# Mathematics Formulary

**By ir. J.C.A. Wevers**

 

Dear reader,

This document contains 66 pages with mathematical equations intended for physicists and engineers. It is intended to be a short reference for anyone who often needs to look up mathematical equations.

This document can also be obtained from the author, Johan Wevers (`johanw@vulcan.xs4all.nl`).

It can also be found on the WWW on `http://www.xs4all.nl/~johanw/index.html`.



The C code for the rootfinding via Newtons method and the FFT in chapter 8 are from "*Numerical Recipes in C* ", 2nd Edition, ISBN 0-521-43108-5.

The Mathematics Formulary is made with teTEX and LATEX version 2.09.

If you prefer the notation in which vectors are typefaced in boldface, uncomment the redefinition of the `\vec` command and recompile the file.

If you find any errors or have any comments, please let me know. I am always open for suggestions and possible corrections to the mathematics formulary.

Johan Wevers

# Contents

# Chapter 1

# Basics

## 1.1 Goniometric functions

For the goniometric ratios for a point $p$ on the unit circle holds:

$$\cos(\phi) = x_p \ , \quad \sin(\phi) = y_p \ , \quad \tan(\phi) = \frac{y_p}{x_p}$$

$\sin^2(x) + \cos^2(x) = 1$ and $\cos^{-2}(x) = 1 + \tan^2(x)$.

$$\cos(a \pm b) = \cos(a)\cos(b) \mp \sin(a)\sin(b) \ , \quad \sin(a \pm b) = \sin(a)\cos(b) \pm \cos(a)\sin(b)$$

$$\tan(a \pm b) = \frac{\tan(a) \pm \tan(b)}{1 \mp \tan(a)\tan(b)}$$

The **sum formulas** are:

$$\begin{aligned}
\sin(p) + \sin(q) &= 2\sin(\tfrac{1}{2}(p+q))\cos(\tfrac{1}{2}(p-q)) \\
\sin(p) - \sin(q) &= 2\cos(\tfrac{1}{2}(p+q))\sin(\tfrac{1}{2}(p-q)) \\
\cos(p) + \cos(q) &= 2\cos(\tfrac{1}{2}(p+q))\cos(\tfrac{1}{2}(p-q)) \\
\cos(p) - \cos(q) &= -2\sin(\tfrac{1}{2}(p+q))\sin(\tfrac{1}{2}(p-q))
\end{aligned}$$

From these equations can be derived that

$$\begin{aligned}
2\cos^2(x) = 1 + \cos(2x) \quad &, \quad 2\sin^2(x) = 1 - \cos(2x) \\
\sin(\pi - x) = \sin(x) \quad &, \quad \cos(\pi - x) = -\cos(x) \\
\sin(\tfrac{1}{2}\pi - x) = \cos(x) \quad &, \quad \cos(\tfrac{1}{2}\pi - x) = \sin(x)
\end{aligned}$$

**Conclusions from equalities**:

$$\begin{aligned}
\underline{\sin(x) = \sin(a)} \quad &\Rightarrow \quad x = a \pm 2k\pi \text{ or } x = (\pi - a) \pm 2k\pi, \ \ k \in I\!N \\
\underline{\cos(x) = \cos(a)} \quad &\Rightarrow \quad x = a \pm 2k\pi \text{ or } x = -a \pm 2k\pi \\
\underline{\tan(x) = \tan(a)} \quad &\Rightarrow \quad x = a \pm k\pi \text{ and } x \neq \frac{\pi}{2} \pm k\pi
\end{aligned}$$

The following relations exist between the inverse goniometric functions:

$$\arctan(x) = \arcsin\left(\frac{x}{\sqrt{x^2+1}}\right) = \arccos\left(\frac{1}{\sqrt{x^2+1}}\right) \ , \quad \sin(\arccos(x)) = \sqrt{1-x^2}$$

## 1.2 Hyperbolic functions

The hyperbolic functions are defined by:

$$\sinh(x) = \frac{e^x - e^{-x}}{2} \ , \quad \cosh(x) = \frac{e^x + e^{-x}}{2} \ , \quad \tanh(x) = \frac{\sinh(x)}{\cosh(x)}$$

From this follows that $\cosh^2(x) - \sinh^2(x) = 1$. Further holds:

$$\operatorname{arsinh}(x) = \ln|x + \sqrt{x^2+1}| \ , \quad \operatorname{arcosh}(x) = \operatorname{arsinh}(\sqrt{x^2-1})$$

## 1.3 Calculus

The derivative of a function is defined as:

$$\frac{df}{dx} = \lim_{h \to 0} \frac{f(x+h) - f(x)}{h}$$

Derivatives obey the following algebraic rules:

$$d(x \pm y) = dx \pm dy \ , \ d(xy) = xdy + ydx \ , \ d\left(\frac{x}{y}\right) = \frac{ydx - xdy}{y^2}$$

For the derivative of the inverse function $f^{\text{inv}}(y)$, defined by $f^{\text{inv}}(f(x)) = x$, holds at point $P = (x, f(x))$:

$$\left(\frac{df^{\text{inv}}(y)}{dy}\right)_P \cdot \left(\frac{df(x)}{dx}\right)_P = 1$$

Chain rule: if $f = f(g(x))$, then holds

$$\frac{df}{dx} = \frac{df}{dg}\frac{dg}{dx}$$

Further, for the derivatives of products of functions holds:

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

For the *primitive function* $F(x)$ holds: $F'(x) = f(x)$. An overview of derivatives and primitives is:

| $\boldsymbol{y = f(x)}$ | $\boldsymbol{dy/dx = f'(x)}$ | $\boldsymbol{\int f(x)dx}$ |
|---|---|---|
| $ax^n$ | $anx^{n-1}$ | $a(n+1)^{-1}x^{n+1}$ |
| $1/x$ | $-x^{-2}$ | $\ln\lvert x\rvert$ |
| $a$ | $0$ | $ax$ |
| $a^x$ | $a^x \ln(a)$ | $a^x / \ln(a)$ |
| $\mathrm{e}^x$ | $\mathrm{e}^x$ | $\mathrm{e}^x$ |
| ${}^a\log(x)$ | $(x\ln(a))^{-1}$ | $(x\ln(x) - x)/\ln(a)$ |
| $\ln(x)$ | $1/x$ | $x\ln(x) - x$ |
| $\sin(x)$ | $\cos(x)$ | $-\cos(x)$ |
| $\cos(x)$ | $-\sin(x)$ | $\sin(x)$ |
| $\tan(x)$ | $\cos^{-2}(x)$ | $-\ln\lvert\cos(x)\rvert$ |
| $\sin^{-1}(x)$ | $-\sin^{-2}(x)\cos(x)$ | $\ln\lvert\tan(\frac{1}{2}x)\rvert$ |
| $\sinh(x)$ | $\cosh(x)$ | $\cosh(x)$ |
| $\cosh(x)$ | $\sinh(x)$ | $\sinh(x)$ |
| $\arcsin(x)$ | $1/\sqrt{1-x^2}$ | $x\arcsin(x) + \sqrt{1-x^2}$ |
| $\arccos(x)$ | $-1/\sqrt{1-x^2}$ | $x\arccos(x) - \sqrt{1-x^2}$ |
| $\arctan(x)$ | $(1+x^2)^{-1}$ | $x\arctan(x) - \frac{1}{2}\ln(1+x^2)$ |
| $(a+x^2)^{-1/2}$ | $-x(a+x^2)^{-3/2}$ | $\ln\lvert x + \sqrt{a+x^2}\rvert$ |
| $(a^2-x^2)^{-1}$ | $2x(a^2+x^2)^{-2}$ | $\frac{1}{2a}\ln\lvert(a+x)/(a-x)\rvert$ |

The *curvature* $\rho$ of a curve is given by: $\rho = \dfrac{(1+(y')^2)^{3/2}}{|y''|}$

The theorem of De 'l Hôpital: if $f(a) = 0$ and $g(a) = 0$, then is $\lim\limits_{x \to a} \dfrac{f(x)}{g(x)} = \lim\limits_{x \to a} \dfrac{f'(x)}{g'(x)}$

## 1.4 Limits

$$\lim_{x\to 0}\frac{\sin(x)}{x}=1\ ,\ \lim_{x\to 0}\frac{\mathrm{e}^x-1}{x}=1\ ,\ \lim_{x\to 0}\frac{\tan(x)}{x}=1\ ,\ \lim_{k\to 0}(1+k)^{1/k}=\mathrm{e}\ ,\ \lim_{x\to\infty}\left(1+\frac{n}{x}\right)^x=\mathrm{e}^n$$

$$\lim_{x\downarrow 0}x^a\ln(x)=0\ ,\ \lim_{x\to\infty}\frac{\ln^p(x)}{x^a}=0\ ,\ \lim_{x\to 0}\frac{\ln(x+a)}{x}=a\ ,\ \lim_{x\to\infty}\frac{x^p}{a^x}=0\ \text{ als } |a|>1.$$

$$\lim_{x\to 0}\left(a^{1/x}-1\right)=\ln(a)\ ,\ \lim_{x\to 0}\frac{\arcsin(x)}{x}=1\ ,\ \lim_{x\to\infty}\sqrt[x]{x}=1$$

## 1.5 Complex numbers and quaternions

### 1.5.1 Complex numbers

The complex number $z=a+bi$ with $a$ and $b\in\mathbb{R}$. $a$ is the *real part*, $b$ the *imaginary part* of $z$. $|z|=\sqrt{a^2+b^2}$. By definition holds: $i^2=-1$. Every complex number can be written as $z=|z|\exp(i\varphi)$, with $\tan(\varphi)=b/a$. The *complex conjugate* of $z$ is defined as $\overline{z}=z^*:=a-bi$. Further holds:

$$\begin{aligned}(a+bi)(c+di) &= (ac-bd)+i(ad+bc)\\ (a+bi)+(c+di) &= a+c+i(b+d)\\ \frac{a+bi}{c+di} &= \frac{(ac+bd)+i(bc-ad)}{c^2+d^2}\end{aligned}$$

Goniometric functions can be written as complex exponents:

$$\begin{aligned}\sin(x) &= \frac{1}{2i}(\mathrm{e}^{ix}-\mathrm{e}^{-ix})\\ \cos(x) &= \frac{1}{2}(\mathrm{e}^{ix}+\mathrm{e}^{-ix})\end{aligned}$$

From this follows that $\cos(ix)=\cosh(x)$ and $\sin(ix)=i\sinh(x)$. Further follows from this that $\mathrm{e}^{\pm ix}=\cos(x)\pm i\sin(x)$, so $\mathrm{e}^{iz}\neq 0\forall z$. Also the theorem of De Moivre follows from this: $(\cos(\varphi)+i\sin(\varphi))^n=\cos(n\varphi)+i\sin(n\varphi)$.

Products and quotients of complex numbers can be written as:

$$\begin{aligned}z_1\cdot z_2 &= |z_1|\cdot|z_2|(\cos(\varphi_1+\varphi_2)+i\sin(\varphi_1+\varphi_2))\\ \frac{z_1}{z_2} &= \frac{|z_1|}{|z_2|}(\cos(\varphi_1-\varphi_2)+i\sin(\varphi_1-\varphi_2))\end{aligned}$$

The following can be derived:

$$|z_1+z_2|\leq|z_1|+|z_2|\ ,\ |z_1-z_2|\geq|\,|z_1|-|z_2|\,|$$

And from $z=r\exp(i\theta)$ follows: $\ln(z)=\ln(r)+i\theta$, $\ln(z)=\ln(z)\pm 2n\pi i$.

### 1.5.2 Quaternions

Quaternions are defined as: $z=a+bi+cj+dk$, with $a,b,c,d\in\mathbb{R}$ and $i^2=j^2=k^2=-1$. The products of $i,j,k$ with each other are given by $ij=-ji=k$, $jk=-kj=i$ and $ki=-ik=j$.

## 1.6 Geometry

### 1.6.1 Triangles

The sine rule is:

$$\frac{a}{\sin(\alpha)} = \frac{b}{\sin(\beta)} = \frac{c}{\sin(\gamma)}$$

Here, $\alpha$ is the angle opposite to $a$, $\beta$ is opposite to $b$ and $\gamma$ opposite to $c$. The cosine rule is: $a^2 = b^2+c^2-2bc\cos(\alpha)$. For each triangle holds: $\alpha + \beta + \gamma = 180°$.

Further holds:

$$\frac{\tan(\frac{1}{2}(\alpha+\beta))}{\tan(\frac{1}{2}(\alpha-\beta))} = \frac{a+b}{a-b}$$

The surface of a triangle is given by $\frac{1}{2}ab\sin(\gamma) = \frac{1}{2}ah_a = \sqrt{s(s-a)(s-b)(s-c)}$ with $h_a$ the perpendicular on $a$ and $s = \frac{1}{2}(a+b+c)$.

### 1.6.2 Curves

**Cycloid**: if a circle with radius $a$ rolls along a straight line, the trajectory of a point on this circle has the following parameter equation:

$$x = a(t+\sin(t)) \ , \ y = a(1+\cos(t))$$

**Epicycloid**: if a small circle with radius $a$ rolls along a big circle with radius $R$, the trajectory of a point on the small circle has the following parameter equation:

$$x = a\sin\left(\frac{R+a}{a}t\right) + (R+a)\sin(t) \ , \ y = a\cos\left(\frac{R+a}{a}t\right) + (R+a)\cos(t)$$

**Hypocycloid**: if a small circle with radius $a$ rolls inside a big circle with radius $R$, the trajectory of a point on the small circle has the following parameter equation:

$$x = a\sin\left(\frac{R-a}{a}t\right) + (R-a)\sin(t) \ , \ y = -a\cos\left(\frac{R-a}{a}t\right) + (R-a)\cos(t)$$

A hypocycloid with $a = R$ is called a **cardioid**. It has the following parameterequation in polar coordinates: $r = 2a[1-\cos(\varphi)]$.

## 1.7 Vectors

The *inner product* is defined by: $\vec{a}\cdot\vec{b} = \sum_i a_i b_i = |\vec{a}|\cdot|\vec{b}|\cos(\varphi)$

where $\varphi$ is the angle between $\vec{a}$ and $\vec{b}$. The *external product* is in $\mathbb{R}^3$ defined by:

$$\vec{a}\times\vec{b} = \begin{pmatrix} a_yb_z - a_zb_y \\ a_zb_x - a_xb_z \\ a_xb_y - a_yb_x \end{pmatrix} = \begin{vmatrix} \vec{e}_x & \vec{e}_y & \vec{e}_z \\ a_x & a_y & a_z \\ b_x & b_y & b_z \end{vmatrix}$$

Further holds: $|\vec{a}\times\vec{b}| = |\vec{a}|\cdot|\vec{b}|\sin(\varphi)$, and $\vec{a}\times(\vec{b}\times\vec{c}) = (\vec{a}\cdot\vec{c})\vec{b} - (\vec{a}\cdot\vec{b})\vec{c}$.

## 1.8 Series

### 1.8.1 Expansion

The Binomium of Newton is:

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

where $\binom{n}{k} := \frac{n!}{k!(n-k)!}$.

By subtracting the series $\sum_{k=0}^{n} r^k$ and $r \sum_{k=0}^{n} r^k$ one finds:

$$\sum_{k=0}^{n} r^k = \frac{1-r^{n+1}}{1-r}$$

and for $|r| < 1$ this gives the *geometric series*: $\sum_{k=0}^{\infty} r^k = \frac{1}{1-r}$.

The *arithmetic series* is given by: $\sum_{n=0}^{N}(a+nV) = a(N+1) + \frac{1}{2}N(N+1)V$.

The expansion of a function around the point $a$ is given by the *Taylor series*:

$$f(x) = f(a) + (x-a)f'(a) + \frac{(x-a)^2}{2}f''(a) + \cdots + \frac{(x-a)^n}{n!}f^{(n)}(a) + R$$

where the remainder is given by:

$$R_n(h) = (1-\theta)^n \frac{h^n}{n!} f^{(n+1)}(\theta h)$$

and is subject to:

$$\frac{mh^{n+1}}{(n+1)!} \leq R_n(h) \leq \frac{Mh^{n+1}}{(n+1)!}$$

From this one can deduce that

$$(1-x)^\alpha = \sum_{n=0}^{\infty} \binom{\alpha}{n} x^n$$

One can derive that:

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}\,, \quad \sum_{n=1}^{\infty} \frac{1}{n^4} = \frac{\pi^4}{90}\,, \quad \sum_{n=1}^{\infty} \frac{1}{n^6} = \frac{\pi^6}{945}$$

$$\sum_{k=1}^{n} k^2 = \tfrac{1}{6}n(n+1)(2n+1)\,, \quad \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n^2} = \frac{\pi^2}{12}\,, \quad \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} = \ln(2)$$

$$\sum_{n=1}^{\infty} \frac{1}{4n^2-1} = \tfrac{1}{2}\,, \quad \sum_{n=1}^{\infty} \frac{1}{(2n-1)^2} = \frac{\pi^2}{8}\,, \quad \sum_{n=1}^{\infty} \frac{1}{(2n-1)^4} = \frac{\pi^4}{96}\,, \quad \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{(2n-1)^3} = \frac{\pi^3}{32}$$

### 1.8.2 Convergence and divergence of series

If $\sum_n |u_n|$ converges, $\sum_n u_n$ also converges.

If $\lim_{n\to\infty} u_n \neq 0$ then $\sum_n u_n$ is divergent.

An alternating series of which the absolute values of the terms drop monotonously to 0 is convergent (Leibniz).

If $\int_p^\infty f(x)dx < \infty$, then $\sum_n f_n$ is convergent.

If $u_n > 0\ \forall n$ then is $\sum_n u_n$ convergent *if* $\sum_n \ln(u_n + 1)$ is convergent.

If $u_n = c_n x^n$ the radius of convergence $\rho$ of $\sum_n u_n$ is given by: $\frac{1}{\rho} = \lim_{n\to\infty} \sqrt[n]{|c_n|} = \lim_{n\to\infty} \left|\frac{c_{n+1}}{c_n}\right|$.

The series $\sum_{n=1}^{\infty} \frac{1}{n^p}$ is convergent if $p > 1$ and divergent if $p \leq 1$.

If: $\lim_{n\to\infty} \frac{u_n}{v_n} = p$, than the following is true: if $p > 0$ than $\sum_n u_n$ and $\sum_n v_n$ are both divergent or both convergent, if $p = 0$ holds: if $\sum_n v_n$ is convergent, than $\sum_n u_n$ is also convergent.

If $L$ is defined by: $L = \lim_{n\to\infty} \sqrt[n]{|n_n|}$, or by: $L = \lim_{n\to\infty} \left|\frac{u_{n+1}}{u_n}\right|$, then is $\sum_n u_n$ divergent if $L > 1$ and convergent if $L < 1$.

### 1.8.3 Convergence and divergence of functions

$f(x)$ is continuous in $x = a$ only if the upper - and lower limit are equal: $\lim_{x\uparrow a} f(x) = \lim_{x\downarrow a} f(x)$. This is written as: $f(a^-) = f(a^+)$.

If $f(x)$ is continuous in $a$ *and*: $\lim_{x\uparrow a} f'(x) = \lim_{x\downarrow a} f'(x)$, *than* $f(x)$ is differentiable in $x = a$.

We define: $\|f\|_W := \sup(|f(x)|\ |x \in W)$, and $\lim_{x\to\infty} f_n(x) = f(x)$. Than holds: $\{f_n\}$ is uniform convergent if $\lim_{n\to\infty} \|f_n - f\| = 0$, or: $\forall(\varepsilon > 0)\exists(N)\forall(n \geq N)\|f_n - f\| < \varepsilon$.

Weierstrass' test: if $\sum \|u_n\|_W$ is convergent, than $\sum u_n$ is uniform convergent.

We define $S(x) = \sum_{n=N}^{\infty} u_n(x)$ and $F(y) = \int_a^b f(x,y)dx := F$. Than it can be proved that:

| **Theorem** | **For** | **Demands on** $W$ | **Than holds on** $W$ |
|---|---|---|---|
| C | rows | $f_n$ continuous,<br>$\{f_n\}$ uniform convergent | $f$ is continuous |
| | series | $S(x)$ uniform convergent,<br>$u_n$ continuous | $S$ is continuous |
| | integral | $f$ is continuous | $F$ is continuous |
| I | rows | $f_n$ can be integrated,<br>$\{f_n\}$ uniform convergent | $f_n$ can be integrated,<br>$\int f(x)dx = \lim_{n\to\infty} \int f_n dx$ |
| | series | $S(x)$ is uniform convergent,<br>$u_n$ can be integrated | $S$ can be integrated, $\int S dx = \sum \int u_n dx$ |
| | integral | $f$ is continuous | $\int F dy = \iint f(x,y)dxdy$ |
| D | rows | $\{f_n\} \in C^{-1}$; $\{f_n'\}$ unif.conv $\to \phi$ | $f' = \phi(x)$ |
| | series | $u_n \in C^{-1}$; $\sum u_n$ conv; $\sum u_n'$ u.c. | $S'(x) = \sum u_n'(x)$ |
| | integral | $\partial f/\partial y$ continuous | $F_y = \int f_y(x,y)dx$ |

## 1.9 Products and quotients

For $a, b, c, d \in \mathbb{R}$ holds:
The **distributive property**: $(a+b)(c+d) = ac + ad + bc + bd$
The **associative property**: $a(bc) = b(ac) = c(ab)$ and $a(b+c) = ab + ac$
The **commutative property**: $a + b = b + a$, $ab = ba$.

Further holds:

$$\frac{a^{2n} - b^{2n}}{a \pm b} = a^{2n-1} \pm a^{2n-2}b + a^{2n-3}b^2 \pm \cdots \pm b^{2n-1} \quad , \quad \frac{a^{2n+1} - b^{2n+1}}{a+b} = \sum_{k=0}^{n} a^{2n-k}b^{2k}$$

$$(a \pm b)(a^2 \pm ab + b^2) = a^3 \pm b^3 \ , \ \ (a+b)(a-b) = a^2 + b^2 \ , \ \ \frac{a^3 \pm b^3}{a+b} = a^2 \mp ba + b^2$$

## 1.10 Logarithms

**Definition**: ${}^a\log(x) = b \Leftrightarrow a^b = x$. For logarithms with base $e$ one writes $\ln(x)$.

**Rules**: $\log(x^n) = n\log(x)$, $\log(a) + \log(b) = \log(ab)$, $\log(a) - \log(b) = \log(a/b)$.

## 1.11 Polynomials

Equations of the type

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

have $n$ roots which may be equal to each other. Each polynomial $p(z)$ of order $n \geq 1$ has at least one root in $\mathbb{C}$. If all $a_k \in \mathbb{R}$ holds: when $x = p$ with $p \in \mathbb{C}$ a root, than $p^*$ is also a root. Polynomials up to and including order 4 have a general analytical solution, for polynomials with order $\geq 5$ there does not exist a general analytical solution.

