Mathematics 2

Hamburg University of Applied Sciences,

Department of Information and Electrical Engineering

Robert Heß

Summer term 2022

August 15, 2022

Contents

1. Integral calculus **5**

1.1. Introduction 5
1.2. Definite integral 6
1.3. Properties of definite integrals 9
1.4. Fundamental theorem of calculus 9
1.5. Indefinite integral 10
1.6. Elementary integrals 11
1.7. Problems 11

2. Application of integrals **13**

2.1. Area of trapezium on abscissa 13
2.2. Area of triangles 14
2.3. Area of polygons 14
2.4. Area of discs 14
2.5. Volume of cones 16
2.6. Volume of spheres 16
2.7. Surface area of spheres 16
2.8. Mean values 17
2.8.1. Arithmetic mean 17
2.8.2. Root mean square 17
2.9. Problems 18

3. Integration techniques **20**

3.1. Improper integrals 20
3.1.1. Infinite domain 20
3.1.2. Integrand with infinite image 20
3.1.3. Summary 21
3.2. Integration by substitution 21
3.2.1. Integration of $f(ax+b)$ 22
3.2.2. Integration of $\frac{\varphi'(x)}{\varphi(x)}$ 23
3.2.3. Integration of $f[\varphi(x)]\varphi'(x)$ 24
3.2.4. Summary 24
3.3. Integration by parts 24
3.4. Integration of absolutes 25
3.5. Integration of rational functions 26
3.5.1. Single real poles 26
3.5.2. Multiple real poles 27
3.5.3. Single complex conjugate poles 27
3.5.4. Summary 28
3.6. Problems 28

4. Functions with multiple arguments and values **30**

4.1. Introduction 30
4.2. Definition of multiple argument functions 30
4.3. Visualization of functions with multiple arguments 30
4.3.1. Parametric plot 30

4.3.2. Surface plot . . . 31
4.3.3. Contour plot . . . 31
4.3.4. Vector plot . . . 32
4.4. Continuity . . . 33
4.5. Partial derivative . . . 33
4.6. Gradient . . . 35
4.7. Multiple partial derivative . . . 35
4.8. Extrema . . . 36
4.9. Multiple value functions . . . 37
4.10. Problems . . . 38

5. Differential equations **40**
5.1. Introduction . . . 40
5.2. Definition of differential equations, DE . . . 40
5.3. Solutions of DEs . . . 41
5.4. Creating DEs . . . 43
5.5. Problems . . . 44

6. First order differential equations **46**
6.1. Introduction . . . 46
6.2. Geometric interpretation . . . 46
6.3. Separation of variables . . . 46
6.4. Variation of parameters . . . 47
6.5. Linear inhomogeneous DE . . . 48
6.6. Summary . . . 49
6.7. Problems . . . 49

7. Higher order linear differential equations **51**
7.1. Introduction . . . 51
7.2. Second order homogeneous DE with constant coefficients . . . 52
7.2.1. Two different real constants . . . 52
7.2.2. Pair of complex conjugate constants . . . 52
7.2.3. Two equal real constants . . . 52
7.2.4. Summary . . . 54
7.3. Higher order homogeneous DE with constant coefficients . . . 54
7.4. Higher order inhomogeneous DE with constant coefficients . . . 55
7.4.1. Polynomial source term . . . 55
7.4.2. Exponential source term . . . 56
7.4.3. Trigonometric source term . . . 57
7.4.4. Source term as a sum of functions . . . 57
7.4.5. DE in resonance . . . 58
7.5. Problems . . . 59

8. Combinatorics **60**
8.1. Permutations . . . 60
8.2. Combinations . . . 60
8.3. Variations . . . 62
8.4. Summary . . . 62
8.5. Problems . . . 62

9. Probability theory **64**
9.1. Introduction . . . 64
9.2. Sample, sample space and event . . . 65

9.3. Probability . . . 66
9.3.1. Frequentist probability . . . 66
9.3.2. Probability axioms . . . 67
9.3.3. Calculating with probabilities . . . 67
9.3.4. Conditional probability . . . 67
9.3.5. Independence . . . 68
9.4. Problems . . . 69

10. Stochastic **71**
10.1. Random variable . . . 71
10.1.1. Definition of random variables . . . 71
10.1.2. Characteristics of random variables . . . 72
10.1.3. Modified random variable . . . 74
10.1.4. Sum of random variables . . . 74
10.2. Discrete random distributions . . . 75
10.2.1. Binomial distribution . . . 75
10.2.2. Poisson distribution . . . 76
10.3. Continuous random distributions . . . 77
10.3.1. Uniform distribution . . . 77
10.3.2. Normal distribution . . . 78
10.3.3. Exponential distribution . . . 80
10.4. Problems . . . 81

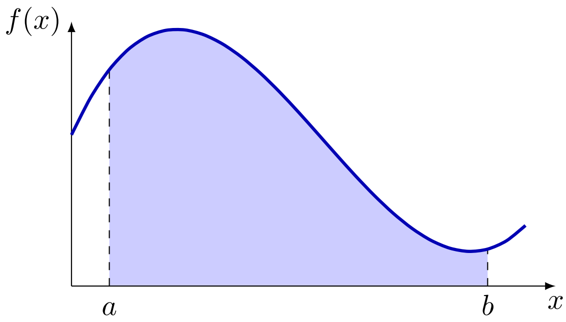
A. Solutions **83**
A.1. Integral calculus . . . 83
A.2. Application of integrals . . . 83
A.3. Integration techniques . . . 84
A.4. Functions with multiple arguments and values . . . 85
A.5. Differential equations . . . 88
A.6. First order differential equations . . . 88
A.7. Higher order linear differential equations . . . 89
A.8. Combinatorics . . . 90
A.9. Probability theory . . . 90
A.10. Stochastic . . . 91

Index **92**

1. Integral calculus

1.1. Introduction

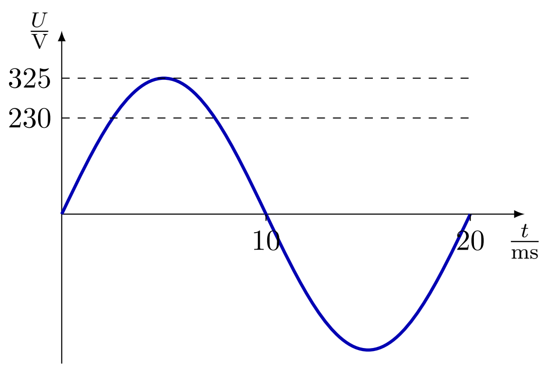
Integral calculus was first motivated by the evaluation of an area underneath a curve of a function. For simple functions like a constant function or a function with a constant slope the evaluation of the surface is obvious. However, for arbitrary functions it becomes difficult.



A way to estimate the area underneath the function is to fill it with a large number of small well defined rectangles or squares of known size. E.g. we may start with squares of $1\,\mathrm{m}^2$ size. If we can't place any more of these $1\,\mathrm{m}^2$ squares we continue with $1\,\mathrm{dm}^2$ squares. We continue in the same manner until we find a sufficient precise solution for the given purpose.

The field of integral calculus is not limited to areas underneath functions curves.

Example 1.1. E.g. the mains voltage in households is said to be 230 V. If you connect an oscilloscope to it (be careful!) you will find a sine-shaped signal oscillating with frequency $f = 50\,\mathrm{Hz}$ and a peak-voltage of $U_p = 325\,\mathrm{V}$. But where do we find the stated of 230 V?



The 230 V is the DC-voltage having the same *effect* to an absorber obeying Ohm's law. The absorbed power in such an absorber is

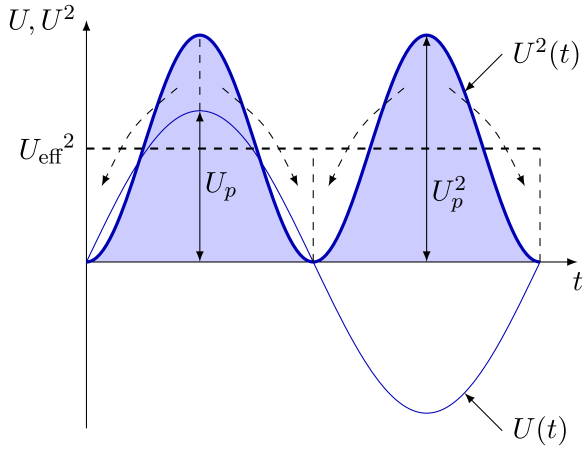
$$P(t) = U(t) \cdot I(t) = U(t) \cdot \frac{U(t)}{R} = \frac{U^2(t)}{R}$$

where t is the time, P the absorbed power, U the voltage, I the current through the absorber and R the resistance of the absorber.

To find this *effective voltage* U_{eff} we need to find the area underneath the function

$$U^2(t) = {U_p}^2 \sin^2(2\pi f\, t) = \frac{{U_p}^2}{2}(1 - \cos(4\pi f\, t))$$

with U_p as the peak-voltage of 325 V and f as the frequency of 50 Hz.



What is the area underneath this function? We are lucky and see that the area above the dashed line fit into the area underneath the dashed line not covered by the function. Hence we get:

$$\begin{aligned} {U_{\text{eff}}}^2 &= \frac{{U_p}^2}{2} \\ U_{\text{eff}} &= \frac{U_p}{\sqrt{2}} = \frac{325\,\mathrm{V}}{\sqrt{2}} \approx 230\,\mathrm{V} \end{aligned}$$

A general approach is to evaluate the integral over a period T of the frequency $f = \frac{1}{T}$:

$$U_{\text{eff}}^2 = \frac{1}{T} \int_0^T U^2(t)\,\mathrm{d}t$$

For a sine wave we get:

$$\begin{aligned}
U_{\text{eff}}^2 &= \frac{1}{T}\int_0^T U_p^2 \sin^2(2\pi f t)\,\mathrm{d}t \\
&= \frac{U_p^2}{T}\int_0^T \frac{1}{2} - \frac{1}{2}\cos(4\pi f t)\,\mathrm{d}t \\
&= \frac{U_p^2}{T}\int_0^T \frac{1}{2}\,\mathrm{d}t - \frac{U_p^2}{T}\int_0^T \frac{1}{2}\cos(4\pi f t)\,\mathrm{d}t \\
&= \frac{U_p^2}{T}\cdot\frac{T}{2} - \frac{U_p^2}{T}\cdot 0 = \frac{U_p^2}{2} \\
U_{\text{eff}} &= \frac{U_p}{\sqrt{2}}
\end{aligned}$$

◁

There are many other applications in science and engineering where we need integration as a standard tool. This and the subsequent chapters give an introduction into integral calculus.

1.2. Definite integral

We approach integral calculus by replacing the investigated function by a series of rectangles. The word *definite* in this context means that we focus on a given interval on the abscissa, i.e. a given start point a and a given end point b.

We define three types of rectangles to replace the area underneath a function:

Definition 1.1 (Riemann sum)**.** Let the domain of the function $f : [a,b] \to \mathbb{R}$ be subdivided in n equal spaced steps $\Delta x = \frac{b-a}{n}$ such that: $x_k = a + k\Delta x$ for $k \in \{0, 1, \ldots, n\}$. We define

$$R_f(n) = \sum_{k=0}^{n-1} \Delta x f(x_k)$$

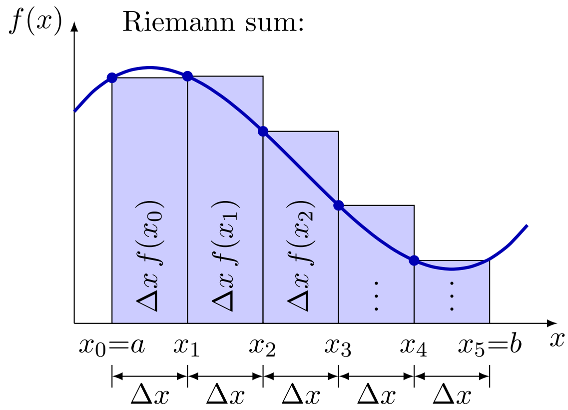
as the *Riemann sum*,

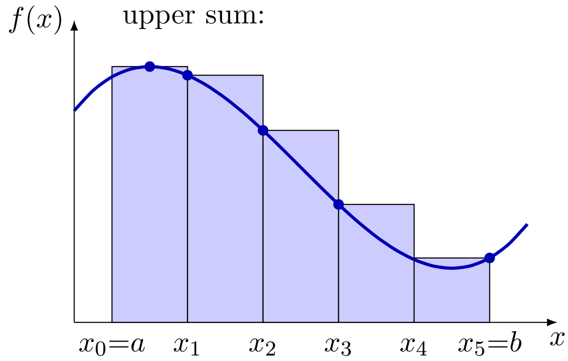
$$U_f(n) = \sum_{k=0}^{n-1} \Delta x \sup_{[x_k, x_{k+1}]}(f(x))$$

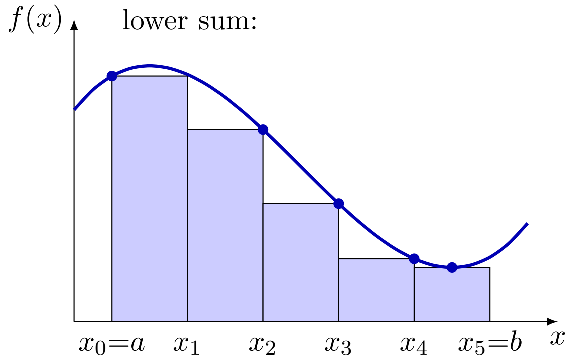
as the *upper sum* and

$$L_f(n) = \sum_{k=0}^{n-1} \Delta x \inf_{[x_k, x_{k+1}]}(f(x))$$

as the *lower sum*. ◁







The subdivisions do not necessary have to have equal size. However, for the sake of easy argument we stick to equal sized subdivisions.

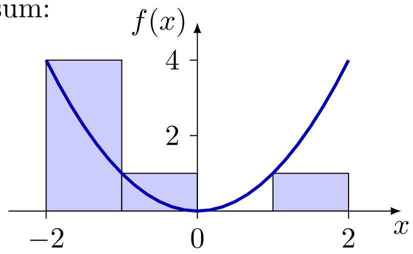
For the Riemann sum we defined the top left corner of each rectangle to meet the function value. However, we may have taken any point at the upper edge to meet the function value.

For rectangles below the abscissa we take negative values. I.e. we may get zero sum if the areas above and below the abscissa "compensate".

Example 1.2. What is the Riemann-, upper- and lower sum for the function $f : \mathbb{R} \to \mathbb{R}, x \mapsto x^2$ on the interval $[-2, 2]$ for four equal sized steps?

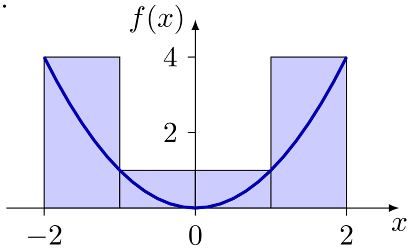
$$\begin{aligned}
\Delta x &= \frac{b-a}{n} = \frac{2-(-2)}{4} = 1 \\
x_k &= a + k\Delta x = -2 + 1\cdot k = k - 2
\end{aligned}$$

Riemann sum:



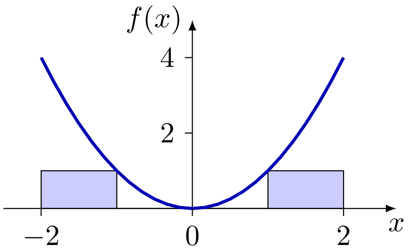
$$R_f(4) = \sum_{k=0}^{n-1} \Delta x f(x_k) = \sum_{k=0}^{3} 1 \cdot (k-2)^2$$
$$= 4 + 1 + 0 + 1 = 6$$

upper sum:



$$U_f(4) = \sum_{k=0}^{n-1} \Delta x \sup_{[x_k, x_{k+1}]} f(x)$$
$$= 4 + 1 + 1 + 4 = 10$$

lower sum:



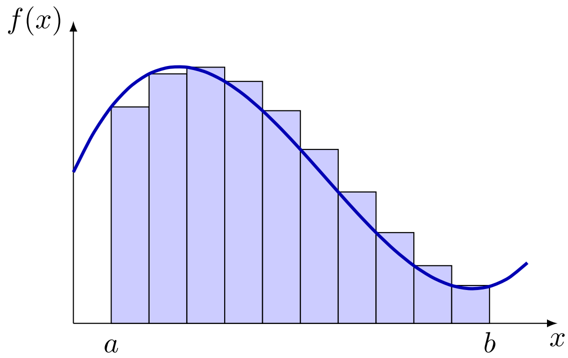
$$L_f(4) = \sum_{k=0}^{n-1} \Delta x \inf_{[x_k, x_{k+1}]} f(x)$$
$$= 1 + 0 + 0 + 1 = 2$$

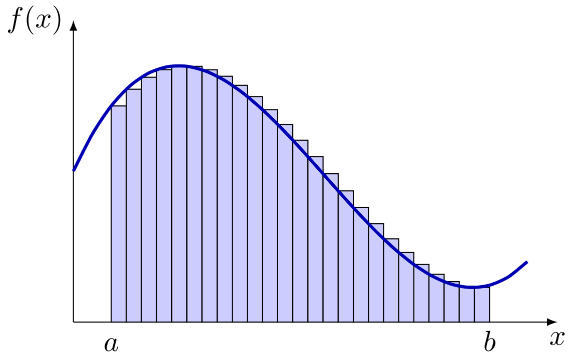
◁

Obviously the lower sum is less or equal to the Riemann sum, which again is less or equal to the upper sum:

$$L_f(n) \le R_f(n) \le U_f(n)$$

If we increase the number of steps n to divide the interval $[a, b]$ the upper sum decreases and the lower sum increases. The larger n becomes the closer the three sums converge towards the area underneath the curve.





Definition 1.2 (Convergent Riemann sum)**.** Let $f : [a, b] \to \mathbb{R}$ with the domain equally subdivided into n intervals of length $\Delta x = \frac{b-a}{n}$. If for increasing n the difference of the upper and lower sum is convergent towards zero we say the sequence of Riemann sums $(R_f(n))$ is convergent. I.e.

$$\lim_{n\to\infty} \{U_f(n) - L_f(n)\} = 0$$

If $(R_f(n))$ is convergent, we say f is an *integrable function*. ◁

Remark: There are integrable functions where the sequence of Riemann sums is not convergent. However, for our purposes the Riemann integral is a sufficient explanation. The interested student may look for *Lebesgue integration* to find more details.

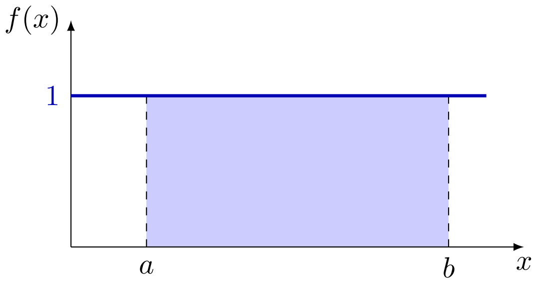
Definition 1.3 (Definite integral)**.** For a convergent sequence of Riemann sums $(R_f(n))$ we define the *definite integral* (or *Riemann integral*) with:

$$\int_a^b f(x)\mathrm{d}x = \lim_{n\to\infty} R_f(n)$$

We call the function $f(x)$ the *integrand*, the interval $[a, b]$ the *domain of integration* and the limits a and b the *limits of integration* or the *lower limit* and *upper limit*, respectively. ◁

Now we want to evaluate the definite integral of some integrands by applying the Riemann sum:

Example 1.3. What is the definite integral of a constant function $f(x) = 1$ over the interval $[a, b]$?

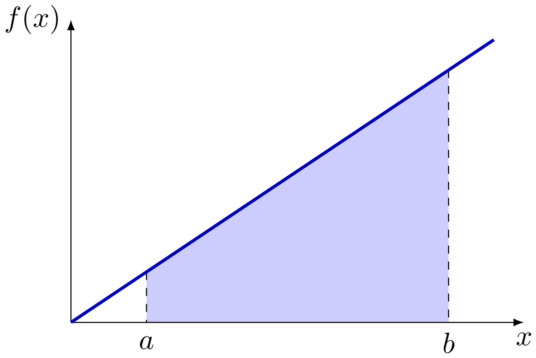


Geometrically we see that the surface underneath the curve equals $b-a$. However, we apply the Riemann integral:

$$\begin{aligned}\int_a^b f(x)\mathrm{d}x &= \lim_{n\to\infty} R_f(n) = \lim_{n\to\infty} \sum_{k=0}^{n-1} \Delta x f(x_k) \\ &= \lim_{n\to\infty} \sum_{k=0}^{n-1} \tfrac{b-a}{n} f(a + k\tfrac{b-a}{n}) \\ &= \lim_{n\to\infty} \sum_{k=0}^{n-1} \tfrac{b-a}{n} = \lim_{n\to\infty} \tfrac{b-a}{n} \sum_{k=0}^{n-1} 1 \\ &= \lim_{n\to\infty} \tfrac{b-a}{n} n = \lim_{n\to\infty} (b-a) \\ &= b-a\end{aligned}$$

I.e. $\int_a^b 1\,\mathrm{d}x = b - a$ ◁

Example 1.4. What is the definite integral of $f(x) = x$ over the interval $[a, b]$?

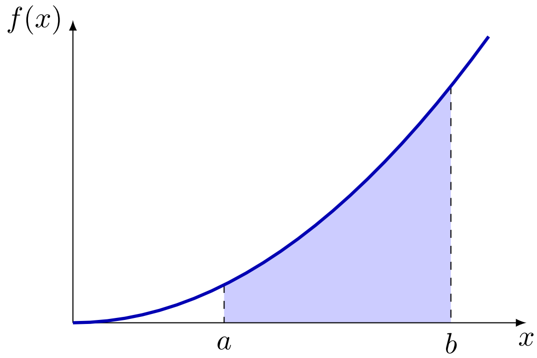


We again apply the Riemann integral:

$$\begin{aligned}R_f(n) &= \sum_{k=0}^{n-1} \Delta x f(x_k) = \sum_{k=0}^{n-1} \Delta x\, x_k \\ &= \sum_{k=0}^{n-1} \Delta x(a + k\Delta x) \\ &= \Delta x\, a \sum_{k=0}^{n-1} 1 + \Delta x^2 \sum_{k=0}^{n-1} k \\ &= \Delta x\, a\, n + \Delta x^2 \frac{n(n-1)}{2} \\ &= \frac{b-a}{n} a\, n + \left(\frac{b-a}{n}\right)^2 \frac{n(n-1)}{2} \\ &= (b-a)a + \frac{(b-a)^2}{2} \cdot \underbrace{\frac{n^2-n}{n^2}}_{\underset{n\to\infty}{\longrightarrow} 1} \\ &\underset{n\to\infty}{\longrightarrow} \frac{2ba - 2a^2}{2} + \frac{b^2 - 2ba + a^2}{2} \\ &= \frac{b^2 - a^2}{2}\end{aligned}$$

I.e. $\int_a^b x\,\mathrm{d}x = \frac{b^2 - a^2}{2}$ ◁

Example 1.5. What is the definite integral of $f(x) = x^2$ over the interval $[a, b]$?



We again apply the Riemann integral:

$$\begin{aligned}R_f(n) &= \sum_{k=0}^{n-1} \Delta x f(x_k) = \sum_{k=0}^{n-1} \Delta x\, {x_k}^2 \\ &= \sum_{k=0}^{n-1} \Delta x(a + k\Delta x)^2 \\ &= \Delta x \sum_{k=0}^{n-1} (a^2 + 2ak\Delta x + k^2\Delta x^2) \\ &= \Delta x a^2 n + 2a\Delta x^2 \sum_{k=0}^{n-1} k + \Delta x^3 \sum_{k=0}^{n-1} k^2\end{aligned}$$

after some conversion:

$$\begin{aligned}&= (b-a)^2 + (b-a)^2 a - \tfrac{(b-a)^2 a}{n} \\ &\quad + \tfrac{(b-a)^3}{3} - \tfrac{(b-a)^3}{2n} + \tfrac{(b-a)^3}{6n^2} \\ &\underset{n\to\infty}{\longrightarrow} (b-a)^2 + (b-a)^2 a + \tfrac{(b-a)^3}{3}\end{aligned}$$

$$= \frac{b^3 - a^3}{3}$$

I.e. $\int_a^b x^2 \, \mathrm{d}x = \frac{b^3 - a^3}{3}$ ◁

Could it be, that we found some sort of pattern

$$\int_a^b x^n \, \mathrm{d}x = \frac{b^{n+1} - a^{n+1}}{n+1} \quad ?$$

We will see later that it is true.

Theorem 1.4 (Continuous functions are integrable)**.** For continuous functions $f : [a, b] \to \mathbb{R}$ the sequence of Riemann sums $(R_f(n))$ is convergent, i.e. continuous functions are integrable. ◁

Remark: However, not all integrable functions are continuous. We look at this later.

1.3. Properties of definite integrals

We now derive some properties of definite integrals:

Theorem 1.5 (Properties of definite integrals)**.** Let $I \subset \mathbb{R}$ be a closed interval, $a, b, c \in I$ with $a < b < c$, $\lambda \in \mathbb{R}$ and $f, g : I \to \mathbb{R}$ be integrable functions. We then have

- $\int_a^b f(x)\mathrm{d}x + \int_b^c f(x)\mathrm{d}x = \int_a^c f(x)\mathrm{d}x$
- $\int_a^b (f(x) + g(x))\mathrm{d}x$
  $$= \int_a^b f(x)\mathrm{d}x + \int_a^b g(x)\mathrm{d}x$$
- $\int_a^b \lambda f(x)\mathrm{d}x = \lambda \int_a^b f(x)\mathrm{d}x$
- $\left| \int_a^b f(x)\mathrm{d}x \right| \leq \int_a^b |f(x)| \, \mathrm{d}x$
- $\int_a^b f(x)\mathrm{d}x = - \int_b^a f(x)\mathrm{d}x$

◁

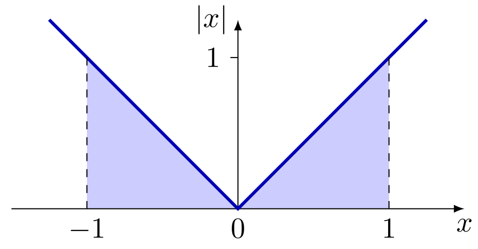
Example 1.6. Integrate the function f over the interval $[-1, 1]$: $\quad f : \mathbb{R} \to \mathbb{R}, x \mapsto |x|$

$$\int_{-1}^1 |x| \, \mathrm{d}x = \int_{-1}^0 -x \, \mathrm{d}x + \int_0^1 x \, \mathrm{d}x$$

$$= - \int_{-1}^0 x \, \mathrm{d}x + \int_0^1 x \, \mathrm{d}x$$

$$= - \frac{0^2 - (-1)^2}{2} + \frac{1^2 - 0^2}{2}$$

$$= \frac{1}{2} + \frac{1}{2} = 1$$

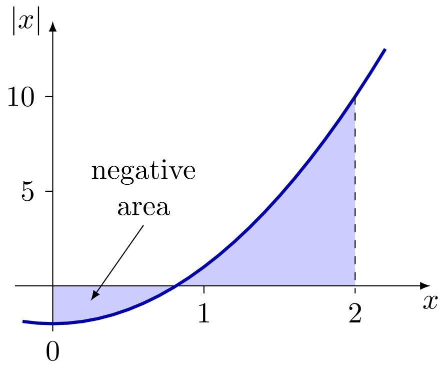


◁

Example 1.7. Integrate the function f over the interval $[0, 2]$: $\quad f : \mathbb{R} \to \mathbb{R}, x \mapsto 3x^2 - 2$

$$\int_0^2 3x^2 - 2 \, \mathrm{d}x = 3 \int_0^2 x^2 \, \mathrm{d}x - 2 \int_0^2 \mathrm{d}x$$

$$= 3 \frac{2^3 - 0^3}{3} - 2(2 - 0) = 4$$



◁

1.4. Fundamental theorem of calculus

We now take the upper limit as a variable x and replace the lower limit a by x_0. Hence, the integral becomes a function of x:

$$F(x) = \int_{x_0}^x f(x')\mathrm{d}x'$$

If we rewrite the results of the previous three examples we get:

ex. 1.3: $\quad \int_{x_0}^x 1 \, \mathrm{d}x' = x - x_0$

ex. 1.4: $\displaystyle\int_{x_0}^{x} x'\,\mathrm{d}x' = \frac{x^2 - {x_0}^2}{2} = \frac{x^2}{2} - \frac{{x_0}^2}{2}$

ex. 1.5: $\displaystyle\int_{x_0}^{x} x'^2\,\mathrm{d}x' = \frac{x^3 - {x_0}^3}{3} = \frac{x^3}{3} - \frac{{x_0}^3}{3}$

An interesting fact is, that the first derivatives of the results bring us back to the functions we integrated in the first place. Please note that this is true for any lower limit x_0.

example 1.3: $\displaystyle\frac{\mathrm{d}(x - x_0)}{\mathrm{d}x} = 1$

example 1.4: $\displaystyle\frac{\mathrm{d}\frac{x^2 - {x_0}^2}{2}}{\mathrm{d}x} = x$

example 1.5: $\displaystyle\frac{\mathrm{d}\frac{x^3 - {x_0}^3}{3}}{\mathrm{d}x} = x^2$

This leads us to the first part of the fundamental theorem of calculus:

Theorem 1.6 (First fundamental theorem of calculus)**.** Let $f : [a, b] \to \mathbb{R}$ be a continuous function. If

$$F(x) = \int_a^x f(x')\,\mathrm{d}x'$$

then we have for any $x \in (a, b)$

$$\frac{\mathrm{d}}{\mathrm{d}x} F(x) = f(x)$$

◁

The fundamental theorem of calculus states, that integration is somehow the inversion of differentiation. The only difference is at some discontinuities where the integral is defined but the derivative of the integral does not lead us back to the original function.

Example 1.8. For the function

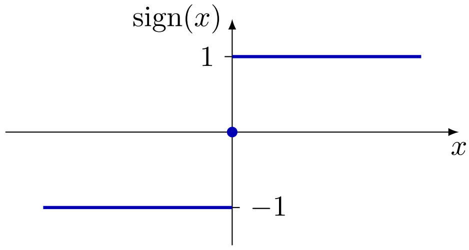
$$\mathrm{sign}(x) : \begin{cases} \mathbb{R} \to \mathbb{R} & \\ x \mapsto 1 & \text{for } x > 0 \\ x \mapsto 0 & \text{for } x = 0 \\ x \mapsto -1 & \text{for } x < 0 \end{cases}$$

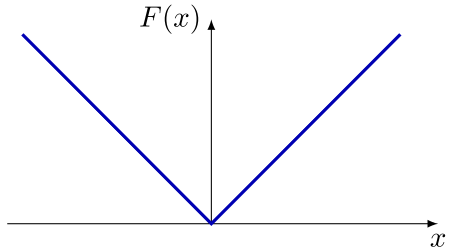
the integral $F(x)$ with lower limit zero is

$$F(x) = \int_0^x \mathrm{sign}(x')\,\mathrm{d}x' = |x|$$

The derivative of F is not defined for at $x = 0$ since is it not possible to evaluate the slope at $x = 0$.

$$\frac{\mathrm{d}F(x)}{\mathrm{d}x} = \mathrm{sign}(x) \quad \text{for } x \neq 0$$





◁

Hence, the first part of the fundamental theorem of calculus is limited to continuous functions.

The second part of the fundamental theorem of calculus avoids differentiation and, hence, can be extended to any integrable function:

Theorem 1.7 (Second fundamental theorem of calculus)**.** Let $f : [a, b] \to \mathbb{R}$ be an integrable function and $F : [a, b] \to \mathbb{R}$,

$$F(x) = \int_{x_0}^x f(x')\mathrm{d}x'$$

be an integral of f. Then we have for any $x_0 \in [a, b]$:

$$\int_a^b f(x)\mathrm{d}x = F(b) - F(a)$$

◁

1.5. Indefinite integral

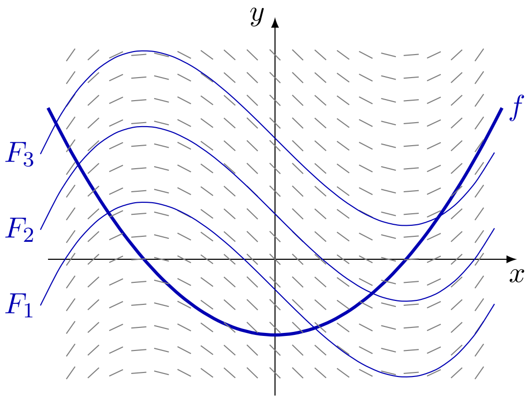
Yet we integrated functions over intervals with lower and upper limits. Now we want to find a more general way to integrate functions. To do so we first need the so called *primitive* of a function:

Definition 1.8 (Primitive function)**.** For a function $f : [a, b] \to \mathbb{R}$ any function $F : [a, b] \to \mathbb{R}$ with derivative equal to f is called a *primitive function* of f, i.e.

$$\frac{\mathrm{d}}{\mathrm{d}x} F(x) = f(x)$$

Other terms for F are *antiderivative* and *inverse derivative*. ◁

When searching for a primitive F of a given function f we remember that the derivative f at any point x gives the slope of F at this point. But since we do not know the value F at this point we may draw little lines indicating the slope at different values of F. A primitive then is a curve having the desired slope at all values of x. Unfortunately there are many primitives fulfilling these requirements:



We take the notation we used earlier:

$$F(x) = \int_{x_0}^{x} f(x')\mathrm{d}x'$$

Since the upper limit is given in the integral function we may write:

$$F(x) = \int_{x_0} f(x)\mathrm{d}x$$

A change of the lower limit x_0 results in an offset C to the primitive. I.e. when changing the lower limit, the primitive remains the same except for a vertical shift.

$$\int_{x_0} f(x)\,\mathrm{d}x = F(x) - F(x_0) = F(x) + C$$

We remove the lower limit and take the constant C as a reminder. This leads us to the indefinite integral:

$$\int f(x)\mathrm{d}x = F(x) + C$$

Remark: From mathematical point of view the constant C acts rather as a reminder than a mathematical variable. E.g. under some circumstances we add two of these constants into a new constant with the same name:

$$\begin{aligned}\int (f_1 + f_2)\mathrm{d}x &= \int f_1\,\mathrm{d}x + \int f_2\,\mathrm{d}x \\ &= (F_1 + C) + (F_2 + C) \\ &= (F_1 + F_2) + C\end{aligned}$$

1.6. Elementary integrals

Elementary integrals can be derived by inverting derivatives. Some of them are listed in the table below. The constant C has been left out and must be added to all integrals except the first.

$f(x)$	$F(x) = \int f(x)\,\mathrm{d}x$
0	C
x^a	$\frac{x^{a+1}}{a+1}$ for $a \in \mathbb{Z}, a \neq -1$ or $a \in \mathbb{R}, a \neq -1, x \geq 0$
$\sqrt{x}$	$\frac{2\sqrt{x^3}}{3}$ for $x \geq 0$
$\frac{1}{x}$	$\ln\lvert x\rvert$
e^x	e^x
a^x	$\frac{a^x}{\ln a}$ for $a \in \mathbb{R}, a > 0$
$\ln x$	$x \ln x - x$ for $x > 0$
$\cos(x)$	$\sin(x)$
$\sin(x)$	$-\cos(x)$
$\cosh(x)$	$\sinh(x)$
$\sinh(x)$	$\cosh(x)$
$\frac{1}{x^2+1}$	$\arctan(x)$
$\frac{-1}{x^2+1}$	$\operatorname{arccot}(x)$

Remark: For a definite integral we write:

$$\int_a^b f(x)\,\mathrm{d}x = \Big[F(x)\Big]_a^b = F(b) - F(a)$$

Here the constant C cancels out.

1.7. Problems

Problem 1.1: For $f : \mathbb{R} \to \mathbb{R}, x \mapsto \sin(x)$ evaluate the Riemann-, upper- and lower-sum on the interval $[0, \pi]$ in four equal sized steps.

Problem 1.2: Evaluate the Riemann-, upper- and lower sum for the function $f : \mathbb{R} \to \mathbb{R}, x \mapsto 2x$ on the interval $[0, 4]$. Divide the domain into 4 sub-intervals of equal size.

Problem 1.3: Repeat the previous problem for 1, 10, 100, 1000 and infinite number of sub-intervals.

Problem 1.4: For $x \in \mathbb{R}$ solve the following integrals by applying Riemann sums.

1. $\int_1^3 2x \, \mathrm{d}x$
2. $\int_0^2 5x \, \mathrm{d}x$
3. $\int_{-1}^4 x \, \mathrm{d}x$
4. $\int_2^3 (x-1) \, \mathrm{d}x$
5. $\int_0^2 (x^2+x) \, \mathrm{d}x$
6. $\int_0^b (2x^2-x) \, \mathrm{d}x$
7. $\int_a^b (x^2+2) \, \mathrm{d}x$
8. $\int_b^a (x^2+2) \, \mathrm{d}x$

Problem 1.5: For $x \in \mathbb{R}$ solve the following integrals. Consider the properties of definite integrals.

1. $\int_{-2}^3 |x| \, \mathrm{d}x$
2. $\int_3^{-2} |x| \, \mathrm{d}x$
3. $\int_{-2}^3 |-x| \, \mathrm{d}x$
4. $\int_{-1}^1 (|x|+x) \, \mathrm{d}x$
5. $\int_{-2}^2 (x^2-|x|) \, \mathrm{d}x$
6. $\int_1^{-1} |x| \, \mathrm{d}x$
7. $\int_a^{2b} x \, \mathrm{d}x$
8. $\int_{-a}^b x^2 \, \mathrm{d}x$

Problem 1.6: For $t \in \mathbb{R}$ solve the following integrals. Consider the fundamental theorem of calculus.

1. $\int_{x_0}^x 3\,t^2 \, \mathrm{d}t$
2. $\int_{x_0}^x \cos(t) \, \mathrm{d}t$
3. $\int_{x_0}^x \sin(t) \, \mathrm{d}t$
4. $\int_{x_0}^x e^t \, \mathrm{d}t$
5. $\int_{x_0}^x 2\pi\cos(2\pi\,t) \, \mathrm{d}t$
6. $\int_{x_0}^x 2\mathrm{j}\pi f\, e^{2\mathrm{j}\pi f\, t} \, \mathrm{d}t$
7. $\int_{x_0}^x 2t\cos\left(t^2\right) \, \mathrm{d}t$
8. $\int_{x_0}^x 3t^2\sin\left(t^3\right) \, \mathrm{d}t$

Problem 1.7: For $x \in \mathbb{R}$ find a primitive for each of the following functions:

$f_1(x) = 1$ $\qquad f_5(x) = \sin(x)$

$f_2(x) = 2x$ $\qquad f_6(x) = \cos(x)$

$f_3(x) = x^2$ $\qquad f_7(x) = e^x$

$f_4(x) = x^3$ $\qquad f_8(x) = x^2 - x + 1$

Problem 1.8: For $x \in \mathbb{R}$ solve the following integrals.

1. $\int 0 \, \mathrm{d}x$
2. $\int a \, \mathrm{d}x$
3. $\int (x^2-1) \, \mathrm{d}x$
4. $\int 3e^x \, \mathrm{d}x$
5. $\int 5^x \, \mathrm{d}x$
6. $\int \pi^x \, \mathrm{d}x$
7. $\int 2\cosh(x) \, \mathrm{d}x$
8. $\int \frac{1}{2}\sinh(x) \, \mathrm{d}x$

Problem 1.9: For $x \in \mathbb{R}$ solve the following integrals.

1. $\int_0^3 |x-1| \, \mathrm{d}x$
2. $\int_{-2}^2 |x^2-1| \, \mathrm{d}x$
3. $\int |x^2| \, \mathrm{d}x$
4. $\int_{-1}^4 e^{|x|} \, \mathrm{d}x$
5. $\int_{-\pi}^{\pi} \sin|x| \, \mathrm{d}x$
6. $\int \cos(2\pi\,x) \, \mathrm{d}x$
7. $\int e^{\mathrm{j}\,x} \, \mathrm{d}x$
8. $\int (1+\tan^2(3x)) \, \mathrm{d}x$

Problem 1.10: For $x \in \mathbb{R}_{>0}$ solve the following integrals.

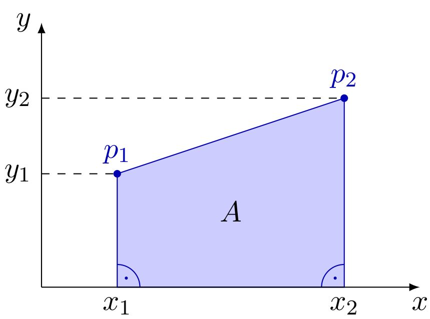
1. $\int \sqrt{x} \, \mathrm{d}x$
2. $\int \sqrt{x^3} \, \mathrm{d}x$
3. $\int \frac{1}{\sqrt{x}} \, \mathrm{d}x$
4. $\int \sqrt[3]{x} \, \mathrm{d}x$
5. $\int \frac{1}{\sqrt[3]{x^2}} \, \mathrm{d}x$
6. $\int \ln(x) \, \mathrm{d}x$
7. $\int \log_{10}(x) \, \mathrm{d}x$
8. $\int \log_2(x) \, \mathrm{d}x$

2. Application of integrals

In this chapter we want to apply integrals to some geometric applications. Some of them can be solved without integral calculus, however, we still try to solve them with integrals.

2.1. Area of trapezium on abscissa

A trapezium has four edges with two of them being parallel. We want to analyse a trapezium with one edge on the abscissa of a Cartesian coordinate system and two other edges being perpendicular on the abscissa. Hence, two corners have a right angle.



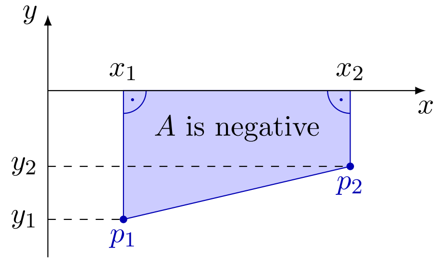
We express the line between the two points p_1 and p_2 by a linear function:

$$
\begin{aligned}
f(x) &= mx + b \\
m &= \frac{y_2 - y_1}{x_2 - x_1} \\
b &= y_1 - mx_1 = \frac{y_1 x_2 - y_2 x_1}{x_2 - x_1}
\end{aligned}
$$

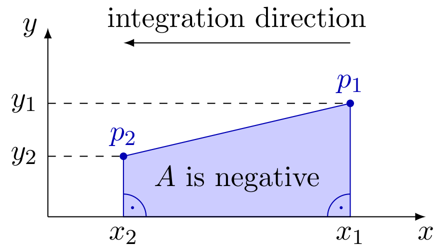
Now we evaluate the area underneath the function by taking the definite integral between x_1 and x_2:

$$
\begin{aligned}
A &= \int_{x_1}^{x_2} f(x)\,\mathrm{d}x = \int_{x_1}^{x_2} (mx + b)\,\mathrm{d}x \\
&= m \int_{x_1}^{x_2} x\,\mathrm{d}x + b \int_{x_1}^{x_2} \mathrm{d}x \\
&= m \frac{{x_2}^2 - {x_1}^2}{2} + b(x_2 - x_1) \\
&\dots \\
A &= \tfrac{1}{2}(y_2 + y_1)(x_2 - x_1)
\end{aligned}
$$

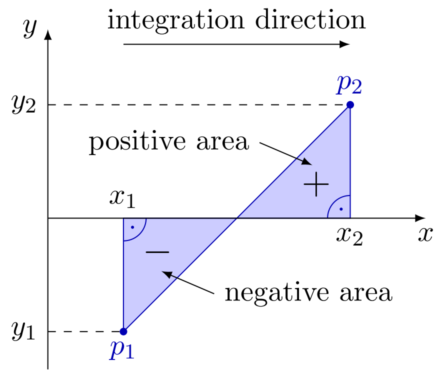
The results must be taken with care: If the points are below the abscissa, the area is negative:

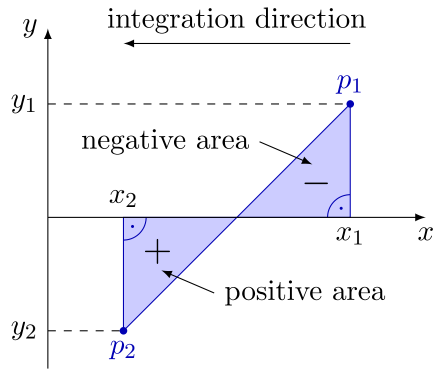


Positive points integrated from right to left, i.e. with $x_1 > x_2$ will also result in a negative area:



Finally, if one of the points is above and the other below the abscissa, parts of the area are negative and other parts are positive:



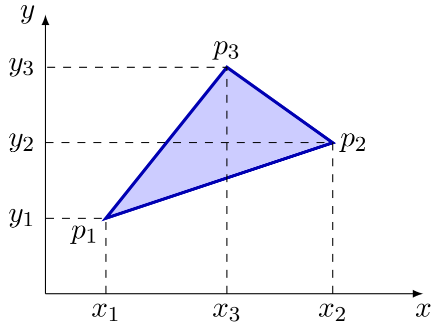


Although the linear function between p_1 and p_2 is not defined for $x_1 = x_2$, the integral remains valid and results in zero:

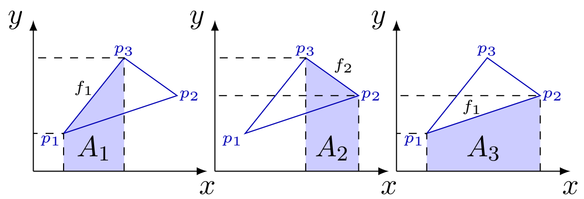
$$x_1 = x_2 \qquad \Rightarrow \qquad A = 0$$

2.2. Area of triangles

What is the area of a triangle with its corners being at arbitrary points p_1, p_2 and p_3?



