

Höhere Mathematik I für die
Fachrichtung Physik
Wintersemester 2023/24

Gerd Herzog, Peer Kunstmann, Christoph Schmoeger
Karlsruher Institut für Technologie
Institut für Analysis

Version: February 13, 2024

1 Logische Grundlagen

Wir beginnen mit Grundlagen der *Aussagenlogik*.

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

Beispiele:

- (1) $2 + 3 = 7$. (f)
- (2) 2 ist eine Primzahl. (w)
- (3) Berechnen Sie bitte $7 - 15$. (keine Aussage)
- (4) Die Lichtgeschwindigkeit beträgt ca. $300000 \frac{km}{s}$. (keine Aussage im mathematischen Sinn).

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

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

(a)

$A \wedge B$ (logisches “und”)	A	w	w	f	f
	B	w	f	w	f
	$A \wedge B$	w	f	f	f

(b)

$A \vee B$ (logisches “oder”)	A	w	w	f	f
	B	w	f	w	f
	$A \vee B$	w	w	w	f

Achtung: Das logische “oder” ist nicht exklusiv, d.h. 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: 9 ist eine Primzahl. (w)

(e)

<i>Äquivalenz</i> $A \Leftrightarrow B$	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:

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

Folgende Regeln sind allgemeingültig:

- $\neg(\neg A) \Leftrightarrow A$ (doppelte Negation)
- $(A \Leftrightarrow B) \Leftrightarrow [(A \Rightarrow B) \wedge (B \Rightarrow A)]$ (Äquivalenz bedeutet zwei Implikationen)
- $\neg(A \wedge B) \Leftrightarrow (\neg A \vee \neg B)$ (Negation von “und”)
- $\neg(A \vee B) \Leftrightarrow (\neg A \wedge \neg B)$ (Negation von “oder”)
- $(A \Rightarrow B) \Leftrightarrow (\neg A \vee B)$ (Umformulierung der Implikation)
- $\neg(A \Rightarrow B) \Leftrightarrow (A \wedge \neg B)$ (Negation der Implikation)
- $(A \Rightarrow B) \Leftrightarrow (\neg B \Rightarrow \neg A)$ (Kontraposition)

Beispiel: Wenn 9 keine Primzahl ist, dann ist $1 \neq 2$. (w)

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 $A \vee (B \wedge C) \Leftrightarrow (A \vee B) \wedge (A \vee C)$.

Ein Drittes gibt es nicht: $A \vee \neg A$ ist stets wahr (*tertium non datur*).

Beispiel: Eine Verknüpfung, welche das “exklusive Oder”

A	w	w	f	f
B	w	f	w	f
$\vee$	f	w	w	f

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).

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.

Beispiele: (1) $A(x) = \text{“}x \text{ ist eine Primzahl“}$. (2) $A(x, y) = \text{“}x + y = 10\text{“}$.

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:

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

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

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

Deren Negation ist

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

also “Es gibt eine Primzahl, die nicht ungerade ist” (w).

Achtung: Bei mehreren Quantoren kommt es i.a. auf die Reihenfolge an!

Beispiel: Betrachtet man für x und y natürliche Zahlen sowie die Aussageform $A(x, y) = \text{“}x < y\text{“}$, so ist

$$\forall x \exists y : x < y$$

wahr, aber

$$\exists y \forall x : x < y$$

ist falsch.

2 Mengen

2.1. Der Begriff der Menge: Wir verwenden die folgende Definition (nach Georg Cantor):

“Eine *Menge* M ist eine Zusammenfassung von bestimmten, wohlunterschiedenen Objekten unserer Anschauung oder unseres Denkens (welche die *Elemente* von M genannt werden) zu einem Ganzen.”

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

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

“ $x \in M$ ” ist also eine Aussageform, d.h. 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: Definition: Es seien M und N Mengen.

- (1) $M \subseteq N :\Leftrightarrow \forall x : (x \in M \Rightarrow x \in N)$ (M ist Teilmenge von N)
- (2) $M = N :\Leftrightarrow (M \subseteq N) \wedge (N \subseteq M)$ (Gleichheit von Mengen)
- (3) $M \subsetneq N :\Leftrightarrow (M \subseteq N) \wedge (N \neq M)$ (echte Teilmenge)
- (4) $M \cap N := \{x : x \in M \wedge x \in N\}$ (Durchschnitt)
- (5) $M \cup N := \{x : x \in M \vee x \in N\}$ (Vereinigung)
- (6) $M \setminus N := \{x : x \in M \wedge x \notin N\}$ (Differenz, “ M ohne N ”)

Beispiele: Wir schreiben $\mathbb{N} = \{1, 2, 3, \dots\}$ Menge der natürlichen Zahlen ($\mathbb{N}$ wird später exakt eingeführt).

$$\{x \in \mathbb{N} : x \text{ ist gerade Primzahl}\} = \{2\},$$

$$\{x \in \mathbb{N} : x \text{ ist Primzahl} \wedge x > 2\} \subseteq \{x \in \mathbb{N} : x \text{ ist ungerade}\}.$$

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

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

Regeln für Durchschnitt und Vereinigung: Wegen 1.3 gelten:

Kommutativität:

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

Assoziativität:

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

Distributivität:

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

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

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

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

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

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

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

$$\text{Pot}(M) := \{N : N \subseteq M\}$$

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

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

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

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

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

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

Beispiele:

- (1) Für $M_1 = \{0, 1\}$, $M_2 = \{1, 2, 3\}$ ist

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

- (2) Für $M = \{0, 1\}$ ist

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

- (3) Ist M eine Menge und $N = \emptyset$, so ist $M \times N = \emptyset$.

2.7. Relationen: Es seien X, Y Mengen und $R \subseteq X \times Y$. Dann heißt R eine *Relation*. Statt $(x, y) \in R$ schreibt man auch xRy (x und y stehen in Beziehungen (Relation) zueinander).

Beispiele:

- (1) $R = \{(n, m) \in \mathbb{N} \times \mathbb{N} : n \text{ ist ein Teiler von } m\}$. Es gilt $2R4$, hingegen $3R4$ gilt nicht.
(2) Es sei X eine Menge und $R = \{(N, M) : N \subseteq M \subseteq X\} \subseteq \text{Pot}(X) \times \text{Pot}(X)$. Hier ist $NRM \Leftrightarrow N \subseteq M$.

2.8. Definition: Es sei X eine Menge und $R \subseteq X^2$ eine Relation.

- (1) R heißt *reflexiv* $:\Leftrightarrow \forall x \in X : xRx$.
(2) R heißt *symmetrisch* $:\Leftrightarrow (xRy \Rightarrow yRx)$,
(3) R heißt *transitiv* $:\Leftrightarrow (xRy \wedge yRz \Rightarrow xRz)$,
(4) R heißt *antisymmetrisch* $:\Leftrightarrow (xRy \wedge yRx \Rightarrow x = y)$,
(5) R heißt eine *Äquivalenzrelation* $:\Leftrightarrow R$ ist reflexiv, symmetrisch und transitiv,
(6) R heißt eine *Ordnungsrelation* $:\Leftrightarrow R$ ist reflexiv, antisymmetrisch und transitiv.

Beispiele:

- (1) R in obigem Beispiel (2) ist eine Ordnungsrelation. R ist nicht symmetrisch, also keine Äquivalenzrelation.
(2) $X = \mathbb{Z}$, $R = \{(n, m) \in \mathbb{Z}^2 : n - m \text{ ist gerade}\}$ ist eine Äquivalenzrelation. R ist nicht antisymmetrisch, also keine Ordnungsrelation.

Ist R eine Äquivalenzrelation, so schreibt man auch $x \sim y$ für xRy .

Ist R eine Ordnungsrelation, so schreibt man auch $x \leq y$ für xRy .

Es sei $R \subseteq X \times X$ eine Äquivalenzrelation. Für jedes $x \in X$ heißt die Menge

$$[x] := [x]_R := \{y \in X : x \sim y\}$$

die Äquivalenzklasse von x bzgl. R und jedes $y \in [x]$ heißt dann ein *Repräsentant* der Äquivalenzklasse $[x]$.

Es gilt für alle $x, y \in X$ dann $x \in [x]$ ($\sim$ ist reflexiv) und

$$[x] \cap [y] \neq \emptyset \Rightarrow [x] = [y], \quad [x] = [y] \Leftrightarrow x \sim y,$$

da R symmetrisch und transitiv ist. Wir setzen

$$X/\sim := \{[x] : x \in X\}.$$

In obigem Beispiel (2) gilt:

1. Fall n ist gerade, also $n \sim 0$ und somit $[n] = [0]$.

2. Fall n ist ungerade, also $n \sim 1$ und somit $[n] = [1]$.

Damit ist $X/\sim := \{[0], [1]\}$.

Beispiel: Für eine endliche Menge M bezeichne $\#M$ die Anzahl ihrer Elemente.

Es sei X eine endliche Menge, $R \subseteq X^2$ eine Relation und $M, N \subseteq X$. Dann gilt

$$\#(M \times N) \cap R = \sum_{x \in M} \#\{y \in N : xRy\} = \sum_{y \in N} \#\{x \in M : xRy\}.$$

Um diese Gleichung einzusehen schreibe man $M \times N$ als rechteckige Punktemenge. Die Elemente von $(M \times N) \cap R$ werden dann spaltenweise bzw. zeilenweise abgezählt.

Nun sei X die Menge der Felder eines Minesweeper Brettes und xRy die Relation “ x ist ein Nachbarfeld von y ”. Ist M eine Belegung mit Minen und $N := X \setminus M$ (N sind also die leeren Felder) so ist

$$\sum_{y \in N} \#\{x \in M : xRy\}$$

die Summe der Zahlen auf dem Brett (für $y \in N$ ist $\#\{x \in M : xRy\}$ die Anzahl der Minen auf den Nachbarfeldern von y). Obige Gleichung zeigt: Geht man zum inversen Brett über (d.h. man vertauscht Minen und leere Felder) so ändert sich die Summe der Zahlen nicht.

*	2	*	1	3	*	5	*
1	3	2	2	*	*	*	*
0	1	*	1	*	*	8	*
0	1	1	1	*	*	*	*

3 Funktionen

3.1. Zum Begriff der Funktion: Es seien X, Y Mengen. Eine *Funktion* (oder *Abbildung*) $f : X \rightarrow Y$ 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 $f : X \rightarrow Y, x \mapsto f(x)$ (“ f von X nach Y , x wird abgebildet auf $f(x)$ ”).

Beispiel: $f : \mathbb{N} \rightarrow \mathbb{N}, x \mapsto 2 \cdot x - 1$, dann ist z.B. $f(1) = 2 \cdot 1 - 1 = 1$, $f(2) = 2 \cdot 2 - 1 = 3$, 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. Eine Funktion von X nach Y ist also eine Relation $R \subseteq X \times Y$, welche die folgende Eigenschaften hat:

$$\begin{aligned} \forall x \in X \exists y \in Y : xRy, \\ \forall x \in X \forall y_1, y_2 \in Y : xRy_1 \wedge xRy_2 \Rightarrow y_1 = y_2. \end{aligned}$$

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

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

und für $B \subseteq Y$ heißt

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

Insbesondere heißt $f(X) = \{f(x) : x \in X\}$ *Bild von f* .

Im Beispiel oben ist z.B.

$$f(\mathbb{N}) = \{y \in \mathbb{N} : y \text{ ist ungerade}\}, \quad f(\{3, 8\}) = \{5, 15\},$$

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

3.2. Definition: Es sei $f : X \rightarrow Y$ eine Funktion.

(a) f heißt *surjektiv*, falls $f(X) = Y$ gilt, d.h. falls jedes $y \in Y$ ein Funktionswert $f(x)$ ist.

(b) f heißt *injektiv*, falls gilt

$$\forall x_1, x_2 \in X : x_1 \neq x_2 \Rightarrow f(x_1) \neq f(x_2),$$

d.h. 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 $f : \mathbb{N} \rightarrow \mathbb{N}$ von oben ist injektiv und nicht surjektiv.

(2) Die Funktion

$$g : \mathbb{N} \rightarrow \{0, 1\}, \quad x \mapsto \begin{cases} 0, & x \text{ ist gerade} \\ 1, & x \text{ ist ungerade} \end{cases}$$

ist surjektiv und nicht injektiv.

Bemerkung: Streng genommen ist Injektivität eine Eigenschaft einer Funktion, aber Surjektivität und damit auch Bijektivität eine Eigenschaft einer Funktion zusammen mit ihrem Wertebereich. Z.B. sind $f : \{1, 2, 3\} \rightarrow \{2, 4, 6\}$, $x \mapsto f(x) = 2x$ und $g : \{1, 2, 3\} \rightarrow \mathbb{N}$, $x \mapsto g(x) = 2x$ dieselbe Funktion, denn f und g haben denselben Graphen, nämlich

$$\{(1, 2), (2, 4), (3, 6)\}.$$

Aber bzgl. des jeweiligen Wertebereichs ist f bijektiv, aber g nicht.

