol{\gamma}_1$, $\boldsymbol{\gamma}_2$ respectively.) Observe $\boldsymbol{\gamma}(c) = P_1$ and $\boldsymbol{\gamma} = \boldsymbol{\gamma}[P_0, P_2]$, so $\boldsymbol{\gamma}$ connects P_0 to P_2 passing through P_1; moreover, the traces Γ,Γ_1 and Γ_2 satisfy $\Gamma = \Gamma_1 \cup \Gamma_2$. The symbol $\boldsymbol{\gamma} \sim \boldsymbol{\gamma}_1 - \boldsymbol{\gamma}_2$ will stand for $\boldsymbol{\gamma} \sim \boldsymbol{\gamma}_1 + (-\boldsymbol{\gamma}_2)$ whenever $\boldsymbol{\gamma}_1 = \boldsymbol{\gamma}_1[P_0, P_1]$ and $\boldsymbol{\gamma}_2 = \boldsymbol{\gamma}_2[P_2, P_1]$. By the additivity of curvilinear integrals, and recalling how they depend upon congruent parametrisations (9.5), we have

$$\int_{\boldsymbol{\gamma}} \boldsymbol{f} \cdot \boldsymbol{\tau} = \int_{\boldsymbol{\gamma}_1} \boldsymbol{f} \cdot \boldsymbol{\tau} \pm \int_{\boldsymbol{\gamma}_2} \boldsymbol{f} \cdot \boldsymbol{\tau} \qquad \text{if} \quad \boldsymbol{\gamma} \sim \boldsymbol{\gamma}_1 \pm \boldsymbol{\gamma}_2\,. \tag{9.25}$$

Now we are in a position to state the result.

Theorem 9.42 *If $\boldsymbol{f}$ is a continuous vector field on the open set $\Omega \subseteq \mathbb{R}^n$, the following are equivalent:*

i) *$\boldsymbol{f}$ is conservative;*

ii) *for any (piecewise-)regular arcs $\boldsymbol{\gamma}_1$, $\boldsymbol{\gamma}_2$ with trace in Ω and common end points,*

$$\int_{\boldsymbol{\gamma}_1} \boldsymbol{f} \cdot \boldsymbol{\tau} = \int_{\boldsymbol{\gamma}_2} \boldsymbol{f} \cdot \boldsymbol{\tau}\,;$$

iii) *for any (piecewise-)regular, closed arc $\boldsymbol{\gamma}$ with trace in Ω,*

$$\oint_{\boldsymbol{\gamma}} \boldsymbol{f} \cdot \boldsymbol{\tau} = 0\,.$$

Proof. The implication *i)* $\Rightarrow$ *ii)* follows easily from Proposition 9.39 on $\boldsymbol{\gamma}_1$, $\boldsymbol{\gamma}_2$ (Fig. 9.13, left). For the converse, we manufacture an explicit potential for $\boldsymbol{f}$. Let Ω_i denote a connected component of Ω and $P_0 \in \Omega_i$ a given point. For any $P \in \Omega_i$ of coordinates $\boldsymbol{x} = (x_1, \ldots, x_n)$, set

$$\varphi(\boldsymbol{x}) = \int_{\boldsymbol{\gamma}} \boldsymbol{f} \cdot \boldsymbol{\tau}$$

where $\boldsymbol{\gamma} = \boldsymbol{\gamma}[P_0, P]$ is an arbitrary (piecewise-)regular arc with trace inside Ω_i, joining P_0 and P. The definition of $\varphi(\boldsymbol{x})$ does not depend on the choice of $\boldsymbol{\gamma}$, by *ii)*. We claim that $\mathbf{grad}\,\varphi = \boldsymbol{f}$, and will prove it only for the first component

$$\frac{\partial \varphi}{\partial x_1}(\boldsymbol{x}) = f_1(\boldsymbol{x})\,.$$

Let $\Delta x_1 \neq 0$ be an increment such that $P + \Delta P = \boldsymbol{x} + \Delta x_1 \boldsymbol{e}_1$ still belongs to Ω_i, and call $\boldsymbol{\gamma}[P, P+\Delta P]$ the curve $\boldsymbol{\gamma}(t) = \boldsymbol{x} + t\boldsymbol{e}_1$ from P to

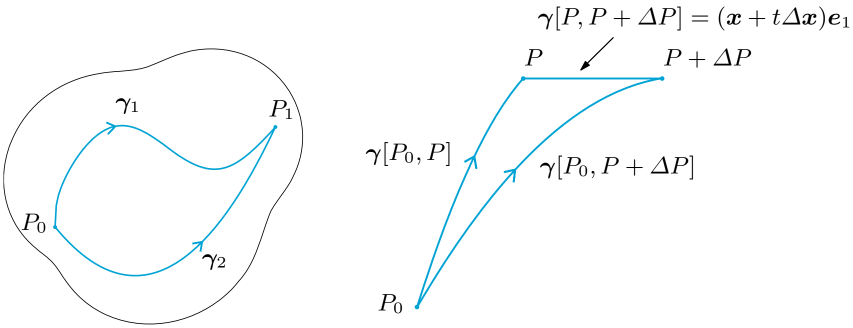


Figure 9.13. Proof of Theorem 9.42

$P + \Delta P$, with $t \in [0, \Delta x_1]$ if $\Delta x_1 > 0$ and $t \in [\Delta x_1, 0]$ if $\Delta x_1 < 0$. Let $\boldsymbol{\gamma}[P_0, P]$ and $\boldsymbol{\gamma}[P_0, P+\Delta P]$ be regular, simple curves from P_0 to P and $P + \Delta P$ respectively (Fig. 9.13, right). Then $\boldsymbol{\gamma}[P_0, P+\Delta P] \sim \boldsymbol{\gamma}[P_0, P] + \boldsymbol{\gamma}[P, P+\Delta P]$, and by (9.25) and (9.8), we have

$$\frac{\varphi(\boldsymbol{x} + \Delta x_1 \boldsymbol{e}_1) - \varphi(\boldsymbol{x})}{\Delta x_1} = \frac{1}{\Delta x_1}\left(\int_{\boldsymbol{\gamma}[P_0, P+\Delta P]} \boldsymbol{f}\cdot\boldsymbol{\tau} - \int_{\boldsymbol{\gamma}[P_0,P]} \boldsymbol{f}\cdot\boldsymbol{\tau}\right)$$

$$= \frac{1}{\Delta x_1}\int_{\boldsymbol{\gamma}[P, P+\Delta P]} \boldsymbol{f}\cdot\boldsymbol{\tau} = \frac{1}{\Delta x_1}\int_0^{\Delta x_1} f_1(x_1 + t, x_2, \ldots, x_n)\, dt\,.$$

The last term is the integral average of the map $t \mapsto f_1(x_1 + t, x_2, \ldots, x_n)$ on the interval between 0 and Δx_1, so the continuity of $\boldsymbol{f}$ and the Mean Value Theorem (Vol. I, Thm. 9.35) force

$$\frac{\varphi(\boldsymbol{x} + \Delta x_1 \boldsymbol{e}_1) - \varphi(\boldsymbol{x})}{\Delta x_1} = f_1(x_1 + \bar{t}, x_2, \ldots, x_n)$$

for a certain $\bar{t}$ with $|\bar{t}| \le |\Delta x_1|$. The limit for $\Delta x_1 \to 0$ proves the claim.

The equivalence of *ii)* and *iii)* is an immediate consequence of (9.25): if $\boldsymbol{\gamma}_1$, $\boldsymbol{\gamma}$ fulfill *ii)* then each $\boldsymbol{\gamma} \sim \boldsymbol{\gamma}_1 - \boldsymbol{\gamma}_2$ satisfies *iii)*, while a $\boldsymbol{\gamma}$ satisfying *iii)*, as seen above, is the difference of $\boldsymbol{\gamma}_1$, $\boldsymbol{\gamma}_2$ with common end points. □

The question remains of how to characterise conservative fields. A necessary condition is the following.

Property 9.43 *Let $\boldsymbol{f}$ be a $\mathcal{C}^1$ vector field on $\Omega \subseteq \mathbb{R}^n$. If it is conservative, we have*

$$\frac{\partial f_i}{\partial x_j} = \frac{\partial f_j}{\partial x_i} \qquad \forall i \neq j\,. \tag{9.26}$$

Proof. Take a potential φ for $\boldsymbol{f}$, so $f_i = \frac{\partial \varphi}{\partial x_i}$ for $i = 1, \ldots, n$; in particular $\varphi \in \mathcal{C}^2(\Omega)$. Then formula (9.26) is nothing but Schwarz's Theorem 5.17. □

We met this property in Sect. 6.3.1 (Proposition 6.8 for dimension two and Proposition 6.7 for dimension three); in fact, it was proven there that a conservative $\mathcal{C}^1$ vector field is necessarily irrotational, $\mathbf{curl}\,\boldsymbol{f} = \mathbf{0}$, on Ω.

At this juncture one would like to know if, and under which conditions, being curl-free is also sufficient to be conservative. From now on we shall suppose Ω is open in $\mathbb{R}^2$ or $\mathbb{R}^3$. By default we will think in three dimensions, and highlight a few peculiarities of the two-dimensional situation. Referring to the equivalent formulation appearing in Theorem 9.42 iii), note that if $\boldsymbol{\gamma}$ is any regular, simple arc whose trace Γ is the boundary of a regular and simple compact surface Σ contained in Ω, Stokes' Theorem 9.37 makes sure that $\boldsymbol{f}$ irrotational implies

$$\oint_\gamma \boldsymbol{f} \cdot \boldsymbol{\tau} = \int_\Sigma \mathbf{curl}\, \boldsymbol{f} \cdot \boldsymbol{n} = \mathbf{0}\,.$$

Nevertheless, not all regular and simple closed arcs Γ in Ω are boundaries of a regular and simple compact surface Σ in Ω: the shape of Ω might prevent this from happening. If, for example, Ω is the complement in $\mathbb{R}^3$ of the axis z, it is self-evident that any compact surface having a closed boundary encircling the axis must also intersect the axis, so it will not be contained entirely in Ω (see Fig. 9.14); in this circumstance Stokes' Theorem does not hold. This fact in itself does not obstruct the vanishing of the circulation of a curl-free field around z; it just says Stokes' Theorem does not apply. In spite of that, if we consider the field

$$\boldsymbol{f}(x,y,z) = -\frac{y}{x^2+y^2}\,\boldsymbol{i} + \frac{x}{x^2+y^2}\,\boldsymbol{j} + 0\,\boldsymbol{k}$$

on Ω, undoubtedly $\mathbf{curl}\, \boldsymbol{f} = \mathbf{0}$ on Ω, whereas

$$\oint_\Gamma \boldsymbol{f} \cdot \boldsymbol{\tau} = 2\pi$$

with Γ the counter-clockwise unit circle on the xy-plane centred at the origin. This is therefore an example (extremely relevant from the physical viewpoint, by the way, $\boldsymbol{f}$ being the magnetic field generated by a current along a wire on the z-axis) of an irrotational vector field that is not conservative on Ω.

The discussion suggests to narrow the class of open sets Ω, in such a way that

$$\boxed{\boldsymbol{f} \quad \text{curl-free} \qquad \Rightarrow \qquad \boldsymbol{f} \quad \text{conservative}}$$

holds. With this in mind, let us give a definition.

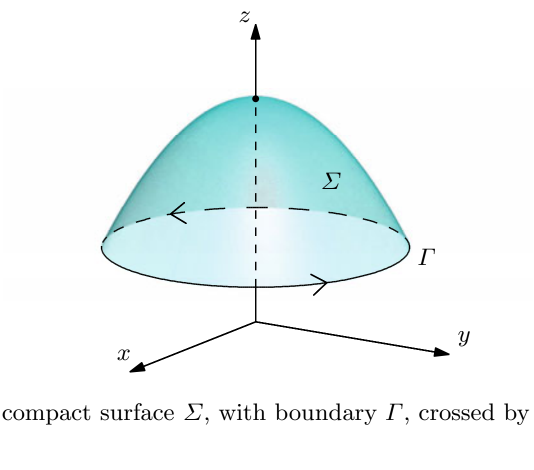


Figure 9.14. A compact surface Σ, with boundary Γ, crossed by the z-axis

Definition 9.44 *An open connected set $\Omega \subseteq \mathbb{R}^n$ is* **simply connected** *if the trace Γ of any arc in Ω can be deformed with continuity to a point, always staying within Ω.*

More precisely: Ω is simply connected if any closed curve $\boldsymbol{\gamma} : I \to \mathbb{R}^n$ belongs to a one-parameter family of closed curves $\boldsymbol{\gamma}_s : I \to \mathbb{R}^n$, $0 \le s \le 1$, with the following properties:

i) the map $(t, s) \mapsto \boldsymbol{\gamma}_s(t)$ from $I \times [0, 1]$ to $\mathbb{R}^n$ is continuous;
ii) each trace $\Gamma_s = \boldsymbol{\gamma}_s(I)$ is contained in Ω;
iii) $\boldsymbol{\gamma}_1 = \boldsymbol{\gamma}$ and $\boldsymbol{\gamma}_0$ is constant, i.e., Γ_0 is a point.

One says $\boldsymbol{\gamma}$ is **homotopic** to a point, and the map $(t, s) \mapsto \boldsymbol{\gamma}_s(t)$ is known as a **homotopy**.

Simply connectedness can be defined alternatively. For instance, in dimension 2 we may equivalently demand that

- the complement of Ω in $\mathbb{R}^2$ is a connected set,

or

- for any Jordan arc γ in Ω, the interior Σ_i is entirely contained in Ω.

Naïvely, an open connected set in the plane is simply connected if it has no 'holes'; an (open) annulus is thus not simply connected (see Fig. 9.15, left).

The situation in three dimension is more intricate. The open domain enclosed by two concentrical spheres is simply connected, whereas an open torus is not (Fig. 9.15, middle and right).

Likewise, the open set obtained by removing one point from $\mathbb{R}^3$ is simply connected, but if we take out a whole line it is not simply connected any longer. It can be proved that an open connected set Ω in $\mathbb{R}^3$ is simply connected if and only if for any (piecewise-)regular Jordan arc Γ in Ω, there exists a regular compact surface Σ in Ω that has Γ as boundary.

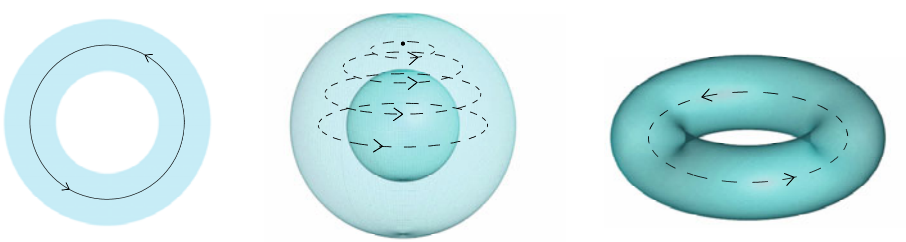


Figure 9.15. A simply connected open set (middle) and non-simply connected ones (left and right)

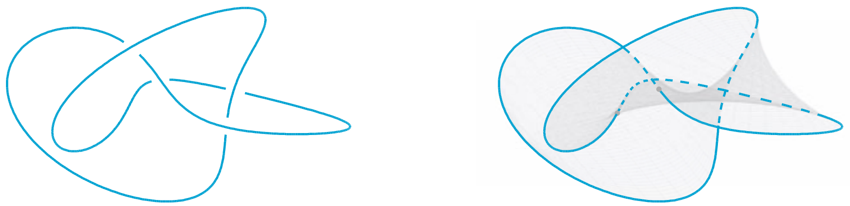


Figure 9.16. A closed simple arc (left) bordering a compact surface (right)

The above characterisation leads to the rather-surprising result for which any curve in space that closes up, even if knotted, is the boundary of a regular compact surface. This is possible because the surface is not required to be simple; as a matter of fact its trace may consist of faces intersecting transversely (as in Fig. 9.16).

Returning to the general set-up, there exist geometrical conditions that guarantee an open set $\Omega \subseteq \mathbb{R}^n$ is simply connected. For instance, **convex** open sets are simply connected, and the same is true for **star-shaped** sets; the latter admit a point $P_0 \in \Omega$ such that the segment $\overline{P_0P}$ from P_0 to an arbitrary $P \in \Omega$ is contained in Ω (see Fig. 9.17). In particular, a convex set is star-shaped with respect to any of its points.

Finally, here is the awaited characterization of conservative fields; the proof is given in Appendix A.2.2, p. 527.

Theorem 9.45 *Let $\Omega \subseteq \mathbb{R}^n$, with $n = 2$ or 3, be open and simply connected. A vector field $\boldsymbol{f}$ of class $\mathcal{C}^1$ on Ω is conservative if and only if it is curl-free.*

A similar result ensures the existence of a potential vector for a field with no divergence.

Remark 9.46 The concepts and results presented in this section may be equivalently expressed by the terminology of *differential forms*. We refer to Appendix A.2.3, p. 529, for further details. □

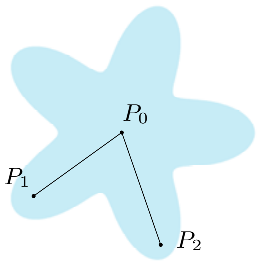


Figure 9.17. A star-shaped set for P_0

9.6.1 Computing potentials explicitly

Suppose $\boldsymbol{f}$ is conservative on an open (without loss of generality, connected) set $\Omega \subseteq \mathbb{R}^n$. We wish to find a potential for $\boldsymbol{f}$, which we already know will be defined up to a constant. Let us explain two different methods for doing this.

The first method uses the representation seen in Theorems 9.42 and 9.45. To be precise, fix a point P_0 and define the potential at every $P \in \Omega$ of coordinate $\boldsymbol{x}$ by

$$\boxed{\varphi(\boldsymbol{x}) = \int_{\Gamma[P_0,P]} \boldsymbol{f} \cdot \boldsymbol{\tau}\,,}$$

where $\Gamma[P_0, P] \subset \Omega$ is a simple and (piecewise-)regular arc from P_0 to P. The idea is to choose the path to make the integral as simple as possible to compute (recall that the integral is independent of the path, by part ii) of Theorem 9.42). In many cases, the best option is a polygonal path with segments parallel to the coordinate axes; if so, over each segment the integrand $\boldsymbol{f} \cdot \boldsymbol{\tau}$ depends only on one component of $\boldsymbol{f}$.

Example 9.47

Consider the field

$$\boldsymbol{f}(x,y) = \frac{y}{\sqrt{1+2xy}}\,\boldsymbol{i} + \frac{x}{\sqrt{1+2xy}}\,\boldsymbol{j}$$

defined on the open set Ω between the branches of the hyperbola $xy = -1/2$. It is not hard to convince oneself that Ω is star-shaped with respect to the origin, and $\mathbf{curl}\,\boldsymbol{f} = \mathbf{0}$ on Ω; hence the vector field is conservative. To find a potential, we may use as path the segment from $(0,0)$ to $P = (x,y)$ given by $\boldsymbol{\gamma}(t) = (tx, ty)$, $0 \le t \le 1$. Then

$$\varphi(x,y) = \int_0^1 \boldsymbol{f}\big(\boldsymbol{\gamma}(t)\big) \cdot \boldsymbol{\gamma}'(t)\,\mathrm{d}t = \int_0^1 \frac{2xyt}{\sqrt{1+2xyt^2}}\,\mathrm{d}t\,.$$

Substituting $u = 1 + 2xyt^2$, so $\mathrm{d}u = 4xyt\,\mathrm{d}t$, we have

$$\varphi(x,y) = \int_1^{1+2xy} \frac{1}{2\sqrt{u}}\,\mathrm{d}u = \Big[\sqrt{u}\Big]_1^{1+2xy} = \sqrt{1+2xy} - 1\,.$$

The generic potential of $\boldsymbol{f}$ on Ω will be

$$\varphi(x,y) = \sqrt{1+2xy} + c\,.$$

□

The second method we propose, often easier to use than the previous one, consists in integrating with respect to the single variables, using one after the other the relationships

$$\frac{\partial \varphi}{\partial x_1} = f_1\,, \quad \frac{\partial \varphi}{\partial x_2} = f_2\,, \quad \ldots \quad , \quad \frac{\partial \varphi}{\partial x_n} = f_n\,.$$

We exemplify the procedure in dimension 2 and give an explicit example for dimension 3.

From

$$\frac{\partial \varphi}{\partial x}(x, y) = f_1(x, y)$$

we obtain

$$\varphi(x, y) = F_1(x, y) + \psi_1(y)\,,$$

where $F_1(x, y)$ is any primitive map of $f_1(x, y)$ with respect to x, i.e., it satisfies $\frac{\partial F_1}{\partial x}(x, y) = f_1(x, y)$, while $\psi_1(y)$, for the moment unknown, is the constant of the previous integration in x, hence depends on y only. To pin down this function, we differentiate the last displayed equation with respect to y

$$\frac{\mathrm{d}\psi_1}{\mathrm{d}y}(y) = \frac{\partial \varphi}{\partial y}(x, y) - \frac{\partial F_1}{\partial y}(x, y) = f_2(x, y) - \frac{\partial F_1}{\partial y}(x, y) = g(y)\,.$$

Note $f_2(x, y) - \frac{\partial F_1}{\partial y}(x, y)$ depends only on y, because its x-derivative vanishes by (9.26), since

$$\frac{\partial f_2}{\partial x}(x, y) - \frac{\partial}{\partial y}\frac{\partial F_1}{\partial x}(x, y) = \frac{\partial f_2}{\partial x}(x, y) - \frac{\partial f_1}{\partial y}(x, y) = 0\,.$$

Calling $G(y)$ an arbitrary primitive of $g(y)$, we have

$$\psi_1(y) = G(y) + c$$

whence

$$\boxed{\varphi(x, y) = F_1(x, y) + G(y) + c\,.}$$

In higher dimension, successive integrations determine one after the other the unknown maps $\psi_1(x_2, \ldots, x_n)$, $\psi_2(x_3, \ldots, x_n), \ldots, \psi_{n-1}(x_n)$ that depend on a decreasing number of variables.

Example 9.48

Consider the vector field in $\mathbb{R}^3$

$$\boldsymbol{f}(x, y, z) = 2yz\,\boldsymbol{i} + 2z(x + 3y)\boldsymbol{j} + \big(y(2x + 3y) + 2z\big)\boldsymbol{k}\,.$$

It is straightforward to check $\mathbf{curl}\,\boldsymbol{f} = \mathbf{0}$, making the field conservative. Integrating

$$\frac{\partial \varphi}{\partial x}(x, y, z) = 2yz$$

produces $\varphi(x, y, z) = 2xyz + \psi_1(y, z)$. Its derivative in y, and $\frac{\partial \varphi}{\partial y}(x, y, z) = 2xz + 6yz$, tells that $\frac{\partial \psi_1}{\partial y}(y, z) = 6yz$, so

$$\psi_1(y, z) = 3y^2 z + \psi_2(z)\,.$$

Differentiating the latter with respect to z, and recalling $\frac{\partial\varphi}{\partial z}(x,y,z) = 2xy + 3y^2 + 2z$, gives
$$\frac{\mathrm{d}\psi_2}{\mathrm{d}z}(z) = 2z\,,$$
hence $\psi_2(z) = z^2 + c$. In conclusion, all potentials for $\boldsymbol{f}$ are of the form
$$\varphi(x,y,z) = 2xyz + 3y^2z + z^2 + c\,. \qquad \square$$

We close the section by calling the attention to a class of conservative fields for which the potential is particularly easy to find. These are the so-called **radial** vector fields,
$$\boxed{\boldsymbol{f}(\boldsymbol{x}) = g(\|\boldsymbol{x}\|)\,\boldsymbol{x}\,,}$$
where $g = g(r)$ is a real, continuous function on an interval I of $[0,+\infty)$. A field is radial when at each point P it is collinear with the vector OP, and its norm depends merely on the distance of P from the origin. A straightforward computation shows $\boldsymbol{f}$ is conservative on $\varOmega = \{\boldsymbol{x} \in \mathbb{R}^n : \|\boldsymbol{x}\| \in I\}$ and gives a potential. Precisely, define
$$\boxed{\varphi(\boldsymbol{x}) = G(\|\boldsymbol{x}\|)}$$
where $G = G(r)$ is an arbitrary primitive map of $rg(r)$ on I. In fact, the chain rule plus Example 5.3 i), give
$$\nabla\varphi(\boldsymbol{x}) = G'(\|\boldsymbol{x}\|)\,\frac{\boldsymbol{x}}{\|\boldsymbol{x}\|} = \|\boldsymbol{x}\|g(\|\boldsymbol{x}\|)\,\frac{\boldsymbol{x}}{\|\boldsymbol{x}\|} = \boldsymbol{f}(\boldsymbol{x})\,.$$
For example, $\boldsymbol{f}(\boldsymbol{x}) = \dfrac{\boldsymbol{x}}{\sqrt{1+\|\boldsymbol{x}\|^2}}$ admits $\varphi(\boldsymbol{x}) = \sqrt{1+\|\boldsymbol{x}\|^2}$ as potential.

9.7 Exercises

1. *Compute the integral of the map*
$$f(x,y,z) = \frac{x^2(1+8y)}{\sqrt{1+y+4x^2y}}$$
along the arc $\boldsymbol{\gamma}$ *defined by* $\boldsymbol{\gamma}(t) = (t, t^2, \log t)\,,\ t \in [1,2]$.

2. *Let* $\varGamma$ *be the union of the parabolic arc* $y = 4 - x^2$ *going from* $A = (-2,0)$ *to* $C = (2,0)$, *and the circle* $x^2 + y^2 = 4$ *from* C *to* A. *Integrate the function* $f(x,y) = x$ *along the closed curve* $\varGamma$.

$\boxed{3.}$ *Integrate $f(x,y) = x+y$ along the closed loop Γ, contained in the first quadrant, that is union of the segment between $O = (0,0)$ and $A = (1,0)$, the elliptical arc $4x^2+y^2 = 4$ from A to $B = (\frac{\sqrt{2}}{2}, \sqrt{2})$ and the segment joining B to the origin.*

4. *Compute the integral of $f(x,y) = \dfrac{1}{x^2+y^2+1}$ along the simple closed arc $\boldsymbol{\gamma}$ whose trace is made of the segment from the origin to $A = (\sqrt{2}, 0)$, the circular arc $x^2+y^2 = 2$ from A to $B = (1,1)$, and the segment from B back to the origin.*

$\boxed{5.}$ *Let $\boldsymbol{\gamma}_1 : [0, \frac{2}{3}\pi] \to \mathbb{R}^2$ be given by $\boldsymbol{\gamma}_1(t) = (t\cos t, t\sin t)$, and $\boldsymbol{\gamma}_2$, $\boldsymbol{\gamma}_3$ parametrise the segments from $B = (-\frac{\pi}{3}, \frac{\pi}{\sqrt{3}})$ to $C = (-\pi, 0)$ and from C to $A = (0,0)$. Compute the length of $\boldsymbol{\gamma} \sim \boldsymbol{\gamma}_1 + \boldsymbol{\gamma}_2 + \boldsymbol{\gamma}_3$ whose trace is the union Γ of the traces of $\boldsymbol{\gamma}_1$, $\boldsymbol{\gamma}_2$, $\boldsymbol{\gamma}_3$, and then find the area of the domain inside Γ.*

6. *Find the centre of mass of the arc Γ, parametrised by $\boldsymbol{\gamma}(t) = \mathrm{e}^t \cos t\, \boldsymbol{i} - \mathrm{e}^t \sin t\, \boldsymbol{j}$, $t \in [0, \pi/2]$, with unit density.*

$\boxed{7.}$ *Consider the arc $\boldsymbol{\gamma}$ whose trace Γ is the union of the segment from $A = (2\sqrt{3}, -2)$ to the origin, and the parabolic arc $y^2 = 2x$ from $(0,0)$ to $B = (\frac{1}{2}k^2, k)$. Knowing it has unit density, determine k so that the centre of gravity of Γ belongs to the x-axis.*

$\boxed{8.}$ *Determine the moment about z of the arc Γ parametrised by $\boldsymbol{\gamma}(t) = t\cos t\, \boldsymbol{i} + t\sin t\, \boldsymbol{j} + t\, \boldsymbol{k}$, $t \in [0, \sqrt{2}]$ and having unit density.*

$\boxed{9.}$ *Integrate the field $\boldsymbol{f}(x,y) = (x^2, xy)$ along $\boldsymbol{\gamma}(t) = (t^2, t)$, $t \in [0,1]$.*

10. *What is the path integral of $\boldsymbol{f}(x,y,z) = (z, y, 2x)$ along the arc $\boldsymbol{\gamma}(t) = (t, t^2, t^3)$, $t \in [0,1]$?*

11. *Integrate $\boldsymbol{f}(x,y,z) = (2\sqrt{z}, x, y)$ along $\boldsymbol{\gamma}(t) = (-\sin t, \cos t, t^2)$, $t \in [0, \frac{\pi}{2}]$.*

$\boxed{12.}$ *Compute the integral of $\boldsymbol{f}(x,y) = (xy^2, x^2y)$ along Γ, the polygonal path joining $A = (0,1)$, $B = (1,1)$, $C = (0,2)$ and $D = (1,2)$.*

13. *Integrate $\boldsymbol{f}(x,y) = (0, y)$ along the following simple closed arc: a segment from the origin to $A = (1,0)$, a circle $x^2+y^2 = 1$ from A to $B = (\frac{\sqrt{2}}{2}, \frac{\sqrt{2}}{2})$, a segment from B to the origin.*

$\boxed{14.}$ *Determine k so that the work done by the force*

$$\boldsymbol{f}(x,y) = (x^2 - xy)\boldsymbol{i} + (y^2 - x^2)\boldsymbol{j}$$

along the parabola $y^2 = 2kx$ from the origin to $P = (k/2, k)$ equals $9/5$.

15. *Compute the work integral of*

$$\boldsymbol{f}(x,y) = (ax^2y - \sin x)\boldsymbol{i} + (x^3 + y\log y)\boldsymbol{j}$$

along Γ, union of the arcs Γ_1, Γ_2, Γ_3 of respective equations $y = x^2$, $y = 1$, $x = 0$. Determine the parameter a so that the work is zero.

$\boxed{16.}$ *Let $f(x,y,z) = z(y-2x)$. Compute the integral of f over the surface $\boldsymbol{\sigma}(u,v) = (u, v, \sqrt{16-u^2-v^2})$, defined on $\mathcal{R} = \{(u,v) \in \mathbb{R}^2 : u \le 0,\ v \ge 0,\ u^2+v^2 \le 16,\ \frac{u^2}{4} + v^2 \ge 1\}$.*

$\boxed{17.}$ *Integrate*

$$f(x,y,z) = \frac{y+1}{\sqrt{1 + \frac{x^2}{4} + 4y^2}}$$

over Σ, which is the portion of the elliptical paraboloid $z = -\frac{x^2}{4} - y^2$ above the plane $z = -1$.

$\boxed{18.}$ *Determine the area of the compact surface Σ, intersection of the surface $z = \frac{1}{2}y^2$ with the prism given by the planes $x+y=4$, $y-x=4$, $y=0$.*

19. *Compute the area of the portion of surface $z = \sqrt{x^2+y^2}$ below the plane $z = \frac{1}{\sqrt{2}}(y+2)$.*

$\boxed{20.}$ *Find the moment, about the axis z, of the surface Σ part of the cone $z^2 = 3(x^2+y^2)$ satisfying $0 \le z \le 3$, $y \ge x$ (assume unit density).*

21. *Let $\mathcal{R}$ in the zy-plane be the union of*

$$\mathcal{R}_1 = \{(y,z) \in \mathbb{R}^2 : y \le 0,\ \sqrt{1-(y+1)^2} \le z \le \sqrt{4-y^2}\}$$

and

$$\mathcal{R}_2 = \{(y,z) \in \mathbb{R}^2 : y \ge 0,\ y \le z \le \sqrt{4-y^2}\}.$$

Find the coordinate x_G of the centre of mass for the part of the surface $x = \sqrt{4-y^2-z^2}$ projecting onto $\mathcal{R}$ (assume unit density).

$\boxed{22.}$ *Using Green's Theorem, compute the work done by*

$$\boldsymbol{f}(x,y) = xy\boldsymbol{i} + x^4y\boldsymbol{j}$$

along the closed counter-clockwise loop Γ, union of Γ_1, Γ_2, Γ_3, Γ_4 of respective equations $x = y^2$, $y = 1$, $x^2+y^2 = 5$, $y = 0$, $(x,\ y \ge 0)$.

$\boxed{23.}$ *Using Green's Theorem, compute the work done by* $\boldsymbol{f}(x,y) = x^2y^2\boldsymbol{i} + ax\boldsymbol{j}$, *as the parameter* a *varies, along the closed counter-clockwise curve* Γ, *union of* Γ_1, Γ_2, Γ_3 *of equations* $x^2+y^2=1$, $x=1$, $y=x^2+1$ $(x,y \geq 0)$.

$\boxed{24.}$ *Let* Γ *be the union of* Γ_1 *parametrised by* $\boldsymbol{\gamma}_1(t) = (\mathrm{e}^{-(t-\pi/4)}\cos t, \sin t)$, $t \in [0,\pi/4]$, *and the segments from the origin to the end points of* Γ_1, $A = (\mathrm{e}^{\pi/4}, 0)$ *and* $B = (\sqrt{2}/2, \sqrt{2}/2)$. *Find the area of the region* Ω *bounded by* Γ.

$\boxed{25.}$ *With the aid of the Divergence Theorem, determine the outgoing flux of*

$$\boldsymbol{f}(x,y,z) = (x^3+yz)\boldsymbol{i} + (xz+y^3)\boldsymbol{j} + (xy+z^3+1)\boldsymbol{k}$$

across Ω, *defined by*

$$x^2+y^2+z^2 \leq 1\,, \quad x^2+y^2-z^2 \leq 0\,, \quad y \geq 0\,, \quad z \geq 0\,.$$

26. *Find the flux of*

$$\boldsymbol{f}(x,y,z) = (xy^2+z^3)\boldsymbol{i} + (x^2+\frac{1}{3}y^3)\boldsymbol{j} + 2(x^2z+\frac{1}{3}z^3+2)\boldsymbol{k}$$

leaving the compact surface

$$x^2+y^2+z^2 \leq 2\,, \quad x^2+y^2-z^2 \geq 0\,, \quad y \geq 0\,, \quad z \geq 0\,.$$

$\boxed{27.}$ *Using Stokes' Theorem, compute the integral of*

$$\boldsymbol{f}(x,y,z) = x\boldsymbol{i} + y\boldsymbol{j} + xy\boldsymbol{k}$$

along the boundary of Σ, *intersection of the cylinder* $x^2+y^2=4$ *and the paraboloid* $z = \frac{x^2}{9}+\frac{y^2}{4}$, *oriented so that the normal points toward the* z*-axis.*

$\boxed{28.}$ *Given the vector field*

$$\boldsymbol{f}(x,y,z) = (y+z)\boldsymbol{i} + 2(x+z)\boldsymbol{j} + 3(x+y)\boldsymbol{k}$$

and the sphere $x^2+y^2+z^2=2$, *find the flux of the curl of* $\boldsymbol{f}$ *going out of the portion of sphere above* $z=y$.

$\boxed{29.}$ *Verify the field* $\boldsymbol{f}(x,y,z) = x\boldsymbol{i} - 2y\boldsymbol{j} + 3z\boldsymbol{k}$ *is conservative and find a potential for it.*

$\boxed{30.}$ *Determine the map* $g \in \mathcal{C}^\infty(\mathbb{R})$ *with* $g(0)=0$ *that makes the vector field*

$$\boldsymbol{f}(x,y) = (y\sin x + xy\cos x + \mathrm{e}^y)\boldsymbol{i} + \big(g(x) + x\mathrm{e}^y\big)\boldsymbol{j}$$

conservative. Find a potential for the resulting f.

31. *Consider*

$$\boldsymbol{f}(x,y) = \big(\frac{g(y)}{x} + \cos y\big)\boldsymbol{i} + (2y\log x - x\sin y)\boldsymbol{j}\,.$$

a) Determine the map $g \in \mathcal{C}^\infty(\mathbb{R})$ with $g(1) = 1$, so that $\boldsymbol{f}$ is conservative.
b) Determine for the field thus obtained the potential φ such that $\varphi(1, \frac{\pi}{2}) = 0$.

32. *Consider the field*

$$\boldsymbol{f}(x,y) = (2x\log y - y\sin x)\boldsymbol{i} + \big(\frac{g(x)}{y} + \cos x\big)\boldsymbol{j}\,.$$

a) Determine $g \in \mathcal{C}^\infty(\mathbb{R})$ such that $g(0) = 1$ and making $\boldsymbol{f}$ conservative.
b) For this field find the potential φ such that $\varphi(\frac{\pi}{3}, 1) = 0$.

33. *Consider*

$$\boldsymbol{f}(x,y,z) = (3y + \cos(x+z^2))\boldsymbol{i} + (3x + y + g(z))\boldsymbol{j} + (y + 2z\cos(x+z^2))\boldsymbol{k}\,.$$

a) Determine the map $g \in \mathcal{C}^\infty(\mathbb{R})$ so that $g(1) = 0$ and $\boldsymbol{f}$ is conservative.
b) For the above field determine the potential φ with $\varphi(0,0,0) = 0$.

34. *Determine the value of λ such that*

$$\boldsymbol{f}(x,y,z) = (x^2 + 5\lambda y + 3yz)\boldsymbol{i} + (5x + 3\lambda xz - 2)\boldsymbol{j} + ((2+\lambda)xy - 4z)\boldsymbol{k}$$

is conservative. Then find the potential φ with $\varphi(3, 1, -2) = 0$.

9.7.1 Solutions

1. For $t \in [1,2]$ we have $f\big(\boldsymbol{\gamma}(t)\big) = \dfrac{t^2(1+8t^2)}{\sqrt{1+t^2+4t^4}}$, and $\boldsymbol{\gamma}'(t) = \big(1, 2t, \frac{1}{t}\big)$, so

$$\int_{\boldsymbol{\gamma}} f = \int_1^2 \frac{t^2(1+8t^2)}{\sqrt{1+t^2+4t^4}}\,\frac{1}{t}\,\sqrt{1+t^2+4t^4}\,\mathrm{d}t = \int_1^2 t(1+8t^2)\,\mathrm{d}t = \frac{63}{2}\,.$$

2. 0.

3. Let us compute first the coordinates of B, in the first quadrant, intersection of the line $y = 2x$ and the ellipse $4x^2 + y^2 = 4$, which are $B = (\frac{\sqrt{2}}{2}, \sqrt{2})$. The piecewise-regular arc $\boldsymbol{\gamma}$ can be divided in three regular arcs $\boldsymbol{\gamma}_1, \boldsymbol{\gamma}_2, \boldsymbol{\gamma}_3$ of respective traces the segment OA, the elliptical arc AB and the segment BO. We can define arcs $\boldsymbol{\delta}_1$, $\boldsymbol{\delta}_2$, $\boldsymbol{\delta}_3$ congruent to $\boldsymbol{\gamma}_1$, $\boldsymbol{\gamma}_2$, $\boldsymbol{\gamma}_3$ as follows:

$$\begin{aligned}
&\boldsymbol{\delta}_1(t) = (t, 0) && 0 \le t \le 1\,, && \boldsymbol{\delta}_1 = \boldsymbol{\gamma}_1\,,\\
&\boldsymbol{\delta}_2(t) = (\cos t, 2\sin t) && 0 \le t \le \frac{\pi}{4}\,, && \boldsymbol{\delta}_2 \sim \boldsymbol{\gamma}_2\,,\\
&\boldsymbol{\delta}_3(t) = (t, 2t) && 0 \le t \le \frac{\sqrt{2}}{2}\,, && \boldsymbol{\delta}_3 \sim -\boldsymbol{\gamma}_3\,,
\end{aligned}$$

so that

$$\int_{\boldsymbol{\gamma}} f = \int_{\boldsymbol{\delta}_1} f + \int_{\boldsymbol{\delta}_2} f + \int_{\boldsymbol{\delta}_3} f\,.$$

Since

$$\begin{aligned} &f\big(\boldsymbol{\delta}_1(t)\big) = t\,, && f\big(\boldsymbol{\delta}_2(t)\big) = \cos t + 2\sin t\,, && f\big(\boldsymbol{\delta}_3(t)\big) = 3t\,,\\ &\boldsymbol{\delta}_1'(t) = (1,0)\,, && \boldsymbol{\delta}_1'(t) = (-\sin t, 2\cos t)\,, && \boldsymbol{\delta}_3'(t) = (1,2)\,,\\ &\|\boldsymbol{\delta}_1'(t)\| = 1\,, && \|\boldsymbol{\delta}_2'(t)\| = \sqrt{\sin^2 t + 4\cos^2 t}\,, && \|\boldsymbol{\delta}_3'(t)\| = \sqrt{5}\,, \end{aligned}$$

it follows

$$\begin{aligned} \int_{\boldsymbol{\gamma}} f &= \int_0^1 t\,dt + \int_0^{\pi/4} \big(\cos t + 2\sin t\big)\sqrt{\sin^2 t + 4\cos^2 t}\,dt + \int_0^{\sqrt{2}/2} 3\sqrt{5}t\,dt\\ &= \frac{1}{2} + \frac{3}{4}\sqrt{5} + \int_0^{\pi/4} \cos t\sqrt{4 - 3\sin^2 t}\,dt + 2\int_0^{\pi/4} \sin t\sqrt{1 + 3\cos^2 t}\,dt\\ &= \frac{1}{2} + \frac{3}{4}\sqrt{5} + I_1 + I_2\,. \end{aligned}$$

For I_1, set $u = \sqrt{3}\sin t$, so $du = \sqrt{3}\cos t\,dt$, and obtain

$$I_1 = \frac{1}{\sqrt{3}}\int_0^{\sqrt{6}/2} \sqrt{4 - u^2}\,du\,.$$

Substituting $v = \frac{u}{2}$ we have

$$I_1 = \frac{1}{\sqrt{3}}\left[\frac{1}{2}u\sqrt{4 - u^2} + 2\arcsin\frac{u}{2}\right]_0^{\sqrt{6}/2} = \frac{\sqrt{5}}{4} + \frac{2}{\sqrt{3}}\arcsin\frac{\sqrt{6}}{4}\,.$$

Similarly for I_2, we set $u = \sqrt{3}\cos t$, so $du = -\sqrt{3}\sin t\,dt$ and

$$I_2 = -\frac{2}{\sqrt{3}}\int_{\sqrt{3}}^{\sqrt{6}/2} \sqrt{1 + u^2}\,du\,.$$

Then

$$\begin{aligned} I_2 &= -\frac{1}{\sqrt{3}}\left[u\sqrt{1 + u^2} + \log\left(\sqrt{1 + u^2} + u\right)\right]_{\sqrt{3}}^{\sqrt{6}/2}\\ &= -\frac{\sqrt{5}}{2} + 2 + \frac{1}{\sqrt{3}}\left(\log(2 + \sqrt{3}) - \log\frac{\sqrt{10} + \sqrt{6}}{2}\right)\,. \end{aligned}$$

4. $2\arctan\sqrt{2} + \frac{\sqrt{2}}{12}\pi$.

5. Since $\ell(\boldsymbol{\gamma}) = \ell(\boldsymbol{\gamma}_1) + \ell(\boldsymbol{\gamma}_2) + \ell(\boldsymbol{\gamma}_3)$, we compute the lengths separately. The last two are elementary, $\ell(\boldsymbol{\gamma}_2) = \frac{\sqrt{7}}{3}\pi$ and $\ell(\boldsymbol{\gamma}_3) = \pi$. As for the first one,

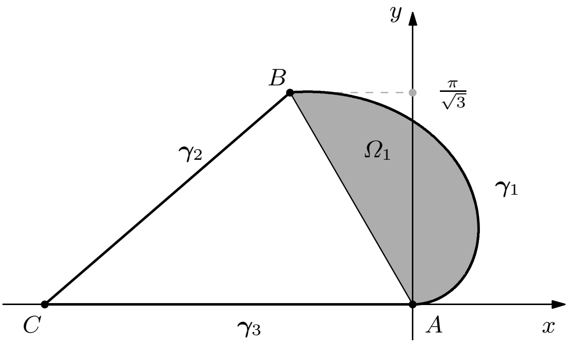


Figure 9.18. The arc $\boldsymbol{\gamma}$ and the region Ω_1 relative to Exercise 5

$$\begin{aligned}\ell(\boldsymbol{\gamma}_1) &= \int_0^{2/3\pi} \|\boldsymbol{\gamma}'(t)\|\,\mathrm{d}t = \int_0^{2/3\pi} \sqrt{1+t^2}\,\mathrm{d}t \\ &= \frac{1}{3}\pi\sqrt{1+\frac{4}{9}\pi^2} + \frac{1}{2}\log\Big(\frac{2}{3}\pi + \sqrt{1+\frac{4}{9}\pi^2}\Big)\,.\end{aligned}$$

The area is then the sum of the triangle ABC and the set Ω_1 of Fig. 9.18. We know that area$(ABC) = \frac{\pi^2}{2\sqrt{3}}$. As Ω_1 reads, in polar coordinates,

$$\Omega_1' = \Big\{(r,\theta\,) : 0 \le r \le \theta,\ 0 \le \theta \le \frac{2}{3}\pi\Big\}\,,$$

we have

$$\text{area}(\Omega_1) = \int_{\Omega_1} \mathrm{d}x\,\mathrm{d}y = \int_0^{(2/3)\pi}\int_0^{\theta} r\,\mathrm{d}r\,\mathrm{d}\theta = \frac{4}{81}\pi^3\,.$$

6. We have

$$x_G = \frac{\mathrm{e}^{\pi}-2}{5(\mathrm{e}^{\pi/2}-1)}\,, \qquad y_G = \frac{2\mathrm{e}^{\pi}+1}{5(1-\mathrm{e}^{\pi/2})}\,.$$

7. The parameter k is fixed by imposing $y_G = 0$. At the same time,

$$y_G = \int_{\gamma_1} y + \int_{\gamma_2} y$$

where $\boldsymbol{\gamma}_1(t) = (-\sqrt{3}t, t)$, $t \in [-2,0]$ and $\boldsymbol{\gamma}_2(t) = (\frac{1}{2}t^2, t)$, $t \in [0,k]$. Then

$$y_G = \int_{-2}^{0} 2t\,\mathrm{d}t + \int_0^k t\sqrt{1+t^2}\,\mathrm{d}t = \frac{1}{3}\big((1+k^2)^{3/2} - 13\big)\,;$$

the latter is zero when $k = \sqrt{13^{2/3}-1}$.

8. As $\boldsymbol{\gamma}'(t) = (\cos t - t\sin t)\boldsymbol{i} + (\sin t + t\cos t)\boldsymbol{j} + \boldsymbol{k}$ and $\|\boldsymbol{\gamma}'(t)\|^2 = 2+t^2$, we have

$$I_z = \int_{\Gamma} (x^2+y^2) = \int_0^{\sqrt{2}} t^2\sqrt{2+t^2}\,\mathrm{d}t = \frac{3}{2}\sqrt{2} - \frac{1}{2}\log(1+\sqrt{2})\,.$$

9. From $\boldsymbol{f}\big(\boldsymbol{\gamma}(t)\big) = (t^4, t^3)$ and $\boldsymbol{\gamma}'(t) = (2t, 1)$ follows

$$\int_{\boldsymbol{\gamma}} \boldsymbol{f} \cdot \mathrm{d}P = \int_0^1 (t^4, t^3) \cdot (2t, 1)\, \mathrm{d}t = \int_0^1 (2t^5 + t^3)\, \mathrm{d}t = \frac{7}{12}\,.$$

10. $\dfrac{9}{4}$; 11. $\dfrac{\pi}{4}$.

12. The piecewise-regular arc $\boldsymbol{\gamma}$ restricts to three regular arcs $\gamma_1,\ \gamma_2,\ \gamma_3$, whose traces are the segments AB, BC, CD. We can define arcs $\boldsymbol{\delta}_1,\ \boldsymbol{\delta}_2$ and $\boldsymbol{\delta}_3$ congruent to $\gamma_1,\ \gamma_2,\ \gamma_3$:

$$\begin{aligned}
&\boldsymbol{\delta}_1(t) = (t, 1) && 0 \le t \le 1\,, && \boldsymbol{\delta}_1 \sim \gamma_1\,,\\
&\boldsymbol{\delta}_2(t) = (t, 2-t) && 0 \le t \le 1\,, && \boldsymbol{\delta}_2 \sim -\gamma_2\,,\\
&\boldsymbol{\delta}_3(t) = (t, 2) && 0 \le t \le 1\,, && \boldsymbol{\delta}_3 \sim \gamma_3\,,
\end{aligned}$$

Then from

$$\begin{aligned}
&\boldsymbol{f}\big(\boldsymbol{\delta}_1(t)\big) = (t, t^2)\,, && \boldsymbol{f}\big(\boldsymbol{\delta}_2(t)\big) = \big(t(2-t)^2, t^2(2-t)\big)\,, && \boldsymbol{f}\big(\boldsymbol{\delta}_3(t)\big) = (4t, 2t^2)\\
&\boldsymbol{\delta}_1'(t) = (1, 0)\,, && \boldsymbol{\delta}_2'(t) = (1, -1)\,, && \boldsymbol{\delta}_3'(t) = (1, 0)\,,
\end{aligned}$$

we have

$$\begin{aligned}
\int_{\boldsymbol{\gamma}} \boldsymbol{f} \cdot \mathrm{d}P &= \int_{\boldsymbol{\delta}_1} \boldsymbol{f} \cdot \mathrm{d}P - \int_{\boldsymbol{\delta}_2} \boldsymbol{f} \cdot \mathrm{d}P + \int_{\boldsymbol{\delta}_3} \boldsymbol{f} \cdot \mathrm{d}P\\
&= \int_0^1 (t, t^2) \cdot (1, 0)\, \mathrm{d}t - \int_0^1 \big(t(2-t)^2, t^2(2-t)\big) \cdot (1, -1)\, \mathrm{d}t\\
&\qquad + \int_0^1 (4t, 2t^2) \cdot (1, 0)\, \mathrm{d}t = 2\,.
\end{aligned}$$

13. 0.

14. Let us impose

$$\int_{\boldsymbol{\gamma}} \boldsymbol{f} \cdot \boldsymbol{\tau} = \frac{9}{5}$$

where $\boldsymbol{\gamma}(t) = \big(\frac{1}{2k}t^2, t\big)$, $t \in [0, k]$ (note $k \neq 0$, otherwise the integral is zero). Since $\boldsymbol{\gamma}'(t) = \big(\frac{1}{k}t, 1\big)$,

$$\int_{\boldsymbol{\gamma}} \boldsymbol{f} \cdot \boldsymbol{\tau} = \int_0^k \Big(\frac{t^5}{4k^3} - \frac{t^4}{2k^2} + t^2 - \frac{t^4}{4k^2}\Big)\, \mathrm{d}t = \frac{9}{40}k^3\,.$$

Therefore $\frac{9}{40}k^3 = \frac{9}{5}$, so $k = 2$.