For a triangle as sketched above we take the area A as a combination of the three partial areas A_1, A_2 and A_3:



$$\begin{aligned}
A &= A_1 + A_2 - A_3 \\
&= \int_{x_1}^{x_3} f_1(x)\,\mathrm{d}x + \int_{x_3}^{x_2} f_2(x)\,\mathrm{d}x \\
&\quad - \int_{x_1}^{x_2} f_3(x)\,\mathrm{d}x \\
&= \int_{x_1}^{x_3} f_1(x)\,\mathrm{d}x + \int_{x_3}^{x_2} f_2(x)\,\mathrm{d}x \\
&\quad + \int_{x_2}^{x_1} f_3(x)\,\mathrm{d}x \\
&= \tfrac{1}{2}\{(y_3 + y_1)(x_3 - x_1) + (y_2 + y_3)(x_2 - x_3) \\
&\quad + (y_1 + y_2)(x_1 - x_2)\} \\
&= \frac{y_1(x_3 - x_2) + y_2(x_1 - x_3) + y_3(x_2 - x_1)}{2}
\end{aligned}$$

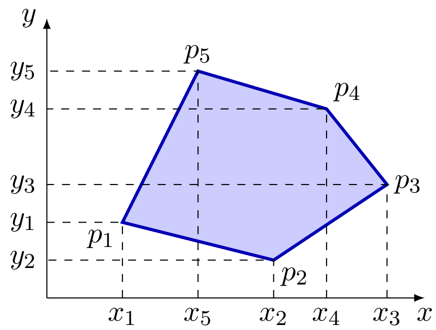
The area is positive if the points are arranged counter clockwise, and negative if the points are arranged clockwise.

We come to the same result, if we write the three points into a 3×3-Matrix and evaluate $\frac{1}{2}$ of the determinant:

$$\begin{aligned}
A &= \frac{1}{2} \det \begin{pmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \end{pmatrix} \\
&= \frac{1}{2} \det \begin{pmatrix} 1 & 1 & 1 \\ x_1 & x_2 & x_3 \\ y_1 & y_2 & y_3 \end{pmatrix}
\end{aligned}$$

2.3. Area of polygons

We may extend the previous result to polygons with an arbitrary number of corners.



We have to add the integrals of all linear functions connecting the points:

$$\begin{aligned}
A &= \int_{x_1}^{x_n} f_n(x)\,\mathrm{d}x + \sum_{k=1}^{n-1} \int_{x_{k+1}}^{x_k} f_k(x)\,\mathrm{d}x \\
&= \frac{1}{2} \Bigg\{ y_1(x_n - x_2) + \sum_{k=2}^{n-1} y_k(x_{k-1} - x_{k+1}) \\
&\qquad + y_n(x_{n-1} - x_1) \Bigg\}
\end{aligned}$$

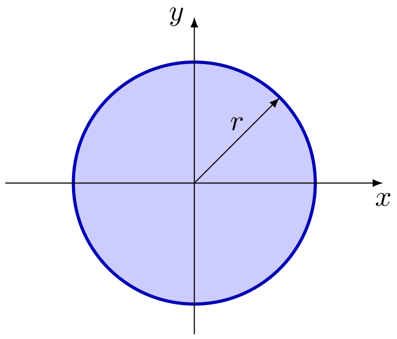
Again, with points arranged counter clockwise we get a positive area and a negative area for points arranged clockwise. If some of the lines connecting the points cross each other, we get positive and negative area fractions.

2.4. Area of discs

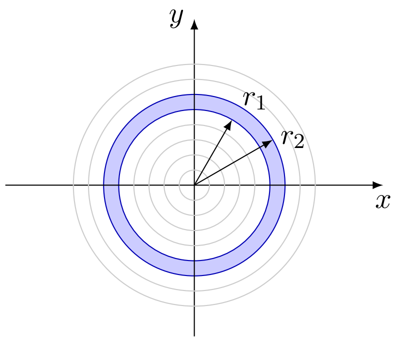
From textbooks we know the area A of a disc with radius r to be:

$$A = \pi r^2$$

where π is the ratio of the circumference of a disc to its diameter. But where does this equation come from? We want to use infinite sums and integral calculus to derive this equation.



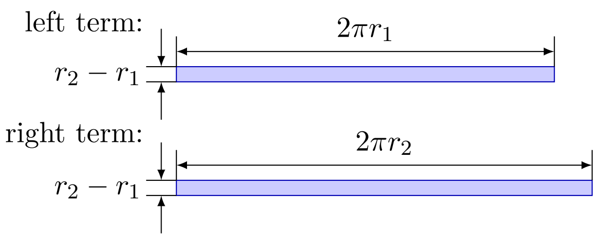
We split the disc into n rings:



The surface of a ring A_r with radii r_1 and r_2 is limited by:

$$2\pi r_1(r_2 - r_1) \leq A_r \leq 2\pi r_2(r_2 - r_1)$$

The left term is the area of an rectangle of length $2\pi r_1$ and width $r_2 - r_1$. The right term is the area of an rectangle of length $2\pi r_2$ and width $r_2 - r_1$.



The areas of the two rectangles act as a lower and upper limit for the true area of the ring.

Adding the area of all rings based on the inner radius leads us to the total area A_L which is not larger than the true area of the disk A_{disc}.

Adding the area of all rings based on the outer radius leads us to the total area A_U which is not less than the true area of the disk A_{disc}.

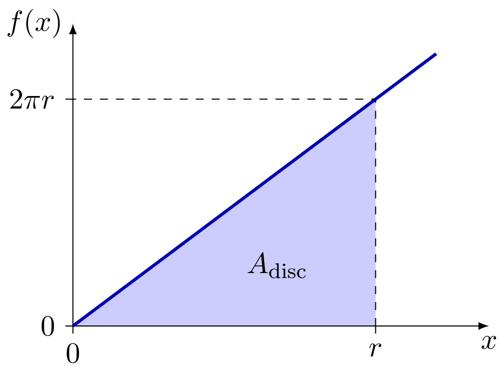
With $\Delta r = \frac{r}{n}$ and $r_k = k\Delta r$ we get:

$$\begin{aligned}
A_L(n) &= \sum_{k=0}^{n-1} 2\pi r_k \Delta r = 2\pi \sum_{k=0}^{n-1} k\Delta r \Delta r \\
&= 2\pi \Delta r^2 \sum_{k=0}^{n-1} k = 2\pi \frac{r^2}{n^2} \cdot \frac{n(n-1)}{2} \\
&= \pi r^2 \frac{n-1}{n} \underset{n\to\infty}{\longrightarrow} \pi r^2 \\
A_U(n) &= \sum_{k=1}^{n} 2\pi r_k \Delta r = 2\pi \sum_{k=1}^{n} k\Delta r \Delta r \\
&= 2\pi \Delta r^2 \sum_{k=1}^{n} k = 2\pi \frac{r^2}{n^2} \cdot \frac{n(n+1)}{2} \\
&= \pi r^2 \frac{n+1}{n} \underset{n\to\infty}{\longrightarrow} \pi r^2
\end{aligned}$$

Since the true surface of the disc A_{disk} is not less than A_L and not larger than A_U we were able to prove:

$$A_{\text{disk}} = \pi r^2$$

Another approach is to plot the circumference as a function the radius:



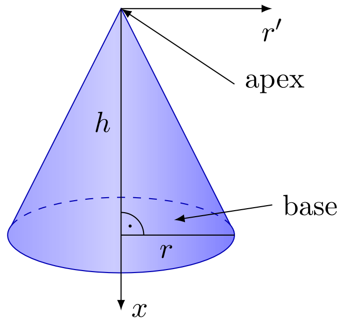
The area under the function equals the area of the disc. Hence, we take the integral from zero to the radius of the disc:

$$\begin{aligned}
A_{\text{disc}} &= \int_0^r 2\pi x \, \mathrm{d}x = 2\pi \int_0^r x \, \mathrm{d}x = 2\pi \frac{r^2 - 0^2}{2} \\
&= \pi r^2
\end{aligned}$$

Again, we find the equation for the area of a disc to be $A_{\text{disc}} = \pi r^2$. The application of integral calculus led us more rapid to the equation under question.

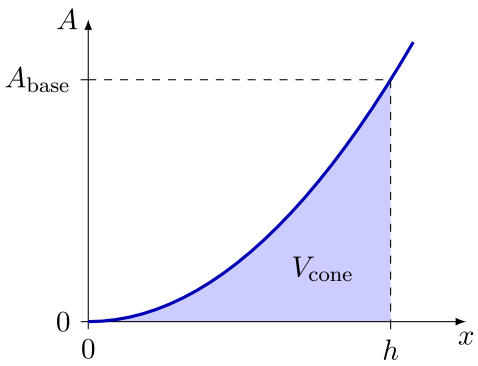
2.5. Volume of cones

We assume a right circular cone with height h and radius r, i.e. a cone with a disc of radius r as base and the apex being with distance h perpendicular above the centre of the base.



We may imagine the cone as a pile of thin discs with decreasing areas towards the apex of the cone. Plotting the area of the discs as a function of distance to the apex results in the parabola where the radius r' of the discs increase with distance to the apex:

$$f(x) = \pi r'^2 = \pi \left(\frac{rx}{h}\right)^2 = \frac{\pi r^2}{h^2} x^2$$



We now integrate from the apex to the base:

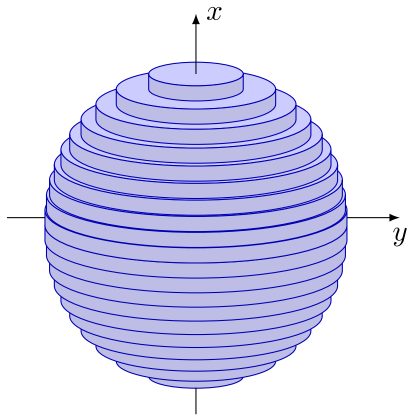
$$\begin{aligned} V_{\text{cone}} &= \int_0^h \frac{\pi r^2}{h^2} x^2 \, \mathrm{d}x = \frac{\pi r^2}{h^2} \int_0^h x^2 \, \mathrm{d}x \\ &= \frac{\pi r^2}{h^2} \cdot \frac{h^3 - 0^3}{3} = \frac{1}{3} \pi r^2 h \end{aligned}$$

This is the equation for the volume of a right circular cone with height h and base radius r. It is one third of the volume of a cylinder with same height and radius.

2.6. Volume of spheres

We imagine a sphere with radius r as a pile of discs. Position and radius of the discs are given by x and y, respectively. Due to the Pythagorean theorem we get

$$y^2 = r^2 - x^2$$



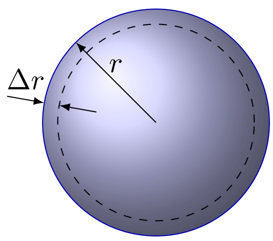
We integrate over all discs from $-r$ to r:

$$\begin{aligned} V_{\text{sphere}} &= \int_{-r}^{r} \pi y^2 \, \mathrm{d}x = \int_{-r}^{r} \pi (r^2 - x^2) \, \mathrm{d}x \\ &= \int_{-r}^{r} \pi r^2 \, \mathrm{d}x - \int_{-r}^{r} \pi x^2 \, \mathrm{d}x \\ &= \pi r^2 \int_{-r}^{r} \mathrm{d}x - \pi \int_{-r}^{r} x^2 \, \mathrm{d}x \\ &= \pi r^2 (r - (-r)) - \pi \frac{r^3 - (-r)^3}{3} \\ &= 2\pi r^3 - \tfrac{2}{3} \pi r^3 = \tfrac{4}{3} \pi r^3 \end{aligned}$$

2.7. Surface area of spheres

Let us first evaluate the volume of the mantle of a sphere (like the skin of a football). The volume of the mantle is the volume of the sphere surrounding the mantle minus the the volume of the sphere inside the mantle. With r as the outer radius of the mantle and Δr as the thickness of the mantle we get:

$$V_{\text{mantle}} = \frac{4\pi}{3} r^3 - \frac{4\pi}{3} (r - \Delta r)^3$$



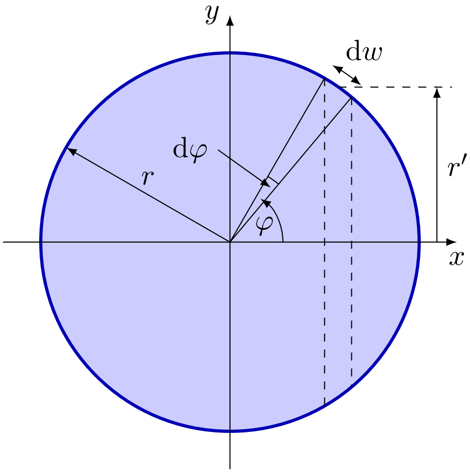
For a thin mantle we may approximate the volume of the mantle by the outer surface area of the sphere multiplied by the thickness of the mantle:

$$V_{\text{mantle}} \approx A_{\text{sphere}} \Delta r \quad \Rightarrow \quad A_{\text{sphere}} \approx \frac{V_{\text{mantle}}}{\Delta r}$$

When reducing Δr towards zero we get a precise relationship:

$$\begin{aligned} A_{\text{sphere}} &= \lim_{\Delta r \to 0} \frac{V_{\text{mantle}}}{\Delta r} \\ &= \lim_{\Delta r \to 0} \frac{4\pi}{3} \cdot \frac{r^3 - (r - \Delta r)^3}{\Delta r} \\ &= \frac{4\pi}{3} \lim_{\Delta r \to 0} \frac{3r^2 \Delta r - 3r \Delta r^2 + \Delta r^3}{\Delta r} \\ &= \frac{4\pi}{3} \lim_{\Delta r \to 0} 3r^2 - 3r\Delta r + \Delta r^2 = 4\pi r^2 \end{aligned}$$

Another approach is to integrate directly over the surface. We subdivide the surface into an infinite number of rings. Each ring has a width $\mathrm{d}w$, a circumference l_C and an area $\mathrm{d}A$. The total surface area of the sphere is the integral over the ring surfaces $\mathrm{d}A$.



$$\begin{aligned} \mathrm{d}w &= r \mathrm{d}\varphi \\ l_C &= 2\pi r' = 2\pi r \sin\varphi \\ \mathrm{d}A &= l_C \mathrm{d}w = 2\pi r \sin\varphi \, r \mathrm{d}\varphi = 2\pi r^2 \sin\varphi \, \mathrm{d}\varphi \\ A_{\text{sphere}} &= \int_0^\pi \mathrm{d}A = 2\pi r^2 \int_0^\pi \sin\varphi \, \mathrm{d}\varphi \\ &= 2\pi r^2 \left[-\cos\varphi\right]_0^\pi = 2\pi r^2 [1 + 1] = 4\pi r^2 \end{aligned}$$

2.8. Mean values

In many situations mean or average values are of interest. Examples are the average salary, the mean temperature, the mean voltage or mean absorbed dose.

Depending on the type of question different types of mean values are required. In this section two types of mean values are investigated.

2.8.1. Arithmetic mean

What is the average income of an electrical engineer in his/her first job? What is the mean temperature in my office during the day?

For questions like this the *arithmetic mean* $\overline{x}$ is the value of interest. For discrete values x_k, $k = 1 \ldots n$ the arithmetic mean is defined by:

$$\overline{x} = \frac{1}{n} \sum_{k=1}^{n} x_k$$

To evaluate the arithmetic mean $\overline{y}$ of a function $y(x)$ an integral must be resolved:

$$\overline{y} = \frac{1}{x_2 - x_1} \int_{x_1}^{x_2} y(x) \, \mathrm{d}x$$

For functions with positive and negative values sometimes the absolute of the functions is used to evaluate the arithmetic mean:

$$\overline{y} = \frac{1}{x_2 - x_1} \int_{x_1}^{x_2} |y(x)| \, \mathrm{d}x$$

E.g. what is the arithmetic mean of a sine wave with amplitude $y = \hat{y} \sin(x)$? The sine wave is periodic, hence, integrating over one period 2π is sufficient. Not taking the absolute of y would result in a zero mean value:

$$\begin{aligned} \overline{y} &= \frac{1}{2\pi} \int_0^{2\pi} \hat{y} \sin(x) \, \mathrm{d}x = \frac{\hat{y}}{2\pi} [-\cos(x)]_0^{2\pi} \\ &= \frac{\hat{y}}{2\pi} (-1 + 1) = 0 \end{aligned}$$

Here we must integrate over the absolute of the sine wave:

$$\begin{aligned} \overline{y} &= \frac{1}{2\pi} \int_0^{2\pi} |\hat{y} \sin(x)| \, \mathrm{d}x \\ &= \frac{\hat{y}}{2\pi} \int_0^{\pi} \sin(x) \, \mathrm{d}x + \frac{\hat{y}}{2\pi} \int_\pi^{2\pi} -\sin(x) \, \mathrm{d}x \\ &= \frac{\hat{y}}{2\pi} [-\cos(x)]_0^\pi + \frac{\hat{y}}{2\pi} [\cos(x)]_\pi^{2\pi} \\ &= \frac{\hat{y}}{2\pi} (1 + 1) + \frac{\hat{y}}{2\pi} (1 + 1) = \frac{2\hat{y}}{\pi} \approx 0.634 \, \hat{y} \end{aligned}$$

2.8.2. Root mean square

In engineering often not a signal itself but the effect of a signal is of interest. In many situations the effect of a signal is proportional to its square value. Here the *root mean square* value is a useful quantity. The term describes the equation to evaluate the root mean square $\tilde{y}$ of a function $y(x)$:

$$\tilde{y} = \sqrt{\frac{1}{x_2 - x_1} \int_{x_1}^{x_2} y^2(x)\,\mathrm{d}x}$$

As an example we want to evaluate the root mean square value of a sine wave with amplitude $\hat{y}$. Since the sine wave is periodic it is sufficient to integrate over one period:

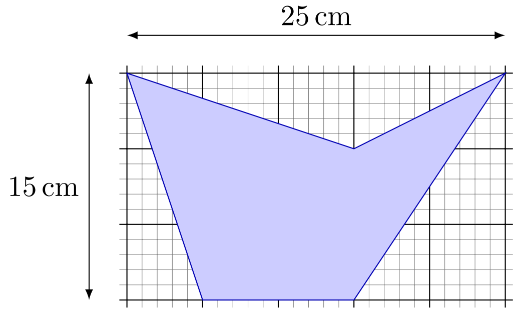
$$\begin{aligned}
\tilde{y} &= \sqrt{\frac{1}{2\pi} \int_0^{2\pi} \hat{y}^2 \sin^2(x)\,\mathrm{d}x} \\
&= \sqrt{\frac{\hat{y}^2}{2\pi} \int_0^{2\pi} \frac{1}{2}(1 - \cos(2x))\,\mathrm{d}x} \\
&= \sqrt{\frac{\hat{y}^2}{4\pi} \left(\int_0^{2\pi} 1\,\mathrm{d}x - \int_0^{2\pi} \cos(2x)\,\mathrm{d}x \right)} \\
&= \sqrt{\frac{\hat{y}^2}{4\pi} \left([x]_0^{2\pi} - \left[\frac{1}{2}\sin(2x)\right]_0^{2\pi} \right)} \\
&= \sqrt{\frac{\hat{y}^2}{4\pi} (2\pi - 0 + 0 - 0)} \\
&= \sqrt{\frac{\hat{y}^2}{2}} = \frac{\hat{y}}{\sqrt{2}} \approx 0.707\,\hat{y}
\end{aligned}$$

This is an important relationship in electrical engineering: Connecting an AC-voltage with amplitude $\hat{u}$ to a resistor will absorb the same amount of power as connecting a DC-voltage with value $\hat{u}/\sqrt{2}$ to the same resistor.

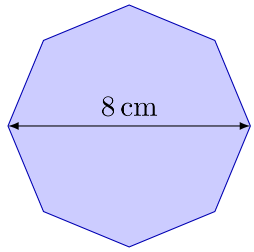
In Germany the mains voltage in households is $\tilde{u} = 230\,\mathrm{V}$. This is the root mean square value of the supplied AC-voltage. Hence, the amplitude of the AC voltage is: $\hat{u} = \tilde{u} \cdot \sqrt{2} \approx 325\,\mathrm{V}$, see example 1.1.

2.9. Problems

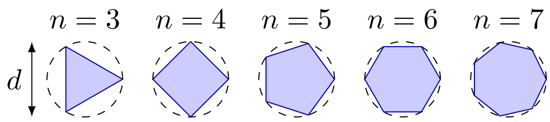
Problem 2.1: You order 500 fancy post cards printed on $120\,\mathrm{g/m^2}$ carton. The shape of the cards are given by the following drawing, where the thin grid lines have a distance of 1 cm. What is the mass m of all postcards in kg?



Problem 2.2: Find the area of a regular octagon with an outer diameter of 8 cm. (Hint: Make use of the symmetry of the octagon.)

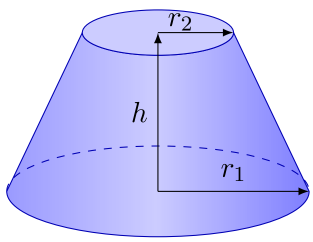


Problem 2.3: Find an equation for the area of a regular polygon with n corners placed on a circle with diameter d. (Hint: Make use of the symmetry of regular polygons.)



Problem 2.4: We derived the volume of a right circular cone with height h and base radius r as $V = \frac{1}{3}\pi r^2 h$ by a definite integral. Verify this equation by applying upper- and lower sums for a pile of discs and decreasing its thickness towards zero (i.e. increasing the number of disks toward infinity).

Problem 2.5: Evaluate the volume of a right circular frustum with height h, base radius r_1 and radius of upper plane r_2 by integral calculus.

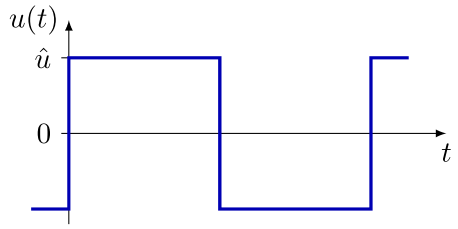


Problem 2.6: For a sphere with radius r we know the surface area to be $A = 4\pi r^2$. The

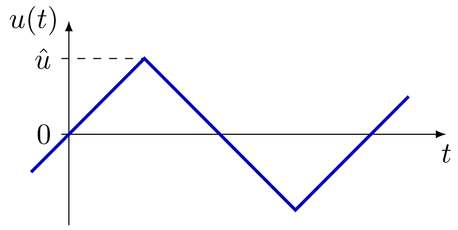
volume of a sphere may be evaluated by a sum of empty spheres with increasing radius r_k and wall thickness Δr.

1. Find the equation for the lower sum $V_L(n)$ by using the surface area $4\pi r_k^2$ and wall thickness Δr.
2. Find the equation for the upper sum $V_U(n)$.
3. Evaluate the limit towards zero wall thickness $\Delta r \to 0$ to verify the volume of a sphere to be $V = \frac{4}{3}\pi r^3$.
4. Derive the volume of the sphere by integrating the surface area over the radius.

Problem 2.7: For a square signal $u(t)$ with amplitude $\pm\hat{u}$ evaluate the arithmetic mean and the root mean square value.



Problem 2.8: For a triangle signal $u(t)$ with amplitude $\pm\hat{u}$ evaluate the arithmetic mean and the root mean square value.



3. Integration techniques

3.1. Improper integrals

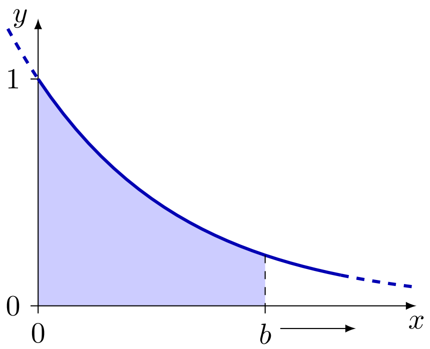
When integrating functions we sometimes have to deal with an infinite domain or image. In both cases the integral may or may not converge. We investigate the limit towards these points to study the behaviour of the integrals.

3.1.1. Infinite domain

To demonstrate the behaviour of integrals with infinite domain we show two examples, one being convergent, the other divergent:

Example 3.1. $\int_0^\infty e^{-x}\,\mathrm{d}x = ?$
We take the limit of the definite integral with the upper limit towards infinity:

$$\begin{aligned}\int_0^\infty e^{-x}\,\mathrm{d}x &= \lim_{b\to\infty}\int_0^b e^{-x}\,\mathrm{d}x\\ &= \lim_{b\to\infty}\left[-e^{-x}\right]_0^b\\ &= \lim_{b\to\infty}\left(-e^{-b}-(-e^{-0})\right)\\ &= \lim_{b\to\infty}\left(1-e^{-b}\right) = 1-0 = 1\end{aligned}$$

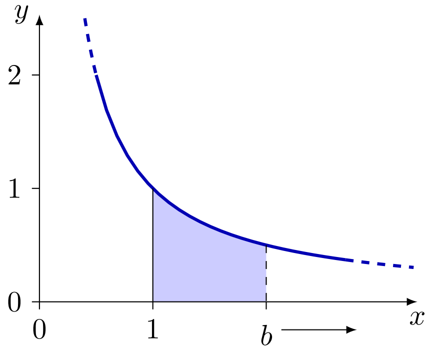


◁

Although the interval on the domain is infinite the integral in this example is finite. Is this because the function approaches zero, i.e. has a limit of zero towards infinity? Let's look at another example:

Example 3.2. $\int_1^\infty \frac{1}{x}\,\mathrm{d}x = ?$
We take the limit of the definite integral with the upper limit towards infinity:

$$\begin{aligned}\int_1^\infty \frac{1}{x}\,\mathrm{d}x &= \lim_{b\to\infty}\int_1^b \frac{1}{x}\,\mathrm{d}x\\ &= \lim_{b\to\infty}\left[\ln|x|\right]_1^b\\ &= \lim_{b\to\infty}\left(\ln|b|-\ln|1|\right)\\ &= \infty - 0 = \infty\end{aligned}$$



◁

In this example we again have a function that approaches zero towards infinity, however, here the integral is infinite. It depends on *how* the function tends towards zero. The exponential function e^{-x} converges faster towards zero than the reciprocal function x^{-1}.

3.1.2. Integrand with infinite image

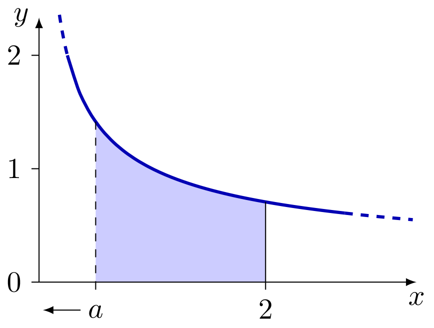
Care must be taken at points of discontinuity. For finite values the integral is convergent. For infinite values the integral may or may not converge.

The next two examples show integrals with infinite values, one of them being convergent the other being divergent. We use limits towards the point of discontinuity.

Example 3.3. $\int_0^2 \frac{1}{\sqrt{x}}\,\mathrm{d}x = ?$
We take the limit of the definite integral with the lower limit towards zero:

$$\begin{aligned}\int_0^2 \frac{1}{\sqrt{x}}\,\mathrm{d}x &= \lim_{a\to 0^+}\int_a^2 \frac{1}{\sqrt{x}}\,\mathrm{d}x\\ &= \lim_{a\to 0^+}\left[2\sqrt{x}\right]_a^2\end{aligned}$$

$$= \lim_{a \to 0^+} \left(2\sqrt{2} - 2\sqrt{a} \right)$$
$$= 2\sqrt{2} - 0 = 2\sqrt{2} \approx 2.828$$

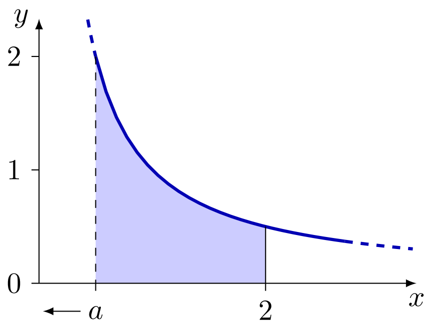


◁

Example 3.4. $\displaystyle\int_0^2 \frac{1}{x}\,\mathrm{d}x = ?$

We take the limit of the definite integral with the lower limit towards zero:

$$\int_0^2 \frac{1}{x}\,\mathrm{d}x = \lim_{a \to 0^+} \int_a^2 \frac{1}{x}\,\mathrm{d}x$$
$$= \lim_{a \to 0^+} \left[\ln|x|\right]_a^2$$
$$= \lim_{a \to 0^+} \left(\ln|2| - \ln|a|\right)$$
$$= \ln|2| - (-\infty) = \infty$$



◁

3.1.3. Summary

To study integrals with infinite domain or infinite values we use limits towards these points. For integrands with points of discontinuity we have to separate the integral into a sum of integrals each of them being continuous.

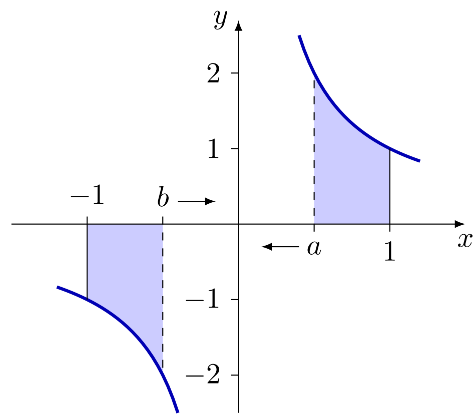
Care must be taken for definite integrals with known primitives containing poles within their domain. We can not simply take the difference of the primitive at the two limits, see the following example:

Example 3.5. $\displaystyle\int_{-1}^1 \frac{\mathrm{d}x}{x} = ?$

We are tempted to use the integral $\ln|x| + C$ and to evaluate the difference of the integral at the upper and lower limit $\ln|1| - \ln|-1| = 0$. But this is wrong! We have to analyse the integral stepwise at its discontinuous points:

$$\int_{-1}^1 \frac{\mathrm{d}x}{x} = \int_{-1}^0 \frac{\mathrm{d}x}{x} + \int_0^1 \frac{\mathrm{d}x}{x}$$
$$= \lim_{b \to 0^-} \int_{-1}^b \frac{\mathrm{d}x}{x} + \lim_{a \to 0^+} \int_a^1 \frac{\mathrm{d}x}{x}$$
$$= -\infty + \infty \Rightarrow \text{ undefined!}$$

Hence, care must be taken when integrating over discontinuous points.



◁

3.2. Integration by substitution

From differential calculus we know a number of basic functions with its derivatives. We further know some techniques to differentiate any combination of these functions.

For integral calculus we also know the a number of basic functions and its integrals. However, not all combinations of these functions are integrable analytically.

In this section we show a technique to integrate more complex integrands called *integration by substitution*. This integration technique is derived from the chain-rule of differential calculus. We start with the basic theorem:

Theorem 3.1 (Integration by substitution)**.** Let I be an interval, $f : I \to \mathbb{R}$ a continuous function and $\varphi : [a, b] \to I$ a continuously differentiable function (i.e. its derivative is continuous). We then have

$$\int_a^b f[\varphi(x)]\varphi'(x)\,\mathrm{d}x = \int_{\varphi(a)}^{\varphi(b)} f(\varphi)\,\mathrm{d}\varphi$$

◁

Proof. Let F be a primitive of f. We differentiate $F[\varphi(x)]$ with respect to x by applying the chain rule:

$$\frac{\mathrm{d}}{\mathrm{d}x}F[\varphi(x)] = F'[\varphi(x)]\varphi'(x) = f[\varphi(x)]\varphi'(x)$$

We now evaluate the definite integral:

$$\begin{aligned}\int_a^b f[\varphi(x)]\varphi'(x)\,\mathrm{d}x &= F[\varphi(b)] - F[\varphi(a)]\\ &= \int_{\varphi(a)}^{\varphi(b)} f(\varphi)\,\mathrm{d}\varphi\end{aligned}$$

□

We show the application of this theorem by examples:

Example 3.6. $\displaystyle\int_a^b 2\sin(2x)\,\mathrm{d}x = ?$

We see that the 2 in front of the sine-function is the first derivative of the argument of the sine-function. Hence, we take $2x$ as the inner function $\varphi(x)$:

$$\begin{aligned}\varphi(x) &= 2x \qquad \varphi'(x) = 2\\ f(\varphi) &= \sin(\varphi)\\ \int_a^b 2\sin(2x)\,\mathrm{d}x &= \int_a^b f(\varphi(x))\varphi'(x)\,\mathrm{d}x\\ &= \int_{\varphi(a)}^{\varphi(b)} f(\varphi)\,\mathrm{d}\varphi\\ &= \int_{\varphi(a)}^{\varphi(b)} \sin(\varphi)\,\mathrm{d}\varphi\\ &= \big[-\cos(\varphi)\big]_{\varphi(a)}^{\varphi(b)}\\ &= \big(-\cos(2b) - (-\cos(2a))\big)\\ &= \cos(2a) - \cos(2b)\end{aligned}$$

◁

Remark: Integration by substitution applies also to indefinite integrals, see the following example.

Example 3.7. $\displaystyle\int 6x\sqrt{3x^2+1}\,\mathrm{d}x = ?$

The first derivative of the argument of the root equals the factor before the root, hence:

$$\begin{aligned}\varphi(x) &= 3x^2+1 \qquad \varphi'(x) = 6x\\ f(\varphi) &= \sqrt{\varphi}\\ \int 6x\sqrt{3x^2+1}\,\mathrm{d}x &= \int \varphi'(x)f[\varphi(x)]\,\mathrm{d}x\\ &= \int f(\varphi)\,\mathrm{d}\varphi = \int \sqrt{\varphi}\,\mathrm{d}\varphi\\ &= \frac{2}{3}\sqrt{\varphi^3} + C\\ &= \frac{2}{3}\sqrt{(3x^2+1)^3} + C\end{aligned}$$

◁

We now leave this general approach and look at two special substitutions before we come back to the general case.

3.2.1. Integration of $f(ax+b)$

If we substitute $\varphi(x) = ax+b$ we get the first derivative $\varphi'(x) = a$ which makes substitution very simple:

Corollary 3.2 (Integration of $f(ax+b)$)**.** Let $f:\mathbb{R}\to\mathbb{R}$ be integrable, $a, b, x_1, x_2 \in \mathbb{R}$ and $\varphi = ax+b$. We then have:

$$\int_{x_1}^{x_2} f(ax+b)\,\mathrm{d}x = \frac{1}{a}\int_{ax_1+b}^{ax_2+b} f(\varphi)\,\mathrm{d}\varphi$$

◁

Proof. With $\varphi:\mathbb{R}\to\mathbb{R}, x\mapsto ax+b$ we have:

$$\varphi'(x) = \frac{\mathrm{d}\varphi}{\mathrm{d}x} = a \quad\Rightarrow\quad \mathrm{d}x = \frac{\mathrm{d}\varphi}{a}$$

Inserting into the integral we get:

$$\begin{aligned}\int_{x_1}^{x_2} f(ax+b)\,\mathrm{d}x &= \frac{1}{a}\int_{x_1}^{x_2} \varphi'(x)f[\varphi(x)]\,\mathrm{d}x\\ &= \frac{1}{a}\int_{ax_1+b}^{ax_2+b} f(\varphi)\,\mathrm{d}\varphi\end{aligned}$$

□

We apply this corollary in two examples:

Example 3.8. $\displaystyle\int_1^2 \sqrt{2x-1}\,\mathrm{d}x = ?$

$$a = 2 \qquad\qquad b = -1$$

$$\begin{aligned}\int_1^2 \sqrt{2x-1}\,\mathrm{d}x &= \int_{x_1}^{x_2} \sqrt{ax+b}\,\mathrm{d}x\\ &= \frac{1}{a}\int_{2x_1-1}^{2x_2-1} \sqrt{\varphi}\,\mathrm{d}\varphi\\ &= \frac{1}{2}\cdot\frac{2}{3}\left[\sqrt{\varphi^3}\right]_1^3\\ &= \frac{1}{3}(\sqrt{27}-\sqrt{1})\\ &= \sqrt{3} - \frac{1}{3} \approx 1.399\end{aligned}$$

◁

Example 3.9. $\int_0^{\frac{\pi}{2}} \cos\left(\frac{x}{3}\right) \mathrm{d}x = ?$

$$a = \frac{1}{3} \qquad b = 0$$

$$\begin{aligned}\int_0^{\frac{\pi}{2}} \cos\left(\frac{x}{3}\right) \mathrm{d}x &= \int_{x_1}^{x_2} \cos(ax)\,\mathrm{d}x \\ &= \frac{1}{a}\int_{ax_1}^{ax_2} \cos(\varphi)\,\mathrm{d}\varphi \\ &= 3\left[\sin(\varphi)\right]_0^{\frac{\pi}{6}} \\ &= 3\sin\left(\frac{\pi}{6}\right) = 1.5\end{aligned}$$

◁

The substitution holds also for indefinite integrals:

Example 3.10. $\int (3x-2)^6 \,\mathrm{d}x = ?$

We may solve this integral by expanding the integrand. However, since this is a tiring process we prefer integration by substitution:

$$\begin{aligned}\int (3x-2)^6\,\mathrm{d}x &= \frac{1}{3}\int \varphi^6\,\mathrm{d}\varphi \\ &= \frac{1}{3}\cdot\frac{\varphi^7}{7} + C \\ &= \frac{(3x-2)^7}{21} + C\end{aligned}$$

◁

3.2.2. Integration of $\frac{\varphi'(x)}{\varphi(x)}$

For this type of expression the outer function is $f(\varphi) = \frac{1}{\varphi}$ with the integral $\ln|\varphi| + C$ which leads us to the following corollary:

Corollary 3.3 (Integration of $\frac{\varphi'(x)}{\varphi(x)}$)**.** With $\varphi : [a,b] \to \mathbb{R}$ being a continuously differentiable function (i.e. its derivative is continuous) and $\varphi(x) \neq 0$ we have:

$$\int \frac{\varphi'(x)}{\varphi(x)}\,\mathrm{d}x = \ln|\varphi(x)| + C$$

◁

Proof. With $f(x) = \dfrac{1}{\varphi(x)}$ we have

$$\begin{aligned}\int \frac{\varphi'(x)}{\varphi(x)}\,\mathrm{d}x &= \int \varphi'(x)f(x)\,\mathrm{d}x \\ &= \int f(\varphi)\,\mathrm{d}\varphi = \int \frac{1}{\varphi}\,\mathrm{d}\varphi \\ &= \ln|\varphi| + C \\ &= \ln|\varphi(x)| + C\end{aligned}$$

□

We show the application of this corollary with some examples:

Example 3.11. $\int \frac{2x}{x^2-1}\,\mathrm{d}x = ?$

With $\varphi(x) = x^2 - 1$ and $\varphi'(x) = 2x$ we get:

$$\begin{aligned}\int \frac{2x}{x^2-1}\,\mathrm{d}x &= \int \frac{\varphi'(x)}{\varphi(x)}\,\mathrm{d}x \\ &= \ln|\varphi(x)| + C \\ &= \ln|x^2-1| + C\end{aligned}$$

◁

Example 3.12. $\int \frac{\tan(x)}{3}\,\mathrm{d}x = ?$

With $\varphi(x) = \cos(x)$ and $\varphi'(x) = -\sin(x)$ we get:

$$\begin{aligned}\int \frac{\tan(x)}{3}\,\mathrm{d}x &= -\frac{1}{3}\int \frac{-\sin(x)}{\cos(x)}\,\mathrm{d}x \\ &= -\frac{1}{3}\int \frac{\varphi'(x)}{\varphi(x)}\,\mathrm{d}x \\ &= C - \frac{1}{3}\ln|\cos(x)|\end{aligned}$$

◁

The corollary may also be applied to definite integrals:

Example 3.13. $\int_1^{\frac{3}{2}} \frac{x-1}{2x-x^2}\,\mathrm{d}x = ?$

With $\varphi(x) = 2x - x^2$ and $\varphi'(x) = 2 - 2x$ we get:

$$\begin{aligned}\int_1^{\frac{3}{2}} \frac{x-1}{2x-x^2}\,\mathrm{d}x &= -\frac{1}{2}\int_1^{\frac{3}{2}} \frac{2-2x}{2x-x^2}\,\mathrm{d}x \\ &= -\frac{1}{2}\int_1^{\frac{3}{2}} \frac{\varphi'(x)}{\varphi(x)}\,\mathrm{d}x \\ &= -\tfrac{1}{2}\left[\ln\left|2x-x^2\right|\right]_1^{\frac{3}{2}}\end{aligned}$$

Since the denominator of the integrand has no zero within the domain we get:

$$\begin{aligned}\int_1^{\frac{3}{2}} \frac{x-1}{2x-x^2}\,\mathrm{d}x &= -\tfrac{1}{2}(\ln|\tfrac{3}{4}| - \ln|1|) \\ &= \tfrac{1}{2}\ln\left(\tfrac{4}{3}\right) \approx 0.1438\end{aligned}$$

◁

3.2.3. Integration of $f[\varphi(x)]\varphi'(x)$

We now go back to the theorem at the beginning of this section:

$$\int_a^b f[\varphi(x)]\varphi'(x)\,\mathrm{d}x = \int_{\varphi(a)}^{\varphi(b)} f(\varphi)\,\mathrm{d}\varphi$$

Once we decided on the term to substitute a safe technique to perform substitution is:

1. Differentiate φ: $\varphi' = \dfrac{\mathrm{d}\varphi}{\mathrm{d}x}$
2. Rewrite this for $\mathrm{d}x$: $\mathrm{d}x = \dfrac{\mathrm{d}\varphi}{\varphi'}$
3. Insert φ and $\mathrm{d}x$ into the integral.
4. For a definite integral replace the limits a and b by $\varphi(a)$ and $\varphi(b)$, respectively.

Now the first derivative should cancel out and the required factors will remain.

