Vorlesungsskript
Mathematik für die Geowissenschaften I

Wintersemester 2023
CAU Kiel

Daniel Bissinger

Version: 8. Februar 2024
Kommentare sind herzlich willkommen.

Inhaltsverzeichnis

1 Grundlagen **1**
- 1.1 Mengen . . . 1
- 1.2 Funktionen . . . 4
- 1.3 Elementare Funktionen und ihre Eigenschaften . . . 8
  - 1.3.1 Potenzen, Wurzeln und Polynomfunktionen . . . 8
  - 1.3.2 Sinus, Cosinus und Tangens . . . 9

2 Grundbegriffe der Analysis **12**
- 2.1 Folgen und Grenzwerte . . . 12
- 2.2 Reihen . . . 15
- 2.3 Grenzwerte von Funktionen und stetige Funktionen . . . 18
- 2.4 Differenzierbarkeit . . . 23
- 2.5 Differenzierbarkeit in mehreren Variablen . . . 29
- 2.6 Integration . . . 33
  - 2.6.1 Integrale und Flächeninhalte . . . 33
  - 2.6.2 Integrale als Mittelwerte . . . 36
  - 2.6.3 Länge von Funktionsgraphen . . . 37
  - 2.6.4 Volumina . . . 39

3 Grundbegriffe der Linearen Algebra **40**
- 3.1 Vektoren . . . 42
- 3.2 Matrizen . . . 44
- 3.3 Die Zeilenstufenform . . . 46
- 3.4 Lineare Abbildungen und Matrizen . . . 51

Kapitel 1

Grundlagen

1.1 Mengen

Definition 1.1. Eine *Menge* M ist eine Zusammenfassung von wohlunterschiedenen Objekten, welche die *Elemente* von M genannt werden.

Definition 1.2. Es seien M und N Mengen.

(1) Wir schreiben $m \in M$, wenn m ein Element von M ist. Ist m kein Element von M, so schreiben wir $m \notin M$.

(2) Wir schreiben $M \subseteq N$ und sagen M ist eine *Teilmenge* von N, falls jedes Element von M schon in N enthalten ist. Gilt dies nicht, so schreiben wir $M \nsubseteq N$.

(3) Zwei Mengen heißen *gleich* (schreibe $M = N$), wenn $M \subseteq N$ und $N \subseteq M$ gilt. Gilt dies nicht, so schreiben wir $M \neq N$.

(4) Gilt $M \subseteq N$ und $M \neq N$, so schreiben wir kurz $M \subsetneq N$, wenn wir betonen wollen, dass keine Mengengleichheit gilt.

Wir stellen Mengen durch geschweifte Klammern $\{\ldots\}$ dar und haben dabei verschiedene Möglichkeiten die Mengen zu beschreiben:

- Auflistung sämtlicher Elemente: $\{2, 4, \frac{1}{2}, \pi, 10\}$.
- Beschreibung durch Eigenschaften: $\{\alpha \mid \alpha$ Buchstabe im deutschen Alphabet[1]$\}$.

Bemerkung 1.3. (1) In einer Menge spielt die Reihenfolge der Elemente keine Rolle, das heißt $\{2, 4, 5\} = \{2, 5, 4\} = \{5, 4, 2\} = \ldots$.

(2) Stellen wir eine Menge durch Auflistung sämtlicher Elemente dar, können wir theoretisch auch Elemente mehrfach nennen ohne die Menge zu verändern, etwa $\{1, 2, 3\} = \{1, 2, 2, 3, 3, 3\}$. Wir treffen aber die Vereinbarung stets auf Mehrfachnennung zu verzichten.
Das heißt $\{1, 2, 2, 3, 3, 3\}$ „reduzieren“ wir zu $\{1, 2, 3\}$.

[1] Wir meinen hier Kleinbuchstaben

(3) Definieren wir ein mathematisches Objekt, so schreiben wir $:=$, etwa

$$\mathcal{A} := \{\alpha \mid \alpha \text{ Buchstabe im Alphabet}\}.$$

Dann gilt $\mathcal{A} = \{a, b, c, \dots, z\}$. Weiter definieren wir

$$X := \{\alpha \in \mathcal{A} \mid \alpha \text{ taucht im Wort „voll“ auf}\}.$$

Dann gilt $X = \{v, o, l, l\} = \{v, o, l\}$.

Beispiel 1.4. (1) Die Menge, die kein Element enthält, notieren wir mit $\varnothing := \{\ \}$ und nennen sie *leere Menge*. Wir bemerken, dass $\varnothing \subseteq A$ für jede beliebige Menge A gilt. Weiter gilt $A \subseteq \varnothing$ nur für $A = \varnothing$.

(2) Es gilt $\{1, 2\} \nsubseteq \{2, 3\}$ und $\{2, 3\} \nsubseteq \{1, 2\}$.

(3) Wir betrachten erneut $\mathcal{A} := \{\alpha \mid \alpha \text{ Buchstabe im Alphabet}\} = \{a, b, c, \dots, z\}$. Weiter sei $\mathcal{V} := \{\alpha \mid \alpha \text{ Vokal im Alphabet}\} = \{a, e, i, o, u\}$. Dann gilt $\mathcal{V} \subseteq \mathcal{A}$ und $\mathcal{V} \subsetneq \mathcal{A}$, sowie $\mathcal{A} \nsubseteq \mathcal{V}$.

In der Vorlesung werden wir stets folgende Abkürzungen für uns wohlbekannte Mengen benutzen:

- Die Menge der *natürlichen Zahlen*: $\mathbb{N} = \{1, 2, 3, 4, \dots\}$, $\mathbb{N}_0 = \{0, 1, 2, \dots\}$.
- Die Menge der *ganzen Zahlen* $\mathbb{Z} = \{\cdots - 5, -4, -3, -2, -1, 0, 1, 2, \dots\}$.
- Die Menge der *rationalen Zahlen* $\mathbb{Q} = \{\frac{p}{q} \mid p, q \in \mathbb{Z},\ q \neq 0\}$.
- Die Menge der *reellen Zahlen* $\mathbb{R}$. Hier fügt man zu $\mathbb{Q}$ weitere Zahlen hinzu, wie etwa $\sqrt{2}$, die Kreiszahl π und die Eulersche Zahl e. Wir definieren weiter $\mathbb{R}_{\geq 0} := \{x \in \mathbb{R} \mid x \geq 0\}$ und $\mathbb{R}_{>0} := \{x \in \mathbb{R} \mid x > 0\}$.
- Sind $a \leq b \in \mathbb{R}$ reelle Zahlen, so definieren wir

$$\begin{aligned}(a, b) &:= \{x \in \mathbb{R} \mid a < x < b\}\\ [a, b] &:= \{x \in \mathbb{R} \mid a \leq x \leq b\}\\ (a, b] &:= \{x \in \mathbb{R} \mid a < x \leq b\}\\ [a, b) &:= \{x \in \mathbb{R} \mid a \leq x < b\}.\end{aligned}$$

Diese Mengen heißen *(beschränkte) Intervalle*. Genauer: (a, b) ist das *offene Intervall* von a und b und $[a, b]$ ist das *abgeschlossene Intervall* von a und b. Die anderen beiden Intervallarten heißen *halboffen*. Weiter definieren wir *unendliche Intervalle*

$$\begin{aligned}[a, \infty) &:= \{x \in \mathbb{R} \mid a \leq x\}\\ (a, \infty) &:= \{x \in \mathbb{R} \mid a < x\}\\ (-\infty, a] &:= \{x \in \mathbb{R} \mid x \leq a\}\\ (-\infty, a) &:= \{x \in \mathbb{R} \mid x < a\}\\ (-\infty, \infty) &:= \mathbb{R}.\end{aligned}$$

Bemerkung 1.5. (1) Für $a = b$ gilt $(a, b) = (a, b] = [a, b) = \varnothing$ und $[a, b] = \{a\} = \{b\}$.

(2) Es gilt $\mathbb{N} \subsetneq \mathbb{N}_0 \subsetneq \mathbb{Z} \subsetneq \mathbb{Q} \subsetneq \mathbb{R}$.

(3) Gilt $a \in \mathbb{N}$, so können wir die Gleichung $a+x = 0$ in $\mathbb{N}$ nicht lösen. Dies ist die Motivation für die ganzen Zahlen. In $\mathbb{Z}$ ist die Gleichung $2x = 1$ nicht lösbar und motiviert die Konstruktion von $\mathbb{Q}$.

Definition 1.6. Es seien M und N Mengen.

(1) Die Menge $M \cup N := \{x \mid x \in M \text{ oder } x \in N\}$ heißt die *Vereinigung* von M und N.

(2) Die Menge $M \cap N := \{x \mid x \in M \text{ und } x \in N\}$ heißt der *Schnitt* von M und N.

(3) Die Menge $M \setminus N := \{x \mid x \in M \text{ und } x \notin N\}$ heißt die *Differenz* von M und N. Wir sagen auch M *ohne* N.

Bemerkung 1.7. Schnitt und Vereinigung sind symmetrische Operationen. Das heißt es gilt stets $M \cup N = N \cup M$ und $M \cap N = N \cap M$. Es gilt jedoch nur $M \setminus N = N \setminus M$, wenn $M = N$ gilt.

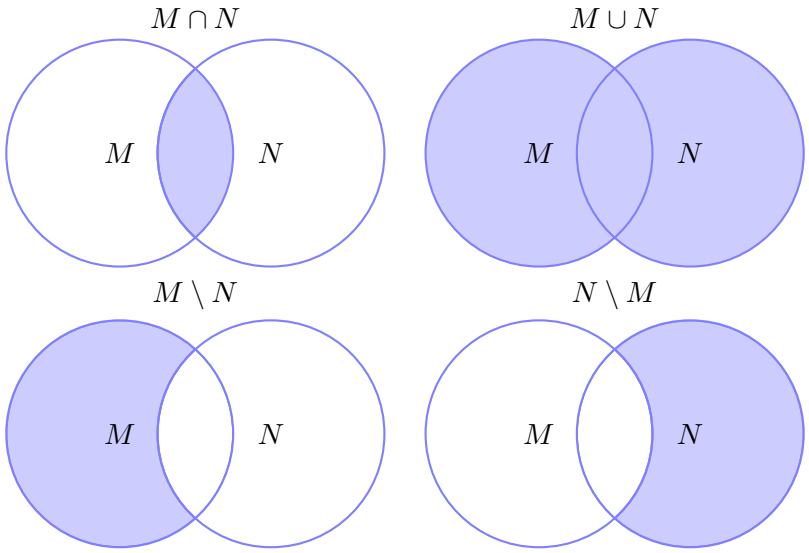


Abbildung 1.1: Veranschaulichung von Mengenoperationen durch Venn-Diagramme.

Beispiel 1.8. (1) $\mathbb{R} \setminus \mathbb{Q}$ heißt die Menge der *irrationalen Zahlen*.

(2) Es sei $A := \{z \in \mathbb{N} \mid z \text{ ungerade}, z \leq 17\}$, $B := \{x^2 \mid x \in \{1, 2, 3, 4, 6\}\}$. Dann gilt

- $A \cup B = \{1, 3, 4, 5, 7, 9, 11, 13, 15, 16, 17, 36\}$,
- $A \cap B = \{1, 9\}$,
- $A \setminus B = \{3, 5, 7, 11, 13, 15, 17\}$ und
- $B \setminus A = \{4, 16, 36\}$.

(3) Wir setzen Beispiel 1.4 fort. Dann gilt

- $\mathcal{A} \cap \mathcal{V} = \mathcal{V}$,
- $\mathcal{A} \cup \mathcal{V} = \mathcal{A}$,
- $\mathcal{A} \setminus \mathcal{V} = \{\alpha \mid \alpha \text{ Konsonant im Alphabet}\}$ und
- $\mathcal{V} \setminus \mathcal{A} = \varnothing$.

1.2 Funktionen

Definition 1.9. Es seien A und B nicht-leere Mengen. Eine *Funktion* oder *Abbildung* $f\colon A \to B$ ist eine Vorschrift, die jedem Element $a \in A$ genau ein $f(a) \in B$ zuordnet. Für $a \in A$ heißt $f(a) \in B$ der *Funktionswert von a unter f*.

Die Funktionsvorschrift kann in unterschiedlicher Weise angegeben werden:

(i) Auflistung: $A = \{1,2,3\}, B = \mathbb{R}$, $f(1) = 1, f(2) = 4, f(3) = \pi$ bzw. $1 \mapsto 2, 2 \mapsto 4, 3 \mapsto \pi$.

(ii) Funktionsgleichung: $A = B = \mathbb{R}$. $f\colon \mathbb{R} \to \mathbb{R}; f(x) = x^2$ bzw. $f\colon \mathbb{R} \to \mathbb{R}; x \mapsto x^2$.

Bemerkung 1.10. (1) Eine Funktion ist durch drei Daten gegeben:

(i) Die Menge A, der sogenannte *Definitionsbereich*,

(ii) die Menge B, der sogenannte *Zielbereich*, und

(iii) die *Funktionsvorschrift*, die die Zuordnung beschreibt.

(2) Zwei Funktionen $f\colon A \to B$, $g\colon A' \to B'$ sind nach Definition genau dann gleich, wenn $A = A'$ und $B = B'$ gelten und für alle $a \in A$ die Gleichung $f(a) = g(a)$ erfüllt ist.

Beispiel 1.11. (1) $f_1\colon \mathbb{R} \to \mathbb{R}; x \mapsto x^2$, $f_2\colon \mathbb{R} \to \mathbb{R}_{\geq 0}; x \mapsto x^2$ und $f_3\colon \mathbb{R}_{\geq 0} \to \mathbb{R}; x \mapsto x^2$ sind Funktionen, die alle paarweise verschieden sind.

(2) $g_1\colon \mathbb{Z} \to \mathbb{N}_0; x \mapsto x^2 - 2$, $g_2\colon \{1\} \to \{2,3\}; 1 \mapsto 2, 1 \mapsto 3$, $g_3\colon \mathbb{Q} \to \mathbb{Z}; \frac{a}{b} \mapsto a$ und $g_4\colon \{1,0\} \to \mathbb{Z}; x \mapsto \frac{1}{x}$ sind allesamt keine Funktionen.

(3) Die Funktion

$$|\cdot|\colon \mathbb{R} \to \mathbb{R}_{\geq 0}; x \mapsto |x| := \begin{cases} x & \text{, falls } x \geq 0 \\ -x & \text{, falls } x < 0 \end{cases}$$

heißt *Betragsfunktion*. Wir bemerken, dass $|a-b|$ den Abstand von a und b auf der reellen Zahlengerade angibt. Weiter gilt die sogenannte *Dreiecksungleichung*: Für alle $x, y \in \mathbb{R}$ gilt $|x+y| \leq |x| + |y|$. Wir beweisen diese Aussage durch eine Fallunterscheidung: Gilt $x + y \geq 0$, so ist $|x+y| = x + y \leq |x| + |y|$. Im anderen Fall gilt $x + y < 0$ und damit $|x+y| = -(x+y) = -x + (-y) \leq |x| + |y|$.

Funktionen erlauben uns Mengen miteinander in Beziehung zu setzen in dem wir jedem Element der Definitionsmenge genau ein Element in der Zielmenge zuordnen. Bevor wir weitere Eigenschaften von Funktionen einführen, betrachten wir zwei motivierende Beispiele.

Beispiel 1.12. (1) Wir betrachten einen Kinosaal des STUDIO Filmtheaters in Kiel. Die Menge der Kinositze im Saal bezeichnen wir mit S. Die Menge der Gäste sei mit G bezeichnet. Jedem Gast $g \in G$ wird bei Besuch genau ein Sitz $f(g) \in S$ zugeordnet. Nachdem alle Gäste vor Ort sind, haben wir also eine Funktion

$$f\colon G \to S; g \mapsto f(g).$$

Wir betrachten folgende Zusatzeigenschaften:

- Für die Besucher ist wünschenswert: Auf jedem Platz sitzt höchstens ein Gast bzw. sitzen Gäste g und g' auf demselben Platz, so gilt schon $g = g'$. Mathematisch: Für alle $g, g' \in G$ folgt aus $f(g) = f(g')$ bereits $g = g'$.
- Für die Betreiber des Kinos ist wünschenswert: Jeder Platz im Saal ist besetzt. Das heißt, auf jedem Platz sitzt (mindestens) ein Gast. Mathematisch: Für alle $s \in S$ existiert ein $g \in G$ mit $f(g) = s$.[a]

(2) Uns werden zwei Eimer A und B gegeben. Eimer A enhält Äpfel und Eimer B enthält Birnen. Wir sollen herausfinden, ob beide Eimer gleichviele Stücke Obst enthalten. Natürlich können wir beide Eimer durchzählen und danach die Anzahl vergleichen. Dabei müssen wir natürlich sehr aufpassen und sicherstellen, dass wir uns nicht verzählen. Da wir nicht an der Anzahl interessiert sind, können wir auch wie folgt vorgehen. Wir entnehmen zeitgleich einen Apfel und eine Birne und legen sie als Paar zur Seite. Dann wiederholen wir diesen Schritt bis einer der Eimer leer ist. Die Eimer enthalten also genau dann gleichviele Stücke Obst, wenn die Eimer zeitgleich leer werden. Im Kontext von Beispiel (1): Es gibt genauso viele Gäste wie Kinositze, wenn auf jedem Platz genau ein Gast sitzt.

[a] Rein mathematisch ist es natürlich erlaubt zwei verschiedene Gäste auf denselben Platz zu setzen. In der Realität würde es hier zu Beschwerden kommen.

Die vorherigen Beispiele motivieren die folgenden Definitionen.

Definition 1.13. Es sei $f\colon A \to B$ eine Funktion. Die Funktion f heißt

(1) *injektiv*, falls für alle $a, a' \in A$ aus $f(a) = f(a')$ bereits $a = a'$ folgt.

(2) *surjektiv*, falls für alle $b \in B$ ein $a \in A$ mit $f(a) = b$ existiert.

(3) *bijektiv*, falls f injektiv und surjektiv ist.

Bemerkung 1.14. Eine Funktion $f\colon A \to B$ ist injektiv (surjektiv/bijektiv), falls für jedes $b \in B$ höchstens (mindestens/genau) ein $a \in A$ mit $f(a) = b$ existiert. Bijektivität in den Worten aus Beispiel 1.12: Wir können jeder Birne aus B genau einen Apfel aus A zuordnen.

Definition 1.15. Zwei Mengen M, N heißen *gleichmächtig*, falls eine bijektive Funktion $f\colon M \to N$ existiert.

Beispiel 1.16. (1) Die Vorschrift $f\colon \{1,2,3\} \to \{4,9,11\}; 1 \mapsto 4, 2 \mapsto 9, 3 \mapsto 9$ ist eine Funktion. Diese ist weder injektiv noch surjektiv.

(2) $f\colon \{1\} \to \{1,2\}; 1 \mapsto 1$ ist eine injektive Funktion, die nicht surjektiv ist, und $g\colon \{1,2\} \mapsto \{1\}; 1 \mapsto 1, 2 \mapsto 1$ ist eine surjektive Funktion, die nicht injektiv ist.

(3) $f\colon \mathbb{R} \to \mathbb{R}; x \mapsto x - 10$ ist eine bijektive Abbildung: Gilt $f(x) = f(y)$, so erhalten wir $x - 10 = y - 10$ und damit $x = (x - 10) + 10 = (y - 10) + 10 = y$. Ist hingegen $b \in \mathbb{R}$ gegeben, so gilt $b + 10 \in \mathbb{R}$ mit $f(b + 10) = b + 10 - 10 = b$.

(4) Ist $f\colon A \to B$ eine Abbildung, so ist $f\colon A \to f(A); a \mapsto f(a)$ eine surjektive Funktion, wobei $f(A) := \{f(a) \mid a \in A\}$ das *Bild von f* bezeichnet.

Beispiel 1.17 (Warnung)**.** Spätestens wenn wir uns mit unendlichen Mengen beschäftigen, treten nicht intuitive Phänome auf: Im Buch „The Fault in Our Stars“ von John Green heißt es „There are infinite numbers between 0 and 1. There's .1 and .12 and .112 and an infinite collection of others. Of course, there is a bigger infinite set of numbers between 0 and 2, or between 0 and a million. Some infinities are bigger than other infinities.“ Es lässt sich in der Tat zeigen, dass es unterschiedliche Unendlichkeiten gibt, zum Beispiel kann man beweisen, dass keine Bijektion zwischen $\mathbb{N}$ und $\mathbb{R}$ existiert. Allerdings sind die obigen Beispiele nicht geeignet. Wir betrachten die Abbildung

$$f\colon [0,1] \to [0,2], x \mapsto 2 \cdot x.$$

Die Funktion f ist bijektiv. Damit sind die Mengen $[0,1]$ und $[0,2]$ gleichmächtig, da wir jedem Element aus $[0,2]$ genau ein Element aus $[0,1]$ zuordnen.
Hinweis: Es existiert sogar eine Bijektion $\mathbb{R} \to [0,1]$.

Definition 1.18. (1) Es seien $n \in \mathbb{N}$ und $A_1, \ldots, A_n$ Mengen. Dann heißt

$$A_1 \times A_2 \times \cdots \times A_n := \{(a_1, \ldots, a_n) \mid a_i \in A_i \text{ für alle } i \in \{1, \ldots, n\}\}$$

das *kartesische Produkt* der Mengen $A_1, \ldots, A_n$.

(2) Es sei $f\colon A \to B$ eine Funktion. Dann heißt die Menge

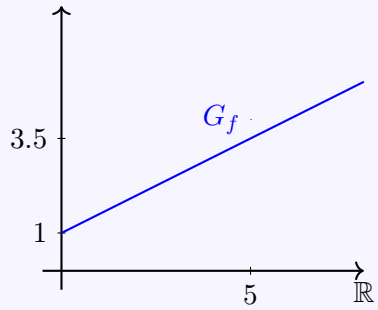
$$G_f := \{(a, f(a)) \mid a \in A\} \subseteq A \times B$$

der *Graph der Funktion* f.

(3) Es sei $n \in \mathbb{N}$. Wir schreiben $\mathbb{R}^n := \underbrace{\mathbb{R} \times \cdots \times \mathbb{R}}_{n-mal}$.

Beispiel 1.19. (1) Es gilt $\{1,2\} \times \{2,3,4\} = \{(1,2), (1,3), (1,4), (2,2), (2,3), (2,4)\}$.

(2) Es sei $f\colon \mathbb{R}_{\geq 0} \to \mathbb{R}; x \mapsto \frac{1}{2}x + 1$. Dann ist $G_f = \{(x, \frac{1}{2}x + 1) \mid x \in \mathbb{R}_{\geq 0}\} \subseteq \mathbb{R}^2$.



