Mathematische Rechenmethoden 1

Stefan Weinzierl

28. Mai 2010

Inhaltsverzeichnis

1 Einführung

Literatur:

- W. Schäfer, K. Georgi, G. Trippler, Mathematik Vorkurs, Teubner 1997
- S. Großmann, Mathematischer Einführungskurs, Teubner 2004
- K. Weltner, Mathematik für Physiker, Springer 2001
- R. Brauner, F. Geiß, Abiturwissen Mathematik, Fischer 2006

Skripten:

- S. Groote, Mathematischer Vorkurs,

  `http://wwwthep.physik.uni-mainz.de/~groote/vorkurs/`

- E. Alt, Mathematischer Einführungskurs,

  `http://wwwthep.physik.uni-mainz.de/~alt/lect-notes_ps/Math_Vorkurs.ps.gz/`

- K. Hefft, Mathematischer Vorkurs,

  `http://www.thphys.uni-heidelberg.de/~hefft/vk1/`

Lehrbücher:

- O. Forster, Analysis 1-3, Vieweg
- G. Fischer, Lineare Algebra, Vieweg
- W. Fischer, I. Lieb, Funktionentheorie, Vieweg

2 Schreibweisen

$\{a,b,c\}$: Menge der Elemente a, b, und c.
Bemerkung: Die Ordnung spielt keine Rolle: $\{a,b,c\} = \{b,a,c\}$

Vereinigung: $A \cup B$ enthält alle Elemente sowohl aus A als auch aus B.
$\{a,b,c\} \cup \{c,d,e\} = \{a,b,c,d,e\}$.

Durchschnitt: $A \cap B$ enhält alle Elemente die sowohl in A als auch in B enthalten sind.
$\{a,b,c\} \cap \{c,d,e\} = \{c\}$.

$a \in A$: a ist ein Element der Menge A.

$A \subset B$: Die Menge A ist eine Teilmenge der Menge B.

$[a,b]$: Intervall, die Grenzen sind im Intervall enthalten: $a \in [a,b]$, $b \in [a,b]$
$]a,b[$: Intervall, die Grenzen sind im Intervall nicht enthalten: $a \notin [a,b]$, $b \notin [a,b]$
Analog: $[a,b[$ und $]a,b]$.

Logisch und: $\wedge$

Logisch oder: $\vee$

Negation: $\neg$

$\exists$: Es existiert

$\forall$: Für alle

∞: Symbol für Unendlich.

$\sum$: Summenzeichen

$$\sum_{j=1}^{n} a_j = a_1 + a_2 + a_3 + \ldots + a_{n-1} + a_n.$$

$\prod$: Produktzeichen

$$\prod_{j=1}^{n} a_j = a_1 \cdot a_2 \cdot a_3 \cdot \ldots \cdot a_{n-1} \cdot a_n.$$

$n!$: Fakultät. Bemerkung $0! = 1$, $1! = 1$.

Binomialkoeffizient:

$$\begin{pmatrix} n \\ k \end{pmatrix} = \frac{n!}{k!(n-k)!}.$$

$\lim\limits_{x\to a}$: Grenzwert für den Fall, daß sich x dem Wert a annähert.

Ableitung: Sei $f(x)$ eine Funktion von x.

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

Integral:

$$\int_a^b f(x)dx = \lim_{n\to\infty} \sum_{j=1}^{n} f(\xi_j)\Delta x_j, \quad \Delta x_j = x_j - x_{j-1},\ x_0 = a,\ x_n = b,\ \xi_j \in [x_{j-1}, x_j].$$

3 Zahlen

3.1 Die natürlichen Zahlen

$\mathbb{N}$: Die natürlichen Zahlen $\mathbb{N} = \{1, 2, 3, 4, ...\}$.

Die **Axiome von Peano** für die natürlichen Zahlen:

- (P1) 1 ist eine natürliche Zahl.
- (P2) Falls a eine natürliche Zahl, so ist die nachfolgende Zahle $a+1$ ebenfalls eine natürliche Zahl.
- (P3) Die natürlichen Zahlen sind die minimale Menge, welche die ersten beiden Axiome erfüllt.

$\mathbb{N}_0$: Die natürlichen Zahlen mit der Null $\mathbb{N}_0 = \{0, 1, 2, 3, 4, ...\}$.

Der **Induktionsbeweis**: Beispiel: Für jede natürliche Zahl n ist die Behauptung

$$\sum_{j=1}^{n} j = \frac{n(n+1)}{2}$$

zu zeigen. Der Induktionsbeweis verläuft in zwei Schritten: Man zeigt zunächst die Behauptung für $n = 1$ (Induktionsanfang):

$$\begin{aligned} \text{linke Seite}: \quad & \sum_{j=1}^{n} j = 1. \\ \text{rechte Seite}: \quad & \frac{1(1+1)}{2} = 1. \end{aligned}$$

Im zweiten Schritt darf man nun verwenden, daß die Behauptung für $n-1$ richtig ist, und zeigt dann, daß sie auch für n gilt (Induktionsschritt). In unserem Fall:

$$\sum_{j=1}^{n} j = \left(\sum_{j=1}^{n-1} j\right) + n = \frac{(n-1)n}{2} + n = \frac{n^2 - n + 2n}{2} = \frac{n(n+1)}{2}$$

3.2 Die ganzen Zahlen

$\mathbb{Z}$: Die ganzen Zahlen $\mathbb{Z} = \{...., -3, -2, -1, 0, 1, 2, 3, ...\}$.

Die ganzen Zahlen bilden bezüglich der Addition eine Gruppe.

Definition einer Gruppe: Eine nicht-leere Menge G mit einer Verknüpfung $\circ$ nennt man Gruppe $(G, \circ)$, falls die folgenden Axioma gelten:

- (G1) $\circ$ ist assoziativ: $a \circ (b \circ c) = (a \circ b) \circ c$
- (G2) Es gibt ein links-neutrales Element : $e \circ a = a$ für alle $a \in G$
- (G3) Zu jedem $a \in G$ gibt es ein links-inverses Element a^{-1} : $a^{-1}a = e$

Eine Gruppe $(G, \circ)$ nennt man kommutativ oder Abelsch, falls $a \circ b = b \circ a$.

Bemerkung: Falls $(G, \circ)$ eine Gruppe ist, dann ist das links-neutrale Element identisch mit dem recht-neutralen Element. Ebenso sind links- und rechts-inverses Element identisch.

Ein Ring ist eine nicht-leere Menge R mit zwei Verknüpfungen, die üblicherweise als $+$ und $\cdot$ geschrieben werden, so daß

- (R1) $(R, +)$ ist eine Abelsche Gruppe.
- (R2) $(R, \cdot)$ ist assoziativ: $a \cdot (b \cdot c) = (a \cdot b) \cdot c$
- (R3) Es gelten die Distributivgesetze:

$$\begin{aligned} a \cdot (b+c) &= (a \cdot b) + (a \cdot c) \\ (a+b) \cdot c &= (a \cdot c) + (a \cdot b) \end{aligned}$$

Die ganzen Zahlen $\mathbb{Z}$ bilden einen Ring.

3.3 Die rationalen Zahlen

$\mathbb{Q}$: Die rationalen Zahlen.

$$\mathbb{Q} = \left\{ \frac{p}{q} \mid p, q \in \mathbb{Z} \right\}.$$

Die rationalen Zahlen sind bezüglich der Division abgeschlossen. Sie bilden einen Körper.

Eine nicht-leere Menge K mit zwei Verknüpfungen $+$ und $\cdot$ nennt man Körper, falls gilt:

- (K1) $(K, +)$ ist eine Abelsche Gruppe.
- (K2) $(K \setminus \{0\}, \cdot)$ ist eine Abelsche Gruppe.
- (K3) Es gelten die Distributivgesetze:

$$\begin{aligned} a \cdot (b+c) &= (a \cdot b) + (a \cdot c) \\ (a+b) \cdot c &= (a \cdot c) + (a \cdot b) \end{aligned}$$

Bruchrechnen:

$$\frac{p_1}{q_1} \cdot \frac{p_2}{q_2} = \frac{p_1 \cdot p_2}{q_1 \cdot q_2}, \quad \frac{p_1}{q_1} \Big/ \frac{p_2}{q_2} = \frac{p_1 \cdot q_2}{q_1 \cdot p_2}, \quad \frac{c \cdot p_1}{c \cdot q_1} = \frac{p_1}{q_1} \quad \frac{p_1}{q_1} + \frac{p_2}{q_2} = \frac{p_1 \cdot q_2 + p_2 \cdot q_1}{q_1 \cdot q_2}.$$

Rechnen mit Potenzen:

$$\begin{aligned}
a^n &= \underbrace{a \cdot a \cdot a \cdot \ldots \cdot a}_{n\ mal}, \\
a^n \cdot b^n &= (a \cdot b)^n, \\
a^n \cdot a^m &= a^{n+m} \\
(a^n)^m &= a^{n \cdot m} \\
\frac{a^n}{b^n} &= \left(\frac{a}{b}\right)^n \\
\frac{a^n}{a^m} &= a^{n-m}
\end{aligned}$$

3.4 Die reellen Zahlen

$\mathbb{R}$: Die reellen Zahlen.

Die reellen Zahlen bilden einen Körper.

Alle rationalen Zahlen sind in den reellen Zahlen enthalten. Darüberhinaus enthält $\mathbb{R}$ Zahlen, die nicht rational sind. Diese nennt man irrational. So ist zum Beispiel $\sqrt{2}$ eine irrationale Zahl. $\sqrt{2}$ ist Lösung der Gleichung $x^2 = 2$. Zahlen, welche Lösungen einer algebraischen Gleichung sind, nennt man algebraisch. Darüberhinaus enthält $\mathbb{R}$ irrationale Zahlen, die keine Lösung einer algebraischen Gleichung sind. Solche Zahlen nennt man transzendental. Die Kreiszahl π oder die Eulersche Konstante e sind transzendental.

Eine Folge $(a_n)_{n \in \mathbb{N}}$ reeller Zahlen nennt man Cauchy-Folge, falls es zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$ gibt, so daß

$$|a_n - a_m| < \varepsilon, \quad \forall n, m \geq N.$$

Vollständigkeitsaxiom: Jede Cauchy-Folge konvergiert.

Darüberhinaus sind die reellen Zahlen angeordnet: Es sind gewisse Elemente als positiv ausgezeichnet $(x > 0)$, so daß die folgenden Axiome erfüllt sind:

- (O1) Es gilt genau eine der Beziehungen $x < 0$, $x = 0$ oder $x > 0$.

- (O2) Aus $x > 0$ und $y > 0$ folgt $x + y > 0$.

- (O3) Aus $x > 0$ und $y > 0$ folgt $x \cdot y > 0$.

Man nennt eine Ordnung archimedisch, falls zu jedem $x > 0$ und $y > 0$ ein natürliche Zahl n exisiert, so daß

$$n \cdot x > y.$$

Axiomatisch lassen sich die reellen Zahlen als ein Körper, der archimedisch angeordnet ist und in dem jede Cauchy-Folge konvergiert, charakterisieren.

3.5 Die komplexen Zahlen

$\mathbb{C}$: Die komplexen Zahlen.

Definition: Man definiert die imaginäre Einheit i als eine Lösung der Gleichung

$$x^2 = -1.$$

Die komplexen Zahlen $\mathbb{C}$ sind die Menge

$$\mathbb{C} = \{x + iy \mid x, y \in \mathbb{R}\}.$$

Wir führen eine Addition und Multiplikation auf $\mathbb{C}$ ein: Sei $z_1 = x_1 + iy_1$ und $z_2 = x_2 + iy_2$.

$$\begin{aligned} z_1 + z_2 &= (x_1 + iy_1) + (x_2 + iy_2) = (x_1 + x_2) + i(y_1 + y_2), \\ z_1 \cdot z_2 &= (x_1 + iy_1) \cdot (x_2 + iy_2) = (x_1 x_2 - y_1 y_2) + i(x_1 y_2 + y_1 x_2) \end{aligned}$$

Mit dieser Addition und Multiplikation bilden die komplexen Zahlen einen Körper. Dieser Körper ist algebraisch abgeschlossen, d.h. die Nullstellen eines jeden Polynoms liegen in dem Körper. Der Körper ist allerdings nicht angeordnet. Das Vollständigkeitsaxiom gilt.

Sei $z = x + iy$ eine komplexe Zahl. Man bezeichnet x als Realteil und y als Imaginärteil.

$$\begin{aligned} \mathrm{Re}\, z &= x, \\ \mathrm{Im}\, z &= y. \end{aligned}$$

Die zu $z = x + iy$ konjugiert komplexe Zahl ist

$$z^* = x - iy.$$

Es ist

$$z \cdot z^* = (x + iy) \cdot (x - iy) = x^2 + y^2.$$

Als Betrag der komplexen Zahl bezeichnet man

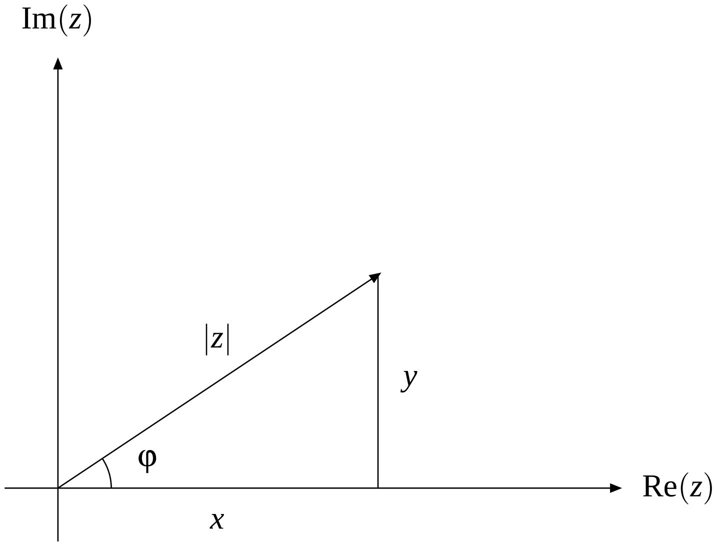
$$|z| = \sqrt{zz^*} = \sqrt{x^2 + y^2}.$$

Polardarstellung einer komplexen Zahl:

$$z \;=\; |z|\cdot(\cos\varphi + i\sin\varphi)\,.$$

φ nennt man das Argument oder die Phase der komplexen Zahl.

Die komplexe Zahlenebene



Umrechnung: Normalform in Polarform:

$$\begin{aligned}
|z| &= \sqrt{x^2+y^2}\\
\tan\varphi &= \frac{y}{x},\quad x\neq 0,\\
\phi &= \frac{\pi}{2}\;\; \text{für} x=0,\; y>0,\\
\phi &= \frac{3\pi}{2}\;\; \text{für} x=0,\; y<0.
\end{aligned}$$

Polarform in Normalform:

$$x = |z|\cos\varphi,\qquad y = |z|\sin\varphi.$$

Rechenregeln:

$$\begin{aligned}
z_1 - z_2 &= (x_1+iy_1)-(x_2+iy_2) = (x_1-x_2)+i\,(y_1-y_2)\,,\\
\frac{z_1}{z_2} &= \frac{x_1+iy_1}{x_2+iy_2} = \frac{(x_1+iy_1)\cdot(x_2-iy_2)}{(x_2+iy_2)\,(x_2-iy_2)} = \frac{(x_1x_2+y_1y_2)+i\,(-x_1y_2+y_1x_2)}{x_2^2+y_2^2}.
\end{aligned}$$

In der Polarform sind Multiplikation und Division besonders einfach:

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

Insbesondere:

$$z^n = |z|^n (\cos n\varphi + i \sin n\varphi)\,.$$

4 Lineare Algebra

4.1 Vektorräume

Sei K ein Körper und $(V,+)$ eine Abelsche Gruppe. Weiter sei eine zusätzliche Verknüpfung gegeben, die man skalare Multiplikation nennt:

$$\begin{aligned} K \times V &\rightarrow V \\ (k,\vec{v}) &\rightarrow k \cdot \vec{v} \end{aligned}$$

V ist ein K-Vektorraum falls gilt:

- (V1) $(K,+,\cdot)$ ist ein Körper
- (V2) $(V,+)$ ist eine Abelsche Gruppe
- (V3) Es gelten die Distributivgesetze:

$$\begin{aligned} k \cdot (\vec{v}_1 + \vec{v}_2) &= (k \cdot \vec{v}_1) + (k \cdot \vec{v}_2) \\ (k_1 + k_2) \cdot \vec{v} &= (k_1 \cdot \vec{v}) + (k_2 \cdot \vec{v}) \end{aligned}$$

- (V4) Es gilt das Assoziativgesetz:

$$k_1 \cdot (k_2 \cdot \vec{v}) = (k_1 \cdot k_2) \cdot \vec{v}$$

  Bemerkung: Bei $(k_1 \cdot k_2)$ ist die Multiplikation im Körper gemeint.

- (V5) Für die Eins gilt:

$$1 \cdot \vec{v} = \vec{v}$$

  für alle $\vec{v} \in V$.

Als Grundkörper treten in der Physik fast immer $\mathbb{R}$ oder $\mathbb{C}$ auf. Beispiele für Vektorräume sind der $\mathbb{R}^n$ (mit Grundkörper $\mathbb{R}$) und der $\mathbb{C}^n$ (mit Grundkörper $\mathbb{C}$).

$$\mathbb{R}^n = \{(x_1,x_2,\ldots,x_n)|x_1,x_2,\ldots,x_n \in \mathbb{R}\}.$$

Man schreibt die Elemente aus dem Vektorraum als Spaltenvektoren, so zum Beispiel:

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3.$$

Ebenso ist die Schreibweise als Zeilenvektor gebräuchlich:

$$(x,y,z) \in \mathbb{R}^{3^*}.$$

V^* bezeichnet den zu V dualen Vektorraum (falls V alle Spaltenvektoren enthält, so enthält V^* die Zeilenvektoren). Bei der Summe zweier Vektoren werden die Vektoren komponentenweise addiert:

$$\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} = \begin{pmatrix} 5 \\ 7 \\ 9 \end{pmatrix}.$$

Bei der skalaren Multiplikation wird jede Komponente mit dem Skalar multipliziert:

$$3 \cdot \begin{pmatrix} 4 \\ 5 \\ 6 \end{pmatrix} = \begin{pmatrix} 12 \\ 15 \\ 18 \end{pmatrix}.$$

Schreibweise: Man bezeichnet mit $\vec{v}^T$ den zu $\vec{v}$ transponierten Vektor (d.h. aus einem Spaltenvektor wird ein Zeilenvektor, und aus einem Zeilenvektor wird ein Spaltenvektor):

$$(x_1, x_2, ..., x_n)^T = \begin{pmatrix} x_1 \\ x_2 \\ ... \\ x_n \end{pmatrix}.$$

Definition: Vektoren, die in fast allen Komponenten eine Null haben, bis auf eine Komponente, in der sie eine Eins haben, spielen eine wichtige Rolle. Hat so ein Vektor in der i-ten Komponente eine Eins,

$$\vec{e}_i = \Big(\underbrace{0, 0, ..., 0}_{i-1}, 1, 0, ..., 0\Big)^T,$$

so bezeichnet man diesen Vektor als den i-ten Einheitsvektor.

Definition: Seien n Vektoren $\vec{v}_1$, $\vec{v}_2$, ..., $\vec{v}_n$ gegeben. Folgt aus

$$\begin{aligned} a_1\vec{v}_1 + a_2\vec{v}_2 + ... + a_n\vec{v}_n &= \vec{0} \\ &\Rightarrow a_1 = a_2 = ... = a_n = 0, \end{aligned}$$

so nennt man die Vektoren linear unabhängig. Anderfalls nennt man sie linear abhängig.

Definition: Sei V ein Vektorraum. Die maximale Anzahl linear unabhängiger Vektoren in V nennt man die Dimension des Vektorraumes. Eine Menge linearer unabhängiger Vektoren, die maximal ist, nennt man eine Basis von V.

Beispiel: $\mathbb{R}^n$ und $\mathbb{C}^n$ haben die Dimension n. Eine Basis von $\mathbb{R}^n$ und $\mathbb{C}^n$ ist zum Beispiel

$$\{\vec{e}_1, \vec{e}_2, ..., \vec{e}_n\}.$$

Man nennt diese Basis die Standardbasis.

4.2 Skalarprodukte

Wir betrachten zunächst den $\mathbb{R}^n$. Seien $\vec{x}, \vec{y} \in \mathbb{R}^n$. Die Komponentendarstellung der beiden Vektoren sei

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{pmatrix}, \qquad \vec{y} = \begin{pmatrix} y_1 \\ y_2 \\ \dots \\ y_n \end{pmatrix}.$$

Wir definieren das Skalarprodukt zwischen zwei Vektoren als eine Abbildung

$$\begin{aligned} V \times V &\rightarrow \mathbb{R}, \\ \vec{x} \cdot \vec{y} &= x_1 y_1 + x_2 y_2 + \dots + x_n y_n. \end{aligned}$$

Das Skalarprodukt ist eine positiv definite symmetrische Bilinearform, es erfüllt die folgenden Bedingungen:

- Linear in der ersten Komponente:

$$\begin{aligned} (\vec{x} + \vec{y}) \cdot \vec{z} &= \vec{x} \cdot \vec{z} + \vec{y} \cdot \vec{z}, \\ (\lambda \cdot \vec{x}) \cdot \vec{y} &= \lambda (\vec{x} \cdot \vec{y}). \end{aligned}$$

- Linear in der zweiten Komponente:

$$\begin{aligned} \vec{x} \cdot (\vec{y} + \vec{z}) &= \vec{x} \cdot \vec{y} + \vec{x} \cdot \vec{z}, \\ \vec{x} \cdot (\lambda \cdot \vec{y}) &= \lambda (\vec{x} \cdot \vec{y}). \end{aligned}$$

- Symmetrisch:

$$\vec{x} \cdot \vec{y} = \vec{y} \cdot \vec{x}.$$

- Positiv definit:

$$\vec{x} \cdot \vec{x} > 0, \quad \text{falls} \quad \vec{x} \neq \vec{0}.$$

Wir betrachten nun $\mathbb{C}^n$. Seien $\vec{x}, \vec{y} \in \mathbb{C}^n$. In diesem Fall definieren wir das Skalarprodukt als

$$\begin{aligned} V \times V &\rightarrow \mathbb{C}, \\ \vec{x} \cdot \vec{y} &= x_1^* y_1 + x_2^* y_2 + \dots + x_n^* y_n. \end{aligned}$$

Dieses Skalarprodukt ist eine positiv definite Hermitische Form, es erfüllt die folgenden Bedingungen:

- Semilinear in der ersten Komponente:

$$\begin{aligned} (\vec{x}+\vec{y})\cdot\vec{z} &= \vec{x}\cdot\vec{z}+\vec{y}\cdot\vec{z}, \\ (\lambda\cdot\vec{x})\cdot\vec{y} &= \lambda^*\,(\vec{x}\cdot\vec{y})\,. \end{aligned}$$

- Linear in der zweiten Komponente:

$$\begin{aligned} \vec{x}\cdot(\vec{y}+\vec{z}) &= \vec{x}\cdot\vec{y}+\vec{x}\cdot\vec{z}, \\ \vec{x}\cdot(\lambda\cdot\vec{y}) &= \lambda\,(\vec{x}\cdot\vec{y})\,. \end{aligned}$$

- Hermitisch:

$$\vec{x}\cdot\vec{y} = (\vec{y}\cdot\vec{x})^*\,.$$

- Positiv definit:

$$\vec{x}\cdot\vec{x} > 0, \quad \text{falls} \quad \vec{x}\neq\vec{0}.$$

Definition: Man bezeichnet mit

$$|\vec{x}| = \sqrt{\vec{x}\cdot\vec{x}}$$

die Länge oder den Betrag von $\vec{x}$.

Sei $\vec{x},\vec{y}\in\mathbb{R}^n$. Der Winkel zwischen den beiden Vektoren ist gegeben durch

$$\vec{x}\cdot\vec{y} = |\vec{x}|\,|\vec{y}|\cos\varphi,$$

also

$$\varphi = \arccos\frac{\vec{x}\cdot\vec{y}}{|\vec{x}|\,|\vec{y}|}.$$

Zwei Vektoren stehen senkrecht aufeinander ($\varphi = 90^\circ$), falls

$$\vec{x}\cdot\vec{y} = 0.$$

4.3 Das Kreuzprodukt

Sei V der Vektorraum $\mathbb{R}^3$ oder $\mathbb{C}^3$. In einem dreidimensionalen Vektorraum ist zusätzlich das Kreuzprodukt als eine Abbildung

$$\begin{aligned} V\times V &\to V, \\ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}\times\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} &= \begin{pmatrix} x_2y_3-x_3y_2 \\ x_3y_1-x_1y_3 \\ x_1y_2-x_2y_1 \end{pmatrix} \end{aligned}$$

definiert.