15. The work integral equals $L = \frac{2}{5} - \frac{2}{15}a$, which vanishes if $a = 3$.

16. The surface is the graph of $\varphi(u, v) = \sqrt{16 - u^2 - v^2}$, so (6.49) gives

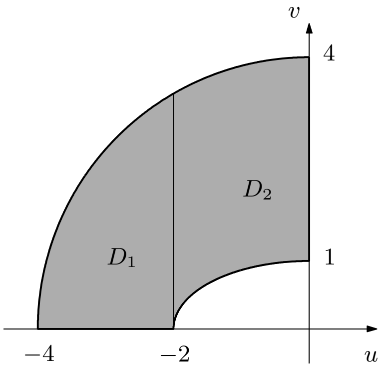


Figure 9.19. The regions D_1 and D_2 relative to Exercise 16

$$\|\boldsymbol{\nu}(u,v)\| = \sqrt{1+\Big(\frac{\partial\varphi}{\partial u}\Big)^2+\Big(\frac{\partial\varphi}{\partial v}\Big)^2} = \frac{4}{\sqrt{16-u^2-v^2}}\,.$$

Therefore

$$\begin{aligned}\int_{\boldsymbol{\sigma}} f &= \int_{\mathcal{R}} \sqrt{16-u^2-v^2}\,(v-2u)\frac{4}{\sqrt{16-u^2-v^2}}\,\mathrm{d}u\,\mathrm{d}v\\ &= 4\int_{\mathcal{R}} (v-2u)\,\mathrm{d}u\,\mathrm{d}v\,.\end{aligned}$$

The region $\mathcal{R}$ is in the second quadrant and lies between the circle with radius 4, centre the origin, and the ellipse with semi-axes $a=2$, $b=1$ (Fig. 9.19). Integrating in v first and dividing the domain into D_1 and D_2, we find

$$\int_{\boldsymbol{\sigma}} f = 4\Big(\int_{-4}^{-2}\int_0^{\sqrt{16-u^2}} (v-2u)\,\mathrm{d}v\,\mathrm{d}u + \int_{-2}^{0}\int_{\sqrt{1-\frac{u^2}{4}}}^{\sqrt{16-u^2}} (v-2u)\,\mathrm{d}v\,\mathrm{d}u\Big) = \frac{728}{3}\,.$$

17. A parametrisation for Σ is (Fig. 9.20)

$$\boldsymbol{\sigma}(u,v) = \Big(u, v, -\frac{u^2}{4}-v^2\Big)\,, \qquad (u,v)\in\mathcal{R} = \{(u,v)\in\mathbb{R}^2 : \frac{u^2}{4}+v^2=1\}\,.$$

Then $\|\boldsymbol{\nu}(u,v)\|^2 = 1+\frac{u^2}{4}+4v^2$ and

$$\int_{\Sigma} f = \int_{\mathcal{R}} (v+1)\,\mathrm{d}u\,\mathrm{d}v\,.$$

Passing to elliptical polar coordinates,

$$\int_{\Sigma} f = 2\int_0^{2\pi}\int_0^1 (r\sin\theta+1)r\,\mathrm{d}r\,\mathrm{d}\theta = 2\pi\,.$$

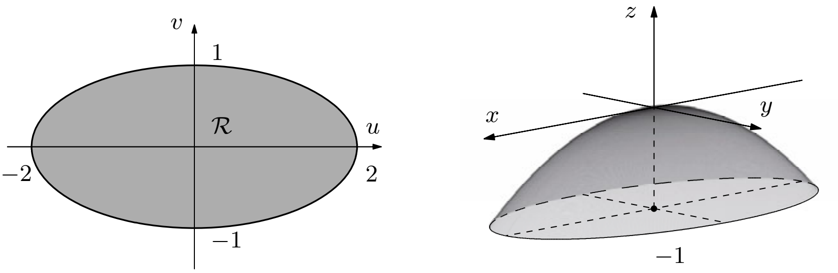


Figure 9.20. The region $\mathcal{R}$ (left) and the surface Σ (right) relative to Exercise 17

18. Parametrise Σ by

$$\boldsymbol{\sigma}(u,v) = \big(u, v, \frac{1}{2}v^2\big), \qquad (u,v) \in \mathcal{R},$$

where $\mathcal{R}$ is the triangle in the uv-plane of vertices $(-4,0)$, $(4,0)$, $(0,4)$ (Fig. 9.21). In this way $\|\boldsymbol{\nu}(u,v)\| = \sqrt{1+v^2}$, so

$$\begin{aligned} \operatorname{area}(\Sigma) &= \int_\Sigma 1 = \int_{\mathcal{R}} \sqrt{1+v^2}\,\mathrm{d}u\,\mathrm{d}v = \int_0^4 \Big(\int_{v-4}^{4-v} \sqrt{1+v^2}\,\mathrm{d}u\Big)\,\mathrm{d}v \\ &= \frac{14}{3}\sqrt{17} + 4\log(4+\sqrt{17}) + \frac{2}{3}. \end{aligned}$$

19. $\operatorname{area}(\Sigma) = 8\pi$.

20. Using cylindrical coordinates Σ reads

$$\boldsymbol{\sigma}(r,\theta) = (r, \theta, \sqrt{3}r), \qquad (r,\theta) \in \mathcal{R} = \{(r,\theta) : 0 \le r \le \sqrt{3}, \frac{\pi}{4} \le \theta \le \frac{5}{4}\pi\}$$

(see Fig. 9.22). Then $\|\boldsymbol{\nu}(r,\theta)\| = 2$ and

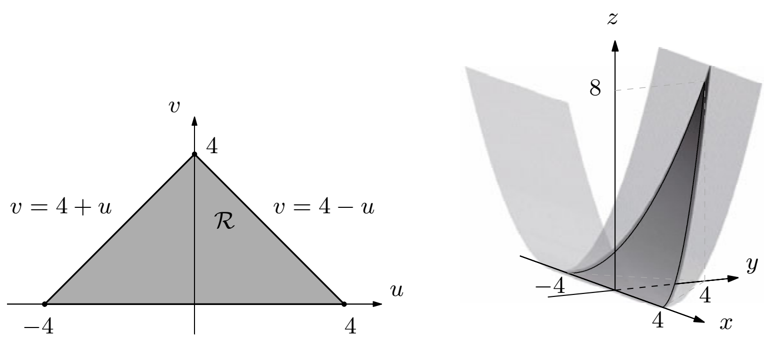
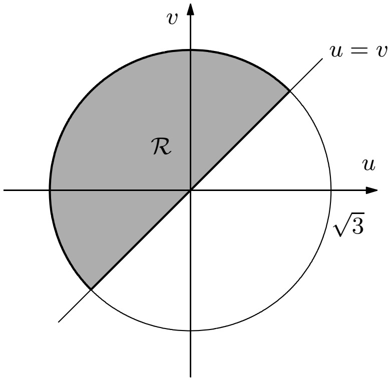


Figure 9.21. The region $\mathcal{R}$ (left) and the surface Σ (right) relative to Exercise 18



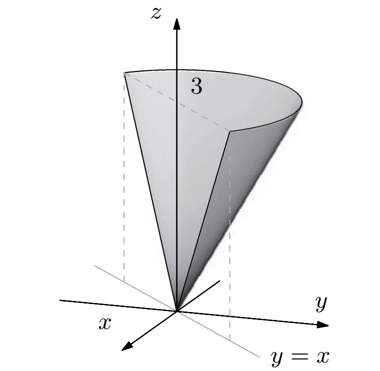


Figure 9.22. The region $\mathcal{R}$ (left) and the surface Σ (right) relative to Exercise 20

$$\int_{\Sigma}(x^2+y^2) = 2\int_{\mathcal{R}} r^2 r\,\mathrm{d}r\,\mathrm{d}\theta = 2\int_{\pi/4}^{(5/4)\pi}\int_0^{\sqrt{3}} r^3\,\mathrm{d}r\,\mathrm{d}\theta = \frac{9}{2}\pi\,.$$

21. $x_G = \frac{2\pi}{\pi+4}$.

22. The arc is shown in Fig. 9.23. The work equals $L = \oint_{\Gamma} \boldsymbol{f}\cdot\boldsymbol{\tau}$, but also

$$L = \int_{\Omega}\Big(\frac{\partial f_2}{\partial x} - \frac{\partial f_1}{\partial y}\Big)\,\mathrm{d}x\,\mathrm{d}y$$

by Green's Theorem, with Ω inside Γ and $f_1(x,y) = xy$, $f_2(x,y) = x^4 y$. Then if we integrate in x first,

$$L = \int_0^1\int_{y^2}^{\sqrt{5-y^2}}(4x^3y - x)\,\mathrm{d}x\,\mathrm{d}y = \frac{47}{6}\,.$$

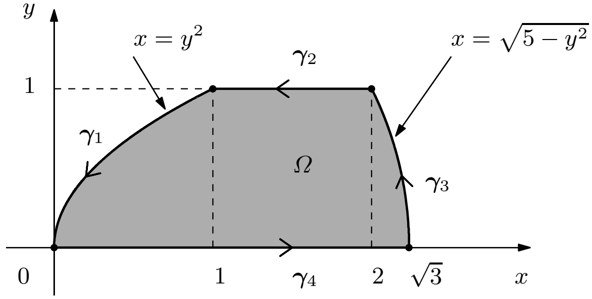


Figure 9.23. The arc Γ and the region Ω relative to Exercise 22

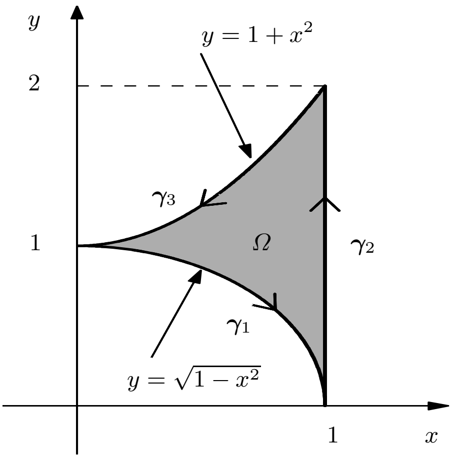


Figure 9.24. The arc Γ and the region Ω relative to Exercise 23

23. The arc Γ is shown in Fig. 9.24. Green's Theorem implies the work is

$$L = \oint_\Gamma \boldsymbol{f} \cdot \boldsymbol{\tau} = \int_\Omega (a - 2x^2 y)\,\mathrm{d}x\,\mathrm{d}y\,,$$

where Ω is inside Γ. Integrating vertically,

$$L = \int_0^1 \int_{\sqrt{1-x^2}}^{x^2+1} (a - 2x^2 y)\,\mathrm{d}y\,\mathrm{d}x = \int_0^1 \Big[ay - x^2 y^2\Big]_{y=\sqrt{1-x^2}}^{y=x^2+1} \mathrm{d}x = a\Big(\frac{4}{3} - \frac{\pi}{4}\Big) - \frac{26}{35}\,.$$

24. We use (9.23) on the arc Γ (see Fig. 9.25). The integral along OA and OB is zero; since $\boldsymbol{\gamma}_1'(t) = (-\mathrm{e}^{-(t-\pi/4)}(\cos t + \sin t), \cos t)$, it follows

$$\text{area}(\Omega) = \frac{1}{2}\int_0^{\pi/4} \big(\sin t(\cos t + \sin t) + \cos^2 t\big)\mathrm{e}^{-(t-\pi/4)}\,\mathrm{d}t = -\frac{11}{20} + \frac{3}{5}\mathrm{e}^{\pi/4}\,.$$

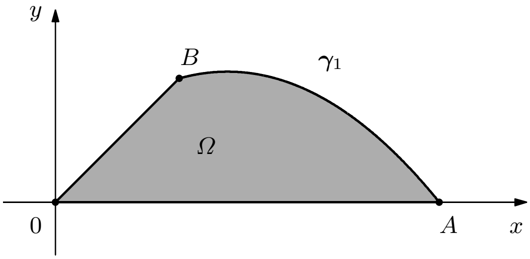


Figure 9.25. The arc Γ and the region Ω relative to Exercise 24

25. The Divergence Theorem tells

$$\int_{\partial\Omega} \boldsymbol{f}\cdot\boldsymbol{n} = \int_{\Omega} \operatorname{div}\boldsymbol{f} = \int_{\Omega} 3(x^2+y^2+z^2)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\,.$$

In spherical coordinates Ω becomes Ω' defined by $0 \le r \le 1$, $0 \le \varphi \le \pi/4$ and $0 \le \theta \le \pi$. Then

$$\int_{\partial\Omega} \boldsymbol{f}\cdot\boldsymbol{n} = \int_0^{\pi}\int_0^{\pi/4}\int_0^1 3r^4\sin\varphi\,\mathrm{d}r\,\mathrm{d}\varphi\,\mathrm{d}\theta = \frac{3}{10}\pi(2-\sqrt{2})\,.$$

26. The flux is $\frac{8}{5}\pi(\sqrt{2}-1)$.

27. Let us parametrise Σ by

$$\boldsymbol{\sigma}(u,v) = \big(u, v, \frac{u^2}{9}+\frac{v^2}{4}\big), \qquad \text{with} \qquad (u,v)\in\mathcal{R} = \big\{(u,v)\in\mathbb{R}^2 : u^2+v^2\le 4\big\}\,.$$

Then $\boldsymbol{\nu}(u,v) = (-\frac{2}{9}u, -\frac{v}{2}, 1)$ is oriented as required. In order to use Stokes' Theorem, notice $\mathbf{curl}\,\boldsymbol{f} = x\boldsymbol{i} - y\boldsymbol{j} + 0\boldsymbol{k}$, and

$$\begin{aligned}\int_{\partial\Sigma} \boldsymbol{f}\cdot\boldsymbol{\tau} &= \int_{\Sigma} \mathbf{curl}\,\boldsymbol{f}\cdot\boldsymbol{n} = \int_{\mathcal{R}} (u\boldsymbol{i} - v\boldsymbol{j} + 0\boldsymbol{k})\cdot\big(-\frac{2}{9}u\,\boldsymbol{i} - \frac{v}{2}\boldsymbol{j} + \boldsymbol{k}\big)\,\mathrm{d}u\,\mathrm{d}v \\ &= \int_{\mathcal{R}} (-\frac{2}{9}u^2 + \frac{1}{2}v^2)\,\mathrm{d}u\,\mathrm{d}v\,.\end{aligned}$$

In polar coordinates,

$$\int_{\partial\Sigma} \boldsymbol{f}\cdot\boldsymbol{\tau} = \int_0^{2\pi}\int_0^2 \big(-\frac{2}{9}\cos^2\theta + \frac{1}{2}\sin^2\theta\big) r^3\,\mathrm{d}r\,\mathrm{d}\theta = \frac{10}{9}\pi\,.$$

28. We use Stokes' Theorem with

$$\Sigma = \big\{(x,y,z)\in\mathbb{R}^3 : x^2+y^2+z^2 = 2,\ z\ge y\big\}\,,$$

whose boundary is $\partial\Sigma = \{(x,y,z)\in\mathbb{R}^3 : x^2+y^2+z^2 = 2,\ z = y\}$. So we parametrise $\partial\Sigma$ by

$$\boldsymbol{\gamma}(t) = (\sqrt{2}\cos t, \sin t, \sin t)\,, \qquad t\in[0,2\pi]\,;$$

then

$$\begin{aligned}&\boldsymbol{f}\big(\boldsymbol{\gamma}(t)\big) = 2\sin t\,\boldsymbol{i} + 2(\sqrt{2}\cos t + \sin t)\,\boldsymbol{j} + 3(\sqrt{2}\cos t + \sin t)\,\boldsymbol{k}\,, \\ &\boldsymbol{\gamma}'(t) = -\sqrt{2}\sin t\,\boldsymbol{i} + \cos t\,\boldsymbol{j} + \cos t\,\boldsymbol{k}\end{aligned}$$

and

$$\int_{\Sigma} \mathbf{curl}\,\boldsymbol{f}\cdot\boldsymbol{n} = \int_{\partial\Sigma} \boldsymbol{f}\cdot\boldsymbol{\tau} = 3\sqrt{2}\pi\,.$$

29. The field $\boldsymbol{f}$ is defined on the whole $\mathbb{R}^3$ (clearly simply connected), and its curl is zero. By Theorem 9.45 then, $\boldsymbol{f}$ is conservative.

To find a potential φ, start from

$$\frac{\partial\varphi}{\partial x} = x\,, \qquad \frac{\partial\varphi}{\partial y} = -2y\,, \qquad \frac{\partial\varphi}{\partial z} = 3z\,. \tag{9.27}$$

From the first we get

$$\varphi(x,y,z) = \frac{1}{2}x^2 + \psi_1(y,z)\,.$$

Differentiating in y and using the second equation in (9.27) gives

$$\frac{\partial\varphi}{\partial y}(x,y,z) = \frac{\partial\psi_1}{\partial y}(y,z) = -2y\,,$$

hence

$$\psi_1(y,z) = -y^2 + \psi_2(z) \qquad \text{and} \qquad \varphi(x,y,z) = \frac{1}{2}x^2 - y^2 + \psi_2(z)\,.$$

Now we differentiate φ in z and use the last in (9.27):

$$\frac{\partial\varphi}{\partial z}(x,y,z) = \frac{\mathrm{d}\psi_2}{\mathrm{d}z}(z) = 3z\,, \quad \text{so} \quad \psi_2(z) = \frac{3}{2}z^2 + c\,.$$

All in all, a potential for $\boldsymbol{f}$ is

$$\varphi(x,y,z) = \frac{1}{2}x^2 - y^2 + \frac{3}{2}z^2 + c\,, \qquad c \in \mathbb{R}\,.$$

30. The field is defined over the simply connected plane $\mathbb{R}^2$. It is thus enough to impose

$$\frac{\partial f_1}{\partial y}(x,y) = \frac{\partial f_2}{\partial x}(x,y)\,,$$

where $f_1(x,y) = y\sin x + xy\cos x + \mathrm{e}^y$ and $f_2(x,y) = g(x) + x\mathrm{e}^y$, for the field to be conservative. This translates to

$$\sin x + x\cos x + \mathrm{e}^y = g'(x) + \mathrm{e}^y\,,$$

i.e., $g'(x) = \sin x + x\cos x$. Integrating,

$$g(x) = \int (\sin x + x\cos x)\,\mathrm{d}x = x\sin x + c\,, \qquad c \in \mathbb{R}\,.$$

The condition $g(0) = 0$ forces $c = 0$, so the required map is $g(x) = x\sin x$.

To find a potential φ for $\boldsymbol{f}$, we must necessarily have

$$\frac{\partial\varphi}{\partial y}(x,y,z) = y\sin x + xy\cos x + \mathrm{e}^y \qquad \text{and} \qquad \frac{\partial\varphi}{\partial y}(x,y,z) = x\sin x + x\mathrm{e}^y\,.$$

The second equation gives

$$\varphi(x,y) = xy\sin x + x\mathrm{e}^y + \psi(x)\,;$$

differentiating in x and using the first equation gives

$$\frac{\partial\varphi}{\partial x}(x,y,z) = y\sin x + xy\cos x + \mathrm{e}^y + \psi'(x) = y\sin x + xy\cos x + \mathrm{e}^y\,.$$

Therefore $\psi'(x) = 0$, so $\psi(x) = c$ for an arbitrary constant c in $\mathbb{R}$. In conclusion,

$$\varphi(x,y) = xy\sin x + x\mathrm{e}^y + c\,, \qquad c \in \mathbb{R}\,.$$

31. a) $g(y) = y^2$; b) $\varphi(x,y) = y^2\log x + x\cos y$.

32. a) $g(x) = x^2 + 1$; b) $\varphi(x,y) = (x^2+1)\log y + y\cos x - \frac{1}{2}$.

33. a) $g(z) = z - 1$; b) $\varphi(x,y,z) = 3xy + \frac{1}{2}y^2 + yz - y + \sin(x+z^2)$.

34. $\lambda = 1$ e $\varphi(x,y,z) = \frac{1}{3}x^3 + 3xyz + 5xy - 2z^2 - 2y + 4$.

10

Ordinary differential equations

Differential equations are among the mathematical instruments most widely used in applications to other fields. The book's final chapter presents in a self-contained way the main notions, theorems and techniques of so-called ordinary differential equations. After explaining the basic principles underlying the theory we review the essential methods for solving several types of equations. We tackle existence and uniqueness issues for a solution to the initial value problem of a vectorial equation, then describe the structure of solutions of linear systems, for which algebraic methods play a central role, and equations of order higher than one. In the end the reader will find a concise introduction to asymptotic stability, with applications to pendulum motion, with and without damping.

Due to the matter's richness and complexity, this exposition is far from exhaustive and will concentrate on the qualitative aspects especially, keeping rigorous arguments to a minimum[1].

10.1 Introductory examples

The reader will presumably be already familiar with differential equation that formalise certain physical principles, like Newton's law of motion. Differential equations are natural tools, hence widely employed to build *mathematical models* describing phenomena and processes of the real world. The reason is that two quantities (the variables) often interact by affecting one another's *variation*, which is expressed by a relationship between one variable and some derivative of the other. Newton's law, for instance, states that a force influences a particle's acceleration, i.e., the change of its velocity in time; and velocity itself is the variation of displacement. If the force is a function of the particle's position, as happens in the case of a spring, we obtain an equation of motion in terms of time t

[1] For proofs, the interested reader may consult, among others, the rich monograph by G. Teschl, *Ordinary Differential Equations and Dynamical Systems*, American Mathematical Society, 2012.

C. Canuto, A. Tabacco: *Mathematical Analysis II,* 2nd Ed.,
UNITEXT – La Matematica per il 3+2 85, DOI 10.1007/978-3-319-12757-6_10,

$$\boxed{m\frac{\mathrm{d}^2x}{\mathrm{d}t^2} = -kx\,;}$$

what this says is that the spring's pulling force is proportional to the displacement, and acts against motion ($k > 0$). A more complicated example is

$$\boxed{m\frac{\mathrm{d}^2x}{\mathrm{d}t^2} = -(k + \beta x^2)x\,,}$$

where the spring's rigidity is enhanced by an increase of the displacement; both these are *differential equations* of second order in the variable t. Another example originating from classical Mechanics, that will accompany us throughout the chapter, is that of a pendulum swinging freely on a vertical plane under the effect of gravity and possibly some friction. The relationship between the angle θ swept by the rod and the bob's angular velocity and acceleration reads

$$\boxed{\frac{\mathrm{d}^2\theta}{\mathrm{d}t^2} + \alpha\frac{\mathrm{d}\theta}{\mathrm{d}t} + k\sin\theta = 0\,.}$$

Multiple mass-spring systems, or compound pendulums, are governed by *systems* of differential equations correlating the positions of the various masses involved, or the angles of the rods forming the pendulum.

The differential equations of concern, like all those treated in the chapter, are *ordinary*, which means the unknowns depend upon *one* independent variable, conventionally called t (the prevailing applications are dynamical, of evolution type, so t is time). *Partial* differential equations, instead, entail unknown functions of several variables (like time and space coordinates), and consequently also partial derivatives.

Ordinary differential equations and systems are fundamental to explain how electric and electronic circuits work. The simplest situation is that of an *LRC* circuit, where the current $i = i(t)$ satisfies a linear equation of second order of the type

$$\boxed{\ell\frac{\mathrm{d}^2i}{\mathrm{d}t^2} + r\frac{\mathrm{d}i}{\mathrm{d}t} + \frac{i}{c} = f\,.}$$

Electric networks can be modelled by systems of as many of the above equations as the number of loops, to which one adds suitable conservation laws at the nodes (Kirchhoff's laws). For electronic circuits with active components, differential models are typically non-linear.

Materials Sciences, Chemistry, Biology, and more generally Life and Social Sciences, are all fertile sources of mathematical models based on ordinary differential equations. For instance,

$$\boxed{y' = ky}$$

describes the radioactive decay of a substance y (the rate of disappearance is proportional, for $k < 0$, to the quantity of matter itself). But the same equation regulates the dynamics of populations (called in this case Malthus' law, it predicts that the rate of growth y'/y of a population in time equals the difference $k = n-m$ between the constant birth rate n and death rate m). Malthus' model becomes realistic by further assuming k is corrected by a braking term that prevents unlimited growth, via the so-called logistic growth equation

$$\boxed{y' = (k - \beta y)y\,.}$$

Another example are the Volterra-Lotka equations

$$\boxed{\begin{cases} p_1' = (k_1 - \beta_1 p_2)p_1 \\ p_2' = (-k_2 + \beta_2 p_1)p_2\,, \end{cases}}$$

that preside over the interactions of two species, labelled preys p_1 and predators p_2, whose relative change rates p_i'/p_i are functions not only of the species' respective features ($k_1, k_2 > 0$), but also of the number of antagonist individuals present on the territory ($\beta_1, \beta_2 > 0$).

At last, we mention a source – intrinsic, so to say, to differential modelling – of very large systems of differential equations: the discretisation with respect to the spatial variables of partial differential equations describing evolution phenomena. A straightforward example is the *heat equation*

$$\frac{\partial u}{\partial t} - k\frac{\partial^2 u}{\partial x^2} = 0\,, \qquad 0 < x < L\,,\ t > 0\,, \tag{10.1}$$

controlling the termperature $u = u(t)$ of a metal bar of length L in time. Dividing the bar in $n+1$ parts of width Δx by points $x_j = j\Delta x$, $0 \le j \le n+1$ (with $(n+1)\Delta x = L$), we can associate to each node the variable $u_j = u_j(t)$ telling how the temperature at x_j changes in time. Using Taylor expansions, the second spatial derivative can be approximated by a difference quotient

$$\frac{\partial^2 u}{\partial x^2}(x_j, t) \sim \frac{u_{j-1}(t) - 2u_j(t) + u_{j+1}(t)}{\Delta x^2}\,,$$

so equation (10.1) is well approximated by a system of n linear ordinary differential equations, one for each internal node:

$$\boxed{u_j' - \frac{k}{\Delta x^2}(u_{j-1} - 2u_j + u_{j+1}) = 0\,, \qquad 1 \le j \le n\,.}$$

The first and last equation of the system contain the temperatures u_0 and u_{n+1} at the bar's ends, which can be fixed by suitable boundary conditions (e.g., $u_0(t) = \phi$, $u_{n+1}(t) = \psi$ if the temperatures at the ends are kept constant by thermostats).

10.2 General definitions

We introduce first of all ‘scalar’ differential equations (those with one unknown), for the benefit of readers who have not seen them in other courses.

An **ordinary differential equation** (**ODE** is the standard acronym) is a relationship between a real independent variable (here, t), an unknown function $y = y(t)$, and the derivatives $y^{(k)}$ up to order n

$$\boxed{\mathcal{F}(t, y, y', ..., y^{(n)}) = 0,} \tag{10.2}$$

where $\mathcal{F}$ is a real map of $n+2$ real variables. The differential equation has **order** n, if n is the highest order of the derivatives of y appearing in (10.2). A **solution** (in the classical sense) of the differential equation on the real interval J is a map $y : J \to \mathbb{R}$, differentiable n times on J, such that

$$\boxed{\mathcal{F}(t, y(t), y'(t), ..., y^{(n)}(t)) = 0 \qquad \text{for all } t \in J.}$$

Using (10.2) it is often useful to write the highest derivative $y^{(n)}$ in terms of t and the other derivatives (in several applications this is precisely the way a differential equation is typically written). In general, the Implicit Function Theorem (Sect. 7.1) makes sure that equation (10.2) can be solved for $y^{(n)}$ when the partial derivative of $\mathcal{F}$ in the last variable is non-zero. If so, (10.2) reads

$$\boxed{y^{(n)} = f(t, y, ..., y^{(n-1)}),} \tag{10.3}$$

with f a real map of $n+1$ real variables. Then one says the differential equation is in **normal form**. The definition of solution for normal ODEs changes accordingly.

As the warm-up examples have shown, beside single equations also **systems** of differential equations are certainly worth studying. The simplest instance is that of a system of order one in n unknowns, written in normal form as

$$\begin{cases} y_1' = f_1(t, y_1, y_2, \ldots, y_n) \\ y_2' = f_2(t, y_1, y_2, \ldots, y_n) \\ \quad\vdots \\ y_n' = f_n(t, y_1, y_2, \ldots, y_n) , \end{cases} \tag{10.4}$$

where f_i is a function of $n+1$ variables. The vectorial notation

$$\boxed{\boldsymbol{y}' = \boldsymbol{f}(t, \boldsymbol{y})} \tag{10.5}$$

where $\boldsymbol{y} = (y_i)_{1 \le i \le n}$ and $\boldsymbol{f} = (f_i)_{1 \le i \le n}$ is quite convenient. A **solution** is now understood as a vector-valued map $\boldsymbol{y} : J \to \mathbb{R}^n$, differentiable on J and satisfying

$$\boldsymbol{y}'(t) = \boldsymbol{f}\big(t, \boldsymbol{y}(t)\big) \qquad \text{for all } t \in J\,.$$

An especially relevant case is that of linear systems

$$\begin{cases} y_1' = a_{11}(t)y_1 + a_{12}(t)y_2 + \ldots + a_{1n}(t)y_n + b_1(t) \\ y_2' = a_{21}(t)y_1 + a_{22}(t)y_2 + \ldots + a_{2n}(t)y_n + b_2(t) \\ \quad\vdots \\ y_n' = a_{n1}(t)y_1 + a_{n2}(t)y_2 + \ldots + a_{nn}(t)y_n + b_n(t)\,, \end{cases}$$

or, in compact form,

$$\boxed{\boldsymbol{y}' = \boldsymbol{A}(t)\boldsymbol{y} + \boldsymbol{b}(t)} \tag{10.6}$$

where $\boldsymbol{A}(t) = \big(a_{ij}(t)\big)_{1\le i,j\le n}$ maps J to the vector space $\mathbb{R}^{n,n}$ or square $n \times n$ matrices, and $\boldsymbol{b}(t) = \big(b_i(t)\big)_{1\le i\le n}$ is a map from J to $\mathbb{R}^n$. First-order linear equations play a particularly important role, both theoretically and in view of the applications: on one hand, they describe many problems that are linear in nature, on the other hand they approximate more complicated, non-linear equations by a linearisation process (see Remark 10.25).

In writing a differential equation like (10.5) or (10.6), it is customary *not* to write the t-dependency of the solution $\boldsymbol{y}$ explicitly.

Systems of first order are the main focus of our study, because they capture many types of differential equations provided one adds the necessary number of unknowns. Each differential equation of order n can be indeed written as a system of order one in n unknonws. To be precise, given equation (10.3), set

$$y_i = y^{(i-1)}\,, \qquad 1 \le i \le n\,,$$

i.e.,

$$\boxed{\begin{aligned} y_1 &= y^{(0)} = y \\ y_2 &= y' = y_1' \\ y_3 &= y'' = (y')' = y_2' \\ &\ \vdots \\ y_n &= y^{(n-1)} = (y^{(n-2)})' = y_{n-1}'\,. \end{aligned}} \tag{10.7}$$

The differential equation then becomes

$$y_n' = f(t, y_1, \ldots, y_n)\,,$$

so we obtain the first-order system

$$\boxed{\begin{cases} y_1' = y_2 \\ \quad\vdots \\ y_{n-1}' = y_n \\ y_n' = f(t, y_1, y_2, \ldots, y_n)\,. \end{cases}} \tag{10.8}$$

It is clear that any solution of equation (10.3) generates a solution $\boldsymbol{y}$ of the system, by (10.7); conversely, if $\boldsymbol{y}$ solves the system, it is easy to convince ourselves the first component y_1 is a solution of the equation. System (10.8) is then equivalent to equation (10.3).

The generalisation of (10.4) is a system of n differential equations in n unknonws, where each equation has order greater or equal than 1. Every equation of order at least two transforms into a system of first order. Altogether, then, we can always reduce to a system of order one in $m \geq n$ equations and unknonws.

Let us go back to equation (10.5), and suppose it is defined on the open set $\Omega = I \times D \subseteq \mathbb{R}^{n+1}$, where $I \subseteq \mathbb{R}$ is open and $D \subseteq \mathbb{R}^n$ open, connected; assume further f is continuous on Ω.

Definition 10.1 *A* **solution** *of the differential equation* (10.5) *is a* $\mathcal{C}^1$ *function* $\boldsymbol{y} = \boldsymbol{y}(t) : J \to D$, *with* J *a non-empty open interval in* I, *such that* $\boldsymbol{y}'(t) = \boldsymbol{f}\big(t, \boldsymbol{y}(t)\big)$, $\forall t \in J$.

We remark that a solution may not be defined on the entire I.

A solution $\boldsymbol{y}(t)$ is therefore a differentiable *curve* defined on J with trace contained in D. The vector $\boldsymbol{f}\big(t_*, \boldsymbol{y}(t_*)\big)$, when not $\mathbf{0}$, is tangent to the curve at each point $t_* \in J$.

The graph of $\boldsymbol{y}(t)$

$$G(\boldsymbol{y}) = \big\{(t, \boldsymbol{y}(t)) \in I \times D \subseteq \mathbb{R}^{n+1} : t \in J\big\}$$

is called an **integral curve** of the differential equation (see Fig. 10.1).

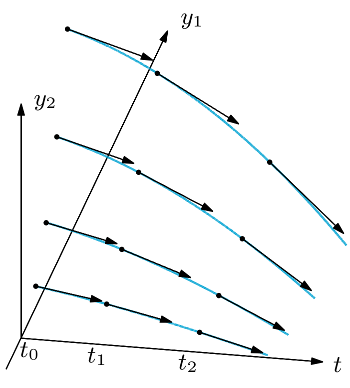


Figure 10.1. Integral curves of a differential equation and corresponding tangent vectors

The differential equation has, in general, infinitely many solutions. Typically these depend, apart from t, upon n arbitrary constants $c_1, \dots, c_n$. We indicate with $\boldsymbol{y}(t; c_1, \dots, c_n)$ the set of solutions, which is called **general integral**.

A natural way to select a special solution is to impose that this solution takes at the time $t_0 \in I$ a given value $\boldsymbol{y}_0 \in D$. We thus consider the problem of finding $\boldsymbol{y} = \boldsymbol{y}(t)$ such that

$$\boxed{\begin{cases} \boldsymbol{y}' = \boldsymbol{f}(t, \boldsymbol{y}) & \text{in } J\,, \\ \boldsymbol{y}(t_0) = \boldsymbol{y}_0\,, \end{cases}} \tag{10.9}$$

where J is an open sub-interval of I containing t_0. Since J in general depends on the solution, it cannot be determined a priori. This kind of question is called **Cauchy problem** for the differential equation, or **initial value problem** because it models the temporal evolution of a physical system, which at t_0, the initial moment of the mathematical simulation, is in the configuration $\boldsymbol{y}_0$. Geometrically speaking, the condition at time t_0 is equivalent to asking that the integral curve pass through the point $(t_0, \boldsymbol{y}_0) \in \Omega$. A Cauchy problem may be solvable locally (on a neighbourhood J of t_0), or globally ($J = I$, in which case one speaks of a *global solution*); these two situations will be treated in Sects. 10.4.1 and 10.4.3, respectively.

A particularly important class of ODEs is that of **autonomous** equations

$$\boxed{\boldsymbol{y}' = \boldsymbol{f}(\boldsymbol{y})\,,}$$

for which $\boldsymbol{f}$ does not depend explicitly on t. The linear system

$$\boldsymbol{y}' = \boldsymbol{A}\boldsymbol{y}$$

with given $\boldsymbol{A} \in \mathbb{R}^{n,n}$, is one such example, and will be studied in Sect. 10.6. For autonomous ODEs, the structure of solutions can be analysed by looking at their traces in the open set $D \subseteq \mathbb{R}^n$. The space $\mathbb{R}^n$ then represents the **phase space** of the equation.

The projection of an integral curve on the phase space (Fig. 10.2)

$$\boxed{\Gamma(\boldsymbol{y}) = \big\{\boldsymbol{y}(t) \in D : t \in J\big\}}$$

is called an **orbit** or **trajectory** of the solution $\boldsymbol{y}$. A solution $\boldsymbol{y}$ of the Cauchy problem (10.9) defines an *orbit passing through* $\boldsymbol{y}_0$, which consists of a *past trajectory* ($t < t_0$) and a *future trajectory* ($t > t_0$). The orbit is **closed** if it is closed viewed as trace of $\boldsymbol{y}$. Conventionally an orbit in phase space is pictured by the streamline of the flow, and the orientation on it describes the generating solution's evolution.

Application: the simple pendulum (I). The simple (gravity) pendulum is the idealised model of a mass m (the bob) fixed to the end of a massless rod of length L suspended from a pivot (the point O). When given an initial push, the pendulum

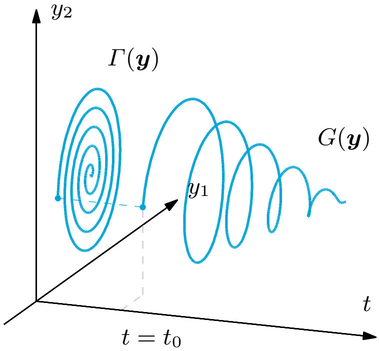


Figure 10.2. An integral curve and the orbit in phase space $y_1 y_2$ of an autonomous equation

swings back and forth on a vertical plane, and the bob describes a circle with centre O and radius L (Fig. 10.3). In absence of external forces, e.g., air drag or friction of sorts, the bob is subject to the force of gravity $\boldsymbol{g}$, a vector lying on the fixed plane of motion and pointing downwards. To describe the motion of m, we use polar coordinates (r,θ) for the point P on the plane where m belongs: the pole is O and the unit vector $\boldsymbol{i} = (\cos 0, \sin 0)$ is vertical and heads downwards; let $\boldsymbol{t}_r = \boldsymbol{t}_r(\theta)$, $\boldsymbol{t}_\theta = \boldsymbol{t}_\theta(\theta)$ be the orthonormal frame, as of (6.31).

The position of P in time is given by the vector $OP(t) = L\boldsymbol{t}_r(\theta(t))$, itself determined by Newton's law

$$m\frac{\mathrm{d}^2 OP}{\mathrm{d}t^2} = \boldsymbol{g}\,. \tag{10.10}$$

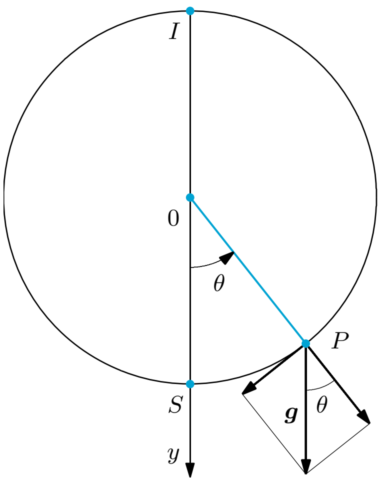


Figure 10.3. The simple pendulum

Successively differentiating $OP(t)$, with the aid of (6.39) we obtain

$$\frac{\mathrm{d}OP}{\mathrm{d}t} = L\frac{\mathrm{d}\boldsymbol{t}_r}{\mathrm{d}\theta}\frac{\mathrm{d}\theta}{\mathrm{d}t} = L\frac{\mathrm{d}\theta}{\mathrm{d}t}\boldsymbol{t}_\theta\,,$$

$$\frac{\mathrm{d}^2OP}{\mathrm{d}t^2} = L\frac{\mathrm{d}^2\theta}{\mathrm{d}t^2}\boldsymbol{t}_\theta + L\frac{\mathrm{d}\theta}{\mathrm{d}t}\frac{\mathrm{d}\boldsymbol{t}_\theta}{\mathrm{d}\theta}\frac{\mathrm{d}\theta}{\mathrm{d}t} = L\frac{\mathrm{d}^2\theta}{\mathrm{d}t^2}\boldsymbol{t}_\theta - L\left(\frac{\mathrm{d}\theta}{\mathrm{d}t}\right)^2\boldsymbol{t}_r\,.$$

On the other hand, if g is the modulus of gravity's pull, then $\boldsymbol{g} = mg\boldsymbol{i} = mg\cos\theta\,\boldsymbol{t}_r - mg\sin\theta\,\boldsymbol{t}_\theta$, so Newton's law becomes

$$mL\left(-\left(\frac{\mathrm{d}\theta}{\mathrm{d}t}\right)^2\boldsymbol{t}_r + \frac{\mathrm{d}^2\theta}{\mathrm{d}t^2}\boldsymbol{t}_\theta\right) = mg\big(\cos\theta\,\boldsymbol{t}_r - \sin\theta\,\boldsymbol{t}_\theta\big)\,.$$

Taking only the angular component, with $k = g/L > 0$, produces the equation of motion

$$\boxed{\frac{\mathrm{d}^2\theta}{\mathrm{d}t^2} + k\sin\theta = 0\,;} \tag{10.11}$$

notice that the mass does not appear anywhere in the formula.

A mathematically more realistic model takes into account the deceleration generated by a damping force at the point O: this will be proportional but opposite to the velocity, and given by the vector

$$\boldsymbol{a} = -\alpha mL\,\frac{\mathrm{d}\theta}{\mathrm{d}t}\,\boldsymbol{t}_\theta\,, \qquad \text{with } \alpha > 0\,,$$

adding to $\boldsymbol{g}$ in Newton's equation. As previously, the equation of motion (10.11) becomes

$$\boxed{\frac{\mathrm{d}^2\theta}{\mathrm{d}t^2} + \alpha\frac{\mathrm{d}\theta}{\mathrm{d}t} + k\sin\theta = 0\,.} \tag{10.12}$$

An ideal (free, undamped) motion is subsumed by taking $\alpha = 0$, so we shall suppose from now on $\alpha \geq 0$ and $k > 0$.

To find the position of P at time t, we first need to know its position and velocity at the starting time, say $t_0 = 0$. Let us then declare $\theta_0 = \theta(0)$ and $\theta_1 = \frac{\mathrm{d}\theta}{\mathrm{d}t}(0)$, which, in essence, is the statement of an initial value problem for an equation of order two:

$$\boxed{\begin{cases} \dfrac{\mathrm{d}^2\theta}{\mathrm{d}t^2} + \alpha\dfrac{\mathrm{d}\theta}{\mathrm{d}t} + k\sin\theta = 0\,, \quad t > 0\,, \\[2mm] \theta(0) = \theta_0\,, \quad \dfrac{\mathrm{d}\theta}{\mathrm{d}t}(0) = \theta_1\,. \end{cases}} \tag{10.13}$$

To capture the model's properties, let us transform this to first-order form (10.34) by setting

$$\boldsymbol{y} = (y_1, y_2) = (\theta, \frac{\mathrm{d}\theta}{\mathrm{d}t}), \quad \boldsymbol{y}_0 = (\theta_0, \theta_1), \quad \boldsymbol{f}(t, \boldsymbol{y}) = \boldsymbol{f}(\boldsymbol{y}) = (\, y_2\,, -k \sin y_1 - \alpha y_2)\,.$$

The system is autonomous, with $I = \mathbb{R}$ and $D = \mathbb{R}^2$.
(The example continues on p. 448.) □

10.3 Equations of first order

Before we undertake the systematic study of differential equations of order n, we examine a few types of first-order ODEs that stand out, in that they can be reduced to the computation of primitive maps.

10.3.1 Equations with separable variables

One speaks of separable variables (or of a separable ODE) for equations of the following sort

$$\boxed{y' = g(t)h(y),} \tag{10.14}$$

where g is continuous in t and h continuous in y. This means the function $f(t, y)$ is a product of a map depending only on t and a map depending only on y, so that the variables are 'separate'.

If $\bar{y} \in \mathbb{R}$ is a zero of h, $h(\bar{y}) = 0$, the constant map $y(t) = \bar{y}$ is a particular integral of (10.14), for the equation becomes $0 = 0$. Therefore a separable ODE has, first of all, as many integrals $y(t) = \mathit{constant}$ as the number of the distinct roots of h. These are called *singular integrals* of the equation.

On every interval J where $h(y)$ does not vanish we can write (10.14) as

$$\frac{1}{h(y)} \frac{\mathrm{d}y}{\mathrm{d}t} = g(t)\,.$$

Let $H(y)$ be a primitive of $\dfrac{1}{h(y)}$ (with respect to y). The formula for differentiating composite functions gives

$$\frac{\mathrm{d}}{\mathrm{d}t} H(y(t)) = \frac{dH}{dy} \frac{\mathrm{d}y}{\mathrm{d}t} = \frac{1}{h(y)} \frac{\mathrm{d}y}{\mathrm{d}t} = g(t)\,,$$

whence $H(y(t))$ is a primitive of $g(t)$. Thus, given any primitive $G(t)$ of $g(t)$ we will have

$$\boxed{H(y(t)) = G(t) + c\,, \qquad c \in \mathbb{R}\,.} \tag{10.15}$$

But since $\dfrac{1}{h(y)} = \dfrac{dH}{dy}$ never vanishes on J by assumption – and thus does not change sign, being continuous – the map $H(y)$ will be strictly monotone on J, hence invertible (Vol. I, Thm. 2.8). This implies we may solve (10.15) for $y(t)$, to get

$$\boxed{y(t) = H^{-1}(G(t) + c),} \tag{10.16}$$

where H^{-1} is the inverse of H. The above is the general integral of (10.14) on every interval where $h(y(t))$ is not zero. But, should we not be able to attain the analytic expression of $H^{-1}(x)$, formula (10.16) would have a paltry theoretical meaning. In such an event one is entitled to stop at the implicit form (10.15).

If equation (10.14) admits singular integrals, these might be of the form (10.16) for suitable constants c. Certain singular integrals can be inferred formally from (10.16) letting c tend to $\pm\infty$.

One recovers expression (10.15) in a formal and easy-to-remember manner by interpreting the derivative $\dfrac{\mathrm{d}y}{\mathrm{d}t}$ as a 'quotient', in Leibniz's notation. In fact, dividing (10.14) by $h(y)$ and 'multiplying' by $\mathrm{d}t$ gives

$$\frac{\mathrm{d}y}{h(y)} = g(t)\,\mathrm{d}t\,,$$

which can be then integrated

$$\boxed{\int \frac{\mathrm{d}y}{h(y)} = \int g(t)\,\mathrm{d}t.}$$

This corresponds exactly to (10.15). At any rate the reader must not forget that the correct proof of the formula is the aforementioned one!

Examples 10.2

i) Let us solve $y' = y(1-y)$. If we set $g(t) = 1$ and $h(y) = y(1-y)$, the zeroes of h determine two singular integrals $y_1(t) = 0$ and $y_2(t) = 1$. Next, assuming $h(y)$ different from 0, we rewrite the equation as

$$\int \frac{\mathrm{d}y}{y(1-y)} = \int \mathrm{d}t,$$

then integrate with respect to y on the left and t on the right to obtain

$$\log\left|\frac{y}{1-y}\right| = t + c$$

Exponentiating yields

$$\left|\frac{y}{1-y}\right| = \mathrm{e}^{t+c} = k\mathrm{e}^t,$$

where $k = \mathrm{e}^c$ is an arbitrary constant > 0. Therefore

$$\frac{y}{1-y} = \pm k\mathrm{e}^t = K\mathrm{e}^t,$$

with K being any non-zero constant. Solving now for y in terms of t gives

$$y(t) = \frac{K\mathrm{e}^t}{1 + K\mathrm{e}^t}\,.$$

In this case the singular integral $y_1(t) = 0$ belongs to the above family of solutions for $K = 0$, a value originally excluded. The other integral $y_2(t) = 1$ arises formally by taking the limit, for K going to infinity, in the general expression.

ii) Consider the ODE

$$y' = \sqrt{y}.$$

It has a singular integral $y_1(t) = 0$. Separating variables gives

$$\int \frac{\mathrm{d}y}{\sqrt{y}} = \int \mathrm{d}t\,,$$

so

$$2\sqrt{y} = t + c\,, \qquad \text{i.e.,} \qquad y(t) = \left(\frac{t}{2} + c\right)^2, \qquad c \in \mathbb{R}\,,$$

where we have written c in place of $c/2$.

iii) The differential equation

$$y' = \frac{\mathrm{e}^t + 1}{\mathrm{e}^y + 1}$$

has $g(t) = \mathrm{e}^t + 1$, $h(y) = \dfrac{1}{\mathrm{e}^y + 1} > 0$ for any y, so there are no singular integrals. The separating recipe gives

$$\int (\mathrm{e}^y + 1)\,\mathrm{d}y = \int (\mathrm{e}^t + 1)\,\mathrm{d}t\,,$$

so

$$\mathrm{e}^y + y = \mathrm{e}^t + t + c\,, \qquad c \in \mathbb{R}\,.$$

But now we are forced to stop for it is not possible to explicitly write y as function of the variable t. □

10.3.2 Homogeneous equations

Homogeneous are ODEs of the type

$$\boxed{y' = \varphi\left(\frac{y}{t}\right)} \tag{10.17}$$

with $\varphi = \varphi(z)$ continuous in z. The map $f(t, y)$ depends on t and y in terms of their ratio $\dfrac{y}{t}$ only; equivalently, $f(\lambda t, \lambda y) = f(t, y)$ for any $\lambda \neq 0$.

A homogeneous equation can be transformed into one with separable variables by the obvious substitution $z = \dfrac{y}{t}$, understood as $z(t) = \dfrac{y(t)}{t}$. Then $y(t) = tz(t)$ and $y'(t) = z(t) + tz'(t)$, so (10.17) becomes

$$z' = \frac{\varphi(z) - z}{t},$$

which has separable variables z and t, as required. The technique of Section 10.3.1 solves it: each solution $\bar{z}$ of $\varphi(z) = z$ gives rise to a singular integral $z(t) = \bar{z}$, hence $y(t) = \bar{z}t$. Assuming $\varphi(z)$ different from z, instead, we have

$$\boxed{\int \frac{\mathrm{d}z}{\varphi(z) - z} = \int \frac{\mathrm{d}t}{t},}$$

giving

$$H(z) = \log|t| + c$$

where $H(z)$ is a primitive of $\dfrac{1}{\varphi(z) - z}$. Indicating by H^{-1} the inverse function, we have

$$z(t) = H^{-1}(\log|t| + c),$$

hence, returning to y, the general integral of (10.17) reads

$$\boxed{y(t) = t\,H^{-1}(\log|t| + c).}$$

Example 10.3

We solve

$$t^2 y' = y^2 + ty + t^2. \tag{10.18}$$

In normal form, this is

$$y' = \left(\frac{y}{t}\right)^2 + \frac{y}{t} + 1,$$

a homogeneous equation with $\varphi(z) = z^2 + z + 1$. The substitution $y = tz$ generates a separable equation

$$z' = \frac{z^2 + 1}{t}.$$

There are no singular integrals, for $z^2 + 1$ is always positive. Integration gives

$$\arctan z = \log|t| + c$$

so the general integral of (10.18) is

$$y(t) = t\tan(\log|t| + c).$$

The constant c can be chosen independently in $(-\infty, 0)$ and $(0, +\infty)$, because of the singularity at $t = 0$. Notice also that the domain of each solution depends on the value of c. □

10.3.3 Linear equations

The differential equation

$$\boxed{y' = a(t)y + b(t)} \tag{10.19}$$

with a, b continuous on I, is called linear.[2] The map $f(t,y) = a(t)y + b(t)$ is a polynomial of degree one in y with coefficients dependent on t. The equation is called **homogeneous** if the source term vanishes identically, $b(t) = 0$, **non-homogeneous** otherwise.

Let us start by solving the homogeneous equation

$$y' = a(t)y. \tag{10.20}$$

It is a special case of equation with separable variables where $g(t) = a(t)$ and $h(y) = y$, see (10.14). Therefore the constant map $y(t) = 0$ is a solution. Apart from that, we can separate variables

$$\int \frac{1}{y}\,\mathrm{d}y = \int a(t)\,\mathrm{d}t.$$

If $A(t)$ indicates a primitive of $a(t)$, i.e., if

$$\int a(t)\,\mathrm{d}t = A(t) + c, \qquad c \in \mathbb{R}, \tag{10.21}$$

then

$$\log|y| = A(t) + c\,,$$

or equivalently

$$|y(t)| = \mathrm{e}^c \mathrm{e}^{A(t)}\,,$$

so

$$y(t) = \pm K \mathrm{e}^{A(t)},$$

where $K = \mathrm{e}^c > 0$. The particular solution $y(t) = 0$ is subsumed by the general formula if we let K be 0. Therefore the solutions of the homogeneous linear equation (10.20) are given by

$$y(t) = K\mathrm{e}^{A(t)}, \qquad K \in \mathbb{R},$$

with $A(t)$ as in (10.21).

Now we tackle the non-homogeneous case. We use the so-called *method of variation of constants*, or *parameters*, which consists in searching for solutions of the form

$$y(t) = K(t)\,\mathrm{e}^{A(t)},$$

where now $K(t)$ is a function to be determined. The representation of $y(t)$ is always possible, for $\mathrm{e}^{A(t)} > 0$. Substituting into (10.19), we obtain

$$K'(t)\mathrm{e}^{A(t)} + K(t)\mathrm{e}^{A(t)}a(t) = a(t)K(t)\mathrm{e}^{A(t)} + b(t),$$

so $\qquad K'(t) = \mathrm{e}^{-A(t)}b(t).$

[2] Contrary to the description of Vol. I, here we prefer to write the linear term y on the right, to be consistent with the theory of linear systems.

Calling $B(t)$ a primitive of $\mathrm{e}^{-A(t)}b(t)$,

$$\int \mathrm{e}^{-A(t)}b(t)\,\mathrm{d}t = B(t) + c, \qquad c \in \mathbb{R}, \tag{10.22}$$

we have

$$K(t) = B(t) + c,$$

so the general solution to (10.19) is

$$\boxed{y(t) = \mathrm{e}^{A(t)}\big(B(t) + c\big),} \tag{10.23}$$

with $A(t)$ and $B(t)$ defined by (10.21) and (10.22). The integral is more often than not found in the form

$$\boxed{y(t) = \mathrm{e}^{\int a(t)\,\mathrm{d}t} \int \mathrm{e}^{-\int a(t)\,\mathrm{d}t}\, b(t)\,\mathrm{d}t,} \tag{10.24}$$

where the steps leading to the solution are clearly spelt out, namely: one has to integrate twice in succession.