Definition 1.20. (1) Es seien $f\colon A \to B$ und $g\colon B \to C$ Funktionen. Dann heißt

$$g \circ f\colon A \to C; a \mapsto g(f(a))$$

das *Kompositum* von f und g.

(2) Es sei A eine Menge. Die Abbildung

$$\mathrm{id}_A\colon A \to A; a \mapsto a$$

heißt die *Identität* von A.

(3) Eine Funktion $f\colon A \to B$ heißt *umkehrbar*, falls eine Funktion $g\colon B \to A$ existiert, die $g \circ f = \mathrm{id}_A$ und $f \circ g = \mathrm{id}_B$ erfüllt. In diesem Fall ist auch g umkehrbar.

Satz 1.21 (Assoziativität des Kompositums). *Es seien $f\colon A \to B$, $g\colon B \to C$ und $h\colon C \to D$ Funktionen. Dann gilt*

$$h \circ (g \circ f) = (h \circ g) \circ f.$$

Beispiel 1.22. Es sei $f\colon A \to B$ eine Funktion.

(1) Es gilt $\mathrm{id}_B \circ f = f = f \circ \mathrm{id}_A$.

(2) Es sei f umkehrbar und $g, h\colon B \to A$ mit $g \circ f = \mathrm{id}_A = h \circ f$ und $f \circ g = \mathrm{id}_B = f \circ h$. Dann gilt

$$g = \mathrm{id}_A \circ g = (h \circ f) \circ g = h \circ (f \circ g) = h \circ \mathrm{id}_B = h.$$

Wir bezeichen daher mit $f^{-1}\colon B \to A$ die eindeutige Funktion mit $f^{-1} \circ f = \mathrm{id}_A$ und $f \circ f^{-1} = \mathrm{id}_B$ und bemerken, dass auch f^{-1} umkehrbar ist. Wir nennen f^{-1} die *Umkehrfunktion von f*.

Satz 1.23. *Es sei $f\colon A \to B$ eine Funktion. Es ist f umkehrbar genau dann, wenn f bijektiv ist. Insbesondere ist mit f auch f^{-1} bijektiv.*

Beispiel 1.24. (1) Die Funktion $f\colon \mathbb{R} \to \mathbb{R}; x \mapsto x - 10$ ist bijektiv. Damit existiert $f^{-1}\colon \mathbb{R} \to \mathbb{R}$ und ist durch $f^{-1}(x) = x + 10$ gegeben.

(2) Wir betrachten die Funktionen $f\colon \{1\} \to \{1,2\}; 1 \mapsto 1$ und $g\colon \{1,2\} \to \{1\}; 1 \mapsto 1, 2 \mapsto 1$. Da f nicht bijektiv ist, ist f nicht umkehrbar. Weiter gilt $g \circ f = \mathrm{id}_{\{1\}}$, aber $f \circ g \neq \mathrm{id}_{\{1,2\}}$.

1.3 Elementare Funktionen und ihre Eigenschaften

In diesem Abschnitt sammeln wir bekannte Funktionen $\mathbb{R} \to \mathbb{R}$ und diskutieren grundlegende Eigenschaften.

Definition 1.25. Es sei $A \subseteq \mathbb{R}$ und $f\colon A \to \mathbb{R}$ eine Funktion. Dann heißt f

(1) *monoton steigend (fallend)*, falls $f(a_1) \leq (\geq) f(a_2)$ für alle $a_1, a_2 \in A$ mit $a_1 \leq a_2$ gilt.

(2) *streng monoton steigend (fallend)*, falls $f(a_1) < (>) f(a_2)$ für alle $a_1, a_2 \in A$ mit $a_1 < a_2$ gilt.

Bemerkung 1.26. Es sei $A \subseteq \mathbb{R}$ und $f\colon A \to \mathbb{R}$ eine Funktion.

(1) Ist $f\colon A \to \mathbb{R}$ streng monoton steigend oder streng monoton fallend, so ist f injektiv: Wir betrachten nur den Fall, dass f streng monoton steigend ist. Seien $x \neq y \in \mathbb{R}$. Im Fall $x < y$ erhalten wir $f(x) < f(y)$ und im Fall $y < x$ erhalten wir $f(y) < f(x)$. In beiden Fällen gilt also $f(x) \neq f(y)$. Wir haben gezeigt, dass $f(x) = f(y)$ die Gleichung $x = y$ impliziert.

(2) Ist f injektiv, so muss f nicht streng monoton steigend oder streng monoton fallend sein:

$$f\colon \mathbb{R} \to \mathbb{R}; x \mapsto \begin{cases} x & \text{, falls } x \in \mathbb{R} \setminus \{1,2\} \\ 2 & \text{, falls } x = 1 \\ 1 & \text{, falls } x = 2. \end{cases}$$

(3) Die konstante Funktion $f\colon \mathbb{R} \to \mathbb{R}; x \mapsto 2$ ist monton steigend und fallend, aber nicht streng monton steigend oder fallend.

Definition 1.27. Es sei $A \subseteq \mathbb{R}$ eine Menge und $f, g\colon A \to \mathbb{R}$ Funktionen. Wir bezeichnen mit

$$f + g, f \cdot g\colon A \to \mathbb{R}$$

die Funktionen $(f+g)(a) = f(a) +_{\mathbb{R}} g(a)$ und $(f \cdot g)(a) = f(a) \cdot_{\mathbb{R}} g(a)$ für alle $a \in A$. Gilt $g(a) \neq 0$ für alle $a \in A$, so definieren wir weiter $\frac{f}{g}\colon A \to \mathbb{R}$ durch $\frac{f}{g}(a) = \frac{f(a)}{g(a)}$ für alle $a \in A$.

1.3.1 Potenzen, Wurzeln und Polynomfunktionen

Definition 1.28 (Ganzzahlige Potenzen und Wurzeln)**.** Es sei $a \in \mathbb{R}$ und $m \in \mathbb{Z}$. Wir definieren die *m-te Potenz von a* durch

$$a^m := \begin{cases} 1 & \text{falls } m = 0, \\ \underbrace{a \cdot a \cdots a \cdot a}_{m-\text{mal}} & \text{falls } m > 0, \\ \frac{1}{a^{|m|}} & \text{falls } m < 0 \text{ und } a \neq 0. \end{cases}$$

Gilt zusätzlich $a > 0$ und $m \in \mathbb{N}$, so bezeichnen wir mit $a^{\frac{1}{m}}$ (auch $\sqrt[m]{a}$) das eindeutig bestimmte Element in $\mathbb{R}_{\geq 0}$ mit $(a^{\frac{1}{m}})^m = a$ und nennen $a^{\frac{1}{m}} = \sqrt[m]{a}$ die *m-te Wurzel aus a*.

Definition 1.29 (Rationale Potenzen)**.** Es sei $a \in \mathbb{R}_{>0}$ und $r = \frac{p}{q} \in \mathbb{Q}$ mit $p \in \mathbb{Z}$, $q \in \mathbb{N}$. Wir definieren

$$a^r = a^{\frac{p}{q}} := (a^{\frac{1}{q}})^p = (\sqrt[q]{a})^p.$$

Für elementare Rechenregeln für Potenzen verweisen wir auf Formelsammlungen oder [2, 3.1.1].

Definition 1.30. Es sei $n \in \mathbb{N}_0$ und $I \subseteq \mathbb{R}$ eine Menge. Ein *Polynom vom Grad n* oder *Polynomfunktion vom Grad n* ist eine Funktion der Form

$$p\colon I \to \mathbb{R}; x \mapsto a_n x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0.$$

mit $a_0, \ldots, a_n \in \mathbb{R}$ und $a_n \neq 0$.[2] Ein Polynom vom Grad 1 nennen wir *lineares Polynom*. Die Nullfunktion $I \to \mathbb{R}; x \mapsto 0$ wird auch als *Nullpolynom* bezeichnet und besitzt nach Definition den Grad $-\infty$.

Satz 1.31. *Es seien $x_1, x_2, y_1, y_2 \in \mathbb{R}$ mit $x_1 \neq x_2$. Dann existiert genau ein lineares Polynom $p\colon \mathbb{R} \to \mathbb{R}; x \mapsto a_1 \cdot x + a_0$ mit $p(x_1) = y_1$ und $p(x_2) = y_2$.*

Beweis. Gesucht sind Elemente $a_1, a_0 \in \mathbb{R}$, sodass

$$a_1 \cdot x_1 + a_0 = y_1$$
$$a_1 \cdot x_2 + a_0 = y_2.$$

Wir erhalten $a_1 \cdot (x_2 - x_1) = (y_2 - y_1)$. Da $x_1 \neq x_2$ gilt, erhalten wir $a_1 = \frac{y_2 - y_1}{x_2 - x_1}$ und damit $a_0 = y_1 - a_1 \cdot x_1 = y_2 - a_1 \cdot x_2$. Die Probe zeigt, dass diese Wahl von a_1 und a_0 beide Gleichungen löst. □

1.3.2 Sinus, Cosinus und Tangens

Wir betrachten den Einheitskreis $K := \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\} \subseteq \mathbb{R}^2$ und erinnern uns, dass der Umfang von K gerade $2\pi \approx 6.28$ beträgt. Es sei $t \in \mathbb{R}$. Wir unterscheiden zwei Fälle:

- $t \geq 0$. Wir starten am Punkt $(1, 0)$ und laufen die Länge t auf dem Einheitskreis gegen den Uhrzeigersinn. Wir landen an einem Punkt $P(t) \in K$.
- $t < 0$. Wir starten am Punkt $(1, 0)$ und laufen die Länge $|t|$ auf dem Einheitskreis im Uhrzeigersinn. Wir landen an einem Punkt $P(t) \in K$.

Wir erhalten also eine Funktion

$$P\colon \mathbb{R} \to K; t \mapsto P(t).$$

Da jeder Punkt $P(t)$ durch eine x- und eine y-Koordinate gegeben ist, erhalten wir zwei Funktionen

$$\cos\colon \mathbb{R} \to \mathbb{R}; t \mapsto \cos(t)$$
$$\sin\colon \mathbb{R} \to \mathbb{R}; t \mapsto \sin(t)$$

mit der Eigenschaft $P(t) = (\cos(t), \sin(t))$ für alle $t \in \mathbb{R}$. Diese Funktionen heißen *Cosinus* und *Sinus*.

[2] Ist $p\colon I \to \mathbb{R}$ eine Polynom, $X \subseteq \mathbb{R}$ eine Menge mit $p(I) \subseteq X$, so nennen wir auch $p\colon I \to X$ ein Polynom.

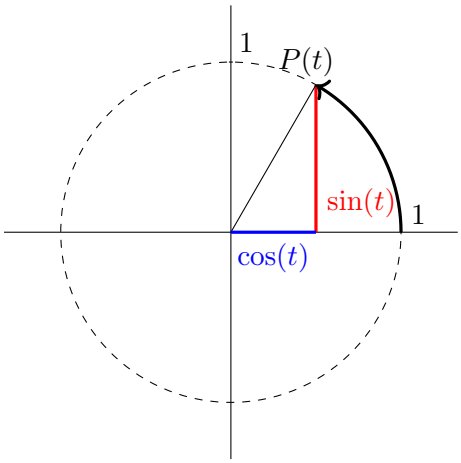


Abbildung 1.2: Veranschaulichung des Einheitskreises.

Satz 1.32. *(1) Die Funktionen* $\sin$ *und* $\cos$ *sind* 2π*-periodisch, das heißt es gilt* $\sin(t+2\pi n) = \sin(t)$ *und* $\cos(t+2\pi n) = \cos(t)$ *für alle* $t \in \mathbb{R}$ *und alle* $n \in \mathbb{Z}$.

(2) Es gilt $\sin(t)^2 + \cos(t)^2 = 1$ *für alle* $t \in \mathbb{R}$.

(3) Es gilt $\cos(\mathbb{R}) = [-1,1] = \sin(\mathbb{R})$.

(4) Für alle $t \in \mathbb{R}$ *gilt* $\sin(-t) = -\sin(t)$ *und* $\cos(t) = \cos(-t)$.

(5) Es gilt $\sin(s \pm t) = \sin(s)\cos(t) \pm \cos(s)\sin(t)$ *und* $\cos(s \pm t) = \cos(s)\cos(t) \mp \sin(s)\sin(t)$ *für alle* $t, s \in \mathbb{R}$.

(6) Es gilt $\sin(t + \frac{\pi}{2}) = \cos(t)$ *für alle* $t \in \mathbb{R}$.

(7) Es gilt $\sin(t) = 0$ *genau dann, wenn* $t \in \{n\pi \mid n \in \mathbb{Z}\}$ *und* $\cos(t) = 0$ *genau dann, wenn* $t \in \{n\pi + \frac{\pi}{2} \mid n \in \mathbb{Z}\}$.

Wir definieren *Tangens* als die Funktion

$$\tan\colon \mathbb{R} \setminus \{n\pi + \frac{\pi}{2} \mid n \in \mathbb{Z}\} \to \mathbb{R}; x \mapsto \frac{\sin(x)}{\cos(x)}.$$

Dies ist wohldefiniert, da $\cos(x) \neq 0$ für alle $x \in \mathbb{R} \setminus \{n\pi + \frac{\pi}{2} \mid n \in \mathbb{Z}\}$ gilt. Weiter gilt

$$\tan(-x) = \frac{\sin(-x)}{\cos(-x)} = \frac{-\sin(x)}{\cos(x)} = -\tan(x)$$

für alle $x \in \mathbb{R} \setminus \{n\pi + \frac{\pi}{2} \mid n \in \mathbb{Z}\}$.

Satz 1.33 (Umkehrfunktionen von $\sin$, $\cos$ und $\tan$)**.** *Es gelten die folgenden Aussagen.*

(1) Die Funktion $\cos\colon [0,\pi] \to [-1,1]$ *ist bijektiv. Die Umkehrfunktion bezeichen wir mit*

$$\arccos\colon [-1,1] \to [0,\pi].$$

(2) Die Funktion $\sin\colon [-\frac{\pi}{2},\frac{\pi}{2}] \to [-1,1]$ *ist bijektiv. Die Umkehrfunktion bezeichnen wir mit* $\arcsin\colon [-1,1] \to [-\frac{\pi}{2},\frac{\pi}{2}]$.

(3) Die Funktion $\tan\colon (-\frac{\pi}{2},\frac{\pi}{2}) \to \mathbb{R}$ *ist bijektiv. Die Umkehrfunktion bezeichnen wir mit*

$$\arctan\colon \mathbb{R} \to (-\frac{\pi}{2},\frac{\pi}{2}).$$

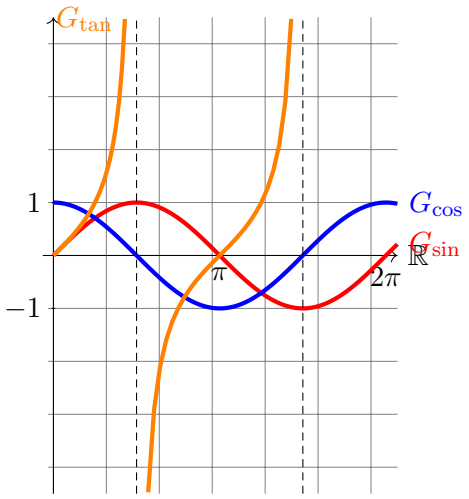


Abbildung 1.3: Graph von $\sin$, $\cos$ und $\tan$ auf dem Intervall $[0, 2\pi]$.

Kapitel 2

Grundbegriffe der Analysis

2.1 Folgen und Grenzwerte

Irrationale Zahlen zeichnen sich dadurch aus, dass ihre Dezimalzahldarstellung unendlich lang ist und sich nicht periodisch wiederholt. Auch wenn diese Eigenschaften sehr speziell klingen, treten irrationale Zahlen häufig auf, etwa hat das Quadrat mit Seitenlänge 1 die Diagonallänge $\sqrt{2} \in \mathbb{R} \setminus \mathbb{Q}$. Um solche Zahlen anzunähern bedient man sich häufig Iterationsverfahren.

Beispiel 2.1. Gesucht ist $\sqrt{3}$. Wir verwenden das Heron-Verfahren[a]. Wir beginnen mit dem Startwert

$$a_1 := 3$$

und definieren iterativ

$$a_{n+1} := \frac{1}{2}(a_n + \frac{3}{a_n}) \in \mathbb{Q}.$$

Man kann zeigen, dass die Zahlenfolge mit Werten in $\mathbb{Q}$ eine immer besser werdende Annäherung an $\sqrt{3}$ liefert. Genauer gilt

$$a_1 > a_2 > a_3 > \cdots > a_n > a_{n+1} > \cdots > \cdots \sqrt{3}.$$

Es stellt sich bei der Annäherung die Frage der „Qualität“: Die Zahlenfolge $b_1 := 1, b_2 := 1.1, b_3 := 1.11$, $b_4 := 1.111$,... ist auch eine immer besser werdende Annäherung an $\sqrt{3}$, allerdings werden wir $\sqrt{3}$ niemals beliebig nahe kommen, da $b_n \leq \frac{5}{4} < \sqrt{3}$ für alle $n \in \mathbb{N}$ gilt. Im Folgenden werden wir uns systematisch mit diesem Problem auseinander setzen. Dies führt auf den Begriff der Folgenkonvergenz.

[a] Wikipedia: https://de.wikipedia.org/wiki/Heron-Verfahren

Definition 2.2. (1) Eine (reelle) Folge ist eine Funktion $a\colon \mathbb{N} \to \mathbb{R}; n \mapsto a_n$. Wir schreiben kurz $a = (a_n)_{n\in\mathbb{N}}$, (a_n) oder $(a_1, a_2, a_3, a_4, \ldots)$.

(2) Eine Folge $(a_n)_{n\in\mathbb{N}}$ ist *konvergent*, falls ein $a \in \mathbb{R}$ existiert, sodass die folgende Bedingung erfüllt ist: Für jedes $\varepsilon \in \mathbb{R}_{>0}$ gibt es ein $n_\varepsilon \in \mathbb{N}$ sodass $|a_n - a| \leq \varepsilon$ für alle $n \geq n_\varepsilon$.

In diesem Fall ist a eindeutig bestimmt und wir nennen a den *Grenzwert* von $(a_n)_{n\in\mathbb{N}}$. Wir schreiben

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

und sagen, dass $(a_n)_{n\in\mathbb{N}}$ gegen a konvergiert. Eine konvergente Folge mit Grenzwert 0 heißt *Nullfolge*.

(3) Eine nicht konvergente Folge heißt *divergent*.

(4) Es sei $(a_n)_{n\in\mathbb{N}}$ eine Folge. Wir schreiben

$$\lim_{n\to\infty} a_n = \infty,$$

falls zu jedem Element $M > 0$ ein $n_M \in \mathbb{N}$ mit $a_n > M$ für alle $n \geq n_M \in \mathbb{N}$ existiert. Analog schreiben wir

$$\lim_{n\to\infty} a_n = -\infty,$$

falls zu jedem Element $M > 0$ ein $n_M \in \mathbb{N}$ mit $a_n < -M$ für alle $n \geq n_M \in \mathbb{N}$ existiert.

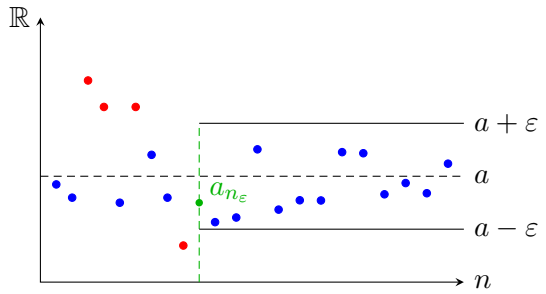


Abbildung 2.1: Veranschaulichung von Konvergenz.

Bemerkung 2.3. (1) Eine alternative Formulierung für Konvergenz gegen a ist: Für jedes $\varepsilon > 0$ ist die Menge $B_\varepsilon(a) := \{n \in \mathbb{N} \mid |a_n - a| > \varepsilon\}$ endlich. Das heißt, wenn wir endlich viele Folgenglieder vernachlässigen, haben alle anderen Folgenglieder den Abstand höchstens ε von a.

(2) Der Grenzwert a einer Folge $(a_n)_{n\in\mathbb{N}}$ muss nicht in der Menge $\{a_n \mid n \in \mathbb{N}\}$ enthalten sein.

(3) Gilt $\lim_{n\to\infty} a_n = \pm\infty$, so ist $(a_n)_{n\in\mathbb{N}}$ nach Definition divergent.

Beispiel 2.4. (1) Die Folge $(\frac{1}{n})_{n\in\mathbb{N}}$ konvergiert gegen 0: Es sei $\varepsilon > 0$. Dann existiert $n_\varepsilon \in \mathbb{N}$ mit $0 < \frac{1}{n_\varepsilon} < \varepsilon$[a]. Wir erhalten für $n \geq n_\varepsilon$ die Ungleichung $0 \leq \frac{1}{n} \leq \frac{1}{n_\varepsilon} < \varepsilon$ und damit $|\frac{1}{n} - 0| \leq \varepsilon$.

(2) Es sei $c \in \mathbb{R}$. Dann ist die konstante Folge $(a_n) = (c, c, c, c, \ldots)$ konvergent mit Grenzwert c: Für $\varepsilon > 0$ und $n \geq n_\varepsilon := 1$ gilt $|a_n - c| = |c - c| = 0 \leq \varepsilon$.

(3) Die Folge $((-1)^n)_{n\in\mathbb{N}} = (-1, 1, -1, 1, \ldots)$ ist divergent: Sei $a \in \mathbb{R}$. Dann gilt $|a - 1| \geq \frac{1}{2}$ oder $|a - (-1)| \geq \frac{1}{2}$. Im ersten Fall gilt für gerades $n \in \mathbb{N}$ $|a - (-1)^n| \geq \frac{1}{2}$ und im

zweiten Fall gilt für $n \in \mathbb{N}$ ungerade $|a - (-1)^n| \geq \frac{1}{2}$. Das heißt, a kann kein Grenzwert von $((-1)^n)_{n\in\mathbb{N}}$ sein. Da a beliebig gewählt war, ist die Folge divergent.

[a] siehe Archimedisches Axiom.