For $a, b, c \in \mathbb{R}$ and $a \neq 0$ holds: the 2nd order equation $ax^2 + bx + c = 0$ has the general solution:

$$x = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

For $a, b, c, d \in \mathbb{R}$ and $a \neq 0$ holds: the 3rd order equation $ax^3 + bx^2 + cx + d = 0$ has the general analytical solution:

$$\begin{aligned} x_1 &= K - \frac{3ac - b^2}{9a^2K} - \frac{b}{3a} \\ x_2 = x_3^* &= -\frac{K}{2} + \frac{3ac - b^2}{18a^2K} - \frac{b}{3a} + i\frac{\sqrt{3}}{2}\left(K + \frac{3ac - b^2}{9a^2K}\right) \end{aligned}$$

with $K = \left(\frac{9abc - 27da^2 - 2b^3}{54a^3} + \frac{\sqrt{3}\sqrt{4ac^3 - c^2b^2 - 18abcd + 27a^2d^2 + 4db^3}}{18a^2}\right)^{1/3}$

## 1.12 Primes

A *prime* is a number $\in \mathbb{N}$ that can only be divided by itself and 1. There are an infinite number of primes. Proof: suppose that the collection of primes $P$ would be finite, than construct the number $q = 1 + \prod_{p \in P} p$, than holds $q = 1(p)$ and so $Q$ cannot be written as a product of primes from $P$. This is a contradiction.

If $\pi(x)$ is the number of primes $\leq x$, than holds:

$$\lim_{x\to\infty} \frac{\pi(x)}{x/\ln(x)} = 1 \quad \text{and} \quad \lim_{x\to\infty} \frac{\pi(x)}{\int_2^x \frac{dt}{\ln(t)}} = 1$$

For each $N \geq 2$ there is a prime between $N$ and $2N$.

The numbers $F_k := 2^k + 1$ with $k \in \mathbb{N}$ are called *Fermat numbers*. Many Fermat numbers are prime.

The numbers $M_k := 2^k - 1$ are called *Mersenne numbers*. They occur when one searches for *perfect numbers*, which are numbers $n \in \mathbb{N}$ which are the sum of their different dividers, for example $6 = 1 + 2 + 3$. There are 23 Mersenne numbers for $k < 12000$ which are prime: for $k \in \{2, 3, 5, 7, 13, 17, 19, 31, 61, 89, 107, 127, 521, 607, 1279, 2203, 2281, 3217, 4253, 4423, 9689, 9941, 11213\}$.

To check if a given number $n$ is prime one can use a sieve method. The first known sieve method was developed by Eratosthenes. A faster method for large numbers are the 4 Fermat tests, who don't prove that a number is prime but give a large probability.

1. Take the first 4 primes: $b = \{2, 3, 5, 7\}$,
2. Take $w(b) = b^{n-1} \bmod n$, for each $b$,
3. If $w = 1$ for each $b$, then $n$ is probably prime. For each other value of $w$, $n$ is certainly not prime.

# Chapter 2

# Probability and statistics

## 2.1 Combinations

The number of possible *combinations* of $k$ elements from $n$ elements is given by

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}$$

The number of *permutations* of $p$ from $n$ is given by

$$\frac{n!}{(n-p)!} = p!\binom{n}{p}$$

The number of different ways to classify $n_i$ elements in $i$ groups, when the total number of elements is $N$, is

$$\frac{N!}{\prod_i n_i!}$$

## 2.2 Probability theory

The probability $P(A)$ that an event $A$ occurs is defined by:

$$P(A) = \frac{n(A)}{n(U)}$$

where $n(A)$ is the number of events when $A$ occurs and $n(U)$ the total number of events.

The probability $P(\neg A)$ that $A$ *does not* occur is: $P(\neg A) = 1 - P(A)$. The probability $P(A \cup B)$ that $A$ and $B$ *both* occur is given by: $P(A \cup B) = P(A) + P(B) - P(A \cap B)$. If $A$ and $B$ are independent, than holds: $P(A \cap B) = P(A) \cdot P(B)$.

The probability $P(A|B)$ that $A$ occurs, given the fact that $B$ occurs, is:

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

## 2.3 Statistics

### 2.3.1 General

The *average* or *mean* value $\langle x \rangle$ of a collection of values is: $\langle x \rangle = \sum_i x_i/n$. The *standard deviation* $\sigma_x$ in the distribution of $x$ is given by:

$$\sigma_x = \sqrt{\frac{\sum_{i=1}^{n} (x_i - \langle x \rangle)^2}{n}}$$

When samples are being used the sample variance $s$ is given by $s^2 = \frac{n}{n-1}\sigma^2$.

The *covariance* $\sigma_{xy}$ of $x$ and $y$ is given by::

$$\sigma_{xy} = \frac{\sum_{i=1}^{n}(x_i - \langle x\rangle)(y_i - \langle y\rangle)}{n-1}$$

The *correlation coefficient* $r_{xy}$ of $x$ and $y$ than becomes: $r_{xy} = \sigma_{xy}/\sigma_x\sigma_y$.

The standard deviation in a variable $f(x,y)$ resulting from errors in $x$ and $y$ is:

$$\sigma^2_{f(x,y)} = \left(\frac{\partial f}{\partial x}\sigma_x\right)^2 + \left(\frac{\partial f}{\partial y}\sigma_y\right)^2 + \frac{\partial f}{\partial x}\frac{\partial f}{\partial y}\sigma_{xy}$$

### 2.3.2 Distributions

1. **The Binomial distribution** is the distribution describing a sampling with replacement. The probability for success is $p$. The probability $P$ for $k$ successes in $n$ trials is then given by:

   $$P(x=k) = \binom{n}{k}p^k(1-p)^{n-k}$$

   The standard deviation is given by $\sigma_x = \sqrt{np(1-p)}$ and the expectation value is $\varepsilon = np$.

2. **The Hypergeometric distribution** is the distribution describing a sampling without replacement in which the order is irrelevant. The probability for $k$ successes in a trial with $A$ possible successes and $B$ possible failures is then given by:

   $$P(x=k) = \frac{\binom{A}{k}\binom{B}{n-k}}{\binom{A+B}{n}}$$

   The expectation value is given by $\varepsilon = nA/(A+B)$.

3. **The Poisson distribution** is a limiting case of the binomial distribution when $p \to 0$, $n \to \infty$ and also $np = \lambda$ is constant.

   $$P(x) = \frac{\lambda^x \mathrm{e}^{-\lambda}}{x!}$$

   This distribution is normalized to $\sum_{x=0}^{\infty} P(x) = 1$.

4. **The Normal distribution** is a limiting case of the binomial distribution for continuous variables:

   $$P(x) = \frac{1}{\sigma\sqrt{2\pi}}\exp\left(-\frac{1}{2}\left(\frac{x-\langle x\rangle}{\sigma}\right)^2\right)$$

5. **The Uniform distribution** occurs when a random number $x$ is taken from the set $a \le x \le b$ and is given by:

   $$\begin{cases} P(x) = \dfrac{1}{b-a} & \text{if} \quad a \le x \le b \\ \\ P(x) = 0 & \text{in all other cases} \end{cases}$$

   $\langle x\rangle = \frac{1}{2}(b-a)$ and $\sigma^2 = \dfrac{(b-a)^2}{12}$.

6. **The Gamma distribution** is given by:

$$\left\{ \begin{array}{l} P(x) = \dfrac{x^{\alpha-1}\mathrm{e}^{-x/\beta}}{\beta^{\alpha}\Gamma(\alpha)} \quad \text{if} \quad 0 \leq y \leq \infty \end{array} \right.$$

with $\alpha > 0$ and $\beta > 0$. The distribution has the following properties: $\langle x \rangle = \alpha\beta$, $\sigma^2 = \alpha\beta^2$.

7. **The Beta distribution** is given by:

$$\left\{ \begin{array}{l} P(x) = \dfrac{x^{\alpha-1}(1-x)^{\beta-1}}{\beta(\alpha,\beta)} \quad \text{if} \quad 0 \leq x \leq 1 \\ \\ P(x) = 0 \quad \text{everywhere else} \end{array} \right.$$

and has the following properties: $\langle x \rangle = \dfrac{\alpha}{\alpha+\beta}$, $\sigma^2 = \dfrac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)}$.

For $P(\chi^2)$ holds: $\alpha = V/2$ and $\beta = 2$.

8. **The Weibull distribution** is given by:

$$\left\{ \begin{array}{l} P(x) = \dfrac{\alpha}{\beta}x^{\alpha-1}\mathrm{e}^{-x^{\alpha}} \quad \text{if} \quad 0 \leq x \leq \infty \wedge \alpha \wedge \beta > 0 \\ \\ P(x) = 0 \quad \text{in all other cases} \end{array} \right.$$

The average is $\langle x \rangle = \beta^{1/\alpha}\Gamma((\alpha+1)\alpha)$

9. For a **two-dimensional distribution** holds:

$$P_1(x_1) = \int P(x_1, x_2)dx_2 \ , \quad P_2(x_2) = \int P(x_1, x_2)dx_1$$

with

$$\varepsilon(g(x_1, x_2)) = \iint g(x_1, x_2)P(x_1, x_2)dx_1dx_2 = \sum_{x_1}\sum_{x_2} g \cdot P$$

## 2.4 Regression analyses

When there exists a relation between the quantities $x$ and $y$ of the form $y = ax + b$ and there is a measured set $x_i$ with related $y_i$, the following relation holds for $a$ and $b$ with $\vec{x} = (x_1, x_2, ..., x_n)$ and $\vec{e} = (1, 1, ..., 1)$:

$$\vec{y} - a\vec{x} - b\vec{e} \in < \vec{x}, \vec{e} >^{\perp}$$

From this follows that the inner products are 0:

$$\left\{ \begin{array}{l} (\vec{y}, \vec{x}\,) - a(\vec{x}, \vec{x}\,) - b(\vec{e}, \vec{x}\,) = 0 \\ (\vec{y}, \vec{e}\,) - a(\vec{x}, \vec{e}\,) - b(\vec{e}, \vec{e}\,) = 0 \end{array} \right.$$

with $(\vec{x}, \vec{x}\,) = \sum_i x_i^2$, $(\vec{x}, \vec{y}\,) = \sum_i x_i y_i$, $(\vec{x}, \vec{e}\,) = \sum_i x_i$ and $(\vec{e}, \vec{e}\,) = n$. $a$ and $b$ follow from this.

A similar method works for higher order polynomial fits: for a second order fit holds:

$$\vec{y} - a\vec{x^2} - b\vec{x} - c\vec{e} \in < \vec{x^2}, \vec{x}, \vec{e} >^{\perp}$$

with $\vec{x^2} = (x_1^2, ..., x_n^2)$.

The *correlation coefficient* $r$ is a measure for the quality of a fit. In case of linear regression it is given by:

$$r = \frac{n\sum xy - \sum x \sum y}{\sqrt{(n\sum x^2 - (\sum x)^2)(n\sum y^2 - (\sum y)^2)}}$$

# Chapter 3

# Calculus

## 3.1 Integrals

### 3.1.1 Arithmetic rules

The primitive function $F(x)$ of $f(x)$ obeys the rule $F'(x) = f(x)$. With $F(x)$ the primitive of $f(x)$ holds for the definite integral

$$\int_a^b f(x)dx = F(b) - F(a)$$

If $u = f(x)$ holds:

$$\int_a^b g(f(x))df(x) = \int_{f(a)}^{f(b)} g(u)du$$

**Partial integration**: with $F$ and $G$ the primitives of $f$ and $g$ holds:

$$\int f(x) \cdot g(x)dx = f(x)G(x) - \int G(x)\frac{df(x)}{dx}dx$$

A derivative can be brought under the intergral sign (see section 1.8.3 for the required conditions):

$$\frac{d}{dy}\left[\int_{x=g(y)}^{x=h(y)} f(x,y)dx\right] = \int_{x=g(y)}^{x=h(y)} \frac{\partial f(x,y)}{\partial y}dx - f(g(y),y)\frac{dg(y)}{dy} + f(h(y),y)\frac{dh(y)}{dy}$$

### 3.1.2 Arc lengts, surfaces and volumes

The arc length $\ell$ of a curve $y(x)$ is given by:

$$\ell = \int \sqrt{1 + \left(\frac{dy(x)}{dx}\right)^2}dx$$

The arc length $\ell$ of a parameter curve $F(\vec{x}(t))$ is:

$$\ell = \int Fds = \int F(\vec{x}(t))|\dot{\vec{x}}(t)|dt$$

with

$$\vec{t} = \frac{d\vec{x}}{ds} = \frac{\dot{\vec{x}}(t)}{|\dot{\vec{x}}(t)|} \quad , \quad |\vec{t}| = 1$$

$$\int (\vec{v},\vec{t})ds = \int (\vec{v},\dot{\vec{t}}(t))dt = \int (v_1dx + v_2dy + v_3dz)$$

The surface $A$ of a solid of revolution is:

$$A = 2\pi \int y\sqrt{1 + \left(\frac{dy(x)}{dx}\right)^2}dx$$

The volume $V$ of a solid of revolution is:

$$V = \pi \int f^2(x)dx$$

### 3.1.3 Separation of quotients

Every rational function $P(x)/Q(x)$ where $P$ and $Q$ are polynomials can be written as a linear combination of functions of the type $(x-a)^k$ with $k \in \mathbb{Z}$, and of functions of the type

$$\frac{px+q}{((x-a)^2+b^2)^n}$$

with $b > 0$ and $n \in \mathbb{N}$. So:

$$\frac{p(x)}{(x-a)^n} = \sum_{k=1}^{n} \frac{A_k}{(x-a)^k} \quad , \quad \frac{p(x)}{((x-b)^2+c^2)^n} = \sum_{k=1}^{n} \frac{A_k x + B}{((x-b)^2+c^2)^k}$$

Recurrent relation: for $n \neq 0$ holds:

$$\int \frac{dx}{(x^2+1)^{n+1}} = \frac{1}{2n}\frac{x}{(x^2+1)^n} + \frac{2n-1}{2n}\int \frac{dx}{(x^2+1)^n}$$

### 3.1.4 Special functions

#### Elliptic functions

Elliptic functions can be written as a power series as follows:

$$\sqrt{1-k^2\sin^2(x)} = 1 - \sum_{n=1}^{\infty} \frac{(2n-1)!!}{(2n)!!(2n-1)} k^{2n} \sin^{2n}(x)$$

$$\frac{1}{\sqrt{1-k^2\sin^2(x)}} = 1 + \sum_{n=1}^{\infty} \frac{(2n-1)!!}{(2n)!!} k^{2n} \sin^{2n}(x)$$

with $n!! = n(n-2)!!$.

#### The Gamma function

The gamma function $\Gamma(y)$ is defined by:

$$\Gamma(y) = \int_0^{\infty} \mathrm{e}^{-x} x^{y-1} dx$$

One can derive that $\Gamma(y+1) = y\Gamma(y) = y!$. This is a way to define faculties for non-integers. Further one can derive that

$$\Gamma(n+\tfrac{1}{2}) = \frac{\sqrt{\pi}}{2^n}(2n-1)!! \text{ and } \Gamma^{(n)}(y) = \int_0^{\infty} \mathrm{e}^{-x} x^{y-1} \ln^n(x) dx$$

#### The Beta function

The betafunction $\beta(p,q)$ is defined by:

$$\beta(p,q) = \int_0^1 x^{p-1}(1-x)^{q-1} dx$$

with $p$ and $q > 0$. The beta and gamma functions are related by the following equation:

$$\beta(p,q) = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}$$

**The Delta function**

The delta function $\delta(x)$ is an infinitely thin peak function with surface 1. It can be defined by:

$$\delta(x) = \lim_{\varepsilon \to 0} P(\varepsilon, x) \text{ with } P(\varepsilon, x) = \begin{cases} 0 & \text{for } |x| > \varepsilon \\ \dfrac{1}{2\varepsilon} & \text{when } |x| < \varepsilon \end{cases}$$

Some properties are:

$$\int_{-\infty}^{\infty} \delta(x)dx = 1 \ , \quad \int_{-\infty}^{\infty} F(x)\delta(x)dx = F(0)$$

### 3.1.5 Goniometric integrals

When solving goniometric integrals it can be useful to change variables. The following holds if one defines $\tan(\frac{1}{2}x) := t$:

$$dx = \frac{2dt}{1+t^2} \ , \quad \cos(x) = \frac{1-t^2}{1+t^2} \ , \quad \sin(x) = \frac{2t}{1+t^2}$$

Each integral of the type $\int R(x, \sqrt{ax^2+bx+c})dx$ can be converted into one of the types that were treated in **section 3.1.3**. After this conversion one can substitute in the integrals of the type:

$$\int R(x, \sqrt{x^2+1})dx \quad : \quad x = \tan(\varphi) \ , dx = \frac{d\varphi}{\cos(\varphi)} \text{ of } \sqrt{x^2+1} = t + x$$

$$\int R(x, \sqrt{1-x^2})dx \quad : \quad x = \sin(\varphi) \ , dx = \cos(\varphi)d\varphi \text{ of } \sqrt{1-x^2} = 1 - tx$$

$$\int R(x, \sqrt{x^2-1})dx \quad : \quad x = \frac{1}{\cos(\varphi)} \ , dx = \frac{\sin(\varphi)}{\cos^2(\varphi)}d\varphi \text{ of } \sqrt{x^2-1} = x - t$$

These definite integrals are easily solved:

$$\int_0^{\pi/2} \cos^n(x)\sin^m(x)dx = \frac{(n-1)!!(m-1)!!}{(m+n)!!} \cdot \begin{cases} \pi/2 & \text{when } m \text{ and } n \text{ are both even} \\ 1 & \text{in all other cases} \end{cases}$$

Some important integrals are:

$$\int_0^{\infty} \frac{xdx}{e^{ax}+1} = \frac{\pi^2}{12a^2} \ , \quad \int_{-\infty}^{\infty} \frac{x^2dx}{(e^x+1)^2} = \frac{\pi^2}{3} \ , \quad \int_0^{\infty} \frac{x^3dx}{e^x-1} = \frac{\pi^4}{15}$$

## 3.2 Functions with more variables

### 3.2.1 Derivatives

The *partial derivative* with respect to $x$ of a function $f(x, y)$ is defined by:

$$\left(\frac{\partial f}{\partial x}\right)_{x_0} = \lim_{h \to 0} \frac{f(x_0+h, y_0) - f(x_0, y_0)}{h}$$

The *directional derivative* in the direction of $\alpha$ is defined by:

$$\frac{\partial f}{\partial \alpha} = \lim_{r \downarrow 0} \frac{f(x_0 + r\cos(\alpha), y_0 + r\sin(\alpha)) - f(x_0, y_0)}{r} = (\vec{\nabla} f, (\sin\alpha, \cos\alpha)) = \frac{\nabla f \cdot \vec{v}}{|\vec{v}|}$$

When one changes to coordinates $f(x(u,v), y(u,v))$ holds:

$$\frac{\partial f}{\partial u} = \frac{\partial f}{\partial x}\frac{\partial x}{\partial u} + \frac{\partial f}{\partial y}\frac{\partial y}{\partial u}$$

If $x(t)$ and $y(t)$ depend only on one parameter $t$ holds:

$$\frac{\partial f}{\partial t} = \frac{\partial f}{\partial x}\frac{dx}{dt} + \frac{\partial f}{\partial y}\frac{dy}{dt}$$

The *total differential* $df$ of a function of 3 variables is given by:

$$df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z}dz$$

So

$$\frac{df}{dx} = \frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\frac{dy}{dx} + \frac{\partial f}{\partial z}\frac{dz}{dx}$$

The *tangent* in point $\vec{x}_0$ at the surface $f(x,y) = 0$ is given by the equation $f_x(\vec{x}_0)(x - x_0) + f_y(\vec{x}_0)(y - y_0) = 0$.

The *tangent plane* in $\vec{x}_0$ is given by: $f_x(\vec{x}_0)(x - x_0) + f_y(\vec{x}_0)(y - y_0) = z - f(\vec{x}_0)$.

### 3.2.2 Taylor series

A function of two variables can be expanded as follows in a Taylor series:

$$f(x_0 + h, y_0 + k) = \sum_{p=0}^{n} \frac{1}{p!}\left(h\frac{\partial^p}{\partial x^p} + k\frac{\partial^p}{\partial y^p}\right) f(x_0, y_0) + R(n)$$

with $R(n)$ the residual error and

$$\left(h\frac{\partial^p}{\partial x^p} + k\frac{\partial^p}{\partial y^p}\right) f(a,b) = \sum_{m=0}^{p} \binom{p}{m} h^m k^{p-m} \frac{\partial^p f(a,b)}{\partial x^m \partial y^{p-m}}$$

### 3.2.3 Extrema

When $f$ is continuous on a compact boundary $V$ there exists a global maximum and a global minumum for $f$ on this boundary. A boundary is called compact if it is limited and closed.

Possible extrema of $f(x,y)$ on a boundary $V \in \mathbb{R}^2$ are:

1. Points on $V$ where $f(x,y)$ is not differentiable,
2. Points where $\vec{\nabla} f = \vec{0}$,
3. If the boundary $V$ is given by $\varphi(x,y) = 0$, than all points where $\vec{\nabla} f(x,y) + \lambda\vec{\nabla}\varphi(x,y) = 0$ are possible for extrema. This is the multiplicator method of Lagrange, $\lambda$ is called a multiplicator.

The same as in $\mathbb{R}^2$ holds in $\mathbb{R}^3$ when the area to be searched is constrained by a compact $V$, and $V$ is defined by $\varphi_1(x,y,z) = 0$ and $\varphi_2(x,y,z) = 0$ for extrema of $f(x,y,z)$ for points (1) and (2). Point (3) is rewritten as follows: possible extrema are points where $\vec{\nabla} f(x,y,z) + \lambda_1\vec{\nabla}\varphi_1(x,y,z) + \lambda_2\vec{\nabla}\varphi_2(x,y,z) = 0$.

