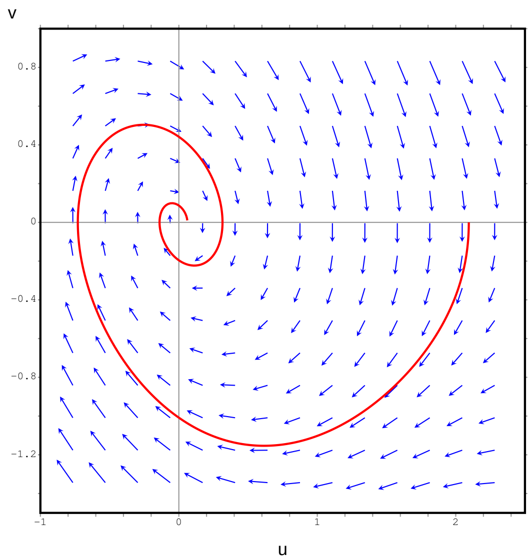
Mathematics 2 for Economics

Analysis and Dynamic Optimization

Josef Leydold



February 13, 2025

Institute for Statistics and Mathematics · WU Wien

Contents

Contents iii

Bibliography vi

1 Introduction 1
1.1 Overview 1

2 Sequences and Series 2
2.1 Limits of Sequences 2
2.2 Series 7
Exercises 11
Problems 11

3 Topology 13
3.1 Open Neighborhood 13
3.2 Convergence 15
3.3 Equivalent Norms 18
3.4 Compact Sets 20
3.5 Continuous Functions 21
Exercises 25
Problems 25

4 Derivatives 27
4.1 Roots of Univariate Functions 27
4.2 Limits of a Function 28
4.3 Derivatives of Univariate Functions 29
4.4 Higher Order Derivatives 31
4.5 The Mean Value Theorem 32
4.6 Gradient and Directional Derivatives 33
4.7 Higher Order Partial Derivatives 35
4.8 Derivatives of Multivariate Functions 36
Exercises 41
Problems 42

5 Taylor Series 46
5.1 Taylor Polynomial 46
5.2 Taylor Series 48
5.3 Landau Symbols 49
5.4 Power Series and Analytic Functions 51

5.5 Defining Functions . . . 52
5.6 Taylor's Formula for Multivariate Functions . . . 53
Exercises . . . 55
Problems . . . 55

6 Inverse and Implicit Functions **56**
6.1 Inverse Functions . . . 56
6.2 Implicit Functions . . . 58
Exercises . . . 61
Problems . . . 62

7 Convex Functions **63**
7.1 Convex Sets . . . 63
7.2 Convex and Concave Functions . . . 64
7.3 Monotone Univariate Functions . . . 68
7.4 Convexity of $\mathscr{C}^2$ Functions . . . 70
7.5 Quasi-Convex Functions . . . 73
Exercises . . . 77
Problems . . . 78

8 Static Optimization **79**
8.1 Extremal Points . . . 79
8.2 The Envelope Theorem . . . 81
8.3 Constraint Optimization – The Lagrange Function . . . 82
8.4 Kuhn-Tucker Conditions . . . 85
Exercises . . . 86
Problems . . . 87

9 Integration **88**
9.1 The Antiderivative of a Function . . . 88
9.2 The Riemann Integral . . . 91
9.3 The Fundamental Theorem of Calculus . . . 93
9.4 The Definite Integral . . . 95
9.5 Improper Integrals . . . 96
9.6 Differentiation under the Integral Sign . . . 97
Exercises . . . 100
Problems . . . 101

10 Multiple Integrals **102**
10.1 The Riemann Integral . . . 102
10.2 Double Integrals over Rectangles . . . 103
10.3 Double Integrals over General Domains . . . 106
10.4 A "Double Indefinite" Integral . . . 107
10.5 Change of Variables . . . 107
10.6 Improper Multiple Integrals . . . 110
Exercises . . . 112
Problems . . . 112

11 Differential Equations **113**
11.1 First-Order Ordinary Differential Equations . . . 113
11.2 A Simple Economic Model . . . 114
11.3 Separable Equations . . . 116
11.4 First-Order Linear Differential Equations . . . 117
11.5 Logistic Differential Equation . . . 122
11.6 Phase Diagrams and Stability . . . 123
11.7 Existence and Uniqueness . . . 125
Exercises . . . 127
Problems . . . 128

12 Second-Order Differential Equations **130**
12.1 Second-Order Differential Equations . . . 130
12.2 Second-Order Linear Differential Equations . . . 130
12.3 Constant Coefficients . . . 132
12.4 Stability for Linear Differential Equations . . . 135
Exercises . . . 137

13 Systems of Differential Equations **138**
13.1 Simultaneous Equations of Differential Equations . . . 138
13.2 Linear Systems with Constant Coefficients . . . 140
13.3 Equilibrium Points for Linear Systems . . . 142
13.4 Stability for Non-Linear Systems . . . 145
13.5 Phase Plane Analysis . . . 149
Exercises . . . 150

14 Control Theory **152**
14.1 The Control Problem . . . 152
14.2 The Standard Problem (T fixed) . . . 153
14.3 The Standard Problem (T variable) . . . 159
Exercises . . . 161

A Complex Numbers **162**
Exercises . . . 165
Problems . . . 165

Solutions **166**

Index **172**

Bibliography

The following books have been used to prepare this course. In particular the course closely follows parts of book [2].

[1] Knut Sydsæter and Peter Hammond. *Essential Mathematics for Economics Analysis*. Prentice Hall, 3rd edition, 2008.

[2] Knut Sydsæter, Peter Hammond, Atle Seierstad, and Arne Strøm. *Further Mathematics for Economics Analysis*. Prentice Hall, 2005.

[3] Alpha C. Chiang and Kevin Wainwright. *Fundamental Methods of Mathematical Economics*. McGraw-Hill, 4th edition, 2005.

1

Introduction

1.1 Overview

- Topology
- Analysis
- Integration
- Multiple integrals
- Ordinary differential equations (ODE) of first and second order
  - initial value problem
  - linear and logistic differential equation
- Autonomous differential equation
  - phase diagram
  - stability of solutions
- Systems of differential equations
  - stationary points (stable, unstable, saddle points)
  - characterization using eigenvalues
  - Saddle path solutions
- Control theory
  - Hamilton function
  - transversality condition

2

Sequences and Series

What happens when we proceed ad infinitum*?*

2.1 Limits of Sequences

Sequence. A **sequence** $(x_n)_{n=1}^{\infty}$ of real numbers is an ordered list of real numbers. Formally it can be defined as a *function* that maps the natural numbers into $\mathbb{R}$. Number x_n is called the nth term of the sequence. We write (x_n) for short to denote a squence if there is no risk if confusion. Sequences can also be seen as vectors of infinite length.

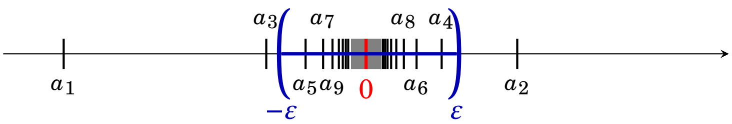
Definition 2.1

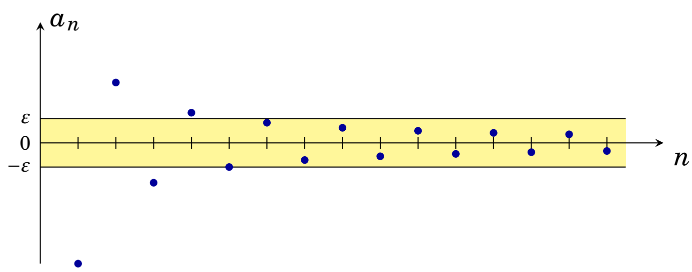
$x\colon \mathbb{N} \to \mathbb{R}, n \mapsto x_n$

Convergence and divergence. A sequence $(x_n)_{n=1}^{\infty}$ in $\mathbb{R}$ **converges** to a number x if for every $\varepsilon > 0$ there exists an index $N = N(\varepsilon)$ such that $|x_n - x| < \varepsilon$ for all $n \geq N$, or equivalently $x_n \in (x-\varepsilon, x+\varepsilon)$. The number x is then called the **limit** of the sequence. We write

Definition 2.2

$$x_n \to x \quad \text{as} \quad n \to \infty, \quad \text{or} \quad \lim_{n\to\infty} x_n = x\,.$$





A sequence that has a limit is called **convergent**. Otherwise it is called **divergent**.

Notice that the limit of a convergent sequence is uniquely determined, see Problem 2.5.

$$\lim_{n\to\infty} c = c \text{ for all } c \in \mathbb{R}$$

$$\lim_{n\to\infty} n^\alpha = \begin{cases} \infty, & \text{for } \alpha > 0, \\ 1, & \text{for } \alpha = 0, \\ 0, & \text{for } \alpha < 0. \end{cases}$$

$$\lim_{n\to\infty} q^n = \begin{cases} \infty, & \text{for } q > 1, \\ 1, & \text{for } q = 1, \\ 0, & \text{for } -1 < q < 1, \\ \nexists, & \text{for } q \leq -1. \end{cases}$$

$$\lim_{n\to\infty} \frac{n^a}{q^n} = \begin{cases} 0, & \text{for } |q| > 1, \\ \infty, & \text{for } 0 < q < 1, \\ \nexists, & \text{for } -1 < q < 0, \end{cases} \qquad \text{for } |q| \notin \{0,1\}.$$

$$\lim_{n\to\infty} \left(1+\tfrac{1}{n}\right)^n = e = 2.7182818\ldots$$

Table 2.5

Limits of important sequences

Example 2.3

The sequences

$$(a_n)_{n=1}^\infty = \left(\frac{1}{2^n}\right)_{n=1}^\infty = \left(\frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \ldots\right) \to 0$$

$$(b_n)_{n=1}^\infty = \left(\frac{n-1}{n+1}\right)_{n=1}^\infty = \left(0, \frac{1}{3}, \frac{2}{4}, \frac{3}{5}, \frac{4}{6}, \frac{5}{7}, \ldots\right) \to 1$$

converge as $n \to \infty$, i.e.,

$$\lim_{n\to\infty} \left(\frac{1}{2^n}\right) = 0 \quad \text{and} \quad \lim_{n\to\infty} \left(\frac{n-1}{n+1}\right) = 1\,. \qquad \diamondsuit$$

Example 2.4

The sequence

$$(c_n)_{n=1}^\infty = \left((-1)^n\right)_{n=1}^\infty = (-1, 1, -1, 1, -1, 1, \ldots)$$

$$(d_n)_{n=1}^\infty = \left(2^n\right)_{n=1}^\infty = (2, 4, 8, 16, 32, \ldots)$$

diverge. However, in the last example the sequence is increasing and not bounded from above. Thus we may write in *abuse of language*

$$\lim_{n\to\infty} 2^n = \infty\,. \qquad \diamondsuit$$

Computing limits can be a very challenging task. Thus we only look at a few examples. Table 2.5 lists limits of some important sequences. Notice that the limit of $\lim_{n\to\infty} \frac{n^a}{q^n}$ just says that in a product of a power sequence with an exponential sequence the latter dominates the limits.

We prove one of these limits in Lemma 2.12 below. For this purpose we need a few more notions.

Definition 2.6

Bounded sequence. A sequence $(x_n)_{n=1}^{\infty}$ of real numbers is called **bounded** if there exists an M such that

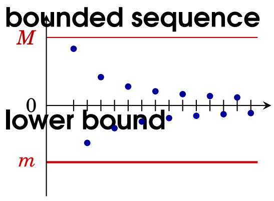
$$|x_n| \le M \quad \text{for all } n \in \mathbb{N}.$$

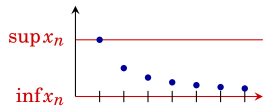
Two numbers m and M are called **lower** and **upper bound**, respectively, if

$$m \le x_n \le M, \quad \text{for all } n \in \mathbb{N}.$$

The greatest lower bound and the smallest upper bound are called **infimum** and **supremum** of the sequence, respectively, denoted by

$$\inf_{n\in\mathbb{N}} x_n \quad \text{and} \quad \sup_{n\in\mathbb{N}} x_n, \quad \text{respectively.}$$





Lemma 2.7

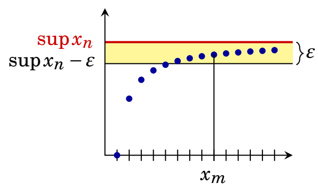
Notice that for a bounded sequence (x_n),

$$x_n \le \sup_{k\in\mathbb{N}} x_k \quad \text{for all } n \in \mathbb{N}$$

and for all $\varepsilon > 0$, there exists an $m \in \mathbb{N}$ such that

$$x_m > \left(\sup_{k\in\mathbb{N}} x_k\right) - \varepsilon$$

since otherwise $\left(\sup_{k\in\mathbb{N}} x_k\right) - \varepsilon$ were a smaller upper bound, a contradiction to the definition of the supremum.

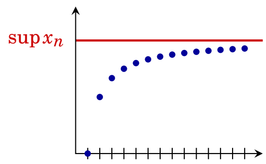


Example 2.8

Do not mix up supremum (or infimum) with the maximal (and minimal) value of a sequence. If a sequence (x_n) has a maximal value, then obviously $\max_{n\in\mathbb{N}} x_n = \sup_{n\in\mathbb{N}} x_n$. However, a maximal value need not exist. The sequence $\left(1 - \frac{1}{n}\right)_{n=1}^{\infty}$ is bounded and we have

$$\sup_{n\in\mathbb{N}} \left(1 - \frac{1}{n}\right) = 1\,.$$

However, 1 is never attained by this sequence and thus it does not have a maximum. ◊



Definition 2.9

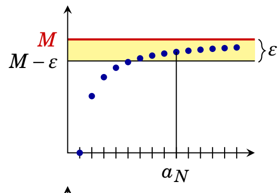
Monotone sequence. A sequence $(a_n)_{n=1}^{\infty}$ is called **monotone** if either $a_{n+1} \ge a_n$ (*increasing*) or $a_{n+1} \le a_n$ (*decreasing*) for all $n \in \mathbb{N}$.

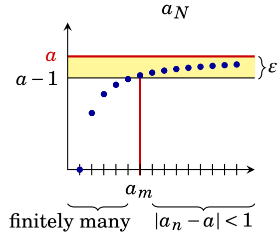
Lemma 2.10

Convergence of a monotone sequence. A monotone sequence $(a_n)_{n=1}^{\infty}$ is convergent if and only if it is bounded. We then find $\lim_{n\to\infty} a_n = \sup_{n\in\mathbb{N}} a_n$ if (a_n) is increasing, and $\lim_{n\to\infty} a_n = \inf_{n\in\mathbb{N}} a_n$ if (a_n) is decreasing.

PROOF IDEA. If (a_n) is increasing and bounded, then there is only a finite number of elements that are less than $\sup_{n\in\mathbb{N}} a_n - \varepsilon$.

If (a_n) is increasing and convergent, then there is only a finite number of elements greater than $\lim_{n\to\infty} a_n + \varepsilon$ or less than $\lim_{n\to\infty} a_n - \varepsilon$. These have a maximum and minimum value, respectively.





PROOF. We consider the case where (a_n) is increasing. Assume that (a_n) is bounded and $M = \sup_{n\in\mathbb{N}} a_n$. Then for every $\varepsilon > 0$, there exists an N such that $a_N > M - \varepsilon$ (Lemma 2.7). Since (a_n) is increasing we find $M \ge a_n > M - \varepsilon$ and thus $|a_n - M| < \varepsilon$ for all $n \ge N$. Consequently, $(a_n) \to M$ as $n \to \infty$.

Conversely, if (a_n) converges to a, then there is only a finite number of elements $a_1, \dots, a_m$ which do not satisfy $|a_n - a| < 1$. Thus $a_n < M = \max\{a+1, a_1, \dots, a_m\} < \infty$ for all $n \in \mathbb{N}$. Moreover, since (a_n) is increasing we also find $a_n \ge a_1$. Thus the sequence is bounded. The case where the sequence is decreasing follows completely analogously. □

Definition 2.11

For any $q \in \mathbb{R}$, the sequence $(q^n)_{n=0}^{\infty}$ is called a **geometric sequence**.

Lemma 2.12

Convergence of geometric sequence. $\lim_{n\to\infty} q^n = 0$ for all $q \in (-1, 1)$.

PROOF. Observe that for $0 \le q < 1$ we find $0 \le q^n = q \cdot q^{n-1} \le q^{n-1}$ for all $n \ge 2$ and hence q^n is decreasing and bounded from below. Hence it converges by Lemma 2.10 and $\lim_{n\to\infty} q^n = \inf_{n\ge 1} q^n$.

Now suppose that $m = \inf_{n\ge 1} q^n > 0$ for some $0 < q < 1$ and let $\varepsilon = m(1/q - 1) > 0$. By Lemma 2.7 there exists a k such that $q^k < m + \varepsilon$. Then $q^{k+1} = q \cdot q^k < q(m + m(1/q - 1)) = m$, a contradiction. Hence $\lim_{n\to\infty} q^n = 0$.

If $-1 < q < 0$, then $\lim_{n\to\infty} |q^n| = 0$ and hence $\lim_{n\to\infty} q^n = 0$ (Problem 2.6). □

Lemma 2.13

Divergence of geometric sequence. For $|q| > 1$ the geometric sequence diverges. Moreover, for $q > 1$ we find $\lim_{n\to\infty} q^n = \infty$.

PROOF. Suppose $M = \sup_{n\in\mathbb{N}} |q^n| < \infty$. Then $|(1/q)^n| \ge 1/M > 0$ for all $n \in \mathbb{N}$ and $M = \inf_{n\in\mathbb{N}} |(1/q)^n| > 0$, a contradiction to Lemma 2.12, as $|1/q| < 1$. □

Limits of sequences with more complex terms can be reduced to the limits listed in Table 2.5 by means of the rules listed in Theorem 2.14 below. Notice that Rule (1) implies that taking the limit of a sequence is a linear operator on the set of all convergent sequences.

Theorem 2.14

Rules for limits. Let $(a_n)_{n=1}^{\infty}$ and $(b_n)_{n=1}^{\infty}$ be convergent sequences in $\mathbb{R}$ and $(c_n)_{n=1}^{\infty}$ be a bounded sequence in $\mathbb{R}$. Then

(1) $\lim_{n\to\infty} (\alpha a_n + \beta b_n) = \alpha \lim_{n\to\infty} a_n + \beta \lim_{n\to\infty} b_n$ for all $\alpha, \beta \in \mathbb{R}$

(2) $\lim_{n\to\infty} (a_n \cdot b_n) = \lim_{n\to\infty} a_n \cdot \lim_{n\to\infty} b_n$

(3) $\lim_{n\to\infty} \frac{a_n}{b_n} = \frac{\lim_{n\to\infty} a_n}{\lim_{n\to\infty} b_n}$ (if $\lim_{n\to\infty} b_n \ne 0$)

(4) $\lim_{n\to\infty} a_n^r = \left(\lim_{n\to\infty} a_n\right)^r$

(5) $\lim_{n\to\infty} (a_n \cdot c_n) = 0$ (if $\lim_{n\to\infty} a_n = 0$)

For the proof of these (and other) properties of sequences the triangle inequality plays a prominent rôle.

Triangle inequality. For two real numbers a and b we find Lemma 2.15

$$|a+b| \le |a| + |b| .$$

PROOF. See Problem 2.4. □

Here we just prove Rule (1) from Theorem 2.14 (see also Problem 2.7). The other rules remain stated without proof.

Sum of covergent sequences. Let (a_n) and (b_n) be two sequences in $\mathbb{R}$ that converge to a and b, resp. Then Lemma 2.16

$$\lim_{n\to\infty} (a_n + b_n) = a + b .$$

PROOF IDEA. Use the triangle inequality for each term $(a_n + b_n) - (a + b)$.

PROOF. Let $\varepsilon > 0$ be arbitrary. Since both $(a_n) \to a$ and $(b_n) \to b$ there exists an $N = N(\varepsilon)$ such that $|a_n - a| < \varepsilon/2$ and $|b_n - b| < \varepsilon/2$ for all $n > N$. Then we find

$$\begin{aligned}|(a_n + b_n) - (a + b)| &= |(a_n - a) + (b_n - b)| \\ &\le |a_n - a| + |b_n - b| < \tfrac{\varepsilon}{2} + \tfrac{\varepsilon}{2} = \varepsilon\end{aligned}$$

for all $n > N$. But this means that $(a_n + b_n) \to (a + b)$, as claimed. □

The rules from Theorem 2.14 allow to reduce limits of composite terms to the limits listed in Table 2.5. Example 2.17

$$\lim_{n\to\infty}\left(2 + \frac{3}{n^2}\right) = 2 + 3 \underbrace{\lim_{n\to\infty} n^{-2}}_{=0} = 2 + 3 \cdot 0 = 2$$

$$\lim_{n\to\infty} (2^{-n} \cdot n^{-1}) = \lim_{n\to\infty} \frac{n^{-1}}{2^n} = 0$$

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

$$\lim_{n\to\infty} \underbrace{\sin(n)}_{\text{bounded}} \cdot \underbrace{\frac{1}{n^2}}_{\to 0} = 0$$ ◊

Exponential function. Theorem 2.14 allows to compute e^x as the limit of a sequence: Example 2.18

$$\begin{aligned}e^x &= \left(\lim_{m\to\infty}\left(1 + \frac{1}{m}\right)^m\right)^x = \lim_{m\to\infty}\left(1 + \frac{1}{m}\right)^{mx} = \lim_{n\to\infty}\left(1 + \frac{1}{n/x}\right)^n \\ &= \lim_{n\to\infty}\left(1 + \frac{x}{n}\right)^n\end{aligned}$$

where we have set $n = mx$. ◊

2.2 Series

Series. Let $(x_n)_{n=1}^{\infty}$ be a sequence of real numbers. Then the associated **series** is defined as the ordered formal sum — Definition 2.19

$$\sum_{n=1}^{\infty} x_n = x_1 + x_2 + x_3 + \dots .$$

The sequence of **partial sums** associated to series $\sum_{n=1}^{\infty} x_n$ is defined as — partial sum

$$S_n = \sum_{i=1}^{n} x_i \qquad \text{for } n \in \mathbb{N}.$$

The series **converges** to a limit S if sequence $(S_n)_{n=1}^{\infty}$ converges to S, i.e., — convergent

$$S = \sum_{i=1}^{\infty} x_i \quad \text{if and only if} \quad S = \lim_{n\to\infty} S_n = \lim_{n\to\infty} \sum_{i=1}^{n} x_i .$$

Otherwise, the series is called **divergent**.

We have already seen that a geometric sequence converges if $|q| < 1$, see Lemma 2.12. The same holds for the associated geometric series.

Geometric series. The *geometric series* converges if and only if $|q| < 1$ and we find — Lemma 2.20

$$\sum_{n=0}^{\infty} q^n = 1 + q + q^2 + q^3 + \dots = \frac{1}{1-q} .$$

PROOF IDEA. We first find a closed form for the terms of the geometric series and then compute the limit.

PROOF. We first show that for any $n \ge 0$,

$$S_n = \sum_{k=0}^{n} q^k = \frac{1-q^{n+1}}{1-q} .$$

In fact,

$$\begin{aligned} S_n(1-q) = S_n - qS_n &= \sum_{k=0}^{n} q^k - q\sum_{k=0}^{n} q^k = \sum_{k=0}^{n} q^k - \sum_{k=0}^{n} q^{k+1} \\ &= \sum_{k=0}^{n} q^k - \sum_{k=1}^{n+1} q^k = q^0 - q^{n+1} = 1 - q^{n+1} \end{aligned}$$

and thus the result follows. Now by the rules for limits of sequences we find by Lemma 2.12 $\lim_{n\to\infty} \sum_{k=0}^{n} q^n = \lim_{n\to\infty} \frac{1-q^{n+1}}{1-q} = \frac{1}{1-q}$ if $|q| < 1$. Conversely, if $|q| > 1$, the sequence diverges by Lemma 2.13. If $q = 1$, the we trivially have $\sum_{n=0}^{\infty} 1 = \infty$. For $q = -1$ the sequence of partial sums is given by $S_n = \sum_{k=0}^{n} (-1)^k = 1 + (-1)^n$ which obviously does not converge. This completes the proof. □

Harmonic series. The so called *harmonic series* diverges, Lemma 2.21

$$\sum_{n=1}^{\infty} \frac{1}{n} = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots = \infty .$$

PROOF IDEA. We construct a new series which is component-wise smaller than or equal to the harmonic series. This series is then transformed by adding some its terms into a series with constant terms which is obviously divergent.

PROOF. We find

$$\begin{aligned}
\sum_{n=1}^{\infty} \frac{1}{n} &= 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8} + \frac{1}{9} + \cdots + \frac{1}{16} + \frac{1}{17} + \dots \\
&> 1 + \frac{1}{2} + \frac{1}{4} + \frac{1}{4} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{16} + \cdots + \frac{1}{16} + \frac{1}{32} + \dots \\
&= 1 + \frac{1}{2} + \left(\frac{1}{4} + \frac{1}{4}\right) + \left(\frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8}\right) + \left(\frac{1}{16} + \cdots + \frac{1}{16}\right) + \left(\frac{1}{32} + \dots\right. \\
&= 1 + \frac{1}{2} + \left(\frac{1}{2}\right) + \left(\frac{1}{2}\right) + \left(\frac{1}{2}\right) + \left(\frac{1}{2}\right) + \dots \\
&= \infty .
\end{aligned}$$

More precisely, we have $\sum_{n=1}^{2^k} \frac{1}{n} > 1 + \frac{k}{2} \to \infty$ as $k \to \infty$. □

The trick from the above proof is called the **comparison test** as we compare our series with a divergent series. Analogously one also may compare the sequence with a convergent one.

Comparison test. Let $\sum_{n=1}^{\infty} a_n$ and $\sum_{n=1}^{\infty} b_n$ be two series with $0 \le a_n \le b_n$ for all $n \in \mathbb{N}$. Lemma 2.22

(a) If $\sum_{n=1}^{\infty} b_n$ converges, then $\sum_{n=1}^{\infty} a_n$ converges.

(b) If $\sum_{n=1}^{\infty} a_n$ diverges, then $\sum_{n=1}^{\infty} b_n$ diverges.

PROOF. (a) Suppose that $B = \sum_{k=1}^{\infty} b_k < \infty$ exists. Then by our assumptions $0 \le \sum_{k=1}^{n} a_k \le \sum_{k=1}^{n} b_k \le B$ for all $n \in \mathbb{N}$. Hence $\sum_{k=1}^{n} a_k$ is increasing and bounded and thus the series converges by Lemma 2.10.

(b) On the other hand, if $\sum_{k=1}^{\infty} a_k$ diverges, then for every M there exists an N such that $M \le \sum_{k=1}^{n} a_k \le \sum_{k=1}^{n} b_k$ for all $n \ge N$. Hence $\sum_{k=1}^{\infty} b_k$ diverges, too. □

Such tests are very important as it allows to verify whether a series converges or diverges by comparing it to a series where the answer is much simpler. However, it does not provide a limit when (b_n) converges (albeit it provides an upper bound for the limit). Nevertheless, the proof of existence is also of great importance. The following example demonstrates that using expressions in a naïve way without checking their existence may result in contradictions.

Example 2.23

Grandi's series. Consider the following series that has been extensively discussed during the 18th century. What is the value of

$$S = \sum_{n=0}^{\infty} (-1)^n = 1 - 1 + 1 - 1 + 1 - 1 + \ldots ?$$

One might argue in the following way:

$$\begin{aligned} 1 - S &= 1 - (1 - 1 + 1 - 1 + 1 - 1 + \ldots) = 1 - 1 + 1 - 1 + 1 - 1 + 1 - \ldots \\ &= 1 - 1 + 1 - 1 + 1 - \ldots = S \end{aligned}$$

and hence $2S = 1$ and $S = \frac{1}{2}$. Notice that this series is just a special case of the geometric series with $q = -1$. Thus we get the same result if we misleadingly use the formula from Lemma 2.20.

However, we also may proceed in a different way. By putting parentheses we obtain

$$\begin{aligned} S &= (1-1) + (1-1) + (1-1) + \ldots = 0 + 0 + 0 + \ldots = 0\,, \quad \text{and} \\ S &= 1 + (-1+1) + (-1+1) + (-1+1) + \ldots = 1 + 0 + 0 + 0 + \ldots = 1\,. \end{aligned}$$

Combining these three computations gives

$$S = \tfrac{1}{2} = 0 = 1$$

which obviously is not what we expect from real numbers. The error in all these computation is that the expression S cannot be treated like a number since the series diverges. ♢

If we are given a convergent sequence $(a_n)_{n=1}^{\infty}$ then the sequence of its absolute values also converges (Problem 2.6). The converse, however, may not hold. For the associated series we have an opposite result.

Lemma 2.24

Let $\sum_{n=1}^{\infty} a_n$ be some series. If $\sum_{n=1}^{\infty} |a_n|$ converges, then $\sum_{n=1}^{\infty} a_n$ also converges.

PROOF IDEA. We split the series into a positive and a negative part.

PROOF. Let $\mathscr{P} = \{n \in \mathbb{N} \colon a_n \geq 0\}$ and $\mathscr{N} = \{n \in \mathbb{N} \colon a_n < 0\}$. Then

$$m_+ = \sum_{n \in \mathscr{P}} |a_n| \leq \sum_{n=1}^{\infty} |a_n| < \infty \quad \text{and} \quad m_- = \sum_{n \in \mathscr{N}} |a_n| \leq \sum_{n=1}^{\infty} |a_n| < \infty$$

and therefore

$$\sum_{n=1}^{\infty} a_n = \sum_{n \in \mathscr{P}} |a_n| - \sum_{n \in \mathscr{N}} |a_n| = m_+ - m_-$$

exists. □

Notice that the converse does not hold. If series $\sum_{n=1}^{\infty} a_n$ converges then $\sum_{n=1}^{\infty} |a_n|$ may diverge.

Example 2.25

It can be shown that the **alternating harmonic series**

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \ldots = \ln 2$$

converges, whereas we already have seen in Lemma 2.21 that the harmonic series $\sum_{n=1}^{\infty} \left|\frac{(-1)^{n+1}}{n}\right| = \sum_{n=1}^{\infty} \frac{1}{n}$ does not not. ♢

Definition 2.26

A series $\sum_{n=1}^{\infty} a_n$ is called **absolutely convergent** if $\sum_{n=1}^{\infty} |a_n|$ converges.

Lemma 2.27

Ratio test. A series $\sum_{n=1}^{\infty} a_n$ converges if there exists a $q < 1$ and an $N < \infty$ such that

$$\left|\frac{a_{n+1}}{a_n}\right| \le q < 1 \quad \text{for all } n \ge N.$$

Similarly, if there exists an $r > 1$ and an $N < \infty$ such that

$$\left|\frac{a_{n+1}}{a_n}\right| \ge r > 1 \quad \text{for all } n \ge N$$

then the series diverges.

PROOF IDEA. We compare the series with a geometric series and apply the comparison test.

PROOF. For the first statement observe that $|a_{n+1}| < |a_n|\, q$ implies $|a_{N+k}| < |a_N| q^k$. Hence

$$\sum_{n=1}^{\infty} |a_n| = \sum_{n=1}^{N} |a_n| + \sum_{k=1}^{\infty} |a_{N+k}| < \sum_{n=1}^{N} |a_n| + |a_N| \sum_{k=1}^{\infty} q^k < \infty$$

where the two inequalities follows by Lemmata 2.22 and 2.20. Thus $\sum_{n=1}^{\infty} a_n$ converges by Lemma 2.24. The second statement follows similarly but requires more technical details and is thus omitted. □

There exist different variants of this test. We give a convenient version for a special case.

Lemma 2.28

Ratio test. Let $\sum_{n=1}^{\infty} a_n$ be a series where $\lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right|$ exists. Then $\sum_{n=1}^{\infty} a_n$ converges if

$$\lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| < 1 .$$

It diverges if

$$\lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| > 1 .$$

PROOF. Assume that $L = \lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right|$ exists and $L < 1$. Then there exists an N such that $\left|\frac{a_{n+1}}{a_n}\right| < q = 1 - \frac{1}{2}(1-L) < 1$ for all $n \ge N$. Thus the series converges by Lemma 2.27. The proof for the second statement is completely analogous. □

— Exercises

2.1 Compute the following limits:

(a) $\lim_{n\to\infty}\left(7+\left(\frac{1}{2}\right)^n\right)$ (b) $\lim_{n\to\infty}\frac{2n^3-6n^2+3n-1}{7n^3-16}$

(c) $\lim_{n\to\infty}\frac{n \bmod 10}{(-2)^n}$ (d) $\lim_{n\to\infty}\frac{n^2+1}{n+1}$

(e) $\lim_{n\to\infty}\left(n^2-(-1)^n n^3\right)$ (f) $\lim_{n\to\infty}\left(\frac{7n}{2n-1}-\frac{4n^2-1}{5-3n^2}\right)$

2.2 Compute the limits of sequence $(a_n)_{n=1}^\infty$ with the following terms:

(a) $a_n=(-1)^n\left(1+\frac{1}{n}\right)$ (b) $a_n=\frac{n}{(n+1)^2}$

(c) $a_n=\left(1+\frac{2}{n}\right)^n$ (d) $a_n=\left(1-\frac{2}{n}\right)^n$

(e) $a_n=\frac{1}{\sqrt{n}}$ (f) $a_n=\frac{n}{n+1}+\frac{1}{\sqrt{n}}$

(g) $a_n=\frac{n}{n+1}+\sqrt{n}$ (h) $a_n=\frac{4+\sqrt{n}}{n}$

2.3 Compute the following limits:

(a) $\lim_{n\to\infty}\left(1+\frac{1}{n}\right)^{nx}$ (b) $\lim_{n\to\infty}\left(1+\frac{x}{n}\right)^n$ (c) $\lim_{n\to\infty}\left(1+\frac{1}{nx}\right)^n$

— Problems

2.4 Prove the triangle inequality in Lemma 2.15.

HINT: Look at all possible cases where $a\ge 0$ or $a<0$ and $b\ge 0$ and $b<0$.

2.5 Let (a_n) be a convergent sequence. Show by means of the triangle inequality (Lemma 2.15) that its limit is uniquely defined.

HINT: Assume that two limits a and b exist and show that $|a-b|=0$.

2.6 Let (a_n) be a convergent sequence with $\lim_{n\to\infty} a_n = a$. Show that

$$\lim_{n\to\infty}|a_n|=|a|\,.$$

State and disprove the converse statement.

HINT: Use inequality $\big||a|-|b|\big|\le|a-b|$.

2.7 Let (a_n) be a sequence in $\mathbb{R}$ that converge to a and $c\in\mathbb{R}$. Show that

$$\lim_{n\to\infty} c\,a_n = c\,a\,.$$

2.8 Let (a_n) be a sequence in $\mathbb{R}$ that converge to 0 and (c_n) be a bounded sequence. Show that

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

2.9 Let (a_n) be a convergent sequence with $a_n \geq 0$. Show that

$$\lim_{n\to\infty} a_n \geq 0\,.$$

Disprove that $\lim_{n\to\infty} a_n > 0$ when all elements of this convergent sequence are positive, i.e., $a_n > 0$ for all $n \in \mathbb{N}$.

2.10 When we inspect the second part of the proof of Lemma 2.10 we find that monotonicity of sequence (a_n) is not required. Show that every convergent sequence (a_n) is bounded.

Also disprove the converse claim that every bounded sequence is convergent.

2.11 Compute $\sum_{k=1}^{\infty} q^n$.

2.12 Show that for any $a \in \mathbb{R}$,

$$\lim_{n\to\infty} \frac{a^n}{n!} = 0\,.$$

HINT: There exists an $N > |a|$.

2.13 **Cauchy's covergence criterion.** A sequence (a_n) in $\mathbb{R}$ is called a *Cauchy sequence* if for every $\varepsilon > 0$ there exists a number N such that $|a_n - a_m| < \varepsilon$ for all $n, m > N$.
Show: If a sequence (a_n) converges, then it is a Cauchy sequence.

HINT: Use the triangle inequality.

(Remark: The converse also holds. If (a_n) is a Cauchy sequence, then it converges.)

2.14 Show that $\sum_{n=1}^{\infty} \frac{1}{n!}$ converges.

HINT: Use the ratio test.

2.15 Someone wants to show the (false!) "theorem":

If $\sum_{n=1}^{\infty} a_n$ converges, then $\sum_{n=1}^{\infty} |a_n|$ also converges.

He argues as follows:

Let $\mathscr{P} = \{n \in \mathbb{N} \colon a_n \geq 0\}$ and $\mathscr{N} = \{n \in \mathbb{N} \colon a_n < 0\}$. Then

$$\sum_{n=1}^{\infty} a_n = \sum_{n\in\mathscr{P}} a_n + \sum_{n\in\mathscr{N}} a_n = \sum_{n\in\mathscr{P}} |a_n| - \sum_{n\in\mathscr{N}} |a_n| < \infty$$

and thus both $m_+ = \sum_{n\in\mathscr{P}} |a_n| < \infty$ and $m_- = \sum_{n\in\mathscr{N}} |a_n| < \infty$. Therefore

$$\sum_{n=1}^{\infty} |a_n| = \sum_{n\in\mathscr{P}} |a_n| + \sum_{n\in\mathscr{N}} |a_n| = m_+ + m_- < \infty$$

exists.

3
Topology

We need the concepts of neighborhood *and* boundary.

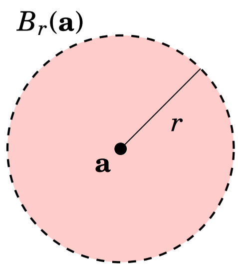
The fundamental idea in analysis can be visualized as *roaming in foggy weather*. We explore a function *locally* around some point by making tiny steps in all directions. However, we then need some conditions that ensure that we do not run against an edge or fall out of our function's world (i.e., its domain). Thus we introduce the concept of an *open neighborhood*.

3.1 Open Neighborhood

Definition 3.1

Interior, exterior and boundary points. Recall that for any point $\mathbf{x} \in \mathbb{R}^n$ the **Euclidean norm** $\|\mathbf{x}\|$ is defined as

$$\|\mathbf{x}\| = \sqrt{\mathbf{x}'\mathbf{x}} = \sqrt{\sum_{i=1}^{n} x_i^2} \, .$$

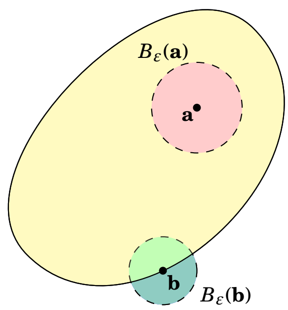


The **Euclidean distance** $d(\mathbf{x},\mathbf{y})$ between any two points $\mathbf{x},\mathbf{y} \in \mathbb{R}^n$ is given as

$$d(\mathbf{x},\mathbf{y}) = \|\mathbf{x}-\mathbf{y}\| = \sqrt{(\mathbf{x}-\mathbf{y})'(\mathbf{x}-\mathbf{y})} \, .$$

These terms allow us to get a notion of points that are "nearby" some point $\mathbf{x}$. The set

$$B_r(\mathbf{a}) = \left\{\mathbf{x} \in \mathbb{R}^n : d(\mathbf{x},\mathbf{a}) < r\right\}$$



is called the **open ball** around $\mathbf{a}$ with radius r (> 0). A point $\mathbf{a} \in D$ is called an **interior point** of a set $D \subseteq \mathbb{R}^n$ if there exists an open ball centered at $\mathbf{a}$ which lies inside D, i.e., there exists an $\varepsilon > 0$ such that $B_\varepsilon(\mathbf{a}) \subseteq D$. An immediate consequence of this definition is that we can move away from some interior point $\mathbf{a}$ in any direction without leaving D provided that the step size is sufficiently small. Notice that every set contains all its interior points.

A point $\mathbf{b} \in \mathbb{R}^n$ is called a **boundary point** of a set $D \subseteq \mathbb{R}^n$ if every open ball centered at $\mathbf{b}$ intersects both D and its complement $D^c = \mathbb{R}^n \setminus D$. Notice that a boundary point $\mathbf{b}$ needs not be an element of D.

A point $\mathbf{x} \in \mathbb{R}^n$ is called an **exterior point** of a set $D \subseteq \mathbb{R}^n$ if it is an interior point of its complement $\mathbb{R}^n \setminus D$.

Definition 3.2

A set $D \subseteq \mathbb{R}^n$ is called an **open neighborhood** of $\mathbf{a}$ if $\mathbf{a}$ is an interior point of D, i.e., if D contains some open ball centered at $\mathbf{a}$.

Definition 3.3

A set $D \subseteq \mathbb{R}^n$ is called **open** if all its members are interior points of D, i.e., if for each $\mathbf{a} \in D$, D contains some open ball centered at $\mathbf{a}$ (that is, have an open neighborhood in D). On the real line $\mathbb{R}$, the simplest example of an open set is an open interval $(a,b) = \{x \in \mathbb{R} \colon a < x < b\}$.

A set $D \subseteq \mathbb{R}^n$ is called **closed** if it contains all its boundary points. On the real line $\mathbb{R}$, the simplest example of a closed set is a closed interval $[a,b] = \{x \in \mathbb{R} \colon a \le x \le b\}$.

Example 3.4

Show that $H = \{(x,y) \in \mathbb{R}^2 \colon x > 0\}$ is an open set.

SOLUTION. Take any point (x_0, y_0) in H and set $\varepsilon = x_0/2$. We claim that $B = B_\varepsilon(x_0,y_0)$ is contained in H. Let $(x,y) \in B$. Then $\varepsilon > \|(x,y)-(x_0,y_0)\| = \sqrt{(x-x_0)^2+(y-y_0)^2} \ge \sqrt{(x-x_0)^2} = |x-x_0|$. Consequently, $x > x_0 - \varepsilon = x_0 - \frac{x_0}{2} = \frac{x_0}{2} > 0$ and thus $(x,y) \in H$ as claimed. □

Lemma 3.5

A set $D \subseteq \mathbb{R}^n$ is closed if and only if its complement D^c is open.

PROOF. See Problem 3.6.

Theorem 3.6

Properties of open sets.

(1) The empty set $\emptyset$ and the whole space $\mathbb{R}^n$ are both open.

(2) Arbitrary unions of open sets are open.

(3) The intersection of finitely many open sets is open.

PROOF IDEA. (1) Every ball of centered at any point is entirely in $\mathbb{R}^n$. Thus $\mathbb{R}^n$ is open. For the empty set observe that it does not contain any element that violates the condition for *"interior point"*.

(2) Every open ball $B_\varepsilon(\mathbf{x})$ remains contained in a set D if we add points to D. Thus interior points of D remain interior points in any superset of D.

(3) If $\mathbf{x}$ is an interior point of open sets $D_1, \dots, D_m$, then there exist open balls $B_i(\mathbf{x}) \subseteq D_i$ centered at $\mathbf{x}$. Since they are only finitely many, there is a smallest one which is thus entirely contained in the intersection of all D_i's.

PROOF. (1) Every ball $B_\varepsilon(\mathbf{a}) \subseteq \mathbb{R}^n$ and thus $\mathbb{R}^n$ is open. All members of the empty set $\emptyset$ are inside balls that are contained entirely in $\emptyset$. Hence $\emptyset$ is open.

(2) Let $\{D_i\}_{i \in I}$ be an arbitrary family of open sets in $\mathbb{R}^n$, and let $D = \bigcup_{i \in I} D_i$ be the union of all these. For each $\mathbf{x} \in D$ there is at least one $i \in I$ such that $\mathbf{x} \in D_i$. Since D_i is open, there exists an open ball $B_\varepsilon(\mathbf{x}) \subseteq D_i \subseteq D$. Hence $\mathbf{x}$ is an interior point of D.

(3) Let $\{D_1, D_2, \ldots, D_m\}$ be a finite collection of open sets in $\mathbb{R}^n$, and let $D = \bigcap_{i=1}^m D_i$ be the intersection of all these sets. Let $\mathbf{x}$ be any point in D. Since all D_i are open there exist open balls $B_i = B_{\varepsilon_i}(\mathbf{x}) \subseteq D_i$ with center $\mathbf{x}$. Let ε be the smallest of all radii ε_i. Then $\mathbf{x} \in B_\varepsilon(\mathbf{x}) = \bigcap_{i=1}^m B_i \subseteq \bigcap_{i=1}^m D_i = D$ and thus D is open. □

ε is the minimum of a finite set of numbers.

The intersection of an infinite number of open sets needs not be open, see Problem 3.10.

Similarly by De Morgan's law we find the following properties of closed sets, see Problem 3.11.

Properties of closed sets. Theorem 3.7

(1) The empty set $\emptyset$ and the whole space $\mathbb{R}^n$ are both closed.

(2) Arbitrary intersections of closed sets are closed.

(3) The union of finitely many closed sets is closed.

Each $\mathbf{y} \in \mathbb{R}^n$ is either an interior, an exterior or a boundary point of some set $D \subseteq \mathbb{R}^n$. As a consequence there is a corresponding partition of $\mathbb{R}^n$ into three mutually disjoint sets.

Definition 3.8

For a set $D \subseteq \mathbb{R}^n$, the set of all interior points of D is called the **interior** of D. It is denoted by D° or $\text{int}(D)$.

The set of all boundary points of a set D is called the **boundary** of D. It is denoted by ∂D or $\text{bd}(D)$.

The union $D \cup \partial D$ is called the **closure** of D. It is denoted by $\overline{D}$ or $\text{cl}(D)$.

Definition 3.9

A point $\mathbf{a}$ is called an **accumulation point** of a set D if every open neighborhood of $\mathbf{a}$ (i.e., open ball $B_\varepsilon(\mathbf{a})$) has non-empty intersection with D (i.e., $D \cap B_\varepsilon(\mathbf{a}) \neq \emptyset$). Notice that $\mathbf{a}$ need not be an element of D.

Lemma 3.10

A set D is closed if and only if D contains all its accumulation points.

PROOF. See Problem 3.12.

3.2 Convergence

Definition 3.11

A **sequence** $(\mathbf{x}_k)_{k=1}^\infty$ in $\mathbb{R}^n$ is a *function* that maps the natural numbers into $\mathbb{R}^n$. A point $\mathbf{x}_k$ is called the kth term of the sequence.
Sequences can also be seen as vectors of infinite length.

$\mathbf{x}\colon \mathbb{N} \to \mathbb{R}^n, k \mapsto \mathbf{x}_k$

Recall that a sequence (x_k) in $\mathbb{R}$ converges to a number x if for every $\varepsilon > 0$ there exists an index N such that $|x_k - x| < \varepsilon$ for all $k > N$. This can be easily generalized.

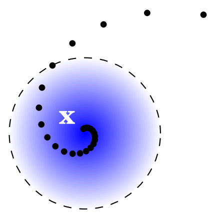
Definition 3.12

Convergence and divergence. A sequence $(\mathbf{x}_k)$ in $\mathbb{R}^n$ **converges** to a point $\mathbf{x}$ if for every $\varepsilon > 0$ there exists an index $N = N(\varepsilon)$ such that $\mathbf{x}_k \in B_\varepsilon(\mathbf{x})$, i.e., $\|\mathbf{x}_k - \mathbf{x}\| < \varepsilon$, for all $k > N$.

Equivalently, $(\mathbf{x}_k)$ converges to $\mathbf{x}$ if $d(\mathbf{x}_k, \mathbf{x}) \to 0$ as $k \to \infty$. The point $\mathbf{x}$ is then called the **limit** of the sequence. We write

$$\mathbf{x}_k \to \mathbf{x} \quad \text{as} \quad k \to \infty, \quad \text{or} \quad \lim_{k\to\infty} \mathbf{x}_k = \mathbf{x}\,.$$

Notice, that the limit of a convergent sequence is uniquely determined. A sequence that is not **convergent** is called **divergent**.



We can look at each of the component sequences in order to determine whether a sequence of points does converge or not. Thus the following theorem allows us to reduce results for convergent sequences in $\mathbb{R}^n$ to corresponding results for convergent sequences of real numbers.

Theorem 3.13

Convergence of each component. A sequence $(\mathbf{x}_k)$ in $\mathbb{R}^n$ converges to the vector $\mathbf{x}$ in $\mathbb{R}^n$ if and only if for each $j = 1, \dots, n$, the real number sequence $\left(x_k^{(j)}\right)_{k=1}^{\infty}$, consisting of the jth component of each vector $\mathbf{x}_k$, converges to $x^{(j)}$, the jth component of $\mathbf{x}$.

PROOF IDEA. For the proof of the necessity of the condition we use the fact that $\max_i |x_i| \leq \|\mathbf{x}\|$. For the sufficiency observe that $\|\mathbf{x}\|^2 \leq n \max_i |x_i|^2$.

PROOF. Assume that $\mathbf{x}_k \to \mathbf{x}$. Then for every $\varepsilon > 0$ there exists an N such that $\|\mathbf{x}_k - \mathbf{x}\| < \varepsilon$ for all $k > N$. Consequently, for each j one has $|x_k^{(j)} - x^{(j)}| \leq \|\mathbf{x}_k - \mathbf{x}\| < \varepsilon$ for all $k > N$, that is, $x_k^{(j)} \to x^{(j)}$.

Now assume that $x_k^{(j)} \to x^{(j)}$ for each j. Then given any $\varepsilon > 0$, for each j there exists a number N_j such that $|x_k^{(j)} - x^{(j)}| \leq \varepsilon/\sqrt{n}$ for all $k > N_j$. It follows that

$$\|\mathbf{x}_k - \mathbf{x}\| = \sqrt{\sum_{i=1}^{n} |x_k^{(i)} - x^{(i)}|^2} < \sqrt{\sum_{i=1}^{n} \varepsilon^2/n} = \sqrt{\varepsilon^2} = \varepsilon$$

for all $k > \max\{N_1, \dots, N_n\}$. Therefore $\mathbf{x}_k \to \mathbf{x}$ as $k \to \infty$. □

We will see in Section 3.3 below that this theorem is just a consequence of the fact that Euclidean norm and supremum norm are equivalent.

The next theorem gives a criterion for convergent sequences. The proof of the necessary condition demonstrates a simple but quite powerful technique.

Definition 3.14

A sequence $(\mathbf{x}_k)$ in $\mathbb{R}^n$ is called a **Cauchy sequence** if for every $\varepsilon > 0$ there exists a number N such that $\|\mathbf{x}_k - \mathbf{x}_m\| < \varepsilon$ for all $k, m > N$.

Theorem 3.15

Cauchy's covergence criterion. A sequence $(\mathbf{x}_k)$ in $\mathbb{R}^n$ is convergent if and only if it is a Cauchy sequence.

PROOF IDEA. For the necessity of the Cauchy sequence we use the trivial equality $\|\mathbf{x}_k - \mathbf{x}_m\| = \|(\mathbf{x}_k - \mathbf{x}) + (\mathbf{x} - \mathbf{x}_m)\|$ and apply the triangle inequality for norms.

For the sufficiency assume that $\|\mathbf{x}_k - \mathbf{x}_m\| \le \frac{1}{j}$ for all $m > k \ge N_j$ and construct closed balls $\overline{B}_{1/j}(\mathbf{x}_{N_j})$ for all $j \in \mathbb{N}$. Their intersection $\bigcap_{j=1}^{\infty} \overline{B}_{1/j}(\mathbf{x}_{N_j})$ is closed by Theorem 3.7 and is either a single point or the empty set. The latter can be excluded by an axiom of the real numbers.

PROOF. Assume that $(\mathbf{x}_k)$ converges to $\mathbf{x}$. Then there exists a number N such that $\|\mathbf{x}_k - \mathbf{x}\| < \varepsilon/2$ for all $k > N$. Hence by the triangle inequality we find

$$\|\mathbf{x}_k - \mathbf{x}_m\| = \|(\mathbf{x}_k - \mathbf{x}) + (\mathbf{x} - \mathbf{x}_m)\| \le \|\mathbf{x}_k - \mathbf{x}\| + \|\mathbf{x} - \mathbf{x}_m\| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

for all $k, m > N$. Thus $(\mathbf{x}_k)$ is a Cauchy sequence.

For the converse assume that for all $\varepsilon = 1/j$ there exists an N_j such that $\|\mathbf{x}_k - \mathbf{x}_m\| \le \frac{1}{j}$ for all $m > k \ge N_j$, i.e., $\mathbf{x}_m \in \overline{B}_{1/j}(\mathbf{x}_{N_j})$ for all $m > N_j$. Let $D_j = \bigcap_{i=1}^{j} \overline{B}_{1/i}(\mathbf{x}_{N_i})$. Then $\mathbf{x}_m \in D_j$ for all $m > N_j$ and thus $D_j \ne \emptyset$ for all $j \in \mathbb{N}$. Moreover, the diameter of $D_j \le 2/j \to 0$ for $j \to \infty$. By Theorem 3.7, $D = \bigcap_{i=1}^{\infty} \overline{B}_{1/i}(\mathbf{x}_{N_i})$ is closed. Therefore, either $D = \{\mathbf{a}\}$ consists of a single point or $D = \emptyset$. The latter can be excluded by a fundamental property (i.e., an axiom) of the real numbers. (However, this step is out of the scope of this course.) □

The next theorem is another example of an application of the triangle inequality.

Theorem 3.16

Sum of covergent sequences. Let $(\mathbf{x}_k)$ and $(\mathbf{y}_k)$ be two sequences in $\mathbb{R}^n$ that converge to $\mathbf{x}$ and $\mathbf{y}$, resp. Then

$$\lim_{k\to\infty} (\mathbf{x}_k + \mathbf{y}_k) = \mathbf{x} + \mathbf{y}\,.$$

PROOF IDEA. Use the triangle inequality for each term $\|(\mathbf{x}_k + \mathbf{y}_k) - (\mathbf{x} + \mathbf{y})\| = \|(\mathbf{x}_k - \mathbf{x}) + (\mathbf{y}_k - \mathbf{y})\|$.

PROOF. Let $\varepsilon > 0$ be arbitrary. Since $(\mathbf{x}_k)$ is convergent, there exists a number N_x such that $\|\mathbf{x}_k - \mathbf{x}\| < \varepsilon/2$ for all $k > N_x$. Analogously there exists a number N_y such that $\|\mathbf{y}_k - \mathbf{y}\| < \varepsilon/2$ for all $k > N_y$. Let N be the greater of the two numbers N_x and N_y. Then by the triangle inequality we find for $k > N$,

$$\begin{aligned}\|(\mathbf{x}_k+\mathbf{y}_k)-(\mathbf{x}+\mathbf{y})\| &= \|(\mathbf{x}_k-\mathbf{x})+(\mathbf{y}_k-\mathbf{y})\|\\ &\le \|\mathbf{x}_k-\mathbf{x}\|+\|\mathbf{y}_k-\mathbf{y}\| < \frac{\varepsilon}{2}+\frac{\varepsilon}{2}=\varepsilon\,.\end{aligned}$$

But this means that $(\mathbf{x}_k+\mathbf{y}_k)\to(\mathbf{x}+\mathbf{y})$, as claimed. □

We can use convergent sequences to characterize closed sets.

Theorem 3.17

Closure and convergence. A set $D \subseteq \mathbb{R}^n$ is closed if and only if every convergent sequence of points in D has its limit in D.

PROOF IDEA. For any sequence in D with limit $\mathbf{x}$ every ball $B_\varepsilon(\mathbf{x})$ contains almost all elements of the sequence. Hence it belongs to the closure of D. So if D is closed then $\mathbf{x} \in D$.

Conversely, if $\mathbf{x} \in \mathrm{cl}(D)$ we can select points $\mathbf{x}_k \in B_{1/k}(\mathbf{x}) \cap D$. Then sequence $(\mathbf{x}_k) \to \mathbf{x}$ converges. If we assume that every convergent sequence of points in D has its limit in D it follows that $\mathbf{x} \in D$ and hence D is closed.

PROOF. Assume that D is closed. Let $(\mathbf{x}_k)$ be a convergent sequence with limit $\mathbf{x}$ such that $\mathbf{x}_k \in D$ for all k. Hence for all $\varepsilon > 0$ there exists an N such that $\mathbf{x}_k \in B_\varepsilon(\mathbf{x})$ for all $k > N$. Therefore $B_\varepsilon(\mathbf{x}) \cap D \neq \emptyset$ and $\mathbf{x}$ belongs to the closure of D. Since D is closed, limit $\mathbf{x}$ also belongs to D.

Conversely, assume that every convergent sequence of points in D has its limit in D. Let $\mathbf{x} \in \mathrm{cl}(D)$. Then $B_{1/k}(\mathbf{x}) \cap D \neq \emptyset$ for every $k \in \mathbb{N}$ and we can choose an $\mathbf{x}_k$ in $B_{1/k}(\mathbf{x}) \cap D$. Then $\mathbf{x}_k \to \mathbf{x}$ as $k \to \infty$ by construction. Thus $\mathbf{x} \in D$ by hypothesis. This shows $\mathrm{cl}(D) \subseteq D$, hence D is closed. □

There is also a smaller brother of the limit of a sequence.

Definition 3.18

A point $\mathbf{a}$ is called an **accumulation point** of a sequence $(\mathbf{x}_k)$ if every open ball $B_\varepsilon(\mathbf{a})$ contains infinitely many elements of the sequence.

Example 3.19

The sequence $\left((-1)^k\right)_{k=1}^{\infty} = (-1, 1, -1, 1 \dots)$ has accumulation points -1 and 1 but neither point is a limit of the sequence. ♢

3.3 Equivalent Norms

Our definition of open sets and convergent sequences is based on the Euclidean norm (or metric) in $\mathbb{R}^n$. However, we have already seen that the concept of *norm* and *metric* can be generalized. Different norms might result in different families of open sets.

Definition 3.20

Two norms $\|\cdot\|$ and $\|\cdot\|'$ are called (topologically) **equivalent** if every open set w.r.t. $\|\cdot\|$ is also an open set w.r.t. $\|\cdot\|'$ and vice versa.

Thus every interior point w.r.t. $\|\cdot\|$ is also an interior point w.r.t. $\|\cdot\|$' and vice versa. That is, there must exist two strictly positive constants c and d such that

$$c\,\|\mathbf{x}\| \leq \|\mathbf{x}\|' \leq d\,\|\mathbf{x}\|$$

for all $\mathbf{x} \in \mathbb{R}^n$.

An immediate consequence is that every sequence that is convergent w.r.t. some norm is also convergent in every equivalent norm.

Theorem 3.21

Euclidean norm $\|\cdot\|_2$, 1-norm $\|\cdot\|_1$, and supremum norm $\|\cdot\|_\infty$ are equivalent in $\mathbb{R}^n$.

PROOF. By a straightforward computation we find

$$\|\mathbf{x}\|_\infty = \max_{i=1,\dots,n} |x_i| \leq \sum_{i=1}^n |x_i| = \|\mathbf{x}\|_1 \leq \sum_{i=1}^n \left(\max_{j=1,\dots,n} |x_j|\right) = n\,\|\mathbf{x}\|_\infty$$

$$\|\mathbf{x}\|_\infty = \max_{i=1,\dots,n} |x_i| = \sqrt{\max_{i=1,\dots,n} |x_i|^2} \leq \sqrt{\sum_{i=1}^n |x_i|^2} = \|\mathbf{x}\|_2$$

$$\|\mathbf{x}\|_2 = \sqrt{\sum_{i=1}^n |x_i|^2} \leq \sqrt{\sum_{i=1}^n \left(\max_{j=1,\dots,n} |x_j|\right)^2} = \sqrt{n}\,\|\mathbf{x}\|_\infty$$

Equivalence of Euclidean norm and 1-norm can be derived from Minkowski's inequality. Using $\mathbf{x} = \sum_{i=1}^n x_i \mathbf{e}_i$ we find

$$\|\mathbf{x}\|_2 = \left\|\sum_{i=1}^n x_i \mathbf{e}_i\right\|_2 \leq \sum_{i=1}^n \|x_i \mathbf{e}_i\|_2 = \sum_{i=1}^n \sqrt{|x_i|^2} = \|\mathbf{x}\|_1 \,.$$

Conversly we find

$$\begin{aligned}
\|\mathbf{x}\|_1 &= \left(\sum_{i=1}^n |x_i|\right)^2 = \sum_{i=1}^n \sum_{j=1}^n |x_i||x_j| \\
&= n \sum_{i=1}^n |x_i|^2 - \left(n \sum_{i=1}^n |x_i|^2 - \sum_{i=1}^n \sum_{j=1}^n |x_i||x_j|\right) \\
&= n \sum_{i=1}^n |x_i|^2 - \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n (|x_i| - |x_j|)^2 \\
&\leq n \sum_{i=1}^n |x_i|^2 = n\,\|\mathbf{x}\|_2^2
\end{aligned}$$

and thus

$$\|\mathbf{x}\|_1 \leq \sqrt{n}\,\|\mathbf{x}\|_2$$

□

Notice that the equivalence of Euclidean norm and supremum norm immediately implies Theorem 3.13.

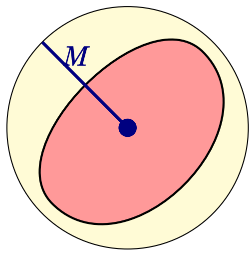
Theorem 3.21 is a corollary of a much stronger result for norms in $\mathbb{R}^n$ which we state without proof.

Theorem 3.22

Finitely generated vector space. All norms in a *finitely generated* vector space are equivalent.

For vector spaces which are not finitely generated this theorem does not hold any more. For example, in probability theory there are different concepts of convergence for sequences of random variates, e.g., convergence in distribution, in probability, almost surely. The corresponding norms or metrices are not equivalent. E.g., a sequence that converges in distribution need not converge almost surely.

3.4 Compact Sets



Bounded set. A set D in $\mathbb{R}^n$ is called **bounded** if there exists a number M such that $\|\mathbf{x}\| \leq M$ for all $x \in D$. A set that is not bounded is called **unbounded**.

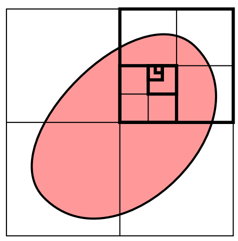
Obviously every convergent sequence is bounded (see Problem 3.15). However, the converse is not true. A sequence in a bounded set need not be convergent. But it always contains an accumulation point and a convergent subsequence.