Wichtig: Das Kreuzprodukt gibt es nur in drei Dimensionen !

Das Kreuzprodukt ist anti-symmetrisch:

$$\vec{x} \times \vec{y} = -\vec{y} \times \vec{x}.$$

Der Vektor $\vec{x} \times \vec{y}$ steht senkrecht auf $\vec{x}$ und $\vec{y}$:

$$\begin{aligned} \vec{x} \cdot (\vec{x} \times \vec{y}) &= 0, \\ \vec{y} \cdot (\vec{x} \times \vec{y}) &= 0, \end{aligned}$$

Für den Betrag von $\vec{x} \times \vec{y}$ gilt:

$$|\vec{x} \times \vec{y}| = |\vec{x}|\,|\vec{y}| \sin\varphi,$$

wobei φ der Winkel zwischen $\vec{x}$ und $\vec{y}$ ist. Sei

$$\vec{z} = \vec{x} \times \vec{y}.$$

Für die Komponenten von $\vec{z}$ gilt:

$$z_i = \sum_{j=1}^{3} \sum_{k=1}^{3} \varepsilon_{ijk} x_j y_k$$

Hier wurde der antisymmetrische Tensor (oder Levi-Civita-Tensor) ε_{ijk} verwendet, der wie folgt definiert ist:

$$\varepsilon_{ijk} = \begin{cases} +1 & \text{für } (i,j,k) \text{ eine gerade Permutation von } (1,2,3), \\ -1 & \text{für } (i,j,k) \text{ eine ungerade Permutation von } (1,2,3), \\ 0 & \text{sonst.} \end{cases}$$

Eine Permutation $(\sigma_1, \sigma_2, ..., \sigma_n)$ nennt man gerade, wenn man sie durch eine gerade Anzahl von paarweisen Vertauschungen aus $(1,2,...,n)$ erzeugen kann. Benötigt man eine ungerade Anzahl von Vertauschungen, so nennt man die Permutation ungerade.
Beispiel:

$$\begin{aligned} &(3,2,1,5,4) && \text{ist eine gerade Permutation (vertausche } 1 \leftrightarrow 3 \text{ und } 4 \leftrightarrow 5), \\ &(1,5,3,4,2) && \text{ist eine ungerade Permutation (vertausche } 2 \leftrightarrow 5). \end{aligned}$$

An dieser Stelle definieren wir auch noch das Kronecker-Delta-Symbol:

$$\delta_{ij} = \begin{cases} +1 & \text{für } i = j, \\ 0 & \text{für } i \neq j. \end{cases}$$

4.4 Vektoren in der Physik

Ein Punkt im Raum wird durch einen Ortsvektor

$$\vec{r} = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$$

beschrieben. Der zugrundeliegende Vektorraum ist der $\mathbb{R}^3$.

In der Physik (insbesondere in der speziellen Relativitätstheorie) fast man oft die Zeit und die drei Ortskoordinaten zu einem Vierervektor zusammen:

$$\begin{pmatrix} ct \\ x \\ y \\ z \end{pmatrix}$$

Die Komponenten eines Vierervektors notiert man als x^μ, wobei $\mu = 0, 1, 2, 3$. Für die Komponenten eines Vierervektors verwendet man üblicherweise griechische Indizes wie zum Beispiel $\mu, \nu, \rho, \sigma, ...$, für die Komponenten eine Dreiervektors hingegen lateinische Indizes $i, j, k, ...$. Der zugrundeliegende Vektorraum ist der $\mathbb{R}^4$, allerdings nicht mit dem euklidischen Skalarprodukt, sondern dem Minkowski-Skalarprodukt. Seien x und y Vierervektoren, so setzt man

$$x \cdot y = x^0 y^0 - x^1 y^1 - x^2 y^2 - x^3 y^3.$$

Dieses Skalarprodukt ist nicht positiv definit !

4.5 Lineare Gleichungssysteme

Definition: Unter einem linearen Gleichungssystem versteht man n Gleichungen mit m Unbekannten x_1, x_2, ..., x_m der Form

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + ... + a_{1m}x_m &= b_1, \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + ... + a_{2m}x_m &= b_2, \\ &\ldots \\ a_{n1}x_1 + a_{n2}x_2 + a_{n3}x_3 + ... + a_{nm}x_m &= b_n. \end{aligned}$$

Die Koeffizienten a_{ij} und b_i sind gegebene reelle oder komplexe Zahlen. Jede Variable kommt nur linear vor und jeder Term auf der linken Seite enthält nur eine Variable, daher der Name "lineares Gleichungssystem". Beispiel:

$$\begin{aligned} 3x_1 + 3x_2 + 9x_3 &= 36, \\ 2x_1 + 3x_2 + 7x_3 &= 29, \\ x_2 + 4x_3 &= 14. \end{aligned}$$

Wir betrachten nun einen Algorithmus um ein Gleichungssystem mit n Gleichungen und m Unbekannten systematisch zu vereinfachen und zu lösen. Wir beginnen mit einer trivialen Beobachtung: Offensichtlich können Zeilen vertauscht werden, d.h. das Gleichungssystem

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \ldots + a_{1m}x_m &= b_1, \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \ldots + a_{2m}x_m &= b_2, \end{aligned}$$

ist äquivalent zu dem Gleichungssystem

$$\begin{aligned} a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \ldots + a_{2m}x_m &= b_2, \\ a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \ldots + a_{1m}x_m &= b_1. \end{aligned}$$

Desweiteren sei $(x_1, x_2, \ldots, x_m)$ ein m-Tupel, welches die Gleichung

$$a_1x_1 + a_2x_2 + a_3x_3 + \ldots + a_mx_m = b,$$

erfüllt. Dann erfüllt es auch die Gleichung

$$(ca_1)x_1 + (ca_2)x_2 + (ca_3)x_3 + \ldots + (ca_m)x_m = cb,$$

Umgekehrt gilt, daß für $c \neq 0$ jedes m-Tupel, welches die zweite Gleichung erfüllt, auch die erste Gleichung erfüllt. Daraus folgt, daß man die linke und rechte Seite einer Gleichung mit einer konstanten Zahl c ungleich Null multiplizieren darf.
Die dritte elementare Umformung ist die folgende: Man darf eine Zeile durch die Summe dieser Zeile mit einer anderen Zeile ersetzen, d.h. die Gleichungssysteme

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \ldots + a_{1m}x_m &= b_1, \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \ldots + a_{2m}x_m &= b_2, \end{aligned}$$

und

$$\begin{aligned} (a_{11}+a_{21})\,x_1 + (a_{12}+a_{22})\,x_2 + (a_{13}+a_{23})\,x_3 + \ldots + (a_{1m}+a_{2m})\,x_m &= b_1, \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \ldots + a_{2m}x_m &= b_2, \end{aligned}$$

haben die gleichen Lösungen.

Mit Hilfe dieser drei elementaren Umformungen läßt sich nun ein Algorithmus zur systematischen Vereinfachung von linearen Gleichungssystemen angeben:

1. Setze $i = 1$ (Zeilenindex), $j = 1$ (Spaltenindex).

2. Falls $a_{ij} = 0$ suche $k > i$, so daß $a_{kj} \neq 0$ und vertausche Zeilen i und k.

3. Falls ein solches k aus Schritt 2 nicht gefunden werden kann, setze $j \to j+1$ und gehe zurück zu Schritt 2.

4. Falls man in Schritt 3 den Wert $j = m+1$ erreicht, beende den Algorithmus.
5. Multipliziere Zeile i mit $1/a_{ij}$.
6. Für alle Zeilen $k \neq i$ addiere zur Zeile k das $-a_{kj}$-fache der i-ten Zeile.
7. Setze $i \to i+1$ und $j \to j+1$ und gehe zurück zu Schritt 2.
8. Falls man in Schritt 6 den Wert $i = n+1$ oder den Wert $j = m+1$ erreicht, beende den Algorithmus.

Man nennt diesen Algorithmus den Gauß'schen Eliminationsalgorithmus. In der Praxis schreibt man das lineare Gleichungssystem

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \ldots + a_{1m}x_m &= b_1, \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \ldots + a_{2m}x_m &= b_2, \\ &\ldots \\ a_{n1}x_1 + a_{n2}x_2 + a_{n3}x_3 + \ldots + a_{nm}x_m &= b_n. \end{aligned}$$

wie folgt auf:

$$\begin{array}{ccccc|c} a_{11} & a_{12} & a_{13} & \ldots & a_{1m} & b_1 \\ a_{21} & a_{22} & a_{23} & \ldots & a_{2m} & b_2 \\ \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & a_{n3} & \ldots & a_{nm} & b_1 \end{array}$$

Dies ist ausreichend, da alle Umformungen nur auf die Koeffizienten a_{ij} und b_i wirken.

Wir betrachten das obige Beispiel:

$$\begin{aligned} 3x_1 + 3x_2 + 9x_3 &= 36, \\ 2x_1 + 3x_2 + 7x_3 &= 29, \\ x_2 + 4x_3 &= 14. \end{aligned}$$

Aufgeschrieben ergibt dies:

$$\begin{array}{ccc|c} 3 & 3 & 9 & 36 \\ 2 & 3 & 7 & 29 \\ 0 & 1 & 4 & 14 \end{array}$$

Wir beginnen nun mit den Umformungen:

$$\begin{array}{ccc|r} 1 & 1 & 3 & 12 \\ 2 & 3 & 7 & 29 \\ 0 & 1 & 4 & 14 \\ & & & \\ 1 & 1 & 3 & 12 \\ 0 & 1 & 1 & 5 \\ 0 & 1 & 4 & 14 \end{array}$$

$$\begin{array}{ccc|c} 1 & 0 & 2 & 7 \\ 0 & 1 & 1 & 5 \\ 0 & 0 & 3 & 9 \\ & & & \\ 1 & 0 & 2 & 7 \\ 0 & 1 & 1 & 5 \\ 0 & 0 & 1 & 3 \\ & & & \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 3 \end{array}$$

Das lineare Gleichungssystem ist somit äquivalent zu dem Gleichungssystem

$$\begin{aligned} x_1 &= 1, \\ x_2 &= 2, \\ x_3 &= 3. \end{aligned}$$

Bemerkung: Durch Umbenennung der Variablen x_1, ..., x_m (dies ist gleichbedeutend mit Spaltenvertauschungen) lässt sich durch den Gauß'schen Eliminationsalgorithmus die folgende Form erreichen:

$$\begin{array}{ccccccc|c} 1 & 0 & ... & 0 & a_{1(r+1)} & ... & a_{1m} & b_1 \\ 0 & 1 & ... & 0 & a_{2(r+1)} & ... & a_{2m} & b_2 \\ ... & ... & ... & ... & ... & ... & ... & ... \\ 0 & 0 & ... & 1 & a_{r(r+1)} & ... & a_{rm} & b_r \\ 0 & 0 & ... & 0 & 0 & ... & 0 & b_{r+1} \\ ... & ... & ... & ... & ... & ... & ... & ... \\ 0 & 0 & ... & 0 & 0 & ... & 0 & b_n \end{array}$$

Man bezeichnet r als den Rank. Wir betrachten nun, unter welchen Bedingungen es eine eindeutige Lösung, keine Lösung oder mehrere Lösungen gibt:

- Ist $r = m$, so gibt es eine eindeutige Lösung.
- Ist $r < m$ und ist eine der Zahlen b_{r+1}, ..., b_n ungleich Null, so hat das lineare Gleichungssystem keine Lösung.
- Ist $r < m$ und $b_{r+1} = ... = b_n = 0$, so gibt es mehrere Lösungen.

Anwendungen: Lineare Unabhängigkeit von Vektoren. Zur Erinnerung: m Vektoren $\vec{v}_1$, $\vec{v}_2$, ..., $\vec{v}_m$ nennt man linear unabhängig, falls die Gleichung

$$\lambda_1 \vec{v}_1 + \lambda_2 \vec{v}_2 + ... + \lambda_m \vec{v}_m \quad = \quad \vec{0}$$

nur die Lösung $(\lambda_1, \lambda_2, ..., \lambda_m) = (0, 0, ..., 0)$ hat. Andernfalls nennt man sie linear abhängig. Ist der zugrundeliegende Vektorraum n-dimensional, so ergibt die obige Gleichung ausgeschrieben

in Komponenten n lineare Gleichungen mit m Unbekannten λ_1,..., λ_m. Man kann nun mit Hilfe des Gauß'schen Eliminationsalgorithmuses feststellen, ob die Vektoren linear abhängig sind.
Beispiel: Sei

$$\vec{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad \vec{v}_2 = \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix}, \quad \vec{v}_3 = \begin{pmatrix} -1 \\ -4 \\ -7 \end{pmatrix}.$$

Dies führt zu dem linearen Gleichungssystem

$$\begin{array}{rcl} \lambda_1 + 2\lambda_2 - \lambda_3 & = & 0, \\ \lambda_1 + 3\lambda_2 - 4\lambda_3 & = & 0, \\ \lambda_1 + 4\lambda_2 - 7\lambda_3 & = & 0. \end{array}$$

Wir formen dieses Gleichungssystem mit Hilfe des Gauß'schen Eliminationsalgorithmuses um:

$$\begin{array}{rrr|r} 1 & 2 & -1 & 0 \\ 1 & 3 & -4 & 0 \\ 1 & 4 & -7 & 0 \\ & & & \\ 1 & 2 & -1 & 0 \\ 0 & 1 & -3 & 0 \\ 0 & 2 & -6 & 0 \\ & & & \\ 1 & 0 & 5 & 0 \\ 0 & 1 & -3 & 0 \\ 0 & 0 & 0 & 0 \end{array}$$

Somit gibt es mehrere Lösungen

$$\lambda_1 = -5t, \quad \lambda_2 = 3t, \quad \lambda_3 = t, \quad t \in \mathbb{R}.$$

Die drei Vektoren sind linear abhängig.

4.6 Matrizen

Definition: Eine rechteckige Anordnung

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix}$$

von Elementen a_{ij} aus einem Körper nennt man Matrix. Die Elemente a_{ij} nennt man die Komponenten der Matrix. Eine Matrix mit n Zeilen und m Spalten bezeichnet man als $n \times m$-Matrix.

Eine Matrix bezeichnet man als quadratisch, falls $n = m$.
Eine Matrix bezeichnet man als Diagonalmatrix, falls sie quadratisch ist und $a_{ij} = 0$ für alle $i \neq j$.
Eine Matrix bezeichnet man als obere Dreiecksmatrix, falls sie quadratisch ist und $a_{ij} = 0$ für alle $i > j$.

Eine Multiplikation zwischen einer $n \times k$-Matrix A und einer $k \times m$-Matrix B ist wie folgt definiert: Das Ergebnis ist eine $n \times m$-Matrix C

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1k} \\ a_{21} & a_{22} & \dots & a_{2k} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nk} \end{pmatrix} \cdot \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1m} \\ b_{21} & b_{22} & \dots & b_{2m} \\ \dots & \dots & \dots & \dots \\ b_{k1} & b_{k2} & \dots & b_{km} \end{pmatrix} = \begin{pmatrix} c_{11} & c_{12} & \dots & c_{1m} \\ c_{21} & c_{22} & \dots & c_{2m} \\ \dots & \dots & \dots & \dots \\ c_{n1} & c_{n2} & \dots & c_{nm} \end{pmatrix},$$

wobei

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \dots + a_{ik}b_{kj}.$$

Ein n-dimensionaler Spaltenvektor kann als eine $n \times 1$-Matrix aufgefasst werden. Ebenso kann ein n-dimensionaler Zeilenvektor als eine $1 \times n$-Matrix betrachtet werden. Setzt man

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix}, \quad \vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \dots \\ x_m \end{pmatrix}, \quad \vec{b} = \begin{pmatrix} b_1 \\ b_2 \\ \dots \\ b_n \end{pmatrix},$$

so läßt sich das lineare Gleichungssystem

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 + \dots + a_{1m}x_m &= b_1, \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 + \dots + a_{2m}x_m &= b_2, \\ &\dots \\ a_{n1}x_1 + a_{n2}x_2 + a_{n3}x_3 + \dots + a_{nm}x_m &= b_n. \end{aligned}$$

auch wie folgt schreiben:

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

Seien A und B zwei $n \times m$-Matrizen. Eine Addition von zwei Matrizen mit gleicher Spalten- und Zeilenanzahl ist definiert durch

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} & \dots & b_{1m} \\ b_{21} & b_{22} & \dots & b_{2m} \\ \dots & \dots & \dots & \dots \\ b_{n1} & b_{n2} & \dots & b_{nm} \end{pmatrix} = \begin{pmatrix} a_{11}+b_{11} & a_{12}+b_{12} & \dots & a_{1m}+b_{1m} \\ a_{21}+b_{21} & a_{22}+b_{22} & \dots & a_{2m}+b_{2m} \\ \dots & \dots & \dots & \dots \\ a_{n1}+b_{n1} & a_{n2}+b_{n2} & \dots & a_{nm}+b_{nm} \end{pmatrix}$$

Eine Multiplikation mit Skalaren ist definiert durch

$$\lambda \cdot \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix} = \begin{pmatrix} \lambda a_{11} & \lambda a_{12} & \dots & \lambda a_{1m} \\ \lambda a_{21} & \lambda a_{22} & \dots & \lambda a_{2m} \\ \dots & \dots & \dots & \dots \\ \lambda a_{n1} & \lambda a_{n2} & \dots & \lambda a_{nm} \end{pmatrix}.$$

Mit dieser Addition und dieser skalaren Multiplikation bilden die $n \times m$-Matrizen einen Vektorraum. Die Dimension dieses Vektorraumes ist $n \cdot m$. Eine Basis ist gegegeben durch die Matrizen e_{ij},

$$e_{ij} = \begin{pmatrix} 0 & \dots & \dots & 0 & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & 0 & \dots & \dots & 0 \\ 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ 0 & \dots & \dots & 0 & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & 0 & \dots & \dots & 0 \end{pmatrix},$$

die nur in dem Eintrag in der i-ten Zeile und j-ten Spalte eine Eins haben, ansonsten nur Nullen.

4.7 Spuren und Determinanten von quadratischen Matrizen

Wir betrachten im folgenden die quadratischen $n \times n$-Matrizen. Seien A und B zwei $n \times n$-Matrizen. In diesem Fall ist das Matrixprodukt

$$A \cdot B$$

wieder eine $n \times n$-Matrix. Für $n \times n$-Matrizen ist die Matrizenmultiplikation also abgeschlossen. Das neutrale Element bezüglich der Matrizenmultiplikation ist offensichtlich die Einheitsmatrix

$$\mathbf{1} = \begin{pmatrix} 1 & 0 & \dots & \dots & 0 \\ 0 & 1 & & & 0 \\ \dots & & \dots & & \dots \\ 0 & & & 1 & 0 \\ 0 & \dots & \dots & 0 & 1 \end{pmatrix}.$$

Wir wollen nun untersuchen, unter welchen Bedingungen auch ein inverses Element existiert. Hierzu führen wir zunächst zwei Begriffe ein: Sei A eine $n \times n$-Matrix.

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}.$$

Unter der Spur (engl. "trace") einer quadratischen Matrix versteht man die Summe der Diagonalelemente:

$$\mathrm{Tr}\,A \;=\; \sum_{i=1}^{n} a_{ii} = a_{11} + a_{22} + \ldots + a_{nn}.$$

Einige Rechenregeln bezüglich der Spur von Matrizen (A, B seien $n \times n$-Matrizen, λ ein Skalar):

$$\begin{aligned} \mathrm{Tr}\,(A+B) &= \mathrm{Tr}\,A + \mathrm{Tr}\,B, \\ \mathrm{Tr}\,(\lambda \cdot A) &= \lambda\,\mathrm{Tr}\,A. \end{aligned}$$

Als Determinante einer quadratischen Matrix definiert man

$$\det A \;=\; \sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \ldots \sum_{i_n=1}^{n} \varepsilon_{i_1 i_2 \ldots i_n} a_{1 i_1} a_{2 i_2} \ldots a_{n i_n},$$

wobei $\varepsilon_{i_1 i_2 \ldots i_n}$ das total antisymmetrische Symbol in n Dimensionen ist

$$\varepsilon_{i_1 i_2 \ldots i_n} \;=\; \begin{cases} +1 & \text{für } (i_1 i_2 \ldots i_n) \text{ eine gerade Permutation von } (1,2,\ldots,n), \\ -1 & \text{für } (i_1 i_2 \ldots i_n) \text{ eine ungerade Permutation von } (1,2,\ldots,n), \\ 0 & \text{sonst.} \end{cases}$$

Für die Determinante existiert auch die folgende Schreibweise

$$\det A \;=\; \begin{vmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nn} \end{vmatrix}.$$

Sei $D = \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_n)$ eine Diagonalmatrix. Dann ist

$$\det D \;=\; \lambda_1 \cdot \lambda_2 \cdot \ldots \cdot \lambda_n.$$

Ein Verfahren zur Berechnung der Determinante ist der Laplace'sche Entwicklungssatz. Zu einer $n \times n$-Matrix A definieren wir zunächst eine $(n-1) \times (n-1)$-Matrix A_{ij}, die dadurch ensteht, daß man die i-te Zeile und die j-te Spalte der Matrix A entfernt. Der Laplace'sche Entwicklungssatz für die Entwicklung nach der i-ten Zeile lautet:

$$\det A \;=\; \sum_{j=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij}.$$

Natürlich kann äquivalent auch nach der j-ten Spalte entwickelt werden:

$$\det A \;=\; \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij}.$$

Dies erlaubt die rekursive Berechnung einer Determinante. Für eine 1×1-Matrix haben wir

$$|a_{11}| \;=\; a_{11}.$$

Rechenregeln für die Determinante:

$$\begin{aligned} \det(A \cdot B) &= (\det A) \cdot (\det B), \\ \det(\lambda \cdot A) &= \lambda^n \cdot \det A. \end{aligned}$$

Wenden wir uns wieder der Existenz eines inversen Elements bezüglich der Matrizenmultiplikation zu. Falls so ein Element existiert bezeichnen wir es mit A^{-1}. Es soll also gelten

$$A \cdot A^{-1} \;=\; \mathbf{1}.$$

Nehmen wir auf beiden Seiten die Determinante, so erhalten wir

$$\det A \cdot \det A^{-1} \;=\; \det \mathbf{1} = 1,$$

also falls $\det A \neq 0$

$$\det A^{-1} \;=\; \frac{1}{\det A}.$$

$\det A \neq 0$ ist eine notwendige Bedingung für die Existenz eines Inversen. Es läßt sich zeigen, daß dies auch hinreichend ist.

Satz: A^{-1} existiert genau dann, wenn $\det A \neq 0$.

Wir betrachten nun die Menge aller quadratischen $n \times n$-Matrizen mit der Eigenschaft $\det A \neq 0$. Wegen $\det(AB) = \det A \det B$ ist diese Menge abgeschlossen bezüglich der Matrizenmultiplikation. Wie gerade diskutiert wurde, existiert zu jeder Matrix auch ein Inverses. Diese Menge bildet daher bezüglich der Matrizenmultiplikation eine Gruppe, die man als

$$GL(n, \mathbb{R}), \quad \text{bzw.} \quad GL(n, \mathbb{C})$$

bezeichnet.