If we want to solve the Cauchy problem

$$\begin{cases} y' = a(t)y + b(t) & \text{on the interval } I, \\ y(t_0) = y_0, & \text{with } t_0 \in I \text{ and } y_0 \in \mathbb{R}, \end{cases} \tag{10.25}$$

it might be convenient to choose as primitive of $a(t)$ the one vanishing at t_0, which we write $A(t) = \displaystyle\int_{t_0}^{t} a(s)\,\mathrm{d}s$, according to the Fundamental Theorem of Integral Calculus; the same can be done for $B(t)$ by putting

$$B(t) = \int_{t_0}^{t} \mathrm{e}^{-\int_{t_0}^{s} a(u)\,\mathrm{d}u}\, b(s)\,\mathrm{d}s$$

(recall that the variables in the definite integral are arbitrary symbols). Using these expressions for $A(t)$ and $B(t)$ in (10.23) we obtain $y(t_0) = c$, hence the solution of the Cauchy problem (10.25) will satisfy $c = y_0$, i.e.,

$$\boxed{y(t) = \mathrm{e}^{\int_{t_0}^{t} a(u)\,\mathrm{d}u} \left(y_0 + \int_{t_0}^{t} \mathrm{e}^{-\int_{t_0}^{s} a(u)\,\mathrm{d}u}\, b(s)\,\mathrm{d}s \right).} \tag{10.26}$$

Examples 10.4

i) Find the general integral of the linear equation

$$y' = ay + b,$$

where $a \neq 0$ and b are real numbers.

Choosing $A(t) = at$ and $B(t) = -\frac{b}{a}\mathrm{e}^{-at}$ generates the integral

$$y(t) = c\mathrm{e}^{at} - \frac{b}{a}.$$

If $a = 1$ and $b = 0$, the formula shows that all solutions of $y' = y$ come in the form $y(t) = c\mathrm{e}^{t}$.

If we want to solve the Cauchy problem

$$\begin{cases} y' = ay + b & \text{on } [1, +\infty), \\ y(1) = y_0, \end{cases}$$

it is better to choose $A(t) = a(t-1)$ and $B(t) = \frac{b}{a}\left(1 - \mathrm{e}^{-a(t-1)}\right)$, so that

$$y(t) = \left(y_0 + \frac{b}{a}\right)\mathrm{e}^{a(t-1)} - \frac{b}{a}.$$

In case $a < 0$, the solution converges to $-b/a$ (independent of the initial datum y_0) as $t \to +\infty$.

ii) We find the integral curves of the ODE

$$ty' + y = t^2$$

that lie in the first quadrant of the plane (t, y). Written as (10.19), the equation is

$$y' = -\frac{1}{t}y + t,$$

so $a(t) = -1/t$, $b(t) = t$. Choose $A(t) = -\log t$, and then $\mathrm{e}^{A(t)} = 1/t$, $\mathrm{e}^{-A(t)} = t$; consequently,

$$\int \mathrm{e}^{-A(t)} b(t)\,\mathrm{d}t = \int t^2\,\mathrm{d}t = \frac{1}{3}t^3 + c.$$

Therefore for $t > 0$, the general integral reads

$$y(t) = \frac{1}{t}\left(\frac{1}{3}t^3 + c\right) = \frac{1}{3}t^2 + \frac{c}{t}.$$

If $c \geq 0$, then $y(t) > 0$ for any $t > 0$, while if $c < 0$, $y(t) > 0$ when $t > \sqrt[3]{3|c|}$. □

Remark 10.5 In the sequel it will turn out useful to consider linear equations of the form

$$\boxed{z' = \lambda z\,,} \tag{10.27}$$

where λ is a complex number. We should determine the solution over the complex field: this will be a differentiable map $z = \mathcal{R}e\, z + i\,\mathcal{I}m\, z : \mathbb{R} \to \mathbb{C}$ (i.e., its real and imaginary parts are differentiable). Recall that the equation

$$\boxed{\frac{\mathrm{d}}{\mathrm{d}t}\,\mathrm{e}^{\lambda t} = \lambda\,\mathrm{e}^{\lambda t}\,, \qquad t \in \mathbb{R}} \tag{10.28}$$

(proved in Vol. I, Eq. (11.30)) is still valid for $\lambda \in \mathbb{C}$. Then we can say that any solution of (10.27) can be written as

$$\boxed{z(t) = c\,\mathrm{e}^{\lambda t}} \tag{10.29}$$

with $c \in \mathbb{C}$ arbitrary. □

10.3.4 Bernoulli equations

This family of equations is characterised by the following form

$$\boxed{y' = p(t)y^{\alpha} + q(t)y\,,\ \alpha \neq 0,\ \alpha \neq 1\,,} \tag{10.30}$$

with p, q continuous on I. If $\alpha > 0$, the constant map 0 is a solution. Supposing then $y \neq 0$ and dividing by y^{α} gives

$$y^{-\alpha}y' = p(t) + q(t)y^{1-\alpha}\,.$$

As $(y^{1-\alpha})' = (1-\alpha)y^{-\alpha}y'$, the substitution $z = y^{1-\alpha}$ transforms the equation into a linear equation in z

$$\boxed{z' = (1-\alpha)p(t) + (1-\alpha)q(t)z\,,}$$

solvable by earlier methods.

Example 10.6

Take $y' = t^3y^2 + 2ty$, a Bernoulli equation where $p(t) = t^3$, $q(t) = 2t$ and $\alpha = 2$. The transformation suggested above reduces the equation to $z' = -(2tz + t^3)$, solved by $z(t) = c\mathrm{e}^{t^2} - (2 + t^2)$. Hence

$$y(t) = \frac{1}{c\mathrm{e}^{t^2} - (2+t^2)}$$

solves the original equation. □

10.3.5 Riccati equations

Equations of Riccati type crop up in optimal control problems, and have the typical form

$$\boxed{y' = p(t)y^2 + q(t)y + r(t)\,,} \tag{10.31}$$

where p, q, r are continuous on I. The general integral can be found provided we know a particular integral $y = u(t)$. In fact, putting

$$y = u(t) + \frac{1}{z}\,, \quad \text{hence} \quad y' = u'(t) - \frac{z'}{z^2}\,,$$

we obtain

$$u'(t) - \frac{z'}{z^2} = p(t)\left(u^2(t) + 2\frac{u(t)}{z} + \frac{1}{z^2}\right) + q(t)\big(u(t) + \frac{1}{z}\big) + r(t)\,.$$

Since u is a solution, this simplifies to

$$\boxed{z' = -\big(2u(t)p(t) + q(t)\big)z - p(t)\,,}$$

and once again we recover a linear equation in the unknown z.

Example 10.7

Consider the Riccati equation

$$y' = ty^2 + \left(\frac{1-2t^2}{t}\right)y + \frac{t^2-1}{t}\,,$$

where

$$p(t) = t\,, \qquad q(t) = \frac{1-2t^2}{t}\,, \qquad r(t) = \frac{t^2-1}{t}\,.$$

There is a particular solution $u(t) = 1$. The aforementioned change of variables gives

$$z' = -\frac{z}{t} - t\,,$$

solved by

$$z(t) = \frac{c-t^3}{3|t|}.$$

It follows the solution is

$$y(t) = 1 + \frac{3|t|}{c-t^3}\,, \qquad c \in \mathbb{R}\,.$$

□

10.3.6 Second-order equations reducible to first order

When a differential equation of order two does not explicitly contain the dependent variable, as in

$$\boxed{y'' = f(t, y')\,,} \tag{10.32}$$

the substitution $z = y'$ leads to the equation of order one

$$z' = f(t, z)$$

in $z = z(t)$. If the latter has a general integral $z(t; c_1)$, all solutions of (10.32) arise from

$$y' = z\,,$$

hence from the primitives of $z(t;c_1)$. This introduces another integration constant c_2. The general integral of (10.32) is thus

$$\boxed{y(t;c_1,c_2) = \int z(t;c_1)\,\mathrm{d}t = Z(t;c_1) + c_2,}$$

where $Z(t;c_1)$ is a particular primitive of $z(t;c_1)$.

Example 10.8

Solve the second-order equation

$$y'' - (y')^2 = 1.$$

Set $z = y'$ to obtain the separable ODE

$$z' = z^2 + 1.$$

The general integral is $\arctan z = t + c_1$, so

$$z(t,c_1) = \tan(t + c_1).$$

Integrating again gives

$$y(t;c_1,c_2) = \int \tan(t + c_1)\,\mathrm{d}t = -\log\big(\cos(t + c_1)\big) + c_2\,, \qquad c_1, c_2 \in \mathbb{R}\,. \qquad \square$$

Consider, at last, a second-order autonomous equation

$$\boxed{y'' = f(y, y')\,.} \tag{10.33}$$

We shall indicate a method for each interval $I \subseteq \mathbb{R}$ where $y'(t)$ has constant sign. In such case $y(t)$ is strictly monotone, so invertible; we may consider therefore $t = t(y)$ as function of y on $J = y(I)$, and consequently everything will depend on y as well; in particular, $z = \frac{\mathrm{d}y}{\mathrm{d}t}$ should be thought of as map in y. Then

$$y'' = \frac{\mathrm{d}z}{\mathrm{d}t} = \frac{\mathrm{d}z}{\mathrm{d}y}\frac{\mathrm{d}y}{\mathrm{d}t} = \frac{\mathrm{d}z}{\mathrm{d}y}z\,,$$

whereby $y'' = f(y, z)$ becomes

$$\boxed{\frac{\mathrm{d}z}{\mathrm{d}y} = g(y, z) = \frac{1}{z}f(y, z)\,.}$$

This is of order one, y being the independent variable and z the unknown. Supposing we are capable of finding all its solutions $z = z(y;c_1)$, the solutions $y(t)$ arise from the autonomous equation of order one

$$\boxed{\frac{\mathrm{d}y}{\mathrm{d}t} = z(y;c_1)\,,}$$

coming from the definition of t.

Example 10.9

Find the solutions of

$$yy'' = (y')^2$$

i.e.,

$$y'' = \frac{(y')^2}{y}\,, \qquad (y \neq 0)\,.$$

Set $f(y,z) = \frac{z^2}{y}$, and solve

$$\frac{\mathrm{d}z}{\mathrm{d}y} = g(y,z) = \frac{z}{y}\,.$$

This gives $z(y) = c_1 y$, and then

$$\frac{\mathrm{d}y}{\mathrm{d}t} = c_1 y\,,$$

whence $y(t) = c_2 \mathrm{e}^{c_1 t}$, $c_2 \neq 0$. □

10.4 The Cauchy problem

We return to differential equations in their most general form (10.5) and discuss the solvability of the Cauchy problem.

As in Sect. 10.2, let $I \subseteq \mathbb{R}$ be an open interval and $D \subseteq \mathbb{R}^n$ an open connected set. Call $\Omega = I \times D \subseteq \mathbb{R}^{n+1}$, and take a map $\boldsymbol{f} : \Omega \to \mathbb{R}^n$ defined on Ω. Given points t_0 in I and $\boldsymbol{y}_0$ in D, we examine the Cauchy problem

$$\boxed{\begin{cases} \boldsymbol{y}' = \boldsymbol{f}(t,\boldsymbol{y})\,, \\ \boldsymbol{y}(t_0) = \boldsymbol{y}_0\,. \end{cases}} \tag{10.34}$$

10.4.1 Local existence and uniqueness

Step one of our study is to determine conditions on $\boldsymbol{f}$ that guarantee the Cauchy problem can be solved locally, by which we mean on a neighbourhood of t_0.

It is remarkable that the continuity of $\boldsymbol{f}$ alone, in t and $\boldsymbol{y}$ simultaneously, is enough to ensure the existence of a solution, as the next result, known as Peano's Existence Theorem, proves.

Theorem 10.10 (Peano) *Suppose $\boldsymbol{f}$ is continuous on Ω. Then there exist a closed neighbourhood $[t_0-\alpha, t_0+\alpha]$ of t_0 and a map $\boldsymbol{y} : [t_0-\alpha, t_0+\alpha] \to D$, differentiable with continuity, that solves the Cauchy problem* (10.34).

There is a way of estimating the size of the interval where the solution exists. Let $\overline{B_a(t_0) \text{ and } B_b(\boldsymbol{y}_0)}$ be neighbourhoods of t_0 and $\boldsymbol{y}_0$ such that the compact set $K = \overline{B_a(t_0) \times B_b(\boldsymbol{y}_0)}$ is contained in Ω, and set $M = \max\limits_{(t,\boldsymbol{y})\in K} \|\boldsymbol{f}(t,\boldsymbol{y})\|$. Then the theorem holds with $\alpha = \min(a, b/M)$.

Peano's Theorem does not claim anything on the solution's uniqueness, and in fact the mere continuity of $\boldsymbol{f}$ does not prevent to have infinitely many solutions to (10.34).

Example 10.11

The initial value problem

$$\begin{cases} y' = \frac{3}{2}\sqrt[3]{y}\,, \\ y(0) = 0\,, \end{cases}$$

solvable by separating variables, admits solutions $y(t) = 0$ (the singular integral) and $y(t) = \sqrt{t^3}$; actually, there are infinitely many solutions, some of which

$$y(t) = y_\alpha(t) = \begin{cases} 0 & \text{if } t \le \alpha\,, \\ \sqrt{(t-\alpha)^3} & \text{if } t > \alpha\,, \end{cases} \qquad \alpha \ge 0\,,$$

are obtained by 'glueing' the previous two in a suitable way (Fig. 10.4). □

That is why we need to add further hypotheses in order to grant uniqueness, besides requiring the solution depend continuously on the initial datum.

A very common assumption is this one:

Definition 10.12 *A map* $\boldsymbol{f}$ *defined on* $\Omega = I \times D$ *is said* **Lipschitz in** $\boldsymbol{y}$**, uniformly in** t**, over** Ω*, if there exists a constant* $L \ge 0$ *such that*

$$\|\boldsymbol{f}(t, \boldsymbol{y}_1) - \boldsymbol{f}(t, \boldsymbol{y}_2)\| \le L\|\boldsymbol{y}_1 - \boldsymbol{y}_2\|\,, \qquad \forall \boldsymbol{y}_1, \boldsymbol{y}_2 \in D,\ \forall t \in I\,. \tag{10.35}$$

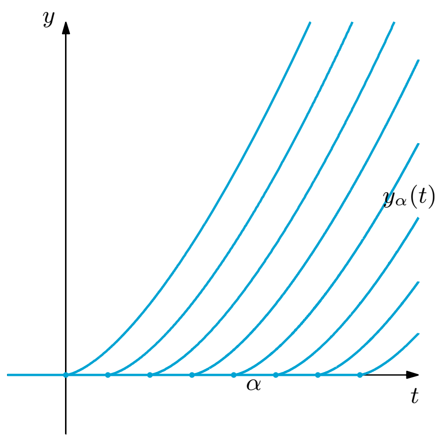


Figure 10.4. The infinitely many solutions of the Cauchy problem relative to Example 10.11

Using Proposition 6.4 for every $t \in I$, it is easy to see that a map is Lipschitz if all its components admit bounded y_j-derivatives on Ω:

$$\sup_{(t,\boldsymbol{y})\in\Omega} \left|\frac{\partial f_i}{\partial y_j}(t,\boldsymbol{y})\right| < +\infty\,, \qquad 1 \le i,j \le n\,.$$

Examples 10.13

i) Let us consider the function $f(t,y) = y^2$, defined and continuous on $\mathbb{R}\times\mathbb{R}$. Then $\forall y_1, y_2 \in \mathbb{R}$ and $\forall t \in \mathbb{R}$,

$$|f(t,y_1) = f(t,y_2)| = |y_1^2 - y_2^2| = |y_1 + y_2|\,|y_1 - y_2|\,.$$

Hence f is Lipschitz in the variable y, uniformly in t, on every domain $\Omega = \Omega_R = \mathbb{R}\times D_R$ with $D_R = (-R,R)$, $R > 0$; the corresponding Lipschitz constant $L = L_R$ equals $L_R = 2R$. Note that the function is *not* Lipschitz on $\mathbb{R}\times\mathbb{R}$, because $|y_1 + y_2|$ tends to $+\infty$ as y_1 and y_2 tend to infinity with the same sign.

ii) Consider now the map $f(t,y) = \sin(ty)$, defined and continuous on $\mathbb{R}\times\mathbb{R}$. It satisfies

$$|\sin(ty_1) - \sin(ty_2)| \le |(ty_1) - (ty_2)| = |t|\,|y_1 - y_2|\,,$$

$\forall y_1, y_2 \in \mathbb{R}$, $\forall t \in \mathbb{R}$, because $\theta \mapsto \sin\theta$ is Lipschitz on $\mathbb{R}$ with constant 1. Therefore f is Lipschitz in y, uniformly in t, on every region $\Omega = \Omega_R = I_R \times \mathbb{R}$, with $I_R = (-R,R)$, $R > 0$, with Lipschitz constant $L = L_R = R$. Again, f is *not* Lipschitz on the whole $\mathbb{R}\times\mathbb{R}$.

iii) Finally, let us consider the affine map $\boldsymbol{f}(t,\boldsymbol{y}) = \boldsymbol{A}(t)\boldsymbol{y} + \boldsymbol{b}(t)$, where $\boldsymbol{A}(t) \in \mathbb{R}^{n,n}$, $\boldsymbol{b}(t) \in \mathbb{R}^n$ are defined and continuous on the open interval $I \subseteq \mathbb{R}$. Then $\boldsymbol{f}$ is defined and continuous on $I \times \mathbb{R}^n$. For any $\boldsymbol{y}_1, \boldsymbol{y}_2 \in \mathbb{R}^n$ and any $t \in I$, we have

$$\begin{aligned}\|\boldsymbol{f}(t,\boldsymbol{y}_1) - \boldsymbol{f}(t,\boldsymbol{y}_2)\| &= \|\boldsymbol{A}(t)\boldsymbol{y}_1 - \boldsymbol{A}(t)\boldsymbol{y}_2\| \\ &= \|\boldsymbol{A}(t)(\boldsymbol{y}_1 - \boldsymbol{y}_2)\| \le \|\boldsymbol{A}(t)\|\,\|\boldsymbol{y}_1 - \boldsymbol{y}_2\|\,,\end{aligned}$$

where $\|\boldsymbol{A}(t)\|$ is the norm of the matrix $\boldsymbol{A}(t)$ defined in (4.9), and the inequality follows from (4.10). Observe that the map $\alpha(t) = \|\boldsymbol{A}(t)\|$ is continuous on I (as composite of continuous maps). If α is bounded, $\boldsymbol{f}$ is Lipschitz in $\boldsymbol{y}$, uniformly in t on $\Omega = I \times \mathbb{R}$, with Lipschitz constant $L = \sup_{t\in I}\alpha(t)$. If not, we can at least say α is bounded on every closed and bounded interval $J \subset I$ (Weierstrass' Theorem), so $\boldsymbol{f}$ is Lipschitz on every $\Omega = J \times \mathbb{R}^n$. □

The next result is of paramount importance. It is known as the Theorem of Cauchy-Lipschitz (or Picard-Lindelöf).

Theorem 10.14 (Cauchy-Lipschitz) *Let $\boldsymbol{f}(t,\boldsymbol{y})$ be continuous on Ω and Lipschitz in $\boldsymbol{y}$, uniformly in t. Then the Cauchy problem* (10.34) *admits one, and only one, solution $\boldsymbol{y} = \boldsymbol{y}(t)$ defined on a closed neighbourhood $[t_0-\alpha, t_0+\alpha]$ of t_0, with values in Ω, and differentiable with continuity on Ω.*

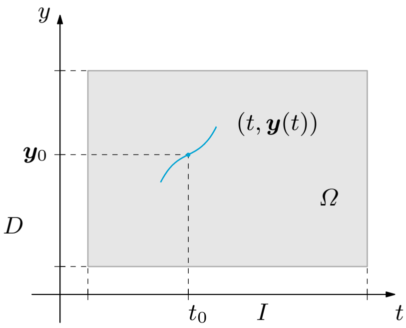


Figure 10.5. Local existence and uniqueness of a solution

The solution's uniqueness and its dependency upon the initial datum are consequences of the following property.

Proposition 10.15 *Under the assumptions of the previous theorem, let* $\boldsymbol{y}$, $\boldsymbol{z}$ *be solutions of the Cauchy problems*

$$\begin{cases} \boldsymbol{y}' = \boldsymbol{f}(t, \boldsymbol{y}), \\ \boldsymbol{y}(t_0) = \boldsymbol{y}_0, \end{cases} \qquad \text{and} \qquad \begin{cases} \boldsymbol{z}' = \boldsymbol{f}(t, \boldsymbol{z}), \\ \boldsymbol{z}(t_0) = \boldsymbol{z}_0, \end{cases}$$

over an interval J *containing* t_0, *for given* $\boldsymbol{y}_0$ *and* $\boldsymbol{z}_0 \in D$. *Then*

$$\|\boldsymbol{y}(t) - \boldsymbol{z}(t)\| \le \mathrm{e}^{L|t-t_0|} \|\boldsymbol{y}_0 - \boldsymbol{z}_0\|, \qquad \forall t \in J. \tag{10.36}$$

In particular, if $\boldsymbol{z}_0 = \boldsymbol{y}_0$, $\boldsymbol{z}$ *coincides with* $\boldsymbol{y}$ *on* J.

What the property is saying is that the solution of (10.34) depends in a continuous way upon the initial datum $\boldsymbol{y}_0$: a small deformation ε of the initial datum perturbs the solution at $t \neq t_0$ by $\mathrm{e}^{L|t-t_0|}\varepsilon$ at most. In other words, the distance between two orbits at time t grows by a factor not larger than $\mathrm{e}^{L|t-t_0|}$. Since this factor is exponential, its magnitude depends on the distance $|t - t_0|$ as well as on the Lipschitz constant of $\boldsymbol{f}$.

For certain equations it is possible to replace $\mathrm{e}^{L|t-t_0|}$ with $\mathrm{e}^{\sigma(t-t_0)}$ if $t > t_0$, or $\mathrm{e}^{\sigma(t_0-t)}$ if $t < t_0$, with $\sigma < 0$ (see Example 10.22 ii) and Sect. 10.8.1). In these cases the solutions move towards one another exponentially.

Remark 10.16 Assume that a mathematical problem is formalised as

$$P(s) = d$$

where d is the datum and s the solution. Whenever s

- exists for any datum d,
- is unique, and

- depends upon d with continuity,

the problem is said to be **well posed** (**in the sense of Hadamard**).

Theorem 10.14 and the subsequent proposition guarantee the Cauchy problem (10.34) is well posed with the assumptions made on $\boldsymbol{f}$. □

From the Theorem of Cauchy-Lipschitz we infer important consequences of qualitative nature regarding the solution set of the differential equation. To be precise, its hypotheses imply:

- two integral curves with a common point must necessarily coincide (by uniqueness of the solution on the neighbourhood of every common point);
- for autonomous systems in particular, the orbits $\Gamma(\boldsymbol{y})$, as $\boldsymbol{y}$ varies among solutions, form a disjoint *partition* of phase space; it can be proved that an orbit is closed if and only if the corresponding solution is periodic in time, $\boldsymbol{y}(t+T) = \boldsymbol{y}(t)$, for all $t \in \mathbb{R}$ and a suitable $T > 0$;
- the solutions on Ω of the ODE depend on n real parameters, as anticipated in Sect. 10.2: given any $t_0 \in I$, there is a distinct solution for every choice of datum $\boldsymbol{y}_0$ on the open set D, and the coordinates $c_1 = y_{01}, \ldots, c_n = y_{0n}$ of $\boldsymbol{y}_0$ should be considered as the free parameters upon which the solution depends.

10.4.2 Maximal solutions

The Theorem of Cauchy-Lipschitz ensures the existence of a closed interval $[t_0 - \alpha, t_0 + \alpha]$, containing t_0, where the problem (10.34) is solvable; we shall call $\boldsymbol{u} = \boldsymbol{u}(t)$ the corresponding solution henceforth. However, this interval might not be the largest interval containing t_0 where the problem can be solved (the *maximal interval* of existence). In fact, suppose $t_1 = t_0 + \alpha$ lies inside I and $\boldsymbol{y}_1 = \boldsymbol{u}(t_1)$ is inside D. Then we can re-ignite the Cauchy problem at t_1

$$\begin{cases} \boldsymbol{y}' = \boldsymbol{f}(t, \boldsymbol{y}) , \\ \boldsymbol{y}(t_1) = \boldsymbol{y}_1 . \end{cases} \tag{10.37}$$

Because the assumptions of Theorem 10.14 are still valid, the new problem has a unique solution $\boldsymbol{v} = \boldsymbol{v}(t)$ defined on an interval $[t_1 - \beta, t_1 + \beta] \subseteq I$. Both functions $\boldsymbol{u}$, $\boldsymbol{v}$ solve the ODE and satisfy the condition $\boldsymbol{u}(t_1) = \boldsymbol{y}_1 = \boldsymbol{v}(t_1)$ on the interval $J = [t_0, t_1] \cap [t_1 - \beta, t_1]$, which contains a left neighbourhood of t_1 (see Fig. 10.6). Therefore they must solve problem (10.37) on J. We can use inequality (10.36) with t_0 replaced by t_1, to obtain

$$\|\boldsymbol{u}(t) - \boldsymbol{v}(t)\| \le \mathrm{e}^{L|t_1 - t_0|} \|\boldsymbol{y}_1 - \boldsymbol{y}_1\| = 0 , \qquad \forall t \in J ;$$

that shows $\boldsymbol{v}$ must coincide with $\boldsymbol{u}$ on J. Since $\boldsymbol{v}$ is also defined on the right of t_1, we may consider the prolongation

$$\boxed{\tilde{\boldsymbol{u}}(t) = \begin{cases} \boldsymbol{u}(t) & \text{in } [t_0 - \alpha, t_0 + \alpha] , \\ \boldsymbol{v}(t) & \text{in } [t_0 + \alpha, t_0 + \alpha + \beta] , \end{cases}}$$

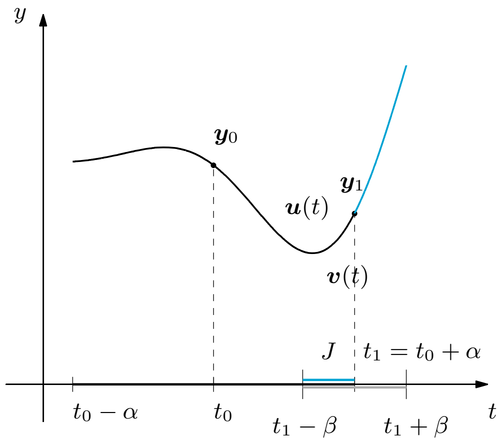


Figure 10.6. Prolongation of a solution beyond t_1

that extends the previous solution to the interval $[t_0 - \alpha, t_0 + \alpha + \beta]$. A similar prolongation to the left of $t_0 - \alpha$ is possible if $t_0 - \alpha$ is inside I and $\boldsymbol{u}(t_0 - \alpha)$ still lies in D.

The procedure thus described may be iterated so that to prolong $\boldsymbol{u}$ to an open interval $J_{\max} = (t_-, t_+)$, where, by definition:

$$\boxed{t_- = \inf\{t_* \in I : \text{problem (10.34) has a solution on } [t_*, t_0]\}}$$

and

$$\boxed{t_+ = \sup\{t^* \in I : \text{problem (10.34) has a solution on } [t_0, t^*]\}\,.}$$

This determines precisely the maximal interval on which a solution $\boldsymbol{u}$ to problem (10.34) lives.

Suppose t_+ is strictly less than the least upper bound of I (so that the solution cannot be extended beyond t_+, despite $\boldsymbol{f}$ is continuous and Lipschitz in $\boldsymbol{y}$). Then

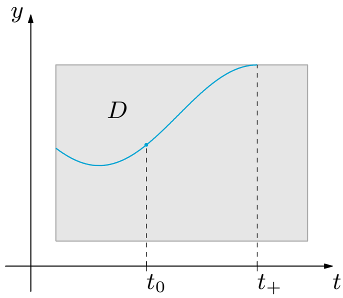


Figure 10.7. Right limit t_+ of the maximal interval $J_{\max}$ for the solution of a Cauchy problem

one could prove $\boldsymbol{u}(t)$ moves towards the boundary ∂D as t tends to t_+ (from the left), see Fig. 10.7. Similar considerations hold for t_-.

Example 10.17

Consider the Cauchy problem

$$\begin{cases} y' = y^2\,, \\ y(0) = y_0\,. \end{cases}$$

The Lipschitz character of $f(t,y) = y^2$ was discussed in Example (10.13) i): for any $R > 0$ we know f is Lipschitz on $\Omega_R = \mathbb{R} \times D_R$, $D_R = (-R, R)$. The problem's solution is

$$y(t) = \frac{y_0}{1 - y_0 t}\,.$$

If $y_0 \in D_R$, it is not hard to check that the maximal interval of existence in D_R, i.e., the set of t for which $y(t) \in D_R$, is

$$J_{\max} = \begin{cases} \left(-\infty, \dfrac{1}{y_0} - \dfrac{1}{R}\right) & \text{if } y_0 > 0\,, \\ (-\infty, +\infty) & \text{if } y_0 = 0\,, \\ \left(\dfrac{1}{y_0} + \dfrac{1}{R}, +\infty\right) & \text{if } y_0 < 0\,. \end{cases}$$

Notice that $\lim_{t \to t_+} y(t) = R$ when $y_0 > 0$, confirming that the solution reaches the boundary of D_R as $t \to t_+$. □

10.4.3 Global existence

We wish to find conditions on $\boldsymbol{f}$ that force every solution of (10.34) to be defined on the whole interval I. A solution for which $J_{\max} = I$ is said to exist globally, and called a **global solution**.

The petty example of an autonomous system in dimension 1,

$$y' = f(y)\,,$$

is already sufficient to show the various possibilities. The map $f(t,y) = f(y) = ay + b$ is Lipschitz on all $\mathbb{R} \times \mathbb{R}$, and the corresponding ODE's solutions

$$y(t) = c\mathrm{e}^{at} - \frac{b}{a}$$

are defined on $\mathbb{R}$ (see Example 10.4 i)). In contrast, $f(t,y) = f(y) = y^2$ is not Lipschitz on the entire $\mathbb{R}^2$ (Example 10.13 i)) and its non-zero solutions exist on semi-bounded intervals (Example 10.17). The linear example is justified by the following sufficient condition for global existence.

Theorem 10.18 *Let $\Omega = I \times \mathbb{R}^n$ and assume Theorem 10.14 holds. Then the solution of the Cauchy problem* (10.34) *is defined everywhere on I.*

This result is coherent with the fact, remarked earlier, that if the supremum t_+ of $J_{\max}$ is not the supremum of I, the solution converges to the boundary of D, as $t \to t_+$ (and analogously for t_-). In the present situation $D = \mathbb{R}^n$ has empty boundary, and therefore the end-points of $J_{\max}$ must coincide with those of I.

Being globally Lipschitz imposes a severe limitation on the behaviour of $\boldsymbol{f}$ as $\boldsymbol{y} \to \infty$: $\boldsymbol{f}$ *can grow, but not more than linearly.* In fact if we choose $\boldsymbol{y}_1 = \boldsymbol{y} \in \mathbb{R}^n$ arbitrarily, and $\boldsymbol{y}_2 = \mathbf{0}$ in (10.35), we deduce

$$\|\boldsymbol{f}(t,\boldsymbol{y}) - \boldsymbol{f}(t,\mathbf{0})\| \leq L\|\boldsymbol{y}\|\,, \qquad \forall \boldsymbol{y} \in \mathbb{R}^n\,.$$

The triangle inequality then yields

$$\begin{aligned}\|\boldsymbol{f}(t,\boldsymbol{y})\| &= \|\boldsymbol{f}(t,\mathbf{0}) + \boldsymbol{f}(t,\boldsymbol{y}) - \boldsymbol{f}(t,\mathbf{0})\| \\ &\leq \|\boldsymbol{f}(t,\mathbf{0})\| + \|\boldsymbol{f}(t,\boldsymbol{y}) - \boldsymbol{f}(t,\mathbf{0})\|\,,\end{aligned}$$

so necessarily

$$\boxed{\|\boldsymbol{f}(t,\boldsymbol{y})\| \leq \|\boldsymbol{f}(t,\mathbf{0})\| + L\|\boldsymbol{y}\|\,, \qquad \boldsymbol{y} \in \mathbb{R}^n\,.}$$

Observe that $\boldsymbol{f}$ continuous on Ω forces the map $\beta(t) = \|\boldsymbol{f}(t,\mathbf{0})\|$ to be continuous on I.

Consider the autonomous equation

$$y' = y\log^2 y\,,$$

for which $f(t,y) = f(y) = y\log^2 y$ grows slightly more than linearly as $y \to +\infty$. The solutions

$$y(t) = \mathrm{e}^{-1/(x+c)}$$

are defined over not all of $\mathbb{R}$, showing that a growth that is at most linear is close to being optimal. Still, we can achieve some level of generalisation in this direction.

We may attain the same result as Theorem 10.18, in fact, imposing that $\boldsymbol{f}$ grows in $\boldsymbol{y}$ at most linearly, and weakening the Lipschitz condition on $I \times \mathbb{R}^n$; precisely, it is enough to have $\boldsymbol{f}$ **locally Lipschitz** (in $\boldsymbol{y}$) over Ω. Equivalently, $\boldsymbol{f}$ is Lipschitz in $\boldsymbol{y}$, uniformly in t, *on every compact set K in Ω*; if so, the Lipschitz constant L_K may depend on K. All maps considered in Examples 10.13 are indeed locally Lipschitz (on $\mathbb{R} \times \mathbb{R}$ for the first two examples, on $I \times \mathbb{R}^n$ for the last).

Theorem 10.19 *Let $\boldsymbol{f}$ be continuous and locally Lipschitz in $\boldsymbol{y}$, uniformly in t, over $\Omega = I \times \mathbb{R}^n$; assume*

$$\|\boldsymbol{f}(t,\boldsymbol{y})\| \leq \alpha(t)\|\boldsymbol{y}\| + \beta(t)\,, \qquad \forall \boldsymbol{y} \in \mathbb{R}^n,\ \forall t \in I\,. \tag{10.38}$$

with α,β continuous and non-negative on I. Then the solution to the Cauchy problem (10.34) *is defined everywhere on I.*

Moreover, if α and β are integrable on I (improperly, possibly), every solution is bounded on I.

A crucial application regards linear systems.

Corollary 10.20 *Let* $\boldsymbol{A}(t) \in \mathbb{R}^{n,n}$*,* $\boldsymbol{b}(t) \in \mathbb{R}^n$ *be continuous on the open interval* $I \subseteq \mathbb{R}$*. Then any solution of the system of ODEs*

$$\boldsymbol{y}' = \boldsymbol{A}(t)\boldsymbol{y} + \boldsymbol{b}(t)$$

is defined on the entire I*. In particular, the corresponding Cauchy problem* (10.34) *with arbitrary* $\boldsymbol{y}_0 \in \mathbb{R}^n$ *admits one, and one only, solution over all of* I*.*

Proof. The map $\boldsymbol{f}(t, \boldsymbol{y}) = \boldsymbol{A}(t)\boldsymbol{y} + \boldsymbol{b}(t)$ is locally Lipschitz on $I \times \mathbb{R}^n$, by Example 10.13 iii). Following that argument,

$$\|\boldsymbol{f}(t, \boldsymbol{y})\| \le \|\boldsymbol{A}(t)\| \, \|\boldsymbol{y}\| + \|\boldsymbol{b}(t)\| \,, \qquad \forall \boldsymbol{y} \in \mathbb{R}^n, \forall t \in I \,,$$

and $\alpha(t) = \|\boldsymbol{A}(t)\|$, $\beta(t) = \|\boldsymbol{b}(t)\|$ are continuous by hypothesis. The claim follows from the previous theorem. □

Application: the simple pendulum (II). Let us resume the example of p. 430. The map $\boldsymbol{f}(\boldsymbol{y})$ has

$$\boldsymbol{Jf}(\boldsymbol{y}) = \begin{pmatrix} 0 & 1 \\ -k\cos y_1 & -\alpha \end{pmatrix}$$

so the Jacobian's components are uniformly bounded on $\mathbb{R}^2$, for

$$\left| \frac{\partial f_i}{\partial y_j}(\boldsymbol{y}) \right| \le \max(k,\alpha) \,, \qquad \forall \boldsymbol{y} \in \mathbb{R}^2 \,, \quad 1 \le i, j \le 2 \,;$$

then $\boldsymbol{f}$ is Lipschitz on the whole of $\mathbb{R}^2$. Theorems 10.14, 10.18 guarantee existence and uniqueness for any $(\theta_0, \theta_1) \in \mathbb{R}^2$, and the solution exists for all $t > 0$.
(Continues on p. 453.) □

10.4.4 Global existence in the future

Many concrete applications require an ODE to be solved 'in the future', rather than 'in the past': it is important, namely, to solve for all $t > t_0$, and the aim is to ensure the solution exists on some interval J bounded by t_0 on the left (e.g., $[t_0, +\infty)$), whereas the solution for $t < t_0$ is of no interest. The next result cashes in on a special feature of $\boldsymbol{f}$ to warrant the *global* existence of the solution *in the future* to an initial value problem.

A simple example will help us to understand. The autonomous problem

$$\begin{cases} y' = -y^3 \,, \\ y(0) = y_0 \,, \end{cases}$$

is determined by $f(y) = -y^3$: this is locally Lipschitz on $\mathbb{R}$, and clearly grows faster than a linear map as $y \to \infty$; therefore it does not fulfill the previous theorems, as the solution

$$y(t) = \frac{y_0}{\sqrt{1 + 2y_0^2 t}}\,, \tag{10.39}$$

not defined on the entire $\mathbb{R}$, confirms. Yet y exists on $[0, +\infty)$, whichever the initial datum y_0. The fact that the solution does not 'blow up', as t increases, can be derived directly from the differential equation, even without solving. Just multiply the ODE by y and observe

$$yy' = \frac{1}{2} 2y \frac{\mathrm{d}y}{\mathrm{d}t} = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{2}y^2\right);$$

then

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{2}y^2\right) = -y^4 \le 0\,,$$

i.e., the quantity $E(y(t)) = \frac{1}{2}|y(t)|^2$ is non-increasing as t grows. In particular,

$$\frac{1}{2}|y(t)|^2 \le \frac{1}{2}|y(0)|^2 = \frac{1}{2}|y_0|^2\,, \qquad \forall t \ge 0\,,$$

so

$$|y(t)| \le |y_0|\,, \qquad \forall t \ge 0\,.$$

At all 'future' instants the solution's absolute value is bounded by the initial value, which can also be established from the analytic expression (10.39).

The argument relies heavily on the *sign* of $f(y)$ (flipping from $-y^3$ to $+y^3$ invalidates the result); this property is not 'seen' by other conditions like (10.35) or (10.38), which involve the norm of $\boldsymbol{f}(t, \boldsymbol{y})$.

This example is somehow generalised by a condition for global existence in the future.

Proposition 10.21 *Let $\boldsymbol{f}$ be continuous and locally Lipschitz in $\boldsymbol{y}$, uniformly in t, on $\Omega = I \times \mathbb{R}^n$. If*

$$\boldsymbol{y} \cdot \boldsymbol{f}(t, \boldsymbol{y}) \le 0\,, \qquad \forall \boldsymbol{y} \in \mathbb{R}^n\,, \quad \forall t \in I\,, \tag{10.40}$$

the solution of the Cauchy problem (10.34) *exists everywhere to the right of t_0, hence on $J = I \cap [t_0, +\infty)$. Moreover, one has*

$$\|\boldsymbol{y}(t)\| \le \|\boldsymbol{y}_0\|\,, \qquad \forall t \in J\,. \tag{10.41}$$

Proof. Dot-multiply the ODE by $\boldsymbol{y}$:

$$\boldsymbol{y} \cdot \boldsymbol{y}' = \boldsymbol{y} \cdot \boldsymbol{f}(t, \boldsymbol{y})\,.$$

As
$$\boldsymbol{y}\cdot\boldsymbol{y}' = \sum_{i=1}^{n} y_i \frac{\mathrm{d}y_i}{\mathrm{d}t} = \frac{1}{2}\sum_{i=1}^{n}\frac{\mathrm{d}}{\mathrm{d}t}y_i^2 = \frac{1}{2}\frac{d}{dt}\|\boldsymbol{y}\|^2\,,$$
we have
$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{2}\|\boldsymbol{y}\|^2\right) \le 0\,,$$
whence $\|\boldsymbol{y}(t)\| \le \|\boldsymbol{y}_0\|$ for any $t > t_0$ where the solution exists. Thus we may use the Local Existence Theorem to extend the solution up to the right end-point of I. □

A few examples will shed light on (10.40).

Examples 10.22

i) Let $\boldsymbol{y}' = \boldsymbol{A}\boldsymbol{y}$, with $\boldsymbol{A}$ a real skew-symmetric matrix, in other words $\boldsymbol{A}$ satisfies $\boldsymbol{A}^T = -\boldsymbol{A}$. Then
$$\boldsymbol{y}\cdot\boldsymbol{f}(\boldsymbol{y}) = \boldsymbol{y}\cdot\boldsymbol{A}\boldsymbol{y} = 0\,, \qquad \forall \boldsymbol{y}\in\mathbb{R}^n\,,$$
because $\boldsymbol{y}\cdot\boldsymbol{A}\boldsymbol{y} = \boldsymbol{y}^T\boldsymbol{A}\boldsymbol{y} = (\boldsymbol{A}\boldsymbol{y})^T\boldsymbol{y} = \boldsymbol{y}^T\boldsymbol{A}^T\boldsymbol{y} = -\boldsymbol{y}^T\boldsymbol{A}\boldsymbol{y}$, so this quantity must be zero. Hence
$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{2}\|\boldsymbol{y}\|^2\right) = 0\,,$$
i.e., the map $E(\boldsymbol{y}(t)) = \frac{1}{2}\|\boldsymbol{y}(t)\|^2$ is constant in time. We call it a *first integral* of the differential equation, or an *invariant of motion*. The differential system is called in this case **conservative**.

ii) Take $\boldsymbol{y}' = -(\boldsymbol{A}\boldsymbol{y} + \boldsymbol{g}(\boldsymbol{y}))$, with $\boldsymbol{A}$ a symmetric, positive-definite matrix, and $\boldsymbol{g}(\boldsymbol{y})$ of components
$$\big(\boldsymbol{g}(\boldsymbol{y})\big)_i = \phi(y_i)\,, \qquad 1 \le i \le n\,,$$
where $\phi:\mathbb{R}\to\mathbb{R}$ is continuous, $\phi(0) = 0$ and $s\phi(s) \ge 0$, $\forall s\in\mathbb{R}$ (e.g., $\phi(s) = s^3$ an in the initial discussion). Recalling (4.18), we have
$$\boldsymbol{y}\cdot\boldsymbol{f}(\boldsymbol{y}) = -\left(\boldsymbol{y}\cdot\boldsymbol{A}\boldsymbol{y} + \boldsymbol{y}\cdot\boldsymbol{g}(\boldsymbol{y})\right) \le -\boldsymbol{y}\cdot\boldsymbol{A}\boldsymbol{y} \;\le\; -\lambda_*\|\boldsymbol{y}\|^2 \le 0\,, \quad \forall \boldsymbol{y}\in\mathbb{R}^n\,,$$
with $\lambda_* > 0$. Setting as above $E(\boldsymbol{y}) = \frac{1}{2}\|\boldsymbol{y}\|^2$, one has
$$\frac{\mathrm{d}}{\mathrm{d}t}E(\boldsymbol{y}(t)) \le -\lambda_*\|\boldsymbol{y}(t)\|^2\,, \qquad \forall t > t_0\,,$$
from which we deduce
$$E(\boldsymbol{y}(t)) \le \mathrm{e}^{-2\lambda_*(t-t_0)}E(\boldsymbol{y}_0)\,, \qquad \forall t > t_0\,;$$
therefore $E(\boldsymbol{y}(t))$ decays exponentially at t increases. Equivalently,
$$\|\boldsymbol{y}(t)\| \le \mathrm{e}^{-\lambda_*(t-t_0)}\|\boldsymbol{y}_0\|\,, \qquad t > t_0\,,$$
and all solutions converge exponentially to $\mathbf{0}$ as $t\to+\infty$. In Sect. 10.8 we shall express this by saying the constant solution $\boldsymbol{y} = \mathbf{0}$ is asymptotically uniformly stable, and speak of a **dissipative** system.

Assuming furthermore ϕ convex, one could prove that two solutions $\boldsymbol{y}$, $\boldsymbol{z}$ starting at $\boldsymbol{y}_0$, $\boldsymbol{z}_0$ satisfy

$$\|\boldsymbol{y}(t) - \boldsymbol{z}(t)\| \leq \mathrm{e}^{-\lambda_*(t-t_0)}\|\boldsymbol{y}_0 - \boldsymbol{z}_0\|, \qquad t > t_0,$$

a more accurate inequality than (10.36) for $t > t_0$.

The map $\boldsymbol{y} \mapsto E(\boldsymbol{y})$ is an example of a **Lyapunov function** for the differential system, relative to the origin. The name denotes any regular map $V : B(\mathbf{0}) \to \mathbb{R}$, where $B(\mathbf{0})$ is a neighbourhood of the origin, satisfying

- $V(\mathbf{0}) = 0, \quad V(\boldsymbol{y}) > 0$ on $B(\mathbf{0}) \setminus \{\mathbf{0}\}$,
- $\dfrac{\mathrm{d}}{\mathrm{d}t}V(\boldsymbol{y}(t)) \leq 0$ for any solution $\boldsymbol{y}$ of the ODE on $B(\mathbf{0})$. □

10.4.5 First integrals

Take an autonomous equation $\boldsymbol{y}' = \boldsymbol{f}(\boldsymbol{y})$, with Lipschitz $\boldsymbol{f} : D \subseteq \mathbb{R}^n \to \mathbb{R}^n$.

Definition 10.23 *A scalar map $\Phi : D \to \mathbb{R}$ is a* **first integral** *of the equation if Φ is constant along every orbit of the differential equation, i.e.,*

$$\Phi(\boldsymbol{y}(t)) = \text{constant} \qquad \textit{for every solution } \boldsymbol{y}(t) \textit{ of the ODE.}$$

A first integral is naturally defined up to an additive constant.

If the equation has a first integral Φ, *each orbit will be contained in a level set of* Φ. The study of level sets can then provide useful informations about the solution's global existence. For instance, if the orbit of a particular solution belongs to a level set that does not touch the boundary of D, we can be sure the solution will exist at any time.

Likewise, the study of the solutions' asymptotic stability (see Sect. 10.8) could profit from what the level sets of a first integral can tell.

To start with, consider the autonomous equation of order two

$$y'' + g(y) = 0, \tag{10.42}$$

where $g : \mathbb{R} \to \mathbb{R}$ is $\mathcal{C}^1$. Such kind of equations play a paramount role in Mechanics (read, Newton's law of motion). Another incarnation is equation (10.11) regulating a simple pendulum.

Let Π be an arbitrary primitive of g on $\mathbb{R}$, so that $\frac{\mathrm{d}\Pi}{\mathrm{d}y}(y) = g(y)$, $\forall y \in \mathbb{R}$. If $y = y(t)$ is a solution for $t \in J$ then, we can multiply (10.42) by y', to get

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{\mathrm{d}y}{\mathrm{d}t}\right)\frac{\mathrm{d}y}{\mathrm{d}t} + \frac{\mathrm{d}\Pi}{\mathrm{d}y}\frac{\mathrm{d}y}{\mathrm{d}t} = 0,$$

i.e.,

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{2}\left(y'(t)\right)^2 + \Pi(y(t))\right) = 0, \qquad \forall t \in J.$$

Thus

$$\boxed{E(y,y') = \frac{1}{2}\,(y')^2 + \varPi(y)}$$

is constant for any solution. One calls $\varPi(y)$ the **potential energy** of the mechanical system described by the equation, and $K(y') = \frac{1}{2}(y')^2$ is the **kinetic energy**. The **total energy** $E(y,y') = K(y') + \varPi(y)$ is therefore *preserved during motion* (the work done equals the gain of kinetic energy).

The function E turns out to be a first integral, in the sense of Definition 10.23, for the autonomous vector equation $\boldsymbol{y}' = \boldsymbol{f}(\boldsymbol{y})$ given by

$$\begin{cases} y_1' = y_2 \\ y_2' = -g(y_1)\,, \end{cases} \tag{10.43}$$

equivalent to (10.42) by putting $y_1 = y$, $y_2 = y'$. In other words $\varPhi(\boldsymbol{y}) = E(y_1,y_2) = \frac{1}{2}y_2^2 + \varPi(y_1)$, and we may study the *level curves* of $\varPhi$ in phase space y_1y_2 to understand the solutions' behaviour. At the end of the section we will see an example, the undamped pendulum.

Generalising this picture we may say that an autonomous equation $\boldsymbol{y}' = \boldsymbol{f}(\boldsymbol{y})$ over a connected open set D of the plane admits infinitely many first integrals *if* $\boldsymbol{f}$ *is the curl of a scalar field* $\varPhi$ on D, i.e., $\boldsymbol{f}(\boldsymbol{y}) = \mathbf{curl}\,\varPhi(\boldsymbol{y})$.

Property 10.24 *Let $\varPhi$ be a $\mathcal{C}^1$ field on $D \subseteq \mathbb{R}^2$. The equation*

$$\boldsymbol{y}' = \mathbf{curl}\,\varPhi \tag{10.44}$$

admits $\varPhi$ as first integral.

Proof. If $\boldsymbol{y} = \boldsymbol{y}(t)$ is a solution defined for $t \in J$, by the chain rule

$$\frac{\mathrm{d}}{\mathrm{d}t}\varPhi(\boldsymbol{y}) = (\mathbf{grad}\,\varPhi)\cdot\frac{\mathrm{d}\boldsymbol{y}}{\mathrm{d}t} = (\mathbf{grad}\,\varPhi)\cdot(\mathbf{curl}\,\varPhi) = 0\,,$$

because gradient and curl are orthogonal in dimension two. □

Bearing in mind Sect. 7.2.1 with regard to level cuves, equation (10.44) forces $\boldsymbol{y}$ to move along a level curve of $\varPhi$.

By Sect. 9.6, a sufficient condition to have $\boldsymbol{f} = \mathbf{curl}\,\varPhi$ is that $\boldsymbol{f}$ is divergence-free on a simply connected, open set D. In the case treated at the beginning, $\boldsymbol{f}(\boldsymbol{y}) = (y_2, -g(y_1)) = \mathbf{curl}\,E(y_1,y_2)$, and $\operatorname{div}\boldsymbol{f} = 0$ over all $D = \mathbb{R}^2$.

To conclude, equation (10.44) extends to dimension $2n$, if $\boldsymbol{y} = (\boldsymbol{y}_1, \boldsymbol{y}_2) \in \mathbb{R}^n \times \mathbb{R}^n = \mathbb{R}^{2n}$ is defined by a system of $2n$ equations

$$\boxed{\begin{cases} \boldsymbol{y}_1' = \dfrac{\partial \Phi}{\partial \boldsymbol{y}_2} \\[2ex] \boldsymbol{y}_2' = -\dfrac{\partial \Phi}{\partial \boldsymbol{y}_1}\,, \end{cases}} \tag{10.45}$$

where Φ is a map in $2n$ variables and each partial derivative symbol denotes the n components of the gradient of Φ with respect to the indicated vectorial variable. Such a system is called a **Hamiltonian system**, and the first integral Φ is known as a **Hamiltonian** (**function**) of the system. An example, that generalises (10.42), is provided by the equation of motion of n bodies

$$\boldsymbol{z}'' + \mathbf{grad}\,\Pi(\boldsymbol{z}) = \mathbf{0}\,, \qquad \boldsymbol{z} \in \mathbb{R}^n\,, \tag{10.46}$$

with Π a function of n variables. Proceeding as for $n = 1$ we can transform the second-order equation into a system of order one like (10.45). The Hamiltonian Φ is the total energy $\Phi(\boldsymbol{y}_1, \boldsymbol{y}_2) = E(\boldsymbol{z}, \boldsymbol{z}') = \frac{1}{2}\|\boldsymbol{z}'\|^2 + \Pi(\boldsymbol{z})$ of the system.