Satz 2.5. *Es seien $(a_n)_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}}$ konvergente Folgen mit Grenzwerten $\lim\limits_{n\to\infty} a_n = a$ und $\lim\limits_{n\to\infty} b_n = b$. Weiter sei $c \in \mathbb{R}$.*

(1) Die Folgen $(a_n + b_n)_{n\in\mathbb{N}}$, $(a_n \cdot b_n)_{n\in\mathbb{N}}$ und $(c \cdot a_n)_{n\in\mathbb{N}}$ sind konvergent und es gilt

$$\lim_{n\to\infty} a_n + b_n = a + b, \lim_{n\to\infty} a_n \cdot b_n = a \cdot b, \lim_{n\to\infty} c \cdot a_n = c \cdot a.$$

(2) Gelten $b \neq 0$ und $b_n \neq 0$ für alle $n \in \mathbb{N}$, so ist $(\frac{a_n}{b_n})_{n\in\mathbb{N}}$ konvergent mit $\lim\limits_{n\to\infty} \frac{a_n}{b_n} = \frac{a}{b}$.

Beispiel 2.6. (1) Es ist $(\frac{1}{n^2})_{n\in\mathbb{N}} = (\frac{1}{n} \cdot \frac{1}{n})_{n\in\mathbb{N}}$. Damit liefert Satz 2.5, dass $(\frac{1}{n^2})_{n\in\mathbb{N}}$ konvergent ist mit $\lim\limits_{n\to\infty} \frac{1}{n^2} = (\lim\limits_{n\to\infty} \frac{1}{n})(\lim\limits_{n\to\infty} \frac{1}{n}) = 0 \cdot 0 = 0$. Analog ergibt sich, dass jede Folge der Form $(\frac{a}{n^\ell})$ mit $a \in \mathbb{R}$ und $\ell \in \mathbb{N}$ eine Nullfolge ist.

(2) Wir untersuchen $(\frac{2n+7}{n+1})_{n\in\mathbb{N}}$ auf Konvergenz. Es gilt

$$\frac{2n+7}{n+1} = \frac{2n+2}{n+1} + \frac{5}{n+1} = 2 + \frac{5}{n+1}.$$

Die konstante Folge $(2)_{n\in\mathbb{N}}$ ist konvergent. Nach Satz 2.5(2) ist $(\frac{5}{n+1})_{n\in\mathbb{N}}$ konvergent mit Grenzwert $\lim\limits_{n\to\infty} \frac{5}{n+1} = \lim\limits_{n\to\infty} 5 \lim\limits_{n\to\infty} \frac{1}{n+1}$. Nach Satz 2.5 ist $(\frac{2n+7}{n+1})_{n\in\mathbb{N}}$ konvergent mit

$$\lim_{n\to\infty} \frac{2n+7}{n+1} = \lim_{n\to\infty} (2 + \frac{5}{n+1}) = \lim_{n\to\infty} 2 + \lim_{n\to\infty} \frac{5}{n+1} = 2 + 0 = 2.$$

Lemma 2.7 (Sandwichlemma). *Es seien $(a_n), (b_n), (c_n)$ Folgen. Es existiere $n_0 \in \mathbb{N}$ mit der Eigenschaft $a_n \leq b_n \leq c_n$ für alle $n \geq n_0$. Dann gelten die folgenden Aussagen.*

(1) Konvergieren $(a_n), (c_n)$ mit $\lim\limits_{n\to\infty} a_n = \lim\limits_{n\to\infty} c_n$, dann ist auch (b_n) konvergent mit $\lim\limits_{n\to\infty} b_n = \lim\limits_{n\to\infty} a_n = \lim\limits_{n\to\infty} c_n$.

(2) Gilt $\lim\limits_{n\to\infty} a_n = \infty$, so gilt auch $\lim\limits_{n\to\infty} b_n = \infty$.

(3) Gilt $\lim\limits_{n\to\infty} (c_n) = -\infty$, so gilt auch $\lim\limits_{n\to\infty} (b_n) = -\infty$.

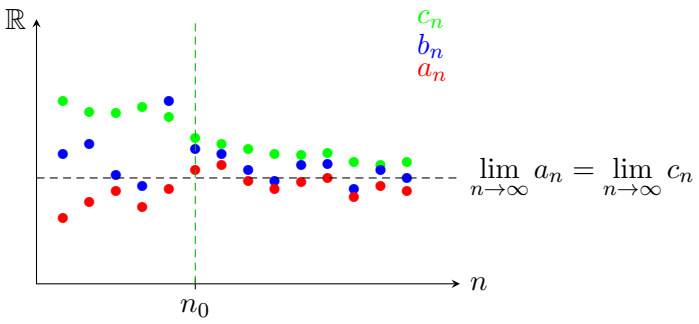


Abbildung 2.2: Veranschaulichung des Sandwich Lemmas.

Beispiel 2.8. (1) Es sei $0 < q < 1$. Dann gilt $1 > q > q^2 > q^3 > \cdots > 0$. Wir zeigen nun, dass $(q^n)_{n\in\mathbb{N}}$ gegen 0 konvergiert. Wir wollen dafür das Sandwichlemma anwenden. Es gilt $\frac{1}{q} > 1$ und wir finden $x \in (1,\infty)$ mit $\frac{1}{q} = 1 + x$. Der binomische Lehrsatz liefert

$$\frac{1}{q^n} = (x+1)^n = \sum_{i=0}^{n} \binom{n}{i} x^i 1^{n-i} \geq \binom{n}{0} x^0 + \binom{n}{1} x^1 = 1 + nx.$$

Umstellen liefert also $q^n \leq \frac{1}{1+nx} \leq \frac{1}{nx}$. Nun wenden wir das Sandwichlemma mit $a_n = 0$ und $c_n = \frac{1}{nx}$ für alle $n \in \mathbb{N}$ an und erhalten $(q^n) \to 0$.

(2) Analog zeigt man, dass (q^n) für $q \in [1,\infty)$ divergent ist mit $\lim_{n\to\infty} q^n = \infty$.

(3) Es sei $b_n := \frac{(-1)^n}{n^2}$ für alle $n \in \mathbb{N}$. Wir definieren $a_n := -\frac{1}{n}$ und $c_n := \frac{1}{n}$. Es gilt $a_n \leq b_n \leq c_n$ für alle $n \in \mathbb{N}$. Weiter sind (a_n) und (c_n) Nullfolgen. Damit gilt dies auch für (b_n).

(4) Wir betrachten die Folge $(\frac{9^n+5^n}{9^{n+1}})_{n\in\mathbb{N}}$. Es gilt $\frac{9^n+5^n}{9^{n+1}} = \frac{1+(\frac{5}{9})^n}{9}$. Anwendung von (1) und Satz 2.5 liefert die Konvergenz und $\lim_{n\to\infty}(\frac{9^n+5^n}{9^{n+1}}) = \frac{1}{9}$.

2.2 Reihen

Es seien $m \leq n \in \mathbb{Z}$ und $x_m, x_{m+1}, \ldots, x_{n-1}, x_n \in \mathbb{R}$. Wir schreiben die endliche Summe der Elemente in der Form

$$\sum_{i=m}^{n} x_i := x_m + x_{m+1} + \ldots + x_n.$$

Bereits in der Schule gehen wir mit unendlichen Summen um, ohne es zu merken. Z.B. lernen wir:

$$\frac{1}{9} = 0,\overline{1} = 0,111111\ldots$$

In Bruchschreibweise ergibt sich die unendliche Summe

$$\frac{1}{9} = 0,111111\ldots = \frac{1}{10} + \frac{1}{100} + \frac{1}{1000} + \cdots = \sum_{i=1}^{\infty} (\frac{1}{10})^i.$$

Wir konkretisieren dies nun mithilfe von Folgen und *Reihen*.

Beispiel 2.9. Es sei $q \in \mathbb{R} \setminus \{1\}$. Wir betrachten die Folge $(q^n)_{n\in\mathbb{N}}$ und definieren $s_n := \sum_{i=1}^{n} q^i = q + q^2 + \ldots + q^n$. Es gilt

$$(1-q)s_n = (1-q)\sum_{i=1}^{n} q^i = \sum_{i=1}^{n} q^i - \sum_{i=1}^{n} q^{i+1} = q - q^{n+1} = q(1-q^n).$$

Da $q \neq 1$ gilt, erhalten wir

$$s_n = \frac{q \cdot (1-q^n)}{1-q}.$$

Wählen wir nun zusätzlich $-1 < q < 1$, so liefert uns Satz 2.5, dass die Folgen im Zähler und Nenner von s_n konvergent sind. Wir erhalten

$$\lim_{n\to\infty} s_n = \frac{\lim\limits_{n\to\infty} q \cdot (1-q^n)}{\lim\limits_{n\to\infty} 1-q} = \frac{q}{1-q}.$$

Anders ausgedrückt: $\sum_{i=1}^{\infty} q^i = \frac{q}{1-q}$ und $\sum_{i=0}^{\infty} q^i = 1 + \frac{q}{1-q} = \frac{1-q+q}{1-q} = \frac{1}{1-q}$. Kehren wir zu unserem Beispiel zurück, so erhalten wir

$$\sum_{i=1}^{\infty} (\frac{1}{10})^i = \frac{1}{9}.$$

Motiviert von diesem Beispiel, betrachten wir in diesem Abschnitt Folgen $(a_n)_{n\in\mathbb{N}}$ und definieren daraus eine neue Folge $(s_n)_{n\in\mathbb{N}}$.

Definition 2.10. Es sei $(a_n)_{n\in\mathbb{N}}$ eine Folge. Dann heißt die Folge $(s_n)_{n\in\mathbb{N}}$ der *Partialsummen* mit

$$s_n := a_1 + a_2 + \ldots + a_n = \sum_{i=1}^{n} a_i$$

$(unendliche) Reihe$. Wir schreiben $\sum_{i=1}^{\infty} a_i$ und meinen damit sowohl die Folge der Partialsummen $(s_n)_{n\in\mathbb{N}}$ als auch den Grenzwert der Folge, falls dieser existiert.

Bemerkung 2.11. Ist $\sum_{i=1}^{\infty} a_i$ konvergent (also die Folge der Partialsummen (s_n)), so ist nach Satz 2.5 für jedes $c \in \mathbb{R}$ auch die Folge der Partialsummen $(c \cdot s_n)$ konvergent, d.h. $\sum_{i=1}^{\infty} c \cdot a_i$ und es gilt $\sum_{i=1}^{\infty} c \cdot a_i = c \cdot \sum_{i=1}^{\infty} a_i$.

Definition 2.12. Es sei $0 \neq q \in (-1, 1)$. Die Reihe $\sum_{n=1}^{\infty} q^n$ heißt *geometrische Reihe* zu q. Es ist $\sum_{n=1}^{\infty} q^n$ konvergent mit $\sum_{n=1}^{\infty} q^n = \frac{q}{1-q}$.

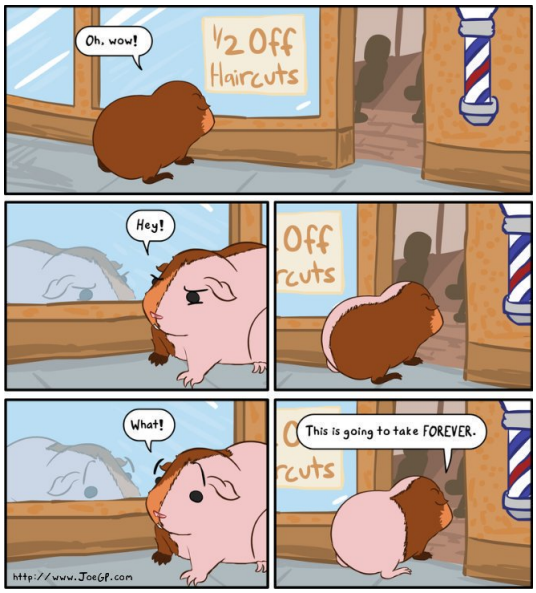


Abbildung 2.3: Es gilt $\sum_{n=1}^{\infty}(\frac{1}{2})^n = 1$.

Beispiel 2.13. Sie gewinnen im Lotto und wollen ihrem besten Freund X einen Geldbetrag G schenken, sodass sich dieser sein Auto im Wert von 60.000€ kaufen kann. Erhält X das Geld, so muss es mit 20 Prozent versteuert werden. Mit

$$G \cdot (1 - 0.20) = 60.000$$

schließen wir, dass wir unserem Freund 75.000€ schenken müssen. Ein komplizierterer Ansatz ist: Schenken wir X 60.000€ müssen wir ihm noch die 20 Prozent von 60.000€ überweisen. Dieser Betrag wird aber ebenfalls mit 20 Prozent versteuert, d.h. wir müssen noch 20 Prozent von 20 Prozent von 60.000€ überweisen. Dieser Prozess setzt sich unendlich lange fort. Mathematisch:

$$G = 60.000 + \sum_{i=1}^{\infty} 60.000 \cdot (\frac{1}{5})^i = 60.000 + 60.000 \cdot \sum_{i=1}^{\infty}(\frac{1}{5})^i = 60.000 + 60.000 \cdot \frac{1}{4} = 75.000.$$

Wir haben gesehen, dass für $q \in (-1, 1) \setminus \{0\}$ die Reihe $\sum_{n=1}^{\infty} q^n$ konvergiert und dass $(q^n)_{n\in\mathbb{N}}$ eine Nullfolge ist. Dies ist kein Zufall:

Satz 2.14. *Es sei $(a_n)_{n\in\mathbb{N}}$, sodass $\sum_{n=1}^{\infty} a_n$ konvergent ist. Dann ist $(a_n)_{n\in\mathbb{N}}$ eine Nullfolge.*

Beweis. Es sei $(a_n)_{n\in\mathbb{N}}$ eine Folge, sodass die Folge $s_n := \sum_{i=1}^{n} a_i$ gegen $s \in \mathbb{R}$ konvergiert. Es sei $s_0 := 0$. Wir betrachten die Folgen $(s_n)_{n\in\mathbb{N}}$ und $(s_{n-1})_{n\in\mathbb{N}}$. Beide Folgen konvergieren gegen s. Nach Satz 2.5 konvergiert auch die Folge

$$(s_n - s_{n-1})_{n\in\mathbb{N}} = (a_n)_{n\in\mathbb{N}}$$

gegen $s - s = 0$, ist also eine Nullfolge. □

Umgekehrt muss die Reihe einer Nullfolge nicht konvergieren:

Beispiel 2.15. Wir betrachten die Nullfolge $(\frac{1}{n})_{n\in\mathbb{N}}$ und die *harmonische Reihe* $\sum_{n=1}^{\infty}\frac{1}{n}$. Wir fassen Summanden wie folgt

$$\underbrace{1}_{a_1:=}+\underbrace{\frac{1}{2}}_{a_2:=}+\underbrace{(\frac{1}{3}+\frac{1}{4})}_{a_3:=}+\underbrace{(\frac{1}{5}+\frac{1}{6}+\frac{1}{7}+\frac{1}{8})}_{a_4:=}+\cdots$$

zusammen und bemerken $a_i \geq \frac{1}{2}$ für alle $i \in \mathbb{N}$. Formal gilt für $m \in \mathbb{N}$:

$$\begin{aligned} a_{m+1} :=& \sum_{k=2^{m-1}+1}^{2^m} \frac{1}{k} \geq \sum_{k=2^{m-1}+1}^{2^m} \frac{1}{2^m} \\ =& \frac{1}{2^m} \sum_{k=2^{m-1}+1}^{2^m} 1 = \frac{1}{2^m}(2^m - 2^{m-1}) = 1 - \frac{1}{2} = \frac{1}{2}. \end{aligned}$$

Damit gilt für $m \in \mathbb{N}$:

$$\sum_{n=1}^{\infty}\frac{1}{n} \geq \sum_{n=1}^{2^m}\frac{1}{n} = \sum_{i=1}^{m+1} a_i \geq \frac{m+1}{2}.$$

Dies zeigt $\sum_{n=1}^{\infty}\frac{1}{n} = \infty$.

2.3 Grenzwerte von Funktionen und stetige Funktionen

Definition 2.16 (Definition aus der Schule)**.** Sei $I \subseteq \mathbb{R}$ ein Intervall. Eine Funktion $f : I \to \mathbb{R}$ heißt *stetig*, wenn der Graph von f keine Sprungstellen hat bzw. wenn wir den Graph von f ohne Absetzen des Stiftes zeichnen können.

Probleme:

(1) Es ist uns gar nicht möglich den Graphen G_f genau zu zeichen, da eine Zahl auf der Zahlengeraden nur ein Punkt ohne Ausdehnung ist.

(2) Es ist uns nicht möglich die Funktion $\tan\colon (\frac{-\pi}{2}, \frac{\pi}{2}) \to \mathbb{R}$ zu zeichnen, da der Funktionsgraph unendlich lang ist.

Im Folgenden sei $I \subseteq \mathbb{R}$ eine Teilmenge von $\mathbb{R}$. Der Folgenbegriff erlaubt es uns Stetigkeit mathematisch sauber zu formulieren:

Definition 2.17. Es sei $f\colon I \to \mathbb{R}$ eine Funktion.

(1) Es sei $a \in \mathbb{R}$ und es existiere eine Folge $(a_n)_{n\in\mathbb{N}}$ mit den Eigenschaften

$$(*) \quad a_n \in I \text{ für alle } n \in \mathbb{N} \text{ und } \lim_{n\to\infty} a_n = a.$$

Ein Element $c \in \mathbb{R}$ heißt *Grenzwert* von f an der Stelle a[1], wenn für jede Folge $(b_n)_{n\in\mathbb{N}}$ mit den Eigenschaften $(*)$

$$\lim_{n\to\infty} f(b_n) = c$$

erfüllt ist. In diesem Fall schreiben wir

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

(2) Es sei $a \in I$. Die Funktion f heißt *stetig* in a, falls

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

gilt.[2]

(3) Ist f in jedem Punkt $a \in I$ stetig, so heißt f *stetig*.

Bemerkung 2.18. Eine äquivalente Definition für Stetigkeit von $f\colon I \to \mathbb{R}$ für $a \in I$ ist: Für jede gegen a konvergente Folge $(a_n)_{n\in\mathbb{N}}$ mit Folgengliedern in I gilt

$$\lim_{n\to\infty} f(a_n) = f(a) = f(\lim_{n\to\infty} a_n).$$

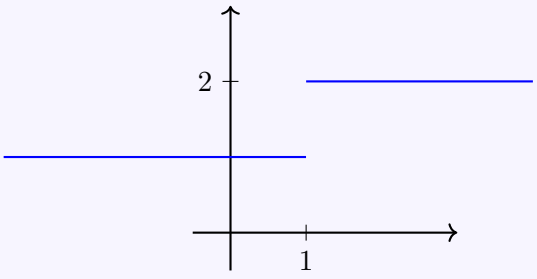
Wir können also den Limes und die Funktionsauswertung vertauschen.

Beispiel 2.19. Wir machen uns an Beispielen klar, dass der neue Stetigkeitsbegriff unsere Intuition mathematisch fasst:

(1) Wir betrachten die Funktion

$$f\colon \mathbb{R} \to \mathbb{R};\ f(x) = \begin{cases} 1 & \text{wenn } x < 1, \\ 2 & \text{wenn } x \geq 1. \end{cases}$$

Der Graph G_f hat die folgende Gestalt.

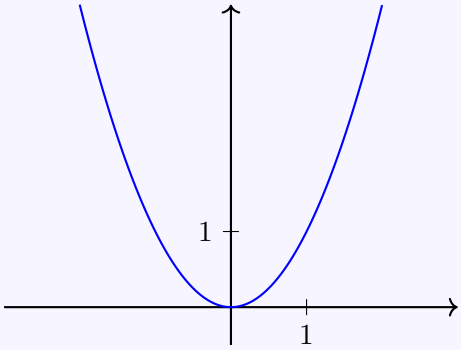


Bei $x = 1$ liegt eine Sprungstelle vor, das heißt die Funktion ist nicht stetig im Sinne der Schuldefinition. Wie nähern uns $x = 1$ von links durch die Folge $(a_n)_{n\in\mathbb{N}}$ mit $a_n = 1 - \frac{1}{n} \in \mathbb{R}$. Es gilt $\lim_{n\to\infty} a_n = 1$. Weiter gilt $a_n < 1$ und damit $f(a_n) = 1$ für alle $n \in \mathbb{N}$. Damit

[1] Beachte, dass a nicht in I enthalten sein muss.
[2] Da $a \in I$, erfüllt die konstante Folge $(a)_{n\in\mathbb{N}}$ die geforderten Eigenschaften aus (1).

ist $(f(a_n))_{n\in\mathbb{N}}$ die konstante Folge mit Grenzwert 1. Insbesondere gilt $\lim_{n\to\infty} f(a_n) = 1 \neq 2 = f(1) = f(\lim_{n\to\infty} a_n)$.

(2) Die Funktion $f\colon \mathbb{R} \to \mathbb{R}; x \mapsto x^2$ ist stetig. Dies ist intuitiv klar, wenn wir uns den Graphen von f ansehen:



Es sei $a \in \mathbb{R}$ beliebig und $(a_n)_{n\in\mathbb{N}}$ eine Folge mit Grenzwert a. Dann gilt $f(a_n) = a_n^2$ und Satz 2.5 liefert $\lim_{n\to\infty} f(a_n) = \lim_{n\to\infty} (a_n \cdot a_n) = \lim_{n\to\infty} a_n \cdot \lim_{n\to\infty} a_n = a^2 = f(a)$. Damit ist f stetig in a. Auf diese Weise zeigt man auch: Jede Polynomfunktion $p\colon I \to \mathbb{R}; x \mapsto \sum_{i=0}^{n} a_i x^i$ vom Grad $n \in \mathbb{N}_0$ ist stetig.

(3) Allgemeiner sind folgende Funktionen stetig: $\sin\colon \mathbb{R} \to \mathbb{R}, \cos\colon \mathbb{R} \to \mathbb{R}$. Wir werden im Laufe des Semesters noch weitere stetige Funktionen kennen lernen.

(4) Für Grenzwerte von Funktionen (sofern diese existieren) gelten dieselben Rechenregeln wie für Grenzwerte von Folgen (vgl. Satz 2.5).

