Höhere Mathematik I (P/ET/AI/MP/DS)

Dr. Agnes Lamacz-Keymling

Fakultät für Mathematik
TU Dortmund

Wintersemester 23/24

Organisatorisches

- Vorlesung Montags 8:15-10:00 Uhr und Dienstags 10:15-12:00 Uhr im Hörsaalgebäude II, Hörsaal 1.
- Übungen in Kleingruppen Montags und Dienstag. Der Übungsbetrieb startet am 16. Oktober.
- Globalübung Freitags 12:15-14:00 Uhr im Audimax.
- Übungsgruppenauswahl via Moodle.
- Vorlesungsmaterialien werden nach und nach auf der Moodle-Seite zur Verfügung gestellt.
- Übungsabgaben online im pdf-Format bei Moodle.
- Gruppenabgaben sind höchstens zu zweit erlaubt. Beide Studierende müssen dazu in <u>einer</u> Übungsgruppe sein.

Viel Erfolg bei der Vorlesung!

Kapitel 1 – Grundlagen und Zahlenmengen

Mengenbegriff nach Cantor

Definition 1.1 (Menge)

Unter einer MENGE *verstehen wir eine Zusammenfassung von bestimmten wohlunterscheidbaren Objekten zu einer Gesamtheit.*
Diese Objekte heißen dann ELEMENTE *der Menge.*

Beschreibung von Mengen durch...

- ...Aufzählen aller Elemente mit Mengenklammern $\{...\}$.
- ...Angabe einer Eigenschaft E, die die Elemente beschreibt:

$$\{x \,|\, x \text{ hat die Eigenschaft } E\}$$

Beispiele:

- $\{1, 2, 3\}$ die Menge, die aus den Zahlen $1, 2, 3$ als Elemente besteht.
- $\{r, m, d, t, o, u, n\} = \{x \,|\, x \text{ ist ein Kleinbuchstabe im Wort "Dortmund"}\}$.
- $\{1, 3, 5, 7, ...\}$ die Menge der ungeraden Zahlen.
- $\emptyset$ ist die *leere Menge*, die kein Element enthält. Alternative Notation: $\{\,\}$.

Schreibweise: Sei M eine Menge.

- Ist x ein Element der Menge M, so schreiben wir kurz $x \in M$.
- Ist x kein Element der Menge M, so schreiben wir kurz $x \notin M$.

Definition 1.2 (Erste Mengenrelationen)

Es seien M und N Mengen.

- *M heißt* TEILMENGE *von N, wenn alle Elemente von M auch in N enthalten sind. Symbolisch:*

$$M \subset N \quad \textit{oder} \quad N \supset M$$

- *M und N sind* GLEICH, *wenn $M \subset N$ und $N \subset M$, wenn also M und N die gleichen Elemente enthalten.*

Beispiele:

- $1 \in \{1, 2, 3\}$, aber $4 \notin \{1, 2, 3\}$.
- $\{1, 2\} \subset \{1, 2, 3\}$ und $\{1, 2\} = \{2, 1\}$.
- Für jede beliebige Menge M gilt $\emptyset \subset M \subset M$. Die leere Menge ist in jeder Menge enthalten. Jede Menge ist in sich selbst enthalten.

Grundlegende Zahlenmengen

Definition 1.3

$\mathbb{N}$ *die Menge der* natürlichen Zahlen*:* $\mathbb{N} := \{1,\ 2,\ 3,\ 4,\ \dots\}$

$\mathbb{N}_0$ *die Menge der* natürlichen Zahlen mit 0*:* $\mathbb{N}_0 := \{0, 1,\ 2,\ 3,\ \dots\}$

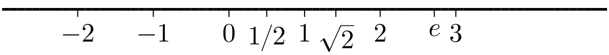
$\mathbb{Z}$ *die Menge der* ganzen Zahlen*:* $\mathbb{Z} := \{\dots,\ -3,\ -2,\ -1,\ 0,\ 1,\ 2,\ 3,\ \dots\}$

$\mathbb{Q}$ *die Menge der* rationalen Zahlen*:* $\mathbb{Q} := \left\{ \frac{m}{n} \,\middle|\, m \in \mathbb{Z},\ n \in \mathbb{N} \right\}$

$\mathbb{R}$ *die Menge der* reellen Zahlen

$\mathbb{C}$ *die Menge der* komplexen Zahlen *(später)*

Zur Veranschaulichung reeller Zahlen dient die Zahlengerade.



Jedem Punkt auf der Zahlengeraden entspricht genau eine reelle Zahl, und umgekehrt.

$$\mathbb{N} \subset \mathbb{N}_0 \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R} \subset \mathbb{C}$$

Definition 1.4 (Summen- und Produktzeichen)

Für ganze Zahlen $m \leqslant n$ definiert man:

$$\sum_{j=m}^{n} a_j = a_m + a_{m+1} + \cdots + a_n$$

$$\prod_{j=m}^{n} a_j = a_m \cdot a_{m+1} \cdots a_n.$$

Für jede reelle (oder komplexe) Zahl x definieren wir $x^0 := 1$.

Als Sonderfall für $m > n$ wird vereinbart:

$$\sum_{j=m}^{n} a_j = 0, \qquad \prod_{j=m}^{n} a_j = 1.$$

Definition 1.5 (Fakultät und Binomialkoeffizient)

- *Für* $n \in \mathbb{N}$ *ist*

$$n! = 1 \cdot 2 \cdot 3 \cdot \dots \cdot n = \prod_{j=1}^{n} j$$

die FAKULTÄT *von* n *(kurz "*n*-Fakultät"); wir setzen außerdem* $0! = 1$.

- *Für ganzzahlige* $n \geqslant 0$ *und* $0 \leqslant k \leqslant n$ *ist*

$$\binom{n}{k} = \frac{n!}{k!(n-k)!} = \frac{n(n-1)\cdot \dots \cdot (n-k+1)}{k!}$$

der BINOMIALKOEFFIZIENT n *über* k.

Die Binomialkoeffizienten erfüllen die Regeln:

(i) $\binom{n}{k} = \binom{n}{n-k}$; speziell $\binom{n}{0} = \binom{n}{n} = 1$

(ii) $\binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1}$ für $n \geqslant 1$ und $0 \leqslant k \leqslant n-1$

Eine wichtige Beweismethode ist das folgende Prinzip.

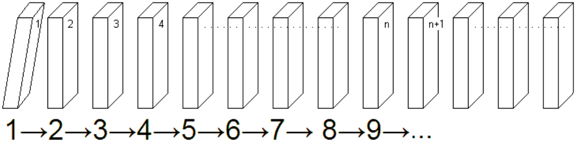
Satz 1.6 (Das Prinzip der vollständigen Induktion)

Für jede natürliche Zahl n sei eine Aussage $A(n)$ formuliert. Wenn wir beweisen können, dass die folgenden beiden Aussagen gelten:

(i) $A(1)$ *ist wahr. (***Induktionsanfang***)*

(ii) *Wenn für eine natürliche Zahl n die Aussage $A(n)$ wahr ist, dann ist auch $A(n+1)$ wahr. (***Induktionsschluss*** *von n auf $n+1$)*

Dann ist bewiesen, dass die Aussage $A(n)$ für jede natürliche Zahl n wahr ist.



Von Joachim Mohr - Eigenes Werk, CC BY-SA 3.0, https://commons.wikimedia.org/w/index.php?curid=37681531

Satz 1.7 (Gaußsche Summenformel)

Für jede natürliche Zahl n *gilt* $\displaystyle\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$.

Satz 1.8 (Anzahl der Permutationen)

Für jede natürliche Zahl n *gibt es genau* $n!$ *verschiedene Möglichkeiten, die Zahlen* $1,\ 2,\ 3, \dots, n$ *anzuordnen.*

Anders ausgedrückt: es gibt genau $n!$ **Permutationen** von n verschiedenen Objekten.)

Satz 1.9 (Anzahl der Kombinationen)

Für jede natürliche Zahl n *und jede ganze Zahl* $0 \leqslant k \leqslant n$ *gibt es genau* $\binom{n}{k}$ *verschiedene Möglichkeiten,* k *verschiedene natürliche Zahlen zwischen* 1 *und* n *auszuwählen (ohne Zurücklegen und ohne Berücksichtigung der Reihenfolge; siehe Lotto 6 aus 49).*

Anders ausgedrückt: Eine Menge mit n Elementen hat genau $\binom{n}{k}$ Teilmengen mit k Elementen.

Satz 1.10 (Binomischer Satz)

Für alle $a, b \in \mathbb{R}$ *und* $n \in \mathbb{N}$ *gilt*

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

Satz 1.11 (Varianten des Induktionsbeweises)

- *Als Induktionsanfang beweist man* $A(n_0)$ *für ein* $n_0 \in \mathbb{Z}$. *Gilt dann der Induktionsschluss von* n *nach* $n+1$ *für jedes* $n \geqslant n_0$, *so ist die Aussage* $A(n)$ *für alle* $n \geqslant n_0$ *bewiesen.*
- *Als Voraussetzung für den Induktionsschluss von* n *nach* $n+1$ *darf man verwenden, dass* $A(k)$ *wahr ist für alle* $n_0 \leqslant k \leqslant n$.

Satz 1.12 (Geometrische Summenformel)

Für beliebiges $a \in \mathbb{R}$, $a \neq 1$, *und* $n \in \mathbb{N}_0$ *gilt*

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

Die Eigenschaften der reellen Zahlen

1.13 ($(\mathbb{R}, +, \cdot)$ ist ein Körper)

Für die Addition reeller Zahlen gilt:

(A1) *Zu beliebigen $a, b \in \mathbb{R}$ gibt es genau eine reelle Zahl $a + b$, die* SUMME *von a und b.*

(A2) $a + b = b + a$ *für beliebige $a, b \in \mathbb{R}$* *(Kommutativität)*

(A3) $(a + b) + c = a + (b + c)$ *für beliebige $a, b, c \in \mathbb{R}$* *(Assoziativität)*

(A4) *Es gibt genau eine reelle Zahl 0 mit der Eigenschaft $a + 0 = 0 + a = a$ für jede reelle Zahl a.*

(A5) *Für jede reelle Zahl a gibt es genau eine reelle Zahl $-a$ mit der Eigenschaft $a + (-a) = (-a) + a = 0$.*

Für die Multiplikation reeller Zahlen gilt:

(M1) Zu beliebigen $a, b \in \mathbb{R}$ gibt es genau eine reelle Zahl $a \cdot b$ (geschrieben ab), das PRODUKT von a und b.

(M2) $ab = ba$ für beliebige $a, b \in \mathbb{R}$ (Kommutativität)

(M3) $(ab)c = a(bc)$ für beliebige $a, b, c \in \mathbb{R}$ (Assoziativität)

(D) $(a + b)c = ac + bc$ für beliebige $a, b, c \in \mathbb{R}$ (Distributivität)

(M4) Es gibt genau eine reelle Zahl 1 mit der Eigenschaft $a \cdot 1 = 1 \cdot a = a$ für jede reelle Zahl a.

(M5) Für jede reelle Zahl $a \neq 0$ gibt es genau eine reelle Zahl a^{-1} (geschrieben $\frac{1}{a}$) mit der Eigenschaft: $a \cdot a^{-1} = a^{-1} \cdot a = 1$.

Bemerkungen:

- Schreibweise der Subtraktion: $a - b := a + (-b)$
- Schreibweise der Division und Brüche: $a/b := \frac{a}{b} := ab^{-1}$, falls $b \neq 0$.
- Das 2. Distributivgesetz $a(b+c) = ab + ac$ folgt aus (D) und (M2).

1.14 (Die Axiome der Anordnung in $\mathbb{R}$)

(O1) *Es gibt eine Relation “ $<$ ”(kleiner) in $\mathbb{R}$, so dass für je zwei reelle Zahlen a, b genau eine der drei folgenden Aussagen gilt:*

$$a < b, \qquad a = b, \qquad b < a$$

Die Relation “ $<$ ”hat die folgenden Eigenschaften:

(O2) *Aus* $a < b$ *und* $b < c$ *folgt* $a < c$ *(Transitivität)*

(O3) *Aus* $a < b$ *folgt für alle reellen* c *:* $a + c < b + c$

(O4) *Aus* $a < b$ *folgt für alle reellen* c *mit* $0 < c$ *:* $ac < bc$

Schreibweisen:

- $b > a$ (b größer a) bedeutet $a < b$.
- $a \leqslant b$ (a kleiner oder gleich b) bedeutet $a < b$ oder $a = b$.
- $b \geqslant a$ (b größer oder gleich a) bedeutet $b > a$ oder $b = a$.

Definition 1.15 (Intervalle)

Für $a, b \in \mathbb{R}$ *mit* $a \leqslant b$ *definiert man*

- *das* ABGESCHLOSSENE INTERVALL $[a,b] := \{x \in \mathbb{R} \mid a \leqslant x \leqslant b\}$
- *das* OFFENE INTERVALL $(a,b) := \{x \in \mathbb{R} \mid a < x < b\}$
- *die* HALBOFFENEN INTERVALLE $(a,b] := \{x \in \mathbb{R} \mid a < x \leqslant b\}$ *und* $[a,b) := \{x \in \mathbb{R} \mid a \leqslant x < b\}$

Weiterhin definiert man

- *die* UNBESCHRÄNKTEN ABGESCHLOSSENEN INTERVALLE $[a,\infty) := \{x \in \mathbb{R} \mid x \geqslant a\}$ *und* $(-\infty,b] := \{x \in \mathbb{R} \mid x \leqslant b\}$,
- *die* UNBESCHRÄNKTEN OFFENEN INTERVALLE $(a,\infty) := \{x \in \mathbb{R} \mid x > a\}$ *und* $(-\infty,b) := \{x \in \mathbb{R} \mid x < b\}$

Für $a \in \mathbb{R}$ *und* $\epsilon > 0$ *heißt das offene Intervall*

$$U_\epsilon(a) := (a-\epsilon, a+\epsilon)$$

ϵ-UMGEBUNG VON a.

Satz 1.16 (Rechenregeln für Ungleichungen)

Für $a, b, c, d \in \mathbb{R}$ folgt aus den Axiomen (O1) bis (O4):

(a) $a < b \quad \wedge \quad c < d \qquad \Longrightarrow \qquad a + c < b + d$

(b) $a < b \quad \wedge \quad c < 0 \qquad \Longrightarrow \qquad ac > bc$

(c) $1 > 0$

(d) $0 < a < b \qquad \Longrightarrow \qquad 0 < \frac{1}{b} < \frac{1}{a}$

(e) $ab > 0 \qquad \Longleftrightarrow \qquad (a > 0 \ \wedge \ b > 0) \ \vee \ (a < 0 \ \wedge \ b < 0)$

$ab < 0 \qquad \Longleftrightarrow \qquad (a > 0 \ \wedge \ b < 0) \ \vee \ (a < 0 \ \wedge \ b > 0)$

(f) *Für jede reelle Zahl $a \neq 0$ ist $a^2 > 0$.*

(g) *Für $a > 0$ und $b > 0$ gilt:* $\quad a < b \quad \Longleftrightarrow \quad a^2 < b^2$

Die Symbole $\wedge$ ("und"), $\vee$ ("oder") sind Verknüpfungen aus der Aussagenlogik.

Die Charakterisierung der reellen Zahlen

1.17 (Archimedisches Axiom)

Zu jeder reellen Zahl a gibt es eine natürliche Zahl $n \in \mathbb{N}$ mit $a < n$.

Äquivalent: Zu jeder reellen Zahl $\epsilon > 0$ gibt es ein $n \in \mathbb{N}$ mit $0 < \frac{1}{n} < \epsilon$.

1.18 (Vollständigkeitsaxiom, "Dedekindscher Schnitt")

Die Mengen A und B seien nichtleere Mengen reeller Zahlen. Für jedes $a \in A$ und jedes $b \in B$ gelte $a \leqslant b$ (anschaulich: A liegt links auf der Zahlengeraden von B.) Dann gibt es (mindestens) eine reelle Zahl c, so dass für alle $a \in A$ und alle $b \in B$ gilt

$$a \leqslant c \leqslant b.$$

Die Körperaxiome 1.13, Anordnungsaxiome 1.14, sowie die beiden letzten Axiome <u>charakterisieren</u> die Menge $\mathbb{R}$; d.h. $\mathbb{R}$ ist der einzige vollständige, archimedische angeordnete Körper.

Frage: $\mathbb{Q}$ ist ein archimedischer angeordneter Körper. Man finde ein Beispiel von Mengen $A, B \subset \mathbb{Q}$, die zeigen, dass das Vollständigkeitsaxiom für $\mathbb{Q}$ nicht gilt.

Definition 1.19 (obere Schranke, Supremum)

M sei eine nichtleere Teilmenge von $\mathbb{R}$.

- *Existiert eine Zahl $b \in \mathbb{R}$, so dass $x \leqslant b$ für alle $x \in M$ gilt, so heißt b eine* OBERE SCHRANKE *von M, und M heißt* NACH OBEN BESCHRÄNKT.

- *Das Vollständigkeitsaxiom 1.18 ist äquivalent zu der folgenden Aussage: Falls die Menge $M \subseteq \mathbb{R}$ nach oben beschränkt ist, so gibt es eine kleinste obere Schranke s von M. Die Zahl s heißt* SUPREMUM *von M, geschrieben $s = \sup M$. Es gilt*

$$x \leqslant \sup M \quad \text{für alle } x \in M,$$

jedoch existiert für jedes $\epsilon > 0$ (mindestens) ein $x \in M$ mit $x > \sup M - \epsilon$.

Analog definiert man:

Definition 1.20 (Untere Schranken, Infimum)

M *sei eine nichtleere Teilmenge von* $\mathbb{R}$.

- *Existiert eine Zahl* $a \in \mathbb{R}$ *so, dass* $x \geqslant a$ *für alle* $x \in M$ *gilt, so heißt* a *eine* UNTERE SCHRANKE *von* M, *und* M *heißt* NACH UNTEN BESCHRÄNKT.
- *Die größte untere Schranke von* M *heißt das* INFIMUM *von* M, *geschrieben* $\inf M$. *Es gilt*
$$x \geqslant \inf M \quad \text{für alle } x \in M,$$
jedoch existiert für jedes $\epsilon > 0$ *(mindestens) ein* $x \in M$ *mit* $x < \inf M + \epsilon$.

Ist M nach oben und unten beschränkt, so heißt M BESCHRÄNKT; dann gilt

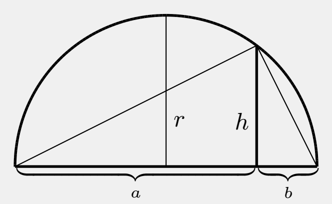
$$\inf M \leqslant x \leqslant \sup M \quad \text{für alle } x \in M.$$

Wichtige Ungleichungen der Analysis

Satz 1.21 (Ungleichung zwischen arithmetischem und geometrischem Mittel)

Sind a und b positive reelle Zahlen, so gilt

$$\sqrt{ab} \leqslant \frac{a+b}{2}.$$



$$r = \frac{a+b}{2}$$
$$h = \sqrt{ab}$$

Satz 1.22 (Bernoullische Ungleichung)

Für alle $x \in \mathbb{R}$ *mit* $x \geqslant -1$ *und alle* $n \in \mathbb{N}$ *gilt*

$$(1+x)^n \geqslant 1 + nx.$$

Satz 1.23 (Cauchy-Schwarz Ungleichung)

Für beliebige reelle Zahlen $a_1, a_2, \ldots, a_n$ *und* $b_1, b_2, \ldots, b_n$ *gilt*

$$\sum_{i=1}^{n} a_i b_i \leqslant \sqrt{a_1^2 + \cdots + a_n^2}\,\sqrt{b_1^2 + \cdots + b_n^2}$$

Definition 1.24 (Absolutbetrag und Signumsfunktion)

Für jede reelle Zahl a definieren wir

(i) *den* ABSOLUTBETRAG *oder* BETRAG

$$|a| := \begin{cases} a, & \text{wenn } a \geqslant 0, \\ -a, & \text{wenn } a < 0. \end{cases}$$

(ii) *das* SIGNUM *(Vorzeichen)*

$$\mathit{sign}(a) := \begin{cases} 1, & \text{wenn } a > 0, \\ 0, & \text{wenn } a = 0, \\ -1, & \text{wenn } a < 0. \end{cases}$$

Bemerkungen:

- Der Absolutbetrag $|a|$ ist der Abstand auf der Zahlengeraden von a zum Nullpunkt.
- Der Abstand von zwei Zahlen a und b auf der Zahlengeraden ist $|a - b|$.
- Für alle $a \in \mathbb{R}$ gilt $|a| = \sqrt{a^2}$.

Satz 1.25 (Rechnen mit Beträgen)

Für alle reellen Zahlen a, b gilt

(a) $|a| \geqslant 0$, *und* $(|a| = 0 \;\Leftrightarrow\; a = 0)$;

(b) $|ab| = |a| \cdot |b|$;

(c) $\left|\frac{a}{b}\right| = \frac{|a|}{|b|}$, *falls* $b \neq 0$;

(d) $|a+b| \leqslant |a| + |b|$ (Dreiecksungleichung)

(e) $|a+b| \geqslant \big||a| - |b|\big|$ (umgekehrte Dreiecksungleichung)

(f) $(|a| = |b| \;\Leftrightarrow\; a^2 = b^2)$ *und* $(|a| < |b| \;\Leftrightarrow\; a^2 < b^2)$

Für reelle Zahlen a_j, $1 \leqslant j \leqslant n$, folgt aus (d) per Induktion die verallgemeinerte Dreiecksungleichung

$$\left|\sum_{j=1}^{n} a_j\right| \leqslant \sum_{j=1}^{n} |a_j|$$

Kapitel 2 – Mengen und Funktionen

Wiederholung

Beschreibung von Mengen durch...

- ...Aufzählen aller Elemente mit Mengenklammern $\{...\}$.
- ...Angabe einer Eigenschaft E, die die Elemente beschreibt:

$$\{x \,|\, x \text{ hat die Eigenschaft } E\}$$

Schreibweisen: Es seien M und N Mengen.

- Ist x ein Element der Menge M, so schreiben wir kurz $x \in M$.
- Ist x kein Element der Menge M, so schreiben wir kurz $x \notin M$.
- M heißt TEILMENGE von N, wenn alle Elemente von M auch in N enthalten sind. Symbolisch:

$$M \subset N \quad \text{oder} \quad N \supset M$$

- M und N sind GLEICH, wenn $M \subset N$ und $N \subset M$, wenn also M und N die gleichen Elemente enthalten.

Definition 2.1 (Mengentheoretische Begriffe)

Seien M und N Mengen.

- *Die Menge M heißt* ECHTE TEILMENGE *von N, geschrieben $M \subsetneq N$, falls*

$$M \subseteq N \text{ und } M \neq N$$

- *Die* VEREINIGUNG *der Mengen M und N ist*

$$M \cup N := \{x \mid x \in M \quad \text{oder} \quad x \in N\}$$

- *Der* DURCHSCHNITT *der Mengen M und N ist*

$$M \cap N := \{x \mid x \in M \quad \text{und} \quad x \in N\}$$

- *Die* MENGENDIFFERENZ *M "ohne" N (auch* KOMPLEMENT *von N in M) ist*

$$M \backslash N := \{x \mid x \in M \quad \text{und} \quad x \notin N\}$$

Satz 2.2 (Rechenregeln für Mengen)

Für Mengen A, B, C *gelten die* Distributivgesetze

$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C), \qquad (A \cap B) \cup C = (A \cup C) \cap (B \cup C)$$

sowie die de Morganschen Regeln

$$C \backslash (A \cup B) = (C \backslash A) \cap (C \backslash B), \qquad C \backslash (A \cap B) = (C \backslash A) \cup (C \backslash B).$$

Definition 2.3 (Kartesisches Produkt)

Zu nichtleeren Mengen A und B definieren wir das KARTESISCHE PRODUKT

$$A \times B = \{(a,b) \mid a \in A,\ b \in B\} \qquad (\textit{sprich "}A \textit{ kreuz } B\textit{"})$$

als die Menge der geordneten Paare (a,b) mit $a \in A$ und $b \in B$.

- Es gilt $(a,b) = (c,d)$ genau dann, wenn $a = c$ und $b = d$ gilt.

Beispiele:

- Das Rechteck $[0,3] \times [1,2]$ enthält geordnete Zahlenpaare. Es ist eine Teilmenge von $\mathbb{R}^2 := \mathbb{R} \times \mathbb{R}$.
- Die Menge $M :=$ {Sonntag, Montag, ..., Samstag} $\times$ {Sonne, Regen, Nebel} enthält 21 geordnete Paare der Form (Wochentag, Wetter).
- Positionen auf einem Schachbrett werden mit einem kartesischen Produkt beschrieben:
$$\{a,b,c,d,e,f,g,h\} \times \{1,2,3,4,5,6,7,8\}$$

Vorsicht: $(a,b) \neq \{a,b\}$

Definition 2.4 (Abbildung, Funktion)

Gegeben seien nichtleere Mengen D und W. Eine FUNKTION *(oder* ABBILDUNG*) f von D nach W ist eine Vorschrift, die jedem $x \in D$ genau ein Element $y = f(x) \in W$ zuordnet. Wir schreiben*

$$f : D \to W, \qquad x \mapsto f(x).$$

- D heißt der DEFINITIONSBEREICH, W heißt der WERTEBEREICH von f.
- $y = f(x)$ heißt das BILD von f an der Stelle x (auch Bild von x unter f).
- Für eine Teilmenge $M \subseteq D$ heißt

  $$f(M) := \{f(x) \mid x \in M\} \subseteq W$$

  das BILD von M unter f. Die Menge $f(D)$ ist die BILDMENGE von f.
- Für eine Teilmenge $N \subseteq W$ heißt

  $$f^{-1}(N) = \{x \in D \mid f(x) \in N\} \subseteq D$$

  das URBILD von N unter f.
- Der GRAPH von f ist die Menge

  $$\mathrm{Graph}(f) := \{(x, f(x)) \mid x \in D\} \subseteq D \times W.$$

Definition 2.5 (Einschränkung, Verkettung)

Es sei $f : D \to U$.

- *Für eine Teilmenge $M \subseteq D$ definieren wir die* Einschränkung *von f auf M*

$$f|_M : M \to U, \qquad x \mapsto f(x).$$

  f heißt dann Fortsetzung *von $f|_M$.*

Sei zusätzlich $g : V \to W$ eine Funktion und es gelte $U \subset V$.

- *Wir definieren die* Hintereinanderausführung *(oder* Verkettung *oder* Komposition*)*

$$g \circ f : D \to W, \qquad x \mapsto g(f(x)) \quad \textit{(sprich "g nach f")}.$$

Achtung: Im Allgemeinen gilt $f \circ g \neq g \circ f$.

Definition 2.6 (injektiv, surjektiv, bijektiv)

Eine Funktion $f : D \to W$ *heißt*

- INJEKTIV, *wenn es zu jedem* $y \in W$ *höchstens ein* $x \in D$ *gibt mit* $f(x) = y$,
- SURJEKTIV, *wenn es zu jedem* $y \in W$ *mindestens ein* $x \in D$ *gibt mit* $f(x) = y$,
- BIJEKTIV, *wenn es zu jedem* $y \in W$ *genau ein* $x \in D$ *gibt mit* $f(x) = y$.

Definition 2.7 (Umkehrfunktion)

Wenn $f : D \to W$ bijektiv ist, so ist die UMKEHRFUNKTION $f^{-1} : W \to D$ *definiert durch*

$$f^{-1}(y) = x \qquad \Longleftrightarrow \qquad f(x) = y.$$

- Der Graph der Umkehrfunktion ist

$$\text{Graph}(f^{-1}) = \{(y, f^{-1}(y)) \mid b \in B\} = \{(f(x), x) \mid x \in D\}.$$

- Für reelle Funktionen (also $D, W \subseteq \mathbb{R}$) ist Graph$(f^{-1})$ die Spiegelung von Graph(f) an der 1. Winkelhalbierenden im (x, y)-Koordinatensystem.

Beispiel:

Für $f : [0, \infty) \to [0, \infty)$ mit $f(x) = x^2$ ist die Umkehrfunktion gegeben durch

$$f^{-1} : [0, \infty) \to [0, \infty), \quad y \mapsto \sqrt{y}$$

Für $f : (-\infty, 0] \to [0, \infty)$ mit $f(x) = x^2$ ist die Umkehrfunktion gegeben durch

$$f^{-1} : [0, \infty) \to (-\infty, 0], \quad y \mapsto -\sqrt{y}$$

Reelle Funktionen

Definition 2.8 (Monotonie)

$D, W \subset \mathbb{R}$ seien nichtleere Mengen. Eine Funktion $f : D \to W$ heißt

- STRENG MONOTON WACHSEND, *wenn für alle $x_1, x_2 \in D$ gilt:*

$$x_1 < x_2 \implies f(x_1) < f(x_2),$$

- STRENG MONOTON FALLEND, *wenn für alle $x_1, x_2 \in D$ gilt:*

$$x_1 < x_2 \implies f(x_1) > f(x_2).$$

- *Die Funktion heißt* MONOTON WACHSEND *(bzw.* FALLEND*), wenn aus $x_1 < x_2$ die Beziehung $f(x_1) \leqslant f(x_2)$ (bzw. $f(x_1) \geqslant f(x_2)$) folgt.*

Satz 2.9 (Streng monotone Funktionen)

Ist die reelle Funktion $f : D \to W$ *streng monoton wachsend (oder streng monoton fallend), so ist sie injektiv.*

Definition 2.10 (gerade, ungerde)

Es sei $I \subset \mathbb{R}$ ein um 0 symmetrisches Intervall. Eine Funktion $f : I \to \mathbb{R}$ heißt

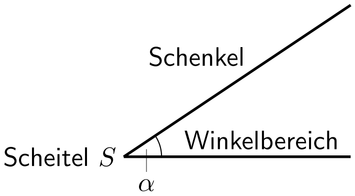
- GERADE *oder* ACHSENSYMMETRISCH, *wenn $f(-x) = f(x)$ für alle $x \in I$.*
- UNGERADE *oder* PUNKTSYMMETRISCH, *wenn $f(-x) = -f(x)$ für alle $x \in I$.*

Beispiel: Sei $k \in \mathbb{N}$. Die Funktion $f : \mathbb{R} \to \mathbb{R}$ mit $f(x) = x^k$ ist

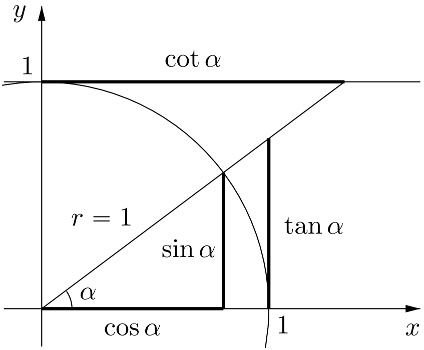
- für k gerade eine gerade Funktion.
- für k ungerade eine ungerade Funktion.