4.8 Berechnung der inversen Matrix

Sei A eine $n \times n$-Matrix mit $\det A \neq 0$. Gesucht ist eine $n \times n$-Matrix X

$$X \;=\; \begin{pmatrix} x_{11} & x_{12} & \dots & x_{1n} \\ x_{21} & x_{22} & \dots & x_{2n} \\ \dots & \dots & \dots & \dots \\ x_{n1} & x_{n2} & \dots & x_{nn} \end{pmatrix},$$

so daß

$$A \cdot X = \mathbf{1}$$

gilt. Wir multiplizieren die linke Seite aus und betrachten danach die j-te Spalte auf beiden Seiten:

$$\begin{aligned}
a_{11}x_{1j} + a_{12}x_{2j} + \ldots + a_{1n}x_{nj} &= 0, \\
\ldots &= 0 \\
a_{j1}x_{1j} + a_{j2}x_{2j} + \ldots + a_{jn}x_{nj} &= 1, \\
\ldots &= 0 \\
a_{n1}x_{1j} + a_{n2}x_{2j} + \ldots + a_{nn}x_{nj} &= 0.
\end{aligned}$$

Diese n Gleichungen bilden eine lineares Gleichungssystem für die Unbekannten x_{1j}, x_{2j}, ..., x_{nj}, welches mit Hilfe des Gaußschen Algorithmus gelöst werden kann. Da dies für jede Spalte j gilt, kann man so alle n^2 Unbekannten x_{ij} bestimmen. Da die Koeffizienten der linken Seite des linearen Gleichungssystems immer gleich sind, verfährt man in der Praxis wie folgt: Man schreibt die Gleichungen wie folgt an

$$\begin{array}{cccc|cccc}
a_{11} & a_{12} & \ldots & a_{1n} & 1 & 0 & \ldots & 0 \\
a_{21} & a_{22} & \ldots & a_{2n} & 0 & 1 & \ldots & 0 \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\
a_{n1} & a_{n2} & \ldots & a_{nn} & 0 & 0 & \ldots & 1
\end{array}$$

und bringt dieses Gleichungssystem mit Hilfe des Gauß'schen Algorithmus auf die Form

$$\begin{array}{cccc|cccc}
1 & 0 & \ldots & 0 & x_{11} & x_{12} & \ldots & x_{1n} \\
0 & 1 & \ldots & 0 & x_{21} & x_{22} & \ldots & x_{2n} \\
\ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots & \ldots \\
0 & 0 & \ldots & 1 & x_{n1} & x_{n2} & \ldots & x_{nn}
\end{array}$$

Die inverse Matrix ist dann gegeben durch

$$A^{-1} = \begin{pmatrix} x_{11} & x_{12} & \ldots & x_{1n} \\ x_{21} & x_{22} & \ldots & x_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ x_{n1} & x_{n2} & \ldots & x_{nn} \end{pmatrix},$$

Beispiel: Sei A die Matrix

$$A = \begin{pmatrix} 1 & 1 & 3 \\ 2 & 3 & 7 \\ 0 & 1 & 4 \end{pmatrix}.$$

Wir bestimmen nun die inverse Matrix:

$$\begin{array}{ccc|ccc}
1 & 1 & 3 & 1 & 0 & 0 \\
2 & 3 & 7 & 0 & 1 & 0 \\
0 & 1 & 4 & 0 & 0 & 1 \\
 & & & & & \\
1 & 1 & 3 & 1 & 0 & 0 \\
0 & 1 & 1 & -2 & 1 & 0 \\
0 & 1 & 4 & 0 & 0 & 1 \\
 & & & & & \\
1 & 0 & 2 & 3 & -1 & 0 \\
0 & 1 & 1 & -2 & 1 & 0 \\
0 & 0 & 3 & 2 & -1 & 1 \\
 & & & & & \\
1 & 0 & 2 & 3 & -1 & 0 \\
0 & 1 & 1 & -2 & 1 & 0 \\
0 & 0 & 1 & \frac{2}{3} & -\frac{1}{3} & \frac{1}{3} \\
 & & & & & \\
1 & 0 & 0 & \frac{5}{3} & -\frac{1}{3} & -\frac{2}{3} \\
0 & 1 & 0 & -\frac{8}{3} & \frac{4}{3} & -\frac{1}{3} \\
0 & 0 & 1 & \frac{2}{3} & -\frac{1}{3} & \frac{1}{3}
\end{array}$$

Somit ist A^{-1}

$$A^{-1} = \frac{1}{3}\begin{pmatrix} 5 & -1 & -2 \\ -8 & 4 & -1 \\ 2 & -1 & 1 \end{pmatrix}.$$

4.9 Eigenwerte und Eigenvektoren

Sei V ein n-dimensionaler Vektorraum und A eine $n \times n$-Matrix. Die Matrix A definiert eine Abbildung

$$\begin{array}{rcl} V & \rightarrow & V, \\ \vec{x} & \rightarrow & \vec{x}' = A \cdot \vec{x}. \end{array}$$

Diese Abbildung ist linear

$$A \cdot (c_1 \vec{x} + c_2 \vec{y}) = c_1 A \cdot \vec{x} + c_2 A \cdot \vec{y}.$$

Lineare Abbildungen zwischen zwei Vektorräumen werden auch als Homomorphismen bezeichnet. Da in unserem Fall die Abbildung von V nach V geht, spricht man auch von einem Epimorphismus.

Wir interessieren uns nun für Vektoren $\vec{x} \in V$, die auf ein Vielfaches ihrer selbst abgebildet werden.

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

In diesem Fall bezeichnet man λ als Eigenwert und $\vec{x}$ als Eigenvektor.

Beispiel: Falls $A = \mathrm{diag}(\lambda_1, \lambda_2, ..., \lambda_n)$ eine Diagonalmatrix ist, wobei alle λ_i paarweise verschieden sind, so sind die Basisvektoren $\vec{e}_j$ die Eigenvektoren. Der zu $\vec{e}_j$ zugehörige Eigenwert ist λ_j.

Sei nun eine $n \times n$-Matrix A gegeben. Wir suchen ein Verfahren zur Bestimmung der Eigenwerte und der Eigenvektoren. Hierzu definieren wir das charakteristische Polynom der Abbildung durch

$$\chi_A(\lambda) = \det\left(A - \lambda \mathbf{1}\right).$$

Dies ist ein Polynom vom Grad n in λ.

Satz: Die Nullstellen von χ_A sind die Eigenwerte von A.

Beispiel: Wir betrachten wieder

$$A = \begin{pmatrix} 1 & 1 & 3 \\ 2 & 3 & 7 \\ 0 & 1 & 4 \end{pmatrix}.$$

Man findet

$$\begin{aligned} \chi_A(\lambda) &= \begin{vmatrix} 1-\lambda & 1 & 3 \\ 2 & 3-\lambda & 7 \\ 0 & 1 & 4-\lambda \end{vmatrix} = -\lambda^3 + 8\lambda^2 - 10\lambda + 3 \\ &= -(\lambda - 1)\left(\lambda - \frac{1}{2}(7+\sqrt{37})\right)\left(\lambda - \frac{1}{2}(7-\sqrt{37})\right) \end{aligned}$$

Somit hat die Matrix A die Eigenwerte

$$1, \quad \frac{1}{2}(7+\sqrt{37}), \quad \frac{1}{2}(7-\sqrt{37}).$$

Sind die Eigenwerte bekannt, so bestimmt man im nächsten Schritt die zugehörigen Eigenvektoren. Sei λ ein (bekannter) Eigenwert der Matrix A. Gesucht werden Vektoren $\vec{x}$, so daß

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

bzw.

$$(A - \lambda \cdot \mathbf{1})\vec{x} = \vec{0}.$$

Dies ist wieder ein lineares Gleichungssystem, welches mit Hilfe des Gauß'schen Algorithmus gelöst werden kann.

Beispiel: Wir bestimmen zu der Matrix A aus dem obigen Beispiel die Eigenvektoren zum Eigenwert $\lambda = 1$:

$$A - \lambda \cdot \mathbf{1} \;=\; \begin{pmatrix} 0 & 1 & 3 \\ 2 & 2 & 7 \\ 0 & 1 & 3 \end{pmatrix}.$$

Der Gauß'sche Algorithmus liefert:

$$\begin{array}{ccc|c} 0 & 1 & 3 & 0 \\ 2 & 2 & 7 & 0 \\ 0 & 1 & 3 & 0 \\ & & & \\ 1 & 1 & \frac{7}{2} & 0 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 0 \\ & & & \\ 1 & 0 & \frac{1}{2} & 0 \\ 0 & 1 & 3 & 0 \\ 0 & 0 & 0 & 0 \end{array}$$

Somit sind alle Vektoren der Form

$$\begin{pmatrix} -\frac{1}{2}t \\ -3t \\ t \end{pmatrix}, \quad t \in \mathbb{R}$$

Eigenvektoren zum Eigenwert $\lambda = 1$. Man bezeichnet diesen Untervektorraum als Eigenraum.

Basiswechsel: Sei V ein n-dimensionaler Vektorraum mit einer Basis $\{\vec{v}_1, ..., \vec{v}_n\}$. Sei

$$\Phi \;:\; V \to V$$

eine Endomorphismus, gegeben in der Koordinatendarstellug der obigen Basis durch eine Matrix A:

$$\begin{pmatrix} x_1 \\ \dots \\ x_n \end{pmatrix} \;\to\; A \cdot \begin{pmatrix} x_1 \\ \dots \\ x_n \end{pmatrix}.$$

Dieser Endomorphismus bildet also den Basisvektor $\vec{v}_i$ auf

$$\vec{v}_j \;\to\; \sum_{i=1}^{n} \vec{v}_i a_{ij}$$

ab. (Das Bild von $\vec{v}_j$ ist die j-te Spalte von A.) Sei nun $\{\vec{w}_1, ..., \vec{w}_n\}$ eine weitere Basis des Vektorraumes. Wir können jeden Basisvektor $\vec{w}_j$ als Linearkombination der $\vec{v}_i$'s schreiben:

$$\vec{w}_j \;=\; \sum_{i=1}^{n} \vec{v}_i m_{ij}.$$

Dies definiert eine $n \times n$-Matrix M

$$M = \left(m_{ij}\right).$$

Die Koordinatendarstellungen eines beliebigen Vektors $\vec{z}$ bezüglich der beiden Basen

$$\vec{z} = x_1\vec{v}_1 + \ldots x_n\vec{v}_n = y_1\vec{w}_1 + \ldots y_n\vec{w}_n$$

transformieren sich wie

$$\begin{pmatrix} x_1 \\ \ldots \\ x_n \end{pmatrix} = M \begin{pmatrix} y_1 \\ \ldots \\ y_n \end{pmatrix}.$$

Da umgekehrt auch die Basisvektoren $\vec{v}_i$ durch die Basisvektoren $\vec{w}_j$ ausgedrückt werden können, ist M invertierbar, also

$$M \in GL(n,K), \qquad K = \mathbb{R}, \mathbb{C},$$

und

$$\vec{v}_j = \sum_{i=1}^{n} \vec{w}_i m_{ij}^{-1}.$$

In der Basis $\{\vec{w}_1, \ldots, \vec{w}_n\}$ wird der Endomorphismus Φ durch eine Matrix B beschrieben

$$\vec{w}_j \to \sum_{i=1}^{n} \vec{w}_i b_{ij}.$$

Wir bestimmen nun den Zusammenhang von A mit B: Es gilt

$$\vec{w}_j = \sum_{k=1}^{n} \vec{v}_k m_{kj} \quad \to \quad \sum_{k=1}^{n}\sum_{l=1}^{n} \vec{v}_l a_{lk} m_{kj} = \sum_{i=1}^{n}\sum_{k=1}^{n}\sum_{l=1}^{n} \vec{w}_i m_{il}^{-1} a_{lk} m_{kj},$$

und somit ist

$$B = M^{-1}AM.$$

Wir betrachten nun wieder den Endomorphismus Φ, der in der Basis $\{\vec{v}_1, \ldots, \vec{v}_n\}$ durch die Matrix A gegeben ist. Wir können nun die Frage stellen, ob es eine neue Basis $\{\vec{w}_1, \ldots, \vec{w}_n\}$ gibt, so daß der Endomorphismus in dieser neuen Basis durch eine besonders einfache Matrix B gegeben ist. Eine besonders einfache Matrix ist eine Diagonalmatrix.

Definition: Einen Endomorphismus nennt man **diagonalisierbar**, falls es eine Basis $\{\vec{w}_1, \ldots, \vec{w}_n\}$ gibt, so daß in dieser neuen Basis die Abbildung durch eine Diagonalmatrix gegeben ist. In diesem Fall ist es unmittelbar einsichtig, daß die $\vec{w}_i$'s Eigenvektoren sind. Da auch die Umkehrung gilt, hat man folgenden Satz:

Eine Matrix A läßt sich genau dann durch einen Basiswechsel auf Diagonalgestalt bringen

$$D = M^{-1}AM, \quad D \text{ Diagonalmatrix}, M \in GL(n,K),$$

falls es eine Basis von V aus Eigenvektoren von A gibt.

Die Berechnung der Eigenwerte und der Eigenvektoren liefert ein konstruktives Verfahren zur Diagonalisierung einer Matrix. Seien $\vec{w}_1$, ..., $\vec{w}_n$ die Eigenvektoren. Dann ergeben diese Vektoren als Spaltenvektoren betrachtet die Matrix M:

$$M = (\vec{w}_1, \ldots, \vec{w}_n)\,.$$

Beispiel: Die Matrix

$$A = \begin{pmatrix} 0 & -1 & 1 \\ -3 & -2 & 3 \\ -2 & -2 & 3 \end{pmatrix}$$

hat hat charakteristische Polynom

$$\chi_A(\lambda) = -(\lambda-1)^2(\lambda+1)\,.$$

Der Eigenraum zum Eigenwert $\lambda = 1$ wird von den Vektoren

$$\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \qquad \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

aufgespannt. Der Vektor

$$\begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix}$$

ist Eigenvektor zum Eigenwert $\lambda = -1$. Somit ist M gegeben durch

$$M = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 3 \\ 1 & 1 & 2 \end{pmatrix}, \qquad M^{-1} = \frac{1}{2}\begin{pmatrix} 1 & -1 & 1 \\ -3 & -1 & 3 \\ 1 & 1 & -1 \end{pmatrix},$$

und

$$D = M^{-1}AM = \mathrm{diag}(1,1,-1).$$

Wir wollen noch genauer untersuchen, unter welchen Bedingungen eine Matrix diagonalisiert werden kann. Bei der Bestimmung der Nullstellen des charakteristischen Polynoms kann der Fall auftreten, daß das Polynom nicht vollständig in Linearfaktoren zerfällt. Dies ist zum Beispiel bei der Matrix

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad \chi_A(\lambda) = \lambda^2 + 1,$$

über dem Körper der reellen Zahlen der Fall, da

$$\lambda^2 = -1$$

keine Lösung in $\mathbb{R}$ hat.
Bemerkung: Die komplexen Zahlen $\mathbb{C}$ sind algebraisch abgeschlossen, d.h. über $\mathbb{C}$ zerfällt jedes Polynom in Linearfaktoren. Im obigen Beispiel findet man

$$\lambda^2 + 1 = (\lambda - i)(\lambda + i).$$

Eine zweite Komplikation kann auftreten, falls das charakteristische Polynom Nullstellen mit höherer Multiplizität hat. Ist die Multiplizität einer Nullstelle größer als die Dimension des zugehörigen Eigenraumes, so ist die Matrix nicht diagonalsierbar. Beispiel: Die Matrix

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \chi_A(\lambda) = \lambda^2,$$

hat den doppelten Eigenwert $\lambda = 0$. Der zugehörige Eigenraum ist aber nur ein-dimensional:

$$\begin{pmatrix} t \\ 0 \end{pmatrix}, \quad t \in \mathbb{R}.$$

A ist nicht diagonalisierbar. Wäre A diagonalisierbar, so wäre die zugehörige Diagonalmatrix die Nullmatrix, was zu einem Widerspruch führt.

Zusammenfassend erhalten wir die folgende Aussage:
Satz: Eine Matrix A mit Einträgen aus dem Körper K ist genau dann diagonalisierbar, falls das charakteristische Polynom in Linearfaktoren zerfällt, und die Dimension des Eigenraumes zum Eigenwert λ_i gleich der Multiplizität der Nullstelle λ_i des charakteristischen Polynoms ist.

Es läßt sich zeigen, daß die Dimension der Eigenräume stets gr̈ßer oder gleich Eins ist. Daher gilt im besonderen: Falls das charakteristische Polynom in Linearfaktoren zerfällt und alle Nullstellen paarweise verschieden sind, so ist die Matrix A diagonalisierbar.

4.10 Das Schmidt'sche Orthonormierungsverfahren

Wir betrachten einen euklidischen bzw. unitären Vektorraum V der Dimension n, d.h einen Vektorraum über den reellen Zahlen mit einer positiv definiten symmetrischen Bilinearform (euklidischer Vektorraum) bzw. einen Vektorraum über den komplexen Zahlen mit eine positive definiten Hermitischen Form. Gegeben seien k lineare unabhängige Vektoren $\{\vec{v}_1, ..., \vec{v}_k\}$. Diese

Vektoren spannen einen Untervektorraum W auf (und bilden eine Basis von W). Gesucht ist nun eine Orthonormalbasis $\{\vec{n}_1, ..., \vec{n}_k\}$ von W. Unter einer Orthonormalbasis versteht man

$$|\vec{n}_i| = 1, \qquad \vec{n}_i \cdot \vec{n}_j = 0, \quad \text{für} \quad i \neq j.$$

Hierzu geht man wie folgt vor: Man setzt am Anfang

$$\vec{n}_1 \;\; = \;\; \frac{1}{|\vec{v}_1|}\vec{v}_1.$$

Seien nun $\vec{n}_1$, ..., $\vec{n}_{j-1}$ schon konstruiert. Dann erhält man $\vec{n}_j$ wie folgt: Man definiert zunächst

$$\vec{w} \;\; = \;\; \vec{v}_j - \sum_{i=1}^{j-1} \left(\vec{v}_j \cdot \vec{n}_i\right) \vec{n}_i.$$

Durch diese Konstruktion steht $\vec{w}$ senkrecht zu allen bisher konstruierten Vektoren:

$$\vec{w} \cdot \vec{n}_i = 0, \quad \text{für} \quad i \in \{1, 2, ..., j-1\}.$$

Wir normieren $\vec{w}$ und erhalten $\vec{n}_j$:

$$\vec{n}_j \;\; = \;\; \frac{1}{|\vec{w}|}\vec{w}.$$

Eine typische Anwendung für das Schmidt'sche Orthonormierungsverfahren ist die Konstruktion einer Orthonormalbasis für einen Eigenraum der Dimension $D > 1$.

5 Analysis

5.1 Folgen

Unter einer Folge (a_n) reeller Zahlen versteht man eine Abbildung $\mathbb{N} \to \mathbb{R}$. Jedem $n \in \mathbb{N}$ wird also ein $a_n \in \mathbb{R}$ zugeordnet.

Eine Folge (a_n) heißt **konvergent** gegen $a \in \mathbb{R}$, falls es zu jedem $\varepsilon > 0$ eine natürliche Zahl N gibt, so daß

$$|a_n - a| \quad < \quad \varepsilon, \quad \forall n \geq N.$$

In diesem Fall schreibt man

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

In anderen Worten liegen für eine konvergente Folge ab einem bestimmten N alle Folgenglieder im Intervall $]a-\varepsilon, a+\varepsilon[$.

Eine Folge nennt man **divergent**, wenn sie gegen keine reelle Zahl konvergiert. Beispiele für divergente Folgen sind

$$\begin{aligned} a_n &= n, \\ b_n &= \begin{cases} 1 & n \text{ gerade} \\ -1 & n \text{ ungerade} \end{cases} \end{aligned}$$

Eine Folge heißt nach oben (bzw. unten) beschränkt, falls es ein $c \in \mathbb{R}$ gibt, so daß $a_n \leq c$ (bzw. $a_n \geq c$) für alle $n \in \mathbb{N}$. Die Folge heißt beschränkt, wenn sie nach oben und unten beschränkt ist. Jede konvergente Folge beschränkt.
Bemerkung: Die Umkehrung gilt nicht, eine beschränkte Folge ist nicht notwendiger Weise konvergent, siehe obiges Beispiel mit der Folge (b_n).

Seien (a_n) und (b_n) zwei konvergente Folgen mit den Grenzwerten

$$\lim_{n\to\infty} a_n = a, \qquad \lim_{n\to\infty} b_n = b.$$

Dann sind auch die Folgen $(a_n + b_n)$, $(a_n - b_n)$, (λa_n), $(a_n b_n)$ konvergent ($\lambda \in \mathbb{R}$) und es gilt

$$\begin{aligned} \lim_{n\to\infty} (a_n + b_n) &= a + b, \\ \lim_{n\to\infty} (a_n - b_n) &= a - b, \\ \lim_{n\to\infty} (\lambda a_n) &= \lambda a, \\ \lim_{n\to\infty} (a_n \cdot b_n) &= a \cdot b. \end{aligned}$$

Ist weiter $b \neq 0$, so gibt es ein $N \in \mathbb{N}$, so daß $b_n \neq 0$ für alle $n \geq N$ und wir können die Folge

$$\left(\frac{a_n}{b_n}\right)_{n \geq N}$$

betrachten. Es gilt

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

Seien (a_n) und (b_n) zwei konvergente Folgen mit $a_n \leq b_n$ für alle n. Dann gilt auch

$$\lim_{n\to\infty} a_n \leq \lim_{n\to\infty} b_n.$$

Bemerkung: Aus $a_n < b_n$ folgt nicht $\lim a_n < \lim b_n$, wie das Beispiel $a_n = 0$ und $b_n = 1/n$ zeigt.

5.2 Reihen

Sei (a_n) eine Folge reeller Zahlen. Man betrachtet nun die Folge der Partialsummen

$$\sum_{j=1}^{n} a_j.$$

Als unendliche Reihe bezeichnet man nun die Folge dieser Partialsummen. Man schreibt

$$\sum_{j=1}^{\infty} a_j = \lim_{n\to\infty} \sum_{j=1}^{n} a_j.$$

Eine unendliche Reihe heißt konvergent, wenn die Folge der Partialsummen konvergiert.

Eine unendliche Reihe heißt **absolut konvergent**, falls die Reihe

$$\sum_{j=1}^{\infty} |a_j|$$

konvergent ist. Wir interessieren uns nun für Kriterien, die eine Konvergenz einer unendlichen Reihe garantieren.

Konvergenzkriterium von Cauchy: Die Reihe $\sum\limits_{j=1}^{\infty} a_j$ konvergiert genau dann, wenn zu jedem $\varepsilon > 0$ ein $N \in \mathbb{N}$ existiert, so daß

$$\left| \sum_{j=m}^{n} a_j \right| < \varepsilon, \quad \forall n \geq m \geq N.$$

Satz: Eine notwendige (aber nicht hinreichende) Bedingung für die Konvergenz einer Reihe ist, daß

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

Satz: Eine Reihe $\sum_{j=1}^{\infty} a_j$ mit $a_j \geq 0$ konvergiert genau dann, wenn die Folge der Partialsummen beschränkt ist.

Konvergenzkriterium von Leibniz für alternierende Reihen: Sei (a_n) eine monoton fallende Folge nicht-negativer Zahlen mit $\lim a_n = 0$. Dann konvergiert die Reihe

$$\sum_{j=1}^{\infty} (-1)^j a_j.$$

Satz: Eine absolut konvergente Reihe konvergiert auch im gewöhnlichen Sinne.

Majorantenkriterium: Sei $\sum_{j=1}^{\infty} c_j$ eine konvergente Reihe mit lauter nicht-negativen Gliedern und (a_n) eine Folge mit $|a_n| \leq c_n$. Dann konvergiert die Reihe

$$\sum_{j=1}^{\infty} a_j$$

absolut. Man nennt $\sum_{j=1}^{\infty} c_j$ eine Majorante von $\sum_{j=1}^{\infty} a_j$.

Quotientenkriterium: Sei $\sum_{j=1}^{\infty} a_j$ eine Reihe mit $a_n \neq 0$ für alle n und x eine reelle Zahl $0 < x < 1$, so daß

$$\left| \frac{a_{n+1}}{a_n} \right| \leq x, \quad \forall n \geq N.$$

Dann konvergiert die Reihe absolut.

Cauchy-Produkt von Reihen: Seien $\sum_{j=1}^{\infty} a_j$ und $\sum_{j=1}^{\infty} b_j$ zwei absolut konvergente Reihen. Für $n \in \mathbb{N}$ setzen wir

$$c_n = \sum_{j=1}^{n-1} a_j b_{n-j}.$$

Dann ist auch die Reihe $\sum_{j=1}^{\infty} c_j$ absolut konvergent und es gilt

$$\sum_{j=1}^{\infty} c_j = \left(\sum_{j=1}^{\infty} a_j \right) \left(\sum_{j=1}^{\infty} b_j \right).$$

Bemerkung: Die absolute Konvergenz ist wesentlich für die Gültigkeit des Satzes ! Im Allgemeinen gilt, daß Umordnungen innerhalb einer Reihe nur erlaubt sind, falls die Reihe absolut konvergiert.

Wir geben noch einige wichtige Reihen an:

$$\begin{aligned}
\exp x &= \sum_{j=0}^{\infty} \frac{x^j}{j!}, \\
-\ln(1-x) &= \sum_{j=1}^{\infty} \frac{x^j}{j}, \quad |x| < 1, \\
\sin x &= \sum_{j=0}^{\infty} (-1)^j \frac{x^{2j+1}}{(2j+1)!}, \\
\cos x &= \sum_{j=0}^{\infty} (-1)^j \frac{x^{2j}}{(2j)!}, \\
\sinh x &= \sum_{j=0}^{\infty} \frac{x^{2j+1}}{(2j+1)!}, \\
\cosh x &= \sum_{j=0}^{\infty} \frac{x^{2j}}{(2j)!},
\end{aligned}$$

sinh und cosh bezeichnet man als Sinus Hyperbolicus bzw. Kosinus Hyperbolicus. Mit Ausnahme der Reihe für $\ln(1-x)$ konvergieren alle Reihen absolut für alle Werte von x. Man sagt die Reihen haben einen unendlichen **Konvergenzradius**. Die Reihe für $\ln(1-x)$ konvergiert absolut für $|x| < 1$. Somit hat diese Reihe den Konvergenzradius 1. Man spricht von einem Konvergenzradius, da die obigen Reihen auch definiert sind, wenn man die reelle Variable x durch eine komplexe Variable z ersetzt.