(5) Wir berechnen den Grenzwert $\lim_{x\to 4} \frac{x^3-1}{x^2+5}$. Nach (2) gilt $\lim_{x\to 4} x^3 - 1 = 4^3 - 1 = 63$ und $\lim_{x\to 4} x^2 + 5 = 4^2 + 5 = 21$. Wir erhalten mit den Grenzwertgesetzen

$$\lim_{x\to 4} \frac{x^3-1}{x^2+5} = \frac{\lim_{x\to 4} x^3 - 1}{\lim_{x\to 4} x^2 + 5} = \frac{63}{21} = 3.$$

Weiter berechnen wir den Grenzwert

$$\lim_{x\to -1} \cos\left(\frac{x^2-1}{x+1}\pi\right).$$

Der Definitionsbereich der Funktion ist $\mathbb{R} \setminus \{-1\}$. Weiter existiert eine Folge mit Folgengliedern in $\mathbb{R} \setminus \{-1\}$, die gegen -1 konvergiert (etwa $(-1 + \frac{1}{n})_{n\in\mathbb{N}}$). Für $x \in \mathbb{R} \setminus \{-1\}$

gilt $\frac{x^2-1}{x+1}\pi = \frac{(x-1)(x+1)}{x+1}\pi = (x-1)\pi$. Wir erhalten $\lim\limits_{x\to -1} \frac{x^2-1}{x+1}\pi = \lim\limits_{x\to -1}(x-1)\pi = -2\pi$. Da die Cosinusfunktion stetig ist, erhalten wir

$$\lim_{x\to -1} \cos(\frac{x^2-1}{x+1}\pi) = \cos(\lim_{x\to -1} \frac{x^2-1}{x+1}\pi) = \cos(-2\pi) = 1.$$

Im Grundlagenkapitel bzw. im Vorkurs haben wir rationale Potenzen kennen gelernt. In [2, 3.1] steht „**Achtung:** Wir haben für $a > 0$ den Ausdruck a^x nur für rationales x definiert. Um die Definition auch auf beliebige $x \in \mathbb{R}$ ausdehnen zu können, braucht man weiteres mathematisches Werkzeug wie den Grenzwert-Begriff." Wir haben nun das nötige Rüstzeug um a^x für $x \in \mathbb{R}$ sauber zu definieren. Dazu nutzen wir den folgenden Satz.

Satz 2.20. *(1) Für jedes $x \in \mathbb{R}$ ist die Reihe $\exp(x) := \sum_{n=0}^{\infty} \frac{x^n}{n!}$ konvergent[a] und es gilt für alle $x, y \in \mathbb{R}$ die Gleichung*

$$\exp(x+y) = \exp(x) \cdot \exp(y).$$

(2) Die Funktion

$$\exp\colon \mathbb{R} \to \mathbb{R}_{>0}; x \mapsto \exp(x)$$

ist stetig, bijektiv und streng mononton wachsend.

[a] Die Konvergenz lässt sich mit dem sogenannten Quotientenkriterium für Reihen nachweisen.

Definition 2.21. (1) Die Zahl $e := \exp(1)$ heißt *Eulersche Zahl* und es gilt $e \approx 2.7183$. Es gilt $e \in \mathbb{R} \setminus \mathbb{Q}$. Für $x \in \mathbb{R}$ definieren wir

$$e^x := \exp(x).$$

In dieser Schreibweise gilt $e^{x+y} = \exp(x+y) = \exp(x) \cdot \exp(y) = e^x \cdot e^y$ für alle $x, y \in \mathbb{R}$.

(2) Die Umkehrfunktion von $\exp$ heißt *natürlicher Logarithmus* und wird mit $\ln\colon \mathbb{R}_{>0} \to \mathbb{R}$ bezeichnet. Es seien $x, y \in \mathbb{R}_{>0}$. Wir bemerken

$$\ln(x \cdot y) = \ln(e^{\ln(x)} \cdot e^{\ln(y)}) = \ln(e^{\ln(x)+\ln(y)}) = \ln(x) + \ln(y).$$

Beispiel 2.22. (1) Es sei $x \in \mathbb{R}_{>0}$ und $n \in \mathbb{N}$. Dann gilt

$$\ln(x) = \ln((x^{\frac{1}{n}})^n) = \ln(\overbrace{x^{\frac{1}{n}} \cdot x^{\frac{1}{n}} \cdots x^{\frac{1}{n}}}^{n\text{-mal}}) = n \ln(x^{\frac{1}{n}})$$

(2) Es gilt $e^0 = \exp(0) = 1$ und damit $0 = \ln(1) = \ln(x \cdot x^{-1}) = \ln(x) + \ln(x^{-1})$ für alle $x \in \mathbb{R}_{>0}$.

(3) Es seien $n, q \in \mathbb{N}$, $p \in \mathbb{Z} \setminus \{0\}$ und $a \in \mathbb{R}_{>0}$. Es gilt

$$e^{n \ln(a)} = e^{\ln(a)+\cdots+\ln(a)} = \overbrace{e^{\ln(a)} \cdot e^{\ln(a)} \cdots e^{\ln(a)}}^{n\text{-mal}} = a^n.$$

und damit

$$(*) \quad e^{\frac{p}{q}\ln(a)} = e^{p\frac{1}{q}\ln(a)} \overset{(1)}{=} e^{p\ln(a^{\frac{1}{q}})} = (a^{\frac{1}{q}})^p,$$

falls $p \in \mathbb{N}$ gilt. Im Fall $p < 0$ ist $-p \in \mathbb{N}$ und wir erhalten

$$e^{\frac{p}{q}\ln(a)} = e^{(-p/q)\cdot(-\ln(a))} \overset{(2)}{=} e^{(-p/q)\cdot\ln(a^{-1})} \overset{(*)}{=} ((a^{-1})^{\frac{1}{q}})^{-p} = a^{\frac{p}{q}}.$$

Definition 2.23 (Allgemeine Potenz)**.** Es sei $a \in \mathbb{R}_{>0}$ und $x \in \mathbb{R}$. Wir definieren

$$a^x := e^{x\ln(a)}.$$[3]

Satz 2.24. *Es sei $1 \neq a \in \mathbb{R}_{>0}$. Die Funktion*

$$p_a \colon \mathbb{R} \to \mathbb{R}_{>0}; x \mapsto a^x$$

ist stetig und bijektiv. Die Umkehrfunktion wird mit

$$\log_a \colon \mathbb{R}_{>0} \to \mathbb{R}$$

bezeichnet und es gilt $\log_a(x) = \frac{\ln(x)}{\ln(a)}$ für alle $x \in \mathbb{R}_{>0}$.

Satz 2.25 (Allgemeine Potenzgesetze)**.** *Es seien $a, b \in \mathbb{R}_{>0}, x, y \in \mathbb{R}$, dann gelten:*

(1) $(a^x)^y = a^{xy}$.

(2) $a^x \cdot b^x = (ab)^x$.

(3) $a^x \cdot a^y = a^{x+y}$.

(4) $\ln(a^x) = x \cdot \ln(a)$.

[3] Nach Beispiel 2.22 ist die Definition mit unserer Definition von a^r für $r \in \mathbb{Q}$ kompatibel.

2.4 Differenzierbarkeit

In der Differentialrechnung studiert man „geeignete“ Funktionen, indem man diese lokal durch Geraden approximiert. Gegeben sei uns eine geeignete Funktion $f\colon I \to \mathbb{R}$ und $a \in \mathbb{R}$. Wenn wir eine Tangente t_a an den Graphen G_f an der Stelle $f(a)$ anlegen können, so unterscheiden sich die Werte auf dem Graphen G_f in einer kleinen Umgebung nicht wesentlich von denen auf der Tangente t_a.

Ziel:

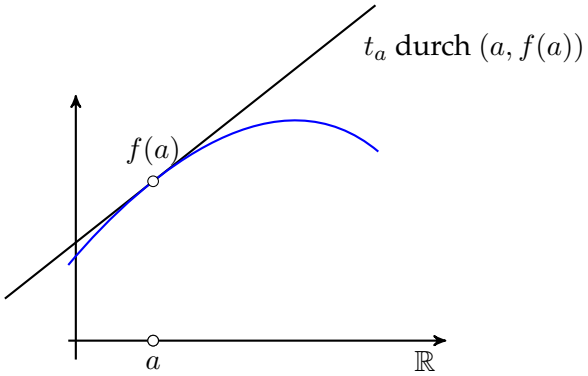


Abbildung 2.4: Tangente an G_f durch $f(a)$.

Probleme:

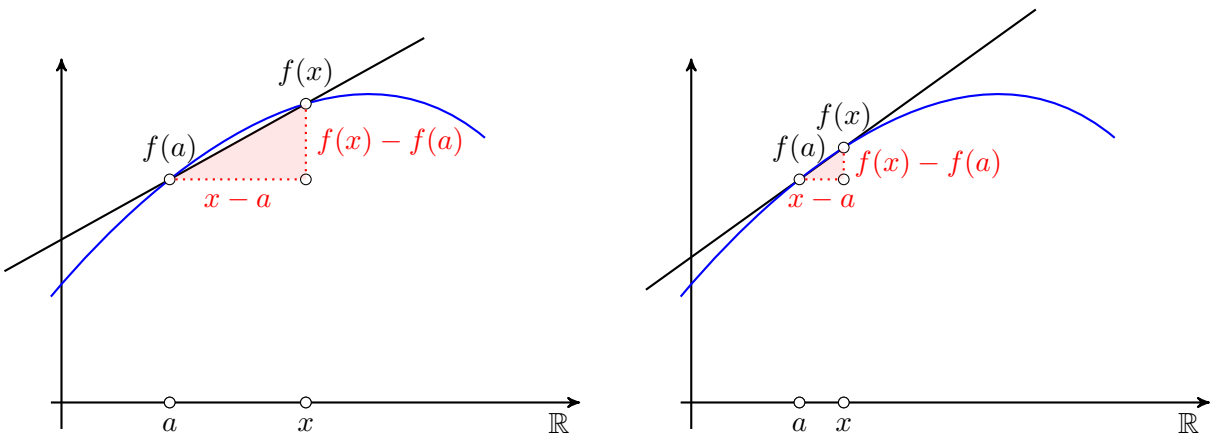
(i) Was bedeutet „geeignet“? Gibt es Funktionen, an die wir keine (eindeutige) Tangente anlegen können?

(ii) Wie berechnet man die Tangente?

Verfahren: Wir geben uns den Punkt $a \in I$ vor. Die Tangente t_a zu f an a ist eine Gerade, die durch den Punkt $(a, f(a))$ läuft. Wir benötigen noch die Steigung von t_a . Dazu betrachten wir für $x \in I \setminus \{a\}$ die Sekante, d.h. die Gerade, die durch die Punkte $(a, f(a))$ und $(x, f(x))$ läuft. Die Steigung dieser Sekante ist durch

$$\frac{f(x) - f(a)}{x - a}$$

gegeben.



Ist die Funktion f „geeignet", so erhalten wir die Steigung der Tangente t_a, indem wir x immer näher an a heran wandern lassen.

Definition 2.26. Es sei $f\colon I \to \mathbb{R}$ eine Funktion und $a \in I$. Die Funktion f heißt *differenzierbar* im Punkt a, falls der Grenzwert

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

existiert. Weiter heißt f *differenzierbar*, falls f differenzierbar in jedem Punkt in I ist.[4]

Beispiel 2.27. (1) Wir zeigen, dass $f\colon \mathbb{R} \to \mathbb{R}; x \mapsto x^3$ differenzierbar in jedem Punkt ist. Es sei $a \in \mathbb{R}$ und $x \in \mathbb{R} \setminus \{a\}$. Es gilt

$$\frac{f(x) - f(a)}{x - a} = \frac{x^3 - a^3}{x - a} = \frac{(x - a)(x^2 + xa + a^2)}{x - a} = x^2 + xa + a^2.$$

Damit existiert der Grenzwert in a und ist gegeben durch

$$f'(a) = \lim_{x \to a} \frac{f(x) - f(a)}{x - a} = \lim_{x \to a} x^2 + xa + a^2 = a^2 + a^2 + a^2 = 3a^2.$$

(2) Mit ähnlichen Argumenten wie in (1) lässt sich zeigen, dass die folgenden elementaren Funktionen differenzierbar sind. Wir listen ohne Beweis die zugehörigen Ableitungen (siehe z.B. [1]).

[4] Es sei $f\colon I \to \mathbb{R}$ differenzierbar in a. Wir berechnen den Grenzwert des *Differenzenquotienten*, also der Funktion

$$I \setminus \{a\} \to \mathbb{R}; x \mapsto \frac{f(x) - f(a)}{x - a}$$

an der Stelle a. Da der Grenzwert existiert, gibt es insbesondere eine Folge $(a_n)_{n \in \mathbb{N}}$ mit Folgengliedern in $I \setminus \{a\}$, sodass $\lim_{n \to \infty} a_n = a$ gilt. Für jede solche Folge gilt weiter

$$\lim_{n \to \infty} \frac{f(a_n) - f(a)}{a_n - a} = f'(a).$$

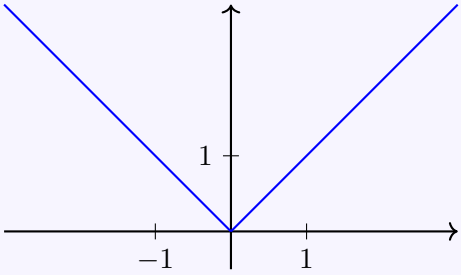
Funktion	Ableitung
$f(x) = x^a$ mit $a \in \mathbb{R}$	$f'(x) = ax^{a-1}$
$f(x) = a^x$ mit $a \in \mathbb{R}_{>0}$	$f'(x) = \ln(a)a^x$
$f(x) = \sin(x)$	$f'(x) = \cos(x)$
$f(x) = \cos(x)$	$f'(x) = -\sin(x)$

Wir bemerken den Spezialfall $f(x) = e^x$. Hier gilt $f'(x) = \ln(e)e^x = e^x$.

(3) Wir betrachten die Betragsfunktion

$$f\colon \mathbb{R} \to \mathbb{R}; x \mapsto |x|$$

und zeigen, dass f nicht differenzierbar in 0 ist.



Anschaulich ist dies klar, da wir in 0 keine eindeutig bestimmte Tangente anlegen können. Wir betrachten zwei Folgen $(a_n)_{n\in\mathbb{N}} = (\frac{-1}{n})_{n\in\mathbb{N}}$ und $(b_n)_{n\in\mathbb{N}} = (\frac{1}{n})_{n\in\mathbb{N}}$. Beide Folgen besitzen den Grenzwert 0 mit Folgengliedern in $\mathbb{R}\backslash\{0\}$. Einerseits gilt $\frac{f(a_n)-f(0)}{a_n-0} = \frac{|\frac{-1}{n}|}{\frac{-1}{n}} = -1$ und andererseits $\frac{f(b_n)-f(0)}{b_n-0} = \frac{|\frac{1}{n}|}{\frac{1}{n}} = 1$. Wir erhalten

$$\lim_{n\to\infty} \frac{f(a_n) - f(0)}{a_n - 0} = -1 \neq 1 = \lim_{n\to\infty} \frac{f(b_n) - f(0)}{b_n - 0}.$$

Damit existiert der Grenzwert nicht und f ist in 0 nicht differenzierbar.

Satz 2.28 (Rechenregeln). *(1) Es seien $f, g\colon I \to \mathbb{R}$ Funktionen, die differenzierbar in x sind. Dann sind $f + g$ und $f \cdot g$ differenzierbar in x und es gilt*

$$(f + g)'(x) = f'(x) + g'(x) \qquad \text{(Linearität)}$$
$$(f \cdot g)'(x) = f'(x)g(x) + f(x)g'(x). \qquad \text{(Produktregel)}$$

(2) Es seien $f, g\colon I \to \mathbb{R}$ Funktionen differenzierbar in x. Ist $g(x) \neq 0$, so ist $\frac{f}{g}$ differenzierbar in x mit

$$(\frac{f}{g})'(x) = \frac{f'(x)g(x) - f(x)g'(x)}{g(x)^2}. \qquad \text{(Quotientenregel)}$$

(3) *Es seien* $h\colon J \to \mathbb{R}$ *und* $f\colon I \to \mathbb{R}$ *Funktionen mit* $f(I) \subseteq J$. *Weiter sei* f *differenzierbar in* x *und* h *in* $f(x)$ *differenzierbar. Dann ist* $h \circ f\colon J \to \mathbb{R}$ *differenzierbar in* x *mit*

$$(h \circ f)'(x) = f'(x) \cdot h'(f(x)) \qquad \text{(Kettenregel)}.$$

(4) *Ist* $f\colon I \to J \subseteq \mathbb{R}$ *bijektiv und differenzierbar in* x *mit* $f'(x) \neq 0$, *so ist* $f^{-1}\colon J \to I$ *differenzierbar in* $y := f(x)$ *mit*

$$(f^{-1})'(y) = \frac{1}{f'(x)} = \frac{1}{f'(f^{-1}(y))}. \qquad \text{(Umkehrregel)}$$

Bemerkung 2.29. Es sei $f\colon \mathbb{R} \to \mathbb{R}; x \mapsto x^2 + e^{2x}$. Streng genommen gilt $f = f_1 + f_2 \circ f_3$ mit $f_1\colon \mathbb{R} \to \mathbb{R}; x \mapsto x^2$, $f_2\colon \mathbb{R} \to \mathbb{R}; x \mapsto e^x$ und $f_3\colon \mathbb{R} \to \mathbb{R}; x \mapsto 2x$. Dann gilt nach den Rechenregeln:

$$f'(x) = (f_1 + f_2 \circ f_3)'(x) = f_1'(x) + (f_2 \circ f_3)'(x) = f_1'(x) + f_3'(x) \cdot f_2'(f_3(x)) = 2x + 2 \cdot e^{2x}.$$

In der Differentialrechnung ist es üblich Funktionen $f\colon I \to \mathbb{R}$ mit ihrer Funktionsvorschrift zu identifizieren: Wir schreiben also

$$f'(x) = (x^2 + e^{2x})' = (x^2)' + (e^{2x})' = 2x + (2x)' \cdot e^{2x}$$

für alle $x \in \mathbb{R}$.
Warnung: Man kann bei dieser Schreibweise auch leicht durcheinander kommmen:
Es gilt $\sin'(2x) = \cos(2x)$, aber $(\sin(2x))' = (2x)'\cos(2x) = 2\cos(2x)$.

Beispiel 2.30. (1) Es sei $f\colon \mathbb{R} \to \mathbb{R}; x \mapsto e^{2x \cdot \cos(x)}$. Es gilt $f = g \circ (h_1 \cdot h_2)$ mit $h_1\colon \mathbb{R} \to \mathbb{R}; x \mapsto 2x$, $h_2\colon \mathbb{R} \to \mathbb{R}; x \mapsto \cos(x)$ und $g\colon \mathbb{R} \to \mathbb{R}; x \mapsto e^x$. Daher gilt

$$\begin{aligned} f'(x) &= (h_1 \cdot h_2)'(x) \cdot g'((h_1 \cdot h_2)(x)) \\ &= h_1'(x)h_2(x) + h_1(x)h_2'(x) \cdot e^{(h_1 \cdot h_2)(x)} \\ &= 2\cos(x) - 2x\sin(x) \cdot e^{2x\cos(x)} = 2(\cos(x) - x\sin(x))e^{2x\cos(x)}. \end{aligned}$$

Nach Identifikation gilt

$$\begin{aligned} f'(x) = (e^{2x\cos(x)})' &= (2x\cos(x))'e^{2x\cos(x)} = ((2x)'\cos(x) + 2x\cos'(x))e^{2x\cos(x)} \\ &= (2\cos(x) - 2x\sin(x))e^{2x\cos(x)} \end{aligned}$$

(2) Es sei $f = \tan$. Es gilt $f = \frac{\sin}{\cos}$. Nach der Quotientenregel gilt für $x \in (\frac{-\pi}{2}, \frac{\pi}{2})$ damit

$$f'(x) = \frac{\sin'(x)\cos(x) - \sin(x)\cos'(x)}{\cos(x)^2} = \frac{\cos(x)^2 + \sin(x)^2}{\cos(x)^2} = \frac{1}{\cos(x)^2}.$$

(3) Es sei $f\colon \mathbb{R} \to \mathbb{R}_{>0}; x \mapsto e^x$. Dann ist f bijektiv und differenzierbar mit $f' = f$. Sei also

$y \in \mathbb{R}_{>0}$. Wir erhalten

$$\ln'(y) = (f^{-1})'(y) = \frac{1}{f'(f^{-1}(y))} = \frac{1}{f(f^{-1}(y))} = \frac{1}{y}.$$

Definition 2.31. Es sei $f\colon I \to \mathbb{R}$ eine differenzierbar, so erhalten wir eine Funktion $f'\colon I \to \mathbb{R}$, die wir als *Ableitung von* f bezeichnen. Ist f' eine differenzierbare Funktion auf I so bezeichnen wir die Ableitung von f' mit f'' und nennen sie *zweite Ableitung von* f. Für $n \in \mathbb{N}$ schreiben wir auch $f^{(n)}$ für die *n-fache Ableitung von* f (falls diese existiert).

Beispiel 2.32. (1) Es gilt $(\sin(x))^{(3)} = (\cos(x))'' = (-\sin(x))' = -\cos(x)$ für alle $x \in \mathbb{R}$.

(2) Wir betrachten

$$f\colon \mathbb{R} \to \mathbb{R}; x \mapsto \begin{cases} x^2, \text{falls } x \geq 0, \\ 0, \text{falls } x \leq 0. \end{cases}$$

Man rechnet nach, dass f auf ganz $\mathbb{R}$ differenzierbar ist und die Ableitung ist durch

$$f'\colon \mathbb{R} \to \mathbb{R}; x \mapsto \begin{cases} 2x, \text{falls } x \geq 0, \\ 0, \text{falls } x \leq 0. \end{cases}$$

gegeben. Die Funktion f' ist differenzierbar auf $\mathbb{R} \setminus \{0\}$, nicht aber in 0.

(3) Ableitungen können auch in der Politik nützlich sein:

> „In the fall of 1972 President Nixon announced that the rate of increase of inflation was decreasing. This was the first time a sitting president used the third derivative to advance his case for reelection." - Hugo Rossi [a]

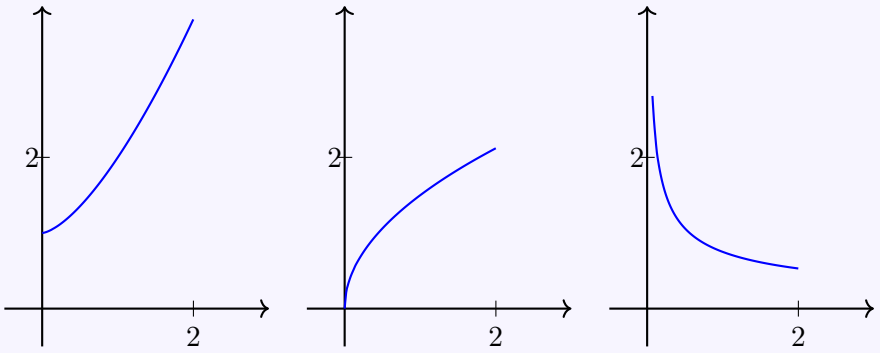
Wir betrachten als (sehr vereinfachtes) Beispiel eine Preisfunktion

$$f\colon [0,4] \to \mathbb{R}; x \mapsto x^{\frac{3}{2}} + 1,$$

wobei x das x-te Jahr seit Amtsantritt von Präsident Nixon bezeichnet. Wir erhalten

- $f(x) = x^{\frac{3}{2}} + 1$: Preisfunktion.
- $f'(x) = \frac{3}{2}x^{\frac{1}{2}}$: Anstieg des Preises = Inflation.
- $f''(x) = \frac{3}{4}x^{\frac{-1}{2}}$: Anstiegsrate der Inflation.
- $f'''(x) = \frac{-3}{8}x^{\frac{-3}{2}}$: Änderung der Anstiegsrate der Inflation.