### 3.2.4 The $\nabla$-operator

In cartesian coordinates $(x, y, z)$ holds:

$$\begin{aligned}
\vec{\nabla} &= \frac{\partial}{\partial x}\vec{e}_x + \frac{\partial}{\partial y}\vec{e}_y + \frac{\partial}{\partial z}\vec{e}_z \\
\mathrm{grad} f &= \frac{\partial f}{\partial x}\vec{e}_x + \frac{\partial f}{\partial y}\vec{e}_y + \frac{\partial f}{\partial z}\vec{e}_z \\
\mathrm{div}\ \vec{a} &= \frac{\partial a_x}{\partial x} + \frac{\partial a_y}{\partial y} + \frac{\partial a_z}{\partial z} \\
\mathrm{curl}\ \vec{a} &= \left(\frac{\partial a_z}{\partial y} - \frac{\partial a_y}{\partial z}\right)\vec{e}_x + \left(\frac{\partial a_x}{\partial z} - \frac{\partial a_z}{\partial x}\right)\vec{e}_y + \left(\frac{\partial a_y}{\partial x} - \frac{\partial a_x}{\partial y}\right)\vec{e}_z \\
\nabla^2 f &= \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}
\end{aligned}$$

In cylindrical coordinates $(r, \varphi, z)$ holds:

$$\begin{aligned}
\vec{\nabla} &= \frac{\partial}{\partial r}\vec{e}_r + \frac{1}{r}\frac{\partial}{\partial \varphi}\vec{e}_\varphi + \frac{\partial}{\partial z}\vec{e}_z \\
\mathrm{grad} f &= \frac{\partial f}{\partial r}\vec{e}_r + \frac{1}{r}\frac{\partial f}{\partial \varphi}\vec{e}_\varphi + \frac{\partial f}{\partial z}\vec{e}_z \\
\mathrm{div}\ \vec{a} &= \frac{\partial a_r}{\partial r} + \frac{a_r}{r} + \frac{1}{r}\frac{\partial a_\varphi}{\partial \varphi} + \frac{\partial a_z}{\partial z} \\
\mathrm{curl}\ \vec{a} &= \left(\frac{1}{r}\frac{\partial a_z}{\partial \varphi} - \frac{\partial a_\varphi}{\partial z}\right)\vec{e}_r + \left(\frac{\partial a_r}{\partial z} - \frac{\partial a_z}{\partial r}\right)\vec{e}_\varphi + \left(\frac{\partial a_\varphi}{\partial r} + \frac{a_\varphi}{r} - \frac{1}{r}\frac{\partial a_r}{\partial \varphi}\right)\vec{e}_z \\
\nabla^2 f &= \frac{\partial^2 f}{\partial r^2} + \frac{1}{r}\frac{\partial f}{\partial r} + \frac{1}{r^2}\frac{\partial^2 f}{\partial \varphi^2} + \frac{\partial^2 f}{\partial z^2}
\end{aligned}$$

In spherical coordinates $(r, \theta, \varphi)$ holds:

$$\begin{aligned}
\vec{\nabla} &= \frac{\partial}{\partial r}\vec{e}_r + \frac{1}{r}\frac{\partial}{\partial \theta}\vec{e}_\theta + \frac{1}{r\sin\theta}\frac{\partial}{\partial \varphi}\vec{e}_\varphi \\
\mathrm{grad} f &= \frac{\partial f}{\partial r}\vec{e}_r + \frac{1}{r}\frac{\partial f}{\partial \theta}\vec{e}_\theta + \frac{1}{r\sin\theta}\frac{\partial f}{\partial \varphi}\vec{e}_\varphi \\
\mathrm{div}\ \vec{a} &= \frac{\partial a_r}{\partial r} + \frac{2a_r}{r} + \frac{1}{r}\frac{\partial a_\theta}{\partial \theta} + \frac{a_\theta}{r\tan\theta} + \frac{1}{r\sin\theta}\frac{\partial a_\varphi}{\partial \varphi} \\
\mathrm{curl}\ \vec{a} &= \left(\frac{1}{r}\frac{\partial a_\varphi}{\partial \theta} + \frac{a_\theta}{r\tan\theta} - \frac{1}{r\sin\theta}\frac{\partial a_\theta}{\partial \varphi}\right)\vec{e}_r + \left(\frac{1}{r\sin\theta}\frac{\partial a_r}{\partial \varphi} - \frac{\partial a_\varphi}{\partial r} - \frac{a_\varphi}{r}\right)\vec{e}_\theta + \\
&\quad \left(\frac{\partial a_\theta}{\partial r} + \frac{a_\theta}{r} - \frac{1}{r}\frac{\partial a_r}{\partial \theta}\right)\vec{e}_\varphi \\
\nabla^2 f &= \frac{\partial^2 f}{\partial r^2} + \frac{2}{r}\frac{\partial f}{\partial r} + \frac{1}{r^2}\frac{\partial^2 f}{\partial \theta^2} + \frac{1}{r^2\tan\theta}\frac{\partial f}{\partial \theta} + \frac{1}{r^2\sin^2\theta}\frac{\partial^2 f}{\partial \varphi^2}
\end{aligned}$$

General orthonormal curvilinear coordinates $(u, v, w)$ can be derived from cartesian coordinates by the transformation $\vec{x} = \vec{x}(u, v, w)$. The unit vectors are given by:

$$\vec{e}_u = \frac{1}{h_1}\frac{\partial \vec{x}}{\partial u}\ ,\ \ \vec{e}_v = \frac{1}{h_2}\frac{\partial \vec{x}}{\partial v}\ ,\ \ \vec{e}_w = \frac{1}{h_3}\frac{\partial \vec{x}}{\partial w}$$

where the terms $h_i$ give normalization to length 1. The differential operators are than given by:

$$\mathrm{grad} f = \frac{1}{h_1}\frac{\partial f}{\partial u}\vec{e}_u + \frac{1}{h_2}\frac{\partial f}{\partial v}\vec{e}_v + \frac{1}{h_3}\frac{\partial f}{\partial w}\vec{e}_w$$

$$
\begin{aligned}
\text{div}\,\vec{a} &= \frac{1}{h_1h_2h_3}\left(\frac{\partial}{\partial u}(h_2h_3a_u)+\frac{\partial}{\partial v}(h_3h_1a_v)+\frac{\partial}{\partial w}(h_1h_2a_w)\right)\\
\text{curl}\,\vec{a} &= \frac{1}{h_2h_3}\left(\frac{\partial(h_3a_w)}{\partial v}-\frac{\partial(h_2a_v)}{\partial w}\right)\vec{e}_u+\frac{1}{h_3h_1}\left(\frac{\partial(h_1a_u)}{\partial w}-\frac{\partial(h_3a_w)}{\partial u}\right)\vec{e}_v+\\
&\quad \frac{1}{h_1h_2}\left(\frac{\partial(h_2a_v)}{\partial u}-\frac{\partial(h_1a_u)}{\partial v}\right)\vec{e}_w\\
\nabla^2 f &= \frac{1}{h_1h_2h_3}\left[\frac{\partial}{\partial u}\left(\frac{h_2h_3}{h_1}\frac{\partial f}{\partial u}\right)+\frac{\partial}{\partial v}\left(\frac{h_3h_1}{h_2}\frac{\partial f}{\partial v}\right)+\frac{\partial}{\partial w}\left(\frac{h_1h_2}{h_3}\frac{\partial f}{\partial w}\right)\right]
\end{aligned}
$$

Some properties of the $\nabla$-operator are:

| | | |
|---|---|---|
| $\text{div}(\phi\vec{v})=\phi\text{div}\vec{v}+\text{grad}\phi\cdot\vec{v}$ | $\text{curl}(\phi\vec{v})=\phi\text{curl}\vec{v}+(\text{grad}\phi)\times\vec{v}$ | $\text{curl grad}\phi=\vec{0}$ |
| $\text{div}(\vec{u}\times\vec{v})=\vec{v}\cdot(\text{curl}\vec{u})-\vec{u}\cdot(\text{curl}\vec{v})$ | $\text{curl curl}\vec{v}=\text{grad div}\vec{v}-\nabla^2\vec{v}$ | $\text{div curl}\vec{v}=0$ |
| $\text{div grad}\phi=\nabla^2\phi$ | $\nabla^2\vec{v}\equiv(\nabla^2v_1,\nabla^2v_2,\nabla^2v_3)$ | |

Here, $\vec{v}$ is an arbitrary vectorfield and $\phi$ an arbitrary scalar field.

### 3.2.5 Integral theorems

Some important integral theorems are:

Gauss: $$\oiint(\vec{v}\cdot\vec{n})d^2A=\iiint(\text{div}\vec{v})d^3V$$

Stokes for a scalar field: $$\oint(\phi\cdot\vec{e}_{\text{t}})ds=\iint(\vec{n}\times\text{grad}\phi)d^2A$$

Stokes for a vector field: $$\oint(\vec{v}\cdot\vec{e}_{\text{t}})ds=\iint(\text{curl}\vec{v}\cdot\vec{n})d^2A$$

this gives: $$\oiint(\text{curl}\vec{v}\cdot\vec{n})d^2A=0$$

Ostrogradsky: $$\oiint(\vec{n}\times\vec{v})d^2A=\iiint(\text{curl}\vec{v})d^3A$$

$$\oiint(\phi\vec{n})d^2A=\iiint(\text{grad}\phi)d^3V$$

Here the orientable surface $\iint d^2A$ is bounded by the Jordan curve $s(t)$.

### 3.2.6 Multiple integrals

Let $A$ be a closed curve given by $f(x,y)=0$, than the surface $A$ inside the curve in $\mathbb{R}^2$ is given by

$$A=\iint d^2A=\iint dxdy$$

Let the surface $A$ be defined by the function $z=f(x,y)$. The volume $V$ bounded by $A$ and the $xy$ plane is than given by:

$$V=\iint f(x,y)dxdy$$

The volume inside a closed surface defined by $z=f(x,y)$ is given by:

$$V=\iiint d^3V=\iint f(x,y)dxdy=\iiint dxdydz$$

### 3.2.7 Coordinate transformations

The expressions $d^2A$ and $d^3V$ transform as follows when one changes coordinates to $\vec{u} = (u, v, w)$ through the transformation $x(u, v, w)$:

$$V = \iiint f(x,y,z)dxdydz = \iiint f(\vec{x}(\vec{u})) \left|\frac{\partial \vec{x}}{\partial \vec{u}}\right| dudvdw$$

In $I\!R^2$ holds:

$$\frac{\partial \vec{x}}{\partial \vec{u}} = \begin{vmatrix} x_u & x_v \\ y_u & y_v \end{vmatrix}$$

Let the surface $A$ be defined by $z = F(x, y) = X(u, v)$. Than the volume bounded by the $xy$ plane and $F$ is given by:

$$\iint_S f(\vec{x})d^2A = \iint_G f(\vec{x}(\vec{u})) \left|\frac{\partial X}{\partial u} \times \frac{\partial X}{\partial v}\right| dudv = \iint_G f(x,y,F(x,y))\sqrt{1+\partial_x F^2 + \partial_y F^2}dxdy$$

## 3.3 Orthogonality of functions

The inner product of two functions $f(x)$ and $g(x)$ on the interval $[a, b]$ is given by:

$$(f,g) = \int_a^b f(x)g(x)dx$$

or, when using a weight function $p(x)$, by:

$$(f,g) = \int_a^b p(x)f(x)g(x)dx$$

The *norm* $\|f\|$ follows from: $\|f\|^2 = (f, f)$. A set functions $f_i$ is *orthonormal* if $(f_i, f_j) = \delta_{ij}$.

Each function $f(x)$ can be written as a sum of orthogonal functions:

$$f(x) = \sum_{i=0}^{\infty} c_i g_i(x)$$

and $\sum c_i^2 \leq \|f\|^2$. Let the set $g_i$ be orthogonal, than it follows:

$$c_i = \frac{f, g_i}{(g_i, g_i)}$$

## 3.4 Fourier series

Each function can be written as a sum of independent base functions. When one chooses the orthogonal basis $(\cos(nx), \sin(nx))$ we have a Fourier series.

A periodical function $f(x)$ with period $2L$ can be written as:

$$f(x) = a_0 + \sum_{n=1}^{\infty} \left[a_n \cos\left(\frac{n\pi x}{L}\right) + b_n \sin\left(\frac{n\pi x}{L}\right)\right]$$

Due to the orthogonality follows for the coefficients:

$$a_0 = \frac{1}{2L}\int_{-L}^{L} f(t)dt \;\;,\;\; a_n = \frac{1}{L}\int_{-L}^{L} f(t)\cos\left(\frac{n\pi t}{L}\right)dt \;\;,\;\; b_n = \frac{1}{L}\int_{-L}^{L} f(t)\sin\left(\frac{n\pi t}{L}\right)dt$$

A Fourier series can also be written as a sum of complex exponents:

$$f(x) = \sum_{n=-\infty}^{\infty} c_n \mathrm{e}^{inx}$$

with

$$c_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) \mathrm{e}^{-inx} dx$$

The *Fourier transform* of a function $f(x)$ gives the transformed function $\hat{f}(\omega)$:

$$\hat{f}(\omega) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} f(x) \mathrm{e}^{-i\omega x} dx$$

The inverse transformation is given by:

$$\frac{1}{2}\left[f(x^+) + f(x^-)\right] = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\infty} \hat{f}(\omega) \mathrm{e}^{i\omega x} d\omega$$

where $f(x^+)$ and $f(x^-)$ are defined by the lower - and upper limit:

$$f(a^-) = \lim_{x \uparrow a} f(x) \ , \ \ f(a^+) = \lim_{x \downarrow a} f(x)$$

For continuous functions is $\frac{1}{2}\left[f(x^+) + f(x^-)\right] = f(x)$.

# Chapter 4

# Differential equations

## 4.1 Linear differential equations

### 4.1.1 First order linear DE

The general solution of a linear differential equation is given by $y_{\mathrm{A}} = y_{\mathrm{H}} + y_{\mathrm{P}}$, where $y_{\mathrm{H}}$ is the solution of the *homogeneous equation* and $y_{\mathrm{P}}$ is a *particular solution*.

A first order differential equation is given by: $y'(x) + a(x)y(x) = b(x)$. Its homogeneous equation is $y'(x) + a(x)y(x) = 0$.

The solution of the homogeneous equation is given by

$$y_{\mathrm{H}} = k \exp\left(\int a(x)dx\right)$$

Suppose that $a(x) = a =$constant.

Substitution of $\exp(\lambda x)$ in the homogeneous equation leads to the *characteristic equation* $\lambda + a = 0$
$\Rightarrow \lambda = -a$.

Suppose $b(x) = \alpha \exp(\mu x)$. Than one can distinguish two cases:

1. $\lambda \neq \mu$: a particular solution is: $y_{\mathrm{P}} = \exp(\mu x)$
2. $\lambda = \mu$: a particular solution is: $y_{\mathrm{P}} = x \exp(\mu x)$

When a DE is solved by *variation of parameters* one writes: $y_{\mathrm{P}}(x) = y_{\mathrm{H}}(x)f(x)$, and than one solves $f(x)$ from this.

### 4.1.2 Second order linear DE

A differential equation of the second order with constant coefficients is given by: $y''(x) + ay'(x) + by(x) = c(x)$. If $c(x) = c =$constant there exists a particular solution $y_{\mathrm{P}} = c/b$.

Substitution of $y = \exp(\lambda x)$ leads to the characteristic equation $\lambda^2 + a\lambda + b = 0$.

There are now 2 possibilities:

1. $\lambda_1 \neq \lambda_2$: than $y_{\mathrm{H}} = \alpha \exp(\lambda_1 x) + \beta \exp(\lambda_2 x)$.
2. $\lambda_1 = \lambda_2 = \lambda$: than $y_{\mathrm{H}} = (\alpha + \beta x) \exp(\lambda x)$.

If $c(x) = p(x) \exp(\mu x)$ where $p(x)$ is a polynomial there are 3 possibilities:

1. $\lambda_1, \lambda_2 \neq \mu$: $y_{\mathrm{P}} = q(x) \exp(\mu x)$.
2. $\lambda_1 = \mu, \lambda_2 \neq \mu$: $y_{\mathrm{P}} = xq(x) \exp(\mu x)$.
3. $\lambda_1 = \lambda_2 = \mu$: $y_{\mathrm{P}} = x^2 q(x) \exp(\mu x)$.

where $q(x)$ is a polynomial of the same order as $p(x)$.

When: $y''(x) + \omega^2 y(x) = \omega f(x)$ and $y(0) = y'(0) = 0$ follows: $y(x) = \int_0^x f(x) \sin(\omega(x-t))dt$.

### 4.1.3 The Wronskian

We start with the LDE $y''(x) + p(x)y'(x) + q(x)y(x) = 0$ and the two initial conditions $y(x_0) = K_0$ and $y'(x_0) = K_1$. When $p(x)$ and $q(x)$ are continuous on the open interval $I$ there exists a unique solution $y(x)$ on this interval.

The general solution can than be written as $y(x) = c_1y_1(x) + c_2y_2(x)$ and $y_1$ and $y_2$ are linear independent. These are also *all* solutions of the LDE.

The *Wronskian* is defined by:

$$W(y_1, y_2) = \begin{vmatrix} y_1 & y_2 \\ y_1' & y_2' \end{vmatrix} = y_1y_2' - y_2y_1'$$

$y_1$ and $y_2$ are linear independent if and only if on the interval $I$ when $\exists x_0 \in I$ so that holds:
$W(y_1(x_0), y_2(x_0)) = 0$.

### 4.1.4 Power series substitution

When a series $y = \sum a_nx^n$ is substituted in the LDE with constant coefficients $y''(x) + py'(x) + qy(x) = 0$ this leads to:

$$\sum_n \left[n(n-1)a_nx^{n-2} + pna_nx^{n-1} + qa_nx^n\right] = 0$$

Setting coefficients for equal powers of $x$ equal gives:

$$(n+2)(n+1)a_{n+2} + p(n+1)a_{n+1} + qa_n = 0$$

This gives a general relation between the coefficients. Special cases are $n = 0, 1, 2$.

## 4.2 Some special cases

### 4.2.1 Frobenius' method

Given the LDE

$$\frac{d^2y(x)}{dx^2} + \frac{b(x)}{x}\frac{dy(x)}{dx} + \frac{c(x)}{x^2}y(x) = 0$$

with $b(x)$ and $c(x)$ analytical at $x = 0$. This LDE has at least one solution of the form

$$y_i(x) = x^{r_i}\sum_{n=0}^{\infty} a_nx^n \quad \text{with } i = 1, 2$$

with $r$ real or complex and chosen so that $a_0 \neq 0$. When one expands $b(x)$ and $c(x)$ as $b(x) = b_0 + b_1x + b_2x^2 + ...$ and $c(x) = c_0 + c_1x + c_2x^2 + ...$, it follows for $r$:

$$r^2 + (b_0 - 1)r + c_0 = 0$$

There are now 3 possibilities:

1. $r_1 = r_2$: than $y(x) = y_1(x)\ln|x| + y_2(x)$.
2. $r_1 - r_2 \in \mathbb{N}$: than $y(x) = ky_1(x)\ln|x| + y_2(x)$.
3. $r_1 - r_2 \neq \mathbb{Z}$: than $y(x) = y_1(x) + y_2(x)$.

### 4.2.2 Euler

Given the LDE

$$x^2\frac{d^2y(x)}{dx^2} + ax\frac{dy(x)}{dx} + by(x) = 0$$

Substitution of $y(x) = x^r$ gives an equation for $r$: $r^2 + (a-1)r + b = 0$. From this one gets two solutions $r_1$ and $r_2$. There are now 2 possibilities:

1. $r_1 \neq r_2$: than $y(x) = C_1x^{r1} + C_2x^{r_2}$.
2. $r_1 = r_2 = r$: than $y(x) = (C_1\ln(x) + C_2)x^r$.

### 4.2.3 Legendre's DE

Given the LDE

$$(1-x^2)\frac{d^2y(x)}{dx^2} - 2x\frac{dy(x)}{dx} + n(n-1)y(x) = 0$$

The solutions of this equation are given by $y(x) = aP_n(x) + by_2(x)$ where the *Legendre polynomials* $P(x)$ are defined by:

$$P_n(x) = \frac{d^n}{dx^n}\left(\frac{(1-x^2)^n}{2^n n!}\right)$$

For these holds: $\|P_n\|^2 = 2/(2n+1)$.

### 4.2.4 The associated Legendre equation

This equation follows from the $\theta$-dependent part of the wave equation $\nabla^2\Psi = 0$ by substitution of $\xi = \cos(\theta)$. Than follows:

$$(1-\xi^2)\frac{d}{d\xi}\left((1-\xi^2)\frac{dP(\xi)}{d\xi}\right) + [C(1-\xi^2) - m^2]P(\xi) = 0$$

Regular solutions exists only if $C = l(l+1)$. They are of the form:

$$P_l^{|m|}(\xi) = (1-\xi^2)^{m/2}\frac{d^{|m|}P^0(\xi)}{d\xi^{|m|}} = \frac{(1-\xi^2)^{|m|/2}}{2^l l!}\frac{d^{|m|+l}}{d\xi^{|m|+l}}(\xi^2-1)^l$$

For $|m| > l$ is $P_l^{|m|}(\xi) = 0$. Some properties of $P_l^0(\xi)$ zijn:

$$\int_{-1}^{1} P_l^0(\xi)P_{l'}^0(\xi)d\xi = \frac{2}{2l+1}\delta_{ll'} \quad , \quad \sum_{l=0}^{\infty} P_l^0(\xi)t^l = \frac{1}{\sqrt{1-2\xi t+t^2}}$$

This polynomial can be written as:

$$P_l^0(\xi) = \frac{1}{\pi}\int_0^{\pi}(\xi + \sqrt{\xi^2-1}\cos(\theta))^l d\theta$$

### 4.2.5 Solutions for Bessel's equation

Given the LDE

$$x^2\frac{d^2y(x)}{dx^2} + x\frac{dy(x)}{dx} + (x^2-\nu^2)y(x) = 0$$

also called *Bessel's equation*, and the Bessel functions of the first kind

$$J_\nu(x) = x^\nu\sum_{m=0}^{\infty}\frac{(-1)^m x^{2m}}{2^{2m+\nu}m!\Gamma(\nu+m+1)}$$

for $\nu := n \in \mathbb{N}$ this becomes:

$$J_n(x) = x^n \sum_{m=0}^{\infty} \frac{(-1)^m x^{2m}}{2^{2m+n} m!(n+m)!}$$

When $\nu \neq \mathbb{Z}$ the solution is given by $y(x) = aJ_\nu(x) + bJ_{-\nu}(x)$. But because for $n \in \mathbb{Z}$ holds: $J_{-n}(x) = (-1)^n J_n(x)$, this does not apply to integers. The general solution of Bessel's equation is given by $y(x) = aJ_\nu(x) + bY_\nu(x)$, where $Y_\nu$ are the *Bessel functions of the second kind*:

$$Y_\nu(x) = \frac{J_\nu(x)\cos(\nu\pi) - J_{-\nu}(x)}{\sin(\nu\pi)} \quad \text{and} \quad Y_n(x) = \lim_{\nu \to n} Y_\nu(x)$$

The equation $x^2y''(x) + xy'(x) - (x^2 + \nu^2)y(x) = 0$ has the modified Bessel functions of the first kind $I_\nu(x) = i^{-\nu}J_\nu(ix)$ as solution, and also solutions $K_\nu = \pi[I_{-\nu}(x) - I_\nu(x)]/[2\sin(\nu\pi)]$.

Sometimes it can be convenient to write the solutions of Bessel's equation in terms of the Hankel functions

$$H_n^{(1)}(x) = J_n(x) + iY_n(x) \ , \ H_n^{(2)}(x) = J_n(x) - iY_n(x)$$

### 4.2.6 Properties of Bessel functions

Bessel functions are orthogonal with respect to the weight function $p(x) = x$.

