Höhere Mathematik I für die Fachrichtung Elektrotechnik und Informationstechnik Wintersemester 2016/17

erstellt von: Peer Christian Kunstmann
Karlsruher Institut für Technologie
Institut für Analysis
Englerstr. 2, 76131 Karlsruhe
e-mail: peer.kunstmann@kit.edu

gehalten von: Wolfgang Reichel
Karlsruher Institut für Technologie
Institut für Analysis
Englerstr. 2, 76131 Karlsruhe
e-mail: wolfgang.reichel@kit.edu

Dies ist eine Vorlesungszusammenfassung, gedacht zur Vorlesungsbegleitung und als Gedächtnisstütze, nicht jedoch als etwas, das für sich selbst stehen könnte (wie etwa ein Lehrbuch). Der Besuch der Vorlesung ist durch die Lektüre in keinem Fall zu ersetzen, es gibt dort noch viel mehr an mündlichen Erklärungen, Erläuterungen und veranschaulichenden Skizzen, die für Verständnis und Einordnung des präsentierten Stoffes unabdingbar sind.

1 Logische Grundlagen

Wir beginnen mit Grundlagen der *Aussagenlogik*. Dafür gibt es mehrere Gründe:

- Mathematische Sätze sind Aussagen (zB "Jede natürliche Zahl läßt sich als Produkt von Primfaktoren schreiben"). Auch wenn man in der Elektrotechnik vor allem rechnen will, so verwendet man dazu mathematische Aussagen und muss diese also verstehen und damit umgehen können.
- Die Verknüpfungen von Aussagen sind dieselben, die man in der Elektrotechnik im Rahmen von logischen Schaltungen (zB zur Steuerung) verwendet.
- Aussagenlogische Verknüpfungen für die Wahrheitswerte "wahr" und "falsch" spielen eine wichtige Rolle in Programmiersprachen.
- Schließlich dient die Beschftigung damit auch dazu, Ordnung in seine Gedanken zu bringen, was generell nicht schadet und häufig äußerst nützlich sein kann.

1.1. Aussagen: Eine *Aussage* ist ein Satz, der entweder **wahr (w)** oder **falsch (f)** ist.

Beispiele:

(1) Der Mond ist ein grüner Käse. (f)
(2) 2 ist eine Primzahl. (w)
(3) Gehen sie geradeaus und dann hinten links! (keine Aussage)

Die Verknüpfungen in 1.2 und die Regeln in 1.3 sind diejenigen aus der *Booleschen Algebra*, die auch bei *logischen Schaltungen* in der Elektrotechnik verwendet werden. Statt "wahr/falsch" verwendet man dann auch "1/0", "H(igh)/L(ow)" oder die Interpretation "Signal/kein Signal".

Im Rahmen der Vorlesung sind wir aber eher an **mathematischen Aussagen** interessiert. Aussagen bezeichnen wir im folgenden mit A, B, C, ...

1.2. Verknüpfung von Aussagen: Wir erklären die logische Verknüpfung von Aussagen durch sogenannte *Wahrheitstafeln*.

(a)

$A \wedge B$ (*logisches "und"* (AND))

$$\begin{array}{r|cccc} A & w & w & f & f \\ B & w & f & w & f \\ \hline A \wedge B & w & f & f & f \end{array}$$

(b)

$A \vee B$ (*logisches "oder"* (OR))

$$\begin{array}{r|cccc} A & w & w & f & f \\ B & w & f & w & f \\ \hline A \vee B & w & w & w & f \end{array}$$

Achtung: Das logische “oder” ist nicht exklusiv, dh es ist zugelassen, dass *beide* Aussagen A und B wahr sind.

(c)

Negation $\neg A$

A	w	f
$\neg A$	f	w

Man sagt: “non A” oder “nicht A”.

(d)

Implikation $A \Rightarrow B$

A	w	w	f	f
B	w	f	w	f
$A \Rightarrow B$	w	f	w	w

Man sagt: “wenn A, dann B”, “A impliziert B”, “aus A folgt B”.

Bemerkung: Aus Falschem folgt Beliebiges (*ex falso quodlibet*).

Beispiel: Aus $1 = 2$ folgt: ich bin der Papst. (w)

(e)

Äquivalenz $A \Leftrightarrow B$

A	w	w	f	f
B	w	f	w	f
$A \Leftrightarrow B$	w	f	f	w

Man sagt: “A ist äquivalent zu B”, “A ist gleichbedeutend mit B”, “A genau dann, wenn B”, “ A dann und nur dann, wenn B”.

1.3. Regeln: Für zusammengesetzte Aussagen gilt die Konvention:

$$\neg \text{ bindet stärker als } \wedge/\vee;\ \wedge/\vee \text{ bindet stärker als } \Rightarrow/\Leftrightarrow.$$

Folgende Regeln sind allgemeingültig:

$$\begin{aligned}
\neg(\neg A) &\Leftrightarrow A \quad \text{(doppelte Negation)}\\
(A \Leftrightarrow B) &\Leftrightarrow [(A \Rightarrow B) \wedge (B \Rightarrow A)] \quad \text{(Äquivalenz bedeutet zwei Implikationen)}\\
\neg(A \wedge B) &\Leftrightarrow (\neg A \vee \neg B) \quad \text{(Negation von “und”)}\\
\neg(A \vee B) &\Leftrightarrow (\neg A \wedge \neg B) \quad \text{(Negation von “oder”)}\\
(A \Rightarrow B) &\Leftrightarrow (\neg A \vee B) \quad \text{(Umformulierung der Implikation)}\\
\neg(A \Rightarrow B) &\Leftrightarrow (A \wedge \neg B) \quad \text{(Negation der Implikation)}\\
(A \Rightarrow B) &\Leftrightarrow (\neg B \Rightarrow \neg A) \quad (\textit{Kontraposition}).
\end{aligned}$$

Beispiele: Wenn ich nicht der Papst bin, dann ist $1 \neq 2$. (w)

Sei A: “es regnet” und B: “die Straße ist nass”. Dann ist $A \Rightarrow B$ wahr (“wenn es regnet, dann ist die Straße nass”) und gleichbedeutend (äquivalent) zu $\neg B \Rightarrow \neg A$ (“wenn die Straße trocken (dh nicht nass) ist, dann regnet es nicht”). Hingegen ist die Aussage “Wenn es nicht regnet, dann ist die Straße nicht nass" (dh also $\neg A \Rightarrow \neg B$) ja falsch, da die Straße zB aus anderen Gründen nass sein kann. Die Kontraposition ist also die korrekte Form des oft anzutreffenden "Umkehrschlusses", der jedoch häufig nicht richtig gebildet wird. Zur Übung können Sie in Alltagssituationen darauf achten.

Weiter gelten folgende Regeln:

> *Kommutativität*: $A \wedge B \Leftrightarrow B \wedge A$ und $A \vee B \Leftrightarrow B \vee A$.
>
> *Assoziativität*: $A \wedge (B \wedge C) \Leftrightarrow (A \wedge B) \wedge C$ und $A \vee (B \vee C) \Leftrightarrow (A \vee B) \vee C$. Deshalb kann man hier die Klammern weglassen.
>
> *Distributivität*: $A \wedge (B \vee C) \Leftrightarrow (A \wedge B) \vee (A \wedge C)$
> und $A \vee (B \wedge C) \Leftrightarrow (A \vee B) \wedge (A \vee C)$.

Tertium non datur: $A \vee \neg A$ ("ein Drittes gibt es nicht").

Beispiel: Eine Verknüpfung, welche das "exklusive Oder" (XOR)

$$\begin{array}{c|cccc} A & w & w & f & f \\ B & w & f & w & f \\ \hline & f & w & w & f \end{array}$$

realisiert, ist z.B. $(\neg A \wedge B) \vee (A \wedge \neg B)$, aber auch $(A \vee B) \wedge \neg(A \wedge B)$ oder $\neg(A \Leftrightarrow B)$. Mit obigen Regeln können diese drei Versionen ineinander umgeformt werden (Übung).

In der Mathematik hat man in der Regel mit Aussagen über gewisse Klassen von Objekten zu tun (zB "für alle natürlichen Zahlen gilt ...", s.o.).

1.4. Quantoren: Eine *Aussageform* $A(x)$, $A(x, y)$, ... ist ein Satz, der eine oder mehrere Variablen x, y, ... enthält und der nach dem Ersetzen dieser Variablen durch konkrete Objekte eine Aussage ist.

Beispiel: x ist eine Primzahl.

Der *Allquantor*

$$\forall x : A(x)$$

bedeutet: für alle Objekte x ist die Aussage $A(x)$ wahr.

Der *Existenzquantor*

$$\exists x : A(x)$$

bedeutet: es gibt (mindestens) ein Objekt x, für das die Aussage $A(x)$ wahr ist.

Negation von Quantoren:

$$\neg(\forall x : A(x)) \Leftrightarrow (\exists x : \neg A(x)) \quad \text{und} \quad \neg(\exists x : A(x)) \Leftrightarrow (\forall x : \neg A(x)).$$

Die Regeln sind für einen Quantor unmittelbar plausibel. Bei mehreren käme es ohne diese Regeln jedoch leicht zu Verwirrung.

In den allermeisten Fällen werden Quantoren *eingeschränkt* und beziehen sich dann nur auf *gewisse* Objekte, z.B.

$$\forall x \text{ mit } A(x) : B(x).$$

Die Negation davon ist dann

$$\exists x \text{ mit } A(x) : \neg B(x).$$

Beispiel: Alle Primzahlen sind ungerade. (f)
Setzen wir $A(x)$: "x ist Primzahl" und $B(x)$: "x ist ungerade", so haben wir die Aussage $\forall x$ mit $A(x)$: $B(x)$. Deren Negation ist $\exists x$ mit $A(x)$: $\neg B(x)$, also "Es gibt eine Primzahl, die nicht ungerade ist" (w).

Achtung: Bei All- und Existenzquantor kommt es i.a. auf die Reihenfolge an! Betrachtet man für x mit Frauen verheiratete Männer und für y mit Männern verheiratete Frauen sowie die Aussageform $A(x, y)$: "x ist verheiratet mit y", so ist $\forall x\, \exists y : A(x, y)$ wahr, aber $\exists y\, \forall x\ : A(x, y)$ ist offensichtlich falsch.

Bemerkung: Häufig schreiben wir Quantoren nicht als Zeichen, sondern sprachlich.

2 Mengen

2.1. Der Begriff der Menge: Wir verwenden die folgende naive "Definition":

> "Eine *Menge* ist die Zusammenfassung wohlbestimmter, wohlunterschiedener Objekte der Anschauung oder des Denkens zu einem neuen Ganzen."

Diese Objekte heißen *Elemente* der Menge.

$x \in M$ bedeutet: x ist Element der Menge M, dh x gehört zu M.

$x \notin M$ bedeutet: x gehört nicht zu M, dh $x \notin M :\Leftrightarrow \neg(x \in M)$.

"$x \in M$" ist also eine Aussageform, dh für jede Menge M und jedes x gilt entweder $x \in M$ oder $x \notin M$.

Die Mengen, mit denen wir uns beschäftigen werden, sind etwa Mengen von Zahlen oder Mengen von Funktionen etc.

Schreibweisen: Ist $A(x)$ eine Aussageform, so kann man schreiben

$$M = \{x : A(x)\} = \text{Menge aller } x, \text{ für die } A(x) \text{ gilt,}$$

also z.B. $N = \{x : x \in N\}$, $P = \{x : x \text{ ist Primzahl}\}$. Häufig schreibt man auch, wenn die Menge M gegeben ist, $\{x \in M : A(x)\}$ für $\{x : x \in M \wedge A(x)\}$.

Eine andere Möglichkeit ist die Aufzählung, etwa $M = \{1, 2, 3, 9\}$.

2.2. Beziehungen zwischen Mengen: Seien M_1, M_2 Mengen.

Definition: "M_1 ist Teilmenge von M_2":

$$M_1 \subseteq M_2 \;:\Leftrightarrow\; \forall x : (x \in M_1 \Rightarrow x \in M_2) \quad (\text{bzw. } \forall x \in M_1 : x \in M_2).$$

Für $M_1 \subseteq M_2$ und $M_1 \neq M_2$ schreibe ich der Deutlichkeit halber $M_1 \subsetneqq M_2$.

Gleichheit von Mengen: $M_1 = M_2$ bedeutet, dass M_1 und M_2 dieselben Elemente enthalten, also $\forall x : (x \in M_1 \Leftrightarrow x \in M_2)$. Nach 1.3 bedeutet $M_1 = M_2$ also $M_1 \subseteq M_2$ und $M_2 \subseteq M_1$.

Beispiele: Wir schreiben $\mathbb{N} = \{1, 2, 3, \ldots\}$ *Menge der natürlichen Zahlen.*
$\{x \in \mathbb{N} : x \text{ ist gerade Primzahl}\} = \{2\}$,
$\{x \in \mathbb{N} : x \text{ ist Primzahl} \wedge x > 2\} \subseteq \{x \in \mathbb{N} : x \text{ ist ungerade}\}$.

2.3. Operationen mit Mengen: Seien M_1, M_2, M_3 und Q Mengen.

(a) *Durchschnitt* $M_1 \cap M_2 := \{x : x \in M_1 \wedge x \in M_2\}$.

(b) *Vereinigung* $M_1 \cup M_2 := \{x : x \in M_1 \vee x \in M_2\}$.

Regeln für Durchschnitt und Vereinigung: Wegen 1.3 gelten:

Kommutativität:

$$M_1 \cap M_2 = M_2 \cap M_1, \qquad M_1 \cup M_2 = M_2 \cup M_1.$$

Assoziativität:

$$M_1 \cap (M_2 \cap M_3) = (M_1 \cap M_2) \cap M_3, \qquad M_1 \cup (M_2 \cup M_3) = (M_1 \cup M_2) \cup M_3.$$

Distributivität:

$$M_1 \cup (M_2 \cap M_3) = (M_1 \cup M_2) \cap (M_1 \cup M_3), \quad M_1 \cap (M_2 \cup M_3) = (M_1 \cap M_2) \cup (M_1 \cap M_3).$$

Außerdem ist $M_1 \subseteq M_1 \cup M_2$, $M_2 \subseteq M_1 \cup M_2$, $M_1 \cap M_2 \subseteq M_1$, $M_1 \cap M_2 \subseteq M_2$.

(c) *Differenz* $M_1 \setminus M_2 := \{x \in M_1 : x \notin M_2\}$ ("M_1 ohne M_2").

Beispiel: $\{x \in \mathbb{N} : x \text{ ist Primzahl}\} \setminus \{x \in \mathbb{N} : x \text{ ist ungerade}\} = \{2\}$.

de Morgansche Regeln: Wegen 1.3 (Negation von "und"/"oder") gilt auch

$$Q \setminus (M_1 \cup M_2) = (Q \setminus M_1) \cap (Q \setminus M_2), \qquad Q \setminus (M_1 \cap M_2) = (Q \setminus M_1) \cup (Q \setminus M_2).$$

2.4. Die leere Menge: Die *leere Menge* $\emptyset$ enthält keine Elemente, dh $\forall x : x \notin \emptyset$.

Regeln: $M \cup \emptyset = M$, $M \setminus \emptyset = M$, $M \cap \emptyset = \emptyset$, $M \setminus M = \emptyset$, $\emptyset \subseteq M$ für jede Menge M.

2.5. Die Potenzmenge: Ist M eine Menge, so heißt die Menge aller Teilmengen von M

$$\mathrm{P}(M) := \{N : N \subseteq M\}$$

die *Potenzmenge von* M (manchmal auch $\mathfrak{P}(M)$).

Bemerkung: Für jede Menge M gilt: $M \in \mathrm{P}(M)$, $\emptyset \in \mathrm{P}(M)$, aber auch $\emptyset \subseteq \mathrm{P}(M)$ (vgl. 2.4).

Beispiel: Für $M = \{1, 2\}$ ist $\mathrm{P}(M) = \{\emptyset, \{1\}, \{2\}, \{1, 2\}\}$.

2.6. Das kartesische Produkt: Sei $n \in \mathbb{N}$. Seien $M_1, M_2, \ldots, M_n$ Mengen. Die Menge der geordneten n-Tupel $(x_1, x_2, \ldots, x_n)$ mit $x_j \in M_j$ für alle $j \in \{1, 2, \ldots, n\}$ heißt das *kartesische Produkt* $M_1 \times M_2 \times \ldots \times M_n$ der Mengen $M_1, M_2, \ldots, M_n$. Also

$$M_1 \times M_2 \times \ldots \times M_n = \{(x_1, x_2, \ldots, x_n) : \text{für alle } j \in \{1, 2, \ldots, n\} \text{ gilt } x_j \in M_j\}.$$

Wir schreiben M^n, falls $M_1 = M_2 = \ldots = M_n = M$ gilt.

Beispiele:

(1) $M_1 = \{0,1\}$, $M_2 = \{1,2,3\}$, $M_1 \times M_2 = \{(0,1),(0,2),(0,3),(1,1),(1,2),(1,3)\}$.

(2) $M = \{0,1\}$, $M^3 = \{(0,0,0),(0,0,1),(0,1,0),(0,1,1),(1,0,0),(1,0,1),(1,1,0),(1,1,1)\}$.

Besonders wichtig ist natürlich der Fall $n = 2$, in dem $M_1 \times M_2$ die Menge aller *geordneten Paare* (x_1, x_2) mit $x_1 \in M_1$, $x_2 \in M_2$ ist.

3 Funktionen

3.1. Zum Begriff der Funktion: Seien X, Y Mengen. Eine *Funktion* (oder *Abbildung*) $f : X \to Y$ ordnet *jedem* $x \in X$ *genau ein* $y \in Y$ zu. Für das einem gegebenen $x \in X$ zugeordnete $y \in Y$ schreiben wir $f(x)$.

Schreibweise $f : X \to Y, x \mapsto f(x)$ ("f von X nach Y, x wird abgebildet auf $f(x)$").

Beispiel: $f : \mathbb{N} \to \mathbb{N}$, $x \mapsto 2 \cdot x - 1$, dann ist etwa $f(1) = 2 \cdot 1 - 1 = 1$, $f(2) = 2 \cdot 2 - 1 = 3$ etc.

Die Menge $\{(x, f(x)) : x \in X\} \subseteq X \times Y$ heißt *Graph von* f. Man kann diesen mit der Funktion f identifizieren.

Für eine Funktion $f : X \to Y$ heißt X *Definitionsbereich* und Y *Wertebereich* von f. Für $A \subseteq X$ heißt

$$f(A) := \{f(x) : x \in A\} \qquad \textit{Bild von } A \textit{ unter } f,$$

und für $B \subseteq Y$ heißt

$$f^{-1}(B) := \{x \in X : f(x) \in B\} \qquad \textit{Urbild von } B \textit{ unter } f.$$

Insbesondere heißt $f(X) = \{f(x) : x \in X\}$ *Bild von* f (Menge aller $y \in Y$, die von f getroffen werden).

Im Beispiel oben ist

$$f(\mathbb{N}) = \{x \in \mathbb{N} : x \text{ ist ungerade}\}$$

und etwa

$$f^{-1}(\{1, 2, 3, 4, 5\}) = f^{-1}(\{1, 3, 5\}) = \{1, 2, 3\}.$$

3.2. Definition: Sei $f : X \to Y$ eine Funktion.

(a) f heißt *surjektiv*, falls $f(X) = Y$ gilt, dh falls jedes $y \in Y$ von f getroffen wird.

(b) f heißt *injektiv*, falls gilt $\forall x_1, x_2 \in X : x_1 \neq x_2 \Rightarrow f(x_1) \neq f(x_2)$, dh falls es zu jedem Element im Bild von f genau ein Urbild gibt.

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

Beispiele: (1) Die Funktion $f : \mathbb{N} \to \mathbb{N}$ von oben ist injektiv und nicht surjektiv.

(2) Die Funktion $g : \mathbb{N} \to \{0, 1\}, x \mapsto \begin{cases} 0, & x \text{ ist gerade} \\ 1, & x \text{ ist ungerade} \end{cases}$ ist surjektiv und nicht injektiv.

3.3. Komposition: Sind $f : X \to Y$, $g : Y \to Z$ Funktionen, so definiert

$$g \circ f : X \to Z, \quad x \mapsto g(f(x))$$

eine Funktion $g \circ f$ ("g nach f"), die *Hintereinanderausführung* oder *Komposition* von f und g.

Satz: Sind $f : X \to Y$, $g : Y \to Z$, $h : Z \to W$ Funktionen, so gilt

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

dh die Hintereinanderausführung von Funktionen ist *assoziativ*.

3.4. Die Umkehrabbildung: Ist $f : X \to Y$ eine bijektive Funktion, so definiert

$$Y \to X, \quad y \mapsto x, \text{ falls } f(x) = y$$

eine Funktion f^{-1}, die *Umkehrabbildung* (oder *Umkehrfunktion*) von f.

Beachte: Da f surjektiv ist, gibt es zu jedem y ein solches x. Da f injektiv ist, ist dieses x eindeutig bestimmt. Somit ist f^{-1} tatsächlich eine Funktion.

Beispiel: Ist $\emptyset \neq X$ eine Menge, so heißt die Funktion $X \to X$, $x \mapsto x$, die *Identität* auf X, geschrieben Id_X oder id_X. Die Funktion id_X ist bijektiv und es ist $(\mathrm{id}_X)^{-1} = \mathrm{id}_X$.

Bemerkung: Ist $f : X \to Y$ bijektiv, so gilt $f^{-1} \circ f = \mathrm{id}_X$ und $f \circ f^{-1} = \mathrm{id}_Y$.

3.5. Satz: Sind $f : X \to Y$ und $g : Y \to X$ Funktionen mit $g \circ f = \mathrm{id}_X$ und $f \circ g = \mathrm{id}_Y$, so ist f bijektiv und es gilt $g = f^{-1}$.

Bemerkung: Durch Vertauschen der Rollen von f und g folgt auch $f = g^{-1}$ und insbesondere $(f^{-1})^{-1} = f$.

3.6. Satz: Sind $f : X \to Y$ und $g : Y \to Z$ bijektive Funktionen, so ist $g \circ f : X \to Z$ bijektiv, und es gilt:

$$(g \circ f)^{-1} = f^{-1} \circ g^{-1} : Z \to X.$$

Das ist klar. Die Formel heißt "Hemd-Jacken-Regel".

4 Die reellen Zahlen

Grundmenge der Analysis ist die Menge $\mathbb{R}$ der *reellen Zahlen*. Wir führen diese Menge durch 15 *Axiome* ein, dh durch grundlegende Eigenschaften, aus denen sich **alle** weiteren Rechenregeln herleiten lassen. Wir nehmen dann $\mathbb{R}$ als mit diesen Axiomen gegeben an. Eine explizite *Konstruktion* (die natürlich möglich ist!), führen wir hier nicht durch.

4.1. Körperaxiome: Es gibt *Verknüpfungen* $+ : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ ("plus", wir schreiben $a+b$) und $\cdot : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ ("mal", wir schreiben ab oder $a \cdot b$) mit

$$\forall a, b, c \in \mathbb{R} : (a+b)+c = a+(b+c) \quad (A1) \qquad (ab)c = a(bc) \quad (A5)$$
$$\exists 0 \in \mathbb{R} \forall a \in \mathbb{R} : a + 0 = a \quad (A2) \qquad \exists 1 \in \mathbb{R} \setminus \{0\} \forall a \in \mathbb{R} : a \cdot 1 = a \quad (A6)$$
$$\forall a \in \mathbb{R} \exists -a \in \mathbb{R} : a + (-a) = 0 \quad (A3) \qquad \forall a \in \mathbb{R} \setminus \{0\} \exists a^{-1} \in \mathbb{R} : a \cdot a^{-1} = 1 \quad (A7)$$
$$\forall a, b \in \mathbb{R} : a + b = b + a \quad (A4) \qquad ab = ba \quad (A8)$$
$$\forall a, b, c \in \mathbb{R} : a(b+c) = ab + ac \quad (A9)$$

Dabei sind $(A1)$ und $(A5)$ die *Assoziativgesetze*, $(A4)$ und $(A8)$ die *Kommutativgesetze*, und $(A9)$ ist das *Distributivgesetz*.

Schreibweisen: Für $a, b \in \mathbb{R}$ setzen wir $a - b := a + (-b)$ und, falls $b \neq 0$ ist, $\frac{a}{b} := ab^{-1}$.

Bemerkung: Aus $(A1) - (A9)$ lassen sich alle Rechenregeln bzgl. "+" und "·" herleiten (insbesondere z.B. "Bruchrechnung"). Diese werden von nun an als bekannt vorausgesetzt.

Beispiele: (1) Die Null in $(A2)$ ist eindeutig, ebenso die Eins in $(A6)$.

Beweis. Ist $\tilde{0} \in \mathbb{R}$ mit $\forall a \in \mathbb{R} : a + 0 = a$, so folgt wegen $(A2)$ und $(A4)$: $0 = 0 + \tilde{0} = \tilde{0} + 0 = \tilde{0}$. □

(2) Die Elemente a in $(A3)$ und a^{-1} in $(A7)$ sind eindeutig bestimmt. Außerdem gilt für jedes $a \in \mathbb{R}$: $-(-a) = a$ und, falls $a \neq 0$ ist, $(a^{-1})^{-1} = a$.

(3) $\forall a \in \mathbb{R}$: $a \cdot 0 = 0$

Beweis. Sei $a \in \mathbb{R}$. Nach $(A2)$ und $(A9)$ gilt $a \cdot 0 = a \cdot (0 + 0) = a \cdot 0 + a \cdot 0$. Setzen wir $b := a \cdot 0$, so folgt $0 = b + (-b) = (b + b) + (-b) = b + (b + (-b)) = b + 0 = b$. □

(4) $\forall a \in \mathbb{R}$: $-a = (-1) \cdot a$.

Beweis. Es gilt (nach $(A6)$, $(A9)$, und (3)): $a + a \cdot (-1) = a \cdot 1 + a \cdot (-1) = a \cdot (1 + (-1)) = a \cdot 0 = 0$, also $-a = a \cdot (-1) = (-1) \cdot a$. □

(5) $\forall a \in \mathbb{R}$: $a^2 = (-a)^2$, wobei $a^2 := a \cdot a$.

Beweis. Es ist, nach (4), ($A5$), ($A8$) und (2): $(-a)^2 = (-a)\cdot(-1)\cdot a = (-(-a))\cdot a = a^2$. $\square$

4.2. Anordnungsaxiome: In $\mathbb{R}$ ist eine *Ordnung* "$\leq$" gegeben mit folgenden Eigenschaften:

$$\begin{array}{ll} (A10) & \forall a,b \in \mathbb{R} : a \leq b \text{ oder } b \leq a, \\ (A11) & \forall a,b,c \in \mathbb{R} : a \leq b \text{ und } b \leq c \Rightarrow a \leq c, \\ (A12) & \forall a,b \in \mathbb{R} : a \leq b \text{ und } b \leq a \Rightarrow a = b, \\ (A13) & \forall a,b,c \in \mathbb{R} : a \leq b \Rightarrow a + c \leq b + c, \\ (A14) & \forall a,b,c \in \mathbb{R} : a \leq b \text{ und } 0 \leq c \Rightarrow ac \leq bc. \end{array}$$

($A11$) heißt Transitivität, ($A12$) heißt Antisymmetrie. ($A10$) bedeutet, dass die Ordnung "total" ist (dh dass man je zwei Elemente vergleichen kann). Außerdem beinhaltet ($A10$) auch, dass $a \leq a$ gilt (Reflexivität).

($A13$) und ($A14$) bedeuten, dass die Ordnung $\leq$ mit den Verknüpfungen "+" und "·" verträglich ist.

Schreibweisen: $b \geq a :\Leftrightarrow a \leq b$; $a < b :\Leftrightarrow a \leq b$ und $a \neq b$; $b > a :\Leftrightarrow a < b$.

Bemerkung: Aus ($A1$) – ($A14$) lassen sich alle Rechenregeln für Ungleichungen herleiten. Diese setzen wir von nun an als bekannt voraus.

Beispiele: (1) Für alle $a \in \mathbb{R}$ gilt $a^2 \geq 0$.

Beweis. Sei $a \in \mathbb{R}$. Fall $a \geq 0$: Dann gilt $a \cdot a \geq 0 \cdot a$ nach ($A14$) und somit $a^2 \geq 0$ nach 4.1(3).
Fall $a < 0$: Dann gilt nach ($A13$): $0 = a + (-a) \leq 0 + (-a) = -a$. Somit ist $(-a)^2 \geq 0$ nach Fall 1. Nach 4.1(5) ist dann $a^2 = (-a)^2 \geq 0$. $\square$

(2) Aus $a \leq b$ und $c \leq 0$ folgt $ac \geq bc$.

Beweis. $c \leq 0 \Rightarrow -c \geq 0$ (siehe Beweis von (1)). Nach ($A14$) ist dann $a(-c) \leq b(-c)$. Beachtet man $a(-c) = -ac$ und $b(-c) = -bc$ und addiert $ac + bc$ zu der Ungleichung (($A13$)!), so erhält man $bc \leq ac$. $\square$

Intervalle: Seien $a, b \in \mathbb{R}$ mit $a < b$. Wir setzen:

$$\begin{aligned}
[a,b] &:= \{x \in \mathbb{R} : a \leq x \leq b\} && \text{abgeschlossenes Intervall,}\\
(a,b) &:= \{x \in \mathbb{R} : a < x < b\} && \text{offenes Intervall,}\\
[a,b) &:= \{x \in \mathbb{R} : a \leq x < b\} && \text{halboffenes Intervall,}\\
(a,b] &:= \{x \in \mathbb{R} : a < x \leq b\} && \text{halboffenes Intervall,}\\
[a,\infty) &:= \{x \in \mathbb{R} : a \leq x\},\\
(a,\infty) &:= \{x \in \mathbb{R} : a < x\},\\
(-\infty,a] &:= \{x \in \mathbb{R} : x \leq a\},\\
(-\infty,a) &:= \{x \in \mathbb{R} : x < a\}.
\end{aligned}$$

Weiter: $[a,a] := \{a\}$ und $(-\infty,\infty) := \mathbb{R}$.

4.3. Der Betrag: Für $a \in \mathbb{R}$ heißt $|a| := \begin{cases} a & ,a \geq 0 \\ -a & ,a < 0 \end{cases}$ der *Betrag von a.*

Beispiele: $|1| = 1$, da $1 = 1 \cdot 1 \geq 0$ (vgl. 4.2(2)), $|-2| = -(-2) = 2$, da $2 = 1 + 1 \geq 1 + 0 = 1 \geq 0$ und somit $-2 \leq 0$.

Beachte: Es gilt $|a| = |-a|$ für alle $a \in \mathbb{R}$, also auch $|a - b| = |b - a|$ für alle $a, b \in \mathbb{R}$.

Anschaulich ist $|a - b|$ der **Abstand** von a und b auf der Zahlengeraden.

Regeln: Seien $a, b, c \in \mathbb{R}$. Dann gilt:

(1) $|a| \geq 0$;

(2) $|a| = 0 \Leftrightarrow a = 0$;

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

(4) $\pm a \leq |a|$ und $\Big(|a| \leq c \;\Leftrightarrow\; (a \leq c \text{ und } -a \leq c)\Big)$;

(5) $|a + b| \leq |a| + |b|$ *Dreiecksungleichung*;

(6) $||a| - |b|| \leq |a - b|$ *umgekehrte Dreiecksungleichung.*

Beweis. (1)-(4) sind leicht. Zu (5): Falls $a + b \geq 0$ ist, so gilt $|a + b| = a + b \leq |a| + |b|$ nach (4). Falls $a + b < 0$ ist, so ist $|a + b| = -(a + b) = -a + (-b) \leq |a| + |b|$ nach (4).

Zu (6): Nach (5) ist $|a| = |a - b + b| \leq |a - b| + |b|$ und $|b| = |b - a + a| \leq |a - b| + |a|$. Es folgt $|a| - |b| \leq |a - b|$ und $|b| - |a| \leq |a - b|$. Nach (4) gilt somit $||a| - |b|| \leq |a - b|$. $\square$

Bemerkung: Setzt man $\max\{a, b\} := \begin{cases} a & ,a \geq b \\ b & ,a < b \end{cases}$, so ist $|a| = \max\{a, -a\}$.

4.4. Supremum und Infimum: Sei $M \subseteq \mathbb{R}$ mit $M \neq \emptyset$.

M heißt *nach oben* [*unten*] *beschränkt* $:\Leftrightarrow \exists \gamma \in \mathbb{R}\, \forall x \in M : x \leq \gamma$ [$x \geq \gamma$].

In diesem Fall heißt γ eine *obere Schranke* (OS) [*untere Schranke* (US)] von M.

Eine obere Schranke [untere Schranke] γ von M mit $\gamma \in M$ heißt *Maximum* [*Minimum*] von M und wird mit $\max M$ [$\min M$] bezeichnet.

Wegen ($A12$) sind $\max M$ und $\min M$ im Falle der Existenz *eindeutig bestimmt*.

Beispiele: Jede endliche nichtleere Menge $M \subseteq \mathbb{R}$ ist nach oben und nach unten beschränkt und besitzt Maximum und Minimum. Für jedes $a \in \mathbb{R}$ gilt $|a| = \max\{a, -a\}$.

$[1, 2]$ ist nach oben und nach unten beschränkt, es ist $1 = \min M$ und $2 = \max M$.

$(1, \infty)$ ist nach unten, aber nicht nach oben beschränkt und hat kein Minimum.

Definition: Ist γ obere Schranke [untere Schranke] von M mit $\gamma \leq \tilde{\gamma}$ [$\gamma \geq \tilde{\gamma}$] für **jede** obere Schranke [untere Schranke] $\tilde{\gamma}$ von M, so heißt γ *Supremum* [*Infimum*] von M (**kleinste** obere Schranke von M [**größte** untere Schranke von M]) und wird mit $\sup M$ [$\inf M$] bezeichnet.

Nach ($A12$) sind Supremum und Infimum im Falle der Existenz eindeutig bestimmt.

Bemerkung: Ein Maximum ist immer auch Supremum, und es gilt $\sup M = \max M$ genau dann, wenn $\sup M \in M$ ist (entsprechend für min und inf).

Beispiele: (1) $M = [1, 2)$. M ist nach unten und nach oben beschränkt. Es ist $1 = \min M = \inf M$, M hat kein Maximum, und es ist $\sup M = 2$.

Beweis. 2 ist obere Schranke von M. Zeige: es gibt keine echt kleinere obere Schranke. Sei $\tilde{\gamma} < 2$. Zeige: $\tilde{\gamma}$ ist nicht obere Schranke von M. Falls $\tilde{\gamma} < 1$ ist, so gilt $\tilde{\gamma} < 1 \in M$, also ist $\tilde{\gamma}$ keine obere Schranke von M. Falls $\tilde{\gamma} \geq 1$ ist, so ist $\tilde{\gamma} < \frac{\tilde{\gamma}+2}{2} \in M$ und $\tilde{\gamma}$ ist keine obere Schranke von M. $\square$

(2) $M = (1, \infty)$: Es ist $1 = \inf M \notin M$ und $\sup M$ existiert nicht.

4.5. Das Vollständigkeitsaxiom:

($A15$) Jede nichtleere nach oben beschränkte Teilmenge von $\mathbb{R}$ hat ein Supremum.

Folgerung: Jede nichtleere nach unten beschränkte Teilmenge von $\mathbb{R}$ hat ein Infimum.

Beweis. Sei $M \subseteq \mathbb{R}$ mit $M \neq \emptyset$. Setze $-M := \{-x : x \in M\}$.

Vorbemerkung: Dann gilt γ ist untere Schranke von $M \Leftrightarrow -\gamma$ ist obere Schranke von $-M$.

Da $M \neq \emptyset$ untere Schranken hat, ist $-M \neq \emptyset$ nach oben beschränkt, hat also ein Supremum $s := \sup(-M)$. Nach der Vorbemerkung ist $-s$ untere Schranke von M, und für jede untere Schranke $\tilde{\gamma}$ von M ist $-\tilde{\gamma}$ eine obere Schranke von $-M$, also $s \leq -\tilde{\gamma}$, dh $\tilde{\gamma} \leq -s$. Somit ist $-s$ größte untere Schranke von M, dh $-s = \inf M$. $\square$

Definition: Eine Menge $\emptyset \neq M \subseteq \mathbb{R}$ heißt *beschränkt*, falls M nach oben und nach unten beschränkt ist.

Bemerkung: M beschränkt $\Leftrightarrow \exists \gamma \geq 0 \forall x \in M : |x| \leq \gamma$.

4.6. Satz: Sei $\emptyset \neq B \subseteq A \subseteq \mathbb{R}$. Dann gilt:

(1) A beschränkt $\Rightarrow \inf A \leq \sup A$.

(2) A nach oben [nach unten] beschränkt $\Rightarrow$ B nach oben [nach unten] beschränkt und $\sup B \leq \sup A$ [$\inf B \geq \inf A$].

(3) Sei A nach oben [nach unten] beschränkt und γ eine obere Schranke [untere Schranke] von A. Dann gilt:

$$\begin{aligned} \gamma = \sup A &\iff \text{für jedes } \varepsilon > 0 \text{ gibt es ein } x \in A \text{ mit } x > \gamma - \varepsilon \\ [\gamma = \inf A &\iff \text{für jedes } \varepsilon > 0 \text{ gibt es ein } x \in A \text{ mit } x < \gamma + \varepsilon]. \end{aligned}$$

Beweis. (1) und (2) sind leicht.

Zu (3): "$\Rightarrow$": Sei $\gamma = \sup A$ und $\varepsilon > 0$. Dann ist $\gamma - \varepsilon$ keine obere Schranke von A.

"$\Leftarrow$" (Kontraposition): Sei $\gamma \neq \sup A =: \tilde{\gamma}$. Dann $\gamma > \tilde{\gamma}$, da γ obere Schranke von A ist, und $\varepsilon := \gamma - \tilde{\gamma} > 0$. Für jedes $x \in A$ gilt dann $x \leq \tilde{\gamma} = \gamma - \varepsilon$. Wir haben gezeigt: $\exists \varepsilon > 0 \forall x \in A : x \leq \gamma - \varepsilon$, dh $\neg(\forall \varepsilon > 0 \exists x \in A : x > \gamma - \varepsilon)$. $\square$

4.7. Natürliche Zahlen: Idee ist $\mathbb{N} = \{1, 1+1, 1+1+1, \ldots\}$.

Definition: $A \subseteq \mathbb{R}$ heißt *Induktionsmenge* (*IM*), falls $1 \in A$ und $\forall x \in A : x + 1 \in A$.

Beispiele: $\mathbb{R}$, $[1, \infty)$, $\{1\} \cup [2, \infty)$ sind Induktionsmengen, $\{1\} \cup (2, \infty)$ ist keine Induktionsmenge.

Definition: $\mathbb{N} := \{x \in \mathbb{R} : \text{für jede Induktionsmenge } A \subseteq \mathbb{R} \text{ gilt: } x \in A\}$ heißt *Menge der natürlichen Zahlen*.

Satz: (1) $\mathbb{N}$ ist eine Induktionsmenge (somit ist $\mathbb{N}$ die kleinste Induktionsmenge).

(2) $\mathbb{N}$ ist nicht nach oben beschränkt.

(3) Für jedes $x \in \mathbb{R}$ gibt es $n \in \mathbb{N}$ mit $n > x$.

(4) Für jedes $b \in \mathbb{R}$ mit $b > 0$ gibt es $n \in \mathbb{N}$ mit $\frac{1}{n} < b$.

Beweis. (1) Es ist $1 \in \mathbb{N}$, da $1 \in A$ für jede Induktionsmenge A. Sei $x \in \mathbb{N}$. Zu zeigen: $x+1 \in \mathbb{N}$. Sei dazu $A \subseteq \mathbb{R}$ eine Induktionsmenge. Dann gilt $x \in \mathbb{N} \subseteq A$ und also $x+1 \in A$ (da A eine Induktionsmenge ist).

(2) **Annahme:** $\mathbb{N}$ ist nach oben beschränkt. Dann existiert $\gamma := \sup \mathbb{N}$. Nach 4.6(3) (für $\varepsilon = 1$) finden wir $n \in \mathbb{N}$ mit $n > \gamma - 1$. Dann gilt $n + 1 > \gamma$ und $n + 1 \in \mathbb{N}$, dh γ ist nicht obere Schranke von $\mathbb{N}$, Widerspruch.

(3) folgt sofort aus (2).

(4) Sei $b > 0$. Nach (3) gibt es $n \in \mathbb{N}$ mit $n > \frac{1}{b} > 0$. Es folgt $\frac{1}{n} < b$. $\square$

4.8. Vollständige Induktion:

Satz: Ist $A \subseteq \mathbb{N}$ und ist A eine Induktionsmenge, dann ist $A = \mathbb{N}$.

Beweis. Nach Voraussetzung gilt $A \subseteq \mathbb{N}$. Da A Induktionsmenge ist, gilt $\mathbb{N} \subseteq A$. $\square$

Beweisverfahren durch Induktion

Für jedes $n \in \mathbb{N}$ sei $A(n)$ eine Aussage. Es gelte

$$\begin{array}{rl} \text{Induktionsanfang } (IA) & A(1) \\ \text{Induktionsschritt } (IS) & \forall n \in \mathbb{N} : (A(n) \Rightarrow A(n+1))\,. \end{array}$$

Dann ist für jedes $n \in \mathbb{N}$ die Aussage $A(n)$ wahr, dh es gilt $\forall n \in \mathbb{N} : A(n)$.

Beweis. Setze $A := \{n \in \mathbb{N} : A(n)\}$. Nach (IA) gilt $1 \in A$. Sei $n \in A$. Dann gilt $A(n)$ und nach (IS) ist auch $A(n+1)$ wahr, dh $n+1 \in A$. Somit ist A eine Induktionsmenge und $A = \mathbb{N}$ folgt aus dem Satz. $\square$

Beispiele: (1) Für alle $n \in \mathbb{N}$ gilt $\underbrace{n \geq 1}_{=:A(n)}$.

Beweis durch Induktion nach n. Induktionsanfang (IA): Es gilt $1 \geq 1$, dh $A(1)$ ist wahr.

Induktionsschluss (IS): Sei $n \in \mathbb{N}$. Es gelte $A(n)$, dh es gelte $n \geq 1$ (Induktionsvoraussetzung (IV)). Dann ist $n+1 \geq 1+1$ nach (IV) und $1+1 \geq 1+0 = 1$ nach 4.2 und 4.1, also $n+1 \geq 1$ und $A(n+1)$ ist wahr. $\square$

(2) Für jedes $n \in \mathbb{N}$ gilt $\underbrace{1+2+\ldots+n = \frac{n(n+1)}{2}}_{=:A(n)}$.

Beweis durch Induktion nach n. (IA) $n = 1$: Es gilt $1 = \frac{1\cdot(1+1)}{2}$, dh $A(1)$ ist wahr.

(IS) Sei $n \in \mathbb{N}$. Es gelte $1+2+\ldots+n = \frac{n(n+1)}{2}$ (IV). Zu zeigen ist $1+2+\ldots+n+(n+1) = \frac{(n+1)(n+2)}{2}$. Es gilt

$$1+2+\ldots+n+(n+1) \overset{(IV)}{=} \frac{n(n+1)}{2} + n + 1 = \frac{(n+1)(n+2)}{2}.$$

$\square$

Wir setzen die natürlichen Zahlen ab jetzt als bekannt voraus.

Definition durch Rekursion (bzw. durch Induktion)

Es sei $G(1)$ definiert, und für jedes $n \in \mathbb{N}$ sei $G(n+1)$ definiert unter der Voraussetzung, dass $G(1)$, $G(2)$, $\ldots$, $G(n)$ schon definiert sind.

Dann hat man $G(n)$ für jedes $n \in \mathbb{N}$ definiert.

Beispiele: (1) *Fakultät*: $1! := 1$ und rekursiv $(n+1)! := (n+1) \cdot n!$ für jedes $n \in \mathbb{N}$, sowie $0! := 1$. Dann ist $n! = n \cdot (n-1) \cdot \ldots \cdot 2 \cdot 1$ für jedes $n \in \mathbb{N}$.

(2) Summenzeichen $\sum$: Seien $a_1, a_2, \ldots \in \mathbb{R}$. Setze

$$\sum_{j=1}^{1} a_j := a_1 \quad \text{und für jedes } n \in \mathbb{N} : \sum_{j=1}^{n+1} a_j := \Big(\sum_{j=1}^{n} a_j \Big) + a_{n+1}.$$

Dann ist $\sum_{j=1}^{n} a_j = a_1 + a_2 + \ldots + a_n$ für jedes $n \in \mathbb{N}$. Die leere Summe ist $\sum_{j=1}^{0} a_j := 0$.

(3) Produktzeichen $\prod$: Seien $a_1, a_2, \ldots \in \mathbb{R}$. Setze

$$\prod_{j=1}^{1} a_j := a_1 \quad \text{und für jedes } n \in \mathbb{N} : \prod_{j=1}^{n+1} a_j := \Big(\prod_{j=1}^{n} a_j \Big) \cdot a_{n+1}.$$

Dann ist $\prod_{j=1}^{n} a_j = a_1 \cdot a_2 \cdot \ldots \cdot a_n$ für jedes $n \in \mathbb{N}$. Das leere Produkt ist $\prod_{j=1}^{0} a_j := 1$.

(4) Potenzen: Setze für $a \in \mathbb{R}$: $a^0 := 1$, $a^1 := a$ und $a^{n+1} := a^n \cdot a$ für jedes $n \in \mathbb{N}$. Dann gilt $a^n = \underbrace{a \cdot a \cdot \ldots \cdot a}_{n \text{ Faktoren}}$ für jedes $n \in \mathbb{N}$.

Varianten der Induktion: Man kann die Induktion auch bei z.B. $n = 5$ beginnen lassen. Zum Beweis von "$\forall n \in \mathbb{N}$ mit $n \geq 5 : A(n)$" hat man dann zu zeigen:

$$\begin{array}{ll} (IA) & A(5) \\ (IS) & \forall n \in \mathbb{N} \text{ mit } n \geq 5 : (A(n) \Rightarrow A(n+1)) \,. \end{array}$$

Bei der *Abschnittsinduktion* zeigt man

$$\begin{array}{ll} (IA) & A(1) \\ (IS) & \forall n \in \mathbb{N} : (A(1) \wedge A(2) \wedge \ldots \wedge A(n) \Rightarrow A(n+1)) \,. \end{array}$$

Auch dann hat man $A(n)$ für jedes $n \in \mathbb{N}$ gezeigt (man führt eigentlich ein Induktion für $B(n) := A(1) \wedge A(2) \wedge \ldots \wedge A(n)$ durch). Vorteil ist hier, dass man zum Nachweis von $A(n+1)$ mehr Aussagen verwenden darf.

Beispiel: Jede natürliche Zahl n lässt sich als (eventuell leeres) Produkt von Primzahlen schreiben. Wir zeigen dies durch Abschnittsinduktion, wobei für jedes $n \in \mathbb{N}$ hier $A(n)$ die

Aussage "n lässt sich als (ev. leeres) Produkt von Primzahlen schreiben" bezeichne (für $n > 1$ kann man "eventuell leeres" weglassen).

(IA) Wir können 1 als leeres Produkt von Primzahlen schreiben, also gilt $A(1)$.

(IS) Sei $n \in \mathbb{N}$ so, dass sich sich jedes $k \in \mathbb{N}$ mit $k \leq n$ als (ev. leeres) Produkt von Primzahlen schreiben lässt (IV). Wir zeigen, dass sich $n+1$ als Produkt von Primzahlen schreiben lässt.
Fall 1: $n+1$ ist Primzahl. Dann sind wir fertig und $A(n+1)$ gilt (das Produkt hat einen Faktor).
Fall 2: $n+1$ ist nicht Primzahl. Wegen $n+1 \neq 1$ gibt es dann ein $k \in \mathbb{N}$ mit $1 < k < n+1$, das $n+1$ teilt. Setze $l := \frac{n+1}{k}$. Dann ist $l \in \mathbb{N}$ mit $1 < l < n+1$. Nach (IV) lassen sich k und l als Produkt von Primzahlen schreiben, also auch $n+1 = kl$. Somit ist auch in diesem Fall $A(n+1)$ gezeigt.

4.9. Ganze und rationale Zahlen:

Definition: $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$, $\mathbb{Z} := \mathbb{N}_0 \cup \{-n : n \in \mathbb{N}\}$ *Menge der ganzen Zahlen* und $\mathbb{Q} := \{\frac{p}{q} : p \in \mathbb{Z}, q \in \mathbb{N}\}$ *Menge der rationalen Zahlen.*

Bemerkung: Die Axiome $(A1) - (A14)$ gelten auch in $\mathbb{Q}$. Hingegen hat die nach oben beschränkte Menge $M := \{x \in \mathbb{Q} : x^2 < 2\}$ kein Supremum in $\mathbb{Q}$, dh das Vollständigkeitsaxiom $(A15)$ gilt **in $\mathbb{Q}$ nicht!**

Satz: Jede nichtleere nach unten beschränkte Teilmenge von $\mathbb{Z}$ hat ein Minimum. (ohne Beweis)

Satz: Sind $x, y \in \mathbb{R}$ mit $x < y$, so gibt es eine rationale Zahl $r \in \mathbb{Q}$ mit $x < r < y$.

Im Fall $y - x > 1$ gibt es ein $p \in \mathbb{Z}$ mit $x < p < y$.

Beweis. Für den Zusatz setze $p := \min\{m \in \mathbb{Z} : m > x\}$. Dann gilt $x < p$ und $p - 1 \leq x$, also $p \leq x + 1 < y$.

Zum Beweis des allgemeinen Falles wählt man zunächst $q \in \mathbb{N}$ so, dass $y - x > 1/q$. Dann ist $qy - qx > 1$ und wir finden nach dem Zusatz ein $p \in \mathbb{Z}$ mit $qx < p < qy$. Es folgt $x < p/q < y$. □

4.10. Binomialkoeffizienten: Für $n, k \in \mathbb{N}_0$ mit $k \leq n$ setzt man

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

("n über k").

Es gilt $\binom{n}{0} = 1 = \binom{n}{n}$ für alle $n \in \mathbb{N}_0$, $\binom{n}{k} = \binom{n}{n-k}$ für alle $n \geq k \geq 0$, $\binom{4}{2} = \frac{4 \cdot 3}{2} = 6$.

Lemma: Für $1 \leq k \leq n$ gilt

$$\binom{n}{k} + \binom{n}{k-1} = \binom{n+1}{k}.$$

Beweis. Es gilt

$$\begin{aligned}\binom{n}{k} + \binom{n}{k-1} &= \frac{n!}{k!(n-k)!}\frac{n+1-k}{n+1-k} + \frac{n!}{(k-1)!(n+1-k)!}\frac{k}{k} \\ &= \frac{(n+1-k+k)n!}{k!(n+1-k)!} = \binom{n+1}{k}.\end{aligned}$$

□

4.11. Potenzen: Für $a \in \mathbb{R}$ und $n \in \mathbb{N}_0$ haben wir a^n in 4.8(4) definiert. Für $a \neq 0$ und $n \in \mathbb{N}$ setzt man $a^{-n} := \frac{1}{a^n}$.
Es gelten die bekannten Rechenregeln, also etwa $(a^m)^n = a^{m \cdot n}$ und $a^m \cdot a^n = a^{m+n}$.

(1) Für alle $a, b \in \mathbb{R}$ und alle $n \in \mathbb{N}$ gilt:

$$a^n - b^n = (a-b)\sum_{k=0}^{n-1} a^{n-1-k}b^k.$$

[Übungsaufgabe]

(2) **Binomialsatz:** Für alle $a, b \in \mathbb{R}$ und $n \in \mathbb{N}_0$ gilt:

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

Etwa (mit $a = b = 1$): $\sum_{k=0}^{n} \binom{n}{k} = 2^n$ oder (mit $a = -1$ und $b = 1$) $\sum_{k=0}^{n}(-1)^k\binom{n}{k} = (-1+1)^n = 0$.

Beweis durch Induktion nach n. IA $n = 0$ ist klar.

IS: Sei $n \in \mathbb{N}$ mit $(a+b)^n = \sum_{k=0}^n \binom{n}{k} a^k b^{n-k}$. Dann gilt

$$\begin{aligned}
(a+b)^{n+1} &= (a+b)(a+b)^n = (a+b) \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k} \\
&= \sum_{k=0}^{n} \binom{n}{k} a^{k+1} b^{n-k} + \sum_{k=0}^{n} \binom{n}{k} a^k b^{n+1-k} \\
(j = k+1) \quad &= \sum_{j=1}^{n+1} \binom{n}{j-1} a^j b^{n+1-j} + \sum_{k=0}^{n} \binom{n}{k} a^k b^{n+1-k} \\
&= a^{n+1} + \sum_{k=1}^{n} \underbrace{\left[\binom{n}{k-1} + \binom{n}{k}\right]}_{=\binom{n+1}{k}} a^k b^{n+1-k} + b^{n+1} \\
&= \sum_{k=0}^{n+1} \binom{n+1}{k} a^k b^{n+1-k}.
\end{aligned}$$

□

(3) **Bernoullische Ungleichung** (BU): Sei $x \geq -1$. Dann gilt für alle $n \in \mathbb{N}$:

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

Beweis durch Induktion nach n. IA $n = 1$: $(1+x)^1 = 1 + x \geq 1 + 1 \cdot x$.
IS: Sei $n \in \mathbb{N}$ mit $(1+x)^n \geq 1 + nx$ (IV). Dann gilt:

$$(1+x)^{n+1} = (1+x)^n(1+x) \geq (1+nx)(1+x) = 1 + nx + x + nx^2 \geq 1 + (n+1)x.$$

□

(4) **Folgerung:** Sei $a \in \mathbb{R}$.
Ist $a > 1$, so gibt es zu jedem $K > 0$ ein $n \in \mathbb{N}$ mit $a^n > K$.
Ist $a \in (0,1)$, so gibt es zu jedem $\varepsilon > 0$ ein $n \in \mathbb{N}$ mit $a^n < \varepsilon$.

Beweis. Ist $a > 1$, so finden wir zu $K > 0$ ein $n \in \mathbb{N}$ mit $n > \frac{K}{a-1}$ und mit (BU) für $x = a - 1$ gilt dann

$$a^n = (1+x)^n \geq 1 + nx = 1 + n(a-1) > 1 + K > K.$$

Ist $a \in (0,1)$, so ist $a^{-1} > 1$ und wir finden zu $\varepsilon > 0$ ein $n \in \mathbb{N}$ mit $a^{-n} > \varepsilon^{-1}$, dh mit $a^n < \varepsilon$. □

(5) Für alle $x, y \geq 0$ und $n \in \mathbb{N}$ gilt:

$$x \leq y \Longleftrightarrow x^n \leq y^n.$$

[Übungsaufgabe, verwende (1)]

4.12. Wurzeln: Sei $n \in \mathbb{N}$.

Satz: Zu jedem $a \in \mathbb{R}$ mit $a \geq 0$ gibt es genau ein $b \in \mathbb{R}$ mit $b \geq 0$ und $b^n = a$.

Bezeichnung: $b = \sqrt[n]{a}$, "n-te Wurzel aus a".

Bemerkung: Also existiert etwa die reelle Zahl $\sqrt{2}$ (bekannt: $\sqrt{2} \notin \mathbb{Q}$). Wir ziehen hier nur Wurzeln aus Zahlen ≥ 0!

Folgerung: Für alle $a, b \geq 0$ gilt wegen 4.11 (5):

$$a \leq b \Longleftrightarrow \sqrt[n]{a} \leq \sqrt[n]{b},$$

und wegen $(\sqrt[n]{a}\sqrt[n]{b})^n = ab$ gilt

$$\sqrt[n]{ab} = \sqrt[n]{a}\sqrt[n]{b}.$$

Beweis des Satzes. Der Fall $a = 0$ ist klar. Sei also $a > 0$. Setze $M := \{x \geq 0 : x^n < a\}$. Dann ist M nichtleer, und $a + 1$ ist obere Schranke von M: Sei $x \geq 0$ mit $x^n < a$. Dann ist $x \leq a + 1$, da die Annahme $x > a + 1$ mithilfe von BU zum Widerspruch

$$x^n > (1 + a)^n \geq 1 + na > na \geq a$$

führt.