Definition 3.23

Subsequence. Let $(\mathbf{x}_k)_{k=1}^{\infty}$ be a sequence in $\mathbb{R}^n$. Consider a strictly increasing sequence $k_1 < k_2 < k_3 < k_4 < \dots$ of natural numbers, and let $\mathbf{y}_j = \mathbf{x}_{k_j}$, for $j \in \mathbb{N}$. Then the sequence $(\mathbf{y}_j)_{j=1}^{\infty}$ is called a **subsequence** of $(\mathbf{x}_k)$. It is often denoted by $(\mathbf{x}_{k_j})_{j=1}^{\infty}$.

Example 3.24

Let $(x_k)_{k=1}^{\infty} = \left((-1)^k \frac{1}{k}\right)_{k=1}^{\infty} = \left(-1, \frac{1}{2}, -\frac{1}{3}, \frac{1}{4}, -\frac{1}{5}, \frac{1}{6}, -\frac{1}{7}, \dots\right)$. Then $(y_k)_{k=1}^{\infty} = \left(\frac{1}{2k}\right)_{k=1}^{\infty} = \left(\frac{1}{2}, \frac{1}{4}, \frac{1}{6}, \frac{1}{8}, \dots\right)$ and $(z_k)_{k=1}^{\infty} = \left(-\frac{1}{2k-1}\right)_{k=1}^{\infty} = \left(-1, -\frac{1}{3}, -\frac{1}{5}, -\frac{1}{7}, \dots\right)$ are two subsequences of (x_k). ◊



Now let $(\mathbf{x}_k)$ be a sequence in a bounded subset $D \subseteq \mathbb{R}^2$. Since D is bounded there exists a bounding square $K_0 \supseteq D$ of edge length L. Divide K_0 into four equal squares, each of which has sides of length $L/2$. At least one of these squares, say K_1, must contain infinitely many elements $\mathbf{x}_k$ of this sequence. Pick one of these, say $\mathbf{x}_{k_1}$. Next divide K_1 into four squares of edge length $L/4$. Again in at least one of them, say K_2, there will still be an infinite number of terms from sequence $(\mathbf{x}_k)$. Take one of these, $\mathbf{x}_{k_2}$, with $k_2 > k_1$.

Repeating this procedure ad infinitum we eventually obtain a subsequence $(\mathbf{x}_{k_j})$ of the original sequence that converges by Cauchy's criterion. It is quite obvious that this approach also works in any $\mathbb{R}^n$ where n may not equal to 2. Then we start with a bounding n-cube which is recursively divided into 2^n subcubes.

We summarize our observations in the following theorem (without giving a stringent formal proof).

Bolzano-Weierstrass. A subset D of $\mathbb{R}^n$ is bounded if and only if every sequence of points in D has a convergent subsequence. Theorem 3.25

A subset D of $\mathbb{R}^n$ is bounded if and only if every sequence has an accumulation point. Corollary 3.26

We now have seen that convergent sequences can be used to characterize *closed sets* (Theorem 3.17) and *bounded sets* (Theorem 3.25).

Compact set. A set D in $\mathbb{R}^n$ is called **compact** if it is closed and bounded. Definition 3.27

Compactness is a central concept in mathematical analysis, see, e.g., Theorems 3.36 and 3.37 below. When we combine the results of Theorems 3.17 and 3.25 we get the following characterization.

Bolzano-Weierstrass. A subset D of $\mathbb{R}^n$ is compact if and only if every sequence of points in D has a subsequence that converges to a point in D. Theorem 3.28

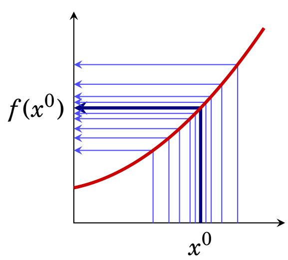
3.5 Continuous Functions

Recall that a univariate function $f:\mathbb{R}\to\mathbb{R}$ is called continuous if (roughly spoken) small changes in the argument cause small changes of the function value. One of the formal definitions reads: f is continuous at a point $x^0\in\mathbb{R}$ if $f(x_k)\to f(x^0)$ for every sequence (x_k) of points that converge to x^0. By our concept of open neighborhood this can easily be generalized for vector-valued functions.

Continuous functions. A function $\mathbf{f}=(f_1,\dots,f_m)\colon D\subseteq\mathbb{R}^n\to\mathbb{R}^m$ is said to be **continuous** at a point $\mathbf{x}^0$ if $\mathbf{f}(\mathbf{x}_k)\to\mathbf{f}(\mathbf{x}^0)$ for every sequence $(\mathbf{x}_k)$ of points in D that converges to $\mathbf{x}^0$. We then have Definition 3.29

$$\lim_{k\to\infty}\mathbf{f}(\mathbf{x}_k)=\mathbf{f}(\lim_{k\to\infty}\mathbf{x}_k)\,.$$

If $\mathbf{f}$ is continuous at every point $\mathbf{x}^0\in D$, we say that $\mathbf{f}$ is continuous on D.



The easiest way to show that a vector-valued function is continuous, is by looking at each of its components. We get the following result by means of Theorem 3.13.

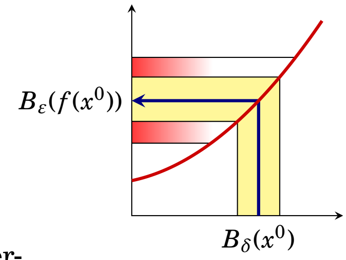
Continuity of each component. A function $\mathbf{f}=(f_1,\dots,f_m)\colon D\subseteq\mathbb{R}^n\to\mathbb{R}^m$ is continuous at a point $\mathbf{x}^0$ if and only if each component function $f_j\colon D\subseteq\mathbb{R}^n\to\mathbb{R}$ is continuous at $\mathbf{x}^0$. Theorem 3.30

There exist equivalent characterizations of continuity which are also used for alternative definitions of continuous functions in the literature. The first one uses open balls.

Theorem 3.31

Continuity and images of balls. A function $\mathbf{f}\colon D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ is continuous at a point $\mathbf{x}^0$ in D if and only if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{x}^0)\| < \varepsilon \quad \text{for all } \mathbf{x} \in D \text{ with } \quad \|\mathbf{x} - \mathbf{x}^0\| < \delta$$



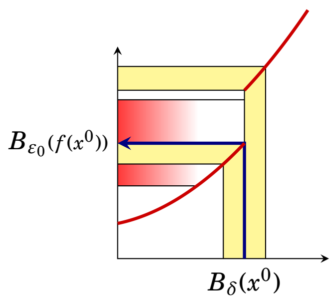
or equivalently,

$$\mathbf{f}\big(B_\delta(\mathbf{x}^0) \cap D\big) \subseteq B_\varepsilon\big(\mathbf{f}(\mathbf{x}^0)\big) .$$

PROOF IDEA. Assume that the condition holds and let $(\mathbf{x}_k)$ be a convergent sequence with limit $\mathbf{x}^0$. Then for every $\varepsilon > 0$ we can find an N such that $\|\mathbf{f}(\mathbf{x}_k) - \mathbf{f}(\mathbf{x}^0)\| < \varepsilon$ for all $k > N$, i.e., $\mathbf{f}(\mathbf{x}_k) \to \mathbf{f}(\mathbf{x}^0)$, which means that $\mathbf{f}$ is continuous at $\mathbf{x}^0$.

Now suppose that there exists an $\varepsilon_0 > 0$ where the condition is violated. Then there exists an $\mathbf{x}_k \in B_\delta(\mathbf{x}^0)$ with $\mathbf{f}(\mathbf{x}_k) \in \mathbf{f}\big(B_\delta(\mathbf{x}^0)\big) \setminus B_{\varepsilon_0}\big(\mathbf{f}(\mathbf{x}^0)\big)$ for every $\delta = \frac{1}{k}$, $k \in \mathbb{N}$. By construction $\mathbf{x}_k \to \mathbf{x}^0$ but $\mathbf{f}(\mathbf{x}_k) \not\to \mathbf{f}(\mathbf{x}^0)$. Thus $\mathbf{f}$ is not continuous at $\mathbf{x}^0$.

PROOF. Suppose that the condition holds. Let $\varepsilon > 0$ be given. Then there exists a $\delta > 0$ such that $\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{x}^0)\| < \varepsilon$ whenever $\|\mathbf{x} - \mathbf{x}^0\| < \delta$. Now let $(\mathbf{x}_k)$ be a sequence in D that converges to $\mathbf{x}^0$. Thus for every $\delta > 0$ there exists a number N such that $\|\mathbf{x}_k - \mathbf{x}^0\| < \delta$ for all $k > N$. But then $\|\mathbf{f}(\mathbf{x}_k) - \mathbf{f}(\mathbf{x}^0)\| < \varepsilon$ for all $k > N$, and consequently $\mathbf{f}(\mathbf{x}_k) \to \mathbf{f}(\mathbf{x}^0)$ for $k \to \infty$, which implies that $\mathbf{f}$ is continuous at $\mathbf{x}^0$.



Conversely, assume that $\mathbf{f}$ is continuous at $\mathbf{x}^0$ but the condition does not hold, that is, there exists an $\varepsilon_0 > 0$ such that for all $\delta = 1/k$, $k \in \mathbb{N}$, there is an $\mathbf{x} \in D$ with $\|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{x}^0)\| \geq \varepsilon_0$ albeit $\|\mathbf{x} - \mathbf{x}^0\| < 1/k$. Now pick a point $\mathbf{x}_k$ in D with this property for all $k \in \mathbb{N}$. Then sequence $(\mathbf{x}_k)$ converges to $\mathbf{x}^0$ by construction but $\mathbf{f}(\mathbf{x}_k) \notin B_{\varepsilon_0}\big(\mathbf{f}(\mathbf{x}^0)\big)$. This means, however, that $(\mathbf{f}(\mathbf{x}_k))$ does not converge to $\mathbf{f}(\mathbf{x}^0)$, a contradiction to our assumption that $\mathbf{f}$ is continuous. □

Continuous functions $\mathbf{f}\colon \mathbb{R}^n \to \mathbb{R}^m$ can also be characterized by their preimages. While the image $\mathbf{f}(D)$ of some open set $D \subseteq \mathbb{R}^n$ need not necessarily be an open set (see Problem 3.18) this always holds for the **preimage** of some open set $U \subseteq \mathbb{R}^m$,

$$\mathbf{f}^{-1}(U) = \{\mathbf{x} \colon \mathbf{f}(\mathbf{x}) \in U\} .$$

For the statement of the general result where the domain of $\mathbf{f}$ is not necessarily open we need the notion of relative open sets.

Definition 3.32

Let D be a subset in $\mathbb{R}^n$. Then

(a) A is **relatively open** in D if $A = U \cap D$ for some open set U in $\mathbb{R}^n$.

(b) A is **relatively closed** in D if $A = F \cap D$ for some closed set F in $\mathbb{R}^n$.

Obviously every open subset of an open set $D \subseteq \mathbb{R}^n$ is relatively open. The usefulness of the concept can be demonstrated by the following example.

Example 3.33

Let $D = [0,1] \subseteq \mathbb{R}$ be the domain of some function f. Then $A = (1/2, 1]$ obviously is not an open set in $\mathbb{R}$. However, A is relatively open in D as $A = (1/2, \infty) \cap [0,1] = (1/2, \infty) \cap D$. ♢

Theorem 3.34

Characterization of continuity. A function $\mathbf{f}\colon D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ is continuous if and only if either of the following equivalent conditions is satisfied:

(a) $\mathbf{f}^{-1}(U)$ is relatively open for each open set U in $\mathbb{R}^m$.

(b) $\mathbf{f}^{-1}(F)$ is relatively closed for each closed set F in $\mathbb{R}^m$.

PROOF IDEA. If $U \subseteq \mathbb{R}^m$ is open, then for all $\mathbf{x} \in \mathbf{f}^{-1}(U)$ there exists an $\varepsilon > 0$ such that $B_\varepsilon(\mathbf{f}(\mathbf{x})) \subseteq U$. If in addition $\mathbf{f}$ is continuous, then $B_\delta(\mathbf{x}) \subseteq \mathbf{f}^{-1}(B_\varepsilon(\mathbf{f}(\mathbf{x}))) \subseteq \mathbf{f}^{-1}(U)$ by Theorem 3.31 and hence $\mathbf{f}^{-1}(U)$ is open.

Conversely, if $\mathbf{f}^{-1}(B_\varepsilon(\mathbf{f}(\mathbf{x})))$ is open for all $\mathbf{x} \in D$ and all $\varepsilon > 0$, then there exists a $\delta > 0$ such that $B_\delta(\mathbf{x}) \subseteq \mathbf{f}^{-1}(B_\varepsilon(\mathbf{f}(\mathbf{x})))$ and thus $\mathbf{f}(B_\delta(\mathbf{x})) \subseteq B_\varepsilon(\mathbf{f}(\mathbf{x}))$, i.e., $\mathbf{f}$ is continuous at $\mathbf{x}$ by Theorem 3.31.

PROOF. For simplicity we only prove the case where $D = \mathbb{R}^n$.

(a) Suppose $\mathbf{f}$ is continuous and U is an open set in $\mathbb{R}^m$. Let $\mathbf{x}$ be any point in $\mathbf{f}^{-1}(U)$. Then $\mathbf{f}(\mathbf{x}) \in U$. As U is open there exists an $\varepsilon > 0$ such that $B_\varepsilon(\mathbf{f}(\mathbf{x})) \subseteq U$. By Theorem 3.31 there exists a $\delta > 0$ such that $\mathbf{f}(B_\delta(\mathbf{x})) \subseteq B_\varepsilon(\mathbf{f}(\mathbf{x})) \subseteq U$. Thus $B_\delta(\mathbf{x})$ belongs to the preimage of U. Therefore $\mathbf{x}$ is an interior point of $\mathbf{f}^{-1}(U)$ which means that $\mathbf{f}^{-1}(U)$ is an open set.

Conversely, assume that $\mathbf{f}^{-1}(U)$ is open for each open set $U \subseteq \mathbb{R}^m$. Let $\mathbf{x}$ be any point in D. Let $\varepsilon > 0$ be arbitrary. Then $U = B_\varepsilon(\mathbf{f}(\mathbf{x}))$ is an open set and by hypothesis the preimage $\mathbf{f}^{-1}(U)$ is open in D. Thus there exists a $\delta > 0$ such that $B_\delta(\mathbf{x}) \subseteq \mathbf{f}^{-1}(U) = \mathbf{f}^{-1}(B_\varepsilon(\mathbf{f}(\mathbf{x})))$ and hence $\mathbf{f}(B_\delta(\mathbf{x})) \subseteq U = B_\varepsilon(\mathbf{f}(\mathbf{x}))$. Consequently, $\mathbf{f}$ is continuous at $\mathbf{x}$ by Theorem 3.31. This completes the proof.

(b) This follows immediately from (a) and Lemma 3.5. □

Example 3.35

Let $U(\mathbf{x}) = U(x_1, \ldots, x_n)$ be a household's real-valued utility function, where $\mathbf{x}$ denotes its commodity vector and U is defined on the whole of $\mathbb{R}^n$. Then for a number a the upper level set $\Gamma_a = \{\mathbf{x} \in \mathbb{R}^n : U(\mathbf{x}) \ge a\}$ consists of all vectors where the household values are at least as much as a. Let F be the closed interval $[a, \infty)$. Then

$$\Gamma_a = \{\mathbf{x} \in \mathbb{R}^n : U(\mathbf{x}) \ge a\} = \{\mathbf{x} \in \mathbb{R}^n : U(\mathbf{x}) \in F\} = U^{-1}(F)\,.$$

According to Theorem 3.34, if U is continuous, then Γ_a is closed for each value of a. Hence, *continuous functions generate close upper level sets*. They also generate closed lower level sets. ♢

Let $\mathbf{f}$ be a continuous function. As already noted the image $\mathbf{f}(D)$ of some open set needs not be open. Similarly neither the image of a closed set is necessarily closed, nor needs the image of a bounded set be bounded (see Problem 3.18). However, there is a remarkable exception.

Continuous functions preserve compactness. Let $\mathbf{f}\colon D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ be continuous. Then the image $\mathbf{f}(K) = \{\mathbf{f}(\mathbf{x})\colon \mathbf{x} \in K\}$ of every compact subset K of D is compact. Theorem 3.36

PROOF IDEA. Take any sequence $(\mathbf{y}_k)$ in $\mathbf{f}(K)$ and a sequence $(\mathbf{x}_k)$ of its preimages in K, i.e., $\mathbf{y}_k = \mathbf{f}(\mathbf{x}_k)$. We now apply the Bolzano-Weierstrass Theorem twice: $(\mathbf{x}_k)$ has a subsequence $(\mathbf{x}_{k_j})$ that converges to some point $\mathbf{x}^0 \in K$. By continuity $\mathbf{y}_{k_j} = \mathbf{f}(\mathbf{x}_{k_j})$ converges to $\mathbf{f}(\mathbf{x}^0) \in \mathbf{f}(K)$. Hence $\mathbf{f}(K)$ is compact by the Bolzano-Weierstrass Theorem.

PROOF. Let $(\mathbf{y}_k)$ be any sequence in $\mathbf{f}(K)$. By definition, for each k there is a point $\mathbf{x}_k \in K$ such that $\mathbf{y}_k = \mathbf{f}(\mathbf{x}_k)$. By Theorem 3.28 (Bolzano-Weierstrass Theorem), there exists a subsequence $(\mathbf{x}_{k_j})$ that converges to a point $\mathbf{x}^0 \in K$. Because $\mathbf{f}$ is continuous, $\mathbf{f}(\mathbf{x}_{k_j}) \to \mathbf{f}(\mathbf{x}^0)$ as $j \to \infty$ where $\mathbf{f}(\mathbf{x}^0) \in \mathbf{f}(K)$. But then $(\mathbf{y}_{k_j})$ is a subsequence of $(\mathbf{y}_k)$ that converges to $\mathbf{f}(\mathbf{x}^0) \in \mathbf{f}(K)$. Thus $\mathbf{f}(K)$ is compact by Theorem 3.28, as claimed. □

We close this section with an important result in optimization theory.

Extreme-value theorem. Let $\mathbf{f}\colon K \subseteq \mathbb{R}^n \to \mathbb{R}$ be a continuous function on a compact set K. Then $\mathbf{f}$ has both a maximum point and a minimum point in K. Theorem 3.37

PROOF IDEA. By Theorem 3.36, $\mathbf{f}(K)$ is compact. Thus $\mathbf{f}(K)$ is bounded and closed, that is, $\mathbf{f}(K) = [a,b]$ for $a, b \in \mathbb{R}$ and $\mathbf{f}$ attains its minimum and maximum in respective points $\mathbf{x}_m, \mathbf{x}_M \in K$.

PROOF. By Theorem 3.36, $f(K)$ is compact. In particular, $f(K)$ is bounded, and so $-\infty < a = \inf_{\mathbf{x}\in K} f(\mathbf{x})$ and $b = \sup_{\mathbf{x}\in K} f(\mathbf{x}) < \infty$. Clearly a and b are boundary points of $f(K)$ which belong to $f(K)$, as $f(K)$ is closed. Hence there must exist points $\mathbf{x}_m$ and $\mathbf{x}_M$ such that $f(\mathbf{x}_m) = a$ and $f(\mathbf{x}_M) = b$. Obviously $\mathbf{x}_m$ and $\mathbf{x}_M$ are minimum point and a maximum point of K, respectively. □

— Exercises

3.1 Is $Q = \{(x, y) \in \mathbb{R}^2 : x > 0, y \geq 0\}$ open, closed, or neither?

3.2 Is $H = \{(x, y) \in \mathbb{R}^2 : x > 0, y \geq 1/x\}$ open, closed, or neither?

HINT: Sketch set H.

3.3 Let $F = \{(1/k, 0) \in \mathbb{R}^2 : k \in \mathbb{N}\}$. Is F open, closed, or neither?

HINT: Is $(0, 0) \in F$?

— Problems

3.4 Show that the open ball $D = B_r(\mathbf{a})$ is an open set.

HINT: Take any point $\mathbf{x} \in B_r(\mathbf{a})$ and an open ball $B_\varepsilon(\mathbf{x})$ of sufficiently small radius ε. (How small is "sufficiently small"?) Show that $B_\varepsilon(\mathbf{x}) \subseteq D$ by means of the triangle inequality.

3.5 Give respective examples for non-empty sets $D \subseteq \mathbb{R}^2$ which are

(a) neither open nor closed, or
(b) both open and closed, or
(c) closed and have empty interior, or
(d) not closed and have empty interior.

3.6 Show that a set $D \subseteq \mathbb{R}^n$ is closed if and only if its complement $D^c = \mathbb{R}^n \setminus D$ is open (Lemma 3.5).

HINT: Look at boundary points of D.

3.7 Show that a set $D \subseteq \mathbb{R}^n$ is open if and only if its complement D^c is closed.

HINT: Use Lemma 3.5.

3.8 Show that closure $\text{cl}(D)$ and boundary ∂D are closed for any $D \subseteq \mathbb{R}^n$.

HINT: Suppose that there is a boundary point of ∂D that is not a boundary point of D.

3.9 Let D and F be subsets of $\mathbb{R}^n$ such that $D \subseteq F$. Show that

$$\text{int}(D) \subseteq \text{int}(F) \quad \text{and} \quad \text{cl}(D) \subseteq \text{cl}(F) .$$

3.10 Recall the proof of Theorem 3.6.

(a) Where exactly do you need the assumption that there is an intersection of *finitely* many open sets in statement (3)?
(b) Let D be the intersection of the infinite family $B_{1/k}(\mathbf{0})$, $k = 1, 2, \ldots$, of open balls centered at $\mathbf{0}$. Is D open or closed?

HINT: Is there any point in D other than $\mathbf{0}$?

3.11 Prove Theorem 3.7.

HINT: Use Theorem 3.6 and De Morgan's law.

3.12 Prove Theorem 3.10.

3.13 Show that the limit of a convergent sequence is uniquely defined.

HINT: Suppose that two limits exist.

3.14 Show that for any points $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and every $j = 1, \ldots, n$,

$$|x_j - y_j| \leq \|\mathbf{x} - \mathbf{y}\|_2 \leq \sqrt{n} \max_{i=1,\ldots,n} |x_i - y_i| .$$

3.15 Show that every convergent sequence $(\mathbf{x}_k)$ in $\mathbb{R}^n$ is bounded.

3.16 Give an example for a bounded sequence that is not convergent.

3.17 For fixed $\mathbf{a} \in \mathbb{R}^n$, show that the function $f: \mathbb{R}^n \to \mathbb{R}$ defined by $f(\mathbf{x}) = \|\mathbf{x} - \mathbf{a}\|$ is continuous.

HINT: Use Theorem 3.31 and inequality $\big|\, \|\mathbf{x}\| - \|\mathbf{y}\| \,\big| \le \|\mathbf{x} - \mathbf{y}\|$.

3.18 Give examples of non-empty subsets D of $\mathbb{R}$ and continuous functions $f: \mathbb{R} \to \mathbb{R}$ such that

(a) D is closed, but $f(D)$ is not closed.
(b) D is open, but $f(D)$ is not open.
(c) D is bounded, but $f(D)$ is not bounded.

3.19 Prove that the set

$$D = \{\mathbf{x} \in \mathbb{R}^n : g_j(\mathbf{x}) \le 0,\ j = 1, \ldots, m\}$$

is closed if the functions g_j are all continuous.

4

Derivatives

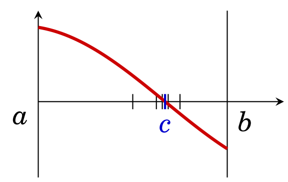
We want to have the best linear approximation *of a function.*

Derivatives are an extremely powerful tool for investigating properties of functions. For univariate functions it allows to check for monotonicity or concavity, or to find candidates for extremal points and verify its optimality. Therefore we want to generalize this tool for multivariate functions.

4.1 Roots of Univariate Functions

The following theorem seems to be trivial. However, it is of great importance as is assures the existence of a root of a continuous function.

Intermediate value theorem (Bolzano). Let $f\colon [a,b] \subseteq \mathbb{R} \to \mathbb{R}$ be a continuous function and assume that $f(a) > 0$ and $f(b) < 0$. Then there exists a point $c \in (a,b)$ such that $f(c) = 0$. Theorem 4.1



PROOF IDEA. We use a technique called *interval bisectioning*: Start with interval $[a_0, b_0] = [a,b]$, split the interval at $c_1 = (a_1 + b_1)/2$ and continue with the subinterval where f changes sign. By iterating this procedure we obtain a sequence of intervals $[a_n, b_n]$ of lengths $(b-a)/2^n \to 0$. By Cauchy's convergence criterion sequence (c_n) converges to some point c with $0 \leq \lim_{n\to\infty} f(c_n) \leq 0$. As f is continuous, we find $f(c) = \lim_{n\to\infty} f(c_n) = 0$.

PROOF. We construct a sequence of intervals $[a_n, b_n]$ by a method called **interval bisectioning**. Let $[a_0, b_0] = [a,b]$. Define $c_n = \frac{a_n+b_n}{2}$ and

$$[a_{n+1}, b_{n+1}] = \begin{cases} [c_n, b_n] & \text{if } f(c_n) \geq 0, \\ [a_n, c_n] & \text{if } f(c_n) < 0, \end{cases} \qquad \text{for } n = 1,2,\dots$$

Notice that $|a_k - a_n| < 2^{-N}(b-a)$ and $|b_k - b_n| < 2^{-N}(b-a)$ for all $k, n \geq N$. Hence (a_i) and (b_i) are Cauchy sequences and thus converge to respective points c_+ and c_- in $[a,b]$ by Cauchy's convergence criterion. Moreover, for every $\varepsilon > 0$, $|c_+ - c_-| \leq |a_k - c_+| + |b_k - c_-| < \varepsilon$ for sufficiently large

k and thus $c_+ = c_- = c$. By construction $f(a_k) \geq 0$ and $f(b_k) \leq 0$ for all k. By assumption f is continuous and thus $f(c) = \lim_{k\to\infty} f(a_k) \geq 0$ and $f(c) = \lim_{k\to\infty} f(b_k) \leq 0$, i.e., $f(c) = 0$ as claimed. $\square$

Interval bisectioning is a brute force method for finding a root of some function f. It is sometimes used as a last resort. Notice, however, that this is a rather slow method. *Newton's method*, *secant method* or *regula falsi* are much faster algorithms.

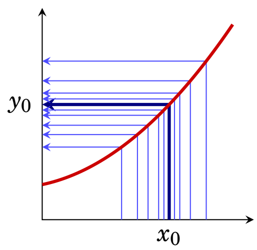
4.2 Limits of a Function

For the definition of derivative we need the concept of *limit* of a function.

Definition 4.2

Limit. Let $\mathbf{f}\colon D \subseteq \mathbb{R} \to \mathbb{R}$ be some function. Then the **limit** of f as x approaches x_0 is y_0 if for every convergent sequence of arguments $x_k \to x_0$, the sequences of images converges to y_0, i.e., $f(x_k) \to y_0$ as $k \to \infty$. We write

$$\lim_{x\to x_0} f(x) = y_0, \quad \text{or} \quad f(x) \to y_0 \quad \text{as} \quad x \to x_0 \,.$$



Notice that x_0 need not be an element of domain D and (in abuse of language) may also be ∞ or $-\infty$.

Thus results for limits of sequences (Theorem 2.14) translates immediately into results on limits of functions.

Theorem 4.3

Rules for limits. Let $f: \mathbb{R} \to \mathbb{R}$ and $g: \mathbb{R} \to \mathbb{R}$ be two functions where both $\lim\limits_{x\to x_0} f(x)$ and $\lim\limits_{x\to x_0} g(x)$ exist. Then

(1) $\lim\limits_{x\to x_0} \big(\alpha f(x) + \beta g(x)\big) = \alpha \lim\limits_{x\to x_0} f(x) + \beta \lim\limits_{x\to x_0} g(x) \qquad$ for all $\alpha, \beta \in \mathbb{R}$

(2) $\lim\limits_{x\to x_0} \big(f(x) \cdot g(x)\big) = \lim\limits_{x\to x_0} f(x) \cdot \lim\limits_{x\to x_0} g(x)$

(3) $\lim\limits_{x\to x_0} \dfrac{f(x)}{g(x)} = \dfrac{\lim_{x\to x_0} f(x)}{\lim_{x\to x_0} g(x)} \qquad$ (if $\lim_{x\to x_0} g(x) \neq 0$)

(4) $\lim\limits_{x\to x_0} \big(f(x)\big)^\alpha = \big(\lim\limits_{x\to x_0} f(x)\big)^\alpha \qquad$ (for $\alpha \in \mathbb{R}$, if $\big(\lim_{x\to x_0} f(x)\big)^\alpha$ is defined)

The notion of *limit* can be easily generalized for arbitrary transformations.

Definition 4.4

Let $\mathbf{f}\colon D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ be some function. Then the **limit** of $\mathbf{f}$ as $\mathbf{x}$ approaches $\mathbf{x}_0$ is $\mathbf{y}_0$ if for every convergent sequence of arguments $\mathbf{x}_k \to \mathbf{x}_0$, the sequences of images converges to $\mathbf{y}_0$, i.e., $\mathbf{f}(\mathbf{x}_0) \to \mathbf{y}_0$. We write

$$\lim_{\mathbf{x}\to \mathbf{x}_0} \mathbf{f}(\mathbf{x}) = \mathbf{y}_0, \quad \text{or} \quad \mathbf{f}(\mathbf{x}) \to \mathbf{y}_0 \quad \text{as} \quad \mathbf{x} \to \mathbf{x}_0 \,.$$

The point $\mathbf{x}_0$ need not be an element of domain D.

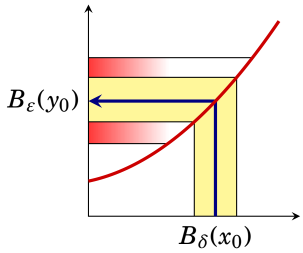
Writing $\lim_{\mathbf{x}\to\mathbf{x}_0} \mathbf{f}(\mathbf{x}) = \mathbf{y}_0$ means that we can make $\mathbf{f}(\mathbf{x})$ as close to $\mathbf{y}_0$ as we want when we put $\mathbf{x}$ sufficiently close to $\mathbf{x}_0$. Notice that a limit at some point $\mathbf{x}_0$ may not exist.

Similarly to our results in Section 3.5 we get the following equivalent characterization of the limit of a function. It is often used as an alternative definition of the term *limit*.

Theorem 4.5

Let $\mathbf{f}\colon D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ be a function. Then $\lim_{\mathbf{x}\to\mathbf{x}_0} \mathbf{f}(\mathbf{x}) = \mathbf{y}_0$ if and only if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that

$$\mathbf{f}(B_\delta(\mathbf{x}_0) \cap D) \subseteq B_\varepsilon(\mathbf{y}_0)) \,.$$



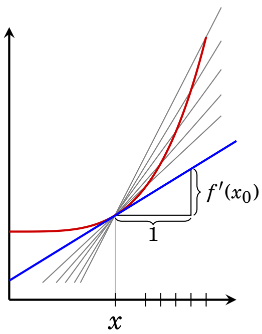
4.3 Derivatives of Univariate Functions

Definition 4.6

Recall that the **derivative** of a function $f : D \subseteq \mathbb{R} \to \mathbb{R}$ at some point x is defined as the limit

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

If this limit exists we say that f is **differentiable** at x. If f is differentiable at every point $x \in D$, we say that f is differentiable on D.



Notice that the term *derivative* is a bit ambiguous. The *derivative at point* x is a *number*, namely the limit of the **difference quotient** of f at point x, that is

$$f'(x) = \frac{d}{dx} f(x) = \lim_{\Delta x\to 0} \frac{\Delta f(x)}{\Delta x} = \lim_{\Delta x\to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x} \,.$$

This number is sometimes called **differential coefficient**. The *differential notation* $\frac{df}{dx}$ is an alternative notation for the derivative which is due to Leibniz. It is very important to remind that *differentiability* is a *local* property of a function.

On the other hand, the **derivative** of f is a *function* that assigns every point x the derivative $\frac{df}{dx}$ at x. Its domain is the set of all points where f is differentiable. Thus $\frac{d}{dx}$ is called the **differential operator** which maps a given function f to its derivative f'. Notice that the differential operator is a linear map, that is

$$\frac{d}{dx}\Big(\alpha f(x) + \beta g(x)\Big) = \alpha \frac{d}{dx} f(x) + \beta \frac{d}{dx} g(x)$$

for all $\alpha, \beta \in \mathbb{R}$, see rules (1) and (2) in Table 4.9.

Differentiability is a stronger property than continuity. Observe that the numerator $f(x+h) - f(x)$ of the difference quotient must coverge to 0 for $h \to 0$ if f is differentiable in x since otherwise the differential quotient would not exist. Thus $\lim_{h\to 0} f(x+h) = f(x)$ and we find:

$f(x)$	$f'(x)$	
c	0	
x^α	$\alpha \cdot x^{\alpha-1}$	(Power rule)
e^x	e^x	
$\ln(x)$	$\frac{1}{x}$	
$\sin(x)$	$\cos(x)$	
$\cos(x)$	$-\sin(x)$	

Table 4.8

Derivatives of some elementary functions.

Lemma 4.7

If $f : D \subseteq \mathbb{R} \to \mathbb{R}$ is differentiable at x, then f is also continuous at x.

See Problem 4.14 for a special case.

Computing limits is a hard job. Therefore, we just list derivatives of some elementary functions in Table 4.8 without proof.

In addition, there exist a couple of rules to reduce the derivative of a given expression to those of elementary functions. Table 4.9 summarizes these rules. Their proofs are straightforward and we given some of these below. See Problem 4.20 for the summation rule and Problem 4.21 for the quotient rule.

PROOF OF RULE (3). Let $F(x) = f(x) \cdot g(x)$. Then we find by Theorem 4.3

$$\begin{aligned}
F'(x) &= \lim_{h\to 0} \frac{F(x+h)-F(x)}{h} \\
&= \lim_{h\to 0} \frac{[f(x+h)\cdot g(x+h)]-[f(x)\cdot g(x)]}{h} \\
&= \lim_{h\to 0} \frac{f(x+h)g(x+h)-f(x)g(x+h)+f(x)g(x+h)-f(x)g(x)}{h} \\
&= \lim_{h\to 0} \frac{f(x+h)-f(x)}{h} g(x+h) + \lim_{h\to 0} f(x)\frac{g(x+h)-g(x)}{h} \\
&= \lim_{h\to 0} \frac{f(x+h)-f(x)}{h} \left[\frac{g(x+h)-g(x)}{h} h + g(x)\right] + \lim_{h\to 0} f(x)\frac{g(x+h)-g(x)}{h} \\
&= f'(x)g(x) + f(x)g'(x)
\end{aligned}$$

as proposed. □

PROOF OF RULE (4). Let $F(x) = (f \circ g)(x) = f(g(x))$. Then

$$F'(x) = \lim_{h\to 0} \frac{F(x+h)-F(x)}{h} = \lim_{h\to 0} \frac{f(g(x+h))-f(g(x))}{h}$$

The change from x to $x+h$ causes the value of g change by the amount $k = g(x+h)-g(x)$. As $\lim_{h\to 0} k = \lim_{h\to 0} \left[\frac{g(x+h)-g(x)}{h}\right] \cdot h = g'(x) \cdot 0 = 0$ we find by

Let g be differentiable at x and f be differentiable at x and $g(x)$. Then sum $f+g$, product $f\cdot g$, composition $f\circ g$, and quotient f/g (for $g(x)\neq 0$) are differentiable at x, and

(1)	$(c\cdot f(x))' = c\cdot f'(x)$	
(2)	$(f(x)+g(x))' = f'(x)+g'(x)$	(Summation rule)
(3)	$(f(x)\cdot g(x))' = f'(x)\cdot g(x)+f(x)\cdot g'(x)$	(Product rule)
(4)	$(f(g(x)))' = f'(g(x))\cdot g'(x)$	(Chain rule)
(5)	$\left(\frac{f(x)}{g(x)}\right)' = \frac{f'(x)\cdot g(x)-f(x)\cdot g'(x)}{(g(x))^2}$	(Quotient rule)

Table 4.9

Rules for differentiation.

Theorem 4.3

$$\begin{aligned} F'(x) &= \lim_{h\to 0} \frac{f(g(x)+k))-f(g(x))}{k}\cdot\frac{k}{h} \\ &= \lim_{h\to 0} \frac{f(g(x)+k))-f(g(x))}{k}\cdot\frac{g(x+h)-g(x)}{h} \\ &= f'(g(x))\cdot g'(x) \end{aligned}$$

as claimed. □

The chain rule can be stated in a quite convenient form by means of differential notation. Let y be function of u, i.e. $y = y(u)$, and u itself is a function of x, i.e., $u = u(x)$, then we find for the derivative of $y(u(x))$,

$$\frac{dy}{dx} = \frac{dy}{du}\cdot\frac{du}{dx}\,.$$

An important application of the chain rule is in the computation of derivatives when variables are changed. Problem 4.24 discusses the case when linear scale is replaces by logarithmic scale.

4.4 Higher Order Derivatives

We have seen that the derivative f' of a function f is again a function. This function may again be differentiable and we then can compute the derivative of derivative f'. It is called the **second derivative** of f and denoted by f''. Recursively, we can compute the third, forth, fifth, ... derivatives denote by f''', f^{iv}, f^{v}, ...

The nth order derivative is denoted by $f^{(n)}$ and we have

$$f^{(n)} = \frac{d}{dx}\left(f^{(n-1)}\right) \quad \text{with} \quad f^{(0)} = f\,.$$

4.5 The Mean Value Theorem

Our definition of the derivative of a function,

$$f'(x) = \lim_{\Delta x \to 0} \frac{f(x+\Delta x) - f(x)}{\Delta x},$$

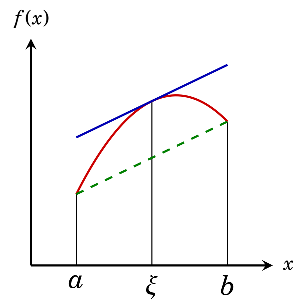
implies for small values of Δx

$$f(x+\Delta x) \approx f(x) + f'(x)\Delta x .$$

The deviation of this *linear* approximation of function f at $x+\Delta x$ becomes small for small values of $|\Delta x|$. We even may improve this approximation.

Mean value theorem. Let f be continuous in the closed bounded interval $[a,b]$ and differentiable in (a,b). Then there exists a point $\xi \in (a,b)$ such that

Theorem 4.10



$$f'(\xi) = \frac{f(b) - f(a)}{b-a} .$$

In particular we find

$$f(b) = f(a) + f'(\xi)(b-a) .$$

PROOF IDEA. We first consider the special case where $f(a) = f(b)$. Then by Theorem 3.37 (and w.l.o.g.) there exists a maximum $\xi \in (a,b)$ of f. We then estimate the limit of the differential quotient when x approaches ξ from the left hand side and from the right hand side, respectively. For the first case we find that $f'(\xi) \geq 0$. The second case implies $f'(\xi) \leq 0$ and hence $f'(\xi) = 0$.

PROOF. Assume first that $f(a) = f(b)$. If f is constant, then we trivially have $f'(x) = 0 = \frac{f(b)-f(a)}{b-a}$ for all $x \in (a,b)$. Otherwise there exists an x with $f(x) \neq f(a)$. Without loss of generality, $f(x) > f(a)$. (Otherwise we consider $-f$.) Let ξ be a maximum of f, i.e., $f(\xi) \geq f(x)$ for all $x \in [a,b]$. By our assumptions, $\xi \in (a,b)$. Now construct sequences $x_k \to \xi$ as $k \to \infty$ with $x_k \in [a,\xi)$ and $y_k \to \xi$ as $k \to \infty$ with $y_k \in (\xi, b]$. Then we find

ξ exists by Theorem 3.37.

$$0 \leq \lim_{k\to\infty} \underbrace{\frac{f(y_k) - f(\xi)}{y_k - \xi}}_{\geq 0} = f'(\xi) = \lim_{k\to\infty} \underbrace{\frac{f(x_k) - f(\xi)}{x_k - \xi}}_{\leq 0} \leq 0 .$$

Consequently, $f'(\xi) = 0$ as claimed.

For the general case consider the function

$$g(x) = f(x) - \frac{f(b) - f(a)}{b-a}(x-a) .$$

Then $g(a) = g(b)$ and there exists a point $\xi \in (a,b)$ such that $g'(\xi) = 0$, i.e., $f'(\xi) - \frac{f(a)-f(b)}{b-a} = 0$. Thus the proposition follows. □

The special case where $f(a) = f(b)$ is also known as **Rolle's theorem**.

4.6 Gradient and Directional Derivatives

Definition 4.11

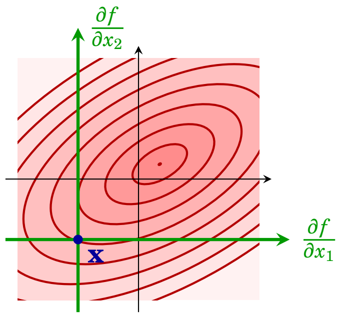
The **partial derivative** of a multivariate function $f(\mathbf{x}) = f(x_1,\ldots,x_n)$ with respect to variable x_i is given as

$$\frac{\partial f}{\partial x_i} = \lim_{h\to 0} \frac{f(\ldots,x_i+h,\ldots)-f(\ldots,x_i,\ldots)}{h}$$

that is, the derivative of f when all variables x_j with $j \neq i$ are held constant.

In the literature there exist several symbols for the partial derivative of f:

$\frac{\partial f}{\partial x_i}$ … derivative w.r.t. x_i
$f_{x_i}(\mathbf{x})$ … derivative w.r.t. variable x_i
$f_i(\mathbf{x})$ … derivative w.r.t. the ith variable
$f'_i(\mathbf{x})$ … ith component of the gradient f'



Notice that the notion of partial derivative is equivalent to the derivative of the univariate function $g(t) = f(\mathbf{x}+t\mathbf{e}_i)$ at $t = 0$, where $\mathbf{e}_i$ denotes the ith unit vector,

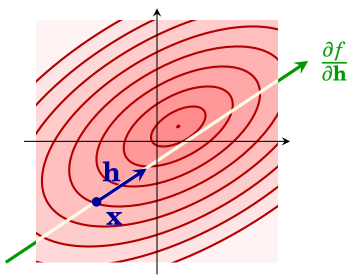
$$f_{x_i}(\mathbf{x}) = \frac{\partial f}{\partial x_i} = \left.\frac{dg}{dt}\right|_{t=0} = \left.\frac{d}{dt} f(\mathbf{x}+t\cdot\mathbf{e}_i)\right|_{t=0}$$

We can, however, replace the unit vectors by arbitrary normalized vectors $\mathbf{h}$ (i.e., $\|\mathbf{h}\| = 1$). Thus we obtain the derivative of f when we move along a straight line through $\mathbf{x}$ in direction $\mathbf{h}$.

Definition 4.12

The **directional derivative** of $f(\mathbf{x}) = f(x_1,\ldots,x_n)$ at $\mathbf{x}$ with respect to $\mathbf{h}$ is given by

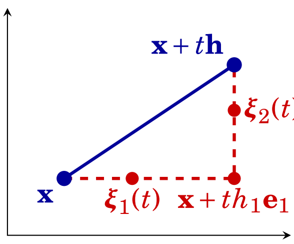
$$f_{\mathbf{h}} = \frac{\partial f}{\partial \mathbf{h}} = \left.\frac{dg}{dt}\right|_{t=0} = \left.\frac{d}{dt} f(\mathbf{x}+t\cdot\mathbf{h})\right|_{t=0}$$



Partial derivatives are special cases of directional derivatives.

The directional derivative can be computed by means of the partial derivatives of f. For the bivariate case ($n = 2$) we find

$$\begin{aligned}\frac{\partial f}{\partial \mathbf{h}} &= \lim_{t\to 0}\frac{f(\mathbf{x}+t\mathbf{h})-f(\mathbf{x})}{t}\\ &= \lim_{t\to 0}\frac{\big(f(\mathbf{x}+t\mathbf{h})-f(\mathbf{x}+th_1\mathbf{e}_1)\big)+\big(f(\mathbf{x}+th_1\mathbf{e}_1)-f(\mathbf{x})\big)}{t}\\ &= \lim_{t\to 0}\frac{f(\mathbf{x}+t\mathbf{h})-f(\mathbf{x}+th_1\mathbf{e}_1)}{t} + \lim_{t\to 0}\frac{f(\mathbf{x}+th_1\mathbf{e}_1)-f(\mathbf{x})}{t}\end{aligned}$$



Notice that $t\mathbf{h} = th_1\mathbf{e}_1 + th_2\mathbf{e}_2$. By the mean value theorem there exists a point $\boldsymbol{\xi}_1(t) \in \{\mathbf{x}+\theta h_1\mathbf{e}_1 : \theta \in (0,t)\}$ such that

$$f(\mathbf{x}+th_1\mathbf{e}_1)-f(\mathbf{x}) = f_{x_1}(\boldsymbol{\xi}_1(t))\cdot th_1$$

and a point $\boldsymbol{\xi}_2(t) \in \{\mathbf{x} + t h_1 \mathbf{e}_1 + \theta h_2 \mathbf{e}_2 : \theta \in (0,t)\}$ such that

$$f(\mathbf{x} + t h_1 \mathbf{e}_1 + t h_2 \mathbf{e}_2) - f(\mathbf{x} + t h_1 \mathbf{e}_1) = f_{x_2}(\boldsymbol{\xi}_2(t)) \cdot t h_2 \,.$$

Consequently,

$$\begin{aligned} \frac{\partial f}{\partial \mathbf{h}} &= \lim_{t \to 0} \frac{f_{x_2}(\boldsymbol{\xi}_2(t)) \cdot t h_2}{t} + \lim_{t \to 0} \frac{f_{x_1}(\boldsymbol{\xi}_1(t)) \cdot t h_1}{t} \\ &= \lim_{t \to 0} f_{x_2}(\boldsymbol{\xi}_2(t)) h_2 + \lim_{t \to 0} f_{x_1}(\boldsymbol{\xi}_1(t)) h_1 \\ &= f_{x_2}(\mathbf{x}) h_2 + f_{x_1}(\mathbf{x}) h_1 \end{aligned}$$

The last equality holds if the partial derivatives f_{x_1} and f_{x_2} are continuous functions of $\mathbf{x}$.

The continuity of the partial derivatives is crucial for our deduction. Thus we define the class of **continuously differentiable functions**, denoted by $\mathscr{C}^1$.

Definition 4.13

A function $f: D \subseteq \mathbb{R}^n \to \mathbb{R}$ belongs to class $\mathscr{C}^m$ if all its partial derivatives of order m or smaller are continuous. The function belongs to class $\mathscr{C}^\infty$ if partial derivatives of all orders exist.

It also seems appropriate to collect all first partial derivatives in a row vector.

Definition 4.14

Gradient. Let $f: D \subseteq \mathbb{R}^n \to \mathbb{R}$ be a $\mathscr{C}^1$ function. Then the **gradient** of f at $\mathbf{x}$ is the row vector

$$f'(\mathbf{x}) = \nabla f(\mathbf{x}) = \big(f_{x_1}(\mathbf{x}), \dots, f_{x_n}(\mathbf{x})\big) \qquad \text{[called "} \textit{nabla } f \text{".]}$$

We can summarize our observations in the following theorem.

Theorem 4.15

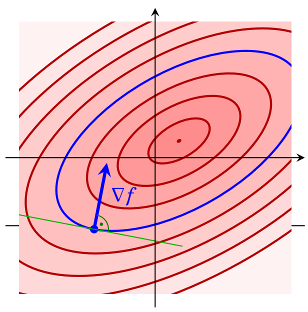
The **directional derivative** of a $\mathscr{C}^1$ function $f(\mathbf{x}) = f(x_1, \dots, x_n)$ at $\mathbf{x}$ with respect to direction $\mathbf{h}$ with $\|\mathbf{h}\| = 1$ is given by

$$\frac{\partial f}{\partial \mathbf{h}}(\mathbf{x}) = f_{x_1}(\mathbf{x}) \cdot h_1 + \dots + f_{x_n}(\mathbf{x}) \cdot h_n = \nabla f(\mathbf{x}) \cdot \mathbf{h} \,.$$

This theorem implies some nice properties of the gradient.

Theorem 4.16

Properties of the gradient. Let $f: D \subseteq \mathbb{R}^n \to \mathbb{R}$ be a $\mathscr{C}^1$ function. Then we find



(1) $\nabla f(\mathbf{x})$ points into the direction of the steepest directional derivative at $\mathbf{x}$.

(2) $\|\nabla f(\mathbf{x})\|$ is the maximum among all directional derivatives at $\mathbf{x}$.

(3) $\nabla f(\mathbf{x})$ is orthogonal to the level set through $\mathbf{x}$.

PROOF. By the Cauchy-Schwarz inequality we have

$$\nabla f(\mathbf{x})\mathbf{h} \le |\nabla f(\mathbf{x})\mathbf{h}| \le \|\nabla f(\mathbf{x})\| \cdot \underbrace{\|\mathbf{h}\|}_{=1} = \nabla f(\mathbf{x}) \frac{\nabla f(\mathbf{x})}{\|\nabla f(\mathbf{x})\|}$$

where equality holds if and only if $\mathbf{h} = \nabla f(\mathbf{x})/\|\nabla f(\mathbf{x})\|$. Thus (1) and (2) follow. For the proof of (3) we need the concepts of level sets and implicit functions. Thus we skip the proof. □

4.7 Higher Order Partial Derivatives

Definition 4.17

The functions f_{x_i} are called **first-order partial derivatives**. Provided that these functions are again differentiable, we can generate new functions by taking their partial derivatives. Thus we obtain **second-order partial derivatives**. They are represented as

$$\frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right) = \frac{\partial^2 f}{\partial x_j \partial x_i} \qquad \text{and} \qquad \frac{\partial}{\partial x_i}\left(\frac{\partial f}{\partial x_i}\right) = \frac{\partial^2 f}{\partial x_i^2}\,.$$

Alternative notations are

$$f_{x_i x_j} \quad \text{and} \quad f_{x_i x_i} \qquad \text{or} \qquad f''_{ij} \quad \text{and} \quad f''_{ii}\,.$$

There are n^2 many second-order derivatives for a function $f(x_1,\dots,x_n)$. Fortunately, for essentially all our functions we need not take care about the succession of particular derivatives. The next theorem provides a sufficient condition. Notice that we again need that all the requested partial derivatives are continuous.

Theorem 4.18

Young's theorem, Schwarz' theorem. Let $f: D \subseteq \mathbb{R}^n \to \mathbb{R}$ be a $\mathscr{C}^m$ function, that is, all the mth order partial derivatives of $f(x_1,\dots,x_n)$ exist and are continuous. If any two of them involve differentiating w.r.t. each of the variables the same number of times, then they are necessarily equal. In particular we find for every $\mathscr{C}^2$ function f,

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}\,.$$

A proof of this theorem is given in most advanced calculus books.

Definition 4.19

Hessian matrix. Let $f: D \subseteq \mathbb{R}^n \to \mathbb{R}$ be a two times differentiable function. Then the $n \times n$ matrix

$$f''(\mathbf{x}) = \mathbf{H}_f(\mathbf{x}) = \begin{pmatrix} f''_{11} & \cdots & f''_{1n} \\ \vdots & \ddots & \vdots \\ f''_{n1} & \cdots & f''_{nn} \end{pmatrix}$$

is called the **Hessian** of f.

By Young's theorem the Hessian is symmetric for $\mathscr{C}^2$ functions.

4.8 Derivatives of Multivariate Functions

We want to generalize the notion of derivative to multivariate functions and transformations. Our starting point is the following observation for univariate functions.

Theorem 4.20

Linear approximation. A function $f: D \subseteq \mathbb{R} \to \mathbb{R}$ is differentiable at an interior point $x_0 \in D$ if and only if there exists a linear function ℓ such that

$$\lim_{h \to 0} \frac{|(f(x_0+h) - f(x_0)) - \ell(h)|}{|h|} = 0\,.$$

We have $\ell(h) = f'(x_0) \cdot h$ (i.e., the differential of f at x_0).

PROOF. Assume that f is differentiable in x_0. Then

$$\begin{aligned}\lim_{h \to 0} \frac{(f(x_0+h) - f(x_0)) - f'(x_0)h}{h} &= \lim_{h \to 0} \frac{f(x_0+h) - f(x_0)}{h} - f'(x_0) \\ &= f'(x_0) - f'(x_0) = 0\,.\end{aligned}$$

Since the absolute value is a continuous function of its argument, the proposition follows.

Conversely, assume that a linear function $\ell(h) = ah$ exists such that

$$\lim_{h \to 0} \frac{|(f(x_0+h) - f(x_0)) - \ell(h)|}{|h|} = 0\,.$$

Then we find

$$\begin{aligned}0 &= \lim_{h \to 0} \frac{|(f(x_0+h) - f(x_0)) - ah|}{|h|} = \lim_{h \to 0} \frac{(f(x_0+h) - f(x_0)) - ah}{h} \\ &= \lim_{h \to 0} \frac{f(x_0+h) - f(x_0)}{h} - a\end{aligned}$$

and consequently

$$\lim_{h \to 0} \frac{f(x_0+h) - f(x_0)}{h} = a\,.$$

But then the limit of the difference quotient exists and f is differentiable at x_0. □

An immediate consequence of Theorem 4.20 is that we can use the existence of such a linear function for the definition of the term *differentiable* and the linear function ℓ for the definition of *derivative*. With the notion of *norm* we can easily extend such a definition to transformations.

Definition 4.21

A function $\mathbf{f}: D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ is **differentiable** at an interior point $\mathbf{x}_0 \in D$ if there exists a linear function $\boldsymbol{\ell}$ such that

$$\lim_{\mathbf{h} \to 0} \frac{\|(\mathbf{f}(\mathbf{x}_0+\mathbf{h}) - \mathbf{f}(\mathbf{x}_0)) - \boldsymbol{\ell}(\mathbf{h})\|}{\|\mathbf{h}\|} = 0\,.$$

The linear function (if it exists) is then given by an $m \times n$ matrix $\mathbf{A}$, i.e., $\boldsymbol{\ell}(\mathbf{h}) = \mathbf{A}\mathbf{h}$. This matrix is called the **(total) derivative** of $\mathbf{f}$ and denoted by $\mathbf{f}'(\mathbf{x}_0)$ or $D\mathbf{f}(\mathbf{x}_0)$.

derivative

Lemma 4.22

A function $\mathbf{f} = (f_1(\mathbf{x}),\dots,f_m(\mathbf{x}))': D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at an interior point $\mathbf{x}_0$ of D if and only if each component function $f_i: D \to \mathbb{R}$ is differentiable.

PROOF. Let $\mathbf{A}$ be an $m\times n$ matrix and $\mathbf{R}(\mathbf{h}) = (\mathbf{f}(\mathbf{x}_0+\mathbf{h})-\mathbf{f}(\mathbf{x}_0))-\mathbf{A}\mathbf{h}$. Then we find for each $j=1,\dots,m$,

$$0 \le |R_j(\mathbf{h})| \le \|\mathbf{R}(\mathbf{h})\|_2 \le \|\mathbf{R}(\mathbf{h})\|_1 = \sum_{i=1}^{m} |R_i(\mathbf{h})|\,.$$

Therefore, $\lim_{\mathbf{h}\to 0} \frac{\|\mathbf{R}(\mathbf{h})\|}{\|\mathbf{h}\|} = 0$ if and only if $\lim_{\mathbf{h}\to 0} \frac{|R_j(\mathbf{h})|}{\|\mathbf{h}\|} = 0$ for all $j = 1,\dots,m$. □

The derivative can be computed by means of the partial derivatives of all the components of $\mathbf{f}$.

Theorem 4.23

Computation of derivative. Let $\mathbf{f} = (f_1(\mathbf{x}),\dots,f_m(\mathbf{x}))': D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ be differentiable at $\mathbf{x}_0$. Then

$$D\mathbf{f}(\mathbf{x}_0) = \begin{pmatrix} \frac{\partial f_1}{x_1}(\mathbf{x}_0) & \dots & \frac{\partial f_1}{x_n}(\mathbf{x}_0) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{x_1}(\mathbf{x}_0) & \dots & \frac{\partial f_m}{x_n}(\mathbf{x}_0) \end{pmatrix} = \begin{pmatrix} \nabla f_1(\mathbf{x}_0) \\ \vdots \\ \nabla f_m(\mathbf{x}_0) \end{pmatrix}$$

This matrix is called the **Jacobian matrix** of $\mathbf{f}$ at $\mathbf{x}_0$.

PROOF IDEA. In order to compute the components of $\mathbf{f}'(\mathbf{x}_0)$ we estimate the change of f_j as function of the kth variable.

PROOF. Let $\mathbf{A} = (\mathbf{a}_1,\dots,\mathbf{a}_m)'$ denote the derivative of $\mathbf{f}$ at $\mathbf{x}_0$ where $\mathbf{a}'_j$ is the jth row vector of $\mathbf{A}$. By Lemma 4.22 each component function f_j is differentiable at $\mathbf{x}_0$ and thus

$$\lim_{\mathbf{h}\to 0} \frac{|(f_j(\mathbf{x}_0+\mathbf{h})-f_j(\mathbf{x}_0))-\mathbf{a}'_j\mathbf{h}|}{\|\mathbf{h}\|} = 0\,.$$

Now set $\mathbf{h} = t\,\mathbf{e}_k$ where $\mathbf{e}_k$ denotes the kth unit vector in $\mathbb{R}^n$. Then

$$\begin{aligned} 0 &= \lim_{t\to 0} \frac{|(f_j(\mathbf{x}_0+t\mathbf{e}_k)-f_j(\mathbf{x}_0))-t\mathbf{a}'_j\mathbf{e}_k|}{|t|} \\ &= \lim_{t\to 0} \frac{f_j(\mathbf{x}_0+t\mathbf{e}_k)-f_j(\mathbf{x}_0)}{t} - \mathbf{a}'_j\mathbf{e}_k \\ &= \frac{\partial f_j}{\partial x_k}(\mathbf{x}_0) - a_{jk}\,. \end{aligned}$$

That is, $a_{jk} = \frac{\partial f_j}{\partial x_k}(\mathbf{x}_0)$, as proposed. □

Notice that an immediate consequence of Theorem 4.23 is that the derivative $\mathbf{f}'(\mathbf{x}_0)$ is uniquely defined (if it exists).

If $f: D \subseteq \mathbb{R}^n \to \mathbb{R}$, then the Jacobian matrix reduces to a row vector and we find $f'(\mathbf{x}) = \nabla f(\mathbf{x})$, i.e., the gradient of f.

The computation by means of the Jacobian matrix suggests that the derivative of a function exists whenever all its partial derivatives exist. However, this need not be the case. Problem 4.25 shows a counterexample. Nevertheless, there exists a simple condition for the existence of the derivative of a multivariate function.

Theorem 4.24

Existence of derivatives. If $\mathbf{f}$ is a $\mathscr{C}^1$ function from an open set $D \subseteq \mathbb{R}^n$ into $\mathbb{R}^m$, then $\mathbf{f}$ is differentiable at every point $\mathbf{x} \in D$.

SKETCH OF PROOF. Similar to the proof of Theorem 4.15 on page 34. □

Differentiability is a stronger property than continuity as the following result shows.

Theorem 4.25

If $\mathbf{f}: D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ is differentiable at an interior point $\mathbf{x}_0 \in D$, then $\mathbf{f}$ is also continuous at $\mathbf{x}_0$.

PROOF. Let $\mathbf{A}$ denote the derivative at $\mathbf{x}_0$. Then we find

$$\begin{aligned}\|\mathbf{f}(\mathbf{x}_0+\mathbf{h})-\mathbf{f}(\mathbf{x}_0)\| &= \|\mathbf{f}(\mathbf{x}_0+\mathbf{h})-\mathbf{f}(\mathbf{x}_0)-\mathbf{A}\mathbf{h}+\mathbf{A}\mathbf{h}\| \\ &\le \underbrace{\|\mathbf{h}\|}_{\to 0} \cdot \underbrace{\frac{\|\mathbf{f}(\mathbf{x}_0+\mathbf{h})-\mathbf{f}(\mathbf{x}_0)-\mathbf{A}\mathbf{h}\|}{\|\mathbf{h}\|}}_{\to 0} + \underbrace{\|\mathbf{A}\mathbf{h}\|}_{\to 0} \to 0 \quad \text{as } \mathbf{h} \to 0.\end{aligned}$$

The ratio tends to 0 since $\mathbf{f}$ is differentiable. Thus $\mathbf{f}$ is continuous at $\mathbf{x}_0$, as claimed. □

Theorem 4.26

Chain rule. Let $\mathbf{f}: D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ and $\mathbf{g}: B \subseteq \mathbb{R}^m \to \mathbb{R}^p$ with $\mathbf{f}(D) \subseteq B$. Suppose $\mathbf{f}$ and $\mathbf{g}$ are differentiable at $\mathbf{x}$ and $\mathbf{f}(\mathbf{x})$, respectively. Then the composite function $\mathbf{g} \circ \mathbf{f}: D \to \mathbb{R}^p$ defined by $(\mathbf{g} \circ \mathbf{f})(\mathbf{x}) = \mathbf{g}(\mathbf{f}(\mathbf{x}))$ is differentiable at $\mathbf{x}$, and

$$(\mathbf{g} \circ \mathbf{f})'(\mathbf{x}) = \mathbf{g}'\big(\mathbf{f}(\mathbf{x})\big) \cdot \mathbf{f}'(\mathbf{x}))\,.$$

PROOF IDEA. A heuristic derivation for the chain rule using linear approximation is obtained in the following way:

$$\begin{aligned}(\mathbf{g} \circ \mathbf{f})'(\mathbf{x})\mathbf{h} &\approx (\mathbf{g} \circ \mathbf{f})(\mathbf{x}+\mathbf{h}) - (\mathbf{g} \circ \mathbf{f})(\mathbf{x}) \\ &= \mathbf{g}\big(\mathbf{f}(\mathbf{x}+\mathbf{h})\big) - \mathbf{g}\big(\mathbf{f}(\mathbf{x})\big) \\ &\approx \mathbf{g}'\big(\mathbf{f}(\mathbf{x})\big)\big[\mathbf{f}(\mathbf{x}+\mathbf{h}) - \mathbf{f}(\mathbf{x})\big] \\ &\approx \mathbf{g}'\big(\mathbf{f}(\mathbf{x})\big)\,\mathbf{f}'(\mathbf{x})\mathbf{h}\end{aligned}$$

for “sufficiently short” vectors $\mathbf{h}$.