Da $f'''(x) < 0$ für $x \in [0,4]$ gilt, ist die Anstiegsrate der Inflation in der gesamten Amtszeit fallend (siehe Satz 2.35), jedoch ist die Preisentwicklung alles andere als positiv für die Bevölkerung. Der Preis steigt in 4 Jahren von 1 Dollar auf 9 Dollar. Es findet also eine Verneunfachung des Preises statt. Die Inflation steigt von Jahr zu Jahr, nur die Geschwindigkeit des Wachstums nimmt ab. Die folgende Abbildung zeigt die Graphen von f,f' und f''.



[a]Mathematics Is an Edifice, Not a Toolbox, Notices of the AMS, v. 43, no. 10, October 1996

Satz 2.33. *Es sei $f : I \to R$ eine Funktion die differenzierbar in $x \in I$ ist. Dann ist f auch stetig in x.*

Definition 2.34. Es sei $f : I \to \mathbb{R}$ eine Funktion.

(1) $a \in I$ heißt globales Maximum (Minimum), falls $f(x) \leq f(a)$ ($f(x) \geq f(a)$) für alle $x \in I$ gilt.

(2) $a \in I$ heißt lokales Maximum (Minimum), falls $f(x) \leq f(a)$ ($f(a) \leq f(x)$) für alle $x \in I$ in einer Umgebung[5] von a gilt.

Definition 2.35. Es sei $f : I \to \mathbb{R}$ differenzierbar in $a \in I$ mit $f'(a) = 0$. Dann nennen wir a einen *kritischen Punkt*.

Satz 2.36. *Es seien I ein offenes Intervall, $f : I \to \mathbb{R}$ differenzierbar und $a \in I$.*

(1) Hat f in a ein lokales Maximum oder ein lokales Minimum, so gilt $f'(a) = 0$. Es handelt sich bei a also um einen kritischen Punkt.

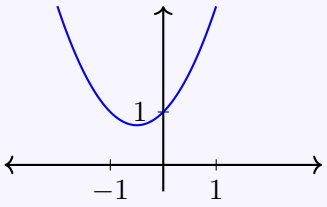
(2) Es sei $f'(a) = 0$ und $f''(a) \neq 0$. Gilt $f''(a) > 0$ ($f''(a) < 0$), so ist a ein lokales Minimum (lokales Maximum).

(3) Gilt $f'(a) > 0$ ($f'(a) < 0$), so ist f in einer Umgebung von a monoton wachsend (fallend).

(4) Es gilt $f'(x) = 0$ für alle $x \in I$ genau dann, wenn f konstant ist.

Beispiel 2.37. (1) Wir untersuchen die nach oben geöffnete Parabelfunktion $\mathbb{R} \to \mathbb{R}; x \mapsto x^2 + x + 1$.

[5]Eine Umgebung ist nach Definition ein Intervall $[a - \varepsilon, a + \varepsilon] \subseteq I$ für $\varepsilon > 0$ geeignet.

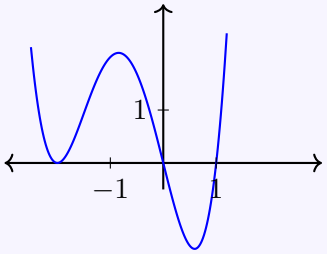


Es gilt $f'(x) = 2x + 1$. Damit kann ein lokales Minimum/Maximum nur bei $-\frac{1}{2}$ vorliegen. Mit $f''(x) = 2 > 0$ folgt, dass es sich um ein lokales Minimum handelt. Weiter besitzt f kein globales Maximum, aber ein globales Minimum in $-\frac{1}{2}$.

(2) Die Funktion $f\colon \mathbb{R} \to \mathbb{R}; x \mapsto x^3$ erfüllt $f'(0) = 0$ und $f''(0) = 0$. Aber 0 ist weder ein lokales Minimum noch ein lokales Maximum. Weiter besitzt f kein globales Maximum oder Minimum.

(3) Die Funktion $f\colon \mathbb{R} \to \mathbb{R}; x \mapsto x^4$ erfüllt $f'(0) = 0, f''(0) = 0$ und 0 ist trotzdem ein lokales Minimum. Weiter besitzt f kein globales Maximum, aber ein globales Minimum in 0.

(4) Die Funktion $f\colon \mathbb{R} \to \mathbb{R}; x \mapsto x^4 + 3x^3 - 4x$ erfüllt $f' = 4x^3 + 9x^2 - 4$ und die Nullstellen von f' sind $-2, -\frac{1}{8} - \frac{\sqrt{33}}{8}, -\frac{1}{8} + \frac{\sqrt{33}}{8}$.



Es gilt $f''(-2), f''(-\frac{1}{8} + \frac{\sqrt{33}}{8}) > 0$ und $f''(-\frac{1}{8} - \frac{\sqrt{33}}{8}) < 0$. Damit sind -2 und $-\frac{1}{8} + \frac{\sqrt{33}}{8}$ lokale Minima und $-\frac{1}{8} - \frac{\sqrt{33}}{8}$ ein lokales Maximum. Weiter besitzt f kein globales Maximum und $-\frac{1}{8} + \frac{\sqrt{33}}{8}$ ist das einzige globale Minimum.

2.5 Differenzierbarkeit in mehreren Variablen

Es sei $f : \mathbb{R}^n \to \mathbb{R}$ eine Funktion mit $n \geq 2$. Dann ist der Graph einer Funktion $f : \mathbb{R}^n \to \mathbb{R}$ eine Teilmenge von

$$\mathbb{R}^n \times \mathbb{R} = \mathbb{R}^{n+1} .$$

Im Fall $n = 2$ ist der Graph einer Funktion also eine Fläche im $\mathbb{R}^3$. Oftmals ist der Graph von f schwer zu zeichnen. Man verwendet daher auch ein Höhenlinienbild[6].

Definition 2.38. Es sei $f : \mathbb{R}^2 \to \mathbb{R}$ eine Funktion und $c \in \mathbb{R}$. Dann heißt

$$H_f(c) := \{(x, y) \in \mathbb{R}^2 \ : \ f(x, y) = c\}$$

[6]Höhenlinien werden oft für Landkarten benutzt: https://de.wikipedia.org/wiki/Hhenlinie

die *Höhenlinie von f zum Wert c.*

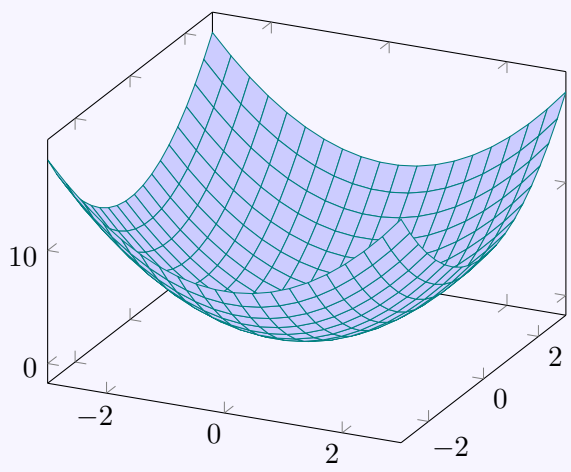
Beispiel 2.39. (1) Wir betrachten die Funktion

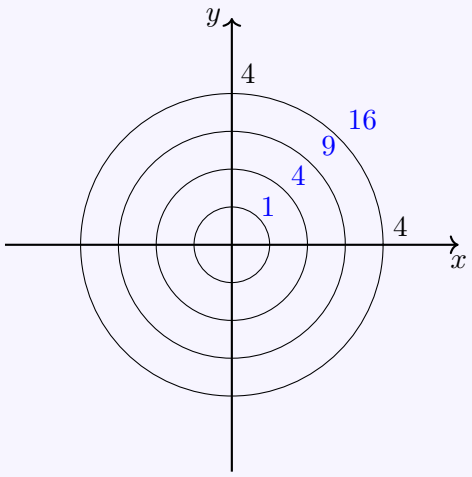
$$f : \mathbb{R}^2 \to \mathbb{R}; (x, y) \mapsto x^2 + y^2$$

Dann ist $G_f = \{(x, y, x^2 + y^2) \in \mathbb{R}^3 \mid x, y \in \mathbb{R}\} \subseteq \mathbb{R}^3$. Dann gilt

$$H_f(c) = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = c\}.$$

Für $c < 0$ gilt $H_f(c) = \varnothing$. Für $c \geq 0$ handelt es sich um einen Kreis mit Radius $\sqrt{c}$. Wir veranschaulichen G_f und ein Höhenlinienbild für $c \in \{1, 4, 9, 16\}$.

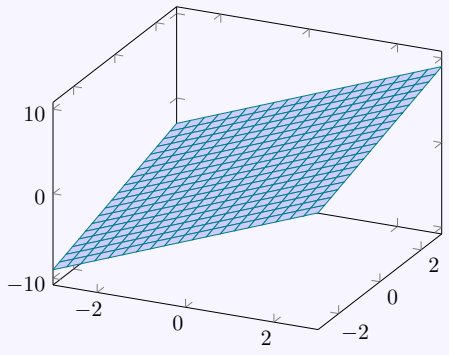


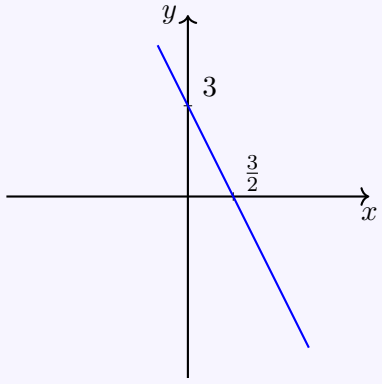


(2) Es sei $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}; f(x, y) = y + 2x$. Ist $c \in \mathbb{R}$ gegeben, so gilt

$$H_f(c) = \{(x, y) \in \mathbb{R}^2 \mid y + 2x = c\}$$

Damit ist $H_f(c)$ der Graph der Geraden $g(x) = -2x+c$. Insbesondere sind alle Höhenlinien parallel mit Steigung -2. Wir veranschaulichen G_f und ein Höhenlinienbild für $c = 3$.





Es sei $f : \mathbb{R}^n \to \mathbb{R}$ eine Funktion und $a = (a_1, \ldots, a_n) \in \mathbb{R}^n$ ein Punkt. Für $i \in \{1, \ldots, n\}$

definieren wir die Funktion

$$f_{a,i} : \begin{cases} \mathbb{R} \to \mathbb{R} \\ x \mapsto f(a_1, \ldots, a_{i-1}, x, a_{i+1}, a_{i+1}, \ldots, a_n) \end{cases}$$

Definition 2.40. Die *i-te partielle Ableitung von f im Punkt a* ist durch

$$\frac{\partial f}{\partial x_i}(a) := f'_{a,i}(a_i)$$

definiert (falls $f'_{a,i}(a_i)$ existiert). Wir nennen f *partiell differenzierbar*, falls für alle $i \in \{1, \ldots, n\}$ die i-ten partiellen Ableitungen existieren.

Beispiel 2.41. (1) Es sei $f : \mathbb{R}^2 \to \mathbb{R}; (x, y) \mapsto x^2 + xy$ eine Funktion. Dann gilt

$$\frac{\partial f}{\partial x}(x, y) = 2x + y \text{ und } \frac{\partial f}{\partial y}(x, y) = x.$$

(2) Es sei $f : \mathbb{R}^2 \to \mathbb{R}; (x, y) \mapsto x \cdot e^{x+y}$ eine Funktion. Wir erhalten

$$\begin{aligned} \frac{\partial f}{\partial x}(x, y) &= e^{x+y} + x\, e^{x+y} = (x + 1)e^{x+y} \\ \frac{\partial f}{\partial y}(x, y) &= x\, e^{x+y}\,. \end{aligned}$$

Definition 2.42. Es sei $f : \mathbb{R}^n \to \mathbb{R}$ und $a \in \mathbb{R}^n$.

(1) Die Funktion f hat ein *lokales Maximum* (*lokales Minimum*) im Punkt a, falls $f(x) \leq f(a)$ ($f(x) \geq f(a)$) für alle x in einer Umgebung von a gilt.

(2) Der Punkt $a \in \mathbb{R}^n$ ist ein *kritischer Punkt von f*, falls $\frac{\partial f}{\partial x_i}(a) = 0$ für $1 \leq i \leq n$ gilt.

Bemerkung 2.43. Lokale Maxima, lokale Minima, kritische Punkte und partielle Differenzierbarkeit lassen sich analog für Funktionen $f : D \to \mathbb{R}$ definieren, wobei $D \subseteq \mathbb{R}^n$ eine nicht-leere *offene* Teilmenge ist. Für uns sind offene Mengen stets von der Form $D = I_1 \times I_2 \times \cdots \times I_n \subseteq \mathbb{R}^n$, wobei $I_1, \ldots, I_n$ offene Intervalle in $\mathbb{R}$ sind. Im Folgenden sei $D \subseteq \mathbb{R}^n$ eine nicht-leere offene Menge.

Satz 2.44. *Es sei $f : D \to \mathbb{R}$ partiell differenzierbar. Hat f im Punkt $a \in D$ ein lokales Maximum oder ein lokales Minimum, so ist 0 ein kritischer Punkt von f.*

Beispiel 2.45. (1) Es sei $f : \mathbb{R}^2 \to \mathbb{R}; (x, y) \mapsto x \cdot e^{x+y}$. Dann gilt

$$\frac{\partial f}{\partial x}(x, y) = (x + 1)e^{x+y} \text{ und } \frac{\partial f}{\partial y}(x, y) = x\, e^{x+y}.$$

Ist also $a = (a_1, a_2)$ ein kritischer Punkt von f, so folgt $(a_1 + 1)e^{a_1+a_2} = 0 = a_1 e^{a_1+a_2}$. Wegen $e^z \neq 0$ für alle $z \in \mathbb{R}$ liefert die erste Gleichung $a_1 = -1$, während die zweite Gleichung zu $a_1 = 0$ führt. Damit besitzt f keine kritischen Punkte.

(2) Es sei $f\colon \mathbb{R}^2 \to \mathbb{R}; (x,y) \mapsto x^2 + 2y^2$. Es gilt

$$\frac{\partial f}{\partial x}(x,y) = 2x \text{ und } \frac{\partial f}{\partial y}(x,y) = 4y.$$

Damit ist $a = (0,0)$ der einzige kritische Punkt von f.

Definition 2.46. Sei $f : D \to \mathbb{R}$ eine Funktion, $a \in D$. Dann definieren wir für alle $i, j \in \{1, \ldots, n\}$

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(a) = \frac{\partial \frac{\partial f}{\partial x_j}}{\partial x_i}(a)$$

die *zweiten partiellen Ableitungen von f im Punkt a* (falls diese existieren) und schreiben auch f_{x_j} für $\frac{\partial f}{\partial x_j}$ für alle $j \in \{1, \ldots, n\}$.

Beispiel 2.47. Wir betrachten die Funktion

$$f : \mathbb{R}^2 \to \mathbb{R}\,;\ f(x,y) = y^2 \sin x.$$

Dann gilt

$$f_x = \frac{\partial f}{\partial x}(x,y) = y^2 \cos x \text{ und } f_y = \frac{\partial f}{\partial y}(x,y) = 2y \sin x.$$

Wir erhalten

$$\begin{aligned} \frac{\partial^2 f}{\partial x \partial y}(x,y) &= \frac{\partial f_y}{\partial x}(x,y) &&= 2y \cos x \text{ und} \\ \frac{\partial^2 f}{\partial y \partial x}(x,y) &= \frac{\partial f_x}{\partial y}(x,y) &&= 2y \cos x. \end{aligned}$$

Bei „schönen“ Funktionen spielt die Reihenfolge der Ableitungen keine Rolle:

Satz 2.48 (Vertauschbarkeit partieller Ableitungen)**.** *Sei $f : D \to \mathbb{R}$ zweimal partiell differenzierbar mit stetigen partiellen Ableitungen erster und zweiter Ordnung. Dann gilt*

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(a) = \frac{\partial^2 f}{\partial x_j \partial x_i}(a)$$

für alle $i, j \in \{1, \ldots, n\}$ und alle $a \in D$.

Wir hatten bereits bei Funktionen in einer Variablen gesehen, dass die Bedingung $f'(a) = 0$ nicht notwendig die Existenz eines lokalen Extremwerts anzeigt. Ebenso sind kritische Punkte im $\mathbb{R}^n$ manchmal keine Extrema.

Beispiel 2.49. Es sei $f\colon \mathbb{R}^2 \to \mathbb{R}; (x,y) \mapsto xy$. Dann gilt $\frac{\partial f}{\partial x}(x,y) = y, \frac{\partial f}{\partial y}(x,y) = x$. Also ist $(0,0)$ der einzige kritische Punkt von f. Wäre $(0,0)$ ein lokales Maximum von f, so wäre 0 ein lokales Maximum der Funktion $g : \mathbb{R} \to \mathbb{R}\ ,\ g(x) = f(x,x) = x^2$. Dies ist aber nicht möglich. Wäre $(0,0)$ ein lokales Minimum von f, so wäre 0 ein lokales Minimum der Funktion $h : \mathbb{R} \to \mathbb{R}\ ,\ g(x) = f(x,-x) = -x^2$. Auch dies ist nicht möglich. Also ist $(0,0)$ kein lokales Maximum oder lokales Minimum von f.

Es sei $D \subseteq \mathbb{R}^2$ offen und nicht-leer. Wir wollen nun ein Kriterium für Extremwerte von Funktionen $f : D \to \mathbb{R}$ angeben, die zweimal differenzierbar sind. Dazu betrachten wir eine neue Funktion

$$D(f) : D \to \mathbb{R}\ ;\ (x,y) \mapsto \frac{\partial^2 f}{\partial x\, \partial x}(x,y)\frac{\partial^2 f}{\partial y\, \partial y}(x,y) - \left(\frac{\partial^2 f}{\partial x\, \partial y}(x,y)\right)^2.$$

Satz 2.50. *Es sei $f : D \to \mathbb{R}$ eine Funktion mit stetigen zweiten partiellen Ableitungen. Ist $(a,b) \in D$ ein kritischer Punkt von f, so gelten die folgenden Aussagen.*

(1) Ist $D(f)(a,b) > 0$ und $\frac{\partial^2 f}{\partial x\, \partial x}(a,b) > 0$, so besitzt f ein lokales Minimum im Punkt (a,b).

(2) Ist $D(f)(a,b) > 0$ und $\frac{\partial^2 f}{\partial x\, \partial x}(a,b) < 0$, so besitzt f ein lokales Maximum im Punkt (a,b).

(3) Ist $D(f)(a,b) < 0$, so besitzt f kein lokales Extremum im Punkt (a,b).[a]

[a] Im Fall $D(f)(a,b) = 0$ macht der Satz keine Aussage.

Beispiel 2.51. Wir betrachten erneut das Beispiel $f : \mathbb{R}^2 \to \mathbb{R}\ ;\ f(x,y) = xy$. Dann gilt $\frac{\partial f}{\partial x}(x,y) = y, \frac{\partial f}{\partial y}(x,y) = x$, $\frac{\partial^2 f}{\partial x\, \partial x}(x,y) = 0$, $\frac{\partial^2 f}{\partial x\, \partial y}(x,y) = 1$ und $\frac{\partial^2 f}{\partial y\, \partial y}(x,y) = 0$. Mithin haben wir $D(f)(x,y) = 0 \cdot 0 - 1 \cdot 1 = -1$ und Satz 2.50(3) zeigt, dass f kein lokales Extremum im kritischen Punkt $(0,0)$ hat.

2.6 Integration

2.6.1 Integrale und Flächeninhalte

Es sei I ein Intervall, $f\colon I \to \mathbb{R}$ eine stetige Funktion und $[a,b] \subseteq I$ ein Intervall mit $f(x) \geq 0$ für alle $x \in [a,b]$. Wir wollen die Fläche berechnen, die durch den Graphen von f und den zur y-Achse parallelen Geraden durch a und b eingeschlossen wird.

Als erste Annäherung approximieren wir den Inhalt durch eine endliche Summme der Flächeninhalte von Rechtecken: Es sei $n \in \mathbb{N}$. Wir definieren $x_i := a + \frac{i}{n} \cdot (b-a)$ für $i \in \{0, \ldots, n\}$. Insbesondere gilt also $x_0 = a$ und $x_n = b$. Wir definieren die *Riemannsche Summe*

$$S(f,n) := \sum_{i=0}^{n-1} f(x_i)(x_{i+1} - x_i).$$

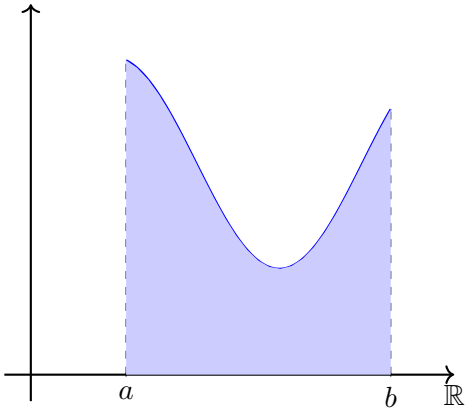


Abbildung 2.5: Veranschaulichung des Flächeninhalts einer Funktion auf dem Intervall $[a, b]$.

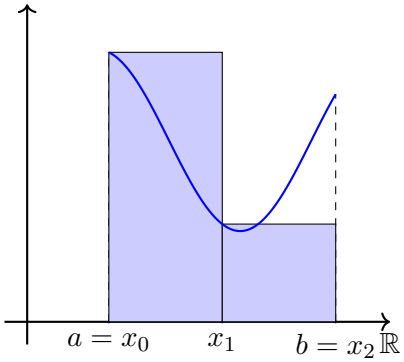


Abbildung 2.6: Illustration der Riemannschen Summen für $n = 2$.

Wählt man eine hinreichend feine Unterteilung, so wird die Annäherung von $S(f, n)$ an den Flächeninhalt besser.