Example 3.14. $\displaystyle\int \frac{2-x^2}{x^3-6x}\,\mathrm{d}x = ?$
We set $\varphi = x^3 - 6x$ and get:

$$\frac{\mathrm{d}\varphi}{\mathrm{d}x} = 3x^2 - 6 \qquad \mathrm{d}x = \frac{\mathrm{d}\varphi}{3x^2-6}$$

$$\begin{aligned}\int \frac{2-x^2}{x^3-6x}\,\mathrm{d}x &= \int \frac{2-x^2}{\varphi}\,\frac{\mathrm{d}\varphi}{3x^2-6}\\ &= -\frac{1}{3}\int \frac{3x^2-6}{\varphi}\,\frac{\mathrm{d}\varphi}{3x^2-6}\\ &= -\frac{1}{3}\int \frac{\mathrm{d}\varphi}{\varphi} = C - \frac{\ln|\varphi|}{3}\\ &= C - \frac{1}{3}\ln|x^3-6x|\end{aligned}$$

◁

If we plan for a definite integral to perform a back-substitution at the end, we may leave the limits as they are, and put brackets around them to indicate that we did not change them:

Example 3.15. $\displaystyle\int_0^{\frac{\pi}{4}} \frac{\sqrt{\tan(x)}}{2\cos^2(x)}\,\mathrm{d}x = ?$

$$\varphi = \tan(x), \quad \frac{\mathrm{d}\varphi}{\mathrm{d}x} = \frac{1}{\cos^2(x)}, \quad \mathrm{d}x = \cos^2(x)\mathrm{d}\varphi$$

$$\begin{aligned}\int_0^{\frac{\pi}{4}} \frac{\sqrt{\tan(x)}}{2\cos^2(x)}\,\mathrm{d}x &= \int_{(0)}^{(\frac{\pi}{4})} \frac{\sqrt{\varphi}}{2\cos^2(x)}\cos^2(x)\,\mathrm{d}\varphi\\ &= \frac{1}{2}\int_{(0)}^{(\frac{\pi}{4})} \sqrt{\varphi}\,\mathrm{d}\varphi\\ &= \frac{1}{2}\left[\frac{2}{3}\sqrt{\varphi^3}\right]_{(0)}^{(\frac{\pi}{4})}\\ &= \frac{1}{3}\left[\sqrt{\tan^3(x)}\right]_0^{\frac{\pi}{4}}\\ &= \frac{1}{3}\left(\sqrt{\tan^3(\tfrac{\pi}{4})} - \sqrt{\tan^3(0)}\right)\\ &= \frac{1}{3}\end{aligned}$$

◁

3.2.4. Summary

We looked at three types of integrals where substitution helps to solve the integral. Depending on the type of integral, we take the choice which part of the integrand to substitute:

type of integral	substitution
$\int f(ax+b)\,\mathrm{d}x$	$\varphi(x) = ax+b$
$\int \frac{f'(x)}{f(x)}\,\mathrm{d}x$	$\varphi(x) = f(x)$
$\int f[g(x)]g'(x)\,\mathrm{d}x$	$\varphi(x) = g(x)$

For definite integrals the limits must be adjusted by the substitution term or kept in brackets until back-substitution.

There are other ways to simplify integrals by substitution. However, we limit ourselves to this three most common substitutions.

3.3. Integration by parts

With most integration techniques we try to convert complicated integrands into one or more easy to integrate parts. We again make use of a rule from differential calculus: the *product rule*.

Theorem 3.4 (Integration by parts)**.** Let $f, g : [a, b] \to \mathbb{R}$ be differentiable and its derivative continuous. We then have:

$$\int f'(x)g(x)\,\mathrm{d}x = f(x)g(x) - \int f(x)g'(x)\,\mathrm{d}x$$

◁

Proof. Integration by parts is the inverse of the product rule in differential calculus:

$$(fg)' = f'g + fg'$$

$$
\begin{aligned}
f'g &= (fg)' - fg' \\
\int f'g \,\mathrm{d}x &= \int \{(fg)' - fg'\} \,\mathrm{d}x \\
&= \int (fg)'\mathrm{d}x - \int fg' \,\mathrm{d}x \\
&= fg - \int fg' \,\mathrm{d}x
\end{aligned}
$$

□

Example 3.16. $\int \cos(x)\sin(x)\,\mathrm{d}x = ?$

With $f' = \cos(x)$ and $g = \sin(x)$ we get:

$$
\begin{aligned}
\int \cos(x)\sin(x)\,\mathrm{d}x &= \int f'g \,\mathrm{d}x \\
&= fg - \int fg' \,\mathrm{d}x \\
&= \sin^2(x) - \int \sin(x)\cos(x)\mathrm{d}x
\end{aligned}
$$

By adding $\int \sin(x)\cos(x)\mathrm{d}x$ to the left side we get:

$$
\begin{aligned}
2\int \sin(x)\cos(x)\mathrm{d}x &= \sin^2(x) \\
\int \sin(x)\cos(x)\mathrm{d}x &= \frac{\sin^2(x)}{2} + C
\end{aligned}
$$

◁

Example 3.17. $\int e^x(x^2+2x)\,\mathrm{d}x = ?$

With $f' = e^x$ and $g = (x^2+2x)$ we get:

$$
\begin{aligned}
\int e^x(x^2+2x)\,\mathrm{d}x = e^x(x^2+2x) \\
- \int e^x(2x+2)\,\mathrm{d}x
\end{aligned}
$$

For the latter integral we again set $f' = e^x$ and $g = 2x+2$ and get:

$$
\int e^x(2x+2)\,\mathrm{d}x = e^x(2x+2) - \int e^x 2\,\mathrm{d}x
$$

The latter integral can be integrated directly:

$$
\int e^x 2\,\mathrm{d}x = 2\int e^x\,\mathrm{d}x = 2e^x + C
$$

Now we combine the integrals:

$$
\begin{aligned}
\int e^x(x^2+2x)\,\mathrm{d}x &= e^x(x^2+2x) \\
&\quad - e^x(2x+2) + 2e^x \\
&= e^x x^2 + C
\end{aligned}
$$

◁

Example 3.18. $\int \ln(x)\,\mathrm{d}x = ?$

With $f' = 1$ and $g = \ln(x)$ we get:

$$
\begin{aligned}
\int \ln(x)\,\mathrm{d}x &= \int 1\cdot\ln(x)\,\mathrm{d}x \\
&= x\ln(x) - \int x\frac{1}{x}\,\mathrm{d}x \\
&= x\ln(x) - x + C
\end{aligned}
$$

◁

Integration by parts should be applied if the remaining integral is less complex than the original integral. E.g. products of polynomials with a sine-, cosine- or exponential functions are candidates for integration by parts.

3.4. Integration of absolutes

For integrals where the integrand contains absolutes the integral must be separated into parts where the argument of the absolute does not change sign. See the following examples:

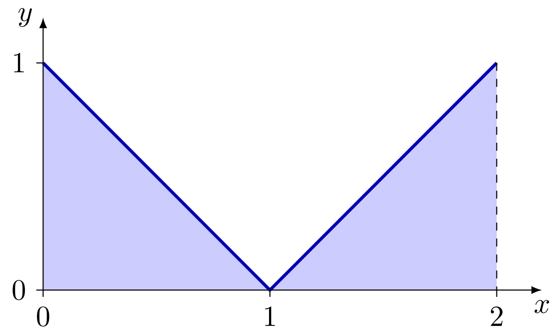
Example 3.19. $\int_0^2 |x-1|\,\mathrm{d}x = ?$

The argument of the absolute changes sign at $x = 1$, hence, we separate the integral in two parts:

$$
\int_0^2 |x-1|\,\mathrm{d}x = \int_0^1 |x-1|\,\mathrm{d}x + \int_1^2 |x-1|\,\mathrm{d}x
$$

For the left integral the argument of the absolute is always negative, hence, we remove the absolute by changing the sign. For the right integral the argument of the absolute is always positive, hence, we remove the absolute without any further change:

$$
\begin{aligned}
\int_0^2 |x-1|\,\mathrm{d}x &= \int_0^1 (1-x)\,\mathrm{d}x + \int_1^2 (x-1)\,\mathrm{d}x \\
&= \left[x - \tfrac{x^2}{2}\right]_0^1 + \left[\tfrac{x^2}{2} - x\right]_1^2 \\
&= (1-\tfrac{1}{2}) + (\tfrac{4}{2} - 2 - \tfrac{1}{2} + 1) = 1
\end{aligned}
$$

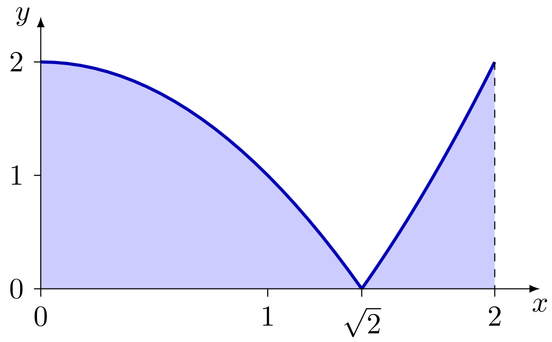


◁

Example 3.20. $\int_0^2 |x^2 - 2| \, \mathrm{d}x = ?$

The argument of the absolute is negative on the interval $[0, \sqrt{2})$ and positive on the interval $(\sqrt{2}, 2]$. Hence we separate the integral in two parts:

$$\begin{aligned}\int_0^2 |x^2-2| \, \mathrm{d}x &= \int_0^{\sqrt{2}} 2 - x^2 \, \mathrm{d}x + \int_{\sqrt{2}}^2 x^2 - 2 \, \mathrm{d}x \\ &= \left[2x - \tfrac{x^3}{3}\right]_0^{\sqrt{2}} + \left[\tfrac{x^3}{3} - 2x\right]_{\sqrt{2}}^2 \\ &= \left(2\sqrt{2} - \tfrac{2^{3/2}}{3}\right) + \left(\tfrac{8}{3} - 4 - \tfrac{2^{3/2}}{3} + 2\sqrt{2}\right) \\ &= \tfrac{4}{3}(2\sqrt{2} - 1) \approx 2.4379\end{aligned}$$



◁

3.5. Integration of rational functions

When analysing the frequency behaviour of systems one often gets to the point of integrating large rational functions of this type:

$$f(x) = \frac{a_n x^n + a_{n-1}x^{n-1} + \ldots + a_1 x + a_0}{b_m x^m + b_{m-1}x^{m-1} + \ldots + b_1 x + b_0}$$

In order to be able to integrate high order rational functions we reduce them into a sum of smaller rational functions by *partial fraction decomposition*. If we then know the integrals of the partial fractions we are able to integrate the whole rational function.

To do so we assume $m > n$, $a_n \neq 0$ and $b_m \neq 0$, i.e. the order of the denominator polynomial is greater than the order of the numerator polynomial. If not we first have to perform a polynomial division which we have seen earlier.

We remember four types of partial fractions depending on the poles of the rational function:

- single real poles
- multiple real poles
- single complex conjugate poles
- multiple complex conjugate poles

We limit ourselves to the first three types of poles.

3.5.1. Single real poles

Theorem 3.5 (Single real pole)**.** For a partial fraction of a single real pole of type $\frac{A}{x - x_0}$, $A, x_0 \in \mathbb{R}$ we have:

$$\int \frac{A}{x - x_0} \, \mathrm{d}x = A \ln |x - x_0| + C$$

◁

If a rational function has only single real poles we separate it into partial fractions of given type and integrate the summands separately.

Example 3.21. $\int \frac{3x+9}{x^2-4} \, \mathrm{d}x = ?$

We find two poles at $x_1 = 2$ and $x_2 = -2$:

$$x^2 - 4 = (x - x_1)(x - x_2) = (x-2)(x+2)$$

Hence, our ansatz for partial fraction decomposition is:

$$\frac{3x+9}{x^2-4} = \frac{A}{x-2} + \frac{B}{x+2}$$

Multiplying by the denominator of the left term $(x-2)(x+2)$ we get:

$$3x + 9 = A(x+2) + B(x-2)$$

We find A and B by setting x to 2 and -2, respectively:

$$\begin{aligned}3 \cdot 2 + 9 &= A(2+2) + B(2-2) \\ \Rightarrow A &= \frac{15}{4} \\ 3 \cdot (-2) + 9 &= A(-2+2) + B(-2-2) \\ \Rightarrow B &= -\frac{3}{4}\end{aligned}$$

This leads us to a simplified integral we are able to integrate:

$$\begin{aligned}\int \frac{3x+9}{x^2-4} \, \mathrm{d}x &= \int \left(\frac{15/4}{x-2} - \frac{3/4}{x+2}\right) \mathrm{d}x \\ &= \int \frac{15/4}{x-2} \, \mathrm{d}x - \int \frac{3/4}{x+2} \, \mathrm{d}x \\ &= \frac{15}{4} \ln |x-2| - \frac{3}{4} \ln |x+2| + C\end{aligned}$$

◁

Remark: For a definite integral care must be taken at poles. If a pole is included in the integration range the integral becomes improper.

3.5.2. Multiple real poles

Theorem 3.6 (Multiple real pole)**.** For a partial fraction of a multiple real pole of type $\frac{A}{(x-x_0)^n}$, $A, x_0 \in \mathbb{R}$, $n \in \mathbb{N}_{>1}$ we have:

$$\begin{aligned}\int \frac{A}{(x-x_0)^n}\,\mathrm{d}x &= C - \frac{A}{(n-1)(x-x_0)^{n-1}} \\ &= \frac{A(x-x_0)^{1-n}}{1-n} + C\end{aligned}$$

◁

If a rational function has only single and multiple real poles we separate it into partial fractions of given type and integrate the summands separately.

Example 3.22. $\int \frac{x^2+2x+1}{x^3-3x+2}\,\mathrm{d}x = ?$
We find a single real pole at $x_1 = -2$ and a double pole at $x_2 = 1$:

$$\begin{aligned}x^3-3x+2 &= (x-x_1)(x-x_2)^2 \\ &= (x+2)(x-1)^2\end{aligned}$$

Hence, our ansatz for partial fraction decomposition is:

$$\frac{x^2+2x+1}{x^3-3x+2} = \frac{A}{x+2} + \frac{B}{x-1} + \frac{C}{(x-1)^2}$$

Multiplying by the denominator of the left term $(x+2)(x-1)^2$ we get:

$$x^2+2x+1 = A(x-1)^2 + B(x+2)(x-1) + C(x+2)$$

We find A and C by setting x to -2 and 1, respectively:

$$\begin{aligned}(-2)^2+2(-2)+1 &= A(-2-1)^2 \\ \Rightarrow A &= \frac{1}{9} \\ 1^2+2\cdot 1+1 &= C(1+2) \\ \Rightarrow C &= \frac{4}{3}\end{aligned}$$

With known A and C we find B by setting x to another value, e.g. $x = 0$:

$$\begin{aligned}1 &= \frac{1}{9}\cdot(-1)^2 + B\cdot 2\cdot(-1) + \frac{4}{3}\cdot 2 \\ \Rightarrow B &= \frac{8}{9}\end{aligned}$$

This leads us to a sum of simple fractions we able to integrate:

$$\begin{aligned}&\int \frac{x^2+2x+1}{x^3-3x+2}\,\mathrm{d}x \\ &= \int \left(\frac{1/9}{x+2} + \frac{8/9}{x-1} + \frac{4/3}{(x-1)^2}\right)\mathrm{d}x \\ &= \frac{1}{9}\int \frac{\mathrm{d}x}{x+2} + \frac{8}{9}\int \frac{\mathrm{d}x}{x-1} + \frac{4}{3}\int \frac{\mathrm{d}x}{(x-1)^2} \\ &= \frac{1}{9}\ln|x+2| + \frac{8}{9}\ln|x-1| - \frac{4}{3(x-1)} + C\end{aligned}$$

◁

3.5.3. Single complex conjugate poles

We combine a pair of complex conjugate poles into a single partial fraction:

$$\frac{Ax+B}{(x-z)(x-\overline{z})} = \frac{Ax+B}{x^2+ax+b}$$

Theorem 3.7 (Single complex conjugate pole)**.** For a partial fraction of a single pair of complex conjugate poles of type $\frac{Ax+B}{x^2+ax+b}$, $A, B, a, b \in \mathbb{R}$ we have:

$$\begin{aligned}\int \frac{Ax+B}{x^2+ax+b}\,\mathrm{d}x &= \frac{A}{2}\ln|x^2+ax+b| \\ &+ \frac{2B-aA}{\sqrt{4b-a^2}}\arctan\left(\frac{2x+a}{\sqrt{4b-a^2}}\right) + C\end{aligned}$$

◁

If the denominator of a rational function contains a pair of complex conjugate zeros, partial fraction decomposition will contain a partial fraction of given type. We are then able to integrate this partial fraction.

Example 3.23. $\int \frac{x^2-4}{x^3-x^2+2}\,\mathrm{d}x = ?$
We find a single real pole at $x_1 = -1$ and complex conjugate poles at $z = 1+\mathrm{j}$ and $\overline{z} = 1-\mathrm{j}$:

$$\begin{aligned}x^3-x^2+2 &= (x-x_1)(x-z)(x-\overline{z}) \\ &= (x+1)(x-1-\mathrm{j})(x-1+\mathrm{j}) \\ &= (x+1)(x^2-2x+2)\end{aligned}$$

Hence, our ansatz for partial fraction decomposition is:

$$\frac{x^2-4}{x^3-x^2+2} = \frac{A}{x+1} + \frac{Bx+C}{x^2-2x+2}$$

Multiplying by the denominator of the left term $(x+1)(x^2-2x+2)$ we get:

$$x^2-4 = A(x^2-2x+2) + (Bx+C)(x+1)$$

We find A by setting x to -1:

$$(-1)^2-4 = A((-1)^2-2(-1)+2)$$

$$\Rightarrow A = -\frac{3}{5}$$

For $x = 0$ we find C:

$$-4 = 2A + C = C - \frac{6}{5}$$
$$\Rightarrow C = -\frac{14}{5}$$

Finally, setting x to one we find B:

$$1 - 4 = -\frac{3}{5} + 2B - \frac{28}{5}$$
$$\Rightarrow B = \frac{8}{5}$$

This leads us to an expression we are able to integrate:

$$\begin{aligned}
&\int \frac{x^2 - 4}{x^3 - x^2 + 2}\,\mathrm{d}x \\
&= \int \left(\frac{-\frac{3}{5}}{x+1} + \frac{\frac{8}{5}x - \frac{14}{5}}{x^2 - 2x + 2} \right) \mathrm{d}x \\
&= -\frac{3}{5} \int \frac{1}{x+1}\,\mathrm{d}x + \frac{1}{5} \int \frac{8x - 14}{x^2 - 2x + 2}\,\mathrm{d}x \\
&= -\frac{3}{5} \ln|x+1| + \frac{4}{5} \ln|x^2 - 2x + 2| \\
&\quad - \frac{6}{5} \arctan(x-1) + C
\end{aligned}$$

◁

3.5.4. Summary

The integration of a large rational function is performed by the following steps:

1. If required perform polynomial division so that the order of the numerator polynomial is less than the order of the denominator polynomial.
2. For the remaining rational function perform partial fraction decomposition.
3. Integrate all partial fractions separately by applying the integrals given above.

3.6. Problems

Problem 3.1: Solve the following improper integrals:

1. $\int_1^\infty \frac{\mathrm{d}x}{x^2}$
2. $\int_1^\infty \frac{\mathrm{d}x}{\sqrt{x^3}}$
3. $\int_0^2 \frac{\mathrm{d}x}{\sqrt[3]{x}}$
4. $\int_0^1 x^{-1/2}\,\mathrm{d}x$
5. $\int_1^\infty \frac{\mathrm{d}x}{x^k}$ for $k > 1$
6. $\int_0^2 x^k\,\mathrm{d}x$ for $k > -1$
7. $\int_{-1}^1 \sqrt{|x|}\,\mathrm{d}x$
8. $\int_{-e}^e \ln(|x|)\,\mathrm{d}x$

Problem 3.2: Integrate by substitution:

1. $\int \cos(x^2)2x\,\mathrm{d}x$
2. $\int e^{\sin\varphi}\cos\varphi\,\mathrm{d}\varphi$
3. $\int 3x^2 \sinh(x^3)\,\mathrm{d}x$
4. $\int \cos(\sin x)\cos x\,\mathrm{d}x$
5. $\int \frac{x}{x^2+1}\,\mathrm{d}x$
6. $\int \exp(ax+b)\,\mathrm{d}x$

Problem 3.3: Integrate by substitution of type $f(ax+b)$:

1. $\int 3\cos(3x+1)\,\mathrm{d}x$
2. $\int \sin(1-x)\,\mathrm{d}x$
3. $\int_{-1}^1 \sqrt[3]{x+2}\,\mathrm{d}x$
4. $\int_1^2 \sqrt{2x-2}\,\mathrm{d}x$
5. $\int_{-\pi/6}^{\pi/6} e^{-\mathrm{j}x}\,\mathrm{d}x$
6. $\int 2\pi\mathrm{j}f e^{2\pi\mathrm{j}ft}\,\mathrm{d}t$

Problem 3.4: Integrate by substitution of type $\frac{\varphi'(x)}{\varphi(x)}$:

1. $\int \frac{x}{x^2+1}\,\mathrm{d}x$
2. $\int \frac{\sinh(x)}{\cosh(x)}\,\mathrm{d}x$
3. $\int_{\pi/6}^{\pi/2} \cot(x)\,\mathrm{d}x$
4. $\int \frac{3\sqrt{x}+4x}{\sqrt{x^3}+x^2}\,\mathrm{d}x$
5. $\int \frac{x-2x^3}{x^4-x^2}\,\mathrm{d}x$
6. $\int_0^{\pi/6} \frac{\cos(x)\,\mathrm{d}x}{2\sin(x)+1}$

Problem 3.5: Integrate by substitution:

1. $\int_0^1 6x\,\cos(3x^2)\,\mathrm{d}x$
2. $\int_0^1 \sqrt{x}\,e^{\sqrt{x^3}}\,\mathrm{d}x$
3. $\int \cos(x)e^{\sin(x)}\,\mathrm{d}x$
4. $\int x^2\,\sin(x^3)\,\mathrm{d}x$
5. $\int_{-1}^1 x\,e^{x^2}\,\mathrm{d}x$
6. $\int \frac{\sin(2x)}{\sin^2(x)}\,\mathrm{d}x$

Problem 3.6: Integrate by parts:

1. $\int xe^x \,\mathrm{d}x$
2. $\int x^3 e^{-x} \,\mathrm{d}x$
3. $\int x^2 \cos(x) \,\mathrm{d}x$
4. $\int x \sinh(x) \,\mathrm{d}x$
5. $\int \pi x \, e^{\pi x} \,\mathrm{d}x$
6. $\int \sin^2(x) \,\mathrm{d}x$

Problem 3.7: Integrate the following functions:

1. $\int_0^1 \left|x - \frac{1}{2}\right| \,\mathrm{d}x$
2. $\int_{-2}^2 \left|x^2 - 1\right| \,\mathrm{d}x$
3. $\int_{-2}^2 \left|1 - x^2\right| \,\mathrm{d}x$
4. $\int_0^3 \left|x^2 - 2x\right| \,\mathrm{d}x$
5. $\int_0^{2\pi} \left|\mathrm{Im}(e^{\mathrm{j}x})\right| \,\mathrm{d}x$
6. $\int_0^{\pi} \left|\sin^2(x) - \frac{1}{2}\right| \,\mathrm{d}x$

Problem 3.8: Integrate by partial fraction decomposition:

1. $f(x) = \dfrac{1}{x^2 - 1}$
2. $f(x) = \dfrac{2x + 3}{x^2 + 4x + 4}$
3. $f(x) = \dfrac{x^2 + 1}{x^3 + 2x^2 + 2x}$
4. $f(x) = \dfrac{x^3 + x}{x^4 + 2x^3 + x^2 - 2x - 2}$

4. Functions with multiple arguments and values

4.1. Introduction

Yet we dealt with functions having one argument and one value, i.e.

$$f : \mathbb{R} \to \mathbb{R}$$

In this chapter we want to extend our view to functions with n arguments and m values, i.e.

$$f : \mathbb{R}^n \to \mathbb{R}^m$$

Example 4.1. The absorbed power in a resistor obeying Ohm's law is given by

$$P = \frac{U^2}{R}$$

The absorbed power depends on both: the applied voltage and the resistance of the resistor. Hence, the function for the absorbed power is a function with two arguments and one value:

$$P = f(U,R) : \begin{cases} \mathbb{R}^2 \to \mathbb{R} \\ U, R \mapsto \frac{U^2}{R} \end{cases}$$

◁

Some further examples are given in table 4.1.

4.2. Definition of multiple argument functions

Definition 4.1 (function with n arguments)**.** We call u a function with n arguments $x_1, x_2, \ldots, x_n$ if u takes a unique value for each valid combination of $x_1, x_2, \ldots, x_n$ and write:

$$u = f(x_1, x_2, \ldots, x_n)$$

◁

Example 4.2. The kinetic energy E_{kin} of a solid, not rotating body with mass m and velocity v is given by:

$$E_{\text{kin}} = f(m,v) = \frac{m}{2}v^2$$

A change of mass or velocity both influence the kinetic energy. For a given combination of m and v we find a unique kinetic energy E_{kin}.

E.g. a car with mass $m = 1\,\text{t} = 1000\,\text{kg}$ and velocity $v = 36\,\text{km/h} = 10\,\text{m/s}$ has a kinetic energy of

$$\begin{aligned} E_{\text{kin}} = f(m,v) &= \frac{m}{2}v^2 \\ &= \frac{1000\,\text{kg}}{2}(10\,\tfrac{\text{m}}{\text{s}})^2 = 50\,\text{kJ} \end{aligned}$$

◁

4.3. Visualization of functions with multiple arguments

There are different ways to visualize a function with more than one argument.

For a continuous function $y = f(x)$ with one in- and output we were able to illustrate it by a line in a Cartesian diagram. I.e. each tuple (x, y) have a unique position in the diagram.

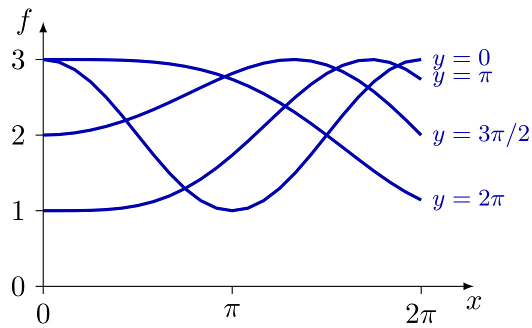
In this section we look at some techniques to visualize functions with more than one input.

4.3.1. Parametric plot

A first technique to visualize a function with more than one input is to keep all inputs except one constant and to plot it as a single argument function. To show the dependence on the other arguments the plot is done several times for different values for the constant arguments.

Example 4.3.

$$f : \begin{cases} \mathbb{R}^2 \to \mathbb{R} \\ (x,y) \mapsto 2 + \cos\sqrt{x^2+y^2} \end{cases}$$



◁

$\mathbb{R} \to \mathbb{R}$	$\mathbb{R} \to \mathbb{R}^2$	$\mathbb{R} \to \mathbb{R}^3$
- voltage over time - altitude over distance - resistance over temperature	- curve on a plane - position on surface over time	- curve in space - 3d positioning over time
$\mathbb{R}^2 \to \mathbb{R}$ - altitude map - grey scale picture	$\mathbb{R}^2 \to \mathbb{R}^2$ - surface vector field - surface transformation - wind on earth's surface	$\mathbb{R}^2 \to \mathbb{R}^3$ - surface in space - surface of geometrical bodies
$\mathbb{R}^3 \to \mathbb{R}$ - temperature in a solid - mass density over volume	$\mathbb{R}^3 \to \mathbb{R}^2$ - 3d to 2d projection - pressure and temperature over space	$\mathbb{R}^3 \to \mathbb{R}^3$ - volume vector field - air flow in volume

Table 4.1.: Examples for functions with multiple arguments and values

4.3.2. Surface plot

We may look at a function with two arguments and one output $z = f(x, y)$ as an infinite set of 3-tuples:

$$\{x, y, f(x, y)\}$$

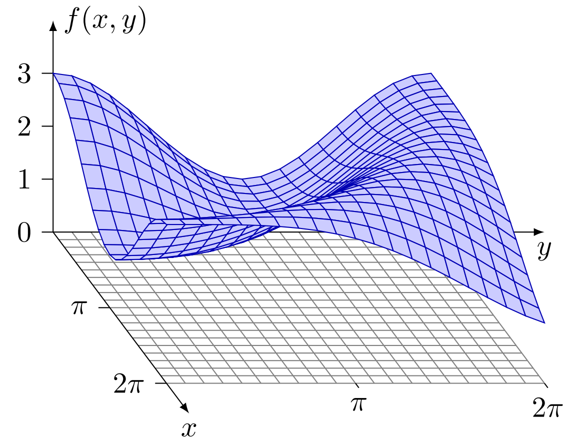
Each tuple represents a point in space. Since a three dimensional model is laborious to create we look for two dimensional plots of such functions.

For continuous functions the set of all tuples result in a surface in space. When drawing lines with equal distance on the x-y plane and plotting the function with $f(x, y)$ above this plane, the function becomes visible.

Example 4.4. The following plot visualizes the function

$$f : \begin{cases} \mathbb{R}^2 \to \mathbb{R} \\ (x, y) \mapsto 2 + \cos\sqrt{x^2 + y^2} \end{cases}$$

over the range $x, y \in [0, 2\pi]$ as a *surface plot*.



◁

4.3.3. Contour plot

Another technique to visualize functions with two arguments and one output is to use level curves like contour lines on maps, see the following example.

Example 4.5. We visualize the function

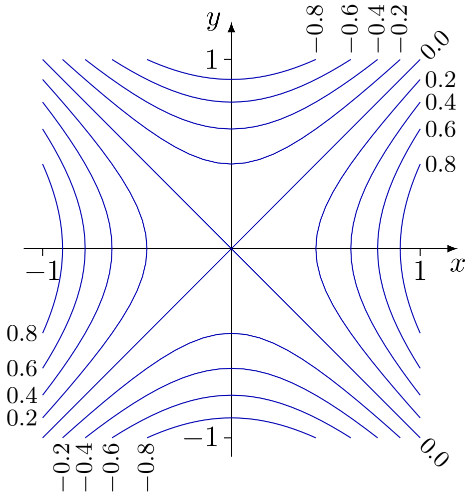
$$f : \begin{cases} \mathbb{R}^2 \to \mathbb{R} \\ (x, y) \mapsto x^2 - y^2 \end{cases}$$

by a *contour plot* over the range $x, y \in [-1, 1]$. To do so we resolve the function $f(x, y)$ to $y(x, f)$:

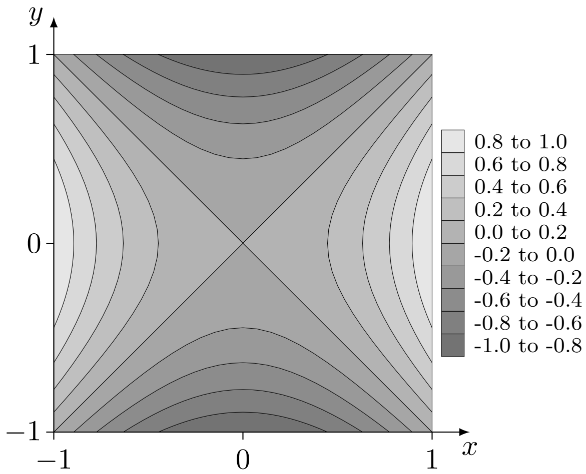
$$f(x, y) = x^2 - y^2 \quad \Rightarrow \quad y(x, f) = \pm\sqrt{x^2 - f}$$

We want to plot level curves at multiples of 0.2. E.g. for the level curve at 0 we insert $f = 0$

into the equation and get $y = \pm x$ which are the diagonals in the following plot:



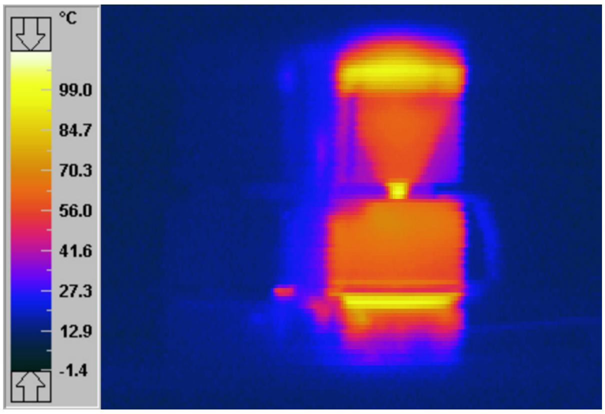
To improve the impression in such a plot we may colour the areas and add a legend to it:



◁

Example 4.6. An application is to measure the thermal emission of a house or a device. Different temperatures are illustrated by different intensities or colours.

The image below shows the thermal emission of a coffee machine.



Wikimedia Commons, Torsten Henning, public domain

◁

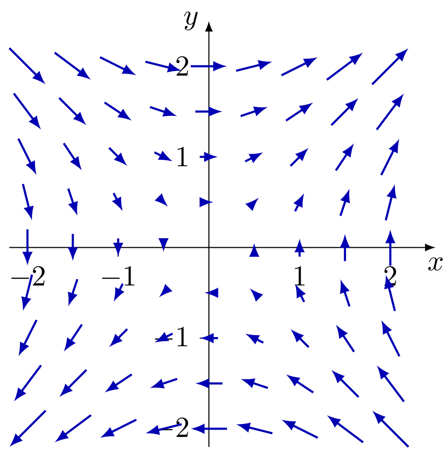
4.3.4. Vector plot

To visualize a function with two input- and two output-values we may distribute arrows on a Cartesian diagram. The position of the arrows show the input values and the horizontal and vertical size represent the output values.

Example 4.7. We visualize the function:

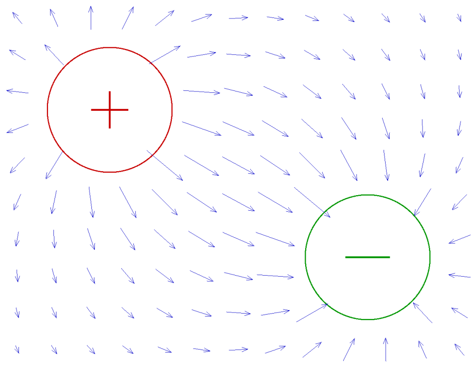
$$f : \begin{cases} \mathbb{R}^2 \to \mathbb{R}^2 \\ (x, y) \mapsto (y, x) \end{cases}$$

by a *vector plot*



◁

Example 4.8. A vector plot is useful to illustrate electric fields in space. The image below shows the electric field of a dipole.



◁

4.4. Continuity

For a function f with one argument we defined a continuous point x_0 by

$$\lim_{x \to {x_0}^+} f(x) = \lim_{x \to {x_0}^-} f(x) = f(x_0)$$

In short we write

$$\lim_{x \to x_0} f(x) = f(x_0)$$

implying that $x \to x_0$ includes both $x \to {x_0}^+$ and $x \to {x_0}^-$. I.e. the function value $f(x_0)$ must equal the limit from left ($x < x_0$) and right ($x > x_0$).

For functions with more than one argument we define continuity by the same technique: The function value must be equal to the limit from any direction. If we combine all arguments to a vector, i.e. $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)^T$ we write in short:

$$\lim_{\boldsymbol{x} \to \boldsymbol{x_0}} f(\boldsymbol{x}) = f(\boldsymbol{x_0})$$

The term $\boldsymbol{x} \to \boldsymbol{x_0}$ means that $\boldsymbol{x}$ approaches $\boldsymbol{x_0}$ from *all* directions and on *any* curve.

Example 4.9. Is the following function continuous at $\boldsymbol{x_0} = (0, 0)$?

$$f : \begin{cases} \mathbb{R}^2 \to \mathbb{R} \\ (x, y) \mapsto 1 \text{ for } x = y^2, y > 0 \\ (x, y) \mapsto 0 \text{ else} \end{cases}$$

The function value at $\boldsymbol{x_0}$ is zero since $y = 0 \not> 0$. We check the x-axis:

$$\lim_{x \to 0} f[(x, 0)] = 0$$

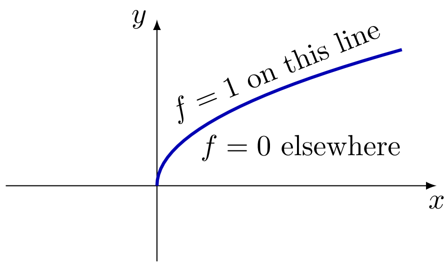
We check the y-axis:

$$\lim_{y \to 0} f[(0, y)] = 0$$

We are tempted to conclude that f is continuous at $\boldsymbol{x_0}$, but this would be wrong: If we take the limit towards $\boldsymbol{x_0}$ along the curve $x = y^2, y > 0$ we get:

$$\lim_{\varepsilon \to 0^+} f[(\varepsilon^2, \varepsilon)] = 1 \neq 0$$

Hence, the investigated function f is not continuous at $\boldsymbol{x_0} = (0, 0)$.



◁

It seems an unsolvable task to check continuity with limits on any curve. A technique to do so is to transform the Cartesian coordinates to polar coordinates around the investigated point $\boldsymbol{x_0}$. However, we do not apply this technique here.

The following theorem helps to analyse continuity for most applications.

Theorem 4.2 (Combined continuous functions)**.** Let $f, g : \mathbb{R}^n \to \mathbb{R}$ be continuous. If defined, then $f \pm g$, $f \cdot g$, f/g and $f \circ g$ are continuous too. ◁

Example 4.10. $f : \begin{cases} \mathbb{R}^2 \to \mathbb{R} \\ (x, y) \mapsto \dfrac{\sin(x) + y^2}{e^y} \end{cases}$
is continuous since the function is a valid combination of continuous functions. ◁

Example 4.11. $f : \begin{cases} \mathbb{R}^2 \to \mathbb{R} \\ (x, y) \mapsto \dfrac{\cos(xy)}{x^2 + y^2} \end{cases}$
is continuous except for the point $(0, 0)$ where the denominator is zero. ◁

4.5. Partial derivative

For functions with one output the derivative somehow represents the slope at any point. For functions with one in- and output we took the derivative only in one direction: towards positive arguments. Taking the derivative towards negative arguments would change the sign. E.g. driving up the hill means a positive slope whereas driving in the other direction down the hill means a negative slope.

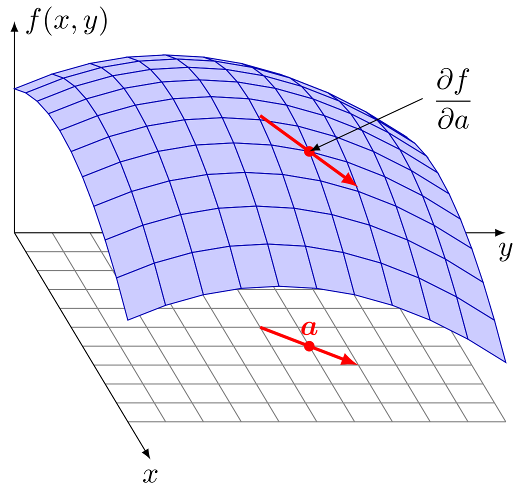
For a multiple input function we not only have a the choice between positive and negative direction but between an infinite number of directions to take the derivative. Hence, we first introduce the directional derivative towards a given direction $\boldsymbol{a}$ before we take a more general approach.

Definition 4.3 (Directional derivative)**.** Let $f : \mathbb{R}^n \to \mathbb{R}, (x_1, \dots, x_n) \mapsto f(x_1, \dots, x_n)$ and $\boldsymbol{a} = (a_1, \dots, a_n) \in \mathbb{R}^n$. If it exits we call

$$\frac{\partial f}{\partial \boldsymbol{a}} = \lim_{h \to 0} \frac{f(x_1{+}ha_1,\dots,x_n{+}ha_n) - f(x_1,\dots,x_n)}{h}$$

directional derivative with respect to $\boldsymbol{a}$. ◁

We may look at the directional derivative as the slope on a hilly ground. Depending on the direction the slope/derivative may turn out quite different.

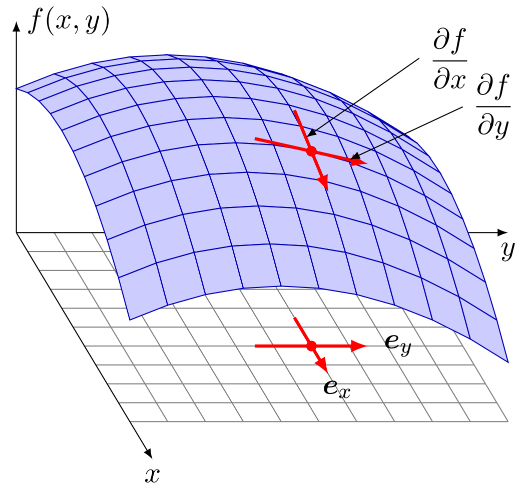


We have a special interest in the direction of the arguments of the function. I.e. we take the derivative w.r.t. one input while keeping the other inputs constant. These are the *partial derivatives* of a function.

Definition 4.4 (Partial derivative)**.** Let $f : \mathbb{R}^n \to \mathbb{R}, (x_1, \dots, x_n) \mapsto f(x_1, \dots, x_n)$ be a function with n arguments. If it exists we call

$$\begin{aligned}\frac{\partial f}{\partial x_k} &= \lim_{h \to 0} \frac{f(\boldsymbol{x} + h\boldsymbol{e_k}) - f(\boldsymbol{x})}{h} \\ &= \lim_{h \to 0} \frac{f(x_1,\dots,x_k{+}h,\dots,x_n) - f(x_1,\dots,x_n)}{h}\end{aligned}$$

the *partial derivative* with respect to x_k. ◁



The partial derivative of a function $f : \mathbb{R}^n \to \mathbb{R}$ again is a function with n arguments and one value.

To derive the partial derivative with respect to one argument we treat the other arguments as constants.

Example 4.12. We evaluate all partial derivatives of $f(x, y, z) = xy + z^2$ and get:

$$\frac{\partial f}{\partial x} = y \qquad \frac{\partial f}{\partial y} = x \qquad \frac{\partial f}{\partial z} = 2z$$

◁

Example 4.13. The partial derivatives of the function $f(x, y, z) = \sqrt{x^2 + y^2} + z^2$ are:

$$\begin{aligned}\frac{\partial f}{\partial x} &= \frac{2x}{2\sqrt{x^2 + y^2}} \\ \frac{\partial f}{\partial y} &= \frac{2y}{2\sqrt{x^2 + y^2}} \\ \frac{\partial f}{\partial z} &= 2z\end{aligned}$$

◁

Example 4.14. We evaluate all partial derivatives of the function $f(x, y, z) = x^2 \sin(y) e^z$ and get:

$$\begin{aligned}\frac{\partial f}{\partial x} &= 2x \sin(y) e^z \\ \frac{\partial f}{\partial y} &= x^2 \cos(y) e^z \\ \frac{\partial f}{\partial z} &= x^2 \sin(y) e^z\end{aligned}$$

◁

Differentiability in general may be derived as for functions with one argument. However, as for continuity we must take all directions and all possible paths towards the investigated point into account.

Another approach is to approximate the function at the point of interest by a plane and to investigate whether the remaining error tends towards zero by a higher order than one.

It is left to the interested reader to study differentiability by other sources.

4.6. Gradient

Definition 4.5 (Gradient)**.** Let $f : \mathbb{R}^n \to \mathbb{R}$ be a differentiable function with n Cartesian arguments. Combining all partial derivatives to a vector gives us

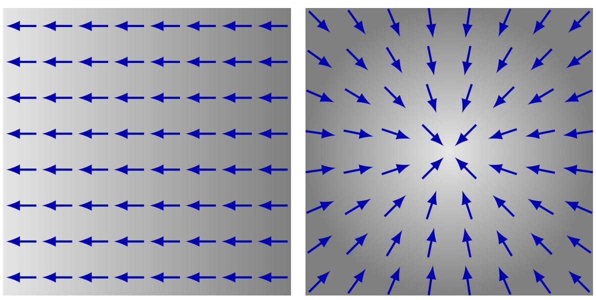
$$\frac{\partial f}{\partial x_1}\boldsymbol{e_1} + \ldots + \frac{\partial f}{\partial x_n}\boldsymbol{e_n} = \begin{pmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{pmatrix}$$

which we call the *gradient of a function* f. In short we write

$$\text{grad} = \nabla = \frac{\partial}{\partial x_1}\boldsymbol{e_1} + \ldots + \frac{\partial}{\partial x_n}\boldsymbol{e_n} = \begin{pmatrix} \frac{\partial}{\partial x_1} \\ \vdots \\ \frac{\partial}{\partial x_n} \end{pmatrix}$$

with ∇ being the *nabla-operator*. We treat the product of $\frac{\partial}{\partial x_k}$ and f as the partial derivative of f with respect to x_k. ◁

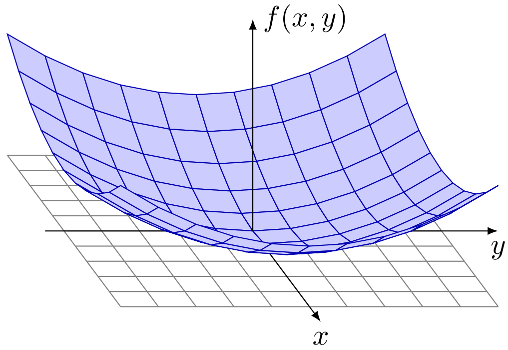
The gradient grad(f) of a function f at point $\boldsymbol{x}$ gives the value and direction of maximum slope of f.