PROOF. Let $\mathbf{R}_f(\mathbf{h}) = \mathbf{f}(\mathbf{x}+\mathbf{h}) - \mathbf{f}(\mathbf{x}) - \mathbf{f}'(\mathbf{x})\mathbf{h}$ and $\mathbf{R}_g(\mathbf{k}) = \mathbf{g}\big(\mathbf{f}(\mathbf{x})+\mathbf{k}\big) - \mathbf{g}\big(\mathbf{f}(\mathbf{x})\big) - \mathbf{g}'\big(\mathbf{f}(\mathbf{x})\big)\mathbf{k}$. As both $\mathbf{f}$ and $\mathbf{g}$ are differentiable at $\mathbf{x}$ and $\mathbf{f}(\mathbf{x})$, respectively, $\lim_{\mathbf{h}\to 0} \|\mathbf{R}_f(\mathbf{h})\| / \|\mathbf{h}\| = 0$ and $\lim_{\mathbf{k}\to 0} \|\mathbf{R}_g(\mathbf{k})\| / \|\mathbf{k}\| = 0$. Define $\mathbf{k}(\mathbf{h}) = \mathbf{f}(\mathbf{x}+\mathbf{h}) - \mathbf{f}(\mathbf{x})$. Then we find

$$\begin{aligned}
\mathbf{R}(\mathbf{h}) &= \mathbf{g}\big(\mathbf{f}(\mathbf{x}+\mathbf{h})\big) - \mathbf{g}\big(\mathbf{f}(\mathbf{x})\big) - \mathbf{g}'\big(\mathbf{f}(\mathbf{x})\big)\mathbf{f}'(\mathbf{x})\mathbf{h} \\
&= \mathbf{g}\big(\mathbf{f}(\mathbf{x}) + \mathbf{k}(\mathbf{h})\big) - \mathbf{g}\big(\mathbf{f}(\mathbf{x})\big) - \mathbf{g}'\big(\mathbf{f}(\mathbf{x})\big)\mathbf{f}'(\mathbf{x})\mathbf{h} \\
&= \mathbf{g}'\big(\mathbf{f}(\mathbf{x})\big)\mathbf{k}(\mathbf{h}) + \mathbf{R}_g\big(\mathbf{k}(\mathbf{h})\big) - \mathbf{g}'\big(\mathbf{f}(\mathbf{x})\big)\mathbf{f}'(\mathbf{x})\mathbf{h} \\
&= \mathbf{g}'\big(\mathbf{f}(\mathbf{x})\big)\big[\mathbf{k}(\mathbf{h}) - \mathbf{f}'(\mathbf{x})\mathbf{h}\big] + \mathbf{R}_g\big(\mathbf{k}(\mathbf{h})\big) \\
&= \mathbf{g}'\big(\mathbf{f}(\mathbf{x})\big)\big[\mathbf{f}(\mathbf{x}+\mathbf{h}) - \mathbf{f}(\mathbf{x}) - \mathbf{f}'(\mathbf{x})\mathbf{h}\big] + \mathbf{R}_g\big(\mathbf{k}(\mathbf{h})\big) \\
&= \mathbf{g}'\big(\mathbf{f}(\mathbf{x})\big)\mathbf{R}_f(\mathbf{h}) + \mathbf{R}_g\big(\mathbf{k}(\mathbf{h})\big) .
\end{aligned}$$

Thus by the triangle inequality we have

$$\frac{\|\mathbf{R}(\mathbf{h})\|}{\|\mathbf{h}\|} \le \frac{\big\|\mathbf{g}'(\mathbf{f}(\mathbf{x}))\mathbf{R}_f(\mathbf{h})\big\|}{\|\mathbf{h}\|} + \frac{\big\|\mathbf{R}_g(\mathbf{k}(\mathbf{h}))\big\|}{\|\mathbf{h}\|} .$$

The right hand side converges to zero as $\mathbf{h} \to 0$ and hence proposition follows[1]. □

Notice that the derivatives in the chain rule are *matrices*. Thus the derivative of a composite function is the composite of linear functions.

Example 4.27

Let $\mathbf{f}(x,y) = \begin{pmatrix} x^2+y^2 \\ x^2-y^2 \end{pmatrix}$ and $\mathbf{g}(x,y) = \begin{pmatrix} e^x \\ e^y \end{pmatrix}$ be two differentiable functions defined on $\mathbb{R}^2$. Compute the derivative of $\mathbf{g}\circ\mathbf{f}$ at $\mathbf{x}$ by means of the chain rule.

SOLUTION. Since $\mathbf{f}'(\mathbf{x}) = \begin{pmatrix} 2x & 2y \\ 2x & -2y \end{pmatrix}$ and $\mathbf{g}'(\mathbf{x}) = \begin{pmatrix} e^x & 0 \\ 0 & e^y \end{pmatrix}$, we have

$$\begin{aligned}
(\mathbf{g}\circ\mathbf{f})'(\mathbf{x}) = \mathbf{g}'(\mathbf{f}(\mathbf{x}))\mathbf{f}'(\mathbf{x}) &= \begin{pmatrix} e^{x^2+y^2} & 0 \\ 0 & e^{x^2-y^2} \end{pmatrix} \cdot \begin{pmatrix} 2x & 2y \\ 2x & -2y \end{pmatrix} \\
&= \begin{pmatrix} 2x\,e^{x^2+y^2} & 2y\,e^{x^2+y^2} \\ 2x\,e^{x^2-y^2} & -2y\,e^{x^2-y^2} \end{pmatrix}
\end{aligned}$$

◇

Example 4.28

Derive the formula for the directional derivative from Theorem 4.15 by means of the chain rule.

SOLUTION. Let $f\colon D \subseteq \mathbb{R}^n \to \mathbb{R}$ some differentiable function and $\mathbf{h}$ a fixed direction (with $\|\mathbf{h}\| = 1$). Then $\mathbf{s}\colon \mathbb{R} \to D \subseteq \mathbb{R}^n,\ t \mapsto \mathbf{x}_0 + t\mathbf{h}$ is a path in $\mathbb{R}^n$ and we find

$$f'(\mathbf{s}(0)) = f'(\mathbf{x}_0) = \nabla f(\mathbf{x}_0) \qquad \text{and} \qquad \mathbf{s}'(0) = \mathbf{h}$$

[1] At this point we need some tools from advanced calculus which we do not have available. Thus we unfortunately still have an heuristic approach albeit on some higher level.

and therefore

$$\frac{\partial f}{\partial \mathbf{h}}(\mathbf{x}_0) = (f \circ \mathbf{s})'(0) = f'(\mathbf{s}(0)) \cdot \mathbf{s}'(0) = \nabla f(\mathbf{x}_0) \cdot \mathbf{h}$$

as claimed. ◊

Example 4.29

Let $f(x_1, x_2, t)$ be a differentiable function defined on $\mathbb{R}^3$. Suppose that both $x_1(t)$ and $x_2(t)$ are themselves functions of t. Compute the total derivative of $z(t) = f\big(x_1(t), x_2(t), t\big)$.

SOLUTION. Let $\mathbf{x}\colon \mathbb{R} \to \mathbb{R}^3$, $t \mapsto \begin{pmatrix} x_1(t) \\ x_2(t) \\ t \end{pmatrix}$. Then $z(t) = (f \circ \mathbf{x})(t)$ and we have

$$\begin{aligned} \frac{dz}{dt} &= (f \circ \mathbf{x})'(t) = f'\big(\mathbf{x}(t)\big) \cdot \mathbf{x}'(t) \\ &= \nabla f\big(\mathbf{x}(t)\big) \cdot \begin{pmatrix} x_1'(t) \\ x_2'(t) \\ 1 \end{pmatrix} = \Big(f_{x_1}\big(\mathbf{x}(t)\big), f_{x_2}\big(\mathbf{x}(t)\big), f_t\big(\mathbf{x}(t)\big)\Big) \cdot \begin{pmatrix} x_1'(t) \\ x_2'(t) \\ 1 \end{pmatrix} \\ &= f_{x_1}\big(\mathbf{x}(t)\big) \cdot x_1'(t) + f_{x_2}\big(\mathbf{x}(t)\big) \cdot x_2'(t) + f_t\big(\mathbf{x}(t)\big) \\ &= f_{x_1}(x_1, x_2, t) \cdot x_1'(t) + f_{x_2}(x_1, x_2, t) \cdot x_2'(t) + f_t(x_1, x_2, t). \quad \diamond \end{aligned}$$

— Exercises

4.1 Estimate the following limits:

(a) $\lim_{x\to\infty} \frac{1}{x+1}$ (b) $\lim_{x\to 0} x^2$ (c) $\lim_{x\to\infty} \ln(x)$

(d) $\lim_{x\to 0} \ln|x|$ (e) $\lim_{x\to\infty} \frac{x+1}{x-1}$

4.2 Sketch the following functions.
Which of these are continuous functions?
In which points are these functions not continuous?

(a) $D = \mathbb{R}, f(x) = x$ (b) $D = \mathbb{R}, f(x) = 3x + 1$

(c) $D = \mathbb{R}, f(x) = e^{-x} - 1$ (d) $D = \mathbb{R}, f(x) = |x|$

(e) $D = \mathbb{R}^+, f(x) = \ln(x)$ (f) $D = \mathbb{R}, f(x) = [x]$

(g) $D = \mathbb{R},\ f(x) = \begin{cases} 1 & \text{for } x \le 0 \\ x+1 & \text{for } 0 < x \le 2 \\ x^2 & \text{for } x > 2 \end{cases}$

HINT: Let $x = p + y$ with $p \in \mathbb{Z}$ and $y \in [0,1)$. Then $[x] = p$.

4.3 Differentiate:

(a) $3x^2 + 5\cos(x) + 1$ (b) $(2x+1)x^2$

(c) $x\ln(x)$ (d) $(2x+1)x^{-2}$

(e) $\frac{3x^2-1}{x+1}$ (f) $\ln(\exp(x))$

(g) $(3x-1)^2$ (h) $\sin(3x^2)$

(i) 2^x (j) $\frac{(2x+1)(x^2-1)}{x+1}$

(k) $2e^{3x+1}(5x^2+1)^2 + \frac{(x+1)^3}{x-1} - 2x$

4.4 Compute the second and third derivatives of the following functions:

(a) $f(x) = e^{-\frac{x^2}{2}}$ (b) $f(x) = \dfrac{x+1}{x-1}$

(c) $f(x) = (x-2)(x^2+3)$

4.5 Compute all first and second order partial derivatives of the following functions at $(1,1)$:

(a) $f(x,y) = x + y$ (b) $f(x,y) = xy$

(c) $f(x,y) = x^2 + y^2$ (d) $f(x,y) = x^2 y^2$

(e) $f(x,y) = x^\alpha y^\beta,\quad \alpha,\beta > 0$ (f) $f(x,y) = \sqrt{x^2 + y^2}$

4.6 Compute gradient and Hessian matrix of the functions in Exercise 4.5 at $(1,1)$.

4.7 Let $f(\mathbf{x}) = \sum_{i=1}^{n} x_i^2$. Compute the directional derivative of f into direction $\mathbf{a}$ using

(a) function $g(t) = f(\mathbf{x} + t\mathbf{a})$;
(b) the gradient ∇f;
(c) the chain rule.

4.8 Let $f(x,y)$ be a differentiable function. Suppose its directional derivative in $(0,0)$ in maximal in direction $\mathbf{a} = (1,3)$ with $\frac{\partial f}{\partial \mathbf{a}} = 4$. Compute the gradient of f in $(0,0)$.

4.9 Let $f(x,y) = x^2 + y^2$ and $\mathbf{g}(t) = \begin{pmatrix} g_1(t) \\ g_2(t) \end{pmatrix} = \begin{pmatrix} t \\ t^2 \end{pmatrix}$. Compute the derivatives of the compound functions $f \circ g$ and $g \circ f$ by means of the chain rule.

4.10 Let $\mathbf{f}(\mathbf{x}) = (x_1^3 - x_2, x_1 - x_2^3)'$ and $\mathbf{g}(\mathbf{x}) = (x_2^2, x_1)'$. Compute the derivatives of the compound functions $\mathbf{f}\circ\mathbf{g}$ and $\mathbf{g}\circ\mathbf{f}$ by means of the chain rule.

4.11 Let $\mathbf{A}$ be a regular $n \times n$ matrix, $\mathbf{b} \in \mathbb{R}^n$ and $\mathbf{x}$ the solution of the linear equation $\mathbf{A}\mathbf{x} = \mathbf{b}$. Compute $\frac{\partial x_i}{\partial b_j}$. Also give the Jacobian matrix of $\mathbf{x}$ as a function of $\mathbf{b}$.

HINT: Use Cramer's rule.

4.12 Let $F(K,L,t)$ be a production function where $L = L(t)$ and $K = K(t)$ are also functions of time t. Compute $\frac{dF}{dt}$.

— Problems

4.13 Prove Theorem 4.5.

HINT: See proof of Theorem 3.31.

4.14 Let $f(x) = x^n$ for some $n \in \mathbb{N}$. Show that $f'(x) = n\,x^{n-1}$ by computing the limit of the difference quotient.

HINT: Use the binomial theorem

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k \cdot b^{n-k}$$

Say "n choose k".

4.15 Show that $f(x) = |x|$ is not differentiable on $\mathbb{R}$.

HINT: Recall that a function is differentiable on D if it is differentiable on every $x \in D$.

4.16 Show that

$$f(x) = \begin{cases} \sqrt{x}, & \text{for } x \geq 0, \\ -\sqrt{-x}, & \text{for } x < 0, \end{cases}$$

is not differentiable on $\mathbb{R}$.

4.17 Construct a function that is differentiable but not twice differentiable.

HINT: Recall that a function is differentiable on D if it is differentiable on every $x \in D$.

4.18 Show that the function

$$f(x) = \begin{cases} x^2 \sin\left(\frac{1}{x}\right), & \text{for } x \neq 0, \\ 0, & \text{for } x = 0, \end{cases}$$

is differentiable in $x = 0$ but not continuously differentiable.

4.19 Compute the derivative of $f(x) = a^x$ $(a > 0)$.

HINT: $a^x = e^{\ln(a)x}$

4.20 Prove the summation rule. (Rule (2) in Table 4.9).

HINT: Let $F(x) = f(x) + g(x)$ and apply Theorem 4.3 for the limit.

4.21 Prove the quotient rule. (Rule (5) in Table 4.9).

HINT: Use chain rule, product rule and power rule.

4.22 Verify the *Square Root Rule*:

$$\left(\sqrt{x}\right)' = \frac{1}{2\sqrt{x}}$$

HINT: Use the rules from Tabs. 4.8 and 4.9.

4.23 Let $f : \mathbb{R} \to (0, \infty)$ be a differentiable function. Show that

$$(\ln(f(x)))' = \frac{f'(x)}{f(x)}$$

4.24 Let $f : (0, \infty) \to (0, \infty)$ be a differentiable function. Then the term

$$\varepsilon_f(x) = x \cdot \frac{f'(x)}{f(x)}$$

is called the **elasticity** of f at x. It describes relative changes of f w.r.t. relative changes of its variable x. We can, however, derive the elasticity by changing from a linear scale to a logarithmic scale. Thus we replace variable x by its logarithm $v = \ln(x)$ and differentiate the logarithm of f w.r.t. v and find

$$\varepsilon_f(x) = \frac{d(\ln(f(x)))}{d(\ln(x))}$$

Derive this formula by means of the chain rule.

HINT: Differentiate $y(v) = \ln(f(e^v))$ and substitute $v = \ln(x)$.

4.25 Let

$$f(x, y) = \begin{cases} \dfrac{xy^2}{x^2 + y^4}, & \text{for } (x, y) \neq 0, \\ 0, & \text{for } (x, y) = 0. \end{cases}$$

(a) Plot the graph of f (by means of the computer program of your choice).

(b) Compute all first partial derivatives for $(x, y) \neq 0$.

(c) Compute all first partial derivatives for $(x, y) = 0$ by computing the respective limits

$$f_x(0,0) = \lim_{t\to 0} \frac{f(t,0) - f(0,0)}{t}$$
$$f_y(0,0) = \lim_{t\to 0} \frac{f(0,t) - f(0,0)}{t}$$

(d) Compute the directional derivative at 0 into some direction $\mathbf{h}' = (h_1, h_2)$,

$$f_{\mathbf{h}}(0,0) = \lim_{t\to 0} \frac{f(th_1, th_2) - f(0,0)}{t}$$

What do you expect if f were differentiable at 0?

4.26 Let $\mathbf{f}\colon \mathbb{R}^n \to \mathbb{R}^m$ be a linear function with $\mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x}$ for some matrix $\mathbf{A}$.

(a) What are the dimensions of matrix $\mathbf{A}$ (number of rows and columns)?

(b) Compute the Jacobian matrix of $\mathbf{f}$.

4.27 Let $\mathbf{A}$ be a symmetric $n \times n$ matrix. Compute the Jacobian matrix of the corresponding quadratic form $\mathbf{q}(\mathbf{x}) = \mathbf{x}'\mathbf{A}\mathbf{x}$.

4.28 A function $f(\mathbf{x})$ is called **homogeneous** of degree k, if

$$f(\alpha\mathbf{x}) = \alpha^k f(\mathbf{x}) \quad \text{for all } \alpha \in \mathbb{R}.$$

(a) Give an example for a homogeneous function of degree 2 and draw level lines of this function.

(b) Show that all first order partial derivatives of a differentiable homogeneous function of degree k ($k \geq 1$) are homogeneous of degree $k - 1$.

(c) Show that the level lines are parallel along each ray from the origin. (A **ray** from the origin in direction $\mathbf{r} \neq 0$ is the halfline $\{\mathbf{x} = \alpha\mathbf{r}\colon \alpha \geq 0\}$.)

HINT: Differentiate both sides of equation $f(\alpha\mathbf{x}) = \alpha^k f(\mathbf{x})$ w.r.t. x_i.

4.29 Let f and g be two n times differentiable functions. Show by induction that

$$(f \cdot g)^{(n)}(x) = \sum_{k=0}^{n} \binom{n}{k} f^{(k)}(x) \cdot g^{(n-k)}(x)\,.$$

HINT: Use the recursion $\binom{n+1}{k+1} = \binom{n}{k} + \binom{n}{k+1}$ for $k = 0, \ldots, n-1$.

4.30 Let

$$f(x) = \begin{cases} \exp\left(-\frac{1}{x^2}\right), & \text{for } x \neq 0, \\ 0, & \text{for } x = 0. \end{cases}$$

(a) Show that f is differentiable in $x = 0$.

(b) Show that $f'(x) = \begin{cases} 2x^{-3} f(x), & \text{for } x \neq 0, \\ 0, & \text{for } x = 0. \end{cases}$

(c) Show that f is continuously differentiable in $x = 0$.

(d) Argue why all derivatives of f vanish in $x = 0$, i.e., $f^{(n)}(0) = 0$ for all $n \in \mathbb{N}$.

HINT: For (a) use $\lim_{x\to 0} f(x) = \lim_{x\to\infty} f\left(\frac{1}{x}\right)$; for (b) use the chain rule for the case where $x \neq 0$; for (d) use the formula from Problem 4.29.

5 Taylor Series

We need a local approximation *of a function that is as simple as possible, but not simpler.*

5.1 Taylor Polynomial

The derivative of a function can be used to find the best linear approximation of a univariate function f, i.e.,

$$f(x) \approx f(x_0) + f'(x_0)(x - x_0).$$

Notice that we evaluate both f and its derivative f' at x_0. By the mean value theorem (Theorem 4.10) we have

$$f(x) = f(x_0) + f'(\xi)(x - x_0)$$

for some appropriate point $\xi \in (x, x_0)$. When we need to improve this *first-order approximation*, then we have to use a polynomial p_n of degree n. We thus select the coefficients of this polynomial such that its first n derivatives at some point x_0 coincides with the first k derivatives of f at x_0, i.e.,

$$p_n^{(k)}(x_0) = f^{(k)}(x_0), \quad \text{for } k = 0, \dots, n.$$

We then find

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k + R_n(x).$$

The term R_n is called the **remainder** and is the error when we approximate function f by this so called Taylor polynomial of degree n.

Definition 5.1

Let f be an n times differentiable function. Then the polynomial

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

is called the nth-order **Taylor polynomial** of f around $x = x_0$. The term $f^{(0)}$ refers to the "0-th derivative", i.e., function f itself.

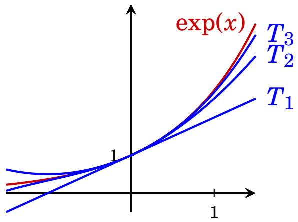
The special case with $x_0 = 0$ is called the **Maclaurin polynomial**.

If we expand the summation symbol we can write the Maclaurin polynomial as

$$T_{f,0,n}(x) = f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \frac{f'''(0)}{3!}x^3 + \cdots + \frac{f^{(n)}(0)}{n!}x^n .$$

Example 5.2

Exponential function. The derivatives of $f(x) = e^x$ at $x_0 = 0$ are given by



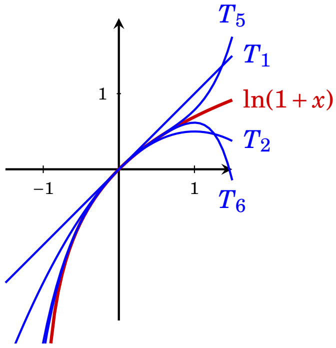
$$f^{(n)}(x) = e^x \quad \text{hence} \quad f^{(n)}(0) = 1 \quad \text{for all } n \geq 0.$$

Therefore we find for the nth order Maclaurin polynomial

$$T_{f,0,n}(x) = \sum_{k=0}^{n} \frac{f^{(k)}(0)}{k!}x^k = \sum_{k=0}^{n} \frac{x^k}{k!} . \qquad \diamondsuit$$

Example 5.3

Logarithm. The derivatives of $f(x) = \ln(1+x)$ at $x_0 = 0$ are given by



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

hence $f^{(n)}(0) = (-1)^{n+1}(n-1)!$ for all $n \geq 1$. As $f(0) = \ln(1) = 0$ we find for the nth order Maclaurin polynomial

$$T_{f,0,n}(x) = \sum_{k=0}^{n} \frac{f^{(k)}(0)}{k!}x^k = \sum_{k=1}^{n} \frac{(-1)^{k+1}(k-1)!}{k!}x^k = \sum_{k=1}^{n} (-1)^{k+1}\frac{x^k}{k} . \qquad \diamondsuit$$

Obviously, the approximation of a function f by its Taylor polynomial is only useful if the remainder $R_n(x)$ is small. Indeed, the error will go to 0 faster than $(x - x_0)^n$ as x tends to x_0.

Theorem 5.4

Taylor's theorem. Let function $f: \mathbb{R} \to \mathbb{R}$ be n times differentiable at the point $x_0 \in \mathbb{R}$. Then there exists a function $h_n: \mathbb{R} \to \mathbb{R}$ such that

$$f(x) = T_{f,x_0,n}(x) + h_n(x)(x - x_0)^n \qquad \text{and} \qquad \lim_{x \to x_0} h_n(x) = 0 .$$

There are even stronger results. The error term can be estimated more precisely. The following theorem gives one such result. Observe that Theorem 5.4 is then just a corollary when the assumptions of Theorem 5.5 are met.

Theorem 5.5

Lagrange's form of the remainder. Suppose f is $n+1$ times differentiable in the interval $[x, x_0]$. Then the remainder for $T_{f,x_0,n}$ can be written as

$$R_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!}(x - x_0)^{n+1}$$

for some point $\xi \in (x, x_0)$.

PROOF IDEA. We construct a function

$$g(t) = R_n(t) - \frac{(t-x_0)^{n+1}}{(x-x_0)^{n+1}} R_n(x)$$

and show that all derivatives $g^{(k)}(x_0) = 0$ vanish for all $k = 0, \dots, n$. Moreover, $g(\xi_0) = 0$ for $\xi_0 = x$ and thus Rolle's Theorem implies that there exists a $\xi_1 \in (\xi_0, x_0)$ such that $g'(\xi_1) = 0$. Repeating this argument recursively we eventually obtain a $\xi = \xi_{n+1} \in (\xi_k, x_0) \subseteq (x, x_0)$ with $g^{(n+1)}(\xi) = f^{(n+1)}(\xi) - \frac{(n+1)!}{(x-x_0)^{n+1}} R_n(x) = 0$ and thus the result follows.

PROOF. Let $R_n(x) = f(x) - T_{f,x_0,n}(x)$ and

$$g(t) = R_n(t) - \frac{(t-x_0)^{n+1}}{(x-x_0)^{n+1}} R_n(x) .$$

We then find $g(x) = 0$. Moreover, $g(x_0) = 0$ and $g^{(k)}(x_0) = 0$ for all $k = 0, \dots, n$ since the first n derivatives of f and $T_{f,x_0,n}$ coincide at x_0 by construction (Problem 5.9). Thus $g(x) = g(x_0)$ and the mean value theorem (Rolle's Theorem, Theorem 4.10) implies that there exists a $\xi_1 \in (x, x_0)$ such that $g'(\xi_1) = 0$ and thus $g'(\xi_1) = g'(x_0) = 0$. Again the mean value theorem implies that there exists a $\xi_2 \in (\xi_1, x_0) \subseteq (x, x_0)$ such that $g''(\xi_2) = 0$. Repeating this argument we find $\xi_1, \xi_2, \dots, \xi_{n+1} \in (x, x_0)$ such that $g^{(k)}(\xi_k) = 0$ for all $k = 1, \dots, n+1$. In particular, for $\xi = \xi_{n+1}$ we then have

$$0 = g^{(n+1)}(\xi) = f^{(n+1)}(\xi) - \frac{(n+1)!}{(x-x_0)^{n+1}} R_n(x)$$

and thus the formula for R_n follows. □

Lagrange's form of the remainder can be seen as a generalization of the mean value theorem for higher order derivatives.

5.2 Taylor Series

Taylor series expansion. The series (Definition 5.6)

$$\sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!} (x-x_0)^n$$

is called the **Taylor series** of f at x_0. We say that we **expand** f into a *Taylor series* around x_0.

If the remainder $R_n(x) \to 0$ as $n \to \infty$, then the Taylor series converges to $f(x)$, i.e., we them have

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!} (x-x_0)^n .$$

Table 5.7 lists Maclaurin series of some important functions. The meaning of ρ is explained in Section 5.4 below.

In some cases it is quite straightforward to show the convergence of the Taylor series.

Table 5.7

Maclaurin series of some elementary functions.

$f(x)$	Maclaurin series		ρ
$\exp(x)$	$= \sum_{n=0}^{\infty} \frac{x^n}{n!}$	$= 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \frac{x^4}{4!} + \cdots$	∞
$\ln(1+x)$	$= \sum_{n=1}^{\infty} (-1)^{n+1} \frac{x^n}{n}$	$= x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots$	1
$\sin(x)$	$= \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}$	$= x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots$	∞
$\cos(x)$	$= \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!}$	$= 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots$	∞
$\frac{1}{1-x}$	$= \sum_{n=0}^{\infty} x^n$	$= 1 + x + x^2 + x^3 + x^4 + \cdots$	1

Theorem 5.8

Convergence of remainder. Assume that *all* derivatives of f are bounded in the interval (x, x_0) by some number M, i.e., $|f^{(k)}(\xi)| \leq M$ for all $\xi \in (x, x_0)$ and all $k \in \mathbb{N}$. Then

$$|R_n(x)| \leq M \frac{|x - x_0|^{n+1}}{(n+1)!} \quad \text{for all } n \in \mathbb{N}$$

and thus $\lim_{n\to\infty} R_n(x) = 0$ as $n \to \infty$.

PROOF. Immediately by Theorem 5.5 and hypothesis of the theorem. □

Example 5.9

We have seen in Example 5.2 that $f^{(n)}(x) = e^x$ for all $n \in \mathbb{N}$. Thus $|f^{(n)}(\xi)| \leq M = \max\{|e^x|, |e^{x_0}|\}$ for all $\xi \in (x, x_0)$ and all $k \in \mathbb{N}$. Then by Theorem 5.8, $e^x = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n$ for all $x \in \mathbb{R}$. ◊

The required order of the Taylor polynomial for the approximation of a function f of course depends on the particular task. A first-order Taylor polynomial may be used to linearize a given function near some point of interest. This also may be sufficient if one needs to investigate local monotonicity of some function. When local convexity or concavity of the function are of interest we need at least a second-order Taylor polynomial.

5.3 Landau Symbols

If all derivatives of f are bounded in the interval (x, x_0), then Lagrange's form of the remainder $R_n(x)$ is expressed as a multiple of the nth power of the distance between the point x of interest and the expansion point x_0, that is, $C|x - x_0|^{n+1}$ for some positive constant C. The constant itself is often hard to compute and thus it is usually not specified. However, in many cases this is not necessary at all.

Suppose we have two terms $C_1|x-x_0|^k$ and $C_2|x-x_0|^{k+1}$ with $C_1, C_2 > 0$, then for values of x *sufficiently close* to x_0 the second term becomes negligible small compared to the first one as

$$\frac{C_2|x-x_0|^{k+1}}{C_1|x-x_0|^k} = \frac{C_2}{C_1} \cdot |x-x_0| \to 0 \quad \text{as } x \to x_0 .$$

More precisely, this ratio can be made as small as desired provided that x is in some sufficiently small open ball around x_0. This observation remains true independent of the particular values of C_1 and C_2. Only the diameter of this *"sufficiently small open ball"* may vary.

Such a situation where we want to describe local or asymptotic behavior of some function up to some non-specified constant is quite common in mathematics. For this purpose the so called *Landau symbol* is used.

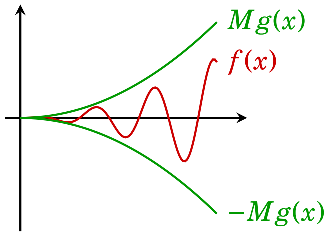
Definition 5.10

Landau symbol. Let $f(x)$ and $g(x)$ be two functions defined on some subset of $\mathbb{R}$. We write

$$f(x) = O\big(g(x)\big) \quad \text{as } x \to x_0 \qquad \text{(say "}f(x)\text{ is big O of }g\text{")}$$

if there exist positive numbers M and δ such that

$$|f(x)| \le M\,|g(x)| \quad \text{for all } x \text{ with } |x-x_0| < \delta .$$



By means of this notation we can write Taylor's formula with the Lagrange form of the remainder as

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k + O(|x-x_0|^{n+1})$$

(provided that f is $n+1$ times differentiable at x_0).

Observe that $f(x) = O\big(g(x)\big)$ implies that there exist positive numbers M and δ such that

$$\left|\frac{f(x)}{g(x)}\right| \le M \quad \text{for all } x \text{ with } |x-x_0| < \delta .$$

We also may have situations where we know that this fraction even converges to 0. Formally, we then write

$$f(x) = o\big(g(x)\big) \quad \text{as } x \to x_0 \qquad \text{(say "}f(x)\text{ is small O of }g\text{")}$$

if for every $\varepsilon > 0$ there exists a positive δ such that

$$|f(x)| \le \varepsilon |g(x)| \quad \text{for all } x \text{ with } |x-x_0| < \delta .$$

Using this notation we can write Taylor's Theorem 5.4 as

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k + o(|x-x_0|^n)$$

The symbols $O(\cdot)$ and $o(\cdot)$ are called **Landau symbols**.

The notation "$f(x) = O\big(g(x)\big)$" is a slight abuse of language as it merely indicates that f belongs to a family of functions that locally behaves similar to $g(x)$. Thus this is sometimes also expressed as

$$f(x) \in O\big(g(x)\big) .$$

5.4 Power Series and Analytic Functions

Taylor series are a special case of so called **power series**

$$p(x) = \sum_{n=1}^{\infty} a_n (x - x_0)^n \, .$$

Suppose that $\lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right|$ exists. Then the ratio test (Lemma 2.28) implies that the power series converges if

$$\lim_{n\to\infty} \left|\frac{a_{n+1}(x-x_0)^{n+1}}{a_n(x-x_0)^n}\right| = \lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| |x - x_0| < 1$$

that is, if

$$|x - x_0| < \lim_{n\to\infty} \left|\frac{a_n}{a_{n+1}}\right| \, .$$

Similarly we find that the series diverges if

$$|x - x_0| > \lim_{n\to\infty} \left|\frac{a_n}{a_{n+1}}\right| \, .$$

Example 5.11

For the exponential function in Example 5.2 we find $a_n = 1/n!$. Thus

$$\lim_{n\to\infty} \left|\frac{a_n}{a_{n+1}}\right| = \lim_{n\to\infty} \frac{(n+1)!}{n!} = \lim_{n\to\infty} n + 1 = \infty \, .$$

Hence the Taylor series converges for all $x \in \mathbb{R}$. ♢

Example 5.12

For function $f(x) = \ln(1+x)$ the situation is different. Recall that for function $f(x) = \ln(1+x)$, we find $a_n = (-1)^{n+1}/n$ (see Example 5.3).

$$\lim_{n\to\infty} \left|\frac{a_n}{a_{n+1}}\right| = \lim_{n\to\infty} \frac{n+1}{n} = 1 \, .$$

Hence the Taylor series converges for all $x \in (-1,1)$; and diverges for $x > 1$ or $x < -1$.
For $x = -1$ we get the divergent harmonic series, see Lemma 2.21. For $x = 1$ we get the convergent alternating harmonic series, see Lemma 2.25. However, a proof requires more sophisticated methods. ♢

Example 5.12 demonstrates that a Taylor series need not converge for all $x \in \mathbb{R}$. Instead there is a maximal distance ρ such that the series converges for all $x \in B_\rho(x_0)$ but diverges for all x with $|x - x_0| > \rho$. The value ρ is called the **radius of convergence** of the power series. Table 5.7 also lists this radius for the given Maclaurin series. $\rho = \infty$ means that the series converges for all $x \in \mathbb{R}$.

There is, however, a subtle difference between Examples 5.9 and 5.11. In the first example we have show that $\sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n = e^x$ for all $x \in \mathbb{R}$ while in the latter we have just shown that $\sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n$ converges.

Similarly, we have shown in Example 5.12 that the Taylor series which we have computed in Example 5.3 converges, but we have not given a proof that $\sum_{k=1}^{n}(-1)^{k+1}\frac{x^k}{k} = \ln(1+x)$.

Indeed functions f exist where the Taylor series converge but do not coincide with $f(x)$.

Example 5.13

The function

$$f(x) = \begin{cases} \exp\left(-\frac{1}{x^2}\right), & \text{for } x \neq 0, \\ 0, & \text{for } x = 0. \end{cases}$$

is infinitely differentiable in $x = 0$ and $f^{(n)}(0) = 0$ for all $n \in \mathbb{N}$ (see Problem 4.30). Consequently, we find for all Maclaurin polynomials $T_{f,n,0}(x) = 0$ for all $x \in \mathbb{R}$. Thus the Maclaurin series converges to 0 for all $x \in \mathbb{R}$. However, $f(x) > 0$ for all $x \neq 0$, i.e., albeit the series converges we find $\sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n \neq f(x)$. ♢

Definition 5.14

Analytic function. An infinitely differentiable function f is called **analytic** in an open interval $B_r(x_0)$ if its Taylor series around x_0 converges and

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n \quad \text{for all } x \in B_r(x_0).$$

5.5 Defining Functions

Computations with power series are quite straightforward. Power series can be

- added or subtracted termwise,
- multiplied,
- divided,
- differentiated and integrated termwise.

Example 5.15

We get the Maclaurin series of the exponential function by differentiating the Maclaurin series of e^x:

$$\begin{aligned} \big(\exp(x)\big)' &= \left(\sum_{n=0}^{\infty} \frac{1}{n!} x^n\right)' = \sum_{n=0}^{\infty} \frac{1}{n!} \left(x^n\right)' = \sum_{n=1}^{\infty} \frac{n}{n!} x^{n-1} = \sum_{n=1}^{\infty} \frac{1}{(n-1)!} x^{n-1} \\ &= \sum_{n=0}^{\infty} \frac{1}{n!} x^n = \exp(x)\,. \end{aligned}$$

Example 5.16

We get the Maclaurin series of $f(x) = x^2 \cdot \sin(x)$ by multiplying the Maclaurin series of $\sin(x)$ by x^2:

$$\begin{aligned} x^2 \cdot \sin(x) &= x^2 \cdot \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} = \sum_{n=0}^{\infty} (-1)^n x^2 \frac{x^{2n+1}}{(2n+1)!} \\ &= \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+3}}{(2n+1)!}\,. \end{aligned}$$

We also can substitute x in the Maclaurin series from Table 5.7 by some polynomial.

Example 5.17

We obtain the Maclaurin series of $\exp(-x^2)$ by substituting $-x^2$ into the Maclaurin series of the exponential function.

$$\exp(-x^2) = \sum_{n=0}^{\infty} \frac{1}{n!}\left(-x^2\right)^n = \sum_{n=0}^{\infty} \frac{(-1)^n}{n!} x^{2n} .$$

For that reason it is quite convenient to define analytic functions by its Taylor series.

$$\exp(x) := \sum_{n=0}^{\infty} \frac{1}{n!} x^n$$

5.6 Taylor's Formula for Multivariate Functions

Taylor polynomials can also be established for multivariate functions. We then construct a polynomial where all its kth order partial derivatives coincide with the corresponding partial derivatives of f at some given expansion point $\mathbf{x}_0$.

In opposition to the univariate case the number of coefficients of a polynomial in two or more variables increases exponentially in the degree of the polynomial. Thus we restrict our interest to the 2nd order Taylor polynomials which can be written as

$$p_2(x_1,\dots,x_n) = a_0 + \sum_{i=1}^{n} a_i x^i + \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} x^i x^j$$

or, using vectors and quadratic forms,

$$p_2(\mathbf{x}) = a_0 + \mathbf{a}'\mathbf{x} + \mathbf{x}'\mathbf{A}\mathbf{x}$$

where $\mathbf{A}$ is an $n \times n$ matrix with $[\mathbf{A}]_{ij} = a_{ij}$ and $\mathbf{a}' = (a_1,\dots,a_n)$.

If we choose the coefficients a_i and a_{ij} such that all first and second order partial derivatives of p_2 at $\mathbf{x}_0 = 0$ coincides with the corresponding derivatives of f we find,

$$p_2(\mathbf{x}) = f(0) + f'(0)\mathbf{x} + \frac{1}{2}\mathbf{x}' f''(0)\mathbf{x}$$

For a general expand point $\mathbf{x}_0$ we get the following analog to Taylor's Theorem 5.4 which we state without proof.

Theorem 5.18

Taylor's formula for multivariate functions. Suppose that f is a $\mathscr{C}^3$ function in an open set containing the line segment $[\mathbf{x}^0, \mathbf{x}^0 + \mathbf{h}]$. Then

$$f(\mathbf{x}_0 + \mathbf{h}) = f(\mathbf{x}_0) + f'(\mathbf{x}_0)\cdot\mathbf{h} + \frac{1}{2}\mathbf{h}'\cdot f''(\mathbf{x}_0)\cdot\mathbf{h} + O(\|\mathbf{h}\|^3) .$$

Example 5.19

Let $f(x,y) = e^{x^2-y^2} + x$. Then gradient and Hessian matrix are given by

$$f'(x,y) = \left(2x\,e^{x^2-y^2} + 1,\ -2y\,e^{x^2-y^2}\right)$$

$$f''(x,y) = \begin{pmatrix} (2+4x^2)\,e^{x^2-y^2} & -4xy\,e^{x^2-y^2} \\ -4xy\,e^{x^2-y^2} & (-2+4y^2)\,e^{x^2-y^2} \end{pmatrix}$$

and thus we get for the 2nd order Taylor polynomial around $\mathbf{x}_0 = 0$

$$\begin{aligned} f(x,y) &= f(0,0) + f'(0,0)(x,y)' + \frac{1}{2}(x,y)f''(0,0)(x,y)' + O(\|(x,y)\|^3) \\ &= 1 + (1,0)(x,y)' + \frac{1}{2}(x,y)\begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}(x,y)' + O(\|(x,y)\|^3) \\ &= 1 + x + x^2 - y^2 + O(\|(x,y)\|^3)\,. \end{aligned}$$

— Exercises

5.1 Expand $f(x) = \frac{1}{2-x}$ into a Maclaurin polynomial of

(a) first order;

(b) second order.

Draw the graph of $f(x)$ and of these two Maclaurin polynomials in the interval $[-3,5]$.
Give an estimate for the radius of convergence.

5.2 Expand $f(x) = (x+1)^{1/2}$ into the 3rd order Taylor polynomial around $x_0 = 0$.

5.3 Expand $f(x) = \sin(x^{10})$ into a Maclaurin polynomial of degree 30.

5.4 Expand $f(x) = \sin(x^2 - 5)$ into a Maclaurin polynomial of degree 4.

5.5 Expand $f(x) = 1/(1+x^2)$ into a Maclaurin series. Compute its radius of convergence.

5.6 Expand the density of the standard normal distribution $f(x) = \exp\left(-\frac{x^2}{2}\right)$ into a Maclaurin series. Compute its radius of convergence.

5.7 Expand $f(x,y) = e^{x^2+y^2}$ into a 2nd order Taylor series around $\mathbf{x}_0 = (0,0)$.

— Problems

5.8 Expand the exponential function $\exp(x)$ into a Taylor series about $x_0 = 0$. Give an upper bound of the remainder $R_n(1)$ as a function of order n. When is this bound less than 10^{-16}?

5.9 Assume that f is n times differentiable in x_0. Show that for the first n derivative of f and of its n-order Taylor polynomial coincide in x_0, i.e.,

$$\left(T_{f,x_0,n}\right)^{(k)}(x_0) = f^{(k)}(x_0)\,, \quad \text{for all } k = 0,\ldots,n.$$

5.10 Verify the Maclaurin series from Table 5.7.

5.11 Show by means of the Maclaurin series from Table 5.7 that

(a) $(\sin(x))' = \cos(x)$ (b) $\left(\ln(1+x)\right)' = \frac{1}{1+x}$

6 Inverse and Implicit Functions

Can we invert the action of some function?

6.1 Inverse Functions

Inverse function. Let $\mathbf{f}\colon D_f \subseteq \mathbb{R}^n \to W_f \subseteq \mathbb{R}^m, \mathbf{x} \mapsto \mathbf{y} = \mathbf{f}(\mathbf{x})$ be some function. Suppose that there exists a function $\mathbf{f}^{-1}\colon W_f \to D_f, \mathbf{y} \mapsto \mathbf{x} = \mathbf{f}^{-1}(\mathbf{y})$ such that — Definition 6.1

$$\mathbf{f}^{-1} \circ \mathbf{f} = \mathbf{f} \circ \mathbf{f}^{-1} = \mathbf{id}$$

that is, $\mathbf{f}^{-1}(\mathbf{f}(\mathbf{x})) = \mathbf{f}^{-1}(\mathbf{y}) = \mathbf{x}$ for all $\mathbf{x} \in D_f$, and $\mathbf{f}(\mathbf{f}^{-1}(\mathbf{y})) = \mathbf{f}(\mathbf{x}) = \mathbf{y}$ for all $\mathbf{y} \in W_f$. Then $\mathbf{f}^{-1}$ is called the **inverse function** of $\mathbf{f}$.

Obviously, the inverse function exists if and only if $\mathbf{f}$ is a *bijection*. — Lemma 6.2

We get the function term of the inverse function by solving equation $\mathbf{y} = \mathbf{f}(\mathbf{x})$ w.r.t. to $\mathbf{x}$.

Affine function. Suppose that $\mathbf{f}\colon \mathbb{R}^n \to \mathbb{R}^m, \mathbf{x} \mapsto \mathbf{y} = \mathbf{f}(\mathbf{x}) = \mathbf{A}\mathbf{x} + \mathbf{b}$ where $\mathbf{A}$ is an $m \times n$ matrix and $\mathbf{b} \in \mathbb{R}^m$. Then we find — Example 6.3

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{b} \quad \Leftrightarrow \quad \mathbf{x} = \mathbf{A}^{-1}\mathbf{y} - \mathbf{A}^{-1}\mathbf{b} = \mathbf{A}^{-1}(\mathbf{y} - \mathbf{b})$$

provided that $\mathbf{A}$ is invertible. In particular we must have $n = m$. Thus we have $\mathbf{f}^{-1}(\mathbf{y}) = \mathbf{A}^{-1}(\mathbf{y} - \mathbf{b})$. Observe that

$$D\mathbf{f}^{-1}(\mathbf{y}) = \mathbf{A}^{-1} = (D\mathbf{f}(\mathbf{x}))^{-1} . \qquad \diamondsuit$$

For an arbitrary function the inverse need not exist. E.g., the function $f\colon \mathbb{R} \to \mathbb{R}, x \mapsto x^2$ is not invertible. However, if we restrict the domain of our function to some (sufficiently small) open interval $D = B_\varepsilon(x_0) \subset$

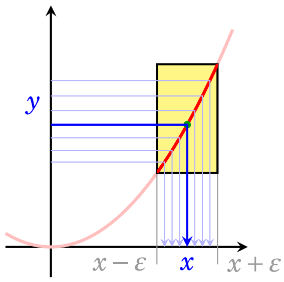
$(0,\infty)$ then the inverse exists. Motivated by Example 6.3 above we expect that this always works whenever $f'(x_0) \neq 0$, i.e., when $\frac{1}{f'(x_0)}$ exists. Moreover, it is possible to compute the derivative $\left(f^{-1}\right)'$ of its inverse in $y_0 = f(x_0)$ as $\left(f^{-1}\right)'(y_0) = \frac{1}{f'(x_0)}$ without having an explicit expression for f^{-1}.

This useful fact is stated in the inverse function theorem.

Theorem 6.4

Inverse function theorem. Let $\mathbf{f}\colon D \subseteq \mathbb{R}^n \to \mathbb{R}^n$ be a $\mathscr{C}^k$ function in some open set D containing $\mathbf{x}^0$. Suppose that the **Jacobian determinant** of $\mathbf{f}$ at $\mathbf{x}^0$ is nonzero, i.e.,

$$\frac{\partial(f_1,\ldots,f_n)}{\partial(x_1,\ldots,x_n)} = \left|\mathbf{f}'(\mathbf{x}^0)\right| \neq 0 \quad \text{for } \mathbf{x} = \mathbf{x}^0 .$$



Then there exists an open set U around $\mathbf{x}^0$ such that $\mathbf{f}$ maps U one-to-one onto an open set V around $\mathbf{y}^0 = \mathbf{f}(\mathbf{x}^0)$. Thus there exists an inverse mapping $\mathbf{f}^{-1}\colon V \to U$ which is also in $\mathscr{C}^k$. Moreover, for all $\mathbf{y} \in V$, we have

$$(\mathbf{f}^{-1})'(\mathbf{y}^0) = (\mathbf{f}'(\mathbf{x}^0))^{-1} .$$

In other words, a $\mathscr{C}^k$ function $\mathbf{f}$ with a nonzero Jacobian determinant at $\mathbf{x}^0$ has a *local inverse* around $\mathbf{f}(\mathbf{x}^0)$ which is again $\mathscr{C}^k$.

This theorem is an immediate corollary of the Implicit Function Theorem 6.11 below, see Problem 6.10. The idea behind the proof is that we can locally replace function $\mathbf{f}$ by its differential in order to get its local inverse.

For the case $n = 1$, that is, a function $f\colon \mathbb{R} \to \mathbb{R}$, we find

$$(f^{-1})'(y_0) = \frac{1}{f'(x_0)} \qquad \text{where } y_0 = f(x_0).$$

Example 6.5

Let $f\colon \mathbb{R} \to \mathbb{R}, x \mapsto y = f(x) = x^2$ and $x_0 = 3$. Then $f'(x_0) = 6 \neq 0$ thus f^{-1} exists in open ball around $y_0 = f(x_0) = 9$. Moreover

$$(f^{-1})'(9) = \frac{1}{f'(3)} = \frac{1}{6} .$$

We remark here that Theorem 6.4 does *not* imply that function f^{-1} does not exist in any open ball around $f(0)$. As $f'(0) = 0$ we simply cannot apply the theorem in this case. ◊

Example 6.6

Let $\mathbf{f}\colon \mathbb{R}^2 \to \mathbb{R}^2, \mathbf{x} \mapsto \mathbf{f}(\mathbf{x}) = \begin{pmatrix} x_1^2 - x_2^2 \\ x_1 x_2 \end{pmatrix}$. Then we find $D\mathbf{f}(\mathbf{x}) = \begin{pmatrix} 2x_1 & -2x_2 \\ x_2 & x_1 \end{pmatrix}$ and thus

$$\frac{\partial(f_1,f_2)}{\partial(x_1,x_2)} = |D\mathbf{f}(\mathbf{x})| = \begin{vmatrix} 2x_1 & -2x_2 \\ x_2 & x_1 \end{vmatrix} = 2x_1^2 + 2x_2^2 \neq 0$$

for all $\mathbf{x} \neq 0$. Consequently, $\mathbf{f}^{-1}$ exists around all $\mathbf{y} = \mathbf{f}(\mathbf{x})$ where $\mathbf{x} \neq 0$. The derivative at $\mathbf{y} = \mathbf{f}(1,1)$ is given by

$$D(\mathbf{f}^{-1})(\mathbf{y}) = (D\mathbf{f}(1,1))^{-1} = \begin{pmatrix} 2 & -2 \\ 1 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} \frac{1}{4} & \frac{2}{4} \\ -\frac{1}{4} & \frac{2}{4} \end{pmatrix} .$$ ◊

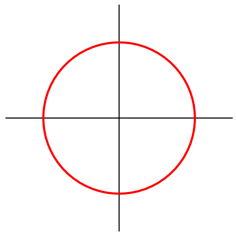
6.2 Implicit Functions

Suppose we are given some function $F(x,y)$. Then equation $F(x,y) = 0$ describes a relation between the two variables x and y. Then if we fix x then y is implicitly given. Thus we call this an **implicit function**. One may ask the question whether it is possible to express y as an *explicit function* of x.

Example 6.7

Linear function. Let $F(x,y) = ax + by = 0$ for $a, b \in \mathbb{R}$. Then we easily find $y = f(x) = -\frac{a}{b}x$ provided that $b \neq 0$. Observe that $F_x = a$ and $F_y = b$. Thus we find

$$y = -\frac{F_x}{F_y}x \quad \text{and} \quad \frac{dy}{dx} = -\frac{F_x}{F_y} \qquad \text{provided that } F_y \neq 0. \qquad \diamondsuit$$



For non-linear functions this need not work. E.g., for

$$F(x,y) = x^2 + y^2 - 1 = 0$$

it is not possible to globally express y as a function of x. Nevertheless, we may try to find such an explicit expression that works locally, i.e., within an open rectangle around a given point (x_0, y_0) that satisfies this equation. Thus we replace F locally by its total derivative

$$dF = F_x dx + F_y dy = d0 = 0$$

and obtain formally the derivative

$$\frac{dy}{dx} = -\frac{F_x}{F_y}.$$

Obviously this only works when $F_y(x_0, y_0) \neq 0$.

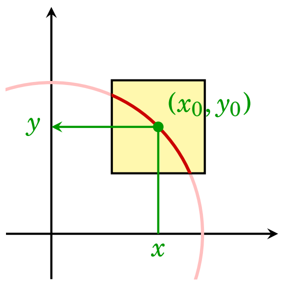
Theorem 6.8

Implicit function theorem. Let $F: D \subseteq \mathbb{R}^2 \to \mathbb{R}$ be a differentiable function in some open set D. Consider an interior point $(x_0, y_0) \in D$ where

$$F(x_0, y_0) = 0 \qquad \text{and} \qquad F_y(x_0, y_0) \neq 0.$$

Then there exists an open rectangle R around (x_0, y_0), such that

- $F(x,y) = 0$ has a unique solution $y = f(x)$ in R, and
- $\frac{dy}{dx} = -\frac{F_x}{F_y}$.



Example 6.9

Let $F(x,y) = x^2 + y^2 - 8 = 0$ and $(x_0, y_0) = (2,2)$. Since $F(x_0, y_0) = 0$ and $F_y(x_0, y_0) = 2y_0 = 4 \neq 0$, there exists a rectangle R around $(2,2)$ such that y can be expressed as an explicit function of x and we find

$$\frac{dy}{dx}(x_0) = -\frac{F_x(x_0, y_0)}{F_y(x_0, y_0)} = -\frac{2x_0}{2y_0} = -\frac{4}{4} = -1.$$

Observe that we cannot apply Theorem 6.8 for the point $(\sqrt{8},0)$ as then $F_y(\sqrt{8},0) = 0$. Thus the hypothesis of the theorem is violated. Notice, however, that this does *not* necessarily imply that the requested local explicit function does not exist at all. ♢

We can generalize Theorem 6.8 to the functions with arbitrary numbers of arguments. Thus we first need a generalization of the partial derivative.

Definition 6.10

Jacobian matrix. Let $\mathbf{F}\colon \mathbb{R}^{n+m} \to \mathbb{R}^m$ be a differentiable function with

$$(\mathbf{x},\mathbf{y}) \mapsto \mathbf{F}(\mathbf{x},\mathbf{y}) = \begin{pmatrix} F_1(x_1,\dots,x_n,y_1,\dots,y_m) \\ \vdots \\ F_m(x_1,\dots,x_n,y_1,\dots,y_m) \end{pmatrix}$$

Then the matrix

$$\frac{\partial \mathbf{F}(\mathbf{x},\mathbf{y})}{\partial \mathbf{y}} = \begin{pmatrix} \frac{\partial F_1}{\partial y_1} & \cdots & \frac{\partial F_1}{\partial y_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial F_m}{\partial y_1} & \cdots & \frac{\partial F_m}{\partial y_m} \end{pmatrix}$$

is called the **Jacobian matrix** of $\mathbf{F}(\mathbf{x},\mathbf{y})$ w.r.t. $\mathbf{y}$.

Theorem 6.11

Implicit function theorem. Let $\mathbf{F}\colon D \subseteq \mathbb{R}^{n+m} \to \mathbb{R}^m$ be $\mathscr{C}^k$ in some open set D. Consider an interior point $(\mathbf{x}^0,\mathbf{y}^0) \in D$ where

$$\mathbf{F}(\mathbf{x}_0,\mathbf{y}_0) = 0 \qquad \text{and} \qquad \left|\frac{\partial \mathbf{F}(\mathbf{x},\mathbf{y})}{\partial \mathbf{y}}\right| \neq 0 \quad \text{for } (\mathbf{x},\mathbf{y}) = (\mathbf{x}_0,\mathbf{y}_0).$$

Then there exist open balls $B(\mathbf{x}^0) \subseteq \mathbb{R}^n$ and $B(\mathbf{y}^0) \subseteq \mathbb{R}^m$ around $\mathbf{x}^0$ and $\mathbf{y}^0$, respectively, with $B(\mathbf{x}^0) \times B(\mathbf{y}^0) \subseteq D$ such that for every $\mathbf{x} \in B(\mathbf{x}^0)$ there exists a unique $\mathbf{y} \in B(\mathbf{y}^0)$ with $\mathbf{F}(\mathbf{x},\mathbf{y}) = 0$. In this way we obtain a $\mathscr{C}^k$ function $\mathbf{f}\colon B(\mathbf{x}^0) \subseteq \mathbb{R}^n \to B(\mathbf{y}^0) \subseteq \mathbb{R}^m$ with $\mathbf{f}(\mathbf{x}) = \mathbf{y}$. Moreover,

$$\frac{\partial \mathbf{y}}{\partial \mathbf{x}} = -\left(\frac{\partial \mathbf{F}}{\partial \mathbf{y}}\right)^{-1} \cdot \left(\frac{\partial \mathbf{F}}{\partial \mathbf{x}}\right)$$

The proof of this theorem requires tools from advanced calculus which are beyond the scope of this course. Nevertheless, the rule for the derivative for the local inverse function (if it exists) can be easily derived by means of the chain rule, see Problem 6.12.

Obviously, Theorem 6.8 is just a special case of Theorem 6.11. For the special case where $F\colon \mathbb{R}^{n+1} \to \mathbb{R}$, $(\mathbf{x}, y) \mapsto F(\mathbf{x}, y) = F(x_1,\dots,x_n, y)$, and some point $(\mathbf{x}_0, y)$ with $F(\mathbf{x}_0, y_0) = 0$ and $F_y(\mathbf{x}_0, y_0) \neq 0$ we then find that there exists an open rectangle around $(\mathbf{x}_0, y)$ such that $y = f(\mathbf{x})$ and

$$\frac{\partial y}{\partial x_i} = -\frac{F_{x_i}}{F_y}\,.$$

Let Example 6.12

$$F(x_1,x_2,x_3,x_4) = x_1^2 + x_2 x_3 + x_3^2 - x_3 x_4 - 1 = 0\,.$$

We are given a point $(x_1,x_2,x_3,x_4) = (1,0,1,1)$. We find $F(1,0,1,1) = 0$ and $F_{x_2}(1,0,1,1) = 1 \neq 0$. Thus there exists an open rectangle where x_2 can be expressed locally by an explicit function of the remaining variables, $x_2 = f(x_1,x_3,x_4)$, and we find for the partial derivative w.r.t. in $(x_1,x_3,x_4) = (1,1,1)$,

$$\frac{\partial x_2}{\partial x_3} = -\frac{F_{x_3}}{F_{x_2}} = -\frac{x_2 + 2x_3 - x_4}{x_3} = -1\,.$$

Notice that we cannot apply the Implicit Function Theorem neither at $(1,1,1,1)$ nor at $(1,1,0,1)$ as $F(1,1,1,1) \neq 0$ and $F_{x_2}(1,1,0,1) = 0$, respectively. ♢

Let Example 6.13

$$\mathbf{F}(\mathbf{x},\mathbf{y}) = \begin{pmatrix} F_1(x_1,x_2,y_1,y_2) \\ F_2(x_1,x_2,y_1,y_2) \end{pmatrix} = \begin{pmatrix} x_1^2 + x_2^2 - y_1^2 - y_2^2 + 3 \\ x_1^3 + x_2^3 + y_1^3 + y_2^3 - 11 \end{pmatrix}$$

and some point $(\mathbf{x}_0,\mathbf{y}_0) = (1,1,1,2)$.

$$\frac{\partial \mathbf{F}}{\partial \mathbf{x}} = \begin{pmatrix} \frac{\partial F_1}{\partial x_1} & \frac{\partial F_1}{\partial x_2} \\ \frac{\partial F_2}{\partial x_1} & \frac{\partial F_2}{\partial x_2} \end{pmatrix} = \begin{pmatrix} 2x_1 & 2x_2 \\ 3x_1^2 & 3x_2^2 \end{pmatrix} \quad \text{and} \quad \frac{\partial \mathbf{F}}{\partial \mathbf{x}}(1,1,1,2) = \begin{pmatrix} 2 & 2 \\ 3 & 3 \end{pmatrix}$$

$$\frac{\partial \mathbf{F}}{\partial \mathbf{y}} = \begin{pmatrix} \frac{\partial F_1}{\partial y_1} & \frac{\partial F_1}{\partial y_2} \\ \frac{\partial F_2}{\partial y_1} & \frac{\partial F_2}{\partial y_2} \end{pmatrix} = \begin{pmatrix} -2y_1 & -2y_2 \\ 3y_1^2 & 3y_2^2 \end{pmatrix} \quad \text{and} \quad \frac{\partial \mathbf{F}}{\partial \mathbf{y}}(1,1,1,2) = \begin{pmatrix} -2 & -4 \\ 3 & 12 \end{pmatrix}$$

Since $\mathbf{F}(1,1,1,2) = 0$ and $\left|\frac{\partial \mathbf{F}(\mathbf{x},\mathbf{y})}{\partial \mathbf{y}}\right| = -12 \neq 0$ we can apply the Implicit Function Theorem and get

$$\frac{\partial \mathbf{y}}{\partial \mathbf{x}} = -\left(\frac{\partial \mathbf{F}}{\partial \mathbf{y}}\right)^{-1} \cdot \left(\frac{\partial \mathbf{F}}{\partial \mathbf{x}}\right) = -\frac{1}{-12}\begin{pmatrix} 12 & 4 \\ -3 & -2 \end{pmatrix} \cdot \begin{pmatrix} 2 & 2 \\ 3 & 3 \end{pmatrix} = \begin{pmatrix} 3 & 3 \\ -1 & -1 \end{pmatrix}.$$ ♢

— Exercises

6.1 Let $\mathbf{f}\colon \mathbb{R}^2 \to \mathbb{R}^2$ be a function with

$$\mathbf{x} \mapsto \mathbf{f}(\mathbf{x}) = \mathbf{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} x_1 - x_1 x_2 \\ x_1 x_2 \end{pmatrix}$$

(a) Compute the Jacobian matrix and determinant of $\mathbf{f}$.

(b) Around which points is it possible to find a local inverse of $\mathbf{f}$?

(c) Compute the Jacobian matrix for the inverse function.

(d) Compute the inverse function (where it exists).

6.2 Let $T\colon \mathbb{R}^2 \to \mathbb{R}^2$ be a function with

$$(x, y) \mapsto (u, v) = (ax + by, cx + dy)$$

where a, b, c, and d are non-zero constants.

Show: If the Jacobian determinant of T equals 0, then the image of T is a straight line through the origin.

6.3 Give a sufficient condition for f and g such that the equations

$$u = f(x, y), \quad v = g(x, y)$$

can be solved w.r.t. x and y.

Suppose we have the solutions $x = F(u, v)$ and $y = G(u, v)$. Compute $\frac{\partial F}{\partial u}$ and $\frac{\partial G}{\partial u}$.