Definition 2.11 (Periodische Funktion)

*Es sei $T > 0$. Eine Funktion $f : \mathbb{R} \to \mathbb{R}$ heißt T-*PERIODISCH, *wenn $f(x+T) = f(x)$ für alle $x \in \mathbb{R}$.*



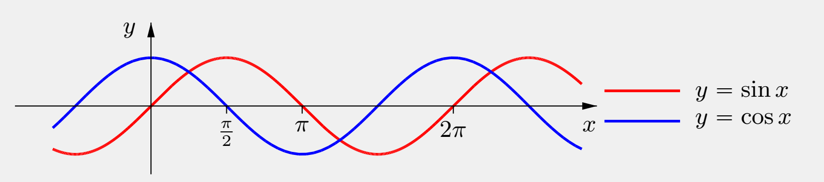
Winkel werden in GRAD oder im BOGENMASS (auch RAD) angegeben: $360^\circ \mathrel{\widehat{=}} 2\pi$.

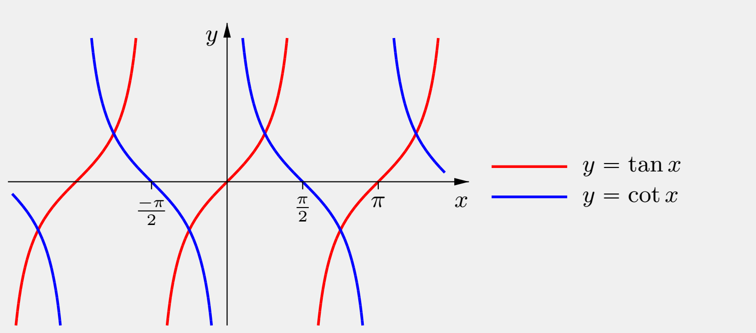


Durch diese Betrachtungen am Einheitskreis werden vier Funktionen definiert.

2.12 (Winkelfunktionen)

Name		D	W
Sinus	$\sin$	$\mathbb{R}$	$[-1,1]$
Cosinus	$\cos$	$\mathbb{R}$	$[-1,1]$
Tangens	$\tan$	$\mathbb{R}\backslash\{\frac{2k+1}{2}\pi \mid k \in \mathbb{Z}\}$	$\mathbb{R}$
Cotangens	$\cot$	$\mathbb{R}\backslash\{k\pi \mid k \in \mathbb{Z}\}$	$\mathbb{R}$





x in Grad	0	30°	45°	60°	90°
x in Rad	0	$\frac{\pi}{6}$	$\frac{\pi}{4}$	$\frac{\pi}{3}$	$\frac{\pi}{2}$
$\sin x$	0	$\frac{1}{2}$	$\frac{1}{2}\sqrt{2}$	$\frac{1}{2}\sqrt{3}$	1
$\cos x$	1	$\frac{1}{2}\sqrt{3}$	$\frac{1}{2}\sqrt{2}$	$\frac{1}{2}$	0
$\tan x$	0	$\frac{1}{3}\sqrt{3}$	1	$\sqrt{3}$	–
$\cot x$	–	$\sqrt{3}$	1	$\frac{1}{3}\sqrt{3}$	0

2.13 Eigenschaften

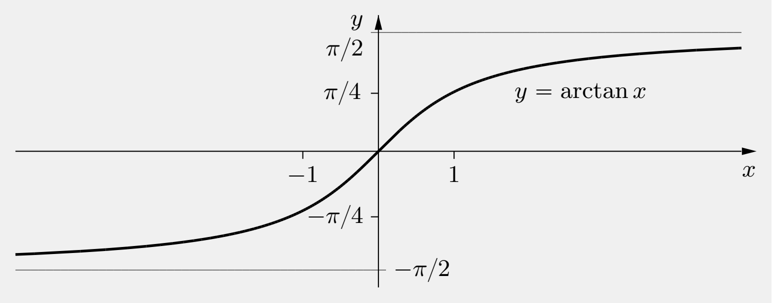
- $\sin$ sowie $\cos$ sind 2π- und $\tan$ sowie $\cot$ sind π-periodisch.
- $\sin(x+\pi) = -\sin x$ und $\cos(x+\pi) = -\cos x$.
- $\sin(x+\frac{\pi}{2}) = \cos x$ und $\cos(x+\frac{\pi}{2}) = -\sin x$.
- $\tan x = \dfrac{\sin x}{\cos x}$ und $\cot x = \dfrac{1}{\tan x} = \dfrac{\cos x}{\sin x}$.
- $\cos$ ist eine gerade Funktion und $\sin$, $\tan$ und $\cot$ sind ungerade Funktionen.
- Für alle $x \in \mathbb{R}$ gilt $|\sin x| \leqslant 1$ und $|\cos x| \leqslant 1$.
- $\sin(x) = 0$ genau dann, wenn $x = k\pi$ mit $k \in \mathbb{Z}$.
  $\cos(x) = 0$ genau dann, wenn $x = \frac{2k+1}{2}\pi$ mit $k \in \mathbb{Z}$.
- $\sin^2 x + \cos^2 x = 1$ der Trigonometrische Pythagoras.
- $\cos^2 x = \dfrac{1}{1+\tan^2 x}$ und $\sin^2 x = \dfrac{1}{1+\cot^2 x}$.
- $\sin(x \pm y) = \sin x \cos y \pm \sin y \cos x \quad$ (Additionstheorem)
- $\cos(x \pm y) = \cos x \cos y \mp \sin x \sin y \quad$ (Additionstheorem)

Um Umkehrfunktionen bilden zu können, wird der Sinus auf $[-\frac{\pi}{2}, \frac{\pi}{2}]$ und der Cosinus auf $[0, \pi]$ eingeschränkt.

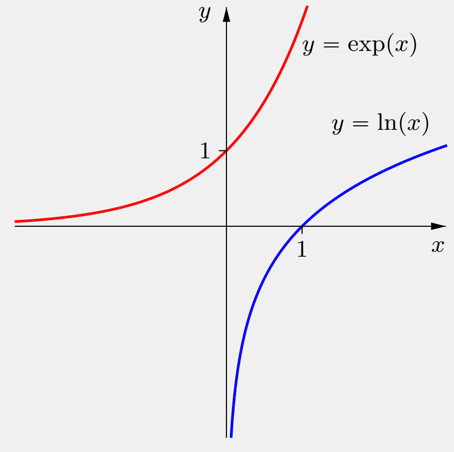
Definition 2.13 (Arcusfunktionen)

Die Umkehrfunktionen der Winkelfunktionen sind die Arcusfunktionen

- $\arcsin : [-1, 1] \to \big[-\frac{\pi}{2}, \frac{\pi}{2}\big]$
- $\arccos : [-1, 1] \to [0, \pi]$
- $\arctan : \mathbb{R} \to \big]-\frac{\pi}{2}, \frac{\pi}{2}\big[$
- $\operatorname{arccot} : \mathbb{R} \to \big]0, \pi\big[$



2.15 Exponentialfunktion und Logarithmus



Diese Funktionen werden später noch exakt definiert. Hier ist nur eine Zusammenstellung der Eigenschaften. Für $\ln : (0, \infty) \to \mathbb{R}$ gilt:

- $\ln 1 = 0$.
- $\ln \frac{1}{x} = -\ln x$.
- $\ln(x \cdot y) = \ln x + \ln y$.
- $\ln(x^n) = n \ln(x)$ für $n \in \mathbb{Z}$.
- $\ln$ ist streng monoton steigend.
- $\ln$ ist unbeschränkt.

Für $\exp : \mathbb{R} \to (0, \infty)$ gilt:

- $\exp$ ist streng monoton wachsend.
- $\exp(\ln x) = \ln(\exp x) = x$.
- $\exp(0) = 1$ und $\exp(x) > 0$.
- $\exp$ ist nach oben unbeschränkt, bei $-\infty$ nähert sich $\exp x$ der Null an.
- $\exp(x) \cdot \exp(y) = \exp(x + y)$
- $\exp(n \cdot x) = \big(\exp(x)\big)^n$ für $n \in \mathbb{Z}$.

Statt $\exp(x)$ schreibt man kürzer e^x.

Definition 2.14 (Eulersche Zahl)

$e := \exp(1) \approx 2,718281828\ldots$ *heißt* Eulersche Zahl.

Kapitel 3 – Komplexe Zahlen

3 Komplexe Zahlen

Die quadratische Gleichung $x^2 + 1 = 0$ hat keine reelle Lösung, denn für jede reelle Zahl x gilt $x^2 + 1 > 0$ (nach Satz 1.16).

Definition 3.1 (Komplexe Zahlen)

Zu den reellen Zahlen fügen wir die "Zahl" i (= "imaginäre Einheit") hinzu, für die $i^2 = -1$ gilt. Dann ist

$$\mathbb{C} := \{a + bi \mid a, b \in \mathbb{R}\}$$

die Menge der KOMPLEXEN ZAHLEN. *Für komplexe Zahlen $z = a + bi$ und $w = c + di$ definieren wie Addition und Multiplikation wie folgt:*

$$\begin{aligned} z + w &:= (a + bi) + (c + di) = (a + c) + (b + d)i, \\ zw &:= (a + bi)(c + di) = (ac - bd) + (ad + bc)i. \end{aligned}$$

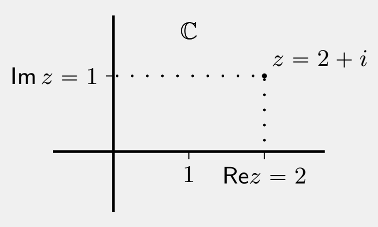
3.2 (Die Gaußsche Zahlenebene)

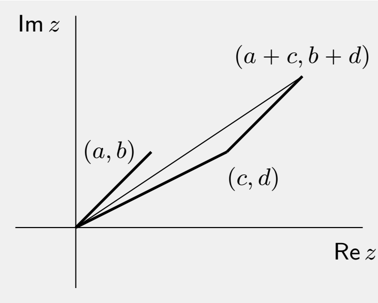
Jede komplexe Zahl $z = a + bi$ *entspricht genau einem geordneten Paar* (a, b) *reeller Zahlen, mit* $a =: \mathit{Re}\,(z)$ *und* $b =: \mathit{Im}\,(z)$.
Dieses Paar (a, b) *wird als Punkt in der Ebene dargestellt. Dabei sind* (a, b) *die* KARTESISCHEN KOORDINATEN *von* $z = a + bi$.

3.3 (Addition in $\mathbb{C}$)

Die Addition $z + w$ *komplexer Zahlen entspricht der "Vektoraddition":*

$$(a + bi) + (c + di) = (a + c) + (b + d)i \quad \sim \quad (a, b) + (c, d) = (a + c, b + d)$$





Es sei $z = a + bi$ eine komplexe Zahl, $a, b \in \mathbb{R}$.

- $a = \text{Re}\,(z)$ heißt REALTEIL von z,
  $b = \text{Im}\,(z)$ heißt IMAGINÄRTEIL von z.

  Zwei komplexe Zahlen sind gleich, wenn Realteil und Imaginärteil übereinstimmen.
- $\overline{z} = a - bi$ heißt die KONJUGIERT KOMPLEXE Zahl (sprich "z-quer"). Es gilt:

$$\text{Re}\,(z) = \frac{1}{2}(z + \overline{z}), \quad \text{Im}\,(z) = \frac{1}{2i}(z - \overline{z})$$

- $|z| = \sqrt{a^2 + b^2}$ heißt der ABSOLUTBETRAG von z. Es gilt $|z| = \sqrt{z\overline{z}}$.
- Es ist $\overline{\overline{z}} = z$ und $|z| = |\overline{z}|$.
- Für $z \neq 0$ ist $|z| > 0$. Durch die Gleichung

$$z \cdot \frac{\overline{z}}{|z|^2} = 1 \quad (= 1 + 0i)$$

  ist der Kehrwert von $z \neq 0$ definiert als $z^{-1} := \frac{1}{z} := \frac{\overline{z}}{|z|^2}$.

Bemerkung: $\mathbb{R} \subset \mathbb{C}$ wird durch die Identität $a = a + 0i$ geklärt. Also ist $z \in \mathbb{C}$ genau dann reell, wenn $z = \overline{z}$ gilt.

Satz 3.4 ($\mathbb{C}$ ist ein Körper)

Die Menge $\mathbb{C}$ der komplexen Zahlen mit der Addition und Multiplikation aus Definition 3.1 ist ein Körper.

Das heißt im Einzelnen:

- es gelten die Kommutativ-, Assoziativ- und Distributivgesetze,
- das neutrale Element der Addition ist $0 = 0 + 0i$, das der Multiplikation ist $1 = 1 + 0i$,
- zu $z = a + bi$ definiert man $-z := (-1)z = -a - bi$. Damit gilt

$$z + (-z) = 0.$$

Falls $z \neq 0$, so gilt für

$$z^{-1} = \frac{1}{z} = \frac{\overline{z}}{|z|^2} = \frac{a}{a^2 + b^2} - \frac{b}{a^2 + b^2} i.$$

die Relation

$$z\, z^{-1} = 1.$$

Beispiele

(a) Berechnen von Real- und Imaginärteil:

$$z = \frac{3+i}{(1-3i)^2} = \frac{3+i}{1-6i-9} = \frac{3+i}{-8-6i}$$

(b) Es gilt $|z| = 1 \;\Leftrightarrow\; \frac{1}{z} = \overline{z}$.

Satz 3.5 (Rechenregeln in $\mathbb{C}$)

Für alle $z, w \in \mathbb{C}$ *gilt*

(a) $0 \cdot z = z \cdot 0 = 0$ *und* $(zw = 0 \;\Leftrightarrow\; z = 0 \vee w = 0)$.

(b) $\overline{z + w} = \overline{z} + \overline{w}$, $\overline{zw} = \overline{z}\,\overline{w}$, $\overline{\left(\frac{z}{w}\right)} = \frac{\overline{z}}{\overline{w}}$, *falls* $w \neq 0$.

(c) $|z| = 0 \iff z = 0$.

(d) $|zw| = |z|\,|w|$, $\left|\frac{z}{w}\right| = \frac{|z|}{|w|}$, *falls* $w \neq 0$.

(e) $|z + w| \leqslant |z| + |w|$ *(Dreiecksungleichung)*

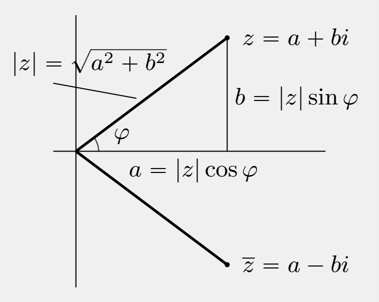
(f) $|z \pm w| \geqslant \big||z| - |w|\big|$ *(umgekehrte Dreiecksungleichung)*

Für die Multiplikation, Division, Potenzen und Wurzeln eignet sich eine andere geometrische Beschreibung der komplexen Zahlen besser.

3.6 (Polarkoordinaten)

Die komplexe Zahl $z = a + bi$ *hat die* POLARKOORDINATEN

- $r = |z| = \sqrt{a^2 + b^2}$ BETRAG
  Das ist der Abstand zu 0.
- $\varphi = \mathit{Arg}(z)$ HAUPTWERT DES ARGUMENTS
  Das ist der Winkel zwischen der positiven reellen Achse und der Strecke von 0 *zu* z. *Dabei ist* $-\pi < \varphi \leqslant \pi$.



Die Darstellung von z *in Polarkoordinaten lautet*

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

Für $z = 0$ *ist* φ *unbestimmt.*

- Alle Winkel werden im BOGENMASS gemessen, wobei π dem Winkel 180^o entspricht. Umrechnung:

$$\alpha^o \mathrel{\widehat{=}} \frac{\alpha\pi}{180}$$

- Für jedes $z \in \mathbb{C}$ und jedes $k \in \mathbb{Z}$ gilt

$$|z|(\cos(\varphi) + i\sin(\varphi)) = |z|(\cos(\varphi + 2k\pi) + i\sin(\varphi + 2k\pi)).$$

  Der Winkel ist nur bis auf Vielfache von 2π eindeutig. Der Hauptwert des Arguments, $\text{Arg}(z) \in (-\pi, \pi]$, ist eindeutig bestimmt.

- Die Zuordnung

$$(a, b) \quad \longleftrightarrow \quad (r, \varphi)$$

  ist für $(a, b) \neq (0, 0)$ und $\varphi \in (-\pi, \pi]$ eineindeutig. Die Umrechnungen lauten

$$a = r\cos\varphi, \qquad b = r\sin\varphi$$

  sowie

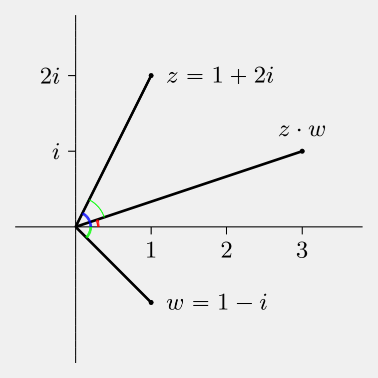
$$\begin{aligned} r &= \sqrt{a^2 + b^2} \\ \varphi &= \begin{cases} \arccos(a/r), & \text{falls } b \geqslant 0, \\ -\arccos(a/r), & \text{falls } b < 0. \end{cases} \end{aligned}$$

Satz 3.7 (Rechenoperationen in $\mathbb{C}$)

Für $z, w \in \mathbb{C}$ mit $z = |z|(\cos(\varphi) + i\sin(\varphi))$ und $w = |w|(\cos(\psi) + i\sin(\psi))$ gilt:

- *Multiplikation:*
$$z \cdot w = |z|\;|w|\,(\cos(\varphi+\psi) + i\sin(\varphi+\psi))\,.$$
(Multiplikation der Beträge und Addition der Winkel)
- *Division für $w \neq 0$:*
$$\frac{z}{w} = \frac{|z|}{|w|}\,(\cos(\varphi-\psi) + i\sin(\varphi-\psi))\,.$$
(Division der Beträge und Subtraktion der Winkel)
- *Konjugation:*
$$\overline{z} = |z|\,(\cos\varphi - i\sin\varphi) = |z|\,(\cos(-\varphi) + i\sin(-\varphi))\,.$$
(Spiegelung an der reellen Achse)

<u>Achtung:</u> Im Allgemeinen gilt $\mathrm{Arg}(zw) \neq \mathrm{Arg}(z) + \mathrm{Arg}(w)$.



Für $z = 1 + 2i$ und $w = 1 - i$ gilt

$$z \cdot w = (1 + 2i)(1 - i) = 3 + i$$

Polarkoordinaten:

$|z| = \sqrt{5}$, $\varphi_z \approx 63^\circ$

$|w| = \sqrt{2}$, $\varphi_w = -45^\circ$

$|zw| = \sqrt{10}$, $\varphi_{zw} \approx 18^\circ$

3.8 (Eulersche Formel)

Wir setzen zunächst nur formal (als Kurzschreibweise)

$$e^{i\varphi} := \cos\varphi + i\sin\varphi,$$

wobei e die Eulersche Zahl ist.

Dann ist durch

$$z = |z|e^{i\varphi}, \quad w = |w|e^{i\psi} \quad \Longrightarrow \quad z \cdot w = |z|\,|w|e^{i(\varphi+\psi)}$$

bereits eine Eigenschaft der "Exponentialfunktion" ausgedrückt, die wir später noch allgemein herleiten werden. Die Exponential-Schreibweise erleichtert den Umgang mit den Polarkoordinaten (es gelten die üblichen Potenzgesetze).

Als direkte Folgerung der Multiplikationsregel ergibt sich:

Satz 3.9 (Moivresche Formel)

Für $z = |z|(\cos\varphi + i\sin\varphi) = |z|e^{i\varphi}$ *und* $n \in \mathbb{N}$ *gilt*

$$z^n = |z|^n \big(\cos(n\varphi) + i\sin(n\varphi)\big) = |z|^n \; e^{in\varphi}.$$

Die komplexen Zahlen vom Betrag 1 haben die Form $e^{i\varphi} = \cos\varphi + i\sin\varphi$. Für spezielle Winkel ergeben sich die sogenannten Einheitswurzeln.

Satz 3.10 (Die n-ten Einheitswurzeln)

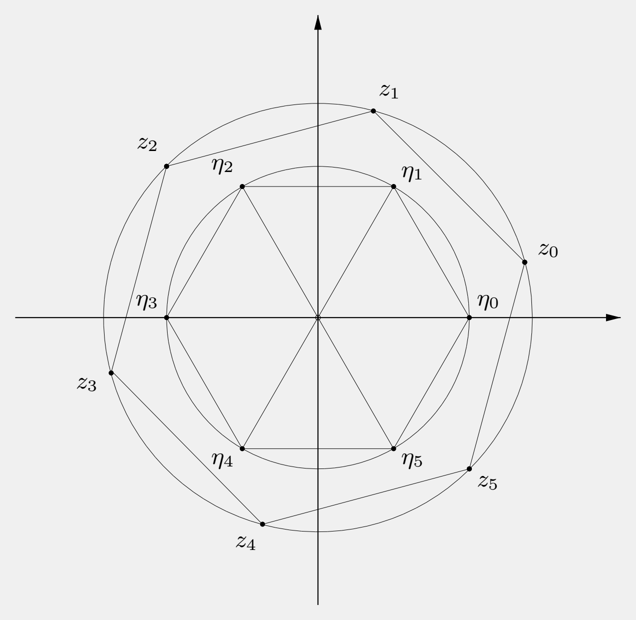
Sei n eine natürliche Zahl. Die komplexen Zahlen

$$\eta_k := e^{i2\pi k/n} = \cos\frac{2\pi k}{n} + i\sin\frac{2\pi k}{n}, \quad k = 0, 1, \dots, n-1,$$

heißen die n-ten Einheitswurzeln; durch sie sind sämtliche komplexe Lösungen der Gleichung

$$z^n = 1$$

gegeben.



Die sechsten Einheitswurzeln η_0 bis η_5 und die Lösungen von $z^6 = 8i$.

Satz 3.11 (Die n-ten Wurzeln in $\mathbb{C}$)

Seien $w = |w|(\cos\varphi + i\sin\varphi) \neq 0$ *und* $n \in \mathbb{N}$. *Dann sind sämtliche Lösungen der Gleichung* $\boxed{z^n = w}$ *gegeben durch die* n *komplexen Zahlen*

$$z_k = |w|^{1/n}\left(\cos\left(\frac{\varphi}{n} + \frac{2\pi k}{n}\right) + i\sin\left(\frac{\varphi}{n} + \frac{2\pi k}{n}\right)\right) \text{ mit } k = 0, 1, \ldots, n-1.$$

Ist z *eine beliebige Zahl mit* $z^n = w$, *so sind alle Lösungen durch* $z_k = z \cdot \eta_k$ *gegeben.*

<u>Schreibweise:</u> Die *komplexe* Wurzel $\sqrt[n]{w}$ oder $w^{1/n}$ bezeichnet die Gesamtheit aller n verschiedenen n-ten Wurzeln von w. (Im Gegensatz zum Reellen: $\sqrt{9} = 3$, und nicht -3.)

Beispiel: Die 3-ten Wurzeln von $w = i = e^{i\pi/2}$ sind

$$z_0 = e^{i\pi/6} = \frac{\sqrt{3}}{2} + \frac{1}{2}i,$$

$$z_1 = e^{i5\pi/6} = -\frac{\sqrt{3}}{2} + \frac{1}{2}i,$$

$$z_2 = e^{i3\pi/2} = e^{-i\pi/2} = -i.$$

Polynome auf $\mathbb{C}$

Definition 3.12 (Polynom)

Sein $n \in \mathbb{N}_0$. Eine Funktion $P : \mathbb{C} \to \mathbb{C}$ mit der Zuordnungsvorschrift

$$x \mapsto P(x) := a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 = \sum_{k=0}^{n} a_k x^k, \quad a_k \in \mathbb{C}$$

heißt POLYNOM.

- *Die a_k mit $0 \leqslant k \leqslant n$ nennt man die* KOEFFIZIENTEN *von P.*
- *Sind alle Koeffizienten reell, so nennt man P ein* REELLES POLYNOM.

Als Definitions- und Wertebereich sind auch $\mathbb{R}$ sowie Teilmengen von $\mathbb{C}$ oder $\mathbb{R}$ zugelassen.

- a_n heißt LEITKOEFFIZIENT oder HÖCHSTKOEFFIZIENT von P.
- Falls $a_n \neq 0$, so hat P den (exakten) GRAD n.
- Ist $a_n = 1$, so heißt P NORMIERT.

Bemerkung: Für das Nullpolynom $P(x) = 0$ ist der Grad als $-\infty$ definiert.

Satz 3.13 (Fundamentalsatz der Algebra)

Jedes Polynom P vom Grad $n \geqslant 1$ hat mindestens eine NULLSTELLE *in $\mathbb{C}$, d.h. es gibt ein $z \in \mathbb{C}$ mit $P(z) = 0$.*

Als Folgerung ergibt sich:

Satz 3.14 (Zerlegung in Linearfaktoren)

Jedes Polynom P vom Grad $n \geqslant 1$ läßt sich schreiben als Produkt von n LINEARFAKTOREN

$$P(z) = a_n(z - z_1)(z - z_2) \cdots (z - z_n).$$

Dabei sind die $z_1, z_2, \ldots, z_n \in \mathbb{C}$ die Nullstellen von P.

- Fasst man gleiche Nullstellen zusammen, so ergibt sich die Produktform

$$P(z) = a_n(z - z_1)^{m_1} \cdots (z - z_r)^{m_r}$$

mit den paarweise verschiedenen Nullstellen $z_1, \ldots, z_r$ von P. Die Zahl $m_k \in \mathbb{N}$ heißt ORDNUNG der Nullstelle z_k, und es ist $\sum_{k=1}^{r} m_k = n$.

3.15 (Horner-Schema)

Die Auswertung des Polynoms

$$P(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0$$

an einer Stelle z erfolgt nach dem Horner-Schema:

$$P(z) = (\cdots((a_n z + a_{n-1})z + a_{n-2})z + \cdots)z + a_0.$$

Das Horner-Schema liefert eine Kurzform zur Polynomdivision durch den Linearfaktor $(z - z_0)$:

- Zum gegebenen Polynom $P(z) = a_n z^n + a_{n-1} z^{n-1} + \cdots + a_1 z + a_0$ und zur Stelle z_0 ist die Division mit Rest, also die Darstellung

$$P(z) = (z - z_0) \cdot (b_n z^{n-1} + b_{n-1} z^{n-2} + \cdots + b_2 z + b_1) + b_0$$

gesucht. Die Koeffizienten $b_0, \ldots, b_n$ liest man aus der letzten Zeile des Hornerschemas ab:

	a_n	a_{n-1}	a_{n-2}	$\cdots$	a_1	a_0
z_0	$-$	$z_0 \cdot b_n$	$z_0 \cdot b_{n-1}$	$\cdots$	$z_0 \cdot b_2$	$z_0 \cdot b_1$
	$b_n = a_n$	b_{n-1}	b_{n-2}	$\cdots$	b_1	$b_0 = P(z_0)$

Hierbei ist $b_k = z_0 \cdot b_{k+1} + a_k$.

- Falls z_0 eine Nullstelle von P ist, so gilt $b_0 = 0$ und wir erhalten die “Abspaltung” des Linearfaktors $(z - z_0)$

$$P(z) = (z - z_0) \cdot (b_n z^{n-1} + b_{n-1} z^{n-2} + \cdots + b_2 z + b_1)$$

Kapitel 4 – Vektoren

4 Vektoren

4.1 (Kartesische Koordinaten in Ebene und Raum)

In der Ebene (mathematisch ist dies die Menge $\mathbb{R}^2$) ist ein KARTESISCHES KOORDINATENSYSTEM *festgelegt durch den Nullpunkt $\mathbf{0}$ sowie zwei Zahlengeraden (die x- und die y-Achse), die sich senkrecht im Nullpunkt schneiden. Ein Punkt P im $\mathbb{R}^2$ hat als x- und y-Koordinate jeweils den Wert, der sich durch orthogonale Projektion auf die entsprechende Achse ergibt.*

Im dreidimensionalen Raum (mathematisch ist dies die Menge $\mathbb{R}^3$) ist ein KARTESISCHES KOORDINATENSYSTEM *festgelegt durch den Nullpunkt $\mathbf{0}$ sowie drei Zahlengeraden (die x-, y- und z-Achse), die sich im Nullpunkt schneiden, paarweise senkrecht stehen und ein* RECHTSSYSTEM *bilden. Ein Punkt P im $\mathbb{R}^3$ hat als x-, y- und z-Koordinate jeweils den Wert, der sich durch orthogonale Projektion auf die entsprechende Achse ergibt.*

Definition 4.2 (Vektor im $\mathbb{R}^2$)

Einem Punkt P im $\mathbb{R}^2$ mit der x-Koordinate a_1 und y-Koordinate a_2 entspricht der VEKTOR

$$\vec{a} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} =: (a_1, a_2)^\top.$$

Geometrische Interpretation: Der Vektor $\vec{a}$ lässt sich (bis auf Parallelverschiebung) als die gerichtete Verbindungsstrecke vom Nullpunkt $\mathbf{0}$ zum Punkt P interpretieren. Der Vektor $\vec{a}$ ist festgelegt durch

- den BETRAG $|\vec{a}| \geqslant 0$ ($=$ Länge des Vektors)
- die RICHTUNG und den RICHTUNGSSINN

Sonderfall: Der NULLVEKTOR $\vec{0} := (0,0)^\top$ hat den Betrag 0, seine Richtung ist nicht definiert.

Analoges gilt für Vektoren im $\mathbb{R}^3$ mit

$$\vec{a} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} =: (a_1, a_2, a_3)^\top \quad \text{und} \quad \vec{0} := \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

4.3 (Addition, Multiplikation mit einem Skalar)

- *Wird der Vektor $\vec{a}$ dargestellt durch die gerichtete Strecke von P nach Q, und der Vektor $\vec{b}$ durch die gerichtete Strecke von Q nach R, so ist die Summe $\vec{a}+\vec{b}$ der Vektor, der durch die gerichtete Strecke von P nach R dargestellt wird. (*DIAGONALREGEL*)*
- *Sei $\vec{a}$ ein Vektor und $\alpha \in \mathbb{R}$.*
    - *Ist $\alpha > 0$, so ist $\alpha\vec{a}$ derjenige Vektor, der dieselbe Richtung und denselben Richtungssinn wie $\vec{a}$ hat und dessen Betrag $\alpha|\vec{a}|$ ist.*
    - *Ist $\alpha < 0$, so ist $\alpha\vec{a}$ derjenige Vektor, der dieselbe Richtung und den entgegengesetzten Richtungssinn wie $\vec{a}$ hat und dessen Betrag $|\alpha||\vec{a}|$ ist.*
    - *Ist $\alpha = 0$, so ist $\alpha\vec{a} = \vec{0}$ der Nullvektor.*

4.4 (Rechnen mit Vektoren im kartesischen Koordinatensystem)

- *Die gewählten Koordinatenachsen definieren die* EINHEITSVEKTOREN

$$\vec{e}_1 = (1,0,0)^\top, \quad \vec{e}_2 = (0,1,0)^\top, \quad \vec{e}_3 = (0,0,1)^\top$$

mit Anfangspunkt $\mathbf{0}$ *und Endpunkt bei der Längeneinheit auf der jeweiligen Koordinatenachse.*

- *Summe und Multiplikation mit Skalaren:*
Seien $\vec{a} = (a_1, a_2, a_3)^\top$, $\vec{b} = (b_1, b_2, b_3)^\top$ *und* $\alpha \in \mathbb{R}$. *Dann gilt*

$$\vec{a} + \vec{b} = \vec{b} + \vec{a} = (a_1 + b_1, a_2 + b_2, a_3 + b_3)^\top$$
$$\alpha\vec{a} = (\alpha a_1, \alpha a_2, \alpha a_3)^\top$$

Der Vektor $\vec{a} = (a_1, a_2, a_3)^\top$ lässt sich als folgende Vektorsumme ausdrücken:

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

Bemerkung: Analoges gilt für Vektoren im $\mathbb{R}^2$.

