# Selected Chapters in Mathematics
# Math Bridge Course for the
# Master of Biomedical Engineering

Andreas Stahel, Jamin Wandel, André Lisibach
Bern University of Applied Sciences, BFH-TI

Version of 4th September 2019

# Contents

# Chapter 1

# Introduction

## 1.1 The Goal

Not all students starting the first semester on the Master in Biomedical Engineering have the mathematical background usually taught at Swiss Universities of Applied Sciences in Engineering departments. We want to give these students the possibility to check whether they have the required skills, and catch up if necessary. The goal cannot be to provide a full bachelor program *Mathematics for Engineers*.

## 1.2 Path to the Goal

- A few key topics were chosen by the faculty of the master program on Biomedical Engineering.
- For each of these topics a few typical questions are given as exercise, including solutions. The students can (try to) solve these problems and will detect possible gaps in their mathematical background. For some topics a few key statements or results are shown in these notes, but the students are expected to consult their lecture notes from their Bachelor program. The students are expected to work through those sections and questions **before** the Question&Answer sessions in class. Then the students should decide whether they already master the topic of if they need to attend the session.
- Some pointers to appropriate literature are provided.
- Some additional topics will be presented. The goal is to expose the students early to those concepts, even if they have not learned about them during their bachelor program. The students are not expected to work through those section before class.

## 1.3 Survey of the Sections

- In chapter 2 : Linear Algebra some basic concept will be illustrated by typical examples and exercises. Some of the keywords are
  - matrices and linear mappings, solving systems of linear equations, row reduction
  - eigenvalues and eigenvectors
- In chapter 3 : Analysis basic concepts of analysis with one and multiple variable are presented. Some of the keywords are
  - Taylor approximation (order 1 and 2), tangent lines
  - multiple variables, partial derivatives, tangent plane

  - line integrals and elementary double integrals
  - vector fields, flow, divergence
  - integration by parts, Green's theorem, conservation laws.

- In chapter 4 : Ordinary Differential Equations some basic concepts on ordinary differential equations are presented. Some of the keywords are
  - ordinary differential equations (ODE) of order 1, visualization, separation of variables
  - ODEs of order 2, linear with constant coefficients, resonators
- In chapter 5 : Partial Differential Equations some basic information on solutions of partial differential equations (PDE) are presented.
  - Types of partial differential equations: elliptic, parabolic, hyperbolic.
  - Possible fields of application: steady state and dynamic heat equations, diffusion equations, wave equation.
  - Qualitative behavior of solutions to PDEs.

  This section will be a short presentation of basic facts.
- In chapter 6 : Fourier Methods basic concepts of Fourier methods are examined. Some of the keywords are
  - Fourier series, basics, spectrum, DFT, FFT
  - Fourier transform, introduction, written as a short presentation
- In chapter 7 : Probability Theory basic concepts of probability theory are examined. Some of the keywords are
  - independent events, Laplace experiment, conditional probability, Bayes' theorem, law of total probability, available
  - random variables, cumulative distribution function, quantile function, probability mass function, density function, expected value, variance.
  - joint distributions, marginal and conditional distributions, independent random variables, covariance, correlation.

Literature:

- Use your notes and/or books from the classes during your bachelor program.
- The book [ONei11] (or an other edition) covers most of the topics presented in these notes.
- In each of the chapters more literature is indicated.

# Chapter 2

# Linear Algebra

## 2.1 Keywords and Literature

Basic matrix operations, matrix notation for linear systems, algorithm of Gauss, augmented matrix, row reduction, homogeneous and inhomogenous systems, eigenvalue, eigenvector.

Literature:

- The book by Anton and Rorres ([AntoRorr10] or other editions) covers the necessary topics. It is a large book, well written on the correct level.
- The book by Koleman and Hill ([KolmHill01] or other editions) covers the necessary topics.
- The book [ONei11] (or a newer edition) has a section on matrix computations.
- Another option is [LandHest92].
- The books by Lothar Papula cover all required topics, but these books are in German.

## 2.2 Solving Systems of Linear Equations

• **2–1 Question:**
Use the matrices

$$\mathbf{A}=\begin{bmatrix} 1 & 2 & 3 \\ 0 & -3 & 4 \end{bmatrix} \quad , \quad \mathbf{B}=\begin{bmatrix} 10 & 3 \\ -1 & 3 \end{bmatrix}$$

and the vectors

$$\vec{x}=\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \quad \text{and} \quad \vec{y}=\begin{pmatrix} -2 \\ 0 \\ 3 \end{pmatrix}$$

to compute the following expressions:

(a) $\vec{a}=\mathbf{A}\,\vec{x}$

(b) $\vec{b}=\vec{x}^T\,\mathbf{A}^T$

(c) $\mathbf{C}=\mathbf{A}\cdot\mathbf{B}$

(d) The scalar product $s=\langle\vec{x}\,,\,\vec{y}\rangle$

(e) The two expressions $p_1=\langle\mathbf{B}\begin{pmatrix} 2 \\ 3 \end{pmatrix}\,,\,\begin{pmatrix} -1 \\ 5 \end{pmatrix}\rangle$ and $p_2=\langle\begin{pmatrix} 2 \\ 3 \end{pmatrix}\,,\,\mathbf{B}^T\begin{pmatrix} -1 \\ 5 \end{pmatrix}\rangle$.

$\diamondsuit$

• **2–2 Question:**
Solve the linear system below by applying the algorithm of Gauss (row reduction) using the notation of an augmented matrix.

$$\begin{array}{rrrcr} & -2\,y & +7\,z & = & 12 \\ 2\,x & -10\,y & +12\,z & = & 28 \\ 2\,x & -5\,y & -5\,z & = & -1 \end{array}$$

$\diamondsuit$

• **2–3 Question:**
Examine the system

$$\begin{array}{rrrcr} x_1 & +x_2 & +a\,x_3 & = & 0 \\ & 2\,x_2 & +4\,x_3 & = & 0 \\ 3\,x_1 & +2\,x_2 & +10\,x_3 & = & 0 \end{array}$$

For what value of the parameter $a$ does the system have multiple solutions? Then determine all solutions for this special value of $a$.

$\diamondsuit$

• **2–4 Question:**
Find all solutions of the linear system

$$\begin{array}{rrrcr} 13\,x & & -39\,z & = & 13 \\ & -2\,y & -4\,z & = & 6 \\ x & +y & -z & = & -2 \end{array}$$

$\diamondsuit$

• **2–5 Question:**
Applying **row operations** to matrices can be written are matrix multiplications from the left by a modification of the identity matrix $\mathbb{I}$. Examine a $3\times 3$ matrix $\mathbf{A}$ given by

$$\mathbf{A} = \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{bmatrix} \quad \text{and} \quad \mathbb{I} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Verify the following statements by setting up simple examples and then performing the operations.

(a) Verify that to multiply row $j$ of the matrix $\mathbf{A}$ by a factor $\alpha$ you can multiply the $j$-th entry on the diagonal of $\mathbb{I}$ by $\alpha$ and then multiply $\mathbf{A}$ from the left by this modified matrix.

(b) Add a multiple of one row to another row in $\mathbb{I}$. Then multiply the matrix $\mathbf{A}$ from the left by this modified matrix. Observe that the same row operation is applied to $\mathbf{A}$.

(c) Swap two rows in $\mathbb{I}$ and then multiply $\mathbf{A}$ from the left by this matrix. Observe that the same rows in $\mathbf{A}$ will be swapped.

The above results are not restricted to $3\times 3$ matrices, but apply to all sizes of matrices.

A similar result can be formulated for **column operations**, but one has to multiply $\mathbf{A}$ from the right by the modifications of $\mathbb{I}$.

$\diamondsuit$

• **2–6 Example:** Examine the matrix $\mathbf{A}$

$$\mathbf{A} = \begin{bmatrix} 1 & 0.5 \\ 0.25 & 0.75 \end{bmatrix}$$

representing a linear mapping from $\mathbb{R}^2$ to $\mathbb{R}^2$, i.e. for an arbitrary vector $\vec{x} \in \mathbb{R}^2$ the image is given by

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \mathbf{A}\,\vec{x} = \begin{bmatrix} 1 & 0.5 \\ 0.25 & 0.75 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

For the standard basis vectors observe

$$\mathbf{A}\,\vec{e}_1 = \begin{bmatrix} 1 & 0.5 \\ 0.25 & 0.75 \end{bmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 0.25 \end{pmatrix} \quad \text{and} \quad \mathbf{A}\,\vec{e}_2 = \begin{bmatrix} 1 & 0.5 \\ 0.25 & 0.75 \end{bmatrix} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 0.5 \\ 0.75 \end{pmatrix}.$$

The columns of $\mathbf{A}$ contain the images of the standard basis vectors. Figure 2.1 visualizes this linear mapping. ♢

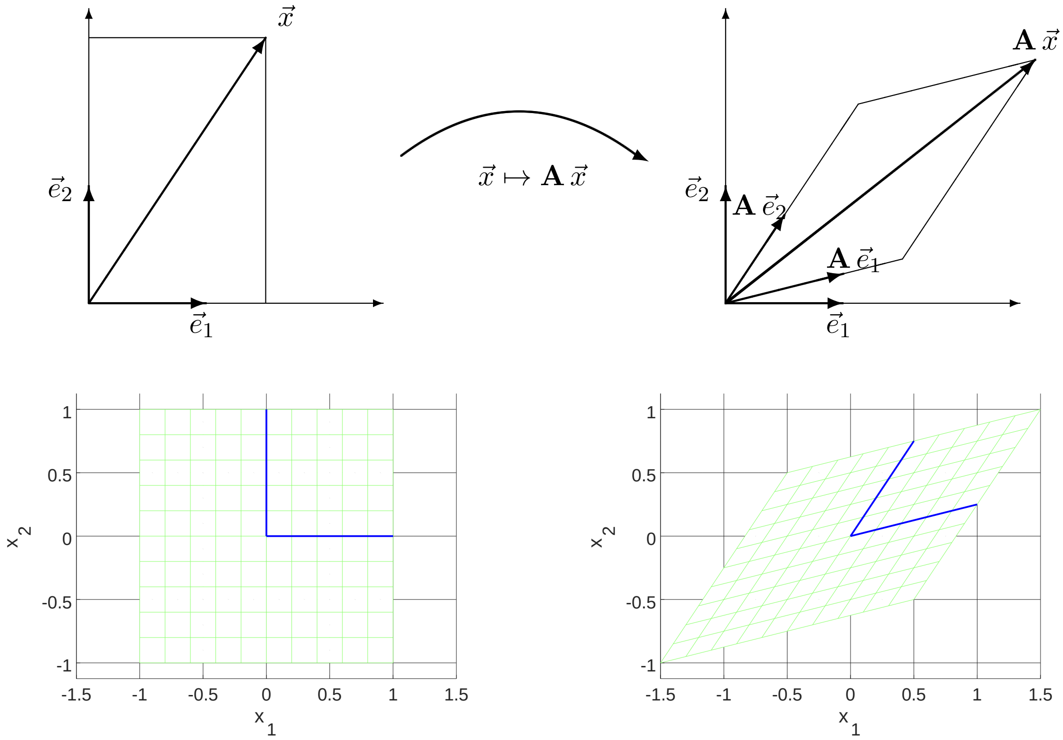



Figure 2.1: A linear mapping applied to a rectangle

• **2–7 Question:**
For a linear mapping $F : \mathbb{R}^2 \to \mathbb{R}^2$ we know that the vector $(1,3)^T$ is stretched by a factor of 3 and the image $(1,1)^T$ is given by $(\frac{-1}{2}, \frac{-1}{2})^T$.

(a) Find the matrix $\mathbf{A}$ for this linear mapping.

(b) Compute the image of the vector $(4,1)^T$.

♢

## 2.3 Eigenvalues and Eigenvectors

For a given square matrix $\mathbf{A} \in \mathbb{R}^{n\times n}$ the **eigenvalues** $\lambda$ and corresponding **eigenvectors** $\vec{v}$ are characterized by

$$\mathbf{A}\,\vec{v} = \lambda\,\vec{v}\,.$$

Thus the nonzero eigenvector $\vec{v}$ is a solution of the homogeneous system $(\mathbf{A} - \lambda\,\mathbb{I})\,\vec{v} = \vec{0}$ . As a consequence we have the key property $\det(\mathbf{A} - \lambda\,\mathbb{I}) = 0$, i.e. the eigenvalues $\lambda$ are solutions of the characteristic equation.

**• 2–8 Question:**
Examine the non-symmetric matrix

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \\ 5 & 4 \end{bmatrix}.$$

(a) Determine the two eigenvalues $\lambda_1$ and $\lambda_2$.

(b) Determine the eigenvector $\vec{v}_1$ for the first eigenvalue $\lambda_1$.

(c) Determine the eigenvector $\vec{v}_2$ for the second eigenvalue $\lambda_2$.

(d) Verify that the two eigenvectors are not orthogonal.

(e) Use `MATLAB` or *Octave* with the command `eig()` to verify the computations.

♢

**• 2–9 Question:**
Examine the symmetric matrix

$$\mathbf{A} = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & 0 \\ 0 & 0 & 7 \end{bmatrix}.$$

(a) Determine the three eigenvalues $\lambda_1$, $\lambda_2$ and $\lambda_3$.

(b) Determine the three eigenvectors $\vec{v}_i$ .

(c) Verify that the three eigenvectors are pairwise orthogonal.

(d) Divide the three eigenvectors by their own length, i.e. normalize them. Then construct the $3 \times 3$ matrix $\mathbf{Q}$ with the normalized eigenvectors as columns. Then verify that

$$\begin{aligned}
\mathbf{Q}^T \cdot \mathbf{Q} &= \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} = \mathbb{I} \\
\mathbf{A} \cdot \mathbf{Q} &= \mathbf{Q} \cdot \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} \\
\mathbf{Q}^T \mathbf{A} \cdot \mathbf{Q} &= \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}
\end{aligned}$$

(e) Use `MATLAB` or *Octave* with the command `eig()` to verify the computations.

$\diamondsuit$

**• 2–10 Question:**

Let $\mathbf{A} \in \mathbb{R}^{n\times n}$ be a real valued matrix with $n$ eigenvalues $\lambda_i$ and the corresponding eigenvectors $\vec{v}_i$. Then examine the homogeneous system of linear differential equations

$$\frac{d}{dt}\,\vec{x}(t) = \mathbf{A}\,\vec{x}(t)\,. \tag{2.1}$$

This type of differential equation often appears when solving dynamic heat equations numerically, see Section 5.3.

(a) Verify that any function $\vec{x}(t) = e^{\lambda_i\,t}\,\vec{v}_i$ solves (2.1).

(b) Verify that any linear combination of above solutions also solve the ODE (2.1), i.e. examine

$$\vec{x}(t) = \sum_{i=1}^{n} c_i\,e^{\lambda_i\,t}\,\vec{v}_i\,.$$

(c) If the above eigenvectors $\vec{v}_i$ are linearly independent, then any initial value $\vec{x}(0) = \vec{x}_0$ can be written as a linear combination of the eigenvectors, i.e.

$$\vec{x}(0) = \vec{x}_0 = \sum_{i=1}^{n} c_i\,\vec{v}_i\,.$$

Verify that solving for the coefficients $c_i$ leads to a system of linear equations. Using this we can solve (2.1) with an initial condition.

Using some more linear algebra one can show that the eigenvectors are linearly independent if the eigenvectors are all different, or if the matrix is symmetric. For nonsymmetric matrices it might happen that the eigenvectors are linearly dependent.

$\diamondsuit$

**• 2–11 Question:**

Let $\mathbf{A} \in \mathbb{R}^{n\times n}$ be a real valued matrix with positive eigenvalues $\lambda_i = \omega_i^2 > 0$ and the corresponding eigenvectors $\vec{v}_i$. Verify that the second order system of linear differential equations

$$\frac{d^2}{dt^2}\,\vec{x}(t) = -\mathbf{A}\,\vec{x}(t)$$

is solved by

$$\vec{x}(t) = \cos(\omega_i\,t)\,\vec{v}_i \quad \text{and} \quad \vec{x}(t) = \sin(\omega_i\,t)\,\vec{v}_i\,.$$

This type of differential equation often appears when solving wave equations numerically, see Section 5.4.

$\diamondsuit$

# Chapter 3

# Analysis

## 3.1 Keywords and Literature

Derivative, tangent, Taylor approximation of order one and two, optimization, partial derivatives, gradient, divergence, line integral, double integral, vector field, flow, divergence, Green's theorem, conservation law.

Literature:

- Any good book on basic calculus should do the job.
- The books by Earl W. Swokowski (e.g. [Swok92] or newer editions) are quite readable.
- The books by Lothar Papula cover all required topics, but these books are in German.

## 3.2 Taylor Approximation

• **3–1 Result:** If $I$ is an interval and $x_0\,,\,x_0+\Delta x\in I$ with a $(n+1)$ differentiable function $f\in\, C^{n+1}\,(I,\mathbb{R})$, then we have the Taylor approximation $T_n(h)$ of order $n$ given by

$$f\,(x_0+\Delta x)=T_n(h)+R_n \;=\; \sum_{k=0}^{n}\frac{1}{k!}\,f^{(k)}\,(x_0)\;(\Delta x)^k+R_n$$

where the remainder $R_n$ satifies

$$R_n=\frac{1}{(n+1)\,!}\,f^{(n+1)}(\xi)\;h^{n+1}\quad\text{for some } \xi \text{ between } x_0 \text{ and } x_0+\Delta x\quad .$$

$\diamondsuit$

Based on the above result one can approximate smooth functions by polynomials. Two approximations are used very frequently.

- Linear approximation, tangent line approximation

  $$f(x_0+\Delta x)\approx f(x_0)+f'(x_0)\,\Delta x$$

  with an error proportional to $(\Delta x)^2$ .

- Quadratic approximation, approximation by a parabola.

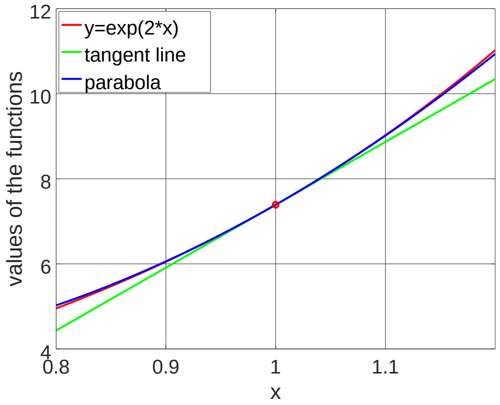
  $$f(x_0+\Delta x)\approx f(x_0)+f'(x_0)\,\Delta x+\frac{1}{2}\,f''(x_0)\,(\Delta x)^2$$

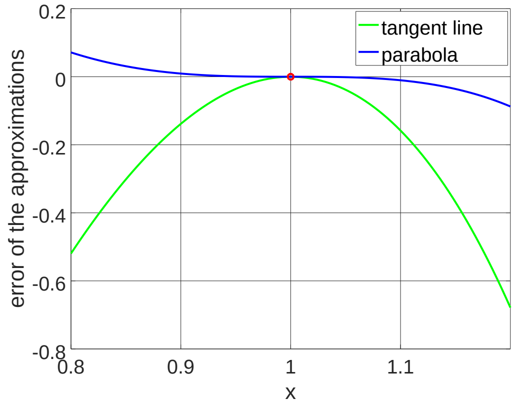
  with an error proportional to $(\Delta x)^3$ .

As an example examine both approximations to the function $f(x) = \exp(2 \cdot x)$ for values close to $x_0 = 1$.

$$\begin{aligned} f(1+\Delta x) &= e^{2+2\,\Delta x} &\approx& \; e^2 + 2\,e^2\,\Delta x \\ f(1+\Delta x) &= e^{2+2\,\Delta x} &\approx& \; e^2 + 2\,e^2\,\Delta x + \frac{1}{2}\,4\,e^2\,(\Delta x)^2 \end{aligned}$$

Below find the graphs of the functions and the resulting approximation errors.




• **3–2 Question:**
Determine $\cos(32^\circ) = \cos(30^\circ + 2^\circ) = \cos(\frac{\pi}{6} + \frac{2\,\pi}{180})$ with the help of $\cos 30^\circ = \cos\frac{\pi}{6} = \frac{\sqrt{3}}{2}$ and $\sin 30^\circ = \sin\frac{\pi}{6} = \frac{1}{2}$.

(a) Use a Taylor approximation of order 1.

(b) Use a Taylor approximation of order 2.

(c) For both of the above approximations determine an upper bound for the error $R_1$, resp.$R_2$.

(d) Generate a graph for the above two approximations for values of $|\Delta x| \leq 0.5$ (in radians).

Hint: It is advisable to express angles in radians instead of degree.

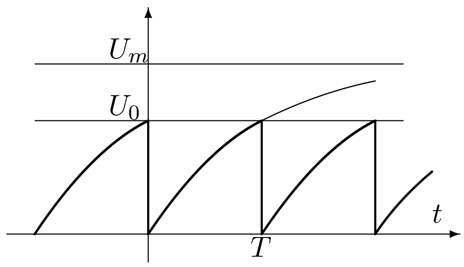
◊

• **3–3 Question:**

A capacitance is loaded through a resistor according to the formula

$$U(t) = U_m\,(1 - e^{-\alpha\,t})\,.$$

At time $T$ the voltage $U(T)$ reaches the critical value $U_0 < U_m$ and with the help of an active element the voltage is reset to 0 and the charging starts again. This leads to a periodic signal with period $T$.




The values of $U_m$ and $T$ are fixed. You may choose the values of $U_0$ and $\alpha$ .

(a) The period $T$ should not be sensitive to small variations of $U_0$, e.g. caused by temperature variations. Verify that this condition is satisfied if the slope of the voltage $U(t)$ at time $t = T$ is maximal

(b) Find the optimal values for $\alpha$ und $U_0$.

$\diamondsuit$

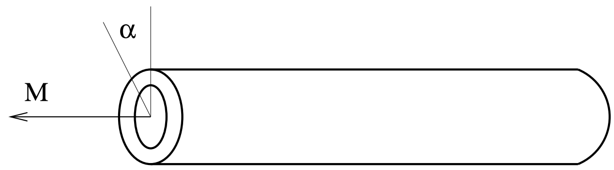
• **3–4 Question:**
A hollow tube of length $L$, inner radius $R_1$ and outer radius $R_2$ is twisted by an angle $\alpha$ by applying a moment $M$. Examine the angle $\alpha$ as function of the inner radius $R_1$.

$$\alpha = \frac{2\,(1+\nu)\;L}{E\;J}\;M \quad \text{where} \quad J = \frac{\pi}{2}\;\left(R_2^4 - R_1^4\right)$$

(a) For a small change of $\Delta R_1 \ll R_1$ determine the resulting change $\Delta\alpha$ for the angle with the help of a linear approximation.

(b) Express the relative change $\frac{\Delta\alpha}{\alpha}$ as function of $R_1$, $R_2$ and $\frac{\Delta R_1}{R_1}$.

(c) Use the values given below for an aluminum tube. Express $\alpha$ in radians and degree. By how much is the radius $R_1$ allowed to vary such that $|\Delta\alpha| \leq 1°$?




| symbol | value | unit |
|---|---|---|
| $M$ | 1 | N m |
| $L$ | 1 | m |
| $R_1$ | 3 | mm |
| $R_2$ | 5 | mm |
| $E$ | $7 \cdot 10^{10}$ | $\frac{\text{N}}{\text{m}^2}$ |
| $\nu$ | 0.34 | |

$\diamondsuit$

## 3.3 Analysis for Functions of Multiple Variables

• **3–5 Definition:**

- For a function $f\,:\,\mathbb{R}^n \longrightarrow \mathbb{R}$ the **gradient** is a function $\nabla f\,:\,\mathbb{R}^n \longrightarrow \mathbb{R}^n$ determined by the partial derivatives[1]

$$\nabla f(\vec{x}) = \operatorname{grad} f(\vec{x}) = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \ldots, \frac{\partial f}{\partial x_n}\right) .$$

- For a function $\vec{u}\,:\,\mathbb{R}^n \longrightarrow \mathbb{R}^n$ the **divergence** is a function $\nabla\vec{u}\,:\,\mathbb{R}^n \longrightarrow \mathbb{R}$ determined by a sum of partial derivatives

$$\nabla\vec{u} = \operatorname{div}\vec{u} = \sum_{i=1}^{n} \frac{\partial u_i}{\partial x_i} = \frac{\partial u_1}{\partial x_1} + \frac{\partial u_2}{\partial x_2} + \ldots + \frac{\partial u_n}{\partial x_n} .$$

[1]Some books use the notation $\vec{\nabla}$ for the gradient, instead of only the nabla symbol $\nabla$. This should not cause confusion, since from the context it is clear whether the divergence $\operatorname{div}\vec{u} = \nabla\vec{u}$ (the argument $\vec{u}$ is a vector) or the gradient $\operatorname{grad} f = \vec{\nabla} f = \nabla f$ (the argument $f$ is a scalar) has to be examined.

- A function $\vec{u} : \mathbb{R} \longrightarrow R^n$ determines a curve in $\mathbb{R}^n$ and the corresponding velocity vector is given by

$$\dot{\vec{u}}(t) = \frac{d}{dt}\,\vec{u}(t) = \begin{pmatrix} \dot{u}_1(t) \\ \dot{u}_2(t) \\ \vdots \\ \dot{u}_n(t) \end{pmatrix} = \begin{pmatrix} \frac{d}{dt}u_1(t) \\ \frac{d}{dt}u_2(t) \\ \vdots \\ \frac{d}{dt}u_n(t) \end{pmatrix}.$$

It is a tangential vector for the curve described by the parametrization $\vec{u}(t)$.

$\diamondsuit$

• **3–6 Result:** For a function $f(\vec{x})$ and a vector $\vec{v} \in \mathbb{R}^n$ the **directional derivative of $f(\vec{x})$ at $\vec{x}_0$ in the direction $\vec{v}$** is defined by

$$\frac{\partial\, f(\vec{x}_0)}{\partial \vec{v}} = D_{\vec{v}} f(\vec{x}_0) := \lim_{h\to 0} \frac{f(\vec{x}_0 + h\,\vec{v}) - f(\vec{x}_0)}{h}$$

and can be computed by

$$D_{\vec{v}} f(\vec{x}_0) = \operatorname{grad} f(\vec{x}_0)\cdot\vec{v} = \left(\frac{\partial\, f}{\partial x_1}, \frac{\partial\, f}{\partial x_2}, \ldots, \frac{\partial\, f}{\partial x_n}\right)\cdot \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix} = \sum_{i=1}^{n} \frac{\partial\, f(\vec{x}_0)}{\partial x_i}\, v_i$$

or by using the notation of the scalar product

$$\frac{\partial\, f(\vec{x}_0)}{\partial \vec{v}} = D_{\vec{v}} f(\vec{x}_0) = \langle \operatorname{grad} f(\vec{x}_0)\,,\,\vec{v}\rangle = \langle \nabla f(\vec{x}_0)\,,\,\vec{v}\rangle\,.$$

Its value represents the slope of the surface given by $f(\vec{x})$ in the direction $\vec{v}$ at the point $\vec{x}_0$. $\diamondsuit$

• **3–7 Result:** For functions $\vec{u} : \mathbb{R} \longrightarrow \mathbb{R}^n$ and $f : \mathbb{R}^n \longrightarrow \mathbb{R}$ the **composition** $h = f \circ \vec{u}$, i.e. $h(t) = f(\vec{u}(t))$, is a mapping $h : \mathbb{R} \longrightarrow \mathbb{R}$ and its derivative is given by the **chain rule**

$$\begin{aligned} \frac{d}{dt}\,h(t) &= \frac{d}{dt}\,f(\vec{u}(t)) = \operatorname{grad} f(\vec{u}(t))\cdot\frac{d\,\vec{u}(t)}{dt} \\ &= \left(\frac{\partial\, f}{\partial u_1}, \frac{\partial\, f}{\partial u_2}, \ldots, \frac{\partial\, f}{\partial u_n}\right)\cdot \begin{pmatrix} \dot{u}_1(t) \\ \dot{u}_2(t) \\ \vdots \\ \dot{u}_n(t) \end{pmatrix} = \sum_{i=1}^{n} \frac{\partial\, f(\vec{u}(t))}{\partial u_i}\,\dot{u}_i(t)\,. \end{aligned}$$

$\diamondsuit$

• **3–8 Question:**
Use the functions

$$\begin{aligned} f(x,y) &= x^2 + \exp(y-x) \\ \vec{x}(t) &= \begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} \cos(t) \\ t^3 \end{pmatrix} \end{aligned}$$

and then determine the following expressions:

$$A = \frac{d\,f}{dx} \quad , \quad B = \operatorname{grad} f(x,y) \quad , \quad C = \frac{d}{dt}\,\vec{x}(t) \quad , \quad D = \frac{d}{dt}\,f(x(t),y(t))\,.$$

$\diamondsuit$

• **3–9 Result:** For a function $f : \mathbb{R} \longrightarrow \mathbb{R}^n$ we use a directional vector $\vec{v}$ with length $\|\vec{v}\| = 1$ and at a point $\vec{x}_0 \in \mathbb{R}^2$ let $\alpha$ be the angle between $\vec{v}$ and the gradient $\nabla f(\vec{x}_0)$. Now examine the curve

$$\vec{u}(t) = \vec{x}_0 + t\,\vec{v} \quad \text{with} \quad \vec{u}(0) = \vec{x}_0 \quad \text{and} \quad \frac{d}{dt}\,\vec{u}(0) = \vec{v}$$

and the composition $h(t) = f(\vec{u}(t))$. Then the chain rule implies

$$\begin{aligned} \frac{d}{dt}\,h(0) &= \operatorname{grad} f(\vec{u}(0)) \cdot \frac{d}{dt}\,u(0) = \operatorname{grad} f(\vec{x}_0) \cdot \vec{v} \\ &= \|\operatorname{grad} f(\vec{x}_0)\|\;\|\vec{v}\|\;\cos\alpha \;=\; \|\operatorname{grad} f(\vec{x}_0)\|\;\cos\alpha \end{aligned}$$

Since $\frac{d}{dt}\,h(0)$ is the slope of the surface given by $f(\vec{x})$ at the point $\vec{x}_0$ in the direction $\vec{v}$ this leads to a geometric interpretation of the gradient.

- The gradient (a vector) is pointing in the **direction of steepest increase** of the surface
- and the length of the gradient is the **slope in that direction**.

♢

• **3–10 Result:** If $\Omega \in \mathbb{R}^2$ is a domain and $(x_0 + \Delta x\,,\, y_0 + \Delta y) \in \Omega$ for $|\Delta x|$ and $|\Delta y|$ small enough, then we have two frequently used approximations: linear and quadratic. There are a few different notations used frequently to state these results.

- Linear approximation, Taylor approximation of order 1, tangent plane approximation.

$$\begin{aligned} f(x_0 + \Delta x\,,\, y_0 + \Delta x) &\approx f(x_0, y_0) + \operatorname{grad} f(x_0, y_0) \cdot \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} \\ &= f(x_0, y_0) + \left(\frac{\partial\, f(x_0, y_0)}{\partial x}\,,\, \frac{\partial\, f(x_0, y_0)}{\partial y}\right) \cdot \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} \\ &= f(x_0, y_0) + \frac{\partial\, f(x_0, y_0)}{\partial x}\,\Delta x + \frac{\partial\, f(x_0, y_0)}{\partial y}\,\Delta y \\ &= f(x_0, y_0) + f_x(x_0, y_0)\,\Delta x + f_y(x_0, y_0)\,\Delta y \end{aligned}$$

  The function $f$ has to be twice differentiable and the approximation error is proportional to $|\Delta x|^2 + |\Delta y|^2$, i.e. of second order.

- Quadratic approximation, Taylor approximation of order 2.

$$\begin{aligned} f(x, y) &= f(x_0 + \Delta x\,,\, y_0 + \Delta x) \\ &\approx f(x_0, y_0) + \operatorname{grad} f(x_0, y_0) \cdot \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} + \frac{1}{2}\,(\Delta x\,,\, \Delta y) \cdot \mathbf{H} \cdot \begin{pmatrix} \Delta x \\ \Delta y \end{pmatrix} \\ &= f(x_0, y_0) + f_x(x_0, y_0)\,\Delta x + f_y(x_0, y_0)\,\Delta y + \\ &\quad + \frac{1}{2}\,\left(f_{xx}(x_0, y_0)\,(\Delta x)^2 + 2\,f_{xy}(x_0, y_0)\,(\Delta x)(\Delta y) + f_{yy}(x_0, y_0)\,(\Delta y)^2\right) \end{aligned}$$

  The **Hesse** matrix $\mathbf{H}$ is given by the partial derivatives of order two.

$$\mathbf{H} = \begin{bmatrix} \frac{\partial^2 f(x_0,y_0)}{\partial x^2} & \frac{\partial^2 f(x_0,y_0)}{\partial x\,\partial y} \\ \frac{\partial^2 f(x_0,y_0)}{\partial x\,\partial y} & \frac{\partial^2 f(x_0,y_0)}{\partial y^2} \end{bmatrix} = \begin{bmatrix} f_{xx}(x_0, y_0) & f_{xy}(x_0, y_0) \\ f_{xy}(x_0, y_0) & f_{yy}(x_0, y_0) \end{bmatrix}$$

  The function $f$ has to be three times differentiable and the approximation error is of third order.

♢

• **3–11 Question:**
Examine the function
$$f(x,y) = 2 - (x+1)^2 - 2\,(y-2)^2\,.$$

(a) Determine the gradient $\operatorname{grad} f(x,y)$

(b) Determine the direction of steepest increase at the origin $(0,0)$.

(c) Determine the equation of the tangent plane at the point $(0\,,\,0\,,\,f(0,0))$.

(d) Determine the Taylor approximation of order 2 at the origin.

♢

• **3–12 Question:**
By measuring the current $I = (15 \pm 0.3)\,A$ and the voltage $U = (110 \pm 2)\,V$ one can use Ohm's law to determine the resistance $R$ and its maximal relative error. Use a linear approximation to estimate the relative error.

♢

• **3–13 Result:** For a smooth function $f\,:\,\mathbb{R}^n \longrightarrow \mathbb{R}$ we have the **necessary condition for an extremal point** (maximum or minimum) that the tangent plane has to be horizontal.

$$f(\vec{x}) \text{ extremal at } \vec{x} = \vec{x}_0 \quad \Longrightarrow \quad \operatorname{grad} f(\vec{x}_0) = \vec{0} \quad \Longleftrightarrow \quad \frac{\partial\, f(\vec{x}_0)}{\partial x_i} = 0 \text{ for } i = 1,2,\ldots,n$$

♢

• **3–14 Question:**
Determine the unique minimum of the function
$$f(x,y) = e^{-y^2+x}\,x\,.$$

♢

## 3.4 Line Integrals and Elementary Double Integrals

The may task in this section is to **set up the integrals** to be computed. For the computation of the resulting definite integrals you may use a tool, e.g. `MATLAB` or *Octave*.

### 3.4.1 Line Integrals

In this section we examine curves $C \in \mathbb{R}^2$ given by a parametrization, i.e.

$$C \quad : \quad \vec{x}\,(t) = \begin{pmatrix} x\,(t) \\ y\,(t) \end{pmatrix} \quad \text{where} \quad a \leq t \leq b\,.$$

All definitions and ideas can easily be extended to curves in space $\mathbb{R}^3$.

There are three different types of line integrals to be examined:

- If $f$ is a scalar function and $C$ a curve, then the integral

  $$\int_C f = \int_C f\; ds$$

  may be examined. As a typical example $f$ might give you the mass per unit length of a cable with variable cross section. Then the above integral will compute the total mass.

- With the notation

  $$\vec{A} = \int_C \vec{F}\; ds = \int_C \begin{pmatrix} F_1 \\ F_2 \end{pmatrix}\; ds$$

  multiple scalar integrals are given in one formula. The notation actually stands for the two scalar integrals

  $$A_1 = \int_C F_1\; ds \quad \text{and} \quad A_2 = \int_C F_2\; ds\,.$$

- If $\vec{F}$ is a vector field and $C$ a curve, then the integral

  $$\int_C \vec{F} \cdot \vec{ds}$$

  may be examined. As typical example the vector $\vec{F}$ might give you the force (strength and direction). Then the above integral will determine the total work required to move a point along the given curve $C$. Sometimes a different notation is used, i.e.

  $$\int_C \vec{F} \cdot \vec{ds} = \int_C F_1\, dx + F_2\, dy\,.$$

• **3–15 Definition:** Examine a curve given by a parametrization at a specific time $t$, and thus position $\vec{x}(t) = (x(t)\,,\, y(t))^T$. For a $0 < \Delta t$ very small use Figure 3.1 as a motivation, leading to

$$\begin{aligned}
\vec{x}(t + \Delta t) &= \begin{pmatrix} x(t+\Delta t) \\ y(t+\Delta t) \end{pmatrix} \approx \begin{pmatrix} x(t) + \dot{x}(t)\Delta t \\ y(t) + \dot{y}(t)\Delta t \end{pmatrix} = \vec{x}(t) + \dot{\vec{x}}(t)\Delta t \\
\Delta\vec{x}(t) &= \vec{x}(t+\Delta t) - \vec{x}(t) = \begin{pmatrix} x(t+\Delta t) - x(t) \\ y(t+\Delta) - y(t) \end{pmatrix} = \begin{pmatrix} \Delta x(t) \\ \Delta y(t) \end{pmatrix} \approx \dot{\vec{x}}(t)\;\Delta t\,.
\end{aligned}$$

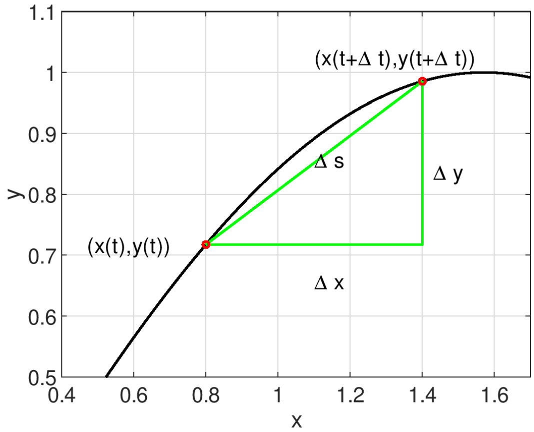



Figure 3.1: A simple curve

This observation leads to the definition of the line elements $ds$ and $\vec{ds}$.

$$\begin{aligned}\vec{ds} &= \dot{\vec{x}}\, dt \\ ds &= \|\vec{ds}\| = \|\dot{\vec{x}}\|\, dt\end{aligned}$$

and the definition of the scalar curve integral

$$\int_C f = \int_C f\, ds = \int_0^b f(\vec{x}(t))\, \|\dot{\vec{x}}(t)\|\, dt\,.$$

One can show that this definition does not depend on the parametrization of the curve $C$, but only on the curve. ♢

• **3–16 Example:** The length $L$ of a curve parametrized by $(x(t)\,,\, y(t))^T$ for $a \leq t \leq b$ is given by

$$L = \int_C 1\, ds = \int_a^b \|(\dot{x}(t)\,,\,\dot{y}(t))^T\|\, dt = \int_a^b \sqrt{(\dot{x}(t))^2 + (\dot{y}(t))^2}\, dt\,.$$

As an example consider a circle of radius $R$, parametrized by

$$\vec{x}(t) = \begin{pmatrix} R\,\cos(t) \\ R\,\sin(t)\end{pmatrix} \quad \text{for} \quad 0 \leq t \leq 2\,\pi\,.$$

Then determine

$$\begin{aligned}\vec{ds} &= \begin{pmatrix}\dot{x}(t) \\ \dot{y}(t)\end{pmatrix} dt = \begin{pmatrix} -R\,\sin(t) \\ +R\,\cos(t)\end{pmatrix} dt \\ ds &= \|\vec{ds}\| = \sqrt{(-R\,\sin(t))^2 + (+R\,\cos(t))^2}\, dt = R\, dt \\ L &= \int_C 1\, ds = \int_0^{2\pi} 1\, R\, dt = 2\,\pi\, R\,.\end{aligned}$$

♢

**• 3–17 Question:**
The shape of a jump rope is given by the function

$$h\,(x) = 3\,(0.5 - x^2) \quad \text{where} \quad \frac{-1}{\sqrt{2}} \leq x \leq \frac{+1}{\sqrt{2}}$$

and has a total mass of 1 kg. The rope is rotating about the $x$-axis with one full rotation in one second.

(a) Determine the total length $L$ of this rope. Then it is easy to find the specific mass $\rho$ in kg/m .

(b) Determine the total kinetic energy of this rotating jump rope. Reminder: the kinetic energy of a mass $m$ moving with speed $v$ is given by $\frac{1}{2}\,m\,v^2$ .

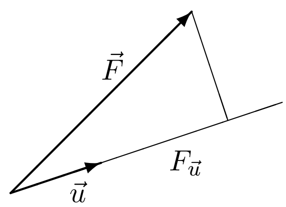
♢

Examine a particle changing its position from a point $\vec{x}$ to a point $\vec{x} + \vec{u}$ and it is exposed to a force $\vec{F}$ with given strength $\|\vec{F}\|$ and direction given by the vector. We want to determine the work performed by this force $\vec{F}$. Only the component $F_{\vec{u}}$ of $\vec{F}$ in the direction of the displacement $\vec{u}$ is relevant, i.e. we have to multiply the strength $\|\vec{F}\|$ by $\cos\alpha$, where $\alpha$ is the angle between $\vec{F}$ and $\vec{u}$.

$$F_{\vec{u}} = \|\vec{F}\|\,\cos\alpha\,.$$

The total work $A$ is then given by the product of this force in the direction of $\vec{u}$ and the traveled distance $\|\vec{u}\|$, i.e. the scalar product of the two vectors $\vec{F}$ and $\vec{u}$.

$$\begin{aligned} A &= F_{\vec{u}}\,\|\vec{u}\| \;=\; \|\vec{F}\|\,\cos\alpha\,\|\vec{u}\| \;=\; \|\vec{F}\|\,\frac{\langle\vec{F},\vec{u}\rangle}{\|\vec{F}\|\,\|\vec{u}\|}\,\|\vec{u}\| \\ &= \langle\vec{F},\vec{u}\rangle \;=\; \vec{F}\cdot\vec{u} \end{aligned}$$




The above serves as a motivation and illustration for the definition of the line integral.

**• 3–18 Definition:** Let $C \in \mathbb{R}^2$ be a curve parametrized by $\vec{x}(t)$ for $a \leq t \leq b$. Then the integral of the vector field $\vec{F}$ along the curve $C$ is defined by

$$\begin{aligned} \int_C \vec{F}\cdot\vec{ds} &= \int_C \vec{F}\,\vec{ds} = \int_C F_1\,dx + F_2\,dy \\ &= \int_a^b \langle\vec{F}(\vec{x}(t))\,,\,\dot{\vec{x}}(t)\rangle\;dt \;=\; \int_a^b \left(F_1(\vec{x}(t))\,\dot{x}(t) + F_2(\vec{x}(t))\,\dot{y}(t)\right)\;dt\,. \end{aligned}$$

♢

**• 3–19 Example:** Let $C$ be the semi circle parametrized by $(R\,\cos\alpha\,,\,R\,\sin\alpha)^T$ for $0 \leq \alpha \leq \pi$ and thus

$$\vec{ds} = \begin{pmatrix} -R\,\sin\alpha \\ +R\,\cos\alpha \end{pmatrix}\,d\alpha\,.$$

Now examine a force given by

$$\vec{F} = \vec{F}(\vec{x}) = \begin{pmatrix} x \\ y \end{pmatrix}$$

and compute the work $A$ performed by this force field along the given semi circle.

$$\begin{aligned} A &= \int_C \vec{F}\cdot\vec{ds} = \int_C x\,dx + y\,dy \\ &= \int_0^\pi \begin{pmatrix} R\,\cos\alpha \\ R\,\sin\alpha \end{pmatrix} \cdot \begin{pmatrix} -R\,\sin\alpha \\ +R\,\cos\alpha \end{pmatrix} d\alpha \\ &= \int_0^\pi -R^2\,\cos\alpha\,\sin\alpha + R^2\,\cos\alpha\,\sin\alpha\,d\alpha = \int_0^\pi 0\,d\alpha = 0 \end{aligned}$$

A graphics will convince you that the force $\vec{F}$ is always orthogonal to the direction in which the point is moving, thus not work generated. ♢

**• 3–20 Question:**
Compute the curve integrals of the vector field

$$\vec{F} = \begin{pmatrix} x^2 - y \\ y^2 + x \end{pmatrix}$$

along different curves.

(a) Along the straight line connecting $(0,1)$ to $(1,2)$.

(b) First along the straight line segment from $(0,1)$ to $(1,1)$, and then along the straight line from $(1,1)$ to $(1,2)$.

(c) Along the parabola given by $x = t$, $y = 1 + t^2$ from $(0,1)$ to $(1,2)$.

♢

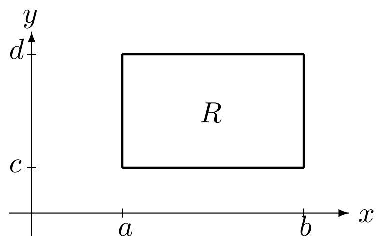
### 3.4.2 Double Integrals

The main point in this section is to **set up the integrals**, and not on how to evaluate the resulting integrals.

**• 3–21 Definition:** For a rectangle $R$ ($a \le x \le b$ and $c \le y \le d$) and a continuous function $f(x,y)$ the double integral over the rectangle $R$ is given by

♢

$$\begin{aligned} \iint\limits_R f(x,y)\,dA &= \int_a^b \left(\int_c^d f(x,y)\,dy\right)\,dx \\ &= \int_c^d \left(\int_a^b f(x,y)\,dx\right)\,dy\,. \end{aligned}$$

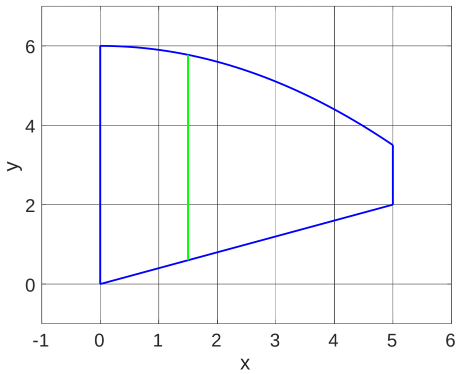



The above definition can be extended to non rectangular domains.

• **3–22 Definition:** If the domain $D \subset \mathbb{R}^2$ is given by $c(x) \leq y \leq d(x)$ for $a \leq x \leq b$ the double integral is determined by

$$\iint\limits_D f\, dA = \iint\limits_D f(x,y)\, dA = \int_a^b \left( \int_{c(x)}^{d(x)} f(x,y)\, dy \right) dx\,.$$

The illustration going along with the above computation is slicing the domain $D \subset \mathbb{R}^2$ into slices of width $\Delta x$ in $y$-direction and then move the slices from $x = a$ to $x = b$.
On the right find the domain of integration given by $\frac{2}{5}\,x \leq y \leq 6 - \frac{1}{10}\,x^2$ for $0 \leq x \leq 5$. The domain is enclosed by the blue curve and one slice in $y$ direction at $x = 1.5$ is given by the green line.




If the domain $D \subset \mathbb{R}^2$ is given by $a(y) \leq x \leq b(y)$ for $c \leq y \leq d$ the the double integral is determined by

$$\iint\limits_D f\, dA = \iint\limits_D f(x,y)\, dA = \int_c^d \left( \int_{a(y)}^{b(y)} f(x,y)\, dx \right) dy\,.$$

The illustration going along with the above computation is slicing the domain $D \subset \mathbb{R}^2$ into slices of height $\Delta y$ in $x$-direction and then move the slices from $y = c$ to $y = d$. $\diamondsuit$

• **3–23 Question:**
Determine the volume between the $xy$ plane and then surface given by $z = 1 + x + y$ for $0 \leq x \leq 2$ and $0 \leq y \leq 3$.

$\diamondsuit$

• **3–24 Question:**
A rectangular plate of thickness $H$ and dimensions $0 \leq x \leq W$ and $0 \leq y \leq D$ is rotated about the $z$-axis with angular velocity $\omega$. Detemine the kinetic energy $E$ if the density is given by $\rho$.

$\diamondsuit$

• **3–25 Question:**
A horizontal circular plate with radius $R$ of thickness $H$ and center at the origin is rotated about the $z$-axis with angular velocity $\omega$. Determine the kinetic energy $E$ if the density is given by $\rho$.

$\diamondsuit$

When integrating over domains in $\mathbb{R}^2$ we used the area element $dA = dx \cdot dy$, which was generated by small rectangles of width $\Delta x$ and height $\Delta y$. The corresponding area is $\Delta A = \Delta x \cdot \Delta y$. Many problems are easier to solve using polar coordinates

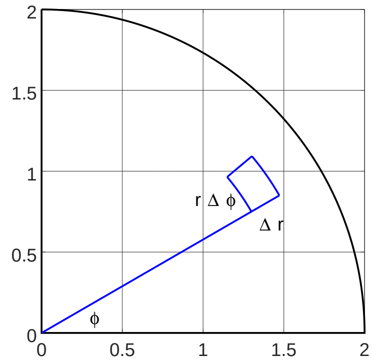
Polar coordinates $r$ and $\phi$ are determined by

$$\begin{aligned} x &= r\cos\phi \\ y &= r\sin\phi \end{aligned}$$

and the corresponding area element is

$$\begin{aligned} \Delta A &\approx r \cdot \Delta\phi \cdot \Delta r \\ dA &= r\,d\phi\,dr\,. \end{aligned}$$

The figure on the right illustrates this identity.



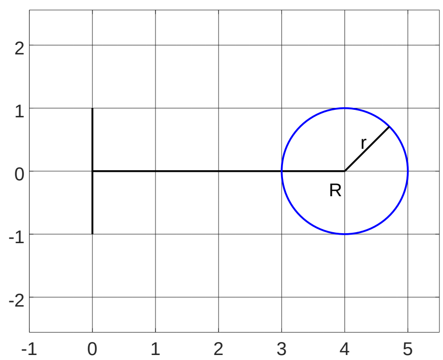

• **3–26 Question:**
A horizontal circular plate with radius $R$ of thickness $H$ and center at the origin is rotated about the $z$-axis with angular velocity $\omega$. Determine the kinetic energy $E$ if the density is given by $\rho$.

♢

• **3–27 Question:**
The disk in the $xz$ plane with center at $(x,z) = (R,0)$ and radius $r$ $(0 < r < R)$ is rotated about the $z$-axis to generate a torus in $\mathbb{R}^3$.

(a) Setup the integral to determine the volume of the generated torus. Slice the circle in narrow bands and rotate those about the $z$ axis.

(b) Use polar coordinates to slice up the circle and setup a double integral.




♢

• **3–28 Question:**
The above torus with total mass $M$ is rotated about the $z$-axis with angular velocity $\omega$. Give the integral to compute the kinetic energy.

♢

### 3.4.3 Volume and Surface Integrals in $\mathbb{R}^3$

• **3–29 Result:** For integrals over domains $D \subset \mathbb{R}^2$ mainly two sets of coordinates are used: Cartesian and polar (cylindrical) coordinates. They are connected by the usual $x = \rho\cos\phi$ and $y = \rho\sin\phi$. The area $dA$ element given by

$$dA = dx \cdot dy = \rho\,d\phi\,d\rho\,.$$

For integrals over domains $D \subset \mathbb{R}^3$ there are three sets of standard coordinates used: Cartesian, cylindrical and spherical coordinates.

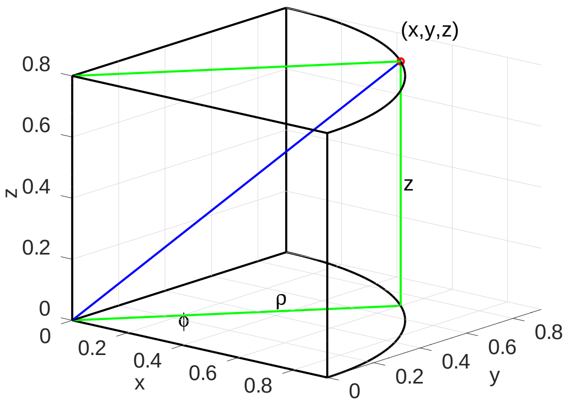
- The Cartesian coordinates are usually denoted by $(x, y, z)$ and the volume element is $dV = dx\,dy\,dz$.
- Cylindrical coordinates: use the angle $\phi$ in the $xy$ plane, the radius $\rho$ in this plane and the height $z$ to describe the point in space.

The cylindrical coordinates are given by

$$\begin{aligned} x &= \rho \cos\phi \\ y &= \rho \sin\phi \\ z &= z \end{aligned}$$

and the volume element is

$$dV = \rho\, d\phi\, d\rho\, dz\,.$$



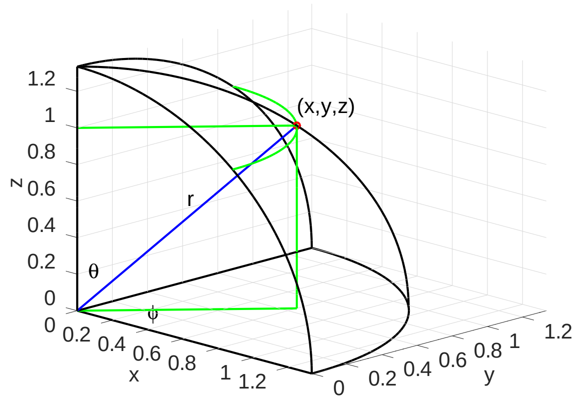

- Spherical coordinates: use the angle $\phi$ in the $xy$ plane, the angle $\theta$ off the $z$ axis and the distance $r$ of the point from the origin to describe the point in space.

The coordinates are given by

$$\begin{aligned} x &= r\,\sin\theta\,\cos\phi \\ y &= r\,\sin\theta\,\sin\phi \\ z &= r\,\cos\theta \end{aligned}$$

where $0 \leq \phi < 2\pi$ and $0 \leq \theta \leq \pi$. The volume element is

$$dV = r^2 \sin\theta\, d\theta\, d\phi\, dr\,.$$




♢

| | Variables | $dV$ |
|---|---|---|
| Cartesian coordinates | $x,\ y,\ z$ | $dx\ dy\ dz$ |
| Cylindrical coordinates | $\rho,\ \phi,\ z$ | $\rho\ d\rho\ d\phi\ dz$ |
| Spherical coordinates | $r,\ \phi,\ \theta$ | $r^2\ \sin\theta\ dr\ d\phi\ d\theta$ |

Table 3.1: Volume elements in $\mathbb{R}^3$

To work with surface integrals in $\mathbb{R}^3$ one needs the normal vectors to the surfaces and the corresponding surface elements. Below find the results for the three most often used sets of coordinates. It is essential to visualize the results with some figures, not shown here.

### Surface Integrals in Cartesian Coordinates

For the Cartesian coordinates $(x, y, z)$ the three directional vectors are

$$\vec{e}_x = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad , \quad \vec{e}_y = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad \text{and} \quad \vec{e}_z = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} .$$

In Table 3.2 find the surface elements, if one of the coordinates is fixed.

### Surface Integrals in Cylindrical Coordinates

For cylindrical coordinates $(\rho, \phi, z)$ the three directional vectors are given by

$$\vec{e}_\rho = \begin{pmatrix} \cos\phi \\ \sin\phi \\ 0 \end{pmatrix} \quad , \quad \vec{e}_\phi = \begin{pmatrix} -\sin\phi \\ \cos\phi \\ 0 \end{pmatrix} \quad \text{and} \quad \vec{e}_z = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} .$$

These vectors are pairwise orthogonal and have length 1. In Table 3.2 find the typical surface elements for cylindrical coordinates, if one of the coordinates is fixed.

### Surface Integrals in Spherical Coordinates

For spherical coordinates $(r, \phi, \theta)$ the three directional vectors are given by

$$\vec{e}_r = \begin{pmatrix} \sin\theta\ \cos\phi \\ \sin\theta\ \sin\phi \\ \cos\theta \end{pmatrix} \quad , \quad \vec{e}_\phi = \begin{pmatrix} -\sin\phi \\ \cos\phi \\ 0 \end{pmatrix} \quad \text{and} \quad \vec{e}_\theta = \begin{pmatrix} \cos\theta\ \cos\phi \\ \cos\theta\ \sin\phi \\ -\sin\theta \end{pmatrix} .$$

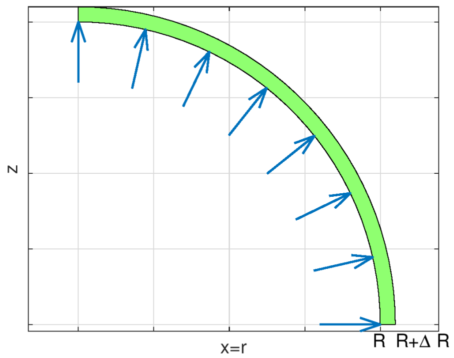
These vectors are pairwise orthogonal and have length 1. In Table 3.2 find the typical surface elements for spherical coordinates, if one of the coordinates is fixed.

| Cartesian coordinates | $x,\ y$ | $x,\ z$ | $y,\ z$ |
|---|---|---|---|
| Surface element $dA$ | $dx\ dy$ | $dx\ dz$ | $dy\ dz$ |
| Normal vector | $\pm\vec{e}_z$ | $\pm\vec{e}_y$ | $\pm\vec{e}_x$ |
| Cylindrical coordinates | $\rho,\ \phi$ | $\phi,\ z$ | $\rho,\ z$ |
| Surface element $dA$ | $\rho\ d\rho\ d\phi$ | $\rho\ d\phi\ dz$ | $d\rho\ dz$ |
| Normal vector | $\pm\vec{e}_z$ | $\pm\vec{e}_\rho$ | $\pm\vec{e}_\phi$ |
| Spherical coodinates | $r,\ \phi$ | $\phi,\ \theta$ | $r,\ \theta$ |
| Surface element $dA$ | $r\ \sin\theta\ dr\ d\phi$ | $r^2\ \sin\theta\ d\phi\ d\theta$ | $r\ dr\ d\theta$ |
| Normal vector | $\pm\vec{e}_\theta$ | $\pm\vec{e}_r$ | $\pm\vec{e}_\phi$ |

Table 3.2: Surface elements in $\mathbb{R}^3$

- **3–30 Question:**

In the figure on the right find a section through half of a thin sphere. To obtain the sphere rotate about the $y$-axis. On the inside a constant pressure $p$ is applied. Use spherical coordinates to examine this situation.




(a) Verify that the vertical pressure component at an angle $\theta$ off the $z$-axis is given by $p\,\cos\theta$ and the horizontal components are given by $p\,\sin\theta\,(\cos\phi\,,\,\sin\phi)$.

(b) Use an integral of the vertical force density (pressure) to determine the total force vertical $F_z$ acting on the semi sphere.

(c) Then determine the area of the cross section at the base $z = 0$ of the thin walled sphere and determine the resulting vertical stress in the wall. Use $\Delta R \ll R$.

(d) Use a similar integral to verify that the total forces in the horizontal directions vanish.

