

Univ.-Prof. Dr. Mathias Richter, EIT
Mathias.Richter@unibw.de

Univ.-Prof. Dr. Dr. Stefan Schäffler, EIT
Stefan.Schaeffler@unibw.de

der Bundeswehr
Universität  **München**

Bachelor BAU, EIT, LRT

Mathematik I, II

Inhaltsverzeichnis

I	Mathematik 1	3
1	Zahlen und Vektoren	4
1.1	Mengen und Abbildungen	4
1.2	Reelle Zahlen	6
1.3	Die vollständige Induktion	8
1.4	Binomialkoeffizienten	10
1.5	Die Ebene	12
1.6	Vektoren	16
1.7	Komplexe Zahlen	23
2	Lineare Algebra	27
2.1	Gleichungssysteme und Matrizen	27
2.2	Matrizenmultiplikation	34
2.3	Vektorräume	39
2.4	Elementarmatrizen	46
2.5	Determinanten	51
2.6	Lineare Abbildungen und Eigenwerte	56
II	Mathematik 2	65
3	Analysis einer reellen Veränderlichen	66
3.1	Funktionen, Grenzwerte, Stetigkeit	66
3.2	Differentiation	82
3.3	Potenzreihen	93
3.4	Integration	98
4	Gewöhnliche Differentialgleichungen	107
4.1	Gewöhnliche Differentialgleichungen n-ter Ordnung	109
4.2	Gewöhnliche Differentialgleichungssysteme	117
4.3	Lineare Differentialgleichungssysteme mit konstanten Koeffizienten	120
4.4	Stabilität	125
5	Transformationen	132
5.1	Die Laplace-Transformation	132
5.2	Die Fourier-Transformation	134

Teil I

Mathematik 1

Kapitel 1

Zahlen und Vektoren

1.1 Mengen und Abbildungen

Der Begriff „Menge“ gehört zu den Grundbegriffen der Mathematik. Eine formale Definition dessen, was Mathematiker unter „Menge“ verstehen, ist sehr problematisch; daher begnügen wir uns mit einer umgangssprachlichen Beschreibung, die von GEORG CANTOR, dem Begründer der Mengenlehre, stammt:

Unter einer Menge verstehen wir jede Zusammenfassung von bestimmten, wohlunterschiedenen Objekten unserer Anschauung oder unseres Denkens zu einem Ganzen.

Wir werden Mengen mit Großbuchstaben bezeichnen. Die Objekte der Menge A werden Elemente von A genannt. Man schreibt $a \in A$ (bzw. $a \notin A$), wenn das Objekt a ein (bzw. kein) Element von A ist.

Neben der Aufzählung von Elementen bei endlichen Mengen zum Beispiel durch

$$A = \{a_1, a_2, \dots, a_n\}, \quad (1.1)$$

wobei die Reihenfolge keine Rolle spielt, werden Mengen sehr häufig durch die Angabe einer Eigenschaft \mathcal{E} , die genau allen Elementen einer zu betrachtenden Menge zukommt, beschrieben. Zum Beispiel könnte man die Menge $\{-\sqrt{2}, +\sqrt{2}\}$ auch durch $\{x \in \mathbb{R}; x^2 = 2\}$ darstellen. Im allgemeinen werden Mengen, deren Elemente eine spezielle Eigenschaft besitzen, folgendermaßen notiert:

$$\{x; x \text{ erfüllt die Eigenschaft } \mathcal{E}\} \quad \text{bzw.} \quad \{x \in G; x \text{ erfüllt die Eigenschaft } \mathcal{E}\}. \quad (1.2)$$

Wie das obige Beispiel zeigt, kann der Term „ x erfüllt die Eigenschaft \mathcal{E} “ auch durch eine mathematische Formel ausgedrückt werden. Die leere Menge \emptyset besitzt kein Element. Die Anzahl der Elemente einer Menge M heißt ihre Mächtigkeit und wird mit $|M|$ bezeichnet. Hat M unendlich viele Elemente, dann schreibt man $|M| = \infty$.

Eine Menge B heißt Teilmenge von A (in Zeichen: $B \subseteq A$), wenn jedes Element von B auch Element von A ist. Zwei Mengen A, B heißen gleich, in Zeichen: $A = B$, falls sowohl $A \subseteq B$ als auch $B \subseteq A$ gilt. Eine Menge B heißt echte Teilmenge von A (in Zeichen: $B \subset A$), wenn $B \subseteq A$ und $B \neq A$ gilt. Für zwei Mengen A, B sind der Durchschnitt $A \cap B$, die Vereinigung $A \cup B$ und das Komplement $A \setminus B$ definiert durch

$$A \cap B := \{x; x \in A \text{ und } x \in B\}, \quad (1.3)$$

$$A \cup B := \{x; x \in A \text{ oder } x \in B\}, \quad (1.4)$$

$$A \setminus B := \{x; x \in A \text{ und } x \notin B\}. \quad (1.5)$$

Das Zeichen „:=“ heißt „definitionsgemäß gleich“, wobei der Doppelpunkt bei dem zu definierenden Ausdruck steht. Besitzen zwei Mengen A, B kein gemeinsames Element, so heißen sie

„disjunkt“, d.h. es gilt $A \cap B = \emptyset$.

Als direktes (oder kartesisches) Produkt von n Mengen A_1, \dots, A_n bezeichnet man die Menge

$$A_1 \times \dots \times A_n := \{(a_1, \dots, a_n); a_i \in A_i \text{ für alle } i = 1, \dots, n\}. \quad (1.6)$$

Die Elemente (a_1, \dots, a_n) dieser Menge werden als „geordnete n -Tupel“ bezeichnet. Der Begriff „geordnet“ ist dadurch gerechtfertigt, dass gilt:

$$(a_1, \dots, a_n) = (a'_1, \dots, a'_n) \text{ genau dann, wenn } a_i = a'_i \text{ für alle } i = 1, \dots, n. \quad (1.7)$$

Beispiel(e) 1.1

- Für $A_1 = \{\alpha, \beta\}$ und $A_2 = \{1, 2, 3\}$ ergibt sich:

$$A_1 \times A_2 = \{(\alpha, 1), (\alpha, 2), (\alpha, 3), (\beta, 1), (\beta, 2), (\beta, 3)\}$$

und

$$A_2 \times A_1 = \{(1, \alpha), (2, \alpha), (3, \alpha), (1, \beta), (2, \beta), (3, \beta)\}$$

- Für $A_1 = \{\text{Latein}, \text{Physik}\}$ und $A_2 = \{1, 2, 3, 4, 5, 6\}$ ergibt sich:

$$\begin{aligned} A_1 \times A_2 = & \{(\text{Latein}, 1), (\text{Latein}, 2), (\text{Latein}, 3), (\text{Latein}, 4), \\ & (\text{Latein}, 5), (\text{Latein}, 6), (\text{Physik}, 1), (\text{Physik}, 2), \\ & (\text{Physik}, 3), (\text{Physik}, 4), (\text{Physik}, 5), (\text{Physik}, 6)\} \end{aligned}$$

Für den Fall $A_1 = A_2 = \dots = A_n = A$ schreibt man

$$A^n \text{ für } \underbrace{A \times \dots \times A}_{n\text{-mal}}. \quad (1.8)$$

Seien A und B zwei Mengen. Eine Abbildung (oder Funktion) f von A nach B wird durch

$$f : A \rightarrow B, \quad x \mapsto f(x) \quad (1.9)$$

dargestellt und ist eine Vorschrift, die jedem $x \in A$ genau ein Element $f(x) \in B$ zuordnet. Man nennt A den Definitionsbereich, B den Wertebereich und die Elemente von A werden als Argumente der Abbildung f bezeichnet; $f(x)$ heißt das Bild von x unter f (oder der Funktionswert von f an der Stelle x). Die Funktion f heißt injektiv, falls aus $f(x_1) = f(x_2)$ folgt: $x_1 = x_2$ für alle $x_1, x_2 \in A$. Die Injektivität einer Funktion hängt nicht nur von der Abbildungsvorschrift ab, sondern auch vom Definitionsbereich, wie die folgenden Beispiele zeigen:

Beispiel(e) 1.2

- $f_1 : \{1, 2, 3, 4\} \rightarrow \{1, 4, 9, 16\}, \quad x \mapsto x^2$.
Diese Funktion ist injektiv.
- $f_2 : \{-1, 1, 2, 3, 4\} \rightarrow \{1, 4, 9, 16\}, \quad x \mapsto x^2$.
Diese Funktion ist nicht injektiv, denn $f_2(-1) = f_2(1)$.

Für jede Teilmenge $C \subseteq A$ bezeichnet man für eine Abbildung f mit Definitionsbereich A die Menge

$$f(C) := \{f(x); x \in C\} \quad (1.10)$$

als Bild von C unter f . Das Bild $f(A)$ des gesamten Definitionsbereichs von f unter f wird als Wertemenge bezeichnet. Ist für unsere Abbildung

$$f : A \rightarrow B, \quad x \mapsto f(x) \tag{1.11}$$

die Wertemenge $f(A)$ gleich der Menge B , so heißt die Funktion f surjektiv.

Beispiel(e) 1.3

- $f_3 : \{1, 2, 3, 4\} \rightarrow \{1, 4, 9, 16\}, \quad x \mapsto x^2$.
Diese Funktion ist surjektiv.
- $f_4 : \{1, 2, 3, 4\} \rightarrow \{1, 2, \dots, 16\}, \quad x \mapsto x^2$.
Diese Funktion ist nicht surjektiv, denn $\{1, 4, 9, 16\} \neq \{1, 2, \dots, 16\}$.
- $f_5 : \{1, 2, 3, 4\} \rightarrow \{1, 4, 9\}, \quad x \mapsto x^2$.
Diese Abbildungsvorschrift ist nicht korrekt formuliert, denn B ist eine echte Teilmenge von $f_5(A)$.

Funktionen, die injektiv und surjektiv sind (in unseren Beispielen die Funktion f_1 bzw. f_3) werden als bijektiv bezeichnet. Für eine bijektive Funktion

$$f : A \rightarrow B, \quad x \mapsto f(x) \tag{1.12}$$

gibt es somit zu jedem $y \in B$ genau ein $x \in A$ mit: $f(x) = y$. Es existiert also eine weitere Abbildung

$$f^{-1} : B \rightarrow A, \quad y \mapsto f^{-1}(y) \tag{1.13}$$

mit: $f^{-1}(f(x)) = x$ für alle $x \in A$ und $f(f^{-1}(y)) = y$ für alle $y \in B$. Man bezeichnet f^{-1} als Umkehrfunktion oder Umkehrabbildung von f . Die auf jeder Menge A definierte Funktion

$$\text{Id}_A : A \rightarrow A, \quad x \mapsto x \tag{1.14}$$

heißt Identität auf A . Zwei Abbildungen

$$f : A \rightarrow B, \quad x \mapsto f(x) \quad \text{und} \quad g : C \rightarrow D, \quad z \mapsto g(z) \tag{1.15}$$

sind genau dann gleich (in Zeichen: $f = g$), falls $A = C$, $B = D$ und $f(x) = g(x)$ für alle $x \in A$. An dieser Definition der Gleichheit für Abbildungen sieht man, dass die Darstellung der Abbildungsvorschrift (etwa mit Hilfe des Zeichens „ x “ oder des Zeichens „ z “) keine Rolle spielt. Durch Einschränkung des Definitionsbereichs einer Abbildung $f : A \rightarrow B, x \mapsto f(x)$ auf eine Teilmenge $A_0 \subseteq A$ erhält man die Restriktion

$$f|_{A_0} : A_0 \rightarrow B, x \mapsto f(x). \tag{1.16}$$

1.2 Reelle Zahlen

Die Menge der reellen Zahlen bezeichnen wir mit \mathbb{R} . Besondere Teilmengen von \mathbb{R} sind

$$\mathbb{N} := \{1, 2, 3, \dots\} \quad \text{Menge der natürlichen Zahlen} \tag{1.17}$$

$$\mathbb{N}_0 := \mathbb{N} \cup \{0\} = \{0, 1, 2, 3, \dots\} \tag{1.18}$$

$$\mathbb{Z} := \{\dots - 3, -2, -1, 0, 1, 2, 3, \dots\} \quad \text{Menge der ganzen Zahlen} \tag{1.19}$$

$$\mathbb{Q} := \left\{ \frac{m}{n}; m, n \in \mathbb{Z}, n \neq 0 \right\} \quad \text{Menge der rationalen Zahlen.} \tag{1.20}$$

Neben den rationalen Zahlen gibt es in \mathbb{R} die irrationalen Zahlen, die durch unendliche, nicht-periodische Dezimalbrüche gekennzeichnet sind. Irrationale Zahlen sind zum Beispiel $\sqrt{2}$ oder

π . Neben der Aufteilung der reellen Zahlen in rationale und irrationale Zahlen gibt es auch die Aufteilung in algebraische und transzendente Zahlen.

Die reellen Zahlen sind durch eine Vergleichsrelation \leq vollständig geordnet, wobei für $x, y \in \mathbb{R}$ gilt:

$$x \leq y \quad :\Leftrightarrow \quad (x - y) \leq 0 \quad \text{bzw.} \quad (1.21)$$

$$x < y \quad :\Leftrightarrow \quad (x - y) < 0. \quad (1.22)$$

Das Zeichen „ $:\Leftrightarrow$ “ bedeutet „definitionsgemäß genau dann, wenn“. Mit Hilfe der Ordnungsstruktur „ \leq “ betrachten wir nun Intervalle reeller Zahlen:

- abgeschlossenes Intervall mit $a, b \in \mathbb{R}$, $a \leq b$:

$$[a, b] := \{x \in \mathbb{R}; a \leq x \leq b\} \quad (1.23)$$

- offenes Intervall mit $a, b \in \mathbb{R}$, $a < b$:

$$(a, b) := \{x \in \mathbb{R}; a < x < b\} \quad (1.24)$$

- halboffene Intervalle mit $a, b \in \mathbb{R}$, $a < b$:

$$(a, b] := \{x \in \mathbb{R}; a < x \leq b\} \quad \text{und} \quad [a, b) := \{x \in \mathbb{R}; a \leq x < b\}. \quad (1.25)$$

Es ist üblich, noch zwei Symbole „ $-\infty$ “ und „ ∞ “ (minus unendlich, unendlich) einzuführen und $-\infty < \infty$ festzulegen. Man definiert für $a \in \mathbb{R}$:

$$(-\infty, a] := \{x \in \mathbb{R}; x \leq a\} \quad (1.26)$$

$$(-\infty, a) := \{x \in \mathbb{R}; x < a\} \quad (1.27)$$

$$[a, \infty) := \{x \in \mathbb{R}; a \leq x\} \quad (1.28)$$

$$(a, \infty) := \{x \in \mathbb{R}; a < x\}. \quad (1.29)$$

Eine Menge S reeller Zahlen heißt nach oben beschränkt, wenn es eine reelle Zahl b gibt, sodass $S \subseteq (-\infty, b]$. In diesem Fall heißt b eine obere Schranke von S . Analog dazu nennt man eine reelle Zahl a eine untere Schranke von S und S nach unten beschränkt, falls $S \subseteq [a, \infty)$. Ist $S \subset \mathbb{R}$ nach oben beschränkt, so nennt man die kleinste obere Schranke s das Supremum von S (in Zeichen: $s = \sup\{S\}$). Offensichtlich braucht $\sup\{S\}$ nicht zur Menge S zu gehören. Es ist nicht selbstverständlich, dass die Menge der oberen Schranken von $S \subseteq (-\infty, b]$ eine kleinste Zahl enthält. Dies gehört zu den Grundannahmen, die man bei der axiomatischen Einführung der reellen Zahlen fordert. Damit hat auch jede nach unten beschränkte Menge $U \subseteq [a, \infty)$ eine größte untere Schranke (ein Infimum $\inf\{U\}$). Hat S ein größtes Element, dann nennt man dies das Maximum von S (in Zeichen: $m = \max\{S\}$), hat S ein kleinstes Element, dann nennt man dies das Minimum von S (in Zeichen $m = \min\{S\}$). Selbst wenn S beschränkt ist, braucht es kein Maximum und kein Minimum von S zu geben.

Beispiel(e) 1.4

- $\sup\{[a, b]\} = \sup\{(a, b)\} = \max\{[a, b]\} = b$. Das Maximum von (a, b) existiert nicht.
- $\sup\{\{x \in \mathbb{Q}; x^2 < 2\}\} = \sqrt{2}$
- Für $S := \{1 + \frac{1}{n}; n \in \mathbb{N}\}$ gilt: $\sup\{S\} = \max\{S\} = 2$, $\inf\{S\} = 1$, $\min\{S\}$ existiert nicht.

Der Betrag $|a|$ einer reellen Zahl ist definiert durch

$$|a| := \begin{cases} a & \text{falls } a \geq 0 \\ -a & \text{falls } a < 0 \end{cases}. \quad (1.30)$$

Offensichtlich gilt für $a, b \in \mathbb{R}$:

$$-|a| \leq a \leq |a|, \quad (1.31)$$

$$|-a| = |a|, \quad (1.32)$$

$$|ab| = |a| \cdot |b|, \quad (1.33)$$

$$\left| \frac{a}{b} \right| = \frac{|a|}{|b|}, \quad b \neq 0, \quad (1.34)$$

$$|a| \leq b \iff -b \leq a \leq b. \quad (1.35)$$

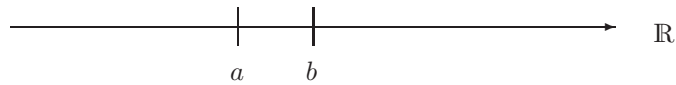
Aus $-|a| \leq a \leq |a|$ und $-|b| \leq b \leq |b|$ folgt:

$$-(|a| + |b|) \leq a + b \leq |a| + |b| \quad (1.36)$$

und damit die Dreiecksungleichung

$$|a + b| \leq |a| + |b|. \quad (1.37)$$

Betrachtet man zwei reelle Zahlen a, b auf der Zahlengeraden,



so ist $|a - b|$ der Abstand der zu a und b gehörenden Punkte auf der Zahlengeraden.

1.3 Die vollständige Induktion

Die vollständige Induktion ist ein Beweisschema, mit dessen Hilfe häufig Aussagen $A(n)$, die für jede ganze Zahl $n \geq n_0$ gelten sollen, bewiesen werden. Ein Beweis mit vollständiger Induktion gliedert sich in zwei Schritte:

1. Induktionsbeginn: Man zeigt, dass $A(n)$ für $n = n_0$ gilt.
2. Schluss von n auf $(n + 1)$: Für beliebiges $n \geq n_0$ setzt man die Gültigkeit von $A(n)$ voraus (Induktionsvoraussetzung) und leitet daraus die Gültigkeit von $A(n + 1)$ ab.

Können beide Schritte ausgeführt werden, dann gilt $A(n)$ für alle $n \geq n_0$; denn nach dem ersten Schritt gilt $A(n_0)$, nach dem zweiten Schritt $A(n_0 + 1)$ und ebenso (immer wieder Schluss von n auf $(n + 1)$) $A(n_0 + 2)$, $A(n_0 + 3)$, ...

Beispiel(e) 1.5

- $1 + 2 + 3 + \dots + n = \frac{n(n+1)}{2} \quad (n \in \mathbb{N})$

Beweis mit vollständiger Induktion:

Induktionsbeginn $n = 1$: $1 = \frac{1 \cdot 2}{2} \quad \checkmark$

Schluss von n auf $(n+1)$: Sei $n \in \mathbb{N}$ beliebig und die Formel richtig für dieses n . Dann folgt:

$$1 + 2 + 3 + \dots + n + (n+1) = \frac{n(n+1)}{2} + (n+1) = \frac{(n+1)(n+2)}{2}$$

Das ist die zu beweisende Formel für $(n+1)$. Sie gilt somit für alle $n \in \mathbb{N}$.

q.e.d.

- Für alle reellen Zahlen $h \geq -1$ und alle $n \in \mathbb{N}$ gilt:

$$(1+h)^n \geq 1+n \cdot h \quad (1.38)$$

Beweis mit vollständiger Induktion:

Induktionsbeginn $n = 1$: $(1+h)^1 \geq 1+h \quad \checkmark$

Schluss von n auf $(n+1)$: Aus $(1+h)^n \geq 1+n \cdot h$ folgt durch Multiplikation mit $(1+h) \geq 0$:

$$(1+h)^{n+1} \geq (1+n \cdot h) \cdot (1+h) = 1 + (n+1) \cdot h + nh^2 \geq 1 + (n+1) \cdot h$$

q.e.d.

Das Prinzip der vollständigen Induktion kann auch zur Definition einer Größe D_n für alle ganzen Zahlen $n \geq n_0$ verwendet werden. Man spricht in diesem Zusammenhang von einer rekursiven Definition:

1. D_{n_0} wird festgelegt
2. Man sieht D_k für $n_0 \leq k \leq n$ als bekannt an und drückt D_{n+1} durch D_{n_0}, \dots, D_n aus.

Beispiel(e) 1.6

- Die Potenzen a^n für eine beliebige Basis $a \in \mathbb{R}$ und Exponenten $n \in \mathbb{N}_0$ werden definiert durch

$$a^0 := 1, \quad a^{n+1} := a \cdot a^n. \quad (1.39)$$

- Die Fakultäten $n!$ (lies: n Fakultät) werden definiert durch

$$0! := 1, \quad (n+1)! := n!(n+1). \quad (1.40)$$

So ist etwa $6! = 720$ und $9! = 362880$

Als Permutation einer Menge A bezeichnet man jede bijektive Abbildung

$$f: A \rightarrow A$$

von A auf sich. Ist A endlich mit den n verschiedenen Elementen $\{a_1, \dots, a_n\}$, so interpretiert man eine Permutation von A als Zuordnung der a_i zu n verschiedenen, mit $1, \dots, n$ durchnu-

rierten Plätzen, so dass auf jeden Platz genau ein Element kommt. Für jedes $n \in \mathbb{N}$ gilt der Satz:

„Jede n -elementige Menge besitzt genau $n!$ verschiedene Permutationen“.

Der Beweis wird mit vollständiger Induktion geführt (Übung).

Für die Summe und das Produkt der Zahlen $a_m, a_{m+1}, \dots, a_n \in \mathbb{R}$ schreibt man:

$$a_m + a_{m+1} + \dots + a_n = \sum_{k=m}^n a_k, \quad a_m \cdot a_{m+1} \cdot \dots \cdot a_n = \prod_{k=m}^n a_k. \quad (1.41)$$

Eine rekursive Definition ersetzt die Punkte:

$$\sum_{k=m}^m a_k := a_m, \quad \sum_{k=m}^{n+1} a_k := \left(\sum_{k=m}^n a_k \right) + a_{n+1}. \quad (1.42)$$

Man beachte die Unabhängigkeit vom Summationsindex

$$\sum_{k=m}^n a_k = \sum_{i=m}^n a_i, \quad (1.43)$$

ferner:

$$\sum_{k=m}^n a_k = \sum_{k=m}^l a_k + \sum_{k=l+1}^n a_k \quad (m \leq l < n). \quad (1.44)$$

Analoges gilt für das Produktzeichen.

1.4 Binomialkoeffizienten

Für zwei ganze Zahlen n, k mit $0 \leq k \leq n$ bezeichnet man die natürliche Zahl

$$\binom{n}{k} := \frac{n!}{k!(n-k)!} \quad (\text{lies: } n \text{ über } k) \quad (1.45)$$

als Binomialkoeffizient. Der Binomialkoeffizient gibt die Anzahl der verschiedenen k -elementigen Teilmengen einer Menge mit n Elementen an. Es gibt also genau $\binom{n}{k}$ Möglichkeiten, aus n Objekten genau k auszuwählen.

Beispiel(e) 1.7

Im Lotto „6 aus 49“ gibt es

$$\binom{49}{6} := 13983816 \quad (1.46)$$

verschiedene Möglichkeiten, „6 Richtige“ zu ziehen.

Aus der Definition der Binomialkoeffizienten folgen sofort einige Eigenschaften:

$$\binom{n}{0} = \binom{n}{n} = 1, \quad \binom{n}{1} = \binom{n}{n-1} = n, \quad \binom{n}{2} = \binom{n}{n-2} = \frac{n(n-1)}{2}, \quad (1.47)$$

$$\binom{n+1}{k} = \binom{n}{k-1} + \binom{n}{k} \quad (1.48)$$

Damit lassen sich die Binomialkoeffizienten durch Addition berechnen:

Beginnend mit $n = 1$ schreibt man die Zahlen $\binom{n}{0}, \binom{n}{1}, \dots, \binom{n}{n}$ in eine Zeile, dabei ist $\binom{n}{0} = \binom{n}{n} = 1$ festgelegt. Die Zahlen $\binom{n+1}{0}, \dots, \binom{n+1}{n+1}$ werden gemäß der Formel

$$\binom{n+1}{k} = \binom{n}{k-1} + \binom{n}{k} \tag{1.49}$$

berechnet. Auf diese Weise entsteht das sogenannte Pascal-Dreieck:

$\binom{0}{0}$					1							
$\binom{1}{0}$	$\binom{1}{1}$				1	1						
		\ + /										
$\binom{2}{0}$	$\binom{2}{1}$	$\binom{2}{2}$				1	2	1				
		\ + / \ + /										
$\binom{3}{0}$	$\binom{3}{1}$	$\binom{3}{2}$	$\binom{3}{3}$			1	3	3	1			
		\ + / \ + / \ + /										
$\binom{4}{0}$	$\binom{4}{1}$	$\binom{4}{2}$	$\binom{4}{3}$	$\binom{4}{4}$			1	4	6	4	1	
		\ + / \ + / \ + / \ + /										
$\binom{5}{0}$	$\binom{5}{1}$	$\binom{5}{2}$	$\binom{5}{3}$	$\binom{5}{4}$	$\binom{5}{5}$	1	5	10	10	5	1	(1.50)

Dieses Schema ist nach **BLAISE PASCAL** (1623-1662), einem bedeutenden Universalgenie, benannt. Zu seinen herausragenden Leistungen gehören:

- Untersuchungen zum Phänomen des Luftdruckes (Barometerversuche). Die physikalische Größe „Druck“ wird in Pascal [1 Pa] gemessen.
- Konstruktion und Herstellung der ersten funktionstüchtigen Rechenmaschine. Die Programmiersprache **PASCAL** ist nach Blaise Pascal benannt.
- Planung und Durchführung der ersten Omnibuslinie (Einweihung am 18. März 1662).
- Zusammen mit Fermat Begründung der Stochastik.
- Pascal ist mit Augustinus bis heute der bedeutendste christliche Philosoph.

In Verbindung mit Binomialkoeffizienten ist die folgende „binomische Formel“ von Interesse: Für beliebige reelle Zahlen $a, b \in \mathbb{R}$ und jede ganze Zahl $n \geq 0$ gilt:

$$(a + b)^n = \sum_{k=0}^n \binom{n}{k} a^{n-k} b^k. \tag{1.51}$$

Der Beweis wird mit vollständiger Induktion geführt:

Induktionsbeginn: $n = 0$: $(a + b)^0 = 1 = \binom{0}{0} a^0 \cdot b^0$

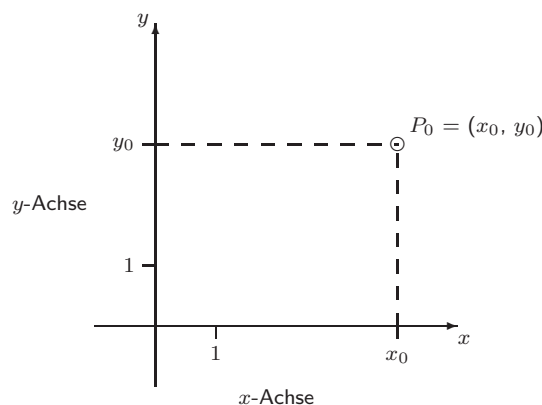
Schluss von n auf $(n + 1)$:

$$\begin{aligned}
(a+b)^{n+1} &= (a+b)^n \cdot (a+b) \stackrel{\text{Ind.-Voraussetzung}}{=} \\
&= \sum_{k=0}^n \binom{n}{k} \cdot a^{n-k} \cdot b^k \cdot (a+b) = \\
&= \sum_{k=0}^n \binom{n}{k} \cdot a^{n-k+1} \cdot b^k + \sum_{k=0}^n \binom{n}{k} \cdot a^{n-k} \cdot b^{k+1} = \\
&= a^{n+1} + \sum_{k=1}^n \binom{n}{k} \cdot a^{n-k+1} \cdot b^k + \sum_{k=0}^{n-1} \binom{n}{k} \cdot a^{n-k} \cdot b^{k+1} + b^{n+1} = \\
&= a^{n+1} + \sum_{k=1}^n \binom{n}{k} \cdot a^{n-k+1} \cdot b^k + \\
&\quad + \sum_{k=1}^n \binom{n}{k-1} \cdot a^{n-(k-1)} \cdot b^k + b^{n+1} = \\
&= a^{n+1} + \sum_{k=1}^n \left\{ \binom{n}{k} + \binom{n}{k-1} \right\} \cdot a^{n-k+1} \cdot b^k + b^{n+1} = \\
&= \binom{n+1}{0} a^{n+1} + \sum_{k=1}^n \binom{n+1}{k} \cdot a^{n-k+1} b^k + \binom{n+1}{n+1} b^{n+1} = \\
&= \sum_{k=0}^{n+1} \binom{n+1}{k} \cdot a^{n+1-k} \cdot b^k
\end{aligned}$$

q.e.d.

1.5 Die Ebene

In einer Ebene E entsteht ein kartesisches Koordinatensystem (benannt nach RENÉ DESCARTES (1596-1650), philosophischer Gegenspieler von Blaise Pascal) durch Vorgabe eines Punktes O und zweier aufeinander senkrechter Zahlengeraden, der x -Achse und der y -Achse, deren Nullpunkt jeweils in O liegt. Dabei muss die y -Achse durch eine positive Drehung (gegen den Uhrzeigersinn) aus der x -Achse hervorgehen. Fällt man für einen beliebigen Punkt P_0 in der Ebene die Lote auf die Achsen, so bestimmen die beiden Fußpunkte die x - bzw. y -Koordinate x_0 bzw. y_0 von P_0 und man schreibt $P_0 = (x_0, y_0)$. Der Punkt $O = (0, 0)$ heißt Nullpunkt oder Ursprung des Koordinatensystems.



(1.52)

Nach Festlegung eines kartesischen Koordinatensystems gibt es zu jedem Zahlenpaar $(x, y) \in \mathbb{R}^2$ genau einen Punkt $P \in E$ mit $P = (x, y)$ und umgekehrt.

Man kann somit Teilmengen von \mathbb{R}^2 (etwa Lösungsmengen von Gleichungen bzw. Ungleichungen) als Punktmengen in E veranschaulichen und umgekehrt geometrische Gebilde durch Funktionen, Gleichungen oder Ungleichungen als Teilmengen des \mathbb{R}^2 beschreiben.

Sei $I \subseteq \mathbb{R}$ und $f : I \rightarrow \mathbb{R}$ eine Funktion. Die Punktmenge

$$G_f := \{(x, y); x \in I \text{ und } y = f(x)\} = \{(x, f(x)); x \in I\} \quad (1.53)$$

in der mit einem kartesischen Koordinatensystem versehenen Ebene heißt der Graph der Funktion f oder die Kurve $y = f(x)$.

Beispiel(e) 1.8

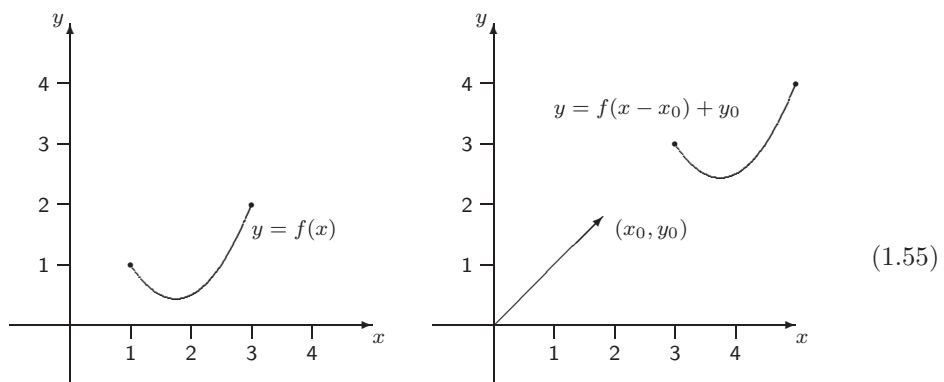
- Für eine gegebene Funktion

$$f : [a, b] \rightarrow \mathbb{R}, \quad x \mapsto f(x)$$

entsteht der Graph der Funktion

$$g : [a + x_0, b + x_0] \rightarrow \mathbb{R} \quad x \mapsto f(x - x_0) + y_0 \quad (1.54)$$

mit $(x_0, y_0) \in \mathbb{R}^2$ durch diejenige Parallelverschiebung des Graphen von f , die einen Punkt (x, y) nach $(x + x_0, y + y_0)$ verschiebt. Dabei bleibt das Koordinatensystem fest.



- Sei $C \subseteq E$ eine Kurve und - nach Wahl eines kartesischen Koordinatensystems - $X = (x, y)$ ein beliebiger Punkt auf C .

Wird die zwischen x und y bestehende Abhängigkeit auf die Form

$$F(x, y) = 0 \quad (1.56)$$

gebracht, d.h. $C = \{(x, y); F(x, y) = 0\}$, dann nennt man „ $F(x, y) = 0$ “ die Gleichung der Kurve C :

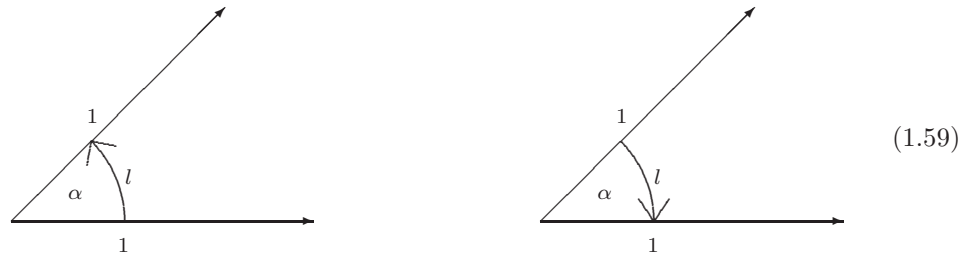
- (i) die Gerade durch die Punkte $A = (a_1, a_2)$ und $B = (b_1, b_2)$ ($A \neq B$) hat die Gleichung:

$$(b_2 - a_2)(x - a_1) - (b_1 - a_1)(y - a_2) = 0. \quad (1.57)$$

- (ii) die Kreisfläche (mit Rand) um $A = (a_1, a_2)$ mit dem Radius r wird beschrieben durch die Ungleichung

$$(x - a_1)^2 + (y - a_2)^2 \leq r^2 \quad (1.58)$$

Ein Winkel α entsteht durch Drehung eines Zeigers um einen gegebenen Punkt der Ebene. Die Länge des zugehörigen Einheitsbogens sei l .



Wir nennen l , bzw. $-l$, das Bogenmaß von α und schreiben $\alpha = l$, bzw. $\alpha = -l$, wenn die Drehung im positiven (gegen Uhrzeigersinn) bzw. im Uhrzeigersinn erfolgt. Ein Winkel von α° besitzt das Bogenmaß

$$x = \frac{\pi}{180^\circ} \cdot \alpha \tag{1.60}$$

mit der Kreiszahl

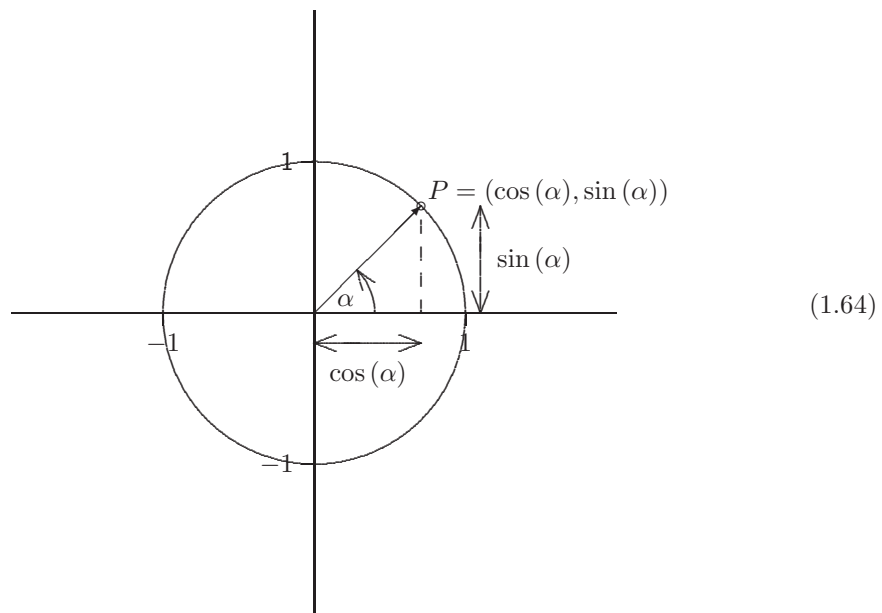
$$\pi = 3.14159265358979 \dots \tag{1.61}$$

Wird in der mit kartesischen Koordinaten versehenen Ebene der vom Ursprung zum Punkt $(1, 0)$ weisende Zeiger um den Winkel α gedreht (α im Bogenmaß), dann bewegt sich die Spitze auf dem Einheitskreis (um O) bis zu einem Punkte P , dessen Koordinaten mit $P = (\cos(\alpha), \sin(\alpha))$ bezeichnet werden. Die derart für alle $\alpha \in \mathbb{R}$ erklärten Funktionen

$$\cos : \mathbb{R} \rightarrow [-1, 1], \quad \alpha \mapsto \cos(\alpha) \tag{1.62}$$

$$\sin : \mathbb{R} \rightarrow [-1, 1], \quad \alpha \mapsto \sin(\alpha) \tag{1.63}$$

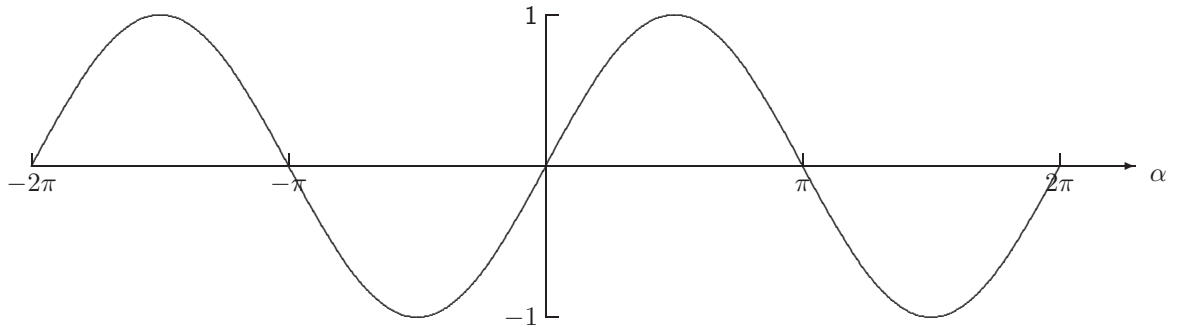
heißen Cosinus- bzw. Sinusfunktionen. Dabei ist zu beachten, dass mit jeder vollen Umdrehung das Bogenmaß des Winkels um 2π , also dem Umfang des Einheitskreises, vergrößert wird.



Für alle $\alpha \in \mathbb{R}$ gilt

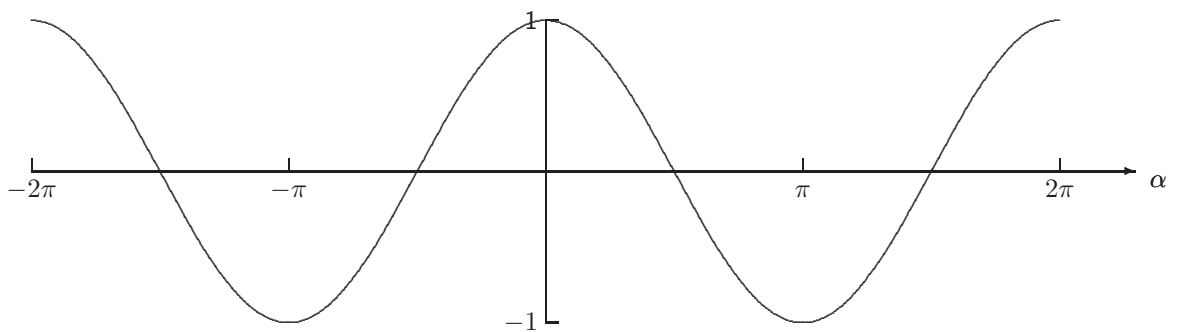
$$(\sin(\alpha))^2 + (\cos(\alpha))^2 = 1. \quad (1.65)$$

Graph der Sinusfunktion:



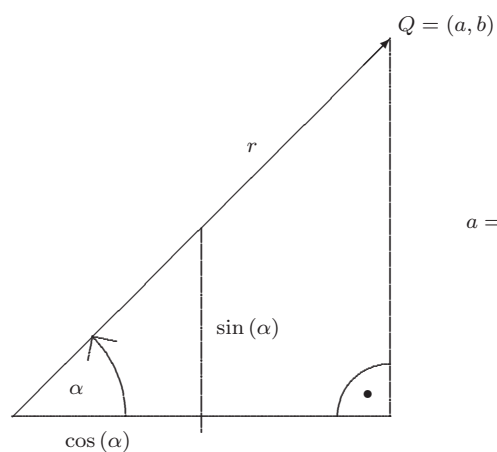
(1.66)

Graph der Cosinusfunktion:



(1.67)

Hat der sich drehende Zeiger die Länge r , so gilt



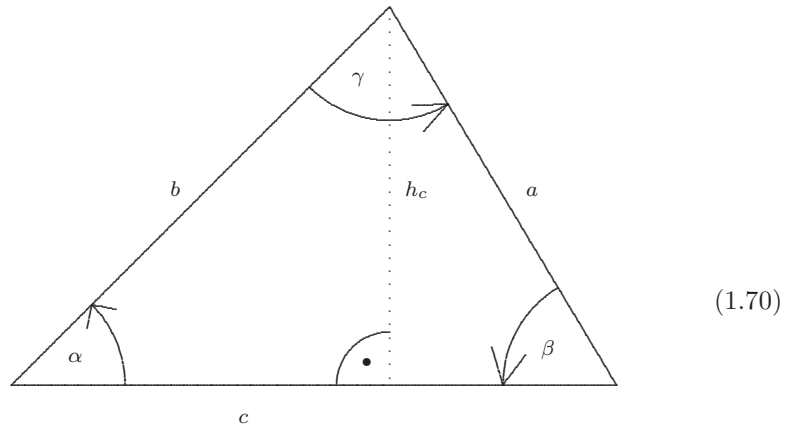
$$a = r \cdot \cos(\alpha), \quad b = r \cdot \sin(\alpha)$$

(1.68)

Somit gilt im rechtwinkligen Dreieck mit der Hypotenuse r , der Ankathete a und der Gegenkathete b :

$$\frac{a}{r} = \cos(\alpha), \quad \frac{b}{r} = \sin(\alpha). \quad (1.69)$$

Für ein beliebiges Dreieck



gilt ferner der Cosinussatz:

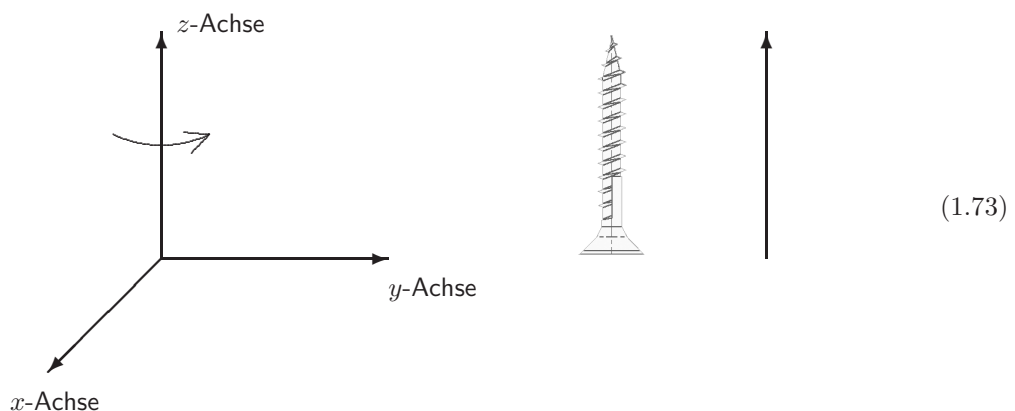
$$a^2 = b^2 + c^2 - 2bc \cos(\alpha) \tag{1.71}$$

und der Sinussatz:

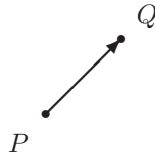
$$\frac{\sin(\alpha)}{a} = \frac{\sin(\beta)}{b} = \frac{\sin(\gamma)}{c} . \tag{1.72}$$

1.6 Vektoren

Kartesische Koordinatensysteme im Raum bestehen aus drei sich in einem Punkt O (Nullpunkt oder Ursprung) rechtwinklig schneidenden Zahlengeraden gleicher Längeneinheit und jeweils mit dem Nullpunkt im Schnittpunkt O . Man bezeichnet sie als x -, y - und z -Achse derart, dass diese Achsen ein Rechtssystem bilden, d.h. die Drehung der positiven x -Achse um 90° in die positive y -Achse muss zusammen mit einer Verschiebung in Richtung der positiven z -Achse eine Rechtsschraubung darstellen.



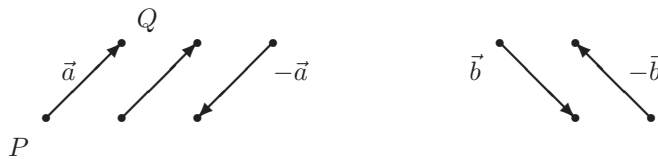
Zu je zwei Punkten P und Q des Raumes gibt es genau eine Parallelverschiebung aller Punkte, die P nach Q abbildet. Diese Verschiebung wird mit \vec{PQ} bezeichnet und heißt „Vektor von P nach Q “. Das lateinische „vector“ bedeutet „Träger“ : \vec{PQ} trägt P nach Q . Der Vektor \vec{PQ} wird dargestellt durch einen Pfeil von P nach Q .



Wird unter \vec{PQ} ein Punkt R nach S verschoben, dann hat offenbar \vec{RS} dieselbe Wirkung wie \vec{PQ} , d.h. $\vec{RS} = \vec{PQ}$. Die Länge des Pfeiles \vec{PQ} ist der Abstand zwischen P und Q . Zwei gleich lange und gleich gerichtete Pfeile im Raum stellen somit denselben Vektor dar.

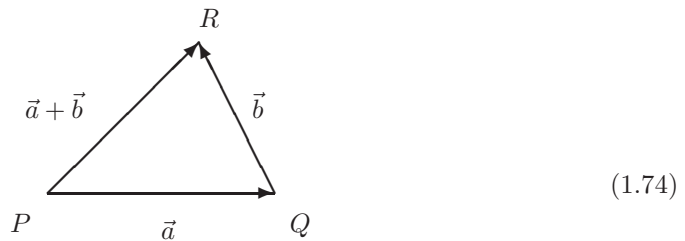
Anstelle „der Pfeil \vec{a} repräsentiert einen Vektor“ sagt man „ \vec{a} ist ein Vektor“ und berücksichtigt, dass \vec{a} im Raum frei parallel verschoben werden kann und nicht an einen festen Anfangspunkt gebunden ist.

Den zu \vec{a} gleich langen, aber entgegengesetzt gerichteten Vektor bezeichnen wir mit $-\vec{a}$; er macht die durch \vec{a} bewirkte Parallelverschiebung rückgängig. Für $\vec{a} = \vec{PQ}$ gilt: $-\vec{a} = \vec{QP}$:



Üblicherweise wird der Nullvektor $\vec{0}$ eingeführt, der eine Verschiebung bezeichnet, bei der sich nichts bewegt; d.h. $\vec{0} = \vec{PP}$.

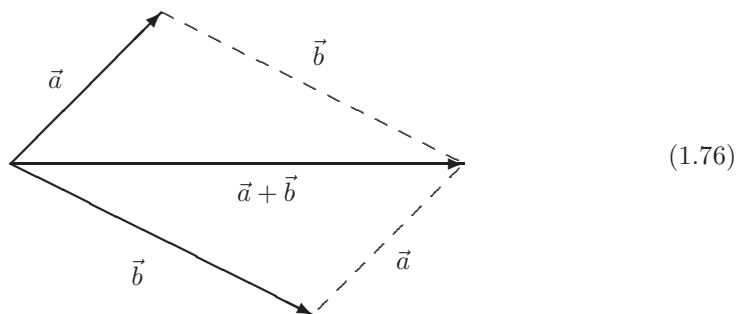
Führt man zwei Parallelverschiebungen, erst $\vec{a} = \vec{PQ}$, dann $\vec{b} = \vec{QR}$, hintereinander aus, so ergibt sich wieder eine Parallelverschiebung, $\vec{c} = \vec{PR}$



Wir nennen \vec{c} die Summe von \vec{a} und \vec{b} und schreiben dafür

$$\vec{c} = \vec{a} + \vec{b}. \tag{1.75}$$

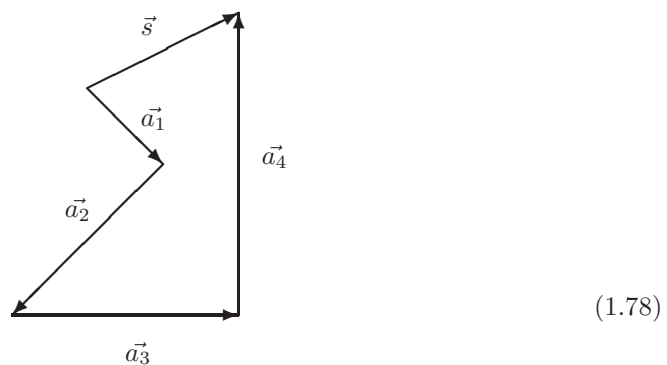
Haben die Pfeile \vec{a} und \vec{b} den gleichen Anfangspunkt, so gewinnt man die Summe $\vec{a} + \vec{b}$ (geometrisch) nach der Parallelogrammregel:



Offenbar gelten für beliebige Vektoren $\vec{a}, \vec{b}, \vec{c}$ die folgenden Rechenregeln:

$$\begin{aligned} \vec{a} + (-\vec{a}) &= \vec{0} \\ \vec{a} + \vec{0} &= \vec{a} \\ \vec{a} + \vec{b} &= \vec{b} + \vec{a} \quad (\text{Kommutativgesetz}) \\ (\vec{a} + \vec{b}) + \vec{c} &= \vec{a} + (\vec{b} + \vec{c}) \quad (\text{Assoziativgesetz}) \end{aligned} \tag{1.77}$$

Die Summe $\vec{a}_1 + \vec{a}_2 + \dots + \vec{a}_n$ ist derjenige Vektor \vec{s} , der vom Anfangspunkt zum Endpunkt der aus $\vec{a}_1, \vec{a}_2, \dots, \vec{a}_n$ gebildeten Vektorkette zeigt.



Die Differenz von Vektoren wird erklärt durch

$$\vec{a} - \vec{b} := \vec{a} + (-\vec{b}). \tag{1.79}$$

Zu einer reellen Zahl $\alpha \geq 0$ und einem Vektor \vec{a} bezeichnet $\alpha \cdot \vec{a}$ (oder kürzer: $\alpha\vec{a}$) denjenigen Vektor, der dieselbe Richtung und Orientierung hat wie \vec{a} , aber die α -fache Länge. Im Fall $\alpha < 0$ setzt man $\alpha \cdot \vec{a} = -(|\alpha| \cdot \vec{a})$. Sonderfälle dieser Festlegung sind $0\vec{a} = \vec{0}$ und $\alpha\vec{0} = \vec{0}$ für jede reelle Zahl α und jeden Vektor \vec{a} . Es gelten die folgenden Rechenregeln ($\alpha, \beta \in \mathbb{R}$):

$$\begin{aligned} \alpha(\beta\vec{a}) &= (\alpha\beta)\vec{a} \\ \alpha(\vec{a} + \vec{b}) &= \alpha\vec{a} + \alpha\vec{b} \\ (\alpha + \beta)\vec{a} &= \alpha\vec{a} + \beta\vec{a}. \end{aligned} \tag{1.80}$$

Die Länge eines Vektors \vec{a} (das ist für $\vec{a} = \vec{PQ}$ die Länge der Strecke $[PQ]$) nennt man seinen „Betrag“ und schreibt dafür $|\vec{a}|$ oder $\|\vec{a}\|$.