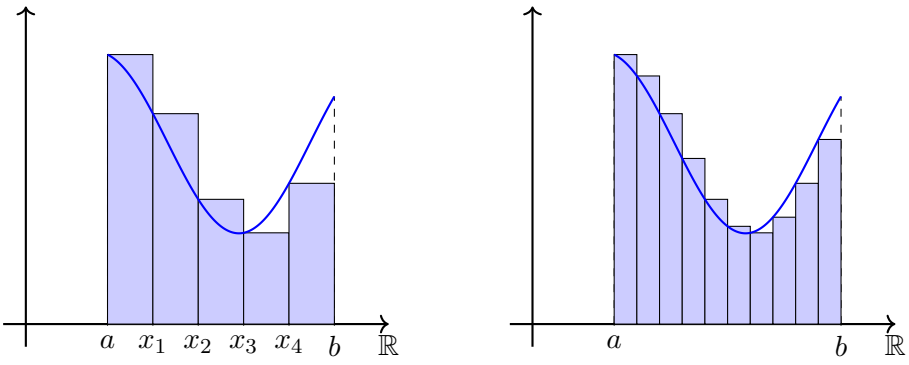


Abbildung 2.7: Illustration der Riemannschen Summen für $n = 5$ und $n = 10$.

Die Konstruktion der Riemann-Summen lässt sich auch durchführen, wenn nicht $f(x) \geq 0$ für alle $x \in [a, b]$ gilt. Wir erhalten also immer eine Folge $(S(f, n))_{n\in\mathbb{N}}$.

Satz 2.52. *Es sei $I \subseteq \mathbb{R}$ ein Intervall, $f\colon I \to \mathbb{R}$ eine stetige Funktion (nicht zwangsweise positiv) und $[a,b] \subseteq I$ ein Intervall.*

(1) Die Folge $(S(f,n))$ ist konvergent. Wir definieren

$$\int_a^b f(x)dx := \lim_{n\to\infty} S(f,n).$$

(2) Gilt $f(x) \geq 0$ für alle $x \in [a,b]$, so ist $\int_a^b f(x)dx$ der Flächeninhalt, der durch den Graphen von f und die x-Achse auf dem Intervall $[a,b]$ eingeschlossen wird.

Ein wichtiges Hilfsmittel zu Berechnung von $\int_a^b f(x)dx$ sind sogenante Stammfunktionen.

Definition 2.53. Sei $f : I \to \mathbb{R}$ eine Funktion. Eine differenzierbare Funktion $F : I \to \mathbb{R}$ heißt *Stammfunktion von f*, falls $F' = f$ gilt.

Satz 2.54 (Hauptsatz der Differential- und Integralrechnung)**.** *Es sei I ein Intervall und $f\colon I \to \mathbb{R}$ eine stetige Funktion.*

(1) Die Funktion f besitzt eine Stammfunktion[a]*.*

(2) Ist F eine Stammfunktion von f und $c \in \mathbb{R}$, so ist auch $F + c\colon I \to \mathbb{R}; x \mapsto F(x) + c$ eine Stammfunktion und alle Stammfunktionen sind von dieser Gestalt.

(3) Ist $F : I \to \mathbb{R}$ eine Stammfunktion und $a \leq b \in I$, so gilt

$$\int_a^b f(x)dx = F(b) - F(a) =: F(x)\Big|_a^b.$$

[a] Die Funktion $f(x) = e^{-x^2}$ ist stetig und besitzt damit eine Stammfunktion. Die Stammfunktion ist aber nicht geschlossen darstellbar.

Beweis. Wir beweisen (2) und gegeben eine Beweisskizze für (3). Ein vollständiger Beweis benötigt den sogenannten *Mittelwertsatz*.

(2) Es gilt $F' = f$. Sei $c \in \mathbb{R}$. Dann gilt $(F + c)' = F' + c' = F' = f$. Damit ist auch $F + c$ eine Stammfunktion. Sei G eine weitere Stammfunktion. Dann gilt $(G - F)' = G' - F' = 0$. Damit ist $G - F$ konstant nach Satz 2.36(4) und wir finden ein $c \in \mathbb{R}$ mit $G = F + c$.

(3) Wir wollen das Integral $\int_a^b f(x)dx$ berechnen und betrachten die Riemann-Summe

$$S(f,n) = \sum_{i=0}^{n-1} f(x_i)(x_{i+1} - x_i).$$

Da f die Ableitung von F ist, gilt für $i \in \{0, \ldots, n-1\}$

$$f(x_i) \approx \frac{F(x_{i+1}) - F(x_i)}{x_{i+1} - x_i} \text{ (vgl. Abschnitt 2.4).}$$

Dies liefert $f(x_i)(x_{i+1} - x_i) \approx F(x_{i+1}) - F(x_i)$ und wir erhalten

$$\begin{aligned} S(f,n) = \sum_{i=0}^{n-1} f(x_i)(x_{i+1} - x_i) &\approx \sum_{i=0}^{n-1} (F(x_{i+1}) - F(x_i)) = -\sum_{i=0}^{n-1} (F(x_i) - F(x_{i+1})) \\ &= -(F(x_0) - F(x_1) + F(x_1) - F(x_2) + \ldots + F(x_{n-1}) - F(x_n)) \\ &= -(F(x_0) - F(x_n)) = F(b) - F(a). \end{aligned}$$

□

Beispiel 2.55. (1) Es sei $f\colon I \to \mathbb{R}; x \mapsto \sum_{i=0}^{n} a_i x^i$ eine Polynomfunktion vom Grad $n \in \mathbb{N}$. Dann ist $F\colon I \to \mathbb{R}; x \mapsto \sum_{i=0}^{n} \frac{a_i}{i+1} x^{i+1}$ eine Stammfunktion von f.

(2) Es gilt $\int_1^2 \frac{1}{x} dx = \ln(x)\Big|_1^2 = \ln(2) - \ln(1) = \ln(2)$.

Satz 2.56. *(1) Es seien $f, g\colon [a,b] \to \mathbb{R}$ stetig und $\lambda \in \mathbb{R}$. Dann gilt*

$$\int_a^b f(x) + \lambda \cdot g(x) dx = \int_a^b f(x) dx + \lambda \int_a^b g(x) dx. \quad \textit{(Linearität des Integrals)}$$

(2) Es seien $f, g\colon [a,b] \to \mathbb{R}$ differenzierbar, sodass $f'\colon [a,b] \to \mathbb{R}$ und $g'\colon [a,b] \to \mathbb{R}$ stetig sind. Dann gilt

$$\int_a^b f(x) g'(x) dx = (f(x)g(x))\Big|_a^b - \int_a^b f'(x) g(x) dx. \quad \textit{(Partielle Integration)}$$

2.6.2 Integrale als Mittelwerte

Sind $n \in \mathbb{N}$ und $y_1, \ldots, y_n \in \mathbb{R}$, so errechnet sich der Mittelwert gemäß

$$\mathrm{MW}(y_1, \ldots, y_n) = \frac{1}{n} \sum_{i=1}^{n} y_i.$$

Nun betrachten wir eine stetige Funktion $f : [a,b] \to \mathbb{R}$ und wollen ein Maß für den Mittelwert der unendlich vielen Funktionswerte von f bestimmen. Für $n \in \mathbb{N}$ erinnern wir uns an die Definition

$$x_i = a + \frac{i}{n} \cdot (b - a)$$

für $i \in \{0, \ldots, n\}$ mit

$$S(f,n) = \sum_{i=0}^{n-1} f(x_i)(x_{i+1} - x_i).$$

Mit $x_{i+1} - x_i = \frac{1}{n}(b-a)$ für $i \in \{0, \ldots, n-1\}$ folgt

$$\begin{aligned} \mathrm{MW}(f,n) := \mathrm{MW}(f(x_0), \ldots, f(x_{n-1})) &= \frac{1}{n}\sum_{i=0}^{n-1} f(x_i) = \sum_{i=0}^{n-1} f(x_i)\frac{1}{n} \\ &= \sum_{i=0}^{n-1} f(x_i)\frac{x_{i+1}-x_i}{b-a} = \frac{1}{b-a}\sum_{i=0}^{n-1} f(x_i)(x_{i+1}-x_i) = \frac{1}{b-a}S(f,n). \end{aligned}$$

Damit folgt

$$\lim_{n\to\infty} \mathrm{MW}(f,n) = \frac{1}{b-a}\int_a^b f(x)dx$$

und wir nennen

$$\mathrm{MW}(f) := \frac{1}{b-a}\int_a^b f(x)dx\,.$$

den *Mittelwert von* f.

Beispiel 2.57. (1) Wir betrachten $f : [a,b] \to \mathbb{R}, x \mapsto c$. Hier würde man $\mathrm{MW}(f) = c$ erwarten. Es gilt

$$\mathrm{MW}(f) = \frac{1}{b-a}\int_a^b c\,dx \quad = \quad \frac{1}{b-a}(cx)\Big|_a^b = \frac{1}{b-a}c\cdot(b-a) = c.$$

(2) Es sei $f : [-1,1] \to \mathbb{R}, x \mapsto x$. Es gilt

$$\mathrm{MW}(f) = \int_{-1}^{1} x\,dx = \frac{1}{2}x^2\Big|_{-1}^{0} \;=\; 0.$$

2.6.3 Länge von Funktionsgraphen

Es sei $f\colon [a,b] \to \mathbb{R}$ eine Funktion. Bekanntlich ist der Abstand der Punkte $(a, f(a))$ und $(b, f(b))$ im $\mathbb{R}^2$ nach dem Satz von Pythagoras durch $\sqrt{(b-a)^2 + (f(b)-f(a))^2}$ gegeben. Wir wollen nun die Länge $\ell \in \mathbb{R}$ des Weges berechnen der entlang G_f zwischen $(a, f(a))$ und $(b, f(b))$ verläuft. Wir nehmen bei der Betrachtung an, dass f differenzierbar ist und geben im Folgenden eine Skizze für den Beweis einer Berechnungsformel an. Es sei $n \in \mathbb{N}$. Wir wenden den Satz von Pythagoras auf unsere Stützstellen an: Für $i \in \{0, \ldots, n-1\}$ sei

$$d_i(n) = \sqrt{(f(x_{i+1}) - f(x_i))^2 + (x_{i+1} - x_i)^2}.$$

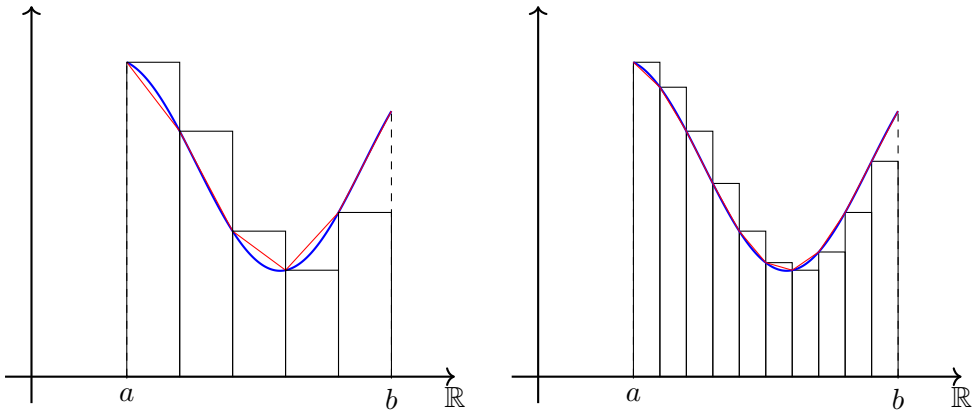


Abbildung 2.8: Illustration der Streckenzüge $d_i(n)$ für $n = 5$ und $n = 10$.

Da f differenzierbar ist, gilt (vgl. Abschnitt 2.4)

$$f'(x_i) \approx \frac{f(x_{i+1}) - f'(x_i)}{x_{i+1} - x_i},$$

das heißt $f(x_{i+1}) - f(x_i) \approx f'(x_i)(x_{i+1} - x_i)$ und wir erhalten

$$\begin{aligned} d_i(n) &\approx \sqrt{(f'(x_i)(x_{i+1} - x_i)^2 + (x_{i+1} - x_i)^2} = \sqrt{f'(x_i) + 1}\sqrt{(x_{i+1} - x_i)^2} \\ &= \sqrt{f'(x_i)^2 + 1} \cdot |x_{i+1} - x_i| = \sqrt{f'(x_i)^2 + 1} \cdot (x_{i+1} - x_i). \end{aligned}$$

Die Länge des Weges wird also durch

$$\sum_{i=0}^{n-1} d_i(n) \approx \sum_{i=0}^{n-1} \sqrt{f'(x_i)^2 + 1}\,(x_{i+1} - x_i) = S(\sqrt{(f')^2 + 1}, n).$$

angenähert. Auf der rechten Seite steht eine Riemann-Summe. Der Übergang zum Grenzwert liefert den folgenden Satz.

Satz 2.58. *Es sei $f\colon [a, b] \to \mathbb{R}$ differenzierbar. Damit ist die Länge ℓ des Weges von a nach b entlang G_f durch*

$$\ell = \int_a^b \sqrt{f'(x)^2 + 1}\, dx$$

gegeben.

2.6.4 Volumina

Mithilfe von Integralen lässt sich auch das Volumen geeigneter Objekte berechnen. Wir betrachten erneut eine stetige Funktion $f\colon [a,b] \to \mathbb{R}$ und nehmen der Einfachheit halber an, dass $0 \leq a < b$ und $f(x) \geq 0$ für alle $x \in [a,b]$ gilt. Wir wollen untersuchen welches Volumen das Objekt besitzt, das wir erhalten, wenn wir den Graphen G_f um die x-Achse rotieren lassen.

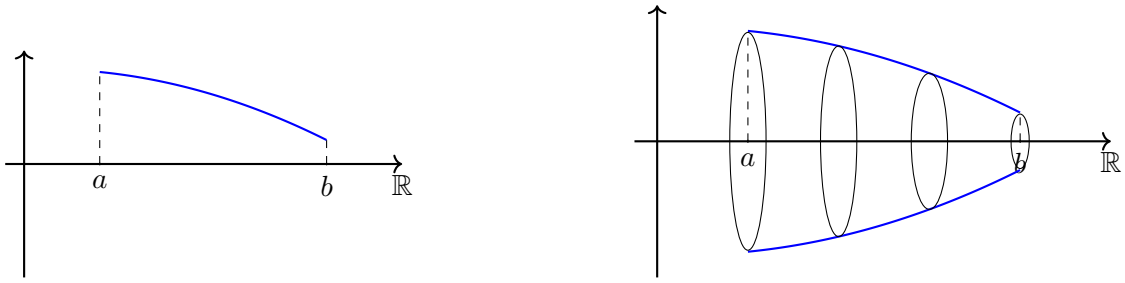


Abbildung 2.9: Illustration von G_f und des zugehörigen Rotationskörpers.

Der enstandene *Rotationskörper* wird durch die Menge

$$\{(x,y,z) \in [a,b] \times \mathbb{R}^2 \mid x^2 + y^2 \leq f(x)^2\}$$

beschrieben. Es sei $n \in \mathbb{N}$. Weiter sei wie vorher $x_i = a + \frac{i}{n} \cdot (b-a)$ für alle $i \in \{0,\dots,n\}$. Wir approximieren das Objekt durch eine Vereinigung von Zylindern $Z_0,\dots,Z_{n-1}$ der Höhe $\frac{1}{n}$ und von Radius $f(x_i)$ für $i \in \{0,\dots,n-1\}$. Der Zylinder Z_i hat damit das Volumen $\pi \cdot (x_{i-1} - x_i) \cdot f(x_i)^2$ und wir definieren

$$V(f,n) := \sum_{i=0}^{n-1} \pi \cdot (x_{i+1} - x_i) \cdot f(x_i)^2 = S(\pi \cdot f^2, n).$$

Dies ist eine Riemann-Summe. Damit gilt für das Volumen

$$V(f) := \lim_{n\to\infty} V(f,n) = \int_a^b \pi \cdot f(x)^2 dx = \pi \int_a^b f(x)^2 dx.$$

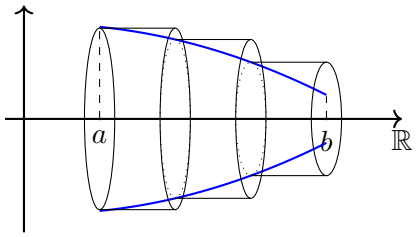


Abbildung 2.10: Illustration der Zylinder für $n = 3$.

Kapitel 3

Grundbegriffe der Linearen Algebra

In der linearen Algebra entwickelt man Verfahren, mit deren Hilfe systematisch lineare Gleichungssysteme gelöst werden können.

Definition 3.1 (Lineares Gleichungsystem)**.** Es seien $m, n \in \mathbb{N}$. Ein lineares Gleichungssystem (kurz LGS) ist ein System

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\ \vdots & & \vdots & & \ddots & & \vdots & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m \end{array}$$

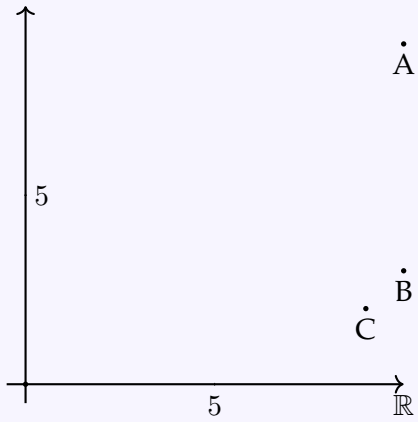
von m Gleichungen in den n unbekannten Elementen $x_1, \ldots, x_n \in \mathbb{R}$. Die Zahlen $a_{ij}, b_i \in \mathbb{R}$ gelten als bekannt.

Das folgende Beispiel zeigt, dass lineare Gleichungssysteme auch im Alltag auftauchen:

Beispiel 3.2. Drei Studierende A, B, C der Geowissenschaften sind gerade nach Kiel gezogen. Alle drei Wohnungen befinden sich im Nord-Ost Quadranten der Kieler Stadtkarte ausgehend vom Steinitz-Hörsaal. Die Entfernung geben wir in Koordinaten (in 100 Meter) an:

- $A = (10, 9)$
- $B = (10, 3)$
- $B = (9, 2)$.

Zum Beispiel liegt die Wohnung von Studentin A 1000 Meter östlich und 900 Meter nördlich vom Steinitz-Hörsaal.



Nachdem Person A, B und C zuhause ihre Übungszettel erfolgreich bearbeitet haben, wollen sie sich abends auf ein gemeinsames Getränk treffen. Was sie trinken wollen ist noch nicht klar, aber bei einer Sache sind sich alle einig: Niemand möchte einen Schritt mehr als die beiden anderen machen. Wir suchen also einen Punkt von dem A, B, C gleichweit entfernt sind. Mathematisch: Ein Kreis auf dem A,B und C liegen. Die (allgemeine) Kreisgleichung lässt sich für jeden beliebigen Kreis mit Mittelpunkt $M = (x_M, y_M)$ und Radius r aufstellen. Die Menge wird beschrieben durch

$$\begin{aligned}&\{(x,y) \in \mathbb{R}^2 \mid (x - x_M)^2 + (y - y_M)^2 = r^2\}\\ =&\{(x,y) \in \mathbb{R}^2 \mid x^2 + y^2 + ax + by + c = 0\}\\ =&\{(x,y) \in \mathbb{R}^2 \mid xa + yb + c = -x^2 - y^2\},\end{aligned}$$

wobei $a = -2x_M$, $b = -2y_M$ und $c = x_M^2 + y_m^2 - r^2$. Da die Punkte A, B, C auf dem Kreis liegen sollen, ergibt sich:

$$\begin{aligned}10a + 9b + c &= -181\\ 10a + 3b + c &= -109\\ 9a + 2b + c &= -85.\end{aligned}$$

Wir erhalten ein LGS in den Unbestimmten a, b, c. Kennen wir a, b und c, so können wir den Treffpunkt $(x_M, y_M) = (\frac{a}{-2}, \frac{b}{-2})$ und die Entfernung $r = \sqrt{c - x_M^2 - y_M^2}$ berechnen.

Fragen:

(1) Gibt es $a, b, c \in \mathbb{R}$ die das obige Gleichungssystem lösen? Anders ausgedrückt: Gibt es überhaupt einen Kreis auf dem A, B und C liegen?

(2) Wenn eine Lösung existiert, ist diese eindeutig bestimmt? Anders ausgedrückt: Ist der Kreis auf dem A, B und C liegen eindeutig bestimmt?

Im einfachsten Fall ($m = 1 = n$) handelt es sich bei einem LGS um

$$ax = b$$

mit $a, b \in \mathbb{R}$ fest. Wir wollen uns zunächst in diesem Fall klar machen, welche Lösungsmengen auftreten können und unterscheiden drei Fälle.

(1) $b \neq 0$ und $a = 0$, dann existiert kein $x \in \mathbb{R}$ mit $ax = b$.

(2) Gilt $a \neq 0$, so ist $\frac{b}{a}$ die einzige Lösung der Gleichung.

(3) Gilt $a = b = 0$, so gibt es unendlich viele Lösungen: Jedes $x \in \mathbb{R}$ löst die Gleichung.

Wir werden uns in diesem Kapitel davon überzeugen, dass die Lösungsmenge eines allgemeinen LGS von ähnlicher Gestalt sind.

3.1 Vektoren

Es sei $n \in \mathbb{N}$. Wir erinnern uns an die Definition

$$\mathbb{R}^n := \underbrace{\mathbb{R} \times \mathbb{R} \times \cdots \times \mathbb{R}}_{n-\text{mal}} = \{(x_1, x_2, \ldots, x_n) \mid x_1, \ldots, x_n \in \mathbb{R}\}.$$

Wir haben bislang in den Fällen $n = 2$ und $n = 3$ die Elemente von $\mathbb{R}^n$ als Punkte in der Ebene bzw. im Raum interpretiert. In der Physik ist die Interpretation als Vektor (gerichteter Pfeil) wichtig.[1] Wir fassen dabei $x = (x_1, \ldots, x_n)$ als den Pfeil von $(0, \ldots, 0)$ nach $(x_1, \ldots, x_n)$ auf.