♢

## 3.5 Vector Fields and Flow Computations

In this section we examine vector field in the plane $\mathbb{R}^2$, but it should be straightforward to adapt the notation to the space $\mathbb{R}^3$.

- **3–31 Definition:** A vector field on a domain $D \subset \mathbb{R}^2$ is a function $\vec{F} : D \to \mathbb{R}^2$

$$\vec{F}(\vec{x}) = \vec{F}(x,y) = \begin{pmatrix} F_1(x,y) \\ F_2(x,y) \end{pmatrix} .$$

To visualize a vector field draw the vectors $\vec{F}(x,y)$ attached at the points $(x,y)$. ♢

- **3–32 Example:** The function

$$\vec{F}(x,y) = \begin{pmatrix} 1-x \\ \sin(y) \end{pmatrix}$$

defines a vector field. Using *Octave*/`MATLAB` we can display this vector field on the restricted domain $0 \le x \le 2$ and $-1 \le y \le 4$. At each point $(x,y)$ in the domain a vector $(1-x\,,\,\sin(y))^T$ is attached. It is very common that the vector field has to be rescaled for good visual results, i.e. instead of $\vec{F}$ the vector field $\alpha\,\vec{F}$ is displayed for a well chosen scalar factor $\alpha > 0$. With the code below Figure 3.2 is generated. In this case `MATLAB`/*Octave* used a reasonable scaling, but you can choose the scaling factor manually, see `help quiver` .

```
[x,y] = meshgrid([0:0.1:2],[-1:0.2:4]);  % generate the grid points
F1 = 1-x; F2 = sin(y);   % evaluate the vector field at the grid points

figure(1)
quiver(x,y,F1,F2)          % display the vector field
axis([0 2 -1 4])           % restrict the domain to be shown
xlabel('x');ylabel('y');   % label the axis
```

◇

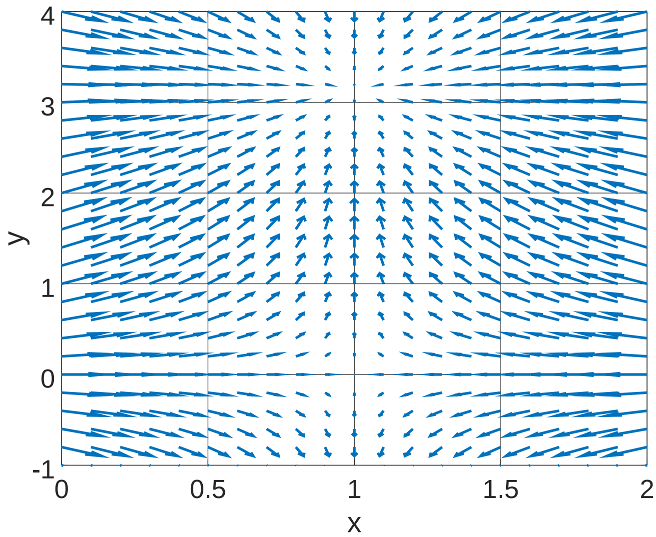
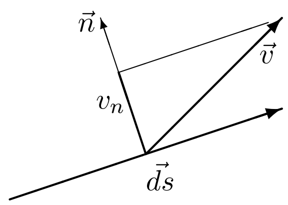



Figure 3.2: The vector field $(x-1\,,\,\sin(y))^T$

A vector field $\vec{v}$ could represent the speed at which a thin layer of liquid is moving in the plane $\mathbb{R}^2$. In Figure 3.2 the vectors would indicate the velocity and the direction of the moving layer of liquid.

If the line segment $\vec{ds}$ is part of the boundary of a domain we can construct the normal vector $\vec{n}$ (length $\|\vec{n}\| = 1$) to this segment and then compute the velocity $v_n$ of the liquid across the line segment.

$$v_n = \|\vec{v}\|\,\|\vec{n}\|\,\cos(\angle(\vec{n},\vec{v})) = \langle \vec{v}\,,\,\vec{n}\rangle\,.$$




Now consider a domain $D \subset \mathbb{R}^2$ with the curve $C$ as a boundary and the outer unit normal vector $\vec{n}$. Based on the above observation we can give an integral formula for the total flux out of the domain.

$$\text{flux} = \oint_C v_n\, d = \oint_C \langle \vec{v}\,,\,\vec{n}\rangle\, ds\,.$$

• **3–33 Example:** As a simple example compute the flux out of the rectangular domain $[0,2] \times [-1,4] \subset \mathbb{R}^2$ of the vector field $\vec{F} = (1-x\,,\,\sin(y))^T$, as shown in Figure 3.2. There are four contributions, each given by a line integral.

$$\text{flux} \;=\; \text{top} + \text{bottom} + \text{left} + \text{right}$$

$$\begin{aligned}
&= \int_{x=0}^{2} \langle \begin{pmatrix} 1-x \\ \sin(4) \end{pmatrix} , \begin{pmatrix} 0 \\ +1 \end{pmatrix} \rangle \, dx + \int_{x=0}^{2} \langle \begin{pmatrix} 1-x \\ \sin(-1) \end{pmatrix} , \begin{pmatrix} 0 \\ -1 \end{pmatrix} \rangle \, dx + \\
&\quad + \int_{y=-1}^{4} \langle \begin{pmatrix} 1-0 \\ \sin(y) \end{pmatrix} , \begin{pmatrix} -1 \\ 0 \end{pmatrix} \rangle \, dy + \int_{y=-1}^{4} \langle \begin{pmatrix} 1-2 \\ \sin(y) \end{pmatrix} , \begin{pmatrix} +1 \\ 0 \end{pmatrix} \rangle \, dy \\
&= \int_{x=0}^{2} \sin(4) \, dx + \int_{x=0}^{2} -\sin(-1) \, dx + \int_{y=-1}^{4} -1 \, dy + \int_{y=-1}^{4} -1 \, dy \\
&= 2 \sin(4) + 2 \sin(1) - 5 - 5
\end{aligned}$$

$\diamondsuit$

**• 3–34 Question:**
Determine the total flux of the velocity field

$$\vec{v} = \begin{pmatrix} 1 \\ x+y \end{pmatrix}$$

out of the disk $D$, a circle with radius 2 and center at the origin.

$\diamondsuit$

Examine the flux out of a rectangular box

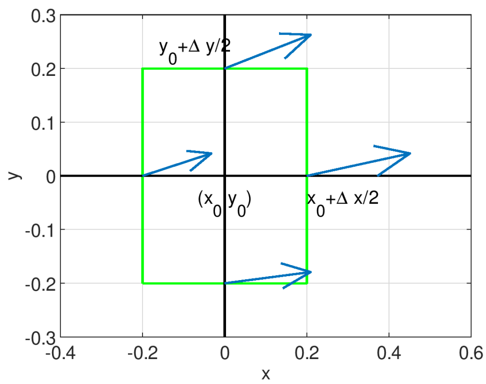
$$[x_0 - \frac{\Delta x}{2}, x_0 + \frac{\Delta x}{2}] \times [y_0 - \frac{\Delta y}{2}, y_0 + \frac{\Delta y}{2}] \subset \mathbb{R}^2$$

by the vector field $\vec{v} = (v_1 \, , \, v_2)$. The boundary consists of four straight line segments of length $\Delta x$, resp. $\Delta y$. Use a Taylor approximation of order 1 of the type

$$f(x_0 \pm \frac{\Delta x}{2}) \approx f(x_0) \pm \frac{\partial \, f(x_0)}{\partial x} \, \frac{\Delta x}{2} \quad \Longrightarrow \quad f(x_0 + \frac{\Delta x}{2}) - f(x_0 - \frac{\Delta x}{2}) \approx \frac{\partial \, f(x_0)}{\partial x} \, \Delta x$$

to estimate the total flux out of this small rectangle.

$$\begin{aligned}
\text{flux} &= \text{right} + \text{left} + \text{top} + \text{bottom} \\
&\approx +v_1(x_0 + \frac{\Delta x}{2}, y_0) \, \Delta y - v_1(x_0 - \frac{\Delta x}{2}, y_0) \, \Delta y + \\
&\quad +v_2(x_0, y_0 + \frac{\Delta y}{2}) \, \Delta x - v_2(x_0, y_0 - \frac{\Delta y}{2}) \, \Delta x \\
&\approx +\frac{\partial \, v_1(x_0, y_0)}{\partial x} \, \Delta x \, \Delta y + \frac{\partial \, v_2(x_0, y_0)}{\partial y} \, \Delta y \, \Delta x \\
&= \left( \frac{\partial \, v_1(x_0, y_0)}{\partial x} + \frac{\partial \, v_2(x_0, y_0)}{\partial y} \right) \, \Delta y \, \Delta x \\
&= \operatorname{div}(\vec{v}) \;\; \Delta y \, \Delta x
\end{aligned}$$




Thus the expression $\operatorname{div}(\vec{v})$ indicates how much flux is generated per unit area, it is the **source density** for this flux. If a domain $D \in \mathbb{R}^2$ is cut up into small rectangles of the above type, one can verify that the total flux out of the domain $D$ equals the sum of the fluxes out of the small rectangles. Based on the above idea this sum approaches the integral of the divergence $\operatorname{div} \vec{v}$ over the domain $D$. This leads to the definition of divergence of a vector field and the divergence theorem.

• **3–35 Definition:** For a vector field $\vec{v} = v_1(x,y)\,,\, v_2(x,y))$ the **divergence** is defined by

$$\text{divergence}(\vec{v}) = \operatorname{div}\vec{v} = \nabla\vec{v} = \vec{\nabla}\cdot\vec{v} = \frac{\partial\, v_1}{\partial x} + \frac{\partial\, v_2}{\partial y}\,.$$

For a vector field in $\mathbb{R}^3$ the correponding definition is

$$\operatorname{div}\vec{v} = \nabla\vec{v} = \frac{\partial\, v_1}{\partial x} + \frac{\partial\, v_2}{\partial y} + \frac{\partial\, v_3}{\partial z}\,.$$

♢

Some of the many applications of divergence are shown in Table 3.3.

| Application | Equation | Expressions |
|---|---|---|
| Fluid dynamics | $\operatorname{div}(\rho\,\vec{v}) = f$ | $\vec{v}$ velocity<br>$\rho$ density<br>$f$ source term |
| Heat conduction | $\rho\,\sigma\,\dot{T} = \operatorname{div}(k\;\operatorname{grad}T)$ | $T$ temperature<br>$\rho$ density<br>$\sigma$ specific heat<br>$k$ thermal conductivity |
| Electrostatics | $\operatorname{div}\vec{D} = \rho$ | $\vec{D} = \varepsilon\,\vec{E}$ electric field<br>$\rho$ charge density |
| Magnetostatics | $\operatorname{div}\vec{B} = 0$ | $\vec{B}$ magnetic field |

Table 3.3: Some equations using divergence

• **3–36 Result: Divergence Theorem**
Examine a domain $D \in \mathbb{R}^2$ with piecewise smooth boundary $C$ and outer unit normal vector $\vec{n}$ and a smooth vector field $\vec{v}$.

$$\text{flux} = \oint_C \langle\vec{n}\,,\,\vec{v}\rangle\; ds = \iint\limits_D \operatorname{div}(\vec{v})\; dA = \iint\limits_D \frac{\partial\, v_1}{\partial x} + \frac{\partial\, v_2}{\partial y}\; dA$$

If $D \in \mathbb{R}^3$ has the surface $S$ and the outer unit normal vector $\vec{n}$. Then the flux out of the domain $D$ can be computed by

$$\text{flux} = \oiint\limits_S \langle\vec{n}\,,\,\vec{v}\rangle\; dS = \iiint\limits_D \operatorname{div}(\vec{v})\; dV = \iiint\limits_D \frac{\partial\, v_1}{\partial x} + \frac{\partial\, v_2}{\partial y} + \frac{\partial\, v_3}{\partial z}\; dV$$

♢

• **3–37 Example:** As a simple example compute the flux out of the domain $[0,2]\times[-1,4]\subset\mathbb{R}^2$ of the vector field $\vec{F} = (1-x\,,\,\sin(y))^T$, as shown in Figure 3.2. Use the divergence theorem.

$$\operatorname{div}\vec{F} = \frac{\partial\,(1-x)}{\partial x} + \frac{\partial\,\sin(y)}{\partial y} = -1+\cos(y)$$

Thus we obtain

$$\begin{aligned} \text{flux} &= \iint\limits_D \operatorname{div} \vec F \, dA = \int_{x=0}^{2} \left( \int_{y=-1}^{4} -1 + \cos(y) \, dy \right) \, dx \\ &= \int_{x=0}^{2} (-5 + \sin(4) - \sin(-1)) \, dx = -10 + 2 \sin(4) - 2 \sin(-1) \, . \end{aligned}$$

This result coincides with Example 3–33. ♢

**• 3–38 Question:**
Determine the total flux of the velocity field

$$\vec v = \begin{pmatrix} 1 \\ x+y \end{pmatrix}$$

out of the disk $D$, a circle with radius 2 and center at the origin. Use the divergence theorem.

♢

The divergence theorem can also be used in different coordinate systems in $\mathbb{R}^2$ and $\mathbb{R}^3$. It is often convenient to express the vector field as a linear combination of the coordinate unit vectors. The formulas to compute div $\vec F$ are shown in Table 3.4.

| Coordinates | $\vec F$ | $\vec\nabla \cdot \vec F = \operatorname{div} \vec F = \text{divergence } \vec F$ |
|---|---|---|
| Cartesian | $F_x \, \vec e_x + F_y \, \vec e_y + F_z \, \vec e_z$ | $\frac{\partial F_x}{\partial x} + \frac{\partial F_y}{\partial y} + \frac{\partial F_z}{\partial z}$ |
| Cylindrical | $F_\rho \, \vec e_\rho + F_\phi \, \vec e_\phi + F_z \, \vec e_z$ | $\frac{1}{\rho} \frac{\partial (\rho F_\rho)}{\partial \rho} + \frac{1}{\rho} \frac{\partial F_\phi}{\partial \phi} + \frac{\partial F_z}{\partial z}$ |
| Spherical | $F_r \, \vec e_r + F_\phi \, \vec e_\phi + F_\theta \, \vec e_\theta$ | $\frac{1}{r^2} \frac{\partial (r^2 F_r)}{\partial r} + \frac{1}{r \sin\theta} \frac{\partial F_\phi}{\partial \phi} + \frac{1}{r \sin\theta} \frac{\partial (\sin\theta \, F_\theta)}{\partial \theta}$ |

Table 3.4: Divergence in different coordinate systems

**• 3–39 Question:**
Examine a ball $B \in \mathbb{R}^3$ with radius $R = 3$ and center at the origin and examine the vector field

$$\vec F(x, y, z) = \begin{pmatrix} x \\ y \\ z \end{pmatrix} .$$

Then determine the total flux of this vector field out of the ball $B$ by

(a) using the divergence theorem.

(b) setting up and computing the surface integral, using spherical coordinates.

♢

• **3–40 Question:**
Examine a ball $B \in \mathbb{R}^3$ with radius $R = 3$ and center at the origin and examine the vector field

$$\vec{F}(x,y,z) = \begin{pmatrix} x \\ y \\ 3 \end{pmatrix} .$$

Then determine the total flux of this vector field out of the ball $B$ by

(a) using the divergence theorem.

(b) setting up and computing the surface integral, using spherical coordinates.

♢

## 3.6 Integration by Parts, Green's Theorem, Conservation Laws

### 3.6.1 Conservation Laws

This section starts with consequences of the divergence theorem 3–36.

• **3–41 Example:** Let $u(\vec{x})$ be a solution of the partial differential equation

$$\nabla(a(\vec{x})\,\nabla u(\vec{x})) = \operatorname{div}(a(\vec{x})\,\operatorname{grad} u(\vec{x})) = f(\vec{x}) \quad \text{for} \quad \vec{x} \in \Omega\,.$$

This could be a steady state heat equation. Then the solution $u$ satisfies

$$\iint\limits_\Omega f\;dA = \iint\limits_\Omega \nabla(a\,\nabla u)\;dA = \oint_{\partial\Omega} a\,\frac{\partial\,u}{\partial\vec{n}}\;ds$$

where

$$\frac{\partial\,u}{\partial\vec{n}} = n_1\,\frac{\partial\,u}{\partial x} + n_2\,\frac{\partial\,u}{\partial y}$$

is the directional derivative of the function $u$ in the direction of the outer unit normal vector $\vec{n}$. Physically this corresponds to the condition that the thermal flux through the boundary has to equal to the heat generated inside the domain. If this condition is violated, there can be no steady state solution. ♢

• **3–42 Result:** Let $u(t,\vec{x})$ be a solution of the dynamic heat equation

$$\frac{\partial}{\partial t}\,u(t,\vec{x}) = \nabla(a(\vec{x})\,\nabla u(t,\vec{x})) + f(t,\vec{x}) \quad \text{for } t > 0 \text{ and } \vec{x} \in \Omega \in \mathbb{R}^2$$

with a positive coefficient function $a(\vec{x})$, then

$$\begin{aligned} \frac{d}{dt}\iint\limits_\Omega u(t,\vec{x})\;dA &= \iint\limits_\Omega \nabla(a(\vec{x})\,\nabla u(t,\vec{x})) + f(t,\vec{x})\;dA \\ &= \oint_{\partial\Omega} a(\vec{x})\,\frac{\partial\,u(t,\vec{x})}{\partial\vec{n}}\;ds + \iint\limits_\Omega f(t,\vec{x})\;dA\,. \end{aligned}$$

This statement is equivalent to the conservation of thermal energy. If the thermal energy in a domain $\Omega$ changes, the energy has either to flow through the boundary $\partial\Omega$ or is added by an external source term $f(t,\vec{x})$. ♢

The above result of conservation is also applicable to most of the applications listed in Table 3.3, and many more The conservation law for a wave equation will be given in Result 3–44 below.

### 3.6.2 From the Product Rule to Green's Identity

Thus usual product rule for derivatives

$$\frac{d}{dx}\,f(x)\,g(x) = f'(x)\;g(x) + f(x)\;g'(x)$$

can be used to find the product rule for the nabla operator $\nabla$, i.e. divergence and gradient.

$$\begin{aligned} \nabla\cdot\left(f(x,y)\begin{pmatrix} v_1(x,y) \\ v_2(x,y)\end{pmatrix}\right) &= \frac{\partial}{\partial x}\,(f(x,y)\,v_1(x,y)) + \frac{\partial}{\partial y}\,(f(x,y)\,v_2(x,y)) \\ &= \frac{\partial\,f(x,y)}{\partial x}\,v_1(x,y) + f(x,y)\,\frac{\partial\,v_1(x,y)}{\partial x} + \end{aligned}$$

$$\begin{aligned}
& +\frac{\partial\, f(x,y)}{\partial y}\, v_2(x,y) + f(x,y)\,\frac{\partial\, v_2(x,y)}{\partial y} \\
= & \ \langle \begin{pmatrix} \frac{\partial\, f(x,y)}{\partial x} \\ \frac{\partial\, f(x,y)}{\partial y} \end{pmatrix} , \begin{pmatrix} v_1(x,y) \\ v_2(x,y) \end{pmatrix} \rangle + f(x,y)\left(\frac{\partial\, v_1(x,y)}{\partial x} + \frac{\partial\, v_2(x,y)}{\partial y}\right) \\
= & \ \langle \nabla f \,,\, \vec{v} \rangle + f \cdot (\nabla \vec{v})
\end{aligned}$$

or with a different notation

$$\operatorname{div}(f\,\vec{v}) = (\operatorname{grad} f)\cdot \vec{v} + f\,(\operatorname{div}\vec{v}) \tag{3.1}$$

or with a notation very similar to the classical product rule

$$\nabla(f\,\vec{v}) = (\nabla f)\cdot\vec{v} + f\,\nabla\vec{v}\,.$$

Using this version of the product rule we can generate a formula similar to the integration by parts identity for functions of one variable.

$$\int_a^b f\cdot g'\,dx = +f(b)\cdot g(b) - f(a)\cdot g(a) - \int_a^b f'\cdot g\,dx$$

Using the above product rule (3.1) and the divergence theorem 3–36 leads to the **Green–Gauss theorem** or **Green's identity**.

$$\begin{aligned}
\iint\limits_G f\,(\operatorname{div}\vec{v})\,dA &= \iint\limits_G \operatorname{div}(f\,\vec{v}) - (\operatorname{grad} f)\cdot\vec{v}\,dA \\
&= \oint_{\partial G} f\,\vec{v}\cdot\vec{n}\,ds - \iint\limits_G (\operatorname{grad} f)\cdot\vec{v}\,dA
\end{aligned} \tag{3.2}$$

In the context of partial differential equations this is often used with $\vec{v} = \operatorname{grad} g$, i.e.

$$\begin{aligned}
\iint\limits_G f\,(\operatorname{div}\operatorname{grad} g)\,dA &= \oint_{\partial G} f\,(\operatorname{grad} g)\cdot\vec{n}\,ds - \iint\limits_G (\operatorname{grad} f)\cdot(\operatorname{grad} g)\,dA \qquad (3.3) \\
\iint\limits_G f\,\Delta g\,dA &= \oint_{\partial G} f\,\nabla g\cdot\vec{n}\,ds - \iint\limits_G \nabla f\cdot\nabla g\,dA\,.
\end{aligned}$$

• **3–43 Result:** The boundary value problem

$$\begin{aligned}
\Delta u = \nabla\cdot\nabla u &= f \quad \text{in } \Omega\subset\mathbb{R}^2 \\
u &= g \quad \text{on } \partial\Omega\subset\mathbb{R}^2
\end{aligned}$$

has at most one solution. ♢

**Proof :** Assume we have two solutions $u_1$ and $u_2$. Then the difference $v = u_1 - u_2$ satisfies

$$\begin{aligned}
\nabla\cdot\nabla v &= 0 \quad \text{in } \Omega\subset\mathbb{R}^2 \\
v &= 0 \quad \text{on } \partial\Omega\subset\mathbb{R}^2
\end{aligned}$$

and Green's identity (3.3) leads to

$$\iint\limits_\Omega \|\nabla v\|^2\,dA = \iint\limits_\Omega \nabla v\cdot\nabla v\,dA = \oint_{\partial G} v\,\nabla v\cdot\vec{n}\,ds - \iint\limits_G v\,\Delta v\,dA = 0$$

and thus the gradient of $v$ has to vanish everywhere. Since $v = 0$ on the boundary $\partial\Omega$ this implies that $v = 0$ everywhere and thus $u_1 = u_2$. □

• **3–44 Result:** Let $u(t,\vec{x})$ be a solution of the wave equation

$$\rho\,\frac{\partial^2}{\partial t^2}\,u(t,\vec{x}) = \nabla(a(\vec{x})\,\nabla u(t,\vec{x})) \quad \text{for } t>0 \text{ and } \vec{x}\in\Omega\in\mathbb{R}^2$$

with $u(t,\vec{x})=0$ or $\frac{\partial\, u(t,\vec{x})}{\partial\vec{n}}=0$ on the boundary $\partial\Omega$ and a positive coefficient function $a(\vec{x})$. Then

$$\begin{aligned}
\frac{d}{dt}\iint\limits_\Omega \frac{\rho(\vec{x})}{2}\,|\dot{u}(t,\vec{x})|^2\,dA &= \iint\limits_\Omega \rho(\vec{x})\,\dot{u}(t,\vec{x})\,\ddot{u}(t,\vec{x})\,dA = \iint\limits_\Omega \dot{u}(t,\vec{x})\,(\nabla(a(\vec{x})\,\nabla u(t,\vec{x}))\,dA \\
&= \oint_{\partial\Omega} \dot{u}(t,\vec{x})\,a(\vec{x})\,\frac{\partial\, u(t,\vec{x})}{\partial\vec{n}}\,ds - \iint\limits_\Omega \nabla\dot{u}(t,\vec{x})\cdot(a(\vec{x})\,\nabla u(t,\vec{x}))\,dA \\
&= -\iint\limits_\Omega a(\vec{x})\;\nabla\dot{u}(t,\vec{x})\cdot\;\nabla u(t,\vec{x}))\,dA \\
&= -\frac{1}{2}\,\frac{d}{dt}\,\iint\limits_\Omega a(x)\,\|\nabla u(t,\vec{x})\|^2\,dA\,.
\end{aligned}$$

As a consequence the energy $E(t)$ is independent on time $t$.

$$E(0)=E(t)=\frac{1}{2}\iint\limits_\Omega \rho(\vec{x})\,|\dot{u}(t,\vec{x})|^2 + a(\vec{x})\,\|\nabla u(t,\vec{x})\|^2\,dA$$

This is the sum of a kinetic energy and a potential energy. The similar result for domains $\Omega\subset\mathbb{R}^3$ in space is correct too. ♢

• **3–45 Result:** Let $u(t,\vec{x})$ be a solution of the dynamic heat equation

$$\frac{\partial}{\partial t}\,u(t,\vec{x}) = \nabla(a(\vec{x})\,\nabla u(t,\vec{x})) + f(t,\vec{x}) \quad \text{for } t>0 \text{ and } \vec{x}\in\Omega\in\mathbb{R}^2$$

with a positive coefficient function $a(\vec{x})$. Then

$$\begin{aligned}
\frac{1}{2}\,\frac{d}{dt}\iint\limits_\Omega |u(t,\vec{x})|^2\,dA &= \iint\limits_\Omega u(t,\vec{x})\,\dot{u}(t,\vec{x})\,dA \\
&= \iint\limits_\Omega u(t,\vec{x})\,(\nabla(a(\vec{x})\,\nabla u(t,\vec{x})) + f(t,\vec{x}))\,dA \\
&= \iint\limits_\Omega -\nabla u(t,\vec{x})\cdot(a(\vec{x})\,\nabla u(t,\vec{x})) + u(t,\vec{x})\,f(t,\vec{x}))\,dA\,+ \\
&\qquad + \oint_{\partial\Omega} u(t,\vec{x})\,a(\vec{x})\,\nabla u(t,\vec{x})\cdot\vec{n}\,ds\,.
\end{aligned}$$

If there is no external heating ($f(t,\vec{x})=0$) and on the boundary $\partial\Omega$ we have either a zero temperature ($u=0$) or a no flux condition ($\frac{\partial u}{\partial\vec{n}}=0$), then this implies

$$\frac{1}{2}\,\frac{d}{dt}\iint\limits_\Omega |u(t,\vec{x})|^2\,dA = -\iint\limits_\Omega a(\vec{x})\,\|\nabla u(t,\vec{x})\|^2\,dA \le 0\,,$$

i.e. the expression $\iint\limits_\Omega |u(t,\vec{x})|^2\,dA$ is decreasing. And it is strictly decreasing, unless the gradient of the function $u$ vanishes everywhere, i.e. the function is constant.

With similar computations we obtain, again using the boudary conditions and $f = 0$,

$$\begin{aligned}
\frac{1}{2}\frac{d}{dt}\iint\limits_{\Omega} a(\vec{x})\,\|\nabla u(t,x)\|^2\,dA &= \iint\limits_{\Omega} \langle a(\vec{x})\,\nabla u(t,x)\,,\,\nabla \dot{u}(t,x)\rangle\,dA \\
&= -\iint\limits_{\Omega} (\nabla(a(\vec{x})\,\nabla u(t,\vec{x}))\,\dot{u}(t,x)\,dA + \\
&\qquad + \oint_{\partial\Omega} a(\vec{x})\,\nabla u(t,x)\cdot\vec{n}\,\dot{u}(t,x)\,ds \\
&= -\iint\limits_{\Omega} |\dot{u}(t,x)|^2\,dA + 0 \leq 0
\end{aligned}$$

Thus solutions of the dynamic heat equation, without external heating, try to elinimate slopes, since the integral of $a(\vec{x})\,\|\nabla u(t,\vec{x})\|^2$ is decreasing.

Both of the above results are also correct for three dimensional domains $\Omega \in \mathbb{R}^3$. $\diamondsuit$

• **3–46 Result:** For the Calculus of Variations we will minimize functionals of the form

$$F(u) = \iint\limits_{\Omega} \frac{1}{2}\,a\,(\nabla u)^2 + \frac{1}{2}\,b\,u^2 + f\cdot u\,dA\,,$$

i.e. we seek the function $u$ which renders the above expression minimal. To derive the Euler–Lagrange equation we need the following identity, which is a direct consequence of Green's identity (3.3). Assume that $u$, $\phi$, $a$ and $f$ are smooth functions.

$$\begin{aligned}
\iint\limits_{\Omega} a\,\nabla u\cdot\nabla\phi + b\,u\,\phi + f\cdot\phi\,dA &= \iint\limits_{\Omega} (-\nabla(\,a\,\nabla u) + b\,u + f)\cdot\phi\,dA + \int_{\partial\Omega} a\,\vec{n}\cdot\nabla u\;\phi\,ds \\
&= \iint\limits_{\Omega} (-\nabla(\,a\,\nabla u) + b\,u + f)\cdot\phi\,dA + \oint_{\partial\Omega} \phi\,a\,\frac{\partial\,u}{\partial\vec{n}}\,ds
\end{aligned}$$

$\diamondsuit$

# Chapter 4

# Ordinary Differential Equations

## 4.1 Keywords and Literature

Ordinary differential equation (ODE), partial differential equation (PDE), vector field, separation of variables, homogeneous problems and their solutions, inhomogeneous problems, particular solution, resonator, phase portrait.

Literature:

- Any good book on Engineering Mathematics should cover the topics on ODEs presented in this section.
- The book [ONei11] (or an other edition) has a section on ODEs.
- The book [CrofDaviHarg92] (or a newer edition) has a a section on ODEs.
- The books by Earl W. Swokowski (e.g. [Swok92] or newer editions) are quite readable.
- The books by Lothar Papula cover all required topics, but these books are in German.
- The lecture notes by A. Stahel cover all of the topics on ODEs, available from the web page at https://web.sha1.bfh.science/Math1.pdf in German.

## 4.2 Ordinary Differential Equations of Order 1

- **4–1 Definition:** Let $f : \mathbb{R}\times\mathbb{R}^n \longrightarrow \mathbb{R}^n$ be a continuous function.
  - A differentiable function $\vec{x}(t)$ is a solution of the **ordinary differential equation** (ODE) of order 1 if the equation
  $$\frac{d}{dt}\,\vec{x}(t) = \vec{f}(t,\vec{x}(t)) \tag{4.1}$$
  is solved. The solution is a function $\vec{x}(t)$ with independent variable $t\in\mathbb{R}$ and dependent variable $\vec{x}\in\mathbb{R}^n$.
  - If $x\in\mathbb{R}$ we have a single ODE and if $\vec{x}\in\mathbb{R}^n$ with $n>1$ we have a system of ODEs.
  - If a starting time $t_0$ and an initial value $\vec{x}(t_0)=\vec{x}_0$ are given, together with (4.1) then we have an **initial value problem**.
  - If the function $f$ is independent on $t$, then we have an **autonomous** differential equation.

$\diamondsuit$

• **4–2 Result:** If the function $f$ is continuous and differentiable with respect to $\vec{x}$, then the above initial value problem has a unique solution $\vec{x}(t)$ for $a < t < b$ with $t_0 \in (a,b)$. Thus we have one (and only one) solution on an interval containing the initial time $t_0$. If $b < \infty$ then $\lim_{t\to b} \|\vec{x}(t)\| = \infty$ and if $-\infty < a$ then $\lim_{t\to a} \|\vec{x}(t)\| = \infty$. ♢

• **4–3 Example:** With $f(x) = 1 + x^2$ we have the autonomous ODE

$$\frac{d}{dt}\, x(t) = 1 + x^2(t) \tag{4.2}$$

with the solutions $x(t) = \tan(t + c)$, where $c$ is an arbitrary constant. To verify this compute

$$\frac{d}{dt}\, x(t) = \frac{d}{dt}\, \tan(t+c) \; = 1 + \tan^2(t+c) \quad \text{and} \quad 1 + x^2(t) = 1 + \tan^2(t+c)\,.$$

Observe that it is (usually) easy to verify that a given function solves the ODE, but it might be difficult to find the explicit formula for the solution. With the initial condition $x(0) = 0$ find the unique solution of (4.2) $x(t) = \tan t$, defined on $a = \frac{-\pi}{2} < t < \frac{+\pi}{2} = b$.

- Visualization
  Since $\dot{x}(t) = f(t, x(t)) = 1 + x^2(t)$ we know that the vector

$$\begin{pmatrix} 1 \\ f(t, x(t)) \end{pmatrix} = \begin{pmatrix} 1 \\ 1 + x^2(t) \end{pmatrix}$$

is a tangent vector to the curve $(t\,,\,x(t))$ in the $tx$–plane. Thus we can draw the vector field $(1\,,\,1 + x^2)$ and a few solutions, as seen in Figure 4.1. The solution in red is $x(t) = \tan(t)$ and it is the solution of the initial value problem with the initial condition $x(0) = 0$. Since the ODE is autonomous the slopes of the vectors in the figure depend on $x$ only, and not on $t$.

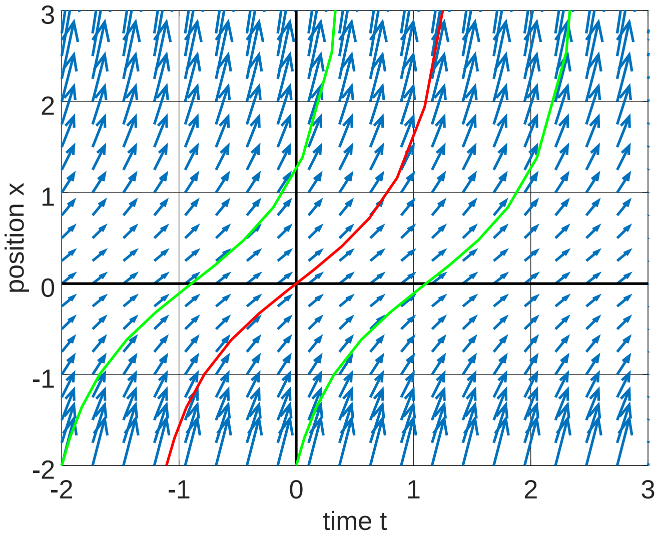



Figure 4.1: Vector field and solutions for the ODE $\dot{x} = 1 + x^2$

Figure 4.1 was generated by the *Octave*/`MATLAB` code below.

```
t = linspace(-2,3,20);   x = linspace(-2,3,21);
[t,x] = meshgrid(t,x);
f_t = ones(size(t)); f_x = 1+x.^2;

figure(1); clf
quiver(t,x,f_t,f_x,3)
xlabel('time t'); ylabel('position x'); axis([-2 3 -2 3])

[t1, x1] = ode45 (@(t,x)1+x^2,[atan(-2) atan(3)],-2);
[t2, x2] = ode45 (@(t,x)1+x^2,[-2 0.5],-2);
[t3, x3] = ode45 (@(t,x)1+x^2,[0 2.5],-2);
hold on
plot([-2 3],[0 0],'k',[0 0],[-2 3],'k')
plot(t1,x1,'r',t2,x2,'g',t3,x3,'g')
hold off
```

- Exact solution
  For the ODE (4.2) we can use the technique of separation of variables to determine the solution formula. Assume that $x(t)$ is a solution, then compute

$$\begin{aligned}
\frac{d}{d\tau}\,x(\tau) &= 1+x^2(\tau) \qquad \text{divide by } 1+x^2(\tau) \text{ and integrate}\\
\int_{\tau=0}^{t} \frac{1}{1+x^2(\tau)}\,\frac{d}{d\tau}\,x(\tau)\;d\tau &= \int_{\tau=0}^{t} 1\;d\tau \qquad \text{substitution } u=x(\tau)\\
\int_{u=x(0)}^{x(t)} \frac{1}{1+u^2}\;du &= t \qquad \text{use } \tfrac{d}{du}\arctan u = \tfrac{1}{1+u^2}\\
\arctan u\,\Big|_{u=x(0)}^{x(t)} &= t\\
\arctan(x(t)) &= t-\arctan(x(0))\\
x(t) &= \tan(t+c)
\end{aligned}$$

$\diamondsuit$

The above computation is known under the name **separation of variables** and there is a simplified notation[1] for this method. By computing two integrals you end up with an implicit equation for $x$ and $t$, which than has to be solved for the dependent variable $x$ as function of $t$.

$$\begin{aligned}
\frac{dx}{dt} &= 1+x^2 \quad \text{separate the variables } x \text{ and } t\\
\frac{dx}{1+x^2} &= 1\;dt \quad \text{integrate}\\
\int \frac{1}{1+x^2}\;dx &= \int 1\;dt\\
\arctan(x) &= t+c \quad \text{solve for } x\\
x(t) = x &= \tan(t+c)
\end{aligned}$$

The method of separation of variables is applicable to differential equations where the right hand side is a product of a function of the dependent variable $x$ and a function of the independent

[1] Never mind that it is not correct to multiply both sides by $dt$ (as $dt$ it is part of the differential operator $\frac{d}{dt}$), and then intgerate the left hand side with respect to $x$ and the right hand side with respect to $t$, and hope for correct results! In this setup twice wrong turns into correct.

variable $t$.

$$\begin{aligned}
\frac{d\,x}{dt} &= h(t)\cdot g(x) \quad \text{separate the variables } x \text{ and } t \\
\frac{d\,x}{g(x)} &= h(t)\;dt \quad \text{integrate} \\
\int \frac{1}{g(x)}\;dx &= \int h(t)\;dt \quad \text{then solve for } x
\end{aligned}$$

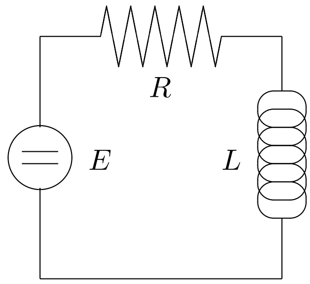
For autonomous ODEs $\dot x(t) = f(x(t))$ the right hand side simplifies to $\int h(t)\;dt = \int 1\;dt = t + c$.

• **4–4 Question:**

For the electric "network" on the right the external voltage source $E(t)$, the resistor $R$ and the inductance $L$ are known. Use Ohm's law to verify that the equation describing the current $I(t)$ through the resistor $R$ is given by

$$L\,\frac{d\,I(t)}{dt} + R\cdot I(t) \;=\; E(t)$$

with an initial condition $I(0) = 0$.




(a) Solve this differential equation for the case $E(t) = E_0$ constant.

(b) Choose some values for $E_0$, $L$ and $R$ and draw the vector field. Then sketch a few solutions and describe their behavior.

♢

Usually the key point for ODE is not solving them, but setting up the correct differential equation. As an example consider the following example from biology.

• **4–5 Example:** To examine the size of a population consider

$$\begin{aligned}
0 < p(t) &= \text{size of the population at time } t\,, \\
0 < \alpha &= \text{birth rate}\,, \\
0 < \beta &= \text{death rate}\,.
\end{aligned}$$

For one time unit we find that $\alpha\,p(t)$ animals will be born and $\beta\,p(t)$ will die. Using this we find a differential equation

$$\dot p = \alpha\,p - \beta\,p = (\alpha-\beta)\,p$$

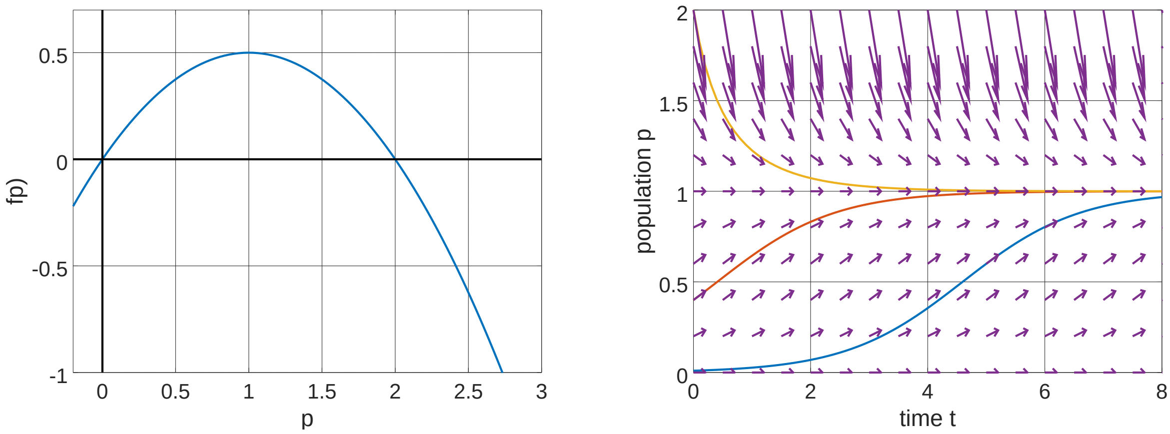
with the general solution

$$p(t) = p(0)\;e^{(\alpha-\beta)\,t}$$

For $\alpha > \beta$ the size of the population will grow exponentially. Once the population is very large this is certain to be wrong. Thus we modify the model by using a larger death rate for large populations. This might be caused by limited food supply. One possible solution is

$$\dot p = (\alpha-\beta)\,p - k\,p^2 = p\;(\alpha-\beta-k\,p) = f(p)$$

for a positive constant $k$.




(a) Graph of $f(p)$

(b) Vector field and solutions

Figure 4.2: Function, vector field and solutions for the logistic equation

This model was proposed by Verhulst[2] and experimentally confirmed. Find the graph of $f(p)$ as function of $p$ in Figure 4.2(a) for the values $\alpha = 2$, $\beta = 1$ and $k = 0.5$. Three solutions of the ODE and the vector field are shown in Figure 4.2(b). This differential equation is also called the **logistic equation**.

The logistic differential equation

$$\dot p = p\;(\alpha - \beta - k\,p) = f(p)$$

is an example of a separabel differential equation and we can construct a solution formula. First separate the variables $p$ and $t$

$$\frac{\dot p(t)}{p(t)\;\;(\alpha - \beta - k\,p(t))} = 1$$

and then integrate with respect to $t$, with a substitution $z = p(t)$

$$\begin{aligned}
\int_0^T \frac{\dot p(t)}{p(t)\;\;(\alpha - \beta - k\,p(t))}\,dt &= \int_0^T 1\;dt \\
\int_{p(0)}^{p(T)} \frac{1}{z\;\;(\alpha - \beta - k\,z)}\,dz &= T\,.
\end{aligned}$$

To evaluate the integral on the left hand side use a partial fraction expansion.

$$\frac{1}{z\;\;(\alpha-\beta-k\,z)} = \frac{A}{z} + \frac{B}{\alpha-\beta-k\,z} = \frac{A\,(\alpha-\beta-k\,z) + B\,z}{z\;\;(\alpha-\beta-k\,z)}\,.$$

To determine the values of $A$ and $B$ we need to solve the equation

$$1 = A\,(\alpha-\beta-k\,z) + B\,z \quad \text{for all } z$$

for $A$ and $B$. Use the special values $z = 0$ and $z = \frac{\alpha-\beta}{k}$ to obtain $A = \frac{1}{\alpha-\beta}$ and $B = \frac{k}{\alpha-\beta}$ and thus

$$\begin{aligned}
\int \frac{1}{z\;\;(\alpha-\beta-k\,z)}\,dz &= \frac{1}{\alpha-\beta}\int \frac{1}{z} + \frac{k}{\alpha-\beta-k\,z}\,dz \\
&= \frac{1}{\alpha-\beta}\int \frac{1}{z} + \frac{1}{\frac{\alpha-\beta}{k} - z}\,dz \\
&= \frac{1}{\alpha-\beta}\left(\ln|z| - \ln|\frac{\alpha-\beta}{k} - z|\right) + c_1\,.
\end{aligned}$$

[2]Pierre–Francois Verhulst (1804–1849), a biologist and mathematician

Use this in the above integral to conclude

$$\frac{1}{\alpha-\beta}\left(\ln|p(T)| - \ln|\frac{\alpha-\beta}{k} - p(T)|\right) = T + c\,.$$

For all solutions with $0 < p(T)$ we find

$$\ln|p(T)| - \ln|\frac{\alpha-\beta}{k} - p(T)| \;=\; \ln p(T) - \ln(\frac{\alpha-\beta}{k} - p(T)) \;=\; \ln\left|\frac{p(T)}{\frac{\alpha-\beta}{k} - p(T)}\right|$$

and thus

$$\left|\frac{p(T)}{\frac{\alpha-\beta}{k} - p(T)}\right| = e^{(\alpha-\beta)\,(T+c)} = c_1\,e^{(\alpha-\beta)\,T}\,.$$

Stick the sign in the constant $c_2$ and replace $T$ by $t$ to find the implicit solution.

$$p(t) = c_2\,e^{(\alpha-\beta)\,t}\left(\frac{\alpha-\beta}{k} - p(t)\right)$$

This can be solved for $p(t)$ with the result

$$p(t) = \frac{c_2\,e^{(\alpha-\beta)\,t}\;\frac{\alpha-\beta}{k}}{1 + c_2\,e^{(\alpha-\beta)\,t}}\,.$$

The value of $c_2$ would have to be determined with an initial condition. If $\alpha > \beta$ the this expression leads to

$$\lim_{t\to\infty} p(t) = \frac{\alpha-\beta}{k}$$

which is consistent with the solutions shown in Figure 4.2(b). ♢

**Linear differential equations of order one**

For linear differential equations of order one with a constant coefficient

$$\dot y(t) + a\,y(t) = f(t)$$

one can verify that any solution $y(t)$ is of the form

$$y(t) = c\,e^{-a\,t} + y_p(t)$$

where

- $y_p(t)$ is one particular solution.
- $c$ is a constant to be determined by a possible initial condition.

For some types of functions $f(t)$ one can use a clever ansatz for the particular solution $y_p(t)$ and then determine the missing constants by using the differential equation.

- If $f(t)$ is a polynomial in $t$, try a polynomial of the same degree.
- If $f(t) = e^{\lambda t}$ with $\lambda \neq -a$, try $y_p(t) = k\,e^{\lambda t}$.
- If $f(t) = \cos(\omega\,t)$ or $f(t) = \sin(\omega\,t)$ try $y_p(t) = A\,\cos(\omega\,t) + B\,\sin(\omega\,t)$.

- **4–6 Question:**

  (a) Find the general solution of $\dot y(t) + 2\,y(t) = 2\,t$ .

  (b) Find the general solution of $\dot y(t) + 2\,y(t) = \cos(3\,t)$.

  (c) Find the general solution of $\dot y(t) + 2\,y(t) = 3\,e^{-3\,t}$.

♢

## 4.3 Ordinary Differential Equations of Order 2

In this section linear ODEs of order 2 with constant coefficients are examined. This very special type of ODE has many important applications.

• **4–7 Definition:** A second order, linear ODE with constants coefficients is of the form

$$a\,\ddot{y}(t) + b\,\dot{y}(t) + c\,y(t) = f(t)\,. \tag{4.3}$$

For a given function $f(t)$ and constants $a \neq 0$, $b$ und $c$.

- If the function $f(t)$ is not identcally zero, then we have an inhomogenoeus problem.
- If $f(t) = 0$ for all $t \in \mathbb{R}$, then we have an homogenoeus problem.

♢

By dividing the equation by $a \neq 0$ we only have to examine equations of the type

$$\ddot{y}(t) + b\,\dot{y}(t) + c\,y(t) = f(t)\,.$$

• **4–8 Result:** Let $b(t)$, $c(t)$ and $f(t)$ be continuous functions.

- The homogeneous ODE
$$\ddot{y}(t) + b(t)\,\dot{y}(t) + c(t)\,y(t) = 0$$
has two linearly independent solutions $y_1(t)$ and $y_2(t)$. Any solution $y_h(t)$ of this homogeneous equation is of the form
$$y_h(t) = c_1\,y_1(t) + c_2\,y_2(t)$$
with constants $c_1$ and $c_2$. If $a$ and $b$ are constant, this type of ODE can be visualized by vector fields, in this case called a phase portrait and examined in Section 4.3.3 on page 48.
- The inhomogeneous ODE
$$\ddot{y}(t) + b(t)\,\dot{y}(t) + c(t)\,y(t) = f(t)$$
has at least one **particular** solution $y_p(t)$. Any solution $y(t)$ is of the form
$$y(t) = y_p(t) + c_1\,y_1(t) + c_2\,y_2(t)$$
with constants $c_1$ and $c_2$.
- For given values of $t_0$, $y_0$ and $y_1$ the initial value problem
$$\left\{\begin{array}{rcl} \ddot{y}\,(t) + b(t)\,\dot{y}(t) + c(t)\,y(t) & = & f(t) \\ y(t_0) & = & y_0 \\ \dot{y}(t_0) & = & y_1 \end{array}\right.$$
has exactly one solution and it is of the above form. The values of $c_1$ and $c_2$ can be determined by using the initial conditions.

♢

Based on the above result there is a clear algorithm to solve this type of ODE.

1. Determine the linearly independent solutions $y_1(t)$ and $y_2(t)$ of the homogeneous problem.
2. If $f(t) \neq 0$ find a particular solution $y_p(t)$.
3. Determine the values of the constants $c_1$ and $c_2$ using the initial conditions.

### 4.3.1 Solutions of the homogeneous problem

As a consequence of the above results it is sufficient to find two (linearly independent) solutions of

$$a\,\ddot{y}(t) + b\,\dot{y}(t) + c\,y(t) = 0$$

to know all solutions. For this use the ansatz

$$y(t) = e^{\lambda\,t}$$

and plug it into the ODE to find

$$a\,\ddot{y}(t) + b\,\dot{y}(t) + c\,y(t) = a\,\lambda^2\,e^{\lambda\,t} + b\,\lambda\,e^{\lambda\,t} + c\,e^{\lambda\,t} = 0\,.$$

This problem can be solved using the **characteristic equation**

$$a\,\lambda^2 + b\,\lambda + c = 0\,. \tag{4.4}$$

Depending on the value of the **discriminant** $D = b^2 - 4\,a\,c$ we find different types of solutions.

1. $D > 0 :\ \lambda_1 \neq \lambda_2$ and both real.
   Then we have two linearly independent solutions
   $$y_1(t) = e^{\lambda_1\,t} \quad \text{and} \quad y_2(t) = e^{\lambda_2\,t}$$
   and thus the general solution is
   $$y(t) = c_1\,y_1(t) + c_2\,y_2(t) = c_1\,e^{\lambda_1\,t} + c_2\,e^{\lambda_2\,t}\,.$$
2. $D = 0 :\ \lambda = \lambda_1 = \lambda_2 \in \mathbb{R}$
   One solution is immediate from the previous case
   $$y_1(t) = e^{\lambda_1\,t}$$
   and one can easily verify the a second solution is given by
   $$y_2(t) = t\;e^{\lambda_1\,t}\,.$$
   Thus the general solution is
   $$y(t) = c_1\,y_1(t) + c_2\,y_2(t) = e^{\lambda_1\,t}\ (c_1 + t\;c_2)\ .$$
3. $D < 0 :\ \lambda_{1,2} = \alpha \pm i\beta$ with $\alpha, \beta \in \mathbb{R}$ and $\beta \neq 0$
   Using the idea form the first case we end up with **complex** solutions
   $$\begin{aligned} u_1(t) &= e^{\lambda_1\,t} = e^{(\alpha+i\beta)\,t} = e^{\alpha t}\ (\cos(\beta t) + i\,\sin(\beta t)) \\ u_2(t) &= e^{\lambda_2\,t} = e^{(\alpha-i\beta)\,t} = e^{\alpha t}\ (\cos(\beta t) - i\,\sin(\beta t))\ . \end{aligned}$$
   Since we can use linear combinations of solutions to the homogeneous problem to generate other solutions we use
   $$\begin{aligned} y_1(t) &= \tfrac{1}{2}\ (u_1(t) + u_2(t)) = e^{\alpha t}\,\cos(\beta t) \\ y_2(t) &= \tfrac{1}{2\,i}\ (u_1(t) - u_2(t)) = e^{\alpha t}\,\sin(\beta t) \end{aligned}$$
   to end up with two linearly independent, real-valued solutions.
   $$y(t) = c_1\,y_1(t) + c_2\,y_2(t) = e^{\alpha\,t}\ (c_1\,\cos(\beta t) + c_2\,\sin(\beta t))\ .$$

• **4–9 Question:**
Find the solution of the initial value problem

$$\left\{\begin{array}{rcl} \ddot y(t) - 2\,\dot y(t) + 5\,y(t) & = & 0 \\ y(0) & = & 2 \\ \dot y(0) & = & 4 \end{array}\right. .$$

♢

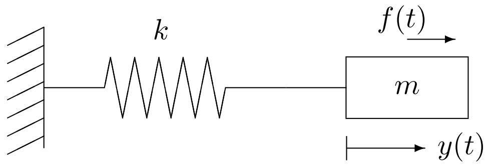
• **4–10 Question:**
Determine the unique solution of the initial value problem

$$\left\{\begin{array}{rcl} \ddot y(t) + 4\,\dot y(t) + 4\,y(t) & = & 0 \\ y(1) & = & 2 \\ \dot y(1) & = & -1 \end{array}\right. .$$

♢

• **4–11 Example:**

Let $y(t)$ denote the horizontal displacement of a mass $m$. This mass is connected to a wall with a spring with spring constant $k$. Consider a frictional force proportional to the velocity of the mass with a constant $\alpha$. An additional external force $f(t)$ is applied to the mass. Using Newton's law the movement of the mass described by the ODE



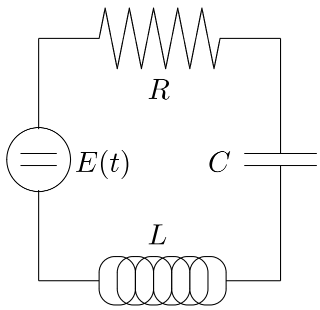

$$m\,\ddot y(t) = -k\,y(t) - \alpha\,\dot y(t) + f(t)\,.$$

♢

• **4–12 Example:**

Let $I$ be the current through a resistor $R$ and $Q(t)$ the electrical charge on the capacitor $C$. Then we use $\frac{d}{dt}\,Q = I$ and Kirchhoff's law to conclude

$$L\,\ddot Q(t) + R\,\dot Q(t) + \frac{1}{C}\,Q(t) = E(t)\,.$$




♢

The above two examples allow to draw analogies between mechanical and electrical systems, shown in Table 4.1.

• **4–13 Result:** The equation of a simple mass-spring system is given by

$$\ddot y + \frac{\alpha}{m}\,\dot y + \frac{k}{m}\,y = 0$$

Consider $m$ and $k$ to be fixed, but consider different cases for the friction $\alpha > 0$ and examine the qualitative behavior of the solutions.

- weak damping, $\alpha^2 < 4\,k\,m$

| electrical | | mechanical | |
|---|---|---|---|
| inductance | $L$ | mass | $m$ |
| resistance | $R$ | coeff. of friction | $\alpha$ |
| capacity | $1/C$ | spring constant | $k$ |
| charge | $Q(t)$ | displacement | $y(t)$ |
| current | $I(t)$ | velocity | $\dot{y}(t)$ |
| external voltage | $E(t)$ | external force | $f(t)$ |

Table 4.1: Analogy between electrical and mechanical resonators

- strong damping, $\alpha^2 > 4\,k\,m$
- critical damping, $\alpha^2 = 4\,k\,m$

The characteristic equation is

$$m\,\lambda^2 + \alpha\,\lambda + k = 0$$

and we have to distinguish three different cases for the discriminant $D = \alpha^2 - 4\,k\,m$ .

1. **weak damping** $\alpha^2 < 4km$, i.e. $D < 0$
   In this case the two solutions of the characteristic equation are

   $$\lambda_{1,2} = -\frac{\alpha}{2\,m} \pm i\sqrt{\frac{k}{m} - \left(\frac{\alpha}{2m}\right)^2} = -\frac{\alpha}{2\,m} \pm i\,\omega$$

   and thus we have the following general solution of the ODE

   $$y(t) = e^{\frac{-\alpha}{2\,m}\,t}\ (c_1\,\cos(\omega\,t) + c_2\,\sin(\omega\,t)) = c\,e^{\frac{-\alpha}{2\,m}\,t}\ \cos(\omega\,t + \delta)$$

   for suitable constants $c > 0$ and $\delta$. Thus we have a periodic contribution $\cos(\omega\,t + \delta)$ with angular velocity

   $$\omega = \sqrt{\frac{k}{m} - \left(\frac{\alpha}{2m}\right)^2}$$

   and an exponentially decaying amplitude with exponent $\frac{-\alpha}{2\,m}$. Find a typical solution in Figure 4.3(a). This solution has infinitely many zeros for $t > 0$, with a fixed distance of $T = \frac{2\,\pi}{\omega}$.

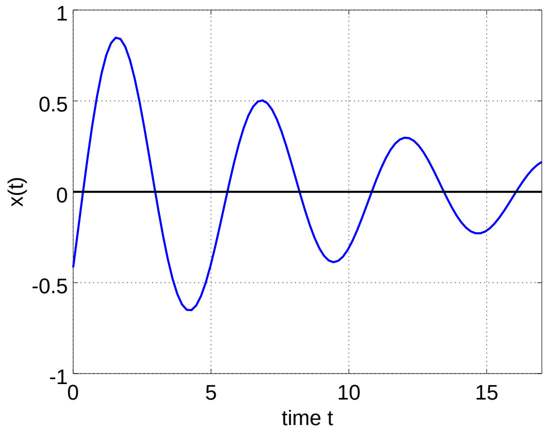
2. **strong damping** $\alpha^2 > 4\,k\,m$, i.e. $D > 0$
   In this case the zeros of the characteristic equation are given by

   $$\lambda_{1,2} = -\frac{\alpha}{2\,m} \pm \frac{1}{2\,m}\sqrt{\alpha^2 - 4\,k\,m} = -\frac{\alpha}{2\,m} \pm \beta$$

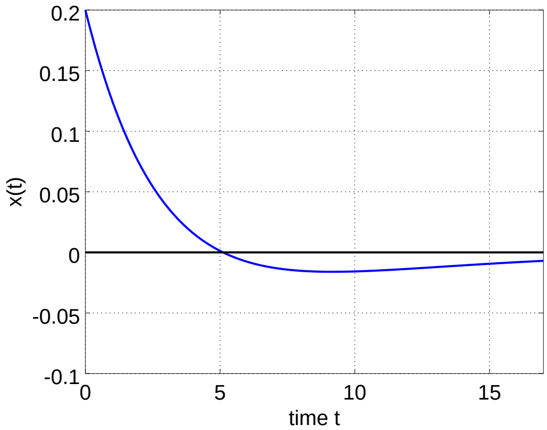
   and thus we find the general solution

   $$\begin{aligned} y(t) &= \exp\left(\frac{-\alpha}{2\,m}\,t\right)\ \left(c_1\,e^{\beta\,t} + c_2\,e^{-\beta\,t}\right) \\ &= \exp\left(-\frac{\alpha - \sqrt{\alpha^2 - 4\,k\,m}}{2\,m}\,t\right)\left(c_1 + c_2\,\exp\left(-\frac{\sqrt{\alpha^2 - 4\,k\,m}}{m}\,t\right)\right)\,. \end{aligned}$$

   Thus the solution converges exponentially to zero, proportional to function $\exp(-\frac{\alpha-\sqrt{\alpha^2-4\,k\,m}}{2\,m}\,t)$. Find a typical solution in Figure 4.3(b). This solution has one or no zeros for $t > 0$.