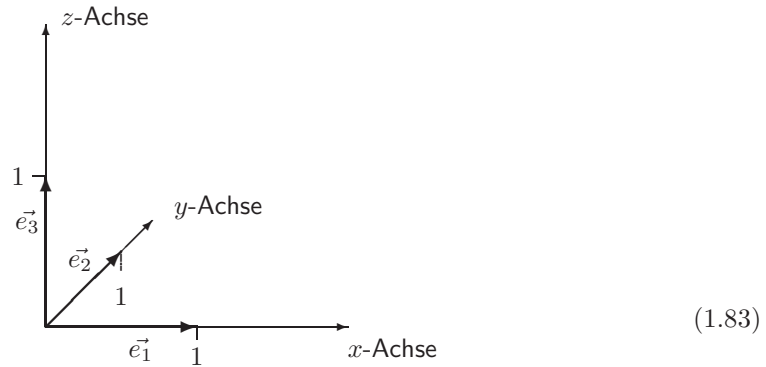
Offensichtlich gilt:

$$|\alpha\vec{a}| = |\alpha| \cdot |\vec{a}|, \quad \text{insbesondere } |-\vec{a}| = |\vec{a}| \tag{1.81}$$

$$|\vec{a} + \vec{b}| \leq |\vec{a}| + |\vec{b}| \quad (\text{Dreiecksungleichung}). \tag{1.82}$$

Der Nullvektor hat keine positive Länge, d.h. $|\vec{0}| = 0$. Ein Vektor vom Betrag 1 heißt „Einheitsvektor“.

Legt man im Raum ein kartesisches Koordinatensystem fest, so werden neben dem Ursprung O gleichzeitig drei Einheitsvektoren $\vec{e}_1, \vec{e}_2, \vec{e}_3$ in Richtung der positiven x -, y - und z -Achse ausgezeichnet.



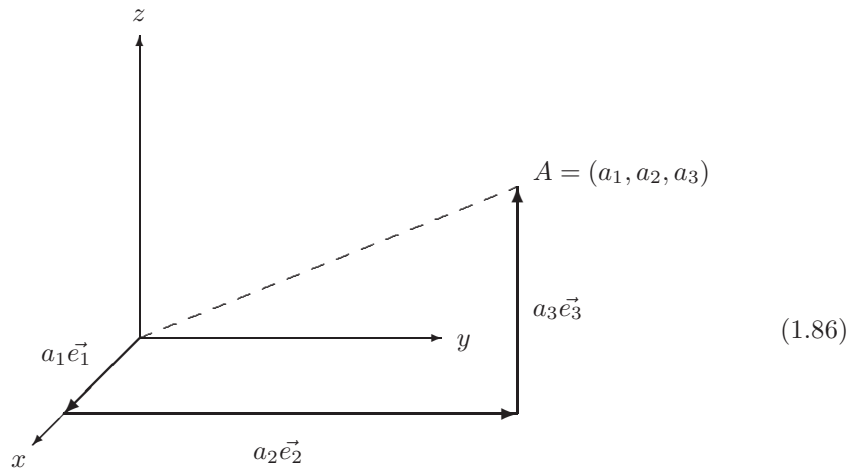
Wir nennen $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$ eine kartesische Basis und bezeichnen das Koordinatensystem mit $(O, \vec{e}_1, \vec{e}_2, \vec{e}_3)$. Der Vektor $\vec{a} = \vec{OA}$ heißt Ortsvektor des Punktes $A = (a_1, a_2, a_3)$; er ist eindeutig zerlegbar als Summe

$$\vec{a} = a_1\vec{e}_1 + a_2\vec{e}_2 + a_3\vec{e}_3. \tag{1.84}$$

Abkürzend schreibt man bei festgelegtem Koordinatensystem

$$\vec{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} : \Leftrightarrow \vec{a} = a_1\vec{e}_1 + a_2\vec{e}_2 + a_3\vec{e}_3 = \vec{OA} \quad \text{mit} \quad A = (a_1, a_2, a_3). \tag{1.85}$$

Man nennt $a_i\vec{e}_i$ die Komponente von \vec{a} in \vec{e}_i -Richtung ($i = 1, 2, 3$) und die Zahlen $a_i \in \mathbb{R}$ die Koordinaten des Vektors \vec{a} bezüglich $(O, \vec{e}_1, \vec{e}_2, \vec{e}_3)$.



Für einen Vektor \vec{a} in allgemeiner Lage, etwa $\vec{a} = \vec{PQ}$ mit $P = (p_1, p_2, p_3)$ und $Q = (q_1, q_2, q_3)$ ergibt sich:

$$\begin{aligned} \vec{PQ} &= \vec{OQ} - \vec{OP} = q_1\vec{e}_1 + q_2\vec{e}_2 + q_3\vec{e}_3 - p_1\vec{e}_1 - p_2\vec{e}_2 - p_3\vec{e}_3 = \\ &= \begin{pmatrix} q_1 - p_1 \\ q_2 - p_2 \\ q_3 - p_3 \end{pmatrix}. \end{aligned} \tag{1.87}$$

Die Summe von Vektoren und die skalaren Vielfachen lassen sich aus der Koordinatendarstellung bezüglich $(O, \vec{e}_1, \vec{e}_2, \vec{e}_3)$ sehr einfach berechnen:

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = \begin{pmatrix} a_1 + b_1 \\ a_2 + b_2 \\ a_3 + b_3 \end{pmatrix}, \quad \alpha \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} \alpha a_1 \\ \alpha a_2 \\ \alpha a_3 \end{pmatrix}. \quad (1.88)$$

Ferner gilt mit dem Satz von Pythagoras:

$$|\vec{a}| = \sqrt{a_1^2 + a_2^2 + a_3^2}, \quad \text{falls } \vec{a} = a_1 \vec{e}_1 + a_2 \vec{e}_2 + a_3 \vec{e}_3. \quad (1.89)$$

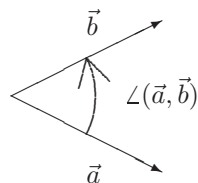
Beispiel(e) 1.9

$$\text{Für } \vec{a} = \begin{pmatrix} 2 \\ 0 \\ -3 \end{pmatrix}, \vec{b} = \begin{pmatrix} 0.5 \\ -1 \\ 6 \end{pmatrix} \text{ gilt: } \vec{a} + \vec{b} = \begin{pmatrix} 2.5 \\ -1 \\ 3 \end{pmatrix},$$

$$5\vec{a} = \begin{pmatrix} 10 \\ 0 \\ -15 \end{pmatrix}, 2\vec{a} - 3\vec{b} = \begin{pmatrix} 2.5 \\ 3 \\ -24 \end{pmatrix} \text{ und}$$

$$|\vec{a}| = \sqrt{13}, |\vec{b}| = \sqrt{37.25}, |\vec{a} + \vec{b}| = \sqrt{16.25}.$$

Trägt man zwei von \vec{O} verschiedene Vektoren \vec{a}, \vec{b} von einem Punkt P aus ab, so bezeichnet man den kleineren der beiden positiv gemessenen Winkel, den die Pfeile \vec{a} und \vec{b} im Scheitel P bilden, als Winkel zwischen \vec{a} und \vec{b} (in Zeichen: $\angle(\vec{a}, \vec{b})$), wobei $0 \leq \angle(\vec{a}, \vec{b}) \leq \pi$.



$$(1.90)$$

Offensichtlich gilt:

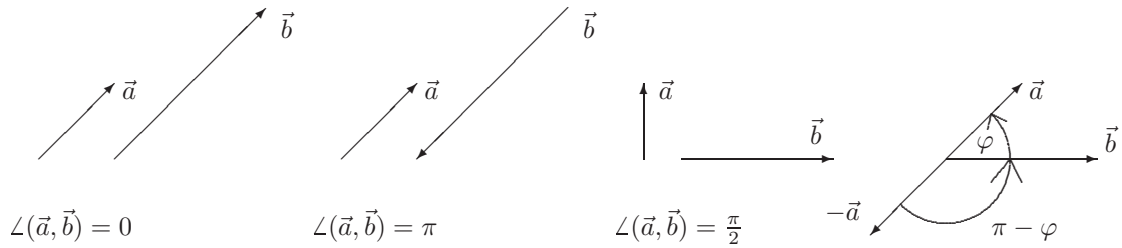
$$\angle(\vec{a}, \vec{b}) = \angle(\vec{b}, \vec{a}) \quad (1.91)$$

$$\angle(\vec{a}, t\vec{a}) = 0, \quad \text{falls } t > 0 \quad (1.92)$$

$$\angle(\vec{a}, t\vec{a}) = \pi, \quad \text{falls } t < 0 \quad (1.93)$$

$$\angle(-\vec{a}, \vec{b}) = \pi - \angle(\vec{a}, \vec{b}) \quad (1.94)$$

Der Vektor \vec{a} heißt orthogonal (senkrecht) zu \vec{b} (in Zeichen: $\vec{a} \perp \vec{b}$), wenn $\angle(\vec{a}, \vec{b}) = \frac{\pi}{2}$. Obwohl zwischen dem Nullvektor \vec{O} und \vec{a} kein Winkel erklärt wird, sagt man dennoch zur Vermeidung von Fallunterscheidungen, dass \vec{O} orthogonal zu jedem beliebigen Vektor \vec{a} ist, also $\vec{O} \perp \vec{a}$.



Das Skalarprodukt $\langle \vec{a} | \vec{b} \rangle$ der Vektoren \vec{a} und \vec{b} ist definiert durch

$$\langle \vec{a} | \vec{b} \rangle := \begin{cases} |\vec{a}| \cdot |\vec{b}| \cos(\angle(\vec{a}, \vec{b})), & \text{falls } \vec{a} \neq \vec{0} \text{ und } \vec{b} \neq \vec{0} \\ 0 & \text{sonst.} \end{cases} \quad (1.95)$$

Das Skalarprodukt wird auch als inneres Produkt bezeichnet.

Rechenregeln für das Skalarprodukt:

$$\begin{aligned} \langle \vec{a} | \vec{b} \rangle &= \langle \vec{b} | \vec{a} \rangle && \text{(Kommutativgesetz)} \\ \langle \alpha \vec{a} | \vec{b} \rangle &= \langle \vec{a} | \alpha \vec{b} \rangle \\ &= \alpha \langle \vec{a} | \vec{b} \rangle && \text{für alle } \alpha \in \mathbb{R} \\ \langle \vec{a} + \vec{b} | \vec{c} \rangle &= \langle \vec{a} | \vec{c} \rangle + \langle \vec{b} | \vec{c} \rangle && \text{(Distributivgesetz)} \\ \langle \vec{a} | \vec{b} \rangle = 0 &\iff \vec{a} \perp \vec{b} \\ \sqrt{\langle \vec{a} | \vec{a} \rangle} &= |\vec{a}|. \end{aligned} \quad (1.96)$$

Vektoren \vec{a}, \vec{b} bezüglich einer kartesischen Basis $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$ ermöglichen eine einfache Berechnung von $\langle \vec{a} | \vec{b} \rangle$ und $\cos(\angle(\vec{a}, \vec{b}))$:

Sei $\vec{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$ und $\vec{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}$, so gilt:

$$\begin{aligned} \langle \vec{a} | \vec{b} \rangle &= \langle a_1 \cdot \vec{e}_1 + a_2 \cdot \vec{e}_2 + a_3 \cdot \vec{e}_3 | b_1 \cdot \vec{e}_1 + b_2 \cdot \vec{e}_2 + b_3 \cdot \vec{e}_3 \rangle \\ &= a_1 b_1 \langle \vec{e}_1 | \vec{e}_1 \rangle + (a_1 b_2 + a_2 b_1) \langle \vec{e}_1 | \vec{e}_2 \rangle + \\ &\quad + (a_1 b_3 + a_3 b_1) \langle \vec{e}_1 | \vec{e}_3 \rangle + (a_2 b_3 + a_3 b_2) \langle \vec{e}_2 | \vec{e}_3 \rangle + \\ &\quad + a_2 b_2 \langle \vec{e}_2 | \vec{e}_2 \rangle + a_3 b_3 \langle \vec{e}_3 | \vec{e}_3 \rangle \\ &= a_1 b_1 + a_2 b_2 + a_3 b_3. \end{aligned} \quad (1.97)$$

$$\cos(\angle(\vec{a}, \vec{b})) = \frac{\langle \vec{a} | \vec{b} \rangle}{|\vec{a}| \cdot |\vec{b}|} \quad (1.98)$$

$$= \frac{a_1 b_1 + a_2 b_2 + a_3 b_3}{\sqrt{a_1^2 + a_2^2 + a_3^2} \cdot \sqrt{b_1^2 + b_2^2 + b_3^2}}. \quad (1.99)$$

Das Vektorprodukt $\vec{a} \times \vec{b}$ zweier Vektoren \vec{a} und \vec{b} ist der Vektor mit folgenden Eigenschaften:

- (i) $\vec{a} \times \vec{b} = \vec{0}$, falls $\vec{a} = \vec{0}$, $\vec{b} = \vec{0}$ oder \vec{a} parallel zu \vec{b} .
- (ii) In allen anderen Fällen ist $\vec{a} \times \vec{b}$ derjenige Vektor, der

- (1) senkrecht auf \vec{a} und \vec{b} steht
- (2) mit dem $(\vec{a}, \vec{b}, \vec{a} \times \vec{b})$ ein Rechtssystem darstellt und
- (3) dessen Betrag gleich dem Flächeninhalt F des von \vec{a}, \vec{b} aufgespannten Parallelogramms ist.

Für den Betrag des Vektors $\vec{a} \times \vec{b}$ gilt somit

$$|\vec{a} \times \vec{b}| = |\vec{a}| \cdot |\vec{b}| \cdot \sin(\angle(\vec{a}, \vec{b})). \quad (1.100)$$

Rechenregeln für das Vektorprodukt:

$$\begin{aligned} \vec{a} \times \vec{a} &= \vec{0} \\ \vec{a} \times \vec{b} &= -(\vec{b} \times \vec{a}) \\ \alpha(\vec{a} \times \vec{b}) &= (\alpha\vec{a}) \times \vec{b} = \vec{a} \times (\alpha\vec{b}) \quad \text{für alle } \alpha \in \mathbb{R} \\ \vec{a} \times (\vec{b} + \vec{c}) &= (\vec{a} \times \vec{b}) + (\vec{a} \times \vec{c}) \\ \vec{a} \times \vec{b} = \vec{0} &\iff \vec{a} = \vec{0} \quad \text{oder} \quad \vec{b} = \vec{0} \quad \text{oder} \quad \vec{a} \text{ parallel zu } \vec{b} \\ |\vec{a} \times \vec{b}|^2 &= |\vec{a}|^2 \cdot |\vec{b}|^2 - (\langle \vec{a}, \vec{b} \rangle)^2. \end{aligned} \quad (1.101)$$

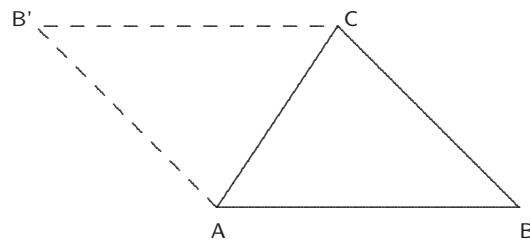
Sei nun $(\vec{e}_1, \vec{e}_2, \vec{e}_3)$ eine kartesische Basis und $\vec{a} = a_1\vec{e}_1 + a_2\vec{e}_2 + a_3\vec{e}_3$, $\vec{b} = b_1\vec{e}_1 + b_2\vec{e}_2 + b_3\vec{e}_3$, so gilt:

$$\begin{aligned} \vec{a} \times \vec{b} &= (a_1 \cdot \vec{e}_1 + a_2 \cdot \vec{e}_2 + a_3 \cdot \vec{e}_3) \times (b_1 \cdot \vec{e}_1 + b_2 \cdot \vec{e}_2 + b_3 \cdot \vec{e}_3) = \\ &= a_1b_1 \underbrace{(\vec{e}_1 \times \vec{e}_1)}_{\vec{0}} + a_1b_2 \underbrace{(\vec{e}_1 \times \vec{e}_2)}_{\vec{e}_3} + a_1b_3 \underbrace{(\vec{e}_1 \times \vec{e}_3)}_{-\vec{e}_2} + \\ &\quad + a_2b_1 \underbrace{(\vec{e}_2 \times \vec{e}_1)}_{-\vec{e}_3} + a_2b_2 \underbrace{(\vec{e}_2 \times \vec{e}_2)}_{\vec{0}} + a_2b_3 \underbrace{(\vec{e}_2 \times \vec{e}_3)}_{\vec{e}_1} + \\ &\quad + a_3b_1 \underbrace{(\vec{e}_3 \times \vec{e}_1)}_{\vec{e}_2} + a_3b_2 \underbrace{(\vec{e}_3 \times \vec{e}_2)}_{-\vec{e}_1} + a_3b_3 \underbrace{(\vec{e}_3 \times \vec{e}_3)}_{\vec{0}} = \\ &= \begin{pmatrix} a_2b_3 - a_3b_2 \\ a_3b_1 - a_1b_3 \\ a_1b_2 - a_2b_1 \end{pmatrix}. \end{aligned} \quad (1.102)$$

Beispiel(e) 1.10

Der Flächeninhalt eines Dreiecks ABC ist

$$F = \frac{1}{2} |\vec{AB} \times \vec{AC}|, \quad (1.103)$$



$$(1.104)$$

denn die Fläche des Parallelogramms $ABCB'$ ist gerade $|\vec{AB} \times \vec{AC}|$.

1.7 Komplexe Zahlen

In der mit einem kartesischen (x, y) -Koordinatensystem versehenen Ebene stellen die Punkte der x -Achse die reellen Zahlen dar. Wir gehen dazu über, auch alle anderen Punkte der Ebene als „Zahlen“ aufzufassen. Dazu schreiben wir den Punkt $z = (x, y)$ in der Form

$$z = x + iy \quad (1.105)$$

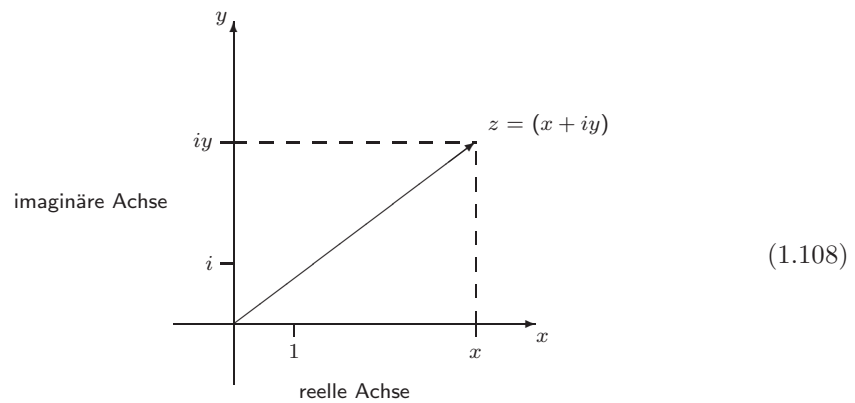
und nennen ihn eine komplexe Zahl mit Realteil

$$\operatorname{Re}(z) := x \quad (1.106)$$

und Imaginärteil

$$\operatorname{Im}(z) := y. \quad (1.107)$$

Die x -Achse heißt reelle Achse, die y -Achse wird imaginäre Achse genannt.



Abkürzend schreibt man

$$x = x + i0 = (x, 0) \quad (1.109)$$

(das sind die reellen Zahlen auf der reellen Achse) und

$$iy = 0 + iy = (0, y) \quad (1.110)$$

(das sind die rein imaginären Zahlen). Die Menge aller komplexen Zahlen wird mit \mathbb{C} bezeichnet:

$$\mathbb{C} := \{x + iy; x, y \in \mathbb{R}\}. \quad (1.111)$$

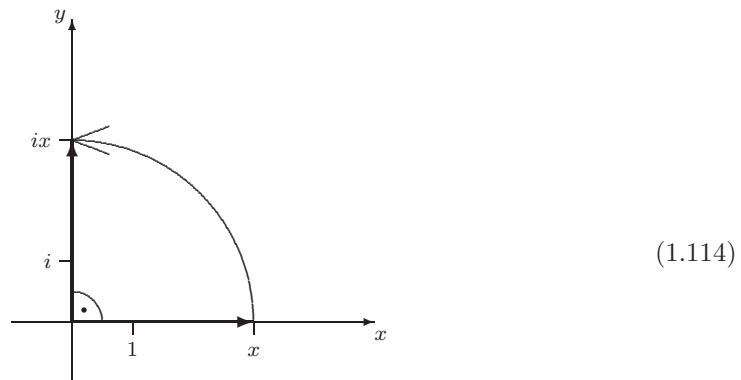
Zwei komplexe Zahlen $z = x + iy$ und $w = u + iv$ sind genau dann gleich, wenn $x = u$ und $y = v$ gilt. Es ist üblich, den vom Ursprung O nach z weisenden Zeiger (Ortsvektor) ebenfalls mit z zu bezeichnen. Die Ebene, deren Punkte als komplexe Zahlen aufgefasst werden, heißt komplexe Zahlenebene.

Die Summe und Differenz komplexer Zahlen ist durch

$$(x + iy) + (u + iv) := (x + u) + i(y + v) \quad (1.112)$$

$$(x + iy) - (u + iv) := (x - u) + i(y - v) \quad (1.113)$$

definiert. Sie entspricht der Vektor-Summe bzw. -Differenz der zugehörigen Ortsvektoren. Die Zahl $ix \in \mathbb{C}$ entsteht aus $x \in \mathbb{R}$ durch eine positive Drehung des entsprechenden Ortsvektors um den Winkel $\frac{\pi}{2}$.



Die Multiplikation einer komplexen Zahl z mit der rein imaginären Zahl $i \in \mathbb{C}$ wird nun so erklärt, dass generell der Ortsvektor $i \cdot z$ aus dem Ortsvektor z durch eine positive Drehung um $\frac{\pi}{2}$ hervorgeht.

Somit erhält man

$$i^2 = i \cdot i = -1. \tag{1.115}$$

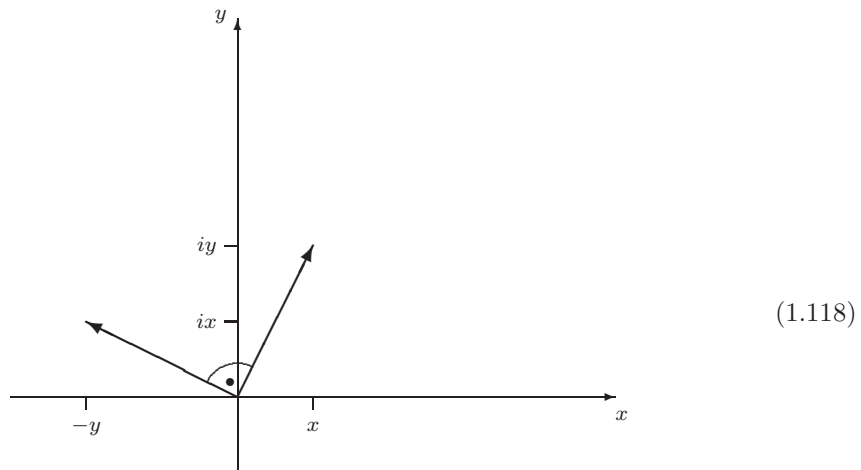
Zusammenfassend wird das Produkt komplexer Zahlen folgendermaßen definiert:

$$(x + iy) \cdot (u + iv) := (xu - yv) + i(xv + yu). \tag{1.116}$$

Diese Definition entspricht dem üblichen „ausmultiplizieren“:

$$(x + iy) \cdot (u + iv) = xu + iyu + ixv - yv = (xu - yv) + i(xv + yu). \tag{1.117}$$

Wegen $i(x + iy) = -y + ix$ stellt $z \mapsto iz$ die gewünschte positive Drehung um den Nullpunkt um $\frac{\pi}{2}$ dar.



Die Potenzen z^n sind wie üblich durch $z^0 := 1$ und $z^{n+1} := z \cdot z^n$ erklärt. Auch in \mathbb{C} gilt die binomische Formel

$$(z + w)^n = \sum_{k=0}^n \binom{n}{k} z^k w^{n-k} \quad (n \in \mathbb{N}_0). \tag{1.119}$$

Beispiel(e) 1.11

$$\begin{aligned} (3 + 0.5i)(2 - 4i) + (-1 + i)^3 &= 6 + i - 12i + 2 + i^3 - 3i^2 + 3i - 1 = \\ &= 10 - 9i. \end{aligned}$$

Die Division ist in \mathbb{C} folgendermaßen erklärt: Man schreibt für $w \neq 0$:

$$q = \frac{z}{w}, \quad \text{falls } w \cdot q = z. \quad (1.120)$$

Für $z = x + iy$ und $w = u + iv \neq 0$ (also $u \neq 0$ oder $v \neq 0$) gilt:

$$\frac{x + iy}{u + iv} = \frac{(x + iy)(u - iv)}{(u + iv)(u - iv)} = \frac{xu + yv}{u^2 + v^2} + i \frac{yu - xv}{u^2 + v^2}. \quad (1.121)$$

Durch $z^{-n} := \frac{1}{z^n}$ sind für $z \neq 0$ und $n \in \mathbb{N}$ die Potenzen mit negativen Exponenten erklärt. Es ist einfach zu zeigen, dass in \mathbb{C} die üblichen, aus \mathbb{R} bekannten Rechengesetze gelten, etwa

$$\begin{aligned} zw &= wz \\ z(ws) &= (zw)s \\ z(w + s) &= zw + zs \\ z^{m+n} &= z^m \cdot z^n. \end{aligned} \quad (1.122)$$

Auch in \mathbb{C} ist eine Division durch 0 nicht möglich.

Man nennt $\bar{z} = x - iy$ die zu $z = x + iy$ konjugiert komplexe Zahl.

Beispiel(e) 1.12

$$\begin{aligned} \overline{-i} &= i \\ \overline{2} &= 2 \\ \overline{4 + 2i} &= 4 - 2i \\ \overline{3 - 2i} &= 3 + 2i. \end{aligned}$$

Für den Übergang $z \mapsto \bar{z}$ zur konjugiert komplexen Zahl (das entspricht der Spiegelung an der reellen Achse) gelten die folgenden Rechenregeln:

$$\begin{aligned} \overline{z + w} &= \bar{z} + \bar{w} \\ \overline{z \cdot w} &= \bar{z} \cdot \bar{w} \\ \overline{\left(\frac{z}{w}\right)} &= \frac{\bar{z}}{\bar{w}} \quad (w \neq 0) \\ \overline{(\bar{z})} &= z \\ \operatorname{Re}(z) &= \frac{1}{2}(z + \bar{z}) \\ \operatorname{Im}(z) &= \frac{1}{2i}(z - \bar{z}). \end{aligned} \quad (1.123)$$

Die Länge des Zeigers $z = x + iy$ wird mit $|z|$ bezeichnet und heißt Betrag der komplexen Zahl. Es gilt:

$$\begin{aligned} |z| &= \sqrt{x^2 + y^2} = \sqrt{z \cdot \bar{z}}, \quad \text{falls } z = x + iy \\ |z \cdot w| &= |z| \cdot |w| \\ |\bar{z}| &= |z| \\ \left|\frac{z}{w}\right| &= \frac{|z|}{|w|} \quad w \neq 0. \end{aligned} \quad (1.124)$$

Ferner gilt die Dreiecksungleichung

$$|z + w| \leq |z| + |w|, \quad (1.125)$$

die sich leicht mit vollständiger Induktion auf n Summanden verallgemeinern läßt:

$$\left| \sum_{k=1}^n z_k \right| \leq \sum_{k=1}^n |z_k|. \quad (1.126)$$

Die Gleichung

$$x^2 + 1 = 0 \quad (1.127)$$

hat in \mathbb{R} offenbar keine Lösung, wohl aber in \mathbb{C} , nämlich $x_{1,2} = \pm i$ und es gilt:

$$x^2 + 1 = (x + i)(x - i). \quad (1.128)$$

Eine quadratische Gleichung

$$ax^2 + bx + c = 0 \quad (\text{mit } a, b, c \in \mathbb{R}, a \neq 0) \quad (1.129)$$

besitzt in \mathbb{C} die beiden Lösungen

$$x_1 = \frac{-b + \sqrt{b^2 - 4ac}}{2a} \quad x_2 = \frac{-b - \sqrt{b^2 - 4ac}}{2a}. \quad (1.130)$$

Für $d < 0$ gilt: $\sqrt{d} = i\sqrt{-d}$ ($-d$ ist dann größer 0) und somit:

$$ax^2 + bx + c = a(x - x_1)(x - x_2). \quad (1.131)$$

Für die quadratische Gleichung

$$x^2 + \sqrt{2}x + 1 = 0 \quad (1.132)$$

ergeben sich zum Beispiel die Lösungen

$$x_1 = \frac{\sqrt{2}}{2}(-1 + i) \quad \text{und} \quad x_2 = \frac{\sqrt{2}}{2}(-1 - i). \quad (1.133)$$

Ist $z_1 \in \mathbb{C} \setminus \mathbb{R}$ die Lösung einer quadratischen Gleichung, so ist nach der obigen Lösungsformel auch \bar{z}_1 eine Lösung dieser Gleichung.

Kapitel 2

Lineare Algebra

2.1 Gleichungssysteme und Matrizen

Zur mathematischen Behandlung vieler Probleme der Technik, etwa zur Netzwerkberechnung in der Elektrotechnik oder zur Berechnung von Fachwerken in der Statik, bedient man sich der Matrizenrechnung.

Definition 2.1 (Matrix)

Eine Matrix vom Typ $m \times n$ (oder eine $(m \times n)$ -Matrix) ist ein rechteckiges Zahlenschema der Form

$$A = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \alpha_{22} & & \alpha_{2n} \\ \vdots & & & \vdots \\ \alpha_{m1} & \alpha_{m2} & \cdots & \alpha_{mn} \end{pmatrix}. \quad (2.1)$$

Die Zahlen $\alpha_{ij} \in \mathbb{R}$ (oder \mathbb{C}) heißen Komponenten (oder Elemente) der Matrix A . Man schreibt abkürzend

$$A = (\alpha_{ij})_{m \times n} \quad (2.2)$$

oder nur $A = (\alpha_{ij})$, wenn der Typ feststeht.

Die $(m \times 1)$ -Matrizen heißen Spaltenmatrizen oder Spaltenvektoren und haben die Form

$$s = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{pmatrix}. \quad (2.3)$$

Die $(1 \times n)$ -Matrizen heißen Zeilenmatrizen oder Zeilenvektoren und haben die Form

$$z = (\alpha_1, \dots, \alpha_n). \quad (2.4)$$

Bei nur einer Spalte oder Zeile benötigt man keinen zusätzlichen Index. Die Matrix $A = (\alpha_{ij})_{m \times n}$ besteht aus m Zeilenvektoren (mit je n Komponenten)

$$z_i = (\alpha_{i1}, \dots, \alpha_{in}) \quad (1 \leq i \leq m) \quad (2.5)$$

bzw. aus n Spaltenvektoren (mit je m Komponenten)

$$s_j = \begin{pmatrix} \alpha_{1j} \\ \vdots \\ \alpha_{mj} \end{pmatrix} \quad (1 \leq j \leq n). \quad (2.6)$$

Je nachdem, ob die Zeilen oder die Spalten von A hervorgehoben werden sollen, schreibt man

$$A = \begin{pmatrix} z_1 \\ \vdots \\ z_m \end{pmatrix} \quad (\text{Zeilendarstellung}) \text{ bzw. } A = (s_1, \dots, s_n) \quad (\text{Spaltendarstellung}). \quad (2.7)$$

Die (i, j) -Komponente α_{ij} von A gehört dem i -ten Zeilenvektor z_i und dem j -ten Spaltenvektor s_j an. Man sagt, α_{ij} steht im Schnittpunkt der i -ten Zeile mit der j -ten Spalte.

Zwei Matrizen $A = (a_{ij})$, $B = (b_{ij})$ heißen gleich, in Zeichen: $A = B$, wenn sie vom gleichen Typ $m \times n$ sind und wenn außerdem $\alpha_{ij} = \beta_{ij}$ gilt für alle i, j mit $1 \leq i \leq m$ und $1 \leq j \leq n$.

Beispiel(e) 2.2

Die Matrix $A = \begin{pmatrix} 1 & 3 & 5 \\ 2 & 0 & 4 \end{pmatrix}$ ist vom Typ 2×3 , sie hat Zeilenvektoren

$$z_1 = (1, 3, 5)$$

$$z_2 = (2, 0, 4)$$

und Spaltenvektoren

$$s_1 = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad s_2 = \begin{pmatrix} 3 \\ 0 \end{pmatrix}, \quad s_3 = \begin{pmatrix} 5 \\ 4 \end{pmatrix}.$$

Es gilt: $A \neq \begin{pmatrix} 1 & 3 & 5 & 0 \\ 2 & 0 & 4 & 0 \end{pmatrix}$.

Die Menge aller $(m \times n)$ -Matrizen mit Komponenten aus \mathbb{R} bezeichnen wir mit $\mathbb{R}^{m \times n}$. Mit \mathbb{R}^n bezeichnen wir üblicherweise die Menge aller Spaltenvektoren $\mathbb{R}^{n \times 1}$. In speziellen Fällen (insbesondere bei Abbildungsvorschriften) wird auch die Menge aller Zeilenvektoren $\mathbb{R}^{1 \times n}$ mit \mathbb{R}^n bezeichnet; dies ist aber die Ausnahme und wird immer explizit kenntlich gemacht. Die Einführung verschiedener Symbole würde mehr zur Verwirrung beitragen, als Klarheit stiften.

Die folgende Definition zeigt, dass man mit Matrizen rechnen kann.

Definition 2.3 (Addition, Subtraktion und Skalar-Multiplikation)

Für $A = (\alpha_{ij})$ und $B = (\beta_{ij})$ aus $\mathbb{R}^{m \times n}$ und jede Zahl $\lambda \in \mathbb{R}$ ist $A + B$, $A - B$ und $\lambda \cdot A$ folgendermaßen definiert:

$$A + B = \begin{pmatrix} \alpha_{11} & \dots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{m1} & \dots & \alpha_{mn} \end{pmatrix} + \begin{pmatrix} \beta_{11} & \dots & \beta_{1n} \\ \vdots & & \vdots \\ \beta_{m1} & \dots & \beta_{mn} \end{pmatrix} \quad (2.8)$$

$$:= \begin{pmatrix} \alpha_{11} + \beta_{11} & \dots & \alpha_{1n} + \beta_{1n} \\ \vdots & & \vdots \\ \alpha_{m1} + \beta_{m1} & \dots & \alpha_{mn} + \beta_{mn} \end{pmatrix}, \quad (2.9)$$

$$A - B = \begin{pmatrix} \alpha_{11} & \dots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{m1} & \dots & \alpha_{mn} \end{pmatrix} - \begin{pmatrix} \beta_{11} & \dots & \beta_{1n} \\ \vdots & & \vdots \\ \beta_{m1} & \dots & \beta_{mn} \end{pmatrix} \quad (2.10)$$

$$:= \begin{pmatrix} \alpha_{11} - \beta_{11} & \dots & \alpha_{1n} - \beta_{1n} \\ \vdots & & \vdots \\ \alpha_{m1} - \beta_{m1} & \dots & \alpha_{mn} - \beta_{mn} \end{pmatrix}, \quad (2.11)$$

$$\lambda \cdot A = \lambda \begin{pmatrix} \alpha_{11} & \dots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{m1} & \dots & \alpha_{mn} \end{pmatrix} := \begin{pmatrix} \lambda \cdot \alpha_{11} & \dots & \lambda \cdot \alpha_{1n} \\ \vdots & & \vdots \\ \lambda \cdot \alpha_{m1} & \dots & \lambda \cdot \alpha_{mn} \end{pmatrix}. \quad (2.12)$$

Mit dieser Definition läßt sich jeder Spaltenvektor $a = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$ folgendermaßen darstellen:

$$a = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ \alpha_2 \\ 0 \end{pmatrix} + \dots + \begin{pmatrix} 0 \\ \vdots \\ \alpha_n \end{pmatrix} = \quad (2.13)$$

$$= \alpha_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + \alpha_2 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \dots + \alpha_n \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \quad (2.14)$$

Somit läßt sich jeder Spaltenvektor $a \in \mathbb{R}^n$ eindeutig in der Form

$$a = \alpha_1 e_1 + \alpha_2 e_2 + \dots + \alpha_n e_n \quad (2.15)$$

mit Vektoren

$$e_1 := \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, e_2 := \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \dots, e_n := \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \quad (2.16)$$

darstellen.

Das System (e_1, \dots, e_n) der Vektoren $e_i \in \mathbb{R}^n$ heißt natürliche Basis des \mathbb{R}^n . Die Matrix

$$O := \begin{pmatrix} 0 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & 0 \end{pmatrix} \in \mathbb{R}^{m \times n}, \quad (2.17)$$

deren sämtliche Komponenten Null sind, heißt Nullmatrix (vom Typ $m \times n$). Die Nullmatrix $O \in \mathbb{R}^{n \times 1}$ (bzw. $O \in \mathbb{R}^{1 \times n}$) wird auch Nullvektor genannt. Mit $-A := (-1) \cdot A$ gelten für Matrizen die folgenden Rechenregeln:

$$\begin{array}{ll} \text{a) } A + B = B + A & \text{für alle } A, B \in \mathbb{R}^{m \times n} \\ \text{b) } (A + B) + C = A + (B + C) & \text{für alle } A, B, C \in \mathbb{R}^{m \times n} \\ \text{c) } A + O = A & \text{für alle } A \in \mathbb{R}^{m \times n} \\ \text{d) } A + (-A) = O & \text{für alle } A \in \mathbb{R}^{m \times n} \\ \text{e) } (\lambda\mu)A = \lambda(\mu A) & \text{für alle } \lambda, \mu \in \mathbb{R}, A \in \mathbb{R}^{m \times n} \\ \text{f) } 1 \cdot A = A & \text{für alle } A \in \mathbb{R}^{m \times n} \\ \text{g) } (\lambda + \mu)A = \lambda A + \mu A & \text{für alle } \lambda, \mu \in \mathbb{R}, A \in \mathbb{R}^{m \times n} \\ \text{h) } \lambda(A + B) = \lambda A + \lambda B & \text{für alle } \lambda \in \mathbb{R}, A, B \in \mathbb{R}^{m \times n}. \end{array} \quad (2.18)$$

Im Folgenden betrachten wir lineare Gleichungssysteme und ihre Darstellung durch Matrizen. Ein lineares Gleichungssystem mit m linearen Gleichungen für n Unbekannte x_1, \dots, x_n hat die Form

$$\begin{array}{cccccc} \alpha_{11}x_1 & + & \alpha_{12}x_2 & + & \dots & + & \alpha_{1n}x_n & = & \beta_1 \\ \alpha_{21}x_1 & + & \alpha_{22}x_2 & + & \dots & + & \alpha_{2n}x_n & = & \beta_2 \\ \vdots & & \vdots & & & & \vdots & & \vdots \\ \alpha_{m1}x_1 & + & \alpha_{m2}x_2 & + & \dots & + & \alpha_{mn}x_n & = & \beta_m \end{array} \quad (2.19)$$

mit den Koeffizienten $\alpha_{ij} \in \mathbb{R}$ und den Absolutgliedern $\beta_i \in \mathbb{R}$.

Kommt eine Unbekannte in einer Gleichung nicht vor, dann hat sie dort den Koeffizient 0.

Für das obige Gleichungssystem schreibt man

$$\begin{pmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \vdots & \vdots & & \vdots \\ \alpha_{m1} & \alpha_{m2} & \dots & \alpha_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_m \end{pmatrix} \quad (2.20)$$

oder kurz

$$Ax = b \quad (2.21)$$

mit der Koeffizientenmatrix $A = (\alpha_{ij}) \in \mathbb{R}^{m \times n}$, dem Spaltenvektor $x \in \mathbb{R}^n$ bestehend aus den unbekannt Komponenten x_1, \dots, x_n und dem Spaltenvektor $b \in \mathbb{R}^m$ bestehend aus den Absolutgliedern.

In der i -ten Zeile der Koeffizientenmatrix stehen die Koeffizienten der i -ten Gleichung, die j -te Spalte von A „gehört“ zur Unbekannten x_j (dort stehen die Koeffizienten von x_j) ($1 \leq i \leq m$, $1 \leq j \leq n$).

Ein lineares Gleichungssystem heißt homogen, falls $b = 0$ (d.h. $\beta_1 = \beta_2 = \dots = \beta_m = 0$), ansonsten inhomogen.

Ein Spaltenvektor $\xi \in \mathbb{R}^n$ mit den Komponenten $\xi_1, \dots, \xi_n \in \mathbb{R}$ heißt eine Lösung des obigen linearen Gleichungssystems, falls für $x_i = \xi_i$, $i = 1, \dots, n$, die m Gleichungen tatsächlich erfüllt sind. Nicht jedes lineare Gleichungssystem ist lösbar. Es können die drei folgenden Fälle auftreten:

A) Das lineare Gleichungssystem besitzt keine Lösung, z.B.:

$$\begin{array}{rcl} 3x_1 & + & 2x_2 = 1 \\ 3x_1 & + & 2x_2 = 5 \end{array} \quad (2.22)$$

B) Das lineare Gleichungssystem besitzt genau eine Lösung, z.B.:

$$\begin{aligned} 3x_1 + 2x_2 &= 1 \\ 3x_1 + x_2 &= 5 \end{aligned} \quad (\text{also: } x_1 = 3, x_2 = -4) \quad (2.23)$$

C) Das lineare Gleichungssystem besitzt unendlich viele Lösungen, z.B.:

$$\begin{aligned} 3x_1 + 2x_2 &= 1 \\ 6x_1 + 4x_2 &= 2 \end{aligned} \quad (2.24)$$

Für jede Zahl $\lambda \in \mathbb{R}$ ist $x_1 = \frac{1}{3}(1 - 2\lambda)$, $x_2 = \lambda$ eine Lösung.

Zwei lineare Gleichungssysteme $Ax = b$ und $Bx = c$ (mit nicht notwendig gleicher Anzahl von Gleichungen, also $A \in \mathbb{R}^{m_1 \times n}$, $B \in \mathbb{R}^{m_2 \times n}$) heißen äquivalent, wenn sie dieselbe Lösungsmenge besitzen.

Ein nach CARL FRIEDRICH GAUSS (1777 – 1855) benanntes Lösungsverfahren für lineare Gleichungssysteme basiert auf der Beobachtung, dass bei folgenden Umformungen ein lineares Gleichungssystem in ein dazu äquivalentes Gleichungssystem übergeht:

- (1) Vertauschung zweier Gleichungen
- (2) Multiplikation einer Gleichung mit einer Zahl $\alpha \neq 0$
- (3) Addition (bzw. Subtraktion) des Vielfachen einer Gleichung zu (bzw. von) einer anderen Gleichung.

Diese Umformungen verändern zwar das lineare Gleichungssystem selbst, aber nicht die entsprechende Lösungsmenge; sie werden übersichtlicher, wenn sie an der sogenannten erweiterten Koeffizientenmatrix

$$(A|b) := \left(\begin{array}{ccc|c} \alpha_{11} & \cdots & \alpha_{1n} & \beta_1 \\ \vdots & & \vdots & \vdots \\ \alpha_{m1} & \cdots & \alpha_{mn} & \beta_n \end{array} \right) \quad (2.25)$$

ausgeführt werden. Man erweitert A um die von den anderen Koeffizienten durch einen senkrechten Strich getrennte Spalte der Absolutglieder. Die den Gleichungsumformungen (1) – (3) entsprechenden Veränderungen der Matrix $(A|b)$ heißen elementare Zeilenumformungen, das sind:

- (1) Vertauschen zweier Zeilen
- (2) Multiplikation einer Zeile mit einer Zahl $\alpha \neq 0$
- (3) Addition (bzw. Subtraktion) des α -fachen einer Zeile zu einer anderen.

Es gilt somit:

Entsteht $(B|c)$ aus $(A|b)$ durch endlich viele elementare Zeilenumformungen, dann sind $Ax = b$ und $Bx = c$ äquivalent.

Das Gauß-Verfahren zur Lösung eines linearen Gleichungssystems $Ax = b$ besteht nun aus drei Teilen:

- (a) der Vorwärtselimination an der erweiterten Matrix $(A|b)$
- (b) einer Lösbarkeitsentscheidung (dieser Schritt entfällt bei homogenen linearen Gleichungssystemen, also bei $b = 0$)
- (c) der Rückwärtssubstitution.

(a): Vorwärtselimination: Man bringt durch eventuelle Zeilenvertauschung eine Zahl ungleich Null an die erste Stelle der ersten Spalte und annulliert die darunter stehenden Zahlen durch Subtraktion eines passenden Vielfachen der neuen ersten Zeile von der zweiten, dritten usw., d.h. ist $\alpha_{11} \neq 0$ (eventuell nach einer Zeilenvertauschung), dann subtrahiert man das $\frac{\alpha_{21}}{\alpha_{11}}$ -fache der ersten Zeile von der zweiten, das $\frac{\alpha_{31}}{\alpha_{11}}$ -fache der ersten Zeile von der dritten usw.. Auf diese Weise entsteht aus $(A|b)$ eine Matrix der Form $(B|c)$ mit:

$$(B|c) = \left(\begin{array}{cccc|c} \diamond & * & * & * & \cdots & * & | & \\ & & & & & & & & \\ & & & & & A_1 & & & \\ & & & & & & & & c \end{array} \right) \quad (2.26)$$

mit einer $(m-1) \times n$ Matrix A_1 , deren erste Spalte aus lauter Nullen besteht. An der \diamond -Stelle steht eine Zahl ungleich Null. Diese Stelle nennt man Pivot-Element. Hat man zu Beginn keine Zeile gefunden, sodass durch Zeilenvertauschung an der ersten Stelle der ersten Spalte ein Zahl ungleich Null steht, so kommt die x_1 -Variable in dem linearen Gleichungssystem nicht vor und muss auch nicht berücksichtigt werden. Im Fall $A_1 = 0$ (Nullmatrix) ist die Vorwärtselimination beendet. Andernfalls wiederholt man dasselbe Vorgehen an der ersten von Null verschiedenen Spalte von A_1 (die erste Zeile von B bleibt unverändert). Diese Vorgehensweise wird so lange wiederholt, bis man zu einer Matrix $(M|d)$ gelangt, die eine Zeilenstufenform besitzt:

$$(M|d) = \left(\begin{array}{cccc|cccc|c} \diamond & * & * & \dots & * & \dots & * & | & \alpha_1 \\ 0 & 0 & \diamond & * & \dots & & * & | & \vdots \\ \vdots & \vdots & 0 & \diamond & * & \dots & * & | & \vdots \\ & & \vdots & 0 & \diamond & * & \dots & * & | & \alpha_r \\ & & & \vdots & 0 & 0 & \dots & 0 & | & \alpha_{r+1} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & | & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & | & \alpha_m \end{array} \right) \cdot \quad (2.27)$$

Kennzeichen der Zeilenstufenform:

- In jeder Zeile stehen links vom Pivot-Element \diamond nur Nullen
- Liest man von oben nach unten, so rückt \diamond pro Zeile um mindestens eine Stelle nach rechts.

Das lineare Gleichungssystem $Mx = d$ ist äquivalent zu $Ax = b$.

Beispiel(e) 2.4

$$\begin{array}{rcl} 2x_1 & - & x_3 = 1 \\ 2x_1 + 4x_2 - x_3 & = & 1 \\ -x_1 + 8x_2 + 3x_3 & = & 2 \end{array} \quad (2.28)$$

$$\left(\begin{array}{ccc|c} 2 & 0 & -1 & 1 \\ 2 & 4 & -1 & 1 \\ -1 & 8 & 3 & 2 \end{array} \right) \xrightarrow[\text{III.} + \frac{1}{2} \cdot \text{I. Zeile}]{\text{II.} - \text{I. Zeile}} \left(\begin{array}{ccc|c} 2 & 0 & -1 & 1 \\ 0 & 4 & 0 & 0 \\ 0 & 8 & 2.5 & 2.5 \end{array} \right) \quad (2.29)$$

$$\xrightarrow{\text{III.} - 2 \cdot \text{II. Zeile}} \left(\begin{array}{ccc|c} 2 & 0 & -1 & 1 \\ 0 & 4 & 0 & 0 \\ 0 & 0 & 2.5 & 2.5 \end{array} \right) \quad (2.30)$$

- (b): Die Lösbarkeitsentscheidung (entfällt bei homogenen Gleichungssystemen, da dort $x = 0$ immer eine Lösung ist):
Ist nach Schritt (a) mit dem Ergebnis

$$(M|d) = \left(\begin{array}{cccccccc|c} \diamond & * & * & \dots & * & \dots & * & \alpha_1 \\ 0 & 0 & \diamond & * & \dots & & * & \vdots \\ \vdots & \vdots & 0 & \diamond & * & \dots & * & \vdots \\ & & \vdots & 0 & \diamond & * & \dots & * & \alpha_r \\ & & & \vdots & 0 & 0 & \dots & 0 & \alpha_{r+1} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \dots & 0 & \alpha_m \end{array} \right) \quad (2.31)$$

eine der Zahlen $\alpha_{r+1}, \dots, \alpha_m$ von Null verschieden, dann ist $Mx = d$ und damit $Ax = b$ nicht lösbar. Denn wäre zum Beispiel $\alpha_{r+1} \neq 0$, so würde die $(r+1)$ -te Zeile des linearen Gleichungssystems $Mx = d$ lauten: $0 = \alpha_{r+1}$.

Für den Fall $\alpha_{r+1} = \dots = \alpha_m = 0$ ist die Rückwärtssubstitution durchführbar.

- (c): Rückwärtssubstitution: Die in (2.31) zu den Spalten ohne \diamond -Stelle gehörenden Unbekannten sind die freien Variablen, die der Reihe nach gleich $\lambda_1, \dots, \lambda_{n-r}$ gesetzt werden. Im (2.31) zugeordneten linearen Gleichungssystem bringt man die freien Variablen auf die rechte Seite, ersetzt sie durch den ihnen zugewiesenen „Wert“ λ_i und berechnet der Reihe nach von unten nach oben die zu \diamond -Stellen gehörenden Variablen (in Abhängigkeit von $\lambda_1, \lambda_2, \dots, \lambda_{n-r}$).

Beispiel(e) 2.5

Wir betrachten das lineare Gleichungssystem

$$\begin{aligned} & x_2 - x_3 - x_4 = 7 \\ x_1 - 4x_2 + 2x_3 & = 0 \\ 2x_1 - 3x_2 - x_3 - 5x_4 & = 35 \\ 3x_1 - 7x_2 + x_3 - 5x_4 & = 35 \end{aligned} \tag{2.32}$$

zu (a):

$$(A|b) = \left(\begin{array}{cccc|c} 0 & 1 & -1 & -1 & 7 \\ 1 & -4 & 2 & 0 & 0 \\ 2 & -3 & -1 & -5 & 35 \\ 3 & -7 & 1 & -5 & 35 \end{array} \right) \tag{2.33}$$

$$\begin{array}{l} \xrightarrow{\text{I. und II.}} \\ \text{Zeile vertauschen} \end{array} \left(\begin{array}{cccc|c} 1 & -4 & 2 & 0 & 0 \\ 0 & 1 & -1 & -1 & 7 \\ 2 & -3 & -1 & -5 & 35 \\ 3 & -7 & 1 & -5 & 35 \end{array} \right) \tag{2.34}$$

$$\begin{array}{l} \text{III.} - 2 \cdot \text{I. Zeile} \\ \text{IV.} - 3 \cdot \text{I. Zeile} \end{array} \left(\begin{array}{cccc|c} 1 & -4 & 2 & 0 & 0 \\ 0 & 1 & -1 & -1 & 7 \\ 0 & 5 & -5 & -5 & 35 \\ 0 & 5 & -5 & -5 & 35 \end{array} \right) \tag{2.35}$$

$$\begin{array}{l} \text{III.} - 5 \cdot \text{II. Zeile} \\ \text{IV.} - 5 \cdot \text{II. Zeile} \end{array} \left(\begin{array}{cccc|c} 1 & -4 & 2 & 0 & 0 \\ 0 & 1 & -1 & -1 & 7 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right) \tag{2.36}$$

$$= (M|d) \tag{2.37}$$

zu (b): $\alpha_3 = \alpha_4 = 0 \Rightarrow$ lösbar.

zu (c): Die Unbekannten x_3 und x_4 sind frei wählbar, also $x_3 = \lambda_1, x_4 = \lambda_2$. Einsetzen in die 2. Zeile und Auflösen nach x_2 ergibt: $x_2 = 7 + \lambda_1 + \lambda_2$. Einsetzen in die erste Zeile und Auflösen nach x_1 ergibt:

$$x_1 = 4(7 + \lambda_1 + \lambda_2) - 2 \cdot \lambda_1 = 28 + 2\lambda_1 + 4\lambda_2.$$

Die allgemeine Lösung lautet:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 28 \\ 7 \\ 0 \\ 0 \end{pmatrix} + \lambda_1 \begin{pmatrix} 2 \\ 1 \\ 1 \\ 0 \end{pmatrix} + \lambda_2 \begin{pmatrix} 4 \\ 1 \\ 0 \\ 1 \end{pmatrix}, \quad \lambda_1, \lambda_2 \in \mathbb{R}. \tag{2.38}$$

2.2 Matrizenmultiplikation

Im Folgenden legen wir das Produkt ab eines Zeilenvektors $a \in \mathbb{R}^n$ und eines Spaltenvektors $b \in \mathbb{R}^n$ mit $a = (\alpha_1, \alpha_2, \dots, \alpha_n)$ und $b = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}$ fest:

$$ab := \alpha_1\beta_1 + \alpha_2\beta_2 + \dots + \alpha_n\beta_n = \sum_{i=1}^n \alpha_i\beta_i. \quad (2.39)$$

Somit ergibt das Produkt „Zeile mal Spalte“ eine Zahl.