Also existiert $b := \sup M$. Dabei ist $b > 0$, denn $x := \left(1 + \frac{1}{na}\right)^{-1} \in M$ wegen

$$\frac{1}{x^n} = \left(1 + \frac{1}{na}\right)^n \overset{BU}{\geq} 1 + \frac{n}{na} = 1 + \frac{1}{a} > \frac{1}{a}.$$

Vorbemerkung: Für $\delta \in (0, b)$ gilt:

$$\begin{array}{rcl} (b+\delta)^n & \leq & b^n + \delta 2^n b^{n-1} \\ (b-\delta)^n & \geq & b^n - \delta 2^n b^{n-1}. \end{array}$$

[Es ist

$$(b+\delta)^n \quad = \quad b^n + \sum_{k=1}^{n} \binom{n}{k} \underbrace{\delta^k b^{n-k}}_{\leq \delta b^{n-1}} \leq b^n + \delta b^{n-1} \underbrace{\sum_{k=1}^{n} \binom{n}{k}}_{\leq 2^n} \leq b^n + \delta 2^n b^{n-1}$$

und die andere Ungleichung zeigt man ähnlich.]

Wir behaupten $b^n = a$.

Annahme: $b^n < a$ [$b^n > a$]. Dann ist $\varepsilon := a - b^n > 0$ [$\varepsilon := b^n - a > 0$] und wir finden $\delta \in (0, b)$ mit $\delta < \frac{\varepsilon}{2^n b^{n-1}}$. Unter Verwendung der Vorbemerkung ist dann

$$\begin{aligned} (b+\delta)^n &\leq b^n + \delta 2^n b^{n-1} < b^n + \varepsilon = a \\ [(b-\delta)^n &\geq b^n - \delta 2^n b^{n-1} > b^n - \varepsilon = a], \end{aligned}$$

also $b + \delta \in M$ und somit $b + \delta \leq b$ [also $b - \delta$ obere Schranke von M und somit $b - \delta \geq b$]: Widerspruch! □

Man kann sehen, dass der Nachweis von $b^n = a$ hier zwar möglich, aber etwas aufwendiger ist. Das liegt daran, dass wir an dieser Stelle noch nicht so viel verwenden können. Später werden wir die Existenz von Wurzeln einfacher zeigen können.

5 Die komplexen Zahlen

5.1. Konstruktion: Auf $\mathbb{R}^2$ erklären wir zwei Verknüpfungen “+” und “$*$” durch

$$\begin{aligned}(x_1, y_1) + (x_2, y_2) &:= (x_1 + x_2, y_1 + y_2) \qquad \text{und} \\ (x_1, y_1) * (x_2, y_2) &:= (x_1 x_2 - y_1 y_2, x_2 y_1 + x_1 y_2).\end{aligned}$$

Man kann nun nachrechnen, dass die Körperaxiome $(A1) - (A9)$ für diese Verknüpfungen gelten (wobei “$*$” die Rolle der Multiplikation übernimmt):

$(A1) - (A4)$ sind leicht. $(A5)$ muss man nur hinschreiben (etwas aufwendiger), $(1, 0)$ ist das neutrale Element bzgl. “$*$”. Das zu $(x, y) \neq (0, 0)$ bzgl. “$*$” inverse Element ist $\left(\frac{x}{x^2+y^2}, -\frac{y}{x^2+y^2}\right)$. $(A8)$ ist klar, $(A9)$ schreibt man sofort hin.

Es gilt $(x_1, 0) + (x_2, 0) = (x_1 + x_2, 0)$ und $(x_1, 0) * (x_2, 0) = (x_1 x_2, 0)$, dh man kann das Paar $(x, 0)$ mit der reellen Zahl x identifizieren. Somit sind die reellen Zahlen $(\mathbb{R}, +, \cdot)$ in unserem Modell $(\mathbb{R}^2, +, *)$ der “komplexen Zahlen” enthalten.

Weiter ist $(0, 1) * (0, 1) = (-1, 0)$, dh $(0, 1)$ ist eine “Zahl”, deren Quadrat $= -1$ ist. Man setzt nun $i := (0, 1)$ und schreibt $x + iy$ statt (x, y). Es ist dann

$$\mathbb{C} := \{x + iy : x, y \in \mathbb{R}\} \quad \text{die Menge der komplexen Zahlen.}$$

Für $z = x + iy \in \mathbb{C}$ mit $x, y \in \mathbb{R}$ heißt x der *Realteil von* z (geschrieben $\operatorname{Re} z$) und y heißt der *Imaginärteil von* z (geschrieben $\operatorname{Im} z$). Komplexe Zahlen z mit $\operatorname{Re} z = 0$ heißen *rein imaginär* und komplexe Zahlen mit $\operatorname{Im} z = 0$ heißen *reell.*

Also: Man kann mit komplexen Zahlen wie gewohnt rechnen und muss nur $i^2 = -1$ berücksichtigen, etwa

$$(x_1 + iy_1)(x_2 + iy_2) = x_1 x_2 + i(y_1 x_2 + x_1 y_2) + i^2 y_1 y_2 = (x_1 x_2 - y_1 y_2) + i(x_2 y_1 + x_1 y_2),$$

vergleiche die Definition von “$*$” oben.

5.2. Konjugation und Betrag: Zu einer komplexen Zahl $z = x + iy$ heißt $\overline{z} := x - iy$ die *konjugiert komplexe Zahl.* Es gilt dann $z \cdot \overline{z} = x^2 + y^2 \geq 0$ und $|z| := \sqrt{z\overline{z}} = \sqrt{x^2 + y^2}$ heißt Betrag der komplexen Zahl z.

Rechenregeln: Für alle $w, z \in \mathbb{C}$ gilt:

$$\overline{(\overline{z})} = z,$$

$$|\overline{z}| = |z|,$$

$$\overline{z + w} = \overline{z} + \overline{w},$$

$$\overline{z \cdot w} = \overline{z} \cdot \overline{w},$$

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

$\max\{|\operatorname{Re} z|, |\operatorname{Im} z|\} \leq |z| \leq |\operatorname{Re} z| + |\operatorname{Im} z|$ [denn $\max\{|x|, |y|\} \leq \sqrt{x^2 + y^2} \leq |x| + |y|$ für $x, y \in \mathbb{R}$],

$|z \cdot w| = |z| \cdot |w|$ [denn $|zw|^2 = zw\overline{zw} = z\overline{z}w\overline{w} = |z|^2|w|^2$],

$|z + w| \leq |z| + |w|$ (Dreiecksungleichung) und $||z| - |w|| \leq |z - w|$. [Es ist nämlich

$$\begin{aligned} |z+w|^2 &= (z+w)(\overline{z}+\overline{w}) \\ &= z\overline{z} + w\overline{w} + z\overline{w} + w\overline{z} \\ &= |z|^2 + |w|^2 + \underbrace{z\overline{w} + \overline{(z\overline{w})}}_{=2\operatorname{Re}(z\overline{w}) \leq 2|z\overline{w}|} \\ &\leq |z|^2 + |w|^2 + 2|z||w| = (|z| + |w|)^2, \end{aligned}$$

woraus die Dreiecksungleichung durch Wurzelziehen folgt.]

Eine komplexe Zahl ist Null genau dann, wenn Real- und Imaginärteil beide Null sind. Ist $z = x + iy \neq 0$, so ist das multiplikativ Inverse gegeben durch

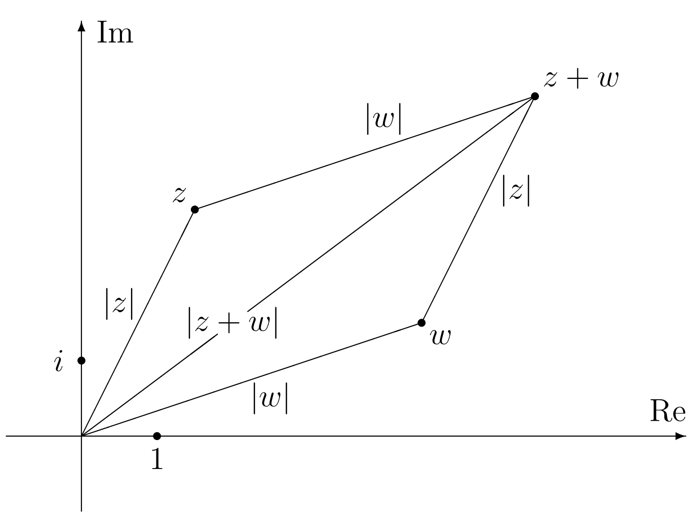
$$\frac{1}{z} = \frac{1}{x+iy} = \frac{x-iy}{(x+iy)(x-iy)} = \frac{\overline{z}}{|z|^2}.$$

5.3. Zur anschaulichen Vorstellung: Man stellt sich komplexe Zahlen gerne in der Ebene vor, also $x + iy$ als den Punkt $(x, y) \in \mathbb{R}^2$.

Addition: Addition mit $a + ib$ bedeutet eine Verschiebung.

Multiplikation: Multiplikation mit i bedeutet eine Drehung um 90^o nach links, Multiplikation mit $a + ib$ bedeutet also eine Drehstreckung.

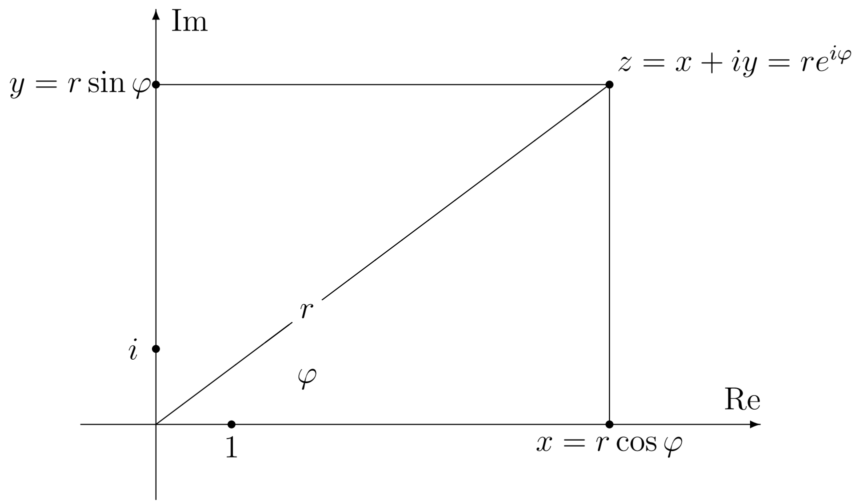
Dreiecksungleichung:



Polarkoordinaten und Eulerformel: Eine komplexe Zahl $z = x + iy$ (wobei $x, y \in \mathbb{R}$) kann man schreiben mithilfe ihres Betrages $r = |z|$ und des Winkels φ zur positiven x-Achse (dh mithilfe von *Polarkoordinaten*). Es ist nämlich $x = r\cos\varphi$ und $y = r\sin\varphi$ nach Schulmathematik. Verwendet man die *Eulerformel*

$$e^{i\varphi} = \cos\varphi + i\sin\varphi \quad (\varphi \in \mathbb{R}),$$

so erhält man $z = x + iy = r\cos\varphi + ir\sin\varphi = re^{i\varphi}$. Multiplikation mit z dreht dann um den Winkel φ.



Sind sin und cos (an dieser Stelle z.B. aus der Schule) bekannt, so kann man die Eulerformel als Definition von $e^{i\varphi}$ verwenden. Wir werden später jedoch umgekehrt vorgehen!

5.4. Polynome: Ein Polynom p (oder $p(z)$) mit komplexen Koeffizienten ist ein formaler Ausdruck $p(z) = a_n z^n + a_{n-1}z^{n-1} + \ldots + a_2 z^2 + a_1 z + a_0$ mit $n \in \mathbb{N}_0$ und $a_j \in \mathbb{C}$ für alle $j \in \{0, 1, \ldots, n\}$. Das Polynom heißt *reell*, wenn alle Koeffizienten a_j reell sind.

Das Polynom heißt *vom Grad n*, falls $a_n \neq 0$ gilt, und zusätzlich *normiert*, falls $a_n = 1$ ist.

Falls $a_n = 0$ ist, so ist auch $p(z) = a_{n-1}z^{n-1} + \ldots + a_2 z^2 + a_1 z + a_0$, dh führende Nullkoeffizienten kann man weglassen.

Es gilt insbesondere

$$p(z) \text{ hat den Grad } 0 \quad \Longleftrightarrow \quad p(z) = a_0 \text{ und } a_0 \neq 0.$$

Das Nullpolynom $p(z) = 0$ hat keinen Grad.

Die Menge aller Polynome mit komplexen Koeffizienten (in der “freien Variablen” z) bezeichnen wir mit $\mathbb{C}[z]$.

Horner-Schema: Bei der Berechnung von $p(z) = a_n z^n + \ldots + a_1 z + a_0$ für ein gegebenes $z \in \mathbb{C}$ geht man ökonomischerweise so vor:

$$p(z) = ((\ldots(((a_n \cdot z + a_{n-1}) \cdot z + a_{n-2}) \cdot z + a_{n-3})\ldots) \cdot z + a_1) \cdot z + a_0.$$

Das spart Multiplikationen und verringert so auch den Einfluss von Rundungsfehlern.

Definition: Ein $z_0 \in \mathbb{C}$ mit $p(z_0) = 0$ heißt *Nullstelle* des Polynoms p.

5.5. Polynomdivision: Seien $p, q \in \mathbb{C}[z] \setminus \{0\}$ Polynome vom Grad n bzw. $k \leq n$. Dann gibt es eindeutig bestimmte Polynome $m \in \mathbb{C}[z] \setminus \{0\}$ und $r \in \mathbb{C}[z]$ mit $\operatorname{Grad} m = n - k$ und $r = 0$ oder $\operatorname{Grad} r < k$ und $p = mq + r$ (*Division mit Rest*).

Beweis. Für $p(z) = a_n z^n + \dots$ und $q(z) = b_k z^k + \dots$ setze $c_{n-k} := a_n / b_k$. Dann ist $p_1(z) := p(z) - c_{n-k} z^{n-k} q(z)$ entweder $= 0$ (dann ist man fertig) oder hat Grad $n_1 \leq n - 1$. Ist $n_1 < k$, so ist man fertig, ansonsten wiederhole man den obigen Schritt mit p_1 statt p. Das Verfahren endet nach endlich vielen Schritten. $\square$

Beispiel: Wir wissen, dass $z_0 = i$ Nullstelle ist von $z^2 + 1$. Division von $p(z) = z^2 + 1$ durch $q(z) = z - i$ ergibt $m(z) = z + i$ mit Rest $r(z) = 0$. Es ist kein Zufall, dass diese Division aufgeht:

Satz: Ist $p \in \mathbb{C}[z]$ Polynom vom Grad $n \geq 1$ und ist z_0 Nullstelle von p, so gibt es ein Polynom $q \in \mathbb{C}[z]$ vom Grad $n - 1$ mit

$$p(z) = q(z) \cdot (z - z_0).$$

Dabei heißt $z - z_0$ *Linearfaktor*.

Beweis. Wir dividieren $p(z)$ durch das Polynom $z - z_0$ (das den Grad 1 hat) und erhalten $p(z) = q(z)(z - z_0) + r(z)$, wobei $\operatorname{Grad} q = n - 1$ und $r = 0$ oder $r(z) = r_0 \neq 0$. Wegen $p(z_0) = 0$ ist $r(z_0) = 0$, also $r = 0$. $\square$

Definition: Die *Vielfachheit* (Vfh) einer Nullstelle z_0 von p gibt an, wie oft man $p(z)$ durch den Linearfaktor $z - z_0$ dividieren kann.

Folgerung: Ein Polynom vom Grad n hat höchstens n Nullstellen.

5.6. Fundamentalsatz der Algebra: Jedes Polynom vom Grad $n \geq 1$ hat mindestens eine Nullstelle in $\mathbb{C}$. (ohne Beweis)

Folgerung: Ist $p(z)$ normiertes Polynom vom Grad $n \geq 1$, so gibt es $z_1, z_2, \dots, z_n \in \mathbb{C}$ mit

$$p(z) = (z - z_1) \cdot (z - z_2) \cdot \ldots \cdot (z - z_n).$$

Eine feste Nullstelle z_0 von $p(z)$ kommt dabei in $z_1, z_2, \dots, z_n$ so oft vor, wie ihre Vielfachheit angibt.

Beispiele: (1) $p(z) = z^3 + z^2 + z + 1$. Eine Nullstelle ist $z_0 = -1$, dann ist $z - z_0 = z + 1$. Durch Polynomdivision findet man $(z^3+z^2+z+1) : (z+1) = z^2+1 = z^2-i^2 = (z-i)(z+i)$. Also ist $p(z) = (z + 1)(z - i)(z + i)$, die Nullstellen sind -1, i und $-i$ und haben jeweils die Vielfachheit 1.

Mittels Polynomdivision (oder auch schon nach 4.11(1)) hat man außerdem $z^4 - 1 = (z - 1)(z^3 + z^2 + z + 1)$, also sind $1, -1, i, -i$ (jeweils mit Vielfachheit 1) die Nullstellen von $p(z) = z^4 - 1$.

(2) $p(z) = z^2 - 2z + 1 = (z - 1)^2$. Einzige Nullstelle ist 1 (mit Vielfachheit 2).

Zur Bedeutung von Polynomen und ihren Nullstellen: Betrachtet man lineare zeitinvariante Systeme (LTI-Systeme) mit Input und Output, so wird (nach *Laplacetransformation*) deren Verhalten durch eine *Übertragungsfunktion* beschrieben: man erhält die Laplacetransformierte des Outputsignals durch Multiplikation der Laplacetransformierten des Inputsignals mit der Übertragungsfunktion. In den einfachsten Fällen hat die Übertragungsfunktion die Gestalt $\frac{1}{p(z)}$, wobei $p \in \mathbb{C}[z]$ ein Polynom ist. Die Nullstellen von p sind Polstellen der Übertragungsfunktion und deren Lage in der komplexen Zahlenebene gibt wichtige Informationen über das System (z.B. zur Stabilität).

6 Folgen und Konvergenz

6.1. Definition: Eine *reelle* [*komplexe*] *Zahlenfolge* ist eine Abbildung $\mathbb{N} \to \mathbb{R}$, $n \mapsto a_n$ [bzw. $\mathbb{N} \to \mathbb{C}$, $n \mapsto a_n$].

Wir schreiben $(a_n)_{n\in\mathbb{N}}$, $(a_n)_{n=1}^{\infty}$ oder kurz (a_n) oder auch $(a_1, a_2, a_3, \ldots)$.

Beispiele: (1) $a_n = \frac{1}{n}$ für alle $n \in \mathbb{N}$, also $(a_n) = (1, \frac{1}{2}, \frac{1}{3}, \ldots)$.

(2) $a_n = (-1)^n$ für alle $n \in \mathbb{N}$, also $(a_n) = (-1, 1, -1, 1, -1, \ldots)$.

(3) $a_n = i^n$ für alle $n \in \mathbb{N}$, also $(a_n) = (i, -1, -i, 1, i, -1, -i, 1, \ldots)$.

Der Begriff der *Konvergenz* ist für die Analysis von zentraler Bedeutung, z.B. für Stetigkeit, Reihendarstellungen, Ableitungen, Integrale, Approximation, ... All dies baut auf dem Konvergenzbegriff für Folgen auf und Zahlenfolgen sind hier der einfachste Fall.

6.2. Konvergenz: Sei (a_n) eine Zahlenfolge und $a \in \mathbb{R}$ [bzw. $a \in \mathbb{C}$]. Wir sagen, dass (a_n) gegen a *konvergiert* und schreiben $\lim_{n\to\infty} a_n = a$, falls gilt:

$$\forall \varepsilon > 0 \; \exists \underbrace{n_0}_{= n_0(\varepsilon)} \in \mathbb{N} \; \forall n \geq n_0 : |a_n - a| < \varepsilon.$$

Das bedeutet: “die a_n kommen a beliebig nahe” oder “der Abstand $|a_n - a|$ wird beliebig klein”.

Die Zahl a heißt dann *Limes* oder *Grenzwert* der Folge (a_n).

Eine Folge (a_n) heißt *konvergent*, falls es ein $a \in \mathbb{R}$ [bzw. $a \in \mathbb{C}$] so gibt, dass (a_n) gegen a konvergiert. Eine Folge, die nicht konvergiert, heißt *divergent*.

Beispiele: (1) $\lim_{n\to\infty} \frac{1}{n} = 0$: Sei $\varepsilon > 0$. Nach 4.7(4) finden wir ein $n_0 \in \mathbb{N}$ mit $\frac{1}{n_0} < \varepsilon$. Für jedes $n \geq n_0$ gilt dann:

$$\left|\frac{1}{n} - 0\right| = \frac{1}{n} \leq \frac{1}{n_0} < \varepsilon.$$

Wir haben $\lim_{n\to\infty} \frac{1}{n} = 0$ nachgewiesen.

Schreibweisen: Statt $\lim_{n\to\infty} a_n = a$ schreibt man auch $\lim a_n = a$, $a_n \to a$ $(n \to \infty)$ oder $a_n \to a$.

(2) Für jede Zahlenfolge (a_n) und jedes a gilt:

$$\lim a_n = a \iff \lim |a_n - a| = 0.$$

Insbesondere ist für $a = 0$:

$$\lim a_n = 0 \iff \lim |a_n| = 0.$$

(3) Die Folge $(a_n)_{n\in\mathbb{N}} := ((-1)^n)_{n\in\mathbb{N}}$ ist divergent: Wir müssen zeigen, dass kein $a \in \mathbb{R}$ Grenzwert von (a_n) ist, dh dass für alle $a \in \mathbb{R}$ gilt:

$$\exists \varepsilon > 0 \, \forall n_0 \in \mathbb{N} \, \exists n \in \mathbb{N} : n \geq n_0 \text{ und } |a_n - a| \geq \varepsilon.$$

Zu gegebenem $a \in \mathbb{R}$ müssen wir also ein $\varepsilon > 0$ so finden, dass (wenn n größer wird) immer mal wieder $|a_n - a| \geq \varepsilon$ gilt.
Sei also $a \in \mathbb{R}$. Setze

$$\varepsilon := \begin{cases} 1/2, & \text{falls } a \in \{1, -1\} \\ 1/2 \cdot \min\{|1-a|, |-1-a|\}, & \text{falls } a \notin \{1, -1\} \end{cases}.$$

Dann gilt $|a_n - a| \geq \varepsilon$ für unendlich viele $n \in \mathbb{N}$ (im Fall $a \notin \{-1, 1\}$ gilt sogar $|a_n - a| \geq \varepsilon$ für alle $n \in \mathbb{N}$). Folglich konvergiert (a_n) nicht gegen a. Da a beliebig war, ist (a_n) divergent.

(4) Für $b \in \mathbb{C}$ gilt:

$$\lim_{n\to\infty} b^n = 0 \iff |b| < 1.$$

Denn für $|b| \geq 1$ gilt $|b^n| = |b|^n \geq 1$ für alle $n \in \mathbb{N}$ und (b^n) konvergiert nicht gegen Null. Für $|b| < 1$ sei $\varepsilon > 0$. Nach 4.11(4) finden wir ein $n_0 \in \mathbb{N}$ mit $|b|^{n_0} < \varepsilon$. Für $n \geq n_0$ gilt dann $0 \leq |b^n| = |b|^n \leq |b|^{n_0} < \varepsilon$. Also ist $\lim_n b^n = 0$.

(5) Zu jedem $r \in \mathbb{R}$ gibt es eine Folge (q_n) mit $q_n \to r$ und $q_n \in \mathbb{Q}$ für alle $n \in \mathbb{N}$. Denn nach 4.9 finden wir zu jedem $n \in \mathbb{N}$ ein $q_n \in (r - 1/n, r + 1/n) \cap \mathbb{Q}$. Wegen $|q_n - r| < 1/n$ für alle $n \in \mathbb{N}$ folgt dann $q_n \to r$.

Umformulierung: Sei $a \in \mathbb{R}$ [oder $a \in \mathbb{C}$] und $\varepsilon > 0$. Dann heißt

$$U_\varepsilon(a) := \{x \in \mathbb{R} : |x - a| < \varepsilon\} \quad [\text{bzw.} \quad U_\varepsilon(a) := \{z \in \mathbb{C} : |z - a| < \varepsilon\}]$$

die *ε-Umgebung von a in* $\mathbb{R}$ [bzw. *in* $\mathbb{C}$].

Die ε-Umgebung von $a \in \mathbb{R}$ in $\mathbb{R}$ ist das Intervall $(a - \varepsilon, a + \varepsilon)$.

Wir sagen, dass eine Eigenschaft "*für fast alle (ffa) $n \in \mathbb{N}$ gilt*", falls sie für alle bis auf endlich viele $n \in \mathbb{N}$ gilt.

Beispiel: Für fast alle $n \in \mathbb{N}$ gilt $n > 5$.

Sind eine Folge (a_n) und eine Zahl a gegeben, so gilt:

$$\lim_{n\to\infty} a_n = a \iff \text{für jedes } \varepsilon > 0 \text{ gilt: } a_n \in U_\varepsilon(a) \text{ für fast alle } n \in \mathbb{N}.$$

Bei Konvergenzfragen kommt es also auf endlich viele Folgenglieder nicht an.

Für $p \in \mathbb{Z}$ bezeichnet man auch $(a_n)_{n=p}^\infty$ als Folge.

Bemerkungen: (a) Der Grenzwert einer Zahlenfolge ist eindeutig bestimmt.

(b) Eine konvergente Folge (a_n) ist *beschränkt*, dh die Menge $\{a_n : n \in \mathbb{N}\}$ ist beschränkt.

(c) Ist (z_n) eine komplexe Zahlenfolge, so ist (z_n) genau dann konvergent, wenn die reellen Zahlenfolgen $(\operatorname{Re} z_n)$ und $(\operatorname{Im} z_n)$ konvergent sind. Genauer gilt für $z_0 \in \mathbb{C}$:

$$z_n \to z_0 \iff \operatorname{Re} z_n \to \operatorname{Re} z_0 \text{ und } \operatorname{Im} z_n \to \operatorname{Im} z_0$$

Das liegt an der Abschätzung (vgl. 5.2)

$$\max\{|\operatorname{Re} z_n - \operatorname{Re} z_0|, |\operatorname{Im} z_n - \operatorname{Im} z_0|\} \leq |z_n - z_0| \leq |\operatorname{Re} z_n - \operatorname{Re} z_0| + |\operatorname{Im} z_n - \operatorname{Im} z_0|.$$

Konvergiert etwa (z_n) gegen $z_0 \in \mathbb{C}$, so zeigt die linke Ungleichung, dass $\operatorname{Re} z_n \to \operatorname{Re} z_0$ und $\operatorname{Im} z_n \to \operatorname{Im} z_0$. Konvergiert hingegen $(\operatorname{Re} z_n)$ gegen $a \in \mathbb{R}$ und $(\operatorname{Im} z_n)$ gegen $b \in \mathbb{R}$, so setze $z_0 := a + ib$, und die rechte Ungleichung zeigt $z_n \to z_0$.

(d) Zur Beruhigung: Ist (a_n) eine reelle Zahlenfolge und $a \in \mathbb{C}$ mit $a_n \to a$, so gilt $a \in \mathbb{R}$. Wäre nämlich $a \in \mathbb{C} \setminus \mathbb{R}$, so hätte man für jedes $n \in \mathbb{N}$:

$$|a_n - a| \geq |\operatorname{Im} a_n - \operatorname{Im} a| = |\operatorname{Im} a| =: \varepsilon > 0,$$

im Widerspruch zu $|a_n - a| \to 0$.

Wir beschränken uns deshalb weitgehend auf reelle Zahlenfolgen.

6.3. Grenzwertsätze: Seien (a_n), (b_n), (α_n) und (c_n) reelle Folgen und $a, b \in \mathbb{R}$.

(1) ($|a_n - a| \leq \alpha_n$ für fast alle $n \in \mathbb{N}$ und $\alpha_n \to 0$) $\Longrightarrow a_n \to a$.

(2) $a_n \to a \Longrightarrow |a_n| \to |a|$.

(3) $a_n \to a$ und $b_n \to b$ und $a_n \leq b_n$ für fast alle $n \Longrightarrow a \leq b$.

(4) $a_n \to a$ und $b_n \to a$ und $a_n \leq c_n \leq b_n$ für fast alle $n \Rightarrow c_n \to a$.

(5) Gilt $a_n \to a$ und $b_n \to b$, so gilt $a_n + b_n \to a + b$ und $a_n \cdot b_n \to a \cdot b$. Ist $b \neq 0$, so ist $b_n \neq 0$ für fast alle n und es gilt $\frac{a_n}{b_n} \to \frac{a}{b}$.

Die Eigenschaften (1), (2) und (5) gelten auch für komplexe Zahlenfolgen.

Beweis. (1) ist leicht. Wir haben das auch schon in Beispiel 6.2(5) und Bemerkung 6.2(c) verwendet.

Für (2) verwendet man $||a_n| - |a|| \leq |a_n - a|$ (vgl. 4.3(6)) und dann (1).

Zu (3): Sonst wäre $a > b$, und mit $\varepsilon := (a - b)/2$ ist $a_n > a - \varepsilon \geq b + \varepsilon > b_n$ für fast alle n, Widerspruch.

Zu (4): Ist $\varepsilon > 0$, so gilt für fast alle n: $a_n, b_n \in U_\varepsilon(a)$, also $a - \varepsilon < a_n \leq c_n \leq b_n < a + \varepsilon$.

Zu (5): Sei $\varepsilon > 0$. Dann $|a_n - a| < \varepsilon/2$ und $|b_n - b| < \varepsilon/2$ für fast alle n. Also

$$|a_n + b_n - (a + b)| \leq |a_n - a| + |b_n - b| < \varepsilon/2 + \varepsilon/2 = \varepsilon$$

für fast alle n. Bei $(a_n \cdot b_n)$ verwendet man

$$|a_n b_n - ab| = |(a_n b_n - a_n b) + (a_n b - ab)| \leq |a_n||b_n - b| + |b||a_n - a|$$

und die Tatsache, dass die konvergente Folge (a_n) beschränkt ist, dh es gibt $M > 0$ mit $|a_n| \leq M$ für alle $n \in \mathbb{N}$. Die Behauptung folgt aus (1) und (5) für "+".

Die Aussage über $(\frac{a_n}{b_n})$ muss man nur für $a_n = 1$ zeigen. Sei dazu $b \neq 0$ und $\delta := |b|/2$. Dann ist $\delta > 0$ und wir finden $n_0 \in \mathbb{N}$ mit $|b_n - b| < \delta$ für alle $n \geq n_0$. Für $n \geq n_0$ gilt dann auch

$$|b_n| = |b - (b - b_n)| \geq |b| - |b - b_n| > 2\delta - \delta = \delta$$

und

$$\left|\frac{1}{b_n} - \frac{1}{b}\right| = \frac{|b - b_n|}{|b||b_n|} \leq \frac{1}{2\delta^2}\,|b_n - b|,$$

woraus die Behauptung folgt. □

6.4. Monotone Folgen: Eine reelle Folge (a_n) heißt

$$\begin{aligned}
&\textit{monoton wachsend}, && \text{falls} \quad \forall n \in \mathbb{N} : a_n \leq a_{n+1},\\
&\textit{monoton fallend}, && \text{falls} \quad \forall n \in \mathbb{N} : a_n \geq a_{n+1},\\
&\textit{streng monoton wachsend}, && \text{falls} \quad \forall n \in \mathbb{N} : a_n < a_{n+1},\\
&\textit{streng monoton fallend}, && \text{falls} \quad \forall n \in \mathbb{N} : a_n > a_{n+1}.
\end{aligned}$$

Satz: Ist eine monoton wachsende [bzw. fallende] reelle Folge (a_n) beschränkt, so konvergiert sie, und zwar gegen $\sup\{a_n : n \in \mathbb{N}\}$ [bzw. $\inf\{a_n : n \in \mathbb{N}\}$].

Beweis. Sei (a_n) monoton wachsend und $s := \sup\{a_n : n \in \mathbb{N}\}$. Zu $\varepsilon > 0$ finden wir $n_0 \in \mathbb{N}$ mit $a_{n_0} > s - \varepsilon$. Für $n \geq n_0$ gilt dann $s - \varepsilon < a_{n_0} \leq a_n \leq s$, also auch $|a_n - s| < \varepsilon$. □

6.5. Beispiele: (1) $a_n \geq 0$ für alle $n \in \mathbb{N}$, $a_n \to a$ und $p \in \mathbb{N} \Rightarrow \sqrt[p]{a_n} \to \sqrt[p]{a}$.

Beweis. Wir stellen zunächst fest, dass für alle $y > x \geq 0$ gilt $\sqrt[p]{y} - \sqrt[p]{x} \leq \sqrt[p]{y - x}$. [Es ist nämlich nach dem binomischen Satz $(\sqrt[p]{y - x} + \sqrt[p]{x})^p = y - x + \underbrace{\ldots}_{\geq 0} + x \geq y$.]

Wir erhalten somit $|\sqrt[p]{a_n} - \sqrt[p]{a}| \leq \sqrt[p]{|a_n - a|}$ für jedes $n \in \mathbb{N}$. Setzen wir $b_n := |a_n - a|$, so gilt $b_n \to 0$ und es reicht zu zeigen, dass $\sqrt[p]{b_n} \to 0$. Zu $\varepsilon > 0$ finden wir n_0 mit $b_n < \varepsilon^p$ für alle $n \geq n_0$. Es gilt dann $0 \leq \sqrt[p]{b_n} < \varepsilon$ für alle $n \geq n_0$. □

(2) $\sqrt[n]{n} \to 1$: Setze $a_n := \sqrt[n]{n} - 1$. Dann ist $a_n \geq 0$ für alle n, und für $n \geq 2$ gilt:

$$n = (1 + a_n)^n = \sum_{k=0}^{n} \binom{n}{k} a_n^k \geq \binom{n}{2} a_n^2 = \frac{n(n-1)}{2}\, a_n^2,$$

also $0 \leq a_n \leq \sqrt{2}/\sqrt{n-1}$. Es folgt $a_n \to 0$.

Für $c > 0$ gilt $\sqrt[n]{c} \to 1$: Für $c \geq 1$ ist $1 \leq \sqrt[n]{c} \leq \sqrt[n]{n}$ für fast alle n, und für $c \in (0,1)$ ist $1/\sqrt[n]{n} \leq \sqrt[n]{c} \leq 1$ für fast alle n. Wende nun 6.3(4) an.

(3) Konvergiert die Folge (a_n) so gilt $a_{n+1} - a_n \to 0$. Denn es gilt auch $a_{n+1} \to 0$, und $a_{n+1} - a_n \to 0$ folgt aus Satz 6.3(5).

Es gilt aber auch z.B. $\sqrt{n+1} - \sqrt{n} \to 0$. Für jedes $n \in \mathbb{N}$ ist nämlich

$$\sqrt{n+1} - \sqrt{n} = \frac{(\sqrt{n+1} - \sqrt{n})(\sqrt{n+1} + \sqrt{n})}{\sqrt{n+1} + \sqrt{n}} = \frac{(n+1) - n}{\sqrt{n+1} + \sqrt{n}} \leq \frac{1}{\sqrt{n}},$$

woraus $\sqrt{n+1} - \sqrt{n} \to 0$ folgt, obwohl $(\sqrt{n})_{n\in\mathbb{N}}$ nicht konvergiert.

(4) Sei $z \in \mathbb{C}$ und $s_n := \sum_{k=0}^{n} z^k = 1 + z + z12 + \ldots + z^n$ für jedes $n \in \mathbb{N}$. Die Folge (s_n) konvergiert genau dann, wenn $|z| < 1$ ist. In diesem Falle gilt $\lim s_n = (1-z)^{-1}$. Für $|z| \geq 1$ ist nämlich $|s_{n+1} - s_n| = |z|^{n+1} \geq 1$, und (s_n) konvergiert nicht (verwende (3) und Bemerkung 6.2(c)). Für $|z| < 1$ ist $s_n = \frac{1-z^{n+1}}{1-z}$ nach 4.11(1) und also $s_n \to (1-z)^{-1}$ nach Beispiel 6.2(4).

6.6. Teilfolgen: Ist (a_n) eine Folge und $k : \mathbb{N} \to \mathbb{N}$ eine Abbildung mit $k(n) < k(n+1)$ für alle $n \in \mathbb{N}$ (dh $(k(n))$ ist eine streng monotone Folge natürlicher Zahlen), so heißt die Folge $(a_{k(n)})$ *Teilfolge (TF)* von (a_n).

Eine Zahl b heißt *Häufungswert (HW) der Folge* (a_n), falls es eine Teilfolge von (a_n) gibt, die gegen b konvergiert.

Man beachte, dass für eine streng monotone Folge $(k(n))_{n\in\mathbb{N}}$ stets $k(n) \geq n$ für jedes $n \in \mathbb{N}$ ist.

Beispiel: $(a_2, a_4, a_6, \ldots)$ ist Teilfolge von (a_n), hier ist $k(n) = 2n$ für alle $n \in \mathbb{N}$. $(a_3, a_1, a_7, a_5, \ldots)$ ist hingegen keine Teilfolge von (a_n).

Bemerkung: Eine Zahl b ist Häufungswert der Folge (a_n) genau dann, wenn für jedes $\varepsilon > 0$ gilt, dass $a_k \in U_\varepsilon(b)$ ist für unendlich viele $k \in \mathbb{N}$.

Beispiele: (1) Gilt $a_n \to a$, so konvergiert jede Teilfolge von (a_n) gegen a, dh a ist einziger Häufungswert von (a_n).

(2) Die Folge $((-1)^n)$ hat genau die Häufungswerte 1 und -1: Wegen $a_{2n} \to 1$ und $a_{2n+1} \to -1$ sind 1 und -1 Häufungswerte. Ist $b \in \mathbb{R} \setminus \{1, -1\}$ und $\varepsilon := \min\{|1-b|, |-1-b|\}/2$, so ist $\varepsilon > 0$ und in $U_\varepsilon(b)$ liegen keine Folgenglieder. Somit ist b kein Häufungswert von (a_n).

(3) Die Folge (a_n) mit $a_n = (1 + (-1)^n)^n$ hat nur den Häufungswert 0. Es gilt nämlich $a_{2n-1} = 0 \to 0$, und $(a_{2n}) = (2^n)$ hat keinen Häufungswert, da jede Teilfolge unbeschränkt ist und somit nicht konvergieren kann.

Satz (Bolzano-Weierstraß): Jede beschränkte Zahlenfolge hat eine konvergente Teilfolge, dh jede beschränkte Zahlenfolge hat einen Häufungswert.

Beweis. Ist $(c_m)_{m\in\mathbb{N}}$ beschränkt, so finden wir $I_1 := [a_1, b_1]$, das alle c_m enthält. Setze $k(1) := 1$, so dass $c_{k(1)} \in I_1$. Ist $I_n = [a_n, b_n]$ konstruiert und $k(n)$ gefunden, so definiere $I_{n+1} := [a_{n+1}, b_{n+1}]$ als linke Hälfte von I_n, falls diese unendlich viele der c_m enthält, sonst als die rechte Hälfte von I_n. Dann findet man $k(n+1) > k(n)$ mit $c_{k(n+1)} \in I_{n+1}$. Offenbar ist (a_n) monoton wachsend und beschränkt, also $a_n \to a$. Weiter ist (b_n) monoton fallend und beschränkt, also $b_n \to b$. Wegen $b_n - a_n = 2^{1-n}(b_1 - a_1) \to 0$, gilt $a = b$. Schließlich folgt $c_{k(n)} \to a$ wegen 6.3(4) (Einschachteln). $\square$

Bemerkung: Der Satz gilt auch für komplexe Zahlenfolgen. Ist (z_n) eine beschränkte Folge in $\mathbb{C}$, so findet man zunächst eine konvergente Teilfolge $(\operatorname{Re} z_{k(n)})$ der Realteile. Da $(\operatorname{Im} z_{k(n)})$ beschränkt ist, findet man davon eine konvergente Teilfolge $(\operatorname{Im} z_{k(l(n))})$. Dann ist $(z_{(k\circ l)(n)})$ eine konvergente Teilfolge von (z_n).

6.7. Rechnen mit ∞: Sei (a_n) eine reelle Folge.

$$\begin{aligned} a_n \to \infty \quad &:\Longleftrightarrow \quad \forall K \in \mathbb{R}\ \exists n_0 \in \mathbb{N}\ \forall n \geq n_0 : a_n > K, \\ a_n \to -\infty \quad &:\Longleftrightarrow \quad \forall K \in \mathbb{R}\ \exists n_0 \in \mathbb{N}\ \forall n \geq n_0 : a_n < K. \end{aligned}$$

Also $a_n \to \infty$ $[a_n \to -\infty]$, falls für jedes $K \in \mathbb{R}$ gilt, dass $a_n > K$ $[a_n < K]$ für fast alle $n \in \mathbb{N}$.

Achtung: Man redet hier nicht von "Konvergenz" und auch nicht von "Grenzwert", denn $\pm\infty \notin \mathbb{R}$, dh ∞ und $-\infty$ sind keine **Zahlen**. Man sagt jedoch etwa "die Folge a_n geht gegen unendlich" und schreibt auch $\lim_n a_n = \infty$.

Bemerkung: (a) Gilt $a_n \to \pm\infty$, so folgt $|a_n| \to \infty$ und $1/a_n \to 0$. [Der erste Teil ist klar. Es gelte $|a_n| \to \infty$. Zu $\varepsilon > 0$ finden wir n_0 mit $\forall n \geq n_0 : |a_n| > 1/\varepsilon$. Es gilt dann $\forall n \geq n_0 : |1/a_n| < \varepsilon$.] Ist $a_n > 0$ für alle n und $a_n \to 0$, so folgt $1/a_n \to \infty$.

(b) Ist die reelle Folge (a_n) nicht nach oben [unten] beschränkt, so gibt es eine Teilfolge $(a_{k(n)})$ von (a_n) mit $a_{k(n)} \to \infty$ [bzw. mit $a_{k(n)} \to -\infty$].

(c) Für jede monoton wachsende [monoton fallende] Folge (a_n) gibt es $a \in \mathbb{R} \cup \{\infty\}$ [bzw. $a \in \mathbb{R} \cup \{-\infty\}$] mit $a_n \to a$.

Konventionen: Für alle $a \in \mathbb{R}$ gilt $-\infty < a < \infty$.

Man setzt für $a \in \mathbb{R}$: $a + \infty = \infty$, $a - \infty = -\infty$, sowie für $a > 0$: $a \cdot \infty = \infty$, $a \cdot (-\infty) = -\infty$ und für $a < 0$: $a \cdot \infty = -\infty$, $a \cdot (-\infty) = \infty$.

Außerdem setzt man $\infty + \infty = \infty$, $-\infty - \infty = -\infty$ und $\infty \cdot \infty = \infty$, $\infty \cdot (-\infty) = -\infty$, $(-\infty) \cdot (-\infty) = \infty$.

Achtung: Die Ausdrücke $0 \cdot \infty$, $0 \cdot (-\infty)$ und $\infty - \infty$ sind **nicht definiert!**

Regeln: Seien (a_n), (b_n) reelle Folgen mit $a_n \to a$ und $b_n \to b$, wobei $a, b \in \mathbb{R} \cup \{\infty, -\infty\}$. Dann gilt:

$$\begin{aligned} a_n + b_n \to a + b, &\quad \text{falls} \quad a + b \text{ definiert ist,} \\ a_n \cdot b_n \to a \cdot b, &\quad \text{falls} \quad a \cdot b \text{ definiert ist.} \end{aligned}$$

Beachte, dass Bemerkung (a) oben das Verhalten von $(1/a_n)$ beschreibt.

Definition: Sei $\emptyset \neq M \subseteq \mathbb{R}$ eine Menge.

$$\begin{aligned} \sup M = \infty \quad &:\Longleftrightarrow \quad M \text{ ist nicht nach oben beschränkt,} \\ \inf M = -\infty \quad &:\Longleftrightarrow \quad M \text{ ist nicht nach unten beschränkt.} \end{aligned}$$

Bemerkung: Manchmal findet man auch $\sup \emptyset := -\infty$ und $\inf \emptyset := \infty$.

Bemerkung: Man hat also für $\emptyset \neq M \subseteq \mathbb{R}$:

$$\begin{aligned} \sup M = \infty \quad &\Longleftrightarrow \quad \forall K > 0 \, \exists x \in M : x > K, \\ \inf M = -\infty \quad &\Longleftrightarrow \quad \forall K > 0 \, \exists x \in M : x < -K. \end{aligned}$$

Beispiele: $\sup \mathbb{N} = \infty$, $\inf \mathbb{Z} = -\infty$.

6.8. Limes superior und Limes inferior: Sei (a_n) eine reelle Folge. Ist (a_n) nach oben beschränkt, so ist die durch $b_n := \sup\{a_k : k \geq n\}$ definierte Folge (b_n) monoton fallend und wir definieren

$$\limsup_{n\to\infty} a_n := \lim_{n\to\infty} b_n = \lim_{n\to\infty} \sup\{a_k : k \geq n\}.$$

Ist (a_n) nach unten beschränkt, so ist die durch $c_n := \inf\{a_k : k \geq n\}$ definierte Folge (c_n) monoton wachsend und wir definieren

$$\liminf_{n\to\infty} a_n := \lim_{n\to\infty} c_n = \lim_{n\to\infty} \inf\{a_k : k \geq n\}.$$

Ist (a_n) nicht nach oben [unten] beschränkt, setzen wir $\limsup_{n\to\infty} a_n := \infty$ [bzw. $\liminf_{n\to\infty} a_n := -\infty$].

Schreib- und Sprechweisen: Man schreibt auch kurz $\limsup_n a_n$ oder $\limsup a_n$, sowie $\overline{\lim}_{n\to\infty} a_n$, $\overline{\lim}_n a_n$, $\overline{\lim} a_n$ und spricht vom "oberen Limes", entsprechend für $\liminf$ mit $\underline{\lim}$ ("unterer Limes").

Bemerkung: (a) Es gibt stets Teilfolgen $(a_{k(n)})$ und $(a_{l(n)})$ von (a_n) mit

$$a_{k(n)} \to \limsup_{m\to\infty} a_m \quad (n \to \infty), \qquad\qquad a_{l(n)} \to \liminf_{m\to\infty} a_m \quad (n \to \infty).$$

(b) Ist (a_n) beschränkt, so ist $\limsup a_n$ der **größte** und $\liminf a_n$ ist der **kleinste** Häufungswert von (a_n).

Beispiele: (1) Für $a_n := (-n)^n$ ist $\limsup a_n = \infty$ und $\liminf a_n = -\infty$. Die Folge hat keine Häufungswerte.

(2) Für $a_n := (-1)^n$ ist $\limsup a_n = 1$ und $\liminf a_n = -1$.

(3) Ist $a_n \geq 0$ für alle $n \in \mathbb{N}$, so gilt $\lim a_n = 0$ genau dann, wenn $\limsup a_n = 0$ ist.

(4) Gilt $a_n \to a \in \mathbb{R} \cup \{\infty, -\infty\}$, so ist $\limsup a_n = \liminf a_n = a$.

Allgemeiner gilt:

> Ist (a_n) eine reelle Folge und
>
> $$A := \{a \in \mathbb{R} \cup \{\infty, -\infty\} : \text{ es gibt eine Teilfolge } (a_{k(n)}) \text{ mit } a_{k(n)} \to a\},$$
>
> so ist $\limsup_n a_n = \max A$ und $\liminf_n a_n = \min A$.

(5) Sei die Folge $(a_n)_{n\in\mathbb{N}}$ gegeben durch $a_n := \begin{cases} (-1)^{\frac{n}{2}}(1 - \frac{1}{n}) & , n \text{ gerade} \\ n(1 + (-1)^{\frac{n+1}{2}}) & , n \text{ ungerade} \end{cases}$. Dann gilt für $k \in \mathbb{N}$:

$$a_{4k} = 1 - \frac{1}{4k} \to 1, \qquad a_{4k+1} = (4k+1)(1-1)0 \to 0,$$
$$a_{4k+2} = -(1 - \frac{1}{4k+2}) \to -1, \qquad a_{4k+3} = (4k+3)(1+1) \to \infty.$$

Wir haben also hier $A = \{1, 0, -1, \infty\}$ und $\limsup_n a_n = \infty$, $\liminf_n a_n = -1$. Man sieht auch z.B. $\limsup_n a_{2n} = 1$, $\liminf_n a_{2n+1} = 0$.

6.9. Ergänzung: Cauchyfolgen: Sei (a_n) eine konvergente Zahlenfolge. Dann gilt:

$$(C) \qquad\qquad \forall \varepsilon > 0 \ \exists n_0 \in \mathbb{N} \ \forall n, m \geq n_0 : |a_n - a_m| < \varepsilon.$$

Beweis. Es gelte $a_n \to a$. Sei $\varepsilon > 0$. Dann finden wir $n_0 \in \mathbb{N}$ mit $\forall n \geq n_0 : |a_n - a| < \varepsilon/2$. Seien nun $n, m \geq n_0$. Dann gilt:

$$|a_n - a_m| = |a_n - a + (a - a_m)| \leq |a_n - a| + |a - a_m| < \varepsilon/2 + \varepsilon/2 = \varepsilon.$$

$\square$

Definition: Eine Folge (a_n), für welche die Eigenschaft (C) gilt, heißt *Cauchyfolge (CF)*.

Bemerkung: Jede Cauchyfolge (a_n) ist beschränkt.

Beweis. Wähle n_0 mit $\forall n, m \geq n_0 : |a_n - a_m| \leq 1$ und setze $M := \max\{|a_1|, |a_2|, \ldots, |a_{n_0}|\} + 1$. Ist nun $n \in \mathbb{N}$, so gilt $|a_n| \leq M$, falls $n \leq n_0$. Für $n > n_0$ ist

$$|a_n| \leq \underbrace{|a_n - a_{n_0}|}_{\leq 1} + |a_{n_0}| \leq M.$$

$\square$

Folgerung: In $\mathbb{R}$ und in $\mathbb{C}$ gilt, dass jede Cauchyfolge konvergiert.

Diese Eigenschaft heißt *Vollständigkeit*. Der Vorteil ist, dass man so Konvergenz zeigen kann, ohne den Grenzwert kennen zu müssen.

Beweis. Da eine Cauchyfolge (a_n) beschränkt ist, hat sie nach Bolzano-Weierstraß eine konvergente Teilfolge $a_{k(n)} \to a$. Die Eigenschaft (C) impliziert dann Konvergenz der gesamten Folge: Zu $\varepsilon > 0$ finden wir ein n_0 so, dass für alle $n, m \geq n_0$ gilt $|a_n - a_m| < \varepsilon/2$ und $|a_{k(n_0)} - a| < \varepsilon/2$. Für jedes $n \geq k(n_0)$ gilt dann wegen $k(n_0) \geq n_0$:

$$|a_n - a| \leq |a_n - a_{k(n_0)}| + |a_{k(n_0)} - a| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

$\square$

7 Reihen

7.1. Definition: Sei $(a_n)_{n\in\mathbb{N}}$ eine Folge und $s_N := \sum_{n=1}^N a_n = a_1 + a_2 + \ldots + a_N$ für jedes $N \in \mathbb{N}$. Die Folge $(s_N)_{N\in\mathbb{N}}$ heißt (*unendliche*) *Reihe* und wird mit $\sum_{n=1}^\infty a_n$ bezeichnet. Die Zahl s_N heißt N*-te Partialsumme* oder N*-te Teilsumme* von $\sum_{n=1}^\infty a_n$.

Die Reihe $\sum_{n=1}^\infty a_n$ heißt *konvergent* [*divergent*], falls die Folge $(s_N)_{N\in\mathbb{N}}$ konvergiert [bzw. divergiert].

Ist $\sum_{n=1}^\infty a_n$ konvergent, so heißt $\lim_{N\to\infty} s_N$ der *Reihenwert* von $\sum_{n=1}^\infty a_n$ und wird ebenfalls mit $\sum_{n=1}^\infty a_n$ bezeichnet.

Bemerkung: (a) Die Folge (a_n) kann hier reell oder komplex sein. Ist (a_n) komplex, so heißt die Reihe $\sum_{n=1}^\infty \operatorname{Re} a_n$ *Realteil* der komplexen Reihe $\sum_{n=1}^\infty a_n$ und $\sum_{n=1}^\infty \operatorname{Im} a_n$ heißt *Imaginärteil* von $\sum_{n=1}^\infty a_n$.

Eine komplexe Reihe konvergiert genau dann, wenn ihr Realteil **und** ihr Imaginärteil **beide** konvergieren. In diesem Fall ist

$$\sum_{n=1}^\infty a_n = \sum_{n=1}^\infty \operatorname{Re} a_n + i \sum_{n=1}^\infty \operatorname{Im} a_n,$$

also

$$\operatorname{Re}\left(\sum_{n=1}^\infty a_n\right) = \sum_{n=1}^\infty \operatorname{Re} a_n \quad \text{und} \quad \operatorname{Im}\left(\sum_{n=1}^\infty a_n\right) = \sum_{n=1}^\infty \operatorname{Im} a_n,$$

vgl. mit 6.2.

(b) Ist $p \in \mathbb{Z}$, so verfährt man für Folgen $(a_n)_{n=p}^\infty$ entsprechend, indem man $s_N := \sum_{n=p}^N$ für $N \ge p$ setzt und $(s_N)_{N=p}^\infty$ betrachtet. Ein wichtiger Fall ist hierbei $p = 0$.

Beispiele: (1) Die *geometrische Reihe* $\sum_{n=0}^\infty z^n$, wobei $z \in \mathbb{C}$, konvergiert genau dann, wenn $|z| < 1$. Für $|z| < 1$ ist $\sum_{n=0}^\infty z^n = (1-z)^{-1}$ (vgl. 6.5(4)).

(2) $\sum_{n=1}^\infty \frac{1}{n(n+1)}$. Für jedes $n \in \mathbb{N}$ gilt: $\frac{1}{n(n+1)} = \frac{1}{n} - \frac{1}{n+1}$, also ist für jedes $N \in \mathbb{N}$:

$$s_N = 1 - \frac{1}{2} + \frac{1}{2} - \frac{1}{3} + \ldots + \frac{1}{N} - \frac{1}{N+1} = 1 - \frac{1}{N+1} \to 1 \ (N \to \infty).$$

Somit ist $\sum_{n=1}^\infty \frac{1}{n(n+1)}$ konvergent und $\sum_{n=1}^\infty \frac{1}{n(n+1)} = 1$.

(3) Die *harmonische Reihe* $\sum_{n=1}^\infty \frac{1}{n}$. Für jedes $N \in \mathbb{N}$ ist

$$s_{2N} = \underbrace{1 + \frac{1}{2} + \ldots + \frac{1}{N}}_{=s_N} + \underbrace{\frac{1}{N+1}}_{\ge \frac{1}{2N}} + \ldots + \underbrace{\frac{1}{2N}}_{\ge \frac{1}{2N}} \ge s_N + N\frac{1}{2N} = s_N + \frac{1}{2}.$$

Also ist (s_N) divergent, denn $s_N \to s \in \mathbb{R}$ würde $s_{2N} - s_N \to s - s = 0$ implizieren im Widerspruch zu $s_{2N} - s_N \ge 1/2$ für alle $N \in \mathbb{N}$.

7.2. Satz: Seien (a_n) und (b_n) Folgen und $s_N := a_1 + \ldots + a_N$, $N \in \mathbb{N}$.

(1) **Monotoniekriterium:** Sind alle $a_n \geq 0$ und ist die Folge $(s_N)_{N\in\mathbb{N}}$ beschränkt, so konvergiert $\sum_{n=1}^{\infty} a_n$.