(a) weak damping




(b) strong damping

Figure 4.3: Qualitative behavior of solution to damped oscillators

3. **critical damping** $\alpha^2 = 4\,k\,m$, i.e. $D = 0$
   In this case the double zero of the characteristic equation is
   $$\lambda = \lambda_{1,2} = -\frac{\alpha}{2\,m}$$
   and the resulting, general solution of the ODE
   $$y(t) = e^{\frac{-\alpha}{2\,m}\,t}\,\left(c_1\, + c_2\,t\right)\,.$$
   The graph of the typical solution is similar to the case of strong damping and the solution has again one or no zero for $t > 0$.

In all three cases we have an exponential decay of the solution. The decay exponent $D(\alpha)$ (as function of $\alpha$ is given by

$$D(\alpha) = \begin{cases} \dfrac{\alpha}{2\,m} & \text{for } \alpha^2 \leq 4\,k\,m \\[2ex] \dfrac{\alpha - \sqrt{\alpha^2 - 4\,k\,m}}{2\,m} & \text{for } \alpha^2 > 4\,k\,m \end{cases}$$

and its graph shown in Figure 4.4. ♢

### 4.3.2 Solutions of the inhomogeneous problem

To determine all solutions of an inhomogeneous problem

$$a\,\ddot{y}(t) + b\,\dot{y}(t) + c\,y(t) = f(t) \tag{4.5}$$

we only need one particular solution $y_p(t)$ and can then use the general solution of the homogeneous problem in the previous section to write down the general solution. One of the most efficient methods to find a $y_p(t)$ is the method of **undetermined coefficients**[3].

1. If the expression $f(t)$ is a sum of different contributions, we can treat them individually and at the end add up the different contributions.

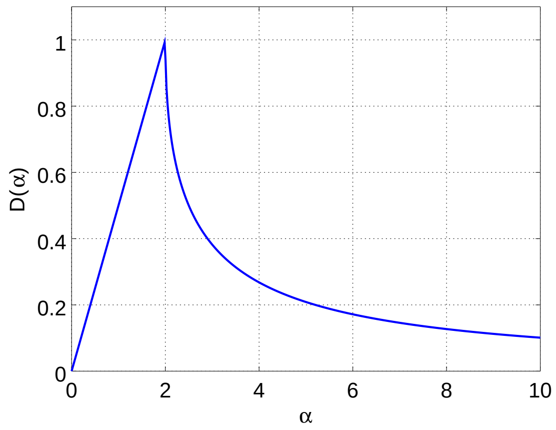



Figure 4.4: Decay exponent as function of the damping constant $\alpha$

| function $f(t)$ | $\longrightarrow$ | ansatz for particular solution $y_p(t)$ |
|---|---|---|
| polynomial of degree $n$ | $\longrightarrow$ | polynomial of degree $n$ |
| exponential function $e^{\alpha\, t}$ | $\longrightarrow$ | exponential function $e^{\alpha\, t}$ |
| trigonometric function with $\omega$ | $\longrightarrow$ | trigonometric function with $\omega$ |
| product of the above | $\longrightarrow$ | product of the above |

Table 4.2: Ansatz for particular solutions

2. To guess a correct form use Table 4.2.
3. Plug your ansatz for $y_p(t)$ into the ODE (4.5). If you obtain

$$a\,\ddot y_p(t) + b\,\dot y_p(t) + c\,y_p(t) = 0$$

then repeat with the new ansatz $t \cdot y_p(t)$. Repeat if necessary.

Once you have the correct ansatz for $y_p(t)$, plug it into the ODE (4.5) and determine the unknown coefficients.

• **4–14 Question:**

(a) Find the general solution of

$$\ddot y(t) - 4\,y(t) = 2\,e^{3\,t}\,.$$

(b) Find a particular solution for the differential equation

$$\ddot y(t) + \dot y(t) + 4\,y(t) = 7\,\sin(t)\,.$$

(c) Find the general solution of

$$\ddot y(t) - 4\,y(t) = 3\,e^{2\,t}\,.$$

◊

[3] I prefer to call it the method of intelligent guesses.

• **4–15 Example:** For a small, positive constant $0 < \delta$ the ODE

$$\ddot{y} + 2\,\delta\,\dot{y} + \omega_0^2\,y = F_0\,\cos(\omega\,t)$$

describes a weakly damped oscillator with natural frequency $\omega_0$. Without damping ($\delta = 0$) the system would oscillate with an angular velocity of $\omega_0$, i.e. a natural frequency of $\nu_0 = \frac{\omega_0}{2\,\pi}$. The inhomogeneous term $F_0\,\cos(\omega\,t)$ describes an external, periodic force with amplitude $F_0$. The goal is to describe the reaction of the system to this external force, as function of $\omega$.

**Homogeneous solution:** The characteristic equation is

$$\lambda^2 + 2\,\delta\,\lambda + \omega_0^2 = 0$$

with the solutions

$$\lambda_{1,2} = \frac{1}{2}\left(-2\,\delta \pm \sqrt{4\delta^2 - 4\omega_0^2}\right) = -\delta \pm i\sqrt{\omega_0^2 - \delta^2} = -\delta \pm i\,\omega_\delta\,.$$

Thus the homogeneous solution is given by

$$y_h(t) = e^{-\delta t}\,\left(c_1\,\cos(\omega_\delta\,t) + c_2\,\sin(\omega_\delta\,t)\right) \underset{t\to\infty}{\longrightarrow} 0\,.$$

**Particular solution:** One possible form for the particular solution is

$$y_p(t) = A\,\cos(\omega\,t) + B\,\sin(\omega\,t) = C\,\cos(\omega\,t + \phi)$$

where the constants $A$ and $B$ have to be determined. It turns out that a computation with complex numbers is easier. Use the complex ansatz

$$y_p(t) = c\,e^{i\,\omega\,t}\,.$$

Using the ODE this leads to

$$c\,\left(-\omega^2 + 2\,i\,\omega\,\delta + \omega_0^2\right)\,e^{i\,\omega\,t} = F_0\,e^{i\,\omega\,t}$$

For the constant $c \in \mathbb{C}$ we find[4]

$$c = \frac{F_0}{-\omega^2 + 2\,i\,\omega\,\delta + \omega_0^2} = F_0\,\frac{1}{\sqrt{\left(\omega_0^2 - \omega^2\right)^2 + 4\,\omega^2\,\delta^2}}\,e^{i\,\phi} = F_0\,A\,e^{i\,\phi}$$

where

$$A = \frac{1}{\sqrt{\left(\omega_0^2 - \omega^2\right)^2 + 4\,\omega^2\,\delta^2}} \quad \text{and} \quad \tan\phi = \frac{-2\,\omega\,\delta}{\omega_0^2 - \omega^2}\,.$$

Thus the complex particular solution is given by

$$y_p(t) = F_0\,A\,e^{i\,\omega\,t + i\,\phi}\,.$$

Using the real part of this complex expression leads to the real-valued solution

$$y_p(t) = F_0\,A\,\cos(\omega\,t + \phi) = F_0\,\frac{1}{\sqrt{\left(\omega_0^2 - \omega^2\right)^2 + 4\,\omega^2\,\delta^2}}\,\cos(\omega\,t + \phi)\,.$$

Comparing with the external contribution $F_0\,\cos(\omega\,t)$ the factor $A$ is called the **amplification factor** and $\phi$ is the **phase shift**.

[4]Use $\omega^2 - \omega_0^2 + 2\,i\,\omega\,\delta = \sqrt{(\omega^2 - \omega_0^2)^2 + 4\,\omega^2\,\delta^2}\,\exp(-i\,\phi)$ with $\tan\phi = \frac{-2\,\omega\,\delta}{\omega^2 - \omega_o^2}$.

**General solution:** The homogeneous contribution $y_h(t)$ in the general solution

$$y(t) = e^{-\delta t}\ (c_1\ \cos(\omega_\delta\, t) + c_2\ \sin(\omega_\delta\, t)) + F_0\ A\ \cos(\omega\, t + \phi)$$

will converges to zero as $t \to \infty$. Thus for large times $t$ we find

$$y(t) \approx y_p(t) = F_0\ \frac{1}{\sqrt{\left(\omega_0^2 - \omega^2\right)^2 + 4\,\omega^2\,\delta^2}}\ \cos(\omega\, t + \phi)\,.$$

Now examine the behavior of this solution, as function of the angular velocity $\omega$ of the external force.

**Phase shift** $\phi$ **:** Instead of the expression $\tan\phi = \frac{-2\,\omega\,\delta}{\omega_0^2 - \omega^2}$ examine the argument of the complex number

$$\phi = \arg\ \left(\omega_0^2 - \omega^2 - 2\,i\,\omega\,\delta\right)$$

as a curve in the complex plane $\mathbb{C}$ for different values of $0 \leq \omega < \infty$. Let $0 < \varepsilon$ be a small, positive number. Then we find

| | | |
|---|---|---|
| $0 < \omega^2 = \varepsilon$ | $-1 \ll \tan\phi < 0$ | $0 > \phi \approx 0$ |
| $0 < \omega^2 < \omega_0^2$ | $\tan\phi < 0$ | $0 > \phi > -\frac{\pi}{2}$ |
| $\omega^2 = \omega_0^2 - \varepsilon$ | $\tan\phi \ll -1$ | $\phi \approx -\frac{\pi}{2}$ |
| $\omega^2 = \omega_0^2 + \varepsilon$ | $\tan\phi \gg 1$ | $\phi \approx -\frac{\pi}{2}$ |
| $\omega_0^2 < \omega^2 < \infty$ | $\infty > \tan\phi > 0$ | $-\frac{\pi}{2} < \phi < -\pi$ |

As the frequency $\omega$ moves from 0 to $\infty$, the phase shift moves from 0 to $-\pi$, i.e. $-90^\circ$.

**Amplification factor** $A$ **:** The amplification factor $A(\omega)$ is given by

$$A(\omega) = \frac{1}{\sqrt{\left(\omega_0^2 - \omega^2\right)^2 + 4\,\omega^2\,\delta^2}}\,.$$

To examine the qualitative behavior of this function first use $z = \omega^2$ and examine

$$f(z) = \left(\omega_0^2 - z\right)^2 + 4\,z\,\delta^2\,.$$

This is a parabola, opened upwards. To find the position of the minimal value set the derivative equals 0

$$\frac{d}{dz}\ f(z) = -2\ \left(\omega_0^2 - z\right) + 4\,\delta^2 = 0$$

and obtain the vertex at

$$z = z_R = \omega_0^2 - 2\,\delta^2$$

Thus the maximal value of $A(\omega)$ is attained at $\omega_R = \sqrt{\omega_0^2 - 2\,\delta^2}$. The minimal value of $f(\omega^2)$ is

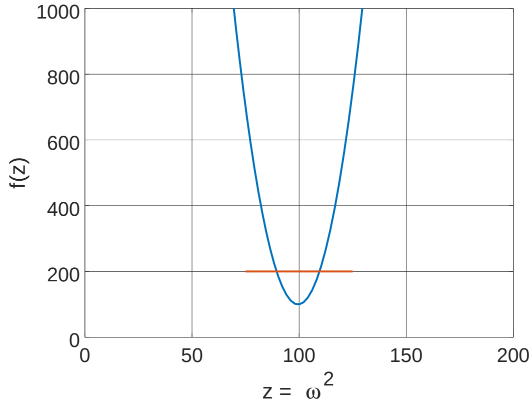
$$f(\omega_R^2) = \left(\omega_0^2 - \omega_R^2\right)^2 + 4\,\omega_R^2\,\delta^2 = \left(2\,\delta^2\right)^2 + 4\ \left(\omega_0^2 - 2\,\delta^2\right)\ \delta^2 = 4\ \left(\omega_0^2 - \delta^2\right)\ \delta^2\,.$$

In Figure 4.5(a) find the graphs of the auxiliary function $f(\omega^2)$ for the numerical example $\omega_0 = 10$ and $\delta = 0.5$. The function $A(\omega) = 1/\sqrt{f(\omega^2)}$ attains the maximal value at the **resonance frequency** $\omega_R/(2\,\pi)$

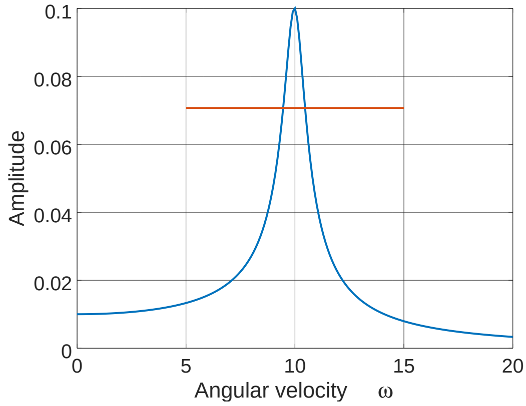
$$\omega_R^2 = \omega_0^2 - 2\,\delta^2$$

and the maximal value is

$$A(\omega_R) = \frac{1}{2\,\delta\,\sqrt{\omega_0^2 - \delta^2}}\,.$$




(a) auxiliary function $f(z) = f(\omega^2)$




(b) amplification factor $A(\omega)$

Figure 4.5: Quality factor of a resonator

Obviously we have $\lim_{\omega\to\infty} A(\omega) = 0$. In Figure 4.5(b) find the graph of the amplification factor $A(\omega)$.

The width of the resonance curve is a measure for the **quality factor** $Q$ of the resonator. The width is determined at $\frac{1}{\sqrt{2}} H \approx 0.71\, H$, where $H$ is the maximal height. Using the function $f(z)$ these values are determined by

$$\begin{aligned} 2\, f(z_R) &= f(z) = f(z_R) + (z - z_R)^2 \\ (z - z_R)^2 &= f(z_R) = 4\,\left(\omega_0^2 - \delta^2\right)\,\delta^2 \\ z - z_R &= 2\,\delta\,\sqrt{\omega_0^2 - \delta^2}\,. \end{aligned}$$

For $\delta \ll \omega_0$ find $\omega_r \approx \omega_0$ and

$$\begin{aligned} z - z_R &= \omega^2 - \omega_R^2 = (\omega - \omega_R)\,(\omega + \omega_R) \approx (\omega - \omega_0)\,2\,\omega_0 \\ \omega - \omega_0 &\approx \frac{2\,\delta\,\sqrt{\omega_0^2 - \delta^2}}{2\,\omega_0} \approx \frac{\delta\,\omega_0}{\omega_0} = \delta\,. \end{aligned}$$

Consequently in Figure 4.5(b) find the relation

$$2\,\delta = \text{ full width at 71\% of maximal height}$$

The quality factor $Q$ is defined by

$$Q = \frac{\pi\, f_R}{\delta} = \frac{\omega_R}{2\,\delta} = \frac{\omega_R}{\text{full width at height } H/\sqrt{2}}\,.$$

Thus a large value for the quality factor $Q$ corresponds to a resonator with little friction. ♢

### 4.3.3 Second order equations and phase portraits

A general second oder equation can be written in the form

$$\ddot{y}(t) = f(t, y(t), \dot{y}(t)) \tag{4.6}$$

and can be transformed to a system of first order ODEs with help of the intermediate variable $v(t) = \dot{y}(t)$,

$$\frac{d}{dt}\begin{pmatrix} y(t) \\ v(t) \end{pmatrix} = \begin{pmatrix} v(t) \\ f(t, y(t), v(t)) \end{pmatrix}$$

or with the original notation

$$\frac{d}{dt}\begin{pmatrix} y(t) \\ \dot y(t)\end{pmatrix} = \begin{pmatrix} \dot y(t) \\ f(t, y(t), \dot y(t))\end{pmatrix} . \tag{4.7}$$

This translation is necessary if using `MATLAB`/*Octave* to solve (4.6), e.g. with the help of the command `ode45()`. If the ODE (4.6) is autonomous (independent on $t$) then we can use a **phase portrait** to examine the solutions. Generate the vector field $(v, f(y, v))^T$ in the $yv$–plane, i.e. in the $y\dot y$–plane.

• **4–16 Example:** The linear ODE

$$\ddot y(t) = -y(t) - 0.2\,\dot y(t)$$

can be solved with the method from the previous section. The corresponding vector field for the phase portrait is

$$\begin{pmatrix} v \\ -y - 0.2\cdot v\end{pmatrix}$$

and is shown in Figure 4.6(a), together with the two solutions starting at $(y(0), \dot y(0)) = (0.4\,,\,0)$ and $(y(0), \dot y(0)) = (3\,,\,0)$. Observe that the vector field is tangential to the solution curves, which is the condition required by the differential equation. It is clearly visible that both solutions converge to zero.

Observe that the independent variable (time $t$) is not visible in the phase portraits 4.6, while in the vector field for first order equation, e.g. Figure 4.1, the value of the independent variable is on the horizontal axis. ♢

• **4–17 Example:** Consider the nonlinear ordinary differential equation

$$\ddot y(t) = -y(t) - 0.2\cdot(|y(t)| - 1)\,\dot y(t)\,.$$

This ODE can **not** be solved with the methods from the previous section. The contribution $-0.2\cdot(|y(t)| - 1)\cdot\dot y$ describes a friction term if $|y| > 1$ and a 'negative' friction if $|y| < 1$. The resulting vector field for the phase portrait is

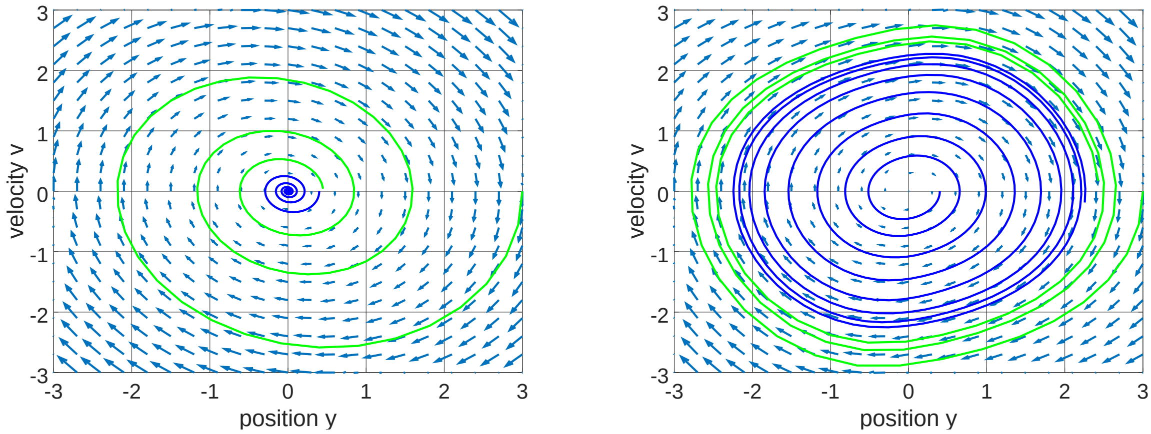
$$\begin{pmatrix} v \\ -y - 0.2\,(|y| - 1)\,v\end{pmatrix}$$

and is shown in Figure 4.6(b), together with the two solutions starting at $(y(0), \dot y(0)) = (0.4\,,\,0)$ and $(y(0), \dot y(0)) = (3\,,\,0)$. Here both solutions tend towards a periodic orbit. The solution starting with a small initial value (in blue) is a spiral with increasing radius and the solution with a large initial value (in green) is a spiral with shrinking radius. The periodic orbit is trapped between the two shown solutions.

To generate numerical approximations to the solutions we use the command `ode45()` from *Octave*/`MATLAB`. The code below generates Figure 4.6(b), the code for Figure 4.6(a) is very similar.

**demo_phaseportrait.m**

```
y = linspace(-3,3,21); v = linspace(-3,3,21);
[y,v] = meshgrid(y,v);  % generate the grid for the vector field
F1 = v;  F2 = -y -0.2*(abs(y)-1).*v; % compute the vector field
figure(1);
quiver(y,v,F1,F2)                    % display the vector field
xlabel('position y'); ylabel('velocity v'); axis(3*[-1 1 -1 1])
% compute approximations to the solutions
```




(a) with constant damping factor $-0.2$

(b) with nonlinear damping factor $-0.2\,(|y|-1)$

Figure 4.6: Phase portrait for second order ODEs

```
[t1,yv1] = ode45(@(t,yv)[yv(2);-yv(1)-0.2*(abs(yv(1))-1)*yv(2)],...
                 linspace(0,6*pi),[3,0]);
[t2,yv2] = ode45(@(t,yv)[yv(2);-yv(1)-0.2*(abs(yv(1))-1)*yv(2)],...
                 linspace(0,16*pi,500),[0.4,0]);
hold on     % display in the same figure as the vector field
plot(yv1(:,1),yv1(:,2),'g',yv2(:,1),yv2(:,2), 'b')
hold off
```

◊

# Chapter 5

# Partial Differential Equations

## 5.1 Keywords and Literature

The goal of this section is to recognize different types of partial differential equations and identify their field of applications. You should become familiar with the typical behavior of the solutions. These notes can not provide a complete presentations and the emphasis is clearly not on solving the equations. It is **not expected** that the students can apply the few solution algorithms presented, but the should know features of the solutions.

Some of the statements in this section are similar to the presentation in [Habe04], [ONei11] and [Farl82] states many of the results in a form readable by engineers. In [TikhSama63] more details are shown. This this section we examine three different type of PDEs: elliptic, parabolic and hyperbolic:

- Typical examples of **elliptic** equations are

$$\begin{aligned} -\frac{\partial^2}{\partial x^2}\,u(x) &= f(x)\,, \\ -(\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}) &= f(x,y)\,. \end{aligned}$$

  Typical applications are steady state heat equations. The name elliptic appears since level curves of $x^2+y^2$ in the $xy$ plane are ellipses.

- Typical examples of **parabolic** equations are

$$\begin{aligned} \frac{\partial\,u(t,x)}{\partial t}-\frac{\partial^2\,u(t,x)}{\partial x^2} &= f(x)\,, \\ \frac{\partial\,u}{\partial t}-(\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}) &= f(x,y)\,. \end{aligned}$$

  Typical applications are dynamic heat equations. The name parabolic appears since level curves of $t-x^2$ in the $tx$ plane are parabolas.

- Typical examples of **hyperbolic** equations are

$$\begin{aligned} \frac{\partial^2\,u(t,x)}{\partial t^2}-\frac{\partial^2\,u(t,x)}{\partial x^2} &= f(x)\,, \\ \frac{\partial^2 u}{\partial t^2}-(\frac{\partial^2 u}{\partial x^2}+\frac{\partial^2 u}{\partial y^2}) &= f\,. \end{aligned}$$

  Typical applications are wave equations or vibrating strings and membranes. The name hyperbolic appears since level curves of $t^2-x^2$ in the $tx$ plane are hyperbolas.

There are many more partial differential equations, some not of the above types. Some examples are Navier-Stokes (fluid dynamics), Korteweg-De Vries (solitons), von Karman (vibrating plates), ... The literature on PDEs is enormous. In the booklet [Stew13] find a non-mathematical description of a few very important equations.

## 5.2 Steady State Heat Equations, Elliptic Problems

### 5.2.1 Different forms of elliptic equations

The simplest form of an elliptic problem is the Laplace equation, i.e. for a given function $f(x,y)$ find a function $u(x,y)$ such that

$$-\Delta u = -\nabla\nabla u = -\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2}\right) = f\,.$$

The general form is given by

$$-\left(a_{11}\,\frac{\partial^2 u}{\partial x^2} + a_{22}\,\frac{\partial^2 u}{\partial y^2} + 2\,a_{12}\,\frac{\partial^2 u}{\partial x\,\partial y}\right) + b_1\,\frac{\partial\, u}{\partial x} + b_2\,\frac{\partial\, u}{\partial y} = f$$

with the conditions[1] $a_{11} > 0$, $a_{22} > 0$ and $a_{12}^2 < a_{11}\,a_{22}$. Equations of this type are called **elliptic** since the level curves of the quadratic expressions

$$a_{11}\,x^2 + a_{22}\,y^2 + 2\,a_{12}\,x\,y \;=\; \langle \begin{bmatrix} a_{11} & a_{12} \\ a_{12} & a_{22} \end{bmatrix} \begin{pmatrix} x \\ y \end{pmatrix} , \begin{pmatrix} x \\ y \end{pmatrix} \rangle$$

are ellipses in the plane $\mathbb{R}^2$.

If we use cylindrical coodinates $(\rho,\phi,z)$ in the space $\mathbb{R}^3$ (see page 22), then the Laplace equation is given by

$$-\Delta u = -\left(\frac{\partial^2 u}{\partial \rho^2} + \frac{1}{\rho}\,\frac{\partial\, u}{\partial \rho} + \frac{1}{\rho^2}\,\frac{\partial^2 u}{\partial \phi^2} + \frac{\partial^2 u}{\partial z^2}\right) = f$$

or often more conveniently

$$-\left(\rho\,\frac{\partial}{\partial \rho}(\rho\,\frac{\partial\, u}{\partial \rho}) + \frac{\partial^2 u}{\partial \phi^2} + \rho^2\,\frac{\partial\, u}{\partial z^2}\right) = \rho^2\,f\,. \tag{5.1}$$

If we use spherical coodinates $(r,\phi,\theta)$ in space $\mathbb{R}^3$ (see page 22), then the Laplace equation is given by

$$-\Delta u = -\left(\frac{\partial^2 u}{\partial r^2} + \frac{2}{r}\,\frac{\partial\, u}{\partial r} + \frac{1}{r^2\,\sin^2\theta}\,\frac{\partial^2 u}{\partial \phi^2} + \frac{\cos\theta}{r^2\,\sin\theta}\,\frac{\partial\, u}{\partial \theta} + \frac{1}{r^2}\,\frac{\partial^2 u}{\partial \theta^2}\right) = f$$

or often more conveniently

$$-\left(\frac{\partial}{\partial r}(r^2\,\frac{\partial\, u}{\partial r}) + \frac{1}{\sin^2\theta}\,\frac{\partial^2 u}{\partial \phi^2} + \frac{1}{\sin\theta}\,\frac{\partial}{\partial \theta}\,(\sin\theta\,\frac{\partial\, u}{\partial \theta})\right) = r^2\,f\,. \tag{5.2}$$

Observe that only in Cartesian coordinates the Laplace operator is a differential operator with constant coefficients. Thus many computations are easier in Cartesian coordinates.

[1] These conditions imply that the eigenvalues of a matrix are positive.

$$\begin{aligned} 0 &= \det \begin{bmatrix} a_{11}-\lambda & a_{12} \\ a_{12} & a_{22}-\lambda \end{bmatrix} = \lambda^2 - (a_{11}+a_{22})\,\lambda + a_{11}\,a_{22} - a_{12}^2 \\ \lambda_{1,2} &= \frac{1}{2}\,(a_{11}+a_{22} \pm \sqrt{(a_{11}+a_{22})^2 - 4\,(a_{11}\,a_{22}-a_{12}^2)} > 0 \end{aligned}$$

For the case of constant coefficients $a_{ij}$ one could use a rotated and rescaled coordinate system and in the new system the transformed differential operator is $\frac{\partial^2}{\partial (x')^2} + \frac{\partial^2}{\partial (y')^2}$, i.e. the Laplace operator.

### 5.2.2 Typical applications and boundary conditions

One of the most common applications of this type of PDE are diffusion problems. If $u(x,y)$ is the temperature of a liquid moving along with a velocity field $\vec{v}$ then there are two effects leading to the movment of energy:

1. Fourier's law for diffusion. The heat flux (unit $\frac{\mathrm{J}}{\mathrm{s\,m^2}}$) is given by $\vec{q} = -a\,\nabla u = -a\ \mathrm{grad}\, u$.
2. The energy flux generated by the liquid moving with velocity $\vec{v}$ is given by $b\,u\,\vec{v}$.

Then use conservation of energy to find the correct PDE. In this setup the given function $f$ describes the external source of energy, e.g. by heating and $a$, $b$ is a material constant, based on the specific heat, the conductivity and the density of the material. For a bounded domain $\Omega \in \mathbb{R}^3$ use the divergence theorem 3–36 to find

$$\begin{aligned}
\text{supply by external sources} &= \text{loss through boundary by diffusion and convection} \\
\iiint\limits_{\Omega} f\, dV &= \oiint\limits_{\partial\Omega} \langle \vec{q}\,,\, \vec{n}\rangle + \langle b\,u\,\vec{v}\,,\, \vec{n}\rangle\; dA \\
&= \oiint\limits_{\partial\Omega} \langle -a\,\nabla u,\, \vec{n}\rangle + \langle b\,u\,\vec{v}\,,\, \vec{n}\rangle\; dA \;=\; \iiint\limits_{\Omega} -\nabla\,(a\,\nabla u) + \nabla(b\,u\,\vec{v})\; dV
\end{aligned}$$

Since this has to be correct for all possible domains $\Omega$ we conclude

$$-\nabla\,(a\,\nabla u) + \nabla(b\,u\,\vec{v}) = f$$

If $a$ and $b$ are constant this simplifies to

$$-a\left(\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2}\right) + b\left(\frac{\partial\,(u\,v_1)}{\partial x} + \frac{\partial\,(u\,v_2)}{\partial y} + \frac{\partial\,(u\,v_3)}{\partial z}\right) = f$$

and for the one-dimensional case we obtain the ordinary differential equation

$$-a\,u'' + b\,v\,u' + b\,v'\,u = f\,.$$

For the above PDE to have a unique solution one has to specify boundary conditions on all parts of the boundary. There are different types of conditions and at each point only one condition should be specified.

| | mathematical name, | physical interpretation |
|---|---|---|
| $u = 0$ | Dirichlet condition, | prescribed zero value |
| $u = g$ | Dirichlet condition, | prescribed value |
| $a\,\frac{\partial u}{\partial \vec{n}} = 0$ | Neuman condition, | prescribed zero flux |
| $a\,\frac{\partial u}{\partial \vec{n}} = g$ | Neuman condition, | prescribed flux |
| $a\,\frac{\partial u}{\partial \vec{n}} = c\,u$ | Robin condition, | flux proportional to value |

### 5.2.3 Applications of elliptic PDEs

In the previous section we use the problem of steady state heat conduction as a typical example leading to an elliptic PDE. There are many other applications leading to the same type of equation. For one dimensional problems the equation takes the form

$$-\frac{d}{dx}\left(a(x)\,\frac{d\,u(x)}{dx}\right) = f(x)\,.$$

| differential equation | problem description | | | constitutive law |
|---|---|---|---|---|
| | | $T =$ | temperature | |
| $\frac{d}{dx}\left(A\,k\frac{dT}{dx}\right)+Q=0$ | one–dimensional | $A =$ | area | Fourier's law |
| | heat flow | $k =$ | thermal conductivity | $q=-k\,\frac{dT}{dx}$ |
| | | $Q =$ | heat supply | |
| | | $u =$ | displacement | Hooke's law |
| $\frac{d}{dx}\left(A\,E\frac{du}{dx}\right)+b=0$ | axially loaded | $A =$ | area | $\sigma=E\,\frac{du}{dx}$ |
| | elastic bar | $E =$ | Young's modulus | $\sigma=$ stress |
| | | $b =$ | axial loading | |
| | transversely loaded | $w =$ | deflection | |
| $\frac{d}{dx}\left(S\frac{dw}{dx}\right)+p=0$ | flexible string | $S =$ | string force | |
| | | $p =$ | lateral loading | |
| | | $c =$ | concentration | Fick's law |
| $\frac{d}{dx}\left(A\,D\frac{dc}{dx}\right)+Q=0$ | one dimensional | $A =$ | area | $q=-D\,\frac{dc}{dx}$ |
| | diffusion | $D =$ | Diffusion coefficient | $q=$ flux |
| | | $Q =$ | external supply | |
| | | $V =$ | voltage | Ohm's law |
| $\frac{d}{dx}\left(A\,\gamma\,\frac{dV}{dx}\right)+Q=0$ | one dimensional | $A =$ | area | $q=-\gamma\,\frac{dV}{dx}$ |
| | electric current | $\gamma =$ | electric conductivity | $q=$ charge flux |
| | | $Q =$ | charge supply | |
| | | $p =$ | pressure | |
| | laminar flow | $A =$ | area | |
| $\frac{d}{dx}\left(A\,\frac{D^2}{32\mu}\,\frac{dp}{dx}\right)+Q=0$ | in a pipe | $D =$ | diameter | $q=\frac{D^2}{32\mu}\,\frac{dp}{dx}$ |
| | (Poisseuille flow) | $\mu =$ | viscosity | $q=$ volume flux |
| | | $Q =$ | fluid supply | |

Table 5.1: Applications of the differential equation $-\frac{\partial}{\partial x}\left(a\,\frac{\partial u}{\partial x}\right)=f$

Table 5.1 (Source: [OttoPete92, p. 63]) shows a few applications related to this type of equation. If there is no diffusion ($\vec{v} = \vec{0}$) then an elliptic problem takes to form of Poisson's equation.

$$-\nabla(a\,\nabla u) = f \quad \text{or} \quad -\operatorname{div}(a\,\operatorname{grad} u) = f\,.$$

A list of typical applications of elliptic equations of second order is shown in Table 5.2 (Source: [Redd84]). This table clearly illustrates the importance of the above type of problem.

### 5.2.4 Properties of solutions of elliptic differential equations

#### Maximum principle

If the external function $f \geq 0$ is positive, i.e. heating only, then the solution $u$ of

$$-\Delta u + \vec{b}\cdot\nabla u = f \qquad \text{in } \Omega$$

can not have a global or local minimum inside the domain $\Omega$. At a minimal point the first order derivatives have to vanish and the second order derivatives have to be positive. Thus the expression on the left in the PDE is negative, which contradicts $f > 0$.

If on the boundary of the domain we require $u = 0$ and inside $f > 0$, then the solution $u$ has to be positive inside the domain. This should coincide with your intuition of the behavior of a steady state temperature distribution.

Similarly on can verify that for an external cooling ($f \leq 0$) the solution $u$ can not have a global or local maximum inside the domain $\Omega$. Results of this type are known as **maximum principles** in the literature.

#### Calculus of variations

Using the Calculus of Variations one can show that minimizers of the functional

$$F(u) = \iiint\limits_{\Omega} \frac{1}{2}\,a\,\|\nabla u\|^2 - f\cdot u\;dV$$

are solutions of the elliptic problem

$$-\nabla\,(a\,\nabla u) = f\,.$$

More details are presented in the class "Numerical Methods", see the lecture notes [Stah08].

#### No net input of energy allowed

If $u(\vec{x})$ is a solution of the steady state heat equation with Neumann boundary condition (no flux)

$$\Delta u(\vec{x}) = f(\vec{x}) \text{ for } \vec{x}\in\Omega \text{ and } \tfrac{\partial}{\partial\vec{n}}\,u(\vec{x}) = 0 \text{ on } \partial\Omega$$

then we have the necessary condition

$$\iiint\limits_{\Omega} f(\vec{x})\;dV = 0\,,$$

i.e. there is no net input of energy by the external source. This is a consequence of the computations in Example 3–41 on page 30. If there were a net input of energy there could be no steady state solution, but the temperature would have to rise. If the above no flux boundary condition is replaced by the Dirichlet condition $u(\vec{x}) = 0$ then the condition is not necessary. In this case the energy put into the domain by $f(\vec{x})$ can leave the domain through the boundary.

| Field of application | Primary variable | Material constant | Source variable | Secondary variables |
|---|---|---|---|---|
| General situation | $u$ | $a$ | $f$ | $\frac{\partial u}{\partial x}$, $\frac{\partial u}{\partial y}$ |
| Heat transfer | Temperature $T$ | Conductivity $k$ | Heat source $Q$ | Heat flow density $\vec{q}$<br>$\vec{q} = -k \nabla T$ |
| Diffusion | Concentration $c$ | Diffusion coefficient | external supply $Q$ | flux $\vec{q} = -D\,\nabla c$ |
| Electrostatics | Scalar potential $\Phi$ | Dielectric constant $\varepsilon$ | Charge density $\rho$ | Electric flux density $D$ |
| Magnetostatics | Magnetic potential $\Phi$ | Permeability $\nu$ | | Magnetic flux density $B$ |
| Transverse deflection of elastic membrane | Transverse deflection $u$ | Tension of membrane $T$ | Transversely distributed load | Normal force $q$ |
| Torsion of a bar | Warping function $\phi$ | 1 | 0 | Stress $\tau$<br>$\tau_{xz} = \frac{E\,\alpha}{2\,(1+\nu)}\,(-y + \frac{\partial\,\phi}{\partial x})$<br>$\tau_{yz} = \frac{E\,\alpha}{2\,(1+\nu)}\,(x + \frac{\partial\,\phi}{\partial y})$ |
| Irrotational flow of an ideal fluid | Stream function $\Psi$<br>Velocity potential $\Phi$ | Density $\rho$ | Mass production $\sigma$ (usually zero) | Velocity $(u, v)^T$<br>$\frac{d\Psi}{\partial x} = -u$<br>$\frac{d\Psi}{\partial y} = v$<br>$\frac{d\Phi}{\partial x} = u$<br>$\frac{d\Phi}{\partial y} = v$ |
| Ground-water flow | Piezometric head $\Phi$ | Permeability $K$ | Recharge $Q$ (or pumping $-Q$) | seepage $q = K\,\frac{\partial\Phi}{\partial n}$<br>velocities<br>$u = -K\,\frac{d\Phi}{\partial x}$<br>$v = -K\,\frac{d\Psi}{\partial y}$ |

Table 5.2: Some examples of Poisson's equation $-\nabla\,(a\,\nabla u) = f$

### 5.2.5 Fourier approximation for the solution of a steady state heat equation on an interval

The typical example for a one dimensional steady state heat equation is

$$-\frac{d^2}{dx^2}\,u(x) = f(x) \quad \text{for } 0 \le x \le L \text{ with } u(0) = u(L) = 0 \quad . \tag{5.3}$$

It is straightforward to verify that the functions $\phi_n(x) = \sin(n\,\frac{\pi}{L}\,x)$ satisfy the boundary conditions and are eigenfunctions with

$$-\frac{d^2}{dx^2}\,\phi_n(x) = (\frac{n\,\pi}{L})^2\,\phi_n(x) \;\; = \lambda_n\,\phi_n(x)\,.$$

The eigenfunction $\phi_n(x)$ has $n$ extrema between 0 and $L$ and for $n$ large it displays very rapid oscillations. Using a Fourier sine series of the given function $f(x)$

$$f(x) = \sum_{n=1}^{\infty} b_n\,\sin(n\,\frac{\pi}{L}\,x)$$

we can write down a solution formula for the static heat equation (5.3)

$$u(x) = \sum_{n=1}^{\infty} \frac{b_n}{\lambda_n}\,\sin(n\,\frac{\pi}{L}\,x) = \sum_{n=1}^{\infty} b_n\,(\frac{L}{\pi\,n})^2\,\sin(n\,\frac{\pi}{L}\,x)\,.$$

Observe that the amplitude of contributions with rapid oscillations is divided by $n^2$ and thus will be very small. This corresponds to the fact that a solution $u(x)$ of static heat equations will typically not show rapid oscillations, even of the heating function $f(x)$ has rapid oscillations. As an example we consider the heating function

$$f(x) = \begin{cases} +1 & \text{if } 0.2 < x < 0.5 \\ -1 & \text{if } 0.6 < x < 0.7 \\ 0 & \text{otherwise} \end{cases} \quad .$$

Find the Fourier sine approximation of this function $f(x)$ by 15 terms in Figure 5.1(a). The resulting solution $u(x)$ is shown in Figure 5.1(b). The first 15 contributions of the Fourier approximation were used.

### 5.2.6 Fourier approximation for the solution of a steady state heat equation on a rectangle

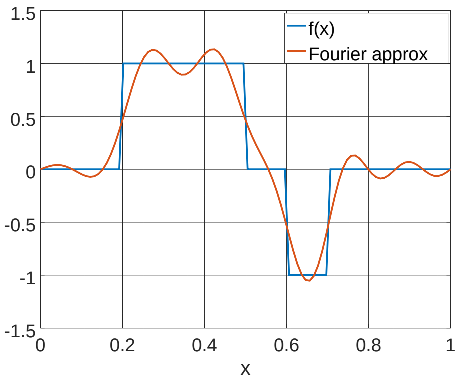
The typical example for a one dimensional steady state heat equation is

$$\begin{aligned} -\Delta u(x,y) &= f(x,y) && \text{for } 0 < x < L \text{ and } 0 < y < W \\ u(x,y) &= 0 && \text{on the boundary of the rectangle} \end{aligned} \quad . \tag{5.4}$$

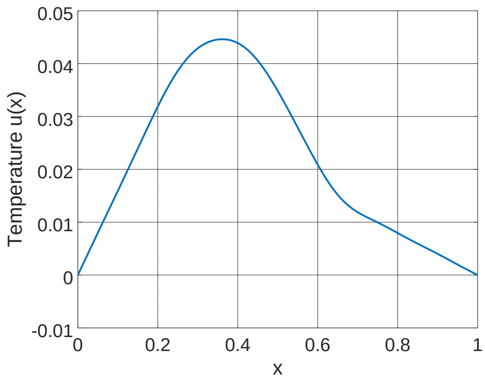
Verify that the functions

$$\begin{aligned} \phi_{n,m}(x,y) &= \sin(n\,\frac{\pi}{L}\,x)\,\sin(m\,\frac{\pi}{W}\,y) \\ -\Delta\phi_{n,m} &= \left((\frac{n\,\pi}{L})^2 + (\frac{m\,\pi}{W})^2\right)\,\phi_{n,m} \;=\; \lambda_{n,m}\,\phi_{n,m} \end{aligned}$$

are eigenfunctions and satisfy the Dirichlet boundary condition. The eigenfunction $\phi_{n,m}(x,y)$ has $n$ extrema in $x$ direction and $m$ extrema in $y$ direction.




(a) Fourier approximation of $f(x)$




(b) solution of the steady state heat equation

Figure 5.1: Fourier approximation for the solution of a steady state heat equation

Using a double Fourier sine series of the given function $f(x,y)$

$$f(x) = \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} b_{n,m}\,\phi_{n,m}(x,y)$$

we can write down a solution formula for the static heat equation (5.4)

$$u(x) = \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} \frac{b_{n,m}}{\lambda_{n,m}}\,\phi_{n,m}(x,y) = \sum_{m=1}^{\infty}\sum_{n=1}^{\infty} \frac{b_{n,m}}{(\frac{n\,\pi}{L})^2 + (\frac{m\,\pi}{W})^2}\ \sin(n\,\frac{\pi}{L}\,x)\ \sin(m\,\frac{\pi}{W}\,y)\,.$$

Observe that the amplitude of contributions with rapid oscillations are divided by large factors and thus will be very small.

### 5.2.7 Fourier approximation on other domains

The above approach does formally work on many types of domains, but the major problem is to find the eigenfunctions such that

$$-\Delta\phi_n(\vec{x}) = \lambda_n\,\phi_n(\vec{x}) \text{ for } \vec{x}\in\Omega \text{ and } \phi_n(\vec{x}) = 0 \text{ on } \partial\Omega\,.$$

On disks the eigenfunctions can be expressed in terms of Bessel functions.

### 5.2.8 Fundamental solution on the unbounded domain $\mathbb{R}^3$

When examining the steady state heat equation on the whole space $\mathbb{R}^3$ the boundary conditions have to be replaced by a decay condition

$$\lim u(\vec{x}) = 0 \quad \text{as} \quad \|\vec{x}\| \to \infty\,.$$

For a given function $f(\vec{x})$ the heat equation

$$-\Delta u(\vec{x}) = f(\vec{x})$$

then has a solution given by

$$u(\vec{x}) = \iiint\limits_{\mathbb{R}^3} \frac{1}{4\,\pi\,\|\vec{x}-\vec{y}\|}\,f(\vec{y})\,dV_y\,. \tag{5.5}$$

The function

$$\Phi(\vec{x}) = \frac{1}{4\,\pi\,\|\vec{x}\|}$$

is called a **fundamental solution** or a **Green's function** for the steady state solution on $\mathbb{R}^3$. It is the solution of the heat equation with a single heat source of strength 1 at the origin. Then the superposition principle is used to derive the formula (5.5).

If the heat source $f(\vec{x})$ is only positive for $\vec{x}$ close to the origin, then (5.5) implies that the temperature $u(\vec{x})$ converges to 0 like $\frac{1}{\|\vec{x}\|}$ as $\|\vec{x}\|$ tends to $+\infty$.

## 5.3 Dynamic Heat Equations, Parabolic Problems

### 5.3.1 Different forms of parabolic equations

A parabolic PDE always contains a first order derivative with respect to time $t$, i.e. we examine dynamic problems. The spatial derivatives are given by an elliptic differential operator. Find a few typical examples below.

$$\begin{aligned}
\frac{\partial}{\partial t}\,u(t,x) &= \frac{\partial^2}{\partial x^2}\,u(t,x) + f(t,x) \\
\frac{\partial}{\partial t}\,u(t,x,y) &= \frac{\partial^2}{\partial x^2}\,u(t,x,y) + \frac{\partial^2}{\partial y^2}\,u(t,x,y) + f(t,x,y) \\
\frac{\partial}{\partial t}\,u(t,\vec{x}) &= \nabla(a(\vec{x})\,\nabla\,u(t,\vec{x})) + f(t,\vec{x}) \\
\frac{\partial}{\partial t}\,u(t,\vec{x}) &= \nabla(a(\vec{x})\,\nabla\,u(t,\vec{x})) + \nabla(u(t,\vec{x})\cdot\vec{v}(\vec{x})) + f(t,\vec{x})
\end{aligned}$$

The last example is the most general one, with given functions $a(\vec{x}) > 0$ and $f(t,\vec{x})$. The term $u\cdot\vec{v}$ describes a convection contribution. To obtain unique solutions for parabolic equations one has to specify boundary conditions for $\vec{x}\in\partial\Omega$ for all times $t > 0$ and an initial condition $u(0,\vec{x}) = u_0(\vec{x})$ for all $\vec{x}\in\Omega$.

### 5.3.2 Applications of parabolic equations

There are two very important applications of the general parabolic equation

$$\frac{\partial}{\partial t}\,u(t,\vec{x}) = \nabla(a(\vec{x})\,\nabla\,u(t,\vec{x})) + \vec{v}\cdot\nabla u(t,\vec{x}) + f(t,\vec{x})\,.$$

- The dynamic heat equation: The basic laws used are conservation of energy and Fourier's law of heat conduction.

$$\begin{aligned}
u &= \text{temperature at time } t \text{ and position } \vec{x} \\
a &= \text{material constant for thermal diffusion, using specific heat, conductivity, ...} \\
-\nabla u &= \text{proportional to the thermal energy flux} \\
f &= \text{external heating or cooling contribution} \\
\vec{v} &= \text{velocity field, if there is also convection}
\end{aligned}$$

- The diffusion equation: The basic laws used are a conservation law for the material examined and Fick's law for diffusion.

$$\begin{aligned}
u &= \text{concentration at time } t \text{ and position } \vec{x} \\
a &= \text{material constant for diffusion} \\
-\nabla u &= \text{proportional to the material flux} \\
f &= \text{external source or sink of material} \\
\vec{v} &= \text{velocity field, if there is also convection}
\end{aligned}$$

### 5.3.3 Properties of solutions of the dynamic heat equation

#### A maximum principle

Let $u(t,\vec{x})$ be a solution of the dynamic heat equation

$$\frac{\partial}{\partial t}\,u(t,\vec{x}) = \nabla(a(\vec{x})\,\nabla u(t,\vec{x})) \quad \text{for } 0 \leq t \leq T \text{ and } \vec{x} \in \Omega \in \mathbb{R}^3$$

with a positive coefficient function $a(\vec{x})$, then $u$ attains its maximal value either at time $t = 0$ or for a time $t > 0$ at an point $\vec{x} \in \partial\Omega$ on the boundary of the domain $\Omega \subset \mathbb{R}^3$.

#### Decaying expressions

Let $u(t,\vec{x})$ be a solution of the dynamic heat equation

$$\frac{\partial}{\partial t}\,u(t,\vec{x}) = \nabla(a(\vec{x})\,\nabla u(t,\vec{x})) \quad \text{for } t > 0 \text{ and } \vec{x} \in \Omega \in \mathbb{R}^3$$

with a positive coefficient function $a(\vec{x})$ and on the boundary $\partial\Omega$ we have either $u(t,\vec{x}) = 0$ or $\frac{\partial\,u(t,\vec{x})}{\partial\vec{n}} = 0$. Then Result 3–45 (page 32) implies

$$\begin{aligned}
\frac{d}{dt}\iint\limits_{\Omega} a(\vec{x})\,\|\nabla u(t,x)\|^2\,dV &= -2\iiint\limits_{\Omega} |\dot{u}(t,x)|^2\,dV + 0 \leq 0 \\
\frac{d}{dt}\iiint\limits_{\Omega} |u(t,\vec{x})|^2\,dV &= -2\iiint\limits_{\Omega} a(\vec{x})\,\|\nabla u(t,\vec{x})\|^2\,dV \leq 0\,.
\end{aligned}$$

Thus without an external heating term $f(t,\vec{x})$ the solution of the dynamic heat equation will make the integrals of $|u|^2$ and $\|\nabla u\|^2$ smaller as time $t$ advances. Both inequalities are consistent with a solution $u(t,\vec{x})$ converging towards a constant temperature as time $t$ advances.

### 5.3.4 Fourier approximation for the solution of a dynamic heat equation on an interval

The typical example for a one dimensional dynamic heat equation is

$$\begin{array}{rcll}
\frac{\partial}{\partial t}\,u(t,x) &=& \frac{\partial^2}{\partial x^2}\,u(t,x) + f(x) & \text{for } 0 \leq x \leq L \text{ and } t \geq 0 \\
u(t,0) = u(t,L) &=& 0 & \text{for } t \geq 0 \\
u(0,x) &=& u_0(x) & \text{for } 0 < x < L
\end{array} \quad . \tag{5.6}$$

Using the same eigenfunctions as for the static problem (5.3) $\phi_n(x) = \sin(n\,\frac{\pi}{L}\,x)$

$$-\frac{\partial^2}{\partial x^2}\,\phi_n(x) = (\frac{n\,\pi}{L})^2\,\phi_n(x) \;= \lambda_n\,\phi_n(x)$$

we assume that the heating term $f$ and the initial temperature are given by their Fourier series.

$$f(x) = \sum_{n=0}^{\infty} f_n\,\sin(n\,\frac{\pi}{L}\,x) \quad \text{and} \quad u_0(x) = \sum_{n=0}^{\infty} b_n\,\sin(n\,\frac{\pi}{L}\,x)$$

and we seek the time dependent Fourier series of the solution $u(t,x)$, i.e. functions $u_n(t)$ such that

$$u(t,x) = \sum_{n=0}^{\infty} u_n(t)\,\sin(n\,\frac{\pi}{L}\,x)\,.$$

Using this ansatz in the dynamic heat equation (5.6) leads to the ODEs

$$\frac{d}{dt}\,u_n(t) = -\lambda_n\,u_n(t) + f_n \quad \text{with} \quad u_n(0) = b_n\,.$$

The solution of this ODE is (see page 39)

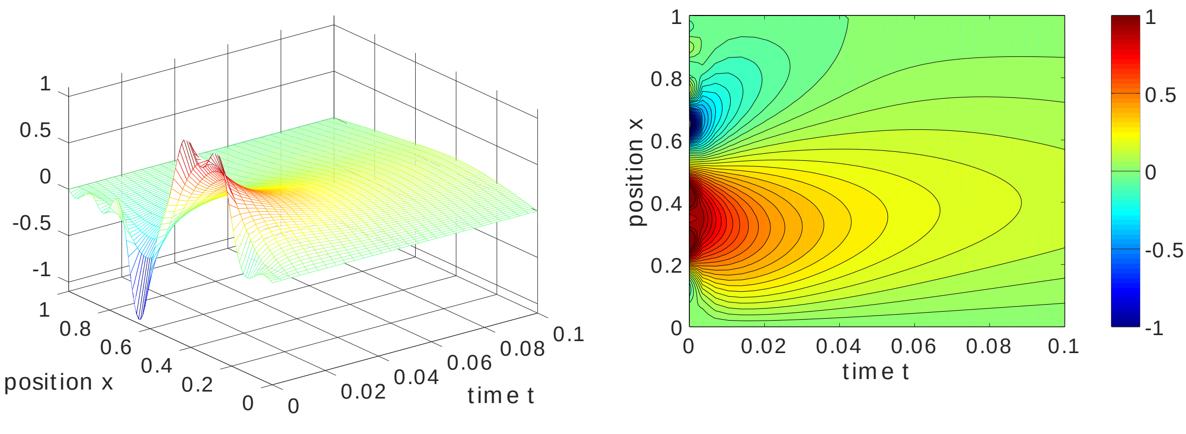
$$u_n(t) = \frac{f_n}{\lambda_n} + (b_n - \frac{f_n}{\lambda_n})\,e^{-\lambda_n\,t}\,.$$

As a consequence we find a formula for solutions of (5.6).

$$\begin{aligned}
u(t,x) &= \sum_{n=1}^{\infty} u_n(t)\,\sin(n\,\frac{\pi}{L}\,x) \\
&= \sum_{n=1}^{\infty} \frac{f_n}{\lambda_n}\,\sin(n\,\frac{\pi}{L}\,x) + \sum_{n=1}^{\infty}(b_n - \frac{f_n}{\lambda_n})\,\sin(n\,\frac{\pi}{L}\,x)\,e^{-\lambda_n\,t} \\
&= \sum_{n=1}^{\infty} \frac{f_n\,L^2}{n^2\,\pi^2}\,\sin(n\,\frac{\pi}{L}\,x) + \sum_{n=1}^{\infty}(b_n - \frac{f_n\,L^2}{n^2\,\pi^2})\,\sin(n\,\frac{\pi}{L}\,x)\,e^{-n^2\,\frac{\pi^2}{L^2}\,t}\,.
\end{aligned}$$

Compare this solution with the solution for the corresponding static heat equation (5.3).

- At $t$ converges to $+\infty$ the second series in the above solution converges to zero. Thus the solution $u(t,x)$ converges to the solution of the steady state problem $0 = u'' + f$, examined in Section 5.2.5 on page 57. The contribution converging slowest to zero has a factor of $\exp(-\frac{\pi^2}{L^2}\,t)$ and thus decays with a time constant of $\frac{L^2}{\pi^2}$.
- Observe that the amplitude of contributions with rapid oscillations converge to zero very rapidly.




(a) surface of $u(t,x)$

(b) contour levels for $u(t,x)$

Figure 5.2: Fourier approximation for the solution of a dynamic heat equation

As an example examine the problem (5.6) with $u_0(x) = f(x)$ given by the function in Figure 5.1(a). The first 15 contributions of the Fourier approximation were used. Find the surface $u(t,x)$ and its level curves in Figure 5.2.

- In Figure 5.2(a) observe that the rapid, spatial oscillations at $t = 0$ become small rather quickly for $t > 0$.

- At the right edge $t = 0.1$ The solution $u(0.1\,,\,x)$ is already rather close to the solution of the static problem, shown in Figure 5.1(b).

Very similar computations allow to write down a solution formula for the dynamic heat equations on a rectangle with Dirichlet boundary conditions.

### 5.3.5 Fundamental solutions on unbounded domains, Green's functions

To find a solution of the dynamic heat equation

$$\begin{aligned} \dot u(t,x) &= \alpha\, u''(t,x) && \text{for } -\infty < x < +\infty \text{ and } t > 0 \\ u(0,x) &= u_0(x) && \text{for } -\infty < x < +\infty \end{aligned} \tag{5.7}$$

one can use the **fundamental solution**

$$\Phi(t,x) = \frac{1}{\sqrt{4\,\alpha\,\pi\,t}}\,\exp(-\frac{x^2}{4\,\alpha\,t})\,. \tag{5.8}$$

Use $\frac{\partial}{\partial t}\,t^{-1/2} = \frac{-1}{2}\,t^{-3/2} = \frac{-1}{2\,t}\,t^{-1/2}$ and the product and chain rule to determine (for $t > 0$)

$$\begin{aligned} \frac{\partial}{\partial t}\,\Phi(t,x) &= \frac{-1}{2\,t}\,\Phi(t,x) + \frac{x^2}{4\,\alpha\,t^2}\,\Phi(t,x) = \left(\frac{-1}{2\,t} + \frac{x^2}{4\,\alpha\,t^2}\right)\Phi(t,x) \\ \frac{\partial}{\partial x}\,\Phi(t,x) &= \frac{-2\,x}{4\,\alpha\,t}\,\Phi(t,x) \\ \frac{\partial^2}{\partial x^2}\,\Phi(t,x) &= \frac{+4\,x^2}{(4\,\alpha\,t)^2}\,\Phi(t,x) + \frac{-2}{4\,\alpha\,t}\,\Phi(t,x) = \frac{1}{\alpha}\left(\frac{x^2}{4\,\alpha\,t^2} - \frac{1}{2\,t}\right)\Phi(t,x) \end{aligned}$$

and thus the differential equation in (5.7) is solved[2]. Since

$$\int_{-\infty}^{+\infty} \Phi(t,x)\;dx = 1 \quad\text{and}\quad \Phi(t,x) \to \begin{cases} 0 & \text{if } x \neq 0 \\ +\infty & \text{if } x = 0 \end{cases} \qquad \text{as } t \to 0+$$

the function $\Phi(t,x)$ converges to the Dirac impulse at $x = 0$ as $t$ approaches 0. This is visualized in Figure 5.3. This figure shows how an initial temperature peak at $x = 0$ spreads out as time $t > 0$ increases. It also implies that the temperature is strictly positive at any point $x$ as soon as $t > 0$, i.e. there there is no finite speed of propagation.

Using this fundamental solution and

$$\lim_{t\to 0+} \int_{-\infty}^{+\infty} \phi(t,x-\xi)\,u_0(\xi)\;d\xi = u_0(x)$$

the linear superposition principle leads to the solution of the initial value problem (5.7) given by

$$u(t,x) = \int_{-\infty}^{+\infty} \Phi(t,x-\xi)\,u_0(\xi)\;d\xi \;=\; \frac{1}{\sqrt{4\,\alpha\,\pi\,t}}\int_{-\infty}^{+\infty} \exp(-\frac{(x-\xi)^2}{4\,\alpha\,t})\,u_0(\xi)\;d\xi\,.$$

For the plane $\mathbb{R}^2$ ($n = 2$) and the space $\mathbb{R}^3$ ($n = 3$) the Green's function is given by

$$\Phi(t,\vec x) = \frac{1}{(4\,\alpha\,\pi\,t)^{n/2}}\,\exp(-\frac{\|\vec x\|^2}{4\,\alpha\,t}) \quad \text{for } \vec x \in \mathbb{R}^n,\text{ where } n = 1,2,3$$

[2]One can use Fourier transforms to derive the fomula, see Example 6–19 on page 93.