When $D = \mathbb{R}^n$, it can be proved that *the ODE's solutions are defined at all times if the potential energy is bounded from below.*

Application: the simple pendulum (III). The example continues from p. 448. The simple pendulum ($\alpha = 0$) is governed by

$$\frac{\mathrm{d}\theta^2}{\mathrm{d}t^2} + k\sin\theta = 0\,.$$

The equation is of type (10.42), with $g(\theta) = k\sin\theta$; therefore the potential energy is $\Pi(\theta) = c - k\cos\theta$, with arbitrary constant c. It is certainly bounded from below, confirming the solutions' existence at any instant of time $t \in \mathbb{R}$. The customary choice $c = k$ makes the potential energy vanish when $\theta = 2\pi\ell$, $\ell \in \mathbb{Z}$, in correspondence with the lowest point S of the bob, and maximises $\Pi(\theta) = 2k > 0$ for $\theta = (2\ell + 1)\pi$, $\ell \in \mathbb{Z}$, when the highest point I is reached. With that choice the total energy is

$$E(\theta,\theta') = \frac{1}{2}(\theta')^2 + k(1 - \cos\theta)\,,$$

or

$$E(y_1, y_2) = \frac{1}{2}y_2^2 + k(1 - \cos y_1) \tag{10.47}$$

in the phase space of coordinates $y_1 = \theta$, $y_2 = \theta'$. Let us examine this map on $\mathbb{R}^2$: it is always ≥ 0 and vanishes at $(2\pi\ell, 0)$, $\ell \in \mathbb{Z}$, which are thus absolute minima: the pendulum is still in the position S. The gradient $\nabla E(y_1, y_2) = (k\sin y_1, y_2)$ is zero at $(\pi\ell, 0)$, $\ell \in \mathbb{Z}$, so we have additional stationary points $\big((2\ell+1)\pi, 0\big)$, $\ell \in \mathbb{Z}$, easily seen to be saddle points. The level curves $E(y_1, y_2) = c \geq 0$, defined by

$$y_2 = \pm\sqrt{2(c-k) + 2k\cos y_1}\,,$$

have the following structure (see Fig. 10.8):

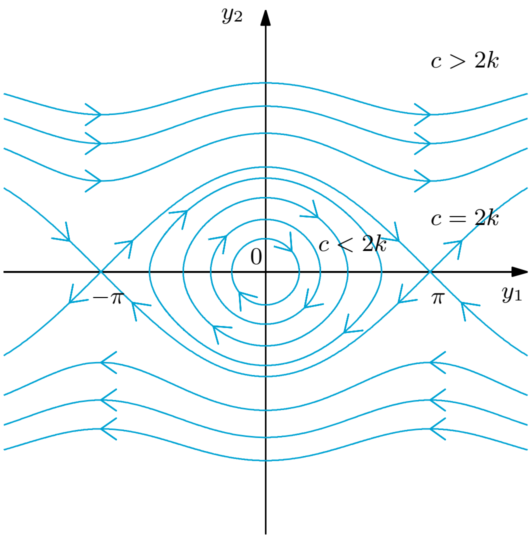


Figure 10.8. The orbits of a simple pendulum coincide with the level curves of the energy $E(y_1, y_2)$

- When $c < 2k$, they are closed curves encircling the minima of E; they correspond to periodic oscillations between $\theta = -\theta_0$ and θ_0 (plus multiples of 2π), where θ_0 is determined by requiring that all energy c be potential, that is $k(1 - \cos\theta_0) = c$.
- When $c = 2k$, the curve's branches connect two saddle points of E, and correspond to the limit situation in which the bob reaches the top equilibrium point I with zero velocity (in an infinite time).
- When $c > 2k$, the curves are neither closed nor bounded; they correspond to a minimum non-zero velocity, due to which the pendulum moves past the top point and then continues to rotate around O, without stopping.

(Continues on p. 486.) $\square$)

10.5 Linear systems of first order

The next two sections concentrate on vectorial linear equations of order one

$$\boxed{\boldsymbol{y}' = \boldsymbol{A}(t)\boldsymbol{y} + \boldsymbol{b}(t)\,,} \tag{10.48}$$

where $\boldsymbol{A}$ is a map from an interval I of the real line to the vector space $\mathbb{R}^{n\times n}$ of square matrices of order n, while $\boldsymbol{b}$ is a function from I to $\mathbb{R}^n$. We shall assume $\boldsymbol{A}$ and $\boldsymbol{b}$ are continuous in t. Equation (10.48) is shorthand writing for a **system** of n differential equations in n unknown functions

$$\begin{cases} y_1' = a_{11}(t)y_1 + a_{12}(t)y_2 + \ldots + a_{1n}(t)y_n + b_1(t) \\ y_2' = a_{21}(t)y_1 + a_{22}(t)y_2 + \ldots + a_{2n}(t)y_n + b_2(t) \\ \quad\vdots \\ y_n' = a_{n1}(t)y_1 + a_{n2}(t)y_2 + \ldots + a_{nn}(t)y_n + b_n(t)\,. \end{cases}$$

Remark 10.25 The concern with linear equations can be ascribed to the fact that many mathematical models are regulated by such equations. Models based on non-linear equations are often approximated by simpler and more practicable linear versions, by the process of **linearisation**. The latter essentially consists in arresting the Taylor expansion of a non-linear map to order one, and disregarding the remaining part.

More concretely, suppose $\boldsymbol{f}$ is a $\mathcal{C}^1$ map on Ω in the variable $\boldsymbol{y}$ and that $\bar{\boldsymbol{y}}(t)$, $t \in J \subseteq I$, is a known solution of the ODE $\boldsymbol{y}' = \boldsymbol{f}(t, \boldsymbol{y})$ (e.g., a constant solution $\bar{\boldsymbol{y}}(t) = \bar{\boldsymbol{y}}_0$). Then the solutions $\boldsymbol{y}(t)$ that are 'close to $\bar{\boldsymbol{y}}(t)$' can be approximated as follows: for any given $t \in J$, the expansion of $\boldsymbol{y} \mapsto \boldsymbol{f}(t, \boldsymbol{y})$ around the point $\bar{\boldsymbol{y}}(t)$ reads

$$\boldsymbol{f}(t, \boldsymbol{y}) = \boldsymbol{f}\big(t, \bar{\boldsymbol{y}}(t)\big) + \boldsymbol{J_y f}\big(t, \bar{\boldsymbol{y}}(t)\big)\big(\boldsymbol{y} - \bar{\boldsymbol{y}}(t)\big) + \boldsymbol{g}(t, \boldsymbol{y}) \tag{10.49}$$

with $\boldsymbol{g}(t, \boldsymbol{y}) = o\big(\boldsymbol{y} - \bar{\boldsymbol{y}}(t)\big)$ for $\boldsymbol{y} \to \bar{\boldsymbol{y}}(t)$; $\boldsymbol{J_y f}$ denotes the Jacobian matrix of $\boldsymbol{f}$ with respect to $\boldsymbol{y}$ only. Set

$$\boldsymbol{A}(t) = \boldsymbol{J_y f}\big(t, \bar{\boldsymbol{y}}(t)\big)$$

and note that $\boldsymbol{f}\big(t, \bar{\boldsymbol{y}}(t)\big) = \bar{\boldsymbol{y}}'(t)$, $\bar{\boldsymbol{y}}$ being a solution. Substituting (10.49) in $\boldsymbol{y}' = \boldsymbol{f}(t, \boldsymbol{y})$ gives

$$(\boldsymbol{y} - \bar{\boldsymbol{y}})' = \boldsymbol{A}(t)(\boldsymbol{y} - \bar{\boldsymbol{y}}) + \boldsymbol{g}(t, \boldsymbol{y})\,;$$

ignoring the infinitesimal part $\boldsymbol{g}$ then, we can approximate the solution $\boldsymbol{y}$ by setting $\boldsymbol{z} \sim \boldsymbol{y} - \bar{\boldsymbol{y}}$ and solving the linear equation

$$\boldsymbol{z}' = \boldsymbol{A}(t)\boldsymbol{z}\,;$$

if $\boldsymbol{y}_0$ is defined by the initial condition $\boldsymbol{y}(t_0) = \boldsymbol{y}_0$ at $t_0 \in J$, $\boldsymbol{z}$ will be determined by the datum $\boldsymbol{z}(t_0) = \boldsymbol{y}_0 - \bar{\boldsymbol{y}}(t_0)$. Once $\boldsymbol{z}$ is known, the approximation of $\boldsymbol{y}(t)$ is $\tilde{\boldsymbol{y}}(t) = \bar{\boldsymbol{y}}(t) + \boldsymbol{z}(t)$.

In the special case of autonomous systems with a constant solution $\bar{\boldsymbol{y}}_0$, the matrix $\boldsymbol{A}$ is not time-dependent. □

We will show in the sequel that equation (10.48) admits exactly n linearly independent solutions; as claimed in Sect. 10.2, therefore, the general integral depends on n arbitrary constants that may be determined by assigning a Cauchy condition at a point $t_0 \in I$. Furthermore, if $\boldsymbol{A}$ does not depend on t, so that the a_{ij} are constants, the general integral can be recovered from the (possibly generalised) eigenvalues and eigenvectors of $\boldsymbol{A}$.

We begin by tackling the homogeneous case, in other words $\boldsymbol{b} = \boldsymbol{0}$.

10.5.1 Homogeneous systems

Consider the homogeneous equation associated to (10.48), namely

$$\boxed{\boldsymbol{y}' = \boldsymbol{A}(t)\boldsymbol{y}\,, \qquad t \in I\,,} \tag{10.50}$$

and let us prove a key fact.

Proposition 10.26 *The set S_0 of solutions to* (10.50) *is a vector space of dimension n.*

Proof. Since the equation is linear, any linear combination of solutions is still a solution, making S_0 a vector space.
To compute its dimension, we shall exhibit an explicit basis. To that end, fix an arbitrary point $t_0 \in I$ and consider the n Cauchy problems

$$\begin{cases} \boldsymbol{y}' = \boldsymbol{A}(t)\boldsymbol{y}\,, & t \in I\,, \\ \boldsymbol{y}(t_0) = \boldsymbol{e}_i\,, \end{cases}$$

where $\{\boldsymbol{e}_i\}_{i=1,\ldots,n}$ is the canonical basis of $\mathbb{R}^n$. Recalling Corollary 10.20, each system admits a unique solution $\boldsymbol{y} = \boldsymbol{u}_i(t)$ of class $\mathcal{C}^1$ on I. The maps $\boldsymbol{u}_i$, $i = 1, \ldots, n$, are linearly independent: in fact, if $\boldsymbol{y} = \sum_{i=1}^{n} \alpha_i \boldsymbol{u}_i$ is the zero map on I, then $\sum_{i=1}^{n} \alpha_i \boldsymbol{u}_i(t) = 0$ for any $t \in I$, so also for $t = t_0$

$$\sum_{i=1}^{n} \alpha_i \boldsymbol{u}_i(t_0) = \sum_{i=1}^{n} \alpha_i \boldsymbol{e}_i = 0\,,$$

whence $\alpha_i = 0$, $\forall i$, proving the linear independence of the $\boldsymbol{u}_i$.
Eventually, if $\boldsymbol{y} = \boldsymbol{y}(t)$ is an element in S_0, we write the vector $\boldsymbol{y}_0 = \boldsymbol{y}(t_0)$ as $\boldsymbol{y}_0 = \sum_{i=1}^{n} y_{0i}\boldsymbol{e}_i$; then

$$\boldsymbol{y}(t) = \sum_{i=1}^{n} y_{0i}\boldsymbol{u}_i(t)\,,$$

because both sides solve the equation and agree on t_0, and the solution is unique. □

The set $\{\boldsymbol{u}_i\}_{i=1,\ldots,n}$ of maps satisfying

$$\begin{cases} \boldsymbol{u}_i' = \boldsymbol{A}(t)\boldsymbol{u}_i\,, & t \in I\,, \\ \boldsymbol{u}_i(t_0) = \boldsymbol{e}_i\,, \end{cases} \tag{10.51}$$

is an example of a fundamental system of solutions.

Definition 10.27 *Any basis of S_0 is called a* **fundamental system** *of solutions of the ODE* (10.50)*.*

There is a useful way to check whether a set of n solutions to (10.50) is a fundamental system.

Proposition 10.28 *Let $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n \in S_0$.*

a) If at a point $t_0 \in I$ the vectors $\boldsymbol{w}_1(t_0), \ldots, \boldsymbol{w}_n(t_0)$ are linearly independent, $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n$ constitute a fundamental system of solutions.
b) If $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n$ form a fundamental system of solutions, at each point $t_0 \in I$ the vectors $\boldsymbol{w}_1(t_0), \ldots, \boldsymbol{w}_n(t_0)$ are linearly independent.

Proof. a) It suffices to show that $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n$ are linearly independent. Suppose then there exist coefficients c_i such that

$$\sum_{i=1}^{n} c_i \boldsymbol{w}_i(t) = 0\,, \qquad \forall t \in I\,.$$

Choosing $t = t_0$, the hypothesis forces every c_i to vanish.
b) Suppose there exist $c_1, \ldots, c_n$ with

$$\sum_{i=1}^{n} c_i \boldsymbol{w}_i(t_0) = 0\,.$$

Define the map $\boldsymbol{z}(t) = \sum_{i=1}^{n} c_i \boldsymbol{w}_i(t)$; since it solves the Cauchy problem

$$\begin{cases} \boldsymbol{z}' = \boldsymbol{A}\boldsymbol{z} \\ \boldsymbol{z}(t_0) = \boldsymbol{0}\,, \end{cases}$$

by uniqueness we have $\boldsymbol{z}(t) = 0$ for any $t \in I$. Therefore all c_i must be zero. □

Owing to the proposition, the linear dependence of n vector-valued maps on I reduces to the (much easier to verify) linear dependence of n vectors in $\mathbb{R}^n$.

We explain now how a fundamental system permits to find all solutions of (10.50) and also to solve the corresponding initial value problem. Given then a fundamental system $\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n\}$, for any $t \in I$ we associate to it the $n \times n$ matrix

$$\boxed{\boldsymbol{W}(t) = \big(\boldsymbol{w}_1(t), \dots, \boldsymbol{w}_n(t)\big)}$$

whose columns are the vectors $\boldsymbol{w}_i(t)$; by part *b)* above, the matrix $\boldsymbol{W}(t)$ is non-singular since its columns are linearly independent. This matrix is called the **fundamental matrix**, and using it we may write the n vectorial equations $\boldsymbol{w}_i' = \boldsymbol{A}(t)\boldsymbol{w}_i$, $i = 1, \dots, n$, in the compact form

$$\boldsymbol{W}' = \boldsymbol{A}(t)\boldsymbol{W} . \tag{10.52}$$

Every solution $\boldsymbol{y} = \boldsymbol{y}(t)$ of (10.50) is represented as

$$\boldsymbol{y}(t) = \sum_{i=1}^{n} c_i \boldsymbol{w}_i(t)$$

for suitable constants $c_i \in \mathbb{R}$. Equivalently, setting $\boldsymbol{c} = (c_1, \dots, c_n)^T$, we have

$$\boxed{\boldsymbol{y}(t) = \boldsymbol{W}(t)\, \boldsymbol{c} .} \tag{10.53}$$

The solution to the generic Cauchy problem

$$\begin{cases} \boldsymbol{y}' = \boldsymbol{A}\boldsymbol{y} & \text{on } I , \\ \boldsymbol{y}(t_0) = \boldsymbol{y}_0 , \end{cases} \tag{10.54}$$

is found by solving the linear system

$$\boxed{\boldsymbol{W}(t_0)\, \boldsymbol{c} = \boldsymbol{y}_0 .} \tag{10.55}$$

Therefore we may write the solution, formally, as

$$\boxed{\boldsymbol{y}(t) = \boldsymbol{W}(t)\boldsymbol{W}^{-1}(t_0)\, \boldsymbol{y}_0 .}$$

We can simplify this by choosing the fundamental matrix $\boldsymbol{U}(t)$ associated with the special basis $\{\boldsymbol{u}_i\}$ defined by (10.51); it arises as solution of the Cauchy problem in matrix form

$$\begin{cases} \boldsymbol{U}' = \boldsymbol{A}(t)\boldsymbol{U} & \text{on } I , \\ \boldsymbol{U}(t_0) = \boldsymbol{I} , \end{cases} \tag{10.56}$$

and allows us to write the solution to (10.54) as

$$\boldsymbol{y}(t) = \boldsymbol{U}(t)\, \boldsymbol{y}_0 .$$

We put on hold, for the time being, the study of equation (10.50) to discuss non-homogeneous systems. We will resume it in Sect. 10.6 under the hypothesis that $\boldsymbol{A}$ be constant.

10.5.2 Non-homogeneous systems

Indicate by S_b the set of all solutions to (10.48). This set can be characterised starting from a particular solution, also known as *particular integral*.

Proposition 10.29 *Given a solution* $\boldsymbol{y}_p$ *of equation* (10.48)*,* S_b *is the affine space* $S_b = \boldsymbol{y}_p + S_0$*.*

Proof. If $\boldsymbol{y} \in S_b$, by linearity $\boldsymbol{y} - \boldsymbol{y}_p$ solves the homogeneous equation (10.50), hence $\boldsymbol{y} - \boldsymbol{y}_p \in S_0$. □

The proposition says that for any given fundamental system $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n$ of solutions to the homogeneous equation, each solution of (10.48) has the form

$$\boxed{\boldsymbol{y}(t) = \boldsymbol{W}(t)\,\boldsymbol{c} + \boldsymbol{y}_p(t) = \sum_{i=1}^{n} c_i \boldsymbol{w}_i(t) + \boldsymbol{y}_p(t)}$$

for some $\boldsymbol{c} = (c_1, \ldots, c_n)^T \in \mathbb{R}^n$.

An alternative route to arrive at the general integral of (10.48) is the *method of variation of constants*, already seen in Sect. 10.3.3 for scalar equations. By (10.53) we look for a solution of the form

$$\boldsymbol{y}(t) = \boldsymbol{W}(t)\,\boldsymbol{c}(t)$$

with $\boldsymbol{c}(t)$ differentiable to be determined. Substituting, we obtain

$$\boldsymbol{W}'(t)\,\boldsymbol{c}(t) + \boldsymbol{W}(t)\,\boldsymbol{c}'(t) = \boldsymbol{A}(t)\boldsymbol{W}(t)\,\boldsymbol{c}(t) + \boldsymbol{b}(t)\,.$$

By (10.52) we arrive at

$$\boldsymbol{W}(t)\,\boldsymbol{c}'(t) = \boldsymbol{b}(t)\,, \qquad \text{from which} \qquad \boldsymbol{c}'(t) = \boldsymbol{W}^{-1}(t)\,\boldsymbol{b}(t)\,,$$

i.e.,

$$\boldsymbol{c}(t) = \int \boldsymbol{W}^{-1}(t)\,\boldsymbol{b}(t)\,\mathrm{d}t\,, \tag{10.57}$$

where the indefinite integral denotes the primitive of the vector $\boldsymbol{W}^{-1}(t)\,\boldsymbol{b}(t)$, component by component. The general integral of (10.48) is then

$$\boxed{\boldsymbol{y}(t) = \boldsymbol{W}(t) \int \boldsymbol{W}^{-1}(t)\,\boldsymbol{b}(t)\,\mathrm{d}t\,.} \tag{10.58}$$

Example 10.30 Consider the equation

$$\boldsymbol{y}' = \begin{pmatrix} 2 & -2 \\ -2 & 2 \end{pmatrix} \boldsymbol{y} + \begin{pmatrix} \dfrac{1}{\sqrt{1-t^2}} - \dfrac{e^{4t}}{1+t^2} \\ \dfrac{1}{\sqrt{1-t^2}} - \dfrac{e^{4t}}{1+t^2} \end{pmatrix} . \tag{10.59}$$

A fundamental system of solutions of the homogeneous equation is

$$\boldsymbol{w}_1(t) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad \boldsymbol{w}_2(t) = e^{4t} \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

(as we shall explain in the next section). The fundamental matrix and its inverse are, thus,

$$\boldsymbol{W}(t) = \begin{pmatrix} 1 & e^{4t} \\ 1 & -e^{4t} \end{pmatrix}, \qquad \boldsymbol{W}^{-1}(t) = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ e^{-4t} & -e^{-4t} \end{pmatrix} .$$

Consequently

$$\boldsymbol{W}^{-1}(t)\boldsymbol{b}(t) = \begin{pmatrix} \dfrac{1}{\sqrt{1-t^2}} \\ \dfrac{1}{1+t^2} \end{pmatrix},$$

so

$$\int \boldsymbol{W}^{-1}(t)\boldsymbol{b}(t)\,dt = \begin{pmatrix} \arcsin t + c_1 \\ \arctan t + c_2 \end{pmatrix} .$$

Due to (10.58), the general integral is

$$\boldsymbol{y}(t) = \begin{pmatrix} \arcsin t + c_1 + e^{4t}(\arctan t + c_2) \\ \arcsin t + c_1 - e^{4t}(\arctan t + c_2) \end{pmatrix} .$$

□

Whenever we have to solve the Cauchy problem

$$\begin{cases} \boldsymbol{y}' = \boldsymbol{A}(t)\boldsymbol{y} + \boldsymbol{b}(t) \quad \text{on } I, \\ \boldsymbol{y}(t_0) = \boldsymbol{y}_0, \end{cases} \tag{10.60}$$

it is convenient to write (10.57) as

$$\boldsymbol{c}(t) = \boldsymbol{c} + \int_{t_0}^{t} \boldsymbol{W}^{-1}(s)\,\boldsymbol{b}(s)\,ds .$$

with arbitrary $\boldsymbol{c} \in \mathbb{R}^n$. Equation (10.58) then becomes

$$\boldsymbol{y}(t) = \boldsymbol{W}(t)\boldsymbol{c} + \int_{t_0}^{t} \boldsymbol{W}(t)\boldsymbol{W}^{-1}(s)\,\boldsymbol{b}(s)\,ds ;$$

putting $t = t_0$ gives $\boldsymbol{y}(t_0) = \boldsymbol{W}(t_0)\boldsymbol{c}$, so $\boldsymbol{c}$ is still determined by solving (10.55).

Altogether, the solution to the Cauchy problem is

$$\boxed{\boldsymbol{y}(t) = \boldsymbol{y}_{\text{hom}}(t) + \boldsymbol{y}_p(t) ,}$$

where $\boldsymbol{y}_{\text{hom}}(t) = \boldsymbol{W}(t)\boldsymbol{W}^{-1}(t_0)\boldsymbol{y}_0$ solves the homogeneous system (10.54), while

$$\boxed{\boldsymbol{y}_p(t) = \int_{t_0}^{t} \boldsymbol{W}(t)\boldsymbol{W}^{-1}(s)\,\boldsymbol{b}(s)\,\mathrm{d}s} \tag{10.61}$$

solves the non-homogeneous problem with null datum at $t = t_0$.

Example 10.31 Let us solve the Cauchy problem (10.60) for the same equation of the previous example, and with $\boldsymbol{y}(0) = (4, 2)^T$. The general integral found earlier is

$$\boldsymbol{y}(t) = \begin{pmatrix} c_1 + \mathrm{e}^{4t} c_2 \\ c_1 - \mathrm{e}^{4t} c_2 \end{pmatrix} + \begin{pmatrix} \arcsin t + \mathrm{e}^{4t} \arctan t \\ \arcsin t - \mathrm{e}^{4t} \arctan t \end{pmatrix};$$

the first bracket represents the solution $\boldsymbol{y}_{\text{hom}}$ of the associated homogeneous equation, and the second is the particular solution $\boldsymbol{y}_p$ vanishing at $t = 0$. The constants c_1, c_2 are found by setting $t = 0$ and solving the linear system

$$\begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 4 \\ 2 \end{pmatrix},$$

whence $c_1 = 3$ and $c_2 = 1$. □

10.6 Linear systems with constant matrix $\boldsymbol{A}$

We are ready to resume equation (10.48), under the additional assumption that $\boldsymbol{A}$ is independent of time. We wish to explain a procedure for finding a fundamental system of solutions for

$$\boxed{\boldsymbol{y}' = \boldsymbol{A}\boldsymbol{y} \qquad \text{on } I = \mathbb{R}\,.} \tag{10.62}$$

The method relies on the computation of the eigenvalues of $\boldsymbol{A}$, the roots of the so-called **characteristic polynomial** of the equation, defined by

$$\chi(\lambda) = \det(\boldsymbol{A} - \lambda \boldsymbol{I}),$$

and the corresponding eigenvectors (perhaps generalised, and defined in the sequel).

At the heart of the whole matter lies an essential property:

> if λ is an eigenvalue of $\boldsymbol{A}$ with eigenvector $\boldsymbol{v}$, the map
>
> $$\boldsymbol{w}(t) = \mathrm{e}^{\lambda t}\boldsymbol{v}$$
>
> solves equation (10.62) (in $\mathbb{C}$, if λ is complex).

Indeed, using (10.28) and $\boldsymbol{A}\boldsymbol{v} = \lambda\boldsymbol{v}$, we have

$$\boldsymbol{w}'(t) = \frac{\mathrm{d}}{\mathrm{d}t}\left(\mathrm{e}^{\lambda t}\boldsymbol{v}\right) = \frac{\mathrm{d}\mathrm{e}^{\lambda t}}{\mathrm{d}t}\,\boldsymbol{v} = \lambda \mathrm{e}^{\lambda t}\,\boldsymbol{v} = \mathrm{e}^{\lambda t}\boldsymbol{A}\boldsymbol{v} = \boldsymbol{A}\left(\mathrm{e}^{\lambda t}\boldsymbol{v}\right) = \boldsymbol{A}\boldsymbol{w}(t)\,.$$

At this point the treatise needs to split in three parts: first, we describe the explicit steps leading to a fundamental system, then explain the procedure by means of examples, and at last provide the theoretical backgrounds.

To simplify the account, the case where $\boldsymbol{A}$ is diagonalisable is kept separate. Diagonalisable matrices include prominent cases, like symmetric matrices (whose eigenvalues and eigenvectors are real) and more generally normal matrices (see Sect. 4.2). Only afterwards we examine the more involved situation of a matrix that cannot be diagonalised.

In the end we illustrate a solving method for non-homogeneous equations, with emphasis on special choices of the source term $\boldsymbol{b}(t)$.

10.6.1 Homogeneous systems with diagonalisable $\boldsymbol{A}$

Suppose $\boldsymbol{A}$ has ℓ real eigenvalues $\lambda_1, \ldots, \lambda_\ell$ and $n-\ell$ complex eigenvalues which, being $\boldsymbol{A}$ real, come in m complex-conjugate pairs $\lambda_{\ell+1}, \overline{\lambda}_{\ell+1}, \ldots, \lambda_{\ell+m}, \overline{\lambda}_{\ell+m}$; eigenvalues are repeated according to their algebraic multiplicity, hence $\ell + 2m = n$. Each real λ_k, $1 \le k \le \ell$, is associated to a real eigenvector $\boldsymbol{v}_k$, while to each pair $(\lambda_{\ell+k}, \overline{\lambda}_{\ell+k})$, $1 \le k \le m$, corresponds a pair $(\boldsymbol{v}_{\ell+k}, \overline{\boldsymbol{v}}_{\ell+k})$ of complex-conjugate eigenvectors. As $\boldsymbol{A}$ is diagonalisable, we can assume the eigenvectors are linearly independent (over $\mathbb{C}$).

- For each real eigenvalue λ_k, $1 \le k \le \ell$, define the map

$$\boxed{\boldsymbol{w}_k(t) = \mathrm{e}^{\lambda_k t}\boldsymbol{v}_k\,.}$$

- For each pair $(\lambda_{\ell+k}, \overline{\lambda}_{\ell+k})$, $1 \le k \le m$, we decompose eigenvalues and eigenvectors into real and imaginary parts: $\lambda_{\ell+k} = \sigma_k + i\omega_k$, $\boldsymbol{v}_{\ell+k} = \boldsymbol{v}_k^{(1)} + i\boldsymbol{v}_k^{(2)}$. Define maps

$$\boxed{\begin{aligned} \boldsymbol{w}_k^{(1)}(t) &= \mathcal{R}e\left(\mathrm{e}^{\lambda_{\ell+k}t}\boldsymbol{v}_{\ell+k}\right) = \mathrm{e}^{\sigma_k t}\left(\boldsymbol{v}_k^{(1)}\cos\omega_k t - \boldsymbol{v}_k^{(2)}\sin\omega_k t\right), \\ \boldsymbol{w}_k^{(2)}(t) &= \mathcal{I}m\left(\mathrm{e}^{\lambda_{\ell+k}t}\boldsymbol{v}_{\ell+k}\right) = \mathrm{e}^{\sigma_k t}\left(\boldsymbol{v}_k^{(1)}\sin\omega_k t + \boldsymbol{v}_k^{(2)}\cos\omega_k t\right). \end{aligned}}$$

Proposition 10.32 *With the above conventions, the set of functions* $\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_\ell, \boldsymbol{w}_1^{(1)}, \boldsymbol{w}_1^{(2)}, \ldots, \boldsymbol{w}_m^{(1)}, \boldsymbol{w}_m^{(2)}\}$ *is a fundamental system of solutions to* (10.62).

The general integral is then of the form

$$\boldsymbol{y}(t) = \sum_{k=1}^{\ell} c_k \boldsymbol{w}_k(t) + \sum_{k=1}^{m} \left(c_k^{(1)} \boldsymbol{w}_k^{(1)}(t) + c_k^{(2)} \boldsymbol{w}_k^{(2)}(t) \right), \tag{10.63}$$

with $c_1, \ldots, c_\ell, c_1^{(1)}, c_1^{(2)}, \ldots, c_m^{(1)}, c_m^{(2)} \in \mathbb{R}$.

The above is nothing but the general formula (10.53) adapted to the present situation; $\boldsymbol{W}(t)$ is the fundamental matrix associated to the system.

Examples 10.33

i) The matrix

$$\boldsymbol{A} = \begin{pmatrix} -1 & 2 \\ 2 & -1 \end{pmatrix}$$

admits eigenvalues $\lambda_1 = 1$, $\lambda_2 = -3$ and corresponding eigenvectors

$$\boldsymbol{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \qquad \text{and} \qquad \boldsymbol{v}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

As a consequence, (10.63) furnishes the general integral of (10.62):

$$\boldsymbol{y}(t) = c_1 \mathrm{e}^t \begin{pmatrix} 1 \\ 1 \end{pmatrix} + c_2 \mathrm{e}^{-3t} \begin{pmatrix} 1 \\ -1 \end{pmatrix} = \begin{pmatrix} c_1 \mathrm{e}^t + c_2 \mathrm{e}^{-3t} \\ c_1 \mathrm{e}^t - c_2 \mathrm{e}^{-3t} \end{pmatrix}.$$

For a particular solution we can consider the Cauchy problem with initial datum

$$\boldsymbol{y}(0) = \boldsymbol{y}_0 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$$

for example, corresponding to the linear system

$$\begin{cases} c_1 + c_2 = 2 \\ c_1 - c_2 = 1 . \end{cases}$$

Therefore $c_1 = 3/2$, $c_2 = 1/2$ and the solution reads

$$\boldsymbol{y}(t) = \frac{1}{2} \begin{pmatrix} 3\mathrm{e}^t + \mathrm{e}^{-3t} \\ 3\mathrm{e}^t - \mathrm{e}^{-3t} \end{pmatrix}.$$

ii) The matrix

$$\boldsymbol{A} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & -10 & 5 \end{pmatrix}$$

has eigenvalues $\lambda_1 = 1$, $\lambda_2 = 2 + i$ and its conjugate $\overline{\lambda}_2 = 2 - i$. The eigenvectors are easy to find:

$$\boldsymbol{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \boldsymbol{v}_2 = \begin{pmatrix} 1 \\ 1+i \\ 2+4i \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} + i \begin{pmatrix} 0 \\ 1 \\ 4 \end{pmatrix}, \quad \overline{\boldsymbol{v}}_2 = \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} - i \begin{pmatrix} 0 \\ 1 \\ 4 \end{pmatrix}.$$

Therefore (10.63) (we write $c_2 = c_2^{(1)}$ and $c_3 = c_2^{(2)}$ for simplicity), tells us that the general integral of (10.62) is

$$
\begin{aligned}
\boldsymbol{y}(t) &= c_1 \mathrm{e}^t \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + c_2 \mathrm{e}^{2t} \left[\begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \cos t - \begin{pmatrix} 0 \\ 1 \\ 4 \end{pmatrix} \sin t \right] + \\
&\qquad\qquad + c_3 \mathrm{e}^{2t} \left[\begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \sin t + \begin{pmatrix} 0 \\ 1 \\ 4 \end{pmatrix} \cos t \right] \\
&= c_1 \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \mathrm{e}^t + \left[c_2 \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} - c_3 \begin{pmatrix} 0 \\ 1 \\ 4 \end{pmatrix} \right] \mathrm{e}^{2t} \cos t + \\
&\qquad\qquad + \left[-c_2 \begin{pmatrix} 0 \\ 1 \\ 4 \end{pmatrix} + c_3 \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} \right] \mathrm{e}^{2t} \sin t \\
&= \begin{pmatrix} c_1 \mathrm{e}^t + c_2 \mathrm{e}^{2t} \cos t + c_3 \mathrm{e}^{2t} \sin t \\ (c_2 - c_3) \mathrm{e}^{2t} \cos t + (c_3 - c_2) \mathrm{e}^{2t} \sin t \\ (c_2 - 4c_3) \mathrm{e}^{2t} \cos t + (2c_3 - 4c_2) \mathrm{e}^{2t} \sin t \end{pmatrix} .
\end{aligned}
$$

□

Explanation. Since eigenvalues and eigenvectors of $\boldsymbol{A}$ may not be real, it is convenient to think (10.62) within the complex field. Thus we view $\boldsymbol{y} = \boldsymbol{y}(t)$ as a map from $\mathbb{R}$ to $\mathbb{C}^n$; the real part $\boldsymbol{y}_{\mathrm{r}}(t)$ and the imaginary part $\boldsymbol{y}_{\mathrm{i}}(t)$ solve the real systems

$$\boldsymbol{y}_{\mathrm{r}}' = \boldsymbol{A}\boldsymbol{y}_{\mathrm{r}} \qquad \text{and} \qquad \boldsymbol{y}_{\mathrm{i}}' = \boldsymbol{A}\boldsymbol{y}_{\mathrm{i}} \qquad \text{on } I = \mathbb{R} .$$

All subsequent operations are intended over $\mathbb{C}$ (clearly, complex arithmetic is not necessary when $\boldsymbol{A}$ is diagonalisable over $\mathbb{R}$, when all eigenvalues–hence all eigenvectors–are real). Call $\boldsymbol{\Lambda} = \mathbf{diag}(\lambda_1, \ldots, \lambda_n)$ the diagonal matrix with the eigenvalues as entries, and $\boldsymbol{P} = (\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n)$ the square matrix with the eigenvectors as columns. Thus $\boldsymbol{A}$ becomes

$$\boxed{\boldsymbol{A} = \boldsymbol{P}\boldsymbol{\Lambda}\boldsymbol{P}^{-1} .} \tag{10.64}$$

Let $\boldsymbol{y}$ be an arbitrary complex solution of (10.62); substituting (10.64) in equation (10.62), and multiplying by $\boldsymbol{P}^{-1}$ on the left, gives

$$\left(\boldsymbol{P}^{-1}\boldsymbol{y}\right)' = \boldsymbol{\Lambda}\left(\boldsymbol{P}^{-1}\boldsymbol{y}\right) ,$$

by associativity. Now setting

$$\boxed{\boldsymbol{z} = \boldsymbol{P}^{-1}\boldsymbol{y} , \quad \text{i.e.,} \quad \boldsymbol{y} = \boldsymbol{P}\boldsymbol{z} ,} \tag{10.65}$$

equation (10.62) becomes diagonal

$$\boxed{\boldsymbol{z}' = \boldsymbol{\Lambda}\boldsymbol{z}\,.} \tag{10.66}$$

In other words we obtain n equations

$$z_k' = \lambda_k z_k\,, \qquad 1 \le k \le n\,.$$

By Remark 10.5 these have solutions

$$z_k(t) = d_k \mathrm{e}^{\lambda_k t}\,, \qquad 1 \le k \le n\,,$$

with $d_k \in \mathbb{C}$ arbitrary constants. Using the diagonal matrix

$$\boxed{\mathrm{e}^{\boldsymbol{\Lambda} t} = \mathbf{diag}\big(\mathrm{e}^{\lambda_1 t}, \ldots, \mathrm{e}^{\lambda_n t}\big)} \tag{10.67}$$

and the constant vector $\boldsymbol{d} = (d_1, \ldots, d_n)^T \in \mathbb{C}^n$, we can write the solutions as

$$\boldsymbol{z}(t) = \mathrm{e}^{\boldsymbol{\Lambda} t}\boldsymbol{d}\,.$$

Writing in terms of the unknown $\boldsymbol{y}$, by the second equation in (10.65), we have

$$\boxed{\boldsymbol{y}(t) = \boldsymbol{P}\mathrm{e}^{\boldsymbol{\Lambda} t}\boldsymbol{d}\,;} \tag{10.68}$$

it is of the general form (10.53) with $\boldsymbol{W}(t) = \boldsymbol{P}\mathrm{e}^{\boldsymbol{\Lambda} t}$. Making the columns of this matrix explicit, we have

$$\boldsymbol{y}(t) = d_1\mathrm{e}^{\lambda_1 t}\boldsymbol{v}_1 + \cdots + d_n\mathrm{e}^{\lambda_n t}\boldsymbol{v}_n = d_1\boldsymbol{w}_1(t) + \cdots + d_n\boldsymbol{w}_n(t)\,, \tag{10.69}$$

where

$$\boldsymbol{w}_k(t) = \mathrm{e}^{\lambda_k t}\boldsymbol{v}_k\,, \qquad 1 \le k \le n\,.$$

The upshot is that every complex solution of equation (10.62) is a linear combination (with coefficients in $\mathbb{C}$) of the n solutions $\boldsymbol{w}_k$, which therefore are a fundamental system in $\mathbb{C}$.

To represent the real solutions, consider the generic eigenvalue λ with eigenvector $\boldsymbol{v}$. Note that if $\lambda \in \mathbb{R}$ (so $\boldsymbol{v} \in \mathbb{R}^n$), the function $\boldsymbol{w}(t) = \mathrm{e}^{\lambda t}\boldsymbol{v}$ is a real solution of (10.62). Instead, if $\lambda = \sigma + i\omega \in \mathbb{C}$, so $\boldsymbol{v} = \boldsymbol{v}^{(1)} + i\boldsymbol{v}^{(2)} \in \mathbb{C}^n$, the functions

$$\boldsymbol{w}^{(1)}(t) = \mathcal{R}e\,\big(\mathrm{e}^{\lambda t}\boldsymbol{v}\big) = \mathrm{e}^{\sigma t}\big(\boldsymbol{v}^{(1)}\cos\omega t - \boldsymbol{v}^{(2)}\sin\omega t\big)\,,$$

$$\boldsymbol{w}^{(2)}(t) = \mathcal{I}m\,\big(\mathrm{e}^{\lambda t}\boldsymbol{v}\big) = \mathrm{e}^{\sigma t}\big(\boldsymbol{v}^{(1)}\sin\omega t + \boldsymbol{v}^{(2)}\cos\omega t\big)\,,$$

are real solutions of (10.62); this is clear by taking real and imaginary parts of the equation and remembering $\boldsymbol{A}$ is real.

In such a way, retaining the notation of Proposition 10.32, we obtain n real solutions

$$\boldsymbol{w}_1(t), \ldots, \boldsymbol{w}_\ell(t) \quad \text{and} \quad \boldsymbol{w}_1^{(1)}(t), \boldsymbol{w}_1^{(2)}(t), \ldots, \boldsymbol{w}_m^{(1)}(t), \boldsymbol{w}_m^{(2)}(t)\,.$$

There remains to verify they are linearly independent (over $\mathbb{R}$). With the help of Proposition 10.28 with $t_0 = 0$, it suffices to show $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_\ell, \boldsymbol{v}_1^{(1)}, \boldsymbol{v}_1^{(2)}, \ldots, \boldsymbol{v}_m^{(1)}, \boldsymbol{v}_m^{(2)}$ are linearly independent over $\mathbb{R}$. An easy Linear Algebra exercise will show that this fact is a consequence of the linear independence of the eigenvectors $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n$ over $\mathbb{C}$ (actually, it is equivalent).

Remark 10.34 Imposing the datum $\boldsymbol{y}(0) = \boldsymbol{y}_0$ for $t_0 = 0$ in (10.68), solution of $\boldsymbol{y}' = \boldsymbol{A}\boldsymbol{y}$, we find $\boldsymbol{P}\boldsymbol{d} = \boldsymbol{y}_0$, so $\boldsymbol{d} = \boldsymbol{P}^{-1}\boldsymbol{y}_0$. The solution of the Cauchy problem

$$\begin{cases} \boldsymbol{y}' = \boldsymbol{A}\boldsymbol{y} & \text{on } I = \mathbb{R}\,, \\ \boldsymbol{y}(0) = \boldsymbol{y}_0 \end{cases}$$

can therefore be represented as

$$\boldsymbol{y}(t) = \mathrm{e}^{\boldsymbol{A}t}\boldsymbol{y}_0\,, \tag{10.70}$$

where we have used the **exponential matrix**

$$\boxed{\mathrm{e}^{\boldsymbol{A}t} = \boldsymbol{P}\mathrm{e}^{\boldsymbol{\Lambda}t}\boldsymbol{P}^{-1}}$$

The above formula generalises the expression $y(t) = \mathrm{e}^{at}y_0$ for the solution to the Cauchy problem $y' = ay$, $y(0) = y_0$.

The representation (10.70) is valid also when $\boldsymbol{A}$ is not diagonalisable. In that case though, the exponential matrix is defined otherwise

$$\boxed{\mathrm{e}^{\boldsymbol{A}t} = \sum_{k=0}^{\infty} \frac{1}{k!}\,(t\boldsymbol{A})^k\ .}$$

This formula is suggested by the power series (2.24) of the exponential function e^x; the series converges for any matrix $\boldsymbol{A}$ and for any $t \in \mathbb{R}$. □

10.6.2 Homogeneous systems with non-diagonalisable $\boldsymbol{A}$

By $\lambda_1, \ldots, \lambda_p$ we number the *distinct* eigenvalues of $\boldsymbol{A}$, while μ_k denotes the algebraic multiplicity of λ_k, i.e., the multiplicity as root of the characteristic polynomial $\det(\boldsymbol{A}-\lambda\boldsymbol{I})$. Then $n = \mu_1+\ldots+\mu_p$. Call $m_k \le \mu_k$ the geometric multiplicity of λ_k, that is the maximum number of linearly independent eigenvectors $\boldsymbol{v}_{k,1}, \ldots, \boldsymbol{v}_{k,m_k}$ relative to λ_k. We remind $\boldsymbol{A}$ is diagonalisable if and only if the algebraic and geometric multiplicities of every eigenvalue coincide; from now on we suppose $m_k < \mu_k$ for at least one λ_k. Then it is possible to prove there are $d_k = \mu_k - m_k$ vectors $\boldsymbol{r}_{k,1}, \ldots, \boldsymbol{r}_{k,d_k}$, called **generalised eigenvectors** associated to λ_k, such that the μ_k vectors $\boldsymbol{v}_{k,1}, \ldots, \boldsymbol{v}_{k,m_k}, \boldsymbol{r}_{k,1}, \ldots, \boldsymbol{r}_{k,d_k}$ are linearly independent. Moreover, the collection of eigenvectors and generalised eigenvectors, for $k = 1, \ldots, p$, builds a basis of $\mathbb{C}^n$.

Let us see how to construct generalised eigenvectors associated to λ_k. For every eigenvector $\boldsymbol{v}_{k,\ell}$, $1 \le \ell \le m_k$, define $\boldsymbol{r}^{(0)}_{k,\ell} = \boldsymbol{v}_{k,\ell}$ and seek whether a solution of

$$\boxed{(\boldsymbol{A} - \lambda_k \boldsymbol{I})\boldsymbol{r} = \boldsymbol{r}^{(0)}_{k,\ell}} \tag{10.71}$$

exists. The matrix $\boldsymbol{A} - \lambda_k \boldsymbol{I}$ is singular, by definition of eigenvalue. The system may not have solutions, in which case the eigenvector $\boldsymbol{v}_{k,\ell}$ will not furnish generalised eigenvectors. If however the system is consistent, a solution $\boldsymbol{r}^{(1)}_{k,\ell}$ will be a generalised eigenvector associated to λ_k. Now we substitute $\boldsymbol{r}^{(1)}_{k,\ell}$ to $\boldsymbol{r}^{(0)}_{k,\ell}$ in (10.71), and repeat the argument. This will produce a cascade of linear systems

$$\begin{aligned} (\boldsymbol{A} - \lambda_k \boldsymbol{I})\boldsymbol{r}^{(1)}_{k,\ell} &= \boldsymbol{r}^{(0)}_{k,\ell}\,, \\ (\boldsymbol{A} - \lambda_k \boldsymbol{I})\boldsymbol{r}^{(2)}_{k,\ell} &= \boldsymbol{r}^{(1)}_{k,\ell}\,, \\ &\vdots \end{aligned} \tag{10.72}$$

that stops once the system $(\boldsymbol{A} - \lambda_k \boldsymbol{I})\boldsymbol{r} = \boldsymbol{r}^{(h)}_{k,\ell}$, $h \ge 0$, has no solutions. As a result, we obtain $q_{k,\ell} = h - 1$ generalised eigenvectors. Varying $\ell = 1, \ldots, m_k$ will yield all the $d_k = q_{k,1} + \ldots + q_{k,m_k}$ generalised eigenvectors relative to λ_k.

Let us now move to fundamental systems of solutions for (10.62).

- For every eigenvalue λ_k $(1 \le k \le p)$ and associated eigenvector $\boldsymbol{v}_{k,\ell}$ $(1 \le \ell \le m_k)$, define

$$\boxed{\boldsymbol{w}^{(0)}_{k,\ell}(t) = \mathrm{e}^{\lambda_k t}\boldsymbol{v}_{k,\ell}\,.}$$

- If $q_{k,\ell} > 0$, so if there are generalised eigenvectors $\boldsymbol{r}^{(h)}_{k,\ell}$ $(1 \le h \le q_{k,\ell})$, build maps

$$\begin{aligned} \boldsymbol{w}^{(1)}_{k,\ell}(t) &= \mathrm{e}^{\lambda_k t}\big(t\,\boldsymbol{r}^{(0)}_{k,\ell} + \boldsymbol{r}^{(1)}_{k,\ell}\big) = \mathrm{e}^{\lambda_k t}\big(t\,\boldsymbol{v}_{k,\ell} + \boldsymbol{r}^{(1)}_{k,\ell}\big)\,, \\ \boldsymbol{w}^{(2)}_{k,\ell}(t) &= \mathrm{e}^{\lambda_k t}\big(\frac{1}{2}t^2\boldsymbol{r}^{(0)}_{k,\ell} + t\,\boldsymbol{r}^{(1)}_{k,\ell} + \boldsymbol{r}^{(2)}_{k,\ell}\big)\,, \end{aligned}$$

and so on; in general, we set

$$\boxed{\boldsymbol{w}^{(h)}_{k,\ell}(t) = \mathrm{e}^{\lambda_k t}\sum_{j=0}^{h}\frac{1}{(h-j)!}\,t^{h-j}\boldsymbol{r}^{(j)}_{k,\ell}\,, \qquad 1 \le h \le q_{k,\ell}\,.}$$

Proposition 10.35 *With the above notations, the set of functions* $\{\boldsymbol{w}^{(h)}_{k,\ell} : 1 \le k \le p,\ 1 \le \ell \le m_k,\ 0 \le h \le q_{k,\ell}\}$ *is a fundamental system of solutions over* $\mathbb{C}$ *of* (10.62).

Hence the general integral reads

$$\boldsymbol{y}(t) = \sum_{k=1}^{p} \sum_{\ell=1}^{m_k} \sum_{h=0}^{q_{k,\ell}} c_{k\ell}^{(h)} \boldsymbol{w}_{k,\ell}^{(h)}(t)\,, \tag{10.73}$$

with $c_{k\ell}^{(h)} \in \mathbb{C}$.

If we wish to represent real solutions only, observe that $\lambda_k \in \mathbb{R}$ implies all eigenvectors (proper and generalised) are real, hence also the maps $\boldsymbol{w}_{k,\ell}^{(h)}(t)$, built from them, are real. In presence of a complex-conjugate pair λ_k, $\lambda_{k'} = \overline{\lambda}_k$ of eigenvalues, the corresponding eigenvectors (generalised or not) will crop up in conjugate pairs. For that reason it is enough to replace each pair of complex-conjugate functions, defined by those vectors, with their real and imaginary parts. The construction of a real fundamental system for (10.62) is thus complete.

Remark 10.36 We have assumed $\boldsymbol{A}$ not diagonalisable. But actually the above recipe works even when the matrix can be diagonalised (when the algebraic multiplicity is greater than 1, one cannot know a priori whether a matrix is diagonalisable or not, except in a few cases, e.g., symmetric matrices). If the matrix is diagonalisable, systems (10.71) will be inconsistent, rendering (10.73) equivalent to the solution of Sect. 10.6.1. □

Examples 10.37

i) The matrix

$$\boldsymbol{A} = \begin{pmatrix} 4 & 0 & -1 \\ 1 & 5 & 1 \\ 1 & 0 & 2 \end{pmatrix}$$

has eigenvalues $\lambda_1 = 5$, $\lambda_2 = 3$. The first has algebraic multiplicity $\mu_1 = 1$, and $\boldsymbol{v}_{11} = (0,1,0)^T$ is the associated eigenvector. The second eigenvalue has algebraic multiplicity $\mu_2 = 2$ and geometric $m_2 = 1$. The vector $\boldsymbol{v}_{21} = (1,-1,1)^T$ is associated to λ_2. By solving $(\boldsymbol{A}-3\boldsymbol{I})\boldsymbol{r} = \boldsymbol{v}_{21}$ we obtain a generalised eigenvector $\boldsymbol{r}_{21}^{(1)} = (1,-1,0)^T$ which, together with $\boldsymbol{v}_{11}$, $\boldsymbol{v}_{21}$, gives a basis of $\mathbb{R}^3$. Therefore, the general integral of (10.62) is

$$\boldsymbol{y}(t) = c_1 \mathrm{e}^{5t} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + c_2 \mathrm{e}^{3t} \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} + c_3 \mathrm{e}^{3t} \left[t \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \right].$$

ii) The matrix

$$\boldsymbol{A} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 1 & 0 & 3 \end{pmatrix}$$

has a unique eigenvalue $\lambda = 3$ with geometric multiplicity $m = 2$, in fact $\boldsymbol{v}_{11} = (0,1,0)^T$ and $\boldsymbol{v}_{12} = (0,0,1)^T$ are linearly independent eigenvectors associated to it. The system

$$(\boldsymbol{A} - 3\boldsymbol{I})\boldsymbol{r} = \boldsymbol{v}_{11}$$

has no solutions, whereas

$$(\boldsymbol{A} - 3\boldsymbol{I})\boldsymbol{r} = \boldsymbol{v}_{12}$$

gives, for instance, $\boldsymbol{r}_{12}^{(1)} = (1,0,0)^T$.
Hence, the general integral of (10.62) is

$$\boldsymbol{y}(t) = c_1 \mathrm{e}^{3t} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + c_2 \mathrm{e}^{3t} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + c_3 \mathrm{e}^{3t} \left[t \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right].$$

iii) Consider

$$\boldsymbol{A} = \begin{pmatrix} 1 & 5 & 0 \\ 0 & 1 & 0 \\ 4 & 0 & 1 \end{pmatrix}$$

with one eigenvalue $\lambda = 1$ and one eigenvector $\boldsymbol{v}_{11} = (0,0,1)^T$ (geometric multiplicity $m = 1$). As the set of all eigenvectors is a basis of $\mathbb{R}^3$, equations (10.72) will spawn two generalised eigenvectors. In fact

$$(\boldsymbol{A} - \boldsymbol{I})\boldsymbol{r}_{11}^{(1)} = \boldsymbol{v}_{11}$$
$$(\boldsymbol{A} - \boldsymbol{I})\boldsymbol{r}_{11}^{(2)} = \boldsymbol{r}_{11}^{(1)}$$

are solved by $\boldsymbol{r}_{11}^{(1)} = (1/4,0,0)^T$ and $\boldsymbol{r}_{11}^{(2)} = (0,1/20,0)^T$. The general integral reads

$$\boldsymbol{y}(t) = c_1 \mathrm{e}^{t} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + c_2 \mathrm{e}^{t} \left[t \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 1/4 \\ 0 \\ 0 \end{pmatrix} \right] +$$

$$+ c_3 \mathrm{e}^{t} \left[\frac{1}{2} t^2 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + t \begin{pmatrix} 1/4 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 1/20 \\ 0 \end{pmatrix} \right].$$

□

Explanation. All the results about eigenvectors, both proper and generalised, can be proved starting from the so-called **Jordan canonical form** of a matrix, whose study goes beyond the reach of the course.

What we can easily do is account for the linear independence of the maps $\boldsymbol{w}_{k,\ell}^{(h)}(t)$ of Proposition 10.35. Taking $t_0 = 0$ in Proposition 10.28 gives

$$\boldsymbol{w}_{k,\ell}^{(h)}(0) = \boldsymbol{r}_{k,\ell}^{(h)}$$

for any k, ℓ, h, so the result follows from the analogue property for $\boldsymbol{A}$.

10.6.3 Non-homogeneous systems

In the light of Proposition 10.29, a non-homogeneous equation

$$\boxed{\boldsymbol{y}' = \boldsymbol{A}\boldsymbol{y} + \boldsymbol{b}(t)} \tag{10.74}$$

can be solved by simply finding a particular integral $\boldsymbol{y}_p$.

The method of variation of parameters, and especially formula (10.61), provides us with a general means. However, in many situations the source term $\boldsymbol{b}(t)$ can be broken into elementary functions like exponentials, polynomials or trigonometric maps. Then a particular integral of the same elementary type is usually easy to find.