Example 4.15. The shape of a parabola antenna may be expressed by:

$$f : \begin{cases} \mathbb{R} \times \mathbb{R} \to \mathbb{R} \\ (x, y) \mapsto x^2 + y^2 \end{cases}$$

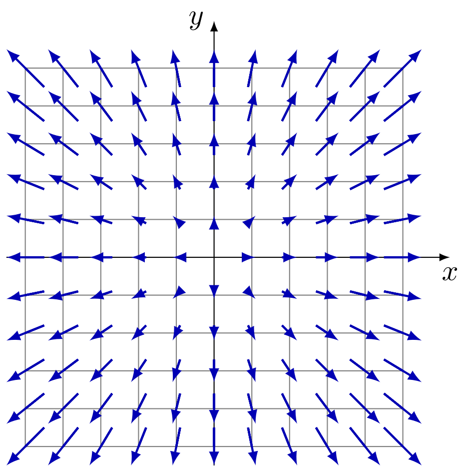
We express this function by a surface plot:



The slope increases with distance to the z-axis and points away from the z-axis.

$$\nabla f = \text{grad}(f) = \frac{\partial}{\partial x} f(x, y)\boldsymbol{e}_x + \frac{\partial}{\partial y} f(x, y)\boldsymbol{e}_y$$
$$= 2x\boldsymbol{e}_x + 2y\boldsymbol{e}_y$$

The gradient grad(f) has two in- and output values. We use a vector plot to illustrate the gradient:



◁

4.7. Multiple partial derivative

We now want to evaluate the second partial derivative of a function. For n arguments we find n^2 possible second partial derivatives.

Example 4.16. For the function

$$f : \begin{cases} \mathbb{R}^3 \to \mathbb{R} \\ (x, y, z) \mapsto xy^2z^3 \end{cases}$$

we want to find all possible second partial derivatives. We first evaluate the first partial derivatives:

$$\tfrac{\partial f}{\partial x} = y^2z^3 \qquad \tfrac{\partial f}{\partial y} = 2xyz^3 \qquad \tfrac{\partial f}{\partial z} = 3xy^2z^2$$

Now we take all second partial derivatives:

$$\begin{array}{lll} \frac{\partial^2 f}{\partial x^2} = 0 & \frac{\partial^2 f}{\partial x \partial y} = 2yz^3 & \frac{\partial^2 f}{\partial x \partial z} = 3y^2z^2 \\ \frac{\partial^2 f}{\partial y \partial x} = 2yz^3 & \frac{\partial^2 f}{\partial y^2} = 2xz^3 & \frac{\partial^2 f}{\partial y \partial z} = 6xyz^2 \\ \frac{\partial^2 f}{\partial z \partial x} = 3y^2z^2 & \frac{\partial^2 f}{\partial z \partial y} = 6xyz^2 & \frac{\partial^2 f}{\partial z^2} = 6xy^2z \end{array}$$

◁

Remark: We read the denominator backward. I.e. for the second derivative below we take first the derivative w.r.t. x and the second w.r.t. y:

$$\frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)$$

Combining all second partial derivatives leads us to the Hessian matrix:

Definition 4.6 (Hessian matrix)**.** The *Hessian matrix* combines all possible second partial derivatives of a multiple argument function into a matrix. I.e. with

$$f : \begin{cases} \mathbb{R}^n \to \mathbb{R} \\ (x_1, \ldots, x_n) \mapsto f(x_1, \ldots, x_n) \end{cases}$$

we get

$$H = \begin{pmatrix} \frac{\partial^2 f}{\partial {x_1}^2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f}{\partial {x_n}^2} \end{pmatrix}$$

◁

Example 4.17. The function

$$f : \begin{cases} \mathbb{R}^3 \to \mathbb{R} \\ (x, y, z) \mapsto xy^2z^3 \end{cases}$$

of the previous example results in the Hessian matrix:

$$H = \begin{pmatrix} 0 & 2yz^3 & 3y^2z^2 \\ 2yz^3 & 2xz^3 & 6xyz^2 \\ 3y^2z^2 & 6xyz^2 & 6xy^2z \end{pmatrix}$$

◁

The Hessian matrix of the previous example is symmetric. This is not always the case. However, if all second partial derivatives of a function are continuous, then the Hessian matrix is symmetric:

Theorem 4.7 (Symmetry of second derivatives)**.** Let

$$f : \begin{cases} \mathbb{R}^n \to \mathbb{R} \\ (x_1, \ldots, x_n) \mapsto f(x_1, \ldots, x_n) \end{cases}$$

be a function with continuous second partial derivatives. Then the order of partial differentiation may be interchanged, i.e.

$$\frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_k}\right) = \frac{\partial}{\partial x_k}\left(\frac{\partial f}{\partial x_j}\right)$$

for $j, k = 1, \ldots, n$. This *symmetry of second derivatives* is also known as the *Schwarz integrability condition.* ◁

4.8. Extrema

We remember, for a single in- and output function $f : \mathbb{R} \to \mathbb{R}$ we have a local maximum at x_0 if all values $f(x)$ in an environment $x \in U_\varepsilon(x_0)$ around the maximum are less or equal than the value of the maximum. I.e.

$$f(x_0) \geq f(x) \qquad \text{for } x \in U_\varepsilon(x_0)$$

The corresponding is true for local minima. The same principle holds for multiple argument functions:

Definition 4.8 (Extrema of multiple argument functions)**.** Let $D \subseteq \mathbb{R}^n$ be the domain of a multiple argument function $f : D \to \mathbb{R}$, $\boldsymbol{x}_0 \in D$ an element of the domain and $U_\varepsilon(\boldsymbol{x}_0)$ be the epsilon neighbourhood around $\boldsymbol{x}_0$.

- If there exist an $\varepsilon > 0$ with $f(\boldsymbol{x}_0) \leq f(\boldsymbol{x})$ for all $\boldsymbol{x} \in U_\varepsilon(\boldsymbol{x}_0)$ we say $\boldsymbol{x}_0$ is a *local minimum* of f.
- If there exist an $\varepsilon > 0$ with $f(\boldsymbol{x}_0) \geq f(\boldsymbol{x})$ for all $\boldsymbol{x} \in U_\varepsilon(\boldsymbol{x}_0)$ we say $\boldsymbol{x}_0$ is a *local maximum* of f.
- If $f(\boldsymbol{x}_0) \leq f(\boldsymbol{x})$ for all $\boldsymbol{x} \in D$ we say $\boldsymbol{x}_0$ is a *global minimum* of f.
- If $f(\boldsymbol{x}_0) \geq f(\boldsymbol{x})$ for all $\boldsymbol{x} \in D$ we say $\boldsymbol{x}_0$ is a *global maximum* of f.

If $\boldsymbol{x}_0$ is a local minimum or a local maximum we call it a *local extremum*. If $\boldsymbol{x}_0$ is a global minimum or a global maximum we call it a *global extremum*. ◁

Theorem 4.9 (Condition for extremum)**.** If a differentiable multiple argument function $f : \mathbb{R}^n \to \mathbb{R}$ has a (local) extremum at $\boldsymbol{x}_0 \in \mathbb{R}^n$ then all partial derivatives are zero at this point, i.e.

$$\left.\frac{\partial f}{\partial x_k}\right|_{\boldsymbol{x}_0} = 0 \qquad \text{for } k = 1, \ldots, n$$

◁

Zeros for all partial derivatives are a necessary condition for an extremum, but not a sufficient condition, see the following example.

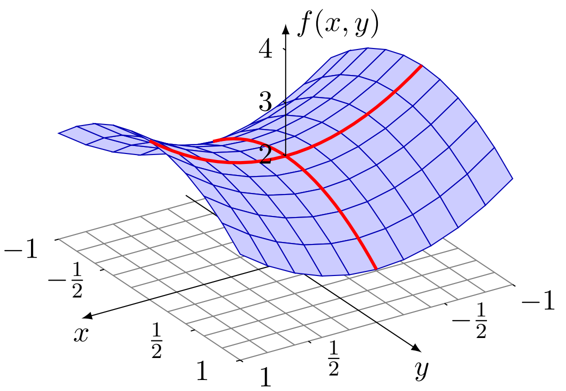
Example 4.18. Does the function

$$f : \begin{cases} \mathbb{R}^2 \to \mathbb{R} \\ (x, y) \mapsto x^2 - y^2 + 2 \end{cases}$$

has an extremum at $\boldsymbol{x}_0 = (0, 0)^T$? We investigate the partial derivatives at this point:

$$\left.\frac{\partial f}{\partial x}\right|_{\boldsymbol{x}_0} = 2x|_{x=0} = 0$$
$$\left.\frac{\partial f}{\partial y}\right|_{\boldsymbol{x}_0} = -2y|_{y=0} = 0$$

The necessary condition is given, however, a plot reveals that the function has no extremum at $\boldsymbol{x}_0 = (0, 0)^T$ but a *saddle point*:



Along the x-axis we find a minimum whereas along the y-axis we find a maximum. Therefore, for any environment $U_\varepsilon(\boldsymbol{x}_0)$ we find values smaller and larger than the value at the investigated point. ◁

How do we find out if a point $\boldsymbol{x}_0$ is an extremum and whether it is a maximum or a minimum? A technique we will not focus on is to check whether the Hessian matrix is positive definite or negative definite. However, for most applications it can easily be checked if the function value $f(\boldsymbol{x}_0)$ is less or greater than its local surrounding $U_\varepsilon(\boldsymbol{x}_0)$.

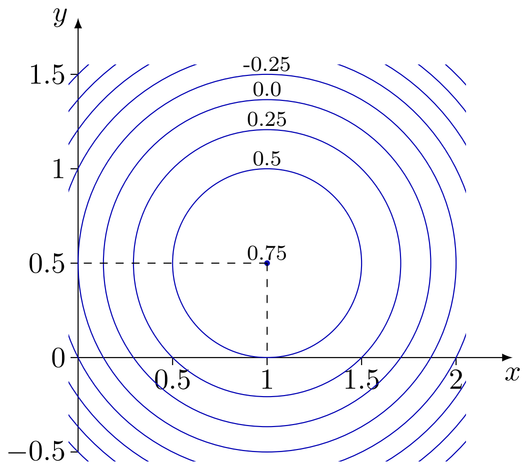
Example 4.19. We want to check the function

$$f : \begin{cases} \mathbb{R}^2 \to \mathbb{R} \\ (x, y) \mapsto 2x - x^2 + y - y^2 - \frac{1}{2} \end{cases}$$

for extrema. As a necessary condition the first partial derivatives must be zero, hence:

$$\frac{\partial f}{\partial x} = 2 - 2x = 0$$
$$\frac{\partial f}{\partial y} = 1 - 2y = 0$$

Both partial derivatives are zero at $\boldsymbol{x}_0 = (1, \frac{1}{2})$. We create a contour plot around $\boldsymbol{x}_0$:



Although a contour plot cannot serve as a proof it becomes obvious that all function values around $\boldsymbol{x}_0$ are less than $f(\boldsymbol{x}_0)$.

Another way to check the candidate $\boldsymbol{x}_0$ is to add small values ε_x and ε_y to the values of $x = 1$ and $y = \frac{1}{2}$, respectively.

$$\begin{aligned} f(1+\varepsilon_x, \tfrac{1}{2}+\varepsilon_y) &= 2(1+\varepsilon_x) - (1+\varepsilon_x)^2 \\ &\quad + (\tfrac{1}{2}+\varepsilon_y) - (\tfrac{1}{2}+\varepsilon_y)^2 - \tfrac{1}{2} \\ &= \ldots = \tfrac{3}{4} - \varepsilon_x^2 - \varepsilon_y^2 \end{aligned}$$

For any combination of ε_x and ε_y except both being zero the function value becomes less than at $\boldsymbol{x}_0$.

Hence, we found a maximum at $\boldsymbol{x}_0 = (1, \frac{1}{2})$ with value $f(\boldsymbol{x}_0) = 0.75$. ◁

4.9. Multiple value functions

Yet we focussed on single valued functions. However, the arguments of a function may influence more than one value.

To differentiate a multiple value function we treat each value of the function separately. For many questions we need the partial derivatives for all values of the function which we combine into a matrix:

Definition 4.10 (Jacobian matrix)**.** Let $f : \mathbb{R}^n \to \mathbb{R}^m$ be a function with n arguments and m values and its component functions $f = (f_1, \dots, f_m)$, $f_k : \mathbb{R}^n \to \mathbb{R}$, $k = 1, \dots, m$. We define the *Jacobian matrix* as the combination of all partial derivatives of the form:

$$J = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \cdots & \frac{\partial f_2}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}$$

◁

Example 4.20. We want to derive the Jacobean matrix for the function

$$f : \begin{cases} \mathbb{R}^2 \to \mathbb{R}^2 \\ (x, y) \mapsto (x \sin(y), y \cos(x)) \end{cases}$$

and get:

$$J = \begin{pmatrix} \sin(y) & x \cos(y) \\ -y \sin(x) & \cos(x) \end{pmatrix}$$

◁

4.10. Problems

Problem 4.1: Plot the following functions for y-values of 0, 1, 2 and 3:

$$f : \begin{cases} (0, 1] \times \mathbb{R} \to \mathbb{R} \\ (x, y) \mapsto x^y \end{cases}$$
$$g : \begin{cases} (0, 1] \times \mathbb{R} \to \mathbb{R} \\ (x, y) \mapsto x^{y/2} \end{cases}$$
$$h : \begin{cases} [0, 2\pi] \times \mathbb{R} \to \mathbb{R} \\ (x, y) \mapsto e^{-y/2} \sin(x) \end{cases}$$
$$i : \begin{cases} [0, 2\pi] \times \mathbb{R} \to \mathbb{R} \\ (x, y) \mapsto \sin(xy/2) \end{cases}$$

Problem 4.2: Create contour plots with at least three contour levels for the following functions:

$$f(x, y) = \sqrt{4 - x^2 - y^2}$$
$$g(x, y) = \sqrt{x^2 + y^2}$$
$$h(x, y) = x^2 - y$$

Problem 4.3: Draw vector plots for the following functions:

$$\boldsymbol{f}(x, y) = (-y, x)$$
$$\boldsymbol{g}(x, y) = (x, y)$$
$$\boldsymbol{h}(x, y) = (x, -y)$$

Problem 4.4: Which of the following functions are continuous in their domain?

1. $f : \begin{cases} \mathbb{R}^2 \to \mathbb{R} \\ (x, y) \mapsto x^2 - y^2 \end{cases}$
2. $f : \begin{cases} \mathbb{R}^2 \to \mathbb{R} \\ (x, y) \mapsto e^x \sin^2(y) \end{cases}$
3. $f : \begin{cases} \mathbb{R} \times \mathbb{R}_{>0} \to \mathbb{R} \\ (x, y) \mapsto \frac{\arctan(\ln(y))}{x^2+1} \end{cases}$
4. $f : \begin{cases} \mathbb{R}^3 \to \mathbb{R} \\ (x, y, z) \mapsto x^2 - y^2 + \arctan(z) \end{cases}$
5. $f : \begin{cases} \mathbb{R}^3 \to \mathbb{R} \\ (x, y, z) \mapsto \frac{\sin^2(x)+\cos^2(y)}{\sin(z)+2} \end{cases}$
6. $f : \begin{cases} \mathbb{R}^4 \to \mathbb{R} \\ (f, t, \delta, \gamma) \mapsto \sin(2\pi f t) e^{-\delta t} + \gamma \end{cases}$

Problem 4.5: Evaluate the first derivatives with respect to all arguments for the following functions:

1. $f(x, y, z) = x + y^2 + z^3$
2. $f(x, y, z) = \sin(xy) + \cos(z^2)$
3. $f(x, y, z) = \dfrac{e^{x+y}}{z^2 + z + 1}$
4. $f(a, b, c, d) = \dfrac{ab}{cd}$
5. $f(u, v, w) = \ln(uw) - e^{vw}$
6. $f(\alpha, \beta, \gamma) = \dfrac{\cos(2\pi\alpha) + \mathrm{j} \sin(2\pi\beta)}{\gamma^2}$

Problem 4.6: Evaluate the first derivatives with respect to all arguments for the following functions:

1. $f(x, y) = \sum_k (x^k + y^k)$

2. $f(x,y) = \sum_k (xy)^k$
3. $f(x,y) = \sum_k e^{-kxy}$
4. $f(a_0, a_1) = \sum_k (a_1 x_k + a_0 - y_k)^2$

Problem 4.7: Evaluate the gradient for the following functions:

1. $f(x,y) = x + y$
2. $f(x,y) = \sqrt{x^2 + y^2}$
3. $f(x,y,z) = x^2 e^y + x^3 \cos(z)$

Problem 4.8: Draw the gradient vector field for the following functions:

1. $f(x,y) = |x + y|$
2. $f(x,y) = \sqrt{x^2 + y^2}$
3. $f(x,y) = \sin(x) + y$

Problem 4.9: Evaluate the Hessian matrix for the following functions:

1. $f(x,y) = x^2 y^2$
2. $f(x,y) = xy$
3. $f(x,y) = \sin(x) + \cos(y)$
4. $f(x,y,z) = x^2 e^y + y^2 \sin(z) + z^2 e^x$

Problem 4.10: Find extrema of the following functions:

1. $f(x,y) = x^2 + y^2 - x - y + \frac{1}{2}$
2. $f(x,y) = e^{-x^2} - y^2 + 2y$
3. $f(x,y) = 2x - x^2 + 4y - y^2$

Problem 4.11: Evaluate the Jacobian matrix for the following functions:

1. $f(x,y) = \begin{pmatrix} x^2 + y^2 \\ xy \end{pmatrix}$
2. $f(x,y) = \begin{pmatrix} \sin(x) + \cos(y) \\ x^2 y^2 \\ e^{xy} \end{pmatrix}$
3. $f(x,y,z) = \begin{pmatrix} x + y^2 + z^3 \\ xe^y \sin(z) \\ x^2 e^{y+z} \end{pmatrix}$

5. Differential equations

5.1. Introduction

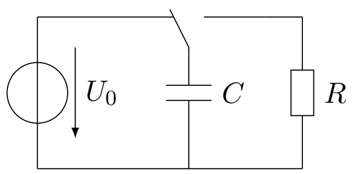
Yet, when solving algebraic equations with unknowns the task was to find values for the unknowns that fulfil the expression. E.g. the equation

$$x^2 - x = 2$$

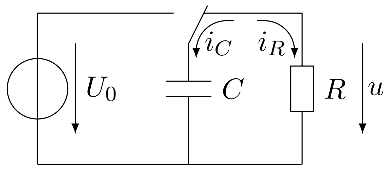
holds for $x = 2$ and $x = -1$.

In this chapter we extend our view on unknowns x to unknown functions $y(x)$.

Example 5.1. Imagine a capacitor with capacitance C that has been charged to a voltage U_0.



At time $t = 0$ the capacitor is switched to a resistor with resistance R. How does the voltage at the capacitor changes over time? I.e. we search for the function $u : \mathbb{R} \to \mathbb{R}, t \mapsto u(t)$.



The current into the resistor is given by:

$$i_R = \frac{u}{R}$$

The current into the capacitor is given by:

$$i_C = C\frac{\mathrm{d}u}{\mathrm{d}t} = Cu'$$

The sum of both currents must be zero, i.e.

$$i_R + i_C = 0$$
$$\frac{u}{R} + Cu' = 0$$

Resolved to u' and $\tau = RC$ we get a differential equation DE:

$$u' = -\tfrac{1}{\tau}u \qquad \text{(DE)}$$

Hence, we are searching for a function with the first derivative being the function itself times a negative factor of $\frac{1}{\tau}$. We find a candidate to be:

$$u = k\,e^{-t/\tau} \qquad u' = -\tfrac{k}{\tau}\,e^{-t/\tau}$$

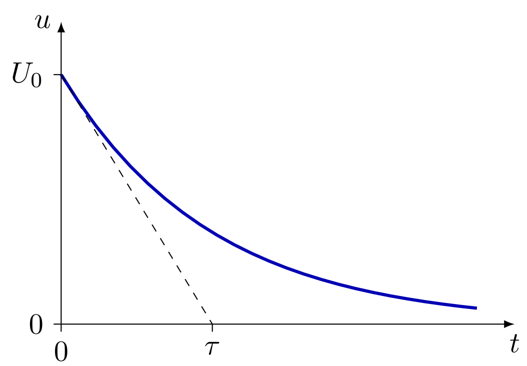
To verify our ansatz we insert u and u' into our DE:

$$-\tfrac{k}{\tau}\,e^{-t/\tau} = -\tfrac{1}{\tau}\,k\,e^{-t/\tau}$$

which is true for all $t \in \mathbb{R}$.

Since the voltage of the capacitor is inert the voltage just after switching to the resistor still is U_0, i.e. $u(0) = U_0$. Hence, $k = U_0$ and we get the solution:

$$u = U_0\,e^{-t/\tau}$$



◁

5.2. Definition of differential equations, DE

In a differential equation (short: DE) the unknown is a function $y(x)$ rather than just a variable x. The DE contains the function $y(x)$ and some derivatives of this function.

Definition 5.1 (Differential equation)**.** For I being an interval a *differential equation, DE* is an equation containing one or more variables $x_1, \ldots, x_p \in I$, one or more functions of these variables $y_1, \ldots, y_q, I^p \to \mathbb{K}$ and derivatives of these functions $y'_k, \ldots, y_k^{(n)}$, $k = 1, \ldots, q$. For one variable and one function we write:

$$y^{(n)} = F(x, y, y', \ldots, y^{(n-1)})$$

We call the order n of the highest occurring derivative the *order of the differential equation*. $\triangleleft$

Example 5.2. Some differential equations:

1. $y'' = y' - y + x$
2. $2\frac{\partial f}{\partial x} - \frac{\partial f}{\partial y} + f^2 + \sin(x) - \cos(y) = 0$
3. $y' + \sqrt{y'} - y + x = 0$
4. $y'' = -\frac{D}{m}y$

$\triangleleft$

Definition 5.2 (Ordinary and partial DE)**.**
A differential equation of a single argument function $f(x)$ is called *ordinary differential equation, ODE*.
A differential equation of a multiple argument function $f(x_1, \ldots, x_n)$ is called a *partial differential equation*. $\triangleleft$

Example 5.3. In the previous example the first, third and fourth DE are ordinary; only the second DE is a partial DE. $\triangleleft$

In this chapter we focus on ordinary differential equations, ODE.

Definition 5.3 (Explicit and implicit notation)**.**
An ODE resolved to its highest derivative is said to be in *explicit notation*, i.e.

$$y^{(n)} = F(x, y, y', \ldots, y^{(n-1)})$$

Otherwise the notation is called *implicit notation* and it is often noted as:

$$F(x, y, y', \ldots, y^{(n)}) = 0$$

$\triangleleft$

Example 5.4. In example 5.2 the first and fourth DE are in explicit notation, whereas the third is in implicit notation. The second is a partial DE where we do not distinguish between explicit and implicit notation. $\triangleleft$

Definition 5.4 (Linear DE)**.** With I being an interval, $a_k, b : I \to \mathbb{R}$, $k = 0, \ldots, n$ being continuous and $y : I \to \mathbb{K}$, the equation

$$\sum_{k=0}^{n} a_k(x) y^{(k)} = b(x), \qquad x \in I$$

or

$$a_n y^{(n)} + \ldots + a_1 y' + a_0 y = b, \qquad x \in I$$

is said to be a *linear differential equation*. $\triangleleft$

Example 5.5. In example 5.2 the first and fourth DE are linear, whereas the second and third are not linear. $\triangleleft$

Definition 5.5 (Homogeneous DE)**.** If a differential equation (DE) can be expressed in the form

$$\sum_{k=0}^{n} a_k(x) y^{(k)}(x) = 0$$

we call it a *homogeneous DE* and an *inhomogeneous DE* otherwise. $\triangleleft$

Remark: A similar definition holds for partial differential equations. There all combinations of derivatives w.r.t. the arguments of the unknown function must be taken into account. However, here the focus is on ODEs only.

Example 5.6. In example 5.2 only the fourth DE is homogeneous; all other DEs are inhomogeneous. $\triangleleft$

5.3. Solutions of DEs

A solution of a DE is a function $y(x)$ for which the DE is true over the investigated interval. However, the solution may include one or more parameters that influence the result.

In integral calculus, when integrating a function we have to add a unknown constant C. Integrating a function n times leads to terms with n unknown constants.

In the same manner the solution of an n^{th} order DE leads to an expression with n unknowns/parameters. The parameters may be resolved by some boundary conditions for the DE.

Definition 5.6 (General and particular solution)**.** We call the solution of an n^{th} order DE with n parameters a *general solution* of the DE.

If one or more of the parameters are set to given values the solution is said to be a *particular solution* of the DE. ◁

Example 5.7. In the introductory example 5.1 we found the solution $u = k\,e^{-t/\tau}$ with the parameter k which we call *general solution* of the DE. We then realized that the voltage at the capacitor at $t = 0$ is U_0 and determined a *particular solution* to be $u = U_0\,e^{-t/\tau}$. ◁

Theorem 5.7 (Number of parameters)**.** An n^{th} order ODE without any further conditions has a general solution with n parameters. ◁

For many applications some conditions are given to further specify the solution of a DE.

For *initial value problems* some initial conditions are given that define some or all of the parameters of the general solution. The introductory example 5.1 belongs to this type of problems.

For *boundary problems* some conditions at the boundary of the domain are given that further specify the solution.

Example 5.8. Radioactive decay.

Radioactive isotopes are unstable and their nuclei change randomly by radioactive processes. E.g. the carbon isotope ^{14}C converts to ^{14}N by beta decay.

$$^{14}_{6}\mathrm{C} \quad \rightarrow \quad ^{14}_{7}\mathrm{N} + \mathrm{e}^- + \bar{\nu}$$

In a given time interval there is a given probability that an atom decays. Experiments show that 50% of a given number of ^{14}C atoms decay in app. 5 730 years. We call this the *half-life* of the isotope since after this time one half of original ^{14}C-isotopes converted into other isotopes/elements.

Let's say N is the number of atoms as a function of time. A change of N means a beta decay including emission of electrons which can be measured. The measured radiation is proportional the number of atoms of ^{14}C (e.g. double mass results in double radiation). We get the differential equation:

$$N' = -\lambda\,N$$

where λ indicates the proportionality. Since the DE is quite similar to the one in the introductory example we choose a similar ansatz:

$$N = c\,e^{-\lambda t} \qquad N' = -\lambda c\,e^{-\lambda t}$$

Inserting into the DE

$$-\lambda c\,e^{-\lambda t} = -\lambda\,c\,e^{-\lambda t}$$

reveals that this is an appropriate solution. λ is related to the half-life $t_{1/2}$:

$$\begin{aligned}
\frac{1}{2} &= \frac{N(t_{1/2})}{N(0)} = \frac{c\,e^{-\lambda t_{1/2}}}{c\,e^{-\lambda \cdot 0}} = e^{-\lambda t_{1/2}} \\
2 &= e^{\lambda t_{1/2}} \\
\lambda &= \frac{\ln(2)}{t_{1/2}}
\end{aligned}$$

Hence, the general solution for the DE is:

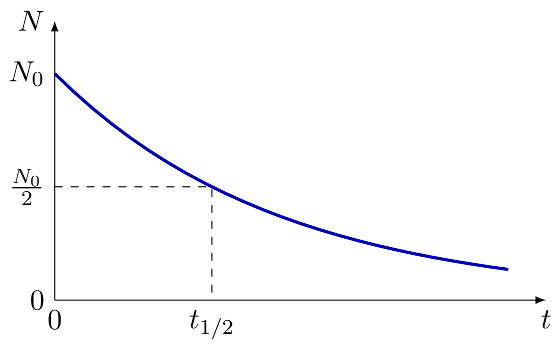
$$N(t) = c\,\exp(-\tfrac{\ln(2)t}{t_{1/2}})$$

Since the DE is of order *one*, the solution has also *one* parameter c. With the condition N_0 as the number of atoms at $t = 0$, i.e.

$$N_0 = N(0) = c\,\exp(-\tfrac{\ln(2)\cdot 0}{t_{1/2}}) = c$$

we get the particular solution:

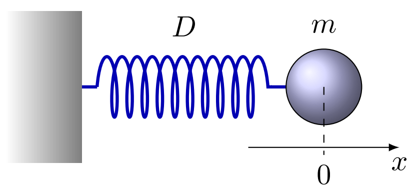
$$N(t) = N_0\,\exp(-\tfrac{\ln(2)t}{t_{1/2}})$$



◁

Example 5.9. Simple harmonic oscillator.

In mechanics a simple harmonic oscillator is made of mass mounted to a spring:



The force acting on the mass is proportional to its displacement. When displaced by distance $x(t)$ the force of the spring acting on the mass is

$$F_s = -Dx(t)$$

with D being the spring constant. In turn, the force of inertia is the negative product of mass m and acceleration $a(t)$ which is the second derivative of the displacement $x(t)$:

$$F_i = -ma(t) = -m\frac{\mathrm{d}^2x(t)}{\mathrm{d}t^2} = -mx''(t)$$

The sum of two the forces must be zero at all times which leads to a second order DE:

$$\begin{aligned} F_s + F_i &= 0 \\ -Dx(t) - mx''(t) &= 0 \\ x'' &= -\tfrac{D}{m}x \end{aligned}$$

As an ansatz we choose:

$$x = \hat{x}\sin(\omega t + \varphi_0)$$

with $\hat{x} \in \mathbb{R}_{>0}$ as the amplitude of the displacement, $\omega \in \mathbb{R}_{>0}$ as the angular frequency and $\varphi_0 \in (-\pi, \pi]$ as the phase at $t = 0$. This is one parameter more than the order of the DE, however, one of them will disappear during further calculations. We evaluate the second derivative

$$\begin{aligned} x' &= \hat{x}\omega\cos(\omega t + \varphi_0) \\ x'' &= -\hat{x}\omega^2\sin(\omega t + \varphi_0) \end{aligned}$$

and insert x and x'' into the DE:

$$\begin{aligned} x'' &= -\tfrac{D}{m}x \\ -\hat{x}\omega^2\sin(\omega t + \varphi_0) &= -\tfrac{D}{m}\hat{x}\sin(\omega t + \varphi_0) \\ \omega^2 &= \tfrac{D}{m} \end{aligned}$$

The chosen ansatz holds only for $\omega^2 = \frac{D}{m}$, i.e. the mass oscillates with an angular frequency:

$$\omega = \sqrt{\frac{D}{m}}$$

Hence we found a general solution with the two remaining parameters $\hat{x}$ and φ_0:

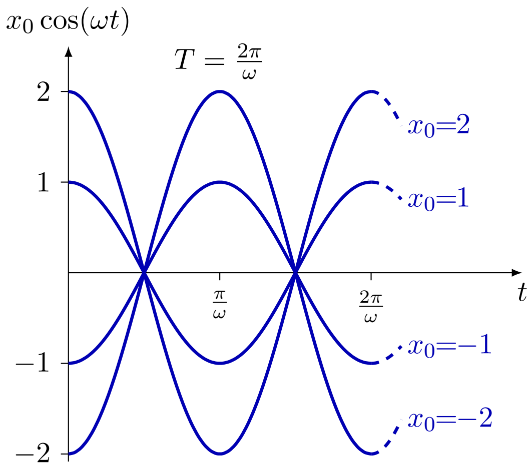
$$x = \hat{x}\sin(\omega t + \varphi_0) \qquad \text{with } \omega = \sqrt{\tfrac{D}{m}}$$

For a particular solution we have to define two independent conditions. E.g. we pull the mass to position x_0 and release it at $t = 0$ with an initial velocity of $v_0 = 0$:

$$\begin{aligned} x(0) &= x_0 \\ v(0) &= x'(0) = v_0 = 0 \end{aligned}$$

We get:

$$x(t) = x_0\sin(\omega t + \tfrac{\pi}{2}) = x_0\cos(\omega t)$$



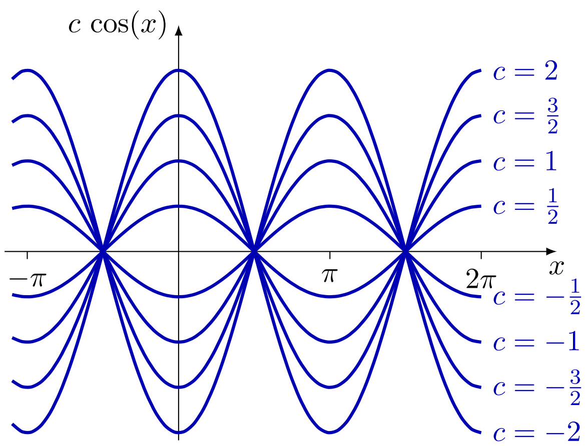
◁

5.4. Creating DEs

We want to define a DE for a given general solution. Although this is somehow the other way round, we want to use this technique to get a deeper understanding on differential equations.

Example 5.10. What is the DE for all cosine functions with angular frequency 1 and amplitude c?

$$y = c\,\cos(x)$$



First we take the first derivative of y:

$$y' = -c\,\sin(x)$$

Then we resolve $y = c\cos(x)$ to c and insert it into the equation for y' in order to eliminate the parameter c:

$$y' = -\frac{y}{\cos(x)}\sin(x) = -y\tan(x)$$
$$y' + y\tan(x) = 0$$

This is a linear homogeneous ODE in implicit notation. (The second last row is the same DE in explicit notation.) ◁

For a given solution with n parameters we create the DE by the following steps:

1. Evaluate n derivatives of the solution resulting in $n+1$ equations.
2. Eliminate the n parameters which results in a single equation without parameters.

Example 5.11. We take the general solution of the harmonic oscillator in problem 5.9 and take the first and second derivative:

$$x = \hat{x}\sin(\omega t + \varphi_0)$$
$$x' = \hat{x}\omega\cos(\omega t + \varphi_0)$$
$$x'' = -\hat{x}\omega^2\sin(\omega t + \varphi_0)$$

The two parameters are $\hat{x}$ and φ_0. We eliminate them by inserting the equation for x into the equation for x'':

$$x'' = -\omega^2 x$$

This is a second order linear homogeneous ODE in explicit notation. ◁

5.5. Problems

Problem 5.1: For I being an interval, $x \in I$ and $y : I \to \mathbb{R}$ which of the following expressions are differential equations?

1. $y'' - y' + xy = 1$
2. $y^2 - y + x = 0$
3. $y^{(2)} - y + x = 0$
4. $\sqrt{y''} + 3xy = \sin(x)$
5. $\sum_{k=0}^{n} a_k y^{(k)} = 0, \quad a_k \in \mathbb{R}$
6. $y^{(n)} + \sum_{k=1}^{n}\left(y^{(k-1)} - y^{(k)}\right) = e^x$

Problem 5.2: Study the following DEs with respect to order of DE, ordinary/partial DE, explicit/implicit notation, linearity and homogeneity:

1. $my'' + ky' + Dy = \sin(x)$
2. $\frac{\partial f}{\partial x} - \frac{\partial f}{\partial y} = 0$
3. $2x + y^2 - \frac{\partial f}{\partial x} + \frac{\partial^2 f}{\partial y^2} = 0$
4. $y'' = \sqrt{x+y}$
5. $\sum_{k=0}^{n-1} a_k y^{(k)} = y^{(n)}$
6. $y' - y^2 + x = 0$

Problem 5.3: Which of the following terms are general solutions and which of them are particular solutions for the given DEs?

1. $y' + ay = 0$

   a) $y = C\,e^{-ax}$ c) $y = 3\,e^{-ax}$
   b) $y = C\,e^{ax}$ d) $y = e^{c-ax}$

2. $y' + y = e^{2x}$

   a) $y = 3e^{-x} + \frac{1}{3}e^{2x}$ c) $y = 4\,e^{-x}$
   b) $y = e^{2x}$ d) $y = Ce^{-x} + \frac{1}{3}e^{2x}$

3. $y' = x\,y^2$

   a) $y = -\frac{2}{x^2}$ c) $y = \frac{2}{C - x^2}$
   b) $y = \frac{C}{x^2}$ d) $y = \frac{2}{C + x^2}$

Problem 5.4: Create a DE for the following functions:

1. $y(x) = c\sin(x)$
2. $y(x) = a\cosh(x)$
3. $y(x) = c\,e^{\omega x}$
4. $y(x) = ax^2$

Problem 5.5: Create a DE for all functions that form straight lines through the origin of a Cartesian coordinate system.

Problem 5.6: Create a DE for all parabolas symmetric around the ordinate and passing through the origin of a Cartesian coordinate system.

Problem 5.7: Create a DE for all circles with radius r around the origin of a Cartesian coordinate system.

6. First order differential equations

6.1. Introduction

In the previous chapter we solved already some DEs. Yet, the ansätze were given by some unclear means.

In this chapter we look at some approaches to find such an ansatz for DEs. We limit ourselves to first order DEs. In the next chapter we extend our view to higher order linear DEs with constant coefficients.

In general we can express a first order differential equation by

$$F(y', y, x) = 0$$

In explicit notation we write

$$y' = F(y, x)$$

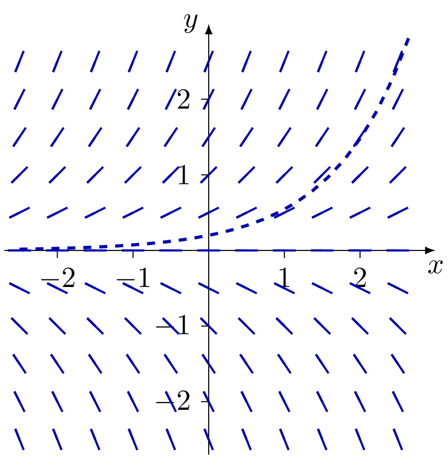
6.2. Geometric interpretation

We write the DE in explicit notation. In this notation we can evaluate the slope (i.e. the first derivative) for any combination of x and y. Plotting the slopes for a number of points gives an impression of the function y.

Example 6.1. We want to solve the DE:

$$y' = y$$

First we plot the slope field and plot a possible solution into it:



The dashed line is an exponential function which seems to fit quite well. Hence we choose as an ansatz:

$$y = c\, e^x \qquad y' = c\, e^x$$

The ansatz solves the DE for any $c \in \mathbb{R}$:

$$y' = y \qquad \Rightarrow \qquad c\, e^x = c\, e^x$$

Hence, $y = c\, e^x$ is the general solution for the DE.

For a particular solution we need a condition like $y(x_0) = y_0$. E.g. for $y(0) = \frac{1}{5}$ we get the particular solution $y_p = \frac{1}{5}\, e^x$ (dashed line in the diagram). ◁

6.3. Separation of variables

If it is possible to express a DE in the form of

$$y' = \frac{\mathrm{d}y}{\mathrm{d}x} = g(x)h(y)$$

we can separate the variables x and y and integrate them separately:

$$\begin{aligned} \frac{1}{h(y)} \frac{\mathrm{d}y}{\mathrm{d}x} &= g(x) \\ \int \frac{1}{h(y)} \frac{\mathrm{d}y}{\mathrm{d}x}\, \mathrm{d}x &= \int g(x)\, \mathrm{d}x \\ \int \frac{1}{h(y)}\, \mathrm{d}y &= \int g(x)\, \mathrm{d}x \end{aligned}$$

Theorem 6.1 (Separation of variables 1)**.** For I being an interval, $g, h : I \to \mathbb{R}$ and $y : I \to \mathbb{K}$, $y \neq 0$ let

$$y' = g(x)h(y)$$

be a DE in explicit notation. Then the general solution can be evaluated by:

$$\int \frac{1}{h(y)}\, \mathrm{d}y = \int g(x)\, \mathrm{d}x$$

◁

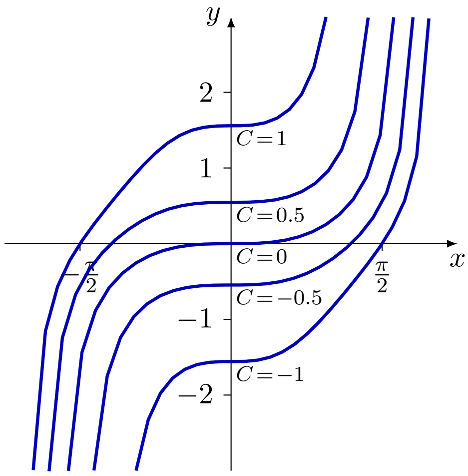
Example 6.2. What is the general solution of

$$y' - x^2 y^2 - x^2 = 0\ ?$$

We bring the equation into explicit notation and separate the variables:

$$\begin{aligned}
y' &= x^2y^2 + x^2 \\
y' &= x^2(y^2+1) \\
\frac{1}{y^2+1}\,\frac{\mathrm{d}y}{\mathrm{d}x} &= x^2 \\
\int \frac{1}{y^2+1}\,\frac{\mathrm{d}y}{\mathrm{d}x}\,\mathrm{d}x &= \int x^2\,\mathrm{d}x \\
\int \frac{1}{y^2+1}\,\mathrm{d}y &= \int x^2\,\mathrm{d}x \\
\arctan(y) &= \tfrac{1}{3}x^3 + C \\
y &= \tan\left(\tfrac{1}{3}x^3 + C\right)
\end{aligned}$$

which is the general solution of the DE.



◁

For a linear homogeneous DE

$$y' = g(x)\,y$$

the function $h(y)$ is the unit function, i.e. $h(y) = y$ and we can further simplify:

$$\begin{aligned}
\frac{1}{y}\,\frac{\mathrm{d}y}{\mathrm{d}x} &= g(x) \\
\int \frac{1}{y}\,\frac{\mathrm{d}y}{\mathrm{d}x}\,\mathrm{d}x &= \int g(x)\,\mathrm{d}x \\
\int \frac{1}{y}\,\mathrm{d}y &= \int g(x)\,\mathrm{d}x \\
\ln|y| + C &= \int g(x)\,\mathrm{d}x \\
y &= \pm\exp\left(\int g(x)\,\mathrm{d}x - C\right) \\
&= \pm e^{-C}\exp\left(\int g(x)\,\mathrm{d}x\right)
\end{aligned}$$

The left factor $\pm e^{-C}$ is the parameter c for the general solution. The constant of the integral in the exponential function may be neglected in this case since it acts like another constant factor outside the exponential function. We get:

Theorem 6.2 (Separation of variables 2)**.** For I being an interval, $g : I \to \mathbb{R}$ and $y : I \to \mathbb{K}$ let

$$y' = g(x)y$$

be a linear homogeneous DE (in explicit notation). We then get the general solution by:

$$y = c\,\exp\left(\int g(x)\,\mathrm{d}x\right)$$

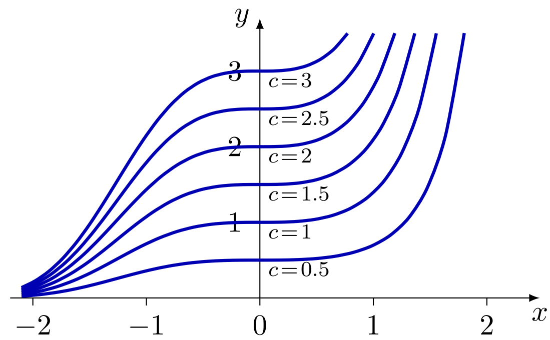
◁

Example 6.3. What is the general solution for

$$y' = x^2y\,?$$

We apply the second theorem for separation of variables:

$$y = c\,\exp\left(\int x^2\,\mathrm{d}x\right) = c\,\exp\left(\tfrac{1}{3}x^3\right) = c\,e^{x^3/3}$$



◁

6.4. Variation of parameters

A *homogeneous* DE is easier to solve than an inhomogeneous one. However, an *inhomogeneous* DE may be solved in two steps:

First we solve the corresponding *homogeneous* DE and get the homogeneous solution y_h.

Second we treat the parameters of the homogeneous solution y_h as functions and and use the modified y_h as ansatz to solve the *inhomogeneous* DE.

For an *inhomogeneous first order DE* we apply the following procedure:

1. Find the general solution y_h of the corresponding homogeneous DE.
2. Treat the parameter c of y_h as a function $c(x)$ (i.e. a *parameter-function*) and evaluate its first derivative.
3. Insert the modified solution and its derivative into the DE and resolve the equation for the parameter-function $c(x)$.

4. Combine the parameter-function $c(x)$ with the general solution y_h of the homogeneous DE as the general solution of the inhomogeneous DE.