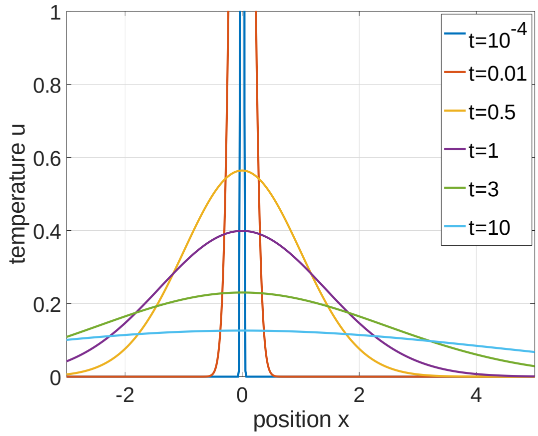



Figure 5.3: Fundamental solution of the dynamic 1D heat equation

and the solution of

$$\begin{aligned} \dot u(t,\vec x) &= \alpha\,\Delta u(t,\vec x) && \text{for } \vec x\in\mathbb{R}^n \text{ and } t>0\\ u(0,\vec x) &= u_0(\vec x) && \text{for } \vec x\in\mathbb{R}^n \end{aligned}$$

is given by

$$u(t,\vec x) = \iint\limits_{\mathbb{R}^n} \Phi(t,\vec x-\vec\xi)\,u_0(\vec\xi)\,dV_\xi = \frac{1}{(4\,\alpha\,\pi\,t)^{n/2}}\iint\limits_{\mathbb{R}^n} \exp\Big(-\frac{\|\vec x-\vec\xi\|^2}{4\,\alpha\,t}\Big)\,u_0(\vec\xi)\,dV_\xi\,. \tag{5.9}$$

Thus the behavior of the solutions is similar to the 1D situation.

## 5.4 Wave Equations, Hyperbolic Problems

### 5.4.1 Different forms of hyperbolic equations

A hyperbolic PDE always contains a second order derivative with respect to time $t$, i.e. we examine dynamic problems. The spatial derivatives are given by an elliptic differential operator. Find a few typical examples below.

$$\begin{aligned} \frac{\partial^2}{\partial t^2}\,u(t,x) &= \frac{\partial^2}{\partial x^2}\,u(t,x)+f(t,x)\\ \frac{\partial^2}{\partial t^2}\,u(t,x,y) &= \frac{\partial^2}{\partial x^2}\,u(t,x,y)+\frac{\partial^2}{\partial y^2}\,u(t,x,y)+f(t,x,y)\\ \frac{\partial^2}{\partial t^2}\,u(t,\vec x) &= \nabla(a(\vec x)\,\nabla\,u(t,\vec x))+f(t,\vec x)\\ \rho(\vec x)\,\frac{\partial^2}{\partial t^2}\,u(t,\vec x) &= \nabla(a(\vec x)\,\nabla\,u(t,\vec x))+\nabla(u(t,\vec x)\cdot\vec v(\vec x))+b(t,\vec x)\,u(t,\vec x)+f(t,\vec x) \end{aligned}$$

The last example is the most general one, with given functions $a(\vec x)>0$, $\rho(\vec x)>0$, $b(t,\vec x)$ and $f(t,\vec x)$. To obtain unique solutions for hyperbolic equations one has to specify boundary conditions for $\vec x\in\partial\Omega$ for all times $t>0$ and the initial value $u(0,\vec x)=u_0(\vec x)$ and the initial velocity $\frac{\partial}{\partial t}\,u(t,\vec x)=u_1(\vec x)$ for all $\vec x\in\Omega$.

### 5.4.2 Applications of hyperbolic equations

- propagation of sound, pressure waves

- ultrasonic waves
- electromagnetic waves
- water waves
- Elastic waves in solid, liquids and gases

### 5.4.3 Properties of solutions of the wave equation

**Traveling wave solutions**

Let $f(z)$ be a twice differentiable function and use the chain rule to conclude

$$\begin{aligned} \frac{\partial^2}{\partial t^2} f(x-c\,t) &= +c^2\, f''(x-c\,t) \\ \frac{\partial^2}{\partial x^2} f(x-c\,t) &= +f''(x-c\,t)\,. \end{aligned}$$

Thus $u(t,x) = f(x-c\,t)$ is a solution of the simplest form of the wave equation.

$$\ddot{u}(t,x) = c^2\, u''(t,x)\,.$$

This is a traveling wave of the shape given by $f$, moving to the right with speed $c$. Similarly one can generate waves moving to the left by $f(x+c\,t)$. Then any linear combination of left and right traveling waves solves the wave equation

$$u(t,x) = f_r(x-c\,t) + f_l(x+c\,t)\,.$$

In Figure 5.4 such a solution is shown with the waves traveling with speed $c = 1$. A $\sin(x)$ curve is moving to the right and a triangular peak is moving to the left, starting at position $x = 8$.

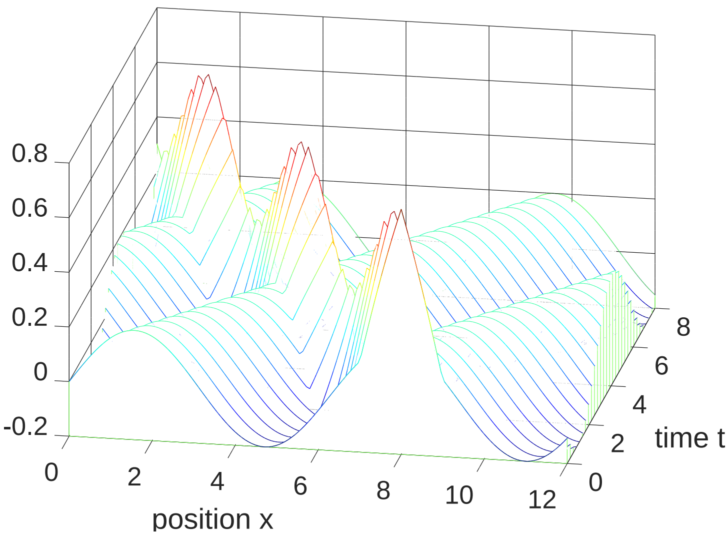



Figure 5.4: Two traveling waves, one to the left and one to the right

The above solution is correct if the wave equation is solved on $-\infty < x < +\infty$. If we have boundary conditions, then the traveling waves will be reflected at the boundaries.

In higher dimensions find similar solutions of

$$\ddot u(t,\vec x) = c^2\,\Delta u(t,\vec x)$$

of the form

$$u(t,\vec x) = f(\langle \vec x,\vec d\rangle - c\,t)\,.$$

This is a wave travelling with speed $c$ in the direction $\vec d$, where $\|\vec d\| = 1$.

**Conservation of energy**

Use Result 3–44 on page 32 to verify that we have conservation of energy. Let $u(t,\vec x)$ be a solution of

$$\rho\,\frac{\partial^2}{\partial t^2}\,u(t,\vec x) = \nabla(a(\vec x)\,\nabla u(t,\vec x)) \quad \text{for } \vec x\in\Omega$$

and either $u(t,\vec x) = 0$ or $\frac{\partial\, u(t,\vec x)}{\partial \vec n} = 0$ on all points of the boundary $\partial\Omega$. Then the total energy is conserved.

$$\begin{array}{rcrcl} & & \text{total energy} = \text{kinetic energy} & + & \text{potential energy} \\ E(0) & = & E(t) = \iint\limits_\Omega \frac{\rho}{2}\,\dot u^2(t,\vec x)\;dV & + & \iint\limits_\Omega \frac{a(\vec x)}{2}\,\|\nabla u(t,\vec x)\|^2\;dV \end{array}$$

This is very different from the behaviour of the dynamic heat equation, where we observed that the two expressions $\iint\limits_\Omega u^2(t,\vec x)\;dV$ and $\iint\limits_\Omega a(x)\,\|\nabla u(t,\vec x)\|^2\;dV$ are decaying, see Section 5.3.3.

### 5.4.4 Fourier approximation for the wave equation on an interval

The typical example for a one dimensional dynamic wave equation is

$$\begin{array}{rcll} \frac{\partial^2}{\partial t^2}\,u(t,x) &=& c^2\,\frac{\partial^2}{\partial x^2}\,u(t,x) & \text{for } 0\le x\le L \text{ and } t\in\mathbb{R} \\ u(t,0) = u(t,L) &=& 0 & \text{for } t\ge 0 \\ u(0,x) &=& u_0(x) & \text{for } 0<x<L \\ \frac{\partial}{\partial t}\,u(0,x) &=& u_1(x) & \text{for } 0<x<L \end{array}\,. \tag{5.10}$$

Using the same eigenfunctions as for the static heat problem (5.3) $\phi_n(x) = \sin(n\,\frac{\pi}{L}\,x)$

$$-\frac{\partial^2}{\partial x^2}\,\phi_n(x) = (\frac{n\,\pi}{L})^2\,\phi_n(x) = \lambda_n\,\phi_n(x)\,.$$

Now the differential equation and boundary condition in (5.10) are solved by

$$u_n(t,x) = \left(A\,\cos(c\,\sqrt{\lambda_n}\,t) + B\,\sin(c\,\sqrt{\lambda_n}\,t)\right)\;\phi_n(x)\,.$$

Observe that the frequencies are given by

$$\nu_n = \frac{\omega_n}{2\,\pi} = \frac{c\,\sqrt{\lambda_n}}{2\,\pi} = n\,\frac{c}{2\,L}\,,$$

i.e. they are multiples of the fundamental frequency $\frac{c}{2\,L}$. The corresponding wavelength $L_n$ is characterized by a period of $\phi_n(n\frac{\pi}{L}\,x) = \sin(n\frac{\pi}{L}\,x)$ and thus given by

$$\text{wavelength } = L_n = \frac{2\,L}{n}\,.$$

Assume that the two initial values are given by their Fourier series

$$u_0(x) = \sum_{n=0}^\infty a_n\,\sin(n\,\frac{\pi}{L}\,x) \quad\text{and}\quad u_1(x) = \sum_{n=0}^\infty b_n\,\sin(n\,\frac{\pi}{L}\,x)\,.$$

Then the solution of (5.10) is

$$u(t,x) = \sum_{n=1}^{\infty} \left( a_n \cos(c\sqrt{\lambda_n}\,t) + \frac{b_n}{c\sqrt{\lambda_n}} \sin(c\sqrt{\lambda_n}\,t) \right) \phi_n(x)\,.$$

Observe that this solution does not converge to zero, while the similar solution for the dynamic heat equation (5.6) converges to zero exponentially.

The above solution also exhibits the behavior of traveling waves, examined in Section 5.4.3, as long as one is away from the boundary points at $x = 0$ and $x = L$.

### 5.4.5 Solutions on unbounded domains

**Fundamental solution in one dimension, d'Alembert's solution**

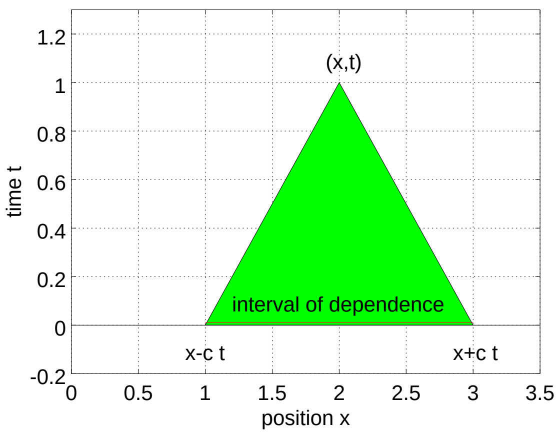
The unique solution of the initial value problem

$$\begin{aligned} \frac{\partial^2}{\partial t^2}\,u(t,x) &= c^2\,\frac{\partial^2}{\partial x^2}\,u(t,x) && \text{for } -\infty < x < \infty \text{ and } t \in \mathbb{R} \\ u(0,x) &= u_0(x) && \text{for } -\infty < x < \infty \\ \frac{\partial}{\partial t}\,u(0,x) &= u_1(x) && \text{for } -\infty < x < \infty \end{aligned} \tag{5.11}$$

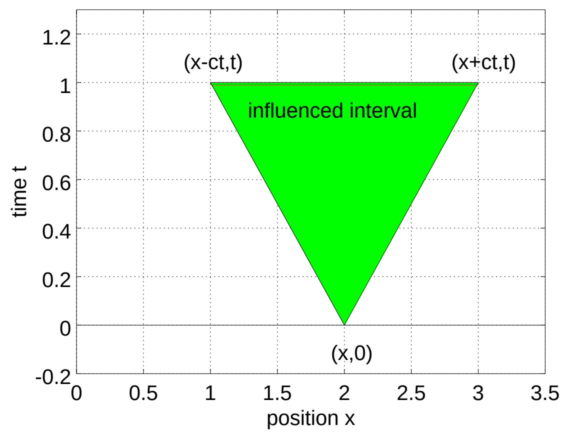
is given by

$$u(t,x) = \frac{1}{2}\,(u_0(x - c\,t) + u_0(x + c\,t)) + \frac{1}{2\,c}\int_{x-c\,t}^{x+c\,t} u_1(\xi)\,d\xi\,. \tag{5.12}$$

D'Alembert's formula implies that the solution of the wave equation at time $t$ and position $x$ is determined by the values in the **cone of dependence**, i.e. all times $\tau < t$ and positions $\tilde{x}$ such that $|\tilde{x} - x| \le c\,(t - \tau)$. This is visualized in Figure 5.5. On the left find the domain having an influence on the solution at time $t$ and position $x$ and on the right the domain influenced by the values at time $t = 0$ and position $x$. This formula and figure also confirm that information is traveling with speed $c$.




(a) cone of dependence




(b) cone of influence

Figure 5.5: D'Alembert's solution of the wave equation

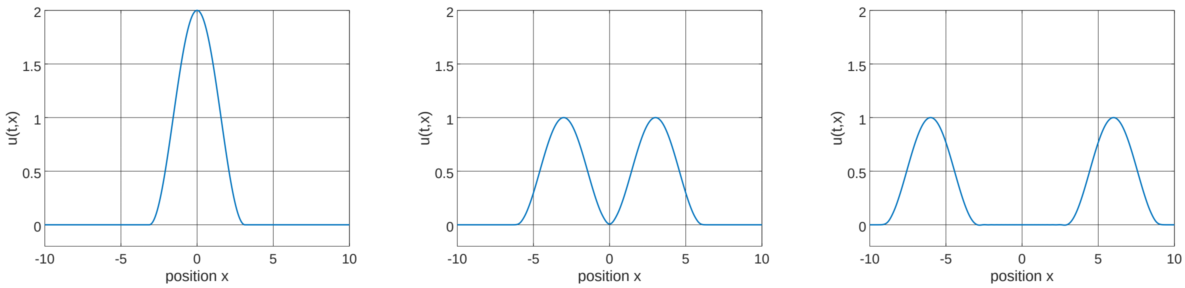
To illustrate the above we use the function

$$f(x) = \begin{cases} 1 + \cos(x) & \text{for } |x| \le \pi \\ 0 & \text{for } |x| \ge \pi \end{cases}$$

and examine the solution of

$$\ddot u(t,x) = u''(t,x) \quad \text{with} \quad u(0,x) = f(x) \quad \text{and} \quad \dot u(0,x) = 0$$

in Figure 5.6. The two pulses moving left and right are clearly visible. Observe the growing gap between the two pulses.



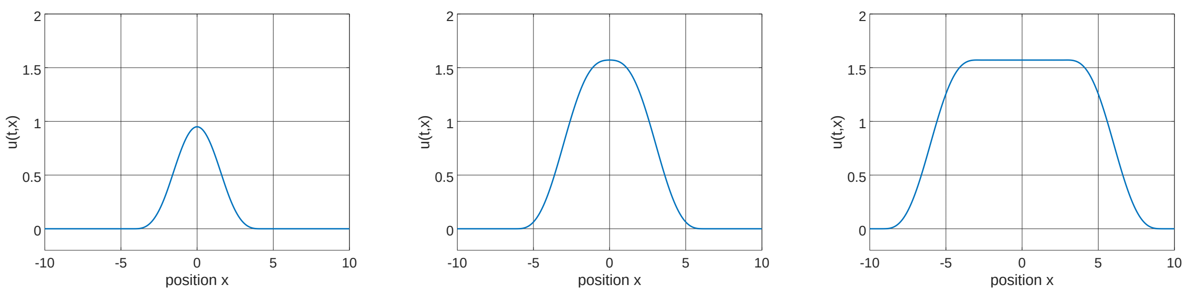

(a) $t = 0$ (b) $t = 3$ (c) $t = 6$

Figure 5.6: Solution of the wave equation with $u(x,0) = f(x)$ and $\dot u(0,x) = 0$

When starting with zero displacement, but a positive initial speed, the picture changes. Examine the solution of

$$\ddot u(t,x) = u''(t,x) \quad \text{with} \quad u(0,x) = 0 \quad \text{and} \quad \dot u(0,x) = \frac{1}{2}\, f(x)$$

in Figure 5.7. The two wave fronts moving left and right are clearly visible. There is no gap between the fronts, but the solution is equal to a nonzero constant for positions $x$ close to 0.




(a) $t = 1$ (b) $t = 3$ (c) $t = 6$

Figure 5.7: Solution of the wave equation with $u(x,0) = 0$ and $\dot u(0,x) = \frac{1}{2}\, f(x)$

A derivation of d'Alembert's formula (5.12) is given in [Farl82, Lesson 17], or many other books. We show the derivation here, for sake of completeness. Try to write the solution as a sum of a right and left traveling wave, i.e. $u(t,x) = f(x - c\,t) + g(x + c\,t)$. Differentiate this expression with respect to time $t$ and plug in $t = 0$ to find the condition

$$-c\, f'(x) + c\, g'(x) = u_1(x)\,.$$

Integrating this expression with respect to $x$ and combining it with $u(0,x) = u_0(x)$ leads to a linear system of equations

$$\begin{array}{rcl} f(x) & + & g(x) = u_0(x) \\ -f(x) & + & g(x) = \frac{1}{c}\int_0^x u_1(\xi)\; d\xi + k \end{array}$$

with the solutions

$$f(x) = \frac{1}{2}\left(u_0(x) - \frac{1}{c}\int_0^x u_1(\xi)\,d\xi - k\right) \quad \text{and} \quad g(x) = \frac{1}{2}\left(u_0(x) + \frac{1}{c}\int_0^x u_1(\xi)\,d\xi + k\right).$$

Then d'Alembert's formula (5.12) results from

$$u(t,x) = f(x-c\,t)+g(x+c\,t) = \frac{u_0(x-c\,t)+u_0(x+c\,t)}{2} + \frac{1}{2\,c}\left(\int_0^{x+c\,t} u_1(\xi)\,d\xi - \int_0^{x-c\,t} u_1(\xi)\,d\xi\right).$$

Observation: One may use the chain rule to verify

$$\frac{\partial}{\partial t}\int_{x-c\,t}^{x+c\,t} u_1(\xi)\,d\xi = c\,u_1(x+c\,t) + c\,u_1(x-c\,t)\,.$$

Thus the solution of $\ddot u(t,x) = c^2\,u''(t,x)$ with initial values $u(0,x) = 0$ and $\dot u(0,x) = u_1(x)$ can be differentiated to construct $v(t,x) = \frac{1}{c}\,\dot u(t,x)$. This function $v(t,x)$ and it will solve the wave equation $\ddot v(t,x) = c^2\,v''(t,x)$ with initial values $v(0,x) = u_1(x)$ and $\dot v(0,x) = 0$.

### Spherical waves in $\mathbb{R}^3$

Assuming that the solution $u$ depends on the radius only (i.e. $u(t,\vec x) = u(t,\|\vec x\|) = u(t,r)$) we use the Laplace operator in spherical coordinates

$$\Delta u = \frac{1}{r^2}\,\frac{\partial}{\partial r}(r^2\frac{\partial\,u}{\partial r}) + \frac{1}{r^2\sin^2\theta}\,\frac{\partial^2\,u}{\partial\phi^2} + \frac{1}{r^2\sin\theta}\,\frac{\partial}{\partial\theta}(\sin\theta\,\frac{\partial\,u}{\partial\theta})\,.$$

Since the solutions depends on time $t$ and radius $r$ only the wave equation $\ddot u = c^2\,\Delta u$ simplifies to

$$\begin{array}{rcll} \frac{\partial^2}{\partial t^2}\,u(t,r) &=& c^2\,\frac{1}{r^2}\,\frac{\partial}{\partial r}(r^2\,\frac{\partial}{\partial r}\,u(t,r)) & \text{for } r>0 \text{ and } t\in\mathbb{R} \\ u(0,r) &=& u_0(r) & \text{for } r>0 \\ \frac{\partial}{\partial t}\,u(0,r) &=& u_1(r) & \text{for } r>0 \end{array}\,. \tag{5.13}$$

Using the magic transformation $v(t,r) = r\,u(t,r)$ and

$$\frac{\partial}{\partial r}\,(r^2\,\frac{\partial\,u(r)}{\partial r}) = 2\,r\,\frac{\partial\,u(r)}{\partial r} + r^2\,\frac{\partial^2\,u(r)}{\partial r^2}$$

we find

$$\begin{aligned} \frac{\partial^2}{\partial r^2}\,v(t,r) &= \frac{\partial^2}{\partial r^2}\,(r\,u(t,r)) = \frac{\partial}{\partial r}\left(u(t,r) + r\,\frac{\partial\,u(t,r)}{\partial r}\right) \\ &= 2\,\frac{\partial\,u(t,r)}{\partial r} + r\,\frac{\partial^2\,u(t,r)}{\partial r^2} = \frac{1}{r}\left(\frac{\partial}{\partial r}\,(r^2\,\frac{\partial\,u(r)}{\partial r})\right) \\ &= \frac{r}{c^2}\,\frac{\partial^2\,u(t,r)}{\partial t^2} = \frac{1}{c^2}\,\frac{\partial^2\,r\,u(t,r)}{\partial t^2} = \frac{1}{c^2}\,\frac{\partial^2\,v(t,r)}{\partial t^2} \end{aligned}$$

and thus the modified function $v(t,r) = r\,u(t,r)$ solves the one dimensional wave equation

$$\begin{array}{rcll} \frac{\partial^2}{\partial t^2}\,v(t,r) &=& c^2\,\frac{\partial^2}{\partial r^2}\,v(t,r) & \text{for } r>0 \\ v(t,0) &=& 0 & \\ v(0,r) &=& r\,u_0(r) = v_0(r) & \text{for } r>0 \\ \frac{\partial}{\partial t}\,v(0,r) &=& r\,u_1(r) = v_1(r) & \text{for } r>0 \end{array}$$

and we can use d'Alembert's solution (5.12) with a modification taking the boundary condition $v(t,0)=0$ into account (see [Farl82, Lesson 18]). For $t\geq 0$ and $r\geq 0$ we obtain

$$\begin{aligned}
v(t,r) &= \frac{1}{2}\left(\operatorname{sign}(r-c\,t)\,v_0(|r-c\,t|)+v_0(r+c\,t)\right)+\frac{1}{2\,c}\int_{|r-c\,t|}^{r+c\,t} v_1(\xi)\;d\xi\\
&= \frac{1}{2}\left(\frac{|r-c\,t|}{\operatorname{sign}(r-c\,t)}\,u_0(|r-c\,t|)+(r+c\,t)\,u_0(r+c\,t)\right)+\frac{1}{2\,c}\int_{r-c\,t}^{r+c\,t}\xi\,u_1(\xi)\;d\xi\\
u(t,r) &= \frac{1}{2\,r}\left((r-c\,t)\,u_0(|r-c\,t|)+(r+c\,t)\,u_0(r+c\,t)\right)+\frac{1}{2\,c\,r}\int_{|r-c\,t|}^{r+c\,t}\xi\,u_1(\xi)\;d\xi\,. \qquad (5.14)
\end{aligned}$$

It is easy to see that this solution satisfies $u(0,r)=u_0(r)$. For $r>0$ and $0<c\,t<r$ the formula is very similar to D'Alembert's formula (5.12).

$$u(t,r)=\frac{r-c\,t}{2\,r}\,u_0(r-c\,t)+\frac{r+c\,t}{2\,r}\,u_0(r+c\,t)+\frac{1}{2\,c\,r}\int_{r-c\,t}^{r+c\,t}\xi\,u_1(\xi)\;d\xi\,.$$

Using $r>0$ and $0<c\,t<r$ find

$$\begin{aligned}
\frac{\partial}{\partial t}\,u(t,r) &= \frac{1}{2\,r}\left(-c\,u_0(r-c\,t)-(r-c\,t)\,c\,u_0'(r-c\,t)+c\,u_0(r+c\,t)+(r+c\,t)\,c\,u_0'(r+c\,t)\right)+\\
&\quad+\frac{1}{2\,c\,r}\left((r+c\,t)\,c\,u_1(r+c\,t)+(r-c\,t)\,c\,u_1(r-c\,t)\right)\\
\frac{\partial}{\partial t}\,u(0,r) &= \frac{1}{r}\left(-c\,u_0(r)-r\,c\,u_0'(r)+c\,u_0(r)+r\,c\,u_0'(r)\right)+\frac{1}{2\,c\,r}\left(r\,c\,u_1(r)+r\,c\,u_1(r)\right)\\
&= u_1(r)
\end{aligned}$$

and thus the initial conditions in (5.13) are satisfied.

Let us examine one consequence of the solution formula (5.14). If the initial values $u_0(r)$ and $u_1(r)$ are concentrated at the origin $r=0$, then the solution is only different form zero if $r=c\,t$ and its amplitude has a factor of $\frac{1}{r}$, i.e. we have traveling waves with radial speed $c$ and amplitudes proportional to $\frac{1}{r}=\frac{1}{c\,t}$. Thus a gun shot at $r=0$ and $t=0$ will be heard as a sharp noise at a distance $r$ at time $t=\frac{r}{c}$.

### Wave equation in $\mathbb{R}^3$

Examine the initial value problem in space $\mathbb{R}^3$:

$$\begin{aligned}
\tfrac{\partial^2}{\partial t^2}\,u(t,\vec{x})=c^2\,\Delta u(t,\vec{x}) &= c^2\left(\tfrac{\partial^2}{\partial x^2}+\tfrac{\partial^2}{\partial^2 y^2}+\tfrac{\partial^2}{\partial z^2}\right)u(t,\vec{x}) &&\text{for } \vec{x}=(x,y,z)\in\mathbb{R}^3\\
u(0,\vec{x})=u_0(\vec{x}) & &&\text{for } \vec{x}\in\mathbb{R}^3 \qquad (5.15)\\
\tfrac{\partial}{\partial t}\,u(0,\vec{x})=u_1(\vec{x}) & &&\text{for } \vec{x}\in\mathbb{R}^3\quad .
\end{aligned}$$

For a given point $\vec{x}\in\mathbb{R}^3$ and a time $t>0$ determine the average value $\Phi(t,\vec{x})$ of the initial displacement $u_0$ over a sphere $S(\vec{x},c\,t)$ with center at $\vec{x}$ and radius $c\,t$, i.e.

$$\begin{aligned}
\Phi(t,\vec{x}) &= \frac{\text{integral } u_0 \text{ over the sphere } S(\vec{x},\,c\,t)}{\text{surface area the sphere } S(\vec{x},\,c\,t)}\\
&= \frac{1}{4\,\pi\,c^2\,t^2}\int_{\phi=0}^{2\,\pi}\int_{\theta=0}^{\pi} u_0(x+c\,t\,\sin\theta\,\cos\phi,\;y+c\,t\,\sin\theta\,\sin\phi,\;z+c\,t\,\cos\theta)\;(c^2\,t^2)\;d\theta\;d\phi
\end{aligned}$$

and similarly the average value $\Psi(t,\vec{x})$ of the initial velocity $u_1$ over the same sphere. For the integration over the sphere use spherical coordinates as shown in Table 3.2 on page 23.

$$\Psi(t,\vec{x})=\frac{1}{4\,\pi\,c^2\,t^2}\int_{\phi=0}^{2\,\pi}\int_{\theta=0}^{\pi} u_1(x+c\,t\,\sin\theta\,\cos\phi,\;y+c\,t\,\sin\theta\,\sin\phi,\;z+c\,t\,\cos\theta)\;(c^2\,t^2)\;d\theta\;d\phi\,.$$

Then the solution of (5.15) is given by

$$u(t,\vec{x}) = t\cdot\Psi(t,\vec{x}) + \frac{\partial}{\partial t}\left(t\cdot\Phi(t,\vec{x})\right). \tag{5.16}$$

The derivation of the above formulas is rather involved and the result is known under the name **Poisson's integral**. Some of the simpler steps are shown in [Farl82, Lesson 24] and more details can be found in [TikhSama63, §5.1].

The main consequence of the above solution formula is that the solution only depends on the values of the initial displacement $u_0$ and the initial velocity, evaluated at a distance $c\,t$ away from the point of observation. This leads to **Huygen's principle**. The results for sound waves are:

- The sound at a position $\vec{x}$ and time $t$ is only determined by the initial values $u_0$ and $u_1$ at the points exactly a distance of $r = c\,t$ away from $\vec{x}$.
- A sharp sound (e.g. a gun shot) at the origin $\vec{0}$ at time 0 will be heard at a position $\vec{x}$ exactly at time $t = \|\vec{x}\|/c$.

The above does not apply to gun shots only, but also to a musical instrument played, which can be heard at any position in the room, without major distortion. This is different in the plane $\mathbb{R}^2$.

With $c\,t = r$ we can estimate the amplitude of the signal and find

$$|t\,\Phi(t,\vec{x})| \sim \frac{t}{c^2\,t^2} = \frac{1}{c\,r}$$

we find the amplitudes decaying like $\frac{1}{r}$ for $r = \|\vec{x}\|$ large.

If $u_0(r)$ and $u_1(r)$ are radially symmetric we find at the origin $\vec{0}$

$$u(t,\vec{0}) \;=\; t\cdot\Psi(t,\vec{0}) + \frac{\partial}{\partial t}\left(t\cdot\Phi(t,\vec{0})\right) \;=\; t\cdot u_0(c\,t) + \frac{\partial}{\partial t}\left(t\cdot u_1(c\,t)\right).$$

The solution (5.14) implies that for $c\,t > r > 0$

$$\begin{aligned}
u(t,r) &= \frac{1}{2\,r}\left((r-c\,t)\,u_0(c\,t-r) + (r+c\,t)\,u_0(r+c\,t)\right) + \frac{1}{2\,c\,r}\int_{c\,t-r}^{r+c\,t} \xi\,u_1(\xi)\,d\xi \\
&= \frac{c\,t\,(-u_0(ct-r) + u_0(c\,t+r))}{2\,r} + \frac{+u_0(c\,t-r) + u_0(c\,t+r)}{2} + \\
&\quad + \frac{1}{2\,c\,r}\int_{-r}^{+r} (c\,t+\xi)\,u_1(c\,t+\xi)\,d\xi \\
\lim_{r\to 0+} u(t,r) &= c\,t\,u_0'(c\,t) + u_0(c\,t) + t\,u_1(c\,t) \;=\; \frac{\partial}{\partial t}\left(t\,u_0(c\,t)\right) + t\,u_1(c\,t)
\end{aligned}$$

and thus coincides with the solution in (5.16).

### Waves in the plane $\mathbb{R}^2$

To examine the initial value problem in the plane $\mathbb{R}^2$

$$\begin{aligned}
\tfrac{\partial^2}{\partial t^2}\,u(t,\vec{x}) = c^2\,\Delta u(t,\vec{x}) \;&=\; c^2\left(\tfrac{\partial^2}{\partial x^2} + \tfrac{\partial^2}{\partial^2 y^2}\right)u(t,\vec{x}) && \text{for } \vec{x} = (x,y,z)\in\mathbb{R}^2 \\
u(0,\vec{x}) = u_0(\vec{x}) & && \text{for } \vec{x}\in\mathbb{R}^2 \\
\tfrac{\partial}{\partial t}\,u(0,\vec{x}) = u_1(\vec{x}) & && \text{for } \vec{x}\in\mathbb{R}^2
\end{aligned} \tag{5.17}$$

we use the solution formula for the situation in $R^3$. Extend the initial values to $R^3$, independent on $z$, i.e. $\tilde{u}_0(x,y,z) = u_0(x,y)$ and $\tilde{u}_1(x,y,z) = u_1(x,y)$. Then use Poisson's integral (5.16) and integrate over $z$. With some tedious computations (e.g. [TikhSama63, §5.1]) this leads to[3]

$$\begin{aligned} u(t,x,y) &= \frac{1}{2\pi c}\int_0^{2\pi}\int_0^{ct}\frac{u_1(x+r\cos\phi, y+r\sin\phi)}{\sqrt{(ct)^2-r^2}}\,r\,dr\,d\phi + \\ &+\frac{1}{2\pi c}\frac{\partial}{\partial t}\int_0^{2\pi}\int_0^{ct}\frac{u_0(x+r\cos\phi, y+r\sin\phi)}{\sqrt{(ct)^2-r^2}}\,r\,dr\,d\phi\,. \end{aligned} \tag{5.18}$$

It is important to observe that the integral is carried out over a disk with center at $(x,y)$ and radius $c\,t$. Thus Huygen's principle is not valid in the plane $\mathbb{R}^2$.

- The sound at a position $\vec{x}$ and time $t$ is determined by all initial values $u_0$ and $u_1$ inside a circle with center at $(x,y)$ and radius $c\,t$.
- A sharp sound (e.g. a gun shot) at the origin $\vec{0}$ at time 0 will be heard at a position $\vec{x}$ at all times $t \geq \|\vec{x}\|/c$.

This would make it impossible to listen to music in a plane $\mathbb{R}^2$, luckily we are living in $\mathbb{R}^3$. When throwing a rock in a calm pond you can observe the same spreading out of an initially concentrated solution.

The situation for the 1D wave equation is even more astonishing. Based on d'Alembert's formula

$$u(t,x) = \frac{1}{2}\,(u_0(x-c\,t) + u_0(c+c\,t)) + \frac{1}{2\,c}\int_{x-c\,t}^{x+c\,t} u_1(\xi)\,d\xi$$

conclude that the contribution to the solution by its initial displacement $u(0,x) = u_0(x)$ displays Huygen's principle, while the contribution by the initial speed $\dot{u}(0,x) = u_1(x)$ does not. This is visualized in Figures 5.6 and 5.7.

### Cylindrical waves

If we assume that the initial values depend on the radius $\rho = \sqrt{x^2+y^2}$ only we use the Laplace operator in polar coodinates

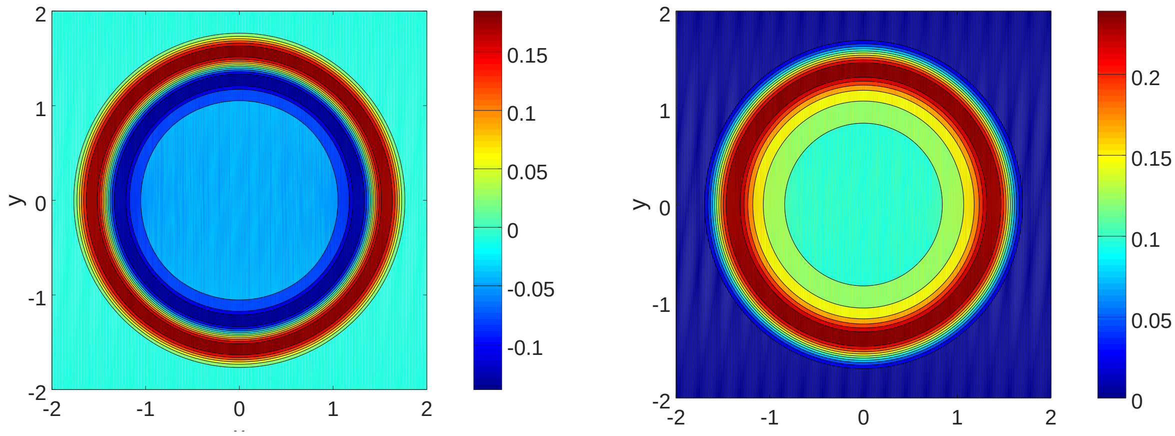
$$\Delta u = \frac{1}{\rho}\,\frac{\partial}{\partial\rho}(\rho\frac{\partial\,u}{\partial\rho}) + \frac{1}{\rho^2}\frac{\partial^2\,u}{\partial\phi^2} + \frac{\partial^2\,u}{\partial z^2}$$

and the wave equation $\ddot{u} = c^2\,\Delta u$ simplifies to

$$\begin{array}{rll} \frac{\partial^2}{\partial t^2}\,u(t,\rho) = c^2\,\frac{1}{\rho}\,\frac{\partial}{\partial\rho}(\rho\frac{\partial}{\partial\rho}\,u(t,\rho)) & \text{for } \rho>0 \text{ and } t\in\mathbb{R} & \\ u(0,\rho) = u_0(\rho) & \text{for } \rho>0 & . \\ \frac{\partial}{\partial t}\,u(0,\rho) = u_1(\rho) & \text{for } \rho>0 & \end{array} \tag{5.19}$$

Unfortunately there is no simple transformation to convert this into the standard 1D wave equation, which was possible for the 3D situation.

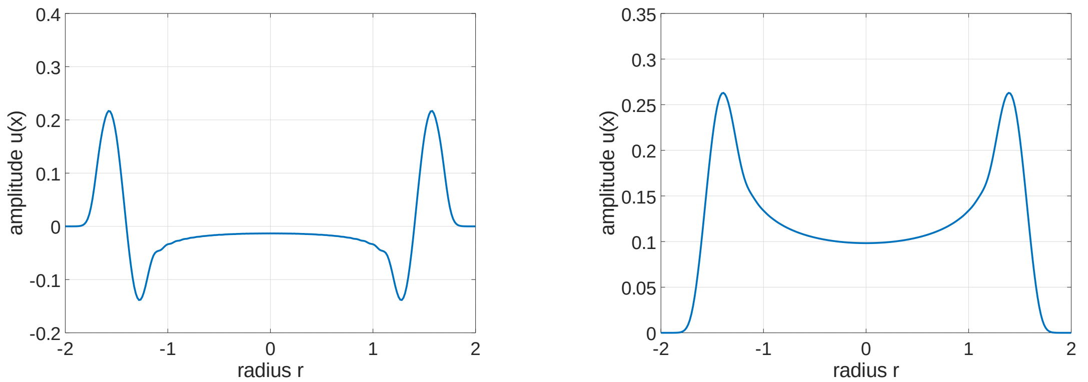
Based on the general solution formula (5.18) one can conclude that a positive initial speed $u_1(r)$ leads to positive solutions, while for $u_0(r) > 0$ no such statement is possible. With the function $f(r) = 1 + \cos(10\,r)$ for $|r| \leq \frac{\pi}{10}$ and zero otherwise we solved the wave equation $\ddot{u} = \Delta u$ with the initial conditions $u(0,r) = f(r)$ and $\dot{u}(0,r) = 0$, resp. $u(0,r) = 0$ and $\dot{u}(0,r) = 5\,f(r)$. Find the level curves in Figures 5.8 and slices through the origin in Figure 5.9. The snapshots were taken at time $t = 1.5$ and thus the leading edge of the waves is at $r = \frac{\pi}{10} + 1.5 \approx 1.8$, which is confirmed in the figures, which were generated by a finite difference method.




(a) with $u(0, r) = f(r) \geq 0$ and $\dot{u}(0, r) = 0$

(b) with $u(0, r) = 0$ and $\dot{u}(0, r) = 5\, f(r) \geq 0$

Figure 5.8: Level curves of solutions of the wave equation $\ddot{u} = \Delta u$ in $\mathbb{R}^2$




(a) with $u(0, r) = f(r) \geq 0$ and $\dot{u}(0, r) = 0$

(b) with $u(0, r) = 0$ and $\dot{u}(0, r) = 5\, f(r) \geq 0$

Figure 5.9: Slices of solutions of the wave equation $\ddot{u} = \Delta u$ in $\mathbb{R}^2$

It has to be pointed out that the different integrals leading to solutions of the wave equation in the plane or in space are not very useful when looking for numerical approximations of the solutions. Most often is is a better idea to use the method of finite differences or finite elements.

**Is Huygen's principle valid in 2D after all?**

Use delta functions[4] at $t = 0$ at the origin to conclude for $c\,t > \sqrt{x^2 + y^2}$

$$\begin{aligned}
2\,\pi\,c\,u(t,x,y) &= \int_0^{2\,\pi}\int_0^{c\,t} \frac{u_1(x + r\,\cos\phi, y + r\,\sin\phi)}{\sqrt{(c\,t)^2 - r^2}}\,r\,dr\,d\phi + \\
&\quad + \frac{\partial}{\partial t}\int_0^{2\,\pi}\int_0^{c\,t} \frac{u_0(x + r\,\cos\phi, y + r\,\sin\phi)}{\sqrt{(c\,t)^2 - r^2}}\,r\,dr\,d\phi \\
&= \frac{u_1}{\sqrt{(c\,t)^2 - x^2 - y^2}} + \frac{\partial}{\partial t}\frac{u_0}{\sqrt{(c\,t)^2 - x^2 - y^2}} \\
&= \frac{u_1}{\sqrt{(c\,t)^2 - x^2 - y^2}} - \frac{c^2\,t\,u_0}{\sqrt{(c\,t)^2 - x^2 - y^2}^3}\,.
\end{aligned}$$

For $c\,t \gg \sqrt{x^2 + y^2}$ obtain

$$2\,\pi\,c\,u(t,x,y) \approx \frac{u_1}{c\,t} - \frac{c^2\,t\,u_0}{(c\,t)^3} = \frac{u_1}{c\,t} - \frac{u_0}{c\,t^2}\,.$$

Thus we observe (hear) no ripples far behind the leading edge of the circular wave. This is confirmed by the flat areas in the center of the maps in Figure 5.8.

[3]This approach to use the solution in $\mathbb{R}^3$ to compute the solution in the plane is called **method of descent**.

[4] If $\vec{x}$ is inside of the domain $\Omega$, then by definition $\iint\limits_{\Omega} \delta(\vec{x} - \vec{\xi})\,f(\vec{\xi})\,dA_\xi = f(\vec{x})$.

# Chapter 6

# Fourier Methods

## 6.1 Keywords and Literature

Fourier series, fundamental frequency, sampling frequency, spectrum, DFT, FFT, aliasing.

Literature:

- The booklet [Jame02] is a nice, readable introduction to Fourier transforms and some of its applications.
- The book [ONei11] (or an other edition) has a section on Fourier methods.
- The book [CrofDaviHarg92] (or a newer edition) has a section on Fourier methods.
- The books by Lothar Papula cover all required topics, but these books are in German.
- The lecture notes by A. Stahel covers most relevant aspect of Fourier methods, available from the web page at https://web.sha1.bfh.science/Math2.pdf in German.
- Fourier methods are frequently used in optics, e.g. [Stew87].

In the book [Stew13] by Ian Stewart find an excellent characterization of the Fourier transform:

- **What does it say?**
  Any pattern in space and time can be thought of as a superposition of sinusoidal pattern with different frequencies.
- **Why is it that important?**
  The component frequencies can be used to analyze the pattern, create them to order, extract important features, and remove random noise.
- **What did it lead to?**
  Fourier's technique is very widely used, for example in image processing and quantum mechanics. It is used to find the structure of large biological molecules like DNA, to compress image data in digital photography, to clean up old or damaged audio records and to analyze earthquakes. Modern variants are used to store fingerprint data efficiently and to improve medical scanners.

## 6.2 Fourier Series

The Fourier series of a function $f(t)$, defined on the interval $0 \le t \le T$, is a $T$ periodic function of the form

$$f_{\text{Fourier}}(t) = \frac{a_0}{2} + \sum_{n=1}^{\infty} \left(a_n \cos(n\omega t) + b_n \sin(n\omega t)\right) , \tag{6.1}$$

where $\omega = \frac{2\pi}{T}$ and the Fourier coefficients $a_n$ and $b_n$ are determined by

$$a_n = \frac{2}{T} \int_0^T f(t) \, \cos\left(n\,\omega\,t\right) \, dt \quad \text{and} \quad b_n = \frac{2}{T} \int_0^T f(t) \, \sin\left(n\,\omega\,t\right) \, dt \, .$$

Observe the following facts about the Fourier series:

- Since all arguments of the trigonometric functions in (6.1) are of the form $n\,\omega\,t = n\,\frac{2\pi}{T}\,t$ the Fourier series $F_{\text{Fourier}}$ is a $T$-periodic function. All occurring frequencies are integer multiples of the **fundamental frequency** $\nu = \frac{1}{T}$.
- It is useful to consider the Fourier series as an approximation of the $T$**-periodic extension** of the original function $f(t)$.
- Since
$$a_n \cos(n\omega t) + b_n \sin(n\omega t) = s_n \, \cos(n\omega t - \phi_n)$$
with
$$s_n \; = \; \sqrt{a_n^2 + b_n^2} \quad \text{and} \quad \tan\phi_n \; = \; \frac{b_n}{a_n}$$
the **amplitude** of the signal contribution with the frequency $n\,\nu$ is given by $s_n = \sqrt{a_n^2 + b_n^2}$ and $\phi_n$ is a phase shift.
- The first contribution
$$\frac{a_0}{2} = \frac{1}{T} \int_0^T f(t) \, dt$$
is the **mean value** of the function $f(t)$ over the interval $[0\,,\,T]$.
- If the original function $f(t)$ is $T$-periodic we can integrate over any period of length $T$ to determine the Fourier coefficients, i.e.
$$a_n = \frac{2}{T} \int_{\text{period}} f(t) \, \cos\left(n\,\omega\,t\right) \, dt \quad \text{and} \quad b_n = \frac{2}{T} \int_{\text{period}} f(t) \, \sin\left(n\,\omega\,t\right) \, dt \, .$$

Based on Euler's formula $e^{i\,\alpha} = \cos\alpha + i\,\sin\alpha$ one can verify that the Fourier series can also be written in a complex form

$$F_{\text{Fourier}}(t) = \sum_{n=-\infty}^{+\infty} c_n \, e^{i\,n\,\omega\,t} \tag{6.2}$$

where the **complex Fourier coefficients** are given by

$$c_n = \frac{1}{T} \int_0^T f(t) \, e^{-in\omega t} \, dt \, .$$

For real valued functions this can be rewritten in the form

$$F_{\text{Fourier}}(t) = c_0 + 2 \text{ Re} \left( \sum_{n=1}^{+\infty} c_n \, e^{i\,n\,\omega\,t} \right) \, . \tag{6.3}$$

For real valued functions the connection between the real Fourier coefficients $a_n$ and $b_n$ and the complex Fourier coefficients is given by

$$\begin{array}{rcl} a_n & = & c_n + c_{-n} \\ b_n & = & i\,(c_n - c_{-n}) \end{array} \qquad \text{or} \qquad \begin{array}{rcl} c_n & = & \frac{1}{2}\,(a_n - i\,b_n) \\ c_{-n} & = & \frac{1}{2}\,(a_n + i\,b_n) \end{array} .$$

• **6–1 Question:**
Determine the Fourier series of the function

$$f\,(x) = x \qquad \text{on the intervall} \quad [-\pi, \pi]\,.$$

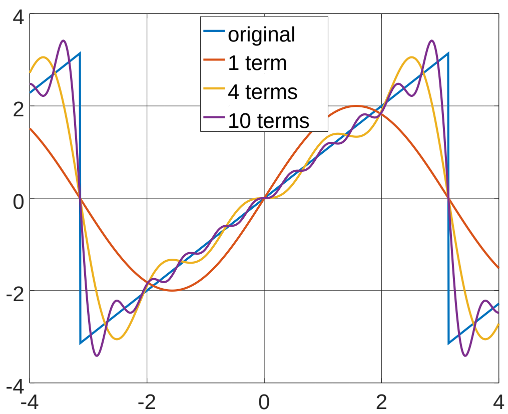
- Determine the real Fourier coefficients $a_n$ and $b_n$.
- Determine the complex Fourier coefficients $c_n$.

Verify that the Fourier series is given by

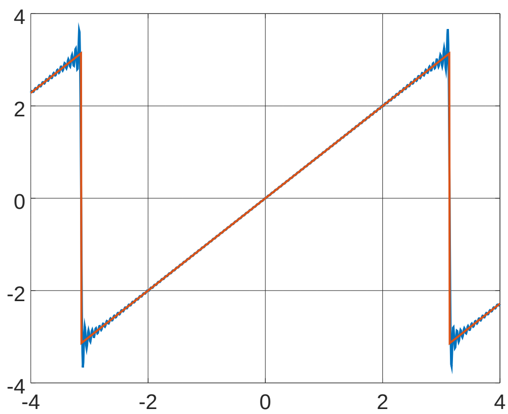
$$\begin{aligned} x \;\sim\; & \sum_{n=1}^{\infty} b_n\,\sin(n\,x) = \sum_{n=1}^{\infty} (-1)^{n+1}\frac{2}{n}\,\sin(n\,x) \\ = \;& 2\,\sin x - \sin(2\,x) + \frac{2\,\sin(3\,x)}{3} - \frac{2\,\sin(4\,x)}{4} + \frac{2\,\sin(5\,x)}{5} - \ldots \end{aligned}$$

The result is illustrated by Figure 6.1.

♢




(a) approximation by 1, 4 and 10 contributions




(b) approximation by many contributions

Figure 6.1: Fourier approximation of $f(x) = x$ on $[-\pi, \pi]$

The result if Figure 6.1 displays a couple of features:

- Since the original function $f(x) = x$ on the symmetric interval $[-\pi\,,\,+\pi]$ is odd, there was no need to calculate the coefficients $a_n$ for the $\cos(n\omega\,t)$ contributions. A simple symmetry argument implies $a_n = 0$.
- If we use many contributions then the approximation is very good on most parts of the interval.
- At the points $x = \pm\pi$ the approximation remains of poor quality. This is the phenomenon of Gibbs. More details are shown in Figure 6.2, where 100 contributions were used.

The observation in Question 6–1 illustrates the general convergence result for Fourier series.

• **6–2 Definition:** A function $f$ is said to be **piecewise continuous** iff

1. The function is continuous on a finite number of open subintervals.
2. If the function is not continous at a point $x = x_0$, then the left and right limits exist, i.e. one sided limits $f(x_0-) = \lim_{x\to x_0-} f(x)$ and $f(x_0+) = \lim_{x\to x_0+} f(x)$ exist.

◇

• **6–3 Result:** Let $f(x)$ be a piecewise differentiable function on the interval $[0\,,\,T)$, resp. its $T$-periodic extension. Then examine the sequence of **Fourier partial sums**

$$f_N\,(x) = \frac{a_0}{2} + \sum_{n=1}^{N}(a_n\cos(n\,x) + b_n\sin(n\,x))\,.$$

1. If $f$ is continuous on a subinterval $[a, b]$, then $f_N$ converges uniformly to $f$ on this subinterval.
2. If $f(x)$ is not continuous at $x = x_0$, then the $f_N(x_0)$ converges to the average of the left and right limit of $f$ at $x = x_0$, i.e.
$$\lim_{N\to\infty} f_N(x_0) = \frac{1}{2}\,(f(x_0-) + f(x_0+)) = \frac{1}{2}\,\left(\lim_{x\to x_0-} f(x) + \lim_{x\to 0+} f(x)\right)\,.$$
3. If the function $f$ is not continuous at $x = x_0$, then the phenomena of Gibbs will show. Even for a Fourier approximation $f_N$ with many terms (i.e. $N$ large) the Fourier partial sum $f_N(x)$ will display bumps on both sides of $x_0$ of an approximate height of 9% of the size of the jump.

A notation for this convergence result is

$$f(x) \sim \lim_{N\to\infty} f_N(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty}\left(a_n\cos(n\omega x) + b_n\sin(n\omega x)\right)\,.$$

◇

The $2\pi$-periodic extension of $f(x)$ on $[-\pi\,,\,+\pi]$ in Exercise 6–1 has jumps of size $2\pi$ at $x = \pm\pi$. Thus the jump has size $2\pi$. In Figure 6.2 the bump of size $0.09 \cdot 2\pi \approx 0.57$ is clearly visible in Figure 6.2.

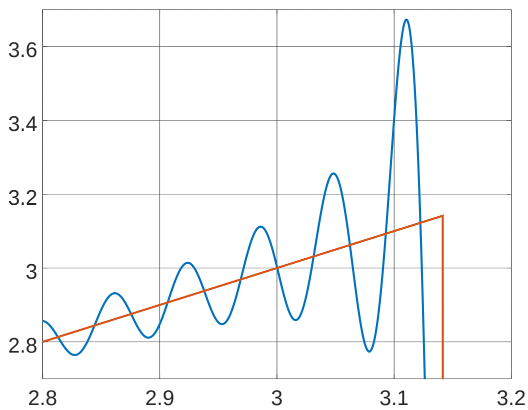



Figure 6.2: Phenomena of Gibbs

If a function $f(x)$ is given on the interval $[0\,,\,L]$ we can examine the Fourier approximation of its **odd extension** $f_o(x)$, given by

$$f_o(x) = \begin{cases} +f(+x) & \text{for} \quad 0 \le x \le L \\ -f(-x) & \text{for} \quad -L < x < 0 \end{cases}$$

resp. its $2L$-periodic extension. Since this new function is odd, we know that $a_n = 0$ and

$$b_n \quad = \quad \frac{2}{2\,L} \int_{-L}^{+L} f_o(x)\,\sin(n\,\frac{2\,\pi}{2\,L}\,x)\;dx = \frac{2}{L} \int_0^{+L} f(x)\,\sin(n\,\frac{2\,\pi}{2\,L}\,x)\;dx$$

and the **Fourier sine series** is given by

$$f(x) \sim \sum_{n=1}^{\infty} b_n \sin(n\frac{\pi}{L}x)\,.$$

This allows to approximate any function on $[0\,,\,L]$ by a series of $\sin(n\,\frac{\pi}{L}\,x)$ functions with zeros at $x = 0$ and $x = L$.

**• 6–4 Question:**
Determine the Fourier sine series of the function $f(x) = 1$ on the interval $[0\,,\,\pi]$. Verify that the series is given by

$$f\,(x) \sim \frac{4}{\pi}\,(\frac{\sin x}{1} + \frac{\sin 3x}{3} + \frac{\sin 5x}{5} + \frac{\sin 7x}{7} + \ldots)$$

Explain the behavior displayed in Figure 6.3.

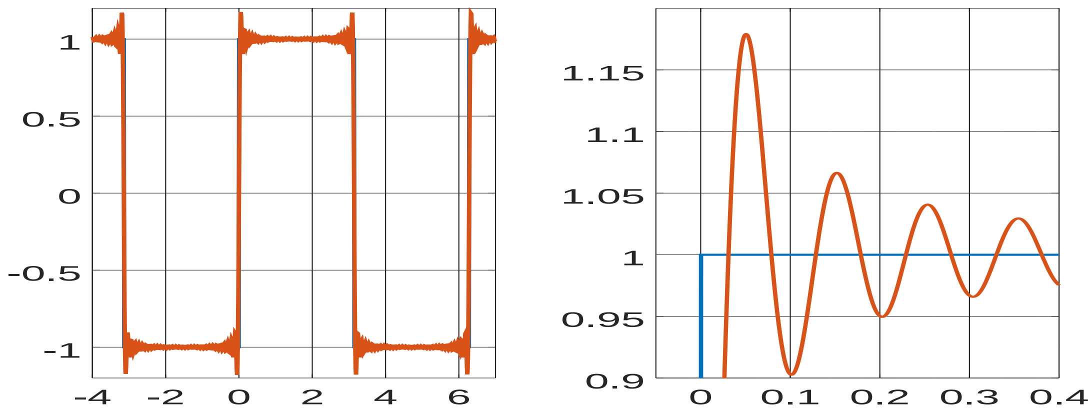



Figure 6.3: Phenomena of Gibbs for the function $f\,(x) = \text{sign}\;x$

$\diamondsuit$

**• 6–5 Question:**
Determine the Fourier sine series of the solution $u(x)$ of the boundary value problem

$$-u''(x) = x \quad \text{for} \quad 0 < x < \pi \quad \text{and} \quad u(0) = u(\pi) = 0\,.$$

Use the result of Question 6–1 and the fact that the boundary value problem

$$-u_n''(x) = \sin(n\,x) \quad \text{for} \quad 0 < x < \pi \quad \text{and} \quad u_n(0) = u_n(\pi) = 0$$

is solved by $u_n(x) = \frac{1}{n^2}\sin(n\,x)$. Using the principle of linear superposition the boundary value problem can easily be solved if the right hand side is a linear combination of functions $\sin(n\,x)$ for different values of $n \in \mathbb{N}$.

$\diamondsuit$

The method used in the above exercise is applicable in many more situations. You can find solutions to steady state heat equations (Section 5.2.5), dynamic heat equations (Section 5.3.4) and wave equations (Section 5.4.4).

For a $T$-periodic function use equations (6.1) and (6.2) to realize that the number

$$s_n = 2\,|c_n| = 2\,\sqrt{\frac{a_n^2}{2^2} + \frac{b_n^2}{2^2}} = \sqrt{a_n^2 + b_n^2}$$

is the amplitude of the contribution

$$2\ \mathrm{Re}\,(c_n\ \exp(n\,\frac{2\,\pi}{T}\,t)) = \sqrt{a_n^2 + b_n^2}\ \cos(n\,\frac{2\,\pi}{T}\,t - \phi_n)$$

with frequency $\frac{n}{T}$. Thus plotting this amplitude as function of the frequency leads to the **spectrum** of the function $f(t)$. For many applications the square of the amplitude measures the power in the signal, thus plotting $s_n^2 = a_n^2 + b_n^2 = 4\,|c_n|^2$ leads to the power spectrum.

• **6–6 Example:** For the $2\pi$-periodic extension of the function $\mathrm{sign}(t)$ on $[-\pi\,,\,+\pi]$ we have the Fourier series (see Exercise 6–4)

$$\mathrm{sign}(t) \sim \frac{4}{\pi}\,(\frac{\sin x}{1} + \frac{\sin 3x}{3} + \frac{\sin 5x}{5} + \frac{\sin 7x}{7} + \ldots)\,.$$

Since the interval has length $2\,\pi$ the fundamental frequency is $\frac{1}{2\pi}$ and thus the frequencies given by $\nu = \frac{n}{2\pi}$. The spectrum $\mathrm{S}(\nu)$ and the power spectrum $\mathrm{PS}(\nu)$ are given by

$$\begin{aligned}
\mathrm{S}(n) &= \begin{cases} \frac{4}{\pi\,n} & \text{for odd values of } n \\ 0 & \text{for even values of } n \end{cases} \\
\mathrm{PS}(n) &= \begin{cases} \frac{4^2}{\pi^2\,n^2} & \text{for odd values of } n \\ 0 & \text{for even values of } n \end{cases}
\end{aligned}\ .$$

The two results are shown in Figure 6.4.