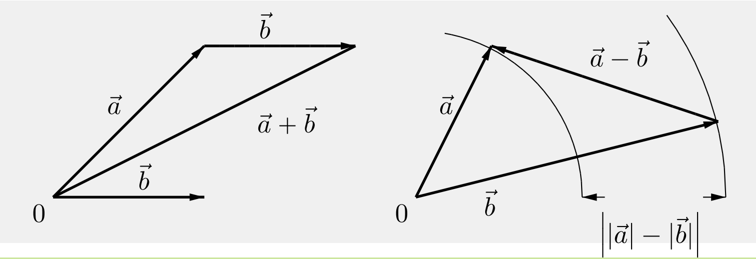
- Für den Betrag ergibt sich:

$$\vec{a} = (a_1, a_2, a_3)^\top \implies |\vec{a}\,| = \sqrt{a_1^2 + a_2^2 + a_3^2}$$
$$\vec{a} = (a_1, a_2)^\top \implies |\vec{a}\,| = \sqrt{a_1^2 + a_2^2}$$

- $|\alpha\vec{a}\,| = |\alpha|\,|\vec{a}\,|$ sowie $\alpha\vec{a} = \vec{0} \iff \alpha = 0 \,\vee\, \vec{a} = \vec{0}.$

- $|\vec{a} + \vec{b}| \leqslant |\vec{a}\,| + |\vec{b}|$ (1. Dreiecksungleichung)
  $\Big||\vec{a}\,| - |\vec{b}|\Big| \leqslant |\vec{a} - \vec{b}|$ (2. Dreiecksungleichung)

- Vektoren der Länge $|\vec{a}\,| = 1$ heißen Einheitsvektoren. Zu beliebigem $\vec{a} \neq \vec{0}$ erhält man durch "Normierung" den Einheitsvektor $\vec{b} = \frac{1}{|\vec{a}\,|}\,\vec{a}$.



Definition 4.5 (Winkel und Skalarprodukt)

Seien $\vec{a}, \vec{b} \neq 0$ *zwei Vektoren, die wir jeweils als gerichtete Strecke mit gemeinsamem Anfangspunkt* P *darstellen. Dann ist*

$$0 \leqslant \angle(\vec{a}, \vec{b}) \leqslant \pi$$

der von den Vektoren $\vec{a}$ *und* $\vec{b}$ EINGESCHLOSSENE WINKEL. *Mit*

$$\vec{a} \cdot \vec{b} := |\vec{a}||\vec{b}| \cos \angle(\vec{a}, \vec{b})$$

wird das SKALARPRODUKT *der Vektoren* $\vec{a}$ *und* $\vec{b}$ *definiert. Für das Skalarprodukt mit dem Nullvektor setzt man*

$$\vec{0} \cdot \vec{a} = \vec{a} \cdot \vec{0} := 0$$

Satz 4.6 (Berechnungsformel für das Skalarprodukt)

$$\vec{a} = (a_1, a_2, a_3)^\top, \vec{b} = (b_1, b_2, b_3)^\top \quad \Rightarrow \quad \vec{a} \cdot \vec{b} = a_1 b_1 + a_2 b_2 + a_3 b_3$$

$$\vec{a} = (a_1, a_2)^\top, \vec{b} = (b_1, b_2)^\top \quad \Rightarrow \quad \vec{a} \cdot \vec{b} = a_1 b_1 + a_2 b_2$$

- Kommutativ- und Distributivgesetz(e)

$$\vec{a} \cdot \vec{b} = \vec{b} \cdot \vec{a}, \qquad \vec{a} \cdot (\vec{b} + \vec{c}) = \vec{a} \cdot \vec{b} + \vec{a} \cdot \vec{c}, \qquad (\vec{a} + \vec{b}) \cdot \vec{c} = \vec{a} \cdot \vec{c} + \vec{b} \cdot \vec{c}$$

- Assoziativität bezüglich Multiplikation mit Skalaren $\alpha \in \mathbb{R}$

$$(\alpha \vec{a}) \cdot \vec{b} = \alpha(\vec{a} \cdot \vec{b}) = \vec{a} \cdot (\alpha \vec{b})$$

Vorsicht: $(\vec{a} \cdot \vec{b})\vec{c}$ ist ein skalares Vielfaches des Vektors $\vec{c}$, während $(\vec{b} \cdot \vec{c})\vec{a}$ ein skalares Vielfaches des Vektors $\vec{a}$ ist. Im Allgemeinen gilt also

$$(\vec{a} \cdot \vec{b})\vec{c} \neq (\vec{b} \cdot \vec{c})\vec{a}$$

- Es gilt $|\vec{a}| = \sqrt{\vec{a} \cdot \vec{a}}$ und

$$|\vec{a} \cdot \vec{b}| \leqslant |\vec{a}|\,|\vec{b}| \quad (\text{Cauchy-Schwarz-Ungleichung})$$

Die Gleichheit $|\vec{a} \cdot \vec{b}| = |\vec{a}|\,|\vec{b}|$ gilt genau dann, wenn die Vektoren $\vec{a}, \vec{b}$ KOLLINEAR sind, d.h. wenn eine Zahl $\alpha \in \mathbb{R}$ existiert mit

$$\vec{b} = \alpha \vec{a} \quad \vee \quad \vec{a} = \alpha \vec{b}$$

- Für die Einheitsvektoren $\vec{e}_1, \vec{e}_2, \vec{e}_3$ gilt

$$\vec{e}_i \cdot \vec{e}_j = \begin{cases} 1 & \text{für } i = j \\ 0 & \text{für } i \neq j \end{cases}$$

- Die Dreiecksungleichung $|\vec{a} + \vec{b}| \leqslant |\vec{a}| + |\vec{b}|$ folgt direkt aus der Cauchy-Schwarz-Ungleichung:

$$\begin{aligned} |\vec{a} + \vec{b}|^2 &= (\vec{a} + \vec{b}) \cdot (\vec{a} + \vec{b}) = \vec{a} \cdot \vec{a} + \vec{a} \cdot \vec{b} + \vec{b} \cdot \vec{a} + \vec{b} \cdot \vec{b} \\ &= |\vec{a}|^2 + 2(\vec{a} \cdot \vec{b}) + |\vec{b}|^2 \leqslant |\vec{a}|^2 + 2|\vec{a}|\,|\vec{b}| + |\vec{b}|^2 \\ &= (|\vec{a}| + |\vec{b}|)^2. \end{aligned}$$

Definition 4.7 (Orthogonalität von Vektoren)

Zwei Vektoren $\vec{a}, \vec{b} \neq \vec{0}$ *sind zueinander* ORTHOGONAL, *wenn*

$$\vec{a} \cdot \vec{b} = 0$$

gilt. Wir schreiben dann $\vec{a} \perp \vec{b}$.

Für Vektoren $\vec{a}, \vec{b} \neq \vec{0}$ *ist damit äquivalent:*

$$\vec{a} \perp \vec{b} \quad \Longleftrightarrow \quad \angle(\vec{a}, \vec{b}) = \frac{\pi}{2}$$

Beispiel 4.8

Die Diagonalen eines Parallelogramms sind genau dann zueinander orthogonal, wenn alle vier Seiten die gleiche Länge haben.

Die Distributivgesetze ergeben zwei wichtige Rechenregeln:

Satz 4.9

a) *Für Vektoren $\vec{a}, \vec{b}$ mit $\vec{a} \perp \vec{b}$ gilt*

$$|\vec{a} + \vec{b}|^2 = |\vec{a}|^2 + |\vec{b}|^2 \qquad \text{(Satz des Pythagoras)}$$

b) *Für alle Vektoren $\vec{a}, \vec{b}$ gilt*

$$|\vec{a} + \vec{b}|^2 + |\vec{a} - \vec{b}|^2 = 2\left(|\vec{a}|^2 + |\vec{b}|^2\right) \qquad \text{(Parallelogrammgleichung)}$$

sowie

$$\vec{a} \cdot \vec{b} = \frac{1}{4}\left(|\vec{a} + \vec{b}|^2 - |\vec{a} - \vec{b}|^2\right) \qquad \text{(Polarisationsformel)}$$

Wir definieren zwei weitere Produkte für Vektoren im $\mathbb{R}^3$.

Definition 4.10 (Vektorprodukt)

Seien $\vec{a}, \vec{b} \in \mathbb{R}^3 \backslash \{\vec{0}\}$. Das VEKTORPRODUKT $\vec{a} \times \vec{b}$ *von $\vec{a}$ und $\vec{b}$ ist derjenige Vektor im $\mathbb{R}^3$,*

1. *für dessen Betrag gilt:*

$$|\vec{a} \times \vec{b}| = |\vec{a}|\,|\vec{b}|\,\sin\angle(\vec{a}, \vec{b}),$$

2. *der orthogonal zu $\vec{a}$ und $\vec{b}$ ist,*
3. *dessen Richtungssinn so gewählt ist, dass im Fall $|\vec{a} \times \vec{b}| \neq 0$ die Vektoren $\vec{a}, \vec{b}, \vec{a} \times \vec{b}$ ein Rechtssystem bilden.*

Zusätzlich setzt man $\vec{a} \times \vec{0} = \vec{0} \times \vec{a} := \vec{0}$.

Satz 4.11

- *Für alle $\vec{a}, \vec{b} \in \mathbb{R}^3$ gilt $\vec{a} \times \vec{b} = -\vec{b} \times \vec{a}$*
- $\vec{e}_1 \times \vec{e}_2 = \vec{e}_3$, $\vec{e}_2 \times \vec{e}_3 = \vec{e}_1$ *und* $\vec{e}_3 \times \vec{e}_1 = \vec{e}_2$

Bemerkungen:

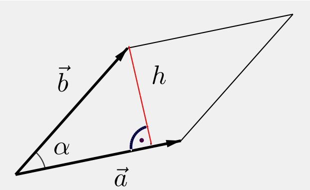
- Es ist $\vec{a} \times \vec{a} = \vec{0}$
- Für $\vec{a} = (a_1, a_2, a_3)^\top$ und $\vec{b} = (b_1, b_2, b_3)^\top$ erhält man

$$\vec{a} \times \vec{b} = (a_2 b_3 - a_3 b_2,\ a_3 b_1 - a_1 b_3,\ a_1 b_2 - a_2 b_1)^\top$$

- Einfaches Nachrechnen ergibt

$$|\vec{a} \times \vec{b}|^2 + (\vec{a} \cdot \vec{b})^2 = |\vec{a}|^2 |\vec{b}|^2$$

- Der Betrag $|\vec{a} \times \vec{b}|$ ist der Flächeninhalt des Parallelogramms mit den Seiten $\vec{a}$ und $\vec{b}$.
  Der Vektor $\vec{w} := \vec{a} \times \vec{b}$ ist senkrecht (orthogonal) zu diesem Parallelogramm, erfüllt also die Orthogonalitätsbedingungen



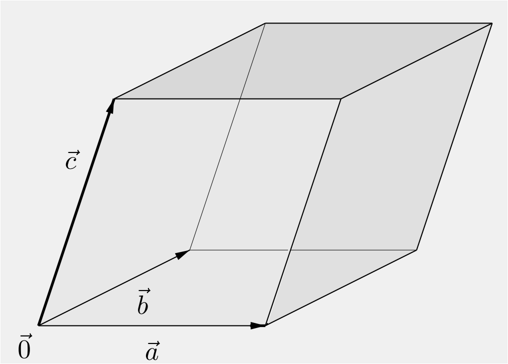
$$\vec{w} \perp \vec{a}, \qquad \vec{w} \perp \vec{b}.$$

- Merkregel: $\vec{c} = \vec{a} \times \vec{b}$ mit "Multiplikation über Kreuz" berechnen.

$$\begin{array}{ccl}
a_1 & b_1 & \\
a_2 & b_2 & \\
 & & c_1 = a_2 b_3 - a_3 b_2 \\
a_3 & b_3 & \\
 & & c_2 = a_3 b_1 - a_1 b_3 \\
a_1 & b_1 & \\
 & & c_3 = a_1 b_2 - a_2 b_1 \\
a_2 & b_2 &
\end{array}$$

Definition 4.12 (Spatprodukt)

Für Vektoren $\vec{a}, \vec{b}, \vec{c} \in \mathbb{R}^3$ *definiert man das* SPATPRODUKT $\vec{a} \cdot (\vec{b} \times \vec{c})$.
Das Spatprodukt ist ein Skalar, dessen Absolutbetrag das Volumen des von den drei Vektoren $\vec{a}, \vec{b}, \vec{c}$ *aufgespannten Spates (auch Parallelepiped genannt) angibt.*



Kapitel 5 – Geraden und Ebenen

Kap. 5: Geraden und Ebenen

Wir behandeln zunächst Geraden im $\mathbb{R}^2$ und $\mathbb{R}^3$. Sei also $n = 2$ oder $n = 3$.

Definition 5.1 (Parameterdarstellung von Geraden)

Die GERADE *durch den Punkt* $\vec{p} \in \mathbb{R}^n$ *mit dem* RICHTUNGSVEKTOR $\vec{v} \in \mathbb{R}^n$, $\vec{v} \neq \vec{0}$, *ist gegeben durch die Menge*

$$G = \{\vec{x} \in \mathbb{R}^n \mid \vec{x} = \vec{p} + t\vec{v},\ t \in \mathbb{R}\}.$$

Die Zahl t *zum Punkt* $\vec{x} = \vec{x}(t) = \vec{p} + t\vec{v} \in G$ *heißt der* PARAMETERWERT *von* $\vec{x}$.

- Um die Parameterdarstellung der Geraden G durch die Punkte $\vec{p}$ und $\vec{q}$ (mit $\vec{p} \neq \vec{q}$) zu bestimmen, berechnet man den Richtungsvektor $\vec{v} = \vec{q} - \vec{p}$.

Satz 5.2 (Lot auf eine Gerade, Abstand Punkt-Gerade)

Die Gerade $G \subset \mathbb{R}^n$ *habe die Parameterdarstellung* $\vec{x}(t) = \vec{p} + t\vec{v}$, $t \in \mathbb{R}$ *(mit* $\vec{v} \neq \vec{0}$*). Sei* $\vec{q} \in \mathbb{R}^n$ *ein beliebiger Punkt.*

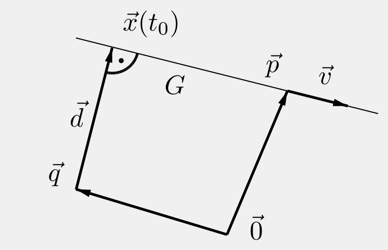
- *Es gibt genau einen Punkt* $\vec{x}(t_0) \in G$*, so dass*

  $$\vec{q} - \vec{x}(t_0) \perp \vec{v}$$

  gilt. Dieser Punkt $\vec{x}(t_0)$ *heißt der* FUSSPUNKT DES LOTS *von* $\vec{q}$ *auf* G.

- $|\vec{q} - \vec{x}(t_0)|$ *ist der Abstand des Punktes* $\vec{q}$ *von* G*, d.h.*

  $$\textit{dist}\,(\vec{q}, G) := |\vec{q} - \vec{x}(t_0)| < |\vec{q} - \vec{x}(t)| \quad \textit{für alle} \quad t \neq t_0.$$



Beispiel:

Gerade G durch die Punkte $(5,1,-1)^\top$ und $(3,-3,3)^\top$; Lot vom Punkt $\vec{q}=(0,0,0)^\top$ auf G (ergibt den Abstand von G zum Nullpunkt)

Spezielle Darstellung von Geraden im $\mathbb{R}^2$

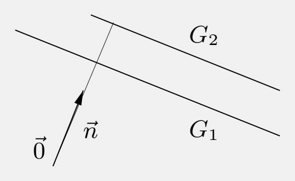
Definition 5.3 (Normalenform, Hesse-Normalform von Geraden im $\mathbb{R}^2$)

Zu gegebenen Zahlen $a, b, c \in \mathbb{R}$ mit $(a,b) \neq (0,0)$ definiert die Menge

$$G = \{(x,y) \in \mathbb{R}^2 \mid ax + by = c\} = \{\vec{x} \in \mathbb{R}^2 \mid \vec{x} \cdot \begin{pmatrix} a \\ b \end{pmatrix} = c\}$$

eine Gerade; der Vektor $(a,b)^\top$ steht senkrecht auf jedem Richtungsvektor $\vec{v}$ von G und heißt Normalenvektor *von G.*

Für $|(a,b)^\top| = 1$ und $c \geqslant 0$ heißt die gegebene Form die Hesse-Normalform *von G. Die Zahl c ist dabei der Abstand von G zum Nullpunkt des $\mathbb{R}^2$. Der* Normalen-Einheitsvektor *$(a,b)^\top$ steht senkrecht auf G.*



$G_1 : \vec{x} \cdot \vec{n} = c$, $G_2 : \vec{x} \cdot \vec{n} = d$
Für $|\vec{n}| = 1$ sind c und d die Abstände der Geraden vom Ursprung, $|c - d|$ ist der Abstand der Geraden.

Die Hesse-Normalform wird bei Abstandsberechnungen eingesetzt:

Beispiel:

Die Gerade $G \subseteq \mathbb{R}^2$ durch die Punkte $\vec{p} = (4, -1)$ und $\vec{q} = (3, 1)$

Ebenen im $\mathbb{R}^3$

Definition 5.4 (Ebenen im $\mathbb{R}^3$)

Die EBENE *durch den Punkt* $\vec{p} \in \mathbb{R}^3$ *mit den nicht-kollinearen* RICHTUNGSVEKTOREN $\vec{v}, \vec{w} \in \mathbb{R}^3$ *ist gegeben durch die Menge*

$$E = \{\vec{x} \in \mathbb{R}^3 \mid \vec{x} = \vec{p} + s\vec{v} + t\vec{w},\ \ s, t \in \mathbb{R}\}.$$

Das Paar (s, t) *zum Punkt* $\vec{x} = \vec{x}(s, t) = \vec{p} + s\vec{v} + t\vec{w} \in E$ *heißt das* PARAMETERPAAR *zum Punkt* $\vec{x}$*.*

- Um die Parameterdarstellung der Ebene E durch drei Punkte $\vec{p}$, $\vec{q}$ und $\vec{r}$ zu bestimmen, die nicht alle auf einer Geraden liegen, berechnet man die Richtungsvektoren $\vec{v} = \vec{q} - \vec{p}$ und $\vec{w} = \vec{r} - \vec{p}$. Diese sind dann NICHT-KOLLINEAR, d.h. keine Vielfache voneinander.

- Die Richtungsvektoren $\vec{v}, \vec{w} \in \mathbb{R}^3$ sind genau dann nicht-kollinear, wenn $\vec{n} := \vec{v} \times \vec{w} \neq \vec{0}$ gilt. $\vec{n}$ ist Normalenvektor der Ebene (d.h. $\vec{n}$ steht senkrecht auf E).

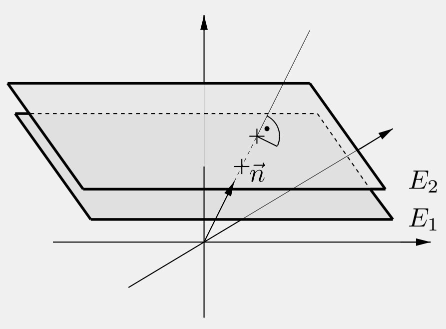
Definition 5.5 (Normalenform, Hesse-Normalform von Ebenen im $\mathbb{R}^3$)

Zu Zahlen $a, b, c, d \in \mathbb{R}$ *mit* $\vec{n}^\top := (a, b, c)^\top \neq (0, 0, 0)^\top$ *definiert die Menge*

$$E = \{(x, y, z)^\top \in \mathbb{R}^3 \mid ax + by + cz = d\} = \{\vec{x} \mid \vec{x} \cdot \vec{n} = d\}$$

eine Ebene; der Vektor $\vec{n}$ *steht senkrecht auf allen Richtungsvektoren* $\vec{v}$ *von* E *und heißt* Normalenvektor *von* E.

Für $|\vec{n}| = 1$ *und* $d \geqslant 0$ *heißt die gegebene Form die* Hesse-Normalform *von* E. *Die Zahl* d *ist dabei der Abstand von* E *zum Nullpunkt des* $\mathbb{R}^3$. *Der* Normalen-Einheitsvektor $\vec{n}$ *steht senkrecht auf* E.



Sei $|\vec{n}| = 1$.

$$E_1 : \vec{x} \cdot \vec{n} = d_1, \quad E_2 : \vec{x} \cdot \vec{n} = d_2$$

Dann ist der Abstand der Ebene E_i zum Ursprung gerade $|d_i|$. Der Abstand der Ebenen ist $|d_1 - d_2|$.

- Die Umwandlung von der Parameterdarstellung in die Hesse-Normalform erfolgt folgendermaßen:
  - Mit dem Normalenvektor $\vec{n} = (a, b, c)^\top = \vec{v} \times \vec{w}$ (beachte $\vec{n} \neq \vec{0}$) und $d := \vec{p} \cdot \vec{n}$ ergibt sich die Normalenform $\vec{x} \cdot \vec{n} = ax + by + cz = d$.
  - Falls $d < 0$, ersetze $\vec{n}$ durch $-\vec{n}$ (und d durch $-d$).
  - Für die Hesse-Normalform ersetzt man $\vec{n}$ durch $\dfrac{1}{|\vec{n}|}\vec{n}$.
- Die Umwandlung der Hesse-Normalform in die Parameterform erfolgt durch Bestimmung dreier Punkte $\vec{p}, \vec{q}, \vec{r} \in E$ (Einsetzen von x-, y- und z-Werten), die nicht auf einer Geraden liegen.
  Alternativ bestimmt man einen Punkt $\vec{p}$ und bestimmt zwei nicht kollineare Richtungsvektoren, die beide auf $\vec{n}$ senkrecht stehen.

Definition 5.6 (Lot auf eine Ebene, Abstand Punkt-Ebene)

Die Ebene $E \subseteq \mathbb{R}^3$ *habe die Parameterdarstellung* $\vec{x}(s,t) = \vec{p} + s\vec{v} + t\vec{w}$, $s,t \in \mathbb{R}$ *(mit* $\vec{v} \times \vec{w} \neq \vec{0}$*). Sei* $\vec{q} \in \mathbb{R}^3$ *ein beliebiger Punkt.*

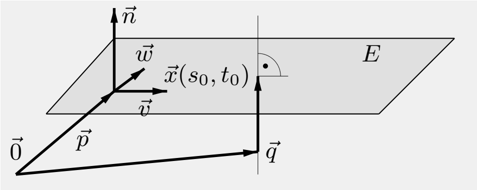
- *Es gibt genau einen Punkt* $\vec{x}(s_0,t_0) \in E$, *so dass*

$$(\vec{q} - \vec{x}(s_0,t_0)) \perp \vec{v} \quad \wedge \quad (\vec{q} - \vec{x}(s_0,t_0)) \perp \vec{w}.$$

  Dieser Punkt $\vec{x}(s_0,t_0)$ *heißt der* FUSSPUNKT DES LOTS *von* $\vec{q}$ *auf die Ebene* E.

- $|\vec{q} - \vec{x}(s_0,t_0)|$ *ist der Abstand des Punktes* $\vec{q}$ *von* E, *d.h.*

$$\textit{dist}\,(\vec{q},E) := |\vec{q} - \vec{x}(s_0,t_0)| < |\vec{q} - \vec{x}(s,t)| \quad \text{für alle} \quad (s,t) \neq (s_0,t_0).$$



Die Schnittmenge zweier Ebenen E_1 und E_2 ist häufig eine Gerade im $\mathbb{R}^3$. Ihre Parameterdarstellung erhält man, indem man die Parameterdarstellung der einen Ebene in die Normalenform der anderen Ebene "einsetzt" und damit einen der Parameter eliminiert.

Definition 5.7 (windschief)

Zwei Geraden G_1 und G_2 im $\mathbb{R}^3$ heißen WINDSCHIEF, *wenn sie sich weder schneiden noch parallel sind.*

Die Geraden seien in Parameterform gegeben:

$$\begin{aligned} G_1: &\quad \vec{x}(t) = \vec{p} + s\vec{v}, &\qquad s \in \mathbb{R}, \\ G_2: &\quad \vec{x}(u) = \vec{q} + t\vec{w}, &\qquad t \in \mathbb{R}. \end{aligned}$$

Dann erfolgt die Berechnung des Abstands durch

- Bestimmung der Lot-Richtung $\vec{n} = \vec{v} \times \vec{w}$ zwischen beiden Geraden
- Bestimmung des Lot-Fußpunkts auf G_2: Schnittpunkt von G_2 mit der Ebene durch $\vec{p}$ mit den Richtungsvektoren $\vec{v}$ und $\vec{n}$.
- Bestimmung des Lot-Fußpunkts auf G_1: Schnittpunkt von G_1 mit der Ebene durch $\vec{q}$ mit den Richtungsvektoren $\vec{w}$ und $\vec{n}$.

Kapitel 6 – Lineare Gleichungssysteme

Kap. 6: Lineare Gleichungssysteme

Wir wollen nun lineare Gleichungssysteme behandeln.

Beispiel: Im $\mathbb{R}^2$ sind z.B. “2 Gleichungen mit 2 Unbekannten” gegeben durch

$$\begin{array}{rcrcr} 2x & - & 3y & = & 1 \\ -x & + & 2y & = & 0 \end{array}$$

Dieses System von 2 Gleichungen hat die eindeutige Lösung $(x,y)^\top = (2,1)^\top$.

Es gibt aber Systeme, die keine Lösung besitzen, wie z.B.

$$\begin{array}{rcrcr} 2x & - & 4y & = & 2 \\ -x & + & 2y & = & 0 \end{array}$$

und auch solche, die unendlich viele Lösungen besitzen:

$$\begin{array}{rcrcr} 2x & - & 4y & = & 2 \\ -x & + & 2y & = & -1 \end{array}$$

Hier sind alle Punkte der Geraden $G: \quad (x,y)^\top = (1+2t,t)^\top$, $t \in \mathbb{R}$, Lösungen.

Definition 6.1 (Lineares Gleichungssystem)

Ein LINEARES GLEICHUNGSSYSTEM *von* m *Gleichungen mit den* n *Unbekannten* $x_1, \dots, x_n$ *ist ein Gleichungssystem der Form*

$$\begin{array}{ccccccccc} a_{1,1}x_1 & + & a_{1,2}x_2 & + & \cdots & + & a_{1,n}x_n & = & b_1 \\ a_{2,1}x_1 & + & a_{2,2}x_2 & + & \cdots & + & a_{2,n}x_n & = & b_2 \\ & & & & \vdots & & & & \\ a_{m,1}x_1 & + & a_{m,2}x_2 & + & \cdots & + & a_{m,n}x_n & = & b_m \end{array}$$

Die $a_{i,k} \in \mathbb{R}$ *heißen die* KOEFFIZIENTEN *und die* $b_i \in \mathbb{R}$ *heißen die* RECHTEN SEITEN *des Gleichungssystems.*

Eine LÖSUNG *des linearen Gleichungssystems ist ein Vektor* $\vec{x} = (x_1, \dots, x_n)^\top \in \mathbb{R}^n$, *dessen Komponenten* x_k *alle* m *Gleichungen erfüllen.*

Das Gleichungssystem heißt

- LÖSBAR (oder KONSISTENT), wenn es mindestens eine Lösung besitzt,
- EINDEUTIG LÖSBAR, wenn es genau eine Lösung besitzt.

Definition 6.2 (Homogenes lineares Gleichungssystem)

Das lineare Gleichungssystem (geschrieben in Kurzform)

$$\sum_{k=1}^{n} a_{i,k} x_k = b_i \qquad (i = 1, \dots, m), \tag{$*$}$$

heißt HOMOGEN, *falls alle rechten Seiten b_i gleich 0 sind. Ansonsten heißt es* INHOMOGEN.

- Ein homogenes lineares Gleichungssystem ist stets lösbar: Es besitzt die TRIVIALE Lösung $\vec{x} = \vec{0} = (0, 0, \dots, 0)^\top$.
- Das "zum linearen Gleichungssystem (*) gehörende homogene System" lautet

$$\sum_{k=1}^{n} a_{i,k} x_k = 0 \qquad (i = 1, \dots, m). \tag{$**$}$$

Die Lösungsmenge eines homogenen linearen Gleichungssystems (**) hat die Struktur eines VEKTORRAUMS: Summen und skalare Vielfache von Lösungen sind selbst wieder Lösungen.

Satz 6.3 (Lösungen homogener linearer Gleichungssysteme)

Sind $\vec{u} = (u_1, \ldots, u_n)^\top$ *und* $\vec{v} = (v_1, \ldots, v_n)^\top$ *Lösungen eines homogenen linearen Gleichungssystems, so ist auch der Vektor*

$$\alpha\vec{u} + \beta\vec{v}$$

mit beliebigen Zahlen $\alpha, \beta \in \mathbb{R}$ *eine Lösung.*

Die Linearität ergibt ein einfaches "Superpositions-Prinzip":

Satz 6.4 (Lösungen inhomogener linearer Gleichungssysteme)

Gegeben sei ein lösbares inhomogenes lineares Gleichungssystem

$$\sum_{k=1}^{n} a_{i,k}x_k = b_i \qquad (i = 1, \ldots, m). \tag{$*$}$$

a) *Sind $\vec{p}$ und $\vec{q}$ Lösungen von (*), so ist $\vec{v} := \vec{p} - \vec{q}$ eine Lösung des zugehörigen homogenen Systems (**).*

b) *Ist $\vec{p}$ eine (spezielle) Lösung von (*), so erhält man alle Lösungen, indem man zu $\vec{p}$ alle Lösungen des zugehörigen homogenen Systems (**) addiert. Aus der Lösungsmenge $\mathbb{L}_h$ des homogenen Systems und der speziellen Lösung $\vec{p}$ ergibt sich also die Lösungsmenge des inhomogenen Systems*

$$\mathbb{L} = \vec{p} + \mathbb{L}_h = \{\vec{x} \in \mathbb{R}^n \mid \vec{x} = \vec{p} + \vec{v},\ \vec{v} \in \mathbb{L}_h\}$$

Zur Lösung linearer Gleichungssysteme verwendet man einfache Äquivalenz-Umformungen des Gleichungssystems.

Satz 6.5 (Elementare Umformungen des LGS)

Die Menge der Lösungen eines linearen Gleichungssystems bleibt unverändert, wenn man

- E1 *die Reihenfolge der Gleichungen vertauscht,*
- E2 *beide Seiten einer Gleichung mit einer Zahl* $\alpha \neq 0$ *multipliziert,*
- E3 *eine Gleichung ersetzt durch die Summe dieser Gleichung und dem Vielfachen einer anderen Gleichung.*
- E4 *Vertauscht man die Reihenfolge der Unbekannten* $x_1, \ldots, x_n$*, setzt also*

$$(y_1, \ldots, y_n) := (x_{\sigma_1}, \ldots, x_{\sigma_n})$$

mit einer Permutation $(\sigma_1, \ldots, \sigma_n)$ *der Zahlen* $(1, \ldots, n)$*, so erhält man die Lösungsmenge des neuen Systems (bzgl.* $\vec{y}$*) aus der Lösungsmenge des alten (bzgl.* $\vec{x}$*) durch entsprechende Vertauschung der Komponenten.*

An drei Beispielen soll erklärt werden, wie man systematisch durch die Äquivalenz-Umformungen (E1)–(E3) sowie die Umformung (E4) eine *reduzierte Stufenform* des Gleichungssystems erhält, um die Lösungsmenge dann leicht zu bestimmen.