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

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

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

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

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

d.h. die Hintereinanderausführung von Funktionen ist *assoziativ*.

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

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

eine Funktion $f^{-1} : Y \rightarrow X$, 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. In diesem Fall hat man eine Doppelbedeutung von f^{-1} . Es gilt dann aber (links Urbild, rechts Umkehrabbildung):

$$f^{-1}(\{y\}) = \{f^{-1}(y)\} \quad (y \in Y).$$

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

Bemerkung: Ist $f : X \rightarrow Y$ bijektiv, so gilt

$$f^{-1} \circ f = \text{id}_X, \quad f \circ f^{-1} = \text{id}_Y.$$

Denn: Für jedes $x \in X$ gilt $f^{-1}(f(x)) = x$ nach Definition von f^{-1} . Für $y \in Y$ und $x := f^{-1}(y)$ ist $f(x) = y$ (vgl. 3.4), also $y = f(x) = f(f^{-1}(y))$.

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

Beweis. Die Funktion $g \circ f$ ist injektiv, also ist f injektiv. Die Funktion $f \circ g$ ist surjektiv, also ist f surjektiv. Somit ist f bijektiv. Nach 3.4 existiert die Umkehrabbildung $f^{-1} : Y \rightarrow X$. Noch zu zeigen: $\forall y \in Y : g(y) = f^{-1}(y)$. Sei $y \in Y$. Setze $x := g(y)$. Dann gilt $f(x) = f(g(y)) = y$ nach Voraussetzung und demnach $x = f^{-1}(y)$ nach 3.4. $\square$

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

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

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

Diese Formel heit auch ‘‘Hemd-Jacken-Regel’’ (Beweis als bung).

4 Die reellen Zahlen

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

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

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

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 “ $\cdot$ ” 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 $\forall a \in \mathbb{R} : a + \tilde{0} = a$, so folgt

$$0 = 0 + \tilde{0} \stackrel{(A4)}{=} \tilde{0} + 0 \stackrel{(A2)}{=} \tilde{0}.$$

□

(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. Es sei $a \in \mathbb{R}$. Wir setzen $b := a \cdot 0$. Dann gilt:

$$b = a \cdot 0 \stackrel{(A2)}{=} a \cdot (0 + 0) \stackrel{(A9)}{=} a \cdot 0 + a \cdot 0 = b + b.$$

Weiter folgt

$$0 \stackrel{(A3)}{=} b + (-b) = (b + b) + (-b) \stackrel{(A1)}{=} b + (b + (-b)) \stackrel{(A3)}{=} b + 0 \stackrel{(A2)}{=} b.$$

□

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

Beweis. Es gilt

$$a + a \cdot (-1) =_{(A6)} a \cdot 1 + a \cdot (-1) =_{(A9)} a \cdot (1 + (-1)) = a \cdot 0 =_{(3)} 0,$$

und mit (2) folgt $-a = a \cdot (-1) =_{(A8)} (-1) \cdot a$. □

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

Beweis. Es ist

$$(-a)^2 =_{(4)} (-a) \cdot ((-1) \cdot a) =_{(A5)} (-(-a)) \cdot a =_{(2)} a^2.$$

□

4.2. Anordnungsaxiome: In $\mathbb{R}$ ist eine *Ordnung* “ $\leq$ ” (d.h. eine Relation in $\mathbb{R} \times \mathbb{R}$) gegeben mit folgenden Eigenschaften:

$$(A10) \quad \forall a, b \in \mathbb{R} : a \leq b \vee b \leq a,$$

$$(A11) \quad \forall a, b, c \in \mathbb{R} : a \leq b \wedge b \leq c \Rightarrow a \leq c,$$

$$(A12) \quad \forall a, b \in \mathbb{R} : a \leq b \wedge b \leq a \Rightarrow a = b,$$

$$(A13) \quad \forall a, b, c \in \mathbb{R} : a \leq b \Rightarrow a + c \leq b + c,$$

$$(A14) \quad \forall a, b, c \in \mathbb{R} : a \leq b \wedge 0 \leq c \Rightarrow ac \leq bc.$$

(A11) heißt Transitivität, (A12) heißt Antisymmetrie. (A10) bedeutet, dass die Ordnung “total” ist (d.h. 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 $a \in \mathbb{R}$.

1. 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).

2. 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 $a^2 = (-a)^2 \geq 0$. □

(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 man $ac + bc$ zu der Ungleichung so folgt mit (A13) die Ungleichung $bc \leq ac$. $\square$

Intervalle: Es seien $a, b \in \mathbb{R}$ mit $a \leq b$. Wir setzen:

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

Beachte: Im Fall $a = b$ ist $[a, b] := \{a\}$ und $(a, b) = [a, b) = (a, b] = \emptyset$.

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(1)), $|-2| = -(-2) = 2$, 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: Es 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 $(|a| \leq c \Leftrightarrow (a \leq c \text{ und } -a \leq c))$;
- (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 $||a| - |b|| \leq |a - b|$. $\square$

4.4. Supremum und Infimum: Es sei $M \subseteq \mathbb{R}$.

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:

(1) 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 $|a| = \max\{a, -a\}$.

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

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

(4) $M = \emptyset$ ist nach oben und nach unten beschränkt. Jede reelle Zahl ist obere und untere Schranke von M .

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 und wird mit $\sup M$ [$\inf M$] bezeichnet.

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

Bemerkungen:

(1) Das Supremum [Infimum] einer Menge M ist also die **kleinste** obere [**größte** untere] Schranke von M .

(2) 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 keine 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. Es sei $\emptyset \neq M \subseteq \mathbb{R}$ nach unten beschränkt. Setze $-M := \{-x : x \in M\}$. Beachte $-M \neq \emptyset$.

Es sei γ eine untere Schranke von M , also $\gamma \leq x$ ($x \in M$). Dann ist $-x \leq -\gamma$ ($x \in M$), d.h. $-M$ ist nach oben beschränkt. Nach (A15) existiert $s := \sup(-M)$. Wegen

$$\forall x \in M : x \geq -s$$

ist $-s$ eine untere Schranke von M . Da $-\gamma$ eine obere Schranke von $-M$ ist folgt $s \leq -\gamma$, also $\gamma \leq -s$, d.h. $-s = \inf M$. $\square$

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

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

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

- (1) A ist beschränkt $\Rightarrow \inf A \leq \sup A$.
- (2) A ist nach oben [nach unten] beschränkt $\Rightarrow B$ ist 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:

$$\begin{aligned} \gamma = \sup A &\iff \forall \varepsilon > 0 \exists x \in A : x > \gamma - \varepsilon \\ [\gamma = \inf A &\iff \forall \varepsilon > 0 \exists x \in A : x < \gamma + \varepsilon]. \end{aligned}$$

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 .

“ $\Leftarrow$ ” Sei $\gamma \neq \sup A =: \tilde{\gamma}$. Dann ist $\gamma > \tilde{\gamma}$, da γ obere Schranke von A ist, und $\varepsilon := \gamma - \tilde{\gamma} > 0$. Dann existiert ein $x \in A$ mit

$$x > \gamma - \varepsilon = \gamma - (\gamma - \tilde{\gamma}) = \tilde{\gamma} = \sup A,$$

ein Widerspruch. $\square$

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

Definition: $A \subseteq \mathbb{R}$ heißt *Induktionsmenge* (IM), falls $1 \in A$ und $\forall x \in A : x + 1 \in A$. Es sei $\mathcal{A}$ die Menge aller Induktionsmengen.

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

Definition: Die Menge

$$\mathbb{N} := \bigcap_{A \in \mathcal{A}} A = \{x \in \mathbb{R} : \forall A \in \mathcal{A} : x \in A\}$$

heißt *Menge der natürlichen Zahlen*. D.h. x ist eine natürlichen Zahl genau dann, wenn für **jede** Induktionsmenge $A \subseteq \mathbb{R}$ gilt: $x \in A$.

Beispiel: $\frac{3}{2} \notin \mathbb{N}$, denn $\frac{3}{2} \notin \{1\} \cup [2, \infty)$.

Satz: (1) $\mathbb{N} \in \mathcal{A}$ (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 $\varepsilon \in \mathbb{R}$ mit $\varepsilon > 0$ gibt es $n \in \mathbb{N}$ mit $\frac{1}{n} < \varepsilon$.

Beweis. (1) Es ist $1 \in \mathbb{N}$, da $1 \in A$ für jedes $A \in \mathcal{A}$. Sei $x \in \mathbb{N}$. Zu zeigen: $x + 1 \in \mathbb{N}$. Sei dazu $A \in \mathcal{A}$. 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 $\gamma := \sup \mathbb{N}$ und $\gamma - 1$ ist keine obere Schranke von $\mathbb{N}$. Also existiert ein $n \in \mathbb{N}$ mit $n > \gamma - 1$. Dann gilt $n + 1 > \gamma$ und $n + 1 \in \mathbb{N}$, d.h. γ ist nicht obere Schranke von $\mathbb{N}$, Widerspruch. (3) folgt sofort aus (2).

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

4.8. Vollständige Induktion:

Satz: Ist $A \subseteq \mathbb{N}$ und ist $A \in \mathcal{A}$, so ist $A = \mathbb{N}$.

Beweis. Nach Voraussetzung gilt $A \subseteq \mathbb{N}$. Wegen $A \in \mathcal{A}$ gilt $\mathbb{N} \subseteq A$. □

Beweisverfahren durch Induktion

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

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

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

Beweis. Setze $A := \{n \in \mathbb{N} : A(n)\}$. Beachte $A \subseteq \mathbb{N}$. Nach (IA) gilt $1 \in A$. Sei $n \in A$. Dann gilt $A(n)$ und nach (IS) ist auch $A(n+1)$ wahr, d.h. $n+1 \in A$. Somit ist $A \in \mathcal{A}$ und $A = \mathbb{N}$ folgt aus obigem Satz. □

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

Beweis. Wir führen den Beweis durch Induktion nach n :

Induktionsanfang (IA): Es gilt $1 \geq 1$, d.h. $A(1)$ ist wahr.

Induktionsschluss (IS): Es sei $n \in \mathbb{N}$ und es gelte $A(n)$, d.h. es gelte $n \geq 1$ (Induktionsvoraussetzung (IV)). Dann ist $n + 1 \geq 1 + 1$ nach (IV) und es gilt $1 + 1 \geq 1 + 0 = 1$, also $n + 1 \geq 1$. Damit ist $A(n + 1)$ ist wahr. $\square$

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

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

(IS) Es sei $n \in \mathbb{N}$ und es gelte

$$1 + 2 + \dots + n = \frac{n(n+1)}{2} \quad (IV).$$

Zu zeigen ist

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

Es gilt

$$1 + 2 + \dots + n + (n+1) \stackrel{(IV)}{=} \frac{n(n+1)}{2} + n + 1 = \frac{(n+1)(n+2)}{2}.$$

$\square$

Wir setzen die natürlichen Zahlen ab jetzt als bekannt voraus, einschließlich ihrer bekannten Eigenschaften, wie z.B. $\forall n, m \in \mathbb{N} : m + n, mn \in \mathbb{N}$, etc.

Definition durch Rekursion (bzw. durch Induktion)

Für jedes $n \in \mathbb{N}$ soll ein mathematisches Objekt (Zahl, Term, etc.) $T(n)$ definiert werden. Es sei $T(1)$ definiert, und für jedes $n \in \mathbb{N}$ sei $T(n+1)$ definiert unter der Voraussetzung, dass $T(1), T(2), \dots, T(n)$ schon definiert sind.

Dann ist $T(n)$ für jedes $n \in \mathbb{N}$ definiert, denn $\{n \in \mathbb{N} : T(1), \dots, T(n) \text{ sind definiert}\} \in \mathcal{A}$.

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

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

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

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

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

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

Dann ist $\prod_{j=1}^n a_j = a_1 \cdot a_2 \cdot \dots \cdot a_n$ 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 $a^n = \underbrace{a \cdot a \cdot \dots \cdot a}_{n \text{ Faktoren}}$ 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:

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

Bei der *Abschnittsinduktion* zeigt man

$$\begin{aligned} (IA) \quad & A(1) \\ (IS) \quad & \forall n \in \mathbb{N} : (A(1) \wedge A(2) \wedge \dots \wedge A(n) \Rightarrow A(n+1)). \end{aligned}$$

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

Beispiel: $\forall n \in \mathbb{N}$ mit $n \geq 5 : 2^n > n^2$.

Beweis durch Induktion. (IA) $n = 5$: Es gilt $2^5 = 32 > 25 = 5^2$.