(2) Ist $\sum_{n=1}^{\infty} a_n$ konvergent und $p \in \mathbb{N}$, so konvergiert $\sum_{n=p+1}^{\infty} a_n$ und es gilt

$$\sum_{n=p+1}^{\infty} a_n = \sum_{n=1}^{\infty} a_n - \sum_{n=1}^{p} a_n.$$

(3) Ist $\sum_{n=1}^{\infty} a_n$ konvergent und definiert man $r_N := \sum_{n=N+1}^{\infty} a_n$ für jedes $N \in \mathbb{N}$, so gilt $r_N \to 0$ $(N \to \infty)$, dh die "Reihenendstücke" konvergieren gegen Null.

(4) Ist $\sum_{n=1}^{\infty} a_n$ konvergent, so folgt $a_n \to 0$.

(5) Sind $\sum_{n=1}^{\infty} a_n$ und $\sum_{n=1}^{\infty} b_n$ konvergent und $\alpha, \beta \in \mathbb{R}$ (oder in $\mathbb{C}$), so konvergiert $\sum_{n=1}^{\infty}(\alpha a_n + \beta b_n)$ und es gilt

$$\sum_{n=1}^{\infty}(\alpha a_n + \beta b_n) = \alpha \sum_{n=1}^{\infty} a_n + \beta \sum_{n=1}^{\infty} b_n.$$

Beweis. (1) folgt aus 6.4, angewandt auf die monton wachsende Folge $(s_N)_{N\in\mathbb{N}}$.

(2) Setzt man $\sigma_N := \sum_{n=p+1}^{N} a_n$ für $N \geq p+1$, so ist $\sigma_N = s_N - s_p \to \sum_{n=1}^{\infty} a_n - s_p$ $(N \to \infty)$.

(3) Wegen (2) gilt $r_N = \sum_{n=1}^{\infty} a_n - s_N \to \sum_{n=1}^{\infty} a_n - \sum_{n=1}^{\infty} a_n = 0$.

(4) Es ist $a_n = s_n - s_{n-1} \to 0$ nach 6.5(3).

(5) folgt aus 6.3(5). □

7.3. Absolut konvergente Reihen: Die Reihe $\sum_{n=1}^{\infty} a_n$ heißt *absolut konvergent*, falls $\sum_{n=1}^{\infty} |a_n|$ konvergiert (dh also, falls die Reihe über die Absolutbeträge der a_n konvergiert).

Beispiel: Für $|z| < 1$ ist die Reihe $\sum_{n=0}^{\infty} z^n$ absolut konvergent (wegen $|z^n| = |z|^n$ und Beispiel 7.1(1)).

Satz: Ist $\sum_{n=1}^{\infty} a_n$ absolut konvergent, so ist $\sum_{n=1}^{\infty} a_n$ konvergent und es gilt die "Dreiecksungleichung für Reihen":

$$\Big|\sum_{n=1}^{\infty} a_n\Big| \leq \sum_{n=1}^{\infty} |a_n|.$$

Beispiel: Die Reihe $\sum_{n=1}^{\infty} \frac{(-1)^n}{n}$ ist konvergent, aber **nicht** absolut konvergent. Setzt man $a_n := (-1)^n/n$ und $s_N := a_1 + \ldots + a_N$ für alle $n, N \in \mathbb{N}$, so stellt man fest, dass $(s_{2N})_{N\in\mathbb{N}}$ monoton fallend ist mit unterer Schranke $s_1 = -1$ und dass $(s_{2N-1})_{N\in\mathbb{N}}$ monoton wachsend ist mit oberer Schranke $s_2 = -1/2$. Nach 6.4 konvergieren beide Folgen, und zwar wegen $s_{2N} - s_{2N-1} = a_{2N} \to 0$ gegen denselben Grenzwert. Somit konvergiert $\sum_{n=1}^{\infty} \frac{(-1)^n}{n}$.

Ergänzung: Der Beweis des Satzes zur absoluten Konvergenz beruht auf der Beobachtung, dass für $N, M \in \mathbb{N}$ mit $N > M$ nach Dreiecksungleichung und 7.2(3) gilt:

$$\Big| \sum_{n=M+1}^{N} a_n \Big| \leq \sum_{n=M+1}^{N} |a_n| \leq \sum_{n=M}^{\infty} |a_n| \to 0 \ (M \to \infty),$$

und dem

Cauchykriterium für Reihen: Eine Reihe $\sum_{n=1}^{\infty} a_n$ konvergiert genau dann, wenn es zu jedem $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ gibt mit

$$\forall N > M \geq n_0 : \Big| \sum_{n=M+1}^{N} a_n \Big| < \varepsilon.$$

Setzt man nämlich $s_N := a_1 + a_2 + \ldots + a_N$ für $N \in \mathbb{N}$, so gilt wegen der Ergänzung 6.9:

$$\begin{aligned}
\sum_{n=1}^{\infty} a_n \text{ konvergiert} \quad &\Longleftrightarrow \quad (s_N)_{N\in\mathbb{N}} \text{ konvergiert} \Longleftrightarrow (s_N)_{N\in\mathbb{N}} \text{ ist Cauchyfolge} \\
&\Longleftrightarrow \quad \forall \varepsilon > 0 \, \exists n_0 \in \mathbb{N} \, \forall N, M \geq n_0 : |s_N - s_M| < \varepsilon \\
&\Longleftrightarrow \quad \forall \varepsilon > 0 \, \exists n_0 \in \mathbb{N} \, \forall N > M \geq n_0 : \Big| \sum_{n=M+1}^{N} a_n \Big| < \varepsilon.
\end{aligned}$$

7.4. Majoranten- und Minorantenkriterium: Seien (a_n) und (b_n) Folgen.

(1) Gilt $|a_n| \leq b_n$ für fast alle $n \in \mathbb{N}$ und ist $\sum_{n=1}^{\infty} b_n$ konvergent, so ist $\sum_{n=1}^{\infty} a_n$ absolut konvergent.

(2) Gilt $a_n \geq b_n \geq 0$ für fast alle $n \in \mathbb{N}$ und ist $\sum_{n=1}^{\infty} b_n$ divergent, so ist auch $\sum_{n=1}^{\infty} a_n$ divergent.

Beweis. Für (1) verwende 7.2(1). (2) folgt aus (1). □

Beispiel: Die Reihe $\sum_{n=1}^{\infty} \frac{1}{(n+1)^2}$ konvergiert, denn $0 \leq \frac{1}{(n+1)^2} \leq \frac{1}{n(n+1)}$ für alle $n \in \mathbb{N}$ und $\sum_{n=1}^{\infty} \frac{1}{n(n+1)}$ konvergiert nach 7.1(2). Somit konvergiert auch $\sum_{n=1}^{\infty} \frac{1}{n^2}$. Weiter ist $\sum_{n=1}^{\infty} \frac{1}{n^q}$ konvergent für jedes $q \in \mathbb{N}$ mit $q \geq 2$.

Bemerkung: (a) Ist (a_n) eine komplexe Folge, so ist $\sum_{n=1}^{\infty} a_n$ genau dann absolut konvergent, wenn beide Reihen $\sum_{n=1}^{\infty} \operatorname{Re} a_n$ und $\sum_{n=1}^{\infty} \operatorname{Im} a_n$ absolut konvergent sind. Ist $\sum_{n=1}^{\infty} a_n$ absolut konvergent, so schreibt man auch $\sum_{n=1}^{\infty} |a_n| < \infty$.

(b) Für eine reelle Folge (a_n) sei $b_n := \max\{a_n, 0\}$ und $c_n := \max\{-a_n, 0\}$, so dass $a_n = b_n - c_n$ und $|a_n| = b_n + c_n$ für jedes $n \in \mathbb{N}$. Dann ist $\sum_{n=1}^{\infty} a_n$ genau dann absolut konvergent, wenn beide Reihen $\sum_{n=1}^{\infty} b_n$ und $\sum_{n=1}^{\infty} c_n$ konvergieren. In diesem Falle ist $\sum_{n=1}^{\infty} a_n = \sum_{n=1}^{\infty} b_n - \sum_{n=1}^{\infty} c_n$.

7.5. Leibnizkriterium für alternierende Reihen: Sei (b_n) eine **monoton fallende** Folge mit $b_n \to 0$. Setzt man $a_n := (-1)^n b_n$, $n \in \mathbb{N}$, so konvergiert $\sum_{n=1}^{\infty} a_n$.

Beweis. Wie im Fall $b_n = 1/n$ im Beispiel in 7.3. □

Bemerkung: Beachte, dass aus den Voraussetzungen folgt: $b_n \geq 0$ für alle $n \in \mathbb{N}$.

Bemerkung: Monotonie ist hier wichtig! Setzt man nur voraus, dass $b_n \geq 0$ für alle $n \in \mathbb{N}$ und $b_n \to 0$, so ist die Aussage i.a. falsch (Beispiele in den Übungen).

7.6. Wurzelkriterium: Sei (a_n) eine Folge und $\alpha := \limsup_{n\to\infty} \sqrt[n]{|a_n|} \in [0, \infty]$.

(1) Ist $\alpha < 1$, so ist $\sum_{n=1}^{\infty} a_n$ absolut konvergent.

(2) Ist $\alpha > 1$, so divergiert $\sum_{n=1}^{\infty} a_n$.

Bemerkung: Ist $\alpha = 1$, so liefert das Wurzelkriterium keine Entscheidung. Denn $\sum_{n=1}^{\infty} \frac{1}{n}$ ist divergent und $\alpha = \limsup \sqrt[n]{1/n} = \lim 1/\sqrt[n]{n} = 1$, hingegen ist $\sum_{n=1}^{\infty} \frac{1}{n^2}$ konvergent mit $\alpha = \limsup \sqrt[n]{1/n^2} = \lim 1/(\sqrt[n]{n})^2 = 1$.

Beweis. Ist $\alpha < 1$, so wählen wir $\beta \in (\alpha, 1)$. Dann gilt $\sqrt[n]{|a_n|} \leq \beta$ für fast alle $n \in \mathbb{N}$, also $|a_n| \leq \beta^n$ für fast alle $n \in \mathbb{N}$. Wegen $\beta \in (0,1)$ konvergiert $\sum_{n=1}^{\infty} \beta^n$, und die Behauptung folgt aus 7.4(1).

Ist $\alpha > 1$, so wählen wir $\gamma \in (1, \alpha)$. Dann gilt $\sqrt[n]{|a_n|} \geq \gamma \geq 1$ für unendlich viele n, also $|a_n| \geq 1$ für unendlich viele $n \in \mathbb{N}$. Somit ist $a_n \not\to 0$, und nach 7.2(4) ist $\sum_{n=1}^{\infty} a_n$ divergent. □

Beispiel: Sei $p \in \mathbb{N}$. Wir untersuchen Konvergenz von $\sum_{n=1}^{\infty} n^p x^n$ für $x \in \mathbb{R}$ mit dem Wurzelkriterium. Hier ist $a_n = n^p x^n$ und

$$\sqrt[n]{|a_n|} = (\sqrt[n]{n})^p |x| \to |x| \ (n \to \infty).$$

Also ist $\sum_{n=1}^{\infty} n^p x^n$ für $|x| < 1$ absolut konvergent und für $|x| > 1$ divergent. Für $|x| = 1$ gilt $|a_n| = n^p \to \infty$ und $\sum_{n=1}^{\infty} n^p x^n$ ist nach 7.2(4) divergent.

7.7. Quotientenkriterium: Sei (a_n) eine Folge mit $a_n \neq 0$ für fast alle $n \in \mathbb{N}$ und $c_n := |\frac{a_{n+1}}{a_n}|$ für n mit $a_n \neq 0$.

(1) Ist $c_n \geq 1$ für fast alle $n \in \mathbb{N}$, so ist $\sum_{n=1}^{\infty} a_n$ divergent.

(2) Ist $\limsup c_n < 1$, so ist $\sum_{n=1}^{\infty} a_n$ absolut konvergent.
Ist $\liminf c_n > 1$, so divergiert $\sum_{n=1}^{\infty} a_n$.

Bemerkung: Konvergiert (c_n) gegen α, so ist für $\alpha < 1$ die Reihe $\sum_{n=1}^{\infty} a_n$ absolut konvergent und für $\alpha > 1$ ist $\sum_{n=1}^{\infty} a_n$ divergent. Im Falle $\alpha = 1$ ist **keine** allgemeine Aussage möglich (man betrachte wieder $a_n = \frac{1}{n}$ bzw. $a_n = \frac{1}{n^2}$).

Beweis. (1) Es sei $c_n \geq 1$ für $n \geq n_0$. Dann ist $|a_n| \geq |a_{n-1}| \geq \ldots \geq |a_{n_0}| > 0$ für alle $n \geq n_0$, dh $|a_n| \not\to 0$. Nach 7.2(4) divergiert $\sum_{n=1}^{\infty} a_n$.

(2) Ist $\alpha := \limsup c_n < 1$, so wähle $\beta \in (\alpha, 1)$. Es gilt dann $c_n \leq \beta$ für fast alle $n \in \mathbb{N}$, dh wir finden n_0 mit $c_n \leq \beta$ für alle $n \geq n_0$. Es folgt

$$|a_n| \leq \beta |a_{n-1}| \leq \ldots \leq \beta^{n-n_0} |a_{n_0}| = \beta^n (|a_{n_0}| \beta^{-n_0})$$

für alle $n \geq n_0$. Wegen $\beta \in (0,1)$ konvergiert $\sum_{n=1}^{\infty} \beta^n$, und nach 7.4(1) ist $\sum_{n=1}^{\infty} a_n$ absolut konvergent.
Ist $\liminf c_n > 1$, so folgt $c_n \geq 1$ für fast alle n und $\sum_{n=1}^{\infty} a_n$ divergiert nach (1). $\square$

7.8. Die Exponentialreihe: Wir betrachten $\sum_{n=0}^{\infty} \frac{z^n}{n!}$ für $z \in \mathbb{C}$. Die Reihe konvergiert für $z = 0$: $\sum_{n=0}^{\infty} \frac{0^n}{n!} = 1$. Setzt man für $z \neq 0$: $a_n := z^n/n!$, so gilt

$$\left| \frac{a_{n+1}}{a_n} \right| = \frac{|z|}{n+1} \to 0.$$

Nach dem Quotientenkriterium konvergiert die Reihe also für jedes $z \in \mathbb{C}$ absolut. Insbesondere ist die Reihe $\sum_{n=0}^{\infty} \frac{x^n}{n!}$ für jedes $x \in \mathbb{R}$ absolut konvergent.

Definition: Die *Eulersche Zahl* ist $e := \sum_{n=0}^{\infty} \frac{1}{n!}$.

Satz: Es gilt $e = \lim_{n\to\infty} \left(1 + \frac{1}{n}\right)^n$.

Beweis. Für jedes $n \in \mathbb{N}$ gilt:

$$\begin{aligned} a_n := \left(1 + \frac{1}{n}\right)^n &= \sum_{k=0}^{n} \binom{n}{k} \frac{1}{n^k} = \sum_{k=0}^{n} \frac{1}{k!} \cdot \frac{n \cdot (n-1) \cdot \ldots \cdot (n-k+1)}{n \cdot n \cdot \ldots \cdot n} \\ &= \sum_{k=0}^{n} \frac{1}{k!} \cdot 1 \cdot \underbrace{\left(1 - \frac{1}{n}\right)}_{\leq 1} \cdot \underbrace{\left(1 - \frac{2}{n}\right)}_{\leq 1} \cdot \ldots \cdot \underbrace{\left(1 - \frac{k-1}{n}\right)}_{\leq 1} \leq \sum_{k=0}^{n} \frac{1}{k!} =: b_n. \end{aligned}$$

Für $n \to \infty$ erhalten wir $\limsup_n a_n \le e$. Andererseits ist für fixiertes $m \in \mathbb{N}$ und $n \ge m$:

$$a_n = \Big(1 + \frac{1}{n}\Big)^n \ge \sum_{k=0}^{m} \frac{1}{k!} \cdot 1 \cdot \Big(1 - \frac{1}{n}\Big) \cdot \Big(1 - \frac{2}{n}\Big) \cdot \ldots \cdot \Big(1 - \frac{k-1}{n}\Big),$$

woraus für $n \to \infty$ folgt: $\liminf_n a_n \ge b_m$. Für $m \to \infty$ erhalten wir $\liminf_n a_n \ge e$. Es folgt $a_n \to e$. $\square$

Bemerkung: Allgemeiner gilt $(1 + \frac{x}{n})^n \to \sum_{k=0}^\infty \frac{x^k}{k!}$ für jedes $x \in \mathbb{R}$ ($\to$ später).

Wir erhalten wegen $n! \ge 2^{n-1}$ für jedes $n \in \mathbb{N}$ die Abschätzung

$$2.5 = 1 + \frac{1}{1!} + \frac{1}{2!} < e < 1 + \sum_{n=1}^{\infty} 2^{1-n} = 1 + \sum_{k=0}^{\infty} 2^{-k} = 3.$$

Bemerkung: Die Folge $\big((1 + \frac{1}{n})^n\big)_{n \in \mathbb{N}}$ ist streng monoton wachsend. Dazu fixieren wir $n \in \mathbb{N}$ und verwenden (BU). Die Ungleichung

$$\Big(1 + \frac{1}{n}\Big)^n < \Big(1 + \frac{1}{n+1}\Big)^{n+1}$$

ist wegen $1 + \frac{1}{n} = \frac{n+1}{n}$ und $1 + \frac{1}{n+1} = \frac{n+2}{n+1}$ äquivalent zu

$$\frac{n+1}{n+2} \le \Big(\frac{n+2}{n+1} \frac{n}{n+1}\Big)^n,$$

also zu

$$1 - \frac{1}{n+2} \le \Big(1 - \frac{1}{(n+1)^2}\Big)^n.$$

Nach (BU) haben wir

$$\Big(1 - \frac{1}{(n+1)^2}\Big)^n \ge 1 - \frac{n}{(n+1)^2}.$$

Wegen $n(n+2) < (n+1)^2$ ist außerdem $\frac{n}{(n+1)^2} < \frac{1}{n+2}$ und dann auch

$$1 - \frac{n}{(n+1)^2} > 1 - \frac{1}{n+2}.$$

Also ist

$$\Big(1 - \frac{1}{(n+1)^2}\Big)^n \ge 1 - \frac{n}{(n+1)^2} > 1 - \frac{1}{n+2},$$

und strenge Monotonie ist somit gezeigt.

7.9. Umordnungen: Sei (a_n) eine Folge und $\varphi : \mathbb{N} \to \mathbb{N}$ eine bijektive Abbildung. Setze $b_n := a_{\varphi(n)}$ für jedes $n \in \mathbb{N}$. Dann heißt (b_n) [bzw. $\sum_{n=1}^{\infty} b_n$] eine *Umordnung* von (a_n) [bzw. von $\sum_{n=1}^{\infty} a_n$].

Beispiel: $(a_2, a_4, a_1, a_3, a_6, a_8, a_5, a_7, \ldots)$ ist eine Umordnung von (a_n) (aber **keine** Teilfolge von (a_n)!).

Satz: Sei (b_n) eine Umordnung von (a_n).

(1) Ist (a_n) konvergent, so konvergiert auch (b_n) und $\lim a_n = \lim b_n$.

(2) Ist $\sum_{n=1}^{\infty} a_n$ **absolut** konvergent, so ist auch $\sum_{n=1}^{\infty} b_n$ absolut konvergent und $\sum_{n=1}^{\infty} a_n = \sum_{n=1}^{\infty} b_n$.

Beweis. (1) Setze $a := \lim a_n$. Sei $\varepsilon > 0$. Dann gilt $|a_n - a| < \varepsilon$ für fast alle $n \in \mathbb{N}$, also auch $|a_{\varphi(n)} - a| < \varepsilon$ für fast alle $n \in \mathbb{N}$.

(2) Wegen der Bemerkung in 7.4 reicht es, den Fall $a_n \in \mathbb{R}$ und $a_n \geq 0$ zu betrachten. Für jedes $N \in \mathbb{N}$ gilt

$$\sum_{n=1}^{N} b_n \leq \sum_{k=1}^{\max \varphi(\{1,\ldots,N\})} a_k \quad \text{und} \quad \sum_{n=1}^{N} a_n \leq \sum_{k=1}^{\max \varphi^{-1}(\{1,\ldots,N\})} b_k.$$

Daraus folgt die Behauptung. □

Riemannscher Umordnungssatz (ohne Beweis): Sei $\sum_{n=1}^{\infty} a_n$ konvergent, aber **nicht** absolut konvergent. Dann gibt es für jedes $s \in \mathbb{R}$ eine Umordnung $\sum_{n=1}^{\infty} b_n$ von $\sum_{n=1}^{\infty} a_n$ mit $\sum_{n=1}^{\infty} b_n = s$. Es gibt außerdem divergente Umordnungen von $\sum_{n=1}^{\infty} a_n$.

Beispiele hierzu gibt es in den Übungen.

7.10. Das Cauchyprodukt: Seien $\sum_{n=0}^{\infty} a_n$ und $\sum_{n=0}^{\infty} b_n$ Reihen. Setze für jedes $n \in \mathbb{N}_0$:

$$c_n := \sum_{k=0}^{n} a_k b_{n-k} = a_0 b_n + a_1 b_{n-1} + \ldots + a_{n-1} b_1 + a_n b_0.$$

Die Reihe $\sum_{n=0}^{\infty} c_n$ heißt das *Cauchyprodukt* der Reihen $\sum_{n=0}^{\infty} a_n$ und $\sum_{n=0}^{\infty} b_n$.

Satz: Sind die Reihen $\sum_{n=0}^{\infty} a_n$ und $\sum_{n=0}^{\infty} b_n$ **absolut konvergent**, so ist auch ihr Cauchyprodukt $\sum_{n=0}^{\infty} c_n$ absolut konvergent und es gilt

$$\sum_{n=0}^{\infty} c_n = \left(\sum_{n=0}^{\infty} a_n \right) \left(\sum_{n=0}^{\infty} b_n \right).$$

Beweis. Auch hier reicht es, den Fall $a_n, b_n \in \mathbb{R}$ und $a_n \geq 0$, $b_n \geq 0$ zu betrachten. Es gilt dann für jedes $N \in \mathbb{N}$:

$$\sum_{n=0}^{N} c_n \leq \left(\sum_{n=0}^{N} a_n\right)\left(\sum_{n=0}^{N} b_n\right) \leq \sum_{n=0}^{2N} c_n,$$

woraus die Behauptung folgt. □

Beispiel: Sei $x \in \mathbb{R}$ mit $|x| < 1$. Dann konvergiert $\sum_{n=0}^{\infty} x^n$ absolut. Das Cauchyprodukt von $\sum_{n=0}^{\infty} x^n$ mit sich selber ist

$$\sum_{n=0}^{\infty}\left(\sum_{k=0}^{n} x^k x^{n-k}\right) = \sum_{n=0}^{\infty}(n+1)x^n.$$

Nach dem Satz konvergiert diese Reihe absolut und es gilt

$$\sum_{n=0}^{\infty}(n+1)x^n = \left(\sum_{n=0}^{\infty} x^n\right)^2 = \frac{1}{(1-x)^2}.$$

7.11. Die Exponentialfunktion: Da die Exponentialreihe nach 7.8 für jedes $z \in \mathbb{C}$ konvergiert, können wir durch

$$E : \mathbb{C} \to \mathbb{C}, \qquad z \mapsto E(z) := \sum_{n=0}^{\infty} \frac{z^n}{n!}$$

eine Funktion definieren, E heißt die *komplexe Exponentialfunktion.*

(0) Es gilt $E(0) = 1$ und $E(1) = e$ (nach 7.8).

(1) Für alle $z, w \in \mathbb{C}$ gilt $E(z)E(w) = E(z+w)$ (*Funktionalgleichung* der Exponentialfunktion).

[Cauchyprodukt für $a_n = z^n/n!$ und $b_n = w^n/n!$; man hat dann

$$c_n = \sum_{k=0}^{n} a_k b_{n-k} = \sum_{k=0}^{n} \frac{z^k}{k!}\frac{w^{n-k}}{(n-k)!} = \frac{1}{n!}\sum_{k=0}^{n}\binom{n}{k} z^k w^{n-k} = \frac{(z+w)^n}{n!}$$

nach dem binomischen Satz, und 7.10 gibt die Behauptung.]

(2) Für alle $z \in \mathbb{C}$ gilt $E(z) \neq 0$ und $E(z)^{-1} = E(-z)$, sowie $E(z)^n = E(nz)$ für alle $n \in \mathbb{Z}$.

[Es ist $1 = E(0) = E(z + (-z)) = E(z)E(-z)$, woraus $E(z) \neq 0$ und $E(z)^{-1} = E(-z)$ folgt. Der Rest folgt aus (1).]

(3) Für alle $x \in \mathbb{R}$ gilt $E(x) \in \mathbb{R}$ und $E(x) > 0$; für $x > 0$ gilt $E(x) > 1$.

[Sei $x \in \mathbb{R}$. Dann ist $E(x) \in \mathbb{R}$ klar und für $x > 0$ gilt

$$E(x) = \sum_{n=0}^{\infty} \underbrace{\frac{x^n}{n!}}_{\geq 0} \geq 1 + x > 1.$$

Also ist für $x < 0$ nach (2): $E(x) = E(-x)^{-1} \in (0, 1)$.]

(4) Für alle $x, y \in \mathbb{R}$ gilt: $x < y \Rightarrow E(x) < E(y)$.

[Ist $x < y$, so gilt nach (1) und (3):

$$E(y) = E(\underbrace{y - x}_{>0}) \underbrace{E(x)}_{>0} > E(x).]$$

(5) Es ist $\sup\{E(x) : x \in \mathbb{R}\} = \infty$ und $\inf\{E(x) : x \in \mathbb{R}\} = 0$.

[Für jedes $K > 0$ gilt $E(K) \geq 1 + K > K$ (woraus die erste Behauptung folgt) und also auch $0 \leq E(-K) = E(K)^{-1} < 1/K$, woraus die zweite Behauptung folgt.]

(6) Für alle $z \in \mathbb{C}$ gilt $E(\overline{z}) = \overline{E(z)}$.

[Dies folgt aus der Definition, sowie der Tatsache, dass $w_n \to w$ für eine komplexe Folge (w_n) impliziert: $\overline{w_n} \to \overline{w}$.]

(7) Für alle $x, y \in \mathbb{R}$ gilt $E(x + iy) = E(x)E(iy)$ und $|E(iy)| = 1$.

[Die erste Gleichung folgt aus (1). Mittels (6), (1) und (0) gilt für $y \in \mathbb{R}$:

$$|E(iy)| = \sqrt{E(iy)\overline{E(iy)}} = \sqrt{E(iy)E(-iy)} = \sqrt{E(iy - iy)} = 1.]$$

(8) Für alle $x \in \mathbb{R}$ gilt $\sqrt[n]{E(x)} = E(\frac{x}{n})$.

[Nach (3) ist $E(x) > 0$, $E(\frac{x}{n}) > 0$, und es gilt nach (2): $E(\frac{x}{n})^n = E(n\frac{x}{n}) = E(x)$.]

(9) Für alle $z \in \mathbb{C}$ gilt $|E(z)| = E(\operatorname{Re} z)$.

[Nach (6), (1) und (8) ist:

$$|E(z)| = \sqrt{E(z)\overline{E(z)}} = \sqrt{E(z + \overline{z})} = \sqrt{E(2\operatorname{Re} z)} = E(\operatorname{Re} z).]$$

Bemerkung und Definition: Wegen (2), (8) und 7.8 gilt für alle $m \in \mathbb{Z}$ und alle $n \in \mathbb{N}$:

$$e^m = E(m), \quad \sqrt[n]{e} = E\Big(\frac{1}{n}\Big), \quad \sqrt[n]{e^m} = E\Big(\frac{m}{n}\Big) = \big(\sqrt[n]{e}\big)^m.$$

Man schreibt deshalb auch $e^z := E(z)$ für jedes $z \in \mathbb{C}$. Eine andere Bezeichnung ist $\exp(z) := E(z)$.

Wir zeigen noch die folgenden **Abschätzungen**:

(10) Für alle $h \in \mathbb{C}$ gilt $|E(h) - 1| \leq |h|E(|h|)$.

(11) Für alle $h \in \mathbb{C} \setminus \{0\}$ gilt $|\frac{E(h)-1}{h} - 1| \leq |h|E(|h|)$.

Beweis. Es ist $E(h) = 1 + h + \frac{h^2}{2!} + \frac{h^3}{3!} + \ldots$. Also

$$|E(h) - 1| = |\sum_{n=1}^{\infty} \frac{h^n}{n!}| \leq |h| \sum_{n=1}^{\infty} \frac{|h|^{n-1}}{n!} \leq |h| \sum_{n=1}^{\infty} \frac{|h|^{n-1}}{(n-1)!} = |h| E(|h|),$$

womit (10) gezeigt ist. Für $h \neq 0$ haben wir

$$|\frac{E(h) - 1}{h} - 1| \leq \sum_{n=1}^{\infty} \frac{|h|^n}{(n+1)!} \leq |h| \sum_{n=1}^{\infty} \frac{|h|^{n-1}}{n!} \leq |h| E(|h|).$$

□

7.12. Sinus und Cosinus: Wir definieren die Funktionen $\sin : \mathbb{C} \to \mathbb{C}$ und $\cos : \mathbb{C} \to \mathbb{C}$ durch

$$\sin z := \frac{1}{2i}(e^{iz} - e^{-iz}) \qquad \text{und} \qquad \cos z := \frac{1}{2}(e^{iz} + e^{-iz}).$$

(0) Es ist $\sin 0 = 0$ und $\cos 0 = 1$. [folgt aus 7.11(0)]

(1) **Reihendarstellungen:** Für jedes $z \in \mathbb{C}$ gilt:

$$\begin{aligned} \sin z &= \sum_{n=0}^{\infty} \frac{(-1)^n z^{2n+1}}{(2n+1)!} = z - \frac{z^3}{3!} + \frac{z^5}{5!} - \ldots, \\ \cos z &= \sum_{n=0}^{\infty} \frac{(-1)^n z^{2n}}{(2n)!} = 1 - \frac{z^2}{2} + \frac{z^4}{4!} - \frac{z^6}{6!} + \ldots. \end{aligned}$$

Das folgt aus

$$\begin{aligned} E(iz) &= 1 + iz - \frac{z^2}{2} - i\frac{z^3}{3!} + \frac{z^4}{4!} + i\frac{z^5}{5!} - \frac{z^6}{6!} - i\frac{z^7}{7!} + \frac{z^8}{8!} + \ldots \\ E(-iz) &= 1 - iz - \frac{z^2}{2} + i\frac{z^3}{3!} + \frac{z^4}{4!} - i\frac{z^5}{5!} - \frac{z^6}{6!} + i\frac{z^7}{7!} + \frac{z^8}{8!} + \ldots. \end{aligned}$$

(2) **Eulerformel:** Für alle $x \in \mathbb{R}$ gilt $\sin x \in \mathbb{R}$ und $\cos x \in \mathbb{R}$, sowie

$$e^{ix} = \cos x + i \sin x \text{ und } (\sin x)^2 + (\cos x)^2 = 1, \text{also } |\sin x| \leq 1, \ |\cos x| \leq 1.$$

[Aus (1) folgt $\sin x, \cos x \in \mathbb{R}$. Weiter ist

$$\operatorname{Re} e^{ix} = \frac{1}{2}(e^{ix} + \overline{e^{ix}}) = \cos x \quad \text{und} \quad \operatorname{Im} e^{ix} = \frac{1}{2i}(e^{ix} - \overline{e^{ix}}) = \sin x.$$

Somit $(\cos x)^2 + (\sin x)^2 = |e^{ix}|^2 = 1$ nach 7.11(7).]

(3) **Additionstheoreme:** Für alle $z, w \in \mathbb{C}$ gilt:

$$\begin{aligned}\sin(z+w) &= \sin z \cos w + \cos z \sin w,\\ \cos(z+w) &= \cos z \cos w - \sin z \sin w.\end{aligned}$$

Das folgt sofort aus 7.11(1) (zur Übung).

Wir zeigen noch die folgenden **Abschätzungen**:

(4) Für alle $h \in \mathbb{C}$ gilt: $|\sin h| \le |h|E(|h|)$.

(5) Für alle $h \in \mathbb{C}$ gilt: $|\cos h - 1| \le |h|E(|h|)$.

(6) Für alle $h \in \mathbb{C} \setminus \{0\}$ gilt: $|\frac{\sin h}{h} - 1| \le |h|E(|h|)$.

(7) Für alle $h \in \mathbb{C} \setminus \{0\}$ gilt: $|\frac{\cos h - 1}{h}| \le |h|E(|h|)$.

Beweis. Die Abschätzungen folgen aus (10) und (11) in 7.11, wenn man beachtet:

$$\begin{aligned}\sin h &= \frac{1}{2i}(E(ih) - 1 - (E(-ih) - 1)),\\ \cos h - 1 &= \frac{1}{2}(E(ih) - 1 + (E(-ih) - 1)),\\ \frac{\sin h}{h} - 1 &= \frac{E(ih) - E(-ih)}{2ih} - 1 = \frac{1}{2}\left(\frac{E(ih)-1}{ih} - 1 + \frac{E(-ih)-1}{-ih} - 1\right),\\ \frac{\cos h - 1}{h} &= \frac{i}{2}\frac{E(ih) - 1 + E(-ih) - 1}{ih} = \frac{i}{2}\left(\frac{E(ih)-1}{ih} - 1 - \left(\frac{E(-ih)-1}{-ih} - 1\right)\right).\end{aligned}$$

□

7.13. Potenzreihen: Sei $z_0 \in \mathbb{C}$. Eine *Potenzreihe (PR) um* z_0 hat die Form $\sum_{n=0}^{\infty} a_n(z - z_0)^n$, wobei $(a_n)_{n\in\mathbb{N}_0}$ eine Folge in $\mathbb{C}$ ist. Die a_n heißen *Koeffizienten* der Potenzreihe und z_0 heißt *Entwicklungspunkt*.

Wir nennen eine Potenzreihe *reell*, falls $z_0 \in \mathbb{R}$ und alle a_n reell sind. Betrachten wir reelle Potenzreihen mit Entwicklungspunkt $x_0 \in \mathbb{R}$ und wollen reelle Zahlen einsetzen, so schreiben wir häufig $\sum_{n=0}^{\infty} a_n(x - x_0)^n$.

Fragen: Für welche $z \in \mathbb{C}$ konvergiert eine gegebene Potenzreihe? Für welche $x \in \mathbb{R}$ konvergiert eine reelle Potenzreihe?

Beispiele: Die Potenzreihen für exp, sin und cos konvergieren für jedes $z \in \mathbb{C}$.

Die geometrische Reihe $\sum_{n=0}^{\infty} z^n$ ist für $|z| < 1$ absolut konvergent und für $|z| \ge 1$ divergent.

Alle diese Reihen haben als Entwicklungspunkt $z_0 = 0$.

Bemerkung: Jede Potenzreihe $\sum_{n=0}^{\infty} a_n(z - z_0)^n$ konvergiert für $z = z_0$, dh im Entwicklungspunkt. Setzt man $z = z_0$ ein, erhält man nämlich $a_0 \cdot 1 + a_1 \cdot 0 + a_2 \cdot 0 + \ldots = a_0$.

7.14. Der Konvergenzradius: Sei $\sum_{n=0}^{\infty} a_n(z-z_0)^n$ eine Potenzreihe. Setzt man

$$R := \frac{1}{\limsup_{n\to\infty} \sqrt[n]{|a_n|}} \qquad \text{(Formel von Cauchy-Hadamard)}$$

mit den Konventionen $\frac{1}{0} = \infty$ und $\frac{1}{\infty} = 0$[1], so gilt:

(a) Die Potenzreihe konvergiert absolut für alle $z \in \mathbb{C}$ mit $|z - z_0| < R$.

(b) Die Potenzreihe divergiert für alle $z \in \mathbb{C}$ mit $|z - z_0| > R$.

Die Zahl R heißt *Konvergenzradius* (KR) der Potenzreihe.

Bemerkung: Im Falle $R = 0$ konvergiert die Potenzreihe also nur für $z = z_0$ und divergiert für alle $z \in \mathbb{C} \setminus \{z_0\}$.

Im Fall $R = \infty$ ist die Potenzreihe für jedes $z \in \mathbb{C}$ absolut konvergent.

Bemerkung: Im Fall $R \in (0, \infty)$ lässt sich für $z \in \mathbb{C}$ mit $|z - z_0| = R$ keine allgemeine Aussage treffen:

(a) Die geometrische Reihe $\sum_{n=0}^{\infty} z^n$ hat Konvergenzradius $R = 1$, sie divergiert für jedes $z \in \mathbb{C}$ mit $|z| = 1$.

(b) Die Potenzreihe $\sum_{n=1}^{\infty} \frac{z^n}{n^2}$ hat Konvergenzradius $R = 1$, für jedes $z \in \mathbb{C}$ mit $|z| = 1$ ist sie absolut konvergent.

Folgerung: (a) Konvergiert die Potenzreihe für ein $z_1 \in \mathbb{C}$, so konvergiert sie absolut für alle $z \in \mathbb{C}$ mit $|z - z_0| < |z_1 - z_0|$ und für den Konvergenzradius R gilt $R \geq |z_1 - z_0|$.

(b) Divergiert die Potenzreihe für ein $z_1 \in \mathbb{C}$, so divergiert sie für alle $z \in \mathbb{C}$ mit $|z - z_0| > |z_1 - z_0|$ und für den Konvergenzradius R gilt $R \leq |z_1 - z_0|$.

Beispiele: (1) Die Potenzreihen für exp, sin, cos haben jeweils Konvergenzradius ∞.

(2) Die geometrische Reihe $\sum_{n=0}^{\infty} z^n$ hat Konvergenzradius $R = 1$, ebenso die Potenzreihen $\sum_{n=0}^{\infty} n^p z^n$ mit $p \in \mathbb{N}$. Die Potenzreihe $\sum_{n=0}^{\infty} n^n z^n$ hat Konvergenzradius $R = 0$.

Beweis des Satzes. Wir haben für $z \in \mathbb{C} \setminus \{z_0\}$:

$$\limsup_{n\to\infty} \sqrt[n]{|a_n(z-z_0)^n|} = \left(\limsup_{n\to\infty} \sqrt[n]{|a_n|}\right) |z - z_0|,$$

und die Behauptungen des Satzes folgen aus dem Wurzelkriterium 7.6. □

[1] Beachte, dass $\alpha := \limsup \sqrt[n]{|a_n|} \in [0, \infty]$ und dass also diese Konventionen für $1/\alpha \in [0, \infty]$ sinnvoll sind, vgl. 6.7.

Bemerkung: Eine reelle Potenzreihe $\sum_{n=0}^{\infty} a_n(x-x_0)^n$ mit Konvergenzradius $R \in (0,\infty]$ ist für $x \in (x_0 - R, x_0 + R)$ absolut konvergent und für $x \in \mathbb{R} \setminus [x_0 - R, x_0 + R]$ divergent. Für $x = x_0 \pm R$ ist keine allgemeine Aussage möglich!

7.15. Satz (Konvergenzradius über Quotienten): Sei $\sum_{n=0}^{\infty} a_n(z-z_0)^n$ eine Potenzreihe mit $a_n \neq 0$ für fast alle $n \in \mathbb{N}$. Gilt $|\frac{a_{n+1}}{a_n}| \to \alpha$, wobei $0 \leq \alpha \leq \infty$, so ist der Konvergenzradius $R = 1/\alpha$.

Beweis. Das folgt aus dem Quotientenkriterium (Bemerkung in 7.7). □

Beispiel: Für die Exponentialreihe ist $a_n = 1/n!$, also $|\frac{a_{n+1}}{a_n}| = \frac{1}{n+1} \to 0$ und $R = 1/0 = \infty$. Auf die Potenzreihen für sin und cos aus 7.12 lässt sich der Satz nicht anwenden.

Bemerkung: Hier erlauben $\limsup |\frac{a_{n+1}}{a_n}|$ und $\liminf |\frac{a_{n+1}}{a_n}|$ i.a. keine Entscheidung!

Beispiel: $a_n = (\frac{2+(-1)^n}{3})^n$ für alle $n \in \mathbb{N}_0$. Dann ist

$$a_n = \begin{cases} 3^{-n} & , n \text{ ungerade}, \\ 1 & , n \text{ gerade}. \end{cases} \qquad \text{und} \qquad \left|\frac{a_{n+1}}{a_n}\right| = \begin{cases} 3^n & , n \text{ ungerade}, \\ 3^{-n-1} & , n \text{ gerade}. \end{cases},$$

also $\limsup |\frac{a_{n+1}}{a_n}| = \infty$ und $\liminf |\frac{a_{n+1}}{a_n}| = 0$, aber der Konvergenzradius ist $R = 1$.

7.16. g-adische Entwicklung reeller Zahlen: Sei $g \in \mathbb{N}$ mit $g \geq 2$. Für jedes $x \in [0,1]$ gibt es eine Folge $(a_n)_{n\in\mathbb{N}}$ in $\{0, 1, \ldots, g-1\}$ so, dass gilt:

$$x = \sum_{n=1}^{\infty} \frac{a_n}{g^n}.$$

Man schreibt dies dann als

$$x = 0.a_1a_2a_3\ldots,$$

wobei aus dem Zusammenhang klar sein muss, was g hier sein soll.

Gebräuchlich sind vor allem die *Dezimaldarstellung* ($g = 10$) und (meist im Zusammenhang mit Computern) die *Dualdarstellung* ($g = 2$).

Warnung: Die Folge (a_n) ist nicht für jedes $x \in [0,1]$ eindeutig bestimmt, z.B. gilt

$$\sum_{n=2}^{\infty} \frac{2}{3^n} = \frac{1}{3} = \frac{1}{3} + \sum_{n=2}^{\infty} \frac{0}{3^n}$$

(links steht $0.02222\ldots$ und rechts steht $0.10000\ldots$, wobei hier $g = 3$). Nichteindeutigkeit tritt aber nur in vergleichbaren Fällen auf.

Bemerkung: Für $x > 1$ findet man ein minimales $k \in \mathbb{N}$ mit $g^k \geq x$ und kann x/g^k wie oben darstellen. Danach verschiebe man den Punkt um k Stellen nach rechts. Für $x < 0$ schreibt man vor die Darstellung von $|x|$ ein Minuszeichen.

7.17. Ergänzung: $\mathbb{R}$ ist überabzählbar: Eine Menge M heißt *unendlich*, wenn es eine injektive Abbildung $\phi : \mathbb{N} \to M$ gibt (M hat dann mindestens so viele Elemente wie $\mathbb{N}$). Eine unendliche Menge heißt *abzählbar*, wenn es eine surjektive Abbildung $\phi : \mathbb{N} \to M$ gibt (tatsächlich gibt es dann auch eine bijektive Abbildung $\mathbb{N} \to M$, so dass M dann "genau so viele" Elemente wie $\mathbb{N}$ hat).

Beispiele: $\mathbb{N}$ ist abzählbar. $\mathbb{Z}$ ist abzählbar: $(1, -1, 2, -2, \ldots)$. $\mathbb{Q} = \{\frac{p}{q} : p \in \mathbb{Z}, q \in \mathbb{N}\}$ ist abzählbar: Das liegt daran, dass $\mathbb{Z} \times \mathbb{N}$ abzählbar ist.

Eine unendliche Menge, die nicht abzählbar ist, heißt *überabzählbar*.

$\mathbb{R}$ ist überabzählbar: Die folgende Teilmenge C von $\mathbb{R}$ ist nämlich überabzählbar:

$$C := \{\sum_{n=1}^{\infty} \frac{a_n}{3^n} : a_n \in \{0, 2\} \text{ für alle } n \in \mathbb{N}\}.$$

Wir setzen

$$M := \{(a_n)_{n\in\mathbb{N}} : a_n \in \{0, 2\} \text{ für alle } n \in \mathbb{N}\}$$

und stellen zunächst fest, dass die Abbildung $M \to C$, $(a_n) \mapsto \sum_{n=1}^{\infty} \frac{a_n}{3^n}$, bijektiv ist (vergleiche 7.16). Mit einem Diagonalschluss sieht man ein, dass M überabzählbar ist: Sei $\phi : \mathbb{N} \to M$ eine Abbildung. Beachte, dass jedes $\phi(j)$ eine Folge $(\phi(j)_n)_{n\in\mathbb{N}}$ in $\{0, 2\}$ ist. Wir schreiben die Folgen $\phi(1), \phi(2), \ldots$ als Zeilen untereinander, definieren für jedes $n \in \mathbb{N}$

$$a_n := \begin{cases} 0 & , \phi(n)_n = 2 \\ 2 & , \phi(n)_n = 0, \end{cases}$$

und erhalten eine Folge $a = (a_n)_{n\in\mathbb{N}} \in M$ mit $a \neq \phi(j)$ für jedes $j \in \mathbb{N}$ (da $a_j \neq \phi(j)_j$). Somit ist ϕ nicht surjektiv. Da $\phi : \mathbb{N} \to M$ beliebig war, ist M überabzählbar.

Also sind auch C und damit $\mathbb{R}$ und $\mathbb{C}$ überabzählbar.

Bemerkung: C ist die *Cantormenge*, die man auch wie folgt erhält: Setze $C_0 := [0, 1]$. C_1 entstehe aus C_0 durch Entfernung des offenen mittleren Drittels $(\frac{1}{3}, \frac{2}{3})$. C_2 entstehe aus C_1 durch Entfernen des mitteleren Drittels aus jedem verbliebenen Teilintervall, und ebenso entstehe C_{n+1} aus C_n für jedes $n \geq 2$. Man erhält eine Folge $C_0 \supset C_1 \supset C_2 \supset \ldots$ und setzt $C := \bigcap_{n\in\mathbb{N}_0} C_n$.

Die einem $x \in C$ zugeordnete Folge (a_n) in $\{0, 2\}$ beschreibt für jedes $n \in \mathbb{N}$, ob x im n-ten Schritt im linken Drittel (für $a_n = 0$) oder im rechten Drittel (für $a_n = 2$) des bisherigen Intervalls zu finden ist, was man sich auch als *dyadischen Baum* vorstellen kann.

7.18. Ergänzung: Gleichmäßige Konvergenz von Potenzreihen: Sei $\sum_{n=0}^{\infty} a_n(x - x_0)^n$ eine reelle Potenzreihe mit Konvergenzradius $R \in (0, \infty]$ und

$$f : (x_0 - R, x_0 + R) \to \mathbb{R}, x \mapsto f(x) := \sum_{n=0}^{\infty} a_n(x - x_0)^n,$$

sowie für jedes $N \in \mathbb{N}_0$:

$$f : (x_0 - R, x_0 + R) \to \mathbb{R}, x \mapsto f(x) := \sum_{n=0}^{N} a_n(x - x_0)^n,$$

Frage: Wie gut ist die Approximation, dh *auf welche Weise* konvergiert f_N für $N \to \infty$ gegen f?

1.Antwort: Es gilt für jedes $x \in (x_0 - R, x_0 + R)$: $f_N(x) \to f(x)$ $(N \to \infty)$, also ist $(|f_N(x) - f(x)|)_{N\in\mathbb{N}}$ Nullfolge für jedes $x \in (x_0 - R, x_0 + R)$. Man sagt, "die Funktionenfolge $(f_N)_{N\in\mathbb{N}}$ *konvergiert punktweise* auf $(x_0 - R, x_0 + R)$ gegen die Funktion f".

2. Antwort: Ist $r \in (0, R)$ und $x \in [x_0 - r, x_0 + r]$, so gilt

$$|f_N(x) - f(x)| \leq \sum_{n=N+1}^{\infty} |a_n||x - x_0|^n \leq \sum_{n=N+1}^{\infty} |a_n|r^n =: \alpha_N$$

und $(\alpha_N)_{N\in\mathbb{N}}$ ist eine Nullfolge. Wir haben also gezeigt, dass es eine Nullfolge (α_N) gibt mit

$$\forall x \in [x_0 - r, x_0 + r] : |f_N(x) - f(x)| \leq \alpha_N.$$

für alle $N \in \mathbb{N}$. Man sagt hierzu: "die Funktionenfolge $(f_N)_{N\in\mathbb{N}}$ *konvergiert auf* $[x_0-r, x_0+r]$ *gleichmäßig* gegen die Funktion f".

Wenn wir $\varepsilon > 0$ als gewünschte Fehlerschranke gegeben haben und N so groß machen, dass $\alpha_N < \varepsilon$ ist, so ist der Fehler $|f_N(x) - f(x)| < \varepsilon$ für **jedes** $x \in [x_0 - R, x_0 + R]$.

Fazit:

(a) Es gilt $f_N \to f$ punktweise auf $(x_0 - R, x_0 + R)$.

(b) Für jedes $r \in (0, R)$ gilt $f_N \to f$ gleichmäßig auf $[x_0 - r, x_0 + r]$.

Beispiele: (1) Für jedes $r > 0$ konvergiert die durch $f_N(x) := \sum_{n=0}^{N} \frac{x^n}{n!}$ definierte Funktionenfolge (f_N) auf $[-r, r]$ gleichmäßig gegen exp. Entsprechendes gilt für die sin- und cos-Reihen aus 7.12.

(2) Für jedes $r \in (0, 1)$ konvergiert die durch $f_N(x) := \sum_{n=0}^{N} x^n$ definierte Funktionenfolge (f_N) auf $[-r, r]$ gleichmäßig gegen die durch $f(x) = \sum_{n=0}^{\infty} x^n = \frac{1}{1-x}$ definierte Funktion f. Auf $[0, 1)$ konvergiert (f_N) *punktweise* gegen f, aber **nicht gleichmäßig!**

8 Vektorräume

In diesem Abschnitt wird der Begriff des Vektorraumes eingeführt und einige Beispiele gegeben. Wir werden uns später im Zusammenhang mit linearen Gleichungssystemen noch genauer mit Vektorräumen beschäftigen. Wir beginnen mit dem einfachsten Beispiel.

8.1. Verknüpfungen in $\mathbb{R}^n$ und $\mathbb{C}^n$: Sei $n \in \mathbb{N}$. $\mathbb{R}^n$ ist die Menge aller n-Tupel $(x_1, x_2, \ldots, x_n)$ reeller Zahlen (dh $x_j \in \mathbb{R}$ für $j = 1, 2, \ldots, n$) und $\mathbb{C}^n$ ist die Menge aller n-Tupel $(z_1, z_2, \ldots, z_n)$ komplexer Zahlen (dh $z_j \in \mathbb{C}$ für $j = 1, 2, \ldots, n$).

Wir schreiben im folgenden oft $\mathbb{K}$ für den *Skalarkörper*, wobei $\mathbb{K}$ für $\mathbb{R}$ oder $\mathbb{C}$ steht. Also

$$\mathbb{K}^n = \{(x_1, x_2, \ldots, x_n) : x_j \in \mathbb{K} \text{ für } j = 1, 2, \ldots, n\}.$$

Wir erklären Verknüpfungen

$$+ : \mathbb{K}^n \times \mathbb{K}^n \to \mathbb{K}^n, \qquad \cdot : \mathbb{K} \times \mathbb{K}^n \to \mathbb{K}^n$$

durch

$$\begin{aligned}(x_1, x_2, \ldots, x_n) + (y_1, y_2, \ldots, y_n) &= (x_1 + y_1, x_2 + y_2, \ldots, x_n + y_n)\\ \alpha \cdot (x_1, x_2, \ldots, x_n) &= (\alpha x_1, \alpha x_2, \ldots, \alpha x_n),\end{aligned}$$

dh also durch *komponentenweise* Addition bzw. komponentenweise Multiplikation mit α. Der Punkt $\cdot$ für die Multiplikation mit Skalaren wird gewöhnlich weggelassen.

Für die **Anschauung** besonders wichtig sind $\mathbb{R}^2$ und $\mathbb{R}^3$. Man beachte, dass für $n = 1$ auch $\mathbb{K}^1 = \mathbb{K}$ ein $\mathbb{K}$-Vektorraum ist.

Beispiele im $\mathbb{R}^2$: $(1, 1) + (0, 1) = (1, 2)$, $2 \cdot (1, -1) = (2, -2)$.

8.2. Vektorraumaxiome: Setzen wir $V = \mathbb{K}^n$, wobei $n \in \mathbb{N}$, so haben die in 8.1 definierten Verknüpfungen $+ : V \times V \to V$ und $\cdot : \mathbb{K} \times V \to V$ die folgenden Eigenschaften:

$(V1)$ $(x + y) + z = x + (y + z)$ für alle $x, y, z \in V$ (Assoziativität),
$(V2)$ $x + y = y + x$ für alle $x, y \in V$ (Kommutativität),
$(V3)$ es gibt eine $0 \in V$ mit $x + 0 = x$ für alle $x \in V$ (Existenz der Null),
$(V4)$ für jedes $x \in V$ gibt es ein $-x \in V$ mit $x + (-x) = 0$ (Existenz des Negativen),
$(V5)$ $(\alpha\beta)x = \alpha(\beta x)$ für alle $\alpha, \beta \in \mathbb{K}$, $x \in V$ (Assoziativität der Multiplikationen),
$(V6)$ $\alpha(x+y) = \alpha x+\alpha y$ und $(\alpha+\beta)x = \alpha x+\beta x$ für alle $\alpha, \beta \in \mathbb{K}$, $x, y \in V$ (Distributivität),
$(V7)$ $1x = x$ für alle $x \in V$ (Kompatibilität).

Definition: Ist $V \neq \emptyset$ eine Menge mit Verknüpfungen $+ : V \times V \to V$ und $\cdot : \mathbb{K} \times V \to V$, für welche die Eigenschaften $(V1)$–$(V7)$ gelten, so heißt V ein *Vektorraum über* $\mathbb{K}$ oder ein $\mathbb{K}$-*Vektorraum* ($\mathbb{K}$-*VR*). Die Elemente eines Vektorraumes heißen *Vektoren.*

Beispiele: $\mathbb{R}^n$ ist ein $\mathbb{R}$-Vektorraum, $\mathbb{C}^n$ ist ein $\mathbb{C}$-Vektorraum, aber auch: $\mathbb{C}^n$ ist ein $\mathbb{R}$-Vektorraum, insbesondere ist $\mathbb{C}$ ein $\mathbb{R}$-Vektorraum.

Allgemein ist jeder $\mathbb{C}$-Vektorraum ist auch ein $\mathbb{R}$-Vektorraum.

Die Menge der komplexen Polynome $\mathbb{C}[z]$ ist ein $\mathbb{C}$-Vektorraum.

Bemerkung: Ist V ein $\mathbb{K}$-Vektorraum, so ist das Element $0 \in V$ in $(V3)$ und für jedes $x \in V$ das Negative $-x$ in $(V4)$ eindeutig bestimmt. Außerdem gilt für alle $x \in V$: $0 \cdot x = 0$ und $-x = (-1) \cdot x$.

Beweis: Die Aussagen zu $(V3)$ und $(V4)$ zeigt man wie in 4.1. Sei nun $x \in V$. Es ist $0 \cdot x = (0+0) \cdot x = (0 \cdot x) + (0 \cdot x)$, also

$$0 \overset{(V4)}{=} 0 \cdot x + (-0 \cdot x) = (0 \cdot x + 0 \cdot x) + (-0 \cdot x) \overset{(V1)}{=} 0 \cdot x + (0 \cdot x + (-0 \cdot x)) \overset{(V4)}{=} 0 \cdot x + 0 \overset{(V3)}{=} 0 \cdot x.$$

Weiter ist

$$x + (-1) \cdot x \overset{(V7)}{=} 1 \cdot x + (-1) \cdot x \overset{(V6)}{=} (1 + (-1)) \cdot x = 0 \cdot x \overset{(s.o.)}{=} 0,$$

also $(-1) \cdot x = -x$.