Henceforth we shall call a polynomial any function $\boldsymbol{p}(t)$, $\boldsymbol{q}(t), \ldots$ from $\mathbb{R}$ to $\mathbb{R}^n$ whose components are real algebraic polynomials; a polynomial has degree m if the maximum degree of the components is m, in which case it can be written as

$$\boldsymbol{q}(t) = \boldsymbol{c}_0 t^m + \boldsymbol{c}_1 t^{m-1} + \ldots + \boldsymbol{c}_{m-1} t + \boldsymbol{c}_m\,,$$

with $\boldsymbol{c}_j \in \mathbb{R}^n$ and $\boldsymbol{c}_0 \neq \boldsymbol{0}$.

Let us suppose that

$$\boxed{\boldsymbol{b}(t) = \mathrm{e}^{\alpha t}\boldsymbol{p}(t)} \tag{10.75}$$

with $\alpha \in \mathbb{R}$ and $\boldsymbol{p}(t)$ a degree-m polynomial. Then there exists a particular integral

$$\boxed{\boldsymbol{y}_p(t) = \mathrm{e}^{\alpha t}\boldsymbol{q}(t)\,,} \tag{10.76}$$

where $\boldsymbol{q}(t)$ is a polynomial of degree

- m if α is not an eigenvalue of $\boldsymbol{A}$,
- $m + \mu$ if α is an eigenvalue of algebraic multiplicity $\mu \geq 1$.

Borrowing from Physics, one refers to the latter situation as **resonance**.

The undetermined coefficients of $\boldsymbol{q}(t)$ are found by substituting (10.76) in equation (10.74), simplifying the exponential terms and matching the coefficients of the corresponding powers of t. This produces a series of linear systems with matrix $\boldsymbol{A}-\alpha\boldsymbol{I}$ that allow to determine solutions $\boldsymbol{c}_0, \boldsymbol{c}_1, \ldots, \boldsymbol{c}_{m+\mu}$ starting from the highest power of t. In case of resonance, $\boldsymbol{A} - \alpha\boldsymbol{I}$ is singular; then the first system just says that $\boldsymbol{c}_0$ is an eigenvector of $\boldsymbol{A}$, and the other systems require compatibility conditions to be solved. Example 10.38 ii) illustrates a situation of this type, detailing the computation to be performed.

If the source term looks like

$$\boxed{\boldsymbol{b}(t) = \mathrm{e}^{\alpha t}\boldsymbol{p}(t)\cos\omega t \qquad \text{or} \qquad \boldsymbol{b}(t) = \mathrm{e}^{\alpha t}\boldsymbol{p}(t)\sin\omega t\,,} \tag{10.77}$$

with $\alpha \in \mathbb{R}$, $\boldsymbol{p}(t)$ a polynomial of degree m, $\omega \neq 0$, then a particular integral will be

$$\boxed{\boldsymbol{y}_p(t) = \mathrm{e}^{\alpha t}\big(\boldsymbol{q}_1(t)\cos\omega t + \boldsymbol{q}_2(t)\sin\omega t\big)} \tag{10.78}$$

with $\boldsymbol{q}_1(t)$, $\boldsymbol{q}_2(t)$ polynomials of degree

- m if $\alpha + i\omega$ is not eigenvalue of $\boldsymbol{A}$,
- $m+\mu$ if $\alpha + i\omega$ is an eigenvalue of algebraic multiplicity $\mu \geq 1$.

The latter case is once again called of resonance. The polynomials' undetermined coefficients are found as before, preliminarily separating terms with $\cos\omega t$ from those with $\sin\omega t$.

Examples 10.38

i) Find a particular integral of (10.74), where

$$\boldsymbol{A} = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 2 \end{pmatrix} \qquad \text{and} \qquad \boldsymbol{b}(t) = \begin{pmatrix} 0 \\ t \\ t^2 \end{pmatrix}.$$

Referring to (10.75), $\alpha = 0$, $\boldsymbol{p}(t) = \boldsymbol{b}_0 t^2 + \boldsymbol{b}_1 t$ with $\boldsymbol{b}_0 = (0,0,1)^T$, $\boldsymbol{b}_1 = (0,1,0)^T$ (and $\boldsymbol{b}_2 = \boldsymbol{0}$). Since α is not a root of the characteristic polynomial $\chi(\lambda) = (2-\lambda)(\lambda^2+1)$, we look for a particular integral

$$\boldsymbol{y}_p(y) = \boldsymbol{q}(t) = \boldsymbol{c}_0 t^2 + \boldsymbol{c}_1 t + \boldsymbol{c}_2\,.$$

Substituting in (10.74), we have

$$2\boldsymbol{c}_0 t + \boldsymbol{c}_1 = \boldsymbol{A}\boldsymbol{c}_0 t^2 + \boldsymbol{A}\boldsymbol{c}_1 t + \boldsymbol{A}\boldsymbol{c}_2 + \boldsymbol{b}_0 t^2 + \boldsymbol{b}_1 t\,,$$

so comparing terms yields the cascade of systems

$$\begin{cases} \boldsymbol{A}\boldsymbol{c}_0 = -\boldsymbol{b}_0 \\ \boldsymbol{A}\boldsymbol{c}_1 = 2\boldsymbol{c}_0 - \boldsymbol{b}_1 \\ \boldsymbol{A}\boldsymbol{c}_2 = \boldsymbol{c}_1\,. \end{cases}$$

These give $\boldsymbol{c}_0 = (0,0,-1/2)^T$, $\boldsymbol{c}_1 = (-1,0,0)^T$ and $\boldsymbol{c}_2 = (0,1,0)^T$, and a particular integral is then

$$\boldsymbol{y}_p(t) = \begin{pmatrix} 0 \\ 0 \\ -1/2 \end{pmatrix} t^2 + \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix} t + \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -t \\ 1 \\ -\frac{1}{2}t^2 \end{pmatrix}.$$

ii) Consider equation (10.74) with

$$\boldsymbol{A} = \begin{pmatrix} 9 & -4 \\ 8 & -3 \end{pmatrix} \qquad \text{and} \qquad \boldsymbol{b}(t) = \mathrm{e}^t \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

As $\alpha = 1$ is a root of the characteristic polynomial $\chi(\lambda) = (\lambda-5)(\lambda-1)$, we are in presence of resonance, so we must try to find a particular integral of the form

$$\boldsymbol{y}_p(t) = \mathrm{e}^t(\boldsymbol{c}_0 t + \boldsymbol{c}_1)\,.$$

Substituting and simplifying, we have

$$\boldsymbol{c}_0 + \boldsymbol{c}_0 t + \boldsymbol{c}_1 = \boldsymbol{A}\boldsymbol{c}_0 t + \boldsymbol{A}\boldsymbol{c}_1 + \boldsymbol{b}_0$$

where $\boldsymbol{b}_0 = (1,0)^T$, and so

$$\begin{cases} (\boldsymbol{A}-\boldsymbol{I})\boldsymbol{c}_0 = \boldsymbol{0} \\ (\boldsymbol{A}-\boldsymbol{I})\boldsymbol{c}_1 = \boldsymbol{c}_0 - \boldsymbol{b}_0 \,. \end{cases}$$

The first system yields an eigenvector $\boldsymbol{c}_0 = (\gamma, 2\gamma)^T$ associated to $\lambda = 1$, for some $\gamma \in \mathbb{R}$. This constant γ is fixed requiring the second system to be consistent, meaning $\boldsymbol{c}_0 - \boldsymbol{b}_0$ is a linear combination of the columns of $\boldsymbol{A}-\boldsymbol{I}$. But that matrix has rank 1 since the columns are linearly dependent, so the condition is that $\boldsymbol{c}_0 - \boldsymbol{b}_0$ is a multiple of another column, $(\gamma-1, 2\gamma)^T = k(1,1)^T$; this implies $\gamma - 1 = 2\gamma$, so $\gamma = -1$. Substituting gives $\boldsymbol{c}_0 = (-1,-2)^T$ and $\boldsymbol{c}_1 = (1, 5/2)^T$, for example. In conclusion, a particular integral has the form

$$\boldsymbol{y}_p(t) = \mathrm{e}^t \left[\begin{pmatrix} -1 \\ -2 \end{pmatrix} t + \begin{pmatrix} 1 \\ 5/2 \end{pmatrix} \right] = \mathrm{e}^t \begin{pmatrix} -t+1 \\ -2t+5/2 \end{pmatrix} .$$

iii) Find a particular integral of (10.74) with

$$\boldsymbol{A} = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & -1 \end{pmatrix} \qquad \text{and} \qquad \boldsymbol{b}(t) = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \sin 2t \,.$$

Referring to (10.77), now $\alpha = 0$, $\boldsymbol{p}(t) = (1,0,0)^T = \boldsymbol{p}_0$ and $\omega = 2$. The complex number $\alpha + i\omega = 2i$ is no eigenvalue of $\boldsymbol{A}$ (these are -1, $\pm i$), so (10.78) will be

$$\boldsymbol{y}_p(t) = \boldsymbol{q}_1 \cos 2t + \boldsymbol{q}_2 \sin 2t$$

with $\boldsymbol{q}_1, \boldsymbol{q}_2 \in \mathbb{R}^3$. Let us substitute in (10.74) to get

$$-2\boldsymbol{q}_1 \sin 2t + 2\boldsymbol{q}_2 \cos 2t = \boldsymbol{A}\boldsymbol{q}_1 \cos 2t + \boldsymbol{A}\boldsymbol{q}_2 \sin 2t + \boldsymbol{p}_0 \sin 2t \,;$$

comparing the corresponding terms we obtain the two linear systems

$$\begin{cases} -2\boldsymbol{q}_1 = \boldsymbol{A}\boldsymbol{q}_2 + \boldsymbol{p}_0 \\ 2\boldsymbol{q}_2 = \boldsymbol{A}\boldsymbol{q}_1 \,, \end{cases}$$

solved by $\boldsymbol{q}_1 = (-2/3, 0, 1/6)^T$, $\boldsymbol{q}_2 = (0, -1/3, -1/12)^T$. We conclude that a particular integral is given by

$$y_p(t) = \begin{pmatrix} -2/3 \\ 0 \\ 1/6 \end{pmatrix} \cos 2t - \begin{pmatrix} 0 \\ 1/3 \\ 1/12 \end{pmatrix} \sin 2t = -\frac{1}{12} \begin{pmatrix} 8\cos 2t \\ 4 \sin 2t \\ \sin 2t - 2\cos 2t \end{pmatrix} .$$

□

The superposition principle

Ultimately, suppose $\boldsymbol{b}(t)$ is the sum of terms like (10.75) or (10.77). By virtue of linearity, a particular solution $\boldsymbol{y}_p$ will be the sum of particular solutions of the single summands:

> if $\boldsymbol{b} = \boldsymbol{b}_1 + \boldsymbol{b}_2 + \ldots + \boldsymbol{b}_K$, and $\boldsymbol{y}_{pk}$ solves $\boldsymbol{y}' = \boldsymbol{A}\boldsymbol{y} + \boldsymbol{b}_k$ for $k = 1, \ldots, K$, then $\boldsymbol{y}_p = \boldsymbol{y}_{p1} + \ldots + \boldsymbol{y}_{pK}$ solves $\boldsymbol{y}' = \boldsymbol{A}\boldsymbol{y} + \boldsymbol{b}$.

In fact,

$$\begin{aligned}\boldsymbol{y}_p' &= \boldsymbol{y}_{p1}' + \ldots + \boldsymbol{y}_{pK}' = (\boldsymbol{A}\boldsymbol{y}_{p1} + \boldsymbol{b}_1) + \ldots + (\boldsymbol{A}\boldsymbol{y}_{pK} + \boldsymbol{b}_K) \\ &= \boldsymbol{A}(\boldsymbol{y}_{p1} + \ldots + \boldsymbol{y}_{pK}) + (\boldsymbol{b}_1 + \ldots + \boldsymbol{b}_K) = \boldsymbol{A}\boldsymbol{y}_p + \boldsymbol{b}\,.\end{aligned}$$

The property is known as *superposition* (or *linearity*) *principle*. Example 10.42 ii) will provide us with a tangible application.

10.7 Linear scalar equations of order n

In this section we tackle linear scalar equations

$$\boxed{y^{(n)} + a_1 y^{(n-1)} + \ldots + a_{n-1}y' + a_n y = b(t)\,,} \tag{10.79}$$

where n is an integer ≥ 2, the coefficients $a_1, \ldots, a_n$ are real constants and b is a real-valued continuous map defined on the real interval I. We abbreviate by $\mathcal{L}y$ the left-hand side; the operator $\mathcal{L} : y \mapsto \mathcal{L}\, y$ is linear because differentiation is a linear operation. Then

$$\boxed{\mathcal{L}y = y^{(n)} + a_1 y^{(n-1)} + \ldots + a_{n-1}y' + a_n y = 0} \tag{10.80}$$

is called the homogeneous equation associated to (10.79).

The background theory can be deduced from the results on linear systems of first-order equations, settled in the previous section. In fact, we remarked in Sect. 10.2 that any ODE of order n is equivalent to a system of n differential equations of order one. In the case at hand, we set $y_i(t) = y^{(i-1)}(t)$, $1 \leq i \leq n$, so that (10.79) is equivalent to the linear system

$$\begin{cases} y_1' = y_2 \\ y_2' = y_3 \\ \quad\vdots \\ y_n' = -a_1 y_n - \ldots - a_{n-1}y_2 - a_n y_1 + b(t)\,; \end{cases}$$

the latter may be written as (10.48), and precisely

$$\boldsymbol{y}' = \boldsymbol{A}\boldsymbol{y} + \boldsymbol{b}(t)\,, \tag{10.81}$$

by putting

$$\boldsymbol{y} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \qquad \boldsymbol{A} = \begin{pmatrix} 0 & 1 & 0 & \ldots & \ldots \\ 0 & 0 & 1 & 0 & \ldots \\ & \vdots & & & \\ \ldots & \ldots & 0 & 0 & 1 \\ -a_n & -a_{n-1} & \ldots & \ldots & -a_1 \end{pmatrix}, \qquad \boldsymbol{b}(t) = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ b(t) \end{pmatrix}.$$

The first component of $\boldsymbol{y}$ is the solution of (10.79).

The homogeneous equation

Equation (10.80) corresponds to the homogeneous system relative to (10.81). From this we recover the following central result.

Proposition 10.39 *i) The set S_0 of solutions to the homogeneous equation* (10.80) *is an n-dimensional vector space.*
ii) The set S_b of solutions to (10.79) *is the affine space $S_b = y_p + S_0$, where y_p is any solution.*

Proof. i) S_0 is a vector space because (10.80) is a linear constraint. With Proposition 10.26 in mind, let $\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n$ be a fundamental system of solutions for $\boldsymbol{y}' = \boldsymbol{A}\boldsymbol{y}$. Denote by $z_i = \boldsymbol{w}_i \cdot \boldsymbol{e}_1$ the first components of the vectors, which clearly solve (10.80).

If $y \in S_0$, the associated vector-valued map $\boldsymbol{y}$ is a linear combination of the $\boldsymbol{w}_i$; in particular its first component is a combination of the z_i. Thus S_0 is spanned by those solutions. The claim follows provided we show that the functions $z_1, \ldots, z_n$ are linearly independent. By successively differentiating

$$\sum_{i=1}^{n} c_i z_i(t) = 0\,, \qquad \forall t \in \mathbb{R}\,,$$

we get, for all $1 \le k \le n-1$,

$$\sum_{i=1}^{n} c_i z_i^{(k)}(t) = 0\,, \qquad \forall t \in \mathbb{R}\,.$$

But $z_i^{(k)}$ is the $(k+1)$th component of $\boldsymbol{w}_i$, so we obtain the vector equation

$$\sum_{i=1}^{n} c_i \boldsymbol{w}_i(t) = 0\,, \qquad \forall t \in \mathbb{R}\,,$$

whence $c_1 = \ldots = c_n = 0$ by linear independence of the $\boldsymbol{w}_i$.

ii) The argument is similar to the one used in Proposition 10.29. □

In order to find a basis for S_0, the linear system associated to (10.80) is not necessary, because one can act directly on the equation: we are looking for solutions $y(t) = \mathrm{e}^{\lambda t}$, with λ constant, possibly complex. Substituting into (10.80), and recalling (10.28), gives

$$\mathcal{L}(\mathrm{e}^{\lambda t}) = (\lambda^n + a_1\lambda^{n-1} + \ldots + a_{n-1}\lambda + a_n)\,\mathrm{e}^{\lambda t} = 0\,,$$

i.e.,

$$\mathcal{L}(\mathrm{e}^{\lambda t}) = \chi(\lambda)\mathrm{e}^{\lambda t} = 0\,,$$

where $\chi(\lambda) = \lambda^n + a_1\lambda^{n-1} + \ldots + a_{n-1}\lambda + a_n$ is the **characteristic polynomial** of the ODE (10.79). Since $\mathrm{e}^{\lambda t}$ is always non-zero, $y(t) = \mathrm{e}^{\lambda t}$ solves the homogeneous equation if and only if λ is a root of the characteristic polynomial, that is if and only if λ satisfies the **characteristic equation**

$$\lambda^n + a_1\lambda^{n-1} + \ldots + a_{n-1}\lambda + a_n = 0\,. \tag{10.82}$$

Hence, the differential problem is reduced to a purely algebraic question, that the Fundamental Theorem of Algebra can handle. We know that (10.82) has p distinct solutions $\lambda_1, \ldots, \lambda_p$, $1 \le p \le n$; each root λ_k, $1 \le k \le p$, has multiplicity $\mu_k \ge 1$, so that overall $\mu_1 + \ldots + \mu_p = n$. The characteristic equation's zeroes are nothing but the eigenvalues of $\boldsymbol{A}$ in (10.81), because one could prove

$$\det(\boldsymbol{A} - \lambda\boldsymbol{I}) = (-1)^n\chi(\lambda)\,.$$

This gives directly p distinct solutions

$$\boxed{\mathrm{e}^{\lambda_1 t}, \ldots, \mathrm{e}^{\lambda_p t}}$$

to (10.80). If $p < n$, any root λ_k of multiplicity $\mu_k > 1$ gives $\mu_k - 1$ further solutions

$$\boxed{t\,\mathrm{e}^{\lambda_k t},\ t^2\,\mathrm{e}^{\lambda_k t}, \ldots,\ t^{\mu_k - 1}\,\mathrm{e}^{\lambda_k t}\,.}$$

The n solutions thus found can be proven to be linearly independent. In summary, the result reads as follows.

Proposition 10.40 *The functions*

$$z_{k,\ell}(t) = t^\ell e^{\lambda_k t}\,, \qquad 1 \le k \le p, \quad 0 \le \ell \le \mu_k - 1\,,$$

form a basis for the space S_0 of solutions to (10.80). *Equivalently, every solution of the equation has the form*

$$y(t) = \sum_{k=1}^{p} q_k(t)\,\mathrm{e}^{\lambda_k t}\,,$$

with q_k a polynomial of degree $\le \mu_k - 1$.

In presence of complex(-conjugate) roots of equation (10.82), the corresponding basis functions are complex-valued. But we can find a real basis if we replace the pair $t^\ell e^{\lambda_k t}$, $t^\ell e^{\overline{\lambda}_k t}$ with the real and imaginary parts of either of them, for any pair of complex-conjugate eigenvalues $\lambda_k = \sigma_k + i\omega_k$, $\overline{\lambda}_k = \sigma_k - i\omega_k$. That is to say,

$$\boxed{t^\ell e^{\sigma_k t}\cos\omega_k t \qquad \text{and} \qquad t^\ell e^{\sigma_k t}\sin\omega_k t}$$

are n linearly independent, real solutions.

Examples 10.41

i) Let us make the previous construction truly explicit for an equation of order two,

$$y'' + a_1 y' + a_2 y = 0\,. \tag{10.83}$$

Let $\Delta = a_1^2 - 4a_2$ be the discriminant of $\lambda^2 + a_1\lambda + a_2 = 0$.
When $\Delta > 0$, there are two distinct real roots $\lambda_{1,2} = (-a_1 \pm \sqrt{\Delta})/2$, so the generic solution to (10.83) is

$$y(t) = c_1\,\mathrm{e}^{\lambda_1 t} + c_2\,\mathrm{e}^{\lambda_2 t}\,. \tag{10.84}$$

When $\Delta = 0$, the real root $\lambda_1 = -a_1/2$ is double, $\mu_1 = 2$; so, the generic solution reads

$$y(t) = (c_1 + c_2 t)\,\mathrm{e}^{\lambda_1 t}\,. \tag{10.85}$$

Eventually, when $\Delta < 0$ we have complex-conjugate roots $\lambda_1 = \sigma + i\omega = -a/2 + i\sqrt{|\Delta|}$ and $\lambda_2 = \sigma - i\omega = -a/2 - i\sqrt{|\Delta|}$, and the solution to (10.83) is

$$y(t) = \mathrm{e}^{\sigma t}(c_1 \cos\omega t + c_2 \sin\omega t)\,. \tag{10.86}$$

ii) Solve the homogeneous equation of fourth order

$$y^{(4)} + y = 0\,.$$

The characteristic zeroes, solving $\lambda^4 + 1 = 0$, are fourth roots of -1, $\lambda_{1,2,3,4} = \frac{\sqrt{2}}{2}(\pm 1 \pm i)$. Therefore y takes the form

$$y(t) = \mathrm{e}^{(\sqrt{2}/2)\,t}\big(c_1 \cos\frac{\sqrt{2}}{2}t + c_2 \sin\frac{\sqrt{2}}{2}t\big) + \mathrm{e}^{-(\sqrt{2}/2)\,t}\big(c_3 \cos\frac{\sqrt{2}}{2}t + c_4 \sin\frac{\sqrt{2}}{2}t\big)\,.$$

□

The non-homogeneous equation

Just as for systems of order one, a particular integral is easy to find when $b(t)$ has a special form. For instance, for

$$\boxed{b(t) = \mathrm{e}^{\alpha t} p(t)} \tag{10.87}$$

with $\alpha \in \mathbb{R}$, $p(t)$ an algebraic polynomial of degree m with real coefficients, there is a particular integral

$$\boxed{y_p(t) = \mathrm{e}^{\alpha t} t^{\mu} q(t)\,,} \tag{10.88}$$

where $\mu \geq 0$ is the multiplicity of α as a zero of $\chi(\lambda) = 0$ (with $\mu = 0$ if α is not a root, while $\mu \geq 1$ gives *resonance*), and $q(t)$ is a polynomial of degree m with unknown coefficients. These coefficients are determined by substituting (10.88) in (10.79), simplifying the common factor $\mathrm{e}^{\alpha t}$ and comparing the polynomial functions in t.

Take another example, like

$$\boxed{b(t) = \mathrm{e}^{\alpha t} p(t) \cos\omega t \qquad \text{or} \qquad b(t) = \mathrm{e}^{\alpha t} p(t) \sin\omega t\,,} \tag{10.89}$$

with $\alpha \in \mathbb{R}$, $p(t)$ a real algebraic polynomial of degree m, $\omega \neq 0$. This gives a particular integral

$$\boxed{y_p(t) = \mathrm{e}^{\alpha t} t^{\mu} \big(q_1(t)\cos\omega t + q_2(t)\sin\omega t\big)\,,} \tag{10.90}$$

where $\mu \geq 0$ is the multiplicity of $\alpha + i\omega$, and q_1, q_2 are unknown polynomials of degree m to be found as above, i.e., separating the terms in $\cos\omega t$ and $\sin\omega t$.

The *superposition principle* is still valid if $b(t)$ is a sum of terms $b_k(t)$ like (10.87) or (10.89): a particular integral for (10.79) will be a sum of particular integrals for the $b_k(t)$.

Examples 10.42

i) Determine the general integral of

$$y'' - y' - 6y = t\mathrm{e}^{-2t}\,.$$

The characteristic equation $\lambda^2 - \lambda - 6 = 0$ has roots $\lambda_1 = -2$, $\lambda_2 = 3$, so the general homogeneous integral is

$$y_{\mathrm{hom}}(t) = c_1\mathrm{e}^{-2t} + c_2\mathrm{e}^{3t}\,.$$

In a situation of resonance between the source term and a component of y_{hom}, the particular integral has to be of type

$$y_p(t) = t(at+b)\mathrm{e}^{-2t}\,.$$

We differentiate and substitute back in the ODE to obtain

$$-10at + 2a - 5b = t$$

(the coefficient of t^2 on the left is zero, as a consequence of resonance), from which $-10a = 1$ and $2a - 5b = 0$, so $a = -1/10$, $b = -1/25$. In conclusion,

$$y(t) = \Big(-\frac{1}{10}t^2 - \frac{1}{25}t + c_1\Big)\mathrm{e}^{-2t} + c_2\mathrm{e}^{3t}\,.$$

ii) Let us find the general integral of

$$y''' + y' = \cos t - 2\mathrm{e}^{3t}\,.$$

By superposition

$$y(t) = y_{\mathrm{hom}}(t) + y_{p_1}(t) - 2y_{p_2}(t)\,,$$

where y_{hom} is the general homogeneous integral, y_{p_1} and y_{p_2} are particular integrals relative to sources $\cos t$ and e^{3t}.

The characteristic equation $\lambda^3 + \lambda = 0$ has roots $\lambda_1 = 0$, $\lambda_{2,3} = \pm i$, hence

$$y_{\mathrm{hom}}(t) = c_1 + c_2\cos t + c_3\sin t\,.$$

Taking resonance into account, we want y_{p_1} of the form $y_{p_1}(t) = t(a\cos t + b\sin t)$. Computing successive derivatives of y_{p_1} and substituting them into

$$y_{p_1}''' + y_{p_1}' = \cos t$$

gives $-2a\cos t - 2b\sin t = \cos t$, so $a = -1/2$ and $b = 0$. Therefore $y_{p_1}(t) = -\frac{1}{2}\cos t$.

Now we search for y_{p_2} of the form $y_{p_2}(t) = d\mathrm{e}^{3t}$. By differentiating and substituting in the equation

$$y'''_{p_2} + y'_{p_2} = \mathrm{e}^{3t},$$

we obtain $d = 1/30$, and then $y_{p_2}(t) = \frac{1}{30}\mathrm{e}^{3t}$. All-in-all,

$$y(t) = c_1 + \big(c_2 - \frac{1}{2}t\big)\cos t + c_3 \sin t - \frac{1}{15}\mathrm{e}^{3t}$$

is the general integral of the given equation. □

10.8 Stability

The long-time behaviour of solutions of an ODE is a problem of great theoretical and applicative importance. A large class of *dynamical systems*, i.e., of systems whose state depends upon time and that are modelled by one or more differential equations, admit solutions at any time t after a given instant t_0. The behaviour of a particular solution can be very diversified: for instance, after a starting transition where it depends strongly on the initial data, it could subsequently converge asymptotically to a limit solution, independent of time; it could, instead, present a periodic, or quasi-periodic, course, or approach such a configuration; a solution could even have an absolutely unpredictable, or chaotic, behaviour in time.

Perhaps more interesting than the single solution is though the behaviour of a family of solutions that differ by slight perturbations either of the initial data, or of the ODE. In fact, the far future of one solution might not be representative of other solutions that kick off nearby. Often the mathematical model described by an ODE is just an approximation of a physically-more-complex system; to establish the model's reliability it is thus fundamental to determine how 'robust' the information that can be extracted from it is, with respect to the possible errors of the model. In the majority of cases moreover, the ODE is solved numerically, and the discretised problem will introduce extra perturbations. These and other reasons lead us to ask ourselves whether solutions that are initially very close stay close at all times, even converge to a limit solution, rather than moving eventually apart from one another. These kinds of issues are generically referred to as concerning (asymptotic) *stability*. The results on continuous dependency upon initial data, discussed in Sect. 10.4.1 (especially Proposition 10.15), do not answer the question satisfactorily. They merely provide information about bounded time intervals: the constant $\mathrm{e}^{L|t-t_0|}$ showing up in (10.36) grows exponentially from the instant t_0. A more specific and detailed analysis is needed to properly understand the matter.

We shall discuss stability exclusively in relationship with stationary solutions; the generalisation to periodic orbits and chaotic behaviours is incredibly fascinating but cannot be dealt with at present.

Let us begin with the Cauchy problem

$$\begin{cases} \boldsymbol{y}' = \boldsymbol{f}(t, \boldsymbol{y}), & t > t_0, \\ \boldsymbol{y}(t_0) = \boldsymbol{y}_0, \end{cases} \tag{10.91}$$

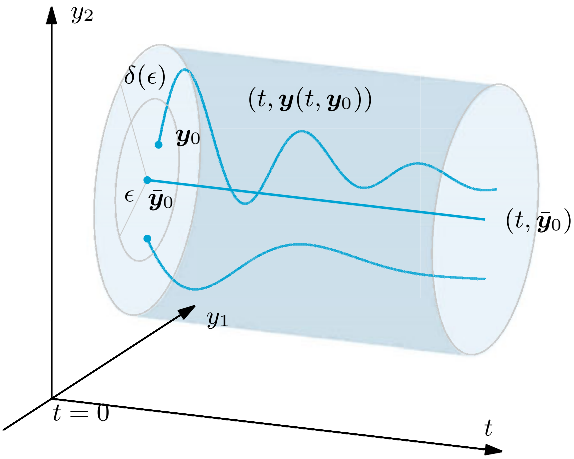


Figure 10.9. Lyapunov stability

and suppose there is a $\bar{\boldsymbol{y}}_0$, belonging to D, such that $\boldsymbol{f}(t, \bar{\boldsymbol{y}}_0) = \mathbf{0}$ for any $t \geq t_0$. Then $\boldsymbol{y}(t) = \bar{\boldsymbol{y}}_0$, $\forall t \geq t_0$, is a constant solution that we shall call *stationary solution*. The point $\bar{\boldsymbol{y}}_0$ is said **critical point**, **stationary point**, or **equilibrium point** for the equation. A further hypothesis will be that the solutions to (10.91) are defined at all times $t > t_0$, whichever the initial datum $\boldsymbol{y}_0$ in a suitable neighbourhood $\bar{B}$ of $\bar{\boldsymbol{y}}_0$; from now on $\boldsymbol{y}(t, \boldsymbol{y}_0)$ will denote such a solution. Therefore, it makes sense to compare these solutions to $\bar{\boldsymbol{y}}_0$ over the interval $[t_0, +\infty)$. To this end, the notions of (Lyapunov) stability and attractive solution are paramount.

Definition 10.43 *The stationary solution $\bar{\boldsymbol{y}}_0 = \boldsymbol{y}(t, \bar{\boldsymbol{y}}_0)$ to problem (10.91) is* **stable** *if, for any neighbourhood $B_\varepsilon(\bar{\boldsymbol{y}}_0)$ of $\bar{\boldsymbol{y}}_0$, there exists a neighbourhood $B_\delta(\bar{\boldsymbol{y}}_0)$ such that*

$$\boldsymbol{y}_0 \in \bar{B} \cap B_\varepsilon(\bar{\boldsymbol{y}}_0) \quad \Longrightarrow \quad \boldsymbol{y}(t, \boldsymbol{y}_0) \in B_\delta(\bar{\boldsymbol{y}}_0) \quad \forall t \geq t_0 \,.$$

This is exemplified in Fig. 10.9.

Definition 10.44 *The stationary solution $\bar{\boldsymbol{y}}_0$ is called* **attractive** (*or an* **attractor**) *if there exists a neighbourhood $B(\bar{\boldsymbol{y}}_0) \subseteq \bar{B}$ such that*

$$\boldsymbol{y}_0 \in B(\bar{\boldsymbol{y}}_0) \quad \Longrightarrow \quad \lim_{t \to +\infty} \boldsymbol{y}(t, \boldsymbol{y}_0) = \bar{\boldsymbol{y}}_0 \,,$$

and **uniformly attractive** (*a* **uniform attractor**) *if the above limit is uniform with respect to* $\boldsymbol{y}_0$, *i.e.,*

$$\lim_{t\to+\infty} \sup_{\boldsymbol{y}_0 \in B(\bar{\boldsymbol{y}}_0)} \|\boldsymbol{y}(t,\boldsymbol{y}_0) - \bar{\boldsymbol{y}}_0)\| = 0\,.$$

The two properties of stability and attractiveness are unrelated. The point $\bar{\boldsymbol{y}}_0$ is **uniformly asymptotically stable** if it is both stable and uniformly attractive.

Example 10.45

The simplest (yet rather meaningful, as we will see) case is the autonomous linear problem

$$\begin{cases} y' = \lambda y\,, \qquad t > 0\,, \qquad \lambda \in \mathbb{R}\,, \\ y(t_0) = y_0\,, \end{cases}$$

whose only stationary solution is $\bar{y}_0 = 0$. The solutions $y(t,y_0) = \mathrm{e}^{\lambda t} y_0$ confirm that 0 is stable if and only if $\lambda \le 0$ (in this case we may choose $\delta = \varepsilon$ in the definition). Moreover, for $\lambda = 0$ the point 0 is not an attractor (all solutions are clearly constant), whereas for $\lambda < 0$, the point 0 is uniformly attractive (hence uniformly asymptotically stable): for example setting $B(0) = B_1(0) = (-1,1)$, we have

$$\sup_{y_0 \in B(0)} |y(t,y_0)| = \mathrm{e}^{\lambda t} \to 0 \qquad \text{as } t \to +\infty\,.$$

As far as stability is concerned, we obtain similar results when $\lambda \in \mathbb{C}$ (complex-valued solutions) provided we replace λ with $\mathcal{R}e\,\lambda$. □

This example generalises directly to autonomous linear systems, now examined.

10.8.1 Autonomous linear systems

Suppose $\boldsymbol{f}(t,\boldsymbol{y}) = \boldsymbol{f}(\boldsymbol{y}) = \boldsymbol{A}\boldsymbol{y} + \boldsymbol{b}$ has $\boldsymbol{A} \in \mathbb{R}^{n,n}$ and $\boldsymbol{b} \in \mathbb{R}^n$ independent of time. A stationary solution of (10.91) corresponds to a solution of the linear system $\boldsymbol{A}\boldsymbol{y} = -\boldsymbol{b}$ (unique if $\boldsymbol{A}$ is non-singular). If $\bar{\boldsymbol{y}}_0$ is one such solution the variable change $\boldsymbol{z} = \boldsymbol{y} - \bar{\boldsymbol{y}}_0$ allows us to study the stability of the zero solution of the homogeneous equation $\boldsymbol{z}' = \boldsymbol{A}\boldsymbol{z}$. There is so no loss of generality in considering, henceforth, the stability of the solution $\boldsymbol{y}(t,\boldsymbol{0}) = \boldsymbol{0}$ of the homogeneous problem

$$\begin{cases} \boldsymbol{y}' = \boldsymbol{A}\boldsymbol{y}\,, \qquad t > 0\,, \\ \boldsymbol{y}(t_0) = \boldsymbol{y}_0\,. \end{cases} \tag{10.92}$$

From Sect. 10.6, in particular Proposition 10.35, we know every solution $\boldsymbol{y}(t,\boldsymbol{y}_0)$ is a linear combination of

$$\boldsymbol{w}(t) = \mathrm{e}^{\lambda t}\boldsymbol{p}(t)\,,$$

where $\lambda \in \mathbb{C}$ is an eigenvalue of $\boldsymbol{A}$, and $\boldsymbol{p}(t)$ is a vector-valued polynomial in t depending on the eigenvector (or one of the eigenvectors) associated to λ. In case the algebraic and geometric multiplicities of λ coincide, every $\boldsymbol{p}(t)$ associated to it has degree 0, and so is constant; otherwise, there exist polynomials of positive degree, which therefore tend to ∞ as $t \to +\infty$.

Consequences:

- if every eigenvalue of $\boldsymbol{A}$ has negative real part, then all basis elements $\boldsymbol{w}(t)$ tend to 0 as $t \to +\infty$ (recall $e^{\sigma t}t^{\beta} \to 0$ for $t \to +\infty$ if $\sigma = \mathcal{R}e\,\lambda < 0$, for any β);
- if all eigenvalues of $\boldsymbol{A}$ have negative or zero real part, and the latter ones have coinciding algebraic and geometric multiplicities, all $\boldsymbol{w}(t)$ are bounded on $[0, +\infty)$;
- if there are eigenvalues of $\boldsymbol{A}$ with positive real part, or zero real part and algebraic multiplicity greater than the geometric multiplicity, then some $\boldsymbol{w}(t)$ tends to ∞ as $t \to +\infty$.

How this translates in the language of stability is easily said.

Proposition 10.46 *a) The origin $\boldsymbol{y}(t, \boldsymbol{0}) = \boldsymbol{0}$ is a stable solution of (10.92) if and only if all eigenvalues λ of $\boldsymbol{A}$ satisfy $\mathcal{R}e\,\lambda \leq 0$, and those with $\mathcal{R}e\,\lambda = 0$ have the same algebraic and geometric multiplicity.*
b) The origin is a uniformly attractive solution (hence, uniformly asymptotically stable) if and only if all eigenvalues of $\boldsymbol{A}$ satisfy $\mathcal{R}e\,\lambda < 0$.

We shall investigate now all the scenarios for *systems of two equations*.

10.8.2 Two-dimensional systems

The generic 2×2 matrix

$$\boldsymbol{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

has determinant $\det\boldsymbol{A} = ad - bc$ and trace $\operatorname{tr}\boldsymbol{A} = a + d$. Its eigenvalues are roots of the characteristic polynomial $\chi(\lambda) = \lambda^2 - \operatorname{tr}\boldsymbol{A}\,\lambda + \det\boldsymbol{A}$, so

$$\lambda = \frac{\operatorname{tr}\boldsymbol{A} \pm \sqrt{(\operatorname{tr}\boldsymbol{A})^2 - 4\det\boldsymbol{A}}}{2}\,.$$

If $(\operatorname{tr}\boldsymbol{A})^2 \neq 4\det\boldsymbol{A}$, then the eigenvalues are distinct, necessarily simple, and the matrix is diagonalisable. If, instead, $(\operatorname{tr}\boldsymbol{A})^2 = 4\det\boldsymbol{A}$, the double eigenvalue $\lambda = \operatorname{tr}\boldsymbol{A}/2$ has geometric multiplicity 2 if and only if $b = c = 0$, i.e., $\boldsymbol{A}$ is diagonal; if the multiplicity is one, $\boldsymbol{A}$ is not diagonalisable.

Let us assume first $\boldsymbol{A}$ is diagonalisable. Then Proposition 10.32 tells us each solution of (10.92) can be written as

$$\boldsymbol{y}(t, \boldsymbol{y}_0) = z_1(t)\boldsymbol{v}_1 + z_2(t)\boldsymbol{v}_2\,, \qquad (10.93)$$

where $\boldsymbol{v}_1$, $\boldsymbol{v}_2$ are linearly independent vectors (in fact, they are precisely the eigenvectors if the eigenvalues are real, or two vectors manufactured from a complex eigenvector's real and imaginary parts if the eigenvalues are complex-conjugate); $z_1(t)$ and $z_2(t)$ denote real maps satisfying, in particular, $z_1(0)\boldsymbol{v}_1 + z_2(0)\boldsymbol{v}_2 = \boldsymbol{y}_0$.

To understand properly the possible asymptotic situations, it proves useful to draw a phase portrait, i.e., a representation of the orbits

$$\Gamma(\boldsymbol{y}) = \{\boldsymbol{y}(t, \boldsymbol{y}_0) = (y_1(t), y_2(t)) \,:\, t \geq 0\}$$

on the phase plane $\mathbb{R}^2$ of coordinates y_1, y_2. We will see straight away that it is possible to eliminate t and obtain an explicit functional relationship between y_1 and y_2. Actually, it will be better to perform such operation on the variables z_1 and z_2 first, represent the orbit $\Gamma(\boldsymbol{z}) = \{(z_1(t), z_2(t)) \,:\, t \geq 0\}$ in the phase plane z_1z_2, and then pass to the plane y_1y_2 using the linear transformation (10.93).

We have to distinguish six cases.

i) Two real non-zero eigenvalues with equal sign: $\lambda_2 \leq \lambda_1 < 0$ *or* $0 < \lambda_1 \leq \lambda_2$.

Then

$$z_1(t) = d_1 \mathrm{e}^{\lambda_1 t}\,, \qquad z_2(t) = d_2 \mathrm{e}^{\lambda_2 t}\,,$$

with d_1, d_2 dependent on $\boldsymbol{y}_0$. If $d_1 = 0$ or $d_2 = 0$, the orbits lie on the coordinate axes $z_1 = 0$ or $z_2 = 0$. If neither is zero,

$$z_2(t) = d_2 \left(\mathrm{e}^{\lambda_1 t}\right)^{\frac{\lambda_2}{\lambda_1}} = d_2 \left(\frac{z_1(t)}{d_1}\right)^{\frac{\lambda_2}{\lambda_1}}\,,$$

so the orbits are graphs of

$$z_2 = d z_1^{\alpha}\,, \qquad \text{with } \alpha \geq 1$$

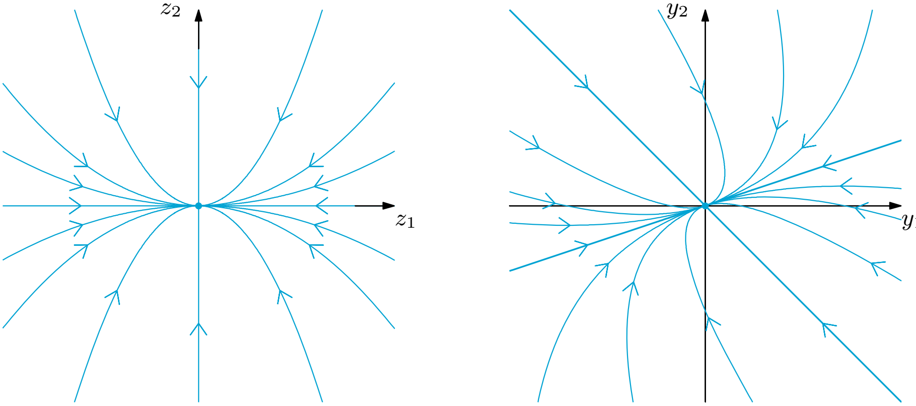


Figure 10.10. Phase portrait for $\boldsymbol{y}' = \boldsymbol{A}\boldsymbol{y}$ in the z_1z_2-plane (left) and in the y_1y_2-plane (right): the case $\lambda_2 < \lambda_1 < 0$ (node)

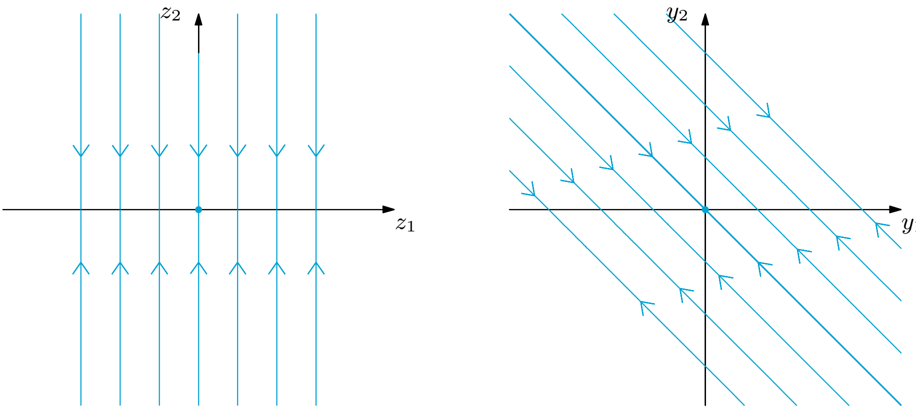


Figure 10.11. Phase portrait for $\boldsymbol{y}' = \boldsymbol{A}\boldsymbol{y}$ in the z_1z_2-plane (left) and in the y_1y_2-plane (right): the case $\lambda_2 < \lambda_1 = 0$

(which are half-lines if $\lambda_1 = \lambda_2$.) Points on the orbits move in time towards the origin, which is thus *uniformly asymptotically stable*, if the eigenvalues are negative (Fig. 10.10, left); the orbits leave the origin if the eigenvalues are positive. On the plane y_1y_2 the corresponding orbits are shown by Fig. 10.10, right, where we took $\boldsymbol{v}_1 = (3,1)$, $\boldsymbol{v}_2 = (-1,1)$.

The origin is called a **node**, and is stable or unstable according to the eigenvalues' signs.

ii) Two real eigenvalues, at least one of which is zero:
$\lambda_2 \leq \lambda_1 = 0 \quad or \quad 0 = \lambda_1 \leq \lambda_2 .$

The function $z_1(t) = d_1$ is constant. Therefore if $\lambda_2 \neq 0$, the orbits are vertical half-lines, oriented towards the axis $z_2 = 0$ if $\lambda_2 < 0$, the other way if $\lambda_2 > 0$ (Fig. 10.11). If $\lambda_2 = 0$, the matrix $\boldsymbol{A}$ is null (being diagonalisable), hence all orbits are constant. Either way, the origin is a *stable equilibrium* point, but *not attractive.*

iii) Two real eigenvalues with opposite signs: $\lambda_2 < 0 < \lambda_1 .$

The z_1z_2-orbits are the four semi-axes, together with the curves

$$z_2 = dz_1^{\alpha}\,, \qquad \text{with } \alpha < 0$$

(hyperbolas, if $\alpha = -1$), oriented as in Fig. 10.12.

The origin, called a **saddle point**, is *neither stable nor attractive.*

iv) Two purely-imaginary eigenvalues: $\lambda = \pm i\omega\,, \omega \neq 0.$

As

$$z_1(t) = d_1 \cos(\omega t + d_2) \qquad \text{and} \qquad z_2(t) = d_1 \sin(\omega t + d_2)\,,$$

the orbits are concentric circles on the phase plane z_1z_2, and concentric ellipses on y_1y_2 (Fig. 10.13).

The origin is called a **centre**, and is *stable but not attractive.*

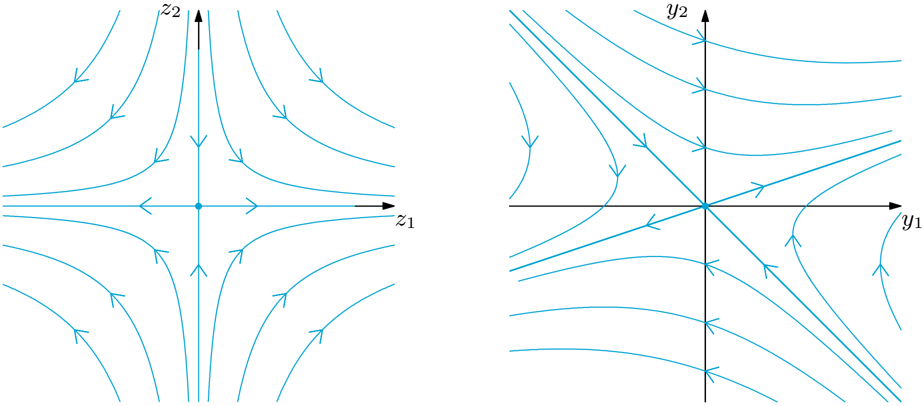


Figure 10.12. Phase portrait for $\boldsymbol{y}' = \boldsymbol{A}\boldsymbol{y}$ in the z_1z_2-plane (left) and in the y_1y_2-plane (right): the case $\lambda_2 < 0 < \lambda_1$ (saddle point)

v) Complex-conjugate eigenvalues, with non-zero real part: $\lambda = \sigma \pm \mathrm{i}\omega$, *with* $\omega \neq 0$ *and* $\sigma < 0$ *or* $\sigma > 0$.

From

$$z_1(t) = d_1 \mathrm{e}^{\sigma t} \cos(\omega t + d_2) \qquad \text{and} \qquad z_2(t) = d_1 \mathrm{e}^{\sigma t} \sin(\omega t + d_2)\,,$$

we see the orbits spiralling towards the origin if $\sigma < 0$, and moving outward if $\sigma > 0$ (Fig. 10.14). In the former case the origin is *uniformly asymptotically stable.*

The origin is a **focus**, stable or unstable according to the sign of σ.

The last case occurs for non-diagonalisable $\boldsymbol{A}$, in other words if there is a double eigenvalue with geometric multiplicity equal 1. Then $\boldsymbol{v}_1$ is the unique eigenvector in (10.93), and $\boldsymbol{v}_2$ is the associated generalised eigenvector, see (10.71).

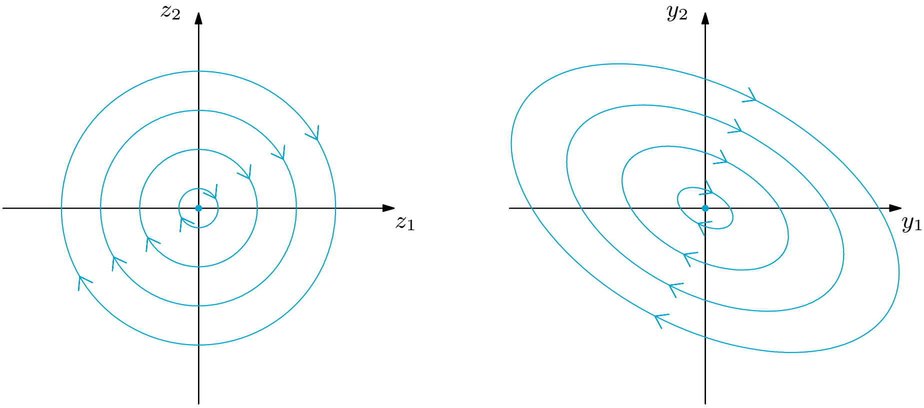


Figure 10.13. Phase portrait for $\boldsymbol{y}' = \boldsymbol{A}\boldsymbol{y}$ in the z_1z_2-plane (left) and in the y_1y_2-plane (right): the case $\lambda = \pm\omega$ (centre)

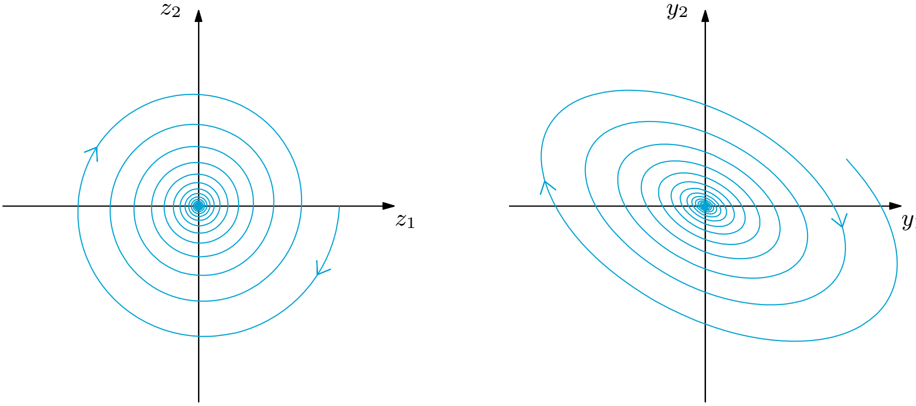


Figure 10.14. Phase portrait for $\boldsymbol{y}' = \boldsymbol{A}\boldsymbol{y}$ in the z_1z_2-plane (left) and in the y_1y_2-plane (right): the case $\lambda = \sigma \pm \omega$ (focus)

vi) One real double eigenvalue λ*, of geometric multiplicity* 1*.*

We have

$$z_1(t) = d_1 \mathrm{e}^{\lambda_1 t}\,, \qquad z_2(t) = \mathrm{e}^{\lambda_2 t}(d_2 + d_1 t)\,.$$

When $\lambda = 0$, and $d_1 \neq 0$, the orbits are vertical straight lines as shown in Fig. 10.15; for $d_1 = 0$ the orbits are fixed points on $z_1 = 0$. Therefore, the origin is neither stable, nor attractive.