6.4 Show that the following equations define y as a function of x in an interval around x_0. Compute $y'(x_0)$.

(a) $y^3 + y - x^3 = 0, \quad x_0 = 0$

(b) $x^2 + y + \sin(xy) = 0, \quad x_0 = 0$

6.5 Compute $\frac{dy}{dx}$ from the implicit function $x^2 + y^3 = 0$.
For which values of x does an explicit function $y = f(x)$ exist locally?
For which values of y does an explicit function $x = g(y)$ exist locally?

6.6 Which of the given implicit functions can be expressed as $z = g(x, y)$ in a neighborhood of the given point (x_0, y_0, z_0).
Compute $\frac{\partial g}{\partial x}$ and $\frac{\partial g}{\partial y}$.

(a) $x^3 + y^3 + z^3 - xyz - 1 = 0, \quad (x_0, y_0, z_0) = (0, 0, 1)$

(b) $\exp(z) - z^2 - x^2 - y^2 = 0, \quad (x_0, y_0, z_0) = (1, 0, 0)$

6.7 Compute the marginal rate of substitution of K for L for the following isoquant of the given production function, that is $\frac{dK}{dL}$:

$$F(K, L) = AK^{\alpha} L^{\beta} = F_0$$

6.8 Compute the derivative $\frac{dx_i}{dx_j}$ of the indifference curve of the utility function:

(a) $u(x_1,x_2) = \left(x_1^{\frac{1}{2}} + x_2^{\frac{1}{2}}\right)^2$

(b) $u(x_1,\dots,x_n) = \left(\sum_{i=1}^{n} x_i^{\frac{\theta-1}{\theta}}\right)^{\frac{\theta}{\theta-1}}$ $\qquad (\theta > 1)$

— Problems

6.9 Prove Lemma 6.2.

6.10 Derive Theorem 6.4 from Theorem 6.11.

HINT: Consider function $\mathbf{F}(\mathbf{x},\mathbf{y}) = \mathbf{f}(\mathbf{x}) - \mathbf{y} = 0$.

6.11 Does the inverse function theorem (Theorem 6.4) provide a necessary or a sufficient condition for the existence of a local inverse function or is the condition both necessary and sufficient?

If the condition is not necessary, give a counterexample.

HINT: Use a function $f: \mathbb{R} \to \mathbb{R}$.

If the condition is not sufficient, give a counterexample.

6.12 Let $\mathbf{f}: \mathbb{R}^n \to \mathbb{R}^n$ be a $\mathscr{C}^1$ function that has a local inverse function $\mathbf{f}^{-1}$ around some point $\mathbf{x}^0$. Show by means of the chain rule that

$$\left(\mathbf{f}^{-1}\right)'(\mathbf{y}^0) = \left(\mathbf{f}'(\mathbf{x}^0)\right)^{-1} \qquad \text{where } \mathbf{y}^0 = \mathbf{f}(\mathbf{x}^0).$$

HINT: Notice that $\mathbf{f}^{-1} \circ \mathbf{f} = \mathrm{id}$, where id denotes the identity function, i.e, $\mathrm{id}(\mathbf{x}) = \mathbf{x}$. Compute the derivatives on either side of the equation. Use the chain rule for the left hand side. What is the derivative of id?

7

Convex Functions

Is there a panoramic view over our entire function?

7.1 Convex Sets

Definition 7.1

Convex set. A set $D \subseteq \mathbb{R}^n$ is called **convex**, if each pair of points $\mathbf{x}, \mathbf{y} \in D$ can be joints by a line segment lying entirely in D, i.e., if

$$(1-t)\mathbf{x} + t\mathbf{y} \in D \quad \text{for all } \mathbf{x}, \mathbf{y} \in D \text{ and all } t \in [0,1].$$

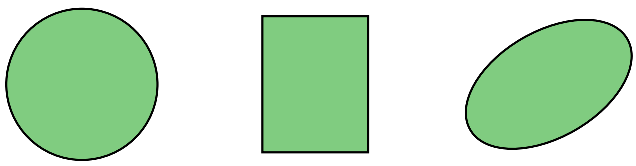
The line segment between $\mathbf{x}$ and $\mathbf{y}$ is the set

$$[\mathbf{x}, \mathbf{y}] = \left\{\mathbf{z} = (1-t)\mathbf{x} + t\mathbf{y} \colon t \in [0,1]\right\}.$$

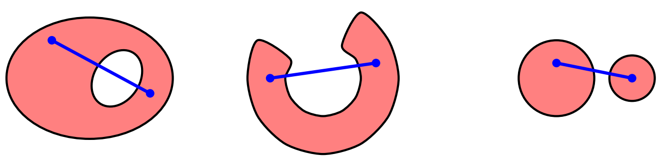
whose elements are so called **convex combinations** of $\mathbf{x}$ and $\mathbf{y}$. Hence $[\mathbf{x}, \mathbf{y}]$ is also called the **convex hull** of these points.

Example 7.2

The following sets are convex:



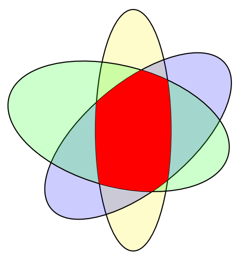
The following sets are not convex:



Theorem 7.3

Intersection. The intersection of convex sets is convex.

PROOF. See Problem 7.7. □



Notice that the union of convex need not be convex.

Example 7.4

Half spaces. Let $\mathbf{p} \in \mathbb{R}^n$, $\mathbf{p} \neq 0$, and $m \in \mathbb{R}$. Then the set

$$H = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{p}' \cdot \mathbf{x} = m\}$$

is called a **hyperplane** in $\mathbb{R}^n$. It divides $\mathbb{R}^n$ into two so called **half spaces**

$$H_+ = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{p}' \cdot \mathbf{x} \geq m\} \quad \text{and} \quad H_- = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{p}' \cdot \mathbf{x} \leq m\}\,.$$

All sets H, H^+, and H^- are convex, see Problem 7.8. ♢

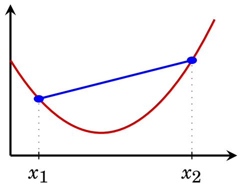
7.2 Convex and Concave Functions

Definition 7.5

Convex and concave function. A function $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ is **convex** if D is convex and

$$f\big((1-t)\mathbf{x}_1 + t\,\mathbf{x}_2\big) \leq (1-t)\,f(\mathbf{x}_1) + t\,f(\mathbf{x}_2)$$

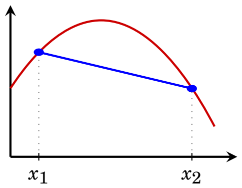
for all $\mathbf{x}_1, \mathbf{x}_2 \in D$ and all $t \in [0,1]$. This is equivalent to the property that the set $\big\{(\mathbf{x}, y) \in \mathbb{R}^{n+1} : y \geq f(\mathbf{x})\big\}$ is convex.



Function f is **concave** if D is convex and

$$f\big((1-t)\mathbf{x}_1 + t\,\mathbf{x}_2\big) \geq (1-t)\,f(\mathbf{x}_1) + t\,f(\mathbf{x}_2)$$

for all $\mathbf{x}_1, \mathbf{x}_2 \in D$ and all $t \in [0,1]$.



Notice that a function f is concave if and only if $-f$ is convex, see Problem 7.9.

Definition 7.6

Strictly convex function. A function $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ is **strictly convex** if D is convex and

$$f\big((1-t)\mathbf{x}_1 + t\,\mathbf{x}_2\big) < (1-t)\,f(\mathbf{x}_1) + t\,f(\mathbf{x}_2)$$

for all $\mathbf{x}_1, \mathbf{x}_2 \in D$ with $\mathbf{x}_1 \neq \mathbf{x}_2$ and all $t \in (0,1)$. Function f is **strictly concave** if this equation holds with "<" replaced by ">".

Example 7.7

Linear function. Let $\mathbf{a} \in \mathbb{R}^n$ be constant. Then $f(\mathbf{x}) = \mathbf{a}' \cdot \mathbf{x}$ is both convex and concave:

$$\begin{aligned} f\big((1-t)\mathbf{x}_1 + t\,\mathbf{x}_2\big) &= \mathbf{a}' \cdot \big((1-t)\mathbf{x}_1 + t\,\mathbf{x}_2\big) = (1-t)\,\mathbf{a}' \cdot \mathbf{x}_1 + t\,\mathbf{a}' \cdot \mathbf{x}_2 \\ &= (1-t)\,f(\mathbf{x}_1) + t\,f(\mathbf{x}_2) \end{aligned}$$

However, it is neither strictly convex nor strictly concave. ♢

Example 7.8

Quadratic function. Function $f(x) = x^2$ is strictly convex:

$$\begin{aligned} & f((1-t)x + t\,y) - \big[(1-t)\,f(x) + t\,f(y)\big] \\ &\qquad = \big((1-t)x + t\,y\big)^2 - \big[(1-t)x^2 + t\,y^2\big] \\ &\qquad = (1-t)^2 x^2 + 2(1-t)t\,xy + t^2 y^2 - (1-t)x^2 - t\,y^2 \\ &\qquad = -t(1-t)x^2 + 2(1-t)t\,xy - t(1-t)\,y^2 \\ &\qquad = -t(1-t)(x-y)^2 < 0 \end{aligned}$$

for $x \neq y$ and $0 < t < 1$. ♢

Theorem 7.9

Convex sum. Let $\alpha_1,\dots,\alpha_k > 0$. If $f_1(\mathbf{x}),\dots,f_k(\mathbf{x})$ are convex (concave) functions, then

$$g(\mathbf{x}) = \sum_{i=1}^{k} \alpha_i f_i(\mathbf{x})$$

is convex (concave). Function $g(\mathbf{x})$ is strictly convex (strictly concave) if at least one of the functions $f_i(\mathbf{x})$ is strictly convex (strictly concave).

PROOF. See Problem 7.12. □

An immediate consequence of this theorem and Example 7.8 is that a quadratic function $f(x) = ax^2 + bx + c$ is strictly convex if $a > 0$, strictly concave if $a < 0$ and both convex and concave if $a = 0$.

Theorem 7.10

Quadratic form. Let $\mathbf{A}$ be a symmetric $n \times n$ matrix. Then quadratic form $q(\mathbf{x}) = \mathbf{x}'\mathbf{A}\mathbf{x}$ is strictly convex if and only if $\mathbf{A}$ is positive definite. It is convex if and only if $\mathbf{A}$ is positive semidefinite.

Similarly, q is strictly concave if and only if $\mathbf{A}$ is negative definite. It is concave if and only if $\mathbf{A}$ is negative semidefinite.

PROOF IDEA. We first show by a straightforward computation that the univariate function $g(t) = q\big((1-t)\mathbf{x}_1 + t\mathbf{x}_2\big)$ is strictly convex for all $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^n$ if and only if $\mathbf{A}$ is positive definite.

PROOF. Let $\mathbf{x}_1$ and $\mathbf{x}_2$ be two distinct points in $\mathbb{R}^n$. Then

$$\begin{aligned}
g(t) &= q\big((1-t)\mathbf{x}_1 + t\mathbf{x}_2\big) = q\big(\mathbf{x}_1 + t(\mathbf{x}_2 - \mathbf{x}_1)\big) \\
&= \big(\mathbf{x}_1 + t(\mathbf{x}_2 - \mathbf{x}_1)\big)'\mathbf{A}\big(\mathbf{x}_1 + t(\mathbf{x}_2 - \mathbf{x}_1)\big) \\
&= t^2(\mathbf{x}_2 - \mathbf{x}_1)'\mathbf{A}(\mathbf{x}_2 - \mathbf{x}_1) + 2t(\mathbf{x}_1'\mathbf{A}\mathbf{x}_2 - \mathbf{x}_1'\mathbf{A}\mathbf{x}_1) + \mathbf{x}_1'\mathbf{A}\mathbf{x}_1 \\
&= q(\mathbf{x}_1 - \mathbf{x}_2)t^2 + 2(\mathbf{x}_1'\mathbf{A}\mathbf{x}_2 - q(\mathbf{x}_1))t + q(\mathbf{x}_1)
\end{aligned}$$

is a quadratic function in t which is strictly convex if and only if $q(\mathbf{x}_1 - \mathbf{x}_2) > 0$. This is the case for each pair of points $\mathbf{x}_1$ and $\mathbf{x}_2$ if and only if $\mathbf{A}$ is positive definite. We then find

$$\begin{aligned}
q\big((1-t)\mathbf{x}_1 + t\mathbf{x}_2\big) &= g(t) = g\big((1-t)0 + t\,1\big) \\
&> (1-t)g(0) + tg(1) = (1-t)q(\mathbf{x}_1) + tq(\mathbf{x}_2)
\end{aligned}$$

for all $t \in (0,1)$ and hence q is strictly convex as well. The cases where q is convex and (strictly) concave follow analogously. □

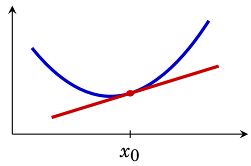
Recall from Linear Algebra that we can determine the definiteness of a symmetric matrix $\mathbf{A}$ by means of the signs of its eigenvalues or by the signs of (leading) principle minors.

Theorem 7.11

Tangents of convex functions. A $\mathscr{C}^1$ function f is convex in an open, convex set D if and only if

$$f(\mathbf{x}) - f(\mathbf{x}_0) \geq \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) \tag{7.1}$$

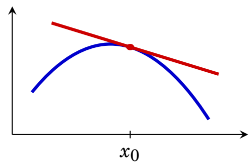
for all $\mathbf{x}$ and $\mathbf{x}_0$ in D, i.e., the tangent is always below the function. Function f is strictly convex if and only if inequality (7.1) is strict for $\mathbf{x} \neq \mathbf{x}_0$.



A $\mathscr{C}^1$ function f is concave in an open, convex set D if and only if

$$f(\mathbf{x}) - f(\mathbf{x}_0) \leq \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0)$$

for all $\mathbf{x}$ and $\mathbf{x}_0$ in D, i.e., the tangent is always above the function.



PROOF IDEA. For the necessity of condition (7.1) we transform the inequality for convexity (see Definition 7.5) into an inequality about difference quotients and apply the Mean Value Theorem. Using continuity of the gradient of f yields inequality (7.1).

We note here that for the case of strict convexity we need some technical trick to obtain the requested strict inequality.

For sufficiency we split an interval $[\mathbf{x}_0, \mathbf{x}]$ into two subintervals $[\mathbf{x}_0, \mathbf{z}]$ and $[\mathbf{z}, \mathbf{x}]$ and apply inequality (7.1) on each.

PROOF. Assume that f is convex, and let $\mathbf{x}_0, \mathbf{x} \in D$. Then we have by definition

$$f\big((1-t)\mathbf{x}_0 + t\mathbf{x}\big) \leq (1-t) f(\mathbf{x}_0) + t f(\mathbf{x})$$

and thus

$$f(\mathbf{x}) - f(\mathbf{x}_0) \geq \frac{f\big(\mathbf{x}_0 + t(\mathbf{x} - \mathbf{x}_0)\big) - f(\mathbf{x}_0)}{t} = \nabla f(\xi(t)) \cdot (\mathbf{x} - \mathbf{x}_0)$$

by the mean value theorem (Theorem 4.10) where $\xi(t) \in [\mathbf{x}_0, \mathbf{x}_0 + t(\mathbf{x} - \mathbf{x}_0)]$. (Notice that the central term is the difference quotient corresponding to the directional derivative.) Since f is a $\mathscr{C}^1$ function we find

$$f(\mathbf{x}) - f(\mathbf{x}_0) \geq \lim_{t \to 0} \nabla f(\xi(t)) \cdot (\mathbf{x} - \mathbf{x}_0) = \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0)$$

as claimed.

Conversely assume that (7.1) holds for all $\mathbf{x}_0, \mathbf{x} \in D$. Let $t \in [0,1]$ and $\mathbf{z} = (1-t)\mathbf{x}_0 + t\mathbf{x}$. Then $\mathbf{z} \in D$ and by (7.1) we find

$$\begin{aligned}(1-t)\big(f(\mathbf{x}_0) - f(\mathbf{z})\big) + t\big(f(\mathbf{x}) - f(\mathbf{z})\big)& \\ \geq (1-t)\, \nabla f(\mathbf{z})(\mathbf{x}_0 - \mathbf{z}) + t\, \nabla f(\mathbf{z})(\mathbf{x} - \mathbf{z})& \\ = \nabla f(\mathbf{z})\big((1-t)\mathbf{x}_0 + t\mathbf{x} - \mathbf{z}\big) = \nabla f(\mathbf{z})\, 0 = 0\,.&\end{aligned}$$

Consequently,

$$(1-t) f(\mathbf{x}_0) + t f(\mathbf{x}) \geq f(\mathbf{z}) = f\big((1-t)\mathbf{x}_0 + t\mathbf{x}\big)$$

and thus f is convex.

The proof for the case where f is strictly convex is analogous. However, in the first part of the proof $f(\mathbf{x}) - f(\mathbf{x}_0) > \nabla f(\xi(t)) \cdot (\mathbf{x} - \mathbf{x}_0)$ does not imply strict inequality in

$$f(\mathbf{x}) - f(\mathbf{x}_0) \geq \lim_{t \to 0} \nabla f(\xi(t)) \cdot (\mathbf{x} - \mathbf{x}_0).$$

So we need a technical trick. Assume $\mathbf{x} \neq \mathbf{x}_0$ and let $\mathbf{x}_1 = (\mathbf{x} + \mathbf{x}_0)/2$. By strict convexity of $\mathbf{f}$ we have $f(\mathbf{x}_1) < \frac{1}{2}(f(\mathbf{x}) + f(\mathbf{x}_0))$. Hence we find

$$2(\mathbf{x}_1 - \mathbf{x}_0) = \mathbf{x} - \mathbf{x}_0 \quad \text{and} \quad 2(f(\mathbf{x}_1) - f(\mathbf{x}_0)) < f(\mathbf{x}) - f(\mathbf{x}_0)$$

and thus

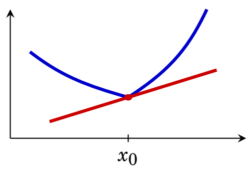
$$f(\mathbf{x}) - f(\mathbf{x}_0) > 2(f(\mathbf{x}_1) - f(\mathbf{x}_0)) \geq 2\nabla f(\mathbf{x}_0) \cdot (\mathbf{x}_1 - \mathbf{x}_0) = \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0)$$

as claimed. □

There also exists a version for functions that are not necessarily differentiable.

Theorem 7.12

Subgradient and supergradient. Let f be a convex function on a convex set $D \subseteq \mathbb{R}^n$, and let $\mathbf{x}_0$ be an interior point of D. If f is convex, then there exists a vector $\mathbf{p}$ such that



$$f(\mathbf{x}) - f(\mathbf{x}_0) \geq \mathbf{p}' \cdot (\mathbf{x} - \mathbf{x}_0) \quad \text{for all } \mathbf{x} \in D.$$

If f is a concave function on D, then there exists a vector $\mathbf{q}$ such that

$$f(\mathbf{x}) - f(\mathbf{x}_0) \leq \mathbf{q}' \cdot (\mathbf{x} - \mathbf{x}_0) \quad \text{for all } \mathbf{x} \in D.$$

The vectors $\mathbf{p}$ and $\mathbf{q}$ are called **subgradient** and **supergradient**, resp., of f at $\mathbf{x}_0$.

We omit the proof and refer the interested reader to [2, Sect. 2.4].

Theorem 7.13

Jensen's inequality, discrete version. A function f on a convex domain $D \subseteq \mathbb{R}^n$ is concave if and only if

$$f\left(\sum_{i=1}^{k} \alpha_i \mathbf{x}_i\right) \geq \sum_{i=1}^{k} \alpha_i f(\mathbf{x}_i)$$

for all $\mathbf{x}_i \in D$ and $\alpha_i \geq 0$ with $\sum_{i=1}^{k} \alpha_i = 1$.

PROOF. See Problem 7.13.

We finish with a quite obvious proposition.

Definition 7.14

Restriction of a function. Let $f: D \subseteq \mathbb{R}^n \to \mathbb{R}^m$ be some function and $S \subset D$. Then the function $f|_S : S \to \mathbb{R}^m$ defined by $f|_S(\mathbf{x}) = f(\mathbf{x})$ for all $\mathbf{x} \in S$ is called the **restriction** of f to S.

Lemma 7.15

Let f be a (strictly) convex function on a convex set $D \subset \mathbb{R}^n$ and $S \subset D$ a convex subset. Then $f|_S$ is (strictly) convex.

We close this section with a few useful results.

Lemma 7.16

Function $f(\mathbf{x})$ is convex if and only if $\{(\mathbf{x}, y) : \mathbf{x} \in D_f, f(\mathbf{x}) \le y\}$ is convex. Function $f(\mathbf{x})$ is concave if and only if $\{(\mathbf{x}, y) : \mathbf{x} \in D_f, f(\mathbf{x}) \ge y\}$ is convex.

PROOF. Observe that $\{(\mathbf{x}, y) : \mathbf{x} \in D_f, f(\mathbf{x}) \le y\}$ is the region above the graph of f. Thus the result follows from Definition 7.5. □

Lemma 7.17

Minimum and maximum of two convex functions.

(a) If $f(\mathbf{x})$ and $g(\mathbf{x})$ are concave, then $\min\{f(\mathbf{x}), g(\mathbf{x})\}$ is concave.

(b) If $f(\mathbf{x})$ and $g(\mathbf{x})$ are convex, then $\max\{f(\mathbf{x}), g(\mathbf{x})\}$ is convex.

PROOF. See Problem 7.14.

Theorem 7.18

Composite functions. Suppose that $f : D_f \subseteq \mathbb{R}^n \to \mathbb{R}$ and $F : D_F \subseteq \mathbb{R} \to \mathbb{R}$ are two functions such that $f(D_f) \subseteq D_F$. Then the following holds:

(a) If $f(\mathbf{x})$ is concave and $F(u)$ is concave and increasing, then $G(\mathbf{x}) = F(f(\mathbf{x}))$ is concave.

(b) If $f(\mathbf{x})$ is convex and $F(u)$ is convex and increasing, then $G(\mathbf{x}) = F(f(\mathbf{x}))$ is convex.

(c) If $f(\mathbf{x})$ is concave and $F(u)$ is convex and decreasing, then $G(\mathbf{x}) = F(f(\mathbf{x}))$ is convex.

(d) If $f(\mathbf{x})$ is convex and $F(u)$ is concave and decreasing, then $G(\mathbf{x}) = F(f(\mathbf{x}))$ is concave.

PROOF. We only show (a). Assume that $f(\mathbf{x})$ is concave and $F(u)$ is concave and increasing. Then a straightforward computation gives

$$\begin{aligned} G\big((1-t)\mathbf{x} + t\mathbf{y}\big) &= F\big(f((1-t)\mathbf{x} + t\mathbf{y})\big) \ge F\big((1-t)f(\mathbf{x}) + tf(\mathbf{y})\big) \\ &\ge (1-t)F(f(\mathbf{x})) + tF(f(\mathbf{y})) = (1-t)G(\mathbf{x}) + tG(\mathbf{y}) \end{aligned}$$

where the first inequality follows from the concavity of f and the monotonicity of F. The second inequality is implied by the concavity of F. □

7.3 Monotone Univariate Functions

We now want to use derivatives to investigate the convexity or concavity of a given function. We start with univariate functions and look at the simpler case of monotonicity.

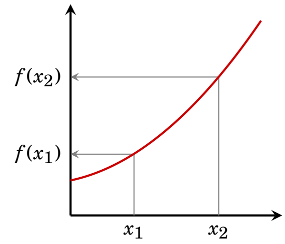
Definition 7.19

Monotone function. A function $f: D \subseteq \mathbb{R} \to \mathbb{R}$ is called **monotonically increasing [monotonically decreasing]** if

$$x_1 \le x_2 \quad \Rightarrow \quad f(x_1) \le f(x_2) \qquad [f(x_1) \ge f(x_2)] .$$

It is called **strictly increasing [strictly decreasing]** if

$$x_1 < x_2 \quad \Rightarrow \quad f(x_1) < f(x_2) \qquad [f(x_1) > f(x_2)] .$$



Notice that a function f is (strictly) monotonically decreasing if and only if $-f$ is (strictly) monotonically increasing. Moreover, the implication in Definition 7.19 can be replaced by an equivalence relation.

Lemma 7.20

A function $f: D \subseteq \mathbb{R} \to \mathbb{R}$ is [strictly] monotonically increasing if and only if

$$x_1 \le x_2 \quad \Leftrightarrow \quad f(x_1) \le f(x_2) \qquad [f(x_1) < f(x_2)] .$$

For a $\mathscr{C}^1$ function f we can use its derivative to verify monotonicity.

Theorem 7.21

Monotonicity and derivatives. Let $f: D \subseteq \mathbb{R} \to \mathbb{R}$ be a $\mathscr{C}^1$ function. Then the following holds.

(1) f is monotonically increasing on its domain D if and only if $f'(x) \ge 0$ for all $x \in D$.

(2) f is strictly increasing if $f'(x) > 0$ for all $x \in D$.

(3) If $f'(x_0) > 0$ for some $x_0 \in D$, then f is strictly increasing in an open neighborhood of x_0.

These statements holds analogously for decreasing functions.

Notice that (2) is a sufficient but not a necessary condition for strict monotonicity, see Problem 7.15.

Condition (2) can be replaced by a weaker condition that we state without proof:

(2') f is strictly increasing if $f'(x) > 0$ for almost all $x \in D$ (i.e., for all but a finite or countable number of points).

PROOF. (1) Assume that $f'(x) \ge 0$ for all $x \in D$. Let $x_1, x_2 \in D$ with $x_1 < x_2$. Then by the mean value theorem (Theorem 4.10) there exists a $\xi \in [x_1, x_2]$ such that

$$f(x_2) - f(x_1) = f'(\xi)(x_2 - x_1) \ge 0 .$$

Hence $f(x_1) \le f(x_2)$ and thus f is monotonically increasing. Conversely, if $f(x_1) \le f(x_2)$ for all $x_1, x_2 \in D$ with $x_1 < x_2$, then

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} \ge 0 \quad \text{and thus} \quad f'(x_1) = \lim_{x_2 \to x_1} \frac{f(x_2) - f(x_1)}{x_2 - x_1} \ge 0$$

for all $x_1 \in D$. For the proof of (2) and (3) see Problem 7.15. $\square$

7.4 Convexity of $\mathscr{C}^2$ Functions

For univariate $\mathscr{C}^2$ functions we can use the second derivative to verify convexity of the function, similar to Theorem 7.21.

Theorem 7.22

Convexity of univariate functions. Let $f: D \subseteq \mathbb{R} \to \mathbb{R}$ be a $\mathscr{C}^2$ function on an open interval $D \subseteq \mathbb{R}$. Then f is convex [concave] in D if and only if $f''(x) \geq 0$ $\left[\, f''(x) \leq 0 \right]$ for all $x \in D$.

PROOF IDEA. In order to verify the necessity of the condition we apply Theorem 7.11 to show that f' is increasing. Thus $f''(x) \geq 0$ by Theorem 7.21.

The sufficiency of the condition immediately follows from the Lagrange form of the reminder of the Taylor polynomial similar to the proof of Theorem 7.21.

PROOF. Assume that f is convex. Then Theorem 7.11 implies

$$f'(x_1) \leq \frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq f'(x_2)$$

for all $x_1, x_2 \in D$ with $x_1 < x_2$. Hence f' is monotonically increasing and thus $f''(x) \geq 0$ for all $x \in D$ by Theorem 7.21, as claimed.

Conversely, assume that $f''(x) \geq 0$ for all $x \in D$. Then the Lagrange's form of the remainder of the first order Taylor series (Theorem 5.5) gives

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(\xi)}{2}(x - x_0)^2 \geq f(x_0) + f'(x_0)(x - x_0)$$

and thus

$$f(x) - f(x_0) \geq f'(x_0)(x - x_0)\,.$$

Hence f is convex by Theorem 7.11. □

Similarly, we obtain a sufficient condition for strict convexity.

Theorem 7.23

Let $f: D \subseteq \mathbb{R} \to \mathbb{R}$ be a $\mathscr{C}^2$ function on an open interval $D \subseteq \mathbb{R}$. If $f''(x_0) > 0$ for some $x_0 \in D$, then f is strictly convex in an open neighborhood of x_0.

PROOF. Since f is a $\mathscr{C}^2$ function there exists an open ball $B_\varepsilon(x_0)$ such that $f''(x) > 0$ for all $x \in B_\varepsilon(x_0)$. Using the same argument as for Theorem 7.22 the statement follows. □

These results can be generalized for multivariate functions.

Theorem 7.24

Convexity of multivariate functions. A $\mathscr{C}^2$ function is convex (concave) on a convex, open set $D \subseteq \mathbb{R}^n$ if and only if the Hessian matrix $f''(\mathbf{x})$ is positive (negative) semidefinite for each $\mathbf{x} \in D$.

PROOF IDEA. We reduce the convexity of f to the convexity of all univariate reductions of f and apply Theorems 7.22 and 7.10.

PROOF. Let $\mathbf{x}, \mathbf{x}_0 \in D$ and $t \in [0,1]$. Define

$$g(t) = f\big((1-t)\mathbf{x}_0 + t\mathbf{x}\big) = f\big(\mathbf{x}_0 + t(\mathbf{x}-\mathbf{x}_0)\big)\,.$$

If g is convex for all $\mathbf{x}, \mathbf{x}_0 \in D$ and $t \in [0,1]$, then

$$\begin{aligned} f\big((1-t)\mathbf{x}_0 + t\mathbf{x}\big) &= g(t) = g\big((1-t)\cdot 0 + t\cdot 1\big) \\ &\le (1-t)g(0) + tg(1) = (1-t)f(\mathbf{x}_0) + tf(\mathbf{x}) \end{aligned}$$

i.e., f is convex. Similarly, if f is convex then g is convex. Applying the chain rule twice gives

$$\begin{aligned} g'(t) &= \nabla f\big(\mathbf{x}_0 + t(\mathbf{x}-\mathbf{x}_0)\big)\cdot(\mathbf{x}-\mathbf{x}_0)\,, \quad \text{and} \\ g''(t) &= (\mathbf{x}-\mathbf{x}_0)' f''\big(\mathbf{x}_0 + t(\mathbf{x}-\mathbf{x}_0)\big)\cdot(\mathbf{x}-\mathbf{x}_0)\,. \end{aligned}$$

By Theorem 7.22, g is convex if and only if $g''(t) \ge 0$ for all t. The latter is the case for all $\mathbf{x}, \mathbf{x}_0 \in D$ if and only if $f''(\mathbf{x})$ is positive semidefinite for each $\mathbf{x} \in D$ by Theorem 7.10. □

By a similar argument we find the multivariate extension of Theorem 7.23.

Theorem 7.25 Let f be a $\mathscr{C}^2$ function on a convex, open set $D \subseteq \mathbb{R}^n$ and $\mathbf{x}_0 \in D$. If $f''(\mathbf{x}_0)$ is positive (negative) definite, then f is strictly convex (strictly concave) in an open ball $B_\varepsilon(\mathbf{x}_0)$ centered at $\mathbf{x}_0$.

PROOF IDEA. Completely analogous to the proof of Theorem 7.24 except that we replace inequalities by strict inequalities. □

We can combine the results from Theorems 7.24 and 7.25 and our results from Linear Algebra as following. Let f be a $\mathscr{C}^2$ function on a convex, open set $D \subseteq \mathbb{R}^n$ and $\mathbf{x}_0 \in D$. Let $H_r(\mathbf{x})$ denotes the rth leading principle minor of $f''(\mathbf{x})$ then we find

- $H_r(\mathbf{x}_0) > 0$ for all $r \implies f$ is strictly convex in some open ball $B_\varepsilon(\mathbf{x}_0)$.
- $(-1)^r H_r(\mathbf{x}_0) > 0$ for all $r \implies f$ is strictly concave in $B_\varepsilon(\mathbf{x}_0)$.

The condition for semidefiniteness requires evaluations of all principle minors. Let $M_{i_1,\dots,i_r}$ denote a generic principle minor of order r of $f''(\mathbf{x})$. Then we have the following sufficient condition:

- $M_{i_1,\dots,i_r} \ge 0$ for all $\mathbf{x} \in D$ and all $i_1 < \dots i_r$ for $r = 1,\dots,n$
  $\iff f$ is convex in D.
- $(-1)^r M_{i_1,\dots,i_r} \ge 0$ for all $\mathbf{x} \in D$ and all $i_1 < \dots i_r$ for $r = 1,\dots,n$
  $\iff f$ is concave in D.

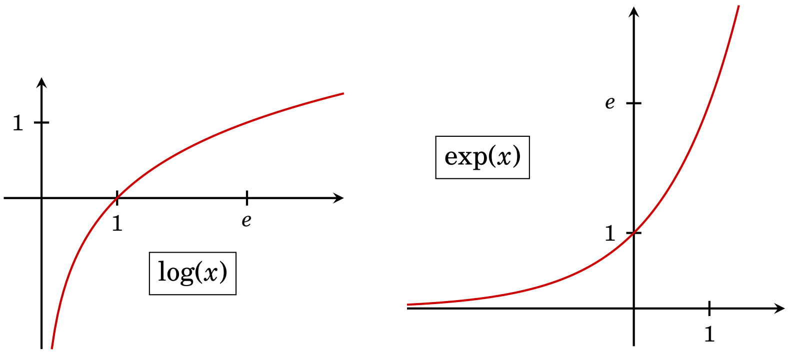
Logarithm and exponential function. The logarithm function — Example 7.26

$$\log\colon D = (0,\infty) \to \mathbb{R},\ x \mapsto \log(x)$$

is strictly concave as its second derivative $(\log(x))'' = -\frac{1}{x} < 0$ is negative for all $x \in D$. The exponential function

$$\exp\colon D = \mathbb{R} \to (0,\infty),\ x \mapsto e^x$$

is a strictly convex as its second derivative $(\exp(x))'' = e^x > 0$ is positive for all $x \in D$. ♢



Function $f(x,y) = x^4 + x^2 - 2xy + y^2$ is strictly convex in $D = \mathbb{R}^2$. — Example 7.27

SOLUTION. Its Hessian matrix is

$$f''(x,y) = \begin{pmatrix} 12x^2+2 & -2 \\ -2 & 2 \end{pmatrix}$$

with leading principle minors $H_1 = 12x^2 + 2 > 0$ and $H_2 = |f''(x,y)| = 24x^2 \geq 0$. Observe that both are positive on $D_0 = \{(x,y)\colon x \neq 0\}$. Hence f is strictly convex on D_0. Since f is a $\mathscr{C}^2$ function and the closure of D_0 is $\overline{D}_0 = D$ we can conclude that f is convex on D. ♢

Cobb-Douglas function. The Cobb-Douglas function — Example 7.28

$$f\colon D = (0,\infty)^2 \to \mathbb{R}, (x,y) \mapsto f(x,y) = x^\alpha y^\beta$$

with $\alpha,\beta \geq 0$ and $\alpha + \beta \leq 1$ is concave.

SOLUTION. The Hessian matrix at (x,y) and its principle minors are

$$f''(x,y) = \begin{pmatrix} \alpha(\alpha-1)x^{\alpha-2}y^\beta & \alpha\beta x^{\alpha-1}y^{\beta-1} \\ \alpha\beta x^{\alpha-1}y^{\beta-1} & \beta(\beta-1)x^\alpha y^{\beta-2} \end{pmatrix},$$

$$M_1 = \alpha(\alpha-1)x^{\alpha-2}y^\beta \leq 0\,,$$

$$M_2 = \beta(\beta-1)x^\alpha y^{\beta-2} \leq 0\,,$$

$$M_{1,2} = \alpha\beta(1-\alpha-\beta)x^{2\alpha-2}y^{2\beta-2} \geq 0\,.$$

The Cobb-Douglas function is strict concave if $\alpha,\beta > 0$ and $\alpha + \beta < 1$. ♢

7.5 Quasi-Convex Functions

Convex and concave functions play a prominent rôle in static optimization. However, in many theorems *convexity* and *concavity* can be replaced by weaker conditions. In this section we introduce a notion that is based on level sets.

Definition 7.29

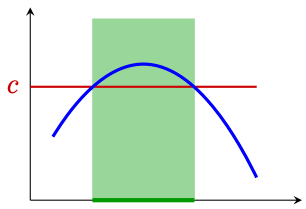
Level set. The set

$$U_c = \{\mathbf{x} \in D : f(\mathbf{x}) \geq c\} = f^{-1}\big([c,\infty)\big)$$

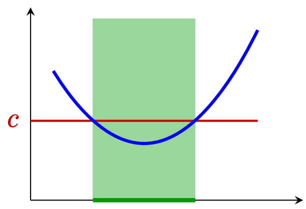
is called a **upper level set** of f. The set

$$L_c = \{\mathbf{x} \in D : f(\mathbf{x}) \leq c\} = f^{-1}\big((-\infty,c]\big)$$

is called a **lower level set** of f.



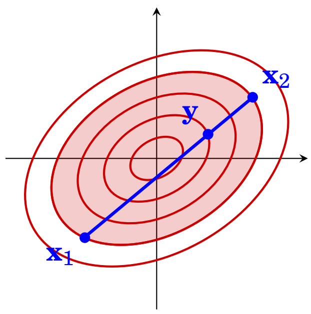
upper level set



lower level set

Lemma 7.30

Level sets of convex functions. Let $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ be a convex function and $c \in \mathbb{R}$. Then the lower level set $L_c = \{\mathbf{x} \in D : f(\mathbf{x}) \leq c\}$ is convex.



PROOF. Let $\mathbf{x}_1, \mathbf{x}_2 \in \{\mathbf{x} \in D : f(\mathbf{x}) \leq c\}$, i.e., $f(\mathbf{x}_1) \leq c$ and $f(\mathbf{x}_2) \leq c$. Then for every $\mathbf{y} = (1-t)\mathbf{x}_1 + t\mathbf{x}_2$ with $t \in [0,1]$ we find

$$f(\mathbf{y}) = f\big((1-t)\mathbf{x}_1 + t\mathbf{x}_2\big) \leq (1-t) f(\mathbf{x}_1) + t f(\mathbf{x}_2) \leq (1-t)c + tc = c$$

that is, $\mathbf{y} \in \{\mathbf{x} \in D : f(\mathbf{x}) \leq c\}$. Thus the lower level set $\{\mathbf{x} \in D : f(\mathbf{x}) \leq c\}$ is convex, as claimed. □

We will see in the next chapter that functions where all its lower level sets are convex behave in many situations similar to convex functions, that is, they are *quasi* convex. This motivates the following definition.

Definition 7.31

Quasi-convex. A function $f : D \subseteq \mathbb{R}^n \to \mathbb{R}$ is called **quasi-convex** if each of its lower level sets $L_c = \{\mathbf{x} \in D : f(\mathbf{x}) \leq c\}$ are convex.

Function f is called **quasi-concave** if each of its upper level sets $U_c = \{\mathbf{x} \in D : f(\mathbf{x}) \geq c\}$ are convex.

Analogously to Problem 7.9 we find that a function f is quasi-concave if and only if $-f$ is quasi-convex, see Problem 7.16.

Obviously every concave function is quasi-concave but not vice versa as the following examples shows.

Example 7.32

Function $f(x) = e^{-x^2}$ is quasi-concave but not concave. ♢

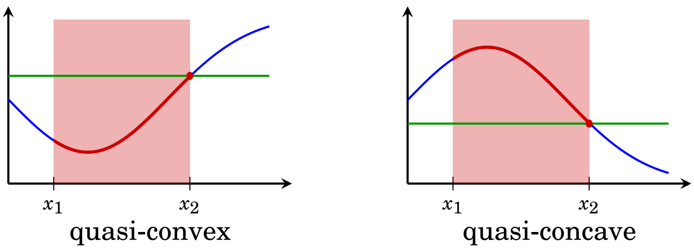
Characterization of quasi-convexity. A function f on a convex set $D \subseteq \mathbb{R}^n$ is quasi-convex if and only if Theorem 7.33

$$f\big((1-t)\mathbf{x}_1 + t\mathbf{x}_2\big) \leq \max\big\{f(\mathbf{x}_1), f(\mathbf{x}_2)\big\}$$

for all $\mathbf{x}_1, \mathbf{x}_2 \in D$ and $t \in [0,1]$. The function is quasi-concave if and only if

$$f\big((1-t)\mathbf{x}_1 + t\mathbf{x}_2\big) \geq \min\big\{f(\mathbf{x}_1), f(\mathbf{x}_2)\big\}$$

for all $\mathbf{x}_1, \mathbf{x}_2 \in D$ and $t \in [0,1]$.



PROOF IDEA. For $c = \max\big\{f(\mathbf{x}_1), f(\mathbf{x}_2)\big\}$ we find that

$$(1-t)\mathbf{x}_1 + t\mathbf{x}_2 \in L_c = \{\mathbf{x} \in D : f(\mathbf{x}) \leq c\}$$

is equivalent to

$$f\big((1-t)\mathbf{x}_1 + t\mathbf{x}_2\big) \leq c = \max\big\{f(\mathbf{x}_1), f(\mathbf{x}_2)\big\}\,.$$

PROOF. Let $\mathbf{x}_1, \mathbf{x}_2 \in D$ and $t \in [0,1]$. Let $c = \max\big\{f(\mathbf{x}_1), f(\mathbf{x}_2)\big\}$ and assume w.l.o.g. that $c = f(x_2) \geq f(x_1)$. If f is quasi-convex, then $(1-t)\mathbf{x}_1 + t\mathbf{x}_2 \in L_c = \{\mathbf{x} \in D : f(\mathbf{x}) \leq c\}$ and thus $f((1-t)\mathbf{x}_1 + t\mathbf{x}_2) \leq c = \max\big\{f(\mathbf{x}_1), f(\mathbf{x}_2)\big\}$. Conversely, if $f((1-t)\mathbf{x}_1 + t\mathbf{x}_2) \leq c = \max\big\{f(\mathbf{x}_1), f(\mathbf{x}_2)\big\}$, then $(1-t)\mathbf{x}_1 + t\mathbf{x}_2 \in L_c$ and thus f is quasi-convex. The case for quasi-concavity is shown analogously. □

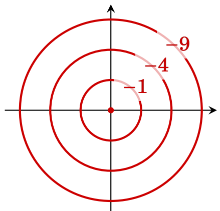
In Theorem 7.18 we have seen that some compositions of functions preserve convexity. Quasi-convexity is preserved under even milder condition.

Composite functions. Suppose that $f: D_f \subseteq \mathbb{R}^n \to \mathbb{R}$ and $F: D_F \subseteq \mathbb{R} \to \mathbb{R}$ are two functions such that $f(D_f) \subseteq D_F$. Then the following holds: Theorem 7.34

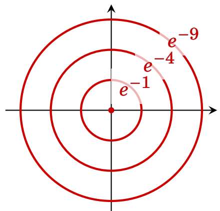
(a) If $f(\mathbf{x})$ is quasi-convex (quasi-concave) and $F(u)$ is increasing, then $G(\mathbf{x}) = F(f(\mathbf{x}))$ is quasi-convex (quasi-concave).

(b) If $f(\mathbf{x})$ is quasi-convex (quasi-concave) and $F(u)$ is decreasing, then $G(\mathbf{x}) = F(f(\mathbf{x}))$ is quasi-concave (quasi-convex).

PROOF IDEA. Monotone transformations preserve (in some sense) level sets of functions.



$f(x,y) = -x^2 - y^2$



$\exp(-x^2 - y^2)$

PROOF. Assume that f is quasi-convex and F is increasing. Thus by Theorem 7.33 $f\big((1-t)\mathbf{x}_1 + t\mathbf{x}_2\big) \le \max\{f(\mathbf{x}_1), f(\mathbf{x}_2)\}$ for all $\mathbf{x}_1, \mathbf{x}_2 \in D$ and $t \in [0,1]$. Moreover, $F(y_1) \le F(y_2)$ if and only if $y_1 \le y_2$. Hence we find for all $\mathbf{x}_1, \mathbf{x}_2 \in D$ and $t \in [0,1]$,

$$F\big(f((1-t)\mathbf{x}_1 + t\mathbf{x}_2)\big) \le F\big(\max\{f(\mathbf{x}_1), f(\mathbf{x}_2)\}\big) \le \max\{F(f(\mathbf{x}_1)), f(f(\mathbf{x}_2))\}$$

and thus $F \circ f$ is quasi-convex. The proof for the other cases is completely analogous. □

Theorem 7.34 allows to determine quasi-convexity or quasi-concavity of some functions. In Example 7.28 we have shown that the Cobb-Douglas function is concave for appropriate parameters. The computation was a bit tedious and it is not straightforward to extend the proof to functions of the form $\sum_{i=1}^{n} x_i^{\alpha_i}$. Quasi-concavity is much easier to show. Moreover, it holds for a larger range of parameters and our computation easily generalizes to many variables.

Example 7.35

Cobb-Douglas function. The Cobb-Douglas function

$$f\colon D = (0,\infty)^2 \to \mathbb{R}, (x,y) \mapsto f(x,y) = x^\alpha y^\beta$$

with $\alpha, \beta \ge 0$ is quasi-concave.

SOLUTION. Observe that $f(x,y) = \exp\big(\alpha\log(x) + \beta\log(y)\big)$. Notice that $\log(x)$ is concave by Example 7.26. Thus $\alpha\log(x) + \beta\log(y)$ is concave by Theorem 7.9 and hence quasi-concave. Since the exponential function exp is monotonically increasing, it follows that the Cobb-Douglas function is quasi-concave if $\alpha, \beta > 0$. ◇

Notice that it is not possible to apply Theorem 7.18 to show concavity of the Cobb-Douglas function when $\alpha + \beta \le 1$.

Example 7.36

CES function. Let $a_1, \dots, a_n \ge 0$. Then function

$$f(\mathbf{x}) = \left(\sum_{i=1}^{n} a_i x_i^r\right)^{1/r}$$

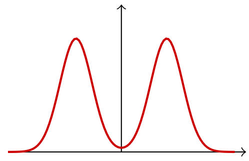
is quasi-concave for all $r \le 1$ and quasi-convex for all $r \ge 1$.

SOLUTION. Since $\big(x_i^r\big)'' = r(r-1)x_i^{r-2}$, we find that x_i^r is concave for $r \in [0,1]$ and convex otherwise. Hence the same holds for $\sum_{i=1}^{n} a_i x_i^r$ by Theorem 7.9. Since $F(y) = y^{1/r}$ is monotonically increasing if $r > 0$ and decreasing if $r < 0$, Theorem 7.34 implies that $f(\mathbf{x})$ is quasi-concave for all $r \le 1$ and quasi-convex for all $r \ge 1$. ◇

In opposition to Theorem 7.9 the sum of quasi-convex functions need not be quasi-convex.

Example 7.37

The two functions $f_1(x) = \exp\big(-(x-2)^2\big)$ and $f_2(x) = \exp\big(-(x+2)^2\big)$ are quasi-concave as each of their upper level sets are intervals (or empty). However, $f_1(x) + f_2(x)$ has two local maxima and thus cannot be quasi-concave.



There is also an analog to strict convexity. However, a definition using lower level set were not useful. So we start with the characterization of quasi-convexity in Theorem 7.33.

Definition 7.38

Strictly quasi-convex. A function f on a convex set $D \subseteq \mathbb{R}^n$ is called **strictly quasi-convex** if

$$f\big((1-t)\mathbf{x}_1 + t\mathbf{x}_2\big) < \max\big\{f(\mathbf{x}_1), f(\mathbf{x}_2)\big\}$$

for all $\mathbf{x}_1, \mathbf{x}_2$ with $\mathbf{x}_1 \neq \mathbf{x}_2$, and $t \in (0,1)$. It is called **strictly quasi-concave** if

$$f\big((1-t)\mathbf{x}_1 + t\mathbf{x}_2\big) > \min\big\{f(\mathbf{x}_1), f(\mathbf{x}_2)\big\}$$

for all $\mathbf{x}_1, \mathbf{x}_2$ with $\mathbf{x}_1 \neq \mathbf{x}_2$, and $t \in (0,1)$.

Our last result shows, that we also can use tangents to characterize quasi-convex function. Again, the condition is weaker than the corresponding condition in Theorem 7.11.

Theorem 7.39

Tangents of quasi-convex functions. A $\mathscr{C}^1$ function f is quasi-convex in an open, convex set D if and only if

$$f(\mathbf{x}) \leq f(\mathbf{x}_0) \quad \text{implies} \quad \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) \leq 0$$

for all $\mathbf{x}, \mathbf{x}_0 \in D$. It is quasi-concave if and only if

$$f(\mathbf{x}) \geq f(\mathbf{x}_0) \quad \text{implies} \quad \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) \geq 0$$

for all $\mathbf{x}, \mathbf{x}_0 \in D$.

PROOF. Assume that f is quasi-convex and $f(\mathbf{x}) \leq f(\mathbf{x}_0)$. Define $g(t) = f\big((1-t)\mathbf{x}_0 + t\mathbf{x}\big) = f\big(\mathbf{x}_0 + t(\mathbf{x}-\mathbf{x}_0)\big)$. Then Theorem 7.33 implies that $g(0) = f(\mathbf{x}_0) \geq g(t)$ for all $t \in [0,1]$ and hence $g'(0) \leq 0$. By the chain rule we find $g'(t) = \nabla f\big(\mathbf{x}_0 + t(\mathbf{x}-\mathbf{x}_0)\big)(\mathbf{x}-\mathbf{x}_0)$ and consequently $g'(0) = \nabla f(\mathbf{x}_0)\cdot(\mathbf{x}-\mathbf{x}_0) \leq 0$ as claimed.

For the converse assume that f is not quasi-convex. Then there exist $\mathbf{x}, \mathbf{x}_0 \in D$ with $f(\mathbf{x}) \leq f(\mathbf{x}_0)$ and a $\mathbf{z} = \mathbf{x}_0 + t(\mathbf{x}-\mathbf{x}_0) \in D$ for some $t \in (0,1)$ such that $f(\mathbf{z}) > f(\mathbf{x}_0)$. Define $g(t)$ as above. Then $g(t) > g(0)$ and there exists a $\tau \in (0,1)$ such that $g'(\tau) > 0$ by the Mean Value Theorem, and thus $\nabla f(\mathbf{z}_0)\cdot(\mathbf{x}-\mathbf{z}_0) > 0$, where $\mathbf{z}_0 = g(\tau)$. We state without proof that we can find such a point $\mathbf{z}_0$ where $f(\mathbf{z}_0) \geq f(\mathbf{x})$. Thus the given condition is violated for some points.

The proof for the second statement is completely analogous. □

— Exercises

7.1 Let

$$f(x) = x^4 + \frac{4}{3}x^3 - 24x^2 + 8\,.$$

(a) Determinte the regions where f is monotonically increasing and monotonically decreasing, respectively.
(b) Determinte the regions where f is concave and convex, respectively.

7.2 A function $f: \mathbb{R} \to (0,\infty)$ is called *log-concave* if $\ln \circ f$ is a concave function.

Which of the following functions is log-concave?

(a) $f(x) = 3\exp(-x^4)$
(b) $g(x) = 4\exp(-x^7)$
(c) $h(x) = 2\exp(x^2)$
(d) $s: (-1,1) \to (0,\infty),\ x \mapsto s(x) = 1 - x^4$

7.3 Determine whether the following functions are convex, concave or neither.

(a) $f(x) = \exp\left(-\sqrt{x}\right)$ on $D = [0,\infty)$.
(b) $f(\mathbf{x}) = \exp\left(-\sum_{i=1}^{n} \sqrt{x_i}\right)$ on $D = [0,\infty)^n$.

HINT: Use Theorem 7.18.

7.4 Determine whether the following functions on $\mathbb{R}^2$ are (strictly) convex or (strictly) concave or neither.

(a) $f(x,y) = x^2 - 2xy + 2y^2 + 4x - 8$
(b) $g(x,y) = 2x^2 - 3xy + y^2 + 2x - 4y - 2$
(c) $h(x,y) = -x^2 + 4xy - 4y^2 + 1$

7.5 Show that function

$$f(x,y) = ax^2 + 2bxy + cy^2 + px + qy + r$$

is strictly concave if $ac - b^2 > 0$ and $a < 0$, and strictly convex if $ac - b^2 > 0$ and $a > 0$.

Find necessary and sufficient conditions for (strict) convexity/concavity of f.

7.6 Show that $f(x,y) = \exp(-x^2 - y^2)$ is quasi-concave in $D = \mathbb{R}^2$ but not concave.

Apply Theorem 7.34.

Is there a domain where f is (strictly) concave? Compute the largest of such domains.

— Problems

7.7 Let $S_1, \dots, S_k$ be convex sets in $\mathbb{R}^n$. Show that their intersection $\bigcap_{i=1}^{k} S_i$ is convex (Theorem 7.3).

Give an example where the union of convex sets is not convex.

7.8 Show that the sets H, H^+, and H^- in Example 7.4 are convex.

7.9 Show that a function $f: D \subseteq \mathbb{R}^n \to \mathbb{R}$ is (strictly) concave if and only if function $g: D \subseteq \mathbb{R}^n \to \mathbb{R}$ with $g(\mathbf{x}) = -f(\mathbf{x})$ is (strictly) convex.

7.10 A function $f: \mathbb{R} \to (0, \infty)$ is called *log-concave* if $\ln \circ f$ is a concave function.

Show that every concave function $f: \mathbb{R} \to (0, \infty)$ is log-concave.

7.11 Let $T: (0, \infty) \to \mathbb{R}$ be a strictly monotonically increasing twice differentiable transformation. A function $f: \mathbb{R} \to (0, \infty)$ is called *T-concave* if $T \circ f$ is a concave function.

Consider the family $T_c(x)$, $c \le 0$, of transformations with $T_0(x) = \ln(x)$ and $T_c(x) = -x^c$ for $c < 0$.

(a) Show that all transformations T_c satisfy the above conditions for all $c \le 0$.
(b) Show that $f(x) = \exp(-x^2)$ is $T_{-1/2}$-concave.
(c) Show that $f(x) = \exp(-x^2)$ is T_c-concave for all $c \le 0$.

7.12 Prove Theorem 7.9.

7.13 Prove Jensen's inequality (Theorem 7.13).

HINT: For $k = 2$ the theorem is equivalent to the definition of concavity. For $k \ge 3$ use induction.

7.14 Prove Lemma 7.17.

HINT: Use Lemma 7.16.

7.15 Prove (2) and (3) of Theorem 7.21.

Condition (2) (i.e., $f'(x) > 0$ for all $x \in D$) is sufficient for f being strictly monotonically increasing. Give a counterexample that shows that this condition is not necessary.

HINT: Give a strictly increasing function f where $f'(0) = 0$.

Suppose one wants to prove the (false!) statement that $f'(x) > 0$ for each $x \in D_f$ for every strictly increasing function f. Thus he or she uses the same argument as in the proof of Theorem 7.21(1). Where does this argument fail?

7.16 Show that a function $f: D \subseteq \mathbb{R}^n \to \mathbb{R}$ is (strictly) quasi-concave if and only if function $g: D \subseteq \mathbb{R}^n \to \mathbb{R}$ with $g(\mathbf{x}) = -f(\mathbf{x})$ is (strictly) quasi-convex.

8

Static Optimization

We want to find the highest peak in our world.

8.1 Extremal Points

We start with so called global extrema.

Extremal points. Let $f: D \subseteq \mathbb{R}^n \to \mathbb{R}$. Then $\mathbf{x}^* \in D$ is called a (global) **maximum** of f if — Definition 8.1

$$f(\mathbf{x}^*) \geq f(\mathbf{x}) \quad \text{for all } \mathbf{x} \in D.$$

It is called a **strict maximum** if the inequality is strict for $\mathbf{x} \neq \mathbf{x}^*$.

Similarly, $\mathbf{x}^* \in D$ is called a (global) **minimum** of f if $f(\mathbf{x}^*) \leq f(\mathbf{x})$ for all $\mathbf{x} \in D$.

A **stationary point** $\mathbf{x}_0$ of a function f is a point where the gradient vanishes, i.e, — Definition 8.2

$$\nabla f(\mathbf{x}_0) = 0\,.$$

Necessary first-order conditions. Let f be a $\mathscr{C}^1$ function on an open set $D \subseteq \mathbb{R}^n$ and let $\mathbf{x}^* \in D$ be an extremal point. Then $\mathbf{x}^*$ is a stationary point of f, i.e., — Theorem 8.3

$$\nabla f(\mathbf{x}^*) = 0\,.$$

PROOF. If $\mathbf{x}^*$ is an extremal point then all directional derivatives are 0 and thus the result follows. □

Sufficient conditions. Let f be a $\mathscr{C}^1$ function on an open set $D \subseteq \mathbb{R}^n$ and let $\mathbf{x}^* \in D$ be a stationary point of f. — Theorem 8.4

If f is (strictly) convex in D, then $\mathbf{x}^*$ is a (strict) minimum of f.

If f is (strictly) concave in D, then $\mathbf{x}^*$ is a (strict) maximum of f.

PROOF. Assume that f is strictly convex. Then by Theorem 7.11

$$f(\mathbf{x}) - f(\mathbf{x}^*) > \nabla f(\mathbf{x}^*) \cdot (\mathbf{x} - \mathbf{x}^*) = 0 \cdot (\mathbf{x} - \mathbf{x}^*) = 0$$

and hence $f(\mathbf{x}) > f(\mathbf{x}^*)$ for all $\mathbf{x} \neq \mathbf{x}^*$, as claimed. The other statements follow analogously. □

Example 8.5

Cobb-Douglas function. We want to find the (global) maxima of

$$f: D = [0, \infty)^2 \to \mathbb{R},\ f(x, y) = 4x^{\frac{1}{4}} y^{\frac{1}{4}} - x - y\,.$$

SOLUTION. A straightforward computation yields

$$f_x = x^{-\frac{3}{4}} y^{\frac{1}{4}} - 1$$
$$f_y = x^{\frac{1}{4}} y^{-\frac{3}{4}} - 1$$

and thus $\mathbf{x}_0 = (1, 1)$ is the only stationary point of this function. As f is strictly concave (see Example 7.28) $\mathbf{x}_0$ is the global maximum of f. ◊

Definition 8.6

Local extremal points. Let $f: D \subseteq \mathbb{R}^n \to \mathbb{R}$. Then $\mathbf{x}^* \in D$ is called a **local maximum** of f if there exists an $\varepsilon > 0$ such that

$$f(\mathbf{x}^*) \geq f(\mathbf{x}) \quad \text{for all } \mathbf{x} \in B_\varepsilon(\mathbf{x}^*).$$

It is called a **strict local maximum** if the inequality is strict for $\mathbf{x} \neq \mathbf{x}^*$.

Similarly, $\mathbf{x}^* \in D$ is called a **local minimum** of f if there exists an $\varepsilon > 0$ such that $f(\mathbf{x}^*) \leq f(\mathbf{x})$ for all $\mathbf{x} \in B_\varepsilon(\mathbf{x}^*)$.

Local extrema necessarily are stationary points.

Theorem 8.7

Sufficient conditions for local extremal points. Let f be a $\mathscr{C}^2$ function on an open set $D \subseteq \mathbb{R}^n$ and let $\mathbf{x}^* \in D$ be a stationary point of f.

If $f''(\mathbf{x}^*)$ is positive definite, then $\mathbf{x}^*$ is a strict local minimum of f.

If $f''(\mathbf{x}^*)$ is negative definite, then $\mathbf{x}^*$ is a strict local maximum of f.

PROOF. Assume that $f''(\mathbf{x}^*)$ is positive definite. Since f'' is continuous, there exists an ε such that $f''(\mathbf{x})$ is positive definite for all $\mathbf{x} \in B_\varepsilon(\mathbf{x}^*)$ and hence f is strictly convex in $B_\varepsilon(\mathbf{x}^*)$. Consequently, $\mathbf{x}^*$ is a strict minimum in $B_\varepsilon(\mathbf{x}^*)$ by Theorem 8.4, i.e., a strict local minimum of f. □

Example 8.8

We want to find all local maxima of

$$f(x, y) = \frac{1}{6}x^3 - x + \frac{1}{4}xy^2\,.$$

SOLUTION. The partial derivative of f are given as

$$f_x = \frac{1}{2}x^2 - 1 + \frac{1}{4}y^2,$$
$$f_y = \frac{1}{2}xy,$$

and hence we find the stationary points $\mathbf{x}_1 = (0,2)$, $\mathbf{x}_2 = (0,-2)$, $\mathbf{x}_3 = (\sqrt{2},0)$, and $\mathbf{x}_4 = (-\sqrt{2},0)$. In order to apply Theorem 7.25 we need the Hessian of f,

$$f''(x,y) = \begin{pmatrix} f_{xx}(\mathbf{x}) & f_{xy}(\mathbf{x}) \\ f_{yx}(\mathbf{x}) & f_{yy}(\mathbf{x}) \end{pmatrix} = \begin{pmatrix} x & \frac{1}{2}y \\ \frac{1}{2}y & \frac{1}{2}x \end{pmatrix}.$$

We then find $f''(\mathbf{x}_3) = \begin{pmatrix} \sqrt{2} & 0 \\ 0 & \frac{\sqrt{2}}{2} \end{pmatrix}$. Its leading principle minors are both positive, $H_1 = \sqrt{2} > 0$ and $H_2 = 1 > 0$, and hence $\mathbf{x}_3$ is a local minimum. Similarly we find that $\mathbf{x}_4$ is a local maximum. ◇

Besides (local) extrema there are also other types of stationary points.

Definition 8.9

Saddle point. Let f be a $\mathscr{C}^2$ function on an open set $D \subseteq \mathbb{R}^n$. A stationary point $\mathbf{x}_0 \in D$ is called a **saddle point** if $f''(\mathbf{x}_0)$ is indefinite, that is, if f is neither convex nor concave in any open ball around $\mathbf{x}^*$.