Für Zeilenvektoren $a, a_1, a_2 \in \mathbb{R}^n$ und Spaltenvektoren $b, b_1, b_2 \in \mathbb{R}^n$ gelten die folgenden Rechenregeln:

$$\begin{aligned} (a_1 + a_2)b &= a_1b + a_2b \\ a(b_1 + b_2) &= ab_1 + ab_2 \\ \alpha(ab) &= (\alpha a)b = a(\alpha b) \quad \text{für alle } \alpha \in \mathbb{R}. \end{aligned} \quad (2.40)$$

Basierend auf dieser Festlegung definieren wir nun die Multiplikation zweier Matrizen.

Definition 2.6 (Matrixmultiplikation)
 Seien $A = (\alpha_{ij})_{m \times n}$ und $B = (\beta_{ij})_{n \times r}$ zwei Matrizen mit der Zeilendarstellung

$$A = \begin{pmatrix} z_1 \\ \vdots \\ z_m \end{pmatrix}, \quad z_i \in \mathbb{R}^n, \quad i = 1, \dots, m, \quad (2.41)$$

für A und der Spaltendarstellung

$$B = (s_1, \dots, s_r), \quad s_i \in \mathbb{R}^n, \quad i = 1, \dots, r, \quad (2.42)$$

für B , so ist das Produkt AB definiert durch

$$AB := \begin{pmatrix} z_1s_1 & \dots & z_1s_r \\ z_2s_1 & \dots & z_2s_r \\ \vdots & & \vdots \\ z_ms_1 & \dots & z_ms_r \end{pmatrix}. \quad (2.43)$$

Das Matrixprodukt AB ist nur erklärt für $A \in \mathbb{R}^{m \times n}$ und $B \in \mathbb{R}^{n \times r}$, d.h. die Spaltenzahl von A muss gleich der Zeilenzahl von B sein.

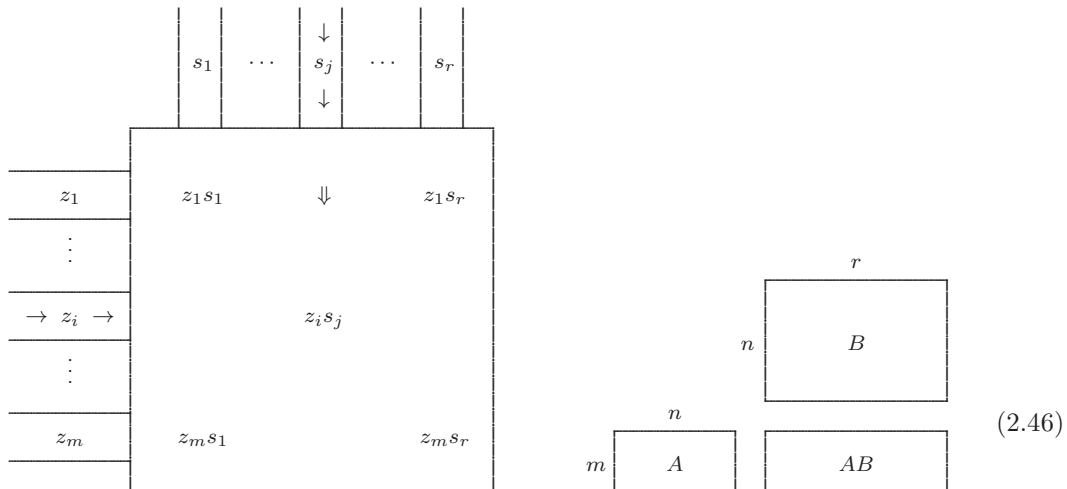
Für $A = (\alpha_{ij})_{m \times n}$, $B = (\beta_{ij})_{n \times r}$ und $AB = (\gamma_{ij})_{m \times r}$ ergibt sich

$$\gamma_{ij} = \sum_{k=1}^n \alpha_{ik} \cdot \beta_{kj}. \quad (2.44)$$

Beispiel(e) 2.7

$$\begin{pmatrix} 4 & 3 & 0 & 1 & 2 \\ 2 & 1 & 4 & 0 & 1 \\ 0 & 0 & 4 & 1 & 0 \\ 2 & 0 & 1 & 0 & 4 \end{pmatrix} \begin{pmatrix} 1 & 2 & 0 \\ 0 & 4 & 2 \\ 3 & 0 & 1 \\ 1 & 1 & 3 \\ 0 & 0 & 4 \end{pmatrix} = \begin{pmatrix} 5 & 21 & 17 \\ 14 & 8 & 10 \\ 13 & 1 & 7 \\ 5 & 4 & 17 \end{pmatrix}. \quad (2.45)$$

Die Matrixmultiplikation läßt sich in folgendem Schema zusammenfassen:



Für das Produkt Ax von $A = (\alpha_{ij})_{m \times n}$ und $x \in \mathbb{R}^n$ erhalten wir

$$Ax = \begin{pmatrix} \alpha_{11} & \dots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{m1} & \dots & \alpha_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \alpha_{11}x_1 + \alpha_{12}x_2 + \dots + \alpha_{1n}x_n \\ \vdots \\ \alpha_{m1}x_1 + \alpha_{m2}x_2 + \dots + \alpha_{mn}x_n \end{pmatrix}. \quad (2.47)$$

In der bisher nur als Abkürzung verstandenen Schreibweise $Ax = b$ für lineare Gleichungssysteme ergibt sich nun durch die obige Gleichung für Ax die richtige Deutung. Unter Verwendung der Matrix

$$E_n := \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \ddots & 0 \\ \vdots & 0 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \dots & 0 & 1 \end{pmatrix} \in \mathbb{R}^{n \times n}, \quad (2.48)$$

die als $n \times n$ -Einheitsmatrix bezeichnet wird, erhalten wir für alle Matrizen $A, A_1, A_2 \in \mathbb{R}^{m \times n}$, $B, B_1, B_2 \in \mathbb{R}^{n \times r}$ und $C \in \mathbb{R}^{r \times s}$:

- a) $(A_1 + A_2) \cdot B = A_1B + A_2B$
 - b) $A(B_1 + B_2) = AB_1 + AB_2$
 - c) $\alpha(AB) = (\alpha A)B = A(\alpha B)$ für alle $\alpha \in \mathbb{R}$
 - d) $A(BC) = (AB)C$
 - e) $E_m A = A E_n = A$.
- (2.49)

Falls $r = m$, gilt im allgemeinen:

$$AB \neq BA. \quad (2.50)$$

Wegen (d) kann man bei Matrixprodukten bestehend aus mehreren Faktoren auf Klammerung verzichten: z.B. $(AB)(CD) = A(BC)D = ABCD$.

Für (quadratische) $(n \times n)$ -Matrizen lassen sich Potenzen A^k , $k \in \mathbb{N}_0$, induktiv definieren:

$$A^0 := E_n, A^{k+1} = A^k A, \text{ d.h. } A^k = \underbrace{A \cdot A \cdot \dots \cdot A}_{k\text{-mal}}. \quad (2.51)$$

Offensichtlich gilt: $A^k \cdot A^l = A^{k+l}, (A^k)^l = A^{k \cdot l}, k, l \in \mathbb{N}_0$.

Jeder $(m \times n)$ -Matrix $A = (\alpha_{ij})$ kann man eine $(n \times m)$ -Matrix A^\top (die transponierte Matrix von A) zuordnen, deren i -te Zeile aus den Elementen der i -ten Spalte von A besteht.

Beispiel(e) 2.8

$$(1, 2, 3, 4)^\top = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}$$

$$\begin{pmatrix} 1 \\ 0 \\ \alpha \end{pmatrix}^\top = (1, 0, \alpha)$$

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 1 & 0 & \alpha \\ \beta & x & 2 & 4 \end{pmatrix}^\top = \begin{pmatrix} 1 & 0 & \beta \\ 2 & 1 & x \\ 3 & 0 & 2 \\ 4 & \alpha & 4 \end{pmatrix}.$$

Es gelten folgende Rechenregeln:

$$\begin{aligned} \text{(a)} \quad (A + B)^\top &= A^\top + B^\top && \text{für alle } A, B \in \mathbb{R}^{m \times n} \\ \text{(b)} \quad (\alpha A)^\top &= \alpha \cdot A^\top && \text{für alle } A \in \mathbb{R}^{m \times n}, \alpha \in \mathbb{R} \\ \text{(c)} \quad (A^\top)^\top &= A && \text{für alle } A \in \mathbb{R}^{m \times n} \\ \text{(d)} \quad (AB)^\top &= B^\top A^\top && \text{für alle } A \in \mathbb{R}^{m \times n}, B \in \mathbb{R}^{n \times r}. \end{aligned} \quad (2.52)$$

Zwei wichtige Klassen von Matrizen werden durch die folgende Definition festgelegt:

Definition 2.9 ((schief-)symmetrische Matrizen)
 Eine $(n \times n)$ -Matrix A heißt symmetrisch, falls $A = A^\top$ gilt,
 sie heißt schiefsymmetrisch, falls $A = -A^\top$.

Für eine symmetrische Matrix $A = (\alpha_{ij})_{n \times n}$ gilt $\alpha_{ij} = \alpha_{ji}$ für alle $i, j \in \{1, \dots, n\}$.
 Für eine schiefsymmetrische Matrix $A = (\alpha_{ij})_{n \times n}$ gilt

$$\alpha_{ij} = -\alpha_{ji} \quad \text{für alle } i, j \in \{1, \dots, n\}; \quad (2.53)$$

insbesondere gilt dann $\alpha_{ii} = 0, i = 1, \dots, n$.

Eine wichtige Menge von Matrizen ist folgendermaßen definiert:

Definition 2.10 (invertierbare Matrizen)

Eine $(n \times n)$ -Matrix A heißt invertierbar, falls es eine $(n \times n)$ -Matrix B gibt, sodass $AB = BA = E_n$ gilt. In diesem Fall ist die Matrix B eindeutig bestimmt und wird im allgemeinen mit A^{-1} bezeichnet. A^{-1} heißt die Inverse oder inverse Matrix von A .

Die obige Definition ist formal nicht ganz korrekt, da nicht nur ein Name vergeben wird, sondern auch eine zu beweisende Behauptung (die Eindeutigkeit von A^{-1}) aufgestellt wird. Daher haben wir den folgenden Satz zu beweisen:

Satz 2.11 (Invertierbarkeit und Eindeutigkeit)

Wenn es zu einer Matrix $A \in \mathbb{R}^{n \times n}$ zwei Matrizen $B, C \in \mathbb{R}^{n \times n}$ gibt mit $BA = AC = E_n$, dann ist A invertierbar und $B = C = A^{-1}$.

Beweis:

$B = BE_n = B(AC) = (BA)C = E_n C = C$,
also ist $BA = AB = E_n$ und somit $B = A^{-1} = C$

q.e.d.

Beispiel(e) 2.12

- E_n ist invertierbar und es gilt $E_n^{-1} = E_n$
- Für $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{R}^{2 \times 2}$ mit $ad - bc \neq 0$ und $B = \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$ gilt $AB = BA = E_2$, also sind A, B invertierbar, $A^{-1} = B$ und $B^{-1} = A$.

Mit dem folgenden Satz fassen wir Eigenschaften invertierbarer Matrizen zusammen.

Satz 2.13 (Eigenschaften invertierbarer Matrizen)

- (i) Die Inverse einer invertierbaren $(n \times n)$ -Matrix ist invertierbar und es gilt:

$$(A^{-1})^{-1} = A. \tag{2.54}$$

- (ii) Das Produkt AB zweier invertierbarer $(n \times n)$ -Matrizen ist invertierbar und es gilt:
 $(AB)^{-1} = B^{-1}A^{-1}$.

- (iii) Die Transponierte A^\top einer $(n \times n)$ -Matrix ist genau dann invertierbar, wenn A invertierbar ist, und es gilt: $(A^\top)^{-1} = (A^{-1})^\top$.

Beweis:

- (i) Da $AA^{-1} = A^{-1}A = E_n$, ist A die Inverse zu A^{-1} .
- (ii) $AB \cdot B^{-1}A^{-1} = A(BB^{-1})A^{-1} = AE_nA^{-1} = AA^{-1} = E_n = B^{-1}A^{-1}AB$. Daher ist $B^{-1}A^{-1}$ die Inverse zu AB .
- (iii) Da $E_n^\top = E_n$, folgt aus $AA^{-1} = E_n$: $E_n = E_n^\top = (AA^{-1})^\top = (A^{-1})^\top A^\top$.
Somit ist $(A^{-1})^\top$ die Inverse zu A^\top , also $(A^{-1})^\top = (A^\top)^{-1}$

q.e.d.

Für endlich viele invertierbare $n \times n$ -Matrizen A_1, \dots, A_k folgt somit:

$$(A_1 A_2 \cdots A_k)^{-1} = A_k^{-1} \cdot \dots \cdot A_1^{-1}. \tag{2.55}$$

In einer quadratischen $(n \times n)$ -Matrix $A = (\alpha_{ij})$ nennt man die Zahlen α_{ii} , $1 \leq i \leq n$, Diagonalelemente. Unterhalb der Diagonalen stehen α_{ij} mit $i > j$, oberhalb die Zahlen α_{ij} mit $i < j$. Man nennt A eine untere bzw. obere Dreiecksmatrix, wenn sie höchstens auf und unterhalb (bzw. oberhalb) der Diagonalen von Null verschiedene Elemente hat.

$$A = \begin{pmatrix} \boxed{\alpha_{11}} & \alpha_{12} & \cdots & \alpha_{1n} \\ \alpha_{21} & \boxed{\alpha_{22}} & \cdots & \vdots \\ \vdots & & \ddots & \vdots \\ \alpha_{n1} & \alpha_{n2} & \cdots & \boxed{\alpha_{nn}} \end{pmatrix} \quad \text{die Diagonale von } A \tag{2.56}$$

$$A = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \cdots & \alpha_{1n} \\ 0 & \alpha_{22} & \cdots & \vdots \\ \vdots & & \ddots & \vdots \\ 0 & \dots & 0 & \alpha_{nn} \end{pmatrix} \quad \text{eine obere Dreiecksmatrix} \tag{2.57}$$

$$A = \begin{pmatrix} \alpha_{11} & 0 & \cdots & 0 \\ \alpha_{21} & \alpha_{22} & \cdots & \vdots \\ \vdots & & \ddots & \vdots \\ \alpha_{n1} & \dots & \dots & \alpha_{nn} \end{pmatrix} \quad \text{eine untere Dreiecksmatrix.} \tag{2.58}$$

Die Transponierte einer oberen Dreiecksmatrix ist eine untere Dreiecksmatrix und umgekehrt. Eine obere (bzw. untere) Dreiecksmatrix ist genau dann invertierbar, wenn alle Diagonalelemente von Null verschieden sind. Für eine $(n \times n)$ -Matrix sind folgende Aussagen äquivalent (d.h., ist eine Aussage wahr, so gilt das auch für alle anderen):

- a) A ist invertierbar
- b) Es gibt eine $(n \times n)$ -Matrix B mit $AB = E_n$
- c) Es gibt eine $(n \times n)$ -Matrix C mit $CA = E_n$
- d) $Ax = 0 \Rightarrow x = 0$.

2.3 Vektorräume

In der Mathematik spielen Mengen eine große Rolle, deren Elemente man addieren und mit einem Zahlenfaktor multiplizieren kann. Mit der Menge der Vektoren in der Ebene (insbesondere den komplexen Zahlen), der Menge der Vektoren im Raum und der Menge der $(m \times n)$ -Matrizen haben wir bereits drei Beispiele kennengelernt. Ein weiteres Beispiel ist etwa die Menge aller auf einem Intervall I definierten Funktionen $f : I \mapsto \mathbb{R}$. Man beobachtet nun, dass für die entsprechenden Rechenoperationen (Addition von Elementen einer Menge und Multiplikation mit einem Zahlenfaktor) unabhängig von der Beschaffenheit der Elemente, die die zugrundegelegte Menge bilden, dieselben Rechengesetze gelten. Zur einheitlichen Herleitung der sich daraus ergebenden

Konsequenzen wird der Begriff des \mathbb{R} -Vektorraumes eingeführt.

Definition 2.14 (\mathbb{R} -Vektorraum)

Eine nichtleere Menge V , in der man zu je zwei Elementen $a, b \in V$ eine Summe $a + b \in V$, zu jedem $a \in V$ und zu jedem $\lambda \in \mathbb{R}$ das λ -fache $\lambda a \in V$ bilden kann, heißt ein \mathbb{R} -Vektorraum (oder Vektorraum über \mathbb{R} , bzw. linearer Raum über \mathbb{R}), wenn folgende acht Rechengesetze (die Vektorraum-Axiome) erfüllt sind:

(V.1) Die Addition ist kommutativ, d.h. für alle $a, b \in V$ gilt

$$a + b = b + a$$

(V.2) Die Addition ist assoziativ, d.h. für alle $a, b, c \in V$ gilt

$$(a + b) + c = a + (b + c)$$

(V.3) Es gibt ein Element $o \in V$, Nullelement genannt, mit

$$a + o = a \quad \text{für alle } a \in V.$$

(V.4) Zu jedem $a \in V$ gibt es genau ein mit $-a$ bezeichnetes Element in V mit $a + (-a) = o$.

(V.5) $1 \cdot a = a$ für alle a (Zahlenfaktor $1 \in \mathbb{R}$).

(V.6) $\lambda(\mu a) = (\lambda\mu)a$ für alle $\lambda, \mu \in \mathbb{R}, a \in V$.

(V.7) $\lambda(a + b) = \lambda a + \lambda b$ für alle $\lambda \in \mathbb{R}, a, b \in V$.

(V.8) $(\lambda + \mu)a = \lambda a + \mu a$ für alle $\lambda, \mu \in \mathbb{R}, a \in V$.

Die Elemente eines Vektorraumes nennt man Vektoren; statt $a + (-b)$ schreibt man $a - b$ (Differenz).

Die Axiome (V.1)-(V.8) garantieren, dass man mit Summen, Differenzen und Vielfachen wie gewohnt rechnen darf. Wegen (V.2) kann man in endlichen Summen $a_1 + a_2 + \dots + a_n$ auf Klammern verzichten und aus den Distributivgesetzen (V.7), (V.8) folgt für alle $\lambda_1, \dots, \lambda_k \in \mathbb{R}, a_1, \dots, a_n \in V$:

$$(\lambda_1 + \lambda_2 + \dots + \lambda_k)(a_1 + \dots + a_n) = \lambda_1 a_1 + \dots + \lambda_k a_n; \quad (2.59)$$

außerdem gelten die Vorzeichenregeln:

$$\begin{aligned} -a &= (-1)a \\ -(-a) &= a \\ -(a + b) &= (-a) + (-b) = -a - b \quad \text{usw.} \end{aligned} \quad (2.60)$$

Das Nullelement wird immer mit o bezeichnet; bei Verwechslungsgefahr mit o_V . Auf das überraschende Axiom (V.5) werden wir später eingehen.

Beispiel(e) 2.15

- Ein Vergleich mit den Rechenregeln für $(m \times n)$ -Matrizen zeigt, dass die Menge aller $(m \times n)$ -Matrizen zusammen mit der komponentenweisen Addition und skalaren Multiplikation ein \mathbb{R} -Vektorraum ist. Insbesondere ist die Menge \mathbb{R}^n für jedes $n \in \mathbb{N}$ ein \mathbb{R} -Vektorraum.
- Für jedes Intervall $I \subseteq \mathbb{R}$ ist die Menge aller Funktionen

$$f : I \rightarrow \mathbb{R} \quad (2.61)$$

durch

$$f + g : I \rightarrow \mathbb{R}, \quad x \mapsto f(x) + g(x) \quad (2.62)$$

und

$$\lambda \cdot f : I \rightarrow \mathbb{R}, \quad x \mapsto \lambda \cdot f(x), \quad \lambda \in \mathbb{R} \quad (2.63)$$

ein \mathbb{R} -Vektorraum.

Häufig erfüllen bereits Teilmengen von \mathbb{R} -Vektorräumen die Vektorraumaxiome.

Definition 2.16 (Unterraum, linearer Teilraum)

Eine nichtleere Teilmenge $U \subseteq V$ eines \mathbb{R} -Vektorraumes V heißt Unterraum oder linearer Teilraum von V , wenn gilt:

$$(U.1) \quad u, v \in U \Rightarrow u + v \in U$$

$$(U.2) \quad u \in U, \lambda \in \mathbb{R} \Rightarrow \lambda \cdot u \in U.$$

Ein Unterraum ist bezüglich der in der Obermenge V definierten Addition und skalaren Multiplikation selbst ein \mathbb{R} -Vektorraum, denn mit $a \in U$ liegt nach (U.2) auch $-a = (-1)a$ in U und nach (U.1) auch $a + (-a) = o$.

Alle anderen Vektorraumaxiome sind selbstverständlich auch für U erfüllt.

Beispiel(e) 2.17

- Jeder \mathbb{R} -Vektorraum V besitzt den „trivialen“ Unterraum $U = \{o\}$.
- Zu jedem Element v eines \mathbb{R} -Vektorraumes V ist

$$\mathbb{R}_v := \{\alpha v; \alpha \in \mathbb{R}\} \tag{2.64}$$

ein Unterraum von V . Für $V = \mathbb{R}^3$ besteht \mathbb{R}_v aus allen zu v parallelen Vektoren (einschließlich dem Nullvektor).

- Im Gegensatz zu

$$V_1 := \left\{ \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{pmatrix} \in \mathbb{R}^4; 2\alpha_1 + 3\alpha_2 + \alpha_4 = 1 \right\} \tag{2.65}$$

ist

$$U := \left\{ \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \end{pmatrix} \in \mathbb{R}^4; 2\alpha_1 + 3\alpha_2 + 4\alpha_4 = 0 \right\} \tag{2.66}$$

ein Unterraum des \mathbb{R}^4 .

- Für jede $(m \times n)$ -Matrix $A \in \mathbb{R}^{m \times n}$ ist

$$\text{Kern}(A) := \{x \in \mathbb{R}^n; Ax = o\}, \quad o \in \mathbb{R}^m \tag{2.67}$$

ein Unterraum des \mathbb{R}^n , denn $o \in \text{Kern}(A)$ und für $u, v \in \text{Kern}(A)$, $\lambda \in \mathbb{R}$ gilt:

$$A(u + v) = Au + Av = o + o = o \tag{2.68}$$

und

$$A(\lambda u) = \lambda(Au) = \lambda \cdot o = o. \tag{2.69}$$

Zwei der zentralen Begriffe der Theorie abstrakter Vektorräume werden in der folgenden Definition festgelegt.

Definition 2.18 (Linearkombination, lineare Hülle)

Sei V ein \mathbb{R} -Vektorraum. Jede aus endlich vielen Vektoren $v_1, \dots, v_k \in V$ gebildete Summe der Form

$$\sum_{i=1}^k \alpha_i v_i = a_1 v_1 + \dots + a_k v_k \tag{2.70}$$

mit den Koeffizienten $\alpha_i \in \mathbb{R}$ heißt eine Linearkombination der v_i .

Eine solche Linearkombination wird trivial genannt, wenn sämtliche α_i gleich Null sind. Die Menge

$$\text{Lin}(\{v_1, \dots, v_k\}) := \left\{ \sum_{i=1}^k \alpha_i v_i; \alpha_i \in \mathbb{R}, i = 1, \dots, k \right\} \tag{2.71}$$

heißt lineare Hülle der v_i .

Ohne Mühe ist einzusehen, dass die lineare Hülle $\text{Lin}(\{v_1, \dots, v_k\})$ ein Unterraum von V ist.

Die folgende Definition stellt einen Zusammenhang zwischen Unterräumen und linearen Hüllen her.

Definition 2.19 (Erzeugendensystem)

Ein Unterraum U eines \mathbb{R} -Vektorraumes V wird von den Vektoren $v_1, \dots, v_k \in V$ erzeugt, oder $\{v_1, \dots, v_k\}$ ist ein Erzeugendensystem von U , wenn

$$U = \text{Lin}(\{v_1, \dots, v_k\}). \quad (2.72)$$

Beispiel(e) 2.20

Die Vektoren $e_1^\top = (1, 0, 0, 0)$, $e_2^\top = (0, 1, 0, 0)$, $e_3^\top = (0, 0, 1, 0)$ und $e_4^\top = (0, 0, 0, 1)$ erzeugen den \mathbb{R}^4 .

Wegen

$$v_i = 0 \cdot v_1 + \dots + 0 \cdot v_{i-1} + 1 \cdot v_i + 0 \cdot v_{i+1} + \dots + 0 \cdot v_k \quad (2.73)$$

(vgl. Axiom (V.5)) gilt

$$v_i \in \text{Lin}(\{v_1, \dots, v_k\}), \quad i = 1, \dots, n. \quad (2.74)$$

Es stellt sich nun die Frage, unter welchen Voraussetzungen zur Erzeugung von $U = \text{Lin}(\{v_1, \dots, v_k\})$ tatsächlich sämtliche Vektoren benötigt werden.

Definition 2.21 (linear abhängig, linear unabhängig)

Endlich viele Vektoren v_1, \dots, v_k eines \mathbb{R} -Vektorraumes V heißen linear abhängig, wenn es Zahlen $\alpha_1, \alpha_2, \dots, \alpha_k \in \mathbb{R}$ gibt, die nicht sämtlich gleich Null sind, sodass gilt:

$$\alpha_1 v_1 + \dots + \alpha_k v_k = o. \quad (2.75)$$

Die Vektoren heißen linear unabhängig, wenn sie nicht linear abhängig sind; d.h.

$$\alpha_1 v_1 + \dots + \alpha_k v_k = o \Rightarrow \alpha_1 = \alpha_2 = \dots = \alpha_k = 0. \quad (2.76)$$

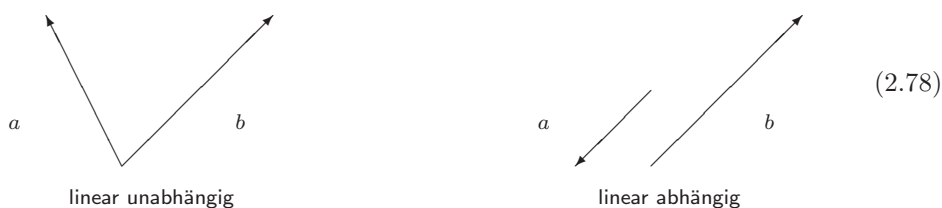
Um Vektoren $v_1, \dots, v_k \in V$ auf lineare Unabhängigkeit zu überprüfen, ist die Gleichung

$$x_1 v_1 + \dots + x_k v_k = o \quad (2.77)$$

zu betrachten. Laut Definition 2.21 gilt:

- (a) v_1, \dots, v_k linear abhängig: obige Gleichung hat unendlich viele Lösungen
- (b) v_1, \dots, v_k linear unabhängig: obige Gleichung hat genau eine Lösung, nämlich $x_1 = x_2 = x_3 = \dots = 0$.

Offensichtlich sind Vektoren $v_1, \dots, v_k \in V$ genau dann linear abhängig, wenn einer von ihnen als Linearkombination der anderen darstellbar ist.



Beispiel(e) 2.22

- $v_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, v_2 = \begin{pmatrix} 0 \\ 2 \\ 1 \end{pmatrix}, v_3 = \begin{pmatrix} 1 \\ 4 \\ 4 \end{pmatrix}$ sind linear abhängig, denn es gilt $v_1 + v_2 - v_3 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$
 bzw. $v_3 = v_1 + v_2$.
 Anschaulich bedeutet dies, dass v_3 parallel zu der von v_1, v_2 aufgespannten Ebene liegt.
- Die Vektoren e_1, e_2, \dots, e_n der natürlichen Basis des \mathbb{R}^n sind linear unabhängig, denn das lineare Gleichungssystem

$$x_1 \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} + x_2 \begin{pmatrix} 0 \\ 1 \\ \vdots \\ 0 \end{pmatrix} + \dots + x_n \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} = o \tag{2.79}$$
 hat als einzige Lösung $x_1 = x_2 = \dots = x_n = 0$.

Betrachtet man eine Matrix in Zeilenstufenform

$$\begin{pmatrix} \diamond & * & * & \dots \\ & \diamond & * & \dots \\ & & \diamond & * & \dots \\ & & & \mathbf{0} & & \diamond & * & * & \dots \end{pmatrix} \tag{2.80}$$

so sind die von Null verschiedenen Zeilen linear unabhängig. Analog dazu sind die von Null verschiedenen Spalten einer Matrix in Spaltenstufenform

$$\begin{pmatrix} \diamond & & & \\ * & \diamond & & \mathbf{0} \\ * & * & \diamond & \\ * & \vdots & * & \\ \vdots & \vdots & * & \diamond \\ * & * & * & * \end{pmatrix} \tag{2.81}$$

linear unabhängig.

Fasst man die zu untersuchenden Vektoren zu einer Matrix zusammen, so ergibt sich der folgende Zusammenhang.

Satz 2.23 (Lineare Unabhängigkeit und Invertierbarkeit)

Für eine $(n \times n)$ -Matrix A sind die folgenden Aussagen äquivalent:

- a) A ist invertierbar
- b) Die Spalten von A sind linear unabhängig
- c) Die Zeilen von A sind linear unabhängig

In der folgenden Definition zeichnen wir spezielle Erzeugendensysteme aus.

Definition 2.24 (Basis)

Eine Menge $\{v_1, \dots, v_n\}$ von Vektoren eines \mathbb{R} -Vektorraumes V heißt eine Basis von V , wenn gilt:

- (B.1) Die Vektoren v_1, \dots, v_n sind linear unabhängig
- (B.2) $V = \text{Lin}(\{v_1, \dots, v_n\})$.

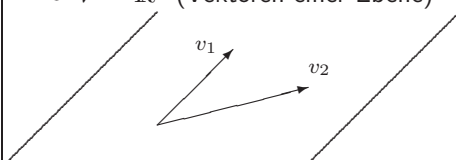
Die Bedeutung einer Basis $\{v_1, \dots, v_n\}$ eines \mathbb{R} -Vektorraumes V ergibt sich aus der Eigenschaft, dass es zu jedem Vektor $v \in V$ genau eine Menge $\{\alpha_1, \dots, \alpha_n\}$ reeller Zahlen gibt mit

$$v = \sum_{i=1}^n \alpha_i v_i. \tag{2.82}$$

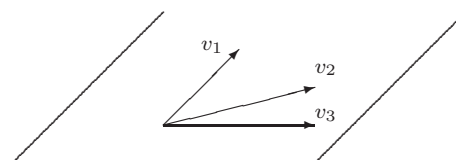
Ferner sind m Vektoren aus V immer linear abhängig, falls $m > n$.

Beispiel(e) 2.25

- $V = \mathbb{R}^2$ (Vektoren einer Ebene)



Basis



keine Basis

(2.83)

- Die Vektoren $e_1, e_2, \dots, e_n \in \mathbb{R}^n$ bilden eine Basis des \mathbb{R}^n
- Die Zeilen (bzw. Spalten) einer invertierbaren $(n \times n)$ -Matrix bilden eine Basis des \mathbb{R}^n .

In der folgenden Definition betrachten wir besondere \mathbb{R} -Vektorräume.

Definition 2.26 (endlichdimensionale \mathbb{R} -Vektorräume)

Ein \mathbb{R} -Vektorraum V heißt endlichdimensional, falls es endlich viele Vektoren $w_1, \dots, w_r \in V$ gibt mit

$$V = \text{Lin}(\{w_1, \dots, w_r\}). \tag{2.84}$$

Die wichtigste Eigenschaft endlichdimensionaler \mathbb{R} -Vektorräume $V \neq \{o\}$ besteht in der Tatsache, dass entweder linear unabhängige Vektoren $w_1, \dots, w_k \in V$ eine Basis von V bilden oder dass diese Vektoren durch Hinzunahme weiterer Vektoren u_1, \dots, u_l zu einer Basis $\{w_1, \dots, w_k, u_1, \dots, u_l\}$ von V ergänzt werden können (Basisergänzungssatz).

Satz und Definition 2.27 (Dimension, Basislänge)

Sei $V \neq \{o\}$ ein endlichdimensionaler \mathbb{R} -Vektorraum, so besitzt V eine Basis $\{v_1, \dots, v_n\}$.

Ist $\{w_1, \dots, w_m\}$ ebenfalls eine Basis von V , so gilt $m = n$. Die gemeinsame Basislänge n aller Basen von V heißt die Dimension von V und wird mit $\text{Dim}(V)$ bezeichnet. Üblicherweise legt man

$$\text{Dim}(\{o\}) = 0 \quad (2.85)$$

fest.

Beispiel(e) 2.28

- $\text{Dim}(\mathbb{R}^n) = n$

- $\text{Dim} \left(\left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} \in \mathbb{R}^3; 3x + y + z = 0 \right\} \right) = 2.$

Ist r die Maximalzahl linear unabhängiger Vektoren aus v_1, \dots, v_k , so gilt:

$$r = \text{Dim}(\text{Lin}(\{v_1, \dots, v_k\})). \quad (2.86)$$

In einem \mathbb{R} -Vektorraum V der Dimension n bilden je n linear unabhängige Vektoren eine Basis von V . Jeder Unterraum U eines endlichdimensionalen \mathbb{R} -Vektorraumes V ist ebenfalls endlichdimensional und im Fall $U \neq V$ gilt:

$$\text{Dim}(U) < \text{Dim}(V). \quad (2.87)$$

2.4 Elementarmatrizen

Für eine gegebene $(m \times n)$ -Matrix A mit Zeilen $z_1, \dots, z_m \in \mathbb{R}^n$ und Spalten $s_1, \dots, s_n \in \mathbb{R}^m$ bezeichnet man mit

$$\text{Lin}(\{z_1, \dots, z_m\}) = \left\{ \sum_{i=1}^m \lambda_i z_i; \lambda_i \in \mathbb{R} \right\} \quad (2.88)$$

bzw.

$$\text{Lin}(\{s_1, \dots, s_n\}) = \left\{ \sum_{i=1}^n \lambda_i s_i; \lambda_i \in \mathbb{R} \right\} \quad (2.89)$$

den Zeilen- bzw. Spaltenraum von A . Offensichtlich gilt

$$\text{Lin}(\{z_1, \dots, z_m\}) = \{y^\top A; y \in \mathbb{R}^m\} \quad (2.90)$$

und

$$\text{Lin}(\{s_1, \dots, s_n\}) = \{Ax; x \in \mathbb{R}^n\}. \quad (2.91)$$

wobei die invertierbare Matrix

$$D_i(\alpha) = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & & \alpha & & \\ & & & & 1 & \\ & \mathbf{0} & & & & \ddots \\ & & & & & & 1 \end{pmatrix} \in \mathbb{R}^{m \times m} \quad (2.95)$$

mit $\alpha_{ii} = \alpha$ Elementarmatrix vom Typ (2) genannt wird. Die Multiplikation $AD_i(\alpha)$ mit $D_i(\alpha) \in \mathbb{R}^{n \times n}$ bewirkt, dass die i -te Spalte von A mit α multipliziert wird. Es ist $D_i(\alpha)^{-1} = D_i(1/\alpha)$.

zu (3) Die Addition des α -fachen der j -ten Zeile von A zur i -ten Zeile von A ist gegeben durch $N_{ij}(\alpha) \cdot A$, wobei

$$N_{ij}(\alpha) = \begin{pmatrix} 1 & & & & & & & & \\ & \ddots & & & & & & & \\ & & 1 & & & & & & \\ & & & 1 & & & & & \\ & & & & 1 & & & & \\ & & & & & \ddots & & & \\ & & & & & & 1 & & \\ & & \alpha & & & & & 1 & \\ & & & & & & & & 1 \\ & \mathbf{0} & & & & & & & \ddots \\ & & & & & & & & & 1 \end{pmatrix} \in \mathbb{R}^{m \times m} \quad (2.96)$$

mit $\alpha_{ij} = \alpha$ Elementarmatrix vom Typ (3) genannt wird. $N_{ij}(\alpha)$ ist invertierbar mit $N_{ij}(\alpha)^{-1} = N_{ij}(-\alpha)$. Die Multiplikation $AN_{ij}(\alpha)$ mit $N_{ij}(\alpha) \in \mathbb{R}^{n \times n}$ bewirkt, dass das α -fache der i -ten (!) Spalte zur j -ten (!) Spalte addiert wird.

Wird nun eine Matrix $(A|b) \in \mathbb{R}^{m \times (n+1)}$ nach dem Gaußverfahren durch die Vorwärtselimination mit elementaren Zeilenumformungen umgewandelt in eine Matrix $(M|d)$ mit

$$(M|d) = \left(\begin{array}{cccc|cccc|c} \diamond & * & \dots & * & & & & & * \\ & \diamond & * & \dots & * & & \mathbf{0} & & \vdots \\ & & \diamond & * & \dots & * & & & \vdots \\ & & & & \diamond & * & \dots & * & \vdots \\ & & & & & & & & * \\ & & & & & & & & * \\ & & & & & & & & * \end{array} \right), \quad (2.97)$$

so bilden die von Null verschiedenen Zeilen von M eine Basis des Zeilenraumes von M und damit auch des Zeilenraumes von A , denn es gilt

$$M = P \cdot A, \quad (2.98)$$

wobei P eine invertierbare Matrix ist, die ihrerseits aus dem Produkt von Elementarmatrizen des Typs (1), (2) oder (3) besteht. Diese Elementarmatrizen repräsentieren gerade die benötigten elementaren Zeilenumformungen der Vorwärtselimination.

Analog dazu entspricht die Multiplikation mit Elementarmatrizen von rechts elementaren Spaltenumformungen und ordnet einer Matrix $A \in \mathbb{R}^{m \times n}$ durch

$$\hat{M} = AQ \quad (2.99)$$

mit

$$\hat{M} = \begin{pmatrix} \diamond & & & & \\ \star & \diamond & \mathbf{0} & & \\ \vdots & \star & \diamond & & \\ & \vdots & \star & & \\ \vdots & \vdots & \vdots & \mathbf{0} & \\ \star & \star & \star & & \end{pmatrix} \quad (2.100)$$

eine Matrix \hat{M} mit demselben Spaltenraum wie A zu.

Die von Null verschiedenen Spalten von \hat{M} bilden eine Basis des Spaltenraumes von A .

Definition 2.30 (Rang einer Matrix)

Sei $A \in \mathbb{R}^{m \times n}$ eine $(m \times n)$ -Matrix, so wird die Dimension des Zeilenraumes als Rang von A bezeichnet.

Der folgende Satz fasst wichtige Ergebnisse über Matrizen zusammen; dabei sei an den Begriff $\text{Kern}(A) := \{x \in \mathbb{R}^n; Ax = o\}$ für eine $(m \times n)$ -Matrix A erinnert.

Satz 2.31 (Eigenschaften von $(m \times n)$ -Matrizen)

Für jede $(m \times n)$ -Matrix A gilt:

- a) Die Dimension des Spaltenraumes von A ist gleich der Dimension des Zeilenraumes von A ($= \text{Rang}(A)$).
- b) $\text{Rang}(A) + \text{Dim}(\text{Kern}(A)) = n$.
- c) Es gibt invertierbare Matrizen $P \in \mathbb{R}^{m \times m}$ und $Q \in \mathbb{R}^{n \times n}$ mit

$$PAQ = \begin{pmatrix} 1 & & & & \\ & \ddots & & & \mathbf{0} \\ & & 1 & & \\ & & & \ddots & \\ \mathbf{0} & & & & \mathbf{0} \end{pmatrix} \quad \text{mit } \text{Rang}(A) = r. \quad (2.101)$$

- d) $\text{Rang}(A^\top) = \text{Rang}(A)$.
- e) $\text{Rang}(PAQ) = \text{Rang}(A)$ für alle invertierbaren Matrizen $P \in \mathbb{R}^{m \times m}$ und $Q \in \mathbb{R}^{n \times n}$.

Nach Satz 2.31 ändert sich der Rang einer Matrix durch elementare Zeilen- und Spaltenumformungen nicht.

Mit elementaren Zeilenumformungen kann man die Inverse einer invertierbaren Matrix $A \in \mathbb{R}^{n \times n}$

berechnen. Man beginnt mit der erweiterten Matrix

$$\left(A \left| \begin{pmatrix} 1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & 1 \end{pmatrix} \right. \right) \quad (2.102)$$

und führt solange elementare Zeilenumformungen durch, bis man die Form

$$\left(\begin{pmatrix} 1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & 1 \end{pmatrix} \left| A^{-1} \right. \right) \quad (2.103)$$

erreicht hat.

Beispiel(e) 2.32
Sei

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 0 \\ 1 & 0 & 2 \end{pmatrix} : \quad (2.104)$$

$$\left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 2 & 0 & 0 & 1 \end{array} \right) \begin{array}{l} II. - 2 \cdot I. \\ III - I \end{array} \quad (2.105)$$

$$\begin{array}{l} II. - 2 \cdot I. \\ III - I \end{array} \left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & -3 & -6 & -2 & 1 & 0 \\ 0 & -2 & -1 & -1 & 0 & 1 \end{array} \right) \begin{array}{l} III. - \frac{2}{3} \cdot II. \end{array} \quad (2.106)$$

$$\begin{array}{l} III. - \frac{2}{3} \cdot II. \end{array} \left(\begin{array}{ccc|ccc} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & -3 & -6 & -2 & 1 & 0 \\ 0 & 0 & 3 & 1/3 & -2/3 & 1 \end{array} \right) \begin{array}{l} I. - III. \\ II. + 2 \cdot III \end{array} \quad (2.107)$$

$$\begin{array}{l} I. - III. \\ II. + 2 \cdot III \end{array} \left(\begin{array}{ccc|ccc} 1 & 2 & 0 & 2/3 & 2/3 & -1 \\ 0 & -3 & 0 & -4/3 & -1/3 & 2 \\ 0 & 0 & 3 & 1/3 & -2/3 & 1 \end{array} \right) \begin{array}{l} -\frac{1}{3} \cdot II. \\ \frac{1}{3} \cdot III. \end{array} \quad (2.108)$$

$$\begin{array}{l} -\frac{1}{3} \cdot II. \\ \frac{1}{3} \cdot III. \end{array} \left(\begin{array}{ccc|ccc} 1 & 2 & 0 & 2/3 & 2/3 & -1 \\ 0 & 1 & 0 & 4/9 & 1/9 & -2/3 \\ 0 & 0 & 1 & 1/9 & -2/9 & 1/3 \end{array} \right) \begin{array}{l} I. - 2 \cdot II. \end{array} \quad (2.109)$$

$$\begin{array}{l} I. - 2 \cdot II. \end{array} \left(\begin{array}{ccc|ccc} 1 & 0 & 0 & -2/9 & 4/9 & 1/3 \\ 0 & 1 & 0 & 4/9 & 1/9 & -2/3 \\ 0 & 0 & 1 & 1/9 & -2/9 & 1/3 \end{array} \right) \quad (2.110)$$

Mit elementaren Spaltenumformungen kann man eine Basis des Kerns einer Matrix $A \in \mathbb{R}^{m \times n}$ berechnen.

Man beginnt mit der erweiterten Matrix $\left(\begin{array}{ccc|ccc} & & & A & & \\ & & & \hline & & & 1 & & \mathbf{0} \\ & & & & \ddots & \\ & & & \mathbf{0} & & 1 \end{array} \right)$ und berechnet die Spaltenstufenform

$$\left(\begin{array}{cccc} \diamond & & & \\ \star & \diamond & & \\ \vdots & \star & \diamond & \mathbf{0} \\ \vdots & \vdots & \vdots & \\ \star & \star & \star & \\ \hline & \star & & v_1, \dots, v_r \end{array} \right) \quad (2.111)$$

Die Vektoren v_1, \dots, v_r bilden dann eine Basis von $\text{Kern}(A)$.

Beispiel(e) 2.33
 Sei

$$A = \begin{pmatrix} 3 & 6 & 9 & -3 \\ 0 & 2 & 6 & 10 \\ 15 & 34 & 58 & 3 \\ -9 & -10 & -3 & 49 + \alpha \end{pmatrix}, \quad \alpha \in \mathbb{R} \quad (2.112)$$

$$\begin{pmatrix} 3 & 6 & 9 & -3 \\ 0 & 2 & 6 & 10 \\ 15 & 34 & 58 & 3 \\ -9 & -10 & -3 & 49 + \alpha \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 2 & 6 & 10 \\ 15 & 4 & 13 & 18 \\ -9 & 8 & 24 & 40 + \alpha \\ 1 & -2 & -3 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix} \rightarrow \dots \quad (2.113)$$

$$\rightarrow \begin{pmatrix} 3 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 15 & 4 & 1 & 0 \\ -9 & 8 & 0 & \alpha \\ 1 & -2 & 3 & 17 \\ 0 & 1 & -3 & -11 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \end{pmatrix} \quad (2.114)$$

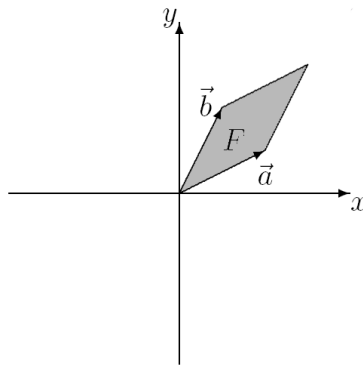
Fall 1: $\alpha \neq 0 \implies \text{Rang}(A) = 4 \implies \text{Kern}(A) = \{o\}$
 Fall 2: $\alpha = 0 \implies \text{Rang}(A) = 3 \implies \text{Kern}(A) = \left\{ t \cdot \begin{pmatrix} 17 \\ -11 \\ 2 \\ 1 \end{pmatrix}; t \in \mathbb{R} \right\}$

2.5 Determinanten

Determinanten wurden von G. W. LEIBNIZ bereits 1678 eingeführt. Sie dienen unter anderem zur Beschreibung elementargeometrischer Lösungen und zur speziellen Darstellung von Lösungen linearer Gleichungssysteme. Betrachtet man zum Beispiel zwei Vektoren $\vec{a} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}$, $\vec{b} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$, so

ist die Fläche des in der (x, y) -Ebene durch \vec{a}, \vec{b} aufgespannten Parallelogrammes gegeben durch

$$F = \left| \begin{pmatrix} a_1 \\ a_2 \\ 0 \end{pmatrix} \times \begin{pmatrix} b_1 \\ b_2 \\ 0 \end{pmatrix} \right| = |a_1 b_2 - b_1 a_2|. \quad (2.115)$$



(2.116)

Fasst man die Vektoren \vec{a}, \vec{b} zu einer (2×2) -Matrix $A = \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix}$ zusammen, so heißt die Zahl

$$\det(A) := a_1 b_2 - b_1 a_2 \quad (2.117)$$

Determinante der (2×2) -Matrix A . Es ist

$$\det(A) = \pm F = 0 \quad (2.118)$$

genau dann, wenn \vec{a}, \vec{b} parallel sind (oder einer der beiden der Nullvektor ist). Damit ist gezeigt:

$$\begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \text{ linear unabhängig} \iff \det \left(\begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix} \right) \neq 0 \iff A \text{ invertierbar.} \quad (2.119)$$

Beispiel(e) 2.34

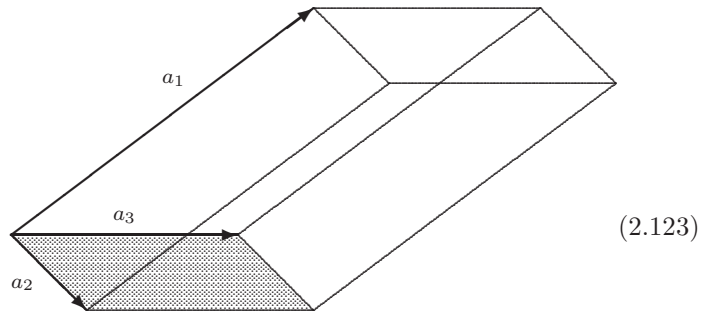
$$\det \left(\begin{pmatrix} 1 & 3 \\ 2 & 5 \end{pmatrix} \right) = 1 \cdot 5 - 2 \cdot 3 = -1 \neq 0. \quad (2.120)$$

Die Matrix ist invertierbar:

$$\begin{pmatrix} 1 & 3 \\ 2 & 5 \end{pmatrix}^{-1} = (-1) \begin{pmatrix} 5 & -3 \\ -2 & 1 \end{pmatrix}. \quad (2.121)$$

Für einen Spat mit den bezüglich einer kartesischen Basis dargestellten Kanten

$$a_1 = \begin{pmatrix} \alpha_{11} \\ \alpha_{21} \\ \alpha_{31} \end{pmatrix}, a_2 = \begin{pmatrix} \alpha_{12} \\ \alpha_{22} \\ \alpha_{32} \end{pmatrix}, a_3 = \begin{pmatrix} \alpha_{13} \\ \alpha_{23} \\ \alpha_{33} \end{pmatrix} \quad (2.122)$$



kann das Volumen berechnet werden. Es ergibt sich:

$$V = |\alpha_{11}(\alpha_{22}\alpha_{33} - \alpha_{32}\alpha_{23}) - \alpha_{21}(\alpha_{12}\alpha_{33} - \alpha_{32}\alpha_{13}) + \alpha_{31}(\alpha_{12}\alpha_{23} - \alpha_{22}\alpha_{13})|. \quad (2.124)$$

Fasst man die Vektoren $\vec{a}_1, \vec{a}_2, \vec{a}_3$ zu einer (3×3) -Matrix $A = \begin{pmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{pmatrix}$ zusammen, so heißt die Zahl

$$\det(A) = \alpha_{11}(\alpha_{22}\alpha_{33} - \alpha_{32}\alpha_{23}) - \alpha_{21}(\alpha_{12}\alpha_{33} - \alpha_{32}\alpha_{13}) + \alpha_{31}(\alpha_{12}\alpha_{23} - \alpha_{22}\alpha_{13}) \quad (2.125)$$

Determinante der (3×3) -Matrix A . Es gilt:

$$\text{Rang}(A) < 3 \iff \det(A) = 0 \quad (2.126)$$

$$\text{Rang}(A) = 3 \iff A \text{ invertierbar} \iff \det(A) \neq 0. \quad (2.127)$$

Ferner gilt:

$$\det(A) = \alpha_{11} \det \begin{pmatrix} \alpha_{22} & \alpha_{23} \\ \alpha_{32} & \alpha_{33} \end{pmatrix} - \alpha_{21} \det \begin{pmatrix} \alpha_{12} & \alpha_{13} \\ \alpha_{32} & \alpha_{33} \end{pmatrix} + \alpha_{31} \det \begin{pmatrix} \alpha_{12} & \alpha_{13} \\ \alpha_{22} & \alpha_{23} \end{pmatrix}. \quad (2.128)$$

Anschaulich

$$\det \begin{pmatrix} \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet \end{pmatrix} = \det \begin{pmatrix} \blacksquare & & \\ & \boxtimes & \\ & & \blacksquare \end{pmatrix} - \det \begin{pmatrix} & \blacksquare & \\ \blacksquare & & \\ & & \boxtimes \end{pmatrix} + \det \begin{pmatrix} & & \blacksquare \\ \blacksquare & & \\ & \boxtimes & \blacksquare \end{pmatrix} \quad (2.129)$$

Beispiel(e) 2.35

$$A = \begin{pmatrix} 0 & -2 & 3 \\ -2 & 1 & -2 \\ 3 & 6 & 5 \end{pmatrix} \quad (2.130)$$

$$\det(A) = 0 \cdot (5 + 12) + 2(-10 - 18) + 3(4 - 3) = -53. \quad (2.131)$$

Analog zur Definition der Determinante für (3×3) -Matrizen lassen sich Determinanten für $(n \times n)$ -Matrizen rekursiv definieren:

Für $n = 1$, d.h. $A = (\alpha_{11})$, ist $\det(A) = \alpha_{11}$.