$J_{-n}(x) = (-1)^n J_n(x)$. The Neumann functions $N_m(x)$ are definied as:

$$N_m(x) = \frac{1}{2\pi} J_m(x)\ln(x) + \frac{1}{x^m} \sum_{n=0}^{\infty} \alpha_n x^{2n}$$

The following holds: $\lim_{x\to 0} J_m(x) = x^m$, $\lim_{x\to 0} N_m(x) = x^{-m}$ for $m \neq 0$, $\lim_{x\to 0} N_0(x) = \ln(x)$.

$$\lim_{r\to\infty} H(r) = \frac{\mathrm{e}^{\pm ikr}\mathrm{e}^{i\omega t}}{\sqrt{r}} \ , \ \lim_{x\to\infty} J_n(x) = \sqrt{\frac{2}{\pi x}}\cos(x - x_n) \ , \ \lim_{x\to\infty} J_{-n}(x) = \sqrt{\frac{2}{\pi x}}\sin(x - x_n)$$

with $x_n = \frac{1}{2}\pi(n + \frac{1}{2})$.

$$J_{n+1}(x) + J_{n-1}(x) = \frac{2n}{x} J_n(x) \ , \ J_{n+1}(x) - J_{n-1}(x) = -2\frac{dJ_n(x)}{dx}$$

The following integral relations hold:

$$J_m(x) = \frac{1}{2\pi}\int_0^{2\pi} \exp[i(x\sin(\theta) - m\theta)]d\theta = \frac{1}{\pi}\int_0^{\pi} \cos(x\sin(\theta) - m\theta)d\theta$$

### 4.2.7 Laguerre's equation

Given the LDE

$$x\frac{d^2y(x)}{dx^2} + (1-x)\frac{dy(x)}{dx} + ny(x) = 0$$

Solutions of this equation are the Laguerre polynomials $L_n(x)$:

$$L_n(x) = \frac{\mathrm{e}^x}{n!}\frac{d^n}{dx^n}\left(x^n\mathrm{e}^{-x}\right) = \sum_{m=0}^{\infty} \frac{(-1)^m}{m!}\binom{n}{m} x^m$$

### 4.2.8 The associated Laguerre equation

Given the LDE

$$\frac{d^2y(x)}{dx^2} + \left(\frac{m+1}{x} - 1\right)\frac{dy(x)}{dx} + \left(\frac{n + \frac{1}{2}(m+1)}{x}\right)y(x) = 0$$

Solutions of this equation are the associated Laguerre polynomials $L_n^m(x)$:

$$L_n^m(x) = \frac{(-1)^m n!}{(n-m)!}\mathrm{e}^{-x}x^{-m}\frac{d^{n-m}}{dx^{n-m}}\left(\mathrm{e}^{-x}x^n\right)$$

### 4.2.9 Hermite

The differential equations of Hermite are:

$$\frac{d^2\mathrm{H}_n(x)}{dx^2} - 2x\frac{d\mathrm{H}_n(x)}{dx} + 2n\mathrm{H}_n(x) = 0 \text{ and } \frac{d^2\mathrm{He}_n(x)}{dx^2} - x\frac{d\mathrm{He}_n(x)}{dx} + n\mathrm{He}_n(x) = 0$$

Solutions of these equations are the Hermite polynomials, given by:

$$\mathrm{H}_n(x) = (-1)^n \exp\left(\frac{1}{2}x^2\right)\frac{d^n(\exp(-\frac{1}{2}x^2))}{dx^n} = 2^{n/2}\mathrm{He}_n(x\sqrt{2})$$

$$\mathrm{He}_n(x) = (-1)^n(\exp\left(x^2\right)\frac{d^n(\exp(-x^2))}{dx^n} = 2^{-n/2}\mathrm{H}_n(x/\sqrt{2})$$

### 4.2.10 Chebyshev

The LDE

$$(1-x^2)\frac{d^2U_n(x)}{dx^2} - 3x\frac{dU_n(x)}{dx} + n(n+2)U_n(x) = 0$$

has solutions of the form

$$U_n(x) = \frac{\sin[(n+1)\arccos(x)]}{\sqrt{1-x^2}}$$

The LDE

$$(1-x^2)\frac{d^2T_n(x)}{dx^2} - x\frac{dT_n(x)}{dx} + n^2T_n(x) = 0$$

has solutions $T_n(x) = \cos(n\arccos(x))$.

### 4.2.11 Weber

The LDE $W_n''(x) + (n + \frac{1}{2} - \frac{1}{4}x^2)W_n(x) = 0$ has solutions: $W_n(x) = \mathrm{He}_n(x)\exp(-\frac{1}{4}x^2)$.

## 4.3 Non-linear differential equations

Some non-linear differential equations and a solution are:

| | |
|---|---|
| $y' = a\sqrt{y^2+b^2}$ | $y = b\sinh(a(x-x_0))$ |
| $y' = a\sqrt{y^2-b^2}$ | $y = b\cosh(a(x-x_0))$ |
| $y' = a\sqrt{b^2-y^2}$ | $y = b\cos(a(x-x_0))$ |
| $y' = a(y^2+b^2)$ | $y = b\tan(a(x-x_0))$ |
| $y' = a(y^2-b^2)$ | $y = b\coth(a(x-x_0))$ |
| $y' = a(b^2-y^2)$ | $y = b\tanh(a(x-x_0))$ |
| $y' = ay\left(\frac{b-y}{b}\right)$ | $y = \frac{b}{1+Cb\exp(-ax)}$ |

## 4.4 Sturm-Liouville equations

Sturm-Liouville equations are second order LDE's of the form:

$$-\frac{d}{dx}\left(p(x)\frac{dy(x)}{dx}\right)+q(x)y(x)=\lambda m(x)y(x)$$

The boundary conditions are chosen so that the operator

$$L=-\frac{d}{dx}\left(p(x)\frac{d}{dx}\right)+q(x)$$

is Hermitean. The normalization function $m(x)$ must satisfy

$$\int_a^b m(x)y_i(x)y_j(x)dx=\delta_{ij}$$

When $y_1(x)$ and $y_2(x)$ are two linear independent solutions one can write the Wronskian in this form:

$$W(y_1,y_2)=\begin{vmatrix} y_1 & y_2 \\ y_1' & y_2' \end{vmatrix}=\frac{C}{p(x)}$$

where $C$ is constant. By changing to another dependent variable $u(x)$, given by: $u(x)=y(x)\sqrt{p(x)}$, the LDE transforms into the *normal form*:

$$\frac{d^2u(x)}{dx^2}+I(x)u(x)=0 \quad \text{with} \quad I(x)=\frac{1}{4}\left(\frac{p'(x)}{p(x)}\right)^2-\frac{1}{2}\frac{p''(x)}{p(x)}-\frac{q(x)-\lambda m(x)}{p(x)}$$

If $I(x)>0$, than $y''/y<0$ and the solution has an oscillatory behaviour, if $I(x)<0$, than $y''/y>0$ and the solution has an exponential behaviour.

## 4.5 Linear partial differential equations

### 4.5.1 General

The *normal derivative* is defined by:

$$\frac{\partial u}{\partial n}=(\vec{\nabla}u,\vec{n})$$

A frequently used solution method for PDE's is *separation of variables*: one assumes that the solution can be written as $u(x,t)=X(x)T(t)$. When this is substituted two ordinary DE's for $X(x)$ and $T(t)$ are obtained.

### 4.5.2 Special cases

#### The wave equation

The *wave equation* in 1 dimension is given by

$$\frac{\partial^2 u}{\partial t^2}=c^2\frac{\partial^2 u}{\partial x^2}$$

When the initial conditions $u(x,0)=\varphi(x)$ and $\partial u(x,0)/\partial t=\Psi(x)$ apply, the general solution is given by:

$$u(x,t)=\frac{1}{2}\left[\varphi(x+ct)+\varphi(x-ct)\right]+\frac{1}{2c}\int_{x-ct}^{x+ct}\Psi(\xi)d\xi$$

**The diffusion equation**

The *diffusion equation* is:

$$\frac{\partial u}{\partial t} = D\nabla^2 u$$

Its solutions can be written in terms of the propagators $P(x, x', t)$. These have the property that $P(x, x', 0) = \delta(x - x')$. In 1 dimension it reads:

$$P(x, x', t) = \frac{1}{2\sqrt{\pi Dt}} \exp\left(\frac{-(x - x')^2}{4Dt}\right)$$

In 3 dimensions it reads:

$$P(x, x', t) = \frac{1}{8(\pi Dt)^{3/2}} \exp\left(\frac{-(\vec{x} - \vec{x}\,')^2}{4Dt}\right)$$

With initial condition $u(x, 0) = f(x)$ the solution is:

$$u(x, t) = \int_{\mathcal{G}} f(x')P(x, x', t)dx'$$

The solution of the equation

$$\frac{\partial u}{\partial t} - D\frac{\partial^2 u}{\partial x^2} = g(x, t)$$

is given by

$$u(x, t) = \int dt' \int dx' g(x', t')P(x, x', t - t')$$

**The equation of Helmholtz**

The equation of Helmholtz is obtained by substitution of $u(\vec{x}, t) = v(\vec{x}) \exp(i\omega t)$ in the wave equation. This gives for $v$:

$$\nabla^2 v(\vec{x}, \omega) + k^2 v(\vec{x}, \omega) = 0$$

This gives as solutions for $v$:

1. In cartesian coordinates: substitution of $v = A \exp(i\vec{k} \cdot \vec{x}\,)$ gives:

$$v(\vec{x}\,) = \int \cdots \int A(k)\mathrm{e}^{i\vec{k}\cdot\vec{x}}dk$$

   with the integrals over $\vec{k}^{\,2} = k^2$.

2. In polar coordinates:

$$v(r, \varphi) = \sum_{m=0}^{\infty} (A_m J_m(kr) + B_m N_m(kr))\mathrm{e}^{im\varphi}$$

3. In spherical coordinates:

$$v(r, \theta, \varphi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} [A_{lm} J_{l+\frac{1}{2}}(kr) + B_{lm} J_{-l-\frac{1}{2}}(kr)] \frac{Y(\theta, \varphi)}{\sqrt{r}}$$

### 4.5.3 Potential theory and Green's theorem

Subject of the potential theory are the *Poisson equation* $\nabla^2 u = -f(\vec{x}\,)$ where $f$ is a given function, and the *Laplace equation* $\nabla^2 u = 0$. The solutions of these can often be interpreted as a potential. The solutions of Laplace's equation are called *harmonic functions*.

When a vector field $\vec{v}$ is given by $\vec{v} = \mathrm{grad}\varphi$ holds:

$$\int_a^b (\vec{v}, \vec{t}\,)ds = \varphi(\vec{b}\,) - \varphi(\vec{a}\,)$$

In this case there exist functions $\varphi$ and $\vec{w}$ so that $\vec{v} = \mathrm{grad}\varphi + \mathrm{curl}\vec{w}$.

The *field lines* of the field $\vec{v}(\vec{x}\,)$ follow from:

$$\dot{\vec{x}}\,(t) = \lambda \vec{v}(\vec{x}\,)$$

The *first theorem of Green* is:

$$\iiint_{\mathcal{G}} [u\nabla^2 v + (\nabla u, \nabla v)]d^3V = \oiint_{\mathcal{S}} u\frac{\partial v}{\partial n}d^2A$$

The *second theorem of Green* is:

$$\iiint_{\mathcal{G}} [u\nabla^2 v - v\nabla^2 u]d^3V = \oiint_{\mathcal{S}} \left(u\frac{\partial v}{\partial n} - v\frac{\partial u}{\partial n}\right) d^2A$$

A harmonic function which is 0 on the boundary of an area is also 0 within that area. A harmonic function with a normal derivative of 0 on the boundary of an area is constant within that area.

The *Dirichlet problem* is:

$$\nabla^2 u(\vec{x}\,) = -f(\vec{x}\,) \ , \ \vec{x} \in R \ , \ u(\vec{x}\,) = g(\vec{x}\,) \ \text{ for all } \ \vec{x} \in S.$$

It has a unique solution.

The *Neumann problem* is:

$$\nabla^2 u(\vec{x}\,) = -f(\vec{x}\,) \ , \ \vec{x} \in R \ , \ \frac{\partial u(\vec{x}\,)}{\partial n} = h(\vec{x}\,) \ \text{ for all } \ \vec{x} \in S.$$

The solution is unique except for a constant. The solution exists if:

$$-\iiint_R f(\vec{x}\,)d^3V = \oiint_S h(\vec{x}\,)d^2A$$

A *fundamental solution* of the Laplace equation satisfies:

$$\nabla^2 u(\vec{x}\,) = -\delta(\vec{x}\,)$$

This has in 2 dimensions in polar coordinates the following solution:

$$u(r) = \frac{\ln(r)}{2\pi}$$

This has in 3 dimensions in spherical coordinates the following solution:

$$u(r) = \frac{1}{4\pi r}$$

The equation $\nabla^2 v = -\delta(\vec{x} - \vec{\xi})$ has the solution

$$v(\vec{x}) = \frac{1}{4\pi|\vec{x} - \vec{\xi}|}$$

After substituting this in Green's 2nd theorem and applying the sieve property of the $\delta$ function one can derive Green's 3rd theorem:

$$u(\vec{\xi}) = -\frac{1}{4\pi} \iiint_R \frac{\nabla^2 u}{r} d^3V + \frac{1}{4\pi} \oiint_S \left[ \frac{1}{r} \frac{\partial u}{\partial n} - u \frac{\partial}{\partial n} \left( \frac{1}{r} \right) \right] d^2A$$

The *Green function* $G(\vec{x}, \vec{\xi})$ is defined by: $\nabla^2 G = -\delta(\vec{x} - \vec{\xi})$, and on boundary $S$ holds $G(\vec{x}, \vec{\xi}) = 0$. Than $G$ can be written as:

$$G(\vec{x}, \vec{\xi}) = \frac{1}{4\pi|\vec{x} - \vec{\xi}|} + g(\vec{x}, \vec{\xi})$$

Than $g(\vec{x}, \vec{\xi})$ is a solution of Dirichlet's problem. The solution of Poisson's equation $\nabla^2 u = -f(\vec{x})$ when on the boundary $S$ holds: $u(\vec{x}) = g(\vec{x})$, is:

$$u(\vec{\xi}) = \iiint_R G(\vec{x}, \vec{\xi}) f(\vec{x}) d^3V - \oiint_S g(\vec{x}) \frac{\partial G(\vec{x}, \vec{\xi})}{\partial n} d^2A$$

# Chapter 5

# Linear algebra

## 5.1 Vector spaces

$\mathcal{G}$ is a group for the operation $\otimes$ if:

1. $\forall a, b \in \mathcal{G} \Rightarrow a \otimes b \in \mathcal{G}$: a group is *closed*.
2. $(a \otimes b) \otimes c = a \otimes (b \otimes c)$: a group is *associative*.
3. $\exists e \in \mathcal{G}$ so that $a \otimes e = e \otimes a = a$: there exists a *unit element*.
4. $\forall a \in \mathcal{G} \exists \overline{a} \in \mathcal{G}$ so that $a \otimes \overline{a} = e$: each element has an *inverse*.

If
5. $a \otimes b = b \otimes a$

the group is called *Abelian* or *commutative*. Vector spaces form an Abelian group for addition and multiplication: $1 \cdot \vec{a} = \vec{a}$, $\lambda(\mu\vec{a}) = (\lambda\mu)\vec{a}$, $(\lambda + \mu)(\vec{a} + \vec{b}) = \lambda\vec{a} + \lambda\vec{b} + \mu\vec{a} + \mu\vec{b}$.

$W$ is a *linear subspace* if $\forall \vec{w}_1, \vec{w}_2 \in W$ holds: $\lambda\vec{w}_1 + \mu\vec{w}_2 \in W$.

$W$ is an *invariant subspace* of $V$ for the operator $A$ if $\forall \vec{w} \in W$ holds: $A\vec{w} \in W$.

## 5.2 Basis

For an orthogonal basis holds: $(\vec{e}_i, \vec{e}_j) = c\delta_{ij}$. For an orthonormal basis holds: $(\vec{e}_i, \vec{e}_j) = \delta_{ij}$.

The set vectors $\{\vec{a}_n\}$ is linear independent if:

$$\sum_i \lambda_i \vec{a}_i = 0 \quad \Leftrightarrow \quad \forall_i \lambda_i = 0$$

The set $\{\vec{a}_n\}$ is a basis if it is 1. independent and 2. $V =< \vec{a}_1, \vec{a}_2, ... >= \sum \lambda_i \vec{a}_i$.

## 5.3 Matrix calculus

### 5.3.1 Basic operations

For the matrix multiplication of matrices $A = a_{ij}$ and $B = b_{kl}$ holds with $^r$ the row index and $^k$ the column index:

$$A^{r_1 k_1} \cdot B^{r_2 k_2} = C^{r_1 k_2} \ , \ (AB)_{ij} = \sum_k a_{ik} b_{kj}$$

where $^r$ is the number of rows and $^k$ the number of columns.

The *transpose* of $A$ is defined by: $a^T_{ij} = a_{ji}$. For this holds $(AB)^T = B^T A^T$ and $(A^T)^{-1} = (A^{-1})^T$. For the *inverse matrix* holds: $(A \cdot B)^{-1} = B^{-1} \cdot A^{-1}$. The inverse matrix $A^{-1}$ has the property that $A \cdot A^{-1} = I\!I$ and can be found by diagonalization: $(A_{ij}|I\!I) \sim (I\!I|A^{-1}_{ij})$.

The inverse of a $2 \times 2$ matrix is:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

The *determinant function* $D = \det(A)$ is defined by:

$$\det(A) = D(\vec{a}_{*1}, \vec{a}_{*2}, ..., \vec{a}_{*n})$$

For the determinant $\det(A)$ of a matrix $A$ holds: $\det(AB) = \det(A) \cdot \det(B)$. Een $2 \times 2$ matrix has determinant:

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - cb$$

The derivative of a matrix is a matrix with the derivatives of the coefficients:

$$\frac{dA}{dt} = \frac{da_{ij}}{dt} \quad \text{and} \quad \frac{dAB}{dt} = B\frac{dA}{dt} + A\frac{dB}{dt}$$

The derivative of the determinant is given by:

$$\frac{d\det(A)}{dt} = D(\frac{d\vec{a}_1}{dt}, ..., \vec{a}_n) + D(\vec{a}_1, \frac{d\vec{a}_2}{dt}, ..., \vec{a}_n) + ... + D(\vec{a}_1, ..., \frac{d\vec{a}_n}{dt})$$

When the rows of a matrix are considered as vectors the *row rank* of a matrix is the number of independent vectors in this set. Similar for the *column rank*. The row rank equals the column rank for each matrix.

Let $\tilde{A} : \tilde{V} \to \tilde{V}$ be the complex extension of the real linear operator $A : V \to V$ in a finite dimensional $V$. Then $A$ and $\tilde{A}$ have the same caracteristic equation.

When $A_{ij} \in I\!R$ and $\vec{v}_1 + i\vec{v}_2$ is an eigenvector of $A$ at eigenvalue $\lambda = \lambda_1 + i\lambda_2$, than holds:

1. $A\vec{v}_1 = \lambda_1\vec{v}_1 - \lambda_2\vec{v}_2$ and $A\vec{v}_2 = \lambda_2\vec{v}_1 + \lambda_1\vec{v}_2$.
2. $\vec{v}^{\,*} = \vec{v}_1 - i\vec{v}_2$ is an eigenvalue at $\lambda^* = \lambda_1 - i\lambda_2$.
3. The linear span $< \vec{v}_1, \vec{v}_2 >$ is an invariant subspace of $A$.

If $\vec{k}_n$ are the columns of $A$, than the transformed space of $A$ is given by:

$$R(A) =< A\vec{e}_1, ..., A\vec{e}_n >=< \vec{k}_1, ..., \vec{k}_n >$$

If the columns $\vec{k}_n$ of a $n \times m$ matrix $A$ are independent, than the nullspace $\mathcal{N}(A) = \{\vec{0}\,\}$.

### 5.3.2 Matrix equations

We start with the equation

$$A \cdot \vec{x} = \vec{b}$$

and $\vec{b} \neq \vec{0}$. If $\det(A) = 0$ the only solution is $\vec{0}$. If $\det(A) \neq 0$ there exists exactly one solution $\neq \vec{0}$.

The equation

$$A \cdot \vec{x} = \vec{0}$$

has exactly one solution $\neq \vec{0}$ if $\det(A) = 0$, and if $\det(A) \neq 0$ the solution is $\vec{0}$.

Cramer's rule for the solution of systems of linear equations is: let the system be written as

$$A \cdot \vec{x} = \vec{b} \equiv \vec{a}_1 x_1 + ... + \vec{a}_n x_n = \vec{b}$$

then $x_j$ is given by:

$$x_j = \frac{D(\vec{a}_1, ..., \vec{a}_{j-1}, \vec{b}, \vec{a}_{j+1}, ..., \vec{a}_n)}{\det(A)}$$

## 5.4 Linear transformations

A transformation $A$ is linear if: $A(\lambda\vec{x} + \beta\vec{y}) = \lambda A\vec{x} + \beta A\vec{y}$.

Some common linear transformations are:

| Transformation type | Equation |
|---|---|
| Projection on the line $< \vec{a} >$ | $P(\vec{x}) = (\vec{a}, \vec{x})\vec{a}/(\vec{a}, \vec{a})$ |
| Projection on the plane $(\vec{a}, \vec{x}) = 0$ | $Q(\vec{x}) = \vec{x} - P(\vec{x})$ |
| Mirror image in the line $< \vec{a} >$ | $S(\vec{x}) = 2P(\vec{x}) - \vec{x}$ |
| Mirror image in the plane $(\vec{a}, \vec{x}) = 0$ | $T(\vec{x}) = 2Q(\vec{x}) - \vec{x} = \vec{x} - 2P(\vec{x})$ |

For a projection holds: $\vec{x} - P_W(\vec{x}) \perp P_W(\vec{x})$ and $P_W(\vec{x}) \in W$.

If for a transformation $A$ holds: $(A\vec{x}, \vec{y}) = (\vec{x}, A\vec{y}) = (A\vec{x}, A\vec{y})$, than $A$ is a projection.