When $\lambda \neq 0$ instead, t can be written as $t = \frac{1}{|\lambda|}\log\left|\frac{z_1}{d_1}\right|$, whence

$$z_2 = \left(\frac{d_1}{d_2} + \frac{1}{|\lambda|}\log\left|\frac{z_1}{d_1}\right|\right) z_1\,.$$

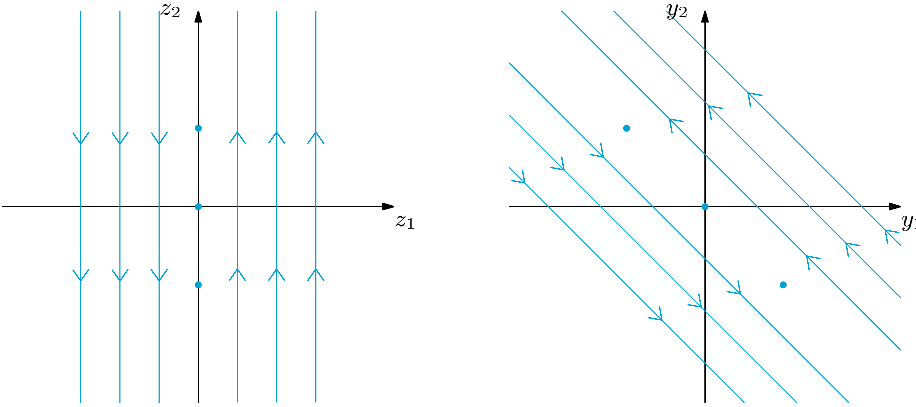


Figure 10.15. Phase portrait for $\boldsymbol{y}' = \boldsymbol{A}\boldsymbol{y}$ in the z_1z_2-plane (left) and in the y_1y_2-plane (right): the case $\lambda_1 = \lambda_2 = 0$, $bc \neq 0$

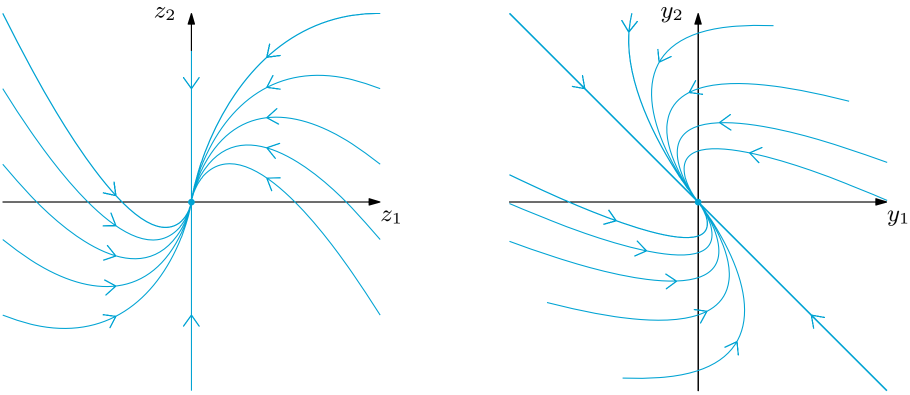


Figure 10.16. Phase portrait for $\boldsymbol{y}' = \boldsymbol{A}\boldsymbol{y}$ in the $z_1 z_2$-plane (left) and in the $y_1 y_2$-plane (right): the case $\lambda_1 = \lambda_2 \neq 0$, $bc \neq 0$ (improper node)

Figure 10.16 shows the orbits oriented to, or from, the origin according to whether $\lambda < 0$, $\lambda > 0$. The origin is *uniformly asymptotically stable*, and still called (degenerate, or improper) **node**; again, it is stable or unstable depending on the eigenvalue sign.

Application: the simple pendulum (IV). The example continues from p. 454. The dynamical system $\boldsymbol{y}' = \boldsymbol{f}(\boldsymbol{y})$ has an infinity of equilibrium points $\bar{\boldsymbol{y}}_0$, because

$$\boxed{\boldsymbol{f}(\bar{\boldsymbol{y}}_0) = \boldsymbol{0} \quad \text{if and only if} \quad \bar{y}_2 = 0 \text{ and } \bar{y}_1 = \ell\pi \quad \text{with } \ell \in \mathbb{Z}\,.}$$

The solutions $\theta(t) = \ell\pi$, with ℓ an even number, correspond to a vertical rod, with P in the lowest point S (Fig. 10.3); when ℓ is odd the bob is in the highest position I. It is physically self-evident that moving P from S a little will make the bob swing back to S; on the contrary, the smallest nudge to the bob placed in I will make it move away and never return, at any future moment. This is precisely the meaning of a *stable* point S and an *unstable* point I.

The discussion of Sect. 10.8.2 renders this intuitive idea precise. Consider a simplified model for the pendulum, obtained by *linearising* equation (10.12) around an equilibrium position.

On a neighbourhood of the point $\bar{\theta} = 0$ we have $\sin\theta \sim \theta$, so equation (10.12) can be replaced by

$$\frac{\mathrm{d}^2\theta}{\mathrm{d}t^2} + \alpha\frac{\mathrm{d}\theta}{\mathrm{d}t} + k\theta = 0\,, \tag{10.94}$$

which describes small oscillations around the equilibrium S; the value $\alpha = 0$ gives the *equation of harmonic motion*. The corresponding solution to the Cauchy problem (10.13) is easy to find by referring to Example 10.4. Equivalently, the initial value problem assumes the form (10.92) with

$$\boldsymbol{A} = \begin{pmatrix} 0 & 1 \\ -k & -\alpha \end{pmatrix} = \boldsymbol{J}\boldsymbol{f}(0,0)\,,$$

whose eigenvalues are $\lambda = \dfrac{-\alpha \pm \sqrt{\alpha^2 - 4k}}{2}$. Owing to Sect. 10.8.2, we have that

- if $\alpha^2 \geq 4k$, the origin $\bar{\boldsymbol{y}}_0 = (0, 0)$ is a uniformly asymptotically stable node [cases *i)* or *vi)*] ;
- if $0 < \alpha^2 < 4k$, the origin is a uniformly asymptotically stable focus [case *v)*] ;
- if $\alpha = 0$, the origin is a centre [case *iv)*] .

Whatever the case, the bottom position S is always stable.

Linearising (10.12) around the equilibrium $\bar{\theta} = \pi$, and changing variables $\varphi = \theta - \pi$ (so that $\sin\theta = -\sin\varphi$) produces a problem of the form (10.92), with

$$\boldsymbol{A} = \begin{pmatrix} 0 & 1 \\ k & -\alpha \end{pmatrix} = \boldsymbol{Jf}(\pi, 0)\,.$$

The eigenvalues $\lambda = \dfrac{-\alpha \pm \sqrt{\alpha^2 + 4k}}{2}$ are always non-zero and of distinct sign. The point $\bar{\boldsymbol{y}}_0 = (\pi, 0)$ is thus a saddle [case *iii)*], and so unstable. This substantiates the claim that I is an unstable equilibrium.

The discussion will end on p. 488. □

10.8.3 Non-linear stability: an overview

Certain stability features of linear systems are inherited by non-linear systems, thought of as deformations of linear ones. Suppose the function $\boldsymbol{f}(t, \boldsymbol{y})$ appearing in (10.91) has the form

$$\boldsymbol{f}(t, \boldsymbol{y}) = \boldsymbol{A}\boldsymbol{y} + \boldsymbol{g}(t, \boldsymbol{y})\,, \tag{10.95}$$

with $\boldsymbol{g}$ continuous and such that

$$\boldsymbol{g}(t, \boldsymbol{y}) = o(\|\boldsymbol{y}\|) \quad \text{as } \boldsymbol{y} \to \boldsymbol{0} \ \text{ uniformly in } t\,; \tag{10.96}$$

this means there exists a continuous map $\phi : \mathbb{R}_+ \to \mathbb{R}_+$ such that $\phi(s) \to 0$, $s \to 0^+$, and

$$\|\boldsymbol{g}(t, \boldsymbol{y})\| \leq \phi(\|\boldsymbol{y}\|)\|\boldsymbol{y}\|\,, \qquad \forall \boldsymbol{y} \in B(\boldsymbol{0})\,, \quad \forall t > t_0\,,$$

on a neighbourhood $B(\boldsymbol{0})$ of the origin. Then, the origin is an equilibrium for equation (10.91). What can we say about its asymptotic stability? One answer is given by the following fact.

Theorem 10.47 *Let $\boldsymbol{f}$ be defined by (10.95), with $\boldsymbol{g}$ as in (10.96).*
a) If all eigenvalues of $\boldsymbol{A}$ have strictly negative real part, the origin is a uniformly asymptotically stable equilibrium for (10.91)*.*
b) If there is an eigenvalue of $\boldsymbol{A}$ with strictly positive real part, the origin is unstable.

Under the given hypotheses the properties of the non-linear system $\boldsymbol{y}' = \boldsymbol{A}\boldsymbol{y} + \boldsymbol{g}(t, \boldsymbol{y})$ are the same of the corresponding linear one $\boldsymbol{y}' = \boldsymbol{A}\boldsymbol{y}$; the latter is nothing but the *linearisation*, around the origin, of the former (compare Remark 10.25). We cannot say much more if all eigenvalues have non-positive real parts and some are purely imaginary: stability in this case depends upon other properties of $\boldsymbol{g}$.

Yet, the theorem has an important consequence for autonomous systems around an equilibrium $\bar{\boldsymbol{y}}_0$. The criterion is known as **Principle of linearised stability**.

Corollary 10.48 *Let $\boldsymbol{f} : D \subseteq \mathbb{R}^n \to \mathbb{R}^n$ be a $\mathcal{C}^1$ map, $\bar{\boldsymbol{y}}_0 \in D$ such that $\boldsymbol{f}(\bar{\boldsymbol{y}}_0) = \boldsymbol{0}$, and $\boldsymbol{A} = \boldsymbol{J}\boldsymbol{f}(\bar{\boldsymbol{y}}_0)$ the Jacobian of $\boldsymbol{f}$ at $\bar{\boldsymbol{y}}_0$. Then the autonomous system $\boldsymbol{y}' = \boldsymbol{f}(\boldsymbol{y})$ has the following stability properties.*
a) If all eigenvalues of $\boldsymbol{A}$ have negative real part, $\bar{\boldsymbol{y}}_0$ is a uniformly asymptotically stable equilibrium for (10.91)*.*
b) If there is an eigenvalue of $\boldsymbol{A}$ with positive real part, $\bar{\boldsymbol{y}}_0$ is unstable.

Proof. Using the Taylor expansion (5.16) at $\bar{\boldsymbol{y}}_0$,

$$\boldsymbol{f}(\boldsymbol{y}) = \boldsymbol{f}(\bar{\boldsymbol{y}}_0) + \boldsymbol{J}\boldsymbol{f}(\bar{\boldsymbol{y}}_0)(\boldsymbol{y} - \bar{\boldsymbol{y}}_0) + \boldsymbol{g}(\boldsymbol{y}) = \boldsymbol{A}(\boldsymbol{y} - \bar{\boldsymbol{y}}_0) + \boldsymbol{g}(\boldsymbol{y}) \,,$$

with $\boldsymbol{g}(\boldsymbol{y}) = o(\|\boldsymbol{y} - \bar{\boldsymbol{y}}_0\|)$ as $\boldsymbol{y} \to \bar{\boldsymbol{y}}_0$. The variable change $\boldsymbol{z} = \boldsymbol{y} - \bar{\boldsymbol{y}}_0$ puts us in the hypotheses of the previous theorem, thus concluding the proof. □

These two results are local in nature. Global information concerning the stability of a stationary point can be obtained, for autonomous systems, from the knowledge of a first integral (Sect. 10.4.5), or a Lyapunov function (Example 10.22 ii)).

A conservative system corresponding to an equation like (10.42) or (10.46) admits a first integral, namely the total energy E; therefore, by a theorem due to Lagrange, we can say that *if the origin is a strict minimum for the potential energy Π, it must be a stable equilibrium for the system.*

The same conclusion follows if the origin is stationary for a system admitting a Lyapunov function V. Furthermore, if the derivative of V along any trajectory is strictly negative (with the exception of the origin), then the origin is an *attractor*.

Application: the simple pendulum (V). The results about the linearised equation (10.94) determine the stability of the stationary points of (10.12) in presence of damping. As a consequence of linearised stability (Corollary 10.48), if $\alpha > 0$ the matrix $\boldsymbol{J}\boldsymbol{f}(\bar{\boldsymbol{y}}_0) = \boldsymbol{A}$ has two eigenvalues with negative real part for $\bar{\boldsymbol{y}}_0 = (0, 0)$, and one eigenvalue with positive real part for $\bar{\boldsymbol{y}}_0 = (\pi, 0)$. As in the linearised problem, the bottom equilibrium point S is uniformly asymptotically stable, whereas the top point I is unstable. Figure 10.17 zooms on the phase portrait. Multiplying equation (10.12) by $\frac{\mathrm{d}\theta}{\mathrm{d}t}$ gives

$$\frac{\mathrm{d}}{\mathrm{d}t} E(\theta, \theta') = -\alpha \left(\frac{\mathrm{d}\theta}{\mathrm{d}t} \right)^2 \leq 0 \,,$$

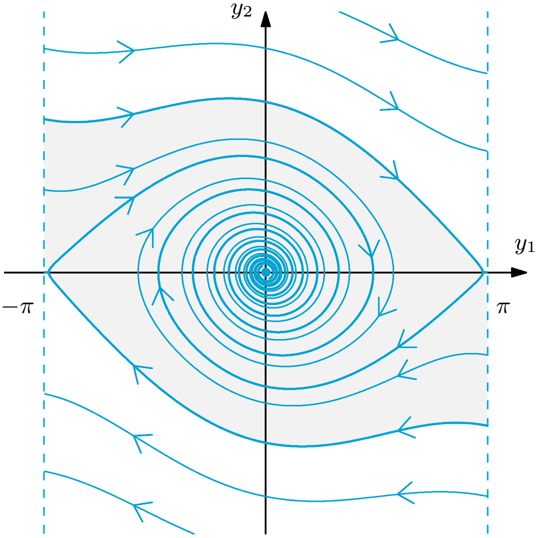


Figure 10.17. Trajectories in the phase plane for the dampled pendulum (restricted to $|y_1| \leq \pi$). The shaded region is the basin of attraction to the origin. The points lying immediately above (resp. below) it, between two orbits, are attracted to the stationary point $(2\pi, 0)$ $((-2\pi, 0))$, and so on.

making the energy a Lyapunov function, which thus decreases along the orbits (compare with Fig. 10.8).

For a free undamped motion, equilibria behave in the same way in the linear and non-linear problems; this, though, can be proved by other methods. In particular, the origin's stability follows from the the fact that it minimises the potential energy, by Lagrange's Theorem. □

10.9 Exercises

1. *Determine the general integral of the following separable ODEs:*

   $\boxed{\text{a)}}$ $y' = \dfrac{(t+2)y}{t(t+1)}$ b) $y' = \dfrac{y^2}{t\log t} - \dfrac{1}{t\log t}$

2. *Tell what the general integral of the following homogeneous equations is:*

   $\boxed{\text{a)}}$ $4t^2y' = y^2 + 6ty - 3t^2$ b) $t^2y' - y^2 e^{t/y} = ty$

3. *Integrate the following linear differential equations:*

   $\boxed{\text{a)}}$ $y' = \dfrac{1}{t}y - \dfrac{3t+2}{t^3}$ b) $ty' = y + \dfrac{2t^2}{1+t^2}$

$\boxed{4.}$ *Find the general integral of the Bernoulli equations:*

a) $y' = \dfrac{1}{t}y - y^2$ b) $y' = \dfrac{1}{t}y + \dfrac{t}{y}\log t$

$\boxed{5.}$ *Determine the particular integral of the ODE*

$$y' = \frac{1 - \mathrm{e}^{-y}}{2t+1}$$

subject to the condition $y(0) = 1$.

$\boxed{6.}$ *Establish if there are solutions to*

$$y' = -2y + \mathrm{e}^{-2t}$$

with null derivative at the origin.

7. *Solve, on* $[\sqrt[4]{\mathrm{e}}, +\infty)$, *the Cauchy problem*

$$\begin{cases} \mathrm{e}^y y' = 4t^3 \log t(1 + \mathrm{e}^y) \\ y(\sqrt[4]{\mathrm{e}}) = 0. \end{cases}$$

8. *Solve on the interval* $(-2,\, 2)$ *the initial value problem*

$$\begin{cases} y' = \dfrac{3t}{t^2 - 4}|y| \\ y(0) = -1. \end{cases}$$

9. *Given the ODE*

$$y' \sin 2t - 2(y + \cos t) = 0, \qquad t \in \left(0,\, \frac{\pi}{2}\right),$$

determine the general integral and write a solution that stays bounded as $t \to \frac{\pi}{2}^-$.

$\boxed{10.}$ *Solve the Cauchy problem*

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}(y^2) = y^2 + \dfrac{t}{y} \\ y(0) = 1\,. \end{cases}$$

11. *Solve the Cauchy problem*

$$\begin{cases} \dfrac{y''}{\big(1+(y')\big)^{3/2}} = \dfrac{8t^3}{(t^4+1)^2} \\ y(1) = 0\,, \quad y'(1) = 0\,. \end{cases}$$

$\boxed{12.}$ *Find the solutions of the initial value problem*

$$\begin{cases} y'' = \dfrac{4}{y^3} \\ y(0) = 2\,, \quad y'(0) = -\sqrt{3}\,. \end{cases}$$

13. *Determine the particular integral of the differential equation*

$$yy'' - (y')^2 = y^2 \log y$$

such that $y(0) = y'(0) = 1$.

$\boxed{14.}$ *As* α *varies in* $\mathbb{R}$, *solve the differential equation*

$$y' = (2+\alpha)y - 2\mathrm{e}^{\alpha t}$$

with initial datum $y(0) = 3$.

$\boxed{15.}$ *Let* a, b *be real numbers. Solve*

$$\begin{cases} y' = a\dfrac{y}{t} + 3t^b \\ y(2) = 1 \end{cases}$$

on the interval $[2, +\infty)$.

$\boxed{16.}$ *Given the ODE*

$$y'(t) = -3ty(t) + kt$$

depending on the real number k, *find the solution vanishing at the origin.*

17. *Given*

$$y' = \frac{y^2 - 2y - 3}{2(1+4t)}\,,$$

a) *determine its general integral;*
b) *find the particular integral* $y_0(t)$ *satisfying* $y_0(0) = 1$;

$\boxed{18.}$ *Given the ODE* $y' = f(t,y) = \sqrt{t+y}$, *determine open sets* $\Omega = I \times D = (\alpha,\beta) \times (\gamma,\delta)$ *inside the domain of* f, *for which Theorem 10.14 is valid.*

$\boxed{19.}$ *Find the maximal interval of existence for the autonomous equation*

$$\boldsymbol{y}' = \boldsymbol{f}(\boldsymbol{y})$$

with:

a) $\boldsymbol{f}(\boldsymbol{y}) = \sqrt{1 + y_1^2 + y_2^2}\,\boldsymbol{i} + \arctan(y_1 + y_2)\,\boldsymbol{j} \quad \text{on } \mathbb{R}^2$

b) $\boldsymbol{f}(\boldsymbol{y}) = \dfrac{\sin \|\boldsymbol{y}\|^2}{\log^3(2 + \|\boldsymbol{y}\|^2)}\,\boldsymbol{y} \quad \text{on } \mathbb{R}^n$

$\boxed{20.}$ *Verify that the autonomous equation* $\boldsymbol{y}' = \boldsymbol{f}(\boldsymbol{y})$, *with*

a) $\boldsymbol{f}(\boldsymbol{y}) = \dfrac{y_2}{1 + 3y_1^2 + 5y_2^4}\,\boldsymbol{i} - \dfrac{y_1}{1 + 3y_1^2 + 5y_2^4}\,\boldsymbol{j}$

b) $\boldsymbol{f}(\boldsymbol{y}) = (4y_1^2 y_2^3 + 2y_1 y_2)\boldsymbol{i} - (2y_1 y_2^4 + y_2^2)\boldsymbol{j}$

admits on $\mathbb{R}^2$ *a first integral, then compute it.*

21. *Determine the general integral of the system* $\boldsymbol{y}' = \boldsymbol{A}\boldsymbol{y}$, *where:*

a) $\boldsymbol{A} = \begin{pmatrix} 9 & -4 \\ 8 & -3 \end{pmatrix}$

b) $\boldsymbol{A} = \begin{pmatrix} 3 & -4 \\ 4 & 3 \end{pmatrix}$

$\boxed{\text{c)}}$ $\boldsymbol{A} = \begin{pmatrix} 13 & 0 & -4 \\ 15 & 2 & -5 \\ 30 & 0 & -9 \end{pmatrix}$

$\boxed{\text{d)}}$ $\boldsymbol{A} = \begin{pmatrix} 4 & 3 & -2 \\ 3 & 2 & -6 \\ 1 & 3 & 1 \end{pmatrix}$

e) $\boldsymbol{A} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & -10 & -6 \end{pmatrix}$

$\boxed{\text{f)}}$ $\boldsymbol{A} = \begin{pmatrix} 4 & 0 & 0 \\ 0 & -2 & 9 \\ 0 & 4 & -2 \end{pmatrix}$

$\boxed{\text{g)}}$ $\boldsymbol{A} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 1 & 0 & 3 \end{pmatrix}$

h) $\boldsymbol{A} = \begin{pmatrix} 1 & 5 & 0 \\ 0 & 1 & 0 \\ 4 & 0 & 1 \end{pmatrix}$

22. *Determine a particular integral of the systems:*

a) $\boldsymbol{y}' = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & 2 \end{pmatrix} \boldsymbol{y} + \begin{pmatrix} 0 \\ t \\ t^2 \end{pmatrix}$

b) $\boldsymbol{y}' = \begin{pmatrix} -1 & -1 & 0 \\ 1/8 & 0 & -1 \\ 0 & 1/8 & -1 \end{pmatrix} \boldsymbol{y} + \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \mathrm{e}^{-2t}$

c) $\boldsymbol{y}' = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & -1 \end{pmatrix} \boldsymbol{y} + \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \sin 2t$

23. *Solve*

$$\begin{cases} y_1' = 2y_1 + y_3 \\ y_2' = y_3 \\ y_3' = 8y_1 \end{cases}$$

with constraints $y_1(0) = y_2(0) = 1$ *and* $y_3(0) = 0$.

$\boxed{24.}$ *Find the solutions of*

$$\begin{cases} y_1' + y_2' = 5y_2 \\ 3y_1' - 2y_2' = 5y_1 \end{cases}$$

with initial data $y_1(0) = 2$, $y_2(0) = -1$.

$\boxed{25.}$ *Determine, in function of the real number* b, *the solutions of*

$$\boldsymbol{y}' = \begin{pmatrix} -1 & b \\ b & -1 \end{pmatrix} \boldsymbol{y}$$

with $\boldsymbol{y}(0) = (1,1)^T$.

26. *As the real parameter* a *varies, find the general integral of*

$$\boldsymbol{y}' = \begin{pmatrix} 1+a & 0 & 1 \\ a-2 & 3a-1 & 1-a \\ 0 & 0 & a \end{pmatrix} \boldsymbol{y}\,.$$

$\boxed{27.}$ *Solve the system*

$$\begin{cases} x' = 2x - y \\ y'' = -10x + 5y - 2y'\,. \end{cases}$$

28. *Write the general integral for the following linear equations of order two:*

a) $y'' + 3y' + 2y = t^2 + 1$

$\boxed{\text{b)}}$ $y'' - 4y' + 4y = \mathrm{e}^{2t}$

$\boxed{\text{c)}}$ $y'' + y = 3\cos t$

d) $y'' - 3y' + 2y = \mathrm{e}^{t}$

$\boxed{\text{e)}}$ $y'' - 9y = \mathrm{e}^{-3t}$

f) $y'' - 2y' - 3y = \sin t$

29. *Solve the Cauchy problems:*

a) $\begin{cases} y'' + 2y' + 5y = 0 \\ y(0) = 0 \\ y'(0) = 2 \end{cases}$

$\boxed{\text{b)}}$ $\begin{cases} y'' - 5y' + 4y = 2t + 1 \\ y(0) = \frac{7}{8} \\ y'(0) = 0 \end{cases}$

30. *Integrate the following linear ODEs of order n:*

$\boxed{\text{a)}}$ $y''' + y'' - 2y = 0$ b) $y''' - 2y'' + y' = 0$

$\boxed{\text{c)}}$ $y^{(4)} - 5y''' + 7y'' - 5y' + 6y = \sin t$

$\boxed{31.}$ *Determine the general integral of $y'' + y' - 6y = e^{kt}$ in function of the real k.*

32. *Determine the general integral of the differential equation $y'' - 2y' + (1+k)y = 0$, as k varies in $\mathbb{R}$.*

$\boxed{33.}$ *Determine the general integral of*

$$y''' - 2y'' + 49y' - 98y = 48\sin t + (\beta^2 + 49)e^{\beta t}$$

for every β in $\mathbb{R}$.

34. *Determine the general integral of the ODE $y''' + 9ay' = \cos 3t$ in function of $a \in \mathbb{R}$.*

$\boxed{35.}$ *Discuss the stability of the origin in $\mathbb{R}^3$ for the equation*

$$\boldsymbol{y}' = \boldsymbol{A}\boldsymbol{y} \qquad \textit{with} \qquad \boldsymbol{A} = \begin{pmatrix} -3 & 0 & -5 \\ 0 & -1 & 0 \\ 5 & 0 & -2 \end{pmatrix}.$$

$\boxed{36.}$ *Study the stability of $\boldsymbol{y}_0 = (-3, 1)$ for the following equation in $\mathbb{R}^2$:*

$$\boldsymbol{y}' = \boldsymbol{f}(\boldsymbol{y}) \qquad \textit{where} \quad \boldsymbol{f}(\boldsymbol{y}) = (3y_1y_2 - 2y_2^2 + 11)\boldsymbol{i} + (y_1 + 3y_2)\boldsymbol{j}\,.$$

10.9.1 Solutions

1. *ODEs with separable variables:*

a) The map $h(y) = y$ has a zero at $y = 0$, which is thus a singular integral. Suppose $y \neq 0$ and separate variables, so that

$$\int \frac{1}{y}\,dy = \int \frac{t+2}{t(t+1)}\,dt = \log\frac{ct^2}{|t+1|}, \qquad c > 0\,.$$

Passing to exponentials,

$$y = y(t) = c\,\frac{t^2}{t+1}, \qquad c \neq 0\,.$$

The singular integral $y = 0$ is obtained by putting $c = 0$ in the general formula.

c) $y = y(t) = \dfrac{1 + c \log^2 t}{1 - c \log^2 t}\,, \quad c \in \mathbb{R}\,.$

2. *Homogeneous ODEs:*

a) Supposing $t \neq 0$ and dividing by $4t^2$ gives

$$y' = \frac{1}{4}\frac{y^2}{t^2} + \frac{3}{2}\frac{y}{t} - \frac{3}{4}\,.$$

Substitute $z = \frac{y}{t}$, so that $y' = z + tz'$ and then

$$z + tz' = \frac{1}{4}z^2 + \frac{3}{2}z - \frac{3}{4}\,,$$

$$4tz' = (z-1)(z+3)\,.$$

Since $\varphi(z) = (z-1)(z+3)$ vanishes at $z = 1$ and $z = -3$, we get $y = t$ and $y = -3t$ as singular integrals. For the general integral we separate the variables,

$$\int \frac{4}{(z-1)(z+3)}\,\mathrm{d}z = \int \frac{1}{t}\,\mathrm{d}t\,.$$

Then exponentiating

$$\log\left|\frac{z-1}{z+3}\right| = \log c|t|\,, \qquad c > 0$$

and solving for z, we have

$$z = \frac{1+3ct}{1-ct}\,, \qquad c \in \mathbb{R}\,,$$

in which the singular integral $z = 1$ is included. Altogether, the general integral of the equation is

$$y = \frac{t+3ct^2}{1-ct}\,, \qquad c \in \mathbb{R}\,.$$

b) $y = y(t) = -\dfrac{t}{\log\log c|t|}\,, \qquad c > 0\,.$

3. *Linear equations:*

a) Formula (10.24), with $a(t) = \frac{1}{t}$ and $b(t) = -\frac{3t+2}{t^3}$, gives

$$y = \mathrm{e}^{\int \frac{1}{t}\,\mathrm{d}t} \int \mathrm{e}^{-\int \frac{1}{t}\,\mathrm{d}t}\left(-\frac{3t+2}{t^3}\right)\mathrm{d}t = \frac{3}{2t} + \frac{2}{3t^2} + ct\,, \qquad c \in \mathbb{R}\,.$$

b) $y = 2t \arctan t + ct\,, \quad c \in \mathbb{R}\,.$

4. *Bernoulli equations:*

a) In the notation of Sect. 10.3.4,

$$p(t) = -1\,, \qquad q(t) = \frac{1}{t}\,, \qquad \alpha = 2\,.$$

As $\alpha = 2 > 0$, $y(t) = 0$ is a solution. Now divide by y^2,

$$\frac{1}{y^2}y' = \frac{1}{ty} - 1\,,$$

and set $z = z(t) = y^{1-2} = \frac{1}{y}$; then $z' = -\frac{1}{y^2}y'$ and the equation reads $z' = 1 - \frac{1}{t}z$. Solving for z,

$$z = z(t) = \frac{t^2 + c}{2t}\,, \qquad c \in \mathbb{R}\,.$$

Therefore,

$$y = y(t) = \frac{1}{z(t)} = \frac{2t}{t^2 + c}\,, \qquad c \in \mathbb{R}$$

to which we have to add $y(t) = 0$.

b) We have

$$p(t) = t\log t\,, \qquad q(t) = \frac{1}{t}\,, \qquad \alpha = -1\,.$$

Set $z = y^2$, so $z' = 2yy'$ and the equation reads

$$z' = \frac{2}{t}z + 2t\log t\,.$$

Integrating the linear equation in z thus obtained, we have

$$z = z(t) = t^2(\log^2 t + c)\,, \qquad c \in \mathbb{R}\,.$$

Therefore,

$$y = y(t) = \pm t\sqrt{\log^2 t + c}\,, \qquad c \in \mathbb{R}\,.$$

5. The ODE is separable. The constant solution $y = 0$ is not valid because it fails the initial condition $y(0) = 1$. By separating variables we get

$$\int \frac{1}{1 - \mathrm{e}^{-y}}\,\mathrm{d}y = \int \frac{1}{2t + 1}\,\mathrm{d}x\,;$$

Then

$$\log|1 - \mathrm{e}^{y}| = \frac{1}{2}\log|2t + 1| + c\,, \qquad c \in \mathbb{R}\,.$$

Solving for y, and noticing that $y = 0$ corresponds to $c = 0$, we obtain the general integral:

$$y = \log\left(1 - c\sqrt{|2t + 1|}\right)\,, \qquad c \in \mathbb{R}.$$

The datum $y(0) = 1$ forces $c = 1 - \mathrm{e}$, so

$$y = \log\left(1 + (\mathrm{e} - 1)\sqrt{|2t + 1|}\right).$$

6. The equation's general integral reads

$$y = \mathrm{e}^{-\int 2\,\mathrm{d}t} \int \mathrm{e}^{\int 2\,\mathrm{d}t} \mathrm{e}^{-2t}\,\mathrm{d}t = \mathrm{e}^{-2t}(t + c)\,, \qquad c \in \mathbb{R}\,.$$

The condition is $y'(0) = 0$. Setting $x = 0$ in $y'(t) = -2y(t) + \mathrm{e}^{-2t}$ gives the new condition $y(0) = \frac{1}{2}$, from which $c = \frac{1}{2}$. Thus,

$$y = \mathrm{e}^{-2t}\left(t + \frac{1}{2}\right).$$

7. $y = \log\left(2\mathrm{e}^{t^4(\log t - \frac{1}{4})} - 1\right)$. 8. $y = -\dfrac{8}{(4 - t^2)^{3/2}}$. 9. $y = \dfrac{\sin t - 1}{\cos t}$.

10. The equation is

$$2yy' = y^2 + \frac{t}{y} \qquad \text{i.e.,} \qquad y' = \frac{1}{2}y + \frac{t}{2y^2}\,.$$

It can be made into a Bernoulli equation by taking

$$p(t) = \frac{t}{2}\,, \qquad q(t) = \frac{1}{2}\,, \qquad \alpha = -2\,.$$

Set $z = z(t) = y^3$, so that $z' = 3y^2y'$ and $z' = \frac{3}{2}z + \frac{3}{2}t$. Solving for z,

$$z = z(t) = c\mathrm{e}^{(3/2)t} - t - \frac{2}{3}\,.$$

Therefore

$$y = y(t) = \sqrt[3]{c\mathrm{e}^{(3/2)t} - t - \frac{2}{3}}\,.$$

The initial condition $y(0) = 1$ finally gives $c = 5/3$, so

$$y = y(t) = \sqrt[3]{\frac{5}{3}\mathrm{e}^{(3/2)t} - t - \frac{2}{3}}\,.$$

11. $y = y(t) = \dfrac{1}{6}t^3 + \dfrac{1}{2t} - \dfrac{2}{3}$.

12. This second-order equation can be reduced to first order by $y' = z(y)$, and observing $z' = \dfrac{1}{z}\dfrac{4}{y^3}$. Then $z = z(y)$ satisfies $z^2 = -\frac{4}{y^2} + c_1$. Using the initial conditions plus $y' = z(y)$ must give

$$(-\sqrt{3})^2 = -1 + c_1 \quad \text{i.e.,} \quad c_1 = 4\,.$$

Therefore

$$z^2 = \frac{4}{y^2}(y^2-1) \qquad \text{hence} \qquad z = -\frac{2}{y}\sqrt{y^2-1}$$

(the minus sign is due to $y'(0) = -\sqrt{3}$). Now the equation is separable:

$$y' = -\frac{2}{y}\sqrt{y^2-1}\,.$$

Solving

$$\sqrt{y^2-1} = -2t + c_2$$

with $y(0) = 2$ produces $c_2 = \sqrt{3}$, and then

$$y = y(t) = \sqrt{1 + (\sqrt{3} - 2t)^2}\,.$$

13. $y = y(t) = \mathrm{e}^{\sinh t}$.

14. The ODE is linear, and the general integral is straightforward

$$y = \mathrm{e}^{\int (2+\alpha)\,\mathrm{d}t} \int \mathrm{e}^{-\int (2+\alpha)\,\mathrm{d}t}(-2\mathrm{e}^{\alpha t})\,\mathrm{d}t = \mathrm{e}^{\alpha t}(1 + c\,\mathrm{e}^{2t})\,, \quad c \in \mathbb{R}.$$

From $y(0) = 3$ we find $3 = 1 + c$, so $c = 2$. The solution is thus

$$y = \mathrm{e}^{\alpha t}(1 + 2\mathrm{e}^{2t}).$$

15. Directly from the formula for linear ODEs,

$$\begin{aligned}
y &= e^{a\int \frac{1}{t}\,\mathrm{d}t}\left(3\int e^{-a\int \frac{1}{t}\,\mathrm{d}t} t^b\,\mathrm{d}t\right) = t^a\left(3\int t^{b-a}\,\mathrm{d}t\right) \\
&= \begin{cases} t^a\left(\dfrac{3}{b-a+1}\,t^{b-a+1} + c\right) & \text{if } b-a \neq -1, \\ t^a\,(3\log t + c) & \text{if } b-a = -1\,, \end{cases} \\
&= \begin{cases} \dfrac{3}{b-a+1} t^{b+1} + c\,t^a & \text{if } b-a \neq -1, \\ 3t^a \log t + ct^a & \text{if } b-a = -1. \end{cases}
\end{aligned}$$

Now, $y(2) = 1$ imposes

$$\begin{cases} \dfrac{3}{b-a+1}\,2^{b+1} + c\,2^a = 1 & \text{if } b-a \neq -1, \\ 3\cdot 2^a \log 2 + c\,2^a = 1 & \text{if } b-a = -1\,, \end{cases}$$

so

$$\begin{cases} c = 2^{-a}\left(1 - \dfrac{3}{b-a+1}\,2^{b+1}\right) & \text{if } b-a \neq -1, \\ c = 2^{-a} - 3\log 2 & \text{if } b-a = -1. \end{cases}$$

The solution is

$$y = \begin{cases} \dfrac{3}{b-a+1}\, t^{b+1} + 2^{-a}\left(1 - \dfrac{3}{b-a+1}\, 2^{b+1}\right) t^a & \text{if } b-a \neq -1, \\ 3t^a \log t + \left(2^{-a} - 3\log 2\right) t^a & \text{if } b-a = -1. \end{cases}$$

16. The integral of this linear equation is

$$y = \mathrm{e}^{-3\int t\,\mathrm{d}t} \int \mathrm{e}^{3\int t\,\mathrm{d}t} kt\ \mathrm{d}t = \frac{k}{3} + c\,\mathrm{e}^{-\frac{3}{2}t^2}, \qquad c \in \mathbb{R}.$$

Condition $y(0) = 0$ implies $c = -\frac{k}{3}$. Therefore

$$y = \frac{k}{3}\left(1 - \mathrm{e}^{-\frac{3}{2}t^2}\right).$$

17. *Solving the ODE* $y' = \frac{y^2-2y-3}{2(1+4t)}$:

a) $y(t) = \dfrac{3 + c\sqrt{|1+4t|}}{1 - c\sqrt{|1+4t|}}$, $c \in \mathbb{R}$, plus the constant solution $y(t) = -1$.

b) $y_0(t) = \dfrac{3 - \sqrt{|1+4t|}}{1 + \sqrt{|1+4t|}}$;

18. We have $\operatorname{dom} f = \{(t,y) \in \mathbb{R}^2 : t + y \geq 0\}$. As

$$\nabla f = \left(\frac{1}{2\sqrt{t+y}}, \frac{1}{2\sqrt{t+y}}\right),$$

f is Lipschitz in y on every Ω where $\frac{1}{2\sqrt{t+y}}$ is bounded. Since

$$\sup_{t\in(\alpha,\beta),\, y\in(\gamma,\delta)} \frac{1}{2\sqrt{t+y}} = \frac{1}{2\sqrt{\alpha+\gamma}},$$

the condition holds for any real α, γ such that $\alpha + \gamma > 0$. The end-points β, δ are completely free, and may also be $+\infty$.

19. *Global solutions:*

a) We have

$$\frac{\partial f_1}{\partial y_j} = \frac{y_j}{1 + y_1^2 + y_2^2}, \qquad \frac{\partial f_2}{\partial y_j} = \frac{1}{1 + (y_1 + y_2)^2}, \quad 1 \leq j \leq 2,$$

so $\left|\dfrac{\partial f_i}{\partial y_j}\right| \leq 1$ on $\mathbb{R}^2$. Hence $\boldsymbol{f}$ is Lipschitz on $D = \mathbb{R}^2$, and every solution exists on the whole $I = \mathbb{R}$.

b) The map $\boldsymbol{f}$ is certainly differentiable on $\mathbb{R}^n$, with continuous partial derivatives (as composites of $\mathcal{C}^1$ elementary functions; besides, the denominator is never zero). Partial derivatives are thus bounded by Weierstrass' Theorem, making $\boldsymbol{f}$ locally Lipschitz on $\mathbb{R}^n$. Moreover,

$$\|\boldsymbol{f}(\boldsymbol{y})\| \le \frac{1}{(\log 2)^3}\,\|\boldsymbol{y}\|\,, \qquad \forall \boldsymbol{y} \in \mathbb{R}^n\,.$$

Therefore Theorem 10.19 holds, and any solution exists on $I = \mathbb{R}$.

20. *Existence of first integrals:*

a) As $\boldsymbol{y}\cdot\boldsymbol{f}(\boldsymbol{y}) = 0$ the equation is conservative, so $\Phi(\boldsymbol{y}) = \frac{1}{2}\|\boldsymbol{y}\|^2$ is a first integral (Example 10.22 i)).

b) Since $\operatorname{div}\boldsymbol{f}(\boldsymbol{y}) = 8y_1y_2^3 + 2y_2 - (8y_1y_2^3 + 2y_2) = 0$ on $\mathbb{R}^2$, $\boldsymbol{f}$ is a curl, and the equation admits a first integral Φ such that $\mathbf{curl}\,\Phi = \boldsymbol{f}$. Then

$$\frac{\partial\Phi}{\partial y_2} = 4y_1^2y_2^3 + 2y_1y_2\,, \qquad -\frac{\partial\Phi}{\partial y_1} = -(2y_1y_2^4 + y_2^2)\,;$$

integrating the first gives

$$\Phi(\boldsymbol{y}) = y_1^2y_2^4 + y_1y_2^2 + c(y_1)\,,$$

and using the second we find $c(y_1) =$ constant. Hence a family of first integrals is

$$\Phi(\boldsymbol{y}) = y_1^2y_2^4 + y_1y_2^2 + c\,.$$

21. *General integrals:*

a) $\boldsymbol{y}(t) = c_1\mathrm{e}^t\begin{pmatrix}1\\2\end{pmatrix} + c_2\mathrm{e}^{5t}\begin{pmatrix}1\\1\end{pmatrix}$.

b) $\boldsymbol{y}(t) = c_1\mathrm{e}^{3t}\begin{pmatrix}\cos 4t\\ \sin 4t\end{pmatrix} + c_2\mathrm{e}^{3t}\begin{pmatrix}-\sin 4t\\ \cos 4t\end{pmatrix}$.

c) The matrix $\boldsymbol{A}$ has eigenvalues $\lambda_1 = 1$, $\lambda_2 = 2$, $\lambda_3 = 3$ corresponding to eigenvectors

$$\boldsymbol{v}_1 = (1,0,3)^T\,, \qquad \boldsymbol{v}_2 = (0,1,0)^T\,, \qquad \boldsymbol{v}_1 = (2,5,5)^T\,.$$

The general integral is thus of type

$$\boldsymbol{y}(t) = c_1\mathrm{e}^t\begin{pmatrix}1\\0\\3\end{pmatrix} + c_2\mathrm{e}^{2t}\begin{pmatrix}0\\1\\0\end{pmatrix} + c_3\mathrm{e}^{3t}\begin{pmatrix}2\\5\\5\end{pmatrix}\,.$$

d) $\boldsymbol{A}$ has eigenvalues $\lambda_1 = 3$, $\lambda_{2,3} = 2 \pm 3i$ with eigenvectors

$$\boldsymbol{v}_1 = (2,0,1)^T\,, \qquad \boldsymbol{v}_{2,3} = (1,\pm i,1)^T\,.$$

Thus

$$\boldsymbol{y}(t) = c_1 \mathrm{e}^{3t}\begin{pmatrix}2\\0\\1\end{pmatrix} + c_2\mathrm{e}^{2t}\begin{pmatrix}\cos 3t\\-\sin 3t\\\cos 3t\end{pmatrix} + c_3\mathrm{e}^{2t}\begin{pmatrix}\sin 3t\\\cos 3t\\\sin 3t\end{pmatrix}$$

is the general integral.

e) $$\boldsymbol{y}(t) = c_1\begin{pmatrix}1\\0\\0\end{pmatrix} + c_2\mathrm{e}^{-3t}\begin{pmatrix}\cos t\\-3\cos t-\sin t\\8\cos t+6\sin t\end{pmatrix} + c_3\mathrm{e}^{-3t}\begin{pmatrix}-\sin t\\-\cos t+3\sin t\\6\cos t-8\sin t\end{pmatrix}\,.$$

f) $\boldsymbol{A}$ has eigenvalues $\lambda_1 = 4$ with multiplicity 2, and $\lambda_2 = -8$. The eigenvectors of λ_1 are $\boldsymbol{v}_1^{(1)} = (1,0,0)^T$ and $\boldsymbol{v}_1^{(2)} = (0,3,2)^T$, while the eigenvector corresponding to λ_2 is $\boldsymbol{v}_2 = (0,3,-2)^T$. Therefore

$$\boldsymbol{y}(t) = c_1\mathrm{e}^{-8t}\begin{pmatrix}0\\3\\-2\end{pmatrix} + c_2\mathrm{e}^{4t}\begin{pmatrix}1\\0\\0\end{pmatrix} + c_3\mathrm{e}^{4t}\begin{pmatrix}0\\3\\2\end{pmatrix}\,.$$

g) The matrix $\boldsymbol{A}$ has one eigenvalue $\lambda = 3$ of multiplicity 3. There are two linearly independent eigenvectors $\boldsymbol{v}_1^{(1)} = (0,1,0)^T$, $\boldsymbol{v}_1^{(2)} = (0,0,1)^T$. Moreover, the latter gives a generalised eigenvector $\boldsymbol{r}_1 = (1,0,0)^T$. Therefore

$$\boldsymbol{y}(t) = c_1\mathrm{e}^{3t}\begin{pmatrix}0\\1\\0\end{pmatrix} + c_2\mathrm{e}^{3t}\begin{pmatrix}0\\0\\1\end{pmatrix} + c_3\mathrm{e}^{3t}\left(t\begin{pmatrix}0\\0\\1\end{pmatrix} + \begin{pmatrix}1\\0\\0\end{pmatrix}\right)\,.$$

h) The matrix $\boldsymbol{A}$ has a unique eigenvalue $\lambda = 1$ with multiplicity 3, and one eigenvector $\boldsymbol{v}_1 = (0,0,1)^T$. This produces two generalised eigenvectors $\boldsymbol{r}_1 = (1/4,0,0)^T$, $\boldsymbol{r}_2 = (0,1/20,0)^T$. Hence

$$\begin{aligned}\boldsymbol{y}(t) &= c_1\mathrm{e}^{t}\begin{pmatrix}0\\0\\1\end{pmatrix} + c_2\mathrm{e}^{t}\left(t\begin{pmatrix}0\\0\\1\end{pmatrix} + \begin{pmatrix}1/4\\0\\0\end{pmatrix}\right)\\ &\quad + c_3\mathrm{e}^{t}\left(\frac{t^2}{2}\begin{pmatrix}0\\0\\1\end{pmatrix} + t\begin{pmatrix}1/4\\0\\0\end{pmatrix} + \begin{pmatrix}0\\1/20\\0\end{pmatrix}\right)\end{aligned}$$

is the general integral.

22. *Particular integrals:*

a) We have

$$\boldsymbol{y}(t) = \begin{pmatrix}0\\0\\-1/2\end{pmatrix}t^2 + \begin{pmatrix}-1\\0\\0\end{pmatrix}t + \begin{pmatrix}0\\1\\0\end{pmatrix}\,.$$

b) We have

$$\boldsymbol{y}(t) = \begin{pmatrix} -25/18 \\ -7/18 \\ 7/144 \end{pmatrix} \mathrm{e}^{-2t} .$$

c) A particular integral is

$$\boldsymbol{y}(t) = \begin{pmatrix} -2/3 \\ 0 \\ 1/6 \end{pmatrix} \cos 2t + \begin{pmatrix} 0 \\ -1/3 \\ -1/12 \end{pmatrix} \sin 2t .$$

23. We have

$$y_1(t) = \frac{1}{3}\mathrm{e}^{-2t} + \frac{2}{3}\mathrm{e}^{4t} , \quad y_2(t) = \frac{2}{3}\mathrm{e}^{-2t} + \frac{1}{3}\mathrm{e}^{4t} , \quad y_3(t) = -\frac{4}{3}\mathrm{e}^{-2t} + \frac{4}{3}\mathrm{e}^{4t} .$$

24. The system reads, in normal form,

$$\boldsymbol{y}' = \boldsymbol{A}\boldsymbol{y} \qquad \text{with} \qquad \boldsymbol{A} = \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix} .$$

The eigenvalues of $\boldsymbol{A}$ are $\lambda_{1,2} = 2 \pm i$ corresponding to $\boldsymbol{v}_{1,2} = (2, 1 \pm i)^T$. The general integral is

$$\boldsymbol{y}(t) = \mathrm{e}^{2t} \left(c_1 \begin{pmatrix} 2\cos t \\ \cos t - \sin t \end{pmatrix} + c_2 \begin{pmatrix} 2\sin t \\ \cos t + \sin t \end{pmatrix} \right) .$$

Imposing the constraints gives $c_1 = 1$, $c_2 = -2$.

25. The eigenvalues of $\boldsymbol{A} = \begin{pmatrix} -1 & b \\ b & -1 \end{pmatrix}$ are $\lambda_{1,2} = -1 \pm b$.

If $b \neq 0$, they are distinct, with corresponding eigenvectors $\boldsymbol{v}_1 = (1,1)^T$, $\boldsymbol{v}_2 = (1,-1)^T$, and the general integral is

$$\boldsymbol{y}(t) = c_1 \mathrm{e}^{(b-1)t} \begin{pmatrix} 1 \\ 1 \end{pmatrix} + c_2 \mathrm{e}^{-(1+b)t} \begin{pmatrix} 1 \\ -1 \end{pmatrix} .$$

If $b = 0$ the eigenvectors stay the same, but the integral reads

$$\boldsymbol{y}(t) = c_1 \mathrm{e}^{-t} \begin{pmatrix} 1 \\ 1 \end{pmatrix} + c_2 \mathrm{e}^{-t} \begin{pmatrix} 1 \\ -1 \end{pmatrix} .$$

In either case the initial datum gives $c_1 = 1$ and $c_2 = 0$. The required solution is, for any b,

$$\boldsymbol{y}(t) = \mathrm{e}^{(b-1)t} \begin{pmatrix} 1 \\ 1 \end{pmatrix} .$$

26. For $a \neq \frac{1}{2}$, $a \neq 1$,

$$\boldsymbol{y}(t) = c_1 \mathrm{e}^{at} \begin{pmatrix} 2a-1 \\ 3-2a \\ 1-2a \end{pmatrix} + c_2 \mathrm{e}^{(1+a)t} \begin{pmatrix} 2a-2 \\ 2-a \\ 0 \end{pmatrix} + c_3 \mathrm{e}^{(3a-1)t} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} ;$$

for $a = 1$,

$$\boldsymbol{y}(t) = c_1 \mathrm{e}^{t} \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} + c_2 \mathrm{e}^{2t} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + c_3 \mathrm{e}^{2t} \left(t \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} -1 \\ 0 \\ 0 \end{pmatrix} \right) ;$$

for $a = 1/2$,

$$\boldsymbol{y}(t) = c_1 \mathrm{e}^{(3/2)t} \begin{pmatrix} -1 \\ 3/2 \\ 0 \end{pmatrix} + c_2 \mathrm{e}^{(1/2)t} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + c_3 \mathrm{e}^{(1/2)t} \left(t \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \begin{pmatrix} -1/2 \\ 0 \\ 1/2 \end{pmatrix} \right) ;$$

27. Calling $y_1 = x$, $y_2 = y$, $y_3 = y'$ we may write $\boldsymbol{y}' = \boldsymbol{A}\boldsymbol{y}$ where

$$\boldsymbol{A} = \begin{pmatrix} 2 & -1 & 0 \\ 0 & 0 & 1 \\ -10 & 5 & -2 \end{pmatrix} .$$

The matrix $\boldsymbol{A}$ has eigenvalues $\lambda_1 = 0$, $\lambda_2 = 3$, $\lambda_3 = -3$ and, correspondingly,

$$\boldsymbol{v}_1 = (1, 2, 0)^T , \qquad \boldsymbol{v}_2 = (1, -1, -3)^T , \qquad \boldsymbol{v}_3 = (1, 5, -15)^T .$$

Consequently, the general integral is

$$\boldsymbol{y}(t) = c_1 \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} + c_2 \mathrm{e}^{3t} \begin{pmatrix} 1 \\ -1 \\ -3 \end{pmatrix} + c_3 \mathrm{e}^{-3t} \begin{pmatrix} 1 \\ 5 \\ -15 \end{pmatrix}$$

i.e.,

$$x(t) = c_1 + c_2 \mathrm{e}^{3t} + c_3 \mathrm{e}^{-3t} , \qquad y(t) = 2c_1 - c_2 \mathrm{e}^{3t} + 5c_3 \mathrm{e}^{-3t} .$$

28. *Second-order linear equations:*

a) $y(t; c_1, c_2) = c_1 \mathrm{e}^{-t} + c_2 \mathrm{e}^{-2t} + \frac{1}{2}t^2 - \frac{3}{2}t + \frac{9}{4}$, $c_1, c_2 \in \mathbb{R}$.

b) We solve first the homogeneous equation. The characteristic equation $\lambda^2 - 4\lambda + 4\lambda = 0$ has one double solution $\lambda = 2$, so the general homogeneous integral will be

$$y_0(t; c_1, c_2) = (c_1 + c_2 t)\mathrm{e}^{2t} , \quad c_1, c_2 \in \mathbb{R} .$$

Since $\mu = \lambda = 2$, we look for a particular integral $y_p(t) = \alpha t^2 \mathrm{e}^{2t}$. This gives

$$2\alpha \mathrm{e}^{2t} = \mathrm{e}^{2t}$$

by substitution, hence $\alpha = \frac{1}{2}$. Therefore $y_p(t) = \frac{1}{2}t^2 \mathrm{e}^{2t}$ and the general integral is

$$y(t; c_1, c_2) = (c_1 + c_2 t)\mathrm{e}^{2t} + \frac{1}{2}t^2 \mathrm{e}^{2t} , \quad c_1, c_2 \in \mathbb{R} .$$

c) The characteristic equation $\lambda^2+1=0$ has discriminant $\Delta=-4$, so $\sigma=0$ and $\omega=1$. The general integral of the homogeneous equation is

$$y_0(t;c_1,c_2)=c_1\cos t+c_2\sin t\,,\quad c_1,c_2\in\mathbb{R}\,.$$

As $\mu=\sigma=0$, we need to find a particular integral $y_p(t)=t(\alpha\cos t+\beta\sin t)$. By substitution,

$$-2\alpha\sin t+2\beta\cos t=3\cos t\,,$$

so $\alpha=0$ and $\beta=\frac{3}{2}$. Therefore $y_p(t)=\frac{3}{2}t\cos t$ and the general integral is

$$y(t;c_1,c_2)=c_1\cos t+c_2\sin t+\frac{3}{2}t\cos t\,,\quad c_1,c_2\in\mathbb{R}\,.$$

d) $y(t;c_1,c_2)=c_1\mathrm{e}^t+c_2\mathrm{e}^{2t}-t\mathrm{e}^t\,,\ c_1,c_2\in\mathbb{R}\,.$

e) The characteristic equation $\lambda^2-9=0$ is solved by $\lambda=\pm3$. Hence the general integral of the homogeneous ODE is

$$y_0(t;c_1,c_2)=c_1\mathrm{e}^{-3t}+c_2\mathrm{e}^{3t}\,,\quad c_1,c_2\in\mathbb{R}\,.$$

The particular integral must have form $y_p(t)=\alpha t\mathrm{e}^{-3t}$, so substituting,

$$-6\alpha\mathrm{e}^{-3t}=\mathrm{e}^{-3t}\,,$$

hence $\alpha=-\frac{1}{6}$. Therefore $y_p(t)=-\frac{1}{6}t\mathrm{e}^{-3t}$ and the general integral of the equation is

$$y(t;c_1,c_2)=c_1\mathrm{e}^{-3t}+c_2\mathrm{e}^{3t}-\frac{1}{6}t\mathrm{e}^{-3t}\,,\quad c_1,c_2\in\mathbb{R}\,.$$

f) $y(t;c_1,c_2)=c_1\mathrm{e}^{-t}+c_2\mathrm{e}^{3t}+\frac{1}{10}\cos t-\frac{1}{5}\sin t\,,\ c_1,c_2\in\mathbb{R}\,.$

29. *Cauchy problems:*

a) $y(t)=\mathrm{e}^{-t}\sin 2t\,.$

b) Let us treat the homogeneous equation first. The characteristic equation $\lambda^2-5\lambda+4=0$ gives $\lambda=1$, $\lambda=4$. Thus the homogeneous general integral is

$$y_0(t;c_1,c_2)=c_1\mathrm{e}^{t}+c_2\mathrm{e}^{4t}\,,\quad c_1,c_2\in\mathbb{R}\,.$$

Substituting the particular integral $y_p(t)=\alpha t+\beta$ in the equation we find

$$-5\alpha+4\alpha t+4\beta=2t+1\,,$$

so $\alpha=\frac{1}{2}$ and $\beta=\frac{7}{8}$. Therefore $y_p(t)=\frac{1}{2}t+\frac{7}{8}$ and the general solution is

$$y(t;c_1,c_2)=c_1\mathrm{e}^{t}+c_2\mathrm{e}^{4t}+\frac{1}{2}t+\frac{7}{8}\,,\quad c_1,c_2\in\mathbb{R}\,.$$

Imposing the initial conditions gives the system

$$\begin{cases} c_1 + c_2 = 0 \\ c_1 + 4c_2 + \dfrac{1}{2} = 0 \end{cases},$$

so $c_1 = \frac{1}{6}$ and $c_2 = -\frac{1}{6}$. Therefore

$$y = \frac{1}{6}e^t - \frac{1}{6}e^{4t} + \frac{1}{2}t + \frac{7}{8}.$$

30. *Linear ODEs of order n:*

a) The characteristic polynomial $\chi(\lambda) = \lambda^3 + \lambda^2 - 2$ has a real root $\lambda_1 = 1$ and two complex-conjugate zeroes $\lambda_{2,3} = -1 \pm i$. Then the general integral, in real form, is

$$y(t) = c_1 e^t + c_2 e^{-t}\cos t + c_3 e^{-t}\sin t, \qquad c_1, c_2, c_3 \in \mathbb{R}.$$

b) $y(t) = c_1 e^t + c_2 t e^t + c_3, \qquad c_1, c_2, c_3 \in \mathbb{R}.$

c) The characteristic polynomial $\chi(\lambda) = \lambda^4 - 5\lambda^3 + 7\lambda^2 - 5\lambda + 6$ has roots $\lambda_1 = 2$, $\lambda_2 = 3$, $\lambda_{3,4} = \pm i$. The homogeneous general integral is thus

$$y_0(t) = c_1 e^{2t} + c_2 t e^{3t} + c_3 \cos t + c_4 \sin t, \qquad c_1, c_2, c_3, c_4 \in \mathbb{R}.$$

Because of resonance, we search for a particular integral of type

$$y_p(t) = t(\alpha \sin t + \beta \cos t).$$

Differentiating, substituting in the ODE and comparing terms produces the system

$$\begin{cases} \alpha + \beta = 0, \\ 10\alpha - 10\beta = 1, \end{cases} \qquad \text{whence} \qquad \begin{cases} \alpha = 1/20, \\ \beta = -1/20. \end{cases}$$

In conclusion,

$$y = y(t) = y_0(t) + \frac{1}{20}t(\sin t - \cos t)$$

is the general integral.