Example 8.10

In Example 8.8 we have found two additional stationary points: $\mathbf{x}_1 = (0,2)$ and $\mathbf{x}_2 = (0,-2)$. However, the Hessian of f at $\mathbf{x}_1$,

$$f''(\mathbf{x}_1) = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

is indefinite as it has leading principle minors $H_1 = 0$ and $H_2 = -1 < 0$. Consequently $\mathbf{x}_1$ is a saddle point. ◇

8.2 The Envelope Theorem

Let $f(\mathbf{x},\mathbf{r})$ be a $\mathscr{C}^1$ function with (endogenous) variable $\mathbf{x} \in D \subseteq \mathbb{R}^n$ and parameter (exogenous variable) $\mathbf{r} \in \mathbb{R}^k$. An extremal point of f may depend on $\mathbf{r}$. So let $\mathbf{x}^*(\mathbf{r})$ denote an extremal point for a given parameter $\mathbf{r}$ and let

$$f^*(\mathbf{r}) = \max_{\mathbf{x} \in D} f(\mathbf{x},\mathbf{r}) = f(\mathbf{x}^*(\mathbf{r}),\mathbf{r})$$

be the **value function**.

Theorem 8.11

Envelope theorem. Let $f(\mathbf{x},\mathbf{r})$ be a $\mathscr{C}^1$ function on $D \times \mathbb{R}^k$ where $D \subseteq \mathbb{R}^n$. Let $\mathbf{x}^*(\mathbf{r})$ denote an extremal point for a given parameter $\mathbf{r}$ and assume that $\mathbf{r} \mapsto \mathbf{x}^*(\mathbf{r})$ is differentiable. Then

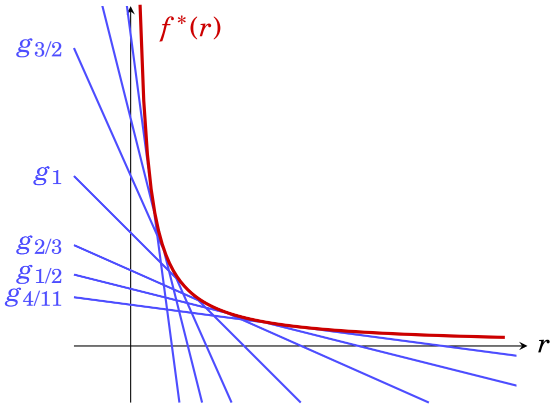
$$\frac{\partial f^*(\mathbf{r})}{\partial r_j} = \left.\frac{\partial f(\mathbf{x},\mathbf{r})}{\partial r_j}\right|_{\mathbf{x}=\mathbf{x}^*(\mathbf{r})}$$

PROOF IDEA. The chain rule implies

$$\begin{aligned} \frac{\partial f^*(\mathbf{r})}{\partial r_j} &= \frac{\partial f(\mathbf{x}^*(\mathbf{r}),\mathbf{r})}{\partial r_j} \\ &= \underbrace{\sum_{i=1}^n f_{x_i}(\mathbf{x}^*(\mathbf{r}),\mathbf{r})}_{=0} \cdot \frac{\partial x_i^*(\mathbf{r})}{\partial r_j} + \left.\frac{\partial f(\mathbf{x},\mathbf{r})}{\partial r_j}\right|_{\mathbf{x}=\mathbf{x}^*(\mathbf{r})} = \left.\frac{\partial f(\mathbf{x},\mathbf{r})}{\partial r_j}\right|_{\mathbf{x}=\mathbf{x}^*(\mathbf{r})} \end{aligned}$$

as claimed. □

The following figure illustrates this theorem. Let $f(x,r) = \sqrt{x} - rx$ and $f^*(r) = \max_x f(x,r)$. $g_x(r) = f(r,x)$ denotes function f with argument x fixed. Observe that $f^*(r) = \max_x g_x(r)$.



See Lecture 11 in *Mathematische Methoden* for further examples.

8.3 Constraint Optimization – The Lagrange Function

In this section we consider the optimization problem

$$\begin{aligned} &\max\ (\min) \quad f(x_1,\dots,x_n) \\ &\text{subject to} \quad g_j(x_1,\dots,x_n) = c_j, \quad j = 1,\dots,m \quad (m < n) \end{aligned}$$

or in vector notation

$$\max\ (\min) \quad f(\mathbf{x}) \quad \text{subject to} \quad \mathbf{g}(\mathbf{x}) = \mathbf{c}\,.$$

Definition 8.12

Lagrange function. Function

$$\mathscr{L}(\mathbf{x};\boldsymbol{\lambda}) = f(\mathbf{x}) - \boldsymbol{\lambda}'(\mathbf{g}(\mathbf{x}) - \mathbf{c}) = f(\mathbf{x}) - \sum_{j=1}^{m} \lambda_j (g_j(\mathbf{x}) - c_j)$$

is called the **Lagrange function** (or **Lagrangian**) of the above constraint optimization problem. The numbers λ_j are called **Lagrange multipliers** .

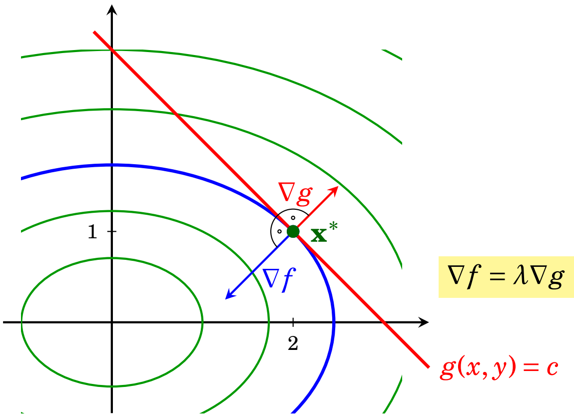
Lagrange multiplier

In order to find candidates for solutions of the constraint optimization problem we have to find stationary points of the Lagrange function. We state this condition without a proof.

Theorem 8.13

Necessary condition. Suppose that f and $\mathbf{g}$ are $\mathscr{C}^1$ functions and $\mathbf{x}^*$ (locally) solves the constraint optimization problem and $\mathbf{g}'(\mathbf{x}^*)$ has maximal rank m, then there exist a unique vector $\boldsymbol{\lambda}^* = (\lambda_1^*,\dots,\lambda_m^*)$ such that $\nabla\mathscr{L}(\mathbf{x}^*,\boldsymbol{\lambda}^*) = 0$.

This necessary condition implies that $\partial f'(\mathbf{x}^*) = \boldsymbol{\lambda}^* \mathbf{g}'(\mathbf{x}^*)$. The following figure illustrates the situation for the case of two variables x and y and one constraint $g(x, y) = c$. Then we have find $\nabla f = \lambda \nabla g$, that is, in an optimal point ∇f is some multiple of ∇g.



Also observe that a point $\mathbf{x}$ is admissible (i.e., satisfies constraint $\mathbf{g}(\mathbf{x}) = \mathbf{c}$) if and only if $\frac{\partial \mathscr{L}}{\partial \lambda}(\mathbf{x}, \boldsymbol{\lambda}) = 0$ for some vector $\boldsymbol{\lambda} = 0$.

Sufficient condition. Let f and $\mathbf{g}$ be $\mathscr{C}^1$. Suppose there exists a $\boldsymbol{\lambda}^* = (\lambda_1^*, \dots, \lambda_m^*)$ and an admissible $\mathbf{x}^*$ such that $(\mathbf{x}^*, \boldsymbol{\lambda}^*)$ is a stationary point of $\mathscr{L}$, i.e., $\nabla \mathscr{L}(\mathbf{x}^*, \boldsymbol{\lambda}^*) = 0$. If $\mathscr{L}(\mathbf{x}, \boldsymbol{\lambda}^*)$ is concave (convex) in $\mathbf{x}$, then $\mathbf{x}^*$ solves the constraint maximization (minimization) problem. Theorem 8.14

PROOF. By Theorem 8.4 these conditions imply that $\mathbf{x}^*$ is a maximum of $\mathscr{L}(\mathbf{x}, \boldsymbol{\lambda}^*)$ w.r.t. $\mathbf{x}$, i.e.,

$$\begin{aligned}\mathscr{L}(\mathbf{x}^*; \boldsymbol{\lambda}^*) &= f(\mathbf{x})^* - \sum_{j=1}^{m} \lambda_j^* (g_j(\mathbf{x}^*) - c_j) \\ &\geq f(\mathbf{x}) - \sum_{j=1}^{m} \lambda_j^* (g_j(\mathbf{x}) - c_j) = \mathscr{L}(\mathbf{x}; \boldsymbol{\lambda}^*) .\end{aligned}$$

However, all admissible $\mathbf{x}$ satisfy $g_j(\mathbf{x}) = c_j$ for all j and thus $f(\mathbf{x}^*) \geq f(\mathbf{x})$ for all admissible $\mathbf{x}$. Hence $\mathbf{x}^*$ solves the constraint maximization problem. □

Similar to Theorem 8.7 we can find sufficient conditions for local solutions of the constraint optimization problem. That is, $(\mathbf{x}^*, \boldsymbol{\lambda}^*)$ is a stationary point of $\mathscr{L}$ and $\mathscr{L}$ w.r.t. $\mathbf{x}$ is strictly concave (strictly convex) in some open ball around $(\mathbf{x}^*, \boldsymbol{\lambda}^*)$, then $\mathbf{x}^*$ solves the local constraint maximization (minimization) problem. Such an open ball exists if the Hessian of $\mathscr{L}$ w.r.t. $\mathbf{x}$ is negative (positive) definite in $(\mathbf{x}^*, \boldsymbol{\lambda}^*)$.

However, such a condition is too strong. There is no need to investigate the behavior of $\mathscr{L}$ for points $\mathbf{x}$ that do not satisfy constraint $\mathbf{g}(\mathbf{x}) = \mathbf{c}$. Hence (roughly spoken) it is sufficient that the Lagrange function $\mathscr{L}$ is strictly concave on the affine subspace spanned by the gradients $\nabla g_1(\mathbf{x}^*), \dots, \nabla g_m(\mathbf{x}^*)$. Again it is sufficient to look at the definiteness of the Hessian $\mathscr{L}''$ at $\mathbf{x}^*$. ($\mathscr{L}''$ denotes the Hessian w.r.t. $\mathbf{x}$.)

Lemma 8.15

Let f and $\mathbf{g}$ be $\mathscr{C}^2$. Suppose there exists a $\boldsymbol{\lambda}^* = (\lambda_1^*, \dots, \lambda_m^*)$ and an admissible $\mathbf{x}^*$ such that $\nabla\mathscr{L}(\mathbf{x}^*, \boldsymbol{\lambda}^*) = 0$. If there exists an open ball around $\mathbf{x}^*$ such that the quadratic form

$$\mathbf{h}'\mathscr{L}''(\mathbf{x}^*; \boldsymbol{\lambda}^*)\mathbf{h}$$

is negative (positive) definite for all $\mathbf{h} \in \operatorname{span}\big(\nabla g_1(\mathbf{x}^*), \dots, \nabla g_m(\mathbf{x}^*)\big)$, then $\mathbf{x}^*$ solves the local constraint maximization (minimization) problem.

This condition can be verified by means of a theorem from Linear Algebra which requires the concept of the bordered Hessian.

Definition 8.16

Bordered Hessian. The matrix

$$\bar{\mathbf{H}}(\mathbf{x}; \boldsymbol{\lambda}) = \begin{pmatrix} 0 & \mathbf{g}'(\mathbf{x}) \\ (\mathbf{g}'(\mathbf{x}))' & \mathscr{L}''(\mathbf{x}; \boldsymbol{\lambda}) \end{pmatrix}$$

$$= \begin{pmatrix} 0 & \dots & 0 & \frac{\partial g_1}{\partial x_1} & \dots & \frac{\partial g_1}{\partial x_n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \frac{\partial g_m}{\partial x_1} & \dots & \frac{\partial g_m}{\partial x_n} \\ \frac{\partial g_1}{\partial x_1} & \dots & \frac{\partial g_m}{\partial x_1} & \mathscr{L}_{x_1x_1} & \dots & \mathscr{L}_{x_1x_n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_1}{\partial x_n} & \dots & \frac{\partial g_m}{\partial x_n} & \mathscr{L}_{x_nx_1} & \dots & \mathscr{L}_{x_nx_n} \end{pmatrix}$$

is called the **bordered Hessian** of $\mathscr{L}(\mathbf{x}; \boldsymbol{\lambda}) = f(\mathbf{x}) - \boldsymbol{\lambda}'(\mathbf{g}(\mathbf{x}) - \mathbf{c})$.
We denote its leading principal minors by

$$B_r(\mathbf{x}) = \begin{vmatrix} 0 & \dots & 0 & \frac{\partial g_1}{\partial x_1}(\mathbf{x}; \boldsymbol{\lambda}) & \dots & \frac{\partial g_1}{\partial x_r}(\mathbf{x}; \boldsymbol{\lambda}) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & \frac{\partial g_m}{\partial x_1}(\mathbf{x}; \boldsymbol{\lambda}) & \dots & \frac{\partial g_m}{\partial x_r}(\mathbf{x}; \boldsymbol{\lambda}) \\ \frac{\partial g_1}{\partial x_1}(\mathbf{x}; \boldsymbol{\lambda}) & \dots & \frac{\partial g_m}{\partial x_1}(\mathbf{x}; \boldsymbol{\lambda}) & \mathscr{L}_{x_1x_1}(\mathbf{x}; \boldsymbol{\lambda}) & \dots & \mathscr{L}_{x_1x_r}(\mathbf{x}; \boldsymbol{\lambda}) \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial g_1}{\partial x_r}(\mathbf{x}; \boldsymbol{\lambda}) & \dots & \frac{\partial g_m}{\partial x_r}(\mathbf{x}; \boldsymbol{\lambda}) & \mathscr{L}_{x_rx_1}(\mathbf{x}; \boldsymbol{\lambda}) & \dots & \mathscr{L}_{x_rx_r}(\mathbf{x}; \boldsymbol{\lambda}) \end{vmatrix}.$$

Theorem 8.17

Sufficient condition for local optimum. Let f and $\mathbf{g}$ be $\mathscr{C}^1$. Suppose there exists a $\boldsymbol{\lambda}^* = (\lambda_1^*, \dots, \lambda_m^*)$ and an admissible $\mathbf{x}^*$ such that $\nabla\mathscr{L}(\mathbf{x}^*, \boldsymbol{\lambda}^*) = 0$.

(a) If $(-1)^r B_r(\mathbf{x}) > 0$ for all $r = m+1, \dots, n$, then $\mathbf{x}^*$ solves the local constraint maximization problem.

(b) If $(-1)^m B_r(\mathbf{x}) > 0$ for all $r = m+1, \dots, n$, then $\mathbf{x}^*$ solves the local constraint minimization problem.

See Lecture 12 in *Mathematische Methoden* for examples.

8.4 Kuhn-Tucker Conditions

In this section we consider the optimization problem

$$\begin{aligned}
&\max \quad f(x_1,\dots,x_n) \\
&\text{subject to} \quad g_j(x_1,\dots,x_n) \le c_j, \quad j=1,\dots,m \quad (m<n), \\
&\quad \text{and} \quad x_i \ge 0, \quad i=1,\dots n \quad \text{(non-negativity constraint)}
\end{aligned}$$

or in vector notation

$$\max \quad f(\mathbf{x}) \quad \text{subject to} \quad \mathbf{g}(\mathbf{x}) \ge \mathbf{c} \quad \text{and} \quad \mathbf{x} \ge 0.$$

Again let

$$\mathscr{L}(\mathbf{x};\boldsymbol{\lambda}) = f(\mathbf{x}) - \boldsymbol{\lambda}'(\mathbf{g}(\mathbf{x}) - \mathbf{c}) = f(\mathbf{x}) - \sum_{j=1}^{m} \lambda_j (g_j(\mathbf{x}) - c_j)$$

denote the Lagrange function of this problem.

Definition 8.18

Kuhn-Tucker condition. The conditions

$$\begin{aligned}
&\frac{\partial \mathscr{L}}{\partial x_j} \le 0, \quad x_j \ge 0 \quad \text{and} \quad x_j \frac{\partial \mathscr{L}}{\partial x_j} = 0 \\
&\frac{\partial \mathscr{L}}{\partial \lambda_i} \ge 0, \quad \lambda_i \ge 0 \quad \text{and} \quad \lambda_i \frac{\partial \mathscr{L}}{\partial \lambda_i} = 0
\end{aligned}$$

are called the **Kuhn-Tucker conditions** of the problem.

Theorem 8.19

Kuhn-Tucker sufficient condition. Suppose that f and $\mathbf{g}$ are $\mathscr{C}^1$ functions and there exists a $\boldsymbol{\lambda}^* = (\lambda_1^*,\dots,\lambda_m^*)$ and an admissible $\mathbf{x}^*$ such that

(1) The objective function f is concave.

(2) The functions g_j are convex for $j = 1,\dots,m$.

(3) The point $\mathbf{x}^*$ satisfies the Kuhn-Tucker conditions.

Then $\mathbf{x}^*$ solves the constraint maximization problem.

See Lecture 13 in *Mathematische Methoden* for examples.

— Exercises

8.1 Compute all local and global extremal points of the functions

(a) $f(x) = (x-3)^6$

(b) $g(x) = \frac{x^2+1}{x}$

8.2 Compute the local and global extremal points of the functions

(a) $f: [0,\infty] \to \mathbb{R}, x \mapsto \frac{1}{x} + x$

(b) $f: [0,\infty] \to \mathbb{R}, x \mapsto \sqrt{x} - x$

(c) $g: \mathbb{R} \to \mathbb{R}, x \mapsto e^{-2x} + 2x$

8.3 Compute all local extremal points and saddle points of the following functions. Are the local extremal points also globally extremal.

(a) $f(x,y) = -x^2 + xy + y^2$

(b) $f(x,y) = \frac{1}{x}\ln(x) - y^2 + 1$

(c) $f(x,y) = 100(y-x^2)^2 + (1-x)^2$

(d) $f(x,y) = 3x + 4y - e^x - e^y$

8.4 Compute all local extremal points and saddle points of the following functions. Are the local extremal points also globally extremal.

$$f(x_1,x_2,x_3) = (x_1^3 - x_1)x_2 + x_3^2 .$$

8.5 We are given the following constraint optimization problem

$$\max(\min) \quad f(x,y) = x^2 y \quad \text{subject to} \quad x + y = 3.$$

(a) Solve the problem graphically.

(b) Compute all stationary points.

(c) Use the bordered Hessian to determine whether these stationary points are (local) maxima or minima.

8.6 Compute all stationary points of the constraint optimization problem

$$\max\ (\min) \quad f(x_1,x_2,x_3) = \frac{1}{3}(x_1-3)^3 + x_2 x_3$$
$$\text{subject to} \quad x_1 + x_2 = 4 \text{ and } x_1 + x_3 = 5.$$

Use the bordered Hessian to determine whether these stationary points are (local) maxima or minima.

8.7 A household has an income m and can buy two commodities with prices p_1 and p_2. We have

$$p_1 x_1 + p_2 x_2 = m$$

where x_1 and x_2 denote the quantities. Assume that the household has a utility function

$$u(x_1, x_2) = \alpha \ln(x_1) + (1-\alpha)\ln(x_2)$$

where $\alpha \in (0,1)$.

(a) Solve this constraint optimization problem.

(b) Compute the change of the optimal utility function when the price of commodity 1 changes.

(c) Compute the change of the optimal utility function when the income m changes.

8.8 We are given the following constraint optimization problem

$$\max \quad f(x,y) = -(x-2)^2 - y \quad \text{subject to} \quad x+y \le 1, \quad x,y \ge 0\,.$$

(a) Solve the problem graphically.

(b) Solve the problem by means of the Kuhn-Tucker conditions.

— Problems

8.9 Our definition of a local maximum (Definition 8.6) is quite simple but has unexpected consequences: There exist non-constant functions where a global minimum is a local maximum. Give an example for such a function. How could Definition 8.6 be "repaired"?

8.10 Let $f\colon \mathbb{R}^n \to \mathbb{R}$ and $T\colon \mathbb{R} \to \mathbb{R}$ be a strictly monotonically increasing transformation. Show that $\mathbf{x}^*$ is a maximum of f if and only if $\mathbf{x}^*$ is a maximum of the transformed function $T \circ f$.

9

Integration

We know the boundary of some domain. What is its area?

In this chapter we deal with two topics that seem to be quite distinct: We want to invert the result of differentiation and we want to compute the area of a region that is enclosed by curves. These two tasks are linked by the Fundamental Theorem of Calculus.

9.1 The Antiderivative of a Function

Definition 9.1

A univariate function F is called an **antiderivative** of some function f if $F'(x) = f(x)$.

Motivated by the Fundamental Theorem of Calculus (p. 93) the antiderivative is usually called the **indefinite integral** (or **primitive integral**) of f and denoted by

$$F(x) = \int f(x)\,dx\,.$$

Finding antiderivatives is quite a hard issue. In opposition to differentiation often no straightforward methods exist. Roughly spoken we have to do the following:

Make an educated guess and verify by differentiation.

Example 9.2

Find the antiderivative of $f(x) = \ln(x)$.

SOLUTION. Guess: $F(x) = x(\ln(x) - 1)$.
Verify: $F'(x) = (x(\ln(x)-1))' = 1\cdot(\ln(x)-1) + x\cdot\frac{1}{x} = \ln(x)$. ♢

It is quite obvious that $F(x) = x(\ln(x)-1) + 123$ is also an antiderivative of $\ln(x)$ as is $F(x) = x(\ln(x)-1) + c$ for every $c \in \mathbb{R}$.

Lemma 9.3

If $F(x)$ is an antiderivative of some function $f(x)$, then $F(x)+c$ is also an antiderivative of $f(x)$ for every $c \in \mathbb{R}$. The constant c is called the **integration constant**.

$f(x)$	$\int f(x)\,dx$	
0	c	
x^α	$\frac{1}{\alpha+1}\cdot x^{\alpha+1}+c$	for $\alpha \neq -1$
e^x	e^x+c	
$\frac{1}{x}$	$\ln\lvert x\rvert+c$	
$\cos(x)$	$\sin(x)+c$	
$\sin(x)$	$-\cos(x)+c$	

Table 9.4

Table of antiderivatives of some elementary functions.

Summation rule:	$\int \alpha f(x)+\beta g(x)\,dx = \alpha\int f(x)\,dx+\beta\int g(x)\,dx$
By parts:	$\int f(x)\cdot g'(x)\,dx = f(x)\cdot g(x)-\int f'(x)\cdot g(x)\,dx$
By substitution:	$\int f\big(g(x)\big)\cdot g'(x)\,dx = \int f(z)\,dz$ where $z=g(x)$ and $dz=g'(x)\,dx$

Table 9.5

Rules for indefinite integrals.

Fortunately there exist some tools that ease the task of "guessing" the antiderivative. Table 9.4 lists basic integrals. Observe that we get these antiderivatives simply by exchanging the columns in our table of derivatives (Table 4.8).

Table 9.5 lists integration rules that allow to reduce the issue of finding indefinite integrals of complicated expressions to simpler ones. Again these rules can be directly derived from the corresponding rules in Table 4.9 for computing derivatives. There exist many other such rules which are, however, often only applicable to special functions. Computer algebra systems like *Maxima* thus use much larger tables for basic integrals and integration rules for finding indefinite integrals.

DERIVATION OF THE INTEGRATION RULES. The *summation rule* is just a consequence of the linearity of the differential operator.

For *integration by parts* we have to assume that both f and g are differentiable. Thus we find by means of the product rule

$$\begin{aligned} f(x)\cdot g(x) &= \int \big(f(x)\cdot g(x)\big)'\,dx = \int \big(f'(x)g(x)+f(x)g'(x)\big)\,dx \\ &= \int f'(x)g(x)\,dx + \int f(x)g'(x)\,dx \end{aligned}$$

and hence the rule follows.

For *integration by substitution* let F denote an antiderivative of f.

Then we find

$$\begin{aligned}\int f(z)\,dz &= F(z)\\ &= F\big(g(x)\big) = \int \Big(F\big(g(x)\big)\Big)'\,dx = \int F'\big(g(x)\big)\,g'(x)\,dx\\ &= \int f\big(g(x)\big)\,g'(x)\,dx\end{aligned}$$

that is, if the integrand is of the form $f\big(g(x)\big)\,g'(x)$ we first compute the indefinite integral $\int f(z)\,dz$ and then substitute $z = g(x)$. □

Example 9.6

Compute the indefinite integral of $f(x) = 4x^3 - x^2 + 3x - 5$.

SOLUTION. By the summation rule we find

$$\begin{aligned}\int f(x)\,dx &= \int 4x^3 - x^2 + 3x - 5\,dx\\ &= 4\int x^3\,dx - \int x^2\,dx + 3\int x\,dx - 5\int dx\\ &= 4\frac{1}{4}x^4 - \frac{1}{3}x^3 + 3\frac{1}{2}x^2 - 5x + c\\ &= x^4 - \frac{1}{3}x^3 + \frac{3}{2}x^2 - 5x + c\,.\end{aligned}$$ ♢

Example 9.7

Compute the indefinite integral of $f(x) = x\,e^x$.

SOLUTION. Integration by parts yields

$$\int f(x)\,dx = \int \underbrace{x}_{f}\cdot\underbrace{e^x}_{g'}\,dx = \underbrace{x}_{f}\cdot\underbrace{e^x}_{g}\,dx - \int \underbrace{1}_{f'}\cdot\underbrace{e^x}_{g}\,dx = x\cdot e^x - e^x + c\,.$$

$$\begin{aligned}f(x) &= x & &\Rightarrow & f'(x) &= 1\\ g'(x) &= e^x & &\Rightarrow & g(x) &= e^x\end{aligned}$$ ♢

Example 9.8

Compute the indefinite integral of $f(x) = 2x\,e^{x^2}$.

SOLUTION. By substitution we find

$$\int f(x)\,dx \int \exp(\underbrace{x^2}_{g(x)})\cdot\underbrace{2x}_{g'(x)}\,dx = \int \exp(z)\,dz = e^z + c = e^{x^2} + c\,.$$

$$z = g(x) = x^2 \quad\Rightarrow\quad dz = g'(x)\,dx = 2x\,dx$$ ♢

Example 9.9

Compute the indefinite integral of $f(x) = x^2\cos(x)$.

SOLUTION. Integration by parts yields

$$\int f(x)\,dx \int \underbrace{x^2}_{f}\cdot\underbrace{\cos(x)}_{g'}\,dx = \underbrace{x^2}_{f}\cdot\underbrace{\sin(x)}_{g} - \int \underbrace{2x}_{f'}\cdot\underbrace{\sin(x)}_{g}\,dx\,.$$

For the last term we have to apply integration by parts again:

$$\int \underbrace{2x}_{f} \cdot \underbrace{\sin(x)}_{g'}\,dx = \underbrace{2x}_{f} \cdot \underbrace{(-\cos(x))}_{g} - \int \underbrace{2}_{f'} \cdot \underbrace{(-\cos(x))}_{g}\,dx$$
$$= -2x \cdot \cos(x) - 2 \cdot (-\sin(x)) + c\,.$$

Therefore we have

$$\int x^2 \cos(x)\,dx = x^2 \sin(x) - \big(-2x\cos(x) + 2\sin(x) + c\big)$$
$$= x^2 \sin(x) + 2x\cos(x) - 2\sin(x) + c\,.$$ ♢

Sometimes the application of these integration rules might not be obvious as the following examples shows.

Example 9.10

Compute the indefinite integral of $f(x) = \ln(x)$.

SOLUTION. We write $f(x) = 1 \cdot \ln(x)$. Integration by parts yields

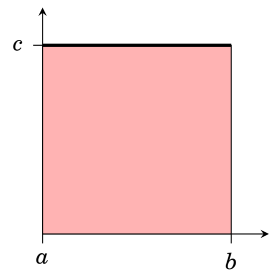
$$\int \underbrace{\ln(x)}_{f} \cdot \underbrace{1}_{g'}\,dx = \underbrace{\ln(x)}_{f} \cdot \underbrace{x}_{g}\,dx - \int \underbrace{\frac{1}{x}}_{f'} \cdot \underbrace{x}_{g}\,dx = \ln(x) \cdot x - x + c$$

$$\begin{aligned} f(x) &= \ln(x) &\Rightarrow\quad f'(x) &= \tfrac{1}{x} \\ g'(x) &= 1 &\Rightarrow\quad g(x) &= x \end{aligned}$$ ♢

We again want to note that there are no simple recipes for finding indefinite integrals. Even with integration rules like those in Table 9.5 there remains still trial and error. (Of course experience increases the change of successful guesses significantly.)

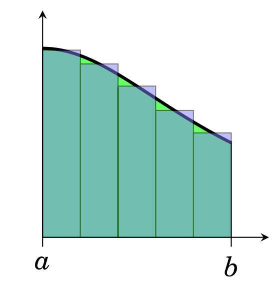
There are even two further obstacles: (1) not all functions have an antiderivative; (2) the indefinite integral may exist but it is not possible to express it in terms of elementary functions. The density of the normal distribution, $\varphi(x) = \exp(-x^2)$, is the most prominent example.

9.2 The Riemann Integral



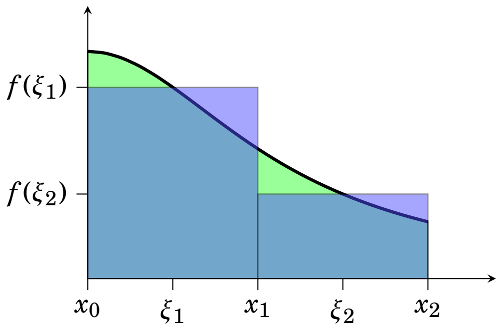
Suppose we are given some nonnegative function f over some interval $[a,b]$ and we have to compute the area A below the graph of f. If $f(x) = c$ is a constant function, then this task is quite simple: The region in question is a rectangle and we find by basic geometry (length of base × height)

$$A = c \cdot (b-a)\,.$$



For general functions with "irregular"-shaped graphs we may approximate the function by a **step function** (or *staircase function*), i.e. a piecewise constant function. The area for the step function can then be computed for each of the rectangles and added up for the total area.

Thus we select points $x_0 = 0 < x_1 < \ldots < x_n = b$ and compute f at intermediate points $\xi \in (x_{i-1}, x_i)$, for $i = 1, \ldots, n$.

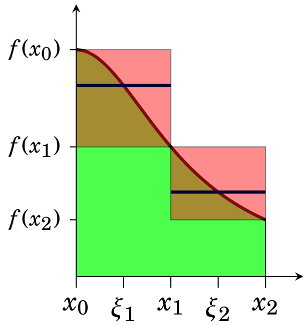


Hence we find for the area

$$A \approx \sum_{i=1}^{n} f(\xi_i)\cdot(x_i - x_{i-1})\,.$$

If f is a monotonically decreasing function and the points $x_0, x_1, \dots, x_n$ are selected equidistant, i.e., $(x_i - x_{i-1}) = \frac{1}{n}(b-a)$, then we find for the approximation error

$$\left| A - \sum_{i=1}^{n} f(\xi_i)\cdot(x_i - x_{i-1}) \right| \le (f_{\max} - f_{\min})(b-a)\frac{1}{n} \to 0 \quad \text{as } n \to \infty.$$



Thus when we increase the number of points n, then this so called *Riemann sum* converges to area A. For a nonmonotone function the limit may not exist. If it exists we get the area under the graph.

Definition 9.11

Riemann integral. Let f be some function defined on $[a,b]$. Let $(Z_k) = \left(\left\{x_0^{(k)}, x_1^{(k)}, \dots, x_{n(k)}^{(k)}\right\}\right)$ be a sequence of point sets such that $x_0^{(k)} = a < x_1^{(k)} < \dots x_{k-1}^{(k)} < x_{n(k)}^{(k)} = b$ for all $k = 1, 2, \dots$ and $\max_{i=1,\dots,n(k)}\left(x_i^{(k)} - x_{i-1}^{(k)}\right) \to 0$ as $k \to \infty$. Let $\xi_i^{(k)} \in \left(x_{i-1}^{(k)}, x_i^{(k)}\right)$. If the **Riemann sum**

$$I_k = \sum_{i=1}^{n(k)} f(\xi_i^{(k)})\cdot\left(x_i^{(k)} - x_{i-1}^{(k)}\right)$$

converges for all such sequences (Z_k) then the function $f\colon [a,b] \to \mathbb{R}$ is **Riemann integrable**. The limit is called the **Riemann integral** of f and denoted by

$$\int_a^b f(x)\,dx = \lim_{k\to\infty} \sum_{i=1}^{n(k)} f(\xi_i^{(k)})\cdot\left(x_i^{(k)} - x_{i-1}^{(k)}\right)\,.$$

This definition requires some remarks.

- This limit (if it exists) is uniquely determined.
- Not all functions are Riemann integrable, that is, there exist functions where this limit does not exist for some choices of sequence (Z_k). However, bounded "nice" (in particular continuous) functions are always Riemann integrable.

Let f and g be integrable functions and $\alpha, \beta \in \mathbb{R}$. Then we find

$$\int_a^b (\alpha f(x) + \beta g(x))\,dx = \alpha \int_a^b f(x)\,dx + \beta \int_a^b g(x)\,dx$$

$$\int_a^b f(x)\,dx = -\int_b^a f(x)\,dx$$

$$\int_a^a f(x)\,dx = 0$$

$$\int_a^c f(x)\,dx = \int_a^b f(x)\,dx + \int_b^c f(x)\,dx$$

$$\int_a^b f(x)\,dx \le \int_a^b g(x)\,dx \qquad \text{if } f(x) \le g(x) \text{ for all } a \le x \le b$$

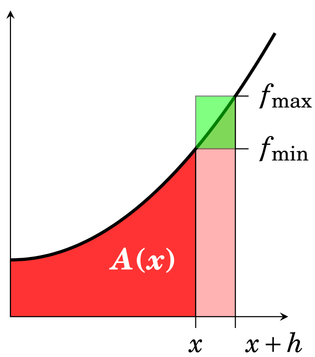
Table 9.12

Properties of definite integrals.

- There also exist other concepts of integration. However, for continuous functions these coincide. Thus we will say **integrable** and **integral** for short.
- As we will see in the next section integrals are usually called *definite integrals*.
- From the definition of the integral we immediately see that for regions where function f is *negative* the integral also is *negative*.
- Similarly, as the definition of *Riemann sum* contains the term $\left(x_i^{(k)} - x_{i-1}^{(k)}\right)$ instead of its absolute value $\left|x_i^{(k)} - x_{i-1}^{(k)}\right|$, the integral of a *positive* function becomes *negative* if the interval (a,b) is traversed from right to left.

Table 9.12 lists important properties of the definite integral. These can be derived from the definition of integrals and the rules for limits (Theorem 4.3 on p. 28).

9.3 The Fundamental Theorem of Calculus



We have defined the integral as the limit of Riemann sums. However, we still need a efficient method to compute the integral. On the other hand we did not establish any condition that ensure the existence of the antiderivative of a given function. Astonishingly these two apparently distinct problems are closely connected.

Let f be some *continuous* function and suppose that the area of f under the graph in the interval $[0,x]$ is given by $A(x)$. We then get the area under the curve of f in the interval $[x, x+h]$ for some h by subtraction, $A(x+h) - A(x)$. As f is continuous it has a maximum $f_{max}(h)$ and a minimum $f_{\min}(h)$ on $[x, x+h]$. Then we find

$$f_{\min}(h)\cdot h \le A(x+h)-A(x) \le f_{\max}(h)\cdot h$$
$$f_{\min}(h) \le \frac{A(x+h)-A(x)}{h} \le f_{\max}(h)$$

If $h \to 0$ we then find by continuity of f,

$$\lim_{h\to 0} f_{\min}(h) = \lim_{h\to 0} f_{\max}(h) = f(x)$$

and hence

$$f(x) \le \underbrace{\lim_{h\to 0} \frac{A(x+h)-A(x)}{h}}_{=A'(x)} \le f(x)\,.$$

Consequently, $A(x)$ is differentiable and we arrive at

$$A'(x) = f(x)$$

that is, $A(x)$ is an antiderivative of f.

This observation is formally stated in the two parts of the Fundamental Theorem of Calculus which we state without a stringent proof.

Theorem 9.13

First fundamental theorem of calculus. Let $f\colon [a,b] \to \mathbb{R}$ be a continuous function that admits an antiderivative F on $[a,b]$, then

$$\int_a^b f(x)\,dx = F(b)-F(a)\,.$$

Theorem 9.14

Second fundamental theorem of calculus. Let $f\colon [a,b] \to \mathbb{R}$ be a continuous function and F defined for all $x \in [a,b]$ as the integral

$$F(x) = \int_a^x f(t)\,dt\,.$$

Then F is differentiable on (a,b) and

$$F'(x) = f(x) \quad \text{for all } x \in (a,b).$$

An immediate corollary is that every continuous function has an antiderivative, namely the integral function F.

Notice that the first part states that we simply can use the indefinite integral to compute the integral of continuous functions, $\int_a^b f(x)\,dx$. In contrast, the second part gives us a sufficient condition for the existence of the antiderivative of a function.

By parts:	$\int_a^b f(x)g'(x)\,dx = f(x)g(x)\Big\|_a^b - \int_a^b f'(x)g(x)\,dx$
By substitution:	$\int_a^b f(g(x))\cdot g'(x)\,dx = \int_{g(a)}^{g(b)} f(z)\,dz$ where $z = g(x)$ and $dz = g'(x)\,dx$

Table 9.17

Rules for definite integrals.

9.4 The Definite Integral

Theorem 9.13 provides a method to compute the integral of a function without dealing with limits of Riemann sums. This motivates the term *definite integral*.

Definition 9.15

Let $f\colon [a,b] \to \mathbb{R}$ be a continuous function and F an antiderivative of f. Then

$$\int_a^b f(x)\,dx = F(x)\Big|_a^b = F(b) - F(a)$$

is called the **definite integral** of f.

Example 9.16

Compute the definite integral of $f(x) = x^2$ in the interval $[0,1]$.

SOLUTION. $\int_0^1 x^2\,dx = \frac{1}{3}x^3\Big|_0^1 = \frac{1}{3}\cdot 1^3 - \frac{1}{3}\cdot 0^3 = \frac{1}{3}$. ♢

The rules for indefinite integrals in Table 9.5 can be easily translated into rules for the definite integral, see Table 9.17

Example 9.18

Compute $\int_e^{10} \frac{1}{\log(x)}\cdot\frac{1}{x}\,dx$.

SOLUTION.

$$\int_e^{10} \frac{1}{\log(x)}\cdot\frac{1}{x}\,dx = \int_1^{\log(10)} \frac{1}{z}\,dz$$

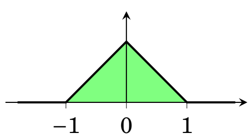
$$z = \log(x) \quad\Rightarrow\quad dz = \frac{1}{x}\,dx$$

$$= \log(z)\Big|_1^{\log(10)} = \log(\log(10)) - \log(1) \approx 0.834\,.$$ ♢

Example 9.19

Compute $\int_{-2}^{2} f(x)\,dx$ where

$$f(x) = \begin{cases} 1+x & \text{for } -1 \le x < 0 \\ 1-x & \text{for } 0 \le x < 1 \\ 0 & \text{for } x < -1 \text{ and } x \ge 1 \end{cases}$$



SOLUTION.

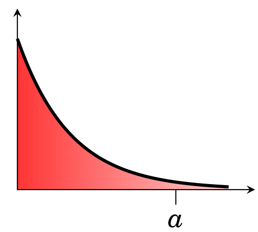
$$\begin{aligned}\int_{-2}^{2} f(x)\,dx &= \int_{-2}^{-1} f(x)\,dx + \int_{-1}^{0} f(x)\,dx + \int_{0}^{1} f(x)\,dx + \int_{1}^{2} f(x)\,dx \\ &= \int_{-2}^{-1} 0\,dx + \int_{-1}^{0} (1+x)\,dx + \int_{0}^{1} (1-x)\,dx + \int_{1}^{2} 0\,dx \\ &= \left(x + \frac{1}{2}x^2\right)\Big|_{-1}^{0} + \left(x - \frac{1}{2}x^2\right)\Big|_{0}^{1} \\ &= \frac{1}{2} + \frac{1}{2} = 1\,. \qquad \diamondsuit\end{aligned}$$

9.5 Improper Integrals

Example 9.20

Suppose we want to compute $\int_0^b e^{-\lambda x}\,dx$. We then get

$$\int_0^b e^{-\lambda x}\,dx = \int_0^{-\lambda b} e^z \left(-\tfrac{1}{\lambda}\right) dz = -\tfrac{1}{\lambda} e^z\Big|_0^{-\lambda b} = \frac{1}{\lambda}\left(1 - e^{-\lambda b}\right). \qquad \diamondsuit$$



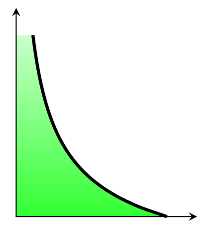
So what happens if b tends to ∞, i.e., when the domain of integration is unbounded. Obviously

$$\lim_{b\to\infty} \int_0^b e^{-\lambda x}\,dx = \lim_{b\to\infty} \frac{1}{\lambda}\left(1 - e^{-\lambda b}\right) = \frac{1}{\lambda}\,.$$

Thus we may use the symbol

$$\int_0^\infty f(x)\,dx$$

for this limit. Similarly we may want to compute the integral $\int_0^1 \frac{1}{\sqrt{x}}\,dx$. But then 0 does not belong to the domain of f as $f(0)$ is not defined. We then replace the lower bound 0 by some $a > 0$, compute the integral and find the limit for $a \to 0$. We again write



$$\int_0^1 \frac{1}{\sqrt{x}}\,dx = \lim_{a\to 0^+} \int_a^1 \frac{1}{\sqrt{x}}\,dx$$

where 0^+ indicates that we are looking at the limit from the right hand side.

Definition 9.21

improper integral

Integrals of functions that are unbounded at a or b or have unbounded domain (i.e., $a = -\infty$ or $b = \infty$) are called **improper integrals** . They are defined as limits of proper integrals. If the limit

$$\int_0^b f(x)\,dx = \lim_{t\to b} \int_0^t f(x)\,dx$$

exists we say that the improper integral *converges*. Otherwise we say that it *diverges*.

For practical reasons we demand that this limit exists if and only if $\lim_{t\to\infty}\int_0^t |f(x)|\,dx$ exists.

Example 9.22

Compute the improper integral $\int_0^1 \frac{1}{\sqrt{x}}\,dx$.

SOLUTION.

$$\int_0^1 \frac{1}{\sqrt{x}}\,dx = \lim_{t\to 0}\int_t^1 x^{-\frac{1}{2}}\,dx = \lim_{t\to 0} 2\sqrt{x}\Big|_t^1 = \lim_{t\to 0}(2-2\sqrt{t}) = 2\,. \qquad \diamondsuit$$

Example 9.23

Compute the improper integral $\int_1^\infty \frac{1}{x^2}\,dx$.

SOLUTION.

$$\int_1^\infty \frac{1}{x^2}\,dx = \lim_{t\to\infty}\int_1^t x^{-2}\,dx = \lim_{t\to\infty} -\frac{1}{x}\Big|_1^t = \lim_{t\to\infty} -\frac{1}{t} - (-1) = 1\,. \qquad \diamondsuit$$

Example 9.24

Compute the improper integral $\int_1^\infty \frac{1}{x}\,dx$.

SOLUTION.

$$\int_1^\infty \frac{1}{x}\,dx = \lim_{t\to\infty}\int_1^t \frac{1}{x}\,dx = \lim_{t\to\infty} \log(x)\Big|_1^t = \lim_{t\to\infty} \log(t) - \log(1) = \infty\,.$$

The improper integral diverges. $\diamondsuit$

9.6 Differentiation under the Integral Sign

We are given some continuous function f with antiderivative F, i.e., $F'(x) = f(x)$. If we differentiate the definite integral $\int_a^x f(t)\,dt = F(x) - F(a)$ w.r.t. its upper bound we obtain

$$\frac{d}{dx}\int_a^x f(t)\,dt = (F(x) - F(a))' = F'(x) = f(x)\,.$$

That is, the derivative of the definite integral w.r.t. the upper limit of integration is equal to the integrand evaluated at that point.

We can generalize this result. Suppose that both lower and upper limit of the definite integral are differentiable functions $a(x)$ and $b(x)$, respectively. Then we find by the chain rule

$$\frac{d}{dx}\int_{a(x)}^{b(x)} f(t)\,dt = (F(b(x)) - F(a(x)))' = f(b(x))\,b'(x) - f(a(x))\,a'(x)\,.$$

Now suppose that $f(x,t)$ is a continuous function of two variables and consider the function $F(x)$ defined by

$$F(x) = \int_a^b f(x,t)\,dt\,.$$

Its derivative $F'(x)$ can be computed as

$$\begin{aligned} F'(x) &= \lim_{h\to 0} \frac{F(x+h)-F(x)}{h} \\ &= \lim_{h\to 0} \int_a^b \frac{f(x+h,t)-f(x,t)}{h}\,dt \\ &= \int_a^b \lim_{h\to 0} \frac{f(x+h,t)-f(x,t)}{h}\,dt \\ &= \int_a^b \frac{\partial f(x,t)}{\partial x}\,dt\,. \end{aligned}$$

That is, in order to get the derivative of the integral with respect to parameter x we differentiate under the integral sign.

Of course the partial derivative $f_x(x,t)$ must be an integrable function which is satisfied whenever it is continuous by the Fundamental Theorem.

It is important to note that both the (Riemann-) integral and the partial derivative are limits. Thus we have to exchange these two limits in our calculation. Notice, however, that this is a problematic step and its validation requires tools from advanced calculus.

In general exchanging limits can change the result!

We now can combine our observations into a single formula.

Theorem 9.25

Leibniz's formula. Let

$$F(x) = \int_{a(x)}^{b(x)} f(x,t)\,dt$$

where the function $f(x,t)$ and its partial derivative $f_x(x,t)$ are continuous in both x and t in the region $\{(x,t)\colon x_0 \le x \le x_1,\, a(x) \le t \le b(x)\}$ and the functions $a(x)$ and $b(x)$ are $\mathscr{C}^1$ functions over $[x_0,x_1]$. Then

$$F'(x) = f(x,b(x))\,b'(x) - f(x,a(x))\,a'(x) + \int_{a(x)}^{b(x)} \frac{\partial f(x,t)}{\partial x}\,dt\,.$$

PROOF. Let $H(x,a,b) = \int_a^b f(x,t)\,dt$. Then $F(x) = H(x,a(x),b(x))$ and we find by the chain rule

$$F'(x) = H_x + H_a a'(x) + H_b b'(x)\,.$$

Since $H_x = \int_a^b f_x(x,t)\,dt$, $H_a = -f(x,a)$ and $H_b = f(x,b)$, the result follows. □

Example 9.26

Compute $F'(x)$, $x \ge 0$, when $F(x) = \int_x^{2x} t x^2\,dt$.

SOLUTION. Let $f(x,t) = tx^2$, $a(x) = x$ and $b(x) = 2x$. By Leibniz's formula we find

$$\begin{aligned} F'(x) &= (2x)\cdot x^2 \cdot 2 - (x)\cdot x^2 \cdot 1 + \int_x^{2x} 2xt\,dt \\ &= 4x^3 - x^3 + (2x\frac{1}{2}t^2)\Big|_x^{2x} = 4x^3 - x^3 + (4x^3 - x^3) = 6x^3\,. \end{aligned}$$

◊

Leibniz formula also works for improper integrals provided that the integral $\int_{a(x)}^{b(x)} f'_x(x,t)\,dt$ converges:

$$\frac{d}{dx}\int_a^\infty f(x,t)\,dt = \int_a^\infty \frac{\partial f(x,t)}{\partial x}\,dt$$

Example 9.27

Let $K(t)$ denote the capital stock of some firm at time t, and let $p(t)$ be the price per unit of capital. Let $R(t)$ denote the rental price per unit of capital and let r be some constant interest rate. In capital theory, one principle for the determining of the correct price of the firm's capital is given by the equation

$$p(t)K(t) = \int_t^\infty R(\tau)K(\tau)e^{-r(\tau-t)}\,d\tau \qquad \text{for all } t.$$

That is, the current cost of capital should equal the discounted present value of the returns from lending it. Find an expression for $R(t)$ by differentiating the equation w.r.t. t.

SOLUTION. By differentiation the left hand side using the product rule and the right hand side using Leibniz's formula we arrive at

$$\begin{aligned} p'(t)K(t) + p(t)K'(t) &= -R(t)K(t) + \int_t^\infty R(\tau)K(\tau)\,r\,e^{-r(\tau-t)}\,d\tau \\ &= -R(t)K(t) + r\,p(t)K(t) \end{aligned}$$

and consequently

$$R(t) = \left(r - \frac{K'(t)}{K(t)}\right)p(t) - p'(t)\,. \qquad \diamondsuit$$

— Exercises

9.1 Compute the following indefinite integrals:

(a) $\int x \ln(x)\,dx$ (b) $\int x^2 \sin(x)\,dx$ (c) $\int 2x\sqrt{x^2+6}\,dx$

(d) $\int e^{x^2} x\,dx$ (e) $\int \frac{x}{3x^2+4}\,dx$ (f) $\int x\sqrt{x+1}\,dx$

(g) $\int \frac{3x^2+4}{x}\,dx$ (h) $\int \frac{\ln(x)}{x}\,dx$

9.2 Compute the following definite integrals:

(a) $\int_1^4 2x^2-1\,dx$ (b) $\int_0^2 3e^x\,dx$

(c) $\int_1^4 3x^2+4x\,dx$ (d) $\int_0^{\frac{\pi}{3}} \frac{-\sin(x)}{3}\,dx$

(e) $\int_0^1 \frac{3x+2}{3x^2+4x+1}\,dx$

9.3 Compute the following improper integrals:

(a) $\int_0^\infty -e^{-3x}\,dx$ (b) $\int_0^1 \frac{2}{\sqrt[4]{x^3}}\,dx$ (c) $\int_0^\infty \frac{x}{x^2+1}\,dx$

9.4 The marginal costs for a cost function $C(x)$ are given by $30-0.05x$. Reconstruct $C(x)$ when the fixed costs are €2000.

9.5 Compute the expectation of a so called *half-normal* distributed random variate which has domain $[0,\infty)$ and probability density function

$$f(x) = \sqrt{\frac{2}{\pi}}\exp\left(-\frac{x^2}{2}\right).$$

HINT: The **expectation** of a random variate X with density f is defined as

$$\mathbb{E}(X) = \int_{-\infty}^{\infty} x f(x)\,dx.$$

9.6 Compute the expectation of a normal distributed random variate with probability density function

$$f(x) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{x^2}{2}\right).$$

HINT: $\int_{-\infty}^{\infty} f(x)\,dx = \int_{-\infty}^{0} f(x)\,dx + \int_{0}^{\infty} f(x)\,dx$

9.7 Compute $F(x) = \int_{-2}^{x} f(t)\,dt$ for function

$$f(x) = \begin{cases} 1+x, & \text{for } -1 \le x < 0, \\ 1-x, & \text{for } 0 \le x < 1, \\ 0, & \text{for } x < -1 \text{ and } x \ge 1. \end{cases}$$

— Problems

9.8 For which value of $\alpha \in \mathbb{R}$ do the following improper integrals converge? What are their values?

(a) $\int_0^1 x^\alpha \, dx$ (b) $\int_1^\infty x^\alpha \, dx$ (c) $\int_0^\infty x^\alpha \, dx$

9.9 Let X be a so called *Cauchy* distributed random variate with probability density function

$$f(x) = \frac{1}{\pi(1+x^2)} .$$

Show that X does not have an expectation.

Why is the following approach incorrect?

$$\mathbb{E}(X) = \lim_{t\to\infty} \int_{-t}^{t} \frac{x}{\pi(1+x^2)} \, dx = \lim_{t\to\infty} 0 = 0 .$$

HINT: Show that the improper integral $\int_{-\infty}^{\infty} x f(x) \, dx$ diverges.

9.10 Compute for $T \geq 0$

$$\frac{d}{dx} \int_0^{g(x)} U(x) e^{-(t-T)} \, dt .$$

Which conditions on $g(x)$ and $U(x)$ must be satisfied?

9.11 Let f be the probability density function of some absolutely continuous distributed random variate X. The *moment generating function* of f is defined as

$$M(t) = \mathbb{E}(e^{tX}) = \int_{-\infty}^{\infty} e^{tx} f(x) \, dx .$$

Show that $M'(0) = \mathbb{E}(X)$, i.e., the expectation of X.

9.12 The gamma function $\Gamma(z)$ is an extension of the factorial function. That is, if n is a positive integer, then

$$\Gamma(n) = (n-1)!$$

For positive real numbers z it is defined as

$$\Gamma(z) = \int_0^\infty t^{z-1} e^{-t} \, dt .$$

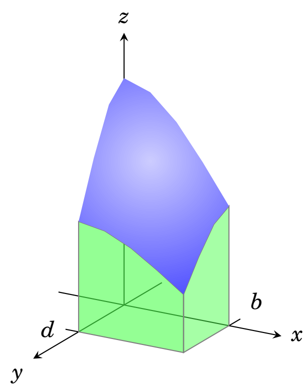
(a) Use integration by parts and show that

$$\Gamma(z+1) = z \, \Gamma(z) .$$

(b) Compute $\Gamma'(z)$ by means of Leibniz's formula.

10

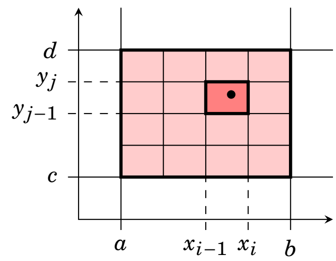
Multiple Integrals



What is the volume *of a smooth mountain?*

The idea of Riemann integration can be extended to the computation of volumes under the graph of bivariate and multivariate functions. However, difficulties arise as the domain of such functions are not simple intervals in general but can be irregular shaped regions.

10.1 The Riemann Integral



Let us start with the simple case where the domain of some bivariate function is the Cartesian product of two closed intervals, i.e., a rectangle

$$R = [a,b] \times [c,d] = \{(x,y) \in \mathbb{R}^2 : a \le x \le b,\ c \le y \le d\}\,.$$

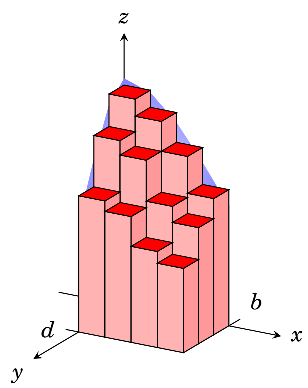
Analogously to Section 9.2 we partition R into rectangular subregions

$$R_{ij} = [x_{i-1}, x_i] \times [y_{j-1}, y_j] \qquad \text{for } 1 \le i \le n \text{ and } 1 \le j \le k$$

where $a = x_0 < x_1 < \ldots < x_n = b$ and $c = y_0 < y_1 < \ldots < y_k = d$.

For $f : R \subset \mathbb{R}^2 \to \mathbb{R}$ we compute $f(\xi_i, \zeta_i)$ for points $(\xi_i, \zeta_j) \in R_{ij}$ and approximate the volume V under the graph of f by the Riemann sum

$$V \approx \sum_{i=1}^{n} \sum_{j=1}^{k} f(\xi_i, \zeta_j)(x_i - x_{i-1})(y_j - y_{j-1})\,,$$



Observe that $(x_i - x_{i-1})(y_j - y_{j-1})$ simply is the area of rectangle R_{ij}. Each term in this sum is just the volume of the bar $[x_{i-1}, x_i] \times [y_{j-1}, y_j] \times [0, f(\xi_i, \zeta_j)]$.

Now suppose that we refine this partition such that the diameter of the largest rectangle tends to 0. If the Riemann sum converges for every such sequence of partitions for arbitrarily chosen points (ξ_i, ζ_i) then this limit is called the **double integral** of f over R and denoted by

$$\iint_R f(x,y)\,dx\,dy = \lim_{\operatorname{diam}(R_{ij}) \to 0} \sum_{i=1}^{n} \sum_{j=1}^{k} f(\xi_i, \zeta_j)(x_i - x_{i-1})(y_j - y_{j-1})\,.$$

Let f and g be integrable functions over some domain D. Let D_1, D_2 be a partition of D, i.e., $D_1 \cup D_2 = D$ and $D_1 \cap D_2 = \emptyset$. Then we find

$$\iint_D (\alpha f(x,y) + \beta g(x,y))\,dx\,dy = \alpha \iint_D f(x,y)\,dx\,dy + \beta \iint_D g(x,y)\,dx\,dy$$

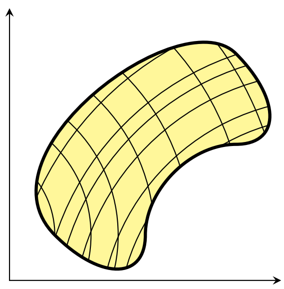
$$\iint_D f(x,y)\,dx = \iint_{D_1} f(x,y)\,dx\,dy + \iint_{D_2} f(x,y)\,dx\,dy$$

$$\iint_D f(x,y)\,dx\,dy \le \iint_D g(x,y)\,dx\,dy$$

if $f(x,y) \le g(x,y)$ for all $(x,y) \in D$

Table 10.1

Properties of double integrals.



It must be noted here that for the definition of the Riemann integral the partition of R need not consist of rectangles. Thus the same idea also works for non-rectangular domains D which may have a quite irregular shape. However, the process of convergence requires more technical details than for the case of univariate functions. For example, the partition has to consist of subdomains D_i of D for which we can determine their areas. Then we have

$$\iint_D f(x,y)\,dx\,dy = \lim_{\text{diam}(D_i)\to 0} \sum_{i=1}^{n} f(\xi_i, \zeta_i) A(D_i)$$

where $A(D_i)$ denotes the area of subdomain D_i. Of course this only works if this limit exists and if it is independent from the particular partition D_i and the choice of the points $(\xi_i, \zeta_i) \in D_i$.

By this definition we immediately get properties that are similar to those of definite integrals, see Table 10.1.

10.2 Double Integrals over Rectangles

As far we only have a concept for the volume below the graph of a bivariate function. However, we also need a convenient method to compute it. So let us again assume that f is a continuous positive function defined on a rectangular domain $R = [a,b] \times [c,d]$. We then write

$$\iint_R f(x,y)\,dx\,dy = \int_a^b \int_c^d f(x,y)\,dy\,dx$$

in analogy to univariate definite integrals.

Let t be an arbitrary point in $[a,b]$. Then let $V(t)$ denote the volume

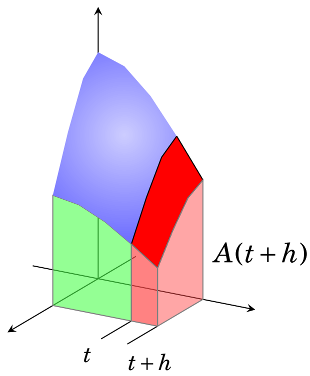
$$V(t) = \int_a^t \int_c^d f(x,y)\,dy\,dx\,.$$

We also obtain a univariate function $g(y) = f(t, y)$ defined on the interval $[c, d]$. Thus

$$A(t) = \int_c^d g(y)\,dy = \int_c^d f(t, y)\,dy$$

is the area of the (2-dimensional) set $\{(t, y, z)\colon 0 \le z \le f(t, y),\ y \in [c, d]\}$. Hence we find

$$V(t+h) - V(t) \approx A(t)\cdot h$$



and consequently

$$V'(t) = \lim_{h\to 0} \frac{V(t+h) - V(t)}{h} = A(t)$$

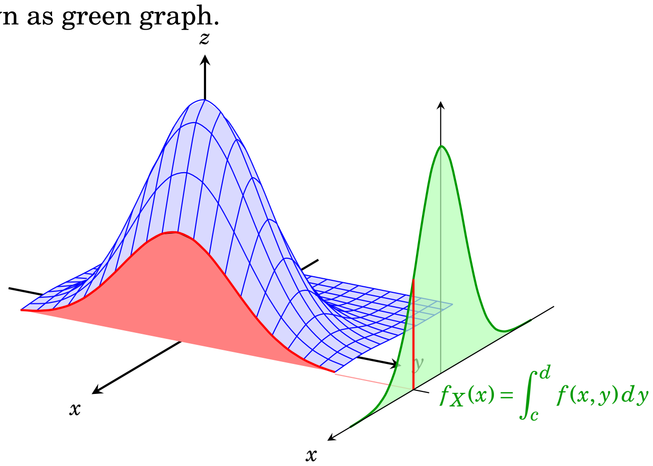
that is, $V(t)$ is an antiderivative of $A(t)$. Here we have used (but did not formally proof) that $A(t)$ is also a continuous function of t.

By this observation we only need to compute the definite integral $\int_c^d f(t, y)\,dy$ for every t and obtain some function $A(t)$. Then we compute the definite integral $\int_a^b A(x)\,dx$. In other words:

For that reason $\iint_R f(x, y)\,dx\,dy$ is called *double integral*.

$$\iint_R f(x, y)\,dx\,dy = \int_a^b \left(\int_c^d f(x, y)\,dy \right) dx\,.$$

This result has the following interpretation in probability theory: Assume that f is a joint density of a bivariate distribution. For fixed x we get a multiple of the density $f(y|X = x)$ for the univariate conditional distribution of Y which is illustrated by the red slice of the graph (blue) of f below. Its integral $f_X(x) = \int_c^d f(x, y)\,dy$ is then the marginal density at x and shown as green graph.



Obviously our arguments remain valid if we exchange the rôles of x and y. Thus

$$\iint_R f(x, y)\,dx\,dy = \int_c^d \left(\int_a^b f(x, y)\,dx \right) dy\,.$$

We summarize our results in the following theorem which we state without a formal proof.

Theorem 10.2

Fubini's theorem. Let $f\colon R=[a,b]\times[c,d]\subset\mathbb{R}^2\to\mathbb{R}$ be a continuous function. Then

$$\iint_R f(x,y)\,dx\,dy=\int_a^b\left(\int_c^d f(x,y)\,dy\right)dx=\int_c^d\left(\int_a^b f(x,y)\,dx\right)dy\,.$$

By this theorem we have the following recipe to compute the double integral of a continuous function $f(x,y)$ defined on the rectangle $[a,b]\times[c,d]$.

(1) Keep y fixed and compute the inner integral w.r.t. x from $x=a$ to $x=b$. This gives $\int_a^b f(x,y)\,dx$, a function of y.

(2) Now integrate $\int_a^b f(x,y)\,dx$ w.r.t. y from $y=c$ to $y=d$ to obtain $\int_c^d\left(\int_a^b f(x,y)\,dx\right)dy$.