Wir betrachten noch $\exp(ix)$:

$$\begin{aligned}
\exp(ix) &= \sum_{j=0}^{\infty} \frac{i^j x^j}{j!} = \sum_{j=0}^{\infty} i^{2j} \frac{x^{2j}}{(2j)!} + \sum_{j=1}^{\infty} i^{2j+1} \frac{x^{2j+1}}{(2j+1)!} \\
&= \sum_{j=0}^{\infty} (-1)^j \frac{x^{2j}}{(2j)!} + i \sum_{j=1}^{\infty} (-1)^j \frac{x^{2j+1}}{(2j+1)!} = \cos x + i \sin x.
\end{aligned}$$

Die Reihendarstellung liefert also einen einfachen Beweis der Formel:

$$\exp(ix) = \cos x + i \sin x.$$

Ebenso findet man

$$\exp x = \cosh x + \sinh x.$$

Man beachte daß für die Umordnung der Reihen die absolute Konvergenz notwendig ist. Man kann die trigometrischen und die hyperbolischen Funktionen auch durch die Exponentialfunktion ausdrücken:

$$\cos x = \frac{1}{2}\left(e^{ix} + e^{-ix}\right), \qquad \sin x = \frac{1}{2i}\left(e^{ix} - e^{-ix}\right),$$
$$\cosh x = \frac{1}{2}\left(e^{x} + e^{-x}\right), \qquad \sinh x = \frac{1}{2}\left(e^{x} - e^{-x}\right).$$

5.3 Funktionen und Stetigkeit

Seien D und W Teilmengen von $\mathbb{R}$. Unter einer reellwertigen Funktion auf D versteht man eine Abbildung

$$\begin{aligned} f \quad : \quad & D \to W, \\ & x \to y = f(x). \end{aligned}$$

Man nennt D den Definitionsbereich und W den Wertebereich der Funktion. Eine Funktion f ordnet jedem $x \in D$ ein $y \in W$ zu. Gibt es zu jedem $y \in W$ genau ein $x \in D$ mit $y = f(x)$, so ist die Funktion f umkehrbar. In diesem Fall bezeichnet man mit f^{-1} die Umkehrfunktion:

$$\begin{aligned} f^{-1} \quad : \quad & W \to D, \\ & y \to x = f^{-1}(y). \end{aligned}$$

Man sagt eine Funktion hat im Punkte a den Grenzwert c, falls es mindestens eine Folge $x_n \in D$ mit $\lim x_n = a$ gibt und es zu jedem $\varepsilon > 0$ ein $\delta_\varepsilon > 0$ gibt, so daß

$$|f(x) - c| \quad < \quad \varepsilon, \quad \forall |x-a| < \delta_\varepsilon \quad \text{und} \quad x \in D.$$

In diesem Fall schreibt man

$$\lim_{x \to a} f(x) \quad = \quad c.$$

Bemerkung: Es wird nicht vorausgesetzt, daß $a \in D$ liegt. Diese Definition macht auch Sinn, falls D ein offenes Intervall ist und der Grenzwert an den Intervallgrenzen betrachtet wird.

Sei nun $a \in D$. Man bezeichnet eine Funktion als stetig im Punkte a, falls

$$\lim_{x \to a} f(x) \quad = \quad f(a)$$

gilt.
Beispiel: Wir betrachten die Heaviside-Funktion, definiert durch

$$\Theta(x) \quad = \quad \left\{ \begin{array}{ll} 1, & x > 0, \\ 0, & x \leq 0. \end{array} \right.$$

Für diese Funktion gilt $\Theta(0) = 0$, aber

$$\lim_{x \to 0+} \Theta(x) = 1.$$

Die Heaviside-Funktion ist im Punkte 0 nicht stetig.

Beispiele von Funktionen, die auf ganz $\mathbb{R}$ stetig sind, sind Polynomfunktionen, $\exp x$, $\sin x$, $\cos x$, $\sinh x$, $\cosh x$.

Satz: Seien $f, g : D \to \mathbb{R}$ Funktionen, die in a stetig sind und sei $\lambda \in \mathbb{R}$. Dann sind auch die Funktionen

$$\begin{aligned} f + g &: D \to \mathbb{R}, \\ \lambda \cdot f &: D \to \mathbb{R}, \\ f \cdot g &: D \to \mathbb{R} \end{aligned}$$

im Punkte a stetig. Ist ferner $g(a) \neq 0$, so ist auch die Funktion

$$\frac{f}{g} : D' \to \mathbb{R}$$

in a stetig, wobei $D' = \{x \in D : g(x) \neq 0\}$.

Wir definieren noch den Begriff der **gleichmäßigen Stetigkeit**: Eine Funktion $f : D \to \mathbb{R}$ heißt in D gleichmäßig stetig, falls es zu jedem $\varepsilon > 0$ ein $\delta > 0$ gibt, so daß

$$|f(x) - f(y)| < \varepsilon \qquad \forall \; |x - y| < \delta.$$

Jede Funktion, die auf D gleichmässig stetig ist, ist auch in jedem Punkte aus D stetig im herkömmlichen Sinne. Die Umkehrung gilt jedoch nicht. Ist eine Funktion in jedem Punkte $x \in D$ stetig im herkömmlichen Sinne, so genügt es für ein vorgegebenes ε für jeden Punkt ein δ_x zu finden. Dieses δ_x darf mit x variieren. Für die gleichmäßige Stetigkeit wird dagegen gefordert, daß δ von x unabhängig ist.

5.3.1 Rationale Funktionen

Seien $p(x)$ und $q(x)$ Polynomfunktionen. Unter einer rationalen Funktion versteht man eine Funktion

$$R(x) = \frac{p(x)}{q(x)}.$$

Der Definitionsbereich einer rationalen Funktion ist gegeben durch $D = \{x \in \mathbb{R}, q(x) \neq 0\}$. Aufgrund des obigen Satzes ist eine rationale Funktion in ihrem Definitionsbereich stetig.

Rationale Funktionen können in **Partialbrüche** zerlegt werden. Ist

$$p(x) = p_n x^n + p_{n-1} x^{n-1} + \dots + p_1 x + p_0,$$
$$q(x) = q_m x^m + q_{m-1} x^{m-1} + \dots + q_1 x + q_0,$$

und ist ausserdem die Faktorisierung des Nennerpolynoms bekannt

$$q(x) = c \prod_{j=1}^{r} \left(x - x_j\right)^{\lambda_j},$$

wobei λ_j die Multiziplität der Nullstelle x_j angibt (wobei $\lambda_1 + \dots + \lambda_r = m$ ist), so läßt sich die rationale Funktion schreiben als

$$\begin{aligned} R(x) &= \frac{p(x)}{q(x)} \\ &= P(x) + \sum_{j=1}^{r} \sum_{k=1}^{\lambda_j} \frac{a_{jk}}{(x - x_j)^k}, \end{aligned}$$

wobei $P(x)$ ein Polynom vom Grad $\deg p(x) - \deg q(x)$ ist und $a_{jk} \in \mathbb{R}$.

Berechnung von $P(x)$ und der Konstanten a_{jk}: $P(x)$ bestimmt sich durch Polynomdivision mit Rest. Wir betrachten als Beispiel die rationale Funktion

$$\frac{x^4 + 3x^3 - 12x^2 - 3x + 18}{(x-2)^2(x+2)}$$

Für das Nennerpolynom haben wir

$$(x-2)^2(x+2) \quad = \quad x^3 - 2x^2 - 4x + 8.$$

Polynomdivision mit Rest liefert

$$\begin{array}{rcl} (x^4 + 3x^3 - 12x^2 - 3x + 18) & : \quad (x^3 - 2x^2 - 4x + 8) \quad = & x + 5 + (2x^2 + 9x - 22)/(x^3 - 2x^2 - 4x + 8) \\ -(x^4 - 2x^3 - 4x^2 + 8x) & & \\ ----------- & & \\ 5x^3 - 8x^2 - 11x + 18 & & \\ -(5x^3 - 10x^2 - 20x + 40) & & \\ ----------- & & \\ 2x^2 + 9x - 22 & & \end{array}$$

Somit ist also $P(x) = x + 5$. Für den Rest verwendet man den Ansatz

$$\frac{2x^2 + 9x - 22}{x^3 - 2x^2 - 4x + 8} \quad = \quad \frac{a_{12}}{(x-2)^2} + \frac{a_{11}}{x-2} + \frac{a_{21}}{x+2}.$$

Man bringt die rechte Seite auf den Hauptnenner

$$\frac{a_{12}}{(x-2)^2}+\frac{a_{11}}{x-2}+\frac{a_{21}}{x+2} \quad = \quad \frac{(a_{11}+a_{21})x^2+(a_{12}-4a_{21})x+(2a_{12}-4a_{11}+4a_{21})}{x^3-2x^2-4x+8}$$

Koeffizientenvergleich liefert ein lineares Gleichungssystem:

$$\begin{aligned} a_{11}+a_{21} &= 2, \\ a_{12}-4a_{21} &= 9, \\ 2a_{12}-4a_{11}+4a_{21} &= -22, \end{aligned}$$

dessen Lösung durch

$$a_{12}=1, \quad a_{11}=4, \quad a_{21}=-2$$

gegeben ist. Somit erhalten wir das Ergebnis

$$\frac{x^4+3x^3-12x^2-3x+18}{(x-2)^2(x+2)} \quad = \quad x+5+\frac{1}{(x-2)^2}+\frac{4}{x-2}-\frac{2}{x+2}.$$

Bemerkung: Die Koeffizienten der Partialbrüche mit der höchsten Potenz einer Nullstelle lassen sich einfacher bestimmen, indem man im Ansatz mit $(x-x_j)^{\lambda_j}$ multipliziert und dann $x=x_j$ setzt. In unserem Beispiel lassen sich so a_{12} und a_{21} bestimmen:

$$\begin{aligned} a_{12} &= \left.\frac{2x^2+9x-22}{(x-2)^2(x+2)}(x-2)^2\right|_{x=2} = \left.\frac{2x^2+9x-22}{x+2}\right|_{x=2} = \frac{8+18-22}{4} = 1, \\ a_{21} &= \left.\frac{2x^2+9x-22}{(x-2)^2(x+2)}(x+2)\right|_{x=-2} = \left.\frac{2x^2+9x-22}{(x-2)^2}\right|_{x=-2} = \frac{8-18-22}{16} = -2. \end{aligned}$$

5.3.2 Trigonometrische Funktionen

Neben den Winkelfunktionen Sinus und Kosinus

$$\cos x = \frac{1}{2}\left(e^{ix}+e^{-ix}\right), \qquad \sin x = \frac{1}{2i}\left(e^{ix}-e^{-ix}\right),$$

gibt es weitere trigonometrische Funktionen:

$$\begin{aligned} \tan x &= \frac{\sin x}{\cos x}, \quad \text{Tangens} \\ \cot x &= \frac{\cos x}{\sin x}, \quad \text{Kotangens} \\ \sec x &= \frac{1}{\cos x}, \quad \text{Sekans} \\ \csc x &= \frac{1}{\sin x}, \quad \text{Kosekans} \end{aligned}$$

Die Umkehrfunktionen werden mit arcsin, arccos, arctan, etc. bezeichnet:

$$\begin{aligned}
\arcsin(x) &= \sin^{-1}(x), \quad \text{Arkussinus} \\
\arccos(x) &= \cos^{-1}(x), \quad \text{Arkuskosinus} \\
\arctan(x) &= \tan^{-1}(x), \quad \text{Arkustangens}
\end{aligned}$$

Diese Umkehrfunktionen lassen sich durch den Logarithmus ausdrücken:

$$\begin{aligned}
\arcsin(x) &= \frac{1}{i} \ln\left(ix + \sqrt{1-x^2}\right), \\
\arccos(x) &= \frac{1}{i} \ln\left(x + i\sqrt{1-x^2}\right), \\
\arctan(x) &= \frac{1}{2i} \ln\left(\frac{1+ix}{1-ix}\right).
\end{aligned}$$

5.3.3 Hyperbolische Funktionen

Neben den bereits eingeführten hyperbolischen Funktionen

$$\cosh x = \frac{1}{2}\left(e^x + e^{-x}\right), \qquad \sinh x = \frac{1}{2}\left(e^x - e^{-x}\right),$$

definiert man auch

$$\tanh x = \frac{\sinh x}{\cosh x}.$$

Bemerkung: Für sinh und cosh gilt

$$\cosh^2 x - \sinh^2 x = 1.$$

Die inversen Funktionen werden als Areafunktionen bezeichnet:

$$\begin{aligned}
\operatorname{arsinh}(x) &= \sinh^{-1}(x), \quad \text{Areasinus Hyperbolicus} \\
\operatorname{arcosh}(x) &= \cosh^{-1}(x), \quad \text{Areakosinus Hyperbolicus} \\
\operatorname{artanh}(x) &= \tanh^{-1}(x), \quad \text{Areatangens Hyperbolicus}
\end{aligned}$$

Diese Umkehrfunktionen lassen sich ebenfalls durch den Logarithmus ausdrücken:

$$\begin{aligned}
\operatorname{arsinh}(x) &= \ln\left(x + \sqrt{x^2+1}\right), \\
\operatorname{arcosh}(x) &= \ln\left(x + \sqrt{x^2-1}\right), \\
\operatorname{artanh}(x) &= \frac{1}{2} \ln\left(\frac{1+x}{1-x}\right).
\end{aligned}$$

Für den Zusammenhang zwischen den trigonometrischen Funktionen und den hyperbolischen Funktionen gilt

$$\begin{aligned}
\sin x &= \frac{1}{i}\sinh(ix),\\
\cos x &= \cosh(ix),\\
\tan x &= \frac{1}{i}\tanh(ix),\\
\arcsin(x) &= \frac{1}{i}\operatorname{arsinh}(ix),\\
\arctan(x) &= \frac{1}{i}\operatorname{artanh}(ix).
\end{aligned}$$

5.4 Differentialrechnung

Sei $D \subset \mathbb{R}$ und $f : D \longrightarrow \mathbb{R}$ eine Funktion. f nennt man im Punkte $x \in D$ **differenzierbar**, falls es mindestens eine Folge $(\xi_n) \in D \backslash x$ gibt, mit $\lim\limits_{n\to\infty} \xi_n = x$ und der Grenzwert

$$f'(x) = \lim_{\xi\to x} \frac{f(\xi) - f(x)}{\xi - x}$$

existiert. Man schreibt auch

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

Seien $f, g : D \longrightarrow \mathbb{R}$ in $x \in D$ differenzierbare Funktionen und $\lambda \in \mathbb{R}$. Dann sind auch die Funktionen $f + g$, $f \cdot g$ und λf in x differenzierbar und es gilt

$$\begin{aligned}
(f+g)'(x) &= f'(x) + g'(x),\\
(f\cdot g)'(x) &= f'(x)g(x) + f(x)g'(x), \quad \text{(Produktregel)}\\
(\lambda f)'(x) &= \lambda f'(x).
\end{aligned}$$

Dies folgt unmittelbar aus den Rechenregeln für Grenzwerte von Folgen. So ist zum Beispiel

$$\begin{aligned}
(f\cdot g)'(x) &= \lim_{h\to 0} \frac{f(x+h)g(x+h) - f(x)g(x)}{h}\\
&= \lim_{h\to 0} \frac{1}{h}\left[f(x+h)g(x+h) - f(x+h)g(x) + f(x+h)g(x) - f(x)g(x)\right]\\
&= \lim_{h\to 0} f(x+h)\frac{g(x+h) - g(x)}{h} + \lim_{h\to 0} \frac{f(x+h) - f(x)}{h} g(x)\\
&= f(x)g'(x) + f'(x)g(x).
\end{aligned}$$

Quotientenregel: Ist $g(x) \neq 0$ für alle $x \in D$, so ist auch die Funktion f/g in x differenzierbar und es gilt

$$\left(\frac{f}{g}\right)'(x) = \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2}.$$

Ableitung der Umkehrfunktion: Sei $D \subset \mathbb{R}$ ein abgeschlossenes Intervall, $f : D \to W$ eine stetige, steng monotone Funktion und $f^{-1} : W \to D$ die Umkehrfunktion. Ist f im Punkt $x \in D$ differenzierbar und ist $f'(x) \neq 0$, so ist f^{-1} im Punkt $y = f(x)$ differenzierbar und es gilt

$$\left(f^{-1}\right)'(y) = \frac{1}{f'(x)} = \frac{1}{f'(f^{-1}(y))}.$$

Kettenregel: Seien $f : D_1 \to W_1$ und $g : D_2 \to W_2$ Funktionen mit $W_1 \subset D_2$. Falls f im Punkte $x \in D_1$ differenzierbar ist und g im Punkte $y = f(x) \in D_2$ differenzierbar ist, so ist die zusammengesetzte Funktion $g \circ f : D_1 \to W_2$ in x differenzierbar und es gilt

$$(g \circ f)'(x) = g'(f(x))\, f'(x).$$

Ableitungen einiger Grundfunktionen:

$$\begin{aligned} f(x) &= x^n, & f'(x) &= nx^{n-1}, \\ f(x) &= e^x, & f'(x) &= e^x. \end{aligned}$$

Die Ableitung des Logarithmus erhält man mit Hilfe der Regel über die Umkehrfunktion: $f(x) = e^x$, $f^{-1}(y) = \ln y$, $(f^{-1})'(y) = 1/f'(f^{-1}(y)) = 1/\exp(\ln y) = 1/y$, also

$$f(x) = \ln x, \qquad f'(x) = \frac{1}{x}.$$

Die Ableitungen von Sinus und Kosinus erhält man aus der Darstellung $\sin(x) = (e^{ix} - e^{-ix})/(2i)$, $\cos(x) = (e^{ix} + e^{-ix})/2$ zu

$$\begin{aligned} f(x) &= \sin(x), & f'(x) &= \cos(x), \\ f(x) &= \cos(x), & f'(x) &= -\sin(x). \end{aligned}$$

Die Ableitung aller weiteren trigonometrischen und hyperbolischen Funktionen lassen sich ebenfalls mit den obigen Regeln bestimmen:

$$\begin{aligned} f(x) &= \tan(x), & f'(x) &= \frac{1}{\cos^2(x)}, \\ f(x) &= \arcsin(x), & f'(x) &= \frac{1}{\sqrt{1-x^2}}, \\ f(x) &= \arctan(x), & f'(x) &= \frac{1}{1+x^2}, \\ f(x) &= \sinh(x), & f'(x) &= \cosh(x), \\ f(x) &= \cosh(x), & f'(x) &= \sinh(x), \\ f(x) &= \tanh(x), & f'(x) &= \frac{1}{\cosh^2(x)}, \\ f(x) &= \operatorname{arsinh}(x), & f'(x) &= \frac{1}{\sqrt{1+x^2}}, \\ f(x) &= \operatorname{artanh}(x), & f'(x) &= \frac{1}{1-x^2}. \end{aligned}$$

5.4.1 Die Regeln von l'Hospital

Seien $f,g : D \to \mathbb{R}$ zwei in $x_0 \in D$ stetige Funktionen mit $f(x_0) = g(x_0) = 0$. Existiert $\lim\limits_{x\to x_0} f'(x)/g'(x)$, so gilt:

$$\lim_{x\to x_0} \frac{f(x)}{g(x)} \quad = \quad \lim_{x\to x_0} \frac{f'(x)}{g'(x)}.$$

Beispiel:

$$\lim_{x\to 0} \frac{1-\cos x}{x^2} \quad = \quad \lim_{x\to 0} \frac{\sin x}{2x} = \frac{1}{2} \lim_{x\to 0} \frac{\cos x}{1} = \frac{1}{2}.$$

Ist $\lim\limits_{x\to x_0} |f(x)| = \infty$ und $\lim\limits_{x\to x_0} |g(x)| = \infty$ und exisitiert $\lim\limits_{x\to x_0} f'(x)/g'(x)$, so gilt ebenfalls

$$\lim_{x\to x_0} \frac{f(x)}{g(x)} \quad = \quad \lim_{x\to x_0} \frac{f'(x)}{g'(x)}.$$

Beispiel:

$$\lim_{x\to 0} \frac{\ln x}{\frac{1}{x}} \quad = \quad \lim_{x\to 0} \frac{\frac{1}{x}}{-\frac{1}{x^2}} = \lim_{x\to 0} (-x) = 0.$$

Bemerkung: Die l'Hospitalschen Regeln gelten auch für $x_0 \to \pm\infty$.

5.5 Integralrechnung

Definition einer Treppenfunktion: Man nennt $t : [a,b] \to \mathbb{R}$ **Treppenfunktion**, falls es eine Unterteilung

$$a = x_0 < x_1 < \ldots < x_n = b$$

gibt, so daß t auf jedem offenen Intervall $]x_{j-1}, x_j[$ konstant ist. Der Wert auf diesem Intervall sei mit c_j bezeichnet.

Das **Integral einer Treppenfunktion** wird definiert als

$$\int\limits_a^b t(x)\, dx \quad = \quad \sum_{j=1}^{n} c_j \left(x_j - x_{j-1}\right).$$

Die Menge aller Treppenfunktionen auf dem Intervall $[a,b]$ bilden einen Vektorraum. Wir bezeichnen diese Vektorraum mit $T[a,b]$. Sei $f : [a,b] \to \mathbb{R}$ eine beliebige beschränkte Funktion und $t \in T[a,b]$. Man schreibt $f \geq t$ falls $f(x) \geq t(x)$ für alle $x \in [a,b]$ gilt.

Wir definieren nun das Ober- und Unterintegral für f:

$$\begin{aligned}
\int_a^{b^*} f(x)\,dx &= \inf\left\{\int_a^b t(x)\,dx;\ t\in T[a,b],\ t\geq f\right\},\\
\int_{a\ *}^{b} f(x)\,dx &= \sup\left\{\int_a^b t(x)\,dx;\ t\in T[a,b],\ t\leq f\right\},
\end{aligned}$$

Definition des Riemann-Integrals: Eine beschränkte Funktion $f:[a,b]\to\mathbb{R}$ heißt **Riemann-integrierbar**, falls

$$\int_a^{b^*} f(x)\,dx = \int_{a\ *}^{b} f(x)\,dx.$$

In diesem Fall setzt man

$$\int_a^b f(x)\,dx = \int_a^{b^*} f(x)\,dx.$$

Satz: Jede stetige Funktion $f:[a,b]\to\mathbb{R}$ ist integrierbar.

Satz: Jede monotone Funktion $f:[a,b]\to\mathbb{R}$ ist integrierbar.

Seien $f,g:[a,b]\to\mathbb{R}$ integrierbare Funktionen und $\lambda\in\mathbb{R}$. Dann sind auch die Funktionen $f+g$ und $\lambda\cdot f$ integrierbar und es gilt

$$\begin{aligned}
\int_a^b (f+g)(x)\,dx &= \int_a^b f(x)\,dx + \int_a^b g(x)\,dx,\\
\int_a^b (\lambda f)(x)\,dx &= \lambda\int_a^b f(x)\,dx.
\end{aligned}$$

Desweiteren ist auch die Funktion fg integrierbar, im allgemeinen ist allerdings

$$\int_a^b (f\cdot g)(x)\,dx \neq \left(\int_a^b f(x)\,dx\right)\cdot\left(\int_a^b g(x)\,dx\right).$$

Eine differenzierbare Funktion $F:I\to\mathbb{R}$ heißt **Stammfunktion** einer Funktion $f:I\to\mathbb{R}$, falls $F'(x)=f(x)$.

Eine weitere Funktion $G : I \to \mathbb{R}$ ist genau dann ebenfalls eine Stammfunktion, falls $F - G$ eine Konstante ist.

Man schreibt auch

$$F(x) \;=\; \int f(x)\,dx.$$

Der Ausdruck auf der rechten Seite wird auch als unbestimmtes Integral bezeichnet.