$\diamondsuit$

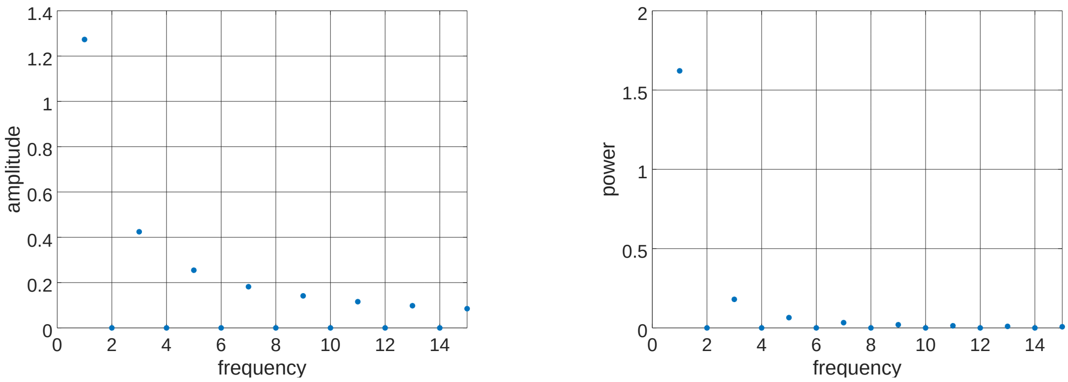
## 6.3 Discrete Fourier Transform, DFT and FFT

The Fourier series in the preceeding section writes “any” $T$-periodic function as a series of trigonometric functions, i.e.

$$f(t) \sim c_0 + 2\ \mathrm{Re}\left(\sum_{n=1}^{+\infty} c_n\,e^{i\,n\,\frac{2\pi}{T}\,t}\right) = \frac{a_0}{2} + \sum_{n=1}^{\infty}\sqrt{a_n^2 + b_n^2}\ \cos(n\,\frac{2\,\pi}{T}\,t - \phi_n)$$

where the complex coefficients are given by the integrals

$$c_n = \frac{1}{T}\int_0^T f(t)\,e^{-i\,n\frac{2\pi}{T}\,t}\,dt$$




(a) spectrum (b) power spectrum

Figure 6.4: Spectrum and power spectrum of the signal sign$(t)$ on $[-\pi\,,\,+\pi]$

If your computations are based on measurements you will only have the data at a discrete set of $N+1$ points $t_k$ where

$$t_k = k\,\frac{T}{N} \quad \text{and} \quad f_k = f(t_k) \quad \text{for } k = 0, 1, 2, 3, \ldots, N$$

Since the function is (assumed to be) $T$ periodic we have $f(0) = f(T)$, or $f_0 = f_N$.

- The **fundamental frequency** is determined by the length $T$ of the interval over which we gather data and is given by $\frac{1}{T}$.
- The function will be approximated by a sum of trigonometric functions whose frequencies are multiples of the fundamental frequency.
- The **sampling frequency** depends on the number $N$ of data points and is given by $\frac{N}{T}$. The sampling interval is $\Delta t = \frac{T}{N}$.

Since the function $f$ is $T$-periodic we have $f_0 = f_N$. Now we determine the Fourier coefficient by a numerical integration with the trapezoidal rule. Use $\frac{\Delta t}{T} = \frac{T}{N\,T} = \frac{1}{N}$ and $e^{i\,0} = e^{i\,n\,2\pi} = 1$ to conclude

$$\begin{aligned}
c_n &= \frac{1}{T}\int_0^T f(t)\,e^{-i\,n\,\frac{2\pi}{T}\,t}\,dt \\
&\approx \frac{1}{T}\left(\frac{1}{2}\,f(0)\,e^{-i\,n\,\frac{2\pi}{T}\,0} + \left(\sum_{k=1}^{N-1} f(k\,\Delta t)\,e^{-i\,n\,\frac{2\pi}{T}\,k\,\Delta t}\right) + \frac{1}{2}\,f(T)\,e^{-i\,n\,\frac{2\pi}{T}\,T}\right)\,\Delta t \\
&= \frac{1}{N}\left(\sum_{k=0}^{N-1} f(k\,\Delta t)\,e^{-i\,\frac{2\pi}{N}\,n\,k}\right) = y_n\,. \qquad (6.4)
\end{aligned}$$

The complex vector $\vec{y} \in \mathbb{C}^N$ is the **discrete Fourier transform** (DFT) of the function $f$, resp. its discretized values in the vector $\vec{f} \in \mathbb{C}^N$. Based on equation (6.4) we have

$$\vec{y} = \begin{pmatrix} y_0 \\ y_1 \\ y_2 \\ y_3 \\ \vdots \\ y_{N-1} \end{pmatrix} \approx \begin{pmatrix} c_0 \\ c_1 \\ c_2 \\ c_3 \\ \vdots \\ c_{N-1} \end{pmatrix} \in \mathbb{C}^N \quad \text{and} \quad \vec{f} = \begin{pmatrix} f_0 \\ f_1 \\ f_2 \\ f_3 \\ \vdots \\ f_{N-1} \end{pmatrix} \in \mathbb{C}^N\,.$$

The complex number $w = \exp(i\,\frac{2\pi}{N}) = \cos(\frac{2\pi}{N}) + i\,\sin(\frac{2\pi}{N})$ satisfies $w^N = 1$. Thus $w$ is the complex

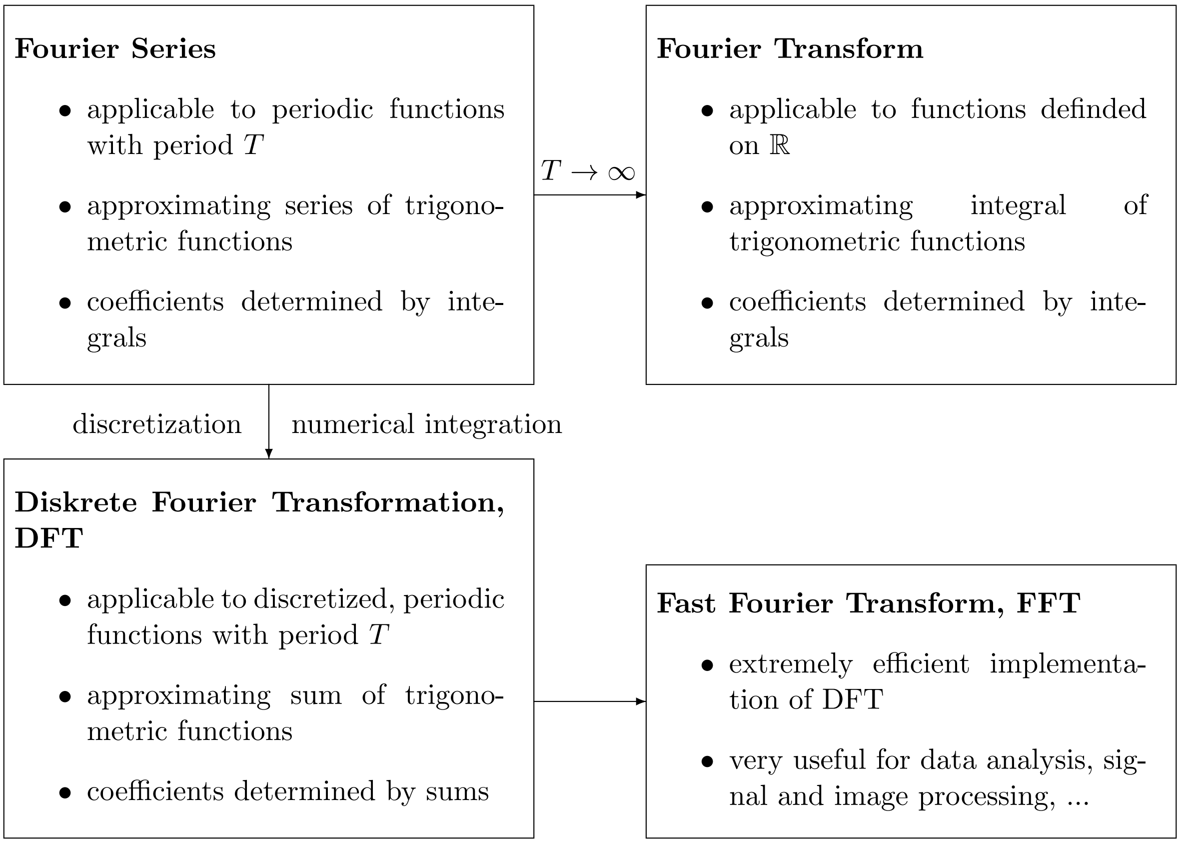



Figure 6.5: Fourier methods

number with absolute value $|w| = 1$ and the angle is $\frac{2\pi}{N} = \frac{360^\circ}{N}$ and $w^N = 1$. With this number we construct the complex $N \times N$ Fourier matrices $\mathbf{F}_N$

$$\mathbf{F}_N = \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ 1 & w & w^2 & \dots & w^{N-1} \\ 1 & w^2 & w^4 & \dots & w^{2\,(N-1)} \\ 1 & w^3 & w^6 & \dots & w^{3\,(N-1)} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & w^{1\,(N-1)} & w^{2\,(N-1)} & \dots & w^{(N-1)\,(N-1)} \end{bmatrix} .$$

Now equation (6.4) implies

$$\vec{y} = \frac{1}{N}\,\overline{\mathbf{F}_N}\,\vec{f}\,.$$

One can verify that

$$\mathbf{F}_N \cdot \overline{\mathbf{F}_N} = \overline{\mathbf{F}_N} \cdot \mathbf{F}_N = N\,\mathbb{I}$$

and this leads to

$$\mathbf{F}_N^{-1} = \frac{1}{N}\,\overline{\mathbf{F}_N} \quad \text{and} \quad \overline{\mathbf{F}_N}^{-1} = \frac{1}{N}\,\mathbf{F}_N\,.$$

Thus we have the **inverse DFT** given by

$$\vec{f} = \mathbf{F}_N\,\vec{y}$$

or if spelled out with a summation

$$f(k\,\Delta t) = f_k = \sum_{n=0}^{N-1} w^{nk}\, y_n = \sum_{n=0}^{N-1} e^{i\,\frac{2\pi}{N}\,n\,k}\, y_n \qquad \text{for} \quad k = 0, 1, 2, \ldots, N-1\,.$$

Based on the above calculations there is a close similarity between Fourier series and the DFT. Use $\Delta t = \frac{T}{N}$ and $\omega = \frac{2\pi}{T}$ to verify the table below.

| DFT<br>discrete | Fourier series<br>continuous |
|---|---|
| $y_n = \frac{1}{N} \sum_{k=0}^{N-1} e^{-i\,\frac{2\pi}{N}\,n\,k}\, f_k$ | $c_n = \frac{1}{T} \int_0^T f(t)\, e^{-i\,n\,\frac{2\pi}{T}\,t}\, dt$ |
| $f_k = \sum_{n=0}^{N-1} e^{+i\,\frac{2\pi}{N}\,n\,k}\, y_n$ | $f\,(t) = \sum_{n=-\infty}^{\infty} c_n e^{+i\,n\,\frac{2\pi}{T}\,t}$ |

Unfortunately there are a few different options to implement the DFT and its inverse. It is not important which set of formulas is used. The only important point is to use a consistent set of formulas for the DFT and the inverse DFT.

- The factor $\frac{1}{N}$ can show at different places:
  - A factor of $\frac{1}{N}$ for the DFT and a factor of 1 for the inverse. This is used in these notes.
  - A factor of 1 for the DFT and a factor of $\frac{1}{N}$ for the inverse. This is used by *Octave*/`MATLAB`.
  - A factor of $\frac{1}{\sqrt{N}}$ for the DFT and its inverse. Because of the symmetry this is often used by mathematicians.
- The negative sign in the complex exponential function can be used in the DFT or the inverse.
  - The exponential expression $\exp(-i\,\frac{2\pi}{N}\,n\,k)$ in the DFT and $\exp(+i\,\frac{2\pi}{N}\,n\,k)$ in the inverse. This is used in these notes.
  - The exponential expression $\exp(+i\,\frac{2\pi}{N}\,n\,k)$ in the DFT and $\exp(-i\,\frac{2\pi}{N}\,n\,k)$ in the inverse. This is used by *Octave*/`MATLAB`.

The next exercise illustrates this variety of formulas.

**• 6–7 Question:**
Examine a DFT with four points only.

(a) Write down the Fourier matrix $\mathbf{F}_4$.

(b) Use the `MATLAB` command `fft(eye(4))` to determine the Fourier matrix used by `MATLAB`. Use `ifft(eye(4))` to obtain the matrix for the inverse DFT.

(c) Verify that `MATLAB` and *Octave* use a different set of formulas for the DFT and its inverse, namely

$$\vec{y} = \mathbf{F}_N\,\vec{f} \quad \text{and} \quad \vec{f} = \frac{1}{N}\,\overline{\mathbf{F}_N}\,\vec{y}$$

or if written with a summation

$$y_n = \sum_{k=0}^{N-1} f_k\, e^{+i\,\frac{2\pi}{N}\,n\,k} \quad \text{and} \quad f_k = \frac{1}{N} \sum_{n=0}^{N-1} y_n\, e^{+i\,\frac{2\pi}{N}\,n\,k}$$

♢

• **6–8 Result:** The above matrix notation for the DFT seems to imply that one needs $N^2$ flops for the matrix multiplication. In 1965 James Cooley and John Tukey found a very ingenious algorithm to compute the DFT, see [CoolTuke65]. The algorithm is known as **FFT** (Fast Fourier Transform) and is one of the main reasons for the many applications of Fourier methods. The number of required flops drops from $N^2$ to $N \log_2 N$. ♢

• **6–9 Result:** To examine the function $f(t) = t$ on the interval $[-\pi\,,\,+\pi]$ with `MATLAB` or *Octave* and determine the spectrum we proceed as follows:

1. Choose the number $N$ of gridpoints to be used for the DFT, preferably a power of 2. Then create the uniformly distibuted values for the times $t_k$. Use $N+1$ points between $-\pi$ and $+\pi$.
2. Evaluate the function $f(t)$ at those points, relpace the first value by the average of the first and last values. Then use the first $N$ points only.
3. Display the first few values of $c_n$, resp. its DFT approximation $y_n$.

Observe that in the notes the numbering starts at 0 (e.g. $c_0$), but `MATLAB`/*Octave* indices always start at 1. Thus the comand `c(1:6)` displays $c_0$ up to $c_5$.

**FFT_t.m**

```
N = 2^7;                      % choose the number of discretization points
t = linspace(-pi,pi,N+1);     % define the time discretization with N+1 points

f = t; f(1) = (f(1)+f(N+1))/2; f = f(1:N);  % determine the values of the function
                   % use average of first and last values, then use N points only
c = fft(f)*2/N;    % determine the Fourier coefficients
cn  = c(1:6)
cni = imag(c(1:6))      % display the imaginary parts of the first few
figure(1)
plot([0:N-1],abs(c),'+')     % display the spectrum of the amplitudes
xlabel('frequency [Hz]'); ylabel('amplitude')
-->
cn =    -0.00000 + 0.00000i    0.00000 + 1.99960i   -0.00000 + 0.99920i
        -0.00000 + 0.66546i    0.00000 + 0.49839i   -0.00000 + 0.39799i

cni =    0.00000    1.99960    0.99920    0.66546    0.49839    0.39799
```

Using Question 6–1 we know

$$x \sim 2\,\sin\,x - \sin\,(2\,x) + \frac{2\,\sin\,(3\,x)}{3} - \frac{2\,\sin\,(4\,x)}{4} + \frac{2\,\sin\,(5\,x)}{5} - \ldots$$

thus

$$c_n = \frac{1}{2}\,(a_b - i\,b_n) = \frac{1}{2}\left(0 - i\,(-1)^{n+1}\,\frac{2}{n}\right) = i\,(-1)^n\,\frac{1}{n}$$

and thus the amplitude of the contribution with frequency $n$ is $2\,|c_n| = \frac{2}{n}$, which is confirmed by the above numerical result. Observe that the result of the command `fft()` is multiplied by $2/N$ and then leads to $\overline{c_n}$.

The amplitude spectrum in Figure 6.6 shows a very surprising effect: in the right half of the figure the amplitudes do **not** decay like $\frac{2}{n}$, but seem to be an mirrored image of the left half of the spectrum. This is not a coincidence, but caused by **aliasing**. One can show that for real functions $f(t)$ we obtain periodic coefficients $y_{n+N} = y_n$ and $y_{-n} = \overline{y_n}$. In the above example with $N = 128$ we find e.g. $|y_{120}| = |y_{-8}| = |y_8|$. Thus only the left half of the spectrum is useful. Since the maximal frequency shown in the full spectrum equals the sampling frequency $\frac{N}{T} = \frac{1}{\Delta t}$ we conclude that only **information up to half of the sampling frequency is valid**. This is called the **Nyquist frequency**. ♢

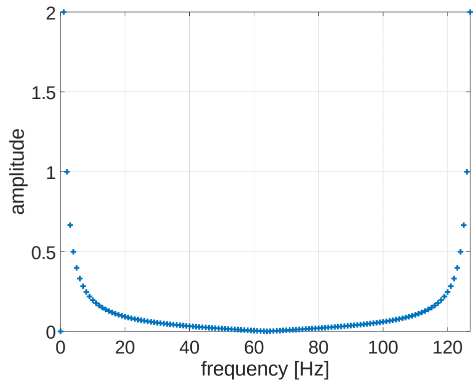



Figure 6.6: Spectrum of the signal $f(t) = t$ on $[-\pi\,, +\pi]$ determined by `fft()`

**• 6–10 Question:**
Examine the function

$$f(t) = 2\,\sin(100.2 \cdot 2\pi \cdot t) - 3\,\cos(102.5 \cdot 2\pi \cdot t) + 0.5\,\sin(150 \cdot 2\pi \cdot t);$$

on the intervall $[0\,, T]$ with $N$ discretization points. Examine different values for $T$ and $N$.

(a) Use $T = 2$ and $N = 2^{10} = 1024$.

(b) Use $T = 5$ and $N = 2^{10} = 1024$.

(c) Use $T = 20$ and $N = 2^{14} = 16384$.

(d) Use $T = 20$ and $N = 2^{16} = 65536$.

Use *Octave*/`MATLAB` to generate the data and add some small random noise. The code might look like

```
func = @(t)2*sin(100.2*2*pi*t) - 3*cos(102.5*2*pi*t) + 0.5*sin(150*2*pi*t);
T = 5;   N = 2^10;              % choose the values
t = linspace(0,T-T/N,N);        % generate the sampling times
f = func(t)+ 0.1*randn(1,N);    % generate the function with added noise
```

For each of the above cases:

- Determine the Nyquist frequency.
- Determine the spectrum of amplitudes and display the spectrum, but only frequencies up to 200 Hz.
- Explain the observed effects in the spectrum.
- Increase the size of the random noise and observe the spectrum.

♢

## 6.4 Fourier Transform

### 6.4.1 From Fourier series to Fourier transform

For a $2T$-periodic function $f(t)$ on the interval $[-T\,,\,+,T]$ use the Fourier series from the previous section and write

$$f(t) \sim \sum_{n=-\infty}^{\infty} c_n\, e^{i\,n\,\frac{2\,\pi}{2\,T}\,t} \quad \text{where} \quad c_n = \frac{1}{2\,T}\int_{-T}^{T} f(\tau)\, e^{-i\,n\,\frac{2\,\pi}{2\,T}\,\tau}\, d\tau\,.$$

Now examine the above formula as $T \to +\infty$. To start out assume that $f(t)$ is only different from 0 on a bounded interval. Then extend $f(t)$ with a period of $2\,T$ where $T$ is chosen larger and larger. This is illustrated in Figure 6.7.

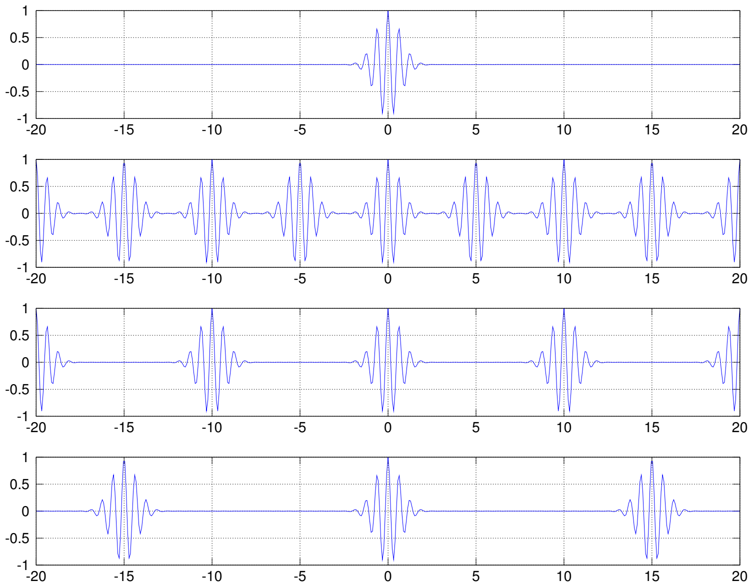



Figure 6.7: A localized function and its periodic extensions with periods 5, 10 and 15

Use

$$\Delta\sigma = \frac{1}{2\,T} \quad , \quad \sigma = n\,\Delta\sigma = \frac{n}{2\,T} \quad \text{and} \quad n\,\omega = 2\,\pi\,\sigma = 2\,\pi\,n\,\Delta\sigma$$

and the above formula can be rewritten in the form

$$f\,(t) \sim \sum_{n=-\infty}^{\infty} \left(\int_{\tau=-T}^{T} f\,(\tau)\; e^{-i\,2\,\pi\,\sigma\,\tau}\; d\tau\right)\; e^{i\,2\,\pi\,n\,\Delta\sigma\,t}\Delta\sigma\,.$$

The series is a numerical evaluation of an improper integral by using the values of the function at the points $n\,\Delta\tau$. For $T \to \infty$ we obtain formally for "nice" functions $f(t)$

$$f\,(t) \sim \int_{\sigma=-\infty}^{\infty} \left(\int_{\tau=-\infty}^{\infty} f\,(\tau)\; e^{-i\,2\,\pi\,\sigma\,\tau}\; d\tau\right)\; e^{i\,2\,\pi\,\sigma\,t}\; d\sigma\,.$$

This leads to the definition of the Fourier transform.

• **6–11 Definition:** Let $f : \mathbb{R} \to \mathbb{R}$ a piecewise continuous function with $\int_{-\infty}^{+\infty} |f(t)|^2 \, dt < \infty$. Then the new function

$$\mathcal{F}[f](\nu) = F(\nu) = \int_{-\infty}^{\infty} f(t) \, e^{-i\,2\,\pi\,\nu\,t} \, dt \tag{6.5}$$

is the **Fourier transform** of $f$. The **inverse Fourier transform** is given by

$$f(t) \sim \int_{-\infty}^{\infty} F(\nu) \, e^{i\,2\,\pi\,\nu\,t} \, d\nu \, .$$

Observe that the forward and inverse Fourier transform only differ in the sign of the complex exponent. ♢

The main idea and feature is that an "arbitrary" function $f(t)$ can be written as an improper integral of periodic functions $e^{i\,2\,\pi\nu t}$ with frequency $\nu$ and (possibly complex) amplitude $F(\nu)$. This is analogous to the idea of a Fourier series in equations (6.1) or (6.2). For continuous functions $f(t)$ the condition $\int_{\mathbb{R}} |f(t)|^2 \, dt < \infty$ implies $\lim_{t\to\pm\infty} f(t) = 0$.

There are multiple, slightly different definitions of the Fourier transform. The factor $2\,\pi$ in the complex exponential function might show at different spots and the combination of plus and minus signs in the forward and inverse transform might be swapped. Some presentations speak of a **Fourier correspondence** and write

$$f(t) \quad \circ\!\!-\!\!\bullet \quad F(\nu)$$

The **inverse Fourier transform** of a function $F(\nu)$ is given by

$$f(t) = \mathcal{F}^{-1}\,[\,F(\nu)\,](t) = \int_{-\infty}^{\infty} F(\nu) \, e^{i\,2\,\pi\,\nu\,t} \, d\nu \, .$$

### 6.4.2 Fourier transform, Fourier series and delta functions

For a $T$–periodic function $f(t)$ we have the Fourier series

$$f(t) \sim \sum_{n=-\infty}^{+\infty} c_n \, e^{i\,n\,\frac{2\pi}{T}\,t}$$

but a periodic function can not satisfy the condition $\int_{-\infty}^{\infty} |f(t)|^2 \, dt < \infty$, since the values will not converge to 0 as $|t| \to \infty$. The theory of Fourier transforms can be extended by using **delta–functions**. These are not functions, but **distributions**. Are careful introduction to distributions is well beyond the scope of these notes, we only use elementary computational rules for distributions.

• **6–12 Definition:** The delta function $\delta(t - t_0)$ at $t = t_0$ is defined by the property

$$\int_{-\infty}^{\infty} f(t) \, \delta\,(t - t_0) \, dt = f(t_0) \, .$$

This identity is correct for any continuous function $f(t)$. ♢

Based on this definition observe that

$$e^{i\,2\,\pi\,\nu_0\,t} = \int_{-\infty}^{\infty} \delta(\nu - \nu_0) \, e^{i\,2\,\pi\,\nu\,t} \, d\nu = \mathcal{F}^{-1}[\delta(\nu - \nu_0)]$$

and thus

$$\mathcal{F}[\,e^{i\,2\,\pi\,\nu_0\,t}\,]\,(\nu) = \delta(\nu - \nu_0) \quad \text{or} \quad e^{i\,2\,\pi\,\nu_0\,t} \circ\!\!-\!\!\bullet \; \delta(\nu - \nu_0) \, .$$

The delta function can also be characterized by this Fourier transform.

An „arbitrary“ $T$–periodic function is represented by its Fourier series

$$f(t) \sim \sum_{n=-\infty}^{\infty} c_n \, e^{i\,n\,\frac{2\pi}{T}\,t} \quad \text{where} \quad c_n = \frac{1}{T}\int_0^T f(\tau)\, e^{-i\,n\,\frac{2\pi}{T}\,\tau}\, d\tau\,.$$

Applying the Fourier transform to this formula and using its linearity we arrive at

$$F(\nu) = \mathcal{F}[f(t)]\,(\nu) = \sum_{n=-\infty}^{\infty} c_n\, \delta(\nu - \frac{n}{T})\,.$$

In this (mathematically imprecise) sense the sequence $c_n$ of Fourier coefficients is connected to a series of delta function, which represents to Fourier transform of the $T$–periodic function $f(t)$.

### 6.4.3 Properties of the Fourier transform and examples

- On can show that the Fourier transform preserves the $L_2$ norm, i.e.

$$\|f\|^2 = \int_{-\infty}^{+\infty} |f(t)|^2\, dt = \int_{-\infty}^{+\infty} |F(\nu)|^2\, d\nu = \|F\|^2\,.$$

  This is called **Parseval's identity**.

- The improper integrals in the Fourier transform hide some limits, e.g.

$$\int_{-\infty}^{+\infty} f(t)\, dt = \lim_{M\to+\infty}\int_{-M}^{+M} f(t)\, dt\,.$$

**• 6–13 Question:**
Verify the following statements:

(a) If the function $f(t)$ is even, i.e. $f(-t) = f(t)$, then

$$F(\nu) = \int_{-\infty}^{+\infty} f(t)\, \cos(2\,\pi\,\nu\,t)\, dt = 2\int_0^{+\infty} f(t)\, \cos(2\,\pi\,\nu\,t)\, dt \in \mathbb{R}$$

and the inverse Fourier transform is given by

$$f(t) \sim \int_{-\infty}^{+\infty} F(\nu)\, \cos(2\,\pi\,\nu\,t)\, d\nu = 2\int_0^{+\infty} F(\nu)\, \cos(2\,\pi\,\nu\,t)\, d\nu\,.$$

This result says that any even function $f(t)$ can be written as an integral of cos functions with different frequencies.

(b) If the function $f(t)$ is odd, i.e. $f(-t) = -f(t)$, then

$$F(\nu) = -i\int_{-\infty}^{+\infty} f(t)\, \sin(2\,\pi\,\nu\,t)\, dt = 2\int_0^{+\infty} f(t)\, \sin(2\,\pi\,\nu\,t)\, dt \in i\,\mathbb{R}$$

and the inverse Fourier transform is given by

$$f(t) \sim i\int_{-\infty}^{+\infty} F(\nu)\, \sin(2\,\pi\,\nu\,t)\, d\nu = 2\int_0^{+\infty} i\,F(\nu)\, \sin(2\,\pi\,\nu\,t)\, d\nu\,.$$

Since $F(\nu) \in i\mathbb{R}$ we have $i\,F(\nu) \in \mathbb{R}$. This result says that any odd function $f(t)$ can be written as an integral of sin functions with different frequencies.

| | $f(t)$ | $\circ\!\!-\!\!\bullet$ | $F(\nu)$ |
|---|---|---|---|
| Linearity | $\alpha\, f(t) + \beta\, g(t)$ | $\circ\!\!-\!\!\bullet$ | $\alpha\, F(\nu) + \beta\, G(\nu)$ |
| Scaling, similarity | $f(c\,t)$ | $\circ\!\!-\!\!\bullet$ | $\frac{1}{\vert c\vert}\, F(\frac{\nu}{c})$ |
| Shift in the time domain | $f(t-a)$ | $\circ\!\!-\!\!\bullet$ | $e^{-i\,2\,\pi\nu\,a}\, F(\nu)$ |
| Shift in the frequency domain | $e^{i\,2\,\pi\,\Omega\,t}\, f(t)$ | $\circ\!\!-\!\!\bullet$ | $F(\nu-\Omega)$ |
| Symmetry, even functions | $f(-t) = f(t)$ | $\Longleftrightarrow$ | $F(-\nu) = F(\nu)$ |
| Symmetry, odd functions | $f(-t) = -f(t)$ | $\Longleftrightarrow$ | $F(-\nu) = -F(\nu)$ |
| Conjugates | $\overline{f(t)}$ | $\circ\!\!-\!\!\bullet$ | $\overline{F(-\nu)}$ |
| Derivative in the time domain | $\frac{d}{dt}\, f(t)$ | $\circ\!\!-\!\!\bullet$ | $i\,2\,\pi\,\nu\, F(\nu)$ |
| Derivative in the frequency domain | $-i\,2\,\pi\; t\, f(t)$ | $\circ\!\!-\!\!\bullet$ | $\frac{d}{d\nu}\, F(\nu)$ |

Table 6.1: Properties of the Fourier transform

Based on the above one can define a Fourier cosine transform and a Fourier sine transform.

$\diamondsuit$

Equation (6.5) and some integral computations lead to Table 6.1 with a few properties of the Fourier transform.

As an example we verify $\mathcal{F}[f'](\nu) = i\,2\,\pi\,\nu\,F(\nu)$ for a smoot function $f(t)$ with $\lim_{t\to\pm\infty} f(t) = 0$.

$$\begin{aligned}
\mathcal{F}[f'](\nu) &= \int_{-\infty}^{\infty} f'(t)\, e^{-i\,2\,\pi\,\nu\,t}\, dt \quad \text{use intgeration by parts} \\
&= f(t)\, e^{-i\,2\,\pi\,\nu\,t}\Big|_{t=-\infty}^{+\infty} - \int_{-\infty}^{\infty} f(t)\,(-i\,2\,\pi\,\nu)\, e^{-i\,2\,\pi\,\nu\,t}\, dt \;=\; 0 + i\,2\,\pi\,\nu\,F(\nu)
\end{aligned}$$

• **6–14 Example:** Examine a rectangular window function $w(t)$ of width $2\,b$.

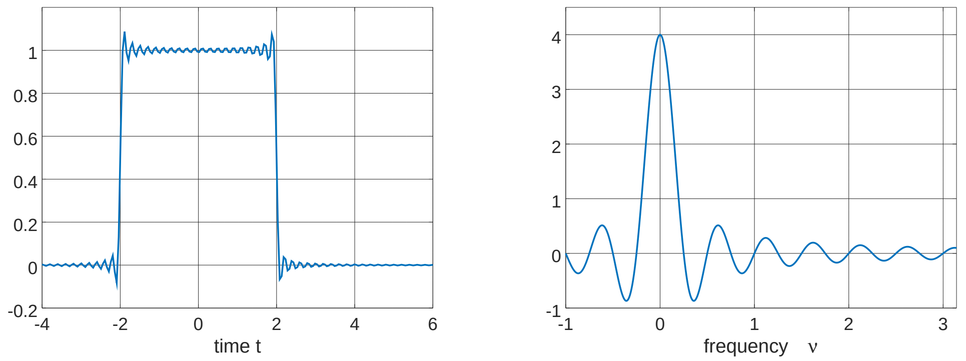
$$w\,(t) = \begin{cases} 1 & \text{for} \quad |t| \le b \\ 0 & \text{for} \quad |t| > b \end{cases}.$$

Its Fourier transform is given by

$$\begin{aligned}
W(\nu) &= \int_{-\infty}^{\infty} w(t)\, e^{-i\,2\,\pi\,\nu\,t}\, dt \;=\; \int_{-b}^{b} e^{-i\,2\,\pi\,\nu\,t}\, dt \\
&= \frac{1}{-i\,2\,\pi\,\nu} e^{-i\,2\,\pi\,\nu\,t}\Big|_{-b}^{b} = \frac{i}{2\,\pi\,\nu}\,(\cos(2\,\pi\,\nu\,t) - i\,\sin(2\,\pi\,\nu\,t))\,\Big|_{-b}^{b} \\
&= \frac{1}{\pi\,\nu}\,\sin(2\,\pi\,\nu\,b)\,.
\end{aligned}$$

The function $W(\nu)$ is of the type $\frac{\sin a\,\nu}{\nu}$, where the zeros closest to $\nu = 0$ are at $\nu = \pm\frac{1}{2\,b}$. We find that the product of the width $2\,b$ of the original function and the width $2\,\frac{1}{2\,b} = b$ leads to a constant $2\,b \cdot 2\,\frac{1}{2\,b} = 2$, i.e. a constant independent on $b$. Thus a wider rectangle will lead to a narrower spectrum. The limit $\frac{\sin x}{x} \to 1$ as $x \to 0$ implies

$$W(0) = \lim_{\nu\to 0} W(\nu) = \lim_{\nu\to 0} \frac{\sin(2\,\pi\,\nu\,b)}{2\,\pi\,\nu\,b}\, 2\,b = 2\,b\,.$$




(a) the window in the time domain (b) the spectrum

Figure 6.8: Fourier transform of a rectangular window and the Fourier reconstruction

The reconstruction of the original window function can be verified by the `MATLAB`/*Octave* code below. The Fourier transform $F(\nu)$ and the by a numerical integration reconstructed signal are shown in Figure 6.8. The phenomena of Gibbs is visible here too. The computations are performed for the value $b = 2$.

**FourierRect.m**

```
b = 2;
fsinc = @(nu)sin(b*2*pi*nu)./(nu*pi);

nu = linspace(-1,pi,500);
figure(1)
plot(nu,fsinc(nu))
xlabel('frequency \nu')
axis([-1 pi, -1 4.5])

t = linspace(-4,6,200); ft = t;
for ii = 1:length(t)
  ft(ii) = quad(@(nu)cos(2*pi*nu*t(ii)).*fsinc(nu),-5,5,1e-5,0);
end%for

figure(2)
plot(t,ft)
xlabel('time t'); axis([-4,6,-0.2,1.2])
```

$\diamondsuit$

The above example illustrates the basic convergence result for Fourier transforms.

• **6–15 Result:** Assume a function $f(t)$ and its derivative $f'(t)$ are piecewise continuous on each bounded interval and the improper integral

$$\int_{-\infty}^{\infty} |\, f(t)\,|\; dt < \infty$$

exists. Then the improper integral of the inverse Fourier transform

$$f\,(t) \sim \int_{-\infty}^{\infty} F(\nu)\; e^{i\,2\,\pi\,\nu\,t}\; d\nu$$

exists for all values of $t \in \mathbb{R}$. If the function $f(t)$ is continuous, then we have equality in the above statement. If the function is discontinuous at a point $x_0$, then the integral converges to the average of the left and right limit of $f$ at $x = x_0$ and the phenomena of Gibbs occurs. $\diamondsuit$

• **6–16 Example:** The Gaussian bell curve with standard deviation $\sigma$ and mean 0 is given by

$$f(t) = \frac{1}{\sigma\sqrt{2\pi}} \exp(-\frac{t^2}{2\sigma^2}) .$$

It has the key properties

$$1 = \int_{-\infty}^{+\infty} f(t)\,dt \quad \text{and} \quad 0 = \int_{-\infty}^{+\infty} t\, f(t)\,dt$$

(a) Verify that its Fourier transform $F(\nu)$ is given by

$$F(\nu) = \exp(-\frac{(\sigma\,2\pi)^2}{2}\,\nu^2) .$$

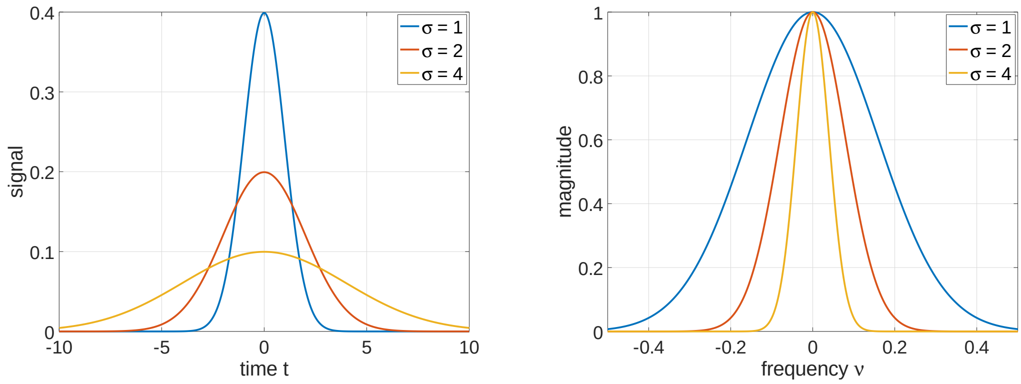
Thus it has the shape of a Gaussian curve, but with standard deviation $b = \frac{1}{2\pi\sigma}$ and the area under the curve is

$$\int_{-\infty}^{+\infty} F(\nu)\,d\nu = \frac{\sigma\,2\pi}{\sqrt{2\pi}} = \sigma\sqrt{2\pi} .$$

(b) Based on the above observe that the product of the standard deviations of the Gauss curve $f(t)$ and its Fourier transform $F(\nu)$ is constant, i.e. independent on $\sigma$.

$$b \cdot \sigma = \frac{1}{2\pi\sigma} \cdot \sigma = \frac{1}{2\pi}$$

Thus it is impossible to obtain Gauss signals that are at the same time localized in the time domain and in the frequency domain. This is illustrated in Figure 6.9.




(a) the signals (b) the spectra

Figure 6.9: Signals and spectra for Gaussian bell curves with different standard deviations $\sigma$

(c) The above also implies

$$\mathcal{F}^{-1}[\exp(-\frac{4\pi^2\sigma^2}{2}\,\nu^2)] = \frac{1}{\sigma\sqrt{2\pi}} \exp(-\frac{t^2}{2\sigma^2}) . \tag{6.6}$$

(d) Based on the computational rule in Table 6.1 on shifts in the time domain

$$f(t-a) \qquad \circ\!\!-\!\!\bullet \qquad e^{-i\,2\,\pi\,\nu\,a}\;F(\nu)$$

we can examine the Fourier transform of a Gauss curve with mean at $t = a$ and standard deviation $\sigma$.

$$\mathcal{F}[f(t-a)](\nu) = \mathcal{F}[\frac{1}{\sigma\,\sqrt{2\,\pi}}\,\exp(-\frac{(t-a)^2}{2\,\sigma^2})](\nu) = e^{-i\,2\,\pi\nu\,a}\;\exp(-\frac{(\sigma\,2\,\pi)^2}{2}\,\nu^2)\,.$$

This implies in particular the the spectrum ($|F(\nu)|$ as function of $\nu$) is independent on the shift $a$.

**Solution:**

(a) Use a straightforward integration and a completion of the square.

$$\begin{aligned}
F\,(\nu) &= \int_{-\infty}^{\infty} f\,(t)\; e^{-i\,2\,\pi\,\nu\,t}\;dt = \frac{1}{\sigma\,\sqrt{2\,\pi}}\int_{-\infty}^{\infty}\exp(-\frac{t^2}{2\,\sigma^2})\;\exp(-i\,2\,\pi\,\nu\,t)\;dt\\
&= \frac{1}{\sigma\,\sqrt{2\,\pi}}\int_{-\infty}^{\infty}\exp(-\frac{1}{2\,\sigma^2}\,(t^2+2\,i\,\sigma^2\,2\,\pi\,\nu\,t))\;dt\\
& \qquad t^2+2\,i\,\sigma^2\,2\,\pi\,\nu\,t = (t+i\,\sigma^2\,2\,\pi\,\nu)^2-(i\,\sigma^2\,2\,\pi\nu)^2 \quad \text{complete the square}\\
&= \frac{1}{\sigma\,\sqrt{2\,\pi}}\int_{-\infty}^{\infty}\exp(-\frac{1}{2\,\sigma^2}\,(t+i\,\sigma^2\,2\,\pi\,\nu)^2)\;dt\;\exp(-\frac{1}{2\,\sigma^2}\,(\sigma^2\,2\,\pi\nu)^2)\\
& \qquad \text{substitution} \quad s = t+i\,\sigma^2\,2\,\pi\,\nu\\
&= \frac{1}{\sigma\,\sqrt{2\,\pi}}\int_{-\infty}^{\infty}\exp(-\frac{1}{2\,\sigma^2}\,s^2)\;ds\;\exp(-\frac{1}{2}\,(\sigma\,2\,\pi\nu)^2)\\
&= \exp(-\frac{(\sigma\,2\,\pi)^2}{2}\,\nu^2)
\end{aligned}$$

$\diamondsuit$

The above simple form of the spectra changes if a sum of shifted Gauss curves is examined. As example consider

$$f(t) = 0.5\cdot\text{Gauss}(t+6,1)+2\cdot\text{Gauss}(t,2)+\text{Gauss}(t-10,4)$$

where

$$\text{Gauss}(t-\text{shift},\sigma) = \frac{1}{\sigma\,\sqrt{2\,\pi}}\,\exp(-\frac{(t-\text{shift})^2}{2\,\sigma^2})\,.$$

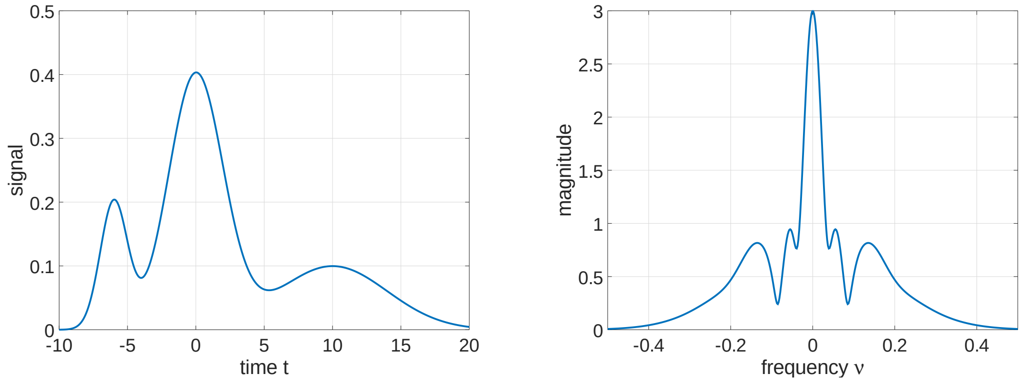
Find the graph of this function $f(t)$ and its spectrum $|F(\nu)|$ in Figure 6.10. Since the function $f(t)$ is real valued, we know $F(-\nu) = \overline{F(+\nu)}$, leading to the visible symmetry $|F(-\nu)| = |F(+\nu)|$ in the spectrum in Figure 6.10.

**• 6–17 Question:**

Examine $f(t)$ a piecewise continuous function, differentiable on the subintervals, with discontinuities at $\tau_k$ for $k = 1, 2, \ldots n$. Then we have

$$f'(t) \qquad \circ\!\!-\!\!\bullet \qquad i\,2\,\pi\,\nu\;F\,(\nu)-\sum_{k=1}^{n}\left(f\,(\tau_k+)-f\,(\tau_k-)\right)\;e^{-i\,2\,\pi\,\tau_k}\,.$$

Tip: integration by parts on the subintervals $[\tau_k, \tau_{k+1}]$.




(a) the signal $f(t)$ (b) the spectum $|F(\nu)|$

Figure 6.10: Signal and spectrum for a linear combination of shifted Gaussian curves

◊

For two functions $f(t)$ und $g(t)$ their **convolution** is given by

$$h\,(t) = (f * g)\,(t) = \int_{-\infty}^{\infty} f\,(t-\tau)\,g\,(\tau)\;d\tau\,.$$

One can verify that the convolution is commutative, i.e. $f * g = g * f$ and we find

$$\begin{aligned}
H\,(\nu) &= \mathcal{F}[\,f * g]\,(\nu) \\
&= \int_{-\infty}^{\infty} e^{-i\,2\,\pi\,\nu\,t}\,\left(\int_{-\infty}^{\infty} f\,(t-\tau)\,g\,(\tau)\;d\tau\right)\;dt \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-i\,2\,\pi\,\nu\,t}\,f\,(t-\tau)\,g\,(\tau)\;d\tau\;dt \\
&\quad \text{substitution } u = t-\tau \\
&= \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-i\,2\,\pi\,\nu\,(u+\tau)}\,f\,(u)\,g\,(\tau)\;d\tau\;du \\
&= \int_{-\infty}^{\infty} e^{-i\,2\,\pi\,\nu\,u}\,f\,(u)\;du\cdot\int_{-\infty}^{\infty} e^{-i\,2\,\pi\,\nu\,\tau}\,g\,(\tau)\;d\tau \\
&= F\,(\nu)\cdot G\,(\nu)\,.
\end{aligned}$$

An identical convolution can also be applied to the Fourier transform in the frequency space, i.e.

$$(F * G)(\nu) = (G * F)(\nu) = \int_{-\infty}^{+\infty} F(\nu-\mu)\,G(\mu)\;d\mu\,.$$

This leads to the convolution theorem for Fourier transforms.

**• 6–18 Result: Convolution theorem**

Examine two functions $f$ and $g$ such that their Fourier transforms exist. Then we have

$$\begin{aligned}
\mathcal{F}[f * g] &= \mathcal{F}[f]\cdot\mathcal{F}[g] \\
\mathcal{F}[f] * \mathcal{F}[g] &= \mathcal{F}[f\cdot g]
\end{aligned}$$

or with a more compact notation $F = \mathcal{F}[f]$ and $G = \mathcal{F}[g]$

$$\mathcal{F}[f * g] = F\cdot G \quad\text{and}\quad F * G = \mathcal{F}[f\cdot g]\,.$$

◊

The above convolution has an important consequence on windowing effects. If a function $f(t)$ originally defined for $t \in \mathbb{R}$ is only measured for $-b \leq t \leq +b$ the result is described by a windowing function $w$ given in Example 6–14.

$$w(t) = \begin{cases} 1 & \text{for} \quad |t| \leq b \\ 0 & \text{for} \quad |t| > b \end{cases} .$$

Thus the original function is modified.

$$\begin{aligned} f(t) &\longrightarrow w(t) \cdot f(t) \\ F(\nu) &\longrightarrow \mathcal{F}[w(t) \cdot f(t)](\nu) \;=\; (W * F)(\nu) \end{aligned}$$

and we have the modified Fourier transform

$$\mathcal{F}[w(t) \cdot f(t)](\nu) = \int_{-\infty}^{+\infty} W(\mu)\, F(\nu - \mu)\, d\mu = \int_{-\infty}^{+\infty} \frac{\sin(2\,\pi\,\mu\,b)}{\pi\,\mu}\, F(\nu - \mu)\, d\mu\,.$$

The graph of $W(\mu)$ is shown in Figure 6.8. As a consequence the spectrum $F(\nu)$ of the original function $f(t)$ is smeared out by the spectrum $W(\nu)$ of the window function $w(t)$. A similar effect shows up with DFT and FFT. As an example consider the function

$$f(t) = 2\,\cos(40 \cdot 2\,\pi\,t) + \sin(35 \cdot 2\,\pi\,t)$$

on the interval $[0\,,\,1]$. This function has contributions with frequencies $\nu_1 = 35$ Hz and $\nu_2 = 40$ Hz. If we apply a window of width $2\,b = 0.3$ the spectrum is modified severely, as shown in Figure 6.11. Thus one has to be very careful when choosing the correct window, size and type. Some windows used frequently are the above rectangular window, and the Hamming and Hanning windows.

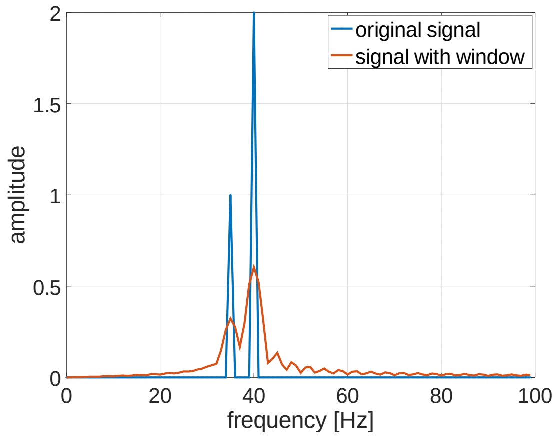



Figure 6.11: The spectra of a function with two trigonometric contributions, with and without windowing effect

• **6–19 Example:** Examine the initial value problem, describing a dynamic heat conduction problem.

$$\begin{aligned} \frac{\partial}{\partial t}\, u(t,x) &= \alpha\, \frac{\partial^2}{\partial x^2}\, u(t,x) \quad \text{for } t > 0 \text{ and } x \in \mathbb{R} \\ u(0,x) &= u_0(x) \quad \text{for } x \in \mathbb{R} \end{aligned}$$

(a) Verify the the Fourier transform of the solution $u(t,x)$ with respect to $x$ satisfies an ordinary differential equation of the the form $\dot{y}(t) = -c\,y(t)$ for each frequency $\nu$.

(b) If the initial value is a Gauss curve

$$u(0,x) = u_0(x) = \frac{1}{\sigma\,\sqrt{2\,\pi}}\exp(-\frac{x^2}{2\,\sigma^2})$$

then the solution is given by

$$u(t,x) = \frac{1}{\sqrt{\sigma^2 + 2\,\alpha\,t}\,\sqrt{2\,\pi}}\;\exp(-\frac{x^2}{2\,(\sigma^2 + 2\,\alpha\,t)})\,.$$

(c) Take the limit as $\sigma \to 0+$ to find the fundamental solution (5.8) shown on page 62.

$$\Phi(t,x) = \frac{1}{\sqrt{4\,\alpha\,\pi\,t}}\;\exp(-\frac{x^2}{4\,\alpha\,t})\,.$$

With this fundamental solution an integral formula for the solution of the initial value problem with arbitrary $u_0(x)$ can be given, see equation (5.9) on page 63.

**Solution:**

(a) Apply a Fourier transform with respect to $x \in \mathbb{R}$ with the notation $U(t,\nu) = \mathcal{F}[u(t,x)](\nu)$. Then use the computational rules in Table 6.1 to conclude

$$\begin{aligned}
\frac{\partial}{\partial t}\,U(t,\nu) &= \alpha\,(i\,2\,\pi\,\nu)^2\,U(t,\nu) \;=\; -\alpha\,4\,\pi^2\,\nu^2\,U(t,\nu)\\
U(t,\nu) &= U(0,\nu)\,\exp(-\alpha\,4\,\pi^2\,\nu^2\,t)\,.
\end{aligned}$$

(b) Use equation (6.6) for the inverse Fourier transform.

$$\begin{aligned}
U(t,\nu) &= U(0,\nu)\,\exp(-\alpha\,4\,\pi^2\,\nu^2\,t)\\
&= \exp(-\frac{(\sigma\,2\,\pi)^2}{2}\,\nu^2)\;\exp(-\alpha\,4\,\pi^2\,\nu^2\,t) \;=\; \exp(-\frac{4\,\pi^2\,(\sigma^2 + 2\,\alpha\,t)}{2}\,\nu^2)\\
u(t,x) &= \mathcal{F}^{-1}[U(t,\nu)](x) \;=\; \frac{1}{\sqrt{\sigma^2 + 2\,\alpha\,t}\,\sqrt{2\,\pi}}\;\exp(-\frac{x^2}{2\,(\sigma^2 + 2\,\alpha\,t)})
\end{aligned}$$

$\diamondsuit$

# Chapter 7

# Probability Theory

## 7.1 Keywords and Literature

Event, sample space, probability measure, Laplace experiment, independent events, conditional probability, Bayes' theorem, law of total probability, random variable, expected value, variance, probability mass function, density function, cumulative distribution function, quantile function, binomial distribution, Poisson distribution, uniform distribution, normal/Gaussian distribution, exponential distribution, joint probability mass function, joint density, marginal and conditional distribution, independent random variables, covariance, correlation.

Literature:

- Any good book on basic probability theory should do the job.
- The book by Montgomery and Runger [MontRung03] covers the necessary topics.
- The books by Lothar Papula cover all required topics, but these books are in German.

## 7.2 Basic Concepts

• **7–1 Definition:**

- An **elementary event** $\omega$ is a possible outcome of an experiment.
- The **sample space** $\Omega$ is the set of all elementary events.
- An **event** $A$ is a subset of the sample space ($A \subset \Omega$).
- The **$\sigma$-algebra** $\mathcal{F}$ is the collection of all events considered.

$\diamondsuit$

• **7–2 Definition:** Let $\Omega$ be a sample space and $\mathcal{F}$ be a $\sigma$-algebra. A **probability measure** is a function $\mathrm{P} : \mathcal{F} \to [0, 1]$ that assigns a value between 0 and 1 to an event $A \subset \Omega$: $\mathrm{P}(A) \in [0, 1]$.
It obeys the following properties (axioms of Kolmogorov):

- $0 \leq \mathrm{P}(A) \leq 1$ for every event $A \subset \Omega$
- $\mathrm{P}(\Omega) = 1$
- $\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B)$ for *disjoint* (mutually exclusive) events $A$ and $B$

$\diamondsuit$

• **7–3 Example:** Events and set operations can be visualised using **Venn diagrams**:

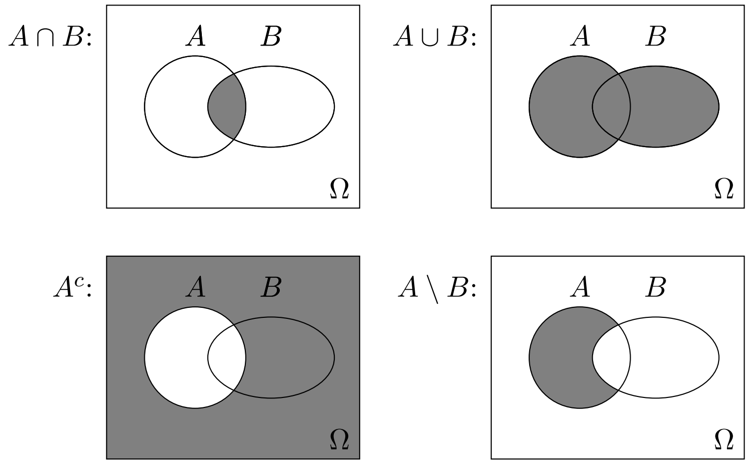



Figure 7.1: Venn diagrams

♢

• **7–4 Result: De Morgan's laws**
Let $A$ and $B$ be events. Then, $(A \cap B)^c = A^c \cup B^c$ and $(A \cup B)^c = A^c \cap B^c$. ♢

• **7–5 Question:**
Prove de Morgan's laws graphically using Venn diagrams.

♢

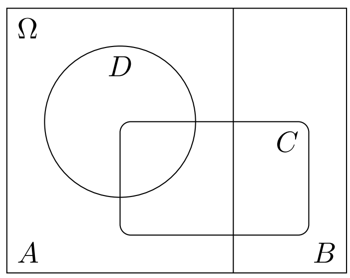
• **7–6 Question:**
$A$, $B$ and $C$ are events.

(a) Which of the following statements are meaningful?

i) $\mathrm{P}(A \cup (B \cap C))$
ii) $\mathrm{P}(A) + \mathrm{P}(B)$
iii) $\mathrm{P}(A^c) \cap \mathrm{P}(B)$
iv) $(\mathrm{P}(B))^c$

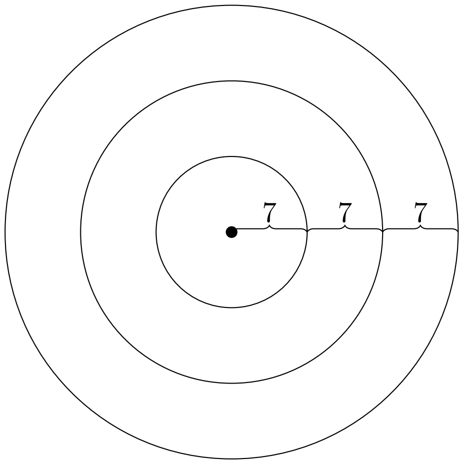
(b) Display the following events in the given diagram.

i) $C \cap D$
ii) $(D \setminus C) \cup (C \cap A)$
iii) $B \cup D$




♢

• **7–7 Question:**
You throw darts at a dartboard consisting of three circular discs (length specifications in centimetres):




On average, every second dart hits the dartboard. Further, the hit probabilities for the single circular discs are proportional to their areas. Let $A$, $B$, $C$ describe the events „the innermost“, „the middle“, „the outermost circular disc is hit“. (Note: If the innermost circular disc is hit, then the middle as well as the outermost circular discs are hit simultaneously.)

(a) Calculate the hit probability for the innermost, the middle and the outermost circular disc.

(b) Calculate the probability to hit the middle circular *ring*.

♢

• **7–8 Result:** Let $A$ and $B$ be events. Then, $\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A \cap B)$.
More general: let $A_1, A_2, \ldots, A_n$ be events. Then,

$$\begin{aligned}\mathrm{P}(A_1 \cup A_2 \cup \ldots \cup A_n) &= \sum_{i_1=1}^{n} \mathrm{P}(A_{i_1}) - \sum_{i_1=1}^{n-1} \sum_{i_2=i_1+1}^{n} \mathrm{P}(A_{i_1} \cap A_{i_2}) \\ &\quad + \sum_{i_1=1}^{n-2} \sum_{i_2=i_1+1}^{n-1} \sum_{i_3=i_2+1}^{n} \mathrm{P}(A_{i_1} \cap A_{i_2} \cap A_{i_3}) - \ldots\end{aligned}$$

♢

• **7–9 Definition:** In a **Laplace experiment** every elementary event is equally probable, i.e. $\mathrm{P}(\{\omega_i\}) = 1/|\Omega|$. As a consequence, the probability of an event $A \subset \Omega$ in a Laplace experiment can be calculated as follows:

$$\mathrm{P}(A) = \frac{\#\text{ favourable cases}}{\#\text{ possible cases}} = \frac{|A|}{|\Omega|}$$

♢

• **7–10 Question:**
In a random experiment two dice are thrown simultaneously. (Assume that the numbers 1 to 6 have equal probability.)

(a) Describe the sample space of the elementary events.

(b) What is the probability of a single elementary event?