31. The associated homogeneous equation has general integral

$$y_0(t) = c_1 e^{2t} + c_2 e^{-3t},$$

as one easily sees. The particular integral of the equation depends on k. In fact,

$$y_p(t) = \begin{cases} \alpha e^{kt} & \text{if } k \neq 2,\ k \neq -3, \\ \alpha t e^{kt} & \text{if } k = 2, \text{ or } k = -3 \end{cases}$$

with the constant α to be determined. Differentiating this expression and substituting,

$$\alpha = \begin{cases} \dfrac{1}{(k-2)(k+3)} & \text{if } k \neq 2,\ k \neq -3\,, \\ \dfrac{1}{2k+1} & \text{if } k = 2, \text{ or } k = -3\,. \end{cases}$$

Therefore,

$$y(t) = \begin{cases} c_1 \mathrm{e}^{2t} + c_2 \mathrm{e}^{-3t} + \dfrac{1}{(k-2)(k+3)} \mathrm{e}^{kt} & \text{if } k \neq 2,\ k \neq -3\,, \\ c_1 \mathrm{e}^{2t} + c_2 \mathrm{e}^{-3t} + \dfrac{1}{2k+1} t\mathrm{e}^{kt} & \text{if } k = 2, \text{ or } k = -3\,. \end{cases}$$

32. We have

$$y(t) = \begin{cases} \mathrm{e}^t \big(c_1 \cos \sqrt{k}\, t + c_2 \sin \sqrt{k}\, t\big) & \text{if } k > 0\,, \\ \mathrm{e}^t (c_1 + t c_2) & \text{if } k = 0\,, \\ c_1 \mathrm{e}^{(1+\sqrt{-k})t} + c_2 \mathrm{e}^{(1-\sqrt{-k})t} & \text{if } k < 0\,, \end{cases}$$

with $c_1, c_2 \in \mathbb{R}$.

33. The characteristic polynomial

$$\chi(\lambda) = \lambda^3 - 2\lambda^2 + 49\lambda - 98 = (\lambda - 2)(\lambda^2 + 49)$$

has roots $\lambda_1 = 2$, $\lambda_{2,3} = \pm 7i$. The homogeneous integral reads

$$y_0(t) = c_1 \mathrm{e}^{2t} + c_2 \cos 7t + c_3 \sin 7t\,, \qquad c_1, c_2, c_3 \in \mathbb{R}\,.$$

Putting $b_1(t) = 48 \sin t$ and $b_2(t) = (\beta^2 + 49)\mathrm{e}^{\beta t}$, by the principle of superposition we begin with a particular integral of the equation with source b_1 of the form $y_{p_1}(t) = a \cos t + b \sin t$. This will give $a = -1/5$ and $b = -2/5$.

The other term b_2 depends on the parameter β. When $\beta \neq 2$, we want a particular integral $y_{p_2}(t) = a\mathrm{e}^{\beta t}$. Proceeding as usual, we obtain $a = 1/(\beta - 2)$. When $\beta = 2$, the particular integral will be $y_{p_2}(t) = at\mathrm{e}^{2t}$ and in this case we find $a = 1$. So altogether, the general integral is

$$y(t) = \begin{cases} y_0(t) + \frac{1}{\beta-2}\mathrm{e}^{\beta t} - \frac{1}{5}\cos t - \frac{2}{5}\sin t & \text{if } \beta \neq 2\,, \\ y_0(t) + t\mathrm{e}^{2t} - \frac{1}{5}\cos t - \frac{2}{5}\sin t & \text{if } \beta = 2\,. \end{cases}$$

34. We have

$$y(t) = \begin{cases} c_1 + c_2 \mathrm{e}^{3\sqrt{-a}\, t} + c_3 \mathrm{e}^{-3\sqrt{-a}\, t} + \dfrac{1}{27(a-1)} \sin 3t & \text{if } a < 0\,, \\ c_1 + c_2 t + c_3 t^2 - \dfrac{1}{27} \sin 3t & \text{if } a = 0\,, \\ c_1 + c_2 \cos 3\sqrt{a}\, t + c_3 \sin 3\sqrt{a}\, t + \dfrac{1}{27(a-1)} \sin 3t & \text{if } a > 0, a \neq 1\,, \\ c_1 + c_2 \cos 3t + c_3 \sin 3t - \dfrac{t}{18} \cos 3t & \text{if } a = 1\,. \end{cases}$$

35. The eigenvalues of $\boldsymbol{A}$ are $\lambda_1 = -1$, $\lambda_{2,3} = -\frac{5\pm\sqrt{163}}{2}$, so $\mathcal{R}e\,\lambda < 0$ for both. The origin is thus uniformly asymptotically stable (see Proposition 10.46).

36. We have

$$\boldsymbol{Jf}(\boldsymbol{y}) = \begin{pmatrix} 3y_2 & 3y_1 - 4y_2 \\ 1 & 3 \end{pmatrix} .$$

The matrix

$$\boldsymbol{Jf}(-3,1) = \begin{pmatrix} -9 & -11 \\ 1 & 3 \end{pmatrix}$$

has eigenvalues $\lambda_1 = -8$, $\lambda_2 = 2$, so the origin is unstable (Corollary 10.48).

Appendices

A.1

Complements on differential calculus

In this appendix, the reader may find the proofs of various results presented in Chapters 5, 6, and 7. In particular, we prove Schwarz's Theorem and we justify the Taylor formulas with Lagrange's and Peano's remainders, as well as the rules for differentiating integrals. At last, we prove Dini's implicit function Theorem in the two dimensional case.

A.1.1 Differentiability and Schwarz's Theorem

▶ Proof of Proposition 5.8, p. 163

Proposition 5.8 *Assume f admits continuous partial derivatives in a neighbourhood of $\boldsymbol{x}_0$. Then f is differentiable at $\boldsymbol{x}_0$.*

Proof. For simplicity we consider only the case $n = 2$.
Let then $\boldsymbol{x} = (x, y)$ be a point in the neighbourhood of $\boldsymbol{x}_0 = (x_0, y_0)$ where the hypotheses hold. Call $(h, k) = (x - x_0, y - y_0)$; we must prove that for $(h, k) \to (0, 0)$,

$$f(x_0 + h, y_0 + k) = f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0)h + \frac{\partial f}{\partial y}(x_0, y_0)k + o(\sqrt{h^2 + k^2})\,.$$

Using the first formula of the finite increment for the map $x \mapsto f(x, y_0)$ gives

$$f(x_0 + h, y_0) = f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0)h + o(h)\,, \qquad h \to 0\,.$$

At the same time, Lagrange's Mean Value Theorem tells us that $y \mapsto f(x_0 + h, y)$ satisfies

$$f(x_0 + h, y_0 + k) = f(x_0 + h, y_0) + \frac{\partial f}{\partial y}(x_0 + h, \overline{y})k$$

C. Canuto, A. Tabacco: *Mathematical Analysis II*, 2nd Ed.,
UNITEXT – La Matematica per il 3+2 85, DOI 10.1007/978-3-319-12757-6_A1,

for some $\overline{y} = \overline{y}(h,k)$ between y_0 and $y_0 + k$. Since $\dfrac{\partial f}{\partial y}$ is continuous on the neighbourhood of (x_0, y_0), we have

$$\frac{\partial f}{\partial y}(x,y) = \frac{\partial f}{\partial y}(x_0,y_0) + o(1)\,, \qquad (x,y) \to (x_0,y_0)\,,$$

and because $(x_0 + h, \overline{y}) \to (x_0, y_0)$ for $(h,k) \to (0,0)$, we may write

$$\frac{\partial f}{\partial y}(x_0 + h, \overline{y}) = \frac{\partial f}{\partial y}(x_0,y_0) + o(1)\,, \qquad (h,k) \to (0,0)\,.$$

In conclusion, when $(h,k) \to (0,0)$,

$$f(x_0+h, y_0+k) = f(x_0,y_0) + \frac{\partial f}{\partial x}(x_0,y_0)h + \frac{\partial f}{\partial y}(x_0,y_0)k + o(h) + o(1)k\,.$$

But $o(h) + o(1)k = o(\sqrt{h^2+k^2})$, $(h,k) \to (0,0)$. In fact, $|h|, |k| \le \sqrt{h^2+k^2}$ implies

$$\frac{|o(h)|}{\sqrt{h^2+k^2}} \le \frac{|o(h)|}{|h|} = \left|\frac{o(h)}{h}\right| \to 0\,, \qquad (h,k) \to (0,0)$$

and

$$\frac{|o(1)k|}{\sqrt{h^2+k^2}} \le \frac{|o(1)k|}{|k|} = |o(1)| \to 0\,, \qquad (h,k) \to (0,0)\,.$$

The claim now follows. □

▶ Proof of Schwarz's Theorem, p. 168

Theorem 5.17 (**Schwarz**) *If the mixed partial derivatives* $\dfrac{\partial^2 f}{\partial x_j \partial x_i}$ *and* $\dfrac{\partial^2 f}{\partial x_i \partial x_j}$ $(j \ne i)$ *exist on a neighbourhood of* $\boldsymbol{x}_0$ *and are continuous at* $\boldsymbol{x}_0$*, they coincide at* $\boldsymbol{x}_0$*.*

Proof. For simplicity let us only consider $n = 2$. We have to prove

$$\frac{\partial^2 f}{\partial y \partial x}(x_0,y_0) = \frac{\partial^2 f}{\partial x \partial y}(x_0,y_0)\,,$$

under the assumption that the derivatives exist on a neighbourhood $B_r(x_0,y_0)$ of $\boldsymbol{x}_0 = (x_0,y_0)$ and are continuous at $\boldsymbol{x}_0$.

Let $\boldsymbol{x} = (x,y) \in B_r(x_0,y_0)$ and set $(h,k) = (x - x_0, y - y_0)$. Consider the function

$$g(h,k) = f(x_0+h, y_0+k) - f(x_0+h, y_0) - f(x_0, y_0+k) + f(x_0,y_0)\,.$$

Putting $\varphi(x) = f(x, y_0 + k) - f(x, y_0)$ we see that

$$g(h,k) = \varphi(x_0 + h) - \varphi(x_0)\,.$$

The second formula of the finite increment for the function $x \mapsto \varphi(x)$ gives a point $\xi = \xi(h,k)$, between x_0 and $x_0 + h$, for which

$$g(h,k) = h\varphi'(\xi) = h\Big(\frac{\partial f}{\partial x}(\xi, y_0 + k) - \frac{\partial f}{\partial x}(\xi, y_0)\Big)\,.$$

Let us use the same formula again, this time to $\psi(y) = \frac{\partial f}{\partial x}(\xi, y)$, to obtain a point $\eta = \eta(h,k)$ between y_0 and $y_0 + k$ such that

$$\frac{\partial f}{\partial x}(\xi, y_0 + k) - \frac{\partial f}{\partial x}(\xi, y_0) = k\frac{\partial^2 f}{\partial y \partial x}(\xi,\eta)$$

and so

$$g(h,k) = hk\frac{\partial^2 f}{\partial y \partial x}(\xi,\eta)\,.$$

Now letting $(h,k) \to (0,0)$ we have $(\xi,\eta) \to (x_0, y_0)$; therefore, since $\frac{\partial^2 f}{\partial y \partial x}$ is continuous at (x_0, y_0), we will have

$$\lim_{(h,k)\to(0,0)} \frac{g(h,k)}{hk} = \lim_{(\xi,\eta)\to(x_0,y_0)} \frac{\partial^2 f}{\partial y \partial x}(\xi,\eta) = \frac{\partial^2 f}{\partial y \partial x}(x_0, y_0)\,. \tag{A.1.1}$$

But we can write g as

$$g(h,k) = f(x_0{+}h, y_0{+}k) - f(x_0, y_0{+}k) - f(x_0{+}h, y_0) + f(x_0, y_0) = \tilde\psi(y_0{+}k) - \tilde\psi(y_0)\,,$$

having put $\tilde\psi(y) = f(x_0 + h, y) - f(x_0, y)$. Swapping the variables and proceeding as before gives

$$\lim_{(h,k)\to(0,0)} \frac{g(h,k)}{hk} = \frac{\partial^2 f}{\partial x \partial y}(x_0, y_0)\,. \tag{A.1.2}$$

Since the limit on the left in $(A.1.1)$ and $(A.1.2)$ is the same, the equality follows. □

A.1.2 Taylor's expansions

▶ Proof of Theorem 5.20, p. 172

Theorem 5.20 *A function f of class $\mathcal{C}^2$ around $\boldsymbol{x}_0$ admits at $\boldsymbol{x}_0$ the* **Taylor expansion of order one with Lagrange's remainder**:

$$f(\boldsymbol{x}) = f(\boldsymbol{x}_0) + \nabla f(\boldsymbol{x}_0)\cdot(\boldsymbol{x} - \boldsymbol{x}_0) + \frac{1}{2}(\boldsymbol{x} - \boldsymbol{x}_0)\cdot \boldsymbol{H}f(\overline{\boldsymbol{x}})(\boldsymbol{x} - \boldsymbol{x}_0)\,,$$

where $\overline{\boldsymbol{x}}$ is interior to the segment $S[\boldsymbol{x}, \boldsymbol{x}_0]$.

Proof. Set $\Delta\boldsymbol{x} = \boldsymbol{x} - \boldsymbol{x}_0 = (\Delta x_i)_{1\le i\le n}$ for simplicity. We consider the function of one real variable $\varphi(t) = f(\boldsymbol{x}_0 + t\Delta\boldsymbol{x})$, defined around $t_0 = 0$, and show it is differentiable twice around 0. In that case we will find its Taylor expansion of second order. Property 5.11 ensures that φ is differentiable on some neighbourhood of the origin, with

$$\varphi'(t) = \nabla f(\boldsymbol{x}_0 + t\Delta\boldsymbol{x}) \cdot \Delta\boldsymbol{x} = \sum_{i=1}^{n} \Delta x_i \frac{\partial f}{\partial x_i}(\boldsymbol{x}_0 + t\Delta\boldsymbol{x})\,.$$

Now set $\psi_i(t) = \frac{\partial f}{\partial x_i}(\boldsymbol{x}_0 + t\Delta\boldsymbol{x})$ and use Property 5.11 on them, to obtain

$$\psi_i'(t) = \nabla\Big(\frac{\partial f}{\partial x_i}\Big)(\boldsymbol{x}_0 + t\Delta\boldsymbol{x}) \cdot \Delta\boldsymbol{x} = \sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_j \partial x_i}(\boldsymbol{x}_0 + t\Delta\boldsymbol{x})\Delta x_j\,.$$

This implies φ can be differentiated twice, and also that

$$\begin{aligned}\varphi''(t) &= \sum_{i=1}^{n} \Delta x_i \sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_j \partial x_i}(\boldsymbol{x}_0 + t\Delta\boldsymbol{x})\Delta x_j \\ &= \Delta\boldsymbol{x} \cdot \boldsymbol{H}f(\boldsymbol{x}_0 + t\Delta\boldsymbol{x})\Delta\boldsymbol{x}\,,\end{aligned}$$

as $\boldsymbol{H}f$ is symmetric.
The Taylor expansion of φ of order two, centred at $t_0 = 0$, and computed at $t = 1$ reads

$$\varphi(1) = \varphi(0) + \varphi'(0) + \frac{1}{2}\varphi''(\bar{t}) \qquad \text{con} \quad 0 < \bar{t} < 1\,;$$

substituting the expressions of $\varphi'(0)$ and $\varphi''(\bar{t})$ found earlier, and putting $\overline{\boldsymbol{x}} = \boldsymbol{x}_0 + \bar{t}\Delta\boldsymbol{x}$, proves the claim. □

▶ Proof of Theorem 5.21, p. 172

Theorem 5.21 *A function f of class $\mathcal{C}^2$ around $\boldsymbol{x}_0$ admits at $\boldsymbol{x}_0$ the following* **Taylor expansion of order two with Peano's remainder***:*

$$\begin{aligned}f(\boldsymbol{x}) = f(\boldsymbol{x}_0) + \nabla f(\boldsymbol{x}_0)\cdot(\boldsymbol{x}-\boldsymbol{x}_0) + \frac{1}{2}(\boldsymbol{x}-\boldsymbol{x}_0)\cdot \boldsymbol{H}f(\boldsymbol{x}_0)(\boldsymbol{x}-\boldsymbol{x}_0)\\ +o(\|\boldsymbol{x}-\boldsymbol{x}_0\|^2)\,, \qquad \boldsymbol{x}\to\boldsymbol{x}_0\,.\end{aligned}$$

Proof. Consider the generic summand

$$\frac{1}{2}\frac{\partial^2 f}{\partial x_j \partial x_i}(\overline{\boldsymbol{x}})\Delta x_i \Delta x_j$$

in the quadratic part on the right-hand side of (5.15). As the second derivative of f is continuous at $\boldsymbol{x}_0$ and $\overline{\boldsymbol{x}}$ belongs to the segment $S[\boldsymbol{x}, \boldsymbol{x}_0]$, we have

$$\frac{\partial^2 f}{\partial x_j \partial x_i}(\overline{\boldsymbol{x}}) = \frac{\partial^2 f}{\partial x_j \partial x_i}(\boldsymbol{x}_0) + \eta_{ij}(\boldsymbol{x}),$$

where $\lim\limits_{\boldsymbol{x}\to\boldsymbol{x}_0} \eta_{ij}(\boldsymbol{x}) = 0$. Hence

$$\frac{\partial^2 f}{\partial x_j \partial x_i}(\overline{\boldsymbol{x}}) \Delta x_i \Delta x_j = \frac{\partial^2 f}{\partial x_j \partial x_i}(\boldsymbol{x}_0) \Delta x_i \Delta x_j + \eta_{ij}(\boldsymbol{x}) \Delta x_i \Delta x_j .$$

We will prove the last term is in fact $o(\|\Delta \boldsymbol{x}\|^2)$; for this, we recall that $0 \le (a-b)^2 = a^2 + b^2 - 2ab$ for any pair of real numbers a, b, so that $ab \le \frac{1}{2}(a^2+b^2)$. Now note that the following inequalities hold

$$|\Delta x_i|\,|\Delta x_j| \le \frac{1}{2}(|\Delta x_i|^2 + |\Delta x_j|^2) \le \frac{1}{2}\|\Delta \boldsymbol{x}\|^2 ,$$

so

$$0 \le \frac{|\eta_{ij}(\boldsymbol{x})\Delta x_i \Delta x_j|}{\|\Delta \boldsymbol{x}\|^2} \le \frac{1}{2}|\eta_{ij}(\boldsymbol{x})|$$

and then

$$\lim_{\boldsymbol{x}\to\boldsymbol{x}_0} \frac{\eta_{ij}(\boldsymbol{x})\Delta x_i \Delta x_j}{\|\Delta \boldsymbol{x}\|^2} = 0 .$$

In summary,

$$\begin{aligned}\frac{1}{2}(\boldsymbol{x}-\boldsymbol{x}_0)\cdot H_f(\overline{\boldsymbol{x}})(\boldsymbol{x}-\boldsymbol{x}_0) &= \frac{1}{2}(\boldsymbol{x}-\boldsymbol{x}_0)\cdot H_f(\boldsymbol{x}_0)(\boldsymbol{x}-\boldsymbol{x}_0) \\ &\qquad + o(\|\boldsymbol{x}-\boldsymbol{x}_0\|^2), \qquad \boldsymbol{x}\to\boldsymbol{x}_0 ,\end{aligned}$$

whence the result. □

A.1.3 Differentiating functions defined by integrals

First, we state the following result, of great importance per se, that will be used below.

Theorem A.1.1 (**Heine-Cantor**) *Let $\boldsymbol{f} : \mathrm{dom}\, \boldsymbol{f} \subseteq \mathbb{R}^n \to \mathbb{R}^m$ be a continuous map on a compact set $\Omega \subseteq \mathrm{dom}\, \boldsymbol{f}$. Then $\boldsymbol{f}$ is uniformly continuous on Ω.*

Proof. The proof is similar to what we saw in the one dimensional case, using the multidimensional version of the Bolzano-Weierstrass Theorem (see Vol. I, Appendix A.3). □

▶ Proof of Theorem 6.17, p. 215

Proposition 6.17 *The function f defined by* (6.14) *is continuous on I. Moreover, if g admits continuous partial derivative $\dfrac{\partial g}{\partial x}$ on $\mathcal{R}$, then f is of class $\mathcal{C}^1$ on I and*

$$f'(x) = \int_a^b \frac{\partial g}{\partial x}(x,y)\,\mathrm{d}y\,.$$

Proof. Suppose x_0 is an interior point of I (the proof can be easily adapted to the case where x_0 is an end-point). Then there is a $\sigma > 0$ such that $[x_0-\sigma, x_0+\sigma] \subset I$; the rectangle $E = [x_0-\sigma, x_0+\sigma] \times J$ isa compact set in $\mathbb{R}^2$.

Let us begin by proving the continuity of f at x_0. As we assumed g continuous on $\mathcal{R}$, hence on the compact subset E, Heine-Cantor's Theorem implies g is uniformly continuous on E. Hence, for any $\varepsilon > 0$ there is a $\delta > 0$ such that $|g(\boldsymbol{x}_1) - g(\boldsymbol{x}_2)| < \varepsilon$ for any pair of points $\boldsymbol{x}_1 = (x_1, y_1)$, $\boldsymbol{x}_2 = (x_2, y_2)$ in E with $\|\boldsymbol{x}_1 - \boldsymbol{x}_2\| < \delta$. We may assume $\delta < \sigma$. Let now $x \in [x_0-\sigma, x_0+\sigma]$ be such that $|x - x_0| < \delta$; for any given y in $[a,b]$, then,

$$\|(x,y) - (x_0,y)\| = |x - x_0| < \delta\,,$$

so $|g(x,y) - g(x_0,y)| < \varepsilon$. Therefore

$$|f(x)-f(x_0)| = \left|\int_a^b \big(g(x,y)-g(x_0,y)\big)\,\mathrm{d}y\right| \le \int_a^b \big|g(x,y)-g(x_0,y)\big|\,\mathrm{d}y < \varepsilon(b-a)\,,$$

proving continuity.

As for differentiability at x_0, in case $\dfrac{\partial g}{\partial x}$ exists and is continuous on $I \times J$, we observe

$$\frac{f(x)-f(x_0)}{x-x_0} = \frac{1}{x-x_0}\int_a^b \big(g(x,y)-g(x_0,y)\big)\,\mathrm{d}y = \int_a^b \frac{g(x,y)-g(x_0,y)}{x-x_0}\,\mathrm{d}y\,.$$

Given $y \in [a,b]$, we can use The Mean Value Theorem (in dimension 1) on $x \mapsto g(x,y)$. This gives a point ξ between x and x_0 for which

$$\frac{g(x,y)-g(x_0,y)}{x-x_0} = \frac{\partial g}{\partial x}(\xi, y)\,.$$

By assumption, $\dfrac{\partial g}{\partial x}$ is uniformly continuous on E (again by Heine-Cantor's Theorem). Consequently, for any $\varepsilon > 0$ there is a $\delta > 0$ ($\delta < \sigma$) such that, if $\boldsymbol{x}_1, \boldsymbol{x}_2 \in E$ with $\|\boldsymbol{x}_1 - \boldsymbol{x}_2\| < \delta$, we have $\left|\dfrac{\partial g}{\partial x}(\boldsymbol{x}_1) - \dfrac{\partial g}{\partial x}(\boldsymbol{x}_2)\right| < \varepsilon$. In particular, when $\boldsymbol{x}_1 = (\xi, y)$ and $\boldsymbol{x}_2 = (x_0, y)$ with $\|\boldsymbol{x}_1 - \boldsymbol{x}_2\| = |\xi - x_0| < \delta$, we have

$$\left|\frac{\partial g}{\partial x}(\xi, y) - \frac{\partial g}{\partial x}(x_0, y)\right| < \varepsilon\,.$$

Therefore

$$\begin{aligned}
&\left|\frac{f(x) - f(x_0)}{x - x_0} - \int_a^b \frac{\partial g}{\partial x}(x_0, y)\,\mathrm{d}y\right| \\
&\qquad = \left|\int_a^b \left(\frac{g(x,y) - g(x_0,y)}{x - x_0} - \frac{\partial g}{\partial x}(x_0, y)\right)\,\mathrm{d}y\right| \\
&\qquad = \left|\int_a^b \left(\frac{\partial g}{\partial x}(\xi, y) - \frac{\partial g}{\partial x}(x_0, y)\right)\,\mathrm{d}y\right| < \varepsilon(b-a)
\end{aligned}$$

which proves differentiability at x_0 and also the formula

$$f'(x_0) = \int_a^b \frac{\partial g}{\partial x}(x_0, y)\,\mathrm{d}y\,. \qquad \square$$

▶ Proof of Theorem 6.18, p. 215

Proposition 6.18 *If α and β are continuous on I, the map f defined by* (6.15) *is continuous on I. If moreover g admits continuous partial derivative $\frac{\partial g}{\partial x}$ on $\mathcal{R}$ and α, β are $\mathcal{C}^1$ on I, then f is $\mathcal{C}^1$ on I, and*

$$f'(x) = \int_{\alpha(x)}^{\beta(x)} \frac{\partial g}{\partial x}(x, y)\,\mathrm{d}y + \beta'(x) g\big(x, \beta(x)\big) - \alpha'(x) g\big(x, \alpha(x)\big)\,.$$

Proof. The only thing to prove is the continuity of f, because the rest is shown on p. 215.

As in the previous argument, we may fix $x_0 \in I$ and assume it is an interior point. Call $E = [x_0 - \sigma, x_0 + \sigma] \times J \subset \mathcal{R}$, the set on which g is uniformly continuous. Let now $\varepsilon > 0$; since g is uniformly continuous on E and by the continuity of the maps α and β on I, there is a number $\delta > 0$ (with $\delta < \sigma$) such that $|x - x_0| < \delta$ implies

$$|g(x,y) - g(x_0,y)| < \varepsilon\,, \qquad \text{for all } y \in J\,,$$

and

$$|\alpha(x) - \alpha(x_0)| < \varepsilon\,, \qquad |\beta(x) - \beta(x_0)| < \varepsilon\,.$$

Then, setting $M = \max_{(x,y)\in E} |g(x,y)|$ gives

$$|f(x) - f(x_0)| = \left|\int_{\alpha(x)}^{\beta(x)} g(x,y)\,\mathrm{d}y - \int_{\alpha(x_0)}^{\beta(x_0)} g(x_0,y)\,\mathrm{d}y\right.$$

$$
\begin{aligned}
&= \int_{\alpha(x)}^{\alpha(x_0)} g(x,y)\,\mathrm{d}y + \int_{\alpha(x_0)}^{\beta(x_0)} g(x,y)\,\mathrm{d}y + \\
&\qquad + \int_{\beta(x_0)}^{\beta(x)} g(x,y)\,\mathrm{d}y - \int_{\alpha(x_0)}^{\beta(x_0)} g(x_0,y)\,\mathrm{d}y \\
&= \int_{\alpha(x)}^{\alpha(x_0)} g(x,y)\,\mathrm{d}y + \int_{\beta(x_0)}^{\beta(x)} g(x,y)\,\mathrm{d}y + \\
&\qquad - \int_{\alpha(x_0)}^{\beta(x_0)} \big(g(x,y) - g(x_0,y)\big)\,\mathrm{d}y \\
&\le M|\alpha(x) - \alpha(x_0)| + M|\beta(x) - \beta(x_0)| + \varepsilon|\beta(x_0) - \alpha(x_0)| \\
&\le \big(2M + |\beta(x_0) - \alpha(x_0)|\big)\varepsilon
\end{aligned}
$$

and f's continuity at x_0 follows immediately. □

A.1.4 The Implicit Function Theorem

▶ Proof of Theorem 7.1, p. 263

Teorema 7.1 *Let Ω be a non-empty open set in $\mathbb{R}^2$ and $f : \Omega \to \mathbb{R}$ a $\mathcal{C}^1$ map. Assume at the point $(x_0, y_0) \in \Omega$ we have $f(x_0, y_0) = 0$. If $\dfrac{\partial f}{\partial y}(x_0, y_0) \neq 0$, there exists a neighbourhood I of x_0 and a function $\varphi : I \to \mathbb{R}$ such that:*

i) $\big(x,\varphi(x)\big) \in \Omega$ *for any* $x \in I$;
ii) $y_0 = \varphi(x_0)$;
iii) $f\big(x,\varphi(x)\big) = 0$ *for any* $x \in I$;
iv) φ *is a* $\mathcal{C}^1$ *map on* I *with derivative*

$$
\varphi'(x) = -\frac{\dfrac{\partial f}{\partial x}\big(x,\varphi(x)\big)}{\dfrac{\partial f}{\partial y}\big(x,\varphi(x)\big)}. \tag{A.1.3}
$$

On a neighbourhood of (x_0, y_0) moreover, the zero set of f coincides with the graph of φ.

Proof. Since the map $f_y = \frac{\partial f}{\partial y}$ is continuous and non-zero at (x_0, y_0), the local invariance of the function's sign (§ 4.5.1) guarantees there exists a neighbourhood $A \subseteq \Omega$ of (x_0, y_0) where $f_y \neq 0$ has constant sign. On a such neighbourhood the auxiliary map $g(x, y) = -f_x(x, y)/f_y(x, y)$ is well defined and continuous.

Consider then the Cauchy problem (Section 10.4)

$$\{ y' = g(x,y)y(x_0) = y_0 \,.$$

This admits, by Peano's Theorem 10.10, a solution $y = \varphi(x)$ defined and of class $\mathcal{C}^1$ on a neighbourhood I of x_0 such that $\big(x,\varphi(x)\big) \in A$ for any $x \in I$. Thus conditions *i)* and *ii)* hold; but then also iv) is satisfied, by definition of solution to the differential equation. As for *iii)*, we define the map $h(x) = f\big(x,\varphi(x)\big)$: it satisfies, on I,

$$\begin{aligned}
h'(x) &= f_x\big(x,\varphi(x)\big) + f_y\big(x,\varphi(x)\big)\varphi'(x) \\
&= f_x\big(x,\varphi(x)\big) + f_y\big(x,\varphi(x)\big)g\big(x,\varphi(x)\big) \\
&= f_x\big(x,\varphi(x)\big) + f_y\big(x,\varphi(x)\big)\left(-\frac{f_x\big(x,\varphi(x)\big)}{f_y\big(x,\varphi(x)\big)}\right) = 0\,,
\end{aligned}$$

so it is constant; but as $h(x_0) = f(x_0, y_0) = 0$, h is necessarily the zero map, wherefore *iii)*.

Concerning the last statement, note that if $(x,y) \in A$ with $x \in I$, then

$$f(x,y) = f\big(x,\varphi(x)\big) + \int_{\varphi(x)}^{y} \frac{\partial f}{\partial y}(x,s)\,\mathrm{d}s = \int_{\varphi(x)}^{y} \frac{\partial f}{\partial y}(x,s)\,\mathrm{d}s$$

as a consequence of the Fundamental Theorem of Calculus (Vol. I, Cor. 9.42) applied to $y \mapsto f(x,y)$. The integral vanishes if and only if $y = \varphi(x)$, because the integrand is always different from 0 (recall the property stated in Vol. I, Thm. 9.33 *iii)*). Therefore on the neighbourhood of (x_0, y_0) where $x \in I$ we have $f(x,y) = 0$ precisely when $y = \varphi(x)$. □

A.2

Complements on integral calculus

In this appendix, we first introduce the notion of norm of a function, illustrated by several examples including norms of integral type. Next, we justify the Theorems of Gauss, Green and Stokes; for the sake of clarity, we confine our discussion to the case of specific, yet representative, geometries. The proof of the equivalence between conservative fields and irrotational fields in simply connected domains is the subsequent result. In the last section, we briefly outline the language of differential forms, and we express various properties of vector fields, discussed in the text, using the corresponding terminology.

A.2.1 Norms of functions

The norm of a function is a non-negative real number that somehow provides a measure of the "size" of the function. For instance, if f is a function and $\tilde{f}$ is another function approximating it, the norm of the difference $f - \tilde{f}$ gives a quantitative indication of the quality of the approximation: a small value of the norm corresponds, in a suitable sense, to a good approximation of f by means of $\tilde{f}$.

Definition A.2.1 *Given a family $\mathcal{F}$ of real functions defined on a set $\Omega \subseteq \mathbb{R}^n$, that forms a vector space, we call* **norm** *a map from $\mathcal{F}$ to $\mathbb{R}$, denoted by $f \mapsto \|f\|$, that fulfills the following properties: for all $f, g \in \mathcal{F}$ and all $\alpha \in \mathbb{R}$ one has*

i) $\|f\| \geq 0$ *and* $\|f\| = 0$ *if and only if* $f = 0$ (**positivity**);

ii) $\|\alpha f\| = |\alpha| \|f\|$ (**homogeneity**);

iii) $\|f + g\| \leq \|f\| + \|g\|$ (**triangle inequality**).

Note the analogy with the properties that define a norm over vectors in $\mathbb{R}^n$.

C. Canuto, A. Tabacco: *Mathematical Analysis II,* 2nd Ed.,
UNITEXT – La Matematica per il 3+2 85, DOI 10.1007/978-3-319-12757-6_A2,

Remarkable examples of norms of functions are as follows. If $\mathcal{F}$ denotes the vector space of the bounded functions on the set Ω, it is easily checked that

$$\boxed{\|f\|_{\infty,\Omega} = \sup_{\boldsymbol{x}\in\Omega} |f(\boldsymbol{x})|}$$

is a norm, called the **supremum norm** or **infinity norm**. It has been introduced in Sect. 2.1 for $\Omega = A \subseteq \mathbb{R}$. If in addition Ω is a compact subset of $\mathbb{R}^n$ and we restrict ourselves to consider the continuous functions on Ω, namely $f \in \mathcal{C}^0(\Omega)$, then the supremum in the previous definition is actually a maximum, as the function $|f(\boldsymbol{x})|$ is continuous on Ω and Weierstrass' Theorem 5.24 applies to it. So we define

$$\boxed{\|f\|_{\mathcal{C}^0(\Omega)} = \max_{\boldsymbol{x}\in\Omega} |f(\boldsymbol{x})| = \|f\|_{\infty,\Omega}\,,}$$

which is called the **maximum norm** in the space $\mathcal{C}^0(\Omega)$.

Other commonly used norms are those of integral type, which measure the size of a function "in the average". If Ω is a measurable set and $\mathcal{F}$ is the vector space of all Riemann-integrable functions on Ω (recall Theorem 8.20), we may define the **absolute-value norm**, or 1**-norm**, as

$$\boxed{\|f\|_{1,\Omega} = \int_\Omega |f(\boldsymbol{x})|\,\mathrm{d}\boldsymbol{x}}$$

as well as the **quadratic norm**, or 2**-norm**, as

$$\boxed{\|f\|_{2,\Omega} = \left(\int_\Omega |f(\boldsymbol{x})|^2\,\mathrm{d}\boldsymbol{x}\right)^{1/2}\,.}$$

The latter norm has been introduced in Sect. 3.2 for $\Omega = [0, 2\pi] \subset \mathbb{R}$. The two integral norms defined above are instances in the family of p**-norms**, with real $1 \leq p < +\infty$, defined as

$$\boxed{\|f\|_{p,\Omega} = \left(\int_\Omega |f(\boldsymbol{x})|^p\,\mathrm{d}\boldsymbol{x}\right)^{1/p}\,.}$$

The quadratic norm is particularly important, since it is associated to a **scalar product** between integrable functions; its definition is

$$\boxed{(f, g)_{2,\Omega} = \int_\Omega f(\boldsymbol{x})g(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x},}$$

and one has $\|f\|_{2,\Omega} = \sqrt{(f,f)_{2,\Omega}}$. In general, the following definition applies.

Definition A.2.2 *A* **scalar product** (*or* **inner product**) *in* $\mathcal{F}$ *is a map from* $\mathcal{F} \times \mathcal{F}$ *to* $\mathbb{R}$*, denoted by* $f, g \mapsto (f, g)$*, that fulfills the following properties: for all* $f, g, f_1, f_2 \in \mathcal{F}$ *and all* $\alpha, \beta \in \mathbb{R}$ *one has*

i) $(f, f) \geq 0$ *and* $(f, f) = 0$ *if and only if* $f = 0$ (**positivity**);

ii) $(f, g) = (g, f)$ (**symmetry**);

iii) $(\alpha f_1 + \beta f_2, g) = \alpha(f_1, g) + \beta(f_2, g)$ (**linearity**).

It is easily checked that the quantity $\|f\| = (f, f)^{1/2}$ is a norm, called the *norm associated with the scalar product* under consideration. It satisfies the **Cauchy-Schwarz inequality**

$$\boxed{|(f, g)| \leq \|f\| \, \|g\| \,, \qquad \forall f, g \in \mathcal{F} \,.}$$

A scalar product allows us to define the concept of orthogonality between functions: two functions $f, g \in \mathcal{F}$ are called **orthogonal** if $(f, g) = 0$. In a vector space endowed with a scalar product, the **Theorem of Pythagoras** holds; it is expressed by the relation

$$\boxed{\|f + g\|^2 = \|f\|^2 + \|g\|^2 \qquad \text{if and only if} \qquad (f, g) = 0 \,.}$$

Indeed, one has

$$\begin{aligned} \|f + g\|^2 &= (f + g, f + g) \\ &= (f, f) + (f, g) + (g, f) + (g, g) \\ &= \|f\|^2 + 2(f, g) + \|g\|^2 \,, \end{aligned}$$

whence the equivalence.

Going back to norms, for a differentiable function it may be useful to measure the size of its derivatives, in addition to that of the function. For instance, if Ω is a bounded open set in $\mathbb{R}^n$, consider the vector space $\mathcal{C}^1(\overline{\Omega})$ of the functions of class $\mathcal{C}^1$ on the compact set $\overline{\Omega}$ (see Sect. 5.4); then, it is natural to define therein the norm

$$\boxed{\|f\|_{\mathcal{C}^1(\overline{\Omega})} = \|f\|_{\mathcal{C}^0(\overline{\Omega})} + \sum_{i=1}^{n} \|D_{x_i} f\|_{\mathcal{C}^0(\overline{\Omega})} \,.}$$

If the maximum norms in this definition are replaced by norms of integral type (such as the quadratic norms of f and its first-order partial derivatives $D_{x_i} f$), we obtain the so-called *Sobolev norms*.

A.2.2 The Theorems of Gauss, Green, and Stokes

▶ Proof of Proposition 9.30, p. 391

Proposition 9.30 *Let the open set $\Omega \subset \mathbb{R}^3$ be G-admissible, and assume $f \in \mathcal{C}^0(\overline{\Omega})$ with $\dfrac{\partial f}{\partial x_i} \in \mathcal{C}^0(\overline{\Omega})$, $i \in \{1,2,3\}$. Then*

$$\int_\Omega \frac{\partial f}{\partial x_i} \,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \int_{\partial\Omega} f n_i \,\mathrm{d}\sigma\,,$$

where n_i is the ith component of the outward normal to $\partial\Omega$.

Proof. Without loss of generality take $i = 3$. As claimed, we shall prove the statement only for open, piecewise-regular sets that are normal for $x_3 = z$. We begin by assuming further that Ω is regular and normal for z, and use the notation introduced in Example 9.27 ii).

Integrating along segments and then by parts we obtain

$$\begin{aligned}\int_\Omega \frac{\partial f}{\partial z}(x,y,z)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z &= \int_D \int_{\alpha(x,y)}^{\beta(x,y)} \frac{\partial f}{\partial z}(x,y,z)\,\mathrm{d}z\,\mathrm{d}x\,\mathrm{d}y \\ &= \int_D f\big(x,y,\beta(x,y)\big)\,\mathrm{d}x\,\mathrm{d}y - \int_D f\big(x,y,\alpha(x,y)\big)\,\mathrm{d}x\,\mathrm{d}y\,.\end{aligned} \tag{A.2.1}$$

As $n_{z|\Sigma_\beta} = 1/\|\boldsymbol{\nu}_\beta\|$, recalling (9.12) and (9.11), we have

$$\begin{aligned}\int_D f\big(x,y,\beta(x,y)\big)\,\mathrm{d}x\,\mathrm{d}y &= \int_D f\big(x,y,\beta(x,y)\big) n_{z|\Sigma_\beta}(x,y)\|\boldsymbol{\nu}_\beta(x,y)\|\,\mathrm{d}x\,\mathrm{d}y \\ &= \int_{\Sigma_\beta} f n_z\,\mathrm{d}\sigma\,.\end{aligned}$$

Likewise, $n_{z|\Sigma_\alpha} = -1/\|\boldsymbol{\nu}_\alpha\|$, hence

$$\begin{aligned}-\int_D f\big(x,y,\alpha(x,y)\big)\,\mathrm{d}x\,\mathrm{d}y &= \int_D f\big(x,y,\alpha(x,y)\big) n_{z|\Sigma_\alpha}(x,y)\|\boldsymbol{\nu}_\alpha(x,y)\|\,\mathrm{d}x\,\mathrm{d}y \\ &= \int_{\Sigma_\alpha} f n_z\,\mathrm{d}\sigma\,.\end{aligned}$$

Eventually, from $n_{z|\Sigma_\ell} = 0$ follows

$$\int_{\Sigma_\ell} f n_z\,\mathrm{d}\sigma = 0\,.$$

We conclude that

$$\int_{\partial\Omega} f n_z \,\mathrm{d}\sigma = \int_{\Sigma_\beta} f n_z \,\mathrm{d}\sigma + \int_{\Sigma_\alpha} f n_z \,\mathrm{d}\sigma + \int_{\Sigma_\ell} f n_z \,\mathrm{d}\sigma$$

$$= \int_D f\big(x, y, \beta(x,y)\big)\,\mathrm{d}x\,\mathrm{d}y - \int_D f\big(x, y, \alpha(x,y)\big)\,\mathrm{d}x\,\mathrm{d}y$$

and the assertion follows from $(A.2.1)$.

Now let us suppose the open set is piecewise regular and normal for z, as in Example 9.27 iii). Call $\{\Omega_k\}_{k=1,\ldots,K}$ a partition of Ω into regular, normal sets for z. Using the above result on each Ω_k, we have

$$\int_\Omega \frac{\partial f}{\partial z}\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \sum_{k=1}^{K} \int_{\partial\Omega_k} f n_z^{(k)} \,\mathrm{d}\sigma\,. \tag{A.2.2}$$

If $\Gamma = \partial\Omega_k \cap \partial\Omega_h$ denotes the intersection of two partition elements (intersection which we assume bigger than a point), then $\boldsymbol{n}^{(k)}_{|\Gamma} = -\boldsymbol{n}^{(h)}_{|\Gamma}$ and

$$\int_\Gamma f n_z^{(k)}\,\mathrm{d}\sigma + \int_\Gamma f n_z^{(h)}\,\mathrm{d}\sigma = 0\,.$$

In other terms, the integrals over the parts of boundary of each Ω_k that are contained in Ω cancel out in pairs; what remains on the right-hand side of $(A.2.2)$ is

$$\sum_{k=1}^{K} \int_{\partial\Omega_k \cap \partial\Omega} f n_z^{(k)}\,\mathrm{d}\sigma = \int_{\partial\Omega} f n_z\,\mathrm{d}\sigma\,,$$

proving the claim. □

► Proof of Green's Theorem, p. 394

Theorem 9.35 (Green) *Let $\Omega \subset \mathbb{R}^2$ be a G-admissible open set whose boundary $\partial\Omega$ is positively oriented. Take a vector field $\boldsymbol{f} = f_1\boldsymbol{i} + f_2\boldsymbol{j}$ in $\big(\mathcal{C}^1(\overline{\Omega})\big)^2$. Then*

$$\int_\Omega \Big(\frac{\partial f_2}{\partial x} - \frac{\partial f_1}{\partial y}\Big)\,\mathrm{d}x\,\mathrm{d}y = \oint_{\partial\Omega} \boldsymbol{f}\cdot\boldsymbol{\tau}\,.$$

Proof. From (6.6) we know that $\operatorname{curl}\boldsymbol{f} = (\mathbf{curl}\,\boldsymbol{\Phi})_3$, $\boldsymbol{\Phi}$ being the three-dimensional vector field $\boldsymbol{f} + 0\boldsymbol{k}$ (constant in z). Setting $Q = \Omega \times (0,1)$ as in the proof of Theorem 9.32,

$$\int_\Omega \Big(\frac{\partial f_2}{\partial x} - \frac{\partial f_1}{\partial y}\Big)\,\mathrm{d}x\,\mathrm{d}y = \int_\Omega \operatorname{curl}\boldsymbol{f}\,\mathrm{d}x\,\mathrm{d}y = \int_0^1 \int_\Omega (\mathbf{curl}\,\boldsymbol{\Phi})_3\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z$$

$$= \int_Q (\mathbf{curl}\,\boldsymbol{\Phi})_3\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z\,.$$

Let us then apply Theorem 9.34 to the field $\boldsymbol{\Phi} \in \big(\mathcal{C}^1(\overline{\Omega})\big)^3$, and consider the third component of equation (9.21), giving

$$\int_Q (\mathbf{curl}\boldsymbol{\Phi}\,)_3 \,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \int_{\partial Q} (\boldsymbol{N} \wedge \boldsymbol{\Phi})_3 \,\mathrm{d}\sigma\,,$$

with $\boldsymbol{N}$ being the unit normal outgoing from ∂Q. It is immediate to check $(\boldsymbol{N} \wedge \boldsymbol{\Phi})_3 = f_2 n_1 - f_1 n_2 = \boldsymbol{f} \cdot \boldsymbol{t}$ on $\partial\Omega \times (0,1)$, whereas $(\boldsymbol{N} \wedge \boldsymbol{\Phi})_3 = 0$ on $\Omega \times \{0\}$ and $\Omega \times \{1\}$. Therefore

$$\int_{\partial Q} (\boldsymbol{N} \wedge \boldsymbol{\Phi})_3 \,\mathrm{d}\sigma = \int_0^1 \int_{\partial\Omega} (\boldsymbol{N} \wedge \boldsymbol{\Phi})_3 \,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \int_{\partial\Omega} \boldsymbol{f} \cdot \boldsymbol{t} \,\mathrm{d}\gamma = \oint_{\partial\Omega} \boldsymbol{f} \cdot \boldsymbol{\tau}\,,$$

and the result follows. □

▶ Proof of Stokes' Theorem, p. 396

Theorem 9.37 (**Stokes**) *Let $\Sigma \subset \mathbb{R}^3$ be an S-admissible compact surface oriented by the unit normal $\boldsymbol{n}$; correspondingly, let the boundary $\partial\Sigma$ be oriented positively. Suppose the vector field $\boldsymbol{f}$, defined on an open set $A \subseteq \mathbb{R}^3$ containing Σ, is such that $\boldsymbol{f} \in \big(\mathcal{C}^1(A)\big)^3$. Then*

$$\int_\Sigma (\mathbf{curl}\,\boldsymbol{f}) \cdot \boldsymbol{n} = \oint_{\partial\Sigma} \boldsymbol{f} \cdot \boldsymbol{\tau}\,.$$

In other words, the flux of the curl of $\boldsymbol{f}$ across the surface equals the path integral of $\boldsymbol{f}$ along the surface's (closed) boundary.