Hauptsatz der Differential- und Integralrechnung:

$$\int\limits_a^b f(x)\,dx \;=\; F(b) - F(a).$$

Man schreibt auch

$$F(x)\big|_a^b \;=\; F(b) - F(a).$$

Substitutionsregel: Sei $f : [a,b] \to W_1$ eine stetig differenzierbare Funktion und $g : D_2 \to W_2$ eine stetige Funktion mit $W_1 \subset D_2$. Dann gilt

$$\int\limits_a^b g\left(f(x)\right) f'(x)\,dx \;=\; \int\limits_{f(a)}^{f(b)} g(x)\,dx.$$

Partielle Integration: Seien $f,g : [a,b] \to \mathbb{R}$ zwei stetig differenzierbare Funktionen. Dann gilt

$$\int\limits_a^b f(x)g'(x)\,dx \;=\; f(x)g(x)\big|_a^b - \int\limits_a^b f'(x)g(x)\,dx.$$

Stammfunktionen einiger Grundfunktionen:

$$\begin{aligned}
f(x) = x^n, &\qquad F(x) = \frac{x^{n+1}}{n+1}, \quad n \neq -1,\\
f(x) = e^x, &\qquad F(x) = e^x,\\
f(x) = \frac{1}{x}, &\qquad F(x) = \ln|x|,\\
f(x) = \sin(x), &\qquad F(x) = -\cos(x),\\
f(x) = \cos(x), &\qquad F(x) = \sin(x),\\
f(x) = \frac{1}{1+x^2}, &\qquad F(x) = \arctan(x).
\end{aligned}$$

Integrale über rationale Funktionen: Wir betrachten als Beispiel

$$I = \int\limits_0^1 \frac{x^4+3x^3-12x^2-3x+18}{(x-2)^2(x+2)}\,dx.$$

Im ersten Schritt zerlegt man den Integranden mit Hilfe der Partialbruchzerlegung:

$$\frac{x^4+3x^3-12x^2-3x+18}{(x-2)^2(x+2)} = x+5+\frac{1}{(x-2)^2}+\frac{4}{x-2}-\frac{2}{x+2}.$$

Somit ist

$$I = \int\limits_0^1 (x+5)\,dx + \int\limits_0^1 \frac{dx}{(x-2)^2} + 4\int\limits_0^1 \frac{dx}{x-2} - 2\int\limits_0^1 \frac{dx}{x+2}.$$

Wir berechnen nun die einzelnen Integrale:

$$\begin{aligned}
\int\limits_0^1 (x+5)\,dx &= 5x+\frac{1}{2}x^2\Big|_0^1 = \frac{11}{2},\\
\int\limits_0^1 \frac{dx}{(x-2)^2} &= -\frac{1}{x-2}\Big|_0^1 = 1-\frac{1}{2} = \frac{1}{2},\\
\int\limits_0^1 \frac{dx}{x-2} &= \ln(|x-2|)|_0^1 = \ln 1 - \ln 2 = -\ln 2,\\
\int\limits_0^1 \frac{dx}{x+2} &= \ln(|x+2|)|_0^1 = \ln 3 - \ln 2,
\end{aligned}$$

Somit erhalten wir

$$I = \frac{11}{2}+\frac{1}{2}+4(-\ln 2)-2(\ln 3-\ln 2) = 6-2\ln 2-2\ln 3 = 6-2\ln 6.$$

5.5.1 Uneigentliche Integrale

Unter einem uneigentlichen Integral versteht man ein Integral, bei dem eine Integrationsgrenze unendlich ist oder bei dem der Integrand an einer Integrationsgrenze nicht definiert ist. Es kann auch eine Kombination der beiden Fälle auftreten.

Wir betrachten zunächst den Fall, daß eine Integrationsgrenze unendlich ist. Sei $f : [a,\infty[\to \mathbb{R}$

eine Funktion, die über jedem Intervall $[a,\Lambda]$ mit $a<\Lambda<\infty$ Riemann-integrierbar ist. Falls der Grenzwert

$$\lim_{\Lambda\to\infty}\int_a^\Lambda f(x)\,dx$$

existiert, nennt man das Integral von a bis Unendlich konvergent und man setzt

$$\int_a^\infty f(x)\,dx \;=\; \lim_{\Lambda\to\infty}\int_a^\Lambda f(x)\,dx.$$

Analog definiert man das Integral für das Intervall $]-\infty,b]$.

Beispiel:

$$\int_1^\infty \frac{dx}{x^2} \;=\; \lim_{\Lambda\to\infty}\int_1^\Lambda \frac{dx}{x^2} = -\lim_{\Lambda\to\infty}\left.\frac{1}{x}\right|_1^\Lambda = 1-\lim_{\Lambda\to\infty}\frac{1}{\Lambda} = 1.$$

Wir betrachten nun den Fall, daß der Integrand an einer Intervallgrenze nicht definiert ist. Sei $f:]a,b]\to\mathbb{R}$ eine Funktion, die über jedem Teilintervall $[a+\varepsilon,b]$ mit $0<\varepsilon<(b-a)$ Riemann-integrierbar ist. Falls der Grenzwert

$$\lim_{\varepsilon\to 0+}\int_{a+\varepsilon}^b f(x)\,dx$$

existiert, nennt man das Integral über $[a,b]$ konvergent und man setzt

$$\int_a^b f(x)\,dx \;=\; \lim_{\varepsilon\to 0+}\int_{a+\varepsilon}^b f(x)\,dx$$

Analog definiert man das Integral für den Fall in der die Funktion an der oberen Intevallgrenze nicht definiert ist.

Beispiel:

$$\int_0^1 \frac{dx}{\sqrt{x}} \;=\; \lim_{\varepsilon\to 0+}\int_\varepsilon^1 \frac{dx}{\sqrt{x}} = \lim_{\varepsilon\to 0+}\int_\varepsilon^1 \frac{dx}{\sqrt{x}} = 2\lim_{\varepsilon\to 0+}\left.\sqrt{x}\right|_\varepsilon^1 = 2-\lim_{\varepsilon\to 0+}\sqrt{\varepsilon} = 2.$$

Allgemeiner Fall: Sei $f:]a,b[\to\mathbb{R}$, $a\in\mathbb{R}\cup\{-\infty\}$, $b\in\mathbb{R}\cup\{\infty\}$, eine Funktion, die über jedem Teilintervall $[\alpha,\beta]\subset]a,b[$ Riemann-integrierbar ist und sei $c\in]a,b[$ beliebig. Falls die beiden uneigentlichen Integrale

$$\int_a^c f(x)\,dx = \lim_{\alpha\to a+}\int_\alpha^c f(x)\,dx, \qquad \int_c^b f(x)\,dx = \lim_{\beta\to b-}\int_c^\beta f(x)\,dx$$

existieren, nennt man das Integral über $]a,b[$ konvergent und man setzt

$$\int_a^b f(x)\,dx \quad = \quad \int_a^c f(x)\,dx + \int_c^b f(x)\,dx.$$

5.6 Taylorreihen

Wir haben bereits die Reihendarstellung einiger Funktionen, wie zum Beispiel der Exponentialfunktion, Sinus oder Kosinus kennengelernt. In diesem Abschnitt geht es um die systematische Entwicklung von Funktionen in Potenzreihen.

Taylorsche Formel: Sei $I \subset \mathbb{R}$ und $f : I \to \mathbb{R}$ eine $(n+1)$-mal stetig differenzierbare Funktion. Dann gilt für $a \in I$ und $x \in I$

$$f(x) \quad = \quad f(a) + \frac{f'(a)}{1!}(x-a) + \frac{f''(a)}{2!}(x-a)^2 + \frac{f'''(a)}{3!}(x-a)^3 + \ldots + \frac{f^{(n)}(a)}{n!}(x-a)^n + R_{n+1}(x).$$

Hierbei bezeichnet $f^{(n)}$ die n-te Ableitung. Für das Restglied gilt

$$R_{n+1}(x) \quad = \quad \frac{1}{n!}\int_a^x (x-t)^n f^{(n+1)}(t)dt.$$

Eine alternative Form für das Restglied geht auf Lagrange zurück: Es gibt ein $\xi \in [a,x]$, so daß

$$R_{n+1}(x) \quad = \quad \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-a)^{n+1}.$$

Bemerkung: Dies ist eine Existenzaussage, ξ ist im allgemeinen schwer zu bestimmen.

In der Praxis verwendet man die ersten n Terme der Taylorentwicklung, um eine Funktion zu approximieren und vernachlässigt das Restglied. Das vernachlässigte Restglied liefert den Fehler dieser Abschätzung.

Sei nun $f : I \to \mathbb{R}$ eine beliebig of differenzierbare Funktion und $a \in I$. Wir definieren die Taylorreihe einer Funktion f um den Entwicklungspunkt a:

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

Bemerkungen:
1. Der Konvergenzradius der Taylorreihe ist nicht notwendig > 0.
2. Falls die Taylorreihe von f konvergiert, konvergiert sie nicht notwendigerweise gegen f.
3. Die Taylorreihe konvergiert genau für diejenigen $x \in I$ gegen $f(x)$, für die das Restglied gegen

Null konvergiert.

Wir geben ein Gegenbeispiel zu Punkt 2 an: Wir betrachten die Taylorreihe der Funktion

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

im Punkte $a = 0$. f ist beliebig oft differenzierbar und es gilt

$$f^{(n)}(0) = 0.$$

Die Taylorreihe von f um den Nullpunkt ist also identisch Null.

5.7 Orthogonale Polynome

Wir betrachten Polynomfunktionen auf einem Intervall $[a,b]$. Die Menge aller Polynomfunktionen $P : [a,b] \to \mathbb{R}$ bilden einen (unendlich dimensionalen) $\mathbb{R}$-Vektorraum. Eine Basis dieses Vektorraumes ist zum Beispiel durch die Monome

$$1, x, x^2, x^3, x^4, \ldots$$

gegeben. Ebenso bildet eine Sequenz von Polynomen $P_0(x)$, $P_1(x)$, ..., wobei $P_j(x)$ vom Grad j ist, eine Basis dieses Vektorraumes.

Wir wollen nun ein Skalarprodukt auf diesem Vektorraum definieren. Zunächst führen wir den Begriff einer **Gewichtsfunktion** ein. Man nennt eine Funktion $w : [a,b] \to \mathbb{R}$ eine Gewichtsfunktion, falls

1. $w(x) \geq 0$ für alle $x \in [a,b]$;
2. $\int\limits_a^b dx\, w(x) > 0$;
3. $\int\limits_a^b dx\, w(x) x^j < \infty$ für alle $j = 0, 1, 2, \ldots$.

Bemerkung: Es ist nicht gefordert, daß die Gewichtsfunktion ein Polynom sein muß.

Seien nun $P_i(x)$ und $P_j(x)$ zwei Polynome. Mit Hilfe einer Gewichtsfunktion definieren wir nun ein Skalarprodukt wie folgt:

$$\langle P_i, P_j \rangle = \int\limits_a^b dx\, w(x) P_i(x) P_j(x).$$

Bemerkung: Diese Definition hängt von der Wahl der Gewichtsfunktion ab !

Man nennt zwei Polynome orthogonal, falls

$$\langle P_i, P_j \rangle = \int_a^b dx\, w(x) P_i(x) P_j(x) = 0.$$

Gesucht ist nun zu einer gegebenen Gewichtsfunktion eine Basis des Vektorraumes der Polynome über $[a,b]$, die orthogonal bezüglich des Skalarproduktes ist. In anderen Worten, gesucht ist eine Sequenz von Polynomen $P_0(x)$, $P_1(x)$, ..., wobei $P_j(x)$ vom Grad j ist, so daß

$$\int_a^b dx\, w(x) P_i(x) P_j(x) = 0 \qquad \text{für } i \neq j.$$

Die folgende Tabelle zeigt die Intervalle und die Gewichtsfunktionen der üblichen orthogonalen Polynome:

Intervall	Gewichtsfunktion	Polynome	Bemerkungen
$[-1,1]$	1	Legendre	
$[-1,1]$	$1/\sqrt{1-x^2}$	Tschebyscheff	
$[-1,1]$	$(1-x^2)^{\mu-1/2}$	Gegenbauer	$\mu > -1/2$
$[-1,1]$	$(1-x)^\alpha (1+x)^\beta$	Jacobi	$\alpha, \beta > -1$
$[0,\infty]$	$x^\alpha e^{-x}$	(verallgemeinerte) Laguerre	$\alpha > -1$
$[-\infty,\infty]$	e^{-x^2}	Hermite	

Die eigentlichen Laguerre-Polynome entsprechen $\alpha = 0$.

Für die Nullstellen von orthogonalen Polynomen gilt: Die Nullstellen von orthogonalen Polynomen sind reell, einfach und befinden sich im Intervall $[a,b]$. Seien $x_1 < x_2 < ... < x_n$ die Nullstellen des Polynoms $P_n(x)$, dann liegt in jedem Intervall $[a,x_1]$, $[x_1,x_2]$, ..., $[x_{n-1},x_n]$, $[x_n,b]$ genau eine Nullstelle des Polynoms $P_{n+1}(x)$.

5.7.1 Legendre-Polynome

Die Legendre-Polynome $P_n(x)$ sind auf dem Intervall $[-1,1]$ mit der Gewichtsfunktion $w(x) = 1$ definiert. Sie sind wie folgt normiert:

$$\int_{-1}^{1} dx P_n(x) P_m(x) = \frac{2}{2n+1} \delta_{nm}.$$

Explizit sind die Polynome gegeben durch

$$P_n(x) = \frac{1}{2^n} \sum_{m=0}^{[n/2]} (-1)^m \begin{pmatrix} n \\ m \end{pmatrix} \begin{pmatrix} 2n-2m \\ n \end{pmatrix} x^{n-2m},$$

wobei $[n/2]$ die größte ganze Zahl kleiner oder gleich $n/2$ bezeichnet. Rekursiv erhält man diese Polynome durch

$$P_0(x) = 1, \quad P_1(x) = x, \qquad (n+1)P_{n+1}(x) = (2n+1)xP_n(x) - nP_{n-1}(x).$$

Die erzeugende Funktion lautet:

$$\frac{1}{\sqrt{1-2xz+z^2}} = \sum_{n=0}^{\infty} P_n(x) z^n, \quad -1 \le x \le 1, \ |z| < 1.$$

Formel von Rodrigues:

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

Die Legendre-Polynome erfüllen die Differentialgleichung

$$(1-x^2)P_n''(x) - 2xP_n'(x) + n(n+1)P_n(x) = 0.$$

5.7.2 Tschebyscheff-Polynome

Die Tschebyscheff-Polynome der ersten Art $T_n(x)$ sind auf dem Intervall $[-1,1]$ mit der Gewichtsfunktion $w(x) = 1/\sqrt{1-x^2}$ definiert. Sie sind wie folgt normiert:

$$\int\limits_{-1}^{1} dx \frac{1}{\sqrt{1-x^2}} T_n(x) T_m(x) = \begin{cases} \frac{\pi}{2}\delta_{nm}, & n \neq 0, \\ \pi\delta_{nm}, & n = 0. \end{cases}$$

Explizit sind die Polynome gegeben durch

$$T_n(x) = \frac{n}{2} \sum_{m=0}^{[n/2]} (-1)^m \frac{(n-m-1)!}{m!(n-2m)!} (2x)^{n-2m}.$$

Rekursiv erhält man diese Polynome durch

$$T_0(x) = 1, \quad T_1(x) = x, \qquad T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x).$$

Die erzeugende Funktion lautet:

$$\frac{1-xz}{1-2xz+z^2} = \sum_{n=0}^{\infty} T_n(x) z^n, \quad -1 \le x \le 1, \ |z| < 1.$$

Formel von Rodrigues:

$$T_n(x) = \frac{(-1)^n (1-x^2)^{1/2} \sqrt{\pi}}{2^n \Gamma\left(n+\frac{1}{2}\right)} \frac{d^n}{dx^n} \left(1-x^2\right)^{n-1/2}.$$

Die Tschebyscheff-Polynome erfüllen die Differentialgleichung

$$(1-x^2)T_n''(x) - xT_n'(x) + n^2 T_n(x) = 0.$$

5.7.3 Laguerre-Polynome

Die Laguerre-Polynome $L_n(x)$ sind auf dem Intervall $[0,\infty]$ mit der Gewichtsfunktion $w(x) = e^{-x}$ definiert. Sie sind wie folgt normiert:

$$\int\limits_0^\infty dx e^{-x} L_n(x) L_m(x) \quad = \quad \delta_{nm}.$$

Explizit sind die Polynome gegeben durch

$$L_n \quad = \quad \sum_{m=0}^{n} \frac{(-1)^m}{m!} \begin{pmatrix} n \\ n-m \end{pmatrix} x^m$$

Rekursiv erhält man diese Polynome durch

$$L_0(x) = 1, \quad L_1(x) = 1 - x, \qquad (n+1) L_{n+1}(x) = ((2n+1) - x) L_n(x) - n L_{n-1}(x).$$

Die erzeugende Funktion lautet:

$$(1-z)^{-1} \exp\left(\frac{xz}{z-1}\right) \quad = \quad \sum_{n=0}^{\infty} L_n(x) z^n, \quad |z| < 1.$$

Formel von Rodrigues:

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

Die Laguerre-Polynome erfüllen die Differentialgleichung

$$x L_n''(x) + (1-x) L_n'(x) + n L_n(x) \quad = \quad 0.$$

5.7.4 Hermite-Polynome

Die Hermite-Polynome $H_n(x)$ sind auf dem Intervall $[-\infty,\infty]$ mit der Gewichtsfunktion $w(x) = e^{-x^2}$ definiert. Sie sind wie folgt normiert:

$$\int\limits_{-\infty}^{\infty} dx e^{-x^2} H_n(x) H_m(x) \quad = \quad 2^n \sqrt{\pi} n! \delta_{nm}.$$

Explizit sind die Polynome gegeben durch

$$H_n(x) \quad = \quad n! \sum_{m=0}^{[n/2]} (-1)^m \frac{(2x)^{n-2m}}{m!(n-2m)!}.$$

Rekursiv erhält man diese Polynome durch

$$H_0(x) = 1, \quad H_1(x) = 2x, \qquad H_{n+1}(x) = 2x H_n(x) - 2n H_{n-1}(x).$$

Die erzeugende Funktion lautet:

$$e^{-t^2+2xt} = \sum_{n=0}^{\infty} \frac{1}{n!} H_n(x) t^n$$

Formel von Rodrigues:

$$H_n(x) = (-1)^n e^{x^2} \frac{d^n}{dx^n} e^{-x^2}$$

Die Hermite-Polynome erfüllen die Differentialgleichung

$$H_n''(x) - 2xH_n'(x) + 2nH_n(x) = 0.$$

6 Differentialgleichungen

Wir wollen nun Gleichungen betrachten, die eine unbekannte Funktion und deren Ableitung enthalten. Diese Gleichungen nennt man **Differentialgleichungen**. Tritt nur die Ableitung nach einer Variablen auf, spricht man von einer **gewöhnlichen Differentialgleichung**. Hängt dagegen die gesuchte Funktion von mehreren Variablen ab, und treten Ableitungen nach verschiedenen Variablen auf, so spricht man von einer **partiellen Differentialgleichung**. Wir behandeln zunächst gewöhnliche Differentialgleichungen. Ein Beispiel für eine Differentialgleichung ist

$$f''(x) + \omega^2 f(x) = 0.$$

Definition: Sei D eine Teilmenge von $\mathbb{R}^2$ und

$$\begin{aligned} G \ : \ & D \to \mathbb{R}, \\ & (x,y) \to G(x,y) \end{aligned}$$

eine stetige Funktion. Dann nennt man

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

eine **Differentialgleichung erster Ordnung**. Unter einer Lösung versteht man eine auf einem Intervall $I \subset \mathbb{R}$ definierte differenzierbare Funktion

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

mit folgenden Eigenschaften:

- Der Graph von f ist in D enthalten, d.h.

$$\Gamma_f = \{(x,y) \in I \times \mathbb{R} : y = f(x)\} \subset D.$$

- Es gilt

$$f'(x) = G(x, f(x)).$$

Beispiel: $G(x,y) = -\lambda y$ führt auf die Differentialgleichung

$$f'(x) = -\lambda f(x).$$

Bemerkung: Hängt die Funktion G nur von x, aber nicht von y ab, so hat man

$$f'(x) = G(x)$$

und man erhält eine Lösung durch Integration:

$$f(x) = c + \int\limits_{x_0}^{x} G(x)\, dx.$$

Wir definieren noch Differentialgleichungen höherer Ordnungen: Hierzu betrachten wir zunächst Systeme von Differentialgleichungen erster Ordnung und zeigen dann, daß sich Differentialgleichungen n-ter Ordnung auf ein System von n Differentialgleichungen erster Ordnung reduzieren lassen.

Sei D eine Teilmenge von $\mathbb{R} \times \mathbb{R}^n$ und

$$\begin{aligned} \vec{G} \;\; : \;\; & D \to \mathbb{R}^n, \\ & (x, \vec{y}) \to \vec{G}(x, \vec{y}) \end{aligned}$$

eine stetige Funktion. Dann nennt man

$$\vec{y}\,' = \vec{G}(x, \vec{y})$$

ein **System von n Differentialgleichungen erster Ordnung**. Unter einer Lösung versteht man eine auf einem Intervall $I \subset \mathbb{R}$ definierte differenzierbare Funktion

$$\vec{f} \;\; : \;\; I \to \mathbb{R}^n$$

mit folgenden Eigenschaften:

- Der Graph von $\vec{f}$ ist in D enthalten.
- Es gilt

$$\vec{f}\,'(x) = \vec{G}\left(x, \vec{f}(x)\right).$$

$\vec{G}$ and $\vec{f}$ sind Vektoren. Schreiben wir diese in Komponenten

$$\vec{G} = \left(\begin{array}{c} G_1 \\ \dots \\ G_n \end{array} \right), \qquad \vec{f} = \left(\begin{array}{c} f_1 \\ \dots \\ f_n \end{array} \right),$$

so ergibt sich ein System von n Differentialgleichungen erster Ordnung:

$$\begin{aligned} f_1'(x) &= G_1(x, f_1(x), \dots, f_n(x)), \\ f_2'(x) &= G_2(x, f_1(x), \dots, f_n(x)), \\ & \dots \\ f_n'(x) &= G_n(x, f_1(x), \dots, f_n(x)). \end{aligned}$$

Beispiel: Für $n = 2$ führt die Abbildung

$$\vec{G}(x, y_0, y_1) = \left(\begin{array}{r} y_1 \\ -\omega^2 y_0 \end{array} \right)$$

auf das System

$$\begin{aligned} f_0'(x) &= f_1(x), \\ f_1'(x) &= -\omega^2 f_0(x). \end{aligned}$$

Sei D eine Teilmenge von $\mathbb{R} \times \mathbb{R}^n$ und

$$\begin{aligned} \tilde{G} \; : \; & D \to \mathbb{R}, \\ & (x,\vec{y}) \to \tilde{G}(x,\vec{y}) \end{aligned}$$

eine stetige Funktion. Dann nennt man

$$y^{(n)} = \tilde{G}\left(x,y,y',...,y^{(n-1)}\right)$$

eine **Differentialgleichungen n-ter Ordnung**. Unter einer Lösung versteht man eine auf einem Intervall $I \subset \mathbb{R}$ definierte differenzierbare Funktion

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

mit folgenden Eigenschaften:

- Die Menge

  $$\left\{(x,y_0,y_1,...,y_{n-1}) \in I \times \mathbb{R}^n : y_j = f^{(j)}(x), 0 \leq j < n\right\}$$

  ist in D enthalten.

- Es gilt

  $$f^{(n)}(x) = \tilde{G}\left(x, f(x), f'(x), ..., f^{(n-1)}(x)\right).$$

Beispiel: $\tilde{G}(x,y_1,y_2) = -\omega^2 y_1$ führt auf die Differentialgleichung zweiter Ordnung

$$f''(x) = -\omega^2 f(x).$$

6.1 Reduktion einer Differentialgleichung höherer Ordnung auf ein System erster Ordnung

Sei nun eine Differentialgleichung n-ter Ordnung gegeben:

$$y^{(n)} = \tilde{G}\left(x,y,y',...,y^{(n-1)}\right)$$

Für

$$\vec{y} = \begin{pmatrix} y_0 \\ ... \\ y_{n-2} \\ y_{n-1} \end{pmatrix}, \quad \vec{G} = \begin{pmatrix} y_1 \\ ... \\ y_{n-1} \\ \tilde{G}(x,\vec{y}) \end{pmatrix},$$

betrachten wir das System von n Differentialgleichungen erster Ordnung:

$$\vec{y}\,' = \vec{G}(x,\vec{y})$$

Ausgeschrieben in Komponenten haben wir

$$\begin{aligned}
y_0' &= y_1, \\
y_1' &= y_2, \\
&\dots \\
y_{n-2}' &= y_{n-1}, \\
y_{n-1}' &= \tilde{G}\left(x, y_0, y_1, \dots, y_{(n-1)}\right).
\end{aligned}$$

Wir betrachten nun die n-te Ableitung von y_0:

$$\begin{aligned}
y_0^{(n)} &= y_1^{(n-1)} = y_2^{n-2)} = \dots = y_{n-1}' = \tilde{G}\left(x, y_0, y_1, \dots, y_{(n-1)}\right) \\
&= \tilde{G}\left(x, y_0, y_0', \dots, y_0^{(n-1)}\right).
\end{aligned}$$

Somit erhält man die Lösung der Differentialgleichung n-ter Ordnung, indem man zunächst das zugehörige System von n Differentialgleichungen erster Ordnung löst. Ist diese Lösung $\vec{f}(x)$, so ist die erste Komponente von $\vec{f}$ Lösung der Differentialgleichung n-ter Ordnung. (Die weiteren Komponenten von $\vec{f}$ sind die Ableitungen der ersten Komponente.)