(c) Calculate the probability of the event $E_1 =$ „the sum of the spots is 7“.

(d) What is the probability that event $E_2 =$ „the sum of spots is smaller than 4“ occurs?

(e) Determine $P(E_3)$ of the event $E_3$ = „both spots are odd".

(f) Calculate $P(E_2 \cup E_3)$.

$\diamondsuit$

• **7–11 Question:**
Calculate the following probabilities.

(a) You draw randomly one of the 26 letters of the alphabet. What is the probability that you draw a vowel?

(b) You toss a fair coin three times. What is the probability that you get head at least once?

(c) A wooden dice with a green surface is cut into $10^3 = 1000$ small dices of equal size. What is the probability for a randomly chosen (small) dice to have exactly two green cube faces?

$\diamondsuit$

• **7–12 Definition:** Two events $A$ and $B$ are called **independent**, if $P(A \cap B) = P(A) \cdot P(B)$.

$\diamondsuit$

• **7–13 Question:**
On Fridays Bonnie and Clyde often miss class. Bonnie is absent with probability 0.3 and Clyde with probability 0.45. The probability that both of them attend class is only 0.4. Are the attendances of Bonnie and Clyde independent events?

$\diamondsuit$

• **7–14 Question:**
An unskilled hunter hits his target only with a probability of 20%. He fires three times at a rabbit. However, the rabbit knows the hunter. The rabbit remains quiet and sits still, as he believes that the probability of being shot is less than 50%. Has the rabbit calculated the probability correctly?

$\diamondsuit$

• **7–15 Definition:** Let $A$ and $B$ be events (with $P(B) > 0$). The **conditional probability of $A$ given $B$** is defined as

$$P(A|B) = \frac{P(A \cap B)}{P(B)} .$$

$\diamondsuit$

• **7–16 Result: Law of total probability**
Assume $B_1, B_2, \ldots, B_k$ are disjoint events with $B_1 \cup B_2 \cup \ldots \cup B_k = \Omega$. Then, the probability of any event $A$ is

$$P(A) = \sum_{i=1}^{k} P(A \cap B_i) = \sum_{i=1}^{k} P(A|B_i)P(B_i) .$$

$\diamondsuit$

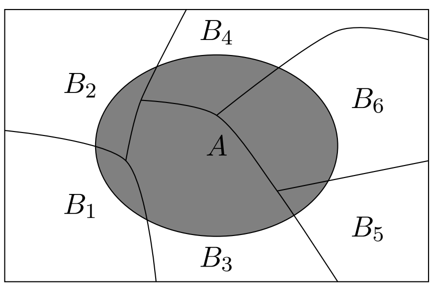



Figure 7.2: Law of total probability

**• 7–17 Result: Bayes' theorem**
Let $A$ and $B$ be events with $\mathrm{P}(A) > 0$ and $\mathrm{P}(B) > 0$. Then, we have

$$\mathrm{P}(B|A) = \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(A)} = \frac{\mathrm{P}(A|B)\mathrm{P}(B)}{\mathrm{P}(A)} \ .$$

In the setting of the law of total probability, we have

$$\mathrm{P}(B_i|A) = \frac{\mathrm{P}(A \cap B_i)}{\mathrm{P}(A)} = \frac{\mathrm{P}(A|B_i)\mathrm{P}(B_i)}{\sum_{j=1}^{k} \mathrm{P}(A|B_j)\mathrm{P}(B_j)} \ .$$

♢

**• 7–18 Question:**
You roll a die. What is the probability that

(a) you roll a six, given that the spots are even?

(b) you roll a six, given that the spots are odd?

(c) you roll even spots, given that the spots are smaller than four?

♢

**• 7–19 Question:**
600 out of a sample of 900 people got vaccinated prophylactically against influenza. After a given time, it was analysed who actually got the flu. You find the results in the following table:

| Group | $B$ (sick) | $B^c$ (healthy) | sum |
|---|---|---|---|
| $A$ (vaccinated) | 60 | 540 | 600 |
| $A^c$ (not vaccinated) | 120 | 180 | 300 |
| sum | 180 | 720 | 900 |

We define the events $A$ = „person is vaccinated" and $B$ = „person got sick". Calculate the following probabilities and describe the meaning of the corresponding events.

(a) $\mathrm{P}(A)$
(b) $\mathrm{P}(B)$
(c) $\mathrm{P}(A \cap B)$
(d) $\mathrm{P}(B|A)$
(e) $\mathrm{P}(A|B)$
(f) $\mathrm{P}(A^c \cap B)$
(g) $\mathrm{P}(B|A^c)$

♢

**• 7–20 Question:**
30% of the Swiss population hold a bachelor or a comparable degree. In this population subgroup the unemployment rate is 5%. For the rest of the Swiss population the unemployment rate is 10%.

(a) What is the probability for a Swiss to be unemployed?

(b) From a specific Swiss guy you know that he is unemployed. What is the probability that he holds a bachelor or a comparable degree?

(c) Are „holding a bachelor or comparable degree" and „being unemployed" independent events?

♢

## 7.3 Random Variables

A random variable is a variable that takes numerical values which depend on the outcome of a random experiment. It is an assignment of an event to a real number.

• **7–21 Definition:** A **random variable** $X$ is a function mapping a sample space $\Omega$ to $\mathbb{R}$ (or a subset or $\mathbb{R}$):

$$X : \Omega \to \mathbb{R}$$

$\diamondsuit$

*Note*: Random variables are denoted by capital letters (e.g. $X$), whereas a realised value of a random variable is denoted by a lower case letter (e.g. $x$):

- Capital letter: description of an experiment (e.g., „measurement of the length of a tibia“)
- Lower case letter: outcome of the experiment (e.g., 38.5 cm)

• **7–22 Definition:** The **cumulative distribution function** (CDF) of a random variable $X$ is defined as $F_X(x) := \mathrm{P}(X \leq x)$. $\diamondsuit$

Properties of a CDF $F_X$:

- $F_X$ is monotonically increasing
- $\lim\limits_{x\to-\infty} F_X(x) = 0$, $\lim\limits_{x\to\infty} F_X(x) = 1$
- $\mathrm{P}(a < X \leq b) = F_X(b) - F_X(a)$

• **7–23 Definition:** Let $X$ be a random variable with distribution function $F_X$ and let $\alpha \in (0, 1)$. The **$\alpha$-quantile** of $X$ fulfils

$$\mathrm{P}(X \leq q) \geq \alpha \qquad \text{and} \qquad \mathrm{P}(X \geq q) \geq 1 - \alpha.$$

$\diamondsuit$

Roughly speaking, an $\alpha$-quantile $q$ is a value where the graph of the CDF crosses (or jumps over) $\alpha$. There is an inverse relation between the quantiles and the values of the CDF. In fact, the quantile function is given by the (generalised) inverse function of the CDF.

### 7.3.1 Discrete Random Variables

A discrete random variable $X$ has a finite (or countable) image (i.e. set of possible values):

$$X : \Omega \to \{x_1, x_2, \ldots\}$$

It is characterized by its probability mass function.

• **7–24 Definition:** The **probability mass function** $p(x_k) := \mathrm{P}(X = x_k)$ of a a (discrete) random variable has the following properties:

- For each set $A \subset \{x_1, x_2, \ldots\}$, we have

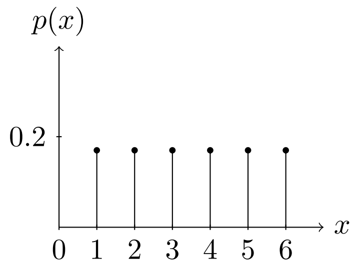
$$\mathrm{P}(X \in A) = \sum_{k:x_k \in A} p(x_k)$$

- Normalisation: $\sum\limits_{k} p(x_k) = 1$
- Connection to CDF: $F_X(x) = \mathrm{P}(X \leq x) = \sum\limits_{k:x_k \leq x} p(x_k)$

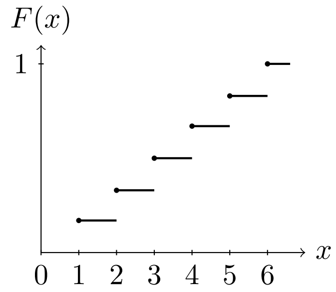
$\diamondsuit$

• **7–25 Example:** A die can take values in $\{1,2,\ldots,6\}$; if it is fair, it takes all values with the same probability. Its probability mass function and CDF look as follows:

**Probability mass function**




**Cumulative distribution function**




♢

• **7–26 Question:**
The random variable $X$ has the following probability mass function:

| $x_i$ | $-3$ | $-1$ | $2$ | $5$ |
|---|---|---|---|---|
| $\mathrm{P}(X = x_i)$ | $3a^2$ | $a$ | $5a$ | $4a^2$ |

Calculate $a$.

♢

• **7–27 Definition:** The **expected value** of a discrete random variable $X$ is defined as

$$\mathrm{E}(X) := \sum_k x_k \cdot p(x_k)\ .$$

*Interpretation*: The expected value $\mathrm{E}(X)$ corresponds to the weighted mean of all possible values of the random variable $X$. The weights are determined by the probability mass function. ♢

• **7–28 Definition:** The **variance** of a discrete random variable $X$ is defined as

$$\mathrm{Var}(X) := \mathrm{E}\left((X - \mathrm{E}(X))^2\right) = \sum_k (x_k - \mathrm{E}(X))^2 \cdot p(x_k)\ .$$

*Interpretation*: The variance of a random variable gives us an idea how strongly the values vary around the expected value (on average). ♢

• **7–29 Question:**
You're being offered the following game: You roll a die and if the number of spots is even, you win the same amount in CHF – otherwise you lose the corresponding amount. Are you gonna play, i.e. do you expect to win?

♢

• **7–30 Question:**
A **discrete** uniformly distributed random variable takes all possible values with equal probability, e.g. $X = \#$ spots when rolling a fair die. All possible values 1,2,…,6 have equal probability 1/6.

(a) Calculate the expected value $\mathrm{E}(X)$ and the variance $\mathrm{Var}(X)$.

(b) Now, generalise the considerations made in (a). What are the expected value and the variance of a discrete uniformly distributed random variable $X$, which takes the values $1,2,\ldots,n$ with equal probability?

*Hint*: The following formulas will help:

$$\sum_{i=1}^{n} i = \frac{n(n+1)}{2} \qquad \text{and} \qquad \sum_{i=1}^{n} i^2 = \frac{n(n+1)(2n+1)}{6}.$$

♢

• **7–31 Question:**
You're being offered the following dice game with two tetrahedra (four-sided dice with number of spots 1 to 4): Show both of the dice the same number of spots, you get five times your stake (i.e. you win 4 times your stake). Otherwise you loose $x$ times your stake, if the difference between the number of spots is $x$.

(a) Calculate the expected value and the variance of your gain.

(b) What would be a fair payout for the described game (indicated as multiple of your stake)?

♢

We will now consider three discrete probability distributions widely used.
• **7–32 Definition:** A discrete random variable $X$ that can only take the values 0 and 1 is said to be **Bernoulli distributed**. The distribution is specified by the probability $\pi := \mathrm{P}(X = 1)$.
We write $X \sim \mathrm{Bernoulli}(\pi)$. ♢
• **7–33 Definition:** A discrete random variable $X \in \{0, 1, \ldots, n\}$ is **binomially distributed**, if

$$p(x) = \mathrm{P}(X = x) = \binom{n}{x} \pi^x (1 - \pi)^{n-x} .$$

We write $X \sim \mathrm{Bin}(n,\pi)$, $n \in \mathbb{N}$, $\pi \in (0, 1)$. ♢

Properties:

- the sum of independent Bernoulli random variables is a binomially distributed random variable
- a binomial random variable models the number of „successes“ of $n$ independent trials with individual success probability $\pi$
- expected value $\mathrm{E}(X) = n\pi$
- variance $\mathrm{Var}(X) = n\pi(1 - \pi)$

In Figure 7.3 you can see the probability mass function of binomial distributions $\mathrm{Bin}(30,\pi)$ for different „success“ probabilities $\pi$:

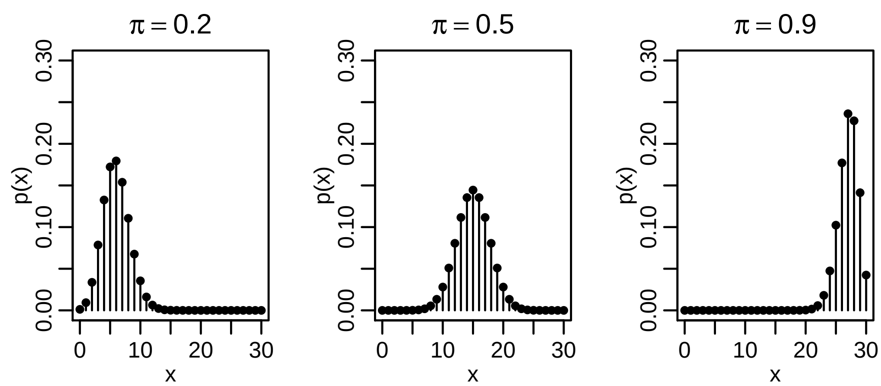



Figure 7.3: Probability mass function of different binomial distributions

• **7–34 Question:**
A friend of you tosses a coin 12 times. He bets on getting exactly six times tail. What is his winning probability?

♢

**• 7–35 Question:**
For an inspection, water samples (10 ml) are tested for contamination. As only 2% of all samples are contaminated, it's proposed to mix ten samples together. From every sample, 5 ml are mixed into a collective sample containing 50 ml. Now the collective sample is tested for contamination. If the sample is not contaminated, the inspection for the ten samples is over. Otherwise, the 10 single samples must be tested separately.

(a) What is the probability that there is no contamination in the collective sample (50 ml) (assuming the ten single samples are independent)?

(b) Let the random variable $Y$ be the number of analyses needed. Which are the possible values for $Y$? Calculate the probability mass function of $Y$.

(c) How many analyses need to be done on average (what is the expected value of $Y$)? How many analyses can be saved by mixing the samples into the collective sample, on average?

♢

**• 7–36 Definition:** A discrete random variable $X \in \mathbb{N}_0$ is **Poisson distributed** with parameter $\lambda$ if

$$p(x) = \mathrm{P}(X = x) = \mathrm{e}^{-\lambda} \frac{\lambda^x}{x!} \ .$$

We write $X \sim \mathrm{Po}(\lambda)$, $\lambda > 0$. ♢

Properties:

- a Poisson random variable models the number of „successes“ during a given time interval
- under the Poisson distribution it is assumed that events are independent and occur at a constant rate $\lambda$
- the binomial distribution converges to a Poisson distribution for $n \to \infty$ and $\mathrm{E}(x) = n\pi = \lambda$, where $\lambda$ is a constant
- expected value $\mathrm{E}(X) = \lambda$
- variance $\mathrm{Var}(X) = \lambda$

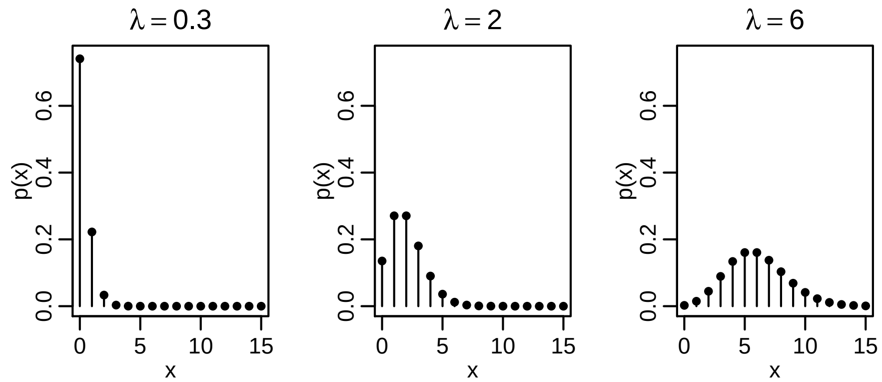



Figure 7.4: Probability mass function of different Poisson distributions

**• 7–37 Question:**
In a city two serious accidents happen per week on average. In particular, we assume that the number of serious accidents is Poisson distributed.

(a) Calculate the probability that more than five serious accidents happen per week

(b) Calculate the probability that more than one serious accident happens per *day*.

♢

• **7–38 Question:**
A technical support centre is being called 12 times per hour on average. We assume that the number of calls is Poisson distributed. Calculate the probability that the support centre

(a) gets no calls during the next five minutes.

(b) gets at least 5 calls during the next 10 minutes.

(c) gets at most 6 calls during the next 20 minutes.

♢

• **7–39 Question:**
A factory produces screws which are defective with probability $p = 0.001$. Calculate the probability that a batch of 500 screws contains at least two defective ones.

(a) Calculate the *exact* probability using the binomial distribution.

(b) Calculate the *approximate* probability using the Poisson distribution.

♢

### 7.3.2 Continuous Random Variables

A continuous random variable's image is $\mathbb{R}$ (or an interval $[a, b] \subset \mathbb{R}$):

$$X : \Omega \to \mathbb{R}$$

It is characterized by its probability density function.
• **7–40 Definition:** A function $f_X : \mathbb{R} \to [0, \infty)$ is called a **probability density function** (PDF), if

$$\int_{-\infty}^{\infty} f_X(x)\,\mathrm{d}x = 1.$$

The PDF is the derivative of the corresponding CDF:

$$f(x) := \frac{\mathrm{d}}{\mathrm{d}x} F_X(x)$$

For any interval $[a, b] \subset \mathbb{R}$

$$\mathrm{P}(a \leq X \leq b) = \int_a^b f_X(x)\,\mathrm{d}x = F_X(b) - F_X(a).$$

(This connection is illustrated in Figure 7.5.)
It follows directly that $\mathrm{P}(X = x) = 0$ for all $x \in \mathbb{R}$. ♢

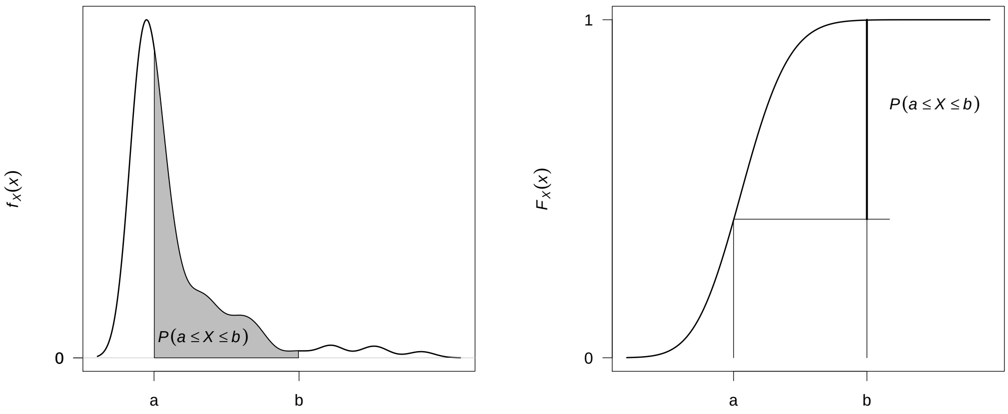



Figure 7.5: Connection between PDF and CDF of a continuous random variable

- **7–41 Definition:** For a continuous random variable $X$ we have:
  - **Expected value**: $\mathrm{E}(X) = \int_{\mathbb{R}} x \cdot f_X(x)\,\mathrm{d}x$
  - **Variance**: $\mathrm{Var}(X) = \int_{\mathbb{R}} (x - \mathrm{E}(X))^2 \cdot f_X(x)\,\mathrm{d}x$

$\diamondsuit$

- **7–42 Question:**

Show that the equality

$$\mathrm{Var}(X) = \mathrm{E}(X^2) - (\mathrm{E}(X))^2$$

holds for a discrete, as well as for a continuous random variable $X$.

$\diamondsuit$

- **7–43 Question:**

Let

$$f_X(x) = \begin{cases} ax, & \text{for } x \in [0,1] \\ 0, & \text{else} \end{cases}$$

with $a \in \mathbb{R}_+$.

(a) Determine the parameter $a$ such that $f_X$ is a probability density function.

(b) Calculate the corresponding CDF $F_X$ and plot the two functions in one graph.

(c) Let $X$ be a random variable with the given density function $f_X$. Calculate the following probabilities and visualise the results in your plot.

$$\mathrm{P}\left(\frac{1}{3} \leq X \leq \frac{3}{4}\right), \quad \mathrm{P}\left(X \leq \frac{1}{2}\right), \quad \mathrm{P}\left(X \geq \frac{3}{4}\right)$$

(d) Calculate the expected value and the variance of $X$.

$\diamondsuit$

Again, we look at some examples of widely used probability density functions.

• **7–44 Definition:** A continuous random variable $X$ is **uniformly distributed** in $[a,b]$, if

$$f_X(x) = \begin{cases} \frac{1}{b-a}, & x \in [a,b], \\ 0, & \text{otherwise.} \end{cases}$$

We write $X \sim \mathcal{U}([a,b])$. ♢

Properties:

- expected value $\mathrm{E}(X) = \frac{a+b}{2}$
- variance $\mathrm{Var}(X) = \frac{(b-a)^2}{12}$

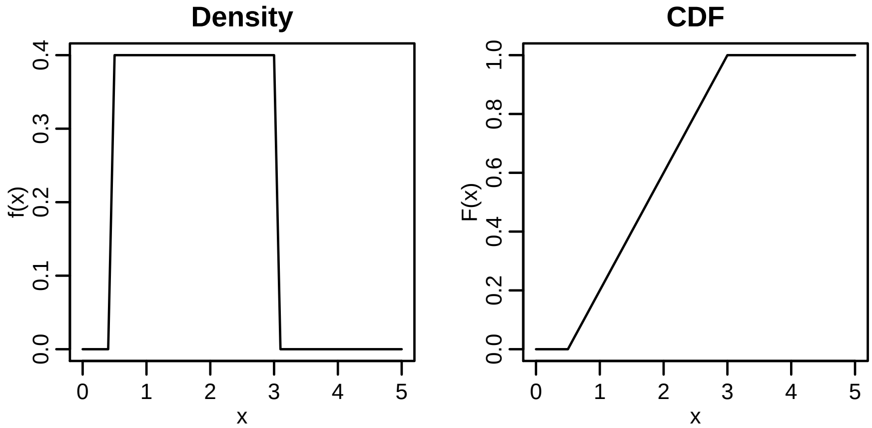



Figure 7.6: Uniform distribution on the interval $[0.5, 3]$

• **7–45 Definition:** A continuous random variable $X$ is **normally distributed** with mean (expected value) $\mu$ and variance $\sigma^2$, if

$$f_X(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{ -\frac{1}{2} \left( \frac{x-\mu}{sigma} \right)^2 \right\}, \; x \in \mathbb{R} \; .$$

We write $X \sim \mathcal{N}(\mu, \sigma^2)$, $\mu \in \mathbb{R}$, $\sigma^2 > 0$. ♢

Properties:

- **standard normal distribution**: $Z \sim \mathcal{N}(0,1)$
- transformation: $X \sim \mathcal{N}(\mu, \sigma^2)$, then $\frac{X-\mu}{\sigma} \sim \mathcal{N}(0,1)$
- as a special case: $\mathrm{P}(-2 \leq Z \leq 2) \approx 0.95$
- the CDF of a normal distribution has no closed form

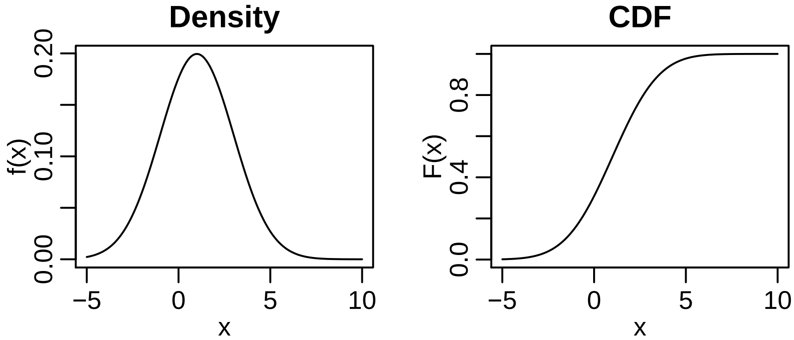



Figure 7.7: Normal distribution with mean $\mu = 1$ and variance $\sigma^2 = 4$

• **7–46 Question:**
Calculate the following probabilities for a standard normally distributed random variable $Z$.

(a) $\mathrm{P}(-1 \leq Z \leq 1)$

(b) $\mathrm{P}(-2 \leq Z \leq 2)$

(c) $\mathrm{P}(-3 \leq Z \leq 3)$

(d) $\mathrm{P}(Z \leq 1)$

(e) $\mathrm{P}(|Z| \geq 0.5)$

(f) $\mathrm{P}(-3 \leq Z \leq 1)$

$\diamondsuit$

• **7–47 Question:**
In the production of condensers, the capacity is a normally distributed random variable with expected value $\mu = 5$ [$\mu$F] and variance $\sigma^2 = 0.02^2$ [$(\mu\mathrm{F})^2$]. What is the expected rejection rate, if the capacity

(a) needs to be at least 4.98 $\mu$F?

(b) is not allowed to be higher than 5.05 $\mu$F?

(c) is allowed to differ from the nominal value $\mu = 5$ $\mu$F by maximally 0.03 $\mu$F?

$\diamondsuit$

• **7–48 Question:**
Thanks to a long-term study it is well-known that the lead content $X$ in a sample of soil is normally distributed. Furthermore, it is known that the expectation is 32 ppb (parts per billion) and the standard deviation is 6 ppb.

(a) Visualise the density of $X$ in a sketch which includes the probability that the sample of soil contains between 26 and 38 ppb of lead.

(b) What is the probability that a sample of soil contains at most 40 ppb of lead?

(c) Calculate the probability that a sample of soil contains at most 27 ppb of lead?

(d) Below which concentration falls the lead content with probability 97.5%? That is, determine $c$ such that $\mathrm{P}(X \leq c) = 97.5\%$.

(e) Below which concentration falls the lead content with probability 10%?

(f) What is the probability of the area you sketched in part (a) of this exercise?

$\diamondsuit$

• **7–49 Definition:** A random variable $X \geq 0$ is **exponentially distributed** with parameter („rate“) $\lambda$, if

$$f_X(x) = \begin{cases} \lambda \mathrm{e}^{-\lambda x}, & x \geq 0, \\ 0, & \text{otherwise.} \end{cases}$$

We write $X \sim \mathrm{Exp}(\lambda)$, $\lambda > 0$. $\diamondsuit$

Properties:

- an exponential random variable models the duration of random time intervals between two events
- expected value $\mathrm{E}(X) = \frac{1}{\lambda}$
- variance $\mathrm{Var}(X) = \frac{1}{\lambda^2}$

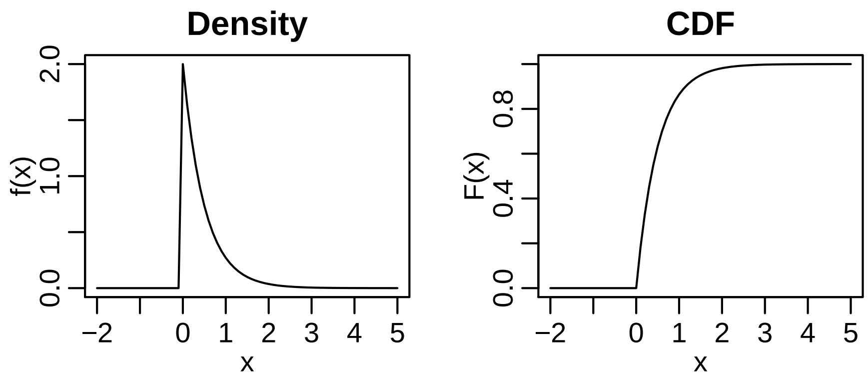



Figure 7.8: Exponential distribution for $\lambda = 2$

**• 7–50 Question:**
You get an email every 15 seconds on average. (For simplicity, assume that this rate is constant over time.)

(a) What is the probability that you wait less than 30 seconds for your next email?

(b) What is the maximum waiting time between two emails with 95% probability?

♢

**• 7–51 Question:**
Car accidents occur on average 4 times a week.

(a) Calculate the probability that more than 5 accidents occur in one week.

(b) Calculate the probability that at least two weeks elapse between two accidents.

♢

## 7.4 Joint Distributions

Up to now we considered the distribution of *single* random variables. In this section we now want to look at the (two-dimensional) distribution of two random variables which in particular do *not* need to be independent. Again, we first focus on discrete random variables and then look at continuous ones.

### 7.4.1 Discrete Random Variables

Let $X : \Omega_X \to W_X$ and $Y : \Omega_Y \to W_Y$ be *discrete* random variables.

• **7–52 Definition:** The **joint cumulative distribution function** of $X$ and $Y$ is given by

$$F_{X,Y}(x,y) := \mathrm{P}(X \leq x, Y \leq y) \ .$$

The **joint probability mass function** of $X$ and $Y$ is

$$p_{X,Y}(x,y) := \mathrm{P}(X = x, Y = y), x \in W_X, y \in W_Y \ .$$

$\diamondsuit$

• **7–53 Example:** A joint probability mass function can be indicated as a $2 \times 2$-table containing the probabilities $p_{X,Y}$ of the joint probability mass function. In this example we look at the pain perception of twins ($X$ = first twin and $Y$ = second twin): 1 = slight pain, 2 = moderate pain, 3 = severe pain.

| $X/Y$ | 1 | 2 | 3 |
|---|---|---|---|
| 1 | 0.31 | 0.08 | 0.02 |
| 2 | 0.11 | 0.18 | 0.06 |
| 3 | 0.04 | 0.09 | 0.11 |

*Note*: Of course, the sum of all the probabilities must equal 1 (just as for a single discrete random variable): $\sum_{x,y} p_{X,Y}(x,y) = 1$! $\diamondsuit$

From the joint probability mass function we can read off directly the so-called marginal probability mass functions – that is the probability mass functions of the single random variables $X$ and $Y$.

• **7–54 Definition:** Let $X$ and $Y$ be two discrete random variables, and $p_{X,Y}$ their joint probability mass function.

The **marginal probability mass function** of $X$ is given by

$$p_X(x) = \mathrm{P}(X = x) = \sum_{y \in W_Y} p_{X,Y}(x,y) \ .$$

Therefore, we get the marginal probability mass function of $X$ by simply summing up the probabilities $\mathrm{P}(X = x, Y = y)$ over all possible values $y$ of the random variable $Y$.

Analogously, the **marginal probability mass function** of $Y$ is given by

$$p_Y(y) = \mathrm{P}(Y = y) = \sum_{x \in W_X} p_{X,Y}(x,y) \ .$$

$\diamondsuit$

In a similar manner as for events, we define the *independence* of random variables.

• **7–55 Definition:** Two discrete random variables $X$ and $Y$ are **independent**, if $p_{X,Y}(x,y) = p_X(x) \cdot p_Y(y)$. $\diamondsuit$

• **7–56 Definition:** Let $X$ and $Y$ be two discrete random variables, and $p_{X,Y}$ their joint probability mass function.

The **conditional probability mass function** of $X$ given $Y = y$ is

$$p_{X|Y=y}(x) = \frac{p_{X,Y}(x,y)}{p_Y(y)}$$

(We can define the conditional probability mass function of $Y$ given $X = x$ in an analogue way: $p_{Y|X=x}(y) = p_{X,Y}(x,y)/p_X(x)$.) The conditional probability function $p_{X|Y}$ allows us to make statements about the distribution of $X$ given that we already know the value of $Y$. $\diamondsuit$

*Note*: The conditional probability mass function is equal to the marginal probability mass function, i.e. $p_{X|Y=y}(x) = p_X(x)$, if $X$ and $Y$ are independent. This makes perfectly sense, as in the case of independence, the value of $Y$ does not give us any information on $X$.

Let us look again at Example 7–53. We can directly read off the marginal probability mass functions by summing up the rows and the columns respectively:

| $x$ | 1 | 2 | 3 |
|---|---|---|---|
| $\mathrm{P}(X = x)$ | 0.41 | 0.35 | 0.24 |

| $y$ | 1 | 2 | 3 |
|---|---|---|---|
| $\mathrm{P}(Y = y)$ | 0.46 | 0.35 | 0.19 |

We can also answer the question, whether the twins have independent pain perceptions or not: As expected, their answers are **not** independent of each other, e.g. $\mathrm{P}(X = 1, Y = 1) = 0.31 \neq \mathrm{P}(X = 1) \cdot \mathrm{P}(Y = 1) = 0.1886$.

Further, given that the first twin has a pain perception of 2, the (conditional) probability mass function of the second twin is given by

$$p_{Y|X=2}(y) = p_{X,Y}(2, y)/p_X(2)$$

which corresponds to

| $y$ | 1 | 2 | 3 |
|---|---|---|---|
| $\mathrm{P}(Y = y \vert X = 2)$ | $0.11/0.35 \approx 0.31$ | $0.18/0.35 \approx 0.51$ | $0.06/0.35 \approx 0.17$ |

(Of course, the probabilities must sum up to 1.)
E.g.: Given the information that the first twin has a pain perception of 2, the second twin has also a pain perception of 2 with probability 0.51.

**• 7–57 Question:**
The pain perception of twins of different gender has been discretised to 3 levels: 1 = slight, 2 = moderate, 3 = severe. The following table shows the joint probability mass function of the pain levels of the twins ($X$ and $Y$, where $X$ describes the pain perception of the female twin):

| $X/Y$ | 1 | 2 | 3 |
|---|---|---|---|
| 1 | 0.05 | 0.08 | 0.12 |
| 2 | 0.14 | 0.19 | 0.09 |
| 3 | 0.22 | 0.08 | 0.03 |

(a) What is the marginal distribution of $X$? And the marginal distribution of $Y$?

(b) Are the random variables $X$ and $Y$ independent?

(c) What is the probability of twin $X$ being on a slight level? Does this probability change, if you know that $Y$ is being on a moderate level?

(d) In the table above, you see that $p_{X,Y}(1, 2) = p_{X,Y}(3, 2) = 0.08$. Does this imply that $p_{Y|X=1}(2) = p_{Y|X=3}(2)$? Comment.

♢

**• 7–58 Question:**
You have in your wallet 1 CHF, 2 CHF and 5 CHF – one coin of each. You draw randomly two coins (out of the three) without replacement. Let $X$ be the value of the first and $Y$ be the value of the second drawn coin.

(a) Calculate the joint probability mass function of $X$ and $Y$.

(b) Calculate the expected value of $Z = X + Y$.

♢

### 7.4.2 Continuous Random Variables

Let now $X : \Omega_X \to \mathbb{R}$ and $Y : \Omega_Y \to \mathbb{R}$ be two *continuous* random variables.
• **7–59 Definition:** The **joint cumulative distribution function** of $X$ and $Y$ is given by

$$F_{X,Y}(x, y) := \mathrm{P}(X \leq x, Y \leq y) \ .$$

Their **joint probability density** is

$$f_{X,Y}(x, y) := \frac{\partial}{\partial x}\frac{\partial}{\partial y} F_{X,Y}(x, y).$$

$\diamondsuit$

With this definition, we can now calculate probabilities like

$$\mathrm{P}(a \leq X \leq b, c \leq Y \leq d) = \int_a^b \int_c^d f_{X,Y}(x, y)\, \mathrm{d}y\, \mathrm{d}x$$

for any real numbers $a < b$ and $c < d$. As an easy graphical representation, we can plot a joint probability density function as a (two-dimensional) contour plot (see Figure 7.10). (Alternatively, we can of course also use the straightforward three-dimensional representation; see Figure 7.11.)

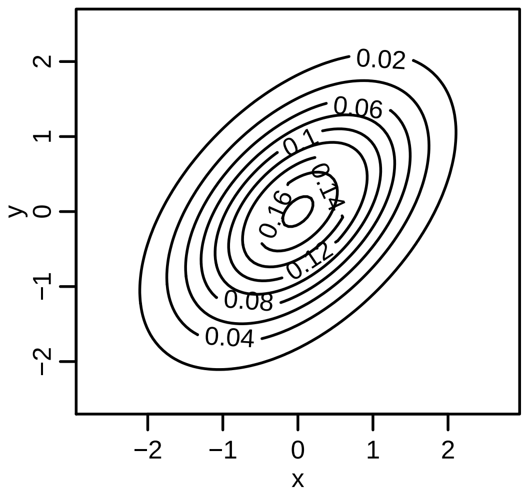



Figure 7.9: Contour plot of a joint probability density function

• **7–60 Definition:** Let $X$ and $Y$ be continuous random variables with joint density $p_{X,Y}$. In analogy to the discrete case, we define: The **marginal probability densities** of $X$ and $Y$ are given by

$$\begin{aligned} f_X(x) &:= \int_{-\infty}^{\infty} f_{X,Y}(x, y)\, \mathrm{d}y \ , \\ f_Y(y) &:= \int_{-\infty}^{\infty} f_{X,Y}(x, y)\, \mathrm{d}x \ . \end{aligned}$$

$\diamondsuit$

Also in the context of continuous random variables it is meaningful to talk about independence of random variables and conditional probability *densities*.
• **7–61 Definition:** Two continuous random variables $X$ and $Y$ are independent, if $f_{X,Y}(x, y) = f_X(x) \cdot f_Y(y)$. $\diamondsuit$

• **7–62 Definition:** Let $X$ and $Y$ be continuous random variables with joint density $p_{X,Y}$. The **conditional probability density** of $X$ given $Y = y$ is

$$f_{X|Y=y}(x) = \frac{f_{X,Y}(x, y)}{f_Y(y)}.$$

◇

The marginal, as well as the conditional density can be represented by simple one-dimensional density curves:

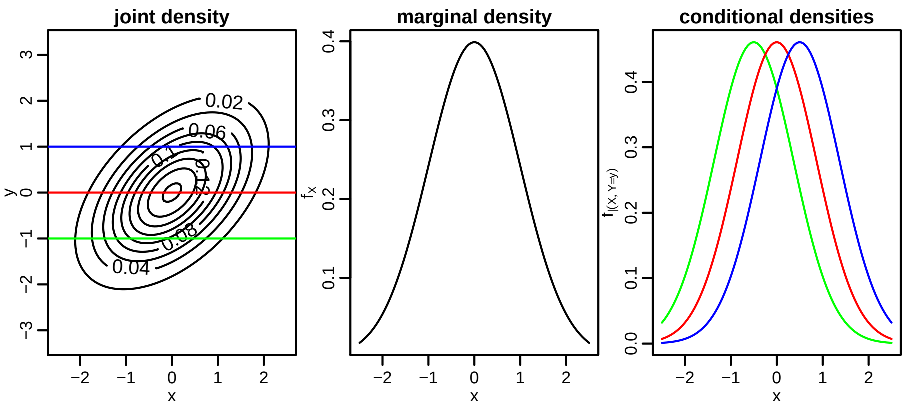



Figure 7.10: Contour plot with corresponding marginal and conditional densities

• **7–63 Example:** The following plot shows a bivariate normal distribution. Note that the marginal densities ($X \sim \mathcal{N}(1, 0.5^2)$ and $Y \sim \mathcal{N}(1.5, 1^2)$), as well as *any* conditional density is univariate normally distributed!

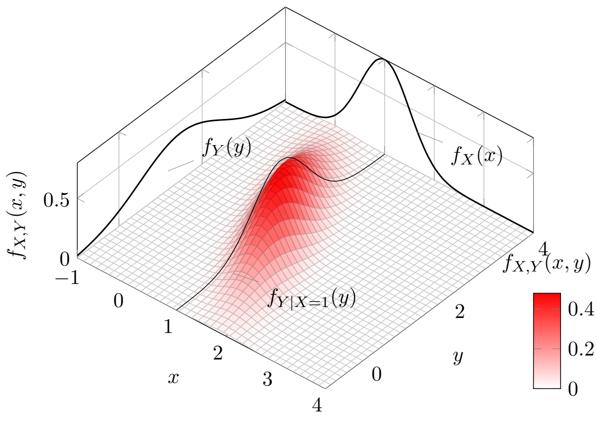
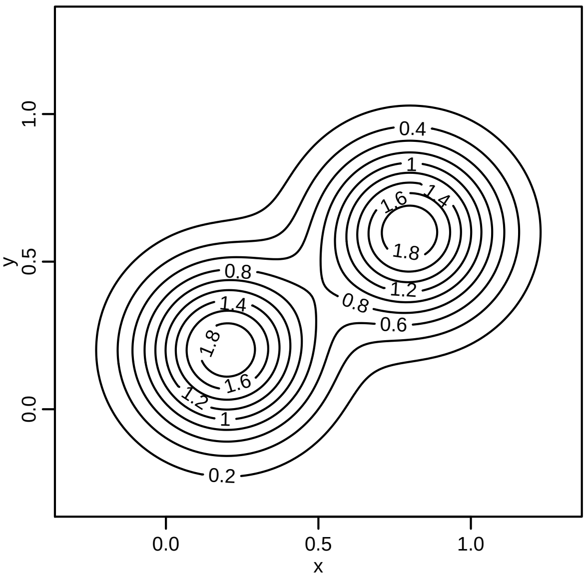



Figure 7.11: Marginal and conditional density of a bivariate normal distribution

◇

• **7–64 Question:**

The following plot shows the contour lines of a two-dimensional joint probability density $f_{X,Y}$:




(a) Sketch the marginal density $f_X(x)$.

(b) Sketch the conditional density $f_{X|Y=y}(x)$ for $y = 0.4$ and for $y = 0.6$.

$\diamondsuit$

• **7–65 Question:**
We look at the function

$$f_{X,Y}(x,y) = \begin{cases} 2e^{-x}e^{-2y}, & \text{for } x, y \geq 0, \\ 0, & \text{else }. \end{cases}$$

(a) Show that $f_{X,Y}$ is a bivariate density function.

(b) Calculate the probability $P(X > 1, Y < 1)$.

(c) Are the random variables $X$ and $Y$ independent?

$\diamondsuit$

### 7.4.3 Covariance and Correlation

• **7–66 Definition:** Let $X$ and $Y$ be two (discrete or continuous) random variables. The **covariance** between $X$ and $Y$ is defined as

$$\begin{aligned} \mathrm{Cov}(X,Y) &:= \mathrm{E}\left((X - \mathrm{E}(X))\,(Y - \mathrm{E}(Y))\right) \\ &= \mathrm{E}(X \cdot Y) - \mathrm{E}(X) \cdot \mathrm{E}(Y). \end{aligned}$$

Their **correlation** is

$$\rho_{XY} := \frac{\mathrm{Cov}(X,Y)}{\sqrt{\mathrm{Var}(X)\,\mathrm{Var}(Y)}} \;.$$

$\diamondsuit$

The covariance of two random variables describes the *linear relationship* between them: The covariance is positive, if for large values of $X$, also $Y$ tends to have large values (with respect to the expected value); on the other hand, the covariance is negative, if $X$ and $Y$ show an inverse linear relationship. The value of the covariance depends on the unit of measurement of the random variables. This fact makes it hard to interpret the value of the covariance. As a consequence, only the sign of it can be meaningfully interpreted. Due to this reason, in practice usually the correlation coefficient is used. The correlation has the advantage of being independent of the unit of measurement and takes values in the standardised interval [-1,1]. Roughly speaking, the correlation is nothing else than the standardised covariance.
Now we can interpret the correlation as **strength of linear relationship** between two random variables:

- If $X$ and $Y$ are independent, $\mathrm{Cov}(X,Y) = 0$ and $\rho_{XY} = 0$, i.e. there exists no linear relationship between $X$ and $Y$. *The other direction is* not *true!*
- Perfect (positive) linear relationship: $\rho_{XY} = 1$ if $Y = a + b \cdot X$ for some $b > 0$.
- Perfect (negative) linear relationship: $\rho_{XY} = -1$ if $Y = a + b \cdot X$ for some $b < 0$.

We can nicely see the strength of the linear relationship between two random variables looking at the contour plot of their bivariate distribution as shown in Figure 7.12.

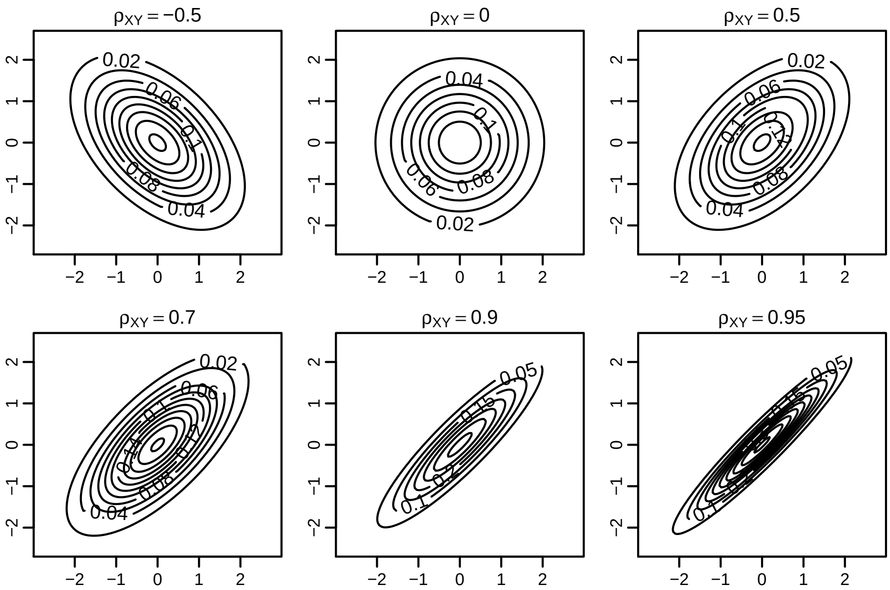



Figure 7.12: Examples of contour plots of correlated random variables

In Example 7–63 the random variables $X$ and $Y$ are not independent, in fact they are negatively correlated ($\rho_{XY} = -0.75$). This can be seen from the fact that the surface is oval-shaped and does not look like a nicely shaped bell.

Let us finish off with a list of a few very helpful calculation rules for $\mathrm{E}(\cdot)$ and $\mathrm{Var}(\cdot)$. Let $X$ and $Y$ be (discrete or continuous) random variables, and $a, b \in \mathbb{R}$ real numbers.

- $\mathrm{E}(X + Y) = \mathrm{E}(X) + \mathrm{E}(Y)$
- if $X$ and $Y$ are independent, $\mathrm{E}(X \cdot Y) = \mathrm{E}(X) \cdot \mathrm{E}(Y)$
- $\mathrm{E}(a \cdot X + b) = a \cdot \mathrm{E}(X) + b$
- $\mathrm{Var}(a \cdot X + b) = a^2 \cdot \mathrm{Var}(X)$

- $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2 \cdot \text{Cov}(X, Y)$
- if $X$ and $Y$ are independent, $\text{Cov}(X, Y) = 0$

• **7–67 Question:**
We claimed that the variance of the sum of two random variables $X$ and $Y$ is

$$\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2 \cdot \text{Cov}(X, Y).$$

Proof this statement.

♢

• **7–68 Question:**

(a) You throw a dice twice and you add the number of spots. Calculate the expected value and the variance of the result.

(b) You throw a dice once and you multiply the number of spots by two. Calculate the expected value and the variance of the result.

(c) You throw a dice once and you take the sum of the number of spots on the top and the bottom side of the dice. Calculate the expected value and the variance of the result.

♢

• **7–69 Question:**
Max and Moritz found a bit of money and decide to split it based on flipping two coins. If the first coin shows head, Max gets 2 CHF; Moritz gets 1 CHF, if the first coin shows tail and 2 CHF, if the first coin shows head *and* the second coin shows tail. (Therefore, the value of the second coin is only relevant, if the first coin shows head.) They donate the rest of the money to a charitable organisation.

(a) Let $X$ be the amount of money Max wins and $Y$ be the amount of money Moritz wins. Calculate the joint distribution of $X$ and $Y$.

(b) Calculate $\text{E}(X)$ and $\text{E}(Y)$.

(c) Are $X$ and $Y$ independent of each other?

(d) Are $X$ and $Y$ uncorrelated?

♢

# Chapter 8

# Solutions to the Questions

**Solution to Question 2–1 :**
(a)

$$\vec{a} = \mathbf{A}\,\vec{x} = \begin{bmatrix} 1 & 2 & 3 \\ 0 & -3 & 4 \end{bmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 1\cdot 1 + 2\cdot 2 + 3\cdot 3 \\ 0\cdot 1 - 3\cdot 2 + 4\cdot 3 \end{pmatrix} = \begin{pmatrix} 14 \\ 6 \end{pmatrix}$$

(b) $\vec{b} = \vec{x}^T \mathbf{A}^T = \vec{a}^T = (14\,,\,6)$

(c)

$$\mathbf{C} = \mathbf{A}\cdot\mathbf{B} = \begin{bmatrix} 10 & 3 \\ -1 & 3 \end{bmatrix} \cdot \begin{bmatrix} 1 & 2 & 3 \\ 0 & -3 & 4 \end{bmatrix} = \begin{bmatrix} 10 & 11 & 42 \\ -1 & -11 & 9 \end{bmatrix}$$

(d)

$$s = \langle \vec{x}\,,\,\vec{y}\rangle = \langle \begin{pmatrix} 1 \\ 2 \\ 7 \end{pmatrix} , \begin{pmatrix} -2 \\ 0 \\ 3 \end{pmatrix} \rangle = 1\cdot(-2) + 2\cdot 0 + 3\cdot 3 = 6$$

(e)

$$\begin{aligned}
p_1 &= \langle \mathbf{B}\begin{pmatrix} 2 \\ 3 \end{pmatrix} , \begin{pmatrix} -1 \\ 5 \end{pmatrix}\rangle = \langle \begin{bmatrix} 10 & 3 \\ -1 & 3 \end{bmatrix}\begin{pmatrix} 2 \\ 3 \end{pmatrix} , \begin{pmatrix} -1 \\ 5 \end{pmatrix}\rangle \\
&= \langle \begin{pmatrix} 29 \\ 7 \end{pmatrix} , \begin{pmatrix} -1 \\ 5 \end{pmatrix}\rangle = +6 \\
p_2 &= \langle \begin{pmatrix} 2 \\ 3 \end{pmatrix} , \mathbf{B}^T\begin{pmatrix} -1 \\ 5 \end{pmatrix}\rangle = \langle \begin{pmatrix} 2 \\ 3 \end{pmatrix} , \begin{bmatrix} 10 & -1 \\ 3 & 3 \end{bmatrix}\begin{pmatrix} -1 \\ 5 \end{pmatrix}\rangle \\
&= \langle \begin{pmatrix} 2 \\ 3 \end{pmatrix} , \begin{pmatrix} -15 \\ 12 \end{pmatrix}\rangle = -30 + 24 = +6
\end{aligned}$$

It should be no surprise that $p_1 = p_2$ since for any matrix $\mathbf{B}$ and vectors $\vec{u}$ and $\vec{v}$ we find

$$\langle \mathbf{B}\,\vec{u}\,,\,\vec{v}\rangle = (\mathbf{B}\vec{u})^T \cdot \vec{v} = \vec{u}^T\,\mathbf{B}^T \cdot \vec{v} = \langle \vec{u}\,,\,\mathbf{B}^T\,\vec{v}\rangle$$

All of the above can be determined (or verified) with *Octave* or `MATLAB`.

```
A = [1  2  3;  0 −3  4];
B = [10  3;  −1  3];
x = [1;2;3];    y = [−2;0;3];

a = A*x
b = x'*A'
C = B*A
s = x'*y
p1 = (B*[2;3])'*[−1;5]
p2 = [2;3]'*(B'*[−1;5])
-->
a =    14
        6
b =    14     6
C =    10    11    42
       −1   −11     9
s =   7
p1 =   6
p2 =   6
```

**Solution to Question 2–2 :** The augmented matrix notation of the system is

$$\left[\begin{array}{ccc|c} 0 & -2 & 7 & 12 \\ 2 & -10 & 12 & 28 \\ 2 & -5 & -5 & -1 \end{array}\right]$$

To solve the system this matrix has to be transformed into an equivalent system by applying row operations. There are three basis operations.

- Multiply a row by a nonzero number. This represents to multiplying the corresponding equation by that number.
- Add a multiple of one row to another row. This represents adding the corresponding equations.
- Swap two rows. This represents swapping the two equations.

The above operations do not change the set of solutions of the original system.

The first row has to be swapped with the second or third, to obtain a nonzero entry in the top left corner.

$$\left[\begin{array}{ccc|c} 2 & -10 & 12 & 28 \\ 0 & -2 & 7 & 12 \\ 2 & -5 & -5 & -1 \end{array}\right]$$

Multiply the first row by $\frac{1}{2}$ to obtain 1.

$$\left[\begin{array}{ccc|c} 1 & -5 & 6 & 14 \\ 0 & -2 & 7 & 12 \\ 2 & -5 & -5 & -1 \end{array}\right]$$

Subtract a multiple of the first row from the third row to obtain all zeros below the leading 1 .

$$\left[\begin{array}{ccc|c} 1 & -5 & 6 & 14 \\ 0 & -2 & 7 & 12 \\ 0 & +5 & -17 & -29 \end{array}\right]$$

From now on the first row can be ignored and we proceed with the smaller matrix. and apply the following steps sequentially.

- Divide the second row by $-2$ .
- Add 5 times the second row to the third row.
- Multiply the third row by 2 .

$$\left[\begin{array}{ccc|c} 1 & -5 & 6 & 14 \\ 0 & 1 & \frac{-7}{2} & -6 \\ 0 & +5 & -17 & -29 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & -5 & 6 & 14 \\ 0 & 1 & \frac{-7}{2} & -6 \\ 0 & +5 & -17 & -29 \end{array}\right]$$

$$\longrightarrow \left[\begin{array}{ccc|c} 1 & -5 & 6 & 14 \\ 0 & 1 & \frac{-7}{2} & -6 \\ 0 & 0 & \frac{1}{2} & 1 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & -5 & 6 & 14 \\ 0 & 1 & \frac{-7}{2} & -6 \\ 0 & 0 & 1 & 2 \end{array}\right]$$

The last matrix corresponds to an equivalent system of linear equations.

$$\begin{array}{cccrr} x & -5\,y & +6\,z & = & 14 \\ & y & -\frac{7}{2}\,z & = & -6 \\ & & z & = & 2 \end{array}$$

This system is now easily solved by starting with the last equation.

$$\begin{array}{ccccccc} z & = & 2 & & & = & 2 \\ y & = & -6 & +\frac{7}{2}\,z & & = & 1 \\ x & = & 14 & -6\,z & +5\,y & = & 7 \end{array}$$

**Solution to Question 2–3 :** Representing the system by an augmented matrix leads to first matrix below. Then the row reduction is performed by

- Subtract 3 times the first row from the third row.
- Add 0.5 times the second row to the third row.

$$\left[\begin{array}{ccc|c} 1 & 1 & a & 0 \\ 0 & 2 & 4 & 0 \\ 3 & 2 & 10 & 0 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 1 & a & 0 \\ 0 & 2 & 4 & 0 \\ 0 & -1 & 10-3\,a & 0 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 1 & a & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 12-3\,a & 0 \end{array}\right]$$

The equation corresponding to the new third row now is given by

$$0 \cdot x_1 + 0 \cdot x_2 + (12 - 3\,a) \cdot x_3 = 0$$

For this system to have infinitely many solutions we need $12 - 3\,a = 0$, i.e. $a = 4$. In this case the new reduced system reads as

$$\begin{array}{cccrr} x_1 & +x_2 & +4\,x_3 & = & 0 \\ & 2\,x_2 & +4\,x_3 & = & 0 \\ 0\,x_1 & +0\,x_2 & +0\,x_3 & = & 0 \end{array}$$

Thus we are free to choose $x_3 = t \in \mathbb{R}$. Then we find $x_2 = -2\,x_3 = -2\,t$ and $x_1 = -x_2 - 4\,x_3 = -2\,t$. Thus all solutions are given by

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = t \begin{pmatrix} -2 \\ -2 \\ 1 \end{pmatrix} \quad \text{where} \quad t \in \mathbb{R}$$

**Solution to Question 2–4 :** Write the system with an augmented matrix and then apply row operations.

- Divide the first row by 13.
- Divide the second row by $-2$.
- Subtract the first row from the third row

$$\left[\begin{array}{ccc|c} 13 & 0 & -39 & 13 \\ 0 & -2 & -4 & 6 \\ 1 & 1 & -1 & -2 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 0 & -3 & 1 \\ 0 & 1 & 2 & -3 \\ 1 & 1 & -1 & -2 \end{array}\right] \longrightarrow \left[\begin{array}{ccc|c} 1 & 0 & -3 & 1 \\ 0 & 1 & 2 & -3 \\ 0 & 0 & 0 & 0 \end{array}\right]$$

Thus the system

$$\begin{array}{rrrcr} x & & -3\,z & = & 1 \\ & y & +2\,z & = & -3 \\ 0\,x & +0\,y & +0\,z & = & 0 \end{array}$$

is equivalent to the original system. Obviously we only have two equations with information. We are free to choose the value of $z = t \in \mathbb{R}$. Then the system reads as

$$\begin{array}{rrcrr} x & & = & 1 & +3\,t \\ & y & = & -3 & -2\,t \end{array}$$

and the general solution is given by

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} 1 \\ -3 \\ 0 \end{pmatrix} + t \begin{pmatrix} 3 \\ -2 \\ 1 \end{pmatrix} \quad \text{where} \quad t \in \mathbb{R}$$

**Solution to Question 2–5 :**

(a) As an example we multiply the second row by $\alpha$.

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & \alpha & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{bmatrix} = \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ \alpha\,a_{2,1} & \alpha\,a_{2,2} & \alpha\,a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{bmatrix}$$

(b) As an example we subtract twice the first row from the third row.

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -2 & 0 & 1 \end{bmatrix} \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{bmatrix} = \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} - 2\,a_{1,1} & a_{3,2} - 2\,a_{1,2} & a_{3,3} - 2\,a_{1,3} \end{bmatrix}$$

(c) As an example we swap rows 2 and 3 .