Example 6.4. What is the general solution for

$$y' = x + y\,?$$

We first evaluate the general solution for the corresponding homogeneous DE $y' = y$ which we found already in example 6.1:

$$y_h = c\,e^x$$

As an ansatz y for the inhomogeneous DE we treat the parameter c as a function $c(x)$ and evaluate the first derivative of y:

$$y = c(x)\,e^x \qquad y' = c'(x)e^x + c(x)e^x$$

Inserting into the inhomogeneous DE results in:

$$\begin{aligned} y' &= x + y \\ c'(x)e^x + c(x)e^x &= x + c(x)\,e^x \\ c'(x) &= x\,e^{-x} \\ c(x) &= \int x\,e^{-x}\,\mathrm{d}x \end{aligned}$$

We integrate by parts with e^{-x} being the derivative:

$$\begin{aligned} c(x) &= \int x\,e^{-x}\,\mathrm{d}x = -x\,e^{-x} + \int e^{-x}\,\mathrm{d}x \\ &= -x\,e^{-x} - e^{-x} + C = C - e^{-x}(x+1) \end{aligned}$$

Inserting into the homogeneous solution results in

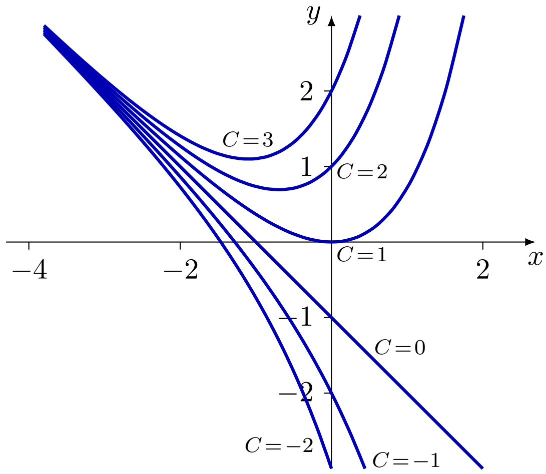
$$\begin{aligned} y &= c(x)\,e^x = (C - e^{-x}(x+1))\,e^x \\ &= C\,e^x - x - 1 \end{aligned}$$

which is the general solution of the inhomogeneous DE.

To remove any doubts we perform a test by inserting the solution and its first derivative into the inhomogeneous DE:

$$\begin{aligned} y &= C\,e^x - x - 1 \\ y' &= C\,e^x - 1 \\ y' &= x + y \\ C\,e^x - 1 &= x + C\,e^x - x - 1 \end{aligned}$$

which is true for any $x \in \mathbb{R}$. Hence, with $y = C\,e^x - x - 1$ we found a general solution for the inhomogeneous DE.



◁

6.5. Linear inhomogeneous DE

Theorem 6.3 (General plus particular solution)**.** Let I be an interval, $a_k, f : I \to \mathbb{R}$ $(k = 0, \dots, n)$ be functions and $y : I \to \mathbb{K}$ be the unknown function. Let further the linear inhomogeneous DE

$$\sum_{k=0}^{n} a_k(x) y^{(k)}(x) = f(x), \qquad x \in I$$

has a particular solution $y(x) = y_p(x)$ and the corresponding homogeneous DE

$$\sum_{k=0}^{n} a_k(x) y^{(k)}(x) = 0, \qquad x \in I$$

has the general solution $y(x) = y_h(x)$. Then the sum the two solutions $y_h(x) + y_p(x)$ is the general solution of the inhomogeneous DE:

$$y(x) = y_h(x) + y_p(x)$$

◁

Proof. We insert the sum of the two solutions into the DE:

$$\begin{aligned} \sum_k a_k (y_h + y_p)^{(k)} &= \sum_k a_k \left(y_h^{(k)} + y_p^{(k)} \right) \\ &= \sum_k \left(a_k y_h^{(k)} + a_k y_p^{(k)} \right) \\ &= \sum_k a_k y_h^{(k)} + \sum_k a_k y_p^{(k)} \\ &= 0 + f(x) = f(x) \end{aligned}$$

□

Example 6.5. We want to solve the DE:

$$y' + xy + x = 0$$

First we solve the corresponding homogeneous DE by separation of variables:

$$\begin{aligned} y' + xy &= 0 \\ y' &= -xy \\ y_h &= c \exp\left(\int -x \,\mathrm{d}x\right) = c\, e^{-x^2/2} \end{aligned}$$

Next we solve the inhomogeneous DE by variation of parameters:

$$\begin{aligned} y &= c(x)\, e^{-x^2/2} \\ y' &= c'(x)\, e^{-x^2/2} - c(x) x\, e^{-x^2/2} \end{aligned}$$

Inserting into the DE:

$$\begin{aligned} c'\, e^{-x^2/2} - c\, x\, e^{-x^2/2} + x\, c\, e^{-x^2/2} + x &= 0 \\ c'\, e^{-x^2/2} + x &= 0 \\ c' &= -x\, e^{x^2/2} \\ c &= -\int x\, e^{x^2/2} \,\mathrm{d}x \end{aligned}$$

We substitute by $u = x^2/2$. Since we search for a particular solution only we neglect the integration constant:

$$u = \frac{x^2}{2} \qquad u' = x \qquad \mathrm{d}x = \frac{\mathrm{d}u}{x}$$

$$c = -\int x\, e^u \frac{\mathrm{d}u}{x} = -\int e^u \,\mathrm{d}u = -e^u = -e^{x^2/2}$$

$$y_p = c\, e^{-x^2/2} = -e^{x^2/2}\, e^{-x^2/2} = -1$$

Inserting y_p into the DE reveals that -1 indeed is a particular solution of the DE.

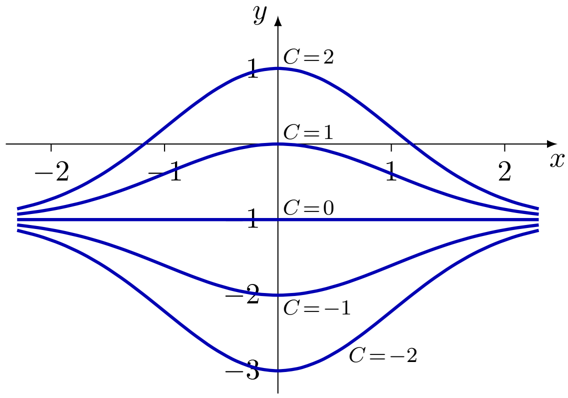
To get the general solution of the inhomogeneous DE we sum up the general solution of the homogeneous DE and the particular solution of the inhomogeneous DE:

$$y = y_h + y_p = c\, e^{-x^2/2} - 1$$

We perform a test:

$$\begin{aligned} y = c\, e^{-x^2/2} - 1 \qquad y' &= -cx\, e^{-x^2/2} \\ y' + xy + x &= 0 \\ -cx\, e^{-x^2/2} + x\,(c\, e^{-x^2/2} - 1) + x &= 0 \\ -cx\, e^{-x^2/2} + xc\, e^{-x^2/2} - x + x &= 0 \end{aligned}$$

which is true for all $x, c \in \mathbb{R}$.



◁

6.6. Summary

In this chapter we looked at some techniques to find solutions for first order DEs.

1. In explicit notation we gained a visual impression by a slope diagram.
2. For DEs of type $y' = g(x)h(y)$ we performed separation of variables.
3. For DEs of type $y' = g(x)y$ we found a special form of separation of variables.
4. For inhomogeneous DEs we performed the variation of parameters.
5. Finally we found the solution of a linear inhomogeneous DE to be $y(x) = y_h(x) + y_p(x)$.

6.7. Problems

Problem 6.1: Draw the slope field for the following DEs:

1. $y'y - 1 = 0$
2. $y'x - x = 0$
3. $y'y + x = 0$

Problem 6.2: Solve the following DEs by separation of variables:

1. $yy' - 1 = 0$
2. $yy' - x = 0$
3. $xy' - y = 0$
4. $e^y y' - 1 = 0$
5. $y' + \dfrac{\sin(2x)}{y} = 0$

6. $y' + 2y^2 = 0$

Problem 6.3: Solve the following DEs by separation of variables:

1. $y' + y = x^2 y$
2. $xy' + 2x^2 y = y$
3. $y' + 3y\sqrt{x} = 0$
4. $y' + y\sin(x) = 0$
5. $\dfrac{y'}{\cosh(x)} - y = 0$
6. $2xy' - y = 0$

Problem 6.4: Solve the following inhomogeneous DEs by variation of parameters:

1. $xy' - y = 1$
2. $xy' + y = e^x$
3. $xy' + 2y = \sin(x)$

Problem 6.5: Solve by adding the general homogeneous solution and a particular inhomogeneous solution:

1. $xy' + y = 1$
2. $y' + 2y = \cos(x)$
3. $y' - y = e^{\mathrm{j}x}$

Problem 6.6: For $k \in \mathbb{R}$, $n \in \mathbb{Z}$, $n \neq -1$ solve the following DEs for $y(x)$:

1. $y' - 3y = 0$
2. $y' + 3y = 0$
3. $y' + x^3 y = 0$
4. $x^2 y' - 2y = 0$
5. $y' - ky = 0$
6. $y' - kx^n y = 0$

Problem 6.7: Solve the following DEs:

1. $y' - y + x = 0$
2. $y' + 2y = x^2$
3. $x^2 y' + y - 1 = 0$
4. $\dfrac{y+1}{y'} = x$
5. $y' + y = x$
6. $\dfrac{y - 2x}{y'} = 1$

7. Higher order linear differential equations

7.1. Introduction

With I being an interval and $a_k, f : I \to \mathbb{R}$ ($k = 0, \ldots, n$) a higher order ($n > 1$) linear DE may be written as

$$\sum_{k=0}^{n} a_k(x) y^{(k)}(x) = f(x)$$

We are particularly interested in linear DEs with constant real coefficients, i.e.

$$\sum_{k=0}^{n} a_k y^{(k)}(x) = f(x), \qquad a_k \in \mathbb{R}$$

Many engineering problems result in DEs of this type.

We call $f(x)$ the *source term* of the inhomogeneous DE. For homogeneous DEs the source term is zero.

With $D = \frac{\mathrm{d}}{\mathrm{d}x}$ as the differential operator we write in short:

$$\begin{aligned} \sum_{k=0}^{n} a_k D^k y &= f \\ \left(\sum_{k=0}^{n} a_k D^k \right) y &= f \\ (a_n D^n + \ldots + a_2 D^2 + a_1 D + a_0) y &= f \end{aligned}$$

In the brackets we find an n^{th} order polynomial with coefficients a_k, $k = 0, \ldots, n$ which we abreviate with $P_n(D)$:

$$P_n(D)\, y = f$$

This is the inhomogeneous linear DE with constant real coefficients. The corresponding homogeneous DE is given by:

$$P_n(D)\, y = 0$$

Theorem 7.1 (Linear combination of solutions)**.** Let y_1 and y_2 be two different solutions of a linear homogeneous DE of order n, i.e.

$$P_n(D) y_1 = 0 \qquad \text{and} \qquad P_n(D) y_2 = 0$$

Then any linear combination of the solutions

$$y = C_1\, y_1 + C_2\, y_2, \qquad C_1, C_2 = \text{const.}$$

is also a solution of the DE, i.e.

$$P_n(D) y = P_n(D)(C_1 y_1 + C_2 y_2) = 0$$

◁

Proof.

$$\begin{aligned} P_n(D) y &= P_n(D) \left(C_1 y_1 + C_2 y_2 \right) \\ &= \sum_{k=0}^{n} a_k (C_1 y_1 + C_2 y_2)^{(k)} \\ &= \sum_{k=0}^{n} a_k (C_1 y_1^{(k)} + C_2 y_2^{(k)}) \\ &= C_1 \sum_{k=0}^{n} a_k y_1^{(k)} + C_2 \sum_{k=0}^{n} a_k y_2^{(k)} \\ &= C_1 P_n(D) y_1 + C_2 P_n(D) y_2 \\ &= C_1 \cdot 0 + C_2 \cdot 0 \\ &= 0 \end{aligned}$$

□

Theorem 7.2 (Complex solution)**.** Let I be an interval, $u, v : I \to \mathbb{R}$ and $y = u + \mathrm{j}v$ be a complex solution of a linear homogeneous DE of order n, i.e.

$$P_n(D) y = P_n(D)(u + \mathrm{j}v) = 0$$

Then the real and imaginary parts are also solutions of the DE, i.e.

$$P_n(D) u = 0 \qquad \text{and} \qquad P_n(D) v = 0$$

◁

Proof.

$$\begin{aligned} P_n(D) y &= P_n(D)(u + \mathrm{j}v) \\ &= P_n(D) u + \mathrm{j} P_n(D) v = 0 \end{aligned}$$

Both, $P_n(D)u$ and $P_n(D)v$ are real. The second term is multiplied by the imaginary unit j becoming a pure imaginary term. A complex number is zero if the real and the imaginary part are zero, hence

$$P_n(D) u = 0 \qquad \text{and} \qquad P_n(D) v = 0$$

□

7.2. Second order homogeneous DE with constant coefficients

It is always possible to bring the DE into this form:

$$y'' + a_1 y' + a_0 y = 0$$

To solve an DE of this type the ansatz is:

$$y(x) = e^{kx}, \qquad k \in \mathbb{K}$$

We take the first and second derivative and insert it into the DE:

$$y = e^{kx} \qquad y' = ke^{kx} \qquad y'' = k^2 e^{kx}$$

$$\begin{aligned} y'' + a_1 y' + a_0 y &= k^2 e^{kx} + a_1 k e^{kx} + a_0 e^{kx} \\ &= e^{kx}(k^2 + a_1 k + a_0) = 0 \end{aligned}$$

We call the second factor the *characteristic polynomial*:

$$p(k) = k^2 + a_1 k + a_0$$

The term e^{kx} is never zero, hence, the characteristic polynomial must be zero:

$$p(k) = k^2 + a_1 k + a_0 = 0$$

We call this the *characteristic equation* of the DE. To find k we solve this second order polynomial and get:

$$k_{1,2} = -\frac{a_1}{2} \pm \sqrt{\frac{{a_1}^2}{4} - a_0} \tag{7.1}$$

For real coefficients this equation can have three different types of solutions:

1. Two different real constants
2. A pair of complex conjugate constants
3. Two equal real constants

We have to treat these three types separately:

7.2.1. Two different real constants

The argument of the root in (7.1) is positive:

$$\frac{{a_1}^2}{4} - a_0 > 0 \qquad \Leftrightarrow \qquad {a_1}^2 > 4a_0$$

In this case we found two independent solutions for the DE, i.e.

$$y_1 = e^{k_1 x} \qquad \text{and} \qquad y_2 = e^{k_2 x}$$

The general solution is any linear combination of these solutions:

$$y = C_1 e^{k_1 x} + C_2 e^{k_2 x} \tag{7.2}$$

$$\text{with} \qquad k_{1,2} = -\frac{a_1}{2} \pm \sqrt{\frac{{a_1}^2}{4} - a_0}$$

7.2.2. Pair of complex conjugate constants

The argument of the root in (7.1) is negative:

$$\frac{{a_1}^2}{4} - a_0 < 0 \qquad \Leftrightarrow \qquad {a_1}^2 < 4a_0$$

For this important case in engineering we may write:

$$k_{1,2} = -\frac{a_1}{2} \pm \sqrt{\frac{{a_1}^2}{4} - a_0} = -\sigma \pm \mathrm{j}\omega$$

with $\sigma = \frac{a_1}{2}$ and $\omega = \sqrt{a_0 - \frac{{a_1}^2}{4}}$. The general solution is:

$$\begin{aligned} y &= C_1\, e^{k_1 x} + C_2\, e^{k_2 x} \\ &= C_1\, e^{(-\sigma + \mathrm{j}\omega)x} + C_2\, e^{(-\sigma - \mathrm{j}\omega)x} \\ &= e^{-\sigma x}(C_1\, e^{\mathrm{j}\omega x} + C_2\, e^{-\mathrm{j}\omega x}) \end{aligned}$$

With Euler's formula $e^{\mathrm{j}\omega x} = \cos(\omega x) + \mathrm{j}\sin(\omega x)$ we get

$$\begin{aligned} y &= e^{-\sigma x}\Big(C_1\,\big(\cos(\omega x) + \mathrm{j}\sin(\omega x)\big) \\ &\qquad + C_2\,\big(\cos(\omega x) - \mathrm{j}\sin(\omega x)\big)\Big) \\ &= e^{-\sigma x}\big((C_1 + C_2)\,\cos(\omega x) \\ &\qquad + (C_1 - C_2)\,\mathrm{j}\sin(\omega x)\big) \\ &= e^{-\sigma x}\big(A\,\cos(\omega x) + \mathrm{j}B\,\sin(\omega x)\big) \end{aligned}$$

where $A = C_1 + C_2$ and $B = C_1 - C_2$. Since for a complex solution the real and imaginary part are both solutions of the DE we may write:

$$y = e^{-\sigma x}\big(A\,\cos(\omega x) + B\,\sin(\omega x)\big) \tag{7.3}$$

$$\text{with} \quad \sigma = \frac{a_1}{2} \quad \text{and} \quad \omega = \sqrt{a_0 - \frac{{a_1}^2}{4}}$$

7.2.3. Two equal real constants

The argument of the root in (7.1) is zero:

$$\frac{{a_1}^2}{4} - a_0 = 0 \qquad \Leftrightarrow \qquad {a_1}^2 = 4a_0$$

For the constants k_1 and k_2 we get:

$$k_1 = k_2 = -\sigma = -\frac{a_1}{2}$$

Applying the same technique as before would result in the solution $C\,e^{-\sigma x}$ with only one parameter. This is not the general solution since we expect for a second order DE a general solution with two parameters.

In order to find a general solution we apply the variation of parameters:

$$\begin{aligned} y &= Ce^{-\sigma x} \\ y' &= C'e^{-\sigma x} - C\sigma\,e^{-\sigma x} = (C' - \sigma C)e^{-\sigma x} \\ y'' &= (C'' - \sigma C')e^{-\sigma x} - (C' - \sigma C)\sigma e^{-\sigma x} \\ &= (C'' - 2\sigma C' + \sigma^2 C)e^{-\sigma x} \end{aligned}$$

Inserting into the DE gives:

$$\begin{aligned} y'' + a_1 y' + a_0 y &= 0 \\ (C'' - 2\sigma C' + \sigma^2 C)e^{-\sigma x} + & \\ a_1(C' - \sigma C)e^{-\sigma x} + a_0 C e^{-\sigma x} &= 0 \\ C'' + (a_1 - 2\sigma)C' + (\sigma^2 - a_1\sigma + a_0)C &= 0 \end{aligned}$$

With $\sigma = a_1/2$ and ${a_1}^2 = 4a_0$ the two brackets become zero and we get

$$C'' = 0$$

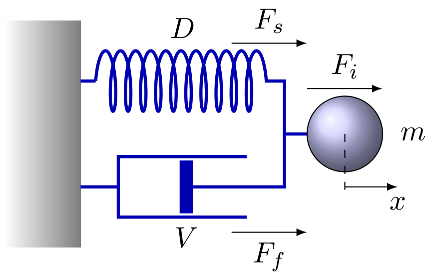
Integrating twice results in

$$C(x) = C_1 x + C_2$$

and we get the general solution

$$y = (C_1 x + C_2)e^{-\sigma x} \quad \text{with} \quad \sigma = \frac{a_1}{2} \qquad (7.4)$$

Example 7.1. Damped harmonic oscillator. We want to analyse a *damped harmonic oscillator* with spring, mass and damper.



The damper adds the force of friction F_f. With the *spring constant* D in N/m, the *viscous damping coefficient* V in Ns/m and the mass m in $\frac{\mathrm{Ns}^2}{\mathrm{m}}$=kg we get three forces which in sum must be zero. With $v = \frac{\mathrm{d}x}{\mathrm{d}t} = x'$ as the velocity and $a = \frac{\mathrm{d}^2x}{\mathrm{d}t^2} = x''$ as the acceleration we get:

$$\begin{aligned} F_i + F_f + F_s &= 0 \\ -ma - Vv - Dx &= 0 \\ -mx'' - Vx' - Dx &= 0 \\ x'' + \tfrac{V}{m}x' + \tfrac{D}{m}x &= 0 \end{aligned}$$

This is a second order linear homogeneous DE with constant real coefficients. For the coefficients k_1 and k_2 we get:

$$\text{char. eq.:} \qquad k^2 + \frac{V}{m}k + \frac{D}{m} = 0$$

$$k_{1,2} = -\frac{V}{2m} \pm \sqrt{\frac{V^2}{4m^2} - \frac{D}{m}}$$

Now we distinguish the three possible solutions:

Two different real coefficients

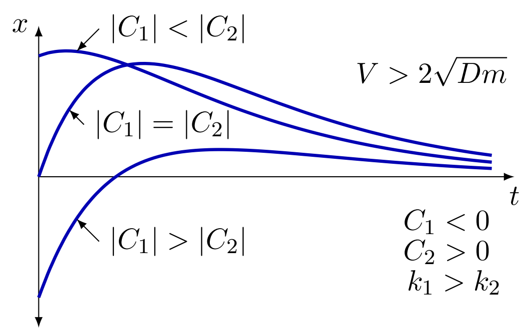
If the viscous damping coefficient V is comparable large, i.e.

$$\frac{V^2}{4m^2} > \frac{D}{m} \quad \Leftrightarrow \quad V > 2\sqrt{Dm}$$

then k_1 and k_2 are two different real numbers. We get

$$x(t) = C_1 e^{k_1 t} + C_2 e^{k_2 t}$$

with C_1 and C_2 being two real numbers.



Pair of complex conjugate constants

If the viscous damping coefficient V is comparable small, i.e.

$$\frac{V^2}{4m^2} < \frac{D}{m} \quad \Leftrightarrow \quad V < 2\sqrt{Dm}$$

then k_1 and k_2 are a pair of complex conjugate numbers. With

$$\sigma = \frac{V}{2m} \qquad \text{and} \qquad \omega = \sqrt{\frac{D}{m} - \frac{V^2}{4m^2}}$$

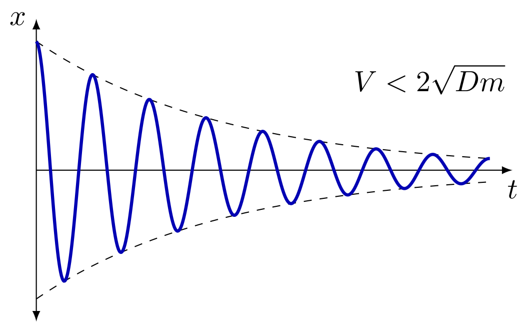
we get

$$x(t) = e^{-\sigma t}(C_1\, e^{j\omega t} + C_2\, e^{-j\omega t})$$

or if we limit ourself to real functions:

$$x(t) = e^{-\sigma t}\{A\cos(\omega t) + B\sin(\omega t)\}$$

with $C_1, C_2, A, B \in \mathbb{R}$.



Two equal real coefficients

If the viscous damping coefficient is $2\sqrt{Dm}$, i.e.

$$\frac{V^2}{4m^2} = \frac{D}{m} \quad \Leftrightarrow \quad V = 2\sqrt{Dm}$$

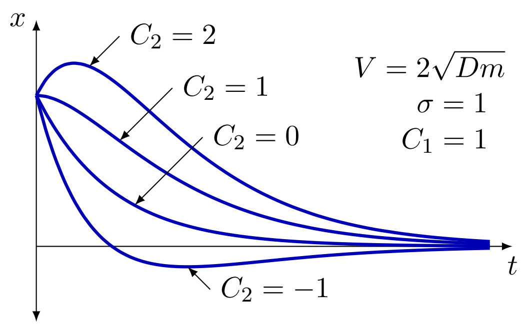
then k_1 and k_2 are two equal real numbers:

$$k_1 = k_2 = -\sigma = -\frac{V}{2m}$$

and we get

$$x(t) = C_1\, e^{-\sigma t} + C_2 t\, e^{-\sigma t} = (C_1 + C_2 t)\, e^{-\sigma t}$$

with C_1 and C_2 being real numbers.



◁

7.2.4. Summary

We solve a second order linear homogeneous DE with constant real coefficients

$$y'' + a_1 y' + a_0 y = 0$$

with the ansatz $y = e^{kx}$ which leads us to the characteristic equation $k^2 + a_1 k + a_0 = 0$ which we solve to:

$$k_{1,2} = -\frac{a_1}{2} \pm \sqrt{\frac{{a_1}^2}{4} - a_0}$$

Depending on the argument of the root we distinguish three types of solutions. With $\sigma = \frac{a_1}{2}$, $\omega = \sqrt{a_0 - \frac{{a_1}^2}{4}}$ and $C_1, C_2, A, B \in \mathbb{R}$ we get:

case	solution
${a_1}^2 > 4a_0$	$y(x) = C_1 e^{k_1 x} + C_2 e^{k_2 x}$
${a_1}^2 < 4a_0$	$y(x) = e^{-\sigma x}(C_1 e^{j\omega x} + C_2 e^{-j\omega x})$
	$y(x) = e^{-\sigma x}(A\cos(\omega x) + B\sin(\omega x))$
${a_1}^2 = 4a_0$	$y(x) = (C_1 + C_2 x)e^{-\sigma x}$

For the case ${a_1}^2 < 4a_0$ the first solution is a complex valued function, i.e. $y : I \to \mathbb{C}$.

7.3. Higher order homogeneous DE with constant coefficients

We apply the same approach as for a second order DE. For an n^{th} order linear homogeneous DE with constant real coefficients we write

$$y^{(n)} + a_{n-1} y^{(n-1)} + \ldots + a_1 y' + a_0 y = 0$$

With the ansatz $y = e^{kx}$ we get an n^{th} order *characteristic polynomial*:

$$p(k) = k^n + a_{n-1}k^{n-1} + \ldots + a_1 k + a_0$$

Setting $p(k)$ to zero gives the *characteristic equation*:

$$p(k) = k^n + a_{n-1}k^{n-1} + \ldots + a_1 k + a_0 = 0$$

We now have to find the roots of the characteristic polynomial, i.e. all possible solutions of the characteristic equation. We know by the fundamental theorem of algebra that the equation must have n zeros. For real coefficients we will find real zeros or pairs of complex conjugate zeros.

We now have to distinguish the different types of solutions for k:

1. For n different real zeros $k_1, k_2, \ldots, k_n$ the general solution is

$$y(x) = C_1 e^{k_1 x} + C_2 e^{k_2 x} + \ldots + C_n e^{k_n x}$$

2. For a pair of complex conjugate zeros

$$k_k = -\sigma + \mathrm{j}\omega \qquad \text{and} \qquad k_{k+1} = -\sigma - \mathrm{j}\omega$$

we combine them to the term

$$e^{-\sigma x}\{C_k \cos(\omega x) + C_{k+1}\sin(\omega x)\}$$

3. For multiple zeros, e.g. an m-fold zero k_k we use the term

$$(C_k + C_{k+1}x + \ldots + C_{k+m-1}x^{m-1})e^{k_k x}$$

Example 7.2. The DE

$$y''' - 2y'' - y' + 2y = 0$$

has the characteristic equation

$$k^3 - 2k^2 - k + 2 = 0$$

with the zeros

$$k_1 = -1, \qquad k_2 = 1, \qquad k_3 = 2$$

and the solution

$$y(x) = C_1\, e^{-x} + C_2\, e^{x} + C_3\, e^{2x}$$

◁

Example 7.3. We want to solve the DE

$$4y^{(4)} + 12y''' + 17y'' + 14y' + 5y = 0$$

Dividing by 4 gives:

$$y^{(4)} + 3y''' + \frac{17}{4}y'' + \frac{7}{2}y' + \frac{5}{4}y = 0$$

The characteristic equation

$$k^4 + 3k^3 + \frac{17}{4}k^2 + \frac{7}{2}k + \frac{5}{4} = 0$$

has zeros at:

$$k_{1,2} = -1, \qquad k_{3,4} = -\frac{1}{2} \pm \mathrm{j}$$

We have a double zero at -1 and a pair of complex conjugate zeros $-\frac{1}{2} \pm \mathrm{j}$ which we have to treat separately. We get

$$y(x) = (C_1 + C_2 x)e^{-x} + e^{-x/2}\{C_3\cos(x) + C_4\sin(x)\}$$

◁

7.4. Higher order inhomogeneous DE with constant coefficients

We write the inhomogeneous linear DE with constant real coefficients as

$$y^{(n)} + a_{n-1}y^{(n-1)} + \ldots + a_1 y' + a_0 y = s(x)$$

where we call $s(x)$ the *source term*.

In the previous chapter we learned: For an inhomogeneous linear DE the sum of a particular solution y_p and the general solution y_h of the corresponding homogeneous DE is the general solution of the investigated DE:

$$y(x) = y_h(x) + y_p(x)$$

This theorem holds for any order of the DE. Hence, we use three steps to solve an inhomogeneous linear DE:

1. Find the general solution y_h of the corresponding homogeneous DE.
2. Find any particular solution y_p of the inhomogeneous DE.
3. Take the sum of the two solutions as the general solution of the inhomogeneous DE.

The solution y_h of the corresponding homogeneous DE may be evaluated as described in the previous section.

The ansatz to find a particular solution y_p of the inhomogeneous DE depends on the source term $s(x)$ and there is no general technique to find an appropriate ansatz. Hence, we concentrate on some typical source terms.

7.4.1. Polynomial source term

If the source term $s(x)$ is an m^{th} order polynomial, i.e.

$$s(x) = S_0 + S_1 x + S_2 x^2 + \ldots + S_m x^m$$

we choose as ansatz for the particular solution y_p an m^{th} order polynomial too:

$$y_p(x) = s_0 + s_1 x + s_2 x^2 + \ldots + s_m x^m$$

We insert this polynomial into the DE to evaluate the coefficients $s_0, \ldots, s_m$. In the ansatz all coefficients must be present, even if the source term does not include all coefficients $S_0, \ldots, S_m$.

Example 7.4. We want to solve the DE

$$y'' + 3y' + 2y = 2x - 1$$

The corresponding homogeneous DE has the characteristic equation $k^2 + 3k + 2 = 0$ with the two zeros at -1 and -2, which leads us to the general homogeneous solution:

$$y_h = C_1\, e^{-x} + C_2\, e^{-2x}$$

Since the source term of the inhomogeneous DE is a first order polynomial we choose as ansatz a first order polynomial too. We insert the ansatz with its derivatives into the DE:

$$y_p = s_0 + s_1 x \qquad y_p' = s_1 \qquad y_p'' = 0$$
$$0 + 3s_1 + 2(s_0 + s_1 x) = 2x - 1$$
$$2s_1 x + (2s_0 + 3s_1) = 2x - 1$$

Comparing the coefficients of the two polynomials results in two equations which we resolve for s_1 and s_0:

$$\begin{aligned} 2s_1 &= 2 \\ 2s_0 + 3s_1 &= -1 \end{aligned} \qquad \Rightarrow \qquad \begin{aligned} s_1 &= 1 \\ s_0 &= -2 \end{aligned}$$

which gives us the particular solution

$$y_p = x - 2$$

The sum of the general homogeneous solution and the particular inhomogeneous solution gives the general inhomogeneous solution:

$$y(x) = C_1\, e^{-x} + C_2\, e^{-2x} + x - 2$$

◁

If the chosen polynomial leads to a contradiction, we increase the order of the polynomial.

Example 7.5. We want to solve the DE

$$y'' + 2y' = x$$

The corresponding homogeneous DE has the characteristic equation $k^2 + 2k = 0$ with the two zeros at -2 and 0, which leads to the general homogeneous solution:

$$y_h = C_1\, e^{-2x} + C_2$$

Since the source term of the inhomogeneous DE is a first order polynomial we choose as ansatz a first order polynomial too. We insert the ansatz with its derivatives into the DE:

$$y_p = s_0 + s_1 x \qquad y_p' = s_1 \qquad y_p'' = 0$$

$$0 + 2s_1 = x$$

This expression cannot be true for all x, hence we choose an ansatz with increased order:

$$\begin{aligned} y_p &= s_0 + s_1 x + s_2 x^2 \\ y_p' &= s_1 + 2s_2 x \\ y_p'' &= 2s_2 \\ 2s_2 + 2(s_1 + 2s_2 x) &= x \\ 4s_2 x + (2s_2 + 2s_1) &= x \end{aligned}$$

Comparing the coefficients of the two polynomials results in two equations which we resolve for s_1 and s_2:

$$\begin{aligned} 4s_2 &= 1 \\ 2s_1 + 2s_2 &= 0 \end{aligned} \qquad \Rightarrow \qquad \begin{aligned} s_2 &= \tfrac{1}{4} \\ s_1 &= -\tfrac{1}{4} \end{aligned}$$

We get a particular solution:

$$y_p = \tfrac{1}{4}x^2 - \tfrac{1}{4}x + s_0$$

The particular solution is valid for any value of s_0. For convenience we set s_0 to zero and get:

$$y_p = \tfrac{1}{4}x^2 - \tfrac{1}{4}x$$

The sum of the general homogeneous solution and the particular inhomogeneous solution gives the general inhomogeneous solution:

$$y(x) = y_h + y_p = C_1\, e^{-2x} + C_2 + \tfrac{1}{4}x^2 - \tfrac{1}{4}x$$

◁

7.4.2. Exponential source term

If the source term is an exponential function, i.e.

$$s(x) = A\, e^{bx}$$

then we choose as ansatz for the particular solution y_p an exponential function with the same exponent, i.e.

$$y_p = a\, e^{bx}$$

The ansatz equals the source term except for another factor a. We insert the ansatz into the DE to derive the factor a.

Example 7.6. We want to solve the DE

$$y'' - y = 4\, e^{-3x}$$

The corresponding homogeneous DE has the characteristic equation $k^2 - 1 = 0$ with the two zeros at ± 1 which leads to the general homogeneous solution:

$$y_h = C_1\, e^{-x} + C_2\, e^{x}$$

We choose $a\, e^{-3x}$ as ansatz for the particular solution of the inhomogeneous DE and take two derivatives:

$$y_p = a\, e^{-3x} \qquad y_p' = -3a\, e^{-3x} \qquad y_p'' = 9a\, e^{-3x}$$

Inserting into the DE gives:

$$\begin{aligned} y'' - y &= 4\, e^{-3x} \\ 9a\, e^{-3x} - a\, e^{-3x} &= 4\, e^{-3x} \\ 9a - a &= 4 \\ a &= \frac{1}{2} \\ y_p &= \frac{1}{2}\, e^{-3x} \end{aligned}$$

This leads us to the general solution of the inhomogeneous DE:

$$y = y_h + y_p = C_1\, e^{-x} + C_2\, e^{x} + \frac{1}{2}\, e^{-3x}$$

◁

7.4.3. Trigonometric source term

If the source term is a trigonometric function, i.e.

$$s(x) = A\cos(\omega x) + B\sin(\omega x)$$

then we choose as ansatz for the particular solution y_p:

$$y_p = a\cos(\omega x) + b\sin(\omega x)$$

The sine and cosine functions must be both present, even if only one of them is present as source term. We evaluate the constants a and b by inserting the ansatz into the DE.

Example 7.7. We want to solve the DE

$$y'' + 3y' + 2y = 5\cos(2x)$$

The corresponding homogeneous DE has the characteristic equation $k^2 + 3k + 2 = 0$ with the two zeros at -1 and -2 which leads us to the general homogeneous solution:

$$y_h = C_1\, e^{-x} + C_2\, e^{-2x}$$

We choose $a\cos(2x) + b\sin(2x)$ as ansatz for the particular solution of the inhomogeneous DE and take two derivatives:

$$\begin{aligned} y_p &= a\cos(2x) + b\sin(2x) \\ y_p' &= -2a\sin(2x) + 2b\cos(2x) \\ y_p'' &= -4a\cos(2x) - 4b\sin(2x) \end{aligned}$$

Inserting into the DE gives:

$$\begin{aligned} y'' + 3y' + 2y &= 5\cos(2x) \\ -4a\cos(2x) - 4b\sin(2x) & \\ +3(-2a\sin(2x) + 2b\cos(2x)) & \\ +2(a\cos(2x) + b\sin(2x)) &= 5\cos(2x) \\ (6b - 2a)\cos(2x) & \\ +(-6a - 2b)\sin(2x) &= 5\cos(2x) \end{aligned}$$

This equation is true if it holds separately for the sine and cosine terms. We get two equations for two unknowns:

$$\begin{matrix} -2a + 6b = 5 \\ -6a - 2b = 0 \end{matrix} \quad \Rightarrow \quad \begin{matrix} a = -\frac{1}{4} \\ b = \frac{3}{4} \end{matrix}$$

$$y_p = -\tfrac{1}{4}\cos(2x) + \tfrac{3}{4}\sin(2x)$$

This leads us to the general solution of the inhomogeneous DE $y = y_h + y_p$:

$$y = C_1\, e^{-x} + C_2\, e^{-2x} - \tfrac{1}{4}\cos(2x) + \tfrac{3}{4}\sin(2x)$$

◁

7.4.4. Source term as a sum of functions

If the source term is a sum of two other known source terms we may treat them separately and add the particular solutions.

Theorem 7.3 (Sum of source terms)**.** Let $y_1(x)$ and $y_2(x)$ be solutions of a linear inhomogeneous DE with constant coefficients and the source terms $s_1(x)$ and $s_2(x)$, respectively:

$$P_n(D)y_1 = s_1 \qquad \text{and} \qquad P_n(D)y_2 = s_2$$

Then $y(x) = y_1(x) + y_2(x)$ is the solution of the DE with the source term $s(x) = s_1(x) + s_2(x)$, i.e.

$$P_n(D)(y_1 + y_2) = s_1 + s_2$$

◁

Proof.

$$\begin{aligned} P_n(D)y &= P_n(D)\,(y_1 + y_2) \\ &= \sum_{k=0}^{n} a_k (y_1 + y_2)^{(k)} \\ &= \sum_{k=0}^{n} a_k (y_1^{(k)} + y_2^{(k)}) \end{aligned}$$

$$
\begin{aligned}
&= \sum_{k=0}^{n} a_k y_1^{(k)} + \sum_{k=0}^{n} a_k y_2^{(k)} \\
&= P_n(D) y_1 + P_n(D) y_2 \\
&= s_1 + s_2
\end{aligned}
$$

□

Example 7.8. We want to solve the DE

$$y'' + 2y' + y = e^{2x} + x - 1$$

The corresponding homogeneous DE has the characteristic equation $k^2 + 2k + 1 = 0$ with a double zero at -1 which leads us to the general homogeneous solution:

$$y_h = (C_1 x + C_2)\, e^{-x}$$

For the first summand of the source term:

$$y_{p1} = a\, e^{2x} \qquad y'_{p1} = 2a\, e^{2x} \qquad y''_{p1} = 4a\, e^{2x}$$

$$
\begin{aligned}
y''_{p1} + 2y'_{p1} + y_{p1} &= e^{2x} \\
4a\, e^{2x} + 2 \cdot 2a\, e^{2x} + a\, e^{2x} &= e^{2x} \\
4a + 4a + a &= 1 \\
a &= \tfrac{1}{9} \\
y_{p1} &= \tfrac{1}{9}\, e^{2x}
\end{aligned}
$$

For the second summand of the source term:

$$y_{p2} = s_0 + s_1 x \qquad y'_{p2} = s_1 \qquad y''_{p2} = 0$$

$$
\begin{aligned}
y''_{p2} + 2y'_{p2} + y_{p2} &= x - 1 \\
0 + 2s_1 + s_0 + s_1 x &= x - 1 \\
s_1 x + 2s_1 + s_0 &= x - 1 \\
\begin{matrix} s_1 = 1 \\ 2s_1 + s_0 = -1 \end{matrix} \quad &\Rightarrow \quad \begin{matrix} s_1 &= 1 \\ s_0 &= -3 \end{matrix} \\
y_{p2} &= x - 3
\end{aligned}
$$

The particular solution of the original DE is

$$y_p = y_{p1} + y_{p2} = \tfrac{1}{9}\, e^{2x} + x - 3$$

The general solution $y = y_h + y_p$:

$$y(x) = (C_1 x + C_2)\, e^{-x} + \tfrac{1}{9}\, e^{2x} + x - 3$$

◁

7.4.5. DE in resonance

If the source term equals one of the particular homogeneous solutions, the DE is said to be in *resonance*.

The ansätze listed above will not lead to the desired result. When inserting into the DE we will end at a contradiction like $0 = f(x)$ or $0 = 1$.

To solve the DE we use the appropriate ansatz from above and perform variation of parameters.

Example 7.9. We want to solve the DE

$$y'' + 3y' + 2y = e^{-x}$$

The corresponding homogeneous DE has the characteristic equation $k^2 + 3k + 2 = 0$ with the two zeros at -1 and -2 which leads to the general homogeneous solution:

$$y_h = C_1\, e^{-x} + C_2\, e^{-2x}$$

We first try the standard ansatz:

$$y_{p1} = a e^{-x} \qquad y'_{p1} = -a\, e^{-x} \qquad y''_{p1} = a\, e^{-x}$$

$$
\begin{aligned}
y''_{p1} + 3y'_{p1} + 2y_{p1} &= e^{-x} \\
a\, e^{-x} - 3a\, e^{-x} + 2a\, e^{-x} &= e^{-x} \\
a - 3a + 2a &= 1 \\
0 &= 1 \qquad ???
\end{aligned}
$$

This is obviously wrong. We take the second ansatz and treat the parameter a as a function:

$$
\begin{aligned}
y_{p2} &= a\, e^{-x} \\
y'_{p2} &= a'\, e^{-x} - a\, e^{-x} \\
&= (a' - a)\, e^{-x} \\
y''_{p2} &= (a'' - a')\, e^{-x} - (a' - a)\, e^{-x} \\
&= (a'' - 2a' + a)\, e^{-x}
\end{aligned}
$$

$$
\begin{aligned}
y''_{p2} + 3y'_{p2} + 2y_{p2} &= e^{-x} \\
(a'' - 2a' + a)\, e^{-x} + 3(a' - a)\, e^{-x} & \\
+2a\, e^{-x} &= e^{-x} \\
(a'' - 2a' + a) + 3(a' - a) + 2a &= 1 \\
a'' + a' &= 1
\end{aligned}
$$

Since we are looking for a particular solution $y_p(x)$ we try to find a simple solution for $a(x)$. Integrating both sides ginves us $a' + a = x + C$.