Für $n \geq 2$ ist (Entwicklung von $\det(A)$ nach der ersten Spalte):

$$\det(A) := \alpha_{11}\det(A_{11}) - \alpha_{21}\det(A_{21}) + \alpha_{31}\det(A_{31}) \pm \dots + (-1)^{n+1}\alpha_{n1}\det(A_{n1}), \quad (2.132)$$

wobei A_{i1} die $((n-1) \times (n-1))$ -Matrix bezeichnet, die aus A durch Entfernen der ersten Spalte und der i -ten Zeile entsteht.

Beispiel(e) 2.36

$$A = \begin{pmatrix} 1 & 2 & 0 & 1 \\ 3 & -2 & 1 & 0 \\ 0 & 6 & 3 & -2 \\ 2 & 4 & 3 & 1 \end{pmatrix} \quad (2.133)$$

$$A_{11} = \begin{pmatrix} -2 & 1 & 0 \\ 6 & 3 & -2 \\ 4 & 3 & 1 \end{pmatrix}, \quad A_{21} = \begin{pmatrix} 2 & 0 & 1 \\ 6 & 3 & -2 \\ 4 & 3 & 1 \end{pmatrix}, \quad (2.134)$$

$$A_{31} = \begin{pmatrix} 2 & 0 & 1 \\ -2 & 1 & 0 \\ 4 & 3 & 1 \end{pmatrix}, \quad A_{41} = \begin{pmatrix} 2 & 0 & 1 \\ -2 & 1 & 0 \\ 6 & 3 & -2 \end{pmatrix}, \quad (2.135)$$

$$\det(A) = 1 \cdot |A_{11}| - 3 \cdot |A_{21}| - 2 \cdot |A_{41}| = -72. \quad (2.136)$$

Für obere Dreiecksmatrizen läßt sich die Determinante einfach berechnen.

Satz 2.37 (Determinante einer oberen Dreiecksmatrix)

Für eine obere Dreiecksmatrix gilt

$$\det \left(\begin{pmatrix} \alpha_{11} & & \star \\ & \ddots & \\ \mathbf{0} & & \alpha_{nn} \end{pmatrix} \right) = \alpha_{11} \cdot \dots \cdot \alpha_{nn} = \prod_{i=1}^n \alpha_{ii}. \quad (2.137)$$

Der Beweis besteht aus einfachem Nachrechnen. Wichtig sind die folgenden Rechenregeln für Determinanten.

Satz 2.38 (Rechenregeln für Determinanten)

Sei A eine $(n \times n)$ -Matrix mit der Zeilendarstellung $A = \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix}$ und der Spaltendarstellung $A = (s_1, \dots, s_n)$, so gilt:

- (i) Besteht die i -te Zeile (bzw. Spalte) von A aus einer Summe, also $z_i = a + b$ (bzw. $s_i = c + d$), $i \in \{1, \dots, n\}$, so gilt:

$$\det \begin{pmatrix} \begin{pmatrix} z_1 \\ \vdots \\ a+b \\ \vdots \\ z_n \end{pmatrix} \end{pmatrix} = \det \begin{pmatrix} \begin{pmatrix} z_1 \\ \vdots \\ a \\ \vdots \\ z_n \end{pmatrix} \end{pmatrix} + \det \begin{pmatrix} \begin{pmatrix} z_1 \\ \vdots \\ b \\ \vdots \\ z_n \end{pmatrix} \end{pmatrix} \quad (2.138)$$

bzw.

$$\det((s_1, \dots, c + d, \dots, s_n)) = \det((s_1, \dots, c, \dots, s_n)) + \det((s_1, \dots, d, \dots, s_n)). \quad (2.139)$$

- (ii) Sei nun $B = E \cdot A$, wobei E eine Elementarmatrix vom Typ (1), (2) oder (3) ist, so gilt:

$$\begin{aligned} \det(B) &= -\det(A), \text{ falls } E \text{ vom Typ (1) ist;} \\ \det(B) &= \alpha \det(A), \text{ falls } E \text{ vom Typ (2) ist;} \\ \det(B) &= \det(A), \text{ falls } E \text{ vom Typ (3) ist;} \end{aligned}$$

- (iii) $\det(A^T) = \det(A)$

- (iv) Sei C eine weitere $(n \times n)$ -Matrix, so gilt

$$\det(AC) = \det(A) \cdot \det(C) \quad (= \det(C) \cdot \det(A) = \det(C \cdot A)) \quad (2.140)$$

- (v) A ist invertierbar $\iff \det(A) \neq 0$.

Da die Determinante rekursiv definiert ist, wird im Beweis des obigen Satzes häufig die vollständige Induktion verwendet.

Sei $A = (a_1, \dots, a_n)$ eine $(n \times n)$ -Matrix in Spaltendarstellung, so entsteht

$$\tilde{A} = (a_j, a_1, a_2, \dots, a_{j-1}, a_{j+1}, \dots, a_n) \quad (2.141)$$

durch $(j - 1)$ sukzessive Vertauschungen benachbarter Spalten. Also gilt

$$\det(\tilde{A}) = (-1)^{j-1} \cdot \det(A). \quad (2.142)$$

Berechnet man nun $\det(\tilde{A})$ wie bisher, so erhält man eine neue Formel für $\det(A)$:

$$\det(A) = \sum_{i=1}^n (-1)^{i+j} \alpha_{ij} \det(A_{ij}), \quad (2.143)$$

wobei die $((n-1) \times (n-1))$ -Matrix A_{ij} durch Streichen der i -ten Zeile und der j -ten Spalte aus der Matrix A entsteht.

Mit Hilfe von Determinanten läßt sich die Lösung $x \in \mathbb{R}^n$ eines linearen Gleichungssystems $Ax = b$ gegeben durch eine invertierbare Matrix $A \in \mathbb{R}^{n \times n}$ darstellen (Cramer-Regel):

$$x_i = \frac{\det((a_1, \dots, a_{i-1}, b, a_{i+1}, \dots, a_n))}{\det(A)}, \quad i = 1, \dots, n, \quad (2.144)$$

wobei $A = (a_1, \dots, a_n)$ die Spaltendarstellung der Matrix A ist (die i -te Spalte von A wird durch b ersetzt).

2.6 Lineare Abbildungen und Eigenwerte

Seien V, W \mathbb{R} -Vektorräume. Mit einer Abbildung $f : V \rightarrow W$ wird jedem Vektor $v \in V$ ein eindeutiger Vektor $w \in W$ zugeordnet.

Definition 2.39 (Lineare Abbildung)

Seien V, W \mathbb{R} -Vektorräume und $f : V \rightarrow W$ eine Abbildung, dann heißt f linear, falls gilt:

(L.1) f ist homogen; d.h. $f(\alpha v) = \alpha \cdot f(v)$ für alle $\alpha \in \mathbb{R}, v \in V$.

(L.2) f ist additiv; d.h. $f(u + v) = f(u) + f(v)$ für alle $u, v \in V$.

Eine lineare Abbildung $f : V \rightarrow W$ wird auch als linearer Operator, lineare Transformation oder Vektorraumhomomorphismus bezeichnet.

Beispiel(e) 2.40

- Die Nullabbildung $N : V \rightarrow V, v \mapsto o$ für alle $v \in V$.
- Die Projektionen $p_i : \mathbb{R}^n \rightarrow \mathbb{R}, x \mapsto x_i$.
- $l : \mathbb{R}^n \rightarrow \mathbb{R}^m, x \mapsto Ax$ mit einer festen Matrix $A \in \mathbb{R}^{m \times n}$.

Keine linearen Abbildungen sind zum Beispiel:

- $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2, (x, y) \mapsto (x^2, x + y)$.
- $t_a : V \rightarrow V, v \mapsto v + a, a \in V, a \neq o$ (Parallelverschiebung).

Sind zwei lineare Abbildungen $f, g : V \rightarrow W$ gegeben, so ist auch die Summe

$$f + g : V \rightarrow W, \quad v \mapsto f(v) + g(v) \quad (2.145)$$

und das α -fache

$$\alpha f : V \rightarrow W, \quad v \mapsto \alpha \cdot f(v), \quad \alpha \in \mathbb{R} \quad (2.146)$$

lineare Abbildungen.

Mit dieser Addition und skalaren Multiplikation ist die Menge aller linearen Abbildungen von V nach W

$$\text{Hom}(V, W) = \{f : V \rightarrow W; f \text{ linear}\} \quad (2.147)$$

selbst wieder ein \mathbb{R} -Vektorraum.

Sind drei \mathbb{R} -Vektorräume U, V, W und zwei lineare Abbildungen

$$g: U \rightarrow V \quad \text{und} \quad f: V \rightarrow W \quad (2.148)$$

gegeben, so ist auch die Hintereinanderausführung („Komposition“)

$$f \circ g: U \rightarrow W, \quad u \mapsto f(g(u)) \quad (2.149)$$

eine lineare Abbildung.

Ist V ein endlichdimensionaler \mathbb{R} -Vektorraum und $B = \{v_1, \dots, v_n\}$ eine Basis von V , dann läßt sich jeder Vektor $v \in V$ eindeutig in der Form

$$v = \alpha_1 v_1 + \dots + \alpha_n v_n \quad (2.150)$$

mit $\alpha_i \in \mathbb{R}$ darstellen. Man nennt

$$v_B := \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \in \mathbb{R}^n \quad \left(\text{mit } v = \sum_{i=1}^n \alpha_i v_i \right) \quad (2.151)$$

den Koordinatenvektor von $v \in V$ bezüglich B . Die lineare Abbildung

$$k_B: V \rightarrow \mathbb{R}^n, \quad v \mapsto v_B \quad (2.152)$$

heißt Koordinatenabbildung bezüglich B .

Mit dieser Abbildung überträgt man Probleme aus V in den \mathbb{R}^n (die Grundidee dieser Vorgehensweise stammt von R. DESCARTES).

Man kann die Abbildung k_B auch umkehren:

$$k_B^{-1}: \mathbb{R}^n \rightarrow V, \quad (\alpha_1, \dots, \alpha_n) \mapsto \sum_{i=1}^n \alpha_i v_i. \quad (2.153)$$

Im Folgenden untersuchen wir lineare Abbildungen $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$.

Satz 2.41 (Darstellung linearer Abbildungen)

Sei $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ eine lineare Abbildung und $F = (f(e_1), \dots, f(e_n))$ diejenige $(n \times n)$ -Matrix in Spaltendarstellung, deren Spalten aus den Bildern der Einheitsvektoren e_1, \dots, e_n (natürliche Basis des \mathbb{R}^n) besteht, dann gilt:

$$f(x) = Fx \quad \text{für alle } x \in \mathbb{R}^n \quad (2.154)$$

Man nennt F die Abbildungsmatrix von f bezüglich der natürlichen Basis e_1, \dots, e_n .

Im Folgenden untersuchen wir die Begriffe „Länge, Winkel und Orthogonalität“ im \mathbb{R}^n .

Definition 2.42 (Skalarprodukt, Betrag, Länge)

Seien $x, y \in \mathbb{R}^n$, $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, $y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$.

Die Zahl

$$x^\top y = \sum_{i=1}^n x_i y_i \quad (2.155)$$

heißt Skalarprodukt der Vektoren x, y und

$$|x| := \sqrt{x^\top x} = \sqrt{x_1^2 + \dots + x_n^2} \quad (2.156)$$

heißt Betrag oder Länge des Vektors x .

Beispiel(e) 2.43

$$x = \begin{pmatrix} 2 \\ 0 \\ 1 \\ 4 \end{pmatrix}, \quad y = \begin{pmatrix} -1 \\ 3 \\ 5 \\ -3 \end{pmatrix} \quad \Rightarrow \quad x^\top y = -9 \quad |x| = \sqrt{21} \quad |y| = \sqrt{44}. \quad (2.157)$$

Ein Vektor x heißt normiert oder Einheitsvektor, falls $|x| = 1$.

Zu jedem Vektor $x \neq o$ ist $\frac{1}{|x|} \cdot x$ normiert.

Wie im \mathbb{R}^2 und \mathbb{R}^3 lassen sich die folgenden Rechenregeln für alle $x, y, z \in \mathbb{R}^n$ und alle $\alpha \in \mathbb{R}$ nachweisen:

- (a) $x^\top y = y^\top x$
- (b) $\alpha(x^\top y) = (\alpha x)^\top y = x^\top (\alpha y)$
- (c) $x^\top (y + z) = x^\top y + x^\top z$
- (d) $x^\top x > 0$ für alle $x \neq o$
- (e) $|x| = 0 \iff x = o$ (2.158)
- (f) $|\alpha x| = |\alpha| \cdot |x|$
- (g) $|x^\top y| \leq |x| \cdot |y|$ (Cauchy-Schwarzsche Ungleichung)
- (h) $|x + y| \leq |x| + |y|$ (Dreiecksungleichung)

Wegen (g) gilt für $x \neq o, y \neq o$ stets $\frac{|x^\top y|}{|x| \cdot |y|} \leq 1$ und somit

$$-1 \leq \frac{x^\top y}{|x| \cdot |y|} \leq 1. \quad (2.159)$$

Deshalb gibt es genau einen Winkel φ mit $0 \leq \varphi \leq \pi$ und

$$\cos(\varphi) = \frac{x^\top y}{|x| \cdot |y|} \quad (2.160)$$

Wir legen daher fest:

Definition 2.44 (Winkel zwischen Vektoren, orthogonale Vektoren)

Sei $x, y \in \mathbb{R}^n$ und $x, y \neq 0$, so nennt man $\angle(x, y) := \varphi = \cos^{-1} \left(\frac{x^\top y}{|x||y|} \right)$ (auch als $\arccos \left(\frac{x^\top y}{|x||y|} \right)$ geschrieben) den Winkel zwischen x und y , wobei $\cos^{-1} : [-1, 1] \rightarrow [0, \pi]$ die Umkehrfunktion der auf $[0, \pi]$ eingeschränkten Cosinus-Funktion bezeichnet.

Ferner nennt man $x, y \in \mathbb{R}^n$ orthogonal, falls $x^\top y = 0$.

Der Nullvektor ist orthogonal zu allen Vektoren im \mathbb{R}^n .

Der wichtige Begriff der Orthogonalität wird nun auf lineare Abbildungen, Matrizen und Basen erweitert.

Definition 2.45 (Orthogonalität von linearen Abbildungen, Matrizen und Basen)

(L.1) Sei $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ eine lineare Abbildung, dann heißt f orthogonal, falls sie das Skalarprodukt invariant läßt, d.h. wenn für alle $x, y \in \mathbb{R}^n$ gilt:

$$f(x)^\top f(y) = x^\top y. \quad (2.161)$$

(L.2) Eine Matrix $A \in \mathbb{R}^{n \times n}$ heißt orthogonal, falls

$$A^\top A = E \quad (\text{also } A^\top = A^{-1}). \quad (2.162)$$

(L.3) Eine Basis $\{b_1, \dots, b_n\}$ heißt orthogonal, falls die Vektoren b_i paarweise orthogonal sind; d.h. wenn für $i \neq j$ gilt

$$b_i^\top b_j = 0. \quad (2.163)$$

Die Basis B heißt orthonormal, wenn sie orthogonal ist und alle Basisvektoren normiert sind.

Nach der Definition orthogonaler Vektoren ist der Winkel zwischen diesen Vektoren gleich $\frac{\pi}{2}$, d.h. diese Vektoren stehen senkrecht aufeinander. Der Begriff „orthogonal“ kommt aus dem Altgriechischen „orthos“ gerade, richtig und „gony“ das Knie. Der „rechte“ Winkel ($= \frac{\pi}{2}$) zwischen Ober- und Unterschenkel ist für das Knie die „richtige“ Sitzposition.

Beispiel(e) 2.46

- Die natürliche Basis $\{e_1, \dots, e_n\}$ ist eine Orthonormalbasis des \mathbb{R}^n

Den Zusammenhang zwischen den Orthonormalitätsbegriffen klärt der folgende

Satz 2.47 (Zusammenhang zwischen den Orthogonalitätsbegriffen)
 Für eine Matrix $A \in \mathbb{R}^{n \times n}$ sind äquivalent:

(L.1) A ist orthogonal

(L.2) $(Ax)^\top Ay = x^\top y$ für alle $x, y \in \mathbb{R}^n$

(L.3) Die Spalten von A bilden eine orthonormale Basis des \mathbb{R}^n .
 Insbesondere ist eine lineare Abbildung $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ genau dann orthogonal, wenn die Abbildungsmatrix $F = (f(e_1), \dots, f(e_n))$ orthogonal ist.

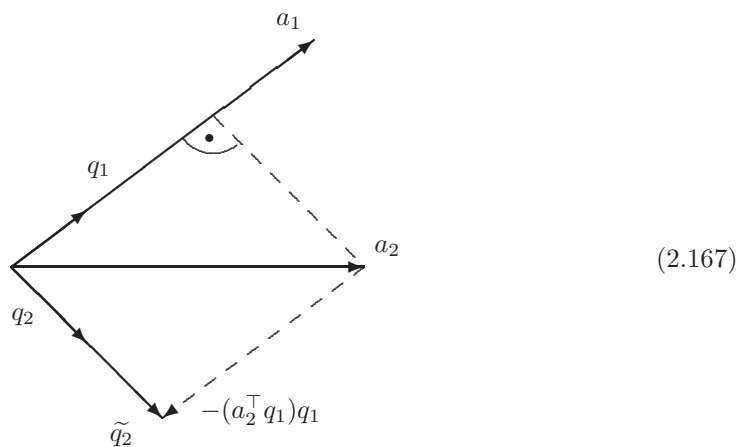
Seien $\{a_1, \dots, a_k\}$ linear unabhängige Vektoren des \mathbb{R}^n , so ermöglicht das nun folgende Gram-Schmidtsche Orthonormalisierungsverfahren die Berechnung von Vektoren $\{q_1, \dots, q_k\}$ mit

$$q_i^\top q_j = \begin{cases} 1 & \text{falls } i = j \\ 0 & \text{sonst} \end{cases} \quad (2.164)$$

und

$$\text{Lin}(\{a_1, \dots, a_k\}) = \text{Lin}(\{q_1, \dots, q_k\}) : \quad (2.165)$$

$$q_1 := \frac{a_1}{|a_1|}, \quad \tilde{q}_i := a_i - \sum_{m=1}^{i-1} (a_i^\top q_m) q_m, \quad q_i = \frac{1}{|\tilde{q}_i|} \tilde{q}_i, \quad i = 2, \dots, k. \quad (2.166)$$



Fasst man eine Basis $\{a_1, \dots, a_n\}$ des \mathbb{R}^n zu einer $(n \times n)$ -Matrix $A = (a_1, \dots, a_n)$ in Spalten-darstellung zusammen, so erhält man bei der Anwendung des Gram-Schmidtschen Orthonormalisierungsverfahrens auf $\{a_1, \dots, a_n\}$ eine Orthonormalbasis $\{q_1, \dots, q_n\}$ des \mathbb{R}^n , welche zu einer Matrix $Q = (q_1, \dots, q_n)$ in Spalten-darstellung führt.

Zwischen den Matrizen A und Q besteht nun der folgende Zusammenhang:

$$Q = A\tilde{R}, \quad (2.168)$$

wobei \tilde{R} eine obere Dreiecksmatrix darstellt, da für die Berechnung der i -ten Spalte von Q nur die ersten i Spalten von A benötigt werden.

Da die Matrix \tilde{R} invertierbar ist, kann man nach A auflösen:

$$A = Q \cdot \tilde{R}^{-1}. \quad (2.169)$$

Da nun $R := \tilde{R}^{-1}$ erneut eine obere Dreiecksmatrix ist, wird die Gleichung

$$A = QR \quad (2.170)$$

als QR -Zerlegung einer invertierbaren Matrix $A \in \mathbb{R}^{n \times n}$ mit einer orthogonalen Matrix Q und einer oberen Dreiecksmatrix R bezeichnet.

Die QR -Zerlegung wird häufig bei numerischen Verfahren angewendet.

Sei $\{b_1, \dots, b_n\}$ eine Basis des \mathbb{R}^n , die wir zu einer invertierbaren $(n \times n)$ -Matrix $B = (b_1, \dots, b_n)$ in Spaltendarstellung zusammenfassen. Die eindeutig bestimmten Koeffizienten $x'_1, \dots, x'_n \in \mathbb{R}$ der Zerlegung

$$x = \sum_{i=1}^n x'_i b_i \quad (2.171)$$

eines Vektors $x \in \mathbb{R}^n$ heißen Koordinaten des Vektors $x \in \mathbb{R}^n$ bezüglich der Basis B . Man nennt

$$x_B := \begin{pmatrix} x'_1 \\ \vdots \\ x'_n \end{pmatrix} = x' \quad \text{mit} \quad x = \sum_{i=1}^n x'_i b_i \quad (2.172)$$

den Koordinatenvektor von $x \in \mathbb{R}^n$ bezüglich B . Die Abbildung $x \mapsto x_B$ ist linear und es gilt:

$$x = Bx_B \quad (\text{Basiswechsel von } \{b_1, \dots, b_n\} \text{ zu } \{e_1, \dots, e_n\}). \quad (2.173)$$

$$x_B = B^{-1}x \quad (\text{Basiswechsel von } \{e_1, \dots, e_n\} \text{ zu } \{b_1, \dots, b_n\}) \quad (2.174)$$

Jeder $(n \times n)$ -Matrix A läßt sich die lineare Abbildung $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$, $x \mapsto Ax$ zuordnen; dabei ist A die Abbildungsmatrix von f bezüglich der natürlichen Basis. Nun betrachten wir die Abbildungsmatrix von f bezüglich einer anderen Basis des \mathbb{R}^n .

Definition 2.48 (Abbildungsmatrix bezüglich B)

Seien $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ eine lineare Abbildung und $B = (b_1, \dots, b_n) \in \mathbb{R}^{n \times n}$ eine invertierbare Matrix in Spaltendarstellung. Die Matrix

$$C := (f(b_1)_B, \dots, f(b_n)_B), \quad (2.175)$$

in deren Spalten die Bilder der Basisvektoren b_1, \dots, b_n bezüglich der Basis B stehen, heißt Abbildungsmatrix von f bezüglich B .

Mit Definition 2.48 können wir nun die Änderung der Abbildungsmatrix einer linearen Abbildung bei Basiswechsel beschreiben.

Satz 2.49 (Änderung der Abbildungsmatrix bei Basiswechsel)

Seien $A \in \mathbb{R}^{n \times n}$ und $B = (b_1, \dots, b_n)$ eine invertierbare $(n \times n)$ -Matrix in Spaltendarstellung. Ein Vektor $y = Ax$ lautet in Koordinaten bezüglich der Basis $\{b_1, \dots, b_n\}$:

$$y_B = (Ax)_B = B^{-1}Ax = B^{-1}ABx_B. \quad (2.176)$$

Für die Abbildungsmatrix C von $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$, $x \mapsto Ax$ bezüglich B gilt daher:

$$C = B^{-1}AB. \quad (2.177)$$

Zu einer linearen Abbildung $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$, $x \mapsto Ax$ gibt es somit je nach Wahl der Basis neben der Matrix A noch verschiedene andere Abbildungsmatrizen, die die lineare Abbildung f repräsentieren.

Definition 2.50 (ähnliche Matrizen)

Zwei $(n \times n)$ -Matrizen A, C heißen ähnlich, falls es eine invertierbare $(n \times n)$ -Matrix B gibt mit $C = B^{-1}AB$

Ähnliche Matrizen repräsentieren dieselbe lineare Abbildung aber unter Verwendung verschiedener Basen. Es stellt sich im Folgenden die Frage, wie eine Basis des \mathbb{R}^n gewählt werden muss, damit die zugehörige Abbildungsmatrix einer gegebenen linearen Abbildung möglichst einfach wird. Diese Fragestellung wird als algebraisches Eigenwertproblem bezeichnet.

Die Behandlung des algebraischen Eigenwertproblems ist eng verbunden mit der Nullstellenbestimmung von Polynomen, also von Funktionen

$$f : \mathbb{C} \rightarrow \mathbb{C}, \quad z \mapsto a_n z^n + a_{n-1} z^{n-1} + \dots + a_1 z + a_0 \quad (2.178)$$

mit $a_i \in \mathbb{C}, i = 1, \dots, n$.

Daher lassen wir im Folgenden auch komplexe Zahlen als Skalare und als Matrixelemente zu. Die bisherigen Untersuchungen über Vektoren und Matrizen ändern sich dadurch nicht.

Definition 2.51 (Eigenwert, Eigenvektor)

Eine Zahl $\lambda \in \mathbb{C}$ heißt Eigenwert einer Matrix $A \in \mathbb{C}^{n \times n}$, falls es wenigstens einen Spaltenvektor $v \in \mathbb{C}^n, v \neq 0$, gibt mit

$$Av = \lambda v. \quad (2.179)$$

Jeder Vektor $v \neq 0$, der diese Gleichung erfüllt, heißt Eigenvektor von A zum Eigenwert λ .

Beispiel(e) 2.52

Die Matrix $A = \begin{pmatrix} 0 & -1 & 0 \\ -1 & -1 & 1 \\ 0 & 1 & 0 \end{pmatrix}$ hat den Eigenwert $\lambda = -2$ mit dem Eigenvektor $v = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}$, denn es gilt:

$$Av = -2 \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}$$

Zur Berechnung der Eigenwerte einer $(n \times n)$ -Matrix betrachtet man das charakteristische Polynom χ_A :

$$\chi_A : \mathbb{C} \rightarrow \mathbb{C}, \quad \lambda \mapsto \det(A - \lambda \cdot E_n) \quad (2.180)$$

Diese Determinante muss berechnet werden.

Definition 2.53 (Eigenwerte und charakteristisches Polynom)

Es ist $\lambda \in \mathbb{C}$ genau dann ein Eigenwert der $(n \times n)$ -Matrix A , falls λ eine Nullstelle des charakteristischen Polynoms χ_A ist, d.h. falls

$$\det(A - \lambda E_n) = 0. \quad (2.181)$$

Die Berechnung der Eigenwerte erfolgt daher in zwei Schritten:

- (1) Berechne die Nullstellen $\lambda_1, \dots, \lambda_r$ der charakteristischen Gleichung

$$\chi_A(\lambda) = 0. \quad (2.182)$$

Die Vielfachheit k_i der Nullstelle λ_i heißt algebraische Vielfachheit des Eigenwerts λ_i .

- (2) Berechne zu jedem Eigenwert λ_i , $i = 1, \dots, r$, den Lösungsraum des homogenen linearen Gleichungssystems $(A - \lambda_i E_n)x = o$.
Jede Lösung $v \neq o$ ist ein Eigenvektor zu λ_i .

$$V(\lambda_i) := \{v \in \mathbb{C}^n; (A - \lambda_i E_n)v = o\} = \text{Kern}(A - \lambda_i E_n) \quad (2.183)$$

heißt Eigenraum zum Eigenwert λ_i . Man nennt $\text{Dim}(V(\lambda_i))$ die geometrische Vielfachheit des Eigenwertes λ_i .

Algebraische und geometrische Vielfachheit stimmen im allgemeinen nicht überein!

Beispiel(e) 2.54

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

$$\det(A - \lambda E_n) = \det \begin{pmatrix} 1 - \lambda & 2 \\ 2 & 1 - \lambda \end{pmatrix} = (1 - \lambda)^2 - 4 = \chi_A(\lambda)$$

$$\chi_A(\lambda) = 0 \iff (1 - \lambda)^2 = 4 \iff \lambda_1 = -1, \quad \lambda_2 = +3.$$

Eigenvektoren zu $\lambda_1 = -1$

$$\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} x = -x \iff \begin{pmatrix} 2 & 2 \\ 2 & 2 \end{pmatrix} x = 0 \iff x = \mu \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \quad \mu \in \mathbb{R},$$

$$V(\lambda_1) = \left\{ \mu \begin{pmatrix} -1 \\ 1 \end{pmatrix}; \mu \in \mathbb{R} \right\}.$$

Eigenvektoren zu $\lambda_2 = +3$

$$\begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} x = 3x \iff \begin{pmatrix} -2 & 2 \\ 2 & -2 \end{pmatrix} x = 0 \iff x = \mu \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \mu \in \mathbb{R},$$

$$V(\lambda_2) = \left\{ \mu \begin{pmatrix} 1 \\ 1 \end{pmatrix}; \mu \in \mathbb{R} \right\},$$

Algebraische Vielfachheit gleich geometrische Vielfachheit.

- $A = \begin{pmatrix} 3 & 0 \\ 0 & 3 \end{pmatrix}$, $\lambda_1 = \lambda_2 = 3$, $V(\lambda_1) = \mathbb{R}^2$.

- $A = \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix}$, $\lambda_1 = \lambda_2 = 2$, $V(\lambda_1) = \left\{ \mu \begin{pmatrix} 1 \\ 0 \end{pmatrix}; \mu \in \mathbb{R} \right\}$.

Algebraische Vielfachheit gleich 2, geometrische Vielfachheit gleich 1.

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

$$\lambda_1 = \cos(\varphi) + i \sin(\varphi), \quad \lambda_2 = \cos(\varphi) - i \sin(\varphi),$$

also im allgemeinen keine reellen Eigenwerte.

$$V(\lambda_1) = \left\{ \mu \begin{pmatrix} 1 \\ -i \end{pmatrix}, \mu \in \mathbb{C} \right\}, \quad V(\lambda_2) = \left\{ \mu \begin{pmatrix} 1 \\ i \end{pmatrix}, \mu \in \mathbb{C} \right\}.$$

Im folgenden Satz stellen wir nützliche Beobachtungen zusammen:

Satz 2.55 (Eigenschaften von Eigenwerten und Eigenvektoren)

Sei $A \in \mathbb{C}^{n \times n}$:

- (a) A und A^\top besitzen dieselben Eigenwerte, jedoch im allgemeinen verschiedene Eigenräume.
- (b) Ähnliche Matrizen A und $B^{-1}AB$ haben dasselbe charakteristische Polynom und deshalb dieselben Eigenwerte. b ist genau dann Eigenvektor von A zum Eigenwert λ , falls $B^{-1}b$ Eigenvektor von $B^{-1}AB$ zum Eigenwert λ ist.
- (c) A ist genau dann invertierbar, falls alle Eigenwerte ungleich Null sind.
Ist λ ein Eigenwert von A mit Eigenvektor b (A sei invertierbar), so ist $\frac{1}{\lambda}$ Eigenwert von A^{-1} mit Eigenvektor b .
- (d) Eigenvektoren b_1, \dots, b_r zu paarweise verschiedenen Eigenwerten $\lambda_1, \dots, \lambda_r$ sind linear unabhängig.
- (e) Besitzt A n linear unabhängige Eigenvektoren b_1, \dots, b_n zu den nicht notwendig verschiedenen Eigenwerten $\lambda_1, \dots, \lambda_n$, so gilt mit $B := (b_1, \dots, b_n)$:

$$B^{-1}AB = \begin{pmatrix} \lambda_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \lambda_n \end{pmatrix}. \quad (2.184)$$

Sei nun $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ eine lineare Abbildung mit der Abbildungsmatrix A bezüglich der natürlichen Basis. Besitzt A n linear unabhängige Eigenvektoren $\{b_1, \dots, b_n\}$ zu den Eigenwerten $\lambda_1, \dots, \lambda_n$, so kann f auch durch die Abbildungsmatrix

$$C = \begin{pmatrix} \lambda_1 & & \mathbf{0} \\ & \ddots & \\ \mathbf{0} & & \lambda_n \end{pmatrix} \quad (2.185)$$

bezüglich der Basis $\{b_1, \dots, b_n\}$ dargestellt werden. Die Matrix C läßt die Eigenschaften von f im allgemeinen besser erkennen als die Matrix A .

Teil II

Mathematik 2

Kapitel 3

Analysis einer reellen Veränderlichen

3.1 Funktionen, Grenzwerte, Stetigkeit

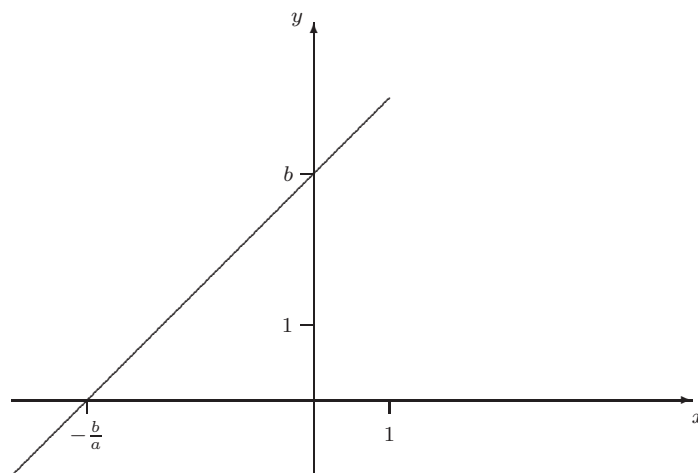
Zunächst betrachten wir reelle Funktionen einer reellen Veränderlichen

$$f : \mathbb{D} \rightarrow \mathbb{R}, \quad x \mapsto f(x) \quad (3.1)$$

mit der Variablen $x \in \mathbb{D} \subseteq \mathbb{R}$. Jede Funktion dieser Art wird durch einen Graph (die Kurve) $y = f(x)$ veranschaulicht, der aus denjenigen Punkten (x, y) einer mit kartesischen (x, y) -Koordinaten versehenen Ebene besteht, für die $x \in \mathbb{D}$ und $y = f(x)$ gilt.

Affin-lineare Funktionen

$f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto ax + b$ mit Konstanten $a, b \in \mathbb{R}$. Die Funktion heißt affin-lineare Funktion. Der Graph dieser Funktion stellt eine Gerade durch die Punkte $(0, b)$ und $(-\frac{b}{a}, 0)$ dar, falls $a \neq 0$. Im Fall $a = 0$ ist f eine konstante Funktion $x \mapsto b$. Ihr Graph stellt die zur x -Achse parallele Gerade $y = b$ dar.



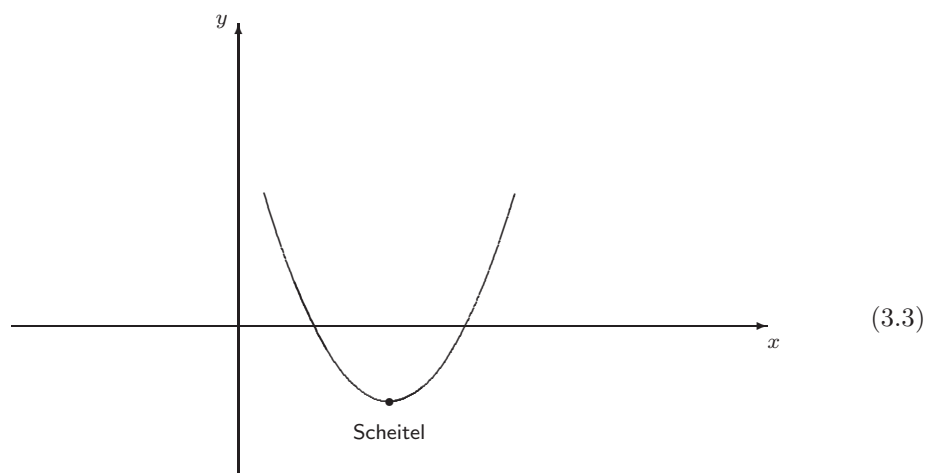
(3.2)

Quadratische Funktionen

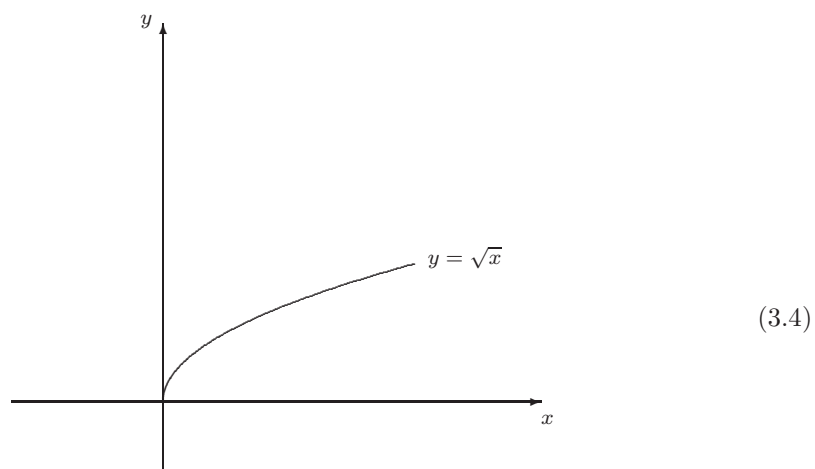
$f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto ax^2 + bx + c$ mit Konstanten $a, b, c \in \mathbb{R}$. Der Graph einer quadratischen Funktion stellt eine Parabel dar, deren Scheitel mittels quadratischer Ergänzung bestimmt wird ($a \neq 0$):

$$\begin{aligned} y &= ax^2 + bx + c \\ &= a \left(x^2 + \frac{b}{a}x \right) + c \\ &= a \left(x^2 + \frac{b}{a}x + \frac{b^2}{4a^2} - \frac{b^2}{4a^2} \right) + c \\ &= a \left(x + \frac{b}{2a} \right)^2 - \frac{b^2}{4a} + c \end{aligned}$$

Es handelt sich somit um eine verschobene Parabel $y = ax^2$, deren Scheitel im Punkt $\left(-\frac{b}{2a}, c - \frac{b^2}{4a}\right)$ liegt.

**Die Wurzelfunktion**

$f : \mathbb{R}_0^+ \rightarrow \mathbb{R}_0^+, x \mapsto \sqrt{x}$

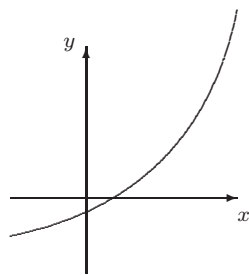


wobei $\mathbb{R}_0^+ := \{x \in \mathbb{R}; x \geq 0\}$.

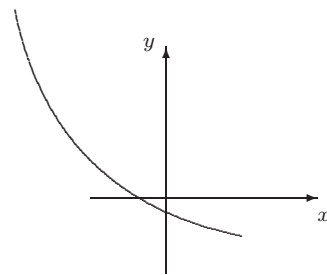
Ein für die genauere Analyse von Funktionen wichtiger Begriff ist die Monotonie.

Definition 3.1 (Monotonie)
 Eine Funktion $f : \mathbb{D} \rightarrow \mathbb{R}$, $x \mapsto f(x)$ heißt

- (a) monoton wachsend (bzw. monoton fallend), falls für alle $x_1, x_2 \in \mathbb{D}$ mit $x_2 > x_1$ gilt:
 $f(x_2) \geq f(x_1)$ (bzw. $f(x_2) \leq f(x_1)$)
- (b) streng monoton wachsend (bzw. streng monoton fallend), falls für alle $x_1, x_2 \in \mathbb{D}$ mit $x_2 > x_1$ gilt: $f(x_2) > f(x_1)$ (bzw. $f(x_2) < f(x_1)$).



streng monoton wachsend



streng monoton fallend

(3.5)

Zu Funktionen $f : \mathbb{D} \rightarrow \mathbb{R}$, $g : \mathbb{D} \rightarrow \mathbb{R}$ mit gleichem Definitionsbereich definiert man die Summe, Differenz und das Produkt folgendermaßen:

$$f + g : \mathbb{D} \rightarrow \mathbb{R}, \quad x \mapsto f(x) + g(x) \tag{3.6}$$

$$f - g : \mathbb{D} \rightarrow \mathbb{R}, \quad x \mapsto f(x) - g(x) \tag{3.7}$$

$$f \cdot g : \mathbb{D} \rightarrow \mathbb{R}, \quad x \mapsto f(x) \cdot g(x) \tag{3.8}$$

Der Quotient $\frac{f}{g} : \mathbb{D} \rightarrow \mathbb{R}$, $x \mapsto \frac{f(x)}{g(x)}$ ist nur erklärt, falls $g(x) \neq 0$ für alle $x \in \mathbb{D}$.

Mit dem Produkt sind auch die Potenzen $f^n : \mathbb{D} \rightarrow \mathbb{R}$, $x \mapsto (f(x))^n$, $n \in \mathbb{N}$, definiert. Das α -fache, $\alpha \in \mathbb{R}$, einer Funktion $f : \mathbb{D} \rightarrow \mathbb{R}$ ist definiert durch $\alpha f : \mathbb{D} \rightarrow \mathbb{R}$, $x \mapsto \alpha \cdot f(x)$.

Zu Funktionen $f : \mathbb{D} \rightarrow \mathbb{R}$ und $g : I \rightarrow \mathbb{R}$ mit $g(I) \subseteq \mathbb{D}$, wobei $g(I) := \{g(x) \in \mathbb{R}; x \in I\}$, kann man die Komposition $h = f \circ g$ definiert durch $h : I \rightarrow \mathbb{R}$, $x \mapsto f(g(x))$ bilden.

Eine Funktion f heißt Polynom vom Grad n , wenn es reelle Zahlen a_0, a_1, \dots, a_n , $a_n \neq 0$, gibt mit

$$f : \mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto a_0 + a_1x + a_2x^2 + \dots + a_nx^n = \sum_{i=0}^n a_i x^i. \tag{3.9}$$

Die a_k heißen Koeffizienten des Polynoms. Das Nullpolynom $x \mapsto 0$ hat keinen Grad. Das Rechnen mit Polynomen in der Summendarstellung ist sehr übersichtlich

$$\sum_{k=0}^n a_k x^k \pm \sum_{k=0}^n b_k x^k = \sum_{k=0}^n (a_k \pm b_k) x^k. \tag{3.10}$$

Zwei Polynome $x \mapsto \sum_{k=0}^n a_k x^k$ und $x \mapsto \sum_{k=0}^n b_k x^k$ sind genau dann gleich, wenn $a_k = b_k$ für alle $k \in \{0, \dots, n\}$ gilt.

Wird ein Funktionswert eines Polynoms $f : \mathbb{R} \rightarrow \mathbb{R}$, $x \mapsto \sum_{k=0}^n a_k x^k$ aus der Summendarstellung berechnet, so benötigt man dafür $2n - 1$ Multiplikationen und n Additionen. Berechnet man die Werte jedoch aus der Horner-Darstellung

$$x \mapsto (\dots ((a_n x + a_{n-1}) \cdot x + a_{n-2}) \cdot x + \dots + a_1) x + a_0 \tag{3.11}$$

- von innen nach außen - so wird der Aufwand auf n Multiplikationen und n Additionen reduziert. Als Nullstelle einer Funktion $f : \mathbb{D} \rightarrow \mathbb{R}$ bezeichnet man jede Lösung der Gleichung $f(x) = 0$ in \mathbb{D} . In praktischen Fällen betrachtet man oft die Nullstellen von Polynomen.

Eine Zahl $b \in \mathbb{R}$ ist genau dann Nullstelle eines Polynoms f , wenn es ein Polynom $h : \mathbb{R} \rightarrow \mathbb{R}$ gibt mit $f(x) = (x - b) \cdot h(x)$ für alle $x \in \mathbb{R}$.

Der Faktor $(x - b)$ kann auch mehrfach vorkommen, nämlich dann, wenn auch $h(b) = 0$ und somit $h(x) = (x - b) \cdot h_1(x)$ gilt. Man nennt b l -fache Nullstelle von f und l die Vielfachheit von b , wenn es ein Polynom $g : \mathbb{R} \rightarrow \mathbb{R}$ gibt mit

$$f(x) = (x - b)^l \cdot g(x) \quad \text{und} \quad g(b) \neq 0. \quad (3.12)$$

Jede weitere Nullstelle $c \neq b$ von f ist auch Nullstelle von g .

Beispiel(e) 3.2

$$x^5 - 5x^4 + 14x^3 - 22x^2 + 17x - 5 = (x - 1)^3(x^2 - 2x + 5); \quad (3.13)$$

$x = 1$ ist dreifache Nullstelle des entsprechenden Polynoms.

Sind nun b_1, \dots, b_r die verschiedenen reellen Nullstellen des Polynoms

$$f : \mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto \sum_{k=0}^n a_k x^k \quad (3.14)$$

mit Grad $f \geq 1$, der jeweiligen Vielfachheit l_1, \dots, l_r , so gibt es ein Polynom $q : \mathbb{R} \rightarrow \mathbb{R}$ vom Grad $(n - l_1 - l_2 - \dots - l_r)$ mit

$$f(x) = (x - b_1)^{l_1} \cdot \dots \cdot (x - b_r)^{l_r} \cdot q(x) \quad \text{für alle } x \in \mathbb{R}. \quad (3.15)$$

Somit hat jedes Polynom vom Grad $n \geq 1$ höchstens n Nullstellen.

Der Quotient

$$\frac{p}{q} : \mathbb{D} \rightarrow \mathbb{R}, \quad x \mapsto \frac{\sum_{k=0}^n a_k x^k}{\sum_{k=0}^m b_k x^k} \quad (\text{mit } a_n \neq 0, b_m \neq 0) \quad (3.16)$$

heißt rationale Funktion. Der Definitionsbereich \mathbb{D} von $\frac{p}{q}$ darf keine Nullstelle des Nennerpolynoms q enthalten. Ist der Grad des Zählerpolynoms p größer oder gleich dem Grad des Nennerpolynoms q (also $n \geq m$), so erhält man mit Hilfe des Polynoms $p_1 : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto p(x) - \frac{a_n}{b_m} x^{n-m} \cdot q(x)$ durch Einsetzen die folgende Darstellung:

$$\frac{p}{q} : \mathbb{D} \rightarrow \mathbb{R}, \quad x \mapsto \frac{a_n}{b_m} x^{n-m} + \frac{p_1(x)}{q(x)}, \quad (3.17)$$

wobei p_1 entweder die Nullfunktion ist oder ein Polynom mit kleinerem Grad als dem Grad von p . Sei nun $p_1 = c_s x^s + \dots + c_1 x^1 + c_0$ mit $c_s \neq 0$ und $s \geq \text{Grad } q$, so können wir die obige Vorgehensweise für die rationale Funktion $\frac{p_1}{q}$ wiederholen und erhalten

$$\frac{p}{q} : \mathbb{D} \rightarrow \mathbb{R}, \quad x \mapsto \frac{a_n}{b_m} x^{n-m} + \frac{c_s}{b_m} x^{s-m} + \frac{p_2(x)}{q(x)} \quad (3.18)$$

mit $p_2 : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto p_1(x) - \frac{c_s}{b_m} x^{s-m} q(x)$. Ist nun erneut $\text{Grad } p_2 \geq \text{Grad } q$, so wiederholt man die Vorgehensweise für $\frac{p_2}{q}$, ansonsten ist das Verfahren beendet und man erhält die Existenz zweier Polynome $g, r : \mathbb{R} \rightarrow \mathbb{R}$ mit

$$\frac{p}{q} : \mathbb{D} \rightarrow \mathbb{R}, \quad x \mapsto g(x) + \frac{r(x)}{q(x)} \quad \text{mit Grad } r < \text{Grad } q. \quad (3.19)$$

Die eben skizzierte Vorgehensweise wird als Polynomdivision mit Rest bezeichnet.

Beispiel(e) 3.3

$$\begin{aligned}
 (x^{12} + x^6 + x + 1) : (2x^4 + 3) &= \underbrace{\frac{1}{2}x^8}_{\frac{a^n}{b^m}x^{n-m}} - \frac{3}{4}x^4 + \frac{1}{2}x^2 + \frac{9}{8} + \frac{-\frac{3}{2}x^2 + x - \frac{19}{8}}{2x^4 + 3} \\
 - \left(x^{12} + \frac{3}{2}x^8 \right) & \\
 \hline
 -\frac{3}{2}x^8 + x^6 + x + 1 & \quad (= p_1(x)) \\
 - \left(-\frac{3}{2}x^8 - \frac{9}{4}x^4 \right) & \\
 \hline
 x^6 + \frac{9}{4}x^4 + x + 1 & \quad (= p_2(x)) \\
 - \left(x^6 + \frac{3}{2}x^2 \right) & \\
 \hline
 \frac{9}{4}x^4 - \frac{3}{2}x^2 + x + 1 & \quad (= p_3(x)) \\
 - \left(\frac{9}{4}x^4 + \frac{27}{8} \right) & \\
 \hline
 -\frac{3}{2}x^2 + x - \frac{19}{8} & \quad (= r(x))
 \end{aligned}$$

Mit Hilfe der Polynomdivision kann bei der bereits betrachteten Zerlegung

$$f : \mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto (x - b)^l \cdot g(x) \quad \text{mit } g(b) \neq 0 \quad (3.20)$$

das Polynom g bei gegebenem Polynom f , bei gegebener Nullstelle b und gegebener Vielfachheit l berechnet werden.

Beispiel(e) 3.4

$$f : \mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto x^4 + 2x^3 - 3x^2 - 4x + 4, \quad b = 1, \quad l = 2$$

$$\begin{array}{r} (x^4 + 2x^3 - 3x^2 - 4x + 4) : \underbrace{(x^2 - 2x + 1)}_{(x-1)^2} = x^2 + 4x + 4 \\ - (x^4 - 2x^3 + x^2) \\ \hline 4x^3 - 4x^2 - 4x + 4 \\ - (4x^3 - 8x^2 + 4x) \\ \hline 4x^2 - 8x + 4 \\ - (4x^2 - 8x + 4) \\ \hline - \end{array}$$

Also:

$$f : \mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto (x - 1)^2 \underbrace{(x^2 + 4x + 4)}_{g(x)}.$$

Ebenso wie für die rationalen Zahlen ist es auch für rationale Funktionen wichtig, den gemeinsamen Teiler des Zählers und des Nenners zu kürzen, wobei ein Polynom $d : \mathbb{R} \rightarrow \mathbb{R}$ Teiler eines Polynoms $p : \mathbb{R} \rightarrow \mathbb{R}$ heißt, falls es ein Polynom $p_0 : \mathbb{R} \rightarrow \mathbb{R}$ gibt mit $p = d \cdot p_0$.

Die Berechnung eventuell vorhandener gemeinsamer Polynomteiler von Polynomen p und q basiert auf der folgenden Beobachtung:

Falls $\text{Grad } p \geq \text{Grad } q$ und $\frac{p}{q} = h + \frac{r}{q}$ mit Polynomen h, r und $\text{Grad } r < \text{Grad } q$, dann ist das Polynom d genau dann gemeinsamer Teiler von p und q , falls d gemeinsamer Teiler von r und q ist. Denn aus $p = d \cdot p_0$ und $q = d \cdot q_0$ folgt $r = p - h \cdot q = d(p_0 - h \cdot q_0)$.