Let $A : W \to W$ define a linear transformation; we define:

- If $S$ is a subset of $V$: $A(S) := \{A\vec{x} \in W | \vec{x} \in S\}$
- If $T$ is a subset of $W$: $A^{\leftarrow}(T) := \{\vec{x} \in V | A(\vec{x}) \in T\}$

Than $A(S)$ is a linear subspace of $W$ and the *inverse transformation* $A^{\leftarrow}(T)$ is a linear subspace of $V$. From this follows that $A(V)$ is the *image space* of $A$, notation: $\mathcal{R}(A)$. $A^{\leftarrow}(\vec{0}) = E_0$ is a linear subspace of $V$, the *null space* of $A$, notation: $\mathcal{N}(A)$. Then the following holds:

$$\dim(\mathcal{N}(A)) + \dim(\mathcal{R}(A)) = \dim(V)$$

## 5.5 Plane and line

The equation of a line that contains the points $\vec{a}$ and $\vec{b}$ is:

$$\vec{x} = \vec{a} + \lambda(\vec{b} - \vec{a}) = \vec{a} + \lambda\vec{r}$$

The equation of a plane is:

$$\vec{x} = \vec{a} + \lambda(\vec{b} - \vec{a}) + \mu(\vec{c} - \vec{a}) = \vec{a} + \lambda\vec{r}_1 + \mu\vec{r}_2$$

When this is a plane in $\mathbb{R}^3$, the *normal vector* to this plane is given by:

$$\vec{n}_V = \frac{\vec{r}_1 \times \vec{r}_2}{|\vec{r}_1 \times \vec{r}_2|}$$

A line can also be described by the points for which the line equation $\ell$: $(\vec{a}, \vec{x}) + b = 0$ holds, and for a plane V: $(\vec{a}, \vec{x}) + k = 0$. The normal vector to V is than: $\vec{a}/|\vec{a}|$.

The distance $d$ between 2 points $\vec{p}$ and $\vec{q}$ is given by $d(\vec{p}, \vec{q}) = \|\vec{p} - \vec{q}\|$.

In $\mathbb{R}^2$ holds: The distance of a point $\vec{p}$ to the line $(\vec{a}, \vec{x}) + b = 0$ is

$$d(\vec{p}, \ell) = \frac{|(\vec{a}, \vec{p}) + b|}{|\vec{a}|}$$

Similarly in $\mathbb{R}^3$: The distance of a point $\vec{p}$ to the plane $(\vec{a}, \vec{x}) + k = 0$ is

$$d(\vec{p}, V) = \frac{|(\vec{a}, \vec{p}) + k|}{|\vec{a}|}$$

This can be generalized for $\mathbb{R}^n$ and $\mathbb{C}^n$ (theorem from Hesse).

## 5.6 Coordinate transformations

The linear transformation $A$ from $\mathbb{K}^n \rightarrow \mathbb{K}^m$ is given by ($\mathbb{K} = \mathbb{R}$ of $\mathbb{C}$):

$$\vec{y} = A^{m \times n}\vec{x}$$

where a column of $A$ is the image of a base vector in the original.

The matrix $A^{\beta}_{\alpha}$ transforms a vector given w.r.t. a basis $\alpha$ into a vector w.r.t. a basis $\beta$. It is given by:

$$A^{\beta}_{\alpha} = (\beta(A\vec{a}_1), ..., \beta(A\vec{a}_n))$$

where $\beta(\vec{x}\,)$ is the representation of the vector $\vec{x}$ w.r.t. basis $\beta$.

The *transformation matrix* $S^{\beta}_{\alpha}$ transforms vectors from coordinate system $\alpha$ into coordinate system $\beta$:

$$S^{\beta}_{\alpha} := \mathbb{I}^{\beta}_{\alpha} = (\beta(\vec{a}_1), ..., \beta(\vec{a}_n))$$

and $S^{\beta}_{\alpha} \cdot S^{\alpha}_{\beta} = \mathbb{I}$

The matrix of a transformation $A$ is than given by:

$$A^{\beta}_{\alpha} = \left(A^{\beta}_{\alpha}\vec{e}_1, ..., A^{\beta}_{\alpha}\vec{e}_n\right)$$

For the transformation of matrix operators to another coordinate system holds: $A^{\delta}_{\alpha} = S^{\delta}_{\lambda}A^{\lambda}_{\beta}S^{\beta}_{\alpha}$, $A^{\alpha}_{\alpha} = S^{\alpha}_{\beta}A^{\beta}_{\beta}S^{\beta}_{\alpha}$ and $(AB)^{\lambda}_{\alpha} = A^{\lambda}_{\beta}B^{\beta}_{\alpha}$.

Further is $A^{\beta}_{\alpha} = S^{\beta}_{\alpha}A^{\alpha}_{\alpha}$, $A^{\alpha}_{\beta} = A^{\alpha}_{\alpha}S^{\alpha}_{\beta}$. A vector is transformed via $X_{\alpha} = S^{\beta}_{\alpha}X_{\beta}$.

## 5.7 Eigen values

The *eigenvalue equation*

$$A\vec{x} = \lambda\vec{x}$$

with *eigenvalues* $\lambda$ can be solved with $(A - \lambda\mathbb{I}) = \vec{0} \Rightarrow \det(A - \lambda\mathbb{I}) = 0$. The eigenvalues follow from this characteristic equation. The following is true: $\det(A) = \prod_i \lambda_i$ and $\mathrm{Tr}(A) = \sum_i a_{ii} = \sum_i \lambda_i$.

The eigen values $\lambda_i$ are independent of the chosen basis. The matrix of $A$ in a basis of eigenvectors, with $S$ the transformation matrix to this basis, $S = (E_{\lambda_1}, ..., E_{\lambda_n})$, is given by:

$$\Lambda = S^{-1}AS = \mathrm{diag}(\lambda_1, ..., \lambda_n)$$

When 0 is an eigen value of $A$ than $E_0(A) = \mathcal{N}(A)$.

When $\lambda$ is an eigen value of $A$ holds: $A^n\vec{x} = \lambda^n\vec{x}$.

## 5.8 Transformation types

**Isometric transformations**

A transformation is *isometric* when: $\|A\vec{x}\| = \|\vec{x}\|$. This implies that the eigen values of an isometric transformation are given by $\lambda = \exp(i\varphi) \Rightarrow |\lambda| = 1$. Than also holds: $(A\vec{x}, A\vec{y}\,) = (\vec{x}, \vec{y}\,)$.

When $W$ is an invariant subspace if the isometric transformation $A$ with $\dim(A) < \infty$, than also $W^{\perp}$ is an invariante subspace.

### Orthogonal transformations

A transformation $A$ is *orthogonal* if $A$ is isometric *and* the inverse $A^{\leftarrow}$ exists. For an orthogonal transformation $O$ holds $O^T O = I\!I$, so: $O^T = O^{-1}$. If $A$ and $B$ are orthogonal, than $AB$ and $A^{-1}$ are also orthogonal.

Let $A : V \to V$ be orthogonal with $\dim(V) < \infty$. Than $A$ is:

**Direct orthogonal** if $\det(A) = +1$. $A$ describes a rotation. A rotation in $I\!R^2$ through angle $\varphi$ is given by:

$$R = \begin{pmatrix} \cos(\varphi) & -\sin(\varphi) \\ \sin(\varphi) & \cos(\varphi) \end{pmatrix}$$

So the rotation angle $\varphi$ is determined by $\mathrm{Tr}(A) = 2\cos(\varphi)$ with $0 \leq \varphi \leq \pi$. Let $\lambda_1$ and $\lambda_2$ be the roots of the characteristic equation, than also holds: $\Re(\lambda_1) = \Re(\lambda_2) = \cos(\varphi)$, and $\lambda_1 = \exp(i\varphi)$, $\lambda_2 = \exp(-i\varphi)$.

In $I\!R^3$ holds: $\lambda_1 = 1$, $\lambda_2 = \lambda_3^* = \exp(i\varphi)$. A rotation over $E_{\lambda_1}$ is given by the matrix

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\varphi) & -\sin(\varphi) \\ 0 & \sin(\varphi) & \cos(\varphi) \end{pmatrix}$$

**Mirrored orthogonal** if $\det(A) = -1$. Vectors from $E_{-1}$ are mirrored by $A$ w.r.t. the invariant subspace $E_{-1}^{\perp}$. A mirroring in $I\!R^2$ in $< (\cos(\frac{1}{2}\varphi), \sin(\frac{1}{2}\varphi)) >$ is given by:

$$S = \begin{pmatrix} \cos(\varphi) & \sin(\varphi) \\ \sin(\varphi) & -\cos(\varphi) \end{pmatrix}$$

Mirrored orthogonal transformations in $I\!R^3$ are rotational mirrorings: rotations of axis $< \vec{a}_1 >$ through angle $\varphi$ and mirror plane $< \vec{a}_1 >^{\perp}$. The matrix of such a transformation is given by:

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & \cos(\varphi) & -\sin(\varphi) \\ 0 & \sin(\varphi) & \cos(\varphi) \end{pmatrix}$$

For all orthogonal transformations $O$ in $I\!R^3$ holds that $O(\vec{x}\,) \times O(\vec{y}\,) = O(\vec{x} \times \vec{y}\,)$.

$I\!R^n$ $(n < \infty)$ can be decomposed in invariant subspaces with dimension 1 or 2 for each orthogonal transformation.

### Unitary transformations

Let $V$ be a complex space on which an inner product is defined. Than a linear transformation $U$ is *unitary* if $U$ is isometric *and* its inverse transformation $A^{\leftarrow}$ exists. A $n \times n$ matrix is unitary if $U^H U = I\!I$. It has determinant $|\det(U)| = 1$. Each isometric transformation in a finite-dimensional complex vector space is unitary.

**Theorem**: for a $n \times n$ matrix $A$ the following statements are equivalent:

1. $A$ is unitary,
2. The columns of $A$ are an orthonormal set,
3. The rows of $A$ are an orthonormal set.

### Symmetric transformations

A transformation $A$ on $I\!R^n$ is *symmetric* if $(A\vec{x}, \vec{y}\,) = (\vec{x}, A\vec{y}\,)$. A matrix $A \in I\!M^{n \times n}$ is symmetric if $A = A^T$. A linear operator is only symmetric if its matrix w.r.t. an arbitrary basis is symmetric. All eigenvalues of a symmetric transformation belong to $I\!R$. The different eigenvectors are mutually perpendicular. If $A$ is symmetric, than $A^T = A = A^H$ on an orthogonal basis.

For each matrix $B \in I\!M^{m \times n}$ holds: $B^T B$ is symmetric.

**Hermitian transformations**

A transformation $H : V \to V$ with $V = \mathbb{C}^n$ is *Hermitian* if $(H\vec{x}, \vec{y}) = (\vec{x}, H\vec{y})$. The *Hermitian conjugated* transformation $A^H$ of $A$ is: $[a_{ij}]^H = [a^*_{ji}]$. An alternative notation is: $A^H = A^\dagger$. The inner product of two vectors $\vec{x}$ and $\vec{y}$ can now be written in the form: $(\vec{x}, \vec{y}) = \vec{x}^H\vec{y}$.

If the transformations $A$ and $B$ are Hermitian, than their product $AB$ is Hermitian if:
$[A, B] = AB - BA = 0$. $[A, B]$ is called the *commutator* of $A$ and $B$.

The eigenvalues of a Hermitian transformation belong to $\mathbb{R}$.

A matrix representation can be coupled with a Hermitian operator $L$. W.r.t. a basis $\vec{e}_i$ it is given by $L_{mn} = (\vec{e}_m, L\vec{e}_n)$.

**Normal transformations**

For each linear transformation $A$ in a complex vector space $V$ there exists exactly one linear transformation $B$ so that $(A\vec{x}, \vec{y}) = (\vec{x}, B\vec{y})$. This $B$ is called the *adjungated transformation* of $A$. Notation: $B = A^*$. The following holds: $(CD)^* = D^*C^*$. $A^* = A^{-1}$ if $A$ is unitary and $A^* = A$ if $A$ is Hermitian.

**Definition**: the linear transformation $A$ is *normal* in a complex vector space $V$ if $A^*A = AA^*$. This is only the case if for its matrix $S$ w.r.t. an orthonormal basis holds: $A^\dagger A = AA^\dagger$.

If $A$ is normal holds:

1. For all vectors $\vec{x} \in V$ and a normal transformation $A$ holds:

$$(A\vec{x}, A\vec{y}) = (A^*A\vec{x}, \vec{y}) = (AA^*\vec{x}, \vec{y}) = (A^*\vec{x}, A^*\vec{y})$$

2. $\vec{x}$ is an eigenvector of $A$ if and only if $\vec{x}$ is an eigenvector of $A^*$.

3. Eigenvectors of $A$ for different eigenvalues are mutually perpendicular.

4. If $E_\lambda$ if an eigenspace from $A$ than the orthogonal complement $E^\perp_\lambda$ is an invariant subspace of $A$.

Let the different roots of the characteristic equation of $A$ be $\beta_i$ with multiplicities $n_i$. Than the dimension of each eigenspace $V_i$ equals $n_i$. These eigenspaces are mutually perpendicular and each vector $\vec{x} \in V$ can be written in exactly one way as

$$\vec{x} = \sum_i \vec{x}_i \quad \text{with} \quad \vec{x}_i \in V_i$$

This can also be written as: $\vec{x}_i = P_i\vec{x}$ where $P_i$ is a projection on $V_i$. This leads to the *spectral mapping theorem*: let $A$ be a normal transformation in a complex vector space $V$ with $\dim(V) = n$. Than:

1. There exist projection transformations $P_i$, $1 \leq i \leq p$, with the properties

   - $P_i \cdot P_j = 0$ for $i \neq j$,
   - $P_1 + ... + P_p = \mathbb{I}$,
   - $\dim P_1(V) + ... + \dim P_p(V) = n$

   and complex numbers $\alpha_1, ..., \alpha_p$ so that $A = \alpha_1 P_1 + ... + \alpha_p P_p$.

2. If $A$ is unitary than holds $|\alpha_i| = 1\ \forall i$.

3. If $A$ is Hermitian than $\alpha_i \in \mathbb{R}\ \forall i$.

**Complete systems of commuting Hermitian transformations**

Consider $m$ Hermitian linear transformations $A_i$ in a $n$ dimensional complex inner product space $V$. Assume they mutually commute.

**Lemma**: if $E_\lambda$ is the eigenspace for eigenvalue $\lambda$ from $A_1$, than $E_\lambda$ is an invariant subspace of all transformations $A_i$. This means that if $\vec{x} \in E_\lambda$, than $A_i\vec{x} \in E_\lambda$.

**Theorem**. Consider $m$ commuting Hermitian matrices $A_i$. Than there exists a unitary matrix $U$ so that all matrices $U^\dagger A_i U$ are diagonal. The columns of $U$ are the common eigenvectors of all matrices $A_j$.

If all eigenvalues of a Hermitian linear transformation in a $n$-dimensional complex vector space differ, than the normalized eigenvector is known except for a phase factor $\exp(i\alpha)$.

**Definition**: a commuting set Hermitian transformations is called *complete* if for each set of two common eigenvectors $\vec{v}_i, \vec{v}_j$ there exists a transformation $A_k$ so that $\vec{v}_i$ and $\vec{v}_j$ are eigenvectors with different eigenvalues of $A_k$.

Usually a commuting set is taken as small as possible. In quantum physics one speaks of commuting observables. The required number of commuting observables equals the number of quantum numbers required to characterize a state.

## 5.9 Homogeneous coordinates

Homogeneous coordinates are used if one wants to combine both rotations and translations in *one* matrix transformation. An extra coordinate is introduced to describe the non-linearities. Homogeneous coordinates are derived from cartesian coordinates as follows:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix}_{\text{cart}} = \begin{pmatrix} wx \\ wy \\ wz \\ w \end{pmatrix}_{\text{hom}} = \begin{pmatrix} X \\ Y \\ Z \\ w \end{pmatrix}_{\text{hom}}$$

so $x = X/w$, $y = Y/w$ and $z = Z/w$. Transformations in homogeneous coordinates are described by the following matrices:

1. Translation along vector $(X_0, Y_0, Z_0, w_0)$:
$$T = \begin{pmatrix} w_0 & 0 & 0 & X_0 \\ 0 & w_0 & 0 & Y_0 \\ 0 & 0 & w_0 & Z_0 \\ 0 & 0 & 0 & w_0 \end{pmatrix}$$
2. Rotations of the $x, y, z$ axis, resp. through angles $\alpha, \beta, \gamma$:
$$R_x(\alpha) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & \cos\alpha & -\sin\alpha & 0 \\ 0 & \sin\alpha & \cos\alpha & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad R_y(\beta) = \begin{pmatrix} \cos\beta & 0 & \sin\beta & 0 \\ 0 & 1 & 0 & 0 \\ -\sin\beta & 0 & \cos\beta & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$
$$R_z(\gamma) = \begin{pmatrix} \cos\gamma & -\sin\gamma & 0 & 0 \\ \sin\gamma & \cos\gamma & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$
3. A perspective projection on image plane $z = c$ with the center of projection in the origin. This transformation has no inverse.
$$P(z = c) = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1/c & 0 \end{pmatrix}$$

## 5.10 Inner product spaces

A complex inner product on a complex vector space is defined as follows:

1. $(\vec{a},\vec{b}\,) = \overline{(\vec{b},\vec{a}\,)}$,
2. $(\vec{a},\beta_1\vec{b}_1+\beta_2\vec{b}_2) = \beta_1(\vec{a},\vec{b}_1)+\beta_2(\vec{a},\vec{b}_2)$ for all $\vec{a},\vec{b}_1,\vec{b}_2 \in V$ and $\beta_1,\beta_2 \in \mathbb{C}$.
3. $(\vec{a},\vec{a}\,) \geq 0$ for all $\vec{a} \in V$, $(\vec{a},\vec{a}\,) = 0$ if and only if $\vec{a} = \vec{0}$.

Due to (1) holds: $(\vec{a},\vec{a}\,) \in \mathbb{R}$. The *inner product space* $\mathbb{C}^n$ is the complex vector space on which a complex inner product is defined by:

$$(\vec{a},\vec{b}\,) = \sum_{i=1}^{n} a_i^* b_i$$

For function spaces holds:

$$(f,g) = \int_a^b f^*(t)g(t)dt$$

For each $\vec{a}$ the length $\|\vec{a}\,\|$ is defined by: $\|\vec{a}\,\| = \sqrt{(\vec{a},\vec{a}\,)}$. The following holds: $\|\vec{a}\,\| - \|\vec{b}\,\| \leq \|\vec{a}+\vec{b}\,\| \leq \|\vec{a}\,\| + \|\vec{b}\,\|$, and with $\varphi$ the angle between $\vec{a}$ and $\vec{b}$ holds: $(\vec{a},\vec{b}\,) = \|\vec{a}\,\| \cdot \|\vec{b}\,\| \cos(\varphi)$.

Let $\{\vec{a}_1,...,\vec{a}_n\}$ be a set of vectors in an inner product space $V$. Than the *Gramian G* of this set is given by: $G_{ij} = (\vec{a}_i,\vec{a}_j)$. The set of vectors is independent if and only if $\det(G) = 0$.

A set is *orthonormal* if $(\vec{a}_i,\vec{a}_j) = \delta_{ij}$. If $\vec{e}_1,\vec{e}_2,...$ form an orthonormal row in an infinite dimensional vector space Bessel's inequality holds:

$$\|\vec{x}\,\|^2 \geq \sum_{i=1}^{\infty} |(\vec{e}_i,\vec{x}\,)|^2$$

The equal sign holds if and only if $\lim_{n\to\infty} \|\vec{x}_n - \vec{x}\,\| = 0$.

The inner product space $\ell^2$ is defined in $\mathbb{C}^\infty$ by:

$$\ell^2 = \left\{ \vec{a} = (a_1,a_2,...) \mid \sum_{n=1}^{\infty} |a_n|^2 < \infty \right\}$$

A space is called a *Hilbert space* if it is $\ell^2$ and if also holds: $\lim_{n\to\infty} |a_{n+1} - a_n| = 0$.

## 5.11 The Laplace transformation

The class LT exists of functions for which holds:

1. On each interval $[0,A]$, $A > 0$ there are no more than a finite number of discontinuities and each discontinuity has an upper - and lower limit,
2. $\exists t_0 \in [0,\infty >$ and $a, M \in \mathbb{R}$ so that for $t \geq t_0$ holds: $|f(t)|\exp(-at) < M$.

Than there exists a Laplace transform for $f$.

The Laplace transformation is a generalisation of the Fourier transformation. The Laplace transform of a function $f(t)$ is, with $s \in \mathbb{C}$ and $t \geq 0$:

$$F(s) = \int_0^\infty f(t)\mathrm{e}^{-st}dt$$

The Laplace transform of the derivative of a function is given by:

$$\mathcal{L}\left(f^{(n)}(t)\right) = -f^{(n-1)}(0) - sf^{(n-2)}(0) - ... - s^{n-1}f(0) + s^nF(s)$$

The operator $\mathcal{L}$ has the following properties:

1. Equal shapes: if $a > 0$ than
$$\mathcal{L}\left(f(at)\right) = \frac{1}{a}F\left(\frac{s}{a}\right)$$
2. Damping: $\mathcal{L}\left(\mathrm{e}^{-at}f(t)\right) = F(s+a)$
3. Translation: If $a > 0$ and $g$ is defined by $g(t) = f(t-a)$ if $t > a$ and $g(t) = 0$ for $t \leq a$, than holds: $\mathcal{L}\left(g(t)\right) = \mathrm{e}^{-sa}\mathcal{L}(f(t))$.

If $s \in I\!R$ than holds $\Re(\lambda f) = \mathcal{L}(\Re(f))$ and $\Im(\lambda f) = \mathcal{L}(\Im(f))$.

For some often occurring functions holds:

| $f(t) =$ | $F(s) = \mathcal{L}(f(t)) =$ |
|---|---|
| $\frac{t^n}{n!}\mathrm{e}^{at}$ | $(s-a)^{-n-1}$ |
| $\mathrm{e}^{at}\cos(\omega t)$ | $\frac{s-a}{(s-a)^2+\omega^2}$ |
| $\mathrm{e}^{at}\sin(\omega t)$ | $\frac{\omega}{(s-a)^2+\omega^2}$ |
| $\delta(t-a)$ | $\exp(-as)$ |

## 5.12 The convolution

The convolution integral is defined by:

$$(f * g)(t) = \int_0^t f(u)g(t-u)du$$

The convolution has the following properties:

1. $f * g \in$LT
2. $\mathcal{L}(f * g) = \mathcal{L}(f) \cdot \mathcal{L}(g)$
3. Distribution: $f * (g + h) = f * g + f * h$
4. Commutative: $f * g = g * f$
5. Homogenity: $f * (\lambda g) = \lambda f * g$

If $\mathcal{L}(f) = F_1 \cdot F_2$, than is $f(t) = f_1 * f_2$.

## 5.13 Systems of linear differential equations

We start with the equation $\dot{\vec{x}} = A\vec{x}$. Assume that $\vec{x} = \vec{v}\exp(\lambda t)$, than follows: $A\vec{v} = \lambda\vec{v}$. In the $2 \times 2$ case holds:

1. $\lambda_1 = \lambda_2$: than $\vec{x}(t) = \sum \vec{v}_i \exp(\lambda_i t)$.
2. $\lambda_1 \neq \lambda_2$: than $\vec{x}(t) = (\vec{u}t + \vec{v})\exp(\lambda t)$.