(IS) Es sei $n \in \mathbb{N}$, $n \geq 5$ und es gelte $2^n \geq n^2$ (IV). Dann gilt:

$$2^{n+1} = 2 \cdot 2^n >_{(IV)} 2n^2 = (n^2 + 2n + 1) + ((n-1)^2 - 2) \geq (n+1)^2 + 14 > (n+1)^2.$$

□

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 z.B. die nichtleere nach oben beschränkte Menge $M := \{x \in \mathbb{Q} : x^2 < 2\}$ kein Supremum in $\mathbb{Q}$, d.h. das Vollständigkeitsaxiom (A15) gilt in $\mathbb{Q}$ **nicht!**

Satz: Ist $\emptyset \neq A \subseteq \mathbb{Z}$ und ist A nach unten beschränkt, so existiert $\min A$. (ohne Beweis)

Satz: Sind $x, y \in \mathbb{R}$ mit $x < y$, so gilt:

- (1) Ist $y - x > 1$, so gibt es ein $p \in \mathbb{Z}$ mit $x < p < y$.
- (2) $\exists r \in \mathbb{Q} : x < r < y$.

Beweis. (1) Die Menge $A := \{m \in \mathbb{Z} : x < m\}$ ist nicht leer und nach unten beschränkt. Somit existiert $p := \min A$. Damit gilt $x < p$ und $p - 1 \leq x$, also $x < p \leq x + 1 < y$.

(2) Wegen $y - x > 0$ gilt: $\exists n \in \mathbb{N} : \frac{1}{n} < y - x$, also $1 < ny - nx$. Nach (1) gilt:

$$\exists p \in \mathbb{Z} : nx < p < ny.$$

Für $r := \frac{p}{n}$ gilt nun $x < r < y$. □

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 ”).

Beispiele:

$$\binom{n}{0} = 1 = \binom{n}{n} \quad (n \in \mathbb{N}_0), \quad \binom{n}{k} = \binom{n}{n-k} \quad (n \geq k \geq 0), \quad \binom{4}{2} = \frac{4 \cdot 3}{2} = 6.$$

Lemma: Für $1 \leq k \leq n$ gilt

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

Beweis. Es gilt

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

□

Aus obigem Lemma ergibt sich das *Pascalsche Dreieck*.

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]

Z.B. $n = 3$: $a^3 - b^3 = (a - b)(a^2 + ab + b^2)$.

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

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

Z.B. (mit $a = b = 1$):

$$\sum_{k=0}^n \binom{n}{k} = 2^n \quad (n \in \mathbb{N}_0),$$

oder (mit $a = -1$ und $b = 1$):

$$\sum_{k=0}^n (-1)^k \binom{n}{k} = (-1 + 1)^n = 0 \quad (n \in \mathbb{N}).$$

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

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

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

□

(3) **Bernoullische Ungleichung** (BU): Es 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 $a \in \mathbb{R}$.

(i) Ist $a > 1$, so gibt es zu jedem $c > 0$ ein $n \in \mathbb{N}$ mit $a^n > c$.

(ii) Ist $a \in (0, 1)$, so gibt es zu jedem $\varepsilon > 0$ ein $n \in \mathbb{N}$ mit $a^n < \varepsilon$.

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

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

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

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

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

[Übungsaufgabe, verwende (1)]

(6) Mit Hilfe der Bernoullische Ungleichung kann gezeigt werden: Für jedes $n \in \mathbb{N}$ gilt

$$\left(1 + \frac{1}{n}\right)^n < \left(1 + \frac{1}{n+1}\right)^{n+1} \quad \text{und} \quad \left(1 + \frac{1}{n}\right)^{n+1} > \left(1 + \frac{1}{n+1}\right)^{n+2}.$$

[Beweis in den Übungen.]

4.12. Wurzeln: Es 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: $\sqrt[n]{a} := b$ heißt die "*n-te Wurzel aus a*".

Beweis. Eindeutigkeit: Es seien $b, c \geq 0$ mit $b^n = c^n = a$. Mit 4.11 (5) folgt $b = c$. Die Existenz folgt später aus dem Zwischenwertsatz 9.10. □

Bemerkungen:

(1) $\sqrt{a} := \sqrt[n]{a}$.

(2) Z.B. existiert die reelle Zahl $\sqrt{2}$ (bekannt: $\sqrt{2} \notin \mathbb{Q}$).

(3) In $\mathbb{R}$ ziehen wir Wurzeln nur aus Zahlen ≥ 0 und Wurzeln sind dann Zahlen ≥ 0 . Z.B. ist $\sqrt{4} = 2$, $\sqrt{4} \neq -2$. Die Gleichung $x^2 = 4$ hat zwei Lösungen: $x = \sqrt{4} = 2$ und

$$x = -\sqrt{4} = -2.$$

(4) **Wichtig:**

$$\forall x \in \mathbb{R} : \sqrt{x^2} = |x|.$$

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

$$a \leq b \iff \sqrt[n]{a} \leq \sqrt[n]{b}, \quad a < b \iff \sqrt[n]{a} < \sqrt[n]{b}$$

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

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

4.13. Ungleichung zwischen harmonischem, geometrischem und arithmetischem

Mittel: Sei $n \in \mathbb{N}$. Sind $a_1, a_2, \dots, a_n > 0$, so gilt

$$\frac{n}{\frac{1}{a_1} + \dots + \frac{1}{a_n}} \leq \sqrt[n]{a_1 \cdot a_2 \cdot \dots \cdot a_n} \leq \frac{a_1 + a_2 + \dots + a_n}{n}.$$

Ohne Beweis.

Bemerkung: Die 2. Ungleichung gilt auch für $a_1, a_2, \dots, a_n \geq 0$.

Beispiel: $a_k = k$ ($k = 1, \dots, n$). Mit der zweiten Ungleichung folgt

$$\sqrt[n]{n!} \leq \frac{1}{n} \sum_{k=1}^n k = \frac{1}{n} \cdot \frac{n(n+1)}{2} = \frac{n+1}{2}.$$

5 Die komplexen Zahlen

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

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

Man kann nachrechnen: $\mathbb{R}^2$ ist mit diesen Verknüpfungen ein Körper (d.h. (A1) – (A9) gelten für diese Verknüpfungen, wobei “*” die Rolle der Multiplikation übernimmt) mit Nullelement $(0, 0)$, Einselement $(1, 0)$, additiver Inverse $-(x, y) = (-x, -y)$ und multiplikativer Inverse

$$\left(\frac{x}{x^2 + y^2}, -\frac{y}{x^2 + y^2} \right) \quad \text{für } (x, y) \neq (0, 0).$$

Es gilt $(x_1, 0) + (x_2, 0) = (x_1 + x_2, 0)$ und $(x_1, 0) * (x_2, 0) = (x_1 x_2, 0)$, d.h. man kann das Paar $(x, 0)$ mit der reellen Zahl x identifizieren. Somit sind die reellen Zahlen $(\mathbb{R}, +, \cdot)$ im Körper $(\mathbb{R}^2, +, *)$ enthalten.

Weiter ist $(0, 1) * (0, 1) = (-1, 0)$, d.h. $(0, 1)$ ist eine “Zahl”, deren Quadrat $= -1$ ist. Man setzt nun $i := (0, 1)$ und schreibt (x, y) als

$$(x, y) = (x, 0) + (0, 1) * (y, 0) = x + iy \quad (= x + yi).$$

Die Menge

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

heißt *Körper 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 $\operatorname{Im} z$).

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

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

5.2. Konjugation und Betrag: Zu einer komplexen Zahl $z = x + iy$ mit $x, y \in \mathbb{R}$ heißt $\bar{z} := x - iy$ die *konjugiert komplexe Zahl*. Es gilt $z\bar{z} = \bar{z}z = x^2 + y^2 \geq 0$ und

$$|z| := \sqrt{z\bar{z}} = \sqrt{x^2 + y^2}$$

heißt *Betrag* der komplexen Zahl z .

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

- (1) $\overline{\bar{z}} = z$, (2) $|\bar{z}| = |z|$, (3) $\overline{z + w} = \bar{z} + \bar{w}$, (4) $\overline{z \cdot w} = \bar{z} \cdot \bar{w}$,
- (5) $\operatorname{Re} z = \frac{1}{2}(z + \bar{z})$ und $\operatorname{Im} z = \frac{1}{2i}(z - \bar{z})$,
- (6) $\max\{|\operatorname{Re} z|, |\operatorname{Im} z|\} \leq |z| \leq |\operatorname{Re} z| + |\operatorname{Im} z|$,
- (7) $|z \cdot w| = |z| \cdot |w|$, (8) $|z + w| \leq |z| + |w|$, (9) $||z| - |w|| \leq |z - w|$.

Beweis: (1)-(5) nachrechnen.

(6) folgt aus

$$\max\{|x|, |y|\} \leq \sqrt{x^2 + y^2} \leq |x| + |y| \quad (x, y \in \mathbb{R}).$$

(7) folgt aus

$$|zw|^2 = zw\overline{zw} = z\overline{z}w\overline{w} = |z|^2|w|^2.$$

(8) Es ist

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

also $|z+w| \leq |z| + |w|$. (9) folgt aus (8). $\square$

Division in $\mathbb{C}$: Für $z \neq 0$ ist

$$\frac{1}{z} = \frac{\overline{z}}{z\overline{z}} = \frac{\overline{z}}{|z|^2}.$$

Beispiel:

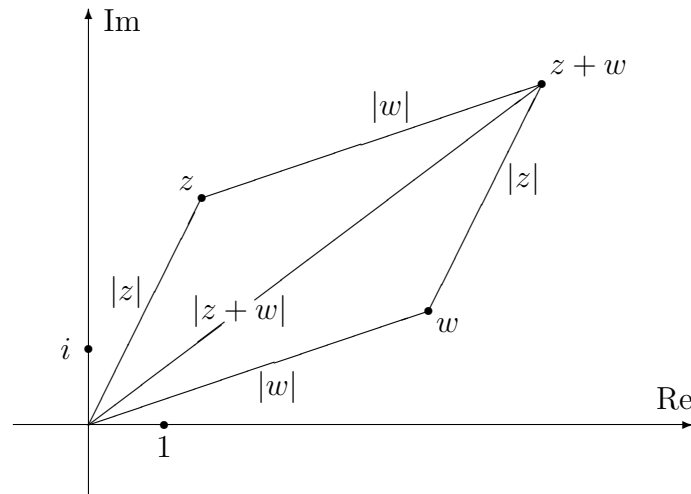
$$\frac{2+3i}{1+i} = \frac{(2+3i)(1-i)}{(1+i)(1-i)} = \frac{2-2i+3i+3}{2} = \frac{5+i}{2} = \frac{5}{2} + \frac{1}{2}i.$$

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° nach links, Multiplikation mit $a+ib$ bedeutet eine Drehstreckung.

Dreiecksungleichung:



5.4. Polynome: Ein Polynom p (oder $p(z)$) mit komplexen Koeffizienten ist ein Ausdruck der Form

$$p(z) = a_n z^n + a_{n-1} z^{n-1} + \dots + a_2 z^2 + a_1 z + a_0$$

mit $n \in \mathbb{N}_0$ und $a_j \in \mathbb{C}$ ($j \in \{0, 1, \dots, n\}$). Das Polynom heißt *reell*, wenn $a_j \in \mathbb{R}$ ($j \in \{0, 1, \dots, n\}$).

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

Es gilt insbesondere

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

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

Die Menge aller Polynome mit komplexen Koeffizienten bezeichnen wir mit $\mathbb{C}[z]$.

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

5.5. Satz und Definition (ohne Beweis): Es sei $p \in \mathbb{C}[z]$ ein Polynom mit Grad $n \geq 1$. Dann gilt:

(1) Ist $z_0 \in \mathbb{C}$ und $p(z_0) = 0$, so existieren eindeutig bestimmte $m \in \{1, \dots, n\}$ und $q \in \mathbb{C}[z]$ mit Grad $n - m$, so dass

$$p(z) = (z - z_0)^m q(z) \text{ und } q(z_0) \neq 0.$$

Dabei heißt $z - z_0$ ein *Linearfaktor* von p und m die *Vielfachheit* der Nullstelle z_0 .

(2) (**Fundamentalsatz der Algebra**): Ist

$$p(z) = a_n z^n + a_{n-1} z^{n-1} + \dots + a_2 z^2 + a_1 z + a_0, \quad a_n \neq 0,$$

so gibt es eindeutig bestimmte Zahlen $z_1, \dots, z_n \in \mathbb{C}$ mit

$$p(z) = a_n \prod_{k=1}^n (z - z_k).$$

Ist z_0 eine Nullstelle von p , so kommt z_0 in der Liste $z_1, \dots, z_n$ so oft vor, wie ihre Vielfachheit angibt.

Folgerung: Ist p ein Polynom vom Grad $n \geq 1$, so besitzt p mindestens eine Nullstelle.