Umgekehrt: Aus $q = d \cdot q_0$ und $r = d \cdot r_0$ folgt

$$p = h \cdot q + r = d(h \cdot q_0 + r_0). \tag{3.21}$$

Mittels Polynomdivision wird also die Bestimmung der gemeinsamen Teiler von p und q auf die Bestimmung der gemeinsamen Teiler der Polynome kleineren Grades q und r reduziert. Dieser Reduktionsschritt muss gegebenenfalls mehrfach wiederholt werden (Euklidischer Algorithmus für Polynome)

Beispiel(e) 3.5
 Wir suchen den gemeinsamen Teiler von

$$p : \mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto x^4 - 2x^3 - 2x^2 - 2x - 3$$

und

$$q : \mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto x^4 - 3x^3 - 7x^2 + 15x + 18 :$$

$p : q:$

$$\begin{array}{r} (x^4 - 2x^3 - 2x^2 - 2x - 3) : (x^4 - 3x^3 - 7x^2 + 15x + 18) = 1 + \frac{x^3 + 5x^2 - 17x - 21}{q(x)} \\ - (x^4 - 3x^3 - 7x^2 + 15x + 18) \\ \hline x^3 + 5x^2 - 17x - 21 \end{array} = r(x)$$

$q : r:$

$$\begin{array}{r} (x^4 - 3x^3 - 7x^2 + 15x + 18) : (x^3 + 5x^2 - 17x - 21) = x - 8 + \frac{50x^2 - 100x - 150}{r(x)} \\ - (x^4 + 5x^3 - 17x^2 - 21x) \\ \hline -8x^3 + 10x^2 + 36x + 18 \\ - (-8x^3 - 40x^2 + 136x + 168) \\ \hline 50x^2 - 100x - 150 \end{array} = r_1(x)$$

$r : r_1:$

$$\begin{array}{r} (x^3 + 5x^2 - 17x - 21) : (50x^2 - 100x - 150) = \frac{1}{50}x + \frac{7}{50} \\ - (x^3 - 2x^2 - 3x) \\ \hline 7x^2 - 14x - 21 \\ - (7x^2 - 14x - 21) \\ \hline - \end{array}$$

Somit ist $r_1 : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto 50x^2 - 100x - 150$ der dem Grad nach größte gemeinsame Teiler von r und r_1 und somit von r und q und damit auch von p und q . Es gilt:

$$p : \mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto \frac{1}{50}(x^2 + 1)(50x^2 - 100x - 150) \tag{3.22}$$

und

$$q : \mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto \frac{1}{50}(x + 2)(x - 3)(50x^2 - 100x - 150). \tag{3.23}$$

Polynome werden häufig zur Interpolation von Punkten benötigt. Ausgehend von Messpunkten $(x_i, y_i) \in \mathbb{R}^2, i = 0, \dots, n$ mit $x_i \neq x_j$ für $i \neq j$ sucht man eine Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$ mit $f(x_i) = y_i$ für alle $i = 0, \dots, n$. Mit Hilfe dieser Funktion kann dann auch einem Argument $\hat{x} \neq x_i, i = 0, \dots, n$ ein Wert $\hat{y} = f(\hat{x})$ zugeordnet werden.

Da es zu $n + 1$ Stützpunkten $(x_0, y_0), \dots, (x_n, y_n)$ genau ein Polynom $p_n : \mathbb{R} \rightarrow \mathbb{R}$ vom Grad kleiner oder gleich n mit $p(x_i) = y_i$ gibt, verwendet man häufig dieses Polynom für die gesuchte Funktion f . Betrachtet man den Ansatz

$$p_n : \mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto \sum_{k=0}^n a_k x^k, \tag{3.24}$$

so können die gesuchten Koeffizienten a_0, \dots, a_n durch das lineare Gleichungssystem in $a = (a_0, \dots, a_n)$:

$$\begin{aligned} \sum_{k=0}^n a_k x_0^k &= y_0 \\ &\vdots \\ \sum_{k=0}^n a_k x_n^k &= y_n \end{aligned} \tag{3.25}$$

beziehungsweise

$$\begin{pmatrix} 1 & x_0 & x_0^2 & \dots & x_0^n \\ 1 & x_1 & x_1^2 & \dots & x_1^n \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \dots & x_n^n \end{pmatrix} \cdot a = \begin{pmatrix} y_0 \\ \vdots \\ y_n \end{pmatrix} \tag{3.26}$$

ausgerechnet werden. Wird ein zusätzlicher Stützpunkt (x_{n+1}, y_{n+1}) hinzugefügt, so muss für die Berechnung von p_{n+1} komplett neu gerechnet werden.

Das Polynom p_n kann einfacher berechnet werden, wenn man es in der Form

$$p_n : \mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto \alpha_0 + \alpha_1(x-x_0) + \alpha_2(x-x_0)(x-x_1) + \dots + \alpha_n(x-x_0) \cdot \dots \cdot (x-x_{n-1}) \tag{3.27}$$

darstellt. Aus der Forderung $p_n(x_i) = y_i, i = 0, \dots, n$, ergibt sich für die Unbekannten $\alpha_0, \alpha_1, \dots, \alpha_n$:

$$\begin{aligned} y_0 &= \alpha_0 \\ y_1 &= \alpha_0 + \alpha_1(x_1 - x_0) \\ &\vdots \\ y_n &= \alpha_0 + \alpha_1(x_n - x_0) + \dots + \alpha_n(x_n - x_0) \cdot \dots \cdot (x_n - x_{n-1}) \end{aligned} \tag{3.28}$$

beziehungsweise

$$\begin{pmatrix} 1 & 0 & 0 & & & \\ 1 & (x_1 - x_0) & 0 & & & \\ 1 & (x_2 - x_0) & (x_2 - x_0)(x_2 - x_1) & & & \\ & & & \ddots & & \\ 1 & (x_n - x_0) & \dots & \dots & (x_n - x_0) \cdot \dots \cdot (x_n - x_{n-1}) & \end{pmatrix} \begin{pmatrix} \alpha_0 \\ \vdots \\ \vdots \\ \vdots \\ \alpha_n \end{pmatrix} = \begin{pmatrix} y_0 \\ \vdots \\ \vdots \\ \vdots \\ y_n \end{pmatrix}. \tag{3.29}$$

Dieses Gleichungssystem kann nun „von oben nach unten“ gelöst werden. Die Idee dieser Vorgehensweise stammt von ISAAC NEWTON (1642-1727) und hat den großen Vorteil, dass problemlos weitere Stützpunkte $(x_{n+1}, y_{n+1}), \dots, (x_{n+m}, y_{n+m})$ in das obige Schema integriert werden können, ohne die Dreiecksform des linearen Gleichungssystems zu verlieren.

Eine der wichtigsten Vorgehensweisen der Analysis ist die Bildung von Grenzwerten. Dabei betrachtet man zunächst spezielle Funktionen $f : \mathbb{N}_0 \rightarrow \mathbb{R}$, die auch als Zahlenfolgen bezeichnet werden. Zahlenfolgen werden oft durch die Notation der Funktionswerte $a_0 = f(0), a_1 = f(1), a_2 = f(2) \dots$ in der Form $\{a_n\}, n \in \mathbb{N}_0$, notiert. Die Zahlen a_n heißen Glieder der Zahlenfolge. Das erste Glied einer Zahlenfolge muss nicht immer a_0 sein. Durch Umbenennung, etwa $b_0 := a_5, b_1 := a_6, b_n := a_{n+5}$ erreicht man, dass auch a_5, a_6, a_7, \dots eine Zahlenfolge im obigen Sinne darstellt. Man bezeichnet sie mit $\{a_n\}, n \geq 5$, (allgemein: $\{a_n\}, n \geq n_0$).

Beispiel(e) 3.6

- $a_n = a_0 + n \cdot d, d \in \mathbb{R}$ fest, $n \in \mathbb{N}_0$ (arithmetische Folge)
- $a_n = a_0 \cdot q^n, q \neq 0$ fest, $n \in \mathbb{N}_0$ (geometrische Folge)
- $a_{n+1} = \frac{1}{2}(a_n + \frac{2}{a_n}), n \geq 1, a_0 := 2$ (rekursiv definierte Zahlenfolge).

Eine Zahlenfolge heißt beschränkt, falls es Konstanten K_1 und K_2 gibt mit

$$K_1 \leq a_n \leq K_2 \quad \text{für alle } n \in \mathbb{N}_0. \quad (3.30)$$

Die Konstanten K_1 bzw. K_2 heißen untere bzw. obere Schranken der Zahlenfolge. Für die Analyse von Zahlenfolgen sind die beiden folgenden Begriffe wichtig:

Definition 3.7 (Häufungspunkt und Grenzwert einer Zahlenfolge)
 Sei $\{a_n\}, n \in \mathbb{N}_0$, eine Zahlenfolge. Eine Zahl $h \in \mathbb{R}$ heißt Häufungspunkt der Zahlenfolge $\{a_n\}$, falls für jedes offene Intervall (α, β) mit $h \in (\alpha, \beta)$ unendlich viele Folgenglieder der Zahlenfolge in (α, β) liegen.
 Eine Zahl $a \in \mathbb{R}$ heißt Grenzwert der Zahlenfolge $\{a_n\}$, falls es zu jedem reellen $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ gibt mit

$$|a_n - a| < \varepsilon \quad \text{für alle } n \geq n_0. \quad (3.31)$$

Offensichtlich ist jeder Grenzwert a einer Zahlenfolge $\{a_n\}, n \in \mathbb{N}_0$, ein Häufungspunkt. Die Umkehrung gilt nicht, wie das folgende Beispiel zeigt:

$$a_n = \begin{cases} \frac{1}{n} & \text{für } n \geq 1 \text{ und } n \text{ gerade} \\ 1 & \text{für } n = 0 \\ 2 - \frac{1}{n} & \text{für } n \geq 1 \text{ und } n \text{ ungerade} \end{cases}. \quad (3.32)$$

Diese Zahlenfolge besitzt die beiden Häufungspunkte $h_1 = 0$ und $h_2 = 2$, aber keinen Grenzwert. Besitzt eine Zahlenfolge $\{a_n\}, n \in \mathbb{N}_0$, einen Grenzwert a , so schreibt man dafür

$$\lim_{n \rightarrow \infty} a_n = a \quad \text{oder} \quad a_n \rightarrow a \quad \text{für } n \rightarrow \infty. \quad (3.33)$$

und man sagt, die Zahlenfolge $\{a_n\}, n \in \mathbb{N}_0$, konvergiert gegen a . Anschaulich bedeutet die Existenz des Grenzwertes a einer Zahlenfolge $\{a_n\}, n \in \mathbb{N}_0$, dass zu jedem $\varepsilon > 0$ nur endlich viele Folgenglieder außerhalb des Intervalls $(a - \varepsilon, a + \varepsilon)$ liegen dürfen. Eine gegen Null konvergierende (man sagt auch konvergente) Zahlenfolge heißt Nullfolge. Nicht konvergente Zahlenfolgen heißen divergent. Konvergiert eine Zahlenfolge $\{a_n\}, n \in \mathbb{N}_0$, gegen a , so ist die Zahlenfolge $\{a_n - a\}, n \in \mathbb{N}_0$, eine Nullfolge. Eine Zahlenfolge $\{a_n\}, n \in \mathbb{N}_0$, heißt „bestimmt divergent gegen ∞ (bzw. $-\infty$)“, falls es zu jedem $K \in \mathbb{R}$ ein $n_0 \in \mathbb{N}$ gibt mit:

$$a_n > K \quad (\text{bzw. } a_n < K) \quad \text{für alle } n \geq n_0. \quad (3.34)$$

Man schreibt dafür

$$\lim_{n \rightarrow \infty} a_n = \infty \quad (\text{bzw. } \lim_{n \rightarrow \infty} a_n = -\infty). \quad (3.35)$$

Beispiel(e) 3.8

- Die Zahlenfolge $a_n = (-1)^n, n \in \mathbb{N}_0$, ist divergent.
- Die Zahlenfolge $a_n = n, n \in \mathbb{N}_0$, ist bestimmt divergent gegen ∞
- Die Zahlenfolge $a_n = -n, n \in \mathbb{N}_0$, ist bestimmt divergent gegen $-\infty$

Offensichtlich sind Zahlenfolgen, die mindestens zwei Häufungspunkte haben, unbestimmt divergent.

Beispiel(e) 3.9

- Sei $a_n = \sum_{k=0}^n q^k$, $q \neq 0$ (geometrische Reihe), so ist $\{a_n\}$ konvergent, falls $|q| < 1$ mit dem Grenzwert $a = \frac{1}{1-q}$, unbestimmt divergent für $q \leq -1$ und bestimmt divergent gegen ∞ für $q \geq 1$. Ist $q \neq 1$, so folgt aus der Gleichung

$$(1-q) \cdot a_n = (1-q) \sum_{k=0}^n q^k = \sum_{k=0}^n q^k - \sum_{k=0}^n q^{k+1} = 1 - q^{n+1} \quad \text{für alle } n \in \mathbb{N}_0:$$

$$a_n = \frac{1 - q^{n+1}}{1 - q}.$$

Für $|q| < 1$ erhalten wir den Grenzwert $a = \frac{1}{1-q}$, denn für jedes $\varepsilon > 0$ gilt:

$$\begin{aligned} |a_n - a| < \varepsilon &\iff \left| \frac{1 - q^{n+1}}{1 - q} - \frac{1}{1 - q} \right| < \varepsilon \\ &\iff \left| \frac{q^{n+1}}{1 - q} \right| < \varepsilon \\ &\iff |q|^{n+1} < (1 - q)\varepsilon \end{aligned}$$

Da $|q| < 1$, ist $|q|^{n+1} < (1 - q)\varepsilon$ ab einem festen n_0 immer erfüllt.

Für $q > 1$ ist $\frac{1 - q^{n+1}}{1 - q} = \frac{q^{n+1} - 1}{q - 1} \rightarrow \infty$. Für $q = 1$ ist $a_n = n + 1$. Somit ist für $q \geq 1$ die Zahlenfolge $\{a_n\}$ bestimmt divergent gegen ∞ . Für $q \leq -1$ gilt

$$a_n = \frac{1 - q^{n+1}}{1 - q} = \begin{cases} \frac{1 - |q|^{n+1}}{1 - q} \leq 0 & \text{für } n \text{ ungerade} \\ \frac{1 + |q|^{n+1}}{1 - q} \geq 1 & \text{für } n \text{ gerade.} \end{cases}$$

Daher ist $\{a_n\}$ für $q \leq -1$ unbestimmt divergent.

- Sei $a_n = \frac{n+2}{n+1}$, $n \in \mathbb{N}_0$, so konvergiert $\{a_n\}$ gegen $a = 1$, denn für jedes $\varepsilon > 0$ gilt:

$$\begin{aligned} |a_n - a| < \varepsilon &\iff \left| \frac{n+2}{n+1} - 1 \right| < \varepsilon \\ &\iff \left| \frac{n+1}{n+1} + \frac{1}{n+1} - 1 \right| < \varepsilon \\ &\iff \frac{1}{n+1} < \varepsilon \\ &\iff n > \frac{1}{\varepsilon} - 1 \end{aligned}$$

Im Folgenden betrachten wir Rechenregeln für Grenzwerte und Konvergenzkriterien. Aus gegebenen Zahlenfolgen $\{a_n\}, n \in \mathbb{N}_0$, und $\{b_n\}, n \in \mathbb{N}_0$, werden durch Addition, Subtraktion, Multiplikation und Division neue Zahlenfolgen $c_n := a_n + b_n$, $d_n := a_n - b_n$, $e_n := a_n \cdot b_n$, $f_n := \frac{a_n}{b_n}$ (falls $b_n \neq 0$ für alle $n \in \mathbb{N}_0$) gewonnen. Ferner können neue Zahlenfolgen dadurch konstruiert werden, dass mit einer gegebenen Zahlenfolge $\{a_n\}, n \in \mathbb{N}_0$, und Indizes $n_0 < n_1 < n_2 < n_3 < \dots$ die Zahlenfolge $\{a_{n_k}\}, k \in \mathbb{N}_0$, betrachtet wird. Man nennt $\{a_{n_k}\}, k \in \mathbb{N}_0$, Teilfolge von $\{a_n\}, n \in \mathbb{N}_0$.

Satz 3.10 (Grenzwertregeln)

Sind $\{a_n\}, n \in \mathbb{N}_0$, und $\{b_n\}, n \in \mathbb{N}_0$, konvergente Folgen mit $\lim_{n \rightarrow \infty} a_n = a$ und $\lim_{n \rightarrow \infty} b_n = b$, dann gilt:

- (a) $\lim_{n \rightarrow \infty} (a_n \pm b_n) = a \pm b$,
- (b) $\lim_{n \rightarrow \infty} (a_n \cdot b_n) = ab$, insbesondere $\lim_{n \rightarrow \infty} (c \cdot a_n) = c \cdot a$ für alle $c \in \mathbb{R}$,
- (c) ist $a \neq 0$, dann gibt es ein $n_1 \in \mathbb{N}$ mit $a_n \neq 0$ für alle $n > n_1$ und für die Zahlenfolgen $\{a_n\}, n \geq n_1$, und $\{b_n\}, n \geq n_1$, gilt:

$$\lim_{n \rightarrow \infty} \frac{b_n}{a_n} = \frac{b}{a},$$

- (d) $\lim_{n \rightarrow \infty} |a_n| = |a|$,
- (e) $\lim_{n \rightarrow \infty} \sqrt{a_n} = \sqrt{a}$ falls alle $a_n \geq 0$.

Beispiel(e) 3.11

$$\lim_{n \rightarrow \infty} \sqrt{\frac{6n^4 + 3n^2 + 2}{7n^4 + 12n^3 + 6}} = \lim_{n \rightarrow \infty} \sqrt{\frac{6 + \frac{3}{n^2} + \frac{2}{n^4}}{7 + \frac{12}{n} + \frac{6}{n^4}}} = \sqrt{\frac{6}{7}} \tag{3.36}$$

Aus der Existenz von $\lim_{n \rightarrow \infty} (a_n \pm b_n)$ (bzw. $\lim_{n \rightarrow \infty} |a_n|$) folgt im allgemeinen nicht die Konvergenz der Folgen $\{a_n\}$ oder $\{b_n\}$ (bzw. $\{a_n\}$). Die Limesbildung erhält schwache Ungleichungen, aber keine strikten Ungleichungen: Seien $\{a_n\}, n \in \mathbb{N}_0$, und $\{b_n\}, n \in \mathbb{N}_0$ konvergente Zahlenfolgen mit $a_n \leq b_n$ für alle $n \geq n_1 \in \mathbb{N}$, so gilt $\lim_{n \rightarrow \infty} a_n \leq \lim_{n \rightarrow \infty} b_n$. Für strikte Ungleichungen gilt dies nicht, wie das Beispiel $a_n = 0, b_n = \frac{1}{1+n}, n \in \mathbb{N}_0$, zeigt, denn $a_n < b_n$ für alle $n \in \mathbb{N}_0$, aber $\lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} b_n$.

Da Zahlenfolgen spezielle Funktionen sind, heißt eine Zahlenfolge $\{a_n\}, n \in \mathbb{N}_0$, genau dann monoton wachsend (bzw. monoton fallend), falls $a_{n+1} \geq a_n$ (bzw. $a_{n+1} \leq a_n$) für alle $n \in \mathbb{N}_0$.

Satz 3.12 (Monotoniekriterium)

Jede monoton wachsende oder monoton fallende beschränkte Zahlenfolge ist konvergent.

Beispiel(e) 3.13

- Die Zahlenfolge $a_n := \sum_{k=0}^n \frac{1}{k!}$, $n \geq 0$, ist monoton wachsend. Da

$$0 < \frac{1}{k!} = \frac{1}{1 \cdot 2 \cdot 3 \cdot \dots \cdot k} \leq \frac{1}{1 \cdot 2 \cdot 2 \cdot \dots \cdot 2} = \frac{1}{2^{k-1}}, \quad (3.37)$$

ergibt sich für $n \geq 2$ die Abschätzung

$$2 \leq 1 + \frac{1}{1!} + \frac{1}{2!} + \dots + \frac{1}{n!} \leq 1 + 1 + \frac{1}{2} + \dots + \frac{1}{2^{n-1}} \leq 1 + \lim_{n \rightarrow \infty} \sum_{k=0}^n \left(\frac{1}{2}\right)^k = 3. \quad (3.38)$$

Somit ist $\{a_n\}, n \in \mathbb{N}_0$, auch beschränkt und damit konvergent.

Durch den Grenzwert dieser Folge ist die Eulersche Zahl e definiert:

$$e := \sum_{k=0}^{\infty} \frac{1}{k!} := \lim_{n \rightarrow \infty} \sum_{k=0}^n \frac{1}{k!}, \quad e = 2.7182818284 \dots \quad (3.39)$$

Die Eulersche Zahl e kann auch durch den Grenzwert der Folge

$$c_n := \left(1 + \frac{1}{n}\right)^n, \quad n \geq 1, \quad (3.40)$$

dargestellt werden.

- Die rekursiv definierte Zahlenfolge $a_0 := 1$, $a_{n+1} = \frac{6(1+a_n)}{7+a_n}$ ist beschränkt ($0 < a_n < 2$) und monoton wachsend (vollständige Induktion). Der noch unbekannte Grenzwert sei a . Aus $a_n \rightarrow a$ und $a_{n+1} \rightarrow a$ und den Rechenregeln für Grenzwerte folgt:

$$a = \frac{6(1+a)}{7+a} \quad (3.41)$$

d.h., $a = 2$ oder $a = -3$. Da $a_n > 0$ für alle $n \in \mathbb{N}_0$, folgt $a = 2$.

Wie die Zahl e kann man nun weitere Zahlen über den Grenzwert von Zahlenfolgen definieren. Dazu betrachten wir für jedes $x \in \mathbb{R}$ die Zahl e^x definiert durch

$$e^x := \lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n \quad (3.42)$$

Die Existenz dieses Grenzwertes für jedes x kann durch geeignete Fallunterscheidungen gezeigt werden. Der Zusammenhang zwischen e^x und der x -ten Potenz von e wird später geklärt.

Die Funktion $\exp : \mathbb{R} \rightarrow \mathbb{R}$, $x \mapsto e^x$ ($:= \lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n$) heißt Exponentialfunktion.

Nach dem Grenzwert von Zahlenfolgen betrachten wir im Folgenden den Grenzwert reellwertiger Funktionen:

Definition 3.14 ((rechtsseitiger, linksseitiger) Grenzwert von Funktionen)

Sei I ein Intervall und $a \in I$. Sei ferner $f : I \setminus \{a\} \rightarrow \mathbb{R}$ oder $f : I \rightarrow \mathbb{R}$ eine gegebene Funktion, so hat f für x gegen a den rechtsseitigen Grenzwert (bzw. linksseitigen Grenzwert) c , in Zeichen

$$\lim_{x \rightarrow a^+} f(x) = c \quad (3.43)$$

bzw.

$$\lim_{x \rightarrow a^-} f(x) = c, \quad (3.44)$$

falls für jede Zahlenfolge $\{x_n\}, n \in \mathbb{N}_0$, aus I mit $x_n \rightarrow a$ und $x_n > a$ für alle $n \in \mathbb{N}_0$ (bzw. $x_n \rightarrow a$ und $x_n < a$ für alle $n \in \mathbb{N}_0$) die Zahlenfolge $\{f(x_n)\}, n \in \mathbb{N}_0$, gegen c konvergiert. Die Funktion f hat für x gegen a den Grenzwert c , in Zeichen $\lim_{x \rightarrow a} f(x) = c$, wenn

$$\lim_{x \rightarrow a^+} f(x) = \lim_{x \rightarrow a^-} f(x) = c. \quad (3.45)$$

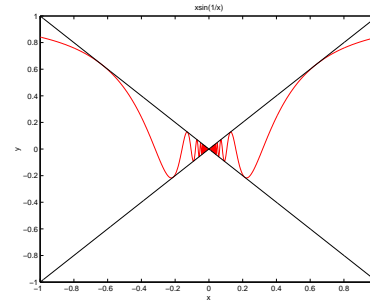
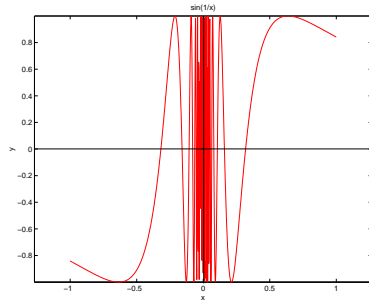
Nur linksseitige bzw. rechtsseitige Grenzwerte lassen sich definieren, wenn man in der Definition $a = +\infty$ bzw. $a = -\infty$ mit entsprechender Funktion f zuläßt. In diesen Fällen schreibt man

$$\lim_{x \rightarrow \infty} f(x) = c \text{ bzw. } \lim_{x \rightarrow -\infty} f(x) = c.$$

Der Begriff „bestimmt divergent“ wird gemäß obiger Definition für Funktionen anstelle von Zahlenfolgen entsprechend übernommen.

Beispiel(e) 3.15

- Die Funktion $f : \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R}, x \mapsto \sin\left(\frac{1}{x}\right)$ ist für alle $x \in \mathbb{R} \setminus \{0\}$ erklärt. Sie hat weder einen linksseitigen noch einen rechtsseitigen Grenzwert für $x \rightarrow 0$. Dagegen besitzt die Funktion $g : \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R}, x \mapsto x \cdot \sin\left(\frac{1}{x}\right)$ den Grenzwert 0 für $x \rightarrow 0$.



$$x \mapsto \sin\left(\frac{1}{x}\right)$$

$$x \mapsto x \sin\left(\frac{1}{x}\right)$$

- $\lim_{x \rightarrow 0^+} \frac{1}{x} = \infty, \lim_{x \rightarrow 0^-} \frac{1}{x} = -\infty$
- $f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto \begin{cases} x^2 & x > 0 \\ 1 & x = 0 \\ x^2 - 1 & x < 0 \end{cases}$
 $\lim_{x \rightarrow 0^+} f(x) = 0, \lim_{x \rightarrow 0^-} f(x) = -1, f(0) = 1$

Grenzwertregeln für Zahlenfolgen lassen sich auch auf Funktionen übertragen:

Satz 3.16 (Grenzwertregeln für Funktionen)

Seien f, g Funktionen gemäß Definition 3.14. Aus $\lim_{x \rightarrow a} f(x) = c$ und $\lim_{x \rightarrow a} g(x) = d, c, d \in \mathbb{R}$, folgt:

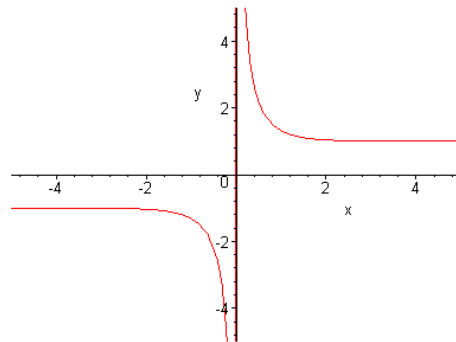
- (i) $\lim_{x \rightarrow a} (f(x) \pm g(x)) = c \pm d,$
- (ii) $\lim_{x \rightarrow a} (f(x) \cdot g(x)) = c \cdot d,$
- (iii) $\lim_{x \rightarrow a} \frac{f(x)}{g(x)} = \frac{c}{d},$ falls $d \neq 0.$

Diese Regeln gelten auch für $a = \pm\infty$, aber nur für $c, d \in \mathbb{R}$.

Beispiel(e) 3.17

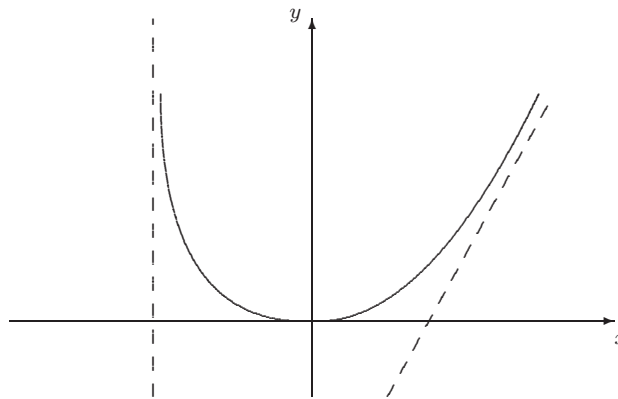
- $\lim_{x \rightarrow 2} \frac{x^3 + 3x + 5}{x^2 - 2x + 1} = \frac{8 + 6 + 5}{4 - 4 + 1} = 19$
- $\lim_{x \rightarrow 1} \frac{x^n - 1}{x - 1} = \lim_{x \rightarrow 1} (x^{n-1} + x^{n-2} + \dots + x + 1) = n$
- $\lim_{x \rightarrow 0} \frac{\sqrt{x+1} - 1}{x} = \lim_{x \rightarrow 0} \frac{(\sqrt{x+1} - 1)(\sqrt{x+1} + 1)}{x(\sqrt{x+1} + 1)} = \lim_{x \rightarrow 0} \frac{x + 1 - 1}{x(\sqrt{x+1} + 1)} = \frac{1}{2}$

Man nennt die Gerade $x = a$ eine vertikale Asymptote der Kurve $y = f(x)$ bei gegebenem f , wenn $\lim_{x \rightarrow a^-} f(x) = \pm\infty$ oder $\lim_{x \rightarrow a^+} f(x) = \pm\infty$. Die Gerade $y = c$ heißt horizontale Asymptote der Kurve $y = f(x)$ bei gegebenem f , falls $\lim_{x \rightarrow \infty} f(x) = c$ oder $\lim_{x \rightarrow -\infty} f(x) = c$.



(3.46)

Als schräge Asymptote der Kurve $y = f(x)$ bezeichnet man die Gerade $y = px + q$, $p \neq 0$, falls $(f(x) - p \cdot x - q) \rightarrow 0$ für $x \rightarrow \infty$ oder $x \rightarrow -\infty$.



(3.47)

Nun zeichnen wir Funktionen aus, bei denen an einer Stelle der rechtsseitige Grenzwert, der linksseitige Grenzwert und der Funktionswert übereinstimmen.

Definition 3.18 (Stetigkeit)

Sei I ein Intervall und $f : I \rightarrow \mathbb{R}$ eine gegebene Funktion, so heißt f in $x_0 \in I$ genau dann stetig, wenn

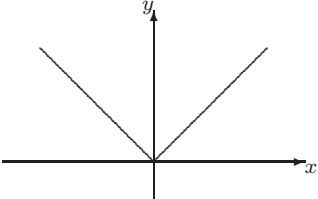
$$\lim_{x \rightarrow x_0} f(x) = f(x_0). \tag{3.48}$$

Ist x_0 Randpunkt von I , so ist $\lim_{x \rightarrow x_0}$ als einseitiger Grenzwert ($\lim_{x \rightarrow x_0^+} f(x)$ oder $\lim_{x \rightarrow x_0^-} f(x)$) zu verstehen. Die Funktion f heißt auf I stetig, falls sie in allen $x_0 \in I$ stetig ist.

Die Stetigkeit von $f : I \rightarrow \mathbb{R}$ auf I bedeutet anschaulich, dass der Graph $y = f(x)$ über I eine zusammenhängende Linie (ohne Lücken und Sprünge, aber durchaus mit Spitzen) darstellt.

Beispiel(e) 3.19

- $f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto |x|$ ist stetig



(3.49)

- $f : \mathbb{R}_0^+ \setminus \{0\} \rightarrow \mathbb{R}, x \mapsto \frac{1}{x}$ ist stetig

Aufgrund der Grenzwertregeln sind für stetige Funktionen $f, g : I \rightarrow \mathbb{R}$ auch $f \pm g, f \cdot g$ stetig. Ferner ist $\frac{f}{g}$ in allen $x \in I$ mit $g(x) \neq 0$ stetig. Seien I, J Intervalle und $f : I \rightarrow \mathbb{R}$ sowie $g : J \rightarrow \mathbb{R}$ mit $g(J) \subseteq I$ stetig, so ist auch die Funktion $f \circ g : J \rightarrow \mathbb{R}, x \mapsto f(g(x))$ stetig. Offensichtlich sind alle Polynome auf \mathbb{R} und alle rationalen Funktionen auf den Teilintervallen des entsprechenden Definitionsbereichs stetig.

Der folgende Satz fasst wichtige Eigenschaften zusammen:

Satz 3.20 (Eigenschaften auf einem abgeschlossenen Intervall stetiger Funktionen)

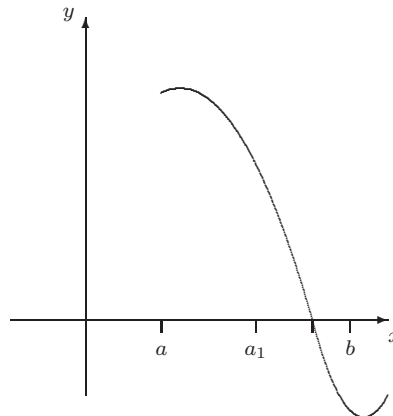
Sei $f : [a, b] \rightarrow \mathbb{R}$ eine auf $[a, b]$ stetige Funktion, so gilt:

- a) Schrankensatz: Es gibt ein $K \in \mathbb{R}$ mit $|f(x)| < K$ für alle $x \in [a, b]$ (man sagt, f ist auf $[a, b]$ beschränkt).
- b) Satz von Minimum und Maximum: Es gibt Werte $x_0, x_1 \in [a, b]$ mit

$$f(x_0) \leq f(x) \leq f(x_1) \tag{3.50}$$
 für alle $x \in [a, b]$.
- c) Nullstellenexistenz: Ist $f(a) \cdot f(b) < 0$, so gibt es ein $\bar{x} \in (a, b)$ mit $f(\bar{x}) = 0$.

Ist $f(a) \cdot f(b) < 0$, so kann eine Stelle \bar{x} mit $f(\bar{x}) = 0$ durch das Bisektionsverfahren berechnet werden:

Wähle $a_1 = \frac{a+b}{2}$. Ist $f(a_1) = 0$, so ist man fertig. Ist $f(a_1) \cdot f(b) < 0$, so ersetze a durch a_1 und wiederhole die Vorgehensweise mit dem Intervall $[a_1, b]$. Ist $f(a) \cdot f(a_1) < 0$, so ersetze b durch a_1 und wiederhole die Vorgehensweise mit $[a, a_1]$.



(3.51)

3.2 Differentiation

Seien $I \subseteq \mathbb{R}$ ein offenes Intervall und $f : I \rightarrow \mathbb{R}$ eine Funktion, so heißt f im Punkt $x_0 \in I$ differenzierbar, wenn der Differenzenquotient $\frac{f(x)-f(x_0)}{x-x_0}$ für $x \rightarrow x_0$ einen endlichen Grenzwert besitzt.

Dieser Grenzwert wird mit $f'(x_0)$ bezeichnet. Mit $h := x - x_0$ gilt also:

$$f'(x_0) = \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h} \tag{3.52}$$

Ist f in jedem Punkt $x_0 \in I$ differenzierbar, so sagt man, f ist auf I differenzierbar. In diesem Fall ist

$$f' : I \rightarrow \mathbb{R}, \quad x \mapsto f'(x) \tag{3.53}$$

eine Funktion, die man die Ableitung von f nennt. Den Übergang von f zu f' nennt man differenzieren oder ableiten. Für $f'(x)$ schreibt man auch $\frac{df}{dx}(x)$ oder $\frac{d}{dx}f(x)$.

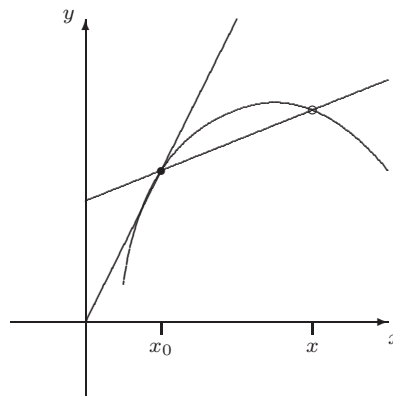
Beispiel(e) 3.21

- $f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto c, f' : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto 0$
- $f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto ax + b, f' : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto a$, denn

$$\lim_{x \rightarrow x_0} \frac{ax + b - (ax_0 + b)}{x - x_0} = \lim_{x \rightarrow x_0} \frac{a(x - x_0)}{x - x_0} = a$$
- $f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto x^n, n \in \mathbb{N}, f' : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto n \cdot x^{n-1}$
- $f : \mathbb{R}_0^+ \setminus \{0\} \rightarrow \mathbb{R}, x \mapsto \sqrt{x}, f' : \mathbb{R}_0^+ \setminus \{0\} \rightarrow \mathbb{R}, x \mapsto \frac{1}{2\sqrt{x}}$, denn $\lim_{x \rightarrow x_0} \frac{\sqrt{x} - \sqrt{x_0}}{x - x_0} =$

$$\lim_{x \rightarrow x_0} \frac{\sqrt{x} - \sqrt{x_0}}{(\sqrt{x} - \sqrt{x_0})(\sqrt{x} + \sqrt{x_0})} = \lim_{x \rightarrow x_0} \frac{1}{(\sqrt{x} + \sqrt{x_0})} = \frac{1}{2\sqrt{x_0}}.$$

Betrachtet man bezüglich kartesischer (x, y) -Koordinaten die Kurve $g = f(x)$, so ist $\frac{f(x)-f(x_0)}{x-x_0}$ die Steigung der Geraden durch die Punkte $(x_0, f(x_0))$ und $(x, f(x))$



(3.54)

Der Grenzwert $f'(x_0) = \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0}$ gibt dann die Steigung der Kurventangente in $(x_0, f(x_0))$ an. Ist also f in x_0 differenzierbar, so ist

$$y = f'(x_0)(x - x_0) + f(x_0) \tag{3.55}$$

die Tangente an den Graphen $y = f(x)$ von f in $(x_0, f(x_0))$. Da die Wurzelfunktion in $x_0 = 0$ eine senkrechte Tangente hat, existiert die Ableitung in $x_0 = 0$ nicht. Will man eine gegebene differenzierbare Funktion $f : I \rightarrow \mathbb{R}$ in der Nähe eines Punktes $x_0 \in I$ durch eine Gerade $g : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto m \cdot (x - x_0) + f(x_0)$ (damit ist $g(x_0) = f(x_0)$) approximieren, so ist der Parameter m zu bestimmen. Fordert man, dass der relative Fehler $\frac{f(x) - g(x)}{x - x_0}$ für $x \rightarrow x_0$ gegen Null konvergiert, so ergibt sich

$$0 = \lim_{x \rightarrow x_0} \frac{f(x) - m(x - x_0) - f(x_0)}{x - x_0} = \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} - m = f'(x_0) - m. \tag{3.56}$$

Somit ist $m = f'(x_0)$ zu wählen. Es gilt also in der Nähe von x_0 :

$$f(x) \approx f'(x_0)(x - x_0) + f(x_0) \tag{3.57}$$

Der folgende Satz stellt eine Beziehung zwischen Differenzierbarkeit und Stetigkeit her:

Satz 3.22 (Differenzierbarkeit und Stetigkeit)
 Seien $I \subseteq \mathbb{R}$ ein offenes Intervall und $f : I \rightarrow \mathbb{R}$ in $x_0 \in I$ differenzierbar, dann ist f in x_0 auch stetig.

Die Umkehrung dieses Satzes gilt nicht, wie das Beispiel $f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto |x|$ an der Stelle $x = 0$ zeigt.

Mit den folgenden Regeln gelingt es, die Ableitung zusammengesetzter Funktionen aus den Ableitungen ihrer einzelnen „Bausteine“ zu ermitteln.

Satz 3.23 (Ableitung zusammengesetzter Funktionen)

Seien I ein offenes Intervall und $f, g : I \rightarrow \mathbb{R}$ Funktionen, die in $x \in I$ differenzierbar sind, dann gilt für alle $x \in I$:

$$(f(x) + g(x))' = f'(x) + g'(x) \quad (3.58)$$

$$(c \cdot f(x))' = c \cdot f'(x) \quad \text{für alle } c \in \mathbb{R} \quad (3.59)$$

$$(f(x) \cdot g(x))' = f'(x) \cdot g(x) + f(x) \cdot g'(x) \quad (\text{Produktregel}) \quad (3.60)$$

$$\left(\frac{f(x)}{g(x)}\right)' = \frac{f'(x) \cdot g(x) - g'(x) \cdot f(x)}{g(x)^2} \quad \text{falls } g(x) \neq 0 \quad (\text{Quotientenregel}) \quad (3.61)$$

speziell:

$$\left(\frac{1}{g(x)}\right)' = -\frac{g'(x)}{g(x)^2}. \quad (3.62)$$

Mit der Produktregel läßt sich nun die bereits erwähnte Formel $(x^n)' = nx^{n-1}$ für alle $n \in \mathbb{N}$, $x \in \mathbb{R}$, einfach mit vollständiger Induktion beweisen.

Jede rationale Funktion $f : \mathbb{D} \rightarrow \mathbb{R}$, $x \mapsto \frac{p}{q}$ kann nun mit der Differentiation für Polynome und der Quotientenregel differenziert werden. Insbesondere gilt für alle $x \in \mathbb{R} \setminus \{0\}$:

$$\left(\frac{1}{x^n}\right)' = -\frac{n}{x^{n+1}} = -n \cdot x^{-n-1}, \quad n \in \mathbb{N}. \quad (3.63)$$

Somit gilt die Formel $(x^n)' = nx^{n-1}$ auch für $n \in \mathbb{Z}$, $x \in \mathbb{R} \setminus \{0\}$.

Für die Sinus- und Cosinusfunktion sowie für die Tangens- und Cotangensfunktion

$$\tan : \mathbb{R} \setminus \{x \in \mathbb{R}; x = (2k+1)\frac{\pi}{2}, k \in \mathbb{Z}\} \rightarrow \mathbb{R}, x \mapsto \frac{\sin(x)}{\cos(x)} \quad (3.64)$$

$$\cot : \mathbb{R} \setminus \{x \in \mathbb{R}; x = k \cdot \pi, k \in \mathbb{Z}\} \rightarrow \mathbb{R}, x \mapsto \frac{\cos(x)}{\sin(x)} \quad (3.65)$$

gilt für alle x aus dem entsprechenden Definitionsbereich:

$$(\sin(x))' = \cos(x) \quad (3.66)$$

$$(\cos(x))' = -\sin(x) \quad (3.67)$$

$$(\tan(x))' = \frac{1}{(\cos(x))^2} \quad (3.68)$$

$$(\cot(x))' = -\frac{1}{(\sin(x))^2} \quad (3.69)$$

Für den Beweis verwendet man die folgenden Additionstheoreme:

$$\cos(x+y) = \cos(x)\cos(y) - \sin(x)\sin(y) \quad (3.70)$$

$$\sin(x+y) = \sin(x)\cos(y) + \cos(x)\sin(y) \quad (3.71)$$

Damit lassen sich auch die folgenden wichtigen Gleichungen herleiten:

$$\cos(x - y) = \cos(x)\cos(y) + \sin(x)\sin(y) \quad (3.72)$$

$$\sin(x - y) = \sin(x)\cos(y) - \cos(x)\sin(y) \quad (3.73)$$

$$\sin(x) + \sin(y) = 2\sin\left(\frac{x+y}{2}\right) \cdot \cos\left(\frac{x-y}{2}\right) \quad (3.74)$$

$$\sin(x) - \sin(y) = 2\sin\left(\frac{x-y}{2}\right) \cdot \cos\left(\frac{x+y}{2}\right) \quad (3.75)$$

$$\cos(x) + \cos(y) = 2\cos\left(\frac{x+y}{2}\right) \cdot \cos\left(\frac{x-y}{2}\right) \quad (3.76)$$

$$\cos(2x) = \cos^2(x) - \sin^2(x) = 2\cos^2(x) - 1 \quad (3.77)$$

$$\sin(2x) = 2\sin(x) \cdot \cos(x) \quad (3.78)$$

$$1 + \cos(x) = 2\cos^2\left(\frac{x}{2}\right) \quad (3.79)$$

$$1 - \cos(x) = 2\sin^2\left(\frac{x}{2}\right). \quad (3.80)$$

Die Komposition $x \mapsto f(g(x))$ zweier differenzierbarer Funktionen definiert auf einem offenen Intervall I ist ebenfalls differenzierbar und es gilt für alle $x \in I$:

$$\frac{d}{dx}f(g(x)) = f'(g(x)) \cdot g'(x), \quad (3.81)$$

denn mit $g(x) \neq g(x_0)$ gilt:

$$\begin{aligned} \lim_{x \rightarrow x_0} \frac{f(g(x)) - f(g(x_0))}{x - x_0} &= \lim_{x \rightarrow x_0} \frac{(f(g(x)) - f(g(x_0))) \cdot (g(x) - g(x_0))}{(g(x) - g(x_0)) \cdot (x - x_0)} = \\ &= \lim_{x \rightarrow g(x_0)} \frac{(f(x) - f(g(x_0)))}{x - g(x_0)} \cdot \lim_{x \rightarrow x_0} \frac{g(x) - g(x_0)}{x - x_0} \quad (3.82) \\ &= f'(g(x_0)) \cdot g'(x_0). \end{aligned}$$

Beispiel(e) 3.24

- $h : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto (x^4 + 6x + 5)^3,$
 $f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto x^3, g : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto x^4 + 6x + 5, h = f \circ g.$
 $h' : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto 3(x^4 + 6x + 5)^2 \cdot (4x^3 + 6)$
- $k : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto \sin^2(x^4 + 2x),$
 $f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto x^2, g : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto \sin(x), h : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto x^4 + 2x,$
 $k = f \circ g \circ h, \text{ also } k : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto f(g(h(x))).$
 $k' : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto 2(\sin(x^4 + 2x)) \cdot \cos(x^4 + 2x) \cdot (4x^3 + 2).$

Die Ableitung der Ableitung bezeichnen wir, falls sie existiert, mit f'' und für $\frac{d}{dx}\left(\frac{d}{dx}f(x)\right)$ schreiben wir $\frac{d^2}{dx^2}f(x)$.

Allgemein definieren wir

$$f^{(0)} = f, \quad f^{(1)} = f', \quad f^{(n)} = \left(f^{(n-1)}\right)', \quad n \in \mathbb{N}. \quad (3.83)$$

Man sagt, f ist n -mal (stetig) differenzierbar, wenn die n -te Ableitung von f existiert (existiert und stetig ist).

Die Funktion $f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto x \cdot |x|$ ist nur einmal stetig differenzierbar.

Eine insbesondere bezüglich der Differentiation interessante Funktion ist die bereits eingeführte

Exponentialfunktion $\exp : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto \lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n$, denn es gilt:

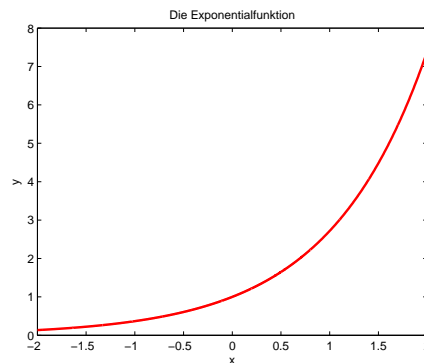
$$\frac{d}{dx} \exp(x) = \exp(x), \quad x \in \mathbb{R}. \quad (3.84)$$

Weitere Eigenschaften der \exp -Funktion, die unter Verwendung der unterschiedlichen Darstellungen der \exp -Funktion bewiesen werden:

$$\begin{aligned} \exp(1) &= e \\ \exp(x) &> 0 && \text{für alle } x \in \mathbb{R} \\ \exp(x + y) &= \exp(x) \cdot \exp(y) && \text{für alle } x, y \in \mathbb{R} \\ \exp(-x) &= \frac{1}{\exp(x)} && \text{für alle } x \in \mathbb{R} \\ \exp(sx) &= (\exp(x))^s && \text{für alle } x \in \mathbb{R}, s \in \mathbb{Q}. \end{aligned}$$

Somit gilt

$$\exp(s) = \exp(s \cdot 1) = \exp(1)^s = e^s \quad \text{für alle } s \in \mathbb{Q}. \quad (3.85)$$



(3.86)

Aus den speziellen Eigenschaften der Exponentialfunktion folgt, dass jede auf einem offenen Intervall $I \subseteq \mathbb{R}$ differenzierbare Funktion f mit $f' : I \rightarrow \mathbb{R}, x \mapsto a \cdot f(x)$ von der Form $f : I \rightarrow \mathbb{R}, x \mapsto c \cdot \exp(ax)$ mit einer Konstanten $c \in \mathbb{R}$ ist.

Beispiel(e) 3.25 (Radioaktiver Zerfall)
 Aufgrund von Beobachtungen geht man davon aus, dass für $N(t)$ Atome eines zerfallenden Stoffes die Anzahl ΔN der Zerfälle in einem kleinen Zeitintervall Δt gegeben ist durch

$$\Delta N = -kN\Delta t. \quad (3.87)$$

Unter der Annahme, dass die Funktion N differenzierbar ist, ergibt sich hieraus

$$\frac{d}{dt}N(t) = -k \cdot N(t) \quad (3.88)$$

und somit das Zerfallsgesetz

$$N(t) = N(0) \exp(-kt). \quad (3.89)$$

Alle in der Natur beobachtbaren Wachstums- und Zerfallsprozesse werden mit Hilfe der Exponentialfunktion beschrieben.

Da die Exponentialfunktion eine streng monoton steigende Funktion mit der Wertemenge $\mathbb{R}_0^+ \setminus \{0\}$ ist, existiert eine Umkehrfunktion, die als natürlicher Logarithmus

$$\ln : \mathbb{R}_0^+ \setminus \{0\} \rightarrow \mathbb{R}, \quad x \mapsto \ln(x) \quad (3.90)$$

bezeichnet wird. Es gilt somit $\exp(\ln(x)) = x$ für alle $x \in \mathbb{R}_0^+ \setminus \{0\}$ und $\ln(\exp(x)) = x$ für alle $x \in \mathbb{R}$. Ist nun $f : I \rightarrow \mathbb{R}$ eine umkehrbare und differenzierbare Funktion, so ist die Umkehrfunktion $f^{-1} : f(I) \rightarrow \mathbb{R}$ in allen $x \in f(I)$ mit $f'(f^{-1}(x)) \neq 0$ differenzierbar und es gilt:

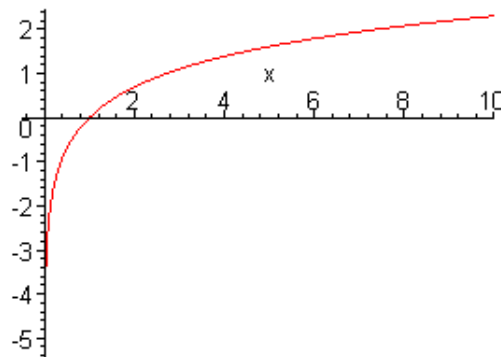
$$\frac{d}{dx} f^{-1}(x) = \frac{1}{f'(f^{-1}(x))}. \tag{3.91}$$

Für den natürlichen Logarithmus erhalten wir somit:

$$\ln'(x) = \frac{1}{\exp(\ln(x))} = \frac{1}{x} \quad \text{für alle } x > 0. \tag{3.92}$$

Ferner gilt:

$$\ln(x \cdot y) = \ln(x) + \ln(y), \quad \ln\left(\frac{x}{y}\right) = \ln(x) - \ln(y), \quad x, y > 0. \tag{3.93}$$



(3.94)

Da $\exp(sx) = (\exp(x))^s$ für alle $x \in \mathbb{R}$ und $s \in \mathbb{Q}$, definieren wir für $x = \ln(a)$, $a > 0$:

$$a^s := \exp(s \ln(a)) \quad \text{für alle } s \in \mathbb{R}. \tag{3.95}$$

(Exponentialfunktion zur Basis a)

Es gilt mit $a, b > 0$, $x, y \in \mathbb{R}$:

$$a^x \cdot a^y = a^{x+y}, \quad (ab)^x = a^x b^x, \quad (a^x)^y = a^{x \cdot y}, \quad \ln(a^x) = x \cdot \ln(a) \tag{3.96}$$

und

$$\frac{d}{dx} a^x = a^x \cdot \ln(a). \tag{3.97}$$

Die Umkehrfunktion zu $f : \mathbb{R} \rightarrow \mathbb{R}_0^+ \setminus \{0\}$, $x \mapsto a^x$ lautet für $a \neq 1$: $\log_a : \mathbb{R}_0^+ \setminus \{0\} \rightarrow \mathbb{R}$ und ist gegeben durch $x \mapsto \frac{\ln(x)}{\ln(a)}$ (Logarithmus zur Basis a : Wichtige Fälle: $a = 2$ (Informationstheorie) mit der Schreibweise ld für \log_2 und $a = 10$ mit der Schreibweise \log für \log_{10} (Nachrichtentechnik)).