Of course we can reverse the order of integration, that is, we first compute $\int_c^d f(x,y)\,dy$ and obtain a function of x which is then integrated w.r.t. x and obtain $\int_a^b\left(\int_c^d f(x,y)\,dy\right)dx$. By Fubini's theorem the results of these two procedures coincide.

Example 10.3

Compute $\displaystyle\int_{-1}^1\int_0^1(1-x-y^2+xy^2)\,dx\,dy$.

SOLUTION. We have to integrate two times.

$$\int_{-1}^1\int_0^1(1-x-y^2+xy^2)\,dx\,dy=\int_{-1}^1\left(x-\frac{1}{2}x^2-xy^2+\frac{1}{2}x^2y^2\Big|_0^1\right)dy$$

$$=\int_{-1}^1\left(\frac{1}{2}-\frac{1}{2}y^2\right)dy=\frac{1}{2}y-\frac{1}{6}y^3\Big|_{-1}^1=\frac{1}{2}-\frac{1}{6}-\left(-\frac{1}{2}+\frac{1}{6}\right)=\frac{2}{3}\,.$$

We can also perform the integration in the reverse order.

$$\int_0^1\int_{-1}^1(1-x-y^2+xy^2)\,dy\,dx=\int_0^1\left(y-xy-\frac{1}{3}y^3+\frac{1}{3}xy^3\Big|_{-1}^1\right)dx$$

$$=\int_0^1\left(1-x-\frac{1}{3}+\frac{1}{3}x-\left(-1+x+\frac{1}{3}-\frac{1}{3}x\right)\right)dx$$

$$=\int_0^1\left(\frac{4}{3}-\frac{4}{3}x\right)dx=\frac{4}{3}x-\frac{4}{6}x^2\Big|_0^1=\frac{2}{3}\,.$$

We obtain the same result by both procedures. ◊

We can extend our results for multivariate functions. Let

$$\Omega=[a_1,b_1]\times\cdots\times[a_n,b_n]=\left\{(x_1,\ldots,x_n)\in\mathbb{R}^n : a_i\le x_i\le b_i,\ i=1,\ldots,n\right\}$$

be the Cartesian product of closed intervals $[a_1,b_1],\ldots,[a_n,b_n]$. We call Ω an ***n*-dimensional rectangle** .

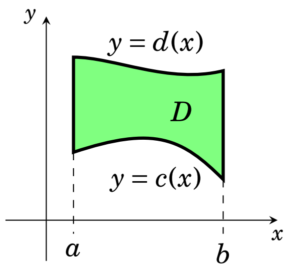
n-dimensional rectangle

If $f:\Omega\to\mathbb{R}$ is a continuous function, then the **multiple integral** of f over Ω is defined as

$$\iint\cdots\int_\Omega f(x_1,\dots,x_n)\,dx_1\dots dx_n = \int_{a_1}^{b_1}\left(\int_{a_2}^{b_2}\cdots\left(\int_{a_n}^{b_n} f(x_1,\dots,x_n)\,dx_n\right)\dots dx_2\right)dx_1\,.$$

It is important to note that the inner integrals are evaluated at first.

10.3 Double Integrals over General Domains



Consider now a domain $D\subseteq\mathbb{R}^2$ defined as

$$D=\{(x,y)\colon a\le x\le b,\ c(x)\le y\le d(x)\}$$

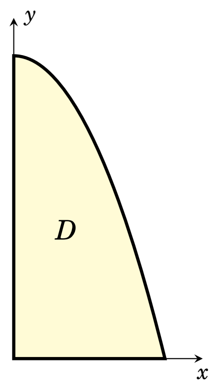
for two functions $c(x)$ and $d(x)$. Let $f(x,y)$ be a continuous function defined over D. As in the case of rectangular domains we can keep x fixed and compute the area

$$A(x)=\int_{c(x)}^{d(x)} f(x,y)\,dy\,.$$

We then can argue in the same way that the volume is given by

$$\iint_D f(x,y)\,dy\,dx=\int_a^b A(x)\,dx=\int_a^b\left(\int_{c(x)}^{d(x)} f(x,y)\,dy\right)dx\,.$$

Example 10.4



Let $D=\{(x,y)\colon 0\le x\le 2,\ 0\le y\le 4-x^2\}$ and let $f(x,y)=x^2y$ be defined on D. Compute $\iint_D f(x,y)\,dy\,dx$.

SOLUTION.

$$\begin{aligned}\iint_D f(x,y)\,dy\,dx &= \int_0^2\int_0^{4-x^2} x^2y\,dy\,dx=\int_0^2\left(\int_0^{4-x^2} x^2y\,dy\right)dx\\ &=\int_0^2\left(\frac{1}{2}x^2y^2\Big|_0^{4-x^2}\right)dx=\int_0^2\left(\frac{1}{2}x^2(4-x^2)^2\right)dx\\ &=\int_0^2\frac{1}{2}\left(x^6-8x^4+16x^2\right)dx\\ &=\frac{1}{14}x^7-\frac{8}{10}x^5+\frac{16}{6}x^3\Big|_0^2=\frac{512}{105}\,.\end{aligned}$$ ◊

It might be convenient if we partition the domain of integration D into two disjoint regions A and B, that is, $A\cup B=D$ and $A\cap B=\emptyset$. We then find

$$\iint_{A\cup B} f(x,y)\,dx\,dy=\iint_A f(x,y)\,dx\,dy+\iint_B f(x,y)\,dx\,dy$$

provided that all integrals exist. The formula extend the corresponding rule for univariate integrals in Table 9.12 on page 93. We can extend this formula to overlapping subsets A and B. We then find

$$\iint_{A\cup B} f(x,y)\,dx\,dy = \\ = \iint_A f(x,y)\,dx\,dy + \iint_B f(x,y)\,dx\,dy - \iint_{A\cap B} f(x,y)\,dx\,dy\,.$$

10.4 A "Double Indefinite" Integral

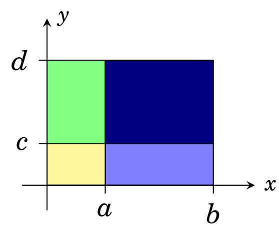
The Fundamental Theorem of Calculus tells us that we can compute a definite integral by the difference of the indefinite integral evaluated at the boundary of the domain of integration. In some sense an equivalent formula exists for double integrals.

Let $f(x,y)$ be an continuous function defined on the rectangle $[a,b]\times[c,d]$. Suppose that $F(x,y)$ has the property that

$$\frac{\partial^2 F(x,y)}{\partial x \partial y} = f(x,y) \quad \text{for all } (x,y) \in [a,b]\times[c,d].$$

Then

$$\int_a^b \int_c^d f(x,y)\,dy\,dx = F(b,d) - F(a,d) - F(b,c) + F(a,c)\,.$$



10.5 Change of Variables

Integration by substitution (see Table 9.17 on page 95) can also be seen as a change of variables. Let $x = g(z)$ where g is a differentiable one-to-one function. Set $z_1 = g^{-1}(a)$ and $z_2 = g^{-1}(b)$. Then

$$\int_a^b f(x)\,dx = \int_{z_1}^{z_2} f(g(z))\cdot g'(z)\,dz\,.$$

That is, instead of expressing f as a function of variable x we introduce a new variable z and a transformation g such that $x = g(z)$. We then integrate $f \circ g$ with respect to z. However, we have to take into account that by this change of variable the domain of integration is deformed. Thus we need the correction factor $g'(z)$.

The same idea of changing variables also works for multivariate functions.

Theorem 10.5

Change of variables in double integrals. Let $f(x,y)$ be a function defined on an open bounded domain $D \subset \mathbb{R}^2$. Suppose that

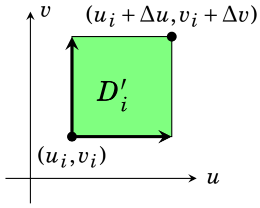
$$x = g(u,v), \quad y = h(u,v)$$

defines a one-to-one $\mathscr{C}^1$ transformation from an open bounded set D' onto D such that the Jacobian determinant $\frac{\partial(g,h)}{\partial(u,v)}$ is bounded and either strictly positive or strictly negative on D'. Then

$$\iint_D f(x,y)\,dx\,dy = \iint_{D'} f(g(u,v),h(u,v)) \left| \frac{\partial(g,h)}{\partial(u,v)} \right| du\,dv\,.$$

This theorem still holds if the set where $\frac{\partial(g,h)}{\partial(u,v)}$ is not bounded or vanishes is a null set, i.e., a set of area 0.

We only give a rough sketch of the proof for this formula. Let $\mathbf{g}$ denote our transformation $(u,v) \mapsto (g(u,v),h(u,v))$. Recall that



$$\iint_D f(x,y)\,dx\,dy = \lim_{\text{diam}(D_i)\to 0} \sum_{i=1}^{n} f(\xi_i,\zeta_i)A(D_i)$$

where the subsets D_i are chosen as the images $\mathbf{g}(D_i')$ of some paraxial rectangle D_i' with vertices

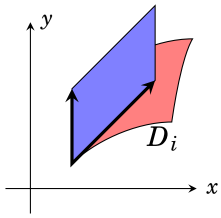
$$(u_i,v_i),\quad (u_i+\Delta u,v_i),\quad (u_i,v_i+\Delta v),\quad \text{and} \quad (u_i+\Delta u,v_i+\Delta v)$$

and $(\xi_i,\zeta_i) = \mathbf{g}(u_i,v_i) \in D_i$. Hence

$$\iint_D f(x,y)\,dx\,dy \approx \sum_{i=1}^{n} f(\xi_i,\zeta_i)A(D_i) = \sum_{i=1}^{n} f(\mathbf{g}(u_i,v_i))A(\mathbf{g}(D_i'))\,.$$

If $\mathbf{g}(D_i')$ were a parallelogram, then we could compute its area by means of the absolute value of the determinant

$$\begin{vmatrix} g(u_i+\Delta u,v_i)-g(u_i,v_i) & g(u_i,v_i+\Delta v)-g(u_i,v_i) \\ h(u_i+\Delta u,v_i)-h(u_i,v_i) & h(u_i,v_i+\Delta v)-h(u_i,v_i) \end{vmatrix}\,.$$



If $\mathbf{g}(D_i')$ is not a parallelogram but Δu is small, then we may use this determinant as an approximation for the area $A(\mathbf{g}(D_i'))$. For small values of Δu we also have

$$g(u_i+\Delta u,v_i)-g(u_i,v_i) \approx \frac{\partial g(u_i,v_i)}{\partial u}\Delta u$$

and thus we find

$$A(\mathbf{g}(D_i')) \approx \left|\begin{vmatrix} \frac{\partial g(u_i,v_i)}{\partial u} & \frac{\partial g(u_i,v_i)}{\partial v} \\ \frac{\partial h(u_i,v_i)}{\partial u} & \frac{\partial h(u_i,v_i)}{\partial v} \end{vmatrix}\right| \Delta u\Delta v = \left|\det(\mathbf{g}'(u_i,v_i))\right|\,\Delta u\Delta v\,.$$

Notice that $\Delta u\Delta v = A(D_i')$ and that we have used the symbol $\frac{\partial(g,h)}{\partial(u,v)}$ to denote the Jacobian determinant. Therefore

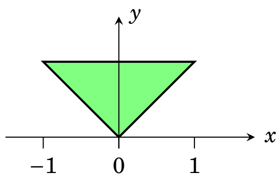
$$\begin{aligned}\iint_D f(x,y)\,dx\,dy &\approx \sum_{i=1}^{n} f(\mathbf{g}(u_i,v_i))\,\left|\det(\mathbf{g}'(u_i,v_i))\right|\,A(D_i') \\ &\approx \iint_{D'} f(g(u,v),h(u,v))\left|\frac{\partial(g,h)}{\partial(u,v)}\right|\,du\,dv\,.\end{aligned}$$

When $\text{diam}(D_i)\to 0$ the approximation errors also converge to 0 and we get the claimed identity. For a stringent proof of Theorem 10.5 the interested reader is referred to literature on advanced calculus.

Example 10.6

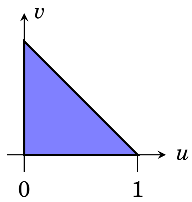
Let $D = \{(x,y)\colon -1 \le x \le 1,\ |x| \le y \le 1\}$ and $f(x,y) = x^2+y^2$ be defined on D. Compute $\iint_D f(x,y)\,dx\,dy$.

SOLUTION. We directly can compute this integral as



$$\begin{aligned}
\iint_D f(x,y)\,dy\,dx &= \int_{-1}^{1}\int_{|x|}^{1} x^2+y^2\,dy\,dx \\
&= \int_{-1}^{0}\int_{-x}^{1} x^2+y^2\,dy\,dx + \int_{0}^{1}\int_{x}^{1} x^2+y^2\,dy\,dx \\
&= \int_{-1}^{0} \left(x^2y+\frac{1}{3}y^3\right)\Big|_{y=-x}^{1} dx + \int_{0}^{1} \left(x^2y+\frac{1}{3}y^3\right)\Big|_{y=x}^{1} dx \\
&= \int_{-1}^{0} x^2+\frac{1}{3}+x^3+\frac{1}{3}x^3\,dx + \int_{0}^{1} x^2+\frac{1}{3}-x^3-\frac{1}{3}x^3\,dx \\
&= \left(\frac{1}{3}x^3+\frac{1}{3}x+\frac{1}{3}x^4\right)\Big|_{-1}^{0} + \left(\frac{1}{3}x^3+\frac{1}{3}x-\frac{1}{3}x^4\right)\Big|_{0}^{1} \\
&= \frac{1}{3}+\frac{1}{3}-\frac{1}{3}+\frac{1}{3}+\frac{1}{3}-\frac{1}{3}=\frac{2}{3}\,.
\end{aligned}$$

We also can first change variables. Let



$$\mathbf{g}(u,v) = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} u \\ v \end{pmatrix}$$

and $D' = \{(u,v)\colon 0 \le u \le 1,\ 0 \le v \le 1-u\}$. Then $\mathbf{g}(D') = D$ and

$$\left|\mathbf{g}'(u,v)\right| = \begin{vmatrix} 1 & -1 \\ 1 & 1 \end{vmatrix} = 2$$

which is constant and thus bounded and strictly positive. Thus we find

$$\begin{aligned}
\iint_D f(x,y)\,dy\,dx &= \iint_{D'} f(\mathbf{g}(u,v))\,|\mathbf{g}'(u,v)|\,dv\,du \\
&= \int_0^1\int_0^{1-u} \left((u-v)^2+(u+v)^2\right)2\,dv\,du \\
&= 4\int_0^1\int_0^{1-u} \left(u^2+v^2\right)\,dv\,du \\
&= 4\int_0^1 \left(u^2v+\frac{1}{3}v^3\right)\Big|_{v=0}^{1-u} du \\
&= 4\int_0^1 \left(-\frac{4}{3}u^3+2u^2-u+\frac{1}{3}\right) du \\
&= 4\left(-\frac{1}{3}u^4+\frac{2}{3}u^3-\frac{1}{2}u^2+\frac{1}{3}u\right)\Big|_0^1 \\
&= 4\left(-\frac{1}{3}+\frac{2}{3}-\frac{1}{2}+\frac{1}{3}\right) = \frac{2}{3}
\end{aligned}$$

which gives (of course) the same result. ◇

Theorem 10.7

Change of variables in multiple integrals. Let $f(\mathbf{x})$ be a function defined on an open bounded domain $D \subset \mathbb{R}^n$. Suppose that $\mathbf{x} = \mathbf{g}(\mathbf{z})$ defines a one-to-one $\mathscr{C}^1$ transformation from an open bounded set $D' \subset \mathbb{R}^n$ onto

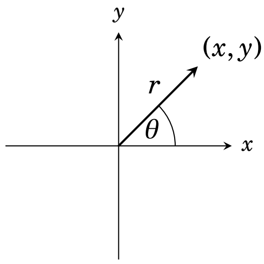
D such that the Jacobian determinant $\frac{\partial(g_1,\dots,g_h)}{\partial(z_1,\dots,z_n)}$ is bounded and either strictly positive or strictly negative on D'. Then

$$\iint_D f(\mathbf{x})d\mathbf{x} = \iint_{D'} f(\mathbf{g}(\mathbf{z}))\left|\frac{\partial(g_1,\dots,g_n)}{\partial(z_1,\dots,z_n)}\right| d\mathbf{z}\,.$$

We also may state this rule analogously to the rule for integration by substitution (Table 9.17)

$$\iint_D f(\mathbf{x})\,d\mathbf{x} = \iint_{D'} f(\mathbf{g}(\mathbf{z}))\left|\det(\mathbf{g}'(\mathbf{z}))\right|\,d\mathbf{z}\,.$$

Polar coordinates are very convenient when we have to deal with circular functions. Thus we represent a point by its distant r from the origin and the angle enclosed by the corresponding vector and the positive x-axis. The corresponding transformation is given by



$$\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{g}(r,\theta) = \begin{pmatrix} r\cos(\theta) \\ r\sin(\theta) \end{pmatrix}$$

where $(r,\theta) \in [0,\infty)\times[0,2\pi)$. It is a $\mathscr{C}^1$ function and its Jacobian determinant is given by

$$\left|\mathbf{g}'(r,\theta)\right| = \begin{vmatrix} \cos(\theta) & -r\sin(\theta) \\ \sin(\theta) & r\cos(\theta) \end{vmatrix} = r\left(\cos^2(\theta)+\sin^2(\theta)\right) = r$$

which is bounded on every bounded domain and it is strictly positive except for the null set $\{(0,\theta)\colon 0\le\theta<2\pi\}$.

Example 10.8

Let $f(x,y) = 1-x^2-y^2$ be a function defined on $D = \{(x,y)\colon x^2+y^2\le 1\}$. Compute $\iint_D f(x,y)\,dx\,dy$.

SOLUTION. A direct computation of this integral is cumbersome:

$$\iint_D (1-x^2-y^2)\,dx\,dy = \int_0^1\int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} (1-x^2-y^2)\,dy\,dx\,.$$

Thus we change to polar coordinates. Then $D' = \{(r,\theta)\colon 0\le r\le 1,\ 0\le\theta< 2\pi\}$ and we find

$$\begin{aligned}\iint_D (1-x^2-y^2)\,dx\,dy &= \int_0^1\int_0^{2\pi}(1-r^2)r\,d\theta\,dr = 2\pi\int_0^1 (r-r^3)\,dr \\ &= 2\pi\left(\frac{1}{2}r^2-\frac{1}{4}r^4\right)\Bigg|_0^1 = \frac{\pi}{2}\,.\end{aligned}$$ ♢

10.6 Improper Multiple Integrals

In Section 9.5 we have extended the concept of integral to unbounded functions or functions with unbounded domains. Using Fubini's theorem the definition of such improper integrals is straight forward by means of limits.

Example 10.9

Compute $\int_0^\infty \int_0^\infty e^{-x^2-y^2}\,dx\,dy$.

SOLUTION. We switch to polar coordinates. $f(x,y) = e^{-x^2-y^2}$ is defined on $D = \{(x,y)\colon x \geq 0,\ y \geq 0\}$. Then $D' = \{(r,\theta)\colon r \geq 0,\ 0 \leq \theta < \pi/2\}$ and we find

$$\begin{aligned}\int_0^\infty \int_0^\infty e^{-x^2-y^2}\,dx\,dy &= \int_0^\infty \int_0^{\pi/2} e^{-r^2} r\,d\theta\,dr = \frac{\pi}{2}\int_0^\infty e^{-r^2} r\,dr \\ &= \lim_{t\to\infty} \frac{\pi}{2}\int_0^t e^{-r^2} r\,dr = \lim_{t\to\infty}\left(-\frac{\pi}{4}e^{-r^2}\right)\Big|_0^t \\ &= \lim_{t\to\infty}\left(-\frac{\pi}{4}\left(e^{-t^2}-1\right)\right) = \frac{\pi}{4}\,. \qquad \diamondsuit\end{aligned}$$

— Exercises

10.1 Evaluate the following double integrals

(a) $\int_0^2 \int_0^1 (2x+3y+4)\,dx\,dy$ (b) $\int_0^a \int_0^b (x-a)(y-b)\,dy\,dx$

(c) $\int_0^1 \int_0^2 (x-y)(x+y)\,dy\,dx$ (d) $\int_0^{1/2} \int_0^{2\pi} y^3 \sin(xy^2)\,dx\,dy$

10.2 Compute

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-x^2-y^2}\,dx\,dy\,.$$

— Problems

10.3 Prove the formula from Section 10.4:

$$\int_a^b \int_c^d f(x,y)\,dy\,dx = F(b,d) - F(a,d) - F(b,c) + F(a,c)\,.$$

where

$$\frac{\partial^2 F(x,y)}{\partial x \partial y} = f(x,y) \quad \text{for all } (x,y) \in [a,b]\times[c,d].$$

HINT: $\int f(x,y)\,dy = \frac{\partial F(x,y)}{\partial x}$

10.4 Let $\Phi(x)$ denote the cumulative distribution function of the (univariate) standard normal distribution. Let

$$f(x,y) = \frac{\sqrt{6}}{\pi}\exp(-2x^2 - 3y^2)$$

be the probability density function of a bivariate normal distribution.

(a) Show that $f(x,y)$ is indeed a probability density function.

HINT: Show that $\iint_{\mathbb{R}^2} f(x,y)dx dy = 1$.

(b) Compute the cumulative distribution function and express the results by means of Φ.
HINT: $F(x,y) = \int_{-\infty}^{x}\int_{-\infty}^{y} \frac{\sqrt{6}}{\pi}\exp(-2s^2-3t^2)\,dt\,ds = \int_{-\infty}^{x}\int_{-\infty}^{y} \frac{\sqrt{2}}{\sqrt{\pi}}\exp(-2s^2)\cdot \frac{\sqrt{3}}{\sqrt{\pi}}\exp(-3t^2)\,dt\,ds$.

10.5 Compute

$$\iint_{\mathbb{R}^2} \exp(-q(x,y))\,dx\,dy$$

where

$$q(x,y) = 2x^2 - 2xy + 2y^2$$

HINT: Observe, that q is a quadratic form with matrix $\mathbf{A} = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$. So change the variables with respect to eigenvectors of $\mathbf{A}$.

11

Differential Equations

We know where we are now and how our system is changing.
Where will we be tomorrow?

11.1 First-Order Ordinary Differential Equations

An **ordinary differential equation (ODE)** is an equation where the unknown is a *function* in one variable and which contains derivatives of this function. It is called of n-th order if it contains a derivative of order n but not higher.

Ordinary means that we have "ordinary" derivatives of a function in one variable in opposition to partial derivatives of functions in two or more variables.

Example 11.1

Three typical examples of first-order differential equations are

$$y' = a\,y\,, \quad y' + a\,y = b\,, \quad y' + a\,y = b\,y^2\,.$$

With suitable chosen constants a and b, these describe exponential (or natural) growth, growth towards a limit and logistic growth. Notice that in this notation the independent variable is omitted. Thus we also can write

$$y'(t) = a\,y(t)\,, \quad y'(t) + a\,y(t) = b\,, \quad y'(t) + a\,y(t) = b\,y^2(t)\,.$$

to stress that y is the dependent variable (i.e., the unknown function) and t is the independent variable. ♢

In economics one often describes variables as a function of time t. Of course our results also hold when the independent variable is denoted by x, u, or whatever.

Example 11.2

$y''(t) + 2y(t) = t^2$ is a (linear) second-order differential equation. ♢

Definition 11.3

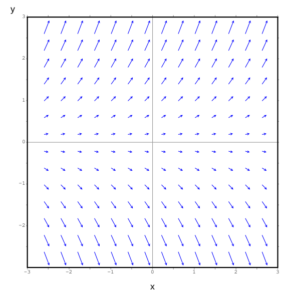
Differential equation. A first-order differential equation is written as

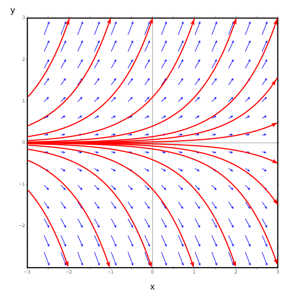
$$y' = F(t, y)$$

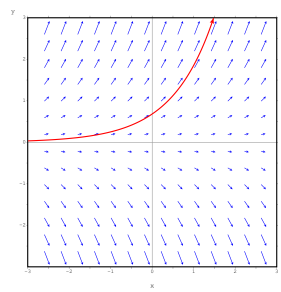
where F is a given function of two variables and $y = y(t)$ is the unknown function. The solution of this equation in an interval I is any differentiable function $y(t)$ defined on I that satisfies this equation for all $t \in I$. The graph of the solution is called **solution curve** or **integral curve**.

Example 11.4

The solution curve of the differential equation $y'(t) = y(t)$ is $y(t) = e^t$. This can easily be verified by differentiating. As for indefinite integrals there exists a family of integral curves: $y(t) = C\,e^t$ is also a solution curve for every $C \in \mathbb{R}$. ♢







A solution curve $y(t)$ has thus the property that the slope $y'(t)$ of the tangent at the point $(t, y(t))$ is just $F(t, y(t))$. We can therefore represent equation $y' = F(t, y)$ by drawing small straight-line segments in the (t, y)-plane. This gives a so called **direction field** (or **slope field**). The task of finding a solution curve is thus equivalent to finding a path in the (t, y)-plane with these line segments as tangents.

Notice that we obtain a set of integral curves which is called the **general solution** of the given ordinary differential equation. However, if we also add an **initial value**, $y(t_0) = y_0$, we only find one solution called the **particular solution** of the so called **initial value problem**,

$$\begin{cases} y' = F(t, y), \\ y(t_0) = y_0. \end{cases}$$

In economic models it is usually not possible to find an explicit solution of the differential equations. This is in particular the case when F is not given explicitly but only some of its properties. Nevertheless, we may still try to answer the following questions:

1. Is there a solution curve for the given equation?
2. Is the solution curve uniquely defined?
3. What can we say about the properties of the solution curve? (*Quality theory*)

Despite this fact we first want to discuss a method for finding explicit solutions. It is applied to some important families of differential equations where we derive closed-form solutions.

There also exist methods to solve differential equations numerically and thus can help us to gain some insight in our models.

11.2 A Simple Economic Model

Example 11.5

Domar growth model. Let $K(t)$ denote the capital stock at time t. Then

$$\frac{dK}{dt} = I$$

where $I(t)$ is the rate of investment flow. In the Domar growth model any change in I will produce a dual effect:

1. An increase in $I(t)$ raises the rate of income flow $Y(t)$:

$$\frac{dY}{dt} = \frac{1}{s} \cdot \frac{dI}{dt} \qquad \text{(D1)}$$

for some constant s (which stands for the marginal propensity to save).

2. A change in $I(t)$ also changes the capacity (or potential output flow) $\kappa(t)$. We assume a constant capacity-capital ratio, i.e.,

$$\frac{\kappa(t)}{K(t)} = \varrho \quad (= \text{a constant}) . \tag{D2}$$

In equilibrium capacity $\kappa(t)$ and income flow $Y(t)$ coincide, i.e.,

$$Y = \kappa . \tag{DE}$$

Now let us assume that our model is in equilibrium at time $t = 0$. Which flow of investment causes our model to remain in equilibrium for all times $t > 0$?

SOLUTION. If equation (DE) holds for every t, then this also holds for the respective derivatives:

$$\frac{dY}{dt} = \frac{d\kappa}{dt}$$

and similarly (D2) implies

$$\frac{d\kappa}{dt} = \varrho \frac{dK}{dt} = \varrho I .$$

Substituting this into (D1) gives

$$\frac{dI}{dt} \cdot \frac{1}{s} = \frac{dY}{dt} = \frac{d\kappa}{dt} = \varrho I$$

or

$$I'(t) = \varrho s I(t) \tag{$*$}$$

that is, a first-order differential equation. Rewriting this equation gives

$$\frac{1}{I}\frac{dI}{dt} = \varrho s$$

which must be satisfied for all times $t > 0$. Thus

$$\int \frac{1}{I}\frac{dI}{dt} dt = \int \varrho s \, dt .$$

For the l.h.s. of this equation we find by substituting $I = I(t)$ and using $dI = \frac{dI}{dt} dt$

$$\int \frac{1}{I}\frac{dI}{dt} dt = \int \frac{1}{I} dI = \ln I + c_1 = \ln I(t) + c_1 .$$

For the r.h.s. we obtain

$$\int \varrho s \, dt = \varrho s t + c_2 .$$

Therefore we find

$$\ln I(t) = \varrho s t + (c_2 - c_1) = \varrho s t + c$$

where we set $c = c_2 - c_1$. Consequently,

$$I(t) = e^{\varrho s t} \cdot e^c = C\, e^{\varrho s t}$$

where we set $C = e^c$. Hence the *general solution* of differential equation (∗) is given by

$$I(t) = C\, e^{\varrho s t}, \qquad C > 0.$$

We can easily verify our solution:

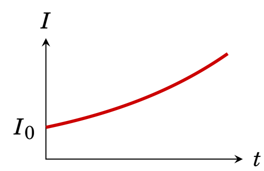
$$\frac{dI}{dt} = \varrho s \cdot C\, e^{\varrho s t} = \varrho s \cdot I(t)\,.$$

Notice that we have obtained an infinite set of solutions as C can be chosen arbitrarily. However, in our model we also have given an initial value for I at time $t = 0$, $I(0) = I_0$. Thus we need the *particular solution* of the *initial value problem*

$$\begin{cases} \frac{dI}{dt} = \varrho s \cdot I, \\ I(0) = I_0. \end{cases}$$

By substituting the initial value we arrive at

$$I_0 = I(0) = C\, e^{\rho s 0} = C$$



and hence

$$I(t) = I_0 e^{\varrho s t}\,.$$

11.3 Separable Equations

The method for finding the solution of the Domar model can be generalized. Suppose that the differential equation can be represented by

$$y'(t) = f(t) \cdot g(y)\,.$$

When $g(y) \neq 0$ this can be rewritten as

$$\frac{dy}{dt} = f(t) \cdot g(y) \quad \Longleftrightarrow \quad \frac{dy}{g(y)} = f(t)\,dt\,.$$

Integration on both sides yields

$$\int \frac{1}{g(y)}\,dy = \int f(t)\,dt + c\,.$$

Evaluating the two integrals gives a solution for the differential equation (possibly in implicit form).

If $g(a) = 0$, then there also is the constant solution $y(t) = a$.

Example 11.6

Solve the differential equation $y' + t y^2 = 0$.

SOLUTION. Separating the variables gives

$$\frac{dy}{dt} = -t y^2 \quad \Rightarrow \quad -\frac{dy}{y^2} = t\,dt\,.$$

Integration yields

Recall that c is just any real number and can equally well be written as $\frac{1}{2}c$.

$$-\int \frac{dy}{y^2} = \int t\,dt + \frac{1}{2}c \quad \Rightarrow \quad \frac{1}{y} = \frac{1}{2}t^2 + \frac{1}{2}c$$

and thus the general solution is

$$y(t) = \frac{2}{t^2 + c}\,. \qquad \diamondsuit$$

Example 11.7

Solve the initial value problem $y' + t y^2 = 0,\ y(0) = 1$.

SOLUTION. By Example 11.6 above the general solution of this ODE is $y(t) = \frac{2}{t^2+c}$. Hence we find

$$1 = y(0) = \frac{2}{0^2 + c} \quad \Rightarrow \quad c = 2$$

and thus the particular solution of the initial value problem is

$$y(t) = \frac{2}{t^2 + 2}\,. \qquad \diamondsuit$$

In the above example it is possible to express y as an explicit function of t. Which conditions guarantee that this is always possible?

Let $G(y) = \int \frac{1}{g(y)}dy$ and $F(t) = \int f(t)\,dt$. Then $G'(y) = \frac{1}{g(y)}$ and $F'(t) = f(t)$ and the solution of $y'(t) = f(t)g(y)$ is given by

$$G(y) = F(t) + c\,.$$

Now if $g(y) \neq 0$ and continuous for all y, then $g(y)$ cannot change sign and $G(y)$ is either strictly monotonically increasing or decreasing. In either case G is invertible and thus there is a solution

$$y = G^{-1}(F(t) + c)\,.$$

11.4 First-Order Linear Differential Equations

Definition 11.8

A **first-order linear differential equation** is one that can be written as

$$y'(t) + a(t)\,y(t) = s(t)$$

where $a(t)$ and $s(t)$ denote continuous functions.

When $s(x) = 0$ the equation is called a **homogeneous linear differential equation**:

$$y'(t) + a(t)\,y(t) = 0\,. \tag{11.1}$$

It then can be easily solved by separation of variables. We find for the general solution

$$y(t) = C\,e^{-A(t)} \quad \text{where} \quad A(t) = \int a(t)\,dt\,.$$

Example 11.9

Solve the differential equation $y' - 3y = 0$.

SOLUTION. By separation of variables we obtain

$$\frac{dy}{dt} = 3y \quad \Rightarrow \quad \frac{1}{y}dy = 3dt \quad \Rightarrow \quad \ln y = 3t + c$$

Thus the general solution is given by

$$y(t) = C\,e^{3t}\,. \qquad \diamondsuit$$

Example 11.10

Solve the differential equation $y' + 3t^2 y = 0$.

SOLUTION. By separation of variables we obtain

$$\frac{dy}{dt} = -3t^2 y \quad \Rightarrow \quad \frac{1}{y}dy = -3t^2\,dt \quad \Rightarrow \quad \ln y = -t^3 + c$$

Thus the general solution is given by

$$y(t) = C\,e^{-t^3}\,. \qquad \diamondsuit$$

Inhomogeneous linear differential equations have a non-zero right-hand side

$$y'(t) + a(t)\,y(t) = s(t)\,. \tag{11.2}$$

The following simple observation shows us how we can obtain general solutions of inhomogeneous equations.

Lemma 11.11

If y_1 and y_2 are two solutions of the inhomogeneous linear equation (11.2), then $y_1 - y_2$ is a solution of homogeneous equation (11.1).

PROOF. As y_1 and y_2 are two solutions of (11.2) we find for $y = y_1 - y_2$

$$\begin{aligned} y'(t) + a(t)\,y(t) &= (y_1(t) - y_2(t))' + a(t)(y_1(t) - y_2(t)) \\ &= \left[y_1'(t) + a(t)\,y_1(t)\right] - \left[y_2'(t) + a(t)\,y_2(t)\right] \\ &= s(t) - s(t) = 0 \end{aligned}$$

i.e., y is a solution of the homogeneous equation (11.1) as claimed. □

An immediate corollary of this lemma is that the general solution y of the inhomogeneous equation (11.2) can be written as

$$y = y_h + y_p$$

where y_h is the general solution of the corresponding homogeneous equation (11.1) and y_p is a particular solution of (11.2). Thus we need a method to get one particular solution.

Linear Equations with Constant Coefficients

If both $a(t)$ and $s(t)$ are constants we set

$$y_p(t) = \frac{s}{a}\,.$$

Then $y_p'(t) = 0$ and the inhomogeneous equation $y_p'(t) + a\,y_p(t) = s$ is satisfied for all t.

Example 11.12

Solve the differential equation $y' - 3y = 6$.

SOLUTION. The general solution of the homogeneous equation $y' - 3y = 0$ is given by (see Example 11.9)

$$y_h(t) = C\,e^{3t}\,.$$

For y_p we use a constant solution, that is,

$$y_p(t) = \frac{s}{a} = \frac{6}{-3} = -2$$

and thus the general solution of the inhomogeneous equation is given by

$$y(t) = y_h(t) + y_p(t) = C\,e^{3t} - 2\,.$$ ♢

Variation of the Constant

When the coefficient $a(t)$ or $s(t)$ is not constant, then **variation of constants** provides a general method for solving inhomogeneous linear differential equations.

$A(t) = \int a(t)\,dt$

Let $y_h(t) = C\,e^{-A(t)}$ be the general solution of the corresponding homogeneous equation, $y'(t) + a(t)\,y(t) = 0$. Then it is possible to replace constant C by some function $C(t)$ such that

$$y_p(t) = C(t)\,e^{-A(t)}$$

becomes a particular solution for the inhomogeneous equation. Its derivative is then

$$y_p'(t) = \big(C'(t) - a(t)\,C(t)\big)\,e^{-A(t)}$$

where we use the fact that $A'(t) = a(t)$. Inserting this into the differential equation $y_p'(t) + a(t)\,y_p(t) = s(t)$ yields

$$\begin{aligned} \big(C'(t) - a(t)\,C(t)\big)\,e^{-A(t)} + a(t)\,C(t)\,e^{-A(t)} &= s(t) \\ C'(t)\,e^{-A(t)} &= s(t) \\ C'(t) &= s(t)\,e^{A(t)}\,. \end{aligned}$$

Hence

$$C(t) = \int s(t)\,e^{A(t)}\,dt$$

and we find for the particular solution

$$y_p(t) = e^{-A(t)} \int s(t)\,e^{A(t)}\,dt\,.$$

Example 11.13

Solve the inhomogeneous differential equation $y' + \frac{y}{t} = t^2 + 4$.

SOLUTION. The general solution y_h of the corresponding homogeneous equation $y' + \frac{y}{t} = 0$ can be found by separating variables:

Write integration constant as $\ln(c)$.

$$\frac{dy}{dt} + \frac{y}{t} = 0 \quad \Rightarrow \quad \frac{dy}{y} = -\frac{dt}{t} \quad \Rightarrow \quad \ln y_h = -\ln(t) + \ln(c) = \ln\left(\frac{c}{t}\right)$$

and hence

$$y_h(t) = \frac{c}{t}\,.$$

For the particular solution y_p we use the method of variation of the constant:

$$y_p = \frac{C(t)}{t} \quad \Rightarrow \quad y_p' = \frac{C'(t)\,t - C(t)}{t^2}$$

and thus

$$\frac{C'(t)\,t - C(t)}{t^2} + \frac{C(t)}{t \cdot t} = t^2 + 4 \quad \Rightarrow \quad \frac{C'(t)}{t} = t^2 + 4 \quad \Rightarrow \quad C'(t) = t^3 + 4t\,.$$

Integration yields

$$C(t) = \frac{1}{4}t^4 + 2t^2 \quad \text{and} \quad y_p(t) = \frac{C(t)}{t} = \frac{1}{4}t^3 + 2t\,.$$

Consequently the general solution of the inhomogeneous differential equation is given by

$$y(t) = y_h(t) + y_p(t) = \frac{c}{t} + \frac{t^3}{4} + 2t\,. \qquad \diamondsuit$$

We can easily verify our solution:

$$\begin{aligned} y' + \frac{y}{t} &= \left(-\frac{c}{t^2} + \frac{3}{4}t^2 + 2\right) + \frac{\frac{c}{t} + \frac{t^3}{4} + 2t}{t} \\ &= -\frac{c}{t^2} + \frac{3}{4}t^2 + 2 + \frac{c}{t^2} + \frac{1}{4}t^2 + 2 = t^2 + 4\,. \end{aligned}$$

Example 11.14

Dynamics of Market Price. Suppose that for a particular commodity, the demand and supply functions are as follows:

$$\begin{aligned} q_d(t) &= \alpha - \beta p(t) \qquad (\alpha, \beta > 0) \\ q_s(t) &= -\gamma + \delta p(t) \qquad (\gamma, \delta > 0) \end{aligned}$$

The rate of price change $p'(t)$ at any moment is proportional to the excess demand $(q_d - q_s)$:

$$\frac{dp}{dt} = j(q_d(t) - q_s(t)) \qquad (j > 0)$$

where j represents the adjustment coefficient.
Which time path $p(t)$ describes the price when $p(0) = p_0$?

SOLUTION. Obviously, the price does not change if and only if $q_d(t) = q_s(t)$, i.e., if we have equilibrium. A straight-forward computation gives the equilibrium price

$$p^* = \frac{\alpha + \gamma}{\beta + \delta}$$

If our model is not in equilibrium we find

$$\frac{dp}{dt} = j(q_d - q_s) = j(\alpha - \beta p - (-\gamma + \delta p)) = j(\alpha + \gamma) - j(\beta + \delta)p$$

and thus

$$\frac{dp}{dt} + j(\beta + \delta)p(t) = j(\alpha + \gamma)$$

i.e., p is described by an inhomogeneous linear differential equation with constant coefficients.

General solution p_h of the homogeneous equation $p' + j(\beta + \delta)p = 0$:

$$\frac{dp}{p} = -j(\beta + \delta)\,dt \quad \Rightarrow \quad \ln p_h = -j(\beta + \delta)t + c$$

and thus

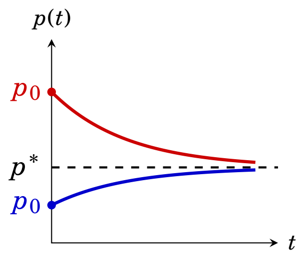
$$p_h(t) = C\,e^{-j(\beta+\delta)t}\,.$$

Particular solution p_p of the inhomogeneous equation:

$$p_p(t) = \frac{j(\alpha + \gamma)}{j(\beta + \delta)} = \frac{\alpha + \gamma}{\beta + \delta} = p^* \quad \text{(= constant).}$$

Therefore

$$p(t) = p_h(t) + p_p(t) = C\,e^{-j(\beta+\delta)t} + p^*\,.$$



We get the solution for the initial value $p(0) = p_0$ by

$$p_0 = p(0) = C\,e^0 + p^* \qquad \text{and thus} \qquad C = p_0 - p^*\,.$$

Consequently

$$p(t) = p^* + (p_0 - p^*)\,e^{-j(\beta+\delta)t}$$ ♢

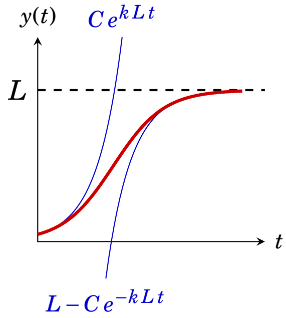
Point p^* in Example 11.14 is a called *equilibrium state* or *stationary state*. We also can see that every solution curve (with initial value $p(0) > 0$) approaches this point as time t tends to ∞. Hence p^* is called *globally asymptotically stable*.

11.5 Logistic Differential Equation

Definition 11.15

A **Logistic differential equation** has the form

$$y'(t) - k\,y(t)(L - y(t)) = 0 \qquad \text{where } k > 0 \text{ and } 0 \le y(t) \le L.$$



Before computing an explicit solution of this equation let us first try some heuristics. For "small" values of y the equation looks like $y'(t) - kL\,y(t) \approx 0$ and the solution is similar to that of a homogeneous linear equation, $y(t) \approx C\,e^{kLt}$. For values of y close to L we find $y'(t) + kL\,y(t) \approx kL^2$ and the solution is similar to that of an inhomogeneous linear equation, $y(t) \approx e^{-kLt}\left(-C + L\,e^{kLt}\right) = L - C\,e^{-kLt}$.

We can solve this equation by separating variables:

$$\frac{dy}{dt} = k\,y(L-y) \quad \Rightarrow \quad \frac{dy}{y(L-y)} = k\,dt$$

The l.h.s. of this equation can be integrated by substitution:

$$z = \frac{L-y}{y} = \frac{L}{y} - 1 \quad \Rightarrow \quad dz = -\frac{L}{y^2}dy$$

$$\begin{aligned}
\int \frac{dy}{y(L-y)} &= -\frac{1}{L}\int \frac{y}{L-y}\cdot\left(-\frac{L}{y^2}\right)dy = -\frac{1}{L}\int \frac{1}{z}\,dz \\
&= -\frac{1}{L}\ln|z| + c = -\frac{1}{L}\ln\left|\frac{L-y}{y}\right| + c \\
&= -\frac{1}{L}\ln\left(\frac{L-y}{y}\right) + c\,.
\end{aligned}$$

Integration of the r.h.s. gives

$$\int k\,dt = k\,t + c$$

and thus

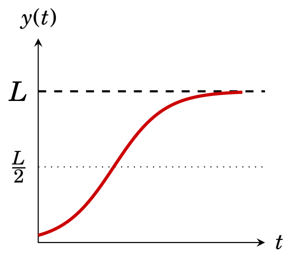
$$-\frac{1}{L}\ln\left(\frac{L-y}{y}\right) = k\,t + c\,.$$

It remains to express y as an explicit function of t:

$$\begin{aligned}
-\frac{1}{L}\ln\left(\frac{L-y}{y}\right) = k\,t + c \quad &\Leftrightarrow \quad \frac{L-y}{y} = e^{-Lkt-Lc} = C\,e^{-Lkt} \\
&\Leftrightarrow \quad L = y\left(1 + C\,e^{-Lkt}\right)
\end{aligned}$$

and consequently

$$y(t) = \frac{L}{1 + C\,e^{-Lkt}}\,.$$



Notice that all solutions of the logistic differential equations (with $0 \le y(t) \le L$) have an inflection point when $y = \frac{L}{2}$. This can be easily seen as the r.h.s. of $\frac{dy}{dt} = k\,y(L-y)$ has a maximum at $y = \frac{L}{2}$.

Example 11.16

There is a serious outbreak of a flu epidemic in a town with 8100 inhabitants. When the disease was first diagnosed, 100 people were infected. 20 days later there are already 1000 infected people. It is expected that eventually all inhabitants will get this flu. How could we describe the spread of the epidemic?

SOLUTION. Let $q(t)$ denote the number of people that got the flu up to time t. As flu is passed by personal contacts and since there are at most $L = 8100$ concerned people we assume a logistic growth model. It has general solution

$$q(t) = \frac{8100}{1 + C e^{-8100kt}}$$

where parameter k and constant C remain to be determined. By means of our initial values we find

$$q(0) = 100 \quad \Rightarrow \quad \frac{8100}{1+C} = 100 \quad \Rightarrow \quad C = 80$$

$$q(20) = 1000 \quad \Rightarrow \quad \frac{8100}{1 + 80\, e^{-8100\cdot 20k}} = 1000 \quad \Rightarrow \quad k = 0.00001495$$

Hence the spread of the epidemic can be described by the function

$$q(t) = \frac{8100}{1 + 80\, e^{-0,121t}}\,. \qquad \diamondsuit$$

11.6 Phase Diagrams and Stability

Definition 11.17

Many differential equations in economics can be expressed in the form

$$y' = F(y)$$

i.e., the r.h.s. does not explicitly depend on the independent variable t. Such an equation is called an **autonomous differential equation**.

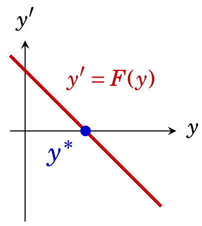
To examine the properties of solutions of such equations, it is useful to study the so called **phase diagram**. This is obtained by plotting y' against y, i.e., by drawing the graph of $F(y)$.

Example 11.18

Linear differential equations with constant coefficients

$$y' = a\,y + s = F(y)$$

are the simplest case of autonomous differential equations.



Now assume that y^* is a root of F, i.e., $F(y^*) = 0$. If $y(t) = y^*$ for some t_0, then $y'(t_0) = 0$ and the integral curve $y(t) = y^*$ for all $t \geq t_0$. Thus y^* is called a **stationary state (equilibrium state**, **fixed point)** of the differential equation.

If we are, however, in some state $y(t_0) = y_0$ that is not a stationary state, then we have two possibilities:

1. If $y'(t_0) = F(t_0) > 0$, then $y(t)$ is strictly monotonically increasing in a sufficiently small interval $(t_0 - \varepsilon, t_0 + \varepsilon)$.

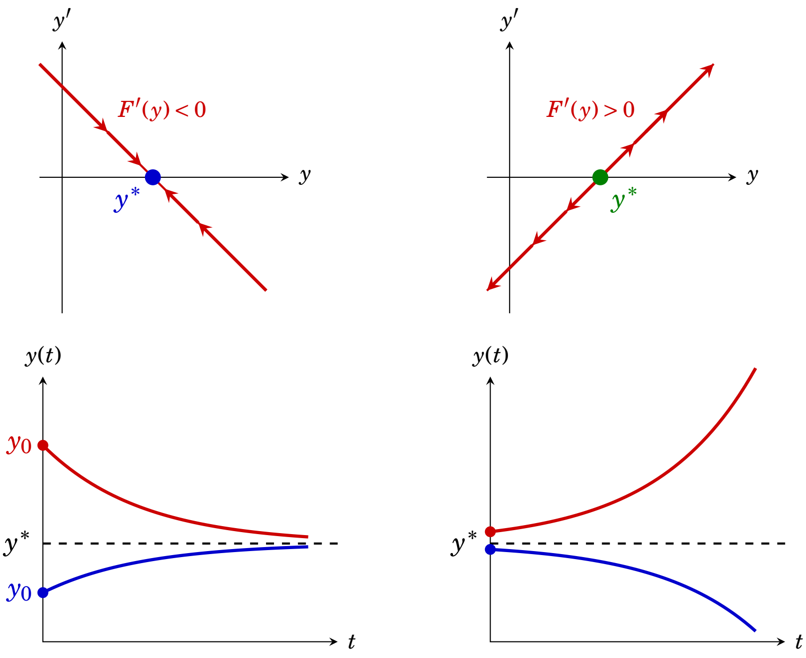
2. If $y'(t_0) = F(t_0) < 0$, then $y(t)$ is strictly monotonically decreasing in a sufficiently small interval $(t_0 - \varepsilon, t_0 + \varepsilon)$.

What happens when we start very close to some stationary point y^*? When we look at Example 11.18 again, then the sign of a and thus of the derivative F' influences the stability of the solution curve.

In the figure on the l.h.s. below the solution curve always moves towards y^*. Thus y^* is called a *locally asymptotically stable equilibrium state*.

y^* is even a *globally* asymptotically stable equilibrium.

In the figure on the r.h.s. below the solution curve always moves away from y^*. Thus y^* is called an *unstable equilibrium state*.



stable equilibrium point unstable equilibrium point

Definition 11.19

Stable and unstable stationary states. A point y^* is called a **stationary state** or **equilibrium point** of the differential equation $y' = F(t, y)$ if $F(t, y^*) = 0$ for all t.

If in addition there exists an $\varepsilon > 0$ such that all solution curves with initial point $y_0 \in B_\varepsilon(y^*)$ converge to y^*, then y^* is called a **locally asymptotically stable equilibrium state**.

If such an ε does not exist, then y^* is called an **unstable equilibrium state**.

Notice that $F'(y) = a < 0$ in the figure on the l.h.s. and $F'(y) = a > 0$ in the figure on the r.h.s. In the case of a non-linear function F it is sufficient to check the sign of $F'(y^*)$. Thus we have the following characterization.

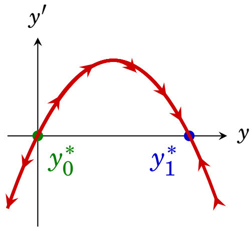
Stability of equilibrium points. Theorem 11.20

(a) If $F(y^*) = 0$ and $F'(y^*) < 0$, then y^* is a *locally asymptotically stable equilibrium*.

(b) If $F(y^*) = 0$ and $F'(y^*) > 0$, then y^* is an *unstable equilibrium*.

Example 11.21

The logistic differential equation $y' = F(y) = k\,y(L-y))$ is an autonomous differential equation. F has two roots $y_0^* = 0$ and $y_1^* = L$. As $F'(y) = L-2y$ we find $F'(0) = L > 0$ and $F(L) = -L < 0$ thus $y_0^* = 0$ is an unstable equilibrium of the logistic differential equation and $y_1^* = L$ is a locally asymptotically stable equilibrium state. ♢



In our examples of autonomous differential equations all nonconstant solution curves never show a local maximum or minimum. Indeed we have the following general result.

Theorem 11.22

If F is a $\mathscr{C}^1$ function, then every solution of the autonomous differential equation $y' = F(y)$ is either constant or strictly monotone on the interval where it is defined.

Theorem 11.23

Suppose that $y = y(t)$ is a solution of $y' = F(y)$, where F is continuous. If $y(t)$ approaches a finite limit y^* as $t \to \infty$, then y^* is an equilibrium state.

11.7 Existence and Uniqueness

So far we have learned some methods for finding explicit solutions of first-order differential equations. We also used phase diagrams to investigate autonomous differential equations. However, until now we do not have any results that guarantee uniqueness of our solution curves or even their existence.

Theorem 11.24

Existence and uniqueness. Consider the first-order differential equation

$$y' = F(t,y)\,.$$

Assume that both $F(t,y)$ and $F_y(t,y)$ are continuous in an open neighborhood in the ty-plane of some point (t_0, y_0). Then there exists exactly one local solution of the equation passing through the point (t_0, y_0).

An immediate corollary of this theorem is, that different solution curves must not intersect each other. If $x(t)$ and $y(t)$ are two solutions to the same differential equation with $x(t_0) = y(t_0)$ for some point t_0, then $x(t) = y(t)$ for all t where these solutions are defined.

Example 11.25

Let $y' = F(t,y) = s(t) - a(t)\,y$ be a linear differential equation where both $s(t)$ and $a(t)$ are continuous. Then for every initial value $y(t_0) = y_0$ there exists a uniquely defined solution in a sufficiently small interval $[t_0, t_0 + \varepsilon)$.

Example 11.26

Let $y' = F(t,y) = f(t)g(y)$ be a separable differential equation. Then existence and uniqueness are ensured if $f(t)$ is continuous and $g(y)$ is continuously differentiable.

Notice that Theorem 11.24 does not give us any information about the length of the interval $[t_0, t_1]$ in which the solution $y(t)$ is defined.

Theorem 11.27

Existence and uniqueness. Consider the initial value problem

$$y' = F(t,y)\,, \qquad y(t_0) = y_0\,. \tag{$*$}$$

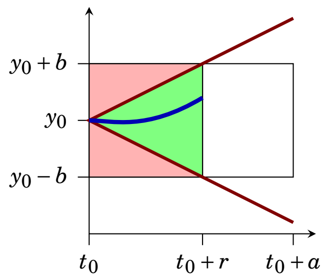
Assume that both $F(t,y)$ and $F_y(t,y)$ are continuous over the rectangle

$$R = \{(t,y) \colon |t - t_0| \le a,\ |y - y_0| \le b\}$$

and let

$$M = \max_{(t,y)\in R} |F(t,y)|\,, \qquad r = \min(a, b/M)\,.$$

Then ($*$) has a unique solution $y(t)$ on $(t_0 - r, t_0 + r)$ and we have $|y(t) - y_0| \le b$ in this interval.



The condition in Theorems 11.24 and 11.27 that $F_y(t,y)$ is continuous can be replaced by a weaker one. Indeed, the existence of partial derivative F_y in an open rectangle around (t_0, y_0) is not required. We only need that for fixed $F(t,y)$ does not change "too much" when we vary argument y. It is sufficient that $F(t,y)$ is continuous in t and **locally Lipschitz continuous** in y, i.e., for each (t,y) there exists an open rectangle R and a constant L such that $|F(t,y_1) - F(t,y_2)| \le L|y_1 - y_2|$ whenever (t,y_1) and (t,y_2) belong to R (Picard–Lindelöf theorem).

When we drop this weaker condition, then a solution will still exist for the differential equation but may not be unique, see Problem 11.10.

— Exercises

11.1 Compute the general solutions of the following ordinary differential equations as well as the particular solution of the corresponding initial value problems with initial values $y(1) = 1$.

(a) $y' - k\frac{y}{t} = 0$

(b) $t\,y' - (1 + y) = 0$

(c) $y' = t\,y$

(d) $y' + e^y = 0$

(e) $y' = y^2$

(f) $y' = \sqrt{t^3 y}$

11.2 Solve the initial value problem

$$y'(t) + 6\,y(t) + e^t = 0\,, \qquad y(0) = 1\,.$$

11.3 Let $U(x)$ be a utility function with the property that the marginal utility $U'(x)$ is indirectly proportional to $U(x)$.

HINT: y is indirectly proportional to z if $y = \alpha\frac{1}{z}$ for some $\alpha \in \mathbb{R}$.

(a) Formulate an ordinary differential equation that describes this property.

(b) Compute the general solution for this differential equation.

(c) Find a sensible particular solution.
(Which value do you suggest for $U(0)$?)

11.4 Suppose that for a particular commodity, the demand and supply functions are as follows:

$$q_d(t) = \alpha - \beta p(t) + \nu p'(t)$$
$$q_s(t) = -\gamma + \delta p(t)$$

where $\alpha, \beta, \gamma, \delta, \nu > 0$ are constants. Assume that the rate of price change $p'(t)$ at any moment is proportional to the excess demand $(q_d - q_s)$:

$$\frac{dp}{dt} = j(q_d(t) - q_s(t))$$

where $j > 0$ represents the adjustment coefficient.
Which time path $p(t)$ describes the price when $p(0) = p_0$?
How does the model differ from Example 11.14?

11.5 The expected number of consumers of a new commodity is 96 000. When a marketing campaign starts there are already 4 000 people who know the product. After two months this number has increased to 12 000.

Assume that $A(t)$ is the number of people that already know the product at time t can be modeled by a logistic differential equation.

(a) Compute function $A(t)$.

(b) How many people know the product after 6 months?

(c) The marketing campaign should be stopped if two third of all potential consumers know the product. How long runs the campaign?

11.6 Consider the following linear first-order differential equation

$$y'(t) + a\,y(t) = s$$

where a and s are constants. Derive a closed form solution of this equation.

11.7 Consider the following initial value problem

$$y'(t) + a\,y(t) = s, \qquad y(t_0) = y_0$$

where a, s, and y_0 are constants. Derive a closed form solution of this equation.

— Problems

11.8 Let y^* be a stationary state of the autonomous differential equation

$$y' = F(y)$$

where F is a continuously differentiable function.

Are the conditions in Theorem 11.20 necessary or sufficient or both or neither for y^* being an unstable or asymptotically stable equilibrium state?

Find examples for the following cases or argue why the respective case cannot happen.

(a) $F'(y^*) = 0$ and y^* is an unstable equilibrium.

(b) $F'(y^*) < 0$ and y^* is an unstable equilibrium.

(c) $F'(y^*) = 0$ and y^* is a locally asymptotically stable equilibrium.

(d) $F'(y^*) > 0$ and y^* is a locally asymptotically stable equilibrium.

11.9 The Solow growth model is based on the differential equation

$$k' = s\,f(k) - \lambda k$$

where $k = k(t)$ denotes capital per worker, $s > 0$ denotes the constant rate of saving, f is a production function, and $\lambda > 0$ denotes the constant proportional rate of growth of the number of workers. Assume that $f(k) = k^\alpha$ for some $\alpha \in (0,1)$.

(a) Sketch the phase diagram for this differential equation.

(b) Compute the stationary states and estimate whether these are unstable or asymptotically stable.

(c) If there is a locally asymptotically stable equilibrium, is it also *globally* asymptotically stable?

11.10 Consider the following initial value problem:

$$y'(t) = 2\sqrt{y}\,, \qquad y(0) = 0\,.$$

For $a > 0$ let

$$y_a(t) = \begin{cases} 0 & \text{for } t \le a, \\ (t-a)^2 & \text{for } t > a. \end{cases}$$

(a) Solve the initial value problem by separation of the variables.

(b) Show that $y_a(t)$ is differentiable. Compute its derivative.

(c) Show that $y_a(t)$ is a solution of the initial value problem.

(d) Why is the solution to this initial value problem not unique?

(e) Is the solution unique if we change the initial value to $y(0) = 1$?

12
Second-Order Differential Equations

How can we model the pork cycle?

12.1 Second-Order Differential Equations

A **second-order ordinary differential equation** is written as   Definition 12.1

$$y'' = F(t, y, y')$$

where F is a given function of three variables and $y = y(t)$ is the unknown function. Second order differential equations are inevitable for modeling phenomenons like the pork cycle. However, analyzing second-order differential equations or even finding explicit solutions is more challenging than for first-order differential equations. Thus we restrict our interest on linear equations.

12.2 Second-Order Linear Differential Equations

A **second-order linear differential equation** is one that can be written as   Definition 12.2

$$y''(t) + a_1(t)\,y'(t) + a_2(t)\,y(t) = s(t)\,. \tag{12.1}$$

Analogously to first-order differential equations we also may give initial values and find a solution curve for this **initial value problem**. However, for second-order differential equations it is necessary to specify *two* such values; usually a value for the function at some time t_0, $y(t_0) = y_0$, as well as its first derivative, $y'(t_0) = y'_0$.

Existence and uniqueness. Suppose that $a_1(t)$, $a_2(t)$, and $s(t)$ are all continuous functions on an open interval (α, β), not necessarily finite.   Theorem 12.3

Let y_0 and y'_0 be two given numbers and $t_0 \in (\alpha, \beta)$. Then differential equation (12.1) has exactly one solution $y(t)$ on the interval (α, β) that satisfies $y(t_0) = y_0$ and $y'(t_0) = y'_0$.

Homogeneous linear equations (i.e., where $s(t) = 0$ in (12.1)) have the nice property that any linear combination of two solutions is again a solution of the equation. Indeed, if both u_1 and u_2 satisfy the homogeneous equation

The set of all solutions forms a 2-dimensional vector space.

$$y''(t) + a_1(t)\,y'(t) + a_2(t)\,y(t) = 0 \tag{12.2}$$

then we find for $u = C_1 u_1 + C_2 u_2$ where $C_1, C_2 \in \mathbb{R}$,

$$\begin{aligned} u'' + a_1 u' + a_2 u & \\ &= (C_1 u_1 + C_2 u_2)'' + a_1 (C_1 u_1 + C_2 u_2)' + a_2 (C_1 u_1 + C_2 u_2) \\ &= C_1 \left[u_1'' + a_1 u_1' + a_2 u_1\right] + C_2 \left[u_2'' + a_1 u_2' + a_2 u_2\right] \\ &= 0 + 0 = 0\,. \end{aligned}$$

That is, $u = C_1 u_1 + C_2 u_2$ is again a solution of (12.2). A consequence of Theorem 12.3 is that any solution u of (12.2) can be expressed as a linear combination of these two solutions u_1 and u_2 provided that they are linearly independent.

Theorem 12.4

The *homogeneous* differential equation

$$y''(t) + a_1(t)\,y'(t) + a_2(t)\,y(t) = 0$$

has general solution

$$y(t) = C_1 u_1(t) + C_2 u_2(t)$$

where $u_1(t)$ and $u_2(t)$ are two independent solutions and C_1 and C_2 are arbitrary constants.