A simple solution is $a(x) = x$ which we insert into the particular solution:

$$y_{p2} = x\, e^{-x}$$

The general solution of the DE $y = y_h + y_{p2}$ is:

$$\begin{aligned} y(x) &= C_1\, e^{-x} + C_2\, e^{-2x} + x\, e^{-x} \\ &= (C_1 + x)\, e^{-x} + C_2\, e^{-2x} \end{aligned}$$

◁

7.5. Problems

Problem 7.1: Solve the following DEs and check your solution:

1. $y'' + y' - 2y = 0$
2. $y'' - y = 0$
3. $y'' - 2y' + y = 0$
4. $y'' + 2y' + 2y = 0$
5. $y'' = 4y$
6. $\dfrac{y''}{y'} + \dfrac{y}{4y'} = 1$
7. $\dfrac{4y'}{y} = \dfrac{4y''}{y} + 5$
8. $2y' = y'' + 10y$

Problem 7.2: Solve the following DEs and check your solution:

1. $y''' + 6y'' + 11y' + 6y = 0$
2. $y^{(4)} - 5y'' + 4y = 0$
3. $y''' - y'' + y' - y = 0$
4. $y''' + 3y'' + 4y' + 2y = 0$
5. $y^{(4)} + y = 2y''$
6. $y^{(4)} + 2y'' + 5y = 8y'$

Problem 7.3: Solve the following DEs:

1. $y'' + 3y' + 2y = 2x + 1$
2. $y'' + 2y' + y = x + 2$
3. $y'' + 2y' + 2y = 2(1 - x)$
4. $y''' + y + 3(y'' + y') + x^2 = 0$
5. $y^{(4)} + 7y'' + 10y = 5 + 18y'$
6. $\dfrac{4y''}{y} + x^3/y = 1$

Problem 7.4: Solve the following DEs:

1. $y'' + y' - 2y = 2\, e^{3x}$
2. $y'' + 4y = 8\, e^{-2x}$
3. $y'' - 4y' + 4y = 3e^{-x}$
4. $y''' + y' = y'' + y + 3\, e^{-2x}$
5. $(2y'' - y' - y)\, e^{2x} = 9$

Problem 7.5: Solve the following DEs:

1. $y'' - 4y = 5\sin(x)$
2. $y'' + y + 4\cos(x) = 2y'$
3. $y'' + 4y = 5\cos(3x)$
4. $y'' - 2y' + 5y = \sin(x) + 2\cos(x)$
5. $y' + y = \sin(-2x)$
6. $y''' + 3(y'' + y') + y = 2\cos(x) - 2\sin(x)$

Problem 7.6: Solve the following DEs:

1. $y'' + 2y' - 3y = e^{-x} - 3x$
2. $y'' + 9y = 8\sin(x) + 6\, e^{3x}$
3. $y'' + 4(x^2 - y) + \cos(x) = 0$
4. $y''' + 4y'' + 5y' + 2y = 2x^2 + 10x + 8 + 12\, e^x$

Problem 7.7: Solve the following DEs:

1. $y' + 2y = e^{-2x}$
2. $y'' + 4(y + \cos(2x)) = 0$
3. $y'' + y' - 6y + 3\, e^{-3x} = 0$
4. $y''' + 4y'' + 5y' + 2y = 2\, e^{-x}$

8. Combinatorics

8.1. Permutations

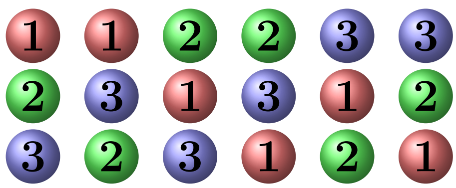
Definition 8.1 (Permutation)**.** We call the arrangement of n items into a definite order *permutation*. ◁

Theorem 8.2 (Permutations of different items)**.** The number of permutations for n different items is:

$$P_n = n!$$

◁

Example 8.1. There are six options to bring three different balls into an order:



◁

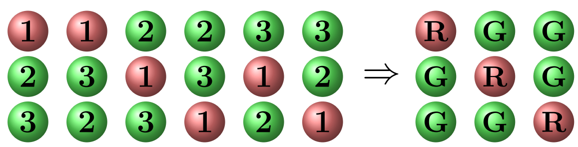
Example 8.2. There are $6! = 720$ different options to place six students on six chairs. ◁

Theorem 8.3 (Permutations with multiples)**.** The number of permutations of n items with k of them being equal is.

$$P_{n,k} = \frac{n!}{k!}$$

◁

Example 8.3. For $n = 3$ different balls there are $n! = 6$ options to bring them into an order. If $k = 2$ of the three balls have the same colour, there are only $n!/k! = 3$ options remaining:



◁

Example 8.4. A waiter wants to place 10 plates on a table. Six plates are white, the other plates are red, green, blue and yellow. There are $\frac{10!}{6!} = 5040$ options to lay out the table. ◁

Theorem 8.4 (Permutations with multiple multiples)**.** The number of permutations for n items containing m k_i-multiples, $i = 1 \dots m$ with $\sum_{i=1}^{m} k_i = n$ is:

$$P_{n,k_1,k_2,\dots,k_m} = \frac{n!}{k_1!k_2!\dots k_m!}$$

◁

Example 8.5. For $n = 4$ balls with different colours there are $\frac{4!}{1!\cdot 1!\cdot 1!\cdot 1!} = 24$ options to bring them into an order, see figure 8.1 left.

If two of them have the same colour the number of permutations reduce to $\frac{4!}{1!\cdot 1!\cdot 2!} = 12$, see figure 8.1 middle.

If another two balls have the same colour the number of permutations reduce further to $\frac{4!}{2!\cdot 2!} = 6$, see figure 8.1 right. ◁

Example 8.6. In Germany playing *lotto* is about choosing six numbers out of 49 where the order of the chosen numbers do not play a role.

Let $n = 49$ be the number of all possible numbers, $k_1 = 6$ be the number of chosen numbers and $k_2 = 43$ the number of not-chosen numbers. The number of possible choices is:

$$n_{\text{choices}} = \frac{n!}{k_1!k_2!} = \frac{49!}{6!\cdot 43!} = 13\,983\,816$$

The chance of being the happy winner is:

$$\frac{1}{n_{\text{choices}}} \approx 7 \cdot 10^{-8} = 0.000\,007\%$$

◁

8.2. Combinations

Definition 8.5 (Combination)**.** A *combination* is a selection of k items out of n different items neglecting the order. We call it a combination of class k with or without replacement. ◁

Remark: *With replacement* means a selected item is put back before the next selection takes place. Hence, the number of selections may be larger than the total number of items.

Unlike to the previous section we neglect the order of the selected items.

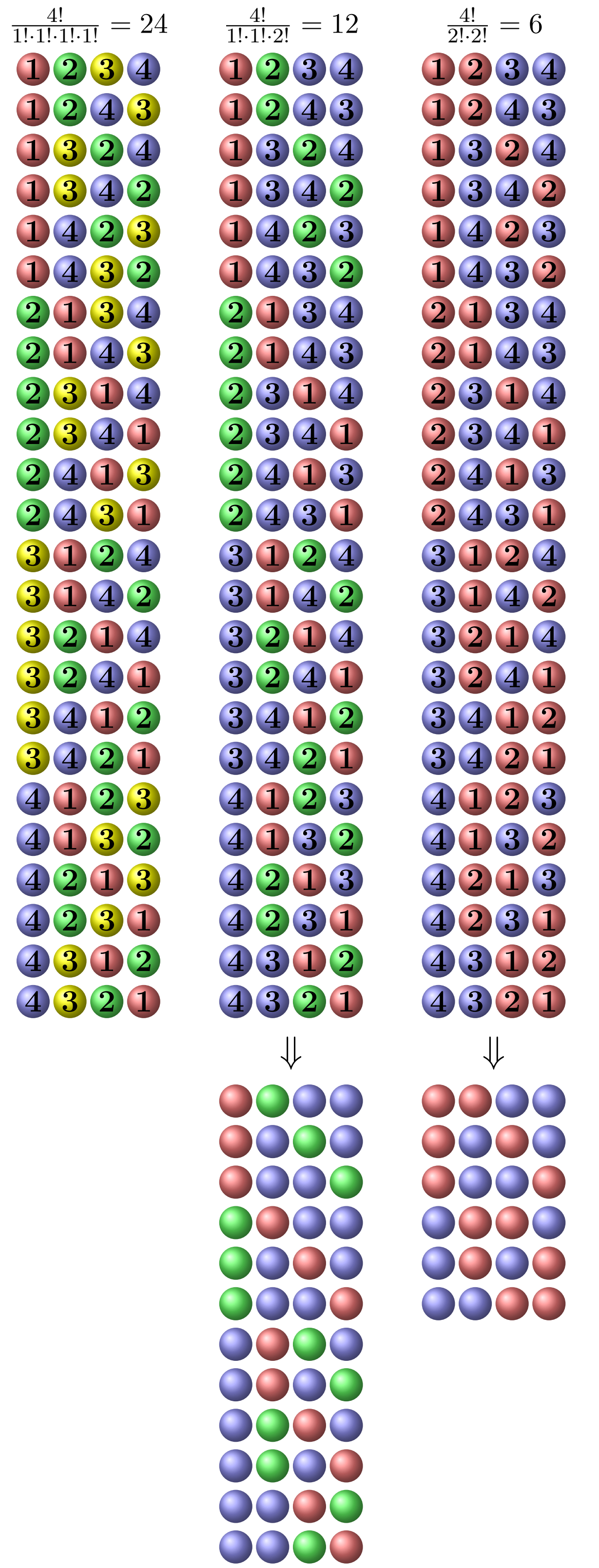


Figure 8.1.: Permutations of 4 balls

Theorem 8.6 (Combinations without replacement)**.** The number of combinations to choose k items out of n different items without replacement is:

$$C_{n,k} = \frac{n!}{k!(n-k)!}$$

◁

Example 8.7. For an election of a steering committee with four members out of 10 candidates the number of different formations of the committee is:

$$C_{10,4} = \frac{10!}{4! \cdot (10-4)!} = 210$$

◁

Definition 8.7 (binomial coefficient)**.** We define

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

as the *binomial coefficient* and read 'n choose k'. ◁

Example 8.8. We analyse again the German *lotto*. Choosing $k=6$ numbers out of $n=49$ different numbers results for the number of combinations:

$$C_{n,k} = \binom{49}{6} = \frac{49!}{6!(49-6)!} = 13\,983\,816$$

◁

Theorem 8.8 (Combinations with replacement)**.** The number of options to choose k items out of n different items with replacement is:

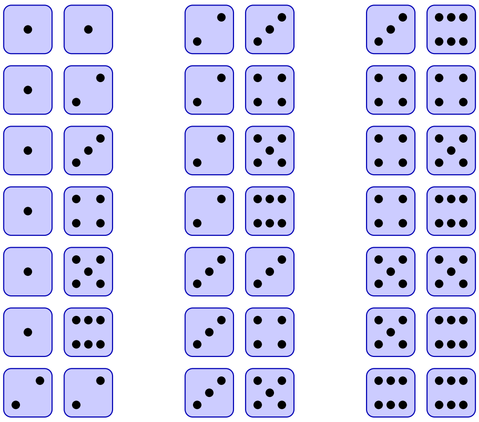
$$C_{n,k} = \binom{n+k-1}{k} = \frac{(n+k-1)!}{k!(n-1)!}$$

◁

Example 8.9. What is the number of different outcomes when throwing two six-sided dice?

Here the items are the numbers one to six, i.e. $n = 6$. The number of choices are the number of throws, i.e. $k = 2$. Hence, the total number of different outcomes is:

$$C_{6,2} = \binom{6+2-1}{2} = \frac{(6+2-1)!}{2!(6-1)!} = 21$$



◁

8.3. Variations

Definition 8.9 (Variations)**.** A *variation* is a selection of k out of n different items with or without replacement minding the order of items. I.e. *variations* are *combinations* minding the *order* of items. ◁

Theorem 8.10 (Variations without replacement)**.** The number of variations for a selection of k items out of n different items without replacement is:

$$V_{n,k} = k!\binom{n}{k} = \frac{n!}{(n-k)!}$$

◁

Example 8.10. We again look at an election of a steering committee with four members out of 10 candidates. This time the roles within the steering committee are *chairman*, his/her *representative*, the *treasurer* and the *secretary*. The number of different possible formations of the committee is:

$$V_{10,4} = 4!\binom{10}{4} = \frac{10!}{(10-4)!} = 5\,040$$

◁

Theorem 8.11 (Variations with replacement)**.** The number of variations for a selection of k items out of n different items with replacement is:

$$V_{n,k} = n^k$$

◁

Example 8.11. A four bit binary number may be looked at as four selections from the set $\{0, 1\}$ with replacement. I.e. either the one or the zero is selected, but then replaced so that both items are available for the next selection. The number of variations is:

$$V_{2,4} = 2^4 = 16$$

0000	0100	1000	1100
0001	0101	1001	1101
0010	0110	1010	1110
0011	0111	1011	1111

◁

8.4. Summary

Summary of the equations for combinatorics:

	recurrence	
	without	with
permutations	$P_n = n!$	$P_{n,k} = \frac{n!}{k!}$
combinations	$C_{n,k} = \binom{n}{k}$	$C_{n,k} = \binom{n+k-1}{k}$
variations	$V_{n,k} = k!\binom{n}{k}$	$V_{n,k} = n^k$

8.5. Problems

Problem 8.1: How many options are there to place eight people on eight chairs?

Problem 8.2: For a race with ten participants how many different placings are there for place one (fastest) to ten (slowest).

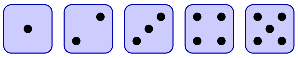
Problem 8.3: How many options are there to place two women and two men on six chairs w.r.t. to their gender?

Problem 8.4: How many different values can a byte-variable (8 bits) have with exactly four bits being one?

Problem 8.5: Assuming a painter has eight different colours available and uses three different of them to mix a new colour. If he always takes the same amount of each colour (e.g. 1 ml from each of the three colours) how many different colours is he able to create?

Problem 8.6: Out of 9 different fruits you choose four different to prepare a fruit salad. How many options do you have?

Problem 8.7: *Yahtzee* (in Germany *Kniffel*) is a famous game with five dice. How many different combinations are there for a single throw?



Problem 8.8: How many different outcomes are expected throwing a coin six times neglecting the order of outcomes?

Problem 8.9: How many different options would the German lotto (6 out of 49 different numbers) have if the order is minded?

Problem 8.10: For a hundred meter sprint how many options are there for eight runners to take the positions one, two and three?

Problem 8.11: For a number system with base n, how many different numbers can be expressed with k digits?

Problem 8.12: A manager has to delegate four different tasks to his twelve staff members. Each task requires only one staff member and a staff member may take more than one task. How many options does the manager have?

9. Probability theory

9.1. Introduction

When dealing with functions any valid input leads to a unique output. E.g. $\sin(\pi/6)$ has the unique value $\frac{1}{2}$ and nothing else. Applying a function to an input is a *deterministic process*.

Throwing an unbiased six-sided die results in an unpredictable number between one and six. We do not know in advance whether the number one will appear, however, we expect on average for every sixth throw a one. Throwing a die is a *stochastic process*.

Example 9.1. Deterministic and stochastic injuries.

In medicine when exposed to some ionizing radiation (e.g. x-rays) we distinguish between *deterministic* and *stochastic* injuries.

Deterministic injuries appear for higher radiation doses and have symptoms like skin redness, skin burn etc. An increase of dose leads to a higher degree of injury.

Stochastic injuries appear also for low dose levels. In many cases the patient will experience no symptoms at all. However, the risk to get an injury like cancer increases. For a given dose level the injury can not be predicted. However, the probabilities are studied and are subject to further investigations. ◁

Although we can not predict the *particular outcome* of a stochastic process, it is possible to analyse the *probability* of such processes. The knowledge of probabilities supports decision making in many situations. Here is one example:

Example 9.2. *Monty Hall problem.*

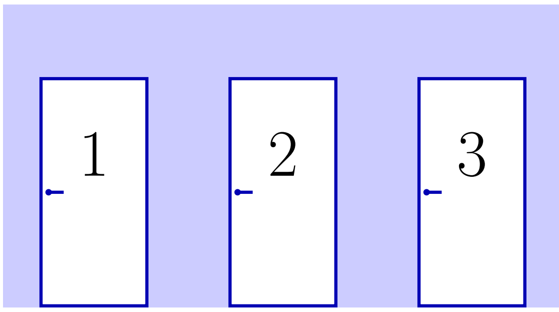
In a famous quiz-show as the winner of the day you get the chance to win a brand new Mercedes by the following game: You enter a room with three doors. Behind one door there is the Mercedes, whereas each of the other two doors hide a goat. The rules of the game are as follows:

- First you choose one of the doors without opening it.
- Then the show-master (who knows where the Mercedes is) opens one of the other doors with a goat behind it.
- Finally you can either open the door of your first choice, or you change your mind and open the other not yet opened door. If you find the Mercedes, it's yours!

The question is: Should you keep your first choice, or should you open the other yet closed door? Which door has the higher probability to hide the Mercedes?

Suppose you chose door 1 and the show-master opened door 2 revealing a goat. You then have the choice between door 1 or 3. Let A_1, A_2 and A_3 be the events that the Mercedes is behind door 1, 2 and 3, respectively. Let further be $\mathrm{P}(A_k)$ the probability that the Mercedes is behind door k ($k = 1, 2, 3$). The question is:

- $\mathrm{P}(A_1) > \mathrm{P}(A_3)$?
- $\mathrm{P}(A_1) = \mathrm{P}(A_3)$?
- $\mathrm{P}(A_1) < \mathrm{P}(A_3)$?



Discuss with you neighbour! ◁

9.2. Sample, sample space and event

When dealing with a stochastic process we perform some sort of experiment. Since the output of the process varies we perform a number of experiments to gain an understanding of the process.

For further discussion we first need to define some terms:

Definition 9.1 (Sample, sample space, event)**.**
We call the outcome of a stochastic experiment a *sample* ω.

The set of all possible different samples forms the *sample space* Ω.

We call any subset of the sample space an *event* A. ◁

Example 9.3. A standard six-sided die has the *sample space*

$$\Omega = \{1, 2, 3, 4, 5, 6\}.$$

We now may define several *events*, e.g.

- the event of throwing an even number, i.e.

$$A_{\text{even}} = \{2, 4, 6\} \subseteq \Omega$$

- the event of throwing an odd number, i.e.

$$A_{\text{odd}} = \{1, 3, 5\} \subseteq \Omega$$

- the event of throwing a particular number:

$$A_k = \{k\} \subseteq \Omega \qquad \text{for } k = 1, \ldots, 6$$

- the event of throwing a number greater than three:

$$A_{>3} = \{4, 5, 6\} \subseteq \Omega$$

Performing experiments give particular samples. E.g. throwing a die ten times may give the samples 1, 4, 3, 2, 6, 4, 4, 6, 3 and 6. ◁

Example 9.4. Throwing a coin with heads and tails (H,T) twice has the sample space

$$\Omega = \{(H, H), (H, T), (T, H), (T, T)\}$$

Some possible events are:

- the event of throwing the same face twice:

$$A_{\text{equal}} = \{(H, H), (T, T)\} \subseteq \Omega$$

- the event of throwing different faces:

$$A_{\text{unequal}} = \{(H, T), (T, H)\} \subseteq \Omega$$

- the event of throwing heads first:

$$A_{HX} = \{(H, H), (H, T)\} \subseteq \Omega$$

Performing five experiments, i.e. throwing a coin ten times in total may result in the samples (T, T), (T, H), (T, H), (H, T) and (T, H). ◁

Example 9.5. For the lifetime of light bulbs in hours we have the sample space

$$\Omega = \mathbb{R}_{\geq 0}$$

We may define the following events:

- $A_{\text{DOA}} = \{x \in \mathbb{R} \mid 0 \leq x < 100\}$
  the event of *dead on arrival.*
- $A_{\text{short}} = \{x \in \mathbb{R} \mid 100 \leq x < 800\}$
  the event of *short lifetime.*
- $A_{\text{typ}} = \{x \in \mathbb{R} \mid 800 \leq x < 1200\}$
  the event of *typical lifetime.*
- $A_{\text{long}} = \{x \in \mathbb{R} \mid x \geq 1200\}$
  the event of *long lifetime.*

Investigating the lifetime of 100 light bulbs of a particular type may result in:

event	samples
DOA	4
short	12
typical	71
long	13

◁

Definition 9.2 (finite, countable, continuous)**.**
We call a sample space with finite (or infinite) number of elements a *finite sample space* (or *infinite sample space*).

We call a infinite sample space Ω with countable (or uncountable) elements *countable sample space* (or *uncountable sample space*).

An infinite uncountable sample space is called *continuous sample space*.

A finite sample space or infinite countable sample space is called *discrete sample space*. ◁

Remark: A finite sample space implies a countable sample space.

Example 9.6. ◁

- The sample space when throwing a coin has the two elements *heads* and *tails*, i.e.

$$\Omega = \{H, T\}$$

  It is a *finite*, hence *discrete* sample space.

- A standard six-sided die has the sample space with six elements:

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

  It is a *finite*, hence *discrete* sample space.

- The sample space for the number of measured quanta per second is the set of natural numbers including zero:

$$\Omega = \{0, 1, 2, 3, \ldots\} = \mathbb{N} \cup \{0\}$$

  It is an *infinite*, *countable* and, hence, *discrete* sample space.

- The time until disintegration of a radioactive atom has the sample space

$$\Omega = \mathbb{R}_{\geq 0}$$

  It is an *infinite*, *uncountable* and, hence, *continuous* sample space.

◁

If we perform an experiment and get a sample ω which is an element of the event A, i.e. $\omega \in A$, then we say: The event A took place.

Since sample space Ω and events A_{xyz} are sets, we may apply the same calculation rules as for sets.

Example 9.7.
With the definitions of example 9.3 we have:

- $A_1 \cup A_{\text{even}} = \{1, 2, 4, 6\}$
- $A_{\text{even}} \cap A_{>3} = \{4, 6\}$
- $A_{\text{odd}} = \overline{A}_{\text{even}} = \Omega \setminus A_{\text{even}} = \{1, 3, 5\}$
- $A_{\text{odd}} \setminus A_1 = \{3, 5\}$
- $2 \in A_{\text{even}}, \qquad 2 \notin A_{\text{odd}}$
- $\{1, 2\} \subseteq \Omega, \qquad \{5, 6, 7\} \nsubseteq \Omega$
- $A_{\text{odd}} \cap A_{\text{even}} = \{\} = \emptyset$
- $A_{\text{odd}} \cup A_{\text{even}} = \Omega$

Definition 9.3 (Mutual exclusivity)**.** If the intersection of two events A and B is an empty set, i.e.

$$A \cap B = \emptyset$$

then we say the two events are *mutually exclusive*. I.e. a sample ω is either an element of A *or* an element of B but never an element of both. ◁

9.3. Probability

Although we can not predict the outcome of a stochastic experiment, we may know the *probability* of a given event.

E.g. for an unbiased six-sided die we expect the probability to throw a six to be $\frac{1}{6}$, i.e. $16\frac{2}{3}\%$. On average every 6^{th} throw should result in a six.

But what is probability and how can we define it?

9.3.1. Frequentist probability

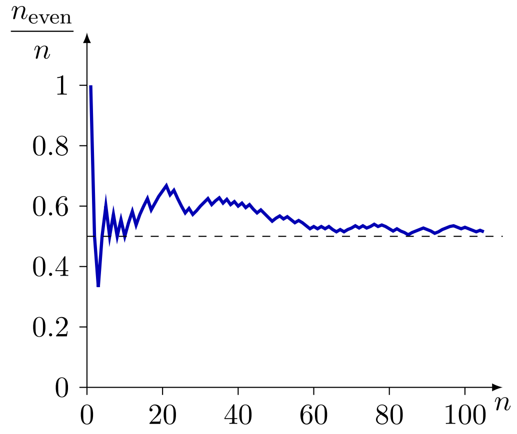
Definition 9.4 (Frequentist probability)**.** For n experiments with n samples let n_k be the number of samples being elements of the event A_k. We define the probability $\mathrm{P}(A_k)$ of the event A_k as

$$\mathrm{P}(A_k) = \lim_{n \to \infty} \frac{n_k}{n}$$

◁

Example 9.8. When throwing an unbiased six-sided die what is the probability A_{even} of throwing even numbers?

We perform an increasing number of experiments n and plot after every throw the quotient of the number of even numbers n_{even} and the total number of experiments n, i.e. $\frac{n_{\text{even}}}{n}$ into a diagram.



The more experiments we perform, the closer we get to the precise probability of $\mathrm{P}(A_{\text{even}}) = 0.5$. ◁

It becomes obvious that the probability has a value between zero an one:

$$0 \leq \mathrm{P}(A) \leq 1$$

9.3.2. Probability axioms

We now build our understanding of probability on the three axioms by Andrey Kolmogorov:

Definition 9.5 (Probability axioms)**.** With X being a stochastic experiment with sample space Ω and events A_k, $k \in \mathbb{N}$ we define the probability $\mathrm{P}(A_k)$ by the following axioms:

Axiom 1: For all $A_k \subseteq \Omega$ we have
$\mathrm{P}(A_k) \in \mathbb{R}_{\geq 0}$

Axiom 2:
$\mathrm{P}(\Omega) = 1$

Axiom 3: If $A_1 \cap A_2 = \emptyset$ then
$\mathrm{P}(A_1 \cup A_2) = \mathrm{P}(A_1) + \mathrm{P}(A_2)$

◁

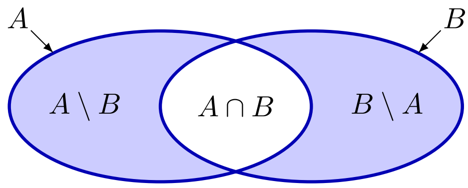
9.3.3. Calculating with probabilities

We now derive a number of calculation rules for probability:

Theorem 9.6 (Calculating with probabilities)**.** For two events A and B of a sample space Ω we have:

- $\mathrm{P}(\overline{A}) = 1 - \mathrm{P}(A)$
- $A \subseteq B \Rightarrow \mathrm{P}(A) \leq \mathrm{P}(B)$
- $\mathrm{P}(A \cup B) = \mathrm{P}(A) + \mathrm{P}(B) - \mathrm{P}(A \cap B)$
- $\mathrm{P}(A \setminus B) = \mathrm{P}(A) - \mathrm{P}(A \cap B)$

◁



Example 9.9. For an unbiased six-sided die we have $\mathrm{P}(A_6) = \frac{1}{6}$ as the probability to throw a six, $\mathrm{P}(A_{\text{even}}) = \frac{1}{2}$ as the probability to throw an even number and $\mathrm{P}(A_{>4}) = \frac{1}{3}$ as the probability to throw a number greater than four.

- The probability to throw a number less or equal to four is
$$\mathrm{P}(\overline{A}_{>4}) = 1 - \mathrm{P}(A_{>4}) = 1 - \tfrac{1}{3} = \tfrac{2}{3}$$
- Since A_6 is a subset of $A_{>4}$ we have
$$\begin{aligned} A_6 \subseteq A_{>4} \quad &\Rightarrow \quad \mathrm{P}(A_6) \leq \mathrm{P}(A_{>4}) \\ \{6\} \subseteq \{5, 6\} \quad &\Rightarrow \quad \tfrac{1}{6} \leq \tfrac{1}{3} \end{aligned}$$
- The probability to throw an even number or a number larger than 4:
$$\begin{aligned} &\mathrm{P}(A_{\text{even}} \cup A_{>4}) \\ &\quad = \mathrm{P}(A_{\text{even}}) + \mathrm{P}(A_{>4}) - \mathrm{P}(A_{\text{even}} \cap A_{>4}) \\ &\quad = \tfrac{1}{2} + \tfrac{1}{3} - \tfrac{1}{6} = \tfrac{2}{3} \end{aligned}$$
- The probability to throw an even number that is not larger than 4:
$$\begin{aligned} &P(A_{\text{even}} \setminus A_{>4}) \\ &\quad = \mathrm{P}(A_{\text{even}}) - \mathrm{P}(A_{\text{even}} \cap A_{>4}) \\ &\quad = \tfrac{1}{2} - \tfrac{1}{6} = \tfrac{1}{3} \end{aligned}$$

◁

9.3.4. Conditional probability

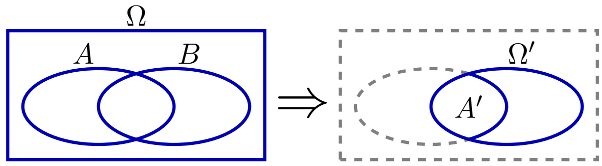
Conditional probability is the probability of an event A under the condition, that event B is given. We denote it by $\mathrm{P}(A|B)$ or $\mathrm{P}_B(A)$.

Theorem 9.7 (Conditional probability)**.** The probability of an event A under the condition of event B is given by

$$\mathrm{P}(A|B) = \frac{\mathrm{P}(A \cap B)}{\mathrm{P}(B)} \qquad \text{for } \mathrm{P}(B) \neq 0$$

◁

We can imagine conditional probability as the probability with reduced sample space. It is a reduced event $A' = A \cap B$ in a reduced sample space $\Omega' = B$.



Example 9.10. What is the probability for an unbiased six-sided die to throw an even number under the condition that the number must be larger than 3?

With A_{even} as the event of throwing an even number and $A_{>3}$ as the event of throwing a number larger than 3 we get:

$$\mathrm{P}(A_{\text{even}}|A_{>3}) = \frac{\mathrm{P}(A_{\text{even}} \cap A_{>3})}{\mathrm{P}(A_{>3})} = \frac{\frac{1}{3}}{\frac{1}{2}} = \frac{2}{3}$$

This is obviously true since in the set of numbers larger than 3 (i.e. $\{4, 5, 6\}$) two of three numbers are even. ◁

There are situations where conditional probabilities are given and unconditional probabilities need to be evaluated. We rewrite the equation of theorem 9.7 to:

$$\mathrm{P}(A \cap B) = \mathrm{P}(A|B) \cdot \mathrm{P}(B)$$

If we want to know the probability of an event A but we know only the probabilities under the condition B and $\overline{B}$ we can use this method. Since

$$A = (A \cap B) \cup (A \cap \overline{B})$$

and

$$(A \cap B) \cap (A \cap \overline{B}) = \emptyset$$

we can write

$$\begin{aligned} \mathrm{P}(A) &= \mathrm{P}((A \cap B) \cup (A \cap \overline{B})) \\ &= \mathrm{P}(A \cap B) + \mathrm{P}(A \cap \overline{B}) \\ \mathrm{P}(A) &= \mathrm{P}(A|B)\mathrm{P}(B) + \mathrm{P}(A|\overline{B})\mathrm{P}(\overline{B}) \end{aligned}$$

Theorem 9.8 (Sum of conditional probabilities)**.** If $B_1, B_2, \ldots$ are mutually exclusive and their union equals the sample space, i.e. for every experiment exactly one of the events $B_1, B_2, \ldots$ must occur, then we have:

$$\mathrm{P}(A) = \sum_k \mathrm{P}(A \cap B_k) = \sum_k \mathrm{P}(A|B_k)\mathrm{P}(B_k)$$

◁

Example 9.11. In a company machines 1 and 2 produce 30% and 70% of the overall production with 5% and 2% failure rate, respectively. a) What is the overall failure rate? b) For a failed sample, what is the probability that it has been produced by machine 1?

We define B_1 and B_2 as the events that a product is produced by machine 1 or 2, respectively. We get

$$\mathrm{P}(B_1) = 0.3 \qquad \mathrm{P}(B_2) = 0.7$$

Any sample is either produced by machine 1 or by machine 2 – never by both, i.e. B_1 and B_2 are mutually exclusive and their union equals the sample space.

With the event of failure A we have the two conditional probabilities

$$\begin{aligned} \mathrm{P}(A|B_1) &= 0.05 \\ \mathrm{P}(A|B_2) &= 0.02 \end{aligned}$$

Question a):

$$\begin{aligned} \mathrm{P}(A) &= \mathrm{P}(A \cap B_1) + \mathrm{P}(A \cap B_2) \\ &= \mathrm{P}(A|B_1)\mathrm{P}(B_1) + \mathrm{P}(A|B_2)\mathrm{P}(B_2) \\ &= 0.05 \cdot 0.3 + 0.02 \cdot 0.7 = 0.029 = 2.9\% \end{aligned}$$

Question b):

$$\begin{aligned} \mathrm{P}(B_1|A) &= \frac{\mathrm{P}(B_1 \cap A)}{\mathrm{P}(A)} = \frac{\mathrm{P}(A|B_1)\mathrm{P}(B_1)}{\mathrm{P}(A)} \\ &= \frac{0.05 \cdot 0.3}{0.029} = \frac{15}{29} \approx 0.52 = 52\% \end{aligned}$$

◁

9.3.5. Independence

In the previous examples we saw that the conditional probability of an event A may differ from its unconditional probability. However, if the unconditional probability of an event A equals the probability of A under the condition of B, then we say the two events A and B are independent.

Definition 9.9 (Independence)**.** Let A and B be two events. If the probability $\mathrm{P}(A)$ equals the conditional probability $\mathrm{P}(A|B)$, i.e.

$$\mathrm{P}(A) = \mathrm{P}(A|B)$$

then we say A and B are *independent events*. ◁

Example 9.12. An aircraft passenger has been caught having a bomb (without fuse) in his luggage. When taken to task he replied: "I took the bomb with me since it is more unlikely to have two bombs on one plane than only one." Where is the mistake in his argument?

Let A be the event that a terrorist places a bomb into the plane and B the event that the passenger takes his bomb with him.

The passenger argues that

$$\mathrm{P}(A \cap B) \leq \mathrm{P}(A)$$

which is true since $(A \cap B) \subseteq A$. However, assuming that the caught passenger and the potential terrorist do not influence each other, the two events A and B are independent and we get:

$$\mathrm{P}(A) = \mathrm{P}(A|B)$$

Hence, the caught passenger is wrong and he can not influence the probability the terrorist placing a bomb into the plane. ◁

9.4. Problems

Problem 9.1: For the sample space and the events

$$\begin{aligned} \Omega &= \{1, 2, 3, A, B, C\} \\ A_L &= \{A, B, C\} \\ A_N &= \{1, 2, 3\} \\ A_k &= \{k\} \qquad \text{for } k = 1, 2, 3, A, B, C \end{aligned}$$

find the following sets in extensional definition.

1. $A_L \cup A_N$
2. $A_L \cap A_N$
3. $\overline{A}_1$
4. $\overline{A}_B \setminus A_N$
5. $\Omega \setminus A_L$
6. $(A_C \cup A_N) \setminus A_2$

Problem 9.2: What does *mutual exclusivity* mean?

Problem 9.3: For the mutually exclusive events A and B with the probabilities $\mathrm{P}(A) = 0.2$ and $\mathrm{P}(B) = 0.3$ resolve the following expressions.

1. $\mathrm{P}(\overline{A})$
2. $\mathrm{P}(A \cup B)$
3. $\mathrm{P}(A \cap B)$
4. $\mathrm{P}(A \setminus B)$
5. $\mathrm{P}(\overline{B} \setminus A)$
6. $\mathrm{P}(\overline{A} \setminus \overline{B})$

Problem 9.4: For the two events A and B with the probabilities $\mathrm{P}(A) = 0.3$, $\mathrm{P}(B) = 0.5$ and $\mathrm{P}(A \cap B) = 0.2$ resolve the following expressions.

1. $\mathrm{P}(A \cup B)$
2. $\mathrm{P}(A \setminus B)$
3. $\mathrm{P}(B \setminus A)$
4. $\mathrm{P}(A \cap \overline{B})$
5. $\mathrm{P}(\overline{A \cap B})$
6. $\mathrm{P}(\overline{A} \setminus \overline{B})$

Problem 9.5: For the two events A and B with the probabilities $\mathrm{P}(A) = 0.3$, $\mathrm{P}(B) = 0.5$ and $\mathrm{P}(A \cap B) = 0.2$ find the conditional probabilities.

1. $\mathrm{P}(A|B)$
2. $\mathrm{P}(B|A)$

Problem 9.6: For the two mutually exclusive events A and B find the conditional probabilities.

1. $\mathrm{P}(A|B)$
2. $\mathrm{P}(B|A)$

Problem 9.7: The machines 1, 2 and 3 of a light bulb manufacturer produce 20%, 30% and 50% of the overall production with 5%, 3% and 2% failure rate, respectively.

a) What is the overall failure rate?

b) For a failed light bulb what are the probabilities that it has been produced by machine 1, 2 or 3?

Problem 9.8: The number of road deaths in Germany split into age groups for 2010 were:

no.	age	population	road deaths
1	< 15	10,941,200	104
2	$15 - 25$	9,136,400	791
3	$25 - 65$	44,829,800	1,842
4	$65 \leq$	16,844,300	910

a) Find the probabilities for a person to be in age group 1 to 4.

b) For each age group find the probability for a person to die due to road death.

c) Find the overall probability of road death by the results of a) and b).

d) By the results of a) to c) what is the probability of a road fatality to be in age group 1 to 4?

Problem 9.9: For the two events A and B we have $\mathrm{P}(A) = 0.5$, $\mathrm{P}(B) = 0.3$ and $\mathrm{P}(A \cap B) = 0.15$. Are the two events A and B independent?

10. Stochastic

10.1. Random variable

10.1.1. Definition of random variables

When dealing with stochastic processes we are often interested in values of the random process in terms of real numbers. E.g. when throwing a six-sided die we are interested in the number on the upper side.

Some stochastic outputs are no numbers, e.g. favourite colours, types of mobiles produced or the two faces of a coin. However, in many situations the outputs can be replaced by numbers. E.g. for a coin heads may be given a one and tails a zero.

Definition 10.1 (Random variable)**.** We call a stochastic process with sample space $\Omega \subseteq \mathbb{R}$ containing real numbers only a *random variable*.

A random variable with *discrete* sample space is called *discrete random variable*.

A random variable with *continuous* sample space is called *continuous random variable*. ◁

We denote a random variable by an upper case letter, e.g. X.

Definition 10.2 (Cumulative distribution function)**.** We define the *cumulative distribution function* or just *distribution function* $F(x)$, $x \in \mathbb{R}$ of a random variable X by

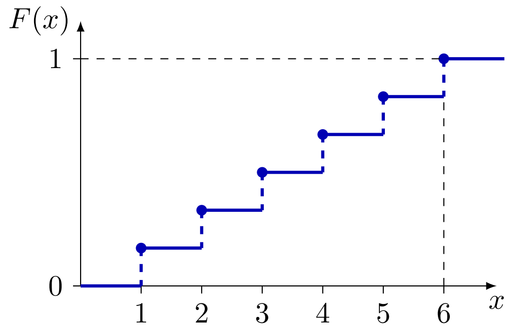
$$F(x) = \mathrm{P}(X \le x)$$

I.e. the cumulative distribution function $F(x)$ is the probability for X being less or equal to x. ◁

Example 10.1. The cumulative distribution function of an unbiased six-sided die with the numbers one to six is given by:

$$F(x) = \begin{cases} 0 & \text{for } \quad x < 1 \\ 1/6 & \text{for } 1 \le x < 2 \\ 1/3 & \text{for } 2 \le x < 3 \\ 1/2 & \text{for } 3 \le x < 4 \\ 2/3 & \text{for } 4 \le x < 5 \\ 5/6 & \text{for } 5 \le x < 6 \\ 1 & \text{for } 6 \le x \end{cases}$$

The dots in the following diagram indicate function values at points of discontinuities:



◁

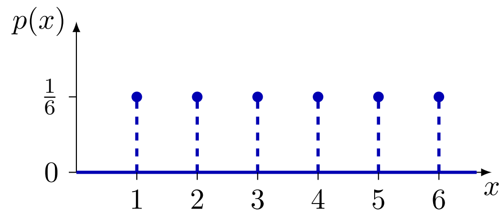
Definition 10.3 (Probability mass function)**.** For a discrete random variable X with sample space Ω we define the *probability mass function* $p(x)$ as:

$$p : \begin{cases} \Omega \to \mathbb{R} \\ x \mapsto \mathrm{P}(X = x) \end{cases}$$

◁

Example 10.2. An unbiased six sided die with the numbers 1 to 6 has the following probability mass function:

$$p(x) = \begin{cases} 1/6 & \text{for } x = 1, 2, \ldots, 6 \\ 0 & \text{otherwise} \end{cases}$$



◁

Definition 10.4 (Probability density function)**.** For a continuous random variable X with sample space $\Omega \subseteq \mathbb{R}$ the *probability density function* $f : \Omega \to \mathbb{R}$ is a function with its integral from $-\infty$ to x being the cumulative distribution function $F(x)$, i.e.

$$F(x) = \int_{-\infty}^{x} f(x')\, \mathrm{d}x'$$

◁

Both, the probability mass function $p(x)$ and the probability density function $f(x)$ are non-negative.

$$p(x) \geq 0 \qquad f(x) \geq 0 \qquad \text{for all } x \in \Omega$$

The sum of all probabilities of a probability mass function $p(x)$ must be one:

$$\sum_{\Omega} p(x) = \sum_{\Omega} P(x = X) = 1$$

The definite integral of a probability density function $f(x)$ over its entire sample space $\Omega \subseteq \mathbb{R}$ must be one:

$$\int_{\Omega} f(x)\,\mathrm{d}x = 1$$

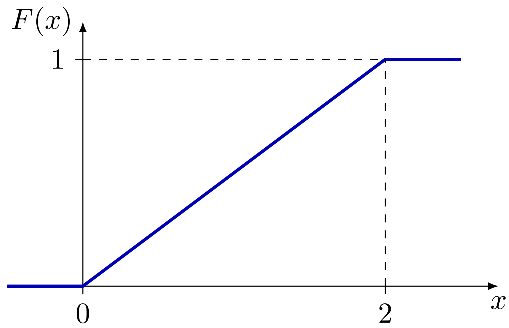
With $f(x) = 0$ for all $x \notin \Omega$ we have:

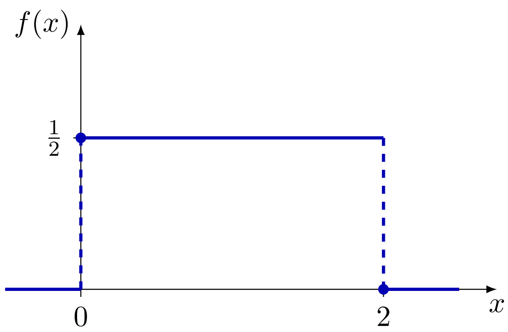
$$\int_{-\infty}^{\infty} f(x)\,\mathrm{d}x = 1$$

Example 10.3. A continuous random variable X has the sample space $\Omega = [0, 2)$ with all values having the same probability. We then have the cumulative distribution function $F(x)$ and the probability density function $f(x)$:

$$F(x) = \begin{cases} 0 & \text{for } x < 0 \\ \frac{x}{2} & \text{for } 0 \leq x < 2 \\ 1 & \text{otherwise} \end{cases}$$

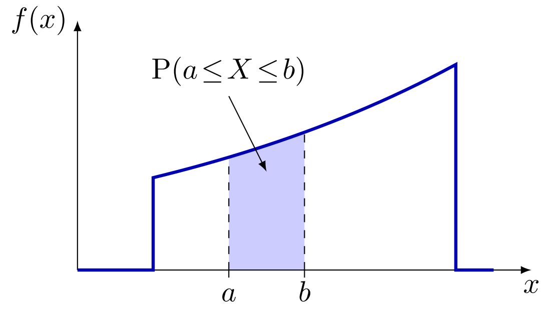
$$f(x) = \begin{cases} \frac{1}{2} & \text{for } 0 \leq x < 2 \\ 0 & \text{otherwise} \end{cases}$$





◁

For a continuous random process X the probability of the event $a \leq X \leq b$ can be visualized as the area between the abscissa and the probability density function on the interval $[a, b]$:



For a probability density function $f(x)$ the probability of a particular value x_0 is zero:

$$\mathrm{P}(X = x_0) = 0$$

Hence, for a continuous random variable X the probability of X being less than a limit x and the probability of X being less or equal to x are the same:

$$\mathrm{P}(X < x) = \mathrm{P}(X \leq x) \qquad \text{for } X \text{ being cont.}$$

10.1.2. Characteristics of random variables

Definition 10.5 (Expectation)**.** The *expectation* μ of a discrete random variable X with sample space $\Omega = \{x_1, x_2, \ldots, x_n\}$ and probability mass function $p(x)$ is defined by

$$\mu = \mathrm{E}(X) = \sum_{k=1}^{n} x_k\,\mathrm{P}(X = x_k) = \sum_{k=1}^{n} x_k\,p(x_k)$$

where E is the *expectation operator*.

The *expectation* μ of a continuous random variable X with sample space $\Omega = \mathbb{R}$ and probability density function $f(x)$ is defined by:

$$\mu = \mathrm{E}(X) = \int_{-\infty}^{\infty} x\,f(x)\,\mathrm{d}x$$

◁

The expectation of of a random variable X yields no information about the variation of X. A useful quantity is the average value of the square of the difference between X and its expectation μ:

Definition 10.6 (Variance and standard deviation)**.** For X being a random variable with expectation μ we define the variance $\sigma^2 = \mathrm{Var}(X)$ as:

$$\sigma^2 = \mathrm{Var}(X) = \mathrm{E}[(X-\mu)^2]$$

We call the square root of the variance $\sigma = \sqrt{\mathrm{Var}(X)}$ the *standard deviation* of X around its expectation μ. ◁

For a discrete random variable X with sample space $\Omega = \{x_1, x_2, \ldots, x_n\}$, probabilities $p(x_k)$, $k = 1, 2, \ldots, n$ and expectation μ we get:

$$\begin{aligned} \sigma^2 = \mathrm{Var}(X) &= \mathrm{E}[(X-\mu)^2] \\ &= \sum_{k=1}^{n} (x_k - \mu)^2 p(x_k) \end{aligned}$$

For a continuous random variable X with sample space $\Omega = \mathbb{R}$, probability density function $f(x)$ and expectation μ we get:

$$\begin{aligned} \sigma^2 = \mathrm{Var}(X) &= \mathrm{E}[(X-\mu)^2] \\ &= \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\,\mathrm{d}x \end{aligned}$$

Both, variance and standard deviation cannot be negative, i.e. $\sigma^2 \geq 0$ and $\sigma \geq 0$.