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{bmatrix} = \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \end{bmatrix}$$

**Solution to Question 2–7 :**

(a) The mapping is described by

$$\mathbf{A} \begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} 3 \\ 9 \end{pmatrix} \quad \text{and} \quad \mathbf{A} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -1/2 \\ -1/2 \end{pmatrix} .$$

Thus the matrix $\mathbf{A}$ is determined by

$$\mathbf{A} \cdot \begin{bmatrix} 1 & 1 \\ 3 & 1 \end{bmatrix} = \begin{bmatrix} 3 & -0.5 \\ 9 & -0.5 \end{bmatrix} \quad \text{or} \quad \mathbf{A} = \begin{bmatrix} 3 & -0.5 \\ 9 & -0.5 \end{bmatrix} \cdot \begin{bmatrix} 1 & 1 \\ 3 & 1 \end{bmatrix}^{-1} .$$

Since

$$\begin{bmatrix} 1 & 1 \\ 3 & 1 \end{bmatrix}^{-1} = \frac{1}{-2} \begin{bmatrix} 1 & -1 \\ -3 & 1 \end{bmatrix} = \begin{bmatrix} -0.5 & 0.5 \\ 1.5 & -0.5 \end{bmatrix}$$

this leads to

$$\mathbf{A} = \begin{bmatrix} -2.25 & 1.75 \\ -5.25 & 4.75 \end{bmatrix} .$$

(b) The image of $(4,1)^T$ is given by

$$\mathbf{A} \begin{pmatrix} 4 \\ 1 \end{pmatrix} = \begin{bmatrix} -2.25 & 1.75 \\ -5.25 & 4.75 \end{bmatrix} \begin{pmatrix} 4 \\ 1 \end{pmatrix} = \begin{pmatrix} -7.25 \\ -16.25 \end{pmatrix}$$

**Solution to Question 2–8 :**

(a) Solve the characteristic equation, which in this case is a quadratic equation.

$$\begin{aligned} 0 &= \det(\mathbf{A} - \lambda\, \mathbb{I}) = \det \begin{bmatrix} 1-\lambda & 2 \\ 5 & 4-\lambda \end{bmatrix} = \lambda^2 - 5\,\lambda - 6 \\ \lambda_{1,2} &= \frac{1}{2}\left(5 \pm \sqrt{25+24}\right) = \begin{cases} -1 \\ +6 \end{cases} \end{aligned}$$

(b) For the first eigenvector we have to examine the system

$$\begin{aligned} \mathbf{A}\,\vec{v} &= \lambda_1\,\vec{v}_1 = -1\,\vec{v}_1 \\ \begin{bmatrix} 1 & 2 \\ 5 & 4 \end{bmatrix} \begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} -x \\ -y \end{pmatrix} \\ \begin{bmatrix} 2 & 2 \\ 5 & 5 \end{bmatrix} \begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} 0 \\ 0 \end{pmatrix} . \end{aligned}$$

Using the notation of an augmented matrix this leads to

$$\left[\begin{array}{cc|c} 2 & 2 & 0 \\ 5 & 5 & 0 \end{array}\right] \longrightarrow \left[\begin{array}{cc|c} 1 & 1 & 0 \\ 0 & 0 & 0 \end{array}\right] .$$

which means $x + y = 0$ and no second equation. An easy solution is

$$\vec{v}_1 = \begin{pmatrix} +1 \\ -1 \end{pmatrix} .$$

but any multiple of this vector is an eigenvector too.

(c) For the second eigenvector we have to examine the system

$$\begin{aligned} \mathbf{A}\,\vec{v} &= \lambda_2\,\vec{v}_2 \;=\; +6\,\vec{v}_2 \\ \begin{bmatrix} 1 & 2 \\ 5 & 4 \end{bmatrix} \begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} 6\,x \\ 6\,y \end{pmatrix} \\ \begin{bmatrix} -5 & 2 \\ 5 & -2 \end{bmatrix} \begin{pmatrix} x \\ y \end{pmatrix} &= \begin{pmatrix} 0 \\ 0 \end{pmatrix} . \end{aligned}$$

Now the solutions are all multiples of

$$\vec{v}_2 = \begin{pmatrix} 2 \\ 5 \end{pmatrix} .$$

(d) The scalar product of the two eigenvectors is different from zero.

$$\langle \vec{v}_1\,,\,\vec{v}_2 \rangle = \langle \begin{pmatrix} +1 \\ -1 \end{pmatrix} , \begin{pmatrix} 2 \\ 5 \end{pmatrix} \rangle = +1 \cdot 2 - 1 \cdot 5 = -3 \neq 0$$

(e) `MATLAB`/*Octave* will return normalized eigenevectors, i.e. with length 1 .

```
[EigenVectors,EigenValues] = eig([1 2;5 4])
-->
EigenVectors =   -0.70711   -0.37139
                 +0.70711   -0.92848

EigenValues =   -1    0
                 0    6
```

**Solution to Question 2–9 :**

(a) Solve the characteristic equation, which in this case is a cubic equation, which can be solved easily.

$$\begin{aligned} 0 &= \det(\mathbf{A} - \lambda\,\mathbb{I}) = \det \begin{bmatrix} 2-\lambda & -1 & 0 \\ -1 & 2-\lambda & 0 \\ 0 & 0 & 7-\lambda \end{bmatrix} = (\lambda^2 - 4\,\lambda + 3)\,(7-\lambda) \\ &= (\lambda-3)\,(\lambda-1)\,(7-\lambda) \\ \lambda_1 &= 3 \quad , \quad \lambda_2 = 1 \quad , \quad \lambda_3 = 7 \end{aligned}$$

(b) For the first eigenvector we have to examine the system

$$\begin{aligned}
\mathbf{A}\,\vec{v} &= \lambda_1\,\vec{v}_1 \;=\; 3\,\vec{v}_1 \\
\begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & 0 \\ 0 & 0 & 7 \end{bmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} &= 3 \begin{pmatrix} x \\ y \\ z \end{pmatrix} \\
\begin{bmatrix} -1 & -1 & 0 \\ -1 & -1 & 0 \\ 0 & 0 & 4 \end{bmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} &= \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} .
\end{aligned}$$

In this system observe that $z = 0$ and $y = -x$ leads to a solution. Thus the first eigenvectors are all multiples of $\vec{v}_1 = (+1\,,\,-1\,,\,0)^T$.

(c) For the second eigenvector examine the system

$$\begin{aligned}
\mathbf{A}\,\vec{v} &= \lambda_2\,\vec{v}_2 \;=\; 1\,\vec{v}_2 \\
\begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & 0 \\ 0 & 0 & 7 \end{bmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} &= 1 \begin{pmatrix} x \\ y \\ z \end{pmatrix} \\
\begin{bmatrix} +1 & -1 & 0 \\ -1 & +1 & 0 \\ 0 & 0 & 6 \end{bmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} &= \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} .
\end{aligned}$$

In this system observe that $z = 0$ and $y = x$ leads to a solution. Thus the second eigenvectors are all multiples of $\vec{v}_2 = (+1\,,\,+1\,,\,0)^T$.

(d) For the third eigenvector we have to examine the system

$$\begin{aligned}
\mathbf{A}\,\vec{v} &= \lambda_3\,\vec{v}_3 \;=\; 7\,\vec{v}_3 \\
\begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & 0 \\ 0 & 0 & 7 \end{bmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} &= 7 \begin{pmatrix} x \\ y \\ z \end{pmatrix} \\
\begin{bmatrix} -5 & -1 & 0 \\ -1 & -5 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} &= \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} .
\end{aligned}$$

Since $\det \begin{bmatrix} -5 & -1 \\ -1 & -5 \end{bmatrix} = 24 \neq 0$ we know that $x = y = 0$ and the third equations reads as $0\,x + 0\,y + 0\,z = 0$, i.e. we are free to choose the value of $z$ . Thus the third eigenvectors are all multiples of $\vec{v}_3 = (0\,,\,0\,,\,1)^T$.

(e) An easy computations shows that the pairwise scalar products all vanish.

$$\langle \vec{v}_1\,,\,\vec{v}_2 \rangle = \langle \vec{v}_1\,,\,\vec{v}_3 \rangle = \langle \vec{v}_2\,,\,\vec{v}_3 \rangle = 0$$

(f) Normalizing the above eigenvectors leads to

$$\vec{v}_1 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \longrightarrow \begin{pmatrix} \frac{-1}{\sqrt{2}} \\ \frac{+1}{\sqrt{2}} \\ 0 \end{pmatrix} \quad , \quad \vec{v}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \longrightarrow \begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{+1}{\sqrt{2}} \\ 0 \end{pmatrix} \quad \text{and} \quad \vec{v}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} .$$

Thus the matrix is given by

$$\mathbf{Q} = \begin{bmatrix} \frac{-1}{\sqrt{2}} & \frac{+1}{\sqrt{2}} & 0 \\ \frac{+1}{\sqrt{2}} & \frac{+1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

and then

$$\mathbf{Q}^T\mathbf{Q} = \begin{bmatrix} \frac{-1}{\sqrt{2}} & \frac{+1}{\sqrt{2}} & 0 \\ \frac{+1}{\sqrt{2}} & \frac{+1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \frac{-1}{\sqrt{2}} & \frac{+1}{\sqrt{2}} & 0 \\ \frac{+1}{\sqrt{2}} & \frac{+1}{\sqrt{2}} & 0 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

Thus the inverse matrix of $\mathbf{Q}$ is given by its transpose $\mathbf{Q}^T$. If the original matrix $\mathbf{A}$ is symmetric this is always the case. But observe that $\mathbf{Q}$ is usually **not symmetric**. For the second statement observe

$$\begin{aligned}
\mathbf{AQ} &= \mathbf{A} \cdot \begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \vec{v}_3 \end{bmatrix} \begin{bmatrix} \lambda_1\vec{v}_1 & \lambda_2\vec{v}_2 & \lambda_3\vec{v}_3 \end{bmatrix} \\
&= \begin{bmatrix} \vec{v}_1 & \vec{v}_2 & \vec{v}_3 \end{bmatrix} \cdot \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix} = \mathbf{Q} \cdot \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}.
\end{aligned}$$

Multiply the above equation from the left by $\mathbf{Q}^T = \mathbf{Q}^{-1}$ to obtain

$$\mathbf{Q}^T\mathbf{A} \cdot \mathbf{Q} = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}.$$

This is called a diagonalization of the symmetric matrix $\mathbf{A}$.

(g) `MATLAB`/*Octave* will return normalized eigenvectors, i.e. with length 1 .

```
[EigenVectors,EigenValues] = eig([2 -1 0;-1 2 0;0 0 7])
-->
EigenVectors =  -0.70711  -0.70711   0.00000
                -0.70711   0.70711   0.00000
                -0.00000   0.00000   1.00000

EigenValues =   1   0   0
                0   3   0
                0   0   7
```

**Solution to Question 2–10 :**

(a) For $\vec{x}(t) = e^{\lambda_i t}\,\vec{v}_i$ compute

$$\begin{aligned}
\frac{d}{dt}\,\vec{x}(t) &= \left(\frac{d}{dt}\,e^{\lambda_i t}\right)\,\vec{v}_i \;=\; \lambda_i\,e^{\lambda_i t}\,\vec{v}_i \\
\mathbf{A}\,\vec{x}(t) &= e^{\lambda_i t}\,\mathbf{A}\,\vec{v}_i \;=\; e^{\lambda_i t}\,\lambda_i\,\vec{v}_i\,.
\end{aligned}$$

and thus the ODE (2.1) is solved.

(b) For $\vec{x}(t) = \sum_{i=1}^n c_i\, e^{\lambda_i\, t}\, \vec{v}_i$ compute

$$\begin{aligned}
\frac{d}{dt}\,\vec{x}(t) &= \sum_{i=1}^n c_i\, \frac{d}{dt}\, e^{\lambda_i\, t}\, \vec{v}_i = \sum_{i=1}^n c_i\, \lambda_i\, e^{\lambda_i\, t}\, \vec{v}_i \\
\mathbf{A}\,\vec{x}(t) &= \sum_{i=1}^n c_i\, e^{\lambda_i\, t}\, \mathbf{A}\,\vec{v}_i = \sum_{i=1}^n c_i\, e^{\lambda_i\, t}\, \lambda_i\, \vec{v}_i\,.
\end{aligned}$$

and thus the ODE (2.1) is solved.

(c) With the notations

$$\vec{v}_i = \begin{pmatrix} v_{1,i} \\ v_{2,i} \\ v_{3,i} \\ \vdots \\ v_{n,i} \end{pmatrix} \quad \text{and} \quad \vec{x}_0 = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix}$$

the equation $\vec{x}_0 = \sum_{i=1}^n c_i\, \vec{v}_i$ leads to

$$\begin{aligned}
\vec{x}_0 &= \sum_{i=1}^n c_i\, \vec{v}_i \\
\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ \vdots \\ x_n \end{pmatrix} &= c_1 \begin{pmatrix} v_{1,1} \\ v_{2,1} \\ v_{3,1} \\ \vdots \\ v_{n,1} \end{pmatrix} + c_2 \begin{pmatrix} v_{1,2} \\ v_{2,2} \\ v_{3,2} \\ \vdots \\ v_{n,2} \end{pmatrix} + c_3 \begin{pmatrix} v_{1,3} \\ v_{2,3} \\ v_{3,3} \\ \vdots \\ v_{n,3} \end{pmatrix} + \cdots + c_n \begin{pmatrix} v_{1,n} \\ v_{2,n} \\ v_{3,n} \\ \vdots \\ v_{n,n} \end{pmatrix} \\
&= \begin{bmatrix} v_{1,1} & v_{1,2} & v_{1,3} & \cdots & v_{1,n} \\ v_{2,1} & v_{2,2} & v_{2,3} & \cdots & v_{2,n} \\ v_{3,1} & v_{3,2} & v_{3,3} & \cdots & v_{3,n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ v_{n,1} & v_{n,2} & v_{n,3} & \cdots & v_{n,n} \end{bmatrix} \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ \vdots \\ c_n \end{pmatrix}
\end{aligned}$$

and thus we have a system of linear equations for the coefficients $c_i$. One can show that this system has a unique solution if and only if the eigenvectors $v_i$ are linearly independent.

**Solution to Question 2–11 :** For $\vec{x}(t) = \cos(\omega_i\, t)\, \vec{v}_i$ compute

$$\begin{aligned}
\frac{d^2}{dt^2}\,\vec{x}(t) &= \left(\frac{d}{dt}\, \cos(\omega_i\, t)\right)\, \vec{v}_i = -\omega_i^2\, \cos(\omega_i\, t)\, \vec{v}_i \\
-\mathbf{A}\,\vec{x}(t) &= -\cos(\omega_i\, t)\, \mathbf{A}\, \vec{v}_i = -\cos(\omega_i\, t)\, \lambda_i\, \vec{v}_i = -\omega_i^2\, \cos(\omega_i\, t)\, \vec{v}_i
\end{aligned}$$

and thus the system of differential equations is solved. The computation for the second solution $\sin(\omega_i\, t)\, \vec{v}_i$ is similar.

**Solution to Question 3–2 :** For the function $f(x) = \cos(x)$ use $f'(x) = -\sin(x)$, $f''(x) = -\cos(x)$ and $f'''(x) = +\sin(x)$ To check your results use the value $\cos 32° \approx 0.848048$.

(a)

$$\cos(32^\circ) = \cos(\frac{\pi}{6}+\frac{\pi}{90}) \approx \cos(\frac{\pi}{6}) - \sin(\frac{\pi}{6})\,\frac{\pi}{90} = \frac{\sqrt{3}}{2} - \frac{\pi}{180} \qquad (\approx 0.848572)$$

(b)

$$\cos(\frac{\pi}{6}+\frac{\pi}{90}) \approx \cos(\frac{\pi}{6}) - \sin(\frac{\pi}{6})\,\frac{\pi}{90} - \frac{1}{2}\,\cos(\frac{\pi}{6})\,(\frac{\pi}{90})^2 = \frac{\sqrt{3}}{2} - \frac{\pi}{180} - \frac{\sqrt{3}\,\pi^2}{4\cdot 90^2} \qquad (\approx 0.848044)$$

(c) For the linear approximation find

$$R_1 = -\frac{1}{2}\,\cos\xi\cdot(\frac{\pi}{90})^2 \quad \text{and thus} \quad |R_1| \leq \frac{1}{2}\,1\,(\frac{\pi}{90})^2 \approx 0.0006$$

For the quadratic approximation find

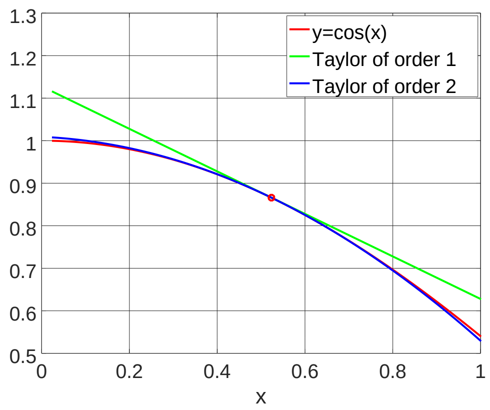
$$R_2 = -\frac{1}{6}\,\sin\xi\cdot(\frac{\pi}{90})^3 \quad \text{thus} \quad |R_2| \leq \frac{1}{6}\,1\,(\frac{\pi}{90})^3 \approx 7\cdot 10^{-6}$$

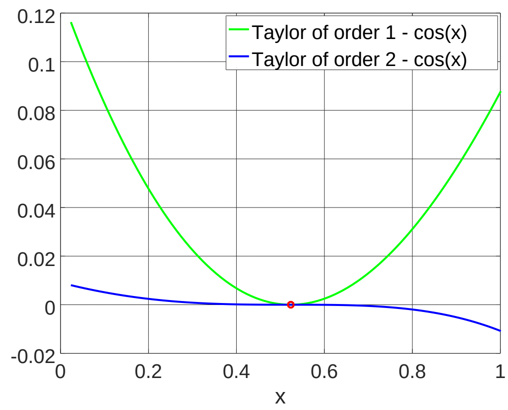
In both cases the actual error is smaller than the upper bound.

(d) Use the two functions

$$\begin{aligned}
\cos(\frac{\pi}{6}+\Delta x) &\approx f_1(\frac{\pi}{6}+\Delta x) &= \frac{\sqrt{3}}{2} - \frac{1}{2}\,\Delta x \\
\cos(\frac{\pi}{6}+\Delta x) &\approx f_2(\frac{\pi}{6}+\Delta x) &= \frac{\sqrt{3}}{2} - \frac{1}{2}\,\Delta x - \frac{1}{2}\frac{\sqrt{3}}{2}\,(\Delta x)^2
\end{aligned}$$

The first is an approximation by a straight line, the second an approximation by a parabola. Below find the functions and the resulting errors.




**Octave**

```
x0 = pi/6; Delta_x = linspace(-1,1,101)*0.5;
x  = x0+Delta_x;
f  = cos(x0+Delta_x);
f1 = cos(x0) - sin(x0)*Delta_x;
f2 = f1 - 0.5*cos(x0)*Delta_x.^2;
figure(1)
plot(x,f,'r',x,f1,'g',x,f2,'b',x0,cos(x0),'ro')
xlabel('x'); xlim([0,1])
legend('y=cos(x)','Taylor of order 1','Taylor of order 2')
print -dpdfwrite CosTaylor1.pdf
```

```
figure(2)
plot(x,f1-f,'g',x,f2-f,'b',x0,0,'ro')
xlabel('x'); xlim([0,1])
legend('Taylor of order 1 - cos(x)','Taylor of order 2 - cos(x)')
print -dpdfwrite CosTaylor2.pdf
```

**Solution to Question 3–3 :**

(a) For small variations $\Delta T$ and $\Delta U_0$ use

$$\Delta U_0 \approx \frac{\partial U}{\partial T}\,\Delta T\,.$$

For $\Delta T$ to be minimal for a given value of $\Delta U_0$ we need thus a large derivative $\frac{\partial U}{\partial T}$.

(b) First determine the slope of the curve

$$\begin{aligned} U\,(t) &= U_m\,(1-e^{-\alpha\,t}) \\ U'\,(t) &= U_m\,\alpha\,e^{-\alpha\,t} \\ U'\,(T) &= U_m\,\alpha\,e^{-\alpha\,T}\,. \end{aligned}$$

Now this new expression has to be as large as possible, as function of the parameter $\alpha$. Thus we need the zeros of the derivative with respect to $\alpha$.

$$\frac{\partial}{\partial\alpha}\,U'(T) = U_m\,e^{-\alpha\,T}\,(1-\alpha\,T)$$

This derivative vanishes for $\alpha = 1/T$ and thus the optimal choice is

$$U_0 = U_m\,(1-e^{-\alpha\,T}) = U_m\,(1-\frac{1}{e})\,.$$

**Solution to Question 3–4 :** Use a linear approximation $f(x_0+\Delta x) \approx f(x_0) + f'(x_0)\cdot\Delta x$ and thus $\Delta f = f(x_0+\Delta x) - f(x_0) \approx f'(x_0)\cdot\Delta x$ . The independent variable is $R_1$ and the dependent variable is the angle $\alpha$ .

(a) Use the chain rule.

$$\begin{aligned} \frac{\partial\,\alpha}{\partial R_1} &= \frac{2\,(1+\nu)\,L\,M}{E}\,\frac{\frac{\pi}{2}\,4\,R_1^3}{J^2} = \frac{2\,(1+\nu)\,L\,M}{E\,J}\,\frac{4\,R_1^3}{\left(R_2^4-R_1^4\right)} = \alpha\,\frac{4\,R_1^3}{\left(R_2^4-R_1^4\right)} \\ \Delta\alpha &\approx \frac{\partial\,\alpha}{\partial R_1}\,\Delta R_1 = \alpha\,\frac{4\,R_1^3}{R_2^4-R_1^4}\,\Delta R_1 \end{aligned}$$

(b) Dividing by $\alpha$ leads to

$$\frac{\Delta\alpha}{\alpha} = \frac{4\,R_1^4}{R_2^4-R_1^4}\,\frac{\Delta R_1}{R_1}\,.$$

(c)

$$\begin{aligned} \alpha &= \frac{2\,(1+\nu)\,L}{E}\,\frac{2}{\pi\,\left(R_2^4-R_1^4\right)} \approx 0.0448 = 2.57^\circ \\ |\Delta\alpha| &\leq \alpha\,\frac{4\,R_1^3}{R_2^4-R_1^4}\,|\Delta R_1| \approx 8.89\,|\Delta R_1| \leq \frac{\pi}{180} \approx 0.017453 \\ |\Delta R_1| &\leq 0.0019\text{ m} = 1.9\text{ mm} \end{aligned}$$

This variation is too large to be approximated by a linear function. It would work just fine with a tolerance of $0.1^\circ$ for $\alpha$.

**Solution to Question 3–8 :**

$$\begin{aligned}
A &= \frac{d\,f}{dx} = \frac{\partial\,f}{\partial x} = 2\,x - e^{y-x} \\
B &= \operatorname{grad} f(x,y) = (\frac{\partial\,f}{\partial x}\,,\,\frac{\partial\,f}{\partial y}) = (2\,x - e^{y-x}\,,\,+e^{y-x}\,) \\
C &= \frac{d}{dt}\,\vec{x}(t) = \begin{pmatrix} \dot{x}(t) \\ \dot{y}(t) \end{pmatrix} = \begin{pmatrix} -\sin(t) \\ 3\,t^2 \end{pmatrix} \\
D &= \frac{d}{dt}\,f(x(t),y(t)) = \operatorname{grad} f(x(t),y(t)) \cdot \begin{pmatrix} \dot{x}(t) \\ \dot{y}(t) \end{pmatrix} \\
&= (2\,x(t) - e^{y(t)-x(t)}\,,\,+e^{y(t)-x(t)}\,) \cdot \begin{pmatrix} \dot{x}(t) \\ \dot{y}(t) \end{pmatrix} \\
&= (2\,\cos(t) - e^{t^3-\cos(t)}\,,\,+e^{t^3-\cos(t)}\,) \cdot \begin{pmatrix} -\sin(t) \\ 3\,t^2 \end{pmatrix} \\
&= -\sin(t)\,(2\,\cos(t) - e^{t^3-\cos(t)}) + 3\,t^2\,e^{t^3-\cos(t)}
\end{aligned}$$

**Solution to Question 3–11 :**

(a) The gradient is given by

$$\nabla f(x,y) = \operatorname{grad} f\,(x,y) = (-2\,(x+1)\,,\,-4\,(y-2))$$

(b) Use the above

$$\operatorname{grad} f\,(0,0) = (-2\,(0+1)\,,\,-4\,(0-2)) = (-2\,,\,+8)$$

and thus the direction is determined by the vector

$$\vec{v} = \frac{1}{\sqrt{2^2+8^2}}\,(-2\,,\,+8) \approx (-0.243\,,\,0.970)$$

i.e. at an angle of $\approx 104°$ off the $x$–axis. The corresponding slope is $\|\operatorname{grad} f(0,0)\| = \sqrt{2^2+8^2} \approx 8.25$. This is visualized in Figure 8.1.

(c)

$$\begin{aligned}
\operatorname{grad} f\,(0,0) &= (-2\,,\,8) \\
f\,(x,y) &\approx f\,(0,0) + \operatorname{grad} f\,(0,0) \cdot \begin{pmatrix} x \\ y \end{pmatrix} = -7 + (-2,8) \begin{pmatrix} x \\ y \end{pmatrix} = -7 - 2\,x + 8\,y
\end{aligned}$$

Find the graph of the surface and the tangent plane in Figure 8.1.

(d) The function $f(x,y)$ is a quadratic function and thus it coincides with the quadratic approximation.

```
[x,y] = meshgrid([-10:0.75:10]);
f = 2-(x+1).^2-2*(y-2).^2;
grad = [-2,8];
f_tan = -7 +grad(1)*x+grad(2)*y;
mesh(x,y,f);
hold on
```

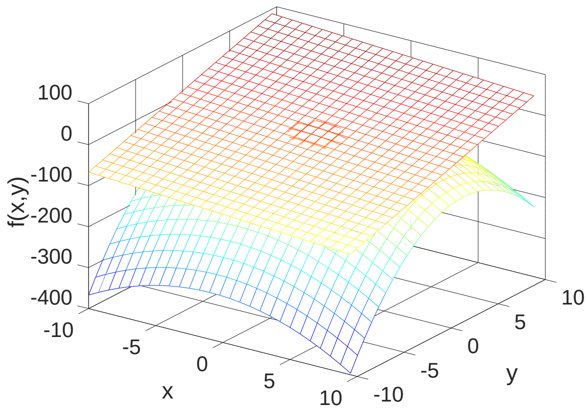



Figure 8.1: Graph of $2-(x+1)^2-2\,(y-2)^2$ and the tangent plane

```
mesh(x,y,f_tan)
hold off
xlabel('x'); ylabel('y'); zlabel('f(x,y)')
view(35,30)
```

**Solution to Question 3–12 :** Since $U=R\,I$ find

$$R=\frac{U}{I}=f\,(U,I)$$

and thus $R=7.3\,\Omega$. To estimate the error use

$$\begin{aligned} R(U+\Delta U,I+\Delta I) &\approx R(U,I)+\frac{\partial\,R}{\partial U}\,\Delta U+\frac{\partial\,R}{\partial I}\,\Delta I = R(U,I)+\frac{1}{I}\,\Delta U-\frac{U}{I^2}\,\Delta I \\ |\Delta R| &\leq \left|\frac{\partial\,R}{\partial U}\,\Delta U\right|+\left|\frac{\partial\,R}{\partial I}\,\Delta I\right|=\left|\frac{1}{I}\,\Delta U\right|+\left|\frac{-U}{I^2}\,\Delta I\right| \end{aligned}$$

For the maximal error we find thus

$$\left|\frac{\Delta R}{R}\right| = \frac{I}{U}\,|\Delta R|\leq\left|\frac{\Delta U}{U}\right|+\left|\frac{\Delta I}{I}\right|=\frac{2}{110}+\frac{0.3}{15}\approx 0.04$$

The maximal relative error is given by $0.04\cdot 7.3\ \Omega\approx 0.3\ \Omega$. In this simple situation the results can be determined as the sum of the relative errors of $U$ and $I$.

**Solution to Question 3–14 :** The necessary condition $\nabla f(x,y)=\vec{0}$ leads to

$$\begin{aligned} \frac{\partial f\,(x_0,y_0)}{\partial x} &= e^{-y_0^2+x_0}(1+x_0) &&=0 \\ \frac{\partial f\,(x_0,y_0)}{\partial y} &= -e^{-y_0^2+x_0}\,2\,y_0\,x_0 &&=0 \end{aligned}$$

with the unique solution $(x_0,y_0)=(-1,0)$. Thus the functions attains its minimal value of $-e^{-1}$ at the point $(x,y)=(-1,0)$. See Figure 8.2.

**Solution to Question 3–17 :** As a parametrization consider

$$\begin{pmatrix} x(t)\\ y(t)\end{pmatrix}=\begin{pmatrix} t\\ 3\,(0.5-t^2)\end{pmatrix}\quad\text{and thus}\quad \vec{ds}=\begin{pmatrix}1\\ -6\,t\end{pmatrix}\,dt$$

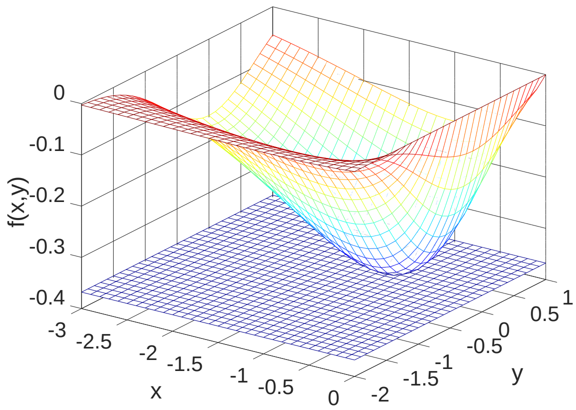



Figure 8.2: Graph of $e^{-y^2+x}\,x$ and the tangent plane at the minimal point

(a) Let $a = \frac{1}{\sqrt{2}}$

$$L = \int_C 1\;ds = \int_{-a}^{a} \sqrt{1+(6\,t)^2}\;dt = \int_{-a}^{a} \sqrt{1+36\,t^2}\;dt \approx 3.44\;[m]$$

This leads to

$$\rho = \frac{M}{L} \approx 0.29\;\left[\frac{kg}{m}\right]$$

```
a = 1/sqrt(2);
L = quad(@(t)sqrt(1+36*t^2),-a,+a)
M = 1; rho = M/L
-->
L = 3.4409
rho =  0.29062
```

(b) One full rotation per second corresponds to an angular velocity of $\omega = 2\,\pi$. At a given value of $x$ is the radius of rotation given by $y$ and thus the velocity by $v = y\,\omega$. Since the basic formula for kinetic energy is $\frac{1}{2}\,m\,v^2$ we obtain a scalar integral for the total kinetic energy.

$$E \;=\; \int_C \frac{1}{2}\,\rho\,y^2\;ds = \frac{\rho}{2}\,\omega^2 \int_{-a}^{a} (3\,(0.5-t^2))^2\,\sqrt{1+(6\,t)^2}\;dt \approx 17.09\;[J]$$

```
omega = 2*pi;
E = rho*omega^2/2*quad(@(t)(3*(1/2-t^2))^2*sqrt(1+36*t^2),-a,a)
-->
E =  17.085
```

**Solution to Question 3–20 :**

(a) One of the possible parametrizations of the straight line is

$$C \quad : \quad \vec{x}\,(t) = \begin{pmatrix} t \\ t+1 \end{pmatrix} \quad \text{where} \quad 0 \le t \le 1$$

and thus

$$\begin{aligned}
\vec{ds}\,(t) &= \dot{\vec{x}}\,(t)\;dt = \begin{pmatrix} 1 \\ 1 \end{pmatrix} dt \\
\int_C \vec{F}\cdot\vec{ds} &= \int_0^1 \begin{pmatrix} t^2-(t+1) \\ (t+1)^2+t \end{pmatrix}\cdot\begin{pmatrix} 1 \\ 1 \end{pmatrix} dt \\
&= \int_0^1 t^2-(t+1)+(t+1)^2+t\;dt \;=\; \int_0^1 2\,t^2+2\,t\;dt \;=\frac{5}{3}\,.
\end{aligned}$$

(b) A parametrization of the first segment is

$$C_1 \quad : \quad \vec{x}\,(t) = \begin{pmatrix} t \\ 1 \end{pmatrix} \quad \text{where} \quad 0\le t\le 1$$

and for the second segment use

$$C_2 \quad : \quad \vec{x}\,(t) = \begin{pmatrix} 1 \\ t \end{pmatrix} \quad \text{where} \quad 1\le t\le 2\,.$$

Then determine

$$\vec{ds}_1\,(t) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} dt \quad \text{and} \quad \vec{ds}_2\,(t) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} dt\,.$$

Thus we obtain the integral

$$\begin{aligned}
\int_C \vec{F}\cdot\vec{ds} &= \int_{C_1} \vec{F}\cdot\vec{ds}_1 + \int_{C_2} \vec{F}\cdot\vec{ds}_2 \\
&= \int_0^1 \begin{pmatrix} t^2-1 \\ 1^2+t \end{pmatrix}\cdot\begin{pmatrix} 1 \\ 0 \end{pmatrix} dt + \int_1^2 \begin{pmatrix} 1^2-t \\ t^2+1 \end{pmatrix}\cdot\begin{pmatrix} 0 \\ 1 \end{pmatrix} dt \\
&= \int_0^1 t^2-1\;dt + \int_1^2 t^2+1\;dt \;=\; -\frac{2}{3}+\frac{10}{3} \;=\; \frac{8}{3}\,.
\end{aligned}$$

(c) One of the parametrizations of the parabola is

$$C \quad : \quad \vec{x}\,(t) = \begin{pmatrix} t \\ t^2+1 \end{pmatrix} \quad \text{where} \quad 0\le t\le 1$$

and thus

$$\begin{aligned}
\vec{ds}\,(t) &= \dot{\vec{x}}\,(t)\;dt = \begin{pmatrix} 1 \\ 2\,t \end{pmatrix} dt \\
\int_C \vec{F}\cdot\vec{ds} &= \int_0^1 \begin{pmatrix} t^2-(t^2+1) \\ (t^2+1)^2+t \end{pmatrix}\cdot\begin{pmatrix} 1 \\ 2\,t \end{pmatrix} dt \\
&= \int_0^1 t^2-(t^2+1)+2\,t\left((t^2+1)^2+t\right)\;dt \\
&= \int_0^1 2\,t^5+4\,t^3+2\,t^2+2\,t-1\;dt \;=\; 2\,.
\end{aligned}$$

Observe that for this example the integral depends on the actual curve, and not the start and end point only. We will later examine that this is caused by

$$\frac{\partial F_1}{\partial y} = -1 \neq +1 = \frac{\partial F_2}{\partial x}\,.$$

**Solution to Question 3–23 :** The height of a narrow column (width $\Delta x$ and depth $\Delta y$) at a point $(x, y)$ is given by $f(x, y) = 1 + x + y$ and then its volume is

$$\Delta V \approx f(x,y)\;\Delta x\,\Delta y = (1+x+y)\;\Delta x\,\Delta y$$

Thus the total volume is given by a double integral

$$\begin{aligned}
V &= \iint\limits_R f(x,y)\;dA \;=\; \iint\limits_R 1+x+y\;dA \\
&= \int_0^2 \left(\int_0^3 1+x+y\;dy\right)\;dx \;=\; \int_0^2 \left(\left(y+x\,y+\frac{1}{2}\,y^2\right)\Big|_{y=0}^{3}\right)\;dx \\
&= \int_0^2 \left(3+x\,3+\frac{1}{2}\,9\right)\;dx \;=\; \left(3\,x+\frac{3}{2}\,x^2+\frac{9}{2}\,x\right)\Big|_{x=0}^{2} \;=\; 3\cdot 2+\frac{3}{2}\,2^2+\frac{9}{2}\,2 \;=\; 21\,.
\end{aligned}$$

The order of the integral can be switched and you will obtain the same result.

$$\begin{aligned}
V &= \int_0^3 \left(\int_0^2 1+x+y\;dx\right)\;dy \;=\; \int_0^3 \left(\left(x+\frac{1}{2}\,x^2+x\,y\right)\Big|_{x=0}^{2}\right)\;dy \\
&= \int_0^3 \left(2+\frac{4}{2}+2\,y\right)\;dy \;=\; \left(2\,y+2\,y+y^2\right)\Big|_{y=0}^{3} \;=\; 2\cdot 3+2\cdot 3+3^2 \;=\; 21
\end{aligned}$$

**Solution to Question 3–24 :** The basic formula for the kinetic energy is $\frac{1}{2}\,m\,v^2$. The velocity of the segment at $(x, y)$ is given by $v = \omega\,r = \omega\,\sqrt{x^2+y^2}$. This leads to the energy of a small sector of

$$\Delta E = \frac{\omega^2\,\rho}{2}\,\sqrt{x^2+y^2}^{\,2}\;h\,\Delta A = \frac{\omega^2\,\rho}{2}\,(x^2+y^2)\,h\,\Delta x\,\Delta y\,.$$

Thus we obtain the kinetic energy $E$ by

$$\begin{aligned}
E &= \iint\limits_R \frac{\omega^2\,\rho}{2}\,(x^2+y^2)\,H\,dA \;=\; \frac{\omega^2\,\rho\,H}{2}\int_0^W \left(\int_0^D x^2+y^2\;dy\right)\;dx \\
&= \frac{\omega^2\,\rho\,H}{2}\int_0^W \left(y\,x^2+\frac{y^3}{3}\Big|_{y=0}^{D}\right)\;dx \;=\; \frac{\omega^2\,\rho\,H}{2}\int_0^W \left(D\,x^2+\frac{D^3}{3}\right)\;dx \\
&= \frac{\omega^2\,\rho\,H}{2}\,\left(\frac{D}{3}\,x^3+\frac{D^3}{3}\,x\right)\Big|_{x=0}^{W} \;=\; \frac{\omega^2\,\rho\,H}{2}\,\left(\frac{D}{3}\,W^3+\frac{D^3}{3}\,W\right) \;=\; \frac{\omega^2}{2}\,\rho\,H\,W\,D\,\frac{W^2+D^2}{3} \\
&= \frac{\omega^2}{2}\,M\,\frac{W^2+D^2}{3}\,.
\end{aligned}$$

**Solution to Question 3–25 :** The basic formula for the kinetic energy is $\frac{1}{2}\,m\,v^2$. The velocity of the segment at $(x, y)$ is given by $v = \omega\,r = \omega\,\sqrt{x^2+y^2}$. This leads to the energy of a small sector of

$$\Delta E = \frac{\omega^2\,\rho}{2}\,\sqrt{x^2+y^2}^{\,2}\;h\,\Delta A = \frac{\omega^2\,\rho}{2}\,(x^2+y^2)\,h\,\Delta x\,\Delta y\,.$$

For a given value of $-R \leq x \leq +R$ the $y$ coordinate satisfies $-\sqrt{R^2 - x^2} \leq y \leq +\sqrt{R^2 - x^2}$. Thus we obtain the kinetic energy $E$ by

$$\begin{aligned} E &= \iint\limits_R \frac{\omega^2\,\rho}{2}\,(x^2+y^2)\,H\,dA = \frac{\omega^2\,\rho\,H}{2}\int_{-R}^{+R}\left(\int_{-\sqrt{R^2-x^2}}^{+\sqrt{R^2-x^2}} x^2+y^2\,dy\right)\,dx \\ &= \frac{\omega^2\,\rho\,H}{2}\int_{-R}^{+R}\left(\left(y\,x^2+\frac{1}{3}\,y^3\right)\Big|_{-\sqrt{R^2-x^2}}^{+\sqrt{R^2-x^2}}\right)\,dx \\ &= \frac{\omega^2\,\rho\,H}{2}\int_{-R}^{+R}\left(2\,\sqrt{R^2-x^2}\,\left(x^2+\frac{1}{3}\,(R^2-x^2)\right)\right)\,dx \end{aligned}$$

and the integral is rather tedious to compute. For this problem you **have to use polar coordinates**, see Question 3–26.

**Solution to Question 3–26 :** For this circular plate the radius varies between 0 and $R$ and for the angle $\phi$ we have $0 \leq \phi \leq 2\,\pi$.

$$\begin{aligned} E &= \iint\limits_R \frac{\omega^2\,\rho}{2}\,r^2\,H\,dA = \frac{\omega^2\,\rho\,H}{2}\int_0^R\left(\int_0^{2\,\pi} r^2\cdot r\,d\phi\right)\,dr \\ &= \frac{\omega^2\,\rho\,H}{2}\,2\,\pi\int_0^R r^3\,dr = \frac{\omega^2\,\rho\,H}{2}\,2\,\pi\,\frac{R^4}{4} = \frac{\omega^2}{2}\,M\,\frac{R^2}{2} \end{aligned}$$

**Solution to Question 3–27 :**

(a) The equation of the circle can be solved for $z = \pm\sqrt{r^2-(x-R)^2}$. Using slices of width $\Delta x \ll r$ at $x$ and rotating them about the $z$-axis will generate a circular band of height $2\,\sqrt{r^2-(x-R)^2}$ and radius $x$ and thus a volume of

$$\Delta V \approx 2\,\pi\,x\,2\,\sqrt{r^2-(x-R)^2}\,\Delta x$$

This leads to the total volume

$$\begin{aligned} V &= \int_{R-r}^{R+r} 2\,\pi\,x\,2\,\sqrt{r^2-(x-R)^2}\,dx = 4\,\pi\int_{R-r}^{R+r} x\,\sqrt{r^2-(x-R)^2}\,dx \\ &= 4\,\pi\int_{-r}^{+r}(u+R)\,\sqrt{r^2-u^2}\,du \quad \text{use the substitution} \quad u = x-R \\ &= 2\,\pi\int_{-r}^{+r} R\,2\,\sqrt{r^2-u^2}\,du = 2\,\pi\,R\;\;(\text{area of disk with radius } r) = 2\,\pi\,R\,\pi\,r^2\,. \end{aligned}$$

(b) We use polar coordinates $(\rho, \phi)$ with

$$\begin{aligned} x &= R+\rho\,\cos\phi \quad \text{radius of rotation} \\ y &= \rho\,\sin\phi \\ dA &= \rho\,d\phi\,d\rho \\ dV &= 2\,\pi\,x\,dA = 2\,\pi\,(R+\rho\,\cos\phi)\,\rho\,d\phi\,d\rho\,. \end{aligned}$$

This leads to the volume $V$ of the torus

$$\begin{aligned} V &= \int_{\rho=0}^{r}\left(\int_{\phi=0}^{2\,\pi} 2\,\pi\,(R+\rho\,\cos\phi)\,\rho\,d\phi\right)\,d\rho = 2\,\pi\int_{\rho=0}^{r}\left((R\,\phi+\rho\,\sin\phi)\Big|_{\phi=0}^{2\,\pi}\right)\,\rho\,d\rho \\ &= 2\,\pi\int_{\rho=0}^{r}(2\,\pi\,R)\,\rho\,d\rho = 4\,\pi^2\,R\,\frac{1}{2}\,\rho^2\Big|_{\rho=0}^{r} = 2\,\pi^2\,R\,r^2\,. \end{aligned}$$

**Solution to Question 3–28 :** Quietly assuming that the density is constant and using the basic formula $\frac{m}{2}\,v^2 = \frac{m}{2}\,\omega^2\,x^2$ for the kinetic energy we find

$$\begin{aligned}
dE &= \frac{1}{2}\,\frac{M}{V}\,\omega^2\,(R+\rho\,\cos\phi)^2\,dV \;=\; \frac{1}{2}\,\frac{M}{V}\,\omega^2\,(R+\rho\,\cos\phi)^2\,2\pi\,(R+\rho\,\cos\phi)\,dA \\
&= \frac{M\,\omega^2\,\pi}{V}\,(R+\rho\,\cos\phi)^3\,\rho\,d\phi\,d\rho \\
E &= \frac{M\,\omega^2\,\pi}{V}\,\int_{\rho=0}^{r}\left(\int_{\phi=0}^{2\,\pi}(R+\rho\,\cos\phi)^3\,d\phi\right)\,\rho\,d\rho \\
&= \frac{M\,\omega^2\,\pi}{V}\,\int_{\rho=0}^{r}\left(\int_{\phi=0}^{2\,\pi} R^3 + 3\,R^2\,\rho\,\cos\phi + 3\,R\,\rho^2\,\cos^2\phi + \rho^3\,\cos^3\phi\,d\phi\right)\,\rho\,d\rho \\
&= \frac{M\,\omega^2\,\pi}{V}\,\int_{\rho=0}^{r}\left(2\,\pi\,R^3 + 3\,R\,\rho^2\,\pi\right)\,\rho\,d\rho \;=\; \frac{M\,\omega^2\,\pi}{V}\,\left(\pi\,R^3\,\rho^2 + \frac{3\,R}{4}\,\rho^4\,\pi\right)\Big|_{\rho=0}^{r} \\
&= \frac{M\,\omega^2\,\pi^2}{V}\,\left(R^3\,r^2 + \frac{3\,R\,r^4}{4}\right) \;=\; \frac{M\,\omega^2\,\pi^2}{V}\,\left(R^2 + \frac{3\,r^2}{4}\right)\,R\,r^2\,.
\end{aligned}$$

To evaluate the above integral we used

$$\int_0^{2\,\pi}\cos\phi\;d\phi = 0 \quad , \quad \int_0^{2\,\pi}\cos^2\phi\;d\phi = \pi \quad \text{and} \quad \int_0^{2\,\pi}\cos^3\phi\;d\phi = 0\,.$$

**Solution to Question 3–30 :**

(a) Use the components $(\sin\theta, 0, \cos\theta)$ of the normal vector on the sphere at $\psi = 0$, i.e. along the $x$ axis.

(b)

$$\begin{aligned}
F_z &= \iint_{\text{sphere}} p\,\cos\theta\,dA \;=\; \int_0^{\pi/2} 2\,\pi\,R\,\sin\theta\;p\,\cos\theta\,R\,d\theta \\
&= 2\,\pi\,R^2\,p\,\frac{\sin^2\theta}{2}\Big|_{\theta=0}^{\pi/2} \;=\; \pi\,R^2\,p
\end{aligned}$$

(c) The area of the section is (approximatively) given by $2\,\pi\,R\,\Delta R$ and thus the stress by

$$\text{stress} = \frac{\text{force}}{\text{area}} = \frac{\pi\,R^2\,p}{2\,\pi\,R\,\Delta R} = \frac{R\,p}{2\,\Delta R}$$

(d)

$$\begin{aligned}
\begin{pmatrix} F_x \\ F_y \end{pmatrix} &= \iint_{\text{sphere}} p\,\sin\theta\,\begin{pmatrix}\cos\phi \\ \sin\phi\end{pmatrix}\,dA \\
&= \int_{\theta=0}^{\pi/2}\int_{\phi=0}^{2\pi} p\,\sin\theta\,\begin{pmatrix}\cos\phi \\ \sin\phi\end{pmatrix}\,R^2\sin\theta\,d\phi\,d\theta \\
&= \int_{\theta=0}^{\pi/2} p\,\sin\theta\,\begin{pmatrix}0 \\ 0\end{pmatrix}\,R^2\sin\theta\,d\theta \;=\; \begin{pmatrix}0 \\ 0\end{pmatrix}
\end{aligned}$$

**Solution to Question 3–34 :**

- The boundary $C$ of this disk $D \in \mathbb{R}^2$ may be parametrized by

$$\begin{pmatrix} x(t) \\ y(t) \end{pmatrix} = \begin{pmatrix} 2\,\cos t \\ 2\,\sin t \end{pmatrix} \quad \text{for} \quad 0 \le t \le 2\,\pi$$

Thus we have

$$\vec{ds} = \begin{pmatrix} -2\,\sin t \\ +2\,\cos t \end{pmatrix} dt \quad \text{and} \quad ds = \|\vec{ds}\| = 2\,dt$$

The outer unit normal vector at time (or angle) $t$ is given by

$$\vec{n}(t) = \begin{pmatrix} \cos t \\ \sin t \end{pmatrix}$$

- This leads to

$$\begin{aligned} \text{flux} &= \oint_C \langle \vec{v}\,,\,\vec{n}\rangle\; ds = \int_{t=0}^{2\,\pi} \langle \begin{pmatrix} 1 \\ 2\,\cos t + 2\,\sin t \end{pmatrix}, \begin{pmatrix} \cos t \\ \sin t \end{pmatrix} \rangle\; 2\,dt \\ &= \int_{t=0}^{2\,\pi} 2\,\cos(t) + 4\,\sin^2(t) + 2\,\cos(t)\,\sin(t)\;dt \;= \ldots = \;4\,\pi \end{aligned}$$

The result can also be determined using the divergence theorem, see Question 3–38.

**Solution to Question 3–38 :** For the vector field $\vec{v}$ we find

$$\text{div}\begin{pmatrix} 1 \\ x+y \end{pmatrix} = \frac{\partial\,1}{\partial x} + \frac{\partial\,(x+y)}{\partial y} = 0 + 1$$

and thus

$$\text{flux} = \oint_C \langle \vec{v}\,,\,\vec{n}\rangle\; ds = \iint\limits_D \text{div}(\vec{v})\; dA \;=\; \iint\limits_D 1\; dA \;=\; 4\,\pi$$

This computation is shorter and easier that the one used in Question 3–34.

**Solution to Question 3–39 :**

(a) For this vector field we have

$$\vec{F} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \sqrt{x^2+y^2+z^2}\;\vec{e}_r = r\,\vec{e}_r$$

and thus $F_r = r$ and $F_\phi = F_\theta = 0$. Using Table 3.4

$$\text{div}\,\vec{F} = \frac{1}{r^2}\,\frac{\partial\,(r^2\,F_r)}{\partial r} \;=\; \frac{1}{r^2}\,\frac{\partial\,(r^2\cdot r)}{\partial r} \;=\; \frac{1}{r^2}\,3\,r^2 \;=\; 3\,.$$

Thus the total flux out of this ball is 3 times the volume of the ball, i.e.

$$\text{flux} = 3\,\frac{4}{3}\,\pi\,R^3 \;=\; \pi\,108\,.$$

This could be computed by a triple integral over the ball, an exercise in integration.

$$\begin{aligned}
\text{flux} &= \iiint\limits_B \operatorname{div} \vec F \, dV = \iiint\limits_B 3 \, dV \\
&= \int_{r=0}^{R} \int_{\theta=0}^{\pi} \int_{\phi=0}^{2\pi} 3\, r^2 \sin\theta \, d\phi \, d\theta \, dr \\
&= \int_{r=0}^{R} \int_{\theta=0}^{\pi} 2\pi\, 3\, r^2 \sin\theta \, d\theta \, dr \\
&= \int_{r=0}^{R} 2\cdot 2\pi\, 3\, r^2 \, dr = 4\pi R^3 = \pi\, 108
\end{aligned}$$

(b) Use spherical coordinates with the variables $0 \le \phi \le 2\pi$ and $0 \le \theta \le \pi$ and the surface element (see Table 3.2 on page 23)

$$dS = R^2 \sin\theta \, d\theta \, d\phi \, .$$

On the sphere with radius $R = 3$ find the outer unit normal vector

$$\vec n = \frac{1}{R} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

and thus

$$\langle \vec n \, , \vec F \rangle = \langle \frac{1}{R} \begin{pmatrix} x \\ y \\ z \end{pmatrix} , \begin{pmatrix} x \\ y \\ z \end{pmatrix} \rangle = \frac{x^2 + y^2 + z^2}{R} = R \, .$$

Thus we find

$$\begin{aligned}
\text{flux} &= \oiint\limits_S \langle \vec n \, , \vec F \rangle \, dS = \int_{\phi=0}^{2\pi} \left( \int_{\theta=0}^{\pi} R \cdot R^2 \sin\theta \, d\theta \right) d\phi \\
&= \int_{\phi=0}^{2\pi} \left( 2\, R \cdot R^2 \right) \, d\phi = 4\pi R^3 = \pi\, 108 \, .
\end{aligned}$$

**Solution to Question 3–40 :**

(a) With

$$\operatorname{div} \vec F = \frac{\partial\, x}{\partial x} + \frac{\partial\, y}{\partial y} + \frac{\partial\, 3}{\partial z} = 2$$

find

$$\text{flux} = \iiint\limits_B \operatorname{div} \vec F \, dV = \iiint\limits_B 2 \, dV = 2 \cdot \frac{4}{3} \pi R^3 = 72\, \pi \, .$$

(b) Use spherical coordinates with the variables $0 \le \phi \le 2\pi$ and $0 \le \theta \le \pi$ and the surface element (see Table 3.2 on page 23)

$$dS = R^2 \sin\theta \, d\theta \, d\phi \, .$$

On the sphere with radius $R = 3$ find the outer unit normal vector

$$\vec{n} = \frac{1}{3} \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

and thus

$$\langle \vec{n}\,,\,\vec{F}\rangle \;=\; \langle \frac{1}{3} \begin{pmatrix} x \\ y \\ z \end{pmatrix}, \begin{pmatrix} x \\ y \\ 3 \end{pmatrix} \rangle \;=\; \frac{x^2+y^2}{3} + z\,.$$

Thus we find

$$\begin{aligned}
\text{flux} \;&=\; \oiint_S \langle \vec{n}\,,\,\vec{F}\rangle\; dS \;=\; \int_{\phi=0}^{2\pi} \left( \int_{\theta=0}^{\pi} \left( \frac{x^2+y^2}{3} + z \right) 3^2 \sin\theta \; d\theta \right) d\phi \\
&=\; \int_{\phi=0}^{2\pi} \left( \int_{\theta=0}^{\pi} \left( \frac{(3\sin\theta\cos\phi)^2 + (3\sin\theta\sin\phi)^2}{3} + 3\cos\theta \right) 3^2 \sin\theta \; d\theta \right) d\phi \\
&=\; \int_{\phi=0}^{2\pi} \left( \int_{\theta=0}^{\pi} \left( 3\sin^2\theta + 3\cos\theta \right) 3^2 \sin\theta \; d\theta \right) d\phi \\
&=\; \int_{\phi=0}^{2\pi} \left( 3^3 \int_{\theta=0}^{\pi} \sin^3\theta \; d\theta + 3^3 \int_{\theta=0}^{\pi} \cos\theta\,\sin\theta \; d\theta \right) d\phi \\
&=\; \int_{\phi=0}^{2\pi} \left( 3^3\,\frac{4}{3} + 0 \right) d\phi \;=\; \int_{\phi=0}^{2\pi} 36 \; d\phi \;=\; 72\,\pi\,.
\end{aligned}$$

**Solution to Question 4–4 :**

(a) First rewrite the differential equation in explicit form

$$\frac{d}{dt}\,I(t) = -\frac{R}{L}\,I(t) + \frac{E_0}{L}$$

and then use separation of variable.

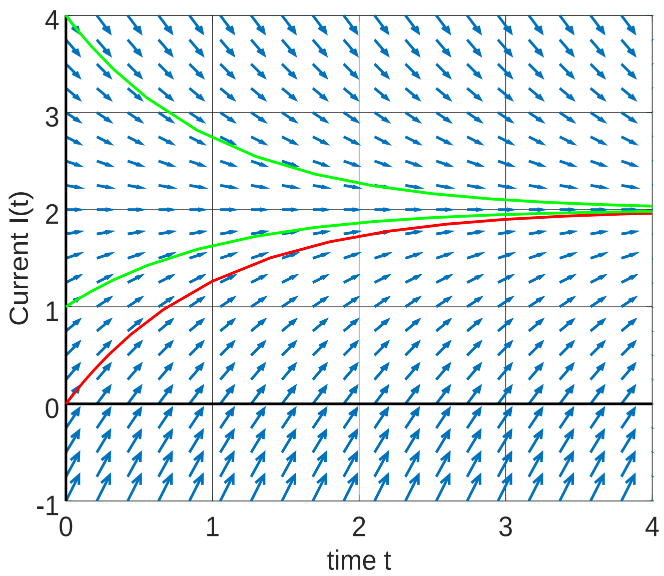
$$\begin{aligned}
\frac{d}{dt}\,I(t) \;&=\; -\frac{R}{L}\,I(t) + \frac{E_0}{L} \\
\frac{d\,I}{-\frac{R}{L}\,I + \frac{E_0}{L}} \;&=\; 1\;dt \\
\int \frac{L}{E_0 - R\,I}\;dI \;&=\; \int 1\;dt \\
-\frac{L}{R}\,\log|E_0 - R\,I| \;&=\; t + c_1 \\
|R\,I - E_0| \;&=\; \frac{R}{L}\,\exp(-\frac{t\,R + c_1}{L}) \;=\; |c_2|\,\exp(-\frac{R}{L}\,t) \\
R\,I \;&=\; E_0 + c_2\,\exp(-\frac{R}{L}\,t) \\
I(t) \;&=\; \frac{E_0}{R} + \frac{c_2}{R}\,\exp(-\frac{R}{L}\,t)
\end{aligned}$$

Using the initial condition $I(0) = 0$ leads to $c_2 = -E_0$ and thus the solution

$$I(t) = \frac{E_0}{R} + \frac{E_0}{R}\,\exp(-\frac{R}{L}\,t) = \frac{E_0}{R}\left(1 - e^{-\frac{R}{L}\,t}\right)$$

(b) With $E_0 = 2$, $L = 1$ and $R = 1$ the differential equation is $\frac{d}{dt}\, I(t) = -\, I(t) + 2$ and the vector field to be examined is $(1\,,\, 2 - I)^T$.

- For values of $I > 2$ we have $\frac{d}{dt}\, I < 0$ and thus the curve $I(t)$ has a negative slope.
- For values of $I < 2$ we have $\frac{d}{dt}\, I > 0$ and thus the curve $I(t)$ has a positive slope.
- For the special case $I = 2$ we find $\frac{d}{dt}\, I = 0$ and thus a zero slope. As a consequence we have the particular solution $I(t) = 2 = \frac{E_0}{R}$.