For **inhomogeneous linear equations** we find a result analogous to Lemma 11.11 on p. 118. Assume that v_1 and v_2 are solutions of (12.1). Then we find for their difference, $v = v_1 - v_2$,

$$\begin{aligned} v'' + a_1 v' + a_2 v &= (v_1 - v_2)'' + a_1 (v_1 - v_2)' + a_2 (v_1 - v_2) \\ &= (v_1'' + a_1 v_1' + a_2 v_1) - (v_2'' + a_1 v_2' + a_2 v_2) \\ &= s - s = 0\,. \end{aligned}$$

That is, $v = v_1 - v_2$ is a solution of the corresponding homogeneous differential equation.

Theorem 12.5

The *inhomogeneous* differential equation

$$y''(t) + a_1(t)\,y'(t) + a_2(t)\,y(t) = s(t)$$

has general solution

$$y(t) = C_1 u_1(t) + C_2 u_2(t) + u_p(t)$$

where $C_1 u_1(t) + C_2 u_2(t)$ is the general solution of the corresponding homogeneous equation and $u_p(t)$ is any **particular** solution of the inhomogeneous equation.

12.3 Constant Coefficients

Homogeneous Differential Equation

Let us consider *homogeneous* linear differential equations

$$y''(t) + a_1\, y'(t) + a_2\, y(t) = 0 \tag{12.3}$$

where the coefficients a_1 and a_2 are constants. Based on our experiences with first-order linear differential equations we try the ansatz

$$y(t) = C\, e^{\lambda t}$$

for some constants C and λ. Computing its first and second derivatives and substituting into equation (12.3) yields

$$y'(t) = \lambda\, C\, e^{\lambda t} \quad \text{and} \quad y''(t) = \lambda^2\, C\, e^{\lambda t}$$

and thus

$$C\, e^{\lambda t}(\lambda^2 + a_1\,\lambda + a_2) = 0\,.$$

Consequently $y(t)$ is a solution of (12.3) if and only if

$$\lambda^2 + a_1 \lambda + a_2 = 0\,. \tag{12.4}$$

Equation (12.4) is called the **characteristic equation** of the homogeneous differential equation (12.3). Its two solutions are given by

$$\lambda_{1,2} = -\frac{a_1}{2} \pm \sqrt{\frac{a_1^2}{4} - a_2}\,.$$

Generally, there are three different cases to consider. Notice that by Theorem 12.4 have to find two basis functions for the vector space of all solutions.

Case $\frac{a_1^2}{4} - a_2 > 0$. We have two distinct real roots λ_1 and λ_2 and the general solution is given by

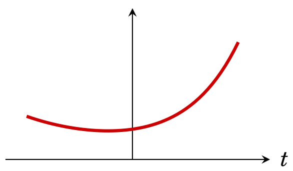
$$y(t) = C_1\, e^{\lambda_1 t} + C_2\, e^{\lambda_2 t}\,, \qquad \text{where} \quad \lambda_{1,2} = -\frac{a_1}{2} \pm \sqrt{\frac{a_1^2}{4} - a_2}\,.$$

Example 12.6

We want to compute the general solution of $y'' - y' - 2y = 0$.

SOLUTION. Using the ansatz $y(t) = e^{\lambda t}$ we find the characteristic equation $\lambda^2 - \lambda - 2 = 0$, with distinct real roots $\lambda_1 = -1$ and $\lambda_2 = 2$. Hence the general solution is given by

$$y(t) = C_1 e^{-t} + C_2 e^{2t}\,.$$ ◇



Case $\frac{a_1^2}{4} - a_2 = 0$. We have only one real (double) root $\lambda = -\frac{a_1}{2}$ and thus one solution is given by

$$y_1(t) = e^{\lambda t} .$$

However we need a second solution as well. Consider

$$y''(t) + a_1 y'(t) + (a_2 - \varepsilon)\, y(t) = 0 \qquad (*)$$

for some $\varepsilon > 0$ with characteristic roots

$$\lambda_{1,2} = -\frac{a_1}{2} \pm \sqrt{\frac{a_1^2}{4} - a_2 + \varepsilon} = \lambda \pm \delta(\varepsilon) .$$

Hence function

$$y_\varepsilon(t) = \frac{e^{(\lambda+\delta(\varepsilon))t} - e^{(\lambda-\delta(\varepsilon))t}}{2\delta(\varepsilon)} = e^{\lambda t}\, \frac{e^{\delta(\varepsilon)t} - e^{-\delta(\varepsilon)t}}{2\delta(\varepsilon)}$$

is a solution of $(*)$. By the Mean Value Theorem there exists a $\tilde{\delta}(\epsilon, t)$ such that

$$y_\varepsilon(t) = e^{\lambda t} t e^{\tilde{\delta}(\varepsilon,t)t} .$$

Observe that $\tilde{\delta}(\varepsilon, t) \in \left[-\delta(\varepsilon)t, \delta(\varepsilon)t\right]$ and that $\delta(\varepsilon) \to 0$ as $\varepsilon \to 0$. Hence we may expect that we may obtain another solution to our original differential differential equation (12.3) by setting $\delta = 0$.

Indeed we claim that

$$y_2(t) = t\, e^{\lambda t}$$

also satisfies (12.3). Differentiating gives

$$y_2'(t) = \lambda\, t\, e^{\lambda t} + e^{\lambda t} = (\lambda t + 1)\, e^{\lambda x} ,$$
$$y_2''(t) = \lambda^2\, t\, e^{\lambda t} + \lambda\, e^{\lambda t} + \lambda\, e^{\lambda t} = (\lambda^2 t + 2\lambda)\, e^{\lambda t} .$$

Substituting into (12.3) yields

$$\begin{aligned}
&\left[(\lambda^2 t + 2\lambda) + a_1(\lambda t + 1) + a_2 t\right] e^{\lambda t} \\
&\quad = \left[\left(\frac{a_1^2}{4} t - a_1\right) + a_1\left(-\frac{a_1}{2} t + 1\right) + a_2\, t\right] e^{\lambda t} \\
&\quad = \left[\frac{a_1^2}{4} t - a_1 - \frac{a_1^2}{2} t + a_1 + \frac{a_1^2}{4} t\right] e^{\lambda t} = 0 , \quad \text{as claimed.}
\end{aligned}$$

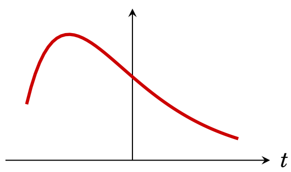
For the first equation we used the fact that $\lambda = -\frac{a_1}{2}$ and for the second equation we used that $\frac{a_1^2}{4} - a_2 = 0$ and thus $a_2 = \frac{a_1^2}{4}$.

Therefore the solution in the case of one double root is given by

$$y(t) = (C_1 + C_2\, t)\, e^{\lambda t} , \qquad \text{where} \quad \lambda = -\frac{a_1}{2} .$$

Example 12.7

We want to compute the general solution of $y'' + 4y' + 4y = 0$.



SOLUTION. Using the ansatz $y(t) = e^{\lambda t}$ we find the characteristic equation $\lambda^2 + 4\lambda + 4 = 0$, with one real double root $\lambda = -2$. Hence the general solution is given by

$$y(t) = (C_1 + C_2 t)e^{-2t} .$$

◇

Case $\frac{a_1^2}{4} - a_2 < 0$. We have two distinct complex roots

$$\lambda = a + bi \quad \text{and} \quad \bar{\lambda} = a - bi$$

where

$$\text{Re}(\lambda) = a = -\frac{a_1}{2} \quad \text{and} \quad \text{Im}(\lambda) = b = \sqrt{\left|\frac{a_1^2}{4} - a_2\right|}$$

and thus the general solution is given by

$$y(t) = \tilde{C}_1 e^{(a+bi)t} + \tilde{C}_2 e^{(a-bi)t} .$$

However, we are only interested in *real* solutions. By Euler's formula (p. 164) we find

$$\begin{aligned} y(t) &= \tilde{C}_1 e^{(a+bi)t} + \tilde{C}_2 e^{(a-bi)t} = e^{at}\left[\tilde{C}_1 e^{bit} + \tilde{C}_2 e^{-bit}\right] \\ &= e^{at}\left[\tilde{C}_1(\cos(bt) + i\sin(bt)) + \tilde{C}_2(\cos(bt) - i\sin(bt))\right] \\ &= e^{at}\left[(\tilde{C}_1 + \tilde{C}_2)\cos(bt) + i(\tilde{C}_1 - \tilde{C}_2)\sin(bt)\right] \\ &= e^{at}\left[C_1\cos(bt) + C_2\sin(bt)\right] \end{aligned}$$

where we set $C_1 = (\tilde{C}_1 + \tilde{C}_2)$ and $C_2 = i(\tilde{C}_1 - \tilde{C}_2)$ for the last equation. We therefore find for the general solution

$$y(t) = e^{at}\left[C_1\cos(bt) + C_2\sin(bt)\right]$$

$$\text{where} \quad a = -\frac{a_1}{2} \quad \text{and} \quad b = \sqrt{\left|\frac{a_1^2}{4} - a_2\right|} .$$

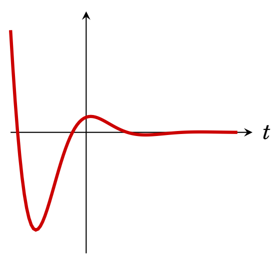
These solutions of second-order linear differential equations describe an oscillating behavior and thus can be used to model cycles in economics.

Example 12.8

We want to compute the general solution of $y'' + y' + y = 0$.

SOLUTION. The characteristic equation $\lambda^2 + \lambda + 1 = 0$ has two complex roots $\lambda_{1,2} = -\frac{1}{2} \pm \frac{\sqrt{3}}{2} i$ with $\text{Re}(\lambda) = a = -\frac{1}{2}$ and $\text{Im}(\lambda) = b = \frac{\sqrt{3}}{2}$. Hence the general solution is given by

$$y(t) = e^{-\frac{1}{2}t}\left[C_1\cos\left(\frac{\sqrt{3}}{2}t\right) + C_2\sin\left(\frac{\sqrt{3}}{2}t\right)\right] . \qquad \diamond$$



Inhomogeneous Differential Equation

Solutions of *inhomogeneous* linear differential equations with constant coefficients a_1, a_2 and s are constants,

$$y''(t) + a_1 y'(t) + a_2 y(t) = s , \qquad (12.5)$$

can be solved by means of Theorem 12.5. Its general solution is given by

$$y = y_h + y_p$$

where y_h is the general solution of the corresponding homogeneous equation (12.3) and y_p is a particular solution of (12.5). As all coefficients are constant we use

$$y_p(t) = \begin{cases} \frac{s}{a_2} & \text{if } a_2 \neq 0, \\ \frac{s}{a_1} t & \text{if } a_2 = 0 \text{ and } a_1 \neq 0. \end{cases}$$

Example 12.9

We want to find the solution of the initial value problem

$$y''(t) + y'(t) - 2y(t) = -10\,, \qquad y(0) = 12, \quad y'(0) = -2\,.$$

SOLUTION. The characteristic equation $\lambda^2 + \lambda - 2 = 0$ has two distinct real roots $\lambda_1 = 1$ and $\lambda_2 = -2$. Thus the solution to the corresponding homogeneous equation is given by

$$y_h(t) = C_1 e^t + C_2 e^{-2t}\,.$$

A particular solution of the inhomogeneous equation is given by

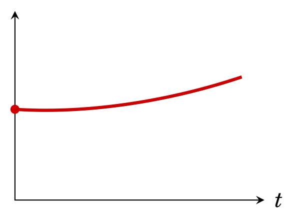
$$y_p(t) = \frac{s}{a_2} = \frac{-10}{-2} = 5$$

and hence the general solution for the inhomogeneous differential equation is

$$y(t) = y_h(t) + y_p(t) = C_1 e^t + C_2 e^{-2t} + 5\,.$$

Substituting the initial values yields the system of linear equations

$$\begin{aligned} 12 &= y(0) = C_1 + C_2 + 5 \\ -2 &= y'(0) = C_1 - 2C_2 \end{aligned}$$



with unique solution $C_1 = 4$ and $C_2 = 3$. Therefore the solution of the initial value problem is given by

$$y(t) = 4e^t + 3e^{-2t} + 5\,.$$ ♢

12.4 Stability for Linear Differential Equations

Let $y'(t) + a\,y(t) = s$ be a first-order linear differential equation with initial value $y(0) = y_0$. It has solution $y(t) = (y_0 - y^*)e^{-at} + y^*$ where $y^* = \frac{s}{a}$ (provided that $a \neq 0$). We have seen in Section 11.6 that y^* is an **equilibrium point** of this differential equation. Moreover, it is **globally asymptotically stable** if and only if $a > 0$, since then the solution curve eventually approaches y^* when t tends to ∞. Notice we also can state this fact in the following way.

Theorem 12.10

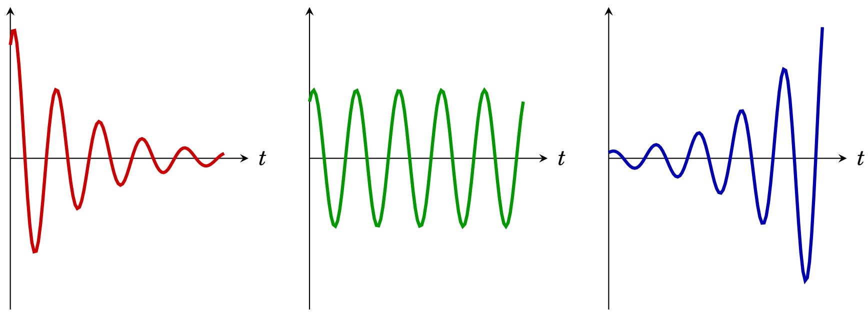
$y^* = \frac{s}{a}$ is a *globally asymptotically stable state* of $y'(t) + a\,y(t) = s$ $(a \neq 0)$ if and only if the root of equation $\lambda + a = 0$ is negative.

For second-order linear differential equations $y''(t)+a_1 y'(t)+a_2 y(t) = s$ we find that $y^* = \frac{s}{a_2}$ is an equilibrium point (provided that $a_2 \neq 0$). In Section 12.3 above we had to distinguish between three cases depending on the roots of the characteristic equation (12.4), $\lambda^2 + a_1\lambda + a_2 = 0$.

If $\frac{a_1^2}{4} - a_2 > 0$ we have solutions $y(t) = C_1 e^{\lambda_1 t} + C_2 e^{\lambda_2 t}$ where $\lambda_{1,2}$ are two distinct real solutions of the characteristic equation. These solutions converge to y^* in general for $t \to \infty$ if and only if both roots are negative.

If $\frac{a_1^2}{4} - a_2 = 0$ we found $y(t) = (C_1 + C_2 t)e^{\lambda t}$ where λ is the real (double) solution root of the characteristic equation. Again these solutions converge to y^* in general for $t \to \infty$ if and only if $\lambda < 0$.

If $\frac{a_1^2}{4} - a_2 < 0$ we have solutions $y(t) = e^{at}\,[C_1 \cos(bt) + C_2 \sin(bt)]$ where $a = -\frac{a_1}{2}$ is the real part of the complex roots of the characteristic equations and $b = \sqrt{\left|\frac{a_1^2}{4} - a_2\right|}$ is the imaginary part. The real part a of root λ controls whether the oscillating solutions is damped or not.



Re(λ) < 0 Re(λ) = 0 Re(λ) > 0

We summarize our observations in the following theorem.

Theorem 12.11 $y^* = \frac{s}{a_2}$ is a *globally asymptotically stable state* of $y''(t)+a_1 y'(t)+a_2 y(t) = s$ $(a_2 \neq 0)$ if and only if the real parts of all roots of the characteristic equation $\lambda^2 + a_1\lambda + a_2 = 0$ are all negative.

— Exercises

12.1 Compute the solution of differential equation $y'' = x^2 + 2x - 5$ with initial values $y(0) = 0$ and $y'(0) = 3$.

12.2 Compute the solution of $y'' + y' - 2y = 3$ with initial values $y(0) = y'(0) = 1$.
Is there an asymptotically stable state for the differential equation?

12.3 Compute the solution of $y'' - 6y' + 9y = 0$ with initial values $y(0) = 2$ and $y'(0) = 0$.
Is there an asymptotically stable state for the differential equation?

12.4 Compute the solution of $y'' + 2y' + 17y = 0$ with initial values $y(0) = 1$ and $y'(0) = 0$.
Sketch (draw) the graph of your solution.
Is there an asymptotically stable state for the differential equation?

12.5 A model by T. Haavelmo leads to an equation of the type

$$p''(t) = \gamma(a - \alpha)\, p(t) + k\,. \qquad (\alpha, \gamma, a, \text{ and } k \text{ are constants})$$

Solve the equation.

13

Systems of Differential Equations

Foxes and rabbits.

13.1 Simultaneous Equations of Differential Equations

We assume that we have two unknown functions y_1 and y_2 where the derivative of each simultaneously depends on both functions, i.e.,

$$\begin{aligned} y_1' &= F_1(t, y_1, y_2)\,, \\ y_2' &= F_2(t, y_1, y_2)\,. \end{aligned} \tag{13.1}$$

Using a vector-valued function this can also be written as

$$\mathbf{y}' = \mathbf{F}(t, \mathbf{y})\,. \tag{13.2}$$

We further assume that all components of $\mathbf{F}$ and all partial derivatives w.r.t. y_i are continuous. A **solution** $\mathbf{y}(t)$ of (13.2) is then a differentiable function where its derivative $\mathbf{y}'$ satisfies this equation for each of its components. Notice that $\mathbf{y}(t)$ is a path in $\mathbb{R}^2$ (or more general in $\mathbb{R}^n$).

One method for finding explicit solutions for a system (13.1) of two differential equation is to reduce it into one second-order differential equation using the following procedure:

Of course we can exchange the rôles of y_1 and y_2.

1. Use $y_1' = F_1(t, y_1, y_2)$ and express y_2 as a function of t, y_1 and y_1': $y_2 = H(t, y_1, y_1')$.
2. Differentiate this equation w.r.t. t and substitute the expressions for y_2 and y_2' into the second equation $y_2' = F_2(t, y_1, y_2)$. We then obtain a second-order differential equation for y_1.
3. Solve this equation and determine $y_1(t)$.
   Then find $y_2(t) = H(t, y_1(t), y_1'(t))$.

Example 13.1

Find the general solution of the system

$$\begin{aligned} y_1' &= 4y_1 + y_2 + 2\,, \\ y_2' &= -2y_1 + y_2 + 8\,. \end{aligned}$$

Also find the particular solution with initial values $y_1(0) = y_2(0) = 0$.

SOLUTION. Solving the first equation for y_2 and differentiation yields

$$\begin{aligned} y_2 &= y_1' - 4y_1 - 2\,, \\ y_2' &= y_1'' - 4y_1'\,. \end{aligned}$$

Substituting these expressions in the second equation gives

$$y_1'' - 4y_1' = -2y_1 + (y_1' - 4y_1 - 2) + 8$$

which simplifies to

$$y_1'' - 5y_1' + 6y_1 = 6\,.$$

Using our methods from Section 12.3 gives the general solution

$$y_1(t) = C_1 e^{2t} + C_2 e^{3t} + 1\,.$$

From $y_2 = y_1' - 4y_1 - 2$ we get

$$\begin{aligned} y_2 &= y_1' - 4y_1 - 2 \\ &= (2C_1 e^{2t} + 3C_2 e^{3t}) - 4(C_1 e^{2t} + C_2 e^{3t} + 1) - 2 \\ &= -2C_1 e^{2t} - C_2 e^{3t} - 6\,. \end{aligned}$$

Thus the general solution of this system of differential equations is given by

$$\begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix} = C_1 e^{2t} \begin{pmatrix} 1 \\ -2 \end{pmatrix} + C_2 e^{3t} \begin{pmatrix} 1 \\ -1 \end{pmatrix} + \begin{pmatrix} 1 \\ -6 \end{pmatrix}.$$

For the particular solution of the initial value problem we have to solve the linear equation

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} = \begin{pmatrix} y_1(0) \\ y_2(0) \end{pmatrix} = C_1 \begin{pmatrix} 1 \\ -2 \end{pmatrix} + C_2 \begin{pmatrix} 1 \\ -1 \end{pmatrix} + \begin{pmatrix} 1 \\ -6 \end{pmatrix}$$

which gives $C_1 = -5$ and $C_2 = 4$ and therefore

$$\begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix} = -5\, e^{2t} \begin{pmatrix} 1 \\ -2 \end{pmatrix} + 4\, e^{3t} \begin{pmatrix} 1 \\ -1 \end{pmatrix} + \begin{pmatrix} 1 \\ -6 \end{pmatrix}.$$ ◇

Notice, however, that this method only works in special cases. As we will see below, it is often easier to analyze a system of first-order differential equation than a single second-order differential equation. Therefore one often uses the reverse of the above procedure to transform a second-order differential equation $y'' = F(t, y, y')$ into a system of two first-order differential equations. Setting $v = y'$ we find

$$\begin{aligned} y' &= v\,, \\ v' &= F(t, y, v)\,. \end{aligned}$$

13.2 Linear Systems with Constant Coefficients

Consider the linear system

$$\begin{pmatrix} y_1' \\ y_2' \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + \begin{pmatrix} s_1 \\ s_2 \end{pmatrix} \quad \text{or equivalently} \quad \mathbf{y}' = \mathbf{A}\mathbf{y} + \mathbf{s} \tag{13.3}$$

where $\mathbf{A}$ and $\mathbf{s}$ are constant.

Example 13.2

The system in Example 13.1 is a linear system:

$$\begin{pmatrix} y_1' \\ y_2' \end{pmatrix} = \begin{pmatrix} 4 & 1 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + \begin{pmatrix} 2 \\ 8 \end{pmatrix}$$

♢

Homogeneous Systems

Again we first start with homogeneous systems, i.e.,

$$\mathbf{y}' = \mathbf{A}\mathbf{y}\,. \tag{13.4}$$

Based on our experiences with single first-order linear equations and motivated by Example 13.1 we try the ansatz

$$\begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix} = \begin{pmatrix} v_1 e^{\lambda t} \\ v_2 e^{\lambda t} \end{pmatrix} = e^{\lambda t} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}.$$

We then find

$$\mathbf{A} e^{\lambda t} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \mathbf{A} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} y_1'(t) \\ y_2'(t) \end{pmatrix} = \lambda e^{\lambda t} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}.$$

Canceling the factor $e^{\lambda t}$ gives the equation

$$\mathbf{A} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \lambda \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}.$$

That is $\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$ is an eigenvector of $\mathbf{A}$ corresponding to eigenvalue λ.

The case in which $\mathbf{A}$ has two distinct real eigenvalues λ_1 and λ_2 is the simplest. Then the respective eigenvectors $\mathbf{v}_1$ and $\mathbf{v}_2$ are linearly independent and we get a general solution for the homogeneous system (13.4) as

Analogously to Theorem 12.4, the set of all solutions forms a vector space.

$$\mathbf{y}(t) = C_1 e^{\lambda_1 t} \mathbf{v}_1 + C_2 e^{\lambda_2 t} \mathbf{v}_2\,.$$

Example 13.3

Find the general solution of the homogeneous linear system

$$\begin{pmatrix} y_1' \\ y_2' \end{pmatrix} = \begin{pmatrix} 4 & 1 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}.$$

SOLUTION. Matrix $\mathbf{A} = \begin{pmatrix} 4 & 1 \\ -2 & 1 \end{pmatrix}$ has eigenvalues $\lambda_1 = 2$ and $\lambda_2 = 3$ with corresponding eigenvectors $\mathbf{v}_1 = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$ and $\mathbf{v}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$. Hence the general solution is given by

$$\begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix} = C_1 e^{2t} \begin{pmatrix} 1 \\ -2 \end{pmatrix} + C_2 e^{3t} \begin{pmatrix} 1 \\ -1 \end{pmatrix}.$$

♢

Inhomogeneous Systems

Consider the inhomogeneous system (13.3),

$$\mathbf{y}' = \mathbf{A}\mathbf{y} + \mathbf{s}$$

Assume that there exists a point $\mathbf{y}^*$ such that $\mathbf{A}\mathbf{y}^* + \mathbf{s} = 0$. Then $\mathbf{y}^*$ is an equilibrium state of (13.3). Analogously to Theorem 12.5, $\mathbf{y} - \mathbf{y}^*$ is then a solution of the homogeneous system (13.4). Thus the general solution is given by

Such a $\mathbf{y}^*$ exists if $\mathbf{A}$ is invertible.

$$\mathbf{y}(t) = \mathbf{y}_h(t) + \mathbf{y}^*$$

where $\mathbf{y}_h$ is the general solution of the corresponding homogeneous equation. If $\mathbf{A}$ has two distinct real eigenvalues λ_1 and λ_2 with respective eigenvectors $\mathbf{v}_1$ and $\mathbf{v}_2$ we obtain the general solution by

$$\mathbf{y}(t) = C_1 e^{\lambda_1 t}\mathbf{v}_1 + C_2 e^{\lambda_2 t}\mathbf{v}_2 + \mathbf{y}^* .$$

Example 13.4

Find the general solution of the inhomogeneous linear system

$$\begin{pmatrix} y_1' \\ y_2' \end{pmatrix} = \begin{pmatrix} 4 & 1 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} + \begin{pmatrix} 2 \\ 8 \end{pmatrix} .$$

Also find the particular solution with initial values $y_1(0) = y_2(0) = 0$.

SOLUTION. The solution of the corresponding homogeneous solution is (see Example 13.3)

$$\mathbf{y}_h(t) = \begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix} = C_1 e^{2t} \begin{pmatrix} 1 \\ -2 \end{pmatrix} + C_2 e^{3t} \begin{pmatrix} 1 \\ -1 \end{pmatrix} .$$

The particular solution $\mathbf{y}^*$ is a solution of the linear equation

$$\mathbf{A}\mathbf{y}^* + \mathbf{s} = \begin{pmatrix} 4 & 1 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} y_1^* \\ y_2^* \end{pmatrix} + \begin{pmatrix} 2 \\ 8 \end{pmatrix} = 0$$

which yields

$$\mathbf{y}^* = \begin{pmatrix} 1 \\ -6 \end{pmatrix} .$$

Thus the general solution is given by

$$\mathbf{y}(t) = \mathbf{y}_h(t) + \mathbf{y}^* = C_1 e^{2t} \begin{pmatrix} 1 \\ -2 \end{pmatrix} + C_2 e^{3t} \begin{pmatrix} 1 \\ -1 \end{pmatrix} + \begin{pmatrix} 1 \\ -6 \end{pmatrix} .$$

This of course coincides with our solution in Example 13.1. (The remaining computation for the particular solution of the initial value problem is thus completely the same.) ♢

13.3 Equilibrium Points for Linear Systems

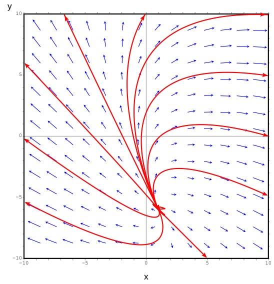
Definition 13.5

Equilibrium point. A solution $\mathbf{y}^*$ of $\mathbf{A}\mathbf{y} = -\mathbf{s}$ induces a constant solution $\mathbf{y}(t) = \mathbf{y}^*$ for the inhomogeneous system (13.3). Hence $\mathbf{y}^*$ is called an **equilibrium state** of system (13.3).

Example 13.6

Source. The general solution in Example 13.4,

$$\mathbf{y}(t) = C_1 e^{2t} \begin{pmatrix} 1 \\ -2 \end{pmatrix} + C_2 e^{3t} \begin{pmatrix} 1 \\ -1 \end{pmatrix} + \begin{pmatrix} 1 \\ -6 \end{pmatrix}$$

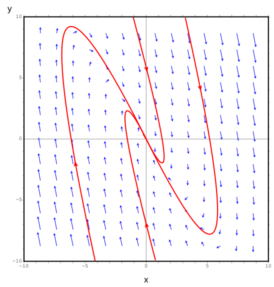


has the property that every solution path that starts close to the equilibrium point $\mathbf{y}^* = \begin{pmatrix} 1 \\ -6 \end{pmatrix}$ moves away from this point. Thus $\mathbf{y}^*$ is called an **unstable equilibrium point** (or **source**). This property is an immediate consequence of the fact that both eigenvalues λ_1 and λ_2 are positive real numbers.

Example 13.7

Sink. Find the general solution of the linear system

$$\begin{pmatrix} y_1' \\ y_2' \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -6 & -4 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} .$$



SOLUTION. Matrix $\mathbf{A} = \begin{pmatrix} 1 & 1 \\ -6 & -4 \end{pmatrix}$ has eigenvalues $\lambda_1 = -1$ and $\lambda_2 = -2$ with respective eigenvectors $\mathbf{v}_1 = \begin{pmatrix} 1 \\ -2 \end{pmatrix}$ and $\mathbf{v}_2 = \begin{pmatrix} 1 \\ -3 \end{pmatrix}$. Hence the general solution is given by

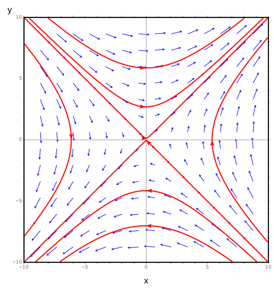
$$\begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix} = C_1 e^{-t} \begin{pmatrix} 1 \\ -2 \end{pmatrix} + C_2 e^{-2t} \begin{pmatrix} 1 \\ -3 \end{pmatrix} . \qquad \diamondsuit$$

In Example 13.7 every solution curve moves towards the point $\mathbf{y}^* = 0$. Therefore $\mathbf{y}^*$ is called a (globally) **asymptotically stable equilibrium state** (also called a **sink**). Notice that in this case both eigenvalues λ_1 and λ_2 are negative. The tangent of the solution curve in the limit point $\mathbf{y}^*$ is given by $\mathbf{v}_1$, i.e., the eigenvector corresponding to the larger of the two eigenvalues $\lambda_1 = -1 > -2 = \lambda_2$.

Example 13.8

Saddle point. Find the general solution of the linear system

$$\begin{pmatrix} y_1' \\ y_2' \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} .$$



SOLUTION. Matrix $\mathbf{A} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ has eigenvalues $\lambda_1 = 1$ and $\lambda_2 = -1$ with respective eigenvectors $\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and $\mathbf{v}_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$. Hence the general solution is given by

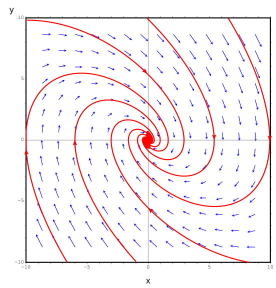
$$\begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix} = C_1 e^{t} \begin{pmatrix} 1 \\ 1 \end{pmatrix} + C_2 e^{-t} \begin{pmatrix} 1 \\ -1 \end{pmatrix} . \qquad \diamondsuit$$

The two eigenvalues in Example 13.8 are both real but have opposite signs. Thus a solution curve converges if and only if $C_1 = 0$, i.e., if and only if $\mathbf{y}(t) = C_2 e^{-t}\begin{pmatrix}1\\-1\end{pmatrix}$. All other solution curves diverge. Such an equilibrium point is called a **saddle point** of the system of differential equations. Notice that the image of the converging curve is a subset of the straight line spanned by the eigenvector $\mathbf{v}_2$ corresponding to the negative eigenvalue. It is called a **saddle path solution**.

Oscillating curve. Find the general solution of the linear system Example 13.9

$$\begin{pmatrix}y_1'\\y_2'\end{pmatrix} = \begin{pmatrix}0 & 2\\-2 & -2\end{pmatrix}\begin{pmatrix}y_1\\y_2\end{pmatrix}.$$

SOLUTION. Matrix $\mathbf{A} = \begin{pmatrix}0 & 2\\-2 & -2\end{pmatrix}$ has complex eigenvalues $\lambda_1 = \lambda = -1 + \sqrt{3}i$ and $\lambda_2 = \bar{\lambda} = -1-\sqrt{3}i$ with respective eigenvectors $\mathbf{v}_1 = \begin{pmatrix}2\\-1+\sqrt{3}i\end{pmatrix}$ and $\mathbf{v}_2 = \begin{pmatrix}2\\-1-\sqrt{3}i\end{pmatrix}$. Hence the general solution is given by



$$\begin{pmatrix}y_1(t)\\y_2(t)\end{pmatrix} = C_1 e^{(-1+\sqrt{3}i)t}\begin{pmatrix}2\\-1+\sqrt{3}i\end{pmatrix} + C_2 e^{(-1-\sqrt{3}i)t}\begin{pmatrix}2\\-1-\sqrt{3}i\end{pmatrix}.$$

A tedious straightforward computation gives the (real-valued) general solution

$$\mathbf{y}(t) = e^{-t}\begin{pmatrix}C_1\cos(\sqrt{3}t) + \frac{2(C_1+C_2)-C_1}{\sqrt{3}}\sin(\sqrt{3}t)\\ C_2\cos(\sqrt{3}t) - \frac{2C_1+C_2}{\sqrt{3}}\sin(\sqrt{3}t)\end{pmatrix}$$ ◊

Example 13.9 shows an example for a system where the equilibrium point $\mathbf{y}^* = 0$ is asymptotically stable. However, in this example we have two complex eigenvalues. Notice that convergence is caused by the negative real parts of both eigenvalues, $\operatorname{Re}(\lambda) = -1$.

Center. Find the general solution of the linear system Example 13.10

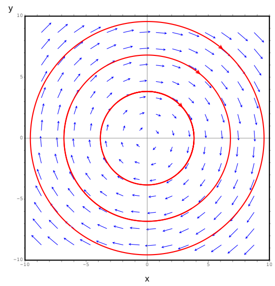
$$\begin{pmatrix}y_1'\\y_2'\end{pmatrix} = \begin{pmatrix}0 & 1\\-1 & 0\end{pmatrix}\begin{pmatrix}y_1\\y_2\end{pmatrix}.$$

SOLUTION. Matrix $\mathbf{A} = \begin{pmatrix}0 & 1\\-1 & 0\end{pmatrix}$ has purely imaginary eigenvalues $\lambda_1 = i$ and $\lambda_2 = -i$ with respective eigenvectors $\mathbf{v}_1 = \begin{pmatrix}1\\i\end{pmatrix}$ and $\mathbf{v}_2 = \begin{pmatrix}1\\-i\end{pmatrix}$. Hence the general solution is given by

$$\begin{pmatrix}y_1(t)\\y_2(t)\end{pmatrix} = \tilde{C}_1 e^{it}\begin{pmatrix}1\\i\end{pmatrix} + \tilde{C}_2 e^{-it}\begin{pmatrix}1\\-i\end{pmatrix}.$$

Using Euler's formula we find

$$\begin{pmatrix} y_1(t) \\ y_2(t) \end{pmatrix} = \begin{pmatrix} (\tilde{C}_1 + \tilde{C}_2)\cos t + i(\tilde{C}_1 - \tilde{C}_2)\sin t \\ i(\tilde{C}_1 - \tilde{C}_2)\cos t - (\tilde{C}_1 + \tilde{C}_2)\sin t \end{pmatrix}$$
$$= \begin{pmatrix} C_1 \cos t + C_2 \sin t \\ C_2 \cos t - C_1 \sin t \end{pmatrix}$$
$$= C_1 \begin{pmatrix} \cos t \\ -\sin t \end{pmatrix} + C_2 \begin{pmatrix} \sin t \\ \cos t \end{pmatrix}$$



where we set $C_1 = (\tilde{C}_1 + \tilde{C}_2)$ and $C_2 = i(\tilde{C}_1 - \tilde{C}_2)$. ♢

In Example 13.10 the real parts of both eigenvalues are 0. None of the solution curve converges towards the equilibrium point $\mathbf{y}^* = 0$ nor do they diverge. However, all curves are periodic with the same period length. All solution curves are ellipses or circles. The point $\mathbf{y}^*$ is called a **center**.

We summarize our observations in the following theorem.

Theorem 13.11

Stability of an equlibrium point. Let $\mathbf{y}^*$ be an equilibrium point of

$$\mathbf{y}' = \mathbf{A}\mathbf{y} + \mathbf{s} \tag{13.5}$$

where both eigenvalues of $\mathbf{A} \in \mathbb{R}^{2\times 2}$ are non-zero.

(a) $\mathbf{y}^*$ is an *asymptotically stable equilibrium state* (or *sink*) of (13.5) if and only if all eigenvalues of $\mathbf{A}$ have negative real parts.

(b) $\mathbf{y}^*$ is an *unstable equilibrium point* (or *source*) if and only if all eigenvalues of $\mathbf{A}$ have positive real parts.

(c) $\mathbf{y}^*$ is a *saddle point* of (13.5) if and only if both eigenvalues of $\mathbf{A}$ are non-zero real numbers of opposite signs.

(d) $\mathbf{y}^*$ is a *center* of (13.5) if and only if all eigenvalues of $\mathbf{A}$ are purely imaginary.

Fortunately it is possible to distinguish between these cases without computing the eigenvalues explicitly. For this purpose we need a result from linear algebra.

Definition 13.12

The **trace** of a matrix $\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$ is the sum of its diagonal entries, i.e., $\mathrm{tr}(\mathbf{A}) = a_{11} + a_{22}$.

Lemma 13.13

Let $\mathbf{A}$ be a 2×2 matrix with (real or complex) eigenvalues λ_1 and λ_2. Then

$$\mathrm{tr}(\mathbf{A}) = \lambda_1 + \lambda_2$$

and

$$\det(\mathbf{A}) = \lambda_1 \cdot \lambda_2 \,.$$

We can distinguish between these types of equilibrium points by means of the trace and the determinant of $\mathbf{A}$.

Theorem 13.14

Let $\mathbf{y}^*$ be an equilibrium point of

$$\mathbf{y}' = \mathbf{A}\mathbf{y} + \mathbf{s} \tag{13.6}$$

where $\mathbf{A} \in \mathbb{R}^{2\times 2}$ has non-zero determinant.

(a) $\mathbf{y}^*$ is an *asymptotically stable equilibrium point* (or *sink*) if and only if $\det(\mathbf{A}) > 0$ and $\operatorname{tr}(\mathbf{A}) < 0$.

(b) $\mathbf{y}^*$ is an *unstable equilibrium point* (or *source*) if and only if $\det(\mathbf{A}) > 0$ and $\operatorname{tr}(\mathbf{A}) > 0$.

(c) $\mathbf{y}^*$ is a *saddle point* of (13.6) if and only if $\det(\mathbf{A}) < 0$.

(d) $\mathbf{y}^*$ is a *center* of (13.6) if and only if $\operatorname{tr}(\mathbf{A}) = 0$ and $\det(\mathbf{A}) > 0$.

13.4 Stability for Non-Linear Systems

A system of differential equations is called **autonomous** if $\mathbf{F}$ does not depend directly on the independent variable t, i.e., if the equation becomes

$$\mathbf{y}' = \mathbf{F}(\mathbf{y})\,. \tag{13.7}$$

Again a point $\mathbf{y}^*$ is called an **equilibrium point** of (13.7) if $\mathbf{F}(\mathbf{y}^*) = 0$. It is called **locally asymptotically stable** if any path starting near $\mathbf{y}^*$ converges to $\mathbf{y}^*$ as $t \to \infty$.

To examine whether an equilibrium point $\mathbf{y}^*$ is locally asymptotically stable it is sufficient to replace $\mathbf{F}$ by a linear approximation, i.e., by its Jacobian,

$$\mathbf{F}(\mathbf{y}) \approx \mathbf{F}'(\mathbf{y}^*)\cdot(\mathbf{y}-\mathbf{y}^*)\,.$$

Thus we find the following sufficient characterization.

Theorem 13.15

Lyapunov's Theorem. Let $\mathbf{F}\colon \mathbb{R}^2 \to \mathbb{R}^2$ be a $\mathscr{C}^1$ function and let $\mathbf{y}^*$ be an equilibrium point of system (13.7). If both eigenvalue of the Jacobian $\mathbf{F}'(\mathbf{y}^*)$ have negative real parts (i.e., $\operatorname{tr}(\mathbf{F}'(\mathbf{y}^*)) < 0$ and $\det(\mathbf{F}'(\mathbf{y}^*)) > 0$), then $\mathbf{y}^*$ is locally asymptotically stable.

Example 13.16

Consider the following system of differential equations.

$$\mathbf{y}' = \mathbf{F}(\mathbf{y}) = \begin{pmatrix} -4y_1^2 - 3y_1 - 2y_1y_2 + y_2^4 \\ 2y_1^2y_2^2 - y_1 - y_2 + y_2^2 \end{pmatrix}$$

Show that $\mathbf{y}^* = (0,0)$ is an asymptotically stable equilibrium point.

SOLUTION. Obviously $\mathbf{F}(0,0) = 0$ and thus $\mathbf{y}^*$ is an equilibrium point. Its Jacobian matrix is given by

$$\mathbf{F}'(0,0) = \begin{pmatrix} -3 & 0 \\ -1 & -1 \end{pmatrix}, \qquad \operatorname{tr}\big(\mathbf{F}'(0,0)\big) = -4 < 0\,, \quad \det\big(\mathbf{F}'(0,0)\big) = 3 > 0\,,$$

and thus $\mathbf{y}^* = (0,0)$ is locally asymptotically stable. ♢

Theorem 13.17

Olech's Theorem. Let $\mathbf{F}\colon \mathbb{R}^2 \to \mathbb{R}^2$ be a $\mathscr{C}^1$ function and let $\mathbf{y}^*$ be an equilibrium point of system (13.7). Assume that for all $\mathbf{y} \in \mathbb{R}^2$ both eigenvalues of the Jacobian $\mathbf{F}'(\mathbf{y})$ have negative real parts (i.e., $\operatorname{tr}(\mathbf{F}'(\mathbf{y})) < 0$ and $\det(\mathbf{F}'(\mathbf{y})) > 0$), and

$$F'_{11}(\mathbf{y})F'_{22}(\mathbf{y}) \neq 0 \text{ for all } \mathbf{y} \in \mathbb{R}^2 \quad \text{or} \quad F'_{12}(\mathbf{y})F'_{21}(\mathbf{y}) \neq 0 \text{ for all } \mathbf{y} \in \mathbb{R}^2,$$

then $\mathbf{y}^*$ is globally asymptotically stable.

Example 13.18

Consider the following system of differential equations.

$$\mathbf{y}' = \mathbf{F}(\mathbf{y}) = \begin{pmatrix} -y_1^3 - y_1 - y_2^3 - y_2 \\ 2y_1 - 3y_2 \end{pmatrix}$$

Show that $\mathbf{y}^* = (0,0)$ is an globally asymptotically stable equilibrium point.

SOLUTION. Obviously $\mathbf{F}(0,0) = 0$ and thus $\mathbf{y}^*$ is an equilibrium point. Its Jacobian matrix is given by

$$\mathbf{F}'(y_1,y_2) = \begin{pmatrix} -3y_1^2 - 1 & -3y_2^2 - 1 \\ 2 & -3 \end{pmatrix}.$$

Thus

$$\begin{aligned} &\operatorname{tr}(\mathbf{F}'(y_1,y_2)) = -3y_1^2 - 4 < 0\,, \\ &\det(\mathbf{F}'(y_1,y_2)) = 9y_1^2 + 6y_2^2 + 5 > 0\,, \\ &F'_{11}(\mathbf{y})F'_{22}(\mathbf{y}) = 9y_1^2 + 3 \neq 0 \end{aligned}$$

for all $\mathbf{y} \in \mathbb{R}^2$ and therefore $\mathbf{y}^* = (0,0)$ is globally asymptotically stable. ♢

Theorem 13.19

Local saddle point. Let $\mathbf{F}\colon \mathbb{R}^2 \to \mathbb{R}^2$ be a $\mathscr{C}^1$ function and let $\mathbf{y}^*$ be an equilibrium point of system (13.7). If both eigenvalue of the Jacobian $\mathbf{F}'(\mathbf{y}^*)$ are non-zero real numbers of opposite signs (i.e., $\det(\mathbf{F}'(\mathbf{y}^*)) < 0$), then $\mathbf{y}^*$ is a **local saddle point**. Moreover, for any given starting point t_0 there exist exactly two solution paths $\mathbf{y}_1(t)$ and $\mathbf{y}_2(t)$ defined on $[t_0,\infty)$ that converge towards $\mathbf{y}^*$ from opposite directions in the phase plane. As $t \to \infty$, both paths become "tangent in the limit" to the line through $\mathbf{y}^*$ with the same direction as the eigenvector corresponding to the negative eigenvalue of $\mathbf{F}'(\mathbf{y}^*)$. These curves are called **saddle path solutions** .

saddle path solution

Example 13.20

Lotka-Volterra Equation. Consider the celebrated Lotka-Volterra predator-prey model.

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \mathbf{F}(x,y) = \begin{pmatrix} x(\alpha-\beta y) \\ -y(\gamma-\delta x) \end{pmatrix} \qquad \text{where } \alpha,\beta,\gamma,\delta>0.$$

Here x is the population of prey (say rabbits) and y is the population of predator (say foxes). Show that $\mathbf{y}^* = (0,0)$ is a local saddle point. Compute the tangent line at $\mathbf{y}^*$ for the saddle path solutions.

SOLUTION. Obviously $\mathbf{F}(0,0) = 0$ and thus $\mathbf{y}^*$ is an equilibrium point. Its Jacobian matrix is given by

$$\mathbf{F}'(x,y) = \begin{pmatrix} \alpha-\beta y & -\beta x \\ \delta y & -\gamma+\delta x \end{pmatrix}, \qquad \mathbf{F}'(0,0) = \begin{pmatrix} \alpha & 0 \\ 0 & -\gamma \end{pmatrix}.$$

As can be easily seen $\mathbf{F}'(0,0)$ has $\lambda_1 = \alpha > 0$ and $\lambda_2 = -\gamma$ and thus $\mathbf{y}^* = (0,0)$ is a local saddle point. Eigenvector $\mathbf{v}_2 = (0,1)$ corresponds to the negative eigenvalue λ_2. Thus $x = 0$ is the tangent line to the saddle path solutions. ♢

Alternatively: $\det(\mathbf{F}'(0,0)) = -\alpha\gamma < 0$.

The other stationary points that we have characterized in Theorem 13.11 for linear differential equations, sources and centers, can occur as well.

If both eigenvalues have positive real parts, then solutions that start close to the equilibrium point $\mathbf{y}^*$ move away from it and the point is a source. However, such solutions may still be bounded.

If the eigenvalues are purely imaginary or 0, no definite statement about the limiting behavior of the solution can be made.

Example 13.21

The point $(0,0)$ is not an interesting point for the Lotka-Volterra model as then there are neither rabbits nor foxes. However, there is a second stationary point of the Lotka-Volterra equation, $\mathbf{y}^* = (\gamma/\delta, \alpha/\beta)$. Its Jacobian matrix is given by

$$\mathbf{F}'(\gamma/\delta, \alpha/\beta) = \begin{pmatrix} 0 & -\frac{\beta\gamma}{\delta} \\ \frac{\alpha\delta}{\beta} & 0 \end{pmatrix},$$

with eigenvalues $\lambda_1 = i\sqrt{\alpha\gamma}$ and $\lambda_2 = -i\sqrt{\alpha\gamma}$. As these eigenvalues are both purely imaginary, no conclusions can be drawn from a linear approximation of $\mathbf{F}$. ♢

Definition 13.22

Stable point. An equilibrium point $\mathbf{y}^*$ of the autonomous system (13.7) is called **stable** if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that $\|\mathbf{y}_0 - \mathbf{y}^*\| < \delta$ implies that every solution $\mathbf{y}(t)$ that satisfies $\mathbf{y}(0) = \mathbf{y}_0$ is defined for all $t > 0$ and satisfies

$$\left\|\mathbf{y}(t) - \mathbf{y}^*\right\| < \varepsilon \quad \text{for all } t > 0.$$

That means that solutions with initial point close to $\mathbf{y}^*$ remain near this point. In opposition to a *locally asymptotically stable* point, solutions need not converge to $\mathbf{y}^*$.

Notice that the *centers* in Theorem 13.11 are *stable* equilibrium points.

Definition 13.23

Lyapunov function. Let $V\colon \mathbb{R}^2 \to \mathbb{R}$ be a positive definite $\mathscr{C}^1$ function in an open neighborhood D of $\mathbf{y}^*$, that is, $V(\mathbf{y}^*) = 0$ and $V(\mathbf{x}) > 0$ for all $\mathbf{x} \in D \setminus \{\mathbf{y}^*\}$. Let $\mathbf{y}(t)$ be a solution of the autonomous system (13.7). For the derivative of $V\big(\mathbf{y}(t)\big)$, i.e., of V along the solution curve, we find by the chain rule

$$\frac{d}{dt}V\big(\mathbf{y}(t)\big) = \nabla V\big(\mathbf{y}(t)\big)\cdot\mathbf{y}'(t) = \nabla V\big(\mathbf{y}(t)\big)\cdot\mathbf{F}\big(\mathbf{y}(t)\big)\,. \tag{13.8}$$

If (13.8) is non-positive for all $\mathbf{y} \in D$, then $V(\mathbf{y})$ is called a **Lyapunov function** for the autonomous system. If (13.8) is negative for all $\mathbf{y} \in D \setminus \{\mathbf{y}^*\}$, then V is called a **strong Lyapunov function** for the autonomous system.

Theorem 13.24

Lyapunov's Theorem. Let $\mathbf{y}^*$ be an equilibrium point for the autonomous system (13.7). If there exists a Lyapunov function in an open neighborhood D of $\mathbf{y}^*$. Then $\mathbf{y}^*$ is a *stable* equilibrium point. If there exists a strong Lyapunov function for the system, then $\mathbf{y}^*$ is a *locally asymptotically stable* equilibrium point.

Example 13.25

Prove that $\mathbf{y}^* = (x^*, y^*) = (\gamma/\delta, \alpha/\beta)$ is a stable equilibrium point of the Lotka-Volterra equation.

SOLUTION. Let

$$H(x,y) = \delta(x - x^* \ln x) + \beta(y - y^* \ln y)\,.$$

We claim that $V(x,y) = H(x,y) - H(x^*,y^*)$ is a Lyapunov function. In fact we find for its gradient

$$V'(x,y) = \Big(\delta(1-\tfrac{x^*}{x}),\quad \beta(1-\tfrac{y^*}{y})\Big)\,, \qquad V'(x^*,y^*) = (0,0)$$

that is, (x^*,y^*) is a stationary point of V. All eigenvalues of its Hessian

$$V''(x,y) = \begin{pmatrix} \delta\frac{x^*}{x^2} & 0 \\ 0 & \beta\frac{y^*}{y^2} \end{pmatrix}$$

are positive and thus V is strict convex and (x^*,y^*) is a strict global minimum of V. Consequently, as

$$V(x^*,y^*) = H(x^*,y^*) - H(x^*,y^*) = 0$$

V is positive definite. Moreover, as $x^* = \gamma/\delta$ and $y^* = \alpha/\beta$ we find

$$\begin{aligned}\frac{d}{dt}V(x(t),y(t)) &= \nabla V(x,y)\cdot\begin{pmatrix} x'(t) \\ y'(t)\end{pmatrix} \\ &= \delta(1 - x^*/x)x(\alpha - \beta y) + \beta(1 - y^*/y)(-y(\gamma - \delta x)) = 0\,.\end{aligned}$$

Thus $\mathbf{y}^* = (x^*,y^*)$ is a stable equilibrium point. In fact $\frac{d}{dt}V(x(t),y(t)) = 0$ even implies that $V(x,y)$ is constant along every solution curve. One can prove that this also implies that these solutions are closed curves. ◊

13.5 Phase Plane Analysis

For an autonomous system of differential equations

$$\mathbf{y}' = \mathbf{F}(\mathbf{y}) \tag{13.9}$$

the derivative of the solution curve $\mathbf{y}(t)$ only depends on the particular point $(y_1(t), y_2(t))$ but not on t. Thus for each point (y_1, y_2) we can draw the corresponding derivative $\mathbf{y}'$ into the y_1y_2-plane. Again we obtain a **direction field** (also called **vector field**) which we can use to analyze the properties of solution curves by visual inspection even when we do not have an explicit solution of the system.

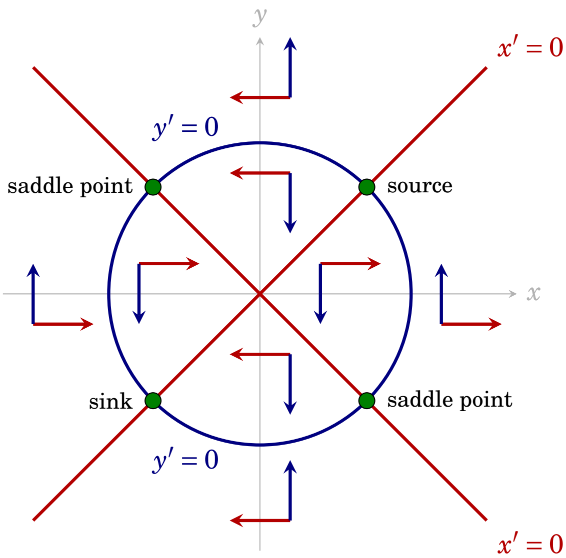
For this purpose we can draw the curves where $y_1' = 0$ and $y_2' = 0$. These curves are called the **nullclines** of the system. Their intersections are just the equilibrium points of the system. Moreover, these nullclines partition the plane into regions where the directions of increase or decrease of each variable remain constant. The resulting diagram is called the **phase diagram** of the system. It provides some useful information about possible solution paths of our system.

Example 13.26

Consider the autonomous system of differential equations:

$$x' = x^2 - y^2$$
$$y' = x^2 + y^2 - 1$$

We find the following phase diagram. Notice that it has four equilibrium points. On each point of a nullcline the derivative of the corresponding variable is 0.



— Exercises

13.1 Find the general solutions of the following systems and draw (sketch) the respective solution curves:

(a) $x' = -3x - 2y$, $y' = -2x - 6y$ (b) $x' = 2x + 3y$, $y' = 4x + 13y$ (c) $x' = -x + 5y$, $y' = 5x - y$

Use both methods:

(1) reduction to a second order differential equation, and

(2) eigenvalues and eigenvectors.

13.2 Find the general solutions of the following systems and draw (sketch) the respective solution curves. Find all equilibrium points and check their stability.

(a) $x' = -3x - 2y + 1$, $y' = -2x - 6y - 4$ (b) $x' = 2x + 3y + 1$, $y' = 4x + 13y - 5$ (c) $x' = -x + 5y + 1$, $y' = 5x - y - 5$

13.3 For which values of the constant a are the following systems globally asymptotically stable?

(a) $x' = ax - y$, $y' = x + ay$ (b) $x' = ax - (2a - 4)y$, $y' = x + 2ay$

13.4 Determine (if possible) the local asymptotic stability of the following systems at the given stationary points. Are these also globally asymptotically stable?

(a) $x' = -x + \frac{1}{2}y^2$, $y' = 2x - 2y$ at $(0,0)$

(b) $x' = -x^3 - y$, $y' = x - y^3$ at $(0,0)$

(c) $x' = x - 3y + 2x^2 + y^2 - xy$, $y' = 2x - y - e^{x-y}$ at $(1,1)$

13.5 Show that $(0,0)$ is a globally asymptotically stable equilibrium point for the system

$$\begin{aligned} x' &= y \\ y' &= -ky - w^2 x \end{aligned} \qquad (k > 0,\ w \neq 0)$$

13.6 Prove that $(0,0)$ is a locally asymptotically stable point of

$$\begin{aligned} x' &= -x^3 - y \\ y' &= x - y^3 \end{aligned}$$

HINT: Show that $V(x, y) = x^2 + y^2$ is a strong Lyapunov function.

13.7 Consider the differential equation for $p > 0$.

$$p'(t) = a\left(\frac{b}{p(t)} - c\right) \qquad \text{where } a, b, c \text{ are positive constants.}$$

Find the equilibrium point and prove that it is locally asymptotically stable using Lyapunov function $V(p) = (p - b/c)^2$.

13.8 Draw the phase diagram for the Lotka-Volterra equation (see Example 13.20).

13.9 Consider the following generalization of the Lotka-Volterra system:

$$\begin{aligned} x' &= x(\alpha - \varepsilon x - \beta y) \\ y' &= y(-\gamma + \delta x - \eta y) \end{aligned} \qquad (\alpha, \beta, \gamma, \delta, \varepsilon, \eta > 0)$$

with $\delta\alpha > \gamma\varepsilon$. Verify that

$$(x^*, y^*) = \left(\frac{\beta\gamma + \alpha\eta}{\beta\delta + \eta\varepsilon}, \frac{\delta\alpha - \gamma\varepsilon}{\beta\delta + \eta\varepsilon}\right)$$

is an equilibrium point. Is it locally asymptotically stable?

13.10 Deduce Theorem 13.14 from Theorem 13.11.

HINT: Use Lemma 13.13.

14

Control Theory

Drive your vehicle as fast as possible but do not leave the road.

14.1 The Control Problem

Example 14.1

Economic growth. We want to maximize the total consumption in a country over time interval $[0,T]$. That is, we have to solve the problem

$$\max_{0\leq s(t)\leq 1} \int_0^T \big(1-s(t)\big) f\big(k(t)\big)\, dt$$

where $f(k)$ denotes the production function, $k(t)$ is the real capital stock of the country at time t, and $s(t)$ is the rate of investment at time t. The integrand $\big(1-s(t)\big) f\big(k(t)\big)$ in our problem is the flow of consumption per unit of time. It is called the *objective function* of this optimization problem.

The capital stock $k(t)$ has to satisfy the initial value problem (differential equation)

$$k'(t) = s(t) f\big(k(t)\big), \quad k(0) = k_0 .$$

Moreover, we may wish to leave some capital stock k_T at time T, i.e.,

$$k(T) \geq k_T .$$

The only quantity that can be chosen freely at any time t is the rate of investment, $s(t)$. It is called the *control function* of our problem. It is quite natural to assume that $s \in [0,1]$. This is called the *control region*.

In summary we have to solve the following optimal control problem:

$$\max_{0\leq s(t)\leq 1} \int_0^T \big(1-s(t)\big) f\big(k(t)\big)\, dt, \qquad s \in [0,1],$$

$$k'(t) = s(t) f\big(k(t)\big), \quad k(0) = k_0, \quad k(T) \geq k_T .$$

Example 14.2

Oil extraction. Suppose that $y(t)$ denotes the amount of oil in some reservoir at time t. The rate of extraction is then given by its first derivative, i.e., $u(t) = -y'(t)$. So if $p(t)$ denotes the market price of oil and $C(t,y,u)$ extraction cost per unit of time, then the instantaneous rate of profit is given by

$$\pi(t,y(t),u(t)) = p(t)u(t) - C(t,y(t),u(t)) .$$

Hence if we denote the (constant) discount rate by r, then total discounted rate of profit over the time interval $[0,T]$ is given by

$$\int_0^T \left[p(t)u(t) - C(t,y(t),u(t))\right]e^{-rt}\,dt .$$

Again we want to maximize our revenue. It is natural to assume that $y(t) \geq 0$ and $u(t) \geq 0$ for all t. Therefore we have to solve the following optimal control problem:

$$\max_{u(t)\geq 0} \int_0^T \left[p(t)u(t) - C(t,y(t),u(t))\right]e^{-rt}\,dt, \qquad u \in [0,\infty),$$

$$y'(t) = -u(t), \quad y(0) = y_0, \quad y(T) \geq 0 .$$

Here the rate of extraction u can be chosen freely (within the given control region) and is thus the control variable in this problem. It is now our task to find an optimal extraction process that optimizes our profit. However, we may distinguish between two case scenarios.

Case 1: The time horizon T is fixed. We plan to stop production at time T.

Case 2: There is no fixed date T when we stop extraction. So we have to find the optimal time stopping time T and together with an optimal extraction process u.

14.2 The Standard Problem (T fixed)

We restrict our interest to the **standard end constraint problem**:

(C1) Find maximum

$$\max_u \int_0^T f(t,y,u)\,dt, \qquad u \in \mathcal{U} \subseteq \mathbb{R}.$$

Function f is called the **objective function**, u is the **control function**, and $\mathcal{U}$ is the **control region**.

(C2) The **state variable** y has to satisfy the following *controlled differential equation* (initial value problem)

$$y' = g(t,y,u), \qquad y(0) = y_0,$$

(C3) and one of the following *terminal conditions*:

(a) $y(T) = y_1$,

(b) $y(T) \geq y_1$ [or: $y(T) \leq y_1$],

(c) $y(T)$ free.

A pair (y, u) that satisfies (C2) and (C3) is called an **admissible pair**. Our task is to find an **optimal pair** among all admissible pairs that maximizes (C1).

Similarly to constraint static optimization we introduce an auxiliary function to gain necessary (and some sufficient) conditions for an optimal pair.

Definition 14.3

Hamiltonian. The function

$$\mathcal{H}(t, y, u, \lambda) = \lambda_0 f(t, y, u) + \lambda(t) g(t, y, u)$$

is called the **Hamiltonian** of our standard problem with fixed time horizon T. The new argument λ is called the **adjoint function** or **co-state variable** associated with the differential equation. λ_0 is either 0 or 1.

Constant $\lambda_0 \in \{0, 1\}$ is in almost all problems equal to 1. Hence we assume for the remaining part of this chapter that $\boldsymbol{\lambda_0 = 1}$, i.e.,

$$\mathcal{H}(t, y, u, \lambda) = f(t, y, u) + \lambda(t) g(t, y, u),$$

in order to keep the presentation simple.

Theorem 14.4

The Maximum principle. Assume that (y^*, u^*) is an optimal pair for the standard problem. Then there exists a continuous function $\lambda(t)$ such that for all $t \in [0, T]$ we have

(i) u^* maximizes $\mathcal{H}$ w.r.t. u, i.e.,

$$\mathcal{H}(t, y^*, u^*, \lambda) \geq \mathcal{H}(t, y^*, u, \lambda) \quad \text{for all } u \in \mathcal{U}.$$

(ii) λ satisfies the differential equation

$$\lambda' = -\frac{\partial}{\partial y} \mathcal{H}(t, y^*, u^*, \lambda).$$

(iii) Corresponding to each of the terminal conditions in (C3) there is a **transversality condition** on $\lambda(T)$:

(a) $y(T) = y_1$: $\lambda(T)$ free,

(b) $y(T) \geq y_1$: $\lambda(T) \geq 0$ [where $\lambda(T) = 0$ if $y^*(T) > y_1$],

(c) $y(T)$ free: $\lambda(T) = 0$.