8.3. Wichtiges Beispiel: Sei $n \in \mathbb{N}$. Ein n-Tupel $x = (x_1, x_2, \ldots, x_n) \in \mathbb{K}^n$ können wir auch auffassen als Abbildung (Funktion)

$$x : \{1, 2, \ldots, n\} \to \mathbb{K} \text{ mittels } x(j) := x_j \text{ für } j = 1, 2, \ldots, n.$$

Allgemeiner sei nun $I \neq \emptyset$ eine Menge (häufig ist zB $I \subset \mathbb{R}$ ein Intervall) und

$$\mathbb{K}^I := \text{ Menge aller Abbildungen } x : I \to \mathbb{K}.$$

Setzt man für $x, y \in \mathbb{K}^I$ und $\alpha \in \mathbb{K}$:

$$\begin{aligned} x + y : I \to \mathbb{K}, &\qquad t \mapsto x(t) + y(t), \\ \alpha \cdot x : I \to \mathbb{K}, &\qquad t \mapsto \alpha x(t), \end{aligned}$$

so ist $\mathbb{K}^I$ bzgl. dieser Verknüpfungen $+ : \mathbb{K}^I \times \mathbb{K}^I \to \mathbb{K}^I$ und $\cdot : \mathbb{K} \times \mathbb{K}^I \to \mathbb{K}^I$ ein $\mathbb{K}$-Vektorraum. Für $f, g \in \mathbb{K}^I$ und $\alpha \in \mathbb{K}$ sind $f + g$ und αf also *punktweise* erklärt, dh indem man die Punkte $t \in I$ in die Funktionen einsetzt und dann die Verknüpfung in $\mathbb{K}$ verwendet.

Beispiele: (1) $I = [a, b]$ mit $a < b$ reell: Die Menge $\mathbb{K}^{[a,b]}$ aller Funktionen $f : [a, b] \to \mathbb{K}$ ist ein $\mathbb{K}$-Vektorraum.

(2) $I = \mathbb{N}$: $\mathbb{K}^{\mathbb{N}}$ ist die Menge aller Folgen in $\mathbb{K}$, dh

$$\mathbb{K}^{\mathbb{N}} = \{(x_1, x_2, \ldots) : \ x_j \in \mathbb{K} \text{ für jedes } j \in \mathbb{N}\}$$

ist ein $\mathbb{K}$-Vektorraum. Die Verknüpfungen sind hier

$$(x_1, x_2, \ldots) + (y_1, y_2, \ldots) = (x_1 + y_1, x_2 + y_2, \ldots), \quad \alpha(x_1, x_2, \ldots) = (\alpha x_1, \alpha x_2, \ldots).$$

8.4. Untervektorräume: Sei V ein $\mathbb{K}$-Vektorraum.

Definition: Eine Teilmenge $U \subseteq V$ heißt *Untervektorraum* oder *linearer Teilraum* von V, falls U bzgl. der Verknüpfungen $+$ und $\cdot$ von V ein $\mathbb{K}$-Vektorraum ist.

Satz: Eine Teilmenge $U \subseteq V$ ist Untervektorraum von V genau dann, wenn

(i) $U \neq \emptyset$ ist und
(ii) für alle $x, y \in U$, $\alpha \in \mathbb{K}$ gilt: $x + y \in U$ und $\alpha x \in U$,

dh genau dann, wenn U nichtleer ist und *abgeschlossen* unter den Verknüpfungen $+$ und $\cdot$.

Bemerkung: Die Bedingung (i) kann man ersetzen durch $0 \in U$.

Beispiele: (0) Jeder Vektorraum V hat die trivialen Untervektorräume $\{0\}$ und V.

(1) $\{(x, 0, 0) : x \in \mathbb{K}\} = \mathbb{K} \times \{0\} \times \{0\}$ ist ein Untervektorraum von $\mathbb{K}^3$.

(2) Sind $x, y \in \mathbb{K}^n$, so ist $\{\alpha x + \beta y : \alpha, \beta \in \mathbb{K}\}$ ein Untervektorraum von $\mathbb{K}^n$.

(3) $K := \{(x_1, x_2) \in \mathbb{R}^2 : x_1^2 + x_2^2 \leq 1\}$ ist **kein** Untervektorraum von $\mathbb{R}^2$, nicht abgeschlossen unter $+$.

(4) $C := \{(x_1, x_2) \in \mathbb{R}^2 : x_1 \geq 0, x_2 \geq 0\}$ ist **kein** Untervektorraum von $\mathbb{R}^2$. C ist zwar abgeschlossen unter $+$, aber $-x \notin C$ für $x \in C \setminus \{0\}$.

(5) Die nichttrivialen Untervektorräume von $\mathbb{R}^2$ sind genau die Geraden durch 0, in $\mathbb{R}^3$ kommen Ebenen durch 0 hinzu.

In den folgenden Beispielen wird 8.3 mit dem Satz oben kombiniert.

(6) Der Raum der konvergenten Folgen $c := \{(x_n) \in \mathbb{K}^{\mathbb{N}} : (x_n) \text{ ist konvergent}\}$ ist ein Untervektorraum von $\mathbb{K}^{\mathbb{N}}$. Die Menge der Nullfolgen $c_0 := \{(x_n) \in c : \lim_n x_n = 0\}$ ist ein Untervektorraum von c.

(7) Der Raum der absolut summierbaren Folgen $l^1 := \{(x_n) \in \mathbb{K}^{\mathbb{N}} : \sum_n |x_n| < \infty\}$ ist ein Untervektorraum von $\mathbb{K}^{\mathbb{N}}$ oder auch von c_0.

(7') Der Raum $s := \{(x_n) \in \mathbb{K}^{\mathbb{N}} : \sum_{n=1}^{\infty} x_n$ konvergiert $\}$ ist ein Untervektorraum von $\mathbb{K}^{\mathbb{N}}$ oder auch von c_0, und l^1 ist ein Untervektorraum von s.

(8) Der Raum der beschränkten Folgen $l^\infty := \{(x_n) \in \mathbb{K}^{\mathbb{N}} : \sup\{|x_n| : n \in \mathbb{N}\} < \infty\}$ ist ein Untervektorraum von $\mathbb{K}^{\mathbb{N}}$. Der Raum c ist ein Untervektorraum von l^∞.

8.5. Lineare Abbildungen: Seien V, W $\mathbb{K}$-Vektorräume.

Definition: Eine Abbildung $\phi : V \to W$ heißt *linear*, (bzw. genauer $\mathbb{K}$-*linear*), falls für alle $x, y \in V$, $\alpha \in \mathbb{K}$ gilt:

$$\phi(x+y) = \phi(x) + \phi(y), \quad \phi(\alpha x) = \alpha\phi(x).$$

Für ein solches ϕ heißt

$$\operatorname{Kern}\phi := \{x \in V : \phi(x) = 0\}$$

der *Kern von* ϕ und

$$\operatorname{Bild}\phi := \{\phi(x) : x \in V\}$$

heißt *Bild von* ϕ.

Ist $\phi : V \to W$ linear, so gilt immer $\phi(0) = 0$. Mithilfe von Satz 8.4 kann man leicht einsehen, dass $\operatorname{Kern}\phi$ ein Untervektorraum von V ist und dass $\operatorname{Bild}\phi$ ein Untervektorraum von W ist. Wir werden das später noch genauer betrachten. Hier geben wir nur einige Beispiele.

Beispiele: (1) Ist V ein $\mathbb{K}$-Vektorraum und $\alpha \in \mathbb{K}$, so ist die Abbildung $V \to V$, $x \mapsto \alpha x$ linear. Ist $y \in V$, so ist die Abbildung $V \to V$, $x \mapsto x + y$ (Verschiebung um y) nur für $y = 0$ linear.

(2) Ist $n \in \mathbb{N}$, so ist für jedes $j \in \{1, 2, \ldots, n\}$ die Abbildung auf die j-te Komponente $\phi_j : \mathbb{K}^n \to \mathbb{K}$, $x \mapsto x_j$ linear. Ähnlich ist das in allgemeineren Situationen.

(3) Die Abbildung $\phi : c \to \mathbb{K}$, $(x_n) \mapsto \lim_{n\to\infty} x_n$, die jede konvergente Folge auf ihren Grenzwert abbildet, ist nach 6.3(5) linear. Hier ist $\operatorname{Kern}\phi = c_0$ der Raum der Nullfolgen.

(4) Die Abbildung $\phi : s \to \mathbb{K}$, $(x_n) \mapsto \sum_{n=1}^{\infty} x_n$ ist linear nach 7.2(5).

(5) Die Abbildung $\phi : s \to \mathbb{K}^{\mathbb{N}}$, $(x_n)_{n\in\mathbb{N}} \mapsto (\sum_{k=1}^{n} x_k)_{n\in\mathbb{N}}$ ist linear und es ist $\operatorname{Bild}\phi = c$: Nach Definition von s in Beispiel 8.4(7') ist $\operatorname{Bild}\phi \subseteq c$. Ist umgekehrt $(y_n)_{n\in\mathbb{N}} \in c$, so definiere $(x_n)_{n\in\mathbb{N}}$ durch $x_1 := y_1$ und $x_n := y_n - y_{n-1}$ für $n \geq 2$. Dann ist $(x_n)_{n\in\mathbb{N}} \in s$ und $\phi((x_n)_{n\in\mathbb{N}}) = (y_n)_{n\in\mathbb{N}}$, also $(y_n)_{n\in\mathbb{N}} \in \operatorname{Bild}\phi$.

9 Stetigkeit

9.1. Definition: Sei $\emptyset \neq D \subseteq \mathbb{R}$ und $f : D \to \mathbb{R}$ eine Funktion. Die Funktion f heißt *stetig in* $x_0 \in D$, falls für jede Folge (x_n) in D mit $x_n \to x_0$ gilt $f(x_n) \to f(x_0)$.

Die Funktion f heißt *stetig in/auf* D, falls sie in jedem $x_0 \in D$ stetig ist.

9.2. Beispiele: (1) exp, sin und cos sind stetig in 0.

Beweis. Sei (x_n) eine Folge mit $x_n \to 0$. Wir finden $M \in \mathbb{R}$ mit $|x_n| \leq M$ für alle $n \in \mathbb{N}$. Nach 7.11(10), 7.12(4) und 7.12(5) gilt

$$\left.\begin{array}{r} |\exp(x_n) - \underbrace{\exp(0)}_{=1}| \\ |\sin x_n - \underbrace{\sin 0}_{=0}| \\ |\cos x_n - \underbrace{\cos 0}_{=1}| \end{array}\right\} \leq \quad |x_n| \exp(|x_n|) \leq \exp(M)\, |x_n| \to 0.$$

□

(2) exp, sin und cos sind stetig auf $\mathbb{R}$.

Beweis. Sei $x_0 \in \mathbb{R}$ und (x_n) eine Folge mit $x_n \to x_0$. Dann gilt nach 7.11, 7.12 und (1):

$$\begin{aligned} |\exp(x_n) - \exp(x_0)| &= \exp(x_0) |\exp(\underbrace{x_n - x_0}_{\to 0}) - 1| \to 0, \\ |\sin(x_n) - \sin(x_0)| &= |\sin(x_n - x_0)\cos x_0 + \cos(x_n - x_0)\sin x_0 - \sin x_0| \\ &\leq |\cos x_0| \underbrace{|\sin(x_n - x_0)|}_{\to 0} + |\sin x_0| \underbrace{|\cos(x_n - x_0) - 1|}_{\to 0} \to 0, \\ |\cos(x_n) - \cos(x_0)| &= |\cos(x_n - x_0)\cos x_0 - \sin(x_n - x_0)\sin x_0 - \cos x_0| \\ &\leq |\cos x_0||\cos(x_n - x_0) - 1| + |\sin x_0||\sin(x_n - x_0)| \to 0. \end{aligned}$$

□

(3) Ist $p \in \mathbb{R}[X]$ ein Polynom, dh $p(X) = a_m X^m + a_{m-1}X^{m-1} + \ldots + a_1 X + a_0$, so ist die Funktion $\mathbb{R} \to \mathbb{R}$, $x \mapsto p(x)$ stetig. Dies folgt aus 6.3, genauso wie der folgende Satz.

9.3. Satz: Seien $f, g : D \to \mathbb{R}$ Funktionen, die in $x_0 \in D$ stetig sind. Dann sind auch $f + g, f \cdot g : D \to \mathbb{R}$ in x_0 stetig. Ist $g(x_0) \neq 0$, so ist auch $\frac{f}{g} : D \setminus \{x \in D : g(x) = 0\} \to \mathbb{R}$ stetig in x_0.

Beachte: Ist g stetig in x_0 und $g(x_0) \neq 0$, so gibt es ein $\delta > 0$ mit

$$(x_0 - \delta, x_0 + \delta) \cap \{x \in D : g(x) = 0\} = \emptyset,$$

dh mit $g(x) \neq 0$ für alle $x \in D$ mit $|x - x_0| < \delta$.

Beweis. Andernfalls findet man zu jedem $n \in \mathbb{N}$ ein $x_n \in D$ mit $g(x_n) = 0$ und $x_n \in (x_0 - 1/n, x_0 + 1/n)$, dh mit $|x_n - x_0| < 1/n$. Wir haben also $x_n \to x_0$ und wegen der Stetigkeit von g in x_0 auch $0 = g(x_n) \to g(x_0) \neq 0$ Widerspruch! $\square$

Folgerung aus dem Satz: Ist $\emptyset \neq D \subseteq \mathbb{R}$, so ist die Menge

$$C(D, \mathbb{R}) := C(D) := \{f : D \to \mathbb{R} : \ f \text{ ist stetig auf } D\}$$

der stetigen Funktionen $D \to \mathbb{R}$ ein reeller Vektorraum (Untervektorraum des Raumes aller Funktionen $D \to \mathbb{R}$).

Beispiel: Sind $p \in \mathbb{R}[X]$ und $q \in \mathbb{R}[X] \setminus \{0\}$ reelle Polynome, so ist die *rationale* Funktion

$$\frac{p}{q} : \mathbb{R} \setminus \{x \in \mathbb{R} : q(x) = 0\} \to \mathbb{R}, \quad x \mapsto \frac{p(x)}{q(x)}$$

stetig.

9.4. Grenzwerte bei Funktionen: Sei $\emptyset \neq D \subseteq \mathbb{R}$ und $f : D \to \mathbb{R}$ eine Funktion. Seien $\alpha, \beta \in \mathbb{R} \cup \{\infty, -\infty\}$ derart, dass eine Folge (x_n) in $D \setminus \{\alpha\}$ mit $x_n \to \alpha$ existiert. Wir schreiben

$$\lim_{x \to \alpha} f(x) = \beta,$$

falls *für jede Folge* (x_n) in $D \setminus \{\alpha\}$ mit $x_n \to \alpha$ gilt: $f(x_n) \to \beta$.

Beispiele: Aus den Abschätzungen in 7.11 und 7.12 folgt etwa:

$$\lim_{x \to 0} \exp(x) = 1, \quad \lim_{x \to \infty} \exp(x) = \infty, \quad \lim_{x \to -\infty} \exp(x) = 0,$$
$$\lim_{x \to 0} \frac{\sin x}{x} = 1, \quad \lim_{x \to 0} \frac{\exp(x) - 1}{x} = 1, \quad \lim_{x \to 0} \frac{\cos x - 1}{x} = 0.$$

Bemerkung: (1) β ist in dieser Definition eindeutig bestimmt (das liegt daran, dass es *mindestens eine* Folge (x_n) in $D \setminus \{\alpha\}$ mit $x_n \to \alpha$ gibt).

(2) Der eventuell vorhandene Funktionswert $f(\alpha)$ spielt **keine** Rolle.

Definition: Ist $\alpha = x_0 \in \mathbb{R}$ und gibt es eine Folge in $D \setminus \{x_0\}$ mit $x_n \to x_0$, so heißt x_0 ein *Häufungspunkt von D*. Für jedes $x_0 \in \mathbb{R}$ und jede Menge $D \subseteq \mathbb{R}$ gilt:

$$x_0 \text{ ist Häufungspunkt von } D \iff \forall \delta > 0 : \ U_\delta(x_0) \cap (D \setminus \{x_0\}) \neq \emptyset.$$

Beispiele: 0 ist ein Häufungspunkt von $(0,1)$ oder von $\{1/n : n \in \mathbb{N}\}$, aber 0 ist kein Häufungspunkt von $\mathbb{Z}$.

Bemerkung: Ist $x_0 \in D$ kein Häufungspunkt von D, so gibt es ein $\delta > 0$ mit

$$U_\delta(x_0) \cap (D \setminus \{x_0\}) = \emptyset, \text{ dh mit } U_\delta(x_0) \cap D = \{x_0\}.$$

Ein solches x_0 heißt *isolierter Punkt von D*. So ist etwa 0 ein isolierter Punkt von $\mathbb{Z}$.

Eine Funktion $f : D \to \mathbb{R}$ ist in jedem isolierten Punkt x_0 von D stetig.

Beweis. Ist (x_n) eine Folge in D mit $x_n \to x_0$, so folgt $x_n = x_0$ für fast alle n, also $f(x_n) = f(x_0)$ für fast alle n und somit $f(x_n) \to f(x_0)$. Also ist f in x_0 stetig. $\square$

Satz: Die Funktion $f : D \to \mathbb{R}$ ist stetig genau dann, wenn für jedes $x_0 \in D$, das Häufungspunkt von D ist, gilt: $\lim_{x \to x_0} f(x) = f(x_0)$.

Beispiel: Sei $D = [0,1] \cup \{2\}$ und $f : D \to \mathbb{R}$, $f(x) := \begin{cases} x^2 & , x \in [0,1) \\ 1/2 & , x = 1 \\ 0 & , x = 2 \end{cases}$.

(i) $x_0 \in [0,1)$: Ist (x_n) eine Folge in D mit $x_n \to x_0$, so gilt $x_n \in [0,1)$ für fast alle n, und somit $f(x_n) = x_n^2 \to x_0^2$. Also ist f stetig in x_0.

(ii) $x_0 = 1$: Ist (x_n) eine Folge in $D \setminus \{1\}$ mit $x_n \to 1$, so gilt ebenfalls $x_n \in [0,1)$ für fast alle n, also $f(x_n) = x_n^2 \to 1$. Somit gilt $\lim_{x \to 1} f(x) = 1 \neq \frac{1}{2} = f(1)$ und f ist in $x_0 = 1$ nicht stetig.

(iii) $x_0 = 2$ ist ein isolierter Punkt von D (man nehme etwa $\varepsilon = 1/2$). Also ist f in $x_0 = 2$ stetig.

9.5. ε-δ-Kriterium: Sei $\emptyset \neq D \subseteq \mathbb{R}$, $f : D \to \mathbb{R}$ eine Funktion und $x_0 \in D$. Dann ist f in x_0 stetig genau dann, wenn es zu jedem $\varepsilon > 0$ ein $\delta > 0$ so gibt, dass

$$\text{für alle } x \in D \text{ mit } |x - x_0| < \delta\text{: } |f(x) - f(x_0)| < \varepsilon.$$

Beweis. Es gelte "ε-δ". Sei (x_n) eine Folge in D mit $x_n \to x_0$ und $\varepsilon > 0$. Dann finden wir ein $\delta > 0$ mit $|f(x) - f(x_0)| < \varepsilon$ für alle $x \in D$ mit $|x - x_0| < \delta$. Wegen $x_n \to x_0$ gilt $|x_n - x_0| < \delta$ für fast alle n. Es folgt $|f(x_n) - f(x_0)| < \varepsilon$ für fast alle n. Damit ist $f(x_n) \to f(x_0)$ gezeigt.

Wenn "ε-δ" nicht gilt, finden wir ein $\varepsilon > 0$ und zu jedem $n \in \mathbb{N}$ ein $x_n \in D$ mit

$$|x_n - x_0| < 1/n [= \delta] \quad \text{und} \quad |f(x_n) - f(x_0)| \geq \varepsilon.$$

Es gilt dann $x_n \to x_0$, aber $f(x_n) \not\to f(x_0)$. Somit ist f in x_0 nicht stetig. $\square$

9.6. Satz zur Komposition stetiger Funktionen: Seien $D_1, D_2 \subseteq \mathbb{R}$, $f : D_1 \to \mathbb{R}$ und $g : D_2 \to \mathbb{R}$ Funktionen. Es sei $f(D_1) \subseteq D_2$, $x_0 \in D_1$ und $y_0 := f(x_0)$. Ist f stetig in x_0 und g stetig in y_0, so ist $g \circ f : D_1 \to \mathbb{R}$ stetig in x_0.

Beweis. Sei (x_n) eine Folge in D_1 mit $x_n \to x_0$. Da f in x_0 stetig ist, folgt $y_n := f(x_n) \to f(x_0) = y_0$. Da g stetig in y_0 ist, folgt

$$(g \circ f)(x_n) = g(f(x_n)) = g(y_n) \to g(y_0) = g(f(x_0)) = (g \circ f)(x_0).$$

Also ist $g \circ f$ in x_0 stetig. □

Beispiel: Sei $p \in \mathbb{N}$. Die Funktion $g_p : [0, \infty) \to \mathbb{R}$, $x \mapsto \sqrt[p]{x}$ ist stetig (nach 6.5(1)). Da $f : x \mapsto x^2$ stetig ist, ist auch $x \mapsto g_2 \circ f(x) = \sqrt{x^2} = |x|$ auf $\mathbb{R}$ stetig.

Die Funktion $\exp \circ (f + \sin) : \mathbb{R} \to \mathbb{R}$, $x \mapsto \exp(x^2 + \sin x)$, ist auf $\mathbb{R}$ stetig.

9.7. Stetigkeit von Potenzreihen: Sei $\sum_{n=0}^{\infty} a_n (x - x_0)^n$ eine reelle Potenzreihe mit Konvergenzradius R, wobei $0 < R \leq \infty$. Dann ist die Funktion

$$f : (x_0 - R, x_0 + R) \to \mathbb{R}, \quad x \mapsto f(x) := \sum_{n=0}^{\infty} a_n (x - x_0)^n$$

auf $D := (x_0 - R, x_0 + R)$ stetig.

Der Satz folgt z.B. aus Ergänzung 7.18 und der Tatsache, dass für eine gleichmäßig konvergente Folge stetiger Funktionen die Grenzfunktion wieder stetig ist. Man kann aber auch direkt argumentieren.

Beweis. (nicht in der Vorlesung) Es reicht, $x_0 = 0$ zu betrachten. Sei $x \in (-R, R)$ und $|x| < r < R$. Dann gilt $|x + h| \leq |x| + |h| \leq r$ für $h \in \mathbb{R}$ mit $|h| \leq r - |x|$, und wir erhalten unter Verwendung von 4.11(1) für $a = x + h$ und $b = x$:

$$\begin{aligned} f(x+h) - f(x) &= \sum_{n=0}^{\infty} a_n (x+h)^n - \sum_{n=0}^{\infty} a_n x^n = \sum_{n=1}^{\infty} a_n [(x+h)^n - x^n] \\ &= \sum_{n=1}^{\infty} \Big(a_n h \sum_{k=0}^{n-1} (x+h)^{n-1-k} x^k \Big) = h \sum_{n=1}^{\infty} \Big(a_n \sum_{k=0}^{n-1} (x+h)^{n-1-k} x^k \Big). \end{aligned}$$

Mithilfe der Dreiecksungleichung ist dann

$$\begin{aligned} |f(x+h) - f(x)| &\leq |h| \sum_{n=1}^{\infty} \Big(|a_n| \sum_{k=0}^{n-1} (|x| + |h|)^{n-1-k} |x|^k \Big) \\ &\leq |h| \sum_{n=1}^{\infty} n |a_n| r^{n-1}. \end{aligned}$$

Die Potenzreihe $\sum_{n=1}^{\infty} n|a_n|z^n$ hat ebenfalls Konvergenzradius R, definiert also eine Funktion $g : (-R, R) \to \mathbb{R}$. Somit

$$|f(x+h) - f(x)| \le |h|g(r) \longrightarrow 0 \quad (h \to 0),$$

und f ist in x stetig. □

Der Satz hat die folgende Ergänzung (hier ohne Beweis).

Abelscher Grenzwertsatz: Ist in der Situation des Satzes $R < \infty$ und konvergiert die Potenzreihe für $x = x_0 + R$, so gilt $\lim_{x \to x_0 + R} f(x) = \sum_{n=0}^{\infty} a_n R^n$. Man kann also f stetig fortsetzen auf $(x_0 - R, x_0 + R]$. Entsprechendes gilt fr Konvergenz in $x = x_0 - R$.

9.8. Zwischenwertsatz (ZWS): Seien $a, b \in \mathbb{R}$ mit $a < b$ und $f : [a, b] \to \mathbb{R}$ stetig. Es sei y_0 *zwischen* $f(a)$ *und* $f(b)$ (dh $y_0 \in [f(a), f(b)]$, falls $f(a) \le f(b)$, und $y_0 \in [f(b), f(a)]$, falls $f(a) > f(b)$). Dann gibt es ein $x_0 \in [a, b]$ mit $f(x_0) = y_0$.

Beweis. Ohne Einschränkung sei $f(a) < f(b)$ (andernfalls betrachte $-f$ und $-y_0$). Wir konstruieren induktiv eine Intervallschachtelung (I_n) mit $I_n = [a_n, b_n]$ so, dass $b_n - a_n = (b-a)/2^n$ und $y_0 \in [f(a_n), f(b_n)]$. Setze dazu $I_0 := [a, b]$ (IA) und im Induktionsschritt $c_n = \frac{a_n + b_n}{2}$ und

$$I_{n+1} := \begin{cases} [a_n, c_n] & , \text{ falls } f(c_n) \ge y_0 \\ [c_n, b_n] & , \text{ falls } f(c_n) < y_0 \end{cases} .$$

Setze $x_0 = \lim_n a_n = \lim_n b_n$. Da f stetig ist, folgt $f(x_0) = \lim_n f(a_n) = \lim_n f(b_n)$ und wegen $y_0 \in [f(a_n), f(b_n)]$ für jedes $n \in \mathbb{N}$ dann auch $y_0 = \lim_n f(a_n) = \lim_n f(b_n)$. □

Folgerung: Ist I ein Intervall und $f : I \to \mathbb{R}$ stetig, so ist $f(I)$ ein Intervall.

Bemerkung: Eine Teilmenge $J \subseteq \mathbb{R}$ ist genau dann ein Intervall, wenn für alle $c, d \in J$ mit $c < d$ gilt: $[c, d] \subseteq J$.

Beweis der Folgerung. Seien $c, d \in f(I)$ mit $c < d$. Sei $y_0 \in [c, d]$. Wir finden $a, b \in I$ mit $f(a) = c$ und $f(b) = d$. Nach dem Zwischenwertsatz gibt es ein x_0 zwischen a und b mit $f(x_0) = y_0$. Somit ist $y_0 \in f(I)$, und $[c, d] \subseteq f(I)$ ist gezeigt. □

Beispiele: (1) Es gilt $\exp(\mathbb{R}) = (0, \infty)$. Nach 7.11(3) ist $e^x > 0$ für alle $x \in \mathbb{R}$, also gilt "$\subset$". Nach der Folgerung ist $\exp(\mathbb{R})$ ein Intervall J und nach 7.11(5) haben wir $\sup J = \infty$, $\inf J = 0$. Folglich ist $(0, \infty) \subset J$.

(2) Ist $p \in \mathbb{R}[X]$ ein normiertes Polynom von ungeradem Grad m, so gilt $p(\mathbb{R}) = \mathbb{R}$.

Beweis. Wegen des Zwischenwertsatzes reicht es zu zeigen, dass $p(x) \to \infty$ für $x \to \infty$ und dass $p(x) \to -\infty$ für $x \to -\infty$. Die Behauptungen sind klar für $m = 1$. Sei also $m \ge 3$

und $p(x) = x^m + a_{m-1}x^{m-1} + \ldots + a_1 x + a_0$. Für $x \geq 1 + |a_{m-1}| + |a_{m-2}| + \ldots + |a_1| + |a_0|$ gilt dann

$$\begin{aligned} p(x) &\geq x^m - |a_{m-1}|x^{m-1} - |a_{m-2}|x^{m-2} - \ldots - |a_1|x - |a_0| \\ &\geq x^m - |a_{m-1}|x^{m-1} - |a_{m-2}|x^{m-1} - \ldots - |a_1|x^{m-1} - |a_0|x^{m-1} \\ &\geq x^{m-1} \underbrace{(x - (|a_{m-1}| + |a_{m-2}| + \ldots + |a_1| + |a_0|))}_{\geq 1} \\ &\geq x^{m-1} \to \infty \quad (x \to \infty). \end{aligned}$$

Die Aussage für $x \to -\infty$ zeigt man ähnlich. $\square$

9.9. Einseitige Grenzwerte: Sei $D \subseteq \mathbb{R}$ und $f : D \to \mathbb{R}$ eine Funktion. Sei $x_0 \in \mathbb{R}$ ein Häufungspunkt von $D \cap (x_0, \infty)$ [bzw. von $D \cap (-\infty, x_0)$]. Wir schreiben

$$\lim_{x \to x_0+} f(x) = \beta \in \mathbb{R} \cup \{\infty, -\infty\} \quad [\text{bzw. } \lim_{x \to x_0-} f(x) = \beta \in \mathbb{R} \cup \{\infty, -\infty\}],$$

falls für jede Folge (x_n) in $D \cap (x_0, \infty)$ [bzw. in $D \cap (-\infty, x_0)$] mit $x_n \to x_0$ gilt: $f(x_n) \to \beta$.

Im Falle der Existenz heißt $\lim_{x \to x_0+} f(x)$ *rechtsseitiger Limes* und $\lim_{x \to x_0-} f(x)$ heißt *linksseitiger Limes* von f in x_0.

Bemerkung: Man kann sich für den rechtsseiten Limes auf monoton fallende Folgen und für den linksseitigen Limes auf monoton wachsende Folgen (x_n) beschränken!

Schreibweisen: $\lim_{x \to x_0+0} f(x)$, $\lim_{x \downarrow x_0} f(x)$ für $\lim_{x \to x_0+} f(x)$ und $\lim_{x \to x_0-0} f(x)$, $\lim_{x \uparrow x_0} f(x)$ für $\lim_{x \to x_0-} f(x)$.

Beispiel: Sei $D = \mathbb{R}$ und $f : \mathbb{R} \to \mathbb{R}$ gegeben durch $f(x) := \begin{cases} 1 & , x \in (0,1) \\ 0 & , x \notin (0,1) \end{cases}$. Dann gilt $\lim_{x \to 0+} f(x) = 1$ und $\lim_{x \to 0-} f(x) = 0$.

Satz: Sei $f : D \to \mathbb{R}$ eine Funktion und $x_0 \in D$ Häufungspunkt sowohl von $D \cap (-\infty, x_0)$ als auch von $D \cap (x_0, \infty)$. f ist stetig in x_0 genau dann, wenn gilt:

$$\lim_{x \to x_0+} f(x) = f(x_0) = \lim_{x \to x_0-} f(x).$$

9.10. Monotone Funktionen und Stetigkeit: Sei $\emptyset \neq D \subseteq \mathbb{R}$ und $f : D \to \mathbb{R}$ eine Funktion. f heißt *monoton wachsend* [bzw. *monoton fallend*], falls für alle $x_1, x_2 \in D$ gilt:

$$x_1 \leq x_2 \Longrightarrow f(x_1) \leq f(x_2) \quad [\text{bzw. } f(x_1) \geq f(x_2)].$$

f heißt *streng monoton wachsend* [bzw. *streng monoton fallend*], falls für alle $x_1, x_2 \in D$ gilt:

$$x_1 < x_2 \Longrightarrow f(x_1) < f(x_2) \quad [\text{bzw. } f(x_1) > f(x_2)].$$

f heißt *monoton*, falls f monoton wachsend oder monoton fallend ist, und f heißt *streng monoton*, falls f streng monoton wachsend oder streng monoton fallend ist.

Beispiele: $\exp : \mathbb{R} \to \mathbb{R}$ ist streng monoton wachsend.
Die Funktion $\mathbb{R} \to \mathbb{R}$, $x \mapsto x^2$ ist nicht monoton.
Die Funktion $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto \begin{cases} 1 & , x \geq 0 \\ 0 & , x < 0 \end{cases}$ ist monoton wachsend, aber *nicht streng* monoton wachsend.

Satz: Sei $I \subseteq \mathbb{R}$ ein Intervall und $f : I \to \mathbb{R}$ streng monoton wachsend [bzw. fallend].

(a) Dann ist f injektiv, also $f : I \to f(I)$ bijektiv, und die Umkehrfunktion $f^{-1} : f(I) \to \mathbb{R}$ ist ebenfalls streng monoton wachsend [bzw. fallend].

(b) Ist f zusätzlich stetig, so ist $f(I)$ ein Intervall und $f^{-1} : f(I) \to \mathbb{R}$ ist stetig.

Beweis. zu (a): Injektivität ist klar. Ist f streng monoton wachsend, erhalten wir durch Kontraposition:

$$\forall x_1, x_2 \in I : \ f(x_1) \geq f(x_2) \Longleftrightarrow x_1 \geq x_2,$$

woraus strenge Monotonie von f^{-1} folgt.

zu (b): Wir verwenden 9.9 zum Nachweis der Stetigkeit. Sei $y_0 = f(x_0) \in f(I)$ und $(y_n) = (f(x_n))$ eine monotone Folge mit $y_n \to y_0$. Nach (a) ist dann auch (x_n) eine monotone Folge, die durch x_1 und x_0 beschränkt ist, und somit gegen ein $\alpha \in I$ konvergiert. Da f stetig ist, folgt $y_n = f(x_n) \to f(\alpha)$. Somit ist $f(\alpha) = y = f(x_0)$ und $x_0 = \alpha$, und wir haben $f^{-1}(y_n) = x_n \to x_0 = f^{-1}(y_0)$ gezeigt. □

9.11. Der Logarithmus: Nach Beispiel 9.8(1) gilt $\exp(\mathbb{R}) = (0, \infty)$. Die Abbildung $\exp : \mathbb{R} \to (0, \infty)$ ist bijektiv und stetig.

Definition: Die Umkehrfunktion der Exponentialfunktion $\exp : \mathbb{R} \to (0, \infty)$ heißt (*natürlicher*) *Logarithmus* $\ln : (0, \infty) \to \mathbb{R}$, dh also $\ln x := \ln x := \exp^{-1}(x)$ für $x \in (0, \infty)$.

Somit ist $\ln(\exp(x)) = x$ für alle $x \in \mathbb{R}$ und $\exp(\ln y) = y$ für alle $y \in (0, \infty)$.

Eigenschaften: ln ist streng monoton wachsend und stetig. Es gilt $\ln((0, \infty)) = \mathbb{R}$, $\ln(1) = 0$ und $\ln(e) = 1$, sowie $\ln x \to \infty$ für $x \to \infty$ und $\ln x \to -\infty$ für $x \to 0+$.

Für alle $x, y > 0$ gilt

$$\ln(xy) = \ln x + \ln y \quad \text{und} \quad \ln(x/y) = \ln x - \ln y.$$

Es gilt nämlich

$$\begin{aligned} \exp(\ln x + \ln y) &= \exp(\ln x)\exp(\ln y) = xy = \exp(\ln(xy)), \\ \exp(\ln x - \ln y) &= \exp(\ln x)\exp(-\ln y) = x\exp(\ln y)^{-1} = x/y = \exp(\ln(x/y)). \end{aligned}$$

9.12. Die allgemeine Potenz: Wir definieren für $a > 0$ die allgemeine Potenz:

$$a^x := \exp(x \ln a) \quad \text{für jedes } x \in \mathbb{R}.$$

Bemerkung: Nach 7.11 stimmt dies für $x \in \mathbb{Z}$ mit der bisherigen Definition überein. Für $x = 1/n$ mit $n \in \mathbb{N}$ ist (ebenfalls nach 7.11) $a^{1/n} = \sqrt[n]{a}$. Für $x = p/q \in \mathbb{Q}$ mit $p \in \mathbb{Z}$, $q \in \mathbb{N}$ folgt aus 7.11:

$$a^{p/q} = \sqrt[q]{a^p} = (\sqrt[q]{a})^p.$$

Für $a = e$ erhalten wir $e^x = \exp(x)$ wie in 7.11. Wir schreiben in Zukunft in der Regel e^x statt $\exp(x)$.

Eigenschaften: Für $a, b > 0$, $x, y \in \mathbb{R}$ gilt:

(1) $a^x > 0$;

(2) die Funktion $x \mapsto a^x$ ist auf $\mathbb{R}$ stetig (wegen 9.6);

(3) $a^{x+y} = e^{(x+y)\ln a} = e^{x \ln a} e^{y \ln a} = a^x a^y$;

(4) $a^{-x} = e^{-x \ln a} = (e^{x \ln a})^{-1} = (a^x)^{-1} = 1/a^x$;

(5) $\ln(a^x) = \ln(e^{x \ln a}) = x \ln a$;

(6) $(a^y)^x = e^{x \ln(a^y)} \overset{(5)}{=} e^{xy \ln a} = a^{xy}$.

(7) $(ab)^x = e^{x \ln(ab)} \overset{8.11}{=} e^{x \ln a} e^{x \ln b} = a^x b^x$.

Der allgemeine Logarithmus: Sei $a > 1$. Dann ist die Funktion $\mathbb{R} \to (0, \infty)$, $x \mapsto a^x$, streng monoton wachsend, stetig und bijektiv.

Die Umkehrfunktion $\log_a : (0, \infty) \to \mathbb{R}$ heißt *Logarithmus zur Basis a*. Es ist also $\log_a(a^x) = x$ für alle $x \in \mathbb{R}$ und $a^{\log_a y} = y$ für alle $y \in (0, \infty)$.

Satz: Für alle $y \in (0, \infty)$ gilt $\log_a y = \frac{\ln y}{\ln a}$.

Denn $a^{\frac{\ln y}{\ln a}} = \exp(\frac{\ln y}{\ln a} \ln a) = \exp(\ln y) = y$.

9.13. Abgeschlossene Mengen: Sei $D \subseteq \mathbb{R}$.

D heißt *abgeschlossen*, falls für jede Folge (x_n) in D mit $x_n \to x_0 \in \mathbb{R}$ gilt: $x_0 \in D$.

Beispiele: (1) Folgende Mengen sind abgeschlossen:

$$\mathbb{R},\ \emptyset,\ [a, b],\ (-\infty, a],\ [a, \infty),\ \{1/n : n \in \mathbb{N}\} \cup \{0\}.$$

(2) Folgende Mengen sind nicht abgeschlossen:

$$(a, b),\ (a, b],\ [a, b),\ (a, \infty),\ (-\infty, a),\ \{1/n : n \in \mathbb{N}\}.$$

Bemerkung: D ist abgeschlossen genau dann, wenn jeder Häufungspunkt von D zu D gehört.

9.14. Satz: Sei $\emptyset \neq D \subseteq \mathbb{R}$, D abgeschlossen und beschränkt, sowie $f : D \to \mathbb{R}$ eine stetige Funktion. Dann ist $f(D)$ abgeschlossen und beschränkt, und es gibt $x_1, x_2 \in D$ mit

$$\text{für alle } x \in D \text{ gilt: } \; f(x_1) \leq f(x) \leq f(x_2)$$

(dh "f nimmt auf D Maximum und Minimum an").

Beweis. (in der Vorlesung als Skizze)

1. Sei (y_n) eine Folge in $f(D)$. Zu jedem $n \in \mathbb{N}$ finden wir $x_n \in D$ mit $f(x_n) = y_n$. Dann ist (x_n) Folge in D. Da D beschränkt ist, gibt es nach Bolzano-Weiersraß eine Teilfolge $(x_{k(n)})$, die gegen ein $x_0 \in \mathbb{R}$ konvergiert. Da D abgeschlossen ist, gilt $x_0 \in D$. Da f stetig ist, folgt $y_{k(n)} = f(x_{k(n)}) \to f(x_0) \in f(D)$.

2. Wir finden (y_n) in $f(D)$ mit $y_n \to \sup f(D)$ bzw. $y_n \to \inf f(D)$. Nach 1. gibt es also $x_1, x_2 \in D$ mit $f(x_2) = \sup f(D)$, $f(x_1) = \inf f(D)$. Insbesondere ist $f(D)$ beschränkt.

3. Ist (y_n) eine Folge in D mit $y_n \to y_0 \in \mathbb{R}$, so finden wir nach 1. ein $x_0 \in D$ mit $y_n \to f(x_0)$. Also ist $y_0 = f(x_0) \in f(D)$, und $f(D)$ ist somit abgeschlossen. $\square$

Folgerung: Sind $a, b \in \mathbb{R}$ mit $a < b$ und ist $f : [a, b] \to \mathbb{R}$ stetig, so gibt es $c, d \in \mathbb{R}$ mit $c \leq d$ und $f([a, b]) = [c, d]$. (Verwende hier außerdem 9.8.)

Beispiel: $D = [1, 53]$ und $f(x) := x^3 \sin(e^x + \ln x)$.

10 Trigonometrische und Hyperbel-Funktionen

10.1. Die Zahl π: Wir beginnen mit einer **Vorbetrachtung** und interessieren uns für Nullstellen der Cosinus-Funktion. Nach 7.12 gilt für jedes $x \in \mathbb{R}$:

$$\cos x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!} = 1 - \frac{x^2}{2} + \frac{x^4}{4!} - \frac{x^6}{6!} + \dots$$

Die Reihe ist alternierend. Wie beim Leibnizkriterium 7.5 können wir durch Abbrechen der Reihe bei $n = n_0$ Abschätzungen nach oben oder nach unten angeben, wenn nur die Folge $(x^{2n}/(2n)!)_{n \geq n_0}$ monoton fallend ist. Wegen

$$\frac{x^{2n}}{(2n)!} \geq \frac{x^{2(n+1)}}{(2(n+1))!} \Leftrightarrow (2n+2)(2n+1) \geq x^2$$

gilt dies für $n_0 = 1$, wenn $x^2 \leq 12$, dh $|x| \leq 2\sqrt{3}$ ist.
Also gilt für alle $x \in [0, 2\sqrt{3}]$:

$$1 - \frac{x^2}{2} + \frac{x^4}{24} \geq \cos x \geq 1 - \frac{x^2}{2}.$$

Insbesondere ist $\cos\sqrt{2} \geq 0$ und $\cos(\sqrt{2}\sqrt{3-\sqrt{3}}) \leq 0$.[2] Nach dem Zwischenwertsatz 9.8 gibt es mindestens ein $x_0 \in [\sqrt{2}, \sqrt{2}\sqrt{3-\sqrt{3}}]$ mit $\cos x_0 = 0$.

Definition: Wir definieren $\pi/2$ als die kleinste Nullstelle des Cosinus im Intervall $[0,2]$, also

$$\frac{\pi}{2} := \inf\{x_0 \in [0,2] : \cos x_0 = 0\} = \min\{x_0 \in [0,2] : \cos x_0 = 0\}.$$

Beachte: Es gilt

$$\sqrt{2} \leq \pi/2 \leq \sqrt{2}\sqrt{3-\sqrt{3}} < 2 < 2\sqrt{3}$$

und $\sqrt{2} \approx 1.41421$, sowie $\sqrt{2}\sqrt{3-\sqrt{3}} \approx 1.59245$.

Bemerkung: Nach 7.12 gilt

$$\sin x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}, \quad \text{für alle } x \in \mathbb{R}.$$

[2] Es ist nämlich $\sqrt{2}$ Nullstelle von $1 - \frac{x^2}{2}$ und es gilt $1 - \frac{x^2}{2} + \frac{x^4}{24} = 0$ genau dann, wenn $(x^2-6)^2 = 12$, dh also wenn $x^2 = 6 \pm \sqrt{12}$. Wir interessieren uns für das Intervall $[0, 2\sqrt{3}]$, also bleibt nur $x = \sqrt{6-\sqrt{12}} = \sqrt{2}\sqrt{3-\sqrt{3}}$ als Nullstelle.

Die Argumente in der Vorbetrachtung zeigen dann

$$\sin x \geq x - \frac{x^3}{3!} \geq 0 \quad \text{für alle } x \in [0, \sqrt{6}].$$

Wegen $\sin^2 x + \cos^2 x = 1$ folgt also $\sin \frac{\pi}{2} = 1$.

10.2. Eigenschaften: Aus den Additionstheoremen in 7.12 erhalten wir:

(1) $\sin \pi = 0$, $\cos \pi = -1$;

(2) $\forall x \in \mathbb{R}$: $\sin(x + \pi) = -\sin x$, $\cos(x + \pi) = -\cos x$;

(3) $\sin 2\pi = 0$, $\cos 2\pi = 1$;

(4) $\forall x \in \mathbb{R}$: $\sin(x + 2\pi) = \sin x$, $\cos(x + 2\pi) = \cos x$ (dh sin und cos sind 2π-*periodisch*).

(5) $\forall x \in \mathbb{R}$: $\sin(\pi + x) = -\sin(\pi - x)$, $\cos(\pi + x) = \cos(\pi - x)$;

(6) $\forall x \in \mathbb{R}$: $\sin(\frac{\pi}{2} + x) = \sin(\frac{\pi}{2} - x)$, $\cos(\frac{\pi}{2} + x) = -\cos(\frac{\pi}{2} - x)$;

(7) $\forall x \in \mathbb{R}$: $\sin(x + \frac{\pi}{2}) = \cos x$.

Hieraus können wir alle Nullstellen von sin und cos bestimmen:

(8) Für $x \in \mathbb{R}$ gilt: $\cos x = 0 \Leftrightarrow$ es gibt $k \in \mathbb{Z}$ mit $x = (2k + 1)\frac{\pi}{2}$.

(9) Für $x \in \mathbb{R}$ gilt: $\sin x = 0 \Leftrightarrow$ es gibt $k \in \mathbb{Z}$ mit $x = k\pi$.

(10) Für $k \in \mathbb{Z}$ gilt: $\cos(k\pi) = (-1)^k$ und $\sin((2k + 1)\pi/2) = (-1)^k$.

Schließlich geben wir noch an:

(11) $\sin \frac{\pi}{4} = \cos \frac{\pi}{4} = \frac{\sqrt{2}}{2}$.

Außerdem erinnern wir an:

(12) Für alle $x \in \mathbb{R}$ gilt $\sin(-x) = -\sin x$ und $\cos(-x) = \cos x$.

(13) Für alle $x \in \mathbb{R}$ gilt $\sin(\frac{\pi}{2} - x) = \cos x$.

10.3. Monotonie bei $\sin$ **und** $\cos$**:** Es reicht, die Funktionen auf dem Intervall $[0, \pi/2]$ zu betrachten. Dabei ist $\sin x > 0$ für $x \in (0, \pi/2]$ und $\cos x > 0$ für $x \in [0, \pi/2)$. Es folgt $\cos x = \sqrt{1 - \sin^2 x} < 1 = \cos 0$ für $x \in (0, \pi/2]$, und für $x, x + h \in (0, \pi/2]$ mit $h > 0$ gilt also

$$\cos(x + h) = \cos x \underbrace{\cos h}_{\leq 1} - \underbrace{\sin x \sin h}_{>0} < \cos x,$$

dh cos **ist auf** $[0, \pi/2]$ **streng monoton fallend**.

Folglich ist $\sin x$ auf $[0, \pi/2]$ streng monoton wachsend (wegen $\sin x = \sqrt{1 - \cos^2 x}$ für $x \in [0, \pi/2]$).

10.4. Arcussinus und Arcuscosinus: Wegen 10.3 und 10.2(6) ist $\cos : [0, \pi] \to [-1, 1]$ streng monoton fallend, nach 9.2(2) ist diese Abbildung stetig und wegen 9.10 ist sie bijektiv.

Die Umkehrabbildung $\arccos : [-1, 1] \to [0, \pi]$ heißt *Arcuscosinus*.

Eigenschaften: $\arccos : [-1, 1] \to [0, \pi]$ ist streng monoton fallend, stetig und bijektiv. Es gilt $\arccos(-1) = \pi$, $\arccos 0 = \pi/2$, $\arccos 1 = 0$.

Wegen 10.3 und 10.2(12) ist $\sin : [-\pi/2, \pi/2] \to [-1, 1]$ streng monoton wachsend. Auch diese Funktion ist stetig und bijektiv.

Die Umkehrfunktion $\arcsin : [-1, 1] \to [-\pi/2, \pi/2]$ heißt *Arcussinus*.

Eigenschaften: Die Funktion $\arcsin : [-1, 1] \to [-\pi/2, \pi/2]$ ist streng monoton wachsend, stetig und bijektiv. Es gilt $\arcsin(1) = \pi/2$, $\arcsin 0 = 0$ und $\arcsin(-x) = -\arcsin x$ für alle $x \in [-1, 1]$.

Achtung! Wegen 10.2 sind für jedes $k \in \mathbb{Z}$ auch die Abbildungen

$$\cos : [k\pi, (k + 1)\pi] \to [-1, 1], \qquad \sin : [(2k - 1)\pi/2, (2k + 1)\pi/2] \to [-1, 1]$$

streng monoton, stetig und bijektiv. Für $k \neq 0$ sind ihre Umkehrabbildungen verschieden von den eben definierten Funktionen arccos und arcsin!

10.5. Der Tangens: Die Funktion

$$\tan : \mathbb{R} \setminus (\frac{\pi}{2} + \pi\mathbb{Z}) \to \mathbb{R}, \quad x \mapsto \tan x := \frac{\sin x}{\cos x}$$

heißt *Tangens*. (Beachte, dass die Menge $\frac{\pi}{2} + \pi\mathbb{Z}$ genau die Nullstellen des Cosinus enthält.)

Eigenschaften: Es gilt $\tan 0 = 0$, $\tan \pi/4 = 1$, sowie für alle x im Definitionsbereich, dh für alle $x \in \mathbb{R} \setminus (\frac{\pi}{2} + \pi\mathbb{Z})$:

$$\tan(-x) = -\tan x \qquad \tan(x + \pi) = \tan x.$$

Somit ist der Tangens eine π-periodische Funktion.

Außerdem ist tan auf $(-\pi/2, \pi/2)$ streng monoton wachsend mit $\tan x \to \infty$ für $x \to \pi/2-$ und $\tan x \to -\infty$ für $x \to -\pi/2+$.

Folglich ist $\tan : (-\pi/2, \pi/2) \to \mathbb{R}$ streng monoton wachsend, stetig und bijektiv. Die Umkehrfunktion $\arctan : \mathbb{R} \to (-\pi/2, \pi/2)$ heißt *Arcustangens*.

Eigenschaften: Es gilt $\arctan 0 = 0$, $\arctan 1 = \pi/4$ und $\arctan(-x) = -\arctan x$ für alle $x \in \mathbb{R}$.

Bemerkung: Sei $k \in \mathbb{Z}$. Man sieht leicht, dass die Umkehrabbildung von $\tan : (k\pi - \pi/2, k\pi + \pi/2) \to \mathbb{R}$ gegeben ist durch

$$\mathbb{R} \to (k\pi - \pi/2, k\pi + \pi/2), x \mapsto k\pi + \arctan x.$$

10.6. Anwendung (Polarkoordinaten): Für jede komplexe Zahl $z \in \mathbb{C} \setminus \{0\}$ gibt es genau ein $r \in (0, \infty)$ und genau einen Winkel $\varphi \in (-\pi, \pi]$ mit $z = re^{i\varphi}$. Dabei heißt $r = |z|$ *Länge von* z und $\varphi =: \arg z$ heißt das *Argument von* z.

Satz: Für jedes Paar $(a, b) \in \mathbb{R}^2$ mit $a^2 + b^2 = 1$ gibt es genau ein $\varphi \in (-\pi, \pi]$ mit $\cos\varphi = a$ und $\sin\varphi = b$.

Bemerkung: Dabei ist $\varphi = \arcsin b$, falls $a \geq 0$ ist, und $\varphi = \arccos a$, falls $b \geq 0$ ist. Sind $a, b < 0$, so ist $\varphi = \arccos |a| - \pi = \arcsin |b| - \pi$.

Alternativ hat man $\varphi = \arctan(b/a)$ für $a > 0$, $\varphi = b\pi/2$ für $a = 0$ (dann $b \in \{-1, 1\}$), sowie $\varphi = \pi + \arctan(b/a)$ für $a < 0$, $b \geq 0$, und $\varphi = -\pi + \arctan(b/a)$ für $a, b < 0$.

Bemerkung: Es gilt $e^{i\psi} = 1$ genau dann, wenn $\psi \in 2\pi\mathbb{Z}$ ist. Insbesondere gibt es zu jedem $z \in \mathbb{C} \setminus \{0\}$ unendlich viele $\psi \in \mathbb{R}$ mit $z = |z|e^{i\psi}$.

Multiplikation komplexer Zahlen: Seien $z = re^{i\varphi}, w = se^{i\psi} \in \mathbb{C} \setminus \{0\}$ mit $\varphi, \psi \in (-\pi, \pi]$. Dann gilt

$$zw = rse^{i(\varphi+\psi)},$$

dh Multiplikation mit w dreht z um den Winkel $\psi = \arg w$. Ist aber $\varphi + \psi \notin (-\pi, \pi]$, so ist $\arg z + \arg w \neq \arg(zw)$! Als Beispiel betrachte man $z = w = -1 + i$ mit $\arg z = \arg w = 3\pi/4$ und $\arg z + \arg w = 3\pi/2 \neq -\pi/2 = \arg(-2i) = \arg(wz)$ oder $w = z = -1$ mit $\arg z = \arg w = \pi$ und $\arg z + \arg w = 2\pi \neq 0 = \arg 1 = \arg(wz)$.

Wurzeln komplexer Zahlen: Sei $z \in \mathbb{C} \setminus \{0\}$ und $n \in \mathbb{N}$ mit $n \geq 2$. Dann gibt es genau n verschiedene komplexe Zahlen $w_1, w_2, \ldots, w_n \in \mathbb{C} \setminus \{0\}$ mit $w_k^n = z$ $(k = 1, 2, \ldots, n)$, dh es n verschiedene n-te Wurzeln von z. Schreibe dazu $z = re^{i\varphi}$ mit $r > 0$ und $\varphi \in (-\pi, \pi]$ und setze

$$w_k := \sqrt[n]{r}\, e^{i\frac{\varphi+(k-1)2\pi}{n}}, \quad k = 1, 2, \ldots, n.$$

Beispiele: 1) $n = 2$, allgemeines $z \in \mathbb{C} \setminus \{0\}$: Die beiden Quadratwurzeln von z sind

$$w_1 = \sqrt{|z|}e^{i\arg(z)/2}, \quad w_2 = \sqrt{|z|}e^{i(\arg(z)+2\pi)/2} = \sqrt{|z|}e^{i\arg(z)}\underbrace{e^{i\pi}}_{=-1} = -w_1.$$

2) $n = 3$, $z = 1$: Die drei dritten Wurzeln von 1 sind

$$w_1 = 1, \qquad w_2 = e^{i\frac{2\pi}{3}}, \qquad w_3 = e^{i\frac{4\pi}{3}} = e^{-i\frac{2\pi}{3}} = \overline{w_2}.$$

3) $n = 4$, $z = 1$: Die vier vierten Wurzeln von 1 sind

$$w_1 = 1, \quad w_2 = e^{i\frac{\pi}{2}} = i, \quad w_3 = e^{i\pi} = 1, \quad w_4 = e^{i\frac{3\pi}{2}} = -i.$$

10.7. Hyperbelfunktionen: Definiere für $x \in \mathbb{R}$:

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

(*Cosinus hyperbolicus* und *Sinus hyperbolicus*).

Dann gilt $\cosh 0 = 1$ und $\sinh 0 = 0$, sowie

$$\forall x \in \mathbb{R} : \cosh(-x) = \cosh x, \quad \sinh(-x) = -\sinh x.$$

Außerdem gilt

$$\begin{array}{lll} \cosh x \to \infty, & \sinh x \to \infty & \text{für} \quad x \to \infty \\ \cosh x \to \infty, & \sinh x \to -\infty & \text{für} \quad x \to -\infty, \end{array}$$

sowie

$$\cosh^2 x - \sinh^2 x = 1 \quad \text{für alle } x \in \mathbb{R}.$$

Reihendarstellungen: Für alle $x \in \mathbb{R}$ gilt

$$\cosh x = \sum_{n=0}^{\infty} \frac{x^{2n}}{(2n)!}, \qquad \sinh x = \sum_{n=0}^{\infty} \frac{x^{2n+1}}{(2n+1)!}.$$

Folgerung: Die Funktionen

$$\cosh : [0, \infty) \to [1, \infty) \quad \text{und} \quad \sinh : \mathbb{R} \to \mathbb{R}$$

sind streng monoton wachsend, stetig und bijektiv.