Beispiel: Zu der Differentialgleichung zweiter Ordnung

$$f''(x) = -\omega^2 f(x).$$

gehört das System von Differentialgleichungen erster Ordnung

$$\begin{aligned}
f_0'(x) &= f_1(x), \\
f_1'(x) &= -\omega^2 f_0(x).
\end{aligned}$$

6.2 Existenz und Eindeutigkeit von Lösungen

In diesem Abschnitt beschäftigen wir uns mit Existenz und Eindeutigkeit von Lösungen eines Systems von Differentialgleichungen erster Ordnung. Wie wir im obigen Abschnitt gesehen haben, lassen sich Differentialgleichungen höherer Ordnung immer auf ein System erster Ordnung zurückführen.

Definition: Sei D eine Teilmenge von $\mathbb{R} \times \mathbb{R}^n$ und

$$\begin{aligned}
\vec{G} \; : \; & D \to \mathbb{R}^n, \\
& (x,\vec{y}) \to \vec{G}(x,\vec{y})
\end{aligned}$$

eine Funktion. Man sagt, $\vec{G}$ genügt in D einer **Lipschitz-Bedingung** mit der Lipschitz-Konstanten $L \geq 0$, falls für alle $(x,\vec{y})$, $(x,\vec{z}) \in D$ gilt

$$\left|\vec{G}(x,\vec{y}) - \vec{G}(x,\vec{z})\right| \quad \leq \quad L\,|\vec{y}-\vec{z}|\,.$$

Man sagt, $\vec{G}$ genügt in D lokal einer Lipschitz-Bedingung, falls jeder Punkt $(x,\vec{y}) \in D$ eine Umgebung U besitzt, so daß $\vec{G}$ in $D \cap U$ einer Lipschitz-Bedingung mit einer von U abhängigen Konstanten L genügt.

Satz: Sei D offen. Ist die Funktion $\vec{G}(x,\vec{y})$ bezüglich der Variablen $\vec{y} = (y_1,\ldots,y_n)$ stetig partiell differenzierbar, so genügt $\vec{G}$ in D lokal einer Lipschitz-Bedingung.

Satz über die **Eindeutigkeit von Lösungen**: Wir setzen voraus, daß die Funktion $\vec{G}$ in D lokal einer Lipschitz-Bedingung genügt. Seien $\vec{f}(x)$ und $\vec{g}(x)$ zwei Lösungen der Differentialgleichung

$$\vec{y}\,' \quad = \quad \vec{G}(x,\vec{y})$$

über einem Intervall $I \subset \mathbb{R}$. Gilt dann

$$\vec{f}(x_0) \quad = \quad \vec{g}(x_0)$$

für ein $x_0 \in I$, so folgt

$$\vec{f}(x) \quad = \quad \vec{g}(x)$$

für alle $x \in I$.

Satz über die **Existenz von Lösungen** von Picard – Lindelöf: Sei D offen und $\vec{G} : D \to \mathbb{R}^n$ eine stetige Funktion, die lokal einer Lipschitz-Bedingung genügt. Dann gibt es zu jedem $(x_0,\vec{y}_0) \in D$ ein $\varepsilon > 0$ und eine Lösung

$$\vec{f} \quad : \quad [x_0-\varepsilon, x_0+\varepsilon] \to \mathbb{R}^n$$

der Differentialgleichung $\vec{y}\,' = \vec{G}(x,\vec{y})$ mit der Anfangsbedinung

$$\vec{f}(x_0) \quad = \quad \vec{y}_0.$$

Bemerkung: Überträgt man dies auf eine Differentialgleichung n-ter Ordnung, so wird eine Lösung eindeutig durch n Anfangsbedingungen $f(x_0)$, $f'(x_0)$, ..., $f^{(n-1)}(x_0)$ bestimmt.

6.3 Elementare Lösungsmethoden

Differentialgleichung mit separierten Variablen: Seien $I,J \subset \mathbb{R}$ offene Intervalle und

$$h : I \to \mathbb{R}, \qquad k : J \to \mathbb{R},$$

zwei stetige Funktionen. Es gelte ausserdem $k(y) \neq 0$ für alle $y \in J$. Wir betrachten eine Differenzialgleichung erster Ordnung

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

Die Funktion G ist auf dem Gebiet $D = I \times J$ definiert und wir nehmen an, daß die Variablen sich trennen lassen:

$$G(x,y) = h(x)k(y).$$

Sei nun $(x_0,y_0) \in I \times J$. Wir setzen

$$H(x) = \int_{x_0}^{x} h(t)dt, \qquad K(y) = \int_{y_0}^{y} \frac{dt}{k(t)}.$$

Es sei $I' \subset I$ ein Intervall mit $x_0 \in I'$ und $H(I') \subset K(J)$. Dann exisitiert genau eine Lösung $f : I' \to \mathbb{R}$ mit

$$f(x_0) = y_0.$$

Diese Lösung erfüllt die Beziehung

$$K(f(x)) = H(x).$$

Beispiel: Wir betrachten die Differentialgleichung

$$y' = 2xe^{-y}$$

und suchen eine Lösung zu der Anfangsbedingung $f(0) = c$. Die Variablen sind klarerweise getrennt. Für dieses Beispiel können wir

$$h(x) = 2x, \qquad k(y) = e^{-y}$$

setzen. Wir erhalten

$$\begin{aligned} H(x) &= 2\int_{0}^{x} t\, dt = x^2, \\ K(y) &= \int_{c}^{y} \frac{dt}{e^{-t}} = e^y - e^c. \end{aligned}$$

Somit

$$e^{f(x)} - e^c = x^2.$$

Umgeformt ergibt sich

$$f(x) = \ln\left(x^2 + e^c\right).$$

Als zweites Beispiel betrachten wir die Differentialgleichung

$$y' = y^2.$$

Gesucht ist eine Lösung zu der Anfangsbedingung $y(0) = 1$. Wir haben

$$\frac{dy}{y^2} = dx,$$

und somit liefert die Integration

$$-\frac{1}{y} = x + c.$$

Durch Auflösen nach y erhält man

$$y = -\frac{1}{x+c}.$$

Die Anfangsbedingung $y(0) = 1$ liefert $c = -1$, somit lautet die Lösung

$$y = \frac{1}{1-x}.$$

Diese Lösung hat einen Pol bei $x = 1$. Dies veranschaulicht die Bedeutung des Satzes von Picard-Lindelöf, der eine Lösung lokal um den Punkt $x_0 = 0$ garantiert: In diesem Fall erhält man eine Lösung auf dem Intervall $]-\infty, 1[$ zu der Anfangsbedingung $y(0) = 1$. Die Lösung läßt sich nicht stetig über den Punkt $x = 1$ hinaus fortsetzen.

Wir betrachten einen weiteren Typ von Differentialgleichungen: Sei $J \subset \mathbb{R}$ ein Intervall, $g : J \to \mathbb{R}$ eine stetige Funktion und

$$D = \left\{(x,y) \in \mathbb{R}^* \times \mathbb{R}, \frac{y}{x} \in J\right\}.$$

Wir betrachten die Differentialgleichung

$$y' = g\left(\frac{y}{x}\right).$$

Diese Differentialgleichung kann auf eine Differentialgleichung mit getrennten Variablen zurückgeführt werden. Hierzu betrachten wir die Differenzialgleichung

$$z' = \frac{1}{x}\left(g(z) - z\right).$$

Sei $I \subset \mathbb{R}^*$ ein Intervall und $(x_0, y_0) \in D$ ein Punkt mit $x_0 \in I$. Eine Funktion $f(x)$ ist genau dann Lösung der ersten Differentialgleichung zum Anfangswert $f(x_0) = y_0$, falls die Funktion

$$\tilde{f}(x) = \frac{f(x)}{x}$$

Lösung der zweiten Differentialgleichung mit der Anfangsbedingung $\tilde{f}(x_0) = y_0/x_0$ ist.

Beweis: Setzen wir $z = y/x$, so ist

$$\frac{dz}{dx} = -\frac{y}{x^2} + \frac{y'}{x} = \frac{1}{x}\left(y' - \frac{y}{x}\right) = \frac{1}{x}\left[g\left(\frac{y}{x}\right) - \frac{y}{x}\right] = \frac{1}{x}\left(g(z) - z\right).$$

Beispiel:

$$y' = 1 + \frac{y}{x} + \left(\frac{y}{x}\right)^2.$$

Die Substitution $z = y/x$ führt auf

$$z' = \frac{1}{x}\left(1 + z^2\right),$$

also

$$\frac{dz}{1+z^2} = \frac{dx}{x}.$$

Mit der Anfangsbedingung $z(x_0) = z_0 = y_0/x_0$ findet man

$$\arctan z - \arctan z_0 = \ln\frac{x}{x_0},$$

also

$$z = \tan\left(\ln\frac{x}{x_0} + \arctan z_0\right).$$

Die Rücksubstitution $y = xz$ liefert

$$y = x\tan\left(\ln\frac{x}{x_0} + \arctan\frac{y_0}{x_0}\right).$$

6.4 Lineare Differentialgleichungen

Wir betrachten nun einen wichtigen Typ von Differentialgleichungen. Ist die Funktion $G(x,y)$ linear in der Variablen y, so spricht man von einer **linearen Differentialgleichung**. Es sei angemerkt, daß nicht gefordert wird, daß $G(x,y)$ auch linear in x ist. Ebenso spricht man bei einem

System von n Differentialgleichungen erster Ordnung von einem linearen Differentialgleichungssystem, falls $\vec{G}(x,\vec{y})$ linear in $\vec{y}$ ist.

Wir betrachten zunächst eine lineare Differentialgleichung erster Ordnung. Da $G(x,y)$ linear in y ist, lässt sich diese Funktion zweier Variablen aufgrund der Linearität mit Hilfe zweier Funktionen $a(x)$ und $b(x)$ wie folgt schreiben:

$$G(x,y) \quad = \quad a(x)y + b(x).$$

Sei I ein Intervall und seien $a, b : I \to \mathbb{R}$ zwei stetige Funktionen. Die lineare Differentialgleichung erster Ordnung lautet

$$y' \quad = \quad a(x)y + b(x).$$

Man spricht von einer **homogenen** linearen Differentialgleichung erster Ordnung, falls die Gleichung von der Form

$$y' \quad = \quad a(x)y$$

ist. Ist im ersten Fall $b(x) \neq 0$, so spricht man auch von einer inhomogenen Gleichung.

6.4.1 Lösung der homogenen Gleichung

Wir betrachten zunächst die homogene Gleichung. Wir werden später sehen, daß die Lösungen der homogenen Gleichung auch eine Rolle bei der Lösung der inhomogenen Gleichung spielen. Die homogene Differentialgleichung ist eine Gleichung mit getrennten Variablen und man erhält als Lösung zmu Anfangswert $y(x_0) = c$

$$y(x) \quad = \quad c \exp\left(\int\limits_{x_0}^{x} a(t)\, dt \right).$$

Beispiel: Die Differentialgleichung

$$y' \quad = \quad 2xy$$

hat zum Anfangswert $y(0) = 1$ die Lösung

$$y(x) \quad = \quad \exp\left(2 \int\limits_{0}^{x} t\, dt \right) = \exp\left(x^2\right).$$

6.4.2 Lösung der inhomogenen Gleichung

Wir betrachten nun eine inhomogene lineare Differentialgleichung erster Ordnung.

$$y' \quad = \quad a(x)y + b(x).$$

Sei $\varphi : I \longrightarrow \mathbb{R}$ eine Lösung der zugehörigen homogenen Differentialgleichung, d.h.

$$\varphi'(x) \quad = \quad a(x)\,\varphi(x).$$

Sei weiter $\varphi(x_0) \neq 0$. Dann ist auch $\varphi(x) \neq 0$ für alle $x \in I$, da die Exponentialfunktion stets positiv ist. Für die gesuchte Lösung der inhomogenen Gleichung machen wir daher den Ansatz

$$f(x) \quad = \quad \varphi(x)\,g(x).$$

Nun ist

$$f'(x) \quad = \quad \varphi'(x)\,g(x) + \varphi(x)\,g'(x) = a(x)\,\varphi(x)\,g(x) + \varphi(x)\,g'(x).$$

Andererseits ist

$$a(x)\,f(x) + b(x) \quad = \quad a(x)\,\varphi(x)\,g(x) + b(x),$$

und somit

$$\varphi(x)\,g'(x) \quad = \quad b(x).$$

Da nach Voraussetzung $\varphi(x) \neq 0$ ist, erhält man

$$g'(x) \quad = \quad \frac{b(x)}{\varphi(x)},$$

und die Funktion $g(x)$ ergibt sich zu

$$g(x) \quad = \quad c + \int\limits_{x_0}^{x} \frac{b(t)}{\varphi(t)}\,dt.$$

Wir fassen zusammen: Die Lösung der inhomogenen Gleichung $y' = a(x)y + b(x)$ mit der Anfangsbedingung $f(x_0) = c$ ergibt sich zu

$$f(x) \quad = \quad \left(c + \int\limits_{x_0}^{x} \frac{b(t)}{\varphi(t)}\,dt \right) \varphi(x),$$

wobei

$$\varphi(x) \quad = \quad \exp\left(\int\limits_{x_0}^{x} a(t)\,dt \right)$$

eine Lösung der zugehörigen homogenen Differentialgleichung ist. Lösungen der homogenen Gleichung haben die Form $c'\varphi(x)$. Bei der Lösung der inhomogenen Gleichung spricht man von der "Variation der Konstanten", da die Konstante c' durch die Funktion $g(x)$ ersetzt wird.

Beispiel: Wir betrachten die inhomogene Gleichung

$$y' = 2xy + x^3$$

und suchen die Lösung zur Anfangsbedinung $f(0) = 1$. Die Lösungen der homogenen Gleichung haben die Form

$$\varphi(x) = c \exp\left(x^2\right).$$

Wir machen nun den Ansatz

$$f(x) = g(x)\, e^{x^2}.$$

Für $g(x)$ erhält man die Differentialgleichung

$$g'(x) = x^3 e^{-x^2},$$

und somit

$$g(x) = c + \int\limits_0^x t^3 e^{-t^2}\, dt.$$

Wir berechnen das Integral mit Hilfe der Substitution $s = t^2$ und partieller Integration

$$\begin{aligned}
\int\limits_0^x t^3 e^{-t^2}\, dt &= \frac{1}{2}\int\limits_0^{x^2} s e^{-s}\, ds = -\frac{1}{2}\int\limits_0^{x^2} s\left(\frac{d}{ds} e^{-s}\right)\, ds = -\frac{1}{2} s e^{-s}\Big|_0^{x^2} + \frac{1}{2}\int\limits_0^{x^2} e^{-s}\, ds \\
&= -\frac{1}{2} x^2 e^{-x^2} - \frac{1}{2} e^{-x^2} + \frac{1}{2} = \frac{1}{2} - \frac{1}{2}\left(1 + x^2\right) e^{-x^2}.
\end{aligned}$$

Man erhält als Lösung der inhomogenen Gleichung

$$f(x) = \left[c + \frac{1}{2} - \frac{1}{2}\left(1 + x^2\right) e^{-x^2}\right] e^{x^2} = -\frac{1}{2}\left(1 + x^2\right) + \left(c + \frac{1}{2}\right) e^{x^2}.$$

Wir bestimmen noch die Konstante c aus der Anfangsbedinung $f(0) = 1$:

$$\begin{aligned}
-\frac{1}{2} + \left(c + \frac{1}{2}\right) &= 1, \\
c &= 1.
\end{aligned}$$

Somit lautet die vollständige Lösung

$$f(x) = -\frac{1}{2}\left(1 + x^2\right) + \frac{3}{2} e^{x^2}.$$

6.4.3 Systeme linearer Differentialgleichungen

Wir betrachten ein System von n linearen Differentialgleichungen erster Ordnung. Es ist zweckmässig, nicht nur reellwertige Funktionen, sondern auch komplexwertige Funktionen zu betrachten. Dies ist eine naheliegende Erweiterung. Ein System von n komplexwertigen Differentialgleichungen ist äquivalent zu einem System von $2n$ reellwertigen Differentialgleichungen, da man eine komplexwertige Funktion immer in zwei reellwertige Funktionen aufspalten kann, die den Real- und Imagimärteil beschreiben. Sei K also $\mathbb{R}$ oder $\mathbb{C}$. Weiter sei $I \subset \mathbb{R}$ ein Intervall. Beachte das I immer eine Teilmenge von $\mathbb{R}$ ist. Gegeben seien n^2 Funktionen

$$\begin{aligned} a_{ij} \;:\; & I \rightarrow K, \qquad 1 \leq i, j \leq n, \\ & x \rightarrow a_{ij}(x), \end{aligned}$$

die wir in einer Matrix $A(x)$ zusammenfassen:

$$A(x) = \begin{pmatrix} a_{11}(x) & a_{12}(x) & ... & a_{1n}(x) \\ a_{21}(x) & a_{22}(x) & ... & a_{2n}(x) \\ ... & ... & ... & ... \\ a_{n1}(x) & a_{n2}(x) & ... & a_{nn}(x) \end{pmatrix}$$

Weiter seien n Funktionen

$$\begin{aligned} b_i \;:\; & I \rightarrow K, \qquad 1 \leq i \leq n, \\ & x \rightarrow b_i(x), \end{aligned}$$

gegeben, die wir in einem Vektor $\vec{b}(x)$ zusammenfassen:

$$\vec{b}(x) = \begin{pmatrix} b_1(x) \\ b_2(x) \\ ... \\ b_n(x) \end{pmatrix}.$$

Dann beschreibt die Gleichung

$$\vec{y}\,' = A(x)\vec{y} + \vec{b}(\vec{x})$$

ein System von n inhomogenen linearen Differentialgleichungen. Die zugehörige homogene Gleichung lautet

$$\vec{y}\,' = A(x)\vec{y}.$$

Satz: Sei I ein offenes Intervall und seien alle Funktionen a_{ij} und b_i stetig. Dann hat die inhomogene Differentialgleichung auf ganz I genau eine Lösung zu der Anfangsbedingung $\vec{y}(x_0) = \vec{c}$.

Bemerkung: Der Satz von Picard-Lindelöf garantiert nur die Existenz einer Lösung in der Umgebung eines Punktes x_0. Die wesentliche Aussage dieses Satzes ist, daß eine Lösung auf ganz I

existiert.

Wir betrachten zunächst wieder die homogene Gleichung. Wir bezeichnen mit L die Menge aller Lösungen $\vec{\varphi} : I \to K^n$ der homogenen Gleichung

$$\vec{\varphi}\,'(x) \;=\; A(x)\vec{\varphi}(x).$$

Satz: Die Menge L aller Lösungen ist ein Vektorraum über K. Der Beweis ist recht einfach: Zunächst ist die Menge L nicht leer, da die Nullfunktion $\vec{0}$ offensichtlich eine (triviale) Lösung der Differentialgleichung ist. Seien nun $\vec{\varphi}, \vec{\psi} \in L$ und $\lambda \in K$. Nun ist

$$\begin{aligned} (\vec{\varphi}+\vec{\psi})' &= \vec{\varphi}\,' + \vec{\psi}' = A\vec{\varphi} + A\vec{\psi} = A\left(\vec{\varphi}+\vec{\psi}\right), \\ (\lambda\vec{\varphi})' &= \lambda\vec{\varphi}\,' = \lambda A\vec{\varphi} = A\left(\lambda\vec{\varphi}\right). \end{aligned}$$

Satz: Der Vektorraum L hat die Dimension n. Seien $\vec{\varphi}_1(x)$, ..., $\vec{\varphi}_n(x)$ Lösungen der homogenen Gleichung. Dann sind die folgenden drei Aussagen äquivalent:

- $\vec{\varphi}_1(x)$, ..., $\vec{\varphi}_n(x)$ sind linear unabhängig und bilden somit eine Basis von L.
- Es existiert ein $x_0 \in I$, so daß die Vektoren $\vec{\varphi}_1(x_0)$, ..., $\vec{\varphi}_n(x_0) \in K^n$ linear unabhängig sind.
- Für alle $x_0 \in I$ sind die Vektoren $\vec{\varphi}_1(x_0)$, ..., $\vec{\varphi}_n(x_0)$ linear unabhängig.

Die wesentliche Aussage dieses Satzes besteht darin, daß es für einen Satz von Lösungsfunktionen $\vec{\varphi}_1(x)$, ..., $\vec{\varphi}_n(x)$ genügt, für einen Punkt x_0 zu überprüfen, ob die Vektoren $\vec{\varphi}_1(x_0)$, ..., $\vec{\varphi}_n(x_0)$ linear unabhängig sind. Damit sind dann auch automatisch die Lösungsfunktionen linear unabhängig. Der Test, ob eine Menge von Lösungsfunktionen linear unabhängig ist, reduziert sich also auf ein Problem der linearen Algebra.

Wir bezeichnen eine Basis des Vektorraumes L als ein **Lösungs-Fundamentalsystem**. Wir können die Lösungen $\vec{\varphi}_j(x)$ in Spaltenform schreiben:

$$\vec{\varphi}_j(x) \;=\; \begin{pmatrix} \varphi_{1j}(x) \\ \dots \\ \varphi_{nj}(x) \end{pmatrix}, \quad 1 \le j \le n.$$

Dies definiert eine Matrix

$$\Phi(x) \;=\; \begin{pmatrix} \varphi_{11}(x) & \dots & \varphi_{1n}(x) \\ \dots & & \dots \\ \varphi_{n1}(x) & \dots & \varphi_{nn}(x) \end{pmatrix}.$$

Mit Hilfe des vorherigen Satzes bilden die Lösungen $\vec{\varphi}_1(x)$, ..., $\vec{\varphi}_n(x)$ eine Basis, falls für einen Punkt $x_0 \in I$

$$\det\,\Phi(x_0) \;\neq\; 0$$

gilt. Die allgemeine Lösung des homogenen Systems läßt sich schreiben als

$$\vec{\varphi}(x) = c_1\vec{\varphi}_1(x) + \ldots + c_n\vec{\varphi}_n(x),$$

also

$$\vec{\varphi}(x) = \Phi(x)\vec{c}, \qquad \vec{c} = \begin{pmatrix} c_1 \\ \ldots \\ c_n \end{pmatrix}.$$

Wir betrachten nun die inhomogene Gleichung

$$\vec{y}\,' = A(x)\vec{y} + \vec{b}(\vec{x}),$$

und bezeichnen mit $\tilde{L}$ die Menge aller Lösungen $\vec{f} : I \to K^n$ der inhomogenen Gleichung. Wie zuvor bezeichne L die Menge der Lösungen der zugehörigen homogenen Gleichung.

Satz: Sei $\vec{f}_0 \in \tilde{L}$ eine Lösung der inhomogenen Gleichung. Dann gilt

$$\tilde{L} = \vec{f}_0 + L.$$

Man erhält also die allgemeine Lösung der inhomogenen Gleichung aus der Summe einer speziellen Lösung der inhomogenen Gleichung plus der allgemeinen Lösung der homogenen Gleichung.

Man bestimmt eine spezielle Lösung der inhomogenen Gleichung mit Hilfe der Variation der Konstanten: Beschreibe $\Phi(x)$ das Fundamentalsystem der homogenen Gleichung. Für die Lösung der inhomogenen Gleichung machen wir den Ansatz

$$\vec{f}(x) = \Phi(x)\vec{g}(x).$$

Dies eingesetzt liefert

$$\vec{f}\,'(x) = \Phi'(x)\vec{g}(x) + \Phi(x)\vec{g}\,'(x) = A(x)\Phi(x)\vec{g}(x) + \Phi(x)\vec{g}\,'(x) \overset{!}{=} A(x)\Phi(x)\vec{g}(x) + \vec{b}(x),$$

also

$$\Phi(x)\vec{g}\,'(x) = \vec{b}(x).$$

Da $\det\Phi(x) \neq 0$ für alle $x \in I$ ist, existiert $\Phi^{-1}(x)$. Somit

$$\vec{g}\,'(x) = \Phi^{-1}(x)\vec{b}(x)$$

und

$$\vec{g}(x) = \vec{c} + \int\limits_{x_0}^{x} \Phi^{-1}(t)\vec{b}(t)\, dt.$$

6.4.4 Lineare Differentialgleichungen höherer Ordnung

Wir hatten schon gezeigt, daß sich eine Differentialgleichung n-ter Ordnung immer auf ein System von n Differentialgleichungen erster Ordnung reduzieren läßt. In diesem Abschnitt soll auf speziellen Fall einer linearen Differentialgleichung n-ter Ordnung näher eingegangen werden.