Beispiele: (1) $p(z) = z^3 + z^2 + z + 1$. Eine Nullstelle ist $z_0 = -1$ (erraten). Nun gilt

$$p(z) = z^3 + z^2 + z + 1 = (z + 1)(z^2 + 1) = (z + 1)(z^2 - i^2) = (z + 1)(z - i)(z + i).$$

Die Nullstellen von p sind also -1 , i und $-i$ und haben jeweils die Vielfachheit 1.

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

6 Folgen und Konvergenz

6.1. Definition: Es sei $X \neq \emptyset$ eine Menge, $p \in \mathbb{Z}$, $Z_p := \{n \in \mathbb{Z} : n \geq p\}$ und $a : Z_p \rightarrow X$ eine Funktion. Dann heißt a eine *Folge in X* . Wir schreiben a_n statt $a(n)$. Statt a schreiben wir $(a_n)_{n \geq p}$, oder $(a_n)_{n=p}^\infty$ oder $(a_n)_{n \in Z_p}$ oder $(a_p, a_{p+1}, a_{p+2}, \dots)$. Oft ist $p = 0$ (also $Z_p = \mathbb{N}_0$) oder $p = 1$ (also $Z_p = \mathbb{N}$). Im Fall $p = 1$ schreiben wir auch kurz (a_n) .

Vereinbarung: Bis auf weiteres sei stets $X = \mathbb{R}$ oder $X = \mathbb{C}$. Die folgenden Definitionen und Sätze formulieren wir nur für Folgen der Form $(a_n)_{n=1}^\infty$ (also $p = 1$). Alles gilt sinngemäß auch im allgemeinen Fall $(a_n)_{n=p}^\infty$. Den allgemeinen Fall kann man auf den Fall $p = 1$ zurückführen: Setze $b_n := a_{p-1+n}$ ($n \in \mathbb{N}$), also

$$(b_1, b_2, b_3, \dots) = (a_p, a_{p+1}, a_{p+2}, \dots).$$

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

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

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

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: Eine Folge (a_n) in X (also $X = \mathbb{R}$ oder $X = \mathbb{C}$) heißt *konvergent* $:\Leftrightarrow$ Es gibt ein $a \in X$ mit:

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

Die Zahl a heißt dann *Limes* oder *Grenzwert* der Folge (a_n) und wir schreiben

$$\lim_{n \rightarrow \infty} a_n = a \text{ oder } a_n \rightarrow a \text{ (} n \rightarrow \infty \text{) oder kurz } a_n \rightarrow a.$$

Ist $\lim_{n \rightarrow \infty} a_n = 0$, so heißt (a_n) eine *Nullfolge*.

Ist (a_n) nicht konvergent, so heißt (a_n) *divergent*.

Bemerkung: Es gilt

$$a_n \rightarrow a \Leftrightarrow |a_n - a| \rightarrow 0, \text{ insbesondere } a_n \rightarrow 0 \Leftrightarrow |a_n| \rightarrow 0.$$

Beispiele: (1) Sei $c \in X$ und $a_n = c$ ($n \in \mathbb{N}$). Dann gilt $|a_n - c| = 0$ ($n \in \mathbb{N}$) und somit $a_n \rightarrow c$ ($n \rightarrow \infty$).

(2) Es gilt

$$\lim_{n \rightarrow \infty} \frac{1}{n} = 0.$$

Beweis: Sei $\varepsilon > 0$. Nach 4.7(4) existiert ein $n_0 \in \mathbb{N}$ mit $\frac{1}{n_0} < \varepsilon$. Für jedes $n \geq n_0$ gilt dann:

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

□

(3) Die Folge $(a_n)_{n \in \mathbb{N}} := ((-1)^n)_{n \in \mathbb{N}}$ ist divergent: Es gilt

$$|a_{n+1} - a_n| = |(-1)^{n+1} - (-1)^n| = |(-1)^n(-1 - 1)| = 2 \quad (n \in \mathbb{N}).$$

Annahme: (a_n) konvergiert. Sei $a := \lim_{n \rightarrow \infty} a_n = a$. Für $\varepsilon = \frac{1}{2}$ gilt dann:

$$\exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : |a_n - a| < \frac{1}{2}.$$

Für $n \geq n_0$ gilt somit

$$2 = |a_{n+1} - a_n| = |a_{n+1} - a + a - a_n| \leq |a_{n+1} - a| + |a - a_n| < \frac{1}{2} + \frac{1}{2} = 1,$$

ein Widerspruch.

(4) Sei $b \in X$ und $a_n = b^n$ ($n \in \mathbb{N}$). Dann gilt:

$$(a_n) \text{ konvergiert} \Leftrightarrow |b| < 1 \vee b = 1.$$

Im Fall $|b| < 1$ gilt $b^n \rightarrow 0$: Sei $\varepsilon > 0$. Nach 4.11(4) existiert ein $n_0 \in \mathbb{N}$ mit $|b|^{n_0} < \varepsilon$. Für $n \geq n_0$ gilt dann

$$|b^n| = |b|^n \leq |b|^{n_0} < \varepsilon.$$

Im Fall $|b| = 1$ und $b \neq 1$ ist (a_n) divergent: Annahme (a_n) konvergiert gegen a . Es gilt

$$|a_{n+1} - a_n| = |b^{n+1} - b^n| = |b|^n |b - 1| = |b - 1| > 0.$$

Für $\varepsilon = \frac{|b-1|}{4}$ gilt:

$$\exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : |a_n - a| < \frac{|b-1|}{4}.$$

Für $n \geq n_0$ gilt somit

$$|b - 1| = |a_{n+1} - a_n| \leq |a_{n+1} - a| + |a - a_n| < \frac{|b-1|}{2},$$

ein Widerspruch. Im Fall $|b| > 1$ ist (a_n) divergent (vgl. folgenden Satz).

(5) Zu jedem $x \in \mathbb{R}$ gibt es eine Folge (r_n) in $\mathbb{Q}$ mit $r_n \rightarrow x$: Nach 4.9 finden wir zu jedem $n \in \mathbb{N}$ ein $r_n \in (x - 1/n, x + 1/n) \cap \mathbb{Q}$, also $|r_n - x| < 1/n$. Sei $\varepsilon > 0$. Nach Beispiel (2) existiert ein $n_0 \in \mathbb{N}$ mit $1/n < \varepsilon$ ($n \geq n_0$), also $|r_n - x| < \varepsilon$ ($n \geq n_0$).

Definition:

(1) Es sei $a \in \mathbb{R}$ [oder $a \in \mathbb{C}$] und $\varepsilon > 0$. Dann heit

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

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 $(a - \varepsilon, a + \varepsilon)$. Die ε -Umgebung von $a \in \mathbb{C}$ in $\mathbb{C}$ ist die (offene) Kreisscheibe mit Mittelpunkt a und Radius ε .

(2) Es sei $A(n)$ eine Aussageform ($n \in \mathbb{N}$).

$A(n)$ (gilt) *fr fast alle* (ffa) $n \in \mathbb{N} :\Leftrightarrow$

$$\exists n_0 \in \mathbb{N} \forall n \geq n_0 : A(n) \text{ ist wahr.}$$

Beispiele:

(1) Fr fast alle $n \in \mathbb{N}$ gilt $n^2 > 10$.

(2) Fr fast alle $n \in \mathbb{N}$ gilt $\frac{1}{n} < \frac{1}{100}$.

(3) Sind $A(n), B(n)$ Aussageformen ($n \in \mathbb{N}$) so gilt:

$$(A(n) \text{ ffa } n \in \mathbb{N}) \wedge (B(n) \text{ ffa } n \in \mathbb{N}) \Leftrightarrow A(n) \wedge B(n) \text{ ffa } n \in \mathbb{N}.$$

Umformulierung der Folgenkonvergenz:

Ist (a_n) eine Folge in X und $a \in X$, so gilt:

$$\lim_{n \rightarrow \infty} a_n = a \iff \forall \varepsilon > 0 : a_n \in U_\varepsilon(a) \text{ ffa } n \in \mathbb{N}.$$

Definition:

(1) Eine Menge $M \subseteq \mathbb{C}$ heit *beschrnkt* $:\Leftrightarrow$

$$\exists \gamma \geq 0 \forall x \in M : |x| \leq \gamma$$

(vgl. 4.5 fr $M \subseteq \mathbb{R}$).

(2) Eine Folge (a_n) in X heit *beschrnkt* $:\Leftrightarrow$

$$\{a_n : n \in \mathbb{N}\} \text{ ist beschrnkt } \Leftrightarrow \exists \gamma \geq 0 \forall n \in \mathbb{N} : |a_n| \leq \gamma.$$

Beispiele: (1) Betrachte $M = U_\varepsilon(a)$. Ist $z \in U_\varepsilon(a)$, so gilt

$$|z| \leq |z - a| + |a| \leq \varepsilon + |a| =: \gamma.$$

Also ist M beschrnkt.

(2) Die Folge $(a_n) = ((-1)^n)$ ist beschrnkt, denn $|a_n| = 1$ ($n \in \mathbb{N}$). Bekannt: (a_n) ist divergent.

(3) Sei $b \in X$ mit $|b| > 1$. Dann ist die Folge $(a_n) = (b^n)$ unbeschränkt: Annahme: (a_n) ist beschränkt. Dann existiert ein $\gamma > 0$ mit $|a_n| = |b|^n \leq \gamma$ ($n \in \mathbb{N}$). Dann ist

$$\left(\frac{1}{|b|}\right)^n \geq \frac{1}{\gamma} > 0 \quad (n \in \mathbb{N}).$$

Wegen $\frac{1}{|b|} < 1$ gilt andererseits

$$\left(\frac{1}{|b|}\right)^n \rightarrow 0 \quad (n \rightarrow \infty),$$

ein Widerspruch.

Satz: Es sei (a_n) eine Folge in X . Dann gilt:

- (1) Ist (a_n) konvergent, so ist ihr Grenzwert eindeutig bestimmt.
- (2) Ist (a_n) konvergent, so ist (a_n) beschränkt.
- (3) Ist $a \in X$ und (α_n) eine Nullfolge in $\mathbb{R}$ mit

$$|a_n - a| \leq \alpha_n \text{ f\"ur } n \in \mathbb{N},$$

so gilt $a_n \rightarrow a$.

- (4) Ist (a_n) eine Folge in $\mathbb{C}$, so gilt

$$a_n \rightarrow a \iff \operatorname{Re} a_n \rightarrow \operatorname{Re} a \wedge \operatorname{Im} a_n \rightarrow \operatorname{Im} a.$$

- (5) Ist (a_n) eine Folge in $\mathbb{C}$ mit $a_n \in \mathbb{R}$ f\"ur $n \in \mathbb{N}$ und $a_n \rightarrow a$, so gilt $a \in \mathbb{R}$.

Beweis: (1) Es gelte $a_n \rightarrow a$ und $a_n \rightarrow b$. Annahme: $a \neq b$. Dann ist $\varepsilon := \frac{|a-b|}{2} > 0$. Es gilt:

$$\exists n_1 \in \mathbb{N} \forall n \geq n_1 : |a_n - a| < \varepsilon \quad \text{und} \quad \exists n_2 \in \mathbb{N} \forall n \geq n_2 : |a_n - b| < \varepsilon.$$

Sei $n_0 := \max\{n_1, n_2\}$ und $n \geq n_0$. Dann ist

$$2\varepsilon = |a - b| = |a - a_n + a_n - b| \leq |a_n - a| + |a_n - b| < 2\varepsilon,$$

ein Widerspruch.

- (2) Es existiert ein $n_0 \in \mathbb{N}$ mit $|a_n - a| < 1$ ($n \geq n_0$), also $|a_n| < 1 + |a|$ ($n \geq n_0$). Setze

$$\gamma := \max\{|a|, |a_1|, \dots, |a_{n_0}|\} + 1.$$

Dann gilt $|a_n| \leq \gamma$ ($n \in \mathbb{N}$).

- (3) Sei $\varepsilon > 0$. Es gilt

$$\exists n_1 \in \mathbb{N} \forall n \geq n_1 : |a_n - a| \leq \alpha_n$$

und

$$\exists n_2 \in \mathbb{N} \forall n \geq n_2 : \alpha_n < \varepsilon.$$

Für $n \geq n_0 := \max\{n_1, n_2\}$ gilt dann $|a_n - a| \leq \alpha_n < \varepsilon$. Somit gilt $a_n \rightarrow a$.

(4) Es gilt (vgl. 5.2):

$$\max\{|\operatorname{Re} a_n - \operatorname{Re} a|, |\operatorname{Im} a_n - \operatorname{Im} a|\} \leq |a_n - a| \leq |\operatorname{Re} a_n - \operatorname{Re} a| + |\operatorname{Im} a_n - \operatorname{Im} a|.$$

Die Behauptung folgt damit aus (3).