This example is typical for linear, inhomogeneous equations of order one. The figure was generated by

```
t = linspace(0,4,20);   x = linspace(-1,4,21);
[t,x] = meshgrid(t,x);

f_t = ones(size(t));   f_x = 2-x;

figure(1); clf
quiver(t,x,f_t,f_x)
xlabel('time t'); ylabel('Current I(t)')
axis([0 4 -1 4])
[t1, x1] = ode45 (@(t,x)2-x,[0 4],0);
[t2, x2] = ode45 (@(t,x)2-x,[0 4],4);
[t3, x3] = ode45 (@(t,x)2-x,[0 4],1);
hold on
plot([0 4],[0 0],'k',[0 0],[-1 4],'k')
plot(t1,x1,'r',t2,x2,'g',t3,x3,'g')
hold off
print -dpdfwrite ODEInductance.pdf
```

**Solution to Question 4–6 :**

(a) To examine $\dot y(t) + 2\,y(t) = 2\,t$ use the ansatz $y_p(t) = c_1\,t + c_0$. Using the differential equation leads to

$$\frac{d}{dt}\,(c_1\,t + c_0) + 2\,(c_1\,t + c_0) \quad = \quad 2\,t$$

$$\begin{aligned}
c_1 + 2\,c_1\,t + 2\,c_0 &= 2\,t \\
&\left\{\begin{array}{rl} 2\,c_1 & = 2 \\ c_1 + 2\,c_0 & = 0 \end{array}\right. \\
y_p(t) &= t - \frac{1}{2}
\end{aligned}$$

Thus the general solution is

$$y(t) = c\,e^{-2\,t} + t - \frac{1}{2}$$

(b) To examine $\dot y(t) + 2\,y(t) = \cos(3\,t)$ use the ansatz $y_p(t) = c_1\,\cos(3\,t) + c_2\,\sin(3\,t)$. Using the differential equation leads to

$$\begin{aligned}
\frac{d}{dt}\,(c_1\,\cos(3\,t) + c_2\,\sin(3\,t)) + 2\,(c_1\,\cos(3\,t) + c_2\,\sin(3\,t)) &= \cos(3\,t) \\
(-3\,c_1\,\sin(3\,t) + 3\,c_2\,\cos(3\,t)) + (2\,c_1\,\cos(3\,t) + 2\,c_2\,\sin(3\,t)) &= \cos(3\,t) \\
\cos(3\,t)\,(3\,c_2 + 2\,c_1) + \sin(3\,t)\,(-3\,c_1 + 2\,c_2) &= \cos(3\,t) \\
\left\{\begin{array}{rl} 2\,c_1 + 3\,c_2 & = 1 \\ -3\,c_1 + 2\,c_2 & = 0 \end{array}\right. \\
c_1 = \frac{2}{13} \qquad c_2 = \frac{3}{13} \\
y_p(t) = \frac{1}{13}\,(2\,\cos(3\,t) + 3\,\sin(3\,t))
\end{aligned}$$

Thus the general solution is

$$y(t) = c\,e^{-2\,t} + \frac{1}{13}\,(2\,\cos(3\,t) + 3\,\sin(3\,t))$$

(c) To examine $\dot y(t) + 2\,y(t) = 3\,e^{-t}$ use the ansatz $y_p(t) = c_1\,e^{-t}$. Using the differential equation leads to

$$\begin{aligned}
\frac{d}{dt}\,\left(c_1\,e^{-t}\right) + 2\,c_1\,e^{-t} &= 3\,e^{-t} \\
c_1\,e^{-t}\,(-1+2) &= 3\,e^{-t} \\
c_1 &= 3 \\
y_p(t) &= 3\,e^{-t}
\end{aligned}$$

Thus the general solution is

$$y(t) = c\,e^{-2\,t} + 3\,e^{-t}$$

**Solution to Question 4–9 :** First examine the characteristic equation

$$\lambda^2 - 2\lambda + 5 = (\lambda - 1)^2 + 2^2 = 0$$

with the solutions $\lambda_{1,2} = 1 \pm i\,2$. Thus the general solution is

$$y(t) = e^t\,(c_1\,\cos(2t) + c_2\,\sin(2t))\ .$$

An elementary calculation leads to

$$\dot y(t) = e^t\,(c_1\,\cos(2t) + c_2\,\sin(2t)) + e^t\,(-2\,c_1\,\sin(2t) + 2\,c_2\,\cos(2t))\ .$$

Use this and the initial conditions to setup a system of equations for $c_1$ and $c_2$

$$c_1 = 2 \quad \text{and} \quad c_1 + 2\,c_2 = 4$$

with the solution $c_1 = 2$ and $c_2 = 1$. The solution of the initial value problem is

$$y(t) = e^t\ (2\,\cos(2\,t) + \sin(2\,t))\ .$$

**Solution to Question 4–10 :** First examine the characteristic equation

$$\lambda^2 + 4\lambda + 4 = (\lambda + 2)^2 = 0$$

with the solutions $\lambda = \lambda_1 = \lambda_2 = -2$. Thus the general solution of the ODE is

$$y(t) = e^{-2\,t}\ (c_1 + c_2\,t)\ .$$

An elementary calculation leads to

$$\dot y(t) = -2\,e^{-2\,t}\ (c_1 + c_2\,t) + c_2\,e^{-2\,t}\ .$$

Use this and the initial conditions to setup a system of equations for $c_1$ and $c_2$.

$$\left\{\begin{array}{rl} e^{-2}\ (c_1 + c_2) & = 2 \\ -2\,e^{-2}\ (c_1 + c_2) + c_2\,e^{-2} & = -1 \end{array}\right.$$

or

$$\left\{\begin{array}{rcrl} c_1 & + & c_2 & = 2\,e^2 \\ -2\,c_1 & - & c_2 & = -e^2 \end{array}\right.$$

with the solutions $c_1 = -e^{+2}$ and $c_2 = 3\,e^{+2}$. The solution of the initial value problem is

$$y(t) = e^{-2t+2}\ (-1 + 3\,t)\ .$$

**Solution to Question 4–14 :**

(a) The homogeneous solution of

$$\ddot y(t) - 4\,y(t) = 2\,e^{3\,t}$$

is determined by the characteristic equation $\lambda^2 - 4 = 0$ with the solutions $\lambda = \pm 2$. Thus the general homogeneous solution is

$$y_h(t) = c_1\,e^{+2\,t} + c_2\,e^{-2\,t}$$

The first ansatz for the particular solution is $y_p(t) = k\,e^{3\,t}$. Plugging this into the ODE leads to

$$3^2\,k\,te^{3\,t} - 4\,k\,e^{3\,t} = 2\,e^{3\,t}$$

or $(9-4)\,k = 2$ with the solution $k = \frac{2}{5}$. This leads to the general solution

$$y(t) = c_1\,e^{+2\,t} + c_2\,e^{-2\,t} + \frac{2}{5}\,e^{3\,t}$$

where the constants $c_1$ and $c_2$ would have to be determined using initial conditions.

(b) Since $f(t) = 7\,\sin(t)$ use the ansatz

$$y_p(t) = a\,\cos\,t + b\,\sin\,t$$

and plug into the ODE

$$(-a\,\cos\,t - b\,\sin\,t) + (-a\,\sin\,t + b\,\cos\,t) + 4\;(a\,\cos\,t + b\,\sin\,t) = 7\,\sin\,t$$

This leads to

$$\sin\,t\;(-b - a + 4\,b) + \cos\,t\;(-a + b + 4\,a) = 7\,\sin\,t$$

Separating the terms with $\sin\,t$ and $\cos\,t$ leads to the system

$$\begin{cases} -a & + & 3\,b = 7 \\ 3\,a & + & b\; = 0 \end{cases}$$

with the solutions $a = -7/10$ and $b = 21/10$. Thus we have a particular solution

$$y_p(t) = \frac{-7}{10}\;\cos\,t + \frac{21}{10}\;\sin\,t\,.$$

(c) The homogeneous solution of

$$\ddot{y}(t) - 4\,y(t) = 3\,e^{2\,t}$$

is determined by the characteristic equation $\lambda^2 - 4 = 0$ with the solutions $\lambda = \pm 2$. Thus the general homogeneous solution is

$$y_h(t) = c_1\,e^{+2\,t} + c_2\,e^{-2\,t}\,.$$

The first ansatz for the particular solution is $y_p(t) = k\,e^{2\,t}$. Plugging this into the ODE leads to

$$2^2\,k\,e^{2\,t} - 4\,k\,e^{2\,t} = 3\,e^{2\,t}\,.$$

This equation **can not** be solved for $k$, since the expression on the left equals zero. Thus we have to modify the ansatz and try with $y_p(t) = k\,t\,e^{2t}$. This leads to

$$\begin{aligned} \ddot{y}(t) - 4\,y(t) &= 3\,e^{2\,t} \\ \left(2^2\,k\,t\,e^{2t} + 4\,k\,e^{2\,t}\right) - 4\,k\,t\,e^{2\,t} &= 3\,e^{2\,t} \\ t\,\left(2^2\,k - 4\,k\right) + 4\,k &= 3 \\ k &= \frac{3}{4}\,. \end{aligned}$$

Thus we find the general solution of the inhomogeneous ODE

$$y(t) = y_h(t) + y_p(t) = c_1\,e^{2\,t} + c_2\,e^{-2\,t} + \frac{3}{4}\,t\,e^{2\,t}$$

where the constants $c_1$ and $c_2$ would have to be determined using initial conditions.

**Solution to Question 6–1 :** The real Fourier coefficients are to be determined by the integrals

$$\begin{aligned} a_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} x\;\cos\,(n\,x)\;dx \qquad n = 0, 1, 2, 3 \ldots \\ b_n &= \frac{1}{\pi}\int_{-\pi}^{\pi} x\;\sin\,(n\,x)\;dx \qquad n = 1, 2, 3, 4 \ldots \end{aligned}$$

Since $x\cos(nx)$ is an odd function for all $n \in \mathbb{N}$ and the domain of integration $[-\pi\,,\,+\pi]$ is symmetric we conclude $a_n = 0$. To evaluate $b_n$ use integration by parts or a table of integrals to conclude

$$\pi\, b_n = \int_{-\pi}^{\pi} x\ \sin(n\,x)\;dx = -x\,\frac{1}{n}\ \cos(n\,x)\Big|_{-\pi}^{\pi} + \int_{-\pi}^{\pi} 1\,\frac{1}{n}\ \cos(n\,x)\;dx = \frac{-2\,\pi}{n}\ (-1)^n + 0\,.$$

Thus we find

$$b_n = (-1)^{n+1}\,\frac{2}{n} \quad \text{for} \quad n \geq 1$$

and the Fourier series is given by

$$\begin{aligned} x &\sim \sum_{n=1}^{\infty} b_n\ \sin(n\,x) = \sum_{n=1}^{\infty} (-1)^{n-1}\,\frac{2}{n}\ \sin(n\,x) \\ &= 2\ \sin x - \sin(2\,x) + \frac{2\ \sin(3\,x)}{3} - \frac{2\ \sin(4\,x)}{4} + \frac{2\ \sin(5\,x)}{5} - \ldots \end{aligned}$$

The complex coefficients are given by the integral

$$\begin{aligned} 2\,\pi\,c_n &= \int_{-\pi}^{+\pi} x\ \exp(-i\,n\,x)\;dx \;=\; x\,\frac{\exp(-i\,n\,x)}{-i\,n}\Big|_{x=-\pi}^{+\pi} - \frac{1}{-i\,n}\int_{-\pi}^{+\pi} 1\ \exp(-i\,n\,x)\;dx \\ &= \pi\,\frac{\exp(-i\,n\,\pi)}{-i\,n} - (-\pi)\,\frac{\exp(+i\,n\,\pi)}{-i\,n} \\ &= \frac{\pi}{-i\,n}\,\left(\cos(n\,\pi) - i\,\sin(n\,\pi) + \cos(n\,\pi) + i\,\sin(n\,\pi)\right) \;=\; i\,\frac{2\,\pi}{n}\,(-1)^n\,. \end{aligned}$$

This result $c_n = \frac{i}{n}\,(-1)^n$ is consistent with

$$c_n = \frac{1}{2}\,(a_n - i\,b_n) = \frac{-i}{2}\,(-1)^{n+1}\,\frac{2}{n}$$

and the complex Fourier approximation is given by

$$\begin{aligned} x &\sim 2\ \mathrm{Re}\sum_{n=1}^{\infty} c_n\ \exp(i\,n\,x) \;=\; 2\ \mathrm{Re}\sum_{n=1}^{\infty} \frac{i}{n}\,(-1)^n\,(\cos(n\,x) + i\,\sin(n\,x)) \\ &= \sum_{n=0}^{\infty} \frac{2}{n}\,(-1)^{n+1}\ \sin(n\,x)\,. \end{aligned}$$

Using *Octave*/`MATLAB` the graph of the function and its Fourier approxumation can be generated by the code below.

**FourierOfx.m**

```
function FourierOfx()
x = linspace(-4,4,1001);

function y = f(x)
  y = mod(x+pi,2*pi)-pi;
end%function

function res = F(x,n)
  res = zeros(size(x));
  for k = 1:n
    res = res - 2*(-1)^k*sin(k*x)/k;
  end%for
end%function
```

```
y1 = F(x,1); y4 = F(x,4); y10 = F(x,10); y100 = F(x,100);
figure(1)
plot(x,f(x),x,[y1;y4;y10])
legend('original','1 term','4 terms','10 terms','location','north'); grid on

figure(2)
plot(x,y100,x,f(x));    grid on
end%function
```

**Solution to Question 6–4 :** We have to examine the odd extension of this function, i.e.

$$f(x) = \begin{cases} -1 & \text{for} \quad -\pi < x \leq 0 \\ +1 & \text{for} \quad 0 < x \leq \pi \end{cases} .$$

Since the function is odd we know that $a_n = 0$ and

$$b_n = \frac{2}{\pi} \int_0^\pi \sin(n\,x)\;dx = \frac{-2}{n\pi}\;\cos(n\,x)\Big|_0^\pi = \begin{cases} \dfrac{4}{n\pi} & \text{for odd} \quad n \\ 0 & \text{for even} \quad n \end{cases} .$$

Thus the Fourier sine approximation is given by

$$f\,(x) \;=\; 1 \sim \frac{4}{\pi}\;(\frac{\sin x}{1} + \frac{\sin 3x}{3} + \frac{\sin 5x}{5} + \frac{\sin 7x}{7} + \ldots)\,.$$

The odd, $2\pi$-periodic extension of this function has jumps of size 2 at all multiples of $\pi$. Thus we have uniform convergence between the multiples of $\pi$ and the phenomenon of Gibbs at the multiples of $\pi$.

**Solution to Question 6–5 :**

$$\begin{aligned} x \;&\sim\; \sum_{n=1}^{\infty}(-1)^{n-1}\frac{2}{n}\;\sin\,(n\,x) \\ &=\; 2\,\sin\,x - \sin\,(2\,x) + \frac{2\,\sin\,(3\,x)}{3} - \frac{2\,\sin\,(4\,x)}{4} + \frac{2\,\sin\,(5\,x)}{5} - \ldots \\ u(x) \;&\sim\; \sum_{n=1}^{\infty}(-1)^{n-1}\frac{1}{n^2}\frac{2}{n}\;\sin\,(n\,x) \\ &=\; 2\,\sin\,x - \frac{2\,\sin\,(2\,x)}{2^3} + \frac{2\,\sin\,(3\,x)}{3^3} - \frac{2\,\sin\,(4\,x)}{4^3} + \frac{2\,\sin\,(5\,x)}{5^3} - \ldots \end{aligned}$$

Since the denominators $n^3$ become quickly very large, this series converges very quickly.

**Solution to Question 6–7 :**

(a) For $N = 4$ we have $w = \exp(i\,\frac{2\,\pi}{4}) = \exp(i\,\frac{\pi}{2} = \cos\frac{\pi}{2} + i\,\sin\frac{\pi}{2} = i$ and thus

$$w = i \quad , \quad w^2 = -1 \quad , \quad w^3 = -i \quad , \quad w^4 = 1 \quad , \quad w^5 = i \quad , \quad w^6 = -1 \quad , \quad w^7 = -i \quad , \quad \ldots$$

This leads to

$$\mathbf{F}_4 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & i & -1 & -i \\ 1 & -1 & 1 & -1 \\ 1 & -i & -1 & i \end{bmatrix}$$

(b) Use

```
fft(eye(4))
-->
   1 + 0i    1 + 0i    1 + 0i    1 + 0i
   1 + 0i    0 - 1i   -1 + 0i    0 + 1i
   1 + 0i   -1 + 0i    1 + 0i   -1 + 0i
   1 - 0i    0 + 1i   -1 - 0i    0 - 1i
```

and this is the matrix $\overline{\mathbf{F}_4}$. For the inverse DFT by `MATLAB` examine

```
ifft(eye(4))
-->
   0.25 + 0.00i    0.25 + 0.00i    0.25 + 0.00i    0.25 + 0.00i
   0.25 + 0.00i    0.00 + 0.25i   -0.25 + 0.00i    0.00 - 0.25i
   0.25 + 0.00i   -0.25 + 0.00i    0.25 + 0.00i   -0.25 + 0.00i
   0.25 + 0.00i    0.00 - 0.25i   -0.25 + 0.00i    0.00 + 0.25i
```

and this is the matrix $\frac{1}{4}\,\mathbf{F}_4$.

**Solution to Question 6–10 :** Use code similar to the one below.

**DFT_test.m**
```
func = @(t)2*sin(100.2*2*pi*t) - 3*cos(102.5*2*pi*t) + 0.5*sin(150*2*pi*t);
T = 2;  N = 2^10;              % choose the values
t = linspace(0,T-T/N,N);       % generate the sampling times
f = func(t)+ 0.1*randn(1,N); % generate the function with added noise
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
NyquistFrequency = N/T/2

figure(1)
plot(t,f)
xlabel('time t [s]'); ylabel('amplitude')
title('signal')

c = abs(fft(f)*2/N);
freq = [0:N-1]/T;
domain = find(freq<200); % display onle up to 200 Hz

figure(2)
plot(freq(domain),c(domain))
xlabel('frequency [Hz]'); ylabel('amplitude')
title('spectrum')
```

(a) Use $T = 2$ and $N = 2^{10} = 1024$.
The Nyquist frequency is 256 Hz and the amplitudes seem to be correct. The two frequencies at 100.2 and 102.5 can not clearly be separated. This is caused by the fundamental frequency of $\frac{1}{2}$ Hz.

(b) Use $T = 5$ and $N = 2^{10} = 1024$.
The Nyquist frequency is 102.4 Hz and thus too small. The amplitudes are obviously wrong.

(c) Use $T = 20$ and $N = 2^{14} = 16384$.
The Nyquist frequency is 409.6 Hz and thus large enough. The two frequencies at 100.2 and 102.5 are clearly separated. This is caused by the fundamental frequency of $\frac{1}{20}$ Hz.

(d) Use $T = 20$ and $N = 2^{16} = 65536$.
Very little difference to the previous case.

Even adding sizable noise does not change these results.

**Solution to Question 6–13 :** The answer is based on

$$\begin{aligned}
F(\nu) &= \int_{-\infty}^{+\infty} f(t)\, e^{-2\,\pi\,\nu\,t}\, dt \\
&= \int_{-\infty}^{+\infty} f(t)\, \cos(2\,\pi\,\nu\,t)\, dt - i \int_{-\infty}^{+\infty} f(t) \sin(2\,\pi\ nu\,t)\, dt \\
f(t) &\sim \int_{-\infty}^{+\infty} F(\nu)\, e^{+2\,\pi\,\nu\,t}\, dt \\
&= \int_{-\infty}^{+\infty} F(\nu)\, \cos(2\,\pi\,\nu\,t)\, dt + i \int_{-\infty}^{+\infty} F(\nu) \sin(2\,\pi\ nu\,t)\, dt
\end{aligned}$$

(a) If $f(-t) = f(t)$ then the integral of the odd function $f(t)\,\sin(2\,\pi\,\nu\,t)$ over the symmetric interval vanishes and we have

$$F(\nu) = 2 \int_0^{+\infty} f(t)\, \cos(2\,\pi\,\nu\,t)\, dt \in \mathbb{R}$$

Since this function satisfies $F(-\nu) = F(\nu)$ we conclude

$$\begin{aligned}
f(t) &\sim \int_{-\infty}^{+\infty} F(\nu)\, (\cos(2\,\pi\,\nu\,t) + i\, \sin(2\,\pi\,\nu\,t))\, dt \\
&= \int_{-\infty}^{+\infty} F(\nu)\, \cos(2\,\pi\,\nu\,t)\, dt = 2 \int_0^{+\infty} F(\nu)\, \cos(2\,\pi\,\nu\,t)\, dt
\end{aligned}$$

(b) If $f(-t) = -f(t)$ then the integral of the odd function $f(t)\,\cos(2\,\pi\,\nu\,t)$ over the symmetric interval vanishes and we have

$$F(\nu) = i\,2 \int_0^{+\infty} f(t)\, \sin(2\,\pi\,\nu\,t)\, dt \in i\mathbb{R}$$

Since this function satisfies $F(-\nu) = -F(\nu)$ we conclude

$$\begin{aligned}
f(t) &\sim \int_{-\infty}^{+\infty} F(\nu)\, (\cos(2\,\pi\,\nu\,t) + i\, \sin(2\,\pi\,\nu\,t))\, dt \\
&= i \int_{-\infty}^{+\infty} F(\nu)\, \sin(2\,\pi\,\nu\,t)\, dt = i\,2 \int_0^{+\infty} F(\nu)\, \sin(2\,\pi\,\nu\,t)\, dt
\end{aligned}$$

**Solution to Question 6–17 :** Since the Fourier transform of the function and its derivative exist we know the limits

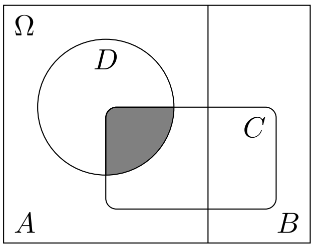
$$\lim_{t\to\pm\infty} f\,(t) = \lim_{t\to\pm\infty} f'\,(t) = 0$$

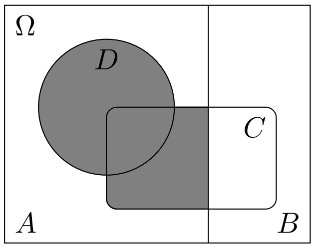
We choose (artificially) $\tau_0 = -\infty$ and $\tau_{n+1} = \infty$ and integrations by part lead to

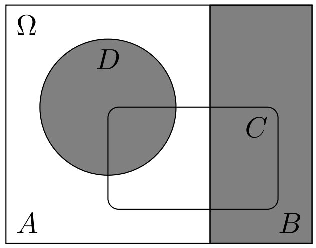
$$\begin{aligned}
\mathcal{F}[\,f'\,(t)\,]\,(\nu) &= \int_{-\infty}^{\infty} f'\,(t)\; e^{-i\,2\,\pi\,\nu\,t}\, dt \\
&= \sum_{k=0}^{n} \int_{\tau_k}^{\tau_{k+1}} f'\,(t)\; e^{-i\,2\,\pi\,\nu\,t}\, dt \\
&= \sum_{k=0}^{n} \left( f\,(t)\; e^{-i\,2\,\pi\,\nu\,t} \Big|_{t=\tau_k}^{\tau_{k+1}} + i\,2\,\pi\,\nu \int_{\tau_k}^{\tau_{k+1}} f\,(t)\; e^{-i\,2\,\pi\,\nu\,t}\, dt \right) \\
&= i\,2\,\pi\;\nu\;F\,(\nu) - \sum_{k=1}^{n} \left(f\,(\tau_k+) - f\,(\tau_k-)\right)\; e^{-i\,2\,\pi\,\tau_k}\,.
\end{aligned}$$

**Solution to Question 7–6 :**

(a) The operators union ($\cup$), intersection ($\cap$) and complement ($^c$) operate on sets and the addition ($+$) on numbers. We get that i) and ii) are meaningful and iii) and iv) are not.

(b) i): 

ii): 

iii): 


**Solution to Question 7–7 :**

(a) We can easily see that $A \subseteq B \subseteq C$ and $A \cup B \cup C = C$. Using the fact that the hit probabilities are proportional to the areas on the dartboard, the ratio of the probabilities must be $7^2\pi : 14^2\pi : 21^2\pi = 1 : 4 : 9$. Additionally, we know that only every second dart hits the board, therefore, $\mathrm{P}(C) = 0.5$. Consequently,

$$\mathrm{P}(B) = \frac{4}{9} \cdot \mathrm{P}(C) = \frac{2}{9}, \qquad \mathrm{P}(A) = \frac{1}{9} \cdot \mathrm{P}(C) = \frac{1}{18}\ .$$

(b) The probability to hit the middle circular ring is $\mathrm{P}(B \setminus A) = \mathrm{P}(B) - \mathrm{P}(A) = \frac{2}{9} - \frac{1}{18} = \frac{1}{6}$, using the fact that $A \subseteq B$.

**Solution to Question 7–10 :** If the elementary events have equal probability we can calculate the probability of event $A$ as the number of elementary events in $A$ divided by the total number of elementary events.

(a) $\Omega = \{(1,1),(1,2),\dots,(1,6),(2,1),(2,2),\dots,(2,6),\dots,(6,6)\}$, $|\Omega| = 36$.

(b) $\mathrm{P}(\{\text{elementary event}\}) = \frac{1}{|\Omega|} = \frac{1}{36}$.

(c) $E_1 = \{(1,6),(2,5),(3,4),(4,3),(5,2),(6,1)\}$;
number of favourable cases: $|E_1| = 6$;
number of possible cases: $|\Omega| = 36$;
$\mathrm{P}(E_1) = \frac{|E_1|}{|\Omega|} = \frac{6}{36} = \frac{1}{6}$.

(d) $E_2 = \{(1,1),(2,1),(1,2)\}$;
$\mathrm{P}(E_2) = \frac{|E_2|}{|\Omega|} = \frac{3}{36} = \frac{1}{12}$.

(e) $E_3 = \{(1,1),(1,3),(1,5),(3,1),(3,3),(3,5),(5,1),(5,3),(5,5)\}$;
$\mathrm{P}(E_3) = \frac{|E_3|}{|\Omega|} = \frac{9}{36} = \frac{1}{4}$.

(f)

$$\begin{aligned} \mathrm{P}(E_2 \cup E_3) = &\ \mathrm{P}(E_2) + \mathrm{P}(E_3) - \mathrm{P}(E_2 \cap E_3) \\ = &\ \mathrm{P}(E_2) + \mathrm{P}(E_3) - \mathrm{P}(\{(1,1)\}) \\ = &\ \tfrac{3}{36} + \tfrac{9}{36} - \tfrac{1}{36} = \tfrac{11}{36}. \end{aligned}$$

**Solution to Question 7–11 :**

(a) $\mathrm{P}(A) = 5/26 \approx 0.19$

(b) Out of the $2^3 = 8$ (equally probable) elementary events 7 are favourable, so, $\mathrm{P}(B) = 7/8 = 0.875$.

(c) For each of the 12 cube borders, only 8 dice have exactly two green faces, so, $\mathrm{P}(C) = (12 \cdot 8)/1000 = 0.096$.

**Solution to Question 7–13 :** P(„Bonnie attends class") = $\mathrm{P}(B) = 0.7$ and P(„Clyde attends class") = $\mathrm{P}(C) = 0.55$. We know that

$$\mathrm{P}(B \cap C) = 0.4 \neq 0.7 \cdot 0.55 = 0.385.$$

So, the attendances of Bonnie and Clyde are *dependent* events.

**Solution to Question 7–14 :** P(„rabbit is shot") = 1 − P(„rabbit survives") = $1 - 0.8 \cdot 0.8 \cdot 0.8 = 1 - 0.512 = 0.488$. So, the rabbit was correct – the probability of being shot is less than 50%!

**Solution to Question 7–18 :**

(a) $\mathrm{P}(6\,|\text{„even spots"}) = \frac{\mathrm{P}(6 \cap \text{„even spots"})}{\mathrm{P}(\text{„even spots"})} = 2 \cdot \frac{1}{6} = \frac{1}{3}$

(b) $\mathrm{P}(6\,|\text{„odd spots"}) = \frac{\mathrm{P}(6 \cap \text{„odd spots"})}{\mathrm{P}(\text{„odd spots"})} = \frac{0}{1/2} = 0$

(c) $\mathrm{P}(\text{„even spots"}|\text{„spots<4"}) = \frac{\mathrm{P}(\text{„even spots"} \cap \text{„spots<4"})}{\mathrm{P}(\text{„spots<4"})} = \frac{1}{6} / \frac{1}{2} = \frac{1}{3}$

**Solution to Question 7–19 :**

(a) $\mathrm{P}(A) = \frac{600}{900} \approx 0.667$: The probability that a randomly selected person from the sample is vaccinated is approx. 0.667.

(b) $\mathrm{P}(B) = \frac{180}{900} = 0.2$: The probability that a randomly selected person from the sample got the flu is 0.2.

(c) $\mathrm{P}(A \cap B) = \frac{60}{900} \approx 0.067$: The probability that a randomly selected person from the sample got sick, even though he or she is vaccinated, is approx. 0.067.

(d) $\mathrm{P}(B|A) = \frac{P(A \cap B)}{P(A)} = \frac{60}{600} = 0.1$: Given that a person is vaccinated, he or she got sick with probability 0.1.

(e) $\mathrm{P}(A|B) = \frac{P(A \cap B)}{P(B)} = \frac{60}{180} \approx 0.333$: Given that a person got the flu, he or she is vaccinated with probability 0.333 (approx.)

(f) $\mathrm{P}(A^c \cap B) = \frac{120}{900} \approx 0.133$: The probability that a randomly selected person from the sample is not vaccinated and got sick is approx. 0.133.

(g) $\mathrm{P}(B|A^c) = \frac{P(A^c \cap B)}{P(A^c)} = \frac{120}{300} = 0.4$: Given that a person is not vaccinated, he or she got the flu with probability 0.4.

**Solution to Question 7–20 :** Let us define the following events: $B$ =„Swiss resident holds a bachelor or comparable degree" and $E$ =„Swiss resident has a job". Therefore, the following information is known:
$\mathrm{P}(B) = 0.3$, also $\mathrm{P}(B^c) = 0.7$, as well as $\mathrm{P}(E^c|B) = 0.05$ and $\mathrm{P}(E^c|B^c) = 0.1$.

(a) $\mathrm{P}(E^c) = \mathrm{P}(E^c|B) \cdot \mathrm{P}(B) + \mathrm{P}(E^c|B^c) \cdot \mathrm{P}(B^c) = 0.05 \cdot 0.3 + 0.1 \cdot 0.7 = 0.085$

(b) $\mathrm{P}(B|E^c) = \frac{\mathrm{P}(B \cap E^c)}{\mathrm{P}(E^c)} = \frac{\mathrm{P}(E^c|B) \cdot \mathrm{P}(B)}{\mathrm{P}(E^c)} = \frac{0.05 \cdot 0.3}{0.085} \approx 0.1765$

(c) No, because $\mathrm{P}(B \cap E^c) = 0.05 \cdot 0.3 = 0.015 \neq \mathrm{P}(B) \cdot \mathrm{P}(E^c) = 0.3 \cdot 0.085 = 0.0255$.

**Solution to Question 7–26 :** The sum of the single probabilities must be 1 in order to define a probability mass function, therefore: $7a^2 + 6a = 1$. It follows that $a = \frac{1}{7}$ (the solution $-1$ of the quadratic equation does obviously not make sense).

**Solution to Question 7–29 :** Let $X$ be your profit with possible values -5, -3, -1, 2, 4, 6. Obviously, $\mathrm{P}(X = x_i) = \frac{1}{6}$ for $i = 1, 2, \ldots, 6$. Therefore, $\mathrm{E}(X) = \frac{1}{6}(-5 - 3 - 1 + 2 + 4 + 6) = 0.5$. Of course you're gonna play – on average you win 0.5 CHF per game!

**Solution to Question 7–30 :**

(a) $\mathrm{E}(X) = \sum_{i=1}^{6} i \cdot \frac{1}{6} = \frac{21}{6} = 3.5$
$\mathrm{Var}(X) = \sum_{i=1}^{6} (i - 3.5)^2 \cdot \frac{1}{6} = 2.92$

(b) $\mathrm{E}(X) = \sum_{i=1}^{n} i \cdot \frac{1}{n} = \frac{1}{n} \cdot \frac{n(n+1)}{2} = \frac{n+1}{2}$
$\mathrm{Var}(X) = \sum_{i=1}^{n} (i - \mathrm{E}(X))^2 \cdot \frac{1}{n} = \frac{1}{n} \left( \frac{n(n+1)(2n+1)}{6} - 2 \cdot \frac{n+1}{2} \cdot \frac{n(n+1)}{2} + \frac{n(n+1)^2}{4} \right) = \frac{n^2-1}{12}$

**Solution to Question 7–31 :**

(a) It is a Laplace experiment with 16 elementary events $(i, k)$ for $i, k = 1, 2, 3, 4$. Your gain is denoted by the random variable $X$ with possible values $4, -1, -2, -3$ and

$$\begin{aligned}
\{X = 4\} &= \{(1,1), (2,2), (3,3), (4,4)\} \\
\{X = -1\} &= \{(1,2), (2,1), (2,3), (3,2), (3,4), (4,3)\} \\
\{X = -2\} &= \{(1,3), (3,1), (2,4), (4,2)\} \\
\{X = -3\} &= \{(1,4), (4,1)\}
\end{aligned}$$

It follows that $\mathrm{E}(X) = 4 \cdot \frac{4}{16} - 1 \cdot \frac{6}{16} - 2 \cdot \frac{4}{16} - 3 \cdot \frac{2}{16} = -\frac{1}{4}$ and $\mathrm{Var} = \left(4 + \frac{1}{4}\right)^2 \cdot \frac{4}{16} + \left(-1 + \frac{1}{4}\right)^2 \cdot \frac{6}{16} + \left(-2 + \frac{1}{4}\right)^2 \cdot \frac{4}{16} + \left(-3 + \frac{1}{4}\right)^2 \cdot \frac{2}{16} \approx 6.438$.

(b) $\mathrm{E}(X) = x \cdot \frac{4}{16} - 1 \cdot \frac{6}{16} - 2 \cdot \frac{4}{16} - 3 \cdot \frac{2}{16} = 0 \implies x = 5$ , i.e. the payout would need to be 6 times your stake.

**Solution to Question 7–34 :** The count of tail is binomially distributed with parameters 12 and 0.5. The winning probability is therefore $\binom{12}{6} \cdot 0.5^{12} \approx 0.2256$.

**Solution to Question 7–35 :**

(a) Let $X$ be the number of contaminated samples in one collective sample. The probability that a sample is contaminated is $\pi = 0.02$. Under the assumption that all samples are independent, $X$ is binomially distributed: $X \sim \mathrm{Bin}(n = 10, \pi = 0.02)$.

The probability to find no contamination in the sample is given by

$$\mathrm{P}(X = 0) = \binom{10}{0} \cdot 0.02^0 \cdot 0.98^{10} = 0.98^{10} = 0.8171.$$

Another possible solution: Each sample is clean with a probability of 0.98, independently of the other samples. Therefore we have

$$\mathrm{P}(\text{all samples are clean}) = \prod_{i=1}^{10} \mathrm{P}(i\text{-th sample is clean}) = 0.98^{10} = 0.8171.$$

(b) The random variable $Y$ can only have the values 1 or 11, because:

1. If all samples are clean, we are done after one analysis: $Y = 1$.
2. If at least one sample is contaminated, then the collective sample is contaminated and we need to check all 10 samples separately: $Y = 11$.

Hence,

$$\begin{aligned} \mathrm{P}(Y=1) &= \mathrm{P}(\text{no sample is contaminated}) = 0.8171 \ , \\ \mathrm{P}(Y=11) &= 1 - \mathrm{P}(Y=1) = 0.1829 \ . \end{aligned}$$

(c) The average number of analyses for one collective sample is given through the expected value of $Y$:

$$\mathrm{E}(Y) = \sum_{k=0}^{\infty} k \cdot \mathrm{P}(Y=k) = 1 \cdot \mathrm{P}(Y=1) + 11 \cdot \mathrm{P}(Y=11) = 1 \cdot 0.8171 + 11 \cdot 0.1829 = 2.8293 \ .$$

On average, we save $10 - 2.8293 = 7.1707 \approx 7$ analyses.

**Solution to Question 7–37 :**

(a) Let $X$ be the number of serious accidents per week: $X \sim \mathrm{Po}(2)$ and the probability is $\mathrm{P}(X > 5) = 1 - \mathrm{P}(X \le 5) \approx 0.017$.

(b) Let $Y$ be the number of serious accidents per day: $Y \sim \mathrm{Po}(2/7)$ and the probability is $\mathrm{P}(Y > 1) = 1 - \mathrm{P}(Y \le 1) \approx 0.034$.

**Solution to Question 7–38 :**

(a) The number of calls per 5 minutes $X$ is Poisson distributed with parameter $\lambda_X = \frac{12}{60} \cdot 5 = 1$. Therefore, the probability is $\mathrm{P}(X = 0) \approx 0.368$.

(b) The number of calls per 10 minutes $Y$ is Poisson distributed with parameter $\lambda_Y = 2 \cdot \lambda_X = 2$. Therefore, the probability is $\mathrm{P}(Y \ge 5) = 1 - \mathrm{P}(Y < 5) = 1 - \mathrm{P}(Y \le 4) \approx 0.053$.

(c) The number of calls per 20 minutes $Z$ is Poisson distributed with parameter $\lambda_Z = 4 \cdot \lambda_X = 2 \cdot \lambda_Y = 4$. Therefore, the probability is $\mathrm{P}(Z \le 6) \approx 0.889$

**Solution to Question 7–39 :**

(a) The number of defective screws $X$ is binomially distributed with parameters $n = 500$ and $\pi = 0.001$:

$$\mathrm{P}(X \ge 2) = 1 - \mathrm{P}(X < 2) = 1 - F_X(1) = 1 - \sum_{k=0}^{1} \binom{500}{k} 0.001^k \cdot 0.999^{500-k} \approx 0.0901$$

(b) The number of defective screws is approximately Poisson distributed with parameter $n \cdot \pi = 500 \cdot 0.001 = 0.5$. Therefore, $\mathrm{P}(X \ge 2) \approx 1 - \sum_{k=0}^{1} \mathrm{e}^{-0.5} \frac{0.5^k}{k!} \approx 0.0902$.

**Solution to Question 7–42 :** We first prove the equality in case of discrete random variables.

$$\begin{aligned} \mathrm{Var}(X) &\overset{\text{def}}{=} \sum_{k=1}^{\infty} (x_k - \mathrm{E}(X))^2 \, p(x_k) = \sum_{k=1}^{\infty} \left(x_k^2 + \mathrm{E}(X)^2 - 2x_k\mathrm{E}(X)\right) p(x_k) \\ &= \underbrace{\sum_{k=1}^{\infty} x_k^2 \, p(x_k)}_{\mathrm{E}(X^2)} + \mathrm{E}(X)^2 - 2\mathrm{E}(X) \underbrace{\sum_{k=1}^{\infty} x_k \, p(x_k)}_{\mathrm{E}(X)} = \mathrm{E}(X^2) - \mathrm{E}(X)^2 \end{aligned}$$

In case of a continuous random variable the proof is similar.

$$\begin{aligned}
\mathrm{Var}(X) &\stackrel{\text{def}}{=} \int_{\mathbb{R}} (x-\mathrm{E}(X))^2 f(x)\,dx = \int_{\mathbb{R}} \left(x^2+\mathrm{E}(X)^2-2x\mathrm{E}(X)\right) f(x)\,dx \\
&= \underbrace{\int_{\mathbb{R}} x^2 f(x)\,dx}_{\mathrm{E}(X^2)} + \mathrm{E}(X)^2 - 2\mathrm{E}(X)\underbrace{\int_{\mathbb{R}} x\, f(x)\,dx}_{\mathrm{E}(X)} = \mathrm{E}(X^2)-\mathrm{E}(X)^2
\end{aligned}$$

**Solution to Question 7–43 :**

(a) The area under the curve $f_X$ needs to be 1, i.e. $\int_{-\infty}^{\infty} f_X\,\mathrm{d}x = \int_0^1 ax\,\mathrm{d}x = 1$. It follows that $a = 2$.

(b)

$$F_X(x) = \begin{cases} 0, & \text{for } x < 0 \\ x^2, & \text{for } x \in [0,1] \\ 1, & \text{for } x > 1 \end{cases}$$

(c) $\mathrm{P}\left(\frac{1}{3} \leq X \leq \frac{3}{4}\right) = \mathrm{P}\left(X \leq \frac{3}{4}\right) - \mathrm{P}\left(X \leq \frac{1}{3}\right) = F_X\left(\frac{3}{4}\right) - F_X\left(\frac{1}{3}\right) \approx 0.451$
$\mathrm{P}\left(X \leq \frac{1}{2}\right) = 0.25$
$\mathrm{P}\left(X \geq \frac{3}{4}\right) \approx 0.438$

(d) $\mathrm{E}(X) = \int_{-\infty}^{\infty} x \cdot f_X(x)\,\mathrm{d}x = \int_0^1 x \cdot 2x\,\mathrm{d}x = \left[\frac{2}{3}x^3\right]_{x=0}^1 = \frac{2}{3}$
$\mathrm{Var}(X) = \int_{-\infty}^{\infty} x^2 \cdot f_X(x)\,\mathrm{d}x - \left(\frac{2}{3}\right)^2 = \int_0^1 2x^3\,\mathrm{d}x - \frac{4}{9} = \frac{1}{2} - \frac{4}{9} = \frac{1}{18}$

**Solution to Question 7–46 :**

(a) $\mathrm{P}(-1 \leq Z \leq 1) = \mathrm{P}(Z \leq 1) - \mathrm{P}(Z \leq -1) = \Phi(1) - \Phi(-1) = 2\Phi(1) - 1 \approx 0.6827$
(*Note*: $\Phi(x) = F_Z(x)$ denotes the CDF of $Z$)

(b) $\mathrm{P}(-2 \leq Z \leq 2) = 2\Phi(2) - 1 \approx 0.9545$

(c) $\mathrm{P}(-3 \leq Z \leq 3) = 2\Phi(3) - 1 \approx 0.9973$

Compare your results to the well-known rule of thumb:

- For a normally distributed variable, approximately **68%** of the measured values lie within the interval mean $\pm$ one standard deviation.
- For a normally distributed variable, approximately **95%** of the measured values lie within the interval mean $\pm$ twice the standard deviation.
- For a normally distributed variable, approximately **99%** of the measured values lie within the interval mean $\pm$ three times the standard deviation.

(d) $\mathrm{P}(Z \leq 1) = \Phi(1) \approx 0.8413$

(e) $\mathrm{P}(|Z| \geq 0.5) = 1 - \mathrm{P}(-0.5 \leq Z \leq 0.5) = 2 - 2\Phi(0.5) \approx 0.6171$

(f) $\mathrm{P}(-3 \leq Z \leq 1) = \Phi(1) - \Phi(-3) = \Phi(1) + \Phi(3) - 1 \approx 0.8400$

**Solution to Question 7–47 :**

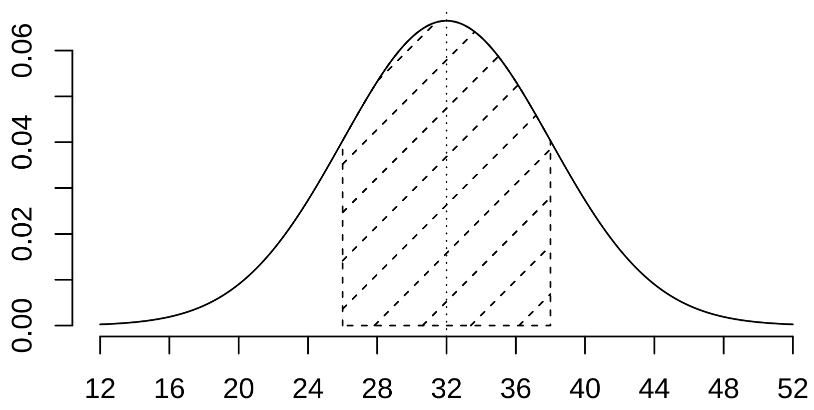
(a) $1 - \mathrm{P}(X \geq 4.98) = \mathrm{P}(X \leq 4.98) = \mathrm{P}\left(\frac{X-5}{0.02} \leq \frac{4.98-5}{0.02}\right) = \mathrm{P}(Z \leq -1) \approx 0.159$

(b) $1 - \mathrm{P}(X \leq 5.05) \approx 0.006$

(c) $\mathrm{P}(|X - 5| \geq 0.03) = 1 - \mathrm{P}(-0.03 \leq X - 5 \leq 0.03) \approx 0.134$

**Solution to Question 7–48 :**

(a)




(b) Let $X$ be the lead content of the sample. It holds that

$$X \sim \mathcal{N}(\mu, \sigma^2) \qquad \text{with } \mu = 32 \text{ and } \sigma^2 = 6^2.$$

Then, $\mathrm{P}(X \leq 40) \approx 0.9088$.

(c) $\mathrm{P}(X \leq 27) \approx 0.2023$

(d) We choose $c$ such that $\mathrm{P}(X \leq c) = 0.975$. Using the quantile function we get that $c \approx 43.76$.

(e) This time we choose $c$ such that $\mathrm{P}(X \leq c) = 0.1 \implies c \approx 24.31$.

(f) $\mathrm{P}(26 \leq X \leq 38) = \mathrm{P}(X \leq 38) - \mathrm{P}(X < 26) \approx 0.6827$

**Solution to Question 7–50 :** Let $T$ be the waiting time in minutes. Therefore, $T$ is exponentially distributed with parameter $\lambda = 4$.

(a)

$$\mathrm{P}(T \leq 0.5) = \int_0^{0.5} 4\mathrm{e}^{-4t}\,\mathrm{d}t = 1 - \mathrm{e}^{-2} \approx 0.8647$$

(b) We need to find $k$ such that $\mathrm{P}(T \leq k) = 0.95$, i.e. we need the 95%-quantile of the random variable $T$. The probability that there will elapse less than 0.7489 minutes (that is about 45 seconds) between two emails is 0.95.

**Solution to Question 7–51 :**

(a) Let $X$ be the number of accidents in one week. We assume that $X \sim \mathrm{Po}(4)$: $\mathrm{P}(X > 5) = 1 - \mathrm{P}(X \leq 5) \approx 0.2149$.

(b) Let $T$ be the waiting time between two accidents. We assume that $T \sim \text{Exp}(4)$

$$\mathrm{P}(T > 2) = \int_2^\infty 4\mathrm{e}^{-4t}\,\mathrm{d}t = \mathrm{e}^{-8} \approx 0.0003$$

**Solution to Question 7–57 :**

(a) If we calculate the sum of each row and each column, we see that the overall value of the probabilities is 1. We get the following table:

| $X/Y$ | 1 | 2 | 3 | $\sum$ |
|---|---|---|---|---|
| 1 | 0.05 | 0.08 | 0.12 | 0.25 |
| 2 | 0.14 | 0.19 | 0.09 | 0.42 |
| 3 | 0.22 | 0.08 | 0.03 | 0.33 |
| $\sum$ | 0.41 | 0.35 | 0.24 | 1.00 |

The marginal distribution of $X$ therefore is
$X = 1$ with probability 0.25
$X = 2$ with probability 0.42 and
$X = 3$ with probability 0.33.

The marginal distribution of Y therefore is
$Y = 1$ with probability 0.41
$Y = 2$ with probability 0.35 and
$Y = 3$ with probability 0.24.

(b) No. If they were independent, the following statement would hold: $p_{X,Y}(x, y) = p_X(x) \cdot p_Y(y)$ for all $x, y = 1, 2, 3$. For example $p_{X,Y}(1, 1) = 0.05 \neq p_X(1) \cdot p_Y(1) = 0.25 \cdot 0.41$.

(c) The probability of $X$ being on a low level is $\mathrm{P}(X = 1) = 0.25$.
If we know the value of $Y$, the probability changes, as we now have a conditional probability. $X$ can still take the same values (1,2 and 3), but as we already know that $Y$ takes the value 2, only the middle column of the table is important to us. As we still need a total probability of 1, we need to adjust the probabilities, with which $X$ takes its values. We do that with the formula

$$p_{X|Y=2}(i) = \mathrm{P}(X = i|Y = 2) = \frac{\mathrm{P}(X = i, Y = 2)}{\mathrm{P}(Y = 2)} = \frac{p_{X,Y}(x, y)}{p_Y(2)} \;.$$

We have $p_{X|Y=2}(1) = \mathrm{P}(X = 1|Y = 2) = \frac{0.08}{0.35} \approx 0.2286$.

(d) No, you can either show that by calculating the two values $p_{Y|X=1}(2)$ and $p_{Y|X=3}(2)$ or by argumentation.

In the cases of $p_{X,Y}(1, 2)$ and $p_{X,Y}(3, 2)$ we want to know the probabilities that $Y$ takes the value 2 and $X$ takes the value 1 (respectively 3). But we have just the probabilities what value $X$ and $Y$ are going to take.

In the case of $p_{Y|X=1}(2)$ and $p_{Y|X=3}(2)$ we already know the value of $X$. So we only need to know with which probability $Y$ has the value 2, when $X$ takes the value 1 (respectively 3).

As the random variables $X$ and $Y$ are dependent, there is a difference.

**Solution to Question 7–58 :**

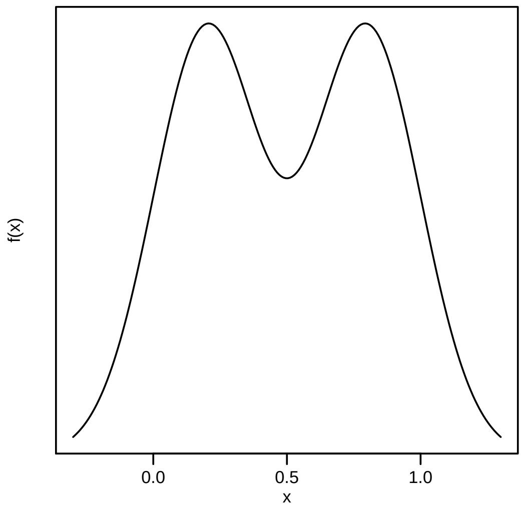
(a)

| $X\backslash Y$ | 1 | 2 | 5 | |
|---|---|---|---|---|
| 1 | 0 | $\frac{1}{6}$ | $\frac{1}{6}$ | $\frac{1}{3}$ |
| 2 | $\frac{1}{6}$ | 0 | $\frac{1}{6}$ | $\frac{1}{3}$ |
| 5 | $\frac{1}{6}$ | $\frac{1}{6}$ | 0 | $\frac{1}{3}$ |
| | $\frac{1}{3}$ | $\frac{1}{3}$ | $\frac{1}{3}$ | 1 |

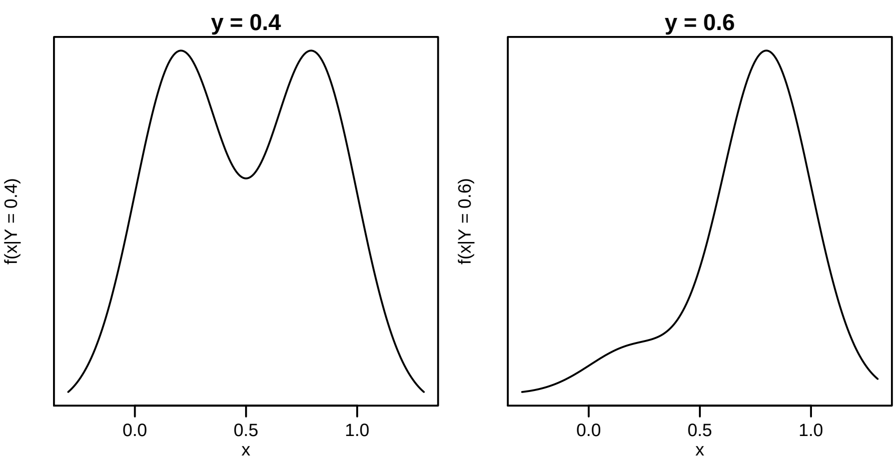
(b) The expected value of the first coin is $\mathrm{E}(X) = \frac{1}{3}(1+2+5) = \frac{8}{3} \approx 2.67$ CHF; the expected value of the second coin $\mathrm{E}(Y)$ is naturally $8/3$ CHF as well. Using the linearity of the expected value, we get $\mathrm{E}(Z) = \mathrm{E}(X) + \mathrm{E}(Y) = \frac{16}{3} \approx 5.33$ CHF.

**Solution to Question 7–64 :**

(a) The plot shows that the joint probability density function has a point symmetry in $(0.5, 0.4)$. This means in particular that the marginal probability density of $X$ must be symmetric around $x = 0.5$. The marginal density attains two maxima, at $x = 0.2$ and at $x = 0.8$:




(b) Conditional probabilities correspond to scaled cuts of the joint probability distribution, in our case at the line $y = 0.4$ (which contains the point of symmetry $(0.5, 0.4)$) and at the line $y = 0.6$ (going through one of the maxima):



**Solution to Question 7–65 :**

(a) It is true that $f_{X,Y}(x,y) \geq 0$ and

$$\begin{aligned}
\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} f_{X,Y}(x,y)\,\mathrm{d}x\,\mathrm{d}y &= \int_{x=0}^{\infty} \mathrm{e}^{-x}\int_{y=0}^{\infty} 2\mathrm{e}^{-2y}\,\mathrm{d}y\,\mathrm{d}x \\
&= \int_{x=0}^{\infty} \mathrm{e}^{-x}\left[-\mathrm{e}^{-2y}\right]_{y=0}^{\infty}\mathrm{d}x \\
&= \int_{x=0}^{\infty} \mathrm{e}^{-x}\,\mathrm{d}x = \left[-\mathrm{e}^{-x}\right]_{x=0}^{\infty} = 1.
\end{aligned}$$

(b)

$$\begin{aligned}
\mathrm{P}(X>1, Y<1) &= \int_{x=1}^{\infty}\int_{y=0}^{1} 2\mathrm{e}^{-2y}\,\mathrm{d}y\,\mathrm{d}x \\
&= \int_{x=1}^{\infty} \mathrm{e}^{-x}\left[-\mathrm{e}^{-2y}\right]_{y=0}^{1}\mathrm{d}x \\
&= \left(1-\mathrm{e}^{-2}\right)\int_{x=1}^{\infty}\mathrm{e}^{-x}\,\mathrm{d}x = \left(1-\mathrm{e}^{-2}\right)\left(\mathrm{e}^{-1}\right) \approx 0.3181
\end{aligned}$$

(c) We need to calculate the marginal distributions first:

$$f_X(x) = \begin{cases} \mathrm{e}^{-x}\int_0^\infty 2\mathrm{e}^{-2y}\,\mathrm{d}y = \mathrm{e}^{-x}, & 0 \leq x < \infty \\ 0, & \text{else} \end{cases}$$

$$f_Y(y) = \begin{cases} 2\mathrm{e}^{-2y}\int_0^\infty \mathrm{e}^{-x}\,\mathrm{d}x = 2\mathrm{e}^{-2y}, & 0 \leq y < \infty \\ 0, & \text{else} \end{cases}$$

It follows that $f_{X,Y}$ is indeed the product of the two marginal distributions $f_X$ and $f_Y$. Therefore, the random variables $X$ and $Y$ are independent!

**Solution to Question 7–67 :** Let $S = X+Y$ with expected value $\mathrm{E}(S) = \mathrm{E}(X)+\mathrm{E}(Y) = a+b$. It follows that

$$\begin{aligned}
\mathrm{Var}(S) &= \mathrm{E}\left(S-(a+b)\right)^2 = \mathrm{E}\left((X-a)+(Y-b)\right)^2 \\
&= \mathrm{E}(X-a)^2 + \mathrm{E}(Y-b)^2 + 2\,\mathrm{E}\left((X-a)(Y-b)\right) \\
&= \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\cdot\mathrm{Cov}(X,Y).
\end{aligned}$$

**Solution to Question 7–68 :**

(a) Let $X_1$ be the number of spots of the first throw and $X_2$ be the number of spots of the second throw. It follows that $\mathrm{E}(X_1+X_2) = \mathrm{E}(X_1)+\mathrm{E}(X_2) = 2\cdot\mathrm{E}(X_1) = 2\cdot 3.5 = 7$. Due to the fact that the two throws are independent of each other, we get for the variance: $\mathrm{Var}(X_1+X_2) = \mathrm{Var}(X_1)+\mathrm{Var}(X_2) = 2\cdot\mathrm{Var}(X_1) \approx 2\cdot 2.9 = 5.8$.

(b) Let $X$ be the number of spots. Then, $\mathrm{E}(2\cdot X) = 2\cdot\mathrm{E}(X) = 7$ and $\mathrm{Var}(2\cdot X) = 4\cdot\mathrm{Var}(X) \approx 11.6$.

(c) Let $X_1$ be the number of spots on the top side and $X_2$ be the number of spots on the bottom side of the dice. Obviously, the sum of the number of spots is always 7. It follow that $X_2 = 7 - X_1$ and the variance of $X_1+X_2$ must be 0. Arithmetically, of course, we get the same: $\mathrm{E}(X_1+X_2) = \mathrm{E}(7) = 7$ and $\mathrm{Var}(X_1+X_2) = \mathrm{Var}(X_1+7-X_1) = \mathrm{Var}(7) = 0$.

**Solution to Question 7–69 :**

(a)

| $X \backslash Y$ | 0 | 1 | 2 | $\mathrm{P}(X = x_i)$ |
|---|---|---|---|---|
| 0 | 0 | 1/2 | 0 | 1/2 |
| 2 | 1/4 | 0 | 1/4 | 1/2 |
| $\mathrm{P}(Y = y_j)$ | 1/4 | 1/2 | 1/4 | 1 |

(b) $\mathrm{E}(X) = 0 \cdot 0.5 + 2 \cdot 0.5 = 1$ and $\mathrm{E}(Y) = 0 \cdot 0.25 + 1 \cdot 0.5 + 2 \cdot 0.25 = 1$.

(c) No, they are not independent, e.g. $\mathrm{P}(\{X = 0\} \cap \{Y = 0\}) = 0 \neq \mathrm{P}(X = 0) \cdot \mathrm{P}(Y = 0) = \frac{1}{2} \cdot \frac{1}{4} = \frac{1}{8}$.

(d) $\mathrm{E}(X \cdot Y) = 4 \cdot \frac{1}{4} = 1$ and $\mathrm{Cov}(X, Y) = \mathrm{E}(X \cdot Y) - \mathrm{E}(X) \cdot \mathrm{E}(Y) = 1 - 1 = 0$, therefore, $X$ and $Y$ are uncorrelated.
*Remember*: Independent random variables are always uncorrelated, however, uncorrelated random variables do not necessarily need to be independent!

# Bibliography

[AntoRorr10] H. Anton and C. Rorres. *Elementary Linear Algebra: Applications Version.* John Wiley & Sons, 10th edition, 2010.

[CoolTuke65] J. W. Cooley and J. W. Tukey. An algorithm for the machine calculation of complex Fourier series. *Math. Comput.*, 19:297–301, 1965.

[CrofDaviHarg92] A. Croft, R. Davison, and M. Hargreaves. *Engineering Mathematics.* Addison–Weseley, Wokingham, 1992.

[Farl82] S. J. Farlow. *Partial Differential Equations for Scientist and Engineers.* Dover, New York, 1982.

[Habe04] R. Haberman. *Applied Partial Differential Equations: With Fourier Series and Boundary Value Problems.* Pearson Prentice Hall, 2004.

[Jame02] J. James. *A Student's Guide to Fourier Transforms: With Applications in Physics and Engineering.* Cambridge University Press, 2002.

[KolmHill01] B. Kolman and D. Hill. *Introductory Linear Algebra with Applications.* Prentice Hall, 7th edition, 2001.

[LandHest92] E. M. Landesman and M. R. Hestenes. *Linear Algebra for Mathematics, Science and Engineering.* Prentice Hall, 1992.

[MontRung03] D. Montgomery and G. Runger. *Applied Statistics and Probability for Engineers.* John Wiley & Sons, third edition, 2003.

[ONei11] P. O'Neil. *Advanced Engineering Mathematics.* Cengage Learning, seventh edition, 2011.

[OttoPete92] N. S. Ottosen and H. Petersson. *Introduction to the Finite Element Method.* Prentice Hall, 1992.

[Redd84] J. N. Reddy. *An Introduction to the Finite Element Analysis.* McGraw–Hill, 1984.

[Stah08] A. Stahel. Numerical Methods. lecture notes, BFH-TI, 2008.

[Stew87] E. G. Steward. *Fourier Optics: An Introduction.* Ellis Horwood Limited, second edition, 1987. republished by Dover.

[Stew13] I. Stewart. *Seventeen Equations that Changed the World.* Profile Books Limited, 2013.

[Swok92] E. W. Swokowski. *Calculus, late Trigonometry Version.* PWS–Kent Publishing Company, Boston, fifth edition, 1992.

[TikhSama63] A. N. Tikhonov and A. A. Samarskii. *Equations of Mathematical Physics.* Pergamon Press, 1963. republished by Dover.

# List of Figures

# List of Tables

# Index