Example 10.4. For an unbiased six-sided die with numbers one to six we have:

$$\begin{aligned} \Omega &= \{1,2,3,4,5,6\} \\ x_k &= k \quad \text{for} \quad k = 1, \ldots, 6 \\ \mathrm{E}(X) = \mu &= \sum_{k=1}^{6} x_k \,\mathrm{P}(X = x_k) = \sum_{k=1}^{6} k \frac{1}{6} \\ &= \frac{1}{6} \cdot \frac{6(6+1)}{2} = 3\tfrac{1}{2} \\ \mathrm{Var}(X) &= \sum_{k=1}^{6} (x_k - \mu)^2 \mathrm{P}(X = x_k) \\ &= \sum_{k=1}^{6} (k - 3\tfrac{1}{2})^2 \frac{1}{6} \\ &= \frac{2}{6}(2.5^2 + 1.5^2 + 0.5^2) = \frac{35}{12} \approx 2.92 \\ \sigma &= \sqrt{\mathrm{Var}(X)} = \sqrt{\frac{35}{12}} \approx 1.71 \end{aligned}$$

◁

Example 10.5. For an unbiased six-sided die with numbers 1, 1, 1, 2, 2 and 3 we have:

$$\Omega = \{1,2,3\}$$

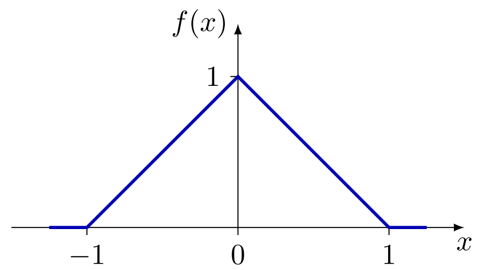
$$\begin{aligned} x_1 = 1 \qquad & x_2 = 2 \qquad x_3 = 3 \\ p(x_1) = \tfrac{1}{2} \qquad & p(x_2) = \tfrac{1}{3} \qquad p(x_3) = \tfrac{1}{6} \end{aligned}$$

$$\begin{aligned} \mu = \mathrm{E}(X) &= \sum_{k=1}^{3} x_k \,\mathrm{P}(X = k) = \sum_{k=1}^{3} k\, p(k) \\ &= 1 \cdot \tfrac{1}{2} + 2 \cdot \tfrac{1}{3} + 3 \cdot \tfrac{1}{6} = 1\tfrac{2}{3} \\ \sigma^2 = \mathrm{Var}(X) &= \sum_{k=1}^{3} (x_k - \mu)^2 \mathrm{P}(X = x_k) \\ &= \sum_{k=1}^{3} (k - 1\tfrac{2}{3})^2 p(k) \\ &= \left(\tfrac{2}{3}\right)^2 \cdot \tfrac{1}{2} + \left(\tfrac{1}{3}\right)^2 \cdot \tfrac{1}{3} + \left(\tfrac{4}{3}\right)^2 \cdot \tfrac{1}{6} \\ &= \frac{5}{9} \approx 0.556 \\ \sigma = \sqrt{\mathrm{Var}(X)} &= \frac{\sqrt{5}}{3} \approx 0.745 \end{aligned}$$

◁

Example 10.6. A continuous random variable X has the probability density function $f(x)$:

$$f(x) = \begin{cases} 1 + x & \text{for } -1 \leq x < 0 \\ 1 - x & \text{for } 0 \leq x < 1 \\ 0 & \text{otherwise} \end{cases}$$



What is the expectation, variance and standard deviation?

$$\begin{aligned} \mu = \mathrm{E}(X) &= \int_{-\infty}^{\infty} x\, f(x)\,\mathrm{d}x \\ &= \int_{-1}^{0} x(1+x)\,\mathrm{d}x + \int_{0}^{1} x(1-x)\,\mathrm{d}x \\ &= \left[\tfrac{1}{2}x^2 + \tfrac{1}{3}x^3\right]_{-1}^{0} + \left[\tfrac{1}{2}x^2 - \tfrac{1}{3}x^3\right]_{0}^{1} \\ &= -\tfrac{1}{2} + \tfrac{1}{3} + \tfrac{1}{2} - \tfrac{1}{3} = 0 \\ \sigma^2 = \mathrm{Var}(X) &= \int_{-\infty}^{\infty} (x-\mu)^2 f(x)\,\mathrm{d}x \\ &= \int_{-1}^{0} x^2(1+x)\,\mathrm{d}x + \int_{0}^{1} x^2(1-x)\,\mathrm{d}x \\ &= \left[\tfrac{1}{3}x^3 + \tfrac{1}{4}x^4\right]_{-1}^{0} + \left[\tfrac{1}{3}x^3 - \tfrac{1}{4}x^4\right]_{0}^{1} \\ &= \tfrac{1}{3} - \tfrac{1}{4} + \tfrac{1}{3} - \tfrac{1}{4} = \tfrac{1}{6} \approx 0.167 \\ \sigma = \sqrt{\mathrm{Var}(X)} &= \frac{1}{\sqrt{6}} \approx 0.408 \end{aligned}$$

◁

10.1.3. Modified random variable

Theorem 10.7 (Modified random variable)**.** Let X be a random variable with expectation μ, variance σ^2 and standard deviation σ. Let further be

$$X' = aX + b \qquad \text{for } a, b \in \mathbb{R}$$

be a random variable equal to X scaled by a with offset b.

We then have for the expectation μ', variance σ'^2 and standard deviation σ' of X':

$$\mu' = a\mu + b \qquad \sigma'^2 = a^2\sigma^2 \qquad \sigma' = |a|\,\sigma$$

◁

Example 10.7. An unbiased six-sided die has the numbers 0, 2, 4, 6, 8 and 10. What is the expectation μ, variance σ^2 and standard deviation σ?

For an unbiased six-sided die X_{prev} with the numbers one to six we know from example 10.4 $\mu_{\text{prev}} = 3\frac{1}{2}$, $\sigma^2_{\text{prev}} = \frac{35}{12}$ and $\sigma_{\text{prev}} = \sqrt{\frac{35}{12}}$.

We realize that:

$$X = aX_{\text{prev}} + b \quad \text{with} \quad a{=}2 \quad \text{and} \quad b{=}{-2}$$

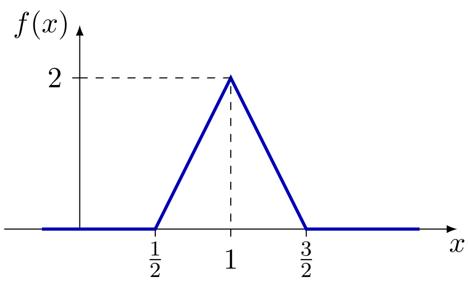
Hence we get:

$$\begin{aligned} \mu &= a\mu_{\text{prev}} + b = 2 \cdot 3\tfrac{1}{2} - 2 = 5 \\ \sigma^2 &= a^2\sigma^2_{\text{prev}} = 2^2 \cdot \tfrac{35}{12} = \tfrac{35}{3} \approx 11.67 \\ \sigma &= |a|\,\sigma_{\text{prev}} = 2\sqrt{\tfrac{35}{12}} = \sqrt{\tfrac{35}{3}} \approx 3.416 \end{aligned}$$

◁

Example 10.8. A continuous random variable X has the probability density function $f(x)$:

$$f(x) = \begin{cases} 4x - 2 & \text{for } \frac{1}{2} \le x < 1 \\ 6 - 4x & \text{for } 1 \le x < \frac{3}{2} \\ 0 & \text{otherwise} \end{cases}$$



What is the expectation μ, variance σ^2 and standard deviation σ?

We realize that the probability density function equals the one from example 10.6 except for a factor $a = \frac{1}{2}$ and an offset of $b = 1$. With μ', σ'^2 and σ' as the expectation, variance and and standard deviation of example 10.6 we get:

$$\begin{aligned} \mu &= \tfrac{1}{2}\mu' + b = 0 + 1 = 1 \\ \sigma^2 &= (\tfrac{1}{2})^2\sigma'^2 = \tfrac{1}{4} \cdot \tfrac{1}{6} = \tfrac{1}{24} \approx 0.042 \\ \sigma &= \tfrac{1}{2}\sigma' = \frac{1}{2\sqrt{6}} \approx 0.204 \end{aligned}$$

◁

10.1.4. Sum of random variables

Theorem 10.8 (Expectation of two random variables)**.** With X and Y being two random variables we have

$$\mathrm{E}(X + Y) = \mathrm{E}(X) + \mathrm{E}(Y)$$

◁

Before analysing the variance of the sum of two random variables X and Y we have to introduce the concept of *covariance*:

Definition 10.9 (Covariance)**.** With two random variables X and Y and their expectations μ_X and μ_Y, respectively, we define

$$\mathrm{Cov}(X, Y) = \mathrm{E}[(X - \mu_X)(Y - \mu_Y)]$$

as the *covariance* of X and Y. ◁

A useful expression for the covariance is given by

$$\begin{aligned} \mathrm{Cov}(X, Y) &= \mathrm{E}[(X - \mu_X)(Y - \mu_Y)] \\ &= \mathrm{E}[XY - X\mu_Y - Y\mu_X + \mu_X\mu_Y] \\ &= \mathrm{E}[XY] - \mu_Y\mathrm{E}[X] - \mu_X\mathrm{E}[Y] + \mu_X\mu_Y \\ &= \mathrm{E}[XY] - \mu_Y\mu_X - \mu_X\mu_Y + \mu_X\mu_Y \\ &= \mathrm{E}[XY] - \mu_X\mu_Y \\ &= \mathrm{E}[XY] - \mathrm{E}[X]\,\mathrm{E}[Y] \end{aligned}$$

Theorem 10.10 (Variance of two random variables)**.** The variance of the sum of two random variables X and Y is given by:

$$\mathrm{Var}(X{+}Y) = \mathrm{Var}(X) + \mathrm{Var}(Y) + 2\,\mathrm{Cov}(X, Y)$$

◁

Theorem 10.11 (Independent random variables)**.** With X and Y as two independent random variables we have:

$$\begin{aligned}\operatorname{Cov}(X,Y) &= 0\\ \operatorname{Var}(X+Y) &= \operatorname{Var}(X)+\operatorname{Var}(Y)\end{aligned}$$

◁

Definition 10.12 (Correlation)**.** With X and Y being two random variables with their standard deviations σ_X and σ_Y, respectively, we define the *correlation* as

$$\operatorname{Corr}(X,Y) = \frac{\operatorname{Cov}(X,Y)}{\sqrt{\operatorname{Var}(X)\operatorname{Var}(Y)}} = \frac{\operatorname{Cov}(X,Y)}{\sigma_X\sigma_Y}$$

We get: $\operatorname{Corr}(X,Y) \in [-1,1]$ ◁

Example 10.9. Let's assume a survey under all first term students about their attitude towards maths and whether they are musicians or not. The outcome of the survey was:

	like math	don't like math
musician	28%	12%
not musician	33%	27%

We define A_X as the event to be a musician and A_Y as the event of loving math. We give *be musician* and *like maths* the value one and the other the value zero to get two random variables X and Y. With p_{XY} we get:

$$\begin{aligned} x_1 = y_1 &= 1 & x_2 = y_2 &= 0\\ p_{11} &= 0.28 & p_{12} &= 0.12\\ p_{21} &= 0.33 & p_{22} &= 0.27\end{aligned}$$

Expectation μ, variance σ^2 and standard deviation σ for X and Y are:

$$\begin{aligned}\mu_X &= \sum_{k=1}^{2} x_k(p_{k1}+p_{k2})\\ &= x_1(p_{11}+p_{12}) + x_2(p_{21}+p_{22})\\ &= 1\cdot 0.4 + 0\cdot 0.6 = 0.4\\ \mu_Y &= \sum_{k=1}^{2} y_k(p_{1k}+p_{2k})\\ &= y_1(p_{11}+p_{21}) + y_2(p_{12}+p_{22})\\ &= 1\cdot 0.61 + 0\cdot 0.39 = 0.61\\ \sigma_X^2 &= \sum_{k=1}^{2} (x_k-\mu_X)^2(p_{k1}+p_{k2})\\ &= (1-0.4)^2\cdot 0.4 + (0-0.4)^2\cdot 0.6 = 0.24\\ \sigma_Y^2 &= \sum_{k=1}^{2} (y_k-\mu_Y)^2(p_{1k}+p_{2k})\\ &= (1-0.61)^2\cdot 0.61 + (0-0.61)^2\cdot 0.39\\ &= 0.2379\\ \sigma_X &= \sqrt{\sigma_X^2} = \sqrt{0.24} \approx 0.490\\ \sigma_Y &= \sqrt{\sigma_Y^2} = \sqrt{0.2379} \approx 0.488\end{aligned}$$

Covariance and correlation are given by:

$$\begin{aligned}\operatorname{Cov}(X,Y) &= \operatorname{E}(XY) - E(X)E(Y)\\ &= x_1y_1p_{11} + x_1y_2p_{12} + x_2y_1p_{21}\\ &\quad + x_2y_2p_{22} - \mu_X\mu_Y\\ &= 1\cdot1\cdot0.28 + 1\cdot0\cdot0.12 + 0\cdot1\cdot0.33\\ &\quad + 0\cdot0\cdot0.27 - 0.61\cdot0.4\\ &= 0.28 - 0.244 = 0.036\\ \operatorname{Corr}(X,Y) &= \frac{\operatorname{Cov}(X,Y)}{\sqrt{\sigma_X^2\sigma_Y^2}} = \frac{0.036}{\sqrt{0.2379\cdot 0.24}}\\ &\approx 0.151\end{aligned}$$

For the sum of X and Y we have:

$$\begin{aligned}\mu_{X+Y} &= \mu_X + \mu_Y = 0.4 + 0.61 = 1.01\\ \sigma_{X+Y}^2 &= \sigma_X^2 + \sigma_Y^2 + 2\operatorname{Cov}(X,Y)\\ &= 0.24 + 0.2379 + 2\cdot 0.036 = 0.5499\\ \sigma_{X+Y} &= \sqrt{\sigma_{X+Y}^2} = \sqrt{0.5499} \approx 0.742\end{aligned}$$

Since $\operatorname{Corr}(X,Y) \neq 0$ the two random variables X and Y are statistically dependent. ◁

Definition 10.13 (independent and identically distributed)**.** We call two random variables X and Y *independent and identically distributed*, *i.i.d.*, if their covariance is zero and they have the same probability density function $f_X(x)$ and $f_Y(x)$ or the same probability mass function $p_X(x)$ and $p_Y(x)$, i.e.

$$\begin{gathered}\operatorname{Cov}(X,Y) = 0\\ f_X(x) = f_Y(x) \quad \text{or} \quad p_X(x) = p_Y(x)\end{gathered}$$

◁

10.2. Discrete random distributions

10.2.1. Binomial distribution

Suppose we perform n i.i.d. experiments each of them having the outcome *success* with a probability p. Let us assume the random variable X gives the number of successes for n trials. I.e. we combine n experiments to one random variable X with sample space $\Omega = \{0, 1, 2, \ldots, n\}$.

Theorem 10.14 (Binomial distribution)**.** Let X_b be a random variable representing the number of successes of n i.i.d. experiments. Each experiment has a success probability p. Then the probability $\mathrm{P}(X_b{=}k)$ for exactly k successes is given by

$$\mathrm{P}(X_b{=}k) = \binom{n}{k} p^k (1-p)^{n-k}$$

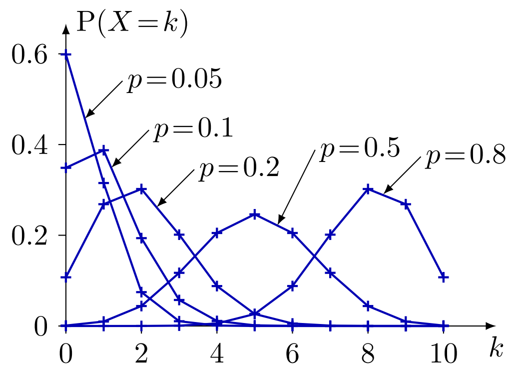
$$\text{where} \qquad \binom{n}{k} = \frac{n!}{k!(n-k)!}$$

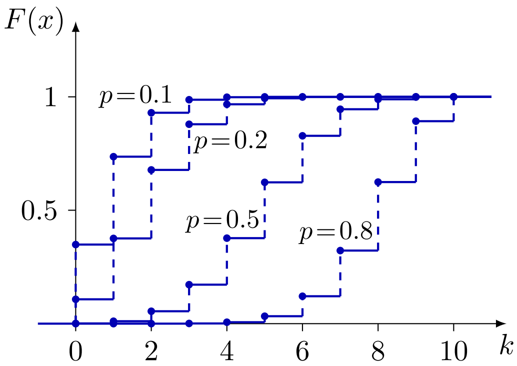
We call this a *binomial random variable* with *binomial distribution*. Expectation and variance are given by

$$\mu_b = \mathrm{E}(X_b) = np$$
$${\sigma_b}^2 = \mathrm{Var}(X_b) = np(1-p)$$

◁

Example 10.10. The following graphs show the binomial distribution for $n = 10$ experiments and different single experiment probabilities p:





◁

Example 10.11. Let X be the number of ones thrown by two unbiased six-sided dice with the numbers 1 − 6. We have $p = \frac{1}{6}$ as the probability of throwing a one with one die. For the events A_0, A_1 and A_2 of throwing zero, one or two ones the probabilities are:

$$\mathrm{P}(A_0) = \binom{2}{0}(\tfrac{1}{6})^0(1-\tfrac{1}{6})^2 = \tfrac{25}{36} \approx 69.4\%$$
$$\mathrm{P}(A_1) = \binom{2}{1}(\tfrac{1}{6})^1(1-\tfrac{1}{6})^1 = \tfrac{10}{36} \approx 27.8\%$$
$$\mathrm{P}(A_2) = \binom{2}{2}(\tfrac{1}{6})^2(1-\tfrac{1}{6})^0 = \tfrac{1}{36} \approx 2.8\%$$

◁

10.2.2. Poisson distribution

As for a binomial distribution we perform n identical experiments with the number of *successes* being a random variable X.

We now increase the number of experiments n and decrease the probability of success of a single experiment p in a way that the product $\lambda = np$ remains constant. For the limit $n \to \infty$ the random variable X takes on the *Poisson distribution*:

Definition 10.15 (Poisson distribution)**.** A random variable X_P with sample space $\Omega = \{0, 1, 2, \ldots\}$, $\lambda \in \mathbb{R}_{>0}$ and probability mass function

$$p(k) = \mathrm{P}(X_P{=}k) = e^{-\lambda}\frac{\lambda^k}{k!}, \qquad k = 0, 1, 2, \ldots$$

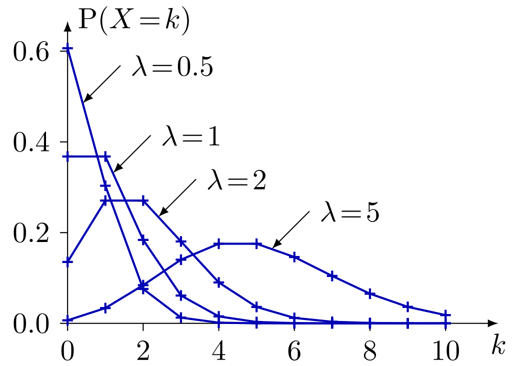
is said to be a *Poisson random variable* with *Poisson distribution*. ◁

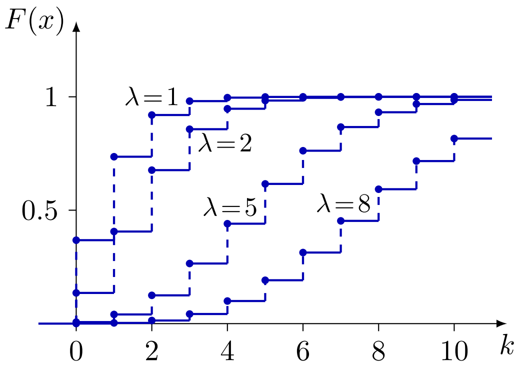
Theorem 10.16 (Poisson distribution)**.** With the conditions of the previous definition expectation and variance are given by:

$$\mathrm{E}(X_P) = \mathrm{Var}(X_P) = \lambda$$

◁

Example 10.12. The following graphs show Poisson distributions for different λ:

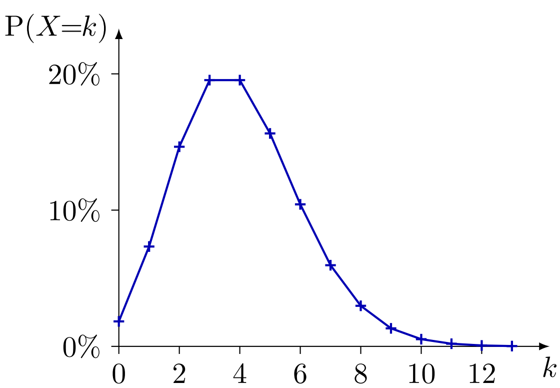




◁

Example 10.13. A shop expects Saturday afternoon on average four customers per minute. The random variable X as the number of customers per minute then has a Poisson distribution. With $\lambda = 4$ we get the following probabilities:

k	$P(X=k)$	k	$P(X=k)$
0	1.83%	7	5.95%
1	7.33%	8	2.98%
2	14.65%	9	1.32%
3	19.54%	10	0.53%
4	19.54%	11	0.19%
5	15.63%	12	0.06%
6	10.42%	13	0.02%



◁

10.3. Continuous random distributions

10.3.1. Uniform distribution

Definition 10.17 (Uniform distribution)**.** For $a, b \in \mathbb{R}$ and $a < b$ the continuous random variable X_u with probability density function

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{for } a \leq x < b \\ 0 & \text{otherwise} \end{cases}$$

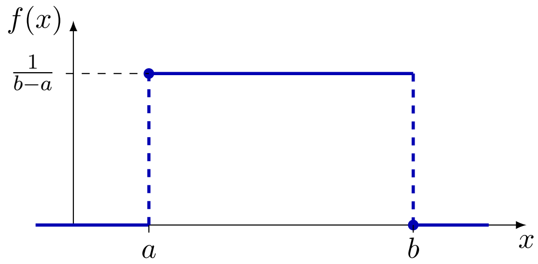
is said to be a *uniform random variable* with *uniform distribution.* ◁

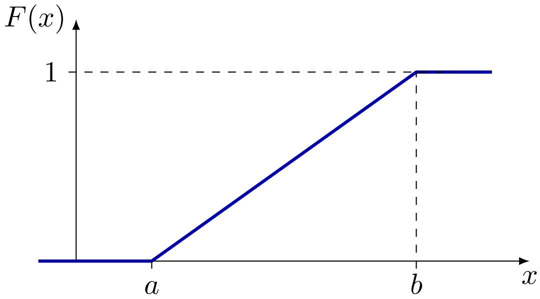
Theorem 10.18 (Uniform distribution)**.** For the definition above expectation and variance are given by

$$\mu_u = \mathrm{E}(X_u) = \frac{a+b}{2}$$

$$\sigma_u{}^2 = \mathrm{Var}(X_u) = \frac{(b-a)^2}{12}$$

◁





Example 10.14. Pseudo random number generator

Many computer languages provide a so called *pseudo random number generator*, PRNG to generate some sort of random numbers. A PRNG provides numbers on a given interval with uniform distribution.

Although talking about random numbers, a PRNG is deterministic, i.e. starting the PRNG a second time with the same parameters will result in exactly the same numbers. Nevertheless, the generated numbers *behave* like random numbers in terms of probability, expectation and variance.

A simple implementation of a PRNG is a *linear congruential generator*: The next number x_{k+1} is evaluated from the previous number x_k by

$$x_{k+1} = (a\,x_k + c) \mod m$$

with $a, c, m \in \mathbb{N}$. The parameters a, c and m must be chosen thoroughly to gain an acceptable quality.

In order to gain unpredictable random numbers a useful technique is to initialize the random number generator by the actual time. I.e. the first random number x_0 is set to a desired value derived from the computer time.

Strictly speaking a PRNG represents a discrete random variable X with discrete sample space $\Omega = \{0, 1, 2, \ldots, x_{\max}\}$. However, for many applications the PRNG is sufficient as input for a continuous random variable.

E.g. dividing the pseudo random variable X by the maximum value $x_{\max}$ provides a random variable with values from zero to one with a high density. ◁

10.3.2. Normal distribution

Definition 10.19 (Normal distribution)**.** For $\mu, \sigma \in \mathbb{R}$ and $\sigma > 0$ a random variable X_n with probability density function

$$f(x) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

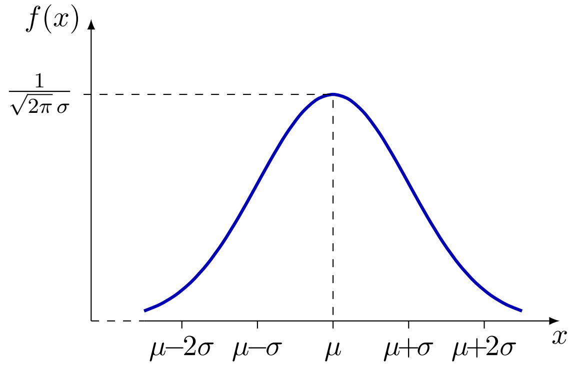
is said to be a *normal random variable* with *normal distribution.* ◁

Theorem 10.20 (Normal distribution)**.** For the definition above expectation and variance are given by

$$\mathrm{E}(X_n) = \mu \qquad \mathrm{Var}(X_n) = \sigma^2$$

◁

The normal probability density function is a bell shaped function symmetric around $x = \mu$:



If X_n is a normal random variable with expectation μ and standard deviation σ, then aX_n+b, $a, b \in \mathbb{R}$, $a > 0$ again is a normal random variable.

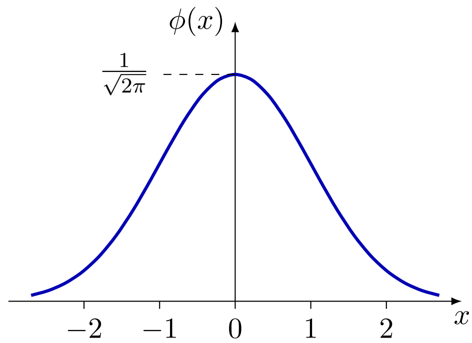
With $a = \frac{1}{\sigma}$ and $b = -\frac{\mu}{\sigma}$ we get the standard normal random variable X_{sn}:

$$X_{sn} = \frac{X_n}{\sigma} - \frac{\mu}{\sigma} = \frac{X_n - \mu}{\sigma}$$

Definition 10.21 (Standard normal distribution)**.** We define a random variable X_{sn} with probability density function

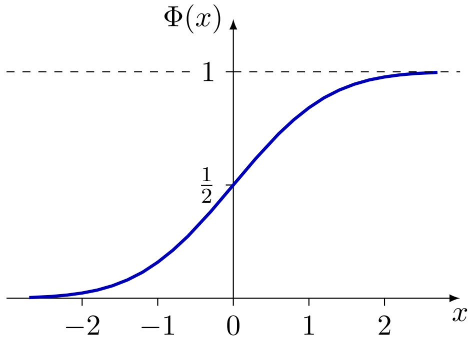
$$f_{sn}(x) = \phi(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

as a *standard normal random variable* with *standard normal distribution.* ◁



The cumulative distribution function $\Phi(x) = F_{sn}(x)$ is the integral of the probability density function $\phi(x)$ from $-\infty$ to x:

$$\Phi(x) = \int_{-\infty}^{x} \phi(x')\,\mathrm{d}x' = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-x'^2/2}\,\mathrm{d}x'$$



For the random variable with standard normal distribution we have:

$$\mu_{sn} = 0 \qquad \sigma_{sn}^2 = \sigma_{sn} = 1$$

Values for the cumulative distribution function may either be evaluated numerically or derived from tables as the one below. Since $\phi(x)$ is symmetric around $x = 0$ we have $\Phi(-x) = 1 - \Phi(x)$ and, hence, we limit the table to arguments $x \geq 0$.

x	$\Phi(x)$	x	$\Phi(x)$	x	$\Phi(x)$
0.0	0.50000	1.2	0.88493	2.4	0.99180
0.1	0.53983	1.3	0.90320	2.5	0.99379
0.2	0.57926	1.4	0.91924	2.6	0.99534
0.3	0.61791	1.5	0.93319	2.7	0.99653
0.4	0.65542	1.6	0.94520	2.8	0.99744
0.5	0.69146	1.7	0.95543	2.9	0.99813
0.6	0.72575	1.8	0.96407	3.0	0.99865
0.7	0.75804	1.9	0.97128	3.2	0.99931
0.8	0.78814	2.0	0.97725	3.4	0.99966
0.9	0.81594	2.1	0.98214	3.6	0.99984
1.0	0.84134	2.2	0.98610	3.8	0.99993
1.1	0.86433	2.3	0.98928	4.0	0.99997

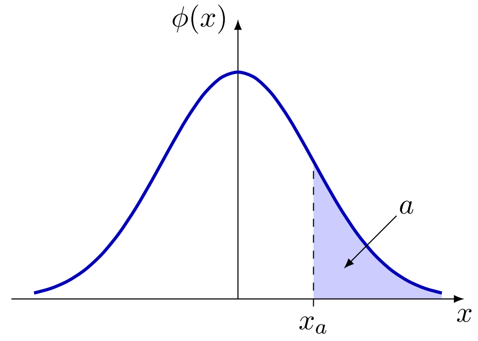
The cumulative distribution function $F_n(x)$ of a normal distributed random variable X_n with expectation μ and standard deviation σ may be expressed by the cumulative distribution function $\Phi(x)$ of the standard normal random variable X_{sn}:

$$\begin{aligned} F_n(x) &= \mathrm{P}(X_n \leq x) = \mathrm{P}\left(X_n - \mu \leq x - \mu\right) \\ &= \mathrm{P}\left(\frac{X_n - \mu}{\sigma} \leq \frac{x - \mu}{\sigma}\right) \\ &= \mathrm{P}\left(X_{sn} \leq \frac{x - \mu}{\sigma}\right) = \Phi\left(\frac{x - \mu}{\sigma}\right) \end{aligned}$$

Typical questions are: What is the probability a for a standard normal random variable X_{sn} to be larger than x_a?

$$a = \mathrm{P}(X_{sn} > x_a) = 1 - \Phi(x_a)$$

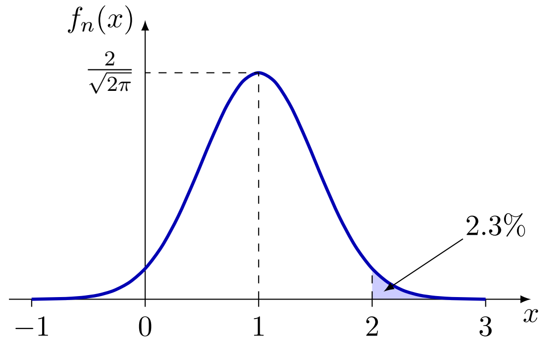
Conversely we may ask: What is the value x_a for a standard normal random variable X_{sn} to be larger than with a probability a? E.g. $x_{0.1} = 1.282$ is the value for a standard normal random variable to be larger than with a probability of 10%.



Example 10.15. What is the probability of a normal random variable with expectation 1 and standard deviation 0.5 to be larger than 2?

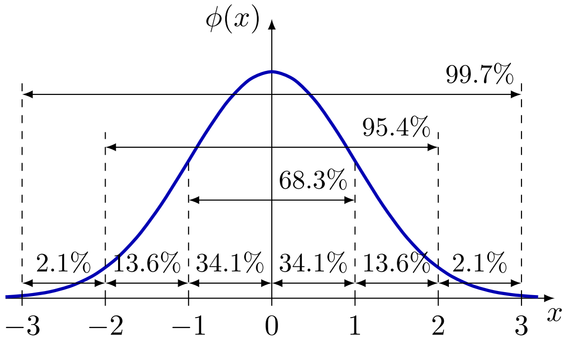
With X_n as the normal random variable, $\mu = 1$ as the expectation, $\sigma = 0.5$ as the standard deviation, $x_a = 2$ as the limit and a as the probability of X_n being greater than x_a we get:

$$\begin{aligned} a &= \mathrm{P}(X_n > x_a) = 1 - \mathrm{P}(X_n \leq x_a) \\ &= 1 - \mathrm{P}(X_n - \mu \leq x_a - \mu) \\ &= 1 - \mathrm{P}\left(\frac{X_n - \mu}{\sigma} \leq \frac{x_a - \mu}{\sigma}\right) \\ &= 1 - \mathrm{P}\left(X_{sn} \leq \frac{x_a - \mu}{\sigma}\right) \\ &= 1 - \Phi\left(\frac{x_a - \mu}{\sigma}\right) = 1 - \Phi\left(\frac{2-1}{0.5}\right) \\ &= 1 - \Phi(2) = 1 - 0.977 = 2.3\% \end{aligned}$$



◁

The probabilities of a normal random variable to be within the limits of $\pm\sigma$, $\pm 2\sigma$ etc. are also of interest. The following graph illustrates some of these probabilities.



E.g. the probabilitiy of X_{sn} to be within $\pm\sigma$, i.e. ± 1 is:

$$\begin{aligned} \mathrm{P}(|X_{sn}| < 1) &= \mathrm{P}(-1 < X_{sn} < 1) \\ &= \mathrm{P}(X_{sn} < 1) - \mathrm{P}(X_{sn} \leq -1) \\ &= \mathrm{P}(X_{sn} \leq 1) - \mathrm{P}(X_{sn} \leq -1) \\ &= \Phi(1) - \Phi(-1) \\ &= \Phi(1) - 1 + \Phi(1) \\ &= 2\,\Phi(1) - 1 \\ &= 2 \cdot 0.84134 - 1 \\ &= 0.68264 \mathrel{\hat{=}} 68.3\% \end{aligned}$$

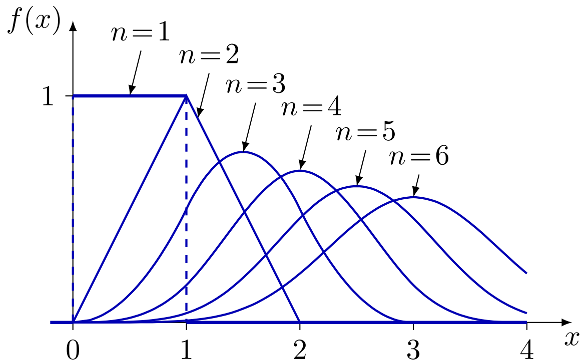
Why do we focus so much on normal random variables? If we deal with the sum of a large number of i.i.d. random variables we approach the normal distribution:

Theorem 10.22 (Central limit theorem)**.** Let $X_1, X_2, \ldots$ be a sequence of i.i.d. random variables with expectation μ and standard deviation σ. Then

$$\lim_{n\to\infty} \mathrm{P}\left(\frac{X_1+X_2+\ldots+X_n - n\mu}{\sigma\sqrt{n}} \leq x\right) = \Phi(x)$$

◁

Example 10.16. The following diagram shows the probability density functions of a sum of i.i.d. uniform distributed random variables each on the interval $[0, 1)$. With increasing number the expectation, variance and standard deviation increase.



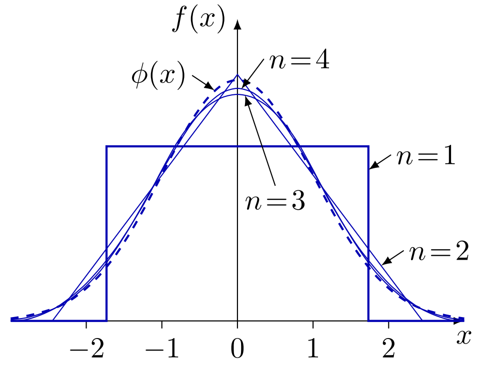
For a single random variable we get:

$$\mu_1 = \frac{1}{2} \qquad \sigma_1^2 = \frac{1}{12} \qquad \sigma_1 = \frac{1}{\sqrt{12}}$$

For the sum of n random variables we get:

$$\begin{aligned} \mu_n &= n\mu_1 = \frac{n}{2} \\ \sigma_n^2 &= n\sigma_1^2 = \frac{n}{12} \\ \sigma_n &= \sqrt{n}\,\sigma_1 = \sqrt{\frac{n}{12}} \end{aligned}$$

If we subtract from the sum of n random variables the expectation μ_n and divide by the standard deviation σ_n we get the graphs illustrated in the following diagram:



◁

10.3.3. Exponential distribution

Definition 10.23 (Exponential distribution)**.** For $\lambda \in \mathbb{R}$, $\lambda > 0$ a random variable X_e with probability density function

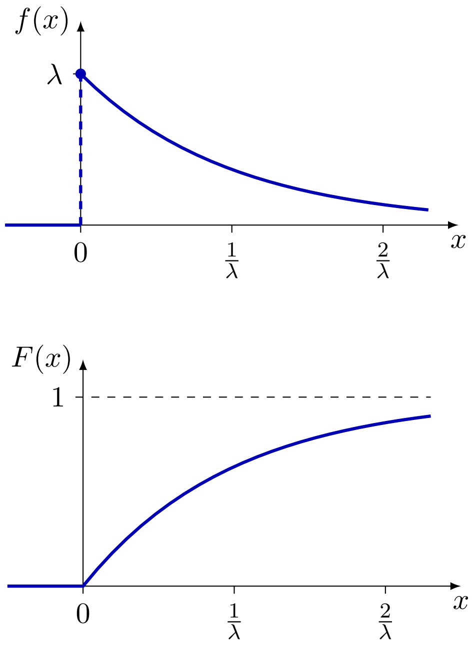
$$f(x) = \begin{cases} \lambda\, e^{-\lambda x} & \text{for } x \geq 0 \\ 0 & \text{otherwise} \end{cases}$$

is said to be an *exponential random variable* with *exponential distribution*. ◁

For $x \geq 0$ the cumulative distribution function becomes:

$$\begin{aligned} F(x) &= \int_{-\infty}^{x} f(x')\,\mathrm{d}x' = \lambda \int_0^x e^{-\lambda x'}\,\mathrm{d}x' \\ &= \lambda \left[\tfrac{1}{-\lambda} e^{-\lambda x'}\right]_0^x = 1 - e^{-\lambda x} \end{aligned}$$

I.e. $$F(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1 - e^{-\lambda x} & \text{for } x \geq 0 \end{cases}$$



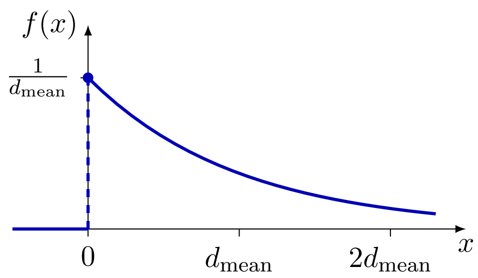
Theorem 10.24 (Exponential distribution)**.** A random variable X_e with exponential distribution has the expectation and variance:

$$E(X_e) = \frac{1}{\lambda} \qquad \text{Var}(X_e) = \frac{1}{\lambda^2}$$

◁

Example 10.17. When shooting randomly into a forest the bullet will hit earlier or later a tree. The *free path* of the bullet d may be looked at as a random variable. On average the bullet will fly a distance d_{mean} which is the *mean free path*. This mean free path d_{mean} depends on the thickness of the trees and the number of trees per area.

The random variable for the free path has *exponential distribution* with $\lambda = \frac{1}{d_{\text{mean}}}$. I.e. Most bullets will be found in the trees close to the shooter. The larger the distance to the shooter the less bullets will be found.



◁

10.4. Problems

Problem 10.1: A random variable has a constant probability over its sample space $\Omega = [-1, 0)$. Plot the cumulative distribution function and the probability density function.

Problem 10.2: An unbiased six-sided die has the numbers 1, 1, 2, 2, 3 and 4 on its six faces. Plot the cumulative distribution function.

Problem 10.3: A continuous random variable X has the probability density function $f(x)$:

$$f(x) = \begin{cases} 0 & \text{for } x < -1 \\ \frac{1}{2} & \text{for } -1 \leq x < 1 \\ 0 & \text{for } 1 \leq x \end{cases}$$

Plot the cumulative distribution function and the probability density function. Evaluate expectation, variance and standard deviation.

Problem 10.4: A continuous random variable X has the probability density function $f(x)$:

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

Evaluate expectation, variance and standard deviation.

Problem 10.5: For a random variable X with sample space $\Omega = [0, 1)$ and constant probability density function $f(x)$ over the whole sample space, i.e.

$$f(x) = \begin{cases} 0 & \text{for } x < 0 \\ 1 & \text{for } 0 \leq x < 1 \\ 0 & \text{for } 1 \leq x \end{cases}$$

plot the cumulative distribution function and the probability density function.

Evaluate expectation, variance and standard deviation. Hint: Make use of the results from the previous problem.

Problem 10.6: A continuous random variable X has the probability density function $f(x)$:

$$f : \begin{cases} \mathbb{R} \to \mathbb{R} & \\ x \mapsto 0 & \text{for } x < 0 \\ x \mapsto a\, e^{-ax} & \text{for } x \geq 0 \end{cases}$$

Plot the cumulative distribution function and the probability density function.

Evaluate expectation, variance and standard deviation. Find these values a) by applying definitions 10.5 and 10.6 and b) by scaling the results from previous problems.

Problem 10.7: For a six-sided unbiased die with the numbers one to six we define two random variables X and Y:

X returns one for even numbers and zero otherwise. Y returns one for the events of throwing a 4, 5 or 6 and zero otherwise.

1. Evaluate expectation μ, variance σ^2 and standard deviation σ for X, Y.
2. Evaluate covariance and correlation for X and Y.
3. For $X+Y$ evaluate expectation μ, variance σ^2 and standard deviation σ.
4. Are X and Y independent?

Problem 10.8: For three unbiased standard dice we define the random variable X as the number of ones and/or sixes thrown. I.e. the sample space of X is $\Omega = \{0, 1, 2, 3\}$.

1. Evaluate the probability of X for all elements of Ω.
2. Evaluate expectation μ, variance σ^2 and standard deviation σ of X.

Problem 10.9: You measure with a Geiger-Mueller counter on average three quanta per second. a) What are the probabilities to measure $0, 1, 2, \ldots, 7$ quanta per second? b) What is the expectation, variance and standard deviation for the number of measured quanta per second?

Problem 10.10: For a uniform random variable X_u on the interval $[a, b)$ prove the equations for expectation μ_u and variance ${\sigma_u}^2$:

$$\mu_u = \mathrm{E}(X_u) = \frac{a+b}{2}$$
$${\sigma_u}^2 = \mathrm{Var}(X_u) = \frac{(b-a)^2}{12}$$

Problem 10.11: For a normal distributed random variable X_n with expectation $\mu = 3$ and standard deviation $\sigma = 1$, what is the probability for X_n being

1. greater than four?
2. less than one?
3. between two and three?
4. outside $\pm 2\sigma$ around μ?

Problem 10.12: You have the sum of 100 i.i.d. random variables each with expectation μ_1 and standard variation σ_1. Approximate the sum by a normal random variable X_n defined by its probability density function $f_n(x)$.

Problem 10.13: A company produces a large number of cylinders with specified diameter of 60 ± 0.02 mm. The actual diameter shows a mean value of 60.01 mm with a standard deviation of 0.01 mm. Assuming normal distribution, what is the failure rate?

Problem 10.14: A company produces resistors that are specified with 1% tolerance. Any resistor with a higher deviation is treated as failure.

Assuming a normal distribution for the actual resistance, what standard deviation σ is acceptable to achieve a failure rate below 1%?

Problem 10.15: For an exponential distributed random variable X with expectation $\mu = 3$, what is the probability for X being

1. greater than four?
2. less than one?
3. between two and three?
4. outside $\pm\sigma/2$ around μ?

Problem 10.16: The half-life of the radioactive plutonium nuclide Pu-239 is 24 110 years. Evaluate expectation, variance and standard deviation of a random variable that describes the radioactivity.