Beispiel 1:

$$\left.\begin{array}{rcrcrcrcr} x_1 & + & 4x_2 & + & 2x_3 & - & x_4 & = & 2 \\ 2x_1 & + & 8x_2 & + & x_3 & - & x_4 & = & 3 \\ -x_1 & + & 2x_2 & + & x_3 & & & = & 2 \\ -x_1 & - & 4x_2 & + & x_3 & + & 2x_4 & = & 2 \end{array}\right\} \longrightarrow \text{Kurzform}$$

x_1	x_2	x_3	x_4	r.S.	
1	4	2	-1	2	
2	8	1	-1	3	Elimination mit E3
-1	2	1	0	2	Elimination mit E3
-1	-4	1	2	2	Elimination mit E3
1	4	2	-1	2	
0	0	-3	1	-1	Zeilentausch E1 $(2 \longrightarrow 3)$
0	6	3	-1	4	und Skalierung E2
0	0	3	1	4	
1	4	2	-1	2	
0	1	$\frac{1}{2}$	$-\frac{1}{6}$	$\frac{2}{3}$	
0	0	-3	1	-1	Glück: keine Elim.
0	0	3	1	4	erforderlich, nur E2

$$
\begin{array}{cccc|cl}
1 & 4 & 2 & -1 & 2 & \\
0 & 1 & \frac{1}{2} & -\frac{1}{6} & \frac{2}{3} & \\
0 & 0 & 1 & -\frac{1}{3} & \frac{1}{3} & \\
0 & 0 & 3 & 1 & 4 & \text{Elimination mit E3} \\
\hline
1 & 4 & 2 & -1 & 2 & \\
0 & 1 & \frac{1}{2} & -\frac{1}{6} & \frac{2}{3} & \\
0 & 0 & 1 & -\frac{1}{3} & \frac{1}{3} & \\
0 & 0 & 0 & 2 & 3 & \text{E2} \\
\hline
1 & 4 & 2 & -1 & 2 & \\
0 & 1 & \frac{1}{2} & -\frac{1}{6} & \frac{2}{3} & \\
0 & 0 & 1 & -\frac{1}{3} & \frac{1}{3} & \\
0 & 0 & 0 & 1 & \frac{3}{2} & \\
\hline
\end{array}
$$

Auflösen durch “Rücksubstitution” (von unten nach oben):

$$
\begin{array}{lrcll}
\text{Gl. 4:} & x_4 & = & \frac{3}{2} & \\
\text{Gl. 3:} & x_3 - \frac{1}{3}x_4 & = & \frac{1}{3} & \Longrightarrow \quad x_3 = \frac{5}{6} \\
\text{Gl. 2:} & x_2 + \frac{1}{2}x_3 - \frac{1}{6}x_4 & = & \frac{2}{3} & \Longrightarrow \quad x_2 = \frac{1}{2} \\
\text{Gl. 1:} & x_1 + 4x_2 + 2x_3 - x_4 & = & 2 & \Longrightarrow \quad x_1 = -\frac{1}{6}, \qquad \text{Probe!}
\end{array}
$$

Beispiel 2:

$$\left.\begin{array}{rcrcrcrcr} x_1 & + & 4x_2 & + & 2x_3 & - & x_4 & = & 2 \\ 2x_1 & + & 8x_2 & + & x_3 & - & x_4 & = & 3 \\ -x_1 & - & 4x_2 & + & x_3 & & & = & -1 \\ -x_1 & - & 4x_2 & + & x_3 & + & 2x_4 & = & 2 \end{array}\right\} \longrightarrow \text{Kurzform}$$

x_1	x_2	x_3	x_4	r.S.	
1	4	2	−1	2	
2	8	1	−1	3	Elimination mit E3
−1	−4	1	0	−1	Elimination mit E3
−1	−4	1	2	2	Elimination mit E3
1	4	2	−1	2	
0	0	−3	1	−1	
0	0	3	−1	1	
0	0	3	1	4	

Vertauschung von x_2 und x_3 (=Spaltentausch $(2 \leftrightarrow 3)$):

x_1	x_3	x_2	x_4	r.S.	
1	2	4	−1	2	
0	−3	0	1	−1	Skalierung E2
0	3	0	−1	1	
0	3	0	1	4	

x_1	x_3	x_2	x_4		
1	2	4	-1	2	
0	1	0	$-\frac{1}{3}$	$\frac{1}{3}$	
0	3	0	-1	1	Elimination mit E3
0	3	0	1	4	Elimination mit E3
1	2	4	-1	2	
0	1	0	$-\frac{1}{3}$	$\frac{1}{3}$	
0	0	0	0	0	
0	0	0	2	3	

Vertauschung von x_2 und x_4 (=Spaltentausch $(3 \leftrightarrow 4)$):

x_1	x_3	x_4	x_2	r.S.	
1	2	-1	4	2	
0	1	$-\frac{1}{3}$	0	$\frac{1}{3}$	
0	0	0	0	0	Zeilentausch E1 $(3 \leftrightarrow 4)$
0	0	2	0	3	und Skalierung E2
1	2	-1	4	2	
0	1	$-\frac{1}{3}$	0	$\frac{1}{3}$	
0	0	1	0	$\frac{3}{2}$	
0	0	0	0	0	

- Feststellung: Das Gleichungssystem ist lösbar (konsistent), weil die letzte Gleichung (Nullzeile) lösbar ist.

- Auflösen durch “Rücksubstitution” der Gleichungen 1–3, wobei die Komponente x_2 (aus Spalte 4) als freie Variable verwendet wird:

  Gl. 3: $x_4 = \frac{3}{2}$

  Gl. 2: $x_3 - \frac{1}{3}x_4 = \frac{1}{3} \Longrightarrow x_3 = \frac{5}{6}$

  Gl. 1: $x_1 + 2x_3 - x_4 + 4x_2 = 2 \Longrightarrow x_1 = \frac{11}{6} - 4x_2,$ Probe!

  Die Lösungsmenge ist eine Gerade im $\mathbb{R}^4$, weil <u>ein</u> “freier” Parameter $t = x_2$ vorliegt:

$$\mathbb{L} = \left\{ \vec{x} = \left(\frac{11}{6} - 4t, t, \frac{5}{6}, \frac{3}{2} \right) \mid t \in \mathbb{R} \right\}.$$

  Eine spezielle Lösung ist $\vec{v} = (\frac{11}{6}, 0, \frac{5}{6}, \frac{3}{2})$.

- Die Lösungsmenge des zugehörigen homogenen Systems ist

$$\mathbb{L}_h = \{ \vec{x} = t(-4, 1, 0, 0) \mid t \in \mathbb{R} \} \subseteq \mathbb{R}^4.$$

Beispiel 3: Abändern der rechten Seite des Gleichungssystems

$$\left.\begin{array}{rcrcrcrcr} x_1 & + & 4x_2 & + & 2x_3 & - & x_4 & = & 2 \\ 2x_1 & + & 8x_2 & + & x_3 & - & x_4 & = & 3 \\ -x_1 & - & 4x_2 & + & x_3 & & & = & 2 \\ -x_1 & - & 4x_2 & + & x_3 & + & 2x_4 & = & 2 \end{array}\right\} \longrightarrow \text{Kurzform}$$

führt zu der Stufenform

x_1	x_3	x_4	x_2	r.S.
1	2	-1	4	2
0	1	$-\frac{1}{3}$	0	$\frac{1}{3}$
0	0	1	0	$\frac{3}{2}$
0	0	0	0	3

- Feststellung: Das Gleichungssystem ist nicht lösbar (inkonsistent), weil die letzte Gleichung der Stufenform nicht lösbar ist: Nullkoeffizienten von $x_1, \ldots, x_4$ treffen auf eine rechte Seite ungleich 0.

Die systematische Vorgehensweise führt zu folgendem Resultat:

Satz 6.6 (Gauß-Algorithmus)

Jedes lineare Gleichungssystem von m Gleichungen mit n Unbekannten $x_1, \ldots, x_n$ kann durch endlich viele Umformungen der Form (E1)–(E4) auf die reduzierte Stufenform gebracht werden:

$$\begin{array}{cccccccc|c}
y_1 & y_2 & y_3 & \cdots & y_r & y_{r+1} & \cdots & y_n & \textit{r.S.} \\
\hline
1 & b_{1,2} & b_{1,3} & \cdots & b_{1,r} & b_{1,r+1} & \cdots & b_{1,n} & c_1 \\
0 & 1 & b_{2,3} & \cdots & b_{2,r} & b_{2,r+1} & \cdots & b_{2,n} & c_2 \\
\vdots & \ddots & \ddots & & \vdots & \vdots & & \vdots & \vdots \\
0 & \cdots & 0 & 1 & b_{r-1,r} & b_{r-1,r+1} & \cdots & b_{r-1,n} & c_{r-1} \\
0 & \cdots & 0 & 0 & 1 & b_{r,r+1} & \cdots & b_{r,n} & c_r \\
0 & \cdots & & \cdots & 0 & 0 & \cdots & 0 & c_{r+1} \\
\vdots & & & & \vdots & \vdots & & \vdots & \vdots \\
0 & \cdots & & \cdots & 0 & 0 & \cdots & 0 & c_m
\end{array}$$

Dabei ist $(y_1, \ldots, y_n)$ eine Permutation der Komponenten des Vektors $(x_1, \ldots, x_n)$.

Die Zahl $r \leqslant \min\{m, n\}$ heißt der Rang *des linearen Gleichungssystems.*

Satz 6.6 (Fortsetzung)

- Das Gleichungssystem ist lösbar (konsistent) genau dann, wenn $c_{r+1} = \cdots = c_m = 0$ gilt; dann wählt man $y_{r+1}, \ldots, y_n$ als “freie Variablen” und bestimmt $y_1, \ldots, y_r$ aus den ersten r Gleichungen der reduzierten Stufenform.

- Das Gleichungssystem ist genau dann eindeutig lösbar, wenn $r = n$ und $c_{r+1} = \cdots = c_m = 0$ gilt.

Bemerkungen:

- Die entscheidende Zahl ist der Rang r des Gleichungssystems. Von vornherein kennt man nur die Abschätzungen $r \leqslant m$ und $r \leqslant n$.

- Wenn $r = m$ gilt ("voller Zeilenrang"), so ist das Gleichungssystem lösbar; denn es gibt keine Gleichungen (Zeilen) mit lauter Null-Koeffizienten in der reduzierten Stufenform.

- Wenn $r = n$ gilt ("voller Spaltenrang"), so existiert höchstens eine Lösung; denn alle Komponenten x_k sind durch die ersten n Gleichungen bereits eindeutig festgelegt. Die weiteren Gleichungen (für $m > n$) entscheiden dann darüber, ob dieser Vektor $\vec{x}$ eine Lösung ist oder nicht.

- Wenn $r < n$ gilt und das Gleichungssystem lösbar ist, gibt es unendlich viele Lösungen. Die Lösungsmenge enthält $n - r$ freie Parameter $y_{r+1}, \ldots, y_n$. Sie ist eine Gerade im Fall $n - r = 1$, Ebene im Fall $n - r = 2$, etc.

Der Spezialfall $m = n$ (Anzahl der Gleichungen ist gleich der Anzahl der Unbekannten) tritt besonders häufig auf.

Satz 6.7 (Alternativsatz)

Für ein lineares Gleichungssystem von n Gleichungen mit n Unbekannten sind die folgenden Aussagen äquivalent:

- *Das Gleichungssystem ist für beliebige rechte Seiten eindeutig lösbar (also universell eindeutig lösbar).*
- *Das zugehörige homogene System hat nur die triviale Lösung $\vec{x} = \vec{0}$.*
- *Das Gleichungssystem hat den Rang $r = n$.*

Beweis: durch Kombination der Aussagen in der vorherigen Bemerkung.

Ein konstruktiver Beweis zur Existenz der reduzierten Stufenform wird durch die Beschreibung des Gauß-Algorithmus angegeben:

Man führt (maximal) n Schritte zur Elimination nach folgenden Regeln durch (erklärt anhand der Kurzform mit Zeilen und Spalten):
Im k-ten Schritt $(1 \leqslant k \leqslant n)$

k_1: betrachte die k-te Spalte ab dem Diagonalelement $b_{k,k}$ nach unten (also Zeilen $k \leqslant i \leqslant m$). Stehen hier nur Nullen (incl. $b_{k,k}$), so können zwei Fälle auftreten:

- 1. Fall: es gibt eine weitere Spalte mit Index $k+1 \leqslant \ell \leqslant n$, die ein von Null verschiedenes Element in mindestens einer Zeile $k \leqslant i \leqslant m$ enthält. Dann tausche die beiden Spalten und nummeriere die Unbekannten um (E4).
- 2. Fall: alle weiteren Spalten mit Index $k+1 \leqslant \ell \leqslant n$ enthalten nur Nullen in den Zeilen $k \leqslant i \leqslant m$. Dann ist die reduzierte Stufenform erreicht und der Algorithmus beendet.

k_2 falls das (neue) Diagonalelement $b_{k,k}$ Null ist, dann tausche Zeile k mit einer Zeile i unterhalb (also $k+1 \leqslant i \leqslant m$), deren Element $b_{i,k}$ derselben Spalte ungleich Null ist (E1).

k_3: dividiere die k-te Zeile (incl. der k-ten Komponente der rechten Seite) durch das (neue) Diagonalelement $b_{k,k} \neq 0$ (E2). Dadurch entsteht das neue Diagonalelement $b_{k,k} = 1$.

k_4: subtrahiere das $b_{i,k}$-fache der Zeile k von Zeile i für $k + 1 \leqslant i \leqslant m$ (E3). Dadurch entstehen Nullen unterhalb des Diagonalelements $b_{k,k} = 1$.

Satz 6.8 (Komplexe Gleichungssysteme)

Alles gilt analog für komplexe Gleichungssysteme.

Kapitel 7 – Vektorräume

Kap. 7: Vektorräume

Als Verallgemeinerung des $\mathbb{R}^2$ und $\mathbb{R}^3$ betrachten wir allgemeine Vektorräume und Skalarprodukte.
Da die Vektoren in allgemeinen Vektorräumen nicht mehr nur Zahlentupel sein müssen, schreiben wir sie ohne den Vektorpfeil.

Definition 7.1 (Vektorraum)

Eine nichtleere Menge V, auf der eine Addition

$$v + w \in V \quad \textit{für alle} \quad v, w \in V$$

sowie eine Multiplikation mit Skalaren

$$\alpha v \in V \quad \textit{für alle} \quad v \in V,\ \alpha \in \mathbb{R} \qquad (\textit{bzw. } \alpha \in \mathbb{C}),$$

definiert ist, heißt REELLER *(bzw.* KOMPLEXER*)* VEKTORRAUM, *wenn die folgenden Gesetze erfüllt sind:*

Für alle $u, v, w \in V$ und $\alpha, \beta \in \mathbb{R}$ (bzw. $\mathbb{C}$) gelte

(a) $u + v = v + u$ (Kommutativgesetz)

(b) $(u + v) + w = u + (v + w)$ (Assoziativgesetz)

(c) Es gibt einen Nullvektor 0 mit der Eigenschaft $u + 0 = 0 + u = u$ für alle $u \in V$

(d) Zu jedem Vektor u gibt es einen Vektor $-u$ mit der Eigenschaft $u + (-u) = (-u) + u = 0$.

(e) $(\alpha + \beta)u = \alpha u + \beta u \quad \alpha(u + v) = \alpha u + \alpha v$ (Distributivgesetze)

(f) $(\alpha\beta)u = \alpha(\beta u)$ und $1u = u$

Definition 7.2 ($\mathbb{R}^n$ und $\mathbb{C}^n$)

Der n-dimensionale EUKLIDISCHE RAUM *ist die Menge aller n-Tupel reeller Zahlen*

$$\mathbb{R}^n = \mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R} = \{(x_1, \dots, x_n) \mid x_1, \dots, x_n \in \mathbb{R}\}.$$

Ein Element $\vec{x} = (x_1, \dots, x_n)^\top \in \mathbb{R}^n$ heißt ein VEKTOR, *und für $1 \leqslant k \leqslant n$ heißt x_k die k-te Koordinate oder* KOMPONENTE *von $\vec{x}$. Der Vektor $\vec{0} = (0, \dots, 0)$ heißt* NULLVEKTOR.

In einem kartesischen Koordinatensystem aus n paarweise senkrechten Koordinatenachsen definiert der Vektor $\vec{x} \in \mathbb{R}^n$ mit Anfangspunkt $(0, \dots, 0)$ (=Koordinatenursprung) und Endpunkt $P = (x_1, \dots, x_n)$ den Ortsvektor von P.

Der Raum $\mathbb{C}^n$ wird analog definiert.

Summe und Multiplikation mit Skalaren wird definiert durch

$$(x_1, \dots, x_n) + (y_1, \dots, y_n) = (x_1 + y_1, \dots, x_n + y_n)$$
$$\alpha(x_1, \dots, x_n) = (\alpha x_1, \dots, \alpha x_n)$$

Wichtigste Beispiele

$\mathbb{R}^n$ ist ein reeller und $\mathbb{C}^n$ ein komplexer Vektorraum.

Definition 7.3 (Unterraum)

Eine Teilmenge U eines Vektorraums V heißt UNTERRAUM, *wenn U mit den Verknüpfungen von V selbst ein Vektorraum ist.*
Dazu reicht es aus, dass U nichtleer ist und dass mit u und v sowie $\alpha, \beta \in \mathbb{R}$ (bzw. $\mathbb{C}$) auch

$$\alpha u + \beta v \in U.$$

Wichtige Beispiele sind $\{\vec{0}\}$, Geraden und Ebenen durch den Ursprung und V.

Definition 7.4 (Skalarprodukt)

Sei V ein Vektorraum. Ein reelles (komplexes) Skalarprodukt *auf V ist eine Abbildung*

$$< .,. >: V \times V \to \mathbb{R} \quad (\mathbb{C})$$

mit den Eigenschaften

(a) $< v, v > \geqslant 0$ *und* $< v, v >= 0 \iff v = 0$ — *Definitheit*

(b) $< \alpha v + \beta u, w >= \alpha < v, w > + \beta < u, w >$ — *Linearität im ersten Faktor*

(c) $< u, v >=< v, u > \quad (< u, v >= \overline{< v, u >})$ — *Symmetrie (Hermitizität)*

Ein Vektorraum mit Skalarprodukt heißt Skalarproduktraum.

Aus (b) und (c) folgt unmittelbar

$$< u, \alpha v + \beta w >= \overline{\alpha} < u, v > + \overline{\beta} < u, w >$$

Im reellen Fall tritt die Konjugation natürlich nicht auf.

Definition 7.5 (Skalarprodukt im $\mathbb{R}^n$ und $\mathbb{C}^n$)

Für Vektoren $\vec{x} = (x_1, \ldots, x_n)$ *und* $\vec{y} = (y_1, \ldots, y_n)$ *des* $\mathbb{R}^n$ *(*$\mathbb{C}^n$*) wird das* (STANDARD-)SKALARPRODUKT *auch mit dem Punkt geschrieben:*

$$\vec{x} \cdot \vec{y} :=< \vec{x}, \vec{y} >:= x_1 y_1 + \ldots + x_n y_n \quad (\vec{x} \cdot \vec{y} :=< \vec{x}, \vec{y} >:= x_1 \overline{y_1} + \ldots + x_n \overline{y_n})$$

Satz 7.6 (Cauchy-Schwarz Ungleichung)

Ist V *ein Skalarproduktraum, so gilt*

$$| < u, v > | \leqslant \sqrt{< u, u >} \sqrt{< v, v >}$$

Definition 7.7 (Norm)

Eine NORM $\|.\|$ *auf einem Vektorraum* V *ist eine Abbildung von* V *in die reellen Zahlen mit den Eigenschaften*

(a) $\|v\| \geqslant 0$ *und* $\|v\| = 0 \Longleftrightarrow v = 0$ *Definitheit*

(b) $\|\alpha v\| = |\alpha| \, \|v\|$ *Homogenität*

(c) $\|u + v\| \leqslant \|u\| + \|v\|$ *Dreiecksungleichung*

Satz 7.8

Ist $< .,. >$ *ein Skalarprodukt auf einem Vektorraum* V*, so ist durch* $\|v\| = \sqrt{< v, v >}$ *eine Norm definiert.*

Definition 7.9 (Winkel)

Der WINKEL $\alpha = \angle(\vec{x}, \vec{y}) \in [0, \pi]$ *zwischen Vektoren* $\vec{x}, \vec{y} \in \mathbb{R}^n \backslash \{\vec{0}\}$ *wird definiert durch die Festlegung*

$$\cos\alpha = \frac{\vec{x} \cdot \vec{y}}{|\vec{x}|\ |\vec{y}|}.$$

Ist $\alpha = \pi/2$*, so heißen* $\vec{x}$ *und* $\vec{y}$ *orthogonal, und wir schreiben* $\vec{x} \perp \vec{y}$*.*

Beachte: Nach der Cauchy-Schwarz Ungleichung gilt

$$-1 \leqslant \frac{\vec{x} \cdot \vec{y}}{|\vec{x}|\ |\vec{y}|} \leqslant 1,$$

also ist der Winkel $\alpha \in [0, \pi]$ eindeutig bestimmt.

Bemerkung: Ob im folgenden die Skalare $\alpha, \beta, \ldots \in \mathbb{R}$ oder $\alpha, \beta, \ldots \in \mathbb{C}$ gewählt werden, hängt davon ab, ob wir einen reellen oder einen komplexen Vektorraum betrachten.

Definition 7.10 (Linearkombination, Aufspann)

Sei V ein reeller oder komplexer Vektorraum.

a) *Ein Vektor $x = \alpha_1 v_1 + \cdots + \alpha_k v_k \in V$ mit Skalaren $\alpha_1, \ldots, \alpha_k$ (aus $\mathbb{R}$ bzw. $\mathbb{C}$) heißt eine* Linearkombination *der Vektoren $v_1, \ldots, v_k$.*

b) *Die Menge aller Linearkombinationen*

$$\mathit{Span}\,(v_1, \ldots, v_k) := \left\{ x = \sum_{j=1}^{k} \alpha_j v_j \mid \alpha_1, \ldots, \alpha_k \ \ \textit{Skalare} \right\}$$

heißt der Spann *oder* Aufspann *der Vektoren $v_1, \ldots, v_k$ (oder der von $v_1, \ldots, v_k$ aufgespannte Unter-Vektorraum von V).*

Definition 7.11 (Basis)

Die Vektoren $v_1, \ldots, v_n$ *heißen* BASIS *von* V, *wenn* *jeder* *Vektor* $x \in V$ *eine* *eindeutige* *Darstellung*

$$x = \alpha_1 v_1 + \cdots + \alpha_n v_n$$

besitzt. Dann heißt

$$\vec{a} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \in \mathbb{R}^n \quad (\mathbb{C}^n)$$

der KOORDINATENVEKTOR *von* x *zur Basis* $v_1, \ldots, v_n$.

Wichtiges Beispiel: Basen des $\mathbb{R}^n$ bzw. $\mathbb{C}^n$

Satz 7.12 (Basen im $\mathbb{R}^n$ und $\mathbb{C}^n$)

Die Vektoren $\vec{v}_1, \ldots, \vec{v}_n \in \mathbb{R}^n$ (bzw. $\mathbb{C}^n$) bilden genau dann eine Basis von $\mathbb{R}^n$ (bzw. $\mathbb{C}^n$), wenn für jeden Vektor $\vec{x}$ das Gleichungssystem

$$\lambda_1 \vec{v}_1 + \cdots + \lambda_n \vec{v}_n = \vec{x}$$

eine eindeutige Lösung hat.

In diesem Fall ist der Koordinatenvektor von $\vec{x} \in \mathbb{R}^n$ (bzw. $\mathbb{C}^n$) gegeben durch $\vec{a} = (\lambda_1, \ldots, \lambda_n)^\top$.

- Wir sehen also, dass jede Basis von $\mathbb{R}^n$ (bzw. $\mathbb{C}^n$) aus genau n Vektoren besteht.

Definition 7.13 (Dimension)

Hat der Vektorraum V eine Basis aus n Vektoren $v_1, \dots, v_n \in V$, so besteht jede andere Basis von V ebenfalls aus n Vektoren.

Die Zahl n heißt dann die DIMENSION *von V, geschrieben $\dim V = n$.*

- Besitzt der Vektorraum V keine endliche Basis, so heißt V UNENDLICH-DIMENSIONAL, geschrieben $\dim V = \infty$.

Lineares homogenes Gleichungssystem:

$$\sum_{k=1}^{n} a_{ik} x_k = 0, \quad (i = 1, \ldots, m) \qquad (A)$$

Nach Satz 6.3 ist die Lösungsmenge von A ein Vektorraum.

Definition 7.14 (Kern)

Die Lösungsmenge eines homogenen linearen Gleichungssystems A wird mit $\textsc{Kern}$, *geschrieben Kern A, bezeichnet.*

Satz 7.15 (Dimensionsformel)

Für ein lineares homogenes Gleichungssystem mit m Gleichungen und n Variablen gelte Rang $A = r$. Dann hat Kern(A) die Dimension $n - r$, d.h.

$$\mathit{dim}(\mathit{Kern}\,A) = n - \mathit{Rang}\,A \qquad \textit{(Dimensionsformel)}$$

Beweis: Wir verwenden die spezielle reduzierte Stufenform des homogenen LGS

$$\begin{array}{cccccccc|cl}
y_1 & y_2 & y_3 & \cdots & y_r & y_{r+1} & \cdots & y_n & \text{r.S.} & \\ \hline
1 & 0 & \cdots & \cdots & 0 & b_{1,r+1} & \cdots & b_{1,n} & 0 & \\
0 & 1 & 0 & \cdots & 0 & b_{2,r+1} & \cdots & b_{2,n} & 0 & \\
\vdots & \ddots & \ddots & \ddots & \vdots & \vdots & & \vdots & \vdots & \\
0 & \cdots & 0 & 1 & 0 & b_{r-1,r+1} & \cdots & b_{r-1,n} & 0 & \\
0 & \cdots & 0 & 0 & 1 & b_{r,r+1} & \cdots & b_{r,n} & 0 & \\
0 & \cdots & & \cdots & 0 & 0 & \cdots & 0 & 0 & \\
\vdots & & & & \vdots & \vdots & & \vdots & \vdots & \left.\vphantom{\begin{array}{c}0\\0\\0\end{array}}\right\} m-r \text{ Gl.} \\
0 & \cdots & & \cdots & 0 & 0 & \cdots & 0 & 0 &
\end{array}$$

Dabei ist $(y_1, \dots, y_n)$ eine Umnummerierung der Komponenten des Vektors $(x_1, \dots, x_n)$, je nach durchgeführten Spaltenvertauschungen.

Falls $r = n$ ist, so ist die Lösungsmenge der Nullraum Kern $A = \{\vec{0}\}$.

Falls $r < n$ ist, so erhält man eine Basis von Kern A durch Einsetzen von Nullen und jeweils einer 1 für die freien Variablen $y_{r+1}, \dots, y_n$:

$$\vec{w}_{r+1} = \begin{pmatrix} -b_{1,r+1} \\ \vdots \\ -b_{r,r+1} \\ 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \vec{w}_{r+2} = \begin{pmatrix} -b_{1,r+2} \\ \vdots \\ -b_{r,r+2} \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \ldots, \quad \vec{w}_n = \begin{pmatrix} -b_{1,n} \\ \vdots \\ -b_{r,n} \\ 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}.$$

Die Basis der Vektoren $\vec{v}_j \in \mathbb{R}^n$ von Kern A ergibt sich dann durch Umnummerierung (Zeilentausch) der Komponenten der Vektoren $\vec{w}_j$, $1 \leqslant j \leqslant n - r$.

Definition 7.16 (lineare Abhängigkeit)

Die Elemente $v_1, \dots, v_k \in V$ *eines Vektorraums* V *heißen* LINEAR ABHÄNGIG, *wenn es eine Linearkombination*

$$\sum_{j=1}^{k} \alpha_j v_j = \vec{0}$$

mit $|\alpha_1| + \dots + |\alpha_k| \neq 0$ *gibt (das ist eine Linearkombination, in der nicht alle Skalare Null sind).*

Andernfalls heißen die Elemente $v_1, \dots, v_k \in V$ LINEAR UNABHÄNGIG.

Die folgenden allgemeinen Aussagen helfen beim Umgang mit Vektorräumen und Unter-Vektorräumen.

Satz 7.17 (Existenz von Basen)

- *Jeder Vektorraum besitzt eine Basis.*
- *Jede Familie* $v_1, \ldots, v_k \in V$ *von linear unabhängigen Vektoren lässt sich zu einer Basis von* V *ergänzen.*

Beispiel 7.18 (Polynome)

Der Vektorraum $\mathcal{P}_n$ der Polynome vom Grad $\leqslant n$ mit komplexen Koeffizienten ist ein komplexer Vektorraum der Dimension $n+1$.

a) Die Basis der Monome ist

$$m_0(x) = 1, \quad m_1(x) = x, \quad \dots, \quad m_n(x) = x^n,$$

und der Koordinatenvektor von $p(x) = \sum_{k=0}^n a_k x^k$ ist $(a_0, \dots, a_n)^\top \in \mathbb{C}^{n+1}$.

b) Es seien $z_0, \dots, z_n \in \mathbb{C}$ paarweise verschieden. Eine andere Basis von $\mathcal{P}_n$ ist die Basis der Lagrange-Grundpolynome

$$L_j(x) = \prod_{\substack{k=0 \\ k \neq j}}^n \frac{x - z_k}{z_j - z_k}.$$

Beachte: L_j ist das eindeutige Polynom in $\mathcal{P}_n$ mit den Nullstellen z_k, $k \neq j$, und dem Wert $L_j(z_j) = 1$. Deshalb gilt für jedes $p \in \mathcal{P}_n$

$$p(x) = \sum_{k=0}^n p(z_k) L_k(x).$$

Der Koordinatenvektor von p zur Basis $L_0, \dots, L_n$ ist der Vektor $(p(z_0), \dots, p(z_n))^\top \in \mathbb{C}^{n+1}$.

Satz 7.19 (Der Vektorraum der reellen Funktionen)

Der Vektorraum $V = \{f \mid f : \mathbb{R} \to \mathbb{R}$ *ist eine Funktion*$\}$ *ist unendlich-dimensional.*

In Vektorräumen mit Skalarprodukt sind Basen mit einer Orthogonalitätsrelation besonders interessant:

Definition 7.20 (Orthonormalsysteme und -basen)

Die Vektoren v_1 *bis* v_n *bilden ein* $\text{Orthonormalsystem (ONS)}$, *wenn gilt*

- *Für* $i \neq j$ *ist* $< v_i, v_j >= 0$
- *Stets ist* $\|v_i\|^2 =< v_i, v_i >= 1$.