Additionstheoreme: Für alle $x, y \in \mathbb{R}$ gilt

$$\begin{aligned} \cosh(x+y) &= \cosh x \cosh y + \sinh x \sinh y, \\ \sinh(x+y) &= \sinh x \cosh y + \cosh x \sinh y. \end{aligned}$$

Definition: Die Funktion

$$\tanh : \mathbb{R} \to \mathbb{R}, x \mapsto \tanh x := \frac{\sinh x}{\cosh x} = \frac{e^x - e^{-x}}{e^x + e^{-x}},$$

heißt *Tangens hyperbolicus*.

Es ist $\tanh(-x) = -\tanh x$ für alle $x \in \mathbb{R}$ und $\tanh 0 = 0$. Außerdem gilt $\tanh x \to 1$ für $x \to \infty$ und $\tanh x \to -1$ für $x \to -\infty$.

Die Funktion $\tanh : \mathbb{R} \to (-1, 1)$ ist streng monoton wachsend, stetig und bijektiv.

10.8. Areafunktionen: Die Umkehrfunktionen von $\sinh : \mathbb{R} \to \mathbb{R}$, $\cosh : [0, \infty) \to [1, \infty)$ und $\tanh : \mathbb{R} \to (-1, 1)$ heißen $\operatorname{Arsinh} : \mathbb{R} \to \mathbb{R}$ (*Areasinus*), $\operatorname{Arcosh} : [1, \infty) \to [0, \infty)$ (*Areacosinus*) und $\operatorname{Artanh} : (-1, 1) \to \mathbb{R}$ (*Areatangens*). Diese Funktionen sind jeweils streng monoton wachsend, stetig und bijektiv.

10.9. Weitere Funktionen: Es gibt noch weitere Funktionen, auf die wir hier nicht näher eingehen, z.B. $\cot x = 1/\tan x$, $\sec x = 1/\cos x$, $\operatorname{sech} x = 1/\cosh x$ etc.

11 Differentialrechnung

In diesem Abschnitt sei $I \subseteq \mathbb{R}$ stets ein Intervall mit mindestens zwei Punkten.

11.1. Differenzierbarkeit: Sei $f : I \to \mathbb{R}$ eine Funktion und $x_0 \in I$.

Idee: Approximation von f "in der Nähe von x_0" durch eine *Gerade* (da Geraden leichter zu behandeln sind):

$$f(x) = f(x_0) + \frac{f(x) - f(x_0)}{x - x_0}(x - x_0) \approx f(x_0) + a(x - x_0).$$

Anschaulich sollte dafür a die "Steigung" von f in x_0 sein.

Definition: f heißt in $x_0 \in I$ *differenzierbar* (*dbar*), falls der Limes $\lim_{x\to x_0} \frac{f(x)-f(x_0)}{x-x_0} \in \mathbb{R}$ existiert. Dieser Grenzwert heißt *die Ableitung von f in x_0*, Bezeichnung: $f'(x_0)$.

Die Funktion f heißt *auf I differenzierbar*, falls f in jedem $x \in I$ differenzierbar ist. In diesem Fall heißt $f' : I \to \mathbb{R}$, $x \mapsto f'(x)$, die *Ableitung von f auf I*.

Bemerkung: Es ist $\lim_{x\to x_0} \frac{f(x)-f(x_0)}{x-x_0} = \lim_{h\to 0} \frac{f(x_0+h)-f(x_0)}{h}$. Außerdem ist $\lim_{x\to x_0} \frac{f(x)-f(x_0)}{x-x_0} = a$ genau dann, wenn

$$\forall \varepsilon > 0\, \exists \delta > 0\, \forall x \in I \text{ mit } |x - x_0| < \delta : |f(x) - f(x_0) - a(x - x_0)| \leq \varepsilon |x - x_0|.$$

Beispiele: (1) $f : I \to \mathbb{R}$, $f(x) = c \in \mathbb{R}$, ist auf I differenzierbar mit $f' = 0$ auf I.

(2) $f : \mathbb{R} \to \mathbb{R}$, $f(x) = |x|$ ist in $x_0 = 0$ **nicht** differenzierbar, denn für $h \neq 0$ ist

$$\frac{f(h) - f(0)}{h} = \frac{|h|}{h} = \begin{cases} 1 & , h > 0 \\ -1 & h < 0 \end{cases}$$

Also existiert $\lim_{h\to 0} \frac{f(h)-f(0)}{h}$ nicht.

(3) $I = \mathbb{R}$, $n \in \mathbb{N}$, $f(x) = x^n$. f ist auf $\mathbb{R}$ differenzierbar mit $f'(x_0) = n x_0^{n-1}$ für jedes $x_0 \in \mathbb{R}$. Für $x \neq x_0$ ist nämlich nach 4.11(1)

$$\frac{x^n - x_0^n}{x - x_0} = \sum_{k=0}^{n-1} x^{n-1-k} x_0^k \to n x_0^{n-1} \ (x \to x_0).$$

(4) exp, sin, cos, sinh, cosh sind auf $\mathbb{R}$ differenzierbar mit $\exp' = \exp$, $\sin' = \cos$, $\cos' = -\sin$, $\sinh' = \cosh$ und $\cosh' = \sinh$. Für $x, h \in \mathbb{R}$ mit $h \neq 0$ gilt nämlich nach 7.11(1) und 9.4:

$$\frac{e^{x+h} - e^x}{h} = e^x \underbrace{\frac{e^h - 1}{h}}_{\to 1} \to e^x \ (h \to 0),$$

sowie nach 7.12(3) (Additionstheoreme) und 9.4:

$$\begin{aligned}\frac{\sin(x+h)-\sin x}{h} &= \frac{\sin x\cos h+\cos x\sin h-\sin x}{h}\\ &= \sin x\underbrace{\frac{\cos h-1}{h}}_{\to 0}+\cos x\underbrace{\frac{\sin h}{h}}_{\to 1}\to\cos x\ (h\to 0).\end{aligned}$$

Die Beweise für cos, sinh und cosh sind analog.

Satz: Ist f in x_0 differenzierbar, so ist f in x_0 stetig.

Beweis. Da I ein Intervall ist, ist $x_0 \in I$ Häufungspunkt von I. Für $x \in I \setminus \{x_0\}$ ist dann

$$f(x)-f(x_0)=\frac{f(x)-f(x_0)}{x-x_0}(x-x_0)\to f'(x_0)\cdot 0=0\ (x\to x_0).$$

Also ist f in x_0 stetig. □

11.2. Ableitungsregeln: Seien $f, g : I \to \mathbb{R}$ Funktionen, die in $x_0 \in I$ differenzierbar sind, und $\alpha, \beta \in \mathbb{R}$. Dann gilt:

(1) $\alpha f + \beta g$ ist differenzierbar in $x_0 \in I$ und

$$(\alpha f+\beta g)'(x_0)=\alpha f'(x_0)+\beta g'(x_0).$$

(2) fg ist differenzierbar in x_0 und

$$(fg)'(x_0)=f'(x_0)g(x_0)+f(x_0)g'(x_0) \qquad \textit{Produktregel.}$$

(3) Ist $g(x_0) \neq 0$, so gibt es ein $\delta > 0$ mit $g(x) \neq 0$ für alle $x \in I \cap U_\delta(x_0) =: J$. Die Funktion $\frac{f}{g} : J \to \mathbb{R}$ ist in x_0 differenzierbar und

$$\left(\frac{f}{g}\right)'(x_0)=\frac{f'(x_0)g(x_0)-f(x_0)g'(x_0)}{g(x_0)^2} \qquad \textit{Quotientenregel.}$$

Bemerkung: (1) besagt, dass $V := \{f : I \to \mathbb{R} : f \text{ differenzierbar in } x_0\}$ ein $\mathbb{R}$-Vektorraum ist und die Abbildung $V \to \mathbb{R}$, $f \mapsto f'(x_0)$ linear ist.

Beweis. (1) ist klar wegen 6.3 und 9.4.

(2) Für $x \in I \setminus \{x_0\}$ gilt

$$\frac{(fg)(x)-(fg)(x_0)}{x-x_0}=\frac{f(x)-f(x_0)}{x-x_0}g(x)+f(x_0)\frac{g(x)-g(x_0)}{x-x_0},$$

woraus für $x \to x_0$ die Behauptung folgt (beachte, dass $g(x) \to g(x_0)$ für $x \to x_0$, da g nach 11.1 in x_0 stetig ist).

(3) Wegen (2) reicht $f = 1$. Die Existenz von δ erhalten wir aus 9.3, da g in x_0 stetig ist. Beachte, dass J ein Intervall ist. Für $x \in J \setminus \{x_0\}$ ist

$$\frac{1/g(x) - 1/g(x_0)}{x - x_0} = \frac{\frac{g(x_0)-g(x)}{x-x_0}}{g(x)g(x_0)},$$

woraus für $x \to x_0$ die Behauptung folgt. Wieder beachte man $g(x) \to g(x_0)$ für $x \to x_0$. $\square$

Beispiele: $\tan = \frac{\sin}{\cos}$ und $\tanh = \frac{\sinh}{\cosh}$ sind auf $\mathbb{R} \setminus (\pi\mathbb{Z} + \pi/2)$ bzw. auf $\mathbb{R}$ differenzierbar, und es gilt:

$$\tan' = \frac{1}{\cos^2} = 1 + \tan^2 \text{ auf } \mathbb{R} \setminus (\pi\mathbb{Z} + \pi/2), \quad \tanh' = \frac{1}{\cosh^2} = 1 - \tanh^2 \text{ auf } \mathbb{R}.$$

Es ist nämlich

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

Der Beweis für tanh ist analog.

11.3. Kettenregel: Sei $f : I \to \mathbb{R}$ in $x_0 \in I$ differenzierbar und $J \subseteq \mathbb{R}$ ein Intervall mit $f(I) \subseteq J$. Sei $g : J \to \mathbb{R}$ in $y_0 := f(x_0)$ differenzierbar. Dann ist $g \circ f : I \to \mathbb{R}$ in x_0 differenzierbar und

$$(g \circ f)'(x_0) = g'(f(x_0)) \cdot f'(x_0)$$

("äußere Ableitung mal innere Ableitung").

Beweis. Die Idee ist zu schreiben

$$\frac{g(f(x)) - g(f(x_0))}{x - x_0} = \frac{g(f(x)) - g(f(x_0))}{f(x) - f(x_0)} \cdot \frac{f(x) - f(x_0)}{x - x_0}.$$

Da man dabei eventuell durch Null dividiert, setzen wir

$$q : J \to \mathbb{R}, q(y) := \begin{cases} \frac{g(y)-g(y_0)}{y-y_0} & , y \neq y_0 \\ g'(y_0) & , y = y_0 \end{cases}.$$

Dann gilt $q(y) \to g'(y_0)$ für $y \to y_0$, also auch $q(f(x)) \to g'(f(x_0))$ für $x \to x_0$, da f in x_0 stetig ist. Außerdem gilt $g(y) - g(y_0) = q(y)(y - y_0)$ für **alle** $y \in J$, also

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

$\square$

Beispiel: Sei $a > 0$ und $h(x) = a^x = e^{x \ln a}$ für $x \in \mathbb{R}$. Hier ist $g(y) = e^y$ und $f(x) = x \ln a$. Dann ist h auf $\mathbb{R}$ differenzierbar und für jedes $x \in \mathbb{R}$ gilt:

$$h'(x) = (a^x)' = g'(x \ln a) f'(x) = a^x \ln a.$$

11.4. Satz über die Umkehrfunktion: Sei $f : I \to \mathbb{R}$ stetig und streng monoton auf I. Ist f in $x_0 \in I$ differenzierbar und $f'(x_0) \neq 0$, so ist die Umkehrfunktion $f^{-1} : f(I) \to \mathbb{R}$ in $y_0 := f(x_0)$ differenzierbar und

$$(f^{-1})'(y_0) = \frac{1}{f'(x_0)} = \frac{1}{f'(f^{-1}(y_0))}.$$

Beweis. Nach 9.8 ist $f(I)$ ein Intervall. Sei $y \in f(I)$ und $x := f^{-1}(y)$. Dann gilt

$$\frac{f^{-1}(y) - f^{-1}(y_0)}{y - y_0} = \frac{x - x_0}{f(x) - f(x_0)} \to \frac{1}{f'(x_0)} \quad (y \to y_0),$$

da wegen der Stetigkeit von f^{-1} aus $y \to y_0$ folgt $x \to x_0$. □

Beispiele: Die Funktionen $\ln : (0, \infty) \to \mathbb{R}$, $\arctan : \mathbb{R} \to (-\pi/2, \pi/2)$, $\operatorname{Arsinh} : \mathbb{R} \to \mathbb{R}$, $\operatorname{Artanh} : (-1, 1) \to \mathbb{R}$ sind differenzierbar. Dabei ist ln die Umkehrfunktion von $f(x) = e^x$ und $f'(x) = e^x$, also

$$\ln' y = (e^{\ln y})^{-1} = \frac{1}{y} \quad \text{für jedes } y > 0.$$

Weiter gilt

$$\begin{aligned}
\arctan'(y) &= \frac{1}{1 + y^2}, \quad \operatorname{Arsinh}'(y) = \frac{1}{\sqrt{1 + y^2}}, \quad y \in \mathbb{R} \\
\operatorname{Artanh}'(y) &= \frac{1}{1 - y^2}, \quad |y| < 1.
\end{aligned}$$

Die Funktionen $\arcsin : (-1, 1) \to (-\pi/2, \pi/2)$, $\arccos : (-1, 1) \to (0, \pi)$ und $\operatorname{Arcosh} : (1, \infty) \to (0, \infty)$ sind differenzierbar mit

$$\begin{aligned}
\arcsin'(y) &= \frac{1}{\sqrt{1 - y^2}}, \quad \arccos'(y) = -\frac{1}{\sqrt{1 - y^2}}, \quad |y| < 1 \\
\operatorname{Arcosh}'(y) &= \frac{1}{\sqrt{y^2 - 1}}, \quad y > 1.
\end{aligned}$$

Anwendung: Für jedes $x \in \mathbb{R}$ gilt:

$$\lim_{n \to \infty} \left(1 + \frac{x}{n}\right)^n = e^x.$$

Beweis. Für $n > -x$ ist die linke Seite $= \exp(n \ln(1 + \frac{x}{n}))$. Also reicht zu zeigen

$$n \ln(1 + \frac{x}{n}) \to x \quad (n \to \infty).$$

Das ist klar für $x = 0$. Für $x \neq 0$ erhalten wir

$$\lim_{n\to\infty} \frac{\ln(1 + \frac{x}{n})}{\frac{x}{n}} = \lim_{h\to 0} \frac{\ln(1+h) - \ln 1}{h} = \ln'(1) = \frac{1}{t}\Big|_{t=1} = 1,$$

woraus die Behauptung folgt. □

Anwendung: Für $\alpha \in \mathbb{R}$ betrachten wir $x \mapsto x^\alpha$ auf $(0, \infty)$. Unter Verwendug der Schreibweise $\frac{d}{dx} f(x)$ für $f'(x)$ haben wir

$$\frac{d}{dx}\Big(x^\alpha\Big) = \frac{d}{dx}\Big(e^{\alpha \ln x}\Big) = e^{\alpha \ln x} \cdot \frac{\alpha}{x} = \alpha x^{\alpha - 1}.$$

11.5. Lokale Extremstellen: Sei $D \subseteq \mathbb{R}$ mit $D \neq \emptyset$, $g : D \to \mathbb{R}$ eine Funktion und $x_0 \in D$.

Definition: g hat in x_0 ein *lokales Maximum* [*Minimum*], falls es ein $\delta > 0$ gibt mit

$$\forall x \in D \cap U_\delta(x_0) : g(x) \leq g(x_0) \quad [\text{ bzw. } g(x) \geq g(x_0)],$$

dh $\forall x \in D$: $|x - x_0| < \delta \Rightarrow g(x) \leq g(x_0)$ [bzw. $g(x) \geq g(x_0)$].

Ein lokales Maximum/Minimum wird auch als *relatives Maximum/Minimum* bezeichnet.

Ein *relatives* oder *lokales Extremum* ist ein lokales Maximum oder ein lokales Minimum.

11.6. Satz: Die Funktion $f : I \to \mathbb{R}$ habe in $x_0 \in I$ ein lokales Extremum und sei in x_0 differenzierbar. Gibt es ein $\delta > 0$ mit $U_\delta(x_0) \subseteq I$, so ist $f'(x_0) = 0$.

Bemerkung: Ist $M \subseteq \mathbb{R}$ eine Menge, so heißen Punkte $x_0 \in M$, für die es ein $\delta > 0$ mit $U_\delta(x_0) \subseteq M$ gibt, *innere Punkte von M*.

Da I ein Intervall ist, ist $x_0 \in I$ genau dann ein innerer Punkt von I, wenn $x_0 \notin \{\sup I, \inf I\}$ gilt.

Beweis des Satzes. Wir nehmen an, dass f in x_0 ein lokales Maximum hat (sonst betrachte $-f$). Wir können weiter annehmen, dass das δ aus 11.5 mit dem δ aus dem Satz übereinstimmt (sonst betrachte das Minimum von beiden). Für $x \in (x_0, x_0 + \delta)$ gilt dann $f(x) \leq f(x_0)$, also $\frac{f(x) - f(x_0)}{x - x_0} \leq 0$ und damit $f'(x_0) = \lim_{x\to x_0+} \frac{f(x)-f(x_0)}{x - x_0} \leq 0$. Für $x \in (x_0 - \delta, x_0)$ gilt $f(x) \leq f(x_0)$ und $\frac{f(x)-f(x_0)}{x-x_0} \geq 0$, woraus $f'(x_0) \geq 0$ folgt. □

11.7. Mittelwertsatz der Differentialrechnung (MWS): Sei $f : [a, b] \to \mathbb{R}$ stetig und auf (a, b) differenzierbar. Dann gibt es ein $\xi \in (a, b)$ mit

$$f'(\xi) = \frac{f(b) - f(a)}{b - a}.$$

Beweis. Wir betrachten zunächst den Fall $f(a) = f(b)$. Nach 9.14 finden wir $x_1, x_2 \in [a, b]$ mit $f(x_1) \le f(x) \le f(x_2)$ für alle $x \in [a, b]$. Somit hat f in x_1 ein lokales Minimum und in x_2 ein lokales Maximum.

Fall 1, $x_1 \in (a, b)$: Dann wenden wir 11.6 an und erhalten $f'(x_1) = 0$.

Fall 2, $x_1 \in \{a, b\}$, $x_2 \in (a, b)$: Wir wenden 11.6 an und erhalten $f'(x_2) = 0$.

Fall 3, $x_1, x_2 \in \{a, b\}$: Dann ist f konstant und $f'(x) = 0$ für jedes $x \in (a, b)$.

Im allgemeinen Fall setzen wir $g(x) := f(x) - \frac{f(b)-f(a)}{b-a}(x - a)$ für $x \in [a, b]$ und wenden das Gezeigte auf g an (es ist $g(a) = f(a) = g(b)$). Wir erhalten $\xi \in (a, b)$ mit

$$0 = g'(\xi) = f'(\xi) - \frac{f(b) - f(a)}{b - a}$$

wie gewünscht. □

11.8. Folgerungen: Seien $f, g : I \to \mathbb{R}$ auf I differenzierbar.

(1) f ist auf I konstant $\iff f' = 0$ auf I.

(2) Ist $f' = g'$ auf I, so gibt es ein $c \in \mathbb{R}$ mit $f = g + c$ auf I.

(3) Ist $f' \ge 0$ [bzw. $f' \le 0$, $f' > 0$, $f' < 0$] auf I, so ist f auf I monoton wachsend [bzw. monoton fallend, streng monoton wachsend, streng monoton fallend].

Beweis. (1) "$\Leftarrow$": Nach MWS ist $f(b) = f(a)$ für alle $a, b \in I$.

(2) Wende (1) an auf $f - g$.

(3) Ist $f' \ge 0$ auf I, so gilt für $x, y \in I$ mit $x < y$:

$$\frac{f(y) - f(x)}{y - x} = f'(\xi) \ge 0$$

für ein $\xi \in (x, y)$. Es folgt $f(y) \ge f(x)$. Die anderen Aussagen beweist man analog. □

Anwendung: Sei $a \in \mathbb{R}$ und $\phi : I \to \mathbb{R}$ eine differenzierbare Funktion mit $\phi' = a\phi$ auf I, sowie $x_0 \in I$. Dann gilt $\phi(x) = \phi(x_0)e^{a(x-x_0)}$ für jedes $x \in I$: Setzt man $\psi(x) := \phi(x)e^{-ax}$, $x \in I$, so ist nämlich ψ differenzierbar auf I mit $\psi' = 0$ auf I, und somit

$$\phi(x)e^{-ax} = \psi(x) = \psi(x_0) = \phi(x_0)e^{-ax_0} \text{ für jedes } x \in I.$$

Somit hat für feste $x_0 \in I$, $c \in \mathbb{R}$ das Anfangswertproblem

$$y' = ay \text{ auf } I, \quad y(x_0) = c,$$

genau eine differenzierbare Lösung $\phi : I \to \mathbb{R}$, gegeben durch $\phi(x) = ce^{a(x-x_0)}$, $x \in I$.

11.9. Die Regeln von de l'Hospital: Seien $a \in \mathbb{R}$, $b \in \mathbb{R} \cup \{\infty\}$ und $L \in \mathbb{R} \cup \{-\infty, \infty\}$. Seien $f, g : (a, b) \to \mathbb{R}$ auf (a, b) differenzierbar mit $g'(x) \neq 0$ für alle $x \in (a, b)$. Weiter sei $\lim_{x \to b} \frac{f'(x)}{g'(x)} = L$.

(a) Ist $\lim_{x\to b} f(x) = \lim_{x\to b} g(x) = 0$, so gilt

$$\frac{f(x)}{g(x)} \to L \quad (x \to b).$$

(b) Ist $\lim_{x\to b} g(x) = \pm\infty$, so gilt

$$\frac{f(x)}{g(x)} \to L \quad (x \to b).$$

Beweis. nur (a) und $b \in \mathbb{R}$: Setze $f(b) = g(b) = 0$. Sei (x_n) Folge in (a, b) mit $x_n \to b$ und $n \in \mathbb{N}$. Setze $h(x) := g(x)f(x_n) - f(x)g(x_n)$ für $x \in [x_n, b]$. Dann ist h auf $[x_n, b]$ stetig und in (x_n, b) differenzierbar mit $h(x_n) = 0 = h(b)$. Nach MWS finden wir ein $\xi_n \in (x_n, b)$ mit $h'(\xi_n) = 0$, dh mit

$$\frac{f(x_n)}{g(x_n)} = \frac{f'(\xi_n)}{g'(\xi_n)}.$$

Die Behauptung folgt nun für $n \to \infty$. $\square$

Beispiele: (1) Für $a, b > 0$ gilt

$$\lim_{x\to 0+} \frac{a^x - b^x}{x} = \lim_{x\to 0+} \frac{a^x \ln a - b^x \ln b}{1} = \ln a - \ln b.$$

(2) Es ist

$$\lim_{x\to\infty} \frac{\ln x}{x} = \lim_{x\to\infty} \frac{1/x}{1} = 0.$$

(3) Es ist

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

(4) Aus (3) folgt mit der Stetigkeit der Exponentialfunktion:

$$\lim_{x\to 0+} x^x = \lim_{x\to 0+} e^{x \ln x} = e^0 = 1.$$

Bemerkung: Hier ist es jeweils so, dass erst die Existenz des letzten Limes die Existenz des ersten Limes garantiert (vgl. die Regeln oben).

Der folgende Satz kann in Anwendungen nützlich sein.

11.10. Satz: Sei $f : [a,b] \to \mathbb{R}$ eine Funktion und $x_0 \in (a,b)$. Weiter sei f auf (a, x_0) und (x_0, b) differenzierbar. Existiert der Grenzwert $\alpha := \lim_{x\to x_0} f'(x) \in \mathbb{R}$ und ist f in x_0 stetig, so ist f in x_0 differenzierbar mit $f'(x_0) = \alpha$ und f' ist in x_0 stetig.

Beweis. Für $x \in (a,b) \setminus \{x_0\}$ gilt nach dem Mittelwertsatz

$$\frac{f(x) - f(x_0)}{x - x_0} = f'(\xi(x))$$

für ein $\xi(x)$ zwischen x und x_0. Für $x \to x_0$ hat man $\xi(x) \to x_0$ und also $f'(\xi(x)) \to \alpha$ nach Voraussetzung. $\square$

Beispiel: Die Funktion $f : \mathbb{R} \to \mathbb{R}$, $f(x) := \begin{cases} e^{-1/x} & ,x > 0 \\ 0 & ,x \le 0 \end{cases}$ ist auf $\mathbb{R}$ stetig differenzierbar. Es gilt

$$f'(x) = \begin{cases} x^{-2}e^{-1/x} & ,x > 0 \\ 0 & ,x < 0 \end{cases}.$$

Wegen $\lim_{x\to 0+} f(x) = \lim_{x\to 0+} e^{-1/x} = 0$ ist f in 0 stetig. Wegen $\lim_{x\to 0+} f'(x) = \lim_{x\to 0+} x^{-2}e^{-1/x} = \lim_{s\to\infty} s^2 e^{-s} = 0$ [3] ist f nach dem Satz in 0 differenzierbar mit $f'(0) = 0$ und f' ist stetig auf $\mathbb{R}$.

Beispiel zur Warnung: Die Funktion $f : \mathbb{R} \to \mathbb{R}$, $f(x) := \begin{cases} x^2 \sin(1/x) & ,x \neq 0 \\ 0 & ,x = 0 \end{cases}$ ist auf $\mathbb{R}$ differenzierbar. In Punkten $x_0 \neq 0$ ist dies klar und es gilt

$$f'(x) = 2x\sin(1/x) - \cos(1/x), \quad x \neq 0.$$

Wir untersuchen f auf Differenzierbarkeit in $x_0 = 0$: Für $x \neq 0$ gilt

$$\frac{f(x) - f(0)}{x - 0} = x\sin(1/x), \quad \left|\frac{f(x) - f(0)}{x - 0}\right| = |x||\sin(1/x)| \le |x| \to 0 \ (x \to 0),$$

also ist f in 0 differenzierbar und $f'(0) = 0$. Man beachte, dass der Limes $\lim_{x\to 0} f'(x)$ nicht existiert: Für $x_n := \frac{1}{n\pi} \neq 0$ gilt nämlich $x_n \to 0$, $2x_n \sin(1/x_n) \to 0$ und $\cos(1/x_n) = \cos(n\pi) = (-1)^n$.

11.11. Höhere Ableitungen: Sei $I \subseteq \mathbb{R}$ ein Intervall und $f : I \to \mathbb{R}$ sei auf I differenzierbar.

Definition: (a) f heißt in $x_0 \in I$ *zweimal differenzierbar*, falls f' in x_0 differenzierbar ist. Dann heißt

$$f''(x_0) := (f')'(x_0)$$

[3] Man kann via 11.9 zeigen, dass $\lim_{s\to\infty} \frac{s^2}{e^s} = 0$ ist. Eine andere Möglichkeit ist die Abschätzung $e^s = \sum_{k=0}^{\infty} s^k/k! \ge s^3/3!$ für $s > 0$, aus der folgt $0 \le s^2 e^{-s} \le 3!/s \to 0$ für $s \to \infty$. Allgemeiner gilt für jedes $n \in \mathbb{N}$: $\lim_{s\to\infty} s^n e^{-s} = 0$.

die 2. *Ableitung von f in x_0*.

f heißt auf I *zweimal differenzierbar*, falls f' auf I differenzierbar ist. Dann heißt $f'' = (f')'$ *zweite Ableitung von f auf I*.

Entsprechend definiert man im Falle der Existenz $f'''(x_0)$, $f^{(4)}(x_0)$ etc. bzw. f''', $f^{(4)}$, ...

(b) Sei $n \in \mathbb{N}$. f heißt *auf I n-mal stetig differenzierbar*, falls f auf I n-mal differenzierbar ist und $f, f', f'', \ldots, f^{(n)} : I \to \mathbb{R}$ stetig sind. Dafür schreiben wir $f \in C^n(I)$.

Außerdem: $f \in C^0(I) = C(I)$, falls f auf I stetig ist, und $f \in C^\infty(I)$, falls $f \in C^n(I)$ für jedes $n \in \mathbb{N}$ gilt, dh wenn f auf I beliebig oft differenzierbar ist.

Beispiele: (1) Für $f : \mathbb{R} \to \mathbb{R}$, $f(x) = e^x$, gilt $f \in C^\infty(\mathbb{R})$. Hier ist $f^{(n)}(x) = e^x$ für jedes $n \in \mathbb{N}$, $x \in \mathbb{R}$.

(2) Es gilt $\sin, \cos \in C^\infty(\mathbb{R})$. Hier ist $\sin' = \cos$, $\sin'' = \cos' = -\sin$, $\sin''' = -\sin' = -\cos$ und $\sin^{(4)} = -\cos' = \sin$, etc. Genauso sind $\sinh, \cosh \in C^\infty(\mathbb{R})$.

(3) Die Funktion $f(x) = \frac{1}{1-x}$ ist auf $\mathbb{R}\backslash\{1\}$ beliebig oft differenzierbar mit $f^{(n)}(x) = \frac{n!}{(1-x)^{n+1}}$ für jedes $n \in \mathbb{N}_0$ (Induktion!).

(3a) Jede Polynomfunktion ist beliebig oft differenzierbar.

(4) Für die Funktion f aus dem Beispiel in 11.10 gilt $f \in C^\infty(\mathbb{R})$: Es ist klar, dass f auf $(-\infty, 0)$ und auf $(0, \infty)$ beliebig oft differenzierbar ist. Für jedes $n \in \mathbb{N}$ gilt:

$$f^{(n)}(x) = \begin{cases} p_n(\frac{1}{x})e^{-1/x} & , x > 0 \\ 0 & , x < 0 \end{cases},$$

wobei p_n ein Polynom ist (für $n = 1$ ist $p_1(s) = s^2$ und also $p_1(\frac{1}{x}) = \frac{1}{x^2}$, vgl. Beispiel in 11.10 oben; die allgemeine Aussage zeigt man durch Induktion nach n: Ist p_n gefunden mit $f^{(n)}(x) = p_n(1/x)e^{-1/x}$ für $x > 0$, so haben wir

$$f^{(n+1)}(x) = p_n'(1/x) \cdot (-x^{-2})e^{-1/x} + p_n(1/x)e^{-1/x} \cdot x^{-2},\ x > 0,$$

also $p_{n+1}(s) = -s^2 p_n'(s) + s^2 p_n(s)$). Wir wenden den Satz 11.10 jetzt wiederholt an (dh sukzessive auf f, f', f'', etc.), wobei wir beachten, dass für jedes $n \in \mathbb{N}$ gilt:

$$\lim_{x \to 0+} f^{(n)}(x) = \lim_{x \to 0+} p_n\left(\frac{1}{x}\right)e^{-1/x} = \lim_{s \to \infty} p_n(s)e^{-s} = 0.$$

Wir erhalten so, dass f, f', f'', etc. auf $\mathbb{R}$ stetig differenzierbar sind.

11.12. Satz von Taylor: Sei $n \in \mathbb{N}_0$, $f \in C^n(I)$ und $f^{(n)}$ sei auf I differenzierbar. Seien $x, x_0 \in I$. Dann gibt es ein ξ zwischen x und x_0 mit

$$\begin{aligned} f(x) &= f(x_0) + \frac{f'(x_0)}{1!}(x - x_0) + \ldots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n + \frac{f^{(n+1)}(\xi)}{(n+1)!}(x - x_0)^{n+1} \\ &= \underbrace{\sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k}_{n\text{-tes Taylorpolynom}} + \underbrace{\frac{f^{(n+1)}(\xi)}{(n+1)!}(x - x_0)^{n+1}}_{\text{Restglied}}. \end{aligned}$$

Bemerkung: Für $n = 0$ ist das der Mittelwertsatz, den wir auch im Beweis verwenden.

Definition: Für $n \in \mathbb{N}_0$ schreiben wir

$$T_n(f; x_0)(x) := \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k, \quad x \in I,$$

für das n-te Taylorpolynom von f bei Entwicklung um x_0.

Beweis des Satzes. Es reicht $x_0 = 0$ und $h := x - x_0 > 0$. Wir definieren $g : [0, h] \to \mathbb{R}$ durch

$$g(x) := f(h) - \sum_{k=0}^{n} \frac{f^{(k)}(x)}{k!}(h - x)^k - m\frac{(h - x)^{n+1}}{(n + 1)!},$$

wobei wir $m \in \mathbb{R}$ so wählen, dass $g(0) = 0$ ist. Dann haben wir $g(h) = 0$ und (wegen $g(0) = 0$)

$$f(h) = \sum_{k=0}^{n} \frac{f^{(k)}(0)}{k!}h^k - m\frac{h^{n+1}}{(n + 1)!}.$$

Nach Voraussetzung ist g in $[0, h]$ differenzierbar , so dass wir nach MWS ein $\xi \in [0, h]$ finden mit $g'(\xi) = 0$. Wegen (nachrechnen!)

$$g'(x) = (f^{(n+1)}(x) - m)\frac{(h - x)^n}{n!}, \quad x \in [0, h]$$

folgt dann $m = f^{(n+1)}(\xi)$ und wir sind fertig. □

Bemerkung: Falls $f \in C^\infty(I)$, so ist f um x_0 in eine auf I konvergente Potenzreihe entwickelbar, falls für jedes $x \in I$ gilt:

$$T_n(f; x_0)(x) \to f(x) \quad (n \to \infty).$$

Gleichbedeutend damit ist, dass für jedes $x \in I$ das entsprechende Restglied für $n \to \infty$ gegen Null geht.

Beispiel: Wir betrachten $f(x) := \ln(1 - x)$ für $x < 1$ und $x_0 = 0$. Hier ist $f'(x) = -\frac{1}{1-x}$ und also nach 11.11(3) für jedes $k \in \mathbb{N}$: $f^{(k)}(x) = -\frac{(k-1)!}{(1-x)^k}$. Wir erhalten somit für jedes $n \in \mathbb{N}$:

$$T_n(f; 0)(x) = -\sum_{k=1}^{n} \frac{x^k}{k}, \quad x < 1.$$

Das Restglied ist hier

$$-\frac{1}{(1 - \xi)^{n+1}} \frac{x^{n+1}}{n + 1}.$$

Da ξ zwischen 0 und x liegt, erhalten wir $0 < \frac{x}{1-\xi} \leq \frac{x}{1-x} \leq 1$ für $x \in (0, \frac{1}{2}]$ und $0 \leq |\frac{x}{1-\xi}| \leq |x| \leq 1$ für $x \in [-1, 0]$. Somit geht das Restglied für $n \to \infty$ zumindest für $x \in [-1, \frac{1}{2}]$ gegen Null, und

$$\ln(1-x) = -\sum_{k=1}^{\infty} \frac{x^k}{k}, \quad x \in [-1, \frac{1}{2}).$$

Tatsächlich gilt dies auch für $x \in (1/2, 1)$ (vgl. 11.14 unten, dort wird Konvergenz gezeigt für $x \in (-1, 1)$; für $x = -1$ kann man auch Ergänzung 11.16 unten verwenden). Wir notieren als Spezialfall ($x = -1$)

$$\sum_{k=1}^{\infty} \frac{(-1)^k}{k} = -\ln 2.$$

Warnung: Nicht jede Funktion $f \in C^\infty(I)$ ist durch ihre Taylorreihe darstellbar!

Beispiel: Wir betrachten $f : \mathbb{R} \to \mathbb{R}$, $f(x) = \begin{cases} e^{-1/x} & , x > 0 \\ 0 & , x < 0 \end{cases}$, in $x_0 = 0$. Es gilt $f^{(n)}(0) = 0$ für jedes $n \in \mathbb{N}_0$, also auch $T_n(f; 0)(x) = 0$ für alle $n \in \mathbb{N}$, $x \in \mathbb{R}$. Für **kein** $x > 0$ gilt somit $T_n(f; 0)(x) \to f(x)$!

11.13. Lokale Extrema: Sei $n \geq 2$, $I \subseteq \mathbb{R}$ ein Intervall, $f \in C^n(I)$ und $x_0 \in I$ ein innerer Punkt von I (dh kein Randpunkt). Weiter sei

$$f'(x_0) = f''(x_0) = \ldots = f^{(n-1)}(x_0) = 0 \quad \textbf{und} \quad f^{(n)}(x_0) \neq 0.$$

(a) Ist n gerade und $f^{(n)}(x_0) > 0$ [bzw. $f^{(n)}(x_0) < 0$], so hat f in x_0 ein lokales Minimum [bzw. ein lokales Maximum].

(b) Ist n ungerade, so hat f in x_0 kein lokales Extremum.

Bemerkung: In der Anwendung ist meist $n = 2$. Ist $f'(x_0) = 0$, so gilt:

für $f''(x_0) > 0$ hat f in ein lokales Minimum,

für $f''(x_0) < 0$ hat f in ein lokales Maximum,

für $f''(x_0) = 0$ erhält man keine Entscheidung.

Beweis. $f^{(n)}$ ist stetig auf I mit $f^{(n)}(x_0) \neq 0$, somit gibt es $\delta > 0$ mit

$$f^{(n)}(\xi) f^{(n)}(x_0) > 0 \quad \text{für alle } \xi \in U_\delta(x_0) \subseteq I.$$

Nach dem Satz von Taylor und der Voraussetzung gibt es für jedes $x \in U_\delta(x_0)$ ein $\xi \in U_\delta(x_0)$ mit

$$f(x) - f(x_0) = \frac{f^{(n)}(\xi)}{n!}(x - x_0)^n.$$

Nun betrachte man das Vorzeichen der rechten Seite. $\square$

Beispiel: Sei $p > 1$, $\alpha > 0$. Bestimme das Maximum von $f(x) := \alpha x - x^p/p$ über $x \geq 0$.

Es gilt $f(0) = 0$ und $f(x) \to -\infty$ für $x \to \infty$. Für $x > 0$ ist $f'(x) = \alpha - x^{p-1}$ und $f''(x) = -(p-1)x^{p-2} < 0$. Weiter ist $f'(x_0) = 0$ genau dann, wenn $x_0 = \alpha^{1/(p-1)}$ ist. Der Funktionswert an dieser Stelle ist $f(\alpha^{1/(p-1)}) = (1 - \frac{1}{p})\alpha^{p/(p-1)} > 0$. Dies ist das gesuchte Maximum.

Bemerkung: Allgemeiner sei $f \in C^2(I)$ mit $f'' \geq 0$ auf I [bzw. mit $f'' \leq 0$ auf I]. Solche Funktionen heißen *konvex* [bzw. *konkav*]. Es gilt dann:

> Ist $x_0 \in I$ mit $f'(x_0) = 0$, so hat f in x_0 ein globales Minimum [bzw. ein globales Maximum].

Es ist dann nämlich f' auf I nach 11.8 monoton wachsend, also $f' \leq 0$ links von x_0 und $f' \geq 0$ rechts von x_0, dh (wieder nach 11.8) also f monoton fallend links von x_0 und monoton wachsend rechts von x_0. Wir erhalten so $f(x) \geq f(x_0)$ für alle $x \in I$.

Außerdem ist $f \in C^2(I)$ konvex genau dann, wenn für alle $x, y \in I$ und $\lambda \in (0,1)$ gilt:

$$f((1-\lambda)x + \lambda y) \leq (1-\lambda)f(x) + \lambda f(y),$$

dh wenn die Funktion auf jedem Teilintervall $[x, y]$ unterhalb der Geraden durch $f(x)$ und $f(y)$ liegt.

Beispiel: $f : \mathbb{R} \to \mathbb{R}$, $f(x) = e^x - x$ ist in $C^2(\mathbb{R})$ mit $f'(x) = e^x - 1$, $f''(x) = e^x > 0$. Die Funktion f ist konvex und hat bei der einzigen Nullstelle $x_0 = 0$ von f' ein globales Minimum.

11.14. Ableitung von Potenzreihen: Sei $\sum_{n=0}^{\infty} a_n(x - x_0)^n$ eine reelle Potenzreihe mit Konvergenradius $R \in (0, \infty]$, $I := (x_0 - R, x_0 + R)$ und

$$f : I \to \mathbb{R}, \quad x \mapsto f(x) := \sum_{n=0}^{\infty} a_n(x - x_0)^n.$$

Dann ist f auf I differenzierbar, und für jedes $x \in I$ gilt

$$f'(x) = \sum_{n=1}^{\infty} n a_n (x - x_0)^{n-1}.$$

Beweis. Wir können $x_0 = 0$ annehmen. Die Potenzreihen $\sum_{n=1}^{\infty} n a_n x^{n-1}$ und $\sum_{n=2}^{\infty} n(n-1)|a_n|x^{n-2}$ haben Konvergenzradius R, und wir bezeichnen die auf I dargestellten Funktionen mit g bzw. φ. Sei nun $x \in I$ und $h \neq 0$ mit $|x| + |h| < R$. Dann gilt

$$\frac{f(x+h) - f(x)}{h} - g(x) = \sum_{n=2}^{\infty} a_n \Big(\frac{(x+h)^n - x^n}{h} - n x^{n-1} \Big).$$

Wir müssen zeigen, dass dies für $h \to 0$ gegen Null geht.

Ist $n \geq 2$, so finden wir nach MWS zunächst ein ξ zwischen x und $x+h$ und dann ein η zwischen ξ und x so, dass

$$\frac{(x+h)^n - x^n}{h} - nx^{n-1} = n(\xi^{n-1} - x^{n-1}) = n(n-1)(\xi - x)\eta^{n-2}.$$

Folglich gilt für jedes $n \geq 2$:

$$\Big|\frac{(x+h)^n - x^n}{h} - nx^{n-1}\Big| \leq n(n-1)|h|(|x|+|h|)^{n-2}.$$

Wir erhalten somit

$$\Big|\frac{f(x+h)-f(x)}{h} - g(x)\Big| \leq |h| \sum_{n=2}^{\infty} n(n-1)|a_n|(|x|+|h|)^{n-2} = |h|\varphi(|x|+|h|).$$

Für $h \to 0$ gilt nun $\varphi(|x|+|h|) \to \varphi(|x|)$, und es folgt $\lim_{h\to 0} \frac{f(x+h)-f(x)}{h} = g(x)$. $\square$

Bemerkung: Der Satz besagt, dass man Potenzreihen im Inneren des Konvergenzintervalls gliedweise differenzieren kann. Da f' wieder eine Potenzreihe mit Konvergenzradius R ist, kann man den Satz wiederholt anwenden und erhält damit: $f \in C^\infty(I)$ und für alle $k \in \mathbb{N}$ gilt:

$$f^{(k)}(x) = \sum_{n=k}^{\infty} n(n-1)(n-2)\cdot \ldots \cdot (n-(k-1))a_n(x-x_0)^{n-k}, \quad x \in I.$$

Für $x = x_0$ erhält man

$$f^{(k)}(x_0) = k(k-1)(k-2)\cdot \ldots \cdot (k-(k-1))a_k = k!a_k.$$

Somit gilt

$$a_k = \frac{f^{(k)}(x_0)}{k!} \quad \text{für alle } k \in \mathbb{N}_0,$$

und wir erhalten die Darstellung

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n, \quad x \in I,$$

so dass die Potenzreihe die *Taylorreihe* der dargestellten Funktion ist.

Beispiel: Für $x \in (-1,1)$ gilt

$$\arctan'(x) = \frac{1}{1+x^2} = \frac{1}{1-(-x^2)} = \sum_{k=0}^{\infty}(-x^2)^k = \sum_{k=0}^{\infty}(-1)^k x^{2k}.$$

Der Konvergenzradius der Potenzreihe ist 1. Der Konvergenzradius der Potenzreihe $\sum_{k=0}^{\infty}(-1)^k \frac{x^{2k+1}}{2k+1}$ ist ebenfalls 1. Setzt man

$$g(x) = \sum_{k=0}^{\infty}(-1)^k \frac{x^{2k+1}}{2k+1}, \quad x \in (-1,1),$$

so ist g nach 11.14 auf $(-1,1)$ differenzierbar, und es gilt

$$g'(x) = \sum_{k=0}^{\infty}(-1)^k x^{2k} = \arctan'(x), \quad x \in (-1,1).$$

Wegen $\arctan 0 = 0 = g(0)$ ist also

$$\arctan(x) = g(x) = \sum_{k=0}^{\infty}(-1)^k \frac{x^{2k+1}}{2k+1}, \quad x \in (-1,1).$$

Eine ähnliche Argumentation zeigt

$$\ln(1-x) = -\sum_{n=1}^{\infty} \frac{x^n}{n}, \quad x \in (-1,1).$$

11.15. Identitätssatz für Potenzreihen: Sei $r > 0$ und $I := (x_0 - r, x_0 + r)$. Seien $f, g : I \to \mathbb{R}$ gegeben durch

$$f(x) := \sum_{n=0}^{\infty} a_n (x-x_0)^n, \quad g(x) := \sum_{n=0}^{\infty} b_n (x-x_0)^n, \quad x \in I,$$

wobei die Potenzreihen auf I konvergieren. Es gebe eine streng monotone Folge $(x_m)_{m\in\mathbb{N}}$ in I mit $x_m \to x_0$ für $m \to \infty$ und $f(x_m) = g(x_m)$ für jedes $m \in \mathbb{N}$. Dann gilt $a_n = b_n$ für alle $n \in \mathbb{N}_0$.

Bemerkung: Insbesondere gilt also:

$$f = g \text{ auf } I \quad \Rightarrow \quad \forall n \in \mathbb{N}_0 : a_n = b_n,$$

aber die Voraussetzung lässt sich sehr abschwächen.

Beweis. Wir dürfen $g = 0$ und $b_n = 0$ annehmen (durch Betrachtung von $f - g$ und $a_n - b_n$). **Annahme:** es gibt $n_0 = \min\{n \in \mathbb{N}_0 : a_n \neq 0\}$. Dann gilt $f(x) = \sum_{n=n_0}^{\infty} a_n (x-x_0)^n$ und

$$\frac{f(x)}{(x-x_0)^{n_0}} = \sum_{n=n_0}^{\infty} a_n (x-x_0)^{n-n_0} \to a_{n_0} \quad (x \to x_0).$$

Also ist nach Voraussetzung $a_{n_0} = \lim_{m\to\infty} \frac{f(x_m)}{(x_m-x_0)^{n_0}} = 0$ im Widerspruch zur Annahme. $\square$

Beispiel: Wir wollen ein Intervall I mit $0 \in I$ und eine differenzierbare Funktion y bestimmen mit $y'(x) = xy(x)$, $x \in I$, und $y(0) = 1$ und machen einen *Potenzreihenansatz* $y(x) = \sum_{n=0}^{\infty} a_n x^n$ für $x \in I = (-R, R)$, wobei R der Konvergenzradius der Potenzreihe sei. Wegen $y(0) = 1$ ist $a_0 = 1$.

Wenn $R > 0$, so ist nach 11.14: $y'(x) = \sum_{n=1}^{\infty} n a_n x^{n-1}$ für $x \in I$. Außerdem ist

$$xy(x) = \sum_{n=0}^{\infty} a_n x^{n+1} = \sum_{n=2}^{\infty} a_{n-2} x^{n-1}, \quad x \in I.$$

Durch *Koeffizientenvergleich* (dh nach 11.15) erhalten wir $a_1 = 0$ und

$$n a_n = a_{n-2}, \quad n \geq 2.$$

Durch Induktion erhält man hieraus $a_3 = a_5 = a_7 = \ldots = 0$ und $a_{2k} = (2^k k!)^{-1}$ für $k = 0, 1, 2, \ldots$. Somit

$$y(x) = \sum_{k=0}^{\infty} \frac{x^{2k}}{2^k k!} = \sum_{k=0}^{\infty} \frac{1}{k!} \left(\frac{x^2}{2} \right)^k = e^{x^2/2}, \quad x \in \mathbb{R},$$

und dies genügt tatsächlich den geforderten Bedingungen.

11.16. Ergänzung: Abelscher Grenzwertsatz: Wir stellen z.B. fest, dass die Arcustangens-Reihe auch noch für $x = \pm 1$ konvergiert (Leibnizkriterium), und würden gerne $x = 1$ einsetzen. Das ermöglicht der folgende Satz:

> Sei $\sum_{n=0}^{\infty} b_n$ konvergent. Dann gilt
>
> $$\lim_{x \to 1-} \sum_{n=0}^{\infty} b_n x^n = \sum_{n=0}^{\infty} b_n.$$

Beachte, dass der Konvergenzradius der Potenzreihe wegen der vorausgesetzten Konvergenz von $\sum_{n=0}^{\infty} b_n$ nicht < 1 sein kann und also ≥ 1 ist.

Für $x = 1$ erhält man also in der Arcustangens-Reihe:

$$\frac{\pi}{4} = \arctan 1 = \lim_{x \to 1-} \arctan(x) = \lim_{x \to 1-} \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k+1}}{2k+1} = \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \ldots.$$

Wenden wir den Satz auf die Logarithmus-Reihe

$$\ln(1-x) = -\sum_{n=1}^{\infty} \frac{x^n}{n}, \quad |x| < 1,$$

an, erhalten wir

$$\sum_{n=1}^{\infty} \frac{(-1)^n}{n} = -\ln 2.$$

Beweis des Abelschen Grenzwertsatzes. Wir fixieren ein $N \in \mathbb{N}$ mit $N \geq 2$. Für $x \in (0,1)$ gilt dann

$$d_N(x) := \sum_{n=0}^{N} b_n - \sum_{n=0}^{N} b_n x^n = \sum_{n=1}^{N} b_n(1-x^n),$$

wobei für $n \geq 1$ nach 4.11(1):

$$1 - x^n = (1-x)\sum_{j=0}^{n-1} x^j.$$

Also haben wir für $x \in (0,1)$ und $N_0 \in \mathbb{N}$ mit $N_0 \leq N$:

$$\begin{aligned} d_N(x) &= (1-x)\sum_{n=1}^{N}\sum_{j=0}^{n-1} b_n x^j = (1-x)\sum_{j=0}^{N-1} x^j \sum_{n=j+1}^{N} b_n \\ &= (1-x)\sum_{j=0}^{N_0-1} x^j \sum_{n=j+1}^{N} b_n + (1-x)\sum_{j=N_0}^{N-1} x^j \sum_{n=j+1}^{N} b_n. \end{aligned}$$

Setzen wir $c_m := \sum_{n=0}^{m} b_n$, so haben wir $c_m \to c := \sum_{n=0}^{\infty} b_n$ für $m \to \infty$. Die Folge $(c_m)_{m\in\mathbb{N}_0}$ ist also beschränkt, etwa $|c_m| \leq M$ für alle $m \in \mathbb{N}_0$. Wir haben nun

$$\Big|\sum_{j=0}^{N_0-1} x^j \sum_{n=j+1}^{N} b_n\Big| \leq \sum_{j=0}^{N_0-1} x^j |c_N - c_j| \leq N_0 \cdot 2M$$

und

$$\Big|(1-x)\sum_{j=N_0}^{N-1} x^j \sum_{n=j+1}^{N} b_n\Big| \leq (1-x)\sum_{j=N_0}^{N-1} x^j |c_N - c_j| \leq \underbrace{(1-x)\sum_{k=0}^{\infty} x^k}_{=1} \max_{j=N_0,\ldots,N} |c_N - c_j|,$$

also

$$|d_N(x)| \leq (1-x)\cdot 2MN_0 + \max_{N_0\leq j\leq N} |c_N - c_j|.$$

Sei $\varepsilon > 0$. Da (c_m) eine Cauchyfolge ist, finden wir ein N_0 so, dass für alle $N \geq j \geq N_0$ gilt: $|c_N - c_j| < \varepsilon/2$. Wir haben dann für $x \in (0,1)$:

$$\underbrace{\Big|\sum_{n=0}^{\infty} b_n - \sum_{n=0}^{\infty} b_n x^n\Big|}_{=:r(x)} = \lim_{N\to\infty} |d_N(x)| \leq (1-x)2MN_0 + \varepsilon/2.$$

Für $x \in (0,1)$ mit $1 - x < \varepsilon/(4MN_0)$ ist dann $|r(x)| < \varepsilon$. Damit ist $\lim_{x\to 1-} r(x) = 0$ gezeigt. $\square$

11.17. Ergänzung: Das Newton-Verfahren: In 11.1 bei der Definition der Differenzierbarkeit haben wir gesagt, dass sich die lineare Approximation einer Funktion oft leichter behandeln lässt als die Funktion selber. Das folgende ist dafür ein Beispiel.

Gegeben sei eine differenzierbare Funktion $f : I \to \mathbb{R}$, für die man eine *Nullstelle* berechnen möchte, dh also ein $x^* \in I$ mit $f(x^*) = 0$ (wir nehmen an, dass f in I *genau eine* Nullstelle hat).

Zu einem gewählten *Startwert* $x_0 \in I$ betrachtet man statt f die linearisierte Funktion $x \mapsto f(x_0) + f'(x_0)(x - x_0)$ und bestimmt für diese eine Nullstelle x_1. Dies ist möglich, wenn $f'(x_0) \neq 0$ ist; dann ist $x_1 = x_0 - f'(x_0)^{-1} f(x_0)$.

Man hofft, dass x_1 eine Näherung für x^* ist, und bestimmt rekursiv

$$x_{k+1} = x_k - f'(x_k)^{-1} f(x_k), \quad k \in \mathbb{N}_0,$$

(unter der Voraussetzung, dass immer $f'(x_k) \neq 0$ ist), in der Hoffnung, dass die Folge (x_k) gegen x^* konvergiert.

Bemerkung: (1) Das Newton-Verfahren konvergiert unter den Voraussetzungen $f \in C^1(I)$ und $f'(x^*) \neq 0$ im allgemeinen nur, wenn x_0 schon genügend nahe bei x^* liegt, dann jedoch "schnell".

(2) Wenn $f \in C^1(I)$ ist mit $f'(x^*) \neq 0$ und (x_k) konvergiert, so ist $\lim x_k$ eine Nullstelle von f.

Satz: Sei $f : I \to \mathbb{R}$ differenzierbar, $x^* \in I$ mit $f(x^*) = 0$ und $f'(x^*) > 0$. Ist $x_0 > x^*$ und f' auf $[x^*, x_0]$ monoton wachsend, so konvergiert das Newton-Verfahren gegen x^*.

Beweis. Wie man leicht sieht, gilt $x_k \geq x^*$ für jedes k. Für jedes $k \in \mathbb{N}_0$ gilt:

$$x_{k+1} - x^* = x_k - x^* - f'(x_k)^{-1} f(x_k)$$

und nach dem MWS

$$f(x_k) = f(x_k) - f(x^*) = f'(\xi)(x_k - x^*)$$

für ein $\xi \in (x^*, x_k)$, also

$$x_{k+1} - x^* = (1 - f'(\xi)/f'(x_k))(x_k - x^*) \leq \underbrace{(1 - f'(x^*)/f'(x_0))}_{=:\alpha \in [0,1)}(x_k - x^*).$$

Wir erhalten $x_k - x^* \leq \alpha^k (x_0 - x^*) \to 0$ für $k \to \infty$. $\square$

Beispiel: $m \in \mathbb{N}$, $a > 0$ und $f(x) = x^m - a$. Für jeden Startwert $x_0 > \sqrt[m]{a}$ konvergiert das Newton-Verfahren gegen $\sqrt[m]{a}$.

12 Integration

12.1. Ober- und Untersummen, oberes und unteres Integral: Seien $a, b \in \mathbb{R}$ mit $a < b$ und sei $f : [a, b] \to \mathbb{R}$ *beschränkt*, dh $f([a, b]) = \{f(x) : x \in [a, b]\}$ beschränkt.

Definition: $Z := \{x_0, x_1, \dots, x_n\}$ heißt eine *Zerlegung* von $[a, b]$, falls $a = x_0 < x_1 < \dots < x_n = b$.