Konvention: Wir schreiben Vektoren $x \in \mathbb{R}^n$ stets als *Spaltenvektoren*[2], das heißt

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

Weiter bezeichnet $x^T = (x_1, \ldots, x_n)$ den zugehörigen Zeilenvektor und

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

Wir können Vektoren addieren und auch mit Skalaren (reellen Zahlen) multiplizieren:

Definition 3.3. Es seien $a = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$, $b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \in \mathbb{R}^n$ und $\lambda \in \mathbb{R}$. Dann heißt

$$a + b := \begin{pmatrix} a_1 + b_1 \\ a_2 + b_2 \\ \vdots \\ a_n + b_n \end{pmatrix}$$

[1] Kräfte sind gerichtete physikalische Größen, die durch Vektoren dargestellt werden können.
[2] Dies wird sich im späteren Verlauf bei der Matrizenmultiplikation als praktisch erweisen.

die *Summe der Vektoren* a, b und

$$\lambda \cdot a := \begin{pmatrix} \lambda a_1 \\ \lambda a_2 \\ \vdots \\ \lambda a_n \end{pmatrix} \in \mathbb{R}^n$$

das λ*-Vielfache von* a. Weiter schreiben wir $-a := (-1) \cdot a$ und mit $0_{\mathbb{R}^n}$ sei der Nullvektor in $\mathbb{R}^n$ bezeichnet.

Bemerkung 3.4. (1) Es seien $x, y, z \in \mathbb{R}^n$ und $\lambda, \mu \in \mathbb{R}$. Dann gelten

$$(x+y)+z = x+(y+z) \qquad \lambda \cdot (x+y) = \lambda \cdot x + \lambda \cdot y$$
$$(\lambda \cdot \mu) \cdot x = \lambda \cdot (\mu \cdot x) \qquad (\lambda + \mu) \cdot x = \lambda \cdot x + \mu \cdot x.$$

Definition 3.5. Es seien $a = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}, b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \in \mathbb{R}^n$. Es heißt

$$\langle a, b \rangle := \sum_{i=1}^{n} a_i b_i \in \mathbb{R}$$

das *Skalarprodukt von* a *und* b. Weiter heißt $|a| := \sqrt{\langle a, a \rangle}$ die *Länge von* a.

Bemerkung 3.6. (1) Es gilt $\langle a, b \rangle \in \mathbb{R}$.

(2) Da $\langle a, a \rangle = \sum_{i=1}^{n} a_i^2 \geq 0$ gilt, dürfen wir die Wurzel ziehen. Weiter ist für $n \in \{2, 3\}$ nach dem Satz von Pythagoras $|a|$ genau die Länge des Streckenzugs zwischen $0_{\mathbb{R}^n}$ und a.

(3) Es seien $a, b, b_1, b_2 \in \mathbb{R}^n$ und $\lambda \in \mathbb{R}$. Es gelten die folgenden Rechenregeln.

(i) $\langle a, b \rangle = \langle b, a \rangle$.
(ii) $\langle a, \lambda b \rangle = \lambda \langle a, b \rangle = \langle \lambda a, b \rangle$.
(iii) $\langle a, b_1 + b_2 \rangle = \langle a, b_1 \rangle + \langle a, b_2 \rangle$.
(iv) $\langle b_1 + b_2, a \rangle = \langle b_1, a \rangle + \langle b_2, a \rangle$.[3]

Satz 3.7. *Seien* $a, b \in \mathbb{R}^n \setminus \{0\}$ *zwei von Null verschiedene Vektoren. Dann gilt* $\frac{\langle a,b \rangle}{|a| \cdot |b|} \in [-1, 1]$. *Damit gib es genau ein* $\alpha \in [0, \pi]$ *mit*

$$\langle a, b \rangle = |a| \cdot |b| \cdot \cos \alpha.$$

Das eindeutig bestimmte $\alpha \in [0, \pi]$ *heißt der Winkel zwischen den Vektoren* a *und* b.

[3] Diese Gleichung folgt schon aus (i) und (ii).

Definition 3.8. Es $a, b \in \mathbb{R}^n$ Vektoren. Dann heißen a und b *orthogonal*, wenn

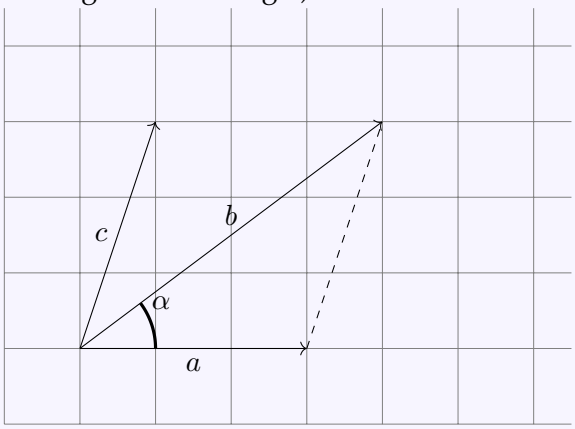
$$\langle a, b\rangle = 0$$

gilt.

Beispiel 3.9. Es seien $a, b \in \mathbb{R}^n$. Wir definieren $c := b - a$. Dann gilt

$$|c|^2 = \langle b-a, b-a\rangle = |a|^2 + |b|^2 - 2\langle a, b\rangle = |a|^2 + |b|^2 - 2 \cdot |a| \cdot |b| \cdot \cos\alpha.$$

Wir betrachten $n = 2$ und das Dreieck mit Eckpunkten $0_{\mathbb{R}^2}, a, b$ (hier fassen wir a, b als Punkte auf). Die Länge der Seite $\overline{0_{\mathbb{R}^2}, a}$ ist $|a|$, die Länge der Seite $\overline{0_{\mathbb{R}^2}, b}$ ist $|b|$ und die Länge der Seite $\overline{a, b}$ ist $|c|$ (Parallelverschiebung von c entlang a). Wir erhalten also den Cosinussatz.



3.2 Matrizen

Ein fundamentales Hilfsmittel für das Studium linearer Gleichungssysteme sind Matrizen.

Definition 3.10. Es seien $m, n \in \mathbb{N}$ natürliche Zahlen. Eine $(m \times n)$-*Matrix* $A = (a_{ij})$ ist ein Zahlenschema

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

mit Koeffizienten $a_{ij} \in \mathbb{R}$. Die Menge aller $(m \times n)$-Matrizen bezeichnen wir mit $\mathrm{Mat}_{m\times n}(\mathbb{R})$. Weiter definieren wir $\mathrm{Mat}_n(\mathbb{R}) := \mathrm{Mat}_{n\times n}(\mathbb{R})$. Definitionsgemäß sind zwei Matrizen $A = (a_{ij}), B = (b_{ij}) \in \mathrm{Mat}_{m\times n}(\mathbb{R})$ genau dann gleich, wenn $a_{ij} = b_{ij}$ für alle $i \in \{1, \dots, m\}$ und alle $j \in \{1, \dots, n\}$ gilt.

Bemerkung 3.11. Eine Matrix $(m \times n)$-Matrix A hat also m Zeilen und n Spalten. Es seien $i \in \{1, \dots, m\}$ und $j \in \{1, \dots, n\}$. Wir bezeichnen mit $A_i = (a_{i1}, \dots, a_{in})$ die i-te Zeile und mit ${}_jA = (a_{1j}, \dots, a_{mj})^T$ die j-te Spalte.

Definition 3.12. (1) Es seien $A = (a_{ij})$, $B = (b_{ij}) \in \mathrm{Mat}_{m\times n}(\mathbb{R})$. Wir definieren die *Summe*

$$A + B := (a_{ij} + b_{ij}) \in \mathrm{Mat}_{m\times n}(\mathbb{R}).$$

Für $\lambda \in \mathbb{R}$ setzen wir

$$\lambda \cdot A := (\lambda a_{ij}).$$

(2) Es seien $A = (a_{ij}) \in \mathrm{Mat}_{m\times n}(\mathbb{R})$, $B = (b_{ij}) \in \mathrm{Mat}_{n\times p}(\mathbb{R})$. Wir definieren das *Produkt* $A \cdot B := (c_{ij}) \in \mathrm{Mat}_{m\times p}(\mathbb{R})$ durch

$$c_{ij} := \sum_{k=1}^{n} a_{ik} b_{kj}$$

für alle $i \in \{1, \ldots, m\}$ und $j \in \{1, \ldots, p\}$.

Beispiel 3.13. Es sei $A := \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \in \mathrm{Mat}_{3\times 3}(\mathbb{R}) = \mathrm{Mat}_3(\mathbb{R})$ und $B := \begin{pmatrix} 10 & 11 \\ 12 & 13 \\ 14 & 15 \end{pmatrix} \in \mathrm{Mat}_{3\times 2}(\mathbb{R})$. Es gilt

$$A{\cdot}B = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \cdot \begin{pmatrix} 10 & 11 \\ 12 & 13 \\ 14 & 15 \end{pmatrix} = \begin{pmatrix} 1\cdot 10 + 2\cdot 12 + 3\cdot 14 & 1\cdot 11 + 2\cdot 13 + 3\cdot 15 \\ 4\cdot 10 + 5\cdot 12 + 6\cdot 14 & 4\cdot 11 + 5\cdot 13 + 6\cdot 15 \\ 7\cdot 10 + 8\cdot 12 + 9\cdot 14 & 7\cdot 11 + 8\cdot 13 + 9\cdot 15 \end{pmatrix} = \begin{pmatrix} 76 & 82 \\ 184 & 199 \\ 292 & 316 \end{pmatrix}.$$

Hingegen ist $B \cdot A$ nicht definiert, da die Anzahl der Spalten von B nicht mit der Anzahl der Zeilen von 3 überein stimmt.

Bemerkung 3.14. (1) Die Matrizen-Multiplikation ist eine weitreichende Verallgemeinerung des Skalarprodukts. Die kompliziert aussehende Formel besagt, dass sich der Eintrag c_{ij} als Skalarprodukt der i-ten Zeile von A und der j-ten Spalte von B ergibt (aufgefasst als Vektoren), genauer $c_{ij} = \langle (A_i)^T, {}_j\, B \rangle$.

(2) Es gelten die folgenden Rechenregeln.

(i) $A + (B + C) = (A + B) + C$ für alle $A, B, C \in \mathrm{Mat}_{m\times n}(\mathbb{R})$.

(ii) $A + B = B + A$ für alle $A, B \in \mathrm{Mat}_{m\times n}(\mathbb{R})$.

(iii) $A \cdot (B + C) = A \cdot B + A \cdot C$ für alle $A \in \mathrm{Mat}_{m\times n}(\mathbb{R})$, $B, C \in \mathrm{Mat}_{n\times p}(\mathbb{R})$.

(iv) $(A + B) \cdot C = A \cdot C + B \cdot C$ für alle $A, B \in \mathrm{Mat}_{m\times n}(\mathbb{R})$, $C \in \mathrm{Mat}_{n\times p}(\mathbb{R})$.

(v) $\lambda \cdot (AB) = (\lambda \cdot A) \cdot B = A \cdot (\lambda \cdot B)$ für alle $A, B \in \mathrm{Mat}_{m\times n}(\mathbb{R})$ und $\lambda \in \mathbb{R}$.

(vi) $(A \cdot B) \cdot C = A \cdot (B \cdot C)$ für alle $A \in \mathrm{Mat}_{m\times n}(\mathbb{R})$, $B \in \mathrm{Mat}_{n\times p}(\mathbb{R})$, $C \in \mathrm{Mat}_{p\times \ell}(\mathbb{R})$.

(3) Da wir Vektoren $x \in \mathbb{R}^n$ als Spalten notieren, können wir diese als Elemente von $\mathrm{Mat}_{n\times 1}(\mathbb{R})$ auffassen und insbesondere mit einer Matrix $A = (a_{ij}) \in \mathrm{Mat}_{m\times n}(\mathbb{R})$ multiplizieren:

$$A \cdot x = A \cdot \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \sum_{i=1}^{n} a_{1i} x_i \\ \sum_{i=1}^{n} a_{2i} x_i \\ \vdots \\ \sum_{i=1}^{n} a_{mi} x_i \end{pmatrix}.$$

Es sei ein LGS

$$\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\ \vdots & & \vdots & & \ddots & & \vdots & & \vdots \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m \end{array}$$

gegeben. Wir interpretieren dies als Matrixgleichung

$$A \cdot x = b$$

mit $A = (a_{ij}) \in \mathrm{Mat}_{m\times n}(\mathbb{R})$, $x = (x_1, \ldots, x_n)^T \in \mathbb{R}^n$ und $b = (b_1, \ldots, b_m)^T \in \mathbb{R}^m$ und definieren die *erweiterte Koeffizientenmatrix*

$$(A|b) := \left(\begin{array}{cccc|c} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_n \end{array}\right) \in \mathrm{Mat}_{m\times(n+1)}(\mathbb{R}),$$

und die *Lösungsmenge von* $(A|b)$

$$\mathbb{L}(A, b) := \{x \in \mathbb{R}^n \mid A \cdot x = b\}.$$

Wir bemerken, dass $\mathbb{L}(A, b)$ aus genau den Lösungen des zugehörigen LGS besteht, wenn wir Lösungen des LGS als Spaltenvektoren notieren.

3.3 Die Zeilenstufenform

Definition 3.15. Es sei $0 \neq A = (a_{ij}) \in \mathrm{Mat}_{m\times n}(\mathbb{R})$ eine Matrix. Wir sagen A ist in *Zeilenstufenform*, falls folgende Bedingungen erfüllt sind.

(1) Es existiert $1 \leq r_A \leq m$, sodass $A_1, \ldots, A_{r_A}$ von null verschiedene Zeilen sind und alle anderen Zeilen der Matrix null sind.

(2) Es gilt $1 \leq s(1) < s(2) < \ldots < s(r_A) \leq n$, wobei $s(i) := \min\{j \in \{1, \ldots, n\} \mid a_{ij} \neq 0\}$ und $a_{is(i)} = 1$ für alle $i \in \{1, \ldots, r_A\}$. Wir nennen die Spalten ${}_{s(1)}A, \ldots, {}_{s(r_A)}A$ *Pivotspalten*.

Die Zahl r_A heißt der *Rang von* A.

Bemerkung 3.16. (1) Da $1 \leq s(1) < s(2) < \ldots < s(r_A) \leq n$ paarweise verschiedene natürliche Zahlen sind, gilt auch $r_A \leq n$, also $r \leq \min\{m, n\}$.

(2) Die Definition besagt: Eine Matrix hat Zeilenstufenform, wenn folgende Bedingungen erfüllt sind:

- Ist eine Zeile eine Nullzeile, so auch alle darauffolgenden Zeilen.
- Ist eine Zeile keine Nullzeile, so ist der erste von null verschieden Eintrag eine 1.
- In jeder nicht Nullzeile steht der erste Eintrag $\neq 0$ weiter rechts als in der Vorgänger-Zeile.

Beispiel 3.17. (1) $\begin{pmatrix} 0 & 0 & \boxed{①} & \pi & 1 \\ 0 & 0 & 0 & 0 & \boxed{①} \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \in \mathrm{Mat}_{3\times 5}(\mathbb{R})$ ist in Zeilenstufenform mit $r_A = 2$, $s(1) = 3$, $s(2) = 5$.

(2) $\begin{pmatrix} \boxed{①} & 4 & 2 & 5 & -1 \\ 0 & \boxed{①} & 0 & 7 & 4 \\ 0 & 0 & 0 & \boxed{①} & 2 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \in \mathrm{Mat}_{4\times 5}(\mathbb{R})$ ist in Zeilenstufenform mit $r_A = 3$, $s(1) = 1$, $s(2) = 2$ und $s(3) = 4$.

(3) Die Matrizen $\begin{pmatrix} \boxed{②} & 0 & 2 & 0 & -1 \\ 0 & \boxed{①} & 0 & 0 & 4 \\ 0 & 0 & 0 & \boxed{①} & 2 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \in \mathrm{Mat}_{4\times 5}(\mathbb{R})$, $\begin{pmatrix} 1 & 0 & 2 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \in \mathrm{Mat}_{4\times 5}(\mathbb{R})$ sind nicht in Zeilenstufenform.

Satz 3.18. *Es sei $0 \neq A \in \mathrm{Mat}_{m\times n}(\mathbb{R})$ eine Matrix in Zeilenstufenform von Rang $r = r_A$ und $b \in \mathbb{R}^m$, dann tritt genau einer der folgenden Fälle auf:*

(1) Es gilt $b_i \neq 0$ für mindestens ein $r < i \leq m$. Dann gilt $\mathbb{L}(A, b) = \varnothing$.

(2) Es gilt $b_{r+1} = b_{r+2} = \ldots = b_m = 0$. Dann ist $\mathbb{L}(A, b) \neq \varnothing$, genauer:

(1) Gilt $n = r$, so ist existiert genaue eine Lösung.

(2) Gilt $r \neq n$, so existieren unendlich viele Lösungen in $n - r$ Freiheitsgraden. Die Freiheitsgrade sind durch die Spalten gegegeben, die keine Pivotspalten sind.

Beispiel 3.19. Die Bestimmung der Lösungsmenge machen wir uns anhand von Beispielen klar.

(1) Es sei $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ und $b = (0, 1)^T$. Es gilt $r_A = 1$. Damit $\mathbb{L}(A, b) = \varnothing$.

(2) Es sei $A = \begin{pmatrix} 1 & -2 & 1 & 0 & 2 \\ 0 & 1 & -1 & 1 & 4 \\ 0 & 0 & 0 & 1 & 3 \end{pmatrix}$ und $b = \begin{pmatrix} 3 & 4 & 5 \end{pmatrix}^T$. Wir erhalten die erweitere Koeffizientenmatrix

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

Es gilt $r_A = 3$ und die Menge der nicht-Pivotspalten von A ist $\{A^3, A^5\}$. Damit haben wir zwei Freiheitsgrade gegeben durch λ_3 und λ_5. Nun errechnen wir aus der letzten

Zeile

$$1 \cdot \lambda_4 + 3 \cdot \lambda_5 = 5$$

und erhalten $\lambda_4 = 5 - 3 \cdot \lambda_5$. Einsetzen in die zweite Zeile ergibt

$$\lambda_2 - \lambda_3 + \lambda_4 + 4 \cdot \lambda_5 = 4$$

und wir erhalten $\lambda_2 = \lambda_3 - \lambda_4 - 4 \cdot \lambda_5 + 4 = \lambda_3 - (5 - 3 \cdot \lambda_5) - 4 \cdot \lambda_5 + 4 = -1 + \lambda_3 - \lambda_5$. Einsetzen in die erste Zeile ergibt

$$\begin{aligned} \lambda_1 = 3 + 2\lambda_2 - \lambda_3 - 2\lambda_5 + 3 &= 3 + 2(-1 + \lambda_3 - \lambda_5) - \lambda_3 - 2\lambda_5 \\ &= 1 + \lambda_3 - 4\lambda_5. \end{aligned}$$

Ingesamt erhalten wir

$$\mathbb{L}(A, b) = \{ \begin{pmatrix} 1 \\ -1 \\ 0 \\ 5 \\ 0 \end{pmatrix} + \begin{pmatrix} \lambda_3 - 4\lambda_5 \\ \lambda_3 - \lambda_5 \\ \lambda_3 \\ -3\lambda_5 \\ \lambda_5. \end{pmatrix} \mid \lambda_3, \lambda_5 \in \mathbb{R} \}.$$

(3) Es sei $A = \begin{pmatrix} 1 & 0 & 2 & 3 \\ 0 & 1 & 4 & 5 \\ 0 & 0 & 1 & 6 \\ 0 & 0 & 0 & 1 \end{pmatrix}$ und $b = (4, 3, 2, 7)^T$. Es gilt $r_A = 4$ und $m = n = r_A$. Wir erhalten die erweitere Koeffizientenmatrix

$$(A|b) = \left(\begin{array}{cccc|c} 1 & 0 & 2 & 3 & 4 \\ 0 & 1 & 4 & 5 & 3 \\ 0 & 0 & 1 & 6 & 2 \\ 0 & 0 & 0 & 1 & 7 \end{array}\right).$$

Die Lösung ist eindeutig bestimmt. Sei diese mit $(x_1, x_2, x_3, x_4)^T$ bezeichnet. Dann folgt $1 \cdot x_4 = 7$, also $x_4 = 7$. Einsetzen in die dritte Zeile liefert: $x_3 + 6x_4 = 2$, das heißt $x_3 = 2 - 42 = -40$. Einsetzen in die zweite Zeile ergibt $x_2 = 3 + 160 - 35 = 128$ und $x_1 = 4 + 80 - 21 = 63$. Das heißt $\mathbb{L}(A, b) = \{(63, 128, -40, 7)^T\}$.

Wir definieren Rechenoperationen auf der Menge der Matrizen:

Definition 3.20 (Elementare Zeilenumformungen)**.** Folgende Matrixoperationen nennen wir *elementare Zeilenumformungen/operationen*:

(i) Vertauschen zweier Zeilen.

(ii) Multiplikation einer Zeile mit einer reellen Zahl λ ungleich 0.

(iii) Addition eines skalaren Vielfachen einer Zeilen zu einer anderen Zeile.

Satz 3.21. *Jede Matrix lässt sich durch eine endliche Anzahl von elementaren Zeilenumformungen in Zeilenstufenform bringen.*

Der Beweis des obigen Satzes ist konstruktiv, das heißt er liefert uns ein Verfahren/Algorithmus den wir anwenden können um eine beliebige Matrix in Zeilenstufenform zu überführen. Das Verfahren ist unter dem Namen *Gauß-Algorithmus* bekannt.

Es sei $A = (a_{ij}) \in \mathrm{Mat}_{m\times n}(\mathbb{R})$ eine Matrix ungleich null. Wir beschreiben den Algorithmus **Konvention**: Der Algorithmus verändert die Einträge der Matrix. Der Einfachheit halber bleiben wir beim Namen A.

(1) Wir suchen die erste Spalte (von links) ungleich 0. Sei diese im Folgenden mit $j \in \{1, \dots, n\}$ bezeichnet. Nach Vertauschung von Zeilen können wir erreichen, dass in der j-ten Spalte der oberste Eintrag $a_{1j} \neq 0$ erfüllt.

(2) Nun multiplizieren wir die erste Zeile mit $\frac{1}{a_{1j}}$. In der neuen Matrix gilt also $a_{1j} = 1$.

(3) Für $i \in \{2, \dots, m\}$ subtrahieren wir von der i-ten Zeile das a_{ij}-fache der ersten Zeile. Wir haben erreicht, dass in der j-ten Spalte $a_{1j} = 1$ gilt und $a_{ij} = 0$ für $i \in \{2, \dots, m\}$.

(4) Nun betrachten wir den Teil der Matrix, der sich unter der obersten Zeile und rechts von der betrachteten Spalte befindet: Wenn es keine solchen Zeilen oder Spalten gibt, ist das Verfahren beendet, andernfalls führen Sie dasselbe Verfahren für die neue Matrix durch.