Mit dem Kroneckersymbol $\delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$ *schreibt sich das kurz als*

$$< v_i, v_j >= \delta_{ij}.$$

Eine $\text{Orthonormalbasis (ONB)}$ *ist eine Basis, die ein ONS ist.*

Die Vektoren eines ONS stehen also paarweise senkrecht aufeinander und haben die Länge 1.

Aus linear unabhängigen Vektoren $u_1, \ldots, u_k \in V$ erhält man mit dem folgenden Algorithmus eine Orthonormalbasis

$$w_1, \ldots, w_k \quad \text{von} \quad \text{Span}(u_1, \ldots, u_k)$$

7.21 (Gram-Schmidt Orthogonalisierung)

1. *Setze* $v_1 = u_1$ *(Initialisierung)*
2. *Für* $\ell = 2, 3, \ldots, k$ *setze* *(Iteration)*

$$v_\ell = u_\ell - \frac{< u_\ell, v_1 >}{< v_1, v_1 >}\, v_1 - \cdots - \frac{< u_\ell, v_{\ell-1} >}{< v_{\ell-1}, v_{\ell-1} >}\, v_{\ell-1}$$

3. *Für* $\ell = 1, 2, \ldots, k$ *setze* *(Normierung)*

$$w_\ell = \frac{1}{\|v_\ell\|}\, v_\ell$$

Der Hauptvorteil einer ONB liegt darin, dass man Koeffizienten in Linearkombinationen berechnen kann, ohne ein Gleichungssystem zu lösen.

Satz 7.22 (Besonderheiten bei ONS und ONB)

- *Die Elemente eines ONS sind stets linear unabhängig.*
- *Sei* $v_1, \ldots, v_k$ *eine ONB,* $w = \alpha_1 v_1 + \cdots + \alpha_k v_k$.
  Dann ist $\alpha_j =< w, v_j >$.
- *Ist* $v_1, \ldots, v_k$ *eine ONB und* $w = \alpha_1 v_1 + \cdots + \alpha_k v_k$.
  Dann ist $\|w\|^2 = \sum_{j=1}^{k} |\alpha_j|^2$.

Definition 7.23 (Lineare Abbildung)

Sind V und W Vektorräume, so heißt eine Abbildung $L : V \to W$ LINEAR, *wenn für Vektoren $v_1, v_2 \in V$ und $\alpha, \beta \in \mathbb{R}$ (bzw. $\mathbb{C}$) stets gilt:*

$$L(\alpha v_1 + \beta v_2) = \alpha L(v_1) + \beta L(v_2).$$

Die Menge der $v \in V$ mit $L(v) = 0$ heißt KERN *von L, geschrieben Kern L.*

Lineare Abbildungen haben die folgenden Eigenschaften:

Satz 7.24 (Eigenschaften linearer Abbildungen)

Es gilt:

a) *Ist $M : W \to U$ eine weitere lineare Abbildung, so ist die zusammengesetzte Abbildung $M \circ L : V \to U$ wieder linear.*
b) *Ist $L : V \to W$ bijektiv, dann ist $L^{-1} : W \to V$ auch linear.*
c) *Kern L ist ein Unterraum von V, das Bild von L ist ein Unterraum von W.*

- Die Abbildung

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} a_{11}x_1 + a_{22}x_2 + \cdots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n \end{pmatrix}$$

  ist linear.

  Bemerkung: Für $m = n = 1$ sind diese Abbildungen Geraden durch den Ursprung.

- Sei V ein Vektorraum mit Skalarprodukt $\langle \cdot, \cdot \rangle$ und $w \in V$ beliebig. Dann ist die Abbildung

$$v \mapsto \langle v, w \rangle$$

  laut Definition 7.4 linear.

Kapitel 8 – Matrizen

Kap. 8: Matrizen

Die Kurzschreibweise von LGS führt zum Begriff der Matrix.

Definition 8.1 (Matrix)

Eine (m,n)-Matrix *ist ein* mn*-Tupel von Zahlen, die zu einem rechteckigen Schema aus* m *Zeilen und* n *Spalten angeordnet sind:*

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix} = (a_{i,k})_{\substack{i=1,\dots,m \\ k=1,\dots,n}}$$

Die $a_{i,k}$ *heißen die* Koeffizienten *(oder* Einträge*) der Matrix* A.

$$\mathbf{0} := \begin{pmatrix} 0 & 0 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 0 \end{pmatrix}_{m\times n} \quad \text{Nullmatrix} \qquad E_n := \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix}_{n\times n} \quad \text{Einheitsmatrix}$$

Bemerkung 8.2

Spaltenvektoren können als Matrizen mit einer Spalte, Zeilenvektoren als Matrizen mit einer Zeile betrachtet werden.

Satz 8.3 (Matrizen als Vektorraum)

Die Menge aller (m, n)-Matrizen mit reellen (bzw. komplexen) Koeffizienten wird mit Mat(m, n) oder $\mathbb{R}^{m\times n}$ (bzw. $\mathbb{C}^{m\times n}$) bezeichnet. Diese Menge ist ein Vektorraum mit den Operationen

$$(a_{i,k})_{m\times n} + (b_{i,k})_{m\times n} = (a_{i,k} + b_{i,k})_{m\times n}, \qquad \alpha(a_{i,k})_{m\times n} = (\alpha a_{i,k})_{m\times n}$$

- Man kann nur Matrizen gleicher Dimensionen addieren.
- Die Nullmatrix $(\mathbf{0})_{m\times n}$ ist das neutrale Element der Addition.
- Für Matrizen $A, B \in$ Mat(m, n) bedeutet $A = B$, dass $a_{i,k} = b_{i,k}$ für alle $1 \leqslant i \leqslant m$, $1 \leqslant k \leqslant n$ gilt.

Wir schreiben eine Matrix $A \in \mathrm{Mat}(m,n)$ häufig mit Hilfe ihrer Spaltenvektoren

$$A = (\vec{a}_1, \dots, \vec{a}_n), \qquad \vec{a}_k = \begin{pmatrix} a_{1,k} \\ \vdots \\ a_{m,k} \end{pmatrix}.$$

Satz 8.4 (Produkte von Matrizen und Vektoren)

- *Die Matrix $A \in \mathrm{Mat}(m,n)$ habe die Spalten $\vec{a}_1$ bis $\vec{a}_n$. Das Produkt von A mit dem Vektor $\vec{x} = (x_1, \dots, x_n)^\top$ ist die Linearkombination der $\vec{a}_i$ mit den x_i als Koeffizienten, also*

$$A\vec{x} = x_1\vec{a}_1 + \cdots + x_n\vec{a}_n.$$

- *Hat die Matrix $B \in \mathrm{Mat}(n,p)$ die Spalten $\vec{b}_1$ bis $\vec{b}_p$, so hat das Matrizenprodukt AB die Spalten $A\vec{b}_1$ bis $A\vec{b}_p$.*

Die einzelnen Einträge werden wie folgt berechnet:

Definition 8.5 (Matrizenprodukt)

Es seien $m, n, p \in \mathbb{N}$ *sowie* $A = (a_{i,k})_{m \times n} \in \mathit{Mat}(m,n)$ *und* $B = (b_{k,j})_{n \times p} \in \mathit{Mat}(n,p)$.

Dann ist das Matrixprodukt $C := AB \in \mathit{Mat}(m,p)$ *mit* $C = (c_{i,j})_{m \times p}$ *gegeben durch*

$$c_{i,j} = \sum_{k=1}^{n} a_{i,k} b_{k,j} \qquad \text{„Zeile } i \text{ mal Spalte } j\text{"}$$

$$B \rightarrow \begin{pmatrix} * & b_{1,j} & * \\ * & b_{2,j} & * \\ * & b_{3,j} & * \\ \vdots & \vdots & \vdots \end{pmatrix}$$

$$\underset{A}{\begin{pmatrix} * & * & * & \cdots \\ a_{i,1} & a_{i,2} & a_{i,3} & \cdots \\ * & * & * & \cdots \end{pmatrix}} \quad \underset{C = AB}{\begin{pmatrix} & & \\ & c_{i,j} & \\ & & \end{pmatrix}}$$

Bemerkung: Die Merkregel "Zeile mal Spalte" ergibt:

- Die i-te Zeile des Produkts AB hängt nur von der i-ten Zeile von A ab. Insbesondere gilt:
- Ist in $(0, \ldots, 0, 1, 0, \ldots, 0)$ die 1 an der i-ten Stelle, so ergibt $(0, \ldots, 0, 1, 0, \ldots, 0)B$ die i-te Zeile von B
- $A\vec{e}_j$ ergibt die j-te Spalte von A.
- Mit den obigen Beobachtungen ergibt sich: Die Einheitsmatrizen sind die neutralen Elemente des Matrixprodukts, d.h. für $A \in \mathrm{Mat}(m, n)$ gilt

$$E_m A = A E_n = A.$$

Satz 8.6 (Rechenregeln)

Gegeben seien Matrizen A, B, C so, dass die entsprechenden Matrizenprodukte definiert sind.

a) *Es gelten das Assoziativ- und das Distributiv-Gesetz:*

$$A(BC) = (AB)C, \quad (A+B)C = AC + BC, \quad A(B+C) = AB + AC.$$

b) *Das Matrixprodukt ist nicht kommutativ.*

Bemerkung: zu (b):

- Wenn $A \in \mathrm{Mat}(m,n)$ und $B \in \mathrm{Mat}(n,p)$, so ist AB definiert, jedoch ist BA nur im Fall $p = m$ definiert.
- Selbst wenn beide Produkte AB und BA definiert sind, so sind die Matrizen oft verschieden.

- Das Produkt mit der Nullmatrix ergibt die Nullmatrix. Aber: Aus $AB = (\mathbf{0})_{m\times p}$ folgt nicht, dass A oder B eine Nullmatrix ist.

- Ebenso gilt im allgemeinen nicht die Kürzungsregel: aus $AB = AC$ folgt i.a. nicht die Gleichheit von B und C. Hierzu muss A eine Zusatzbedingung erfüllen (siehe invertierbare Matrizen, reguläre Matrizen).

- Wenn man das Produkt eines Vektors mit einer Zahl als Matrizenprodukt auffassen will, muss die Zahl rechts vom Vektor stehen.

Satz 8.7 (Matrix-Form des linearen Gleichungssystems)

Für $A \in$ Mat(m, n), einen Spaltenvektor $\vec{b} \in \mathbb{R}^m$ sowie einen Spaltenvektor $\vec{x} \in \mathbb{R}^n$ lässt sich das inhomogene Gleichungssystem (*) in 6.2 schreiben als

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

A heißt die KOEFFIZIENTENMATRIX *des linearen Gleichungssystems und $\vec{b}$ die* RECHTE SEITE.

- Das lineare Gleichungssystem $A\vec{x} = \vec{b}$ ist genau dann lösbar, wenn die rechte Seite $\vec{b}$ eine Linearkombination der Spaltenvektoren von A ist.
- Die elementaren Umformungen (E1)–(E4) lassen sich auf die Zeilen und Spalten der Matrix A anwenden, um eine REDUZIERTE STUFENFORM von A zu erhalten.
- Der RANG der Matrix A ist dasselbe wie der Rang r der reduzierten Stufenform von A.

- Die Matrix $\tilde{A} = (A, \vec{b})_{m \times (n+1)}$ in der Kurzform des Gleichungssystems heißt die ERWEITERTE KOEFFIZIENTENMATRIX. Wir haben in Satz 6.6 festgestellt, dass das inhomogene lineare Gleichungssystem genau dann lösbar ist, wenn
$$\operatorname{Rang} A = \operatorname{Rang} \tilde{A}.$$

Weitere Aussagen in Satz 6.6 erhalten nun die folgende Form:

- $\operatorname{Rang} A = m \implies$ das Gleichungssystem $A\vec{x} = \vec{b}$ ist lösbar
- $\operatorname{Rang} A = n \implies$ das Gleichungssystem $A\vec{x} = \vec{b}$ hat höchstens eine Lösung
- $\operatorname{Rang} A = m = n \implies$ das Gleichungssystem $A\vec{x} = \vec{b}$ ist für jede rechte Seite $\vec{b}$ eindeutig lösbar (universell eindeutig lösbar)

Definition 8.8 (Kern und Bildraum)

Die Lösungsmenge des homogenen LGS $A\vec{x} = \vec{0}$ zur Matrix $A \in Mat(m,n)$ ist ein Vektorraum. Dieser Vektorraum ist ein Untervektorraum von $\mathbb{R}^n$ bzw. $\mathbb{C}^n$ und wird wie in Definition 7.14 mit KERN *bezeichnet.*

$$\mathit{Kern}\, A := \{\vec{x} \in \mathbb{R}^n \mid A\vec{x} = \vec{0}\}$$

Für eine Matrix $A = (\vec{a}_1, \ldots, \vec{a}_n) \in Mat(m,n)$ mit Spaltenvektoren $\vec{a}_k$ heißt der Unter-Vektorraum

$$\mathit{Bild}\, A = \mathit{Span}(\vec{a}_1, \ldots, \vec{a}_n)$$

der BILDRAUM *von A. Es gilt $\mathit{Rang}\, A = \dim(\mathit{Bild}\, A)$, also mit der Dimensionsformel*

$$\dim(\mathit{Kern}\, A) + \dim(\mathit{Bild}\, A) = n. \qquad (*)$$

Bemerkung: Hieraus folgt, dass sich der Rang einer Matrix auf zwei verschiedene Weisen darstellen lässt:

- $r = \operatorname{Rang} A$ ist die maximale Anzahl von linear unabhängigen Zeilen von A; das bedeutet $\dim(\operatorname{Kern} A) = n - r$.
- $r = \operatorname{Rang} A$ ist die maximale Anzahl von linear unabhängigen Spalten von A; das bedeutet $\dim(\operatorname{Bild} A) = r$.

Weiter gilt: es gibt eine Basis $\vec{v_1}$ bis $\vec{v}_n$ des $\mathbb{R}^n$ mit:

- $\vec{v}_1$ bis $\vec{v}_{n-r}$ ist eine Basis des Kerns von A
- $A\vec{v}_{n-r+1}$ bis $A\vec{v}_n$ ist eine Basis des Bildes von A

Eine weitere Operation für Matrizen:

Definition 8.9 (Transponierte)

Die zu

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \cdots & a_{1,n} \\ a_{2,1} & a_{2,2} & \cdots & a_{2,n} \\ \vdots & \vdots & & \vdots \\ a_{m,1} & a_{m,2} & \cdots & a_{m,n} \end{pmatrix} \in \mathit{Mat}(m,n)$$

TRANSPONIERTE MATRIX *ist*

$$A^\top = \begin{pmatrix} a_{1,1} & a_{2,1} & \cdots & a_{m,1} \\ a_{1,2} & a_{2,2} & \cdots & a_{m,2} \\ \vdots & \vdots & & \vdots \\ a_{1,n} & a_{2,n} & \cdots & a_{m,n} \end{pmatrix} \in \mathit{Mat}(n,m).$$

Die Adjungierte A^* einer komplexen Matrix ist definiert durch

$$A^* = \overline{A^\top} = \overline{A}^\top$$

Satz 8.10 (Transponierte, Adjungierte und Skalarprodukt)

Sei $A \in$ Mat(m, n). Dann gibt es genau eine Matrix $B \in$ Mat(n, m), so dass für alle $\vec{x} \in \mathbb{R}^n$ und $\vec{y} \in \mathbb{R}^m$ gilt

$$\langle A\vec{x}, \vec{y}\rangle = \langle \vec{x}, B\vec{y}\rangle.$$

Es ist $B = A^\top$.
Eine entsprechende Aussage gilt für komplexe Vektorräume und $B = A^*$.

Satz 8.11 (Rechenregeln)

- *Addition: Für $A, B \in \mathit{Mat}(m, n)$ und Skalare α gilt*

$$(A+B)^\top = A^\top + B^\top, \qquad (\alpha A)^\top = \alpha A^\top.$$

$$(A+B)^* = A^* + B^*, \qquad (\alpha A)^* = \overline{\alpha} A^*.$$

- *Matrixprodukt: Für $A \in \mathit{Mat}(m, n)$ und $B \in \mathit{Mat}(n, p)$ gilt*

$$(AB)^\top = B^\top A^\top \qquad \text{und} \qquad (AB)^* = B^* A^*$$

- *Für jede Matrix A gilt $(A^\top)^\top = A$ und $(A^*)^* = A$.*
- *Es gilt Rang A = Rang $A^\top$ und Rang A = Rang A^*.*
- *Es ist $\langle \vec{x}, \vec{y} \rangle = \vec{y}^\top \vec{x}$ bzw. $\langle \vec{x}, \vec{y} \rangle = \vec{y}^* \vec{x}$ (Skalarprodukt als Matrizenprodukt).*

Satz 8.12 (Rang von Produkten)

Für alle $A \in \mathit{Mat}(m, n)$ *und* $B \in \mathit{Mat}(n, p)$ *gilt*

$$\mathit{Rang}(AB) \leqslant \min\{\mathit{Rang}\,A, \mathit{Rang}\,B\}.$$

Wir betrachten ab jetzt **quadratische** Matrizen $A \in \text{Mat}(n, n)$ (Zeilenzahl = Spaltenzahl)

Definition 8.13 (Inverse)

Gegeben sei $A \in \text{Mat}(n, n)$. Falls es eine Matrix $X \in \text{Mat}(n, n)$ mit

$$AX = XA = E_n$$

gibt, so heißt A INVERTIERBAR *(oder* REGULÄR*) und $A^{-1} := X$ heißt die* INVERSE *von A.*

Satz 8.14

a) *Die Matrix $A \in \text{Mat}(n, n)$ besitzt genau dann eine Inverse A^{-1}, wenn $\text{Rang}\, A = n$ gilt.*

b) *Falls A invertierbar ist, so ist die Inverse A^{-1} eindeutig bestimmt.*

Bemerkung; Es gibt viele Matrizen $A \in \text{Mat}(n, n)$, die keine Inverse besitzen (nämlich alle Matrizen vom Rang $r < n$). Wir werden später sehen, dass die Bedingung $\det A \neq 0$ notwendig und hinreichend für die Invertierbarkeit von A ist.

Beispiel

Bestimme die Inverse (falls möglich) zu $A = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 1 & 2 \\ 1 & 1 & 1 \end{pmatrix}$

Satz 8.15

- *Die Umformungen (E1) - (E3) des Gauß-Algorithmus lassen sich durch Multiplikation der (erweiterten) Matrix von links mit invertierbaren Matrizen darstellen.*
- *Eine wichtige Konsequenz aus Satz 8.12 ist die folgende:*
  Ist $A \in \mathit{Mat}(n,n)$ *regulär, so gilt*

$$\mathit{Rang}(AB) = \mathit{Rang}\,B \qquad \textit{für alle} \quad B \in \mathit{Mat}(n,p),$$

$$\mathit{Rang}(CA) = \mathit{Rang}\,C \qquad \textit{für alle} \quad C \in \mathit{Mat}(p,n).$$

Satz 8.16 (Lösbarkeit von LGS)

a) *Falls $A \in$ Mat(n,n) regulär ist, so ist das lineare Gleichungssystem $A\vec{x} = \vec{b}$ universell eindeutig lösbar und die Lösung lautet $\vec{x} = A^{-1}\vec{b}$.*

b) *Falls $A \in$ Mat(n,n) regulär ist, so gilt für Matrizen $B, C \in$ Mat(n,p) die Kürzungsregel*

$$AB = AC \quad \Longleftrightarrow \quad B = C.$$

c) *Falls $A, B \in$ Mat(n,n) beide regulär sind, so ist auch $C = AB$ regulär und es gilt*

$$(AB)^{-1} = B^{-1}A^{-1}.$$

d) *Für reguläre Matrizen $A \in$ Mat(n,n) gilt $(A^\top)^{-1} = (A^{-1})^\top$.*

- Die Kürzungsregel besagt: Multiplikation beider Seiten eines linearen Gleichungssystems mit einer regulären Matrix ist eine Äquivalenzumformung:

$$A\vec{x} = \vec{b} \iff MA\vec{x} = M\vec{b}, \qquad \text{falls } M \in \mathrm{Mat}(m,m) \text{ regulär}$$

Satz 8.17 (Gruppe der regulären Matrizen)

Die regulären (n, n)-Matrizen bilden mit der Multiplikation die Gruppe *Gl*(n), *das heißt:*

(G1) *das Produkt $C = AB$ zweier regulärer (n, n)-Matrizen ist eine reguläre (n, n)-Matrix,*

(G2) *es gilt das Assoziativgesetz der Matrix-Multiplikation,*

(G3) *die Einheitsmatrix E_n ist das eindeutige "neutrale" Element der Multiplikation, also $AE_n = E_nA = A$ für jede reguläre (n, n)-Matrix,*

(G4) *zu jeder regulären (n, n)-Matrix A gibt es genau eine reguläre (n, n)-Matrix X mit $AX = XA = E_n$, nämlich $X = A^{-1}$.*

- **Beachte:** Das Kommutativgesetz der Multiplikation gilt in Mat(n, n) **nicht**.
- $E_n^{-1} = E_n$, denn $E_nE_n = E_n$

Definition 8.18 (Determinante)

Die DETERMINANTE *ist eine Funktion, die n Vektoren (bzw. einer $n \times n$-Matrix) eine reelle Zahl zuordnet. Sie lässt sich durch folgende Bedingungen charakterisieren:*

- $\det : \underbrace{\mathbb{R}^n \times \cdots \times \mathbb{R}^n}_{n \text{ Faktoren}} \to \mathbb{R}$ *ist linear in jedem Faktor, d.h.*

$$\det(\vec{v}_1, \ldots, (\alpha\vec{u} + \beta\vec{w}), \ldots, \vec{v}_n) = \alpha \det(\vec{v}_1, \ldots, \vec{u}, \ldots, \vec{v}_n) + \beta \det(\vec{v}_1, \ldots, \vec{w}, \ldots, \vec{v}_n)$$

- *Die Determinante ist alternierend, d.h.*

$$\det(\vec{v}_1, \ldots, \vec{v}_i, \ldots, \vec{v}_j, \ldots, \vec{v}_n) = -\det(\vec{v}_1, \ldots, \vec{v}_j, \ldots, \vec{v}_i, \ldots, \vec{v}_n)$$

- *Die Determinante ist normiert:* $\det(\vec{e}_1, \ldots, \vec{e}_n) = 1.$

Diese Bedingungen legen die Determinante eindeutig fest (ohne Beweis).

Folgerungen:

- $\det(\vec{v}_1, \ldots, \vec{u}, \ldots, \vec{u}, \ldots, \vec{v}_n) = 0$
- $\det(\vec{v}_1, \ldots, \vec{v}_i + \alpha\vec{v}_j, \ldots, \vec{v}_j, \ldots, \vec{v}_n) = \det(\vec{v}_1, \ldots, \vec{v}_i, \ldots, \vec{v}_j, \ldots, \vec{v}_n).$

Alternative Schreibweise: $|A|$ statt $\det A$

Satz 8.19 (Berechnung von Determinanten)

- *Für eine* $(1,1)$*-Matrix* $A = (a)$ *ist* $\det A = a$.
- *Für eine* $(2,2)$*-Matrix* $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ *ist* $\det A = ad - bc$.
- *Für eine* (n,n)*-Matrix* $A = (a_{i,k})_{n\times n}$ *wird* $\det A$ *berechnet durch die Rekursion ("Entwicklung nach der ersten Spalte")*

$$\det A = \sum_{k=1}^{n} (-1)^{k+1} a_{k,1} \det(A_{k,1}),$$

wobei die $(n-1,n-1)$*-Matrix* $A_{k,m}$ *aus* A *durch Streichen der* k*-ten Zeile und der ersten Spalte hervorgeht.*

Bemerkung: Die Rekursion kann nach jeder Zeile und jeder Spalte gebildet werden und führt zum gleichen Wert:

$$\det A = \sum_{k=1}^{n} (-1)^{j+k} a_{j,k} \det(A_{j,k}), \qquad \text{Entwicklung nach der } j\text{-ten Zeile}$$

$$\det A = \sum_{i=1}^{n} (-1)^{i+\ell} a_{i,\ell} \det(A_{i,\ell}), \qquad \text{Entwicklung nach der } \ell\text{-ten Spalte}$$

Man beachte das Schachbrettmuster des Vorzeichenfaktors $(-1)^{i+j}$

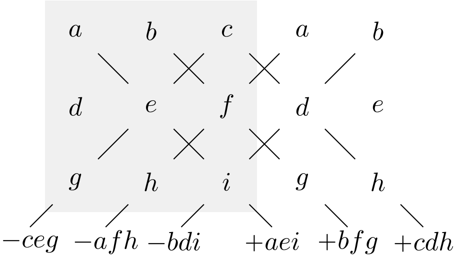
$$\begin{pmatrix} + & - & + & - & \cdots \\ - & + & - & + & \cdots \\ + & - & + & - & \cdots \\ \vdots & \vdots & \vdots & \vdots & \end{pmatrix}$$

Beispiele:

- Für $(3,3)$-Matrizen ergibt sich die Determinante

$$\det \begin{pmatrix} a & b & c \\ d & e & f \\ g & h & i \end{pmatrix} = a(ei - fh) - b(di - fg) + c(dh - eg)$$

auch mit der Merkregel von Sarrus



Achtung! Die Sarrus-Regel kann nur bei 3×3-Matrizen angewendet werden!

- Für $A = \begin{pmatrix} 1 & -1 & 2 & 3 \\ 5 & 2 & 0 & 0 \\ 3 & 1 & 0 & 4 \\ 6 & 7 & -2 & 6 \end{pmatrix}$ $\longleftarrow$ Entw. n. d. 2. Zeile

  ist $\det A = -5 \det A_{2,1} + 2 \det A_{2,2} = -5 \cdot 30 + 2 \cdot 2 = -146.$

Die Berechnung von $\det A$ erfolgt bei größeren Matrizen mit dem Gauß-Algorithmus:

8.20 (Determinantenberechnung mit dem Gauß-Algorithmus)

Gegeben sei $A \in$ Mat(n, n).

a) *Vertauschen zweier Zeilen oder Spalten von A (Umformungen E1 und E4) ändert $\det A$ um den Faktor -1.*

b) *Addition eines Vielfachen der i-ten Zeile zur j-ten Zeile (mit $i \neq j$)* **ändert** $\det A$ **nicht** *(Umformung E3). Dasselbe gilt für die entsprechende Umformung für die Spalten.*

c) *Multiplikation* **einer** *Zeile (oder Spalte) von A mit einer Zahl α ändert die Determinante um diesen Faktor α (Umformung E2).*

Vorgehen

Bestimmung von $\det A$: Man bringt A auf reduzierte Stufenform und merkt sich in jedem Schritt, wie sich die Determinante ändert.

Es muss also nur noch die Determinante einer reduzierten Stufenform bestimmt werden. Man nennt eine Matrix der Form

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & a_{1,3} & \cdots & a_{1,n} \\ 0 & a_{2,2} & a_{2,3} & \cdots & a_{2,n} \\ \vdots & \ddots & \ddots & & \vdots \\ \vdots & & \ddots & \ddots & a_{n-1,n} \\ 0 & \cdots & \cdots & 0 & a_{n,n} \end{pmatrix}, \qquad \text{also} \quad a_{i,k} = 0 \quad \text{für} \quad i > k,$$

eine OBERE DREIECKSMATRIX.

Satz 8.21 (Determinante oberer Dreiecksmatrizen)

Für eine obere Dreiecksmatrix $A \in Mat(n,n)$ ist $\det A$ das Produkt der Diagonal-Elemente, also

$$\det A = a_{1,1} a_{2,2} \cdots a_{n,n}.$$

Beispiel: Erneute Berechnung der Determinante von $A = \begin{pmatrix} 1 & -1 & 2 & 3 \\ 5 & 2 & 0 & 0 \\ 3 & 1 & 0 & 4 \\ 6 & 7 & -2 & 6 \end{pmatrix}$.

$$
\begin{array}{rrrrl}
1 & -1 & 2 & 3 & \text{tausche Gl.1 und Gl.2} \\
5 & 2 & 0 & 0 & \text{dividiere Gl.2 durch 5} \\
3 & 1 & 0 & 4 & \\
6 & 7 & -2 & 6 & \\
\hline
\boxed{1} & \frac{2}{5} & 0 & 0 & \\
1 & -1 & 2 & 3 & \text{subtrahiere Gl.1} \\
3 & 1 & 0 & 4 & \text{subtrahiere 3*Gl.1} \\
6 & 7 & -2 & 6 & \text{subtrahiere 6*Gl.1} \\
\hline
1 & \frac{2}{5} & 0 & 0 & \\
0 & -\frac{7}{5} & 2 & 3 & \text{tausche Gl.2 und Gl.3} \\
0 & -\frac{1}{5} & 0 & 4 & \text{multipliziere Gl.3 mit -5} \\
0 & \frac{23}{5} & -2 & 6 & \\
\hline
\end{array}
$$

1	$\frac{2}{5}$	0	0	
0	$\boxed{1}$	0	-20	
0	$-\frac{7}{5}$	2	3	addiere $\frac{7}{5}$*Gl.2
0	$\frac{23}{5}$	-2	6	subtrahiere $\frac{23}{5}$*Gl.2
1	$\frac{2}{5}$	0	0	
0	1	0	-20	
0	0	$\boxed{2}$	-25	
0	0	-2	98	addiere Gl.3
1	$\frac{2}{5}$	0	0	
0	1	0	-20	
0	0	2	-25	
0	0	0	73	

Mit 2 Vertauschungen und den Multiplikationen ergibt sich

$$\det A = (-1)^2 \cdot 5 \cdot (-\frac{1}{5}) \cdot 146 = -146.$$

Satz 8.22 (Weitere Regeln zur Determinatenberechnung)

- *Enthält A eine Nullzeile (oder Nullspalte), so ist $\det A = 0$ (klar durch reduzierte Stufenform). Insbesondere gilt für eine (n, n)-Matrix A:*
  $\det A \neq 0 \iff$ *Rang* $A = n$.
- *Falls eine Zeile (oder Spalte) von A das Vielfache einer anderen Zeile (bzw. Spalte) ist, so ist $\det A = 0$.*
- *Es gelten die Rechenregeln*

$$\det A^\top = \det A, \qquad \det A^{-1} = \frac{1}{\det A} \quad (\textit{falls } \det A \neq 0),$$

und der Multiplikationssatz

$$\det(AB) = \det A \ \det B \quad \textit{und} \quad \det(\alpha A) = \alpha^n \det A$$

Achtung: *Es gibt keine Regel für $\det(A + B)$.*

- *Für $n = 2$ und $n = 3$ gilt: Der Absolutbetrag von $\det A$ ist der Flächeninhalt des Parallelogramms bzw. das Volumen des Spats, das von den* **Spaltenvektoren** *von A aufgespannt wird. Dies besitzt eine Verallgemeinerung auf n-dimensionale Körper.*

Wir können nun die folgenden Aussagen zur Lösbarkeit des linearen Gleichungssystems $A\vec{x} = \vec{b}$ mit einer (n, n)-Matrix A formulieren.

Satz 8.23 (Charakterisierung regulärer Matrizen)

Für $A \in Mat(n, n)$ sind folgende Aussagen äquivalent:

(i) *A ist regulär.*

(ii) *Rang $A = n$.*

(iii) $\det A \neq 0$.

(iv) *Das homogene LGS $A\vec{x} = \vec{0}$ hat als einzige Lösung $\vec{x} = \vec{0}$.*

(v) *Das LGS $A\vec{x} = \vec{b}$ ist universell eindeutig lösbar.*

(v) $\dim$ *span* $(\vec{a}_1, \ldots, \vec{a}_n) = n$.