Sei $m := \inf f([a, b])$ und $M := \sup f([a, b])$. Sei $Z = \{x_0, x_1, \dots, x_n\}$ ein Zerlegung von $[a, b]$. Für $j = 1, 2, \dots, n$ setze

$$\begin{aligned} I_j &:= [x_{j-1}, x_j], \quad |I_j| := x_j - x_{j-1}, \quad m_j := \inf f(I_j), \quad M_j := \sup f(I_j), \text{ sowie} \\ s_f(Z) &:= \sum_{j=1}^{n} m_j |I_j| \qquad \textit{Untersumme} \text{ von } f \text{ bzgl. } Z, \\ S_f(Z) &:= \sum_{j=1}^{n} M_j |I_j| \qquad \textit{Obersumme} \text{ von } f \text{ bzgl. } Z. \end{aligned}$$

Es ist $m \leq m_j \leq M_j \leq M$, also wegen $|I_j| > 0$:

$$\sum_{j=1}^{n} m|I_j| \leq \sum_{j=1}^{n} m_j|I_j| \leq \sum_{j=1}^{n} M_j|I_j| \leq \sum_{j=1}^{n} M|I_j|.$$

Somit gilt für jede Zerlegung Z von $[a, b]$:

$$(*) \qquad m(b-a) \leq s_f(Z) \leq S_f(Z) \leq M(b-a).$$

Satz: Seien Z_1, Z_2 Zerlegungen.

(1) Ist $Z_1 \subseteq Z_2$, so gilt $s_f(Z_1) \leq s_f(Z_2) \leq S_f(Z_2) \leq S_f(Z_1)$.

(2) Es gilt $s_f(Z_1) \leq S_f(Z_2)$.

Beweis. (1) Wende $(*)$ an auf diejenigen Teilintervalle von Z_1, die durch Punkte von Z_2 weiter unterteilt werden.

(2) Setze $Z := Z_1 \cup Z_2$. Wegen $Z_1 \subseteq Z$ und $Z_2 \subseteq Z$ gilt dann nach (1) und $(*)$:

$$s_f(Z_1) \leq s_f(Z) \leq S_f(Z) \leq S_f(Z_2).$$

□

Nach $(*)$ können wir definieren:

$$s_f := \sup\{s_f(Z) : Z \text{ ist Zerlegung von } [a, b]\},$$

das *untere Integral von f* über $[a, b]$ und

$$S_f := \inf\{S_f(Z) : Z \text{ ist Zerlegung von } [a, b]\}$$

das *obere Integral von f* über $[a, b]$.

Wegen $(*)$ und (2) gilt dann

$$(**) \qquad m(b-a) \leq s_f \leq S_f \leq M(b-a).$$

[Zunächst folgt aus (2) durch Supremumsbildung über Z_1: $s_f \leq S_f(Z_2)$ für jede Zerlegung Z_2. Dann bilde man das Infimum über alle Z_2.]

12.2. Definition: Seien $a, b \in \mathbb{R}$ mit $a < b$ und sei $f : [a, b] \to \mathbb{R}$ *beschränkt*. Dann heißt f (Riemann-)*integrierbar* (*ib*), falls $s_f = S_f$ gilt.

In diesen Falle heißt

$$\int_a^b f\,dx := \int_a^b f(x)\,dx := S_f (= s_f)$$

das (Riemann-)*Integral* von f über $[a, b]$.

Wir setzen

$$R[a, b] := \{f : [a, b] \to \mathbb{R} : \ f \text{ ist beschränkt und über } [a, b] \text{ integrierbar}\}$$

(Menge der über $[a, b]$ integrierbaren Funktionen).

Beispiele: (1) Sei $c \in \mathbb{R}$ und $f(x) = c$, $x \in [a, b]$. Dann ist $m = M = c$ und aus $(**)$ folgt $s_f = S_f = c(b-a)$. Also ist f integrierbar und

$$\int_a^b c\,dx = c(b-a).$$

(2) Sei $a = 0$, $b = 1$ und $f(x) := \begin{cases} 1 & , x \in [0,1] \cap \mathbb{Q} \\ 0 & , x \in [0,1] \setminus \mathbb{Q} \end{cases}$. Ist $Z = \{x_0, x_1, \ldots, x_n\}$ Zerlegung von $[0, 1]$ und sind m_j, M_j, I_j wie oben, so haben wir $m_j = 0$, $M_j = 1$ für alle j [4], also $s_f(Z) = 0$ und $S_f(Z) = 1$. Folglich ist

$$s_f = 0 \neq 1 = S_f,$$

und f ist nicht integrierbar über $[0, 1]$.

[4]Hier wird benutzt, dass jedes Intervall (mit mindestes zwei Punkten) rationale und irrationale Zahlen enthält.

12.3. Satz: Seien $f, g \in R[a,b]$.

(1) Gilt $f \leq g$ auf $[a,b]$, so ist $\int_a^b f\,dx \leq \int_a^b g\,dx$ (Monotonie des Integrals).

(2) Linearität des Integrals: Für $\alpha, \beta \in \mathbb{R}$ ist $\alpha f + \beta g \in R[a,b]$ und

$$\int_a^b (\alpha f + \beta g)\,dx = \alpha \int_a^b f\,dx + \beta \int_a^b g\,dx.$$

Insbesondere ist $R[a,b]$ ein $\mathbb{R}$-Vektorraum und die Abbildung $R[a,b] \to \mathbb{R}$, $f \mapsto \int_a^b f\,dx$, ist linear.

(3) Setzt man $\gamma := \sup\{|f(x)| : x \in [a,b]\}$, so gilt

$$\Big| \int_a^b f(x)\,dx \Big| \leq \gamma(b-a).$$

Beweis. (1) ist klar. (2) ist klar für $g = 0$ und $\alpha \geq 0$. Für $g = 0$ und $\alpha = -1$ folgt die Aussage aus $S_{-f}(Z) = -s_f(Z)$, $s_{-f}(Z) = -S_f(Z)$. Für $\alpha = \beta = 1$ haben wir

$$s_f(Z) + s_g(Z) \leq s_{f+g}(Z) \leq S_{f+g}(Z) \leq S_f(Z) + S_g(Z)$$

für jede Zerlegung, woraus

$$s_f + s_g \leq s_{f+g} \leq S_{f+g} \leq S_f + S_g$$

folgt. Hieraus folgt die Aussage im Fall $\alpha = \beta = 1$.

(3) Sind M und m wie in 12.1, so folgt aus $(**)$:

$$\Big| \int_a^b f(x)\,dx \Big| \leq \max\{|M|, |m|\}(b-a).$$

Beachte nun $\gamma = \max\{|M|, |m|\}$. □

12.4. Satz: Sei $f : [a,b] \to \mathbb{R}$ eine beschränkte Funktion. Es gilt $f \in R[a,b]$ genau dann, wenn es zu jedem $\varepsilon > 0$ eine Zerlegung Z gibt mit $S_f(Z) - s_f(Z) < \varepsilon$ bzw. genau dann, wenn es eine Folge $(Z_l)_{l \in \mathbb{N}}$ von Zerlegungen gibt mit $\lim_l S_f(Z_l) - s_f(Z_l) = 0$. In diesem Fall ist

$$\int_a^b f(x)\,dx = \lim_{l \to \infty} S_f(Z_l) = \lim_{l \to \infty} s_f(Z_l).$$

Beweis. Das folgt aus der Definition mithilfe von 12.1 Satz (1). □

Wir betrachten insbesondere $Z_n := \{x_0, x_1, \ldots, x_n\}$ für $n \in \mathbb{N}$, wobei $x_j := a + j(b-a)/n$ für $j = 0, 1, \ldots, n$. Sind I_j, M_j und m_j wie in 12.1, so gilt $|I_j| = (b-a)/n$ für jedes j und

$$(***) \quad S_f(Z_n) - s_f(Z_n) = \frac{b-a}{n} \sum_{j=1}^{n} (M_j - m_j) \le (b-a) \max_{j=1,\ldots,n} (M_j - m_j).$$

Das werden wir in 12.5 und 12.6 ausnutzen.

12.5. Satz über stetige Funktionen: Ist $f \in C[a,b]$, so ist f integrierbar, dh $f \in R[a,b]$.

Beweis. Wir verwenden 12.4 und $(***)$. Es reicht also zu zeigen $\max_{j=1,\ldots,n}(M_j - m_j) \to 0$ für $n \to \infty$. Nach 9.14 finden wir jeweils $x_j, \tilde{x}_j \in I_j$ mit $M_j - m_j = f(x_j) - f(\tilde{x}_j)$. Wegen $|x_j - \tilde{x}_j| \le |I_j| = \frac{b-a}{n}$ reicht es somit zu zeigen, dass f *gleichmäßig stetig* ist, dh zu jedem $\varepsilon > 0$ gibt es ein $\delta > 0$ so, dass

$$\forall x, \widetilde{x} \in [a,b] \text{ mit } |x - \widetilde{x}| < \delta : \quad |f(x) - f(\widetilde{x})| < \varepsilon.$$

Sei $\varepsilon > 0$. Ist die Aussage falsch, so finden wir zu jedem $n \in \mathbb{N}$ Punkte $x_n, \widetilde{x}_n \in [a,b]$ mit $|x_n - \widetilde{x}_n| < 1/n$ und $|f(x_n) - f(\widetilde{x}_n)| \ge \varepsilon$. Nach Bolzano-Weierstraß finden wir eine konvergente Teilfolge $(x_{k(n)})$ von (x_n) mit Limes $x_0 \in [a,b]$. Dann konvergiert $(\widetilde{x}_{k(n)})$ ebenfalls gegen x_0. Da f in x_0 stetig ist, folgt $f(x_{k(n)}) - f(\widetilde{x}_{k(n)}) \to 0$ im Widerspruch zu $|f(x_{k(n)}) - f(\widetilde{x}_{k(n)})| \ge \varepsilon$ für alle n. □

12.6. Satz: Ist $f : [a,b] \to \mathbb{R}$ monoton, so ist $f \in R[a,b]$.

Beweis. Sei f monoton wachsend (sonst betrachte $-f$ und verwende 12.3(2)). Wir verwenden wieder $(***)$. Hier ist $M_j = f(x_j)$ und $m_j = f(x_{j-1})$. Somit haben wir

$$S_f(Z_n) - s_f(Z_n) = \frac{b-a}{n} \sum_{j=1}^{n} (f(x_j) - f(x_{j-1})) = \frac{b-a}{n}(f(b) - f(a)) \to 0$$

für $n \to \infty$, und die Behauptung folgt aus 12.4. □

Beispiele: (1) Sei $b > 0$ und $f : [0,b] \to \mathbb{R}$, $f(x) = x$. Nach 12.6 ist $f \in R[0,b]$. Wir berechnen das Integral. Sei Z_n wie eben im Beweis, also $x_j = jb/n$ für $j = 0, 1, \ldots, n$. Dann gilt $M_j = f(x_j) = x_j = jb/n$ und $m_j = (j-1)b/n$, sowie

$$\begin{aligned} S_f(Z_n) &= \sum_{j=1}^{n} f(x_j)|I_j| = \sum_{j=1}^{n} j\frac{b^2}{n^2} = \frac{b^2}{n^2}\frac{n(n+1)}{2} \longrightarrow \frac{b^2}{2} \ (n \to \infty), \\ s_f(Z_n) &= \sum_{j=1}^{n} f(x_{j-1})|I_j| = \sum_{j=1}^{n} (j-1)\frac{b^2}{n^2} = \frac{b^2}{n^2}\frac{n(n-1)}{2} \longrightarrow \frac{b^2}{2} \ (n \to \infty), \end{aligned}$$

und mithilfe von 12.4 folgt $\int_0^b x\,dx = \frac{b^2}{2}$.

(2) Für $n \in \mathbb{N}$ und $b > 0$ ist $x \mapsto x^n$ über $[0, b]$ integrierbar. Für $0 < a < b$ und $n \in \mathbb{N}$ sind $x \mapsto x^{-n}$ und $x \mapsto \ln x$ über $[a, b]$ integrierbar.

12.7. Riemann-Summen: Sei $f : [a, b] \to \mathbb{R}$ beschränkt und $Z = \{x_0, x_1, \ldots, x_n\}$ eine Zerlegung, sowie I_j, M_j, m_j wie in 12.1. Dann heißt

$$\|Z\| := \max\{|I_j| : j = 1, 2, \ldots, n\}$$

die *Feinheit* von Z.

Ein n-Tupel $\xi = (\xi_1, \ldots, \xi_n)$ heißt *passender Zwischenvektor*, wenn $\xi_j \in I_j$ für jedes $j = 1, \ldots, n$ gilt. Für einen solchen heißt

$$\sigma_f(Z, \xi) := \sum_{j=1}^{n} f(\xi_j)|I_j|$$

eine *Riemannsche Summe*. Wegen $m_j \le f(\xi_j) \le M_j$ gilt dabei $s_f(Z) \le \sigma_f(Z, \xi) \le S_f(Z)$.

Satz: Sei $f \in R[a, b]$ und $(Z_l)_{l\in\mathbb{N}}$ eine Folge von Zerlegungen mit $\|Z_l\| \to 0$, sowie $(\xi^{(l)})$ eine Folge von passenden Zwischenvektoren. Dann gilt

$$\lim_{l\to\infty} \sigma_f(Z_l, \xi^{(l)}) = \int_a^b f\,dx.$$

(ohne Beweis)

Bemerkung: Haben *alle* Folgen Riemannscher Summen, deren Feinheit gegen Null geht, den Limes $\alpha \in \mathbb{R}$, so gilt $f \in R[a, b]$ und $\int_a^b f(x)\,dx = \alpha$.

Beweisidee. Zeige für eine Folge (Z_l) mit $\|Z_l\| \to 0$, dass gilt

$$S_f(Z_l) \to \alpha \ (l \to \infty), \quad s_f(Z_l) \to \alpha \ (l \to \infty),$$

und verwende 12.4. Beachte dazu, dass man $\sigma_f(Z, \xi) - s_f(Z)$ bzw. $S_f(Z) - \sigma_f(Z, \xi)$ durch Wahl geeigneter ξ beliebig klein machen kann. □

12.8. Satz: Sei I ein Intervall und $f : I \to \mathbb{R}$ eine Funktion.

(1) Seien $a, b, c \in I$ mit $a < c < b$. Dann gilt $f \in R[a, b]$ genau dann, wenn $f \in R[a, c]$ und $f \in R[c, b]$. In diesem Fall ist

$$\int_a^b f\,dx = \int_a^c f\,dx + \int_c^b f\,dx.$$

(2) Sei $f \in R[a,b]$ für alle $[a,b] \subseteq I$, und sei $c \in I$. Setzt man

$$F : I \to \mathbb{R}, y \mapsto F(y) := \int_c^y f(x)\,dx,$$

wobei $\int_c^y f(x)\,dx := -\int_y^c f(x)\,dx$ für $y < c$ und $\int_c^c f(x)\,dx := 0$ gesetzt ist, so gilt für alle $a, b \in I$:

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

Beweis. (2) folgt aus (1). Zum Beweis von (1) sei zunächst $f \in R[a,b]$ und $\varepsilon > 0$. Wir finden eine Zerlegung Z von $[a,b]$ mit $S_f(Z) - s_f(Z) < \varepsilon$. Wegen Satz 12.1(1) können wir davon ausgehen, dass $c \in Z$ gilt. Setzen wir dann $Z_1 := Z \cap [a,c]$ und $Z_2 := Z \cap [c,b]$, so gilt für $k = 1, 2$:

$$S_f(Z_k) - s_f(Z_k) \leq S_f(Z) - s_f(Z) < \varepsilon.$$

Nach 12.4 ist also $f \in R[a,c]$ und $f \in R[c,b]$.

Ist umgekehrt $f \in R[a,c]$ und $f \in R[c,b]$, sowie $\varepsilon > 0$, so finden wir Zerlegungen Z_1 von $[a,c]$ und Z_2 von $[c,b]$ mit $S_f(Z_k) - s_f(Z_k) < \varepsilon/2$ für $k = 1, 2$. Wir setzen $Z := Z_1 \cup Z_2$ und erhalten

$$S_f(Z) - s_f(Z) = S_f(Z_1) + S_f(Z_2) - s_f(Z_1) - s_f(Z_2) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Die Behauptung folgt also wieder mithilfe von 12.4. Schließlich haben wir

$$\begin{aligned}
\int_a^b f\,dx &= \sup\{s_f(Z) : Z \text{ Zerl. von } [a,b]\} = \sup\{s_f(Z) : Z \text{ Zerl. von } [a,b] \text{ mit } c \in Z\} \\
&= \sup\{s_f(Z_1) + s_f(Z_2) : Z_1 \text{ Zerl. von } [a,c],\ Z_2 \text{ Zerl. von } [c,b]\} \\
&= \sup\{s_f(Z_1) : Z_1 \text{ Zerl. von } [a,c]\} + \sup\{s_f(Z_2) : Z_2 \text{ Zerl. von } [c,b]\} \\
&= \int_a^c f\,dx + \int_c^b f\,dx,
\end{aligned}$$

wobei die erste Gleichheit die Definition ist, die zweite aus Satz 12.1(1) folgt, die dritte die Überlegungen im Beweis oben benutzt und die vierte die Gleichung $\sup(A + B) = \sup A + \sup B$ für $A, B \subseteq \mathbb{R}$ verwendet. □

Bemerkung: Ist $f \in R[a,b]$ und $g : [a,b] \to \mathbb{R}$ so, dass $g(x) = f(x)$ für alle bis auf endlich viele $x \in [a,b]$, dann gilt $g \in R[a,b]$ und $\int_a^b f(x)\,dx = \int_a^b g(x)\,dx$. Ändert man also eine integrierbare Funktion an endlich vielen Stellen, so bleiben Integrierbarkeit und Wert des Integrals erhalten. [Es reicht $f = 0$. Wegen 12.8 reicht es, $g(x) = 0$ für alle $x \in (a,b]$ bzw. alle $x \in [a,b)$ anzunehmen. Dieser Fall ist leicht.]

Folgerung: Sei $g : [a,b] \to \mathbb{R}$ so, dass es $n \in \mathbb{N}$ und eine Zerlegung $a = x_0 < x_1 < x_2 < \ldots < x_n = b$ mit

- $g : (x_{j-1}, x_j) \to \mathbb{R}$ ist stetig für $j = 1, 2, \ldots, n$,
- die einseitigen Grenzwerte $\lim_{x \to x_j -} g(x)$ und $\lim_{x \to x_j +} g(x)$ existieren und sind reell für $j = 1, \ldots, n$ bzw. $j = 0, \ldots, n-1$.

Dann ist $g \in R[a, b]$ und

$$\int_a^b g(x)\,dx = \sum_{j=1}^{n} \int_{x_{j-1}}^{x_j} g(x)\,dx.$$

Die Voraussetzungen sorgen nämlich dafür, dass man $g : [x_{j-1}, x_j] \to \mathbb{R}$ durch Ändern in den Randpunkten zu einer stetigen Funktion auf $[x_{j-1}, x_j]$ machen kann. Dann verwende 12.8.

Solche Funktionen $g : [a, b] \to \mathbb{R}$ heißen *stückweise stetig*. Sie spielen in der Regelungstechnik z.B. als "Steuerfunktionen" eine Rolle.

12.9. Satz: Sei $f \in R[a, b]$ und F wie in 12.8(2), sowie $x_0 \in [a, b]$. Dann gilt:

(1) F ist stetig in x_0, dh

$$\int_a^{x_0+h} f(x)\,dx \to \int_a^{x_0} f(x)\,dx \quad (h \to 0).$$

(2) Ist f stetig in x_0, so ist F differenzierbar in x_0 mit $F'(x_0) = f(x_0)$, dh es gilt

$$\frac{1}{h} \int_{x_0}^{x_0+h} f(x)\,dx \to f(x_0) \quad (h \to 0).$$

Hierbei betrachten wir für $x_0 = a$ nur $h \to 0+$ und für $x_0 = b$ nur $h \to 0-$.

Beweis. (1) Wir wählen $\gamma > 0$ mit $|f(x)| \le \gamma$ für alle $x \in [a, b]$. Für $h \neq 0$ mit $x_0, x_0 + h \in [a, b]$ haben wir nach 12.8 und 12.3(3):

$$\Big| \int_a^{x_0+h} f(x)\,dx - \int_a^{x_0} f(x)\,dx \Big| = \Big| \int_{x_0}^{x_0+h} f(x)\,dx \Big| \le \gamma |h| \to 0 \ (h \to 0).$$

(2) Sei $\varepsilon > 0$. Da f in x_0 stetig ist, finden wir $\delta > 0$ mit $|f(x) - f(x_0)| \le \varepsilon$ für $x \in [a, b]$ mit $|x - x_0| < \delta$. Für $|h| < \delta$ mit $x_0, x_0 + h \in [a, b]$ gilt wegen $h^{-1} \int_{x_0}^{x_0+h} 1\,dx = 1$ und 12.3(3):

$$\Big| \frac{1}{h} \int_{x_0}^{x_0+h} f(x)\,dx - f(x_0) \Big| = \Big| \frac{1}{h} \int_{x_0}^{x_0+h} (f(x) - f(x_0))\,dx \Big| \le \frac{1}{|h|} \cdot |h| \varepsilon \le \varepsilon.$$

Daraus folgt die Behauptung. □

Beispiel: Sei $n \in \mathbb{N}$ und $f(x) := x^n$ sowie $G(x) := \frac{x^{n+1}}{n+1}$ für $x \in \mathbb{R}$. Dann ist G auf $\mathbb{R}$ differenzierbar mit $G'(x) = f(x)$, $x \in \mathbb{R}$. Für $[a, b] \subseteq \mathbb{R}$ ist $f \in R[a, b]$. Definiert man F

wie in 12.8(2), so ist F auf $[a, b]$ differenzierbar mit $F' = f$ auf $[a, b]$. Nach 11.8 ist $G - F$ konstant, also

$$\int_a^b x^n\,dx = F(b) - F(a) = G(b) - G(a) = \frac{b^{n+1} - a^{n+1}}{n+1}.$$

12.10. Hauptsatz der Differential- und Integralrechnung: Sei $f : [a, b] \to \mathbb{R}$ stetig.
(1) Ist $G : [a, b] \to \mathbb{R}$ eine differenzierbare Funktion mit $G' = f$ auf $[a, b]$, so gilt

$$\int_a^b f(x)\,dx = G(b) - G(a) \ \left(=: G(x)\Big|_a^b =: [G(x)]_a^b \right).$$

(2) Setzt man $F : [a, b] \to \mathbb{R}$, $x \mapsto F(x) := \int_a^x f(\xi)\,d\xi$, so ist F auf $[a, b]$ differenzierbar mit $F' = f$ auf $[a, b]$.

Beweis. (2) haben wir in 12.9 gezeigt. (1) folgt aus (2) wie im Beispiel 12.9. □

12.11. Stammfunktionen: Sind $f, F : I \to \mathbb{R}$ Funktionen, wobei F auf I differenzierbar ist mit $F' = f$ auf I, so heißt F eine *Stammfunktion von f auf I*.

Bemerkung: Der Hauptsatz 12.10 besagt, dass man zur Berechnung von Integralen stetiger Funktionen Stammfunktionen verwenden kann und dass stetige Funktionen Stammfunktionen besitzen. Für eine Stammfunktion von $f : I \to \mathbb{R}$ schreibt man wegen 12.10(1) auch $\int f(x)\,dx$ (*unbestimmtes Integral*).

Zwei Stammfunktionen von $f : I \to \mathbb{R}$ unterscheiden sich nach 11.8(2) nur durch eine additive Konstante.

Beispiele: Auf $I = \mathbb{R}$ gilt $\int \cos x\,dx = \sin x + c$, wobei $c \in \mathbb{R}$ eine feste Konstante ist.

Auf $I = (-\infty, 0)$ oder $I = (0, \infty)$ gilt $\int \frac{1}{x}\,dx = \ln|x|$. Dies ist für $x > 0$ klar. Für $x < 0$ ist nach Kettenregel:

$$\frac{d}{dx}\ln|x| = \frac{1}{|x|} \cdot (-1) = \frac{1}{x}.$$

Erinnerung: Eine Funktion $f : I \to \mathbb{R}$ heißt *stetig differenzierbar auf I* und wir schreiben $f \in C^1(I)$, falls f auf I differenzierbar ist und $f' : I \to \mathbb{R}$ stetig ist.

Bemerkung: Ist $f \in C^1(I)$, so gilt $\int f'\,dx = f$ auf I.

Beispiel: Sei $f : [0, 1] \to \mathbb{R}$ gegeben durch $f(x) := \begin{cases} x^{3/2}\sin(\frac{1}{x}) & , x \in (0, 1] \\ 0 & , x = 0 \end{cases}$. Für $x \in (0, 1]$ ist f in x differenzierbar mit

$$f'(x) = \frac{3}{2}\sqrt{x}\sin(\frac{1}{x}) - \frac{1}{\sqrt{x}}\cos(\frac{1}{x}).$$

Beachte, dass diese Funktion auf $(0,1]$ nicht beschränkt ist.

Die Funktion f ist in 0 differenzierbar mit $f'(0)=0$, denn für jedes $h \in (0,1]$ gilt

$$\left|\frac{f(h)-f(0)}{h}\right| = \sqrt{h}\left|\sin\left(\frac{1}{h}\right)\right| \le \sqrt{h} \to 0 \quad (h \to 0+),$$

dh $\lim_{h\to 0+} \frac{f(h)-f(0)}{h} = 0$.

Die Funktion f ist auf $[0,1]$ differenzierbar, aber $f \notin C^1([0,1])$, da f' nicht stetig ist. f' ist hier sogar unbeschränkt, insbesondere also $f' \notin R[0,1]$.

Die beiden folgenden Integrationsregeln ergeben sich über den Hauptsatz aus Differentiationsregeln.

12.12. Partielle Integration: Seien $f, g \in C^1(I)$. Dann gilt

$$\int f'g\,dx = fg - \int fg'\,dx \quad \text{auf } I.$$

Ist $I=[a,b]$, so gilt

$$\int_a^b f'(x)g(x)\,dx = f(x)g(x)\Big|_a^b - \int_a^b f(x)g'(x)\,dx.$$

Beweis. Es gilt $(fg)' = f'g + fg'$ auf I. Somit hat $f'g+fg'$ die Stammfunktion fg auf I, woraus die Behauptungen folgen (für die zweite Formel verwenden wir den Hauptsatz 12.10). □

Beispiele: (1) $\int \underbrace{x}_{g}\underbrace{e^x}_{f'}\,dx = \underbrace{x}_{g}\underbrace{e^x}_{f} - \int \underbrace{1}_{g'}\cdot\underbrace{e^x}_{f}\,dx = xe^x - e^x$.

(2) $\int \ln x\,dx = \int \underbrace{1}_{f'}\cdot\underbrace{\ln x}_{g}\,dx = \underbrace{x}_{f}\underbrace{\ln x}_{g} - \int \underbrace{x}_{f}\underbrace{\frac{1}{x}}_{g'}\,dx = x\ln x - x$.

12.13. Integration durch Substitution: Seien $I, J \subseteq \mathbb{R}$ Intervalle, $f : I \to \mathbb{R}$ stetig und $g \in C^1(J)$ mit $g(J) \subseteq I$. Dann ist

$$\int f(g(t))g'(t)\,dt = \int f(x)\,dx\Big|_{x=g(t)} \quad \text{auf } J.$$

Ist $g'(t) \neq 0$ für alle $t \in J$, so ist g auf J streng monoton und

$$\int f(x)\,dx = \int f(g(t))g'(t)\,dt\Big|_{t=g^{-1}(x)} \quad \text{auf } g(J).$$

Ist $I = [a, b]$ und $g(\alpha) = a$, $g(\beta) = b$ sowie $J = \begin{cases} [\alpha, \beta] & , \alpha \leq \beta \\ [\beta, \alpha] & , \alpha > \beta \end{cases}$, so gilt

$$\int_a^b f(x)\,dx = \int_\alpha^\beta f(g(t))g'(t)\,dt$$

Beweis. Wähle eine Stammfunktion F von f auf I. Dann ist $G := F \circ g$ eine Stammfunktion von $h := (f \circ g) \cdot g'$ auf J (denn $G' = (F \circ g)' = (F' \circ g)g' = (f \circ g)g'$). Also ist

$$\int h(t)\,dt = G(t) = F(g(t)) = \int f(x)\,dx|_{x=g(t)}$$

auf J. Ist $g'(t) \neq 0$ für alle $t \in J$, so ist nach dem Zwischenwertsatz entweder $g' > 0$ auf J oder es ist $g' < 0$ auf J. In jedem Fall ist g streng monoton auf J und besitzt also eine Umkehrfunktion $g^{-1} : g(J) \to J$ (beachte auch, dass $g(J)$ ein Intervall ist). Dann ist

$$\int h(t)\,dt\Big|_{t=g^{-1}(x)} = G(g^{-1}(x)) = F(g(g^{-1}(x)) = F(x) = \int f(x)\,dx$$

auf $g(J)$. Schließlich verwenden wir den Hauptsatz. □

Merkregel: Ist $y = y(x)$ eine differenzierbare Funktion, so schreibt man für die Ableitung y' auch $\frac{dy}{dx}$. In $\int f(x)\,dx$ substituiere nun $x = g(t)$, dh fasse x als Funktion von t auf. Dann ist $\frac{dx}{dt} = g'(t)$ und man erhält (formal!) "$dx = g'(t)\,dt$" (dies ist nur eine Schreibweise, da "dx" oder "dt" hier **keine mathematische Bedeutung** tragen).

Beispiele: (1) $\int_0^1 \frac{e^{2x}+1}{e^x}\,dx$, substituiere $t = e^x$, also $x = \ln t$. Dann ist $dx = dt/t$ und aus $x : 0 \to 1$ ergibt sich $t : 1 \to e$. Wir erhalten:

$$\int_0^1 \frac{e^{2x}+1}{e^x}\,dx = \int_1^e (1 + t^{-2})\,dt = (t - t^{-1})|_1^e = e - 1/e.$$

(2) $\int_0^1 \sqrt{1 - x^2}\,dx$, das ist der Flächeninhalt eines Viertelkreises mit Radius 1. Substituiere $x = \sin t$. Dann ist $dx = \cos t\,dt$ und aus $x : 0 \to 1$ ergibt sich $t : 0 \to \pi/2$. Wir erhalten

$$\int_0^1 \sqrt{1 - x^2}\,dx = \int_0^{\pi/2} \underbrace{\sqrt{1 - \sin^2 t}}_{=\sqrt{\cos^2 t}} \cos t\,dt = \int_0^{\pi/2} \cos^2 t\,dt,$$

da $\cos \geq 0$ auf $[0, \pi/2]$ ist. Nun schreiben wir

$$\cos(2t) = \cos^2 t - \sin^2 t = \cos^2 t - (1 - \cos^2 t) = 2\cos^2 t - 1,$$

also $\cos^2 t = \frac{1}{2}(1 + \cos(2t))$, und erhalten

$$\int_0^{\pi/2} \cos^2 t\,dt = \frac{1}{2}\Big(\int_0^{\pi/2} 1\,dt + \underbrace{\int_0^{\pi/2} \cos(2t)\,dt}_{=0}\Big) = \frac{\pi}{4}.$$

Der Flächeninhalt eines Kreises mit Radius 1 is also π.

(3) $\int xe^{-x^2}\,dx$, substituiere $u = x^2$, also $2x\,dx = du$ bzw. $x\,dx = \frac{du}{2}$. Dann ist

$$\int xe^{-x^2}\,dx = \frac{1}{2}\int e^{-u}\,du\Big|_{u=x^2} = -\frac{1}{2}e^{-u}\Big|_{u=x^2} = -\frac{1}{2}e^{-x^2}.$$

Mittels partieller Integration kann man nun berechnen:

$$\int x^3e^{-x^2}\,dx = \int \underbrace{x^2}_{g}\cdot\underbrace{xe^{-x^2}}_{f'}\,dx = -\frac{x^2}{2}e^{-x^2} + \int xe^{-x^2}\,dx = -\frac{x^2}{2}e^{-x^2} - \frac{1}{2}e^{-x^2}.$$

Man kann auf diese Weise Stammfunktionen bestimmen von $x^n e^{-x^2}$ für ungerade $n \in \mathbb{N}$. Für gerade $n \in \mathbb{N}$ gibt es diese Stammfunktionen nicht in geschlossener Form.

(4) $\int \frac{1}{1+e^x}\,dx$, substituiere $e^x = y$, also $x = \ln y$ und $dx = \frac{dy}{y}$, wobei man $y > 0$ beachte:

$$\int \frac{1}{1+e^x}\,dx = \int \frac{1}{(1+y)y}\,dy\Big|_{y=e^x}.$$

Zur Umformung des Integranden macht man den Ansatz

$$\frac{1}{(1+y)y} = \frac{A}{y} + \frac{B}{1+y}.$$

Multipliziert man mit y und setzt $y = 0$, so erhält man $A = 1$. Multipliziert man mit $1+y$ und setzt $y = -1$, so erhält man $B = -1$. Das ist ein Spezialfall der Partialbruchzerlegung ($\rightarrow$ nächstes Semester). Wir haben also

$$\int \frac{1}{(1+y)y}\,dy\Big|_{y=e^x} = \int \frac{1}{y} - \frac{1}{1+y}\,dy\Big|_{y=e^x} = (\ln y - \ln(1+y))\Big|_{y=e^x} = x - \ln(1+e^x).$$

12.14. Satz: Seien $f, g \in R[a,b]$ und $D := f([a,b])$.

(1) Sei $L \geq 0$ und $h : D \to \mathbb{R}$ eine Funktion mit

$$\forall s,t \in D : |h(t) - h(s)| \leq L|t-s|. \tag{L}$$

Dann gilt $h \circ f \in R[a,b]$.

(2) Es ist $|f| \in R[a,b]$ und

$$\Big|\int_a^b f\,dx\Big| \leq \int_a^b |f|\,dx \qquad \text{(Dreiecksungleichung für Integrale).}$$

(3) Es ist $fg \in R[a,b]$.

(4) Ist $c > 0$ und gilt $|f(x)| \geq c$ für alle $x \in [a,b]$, so ist $1/f \in R[a,b]$.

Bemerkung: Eine Funktion h mit (L) heißt *Lipschitz-stetig* mit *Lipschitz-Konstante* L. Eine solche Funktion ist insbesondere stetig (sogar gleichmäßig stetig).

Beweis. (1) Sei $Z = \{x_0, x_1, \dots, x_n\}$ eine Zerlegung und I_j, M_j, m_j wie in 12.1, sowie

$$\tilde{M}_j := \sup(h \circ f)(I_j) \quad \text{und} \quad \tilde{m}_j := \inf(h \circ f)(I_j).$$

Wir haben nun

$$\tilde{M}_j - \tilde{m}_j = \sup\{\underbrace{h(f(x)) - h(f(\tilde{x}))}_{\leq L|f(x)-f(\tilde{x})|} : x, \tilde{x} \in I_j\} \leq L(M_j - m_j).$$

Man erhält $S_{h\circ f}(Z) - s_{h\circ f}(Z) \leq L(S_f(Z) - s_f(Z))$, und $h \circ f \in R[a,b]$ folgt mithilfe von 12.4.

(2) Wende (1) an auf $h(t) = |t|$. Es gilt $||t| - |s|| \leq |t - s|$ für alle $t, s \in \mathbb{R}$, dh (L) gilt mit $L = 1$. Also ist $|f| \in R[a,b]$. Weiter ist

$$\Big|\int_a^b f\,dx\Big| = \max\Big\{\int_a^b \underbrace{f}_{\leq|f|}\,dx, \int_a^b \underbrace{(-f)}_{\leq|f|}\,dx\Big\} \leq \int_a^b |f|\,dx.$$

(3) Wegen $fg = \frac{1}{4}((f+g)^2 - (f-g)^2)$ reicht es, $f^2 \in R[a,b]$ zu zeigen. Da f beschränkt ist, gibt es $\gamma > 0$ mit $|f(x)| \leq \gamma$ für alle $x \in [a,b]$. Also gilt $D \subseteq [-\gamma, \gamma]$. Für $s, t \in D$ gilt somit:

$$|t^2 - s^2| = |t+s||t-s| \leq (|t| + |s|)|t-s| \leq 2\gamma|t-s|.$$

Mit $h(t) = t^2$ in (1) folgt $f^2 \in R[a,b]$.

(4) Nach Voraussetzung gilt $D \subseteq (-\infty, -c] \cup [c, \infty)$. Somit ist für $t, s \in D$:

$$\left|\frac{1}{t} - \frac{1}{s}\right| = \frac{|s-t|}{|s||t|} \leq \frac{1}{c^2}|t-s|.$$

Mit $h(t) := 1/t$ in (1) folgt die Behauptung. □

Zum Abschluss noch ein Beispiel: Wir wollen den ohmschen Widerstand eines kegelstumpfartigen Kontaktes berechnen unter Verwendung der folgenden Formel für einen zylinderförmigen Draht:

$$R = \rho\frac{l}{A},$$

wobei ρ der spezifische Widerstand des (homogenen) Drahtes sei, l die Länge und A die Querschittsfläche. Der Kegelstumpf habe ebenfalls die Länge l und und kreisförmige Querschnitte, an den Enden mit den Radien r_1 bzw. r_2, wobei $0 < r_1 < r_2$ sei.

Wir verwenden die Variable $x \in [0, l]$ zur Parametrisierung der Querschnittsmittelpunkte. An der Stelle x hat der Querschnitt die Fläche $A(x) = \pi r(x)^2$, wobei $r(x) = \frac{r_2 - r_1}{l} x + r_1$ ist (beachte $r(0) = r_1$, $r(l) = r_2$). An der Stelle x hat eine Zylinderscheibe der Dicke Δx nach der obigen Formel den ohmschen Widerstand

$$\Delta R = \rho \frac{\Delta x}{A(x)} = \rho \frac{\Delta x}{\pi r(x)^2}.$$

Summation über viele solche Scheiben ergibt mit $\Delta x \to 0$ das Integral[5]

$$R = \frac{\rho}{\pi} \int_0^l \frac{dx}{r(x)^2}.$$

Wir substituieren hier $r = r(x)$, also $dr = r'(x)\, dx = \frac{r_2 - r_1}{l}\, dx$ und erhalten

$$R = \frac{\rho}{\pi} \int_0^l \frac{dx}{r(x)^2} = \frac{\rho\, l}{\pi(r_2 - r_1)} \int_{r_1}^{r_2} \frac{dr}{r^2} = \frac{\rho\, l}{\pi(r_2 - r_1)} \Big(\frac{1}{r_1} - \frac{1}{r_2} \Big) = \frac{\rho\, l}{\pi r_1 r_2}$$

für den ohmschen Widerstand des Kegelstumpfes.

[5]Tatsächlich ist eine solche Summe eine Riemannsche Summe für das angegebene Integral.

13 Ergänzungen zur Integration

Wir erinnern an den Begriff der gleichmäßigen Konvergenz aus 7.18: Eine Folge $(f_n)_{n\in\mathbb{N}}$ von Funktionen $f_n : [a,b] \to \mathbb{R}$ heißt *auf* $[a,b]$ *gleichmäßig konvergent* gegen eine Funktion $f : [a,b] \to \mathbb{R}$, falls es eine Nullfolge (α_n) so gibt, dass für jedes $n \in \mathbb{N}$ gilt:

$$\forall x \in [a,b] : |f_n(x) - f(x)| \leq \alpha_n.$$

13.1. Konvergenzsatz für Integrale: Sei $f : [a,b] \to \mathbb{R}$ eine Funktion und $(f_n)_{n\in\mathbb{N}}$ eine Folge von Funktionen $f_n \in R[a,b]$, die auf $[a,b]$ gleichmäßig gegen f konvergiert. Dann gilt $f \in R[a,b]$ und

$$\lim_{n\to\infty} \int_a^b f_n(x)\,dx = \int_a^b f(x)\,dx \quad \left[= \int_a^b \lim_{n\to\infty} f_n(x)\,dx\right].$$

Bei gleichmäßiger Konvergenz kann man also Limes und Integral vertauschen.

Beweis. Wir verwenden 12.4. Sei $\varepsilon > 0$. Wähle $n \in \mathbb{N}$ mit

$$\forall x \in [a,b] : |f_n(x) - f(x)| < \frac{\varepsilon}{3(b-a)} =: \delta$$

und eine Zerlegung Z mit $S_{f_n}(Z) - s_{f_n}(Z) < \varepsilon/3$. Sind I_j, M_j und m_j wie in 12.1, so haben wir für jedes j (vgl. Beweis von 12.5):

$$M_j - m_j = \sup\{|f(x) - f(\tilde{x})| : x, \tilde{x} \in I_j\}.$$

Wegen

$$|f(x) - f(\tilde{x})| \leq |f(x) - f_n(x)| + |f_n(x) - f_n(\tilde{x})| + |f_n(\tilde{x}) - f(\tilde{x})| \leq 2\delta + |f_n(x) - f_n(\tilde{x})|$$

erhalten wir

$$S_f(Z) - s_f(Z) \leq 2\delta(b-a) + S_{f_n}(Z) - s_{f_n}(Z) = \frac{2\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon,$$

so dass $f \in R[a,b]$ aus 12.4 folgt. Wegen 12.14(2) haben wir

$$\Big| \int_a^b f_n\,dx - \int_a^b f\,dx \Big| \leq \int_a^b \underbrace{|f_n - f|}_{\leq \alpha_n}\,dx \leq (b-a)\alpha_n,$$

woraus Konvergenz der Integrale folgt. $\square$

Beispiele: (1) Sei $b > 0$ und $f_n(x) := e^{-x^2/n}$ für $x \in [-b, b]$ und $n \in \mathbb{N}$. Dann konvergiert (f_n) auf $[-b, b]$ punktweise gegen $f : [-b, b] \to \mathbb{R}$, $f(x) = 1$. Für jedes $x \in [-b, b]$ und $n \in \mathbb{N}$ gilt dabei

$$|f_n(x) - f(x)| = 1 - e^{x^2/n} \leq 1 - e^{-b^2/n} =: \alpha_n.$$

Wegen $\alpha_n \to 0$ konvergiert (f_n) auf $[-b, b]$ gleichmäßig gegen f, also

$$\int_{-b}^{b} e^{-x^2/n}\, dx \;\longrightarrow\; \int_{-b}^{b} 1\, dx = 2b \quad (n \to \infty).$$

(2) Sei $f_n : [0, 2] \to \mathbb{R}$ definiert durch

$$f_n(x) = \begin{cases} n^2 x & , x \in [0, 1/n] \\ n^2(2/n - x) & , x \in (1/n, 2/n] \\ 0 & , x \in (2/n, 2] \end{cases} .$$

Dann gilt $\int_0^2 f_n(x)\, dx = 1$ für jedes $n \in \mathbb{N}$, und für jedes $x \in [0, 2]$ konvergiert die Folge $(f_n(x))$ gegen 0. Setzt man $f : [0, 2] \to \mathbb{R}$, $f(x) = 0$, so konvergiert die Folge (f_n) auf $[0, 2]$ *punktweise* gegen f. Die Folge der Integrale $(\int_0^2 f_n(x)\, dx)$ konvergiert jedoch *nicht* gegen $\int_0^2 f(x)\, dx = 0$. Insbesondere konvergiert (f_n) also auf $[0, 2]$ *nicht gleichmäßig* gegen f (das kann man natürlich auch direkt einsehen). Dieses Beispiel zeigt, dass Satz 13.1 ohne Gleichmäßigkeit der Konvergenz im allgemeinen falsch ist.

(3) Sei $f_n : [0, \pi] \to \mathbb{R}$, $f_n(x) = \sin(nx)$. Dann gilt

$$(f_n(\pi/2)) = (1, 0, -1, 0, 1, 0, -1, 0, \ldots),$$

insbesondere ist (f_n) auf $[0, \pi]$ nicht punktweise konvergent. Trotzdem gilt

$$\int_0^{\pi} f_n(x)\, dx = \int_0^{\pi} \sin(nx)\, dx = \frac{1}{n}\Big[- \cos(nx) \Big]_0^{\pi} = \frac{1}{n}(1 - (-1)^n) \;\longrightarrow\; 0 \quad (n \to \infty).$$

Aus Konvergenz der Integrale kann also nicht auf punktweise Konvergenz der Integranden geschlossen werden.

13.2. Anwendung (Vertauschen von Limes und Differentiation): Sei (f_n) eine Folge von Funktionen $f_n \in C^1[a, b]$ und $f, g : [a, b] \to \mathbb{R}$ derart, dass (f_n') auf $[a, b]$ gleichmäßig gegen g konvergiert und dass (f_n) auf $[a, b]$ punktweise gegen f konvergiert. Dann gilt $f \in C^1[a, b]$ und $f' = g$ auf $[a, b]$.

Beweis. Zunächst ist g nach 13.1 integrierbar, als gleichmäßiger Limes stetiger Funktionen aber auch stetig. Für jedes $x \in [a, b]$ gilt nach 13.1:

$$\int_a^x f_n'(t)\, dt \;\longrightarrow\; \int_a^x g(t)\, dt$$

Andererseits steht links nach dem Hauptsatz $f_n(x) - f_n(a)$ und nach Voraussetzung gilt

$$\lim_{n\to\infty} f_n(x) - f_n(a) = f(x) - f(a).$$

Also ist

$$f(x) - f(a) = \int_a^x g\,dt, \quad x \in [a,b],$$

und nach dem Hauptsatz ist $f \in C^1[a,b]$ mit $f' = g$ auf $[a,b]$. $\square$

Man kann diesen Satz dazu verwenden, die gliedweise Differenzierbarkeit von Potenzreihen einzusehen. Dann braucht man den Beweis nicht direkt zu führen wie in 11.14.

14 Uneigentliche Integrale

Die folgenden Vereinbarungen sollen für den Rest des Kapitels gelten:

> Ist $I \subseteq \mathbb{R}$ ein Intervall und $f : I \to \mathbb{R}$ eine Funktion, so sei f über jedes Intervall $[a, b] \subseteq I$ integrierbar.
>
> Es seien stets $a, b \in \mathbb{R}$, $\alpha \in \mathbb{R} \cup \{-\infty\}$ und $\beta \in \mathbb{R} \cup \{\infty\}$ mit $\alpha < b$ und $a < \beta$.

14.1. Konvergenz uneigentlicher Integrale: Sei $f : [a, \beta) \to \mathbb{R}$ (bzw. $(\alpha, b] \to \mathbb{R}$) eine Funktion. Das *uneigentliche Integral* $\int_a^\beta f(x)\,dx$ (bzw. $\int_\alpha^b f(x)\,dx$) heißt *konvergent*, falls der Limes $\lim_{r\to\beta-} \int_a^r f(x)\,dx$ (bzw. $\lim_{r\to\alpha+} \int_r^b f(x)\,dx$) existiert und reell ist. In diesem Fall setzt man

$$\int_a^\beta f(x)\,dx := \lim_{r\to\beta-} \int_a^r f(x)\,dx \quad \Big(\text{bzw.} \quad \int_\alpha^b f(x)\,dx := \lim_{r\to\alpha+} \int_r^b f(x)\,dx\Big).$$

Ein nicht konvergentes uneigentliches Integral heißt *divergent*.

Beispiele: (1) Sei $\gamma > 0$. Dann gilt für jedes $r > 1$:

$$\int_1^r \frac{1}{x^\gamma}\,dx = \begin{cases} \ln r & ,\gamma = 1 \\ \frac{r^{1-\gamma}-1}{1-\gamma} & ,\gamma \neq 1 \end{cases},$$

also ist $\int_1^\infty x^{-\gamma}\,dx$ konvergent genau dann, wenn $\gamma > 1$ ist. In diesem Fall ist das Integral $= 1/(\gamma - 1)$.

(2) Für $r > 0$ gilt

$$\int_0^r \frac{1}{1+x^2}\,dx = \arctan r \to \frac{\pi}{2} \quad (r \to \infty).$$

Also ist $\int_0^\infty \frac{1}{1+x^2}\,dx$ konvergent und $= \pi/2$.

(3) Sei $\gamma > 0$. Dann gilt für $r \in (0, 1)$:

$$\int_r^1 \frac{1}{x^\gamma}\,dx = \begin{cases} -\ln r & ,\gamma = 1 \\ \frac{1-r^{1-\gamma}}{1-\gamma} & ,\gamma \neq 1 \end{cases},$$

also ist $\int_0^1 x^{-\gamma}\,dx$ konvergent genau dann, wenn $\gamma < 1$ ist. In diesem Fall ist das Integral $= 1/(1 - \gamma)$.

(4) Analog zu (2): $\int_{-\infty}^0 \frac{1}{1+x^2}\,dx$ ist konvergent und $= \pi/2$.

Definition: Sei $f:(\alpha,\beta)\to\mathbb{R}$ eine Funktion. Das *uneigentliche Integral* $\int_\alpha^\beta f(x)\,dx$ heißt *konvergent*, falls es ein $c\in(\alpha,\beta)$ so gibt, dass die uneigentlichen Integrale $\int_\alpha^c f(x)\,dx$ und $\int_c^\beta f(x)\,dx$ **beide** konvergent sind. In diesem Fall setzt man

$$\int_\alpha^\beta f(x)\,dx := \int_\alpha^c f(x)\,dx + \int_c^\beta f(x)\,dx,$$

und die Definition ist unabhängig von $c\in(\alpha,\beta)$.

Das Integral $\int_\alpha^\beta f(x)\,dx$ heißt *divergent*, falls $\int_\alpha^\beta f(x)\,dx$ nicht konvergent ist.

Beispiele: (5) Sei $\gamma>0$. Nach den Beispielen (1) und (3) ist $\int_0^\infty x^{-\gamma}\,dx$ divergent.

(6) Nach den Beispielen (2) und (4) ist $\int_{-\infty}^\infty \frac{1}{1+x^2}\,dx$ konvergent und $=\pi$.

(7) $\int_{-\infty}^\infty x\,dx$ ist divergent, da $\int_0^\infty x\,dx$ divergent ist. Z.B. konvergiert aber $\int_{-b}^b x\,dx=0$ für $b\to\infty$ gegen 0.

Bemerkung: Wir betrachten im folgenden nur den Fall von Funktionen $f:[a,\beta)\to\mathbb{R}$. Entsprechendes gilt jeweils auch für Funktionen $f:(\alpha,b]\to\mathbb{R}$ bzw. $f:(\alpha,\beta)\to\mathbb{R}$.

14.2. Satz (Cauchykriterium): $\int_a^\beta f(x)\,dx$ konvergiert genau dann, wenn es zu jedem $\varepsilon>0$ ein $c=c(\varepsilon)\in(a,\beta)$ gibt mit

$$\forall u,v\in(c,\beta):\left|\int_u^v f(x)\,dx\right|<\varepsilon.$$

(ohne Beweis)

Beispiele: $\int_0^\infty \frac{\sin x}{x}\,dx$ ist konvergent.

Zunächst beachten wir $\frac{\sin x}{x}\to 1$ für $x\to 0+$, weshalb das Integral bei 0 nicht uneigentlich ist. Für $v>u>0$ gilt:

$$\begin{aligned}
\Big|\int_u^v \frac{\sin x}{x}\,dx\Big| &= \Big|\int_u^v \underbrace{\frac{1}{x}}_{g}\underbrace{\sin x}_{f'}\,dx\Big| \\
&= \Big|\Big[-\frac{\cos x}{x}\Big]_u^v - \int_u^v \frac{\cos x}{x^2}\,dx\Big| \\
&\le \frac{|\cos u|}{u}+\frac{|\cos v|}{v}+\int_u^v\frac{|\cos x|}{x^2}\,dx \\
&\le \frac{1}{u}+\frac{1}{v}+\Big[-\frac{1}{x}\Big]_u^v=\frac{2}{u}.
\end{aligned}$$

Sei $\varepsilon>0$. Für $v>u>2/\varepsilon$ gilt dann $|\int_u^v \frac{\sin x}{x}\,dx|\le 2/u<\varepsilon$.

(2) Für $\beta \in [0,1)$ ist $\int_0^\infty x^\beta \sin(x^2)\,dx$ konvergent. Beachte, dass der Integrand hier nicht gegen Null geht für $x \to \infty$ und für $\beta \in (0,1)$ sogar unbeschränkt ist. Das Integral ist nur bei $\beta = \infty$ uneigentlich. Wieder betrachten wir $v > u > 0$ und substituieren zunächst $y = x^2$, dh $x = \sqrt{y}$, $dx = \frac{dy}{2\sqrt{y}}$, bevor wir partiell integrieren:

$$\begin{aligned}\int_u^v x^\beta \sin(x^2)\,dx &= \frac{1}{2}\int_{u^2}^{v^2} \underbrace{y^{\frac{\beta-1}{2}}}_{\downarrow}\underbrace{\sin y}_{\uparrow}\,dy \\ &= \frac{1}{2}\Big[-y^{\frac{\beta-1}{2}}\cos y\Big]_{u^2}^{v^2} + \frac{\beta-1}{4}\int_{u^2}^{v^2} y^{\frac{\beta-3}{2}}\cos y\,dy.\end{aligned}$$

Nun schätzen wir ab:

$$\begin{aligned}\Big|\int_u^v x^\beta \sin(x^2)\,dx\Big| &\le \frac{u^{\beta-1}}{2} + \frac{v^{\beta-1}}{2} + \frac{1-\beta}{4}\int_{u^2}^{v^2} y^{\frac{\beta-3}{2}}\,dy \\ &= \frac{u^{\beta-1}}{2} + \frac{v^{\beta-1}}{2} + \Big[-\frac{1}{2}y^{\frac{\beta-1}{2}}\Big]_{u^2}^{v^2} = \frac{1}{u^{1-\beta}}.\end{aligned}$$

Wegen $1-\beta > 0$ gilt $\frac{1}{u^{1-\beta}} \to 0$ für $u \to \infty$, und man kann wie unter (1) argumentieren.

14.3. Absolut konvergente uneigentliche Integrale: $\int_a^\beta f(x)\,dx$ heißt *absolut konvergent*, falls $\int_a^\beta |f(x)|\,dx$ konvergent ist.

Beispiel: $\int_0^\infty \frac{\sin x}{x}\,dx$ ist **nicht** absolut konvergent. Für $k \in \mathbb{N}_0$ gilt nämlich

$$\int_{k\pi}^{(k+1)\pi} \frac{|\sin x|}{x}\,dx \ge \underbrace{\int_0^\pi |\sin x|\,dx}_{=2}\,\frac{1}{(k+1)\pi} = \frac{2}{(k+1)\pi},$$

also für $n \in \mathbb{N}$:

$$\int_0^{(n+1)\pi} \frac{|\sin x|}{x}\,dx \ge \sum_{k=0}^n \frac{2}{(k+1)\pi} \to \infty \quad (n \to \infty).$$

Mithilfe von 14.2 kann man zeigen (ähnlich wie bei Reihen, vgl. 7.3 und 7.4):

14.4. Satz: (1) Ist $\int_a^\beta f(x)\,dx$ absolut konvergent, so ist $\int_a^\beta f(x)\,dx$ konvergent und

$$\Big|\int_a^\beta f(x)\,dx\Big| \le \int_a^\beta |f(x)|\,dx.$$

(2) **Majorantenkriterium:**
Ist $|f| \le g$ auf $[a,\beta)$ und $\int_a^\beta g(x)\,dx$ konvergent, so ist $\int_a^\beta f(x)\,dx$ absolut konvergent.

(3) **Minorantenkriterium:**
Ist $f \ge g \ge 0$ auf $[a,\beta)$ und $\int_a^\beta g(x)\,dx$ divergent, so ist $\int_a^\beta f(x)\,dx$ divergent.

Beispiele: (1) $\int_1^\infty \frac{x}{\sqrt{1+x^5}}\,dx$; setze $f(x) := \frac{x}{\sqrt{1+x^5}}$ und $g(x) := x^{-3/2}$. Dann gilt $|f(x)| = f(x) \le g(x)$ für jedes $x \ge 1$. Da $\int_1^\infty x^{-3/2}\,dx$ konvergiert, konvergiert auch $\int_1^\infty f(x)\,dx$ und es gilt:

$$\int_1^\infty f(x)\,dx \le \int_1^\infty g(x)\,dx = 2.$$

(2) $\int_1^\infty \underbrace{\frac{x}{x^2+7x}}_{=:f(x)}\,dx$; setze $g(x) = 1/x$. Dann gilt $f(x)/g(x) = \frac{x^2}{x^2+7x} \to 1$ $(x \to \infty)$ und wir finden $c \ge 1$ mit $f/g \ge 1/2$ auf $[c,\infty)$, dh es gilt $f(x) \ge g(x)/2 = (2x)^{-1}$ für alle $x \ge c$ (man kann hier $c = 7$ nehmen, wie man direkt einsieht). Da $\int_c^\infty (2x)^{-1}\,dx$ divergiert, divergiert auch $\int_c^\infty f(x)\,dx$, und somit divergiert auch $\int_1^\infty f(x)\,dx$.