(5) Es gilt $\operatorname{Re} a_n = a_n$ ffa $n \in \mathbb{N}$. Wir haben $a_n \rightarrow a$ und mit (4) folgt $a_n \rightarrow \operatorname{Re} a$. Mit (1) folgt $\operatorname{Re} a = a$. $\square$

Bemerkung: Wegen (4) können wir uns bei der weiteren Untersuchung der Folgenkonvergenz weitgehend auf Folgen in $\mathbb{R}$ beschränken.

6.3. Grenzwertsätze: Es seien (a_n) , (b_n) und (c_n) Folgen in $\mathbb{R}$ und $a, b \in \mathbb{R}$. Dann gilt:

(1) $a_n \rightarrow a \implies |a_n| \rightarrow |a|$.

(2) $a_n \rightarrow a$, $b_n \rightarrow b$ und $a_n \leq b_n$ ffa $n \in \mathbb{N} \implies a \leq b$.

(3) $a_n \rightarrow a$, $b_n \rightarrow a$ und $a_n \leq c_n \leq b_n$ ffa $n \in \mathbb{N} \implies c_n \rightarrow a$.

(4) Gilt $a_n \rightarrow a$ und $b_n \rightarrow b$, so gilt:

(a) $a_n + b_n \rightarrow a + b$.

(b) $a_n \cdot b_n \rightarrow a \cdot b$.

(c) Ist $b \neq 0$, so ist $b_n \neq 0$ ffa $n \in \mathbb{N}$ und es gilt $\frac{a_n}{b_n} \rightarrow \frac{a}{b}$.

Die Eigenschaften (1) und (4) gelten auch für Folgen in $\mathbb{C}$.

Beispiele: (1) Sei $p \in \mathbb{N}$. Bekannt: $\frac{1}{n} \rightarrow 0$. Mit 6.3(4)(b) folgt

$$\frac{1}{n^p} = \left(\frac{1}{n}\right)^p \rightarrow 0^p = 0 \quad (n \rightarrow \infty).$$

(2) Mit 6.3(4) folgt

$$\frac{n^2 + 4n + 5}{3n^2 + n - 1} = \frac{1 + \frac{4}{n} + \frac{5}{n^2}}{3 + \frac{1}{n} - \frac{1}{n^2}} \rightarrow \frac{1 + 0 + 0}{3 + 0 + 0} = \frac{1}{3} \quad (n \rightarrow \infty).$$

Beweis von 6.3: (1) Es gilt $||a_n| - |a|| \leq |a_n - a| =: \alpha_n$ ($n \in \mathbb{N}$) und $\alpha_n \rightarrow 0$. Damit folgt $|a_n| \rightarrow |a|$.

(2) Annahme: $a > b$. Setze $\varepsilon := (a - b)/2$. Nun gilt

$$b_n < b + \varepsilon = \frac{a + b}{2} = a - \varepsilon < a_n \text{ ffa } n \in \mathbb{N},$$

Widerspruch.

(3) Es sei $\varepsilon > 0$. Es gilt

$$a - \varepsilon < a_n \leq c_n \leq b_n < a + \varepsilon \text{ ffa } n \in \mathbb{N},$$

also $|c_n - a| < \varepsilon$ ffa $n \in \mathbb{N}$.

(4) (a) Es sei $\varepsilon > 0$. Es gilt $|a_n - a| < \varepsilon/2$ und $|b_n - b| < \varepsilon/2$ ffa $n \in \mathbb{N}$. Also ist

$$|a_n + b_n - (a + b)| \leq |a_n - a| + |b_n - b| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \text{ ffa } n \in \mathbb{N}.$$

(b) Die Folge (a_n) ist beschränkt (da konvergent). Für ein $\gamma \geq 0$ gilt also $|a_n| \leq \gamma$ ($n \in \mathbb{N}$). Es folgt

$$\begin{aligned} |a_n b_n - ab| &= |(a_n b_n - a_n b) + (a_n b - ab)| \leq |a_n| |b_n - b| + |b| |a_n - a| \\ &\leq \gamma |b_n - b| + |b| |a_n - a| =: \alpha_n \quad (n \in \mathbb{N}). \end{aligned}$$

Es gilt $\alpha_n \rightarrow 0$, also $a_n b_n \rightarrow ab$.

(c) Wegen $b \neq 0$ ist $\delta := |b|/2 > 0$. Also (bea. $|b_n - b| < \delta \Rightarrow |b_n| > |b| - \delta$)

$$\exists n_0 \in \mathbb{N} \forall n \geq n_0 : |b_n| > |b| - \delta = \delta > 0.$$

Somit ist die Folge $(\frac{a_n}{b_n})_{n=n_0}^\infty$ definiert. Für $n \geq n_0$ gilt nun

$$\left| \frac{1}{b_n} - \frac{1}{b} \right| = \frac{|b - b_n|}{|b| |b_n|} \leq \frac{1}{2\delta^2} |b_n - b| \rightarrow 0 \quad (n \rightarrow \infty),$$

woraus mit (b) die Behauptung folgt. □

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

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

Satz (Monotoniekriterium): 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. Es sei (a_n) monoton wachsend und $s := \sup\{a_n : n \in \mathbb{N}\}$. Zu $\varepsilon > 0$ existiert ein $n_0 \in \mathbb{N}$ mit $a_{n_0} > s - \varepsilon$. Für $n \geq n_0$ gilt dann

$$s - \varepsilon < a_{n_0} \leq a_n \leq s < s + \varepsilon,$$

also $|a_n - s| < \varepsilon$. □

Beispiel: Sei (a_n) rekursiv definiert durch $a_1 := \sqrt{6}$, $a_{n+1} := \sqrt{6 + a_n}$ ($n \in \mathbb{N}$). Dann gilt: (a_n) ist streng wachsend und $0 < a_n < 3$ ($n \in \mathbb{N}$).

Nachweis (induktiv) $A(n) := [a_n < a_{n+1} \wedge 0 < a_n < 3]$:

(IA) $n = 1$: $a_2 = \sqrt{6 + \sqrt{6}} > \sqrt{6} = a_1$ und $0 < a_1 < \sqrt{6 + 3} = 3$.

(IS) : Sei $n \in \mathbb{N}$ und es gelte $A(n)$. Dann ist

$$a_{n+1} = \sqrt{6 + a_n} < \sqrt{6 + a_{n+1}} = a_{n+2} \wedge 0 < a_{n+1} = \sqrt{6 + a_n} < \sqrt{6 + 3} = 3.$$

Also gilt $A(n)$ ($n \in \mathbb{N}$) und nach dem Monotoniekriterium ist (a_n) konvergent; setze $a := \lim_{n \rightarrow \infty} a_n$. Aus $a_{n+1} := \sqrt{6 + a_n}$ folgt $a_{n+1}^2 := 6 + a_n$ ($n \in \mathbb{N}$) und $n \rightarrow \infty$ liefert $a^2 = 6 + a$, d.h. $0 = a^2 - a - 6 = (a + 2)(a - 3)$. Also gilt $a = -2$ oder $a = 3$. Wegen $a_n > 0$ ($n \in \mathbb{N}$) gilt $a \geq 0$, also ist $a = 3$.

6.5. Wichtige Beispiele: (1) Sei $a_n \geq 0$ ($n \in \mathbb{N}$) und $a_n \rightarrow a$. Dann gilt für jedes $p \in \mathbb{N}$:

$$\sqrt[p]{a_n} \rightarrow \sqrt[p]{a} \quad (n \rightarrow \infty).$$

Beweis. Vorbemerkung: Für alle $y \geq x \geq 0$ gilt $\sqrt[p]{y} - \sqrt[p]{x} \leq \sqrt[p]{y - x}$, denn

$$(\sqrt[p]{y - x} + \sqrt[p]{x})^p = \sum_{k=0}^p \binom{p}{k} (\sqrt[p]{y - x})^k (\sqrt[p]{x})^{p-k} \geq x + y - x = y.$$

Wir erhalten somit $|\sqrt[p]{a_n} - \sqrt[p]{a}| \leq \sqrt[p]{|a_n - a|}$ ($n \in \mathbb{N}$). Für $b_n := |a_n - a|$ gilt $b_n \rightarrow 0$ und es reicht zu zeigen, dass $\sqrt[p]{b_n} \rightarrow 0$:

Zu $\varepsilon > 0$ existiert ein $n_0 \in \mathbb{N}$ mit $0 \leq b_n < \varepsilon^p$ ($n \geq n_0$), also $0 \leq \sqrt[p]{b_n} < \varepsilon$ ($n \geq n_0$). □

(2) Es gilt $\sqrt[n]{n} \rightarrow 1$.

Beweis. Setze $a_n := \sqrt[n]{n} - 1$. Dann ist $a_n \geq 0$ ($n \in \mathbb{N}$), 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 (beachte: nach Bsp. (1) gilt $\sqrt{1/n} \rightarrow 0$)

$$0 \leq a_n \leq \frac{\sqrt{2}}{\sqrt{n-1}} \rightarrow 0 \quad (n \rightarrow \infty)$$

Es folgt $a_n \rightarrow 0$. □

Folgerung: Für $c > 0$ gilt $\sqrt[n]{c} \rightarrow 1$:

Für $c \geq 1$ ist $1 \leq \sqrt[n]{c} \leq \sqrt[n]{n}$ ffa $n \in \mathbb{N}$, und für $c \in (0, 1)$ ist $1/\sqrt[n]{n} \leq \sqrt[n]{c} \leq 1$ ffa $n \in \mathbb{N}$. Die Behauptung folgt nun aus 6.3(3).

(3) Konvergiert eine Folge (a_n) so gilt $a_{n+1} - a_n \rightarrow 0$, denn für $a_n \rightarrow a$ gilt auch $a_{n+1} \rightarrow a$, und $a_{n+1} - a_n \rightarrow 0$ folgt aus Satz 6.3(4).

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

$$0 < \sqrt{n+1} - \sqrt{n} = \frac{(\sqrt{n+1} - \sqrt{n})(\sqrt{n+1} + \sqrt{n})}{\sqrt{n+1} + \sqrt{n}} = \frac{(n+1) - n}{\sqrt{n+1} + \sqrt{n}} \leq \frac{1}{\sqrt{n}},$$

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

(4) Es sei $z \in \mathbb{C}$ und

$$s_n := \sum_{k=0}^n z^k = 1 + z + z^2 + \dots + z^n \quad (n \in \mathbb{N}_0).$$

1. Fall $|z| < 1$: Nach 4.11(1) ist dann

$$s_n = \frac{1 - z^{n+1}}{1 - z} \quad (n \in \mathbb{N}_0).$$

Wegen $|z| < 1$ gilt $z^n \rightarrow 0$, also

$$s_n \rightarrow \frac{1}{1 - z} \quad (n \rightarrow \infty).$$

2. Fall $|z| \geq 1$. Dann ist

$$|s_{n+1} - s_n| = |z|^{n+1} \geq 1 \quad (n \in \mathbb{N}_0)$$

und somit ist $(s_n)_{n=0}^\infty$ divergent (vgl. Bsp. (3)).

6.6. Die Eulersche Zahl e : Für jedes $n \in \mathbb{N}$ sei

$$a_n := \left(1 + \frac{1}{n}\right)^n, \quad b_n := \left(1 + \frac{1}{n}\right)^{n+1}.$$

Nach 4.11(6) ist (a_n) (streng) monoton wachsend und (b_n) ist (streng) monoton fallend. Außerdem gilt $a_n < b_n$ für jedes $n \in \mathbb{N}$. Also

$$\forall n \in \mathbb{N} : a_1 \leq a_n < a_{n+1} < b_{n+1} < b_n \leq b_1.$$

Nach 6.4 Satz sind die Folgen (a_n) und (b_n) konvergent wegen

$$b_n = \left(1 + \frac{1}{n}\right)a_n \quad (n \in \mathbb{N})$$

gilt $\lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} b_n$.

Der Limes

$$e := \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n$$

heißt *Eulersche Zahl*.

Bemerkung: Es gilt $e \approx 2.718$. Wir haben aus dem Beweis z.B. die Abschätzung $2 = a_1 < e < b_1 = 4$, aber auch $2.25 = \frac{9}{4} = a_2 < e < b_2 = \frac{27}{8} = 3.375$ etc.

6.7. Intervallschachtelung: Eine *Intervallschachtelung* ist eine Folge $(I_n)_{n \in \mathbb{N}}$ von abgeschlossenen Intervallen $I_n = [a_n, b_n]$, mit $a_n \leq b_n$, $I_{n+1} \subseteq I_n$ ($n \in \mathbb{N}$) und $b_n - a_n \rightarrow 0$.