Weitere für die Anwendungen wichtige Funktionen sind der sinus hyperbolicus:

$$\sinh : \mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto \frac{e^x - e^{-x}}{2} \tag{3.98}$$

mit der Umkehrfunktion area sinus hyperbolicus:

$$\text{arsinh} : \mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto \ln(x + \sqrt{x^2 + 1}), \tag{3.99}$$

cosinus hyperbolicus:

$$\cosh : \mathbb{R} \rightarrow [1, \infty), \quad x \mapsto \frac{e^x + e^{-x}}{2} \tag{3.100}$$

mit der Umkehrfunktion area cosinus hyperbolicus des zu $[0, \infty)$ gehörenden Zweiges:

$$\operatorname{arcosh} : [1, \infty) \rightarrow [0, \infty), \quad x \mapsto \ln(x + \sqrt{x^2 - 1}) \quad (3.101)$$

und der tangens hyperbolicus:

$$\operatorname{tanh} : \mathbb{R} \rightarrow (-1, 1), \quad x \mapsto \frac{\sinh(x)}{\cosh(x)} \quad (3.102)$$

mit der Umkehrfunktion area tangens hyperbolicus:

$$\operatorname{artanh} : (-1, 1) \rightarrow \mathbb{R}, \quad x \mapsto \frac{1}{2} \ln\left(\frac{1+x}{1-x}\right). \quad (3.103)$$

Die jeweiligen Ableitungen berechnen sich durch die Kettenregel. Es gilt zum Beispiel:

$$(\operatorname{artanh}(x))' = \left(\frac{1}{2} \ln\left(\frac{1+x}{1-x}\right)\right)' = \frac{1}{2} \left(\frac{1-x}{1+x}\right) \frac{(1-x) + (1+x)}{(1-x)^2} = \frac{1}{1-x^2} \quad (3.104)$$

$$(\sinh(x))' = \frac{e^x + e^{-x}}{2} = \cosh(x) \quad (3.105)$$

$$(\cosh(x))' = \frac{e^x - e^{-x}}{2} = \sinh(x) \quad (3.106)$$

Die trigonometrischen Funktionen \sin, \cos sind nicht über dem ganzen Definitionsbereich umkehrbar. Die Sinusfunktion ist über dem Intervall $-\frac{\pi}{2} \leq x \leq \frac{\pi}{2}$ streng monoton wachsend und daher dort umkehrbar mit der Umkehrfunktion arcus sinus:

$$\operatorname{arcsin} : [-1, 1] \rightarrow \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \quad (3.107)$$

Für die Ableitung erhalten wir:

$$(\operatorname{arcsin}(x))' = \frac{1}{\cos(\operatorname{arcsin}(x))} = \frac{1}{\sqrt{1 - \sin^2(\operatorname{arcsin}(x))}} = \frac{1}{\sqrt{1 - x^2}}, \quad |x| < 1 \quad (3.108)$$

Die Cosinusfunktion ist im Intervall $[0, \pi]$ streng monoton fallend und daher dort umkehrbar mit der Umkehrfunktion

$$\operatorname{arccos} : [-1, 1] \rightarrow [0, \pi]. \quad (3.109)$$

Für die Ableitung gilt mit $|x| < 1$:

$$(\operatorname{arccos}(x))' = -\frac{1}{\sin(\operatorname{arccos}(x))} = -\frac{1}{\sqrt{1 - \cos^2(\operatorname{arccos}(x))}} = -\frac{1}{\sqrt{1 - x^2}}. \quad (3.110)$$

Analog erhält man

$$\operatorname{arctan} : \mathbb{R} \rightarrow \left(-\frac{\pi}{2}, \frac{\pi}{2}\right) \quad (\operatorname{arctan}(x))' = \frac{1}{1+x^2}, \quad x \in \mathbb{R} \quad (3.111)$$

$$\operatorname{arccot} : \mathbb{R} \rightarrow (0, \pi) \quad (\operatorname{arccot}(x))' = -\frac{1}{1+x^2}, \quad x \in \mathbb{R}. \quad (3.112)$$

Für eine gegebene komplexe Zahl $z = x + iy$ definieren wir nun die Exponentialfunktion folgendermaßen:

$$\exp : \mathbb{C} \rightarrow \mathbb{C}, \quad z \mapsto e^z := e^x \cdot (\cos(y) + i \sin(y)). \quad (3.113)$$

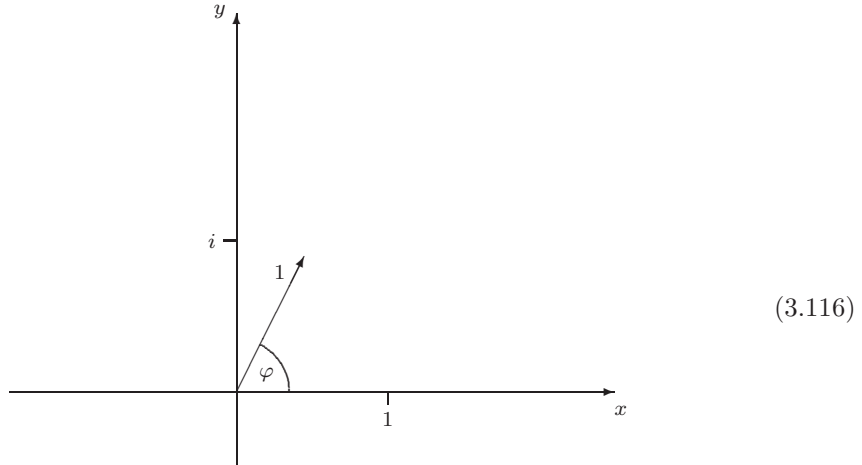
Speziell für $z = iy$ folgt somit:

$$e^{iy} = \cos(y) + i \sin(y), \quad |e^{iy}| = \cos^2(y) + \sin^2(y) = 1. \quad (3.114)$$

Jede komplexe Zahl mit $|z| = 1$ läßt sich also in der Form

$$z = \cos(\varphi) + i \sin(\varphi) \tag{3.115}$$

darstellen. Der Winkel φ kann dabei auch durch $\varphi + 2\pi k$, $k \in \mathbb{Z}$, ersetzt werden.

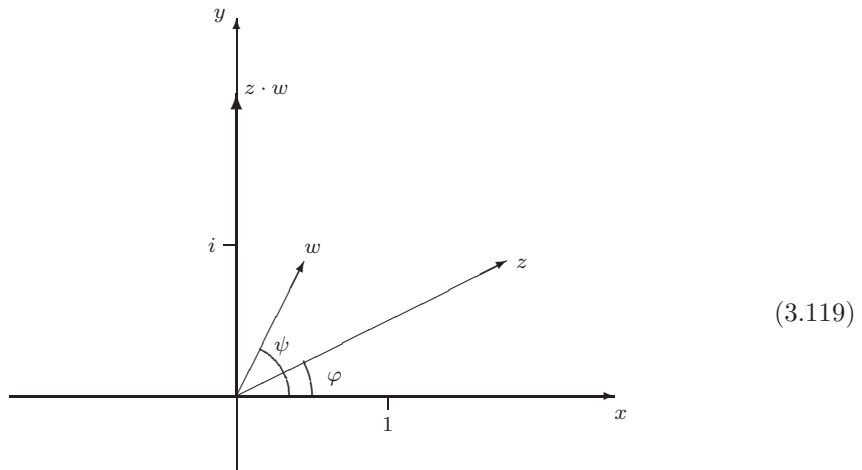


Somit erhalten wir für jede beliebige komplexe Zahl $z = x + iy$ die Darstellung

$$z = |z| \cdot (\cos(\varphi) + i \sin(\varphi)) = |z| \cdot e^{i\varphi}. \tag{3.117}$$

Den kartesischen Koordinaten x und y entsprechen dabei die polaren Koordinaten $|z|$ und φ . Während die Addition $z + w$ und die Subtraktion $z - w$ komplexer Zahlen in kartesischen Koordinaten $z = x + iy$, $w = u + iv$ zu naheliegenden Formeln führte, erhält man nun für die Multiplikation und für die Division komplexer Zahlen $z = |z| \cdot e^{i\varphi}$, $w = |w| \cdot e^{i\psi}$ ($|w| \neq 0$) einfache Formeln:

$$z \cdot w = |z| \cdot e^{i\varphi} \cdot |w| \cdot e^{i\psi} = \underbrace{|z| \cdot |w|}_{=|z \cdot w|} \cdot e^{i(\varphi+\psi)} \tag{3.118}$$



$$\frac{z}{w} = |z| \cdot e^{i\varphi} \cdot \frac{1}{|w|} \cdot e^{-i\psi} = \frac{|z|}{|w|} \cdot e^{i(\varphi-\psi)}, \quad w \neq 0. \tag{3.120}$$

Der Logarithmus naturalis (ln) einer komplexen Zahl $z = |z| \cdot e^{i\varphi} \neq 0$ ist definiert durch

$$\ln(z) := \ln(|z|) + i\varphi. \tag{3.121}$$

Da φ in der Darstellung $z = |z| \cdot e^{i\varphi}$ nicht eindeutig ist, fordert man $-\pi < \varphi \leq \pi$. Der entsprechende Winkel φ wird Hauptargument genannt und mit $\arg(z)$ bezeichnet. Ein wichtiger Satz im Zusammenhang mit komplexen Zahlen ist der nun folgende: Fundamentalsatz der Algebra: Seien $n \in \mathbb{N}$ und $a_0, a_1, \dots, a_n \in \mathbb{C}$ mit $a_n \neq 0$, so gibt es komplexe Zahlen z_1, z_2, \dots, z_n mit

$$a_n z^n + a_{n-1} z^{n-1} + \dots + a_2 z^2 + a_1 z + a_0 = a_n (z - z_1) \cdot \dots \cdot (z - z_n) \quad (3.122)$$

für alle $z \in \mathbb{C}$.

Beispiel(e) 3.26 (*n*-te Einheitswurzel)
 Die Gleichung

$$z^n = 1 \quad (3.123)$$

kann mit dem Ansatz $z = e^{i\varphi}$ aufgelöst werden:

$$z^n = 1 \iff n\varphi = k2\pi, \quad k = 0, 1, \dots, (n - 1). \quad (3.124)$$

Somit ist

$$z_{k+1} = e^{i2\pi \frac{k}{n}}, \quad k = 0, 1, \dots, (n - 1). \quad (3.125)$$

Den wichtigen Zusammenhang zwischen den trigonometrischen Funktionen \cos, \sin und der komplexen Exponentialfunktion zeigt die folgende Rechnung:

$$\begin{aligned} e^{i(\varphi+\psi)} &= \cos(\varphi + \psi) + i \sin(\varphi + \psi) = e^{i\varphi} e^{i\psi} = \\ &= (\cos(\varphi) + i \sin(\varphi))(\cos(\psi) + i \sin(\psi)) = \\ &= \cos(\varphi) \cos(\psi) - \sin(\varphi) \sin(\psi) + i(\sin(\varphi) \cos(\psi) + \sin(\psi) \cos(\varphi)). \end{aligned} \quad (3.126)$$

Somit reproduzieren sich die Additionstheoreme für die Funktionen \cos und \sin :

$$\cos(\varphi + \psi) = \cos(\varphi) \cos(\psi) - \sin(\varphi) \sin(\psi) \quad (3.127)$$

$$\sin(\varphi + \psi) = \sin(\varphi) \cos(\psi) + \sin(\psi) \cos(\varphi). \quad (3.128)$$

Zu den wichtigsten Anwendungen der Differentiation gehört die Berechnung von Maxima und Minima gegebener Funktionen. Eine auf $\mathbb{D} \subseteq \mathbb{R}$ definierte Funktion $f : \mathbb{D} \rightarrow \mathbb{R}$ hat in $a \in \mathbb{D}$ ein globales Maximum (Minimum), wenn $f(x) \leq f(a)$ für alle $x \in \mathbb{D}$ gilt ($f(a) \leq f(x)$ für alle $x \in \mathbb{D}$). In diesem Fall heißt a ein globaler Maximierer (Minimierer) von f . Eine Stelle $b \in \mathbb{D}$ heißt lokaler Maximierer (Minimierer) von f , falls es ein $\delta > 0$ gibt mit $f(x) \leq f(b)$ für alle $x \in \mathbb{D} \cap (b - \delta, b + \delta)$ ($f(b) \leq f(x)$ für alle $x \in \mathbb{D} \cap (b - \delta, b + \delta)$). Der Begriff „global“ bezieht sich also immer auf den ganzen Definitionsbereich, der Begriff „lokal“ nur auf die unmittelbare Umgebung von b . Ist x_0 ein Maximierer von f , dann ist x_0 ein Minimierer von $-f$ und umgekehrt. Der folgende Satz beschreibt einen Zusammenhang zwischen der Lage von Extremalstellen (Maximierer und Minimierer) und der Ableitung der gegebenen Funktion f .

Satz 3.27 (Differentiation und Extremalstellen)
 Sei $f : I \rightarrow \mathbb{R}$ eine auf dem offenen Intervall I differenzierbare Funktion, so gilt:
 Ist $x_0 \in I$ (lokale) Extremalstelle von f , so gilt: $f'(x_0) = 0$

Die Bedingung $f'(x_0) = 0$ (waagerechte Tangente des Graphen von f in $(x_0, f(x_0))$) ist zwar notwendig für eine Extremalstelle, aber nicht hinreichend, wie das Beispiel $f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto x^3$ mit $x_0 = 0$ zeigt. Der Satz gibt auch keine Auskunft über Extremalstellen an Intervallenden, an Spitzen oder anderen Stellen, an denen f nicht differenzierbar ist.

Als Kandidaten für Extremalstellen einer Funktion $f : I \rightarrow \mathbb{R}$ kommen also in Frage:

- a) die Randpunkte von I , falls I nicht offen ist
- b) die Punkte aus I , an denen f nicht differenzierbar ist
- c) die Punkte aus I , für die $f'(x) = 0$ gilt (stationäre Punkte).

Beispiel(e) 3.28

- Aus einer rechtwinkligen Blechplatte der Seitenlängen 16cm und 10cm soll eine quaderförmige, oben offene Wanne mit maximalem Volumen geformt werden.

Für das Volumen $V(x)$ in Abhängigkeit von der Höhe x der Wanne erhalten wir:

$$V : [0, 5] \rightarrow \mathbb{R}, \quad x \mapsto (16 - 2x) \cdot (10 - 2x) \cdot x$$

Wegen $V(0) = V(5) = 0$ sind die Randpunkte keine Maximierer. Es gilt:

$$V' : (0, 5) \rightarrow \mathbb{R}, \quad x \mapsto 12x^2 - 104x + 160. \tag{3.129}$$

$$V'(x) = 0 \iff 12x^2 - 104x + 160 = 0 \iff x = 2, \text{ da } x \in [0, 5].$$

Wir erhalten den globalen Maximierer $x_0 = 2$ mit $V(2) = 144\text{cm}^3$.

- In der Informationstheorie betrachtet man Folgen $\{x_i\}$ von Bits $x_i \in \{\pm 1\}$, $i \in \mathbb{N}_0$, mit Wahrscheinlichkeiten $P(x_i = 1) = w$, $P(x_i = -1) = 1 - w$, $w \in [0, 1]$.

Der Informationsgehalt eines Zeichens x_i wird gemessen durch $E : [0, 1] \rightarrow \mathbb{R}$, $w \mapsto -w \cdot \log_2(w) - (1 - w) \cdot \log_2(1 - w)$ für $w \in (0, 1)$, und $E(0) = E(1) = 0$.

Für welches w wird der Informationsgehalt maximal?

$$E' : (0, 1) \rightarrow \mathbb{R}, \quad w \mapsto -\log_2(w) - w \cdot \frac{1}{\ln(2) \cdot w} + \log_2(1 - w) + \frac{1 - w}{\ln(2)(1 - w)}. \tag{3.130}$$

$$E'(w) = 0 \iff \log_2(w) = \log_2(1 - w) \iff w = 1 - w \iff w = \frac{1}{2}. \tag{3.131}$$

Für $w = \frac{1}{2}$ erhält man den maximalen Informationsgehalt.

Ein wichtiges Hilfsmittel für die Analyse von gegebenen Funktionen bildet der folgende Satz:

Satz 3.29 (Mittelwertsatz)

Ist eine Funktion f auf einem abgeschlossenen Intervall $[a, b]$ definiert und stetig sowie auf dem offenen Intervall (a, b) differenzierbar, dann gibt es mindestens ein $x_0 \in (a, b)$ mit

$$f'(x_0) = \frac{f(b) - f(a)}{b - a}, \quad a \neq b. \tag{3.132}$$

Im Folgenden fassen wir wichtige Anwendungen des Mittelwertsatzes zusammen. Dazu sei $f : I \rightarrow \mathbb{R}$ eine auf dem offenen Intervall I differenzierbare Funktion. Es gilt:

- (a) $f'(x) > 0$ auf $I \implies f$ ist auf I streng monoton wachsend
 - (b) $f'(x) < 0$ auf $I \implies f$ ist auf I streng monoton fallend
 - (c) $f'(x) \geq 0$ auf $I \implies f$ ist auf I monoton wachsend
 - (d) $f'(x) \leq 0$ auf $I \implies f$ ist auf I monoton fallend
 - (e) $f'(x) = 0$ auf $I \implies f$ ist auf I konstant.
- (3.133)

Beispiel(e) 3.30

$$f : \mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto \frac{x}{\sqrt{1+x^2}}, \quad f' : \mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto \frac{1}{(\sqrt{1+x^2})^3}, \quad (3.134)$$

somit ist f streng monoton wachsend.

Der Mittelwertsatz kann auch dazu verwendet werden, Kandidaten für Extremalstellen zu klassifizieren.

Satz 3.31 (Extremwert-Test)
 Sei $f : (a, b) \rightarrow \mathbb{R}$ eine differenzierbare Funktion mit $f'(x_0) = 0$, so hat f in x_0 ein lokales Maximum (Minimum), falls es ein $\varepsilon > 0$ gibt mit $f'(y) > 0$ ($f'(y) < 0$) für alle $y \in (x_0 - \varepsilon, x_0)$ und $f'(y) < 0$ ($f'(y) > 0$) für alle $y \in (x_0, x_0 + \varepsilon)$.

Steht für die Funktion f die zweite Ableitung zur Verfügung, so kann der Extremwert-Test auch folgendermaßen durchgeführt werden:

$$\begin{aligned} f'(x_0) = 0 \text{ und } f''(x_0) > 0 &\implies x_0 \text{ ist lokales Minimum} \\ f'(x_0) = 0 \text{ und } f''(x_0) < 0 &\implies x_0 \text{ ist lokales Maximum} \end{aligned} \quad (3.135)$$

Die zweite Ableitung bestimmt die Krümmung des Graphen $y = f(x)$.
 Ist $f'' > 0$, so ist f konvex (Linkskrümmung), etwa $f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto x^2$.
 Ist $f'' < 0$, so ist f konkav (Rechtskrümmung), etwa $f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto -x^2$.
 Eine Stelle, an der sich die Krümmung ändert, wird als Wendepunkt bezeichnet. Als Kandidaten für Wendepunkte erhalten wir

- a) alle Punkte an denen zwar f , aber nicht f'' definiert ist
- b) alle Punkte mit $f''(x_0) = 0$.

Ist eine Funktion $f : I \rightarrow \mathbb{R}$ im offenen Intervall I dreimal differenzierbar und gilt für $x_0 \in I$:
 $f''(x_0) = 0, f'''(x_0) \neq 0$, so ist x_0 ein Wendepunkt.
 Gilt zudem $f'(x_0) = 0$, so heißt x_0 Sattelpunkt, zum Beispiel $f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto x^3, x_0 = 0$.
 Der Mittelwertsatz läßt sich folgendermaßen verallgemeinern:

Satz 3.32 (verallgemeinerter Mittelwertsatz)
 Seien $f, g : \mathbb{D} \subseteq \mathbb{R} \rightarrow \mathbb{R}$ in einem offenen Intervall $(a, b) \subset \mathbb{D}$ differenzierbar, in $[a, b] \subseteq \mathbb{D}$ stetig und $g'(x) \neq 0$ in (a, b) , dann gibt es mindestens eine Stelle $\xi \in (a, b)$ mit

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(\xi)}{g'(\xi)} \quad (g(b) \neq g(a)). \quad (3.136)$$

Dieser Satz ist Grundlage für die wichtige Regel von L'Hospital:
 Sind $f, g : (a, b) \rightarrow \mathbb{R}$ differenzierbar mit:

- a) $g'(x) \neq 0, x \in (a, b)$,
- b) $f(x) \rightarrow 0, g(x) \rightarrow 0$ oder $f(x) \rightarrow \pm\infty, g(x) \rightarrow \pm\infty$ für $x \rightarrow b-$
- c) $\lim_{x \rightarrow b-} \frac{f'(x)}{g'(x)} = L$ mit $L \in \mathbb{R} \cup \{\pm\infty\}$,

dann gilt:

$$\lim_{x \rightarrow b^-} \frac{f(x)}{g(x)} = \lim_{x \rightarrow b^-} \frac{f'(x)}{g'(x)}. \quad (3.137)$$

Entsprechendes gilt für $x \rightarrow a+$ bzw. $x \rightarrow \infty$ (dann ist $(a, b) = (a, \infty)$) oder $x \rightarrow -\infty$ (dann ist $(a, b) = (-\infty, b)$).

Beispiel(e) 3.33

- $\lim_{x \rightarrow 0} \frac{\sin(x)}{x} = \lim_{x \rightarrow 0} \frac{\cos(x)}{1} = 1$
- $\lim_{x \rightarrow \infty} x \cdot \ln\left(\frac{x+1}{x-1}\right) = \lim_{x \rightarrow \infty} \frac{\ln(x+1) - \ln(x-1)}{\frac{1}{x}} = \lim_{x \rightarrow \infty} \frac{\frac{1}{x+1} - \frac{1}{x-1}}{-\frac{1}{x^2}} = \lim_{x \rightarrow \infty} \frac{2x^2}{x^2-1} = 2$
- $\lim_{x \rightarrow 0} \left(\frac{1}{x} - \frac{1}{\sin(x)}\right) = \lim_{x \rightarrow 0} \frac{\sin(x) - x}{x \cdot \sin(x)} = \lim_{x \rightarrow 0} \frac{\cos(x) - 1}{\sin(x) + x \cdot \cos(x)} = \lim_{x \rightarrow 0} \frac{-\sin(x)}{\cos(x) + \cos(x) - x \cdot \sin(x)} = 0$

3.3 Potenzreihen

Die aus einer Zahlenfolge $\{a_k\}$, $k \in \mathbb{N}_0$, gebildete Summenfolge $\{s_n\}$, $n \in \mathbb{N}_0$, mit

$$s_n := \sum_{k=0}^n a_k \quad (3.138)$$

heißt unendliche Reihe und wird mit $\sum_{k=0}^{\infty} a_k$ bezeichnet. Die Folgenglieder s_i werden als Partialsummen bezeichnet. Die Konvergenz- und Divergenzbegriffe für Zahlenfolgen lassen sich durch die Betrachtung von $\{s_n\}$ auf unendliche Reihen übertragen.

Beispiel(e) 3.34

$\{s_n\} := \sum_{k=0}^{\infty} q^k$ mit $|q| < 1$. Es gilt:

$$\lim_{n \rightarrow \infty} s_n = \frac{1}{1-q}. \quad (3.139)$$

Somit können Konvergenzkriterien für Zahlenfolgen auch für unendliche Reihen verwendet werden. Speziell für unendliche Reihen ist das folgende Leibniz-Kriterium wichtig:

- Sei $\{a_k\}$, $k \in \mathbb{N}_0$, eine monoton fallende Nullfolge, so konvergiert die unendliche Reihe $\sum_{k=0}^{\infty} (-1)^k a_k$.

Elementare Manipulationen, die bei endlichen Summen erlaubt sind, sind bei unendlichen Reihen nicht uneingeschränkt möglich.

Beispiel(e) 3.35

Die unendliche Reihe $\sum_{k=0}^{\infty} a_k = (1-1) + (1-1) + \dots$ mit $a_k = (1-1) = 0$ konvergiert gegen 0. Die unendliche Reihe $\sum_{k=0}^{\infty} (-1)^k = 1 - 1 + 1 - 1 + 1 \dots$, die aus $\sum_{k=0}^{\infty} a_k$ durch Entfernen der Klammern entsteht, ist divergent.

Oft untersucht man, ob eine gegebene unendliche Reihe $\sum_{k=0}^{\infty} a_k$ absolut konvergiert, ob also die unendliche Reihe $\sum_{k=0}^{\infty} |a_k|$ einen Grenzwert besitzt. Dann konvergiert auch $\sum_{k=0}^{\infty} a_k$. Eine direkte Folge des Monotoniekriteriums für Zahlenfolgen ist der folgende

Satz 3.36 (absolute Konvergenz)

Eine unendliche Reihe $\sum_{k=0}^{\infty} a_k$ ist genau dann absolut konvergent, wenn die Zahlenfolge $\{b_n\}, n \in \mathbb{N}_0$, mit $b_n = \sum_{k=0}^n |a_k|$ beschränkt ist.

Ein wichtiges Kriterium für die Konvergenz unendlicher Reihen ist das **Quotientenkriterium**: Sei $\sum_{k=0}^{\infty} a_k$ eine unendliche Reihe mit $a_k \neq 0$ für alle $k \geq k_0 \in \mathbb{N}_0$ und sei $\left\{ \left| \frac{a_{k+1}}{a_k} \right| \right\}, k \geq k_0$, konvergent, so gilt:

$$\lim_{k \rightarrow \infty} \left| \frac{a_{k+1}}{a_k} \right| < 1 \implies \sum_{k=0}^{\infty} a_k \text{ ist absolut konvergent} \quad (3.140)$$

$$\lim_{k \rightarrow \infty} \left| \frac{a_{k+1}}{a_k} \right| > 1 \implies \sum_{k=0}^{\infty} a_k \text{ konvergiert nicht.} \quad (3.141)$$

Analog zu Zahlenfolgen kann man auch Folgen $\{f_n\}, n \in \mathbb{N}_0$, von Funktionen $f_i : \mathbb{D} \rightarrow \mathbb{R}$, betrachten. Für jedes $x \in \mathbb{D}$ ergibt sich dann eine Zahlenfolge $\{f_n(x)\}, n \in \mathbb{N}_0$. Eine Folge $\{p_n\}, n \in \mathbb{N}_0$,

$$p_n : \mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto \sum_{k=0}^n a_k x^k \quad (3.142)$$

heißt Potenzreihe und wird durch $\sum_{k=0}^{\infty} a_k x^k$ notiert. Dabei ist wichtig festzuhalten, dass die Koeffizienten von p_n und p_{n+1} für $x^i, i = 0, \dots, n$, übereinstimmen.

An einer Potenzreihe interessiert die Menge M aller $x \in \mathbb{R}$, sodass die unendliche Reihe $\{p_n(x)\}, n \in \mathbb{N}_0$, konvergiert.

Die Größe

$$R := \begin{cases} \sup\{|x|, x \in M\}, & \text{falls } M \text{ beschränkt ist} \\ \infty & \text{sonst} \end{cases} \quad (3.143)$$

heißt Konvergenzradius.

Beispiel(e) 3.37

- Die Potenzreihe $\sum_{k=0}^{\infty} k!x^k$ hat den Konvergenzradius 0. Sie konvergiert nur für $x = 0$.
- Die Potenzreihe $\sum_{k=0}^{\infty} x^k$ hat den Konvergenzradius 1, denn $M = \{x \in \mathbb{R}; |x| < 1\}$

Eine Potenzreihe $\sum_{k=0}^{\infty} a_k x^k$ mit Konvergenzradius $R > 0$ konvergiert absolut für alle $-R < x < R$.

Dies gilt auch für die Potenzreihe $\sum_{k=1}^{\infty} k a_k x^{k-1}$ und damit für

$$\sum_{k=l}^{\infty} k(k-1) \cdot \dots \cdot (k-(l-1)) a_k \cdot x^{k-l}, \quad l \in \mathbb{N}. \quad (3.144)$$

Ist für eine Potenzreihe $\sum_{k=0}^{\infty} a_k x^k$ $a_k \neq 0$ für $k \geq k_0 \in \mathbb{N}_0$, so läßt sich der Konvergenzradius durch

$$R = \lim_{k \rightarrow \infty} \left| \frac{a_k}{a_{k+1}} \right| \quad (3.145)$$

berechnen, falls $\left\{ \left| \frac{a_k}{a_{k+1}} \right| \right\}, k \geq k_0$, konvergiert oder bestimmt divergiert.

Beispiel(e) 3.38

Die Potenzreihe $\sum_{k=0}^{\infty} \frac{k^k}{k!} x^k$ hat den Konvergenzradius

$$R = \lim_{k \rightarrow \infty} \left| \frac{a_k}{a_{k+1}} \right| = \lim_{k \rightarrow \infty} \left(\frac{k}{k+1} \right)^k = \frac{1}{e}. \quad (3.146)$$

Sei nun $\sum_{k=0}^{\infty} a_k x^k$ eine Potenzreihe mit Konvergenzradius $R > 0$, so konvergieren die Zahlenfolgen

$\left\{ \sum_{k=0}^n a_k x^k \right\}$ und $\left\{ \sum_{n=l}^n k \cdot \dots \cdot (k-(l-1)) a_k x^{k-l} \right\}$ für alle $x \in (-R, R)$ und für alle $l \in \mathbb{N}$.

Somit existiert eine Funktion $f : (-R, R) \rightarrow \mathbb{R}$, $x \mapsto \lim_{n \rightarrow \infty} \sum_{k=0}^n a_k x^k$ und es gilt (**Differentiation von Potenzreihen**):

$$\begin{aligned} f' : (-R, R) &\rightarrow \mathbb{R}, & x &\mapsto \lim_{n \rightarrow \infty} \sum_{k=1}^n k \cdot a_k x^{k-1} \\ f'' : (-R, R) &\rightarrow \mathbb{R}, & x &\mapsto \lim_{n \rightarrow \infty} \sum_{k=2}^n k(k-1) \cdot a_k x^{k-2} \\ f^{(l)} : (-R, R) &\rightarrow \mathbb{R}, & x &\mapsto \lim_{n \rightarrow \infty} \sum_{k=l}^n k(k-1) \cdot \dots \cdot (k-(l-1)) a_k x^{k-l}, \quad l \in \mathbb{N}. \end{aligned} \quad (3.147)$$

Beispiel(e) 3.39

- $\exp : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto \lim_{n \rightarrow \infty} \sum_{k=0}^n \frac{1}{k!} x^k$
- $\sin : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto \lim_{n \rightarrow \infty} \sum_{k=0}^n \frac{(-1)^k}{(2k+1)!} x^{2k+1}$
- $\cos : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto \lim_{n \rightarrow \infty} \sum_{k=0}^n \frac{(-1)^k}{(2k)!} x^{2k}$
- $\ln(1 + \bullet) : (-1, 1) \rightarrow \mathbb{R}, x \mapsto \lim_{n \rightarrow \infty} \sum_{k=0}^n \frac{(-1)^k}{k+1} x^{k+1} (= \ln(1 + x))$

Eine unendliche Reihe der Form

$$\sum_{k=0}^{\infty} a_k (x - a)^k \quad (3.148)$$

heißt Potenzreihe mit dem Zentrum (Entwicklungspunkt) a . Die bisherigen Potenzreihen haben das Zentrum $a = 0$. Für theoretische Überlegungen genügt es, sich auf den Fall $a = 0$ zu beschränken, denn eine Potenzreihe mit Zentrum a geht durch die Substitution $z := x - a$ in die Potenzreihe $\sum_{k=0}^{\infty} a_k z^k$ über. Als Konvergenzradius einer Potenzreihe mit Zentrum a wird der Konvergenzradius der entsprechenden Potenzreihe mit Zentrum 0 betrachtet. Es gilt dann:

$$x \in (a - R, a + R) \implies \sum_{k=0}^{\infty} a_k (x - a)^k \text{ konvergiert} \quad (3.149)$$

$$x \notin [a - R, a + R] \implies \sum_{k=0}^{\infty} a_k (x - a)^k \text{ divergiert.} \quad (3.150)$$

Falls eine Funktion f über $(a - R, a + R)$, $R > 0$, als Potenzreihe

$$f(x) = \lim_{n \rightarrow \infty} \sum_{k=0}^n a_k (x - a)^k \quad (3.151)$$

mit Zentrum a darstellbar ist, gilt:

$$a_k := \frac{f^{(k)}(a)}{k!} \quad (f^{(k)}: \text{ die } k\text{-te Ableitung von } f; \quad f^{(0)} = f) \quad (3.152)$$

Die Potenzreihe $\sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!} (x - a)^k$ heißt Taylor-Reihe einer beliebig oft differenzierbaren Funktion $f : \mathbb{D} \rightarrow \mathbb{R}$ im Entwicklungspunkt a .

Ist R der Konvergenzradius von $\sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!} (x - a)^k$, $(a - R, a + R) \subseteq \mathbb{D}$, und

$$f(x) = \lim_{n \rightarrow \infty} \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (x - a)^k \quad (3.153)$$

für alle $a - R < x < a + R$, so sagt man: f läßt sich um a als Taylor-Reihe darstellen.

Beispiel(e) 3.40

Es kann passieren, dass die Taylor-Reihe für alle $x \in \mathbb{R}$ konvergiert; aber nur für $x = a$ gleich $f(a)$ ist.

$$\text{Sei } f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto \begin{cases} e^{-\frac{1}{x^2}} & \text{falls } x \neq 0 \\ 0 & x = 0, \end{cases}$$

so ist $\lim_{n \rightarrow \infty} \sum_{k=0}^n \frac{f^{(k)}(0)}{k!} x^k = 0$ für alle $x \in \mathbb{R}$, aber $f(x) \neq 0$ für $x \neq 0$.

Um festzustellen, wann eine Taylor-Reihe gegen den Funktionswert konvergiert, benötigen wir die Taylor-Formel:

Für jede auf einem offenen Intervall $I \subseteq \mathbb{R}$ $(n+1)$ -mal stetig differenzierbare Funktion $f : I \rightarrow \mathbb{R}$ und $a, x \in I$ gilt:

$$f(x) = \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (x-a)^k + R_{n+1}(x, a) \text{ mit} \quad (3.154)$$

$$R_{n+1}(x, a) = \frac{f^{(n+1)}(\xi)}{(n+1)!} (x-a)^{n+1} \text{ mit } \xi \text{ zwischen } x \text{ und } a. \quad (3.155)$$

Ist nun f auf dem Intervall I beliebig oft stetig differenzierbar und $a \in I$, dann konvergiert die Taylor-Reihe $\sum_{k=0}^{\infty} \frac{f^{(k)}(a)}{k!} (x-a)^k$ genau dann gegen $f(x)$ für $x \in I$, wenn $\lim_{n \rightarrow \infty} R_n(x, a) \rightarrow 0$.

Sind die Werte einer Funktion $f : I \rightarrow \mathbb{R}$ und ihrer ersten n Ableitungen in einem Punkt $x = a$ aus dem Innern des Intervalls I bekannt, dann wird mit

$$p(x) = f(a) + f'(a)(x-a) + \dots + \frac{f^{(n)}(a)}{n!} (x-a)^n \quad (\text{Taylor-Polynom}) \quad (3.156)$$

ein Polynom bestimmt, das die Funktion in einer Umgebung des Punktes $x = a$ gut approximiert. Es gilt:

$$p^{(k)}(a) = f^{(k)}(a) \quad 0 \leq k \leq n \quad (f^{(0)} = f) \quad (3.157)$$

Die Bedeutung der Taylor-Formel besteht darin, dass der Fehler $|f(x) - p(x)|$ durch

$$|f(x) - p(x)| = |R_{n+1}(x, a)| \quad (3.158)$$

darstellbar ist.

Beispiel(e) 3.41

- $e^x = 1 + x + \frac{x^2}{2!} + \dots + \frac{x^n}{n!} + \frac{e^\xi}{(n+1)!} x^{n+1}$.

Für $|x| \leq 1$ gilt:

$$\frac{e^\xi}{(n+1)!} |x|^{n+1} \leq \frac{e}{(n+1)!} |x|^{n+1} \quad (3.159)$$

Soll der Fehler kleiner oder gleich 10^{-8} sein, so genügt für $x = \frac{1}{10}$ zum Beispiel $n = 5$.

- $\sin(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \frac{\sin(\xi)}{8!} \cdot x^8$.

3.4 Integration

Mit der Integration wird das Problem gelöst, aus der Ableitung f' die ursprüngliche Funktion f zu rekonstruieren. Dazu verwendet man den Mittelwertsatz. Zu jeder Zerlegung

$$a = x_0 < x_1 < \dots < x_n = x \tag{3.160}$$

des Intervalls $[a, x]$, auf dem die Funktion f stetig und in (a, x) differenzierbar ist, gibt es Zwischenpunkte $\xi_i \in [x_{i-1}, x_i]$ mit

$$f(x_i) - f(x_{i-1}) = f'(\xi_i)(x_i - x_{i-1}), \quad i = 1, \dots, n \tag{3.161}$$

beziehungsweise

$$f(x) - f(x_0) = \sum_{i=1}^n f'(\xi_i)(x_i - x_{i-1}). \tag{3.162}$$

Die letzte Gleichung bildet die Grundidee zur Einführung des bestimmten Integrals. Sei f eine auf dem Intervall $[a, b]$ definierte beschränkte Funktion, die an höchstens endlich vielen Stellen nicht stetig ist (derartige Funktionen nennt man stückweise stetig).

Durch Einfügen von $(n - 1)$ Teilpunkten

$$a = x_0 < x_1 < \dots < x_n = b \tag{3.163}$$

wird $[a, b]$ in n Teilintervalle zerlegt.

Für jedes Teilintervall wählt man einen Zwischenpunkt $\xi_i \in [x_{i-1}, x_i]$ und betrachtet die Summe

$$z_n := \sum_{i=1}^n f(\xi_i)(x_i - x_{i-1}) \tag{3.164}$$

Man bezeichnet z_n nach dem Mathematiker B. RIEMANN (1826-1866) als Riemannsche Summe. Man kann nun zeigen, dass $\lim_{n \rightarrow \infty} z_n$ existiert, falls die maximale Intervallbreite der einzelnen Unterteilungen mit $n \rightarrow \infty$ gegen Null konvergiert. Außerdem ist dieser Grenzwert $\lim_{n \rightarrow \infty} z_n$ unabhängig von der Wahl der Teilpunkte und Zwischenpunkte und wird mit

$$\int_a^b f(x)dx := \lim_{n \rightarrow \infty} z_n \tag{3.165}$$

(das bestimmte Integral von f über $[a, b]$) bezeichnet. Die Randpunkte a, b heißen Integrationsgrenzen und $\int_a^b f(x)dx$ ist immer eine Zahl.

Üblicherweise setzt man $\int_a^a f(x)dx = 0, \int_b^a f(x)dx = -\int_a^b f(x)dx$ für $a < b$.

Ist $f : [a, b] \rightarrow \mathbb{R}$ stetig und $f(x) \geq 0$ für alle $x \in [a, b]$, so ist $\int_a^b f(x)dx$ der Flächeninhalt des von der Kurve $y = f(x)$, der x -Achse und den Geraden $x = a, x = b$ begrenzten Flächenstücks. Verläuft die Kurve $y = f(x)$ ganz unterhalb der x -Achse ($f(x) \leq 0, x \in [a, b]$), so kann man den Flächeninhalt I berechnen durch

$$I = \int_a^b (-f(x))dx = -\int_a^b f(x)dx, \quad \text{d.h.} \quad \int_a^b f(x)dx = -I \tag{3.166}$$

Begrenzt $y = f(x)$ Flächenstücke oberhalb und unterhalb der x -Achse, dann ist $\int_a^b f(x)dx$ die Summe der mit Vorzeichen versehenen Flächeninhalte.

Beispiel(e) 3.42

- $\int_a^b c dx = c(b - a)$
- $\int_a^b x dx = \frac{1}{2}(b^2 - a^2)$
- $\int_{-1}^1 x dx = 0.$

Für auf dem abgeschlossenen Intervall $[a, b]$ stückweise stetige Funktionen $f, g : [a, b] \rightarrow \mathbb{R}$ erhalten wir:

$$\int_a^b (\alpha f(x) + \beta g(x)) dx = \alpha \int_a^b f(x) dx + \beta \int_a^b g(x) dx, \quad \alpha, \beta \in \mathbb{R} \quad (3.167)$$

$$\int_a^b f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx \quad (a \leq c \leq b) \quad (3.168)$$

$$f(x) \leq g(x) \text{ für alle } a \leq x \leq b \implies \int_a^b f(x) dx \leq \int_a^b g(x) dx. \quad (3.169)$$

Für eine stetige Funktion $f : [a, b] \rightarrow \mathbb{R}$ gilt:

$$\left| \int_a^b f(x) dx \right| \leq \int_a^b |f(x)| dx \quad (3.170)$$

und aus $m \leq f(x) \leq M$ für alle $x \in [a, b]$ folgt $m(b - a) \leq \int_a^b f(x) dx \leq M(b - a)$.

Mit der letzten Doppelungleichung läßt sich der folgende Satz beweisen:

Satz 3.43 (Mittelwertsatz der Integralrechnung)

Sind die Funktionen $f, g : \mathbb{D} \rightarrow \mathbb{R}$ mit $[a, b] \subseteq \mathbb{D}$ auf $[a, b]$ stetig und $g(x) \geq 0$ für alle $x \in [a, b]$, dann gibt es mindestens ein $\xi \in [a, b]$ mit

$$\int_a^b f(x) \cdot g(x) dx = f(\xi) \int_a^b g(x) dx. \quad (3.171)$$

Der zu Beginn dieses Kapitels betrachtete Weg, aus der Ableitung f' einer Funktion f die Funktion f zu konstruieren, führt zu einer einfachen Methode, Integrale zu berechnen.

Definition 3.44 (Stammfunktion)

Man nennt eine auf einem offenen Intervall I differenzierbare Funktion F eine Stammfunktion einer Funktion $f : I \rightarrow \mathbb{R}$, falls $F'(x) = f(x)$ für alle $x \in I$ gilt.

Offensichtlich sind Stammfunktionen zu f nicht eindeutig, da mit F auch $F + c$, $c \in \mathbb{R}$, eine Stammfunktion von f ist:

$$(F(x) + c)' = F'(x) + 0 = f(x), \quad x \in I. \quad (3.172)$$

Die bereits angedeutete Beziehung zwischen Differentiation und Integration wird durch den folgenden Satz spezifiziert.

Satz 3.45 (Hauptsatz der Differential- und Integralrechnung)

Sei $f : I \rightarrow \mathbb{R}$ eine auf dem Intervall I stetige Funktion und $a, b \in I$, $a < b$, dann gilt:

- a) Die durch $F_a : I \rightarrow \mathbb{R}$, $x \mapsto \int_a^x f(t)dt$ definierte Funktion ist eine Stammfunktion von f , d.h.

$$\frac{d}{dx} \left(\int_a^x f(t)dt \right) = f(x) \quad \text{für alle } x \in (a, b). \quad (3.173)$$

Man nennt F_a eine Integralfunktion. Jede andere Stammfunktion von f hat die Form

$$F : I \rightarrow \mathbb{R}, \quad x \mapsto F_a(x) + c, \quad c \in \mathbb{R}. \quad (3.174)$$

- b) Sei F eine Stammfunktion von f , so gilt:

$$\int_a^b f(x)dx = F(b) - F(a) =: F(x)|_a^b. \quad (3.175)$$

Der Unterschied zwischen Integral- und Stammfunktion wird im folgenden Beispiel deutlich:

Sei $f : \mathbb{R} \rightarrow \mathbb{R}$, $x \mapsto x$, so ist die Menge aller Stammfunktionen F gegeben durch $F : \mathbb{R} \rightarrow \mathbb{R}$, $x \mapsto \frac{1}{2}x^2 + c$, $c \in \mathbb{R}$. Eine Integralfunktion $F_a : \mathbb{R} \rightarrow \mathbb{R}$, $x \mapsto \int_a^x tdt$ für $a \in \mathbb{R}$ ist von der Form $x \mapsto \frac{1}{2}x^2 - \frac{1}{2}a^2$. Somit sind alle Stammfunktionen mit $c \leq 0$ auch Integralfunktionen.

Definition 3.46 (Unbestimmtes Integral)

Die Menge aller Stammfunktionen von f wird mit $\int f(x)dx$ bezeichnet und heißt unbestimmtes Integral.

Beispiel(e) 3.47

mit $I = \mathbb{R}$ ($c \in \mathbb{R}$)

- $\int \cos(x)dx : \mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto \sin(x) + c$
- $\int x^n dx : \mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto \frac{x^{n+1}}{n+1} + c$
- $\int \sinh(x)dx : \mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto \cosh + c$
- $\int \cosh(x)dx : \mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto \sinh(x) + c$
- $\int_0^{\frac{\pi}{2}} \sin(x)dx = -\cos(x)|_0^{\frac{\pi}{2}} = -\cos(\frac{\pi}{2}) + \cos(0) = 1.$

Häufig wird die Konstante c weggelassen.

Im Folgenden betrachten wir Integrationsregeln, die direkt aus dem Hauptsatz der Differential- und Integralrechnung folgen:

1. Linearität:

$$\int (\alpha f(x) + \beta g(x)) dx = \alpha \int f(x) dx + \beta \int g(x) dx. \quad (3.176)$$

Beispiel:

$$\int \sum_{i=0}^n a_i x^i dx : \mathbb{R} \rightarrow \mathbb{R} \quad x \mapsto \sum_{i=0}^n \frac{a_i}{i+1} x^{i+1} + c. \quad (3.177)$$

2. Partielle Integration:

$$\int u'(x) \cdot v(x) dx = u \cdot v - \int u(x) \cdot v'(x) dx. \quad (3.178)$$

Beispiel(e) 3.48

- $\int \underbrace{x}_v \cdot \underbrace{e^x}_{u'} dx : \mathbb{R} \rightarrow \mathbb{R} \quad x \mapsto x \cdot e^x - \underbrace{(e^x + c)}_{\int u(x) \cdot v'(x) dx} = (x - 1)e^x + \bar{c}$
- $\int \underbrace{x}_v \cdot \underbrace{\cos(x)}_{u'} dx : \mathbb{R} \rightarrow \mathbb{R} \quad x \mapsto x \cdot \sin(x) + \underbrace{(\cos(x) + c)}_{\int u(x) \cdot v'(x) dx}$

3. Substitutionsmethode:

$$\int f(g(x)) \cdot g'(x) dx : \mathbb{D} \rightarrow \mathbb{R}, \quad x \mapsto F(g(x)) + c, \quad \text{wobei } F \text{ eine Stammfunktion von } f \text{ ist.} \quad (3.179)$$

Für das bestimmte Integral gilt somit:

$$\int_a^b f(g(x))g'(x)dx = F(g(x))\Big|_a^b = F(g(b)) - F(g(a)) = \int_{g(a)}^{g(b)} f(t)dt. \quad (3.180)$$

Gemäß der letzten Gleichungen gibt es zwei mögliche Anwendungen der Substitutionsregel:

1. Version: Berechne $\int f(g(x))g'(x)dx$ durch $F \circ g + c$.

Beispiel(e) 3.49

- $\int \frac{g'(x)}{g(x)} dx = \int f(g(x))g'(x)dx$ mit $f : \mathbb{R}_0^+ \setminus \{0\} \rightarrow \mathbb{R}, x \mapsto \frac{1}{x}, g > 0$,
also $\int \frac{g'(x)}{g(x)} dx : \mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto \ln(g(x)) + c$.
- $\int \frac{(\ln(x))^2}{x} dx = \int f(g(x))g'(x)dx$ mit $f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto x^2, g : \mathbb{R}_0^+ \setminus \{0\} \rightarrow \mathbb{R}, x \mapsto \ln(x)$,
also: $\int \frac{(\ln(x))^2}{x} dx : \mathbb{R}_0^+ \setminus \{0\} \rightarrow \mathbb{R}, \quad x \mapsto \frac{(\ln(x))^3}{3} + c$,
- $\int e^{\sin(x)} \cdot \cos(x) dx : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto e^{\sin(x)} + c$.

2. Version: Sei $g : \mathbb{D}_1 \rightarrow \mathbb{R}$ bijektiv und stetig differenzierbar, $f : \mathbb{D}_2 \rightarrow \mathbb{R}$ stückweise stetig und $g(\mathbb{D}_1) \subseteq \mathbb{D}_2$:

$$\int f(x)dx = H \circ g^{-1} + c, \quad \text{wobei } H \text{ Stammfunktion von } (f \circ g) \cdot g'. \quad (3.181)$$

„ersetze x durch $g(t)$, erweitere mit $g'(t)$ und ersetze in den Stammfunktionen t durch $g^{-1}(x)$ “.

Beispiel(e) 3.50

- Berechne:

$$\int \frac{e^{3x}}{e^{2x} - 1} dx : \mathbb{R}_0^+ \setminus \{0\} \rightarrow \mathbb{R}.$$

Wähle: $g = \ln$:

$$\begin{aligned} & \int \frac{t^3}{t^2 - 1} \frac{1}{t} dt = \int \frac{t^2}{t^2 - 1} dt = \\ & = \int \left(1 + \frac{1}{t^2 - 1} \right) dt : (1, \infty) \rightarrow \mathbb{R}, \quad t \xrightarrow{\text{Formelsammlung}} t + \frac{1}{2} \ln \left(\frac{t-1}{t+1} \right) + c, \end{aligned}$$

also:

$$\int \frac{e^{3x}}{e^{2x} - 1} dx : \mathbb{R}_0^+ \setminus \{0\} \rightarrow \mathbb{R}, \quad x \mapsto e^x + \frac{1}{2} \ln \left(\frac{e^x - 1}{e^x + 1} \right) + c.$$

- Berechne:

$$\int \frac{1}{\sqrt{x^2 + 1}} dx.$$

Wähle $g = \sinh$:

$$\begin{aligned} & \int \frac{1}{\sqrt{\sinh^2(t) + 1}} \cdot \cosh(t) dt = \\ & \quad \quad \quad \sqrt{\underbrace{\sinh^2(t) + 1}_{=\cosh^2(t)}} \\ & = \int \frac{\cosh(t)}{\cosh(t)} dt : \mathbb{R} \rightarrow \mathbb{R}, \quad t \mapsto t + c, \end{aligned}$$

also:

$$\int \frac{1}{\sqrt{x^2 + 1}} dx : \mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto \operatorname{arsinh}(x) + c.$$

4. Die Integration rationaler Funktionen

Für die Integration rationaler Funktionen ist deren Zerlegung in eine Summe von Partialbrüchen wichtig. Sei $f : \mathbb{D} \rightarrow \mathbb{R}, x \mapsto \frac{p(x)}{q(x)}$ mit Polynomen p, q ohne gemeinsamen Faktor, $\text{Grad } p < \text{Grad } q$. Sei ferner $q : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto c(x - b_1)^{k_1} (x - b_2)^{k_2} \dots (x - b_r)^{k_r} \cdot q_1(x)^{l_1} \cdot \dots \cdot q_s(x)^{l_s}$ mit den paarweise verschiedenen reellen Nullstellen b_i der Vielfachheit k_i und verschiedenen quadratischen Polynomen q_j , die in \mathbb{R} keine Nullstelle haben. Dann existieren reelle Zahlen $A_1^{(1)}, \dots, A_1^{(k_1)}, \dots, A_r^{(1)}, \dots, A_r^{(k_r)}, B_1^{(1)}, \dots, B_1^{(l_1)}, \dots, B_s^{(1)}, \dots, B_s^{(l_s)}, C_1^{(1)}, \dots, C_1^{(l_1)}, \dots, C_s^{(1)}, \dots, C_s^{(l_s)}$ mit:

$$\frac{p}{q} : \mathbb{D} \rightarrow \mathbb{R}, \quad x \mapsto \sum_{i=1}^r \sum_{j=1}^{k_i} \frac{A_i^{(j)}}{(x - b_i)^j} + \sum_{i=1}^s \sum_{j=1}^{l_i} \frac{B_i^{(j)} + C_i^{(j)} x}{(q_i(x))^j}. \quad (3.182)$$

(Partialbruchzerlegung)

Beispiel(e) 3.51

$$f : \mathbb{D} \rightarrow \mathbb{R}, \quad x \mapsto \frac{x^2 + x + 1}{(x-1)^3(x-2)} = \frac{A_1^{(1)}}{(x-1)} + \frac{A_1^{(2)}}{(x-1)^2} + \frac{A_1^{(3)}}{(x-1)^3} + \frac{A_2^{(1)}}{(x-2)}$$

Koeffizientenvergleich:

$$\begin{aligned} A_1^{(1)}(x-1)^2(x-2) + A_1^{(2)}(x-1)(x-2) + A_1^{(3)}(x-2) + A_2^{(1)}(x-1)^3 &= x^2 + x + 1 \\ \underbrace{(A_1^{(1)} + A_2^{(1)})}_{=0} x^3 + \underbrace{(-4A_1^{(1)} + A_1^{(2)} - 3A_2^{(1)})}_1 x^2 + \\ + \underbrace{(5A_1^{(1)} - 3A_1^{(2)} + A_1^{(3)} + 3A_2^{(1)})}_1 x + \underbrace{(-2A_1^{(1)} + 2A_1^{(2)} - 2A_1^{(3)} - A_2^{(1)})}_1 &= \\ = x^2 + x + 1 \implies A_1^{(1)} = -7, \quad A_1^{(2)} = -6, \quad A_1^{(3)} = -3, \quad A_2^{(1)} = 7. \end{aligned}$$

Hat man nun eine rationale Funktion f zu integrieren, so schreibt man zunächst f in der Form (Polynomdivision):

$$f : \mathbb{D} \rightarrow \mathbb{R}, \quad x \mapsto g(x) + \frac{p(x)}{q(x)}, \tag{3.183}$$

wobei g, p, q Polynome sind und $\text{Grad } p < \text{Grad } q$. Mit der Partialbruchzerlegung für $\frac{p}{q}$ sind somit lediglich Funktionen der Form $x \mapsto \sum_{i=0}^n a_i x_i$, $x \mapsto \frac{A}{(x-b)^k}$, $x \mapsto \frac{B+Cx}{q_i(x)^m}$, $k, m \in \mathbb{N}$ zu integrieren, wobei q_i ein quadratisches Polynom ohne reelle Nullstelle ist. Es gilt:

- $$\int \frac{1}{(x-b)} dx : \mathbb{R} \setminus \{b\} \rightarrow \mathbb{R}, \quad x \mapsto \ln(|x-b|) + c \tag{3.184}$$

- $$\int \frac{1}{(x-b)^k} dx : \mathbb{R} \setminus \{b\} \rightarrow \mathbb{R}, \quad x \mapsto -\frac{1}{k-1} \frac{1}{(x-b)^{k-1}} + c, \quad k > 1 \tag{3.185}$$

- $$\int \frac{1}{x^2 + \alpha x + \beta} dx : \mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto \frac{2}{\sqrt{4\beta - \alpha^2}} \arctan\left(\frac{2x + \alpha}{\sqrt{4\beta - \alpha^2}}\right) + c \tag{3.186}$$

wobei $\alpha^2 - 4\beta < 0$

- $$\begin{aligned} \int \frac{1}{(x^2 + \alpha x + \beta)^k} dx : \mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto & \frac{2x + \alpha}{(k-1)(4\beta - \alpha^2)(x^2 + \alpha x + \beta)^{k-1}} + \\ & + \frac{2(2k-3)}{(k-1)(4\beta - \alpha^2)} s(x) + c, \end{aligned} \tag{3.187}$$

wobei s eine Stammfunktion von

$$\mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto \frac{1}{(x^2 + \alpha x + \beta)^{k-1}}$$

darstellt ($\alpha^2 - 4\beta < 0, k > 1$).