Es sei $I \subset \mathbb{R}$ ein Intervall und $a_k : I \to K$ $(0 \leq k < n)$ und $b : I \to K$ stetige Funktionen. Wie zuvor betrachten wir entweder reellwertige oder komplexwertige Funktionen. Man nennt

$$y^{(n)} + a_{n-1}(x)y^{(n-1)} + \ldots + a_1(x)y' + a_0(x)y \quad = \quad b(x)$$

lineare Differentialgleichung n-ter Ordnung. Ist $b(x) = 0$, so spricht man von einer homogenen Gleichung. Das zugehörige System von n Differentialgleichungen erster Ordnung ist gegeben durch

$$\begin{aligned}
y_0' &= y_1, \\
y_1' &= y_2, \\
&\ldots \\
y_{n-2}' &= y_{n-1}, \\
y_{n-1}' &= -a_{n-1}(x)y_{n-1} - \ldots - a_1(x)y_1 - a_0(x)y_0 + b(x),
\end{aligned}$$

also

$$\vec{y} = \begin{pmatrix} y_0 \\ y_1 \\ \ldots \\ y_{n-2} \\ y_{n-1} \end{pmatrix}, \quad A = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & & 0 \\ \ldots & & & & \ldots \\ 0 & 0 & 0 & \ldots & 1 \\ -a_0(x) & -a_1(x) & -a_2(x) & \ldots & -a_{n-1}(x) \end{pmatrix}, \quad \vec{b} = \begin{pmatrix} 0 \\ 0 \\ \ldots \\ 0 \\ b(x) \end{pmatrix}.$$

Wir übersetzen nun die Aussagen über Systeme linearer Differentialgleichungen erster Ordnung auf die lineare Differentialgleichung n-ter Ordnung:

Satz: Sei L die Menge aller Lösungen der homogenen linearen Differentialgleichung n-ter Ordnung und $\tilde{L}$ die Menge aller Lösungen der inhomogenen Gleichung. Dann gilt:

1. L ist ein Vektorraum der Dimension n.
2. Für ein beliebiges $f_0(x) \in \tilde{L}$ gilt
$$\tilde{L} \quad = \quad f_0 + L.$$
3. Ein n-Tupel $\varphi_1(x)$, ..., $\varphi_n(x) \in L$ von Lösungen der homogenen gleichung bildet genau dann eine Basis von L, falls für ein $x_0 \in I$ (und damit für alle $x \in I$) gilt:
$$W(x_0) \quad = \quad \det \begin{pmatrix} \varphi_1(x_0) & \ldots & \varphi_n(x_0) \\ \varphi_1'(x_0) & \ldots & \varphi_n'(x_0) \\ \ldots & & \ldots \\ \varphi_1^{(n-1)}(x_0) & \ldots & \varphi_n^{(n-1)}(x_0) \end{pmatrix} \neq 0.$$

Man bezeichnet $W(x)$ als **Wronski-Determinante**.

Beispiel: Die Differentialgleichung

$$y'' - \frac{1}{2x}y' + \frac{1}{2x^2}y = 0$$

besitzt auf dem Intervall $I =]0,\infty[$ die Lösungen

$$\varphi_1(x) = x, \qquad \varphi_2(x) = \sqrt{x},$$

wie man leicht durch Einsetzen überprüft. Die Wronski-Determinante ist

$$W(x) = \begin{vmatrix} x & \sqrt{x} \\ 1 & \frac{1}{2\sqrt{x}} \end{vmatrix} = -\frac{1}{2}\sqrt{x}.$$

Es ist $W(x) \neq 0$ für alle $x \in I$, und somit spannen φ_1 und φ_2 den Lösungsraum auf.

6.5 Lineare Differentialgleichungen mit konstanten Koeffizienten

6.5.1 Systeme von linearen Differentialgleichungen mit konstanten Koeffizienten

Wir betrachten das lineare Differentialgleichunbgssystem

$$\vec{y}\,' = A\vec{y} + \vec{b}(x),$$

wobei A eine $n \times n$-Matrix ist. "Konstante Koeffizienten" bedeutet, daß A von x unabhängig ist, $\vec{b}(x)$ darf dagegen von x abhängen. Wir suchen ein Fundamentalsystem für die homogene Gleichung

$$\vec{y}\,' = A\vec{y},$$

eine Lösung für die inhomogene Gleichung kann dann mittels der Technik der Variation der Konstanten erhalten werden.

Satz: Sei $\vec{v}_\lambda$ ein Eigenvektor der Matrix A zum Eigenwert λ. Dann ist

$$\vec{\varphi}(x) = \vec{v}_\lambda e^{\lambda x}$$

eine Lösung der homogenen Gleichung.

Beweis: Für einen Eigenvektor gilt per Definition $A\vec{v}_\lambda = \lambda\vec{v}_\lambda$. Wir berechnen $\vec{\varphi}\,'(x)$:

$$\vec{\varphi}\,'(x) = \vec{v}_\lambda \frac{d}{dx} e^{\lambda x} = \lambda \vec{v}_\lambda e^{\lambda x} = A\vec{v}_\lambda e^{\lambda x} = A\vec{\varphi}(x).$$

Satz: Besitzt die Matrix A eine Basis $\vec{v}_{\lambda_1}$, ..., $\vec{v}_{\lambda_n}$ von Eigenvektoren mit den Eigenwerten λ_1, ..., λ_n, so bilden die Funktionen

$$\vec{\varphi}_k(x) \;=\; \vec{v}_{\lambda_k} e^{\lambda_k x}, \quad 1 \leq k \leq n,$$

ein Fundamentalsystem von Lösungen der homogenen Differentialgleichung.

Beweis: Oben wurde bereits gezeigt, daß jede dieser Funktionen eine Lösung der homogenen Gleichung ist. Es bleibt zu zeigen, daß diese n Lösungen den gesamten Lösungsraum aufspannen. Hierzu genügt es zu zeigen, daß die Vektoren $\vec{\varphi}_k(x=0) = \vec{v}_{\lambda_k}$, $k = 1, \dots, n$, linear unabhängig sind. Nach Voraussetzung bilden die $\vec{v}_{\lambda_1}$, ..., $\vec{v}_{\lambda_n}$ eine Basis des K^n, sie sind daher linear unabhängig.

Fazit: Die Bestimmung eines Fundamentalsystems reduziert sich für ein System linearer Differentialgleichungen mit konstanten Koeffizienten auf die Bestimmung von Eigenwerten und Eigenvektoren einer Matrix.

Bemerkung: Nicht immer exisiert eine Basis aus Eigenvektoren der Matrix A. Eine solche Basis existiert genau dann, wenn A diagonalisierbar ist. Ein Matrix ist aber immer triagonalisierbar, d.h. sie kann durch einen Basiswechsel auf eine obere Dreiecksgestalt gebracht werden. Hat man ein lineares Differentialgleichungssystem mit einer nicht-diagonalisierbaren Matrix, so bringt man das System zunächst auf eine obere Dreiecksform. Man löst dann die letzte Gleichung und setzt diese Lösung dann in die vorletzte Gleichung ein. Dies ergibt eine inhomogene Differentialgleichung (in einer Variablen) die man mit den bekannten Methoden lösen kann. Man iteriert nun diese Prozedur, bis alle Gleichungen gelöst sind.

6.5.2 Die lineare Differentialgleichung n-ter Ordnung mit konstanten Koeffizienten

In diesem Abschnitt betrachten wir eine lineare Differentialgleichung n-ter Ordnung mit konstanten Koeffizienten. Dieser Fall kommt in der Anwendung oft vor. Für diesen Typ von Differentialgleichungen gibt es eine gute Lösungstheorie über den komplexen Zahlen. Die Bestimmung eines Fundamentalsystems ist wesentlich einfacher im Vergleich zum vorherigen Abschnitt. Daher betrachten wir in diesem Abschnitt nur die komplexen Zahlen.

Es seien $(n+1)$ komplexe Zahlen $a_i \in \mathbb{C}$, $(0 \leq i \leq n)$ gegeben. Wir betrachten die homogene Differentialgleichung

$$a_n y^{(n)} + a_{n-1} y^{(n-1)} + \dots + a_1 y' + a_0 y \;=\; 0.$$

Zu dieser Differentialgleichung betrachten wir das Polynom

$$P(\lambda) = a_n \lambda^n + a_{n-1} \lambda^{n-1} + \dots + a_1 \lambda + a_0 \;\in\; \mathbb{C}[\lambda].$$

Satz: Ist λ_0 eine Nullstelle von $P(\lambda)$, so ist

$$\varphi(x) \;=\; e^{\lambda_0 x}$$

eine Lösung der homogenen Differentialgleichung.

Zum Beweis betrachten wir zunächst die j-te Ableitung von $\varphi(x)$:

$$\frac{d^j}{dx^j}\varphi(x) \quad = \quad \lambda_0^j\,\varphi(x).$$

Eingesetzt in die linke Seite der Differentialgleichung erhält man

$$\begin{aligned} & a_n\varphi^{(n)} + a_{n-1}\varphi^{(n-1)} + \ldots + a_1\varphi' + a_0\varphi = a_n\lambda_0^n\varphi + a_{n-1}\lambda_0^{n-1}\varphi + \ldots + a_1\lambda\varphi + a_0\varphi \\ = \quad & \left(a_n\lambda_0^n + a_{n-1}\lambda_0^{n-1} + \ldots + a_1\lambda + a_0\right)\varphi = P(\lambda_0)\,\varphi = 0. \end{aligned}$$

Die Suche nach den Lösungen der homogenen Differentialgleichung reduziert sich daher auf die Bestimmung der Nullstellen des Polynoms $P(\lambda)$.

Satz: Falls die Nullstellen λ_1, ..., λ_n des Polynoms $P(\lambda)$ paarweise verschieden sind, so bilden die Funktionen

$$\varphi_i(x) \quad = \quad e^{\lambda_i x}, \quad 1 \leq i \leq n,$$

ein Fundamentalsystem für die Lösungen der homogenen Differentialgleichung.

Beweis: Wir müssen überprüfen, daß die Lösungen linear unabhängig sind. Hierzu betrachten wir die Wronski-Determinante am Punkte $x_0 = 0$:

$$W(0) \quad = \quad \left| \begin{array}{llll} 1 & 1 & \ldots & 1 \\ \lambda_1 & \lambda_2 & \ldots & \lambda_n \\ \lambda_1^2 & \lambda_2^2 & \ldots & \lambda_n^2 \\ \ldots & & & \ldots \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \ldots & \lambda_n^{n-1} \end{array} \right|$$

Eine Determinante dieser Form bezeichnet man als Vandermondesche Determinante. Für die Vandermondesche Determinante gilt

$$W(0) \quad = \quad \prod_{i>j}\left(\lambda_i - \lambda_j\right) \neq 0,$$

da nach Voraussetzung alle Nullstellen paarweise verschieden sind.

Wir müssen noch den Fall betrachten, daß einige Nullstellen mehrfach auftreten. Das Polynom $P(\lambda)$ habe also die paarweise verschiedenen Nullstellen λ_i mit Vielfachheiten ν_i, wobei $1 \leq i \leq r$ und $\nu_1 + \ldots + \nu_r = n$ gilt.

Satz: Die homogene lineare Differentialgleichung

$$P\left(\frac{d}{dx}\right)y \quad = \quad 0$$

besitzt ein Lösungs-Fundamentalsystem aus folgenden Funktionen:

$$\varphi_{ij}(x) \quad = \quad x^j e^{\lambda_i x}, \quad 1 \le i \le r, \quad 0 \le j \le \nu_i - 1.$$

Beispiel: Gesucht ist ein Lösungs-Fundamentalsystem der Differentialgleichung

$$y'' - 2y' + y \quad = \quad 0.$$

Das zugehörige Polynom lautet

$$P(\lambda) \quad = \quad \lambda^2 - 2\lambda + 1 = (\lambda - 1)^2 .$$

Also ist ein Fundamentalsystem gegeben durch

$$\{e^x, xe^x\}$$

und die allgemeine Lösung lautet

$$y \quad = \quad (c_1 x + c_0)\, e^x.$$

Den obigen Satz sieht man leicht ein, wenn man

$$\left(\frac{d}{dx} - \lambda\right)^n y \quad = \quad 0$$

betrachtet. Für $j \in \mathbb{N}_0$ ist

$$\left(\frac{d}{dx} - \lambda\right) x^j e^{\lambda x} \quad = \quad j x^{j-1} e^{\lambda x}$$

und somit

$$\left(\frac{d}{dx} - \lambda\right)^n x^j e^{\lambda x} \quad = \quad j(j-1)\ldots(j-n+1) x^{j-n} e^{\lambda x}$$

Für $j < n$ ist einer der Vorfaktoren Null, und somit ist $x^j e^{\lambda x}$ Lösung der homogenen Differentialgleichung für alle $0 \le j < n$.

Die inhomogene Gleichung

$$a_n y^{(n)} + a_{n-1} y^{(n-1)} + \ldots + a_1 y' + a_0 y \quad = \quad b(x).$$

löst man wie üblich, indem man zuerst die allgemeine Lösung der homogenen Gleichung bestimmt. Man führt dann die Differentialgleichung n-ter Ordnung auf ein System von n Gleichungen erster Ordnung zurückführt und bestimmt hierfür eine spezielle Lösung. Die allgemeine Lösung ist dann gegeben als die Summe einer speziellen Lösung der inhomogenen Gleichung plus die allgemeine Lösung der homogenen Gleichung.

Für den speziellen Fall, daß der inhomogene Term von der Form

$$b(x) \;=\; ce^{\kappa x}$$

ist, wobei κ **keine** Nullstelle des zur homogenen Gleichung assozierten Polynoms $P(\lambda)$ ist, d.h.

$$P(\kappa) \;\neq\; 0,$$

gibt es ein vereinfachtes Verfahren: Die Funktion

$$f(x) \;=\; \frac{c}{P(\kappa)} e^{\kappa x}$$

ist **eine** spezielle Lösung der inhomogenen Gleichung.

6.5.3 Der harmonische Oszillator

Gesucht ist eine reellwertige Funktion $x(t)$ die folgende Differentialgleichung erfüllt:

$$\frac{d^2}{dt^2} x + 2\mu \frac{d}{dt} x + \omega_0^2 x \;=\; A \cos\left(\omega_{ext} t\right)$$

In der Physik ist es üblich, Ableitungen nach der Zeit mit einem Punkt zu kennzeichnen:

$$\dot{x}(t) \;=\; \frac{d}{dt} x(t).$$

Somit schreibt man die obige Gleichug auch oft als

$$\ddot{x} + 2\mu\dot{x} + \omega_0^2 x \;=\; A \cos\left(\omega_{ext} t\right)$$

ω_0 ist die Eigenfrequenz des harmonischen Oszillators, der Term proportional zu μ beschreibt die Dämpfung, die rechte Seite eine äußere treibende Kraft. Wir können $\mu \geq 0$ und $\omega_0 > 0$ annehmen.

Es ist einfacher, zunächst eine komplexwertige Funktion $z(t)$ zu suchen, die die Differentialgleichung

$$\ddot{z} + 2\mu\dot{z} + \omega_0^2 z \;=\; Ae^{i\omega_{ext} t}$$

erfüllt, und dann den Realteil zu nehmen:

$$x(t) \;=\; \mathrm{Re}\, z(t).$$

Wir betrachten zunächst die homogene Gleichung:

$$\ddot{z} + 2\mu\dot{z} + \omega_0^2 z \;=\; 0.$$

Das zugehörige Polynom lautet

$$P(\lambda) \quad = \quad \lambda^2 + 2\mu\lambda + \omega_0^2.$$

Wir bestimmen die Nullstellen von P. Hierzu betrachten wir die Fälle $\mu^2 < \omega_0^2$, $\mu^2 = \omega_0^2$ und $\mu^2 > \omega_0^2$ getrennt.

1. Fall: $\mu^2 < \omega_0^2$. Dieser Fall beschreibt eine kleine Dämpfung.

$$\lambda \quad = \quad -\mu \pm i\sqrt{\omega_0^2 - \mu^2} = -\mu \pm i\omega,$$

wobei wir

$$\omega \quad = \quad \sqrt{\omega_0^2 - \mu^2}$$

gesetzt haben. Wir erhalten somit das Lösungsfundamentalsystem

$$\varphi_1(t) = e^{-\mu t}e^{i\omega t}, \qquad \varphi_2(t) = e^{-\mu t}e^{-i\omega t},$$

und die allgemeine Lösung der homogenen Gleichung lautet in diesem Fall

$$\varphi(t) \quad = \quad c_1 e^{-\mu t}e^{i\omega t} + c_2 e^{-\mu t}e^{-i\omega t}.$$

2. Fall: $\mu^2 = \omega_0^2$. In diesem Fall erhalten wir eine doppelte Nullstelle

$$\lambda \quad = \quad -\mu = -\omega.$$

Das Lösungsfundamentalsystem lautet

$$\varphi_1(t) = e^{-\mu t}, \qquad \varphi_2(t) = te^{-\mu t},$$

und die allgemeine Lösung der homogenen Gleichung lautet in diesem Fall

$$\varphi(t) \quad = \quad c_1 e^{-\mu t} + c_2 t e^{-\mu t}.$$

Den Fall $\mu^2 = \omega_0^2$ bezeichnet man als aperiodischen Grenzfall.

3. Fall: $\mu^2 > \omega_0^2$. In diesem Fall erhalten wir reelle Nullstellen

$$\lambda \quad = \quad -\mu \pm \sqrt{\mu^2 - \omega_0^2}.$$

Beide Nullstellen sind negativ. Wir setzen

$$\mu_1 = \mu + \sqrt{\mu^2 - \omega_0^2}, \qquad \mu_2 = \mu - \sqrt{\mu^2 - \omega_0^2}.$$

Das Lösungsfundamentalsystem lautet

$$\varphi_1(t) = e^{-\mu_1 t}, \qquad \varphi_2(t) = e^{-\mu_2 t},$$

und die allgemeine Lösung der homogenen Gleichung lautet in diesem Fall

$$\varphi(t) \;\; = \;\; c_1 e^{-\mu_1 t} + c_2 e^{-\mu_2 t}.$$

Den Fall $\mu^2 > \omega_0^2$ bezeichnet man als Kriechfall.

Wir betrachten nun den Fall einer schwachen Dämpfung mit einer äußeren treibenden Kraft, d.h. für $\mu^2 < \omega_0^2$ die inhomogene Differentialgleichung

$$\ddot{z} + 2\mu\dot{z} + \omega_0^2 z \;\; = \;\; A e^{i\omega_{ext} t}.$$

Wir können annehmen, daß ω_{ext} reell und positiv ist. Wie bereits oben gezeigt, ist das Fundamentalsystem der homogenen Gleichung gegeben durch

$$\varphi_1(t) = e^{-\mu t} e^{i\omega t}, \qquad \varphi_2(t) = e^{-\mu t} e^{-i\omega t}, \quad \omega = \sqrt{\omega_0^2 - \mu^2}.$$

Für $\mu > 0$ haben wir immer

$$i\omega_{ext} \;\; \neq \;\; -\mu \pm i\omega,$$

somit ist $(i\omega_{ext})$ keine Nullstelle von $P(\lambda)$ und eine spezielle Lösung der inhomogenen Differentialgleichung ist

$$f_0(t) \;\; = \;\; \frac{A}{P(i\omega_{ext})} e^{i\omega_{ext} t} = \frac{A}{\omega_0^2 - \omega_{ext}^2 + 2i\mu\omega_{ext}} e^{i\omega_{ext} t} = A \frac{\omega_0^2 - \omega_{ext}^2 - 2i\mu\omega_{ext}}{\left(\omega_0^2 - \omega_{ext}^2\right)^2 + 4\mu^2\omega_{ext}^2} e^{i\omega_{ext} t}$$

Die allgemeine Lösung der inhomogenen Gleichung ist somit

$$z(t) \;\; = \;\; A \frac{\omega_0^2 - \omega_{ext}^2 - 2i\mu\omega_{ext}}{\left(\omega_0^2 - \omega_{ext}^2\right)^2 + 4\mu^2\omega_{ext}^2} e^{i\omega_{ext} t} + c_1 e^{-\mu t} e^{i\omega t} + c_2 e^{-\mu t} e^{-i\omega t}.$$

Es bleibt noch der Spezialfall $\mu = 0$ und $\omega = \omega_0 = \omega_{ext}$ zu diskutieren. Man bezeichnet diesen Fall als Resonanzfall. Dies entspricht der Differentialgleichung

$$\ddot{z} + \omega^2 z \;\; = \;\; A e^{i\omega t}.$$

Das Fundamentalsystem der homogenen Gleichung ist

$$\varphi_1(t) = e^{i\omega t}, \qquad \varphi_2(t) = e^{-i\omega t}.$$

Wir schreiben die Differentialgleichung zweiter Ordnung um auf ein System zweier Gleichungen erster Ordnung:

$$\begin{pmatrix} \dot{z}_0 \\ \dot{z}_1 \end{pmatrix} \;\; = \;\; \begin{pmatrix} 0 & 1 \\ -\omega^2 & 0 \end{pmatrix} \begin{pmatrix} z_0 \\ z_1 \end{pmatrix} + \begin{pmatrix} 0 \\ A e^{i\omega t} \end{pmatrix}.$$

Das Lösungsfundamentalsystem für dieses System lautet

$$\Phi(t) = \begin{pmatrix} e^{i\omega t} & e^{-i\omega t} \\ i\omega e^{i\omega t} & -i\omega e^{-i\omega t} \end{pmatrix}$$

Das Inverse dieser Matrix lautet

$$\Phi^{-1}(t) = \frac{i}{2\omega} \begin{pmatrix} -i\omega e^{-i\omega t} & -e^{-i\omega t} \\ -i\omega e^{i\omega t} & e^{i\omega t} \end{pmatrix}$$

Wir berechnen nun

$$\begin{aligned} \int_0^t \Phi^{-1}(\tilde{t})\vec{b}(\tilde{t})d\tilde{t} &= \frac{i}{2\omega}\int_0^t d\tilde{t} \begin{pmatrix} -i\omega e^{-i\omega\tilde{t}} & -e^{-i\omega\tilde{t}} \\ -i\omega e^{i\omega\tilde{t}} & e^{i\omega\tilde{t}} \end{pmatrix} \begin{pmatrix} 0 \\ Ae^{i\omega\tilde{t}} \end{pmatrix} \\ &= \frac{iA}{2\omega}\int_0^t d\tilde{t} \begin{pmatrix} -1 \\ e^{2i\omega\tilde{t}} \end{pmatrix} = A \begin{pmatrix} -\frac{i}{2\omega}t \\ -\frac{1}{4\omega^2}\left(1-e^{2i\omega t}\right) \end{pmatrix} \end{aligned}$$

Damit ergibt sich eine spezielle Lösung zu

$$\Phi(t)\int_0^t \Phi^{-1}(\tilde{t})\vec{b}(\tilde{t})d\tilde{t} = A \begin{pmatrix} -\frac{i}{2\omega}te^{i\omega t} - \frac{1}{4\omega^2}e^{-i\omega t} + \frac{1}{4\omega^2}e^{i\omega t} \\ \frac{1}{2}te^{i\omega t} + \frac{i}{4\omega}e^{-i\omega t} - \frac{i}{4\omega}e^{i\omega t} \end{pmatrix}$$

Wir benötigen die erste Komponente dieses Vektors, außerdem können wir Terme, die nur Linearkombinationen der homogenen Lösungen sind, zur Bestimmung der speziellen Lšoung weglassen. Wir erhalten also

$$f(t) = -\frac{i}{2\omega}Ate^{i\omega t}$$

als eine spezielle Lösung der inhomogenen Differentialgleichung. Die allgemeine Lösung im Resonanzfall ist somit

$$z(t) = -\frac{i}{2\omega}Ate^{i\omega t} + c_1e^{i\omega t} + c_2e^{-i\omega t}.$$

Die Amplitude des ersten Terms wächst linear mit t an.