(vi) *$\vec{a}_1, \ldots, \vec{a}_n$ sind linear unabhängig.*

(vii) *$\vec{a}_1, \ldots, \vec{a}_n$ sind Basis des $\mathbb{R}^n$.*

(viii) *Alle Eigenwerte von A sind ungleich Null (später).*

Für kleine Matrizen und weitere Anwendungen ist die folgende Regel wichtig.

Satz 8.24 (Cramersche Regel)

Die Matrix $A \in$ Mat(n,n) sei regulär. Die Komponenten des eindeutigen Lösungsvektors $\vec{x}$ von $A\vec{x} = \vec{b}$ haben die Darstellung

$$x_j = \frac{1}{\det A}\ \det \begin{pmatrix} a_{1,1} & \cdots & a_{1,j-1} & b_1 & a_{1,j+1} & \cdots & a_{1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,j-1} & b_n & a_{n,j+1} & \cdots & a_{n,n} \end{pmatrix}$$

Mit der Cramerschen Regel kann man auch eine Darstellung für die Inverse von A erhalten.

Satz 8.25 (Adjunktenform der Inversen)

Die Inverse einer regulären Matrix A hat die Form

$$A^{-1} = \frac{1}{\det A} \left(\alpha_{j,k}\right)_{n\times n} \quad \textit{mit} \quad \alpha_{j,k} = (-1)^{j+k} \det A_{k,j}.$$

Beispiel zur Cramerschen Regel:

Löse $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \vec{x} = \begin{pmatrix} 7 \\ 8 \end{pmatrix}$.

Möglichkeit 1: Es ist $\det \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = -2$, also nach der Cramerschen Regel

$$x_1 = \frac{\det \begin{pmatrix} 7 & 2 \\ 8 & 4 \end{pmatrix}}{-2} = -6, \quad x_2 = \frac{\det \begin{pmatrix} 1 & 7 \\ 3 & 8 \end{pmatrix}}{-2} = \frac{13}{2}$$

Möglichkeit 2: $\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}^{-1} = \frac{1}{-2} \begin{pmatrix} 4 & -2 \\ -3 & 1 \end{pmatrix}$ und damit

$$\vec{x} = \frac{1}{-2} \begin{pmatrix} 4 & -2 \\ -3 & 1 \end{pmatrix} \begin{pmatrix} 7 \\ 8 \end{pmatrix} = \begin{pmatrix} -6 \\ 13/2 \end{pmatrix}$$

Allgemein: $\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad - bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$, falls $ad - bc \neq 0$; sonst ist die Matrix nicht invertierbar.

Kapitel 9 – Lineare Abbildungen

Kap. 9: Lineare Abbildungen

Die wichtigste Klasse von Funktionen zwischen Vektorräumen sind die LINEAREN ABBILDUNGEN. Wir wiederholen daher Definition 7.23:

Definition 9.1 (Lineare Abbildung)

Eine Funktion $f : V \to W$ zwischen Vektorräumen V und W heißt LINEAR, *wenn für alle $v_1, v_2 \in V$ und alle Skalare α, β gilt*

$$f(\alpha v_1 + \beta v_2) = \alpha f(v_1) + \beta f(v_2).$$

Satz 9.2 (Linearität und Basen)

Gegeben seien Vektorräume V und W sowie eine Basis $v_1, \ldots, v_n$ von V. Jede lineare Abbildung $f : V \to W$ ist durch die Bilder der Basiselemente $f(v_k) \in W$, $1 \leqslant k \leqslant n$, eindeutig festgelegt.

Begründung: *Für einen beliebigen Vektor $x \in V$ mit der eindeutigen Darstellung*

$$x = \sum_{k=1}^{n} \alpha_k v_k$$

gilt aufgrund der Linearität von f:

$$f(x) = \sum_{k=1}^{n} \alpha_k f(v_k).$$

Definition 9.3 (Matrixdarstellung linearer Abbildungen)

Sei $v_1, \dots, v_n$ *eine Basis von* V *und* $w_1, \dots, w_m$ *eine Basis von* W.
Sei $f : V \to W$ *eine lineare Abbildung.*

- *Für die Bilder* $f(v_k)$ *existiert eine eindeutige Darstellung*

  $$f(v_k) = \beta_{1,k} w_1 + \cdots + \beta_{m,k} w_m$$

  mit Skalaren $\beta_{i,k}$, $1 \leqslant i \leqslant m$, $1 \leqslant k \leqslant n$.

- *Die Matrix*

  $$A = \begin{pmatrix} \beta_{1,1} & \cdots & \beta_{1,n} \\ \vdots & & \vdots \\ \beta_{m,1} & \cdots & \beta_{m,n} \end{pmatrix}$$

  heißt die DARSTELLUNGSMATRIX *der linearen Abbildung* f *bzgl. der gegebenen Basen von* V *und* W.

- *Besitzt* $x \in V$ *den Koordinatenvektor* $\vec{a} \in \mathbb{R}^n$ *(oder* $\mathbb{C}^n$*) bzgl. der Basis* $v_1, \dots, v_n$ *von* V, *so ist* $\vec{b} = A\vec{a}$ *der Koordinatenvektor von* $f(x)$ *bzgl. der Basis* $w_1, \dots, w_m$ *von* W.

Wie lassen sich Eigenschaften von f durch Eigenschaften der Darstellungsmatrix ausdrücken?

Satz 9.4 (Eigenschaften der Darstellungsmatrix)

Gegeben seien Vektorräume V und W der Dimensionen $\dim V = n$, $\dim W = m$. Die lineare Abbildung $f : V \to W$ habe die Darstellungsmatrix $A \in \mathrm{Mat}(m, n)$ bzgl. gegebener Basen von V und W. Dann gilt:

- *Der Bildraum $\mathrm{Bild}(f) = \{f(\vec{v}) \mid \vec{v} \in V\}$ ist ein Unter-Vektorraum von W der Dimension $\dim(\mathrm{Bild}(f)) = \mathrm{Rang}\, A$.*
- *f ist genau dann surjektiv, wenn $\mathrm{Rang}\, A = m$ gilt.*
- *f ist genau dann injektiv, wenn $\mathrm{Rang}\, A = n$ gilt.*
- *f ist genau dann bijektiv, wenn $m = n$ und $\mathrm{Rang}\, A = n$ gilt. (Dies ist äquivalent zu $\det A \neq 0$.)*

Wir betrachten ab jetzt lineare Abbildungen $f : V \to V$ und $V = \mathbb{R}^n$ oder $\mathbb{C}^n$. Im Definitions- und im Wertebereich verwenden wir die gleiche Basis $\vec{v}_1, \dots, \vec{v}_n$ zur Bestimmung der Matrixdarstellung von f.

Satz 9.5 (Basiswechsel)

Gegeben sei eine lineare Abbildung $f : V \to V$ (mit $V = \mathbb{R}^n$ oder $\mathbb{C}^n$) sowie die Darstellungsmatrix $A \in Mat(n,n)$ von f bzgl. der Basis $\vec{v}_1, \dots, \vec{v}_n$ (sowohl im Definitions- als auch im Wertebereich). Dann gilt:

- *die Darstellungsmatrix von f bzgl. einer weiteren Basis $\vec{w}_1, \dots, \vec{w}_n$ von V ist gegeben durch*

$$B = S^{-1}AS.$$

Hierbei ist $S \in Mat(n,n)$ die reguläre Matrix, deren Spaltenvektor $\vec{s}_k = (s_{1,k}, \dots, s_{n,k})^\top$ der Koordinatenvektor von $\vec{w}_k$ bzgl. der Basis $\vec{v}_1, \dots, \vec{v}_n$ ist, also

$$\vec{w}_k = s_{1,k}\vec{v}_1 + \dots + s_{n,k}\vec{v}_n, \qquad 1 \leqslant k \leqslant n.$$

Die Darstellungsmatrizen A und $B = S^{-1}AS$ heißen ÄHNLICH.

Bemerkung: Ähnliche Matrizen haben den gleichen Rang und dieselbe Determinante:

$$\text{Rang}(S^{-1}AS) = \text{Rang}(AS) = \text{Rang}\, A,$$

weil die Multiplikation mit regulären Matrizen (hier S bzw. S^{-1}) den Rang nicht verändert. Mit dem Multiplikationssatz aus Satz 8.22 folgt

$$\det(S^{-1}AS) = \det S^{-1} \ \det A \ \det S = \frac{1}{\det S} \ \det A \ \det S = \det A.$$

Motivation für die Eigenvektoren

Eine schöne Interpretation vieler linearer Abbildungen f erhält man dadurch, dass man durch geschickte Wahl der Basis eine DIAGONALMATRIX

$$B = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}, \qquad \text{also} \quad b_{i,k} = 0 \quad \text{für} \quad i \neq k,$$

als Darstellungsmatrix erhält. Dann lässt f die Richtung der Basisvektoren $\vec{v}_k$ (bis auf Umkehrung für $\lambda_k < 0$) unverändert. Dadurch erhält man eine geometrische Interpretation von f.

Definition 9.6 (Eigenwert, Eigenvektor)

Gegeben sei eine lineare Abbildung $f : V \to V$. *Ein Vektor* $\vec{v} \in V$, $\vec{v} \neq \vec{0}$ *heißt* EIGENVEKTOR *(kurz: EV) von* f, *wenn es ein Skalar* λ *gibt mit*

$$f(\vec{v}) = \lambda \vec{v}.$$

Das Skalar λ *heißt dann der* EIGENWERT *(kurz EW) zum* EIGENVEKTOR $\vec{v} \neq \vec{0}$.

Die Berechnung von Eigenwerten und Eigenvektoren erfolgt über eine Matrix-Darstellung von f in zwei Schritten.

Satz 9.7 (Berechnung von EW und EV)

Gegeben sei eine lineare Abbildung $f : V \to V$ sowie eine Basis $\vec{v}_1, \ldots, \vec{v}_n$ von V (also $\dim V = n$).

- *Das Skalar λ ist genau dann ein Eigenwert von f, wenn für die Darstellungsmatrix $A \in \mathit{Mat}(n,n)$ von f gilt*

$$\det(A - \lambda E_n) = 0. \tag{$*$}$$

- *Zu einem Eigenwert λ ist $\mathit{Kern}(A - \lambda E_n)$ der Unter-Vektorraum der Koordinatenvektoren aller Eigenvektoren von f zum Eigenwert λ (sowie zusätzlich $\vec{0}$):*

$$\begin{aligned} V_\lambda &= \{\vec{x} \in V \mid f(\vec{x}) = \lambda \vec{x}\} \\ &= \left\{ \vec{x} \in V \mid \vec{x} = \sum_{k=1}^{n} \alpha_k \vec{v}_k,\ (\alpha_1, \ldots, \alpha_n)^\top \in \mathit{Kern}(A - \lambda E_n) \right\}. \end{aligned}$$

Bemerkung:

- Im ersten Schritt bestimmt man alle Eigenwerte von f, indem man alle Nullstellen des CHARAKTERISTISCHEN POLYNOMS

$$p_A(\lambda) = \det(A - \lambda E_n)$$

der Darstellungsmatrix A berechnet. Nach dem Fundamentalsatz der Algebra gibt es (in $\mathbb{C}$) genau n Eigenwerte unter Berücksichtigung der Vielfachheit m_λ der Nullstelle λ. Die Zahl m_λ heißt die ALGEBRAISCHE VIELFACHHEIT des Eigenwerts λ.

- Im zweiten Schritt berechnet man zu jedem Eigenwert λ eine Basis von Kern$(A - \lambda E_n)$ durch Lösen des homogenen LGS

$$(A - \lambda E_n)\vec{x} = \vec{0}.$$

Wegen $\det(A - \lambda E_n) = 0$ gibt es nicht-triviale Lösungen. Die Zahl

$$\widetilde{m}_\lambda = \mathsf{dim}(\mathsf{Kern}(A - \lambda E_n)) = n - \mathsf{Rang}(A - \lambda E_n)$$

nennt man die GEOMETRISCHE VIELFACHHEIT des Eigenwerts λ.

- Es gilt stets $\widetilde{m}_\lambda \leqslant m_\lambda$. (o.B.)

Weitere Eigenschaften des charakteristschen Polynoms

(i) Die Summe der Eigenwerte ist die SPUR der Matrix A: der Koeffizient von $(-1)^{n-1}\lambda^{n-1}$ ist gleich spur $A := \sum_{i=1}^{n} a_{ii}$.

(ii) Das Produkt der Eigenwerte ist die Determinate von A, daher ist eine Matrix genau dann invertierbar, wenn Null kein Eigenwert ist.

(iii) Die Eigenwerte der inversen Matrix sind die Inversen der Eigenwerte.

(iv) Für eine 2×2-Matrix ist stets $p(\lambda) = \lambda^2 - \operatorname{spur} A \cdot \lambda + \det A$.

(v) Allgemein ist bei einer $n \times n$-Matrix A

$$p(\lambda) = (-1)^n \lambda^n + (-1)^{n-1} \operatorname{spur} A \lambda^{n-1} + \cdots + \det A$$

Satz 9.8 (Eigenwerte und Basiswechsel)

Die Eigenwerte und Eigenvektoren der linearen Abbildung $f : V \to V$ hängen nicht von der gewählten Basis von V ab; insbesondere gilt für zwei ähnliche Matrizen A und $B = S^{-1}AS$ (mit regulärer Matrix S) die Identität

$$p_A(\lambda) = p_B(\lambda).$$

Für die Koordinatenvektoren gilt: Ist $\vec{a}$ der Koordinatenvektor des Eigenvektors $\vec{v}$ zu der Basis, die die Darstellungsmatrix A von f definiert, so ist $\vec{b} = S^{-1}\vec{a}$ der Koordinatenvektor desselben Eigenvektors $\vec{v}$ zu der Basis, die die Darstellungsmatrix $B = S^{-1}AS$ definiert (siehe Basiswechsel 9.5).

Die anfangs erwähnte “schöne” Darstellung einer linearen Abbildung durch eine Diagonalmatrix kann nur dann erzielt werden, wenn die folgende Eigenschaft gilt.

Definition 9.9 (diagonalisierbar)

Eine lineare Abbildung $f : V \to V$ (mit $\dim V = n$) heißt DIAGONALISIERBAR, *wenn es eine Basis von Eigenvektoren $\vec{v}_1, \ldots, \vec{v}_n$ von f gibt.*

Dann gilt:

- Die Darstellungsmatrix bzgl. der Basis der Eigenvektoren ist die Diagonalmatrix der Eigenwerte

$$J = \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & \lambda_n \end{pmatrix}.$$

- Die Darstellungsmatrix A bzgl. irgendeiner Basis $\vec{w}_1, \ldots, \vec{w}_n$ ist ähnlich zu dieser Diagonalmatrix, d.h. es gibt eine reguläre Matrix S mit $S^{-1}AS = J$.
- Die hierzu äquivalente Beziehung $AS = SJ$ drückt aus, dass die Spalten von S die *Eigenvektoren der Matrix* A mit zugehörigen Eigenwerten $\lambda_1, \ldots, \lambda_n$ sind.

Wir behandeln zwei Fälle, in denen die Diagonalisierbarkeit vorliegt.

Satz 9.10 (Kriterium für Diagonalisierbarkeit)

Gegeben sei eine lineare Abbildung $f : V \to V$ mit $\dim V = n$. Besitzt f paarweise verschiedene Eigenwerte $\lambda_1, \ldots, \lambda_n$, so ist f diagonalisierbar.

Als Beweis dient die folgende Aussage:

Satz 9.11 (lineare Unabhängigkeit der Eigenvektoren)

Eigenvektoren $\vec{v}_1, \ldots, \vec{v}_k$ einer linearen Abbildung $f : V \to V$ zu paarweise verschiedenen Eigenwerten $\lambda_1, \ldots, \lambda_k$ sind linear unabhängig.

Eine Erweiterung von Satz 9.11 ergibt die folgende Aussage.

Satz 9.12 (Ein weiteres Kriterium)

Die lineare Abbildung $f : V \to V$ (mit $\dim V = n$) ist genau dann diagonalisierbar, wenn für jeden Eigenwert λ von f die algebraische Vielfachheit und die geometrische Vielfachheit übereinstimmen, wenn also $m_\lambda = \tilde{m}_\lambda$ gilt.

Eine Basis von Eigenvektoren von f erhält man, indem man für jeden Eigenraum V_λ eine Basis aus m_λ Vektoren bestimmt. Die Gesamtheit dieser Vektoren ist dann eine Basis des Vektorraums V.

Die zweite Klasse diagonalisierbarer Matrizen tritt in vielen Anwendungen auf.

Definition 9.13 (symmetrische und hermitesche Matrizen)

- *Eine reelle $n \times n$-Matrix A heißt* SYMMETRISCH, *wenn $A^\top = A$ gilt.*
- *Eine komplexe $n \times n$-Matrix A heißt* HERMITESCH, *wenn $A^* = A$ gilt.*

Satz 9.14 (Diagonalisierbarkeit symmetrischer und hermitescher Matrizen)

Sei A eine reell-symmetrische oder einer komplex-hermitesche $n \times n$-Matrix. Dann gilt:

- *Alle Eigenwerte von A sind reell und A ist diagonalisierbar.*
- *Es existiert eine Orthonormalbasis $\vec{v}_1, \ldots, \vec{v}_n$ von Eigenvektoren von A.*

Das bedeutet, dass A eine Darstellung

$$A = UDU^\top \qquad \textit{bzw.} \qquad A = UDU^*$$

mit einer reellen Diagonalmatrix D und einer orthogonalen bzw. unitären Matrix U hat.

ORTHOGONAL bzw. UNITÄR nennt man Matrizen U, deren Spalten eine ONB bilden. Für solche Matrizen gilt, dass

$$U^\top = U^{-1} \quad \text{bzw.} \quad U^* = U^{-1}$$

ist. Dieses wird in Kapitel 10 genauer behandelt.

Zur Berechnung der Eigenvektoren und zur Probe sind die folgenden Aussagen hilfreich, die eine Verschärfung von Satz 9.14 darstellen:

Satz 9.15 (Weitere Eigenschaften)

(a) *Eigenvektoren $\vec{v}_1, \ldots, \vec{v}_k$ einer reell-symmetrischen (oder hermiteschen) Matrix zu paarweise verschiedenen Eigenwerten sind paarweise orthogonal.*

(b) *Für einen Eigenwert λ einer reell-symmetrischen (oder hermiteschen) Matrix $A \in \mathit{Mat}(n,n)$ bezeichne m_λ die algebraische Vielfachheit (als Nullstelle des charakteristischen Polynoms). Dann gilt*

$$\mathit{dim}(\mathit{Kern}(A - \lambda E_n)) = m_\lambda,$$

d.h. die algebraische und die geometrische Vielfachheit des Eigenwerts stimmen überein.

Zusammenfassend erhalten wir die folgende Aussage.

Satz 9.16 (Entwicklungssatz)

Sei A eine symmetrische oder hermitesche Matrix mit einer ONB aus Eigenvektoren $\vec{v}_1$ bis $\vec{v}_n$.

Dann lässt sich jeder Vektor des $\mathbb{R}^n$ ($\mathbb{C}^n$) nach $\vec{v}_1$ bis $\vec{v}_n$ ENTWICKEN:

$$\vec{x} = \sum_{k=1}^{n} < \vec{x}, \vec{v_k} > \vec{v}_k.$$

Weiter gilt mit den Eigenwerten λ_1 bis λ_n:

$$A\vec{x} = \sum_{k=1}^{n} \lambda_k < \vec{x}, \vec{v_k} > \vec{v}_k.$$

Es folgen ergänzende Aussagen zu linearen Abbildungen.

Satz 9.17 (Der Vektorraum der linearen Abbildungen)

Gegeben seien Vektorräume V und W. Die Menge aller linearen Abbildungen

$$\mathcal{L}(V, W) = \{f : V \to W \mid f \text{ ist linear}\}$$

ist ein Vektorraum. (Die Addition $f + g$ und die Multiplikation mit Skalaren αf ist wie üblich bei Funktionen definiert.)

Satz 9.18 (Verkettung linearer Abbildungen)

Es seien U, V, W Vektorräume und $f : U \to V$ sowie $g : V \to W$ lineare Abbildungen. Dann ist auch die Hintereinanderausführung

$$g \circ f : U \to W, \quad g \circ f(\vec{u}) = g(f(\vec{u}))$$

linear.

Weiterhin sei $\mathcal{U} = \vec{u}_1, \ldots, \vec{u}_n$ eine Basis von U, $\mathcal{V} = \vec{v}_1, \ldots, \vec{v}_m$ eine Basis von V sowie $\mathcal{W} = \vec{w}_1, \ldots, \vec{w}_p$ eine Basis von W. Mit den Darstellungsmatrizen $A \in Mat(m, n)$ von f (bzgl. $\mathcal{U}, \mathcal{V}$) und $B \in Mat(p, m)$ von g (bzgl. $\mathcal{V}, \mathcal{W}$) ist dann

$$C = BA$$

die Darstellungsmatrix von $g \circ f$ bzgl. der Basen $\mathcal{U}, \mathcal{W}$.

Als "Multiplikation" auf dem Vektorraum $\mathcal{L}(V,V)$ der Endomorphismen (=lineare Abbildungen von V nach V) verwenden wir die Hintereinanderausführung. Dann gilt:

Satz 9.19 (Der Ring der linearen Abbildungen)

Der Vektorraum $\mathcal{L}(V,V)$ ist ein $\textsc{Ring}$, d.h.

- *Die Hintereinanderausführung erfüllt das Assoziativ- und Distributivgesetz.*
- *Das neutrale Element dieser Multiplikation ist die Einheitsmatrix E_n.*

Kapitel 10 – Spezielle Matrizen und Abbildungen

Interessant sind lineare Abbildungen, die die Länge von Vektoren nicht verändern. Wichtige Beispiele sind Drehungen und Spiegelungen.

Definition 10.1 (Matrix)

- *Eine (quadratische) reelle Matrix heißt* ORTHOGONAL, *falls* $A^\top = A^{-1}$ *ist.*
- *Eine (quadratische) komplexe Matrix heißt* UNITÄR, *falls* $A^* = A^{-1}$ *ist.*

Äquivalent ist

Satz 10.2 (Eigenschaften orthogonaler Matrizen)

(i) A *ist orthogonal.*

(ii) $A^\top = A^{-1}$.

(iii) *Die Spalten von* A *bilden eine ONB.*

(iv) *Die Zeilen von* A *bilden eine ONB.*

(v) *Für* $\vec{v}$, $\vec{w} \in \mathbb{R}^n$ *gilt* $< \vec{v}, \vec{w} > = < A\vec{v}, A\vec{w} >$.

(vi) *Für* $\vec{v} \in \mathbb{R}^n$ *gilt* $\|A\vec{v}\| = \|\vec{v}\|$.

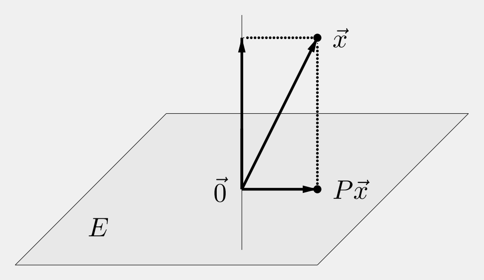
Satz 10.3 (Weitere Eigenschaften)

Sei A orthogonal.

- *Für $\vec{v}, \vec{w}$ gilt $\sphericalangle(\vec{v}, \vec{w}) = \sphericalangle(A\vec{v}, A\vec{w})$*
- *Die Determinante von A ist ± 1*

Satz 10.4 (Orthogonale Projektionen)

Eine ORTHOGONALE PROJEKTION P *ist dadurch gekennzeichnet, dass gilt* $P^2 = P$ *und* $P = P^\top$.



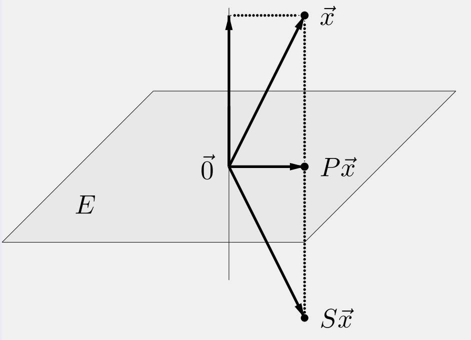
Spezialfall:
Projektion auf eine Hyperebene.
$\vec{n}$ ist ein Vektor mit $\|\vec{n}\| = 1$, der senkrecht auf der Hyperebene steht.
$P\vec{x} = \vec{x} - < \vec{x}, \vec{n} > \vec{n} = (E - \vec{n}\vec{n}^\top)\vec{x}$.
Die Matrix ist also $A = E - \vec{n}\vec{n}^\top$.

Die allgemeine Form einer orthogonalen Projektion ist wie folgt:

Sei $\vec{v}_1$ bis $\vec{v}_n$ eine ONB, so dass $\vec{v}_1$ bis $\vec{v}_k$ Basis eines Unterraums U ist.
Dann ist die orthogonale Projektion auf U gegeben durch

$$P_U\vec{x} = \sum_{j=1}^{k} < \vec{x}, \vec{v}_j > \vec{v}_j = \left(\sum_{j=1}^{k} \vec{v}_j\vec{v}_j^\top\right)\vec{x}$$

Satz 10.5 (Spiegelungen)



Verdoppelt man die Strecke auf den Projektionspunkt, so sieht man die Form der Spiegelungsmatrix an einer Hyperebene:

$$S\vec{x} = (E - 2\vec{n}\vec{n}^\top)\vec{x}$$

$$S = E - 2\vec{n}\vec{n}^\top$$

Mit denselben Voraussetzungen wie bei der Projektion ist die allgemeine Form

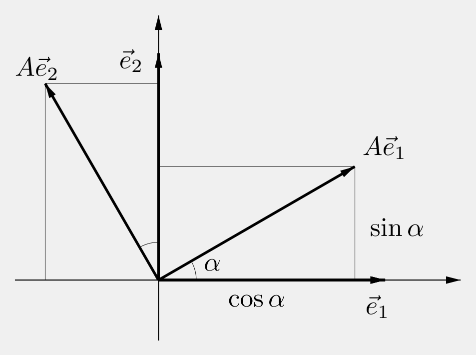
$$S_U\vec{x} = \sum_{j=1}^{k} < \vec{x}, \vec{v}_j > \vec{v}_j - \sum_{j=k+1}^{n} < \vec{x}, \vec{v}_j > \vec{v}_j = \Big(\sum_{j=1}^{k} \vec{v}_j\vec{v}_j^\top - \sum_{j=k+1}^{n} \vec{v}_j\vec{v}_j^\top \Big)\vec{x}$$

Satz 10.6 (Drehungen)

Drehungen um den Winkel α im $\mathbb{R}^2$ haben die Matrixdarstellung

$$A = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix}$$

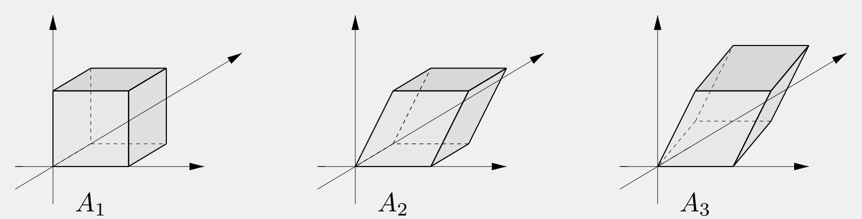
Drehmatrizen sind orthogonal.



Satz 10.7 (Scherungen)

Lineare Abbildungen, bei denen die geometrische Vielfachheit eines Eigenwerts kleiner als die algebraische ist, sind SCHERUNGEN.

Beispiele für Scherungen:



$$A_1 = E_3 \quad A_2 = \begin{pmatrix} 1 & 0.5 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \text{ und } A_3 = \begin{pmatrix} 1 & 0.5 & 0 \\ 0 & 1 & 0.3 \\ 0 & 0 & 1 \end{pmatrix}$$

Die Skizzen wurden mit der Projektionsmatrix $P = \begin{pmatrix} 1 & 0 & 0.5 \\ 0 & 1 & 0.3 \\ 0 & 0 & 0 \end{pmatrix}$ erstellt.

Kapitel 11 – Folgen und Reihen

Definition 11.1 (Zahlenfolge)

Eine Funktion

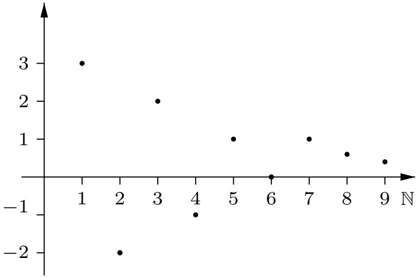
$$a : \mathbb{N} \to \mathbb{R} \qquad (\textit{oder } \mathbb{C})$$

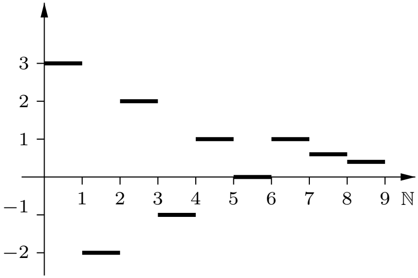
heißt reelle (oder komplexe) ZAHLENFOLGE. *Man nennt* $a_n = a(n)$ *das* n*-te* FOLGENGLIED *und schreibt kurz* $(a_n)_{n\in\mathbb{N}}$ *oder* (a_n).

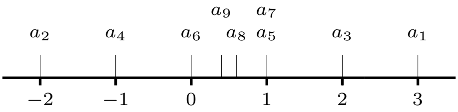
- Ebenso sind $(a_n)_{n\in\mathbb{N}_0}$ oder $(a_n)_{n\geqslant n_0}$ mit festem $n_0 \in \mathbb{Z}$ Zahlenfolgen.
- Eine TEILFOLGE entsteht, indem man (endlich oder unendlich viele) Folgenglieder weglässt, wobei noch unendlich viele Glieder übrigbleiben müssen: Für natürliche Zahlen $1 \leqslant n_1 < n_2 < n_3 < \cdots$ ist

$$(a_{n_k})_{k\in\mathbb{N}} \text{ eine Teilfolge von } (a_n)_{n\in\mathbb{N}}.$$

- Später behandeln wir auch Folgen von Vektoren $(\vec{v}_n)_{n\in\mathbb{N}}$ mit $\vec{v}_n \in \mathbb{R}^d$, Folgen von Matrizen $(A_n)_{n\in\mathbb{N}}$ mit $A_n \in \mathrm{Mat}(p,q)$ oder Funktionenfolgen $(f_n)_{n\in\mathbb{N}}$ mit $f_n : [a,b] \to \mathbb{R}$.







Definition 11.2 (Konvergenz, Grenzwert)

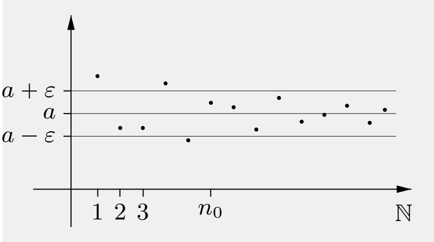
Eine Zahlenfolge $(a_n)_{n\in\mathbb{N}}$ *heißt* KONVERGENT, *wenn es eine Zahl* $a \in \mathbb{R}$ *(oder* $\mathbb{C}$*) mit der folgenden Eigenschaft gibt: zu jedem* $\epsilon > 0$ *gibt es ein* $n_0 \in \mathbb{N}$*, so dass für alle* $n \geqslant n_0$ *die Ungleichung*

$$|a_n - a| < \epsilon$$

gilt. Dann heißt a *der* GRENZWERT *der Folge* $(a_n)_{n\in\mathbb{N}}$*, die Folge "*KONVERGIERT GEGEN a*" und wir schreiben*

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

Ist die Zahlenfolge $(a_n)_{n\in\mathbb{N}}$ *nicht konvergent, so heißt sie* DIVERGENT.