•

$$\int \frac{ax + b}{(x^2 + \alpha x + \beta)^k} dx : \mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto -\frac{a}{2(k-1)(x^2 + \alpha x + \beta)^{k-1}} + \left(b - \frac{a\alpha}{2}\right) s(x) + c, \quad (3.188)$$

wobei s eine Stammfunktion von

$$\mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto \frac{1}{(x^2 + \alpha x + \beta)^k}$$

darstellt ($\alpha^2 - 4\beta < 0, k > 1$).

•

$$\int \frac{ax + b}{(x^2 + \alpha x + \beta)} dx : \mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto \frac{a}{2} \ln(|x^2 + \alpha x + \beta|) + \left(b - \frac{a\alpha}{2}\right) s(x) + c, \quad (3.189)$$

wobei s eine Stammfunktion von

$$\mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto \frac{1}{(x^2 + \alpha x + \beta)}$$

darstellt ($\alpha^2 - 4\beta < 0, k > 1$).

Falls $\alpha^2 - 4\beta \geq 0$, hat $x \mapsto x^2 + \alpha x + \beta$ reelle Nullstellen.

Beispiel(e) 3.52

$$\begin{aligned} & \int \frac{3x^5 - 2x^4 + 4x^3 + 4x^2 - 7x + 6}{(x-1)^2(x^2+1)^2} dx = \\ & = \int \left(\frac{1}{x-1} + \frac{2}{(x-1)^2} + \frac{2x+1}{x^2+1} + \frac{4}{(x^2+1)^2} \right) dx : \mathbb{R} \setminus \{1\} \rightarrow \mathbb{R} \\ x & \mapsto \ln(|x-1|) - \frac{2}{x-1} + \ln(x^2+1) + \arctan(x) + \frac{2x}{x^2+1} + 2 \arctan(x) + c. \end{aligned}$$

Ist eine Funktion $f : [a, b) \rightarrow \mathbb{R}$ auf jedem abgeschlossenen Intervall $[a, c]$, $a < c < b$, $b \in \mathbb{R} \cup \{\infty\}$ stückweise stetig, so heißt das Integral

$$\int_a^b f(x) dx := \lim_{c \rightarrow b^-} \int_a^c f(x) dx, \quad (3.190)$$

bzw.

$$\int_a^\infty f(x) dx := \lim_{c \rightarrow \infty} \int_a^c f(x) dx \quad (3.191)$$

uneigentliches Integral. Analog definiert man für $f : (a, b] \rightarrow \mathbb{R}$:

$$\int_a^b f(x) dx := \lim_{c \rightarrow a^+} \int_c^b f(x) dx \quad (3.192)$$

bzw.

$$\int_{-\infty}^b f(x) dx := \lim_{c \rightarrow -\infty} \int_c^b f(x) dx. \quad (3.193)$$

Beispiel(e) 3.53

- $\int_1^{\infty} \frac{1}{x} dx = \lim_{c \rightarrow \infty} \int_1^c \frac{1}{x} dx = \lim_{c \rightarrow \infty} \ln(c) = \infty$
- $\int_0^1 \frac{1}{x} dx = \lim_{c \rightarrow 0^+} \int_c^1 \frac{1}{x} dx = \lim_{c \rightarrow 0^+} (-\ln(c)) = \infty$
- $\int_1^{\infty} \frac{1}{x^\alpha} dx = \begin{cases} \frac{1}{\alpha-1} & \text{falls } \alpha > 1 \\ \infty & \text{falls } \alpha \leq 1 \end{cases}$
- $\int_0^1 \frac{1}{x^\alpha} dx = \begin{cases} \infty & \text{falls } \alpha \geq 1 \\ \frac{1}{1-\alpha} & \text{falls } \alpha < 1 \end{cases}$

Ein an beiden Grenzen uneigentliches Integral ist definiert durch

$$\int_a^b f(x) dx = \lim_{u \rightarrow a^+} \int_u^c f(x) dx + \lim_{v \rightarrow b^-} \int_c^v f(x) dx \quad \text{mit } a < c < b. \quad (3.194)$$

Die Grenzwerte auf der rechten Seite sind voneinander unabhängig!

Es gilt also: $\int_{-\infty}^{\infty} x dx$ ist nicht definiert!

Beispiel(e) 3.54

-

$$\begin{aligned} \int_{-\infty}^{\infty} \frac{1}{1+x^2} dx &= \lim_{a \rightarrow -\infty} \int_a^0 \frac{1}{1+x^2} dx + \\ &+ \lim_{b \rightarrow \infty} \int_0^b \frac{1}{1+x^2} dx = \\ &= \arctan(0) - \lim_{a \rightarrow -\infty} \arctan(a) + \lim_{b \rightarrow \infty} \arctan(b) - \\ &- \arctan(0) = \frac{\pi}{2} + \frac{\pi}{2} = \pi. \end{aligned}$$

- $\Gamma : (0, \infty) \rightarrow \mathbb{R}, x \mapsto \int_0^{\infty} e^{-t} t^{x-1} dt$

Die Gammafunktion Γ ist über ein an beiden Seiten uneigentliches Integral definiert. Es gilt:

$$\Gamma(x+1) = x \cdot \Gamma(x), \quad \Gamma(n+1) = n! \quad \text{für alle } n \in \mathbb{N}. \quad (3.195)$$

Anwendung der Integrationstheorie:

Funktionsapproximation:

Die bereits betrachtete Taylorformel für eine auf dem offenen Intervall I $(n+1)$ -mal stetig diffe-

renzierbare Funktion

$$f : I \rightarrow \mathbb{R}, \quad x \mapsto \sum_{i=0}^n \frac{1}{i!} f^{(i)}(a)(x-a)^i + \frac{f^{(n+1)}(\xi)}{(n+1)!} (x-a)^{n+1} \quad (a \in I) \quad (3.196)$$

kann auch mit Hilfe der Integrationstheorie in der Form

$$f : I \rightarrow \mathbb{R}, \quad x \mapsto \sum_{i=0}^n \frac{1}{i!} f^{(i)}(a)(x-a)^i + \underbrace{\frac{1}{n!} \int_a^x (x-t)^n f^{(n+1)}(t) dt}_{R_{n+1}(x,a)} \quad (3.197)$$

geschrieben werden. Diese Form des Restgliedes ist häufig praktikabler als die differentielle Form. Weitere Anwendungen der Integrationstheorie folgen in den folgenden Kapiteln.

Kapitel 4

Gewöhnliche Differentialgleichungen

Es sei $y : \mathbb{R}^+ \rightarrow \mathbb{R}$, $t \mapsto y(t)$ eine Funktion mit der Interpretation: $y(t)$ ist die Menge eines zum Zeitpunkt t vorhandenen radioaktiven Materials. Ein radioaktiver Zerfallsprozess kann mathematisch durch folgende Gleichung beschrieben werden:

$$\dot{y}(t) = a \cdot y(t) \quad \text{mit} \quad a < 0. \quad (4.1)$$

Hier bezeichnen wir mit \dot{y} , wie in der Physik anstelle von y' üblich, die Ableitung der zeitabhängigen Funktion y . Gleichung (4.1) sagt aus, dass die Abnahme $\dot{y}(t)$ des Materials zum Zeitpunkt t abhängt von der zu diesem Zeitpunkt noch vorhandenen Restmenge an Material. Gleichung (4.1) stellt also einen funktionalen Zusammenhang her zwischen einer (gesuchten) Funktion und ihren Ableitungen und heißt deshalb **Differentialgleichung**.

In Gleichung (4.1) tritt nur die erste Ableitung von y auf — es handelt sich um eine Differentialgleichung *erster Ordnung*. Die allgemeinste Differentialgleichung erster Ordnung für eine Funktion $y : t \mapsto y(t)$ hat die Form

$$F(t, y(t), \dot{y}(t)) = 0, \quad (4.2)$$

wobei F im Fall des Beispiels (4.1) die durch $F(u, v, w) = w - av$ gegebene lineare Funktion ist. Eine Funktion $y : t \mapsto y(t)$ heißt **Lösung** der Differentialgleichung (4.2), wenn y auf einem Intervall $I \subseteq \mathbb{R}$ differenzierbar ist und die Gleichung (4.2) für alle $t \in I$ erfüllt. Hier wie in allen folgenden Abschnitten verstehen wir unter einem Intervall $I \subseteq \mathbb{R}$ immer eine Menge der Form $I = (a, b)$ (offenes Intervall), $I = [a, b]$ (abgeschlossenes Intervall), $I = (a, b]$ oder $I = [a, b)$ (halboffenes Intervall) mit $a < b$. Ist I bei a offen, dann ist auch $a = -\infty$ erlaubt und ist I bei b offen, dann ist auch $b = \infty$ erlaubt. An abgeschlossenen Intervallrändern sind Stetigkeits- und Differenzierbarkeitsforderungen stets im Sinn einseitiger Grenzwerte zu verstehen.

Im Beispiel (4.1) kann man sofort nachprüfen, dass die Funktionen

$$y_C : t \mapsto y_C(t) = C \cdot e^{at}, \quad C \in \mathbb{R},$$

auf $I = \mathbb{R}$ definierte Lösungen der Differentialgleichung ist — es sind dies auch die einzigen Lösungen, wie wir bereits gesehen haben. Kennt man zusätzlich die Menge y_0 des zum Zeitpunkt t_0 vorhandenen Materials, kann man die unbestimmte Konstante C aus der Gleichung $y_0 = C \exp(at_0)$ bestimmen und bekommt eine *eindeutige* Lösung der Differentialgleichung. Man spricht von einem **Anfangswertproblem**, wenn neben einer Differentialgleichung noch der Funktionswert vorgegeben wird, den die Lösung an einer Stelle annehmen soll.

Betrachten wir ein weiteres Beispiel, bei dem es nicht mehr um eine skalare, sondern um eine vektorwertige Funktion (eine Kurve) $\mathbf{r} : \mathbb{R}_0^+ \rightarrow \mathbb{R}^3$ geht: Eine Aufgabenstellung der Mechanik

besteht darin, für Zeitpunkte $t \geq 0$ die Bahnkurve $\mathbf{r}(t) \in \mathbb{R}^3$ eines Massenpunkts zu bestimmen. Es ist also

$$\mathbf{r} : \mathbb{R}_0^+ \rightarrow \mathbb{R}^3, \quad t \mapsto \mathbf{r}(t) = \begin{pmatrix} r_1(t) \\ r_2(t) \\ r_3(t) \end{pmatrix},$$

wobei $r_1(t)$ „die x -Position“, $r_2(t)$ „die y -Position“ und $r_3(t)$ „die z -Position“ des Massepunkts zur Zeit t bezeichnet. Geschwindigkeit und Beschleunigung des Massepunkts sind durch komponentenweise definierte erste und zweite Ableitungen gegeben, also

$$\dot{\mathbf{r}}(t) = \begin{pmatrix} \dot{r}_1(t) \\ \dot{r}_2(t) \\ \dot{r}_3(t) \end{pmatrix}, \quad \ddot{\mathbf{r}}(t) = \begin{pmatrix} \ddot{r}_1(t) \\ \ddot{r}_2(t) \\ \ddot{r}_3(t) \end{pmatrix}.$$

Ein simpler Fall liegt vor, wenn Beschleunigungswerte $\ddot{\mathbf{r}}(t) = \mathbf{a}(t)$ explizit gegeben sind. In diesem einfachen Fall erhält man \mathbf{r} durch 2-malige Integration der Komponenten von \mathbf{a} . Dabei entsteht ein unbestimmter linearer Term der Form $\mathbf{u} + t \cdot \mathbf{v}$ mit konstanten Vektoren $\mathbf{u}, \mathbf{v} \in \mathbb{R}^3$, der wiederum durch Zusatzbedingungen festgelegt werden kann, am einfachsten durch die Anfangsbedingungen $\mathbf{u} = \mathbf{r}(0)$ (bekannter Ort des Massenpunkts zum Zeitpunkt 0) und $\mathbf{v}(0) = \dot{\mathbf{r}}(0)$ (bekannte Anfangsgeschwindigkeit des Massenpunkts).

Im allgemeinen wird die Beschleunigung \mathbf{a} nicht explizit gegeben sein, man weiß jedoch, dass sie durch eine Kraft

$$\mathbf{F} = \begin{pmatrix} F_1 \\ F_2 \\ F_3 \end{pmatrix}$$

(mit Komponenten in x -, y - und z -Richtung) verursacht wird, so dass nach dem 2. Newtonschen Bewegungsgesetz (konstante Masse m vorausgesetzt)

$$\mathbf{F} = m\ddot{\mathbf{r}} = m\mathbf{a}$$

gilt. Typischerweise ist die Kraft eine Funktion der Gestalt

$$\mathbf{F} = \mathbf{F}(\mathbf{r}, \dot{\mathbf{r}}, t)$$

(zum Beispiel wirken Gravitationskräfte in Abhängigkeit vom Ort \mathbf{r} des Massenpunkts und Reibungskräfte in Abhängigkeit von der Geschwindigkeit $\dot{\mathbf{r}}$) und damit lautet die **Bewegungsgleichung** für die Funktion \mathbf{r} :

$$m\ddot{\mathbf{r}} = \mathbf{F}(\mathbf{r}, \dot{\mathbf{r}}, t).$$

Die Bewegungsgleichung besteht hier aus einem ganzen System von gekoppelten Differentialgleichungen, ausgeschrieben

$$\begin{pmatrix} m\ddot{r}_1(t) \\ m\ddot{r}_2(t) \\ m\ddot{r}_3(t) \end{pmatrix} = \begin{pmatrix} F_1(t, r_1(t), r_2(t), r_3(t), \dot{r}_1(t), \dot{r}_2(t), \dot{r}_3(t)) \\ F_2(t, r_1(t), r_2(t), r_3(t), \dot{r}_1(t), \dot{r}_2(t), \dot{r}_3(t)) \\ F_3(t, r_1(t), r_2(t), r_3(t), \dot{r}_1(t), \dot{r}_2(t), \dot{r}_3(t)) \end{pmatrix}.$$

Wiederum besteht ein Gleichung zwischen (den Komponenten) einer (gesuchten) Funktion und deren Ableitungen. Bewegungsgleichungen oder allgemeiner Energieerhaltungssätze der Physik sind der wichtigste Grund für das Auftreten von Differentialgleichungen.

Wir geben nun die formale Definition zunächst von skalaren und im Anschluss daran von Systemen von Differentialgleichungen an. Dabei kehren wir wieder zu der gewohnten Schreibweise y', y'', \dots für erste, zweite und höhere Ableitungen einer Funktion y zurück und nennen die unabhängige Variable wieder x statt t .

4.1 Gewöhnliche Differentialgleichungen n-ter Ordnung

Definition 4.1 (gewöhnliche Differentialgleichungen n-ter Ordnung)

Seien $n \in \mathbb{N}$, $F : \mathbb{D} \rightarrow \mathbb{R}$, $\mathbb{D} \subseteq \mathbb{R}^{n+2}$, $f : \mathbb{G} \rightarrow \mathbb{R}$, $\mathbb{G} \subseteq \mathbb{R}^{n+1}$.

Eine Bestimmungsgleichung

$$F(x, y(x), y'(x), \dots, y^{(n)}(x)) = 0 \quad \text{für alle } x \in I \subseteq \mathbb{R} \text{ offen.} \quad (4.3)$$

für eine gesuchte Funktion $y : I \rightarrow \mathbb{R}$, $x \mapsto y(x)$ heißt implizite gewöhnliche Differentialgleichung n-ter Ordnung.

Erlaubt F die Gestalt

$$y^{(n)}(x) = f(x, y(x), y'(x), \dots, y^{(n-1)}(x)), \quad (4.4)$$

so spricht man von einer expliziten gewöhnlichen Differentialgleichung n-ter Ordnung.

Eine n-mal stetig differenzierbare Funktion $y : I \rightarrow \mathbb{R}$ heißt explizite Lösung von (4.3) bzw. (4.4), falls

$$(x, y(x), y'(x), \dots, y^{(n)}(x)) \in \mathbb{D} \quad \text{für alle } x \in I \text{ und} \quad (4.5)$$

$$F(x, y(x), y'(x), \dots, y^{(n)}(x)) = 0 \quad \text{für alle } x \in I \quad (4.6)$$

beziehungsweise

$$(x, y(x), y'(x), \dots, y^{(n-1)}(x)) \in \mathbb{G} \quad \text{für alle } x \in I \text{ und} \quad (4.7)$$

$$y^{(n)}(x) = f(x, y(x), y'(x), \dots, y^{(n-1)}(x)) = 0 \quad \text{für alle } x \in I. \quad (4.8)$$

Eine implizite Lösung ist eine durch $g : \mathbb{B} \rightarrow \mathbb{R}$, $\mathbb{B} \subseteq \mathbb{R}^2$ mit

$$g(x, y) = 0 \quad \text{für alle } x, y \in \mathbb{B} \quad (4.9)$$

gegebene Funktion $y : I \rightarrow \mathbb{R}$, die die entsprechende Differentialgleichung löst und die wenigstens auf einem Teilintervall $\tilde{I} \subseteq I$ explizit in der Form $x \mapsto y(x)$ darstellbar ist.

Beispiel(e) 4.2

- $y'(x) = xy^2(x)$ ist eine explizite gewöhnliche Differentialgleichung erster Ordnung. Eine Lösung für $I = \mathbb{R}$ ist

$$y : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto -\frac{2}{1+x^2}.$$

- $(y(x) \cdot y^{(5)}(x))^2 + xy^{(2)}(x) + \ln(y(x)) = 0$ ist eine implizite gewöhnliche Differentialgleichung 5. Ordnung. Eine Lösung wäre z.B. $y : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto 1$.

- $y(x) \cdot y'(x) + x = 0$ ist eine implizite gewöhnliche Differentialgleichung 1. Ordnung mit einer impliziten Lösung $g(x, y) = 0$, wobei

$$g : \mathbb{R}^2 \rightarrow \mathbb{R}, (x, y) \mapsto x^2 + y^2 - c, \quad c \geq 0.$$

- Die implizite gewöhnliche Differentialgleichung erster Ordnung

$$(y'(x))^2 + 1 = 0$$

besitzt keine Lösung.

Wie man bereits am Beispiel

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

mit den Lösungsfunktionen $y_{a,b} : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto ax + b, a, b \in \mathbb{R}$ sieht, kann die Lösungsmenge einer gewöhnlichen Differentialgleichung mehrparametrische Kurvenscharen enthalten. Man fasst jede r -parametrische Kurvenschar $y_{c_1, \dots, c_r} : I \rightarrow \mathbb{R}, c_i \in I_i \subseteq \mathbb{R}$ als eine Lösung mit r freien Parametern (den sogenannten Integrationskonstanten c_1, \dots, c_r) auf. Jede einzelne Lösung, die keine wählbaren Konstanten enthält, wird als spezielle oder partikuläre Lösung bezeichnet. Gehört sie keiner Lösungsschar an, so nennt man sie singuläre Lösung. Eine parameterabhängige Lösung einer gewöhnlichen Differentialgleichung n -ter Ordnung heißt allgemein, wenn sie n frei wählbare Konstanten enthält. Sie heißt vollständig, wenn dadurch sämtliche Lösungen erfasst werden. Entsprechende Begriffe gelten auch für implizite Lösungen.

Beispiel(e) 4.3

- $y'(x) - 5y(x) = 0$ hat auf $I = \mathbb{R}$ die allgemeine Lösung

$$y_c : \mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto c \cdot e^{5x}, \quad c \in \mathbb{R}.$$

Diese Lösung ist vollständig.

- $y''(x) + 9y(x) - 9 = 0$ besitzt auf \mathbb{R} die spezielle Lösung

$$y : \mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto 1$$

und die allgemeine Lösung

$$y_{c_1, c_2} : \mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto 1 + c_1 \sin(3x) + c_2 \cos(3x)$$

mit $c_1, c_2 \in \mathbb{R}$. Diese ist vollständig.

- $|y'(x)| + |y| = 0$ hat keine allgemeine Lösung, sondern nur die singuläre Lösung

$$y : \mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto 0.$$

- $y'(x)^2 - 4xy'(x) + 4y(x) = 0$ ist eine implizite gewöhnliche Differentialgleichung erster Ordnung mit der allgemeinen Lösung

$$y_c : \mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto 2cx - c^2, \quad c \in \mathbb{R}$$

und der singulären Lösung $y : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto x^2$.

Fügt man einer gewöhnlichen Differentialgleichung zusätzliche Bedingungen hinzu, um die Integrationskonstanten festzulegen, so erhält man als einfachsten Fall Anfangswertprobleme:

$$y^{(n)}(x) = f(x, y(x), \dots, y^{(n-1)}(x)) \quad (4.10)$$

mit vorgegebenen Werten $y(x_0) = y_0, y'(x_0) = y_1, \dots, y^{(n-1)}(x_0) = y_{n-1}$, wobei

$$(x_0, y_0, y_1, \dots, y_{n-1})$$

im Definitionsbereich von f liegen muss.

Die Bezeichnung „Anfangswertproblem“ stammt aus technischen Anwendungen, in denen man Zeitfunktionen $x : [a, b] \rightarrow \mathbb{R}$ betrachtet, die durch

$$x^{(n)}(t) = f(t, x(t), \dot{x}(t), \dots, x^{(n-1)}(t)) \quad (4.11)$$

mit Anfangszustand $x(t_0) = x_0, \dot{x}(t_0) = \dot{x}_0, \dots, x^{(n-1)}(t_0) = x_0^{(n-1)}, t_0 \in]a, b[$ gegeben sind ($(t_0, x_0, \dots, x_0^{(n-1)})$ im Definitionsbereich von f).

Beispiel(e) 4.4

Für ein lineares Pendel modelliert durch

$$\ddot{x}(t) + \omega^2 x(t) = 0, \quad x(0) = x_0, \dot{x}(0) = v_0 \quad (4.12)$$

gibt es die eindeutige Lösung:

$$x : \mathbb{R}_0^+ \rightarrow \mathbb{R}, \quad t \mapsto x_0 \cos(\omega t) + \frac{v_0}{\omega} \sin(\omega t), \quad \omega \neq 0. \quad (4.13)$$

Ein Anfangswertproblem heißt lokal lösbar, falls es ein $\epsilon > 0$ gibt, sodass auf $I = (x_0 - \epsilon, x_0 + \epsilon)$ eine Lösung $y : (x_0 - \epsilon, x_0 + \epsilon) \rightarrow \mathbb{R}$ der gewöhnlichen Differentialgleichung

$$y^{(n)}(x) = f\left(x, y(x), \dots, y^{(n-1)}(x)\right) \quad (4.14)$$

mit den Anfangsbedingungen

$$y(x_0) = y_0, y'(x_0) = y_1, \dots, y^{(n-1)}(x_0) = y_{n-1}, \quad (4.15)$$

existiert ($(x_0, y_0, y_1, \dots, y_{n-1})$ im Definitionsbereich von f). Diese Lösung heißt lokale Lösung. Ein Anfangswertproblem heißt „sachgemäß gestellt“, falls

- die Existenz und Eindeutigkeit einer lokalen Lösung
- die stetige Abhängigkeit der lokalen Lösung von den Anfangswerten

gewährleistet ist.

Für sachgemäß gestellte Probleme stellen sich zusätzliche Fragen:

- globale Existenz,
- explizite Bestimmung einer Lösung.

Beispiel(e) 4.5

Jede kubische Funktion $y : \mathbb{R} \rightarrow \mathbb{R}$, $x \mapsto (x - c)^3$, $c \in \mathbb{R}$, genügt der gewöhnlichen Differentialgleichung

$$y'(x) = 3\sqrt[3]{y^2(x)}. \quad (4.16)$$

Da auch $y : \mathbb{R} \rightarrow \mathbb{R}$, $x \mapsto 0$ eine singuläre Lösung dieser gewöhnlichen Differentialgleichung ist, hat das Anfangswertproblem

$$y'(x) = 3\sqrt[3]{y^2(x)}, \quad y(0) = 0 \quad (4.17)$$

zumindest zwei Lösungen: $\hat{y} : \mathbb{R} \rightarrow \mathbb{R}$, $x \mapsto x^3$ und $y : \mathbb{R} \rightarrow \mathbb{R}$, $x \mapsto 0$.

Im Folgenden betrachten wir spezielle Klassen gewöhnlicher Differentialgleichungen. Eine gewöhnliche Differentialgleichung der Form

$$y'(x) = h(x)g(y(x)) \quad (4.18)$$

mit stetigen, auf Intervallen $I \subseteq \mathbb{R}$ und $J \subseteq \mathbb{R}$ definierten Funktionen $h : I \rightarrow \mathbb{R}$ und $g : J \rightarrow \mathbb{R}$ heißt trennbar. Eine Lösung der obigen trennbaren gewöhnlichen Differentialgleichung ist für $\xi \in I$, $\eta \in J$ und dem Anfangswert $y(\xi) = \eta$ implizit gegeben durch

$$\int_{\eta}^{y(x)} \frac{1}{g(t)} dt = \int_{\xi}^x h(t) dt, \quad (4.19)$$

falls $g(t) \neq 0$ für alle $t \in J$.

Ist $g(\eta) = 0$ für ein $\eta \in J$, dann ist die Funktion $y : I \rightarrow \mathbb{R}$, $x \mapsto \eta$ eine konstante Lösung des Anfangswertproblems

$$y'(x) = h(x)g(y(x)), \quad y(\xi) = \eta, \quad \xi \in I. \quad (4.20)$$

Beispiel(e) 4.6

$$\bullet y'(x) = y^2(x), \quad y(0) = 1. \quad \int_1^{y(x)} \frac{1}{t^2} dt = \int_0^x dt \Leftrightarrow -\frac{1}{y(x)} + 1 = x,$$

$$\text{also: } y : (-\infty, 1) \rightarrow \mathbb{R}, \quad x \mapsto \frac{1}{1-x}.$$

$$\bullet y'(x) = \frac{1-y(x)}{x}, \quad y(\xi) = \eta, \quad \eta \neq 1, \quad \xi \neq 0.$$

$$\int_{\eta}^{y(x)} \frac{1}{1-t} dt = \int_{\xi}^x \frac{1}{t} dt \Leftrightarrow -\ln(|1-y(x)|) + \ln(|1-\eta|) = \ln(|x|) - \ln(|\xi|), \quad (4.21)$$

$$\text{also: } |1-y(x)| = \left| \frac{\xi(1-\eta)}{x} \right|.$$

$$\text{Für } \xi > 0, \quad |\eta| < 1 \text{ folgt: } y : (0, \infty) \rightarrow \mathbb{R}, \quad x \mapsto 1 - \frac{\xi(1-\eta)}{x}.$$

Lineare gewöhnliche Differentialgleichungen erster Ordnung sind von der Form

$$y'(x) + a(x)y(x) = f(x) \quad (4.22)$$

mit Funktionen definiert auf einem Intervall $I \subseteq \mathbb{R}$.

Diese gewöhnliche Differentialgleichung heißt homogen, falls $f \equiv 0$; f wird Störfunktion genannt. Ist die Funktion a auf I stetig und ist $f \equiv 0$, so erhält man durch

$$y : I \rightarrow \mathbb{R}, \quad x \mapsto \exp\left(-\int a(x)dx\right), \quad (4.23)$$

eine allgemeine Lösung der gewöhnlichen Differentialgleichung

$$y'(x) + a(x)y(x) = 0 \quad (4.24)$$

(die Konstante c ist in der Menge aller Stammfunktionen $\int a(x)dx$ versteckt). Diese gewöhnliche Differentialgleichung wird als Modell für Wachstumsprozesse gedeutet.

Beispiel(e) 4.7

Für $a(x) \equiv -a$ hat das gewöhnliche Anfangswertproblem

$$y'(x) = a \cdot y(x), \quad y(\xi) = \eta \quad (4.25)$$

die Lösung

$$y : \mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto \eta \cdot e^{a(x-\xi)}. \quad (4.26)$$

Für $a < 0$ handelt es sich um einen Zerfallsprozess, für $a > 0$ um einen Wachstumsprozess.

Für die lineare gewöhnliche Differentialgleichung

$$y'(x) + a(x)y(x) = f(x) \quad (4.27)$$

mit nichtverschwindender Störfunktion f erhält man für stetige Funktionen a, f eine vollständige allgemeine Lösung

$$y : I \rightarrow \mathbb{R}, \quad x \mapsto e^{-A(x)} \left(\int_{\xi}^x e^{A(t)} f(t) dt + c \right), \quad c \in \mathbb{R}, \quad (4.28)$$

wobei $A(x) := \int_{\xi}^x a(t)dt$ und $\xi \in I$ beliebig.

Es gilt somit für gewöhnliche Anfangswertprobleme der Form

$$y'(x) + a(x)y(x) = f(x), \quad y(\xi) = \eta : \quad (4.29)$$

- Es existiert eine eindeutige Lösung

$$y : I \rightarrow \mathbb{R}, \quad x \mapsto e^{-A(x)} \left(\int_{\xi}^x e^{A(t)} f(t) dt + \eta \right), \quad (4.30)$$

wobei $A(x) = \int_{\xi}^x a(t)dt$.

Die vollständige allgemeine Lösung y der inhomogenen gewöhnlichen Differentialgleichung

$$y'(x) + a(x)y(x) = f(x) \quad (4.31)$$

ist gegeben durch

$$y = y_h + y_p, \quad (4.32)$$

wobei y_p eine spezielle (partikuläre) Lösung der obigen gewöhnlichen Differentialgleichung ist (etwa eine Lösung zu fest gewählten Anfangswerten) und y_h die allgemeine Lösung der zugehörigen homogenen gewöhnlichen Differentialgleichung darstellt.

Beispiel(e) 4.8

- Betrachte

$$y'(x) + \frac{1}{x}y(x) = x^3, \quad x > 0: \quad (4.33)$$

$$\int \frac{1}{t} dt : (0, \infty) \rightarrow \mathbb{R}, \quad x \mapsto \ln(x) + c, \quad y_h : (0, \infty) \rightarrow \mathbb{R}, \quad x \mapsto e^{-\ln(x)-c} = \frac{\tilde{c}}{x}. \quad (4.34)$$

Mit $\xi = 1, \eta = 0$ erhält man:

$$y_p : I \rightarrow \mathbb{R}, \quad x \mapsto e^{-\ln(x)} \left(0 + \int_1^x e^{\ln(t)} t^3 dt \right) = \frac{x^4}{5} - \frac{1}{5x}. \quad (4.35)$$

Somit ist

$$y : I \rightarrow \mathbb{R}, \quad x \mapsto \frac{\tilde{c}}{x} + \frac{x^4}{5} \quad (4.36)$$

die allgemeine Lösung.

- RL-Stromkreis:

Wird in einem Stromkreis eine Spule der Induktivität L und ein Widerstand R in Serie geschaltet an eine Spannungsquelle $U(t)$ angeschlossen, so ergibt sich für die Stromstärke $I(t)$:

$$L \cdot \dot{I}(t) + R \cdot I(t) = U(t). \quad (4.37)$$

$I_h : [0, \infty) \rightarrow \mathbb{R}, t \mapsto c \cdot \exp\left(-\frac{R}{L}t\right)$ (exponentielles Abklingen der Stromstärke nach dem Abschalten der Spannungsquelle).

Allgemeine Lösung:

1. Fall: Gleichspannung $U : [0, \infty) \rightarrow \mathbb{R}, t \mapsto U_0$:

$$I : [0, \infty) \rightarrow \mathbb{R}, t \mapsto \frac{U_0}{R} - \left(\frac{U_0}{R} - I_0 \right) \exp\left(-\frac{R}{L}t\right), \quad (4.38)$$

wobei $I_p : [0, \infty) \rightarrow \mathbb{R}, t \mapsto \frac{U_0}{R}$ und $I(0) = I_0$.

2. Fall: Wechselspannung $U : [0, \infty) \rightarrow \mathbb{R}, t \mapsto U_0 \cos(\omega t)$:

$$I : [0, \infty) \rightarrow \mathbb{R}, t \mapsto c \cdot \exp\left(-\frac{R}{L}t\right) + \frac{U_0}{\sqrt{R^2 + \omega^2 L^2}} \cos\left(\omega t - \arctan\left(\frac{\omega L}{R}\right)\right), \quad (4.39)$$

wobei

$$I_p : [0, \infty) \rightarrow \mathbb{R}, t \mapsto \frac{U_0}{\sqrt{R^2 + \omega^2 L^2}} \cos\left(\omega t - \arctan\left(\frac{\omega L}{R}\right)\right). \quad (4.40)$$

In den Anwendungen spielen spezielle gewöhnliche Differentialgleichungen zweiter Ordnung eine wichtige Rolle. Beginnen wir zunächst mit der homogenen linearen Differentialgleichung zweiter Ordnung

$$y''(x) + ay'(x) + by(x) = 0, \quad a, b \in \mathbb{R}. \quad (4.41)$$

Der Ansatz $y : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto e^{\lambda x}$ liefert:

$$\lambda^2 e^{\lambda x} + a\lambda e^{\lambda x} + b e^{\lambda x} = 0, \quad (4.42)$$

also:

$$\lambda^2 + a\lambda + b = 0, \quad \lambda_{1,2} = \frac{-a \pm \sqrt{a^2 - 4b}}{2}. \quad (4.43)$$

1. Fall: λ_1, λ_2 reell und $\lambda_1 \neq \lambda_2$:

Die vollständige allgemeine Lösung lautet:

$$y : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto c_1 e^{\lambda_1 x} + c_2 e^{\lambda_2 x}. \quad (4.44)$$

2. Fall: λ_1, λ_2 komplex, also $\lambda_1 = \bar{\lambda}_2$:

Die vollständige allgemeine Lösung lautet mit $\lambda_1 = \alpha + \beta i$, $\beta \neq 0$:

$$\begin{aligned} y : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto c_1 \operatorname{Re}(e^{\lambda_1 x}) + c_2 \operatorname{Im}(e^{\lambda_1 x}) &= \\ &= c_1 e^{\alpha x} \cos(\beta x) + c_2 e^{\alpha x} \sin(\beta x) \\ &= c_1 \operatorname{Re}(e^{\lambda_2 x}) - c_2 \operatorname{Im}(e^{\lambda_2 x}). \end{aligned} \quad (4.45)$$

3. Fall: λ_1, λ_2 reell und $\lambda_1 = \lambda_2 = -\frac{\alpha}{2}$:

Ansatz: Wähle $y : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto c(x)e^{\lambda_1 x}$, so folgt:

$$y' : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto c'(x)e^{\lambda_1 x} + c(x)e^{\lambda_1 x} \lambda_1 \quad (4.46)$$

und

$$y'' : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto c''(x)e^{\lambda_1 x} + c'(x)e^{\lambda_1 x} \lambda_1 + c'(x)e^{\lambda_1 x} \lambda_1 + c(x)e^{\lambda_1 x} \lambda_1^2. \quad (4.47)$$

Einsetzen in die gegebene gewöhnliche Differentialgleichung liefert $c''(x) = 0$ und damit:

$$c : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto c_1 + c_2 x, \quad c_1, c_2 \in \mathbb{R} \quad (4.48)$$

sowie

$$y : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto c_1 e^{-\frac{\alpha x}{2}} + c_2 x e^{-\frac{\alpha x}{2}}. \quad (4.49)$$

Beispiel(e) 4.9

- $y''(x) - 4y'(x) + 4y(x) = 0$.

Wegen $\lambda_1 = \lambda_2 = 2$ ergibt sich

$$y : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto c_1 e^{2x} + c_2 x e^{2x}. \quad (4.50)$$

- $y''(x) - 6y'(x) + 34y(x) = 0$.

Wegen $\lambda_1 = 3 + 5i$, $\lambda_2 = 3 - 5i$ ergibt sich

$$y : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto c_1 e^{3x} \cos(5x) + c_2 e^{3x} \sin(5x). \quad (4.51)$$

Für die vollständige allgemeine Lösung der inhomogenen linearen gewöhnlichen Differentialgleichung

$$y''(x) + ay'(x) + by(x) = f(x) \quad (4.52)$$

mit konstanten Koeffizienten $a, b \in \mathbb{R}$ und der Störfunktion $f : I \rightarrow \mathbb{R}$, $I \subseteq \mathbb{R}$ ein Intervall, genügt es, die vollständige allgemeine Lösung y_h der homogenen gewöhnlichen Differentialgleichung mit einer partikulären Lösung y_p der inhomogenen gewöhnlichen Differentialgleichung zu kombinieren:

$$y = y_h + y_p. \quad (4.53)$$

Sei nun f stetig und $y_h : I \rightarrow \mathbb{R}, x \mapsto c_1 y_1(x) + c_2 y_2(x)$, wobei

$$\begin{aligned} y_1(x) &= e^{\lambda_1 x}, \quad y_2(x) = e^{\lambda_2 x} && \text{im 1. Fall (4.44),} \\ y_1(x) &= e^{\alpha x} \cos(\beta x), \quad y_2(x) = e^{\alpha x} \sin(\beta x) && \text{im 2. Fall (4.45) und} \\ y_1(x) &= e^{\lambda_1 x}, \quad y_2(x) = x e^{\lambda_1 x} && \text{im 3. Fall (4.49).} \end{aligned}$$

Es ist dann

$$y_1(x)y_2'(x) - y_2(x)y_1'(x) \neq 0 \quad \text{für alle } x \in I, \quad (4.54)$$

und man kann eine partikuläre Lösung der inhomogenen gewöhnlichen Differentialgleichung in der Form

$$y_p : I \rightarrow \mathbb{R}, x \mapsto -y_1(x) \int \frac{y_2(x)f(x)}{y_1(x)y_2'(x) - y_2(x)y_1'(x)} dx + y_2(x) \int \frac{y_1(x)f(x)}{y_1(x)y_2'(x) - y_2(x)y_1'(x)} dx \quad (4.55)$$

angeben. Diese Formel ergibt sich, wenn man den Ansatz

$$y_p : I \rightarrow \mathbb{R}, x \mapsto c_1(x)y_1(x) + c_2(x)y_2(x) \quad (4.56)$$

in die Differentialgleichung einsetzt und die zusätzliche Forderung $c_1'(x)y_1(x) + c_2'(x)y_2(x) = 0$ aufstellt. Man nennt diesen Ansatz „Variation der Konstanten“.

Beispiel(e) 4.10

$$y''(x) - 6y'(x) + 5y(x) = \ln(x), \quad x > 0. \quad (4.57)$$

$$y_h : (0, \infty) \rightarrow \mathbb{R}, x \mapsto c_1 e^x + c_2 e^{5x}, \quad (4.58)$$

$$\begin{aligned} y_p : (0, \infty) \rightarrow \mathbb{R}, x \mapsto & -e^x \int_{\xi}^x \frac{e^{5t} \ln(t)}{5e^t e^{5t} - e^{5t} e^t} dt + e^{5x} \int_{\xi}^x \frac{e^t \ln(t)}{5e^t e^{5t} - e^{5t} e^t} dt = \\ & = -\frac{1}{4} e^x \int_{\xi}^x e^{-t} \ln(t) dt + \frac{1}{4} e^{5x} \int_{\xi}^x e^{-5t} \ln(t) dt. \end{aligned} \quad (4.59)$$

Für spezielle rechte Seiten $f(x)$ der Differentialgleichung (4.52) gibt es spezielle Ansätze, mit denen sich eine partikuläre Lösung $y_p(x)$ oft einfacher berechnen lässt als mit (4.55). Diese Ansätze sind insbesondere möglich, wenn f ein Polynom, eine trigonometrische oder eine Exponentialfunktion ist. Da der Ansatz für y_p dann ebenfalls aus Polynomen, trigonometrischen oder Exponentialfunktionen besteht, spricht man von einem „Ansatz vom Typ der rechten Seite“.

4.2 Gewöhnliche Differentialgleichungssysteme

In vielen Anwendungen betrachtet man Funktionen $y_1, \dots, y_n : I \rightarrow \mathbb{R}$, deren Ableitungen $y_i'(x)$ neben $y_i(x)$ und x auch von $y_1(x), \dots, y_{i-1}(x), y_{i+1}(x), \dots, y_n(x)$ abhängen. Dies führt zu einer vektoriellen gewöhnlichen Differentialgleichung erster Ordnung:

$$(y_1'(x), \dots, y_n'(x))^{\top} = \mathbf{f}(x, y_1(x), \dots, y_n(x)), \quad (4.60)$$

wobei $\mathbf{f} : \mathbb{D} \rightarrow \mathbb{R}^n$, $\mathbb{D} \subseteq \mathbb{R}^{n+1}$.

Definition 4.11 (*n*-dim. gewöhnliches Differentialgleichungssystem)
 Seien $n \in \mathbb{N}$, $\mathbf{f} : \mathbb{D} \rightarrow \mathbb{R}^n$, $\mathbb{D} \subseteq \mathbb{R}^{n+1}$ und $I \subseteq \mathbb{R}$ ein Intervall, so heißt eine Funktion

$$\mathbf{y} : I \rightarrow \mathbb{R}^n, \quad x \mapsto \mathbf{y}(x) \tag{4.61}$$

Lösungskurve des *n*-dim. gewöhnlichen Differentialgleichungssystems

$$\mathbf{y}'(x) := (y_1'(x), \dots, y_n'(x))^\top = \mathbf{f}(x, \mathbf{y}(x)) := \mathbf{f}(x, y_1(x), \dots, y_n(x)), \tag{4.62}$$

falls $(x, \mathbf{y}(x)) \in \mathbb{D}$ und $\mathbf{y}'(x) = \mathbf{f}(x, \mathbf{y}(x))$ für alle $x \in I$.
 Ist $\mathbf{y}(x_0) = \mathbf{y}_0$, so ist \mathbf{y} eine Lösung des gewöhnlichen Anfangswertproblems

$$\mathbf{y}'(x) = \mathbf{f}(x, \mathbf{y}(x)), \quad \mathbf{y}(x_0) = \mathbf{y}_0. \tag{4.63}$$

Analog zum eindimensionalen Fall heißt \mathbf{y} allgemeine Lösung, falls \mathbf{y} *n* freie Parameter enthält. Eine allgemeine Lösung heißt vollständig, falls damit alle möglichen Lösungen erfasst werden.

Eine wichtige Klasse von gewöhnlichen Differentialgleichungssystemen erster Ordnung ist durch eindimensionale gewöhnliche Differentialgleichungen *n*-ter Ordnung

$$y^{(n)}(x) = a_{n-1}(x)y^{(n-1)}(x) + \dots + a_1(x)y'(x) + a_0(x)y(x) + b(x) \tag{4.64}$$

mit $a_{n-1}, \dots, a_0, b : I \rightarrow \mathbb{R}$ gegeben, denn es gilt mit

$$\mathbf{f} : \mathbb{D} \rightarrow \mathbb{R}^n, \quad (x, \mathbf{z}) \mapsto \begin{pmatrix} z_2 \\ z_3 \\ \vdots \\ z_n \\ a_{n-1}(x)z_n + \dots + a_1(x)z_2 + a_0(x)z_1 + b(x) \end{pmatrix} =$$

$$= \underbrace{\begin{pmatrix} 0 & 1 & \dots & \dots & 0 \\ 0 & 0 & \ddots & & 0 \\ \vdots & \vdots & \ddots & & \vdots \\ 0 & 0 & \ddots & \ddots & 1 \\ a_0(x) & a_1(x) & \dots & a_{n-2}(x) & a_{n-1}(x) \end{pmatrix}}_{M(x)} \mathbf{z} + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ b(x) \end{pmatrix}, \tag{4.65}$$

$$\mathbf{y}'(x) = M(x) \begin{pmatrix} y_1(x) \\ y_2(x) \\ \vdots \\ y_n(x) \end{pmatrix} + \begin{pmatrix} 0 \\ \vdots \\ 0 \\ b(x) \end{pmatrix} \quad \text{und} \quad \mathbf{y} : I \rightarrow \mathbb{R}^n, \quad x \mapsto \begin{pmatrix} y(x) \\ y'(x) \\ \vdots \\ y^{(n-1)}(x) \end{pmatrix} : \tag{4.66}$$

$$\mathbf{y}'(x) = M(x)\mathbf{y}(x) + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ b(x) \end{pmatrix} \tag{4.67}$$

\Leftrightarrow

$$y^{(n)}(x) = a_{n-1}(x)y^{(n-1)}(x) + \dots + a_1(x)y'(x) + a_0(x)y(x) + b(x). \tag{4.68}$$

Beispiel(e) 4.12

- Van der Pol-Differentialgleichung

$$y''(x) = (\alpha - \beta x^2)y'(x) - y(x) \quad (4.69)$$

beziehungsweise

$$\mathbf{y}'(x) = \begin{pmatrix} 0 & 1 \\ -1 & \alpha - \beta x^2 \end{pmatrix} \mathbf{y}(x). \quad (4.70)$$

Beschreibung der Änderung der Gittervorspannung in Triodenschaltungen.

- Räuber-Beute-Modell von Lotka-Volterra:

$$y_1'(x) = -(\alpha - \beta y_2(x))y_1(x) \quad (4.71)$$

$$y_2'(x) = (\gamma - \delta y_1(x))y_2(x). \quad (4.72)$$

Die Population „Räuber“ ($\hat{=} y_1$) lebt von der „Beute“ ($\hat{=} y_2$). Die Sterberate der Räuber ohne Beute ($y_2 \equiv 0$) ist α . Die Geburtsrate der Beute ohne Räuber ($y_1 \equiv 0$) ist γ .

Im Folgenden betrachten wir die Existenz und Eindeutigkeit der Lösung gewöhnlicher Differentialgleichungssysteme

Satz 4.13 (Existenz und Eindeutigkeit)

Sei $f : [a, b] \times \mathbb{R}^n \rightarrow \mathbb{R}^n$, $(x, \mathbf{z}) \mapsto f(x, \mathbf{z})$, eine Abbildung mit

$$|f(x, \mathbf{z}_1) - f(x, \mathbf{z}_2)| \leq L|\mathbf{z}_1 - \mathbf{z}_2| \quad \text{für } x \in [a, b]; \mathbf{z}_1, \mathbf{z}_2 \in \mathbb{R}^n \quad (4.73)$$

(Lipschitz-Bedingung), die in x stetig ist. Sei ferner $\eta \in \mathbb{R}^n$, $n \in \mathbb{N}$, so hat jedes Anfangswertproblem

$$\mathbf{y}'(x) = \mathbf{f}(x, \mathbf{y}(x)), \quad \mathbf{y}(x_0) = \eta \quad (4.74)$$

mit $x_0 \in (a, b)$ eine eindeutige Lösung $\mathbf{y} : [a, b] \rightarrow \mathbb{R}^n$.

Die wichtigsten Systeme gewöhnlicher Differentialgleichungen sind lineare gewöhnliche Differentialgleichungssysteme der Form:

$$\mathbf{y}'(x) = A(x)\mathbf{y}(x) + \tilde{\mathbf{b}}(x) \quad (4.75)$$

mit der Koeffizientenmatrix $A(x) \in \mathbb{R}^{n,n}$ und einer Störfunktion $\tilde{\mathbf{b}}$. Im Folgenden nehmen wir an, dass $A : I \rightarrow \mathbb{R}^{n,n}$ und $\tilde{\mathbf{b}} : I \rightarrow \mathbb{R}^n$ über einem Intervall $I \subseteq \mathbb{R}$ definiert sind. Für $\tilde{\mathbf{b}} \equiv 0$ auf I spricht man von einem homogenen linearen gewöhnlichen Differentialgleichungssystem.

Beispiel(e) 4.14

$$\mathbf{y}'(x) = \begin{pmatrix} 3 & 2 & x^2 \\ \sin(x) & 0 & 5 \\ x & x^2 & x^3 \end{pmatrix} \mathbf{y}(x) + \begin{pmatrix} \cos(\omega x) \\ 0 \\ \sqrt{x} \end{pmatrix}. \quad (4.76)$$

Für lineare Systeme gewöhnlicher Differentialgleichungen gelten wichtige Eigenschaften:

- Aus

$$\mathbf{y}'(x) = A(x)\mathbf{y}(x) + \tilde{\mathbf{b}}_1(x) \quad (4.77)$$

$$\mathbf{w}'(x) = A(x)\mathbf{w}(x) + \tilde{\mathbf{b}}_2(x) \quad (4.78)$$

folgt:

$$\mathbf{v} := \alpha\mathbf{y} + \beta\mathbf{w} \quad (4.79)$$

ist für $\alpha, \beta \in \mathbb{R}$ die Lösung von

$$\mathbf{v}'(x) = A(x)\mathbf{v}(x) + \alpha\tilde{\mathbf{b}}_1(x) + \beta\tilde{\mathbf{b}}_2(x). \quad (4.80)$$

- $\mathbf{y} = \mathbf{y}_h + \mathbf{y}_p$.
- $\{\mathbf{y}(x) : I \rightarrow \mathbb{R}^n; \mathbf{y}'(x) = A(x)\mathbf{y}(x)\}$ ist ein Vektorraum über \mathbb{R} der Dimension n .
- Sind n Lösungen $\mathbf{y}_1, \dots, \mathbf{y}_n$ des homogenen linearen Differentialgleichungssystems

$$\mathbf{y}'(x) = A(x)\mathbf{y}(x) \quad (4.81)$$

gefunden, die in einem Punkt x_0 linear unabhängig sind (d.h. die Vektoren $\mathbf{y}_1(x_0), \dots, \mathbf{y}_n(x_0)$ sind linear unabhängig), dann sind die Vektoren $\mathbf{y}_1(x), \dots, \mathbf{y}_n(x)$ in jedem Punkt $x \in I$ linear unabhängig. Man nennt diese n Lösungen dann ein **Fundamentalsystem** und kann zeigen, dass ein Fundamentalsystem stets existieren muss.

- Hat man ein Fundamentalsystem, so ist mit $Y : I \rightarrow \mathbb{R}^{n,n}$, $x \mapsto (\mathbf{y}_1(x) \dots \mathbf{y}_n(x)) \in \mathbb{R}^{n,n}$ durch

$$\mathbf{y} : I \rightarrow \mathbb{R}^n, x \mapsto Y(x) \left(\int_{\xi}^x Y^{-1}(t)\tilde{\mathbf{b}}(t)dt + c \right), \quad c \in \mathbb{R}^n \quad (4.82)$$

eine vollständige allgemeine Lösung des inhomogenen gewöhnlichen Differentialgleichungssystems gegeben.