Observe that the maximum principle gives a *necessary* condition for an optimal pair of our standard problem. That is, each *admissible* pair for which we can find such a function λ is a good candidate for an optimal pair. The next theorem provides a *sufficient condition* for an optimal pair.

Mangasarian's theorem. Assume that (y^*, u^*) is an admissible pair for the standard problem with corresponding adjoint function $\lambda(t)$ that satisfies conditions (i)–(iii) from Theorem 14.4. If $\mathscr{U}$ is convex and $\mathscr{H}(t, y, u, \lambda)$ is concave in (y, u) for all $t \in [0, T]$, then (y^*, u^*) is an optimal pair. Theorem 14.5

We can derive a recipe for finding optimal pairs for "nice" control problems from Theorems 14.4 and 14.5:

1. For each triple (t, y, λ) find a (global) maximum $\hat{u}(t, y, \lambda)$ of $\mathscr{H}(t, y, u, \lambda)$ w.r.t. u.
2. Solve the differential equations
$$y' = g\big(t, y, \hat{u}(t, y, \lambda)\big) \text{ and}$$
$$\lambda' = -\mathscr{H}_y\big(t, y, \hat{u}(t, y, \lambda), \lambda\big).$$
3. Find a particular solution $y^*(t)$ which satisfies $y^*(0) = y_0$ and the given terminal condition.
4. Find a particular solution $\lambda^*(t)$ that satisfies the corresponding transversality condition.
5. Thus we obtain candidates (y^*, u^*) for optimal pairs using $y^*(t)$ and $u^*(t) = \hat{u}(t, y^*, \lambda^*)$.
6. If $\mathscr{U}$ is convex and $\mathscr{H}(t, y, u, \lambda^*)$ is concave in (y, u), then (y^*, u^*) is an optimal pair.

It is important to note that these steps need not be computed in the given ordering. Indeed, often a different ordering is more appropriate.

Find an optimal control u^* for Example 14.6

$$\max \int_0^1 y(t)\,dt, \quad u \in [0, 1], \quad y' = y + u, \quad y(0) = 0, \quad y(1) \text{ free.}$$

Heuristically we can argue that the objective function $y(t)$ should be as large as possible and hence we should set $u^*(t) = 1$ for all t.

SOLUTION. The Hamiltonian for our problem is given by

$$\mathscr{H}(t, y, u, \lambda) = f(t, y, u) + \lambda g(t, y, u) = y + \lambda(y + u).$$

Maximum $\hat{u}$ of $\mathscr{H}$ w.r.t. u is then

$$\hat{u} = \begin{cases} 1 & \text{if } \lambda \geq 0, \\ 0 & \text{if } \lambda < 0 \end{cases}$$

The solutions of the (inhomogeneous linear) differential equation

$$\lambda' = -\mathscr{H}_y = -(1 + \lambda), \quad \lambda(1) = 0$$

is given by

$$\lambda^*(t) = e^{1-t} - 1\,.$$

As $\lambda^*(t) = e^{1-t} - 1 \geq 0$ for all $t \geq 0$ we have $\hat{u}(t) = 1$.
The solution of the (inhomogeneous linear) differential equation

$$y' = y + \hat{u} = y + 1\,, \quad y(0) = 0$$

is given by

$$y^*(t) = e^t - 1\,.$$

Thus we obtain $u^*(t) = \hat{u}(t) = 1$ in accordance with our heuristic approach.

At last observe that the Hamiltonian $\mathscr{H}(t, y, u, \lambda) = y + \lambda(y + u)$ is linear and thus concave in (y, u). Consequently, $u^*(t) = 1$ is the sought optimal control. ♢

Example 14.7

Find an optimal control u^* for

$$\min \int_0^T \left[y^2(t) + cu^2(t)\right] dt, \quad u \in \mathbb{R}, \quad y' = u, \quad y(0) = y_0, \quad y(T) \text{ free.}$$

where $c > 0$ is some positive constant.

SOLUTION. We do not have a tools for solving this problem directly. So we solve the *maximization* problem instead:

$$\max \int_0^T -\left[y^2(t) + cu^2(t)\right] dt\,.$$

The Hamiltonian for this is given by

$$\mathscr{H}(t, y, u, \lambda) = f(t, y, u) + \lambda g(t, y, u) = -y^2 - cu^2 + \lambda u\,.$$

Maximum $\hat{u}$ of $\mathscr{H}$ w.r.t. u is obtained from

$$0 = \mathscr{H}_u = -2c\hat{u} + \lambda \quad \text{and thus} \quad \hat{u} = \frac{\lambda}{2c}\,.$$

Next we have to find solutions of the differential equations

$$y' = \hat{u} = \frac{\lambda}{2c}\,,$$
$$\lambda' = -\mathscr{H}_y = 2y\,.$$

Differentiating the second equation yields the second order homogeneous linear differential equation

$$\lambda'' = 2y' = \frac{\lambda}{c} \quad \text{and thus} \quad \lambda'' - \frac{1}{c}\lambda = 0$$

with general solution

$\pm\frac{1}{\sqrt{c}}$ are the roots of the characteristic equation.

$$\lambda^*(t) = C_1 e^{rt} + C_2 e^{-rt} \quad \text{where} \quad r = \frac{1}{\sqrt{c}}\,.$$

Initial value and transversality condition result in the constraints

$$\lambda^{*\prime}(0) = 2y(0) = 2y_0 \quad \text{and} \quad \lambda^*(T) = 0$$

and hence

$$\begin{aligned} r(C_1 - C_2) &= 2y_0, \\ C_1 e^{rT} + C_2 e^{-rT} &= 0, \end{aligned}$$

with solution

$$C_1 = \frac{2y_0 e^{-rT}}{r(e^{rT} + e^{-rT})} \quad \text{and} \quad C_2 = -\frac{2y_0 e^{rT}}{r(e^{rT} + e^{-rT})}\,.$$

Therefore we obtain the optimal control by

$$\begin{aligned} \lambda^*(t) &= \frac{2y_0}{r(e^{rT} + e^{-rT})}\left(e^{-r(T-t)} - e^{r(T-t)}\right) \\ y^*(t) &= \frac{1}{2}\lambda^*(t) = y_0 \frac{e^{-r(T-t)} - e^{r(T-t)}}{r(e^{rT} + e^{-rT})} \\ u^*(t) &= \hat{u}(t, y^*, \lambda^*) = \frac{1}{2c}\lambda^*(t) = \frac{y_0}{c}\,\frac{e^{-r(T-t)} - e^{r(T-t)}}{r(e^{rT} + e^{-rT})} \end{aligned}$$

It is easy to see that the Hamiltonian $\mathscr{H}(t, y, u, \lambda) = -y^2 - cu^2 + \lambda u$ is concave in (y, u) and thus $u^*(t) = \frac{y_0}{c}\frac{e^{-r(T-t)} - e^{r(T-t)}}{r(e^{rT}+e^{-rT})}$ is the optimal control. ♢

Optimal consumption. Consider a consumer who expects to live from present time, $t = 0$, until time T. He or she wants to maximize his or her "lifetime intertemporal utility function" Example 14.8

$$\int_0^T e^{-\alpha t} u\big(c(t)\big)\, dt$$

where $u(c)$ is his or her utility function with "intertemporal discount factor" $\alpha > 0$ and consumption expenditure $c(t)$ at time t. We may assume that

$$u'(c) > 0 \quad \text{and} \quad u''(c) < 0 \quad \text{for all } c \geq 0.$$

The wealth $w(t)$ of the consumer follows the differential equation

$$w'(t) = r(t)w(t) + y(t) - c(t), \quad w(0) = w_0, \quad \text{and} \quad w(T) \geq 0$$

where $r(t)$ is the instantaneous rate of interest at time t and $y(t)$ is the predicted income. This is an optimal control problem with state variable w and control variable c. The control region obviously is $c \geq 0$.

SOLUTION. The Hamiltonian for this optimal control problem is given by

$$\mathcal{H}(t,w,c,\lambda) = e^{-\alpha t}u(c) + \lambda(rw + y - c)\,.$$

If c^* is the optimal consumption, then

$$\mathcal{H}_c = e^{-\alpha t}u'(c^*) - \lambda = 0 \quad \text{and thus} \quad \lambda(t) = e^{-\alpha t}u'(c^*)\,.$$

By the maximum principle we get the differential equation

$$\lambda'(t) = -\mathcal{H}_w = -\lambda(t)r$$

with solution

$$\lambda(t) = \lambda(0)\exp\left(-\int_0^t r(s)\,ds\right)\,.$$

Unfortunately, more explicit formula is not possible. So we only look at a special case. Assume that $r(t) = r$ is independent of time and that $r = \alpha$. Then the last formulæ reduce to

$$\lambda(t) = e^{-rt}u'(c^*) \quad \text{and} \quad \lambda(t) = \lambda(0)e^{-rt},$$

respectively and hence

$$u'(c^*) = \lambda(0) = \text{constant}\,.$$

Since by our assumptions $u''(c) < 0$, we conclude for the optimal control

$$c^*(t) = \bar{c} = \text{constant}\,.$$

We arrive at the differential equation

$$w'(t) = rw(t) + y(t) - \bar{c}, \quad w(0) = w_0$$

whose solution is

$$w^*(t) = e^{rt}\left[w_0 + \int_0^t e^{-rs}y(s)\,ds - \frac{\bar{c}}{r}\left(1 - e^{-rt}\right)\right]\,.$$

The transversality condition (iiib) in Theorem 14.4 for $w(T) \geq 0$ implies that

$$\lambda(T) \geq 0 \quad \text{with} \quad \lambda(T) = 0 \text{ if } w^*(T) > 0\,.$$

However, $\lambda(T) = 0$ implies $\lambda(T) = e^{-rT}u'(c^*) = 0$ and thus $u'(c^*) = 0$, a contradiction to our assumption that $u'(c) > 0$. Thus $w^*(T) = 0$, that is, it is optimal for the consumer to leave no legacy. We then find

$$0 = w^*(T) = e^{rT}\left[w_0 + \int_0^T e^{-rs}y(s)\,ds - \frac{\bar{c}}{r}\left(1 - e^{-rT}\right)\right]$$

and consequently we find for the optimal level of consumption

$$\bar{c} = \frac{r}{1 - e^{-rT}}\left[w_0 + \int_0^T e^{-rs}y(s)\,ds\right]\,.$$

Observe that the term in the square brackets is the initial wealth of the consumer plus the total discounted income.

At last we note that $\mathcal{H}$ is concave in (w,c) as $u''(c) < 0$ and thus $\bar{c}$ indeed is an optimal control. ♢

14.3 The Standard Problem (T variable)

Suppose that the time horizon $[0, T]$ is not fixed in advance. So in addition to the optimal control function we also have to find an optimal value T^*.

The **variable final time problem** can be shortly formulated as

$$\max_{u,T} \int_0^T f(t,y,u)\,dt, \qquad u \in \mathscr{U}, \quad y' = g(t,y,u), \quad y(0) = y_0,$$

$$\text{(a) } y(T) = y_1, \text{ or} \quad \text{(b) } y(T) \geq y_1, \text{ or} \quad \text{(c) } y(T) \text{ free.}$$

Notice that T can now be chosen freely and thus the maximum is taken over all feasible u and T.

Theorem 14.9

The Maximum principle with variable time. Assume that (y^*, u^*) is an admissible pair defined on $[0, T^*]$ that solves the variable final time problem from above with T free. Then all the conditions (i)–(iii) in the maximum principle (Theorem 14.4) are satisfied on $[0, T^*]$, and in addition,

(iv) $\mathscr{H}(T^*, y^*(T^*), u^*(T^*), \lambda(T^*)) = 0\,.$

Example 14.10

Oil extraction. We consider the special case of the optimal control problem from Example 14.2 where $C = C(t,u)$ is independent of the remaining amount y of oil in the reservoir and strictly convex in the rate of extraction u, i.e., $C_{uu} > 0$. When time horizon T can be chosen freely the optimal control problem is

$$\max_{u,T} \int_0^T [p(t)u(t) - C(t,u(t))]e^{-rt}\,dt, \qquad u(t) \geq 0,$$

$$y'(t) = -u(t), \quad y(0) = y_0, \quad y(T) \geq 0\,.$$

What does the maximum principle imply for this problem?

SOLUTION. Suppose (y^*, u^*) solves our problem. The Hamiltonian is given by

$$\mathscr{H}(t,y,u,\lambda) = \left[p(t)u - C(t,u)\right]e^{-rt} + \lambda(-u)\,.$$

The maximum principle implies that there exists a continuous function $\lambda(t)$ such that

(1) u^* maximizes $\mathscr{H}(t, y^*, u, \lambda)$ subject to $u \geq 0$.

(2) $\lambda'(t) = -\mathscr{H}_y = 0$.

(3) $\lambda(T^*) \geq 0$ where $\lambda(T^*) = 0$ if $y^*(T) > 0$.

(4) $\left[p(T^*)u(T^*) - C(T^*, u^*(T^*))\right]e^{-rT^*} = \lambda(T^*)u(T^*)$ (from (iv)).

Properties (2) and (3) imply that $\lambda(t) = \bar{\lambda} \geq 0$ is constant.

Now assume that $u^*(t) > 0$. Then Property (1) implies

$$0 = \mathscr{H}_u(t, y, u^*(t), \lambda) = \left[p(t) - C_u(t, u^*(t))\right] e^{-rt} - \bar{\lambda}$$

and thus

$$p(t) - C_u(t, u^*(t)) = \bar{\lambda} e^{rt}$$

and in particular

$$p(T^*) - C_u(T^*, u^*(T^*)) = \bar{\lambda} e^{rT^*}.$$

Since $C_{uu} > 0$ by our assumptions, $\mathscr{H}(t, y, u, \lambda)$ is concave in u and thus this condition is also sufficient for u^* being a maximum of $\mathscr{H}(t, y, u, \lambda)$ w.r.t. u. Observe that the l.h.s. of this equation is just the marginal profit. On the other hand Property (4) implies

$$p(T^*) - \frac{C(T^*, u^*(T^*))}{u(T^*)} = \bar{\lambda} e^{rT^*}$$

which implies

$$C_u(T^*, u^*(T^*)) = \frac{C(T^*, u^*(T^*))}{u(T^*)}.$$

We can conclude that one should terminate extraction at a time when the marginal cost of extraction is equal to average cost. ♢

It is important to note that we did *not* prove that there exists an optimal solution.

In fact concavity of the Hamiltonian in (y, u) is *not* sufficient for optimality if T is free!

— Exercises

14.1 Solve the following control problem:

$$\max_{u(t)\in(-\infty,\infty)} \int_0^2 [e^t x(t) - u(t)^2]dt$$

$$x'(t) = -u(t), \quad x(0) = 0, \quad x(2) \text{ free}$$

14.2 Solve the following control problem:

$$\max_{u(t)\in(-\infty,\infty)} \int_0^1 [1 - u(t)^2]dt$$

$$x'(t) = x(t) + u(t), \quad x(0) = 1, \quad x(1) \text{ free}$$

14.3 Solve the following control problem:

$$\min_{u(t)\in(-\infty,\infty)} \int_0^1 [x(t) + u(t)^2]dt$$

$$x'(t) = -u(t), \quad x(0) = 0, \quad x(1) \text{ free}$$

14.4 Solve the following control problem:

$$\max_{u(t)\in(-\infty,\infty)} \int_0^{10} [1 - 4x(t) - 2u(t)^2]dt$$

$$x'(t) = u(t), \quad x(0) = 0, \quad x(10) \text{ free}$$

14.5 Solve the following control problem:

$$\max_{u(t)\in(-\infty,\infty)} \int_0^T [x(t) - u(t)^2]dt$$

$$x'(t) = x(t) + u(t), \quad x(0) = 0, \quad x(T) \text{ free}$$

A

Complex Numbers

We need an imaginary extension
when we want to understand the real world.

Imaginary numbers and complex numbers do not seem of any practical relevance for economics. Nevertheless, they are inevitable when we want to understand mathematical structures.

A quadratic equation has exactly two convex roots.

Thus we shortly describe the basic properties of these numbers.

Imaginary and Complex Numbers

Definition A.1

We introduce a new number i (**imaginary unit**) with the property

$$i^2 = -1$$

and assume that we can add and multiply i like ordinary real numbers.

We then immediately find

(1) $i \notin \mathbb{R}$,

(2) $\sqrt{-1} = i$,

(3) $i^2 = -1, \quad i^3 = -i, \quad i^4 = 1, \quad i^5 = i, \quad \dots$

Definition A.2

Numbers of the form bi, $b \in \mathbb{R}$ (e.g., $5i$), are called **imaginary numbers**. Numbers of the form $z = a + bi$, $a, b \in \mathbb{R}$ (e.g., $3 + 5i$), are called **complex numbers**. $\text{Re}(z) = a$ is then called the **real part** of z, $\text{Im}(z) = b$ is called the **imaginary part** of z.
The set of all complex number is denoted by $\mathbb{C}$.

Complex variables are often denoted by z.

Notice that all real numbers are complex numbers with imaginary part $\text{Im}(z) = 0$. Thus $\mathbb{R} \subset \mathbb{C}$.

Example A.3

$z = 3+7i$ is a complex number with real part $\text{Re}(3+7i) = 3$, and imaginary part $\text{Im}(3+7i) = 7$. ♢

Computing with Complex Numbers

We can compute with complex numbers in a similar way as with real numbers. We have the following rules.

(a) $z_1 + z_2 = (a_1 + b_1 i) + (a_2 + b_2 i) = (a_1 + a_2) + (b_1 + b_2)i$

(b) $z_1 \cdot z_2 = (a_1 + b_1 i) \cdot (a_2 + b_2 i) = (a_1 a_2 - b_1 b_2) + (a_1 b_2 + a_2 b_1)i$

(c) $\dfrac{a_1 + b_1 i}{a_2 + b_2 i} = \left(\dfrac{a_1 a_2 + b_1 b_2}{a_2^2 + b_2^2}\right) + \left(\dfrac{b_1 a_2 - a_1 b_2}{a_2^2 + b_2^2}\right) i$

Notice, however, that there is no ordering of the complex numbers that is compatible with multiplication. That is, it is not possible to define a relation $<$ such that $u < v$ and $c > 0$ implies $uc < vc$. In particular there is no such thing called a *positive complex number*.

Example A.4

$(5 + 4i) + (3 - 2i) = 8 + 2i$

$2 \cdot (3 - i) = 6 - 2i$

$(1 + i) \cdot (2 - 2i) = 2 - 2i + 2i - 2i^2 = 2 - 2i + 2i + 2 = 4$

$\dfrac{3 + 2i}{2 - 3i} = \dfrac{6 - 6}{4 + 9} + \dfrac{4 + 9}{4 + 9} i = i$ ◇

Complex Conjugate

Definition A.5

Let $z = a + bi$ be some complex number. Then $\bar{z} = a - bi$ is called the **complex conjugate** of z.

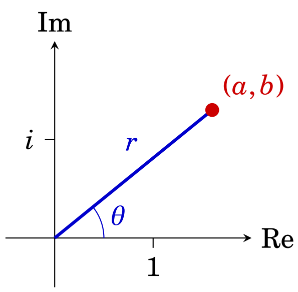
If z is a complex root of $x^2 + a_1 x + a_2$, $(a_1, a_2 \in \mathbb{R})$, then $\bar{z}$ is also a root of this equation.

Example A.6

The roots of the quadratic equation $z^2 - 4z + 5 = 0$ are

$$z_{1,2} = 2 \pm \sqrt{4 - 5} = 2 \pm i \qquad \Rightarrow \qquad z_1 = 2 + i, \quad z_2 = 2 - i = \overline{z_1}$$ ◇

The Gaussian Plane



Complex numbers $z = a + bi$ can be identified with points $(a, b) \in \mathbb{R}^2$ in the real plane. The axes are then called the **real axis** and **imaginary axis**, respectively.

The **modulus** or **absolute value** $|z|$ of z is defined as

$$|z| = \sqrt{z \bar{z}} = \sqrt{a^2 + b^2}$$

which is just the Euclidean distance of the point representing z to the origin.

Complex numbers also can be represented by polar coordinates (r, θ) where $r = |z|$ and $\theta \in [0, 2\pi)$ is the angle to the real axis and called the **argument** of z, $\arg(z)$. Then

$$\mathrm{Re}(z) = r \cos\theta \qquad \text{and} \qquad \mathrm{Im}(z) = r \sin\theta$$

and hence z can be represented in **polar form** as

$$z = r(\cos\theta + i\,\sin\theta)\,.$$

Recall that $\tan(\theta) = \frac{b}{a}$ and thus

$$\arg(z) = \theta = \arctan(b/a)\,.$$

The representation $z = a + bi$ is also called the **rectangular form** of the complex number z.

Example A.7

$z = 2\left((\cos\left(\frac{\pi}{3}\right) + i\,\sin\left(\frac{\pi}{3}\right)\right)$ has rectangular form $z = 1 + \sqrt{3}\,i$.
$z = 2 - 2\sqrt{3}i$ has polar form $z = 4\left(\cos\left(\frac{5}{3}\pi\right) + i\,\sin\left(\frac{5}{3}\pi\right)\right)$. ♢

Multiplication and division can be performed in polar form by the following rules

$$z_1 \cdot z_2 = r_1 \cdot r_2 \left(\cos(\theta_1 + \theta_2) + i\,\sin(\theta_1 + \theta_2)\right)$$
$$\frac{z_1}{z_2} = \frac{r_1}{r_2}\left(\cos(\theta_1 - \theta_2) + i\,\sin(\theta_1 - \theta_2)\right)$$

Powers of integer order n can be computed by means of **de Moivre's formula**:

De Moivre's formula does not hold in general for $n \notin \mathbb{Z}$.

$$z^n = r^n\left(\cos(n\theta) + i\,\sin(n\theta)\right) \qquad (n \in \mathbb{Z}).$$

This formula also can be used to derive an explicit expression for the nth roots of a complex number z.

$$z^{1/n} = [r(\cos\theta + i\sin\theta)]^{1/n} = r^{1/n}\left[\cos\left(\frac{\theta + 2k\pi}{n}\right) + i\sin\left(\frac{\theta + 2k\pi}{n}\right)\right]$$

where k is an integer. To get the n different nth roots of z one needs to consider all values of $k = 0, 1, \dots, n-1$.

Euler's Formula

There is a very important relation, called **Euler's formula**, between the complex exponential function and trigonometric functions,

$$e^{ix} = \cos x + i\,\sin x\,.$$

In particular we may express the polar form of a complex number z using the exponential function:

$$z = a + bi = |z|\,e^{i\theta}\,.$$

Furthermore, Euler's formula provides a tool to evaluate the exponential function for complex arguments.

$$e^z = e^{a+bi} = e^a\,[\cos(b) + i\,\sin(b)]\,.$$

— Exercises

A.1 Compute

(a) $(2i-1)+i$ (b) $(2i-1)\cdot i$ (c) $(4-2i)\cdot\overline{(4-2i)}$

(d) $(3+i)^2$ (e) i^{23} (f) $(3-i):(3+i)$

A.2 Transform into polar form and rectangular form, respectively.

(a) $2(\cos\frac{\pi}{4}+i\,\sin\frac{\pi}{4})$ (b) $4e^{-i\pi/3}$ (c) $5(\cos\frac{\pi}{2}+i\,\sin\frac{\pi}{2})$

(d) i (e) $1-i$ (f) $\frac{3}{2}+\frac{\sqrt{3}}{2}i$

A.3 Solve the quadratic equations:

(a) $x^2-3x+9=0$ (b) $2x^2+x+8=0$

(c) $x^2+2x+17=0$ (d) $2x^2-4x+4=0$

A.4 Compute

(a) e^i (b) e^{2-i} (c) $e^{5+\pi i}$

— Problems

A.5 Taylor's theorem also holds for complex numbers and complex valued functions. Use the MacLaurin series for exp, sin and cos and verify Euler's formula $e^{ix}=\cos x+i\,\sin x$.

A.6 Derive Euler's identity

$$e^{i\pi}+1=0$$

from Euler's formula.

Euler's identity links the five most fundamental mathematical constants into one equation. It has been elected to be the *most beautiful theorem in mathematics*.

Solutions

2.1 (a) 7; (b) $\frac{2}{7}$; (c) 0; (d) divergent with $\lim_{n\to\infty}\frac{n^2+1}{n+1}=\infty$; (e) divergent; (f) $\frac{29}{6}$.

2.2 (a) divergent; (b) 0; (c) $e^2 \approx 7{,}38906$; (d) $e^{-2}\approx 0.135335$; (e) 0; (f) 1;
(g) divergent with $\lim_{n\to\infty}\frac{n}{n+1}+\sqrt{n}=\infty$; (h) 0.

2.3 (a) e^x; (b) e^x; (c) $e^{1/x}$.

2.11 By Lemma 2.20 we find $\sum_{k=1}^{\infty} q^n = q\sum_{k=0}^{\infty} q^n = \frac{q}{1-q}$.

4.1 (a) 0, (b) 0, (c) ∞, (d) $-\infty$, (e) 1.

4.2 The functions are continuous in
(a) D, (b) D, (c) D, (d) D, (e) D, (f) $\mathbb{R}\setminus\mathbb{Z}$, (g) $\mathbb{R}\setminus\{2\}$.

4.3 (a) $6x-5\sin(x)$; (b) $6x^2+2x$; (c) $1+\ln(x)$; (d) $-2x^{-2}-2x^{-3}$; (e) $\frac{3x^2+6x+1}{(x+1)^2}$; (f) 1; (g) $18x-6$; (h) $6x\cos(3x^2)$; (i) $\ln(2)\cdot 2^x$; (j) $4x-1$; (k) $6e^{3x+1}(5x^2+1)^2+40e^{3x+1}(5x^2+1)x+\frac{3(x-1)(x+1)^2-(x+1)^3}{(x-1)^2}-2$.

4.4

	$f'(x)$	$f''(x)$	$f'''(x)$
(a)	$-x\,e^{-\frac{x^2}{2}}$	$(x^2-1)\,e^{-\frac{x^2}{2}}$	$(3x-x^3)\,e^{-\frac{x^2}{2}}$
(b)	$\frac{-2}{(x-1)^2}$	$\frac{4}{(x-1)^3}$	$\frac{-12}{(x-1)^4}$
(c)	$3x^2-4x+3$	$6x-4$	6

4.5 Derivatives:

	(a)	(b)	(c)	(d)	(e)	(f)
f_x	1	y	$2x$	$2xy^2$	$\alpha x^{\alpha-1}y^{\beta}$	$x(x^2+y^2)^{-1/2}$
f_y	1	x	$2y$	$2x^2y$	$\beta x^{\alpha}y^{\beta-1}$	$y(x^2+y^2)^{-1/2}$
f_{xx}	0	0	2	$2y^2$	$\alpha(\alpha-1)x^{\alpha-2}y^{\beta}$	$(x^2+y^2)^{-1/2}-x^2(x^2+y^2)^{-3/2}$
$f_{xy}=f_{yx}$	0	1	0	$4xy$	$\alpha\beta x^{\alpha-1}y^{\beta-1}$	$-xy(x^2+y^2)^{-3/2}$
f_{yy}	0	0	2	$2x^2$	$\beta(\beta-1)x^{\alpha}y^{\beta-2}$	$(x^2+y^2)^{-1/2}-y^2(x^2+y^2)^{-3/2}$

Derivatives at $(1,1)$:

	(a)	(b)	(c)	(d)	(e)	(f)
f_x	1	1	2	2	α	$\sqrt{2}/2$
f_y	1	1	2	2	β	$\sqrt{2}/2$
f_{xx}	0	0	2	2	$\alpha(\alpha-1)$	$\sqrt{2}/4$
$f_{xy}=f_{yx}$	0	1	0	4	$\alpha\beta$	$-\sqrt{2}/4$
f_{yy}	0	0	2	2	$\beta(\beta-1)$	$\sqrt{2}/4$

4.6 (a) $f'(1,1)=(1,1)$, $f''(1,1)=\begin{pmatrix}0&0\\0&0\end{pmatrix}$;

(b) $f'(1,1)=(1,1)$, $f''(1,1)=\begin{pmatrix}0&1\\1&0\end{pmatrix}$;

(c) $f'(1,1)=(2,2)$, $f''(1,1)=\begin{pmatrix}2&0\\0&2\end{pmatrix}$;

(d) $f'(1,1) = (2,2)$, $f''(1,1) = \begin{pmatrix} 2 & 4 \\ 4 & 2 \end{pmatrix}$;

(e) $f'(1,1) = (\alpha,\beta)$, $f''(1,1) = \begin{pmatrix} \alpha(\alpha-1) & \alpha\beta \\ \alpha\beta & \beta(\beta-1) \end{pmatrix}$;

(f) $f'(1,1) = (\sqrt{2}/2, \sqrt{2}/2)$, $f''(1,1) = \begin{pmatrix} \sqrt{2}/4 & -\sqrt{2}/4 \\ -\sqrt{2}/4 & \sqrt{2}/4 \end{pmatrix}$.

4.7 $\frac{\partial f}{\partial \mathbf{a}} = 2\mathbf{x}' \cdot \mathbf{a}$.

4.8 $\nabla f(0,0) = (4/\sqrt{10}, 12/\sqrt{10})$.

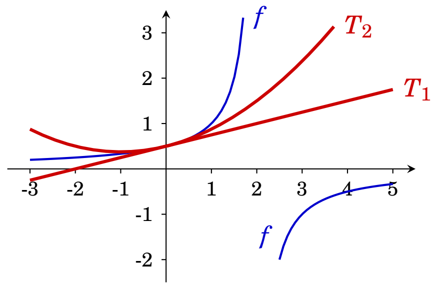
4.9 $D(f \circ g)(t) = 2t + 4t^3$; $D(g \circ f)(x,y) = \begin{pmatrix} 2x & 2y \\ 4x^3 + 4xy^2 & 4x^2y + 4y^3 \end{pmatrix}$.

4.10 $D(\mathbf{f} \circ \mathbf{g})(\mathbf{x}) = \begin{pmatrix} -1 & 6x_2^5 \\ -3x_1^2 & 2x_2 \end{pmatrix}$; $D(\mathbf{g} \circ \mathbf{f})(\mathbf{x}) = \begin{pmatrix} 2(x_1 - x_2^3) & 6(-x_1x_2^2 + x_2^5) \\ 3x_2^2 & -1 \end{pmatrix}$.

4.11 $D\mathbf{x}(\mathbf{b}) = \mathbf{A}^{-1}$ and thus (by Cramer's rule) $\frac{\partial x_i}{\partial b_j} = (-1)^{i+j} M_{ji}/|\mathbf{A}|$ where M_{ki} denotes the the (k,i) minor of $\mathbf{A}$.

4.12 $\frac{d}{dt} F(K(t), L(t), t) = F_K(K,L,t)K'(t) + F_L(K,L,t)L'(t) + F_t(K,L,t)$.

5.1



(a) $f(x) \approx T_1(x) = \frac{1}{2} + \frac{1}{4}x$,

(b) $f(x) \approx T_2(x) = \frac{1}{2} + \frac{1}{4}x + \frac{1}{8}x^2$.

radius of convergence $\rho = 2$.

5.2 $T_{f,0,3}(x) = 1 + \frac{1}{2}x - \frac{1}{8}x^2 + \frac{1}{16}x^3$.

5.3 $T_{f,0,30}(x) = x^{10} - \frac{1}{6}x^{30}$.

5.4 $T_{f,0,4}(x) \approx 0.959 + 0.284x^2 - 0.479x^4$.

5.5 $f(x) = \sum_{n=0}^{\infty} (-1)^n x^{2n}$; $\rho = 1$.

5.6 $f(x) = \sum_{n=0}^{\infty} (-\frac{1}{2})^n \frac{1}{n!} x^{2n}$; $\rho = \infty$.

5.7 $f(x,y) = 1 + x^2 + y^2 + O(\|(x,y)\|^3)$.

5.8 $|R_n(1)| \le \frac{e}{(n+1)!}$; $|R_n(1)| < 10^{-16}$ if $n \ge 18$.

6.1 (a) $D\mathbf{f}(\mathbf{x}) = \begin{pmatrix} 1 - x_2 & -x_1 \\ x_2 & x_1 \end{pmatrix}$, $\frac{\partial(y_1,y_2)}{\partial(x_1,x_2)} = x_1$;

(b) for all images of points (x_1, x_2) with $x_1 \ne 0$;

(c) $D(\mathbf{f}^{-1})(\mathbf{y}) = (D\mathbf{f}(\mathbf{x}))^{-1} = \begin{pmatrix} 1 - x_2 & -x_1 \\ x_2 & x_1 \end{pmatrix}^{-1} = \frac{1}{x_1} \begin{pmatrix} x_1 & x_1 \\ -x_2 & 1 - x_2 \end{pmatrix}$;

(d) in order to get the inverse function we have to solve equation $\mathbf{f}(\mathbf{x}) = \mathbf{y}$: $x_1 = y_1 + y_2$ and $x_2 = y_2/(y_1 + y_2)$, if $y_1 + y_2 \ne 0$.

6.2 T is the linear map given by the matrix $\mathbf{T} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$. Hence its Jacobian matrix is just $\det(\mathbf{T})$. If $\det(\mathbf{T}) = 0$, then the columns of $\mathbf{T}$ are linearly dependent. Since the constants are non-zero, $\mathbf{T}$ has rank 1 and thus the image is a linear subspace of dimension 1, i.e., a straight line through the origin.

6.3 Let $J = \begin{vmatrix} \frac{\partial f}{\partial x} & \frac{\partial f}{\partial y} \\ \frac{\partial g}{\partial x} & \frac{\partial g}{\partial y} \end{vmatrix}$ be the Jacobian determinant of this function. Then the equation can be solved locally if $J \neq 0$. We then have $\frac{\partial F}{\partial u} = \frac{1}{J}\frac{\partial g}{\partial y}$ and $\frac{\partial G}{\partial u} = -\frac{1}{J}\frac{\partial g}{\partial x}$.

6.4 (a) $F_y = 3y^2 + 1 \neq 0$, $y' = -F_x/F_y = 3x^2/(3y^2+1) = 0$ for $x = 0$;
(b) $F_y = 1 + x\cos(xy) = 1 \neq 0$ for $x = 0$, $y'(0) = 0$.

6.5 $\frac{dy}{dx} = -\frac{2x}{3y^2}$, $y = f(x)$ exists locally in an open rectangle around $\mathbf{x}_0 = (x_0, y_0)$ if $y_0 \neq 0$; $x = g(y)$ exists locally if $x_0 \neq 0$.

6.6 (a) $z = g(x,y)$ can be locally expressed since $F_z = 3z^2 - xy$ and $F_z(0,0,1) = 3 \neq 0$; $\frac{\partial g}{\partial x} = -\frac{F_x}{F_z} = -\frac{3x^2 - yz}{3z^2 - xy} = -\frac{0}{3} = 0$ for $(x_0, y_0, z_0) = (0,0,1)$; $\frac{\partial g}{\partial y} = -\frac{F_y}{F_z} = -\frac{3y^2 - xz}{3z^2 - xy} = -\frac{0}{3} = 0$.
(b) $z = g(x,y)$ can be locally expressed since $F_z = \exp(z) - 2z$ and $F_z(1,0,0) = 1 \neq 0$; $\frac{\partial g}{\partial x} = -\frac{F_x}{F_z} = -\frac{-2x}{\exp(z)-2z} = 2$ for $(x_0, y_0, z_0) = (1,0,0)$; $\frac{\partial g}{\partial y} = -\frac{F_y}{F_z} = -\frac{-2y}{\exp(z)-2z} = 0$ for $(x_0, y_0, z_0) = (1,0,0)$.

6.7 $\frac{dK}{dL} = -\frac{\beta K}{\alpha L}$.

6.8 (a) $\frac{dx_i}{dx_j} = -\frac{u_{x_j}}{u_{x_i}} = -\frac{\left(x_1^{\frac{1}{2}} + x_2^{\frac{1}{2}}\right) x_j^{-\frac{1}{2}}}{\left(x_1^{\frac{1}{2}} + x_2^{\frac{1}{2}}\right) x_i^{-\frac{1}{2}}} = -\frac{x_i^{\frac{1}{2}}}{x_j^{\frac{1}{2}}}$;

(b) $\frac{dx_i}{dx_j} = -\frac{u_{x_j}}{u_{x_i}} = -\frac{\frac{\theta}{\theta-1}\left(\sum_{i=1}^n x_i^{\frac{\theta-1}{\theta}}\right)^{\frac{1}{\theta-1}} \frac{\theta-1}{\theta} x_j^{-\frac{1}{\theta}}}{\frac{\theta}{\theta-1}\left(\sum_{i=1}^n x_i^{\frac{\theta-1}{\theta}}\right)^{\frac{1}{\theta-1}} \frac{\theta-1}{\theta} x_i^{-\frac{1}{\theta}}} = -\frac{x_i^{\frac{1}{\theta}}}{x_j^{\frac{1}{\theta}}}$.

7.1 (a) decreasing in $(-\infty, -4] \cup [0,3]$, increasing in $[-4,0] \cup [3,\infty)$; (b) concave in $[-2 - \sqrt{148})/6, -2 + \sqrt{148})/6]$, convex otherwise.

7.2 (a) log-concave; (b) not log-concave; (c) not log-concave; (d) log-concave on $(-1,1)$.

7.3 (a) convex; (b) convex.

8.1 (a) global minimum at $x = 3$ ($f''(x) \geq 0$ for all $x \in \mathbb{R}$), no local maximum;
(b) local minimum at $x = 1$, local maximum at $x = -1$, no global extrema.

8.2 (a) global minimum in $x = 1$, no local maximum;
(b) global maximum in $x = \frac{1}{4}$, no local minimum;
(c) global minimum in $x = 0$, no local maximum.

8.3 (a) stationary point: $\mathbf{p}_0 = (0,0)$, $\mathbf{H}_f = \begin{pmatrix} -2 & 1 \\ 1 & 2 \end{pmatrix}$,
$H_2 = -5 < 0$, $\Rightarrow$ $\mathbf{p}_0$ is a saddle point;
(b) stationary point: $\mathbf{p}_0 = (e,0)$, $\mathbf{H}_f(\mathbf{p}_0) = \begin{pmatrix} -e^{-3} & 0 \\ 0 & -2 \end{pmatrix}$,
$H_1 = -e^{-3} < 0$, $H_2 = 2e^{-3} > 0$, $\Rightarrow$ $\mathbf{p}_0$ is local maximum;
(c) stationary point: $\mathbf{p}_0 = (1,1)$, $\mathbf{H}_f(\mathbf{p}_0) = \begin{pmatrix} 802 & -400 \\ -400 & 200 \end{pmatrix}$,
$H_1 = 802 > 0$, $H_2 = 400 > 0$, $\Rightarrow$ $\mathbf{p}_0$ is local minimum;
(d) stationary point: $\mathbf{p}_0 = (\ln(3), \ln(4))$, $\mathbf{H}_f = \begin{pmatrix} -e^{x_1} & 0 \\ 0 & -e^{x_2} \end{pmatrix}$,
$H_1 = -e^{x_1} < 0$, $H_2 = e^{x_1} \cdot e^{x_2} > 0$, $\Rightarrow$ local maximum in $\mathbf{p}_0 = (\ln(3), \ln(4))$.

8.4 stationary points: $\mathbf{p}_1 = (0,0,0)$, $\mathbf{p}_2 = (1,0,0)$, $\mathbf{p}_3 = (-1,0,0)$,
$\mathbf{H}_f = \begin{pmatrix} 6x_1x_2 & 3x_1^2-1 & 0 \\ 3x_1^2-1 & 0 & \\ 0 & 0 & 2 \end{pmatrix}$,
leading principle minors: $H_1 = 6x_1x_2 = 0$, $H_2 = -(3x_1^2-1)^2 < 0$ (da $x_1 \in \{0,-1,1\}$), $H_3 = -2(3x_1^2-1)^2 < 0$,
$\Rightarrow$ all three stationary points are saddle points. The function is neither convex nor concave.

8.5 (b) Lagrange function: $\mathscr{L}(x,y;\lambda) = x^2y + \lambda(3-x-y)$,
stationary points $\mathbf{x}_1 = (2,1;4)$ and $\mathbf{x}_2 = (0,3;0)$,
(c) bordered Hessian: $\bar{\mathbf{H}} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 2y & 2x \\ 1 & 2x & 0 \end{pmatrix}$,
$\bar{\mathbf{H}}(\mathbf{x}_1) = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 2 & 4 \\ 1 & 4 & 0 \end{pmatrix}$, $\det(\bar{\mathbf{H}}(\mathbf{x}_1)) = 6 > 0$, $\Rightarrow \mathbf{x}_1$ is a local maximum,
$\bar{\mathbf{H}}(\mathbf{x}_2) = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 6 & 0 \\ 1 & 0 & 0 \end{pmatrix}$, $\det(\bar{\mathbf{H}}(\mathbf{x}_2)) = -6 \Rightarrow \mathbf{x}_2$ is a local minimum.

8.6 Lagrange function: $\mathscr{L}(x_1,x_2,x_3;\lambda_1,\lambda_2) = f(x_1,x_2,x_3) = \frac{1}{3}(x_1-3)^3 + x_2x_3 + \lambda_1(4-x_1-x_2) + \lambda_2(5-x_1-x_3)$,
stationary points: $\mathbf{x}_1 = (0,4,5;5,4)$ and $\mathbf{x}_2 = (4,0,1;1,0)$; $\det(\bar{\mathbf{H}}(x_1,x_2,x_3)) = 2x_1-4$; $\mathbf{x}_1$ is a local maximum; $\mathbf{x}_2$ is a local minimum.

8.7 (a) $x_1 = \alpha\frac{m}{p_1}$, $x_2 = (1-\alpha)\frac{m}{p_2}$ and $\lambda = \frac{1}{m}$, (c) marginal change for optimum: $\frac{1}{m}$.

8.8 Kuhn-Tucker theorem: $\mathscr{L}(x,y;\lambda) = -(x-2)^2 - y + \lambda(1-x-y)$, $x = 1$, $y = 0$, $\lambda = 2$.

9.1 (a) integration by parts (P): $\frac{1}{4}x^2(2\ln x - 1) + c$;
(b) 2×P: $2\cos(x) - x^2\cos(x) + 2x\sin(x) + c$;
(c) by substitution (S), $z = x^2+6$: $\frac{2}{3}\left(x^2+6\right)^{\frac{3}{2}} + c$;
(d) S, $z = x^2$: $\frac{1}{2}e^{x^2} + c$;
(e) S, $z = 3x^2+4$: $\frac{1}{6}\ln(4+3x^2) + c$;
(f) P or S, $z = x+1$: $\frac{2}{5}(x+1)^{\frac{5}{2}} - \frac{2}{3}(x+1)^{\frac{3}{2}} + c$;
(g) $= \int 3x + \frac{4}{x}\,dx = \frac{3}{2}x^2 + 4\ln(x) + c$; S not suitable;
(h) S, $z = \ln(x)$: $\frac{1}{2}(\ln(x))^2 + c$.

9.2 (a) 39, (b) $3e^2 - 3 \approx 19.17$, (c) 93, (d) $-\frac{1}{6}$ (use radiant instead of degree), (e) $\frac{1}{2}\ln(8) \approx 1.0397$

9.3 (a) $\int_0^\infty -e^{-3x}\,dx = \lim_{t\to\infty}\int_0^t -e^{-3x}\,dx = \lim_{t\to\infty}\frac{1}{3}e^{-3t} - \frac{1}{3} = -\frac{1}{3}$;
(b) $\int_0^1 \frac{2}{\sqrt[4]{x^3}}\,dx = \lim_{t\to 0}\int_t^1 \frac{2}{\sqrt[4]{x^3}}\,dx = \lim_{t\to 0} 8 - 8t^{\frac{1}{4}} = 8$;
(c) $= \lim_{t\to\infty}\int_0^t \frac{x}{x^2+1}\,dx = \lim_{t\to\infty}\frac{1}{2}\int_2^{t^2+1}\frac{1}{z}\,dz = \lim_{t\to\infty}\frac{1}{2}(\ln(t^2+1) - \ln(2)) = \infty$,
the improper integral does not exist.

9.4 We need the antiderivative $C(x)$ of $C'(x) = 30 - 0.05x$ with $C(0) = 2000$: $C(x) = 2000 + 30x - 0{,}025x^2$.

9.5 $\mathbb{E}(X) = \sqrt{\frac{2}{\pi}}$.

9.6 $\mathbb{E}(X) = -\sqrt{\frac{2}{\pi}} + \sqrt{\frac{2}{\pi}} = 0.$

9.7 $F(x) = \begin{cases} 0, & \text{for } x \le -1, \\ \frac{1}{2} + \frac{2x+x^2}{2}, & \text{for } -1 < x \le 0, \\ \frac{1}{2} + \frac{2x-x^2}{2}, & \text{for } 0 < x \le 1, \\ 1, & \text{for } x > 1. \end{cases}$

9.8 (a) The improper integral exists if and only if $\alpha > -1$;
(b) the improper integral exists if and only if $\alpha < -1$;
(c) the improper integral never converges.

10.1 (a) 16; (b) $\frac{a^2b^2}{4}$; (c) -2; (d) $\frac{\pi-2}{8\pi}$.

10.2 π.

11.1 Method: Separation of variables.
(a) $y = C\,t^k$, $y = t^k$;
(b) $y = C\,t - 1$, $y = 2t - 1$;
(c) $y = C\,e^{\frac{1}{2}t^2}$, $y = \frac{1}{\sqrt{e}}e^{\frac{t^2}{2}}$;
(d) $y = -\ln(t+c)$, $y = -\ln\left(t - 1 + \frac{1}{e}\right)$;
(e) $y = -\frac{1}{t+c}$, $y = -\frac{1}{t-2}$;
(f) $y = \left(\frac{1}{5}t^{\frac{5}{2}} + c\right)^2$, $y = \left(\frac{1}{5}t^{\frac{5}{2}} + \frac{4}{5}\right)^2$; (observe that $y = \left(\frac{1}{5}t^{\frac{5}{2}} - \frac{6}{5}\right)^2$ also satisfies $y(1) = 1$ but does not solve the differential equation).

11.2 Solution of initial value problem: $y(t) = \frac{8}{7}e^{-6t} - \frac{1}{7}e^t$.

11.3 (a) $U' = \frac{\alpha}{U}$, (b) $U(x) = \sqrt{2\alpha x + c}$, (c) $U(0) = 0$, $U(x) = \sqrt{2\alpha x}$.

11.4 $p(t) = (p_0 - \bar{p})\exp(-j\frac{\beta+\delta}{1-j\nu}t) + \bar{p}$. $p_0 = p(0)$, $\bar{p} = \frac{\alpha+\gamma}{\beta+\delta}$ is equilibrium price.

11.5 (a) $A(t) = \frac{L}{1+C\exp(-Lkt)}$, t number of months. $L = 96000$, using $A(0) = 4000$ and $A(2) = 12000$: $C = 23$ and $k = 0{,}000006196$.
$A(t) = \frac{96000}{1+23\exp(-0{,}59479t)}$. (b) $A(6) = 58238$, (c) $t_{\frac{2}{3}} = 7{,}12$ months.

11.6 Case $a \ne 0$: $y(t) = C\,e^{-at} + \frac{s}{a}$ for some $C > 0$;
case $a = 0$: $y(t) = st + c$ for some $c \in \mathbb{R}$.

11.7 Case $a \ne 0$: $y(t) = (y_0 - \frac{s}{a})e^{-a(t-t_0)} + \frac{s}{a}$;
case $a = 0$: $y(x) = s(t - t_0) + y_0$.

11.10 (a) $y(t) = t^2$; (b) $y_a'(t) = 0$ for $t \le a$ and $y_a'(t) = 2(t-a)$ for $t > a$; (c) $2\sqrt{y_a(t)} = 2\sqrt{(t-a)^2} = 2(t-a) = y_a'(t)$.

12.1 $y(x) = \frac{1}{12}(x^4 + 4x^3 - 30x^2 + 36x)$ (by integrating two times).

12.2 General solution: $y = y_h + y_p = C_1e^t + C_2e^{-2t} - \frac{3}{2}$;
particular solution: $y(t) = 2e^t + \frac{1}{2}e^{-2t} - \frac{3}{2}$.

12.3 General solution: $y = (C_1 + C_2t)e^{3t}$, particular solution: $y(t) = (2 - 6t)e^{3t}$.

12.4 $y(t) = e^{-t}[C_1\cos 4t + C_2\sin 4t]$; $\lim_{t\to\infty} y(t) = 0$.

12.5 Case 1 ($\gamma(a-\alpha) > 0$): $p(t) = C_1\exp(-t\sqrt{\gamma(a-\alpha)}) + C_1\exp(t\sqrt{\gamma(a-\alpha)}) - \frac{k}{\gamma(a-\alpha)}$;
Case 2 ($\gamma(a-\alpha) = 0$): $p(t) = C_1 + C_2t + \frac{1}{2}kt^2$;
Case 3 ($\gamma(a-\alpha) < 0$): $p(t) = C_1\cos(t\sqrt{|\gamma(a-\alpha)|}) + C_2\sin(t\sqrt{|\gamma(a-\alpha)|}) - \frac{k}{\gamma(a-\alpha)}$.

13.1 (a) $\mathbf{y}(t) = C_1 e^{-2t}\begin{pmatrix}2\\-1\end{pmatrix} + C_2 e^{-7t}\begin{pmatrix}1\\2\end{pmatrix}$, (b) $\mathbf{y}(t) = C_1 e^{t}\begin{pmatrix}3\\-1\end{pmatrix} + C_2 e^{14t}\begin{pmatrix}1\\4\end{pmatrix}$,
(c) $\mathbf{y}(t) = C_1 e^{4t}\begin{pmatrix}1\\1\end{pmatrix} + C_2 e^{-6t}\begin{pmatrix}1\\-1\end{pmatrix}$.

13.2 (a) $\mathbf{y}(t) = C_1 e^{-2t}\begin{pmatrix}2\\-1\end{pmatrix} + C_2 e^{-7t}\begin{pmatrix}1\\2\end{pmatrix} + \begin{pmatrix}1\\-1\end{pmatrix}$,
$\mathbf{y}^* = \begin{pmatrix}1\\-1\end{pmatrix}$ is asymptotically stable (sink);
(b) $\mathbf{y}(t) = C_1 e^{t}\begin{pmatrix}3\\-1\end{pmatrix} + C_2 e^{14t}\begin{pmatrix}1\\4\end{pmatrix} + \begin{pmatrix}-2\\1\end{pmatrix}$,
$\mathbf{y}^* = \begin{pmatrix}-2\\-1\end{pmatrix}$ is unstable equilibrium point (source);
(c) $\mathbf{y}(t) = C_1 e^{4t}\begin{pmatrix}1\\1\end{pmatrix} + C_2 e^{-6t}\begin{pmatrix}1\\-1\end{pmatrix} + \begin{pmatrix}1\\0\end{pmatrix}$, $\mathbf{y}^* = \begin{pmatrix}1\\0\end{pmatrix}$ is a saddle point.

13.3 (a) for all $a < 0$; (b) for all $a < -2$. (Olech's theorem)

13.4 (a) $\mathbf{F}'(x,y) = \begin{pmatrix}-1 & y\\ 2 & -2\end{pmatrix}$, $\mathbf{F}'(0,0) = \begin{pmatrix}-1 & 0\\ 2 & -2\end{pmatrix}$, all eigenvales negative, locally asymptotically stable by Theorem 13.15; Theorem 13.17 cannot be applied;
(b) $\mathbf{F}'(x,y) = \begin{pmatrix}-3x^2 & -1\\ 1 & -2y^2\end{pmatrix}$, eigenvalues of $\mathbf{F}'(0,0) = \begin{pmatrix}0 & -1\\ 1 & 0\end{pmatrix}$ are purely imaginary, none of our theorems can be applied;
(c) $\mathbf{F}'(x,y) = \begin{pmatrix}1+4x-y & -3-x+2y\\ 2-e^{x-y} & -1+e^{x-y}\end{pmatrix}$, $\mathbf{F}'(1,1) = \begin{pmatrix}4 & -2\\ 1 & 0\end{pmatrix}$, $\mathrm{tr}(\mathbf{F}'(1,1)) = 4 > 0$, not stable.

13.9 $\mathbf{F}'(x^*,y^*) = \begin{pmatrix}-\varepsilon x^* & -\beta x^*\\ \delta y^* & -\eta y^*\end{pmatrix}$, $\mathrm{tr}(\mathbf{F}'(x^*,y^*)) = -\varepsilon x^* - \eta y^* < 0$ and $\det(\mathbf{F}'(x^*,y^*)) = (\varepsilon\eta + \beta\delta)x^* y^* > 0$. Hence (x^*,y^*) is a locally asymptotically stable point by Theorem 13.15.

14.1 Optimal solution: $x^*(t) = \frac{1}{2}(e^2 t - e^t + 1)$; optimal control: $u^*(t) = \frac{1}{2}e^t - \frac{1}{2}e^2$.

14.2 Optimal solution: $x^*(t) = e^t$; optimal control: $u^*(t) = 0$.

14.3 Optimal solution: $x^*(t) = \frac{1}{4}t^2 - \frac{1}{2}t$; optimal control: $u^*(t) = -\frac{1}{2}\lambda(t) = -\frac{1}{2}t + \frac{1}{2}$.

14.4 Optimal solution: $x^*(t) = \frac{1}{2}t^2 - 10t$; optimal control: $u^*(t) = t - 10$.

14.5 Optimal solution: $x^*(t) = \frac{1}{4}e^{T+t} - \frac{1}{4}e^{T-t} - \frac{1}{2}e^t + \frac{1}{2}$; optimal control: $u^*(t) = \frac{1}{2}(e^{T-t} - 1)$.

A.1 (a) $-1+3i$, (b) $-2-i$, (c) 20, (d) $8+6i$, (e) $= i^3 \cdot i^{20} = -i$, (f) $\frac{4}{5} - \frac{3}{5}i$.

A.2 (a) $\sqrt{2}+\sqrt{2}\,i$, (b) $4(\cos(-\frac{\pi}{3}) + i\sin(-\frac{\pi}{3})) = 2 - 2\sqrt{3}\,i$, (c) $5\,i$, (d) $(\cos\frac{\pi}{2} + i\sin\frac{\pi}{2})$, (e) $\sqrt{2}(\cos\frac{\pi}{4} - i\sin\frac{\pi}{4})$, (f) $\sqrt{3}(\cos\frac{\pi}{6} + i\sin\frac{\pi}{6})$.

A.3 (a) $\frac{3}{2} \pm \frac{3\sqrt{3}}{2}i$, (b) $-\frac{1}{4} \pm \frac{3\sqrt{7}}{4}i$, (c) $-1 \pm 4i$, (d) $1 \pm i$.

A.4 (a) $e^i = \cos 1 + i\sin 1 \approx 0.54 + 0.841i$; (b) $e^{2-i} = e^2(\cos(-1) + i\sin(-1)) = e^2(\cos 1 - i\sin 1) \approx 3.992 - 6.218i$; (c) $e^{5+\pi i} = e^5(\cos\pi + i\sin\pi) = -e^5 \approx -148.4$.

Index

Entries in *italics* indicate lemmata and theorems.

absolute value, 163
absolutely convergent, 10
accumulation point, 15, 18
adjoint function, 154
admissible pair, 154
alternating harmonic series, 10
analytic, 52
antiderivative, 88
argument, 163
asymptotically stable equilibrium state, 142
autonomous, 145
autonomous differential equation, 123

Bolzano-Weierstrass, 21
bordered Hessian, 84
boundary, 15
boundary point, 14
bounded, 20
bounded sequence, 4

Cauchy sequence, 17
Cauchy's covergence criterion, 17
center, 144
Chain rule, 38
Change of variables in double integrals, 107
Change of variables in multiple integrals, 109
characteristic equation, 132
Characterization of continuity, 23
Characterization of quasi-convexity, 74
closed, 14
closure, 15
Closure and convergence, 18
co-state variable, 154
compact, 21
Comparison test, 8
comparison test, 8
complex conjugate, 163
complex numbers, 162
Composite functions, 68, 74
Computation of derivative, 37
concave, 64
Continuity and images of balls, 22
Continuity of each component, 21
continuous, 21
Continuous functions preserve compactness, 24
continuously differentiable functions, 34
control function, 153
control region, 153
Convergence of a monotone sequence, 4
Convergence of each component, 16
Convergence of geometric sequence, 5
Convergence of remainder, 49
convergent, 2, 7, 16
converges, 2, 16
convex, 63, 64
convex combinations, 63
convex hull, 63
Convex sum, 65
Convexity of multivariate functions, 70
Convexity of univariate functions, 70

de Moivre's formula, 164
definite integral, 95
derivative, 29, 36
difference quotient, 29
differentiable, 29, 36
differential coefficient, 29
differential operator, 29
direction field, 114, 149
directional derivative, 33, 34
Divergence of geometric sequence, 5
divergent, 2, 7, 16
double integral, 102

elasticity, 43
Envelope theorem, 81
equilibrium point, 124, 135, 145
equilibrium state, 123, 142
equivalent, 18
Euclidean distance, 13
Euclidean norm, 13
Euler's formula, 164
Existence and uniqueness, 125, 126, 130
Existence of derivatives, 38
expand, 48
expectation, 100
exterior point, 14
Extreme-value theorem, 24

Finitely generated vector space, 20
First fundamental theorem of calculus, 94
first-order linear differential equation, 117
first-order partial derivatives, 35
fixed point, 123
Fubini's theorem, 105

general solution, 114
geometric sequence, 5
Geometric series, 7
globally asymptotically stable, 135
gradient, 34

half spaces, 64
Hamiltonian, 154
Harmonic series, 8
Hessian, 35
homogeneous, 44
homogeneous linear differential equation, 118
hyperplane, 64

imaginary axis, 163
imaginary numbers, 162
imaginary part, 162
imaginary unit, 162
implicit function, 58
Implicit function theorem, 58, 59
improper integral, 96
indefinite integral, 88
infimum, 4
Inhomogeneous linear differential equations, 118
initial value, 114
initial value problem, 114, 130
integrable, 93
integral, 93
integral curve, 113
integration constant, 88
interior, 15
interior point, 13
Intermediate value theorem (Bolzano), 27
Intersection, 63
interval bisectioning, 27
inverse function, 56
Inverse function theorem, 57

Jacobian determinant, 57
Jacobian matrix, 37, 59
Jensen's inequality, discrete version, 67

Kuhn-Tucker conditions, 85
Kuhn-Tucker sufficient condition, 85

Lagrange function, 82
Lagrange multiplier, 82
Lagrange's form of the remainder, 47
Lagrangian, 82
Landau symbols, 50

Leibniz's formula, 98
Level sets of convex functions, 73
limit, 2, 16, 28
Linear approximation, 36
local maximum, 80
local minimum, 80
Local saddle point, 146
local saddle point, 146
locally asymptotically stable, 145
locally asymptotically stable equilibrium state, 124
locally Lipschitz continuous, 126
Logistic differential equation, 122
lower bound, 4
lower level set, 73
Lyapunov function, 148
Lyapunov's Theorem, 145, 148

Maclaurin polynomial, 47
Mangasarian's theorem, 155
maximum, 79
Mean value theorem, 32
minimum, 79
Minimum and maximum of two convex functions, 68
modulus, 163
monotone, 4
monotonically decreasing, 69
monotonically increasing, 69
Monotonicity and derivatives, 69
multiple integral, 106

n-dimensional rectangle, 105
Necessary condition, 82
Necessary first-order conditions, 79
nullclines, 149

objective function, 153
ODE, 113
Olech's Theorem, 146
open, 14
open ball, 13
open neighborhood, 14
optimal pair, 154
ordinary differential equation, 113

partial derivative, 33
partial sum, 7
particular, 131
particular solution, 114
phase diagram, 123, 149
Polar coordinates, 110
polar form, 164
power series, 51
preimage, 22
primitive integral, 88
Properties of closed sets, 15
Properties of open sets, 14
Properties of the gradient, 34

Quadratic form, 65
quasi-concave, 73
quasi-convex, 73

radius of convergence, 51
Ratio test, 10
ray, 44
real axis, 163
real part, 162
rectangular form, 164
relatively closed, 23
relatively open, 23
remainder, 46
restriction, 67
Riemann integrable, 92
Riemann integral, 92
Riemann sum, 92
Rolle's theorem, 32
Rules for limits, 5, 28

saddle path solution, 143, 146
saddle point, 81, 143
second derivative, 31
Second fundamental theorem of calculus, 94
second-order linear differential equation, 130
second-order ordinary differential equation, 130
second-order partial derivatives, 35
sequence, 2, 15
series, 7
sink, 142
slope field, 114

solution, 138
solution curve, 113
source, 142
Stability of an equlibrium point, 144
Stability of equilibrium points, 125
stable, 147
standard end constraint problem, 153
state variable, 153
stationary point, 79
stationary state, 123, 124
step function, 91
strict local maximum, 80
strict maximum, 79
strictly concave, 64
strictly convex, 64
strictly decreasing, 69
strictly increasing, 69
strictly quasi-concave, 76
strictly quasi-convex, 76
strong Lyapunov function, 148
subgradient, 67
Subgradient and supergradient, 67
subsequence, 20
Sufficient condition, 83
Sufficient condition for local optimum, 84
Sufficient conditions, 79
Sufficient conditions for local extremal points, 80
Sum of covergent sequences, 6, 17
supergradient, 67
supremum, 4

Tangents of convex functions, 66
Tangents of quasi-convex functions, 76
Taylor polynomial, 46
Taylor series, 48
Taylor's formula for multivariate functions, 53
Taylor's theorem, 47
The Maximum principle, 154
The Maximum principle with variable time, 159
trace, 144
transversality condition, 154
Triangle inequality, 6

unbounded, 20
unstable equilibrium point, 142
unstable equilibrium state, 124
upper bound, 4
upper level set, 73

value function, 81
variable final time problem, 159
variation of constants, 119
vector field, 149

Young's theorem, Schwarz' theorem, 35