Einfache Hilfsmittel zur Berechnung von Grenzwerten:

Satz 11.3 (Grenzwertsatz)

Die Zahlenfolgen $(a_n)_{n\in\mathbb{N}}$ *und* $(b_n)_{n\in\mathbb{N}}$ *seien konvergent, und* $a = \lim\limits_{n\to\infty} a_n$ *sowie* $b = \lim\limits_{n\to\infty} b_n$ *seien die Grenzwerte. Dann gilt:*

a) $(a_n + b_n)_{n\in\mathbb{N}}$ *ist konvergent und* $\lim\limits_{n\to\infty}(a_n + b_n) = a + b.$

b) *Für beliebiges* $\alpha \in \mathbb{C}$ *ist* $(\alpha a_n)_{n\in\mathbb{N}}$ *konvergent und* $\lim\limits_{n\to\infty}(\alpha a_n) = \alpha a.$

c) $(a_n\, b_n)_{n\in\mathbb{N}}$ *ist konvergent und* $\lim\limits_{n\to\infty}(a_n\, b_n) = ab.$

d) *Sind alle* $b_n \neq 0$ *und gilt* $b \neq 0$, *so ist* $\left(\dfrac{a_n}{b_n}\right)_{n\in\mathbb{N}}$ *konvergent und*

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

Definition 11.4 (Nullfolge)

a) *Eine konvergente Folge* $(a_n)_{n\in\mathbb{N}}$ *mit* $\lim\limits_{n\to\infty} a_n = 0$ *heißt* NULLFOLGE.

b) $a_n \to 0 \iff |a_n| \to 0$

c) *Eine reelle Zahlenfolge heißt* BESTIMMT DIVERGENT GEGEN ∞ *(bzw. gegen* $-\infty$*), wenn für jedes* $r > 0$ *ein* $n_0 \in \mathbb{N}$ *existiert, so dass für alle* $n \geqslant n_0$ *die Ungleichung*

$$a_n > r \qquad \textit{(bzw. } a_n < -r\textit{)}$$

gilt. Dann heißt ∞ *(bzw.* $-\infty$ *) der* UNEIGENTLICHE GRENZWERT *der Folge und wir schreiben*

$$\lim_{n\to\infty} a_n = \infty \qquad (\textit{ bzw. } \lim_{n\to\infty} a_n = -\infty) \quad \textit{oder} \quad a_n \to \pm\infty.$$

Bemerkung:

- $(a_n)_{n\in\mathbb{N}}$ konvergiert gegen a genau dann, wenn $(a_n - a)_{n\in\mathbb{N}}$ eine Nullfolge ist.
- $a_n \to \infty \iff$ Für $n \geqslant n_0$ ist $a_n > 0$ und $\dfrac{1}{a_n} \to 0$.

Satz 11.5 (Konvergenz und Ordnung)

a) *Sei* $\lim\limits_{n\to\infty} a_n = a$ *und* $\lim\limits_{n\to\infty} b_n = b$. *Falls es ein* $n_0 \in \mathbb{N}$ *mit* $a_n \leqslant b_n$ *für alle* $n \geqslant n_0$ *gibt, so ist auch* $a \leqslant b$.

b) *Gegeben seien drei reelle Zahlenfolgen* (a_n), (b_n) *und* (c_n) *und es gelte* $\lim\limits_{n\to\infty} a_n = \lim\limits_{n\to\infty} c_n = s$. *Falls es ein* $n_0 \in \mathbb{N}$ *mit*

$$a_n \leqslant b_n \leqslant c_n \qquad \text{für alle} \quad n \geqslant n_0$$

gibt, so ist auch (b_n) *konvergent und* $\lim\limits_{n\to\infty} b_n = s$.

Bemerkung: Kriterium (b) heißt Einschliessungskriterium oder Sandwich-Lemma.

Beispiel 11.6

a) $\left(\sqrt{n+1}-\sqrt{n}\right)_{n\in\mathbb{N}}$ *konvergiert gegen* 0.

b) $(\sqrt[n]{n})_{n\in\mathbb{N}}$ *konvergiert gegen* 1.

c) *Für jedes* $c>0$ *konvergiert* $(\sqrt[n]{c})_{n\in\mathbb{N}}$ *gegen* 1.

d) *Für jedes feste* $\ell\in\mathbb{N}$ *konvergiert* $\left(\sqrt[n]{n^\ell}\right)_{n\in\mathbb{N}}$ *gegen* 1 *(folgt aus b) und dem Grenzwertsatz).*

e) *Die Folge* $(z^n)_{n\in\mathbb{N}}$ *mit* $z\in\mathbb{C}$ *und* $|z|<1$ *ist eine Nullfolge.*

Hilfreich ist die folgende Wachstumshierarchie. Weiter rechts stehende Folgen gehen schneller gegen Unendlich als linksstehende. Dabei sei $\alpha>0$ und $q>1$.

1	$\ln n$	n^α	q^n	$n!$	n^n

Beispiel: $\lim\limits_{n\to\infty}\dfrac{n^7}{2^n}=0.$

Wichtiges Hilfsmittel ist die Stirling-Formel

$$n! \approx \left(\frac{n}{e}\right)^n \sqrt{2\pi n}$$

- Der Quotient der beiden Terme hat den Grenzwert 1.
- Die Differenz geht gegen Unendlich. Es gibt noch weit genauere Abschätzungen.

Zur Klarstellung dient der folgende Satz.

Satz 11.7

Die Folge $(a_n)_{n\in\mathbb{N}}$ *sei konvergent.*

a) *Der Grenzwert* a *von* $(a_n)_{n\in\mathbb{N}}$ *ist eindeutig bestimmt.*
b) *Jede Teilfolge* $(a_{n_k})_{k\in\mathbb{N}}$ *ist konvergent und hat denselben Grenzwert* a.

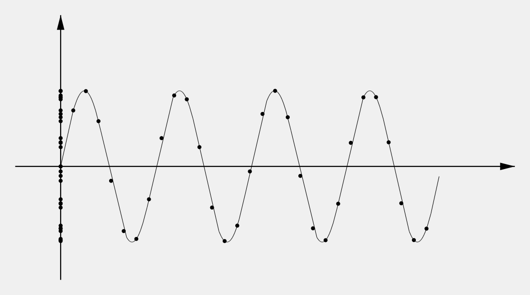
Ein wichtiger Begriff im Zusammenhang mit Teilfolgen:

Definition 11.8 (Häufungswert)

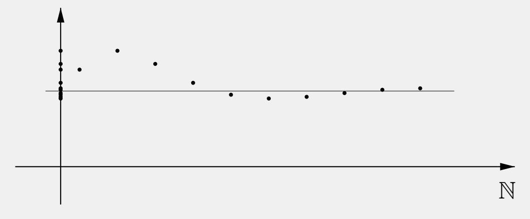
Die Zahl b *heißt ein* HÄUFUNGSWERT *der Folge* $(a_n)_{n\in\mathbb{N}}$, *wenn es eine konvergente Teilfolge* $(a_{n_k})_{k\in\mathbb{N}}$ *mit Grenzwert* b *gibt, d.h.*

$$\lim_{k\to\infty} a_{n_k} = b.$$

Beispiele



Die durch $a_n = \sin \frac{5}{6} n$ definierte Folge hat jedes $x \in [-1, 1]$ als Häufungswert.



Bei einer konvergenten Folge ist der Grenzwert der einzige Häufungswert.

Definition 11.9 (beschränkte und monotone Folge)

a) *Eine Folge* $(a_n)_{n\in\mathbb{N}}$ *heißt* BESCHRÄNKT, *wenn es ein* $r > 0$ *gibt mit* $|a_n| \leqslant r$ *für alle* $n \in \mathbb{N}$.

b) *Eine reelle Zahlenfolge* $(a_n)_{n\in\mathbb{N}}$ *heißt*

- MONOTON WACHSEND, *wenn* $a_{n+1} \geqslant a_n$ *für alle* $n \in \mathbb{N}$ *gilt,*
- MONOTON FALLEND, *wenn* $a_{n+1} \leqslant a_n$ *für alle* $n \in \mathbb{N}$ *gilt,*
- MONOTON, *wenn sie monoton wachsend oder monoton fallend ist.*

Neben Satz 11.3 lautet das wichtigste “Konvergenz-Kriterium” wie folgt:

Satz 11.10 (Monotoniekriterium)

Jede monotone und beschränkte reelle Zahlenfolge ist konvergent.

Bemerkung: Eine andere Formulierung, die auch für komplexe Zahlenfolgen gültig ist, lautet:

Satz 11.11 (Satz von Bolzano und Weierstraß)

Jede beschränkte Folge (in $\mathbb{R}$ oder $\mathbb{C}$) besitzt eine konvergente Teilfolge.

Genauer: Ist die Folge $(a_n)_{n\in\mathbb{N}}$ beschränkt, so existiert eine Teilfolge $(a_{n_k})_{k\in\mathbb{N}}$ und eine Zahl b mit

$$\lim_{k\to\mathbb{N}} a_{n_k} = b.$$

Satz 11.12 (Intervallschachtelung)

Gegeben seien eine monoton wachsende reelle Zahlenfolge $(a_n)_{n\in\mathbb{N}}$ *und eine monoton fallende reelle Zahlenfolge* $(b_n)_{n\in\mathbb{N}}$*. Es gelte*

(i) $a_n \leqslant b_n$ *für alle* $n \in \mathbb{N}$ *und*

(ii) $\lim\limits_{n\to\infty}(b_n - a_n) = 0$.

Man sagt, dass die abgeschlossenen Intervalle $[a_n, b_n] \subseteq [a_{n-1}, b_{n-1}]$ *eine* INTERVALLSCHACHTELUNG *definieren. Dann enthält der Durchschnitt*

$$\bigcap_{n=1}^{\infty}[a_n, b_n]$$

genau eine reelle Zahl s*, nämlich*

$$s = \lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n.$$

Beispiel 11.13 (Eulersche Zahl e)

Die Eulersche Zahl $e = 2.718281828459 04...$ *kann wie folgt definiert werden: Wir betrachten die Folgen* $(a_n)_{n\in\mathbb{N}}$ *und* $(b_n)_{n\in\mathbb{N}}$ *mit den Folgengliedern*

$$a_n = \left(1 + \frac{1}{n}\right)^n, \qquad b_n = \left(1 + \frac{1}{n}\right)^{n+1}.$$

Dann gilt:

1. *Die Folge* $(a_n)_{n\in\mathbb{N}}$ *ist monoton wachsend.*
2. *Die Folge* $(b_n)_{n\in\mathbb{N}}$ *ist monoton fallend.*
3. *Es ist* $b_n - a_n \geqslant 0$ *und* $\lim\limits_{n\to\infty}(b_n - a_n) = 0$*.*
4. *Das Prinzip der Intervallschachtelung ergibt:*

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n =: e.$$

Zahlenwerte: "sehr langsame Konvergenz" gegen e

n	a_n	b_n
1	2	4
10	2.59374246010000	2.85311670611000
100	2.70481382942153	2.73186196771574
1000	2.71692393223559	2.71964085616783
1000000	$\underline{2.7182}6823719230$	$\underline{2.7182}8318737622$

Definition 11.14 (Limes superior und Limes inferior)

Gegeben sei eine reelle Zahlenfolge $(a_n)_{n\in\mathbb{N}}$.

- *Falls* $(a_n)_{n\in\mathbb{N}}$ *beschränkt ist, sind der größte Häufungswert (*Limes superior*)*

$$\limsup_{n\to\infty} a_n := \max\{b | b \textit{ ist Häufungswert von } (a_n)_{n\in\mathbb{N}}\}$$

  *und der kleinste Häufungswert (*Limes inferior*)*

$$\liminf_{n\to\infty} a_n := \min\{b | b \textit{ ist Häufungswert von } (a_n)_{n\in\mathbb{N}}\}$$

  definiert.

- *Falls* $(a_n)_{n\in\mathbb{N}}$ *nicht nach oben beschränkt ist (d.h. zu jedem* $r > 0$ *existiert ein* $n \in \mathbb{N}$ *mit* $a_n > r$*), so setzen wir* $\limsup_{n\to\infty} a_n := \infty$.

  Falls $(a_n)_{n\in\mathbb{N}}$ *nicht nach unten beschränkt ist (d.h. zu jedem* $r > 0$ *existiert ein* $n \in \mathbb{N}$ *mit* $a_n < -r$*), so setzen wir* $\liminf_{n\to\infty} a_n := -\infty$.

Als partielle Umkehrung des Satzes von Bolzano-Weierstraß geben wir noch folgendes Resultat an:

Satz 11.15 (Aus konvergent folgt beschränkt)

Jede konvergente Folge $(a_n)_{n\in\mathbb{N}}$ *ist beschränkt.*

Als wichtiges (abstraktes) Konvergenz-Kriterium dient:

Satz 11.16 (Cauchy-Kriterium)

Eine Folge $(a_n)_{n\in\mathbb{N}}$ *ist genau dann konvergent, wenn zu jedem* $\epsilon > 0$ *ein* $n_0 \in \mathbb{N}$ *existiert, so dass für alle* $n > m \geqslant n_0$ *die folgende Ungleichung erfüllt ist:*

$$|a_n - a_m| < \epsilon$$

Definition 11.17 (Reihen)

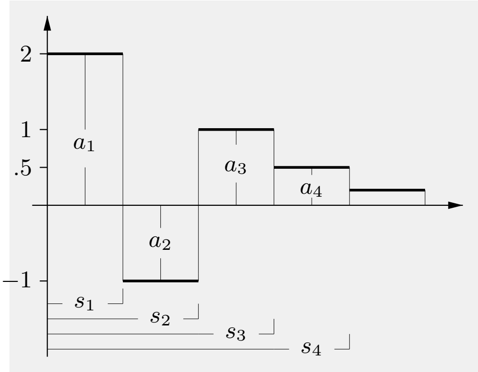
- *Die zu einer gegebenen Zahlenfolge* $(a_n)_{n\in\mathbb{N}}$ *gebildete Folge* $(s_n)_{n\in\mathbb{N}}$ *der* PARTIALSUMMEN
$$s_n = \sum_{k=1}^{n} a_k$$
heißt eine UNENDLICHE REIHE. *Der Summand* a_k *heißt* k*-tes Reihenglied.*

- *Anstatt* $(s_n)_{n\in\mathbb{N}}$ *schreibt man kurz* $\sum\limits_{n=1}^{\infty} a_n$ *für die unendliche Reihe.*

- *Falls die unendliche Reihe* $\sum\limits_{n=1}^{\infty} a_n$ *gegen die Zahl* $s \in \mathbb{C}$ *konvergiert, so schreiben wir*
$$s = \lim_{n\to\infty} s_n = \sum_{n=1}^{\infty} a_n.$$

- *Eine Reihe heißt* ABSOLUT KONVERGENT, *wenn* $\sum\limits_{k=1}^{n} |a_k|$ *konvergiert.*

Beispiel:



- Mit $a_1 = 2$, $a_2 = -1$, $a_3 = 1$, $a_4 = 0.5$, $\dots$ entspricht die Partialsumme s_n dem orientierten Flächeninhalt unter den ersten n Kästchen.
- In diesem Beispiel: $s_1 = 2$, $s_2 = 1$, $s_3 = 2$, $s_4 = 2.5$, $s_5 = \dots$. Die Reihe konvergiert, wenn die Folge der Partialsummen einen Grenzwert hat.

Bemerkung:

- Unendliche Reihen sind also spezielle Zahlenfolgen. Die Begriffe Konvergenz und Grenzwert, Beschränktheit und Monotonie einer unendlichen Reihe beziehen sich immer auf die Folge der Partialsummen $(s_n)_{n\in\mathbb{N}}$ (bzw. $(s_n)_{n\geqslant n_0}$ mit $n_0 \in \mathbb{Z}$).
- Die Konvergenz kann mit allen bisherigen Methoden untersucht werden.
  - Ist die unendliche Reihe z.B. monoton und beschränkt, so ist sie konvergent (Satz von Bolzano-Weierstraß).
  - Insbesondere ist eine Reihe genau dann absolut konvergent, wenn die Reihe der Absolutbeträge beschränkt ist.
  - Die unendliche Reihe ist genau dann konvergent, wenn das Cauchy-Kriterium gilt: Zu jedem $\epsilon > 0$ existiert ein $n_0 \in \mathbb{N}$ so, dass für alle $n > m \geqslant n_0$ die Ungleichung
  $$|s_n - s_m| = \left|\sum_{k=m+1}^{n} a_k\right| < \epsilon \qquad \text{gilt.}$$

Satz 11.18 (Konvergenz und absolute Konvergenz)

Eine absolut konvergente Reihe konvergiert.

Die Umkehrung gilt nicht.

11.19 (Harmonische und geometrische Reihe)

a) *Die Reihe*

$$\sum_{n=1}^{\infty} \frac{1}{n} = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \cdots$$

nennt man die HARMONISCHE REIHE. *Die harmonische Reihe divergiert.*

b) *Die Reihe*

$$\sum_{n=0}^{\infty} z^n = 1 + z + z^2 + z^3 + \cdots$$

mit $z \in \mathbb{C}$ *nennt man die* GEOMETRISCHE REIHE. *Für* $z \neq 1$ *lautet die Partialsumme (nach 1.12, geometrische Summenformel)*

$$s_n = \sum_{k=0}^{n} z^k = \frac{1 - z^{n+1}}{1 - z}.$$

Die geometrische Reihe konvergiert genau für $|z| < 1$, *der Grenzwert ist dann* $\dfrac{1}{1-z}$.

Der Grenzwertsatz 11.3 ergibt:

11.20 (Rechenregeln für konvergente Reihen)

Die unendlichen Reihen $\sum_{n=1}^{\infty} a_n$ *und* $\sum_{n=1}^{\infty} b_n$ *seien konvergent, ihre Grenzwerte seien* s_a *und* s_b. *Dann ist für beliebige* $\alpha, \beta \in \mathbb{C}$ *die Reihe* $\sum_{n=1}^{\infty} (\alpha a_n + \beta b_n)$ *konvergent und ihr Grenzwert ist* $\alpha s_a + \beta s_b$, *also*

$$\sum_{n=1}^{\infty} (\alpha a_n + \beta b_n) = \alpha \sum_{n=1}^{\infty} a_n + \beta \sum_{n=1}^{\infty} b_n.$$

Weitere Kriterien zur Konvergenz- bzw. Divergenz-Untersuchung von Reihen sind:

Satz 11.21 (Notwendiges Kriterium für die Konvergenz)

Falls die Reihe $\sum_{n=1}^{\infty} a_n$ *konvergiert, so muss die Folge* $(a_n)_{n \in \mathbb{N}}$ *der Reihenglieder eine Nullfolge sein.*

Beispiel: Die geometrische Reihe $\sum_{n=1}^{\infty} z^n$ mit $z \in \mathbb{C}$ und $|z| \geqslant 1$ ist divergent.

Das wichtigste Kriterium für Konvergenz bzw. Divergenz von Reihen ist:

Satz 11.22 (Majorantenkriterium)

Für zwei Reihen $\sum_{n=1}^{\infty} a_n$ *und* $\sum_{n=1}^{\infty} b_n$ *und ein festes* $n_0 \in \mathbb{N}$ *gelte*

$$|a_n| \leqslant b_n \qquad \textit{für alle} \quad n \geqslant n_0.$$

Falls dann $\sum_{n=1}^{\infty} b_n$ *konvergiert, so konvergieren auch* $\sum_{n=1}^{\infty} a_n$ *und* $\sum_{n=1}^{\infty} |a_n|$.

- Man beachte, dass die Reihe $\sum_{n=1}^{\infty} b_n$ absolut konvergieren muss, da stets $b_n \geqslant 0$ und daher $b_n = |b_n|$ ist.
- Die Reihe $\sum_{n=1}^{\infty} b_n$ heißt eine Majorante von $\sum_{n=1}^{\infty} a_n$.

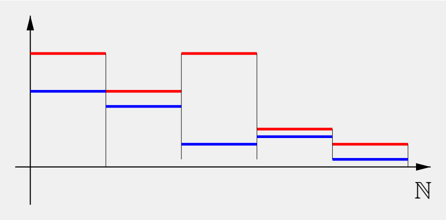
Satz 11.23 (Minorantenkriterium)

Für zwei Reihen $\sum_{n=1}^{\infty} a_n$ *und* $\sum_{n=1}^{\infty} b_n$ *und ein festes* $n_0 \in \mathbb{N}$ *gelte*

$$a_n \geqslant b_n \geqslant 0 \qquad \text{für alle} \quad n \geqslant n_0.$$

Falls dann $\sum_{n=1}^{\infty} b_n$ *divergiert, so divergiert auch* $\sum_{n=1}^{\infty} a_n$.

- Die Reihe $\sum_{n=1}^{\infty} b_n$ heißt eine Minorante von $\sum_{n=1}^{\infty} a_n$.



Eine Folge mit positiven Gliedern konvergiert genau dann, wenn die Partialsummen beschränkt sind.
Sind die Partialsummen bei der roten Majorante beschränkt, sind sie es bei der blauen Minorante auch, sind die Partialsummen der Minorante unbeschränkt, können die der Majorante nicht beschränkt sein.

Wichtige Konsequenz von Minoranten- und Majorantenkriterium ist das VERGLEICHSKRITERIUM.

Satz 11.24 (Vergleichskriterium)

Es gebe eine Zahl $c > 0$ *mit* $\lim_{n\to\infty} \frac{|a_n|}{|b_n|} = c$

Dann konvergiert $\sum_{n=1}^{\infty} a_n$ *genau dann absolut, wenn* $\sum_{n=1}^{\infty} b_n$ *absolut konvergiert.*

Beispiel 11.25

a) *Die Reihe* $\sum_{n=1}^{\infty} \frac{1}{n^2}$ *konvergiert.*

Der Grenzwert

$$s = \sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6} = 1.64493406684822...$$

kann erst viel später berechnet werden.

b) *Die Reihe* $\sum_{n=1}^{\infty} \frac{1}{\sqrt{n}}$ *divergiert.*

c) *Man merke sich schon jetzt als wichtige Reihen zum Vergleich: Die Reihe*

$$\sum_{n=1}^{\infty} \frac{1}{n^\alpha} \quad \textit{ist} \quad \begin{cases} \textit{konvergent, falls } \alpha > 1 \textit{ ist,} \\ \textit{bestimmt divergent gegen } \infty, \textit{ falls } 0 < \alpha \leqslant 1 \textit{ ist.} \end{cases}$$

Satz 11.26 (Wurzelkriterium)

- *Falls ein* $q \in \mathbb{R}$ *mit* $0 < q < 1$ *und ein* $n_0 \in \mathbb{N}$ *existieren, so dass*

$$\sqrt[n]{|a_n|} \leqslant q \qquad \text{für alle} \quad n \geqslant n_0, \tag{$*$}$$

  so konvergieren $\sum_{n=1}^{\infty} a_n$ *und* $\sum_{n=1}^{\infty} |a_n|$.

- *Falls*

$$\sqrt[n]{|a_n|} \geqslant 1 \qquad \text{für unendlich viele} \quad n \in \mathbb{N}$$

  gilt, so divergieren $\sum_{n=1}^{\infty} a_n$ *und* $\sum_{n=1}^{\infty} |a_n|$.

Bemerkung: Falls die Bedingung (*) nur mit $q = 1$ gilt (z.B. harmonische Reihe), so kann nicht auf Konvergenz geschlossen werden. Dann muss die Reihe mit anderen Kriterien untersucht werden.

Satz 11.27 (Quotientenkriterium)

- *Falls ein* $q \in \mathbb{R}$ *mit* $0 < q < 1$ *und ein* $n_0 \in \mathbb{N}$ *existieren, so dass*

$$a_n \neq 0 \qquad \text{und} \qquad \left|\frac{a_{n+1}}{a_n}\right| \leqslant q \qquad \text{für alle} \quad n \geqslant n_0 \qquad (**)$$

  gilt, so konvergieren $\sum_{n=1}^{\infty} a_n$ *und* $\sum_{n=1}^{\infty} |a_n|$.

- *Falls ein* $n_0 \in \mathbb{N}$ *existiert, so dass*

$$a_n \neq 0 \qquad \text{und} \qquad \left|\frac{a_{n+1}}{a_n}\right| \geqslant 1 \qquad \text{für alle} \quad n \geqslant n_0 \qquad (**)$$

  gilt, so divergieren $\sum_{n=1}^{\infty} a_n$ *und* $\sum_{n=1}^{\infty} |a_n|$.

Bemerkung:

- Falls die Bedingung (**) nur mit $q = 1$ gilt, muss die Reihe wieder mit anderen Kriterien untersucht werden.

- Das Quotientenkriterium ist oft leichter zu handhaben als das Wurzelkriterium. Aber: es verlangt $a_n \neq 0$ ab einem Index n_0.

Beispiel 11.28

a) *Die Reihe* $\displaystyle\sum_{n=0}^{\infty} \frac{1}{n!}$ *ist konvergent.*

Anmerkung: Der Grenzwert ist die Eulersche Zahl $e = 2.71828182845904\ldots$. *Die Folge der Partialsummen* $(s_n)_{n\geqslant 0}$ *konvergiert viel schneller gegen* e *als die Folge* $\big((1+\frac{1}{n})^n\big)_{n\in\mathbb{N}}$ *in 11.13.*

Zahlenwerte:

n	s_n		
0	1	6	2.71805555555556
1	2	7	2.71825396825397
2	2.5	8	2.71827876984127
3	2.66666666666667	9	2.71828152557319
4	2.70833333333333	10	2.71828180114638
5	2.71666666666667		

b) Für die Reihen $\sum_{n=1}^{\infty} \frac{1}{n}$ und $\sum_{n=1}^{\infty} \frac{1}{n^2}$ ergibt sich jeweils

$$\lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right| = 1.$$

Also kann mit dem Quotientenkriterium keine Aussage zur Konvergenz der Reihen getroffen werden. Ebenso erhalten wir

$$\lim_{n\to\infty} \sqrt[n]{|a_n|} = 1,$$

also hilft auch das Wurzelkriterium nicht weiter.

Oft prüft man die Existenz von $0 < q < 1$ in den Kriterien 11.26 und 11.27 so:

Satz 11.29 (Limesversion)

$\sum_{n=1}^{\infty} a_n$ *sei eine Reihe mit*

$$\limsup_{n\to\infty} \sqrt[n]{|a_n|} < 1 \qquad \textit{oder} \qquad \limsup_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right| < 1.$$

Dann konvergieren die Reihen $\sum_{n=1}^{\infty} a_n$ *und* $\sum_{n=1}^{\infty} |a_n|$.

Beispiel: für $k \in \mathbb{Z}$ und $0 \leqslant q < 1$ ist $\lim_{n\to\infty} \sqrt[n]{n^k q^n} = q < 1$.
Daher gilt nach 11.21 $n^k q^n \to 0$.

Wir nennen die reelle Zahlenreihe $\sum_{n=1}^{\infty} a_n$ ALTERNIEREND, wenn die Reihenglieder $a_n \in \mathbb{R}$ abwechselnd positiv und negativ sind, also

$$\operatorname{sign}(a_n) = -\operatorname{sign}(a_{n+1}) \qquad \text{für alle } \ n \in \mathbb{N} \ \text{ gilt.}$$

Satz 11.30 (Leibnizkriterium)

Die Reihe $\sum_{n=1}^{\infty} a_n$ *sei alternierend. Falls die Folge der Absolutbeträge* $(|a_n|)_{n\in\mathbb{N}}$ *eine monoton fallende Nullfolge ist, so konvergiert die Reihe* $\sum_{n=1}^{\infty} a_n$.

Zusatz: Der Grenzwert s erfüllt für jedes $n \in \mathbb{N}$ mit $\operatorname{sign}(a_n) = 1$

$$\sum_{k=1}^{n-1} a_k \leqslant s \leqslant \sum_{k=1}^{n} a_k .$$

Beispiel 11.31 (Leibniz'sche Reihe)

Die Reihe

$$\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \cdots$$

heißt die Leibniz'sche Reihe. *Sie ist alternierend und die Folge der Absolutbeträge* $(|a_n|)_{n\in\mathbb{N}} = \left(\frac{1}{n}\right)_{n\in\mathbb{N}}$ *ist eine monoton fallende Nullfolge. Also konvergiert die Leibniz'sche Reihe. Wir zeigen (viel) später:*

$$\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} = \ln 2 = 0.6914718055994....$$

Kapitel 12 – Grenzwerte von Funktionen, Stetigkeit

12 Grenzwerte von Funktionen, Stetigkeit

Bemerkung 12.1 (Festlegung des Definitionsbereichs)

Wir betrachten Funktionen $f : D \to \mathbb{R}$, *deren Definitionsbereich eine endliche Vereinigung von Intervallen ist, also z.B.*

$$D = [a, b], \quad D = [-5, 1) \cup (3, \infty), \quad D = \mathbb{R} \setminus \{0\}.$$

Der Abschluss $\overline{D}$ *ist die Menge, die durch Hinzunahme der Intervallränder entsteht, in den obigen Beispielen also*

$$\overline{D} = [a, b], \quad \overline{D} = [-5, 1] \cup [3, \infty), \quad \overline{D} = \mathbb{R}.$$

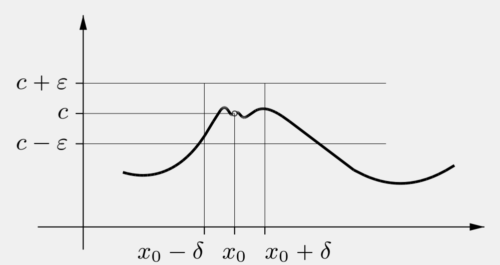
Definition 12.2 (Grenzwert)

Gegeben sei eine Funktion $f : D \to \mathbb{R}$ *und ein* $x_0 \in \overline{D}$.

(a) f *besitzt den Grenzwert* $c \in \mathbb{R}$ *an der Stelle* x_0*, wenn zu jedem* $\epsilon > 0$ *ein* $\delta > 0$ *existiert, so dass*

$$|f(x) - c| < \epsilon \quad \text{für alle} \quad x \in D \backslash \{x_0\} \text{ mit } |x - x_0| < \delta$$

gilt. Schreibweise: $\lim\limits_{x \to x_0} f(x) = c$.



Bemerkung: Der Funktionswert $f(x_0)$ taucht in der Bedingung an den Grenzwert von f an der Stelle x_0 nicht auf! Dieser wird erst später für den Begriff der Stetigkeit von f verwendet.

(b) f besitzt den uneigentlichen Grenzwert ∞ (bzw. $-\infty$) an der Stelle x_0, wenn zu jedem $r > 0$ ein $\delta > 0$ existiert, so dass

$$f(x) > r \quad \text{(bzw. } f(x) < -r\text{)} \qquad \text{für alle } \; x \in D\backslash\{x_0\} \; \text{ mit } \; |x - x_0| < \delta$$

gilt. Schreibweise: $\lim\limits_{x\to x_0} f(x) = \infty$ (bzw. $\lim\limits_{x\to x_0} f(x) = -\infty$).