Assume that $\lambda = \alpha + i\beta$ is an eigenvalue with eigenvector $\vec{v}$, than $\lambda^*$ is also an eigenvalue for eigenvector $\vec{v}^*$. Decompose $\vec{v} = \vec{u} + i\vec{w}$, than the real solutions are

$$c_1[\vec{u}\cos(\beta t) - \vec{w}\sin(\beta t)]\mathrm{e}^{\alpha t} + c_2[\vec{v}\cos(\beta t) + \vec{u}\sin(\beta t)]\mathrm{e}^{\alpha t}$$

There are two solution strategies for the equation $\ddot{\vec{x}} = A\vec{x}$:

1. Let $\vec{x} = \vec{v}\exp(\lambda t) \Rightarrow \det(A - \lambda^2 I\!I) = 0$.
2. Introduce: $\dot{x} = u$ and $\dot{y} = v$, this leads to $\ddot{x} = \dot{u}$ and $\ddot{y} = \dot{v}$. This transforms a $n$-dimensional set of second order equations into a $2n$-dimensional set of first order equations.

## 5.14 Quadratic forms

### 5.14.1 Quadratic forms in $I\!R^2$

The general equation of a quadratic form is: $\vec{x}^T A\vec{x} + 2\vec{x}^T P + S = 0$. Here, $A$ is a symmetric matrix. If $\Lambda = S^{-1}AS = \mathrm{diag}(\lambda_1, ..., \lambda_n)$ holds: $\vec{u}^T\Lambda\vec{u} + 2\vec{u}^T P + S = 0$, so all cross terms are 0. $\vec{u} = (u, v, w)$ should be chosen so that $\det(S) = +1$, to maintain the same orientation as the system $(x, y, z)$.

Starting with the equation

$$ax^2 + 2bxy + cy^2 + dx + ey + f = 0$$

we have $|A| = ac - b^2$. An ellipse has $|A| > 0$, a parabola $|A| = 0$ and a hyperbole $|A| < 0$. In polar coordinates this can be written as:

$$r = \frac{ep}{1 - e\cos(\theta)}$$

An ellipse has $e < 1$, a parabola $e = 1$ and a hyperbola $e > 1$.

### 5.14.2 Quadratic surfaces in $I\!R^3$

Rank 3:

$$p\frac{x^2}{a^2} + q\frac{y^2}{b^2} + r\frac{z^2}{c^2} = d$$

- Ellipsoid: $p = q = r = d = 1$, $a, b, c$ are the lengths of the semi axes.
- Single-bladed hyperboloid: $p = q = d = 1$, $r = -1$.
- Double-bladed hyperboloid: $r = d = 1$, $p = q = -1$.
- Cone: $p = q = 1$, $r = -1$, $d = 0$.

Rank 2:

$$p\frac{x^2}{a^2} + q\frac{y^2}{b^2} + r\frac{z}{c^2} = d$$

- Elliptic paraboloid: $p = q = 1$, $r = -1$, $d = 0$.
- Hyperbolic paraboloid: $p = r = -1$, $q = 1$, $d = 0$.
- Elliptic cylinder: $p = q = -1$, $r = d = 0$.
- Hyperbolic cylinder: $p = d = 1$, $q = -1$, $r = 0$.
- Pair of planes: $p = 1$, $q = -1$, $d = 0$.

Rank 1:

$$py^2 + qx = d$$

- Parabolic cylinder: $p, q > 0$.
- Parallel pair of planes: $d > 0$, $q = 0$, $p \neq 0$.
- Double plane: $p \neq 0$, $q = d = 0$.

# Chapter 6

# Complex function theory

## 6.1 Functions of complex variables

Complex function theory deals with complex functions of a complex variable. Some definitions:

$f$ is *analytical* on $\mathcal{G}$ if $f$ is continuous and differentiable on $\mathcal{G}$.

A *Jordan curve* is a curve that is closed and singular.

If K is a curve in $\mathbb{C}$ with parameter equation $z = \phi(t) = x(t) + iy(t)$, $a \leq t \leq b$, than the length $L$ of K is given by:

$$L = \int_a^b \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2}\, dt = \int_a^b \left|\frac{dz}{dt}\right| dt = \int_a^b |\phi'(t)| dt$$

The derivative of $f$ in point $z = a$ is:

$$f'(a) = \lim_{z \to a} \frac{f(z) - f(a)}{z - a}$$

If $f(z) = u(x, y) + iv(x, y)$ the derivative is:

$$f'(z) = \frac{\partial u}{\partial x} + i\frac{\partial v}{\partial x} = -i\frac{\partial u}{\partial y} + \frac{\partial v}{\partial y}$$

Setting both results equal yields the equations of Cauchy-Riemann:

$$\frac{\partial u}{\partial x} = \frac{\partial v}{\partial y} \quad , \quad \frac{\partial u}{\partial y} = -\frac{\partial v}{\partial x}$$

These equations imply that $\nabla^2 u = \nabla^2 v = 0$. $f$ is analytical if $u$ and $v$ satisfy these equations.

## 6.2 Complex integration

### 6.2.1 Cauchy's integral formula

Let $K$ be a curve described by $z = \phi(t)$ on $a \leq t \leq b$ and $f(z)$ is continuous on $K$. Than the integral of $f$ over $K$ is:

$$\int_K f(z)dz = \int_a^b f(\phi(t))\dot{\phi}(t)dt \overset{f\text{continuous}}{=} F(b) - F(a)$$

**Lemma**: let $K$ be the circle with center $a$ and radius $r$ taken in a positive direction. Than holds for integer $m$:

$$\frac{1}{2\pi i} \oint_K \frac{dz}{(z-a)^m} = \begin{cases} 0 \text{ if } m \neq 1 \\ 1 \text{ if } m = 1 \end{cases}$$

**Theorem**: if $L$ is the length of curve $K$ and if $|f(z)| \leq M$ for $z \in K$, than, if the integral exists, holds:

$$\left|\int_K f(z)dz\right| \leq ML$$

**Theorem**: let $f$ be continuous on an area $G$ and let $p$ be a fixed point of $G$. Let $F(z) = \int_p^z f(\xi)d\xi$ for all $z \in G$ only depend on $z$ and not on the integration path. Than $F(z)$ is analytical on $G$ with $F'(z) = f(z)$.

This leads to two equivalent formulations of the *main theorem of complex integration*: let the function $f$ be analytical on an area $G$. Let $K$ and $K'$ be two curves with the same starting - and end points, which can be transformed into each other by continous deformation within $G$. Let $B$ be a Jordan curve. Than holds

$$\int_K f(z)dz = \int_{K'} f(z)dz \Leftrightarrow \oint_B f(z)dz = 0$$

By applying the main theorem on $e^{iz}/z$ one can derive that

$$\int_0^\infty \frac{\sin(x)}{x}dx = \frac{\pi}{2}$$

### 6.2.2 Residue

A point $a \in \mathbb{C}$ is a *regular point* of a function $f(z)$ if $f$ is analytical in $a$. Otherwise $a$ is a *singular point* or *pole* of $f(z)$. The *residue* of $f$ in $a$ is defined by

$$\operatorname*{Res}_{z=a} f(z) = \frac{1}{2\pi i}\oint_K f(z)dz$$

where $K$ is a Jordan curve which encloses $a$ in positive direction. The residue is 0 in regular points, in singular points it can be both 0 and $\neq 0$. Cauchy's residue proposition is: let $f$ be analytical within and on a Jordan curve $K$ except in a finite number of singular points $a_i$ within $K$. Than, if $K$ is taken in a positive direction, holds:

$$\frac{1}{2\pi i}\oint_K f(z)dz = \sum_{k=1}^{n} \operatorname*{Res}_{z=a_k} f(z)$$

**Lemma**: let the function $f$ be analytical in $a$, than holds:

$$\operatorname*{Res}_{z=a} \frac{f(z)}{z-a} = f(a)$$

This leads to Cauchy's integral theorem: if $F$ is analytical on the Jordan curve $K$, which is taken in a positive direction, holds:

$$\frac{1}{2\pi i}\oint_K \frac{f(z)}{z-a}dz = \begin{cases} f(a) \text{ if } a \text{ inside } K \\ 0 \text{ if } a \text{ outside } K \end{cases}$$

**Theorem**: let $K$ be a curve ($K$ need not be closed) and let $\phi(\xi)$ be continuous on $K$. Than the function

$$f(z) = \int_K \frac{\phi(\xi)d\xi}{\xi - z}$$

is analytical with $n$-th derivative

$$f^{(n)}(z) = n!\int_K \frac{\phi(\xi)d\xi}{(\xi - z)^{n+1}}$$

**Theorem**: let $K$ be a curve and $G$ an area. Let $\phi(\xi, z)$ be defined for $\xi \in K$, $z \in G$, with the following properties:

1. $\phi(\xi, z)$ is limited, this means $|\phi(\xi, z)| \leq M$ for $\xi \in K$, $z \in G$,
2. For fixed $\xi \in K$, $\phi(\xi, z)$ is an analytical function of $z$ on $G$,

3. For fixed $z \in G$ the functions $\phi(\xi, z)$ and $\partial\phi(\xi, z)/\partial z$ are continuous functions of $\xi$ on $K$.

Than the function

$$f(z) = \int_K \phi(\xi, z) d\xi$$

is analytical with derivative

$$f'(z) = \int_K \frac{\partial \phi(\xi, z)}{\partial z} d\xi$$

**Cauchy's inequality**: let $f(z)$ be an analytical function within and on the circle $C : |z-a| = R$ and let $|f(z)| \leq M$ for $z \in C$. Than holds

$$\left| f^{(n)}(a) \right| \leq \frac{Mn!}{R^n}$$

## 6.3 Analytical functions definied by series

The series $\sum f_n(z)$ is called *pointwise convergent* on an area $G$ with sum $F(z)$ if

$$\forall_{\varepsilon>0} \forall_{z\in G} \exists_{N_0 \in \mathbb{R}} \forall_{n>n_0} \left[ \left| f(z) - \sum_{n=1}^{N} f_n(z) \right| < \varepsilon \right]$$

The series is called *uniform convergent* if

$$\forall_{\varepsilon>0} \exists_{N_0 \in \mathbb{R}} \forall_{n>n_0} \exists_{z\in G} \left[ \left| f(z) - \sum_{n=1}^{N} f_n(z) \right| < \varepsilon \right]$$

Uniform convergence implies pointwise convergence, the opposite is not necessary.

**Theorem**: let the power series $\sum_{n=0}^{\infty} a_n z^n$ have a radius of convergence $R$. $R$ is the distance to the first non-essential singularity.

- If $\lim_{n\to\infty} \sqrt[n]{|a_n|} = L$ exists, than $R = 1/L$.
- If $\lim_{n\to\infty} |a_{n+1}|/|a_n| = L$ exists, than $R = 1/L$.

If these limits both don't exist one can find $R$ with the formula of Cauchy-Hadamard:

$$\frac{1}{R} = \lim_{n\to\infty} \sup \sqrt[n]{|a_n|}$$

## 6.4 Laurent series

**Taylor's theorem**: let $f$ be analytical in an area $G$ and let point $a \in G$ has distance $r$ to the boundary of $G$. Than $f(z)$ can be expanded into the Taylor series near $a$:

$$f(z) = \sum_{n=0}^{\infty} c_n (z-a)^n \quad \text{with} \quad c_n = \frac{f^{(n)}(a)}{n!}$$

valid for $|z-a| < r$. The radius of convergence of the Taylor series is $\geq r$. If $f$ has a pole of order $k$ in $a$ than $c_1, ..., c_{k-1} = 0$, $c_k \neq 0$.

**Theorem of Laurent**: let $f$ be analytical in the circular area $G : r < |z-a| < R$. Than $f(z)$ can be expanded into a Laurent series with center $a$:

$$f(z) = \sum_{n=-\infty}^{\infty} c_n (z-a)^n \quad \text{with} \quad c_n = \frac{1}{2\pi i} \oint_K \frac{f(w)dw}{(w-a)^{n+1}} \quad , \quad n \in \mathbb{Z}$$

valid for $r < |z-a| < R$ and $K$ an arbitrary Jordan curve in $G$ which encloses point $a$ in positive direction.

The *principal part* of a Laurent series is: $\sum_{n=1}^{\infty} c_{-n}(z-a)^{-n}$. One can classify singular points with this. There are 3 cases:

1. There is no principal part. Than $a$ is a non-essential singularity. Define $f(a) = c_0$ and the series is also valid for $|z-a| < R$ and $f$ is analytical in $a$.
2. The principal part contains a finite number of terms. Than there exists a $k \in I\!N$ so that $\lim_{z\to a}(z-a)^k f(z) = c_{-k} \neq 0$. Than the function $g(z) = (z-a)^k f(z)$ has a non-essential singularity in $a$. One speaks of a pole of order $k$ in $z = a$.
3. The principal part contains an infinite number of terms. Then, $a$ is an essential singular point of $f$, such as $\exp(1/z)$ for $z = 0$.

If $f$ and $g$ are analytical, $f(a) \neq 0$, $g(a) = 0$, $g'(a) \neq 0$ than $f(z)/g(z)$ has a simple pole (i.e. a pole of order 1) in $z = a$ with

$$\operatorname*{Res}_{z=a} \frac{f(z)}{g(z)} = \frac{f(a)}{g'(a)}$$

## 6.5 Jordan's theorem

Residues are often used when solving definite integrals. We define the notations $C_\rho^+ = \{z||z| = \rho, \Im(z) \geq 0\}$ and $C_\rho^- = \{z||z| = \rho, \Im(z) \leq 0\}$ and $M^+(\rho, f) = \max_{z\in C_\rho^+} |f(z)|$, $M^-(\rho, f) = \max_{z\in C_\rho^-} |f(z)|$. We assume that $f(z)$ is analytical for $\Im(z) > 0$ with a possible exception of a finite number of singular points which do not lie on the real axis, $\lim_{\rho\to\infty} \rho M^+(\rho, f) = 0$ and that the integral exists, than

$$\int_{-\infty}^{\infty} f(x)dx = 2\pi i \sum \operatorname{Res} f(z) \quad \text{in} \quad \Im(z) > 0$$

Replace $M^+$ by $M^-$ in the conditions above and it follows that:

$$\int_{-\infty}^{\infty} f(x)dx = -2\pi i \sum \operatorname{Res} f(z) \quad \text{in} \quad \Im(z) < 0$$

*Jordan's lemma*: let $f$ be continuous for $|z| \geq R$, $\Im(z) \geq 0$ and $\lim_{\rho\to\infty} M^+(\rho, f) = 0$. Than holds for $\alpha > 0$

$$\lim_{\rho\to\infty} \int_{C_\rho^+} f(z)\mathrm{e}^{i\alpha z}dz = 0$$

Let $f$ be continuous for $|z| \geq R$, $\Im(z) \leq 0$ and $\lim_{\rho\to\infty} M^-(\rho, f) = 0$. Than holds for $\alpha < 0$

$$\lim_{\rho\to\infty} \int_{C_\rho^-} f(z)\mathrm{e}^{i\alpha z}dz = 0$$

Let $z = a$ be a simple pole of $f(z)$ and let $C_\delta$ be the half circle $|z-a| = \delta, 0 \leq \arg(z-a) \leq \pi$, taken from $a+\delta$ to $a-\delta$. Than is

$$\lim_{\delta\downarrow 0} \frac{1}{2\pi i} \int_{C_\delta} f(z)dz = \tfrac{1}{2} \operatorname*{Res}_{z=a} f(z)$$

# Chapter 7

# Tensor calculus

## 7.1 Vectors and covectors

A finite dimensional vector space is denoted by $\mathcal{V}, \mathcal{W}$. The vector space of linear transformations from $\mathcal{V}$ to $\mathcal{W}$ is denoted by $\mathcal{L}(\mathcal{V}, \mathcal{W})$. Consider $\mathcal{L}(\mathcal{V},\mathbb{R}) := \mathcal{V}^*$. We name $\mathcal{V}^*$ the *dual space* of $\mathcal{V}$. Now we can define *vectors* in $\mathcal{V}$ with basis $\vec{c}$ and *covectors* in $\mathcal{V}^*$ with basis $\hat{\vec{c}}$. Properties of both are:

1. Vectors: $\vec{x} = x^i \vec{c}_i$ with basis vectors $\vec{c}_i$:

$$\vec{c}_i = \frac{\partial}{\partial x^i}$$

   Transformation from system $i$ to $i'$ is given by:

$$\vec{c}_{i'} = A^i_{i'} \vec{c}_i = \partial_i \in \mathcal{V} \ , \ x^{i'} = A^{i'}_i x^i$$

2. Covectors: $\hat{\vec{x}} = x_i \hat{\vec{c}}^{\,i}$ with basis vectors $\hat{\vec{c}}^{\,i}$

$$\hat{\vec{c}}^{\,i} = dx^i$$

   Transformation from system $i$ to $i'$ is given by:

$$\hat{\vec{c}}^{\,i'} = A^{i'}_i \hat{\vec{c}}^{\,i} \in \mathcal{V}^* \ , \ \vec{x}_{i'} = A^i_{i'} \vec{x}_i$$

Here the *Einstein convention* is used:

$$a^i b_i := \sum_i a^i b_i$$

The coordinate transformation is given by:

$$A^i_{i'} = \frac{\partial x^i}{\partial x^{i'}} \ , \ A^{i'}_i = \frac{\partial x^{i'}}{\partial x^i}$$

From this follows that $A^i_k \cdot A^k_l = \delta^k_l$ and $A^i_{i'} = (A^{i'}_i)^{-1}$.

In differential notation the coordinate transformations are given by:

$$dx^i = \frac{\partial x^i}{\partial x^{i'}} dx^{i'} \quad \text{and} \quad \frac{\partial}{\partial x^{i'}} = \frac{\partial x^i}{\partial x^{i'}} \frac{\partial}{\partial x^i}$$

The general transformation rule for a tensor $T$ is:

$$T^{q_1 \dots q_n}_{s_1 \dots s_m} = \left| \frac{\partial \vec{x}}{\partial \vec{u}} \right|^\ell \frac{\partial u^{q_1}}{\partial x^{p_1}} \cdots \frac{\partial u^{q_n}}{\partial x^{p_n}} \cdot \frac{\partial x^{r_1}}{\partial u^{s_1}} \cdots \frac{\partial x^{r_m}}{\partial u^{s_m}} T^{p_1 \dots p_n}_{r_1 \dots r_m}$$

For an *absolute tensor* $\ell = 0$.

## 7.2 Tensor algebra

The following holds:

$$a_{ij}(x_i + y_i) \equiv a_{ij}x_i + a_{ij}y_i, \quad \text{but: } a_{ij}(x_i + y_j) \not\equiv a_{ij}x_i + a_{ij}y_j$$

and

$$(a_{ij} + a_{ji})x_i x_j \equiv 2a_{ij}x_i x_j, \quad \text{but: } (a_{ij} + a_{ji})x_i y_j \not\equiv 2a_{ij}x_i y_j$$

en $(a_{ij} - a_{ji})x_i x_j \equiv 0$.

The sum and difference of two tensors is a tensor of the same rank: $A^p_q \pm B^p_q$. The *outer tensor product* results in a tensor with a rank equal to the sum of the ranks of both tensors: $A^{pr}_q \cdot B^m_s = C^{prm}_{qs}$. The *contraction* equals two indices and sums over them. Suppose we take $r = s$ for a tensor $A^{mpr}_{qs}$, this results in: $\sum_r A^{mpr}_{qr} = B^{mp}_q$. The *inner product* of two tensors is defined by taking the outer product followed by a contraction.

## 7.3 Inner product

**Definition**: the bilinear transformation $B : \mathcal{V} \times \mathcal{V}^* \to I\!R$, $B(\vec{x}, \hat{\vec{y}}) = \hat{\vec{y}}(\vec{x})$ is denoted by $< \vec{x}, \hat{\vec{y}} >$. For this *pairing operator* $< \cdot, \cdot >= \delta$ holds:

$$\hat{\vec{y}}(\vec{x}) =< \vec{x}, \hat{\vec{y}} >= y_i x^i \quad , \quad < \hat{\vec{c}}^{\,i}, \vec{c}_j >= \delta^i_j$$

Let $G : \mathcal{V} \to \mathcal{V}^*$ be a linear bijection. Define the bilinear forms

$$\begin{aligned} g : \mathcal{V} \times \mathcal{V} \to I\!R \qquad & g(\vec{x}, \vec{y}) =< \vec{x}, G\vec{y} > \\ h : \mathcal{V}^* \times \mathcal{V}^* \to I\!R \qquad & h(\hat{\vec{x}}, \hat{\vec{y}}) =< G^{-1}\hat{\vec{x}}, \hat{\vec{y}} > \end{aligned}$$

Both are not degenerated. The following holds: $h(G\vec{x}, G\vec{y}) =< \vec{x}, G\vec{y} >= g(\vec{x}, \vec{y})$. If we identify $\mathcal{V}$ and $\mathcal{V}^*$ with $G$, than $g$ (or $h$) gives an inner product on $\mathcal{V}$.

The inner product $(,)_\Lambda$ on $\Lambda^k(\mathcal{V})$ is defined by:

$$(\Phi, \Psi)_\Lambda = \frac{1}{k!}(\Phi, \Psi)_{T^0_k(\mathcal{V})}$$

The inner product of two vectors is than given by:

$$(\vec{x}, \vec{y}) = x^i y^i < \vec{c}_i, G\vec{c}_j >= g_{ij}x^i x^j$$

The matrix $g_{ij}$ of $G$ is given by

$$g_{ij}\hat{\vec{c}}^{\,j} = G\vec{c}_i$$

The matrix $g^{ij}$ of $G^{-1}$ is given by:

$$g^{kl}\vec{c}_l = G^{-1}\hat{\vec{c}}^{\,k}$$

For this *metric tensor* $g_{ij}$ holds: $g_{ij}g^{jk} = \delta^k_i$. This tensor can raise or lower indices:

$$x_j = g_{ij}x^i \quad , \quad x^i = g^{ij}x_j$$

and $du^i = \hat{\vec{c}}^{\,i} = g^{ij}\vec{c}_j$.

## 7.4 Tensor product

**Definition**: let $\mathcal{U}$ and $\mathcal{V}$ be two finite dimensional vector spaces with dimensions $m$ and $n$. Let $\mathcal{U}^* \times \mathcal{V}^*$ be the cartesian product of $\mathcal{U}$ and $\mathcal{V}$. A function $t : \mathcal{U}^* \times \mathcal{V}^* \to I\!R$; $(\hat{\vec{u}}; \hat{\vec{v}}) \mapsto t(\hat{\vec{u}}; \hat{\vec{v}}) = t^{\alpha\beta} u_\alpha u_\beta \in I\!R$ is called a tensor if $t$ is linear in $\hat{\vec{u}}$ and $\hat{\vec{v}}$. The tensors $t$ form a vector space denoted by $\mathcal{U} \otimes \mathcal{V}$. The elements $T \in \mathcal{V} \otimes \mathcal{V}$ are called contravariant 2-tensors: $T = T^{ij} \vec{c}_i \otimes \vec{c}_j = T^{ij} \partial_i \otimes \partial_j$. The elements $T \in \mathcal{V}^* \otimes \mathcal{V}^*$ are called covariant 2-tensors: $T = T_{ij} \hat{\vec{c}}^{\,i} \otimes \hat{\vec{c}}^{\,j} = T_{ij} dx^i \otimes dx^j$. The elements $T \in \mathcal{V}^* \otimes \mathcal{V}$ are called mixed 2 tensors: $T = T_i^{\cdot j} \hat{\vec{c}}^{\,i} \otimes \vec{c}_j = T_i^{\cdot j} dx^i \otimes \partial_j$, and analogous for $T \in \mathcal{V} \otimes \mathcal{V}^*$.