6.6 Exakte Differentialgleichungen und integrierende Faktoren

Eine gewöhnliche Differentialgleichung der Form

$$a(x,y) + b(x,y)y' = 0$$

nennt man **exakt**, falls es eine stetig differenzierbare Funktion $F(x,y)$ gibt, so daß

$$\begin{aligned} a(x,y) &= \frac{\partial F(x,y)}{\partial x}, \\ b(x,y) &= \frac{\partial F(x,y)}{\partial y} \end{aligned}$$

ist. Ist die Funktion $F(x,y)$ bekannt, so erhält man eine Lösung der Differentialgleichung durch Auflösen der impliziten Gelichung

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

nach y, wobei c eine Integrationskonstante ist. Es läßt sich zeigen, daß eine Funktion $F(x,y)$ genau dann existiert, falls

$$\frac{\partial}{\partial y}a(x,y) = \frac{\partial}{\partial x}b(x,y)$$

ist.

Beispiel: Wir betrachten die Differentialgleichung

$$3x^2 \tan y + x^3 \left(1+\tan^2 y\right) y' = 0.$$

Diese Differentialgleichung ist exakt, die Funktion $F(x,y)$ lautet

$$F(x,y) = x^3 \tan y,$$

wie man leicht durch Nachrechnen überprüft:

$$\begin{aligned} \frac{\partial}{\partial x}F(x,y) &= 3x^2 \tan y, \\ \frac{\partial}{\partial y}F(x,y) &= x^3 \left(1+\tan^2 y\right). \end{aligned}$$

Somit müssen wir die Gleichung

$$x^3 \tan y = c$$

lösen. Dies ergibt

$$y = \arctan\left(\frac{c}{x^3}\right).$$

Gilt für eine Differentialgleichung

$$a(x,y) + b(x,y)y' = 0$$

$$\frac{\partial}{\partial y}a(x,y) \neq \frac{\partial}{\partial x}b(x,y),$$

so ist sie nicht exakt. Findet man aber eine Funktion $c(x,y) \neq 0$, so daß

$$\frac{\partial}{\partial y}\left(a(x,y)c(x,y)\right) = \frac{\partial}{\partial x}\left(b(x,y)c(x,y)\right),$$

gilt, so ist die Differentialgleichung

$$a(x,y)c(x,y) + b(x,y)c(x,y)y' \quad = \quad 0$$

exakt. Da nach Voraussetzung $c(x,y) \neq 0$ für alle x und y ist, so ist eine Lösung dieser Differentialgleichung auch eine Lösung der ursprünglichen Gleichung und umgekehrt. Man bezeichnet die Funktion $c(x,y)$ als integrierenden Faktor.

Beispiel: Wir betrachten für $x > 0$ die Differentialgleichung

$$3\tan y + x\left(1 + \tan^2 y\right) y' \quad = \quad 0.$$

Diese Differentialgleiching ist nicht exakt, da

$$\frac{\partial}{\partial y}(3\tan y) = 3\left(1 + \tan^2 y\right) \quad \neq \quad \left(1 + \tan^2 y\right) = \frac{\partial}{\partial y} x\left(1 + \tan^2 y\right).$$

Allerdings ist die Funktion

$$c(x,y) \quad = \quad x^2$$

ein integrierender Faktor (und ungleich Null für $x > 0$). Multiplikation mit $c(x,y) = x^2$ ergibt die exakte Differentialgleichung

$$3x^2 \tan y + x^3\left(1 + \tan^2 y\right) y' \quad = \quad 0,$$

deren Lösung wir schon oben betrachtet haben.

7 Die Eulersche Gamma-Funktion

In diesem Abschnitt behandeln wir die Eulersche Gamma-Funktion.

Definition: Sei $x > 0$. Man definiert die Eulersche Gamma-Funktion durch das uneigentliche Integral

$$\Gamma(x) \quad = \quad \int\limits_0^\infty t^{x-1} e^{-t} dt.$$

Diese Funktion hat die Eigenschaft, daß

$$\begin{aligned} \Gamma(n+1) &= n! \quad \text{für alle } n \in \mathbb{N}, \\ \Gamma(x+1) &= x\Gamma(x) \quad \text{für alle } x > 0. \end{aligned}$$

Die zweite Behauptung beweisen wir mit partieller Integration:

$$\begin{aligned} \Gamma(x+1) &= \lim_{a\to 0, b\to\infty} \int\limits_a^b t^x e^{-t} dt = \lim_{a\to 0, b\to\infty} -t^x e^{-t}\Big|_a^b + \lim_{a\to 0, b\to\infty} x \int\limits_a^b t^{x-1} e^{-t} dt = x \int\limits_0^\infty t^{x-1} e^{-t} dt \\ &= x\Gamma(x). \end{aligned}$$

Um $\Gamma(n+1) = n!$ zu zeigen, genügt es nun, da die Funktionalgleichung bereits bewiesen wurde, $\Gamma(1) = 1$ zu zeigen.

$$\Gamma(1) \quad = \quad \lim_{a\to 0, b\to\infty} \int\limits_a^b e^{-t} dt = -\lim_{a\to 0, b\to\infty} e^{-t}\Big|_a^b = 1.$$

Die Eulersche Gamma-Funktion läßt sich auch axiomatisch definieren. Hierzu betrachten wir zunächst den Begriff "logarithmisch konvex". Sei $I \subset \mathbb{R}$ ein Intervall und bezeichne $\mathbb{R}_+^*$ alle Zahlen $x \in \mathbb{R}$ mit $x > 0$. Wir betrachten eine Funktion $F : I \to \mathbb{R}_+^*$. Die Funktion F nennt man **logarithmisch konvex**, falls die Funktion $\ln F : I \to \mathbb{R}$ konvex ist. Übersetzt bedeutet dies, daß für alle $x, y \in I$ und $0 < \lambda < 1$ gilt:

$$F(\lambda x + (1-\lambda)y) \quad \leq \quad F(x)^\lambda F(y)^{1-\lambda}.$$

Satz von H. Bohr: Sei $F : \mathbb{R}_+^* \to \mathbb{R}_+^*$ eine Funktion mit den folgenden Eigenschaften:

1. $F(1) = 1$,
2. $F(x+1) = xF(x)$ für alle $x \in \mathbb{R}_+^*$,
3. F ist logarithmisch konvex.

Dann gilt $F(x) = \Gamma(x)$ für alle $x \in \mathbb{R}_+^*$.

Es läßt sich zeigen, daß sich die Eulersche Gamma-Funktion auf $\mathbb{C}$ fortsetzen läßt. Als Funktion auf $\mathbb{C}$ hat $\Gamma(z)$ Pole entlang der negativen reellen Achse bei $z = 0, -1, -2, -3, ...$.

Oft benötigt man die Taylorentwicklung der Gamma-Funktion um einen ganzzahligen Wert. Diese Entwicklung enthält man aus der Funktionalgleichung $\Gamma(x+1) = x\Gamma(x)$ indem man die Entwicklung um $x = n$, $n \in \mathbb{N}$ auf eine Einwicklung um $n = 1$ zurückführt. Die Entwicklung um $x = 1$ gewinnt man aus der Formel

$$\Gamma(1+x) = \exp\left(-\gamma_E x + \sum_{n=2}^{\infty} \frac{(-1)^n}{n} \zeta_n x^n\right).$$

Hierin ist γ_E die Eulersche Konstante definiert durch

$$\gamma_E = \lim_{n\to\infty}\left(\sum_{j=1}^{n} \frac{1}{j} - \ln n\right) = 0.5772156649...$$

und ζ_n der Wert der Riemannschen Zetafunktion an der Stelle $s = n$.

$$\zeta_n = \sum_{j=1}^{\infty} \frac{1}{j^n}.$$

Bemerkung: Die Riemannsche Zetafunktion ist für $s > 1$ definiert durch

$$\zeta(s) = \sum_{j=1}^{\infty} \frac{1}{j^s}.$$

Diese Funktion läßt sich auf $\mathbb{C}$ fortsetzen. Die Lage der Nullstellen dieser Funktion ist ein ungelöstes Problem der Mathematik.

Wir betrachten noch die Eulersche Beta-Funktion. Sie ist definiert für $x > 0$ und $y > 0$ durch

$$B(x,y) = \int_0^{\infty} \frac{t^{x-1}}{(1+t)^{x+y}} dt.$$

Durch Variablensubstitution läßt sich dieses Integral umschreiben auf

$$B(x,y) = \int_0^{1} t^{x-1}(1-t)^{y-1} dt.$$

Die Eulersche Beta-Funktion läßt sich durch Eulersche Gamma-Funktionen ausdrücken:

$$B(x,y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)}.$$

Wir betrachten noch eine Anwendung der Eulerschen Gamma-Funktion. Wir betrachten zwei Fogen (a_n) und (b_n) deren Folgenglieder alle ungleich Null sind. Wir bezeichnen diese Folgen als **asymptotisch gleich**, falls

$$\lim_{n\to\infty} \frac{a_n}{b_n} \quad = \quad 1$$

gilt. In diesem Fall schreiben wir

$$a_n \quad \sim \quad b_n.$$

Bemerkung: Es wird nicht vorausgesetzt, daß die Folgen konvergieren.

Satz (Stirling): Die Fakultät hat das asymptotische Verhalten

$$n! \quad \sim \quad \sqrt{2\pi n} \left(\frac{n}{e}\right)^n.$$

8 Asymptotisches Verhalten

Im letzten Abschnitt hatten wir bereits asymptotisch gleiche Folgen betrachtet. In diesem Abschnitt wollen wir dies noch etwas vertiefen. Wir betrachten das Verhalten zweier Funktionen $f(x)$ und $g(x)$ in der Umgebung eines Punktes $x = x_0$. Wir sind im wesentlichen an dem Fall interessiert, in dem beide Funktionen an diesem Punkte divergieren. Wir betrachten daher den Grenzwert $x \to x_0$. Der Fall $x \to \pm\infty$ ist ein Spezialfall mit $x_0 = \pm\infty$. Wir sagen, die beiden Funktionen haben an diesem Punkte das gleiche asymptotische Verhalten, falls

$$\lim_{x \to x_0} \frac{f(x)}{g(x)} = 1$$

gilt. Wir schreiben in diesem Fall

$$f(x) \sim g(x) \quad \text{für } x \to x_0.$$

Bemerkung: Sind beide Funktionen an der Stelle $x = x_0$ endlich und haben sie den gleichen Wert $f(x_0) = g(x_0)$, so haben sie natürlich auch trivialerweise das gleiche asymptotische Verhalten an dieser Stelle.

Wir führen noch die Notation mit einem großen O und einem kleinen o ein. Falls

$$\lim_{x \to x_0} \frac{f(x)}{g(x)} = 0$$

gilt, so schreibt man

$$f(x) = o(g(x)) \quad \text{für } x \to x_0.$$

Gibt es eine Konstante K, so daß

$$\lim_{x \to x_0} \frac{|f(x)|}{|g(x)|} \leq K$$

gilt, so schreibt man

$$f(x) = O(g(x)) \quad \text{für } x \to x_0.$$

Bemerkung: Die Relationen $\sim$ und o implizieren die Relation O.

Die formale Reihe

$$\sum_{j=0}^{\infty} f_j(x)$$

nennt man eine **asymptotische Entwicklung** der Funktion $f(x)$ um den Punkt x_0 falls für jedes n für $x \to x_0$

$$f(x) - \sum_{j=0}^{n} f_j(x) = o(f_n(x))$$

gilt. Dies bedeutet

$$\lim_{x \to x_0} \frac{f(x) - \sum_{j=0}^{n} f_j(x)}{f_n(x)} \quad = \quad 0, \quad \forall\, n.$$

9 Fehlerrechnung

Wir betrachten den Fall, daß eine Person A eine bestimmte Größe experimentell mißt. Sie führt diese Messung öfters durch. Aufgrund der experimentellen Meßungenauigkeit ergeben sich leicht unterschiedliche Werte. Zum Beispiel:

Meßreihe 1:

Messung	1	2	3	4	5	6	7	8	9	10	11	12	13
Ergebnis	2.6	2.3	2.5	2.3	2.6	2.4	2.2	2.3	2.4	2.5	2.6	2.8	2.7

Wir definieren den Mittelwert einer Meßreihe mit n Meßpunkten als

$$\bar{x} = \frac{1}{n}\sum_{j=1}^{n} x_j .$$

Für die obige Meßreihe ergibt sich

$$\bar{x} = 2.48$$

Wir betrachten nun weiter den Fall, daß eine Person B die gleiche Größe experimentell bestimmt. Person B verwendet allerdings eine schlechtere Meßapparatur und führt weniger Messungen durch. Person B erhält die folgenden Meßwerte:

Meßreihe 2:

Messung	1	2	3	4
Ergebnis	0.3	5.2	3.1	1.4

Der Mittelwert ergibt sich zu

$$\bar{x} = 2.48$$

Im zweiten Fall streuen die einzelnen Messungen wesentlich stärker als im ersten Fall. Es ist daher offensichtlich, daß das Ergebnis von Person A vertauenswürdiger als das Ergebnis von Person B ist. Wir wollen nun diese Aussage quantitativ machen und suchen ein Maß für die Streuung der Meßpunkte.

Wir beginnen mit einigen Definitionen:

$$\begin{aligned} \Omega &: \text{Ergebnismenge eines Zufallsexperiments,} \\ \text{Zufallsfunktion} &: \text{Funktion } X : \Omega \to \mathbb{R}, \end{aligned}$$

Wahrscheinlichkeitsfunktion einer Zufallsgröße X:

$$W : x \rightarrow P(\omega | X(\omega) = x) .$$

Erwartungswert einer Zufallsgröße: Nimmt die Zufallsgröße X die Werte x_1, x_2, ..., x_n an, so bezeichnet man mit

$$\mu(X) = \sum_{j=1}^{n} x_j W(x_j)$$

den Erwartungswert von X.

Satz: Entsprechen die einzelnen Messungen einzelnen unabhängigen Realisierungen eines Zufallsexperiments, so ist der Mittelwert $\bar{x}$ eine Schätzung für $\mu(X)$.

Die Varianz einer Zufallsgröße ist definiert durch

$$\mathrm{Var}(X) = \sum_{j=1}^{n} \left(x_j - \mu\right)^2 W(x_j).$$

Die Standardabweichung einer Zufallsgröße ist definiert durch

$$\sigma(X) = \sqrt{\mathrm{Var}(X)}.$$

Kennen wir den Erwartungswert μ einer Zufallsgröße und machen n Messungen x_j, so ist

$$\frac{1}{n} \sum_{j=1}^{n} \left(x_j - \mu\right)^2$$

eine Schätzfunktion für die Varianz.

Im allgemeinen ist μ aber nicht bekannt und man verwendet $\bar{x}$ als Schätzung für μ. In diesem Fall ist

$$S^2 = \frac{1}{n-1} \sum_{j=1}^{n} \left(x_j - \bar{x}\right)^2$$

eine Schätzfunktion für die Varianz der Zufallsgröße X.

Sätze über die Varianz: Sei $c \in \mathbb{R}$ und seien X_1, X_2, ..., X_n unabhängige Zufallsgrößen. Dann gilt:

$$\begin{aligned} \mathrm{Var}(cX) &= c^2\, \mathrm{Var}(X), \\ \mathrm{Var}(X_1 + X_2 + \ldots + X_n) &= \mathrm{Var}(X_1) + \mathrm{Var}(X_2) + \ldots + \mathrm{Var}(X_n). \end{aligned}$$

Insbesondere gilt

$$\mathrm{Var}\left(\frac{1}{n} \underbrace{(X + X + \ldots + X)}_{n \text{ mal}}\right) = \frac{1}{n^2} \left(\mathrm{Var}(X) + \mathrm{Var}(X) + \ldots + \mathrm{Var}(X)\right) = \frac{1}{n} \mathrm{Var}(X).$$

Varianz des Mittelwertes: Es interessiert in erster Linie nicht die Varianz der einzelnen Messungen $\mathrm{Var}(X)$, sondern die Varianz des Mittelwertes $\mathrm{Var}(\bar{X})$. Bei n Messungen gilt:

$$\mathrm{Var}(\bar{X}) = \frac{1}{n}\mathrm{Var}(X).$$

Somit erhält man als Schätzung für die Varianz des Mittelwertes

$$S_{\bar{X}}^2 = \frac{1}{n}S^2 = \frac{1}{n(n-1)}\sum_{j=1}^{n}\left(x_j - \bar{x}\right)^2.$$

Für die Standardabweichung erhält man

$$\sigma_{\bar{X}} = \sqrt{\frac{1}{n(n-1)}\sum_{j=1}^{n}\left(x_j - \bar{x}\right)^2}.$$

Somit findet man für die beiden oben aufgeführten Meßreihen:

$$\begin{aligned} \text{Meßreihe 1}: \quad & \sigma_{\bar{X}} = 0.05, \\ \text{Meßreihe 2}: \quad & \sigma_{\bar{X}} = 1.07. \end{aligned}$$

Es ist üblich mit dem Mittelwert auch immer die Standardabweichung anzugeben, also

$$\begin{aligned} \text{Meßreihe 1}: \quad & x = 2.48 \pm 0.05, \\ \text{Meßreihe 2}: \quad & x = 2.48 \pm 1.07. \end{aligned}$$

Zur Interpretation der Standardabweichung betrachten wir zunächst kontinuierliche Zufallsgrößen. Die Wahrscheinlichkeitsdichtefunktion $p(x)$ für eine kontinuierliche Zufallsgröße beschreibt

$$P\left(a < X \leq b\right) = \int_a^b p(x)\, dx.$$

Definition: Man nennt eine kontinuierliche Zufallsgröße **normalverteilt**, falls sie die Wahrscheinlichkeitsdichtefunktion

$$p(x) = \frac{1}{\sqrt{2\pi}}\frac{1}{\sigma}\exp\left(-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\right)$$

besitzt. Der Erwartungswert dieser normalverteilten Zufallsgröße ist μ, die Standardabweichung ist σ. Für eine normalverteilte Zufallsgröße gilt:

$$\begin{aligned} P\left(\mu - \sigma < X \leq \mu + \sigma\right) &\approx 68.27\%, \\ P\left(\mu - 2\sigma < X \leq \mu + 2\sigma\right) &\approx 95.45\%, \\ P\left(\mu - 3\sigma < X \leq \mu + 3\sigma\right) &\approx 99.73\%. \end{aligned}$$

9.1 Fehlerfortpflanzung

Problemstellung: Gesucht wird eine Größe $f = f(x,y)$ die von zwei weiteren Größen x und y abhängt. Die Funktion f wird als bekannt vorausgesetzt, die Größen x und y werden durch eine Messung mit Fehlern $x \pm \Delta x$ und $y \pm \Delta y$ bestimmt. Gesucht ist nun der Fehler für die Größe f.

Für die Größe f beginnen wir mit der Taylorentwicklung:

$$f(x+\Delta x, y+\Delta y) = f(x,y) + \frac{\partial f(x,y)}{\partial x}\Delta x + \frac{\partial f(x,y)}{\partial y}\Delta y + \ldots$$

Wir nehmen an, daß wir n Messungen für die Größen x und y haben, die einzelnen Meßwerte seien mit x_j und y_j bezeichnet. Somit haben wir auch n Ergebnisse für f. Für die Abweichung eines Einzelergebnisses vom Mittelwert gilt für kleine Abweichungen

$$f_j - \bar{f} = \frac{\partial f}{\partial x} \cdot (x_j - \bar{x}) + \frac{\partial f}{\partial y} \cdot (y_j - \bar{y}) + \ldots$$

Somit gilt für die Varianz:

$$\begin{aligned}
\sigma_f^2 &= \lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} \left(f_j - \bar{f}\right)^2 \\
&= \lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} \left[(x_j - \bar{x})^2 \left(\frac{\partial f}{\partial x}\right)^2 + (y_j - \bar{y})^2 \left(\frac{\partial f}{\partial y}\right)^2 + 2\,(x_j - \bar{x})\,(y_j - \bar{y}) \left(\frac{\partial f}{\partial x}\right) \left(\frac{\partial f}{\partial y}\right) \right]
\end{aligned}$$

Wir definieren die **Kovarianz** als

$$\mathrm{Cov}(x,y) = \sigma_{xy} = \lim_{n\to\infty} \frac{1}{n} \sum_{j=1}^{n} (x_j - \bar{x})\,(y_j - \bar{y})$$

Somit haben wir

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

Falls x und y unkorreliert sind, gilt $\sigma_{xy} = 0$ und somit

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

bzw.

$$\sigma_f = \sqrt{\left(\frac{\partial f}{\partial x}\right)^2 \sigma_x^2 + \left(\frac{\partial f}{\partial y}\right)^2 \sigma_y^2}.$$

Wir betrachten nun einige Beispiele für diese Formel:

1. $f = x + y$. In diesem Fall haben wir

$$\sigma_f = \sqrt{\sigma_x^2 + \sigma_y^2},$$

man sagt, die (absoluten) Fehler addieren sich quadratisch. Für $x = 15 \pm 3$ und $y = 17 \pm 4$ ergibt sich also $f = 32 \pm 5$.

2. $f = x \cdot y$. In diesem Fall findet man

$$\sigma_f = \sqrt{y^2\sigma_x^2 + x^2\sigma_y^2},$$

oder anders geschrieben

$$\frac{\sigma_f}{f} = \sqrt{\left(\frac{\sigma_x}{x}\right)^2 + \left(\frac{\sigma_y}{y}\right)^2}.$$

Bei einem Produkt addieren sich die relativen Fehler quadratisch.

3. $f = x - y$. Hier findet man wie bei einer Summe

$$\sigma_f = \sqrt{\sigma_x^2 + \sigma_y^2}.$$

4. $f = \frac{x}{y}$. In diesem Fall findet man

$$\sigma_f = \sqrt{\frac{1}{y^2}\sigma_x^2 + \frac{x^2}{y^4}\sigma_y^2}.$$

Schreibt man dies mit Hilfe der relativen Fehler erhält man wie beim Produkt

$$\frac{\sigma_f}{f} = \sqrt{\left(\frac{\sigma_x}{x}\right)^2 + \left(\frac{\sigma_y}{y}\right)^2}.$$

5. Zum Abschluss betrachten wir noch $f = x^a y^b$. Man erhält

$$\sigma_f = \sqrt{\left(ax^{a-1}y^b\right)^2\sigma_x^2 + \left(bx^a y^{b-1}\right)^2\sigma_y^2}$$

Auch hier empfiehlt es sich wieder, die Formel in relativen Fehler zu schreiben:

$$\frac{\sigma_f}{f} = \sqrt{a^2\left(\frac{\sigma_x}{x}\right)^2 + b^2\left(\frac{\sigma_y}{y}\right)^2}.$$

Wir hatten zuvor den Fall betrachtet, daß eine Größe durch zwei Meßreihen experimentell bestimmt wird:

$$\begin{aligned} \text{Meßreihe 1}: \quad & x = 2.48 \pm 0.05, \\ \text{Meßreihe 2}: \quad & x = 2.48 \pm 1.07. \end{aligned}$$

Es stellt sich nun die Frage, wie man diese Ergebnisse miteinander kombiniert. Etwas allgemeiner seien für eine Größe x n Messungen x_j mit Fehlern σ_j gegeben. Dann setzt man

$$x = \frac{\frac{1}{\sigma_1^2} x_1 + \frac{1}{\sigma_2^2} x_2 + \ldots + \frac{1}{\sigma_n^2} x_n}{\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} + \ldots + \frac{1}{\sigma_n^2}}, \qquad \sigma = \frac{1}{\sqrt{\frac{1}{\sigma_1^2} + \frac{1}{\sigma_2^2} + \ldots + \frac{1}{\sigma_n^2}}}.$$

Für das Beispiel hat man

$$\begin{aligned} x_1 &= 2.48, \qquad \sigma_1 = 0.05, \\ x_2 &= 2.48, \qquad \sigma_2 = 1.07. \end{aligned}$$

Man findet somit

$$x = 2.48, \qquad \sigma = 0.04995.$$

Die zweite Meßreihe liefert keinen wesentlichen Beitrag zur Verbesserung des Fehlers.

9.2 Fitten von Parametern

Wir betrachten die Situation, in dem in einem Experiment eine Meßkurve bestimmt wird, d.h. zu n bekannten Werten $x_1,...,x_n$ werden die Meßpunkte $y_1 \pm \sigma_1,...,y_n \pm \sigma_n$ mit den Fehlern $\sigma_1,...,\sigma_n$ bestimmt. Wir nehmen an, daß diese n Messungen unabhängig sind. Man sucht nun eine Funktion $y = f(x)$, die diese Meßreihe möglichst gut beschreibt. Da diese Funktion nicht bekannt ist, verwendet man einen Ansatz für diese Funktion mit freien Parametern α_1, ..., α_r:

$$y = f(x, \alpha).$$

Nun bildet man die Größe χ^2 wie folgt:

$$\chi^2(\alpha) = \sum_{j=1}^{n} \frac{\left(y_j - f(x_j, \alpha)\right)^2}{\sigma_j^2}.$$

Offensichtlich nimmt χ^2 Werte zwischen 0 und Unendlich an. Die Werte von α, die die Größe χ^2 minimieren, bezeichnet man als den besten Fit.