Bemerkung 3.22. Der Gauß-Algorithmus liefert eine Zeilenstufenform, es gibt aber stets viele Wege wie man auf eine Zeilenstufenform gelangen kann und diese ist nicht eindeutig bestimmt (siehe folgende Beispiele).

Beispiel 3.23. (1) Die Zeilenstufenform ist selbst bei kleinen Matrizen nicht eindeutig: Die Matrix $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$ ist in Zeilenstufenform. Wenn wir von der ersten Zeile die zweite Zeile subtrahieren, bekommen wir mit $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ eine Matrix, die ebenfalls in Zeilenstufenform ist.

(2)

$$\begin{pmatrix} -6 & 12 & -6 & 0 \\ 2 & 1 & -3 & 5 \\ 4 & -7 & 1 & -1 \end{pmatrix} \overset{Z_1 \cdot \frac{1}{-6}}{\rightsquigarrow} \begin{pmatrix} 1 & -2 & 1 & 0 \\ 2 & 1 & -3 & 5 \\ 4 & -7 & 1 & -1 \end{pmatrix} \overset{Z_2 - 2\cdot Z_1}{\underset{Z_3 - 4\cdot Z_1}{\rightsquigarrow}} \begin{pmatrix} 1 & -2 & 1 & 0 \\ 0 & 5 & -5 & 5 \\ 0 & 1 & -3 & -1 \end{pmatrix}$$

Jetzt fahren wir mit Schritt (4) fort und führen das Verfahren für die Teilmatrix durch.

$$\begin{pmatrix} 1 & -2 & 1 & 0 \\ 0 & 5 & -5 & 5 \\ 0 & 1 & -3 & -1 \end{pmatrix} \overset{Z_2 \cdot \frac{1}{5}}{\rightsquigarrow} \begin{pmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -1 & 1 \\ 0 & 1 & -3 & -1 \end{pmatrix} \overset{}{\underset{Z_3 - 1\cdot Z_2}{\rightsquigarrow}} \begin{pmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & -2 & -2 \end{pmatrix}$$

Wir fahren mit Schritt (4) fort und erhalten eine Zeilenstufenform:

$$\begin{pmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & -2 & -2 \end{pmatrix} \underset{Z_3 \cdot \frac{1}{-2}}{\rightsquigarrow} \begin{pmatrix} 1 & -2 & 1 & 0 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix}.$$

Satz 3.24. *Es sei $A, A' \in \mathrm{Mat}_{m\times n}(\mathbb{R})$ und $b, b' \in \mathbb{R}^n$. Entsteht die Matrix $(A'|b') \in \mathrm{Mat}_{m\times(n+1)}(\mathbb{R})$ aus $(A|b) \in \mathrm{Mat}_{m\times(n+1)}(\mathbb{R})$ durch elementare Zeilenumformungen, so gilt $\mathbb{L}(A,b) = \mathbb{L}(A',b')$.*

Jetzt haben wir ein systematisches Verfahren um ein LGS zu lösen: Ist uns ein LGS mit m Gleichungen in n Unbestimmten gegeben, so stellen wir die zugehörige Koeffizientenmatrix $(A|b) \in \mathrm{Mat}_{m\times(n+1)}(\mathbb{R})$ auf. Anschließend überführen wir $(A|b)$ durch elementare Zeilenumformungen in Zeilenstufenform. Das erhaltene System sei mit $(A'|b')$ bezeichnet. Dann gilt nach Satz 3.24

$$\mathbb{L}(A,b) = \mathbb{L}(A',b').$$

Wir bemerken, dass mit $(A'|b')$ auch $A' \in \mathrm{Mat}_{m\times n}(\mathbb{R})$ in Zeilenstufenform ist. Nun wenden wir Satz 3.18 an um $\mathbb{L}(A',b')$ zu berechnen. Zur Veranschaulichung kehren wir zu Beispiel 3.2 zurück:

Beispiel 3.25. Wir waren bei dem LGS

$$\begin{aligned} 10a + 9b + c &= -181 \\ 10a + 3b + c &= -109 \\ 9a + 2b + c &= -85. \end{aligned}$$

stehen geblieben. Wir betrachten die erweitere Koeffizientenmatrix $(A|b) := \left(\begin{array}{ccc|c} 10 & 9 & 1 & -181 \\ 10 & 3 & 1 & -109 \\ 9 & 2 & 1 & -85 \end{array}\right)$.
Wir überführen die Matrix durch elementare Zeilenumformungen in Zeilenstufenform und erhalten

$$(A'|b') := \left(\begin{array}{ccc|c} 1 & \frac{9}{10} & \frac{1}{10} & -18.1 \\ 0 & 1 & 0 & -12 \\ 0 & 0 & 1 & 47 \end{array}\right).$$

Das heißt $\mathbb{L}(A,b) = \mathbb{L}(A',b') = \{(-12,-12,47)^T\}$. Damit ergibt sich $x_M = 6 = y_M$ und $r = 5$. Das heißt, die Studierenden müssen sich am Punkt $(6,6)$ treffen und alle müssen genau 500 Meter gehen.

Bemerkung 3.26. Es hängt stark an den Koordinaten der Wohnungen A, B und C, welche Treffpunkte sich ergeben:

- Gilt $A = B = C$ so gibt es unendlich viele Kreise auf denen A, B und C liegen.
- Gilt z.B. $A = (0,0), B = (1,1), C = (2,2)$, so existiert kein Kreis auf dem A, B und C liegen.

3.4 Lineare Abbildungen und Matrizen

Nach Satz 3.18 und Satz 3.24 kann für ein LGS mit m Zeilen und n Unbekannten nur dann eine eindeutige Lösung existieren, wenn $m = n$ gilt. Im Folgenden untersuchen wir wie man entscheiden kann, ob eine Lösung eindeutig bestimmt ist.

Bisher würden wir wie folgt verfahren: Wir stellen die zugehörige Koeffizientenmatrix $(A|b) \in \mathrm{Mat}_{n\times(n+1)}(\mathbb{R})$ auf und überführen diese durch elementare Zeilenumformungen in Zeilenstufenform. Sei diese mit $(A'|b') \in \mathrm{Mat}_{n\times(n+1)}(\mathbb{R})$ bezeichnet. Dann gilt nach Satz 3.24

$$\mathbb{L}(A, b) = \mathbb{L}(A', b').$$

Nach Satz 3.18 existiert genau dann eine eindeutige Lösung für $\mathbb{L}(A', b')$, $A' \in \mathrm{Mat}_n(\mathbb{R})$ wenn, $r_{A'} = n$ gilt, das heißt jeder Diagonaleintrag von A' ist $1 \in \mathbb{R}$. In diesem Abschnitt werden wir eine weniger aufwändiges Verfahren kennen lernen.

Definition 3.27. Eine Matrix $A \in \mathrm{Mat}_n(\mathbb{R}) = \mathrm{Mat}_{n\times n}(\mathbb{R})$ heißt *invertierbar*, falls eine Matrix $A^{-1} \in \mathrm{Mat}_n(\mathbb{R})$ mit $A^{-1} \cdot A = I_n = A \cdot A^{-1}$ existiert, wobei

$$I_n := \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \\ 0 & 0 & 0 & \cdots & 1 \end{pmatrix} \in \mathrm{Mat}_n(\mathbb{R}).$$

In diesem Fall ist A^{-1} eindeutig bestimmt, da $B \cdot I_n = B = I_n \cdot B$ für alle $B \in \mathrm{Mat}_n(\mathbb{R})$ gilt.

Beispiel 3.28. (1) Eine (1×1)-Matrix $A = (a)$ ist genau dann invertierbar, wenn $a \neq 0$ gilt. In dem Fall ist $A^{-1} = (\frac{1}{a})$.

(2) Die (2×2)-Matrix $A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ ist nicht invertierbar: Es sei $B = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathrm{Mat}_2(\mathbb{R})$. Dann gilt $A \cdot B = \begin{pmatrix} a & b \\ 0 & 0 \end{pmatrix} \neq I_2$.

Satz 3.29. *Es sei $A \in \mathrm{Mat}_n(\mathbb{R})$ und $b \in \mathbb{R}^n$. Die Lösungmenge $\mathbb{L}(A, b)$ ist genau dann einelementig, wenn A invertierbar ist. In diesem Fall gilt $\mathbb{L}(A, b) = \{A^{-1} \cdot b\}$.*

Ein wichtiges Hilfmittel, um eine quadratische Matrix auf Invertierbarkeit zu prüfen, ist die sogenannte Determinate.

Definition 3.30. Es sei $n \in \mathbb{N}$. Für $A \in \mathrm{Mat}_n(\mathbb{R})$ ist $\det(A)$ eine reelle Zahl, die wir rekursiv definieren:

$n = 1$: Es sei $A = (a) \in \mathrm{Mat}_1(\mathbb{R})$. Wir definieren $\det(A) := a$.

$n > 1$: Es sei $A = (a_{ij}) \in \mathrm{Mat}_n(\mathbb{R})$. Für $i, j \in \{1, \dots, n\}$ sei $A_{ij} \in \mathrm{Mat}_{n-1}(\mathbb{R})$ die Matrix, die aus A durch das Streichen der i-ten Zeile und j-ten Spalte entsteht. Wir definieren

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

Wir nennen $\det(A)$ *Determinate von* A.

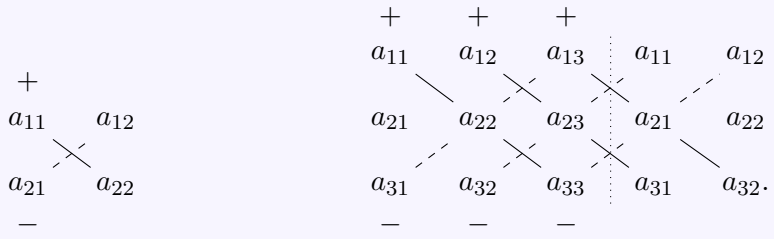
Beispiel 3.31. (1) Es sei $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$. Dann gilt

$$\det(A) = (-1)^{1+1} a_{11} \det(A_{11}) + (-1)^{1+2} a_{12} \det(A_{12}) = a_{11}a_{22} - a_{12}a_{21}.$$

(2) Es sei $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$. Dann gilt

$$\begin{aligned} \det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} =& (-1)^2 a_{11} \det \begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} + (-1)^3 a_{12} \det \begin{pmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{pmatrix} \\ &+ (-1)^4 a_{13} \det \begin{pmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix} \\ =& a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} \\ &- (a_{31}a_{22}a_{13} + a_{32}a_{23}a_{11} + a_{33}a_{21}a_{12}). \end{aligned}$$

Beide Formeln lassen sich gut veranschaulichen:



Satz 3.32 (Laplascher Entwicklungssatz)**.** *Es sei* $n \in \mathbb{N}$ *und* $A \in \mathrm{Mat}_n(\mathbb{R})$ *eine Matrix. Für jedes* $i \in \{1, \dots, n\}$ *gilt*

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

Bemerkung 3.33. Der Laplasche Entwicklungssatz ist insbesondere dann hilfreich, wenn eine Zeile/Spalte vielen Nulleinträge enthält. Die linke Formel wird auch *Entwicklung nach der* i*-ten Zeile* genannt und die rechte Formel *Entwicklung nach der* i*-ten Spalte*.

Beispiel 3.34. Es sei $A = \begin{pmatrix} 1 & 2 & 0 & 0 \\ 1 & 3 & 0 & 0 \\ 1 & 2 & 2 & 4 \\ 1 & 2 & 0 & 4 \end{pmatrix}$. Wir entwickeln nach der 3-ten Spalte und erhalten

$$\begin{aligned} \det(A) &= \sum_{j=1}^{4} (-1)^{j+3} a_{j3} \det(A_{j3}) \\ &= (-1)^4 \cdot a_{13} \det(A_{13}) + (-1)^5 a_{23} \det(A_{23}) + (-1)^6 a_{33} \det(A_{33}) + (-1)^7 a_{43} \det(A_{43}) \\ &= (-1)^6 a_{33} \det(A_{33}) = 2 \det(A_{33}) = 2 \det \begin{pmatrix} 1 & 2 & 0 \\ 1 & 3 & 0 \\ 1 & 2 & 4 \end{pmatrix}. \end{aligned}$$

Wir entwickeln diese (3×3)-Matrix nach der dritten Spalte und erhalten

$$\det(A) = 2 \cdot (-1)^6 \cdot 4 \det \begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix} = 2 \cdot 4 \cdot (3 - 2) = 8.$$

Satz 3.35. *Es sei $n \in \mathbb{N}$ und $A \in \mathrm{Mat}_n(\mathbb{R})$. Folgende Aussagen sind äquivalent:*

(1) A ist invertierbar.

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

(3) Für ein $b \in \mathbb{R}^n$ ist $\mathbb{L}(A, b)$ einelementig.

(4) Für jedes $b \in \mathbb{R}^n$ ist $\mathbb{L}(A, b)$ einelementig.

Beispiel 3.36. Wir betrachten erneut Beispiel 3.2. Es gilt

$$\det \begin{pmatrix} 10 & 9 & 1 \\ 10 & 3 & 1 \\ 9 & 2 & 1 \end{pmatrix} = 30 + 81 + 20 - (27 + 20 + 90) = -6 \neq 0.$$

Also existiert eine eindeutige Lösung des LGS.

Definition 3.37. Es seien $n, m \in \mathbb{N}$. Eine Abbildung $f \colon \mathbb{R}^n \to \mathbb{R}^m$ heißt *linear* falls

(1) $f(x + y) = f(x) + f(y)$ für alle $x, y \in \mathbb{R}^n$, und

(2) $f(\lambda \cdot x) = \lambda \cdot f(x)$ für alle $x \in \mathbb{R}^n$ und $\lambda \in \mathbb{R}$ gilt.

Bemerkung 3.38. Es sei f linear, $k \in \mathbb{N}$, $x_1, \ldots, x_k \in \mathbb{R}^n$ und $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$. Dann gilt

$$f(\sum_{i=1}^{k} \lambda_i x_i) = \sum_{i=1}^{k} \lambda_i f(x_i).$$

Beispiel 3.39. (1) Die Abbildung $f\colon \mathbb{R}^2 \to \mathbb{R}^3; (x_1, x_2)^T \mapsto (3x_1 - x_2, 5x_2, 7x_1)^T$ ist linear: Für $x, y \in \mathbb{R}^2$ und $\lambda \in \mathbb{R}$ gilt

$$\begin{aligned} f(x+y) &= (3(x_1+y_1) - (x_1+y_1), 5(x_2+y_2), 7(x_1+y_1))^T \\ &= (3x_1 - x_2, 5x_2, 7x_1)^T + (3y_1 - y_2, 5y_2, 7y_1)^T = f(x) + f(y) \end{aligned}$$

und

$$f(\lambda \cdot x) = (3(\lambda x_1) - \lambda \cdot x_2, 5(\lambda x_2), 7(\lambda x_1))^T = \lambda(3x_1 - x_2, 5x_2, 7x_1)^T = \lambda \cdot f(x).$$

(2) Die Abbildung $f\colon \mathbb{R} \to \mathbb{R}^2; x \mapsto (x, 5)$ ist nicht linear, da

$$f(1+1) = f(2) = (2,5) \neq (2,10) = (1,5) + (1,5) = f(1) + f(1).$$

(3) Es sei $f\colon \mathbb{R}^n \to \mathbb{R}^m$ linear. Dann gilt $f(0_{\mathbb{R}^n}) = f(0_{\mathbb{R}^n} + 0_{\mathbb{R}^n}) = f(0_{\mathbb{R}^n}) + f(0_{\mathbb{R}^n})$, also $f(0_{\mathbb{R}^n}) = 0_{\mathbb{R}^m}$.

Es sei $j \in \{1, \dots, n\}$. Wir bezeichnen mit $e_j \in \mathbb{R}^n$ definiert durch

$$\overset{j\text{-te Stelle}}{\downarrow}$$
$$e_j := (0, 0, \dots, 1, 0 \dots 0)^T \qquad 1 \leq j \leq n$$

den j-ten *Standardbasisvektor*. Es sei $a = (a_1, \dots, a_n)^T$, dann lässt sich a in eindeutiger Weise durch die Basisvektoren $e_1, \dots, e_n$ schreiben. Es gilt nämlich

$$a = \sum_{i=1}^{n} a_i e_i.$$

Satz 3.40. *(1) Es sei A eine $(m \times n)$-Matrix. Dann ist $L_A\colon \mathbb{R}^n \to \mathbb{R}^m; x \mapsto A \cdot x$ eine lineare Abbildung.*

(2) Ist $e_i \in \mathbb{R}^n$ der i-te Einheitsvektor und $A = (a_{ij}) \in \mathrm{Mat}_{m\times n}(\mathbb{R})$, so ist $L_A(e_i) = A \cdot x = {}_iA$ die i-te Spalte von A.

(3) Zu jeder linearen Abbildung $f\colon \mathbb{R}^n \to \mathbb{R}^m$ existiert genau eine $(m \times n)$-Matrix A mit $f = L_A$.

Beweis.

(1) Es seien $x, y \in \mathbb{R}^n$ und $\lambda \in \mathbb{R}$. Es gilt aufgrund von Bemerkung 3.14

$$L_A(x+y) = A \cdot (x+y) = A \cdot x + A \cdot y = L_A(x) + L_A(y)$$

und

$$L_A(\lambda \cdot x) = A \cdot (\lambda \cdot x) = \lambda \cdot (A \cdot x) = \lambda L_A(x).$$

(2) Folgt durch direktes Nachrechnen.

(3) Wir betrachten die Einheitsvektoren $e_1, \dots, e_n \in \mathbb{R}^n$. Für $j \in \{1, \dots, n\}$ ist $f(e_i) \in \mathbb{R}^m$ ein Vektor mit m Einträgen. Wir definieren eine Matrix $A \in \mathrm{Mat}_{m\times n}(\mathbb{R})$ deren i-te Spalte durch $f(e_i)$ gegeben ist. Es sei nun $a \in \mathbb{R}^n$ beliebig. Wir schreiben

$$x = (a_1, \dots, a_n)^T = \sum_{i=1}^{n} a_i e_i.$$

Aufgrund der Linearität erhalten wir

$$L_A(x) = \sum_{i=1}^{n} a_i L_A(e_i) = \sum_{i=1}^{n} a_i f(e_i) = f(\sum_{i=1}^{n} a_i e_i) = f(x).$$

Ist B eine weitere Matrix mit $L_A = f = L_B$, so erhalten wir insbesondere ${}_iA = L_A(e_i) = L_B(e_i) = {}_iB$ für alle $i \in \{1, \dots, n\}$. Das heißt alle Spalten von A und B stimmen überein und es folgt $A = B$.

□

Bemerkung 3.41. (1) Der Satz besagt, dass die Abbildung

$$\mathrm{Mat}_{m\times n}(\mathbb{R}) \to \{g \colon \mathbb{R}^n \to \mathbb{R}^m \mid g \text{ lineare Abbildung}\}; A \mapsto L_A$$

eine Bijektion ist.

(2) Es sei f eine lineare Abbildung. Nach Satz 3.40 existiert genau eine Matrix A mit $L_A = f$. Damit legen schon die endlich vielen Einträge von A die lineare Abbildung f mit unendlich vielen Funktionswerten (für $f \neq 0$) eindeutig fest.

Beispiel 3.42. Die lineare Abbildung $f \colon \mathbb{R}^2 \to \mathbb{R}^3; (x_1, x_2)^T \mapsto (3x_1 - x_2, 5x_2, 7x_1)^T$ wir durch die Matrix $A = \begin{pmatrix} 3 & -1 \\ 0 & 5 \\ 7 & 0 \end{pmatrix} \in \mathrm{Mat}_{3\times 2}(\mathbb{R})$ beschrieben, das heißt $f = L_A$.

Definition 3.43. Es seien $m \in \mathbb{N}$ und $x_1, \dots, x_m \in \mathbb{R}^n$.

(1) Die Vektoren $x_1, \dots, x_m$ heißen *linear unabhängig*, falls für alle $\lambda_1, \dots, \lambda_m \in \mathbb{R}$ mit

$$\sum_{i=1}^{m} \lambda_i \cdot x_i = 0_{\mathbb{R}^n}$$

schon $\lambda_1 = \lambda_2 = \cdots = \lambda_m = 0_{\mathbb{R}}$ gilt.

(2) Die Menge $\{x_1, \dots, x_m\}$ heißt *Basis*, falls sich jedes Element $x \in \mathbb{R}^n$ in eindeutiger Weise als Linearkombination in $x_1, \dots, x_m$ schreiben lässt. Das heißt es existieren eindeutig bestimmte $\lambda_1, \dots, \lambda_m \in \mathbb{R}$ mit $\sum_{i=1}^{m} \lambda_i \cdot x_i = x$.[4]

[4] Insbesondere sind $x_1, \dots, x_m$ linear unabhängig, da $0_{\mathbb{R}^n} = \sum_{i=1}^{m} 0_{\mathbb{R}} \cdot x_i$ eine Darstellung von $0_{\mathbb{R}^n}$ ist.

Satz 3.44. *Es seien $m \in \mathbb{N}$ und $x_1, \dots, x_m \in \mathbb{R}^n$ Vektoren. Es gelten die folgenden Aussagen*

(1) Sind $x_1, \dots, x_m$ linear unabhängig, so gilt $m \leq n$.

(2) Ist $\{x_1, \dots, x_m\}$ eine Basis, so gilt $m = n$.

(3) Sind $x_1, \dots, x_m$ linear unabhängig und $m = n$, so ist $\{x_1, \dots, x_m\}$ eine Basis.

Satz 3.45. *Es sei $n \in \mathbb{N}$ und $A \in \mathrm{Mat}_n(\mathbb{R})$. Folgende Aussagen sind äquivalent.*

(1) A ist invertierbar.

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

(3) Für ein $b \in \mathbb{R}^n$ ist $\mathbb{L}(A, b)$ einelementig.

(4) Für jedes $b \in \mathbb{R}^n$ ist $\mathbb{L}(A, b)$ einelementig.

(5) Die Abbildung $L_A \colon \mathbb{R}^n \to \mathbb{R}^n; x \mapsto Ax$ ist injektiv.

(6) Die Abbildung $L_A \colon \mathbb{R}^n \to \mathbb{R}^n; x \mapsto Ax$ ist surjektiv.

(7) Die Zeilen (bzw. Spalten) von A sind eine Basis von $\mathbb{R}^n$.

Literaturverzeichnis

[1] Ina Kersten, *Mathematische Grundlagen in Biologie und Geowissenschaften*, 2004.

[2] Alexander Ullmann, *Vorkurs Mathematik für Studierende der Geowissenschaften*, 2023.