The numbers given by

$$t^{\alpha\beta} = t(\hat{\vec{c}}^{\,\alpha}, \hat{\vec{c}}^{\,\beta})$$

with $1 \leq \alpha \leq m$ and $1 \leq \beta \leq n$ are the components of $t$.

Take $\vec{x} \in \mathcal{U}$ and $\vec{y} \in \mathcal{V}$. Than the function $\vec{x} \otimes \vec{y}$, definied by

$$(\vec{x} \otimes \vec{y})(\hat{\vec{u}}, \hat{\vec{v}}) = < \vec{x}, \hat{\vec{u}} >_U < \vec{y}, \hat{\vec{v}} >_V$$

is a tensor. The components are derived from: $(\vec{u} \otimes \vec{v})_{ij} = u_i v^j$. The tensor product of 2 tensors is given by:

$$\binom{2}{0} \text{ form:} \qquad (\vec{v} \otimes \vec{w})(\hat{\vec{p}}, \hat{\vec{q}}) = v^i p_i w^k q_k = T^{ik} p_i q_k$$

$$\binom{0}{2} \text{ form:} \qquad (\hat{\vec{p}} \otimes \hat{\vec{q}})(\vec{v}, \vec{w}) = p_i v^i q_k w^k = T_{ik} v^i w^k$$

$$\binom{1}{1} \text{ form:} \qquad (\vec{v} \otimes \hat{\vec{p}})(\hat{\vec{q}}, \vec{w}) = v^i q_i p_k w^k = T_k^i q_i w^k$$

## 7.5 Symmetric and antisymmetric tensors

A tensor $t \in \mathcal{V} \otimes \mathcal{V}$ is called symmetric resp. antisymmetric if $\forall \hat{\vec{x}}, \hat{\vec{y}} \in \mathcal{V}^*$ holds: $t(\hat{\vec{x}}, \hat{\vec{y}}) = t(\hat{\vec{y}}, \hat{\vec{x}})$ resp. $t(\hat{\vec{x}}, \hat{\vec{y}}) = -t(\hat{\vec{y}}, \hat{\vec{x}})$.

A tensor $t \in \mathcal{V}^* \otimes \mathcal{V}^*$ is called symmetric resp. antisymmetric if $\forall \vec{x}, \vec{y} \in \mathcal{V}$ holds: $t(\vec{x}, \vec{y}) = t(\vec{y}, \vec{x})$ resp. $t(\vec{x}, \vec{y}) = -t(\vec{y}, \vec{x})$. The linear transformations $\mathcal{S}$ and $\mathcal{A}$ in $\mathcal{V} \otimes \mathcal{W}$ are defined by:

$$\begin{aligned} \mathcal{S}t(\hat{\vec{x}}, \hat{\vec{y}}) &= \tfrac{1}{2}(t(\hat{\vec{x}}, \hat{\vec{y}}) + t(\hat{\vec{y}}, \hat{\vec{x}})) \\ \mathcal{A}t(\hat{\vec{x}}, \hat{\vec{y}}) &= \tfrac{1}{2}(t(\hat{\vec{x}}, \hat{\vec{y}}) - t(\hat{\vec{y}}, \hat{\vec{x}})) \end{aligned}$$

Analogous in $\mathcal{V}^* \otimes \mathcal{V}^*$. If $t$ is symmetric resp. antisymmetric, than $\mathcal{S}t = t$ resp. $\mathcal{A}t = t$.

The tensors $\vec{e}_i \vee \vec{e}_j = \vec{e}_i \vec{e}_j = 2\mathcal{S}(\vec{e}_i \otimes \vec{e}_j)$, with $1 \leq i \leq j \leq n$ are a basis in $\mathcal{S}(\mathcal{V} \otimes \mathcal{V})$ with dimension $\frac{1}{2}n(n+1)$.

The tensors $\vec{e}_i \wedge \vec{e}_j = 2\mathcal{A}(\vec{e}_i \otimes \vec{e}_j)$, with $1 \leq i \leq j \leq n$ are a basis in $\mathcal{A}(\mathcal{V} \otimes \mathcal{V})$ with dimension $\frac{1}{2}n(n-1)$.

The complete antisymmetric tensor $\varepsilon$ is given by: $\varepsilon_{ijk}\varepsilon_{klm} = \delta_{il}\delta_{jm} - \delta_{im}\delta_{jl}$.

The permutation-operators $e_{pqr}$ are defined by: $e_{123} = e_{231} = e_{312} = 1$, $e_{213} = e_{132} = e_{321} = -1$, for all other combinations $e_{pqr} = 0$. There is a connection with the $\varepsilon$ tensor: $\varepsilon_{pqr} = g^{-1/2} e_{pqr}$ and $\varepsilon^{pqr} = g^{1/2} e^{pqr}$.

## 7.6 Outer product

Let $\alpha \in \Lambda^k(\mathcal{V})$ and $\beta \in \Lambda^l(\mathcal{V})$. Than $\alpha \wedge \beta \in \Lambda^{k+l}(\mathcal{V})$ is defined by:

$$\alpha \wedge \beta = \frac{(k+l)!}{k!l!} \mathcal{A}(\alpha \otimes \beta)$$

If $\alpha$ and $\beta \in \Lambda^1(\mathcal{V}) = \mathcal{V}^*$ holds: $\alpha \wedge \beta = \alpha \otimes \beta - \beta \otimes \alpha$

The outer product can be written as: $(\vec{a} \times \vec{b})_i = \varepsilon_{ijk} a^j b^k$, $\vec{a} \times \vec{b} = G^{-1} \cdot *(G\vec{a} \wedge G\vec{b})$.

Take $\vec{a}, \vec{b}, \vec{c}, \vec{d} \in I\!R^4$. Than $(dt \wedge dz)(\vec{a}, \vec{b}) = a_0 b_4 - b_0 a_4$ is the oriented surface of the projection on the $tz$-plane of the parallelogram spanned by $\vec{a}$ and $\vec{b}$.

Further

$$(dt \wedge dy \wedge dz)(\vec{a}, \vec{b}, \vec{c}) = \det \begin{vmatrix} a_0 & b_0 & c_0 \\ a_2 & b_2 & c_2 \\ a_4 & b_4 & c_4 \end{vmatrix}$$

is the oriented 3-dimensional volume of the projection on the $tyz$-plane of the parallelepiped spanned by $\vec{a}$, $\vec{b}$ and $\vec{c}$.

$(dt \wedge dx \wedge dy \wedge dz)(\vec{a}, \vec{b}, \vec{c}, \vec{d}) = \det(\vec{a}, \vec{b}, \vec{c}, \vec{d})$ is the 4-dimensional volume of the hyperparellelepiped spanned by $\vec{a}$, $\vec{b}$, $\vec{c}$ and $\vec{d}$.

## 7.7 The Hodge star operator

$\Lambda^k(\mathcal{V})$ and $\Lambda^{n-k}(\mathcal{V})$ have the same dimension because $\binom{n}{k} = \binom{n}{n-k}$ for $1 \leq k \leq n$. Dim$(\Lambda^n(\mathcal{V})) = 1$. The choice of a basis means the choice of an oriented measure of volume, a volume $\mu$, in $\mathcal{V}$. We can gauge $\mu$ so that for an orthonormal basis $\vec{e}_i$ holds: $\mu(\vec{e}_i) = 1$. This basis is than by definition positive oriented if $\mu = \hat{\vec{e}}^{\,1} \wedge \hat{\vec{e}}^{\,2} \wedge ... \wedge \hat{\vec{e}}^{\,n} = 1$.

Because both spaces have the same dimension one can ask if there exists a bijection between them. If $\mathcal{V}$ has no extra structure this is not the case. However, such an operation does exist if there is an inner product defined on $\mathcal{V}$ and the corresponding volume $\mu$. This is called the *Hodge star operator* and denoted by $*$. The following holds:

$$\forall_{w \in \Lambda^k(\mathcal{V})} \exists_{*w \in \Lambda^{k-n}(\mathcal{V})} \forall_{\theta \in \Lambda^k(\mathcal{V})} \;\; \theta \wedge *w = (\theta, w)_\lambda \mu$$

For an orthonormal basis in $I\!R^3$ holds: the volume: $\mu = dx \wedge dy \wedge dz$, $*dx \wedge dy \wedge dz = 1$, $*dx = dy \wedge dz$, $*dz = dx \wedge dy$, $*dy = -dx \wedge dz$, $*(dx \wedge dy) = dz$, $*(dy \wedge dz) = dx$, $*(dx \wedge dz) = -dy$.

For a Minkowski basis in $I\!R^4$ holds: $\mu = dt \wedge dx \wedge dy \wedge dz$, $G = dt \otimes dt - dx \otimes dx - dy \otimes dy - dz \otimes dz$, and $*dt \wedge dx \wedge dy \wedge dz = 1$ and $*1 = dt \wedge dx \wedge dy \wedge dz$. Further $*dt = dx \wedge dy \wedge dz$ and $*dx = dt \wedge dy \wedge dz$.

## 7.8 Differential operations

### 7.8.1 The directional derivative

The *directional derivative* in point $\vec{a}$ is given by:

$$\mathcal{L}_{\vec{a}} f = < \vec{a}, df > = a^i \frac{\partial f}{\partial x^i}$$

### 7.8.2 The Lie-derivative

The *Lie-derivative* is given by:

$$(\mathcal{L}_{\vec{v}} \vec{w})^j = w^i \partial_i v^j - v^i \partial_i w^j$$

### 7.8.3 Christoffel symbols

To each curvelinear coordinate system $u^i$ we add a system of $n^3$ functions $\Gamma^i_{jk}$ of $\vec{u}$, defined by

$$\frac{\partial^2 \vec{x}}{\partial u^i \partial u^k} = \Gamma^i_{jk} \frac{\partial \vec{x}}{\partial u^i}$$

These are *Christoffel symbols of the second kind*. Christoffel symbols are no tensors. The Christoffel symbols of the second kind are given by:

$$\left\{ \begin{matrix} i \\ jk \end{matrix} \right\} := \Gamma^i_{jk} = \left\langle \frac{\partial^2 \vec{x}}{\partial u^k \partial u^j}, dx^i \right\rangle$$

with $\Gamma^i_{jk} = \Gamma^i_{kj}$. Their transformation to a different coordinate system is given by:

$$\Gamma^{i'}_{j'k'} = A^i_{i'} A^j_{j'} A^k_{k'} \Gamma^i_{jk} + A^{i'}_i (\partial_{j'} A^i_{k'})$$

The first term in this expression is 0 if the primed coordinates are cartesian.

There is a relation between Christoffel symbols and the metric:

$$\Gamma^i_{jk} = \tfrac{1}{2} g^{ir} (\partial_j g_{kr} + \partial_k g_{rj} - \partial_r g_{jk})$$

and $\Gamma^\alpha_{\beta\alpha} = \partial_\beta (\ln(\sqrt{|g|}))$.

Lowering an index gives the *Christoffel symbols of the first kind*: $\Gamma^i_{jk} = g^{il} \Gamma_{jkl}$.

### 7.8.4 The covariant derivative

The *covariant derivative* $\nabla_j$ of a vector, covector and of rank-2 tensors is given by:

$$\begin{aligned}
\nabla_j a^i &= \partial_j a^i + \Gamma^i_{jk} a^k \\
\nabla_j a_i &= \partial_j a_i - \Gamma^k_{ij} a_k \\
\nabla_\gamma a^\alpha_\beta &= \partial_\gamma a^\alpha_\beta - \Gamma^\varepsilon_{\gamma\beta} a^\alpha_\varepsilon + \Gamma^\alpha_{\gamma\varepsilon} a^\varepsilon_\beta \\
\nabla_\gamma a_{\alpha\beta} &= \partial_\gamma a_{\alpha\beta} - \Gamma^\varepsilon_{\gamma\alpha} a_{\varepsilon\beta} - \Gamma^\varepsilon_{\gamma\beta} a_{\alpha\varepsilon} \\
\nabla_\gamma a^{\alpha\beta} &= \partial_\gamma a^{\alpha\beta} + \Gamma^\alpha_{\gamma\varepsilon} a^{\varepsilon\beta} + \Gamma^\beta_{\gamma\varepsilon} a^{\alpha\varepsilon}
\end{aligned}$$

Ricci's theorem:

$$\nabla_\gamma g_{\alpha\beta} = \nabla_\gamma g^{\alpha\beta} = 0$$

## 7.9 Differential operators

**The Gradient**

is given by:

$$\mathrm{grad}(f) = G^{-1} df = g^{ki} \frac{\partial f}{\partial x^i} \frac{\partial}{\partial x^k}$$

**The divergence**

is given by:

$$\mathrm{div}(a^i) = \nabla_i a^i = \frac{1}{\sqrt{g}} \partial_k (\sqrt{g}\, a^k)$$

**The curl**

is given by:

$$\mathrm{rot}(a) = G^{-1} \cdot * \cdot d \cdot G\vec{a} = -\varepsilon^{pqr} \nabla_q a_p = \nabla_q a_p - \nabla_p a_q$$

**The Laplacian**

is given by:

$$\Delta(f) = \mathrm{div}\ \mathrm{grad}(f) = *d * df = \nabla_i g^{ij} \partial_j f = g^{ij} \nabla_i \nabla_j f = \frac{1}{\sqrt{g}} \frac{\partial}{\partial x^i} \left( \sqrt{g}\, g^{ij} \frac{\partial f}{\partial x^j} \right)$$

## 7.10 Differential geometry

### 7.10.1 Space curves

We limit ourselves to $\mathbb{R}^3$ with a fixed orthonormal basis. A point is represented by the vector $\vec{x} = (x^1, x^2, x^3)$. A space curve is a collection of points represented by $\vec{x} = \vec{x}(t)$. The arc length of a space curve is given by:

$$s(t) = \int_{t_0}^{t} \sqrt{\left(\frac{dx}{d\tau}\right)^2 + \left(\frac{dy}{d\tau}\right)^2 + \left(\frac{dz}{d\tau}\right)^2}\, d\tau$$

The derivative of $s$ with respect to $t$ is the length of the vector $d\vec{x}/dt$:

$$\left(\frac{ds}{dt}\right)^2 = \left(\frac{d\vec{x}}{dt}, \frac{d\vec{x}}{dt}\right)$$

The *osculation plane* in a point $P$ of a space curve is the limiting position of the plane through the tangent of the plane in point $P$ and a point $Q$ when $Q$ approaches $P$ along the space curve. The osculation plane is parallel with $\dot{\vec{x}}(s)$. If $\ddot{\vec{x}} \neq 0$ the osculation plane is given by:

$$\vec{y} = \vec{x} + \lambda\dot{\vec{x}} + \mu\ddot{\vec{x}} \quad \text{so} \quad \det(\vec{y} - \vec{x}, \dot{\vec{x}}, \ddot{\vec{x}}\,) = 0$$

In a bending point holds, if $\dddot{\vec{x}} \neq 0$:

$$\vec{y} = \vec{x} + \lambda\dot{\vec{x}} + \mu\,\dddot{\vec{x}}$$

The *tangent* has unit vector $\vec{\ell} = \dot{\vec{x}}$, the *main normal* unit vector $\vec{n} = \ddot{\vec{x}}$ and the *binormal* $\vec{b} = \dot{\vec{x}} \times \ddot{\vec{x}}$. So the main normal lies in the osculation plane, the binormal is perpendicular to it.

Let $P$ be a point and $Q$ be a nearby point of a space curve $\vec{x}(s)$. Let $\Delta\varphi$ be the angle between the tangents in $P$ and $Q$ and let $\Delta\psi$ be the angle between the osculation planes (binormals) in $P$ and $Q$. Then the *curvature* $\rho$ and the *torsion* $\tau$ in $P$ are defined by:

$$\rho^2 = \left(\frac{d\varphi}{ds}\right)^2 = \lim_{\Delta s \to 0}\left(\frac{\Delta\varphi}{\Delta s}\right)^2 \quad , \quad \tau^2 = \left(\frac{d\psi}{ds}\right)^2$$

and $\rho > 0$. For plane curves $\rho$ is the ordinary curvature and $\tau = 0$. The following holds:

$$\rho^2 = (\vec{\ell}, \vec{\ell}) = (\ddot{\vec{x}}, \ddot{\vec{x}}\,) \quad \text{and} \quad \tau^2 = (\dot{\vec{b}}, \dot{\vec{b}})$$

Frenet's equations express the derivatives as linear combinations of these vectors:

$$\dot{\vec{\ell}} = \rho\vec{n} \;\; , \;\; \dot{\vec{n}} = -\rho\vec{\ell} + \tau\vec{b} \;\; , \;\; \dot{\vec{b}} = -\tau\vec{n}$$

From this follows that $\det(\dot{\vec{x}}, \ddot{\vec{x}}, \dddot{\vec{x}}\,) = \rho^2\tau$.

Some curves and their properties are:

| | |
|---|---|
| Screw line | $\tau/\rho$ =constant |
| Circle screw line | $\tau$ =constant, $\rho$ =constant |
| Plane curves | $\tau = 0$ |
| Circles | $\rho$ =constant, $\tau = 0$ |
| Lines | $\rho = \tau = 0$ |

### 7.10.2 Surfaces in $\mathbb{R}^3$

A surface in $\mathbb{R}^3$ is the collection of end points of the vectors $\vec{x} = \vec{x}(u, v)$, so $x^h = x^h(u^\alpha)$. On the surface are 2 families of curves, one with $u$ =constant and one with $v$ =constant.

The tangent plane in a point $P$ at the surface has basis:

$$\vec{c}_1 = \partial_1\vec{x} \quad \text{and} \quad \vec{c}_2 = \partial_2\vec{x}$$

### 7.10.3 The first fundamental tensor

Let $P$ be a point of the surface $\vec{x} = \vec{x}(u^\alpha)$. The following two curves through $P$, denoted by $u^\alpha = u^\alpha(t)$, $u^\alpha = v^\alpha(\tau)$, have as tangent vectors in $P$

$$\frac{d\vec{x}}{dt} = \frac{du^\alpha}{dt}\partial_\alpha \vec{x} \quad , \quad \frac{d\vec{x}}{d\tau} = \frac{dv^\beta}{d\tau}\partial_\beta \vec{x}$$

The *first fundamental tensor* of the surface in $P$ is the inner product of these tangent vectors:

$$\left(\frac{d\vec{x}}{dt}, \frac{d\vec{x}}{d\tau}\right) = (\vec{c}_\alpha, \vec{c}_\beta)\frac{du^\alpha}{dt}\frac{dv^\beta}{d\tau}$$

The covariant components w.r.t. the basis $\vec{c}_\alpha = \partial_\alpha \vec{x}$ are:

$$g_{\alpha\beta} = (\vec{c}_\alpha, \vec{c}_\beta)$$

For the angle $\phi$ between the parameter curves in $P$: $u = t, v =$constant and $u =$constant, $v = \tau$ holds:

$$\cos(\phi) = \frac{g_{12}}{\sqrt{g_{11}g_{22}}}$$

For the arc length $s$ of $P$ along the curve $u^\alpha(t)$ holds:

$$ds^2 = g_{\alpha\beta}du^\alpha du^\beta$$

This expression is called the *line element*.

### 7.10.4 The second fundamental tensor

The 4 derivatives of the tangent vectors $\partial_\alpha\partial_\beta\vec{x} = \partial_\alpha\vec{c}_\beta$ are each linear independent of the vectors $\vec{c}_1$, $\vec{c}_2$ and $\vec{N}$, with $\vec{N}$ perpendicular to $\vec{c}_1$ and $\vec{c}_2$. This is written as:

$$\partial_\alpha\vec{c}_\beta = \Gamma^\gamma_{\alpha\beta}\vec{c}_\gamma + h_{\alpha\beta}\vec{N}$$

This leads to:

$$\Gamma^\gamma_{\alpha\beta} = (\vec{c}^{\,\gamma}, \partial_\alpha\vec{c}_\beta) \quad , \quad h_{\alpha\beta} = (\vec{N}, \partial_\alpha\vec{c}_\beta) = \frac{1}{\sqrt{\det|g|}}\det(\vec{c}_1, \vec{c}_2, \partial_\alpha\vec{c}_\beta)$$

### 7.10.5 Geodetic curvature

A curve on the surface $\vec{x}(u^\alpha)$ is given by: $u^\alpha = u^\alpha(s)$, than $\vec{x} = \vec{x}(u^\alpha(s))$ with $s$ the arc length of the curve. The length of $\ddot{\vec{x}}$ is the curvature $\rho$ of the curve in $P$. The projection of $\ddot{\vec{x}}$ on the surface is a vector with components

$$p^\gamma = \ddot{u}^\gamma + \Gamma^\gamma_{\alpha\beta}\dot{u}^\alpha\dot{u}^\beta$$

of which the length is called the *geodetic curvature* of the curve in $p$. This remains the same if the surface is curved and the line element remains the same. The projection of $\ddot{\vec{x}}$ on $\vec{N}$ has length

$$p = h_{\alpha\beta}\dot{u}^\alpha\dot{u}^\beta$$

and is called the *normal curvature* of the curve in $P$. The *theorem of Meusnier* states that different curves on the surface with the same tangent vector in $P$ have the same normal curvature.

A *geodetic line* of a surface is a curve on the surface for which in each point the main normal of the curve is the same as the normal on the surface. So for a geodetic line is in each point $p^\gamma = 0$, so

$$\frac{d^2u^\gamma}{ds^2} + \Gamma^\gamma_{\alpha\beta}\frac{du^\alpha}{ds}\frac{du^\beta}{ds} = 0$$

The covariant derivative $\nabla/dt$ in $P$ of a vector field of a surface along a curve is the projection on the tangent plane in $P$ of the normal derivative in $P$.