Beispiel: Nach 6.6 definiert $I_n := [(1 + \frac{1}{n})^n, (1 + \frac{1}{n})^{n+1}]$ eine Intervallschachtelung, für die $\bigcap_{n \in \mathbb{N}} I_n = \{e\}$ gilt.

Satz: Ist (I_n) eine Intervallschachtelung, so gibt es genau eine Zahl $r \in \mathbb{R}$ mit

$$\bigcap_{n \in \mathbb{N}} I_n = \{r\}.$$

Es gilt $\max I_n \rightarrow r$ und $\min I_n \rightarrow r$.

Beweis. Es gilt $\min I_n = a_n$ und $\max I_n = b_n$ ($n \in \mathbb{N}$). Wegen

$$I_{n+1} = [a_{n+1}, b_{n+1}] \subseteq [a_n, b_n] = I_n \quad (n \in \mathbb{N})$$

ist (b_n) monoton fallend und beschränkt und (a_n) ist monoton wachsend und beschränkt. Somit existieren $a, b \in \mathbb{R}$ mit $a_n \rightarrow a$ und $b_n \rightarrow b$. Wegen $b_n - a_n \rightarrow 0$ ist $a = b =: r$. Wegen $a_n \leq r \leq b_n$ ($n \in \mathbb{N}$) ist $\bigcap_{n \in \mathbb{N}} I_n = \{r\}$. $\square$

6.8. Teilfolgen: Ist (a_n) eine Folge in einer beliebigen Menge und $k : \mathbb{N} \rightarrow \mathbb{N}$ eine Abbildung mit

$$\forall n \in \mathbb{N} : k(n) < k(n+1)$$

(d.h. $(k(n))$ ist eine streng wachsende Folge natürlicher Zahlen), so heißt die Folge $(a_{k(n)})$ *Teilfolge (TF)* von (a_n) . Beachte: $k(n) \geq n$ ($n \in \mathbb{N}$).

Ist (a_n) eine Folge in $X \in \{\mathbb{R}, \mathbb{C}\}$ so heißt $\alpha \in X$ ein *Häufungswert (HW)* der Folge (a_n) , falls es eine Teilfolge von (a_n) gibt, die gegen α konvergiert. Wir setzen

$$H(a_n) := \{\alpha \in X : \alpha \text{ ist ein HW von } (a_n)\}.$$

Beispiel: Die Folge $(a_2, a_4, a_6, \dots)$ ist eine Teilfolge von (a_n) , hier ist $k(n) = 2n$ ($n \in \mathbb{N}$). Hingegen ist $(a_3, a_1, a_7, a_5, \dots)$ keine Teilfolge von (a_n) .

Satz: Es sei (a_n) eine Folge. Dann gilt:

(1) α ist ein Häufungswert von $(a_n) \Leftrightarrow$

$$\forall \varepsilon > 0 : a_n \in U_\varepsilon(\alpha) \text{ für unendlich viele } n \in \mathbb{N}.$$

(2) Ist (a_n) konvergent mit Grenzwert a , und $(a_{k(n)})$ eine Teilfolge von (a_n) , so gilt $a_{k(n)} \rightarrow a$. Insbesondere gilt dann $H(a_n) = \{a\}$ (d.h. eine konvergente Folge hat genau einen Häufungswert).

(3) Sind die Teilfolgen (a_{2n}) und (a_{2n-1}) konvergent mit

$$a := \lim_{n \rightarrow \infty} a_{2n} = \lim_{n \rightarrow \infty} a_{2n-1}$$

so gilt $a_n \rightarrow a$.

Beweis: (1) “ $\Rightarrow$ ” Für eine Teilfolge $(a_{k(n)})$ von (a_n) gilt $a_{k(n)} \rightarrow \alpha$. Sei $\varepsilon > 0$. Dann gilt

$$a_{k(n)} \in U_\varepsilon(\alpha) \text{ f\"ur } n \in \mathbb{N},$$

also (bea. k ist streng wachsend) $a_n \in U_\varepsilon(\alpha)$ für unendlich viele $n \in \mathbb{N}$.

“ $\Leftarrow$ ” Zu $\varepsilon = 1$ existiert ein $k(1) \in \mathbb{N}$ mit $a_{k(1)} \in U_1(\alpha)$;

Zu $\varepsilon = 1/2$ existiert ein $k(2) \in \mathbb{N}$ mit $a_{k(2)} \in U_{1/2}(\alpha)$ und $k(2) > k(1)$;

Zu $\varepsilon = 1/3$ existiert ein $k(3) \in \mathbb{N}$ mit $a_{k(3)} \in U_{1/3}(\alpha)$ und $k(3) > k(2)$; etc.

Es existiert also eine Teilfolge $(a_{k(n)})$ mit $|a_{k(n)} - \alpha| < 1/n$ ($n \in \mathbb{N}$), also mit $a_{k(n)} \rightarrow \alpha$.

(2) Sei $\varepsilon > 0$. Es gilt $a_n \in U_\varepsilon(a)$ f\"ur $n \in \mathbb{N}$, also auch $a_{k(n)} \in U_\varepsilon(a)$ f\"ur $n \in \mathbb{N}$. Somit gilt $a_{k(n)} \rightarrow a$.

(3) Übung. □

Beispiele: (1) Es gilt $H((-1)^n) = \{-1, 1\}$:

Wegen $a_{2n} \rightarrow 1$ und $a_{2n+1} \rightarrow -1$ gilt $-1, 1 \in H((-1)^n)$. Ist $\alpha \in \mathbb{R} \setminus \{1, -1\}$ und $\varepsilon := \min\{|1 - \alpha|, |-1 - \alpha|\}/2$, so ist $\varepsilon > 0$ und in $U_\varepsilon(\alpha)$ liegen keine Folgenglieder. Somit ist $\alpha \notin H((-1)^n)$.

(2) Für $(a_n) = ((1 + (-1)^n)^n)$ gilt $H(a_n) = \{0\}$: Wegen $a_{2n-1} = 0 \rightarrow 0$ ist $0 \in H(a_n)$ und $(a_{2n}) = (2^n)$ hat keinen Häufungswert, da jede Teilfolge unbeschränkt ist und somit nicht konvergiert.

(3) Übung: Es gilt $H(i^n) = \{i, -1, -i, 1\}$.

Satz (Bolzano-Weierstraß): Jede beschränkte Folge (c_n) in $\mathbb{R}$ hat eine konvergente Teilfolge, d.h. $H(c_n) \neq \emptyset$.

Beweis. Da (c_n) beschränkt ist existiert ein Intervall $I_1 := [a_1, b_1]$ mit $c_n \in I_1$ ($n \in \mathbb{N}$). Setze $k(1) := 1$. Es gilt $c_{k(1)} \in I_1$.

Sind $I_n = [a_n, b_n]$ und $k(n)$ konstruiert und

$$J_l := [a_n, \frac{a_n + b_n}{2}], \quad J_r := [\frac{a_n + b_n}{2}, b_n],$$

so gilt

$$(c_m \in J_l \text{ für unendlich viele } m \in \mathbb{N}) \vee (c_m \in J_r \text{ für unendlich viele } m \in \mathbb{N}).$$

Definiere $I_{n+1} = [a_{n+1}, b_{n+1}]$ als J_l oder J_r , so dass I_{n+1} unendlich viele der c_m enthält. Dann existiert ein $k(n+1) > k(n)$ mit $c_{k(n+1)} \in I_{n+1}$. Nun ist (I_n) eine Intervallschachtelung, denn

$$b_n - a_n = \frac{b_1 - a_1}{2^{n-1}} \rightarrow 0 \quad (n \rightarrow \infty),$$

also $\bigcap_{n \in \mathbb{N}} I_n = \{\alpha\}$ und $a_n \rightarrow \alpha$, $b_n \rightarrow \alpha$ nach 6.7. Es folgt $c_{k(n)} \rightarrow \alpha$ nach 6.3(3). $\square$

Bemerkung: Der Satz gilt auch für Folgen in $\mathbb{C}$. 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 $(z_{(kol)(n)})$ eine konvergente Teilfolge von (z_n) .

6.9. Rechnen mit ∞ : Es sei (a_n) eine Folge in $\mathbb{R}$.

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

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

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

Bemerkung: (a) Gilt $a_n \rightarrow \pm\infty$, so folgt offensichtlich $|a_n| \rightarrow \infty$. Damit folgt $1/a_n \rightarrow 0$: Sei $\varepsilon > 0$. Dann

$$\exists n_0 \in \mathbb{N} \forall n \geq n_0 : |a_n| > 1/\varepsilon,$$

also

$$\forall n \geq n_0 : |1/a_n| < \varepsilon.$$

Ist $a_n > 0$ [bzw. $a_n < 0$] ($n \in \mathbb{N}$) und gilt $a_n \rightarrow 0$, so folgt $1/a_n \rightarrow \infty$ [bzw. $1/a_n \rightarrow -\infty$].

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

(c) Für jede monoton wachsende [monoton fallende] reelle Folge (a_n) existiert genau ein $a \in \mathbb{R} \cup \{\infty\}$ [bzw. $a \in \mathbb{R} \cup \{-\infty\}$] mit $a_n \rightarrow a$.

Beispiel: Sei $b > 1$ und $a_n := b^n$ ($n \in \mathbb{N}$). Dann gilt $a_n \rightarrow \infty$.

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

Man setzt

$$\forall a \in \mathbb{R} : a + \infty := \infty, \quad a - \infty = -\infty,$$

$$\begin{aligned}
\forall a > 0 : a \cdot \infty &:= \infty, \quad a \cdot (-\infty) := -\infty, \\
\forall a < 0 : a \cdot \infty &:= -\infty, \quad a \cdot (-\infty) := \infty, \\
\infty + \infty &:= \infty, \quad -\infty - \infty := -\infty, \\
\infty \cdot \infty &:= \infty, \quad \infty \cdot (-\infty) := -\infty, \quad (-\infty) \cdot (-\infty) := \infty.
\end{aligned}$$

Achtung: Die Ausdrücke $0 \cdot \infty$, $0 \cdot (-\infty)$ und $\infty - \infty$ sind (aus gutem Grund) **nicht definiert!** Beispiele zu “ $0 \cdot \infty$ ”:

$$\frac{1}{n+1}n \rightarrow 1, \quad \frac{1}{2n+1}n \rightarrow \frac{1}{2}, \quad \frac{1}{n^2+1}n \rightarrow 0, \quad \frac{1}{n+1}n^2 \rightarrow \infty.$$

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

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

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

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

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

$$\sup \emptyset := -\infty, \quad \inf \emptyset := \infty.$$

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

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

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

6.10. Limes superior und Limes inferior: Es sei (a_n) eine Folge in $\mathbb{R}$. Für $n \in \mathbb{N}$ sei

$$A_n := \{a_k : k \geq n\} = \{a_n, a_{n+1}, a_{n+2}, \dots\}.$$

Es gilt:

$$(*) \quad \forall n \in \mathbb{N} : A_{n+1} \subseteq A_n.$$

Fall 1: Ist (a_n) nicht nach oben [unten] beschränkt, so setzen wir

$$\limsup_{n \rightarrow \infty} a_n := \infty, \quad [\liminf_{n \rightarrow \infty} a_n := -\infty].$$

Fall 2: Ist (a_n) nach oben [unten] beschränkt, so sind alle A_n nach oben [unten] beschränkt. Wir setzen $b_n := \sup A_n$ [$b_n := \inf A_n$] ($n \in \mathbb{N}$). Wegen (*) ist (b_n) fallend [wachsend]. Setze

$$\limsup_{n \rightarrow \infty} a_n := \lim_{n \rightarrow \infty} b_n \in \mathbb{R} \cup \{-\infty\}$$

$$[\liminf_{n \rightarrow \infty} a_n := \lim_{n \rightarrow \infty} b_n \in \mathbb{R} \cup \{\infty\}].$$

In beiden Fällen heißt $\limsup_{n \rightarrow \infty} a_n$ [$\liminf_{n \rightarrow \infty} a_n$] der Limes superior [Limes inferior] von (a_n) .

Schreib- und Sprechweisen: Man schreibt auch $\overline{\lim}_{n \rightarrow \infty} a_n$ [$\underline{\lim}_{n \rightarrow \infty} a_n$] und spricht vom “oberen Limes” [“unteren Limes”] von (a_n) .

Beispiele: (1) Die Folge $(a_n) = ((-1)^n n) = (-1, 2, -3, 4, -5, \dots)$ ist weder nach oben noch nach unten beschränkt. Also ist $\limsup_{n \rightarrow \infty} a_n = \infty$ und $\liminf_{n \rightarrow \infty} a_n := -\infty$. (2) Die Folge $(a_n) = (n)$ ist nicht nach oben beschränkt. Also ist $\limsup_{n \rightarrow \infty} a_n = \infty$. Sie ist aber nach unten beschränkt. Es gilt

$$A_n = \{n, n+1, n+2, \dots\} \text{ also } b_n := \inf A_n = n \quad (n \in \mathbb{N}).$$

Somit ist $\liminf_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} b_n = \infty$.