Äquivalent zur "ϵ-δ-Definition" 12.2 des Grenzwerts ist die folgende:

Satz 12.3 (Definition mit Folgen)

Die Funktion $f : D \to \mathbb{R}$ *hat an der Stelle* $x_0 \in \overline{D}$ *den Grenzwert* c, *wenn für jede Folge* $(x_n)_{n\in\mathbb{N}}$ *mit Folgengliedern* $x_n \in D\backslash\{x_0\}$ *und mit dem Grenzwert* $\lim\limits_{n\to\infty} x_n = x_0$ *gilt*

$$\lim_{n\to\infty} f(x_n) = c.$$

Zwei Varianten von Grenzwerten werden in den nächsten beiden Abschnitten definiert:

Definition 12.4 (Grenzwert von f bei ∞ und $-\infty$)

Der Definitionsbereich D enthalte ein Intervall (a, ∞). Dann hat f in ∞ den Grenzwert $c \in \mathbb{R}$, wenn zu jedem $\epsilon > 0$ ein $r > a$ existiert, so dass

$$|f(x) - c| < \epsilon \quad \text{für alle} \quad x > r$$

gilt. Schreibweise: $\lim\limits_{x\to\infty} f(x) = c$.

Entsprechend: $\lim\limits_{x\to-\infty} f(x) = c$ *sowie uneigentliche Grenzwerte* $\lim\limits_{x\to\infty} f(x) = \infty$ *etc.*

Bemerkung: Endliche Grenzwerte c für $x \to \infty$ bzw. $x \to -\infty$ sind bei der Kurvendiskussion die *horizontalen Asymptoten* von f.

Definition 12.5 (rechtsseitiger und linksseitiger Grenzwert)

a) f *besitzt den* RECHTSSEITIGEN GRENZWERT $c \in \mathbb{R}$ *an der Stelle* x_0, *wenn zu jedem* $\epsilon > 0$ *ein* $\delta > 0$ *existiert, so dass*

$$|f(x) - c| < \epsilon \quad \text{für alle} \quad x \in D \text{ mit } x_0 < x < x_0 + \delta$$

gilt. Schreibweise: $\lim\limits_{x \to x_0+} f(x) = c$.

b) f *besitzt den* LINKSSEITIGEN GRENZWERT $c \in \mathbb{R}$ *an der Stelle* x_0, *wenn zu jedem* $\epsilon > 0$ *ein* $\delta > 0$ *existiert, so dass*

$$|f(x) - c| < \epsilon \quad \text{für alle} \quad x \in D \text{ mit } x_0 - \delta < x < x_0$$

gilt. Schreibweise: $\lim\limits_{x \to x_0-} f(x) = c$.

Ebenso: $\lim\limits_{x \to x_0+} f(x) = \infty$ *etc.*

Bemerkung:

- Am Randpunkt a eines Intervalls (a, b) oder $[a, b]$ ist der Grenzwert aus Definition 12.2 das gleiche wie der rechtsseitige Grenzwert.
- An einer Stelle $x_0 \in (a, b)$ (also im Inneren des Intervalls) existiert der Grenzwert c von f genau dann, wenn sowohl der rechtsseitige als auch der linksseitige Grenzwert existieren und gleich c sind, also

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

  gilt.

Die Rechenregeln für Grenzwerte von Folgen lassen sich auf Grenzwerte von Funktionen übertragen.

Regel 12.6 (Rechenregeln für Grenzwerte)

Gegeben seien zwei Funktionen $f : D \to \mathbb{R}$ und $g : D \to \mathbb{R}$ sowie ein $x_0 \in \overline{D}$. Weiter seien $\lim\limits_{x \to x_0} f(x) = c$ und $\lim\limits_{x \to x_0} g(x) = d$. Dann gilt

a) $\lim\limits_{x \to x_0} (\alpha f(x) + \beta g(x)) = \alpha c + \beta d$ *für alle* $\alpha, \beta \in \mathbb{R}$.

b) $\lim\limits_{x \to x_0} (f(x) g(x)) = cd$.

c) *Falls* $g(x) \neq 0$ *für alle* $x \in D$ *und* $d \neq 0$ *gilt, so ist* $\lim\limits_{x \to x_0} \dfrac{f(x)}{g(x)} = \dfrac{c}{d}$.

Definition 12.7 (Stetigkeit)

Gegeben sei eine Funktion $f : D \to \mathbb{R}$.

a) *f heißt* STETIG AN DER STELLE *$x_0 \in D$ (auch: "stetig in x_0"), wenn $\lim\limits_{x \to x_0} f(x) = f(x_0)$ gilt.*

b) *f heißt* STETIG, *wenn f in jedem Punkt $x_0 \in D$ stetig ist.*

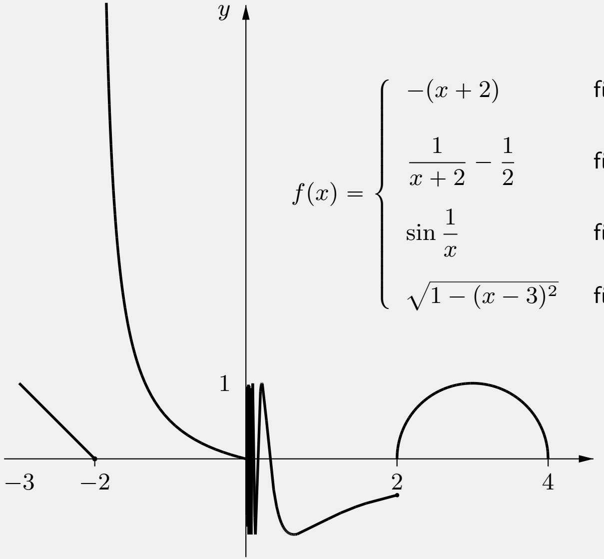
Beachte: Die Grenzwertdefinition 12.2 und 12.3 besagt:

- $f : D \to \mathbb{R}$ ist stetig an der Stelle $x_0 \in D$ genau dann, wenn

$$\forall \epsilon > 0 \; \exists \delta > 0 \; \forall x \in D : \quad (|x - x_0| < \delta \quad \Longrightarrow \quad |f(x) - f(x_0)| < \epsilon).$$

- $f : D \to \mathbb{R}$ ist stetig an der Stelle $x_0 \in D$ genau dann, wenn für jede Folge $(x_n)_{n \in \mathbb{N}}$ mit Folgengliedern $x_n \in D$ und dem Grenzwert $\lim\limits_{n \to \infty} x_n = x_0$ gilt:

$$\lim_{n \to \infty} f(x_n) = f(x_0).$$



$$f(x) = \begin{cases} -(x+2) & \text{für } -3 < x \leqslant -2 \\ \dfrac{1}{x+2} - \dfrac{1}{2} & \text{für } -2 < x \leqslant 0 \\ \sin \dfrac{1}{x} & \text{für } 0 < x \leqslant 2 \\ \sqrt{1-(x-3)^2} & \text{für } 2 < x < 4 \end{cases}$$

Verschiedene Typen von Unstetigkeit

Bemerkung: Die letzte Aussage wird häufig zur Grenzwertberechnung von Folgen verwendet: Falls die Folge $(a_n)_{n\in\mathbb{N}}$ gegen a konvergiert und f stetig an der Stelle a ist, so darf der Grenzwert "hineingezogen' werden:

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

Als direkte Folgerung aus 12.6 ergibt sich

Regel 12.8 (Rechenregeln für stetige Funktionen)

Die Funktionen $f, g : D \longrightarrow \mathbb{R}$ seien stetig in x_0. Dann gilt:

a) *Für beliebige Skalare $\alpha, \beta \in \mathbb{R}$ ist die Funktion $\alpha f + \beta g$ stetig in x_0.*

b) *Das Produkt fg ist stetig in x_0.*

c) *Falls $g(x) \neq 0$ für alle $x \in D$ gilt, so ist der Quotient $\frac{f}{g}$ stetig in x_0.*

Die Summe von Funktionen ist wie in Kapitel 7 definiert, Produkt und Quotient entsprechend.

Satz 12.9 (Vektorraum stetiger Funktionen)

Die Menge aller stetigen Funktionen $f : D \to \mathbb{R}$ *bildet einen reellen Vektorraum,*

$$C(D) := \{f : D \to \mathbb{R} \mid f \text{ ist stetig}\}.$$

Hintereinanderausführung von Funktionen:

- Gegeben sind zwei Funktionen $f : D_f \to \mathbb{R}$ und $g : D_g \to \mathbb{R}$. Der Bildbereich $f(D_f)$ sei eine Teilmenge von D_g. Dann ist die zusammengesetzte Funktion

$$g \circ f : D_f \to \mathbb{R}, \quad (g \circ f)(x) = g(f(x))$$

  definiert.

Satz 12.10 (Zusammengesetzte Funktionen)

Falls $f : D_f \to \mathbb{R}$ *stetig an der Stelle* x_0 *und* $g : D_g \to \mathbb{R}$ *stetig an der Stelle* $f(x_0)$ *ist, so ist die zusammengesetzte Funktion* $g \circ f : D_f \to \mathbb{R}$ *stetig an der Stelle* x_0.

Die folgende Aussage ist in vielen Situationen nützlich.

Lemma 12.11

Die Funktion $f : D \to \mathbb{R}$ sei stetig an der Stelle $x_0 \in D$ und es gelte $f(x_0) > c$ für irgendeine Zahl $c \in \mathbb{R}$. Dann existiert ein $\delta > 0$ so, dass

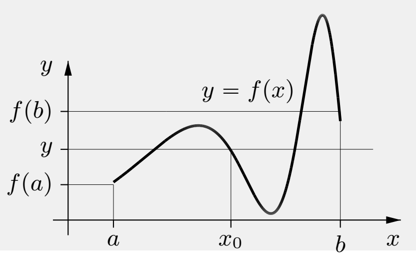
$$f(x) > c \quad \textit{für alle} \quad x \in D \textit{ mit } |x - x_0| < \delta$$

gilt. (Entsprechende Aussagen gelten für $f(x_0) < c$ bzw. $f(x_0) \neq c$.)

Bemerkung: Ist z.B. $f(x_0) \neq 0$, so gibt es sogar eine *offene Umgebung* $(x_0 - \delta, x_0 + \delta) \cap D$ von x_0, in der keine Nullstelle von f liegt.

Satz 12.12 (Zwischenwertsatz)

Sei $f : D \to \mathbb{R}$ stetig und $[a, b] \subseteq D$ ein abgeschlossenes beschränktes Intervall. Dann nimmt f auf dem Intervall $[a, b]$ jeden Wert $y \in \mathbb{R}$ zwischen $f(a)$ und $f(b)$ an.



Anwendung: Der Nullstellensatz von Bolzano

Die Funktion $f : [a, b] \to \mathbb{R}$ sei stetig und es gelte $f(a) \cdot f(b) < 0$ (d.h. $f(a)$ und $f(b)$ haben unterschiedliches Vorzeichen).

Dann existiert ein $x_0 \in (a, b)$ mit $f(x_0) = 0$.

Ein abgeschlossenes beschränktes Intervall $[a, b]$ nennt man *kompakt*.

Satz 12.13 (Satz vom Maximum)

$D = [a, b]$ sei ein kompaktes Intervall und die Funktion $f : D \to \mathbb{R}$ sei stetig. Dann nimmt f auf $[a, b]$ sein Maximum und sein Minimum an; d.h. der Bildbereich $f(D)$ ist das kompakte Intervall

$$f(D) = [c, d]$$

mit

$$c = \min_{x \in D} f(x), \qquad d = \max_{x \in D} f(x).$$

D sei ein Intervall und $f : D \to \mathbb{R}$ sei stetig. Der Bildbereich $J = f(D)$ ist wiederum ein Intervall (siehe 12.13). Wir setzen $\tilde{f} : D \to J$ mit $\tilde{f}(x) = f(x)$ für alle $x \in D$ und erhalten so eine surjektive Funktion.

Satz 12.14 (Satz zur Umkehrfunktion)

D sei ein Intervall und $f : D \to J$ mit $J = f(D)$ sei stetig. Dann gilt:

a) *Die Funktion f ist genau dann bijektiv, wenn sie streng monoton ist.*

b) *Ist f streng monoton wachsend (also auch bijektiv), so ist die Umkehrfunktion f^{-1} ebenfalls streng monoton wachsend und stetig.*

   Die entsprechende Aussage gilt für streng monoton fallendes f.

13 Differenzierbare Funktionen

13.1 Physikalisches Experiment Eine Person wirft zum Zeitpunkt $t = 0$ einen Ball senkrecht in die Höhe. Die Funktion $h : [0, T] \to \mathbb{R}$ mit

$$h(t) = h_0 + v_0 t - \frac{1}{2}\, g t^2, \qquad t \geqslant 0,$$

beschreibt näherungsweise die Höhe (in Metern) des Balls zur Zeit $t \geqslant 0$ (in Sekunden). Dabei ist h_0 die Anfangshöhe (ca. die Körpergröße), v_0 die Anfangsgeschwindigkeit und $g \approx 9.81\ m/s^2$ die Fallbeschleunigung. (Weitere Einflüsse wie der Luftwiderstand werden ignoriert.)

- Der Differenzenquotient
$$\frac{h(t_1) - h(t_0)}{t_1 - t_0}$$
gibt die Durchschnitts-Geschwindigkeit im Zeitintervall $[t_0, t_1]$ an. Dies ist die konstante Geschwindigkeit, die ein Ball haben müsste, um die Weglänge $h(t_1) - h(t_0)$ im gleichen Zeitintervall zurückzulegen:
$$S(t) = h(t_0) + \frac{h(t_1) - h(t_0)}{t_1 - t_0}\,(t - t_0). \qquad \text{(Sekanten-Gleichung)}$$

- Der Differential-Quotient
$$\lim_{t \to t_0} \frac{h(t) - h(t_0)}{t - t_0} =: h'(t_0) \qquad \text{(Ableitung)}$$
gibt die Momentan-Geschwindigkeit im Zeitpunkt t_0 an. Ein Ball mit dieser konstanten Geschwindigkeit würde zur Zeit t die Höhe $T(t)$ erreichen.
$$T(t) = h(t_0) + h'(t_0)(t - t_0) \qquad \text{(Tangenten-Gleichung)}$$

Festlegung wie in im Abschnitt über Stetigkeit: Der Definitionsbereich D ist die endliche Vereinigung von Intervallen.

Definition 13.1 (Ableitung)

Die Funktion $f : D \to \mathbb{R}$ *heißt* DIFFERENZIERBAR AN DER STELLE $x_0 \in D$, *wenn der Grenzwert*

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

existiert. Dieser Grenzwert heißt die ABLEITUNG $f'(x_0)$ *oder auch der* DIFFERENTIALQUOTIENT $\frac{df}{dx}(x_0)$ *von* f *an der Stelle* x_0.

f *heißt* DIFFERENZIERBAR, *wenn* f *an jeder Stelle von* D *differenzierbar ist. Dann heißt die Funktion*

$$f' : D \to \mathbb{R}, \quad x \mapsto f'(x)$$

die ABLEITUNG *von* f.

- EINSEITIGE ABLEITUNGEN: (siehe einseitige Grenzwerte 12.5)

$$f'(x_0+) = \lim_{x \to x_0+} \frac{f(x) - f(x_0)}{x - x_0}, \qquad \text{(rechtsseitige Ableitung)}$$

$$f'(x_0-) = \lim_{x \to x_0-} \frac{f(x) - f(x_0)}{x - x_0}, \qquad \text{(linksseitige Ableitung)}$$

  Im Fall $D = [a, b]$ ist $f'(a) = f'(a+)$ als die rechtsseitige und $f'(b) = f'(b-)$ als die linksseitige Ableitung definiert.

- Ist f in x_0 differenzierbar, so heißt die Gerade mit der Gleichung

$$y = T_1(x; x_0) = f(x_0) + f'(x_0)(x - x_0), \qquad x \in \mathbb{R},$$

  die TANGENTE an den Graphen von f in x_0.

- Es sei $(x_0 - \delta, x_0 + \delta) \subseteq D$ für ein $\delta > 0$. f ist genau dann an der Stelle x_0 differenzierbar, wenn beide einseitigen Ableitungen an dieser Stelle existieren und denselben Wert $f'(x_0+) = f'(x_0-) = f'(x_0)$ haben.

Satz 13.2 (Aus differenzierbar folgt stetig)

Ist $f : D \to \mathbb{R}$ differenzierbar an der Stelle $x_0 \in D$, so ist f dort auch stetig.

Definition 13.3 (Landau-Symbole)

Sind f und g zwei in einer Umgebung von $a \in \mathbb{R}$ definierte Funktionen, so bedeutet

- $f = O(g)$ *für* $x \to a$, *dass* $\dfrac{f(x)}{g(x)}$ *in einer Umgebung von* a *beschränkt ist*
- $f = o(g)$ *für* $x \to a$, *dass* $\lim\limits_{x \to a} \dfrac{f(x)}{g(x)} = 0$ *ist.*

Eine analoge Definition gilt für Folgen und für uneigentliche Grenzwerte $x \to \pm\infty$.
Bemerkung: $f = o(g)$ für $x \to a$ bedeutet, dass man $f(x) = g(x) \cdot q(x)$ mit

$$\lim_{x \to a} q(x) = 0$$

schreiben kann.

Definition 13.4 (Lineare Approximation)

Das Polynom $T_1(x;x_0) = f(x_0) + f'(x_0)(x - x_0)$ *vom Grad* 1 *heißt die lineare Approximation von* f *an der Stelle* x_0. *Der "Approximationsfehler"*

$$R_1(x;x_0) = f(x) - T_1(x;x_0)$$

erfüllt die Beziehungen

$$R_1(x_0;x_0) = 0, \qquad \lim_{x\to x_0} \frac{R_1(x;x_0)}{x - x_0} = 0 \Longleftrightarrow R_1(x;x_0) = o(x - x_0).$$

Satz 13.5 (Charakterisierung von Differenzierbarkeit)

$f : D \to \mathbb{R}$ *ist genau dann in* $x_0 \in D$ *differenzierbar mit der Ableitung* $f'(x_0)$*, wenn gilt*

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + o(x - x_0).$$

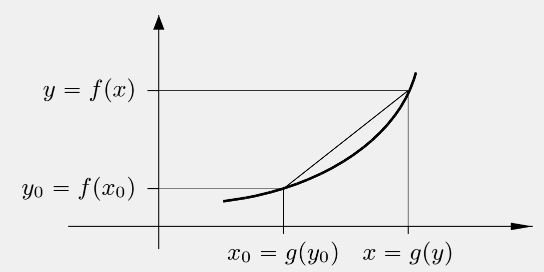
Satz 13.6 (Kettenregel und Ableitung der Umkehrfunktion)

(a) *I und J seien Intervalle, $f : I \to J$ sei differenzierbar an der Stelle x_0 und $g : J \to \mathbb{R}$ sei differenzierbar an der Stelle $y_0 = f(x_0)$. Dann ist die zusammengesetzte Funktion $g \circ f : I \to \mathbb{R}$ differenzierbar an der Stelle x_0 und es gilt*

$$(g \circ f)'(x_0) = g'(f(x_0))\ f'(x_0) = g'(y_0)\ f'(x_0). \qquad \text{(Kettenregel)}$$

(b) *Ableitung der Umkehrfunktion:*
I und J seien Intervalle, $f : I \to J$ sei bijektiv und $g := f^{-1} : J \to I$ sei die Umkehrfunktion.
Wenn f an der Stelle $x_0 \in I$ differenzierbar ist und $f'(x_0) \neq 0$ gilt, dann ist die Umkehrfunktion g an der Stelle $y_0 = f(x_0) \in J$ ebenfalls differenzierbar und es gilt

$$g'(y_0) = \frac{1}{f'(x_0)} = \frac{1}{f'(g(y_0))}.$$


$y = f(x)$
$y_0 = f(x_0)$
$x_0 = g(y_0)$
$x = g(y)$

Regel 13.7 (Rechenregeln für die Ableitung)

f und g seien differenzierbar an der Stelle x_0, α, β seien Skalare. Dann existieren die folgenden Ableitungen und es gilt

$$
\begin{aligned}
(\alpha f + \beta g)'(x_0) &= \alpha f'(x_0) + \beta g'(x_0) \\
(fg)'(x_0) &= f'(x_0)g(x_0) + f(x_0)g'(x_0) && \textit{(Produktregel)} \\
\left(\frac{1}{g}\right)'(x_0) &= -\frac{g'(x_0)}{(g(x_0))^2} \\
\left(\frac{f}{g}\right)'(x_0) &= \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{(g(x_0))^2} && \textit{(Quotientenregel)}
\end{aligned}
$$

Bei den letzten beiden Regeln muss $g(x_0) \neq 0$ vorausgesetzt werden.

Beispiel 13.8

(a) *Jedes Polynom ist differenzierbar.*

(b) $f : (0, \infty) \to (0, \infty)$, $f(x) = \sqrt{x}$ *über den Differentialquotienten*

(c) *Produktregel:* $f(x) = x^2 \sin x$

(d) *Quotientenregel:* $f(x) = \frac{x^2-1}{x^2+1}$

(e) *Quotientenregel:* $\tan x$ *und* $\cot x$

(f) *Kettenregel:* $\sin(x^2 + 1)$, $\cos^5(\sqrt{x})$

(g) *Umkehrfunktionen:* $f(x) = x^2$ *auf* $(0, \infty)$, $g(x) = \sin x$ *auf* $(-\pi/2, \pi/2)$, $h(x) = \cos x$ *auf* $(0, \pi)$.

In Satz 12.13 haben wir das ABSOLUTE (oder globale) Maximum bzw. Minimum einer stetigen Funktion $f : I \to \mathbb{R}$ behandelt.

Definition 13.9 (relative (oder lokale) Extrema)

Sei I ein Intervall und $f : I \to \mathbb{R}$ eine Funktion.
f hat an der Stelle x_0 ein RELATIVES MAXIMUM *(bzw. ein* RELATIVES MINIMUM*) $f(x_0)$, wenn es ein Intervall $(x_0 - \delta, x_0 + \delta) \subseteq I$ gibt, so dass*

$$f(x) \leqslant f(x_0) \quad \text{(bzw. } f(x) \geqslant f(x_0) \text{)} \qquad \text{für alle } x \in (x_0 - \delta, x_0 + \delta).$$

Ein relatives Maximum oder Minimum heißt auch ein RELATIVER EXTREMWERT *von f.*

Beachte: An einem Endpunkt des Intervalls $I = [a, b]$ kann laut der Definition kein RELATIVER Extremwert vorliegen. Hier kann jedoch das ABSOLUTE Maximum oder Minimum von f angenommen werden (Beispiele: monotone Funktionen)

Hauptsatz zur Kurvendiskussion:

Satz 13.10 (Notwendiges Kriterium)

Sei I ein Intervall und $f : I \to \mathbb{R}$ an der Stelle $x_0 \in I$ differenzierbar.
Wenn f an der Stelle x_0 einen relativen Extremwert hat, so gilt $f'(x_0) = 0$.

Beispiel 13.11

(a) $f : [-1,1] \to \mathbb{R}$, $f(x) = |x|$ (*"es geht auch ohne Differenzierbarkeit"*)

(b) $f : [-2,2] \to \mathbb{R}$, $f(x) = x^3 - x$ *mit relativem Maximum bei* $x_1 = -\frac{\sqrt{3}}{3}$ *und relativem Minimum bei* $x_2 = \frac{\sqrt{3}}{3}$ *und absolutem Maximum bei* $x = 2$ *sowie absolutem Minimum bei* $x = -2$.

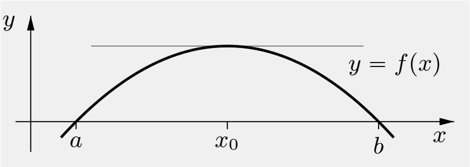
(c) $f : [-1,1] \to \mathbb{R}$, $f(x) = x^3$. *Der einzige Kandidat für einen relativen Extremwert ist* $x_0 = 0$. *Wegen*

$$f(x_1) < f(0) < f(x_2) \qquad \textit{für alle} \quad x_1 < 0 < x_2$$

hat f *jedoch keinen relativen Extremwert an der Stelle* $x_0 = 0$.

Satz 13.12 (Satz von Rolle)

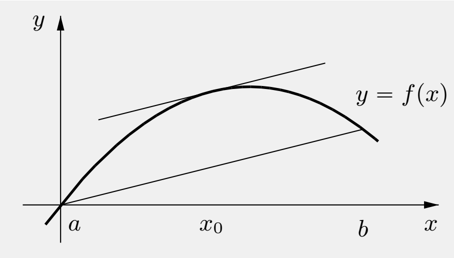
Es sei $I = [a, b]$ ein Intervall, $a < b$. Die Funktion $f : [a, b] \to \mathbb{R}$ sei stetig in I und differenzierbar in (a, b).
Falls $f(a) = f(b)$ gilt, so gibt es ein $x_0 \in (a, b)$ mit $f'(x_0) = 0$.



Satz 13.13 (Mittelwertsatz der Differentialrechnung)

Es sei $I = [a,b]$ ein Intervall, $a < b$. Die Funktion $f : [a,b] \to \mathbb{R}$ sei stetig in I und differenzierbar in (a,b).

Dann gibt es ein $x_0 \in (a,b)$ mit $f'(x_0) = \dfrac{f(b) - f(a)}{b - a}$.



Geometrisch: Es existiert eine Stelle $x_0 \in (a,b)$, an der die Steigung $f'(x_0)$ der Tangente gleich der Steigung der Sekante durch $(a, f(a))$ und $(b, f(b))$ ist.

Direkte Folgerungen aus dem Mittelwertsatz:

Satz 13.14 (Satz von der konstanten Funktion, Monotonie)

Es sei $I = [a, b]$ ein Intervall, $a < b$. Die Funktion $f : [a, b] \to \mathbb{R}$ sei stetig in I und differenzierbar in (a, b).

(a) *Wenn $f'(x) = 0$ für alle $x \in (a, b)$ gilt, so ist f konstant.*

(b) *Wenn $f'(x) > 0$ (bzw. $f'(x) < 0$) für alle $x \in (a, b)$ gilt, so ist f streng monoton wachsend (bzw. streng monoton fallend).*

Wir betrachten nun Funktionen $f : D \to \mathbb{R}$, die (auf ganz D) differenzierbar sind. In 13.1 wurde dann die Ableitung, also die Funktion

$$f' : D \to \mathbb{R}, \quad x \mapsto f'(x)$$

definiert.

Definition 13.15 (Höhere Ableitungen)

Die Funktion $f : D \to \mathbb{R}$ sei differenzierbar.

- *Falls f' differenzierbar ist, nennen wir $f'' := (f')'$ die 2. Ableitung von f, und f heißt zweimal differenzierbar.*
  (Andere Schreibweisen: $f^{(2)}$ oder $\frac{d^2 f}{dx^2}$)
- *Analog erhält man die höheren Ableitungen $f''' = f^{(3)}$, $f^{(4)}$, ..., und allgemein für ein $n \in \mathbb{N}$ die n-te Ableitung*

  $$f^{(n)} = \frac{d^n f}{dx^n},$$

  falls $f^{(n-1)}$ weiterhin differenzierbar ist. Die Funktion f heißt dann n-mal differenzierbar.

Zusatz:

- Falls die Ableitung f' stetig ist, heißt f STETIG DIFFERENZIERBAR.
- Falls f n-mal differenzierbar ist und die n-te Ableitung $f^{(n)}$ stetig ist, so heißt f n-MAL STETIG DIFFERENZIERBAR. (Siehe hierzu Satz 13.2)
- Die Menge aller n-mal stetig differenzierbaren Funktionen $f : D \to \mathbb{R}$ ist ein Vektorraum, den wir mit $C^n(D)$ bezeichnen. Zur Vollständigkeit setzen wir $C^0(D) := C(D)$.

n-mal differenzierbare Funktionen erlauben eine bessere Approximation als durch lineare Polynome in 13.4:

Definition 13.16 (Taylor-Polynom)

Es sei $n \in \mathbb{N}$, $I = (a, b)$ ein Intervall, $a < x_0 < b$ und $f \in C^n(I)$. Dann heißt das Polynom

$$T_n(x; x_0) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k$$

das n-te Taylor-Polynom *von f im Entwicklungspunkt x_0.*

T_n ist durch $T_n^{(k)}(x_0; x_0) = f^{(k)}(x_0)$ für $k = 0, \ldots, n$ charakterisiert.

Beispiele:

$$\begin{aligned}
T_0(x; x_0) &= f(x_0), \\
T_1(x; x_0) &= f(x_0) + f'(x_0)(x - x_0), \\
T_2(x; x_0) &= f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2} (x - x_0)^2, \\
T_3(x; x_0) &= f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2} (x - x_0)^2 + \frac{f'''(x_0)}{6} (x - x_0)^3, etc.
\end{aligned}$$

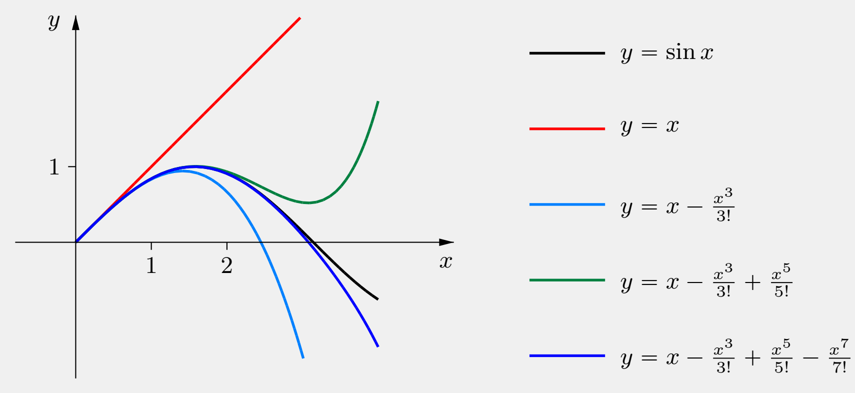
Die Abweichung $f(x) - T_n(x; x_0)$ lässt sich folgendermaßen ausdrücken:

Satz 13.17 (Satz: Taylorsche Formel)

Es sei $n \in \mathbb{N}$, $I = (a, b)$ *ein Intervall,* $a < x_0 < b$ *und* $f \in C^{n+1}(I)$.
Dann gilt für alle $x \in I$

$$f(x) - T_n(x; x_0) = \frac{f^{(n+1)}(\xi)}{(n+1)!} \, (x - x_0)^{n+1}$$

mit einer Stelle $\xi = x_0 + \alpha(x - x_0)$ *und* $0 < \alpha < 1$. *(Restglied von Lagrange)*



Geogebra Taylor

Wir benötigen noch:

Satz 13.18 (Verallgemeinerter Mittelwertsatz)

Es sei $I = [a, b]$ ein Intervall, $a < b$, und die Funktionen $f, g : I \to \mathbb{R}$ seien stetig und in (a, b) differenzierbar; weiter sei $g'(x) \neq 0$ für alle $x \in (a, b)$.
Dann gibt es ein $\xi \in (a, b)$ mit

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(\xi)}{g'(\xi)}.$$