For two vector fields $\vec{v}(t)$ and $\vec{w}(t)$ along the same curve of the surface follows Leibniz' rule:

$$\frac{d(\vec{v},\vec{w}\,)}{dt} = \left(\vec{v}, \frac{\nabla\vec{w}}{dt}\right) + \left(\vec{w}, \frac{\nabla\vec{v}}{dt}\right)$$

Along a curve holds:

$$\frac{\nabla}{dt}(v^\alpha \vec{c}_\alpha) = \left(\frac{dv^\gamma}{dt} + \Gamma^\gamma_{\alpha\beta}\frac{du^\alpha}{dt}v^\beta\right)\vec{c}_\gamma$$

## 7.11 Riemannian geometry

The *Riemann tensor* $R$ is defined by:

$$R^\mu_{\nu\alpha\beta}T^\nu = \nabla_\alpha\nabla_\beta T^\mu - \nabla_\beta\nabla_\alpha T^\mu$$

This is a $\binom{1}{3}$ tensor with $n^2(n^2-1)/12$ independent components not identically equal to 0. This tensor is a measure for the curvature of the considered space. If it is 0, the space is a flat manifold. It has the following symmetry properties:

$$R_{\alpha\beta\mu\nu} = R_{\mu\nu\alpha\beta} = -R_{\beta\alpha\mu\nu} = -R_{\alpha\beta\nu\mu}$$

The following relation holds:

$$[\nabla_\alpha, \nabla_\beta]T^\mu_\nu = R^\mu_{\sigma\alpha\beta}T^\sigma_\nu + R^\sigma_{\nu\alpha\beta}T^\mu_\sigma$$

The Riemann tensor depends on the Christoffel symbols through

$$R^\alpha_{\beta\mu\nu} = \partial_\mu\Gamma^\alpha_{\beta\nu} - \partial_\nu\Gamma^\alpha_{\beta\mu} + \Gamma^\alpha_{\sigma\mu}\Gamma^\sigma_{\beta\nu} - \Gamma^\alpha_{\sigma\nu}\Gamma^\sigma_{\beta\mu}$$

In a space and coordinate system where the Christoffel symbols are 0 this becomes:

$$R^\alpha_{\beta\mu\nu} = \tfrac{1}{2}g^{\alpha\sigma}(\partial_\beta\partial_\mu g_{\sigma\nu} - \partial_\beta\partial_\nu g_{\sigma\mu} + \partial_\sigma\partial_\nu g_{\beta\mu} - \partial_\sigma\partial_\mu g_{\beta\nu})$$

The *Bianchi identities* are: $\nabla_\lambda R_{\alpha\beta\mu\nu} + \nabla_\nu R_{\alpha\beta\lambda\mu} + \nabla_\mu R_{\alpha\beta\nu\lambda} = 0$.

The *Ricci tensor* is obtained by contracting the Riemann tensor: $R_{\alpha\beta} \equiv R^\mu_{\alpha\mu\beta}$, and is symmetric in its indices: $R_{\alpha\beta} = R_{\beta\alpha}$. The *Einstein tensor* $G$ is defined by: $G^{\alpha\beta} \equiv R^{\alpha\beta} - \frac{1}{2}g^{\alpha\beta}$. It has the property that $\nabla_\beta G^{\alpha\beta} = 0$. The Ricci-scalar is $R = g^{\alpha\beta}R_{\alpha\beta}$.

# Chapter 8

# Numerical mathematics

## 8.1 Errors

There will be an error in the solution if a problem has a number of parameters which are not exactly known. The dependency between errors in input data and errors in the solution can be expressed in the *condition number* $c$. If the problem is given by $x = \phi(a)$ the first-order approximation for an error $\delta a$ in $a$ is:

$$\frac{\delta x}{x} = \frac{a\phi'(a)}{\phi(a)} \cdot \frac{\delta a}{a}$$

The number $c(a) = |a\phi'(a)|/|\phi(a)|$. $c \ll 1$ if the problem is well-conditioned.

## 8.2 Floating point representations

The floating point representation depends on 4 natural numbers:

1. The basis of the number system $\beta$,
2. The length of the mantissa $t$,
3. The length of the exponent $q$,
4. The sign $s$.

Than the representation of machine numbers becomes: $\boxed{\mathrm{rd}(x) = s \cdot m \cdot \beta^e}$ where mantissa $m$ is a number with $t$ $\beta$-based numbers and for which holds $1/\beta \leq |m| < 1$, and $e$ is a number with $q$ $\beta$-based numbers for which holds $|e| \leq \beta^q - 1$. The number 0 is added to this set, for example with $m = e = 0$. The largest machine number is

$$a_{\max} = (1 - \beta^{-t})\beta^{\beta^q - 1}$$

and the smallest positive machine number is

$$a_{\min} = \beta^{-\beta^q}$$

The distance between two successive machine numbers in the interval $[\beta^{p-1}, \beta^p]$ is $\beta^{p-t}$. If $x$ is a real number and the closest machine number is $\mathrm{rd}(x)$, than holds:

$$\begin{aligned} \mathrm{rd}(x) &= x(1+\varepsilon) \quad \text{with} \quad |\varepsilon| \leq \tfrac{1}{2}\beta^{1-t} \\ x &= \mathrm{rd}(x)(1+\varepsilon') \quad \text{with} \quad |\varepsilon'| \leq \tfrac{1}{2}\beta^{1-t} \end{aligned}$$

The number $\eta := \frac{1}{2}\beta^{1-t}$ is called the machine-accuracy, and

$$\varepsilon, \varepsilon' \leq \eta \left| \frac{x - \mathrm{rd}(x)}{x} \right| \leq \eta$$

An often used 32 bits float format is: 1 bit for $s$, 8 for the exponent and 23 for de mantissa. The base here is 2.

## 8.3 Systems of equations

We want to solve the matrix equation $A\vec{x} = \vec{b}$ for a non-singular $A$, which is equivalent to finding the inverse matrix $A^{-1}$. Inverting a $n \times n$ matrix via Cramer's rule requires too much multiplications $f(n)$ with $n! \leq f(n) \leq (e-1)n!$, so other methods are preferable.

### 8.3.1 Triangular matrices

Consider the equation $U\vec{x} = \vec{c}$ where $U$ is a right-upper triangular, this is a matrix in which $U_{ij} = 0$ for all $j < i$, and all $U_{ii} \neq 0$. Than:

$$\begin{aligned} x_n &= c_n/U_{nn} \\ x_{n-1} &= (c_{n-1} - U_{n-1,n}x_n)/U_{n-1,n-1} \\ \vdots &\quad \vdots \\ x_1 &= (c_1 - \sum_{j=2}^{n} U_{1j}x_j)/U_{11} \end{aligned}$$

In code:

```
for (k = n; k > 0; k--)
{
  S = c[k];
  for (j = k + 1; j < n; j++)
  {
    S -= U[k][j] * x[j];
  }
  x[k] = S / U[k][k];
}
```

This algorithm requires $\frac{1}{2}n(n+1)$ floating point calculations.

### 8.3.2 Gauss elimination

Consider a general set $A\vec{x} = \vec{b}$. This can be reduced by Gauss elimination to a triangular form by multiplying the first equation with $A_{i1}/A_{11}$ and than subtract it from all others; now the first column contains all 0's except $A_{11}$. Than the 2nd equation is subtracted in such a way from the others that all elements on the second row are 0 except $A_{22}$, etc. In code:

```
for (k = 1; k <= n; k++)
{
  for (j = k; j <= n; j++) U[k][j] = A[k][j];
  c[k] = b[k];

  for (i = k + 1; i <= n; i++)
  {
    L = A[i][k] / U[k][k];
    for (j = k + 1; j <= n; j++)
    {
      A[i][j] -= L * U[k][j];
    }
    b[i] -= L * c[k];
  }
}
```

This algorithm requires $\frac{1}{3}n(n^2-1)$ floating point multiplications and divisions for operations on the coefficient matrix and $\frac{1}{2}n(n-1)$ multiplications for operations on the right-hand terms, whereafter the triangular set has to be solved with $\frac{1}{2}n(n+1)$ operations.

### 8.3.3 Pivot strategy

Some equations have to be interchanged if the corner elements $A_{11}, A_{22}^{(1)}, ...$ are not all $\neq 0$ to allow Gauss elimination to work. In the following, $A^{(n)}$ is the element after the $n$th iteration. One method is: if $A_{kk}^{(k-1)} = 0$, than search for an element $A_{pk}^{(k-1)}$ with $p > k$ that is $\neq 0$ and interchange the $p$th and the $n$th equation. This strategy fails only if the set is singular and has no solution at all.

## 8.4 Roots of functions

### 8.4.1 Successive substitution

We want to solve the equation $F(x) = 0$, so we want to find the root $\alpha$ with $F(\alpha) = 0$.

Many solutions are essentially the following:

1. Rewrite the equation in the form $x = f(x)$ so that a solution of this equation is also a solution of $F(x) = 0$. Further, $f(x)$ may not vary too much with respect to $x$ near $\alpha$.

2. Assume an initial estimation $x_0$ for $\alpha$ and obtain the series $x_n$ with $x_n = f(x_{n-1})$, in the hope that $\lim_{n\to\infty} x_n = \alpha$.

Example: choose

$$f(x) = \beta - \varepsilon\frac{h(x)}{g(x)} = x - \frac{F(x)}{G(x)}$$

than we can expect that the row $x_n$ with

$$\begin{aligned} x_0 &= \beta \\ x_n &= x_{n-1} - \varepsilon\frac{h(x_{n-1})}{g(x_{n-1})} \end{aligned}$$

converges to $\alpha$.

### 8.4.2 Local convergence

Let $\alpha$ be a solution of $x = f(x)$ and let $x_n = f(x_{n-1})$ for a given $x_0$. Let $f'(x)$ be continuous in a neighbourhood of $\alpha$. Let $f'(\alpha) = A$ with $|A| < 1$. Than there exists a $\delta > 0$ so that for each $x_0$ with $|x_0 - \alpha| \leq \delta$ holds:

1. $\lim_{n\to\infty} n_n = \alpha$,

2. If for a particular $k$ holds: $x_k = \alpha$, than for each $n \geq k$ holds that $x_n = \alpha$. If $x_n \neq \alpha$ for all $n$ than holds

$$\lim_{n\to\infty}\frac{\alpha - x_n}{\alpha - x_{n-1}} = A \quad , \quad \lim_{n\to\infty}\frac{x_n - x_{n-1}}{x_{n-1} - x_{n-2}} = A \quad , \quad \lim_{n\to\infty}\frac{\alpha - x_n}{x_n - x_{n-1}} = \frac{A}{1-A}$$

The quantity $A$ is called the *asymptotic convergence factor*, the quantity $B = -^{10}\log|A|$ is called the *asymptotic convergence speed*.

### 8.4.3 Aitken extrapolation

We define

$$A = \lim_{n\to\infty} \frac{x_n - x_{n-1}}{x_{n-1} - x_{n-2}}$$

$A$ converges to $f'(a)$. Than the row

$$\alpha_n = x_n + \frac{A_n}{1 - A_n}(x_n - x_{n-1})$$

will converge to $\alpha$.

### 8.4.4 Newton iteration

There are more ways to transform $F(x) = 0$ into $x = f(x)$. One essential condition for them all is that in a neighbourhood of a root $\alpha$ holds that $|f'(x)| < 1$, and the smaller $f'(x)$, the faster the series converges. A general method to construct $f(x)$ is:

$$f(x) = x - \Phi(x)F(x)$$

with $\Phi(x) \neq 0$ in a neighbourhood of $\alpha$. If one chooses:

$$\Phi(x) = \frac{1}{F'(x)}$$

Than this becomes Newtons method. The iteration formula than becomes:

$$x_n = x_{n-1} - \frac{F(x_{n-1})}{F'(x_{n-1})}$$

Some remarks:

- This same result can also be derived with Taylor series.
- Local convergence is often difficult to determine.
- If $x_n$ is far apart from $\alpha$ the convergence can sometimes be very slow.
- The assumption $F'(\alpha) \neq 0$ means that $\alpha$ is a simple root.

For $F(x) = x^k - a$ the series becomes:

$$x_n = \frac{1}{k}\left((k-1)x_{n-1} + \frac{a}{x_{n-1}^{k-1}}\right)$$

This is a well-known way to compute roots.

The following code finds the root of a function by means of Newton's method. The root lies within the interval `[x1, x2]`. The value is adapted until the accuracy is better than $\pm$`eps`. The function `funcd` is a routine that returns both the function and its first derivative in point `x` in the passed pointers.

```
float SolveNewton(void (*funcd)(float, float*, float*), float x1, float x2, float eps)
{
  int   j, max_iter = 25;
  float df, dx, f, root;

  root = 0.5 * (x1 + x2);
  for (j = 1; j <= max_iter; j++)
  {
    (*funcd)(root, &f, &df);
    dx  = f/df;
```

```
    root = -dx;
    if ( (x1 - root)*(root - x2) < 0.0 )
    {
      perror("Jumped out of brackets in SolveNewton.");
      exit(1);
    }
    if ( fabs(dx) < eps ) return root; /* Convergence */
  }
  perror("Maximum number of iterations exceeded in SolveNewton.");
  exit(1);
  return 0.0;
}
```

### 8.4.5 The secant method

This is, in contrast to the two methods discussed previously, a two-step method. If two approximations $x_n$ and $x_{n-1}$ exist for a root, than one can find the next approximation with

$$x_{n+1} = x_n - F(x_n)\frac{x_n - x_{n-1}}{F(x_n) - F(x_{n-1})}$$

If $F(x_n)$ and $F(x_{n-1})$ have a different sign one is interpolating, otherwise extrapolating.

## 8.5 Polynomial interpolation

A base for polynomials of order $n$ is given by *Lagrange's interpolation polynomials*:

$$L_j(x) = \prod_{\substack{l=0 \\ l \neq j}}^{n} \frac{x - x_l}{x_j - x_l}$$

The following holds:

1. Each $L_j(x)$ has order $n$,
2. $L_j(x_i) = \delta_{ij}$ for $i, j = 0, 1, ..., n$,
3. Each polynomial $p(x)$ can be written uniquely as

$$p(x) = \sum_{j=0}^{n} c_j L_j(x) \quad \text{with} \quad c_j = p(x_j)$$

This is not a suitable method to calculate the value of a ploynomial in a given point $x = a$. To do this, the Horner algorithm is more usable: the value $s = \sum_k c_k x^k$ in $x = a$ can be calculated as follows:

```
float GetPolyValue(float c[], int n)
{
  int i; float s = c[n];
  for (i = n - 1; i >= 0; i--)
  {
    s = s * a + c[i];
  }
  return s;
}
```

After it is finished `s` has value $p(a)$.

## 8.6 Definite integrals

Almost all numerical methods are based on a formula of the type:

$$\int_a^b f(x)dx = \sum_{i=0}^{n} c_i f(x_i) + R(f)$$

with $n$, $c_i$ and $x_i$ independent of $f(x)$ and $R(f)$ the error which has the form $R(f) = Cf^{(q)}(\xi)$ for all common methods. Here, $\xi \in (a, b)$ and $q \geq n + 1$. Often the points $x_i$ are chosen equidistant. Some common formulas are:

- The trapezoid rule: $n = 1$, $x_0 = a$, $x_1 = b$, $h = b - a$:

$$\int_a^b f(x)dx = \frac{h}{2}[f(x_0) + f(x_1)] - \frac{h^3}{12}f''(\xi)$$

- Simpson's rule: $n = 2$, $x_0 = a$, $x_1 = \frac{1}{2}(a + b)$, $x_2 = b$, $h = \frac{1}{2}(b - a)$:

$$\int_a^b f(x)dx = \frac{h}{3}[f(x_0) + 4f(x_1) + f(x_2)] - \frac{h^5}{90}f^{(4)}(\xi)$$

- The midpoint rule: $n = 0$, $x_0 = \frac{1}{2}(a + b)$, $h = b - a$:

$$\int_a^b f(x)dx = hf(x_0) + \frac{h^3}{24}f''(\xi)$$

The interval will usually be split up and the integration formulas be applied to the partial intervals if $f$ varies much within the interval.

A Gaussian integration formula is obtained when one wants to get both the coefficients $c_j$ and the points $x_j$ in an integral formula so that the integral formula gives exact results for polynomials of an order as high as possible. Two examples are:

1. Gaussian formula with 2 points:

$$\int_{-h}^{h} f(x)dx = h\left[f\left(\frac{-h}{\sqrt{3}}\right) + f\left(\frac{h}{\sqrt{3}}\right)\right] + \frac{h^5}{135}f^{(4)}(\xi)$$

2. Gaussian formula with 3 points:

$$\int_{-h}^{h} f(x)dx = \frac{h}{9}\left[5f\left(-h\sqrt{\tfrac{3}{5}}\right) + 8f(0) + 5f\left(h\sqrt{\tfrac{3}{5}}\right)\right] + \frac{h^7}{15750}f^{(6)}(\xi)$$

## 8.7 Derivatives

There are several formulas for the numerical calculation of $f'(x)$:

- Forward differentiation:

$$f'(x) = \frac{f(x + h) - f(x)}{h} - \tfrac{1}{2}hf''(\xi)$$

- Backward differentiation:
$$f'(x) = \frac{f(x) - f(x-h)}{h} + \tfrac{1}{2}hf''(\xi)$$
- Central differentiation:
$$f'(x) = \frac{f(x+h) - f(x-h)}{2h} - \frac{h^2}{6}f'''(\xi)$$
- The approximation is better if more function values are used:
$$f'(x) = \frac{-f(x+2h) + 8f(x+h) - 8f(x-h) + f(x-2h)}{12h} + \frac{h^4}{30}f^{(5)}(\xi)$$

There are also formulas for higher derivatives:

$$f''(x) = \frac{-f(x+2h) + 16f(x+h) - 30f(x) + 16f(x-h) - f(x-2h)}{12h^2} + \frac{h^4}{90}f^{(6)}(\xi)$$

## 8.8 Differential equations

We start with the first order DE $y'(x) = f(x, y)$ for $x > x_0$ and initial condition $y(x_0) = x_0$. Suppose we find approximations $z_1, z_2, ..., z_n$ for $y(x_1)$, $y(x_2)$,..., $y(x_n)$. Than we can derive some formulas to obtain $z_{n+1}$ as approximation for $y(x_{n+1})$:

- Euler (single step, explicit):
$$z_{n+1} = z_n + hf(x_n, z_n) + \frac{h^2}{2}y''(\xi)$$
- Midpoint rule (two steps, explicit):
$$z_{n+1} = z_{n-1} + 2hf(x_n, z_n) + \frac{h^3}{3}y'''(\xi)$$
- Trapezoid rule (single step, implicit):
$$z_{n+1} = z_n + \tfrac{1}{2}h(f(x_n, z_n) + f(x_{n+1}, z_{n+1})) - \frac{h^3}{12}y'''(\xi)$$

Runge-Kutta methods are an important class of single-step methods. They work so well because the solution $y(x)$ can be written as:

$$y_{n+1} = y_n + hf(\xi_n, y(\xi_n)) \quad \text{with} \quad \xi_n \in (x_n, x_{n+1})$$

Because $\xi_n$ is unknown some "measurements" are done on the increment function $k = hf(x, y)$ in well chosen points near the solution. Than one takes for $z_{n+1} - z_n$ a weighted average of the measured values. One of the possible 3rd order Runge-Kutta methods is given by:

$$\begin{aligned} k_1 &= hf(x_n, z_n) \\ k_2 &= hf(x_n + \tfrac{1}{2}h, z_n + \tfrac{1}{2}k_1) \\ k_3 &= hf(x_n + \tfrac{3}{4}h, z_n + \tfrac{3}{4}k_2) \\ z_{n+1} &= z_n + \tfrac{1}{9}(2k_1 + 3k_2 + 4k_3) \end{aligned}$$

and the classical 4th order method is:

$$\begin{aligned} k_1 &= hf(x_n, z_n) \\ k_2 &= hf(x_n + \tfrac{1}{2}h, z_n + \tfrac{1}{2}k_1) \\ k_3 &= hf(x_n + \tfrac{1}{2}h, z_n + \tfrac{1}{2}k_2) \\ k_4 &= hf(x_n + h, z_n + k_3) \\ z_{n+1} &= z_n + \tfrac{1}{6}(k_1 + 2k_2 + 2k_3 + k_4) \end{aligned}$$

Often the accuracy is increased by adjusting the stepsize for each step with the estimated error. Step doubling is most often used for 4th order Runge-Kutta.

## 8.9 The fast Fourier transform

The Fourier transform of a function can be approximated when some discrete points are known. Suppose we have $N$ successive samples $h_k = h(t_k)$ with $t_k = k\Delta$, $k = 0, 1, 2, ..., N-1$. Than the discrete Fourier transform is given by:

$$H_n = \sum_{k=0}^{N-1} h_k \mathrm{e}^{2\pi i k n/N}$$

and the inverse Fourier transform by

$$h_k = \frac{1}{N} \sum_{n=0}^{N-1} H_n \mathrm{e}^{-2\pi i k n/N}$$

This operation is order $N^2$. It can be faster, order $N \cdot^2 \log(N)$, with the fast Fourier transform. The basic idea is that a Fourier transform of length $N$ can be rewritten as the sum of two discrete Fourier transforms, each of length $N/2$. One is formed from the even-numbered points of the original $N$, the other from the odd-numbered points.

This can be implemented as follows. The array `data[1..2*nn]` contains on the odd positions the real and on the even positions the imaginary parts of the input data: `data[1]` is the real part and `data[2]` the imaginary part of $f_0$, etc. The next routine replaces the values in `data` by their discrete Fourier transformed values if `isign` $= 1$, and by their inverse transformed values if `isign` $= -1$. `nn` must be a power of 2.

```
#include <math.h>
#define SWAP(a,b) tempr=(a);(a)=(b);(b)=tempr

void FourierTransform(float data[], unsigned long nn, int isign)
{
  unsigned long n, mmax, m, j, istep, i;
  double        wtemp, wr, wpr, wpi, wi, theta;
  float         tempr, tempi;

  n = nn << 1;
  j = 1;
  for (i = 1; i < n; i += 2)
  {
    if ( j > i )
    {
      SWAP(data[j], data[i]);
      SWAP(data[j+1], data[i+1]);
    }
    m = n >> 1;
    while ( m >= 2 && j > m )
    {
      j -= m;
      m >>= 1;
    }
    j += m;
  }
  mmax = 2;
  while ( n > mmax ) /* Outermost loop, is executed log2(nn) times */
  {
    istep = mmax << 1;
    theta = isign * (6.28318530717959/mmax);
    wtemp = sin(0.5 * theta);
    wpr   = -2.0 * wtemp * wtemp;
    wpi   = sin(theta);
```

```
    wr    = 1.0;
    wi    = 0.0;
    for (m = 1; m < mmax; m += 2)
    {
      for (i = m; i <= n; i += istep) /* Danielson-Lanczos equation */
      {
        j          = i + mmax;
        tempr      = wr * data[j]   - wi * data[j+1];
        tempi      = wr * data[j+1] + wi * data[j];
        data[j]    = data[i]   - tempr;
        data[j+1]  = data[i+1] - tempi;
        data[i]   += tempr;
        data[i+1] += tempi;
      }
      wr = (wtemp = wr) * wpr - wi * wpi + wr;
      wi = wi * wpr + wtemp * wpi + wi;
    }
    mmax=istep;
  }
}
```