(3) Die Folge $(a_n) = (1, 2, 1, 3, 1, 4, \dots)$ ist nicht nach oben beschränkt. Also ist $\limsup_{n \rightarrow \infty} a_n = \infty$. Sie ist aber nach unten beschränkt. Es gilt $\inf A_n = 1$ ($n \in \mathbb{N}$), also $\liminf_{n \rightarrow \infty} a_n = 1$.

Satz (ohne Beweis): Es sei (a_n) eine Folge in $\mathbb{R}$. Dann gilt:

$$(1) \quad \liminf_{n \rightarrow \infty} a_n \leq \limsup_{n \rightarrow \infty} a_n.$$

(2) Es gibt Teilfolgen $(a_{k(n)})$ und $(a_{l(n)})$ von (a_n) mit

$$\lim_{n \rightarrow \infty} a_{k(n)} = \limsup_{n \rightarrow \infty} a_n, \quad \lim_{n \rightarrow \infty} a_{l(n)} = \liminf_{n \rightarrow \infty} a_n.$$

(3) Sei $\alpha := \limsup_{n \rightarrow \infty} a_n$ und $\beta, \gamma \in \mathbb{R}$ mit $\gamma < \alpha < \beta$. Dann gilt:

$$a_n < \beta \text{ ffa } n \in \mathbb{N} \quad \text{und} \quad a_n > \gamma \text{ für unendlich viele } n \in \mathbb{N}.$$

(4) Sei $\alpha := \liminf_{n \rightarrow \infty} a_n$ und $\beta, \gamma \in \mathbb{R}$ mit $\gamma < \alpha < \beta$. Dann gilt:

$$a_n < \beta \text{ für unendlich viele } n \in \mathbb{N} \quad \text{und} \quad a_n > \gamma \text{ ffa } n \in \mathbb{N}.$$

(5) Ist (a_n) beschränkt, so ist $H(a_n)$ beschränkt (und nicht leer; vgl. Satz von Bolzano Weierstraß) und

$$\liminf_{n \rightarrow \infty} a_n = \min H(a_n), \quad \limsup_{n \rightarrow \infty} a_n = \max H(a_n).$$

(6) Ist (a_n) konvergent, so ist

$$\liminf_{n \rightarrow \infty} a_n = \limsup_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} a_n.$$

(7) Gilt $a_n \geq 0$ ($n \in \mathbb{N}$) und $\limsup_{n \rightarrow \infty} a_n = 0$, so gilt $\lim_{n \rightarrow \infty} a_n = 0$.

Bemerkung: In Verallgemeinerung von (5) gilt: Ist (a_n) eine reelle Folge und

$$H(a_n) := \{\alpha \in \mathbb{R} \cup \{-\infty, \infty\} : \text{es gibt eine Teilfolge } (a_{k(n)}) \text{ mit } a_{k(n)} \rightarrow \alpha\},$$

so ist

$$\liminf_{n \rightarrow \infty} a_n = \min A, \quad \limsup_{n \rightarrow \infty} a_n = \max A.$$

Beispiele: (1) Bekannt: Die Folge $(a_n) = ((-1)^n)$ ist beschränkt und $H(a_n) = \{-1, 1\}$. Also gilt

$$\liminf_{n \rightarrow \infty} a_n = -1, \quad \limsup_{n \rightarrow \infty} a_n = 1.$$

(2) Es sei die Folge (a_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:

$$\begin{aligned} a_{4n} &= 1 - \frac{1}{4n} \rightarrow 1, & a_{4n+1} &= (4n+1)(1-1) = 0 \rightarrow 0, \\ a_{4n+2} &= -(1 - \frac{1}{4n+2}) \rightarrow -1, & a_{4n+3} &= (4n+3)(1+1) \rightarrow \infty. \end{aligned}$$

Wir haben also $H(a_n) = \{-1, 0, 1, \infty\}$ und somit $\limsup_{n \rightarrow \infty} a_n = \infty$, $\liminf_{n \rightarrow \infty} a_n = -1$.

6.11. Cauchyfolgen: Es sei (a_n) eine konvergente Zahlenfolge. Dann gilt:

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

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

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

□

Definition: Eine Folge (a_n) in $\mathbb{R}$ oder $\mathbb{C}$, für welche (C) gilt, heißt *Cauchyfolge* (CF).

Satz (Cauchy Kriterium): Es sei (a_n) eine Folge in $\mathbb{R}$ oder $\mathbb{C}$. Dann gilt:

$$(a_n) \text{ ist konvergent} \Leftrightarrow (a_n) \text{ ist eine Cauchyfolge}$$

Beweis. “ $\Rightarrow$ ”: Siehe oben.

“ $\Leftarrow$ ”: Wir zeigen zuerst, dass (a_n) beschränkt ist: Es existiert ein $n_0 \in \mathbb{N}$ mit

$$\forall n, m \geq n_0 : |a_n - a_m| < 1.$$

Für $\gamma := \max\{|a_1|, |a_2|, \dots, |a_{n_0}|\} + 1$ gilt nun: $|a_n| \leq \gamma$ ($1 \leq n \leq n_0$) und

$$|a_n| \leq |a_n - a_{n_0}| + |a_{n_0}| < 1 + |a_{n_0}| \leq \gamma \quad (n > n_0).$$

Nach dem Satz von Bolzano-Weierstraß hat (a_n) eine konvergente Teilfolge $(a_{k(n)})$. Setze $a := \lim_{n \rightarrow \infty} a_{k(n)}$. Sei $\varepsilon > 0$. Nun existiert ein $n_0 \in \mathbb{N}$ so, dass für alle $n, m \geq n_0$ gilt

$$|a_n - a_m| < \varepsilon/2 \text{ und } |a_{k(n_0)} - a| < \varepsilon/2.$$

Für jedes $n \geq k(n_0)$ gilt dann wegen $k(n_0) \geq n_0$:

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

□

6.12. Abzählbarkeit von Mengen:

Definition: Es sei M eine Menge.

(1) M heißt *endlich* : $\Leftrightarrow M = \emptyset$, oder es existiert ein $n \in \mathbb{N}$ und eine surjektive Funktion

$$f : \{1, 2, \dots, n\} \rightarrow M, \text{ also } M = \{f(1), f(2), \dots, f(n)\}.$$

(2) M heißt *unendlich* : $\Leftrightarrow M$ ist nicht endlich.

(3) M heißt *abzählbar* : $\Leftrightarrow M = \emptyset$, oder es existiert eine surjektive Abbildung von $\mathbb{N}$ nach M , also eine Folge $(a_n)_{n \in \mathbb{N}}$ in M mit

$$M = \{a_n : n \in \mathbb{N}\} = \{a_1, a_2, \dots\}.$$

Beachte: Ist M endlich, so ist M abzählbar. Z.B.

$$M = \{-1, 1\} = \{a_n : n \in \mathbb{N}\} \text{ für } (a_n) = ((-1)^n).$$

(4) M heißt *überabzählbar* : $\Leftrightarrow M$ ist nicht abzählbar.

Beispiele: (1) $\mathbb{N}$ ist abzählbar: $(a_n) = (1, 2, 3, \dots)$.

(2) $\mathbb{Z}$ ist abzählbar: $(a_n) = (0, 1, -1, 2, -2, \dots)$.

(3) $\mathbb{Q} = \{\frac{p}{q} : p \in \mathbb{Z}, q \in \mathbb{N}\}$ ist abzählbar: Das liegt daran, dass $\mathbb{Z} \times \mathbb{N}$ abzählbar ist (Cantorsches Diagonalisierungsverfahren).

Satz Jedes Intervall mit mindestens 2 Punkten ist überabzählbar. Insbesondere ist $\mathbb{R}$ überabzählbar.

Beweis. Wir betrachten o.B.d.A. $I = [0, 1]$. Angenommen es gibt eine Folge (a_n) mit

$$I = \{a_n : n \in \mathbb{N}\}.$$

Wir konstruieren induktiv eine Intervallschachtelung (I_n) wie folgt: Sei I_1 eines der Intervalle $[0, 1/3]$, $[1/3, 2/3]$, $[2/3, 1]$ mit $a_1 \notin I_1$. Ist I_n konstruiert, so zerlegt man I_n in 3 gleichlange Intervalle und wählt als I_{n+1} eines dieser Intervalle mit $a_{n+1} \notin I_{n+1}$. Die Länge von I_n ist $1/3^n$ ($n \in \mathbb{N}$). Nach 6.7 existiert ein $r \in [0, 1]$ mit

$$\bigcap_{n \in \mathbb{N}} I_n = \{r\},$$

und nach Konstruktion ist $r \notin \{a_n : n \in \mathbb{N}\}$, ein Widerspruch. $\square$

Satz Ist $M \neq \emptyset$ und $f : M \rightarrow \text{Pot}(M)$ eine Funktion, so ist f nicht surjektiv.

Beweis. Annahme: f ist surjektiv. Sei $N := \{x \in M : x \notin f(x)\}$. Dann ist $N \in \text{Pot}(M)$, also $N = f(x_0)$ für ein $x_0 \in M$.

1.Fall: $x_0 \in N$. Dann gilt $x_0 \notin f(x_0) = N$. Widerspruch.

2.Fall: $x_0 \notin N$. Dann gilt $x_0 \in f(x_0) = N$. Widerspruch. $\square$

Folgerung: $\text{Pot}(\mathbb{N})$ ist überabzählbar.

7 Reihen

7.1. Definition: Sei $(a_n)_{n \in \mathbb{N}}$ eine Folge in $\mathbb{R}$ oder $\mathbb{C}$ und

$$s_N := \sum_{n=1}^N a_n = a_1 + a_2 + \dots + a_N \quad (N \in \mathbb{N}).$$

Die Folge $(s_N)_{N \in \mathbb{N}}$ heißt eine (*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 $\lim_{N \rightarrow \infty} s_N$ der *Reihenwert* von $\sum_{n=1}^{\infty} a_n$ und wird ebenfalls mit $\sum_{n=1}^{\infty} a_n$ bezeichnet (**Vorsicht:** Doppelbedeutung des Symbols $\sum_{n=1}^{\infty} a_n$).

Bemerkung: (a) Ist $p \in \mathbb{Z}$ so verfährt man mit Folgen der Form $(a_n)_{n=p}^{\infty}$ entsprechend und setzt

$$s_N := \sum_{n=p}^N a_n = a_p + a_{p+1} + \dots + a_N \quad (N \geq p).$$

Meistens: $p = 1$ oder $p = 0$.

(b) Ist (a_n) komplex, so gilt nach 6.2:

$$\sum_{n=1}^{\infty} a_n \text{ ist konvergent} \Leftrightarrow \sum_{n=1}^{\infty} \operatorname{Re} a_n \text{ und } \sum_{n=1}^{\infty} \operatorname{Im} a_n \text{ sind konvergent.}$$

In diesem Fall ist

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

Beispiele: (1) Die *geometrische Reihe* $\sum_{n=0}^{\infty} z^n$, $z \in \mathbb{C}$, konvergiert genau dann, wenn $|z| < 1$. Es gilt

$$\sum_{n=0}^{\infty} z^n = \frac{1}{1-z} \quad (|z| < 1)$$

(vgl. 6.5(4)).

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

$$s_N = 1 - \frac{1}{2} + \frac{1}{2} - \frac{1}{3} + \dots + \frac{1}{N} - \frac{1}{N+1} = 1 - \frac{1}{N+1} \rightarrow 1 \quad (N \rightarrow \infty).$$

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* $\sum_{n=1}^{\infty} \frac{1}{n}$ ist divergent: Für jedes $N \in \mathbb{N}$ ist

$$s_{2N} = 1 + \underbrace{\frac{1}{2} + \dots + \frac{1}{N}}_{=s_N} + \underbrace{\frac{1}{N+1} + \dots + \frac{1}{2N}}_{\geq \frac{1}{2N}} \geq s_N + N \frac{1}{2N} = s_N + \frac{1}{2}.$$

Somit ist (s_N) divergent, denn aus $s_N \rightarrow s \in \mathbb{R}$ folgt $s_{2N} \rightarrow s$, also $s \geq s + \frac{1}{2}$, ein Widerspruch.

Bemerkung: Da (s_N) monoton wachsend ist folgt $s_N \rightarrow \infty$ ($N \rightarrow \infty$). Man schreibt dafür auch

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

7.2. Satz: Es seien (a_n) und (b_n) Folgen und $s_N := a_1 + \dots + a_N$ ($N \in \mathbb{N}$).

(1) **Monotoniekriterium:** Gilt $a_n \geq 0$ ($n \in \mathbb{N}$) und ist $(s_N)_{N \in \mathbb{N}}$ beschränkt, so ist

$$\sum_{n=1}^{\infty} a_n \text{ konvergent.}$$

(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

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

Definiert man

$$r_p := \sum_{n=p+1}^{\infty} a_n \quad (p \in \mathbb{N}),$$

so gilt $r_p \rightarrow 0$ ($p \rightarrow \infty$), d.h. die Folge der "Reihenreste" konvergiert gegen Null.