A. Solutions

A.1. Integral calculus

Solution 1.1:

$$R_f(4) = \tfrac{\pi}{4}(1+\sqrt{2})$$
$$U_f(4) = \tfrac{\pi}{4}(2+\sqrt{2})$$
$$L_f(4) = \tfrac{\pi}{4}\sqrt{2}$$

Solution 1.2:

$$R_f(4) = 12 \qquad U_f(4) = 20 \qquad L_f(4) = 12$$

Solution 1.3:

$$R_f(n) = \tfrac{16(n-1)}{n}$$
$$U_f(n) = \tfrac{16(n+1)}{n}$$
$$L_f(n) = R_f(n) = \tfrac{16(n-1)}{n}$$

n	1	10	100	1000	∞
$R_f(n)$	0	14.4	15.84	15.984	16
$U_f(n)$	32	17.6	16.16	16.016	16
$L_f(n)$	0	14.4	15.84	15.984	16

Solution 1.4:

1. 8
2. 10
3. $\frac{15}{2}$
4. $\frac{3}{2}$
5. $\frac{14}{3}$
6. $\frac{2}{3}b^3 - \frac{1}{2}b^2$
7. $2(b-a) + \frac{1}{3}b^3 - \frac{1}{3}a^3$
8. $2(a-b) + \frac{1}{3}a^3 - \frac{1}{3}b^3$

Solution 1.5:

1. $\frac{13}{2}$
2. $-\frac{13}{2}$
3. $\frac{13}{2}$
4. 1
5. $\frac{4}{3}$
6. -1
7. $\frac{4b^2-a^2}{2}$
8. $\frac{b^3+a^3}{3}$

Solution 1.6:

1. $x^3 - {x_0}^3$
2. $\sin(x) - \sin(x_0)$
3. $\cos(x_0) - \cos(x)$
4. $e^x - e^{x_0}$
5. $\sin(2\pi x) - \sin(2\pi x_0)$
6. $e^{2\mathrm{j}\pi f\, x} - e^{2\mathrm{j}\pi f\, x_0}$
7. $\sin\left(x^2\right) - \sin\left({x_0}^2\right)$
8. $\cos\left({x_0}^3\right) - \cos\left(x^3\right)$

Solution 1.7: For any constant C:

$$f_1(x) = x + C \qquad f_5(x) = -\cos(x) + C$$
$$f_2(x) = x^2 + C \qquad f_6(x) = \sin(x) + C$$
$$f_3(x) = \tfrac{1}{3}x^3 + C \qquad f_7(x) = e^x + C$$
$$f_4(x) = \tfrac{1}{4}x^4 + C \qquad f_8(x) = \tfrac{1}{3}x^3 - \tfrac{1}{2}x^2 + x + C$$

Solution 1.8:

1. C
2. $ax + C$
3. $\frac{1}{3}x^3 - x + C$
4. $3e^x + C$
5. $\frac{1}{\ln(5)}5^x + C$
6. $\frac{1}{\ln(\pi)}\pi^x + C$
7. $2\sinh(x) + C$
8. $\frac{1}{2}\cosh(x) + C$

Solution 1.9:

1. $\frac{5}{2}$
2. 4
3. $\frac{1}{3}x^3 + C$
4. $e^4 + e^1 - 2 \approx 55.3$
5. 4
6. $\frac{1}{2\pi}\sin(2\pi x) + C$
7. $-\mathrm{j}e^{\mathrm{j}x} + C$
8. $\frac{1}{3}\tan(3x) + C$

Solution 1.10:

1. $\frac{2}{3}\sqrt{x^3} + C$
2. $\frac{2}{5}\sqrt{x^5} + C$
3. $2\sqrt{x} + C$
4. $\frac{3}{4}\sqrt[3]{x^4} + C$
5. $3\sqrt[3]{x} + C$
6. $x(\ln(x) - 1) + C$
7. $\frac{1}{\ln(10)}x(\ln(x) - 1) + C$
8. $\frac{1}{\ln(2)}x(\ln(x) - 1) + C$

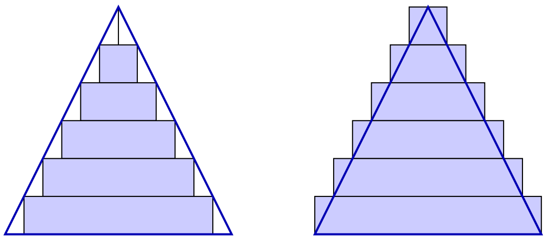
A.2. Application of integrals

Solution 2.1: $m = 1.2\,\mathrm{kg}$

Solution 2.2: $A = 32\sqrt{2}\,\mathrm{cm}^2 \approx 45.3\,\mathrm{cm}^2$

Solution 2.3: $A = \frac{nd^2}{8} \sin\left(\frac{2\pi}{n}\right)$

Solution 2.4: Ansatz:



$V_L(n) = \sum_{k=0}^{n-1} \frac{h}{n}\pi\left(\frac{kr}{n}\right)^2$ $\quad$ $V_U(n) = \sum_{k=1}^{n} \frac{h}{n}\pi\left(\frac{kr}{n}\right)^2$

Solution 2.5: Possible ansatz:

$$
\begin{aligned}
V &= \int_0^h \pi\left(r_1 - (r_1 - r_2)\tfrac{x}{h}\right)^2 \,\mathrm{d}x \\
&\cdots \\
&= \tfrac{\pi h}{3}({r_1}^2 + r_1 r_2 + {r_2}^2)
\end{aligned}
$$

Solution 2.6:

1. $V_L(n) = \sum_{k=0}^{n-1} 4\pi\left(\frac{kr}{n}\right)^2 \frac{r}{n}$
2. $V_U(n) = \sum_{k=1}^{n} 4\pi\left(\frac{kr}{n}\right)^2 \frac{r}{n}$
3. $V = \lim_{n\to\infty} V_L(n) = \lim_{n\to\infty} V_U(n) = \frac{4}{3}\pi r^3$
4. $V = \int_0^r 4\pi x^2 \,\mathrm{d}x = \frac{4}{3}\pi r^3$

Solution 2.7:

$$\overline{u} = \tilde{u} = \hat{u}$$

Solution 2.8: Hint: due to symmetry it is sufficient to concentrate on the first quarter of a period only.

$$\overline{u} = \frac{\hat{u}}{2} \qquad \tilde{u} = \frac{\hat{u}}{\sqrt{3}}$$

A.3. Integration techniques

Solution 3.1:

1. 1
2. 2
3. $\frac{3}{\sqrt[3]{2}}$
4. 2
5. $\frac{1}{k-1}$
6. $\frac{2^{k+1}}{k+1}$
7. $\frac{4}{3}$
8. 0

Solution 3.2:

1. $\sin(x^2) + C$
2. $e^{\sin\varphi} + C$
3. $\cosh(x^3) + C$
4. $\sin(\sin x) + C$
5. $\frac{1}{2}\ln(x^2+1) + C$
6. $\frac{1}{a}e^{ax+b} + C$

Solution 3.3:

1. $\sin(3x+1) + C$
2. $\cos(1-x) + C$
3. $\frac{3}{4}\left(\sqrt[3]{81} - 1\right) \approx 2.495$
4. $\frac{\sqrt{8}}{3}$
5. 1
6. $e^{2\pi \mathrm{j} f t} + C$

Solution 3.4:

1. $\frac{1}{2}\ln|x^2+1| + C$
2. $\ln|\cosh(x)| + C$
3. $\ln(2) \approx 0.6931$
4. $2\ln\left|\sqrt{x^3} + x^2\right| + C$
5. $-\frac{1}{2}\ln|x^4 - x^2| + C$
6. $\frac{\ln(2)}{2} \approx 0.3466$

Solution 3.5:

1. $\sin(3) \approx 0.1411$
2. $\frac{2e-2}{3} \approx 1.146$
3. $e^{\sin(x)} + C$
4. $C - \frac{1}{3}\cos(x^3)$
5. 0
6. $2\ln|\sin(x)| + C$

Solution 3.6:

1. $(x-1)e^x + C$
2. $C - (x^3 + 3x^2 + 6x + 6)e^{-x}$

3. $(x^2 - 2)\sin(x) + 2x\cos(x) + C$
4. $x\cosh(x) - \sinh(x) + C$
5. $\left(x - \frac{1}{\pi}\right) e^{\pi x} + C$
6. $\frac{1}{2}(x - \sin(x)\cos(x)) + C$

Solution 3.7:

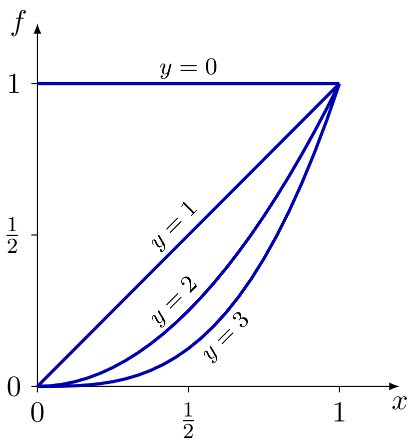
1. $\frac{1}{4}$	3. 4	5. 4
2. 4	4. $\frac{8}{3}$	6. 1

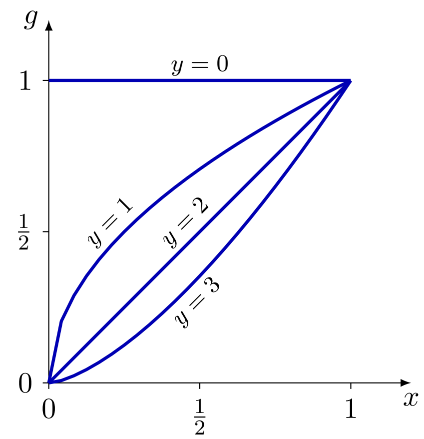
Solution 3.8:

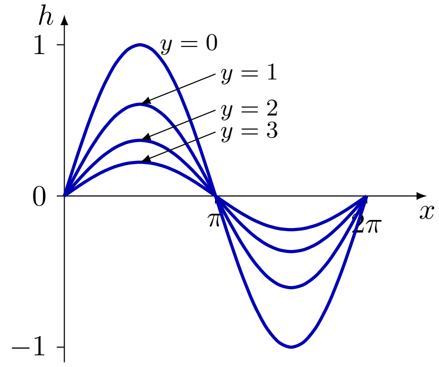
1. $\frac{1}{2}\{\ln|x-1| - \ln|x+1|\} + C$
2. $\dfrac{1}{x+2} + 2\ln|x+2| + C$
3. $\frac{1}{2}\ln|x| + \frac{1}{4}\ln|x^2+2x+2|$
   $- \frac{3}{2}\arctan(x+1) + C$
4. $\frac{1}{5}\ln|x-1| + \ln|x+1| - \frac{7}{5}\arctan(x+1)$
   $- \frac{1}{10}\ln|x^2+2x+2| + C$

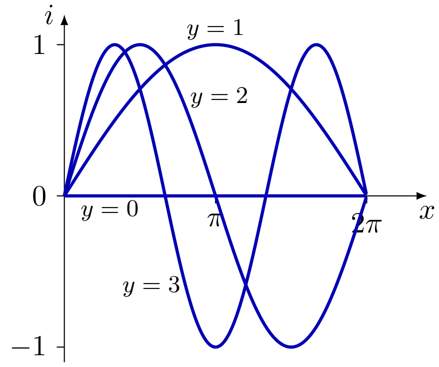
A.4. Functions with multiple arguments and values

Solution 4.1:

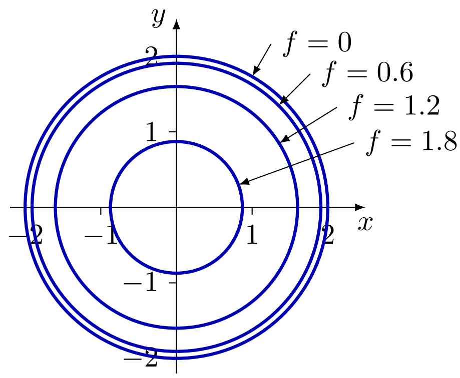


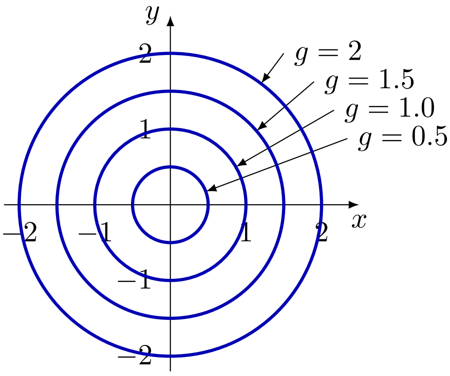


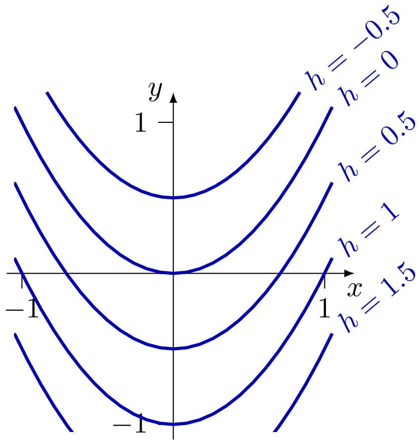




Solution 4.2:

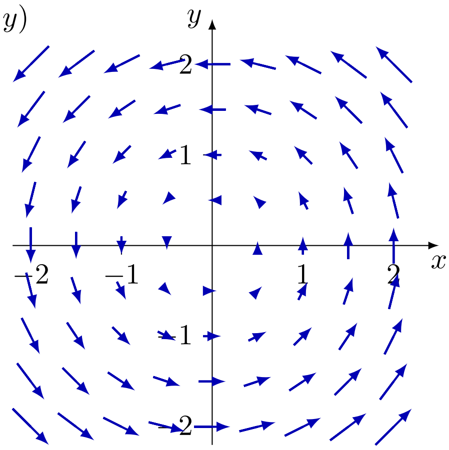




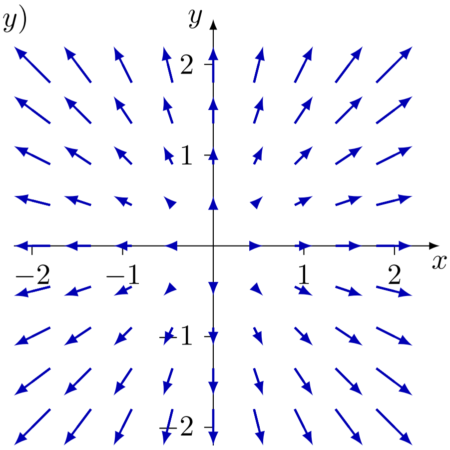


Solution 4.3:

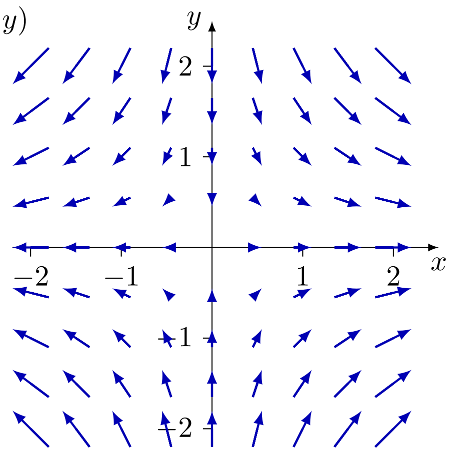
$\boldsymbol{f}(x, y)$



$\boldsymbol{g}(x, y)$



$\boldsymbol{h}(x, y)$



Solution 4.4: All the functions are continuous in their domain.

Solution 4.5:

1. $\frac{\partial f}{\partial x} = 1 \qquad \frac{\partial f}{\partial y} = 2y \qquad \frac{\partial f}{\partial z} = 3z^2$
2. $\frac{\partial f}{\partial x} = y\cos(xy)$
   $\frac{\partial f}{\partial y} = x\cos(xy)$
   $\frac{\partial f}{\partial z} = -2z\sin(z^2)$
3. $\frac{\partial f}{\partial x} = \frac{e^{x+y}}{z^2 + z + 1}$
   $\frac{\partial f}{\partial y} = \frac{e^{x+y}}{z^2 + z + 1}$
   $\frac{\partial f}{\partial z} = -\frac{(2z+1)e^{x+y}}{(z^2 + z + 1)^2}$
4. $\frac{\partial f}{\partial a} = \frac{b}{cd}$
   $\frac{\partial f}{\partial b} = \frac{a}{cd}$
   $\frac{\partial f}{\partial c} = -\frac{ab}{c^2 d}$
   $\frac{\partial f}{\partial d} = -\frac{ab}{cd^2}$
5. $\frac{\partial f}{\partial u} = \frac{1}{u}$
   $\frac{\partial f}{\partial v} = -we^{vw}$
   $\frac{\partial f}{\partial w} = \frac{1}{w} - ve^{vw}$
6. $\frac{\partial f}{\partial \alpha} = -\frac{2\pi}{\gamma^2}\sin(2\pi\alpha)$
   $\frac{\partial f}{\partial \beta} = \frac{2\pi \mathrm{j}}{\gamma^2}\cos(2\pi\beta)$
   $\frac{\partial f}{\partial \beta} = -2\frac{\cos(2\pi\alpha) + \mathrm{j}\sin(2\pi\beta)}{\gamma^3}$

Solution 4.6:

1. $\frac{\partial f}{\partial x} = \sum_k k x^{k-1} \qquad \frac{\partial f}{\partial y} = \sum_k k y^{k-1}$
2. $\frac{\partial f}{\partial x} = \sum_k k x^{k-1} y^k \qquad \frac{\partial f}{\partial y} = \sum_k k x^k y^{k-1}$
3. $\frac{\partial f}{\partial x} = -y \sum_k k e^{-kxy}$

   $\frac{\partial f}{\partial y} = -x \sum_k k e^{-kxy}$
4. $\frac{\partial f}{\partial a_0} = 2 \sum_k (a_1 x_k + a_0 - y_k)$

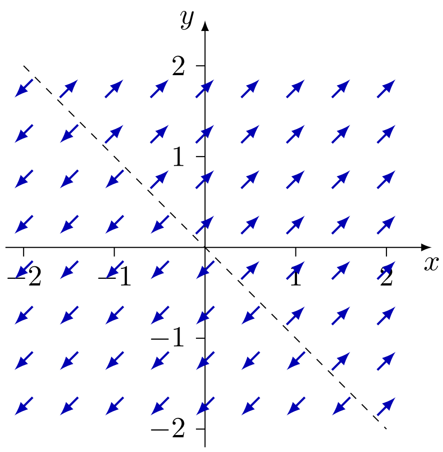
   $\frac{\partial f}{\partial a_1} = 2 \sum_k x_k (a_1 x_k + a_0 - y_k)$

Solution 4.7:

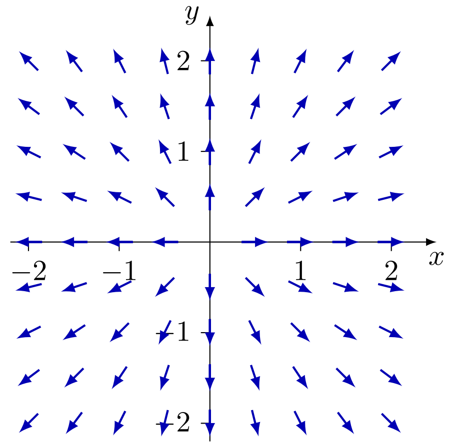
1. $\operatorname{grad}(f) = \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \boldsymbol{e_x} + \boldsymbol{e_y}$
2. $\operatorname{grad}(f) = \begin{pmatrix} \frac{x}{\sqrt{x^2+y^2}} \\ \frac{y}{\sqrt{x^2+y^2}} \end{pmatrix} = \frac{x\boldsymbol{e_x} + y\boldsymbol{e_y}}{\sqrt{x^2+y^2}}$
3. $\operatorname{grad}(f) = \begin{pmatrix} 2xe^y + 3x^2\cos(z) \\ x^2 e^y \\ -x^3 \sin(z) \end{pmatrix}$

Solution 4.8:

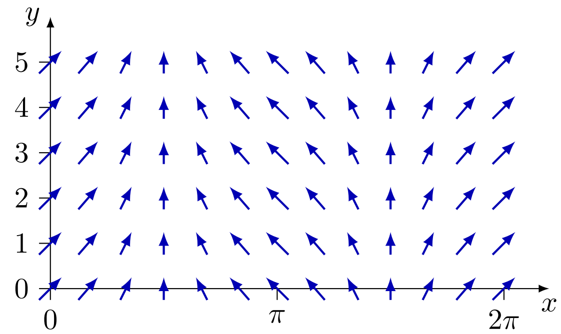
1. $\operatorname{grad}(f) = \frac{x+y}{|x+y|}(\boldsymbol{e}_x + \boldsymbol{e}_y) \quad \text{for} \quad x \neq -y$



2. $\operatorname{grad}(f) = \frac{x\boldsymbol{e}_x + y\boldsymbol{e}_y}{\sqrt{x^2+y^2}} \quad \text{for} \quad x^2 + y^2 \neq 0$



3. $\operatorname{grad}(f) = \cos(x)\boldsymbol{e}_x + \boldsymbol{e}_y$



Solution 4.9:

1. $\begin{pmatrix} 2y^2 & 4xy \\ 4xy & 2x^2 \end{pmatrix}$
2. $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$
3. $\begin{pmatrix} -\sin(x) & 0 \\ 0 & -\cos(y) \end{pmatrix}$
4. $\begin{pmatrix} 2e^y + z^2 e^x & 2xe^y & 2ze^x \\ 2xe^y & x^2 e^y + 2\sin(z) & 2y\cos(z) \\ 2ze^x & 2y\cos(z) & 2e^x - y^2\sin(z) \end{pmatrix}$

Solution 4.10:

1. minimum at $f(\frac{1}{2}, \frac{1}{2}) = 0$
2. maximum at $f(0, 1) = 2$
3. maximum at $f(1, 2) = 5$

Solution 4.11:

1. $\begin{pmatrix} 2x & 2y \\ y & x \end{pmatrix}$
2. $\begin{pmatrix} \cos(x) & -\sin(y) \\ 2xy^2 & 2x^2 y \\ ye^{xy} & xe^{xy} \end{pmatrix}$

3. $\begin{pmatrix} 1 & 2y & 3z^2 \\ e^y\sin(z) & xe^y\sin(z) & xe^y\cos(z) \\ 2xe^{y+z} & x^2e^{y+z} & x^2e^{y+z} \end{pmatrix}$

A.5. Differential equations

Solution 5.1: Equations 1, 3, 4 and 5 are differential equations. (In expression 6 the derivatives cancel out.)

Solution 5.2:

	order	ordinary	explicit	linear	homogeneous
1.	2	yes	no	yes	no
2.	1	no	n.a.	yes	yes
3.	2	no	n.a.	yes	no
4.	2	yes	yes	no	no
5.	n	yes	yes	yes	yes
6.	1	yes	no	no	no

n.a. = not applicable

Solution 5.3:

Nr.	general	particular
1.	a), d)	c)
2.	d)	a)
3.	c)	a)

Solution 5.4:

1. $y' = y\cot(x)$
2. $y' = y\tanh(x)$
3. $y''y = (y')^2$
4. $xy' = 2y$

Solution 5.5: $xy' = y$

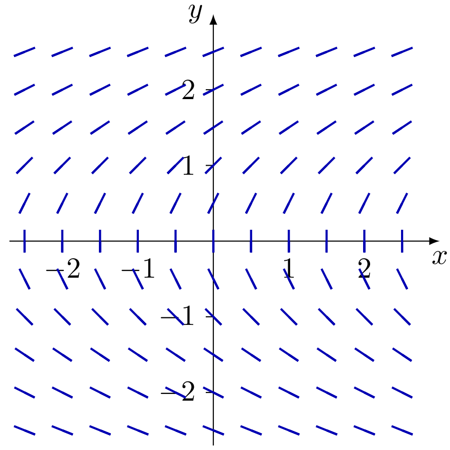
Solution 5.6: $xy' = 2y$

Solution 5.7: $y'y + x = 0$

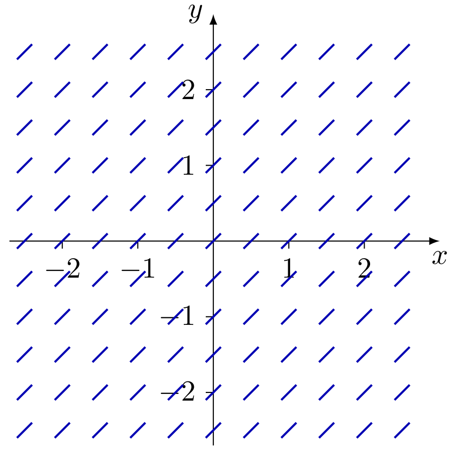
A.6. First order differential equations

Solution 6.1:

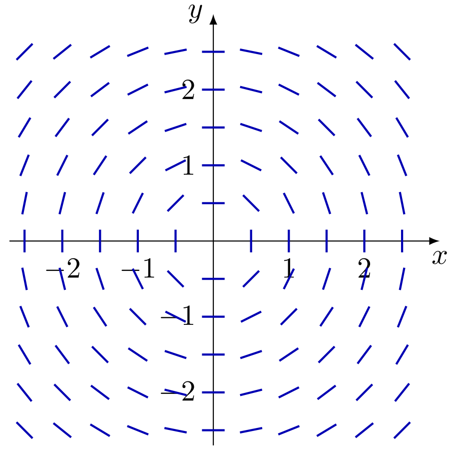
1.



2.



3.



Solution 6.2:

1. $y = \sqrt{2x + C}$
2. $y = \sqrt{x^2 + C}$
3. $y = Cx$
4. $y = \ln(x + C)$
5. $y = \sqrt{\cos(2x) + C}$
6. $y = \dfrac{1}{2x + C}$

Solution 6.3:

1. $y = c\,e^{x^3/3-x}$
2. $y = c\,x\,e^{-x^2}$
3. $y = c\,e^{-2\sqrt{x^3}}$
4. $y = c\,e^{\cos(x)}$
5. $y = c\,e^{\sinh(x)}$
6. $y = c\,\sqrt{|x|}$

Solution 6.4:

1. $y = cx - 1$
2. $y = \dfrac{e^x + c}{x}$
3. $y = \dfrac{\sin(x) + c}{x^2} - \dfrac{\cos(x)}{x}$

Solution 6.5:

1. $y = \dfrac{c}{x} + 1$
2. $y = \dfrac{\sin(x) + 2\cos(x)}{5} + c\,e^{-2x}$
3. $y = c\,e^x - \dfrac{1+\mathrm{j}}{2}\,e^{\mathrm{j}x}$

Solution 6.6:

1. $y = c\,e^{3x}$
2. $y = c\,e^{-3x}$
3. $y = c\,e^{-x^4/4}$
4. $y = c\,e^{-2/x}$
5. $y = c\,e^{kx}$
6. $y = c\,\exp\left(\dfrac{kx^{n+1}}{n+1}\right)$

Solution 6.7:

1. $y = c\,e^x + x + 1$
2. $y = c\,e^{-2x} + \frac{1}{4}(2x^2 - 2x + 1)$
3. $y = c\,e^{1/x} + 1$
4. $y = c\,x - 1$
5. $y = c\,e^{-x} + x - 1$
6. $y = c\,e^x + 2x + 2$

A.7. Higher order linear differential equations

Solution 7.1:

1. $y = C_1\,e^{-2x} + C_2\,e^x$
2. $y = C_1\,e^{-x} + C_2\,e^x$
3. $y = (C_1x + C_2)\,e^x$
4. $y = (C_1\cos(x) + C_2\sin(x))\,e^{-x}$
5. $y = C_1\,e^{-2x} + C_2\,e^{2x}$
6. $y = (C_1x + C_2)\,e^{x/2}$
7. $y = (C_1\cos(x) + C_2\sin(x))\,e^{x/2}$
8. $y = (C_1\cos(3x) + C_2\sin(3x))\,e^x$

Solution 7.2:

1. $y = C_1\,e^{-x} + C_2\,e^{-2x} + C_3\,e^{-3x}$
2. $y = C_1\,e^{-2x} + C_2\,e^{-x} + C_3\,e^x + C_4\,e^{2x}$
3. $y = C_1\,e^x + C_2\cos(x) + C_3\sin(x)$
4. $y = (C_1 + C_2\cos(x) + C_3\sin(x))\,e^{-x}$
5. $y = (C_1x + C_2)\,e^{-x} + (C_3x + C_4)\,e^x$
6. $y = (C_1x + C_2)\,e^x$
   $+ (C_3\cos(2x) + C_4\sin(2x))\,e^{-x}$

Solution 7.3:

1. $y = C_1\,e^{-2x} + C_2\,e^{-x} + x - 1$
2. $y = (C_1x + C_2)\,e^{-x} + x$
3. $y = (C_1\cos(x) + C_2\sin(x))\,e^{-x} - x + 2$
4. $y = (C_1x^2 + C_2x + C_3)\,e^{-x} - x^2 + 6x - 12$
5. $y = (C_1x + C_2)\,e^x$
   $+ (C_3\cos(3x) + C_4\sin(3x))\,e^{-x} + \frac{1}{2}$
6. $y = C_1\,e^{x/2} + C_2\,e^{-x/2} + x^3 + 24x$

Solution 7.4: Solve the following DEs:

1. $y = C_1\,e^{-2x} + C_2\,e^x + \frac{1}{5}\,e^{3x}$
2. $y = C_1\cos(2x) + C_2\sin(2x) + e^{-2x}$
3. $y = (C_1x + C_2)\,e^{2x} + \frac{1}{3}\,e^{-x}$
4. $y = C_1\,e^x + C_2\cos(x) + C_3\sin(x) - \frac{1}{5}\,e^{-2x}$
5. $y = C_1\,e^{-x/2} + C_2\,e^x + e^{-2x}$

Solution 7.5: Solve the following DEs:

1. $y = C_1\, e^{2x} + C_2\, e^{-2x} - \sin(x)$
2. $y = (C_1 x + C_2)\, e^{x} + 2\sin(x)$
3. $y = C_1\cos(2x) + C_2\sin(2x) - \cos(3x)$
4. $y = (C_1\cos(2x) + C_2\sin(2x))\, e^{x} + \frac{1}{2}\cos(x)$
5. $y = C\, e^{-x} + \frac{2}{5}\cos(2x) - \frac{1}{5}\sin(2x)$
6. $y = (C_1 x^2 + C_2 x + C_3)\, e^{-x} + \sin(x)$

Solution 7.6: Solve the following DEs:

1. $y = C_1\, e^{-3x} + C_2\, e^{x} - \frac{1}{4}\, e^{-x} + x + \frac{2}{3}$
2. $y = C_1\cos(3x) + C_2\sin(3x) + \frac{1}{3}\, e^{3x} + \sin(x)$
3. $y = C_1\, e^{-2x} + C_2\, e^{2x} + \frac{1}{5}\cos(x) + x^2 + \frac{1}{2}$
4. $y = (C_1 x + C_2)\, e^{-x} + C_3\, e^{-2x} + x^2 + e^{x}$

Solution 7.7: Solve the following DEs:

1. $y = (C_1 + x)\, e^{-2x}$
2. $y = C_1\cos(2x) + (C_2 - x)\sin(2x)$
3. $y = (C_1 + \frac{3}{5}x)\, e^{-3x} + C_2\, e^{2x}$
4. $y = (C_1 x + C_2 + x^2)\, e^{-x} + C_3\, e^{-2x}$

A.8. Combinatorics

Solution 8.1: 40 320

Solution 8.2: 3 628 800

Solution 8.3: 90

Solution 8.4: 70

Solution 8.5: 56

Solution 8.6: 126

Solution 8.7: 252

Solution 8.8: 7

Solution 8.9: 10 068 347 520

Solution 8.10: 336

Solution 8.11: n^k

Solution 8.12: 20 736

A.9. Probability theory

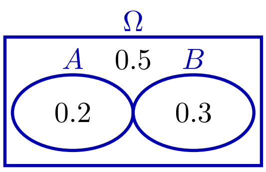
Solution 9.1:

1. $\{1, 2, 3, A, B, C\}$
2. $\{\}$
3. $\{2, 3, A, B, C\}$
4. $\{A, C\}$
5. $\{1, 2, 3\}$
6. $\{1, 3, C\}$

Solution 9.2: If two events have no common elements (i.e. their intersection is empty) they are said to be *mutually exclusive*:

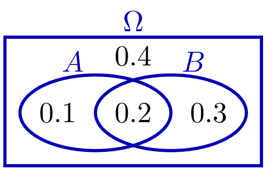
Solution 9.3:

1. 0.8
2. 0.5
3. 0.0
4. 0.2
5. 0.5
6. 0.3



Solution 9.4:

1. 0.6
2. 0.1
3. 0.3
4. 0.1
5. 0.8
6. 0.3



Solution 9.5:

1. 0.4
2. $\frac{2}{3}$

Solution 9.6:

1. 0
2. 0

Solution 9.7: With A_k, $k = 1, 2, 3$ as the event of an product to be produced by machine 1 to 3, respectively, and B as the event of a failure:

a) $P(B) = 2.9\ \%$

b) $P(A_1|B) = 34.5\ \%$
$P(A_2|B) = 31.0\ \%$
$P(A_3|B) = 34.5\ \%$

Solution 9.8: With A_k, $k = 1, 2, 3, 4$ as the event of age group 1 to 4, respectively, and B as the event of road fatality:

a)

$$P(A_1) = 13.4\ \% \qquad P(A_3) = 54.8\ \%$$
$$P(A_2) = 11.2\ \% \qquad P(A_4) = 20.6\ \%$$

b) (in parts per million, ppm)

$$P(B|A_1) = 9.51 \qquad P(B|A_3) = 41.1$$
$$P(B|A_2) = 86.6 \qquad P(B|A_4) = 54.0$$

c)

$$P(B) = 44.6 \text{ ppm}$$

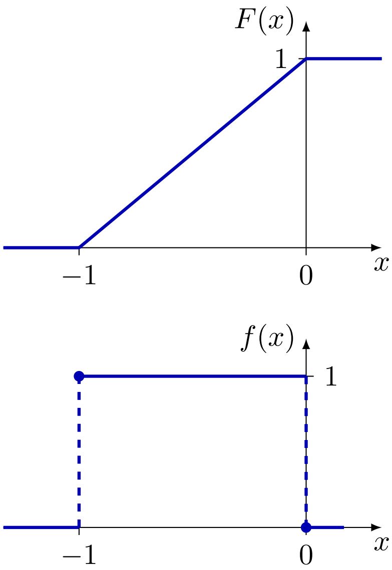
d)

$$P(A_1|B) = 2.9\ \% \qquad P(A_3|B) = 50.5\ \%$$
$$P(A_2|B) = 21.7\ \% \qquad P(A_4|B) = 25.0\ \%$$

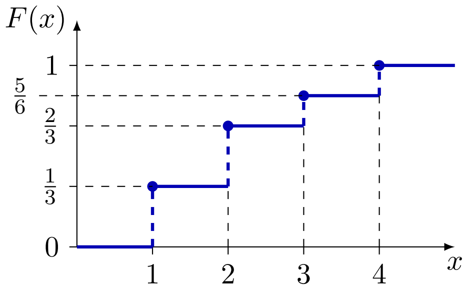
Solution 9.9: The two events A and B are independent.

A.10. Stochastic

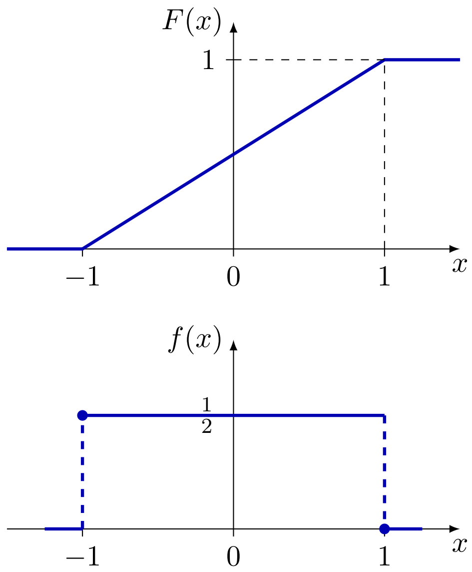
Solution 10.1:



Solution 10.2:



Solution 10.3:

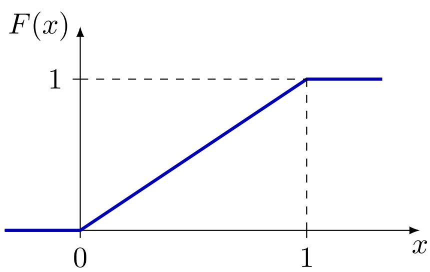


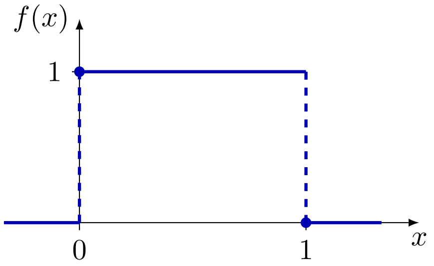
$$\mu = 0 \quad \sigma^2 = \tfrac{1}{3} \approx 0.333 \quad \sigma = \tfrac{1}{\sqrt{3}} \approx 0.577$$

Solution 10.4:

$$\mu = \sigma^2 = \sigma = 1$$

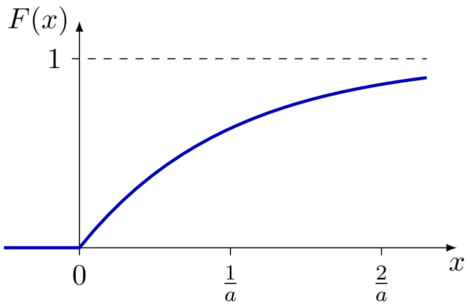
Solution 10.5:

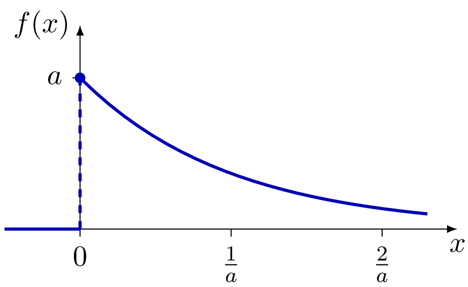




$$\mu = \tfrac{1}{2} = 0.5$$
$$\sigma^2 = \tfrac{1}{12} \approx 0.083$$
$$\sigma = \tfrac{1}{2\sqrt{3}} \approx 0.289$$

Solution 10.6:





$$\mu = \tfrac{1}{a} \qquad \sigma^2 = \tfrac{1}{a^2} \qquad \sigma = \tfrac{1}{a}$$

Solution 10.7:

1. $\mu_X = \frac{1}{2}$, $\quad {\sigma_X}^2 = \frac{1}{4}$, $\quad \sigma_X = \frac{1}{2}$
   $\mu_Y = \frac{1}{2}$, $\quad {\sigma_Y}^2 = \frac{1}{4}$, $\quad \sigma_Y = \frac{1}{2}$
2. $\mathrm{Cov}(X,Y) = \frac{1}{12}$, $\quad \mathrm{Corr}(X,Y) = \frac{1}{3}$.
3. $\mu = 1$, $\quad \sigma^2 = \frac{2}{3}$, $\quad \sigma = \sqrt{\frac{2}{3}}$
4. No

Solution 10.8:

1. $P(X=0) = \frac{8}{27} \qquad P(X=1) = \frac{4}{9}$
   $P(X=2) = \frac{2}{9} \qquad P(X=3) = \frac{1}{27}$
2. $\mu = 1 \qquad \sigma^2 = \frac{2}{3} \qquad \sigma = \sqrt{\frac{2}{3}}$

Solution 10.9:

a)

k	$P(X=k)$
0	5.0 %
1	14.9 %
2	22.4 %
3	22.4 %

k	$P(X=k)$
4	16.8 %
5	10.1 %
6	5.0 %
7	2.2 %

b) $\mu = 3 \qquad \sigma^2 = 3 \qquad \sigma = \sqrt{3}$

Solution 10.10: For results, see problem.

Solution 10.11:

1. $P(X_n > 4) = 15.9\,\%$
2. $P(X_n < 1) = 2.3\,\%$
3. $P(2 < X_n < 3) = 34.1\,\%$
4. $P(|X_n - \mu| > 2) = 4.6\,\%$

Solution 10.12:

$$f_n(x) = \frac{1}{\sqrt{2\pi}\,\sigma} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

with $\quad \mu = 100\,\mu_1 \quad$ and $\quad \sigma = 10\,\sigma_1$

Solution 10.13: failure rate: 16.0 %

Solution 10.14: $\sigma = 0.004 \cdot R$

Solution 10.15:
1. 26.4 %, 2. 28.3 %, 3. 14.6 %, 4. 61.7 %

Solution 10.16:
$\mu = \sigma = 34\,783\,\mathrm{a} \qquad \sigma^2 = 1.2099 \times 10^9\,\mathrm{a}^2$

Index

absolutes
integration of, 25
absorbed power, 30
antiderivative, 10
area of a trapezium, 13
area of discs, 14
area of polygons, 14
area of triangles, 14
arithmetic mean, 17
axioms of probability, 67

binomial coefficient, 61
binomial distribution, 75, 76
binomial random variable, 76
boundary problems, 42

capacitor discharge, 40
characteristic equation, 52, 54
characteristic polynomial, 52, 54
combination, 60
combined continuous functions, 33
complex solution of a DE, 51
conditional probability, 67
continuity of multiple argument functions, 33
continuous random variable, 71
continuous sample space, 65
contour plot, 31
correlation, 75
countable sample space, 65
covariance, 74
creating differential equations, 43
cumulative distribution function, 71

damped harmonic oscillator, 53
DE in resonance, 58
DE, differential equation, 40
definite integral, 6, 7
deterministic process, 64
differential equation, DE, 40
directional derivative, 34
discrete random variable, 71
discrete sample space, 65
distribution function, 71
domain of integration, 7

effective voltage, 5
elementary integrals, 11
event, 65
expectation, 72
expectation operator, 72
explicit notation, 41
exponential distribution, 80
exponential random variable, 80
exponential source term if a DE, 56
extrema of multiple argument functions, 36
condition for, 37

finite sample space, 65
first fundamental theorem of calculus, 10
first order differential equations, 46
geometric interpretation, 46
free path, 81
frequentist probability, 66
function with n arguments, 30
fundamental theorem of calculus, 9
first, 10
second, 10

general plus particular solution, 48
general solution, 42
global extremum, 36
global maximum, 36
global minimum, 36
gradient, 35
gradient of a function, 35

half-life, 42
harmonic oscillator
damped, 53
simple, 42
Hessian matrix, 36
higher order hom. DE with const. coeff., 54
higher order inhom. DE with const. coeff., 55
homogeneous DE, 41

i.i.d., 75
implicit notation, 41
improper integrals, 20
indefinite integral, 10
independence of probabilities, 68
independent and identically distributed, 75
independent events, 68

infinite sample space, 65
inhomogeneous DE, 41
initial value problems, 42
integrable function, 7
integrals with infinite domain, 20
integrals with infinite image, 20
integrand, 7
integration by parts, 24
integration by substitution, 21
integration of absolutes, 25
integration of complex conjugate poles, 27
integration of multiple real poles, 27
integration of rational functions, 26
integration of single real poles, 26
inverse derivative, 10

Jacobian matrix, 38

kinetic energy, 30

Lebesgue integration, 7
limits of integration, 7
linear combination of solutions of a DE, 51
linear congruential generator, 78
linear differential equation, 41
linear inhomogeneous DE, 48
local extremum, 36
local maximum, 36
local minimum, 36
lotto, 60, 61
lower limit, 7
lower sum, 6

maxima of multiple argument functions, 36
mean free path, 81
minima of multiple argument functions, 36
Monty Hall problem, 64
multiple argument function, 30
multiple partial derivative, 35
multiple value functions, 37
mutually exclusive, 66

nabla-operator, 35
normal distribution, 78
normal random variable, 78

ODE, ordinary differential equation, 41
order of the differential equation, 41
ordinary differential equation, ODE, 41
oscillator
  damped harmonic, 53
  simple harmonic, 42

parameter-function, 47
parametric plot, 30
partial derivative, 33, 34
partial differential equation, 41
partial fraction decomposition, 26
particular solution, 42
permutation, 60
Poisson distribution, 76
Poisson random variable, 76
polynomial source term of a DE, 55
primitive function, 10
probability, 64
probability axioms, 67
probability density function, 71
probability mass function, 71
product rule, 24
pseudo random number generator, 77

radioactive decay, 42
random variable, 71
Riemann integral, 7
Riemann sum, 6
root mean square, 17

saddle point, 37
sample, 65
sample space, 65
  continuous, 65
  countable, 65
  discrete, 65
  finite, 65
  infinite, 65
  uncountable, 65
Schwarz integrability condition, 36
second fundamental theorem of calculus, 10
second order hom. DE with const. coeff., 52
  pair of complex conjugate constants, 52
  two different real constants, 52
  two equal real constants, 52
separation of variables, 46
separation of variables 1, 46
separation of variables 2, 47
simple harmonic oscillator, 42
source term, 51, 55
source term of an DE as a sum of functions, 57
standard deviation, 73
standard normal distribution, 78
standard normal random variable, 78
stochastic process, 64
sum of conditional probabilities, 68
surface area of spheres, 16
surface plot, 31
symmetry of second derivatives, 36

trigonometric source term of a DE, 57

uncountable sample space, 65
uniform distribution, 77
uniform random variable, 77
upper limit, 7
upper sum, 6

variation, 62
variation of parameters, 47
vector plot, 32
visualization of functions, 30
volume of cones, 16
volume of spheres, 16