Warnung: Sind $\int_a^\beta f\,dx$ und $\int_a^\beta g\,dx$ konvergent, so muss $\int_a^\beta fg\,dx$ **nicht** konvergieren!

Ausblick Laplacetransformation: Sei $f : [0,\infty) \to \mathbb{R}$, $t \mapsto f(t)$, eine Funktion, die auf jedem Intervall $[0,b]$ stückweise stetig ist (insbesondere ist also f über jedem Intervall $[0,b]$ integrierbar). Weiter sei f beschränkt, also etwa $|f(t)| \le M$ für alle $t \ge 0$. Ist $s > 0$, so ist das Integral $\mathscr{L}\{f\}(s) := \int_0^\infty e^{-st} f(t)\,dt$ absolut konvergent: Es gilt nämlich $|e^{-st} f(t)| \le M e^{-st}$ und für $r > 0$:

$$\int_0^r M e^{-st}\,dt = M\Big[-\frac{e^{-st}}{s}\Big]_0^r = \frac{M}{s}(1 - e^{-sr}),$$

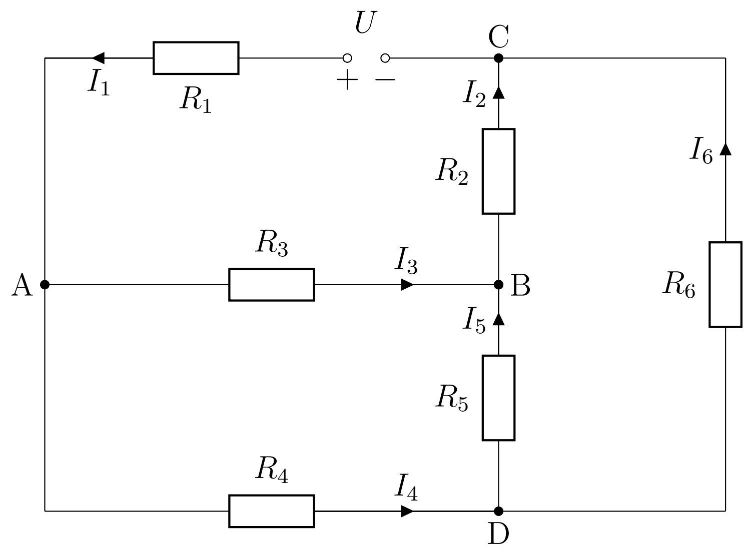
und $e^{-sr} \to 0$ für $r \to \infty$ wegen $s > 0$. Wir erhalten

$$\Big|\int_0^\infty e^{-st} f(t)\,dt\Big| \le \frac{M}{s}.$$

Zur Analyse von Zeitsignalen f betrachtet man nun $\mathscr{L}\{f\}(s)$ als Funktion von s.

15 Grundzüge der linearen Algebra

15.1. Beispiel: Zur Motivation betrachten wir das folgende elektrische Netzwerk mit elektromotorischer Kraft U:



Wir stellen die beschreibenden Gleichungen nach den Kirchhoffschen Regeln auf:

- *Knotenregel:* In jedem Knoten ist die Summe der zufließenden Ströme gleich der Summe der abfließenden Ströme.
- *Maschenregel:* In jeder Masche ist die Summe der Spannungsabfälle über den Widerständen gleich der Summe der elektromotorischen Kräfte.

Dabei nehmen wir an, dass U und $R_1, R_2, \ldots, R_6$ bekannt sind, und wollen die Ströme $I_1, I_2, \ldots, I_6$ bestimmen. Durch Betrachtung der Knoten A, B, C, D bzw. der Maschen ABC, ABD, BCD, ADC erhalten wir die Gleichungen

$$\begin{array}{cccccccc}
I_1 & & -I_3 & -I_4 & & & = & 0 \\
 & -I_2 & +I_3 & & +I_5 & & = & 0 \\
-I_1 & +I_2 & & & & +I_6 & = & 0 \\
 & & & I_4 & -I_5 & -I_6 & = & 0 \\
R_1 I_1 & +R_2 I_2 & +R_3 I_3 & & & & = & U \\
 & & R_3 I_3 & -R_4 I_4 & -R_5 I_5 & & = & 0 \\
 & R_2 I_2 & & & +R_5 I_5 & -R_6 I_6 & = & 0 \\
R_1 I_1 & & & +R_4 I_4 & & +R_6 I_6 & = & U.
\end{array}$$

Das ist ein *lineares Gleichungssystem* (LGS) mit acht Gleichungen für die sechs Unbekannten $I_1, \ldots, I_6$. Die natürlichen Fragen sind zunächst die nach **Existenz** und **Eindeutigkeit** der Lösung und ggf. die nach deren **Berechnung**.

Bemerkung: Man kann hier sehen, dass zB Addition der ersten drei Gleichungen (bis aufs Vorzeichen) die vierte Gleichung ergibt und Addition der letzten drei Gleichungen die fnfte ergibt. Es gibt hier also redundante Gleichungen. Wir kommen auf dieses Phänomen zurück.

Lineare Gleichungssysteme treten in vielen Zusammenhängen auf. Wir werden sie in der Form $A\vec{x} = \vec{b}$ schreiben, wobei der *Vektor* $\vec{x}$ gesucht wird und der Vektor $\vec{b}$ und die *Matrix* A gegeben sind. Im Beispiel ist $\vec{x} = (I_1, I_2, \ldots, I_6) \in \mathbb{R}^6$ gesucht, die rechte Seite ist $\vec{b} = (0,0,0,0,U,0,0,U) \in \mathbb{R}^8$ und A ist die *Matrix*

$$\begin{pmatrix} 1 & 0 & -1 & -1 & 0 & 0 \\ 0 & -1 & 1 & 0 & 1 & 0 \\ -1 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & -1 & -1 \\ R_1 & R_2 & R_3 & 0 & 0 & 0 \\ 0 & 0 & R_3 & -R_4 & -R_5 & 0 \\ 0 & R_2 & 0 & 0 & R_5 & -R_6 \\ R_1 & 0 & 0 & R_4 & 0 & R_6 \end{pmatrix} \in \mathbb{R}^{8\times 6}$$

mit acht Zeilen und sechs Spalten.

15.2. Lineare Gleichungssysteme: Seien $n, m \in \mathbb{N}$. Wir schreiben ein *lineares Gleichungssystem* (LGS)

$$\begin{array}{ccccc} a_{11}x_1+ & a_{12}x_2+ & \ldots & +a_{1m}x_m & = b_1 \\ a_{21}x_1+ & a_{22}x_2+ & \ldots & +a_{2m}x_m & = b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1}x_1+ & a_{n2}x_2+ & \ldots & +a_{nm}x_m & = b_n \end{array}$$

mit gegebenen $a_{jk}, b_j \in \mathbb{K}$ und gesuchten $x_1, x_2, \ldots, x_m \in \mathbb{K}$ in Matrixschreibweise als

$$\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1m} \\ a_{21} & a_{22} & \ldots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \ldots & a_{nm} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

oder als

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

wobei $A \in \mathbb{K}^{n\times m}$ und $\vec{b} \in \mathbb{K}^n$ gegeben und $\vec{x} \in \mathbb{K}^m$ gesucht ist. Hierbei verwenden wir das *Matrix-Vektor-Produkt*:

Für $A = (a_{jk})_{j=1}^{n}{}_{k=1}^{m} \in \mathbb{K}^{n\times m}$ und $\vec{x} = (x_k)_{k=1}^m \in \mathbb{K}^m$ ist

$$A\vec{x} = \Big(\sum_{k=1}^m a_{jk}x_k\Big)_{j=1}^n \in \mathbb{K}^n.$$

Beachte: Wir schreiben Vektoren $\vec{x} = (x_1, x_2, \ldots, x_m) \in \mathbb{K}^m$, die wir bisher als m-Tupel geschrieben haben, ab jetzt ggf. als **Spaltenvektoren**! Dies ist die übliche Konvention.

Beispiel: Sei $l \in \{1, 2, \ldots, m\}$. Wir bezeichnen den Vektor $\vec{e}_l = (\delta_{kl})_{k=1}^m \in \mathbb{K}^m$ als den *l-ten Einheitsvektor*, wobei $\delta_{kl} = \begin{cases} 1 & , k = l \\ 0 & , k \neq l \end{cases}$ das *Kroneckersymbol* ist. Wir erhalten dann

$$A\vec{e}_l = \Big(\sum_{k=1}^m a_{jk}\delta_{kl} \Big)_{j=1}^n = (a_{jl})_{j=1}^n \in \mathbb{K}^n,$$

dh $A\vec{e}_l$ **ist die l-te Spalte der Matrix** A. Für $n = m = 3$ und $l = 2$ ist etwa

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \end{pmatrix}.$$

15.3. Eigenschaften des Matrix-Vektor-Produktes:
Für alle $A, B \in \mathbb{K}^{n\times m}$ und $\vec{x}, \vec{y} \in \mathbb{K}^m$, $\alpha \in \mathbb{K}$ gilt:

$$A(\vec{x} + \vec{y}) = A\vec{x} + A\vec{y}, \qquad (A + B)\vec{x} = A\vec{x} + B\vec{x}, \qquad A(\alpha\vec{x}) = \alpha(A\vec{x}) = (\alpha A)\vec{x},$$

wobei für $A = (a_{jk})$, $B = (b_{jk})$ die Matrizen $A + B, \alpha A \in \mathbb{K}^{n\times m}$ gegeben sind durch

$$A + B = (a_{jk} + b_{jk})_{j=1}^n{}_{k=1}^m, \qquad \alpha A = (\alpha a_{jk})_{j=1}^n{}_{k=1}^m,$$

dh an jeder Stelle (j, k) werden die Einträge addiert bzw. mit α multipliziert.

Satz: 1) Mit diesen Verknüpfungen ist $\mathbb{K}^{n\times m}$ ein $\mathbb{K}$-Vektorraum.
2) Für festes $A \in \mathbb{K}^{n\times m}$ ist die Abbildung $\mathbb{K}^m \to \mathbb{K}^n$, $\vec{x} \mapsto A\vec{x}$ linear.
3) Für festes $\vec{x} \in \mathbb{K}^m$ ist die Abbildung $\mathbb{K}^{m\times n} \to \mathbb{K}^n$, $A \mapsto A\vec{x}$ linear.

Bemerkung: Das bedeutet, dass das lineare Gleichungssystem

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

wobei $A \in \mathbb{K}^{n\times m}$ und $\vec{b} \in \mathbb{K}^m$ gegeben sind und $\vec{x} \in \mathbb{K}^n$ gesucht ist, ein Spezialfall der folgenden Situation ist:

15.4. Lineare Gleichungen: Sei V und W gegebene $\mathbb{K}$-Vektorräume, $\phi : V \to W$ sei eine gegebene lineare Abbildung und $w \in W$ sei gegeben. Man sucht $v \in V$ mit

$$\phi(v) = w. \tag{LG}$$

Ist $w = 0$, so heißt die lineare Gleichung (LG) *homogen*, sonst *inhomogen*, wobei $w \neq 0$ als *Inhomogenität* bezeichnet wird.

Erinnerung: Wir hatten in dieser Situation die Notationen $\operatorname{Kern} \phi = \{v \in V : \phi(v) = 0\}$ und $\operatorname{Bild} \phi = \{\phi(v) : v \in V\}$ eingeführt und gezeigt, dass $\operatorname{Kern} \phi$ ein Untervektorraum von V und $\operatorname{Bild} \phi$ ein Untervektorraum von W ist (vgl. 8.5).[6] Beachte, dass $\operatorname{Kern} \phi$ die Menge der Lösungen der homogenen Gleichung ist.

Satz: 1) Die Gleichung (LG) ist genau dann lösbar, wenn $w \in \operatorname{Bild} \phi$ ist.

2) Ist (LG) lösbar und $v_0 \in V$ eine Lösung, so gilt für die Lösungsmenge

$$\{v \in V : \phi(v) = w\} = v_0 + \operatorname{Kern} \phi.$$

Insbesondere ist v_0 genau dann die einzige Lösung, wenn $\operatorname{Kern} \phi = \{0\}$ ist.

3) ϕ ist injektiv genau dann, wenn $\operatorname{Kern} \phi = \{0\}$ gilt.

Beweis. 1) ist klar. zu 2): Gilt $\phi(v) = w$, so ist $\phi(v - v_0) = \phi(v) - \phi(v_0) = w - w = 0$ und somit $v - v_0 \in \operatorname{Kern} \phi$, dh $v \in v_0 + \operatorname{Kern} \phi$. Ist umgekehrt $v = v_0 + v_1$ mit $v_1 \in \operatorname{Kern} \phi$, so gilt $\phi(v) = \phi(v_0) + \phi(v_1) = w + 0 = w$. Somit ist 2) gezeigt. 3) folgt aus 2). $\square$

Bemerkung: Teil 2) des Satzes bedeutet, dass man die Lösungen der inhomogenen Gleichung erhält, wenn man zu *einer* Lösung der inhomogenen Gleichung alle Lösungen der homogenen Gleichung addiert. Anders ausgedrückt besagt Teil 2) des Satzes, dass die Lösungsmenge von (LG), wenn sie nicht leer ist, ein *affiner Teilraum* von V ist im Sinne der folgenden Definition.

Definition: Sei V ein $\mathbb{K}$-Vektorraum. Eine Teilmenge $F \subseteq V$ heißt *affiner Teil-* oder *Unterraum* von V, falls es einen Untervektorraum U von V und ein $v_0 \in V$ gibt mit

$$F = v_0 + U = \{v_0 + u : u \in U\}.$$

Bemerkung: Ist $F = v_0 + U$ ein affiner Teilraum, so folgt

$$U = \{v - v_0 : v \in F\} = \{v_1 - v_2 : v_1, v_2 \in F\}.$$

Ist umgekehrt $F \subseteq V$ nicht-leer und

$$U := \{v_1 - v_2 : v_1, v_2 \in F\}$$

ein Untervektorraum von V, so ist F ein affiner Teilraum von V und $F = v + U$, wobei man $v \in F$ beliebig wählen kann.

Bemerkung: In $\mathbb{R}^2$ sei $(a, b) \neq (0, 0)$. Dann ist $U := \{(x, y) \in \mathbb{R}^2 : ax + by = 0\}$ ein Untervektorraum (Gerade durch $(0, 0)$ und $F := \{(x, y) \in \mathbb{R}^2 : ax + by = 1\}$ ein affiner

[6] Erinnerung: Eine Teilmenge U eines $\mathbb{K}$-Vektorraumes V ist genau dann ein Untervektorraum von V, falls $0 \in U$ und für $u_1, u_2 \in U$, $\alpha \in \mathbb{K}$ gilt $u_1 + u_2, \alpha u_1 \in U$.

Teilraum. Ist speziell $(a,b)=(1,1)$, so ist U die durch $y=-x$ gegebene Gerade und F ist die durch $y=-x+1$ gegebene Gerade, man sieht hier z.B. $F=(1,0)+U$.

Die affinen Teilräume von $\mathbb{R}^3$ sind Punkte, Geraden, Ebenen und $\mathbb{R}^3$ selber.

Spezialfall $A\vec{x}=\vec{b}$, wobei $A\in\mathbb{K}^{n\times m}$ und $\vec{b}\in\mathbb{K}^n$: Dann sind

$$\operatorname{Kern} A := \{\vec{x}\in\mathbb{K}^m : A\vec{x}=\vec{0}\}\subseteq\mathbb{K}^m \quad\text{und}\quad \operatorname{Bild} A := \{A\vec{x} : \vec{x}\in\mathbb{K}^m\}\subseteq\mathbb{K}^n$$

Untervektorräume von $\mathbb{K}^m$ bzw. $\mathbb{K}^n$. $\operatorname{Kern} A$ ist die Lösungsmenge der *homogenen Gleichung* $A\vec{x}=\vec{0}$.

Die Gleichung $A\vec{x}=\vec{b}$ ist lösbar genau dann, wenn $\vec{b}\in\operatorname{Bild} A$ ist. Ist $\vec{z}\in\mathbb{K}^m$ eine Lösung, so ist die Menge aller Lösungen $\vec{z}+\operatorname{Kern} A$ ein affiner Teilraum von $\mathbb{K}^m$.

15.5. Der lineare Aufspann: Ist V ein $\mathbb{K}$-Vektorraum, $n\in\mathbb{N}$ und sind $v_1,v_2,\ldots,v_n$ Vektoren aus V, so heißt jeder Vektor der Form

$$\sum_{j=1}^{n}\alpha_j v_j \quad \text{mit } \alpha_1,\alpha_2,\ldots,\alpha_n\in\mathbb{K}$$

eine *Linearkombination* (*LK*) der Vektoren $v_1,v_2,\ldots,v_n$.

Satz: Ist V ein $\mathbb{K}$-Vektorraum und $\emptyset\neq M\subseteq V$, so ist

$$\operatorname{lin}(M) := \Big\{\sum_{j=1}^{n}\alpha_j v_j : n\in\mathbb{N}_0, \alpha_j\in\mathbb{K}, v_j\in M\Big\}$$

ein Untervektorraum von V, genannt der *lineare Aufspann von M* (oder *der von M erzeugte Untervektorraum*).

Bemerkung: (1) $\operatorname{lin}(M)$ besteht aus allen Linearkombinationen von Vektoren aus M.

(2) Offenbar ist $\operatorname{lin}(M)$ der *kleinste* Untervektorraum von V, der M enthält. Insbesondere ist U ein Untervektorraum von V genau dann, wenn $\operatorname{lin}(U)=U$ gilt.

Beispiele: (1) In $V=\mathbb{R}^2$ gilt $\operatorname{lin}(\{(1,0)\})=\mathbb{R}\times\{0\}$.

(2) In $V=\mathbb{R}^2$ gilt $\operatorname{lin}(\{(1,0),(0,1)\})=\mathbb{R}^2=\operatorname{lin}(\{(1,0),(0,1),(1,1)\}$.

(3) Für $K=\{(x_1,x_2)\in\mathbb{R}^2 : x_1^2+x_2^2\leq 1\}$ aus Beispiel 8.4(3) gilt $\operatorname{lin}(K)=\mathbb{R}^2$.

(4) Für $C=[0,\infty)\times[0,\infty)$ aus Beispiel 8.4(4) gilt $\operatorname{lin}(C)=\mathbb{R}^2$.

(5) In $V=\mathbb{K}^m$ gilt $\mathbb{K}^m=\operatorname{lin}(\{\vec{e}_1,\vec{e}_2,\ldots,\vec{e}_m\})$, also ist für $A\in\mathbb{K}^{n\times m}$:

$$\operatorname{Bild} A = A\operatorname{lin}(\{\vec{e}_1,\vec{e}_2,\ldots,\vec{e}_m\})=\operatorname{lin}(\{A\vec{e}_1,A\vec{e}_2,\ldots,A\vec{e}_m\}).$$

Wegen des Beispiels in 15.3: **Bild A ist der lineare Aufspann der *Spalten* von A.**

15.6. Lineare Unabhängigkeit: Sei V ein $\mathbb{K}$-Vektorraum.

Definition: Man nennt n Vektoren $v_1, v_2, \ldots, v_n \in V$ *linear unabhängig*, falls für alle $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{K}$ gilt:

$$\sum_{j=1}^{n} \alpha_j v_j = 0 \quad \Longrightarrow \quad \alpha_1 = \alpha_2 = \ldots = \alpha_n = 0.$$

Man sagt dazu, dass man den Nullvektor **nur** als *triviale Linearkombination* der $v_1, v_2, \ldots, v_n$ erhalten kann (wenn nämlich alle $\alpha_j = 0$ sind, ist natürlich auch $\sum_{j=1}^{n} \alpha_j v_j = 0$).

Sind $v_1, v_2, \ldots, v_n$ nicht linear unabhängig, so heißen sie *linear abhängig*. Die Vektoren $v_1, v_2, \ldots, v_n$ sind also genau dann linear abhängig, wenn man $\alpha_1, \alpha_2, \ldots, \alpha_n$ findet, die nicht alle $= 0$ sind, mit $\sum_{j=1}^{n} \alpha_j v_j = 0$.

Beispiele: (1) v ist linear unabhängig genau dann, wenn $v \neq 0$ gilt.

(2) In $\mathbb{R}^2$ sind $(1,0), (0,1)$ linear unabhängig, aber $(1,0), (0,1), (1,1)$ sind linear abhängig, denn $(1,0) + (0,1) - (1,1) = (0,0)$.

(3) In $\mathbb{K}^n$ sind die *Einheitsvektoren* $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$ linear unabhängig. Hierbei ist

$$\vec{e}_j := (0, \ldots, 0, \underbrace{1}_{j\text{-te Stelle}}, 0, \ldots, 0) \quad \text{für } j = 1, 2, \ldots, n.$$

(4) In $C^2(\mathbb{R})$ sind $\sin, \cos$ linear unabhängig. Sind nämlich $\alpha, \beta \in \mathbb{R}$ mit

$$\alpha \sin + \beta \cos = 0, \quad \text{dh mit} \quad \alpha \sin x + \beta \cos x = 0 \quad \text{für alle } x \in \mathbb{R},$$

so erhält man $\alpha = 0$ durch Einsetzen von $x = \frac{\pi}{2}$ und $\beta = 0$ durch Einsetzen von $x = 0$.

(5) Betrachtet man in Beispiel 15.1 die ersten vier Zeilen der Matrix als Vektoren im $\mathbb{R}^6$, so sind diese linear abhängig, da ihre Summe den Nullvektor ergibt. Ebenso sind die letzten vier Zeilen der Matrix als Vektoren linear abhängig, da die Summe der letzten drei die fünfte ergibt.

(6) Ist $\phi : V \to W$ linear und **injektiv** und sind $v_1, v_2, \ldots, v_n \in V$ linear unabhängig, so sind $\phi(v_1), \phi(v_2), \ldots, \phi(v_n)$ linear unabhängig in W: Sind nämlich $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{K}$ mit $\sum_{j=1}^{n} \alpha_j \phi(v_j) = 0$, so folgt $\phi(\sum_{j=1}^{n} \alpha_j v_j) = 0$. Da ϕ injektiv ist, folgt $\sum_{j=1}^{n} \alpha_j v_j = 0$, und aus linearen Unabhängigkeit von $v_1, v_2, \ldots, v_n$ folgt dann $\alpha_1 = \alpha_2 = \ldots = \alpha_n = 0$.

15.7. Zeilenumformungen: Seien n Vektoren $v_1, v_2, \ldots, v_n$ aus dem $\mathbb{K}^m$ gegeben, deren lineare Unabhängigkeit wir untersuchen wollen. Dabei sei $v_j = (v_{j1}, v_{j2}, \ldots, v_{jm})$ für $j = 1, 2, \ldots, n$. Wir schreiben die Vektoren $v_1, v_2, \ldots, v_n$ als **Zeilen** in eine Matrix A, dh

$$A = \begin{pmatrix} v_{11} & v_{12} & \ldots & v_{1m} \\ v_{21} & v_{22} & \ldots & v_{2m} \\ \ldots & \ldots & \ldots & \ldots \\ v_{n1} & v_{n2} & \ldots & v_{nm} \end{pmatrix} \in \mathbb{K}^{n \times m}.$$

Die Matrix A ist eine $n \times m$-Matrix mit n Zeilen und m Spalten.

Die Matrix A bringen wir in eine Form, an der wir die lineare Unabhängigkeit der Zeilen ablesen können, mittels der folgenden *Zeilenumformungen* für eine gegebene $n \times m$-Matrix B mit Zeilen $w_1, w_2, \ldots, w_n \in \mathbb{K}^m$:

(Z1) Ersetze eine Zeile w_j durch αw_j, wobei $\alpha \in \mathbb{K} \setminus \{0\}$;

(Z2) Ersetze eine Zeile w_j durch $w_j + \beta w_k$, wobei $\beta \in \mathbb{K}$ und $k \neq j$;

(Z3) Vertausche die Zeilen w_j und w_k, wobei $j \neq k$.

Bemerkung: Die Zeilen von B sind genau dann linear unabhängig, wenn die Zeilen der umgeformten Matrix linear unabhängig sind.

Satz: Jede Matrix $A \in \mathbb{K}^{n\times m}$ kann man durch endlich viele Zeilenumformungen in eine Matrix $C \in \mathbb{K}^{n\times m}$ überführen, die in *Zeilenstufenform* (*ZSF*) ist. Dabei heißt

$$C = \begin{pmatrix} c_{11} & c_{12} & \ldots & c_{1m} \\ c_{21} & c_{22} & \ldots & c_{2m} \\ \ldots & \ldots & \ldots & \ldots \\ c_{n1} & c_{n2} & \ldots & c_{nm} \end{pmatrix}$$

in *Zeilenstufenform*, falls es ein $r \in \{0, 1, \ldots, n\}$ und $1 \leq k_1 < k_2 < \ldots < k_r \leq m$ gibt mit

für $j = 1, 2, \ldots, r$ gilt $c_{jk} = 0$ für $k = 1, 2, \ldots, k_j - 1$ und $\gamma_j := c_{jk_j} \neq 0$;

für $j = r + 1, \ldots, n$ gilt $c_{jk} = 0$ für alle $k = 1, 2, \ldots, m$.

Man braucht dafür sogar nur Umformungen der Art (Z2) und (Z3).

Beispiel für Zeilenstufenform mit $n = 5$ und $m = 8$ und $r = 4$, hierbei seien $\gamma_1, \gamma_2, \gamma_3, \gamma_4 \neq 0$ und $*$ ist ein beliebiger Eintrag:

$$\begin{pmatrix} 0 & \gamma_1 & * & * & * & * & * & * \\ 0 & 0 & 0 & \gamma_2 & * & * & * & * \\ 0 & 0 & 0 & 0 & \gamma_3 & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & \gamma_4 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Zum Beweis des Satzes betrachtet man zunächst die erste Spalte von C.

1) Ist $c_{11} \neq 0$, so setzt man $\gamma_1 := c_{11}$ und zieht für $j = 2, \ldots, n$ von der j-ten Zeile das c_{j1}/γ_1-fache der ersten Zeile ab (dh man wendet $(Z3)$ an). Man erhält so eine neue

Matrix $\tilde{C}$ mit Nullen in der ersten Spalte unterhalb von γ_1:

$$\tilde{C} = \begin{pmatrix} \gamma_1 & c_{12} & \cdots & c_{1m} \\ 0 & \tilde{c}_{22} & \cdots & \tilde{c}_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \tilde{c}_{n2} & \cdots & \tilde{c}_{nm} \end{pmatrix},$$

und macht mit der Matrix weiter, die man durch “Streichen” der ersten Zeile und ersten Spalte von $\tilde{C}$ erhält.

2) Ist $c_{11} = 0$ und $c_{j1} \neq 0$ für ein $j \in \{2, \ldots, n\}$, so vertauscht man die erste und die j-te Zeile (dh man wendet $(Z2)$ an), und mache dann bei 1) weiter.

3) Ist $c_{j1} = 0$ für alle $j = 1, 2, \ldots, n$, so wendet man das Verfahren an auf die Matrix, die durch “Streichen” der ersten Spalte von C entsteht.

Das Verfahren wird nun solange wiederholt, bis Zeilenstufenform erreicht ist. “Streichen” soll hier nur bedeuten, dass die entsprechenden Zeilen bzw. Spalten nicht mehr verändert werden.

Folgerung aus dem Satz: Die n Zeilen der Matrix A sind genau dann linear unabhängig, wenn für eine zugehörige Matrix in Zeilenstufenform $r = n$ gilt, dh also genau dann, wenn in der Zeilenstufenform keine Nullzeilen auftreten. Dies kann wegen $r \leq m$ höchstens dann sein, wenn $n \leq m$ ist. Das bedeutet, dass $m+1$ oder mehr Vektoren im $\mathbb{K}$-Vektorraum $\mathbb{K}^m$ immer linear abhängig sind.

Beispiel: Wir untersuchen die Vektoren $v_1 = (1, 3, -2, 4)$, $v_2 = (-1, -1, 5, -9)$, $v_3 = (2, 0, -13, 23)$, $v_4 = (1, 5, 1, -2) \in \mathbb{R}^4$ und erhalten nacheinander die Matrizen

$$\begin{pmatrix} 1 & 3 & -2 & 4 \\ -1 & -1 & 5 & -9 \\ 2 & 0 & -13 & 23 \\ 1 & 5 & 1 & -2 \end{pmatrix} \quad \begin{pmatrix} 1 & 3 & -2 & 4 \\ 0 & 2 & 3 & -5 \\ 0 & -6 & -9 & 15 \\ 0 & 2 & 3 & -6 \end{pmatrix} \quad \begin{pmatrix} 1 & 3 & -2 & 4 \\ 0 & 2 & 3 & -5 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \quad \begin{pmatrix} 1 & 3 & -2 & 4 \\ 0 & 2 & 3 & -5 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Die Vektoren v_1, v_2, v_3, v_4 sind nicht linear unabhängig.

Bemerkung: Man sieht aber auch, dass Zeilenumformungen den linearen Aufspann der Zeilen einer Matrix nicht ändern. Ein Blick auf die dritte Matrix zeigt, dass v_3 eine Linearkombination der Vektoren v_1, v_2 ist (es wurden bis dahin nur Umformungen der Art (Z2) vorgenommen, dh es wurden keine Zeilen vertauscht!): Die erste Zeile ist immer v_1, die zweite Zeile ist beim ersten Mal v_2 und danach immer $v_2 + v_1$, die dritte Zeile ist erst v_3, dann $v_3 - 2v_1$ und dann $(v_3 - 2v_1) + 3(v_2 + v_1) = v_3 + v_1 + 3v_2 = 0$, also gilt $v_3 = -v_1 - 3v_2$. Wir haben deshalb

$$\operatorname{lin}\{v_1, v_2, v_3\} = \operatorname{lin}\{v_1, v_2\},$$

und weiter also auch

$$\operatorname{lin}\{v_1, v_2, v_3, v_4\} = \operatorname{lin}\{v_1, v_2, v_4\}.$$

Außerdem sieht man, dass v_1, v_2, v_4 linear unabhängig sind.

Beispiel: Wir wenden das Verfahren an auf die ersten vier Zeilen der Matrix A aus 15.1:

$$\begin{pmatrix} 1 & 0 & -1 & -1 & 0 & 0 \\ 0 & -1 & 1 & 0 & 1 & 0 \\ -1 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & -1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 & -1 & -1 & 0 & 0 \\ 0 & -1 & 1 & 0 & 1 & 0 \\ 0 & 1 & -1 & -1 & 0 & 1 \\ 0 & 0 & 0 & 1 & -1 & -1 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 0 & -1 & -1 & 0 & 0 \\ 0 & -1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1 & 1 \\ 0 & 0 & 0 & 1 & -1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 & -1 & -1 & 0 & 0 \\ 0 & -1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Es wurden hierbei keine Zeilen vertauscht. Die ersten drei Zeilen sind also linear unabhängig, alle vier Zeilen sind linear abhängig. Die vierte liegt somit im linearen Aufspann der ersten drei Zeilen.

15.8. Zeilennormalform: Eine Matrix, die in Zeilenstufenform ist, kann man durch Zeilenumformungen auf *Zeilennormalform* (*ZNF*) bringen. Dabei heißt $C = (c_{jk}) \in \mathbb{K}^{n\times m}$ in *Zeilennormalform*, falls zusätzlich (mit den Bezeichnungen von 15.7) gilt:

> für $j = 1, 2, \ldots, r$ gilt: $\gamma_j = c_{jk_j} = 1$ und $c_{lk_j} = 0$ für $l = 1, 2, \ldots, j-1$ (dh oberhalb von c_{jk_j} stehen auch nur Nullen).

Beispiel mit $n = 5$ und $m = 8$, $r = 4$ aus 15.7:

$$\begin{pmatrix} 0 & 1 & * & 0 & 0 & * & 0 & * \\ 0 & 0 & 0 & 1 & 0 & * & 0 & * \\ 0 & 0 & 0 & 0 & 1 & * & 0 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Beispiel: Für die konkrete Matrix im Beispiel in 15.7 formen wir weiter um: Zunächst bringen wir γ_j auf 1 und gehen dann von rechts nach links vor:

$$\begin{pmatrix} 1 & 3 & -2 & 4 \\ 0 & 2 & 3 & -5 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 3 & -2 & 4 \\ 0 & 1 & 3/2 & -5/2 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & -2 & 0 \\ 0 & 1 & 3/2 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & -13/2 & 0 \\ 0 & 1 & 3/2 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

15.9. Basen und Dimension: Sei V ein $\mathbb{K}$-Vektorraum.

Definition: Sei $n \in \mathbb{N}$. Die Vektoren $b_1, b_2, \ldots, b_n \in V$ heißen *(geordnete) Basis von V*, falls sie linear unabhängig sind und $\operatorname{lin}\{b_1, b_2, \ldots, b_n\} = V$ gilt. V heißt *endlich-dimensional*, falls V in diesem Sinne eine Basis hat, und V heißt *unendlich-dimensional* sonst.

Satz und Definition: Ist V endlich-dimensional, so enthalten je zwei Basen von V die gleiche Anzahl von Elementen. Diese Zahl heißt die *Dimension von V*, geschrieben $\dim V$.

Beispiele: Es gilt $\dim \mathbb{K}^n = n$ und eine Basis ist gegeben durch die *Einheitsvektoren* $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$.

Der $\mathbb{K}$-Vektorraum $\mathbb{K}^{n \times m}$ hat die Dimension $n \cdot m$. Eine Basis ist gegeben durch die Matrizen

$$B_{pq} := (\delta_{jp}\delta_{kq})_{j=1}^{n}{}_{k=1}^{m}, \qquad p = 1, \ldots, n,\ q = 1, \ldots, m.$$

Die Matrix B_{pq} hat gerade an der Stelle (p, q) eine Eins und sonst Nullen.

Bemerkung: V ist unendlich-dimensional genau dann, wenn es zu jedem $n \in \mathbb{N}$ linear unabhängige Vektoren $v_1, v_2, \ldots, v_n \in V$ gibt.

Beispiel: Die Menge $P[a, b]$ der Polynomfunktionen $p : [a, b] \to \mathbb{R}$ ist ein unendlich-dimensionaler $\mathbb{R}$-Vektorraum: Da Summen und Produkte von Polynomen wieder Polynome sind, ist $P[a, b]$ ein Untervektorraum von $\mathbb{R}^{[a,b]}$. Für jedes $k \in \mathbb{N}_0$ sei p_k gegeben durch $p_k(x) := x^k$. Dann sind für jedes $n \in \mathbb{N}$ die Polynome $p_0, p_1, \ldots, p_n$ linear unabhängig, denn für $\alpha_0, \alpha_1, \ldots, \alpha_n \in \mathbb{R}$ mit

$$\alpha_0 p_0 + \alpha_1 p_1 + \ldots + \alpha_n p_n = 0 \quad \text{auf } [a, b]$$

folgt $\alpha_0 = \alpha_1 = \ldots = \alpha_n = 0$, da nur das Nullpolynom unendlich viele Nullstellen hat (vgl. 5.4).

Bemerkung: Ist $n \in \mathbb{N}$ und $b_1, b_2, \ldots, b_n$ eine Basis von V, so gibt es zu jedem $v \in V$ eindeutig bestimmte $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{K}$ mit

$$v = \alpha_1 b_1 + \alpha_2 b_2 + \ldots + \alpha_n b_n = \sum_{j=1}^{n} \alpha_j b_j,$$

die $\alpha_1, \alpha_2, \ldots, \alpha_n$ heißen *Koordinaten* von v bzgl. der Basis $b_1, b_2, \ldots, b_n$.

Beweis. Existenz folgt aus $v \in V = \operatorname{lin}\{b_1, b_2, \ldots, b_n\}$. Sind $\alpha_j, \tilde{\alpha}_j \in \mathbb{K}$ für $j = 1, \ldots, n$ mit $\sum_{j=1}^{n} \alpha_j b_j = v = \sum_{j=1}^{n} \tilde{\alpha}_j b_j$, so folgt

$$0 = v - v = \sum_{j=1}^{n} (\alpha_j - \tilde{\alpha}_j) b_j$$

und weiter $\alpha_j = \tilde{\alpha}_j$ für alle $j = 1, \ldots, n$, da $b_1, b_2, \ldots, b_n$ linear unabhängig sind. $\square$

Satz: Sei V ein $\mathbb{K}$-Vektorraum und seien $v_1, v_2, \ldots, v_n \in V$. Dann ist die Dimension $\dim \operatorname{lin}\{v_1, v_2, \ldots, v_n\}$ des linearen Aufspanns von $v_1, v_2, \ldots, v_n$ gerade die Maximalzahl linear unabhängiger Vektoren unter den $v_1, v_2, \ldots, v_n$.

Beispiel: Im $\mathbb{R}^3$ sei $\vec{v}_1 := (1,1,0)$, $\vec{v}_2 := (1,0,0)$ und $\vec{v}_3 := (0,1,0)$. Dann sind $\vec{v}_1, \vec{v}_2$ linear unabhängig, aber $\vec{v}_3 \in \operatorname{lin}\{\vec{v}_1, \vec{v}_2\}$ und $\vec{v}_1, \vec{v}_2$ ist Basis von $\operatorname{lin}\{\vec{v}_1, \vec{v}_2, \vec{v}_3\} = \operatorname{lin}\{\vec{v}_1, \vec{v}_2\}$. Es ist $\dim \operatorname{lin}\{\vec{v}_1, \vec{v}_2, \vec{v}_3\} = 2$. Auch $\vec{v}_1, \vec{v}_3$ bzw. $\vec{v}_2, \vec{v}_3$ sind hier Basen von $\operatorname{lin}\{\vec{v}_1, \vec{v}_2, \vec{v}_3\}$.

Warnung: Für einen $\mathbb{C}$-Vektorraum V ist es für die Dimensionsbestimmung wichtig, ob man ihn als $\mathbb{C}$-Vektorraum oder als $\mathbb{R}$-Vektorraum betrachtet, dh ob man Linearkombinationen mit komplexen oder nur mit reellen Koeffizienten zulässt. So gilt $\mathbb{C}\text{-}\dim \mathbb{C} = 1$, aber $\mathbb{R}\text{-}\dim \mathbb{C} = 2$ (wie schon an der Veranschaulichung von $\mathbb{C}$ als komplexe Zahlen*ebene* zu sehen ist). Eine Basis des $\mathbb{R}$-Vektorraumes $\mathbb{C}$ ist gegeben durch $1, i$.

Ebenso ist die Dimension von $\mathbb{C}^n$ als $\mathbb{C}$-Vektorraum $= n$ und eine Basis ist $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$. Die Dimension von $\mathbb{C}^n$ als $\mathbb{R}$-Vektorraum ist hingegen $= 2n$ und eine Basis ist z.B. $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n, i\vec{e}_1, i\vec{e}_2, \ldots, i\vec{e}_n$.

15.10. Die Lösungsmenge eines linearen Gleichungssystems: Seien $A \in \mathbb{K}^{n\times m}$ und $\vec{b} \in \mathbb{K}^n$ gegeben. Bisher haben wir eingesehen:

(1) $\operatorname{Kern} A$ ist ein Untervektorraum von $\mathbb{K}^m$ und $\operatorname{Bild} A$ ist ein Untervektorraum von $\mathbb{K}^n$.

(2) $\operatorname{Bild} A$ ist der lineare Aufspann der Spalten von A.

(3) Hat die Gleichung $A\vec{x} = \vec{b}$ eine Lösung $\vec{z}$, so ist die Lösungsmenge der Gleichung der affine Teilraum $\vec{z} + \operatorname{Kern} A$. Insbesondere ist die Lösung eindeutig genau dann, wenn $\operatorname{Kern} A = \{\vec{0}\}$ gilt.

(4) Es gilt

$$A\vec{x} = \vec{b} \text{ ist lösbar} \iff \vec{b} \in \operatorname{Bild} A \iff \vec{b} \text{ ist LK der Spalten von } A.$$

Beispiele: (1) $A = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$, $\vec{b} = \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right)$. Dann hat $A\vec{x} = \vec{b}$ keine Lösung.

(2) Sei A wie eben und $\vec{b} = \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$. Dann ist $A\vec{x} = \vec{b}$ lösbar und die Menge aller Lösungen ist der affine Teilraum

$$\left\{ \begin{pmatrix} \alpha \\ 1 \end{pmatrix} : \alpha \in \mathbb{K} \right\} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \underbrace{\operatorname{lin}\begin{pmatrix} 1 \\ 0 \end{pmatrix}}_{= \operatorname{Kern}\ A} .$$

(3) Sei $A = \left(\begin{smallmatrix} 2 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ und $\vec{b} = \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right)$. Dann ist $\vec{x} := \left(\begin{smallmatrix} -1/2 \\ 1 \end{smallmatrix}\right)$ die eindeutige Lösung von $A\vec{x} = \vec{b}$.

Bemerkung: Ein lineares Gleichungssystem kann also keine, genau eine oder unendlich viele Lösungen haben.

15.11. Lösungsalgorithmus: Seien $A \in \mathbb{K}^{n\times m}$ und $\vec{b} \in \mathbb{K}^n$ gegeben. Zur Lösung erweitern wir die Matrix A um $\vec{b}$ als $(m+1)$-te Spalte, betrachten also $(A\,|\,\vec{b}) \in \mathbb{K}^{n\times(m+1)}$. Für $A = (a_{jk})$ und $\vec{b} = (b_j)$ ist

$$(A\,|\,\vec{b}) = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1m} & b_1 \\ a_{21} & a_{22} & \dots & a_{2m} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} & b_n \end{array}\right)$$

Wir werden Zeilenumformungen verwenden (vgl. 15.7).

Beobachtung: Geht die Matrix $(\tilde{B}\,|\,\vec{\tilde{c}})$ aus der Matrix $(B\,|\,\vec{c})$ durch eine Zeilenumformung hervor, so gilt

$$\{\vec{x} \in \mathbb{K}^m : \tilde{B}\vec{x} = \vec{\tilde{c}}\} = \{\vec{x} \in \mathbb{K}^m : B\vec{x} = \vec{c}\},$$

dh die Lösungsmenge der entsprechenden linearen Gleichungssysteme ändert sich nicht.

Allgemein gilt: Geht $\tilde{B}$ aus B durch eine Zeilenumformung hervor und ist die k-te Spalte von B Linearkombination anderer Spalten von B, so ist die k-te Spalte von $\tilde{B}$ Linearkombination der entsprechenden Spalten von $\tilde{B}$ und zwar *mit denselben Koeffizienten*.

Algorithmus (Eliminationsverfahren nach Gauß):

(1) Matrix A um $\vec{b}$ als letzte Spalte erweitern.
(2) Die erweiterte Matrix $(A\,|\,\vec{b})$ durch Zeilenumformungen auf ZNF (oder ZSF) bringen.
(3) Lösbarkeit und Lösung ablesen.

Zu Schritt (2) vergleiche 15.7 und 15.8.

Beispiel: Wir rechnen das Beispiel 15.1 mit den folgenden Werten:

$$R_1 = 5\,\Omega,\; R_2 = R_6 = 10\,\Omega,\; R_3 = R_4 = 20\,\Omega,\; R_5 = 40\,\Omega,\; U = 40\,V.$$

Dabei lassen wir die dritte und die vierte Zeile weg, da diese LK der ersten drei Zeilen bzw. der letzten drei Zeilen sind und also die Lösungsmenge nicht geändert wird.

$$\left(\begin{array}{cccccc|c} 1 & 0 & -1 & -1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & -1 & -1 & 0 \\ R_1 & R_2 & R_3 & 0 & 0 & 0 & U \\ 0 & 0 & R_3 & -R_4 & -R_5 & 0 & 0 \\ 0 & R_2 & 0 & 0 & R_5 & -R_6 & 0 \\ R_1 & 0 & 0 & R_4 & 0 & R_6 & U \end{array}\right) \quad \to \quad \left(\begin{array}{cccccc|c} \underline{1} & 0 & -1 & -1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 20 & -20 & -40 & 0 & 0 \\ 0 & 10 & 0 & 0 & 40 & -10 & 0 \\ 5 & 0 & 0 & 20 & 0 & 10 & 40 \end{array}\right)$$

Wir addieren Z_1 zu Z_3 und $(-5) \cdot Z_1$ zu Z_6 und erhalten:

$$\left(\begin{array}{cccccc|c} 1 & 0 & -1 & -1 & 0 & 0 & 0 \\ 0 & \underline{-1} & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & -1 & -1 & 0 & 1 & 0 \\ 0 & 0 & 20 & -20 & -40 & 0 & 0 \\ 0 & 10 & 0 & 0 & 40 & -10 & 0 \\ 0 & 0 & 5 & 25 & 0 & 10 & 40 \end{array}\right)$$

Wir addieren Z_2 zu Z_3 und $10 \cdot Z_2$ zu Z_5 und erhalten:

$$\left(\begin{array}{cccccc|c} 1 & 0 & -1 & -1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 1 & 1 & 0 \\ 0 & 0 & 20 & -20 & -40 & 0 & 0 \\ 0 & 0 & 10 & 0 & 50 & -10 & 0 \\ 0 & 0 & \underline{5} & 25 & 0 & 10 & 40 \end{array}\right)$$

Wir addieren $(-2) \cdot Z_6$ zu Z_5 und $(-4) \cdot Z_6$ zu Z_4, anschließend schieben wir die sechste Zeile hinter die zweite:

$$\left(\begin{array}{cccccc|c} 1 & 0 & -1 & -1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 5 & 25 & 0 & 10 & 40 \\ 0 & 0 & 0 & \underline{-1} & 1 & 1 & 0 \\ 0 & 0 & 0 & -120 & -40 & -40 & -160 \\ 0 & 0 & 0 & -50 & 50 & -30 & -80 \end{array}\right)$$

Wir addieren $(-120) \cdot Z_4$ zu Z_5 und $(-50) \cdot Z_4$ zu Z_6 und erhalten:

$$\left(\begin{array}{cccccc|c} 1 & 0 & -1 & -1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 5 & 25 & 0 & 10 & 40 \\ 0 & 0 & 0 & -1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & -160 & -160 & -160 \\ 0 & 0 & 0 & 0 & 0 & -80 & -80 \end{array}\right)$$

Diese Matrix ist in ZSF, das LGS ist eindeutig lösbar, und wir können von unten nach oben nacheinander ablesen bzw. berechnen:

$$I_6 = 1\,A,\ I_5 = 0\,A,\ I_4 = 1\,A,\ I_3 = 1\,A,\ I_2 = 1\,A,\ I_1 = 2\,A.$$

Algorithmus Schritt (3), Ablesen der Lösbarkeit: Das lineare Gleichungssystem $A\vec{x} = \vec{b}$ ist **nicht lösbar** genau dann, wenn eine ZSF von $(A\,|\,\vec{b})$ die Form $C = (c_{jk}) \in \mathbb{K}^{n\times(m+1)}$ hat und es ein $j \in \{1, \ldots, n\}$ gibt mit $c_{jk} = 0$ für $k = 1, \ldots, m$ und $c_{j,m+1} \neq 0$.

Mit den Bezeichnungen aus 15.7 ist dies genau dann der Fall, wenn $k_r = m+1$ gilt.

Beispiel: siehe Beispiel 15.10(1). Dort ist $n = m = 2$, $r = 2$, $k_1 = 2$, $k_2 = 3$.

Algorithmus Schritt (3), Ablesen der Lösung im allgemeinen Fall: Man bringe die Matrix auf Zeilennormalform. Wir zeigen am Beispiel, wie man im Falle der Lösbarkeit die Lösungen ablesen kann. Dazu nehmen wir an, dass die berechnete Zeilennormalform von $(A \,|\, \vec{b})$ die folgende Gestalt hat (hier ist $n = 4$, $m = 5$):

$$\left(\begin{array}{ccccc|c} 1 & 0 & \alpha_1 & 0 & \beta_1 & c_1 \\ 0 & 1 & \alpha_2 & 0 & \beta_2 & c_2 \\ 0 & 0 & 0 & 1 & \beta_3 & c_3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Man nehme die Variablen zu den Spalten, die "hinter" den Stufen stehen (im Beispiel die dritte und die fünfte Spalte), als *freie Parameter* (im Beispiel also etwa $s = x_3$ und $t = x_5$). Schreibt man die Gleichungen wieder aus, so erhält man

$$x_1 = c_1 - s\alpha_1 - t\beta_1, \quad x_2 = c_2 - s\alpha_2 - t\beta_2, \quad x_4 = c_3 - t\beta_3.$$

Also ist die Lösungsmenge von $Ax = b$ gegeben durch

$$\left\{ \begin{pmatrix} c_1 \\ c_2 \\ 0 \\ c_3 \\ 0 \end{pmatrix} + s \begin{pmatrix} -\alpha_1 \\ -\alpha_2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + t \begin{pmatrix} -\beta_1 \\ -\beta_2 \\ 0 \\ -\beta_3 \\ 1 \end{pmatrix} : s, t \in \mathbb{K} \right\}.$$

Hier gilt $\operatorname{Kern} A = \operatorname{lin}\{ \begin{pmatrix} -\alpha_1 \\ -\alpha_2 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -\beta_1 \\ -\beta_2 \\ 0 \\ -\beta_3 \\ 1 \end{pmatrix} \}$ und $\dim(\operatorname{Kern} A) = 2$. Die Lösungsmenge ist eine Ebene durch den Punkt $\begin{pmatrix} c_1 \\ c_2 \\ 0 \\ c_3 \\ 0 \end{pmatrix}$.

Zum selben Ergebnis gelangt man mit dem (-1)**-Ergänzungstrick:**

Ausgehend von der Zeilennormalform lasse man zunächst die Nullzeilen weg. Dann ergänze man unter den Zeilen mit "längeren" Stufen eine Zeile mit -1 und sonst Nullen (ggf.

mehrere solcher Zeilen s.u.) so, dass auf der Diagonalen nur ± 1 steht. Im Beispiel

$$\left(\begin{array}{ccccc|c} 1 & 0 & \alpha_1 & 0 & \beta_1 & c_1 \\ 0 & 1 & \alpha_2 & 0 & \beta_2 & c_2 \\ 0 & 0 & 0 & 1 & \beta_3 & c_3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right) \quad \left(\begin{array}{ccccc|c} 1 & 0 & \alpha_1 & 0 & \beta_1 & c_1 \\ 0 & 1 & \alpha_2 & 0 & \beta_2 & c_2 \\ 0 & 0 & 0 & 1 & \beta_3 & c_3 \end{array}\right) \quad \left(\begin{array}{ccccc|c} 1 & 0 & \alpha_1 & 0 & \beta_1 & c_1 \\ 0 & 1 & \alpha_2 & 0 & \beta_2 & c_2 \\ 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & \beta_3 & c_3 \\ 0 & 0 & 0 & 0 & -1 & 0 \end{array}\right).$$

Dann nimmt man für jede (-1)-Spalte einen freien Parameter und kann die Lösungsmenge hinschreiben: letzte Spalte plus jeweils freier Parameter mal entsprechender Spalte, also

$$\left\{ \begin{pmatrix} c_1 \\ c_2 \\ 0 \\ c_3 \\ 0 \end{pmatrix} + s \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ -1 \\ 0 \\ 0 \end{pmatrix} + t \begin{pmatrix} \beta_1 \\ \beta_2 \\ 0 \\ \beta_3 \\ -1 \end{pmatrix} : s, t \in \mathbb{K} \right\}.$$

Man vergleiche dies mit der Darstellung oben.

Weiteres Beispiel zum (-1)-Ergänzungstrick:

$$\left(\begin{array}{cccc|c} 1 & 0 & \alpha_1 & \beta_1 & c_1 \\ 0 & 1 & \alpha_2 & \beta_2 & c_2 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right) \quad \left(\begin{array}{cccc|c} 1 & 0 & \alpha_1 & \beta_1 & c_1 \\ 0 & 1 & \alpha_2 & \beta_2 & c_2 \end{array}\right) \quad \left(\begin{array}{cccc|c} 1 & 0 & \alpha_1 & \beta_1 & c_1 \\ 0 & 1 & \alpha_2 & \beta_2 & c_2 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \end{array}\right).$$

Die Lösungsmenge ist hier also gegeben durch

$$\left\{ \begin{pmatrix} c_1 \\ c_2 \\ 0 \\ 0 \end{pmatrix} + s \begin{pmatrix} \alpha_1 \\ \alpha_2 \\ -1 \\ 0 \end{pmatrix} + t \begin{pmatrix} \beta_1 \\ \beta_2 \\ 0 \\ -1 \end{pmatrix} : s, t \in \mathbb{K} \right\}.$$

Bemerkung: Enthält die ZNF links Nullspalten, so ergänze man für diese die (-1)-Zeilen *oben*. Im Beispiel 15.3(2) also

$$\left(\begin{array}{cc|c} 0 & 1 & 1 \\ 0 & 0 & 0 \end{array}\right) \quad \left(\begin{array}{cc|c} 0 & 1 & 1 \end{array}\right) \quad \left(\begin{array}{cc|c} -1 & 0 & 0 \\ 0 & 1 & 1 \end{array}\right)$$

mit der abgelesenen Lösungsmenge

$$\left\{ \begin{pmatrix} 0 \\ 1 \end{pmatrix} + s \begin{pmatrix} -1 \\ 0 \end{pmatrix} : s \in \mathbb{K} \right\}.$$

15.12. Der Rang einer Matrix: Für eine Matrix $A \in \mathbb{K}^{n\times m}$ ist die Maximalzahl r linear unabhängiger Zeilen gleich der Maximalzahl linear unabhängiger Spalten (man sagt auch "Zeilenrang = Spaltenrang"). Die Zahl r heißt *Rang der Matrix A*, geschrieben $\operatorname{Rang} A$.

Beweis. Man betrachte die ZNF von A. $\square$

Bemerkung: Es gilt $\operatorname{Rang} A = \dim \operatorname{Bild} A$.

Satz: Sind $A \in \mathbb{K}^{n\times m}$ und $\vec{b} \in \mathbb{K}^n$, so gilt:

$$A\vec{x} = \vec{b} \text{ ist lösbar} \quad \Longleftrightarrow \quad \operatorname{Rang} A = \operatorname{Rang}(A \,|\, \vec{b}),$$

dh ein lineares Gleichungssystem ist lösbar genau dann, wenn der Rang der Matrix A gleich dem Rang der erweiterten Matrix $(A \,|\, \vec{b})$ ist.

15.13. Dimensionsformel: Sei $A \in \mathbb{K}^{n\times m}$. Dann gilt

$$m = \dim \operatorname{Bild} A + \dim \operatorname{Kern} A = \operatorname{Rang} A + \dim \operatorname{Kern} A.$$

Man vergleiche mit 15.11 (insbesondere mit der Form der Matrix, die beim (-1)-Erweiterungstrick entsteht).