Proof. We start with the case in which Σ is the surface (9.15) from Example 9.29, whose notations we retain; additionally, let us assume the function φ belongs to $\mathcal{C}^2(\mathcal{R})$. Where possible, partial derivatives will be denoted using subscripts x, y, z, and likewise for the components of normal and tangent vectors. Recalling (9.13) and the expression (9.16) for the unit normal of Σ, we have

$$\begin{aligned}\int_\Sigma (\mathbf{curl}\,\boldsymbol{f}) \cdot \boldsymbol{n} &= \int_{\mathcal{R}} (f_{2,z} - f_{3,y})\varphi_x + (f_{3,x} - f_{1,z})\varphi_y + (f_{2,x} - f_{1,y}) \,\mathrm{d}x\,\mathrm{d}y \\ &= \int_{\mathcal{R}} -(f_{1,y} + f_{1,z}\varphi_y) + (f_{2,x} + f_{3,z}\varphi_x) + (f_{3,x}\varphi_y - f_{3,y}\varphi_x) \,\mathrm{d}x\,\mathrm{d}y\,,\end{aligned}$$

where the derivatives of the components of $\boldsymbol{f}$ are taken on $\big(x, y, \varphi(x,y)\big)$ as (x,y) varies in $\mathcal{R}$. By the chain rule $f_{1,y} + f_{1,z}\varphi_y$ is the partial derivative in y of the map $f_1\big(x, y, \varphi(x,y)\big)$; similarly, $f_{2,x} + f_{3,z}\varphi_x$ represents the partial x-derivative of the map $f_2\big(x, y, \varphi(x,y)\big)$. Moreover, adding and subtracting $f_{3,z}\varphi_x\varphi_y + f_{3,y}\varphi^2_{xy}$ to formula $f_{3,x}\varphi_y - f_{3,y}\varphi_x$, we see that this expression equals

$$\frac{\partial}{\partial x}\Big(f_3\big(x,y,\varphi\,(x,y)\big)\varphi_y(x,y)\Big) - \frac{\partial}{\partial y}\Big(f_3\big(x,y,\varphi\,(x,y)\big)\varphi_x(x,y)\Big)\,.$$

Therefore we can use the two-dimensinal analogue of Proposition 9.30 to obtain

$$\int_\Sigma (\mathbf{curl}\,\boldsymbol{f})\cdot\boldsymbol{n} = \int_{\partial\mathcal{R}} \big(-f_1 n_y + f_2 n_x + f_3(\varphi_y n_x - \varphi_x n_y)\big)\,\mathrm{d}\gamma\,, \tag{A.2.3}$$

where n_x, n_y are the components of the outgoing unit normal of $\partial\mathcal{R}$.

If $\boldsymbol{\gamma} : I \to \mathbb{R}^2$ denotes a positive parametrisation of the boundary of $\mathcal{R}$, then $n_y = -\gamma_1'/\|\boldsymbol{\gamma}'\|$ and $n_x = \gamma_2'/\|\boldsymbol{\gamma}'\|$. Furthermore the arc $\boldsymbol{\eta} : I \to \mathbb{R}^3$ given by $\boldsymbol{\eta}(t) = \big(\gamma_1(t), \gamma_2(t), \varphi(\gamma_1(t), \gamma_2(t))\big)$ is a positive parametrisation of $\partial\Sigma$ with respect to the chosen orientation of Σ. The corresponding tangen vector is given by (9.17). Overall then, recalling the definitions of integral along a curve and line integral,

$$\begin{aligned}\int_{\partial\mathcal{R}} \big(-f_1 n_y + f_2 n_x + f_3(\varphi_y n_x - \varphi_x n_y)\big)\,\mathrm{d}\gamma &= \int_I (f_1\eta_1' + f_2\eta_2' + f_3\eta_3')\,\mathrm{d}t \\ &= \oint_{\partial\Sigma} \boldsymbol{f}\cdot\boldsymbol{\tau}\,.\end{aligned}$$

Equation (9.24) in the present case follows from $(A.2.3)$.

Let us now suppose Σ is made by K faces $\Sigma_1, \ldots, \Sigma_K$, each as above. Using the result just found on every Σ_k, and summing over k, we find

$$\int_\Sigma (\mathbf{curl}\,\boldsymbol{f})\cdot\boldsymbol{n} = \sum_{k=1}^K \oint_{\partial\Sigma_k} \boldsymbol{f}\cdot\boldsymbol{\tau}^{(k)}\,. \tag{A.2.4}$$

If $\Gamma = \partial\Sigma_h \cap \partial\Sigma_k$ is the intersection of two faces (suppose not a point), the unit tangents to $\partial\Sigma_h$ and $\partial\Sigma_k$ satisfy $\boldsymbol{t}^{(h)}_{|\Gamma} = -\boldsymbol{t}^{(k)}_{|\Gamma}$, hence

$$\int_\Gamma \boldsymbol{f}\cdot\boldsymbol{\tau}^{(h)} + \int_\Gamma \boldsymbol{f}\cdot\boldsymbol{\tau}^{(k)} = 0\,.$$

That is, the integrals over the boundary parts common to two faces cancel out; the right-hand side of $(A.2.2)$ thus reduces to

$$\sum_{k=1}^K \int_{\partial\Sigma_k\cap\partial\Sigma} \boldsymbol{f}\cdot\boldsymbol{\tau}^{(k)} = \oint_{\partial\Sigma} \boldsymbol{f}\cdot\boldsymbol{\tau}\,,$$

proving the assertion. □

▶ Proof of Theorem 9.45, p. 403

Theorem 9.45 *Let $\Omega \subseteq \mathbb{R}^n$, with $n = 2$ or 3, be open and simply connected. A vector field $\boldsymbol{f}$ of class $\mathcal{C}^1$ on Ω is conservative if and only if it is curl-free.*

Proof. We have shown, in Proposition 9.43, the arrow $\boldsymbol{f}$ conservative $\Rightarrow$ $\boldsymbol{f}$ curl-free. Let us deal with the opposite implication.

First, though, we handle the two-dimensional case, and prove condition iii) of Theorem 9.42. Suppose $\boldsymbol{\gamma}$ is a simple closed (i.e., Jordan) arc, (piecewise) regular and with trace Γ contained in Ω; its interior Σ is all contained Ω. Hence we can use Green's Theorem 9.35 and conclude

$$\int_{\boldsymbol{\gamma}} \boldsymbol{f} \cdot \boldsymbol{\tau} = \pm \int_{\Sigma} \mathbf{curl}\, \boldsymbol{f} \, \mathrm{d}x \, \mathrm{d}y = 0$$

(the sign on the second integral is determined by the orientation of Γ). Should the arc be not simple, we can decompose it in closed simple arcs to which the result applies.

Now let us consider the three-dimensional picture, for which we discuss only the case of star-shaped sets; build explicitly the potential $\boldsymbol{f}$, in analogy to–and using the notation of–the proof of Theorem 9.42. Precisely, if P_0 is the point for which Ω is star-shaped, we define the potential at P of coordinates $\boldsymbol{x}$ by setting

$$\varphi(\boldsymbol{x}) = \int_{\Gamma_{[P_0,P]}} \boldsymbol{f} \cdot \boldsymbol{\tau}$$

where $\Gamma_{[P_0,P]}$ is the segment joining P_0 and P. We claim $\mathbf{grad}\, \varphi = \boldsymbol{f}$, and will prove it for the first component only. So let $P + \Delta P = \boldsymbol{x} + \Delta x_1 \, \boldsymbol{e}_1 \in \Omega$ be a nearby point to P, $\Gamma_{[P_0,P+\Delta P]}$ the segment joining P_0 to $P + \Delta P$, and $\Gamma_{[P,P+\Delta P]}$ the (horizontal) segment from P to $P + \Delta P$. We wish to prove

$$\int_{\Gamma_{[P_0,P+\Delta P]}} \boldsymbol{f} \cdot \boldsymbol{\tau} - \int_{\Gamma_{[P_0,P]}} \boldsymbol{f} \cdot \boldsymbol{\tau} = \int_{\Gamma_{[P,P+\Delta P]}} \boldsymbol{f} \cdot \boldsymbol{\tau} \,. \qquad \text{(A.2.5)}$$

But this is straightforward if P_0, P and $P + \Delta P$ are collinear. If not, they form a triangle Σ, which is entirely contained in Ω, the latter being star-shaped. Calling Γ the boundary of Σ oriented from P_0 to P along $\Gamma_{[P_0,P]}$, we invoke Stokes' Theorem and have

$$\int_{\Gamma} \boldsymbol{f} \cdot \boldsymbol{\tau} = \int_{\Gamma} \mathbf{curl}\, \boldsymbol{f} \wedge \boldsymbol{n} = 0 \,,$$

i.e.,

$$\int_{\Gamma_{[P_0,P]}} \boldsymbol{f} \cdot \boldsymbol{\tau} + \int_{\Gamma_{[P_0,P+\Delta P]}} \boldsymbol{f} \cdot \boldsymbol{\tau} - \int_{\Gamma_{[P,P+\Delta P]}} \boldsymbol{f} \cdot \boldsymbol{\tau} = 0 \,,$$

whence $(A.2.5)$. Therefore

$$\begin{aligned} \frac{\varphi(\boldsymbol{x} + \Delta x_1 \boldsymbol{e}_1) - \varphi(\boldsymbol{x})}{\Delta x_1} &= \frac{1}{\Delta x_1} \int_{\Gamma_{[P,P+\Delta P]}} \boldsymbol{f} \cdot \boldsymbol{\tau} \\ &= \frac{1}{\Delta x_1} \int_0^{\Delta x_1} f_1(\boldsymbol{x} + t\boldsymbol{e}_1) \, \mathrm{d}t \,, \end{aligned}$$

and we conclude as in Theorem 9.42. □

A.2.3 Differential forms

In this section, we introduce a few essential notions about differential forms; through them, it is possible to reformulate various definitions and relevant properties of vector fields, that we encountered in previous chapters. The forthcoming exposition is deliberately informal and far from being complete; our goal indeed is just to establish a relation between certain notations adopted in this textbook and the language of differential forms, which is commonly used in various applications.

Let us assume to be in dimension 3; as usual, the reduction to dimension 2 is straightforward, while several concepts may actually be formulated in any dimension n. So, in the sequel Ω will be an open set in $\mathbb{R}^3$.

A **differential 0-form** F is, in our notation, a *real-valued function* (also called a *scalar field*) $\varphi = \varphi(x, y, z)$ defined in Ω, i.e.,

$$\boxed{F = \varphi \, .}$$

A **differential 1-form** ω is an expression like

$$\boxed{\omega = P \, \mathrm{d}x + Q \, \mathrm{d}y + R \, \mathrm{d}z \, ,}$$

where $P = P(x, y, z)$, $Q = Q(x, y, z)$ and $R = R(x, y, z)$ are scalar fields defined in Ω. In our notation, it corresponds to the *vector field*

$$\boldsymbol{f} = P \, \boldsymbol{i} + Q \, \boldsymbol{j} + R \, \boldsymbol{k} \, .$$

Thus, the symbols $\mathrm{d}x$, $\mathrm{d}y$ and $\mathrm{d}z$ denote particular 1-forms, that span all the others by linear combinations. They correspond to the vectors $\boldsymbol{i}$, $\boldsymbol{j}$ and $\boldsymbol{k}$, respectively, of the canonical basis in $\mathbb{R}^3$.

An expression like

$$\boxed{\int_\Gamma \omega = \int_\Gamma (P \, \mathrm{d}x + Q \, \mathrm{d}y + R \, \mathrm{d}z) \, ,}$$

where Γ is an arc contained in Ω, corresponds in our notation to the *path integral*

$$\int_\Gamma \boldsymbol{f} \cdot \boldsymbol{\tau} \, .$$

Formally, the relationship may be motivated by multiplying and dividing by $\mathrm{d}t$ under the integral sign,

$$\int_\Gamma \omega = \int_\Gamma \left(P \, \frac{\mathrm{d}x}{\mathrm{d}t} + Q \, \frac{\mathrm{d}y}{\mathrm{d}t} + R \, \frac{\mathrm{d}z}{\mathrm{d}t} \right) \, \mathrm{d}t$$

and thinking of $\boldsymbol{\gamma}(t) = (x(t), y(t), z(t))$ as the parametrization of the arc Γ (recall (9.8) and (9.10)).

The **derivative** of a 0-form F is defined as the 1-form

$$\boxed{\mathrm{d}F = \frac{\partial\varphi}{\partial x}\,\mathrm{d}x + \frac{\partial\varphi}{\partial y}\,\mathrm{d}y + \frac{\partial\varphi}{\partial z}\,\mathrm{d}z\,,}$$

obviously assuming the scalar field φ differentiable in Ω. In our notation, this relation is expressed as

$$\boldsymbol{f} = \mathbf{grad}\,\varphi\,.$$

A differential 1-form ω is called **exact** if there exists a 0-form F such that

$$\boxed{\omega = \mathrm{d}F\,.}$$

In our notation, this is equivalent to the property that the vector field $\boldsymbol{f}$ associated with the form ω is *conservative*.

It is possible to define the **derivative** of a 1-form $\omega = P\,\mathrm{d}x + Q\,\mathrm{d}y + R\,\mathrm{d}z$ as the 2-form

$$\boxed{\mathrm{d}\omega = \left(\frac{\partial R}{\partial y} - \frac{\partial Q}{\partial z}\right)\mathrm{d}y\wedge\mathrm{d}z + \left(\frac{\partial P}{\partial z} - \frac{\partial R}{\partial x}\right)\mathrm{d}z\wedge\mathrm{d}x + \left(\frac{\partial Q}{\partial x} - \frac{\partial P}{\partial y}\right)\mathrm{d}x\wedge\mathrm{d}y\,.}$$

If we identify the symbols $\mathrm{d}x$, $\mathrm{d}y$ and $\mathrm{d}z$ with the vectors of the canonical basis in $\mathbb{R}^3$ and we formally interpret the symbol $\wedge$ as the external product of two vectors (recall (4.5)), then we have $\mathrm{d}y \wedge \mathrm{d}z = \mathrm{d}x$, $\mathrm{d}z \wedge \mathrm{d}x = \mathrm{d}y$ and $\mathrm{d}x \wedge \mathrm{d}y = \mathrm{d}z$; this shows that, with respect to our notation, the differential form $\mathrm{d}\omega$ is associated with the vector field

$$\boldsymbol{g} = \mathbf{curl}\,\boldsymbol{f}$$

(recall (6.4)). In general, for a differential 2-form

$$\Psi = S\,\mathrm{d}y\wedge\mathrm{d}z + T\,\mathrm{d}z\wedge\mathrm{d}x + U\,\mathrm{d}x\wedge\mathrm{d}y\,,$$

where $S = S(x, y, z)$, $T = T(x, y, z)$ and $U = U(x, y, z)$ are scalar fields, it is possible to define the integral

$$\int_\Sigma \Psi$$

over a surface Σ contained in Ω; it corresponds, in our notation, to the *flux integral*

$$\int_\Sigma \boldsymbol{g}\cdot\boldsymbol{n}$$

(recall (9.14)), where $\boldsymbol{g} = S\boldsymbol{i} + T\boldsymbol{j} + U\boldsymbol{k}$ is the vector field associated with the form Ψ. In particular, we have

$$\int_\Sigma \mathrm{d}\omega = \int_\Sigma (\mathbf{curl}\,\boldsymbol{f})\cdot\boldsymbol{n}\,.$$

Hence, in the language of differential forms, Stokes' Theorem 9.37 takes the elegant expression

$$\boxed{\int_{\Sigma} \mathrm{d}\omega = \int_{\partial\Sigma} \omega \, .}$$

A differential 1-form ω is called **closed** if

$$\boxed{\mathrm{d}\omega = 0 \, ;}$$

in our notation, this corresponds to the property that the vector field $\boldsymbol{f}$ associated with the form is *irrotational*, i.e., it satisfies

$$\mathbf{curl}\, \boldsymbol{f} = \mathbf{0} \, .$$

One has the property

$$\boxed{\mathrm{d}^2 F = \mathrm{d}\,(\mathrm{d}F) = 0}$$

that in our notation is equivalent to the identity

$$\mathbf{curl}\,\mathbf{grad}\,\varphi = \mathbf{0}$$

(recall Proposition 6.7). Such a property may be formulated as

$$\boxed{\textit{an exact differential 1-form is closed}}$$

that corresponds to state that a *conservative vector field is irrotational* (see Property 9.43). If the domain Ω is simply connected, then we have the equivalence

$$\boxed{\textit{a differential 1-form is exact if and only if it is closed}}$$

that translates into the language of differential forms our Theorem 9.45, according to which in such a domain *a vector field is conservative if and only if it is irrotational*.

Basic definitions and formulas

Sequences and series

Geometric sequence (p. 3):

$$\lim_{n\to\infty} q^n = \begin{cases} 0 & \text{if } |q| < 1, \\ 1 & \text{if } q = 1, \\ +\infty & \text{if } q > 1, \\ \text{does not exist} & \text{if } q \le -1 \end{cases}$$

The number e (p. 3):

$$\mathrm{e} = \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n = \sum_{k=0}^{\infty}\frac{1}{n!}$$

Geometric series (p. 6):

$$\sum_{k=0}^{\infty} q^k \quad \begin{cases} \text{converges to } \dfrac{1}{1-q} & \text{if } |q| < 1\,, \\ \text{diverges to } \; +\infty & \text{if } q \ge 1\,, \\ \text{is indeterminate} & \text{if } q \le -1 \end{cases}$$

Mengoli's series (p. 7):

$$\sum_{k=2}^{\infty}\frac{1}{(k-1)k} = 1$$

Generalised harmonic series (p. 15):

$$\sum_{k=1}^{\infty}\frac{1}{k^\alpha} \quad \begin{cases} \text{converges if } \alpha > 1\,, \\ \text{diverges if } \alpha \le 1 \end{cases}$$

C. Canuto, A. Tabacco: *Mathematical Analysis II*, 2nd Ed.,
UNITEXT – La Matematica per il 3+2 85, DOI 10.1007/978-3-319-12757-6,

Power series

Convergence radius (p. 48):

$$R = \sup \Big\{ x \in \mathbb{R} : \sum_{k=0}^{\infty} a_k x^k \text{ converges} \Big\}$$

Ratio Test (p. 49):

$$\lim_{k\to\infty} \left| \frac{a_{k+1}}{a_k} \right| = \ell \quad \Longrightarrow \quad R = \begin{cases} 0 & \text{if } \ell = +\infty\,, \\ +\infty & \text{if } \ell = 0\,, \\ 1/\ell & \text{if } 0 < \ell < +\infty \end{cases}$$

Root Test (p. 50):

$$\lim_{k\to\infty} \sqrt[k]{|a_k|} = \ell \quad \Longrightarrow \quad R = \begin{cases} 0 & \text{if } \ell = +\infty\,, \\ +\infty & \text{if } \ell = 0\,, \\ 1/\ell & \text{if } 0 < \ell < +\infty \end{cases}$$

Power series for analytic functions (p. 56):

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k$$

Special power series (p. 54 and 58):

$$\frac{1}{1-x} = \sum_{k=0}^{\infty} x^k\,, \qquad x \in (-1,1)$$

$$\frac{1}{1+x^2} = \sum_{k=0}^{\infty} (-1)^k x^{2k}\,, \qquad x \in (-1,1)$$

$$(1+x)^\alpha = \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k\,, \qquad x \in (-1,1)$$

$$\mathrm{e}^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}\,, \qquad x \in \mathbb{R}$$

$$\log(1+x) = \sum_{k=0}^{\infty} \frac{(-1)^k}{k+1} x^{k+1} = \sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k} x^k\,, \qquad x \in (-1,1)$$

$$\sin x = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k+1)!}\, x^{2k+1}\,, \qquad x \in \mathbb{R}$$

$$\cos x = \sum_{k=0}^{\infty} \frac{(-1)^k}{(2k)!}\, x^{2k}\,, \qquad x \in \mathbb{R}$$

Fourier series

Fourier coefficients of a map f (p. 82):

$$a_0 = \frac{1}{2\pi}\int_0^{2\pi} f(x)\,\mathrm{d}x$$

$$a_k = \frac{1}{\pi}\int_0^{2\pi} f(x)\cos kx\,\mathrm{d}x\,, \qquad k \geq 1$$

$$b_k = \frac{1}{\pi}\int_0^{2\pi} f(x)\sin kx\,\mathrm{d}x\,, \qquad k \geq 1$$

$$c_k = \frac{1}{2\pi}\int_0^{2\pi} f(x)e^{-ikx}\,\mathrm{d}x\,, \qquad k \in \mathbb{Z}$$

Fourier series of a map $f \in \tilde{\mathcal{C}}_{2\pi}$ (p. 85):

$$f \approx a_0 + \sum_{k=1}^{\infty}(a_k \cos kx + b_k \sin kx) \approx \sum_{k=-\infty}^{+\infty} c_k e^{ikx}$$

Parseval's formula (p. 91):

$$\int_0^{2\pi} |f(x)|^2\,\mathrm{d}x = 2\pi a_0^2 + \pi\sum_{k=1}^{+\infty}(a_k^2 + b_k^2) = 2\pi\sum_{k=-\infty}^{+\infty}|c_k|^2$$

Square wave (p. 85):

$$f(x) = \begin{cases} -1 & \text{if } -\pi < x < 0\,, \\ 0 & \text{if } x = 0,\ \pm\pi\,, \\ 1 & \text{if } 0 < x < \pi\,, \end{cases} \qquad f \approx \frac{4}{\pi}\sum_{m=0}^{\infty}\frac{1}{2m+1}\sin(2m+1)x$$

Rectified wave (p. 87):

$$f(x) = |\sin x|\,, \qquad f \approx \frac{2}{\pi} - \frac{4}{\pi}\sum_{m=1}^{\infty}\frac{1}{4m^2-1}\cos 2mx$$

Sawtooth wave (p. 87):

$$f(x) = x\,, \quad x \in (-\pi,\pi)\,, \qquad f \approx \sum_{k=1}^{\infty}\frac{2}{k}(-1)^{k+1}\sin kx$$

Real-valued functions

Partial derivative (p. 157):

$$\frac{\partial f}{\partial x_i}(\boldsymbol{x}_0) = \lim_{\Delta x \to 0} \frac{f(\boldsymbol{x}_0 + \Delta x\, \boldsymbol{e}_i) - f(\boldsymbol{x}_0)}{\Delta x}$$

Gradient (p. 157):

$$\nabla f(\boldsymbol{x}_0) = \mathbf{grad}\, f(\boldsymbol{x}_0) = \left(\frac{\partial f}{\partial x_i}(\boldsymbol{x}_0)\right)_{1 \le i \le n}$$

Differential (p. 162):

$$\mathrm{d}f_{\boldsymbol{x}_0}(\Delta \boldsymbol{x}) = \nabla f(\boldsymbol{x}_0) \cdot \Delta \boldsymbol{x}$$

Directional derivative (p. 163):

$$\frac{\partial f}{\partial \boldsymbol{v}}(\boldsymbol{x}_0) = \nabla f(\boldsymbol{x}_0) \cdot \boldsymbol{v} = \frac{\partial f}{\partial x_1}(\boldsymbol{x}_0)\, v_1 + \cdots + \frac{\partial f}{\partial x_n}(\boldsymbol{x}_0)\, v_n$$

Second partial derivative (p. 168):

$$\frac{\partial^2 f}{\partial x_j \partial x_i}(\boldsymbol{x}_0) = \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right)(\boldsymbol{x}_0)$$

Hessian matrix (p. 169):

$$\boldsymbol{H}f(\boldsymbol{x}_0) = (h_{ij})_{1 \le i,j \le n} \qquad \text{with} \qquad h_{ij} = \frac{\partial^2 f}{\partial x_j \partial x_i}(\boldsymbol{x}_0)$$

Taylor expansion with Peano's remainder (p. 172):

$$f(\boldsymbol{x}) = f(\boldsymbol{x}_0) + \nabla f(\boldsymbol{x}_0) \cdot (\boldsymbol{x} - \boldsymbol{x}_0) + \frac{1}{2}(\boldsymbol{x} - \boldsymbol{x}_0) \cdot \boldsymbol{H}f(\boldsymbol{x}_0)(\boldsymbol{x} - \boldsymbol{x}_0) + o(\|\boldsymbol{x} - \boldsymbol{x}_0\|^2)$$

Curl in dimension 2 (p. 206):

$$\mathbf{curl}\, f = \frac{\partial f}{\partial x_2}\boldsymbol{i} - \frac{\partial f}{\partial x_1}\boldsymbol{j}$$

Fundamental identity (p. 209):

$$\mathbf{curl}\,\mathbf{grad}\, f = \nabla \wedge (\nabla f) = \mathbf{0}$$

Laplace operator (p. 211):

$$\Delta f = \operatorname{div} \mathbf{grad}\, f = \nabla \cdot \nabla f = \sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_j^2}$$

Vector-valued functions

Jacobian matrix (p. 202):

$$\boldsymbol{J}\boldsymbol{f}(\boldsymbol{x}_0) = \Big(\frac{\partial f_i}{\partial x_j}(\boldsymbol{x}_0)\Big)_{\substack{1 \le i \le m \\ 1 \le j \le n}} = \begin{pmatrix} \nabla f_1(\boldsymbol{x}_0) \\ \vdots \\ \nabla f_m(\boldsymbol{x}_0) \end{pmatrix}$$

Differential (p. 203):

$$\mathbf{d}\boldsymbol{f}_{\boldsymbol{x}_0}(\Delta\boldsymbol{x}) = \boldsymbol{J}\boldsymbol{f}(\boldsymbol{x}_0)\Delta\boldsymbol{x}$$

Divergence (p. 205):

$$\operatorname{div}\boldsymbol{f} = \nabla\cdot\boldsymbol{f} = \frac{\partial f_1}{\partial x_1} + \cdots + \frac{\partial f_n}{\partial x_n} = \sum_{j=1}^{n}\frac{\partial f_j}{\partial x_j}$$

Curl in dimension 3 (p. 205):

$$\mathbf{curl}\,\boldsymbol{f} = \nabla\wedge\boldsymbol{f} = \Big(\frac{\partial f_3}{\partial x_2} - \frac{\partial f_2}{\partial x_3}\Big)\boldsymbol{i} + \Big(\frac{\partial f_1}{\partial x_3} - \frac{\partial f_3}{\partial x_1}\Big)\boldsymbol{j} + \Big(\frac{\partial f_2}{\partial x_1} - \frac{\partial f_1}{\partial x_2}\Big)\boldsymbol{k}$$

Curl in dimension 2 (p. 205):

$$\operatorname{curl}\boldsymbol{f} = \frac{\partial f_2}{\partial x_1} - \frac{\partial f_1}{\partial x_2}$$

Fundamental identity (p. 209):

$$\operatorname{div}\mathbf{curl}\,\boldsymbol{f} = \nabla\cdot(\nabla\wedge\boldsymbol{f}) = 0$$

Derivative of a composite map – Chain rule (p. 212):

$$\boldsymbol{J}(\boldsymbol{g}\circ\boldsymbol{f})(\boldsymbol{x}_0) = \boldsymbol{J}\boldsymbol{g}(\boldsymbol{y}_0)\,\boldsymbol{J}\boldsymbol{f}(\boldsymbol{x}_0)$$

Tangent line to a curve (p. 217):

$$T(t) = \boldsymbol{\gamma}(t_0) + \boldsymbol{\gamma}'(t_0)(t-t_0)\,, \qquad t\in\mathbb{R}$$

Length of a curve (p. 222):

$$\ell(\boldsymbol{\gamma}) = \int_a^b \|\boldsymbol{\gamma}'(t)\|\,\mathrm{d}t$$

Tangent plane to a surface (p. 237):

$$\Pi(u,v) = \boldsymbol{\sigma}(u_0,v_0) + \frac{\partial\boldsymbol{\sigma}}{\partial u}(u_0,v_0)\,(u-u_0) + \frac{\partial\boldsymbol{\sigma}}{\partial v}(u_0,v_0)\,(v-v_0)$$

Normal vector of a surface (p. 238):

$$\boldsymbol{\nu}(u_0,v_0) = \frac{\partial\boldsymbol{\sigma}}{\partial u}(u_0,v_0)\wedge\frac{\partial\boldsymbol{\sigma}}{\partial v}(u_0,v_0)$$

Polar coordinates

From polar to Cartesian coordinates (p. 230):

$$\mathbf{\Phi} : [0, +\infty) \times \mathbb{R} \to \mathbb{R}^2 , \qquad (r,\theta) \mapsto (x, y) = (r\cos\theta, r\sin\theta)$$

Jacobian matrix and determinant (p. 230):

$$\boldsymbol{J\Phi}(r,\theta) = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix} , \qquad \det \boldsymbol{J\Phi}(r,\theta) = r$$

Partial derivatives in polar coordinates (p. 231):

$$\frac{\partial f}{\partial r} = \frac{\partial f}{\partial x}\cos\theta + \frac{\partial f}{\partial y}\sin\theta , \qquad \frac{\partial f}{\partial \theta} = -\frac{\partial f}{\partial x} r\sin\theta + \frac{\partial f}{\partial y} r\cos\theta$$

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial r}\cos\theta - \frac{\partial g}{\partial \theta}\frac{\sin\theta}{r} , \qquad \frac{\partial f}{\partial y} = \frac{\partial g}{\partial r}\sin\theta + \frac{\partial f}{\partial \theta}\frac{\cos\theta}{r}$$

Variable change in double integrals (p. 320):

$$\int_{\Omega} f(x, y)\,\mathrm{d}x\,\mathrm{d}y = \int_{\Omega'} f(r\cos\theta, r\sin\theta)\, r\,\mathrm{d}r\,\mathrm{d}\theta$$

Cylindrical coordinates

From cylindrical to Cartesian coordinates (p. 233):

$$\mathbf{\Phi} : [0, +\infty) \times \mathbb{R}^2 \to \mathbb{R}^3 , \qquad (r, \theta, t) \mapsto (x, y, z) = (r\cos\theta, r\sin\theta, t)$$

Jacobian matrix and determinant (p. 233):

$$\boldsymbol{J\Phi}(r, \theta, t) = \begin{pmatrix} \cos\theta & -r\sin\theta & 0 \\ \sin\theta & r\cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix} , \qquad \det \boldsymbol{J\Phi}(r, \theta, t) = r$$

Partial derivatives in cylindrical coordinates (p. 233):

$$\frac{\partial f}{\partial r} = \frac{\partial f}{\partial x}\cos\theta + \frac{\partial f}{\partial y}\sin\theta , \qquad \frac{\partial f}{\partial \theta} = -\frac{\partial f}{\partial x} r\sin\theta + \frac{\partial f}{\partial y} r\cos\theta , \qquad \frac{\partial f}{\partial t} = \frac{\partial f}{\partial z}$$

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial r}\cos\theta - \frac{\partial f}{\partial \theta}\frac{\sin\theta}{r} , \qquad \frac{\partial f}{\partial y} = \frac{\partial f}{\partial r}\sin\theta + \frac{\partial f}{\partial \theta}\frac{\cos\theta}{r} , \qquad \frac{\partial f}{\partial z} = \frac{\partial f}{\partial t}$$

Variable change in triple integrals (p. 328):

$$\int_{\Omega} f(x, y, z)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \int_{\Omega'} f(r\cos\theta, r\sin\theta, t)\, r\,\mathrm{d}r\,\mathrm{d}\theta\,\mathrm{d}t$$

Spherical coordinates

From spherical to Cartesian coordinates (p. 234):

$\mathbf{\Phi} : [0, +\infty) \times \mathbb{R}^2 \to \mathbb{R}^3$,

$(r, \varphi, \theta) \mapsto (x, y, z) = (r \sin\varphi \cos\theta, r \sin\varphi \sin\theta, r \cos\varphi)$

Jacobian matrix (p. 234):

$$\boldsymbol{J\Phi}(r, \varphi, \theta) = \begin{pmatrix} \sin\varphi\cos\theta & r\cos\varphi\cos\theta & -r\sin\varphi\sin\theta \\ \sin\varphi\sin\theta & r\cos\varphi\sin\theta & r\sin\varphi\cos\theta \\ \cos\varphi & -r\sin\varphi & 0 \end{pmatrix}$$

Jacobian determinant (p. 234):

$\det \boldsymbol{J\Phi}(r, \varphi, \theta) = r^2 \sin\varphi$

Partial derivatives in spherical coordinates (p. 233):

$$\frac{\partial f}{\partial r} = \frac{\partial f}{\partial x}\sin\varphi\cos\theta + \frac{\partial f}{\partial y}\sin\varphi\sin\theta + \frac{\partial f}{\partial z}\cos\varphi$$
$$\frac{\partial f}{\partial \varphi} = \frac{\partial f}{\partial x} r\cos\varphi\cos\theta + \frac{\partial f}{\partial y} r\cos\varphi\sin\theta - \frac{\partial f}{\partial z} r\sin\varphi$$
$$\frac{\partial f}{\partial \theta} = -\frac{\partial f}{\partial x} r\sin\varphi\sin\theta + \frac{\partial f}{\partial y} r\sin\varphi\cos\theta$$

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial r}\sin\varphi\cos\theta + \frac{\partial f}{\partial \varphi}\frac{\cos\varphi\cos\theta}{r} - \frac{\partial f}{\partial \theta}\frac{\sin\theta}{r\sin\varphi}$$
$$\frac{\partial f}{\partial y} = \frac{\partial f}{\partial r}\sin\varphi\sin\theta + \frac{\partial f}{\partial \varphi}\frac{\cos\varphi\sin\theta}{r} + \frac{\partial f}{\partial \theta}\frac{\cos\theta}{r\sin\varphi}$$
$$\frac{\partial f}{\partial z} = \frac{\partial f}{\partial r}\cos\varphi - \frac{\partial f}{\partial \varphi}\frac{\sin\varphi}{r}$$

Variable change in triple integrals (p. 329):

$$\int_{\Omega} f(x, y, z)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = \int_{\Omega'} f(r\sin\varphi\cos\theta, r\sin\varphi\sin\theta, r\cos\varphi)\, r^2 \sin\varphi\,\mathrm{d}r\,\mathrm{d}\varphi\,\mathrm{d}\theta$$

Multiple integrals

Vertical integration (p. 309):

$$\int_{\Omega} f = \int_a^b \left(\int_{g_1(x)}^{g_2(x)} f(x, y)\, \mathrm{d}y \right) \mathrm{d}x$$

Horizontal integration (p. 309):

$$\int_{\Omega} f = \int_c^d \left(\int_{h_1(y)}^{h_2(y)} f(x, y)\, \mathrm{d}x \right) \mathrm{d}y$$

Iterated integral (p. 324):

$$\int_{\Omega} f = \int_D \left(\int_{g_1(x,y)}^{g_2(x,y)} f(x, y, z)\, \mathrm{d}z \right) \mathrm{d}x\, \mathrm{d}y$$

Iterated integral (p. 325):

$$\int_{\Omega} f = \int_{\alpha}^{\beta} \left(\int_{A_z} f(x, y, z)\, \mathrm{d}x\, \mathrm{d}y \right) \mathrm{d}z$$

Variable change in multiple integrals (p. 328):

$$\int_{\Omega} f(\boldsymbol{x})\, \mathrm{d}\Omega = \int_{\Omega'} f\big(\boldsymbol{\Phi}(\boldsymbol{u})\big) |\det \boldsymbol{J\Phi}(\boldsymbol{u})|\, \mathrm{d}\Omega'$$

Pappus' Centroid Theorem (p. 333):

$$\mathrm{vol}(\Omega) = 2\pi y_G\, \mathrm{area}(T)$$

Integrals on curves and surfaces

Integral along a curve (p. 368):

$$\int_{\boldsymbol{\gamma}} f = \int_a^b f\big(\boldsymbol{\gamma}(t)\big)\,\|\boldsymbol{\gamma}'(t)\|\,\mathrm{d}t$$

Path integral (p. 375):

$$\int_{\boldsymbol{\gamma}} \boldsymbol{f}\cdot\boldsymbol{\tau} = \int_{\boldsymbol{\gamma}} f_\tau = \int_a^b \boldsymbol{f}\big(\boldsymbol{\gamma}(t)\big)\cdot\boldsymbol{\gamma}'(t)\,\mathrm{d}t$$

Integral on a surface (p. 378):

$$\int_{\boldsymbol{\sigma}} f = \int_{\mathcal{R}} f\big(\boldsymbol{\sigma}(u,v)\big)\|\boldsymbol{\nu}(u,v)\|\,\mathrm{d}u\,\mathrm{d}v$$

Flux integral (p. 384):

$$\int_{\boldsymbol{\sigma}} \boldsymbol{f}\cdot\boldsymbol{n} = \int_{\boldsymbol{\sigma}} f_n = \int_{\mathcal{R}} \boldsymbol{f}\big(\boldsymbol{\sigma}(u,v)\big)\cdot\boldsymbol{\nu}(u,v)\,\mathrm{d}u\,\mathrm{d}v$$

Divergence Theorem (p. 391):

$$\int_{\Omega} \operatorname{div}\boldsymbol{f} = \int_{\partial\Omega} \boldsymbol{f}\cdot\boldsymbol{n}$$

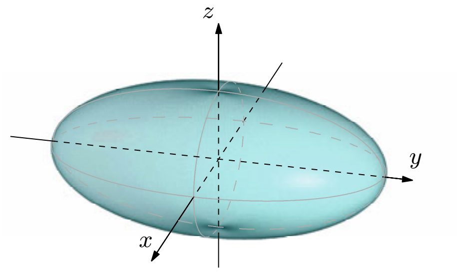
Green's Theorem (p. 394):

$$\int_{\Omega}\Big(\frac{\partial f_2}{\partial x} - \frac{\partial f_1}{\partial y}\Big)\,\mathrm{d}x\,\mathrm{d}y = \oint_{\partial\Omega} \boldsymbol{f}\cdot\boldsymbol{\tau}$$

Stokes' Theorem (p. 396):

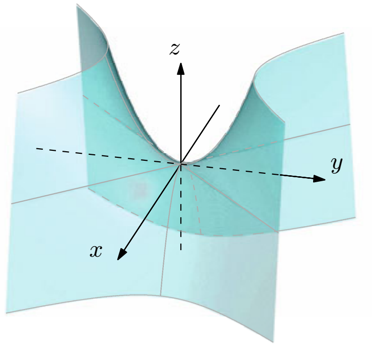
$$\int_{\Sigma} (\mathbf{curl}\,\boldsymbol{f})\cdot\boldsymbol{n} = \oint_{\partial\Sigma} \boldsymbol{f}\cdot\boldsymbol{\tau}$$

Examples of quadrics



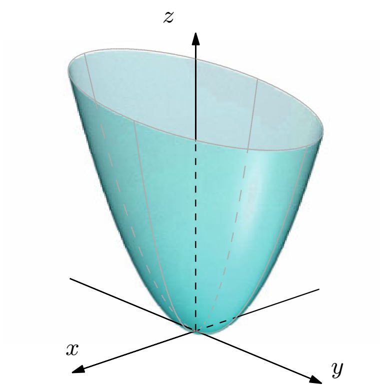
Ellipsoid

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}+\frac{z^2}{c^2}=1$$



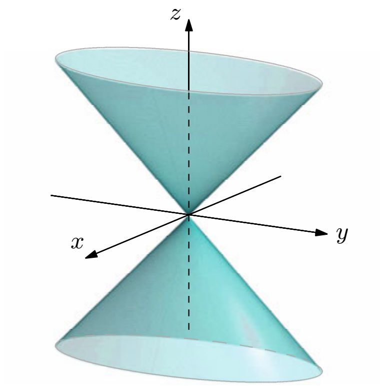
Hyperbolic paraboloid

$$\frac{z}{c}=-\frac{x^2}{a^2}+\frac{y^2}{b^2}$$



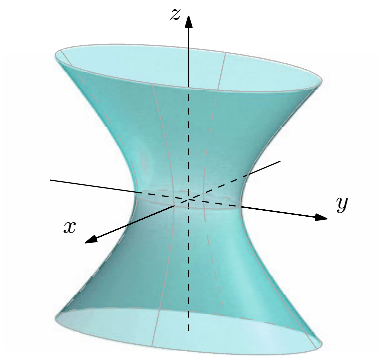
Elliptic paraboloid

$$\frac{z}{c}=\frac{x^2}{a^2}+\frac{y^2}{b^2}$$



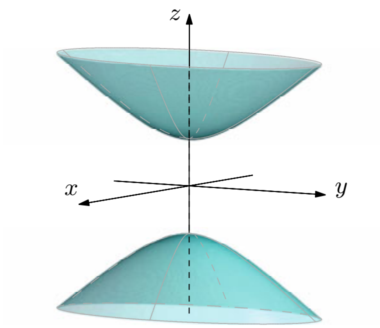
Cone

$$\frac{z^2}{c^2} = \frac{x^2}{a^2} + \frac{y^2}{b^2}$$



One-sheeted hyperboloid

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} - \frac{z^2}{c^2} = 1$$



Two-sheeted hyperboloid

$$\frac{z^2}{c^2} - \frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$$

Index

C. Canuto, A. Tabacco: *Mathematical Analysis II,* 2nd Ed.,
UNITEXT – La Matematica per il 3+2 85, DOI 10.1007/978-3-319-12757-6,

Collana Unitext – La Matematica per il 3+2

As of 2004, the books published in the series have been given a volume number. Titles in grey indicate editions out of print.
As of 2011, the series also publishes books in English.

A. Bernasconi, B. Codenotti
Introduzione alla complessità computazionale
1998, X+260 pp, ISBN 88-470-0020-3

A. Bernasconi, B. Codenotti, G. Resta
Metodi matematici in complessità computazionale
1999, X+364 pp, ISBN 88-470-0060-2

E. Salinelli, F. Tomarelli
Modelli dinamici discreti
2002, XII+354 pp, ISBN 88-470-0187-0

S. Bosch
Algebra
2003, VIII+380 pp, ISBN 88-470-0221-4

S. Graffi, M. Degli Esposti
Fisica matematica discreta
2003, X+248 pp, ISBN 88-470-0212-5

S. Margarita, E. Salinelli
MultiMath – Matematica Multimediale per l'Università
2004, XX+270 pp, ISBN 88-470-0228-1

A. Quarteroni, R. Sacco, F.Saleri
Matematica numerica (2a Ed.)
2000, XIV+448 pp, ISBN 88-470-0077-7
2002, 2004 ristampa riveduta e corretta
(1a edizione 1998, ISBN 88-470-0010-6)

13. A. Quarteroni, F. Saleri
Introduzione al Calcolo Scientifico (2a Ed.)
2004, X+262 pp, ISBN 88-470-0256-7
(1a edizione 2002, ISBN 88-470-0149-8)

14. S. Salsa
Equazioni a derivate parziali - Metodi, modelli e applicazioni
2004, XII+426 pp, ISBN 88-470-0259-1

15. G. Riccardi
Calcolo differenziale ed integrale
2004, XII+314 pp, ISBN 88-470-0285-0

16. M. Impedovo
Matematica generale con il calcolatore
2005, X+526 pp, ISBN 88-470-0258-3

17. L. Formaggia, F. Saleri, A. Veneziani
Applicazioni ed esercizi di modellistica numerica
per problemi differenziali
2005, VIII+396 pp, ISBN 88-470-0257-5

18. S. Salsa, G. Verzini
Equazioni a derivate parziali – Complementi ed esercizi
2005, VIII+406 pp, ISBN 88-470-0260-5
2007, ristampa con modifiche

19. C. Canuto, A. Tabacco
Analisi Matematica I (2a Ed.)
2005, XII+448 pp, ISBN 88-470-0337-7
(1a edizione, 2003, XII+376 pp, ISBN 88-470-0220-6)

20. F. Biagini, M. Campanino
Elementi di Probabilità e Statistica
2006, XII+236 pp, ISBN 88-470-0330-X

21. S. Leonesi, C. Toffalori
Numeri e Crittografia
2006, VIII+178 pp, ISBN 88-470-0331-8

22. A. Quarteroni, F. Saleri
Introduzione al Calcolo Scientifico (3a Ed.)
2006, X+306 pp, ISBN 88-470-0480-2

23. S. Leonesi, C. Toffalori
Un invito all'Algebra
2006, XVII+432 pp, ISBN 88-470-0313-X

24. W.M. Baldoni, C. Ciliberto, G.M. Piacentini Cattaneo
Aritmetica, Crittografia e Codici
2006, XVI+518 pp, ISBN 88-470-0455-1

25. A. Quarteroni
Modellistica numerica per problemi differenziali (3a Ed.)
2006, XIV+452 pp, ISBN 88-470-0493-4
(1a edizione 2000, ISBN 88-470-0108-0)
(2a edizione 2003, ISBN 88-470-0203-6)

26. M. Abate, F. Tovena
Curve e superfici
2006, XIV+394 pp, ISBN 88-470-0535-3

27. L. Giuzzi
Codici correttori
2006, XVI+402 pp, ISBN 88-470-0539-6

28. L. Robbiano
Algebra lineare
2007, XVI+210 pp, ISBN 88-470-0446-2

29. E. Rosazza Gianin, C. Sgarra
Esercizi di finanza matematica
2007, X+184 pp, ISBN 978-88-470-0610-2

30. A. Machì
Gruppi – Una introduzione a idee e metodi della Teoria dei Gruppi
2007, XII+350 pp, ISBN 978-88-470-0622-5
2010, ristampa con modifiche

31 Y. Biollay, A. Chaabouni, J. Stubbe
Matematica si parte!
A cura di A. Quarteroni
2007, XII+196 pp, ISBN 978-88-470-0675-1

32. M. Manetti
Topologia
2008, XII+298 pp, ISBN 978-88-470-0756-7

33. A. Pascucci
Calcolo stocastico per la finanza
2008, XVI+518 pp, ISBN 978-88-470-0600-3

34. A. Quarteroni, R. Sacco, F. Saleri
Matematica numerica (3a Ed.)
2008, XVI+510 pp, ISBN 978-88-470-0782-6

35. P. Cannarsa, T. D'Aprile
Introduzione alla teoria della misura e all'analisi funzionale
2008, XII+268 pp, ISBN 978-88-470-0701-7

36. A. Quarteroni, F. Saleri
Calcolo scientifico (4a Ed.)
2008, XIV+358 pp, ISBN 978-88-470-0837-3

37. C. Canuto, A. Tabacco
Analisi Matematica I (3a Ed.)
2008, XIV+452 pp, ISBN 978-88-470-0871-3

38. S. Gabelli
Teoria delle Equazioni e Teoria di Galois
2008, XVI+410 pp, ISBN 978-88-470-0618-8

39. A. Quarteroni
Modellistica numerica per problemi differenziali (4a Ed.)
2008, XVI+560 pp, ISBN 978-88-470-0841-0

40. C. Canuto, A. Tabacco
Analisi Matematica II
2008, XVI+536 pp, ISBN 978-88-470-0873-1
2010, ristampa con modifiche

41. E. Salinelli, F. Tomarelli
Modelli Dinamici Discreti (2a Ed.)
2009, XIV+382 pp, ISBN 978-88-470-1075-8

42. S. Salsa, F.M.G. Vegni, A. Zaretti, P. Zunino
Invito alle equazioni a derivate parziali
2009, XIV+440 pp, ISBN 978-88-470-1179-3

43. S. Dulli, S. Furini, E. Peron
Data mining
2009, XIV+178 pp, ISBN 978-88-470-1162-5

44. A. Pascucci, W.J. Runggaldier
Finanza Matematica
2009, X+264 pp, ISBN 978-88-470-1441-1

45. S. Salsa
Equazioni a derivate parziali – Metodi, modelli e applicazioni (2a Ed.)
2010, XVI+614 pp, ISBN 978-88-470-1645-3

46. C. D'Angelo, A. Quarteroni
Matematica Numerica – Esercizi, Laboratori e Progetti
2010, VIII+374 pp, ISBN 978-88-470-1639-2

47. V. Moretti
Teoria Spettrale e Meccanica Quantistica – Operatori in spazi di Hilbert
2010, XVI+704 pp, ISBN 978-88-470-1610-1

48. C. Parenti, A. Parmeggiani
Algebra lineare ed equazioni differenziali ordinarie
2010, VIII+208 pp, ISBN 978-88-470-1787-0

49. B. Korte, J. Vygen
Ottimizzazione Combinatoria. Teoria e Algoritmi
2010, XVI+662 pp, ISBN 978-88-470-1522-7

50. D. Mundici
Logica: Metodo Breve
2011, XII+126 pp, ISBN 978-88-470-1883-9

51. E. Fortuna, R. Frigerio, R. Pardini
Geometria proiettiva. Problemi risolti e richiami di teoria
2011, VIII+274 pp, ISBN 978-88-470-1746-7

52. C. Presilla
Elementi di Analisi Complessa. Funzioni di una variabile
2011, XII+324 pp, ISBN 978-88-470-1829-7

53. L. Grippo, M. Sciandrone
Metodi di ottimizzazione non vincolata
2011, XIV+614 pp, ISBN 978-88-470-1793-1

54. M. Abate, F. Tovena
Geometria Differenziale
2011, XIV+466 pp, ISBN 978-88-470-1919-5

55. M. Abate, F. Tovena
Curves and Surfaces
2011, XIV+390 pp, ISBN 978-88-470-1940-9

56. A. Ambrosetti
Appunti sulle equazioni differenziali ordinarie
2011, X+114 pp, ISBN 978-88-470-2393-2

57. L. Formaggia, F. Saleri, A. Veneziani
Solving Numerical PDEs: Problems, Applications, Exercises
2011, X+434 pp, ISBN 978-88-470-2411-3

58. A. Machì
Groups. An Introduction to Ideas and Methods of the Theory of Groups
2011, XIV+372 pp, ISBN 978-88-470-2420-5

59. A. Pascucci, W.J. Runggaldier
Financial Mathematics. Theory and Problems for Multi-period Models
2011, X+288 pp, ISBN 978-88-470-2537-0

60. D. Mundici
Logic: a Brief Course
2012, XII+124 pp, ISBN 978-88-470-2360-4

61. A. Machì
Algebra for Symbolic Computation
2012, VIII+174 pp, ISBN 978-88-470-2396-3

62. A. Quarteroni, F. Saleri, P. Gervasio
Calcolo Scientifico (5a ed.)
2012, XVIII+450 pp, ISBN 978-88-470-2744-2

63. A. Quarteroni
Modellistica Numerica per Problemi Differenziali (5a ed.)
2012, XVIII+628 pp, ISBN 978-88-470-2747-3

64. V. Moretti
Spectral Theory and Quantum Mechanics
With an Introduction to the Algebraic Formulation
2013, XVI+728 pp, ISBN 978-88-470-2834-0

65. S. Salsa, F.M.G. Vegni, A. Zaretti, P. Zunino
A Primer on PDEs. Models, Methods, Simulations
2013, XIV+482 pp, ISBN 978-88-470-2861-6

66. V.I. Arnold
Real Algebraic Geometry
2013, X+110 pp, ISBN 978-3-642–36242-2

67. F. Caravenna, P. Dai Pra
Probabilità. Un'introduzione attraverso modelli e applicazioni
2013, X+396 pp, ISBN 978-88-470-2594-3

68. A. de Luca, F. D'Alessandro
Teoria degli Automi Finiti
2013, XII+316 pp, ISBN 978-88-470-5473-8

69. P. Biscari, T. Ruggeri, G. Saccomandi, M. Vianello
Meccanica Razionale
2013, XII+352 pp, ISBN 978-88-470-5696-3

70. E. Rosazza Gianin, C. Sgarra
Mathematical Finance: Theory Review and Exercises. From Binomial Model to Risk Measures
2013, X+278pp, ISBN 978-3-319-01356-5

71. E. Salinelli, F. Tomarelli
Modelli Dinamici Discreti (3a Ed.)
2014, XVI+394pp, ISBN 978-88-470-5503-2

72. C. Presilla
Elementi di Analisi Complessa. Funzioni di una variabile (2a Ed.)
2014, XII+360pp, ISBN 978-88-470-5500-1

73. S. Ahmad, A. Ambrosetti
A Textbook on Ordinary Differential Equations
2014, XIV+324pp, ISBN 978-3-319-02128-7

74. A. Bermúdez, D. Gómez, P. Salgado
Mathematical Models and Numerical Simulation in Electromagnetism
2014, XVIII+430pp, ISBN 978-3-319-02948-1

75. A. Quarteroni
Matematica Numerica. Esercizi, Laboratori e Progetti (2a Ed.)
2013, XVIII+406pp, ISBN 978-88-470-5540-7

76. E. Salinelli, F. Tomarelli
Discrete Dynamical Models
2014, XVI+386pp, ISBN 978-3-319-02290-1

77. A. Quarteroni, R. Sacco, F. Saleri, P. Gervasio
Matematica Numerica (4a Ed.)
2014, XVIII+532pp, ISBN 978-88-470-5643-5

78. M. Manetti
Topologia (2a Ed.)
2014, XII+334pp, ISBN 978-88-470-5661-9

79. M. Iannelli, A. Pugliese
An Introduction to Mathematical Population Dynamics. Along the trail of Volterra and Lotka
2014, XIV+338pp, ISBN 978-3-319-03025-8

80. V. M. Abrusci, L. Tortora de Falco
Logica. Volume 1
2014, X+180pp, ISBN 978-88-470-5537-7

81. P. Biscari, T. Ruggeri, G. Saccomandi, M. Vianello
Meccanica Razionale (2a Ed.)
2014, XII+390pp, ISBN 978-88-470-5725-8

82. C. Canuto, A. Tabacco
Analisi Matematica I (4a Ed.)
2014, XIV+508pp, ISBN 978-88-470-5722-7

83. C. Canuto, A. Tabacco
Analisi Matematica II (2a Ed.)
2014, XII+576pp, ISBN 978-88-470-5728-9

84. C. Canuto, A. Tabacco
Mathematical Analysis I (2nd Ed.)
2015, XIV+484pp, ISBN 978-3-319-12771-2

85. C. Canuto, A. Tabacco
Mathematical Analysis II (2nd Ed.)
2015, XII+550pp, ISBN 978-3-319-12756-9

The online version of the books published in this series is available at SpringerLink.
For further information, please visit the following link:
http://www.springer.com/series/5418