- Ist speziell $A(x) \equiv A$ unabhängig von x , so erhält man mit

$$e^{\bullet A} : I \rightarrow \mathbb{R}^{n,n}, x \mapsto \sum_{k=0}^{\infty} \frac{x^k}{k!} A^k = E + xA + \frac{x^2}{2} A \cdot A + \dots \quad (4.83)$$

durch

$$\mathbf{y}(x) = e^{xA}c, \quad c \in \mathbb{R}^n \quad (4.84)$$

die vollständige allgemeine Lösung des homogenen Systems linearer gewöhnlicher Differentialgleichungen

$$\mathbf{y}'(x) = A\mathbf{y}(x). \quad (4.85)$$

Für von x abhängige Matrizen $A(x)$ ist die Lösung nur in Spezialfällen möglich.

4.3 Lineare Differentialgleichungssysteme mit konstanten Koeffizienten

Es wird noch einmal der in der Praxis häufig auftretende Fall des linearen DGL-Systems mit konstanten Koeffizienten

$$\mathbf{y}'(x) = A\mathbf{y}(x) + \mathbf{b}(x), \quad A \in \mathbb{R}^{n,n}, \quad (4.86)$$

betrachtet. Anfangswertprobleme zu solchen Systemen kann man günstig mit der Laplace-Transformation berechnen – siehe das spätere Kapitel über Transformationen.

Die allgemeine Lösung von (4.86) hat die Form $\mathbf{y} = \mathbf{y}_h + \mathbf{y}_p$ mit einer partikulären Lösung \mathbf{y}_p von (4.86) und der allgemeinen Lösung \mathbf{y}_h des homogenen Systems

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

die sich theoretisch wie in (4.84) mithilfe der Matrix-Exponentialfunktion angeben lässt. Praktischer ist folgender Zugang: Wenn \mathbf{v} ein Eigenvektor von A zum Eigenwert λ ist, dann ist die Funktion $\mathbf{y}(x) = \exp(\lambda x)\mathbf{v}$ eine Lösung des homogenen Systems, denn

$$\mathbf{y}'(x) = \lambda e^{\lambda x}\mathbf{v} \quad \text{und} \quad A\mathbf{y}(x) = e^{\lambda x}A\mathbf{v} = \lambda e^{\lambda x}\mathbf{v}.$$

Als Nullstellen des charakteristischen Polynoms können Eigenwerte und damit auch Eigenvektoren komplex sein. Solange jedoch die Matrix A reellwertig ist, erhält man durch Übergang zum konjugiert Komplexen:

$$A\mathbf{v} = \lambda\mathbf{v} \quad \iff \quad A\bar{\mathbf{v}} = \bar{\lambda}\bar{\mathbf{v}},$$

das heißt mit λ und \mathbf{v} ist stets auch $\bar{\lambda}$ ein Eigenwert und $\bar{\mathbf{v}}$ ein Eigenvektor zu A . Mit der komplexen Lösung $\exp(\lambda x)\mathbf{v}$ erhält man die beiden linear unabhängigen reellen Lösungen $\operatorname{Re}(\exp(\lambda x)\mathbf{v})$ und $\operatorname{Im}(\exp(\lambda x)\mathbf{v})$.

Kann man n linear unabhängige Eigenvektoren $\mathbf{v}_1, \dots, \mathbf{v}_n$ zu Eigenwerten $\lambda_1, \dots, \lambda_n$ (die nicht notwendig unterschiedlich sein müssen) finden, so hat man auf diese Art die allgemeine homogene Lösung in der Form

$$\mathbf{y}_h(x) = c_1 e^{\lambda_1 x} \mathbf{v}_1 + \dots + c_n e^{\lambda_n x} \mathbf{v}_n$$

mit beliebigen Konstanten c_1, \dots, c_n gefunden. In manchen Fällen ist das möglich, so zum Beispiel bei symmetrischen Matrizen, bei denen sogar immer garantiert ist, dass alle Eigenwerte reellwertig sind. Aber auch bei komplexen Eigenwerten lassen sich durch Übergang zu Real- und Imaginärteil immer reelle Lösungen finden.

Beispiel(e) 4.15

$$\mathbf{y}'(x) = \begin{pmatrix} 0 & 2 & 0 \\ 0 & 0 & 2 \\ -1 & 1 & 0 \end{pmatrix} \mathbf{y}(x). \quad (4.87)$$

Hier ergeben sich die Eigenwerte aus

$$\det(A - \lambda E) = -\lambda^3 + 2\lambda - 4 = 0 \implies \lambda = -2, \lambda = 1 \pm i$$

und damit die Eigenvektoren

$$\lambda = -2 : (A + 2E)\mathbf{v} = 0 \implies \mathbf{v} = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}$$

sowie

$$\lambda = 1 + i : (A - (1 + i)E)\mathbf{v} = 0 \implies \mathbf{v} = \begin{pmatrix} 2 - 2i \\ 2 \\ 1 + i \end{pmatrix}.$$

Drei linear unabhängige komplexe Lösungen sind damit $\mathbf{y}_1(x)$, $\mathbf{y}_2(x)$ und $\mathbf{y}_3(x)$ mit $\mathbf{y}_3(x) = \bar{\mathbf{y}}_2(x)$ und

$$\mathbf{y}_1(x) = e^{-2x} \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \quad \mathbf{y}_2(x) = e^{(1+i)x} \begin{pmatrix} 2 - 2i \\ 2 \\ 1 + i \end{pmatrix}.$$

Drei linear unabhängige reelle Lösungen sind \mathbf{y}_1 , $\tilde{\mathbf{y}}_2 = \operatorname{Re} \mathbf{y}_2$ und $\tilde{\mathbf{y}}_3 = \operatorname{Im} \mathbf{y}_2$, also

$$\tilde{\mathbf{y}}_2(x) = e^x \left[\cos x \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix} - \sin x \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix} \right], \quad \tilde{\mathbf{y}}_3(x) = e^x \left[\sin x \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix} + \cos x \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix} \right].$$

Es gibt durchaus Matrizen $A \in \mathbb{R}^{n,n}$, die weniger als n linear unabhängige Eigenvektoren besitzen, etwa

$$A = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix},$$

die den dreifachen Eigenwert $\lambda = 1$ besitzt, dessen Eigenraum

$$\operatorname{Kern}(A - E) = \left\{ c \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}; c \in \mathbb{R} \right\}$$

jedoch nur die Dimension 1 besitzt. Auch in solchen Fällen ist es jedoch möglich, die allgemeine Lösung von $y' = Ay$ anzugeben. Man benötigt dafür die folgende Verallgemeinerung von Eigenvektoren.

Definition 4.16 (Hauptvektoren)

Ein Vektor $\mathbf{v} \in \mathbb{C}^n$ heißt **Hauptvektor der Stufe** $\ell \in \mathbb{N}$ zum Eigenwert $\lambda \in \mathbb{C}$ von $A \in \mathbb{R}^{n,n}$, wenn

$$(A - \lambda E)^\ell \mathbf{v} = 0 \quad \text{und} \quad (A - \lambda E)^{\ell-1} \mathbf{v} \neq 0.$$

Beispiel(e) 4.17

(a): Jeder Eigenvektor \mathbf{v} von A ist ein Hauptvektor der Stufe 1, denn

$$(A - \lambda E)\mathbf{v} = 0 \quad \text{und} \quad (A - \lambda E)^0 \mathbf{v} = \mathbf{v} \neq 0.$$

(b): Die Matrix

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

hat den dreifachen Eigenwert $\lambda = 1$. Der Vektor e_1 ist Eigenvektor (und damit Hauptvektor der Stufe 1). Wegen

$$(A - E)e_2 = e_1 \quad \implies \quad (A - E)^2 e_2 = 0$$

ist dann e_2 Hauptvektor der Stufe 2. Ebenso ist e_3 wegen $(A - E)e_3 = e_1 + e_2$ Hauptvektor der Stufe 3.

(c): Allgemein sind für einen Hauptvektor \mathbf{v} der Stufe ℓ zum Eigenwert λ die Vektoren

$$\mathbf{v}, \quad (A - \lambda E)\mathbf{v}, \quad (A - \lambda E)^2 \mathbf{v}, \quad \dots, \quad (A - \lambda E)^{\ell-1} \mathbf{v}$$

linear unabhängige Hauptvektoren der Stufen $\ell, \ell - 1, \dots, 1$.

Wenn es auch nicht zu jeder Matrix $A \in \mathbb{R}^{n,n}$ n linear unabhängige Eigenvektoren gibt, so gilt dennoch

Satz 4.18 (Hauptvektor-Basis)

Zu jedem k -fachen Eigenwert λ von $A \in \mathbb{R}^{n,n}$ gibt es k linear unabhängige Hauptvektoren. Die Hauptvektoren zu verschiedenen Eigenwerten sind linear unabhängig.

Damit gibt es zu jeder Matrix $A \in \mathbb{R}^{n,n}$ n linear unabhängige Hauptvektoren.

Beispiel(e) 4.19

Die Matrix

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

hat Eigenwerte

$$\det(A - \lambda E) = -\lambda^3 + 4\lambda^2 - 5\lambda + 2 = 0 \implies \lambda = 2, \lambda = 1 \quad (\text{doppelt}).$$

Zum ersten Eigenwert ergibt sich der Eigenvektor

$$\lambda = 2 : (A - 2E)\mathbf{v} = 0 \implies \mathbf{v}_1 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

und zum zweiten Eigenwert der Eigenvektor

$$\lambda = 1 : (A - E)\mathbf{v} = 0 \implies \mathbf{v}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}.$$

Es gibt keinen zu \mathbf{v}_2 linear unabhängigen Eigenvektor zu A , jedoch bekommt man

$$(A - E)^2\mathbf{v} = \begin{pmatrix} 0 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & -1 & 0 \end{pmatrix} \mathbf{v} = 0 \implies \mathbf{v}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \mathbf{v}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

 \mathbf{v}_3 ist ein Hauptvektor der Stufe 2 und $\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}$ ist eine Basis aus Hauptvektoren.

In Verallgemeinerung des Sachverhalts bei Eigenvektoren gilt nun der

Satz 4.20 (Allgemeine Lösung linearer DGL-Systeme)Es sei \mathbf{v} ein Hauptvektor der Stufe ℓ zum Eigenwert λ der Matrix $A \in \mathbb{R}^{n,n}$. Dann ist

$$\mathbf{y}_v(x) := e^{\lambda x} \left[\mathbf{v} + x(A - \lambda E)\mathbf{v} + \frac{x^2}{2!}(A - \lambda E)^2\mathbf{v} + \dots + \frac{x^{\ell-1}}{(\ell-1)!}(A - \lambda E)^{\ell-1}\mathbf{v} \right]$$

eine Lösung des homogenen Differentialgleichungssystems $\mathbf{y}'(x) = A\mathbf{y}(x)$.Mit einer nach Satz 4.18 existierenden Basis $\{\mathbf{v}_1, \dots, \mathbf{v}_n\}$ aus Hauptvektoren von A ist dann

$$\mathbf{y}_h(x) = c_1\mathbf{y}_{v_1}(x) + \dots + c_n\mathbf{y}_{v_n}(x),$$

die allgemeine Lösung des homogenen Systems.

Beispiel(e) 4.21

Im Fall von Beispiel 4.19 bekommt man die Lösungsbasis $\{\mathbf{y}_{v_1}, \mathbf{y}_{v_2}, \mathbf{y}_{v_3}\}$ mit

$$\mathbf{y}_{v_1}(x) = e^{2x} \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \mathbf{y}_{v_2}(x) = e^x \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \mathbf{y}_{v_3}(x) = e^x \left[\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + x \begin{pmatrix} -1 \\ -1 \\ 0 \end{pmatrix} \right].$$

4.4 Stabilität

In diesem Abschnitt betrachten wir sogenannte **autonome Differentialgleichungssysteme**. Das sind spezielle Systeme der Bauart

$$\mathbf{y}'(x) = \mathbf{f}(\mathbf{y}(x))$$

mit $\mathbf{f}: D \rightarrow \mathbb{R}^n$ und $D \subseteq \mathbb{R}^n$. Im Unterschied zum allgemeinen Fall hängt die rechte Seite nicht explizit von x ab. Die Interpretation hiervon wird deutlicher, wenn man die unabhängige Variable nicht mit x , sondern mit t („Zeit“) bezeichnet, also zu

$$\dot{\mathbf{y}}(t) = \mathbf{f}(\mathbf{y}(t)) \quad (4.88)$$

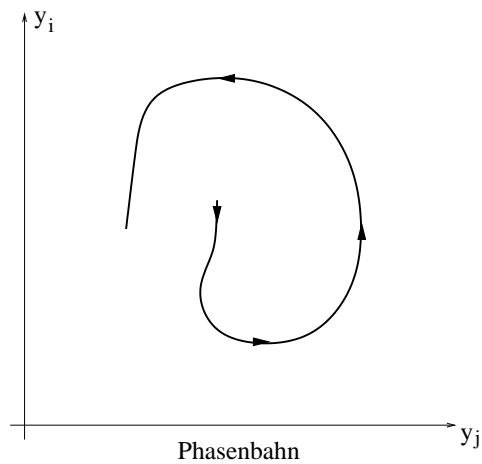
übergeht mit der in der Physik üblichen Konvention, die Ableitung nach der Zeit mit einem Punkt zu kennzeichnen. $\mathbf{y}(t)$ ist der Zustand eines Systems zum Zeitpunkt t . Zum Beispiel könnte man die Flugbahn einer Raumkapsel modellieren, die durch die Komponenten **Tangentialgeschwindigkeit** $v(t)$, **Bahnneigungswinkel** $\gamma(t)$ und **Flughöhe** $h(t)$ gekennzeichnet ist, so dass in diesem Fall

$$\mathbf{y}(t) = \begin{pmatrix} y_1(t) \\ y_2(t) \\ y_3(t) \end{pmatrix} = \begin{pmatrix} v(t) \\ \gamma(t) \\ h(t) \end{pmatrix}.$$

Der Zustand des Systems ändert sich mit der Zeit und die spezielle Interpretation des Differentialgleichungssystems (4.88) ist, dass Zustandsänderungen $\dot{\mathbf{y}}(t)$ nur vom augenblicklichen Zustand $\mathbf{y}(t)$ selbst abhängen, aber nicht explizit vom Zeitpunkt, wann dieser Zustand angenommen wird. Dies ist ein in der Steuerungs- und Regelungstechnik wichtiger Fall.

Die Menge der Zustände, die eine Lösung $\mathbf{y}(t)$ von (4.88) ausgehend von einem bestimmten Startzustand $\mathbf{y}(t_0) = \mathbf{y}_0$ annehmen kann, nennt man eine **Phasenbahn**. Interessiert man sich nur für die Zustände selbst, aber nicht dafür, wann sie angenommen werden, dann ist es oft aufschlussreicher, Phasenbahnen zu skizzieren als Lösungskurven.

Schematisch sieht das so aus:



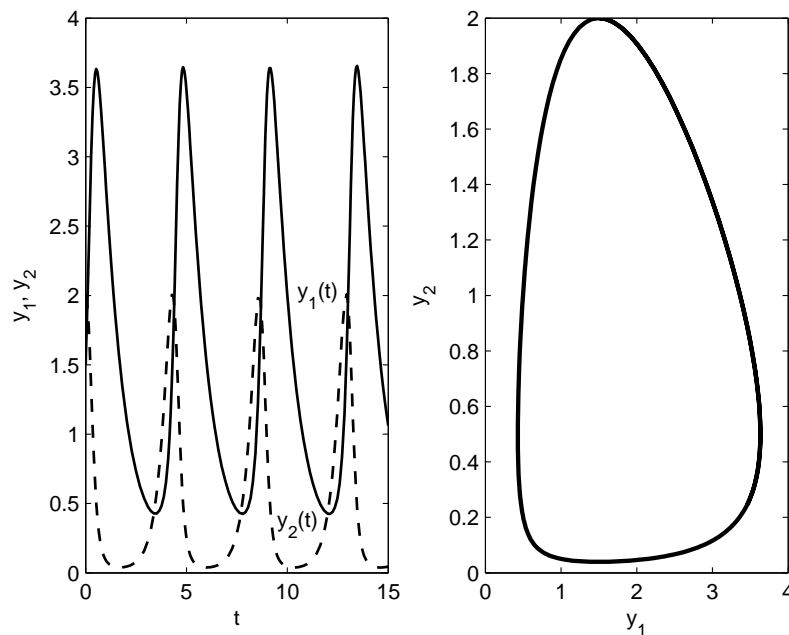
Mit den Pfeilen wird angedeutet, in welcher Richtung die Phasenbahn durchlaufen wird.

Beispiel(e) 4.22

Das schon betrachtete Räuber-Beute-Modell ist ein autonomes DGL-System:

$$\begin{aligned} \dot{y}_1(t) &= -[\alpha - \beta y_2(t)] y_1(t) \\ \dot{y}_2(t) &= +[\gamma - \delta y_1(t)] y_2(t) \end{aligned} \tag{4.89}$$

Zur Erinnerung: $y_1(t)$ ist die Anzahl von Individuen einer Population von Raubtieren und $y_2(t)$ ist die Anzahl von Individuen einer Population von Beutetieren, jeweils zur Zeit t . In der folgenden Graphik sind links die Graphen der Lösungen y_1 und y_2 als Funktionen der Zeit bei Wahl der Parameter $\alpha = 1$, $\beta = 2$, $\gamma = 3$ und $\delta = 2$ und für die Wahl $y_1(0) = 3/2$ und $y_2(0) = 2$ von Anfangswerten gezeichnet. Rechts daneben sieht man die entsprechende Phasenbahn, in diesem Fall eine geschlossene Kurve, welche rechts herum durchlaufen wird.



Für andere Anfangswerte ergäben sich andere Lösungen und eine andere Phasenbahn.

Beispiel(e) 4.23

Wir betrachten das homogene System

$$\dot{\mathbf{y}}(t) = \begin{pmatrix} -4 & 2 \\ -3 & 1 \end{pmatrix} \mathbf{y}(t).$$

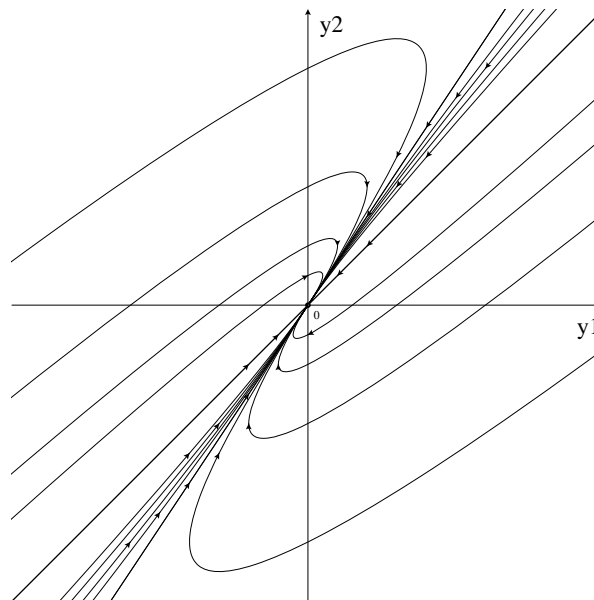
A hat Eigenwerte $\lambda_1 = -2$ und $\lambda_2 = -1$ mit Eigenvektoren

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \quad \text{und} \quad \mathbf{v}_2 = \begin{pmatrix} 2 \\ 3 \end{pmatrix}.$$

Die allgemeine Lösung ist dann

$$\mathbf{y}(t) = c_1 e^{-2t} \begin{pmatrix} 1 \\ 1 \end{pmatrix} + c_2 e^{-t} \begin{pmatrix} 2 \\ 3 \end{pmatrix}, \quad c_1, c_2 \in \mathbb{R}.$$

Wir skizzieren einige Phasenbahnen (zu jeder einzelnen gehört eine spezielle Wahl von c_1 und c_2):



Man beobachtet und erkennt auch an der allgemeinen Lösung, dass hier sämtliche Phasenbahnen in die spezielle Lösung $\mathbf{y}(t) \equiv 0$ einmünden.

Konstante Lösungen wie im letzten Beispiel die Nulllösung $\mathbf{y}(t) = 0$ bekommen einen speziellen Namen.

Definition 4.24 (Stationäre Lösungen)

Eine konstante Funktion $\mathbf{y}(t) = \mathbf{a} \in \mathbb{R}^n$ heißt **stationäre Lösung** oder **Gleichgewichtslösung** (GGL) von (4.88), wenn $\mathbf{f}(\mathbf{a}) = 0$.

GGL entsprechen Zuständen, in denen sich ein System nicht mehr ändert. Oft interessiert man sich für das sogenannte Stabilitätsverhalten eines Systems „in der Nähe“ einer GGL: kann das

System „ausbrechen“ oder bleibt es für alle Zeiten „stabil“ in der Nähe der GGL (oder – wie im letzten Beispiel – konvergiert es sogar gegen die GGL)? Dazu die folgende

Definition 4.25 (Stabilität)

Eine GGL \mathbf{a} von $\dot{\mathbf{y}}(t) = \mathbf{f}(\mathbf{y})$ heißt

(a): **stabil**, wenn es für jedes $\varepsilon > 0$ ein $\delta > 0$ gibt, so dass für jede Lösung gilt

$$|\mathbf{y}(0) - \mathbf{a}| < \delta \implies |\mathbf{y}(t) - \mathbf{a}| < \varepsilon \quad \text{für alle } t > 0,$$

(b): **asymptotisch stabil**, wenn ein $\delta > 0$ existiert, so dass für jede Lösung gilt

$$|\mathbf{y}(0) - \mathbf{a}| < \delta \implies \lim_{t \rightarrow \infty} |\mathbf{y}(t) - \mathbf{a}| = 0,$$

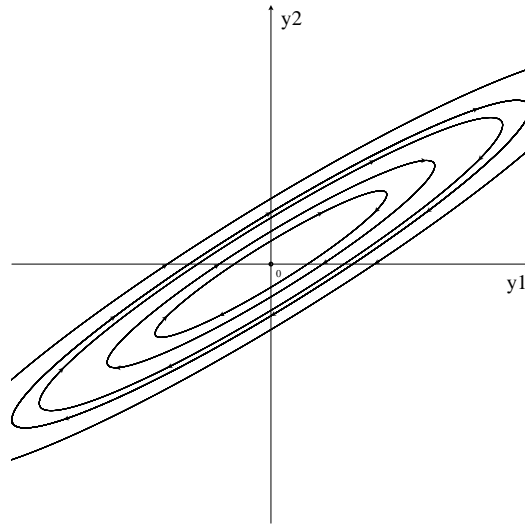
(c): **instabil**, wenn sie nicht stabil ist.

Beispiel(e) 4.26

Das Beispiel (4.23) zeigt eine asymptotisch stabile GGL. Beim homogenen System

$$\dot{\mathbf{y}}(t) = \begin{pmatrix} -3 & 5 \\ -2 & 3 \end{pmatrix} \mathbf{y}(t)$$

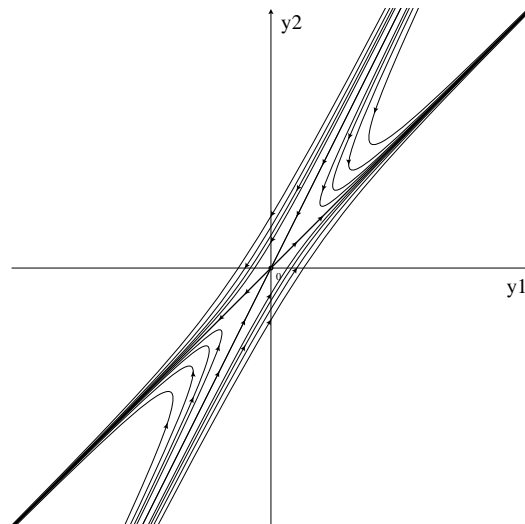
lauten die Eigenwerte von A $\lambda_{1,2} = \pm i$ und wir haben folgende Phasenbahnen



Die GGL $\mathbf{y}(t) = 0$ ist in diesem Fall stabil. Die GGL $\mathbf{y}(t) = 0$ ist instabil im Fall von

$$\dot{\mathbf{y}}(t) = \begin{pmatrix} 4 & -3 \\ 6 & -5 \end{pmatrix} \mathbf{y}(t)$$

in dem A die Eigenwerte $\lambda_1 = 1$ und $\lambda_2 = -2$ hat. Wir skizzieren einige Phasenbahnen:



Entscheidend für die Stabilität in den gezeigten Beispielen sind allein die Eigenwerte der Matrix A . Dass das immer so ist, zeigt der folgende

Satz 4.27 (Stabilitätssatz für lineare Systeme)

Es seien $\lambda_1, \dots, \lambda_n$ die Eigenwerte der Matrix $A \in \mathbb{R}^{n,n}$ des linearen DGL-Systems $\dot{\mathbf{y}} = A\mathbf{y}$.

- (a): Genau dann ist eine GGL asymptotisch stabil, wenn $\operatorname{Re}(\lambda_k) < 0$ für alle k .
- (b): Wenn $\operatorname{Re}(\lambda_k) > 0$ für ein k , dann ist eine GGL instabil.
- (c): Genau dann ist eine GGL stabil, wenn erstens $\operatorname{Re}(\lambda_k) \leq 0$ für alle k und wenn zweitens zu jedem Eigenwert λ_k mit $\operatorname{Re}(\lambda_k) = 0$ nur Eigenvektoren, aber keine Hauptvektoren der Stufe 2 oder höher existieren.

Der **Beweis** dieses Satzes ergibt sich fast unmittelbar aus Satz 4.20.

Bemerkung. Das inhomogene lineare DGL-System mit konstanten Koeffizienten $\dot{\mathbf{y}} = A\mathbf{y} + \mathbf{b}$ hat als GGL die Lösungen \mathbf{a} des linearen Gleichungssystems $A\mathbf{a} + \mathbf{b} = 0$. Wenn eine GGL \mathbf{a} existiert, dann kann man das DGL-System umschreiben in der Form

$$\dot{\mathbf{y}} = A(\mathbf{y} - \mathbf{a}) \iff \dot{\mathbf{z}} = A\mathbf{z} \quad \text{mit} \quad \mathbf{z} = \mathbf{y} - \mathbf{a}.$$

Folglich wird auch in diesem Fall die Art der GGL unabhängig von \mathbf{b} von den Eigenwerten von A bestimmt.

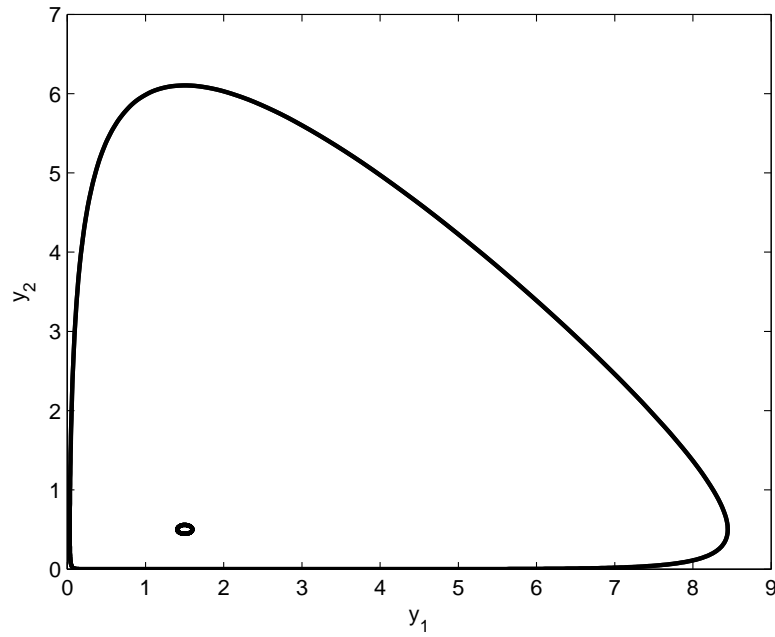
Den Fall nichtlinearer autonomer DGL-Systeme versucht man durch „Linearisierung“ auf den linearen Fall zurückzuspielen. Dazu benötigt man die Ableitung der vektorwertigen Funktion f in den Gleichgewichtslagen \mathbf{a} : diese Ableitung hat die Form einer Matrix A , was aber erst in der Vorlesung „Mathematik 3“ behandelt wird. Wenn A nur Eigenwerte mit negativem Realteil hat, dann ist \mathbf{a} eine asymptotisch stabile GGL, tritt ein Eigenwert mit positivem Realteil auf, dann ist \mathbf{a} instabil. Im nichtlinearen Fall können verschiedene GGL desselben DGL-Systems durchaus unterschiedliches Stabilitätsverhalten aufweisen.

Beispiel(e) 4.28

Das Räuber-Beute-Modell aus Beispiel 4.22 ist nichtlinear. Es gibt zwei GGL, nämlich

$$\mathbf{a}_1 = (0, 0) \quad \text{und} \quad \mathbf{a}_2 = (\gamma/\delta, \alpha/\beta).$$

Die folgende Graphik zeigt zwei Phasenbahnen, welche in der Nähe einer GGL starten (Parameter $\alpha = 1$, $\beta = 2$, $\gamma = 3$ und $\delta = 2$).



Es zeigt sich, dass die GGL $(0, 0)$ instabil, die GGL $(\gamma/\delta, \alpha/\beta)$ stabil ist.

Kapitel 5

Transformationen

In diesem Kapitel haben wir es häufig mit „komplexwertigen Funktionen mit reellen Argumenten“ zu tun. Dabei handelt es sich um Folgendes. Es sei $I \subseteq \mathbb{R}$ ein reelles Intervall und $f : I \rightarrow \mathbb{C}$ eine Funktion mit komplexen Funktionswerten (aber reellen Argumenten). Dann ist für jedes $x \in I$ die Zahl $f(x) \in \mathbb{C}$ zerlegbar in Real- und Imaginärteil:

$$f(x) = u(x) + iv(x) \quad \text{mit} \quad u(x) = \operatorname{Re} f(x) \quad \text{und} \quad v(x) = \operatorname{Im} f(x).$$

Die komplexwertige Funktion f definiert also zwei reellwertige Funktion $u : I \rightarrow \mathbb{R}$ und $v : I \rightarrow \mathbb{R}$ mit $f(x) = u(x) + iv(x)$ für alle $x \in I$. Mit der üblichen Identifikation von \mathbb{C} mit \mathbb{R}^2 können wir uns komplexwertige Funktionen mit reellen Argumenten also immer als Kurven $I \rightarrow \mathbb{R}^2$ vorstellen:

$$f : I \rightarrow \mathbb{C}, \quad f(x) = u(x) + iv(x) = \begin{pmatrix} u(x) \\ v(x) \end{pmatrix}.$$

Gemäß dieser Vorstellung definieren wir:

- f ist stetig in $x_0 \in I$, wenn u und v in x_0 stetig sind
- f ist differenzierbar in $x_0 \in I$, wenn u und v in x_0 differenzierbar sind. Für die Ableitung gilt $f'(x_0) = u'(x_0) + iv'(x_0)$.
- f ist Riemann-integrierbar auf $[a, b] \subseteq I$, wenn u und v es sind. Dann ist $\int_a^b f(x) dx = \int_a^b u(x) dx + i \int_a^b v(x) dx$.

5.1 Die Laplace-Transformation

Für $\alpha > 0$ sei Φ_α die Menge aller Funktionen $f : [0, \infty) \rightarrow \mathbb{R}$ mit:

- f ist stückweise stetig (unstetig nur an endlich vielen Punkten, in denen jedoch die einseitigen Grenzwerte existieren),
- es existiert ein $C > 0$ mit

$$|f(x)| \leq Ce^{\alpha x}, \quad x \geq 0. \quad (5.1)$$

Jeder Funktion $f \in \Phi_\alpha$ wird eine Funktion

$$F : (\alpha, \infty) \rightarrow \mathbb{R}, \quad s \mapsto \int_0^\infty e^{-st} f(t) dt \quad (5.2)$$

zugeordnet. Dank der obigen zwei Bedingungen ist diese Zuordnung möglich. Die Zuordnung $f \mapsto F$ heißt Laplace-Transformation und wird mit \mathcal{L} bezeichnet, also

$$\mathcal{L}(f) = F. \tag{5.3}$$

Rechenregeln:

- Linearität:

$$\mathcal{L}(\lambda f + \mu g) = \lambda \mathcal{L}(f) + \mu \mathcal{L}(g), \quad \lambda, \mu \in \mathbb{R}, f, g \in \Phi_\alpha. \tag{5.4}$$

-

$$\mathcal{L}(f^{(n)})(s) = s^n F(s) - s^{n-1} \lim_{x \rightarrow 0} f(x) - s^{n-2} \lim_{x \rightarrow 0} f'(x) - \dots \tag{5.5}$$

$$- s \lim_{x \rightarrow 0} f^{(n-2)}(x) - \lim_{x \rightarrow 0} f^{(n-1)}(x), \tag{5.6}$$

wobei $f^{(n)} \in \Phi_\alpha$.

-

$$\mathcal{L}(g)(s) = (-1)^n F^{(n)}(s) \quad \text{für } g : [0, \infty) \rightarrow \mathbb{R}, \quad x \mapsto x^n f(x). \tag{5.7}$$

-

$$\mathcal{L}(g)(s) = \frac{1}{a} F\left(\frac{s}{a}\right) \quad \text{für } g : [0, \infty) \rightarrow \mathbb{R}, \quad x \mapsto f(ax), \quad a > 0. \tag{5.8}$$

-

$$\mathcal{L}(g)(s) = F(s + b) \quad \text{für } g : [0, \infty) \rightarrow \mathbb{R}, \quad x \mapsto e^{-bx} f(x), \quad b \in \mathbb{R}. \tag{5.9}$$

- Sei

$$f_1 * f_2 : [0, \infty) \rightarrow \mathbb{R}, \quad x \mapsto \int_0^x f_1(\tau) f_2(x - \tau) d\tau \quad (\text{Faltung}) \tag{5.10}$$

und seien $\mathcal{L}(f_1) = F_1, \mathcal{L}(f_2) = F_2$, so gilt:

$$\mathcal{L}(f_1 * f_2) = F_1 \cdot F_2. \tag{5.11}$$

Die Laplace-Transformation läßt sich auch umkehren: sind $f, g \in \Phi_\alpha$ mit $\mathcal{L}(f) = \mathcal{L}(g)$, dann gilt $f(x) = g(x)$ an allen Stellen x , wo f und g stetig sind.

Anwendung: Gewöhnliche Differentialgleichungen:

$$y''(x) + 4y(x) = \sin(\omega x), \quad \omega > 0, \quad y(0) = 1, \quad y'(0) = 2. \tag{5.12}$$

Die Laplace-Transformation angewendet auf diese Gleichung ergibt:

$$s^2 Y(s) - sy(0) - y'(0) + 4Y(s) = \mathcal{L}(f) \quad \text{mit } f : \mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto \sin(\omega x), \tag{5.13}$$

also (mit Tabellen für die Laplace-Transformation):

$$s^2 Y(s) - sy(0) - y'(0) + 4Y(s) = \frac{\omega}{s^2 + \omega^2}. \tag{5.14}$$

Somit folgt:

$$Y(s) = \left(\frac{\omega}{s^2 + \omega^2} + s + 2 \right) \frac{1}{s^2 + 4} = \tag{5.15}$$

$$= \frac{\omega}{(s^2 + \omega^2)(s^2 + 4)} + \frac{s}{s^2 + 4} + \frac{2}{s^2 + 4}. \tag{5.16}$$

Rücktransformation mit entsprechender Tabelle:

$$y(x) = \cos(2x) + \sin(2x) + \begin{cases} \frac{1}{2(\omega^2 - 4)} (\omega \sin(2x) - 2 \sin(\omega x)) & \omega \neq 2 \\ \frac{1}{8} (\sin(2x) - 2x \cos(2x)) & \omega = 2 \end{cases}. \tag{5.17}$$

5.2 Die Fourier-Transformation

Vor der Fourier-Transformation betrachten wir Fourier-Reihen, welche eine Darstellung periodischer, integrierbarer Funktionen mittels abzählbar vieler Parameter ermöglichen. Fourier-Reihen bilden das wichtigste Werkzeug der Mathematik zur Approximation von Funktionen durch Schwingungen; daher ist die Theorie der Fourier-Reihen zum Beispiel in der Kommunikationstechnik unentbehrlich.

Definition 5.1 (periodische Funktion)

Seien $L > 0$ eine reelle Zahl und

$$f : \mathbb{R} \rightarrow \mathbb{C} \quad (5.18)$$

eine Funktion mit

$$f(x + L) = f(x) \quad \text{für alle } x \in \mathbb{R}, \quad (5.19)$$

dann heißt die Funktion f L -periodisch bzw. periodisch mit der Periode L .

Für jede L -periodische Funktion gilt offensichtlich

$$f(x + nL) = f(x) \quad \text{für alle } n \in \mathbb{Z} \quad \text{und alle } x \in \mathbb{R}. \quad (5.20)$$

Durch die Transformation

$$F : \mathbb{R} \rightarrow \mathbb{C}, \quad x \mapsto f\left(\frac{L}{2\pi}x\right) \quad (5.21)$$

wird aus einer L -periodischen Funktion f eine 2π -periodische Funktion F . Aus F gewinnt man die Funktion f durch

$$f(x) = F\left(\frac{2\pi}{L}x\right) \quad \text{für alle } x \in \mathbb{R} \quad (5.22)$$

zurück. Wir werden uns daher im Folgenden auf 2π -periodische Funktionen beschränken.

Beispiel(e) 5.2

Die reellen trigonometrischen Polynome der Ordnung n gegeben durch

$$f : \mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto \frac{a_0}{2} + \sum_{k=1}^n (a_k \cos(kx) + b_k \sin(kx)), \quad a_0, a_i, b_i \in \mathbb{R}, \quad i \in \mathbb{N} \quad (5.23)$$

sind 2π -periodische Funktionen. Da mit partieller Integration gezeigt werden kann, dass

$$\int_0^{2\pi} \cos(kx) \sin(lx) dx = 0 \quad \text{für alle } k, l \in \mathbb{N}_0 \quad (5.24)$$

$$\int_0^{2\pi} \cos(kx) \cos(lx) dx = \int_0^{2\pi} \sin(kx) \sin(lx) dx = 0 \quad \text{für alle } k, l \in \mathbb{N}_0, \quad k \neq l \quad (5.25)$$

$$\int_0^{2\pi} \cos^2(kx) dx = \int_0^{2\pi} \sin^2(kx) dx = \pi \quad \text{für alle } k \in \mathbb{N}, \quad (5.26)$$

sind für

$$f : \mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto \frac{a_0}{2} + \sum_{k=1}^n (a_k \cos(kx) + b_k \sin(kx)), \quad a_0, a_i, b_i \in \mathbb{R}, \quad i \in \mathbb{N} \quad (5.27)$$

die Konstanten

$$a_0, a_i, b_i \in \mathbb{R}, \quad i \in \mathbb{N} \quad (5.28)$$

durch f eindeutig bestimmt. Zum Beispiel erhält man aus der Gleichung

$$f(x) \cdot \cos(mx) = \frac{a_0}{2} \cos(mx) + \sum_{k=1}^n (a_k \cos(kx) \cos(mx) + b_k \sin(kx) \cos(mx)), \quad m \in \mathbb{N}_0 \quad (5.29)$$

durch Integration:

$$\int_0^{2\pi} f(x) \cos(mx) dx = \pi a_m \quad \text{bzw.} \quad a_m = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos(mx) dx. \quad (5.30)$$

Analog erhält man für $n \in \mathbb{N}$:

$$\int_0^{2\pi} f(x) \sin(nx) dx = \pi b_n \quad \text{bzw.} \quad b_n = \frac{1}{\pi} \int_0^{2\pi} f(x) \sin(nx) dx. \quad (5.31)$$

Die Parameter a_m und b_n lassen sich nach den Formeln (5.30) und (5.31) für beliebige, 2π -periodische, integrierbare Funktionen ausrechnen und man kann sich fragen, ob dann auch die Identität

$$f(x) = \frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos(kx) + b_k \sin(kx)) \quad (5.32)$$

gilt in dem Sinn, dass für jedes $x \in \mathbb{R}$ die rechts stehende Reihe gegen den Funktionswert $f(x)$ konvergiert. Dies ist zum Beispiel richtig, wenn f stetig differenzierbar ist. Bevor wir darauf zurückkommen, vereinfachen wir die Darstellung trigonometrischer Polynome: Verwendet man in

der Abbildungsvorschrift

$$f : \mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto \frac{a_0}{2} + \sum_{k=1}^n (a_k \cos(kx) + b_k \sin(kx)), \quad a_0, a_i, b_i \in \mathbb{R}, \quad i \in \mathbb{N} \quad (5.33)$$

die bekannten Formeln

$$\cos(kx) = \frac{1}{2} (e^{ikx} + e^{-ikx}) \quad k \in \mathbb{N}_0 \quad (5.34)$$

$$\sin(kx) = \frac{1}{2i} (e^{ikx} - e^{-ikx}) \quad k \in \mathbb{N}_0, \quad (5.35)$$

so ergibt sich für alle $x \in \mathbb{R}$:

$$f(x) = \sum_{k=1}^n c_{-k} e^{-ikx} + \sum_{k=0}^n c_k e^{ikx} =: \sum_{k=-n}^n c_k e^{ikx} \quad (5.36)$$

mit $c_0 = \frac{a_0}{2}$, $c_{-k} = \frac{1}{2} (a_k + ib_k)$ und $c_k = \frac{1}{2} (a_k - ib_k)$, $k \in \mathbb{N}$. Diese Beobachtung führt zu der Idee, auch komplexwertige trigonometrische Funktionen

$$f : \mathbb{R} \rightarrow \mathbb{C}, \quad x \mapsto \sum_{k=-n}^n c_k e^{ikx}, \quad c_k \in \mathbb{C}, \quad k \in \{-n, -n+1, \dots, -1, 0, 1, \dots, n\}, \quad (5.37)$$

die auch 2π -periodisch sind, zu betrachten.

Speziell für Funktionen

$$\phi : [a, b] \rightarrow \mathbb{C}, \quad x \mapsto e^{imx}, \quad m \in \mathbb{Z} \setminus \{0\} \quad (5.38)$$

ergibt sich

$$\int_a^b e^{imx} dx = \frac{1}{im} e^{imx} \Big|_a^b \quad (5.39)$$

also für $a = 0$ und $b = 2\pi$:

$$\int_0^{2\pi} e^{imx} dx = 0. \quad (5.40)$$

Für

$$f : \mathbb{R} \rightarrow \mathbb{C}, \quad x \mapsto \sum_{k=-n}^n c_k e^{ikx}, \quad c_k \in \mathbb{C}, \quad k \in \{-n, -n+1, \dots, -1, 0, 1, \dots, n\} \quad (5.41)$$

gilt

$$f(x) e^{-imx} = \sum_{k=-n}^n c_k e^{i(k-m)x} \quad \text{für alle } x \in \mathbb{R} \quad (5.42)$$

und somit durch Integration

$$c_m = \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-imx} dx \quad c_m \in \mathbb{C}, \quad m \in \{-n, -n+1, \dots, -1, 0, 1, \dots, n\}. \quad (5.43)$$

Im Folgenden soll eine gegebene 2π -periodische Riemann-integrierbare Funktion g durch eine Folge von komplexen trigonometrischen Polynomen approximiert werden. Daher definiert man

Definition 5.3 (Fourier-Koeffizienten, Fourier-Reihe)

Sei $f : \mathbb{R} \rightarrow \mathbb{C}$ eine 2π -periodische und über dem Intervall $[0, 2\pi]$ Riemann-integrierbare Funktion, dann heißen die Zahlen

$$c_m = \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-imx} dx \quad c_m \in \mathbb{C}, m \in \mathbb{Z} \quad (5.44)$$

Fourier-Koeffizienten von f und mit

$$\mathcal{F}_n[f] : \mathbb{R} \rightarrow \mathbb{C}, \quad x \mapsto \sum_{k=-n}^n c_k e^{ikx} \quad (5.45)$$

die Funktionenfolge $(\mathcal{F}_n[f])$, $n \geq 0$, Fourier-Reihe von f .

Im allgemeinen konvergiert eine Folge $(\mathcal{F}_n[g](x))$, $n \geq 0$, $x \in \mathbb{R}$, von Funktionswerten einer Fourier-Reihe von g nicht gegen $g(x)$. Man muss dazu weitere Forderungen an g stellen, zum Beispiel die der stetigen Differenzierbarkeit. Die Verhältnisse werden wesentlich einfacher, wenn man auf einen anderen Konvergenzbegriff, den der Konvergenz in quadratischen Mittel, ausweicht.

Definition 5.4 (Konvergenz im quadratischen Mittel)

Seien

$$f, f_n : \mathbb{R} \rightarrow \mathbb{C} \quad (5.46)$$

für alle $n \in \mathbb{N}_0$ 2π -periodische und über dem Intervall $[0, 2\pi]$ Riemann-integrierbare Funktionen, dann heißt die Funktionenfolge (f_n) , $n \geq 0$, konvergent im quadratischen Mittel gegen f , falls

$$\lim_{n \rightarrow \infty} \int_0^{2\pi} |f(x) - f_n(x)|^2 dx = 0. \quad (5.47)$$

Bei diesem Konvergenzbegriff wird die quadratische Abweichung der zu betrachtenden Funktion und ihrer entsprechenden Approximation über dem ganzen Intervall $[0, 2\pi]$ integriert.

Satz 5.5 (Konvergenz der Fourier-Reihe)

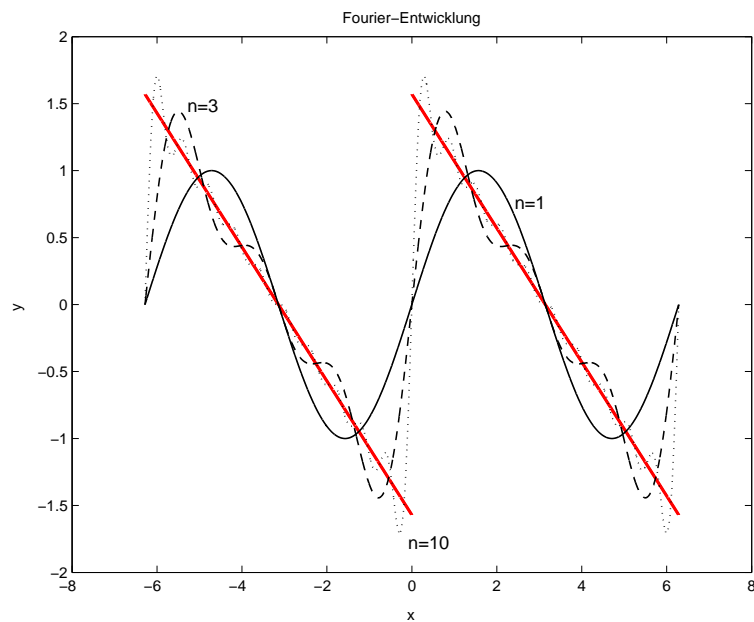
Sei $g : \mathbb{R} \rightarrow \mathbb{C}$ eine 2π -periodische und über dem Intervall $[0, 2\pi]$ Riemann-integrierbare Funktion, dann konvergiert die Fourier-Reihe von g im quadratischen Mittel gegen g und für die Fourierkoeffizienten c_k gilt:

$$\lim_{n \rightarrow \infty} \sum_{k=-n}^n |c_k|^2 = \frac{1}{2\pi} \int_0^{2\pi} |f(x)|^2 dx. \quad (5.48)$$

Die folgende Abbildung zeigt die 2π -periodische Funktion g mit

$$g|_{[0, 2\pi]} : [0, 2\pi] \rightarrow \mathbb{R}, \quad x \mapsto \frac{\pi}{2} - \frac{1}{2}x \quad (5.49)$$

und die Funktionen $\mathcal{F}_1[g]$, $\mathcal{F}_3[g]$ und $\mathcal{F}_{10}[g]$ im Intervall $[-2\pi, 2\pi]$.



Lässt man bei den Fourier-Reihen neben ganzzahligen Frequenzen in e^{ikx} , $k \in \mathbb{Z}$, auch reelle Frequenzen zu, so kommt man zur Fourier-Transformation:

Sei $f : \mathbb{R} \rightarrow \mathbb{C}$ eine stückweise stetige, beschränkte Funktion mit der Eigenschaft der „absoluten Integrierbarkeit“, das heißt:

$$\int_{-\infty}^{\infty} |f(t)| dt < \infty, \tag{5.50}$$

dann kann man f eine Funktion F zuordnen, definiert durch

$$F : \mathbb{R} \rightarrow \mathbb{C}, \quad \omega \mapsto \int_{-\infty}^{\infty} f(t)e^{-i\omega t} dt. \tag{5.51}$$

F heißt Fourier-Transformierte von f , die Zuordnung von f zu F heißt Fourier-Transformation oder Fourier-Analyse. Man schreibt dafür:

$$\mathcal{F}(f) = F. \tag{5.52}$$

Wenn f zusätzliche Eigenschaften hat, dann kann man f auch wieder aus F rekonstruieren mittels der Formel

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega)e^{i\omega t} d\omega. \tag{5.53}$$

Die Rekonstruktion von f aus F heißt inverse Fourier-Transformation oder Fourier-Synthese, man schreibt dafür auch $f = \mathcal{F}^{-1}(F)$. Diese Rekonstruktion ist zum Beispiel unter den beiden folgenden zusätzlichen Bedingungen an f (neben der absoluten Integrierbarkeit) möglich:

- $f : \mathbb{R} \rightarrow \mathbb{C}$ ist stückweise differenzierbar, das heißt die Ableitung existiert bis auf endlich viele Punkte, wo jedoch wenigstens die entsprechenden einseitigen Grenzwerte existieren und
- es gilt:

$$f(t) = \frac{1}{2} \left(\lim_{x \rightarrow t^-} f(x) + \lim_{x \rightarrow t^+} f(x) \right) \quad \text{für alle } t \in \mathbb{R}. \tag{5.54}$$

An Stellen t , an denen die letzte Bedingung nicht erfüllt ist, muss die Formel (5.53) nicht gelten. Seien nun im folgenden $\mathcal{F}(f) = F$ und $\mathcal{F}(g) = G$, also insbesondere für f, g die obigen Voraussetzungen erfüllt, so erhalten wir folgende Eigenschaften:

- Linearität:

$$\mathcal{F}(af + bg) = aF + bG, \quad \text{für alle } a, b \in \mathbb{C}. \quad (5.55)$$

- Mit

$$f * g : \mathbb{R} \rightarrow \mathbb{C}, \quad t \mapsto \int_{-\infty}^{\infty} f(t - \tau)g(\tau)d\tau \quad (\text{Faltung}) \quad (5.56)$$

gilt das Faltungstheorem:

$$\mathcal{F}(f * g) = F \cdot G. \quad (5.57)$$

-

$$\int_{-\infty}^{\infty} f(t)\overline{g(t)}dt = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega)\overline{G(\omega)}d\omega. \quad (5.58)$$

- Falls $f^{(n)}$ absolut integrierbar ist, dann gilt

$$\mathcal{F}(f^{(n)}) = (i\omega)^n \mathcal{F}(f). \quad (5.59)$$

Eine entscheidende Bedeutung in der Nachrichtentechnik kommt dem folgenden Resultat zu:

Sei F die Fourier-Transformierte von f und sei

$$F(\omega) = 0 \quad \text{für alle } |\omega| > \alpha > 0, \quad (5.60)$$

so gilt:

$$f(t) = \sum_{n=-\infty}^{\infty} f\left(\frac{n\pi}{\alpha}\right) \frac{\sin(\alpha t - n\pi)}{\alpha t - n\pi}. \quad (5.61)$$

Die Funktion f besteht also aus einer Superposition von Schwingungen. Mit der maximalen Kreisfrequenz

$$\omega_{\max} = \alpha$$

ist die größte auftretende Frequenz gegeben durch

$$\text{fr}_{\max} = \frac{\alpha}{2\pi}. \quad (5.62)$$

Die Funktion f ist somit durch alle Funktionswerte $f\left(\frac{n\pi}{\alpha}\right), n \in \mathbb{Z}$, an den Vielfachen von

$$\frac{1}{2 \cdot \text{fr}_{\max}}$$

festgelegt. Dieses Resultat wird als Abtasttheorem bezeichnet. Es erlaubt die Rekonstruktion bandbegrenzter Signale ($F(\omega) = 0$ für alle $|\omega| > \alpha > 0$) durch „Abtastung“ der Funktionswerte $f\left(\frac{n\pi}{\alpha}\right), n \in \mathbb{Z}$.