Beispiele: (1) $A = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix}$. Hier ist $n = m = 3$ und $\operatorname{Rang} A = 3$, da die Zeilen (oder auch die Spalten) linear unabhängig sind. Somit ist $\dim(\operatorname{Kern} A) = 3 - 3 = 0$ und $\operatorname{Kern} A = \{\vec{0}\}$. Ablesen lässt sich das natürlich auch an einer ZSF:

$$\begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 1 & 0 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}.$$

(2) $A = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$. Hier ist $n = m = 3$ und $\operatorname{Rang} A = 2$, da die ersten beiden Spalten linear unabhängig sind, die dritte Spalte jedoch Summe der ersten beiden. Somit ist $\dim(\operatorname{Kern} A) = 3 - 2 = 1$. Durch Probieren findet man $\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} \in \operatorname{Kern} A$. Folglich ist $\operatorname{Kern} A = \operatorname{lin}\{\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}\}$. Zum selben Ergebnis kommt man mit einer ZNF und anschließender (-1)-Ergänzung:

$$\begin{pmatrix} 1 & 1 & 2 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 & 1 \\ 0 & -1 & -1 \\ 0 & 1 & 1 \end{pmatrix} \quad \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix} \quad \xrightarrow{(-1)} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & -1 \end{pmatrix}.$$

15.14. Lineare Abbildungen als Matrizen: Seien V, W endlich-dimensionale $\mathbb{K}$-Vektorräume und $\phi : V \to W$ eine lineare Abbildung. In V und W seien Basen $b_1, b_2, \ldots, b_m$ bzw. $c_1, c_2, \ldots, c_n$ gegeben. Dann gibt es genau eine Matrix $A \in \mathbb{K}^{n\times m}$ mit der Eigenschaft

$$\phi\Big(\sum_{k=1}^{m} \alpha_k b_k\Big) = \sum_{j=1}^{n} \beta_j c_j \iff A \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{pmatrix} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}.$$

Durch die Matrix A werden also die Koordinatentupel $(\alpha_1, \alpha_2, \ldots, \alpha_m) \in \mathbb{K}^m$ von Vektoren $v \in V$ bzgl. der Basis $b_1, b_2, \ldots, b_m$ auf die Koordinatentupel $(\beta_1, \beta_2, \ldots, \beta_n) \in \mathbb{K}^n$ von $\phi(v)$ bzgl. der Basis $c_1, c_2, \ldots, c_n$ abgebildet.

A heißt dann *Darstellung von ϕ bzgl. der Basen $b_1, b_2 \ldots, b_m$ und $c_1, c_2, \ldots, c_n$.*

Man erhält die Matrix A wie folgt:

> Zu jedem Basisvektor b_k, $k = 1, \ldots, m$ bestimme man die Koordinaten von $\phi(b_k) \in W$ bzgl. der Basis $c_1, c_2, \ldots, c_n$. Diese bilden die k-te Spalte von A.

Beispiele: (1) Sei $A \in \mathbb{K}^{n\times m}$. Die Darstellungsmatrix von $\phi_A : \mathbb{K}^m \to \mathbb{K}^n$, $\phi_A(x) := Ax$, bzgl. der Standardbasen in $\mathbb{K}^m$ bzw. $\mathbb{K}^n$ ist A.

(2) Die lineare Abbildung $\phi : \mathbb{K}^3 \to \mathbb{K}^3$ sei gegeben durch

$$\phi(x_1, x_2, x_3) = (2x_1 - x_2 + x_3, -x_1 + 2x_2 + x_3, x_1 - x_2 + 2x_3).$$

Wir bestimmen die Matrix A zu ϕ bzgl. der Basis $\vec{b}_1, \vec{b}_2, \vec{b}_3$ im Argument- und Zielraum, wobei

$$\vec{b}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \vec{b}_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad \vec{b}_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}.$$

Es ist

$$\begin{aligned}
\phi(\vec{b}_1) &= \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = 1 \cdot \vec{b}_1 + 0 \cdot \vec{b}_2 + 0 \cdot \vec{b}_3, \\
\phi(\vec{b}_2) &= \begin{pmatrix} 3 \\ 0 \\ 3 \end{pmatrix} = 0 \cdot \vec{b}_1 + 3 \cdot \vec{b}_2 + 0 \cdot \vec{b}_3, \\
\phi(\vec{b}_3) &= \begin{pmatrix} 0 \\ 3 \\ 1 \end{pmatrix} = 1 \cdot \vec{b}_1 + (-1) \cdot \vec{b}_2 + 2 \cdot \vec{b}_3,
\end{aligned}$$

also
$$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 3 & -1 \\ 0 & 0 & 2 \end{pmatrix}.$$
Verwenden wir hingegen $\vec{b}_1, \vec{b}_2, \vec{b}_3$ nur "vorne" (dh im Argumentraum von ϕ) und die Standardbasis "hinten" (dh im Zielraum von ϕ), so erhalten wir als Darstellungsmatrix
$$\begin{pmatrix} 1 & 3 & 0 \\ 1 & 0 & 3 \\ 0 & 3 & 1 \end{pmatrix}.$$
Verwenden wir in Argument- und Zielraum die Standardbasis, so erhalten wir als Darstellungsmatrix
$$\begin{pmatrix} 2 & -1 & 1 \\ -1 & 2 & 1 \\ 1 & -1 & 2 \end{pmatrix}.$$
(3) Ein besonderer Fall ist die Darstellung der Identität $\mathrm{Id} : \mathbb{K}^n \to \mathbb{K}^n$, $\vec{x} \mapsto \vec{x}$, wenn man in Argument- und Zielraum verschiedene Basen verwendet. Nehmen wir im Argumentraum die Basis $\vec{b}_1, \vec{b}_2, \vec{b}_3$ aus Beispiel (2) und im Zielraum die Standardbasis $\vec{e}_1, \vec{e}_2, \vec{e}_3$, so ist die k-te Spalte der Darstellungsmatrix gerade $\vec{b}_k$, dh die Darstellungsmatrix ist
$$B := \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}.$$
Nehmen wir hingegen im Argumentraum die Standardbasis und im Zielraum die Basis $\vec{b}_1, \vec{b}_2, \vec{b}_3$ aus Beispiel (2), so ist wegen
$$\vec{e}_1 = \frac{1}{2}(\vec{b}_1 + \vec{b}_2 - \vec{b}_3), \quad \vec{e}_2 = \frac{1}{2}(\vec{b}_1 - \vec{b}_2 + \vec{b}_3), \quad \vec{e}_3 = \frac{1}{2}(-\vec{b}_1 + \vec{b}_2 + \vec{b}_3)$$
die Darstellungsmatrix gegeben durch
$$C := \frac{1}{2} \begin{pmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \\ -1 & 1 & 1 \end{pmatrix}.$$

15.15. Das Produkt von Matrizen: Seien $n, m, q \in \mathbb{N}$ und $A \in \mathbb{K}^{n \times m}$, $B \in \mathbb{K}^{m \times q}$ Matrizen. Dann sind die Abbildungen $\phi_A : \mathbb{K}^m \to \mathbb{K}^n$, $\phi_A(\vec{x}) = A\vec{x}$, und $\phi_B : \mathbb{K}^q \to \mathbb{K}^m$, $\phi_B(\vec{x}) = B\vec{x}$, linear, und $\phi_A \circ \phi_B : \mathbb{K}^q \to \mathbb{K}^n$ ist auch linear. Die Darstellungsmatrix von $\phi_A \circ \phi_B$ bzgl. der Standardbasen ist gegeben durch das *Matrixprodukt* $AB \in \mathbb{K}^{n \times q}$, wobei für $A = (a_{jk})_{j=1}^{n}{}_{k=1}^{m}$ und $B = (b_{kl})_{k=1}^{m}{}_{l=1}^{q}$ die Matrix AB Einträge $(c_{jl})_{j=1}^{n}{}_{l=1}^{q}$ hat mit
$$c_{jl} = \sum_{k=1}^{m} a_{jk} b_{kl}, \qquad j = 1, \ldots, n, \; l = 1, \ldots, q.$$

Für $\vec{x} = (x_1, x_2, \ldots, x_q) \in \mathbb{K}^q$ ist nämlich

$$\begin{aligned} A(B\vec{x}) &= (a_{jk})_{j,k}\Big(\sum_{l=1}^{q} b_{kl}x_l\Big)_k = \Big(\sum_{k=1}^{m} a_{jk}\sum_{l=1}^{q} b_{kl}x_l\Big)_j \\ &= \Big(\sum_{l=1}^{q}\Big(\sum_{k=1}^{m} a_{jk}b_{kl}\Big)x_l\Big)_j = \Big(\sum_{k=1}^{m} a_{jk}b_{kl}\Big)_{j,l}(x_l)_l. \end{aligned}$$

Man beachte, dass hierbei die Anzahl der Spalten von A gleich der Anzahl der Zeilen von B ist (nämlich m), andernfalls ist das Matrixprodukt **nicht definiert**.

Bemerkung: Man erhält die l-te Spalte von AB, indem man die Matrix A mit der l-ten Spalte von B multipliziert (im Sinne des Matrix-Vektor-Produktes aus 15.2 und 15.3).

Beispiel mit $n = m = 2$:

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

Eigenschaften des Produktes von Matrizen: Für alle $A_1, A_2 \in \mathbb{K}^{n\times m}$, $B_1, B_2 \in \mathbb{K}^{m\times q}$ und alle $\alpha \in \mathbb{K}$ gilt:

$$(A_1 + \alpha A_2)B_1 = A_1B_1 + \alpha A_2B_1, \qquad A_1(B_1 + \alpha B_2) = A_1B_1 + \alpha A_1B_2,$$

dh für festes $B \in \mathbb{K}^{m\times q}$ ist

$$\mathbb{K}^{n\times m} \to \mathbb{K}^{n\times q}, \quad A \mapsto AB$$

linear, und für festes $A \in \mathbb{K}^{n\times m}$ ist

$$\mathbb{K}^{m\times q} \to \mathbb{K}^{n\times q}, \quad B \mapsto AB$$

linear.

Außerdem ist das Produkt von Matrizen assoziativ.

Warnung: Im allgemeinen gilt $AB \neq BA$! Damit beide Produkte existieren, muss zunächst $n = q$ sein. Dann ist AB eine $n \times n$-Matrix und BA eine $m \times m$-Matrix. Gleichheit kann also höchstens für $n = m = q$ gelten. Für $n = m = q = 2$ ist aber z.B.

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix} \neq \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

Anwendung Basiswechsel: Im Beispiel 15.14(2) ist die lineare Abbildung $\phi : \mathbb{K}^3 \to \mathbb{K}^3$ bzgl. der Standardbasis gegeben durch die Matrix

$$F := \begin{pmatrix} 2 & -1 & 1 \\ -1 & 2 & 1 \\ 1 & -1 & 2 \end{pmatrix}.$$

Gehen wir im Argumentraum zur Basis $\vec{b}_1, \vec{b}_2, \vec{b}_3$ über, so erhalten wir die entsprechende Darstellungsmatrix für ϕ, wenn wir mit der entsprechenden Darstellungsmatrix B der Identität $\mathrm{Id} : \mathbb{K}^3 \to \mathbb{K}^3$ aus Beispiel 15.14(3) von rechts multiplizieren:

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

man vergleiche mit der Matrix aus 15.14(2).

15.16. Invertierbare Matrizen: Für $n \in \mathbb{N}$ heißt $I_n := (\delta_{jk})_{j,k=1}^n \in \mathbb{K}^{n\times n}$ die *Einheitsmatrix* (wir schreiben auch kurz I, wenn n aus dem Zusammenhang klar ist).

Es gilt dann

$$IA = AI = A \quad \text{für alle } A \in \mathbb{K}^{n\times n}$$

und die zugehörige Abbildung $\phi_I : \mathbb{K}^n \to \mathbb{K}^n$ ist die Identität $\vec{x} \mapsto \vec{x}$.

Definition: Eine Matrix $A \in \mathbb{K}^{n\times n}$ heißt *regulär*, falls $\operatorname{Rang} A = n$ ist, und *invertierbar*, falls es eine Matrix $B \in \mathbb{K}^{n\times n}$ gibt mit $AB = BA = I$.

Bemerkung und Definition: Ist A invertierbar, so ist die Matrix B in obiger Definition eindeutig bestimmt, denn für eine weitere Matrix $\tilde{B}$ mit diesen Eigenschaften folgt:

$$B = BI = B(A\tilde{B}) = (BA)\tilde{B} = I\tilde{B} = \tilde{B}. \qquad (*)$$

Diese Matrix B heißt *Inverse von A* (oder *zu A inverse Matrix*) und wird mit A^{-1} bezeichnet.

Rechenregeln: Seien $A, B \in \mathbb{K}^{n\times n}$ invertierbar. Dann sind A^{-1} und AB invertierbar, und es gilt:

$$(A^{-1})^{-1} = A \qquad \text{und} \qquad (AB)^{-1} = B^{-1}A^{-1}.$$

Die zweite Regel ist die *Hemd-Jacken-Regel* (vgl. 3.6).

Satz: Sei $A \in \mathbb{K}^{n\times n}$. Dann gilt:

$$A \text{ regulär} \iff \operatorname{Bild} A = \mathbb{K}^n \iff \operatorname{Kern} A = \{\vec{0}\} \iff A \text{ invertierbar}.$$

Für die induzierte lineare Abbildung $\phi_A : \mathbb{K}^n \to \mathbb{K}^n$ gilt:

$$\phi_A \text{ surjektiv} \iff \phi_A \text{ injektiv} \iff \phi_A \text{ bijektiv}.$$

Bemerkung: Sind $A, B \in \mathbb{K}^{n\times n}$ mit $AB = I$, so sind A und B invertierbar und es gilt $A^{-1} = B$. So ist etwa $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)^{-1} = \left(\begin{smallmatrix} 1 & -1 \\ 0 & 1 \end{smallmatrix}\right)$ (vgl. Beispiel in 15.15).

Beweis. Gilt $AB = I$, so ist $\operatorname{Kern} B = \{\vec{0}\}$ und $\operatorname{Bild} A = \mathbb{K}^n$, und nach dem Satz sind A, B invertierbar. Für $A^{-1} = B$ verwenden wir ein Argument wie in (*). $\square$

Berechnung von A^{-1}: Man löse für jedes $l \in \{1, 2, \dots, n\}$ das Gleichungssytem $A\vec{x} = \vec{e}_l$. Die Lösung ist die l-te Spalte von A^{-1}. Man kann diese Gleichungssysteme auch simultan lösen.

Beispiel: Sei $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathbb{K}^{2\times 2}$ invertierbar und $a \neq 0$. Wir setzen $\delta := ad - bc$ und sehen an der dritten Matrix, dass $\delta \neq 0$ ist.

$$\left(\begin{array}{cc|cc} a & b & 1 & 0 \\ c & d & 0 & 1 \end{array}\right) \quad \left(\begin{array}{cc|cc} 1 & b/a & 1/a & 0 \\ 0 & d - bc/a & -c/a & 1 \end{array}\right) \quad \left(\begin{array}{cc|cc} 1 & b/a & 1/a & 0 \\ 0 & \delta/a & -c/a & 1 \end{array}\right)$$
$$\left(\begin{array}{cc|cc} 1 & b/a & 1/a & 0 \\ 0 & 1 & -c/\delta & a/\delta \end{array}\right) \quad \left(\begin{array}{cc|cc} 1 & 0 & \frac{1}{a} + \frac{bc}{\delta a} & -b/\delta \\ 0 & 1 & -c/\delta & a/\delta \end{array}\right).$$

Wegen $\frac{1}{a} + \frac{bc}{\delta a} = d/\delta$ ist also

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

Den Fall $c \neq 0$ behandelt man analog, er führt auf dieselbe Formel. Wir sehen auch, dass Invertierbarkeit äquivalent ist zu $\delta \neq 0$.

Anwendung Basiswechsel: Ist $\vec{b}_1, \dots, \vec{b}_n$ eine weitere Basis und B die Matrix mit Spalten $\vec{b}_k$, $k = 1, \dots, n$, so ist B die Darstellungsmatrix von $\mathrm{Id} : \mathbb{K}^n \to \mathbb{K}^n$ bzgl. $\vec{b}_1, \dots, \vec{b}_n$ im Argumentraum und Standardbasis im Zielraum (vergleiche Beispiel 15.14(3)). Die Matrix B ist invertierbar und B^{-1} ist gerade die Darstellungsmatrix von $\mathrm{Id} : \mathbb{K}^n \to \mathbb{K}^n$ bzgl. Standardbasis im Argumentraum und $\vec{b}_1, \dots, \vec{b}_n$ im Zielraum (im Beispiel 15.14(3) gilt $BC = I$, also $C = B^{-1}$).

Ist $\phi : \mathbb{K}^n \to \mathbb{K}^n$ eine lineare Abbildung mit Darstellungsmatrix F bzgl. der Standardbasis, so ist die Darstellungsmatrix von ϕ bzgl. der Basis $\vec{b}_1, \dots, \vec{b}_n$ in Argument- und Zielraum gegeben durch $B^{-1}FB$. Mit F aus Beispiel 15.15 und B aus Beispiel 15.14(3) gilt etwa

$$B^{-1}FB = A,$$

mit der Matrix A aus Beispiel 15.14(2).

Beispiel: Sei V der $\mathbb{C}$-Vektorraum der Polynome $p = p(x)$ mit komplexen Koeffizienten vom Grad ≤ 2 und die lineare Abbildung $\phi : V \to V$ gegeben durch $\phi(p) = p' - p$. Die Darstellungsmatrix F von ϕ bzgl. der Basis $e_1(x) = 1, e_2(x) = x, e_3(x) = x^2$ von V ist wegen $\phi(1) = -1$, $\phi(x) = 1 - x$, $\phi(x^2) = 2x - x^2$ gegeben durch

$$F = \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 2 \\ 0 & 0 & -1 \end{pmatrix}.$$

Eine andere Basis von V ist gegeben durch

$$b_1(x) = (x+i)^2 = x^2 + 2ix - 1, \quad b_2(x) = (x-i)^2 = x^2 - 2ix - 1, \quad b_3(x) = x^2 + 1,$$

und die Darstellungsmatrix von $\mathrm{Id}: V \to V$ mit b-Basis im Argumentraum und e-Basis im Zielraum ist

$$B = \begin{pmatrix} -1 & -1 & 1 \\ 2i & -2i & 0 \\ 1 & 1 & 1 \end{pmatrix}.$$

Wir invertieren B nach dem oben angegebenen Verfahren:

$$\left(\begin{array}{rrr|rrr} -1 & -1 & 1 & 1 & 0 & 0 \\ 2i & -2i & 0 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 0 & 1 \end{array}\right) \left(\begin{array}{rrr|rrr} 1 & 1 & -1 & -1 & 0 & 0 \\ 0 & -4i & 2i & 2i & 1 & 0 \\ 0 & 0 & 2 & 1 & 0 & 1 \end{array}\right)$$

$$\left(\begin{array}{rrr|rrr} 1 & 1 & 0 & -1/2 & 0 & 1/2 \\ 0 & -4i & 0 & i & 1 & -i \\ 0 & 0 & 1 & 1/2 & 0 & 1/2 \end{array}\right) \left(\begin{array}{rrr|rrr} 1 & 0 & 0 & -1/4 & -i/4 & 1/4 \\ 0 & 1 & 0 & -1/4 & i/4 & 1/4 \\ 0 & 0 & 1 & 1/2 & 0 & 1/2 \end{array}\right)$$

(man überzeuge sich am Ende unbedingt durch die Probe). Die Inverse von B ist also gegeben durch

$$B^{-1} = \begin{pmatrix} -1/4 & -i/4 & 1/4 \\ -1/4 & i/4 & 1/4 \\ 1/2 & 0 & 1/2 \end{pmatrix},$$

und die Darstellungsmatrix von ϕ bzgl. der b-Basis in Argument- und Zielraum ist also

$$\begin{aligned} B^{-1}FB &= \begin{pmatrix} -1/4 & -i/4 & 1/4 \\ -1/4 & i/4 & 1/4 \\ 1/2 & 0 & 1/2 \end{pmatrix} \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 2 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} -1 & -1 & 1 \\ 2i & -2i & 0 \\ 1 & 1 & 1 \end{pmatrix} \\ &= \begin{pmatrix} -1-i & 0 & -i/2 \\ 0 & -1+i & i/2 \\ i & -i & -1 \end{pmatrix}. \end{aligned}$$

Das bedeutet $b_1' - b_1 = (-1-i)b_1 + ib_3$, $b_2' - b_2 = (-1+i)b_2 - ib_3$, $b_3' - b_3 = -\frac{i}{2}b_1 + \frac{i}{2}b_2 - b_3$, bzw.

$$b_1' = -ib_1 + ib_3, \quad b_2' = ib_2 - ib_3, \quad b_3' = -\frac{i}{2}b_1 + \frac{i}{2}b_2$$

(man mache auch hier die Probe).

16 Skalarprodukt und Orthogonalität

16.1. Skalarprodukte: Sei V ein $\mathbb{K}$-Vektorraum. Eine Abbildung $(\cdot|\cdot) : V \times V \to \mathbb{K}$ mit den Eigenschaften

(S1) $\forall x, y \in V$: $(x|y) = \overline{(y|x)}$,

(S2) $\forall x, y, z \in V$, $\alpha \in \mathbb{K}$: $(\alpha x + y|z) = \alpha(x|z) + (y|z)$,

(S3) $\forall x \in V \setminus \{0\}$: $(x|x) > 0$.

heißt ein *Skalarprodukt auf* V.

Eigenschaften eines Skalarproduktes auf einem $\mathbb{K}$-Vektorraum V sind

$$\forall x, y \in V, \alpha \in \mathbb{K}\colon (x|\alpha y + z) = \overline{\alpha}(x|y) + (x|z),$$
$$\forall x \in V\colon (x|0) = (0|x) = 0,$$
$$\forall x, y \in V : |(x|y)| \leq \sqrt{(x|x)}\sqrt{(y|y)} \text{ (Cauchy-Schwarz-Ungleichung).}$$

Zum Beweis der Cauchy-Schwarz-Ungleichung für $y \neq 0$ schreibe man

$$0 \leq (x - \alpha y|x - \alpha y) = (x|x) - \alpha(y|x) - \overline{\alpha}(x|y) + |\alpha|^2(y|y) = (x|x) - \frac{|(x|y)|^2}{(y|y)},$$

wobei $\alpha = (x|y)/(y|y)$. Wegen (S3) gilt Gleichheit in der Cauchy-Schwarz-Ungleichung nur, wenn $x = \alpha y$, dh also nur dann, wenn x, y linear abhängig sind.

Bemerkung: Ist $\mathbb{K} = \mathbb{R}$ und also V ein $\mathbb{R}$-Vektorraum, so kann auf die komplexe Konjugation verzichtet werden, da $\overline{r} = r$ für alle $r \in \mathbb{R}$.

Beispiele: (1) Das *gewöhnliche Skalarprodukt* auf dem $\mathbb{K}^n$ (für $\mathbb{K} = \mathbb{R}$ heißt es auch *euklidisch*) ist gegeben durch

$$(\vec{x}|\vec{y}) := \sum_{j=1}^{n} x_j \overline{y_j} \quad \text{für } \vec{x} = (x_1, x_2, \ldots, x_n), \vec{y} = (y_1, y_2, \ldots, y_n) \in \mathbb{K}^n.$$

Für das Skalarprodukt von $\vec{x}, \vec{y} \in \mathbb{R}^n$ schreibt man häufig auch $\vec{x} \cdot \vec{y}$ (und gelegentlich sogar nur $\vec{x}\vec{y}$).

(2) Sind $a_1, a_2, \ldots, a_n > 0$, so definiert auch

$$(\vec{x}|\vec{y}) := \sum_{j=1}^{n} a_j x_j \overline{y_j} \quad \text{für } \vec{x} = (x_1, x_2, \ldots, x_n), \vec{y} = (y_1, y_2, \ldots, y_n) \in \mathbb{K}^n$$

ein Skalarprodukt auf $\mathbb{K}^n$: (S1) und (S2) sind leicht, für (S3) beachte man, dass

$$(\vec{x}|\vec{x}) = \sum_{j=1}^{n} a_j |x_j|^2 \geq 0$$

wegen $a_j > 0$ nur dann $= 0$ ist, wenn alle $x_j = 0$ sind.

(2') Sei $A \in \mathbb{K}^{n\times n}$ invertierbar und $(\cdot|\cdot)$ das Standardskalarprodukt. Dann definiert

$$(\!(\vec{x}|\vec{y})\!) := (A\vec{x}|A\vec{y})$$

ein Skalarprodukt auf $\mathbb{K}^n$. Auch hier sind (S1) und (S2) leicht. Für (S3) benötigt man die Invertierbarkeit von A.

(3) Ist $V = C([a,b],\mathbb{C})$ (hier ist $\mathbb{K} = \mathbb{C}$), so wird durch

$$(f|g) := \int_a^b f(t)\overline{g(t)}\,dt, \quad f,g \in C([a,b],\mathbb{C}),$$

ein Skalarprodukt auf $C([a,b],\mathbb{C})$ definiert. Zunächst führen wir die nötigen Begriffe ein: $C([a,b],\mathbb{C})$ ist der Raum der stetigen Funktionen $[a,b] \to \mathbb{C}$, und eine Funktion $f : [a,b] \to \mathbb{C}$ heißt *stetig*, falls die beiden reellwertigen Funktionen

$$\operatorname{Re} f : [a,b] \to \mathbb{R}, t \mapsto (\operatorname{Re} f(t)), \quad \operatorname{Im} f : [a,b] \to \mathbb{R}, t \mapsto (\operatorname{Im} f(t)),$$

stetig sind[7] Entsprechend definieren wir $R([a,b],\mathbb{C})$ als Raum aller intergierbaren Funktionen $f : [a,b] \to \mathbb{C}$, wobei f *integrierbar* heißt, wenn $\operatorname{Re} f$ und $\operatorname{Im} f$ beide integrierbar sind. In diesem Fall setzt man

$$\int_a^b f(t)\,dt := \int_a^b (\operatorname{Re} f)(t)\,dt + i\int_a^b (\operatorname{Im} f)(t)\,dt.$$

$C([a,b],\mathbb{C})$ und $R([a,b],\mathbb{C})$ sind $\mathbb{C}$-Vektorräume, und die Abbildung

$$R([a,b],\mathbb{C}) \to \mathbb{C}, \quad f \mapsto \int_a^b f(t)\,dt,$$

ist $\mathbb{C}$-linear.
Nun zurück zu $(\cdot|\cdot)$: Die Eigenschaften (S1) und (S2) sind leicht. Zum Nachweis von (S3) stellt man fest, dass für $h \in C[a,b]$ mit $h \geq 0$ und $\int_a^b h(t)\,dt = 0$ zunächst folgt $\int_a^x h(t)\,dt$ für alle $x \in [a,b]$, und dann $h = 0$ nach dem Hauptsatz (alternativ kann man die entsprechende Übungsaufgabe verwenden). Diese Beobachtung wendet man auf $h := |f|^2$ an.

Bemerkung: Man kann $(f|g)$ auch für $f,g \in R([a,b],\mathbb{C})$ definieren (beachte, dass $f\overline{g} \in R[a,b]$ nach einer komplexen Version von 12.14(3)). Die Eigenschaften (S1), (S2) gelten weiterhin, aber (S3) ist falsch. Definiert man $f(t) = 0$ für $t \in [a,b] \setminus \{\frac{a+b}{2}\}$ und $f(\frac{a+b}{2}) = 1$, so gilt $f \neq 0$, aber $(f|f) = \int_a^b |f(t)|^2\,dt = 0$ (das Integral "merkt" nicht, wenn man die Nullfunktion in einem Punkt abändert).

[7] Das ist äquivalent zu $f(t_n) \to f(t_0)$ für jede Folge (t_n) in $[a,b]$ mit $t_n \to t_0 \in [a,b]$, vgl. 6.2 und 9.1.

(4) Durch Grenzübergang von (1) erhält man: Sei $l^2(\mathbb{N})$ der Raum der quadratsummierbaren Folgen in $\mathbb{K}$, also

$$l^2(\mathbb{N}) = \{x = (x_j)_{j\in\mathbb{N}} \in \mathbb{K}^{\mathbb{N}} : \sum_{j=1}^{\infty} |x_j|^2 < \infty\}.$$

Dann ist $l^2(\mathbb{N})$ ein Untervektorraum von $\mathbb{K}^{\mathbb{N}}$ und durch

$$(x|y) := \sum_{j=1}^{\infty} x_j \overline{y_j}, \quad x, y \in l^2(\mathbb{N}),$$

wird ein Skalarprodukt auf $l^2(\mathbb{N})$ definiert. Dabei ist die Reihe im Skalarprodukt absolut konvergent: Zunächst gilt für jedes $N \in \mathbb{N}$ nach Cauchy-Schwarz:

$$\begin{aligned} \sum_{j=1}^{N} |x_j\overline{y_j}| = \sum_{j=1}^{N} |x_j||y_j| &\leq \Big(\sum_{j=1}^{N} |x_j|^2\Big)^{1/2}\Big(\sum_{j=1}^{N} |y_j|^2\Big)^{1/2} \\ &\leq \Big(\sum_{j=1}^{\infty} |x_j|^2\Big)^{1/2}\Big(\sum_{j=1}^{\infty} |y_j|^2\Big)^{1/2} =: M < \infty, \end{aligned}$$

und $N \to \infty$ impliziert absolute Konvergenz. Weiter haben wir für $x, y \in l^2(\mathbb{N})$:

$$\sum_{j=1}^{N} |x_j + y_j|^2 = \sum_{j=1}^{\infty} |x_j|^2 + \sum_{j=1}^{\infty} |y_j|^2 + \underbrace{\sum_{j=1}^{N} (x_j\overline{y_j} + \overline{x_j}y_j)}_{\leq 2M},$$

was für $N \to \infty$ impliziert, dass $x + y \in l^2(\mathbb{N})$ gilt. Nun sind die Eigenschaften leicht nachzuweisen.

Entsprechendes gilt, wenn man statt $\mathbb{N}$ als Indexmenge $\mathbb{Z}$ nimmt und

$$l^2(\mathbb{Z}) := \{x = (x_j)_{j\in\mathbb{Z}} \in \mathbb{K}^{\mathbb{Z}} : \sum_{j=-\infty}^{\infty} |x_j|^2 < \infty\}$$

betrachtet.

16.2. Normen: Ist V ein $\mathbb{K}$-Vektorraum und $(\cdot|\cdot) : V \times V \to \mathbb{K}$ ein Skalarprodukt, so hat die Abbildung $\|\cdot\| : V \to [0, \infty)$, $v \mapsto \sqrt{(v|v)}$ folgende Eigenschaften:

(N1) $\forall v \in V$: $\|v\| = 0 \Rightarrow v = 0$,

(N2) $\forall v \in V$, $\alpha \in \mathbb{K}$: $\|\alpha v\| = |\alpha|\|v\|$,

(N3) $\forall u, v \in V$: $\|u + v\| \leq \|u\| + \|v\|$ (Dreiecksungleichung).

Beweis der Dreiecksungleichung. Es gilt (unter Verwendung der Cauchy-Schwarz-Ungleichung):

$$\|u+v\|^2 = (u+v|u+v) = \|u\|^2+2\mathrm{Re}\,(u|v)+\|v\|^2 \le \|u\|^2+2|(u|v)|+\|v\|^2 \le (\|u\|+\|v\|)^2,$$

und die Ungleichung folgt durch Wurzelziehen. □

Definition: Sei V ein $\mathbb{K}$-Vektorraum. Eine Abbildung $\|\cdot\| : V \to [0,\infty)$ mit den Eigenschaften (N1)–(N3) heißt eine *Norm auf* V.

Ist $\|\cdot\|$ eine Norm auf V, so wird für $u, v \in V$ die Zahl $\|u - v\| \ge 0$ als **Abstand** von u und v interpretiert, und es gilt $\|0\| = 0$, sowie

$$|\|u\| - \|v\|| \le \|u - v\| \quad \text{(umgekehrte Dreiecksungleichung).}$$

Beispiel: Durch $\|(x_1, x_2)\|_1 := |x_1| + |x_2|$ wird auf dem $\mathbb{R}^2$ eine Norm definiert, zu der es kein Skalarprodukt gibt.

Bemerkung: Durch eine Norm hat man also einen *Abstandsbegriff* auf einem Vektorraum. Durch ein Skalarprodukt hat man aber außerdem noch die Möglichkeit, *Winkel* zu betrachten:

Sind z.B. $\vec{x}, \vec{y} \in \mathbb{R}^n \setminus \{\vec{0}\}$, so ist

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

wobei φ den von $\vec{x}, \vec{y}$ eingeschlossenen **Winkel** bezeichne. Da das Skalarprodukt in jeder Komoponente linear ist, muss man das nur für $\|\vec{x}\| = \|\vec{y}\| = 1$ einsehen. Am besten geht das für $n = 2$ und etwa $\vec{x} = \vec{e}_1$, $\vec{y} = (y_1, y_2)$. Dann ist

$$(\vec{x}|\vec{y}) = (\vec{e}_1|\vec{y}) = y_1.$$

16.3. Orthogonalität: Sei V ein $\mathbb{K}$-Vektorraum mit Skalarprodukt. Vektoren $v_1, v_2, \ldots, v_m \in V$ heißen *orthogonal*, falls für alle $j, k \in \{1, 2 \ldots, m\}$ mit $j \ne k$ gilt $(v_j|v_k) = 0$. Statt $(v|w) = 0$ schreibt man auch $v \perp w$.

Die Vektoren $v_1, v_2, \ldots, v_m$ heißen *orthonormal* oder ein *Orthonormalsystem* (*ONS*), falls für alle j, k gilt: $(v_j|v_k) = \delta_{jk}$, dh also falls die Vektoren orthogonal sind und zusätzlich alle Norm 1 haben.

Ist V endlich-dimensional, so ist eine *Orthonormalbasis* (*ONB*) von V eine Basis von V, die ein Orthonormalsystem ist.

Beispiel: Die Standardbasis $\vec{e}_1, \vec{e}_2, \ldots, \vec{e}_n$ ist eine Orthonormalbasis von $\mathbb{K}^n$, denn es gilt $(\vec{e}_j|\vec{e}_k) = \delta_{jk}$ für alle $j, k = 1, \ldots, n$.

Satz: Sei V ein $\mathbb{K}$-Vektorraum mit Skalarprodukt und seien $v_1, v_2, \ldots, v_n \in V \setminus \{0\}$ orthogonal. Dann sind $v_1, v_2, \ldots, v_n$ linear unabhängig.

Beweis. Seien $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{K}$ mit

$$\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n = 0.$$

Sei nun $k \in \{1, 2, \ldots, n\}$. Wir nehmen das Skalarprodukt der Gleichung mit v_k und erhalten

$$0 = (0|v_k) = (\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n | v_k) = \alpha_k (v_k|v_k)$$

wegen der vorausgesetzten Orthogonalität. Wegen $v_k \neq 0$ ist auch $(v_k|v_k) \neq 0$, und wir erhalten $\alpha_k = 0$. Da k beliebig war, sind $v_1, v_2, \ldots, v_n$ linear unabhängig. □

Bemerkung: Ist $v_1, v_2, \ldots, v_m$ ein Orthonormalsystem in V und $v \in \operatorname{lin}\{v_1, v_2, \ldots, v_m\}$, so lassen sich die Koordinaten von v bzgl. $v_1, v_2, \ldots, v_m$ leicht bestimmen. Es gilt nämlich

$$v = \sum_{j=1}^{m} (v|v_j) v_j.$$

Zum Beweis schreibt man $v = \sum_{j=1}^{m} \alpha_j v_j$ und bildet das Skalarprodukt mit v_k, $k = 1, \ldots, m$:

$$(v|v_k) = \sum_{j=1}^{m} \alpha_j \underbrace{(v_j|v_k)}_{=\delta_{jk}} = \alpha_k.$$

Beispiel: Sei $e_k : [-\pi, \pi] \to \mathbb{C}$, $e_k(t) := e^{ikt}$. In $V = C([-\pi, \pi], \mathbb{C})$ bilden für jedes $n \in \mathbb{N}$ die Funktionen $(e_k)_{|k| \leq n}$ ein Orthonormalsystem bezüglich des durch

$$(f|g) := \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) \overline{g(t)} \, dt$$

auf V definierten Skalarprodukts, denn für $k, l \in \mathbb{Z}$ gilt

$$(e_k|e_l) = \frac{1}{2\pi} \int_{-\pi}^{\pi} e^{ikt} e^{-ilt} \, dt = \begin{cases} 1 & , k = l \\ 0 & , k \neq l \end{cases} .$$

Für jede Funktion $f \in C([-\pi, \pi], \mathbb{C})$ definieren wir die *Fourierkoeffizienten* von f durch

$$\hat{f}(k) := (f|e_k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) e^{-ikt} \, dt, \quad k \in \mathbb{Z}.$$

Jede Linearkombination der Funktionen e_k, $|k| \leq n$, heißt *trigonometrisches Polynom vom Grad* $\leq n$. Ist $f = \sum_{|j| \leq n} c_j e_j$ ein trigonometrisches Polynom vom Grad $\leq n$, so gilt für jedes $k \in \mathbb{Z}$

$$\hat{f}(k) = (\sum_{|j| \leq n} c_j e_j | e_k) = \sum_{|j| \leq n} c_j (e_j|e_k) = \sum_{|j| \leq n} c_j \delta_{jk} = \begin{cases} c_k & , |k| \leq n \\ 0 & , |k| > n \end{cases} .$$

Die Koordinaten sind hier also gerade die Fourierkoeffizienten, und es gilt

$$f(t) = \sum_{|k|\le n} \hat{f}(k) e_k(t) = \sum_{|k|\le n} \hat{f}(k) e^{ikt}, \quad t \in [-\pi, \pi].$$

Wenn wir also folgenden Fourierreihen betrachten, beziehen wir uns immer auf das hier definierte Skalarprodukt. Die zugehörige Norm ist gegeben durch

$$\|f\| = \sqrt{(f|f)} = \Big(\frac{1}{2\pi}\int_{-\pi}^{\pi} |f(t)|^2\Big)^{1/2}.$$

16.4. Das Gram-Schmidt-Verfahren: Gegeben sei ein $\mathbb{K}$-Vektorraum mit Skalarprodukt und $v_1, v_2, \ldots, v_m \in V$ seien linear unabhängig. Wir werden ein Orthonormalsystem $b_1, b_2, \ldots, b_m$ konstruieren mit

$$\operatorname{lin}\{v_1, v_2, \ldots, v_m\} = \operatorname{lin}\{b_1, b_2, \ldots, b_m\}.$$

Falls $v_1, v_2, \ldots, v_m$ eine Basis von V ist, so ist $b_1, b_2, \ldots, b_m$ eine Orthonormalbasis von V. Die Vektoren $b_1, \ldots, b_m$ werden sukzessiv so konstruiert, dass gilt:

$$\operatorname{lin}\{v_1, v_2, \ldots, v_k\} = \operatorname{lin}\{b_1, b_2, \ldots, b_k\} \quad \text{für alle } k = 1, \ldots, m.$$

Wir setzen $b_1 := v_1/\|v_1\|$ und für $k = 2, \ldots, m$:

$$c_k := v_k - \sum_{j=1}^{k-1} (v_k|b_j) b_j, \quad b_k := \frac{c_k}{\|c_k\|},$$

oder gleichbedeutend

$$c_1 := v_1, \quad c_k := v_k - \sum_{j=1}^{k-1} \frac{(v_k|c_j)}{(c_j|c_j)} c_j \text{ für } k = 2, \ldots, m, \quad b_k := \frac{c_k}{\|c_k\|} \text{ für } k = 1, \ldots, m.$$

Beispiele: (1) Wir betrachten $n = 3$, $V = \mathbb{C}^3$ und

$$\vec{v}_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \quad \vec{v}_2 = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \quad \vec{v}_3 = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}.$$

Dann ist

$$\begin{aligned}
\vec{b}_1 &= \frac{\vec{v}_1}{\|\vec{v}_1\|} = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\0\\-1\end{pmatrix}\\
\vec{c}_2 &= \begin{pmatrix}2\\1\\0\end{pmatrix} - \frac{1}{2}\cdot 2\cdot\begin{pmatrix}1\\0\\-1\end{pmatrix} = \begin{pmatrix}1\\1\\1\end{pmatrix}\\
\vec{b}_2 &= \frac{1}{\sqrt{3}}\begin{pmatrix}1\\1\\1\end{pmatrix}\\
\vec{c}_3 &= \begin{pmatrix}0\\-1\\1\end{pmatrix} - \frac{1}{2}\cdot(-1)\cdot\begin{pmatrix}1\\0\\-1\end{pmatrix} - \frac{1}{3}\cdot 0\cdot\begin{pmatrix}1\\1\\1\end{pmatrix} = \begin{pmatrix}1/2\\-1\\1/2\end{pmatrix}\\
\vec{b}_3 &= \frac{1}{\sqrt{6}}\begin{pmatrix}1\\-2\\1\end{pmatrix}.
\end{aligned}$$

(2) $\vec{v}_1, \vec{v}_2$ wie eben, aber $\vec{v}_3 = \begin{pmatrix}0\\1\\1\end{pmatrix}$. Dann sind $\vec{b}_1, \vec{b}_2$ wie eben, aber

$$\begin{aligned}
\vec{c}_3 &= \begin{pmatrix}0\\1\\1\end{pmatrix} - \frac{1}{2}\cdot(-1)\cdot\begin{pmatrix}1\\0\\-1\end{pmatrix} - \frac{1}{3}\cdot 2\cdot\begin{pmatrix}1\\1\\1\end{pmatrix} = \begin{pmatrix}-1/6\\1/3\\-1/6\end{pmatrix}\\
\vec{b}_3 &= \frac{1}{\sqrt{6}}\begin{pmatrix}-1\\2\\-1\end{pmatrix}.
\end{aligned}$$

Man beachte, dass sich dieses $\vec{b}_3$ von dem in Beispiel (1) nur durch das Vorzeichen unterscheidet. Dies ist nicht erstaunlich, da es genau zwei Möglichkeiten gibt, $\vec{b}_1, \vec{b}_2$ zu einer Orthonormalbasis zu ergänzen.

Folgerung: Jeder endlichdimensionale Vektorraum V mit Skalarprodukt besitzt eine Orthonormalbasis.

16.5. Transponierte und adjungierte Matrizen: Für eine Matrix $A = (a_{jk}) \in \mathbb{K}^{m\times n}$ heißt die Matrix $\in \mathbb{K}^{n\times m}$, die durch Vertauschen von Zeilen und Spalten entsteht, die *transponierte Matrix zu* A und wird mit A^T bezeichnet. Für alle $k \in \{1,\dots,n\}$, $j \in \{1,\dots,m\}$ steht an der Stelle (k,j) in der Matrix A^T also der Eintrag a_{jk}, der in der Matrix A an der Stelle (j,k) steht. Setzen wir $B := A^T$ mit $B = (b_{kj})_{k=1}^{n}{}_{j=1}^{m} \in \mathbb{K}^{n\times m}$, so gilt also

$$b_{kj} = a_{jk}, \quad k \in \{1,\dots,n\}, j \in \{1,\dots,m\}.$$

Im Falle $\mathbb{K} = \mathbb{C}$ heißt die Matrix $\in \mathbb{C}^{n\times m}$, für die an jeder Stelle (k,j) der Eintrag $\overline{a_{jk}}$ steht, die *adjungierte Matrix zu* A und wird mit A^* bezeichnet.

Bemerkung: Setzt man $\overline{A} := (\overline{a_{jk}}) \in \mathbb{C}^{m\times n}$ (*konjugiert komplexe Matrix zu* A), so gilt also

$$A^* = \overline{A^T} = \left(\overline{A}\right)^T.$$

Schreibweisen des Skalarprodukts:

Für $\vec{x}, \vec{y} \in \mathbb{R}^n = \mathbb{R}^{n\times 1}$ gilt $(\vec{x}|\vec{y}) = \vec{y}^T\vec{x} = \vec{x}^T\vec{y}$.
Für $\vec{x}, \vec{y} \in \mathbb{C}^n = \mathbb{C}^{n\times 1}$ gilt $(\vec{x}|\vec{y}) = \vec{y}^*\vec{x} = \overline{\vec{x}^*\vec{y}}$.

Rechenregeln: Für Matrizen A, B, deren Produkt erklärt ist, gilt:

$$(AB)^T = B^T A^T \quad \text{und} \quad (AB)^* = B^* A^*.$$

Für eine invertierbare Matrix A sind auch A^T und A^* invertierbar, und es gilt

$$(A^T)^{-1} = (A^{-1})^T, \qquad (A^*)^{-1} = (A^{-1})^*.$$

Wende dazu die Rechenregeln auf $B = A^{-1}$ an und beachte $I^T = I = I^*$, wobei I die jeweilige Einheitsmatrix sei.

Folgerung: Sei $A \in \mathbb{K}^{m\times n}$. Dann gilt:

(a) Im Fall $\mathbb{K} = \mathbb{R}$ ist $(A\vec{x}|\vec{y}) = (\vec{x}|A^T\vec{y})$ für alle $\vec{x} \in \mathbb{R}^n$, $\vec{y} \in \mathbb{R}^m$.

(b) Im Fall $\mathbb{K} = \mathbb{C}$ ist $(A\vec{x}|\vec{y}) = (\vec{x}|A^*\vec{y})$ für alle $\vec{x} \in \mathbb{C}^n$, $\vec{y} \in \mathbb{C}^m$.

Beweis. Man schreibe das Skalarprodukt wie oben angegeben und benutze die Rechenregeln. □

16.6. Orthogonale und unitäre Matrizen: Eine wichtige Rolle spielen Matrizen $A \in \mathbb{K}^{n\times n}$, deren zugehörige lineare Abbildung $\mathbb{K}^n \to \mathbb{K}^n$, $\vec{x} \mapsto A\vec{x}$, das Skalarprodukt invariant lässt, dh für die gilt:

$$(A\vec{x}|A\vec{y}) = (\vec{x}|\vec{y}) \quad \text{für alle } \vec{x}, \vec{y} \in \mathbb{K}^n.$$

Damit verändert A auch *Winkel* und *Abstände* nicht. Eine solche Matrix A hat Kern $A = \{\vec{0}\}$, ist also invertierbar. Aus den Rechenregeln folgt

$$A^T A = I_n \quad (\text{für } \mathbb{K} = \mathbb{R}), \qquad A^* A = I_n \quad (\text{für } \mathbb{K} = \mathbb{C}).$$

Eine Matrix $A \in \mathbb{K}^{n\times n}$ mit dieser Eigenschaft heißt *orthogonal* (für $\mathbb{K} = \mathbb{R}$) bzw. *unitär* (für $\mathbb{K} = \mathbb{C}$). Somit gilt

$$A^T = A^{-1} \text{ falls } A \in \mathbb{R}^{n\times n} \text{ orthogonal}, A^* = A^{-1} \text{ falls } A \in \mathbb{C}^{n\times n} \text{ unitär}.$$

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

1) A ist genau dann unitär, wenn die Spalten von A eine Orthonormalbasis von $\mathbb{C}^n$ bilden.

2) A ist genau dann unitär, wenn für jedes Orthonormalsystem $v_1, \dots, v_m$ in $\mathbb{C}^n$ auch $Av_1, \dots, Av_m$ ein Orthonormalsystem von $\mathbb{C}^n$ ist.

3) Produkte, Inverse, Transponierte und Adjungierte von unitären Matrizen sind unitär.

Beispiele: 1) Spiegelungen in $\mathbb{C}^n$: etwa $A\vec{e}_1 = -\vec{e}_1$, $A\vec{e}_j = \vec{e}_j$ für $j = 2, \dots, n$.

2) Rotation in $\mathbb{R}^2$ um den Winkel $\theta \in \mathbb{R}$: $A = \left(\begin{smallmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{smallmatrix}\right)$.

3) Im $\mathbb{R}^3$ Rotation um die z-Achse bei Spiegelung an der (x, y)-Ebene:

$$A = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

16.7. Orthogonalprojektionen: Sei V ein $\mathbb{K}$-Vektorraum mit Skalarprodukt, $b_1, b_2, \dots, b_m$ ein Orthonormalsystem in V und $U := \mathrm{lin}\{b_1, b_2, \dots, b_m\}$. Die lineare Abbildung

$$P : V \to U, v \mapsto Pv = \sum_{j=1}^{m} (v|b_j) b_j$$

hat folgende Eigenschaften:

$$P \circ P = P, \quad \mathrm{Bild}\, P = U, \quad \mathrm{Kern}\, P = \{v \in V : v \perp u \text{ für alle } u \in U\}.$$

Es gilt $(v - Pv|u) = 0$ für alle $u \in U$, $v \in V$, und

$$\|v - Pv\| = \min\{\|v - u\| : u \in U\} \quad \text{für jedes } v \in V,$$

dh Pv ist die (eindeutig bestimmte) *Bestapproximation* von v in U. Die Abbildung P heißt *Orthogonalprojektion* von V auf U.

Beweis. Klar ist $\mathrm{Bild}\, P \subseteq U$. Nach der Bemerkung in 16.3 gilt $Pu = u$ für alle $u \in U$, also ist $P \circ P = P$ und $\mathrm{Bild}\, P = U$. Für jedes $v \in V$ gilt:

$$Pv = 0 \iff (v|b_j) = 0 \text{ für alle } j \iff (v|u) = 0 \text{ für alle } u \in U.$$

Außerdem ist für $k = 1, \dots, m$:

$$(v - Pv|b_k) = (v|b_k) - (Pv|b_k) = (v|b_k) - \sum_{j=1}^{m} (v|b_j) \underbrace{(b_j|b_k)}_{=\delta_{jk}} = (v|b_k) - (v|b_k) = 0,$$

und folglich auch $(v - Pv|u) = 0$ für alle $u \in U$. Schließlich gilt für $u \in U$:

$$\begin{aligned}
\|v-u\|^2 &= \|v - Pv + Pv - u\|^2 = \|v - Pv\|^2 + \|Pv - u\|^2 + 2\mathrm{Re}\,(v - Pv|Pv - u) \\
&= \|v - Pv\|^2 + \|Pv - u\|^2.
\end{aligned}$$

Hieran sieht man, dass Pv die eindeutige Bestapproximation von v in U ist. $\square$

Bemerkung: Im Gram-Schmidt-Verfahren in 16.4 hat man im k-ten Schritt (für $k = 2, \ldots, m$) und für $U = \mathrm{lin}\{b_1, b_2, \ldots, b_{k-1}\}$, dass $c_k = v_k - Pv_k$ und somit $c_k \perp u$ für jedes $u \in U$ wie gewünscht.

Beispiel: Sei $V := C([-\pi, \pi], \mathbb{C})$ mit dem Skalarprodukt aus dem Beispiel in 16.3 und sei $n \in \mathbb{N}$. Für eine gegebene Funkton $f \in V$ ist die Funktion

$$t \mapsto \sum_{|k| \le n} (f|e_k) e_k(t) = \sum_{|k| \le n} \hat{f}(k) e^{ikt}$$

die Bestapproximation von f in $U := \mathrm{lin}\{e_k : |k| \le n\}$, also $(Pf)(t) = \sum_{|k|\le n} \hat{f}(k)e^{ikt}$.

Einschub zu Fourierreihen: Wir haben bisher für stetige Funktionen $f : [-\pi, \pi] \to \mathbb{C}$ die Fourierkoeffizienten

$$\hat{f}(k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(t) e^{-ikt}\, dt, \quad k \in \mathbb{Z},$$

erklärt, wobei $\hat{f}(k) = (f|e_k)$ für das betrachtete Skalarprodukt und $e_k(t) = e^{ikt}$ gilt. In Beispiel 16.1(3) waren wir auf das Problem der Definitheit eingegangen, wenn man das Skalarprodukt in dem größeren Raum $R([-\pi, \pi], \mathbb{C})$ betrachten möchte.

Man kann das Problem für stückweise stetige Funktionen (siehe Folgerung in 12.8) noch beheben, wenn man nur *normalisierte* Funktionen $f : [-\pi, \pi] \to \mathbb{C}$ betrachtet, dh solche, für die an den Unstetigkeitsstellen $t_j \in (-\pi, \pi)$ gilt $f(t_j) = \frac{f(t_j-)+f(t_j+)}{2}$ und außerdem $f(\pi) = f(-\pi) = \frac{f(\pi-)+f(-\pi+)}{2}$ (auf den Sinn dieser Bedingung gehen wir später ein). Auch auf dem Raum dieser Funktionen ist $(\cdot|\cdot)$ ein Skalarprodukt.

Wir betrachten als Beispiel $f : [-\pi, \pi] \to \mathbb{C}$, $f(t) = t$ für $t \in (-\pi, \pi)$ und $f(\pm\pi) = 0$. Dann gilt $\hat{f}(0) = 0$ und

$$\hat{f}(k) = \frac{1}{2\pi} \int_{-\pi}^{\pi} t e^{-ikt}\, dt = (-1)^k \frac{i}{k}, \quad k \in \mathbb{Z} \setminus \{0\}.$$

Für reellwertige Funktionen $f : [-\pi, \pi] \to \mathbb{R}$ ist es eventuell angenehmer, statt der komplexen Fourierkoeffizienten $(\hat{f}(k))_{k\in\mathbb{Z}}$ die reellen Folgen

$$\begin{aligned}
a_k(f) &:= \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \cos(kt)\, dt, \quad k \in \mathbb{N}_0, \\
b_k(f) &:= \frac{1}{\pi} \int_{-\pi}^{\pi} f(t) \sin(kt)\, dt, \quad k \in \mathbb{N},
\end{aligned}$$

zu betrachten. Wegen der Definitionen von cos und sin (siehe 7.12) haben wir $a_0(f) = 2\hat{f}(0)$ und für $k \in \mathbb{N}$

$$a_k(f) = \hat{f}(k) + \hat{f}(-k), \quad b_k(f) = i(\hat{f}(k) - \hat{f}(-k),$$

also für $k \in \mathbb{N}$

$$\hat{f}(k) = \frac{a_k(f) - ib_k(f)}{2}, \quad \hat{f}(-k) = \frac{a_k(f) + ib_k(f)}{2}.$$

Wir schreiben noch die Darstellung der Bestapproximation von f aus dem Beispiel 16.7 oben um:

$$\begin{aligned}
\sum_{|k|\leq n} \hat{f}(k)e^{ikt} &= \sum_{|k|\leq n} \hat{f}(k)(\cos(kt) + i\sin(kt)) \\
&= \hat{f}(0) + \sum_{k=1}^{n}(\hat{f}(k) + \hat{f}(-k))\cos(kt) + \sum_{k=1}^{n} i(\hat{f}(k) - \hat{f}(-k))\sin(kt) \\
&= \frac{a_0(f)}{2} + \sum_{k=1}^{n}(a_k(f)\cos(kt) + b_k(f)\sin(kt)).
\end{aligned}$$

Die Frage ist, ob und ggf. in welchem Sinne dies für $n \to \infty$ gegen die Funktion f konvergiert. Für $n = \infty$ heißen diese Reihen *Fourierreihe von f*.

16.8. Besselsche Ungleichung und Parsevalsche Gleichung: Wir setzen das Beispiel von eben fort: wegen $(f - Pf) \perp Pf$ (vgl. 16.7) gilt

$$\|f\|^2 = \|f - Pf + Pf\|^2 = \|f - Pf\|^2 + \|Pf\|^2 = \|f - Pf\|^2 + \sum_{|k|\leq n} |(f|e_k)|^2.$$

Daraus folgt zum einen die *Besselsche Ungleichung*

$$\sum_{k=-\infty}^{\infty} |(f|e_k)|^2 \leq \|f\|^2,$$

die analog für jedes durch $\mathbb{N}$ oder $\mathbb{Z}$ indizierte Orthonormalsystem gilt. Bezeichnen wir die Orthogonalprojektion auf U_n mit P_n, so sieht man zum anderen, dass $\lim_{n\to\infty} \|f - P_n f\| = 0$ äquivalent ist zur *Parsevalschen Gleichung*

$$\|f\|^2 = \sum_{k=-\infty}^{\infty} |(f|e_k)|^2.$$

Ein Orthonormalsystem heißt *vollständig* in V, wenn dies für alle $f \in V$ gilt.

Bemerkung: Das Orthonormalsystem $(e_k)_{k\in\mathbb{Z}}$ ist vollständig im Raum der stückweise stetigen und normalisierten Funktionen $f : [-\pi, \pi) \to \mathbb{C}$, dh für alle diese Funktionen gilt:

$$\|f\|^2 = \frac{1}{2\pi}\int_{-\pi}^{\pi} |f(t)|^2\,dt = \sum_{k=-\infty}^{\infty} |\hat{f}(k)|^2$$

und

$$\|f - \sum_{|k| \leq n} \hat{f}(k) e_k\| \to 0 \; (n \to \infty).$$

Letzteres bezeichnet man als Konvergenz *im quadratischen Mittel* und schreibt dann auch $f = \sum_{k=-\infty}^{\infty} \hat{f}(k) e_k$.