(3) Ist $\sum_{n=1}^{\infty} a_n$ konvergent, so gilt $a_n \rightarrow 0$.

(4) Sind $\sum_{n=1}^{\infty} a_n$ und $\sum_{n=1}^{\infty} b_n$ konvergent und $\alpha, \beta \in \mathbb{R}$ (bzw. $\mathbb{C}$), so ist

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

und es gilt

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

(5) **Cauchy Kriterium:** $\sum_{n=1}^{\infty} a_n$ konvergiert genau dann, wenn gilt:

$$\forall \varepsilon > 0 \exists n_0 \in \mathbb{N} \forall N > M \geq n_0 : \left| \sum_{n=M+1}^N a_n \right| < \varepsilon.$$

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

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

$$\sum_{n=p+1}^{\infty} a_n = \sum_{n=1}^{\infty} a_n - s_p \quad (N \rightarrow \infty).$$

Somit gilt

$$r_p = \sum_{n=1}^{\infty} a_n - s_p \rightarrow \sum_{n=1}^{\infty} a_n - \sum_{n=1}^{\infty} a_n = 0 \quad (p \rightarrow \infty).$$

(3) Es gilt $a_n = s_n - s_{n-1} \rightarrow 0$ nach 6.5(3).

(4) folgt aus 6.3(4).

(5) Für $N \in \mathbb{N}$, gilt wegen 6.11:

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

□

Bemerkung: Aus 7.2(3) folgt: Ist (a_n) keine Nullfolge, so ist $\sum_{n=1}^{\infty} a_n$ divergent. Die Umkehrung von 7.2(3) gilt nicht:

$$\sum_{n=1}^{\infty} \frac{1}{n} \text{ ist divergent, aber } \frac{1}{n} \rightarrow 0 \ (n \rightarrow \infty).$$

7.3. Absolut konvergente Reihen und Leibnizkriterium:

Definition: Die Reihe $\sum_{n=1}^{\infty} a_n$ heißt *absolut konvergent*, falls $\sum_{n=1}^{\infty} |a_n|$ konvergiert.

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”:

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

Beweis. Sei $\varepsilon > 0$. Mit der Dreiecksungleichung und 7.2(2) folgt: Es existiert ein $n_0 \in \mathbb{N}$ so, dass

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

Mit 7.2(5) folgt die Konvergenz von $\sum_{n=1}^{\infty} a_n$. Die Dreiecksungleichung für Reihen folgt mit $N \rightarrow \infty$ aus

$$\left| \sum_{n=1}^N a_n \right| \leq \sum_{n=1}^N |a_n|.$$

□

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

Beweis. Da (b_n) eine fallende Nullfolge ist gilt $b_n \geq 0$ ($n \in \mathbb{N}$). Sei $s_N := a_1 + \dots + a_N$ ($N \in \mathbb{N}$). Nachrechnen:

(s_{2N}) ist nach unten beschränkt und fallend,

(s_{2N-1}) ist nach oben beschränkt und wachsend.

Somit sind (s_{2N}) und (s_{2N-1}) konvergent. Sei $a := \lim_{N \rightarrow \infty} s_{2N}$ und $b := \lim_{N \rightarrow \infty} s_{2N-1}$. Es gilt

$$|s_{2N} - s_{2N-1}| = |a_{2N}| = b_{2N} \rightarrow 0 \quad (N \rightarrow \infty),$$

also $a = b$. Mit 6.8(3) folgt $s_N \rightarrow a$ ($N \rightarrow \infty$).

□

Beispiel: (1) Die *alternierende harmonische Reihe*

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} = - \sum_{n=1}^{\infty} \frac{(-1)^n}{n}$$

ist konvergent (denn $(1/n)$ ist eine fallende Nullfolge), aber **nicht** absolut konvergent. Später werden wir sehen:

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} = \log(2).$$

(2) Die geometrische Reihe $\sum_{n=0}^{\infty} z^n$ ist für jedes $z \in \mathbb{C}$ mit $|z| < 1$ absolut konvergent (beachte: $|z^n| = |z|^n$).

Bemerkungen: (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

$$a_n = b_n - c_n, \quad |a_n| = b_n + c_n \quad (n \in \mathbb{N}).$$

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 $\sum_{n=1}^{\infty} a_n = \sum_{n=1}^{\infty} b_n - \sum_{n=1}^{\infty} c_n$.

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

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

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

Beweis. (1) O.B.d.A. sei $|a_n| \leq b_n$ ($n \in \mathbb{N}$). Setze

$$\alpha_N := \sum_{n=1}^N |a_n|, \quad \beta_N := \sum_{n=1}^N b_n \quad (N \in \mathbb{N}).$$

Dann gilt $0 \leq \alpha_N \leq \beta_N$ ($N \in \mathbb{N}$). Nach Voraussetzung ist (β_N) beschränkt. Also ist (α_N) beschränkt und die Behauptung folgt mit 7.2(1).

(2) folgt aus (1). □

Beispiel: (1) Die Reihe

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

konvergiert: Es gilt

$$0 \leq \frac{1}{(n+1)^2} \leq \frac{1}{n(n+1)} \quad (n \in \mathbb{N})$$

und $\sum_{n=1}^{\infty} \frac{1}{n(n+1)}$ konvergiert nach 7.1(2). Somit konvergiert auch $\sum_{n=1}^{\infty} \frac{1}{n^2}$. In HMII werden wir sehen:

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}.$$

(2) Ist $q \in \mathbb{N}$ mit $q \geq 2$ so gilt $1/n^q \leq 1/n^2$ ($n \in \mathbb{N}$). Also ist

$$\sum_{n=1}^{\infty} \frac{1}{n^q}$$

konvergent.

(3) Es gilt

$$\frac{1}{\sqrt[n]{n}} \geq \frac{1}{n} \geq 0 \quad (n \in \mathbb{N}).$$

Da $\sum_{n=1}^{\infty} \frac{1}{n}$ divergiert ist auch $\sum_{n=1}^{\infty} \frac{1}{\sqrt[n]{n}}$ divergent.

(4) Ist später die allgemeine Potenz definiert, so sieht man für $q \in \mathbb{R}$:

$$\sum_{n=1}^{\infty} \frac{1}{n^q} \text{ konvergiert} \Leftrightarrow q > 1.$$

7.5. Wurzelkriterium: Es sei (a_n) eine Folge und $\alpha := \limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|} \in [0, \infty]$.

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

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

Bemerkung: Ist $\alpha = 1$, so liefert das Wurzelkriterium keine Entscheidung.

Z.B. ist $\sum_{n=1}^{\infty} \frac{1}{n}$ divergent und

$$\alpha = \limsup_{n \rightarrow \infty} \sqrt[n]{1/n} = \lim_{n \rightarrow \infty} 1/\sqrt[n]{n} = 1,$$

hingegen ist $\sum_{n=1}^{\infty} \frac{1}{n^2}$ konvergent und

$$\alpha = \limsup_{n \rightarrow \infty} \sqrt[n]{1/n^2} = \lim_{n \rightarrow \infty} 1/(\sqrt[n]{n})^2 = 1.$$

Beweis. Ist $\alpha < 1$, so wählen wir $\beta \in (\alpha, 1)$. Dann gilt

$$\sqrt[n]{|a_n|} \leq \beta \text{ ffa } n \in \mathbb{N}, \text{ also } |a_n| \leq \beta^n \text{ ffa } n \in \mathbb{N}.$$

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

$$\sqrt[n]{|a_n|} \geq \gamma \geq 1 \text{ für unendlich viele } n \in \mathbb{N}, \text{ also } |a_n| \geq 1 \text{ für unendlich viele } n \in \mathbb{N}.$$

Somit ist $a_n \not\rightarrow 0$, und nach 7.2(3) ist $\sum_{n=1}^{\infty} a_n$ divergent. □

Beispiel: Sei $p \in \mathbb{N}$. Wir betrachten $\sum_{n=1}^{\infty} n^p z^n$ für $z \in \mathbb{C}$. Hier ist also $a_n = n^p z^n$ und

$$\sqrt[n]{|a_n|} = (\sqrt[n]{n})^p |z| \rightarrow |z| \quad (n \rightarrow \infty).$$

Also ist $\sum_{n=1}^{\infty} n^p z^n$ für $|z| < 1$ absolut konvergent und für $|z| > 1$ divergent. Für $|z| = 1$ gilt $|a_n| = n^p \rightarrow \infty$ und $\sum_{n=1}^{\infty} n^p z^n$ ist nach 7.2(3) divergent.

7.6. Quotientenkriterium: Sei (a_n) eine Folge mit $a_n \neq 0$ ($n \geq n_0 \in \mathbb{N}$) und

$$c_n := \left| \frac{a_{n+1}}{a_n} \right| \quad (n \geq n_0).$$

- (1) Ist $c_n \geq 1$ für fast alle $n \in \mathbb{N}$, so ist $\sum_{n=1}^{\infty} a_n$ divergent.
- (2) Ist $\limsup_{n \rightarrow \infty} c_n < 1$, so ist $\sum_{n=1}^{\infty} a_n$ absolut konvergent.
- (3) Ist $\liminf_{n \rightarrow \infty} c_n > 1$, so ist $\sum_{n=1}^{\infty} a_n$ divergent.
- (4) Existiert $\alpha := \lim_{n \rightarrow \infty} c_n$, 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.

Bemerkung: In (4) ist im Fall $\alpha = 1$ **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 $n \geq n_1 \geq n_0$. Dann ist

$$|a_n| \geq |a_{n-1}| \geq \dots \geq |a_{n_1}| > 0 \quad (n \geq n_1),$$

d.h. $|a_n| \not\rightarrow 0$. Nach 7.2(3) divergiert $\sum_{n=1}^{\infty} a_n$.

(2) Ist $\alpha = \limsup_{n \rightarrow \infty} c_n < 1$, so wähle $\beta \in (\alpha, 1)$. Es gilt dann $c_n \leq \beta$ ffa $n \in \mathbb{N}$, d.h. wir finden $n_1 \geq n_0$ mit $c_n \leq \beta$ ($n \geq n_1$). Es folgt

$$|a_n| \leq \beta |a_{n-1}| \leq \dots \leq \beta^{n-n_1} |a_{n_1}| = \beta^n (|a_{n_1}| \beta^{-n_1}) \quad (n \geq n_1).$$

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.

(3) Ist $\alpha = \liminf_{n \rightarrow \infty} c_n > 1$, so folgt $c_n \geq 1$ ffa $n \in \mathbb{N}$ und $\sum_{n=1}^{\infty} a_n$ divergiert nach (1).

(4) folgt aus (2) und (3). □

7.7. Die Exponentialreihe: Die Reihe

$$\sum_{n=0}^{\infty} \frac{z^n}{n!} = 1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + \dots \quad (z \in \mathbb{C})$$

konvergiert absolut für jedes $z \in \mathbb{C}$:

Beweis. Für $z = 0$ ist $\sum_{n=0}^{\infty} \frac{0^n}{n!}$ offensichtlich absolut konvergent (und $= 1$). Es sei $z \neq 0$ und $a_n := z^n/n!$ ($n \in \mathbb{N}_0$). Dann gilt:

$$\left| \frac{a_{n+1}}{a_n} \right| = \left| \frac{z^{n+1}}{(n+1)!} \cdot \frac{n!}{z^n} \right| = \frac{|z|}{n+1} \rightarrow 0 \quad (n \rightarrow \infty).$$

Nach dem 7.6(4) konvergiert die Reihe also absolut. □

Satz: Es gilt

$$e = \sum_{n=0}^{\infty} \frac{1}{n!} \quad (\text{Eulersche Zahl, vgl. 6.6}).$$

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

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

Für $n \rightarrow \infty$ erhalten wir

$$e \leq \sum_{k=0}^{\infty} \frac{1}{k!}.$$

Andererseits ist für festes $m \in \mathbb{N}$ und $n \geq m$:

$$a_n = \left(1 + \frac{1}{n}\right)^n \geq \sum_{k=0}^m \frac{1}{k!} \cdot 1 \cdot \left(1 - \frac{1}{n}\right) \cdot \left(1 - \frac{2}{n}\right) \cdot \dots \cdot \left(1 - \frac{k-1}{n}\right),$$

woraus für $n \rightarrow \infty$ folgt: $e \geq b_m$. Für $m \rightarrow \infty$ erhalten wir

$$e \geq \sum_{k=0}^{\infty} \frac{1}{k!}.$$

□

Bemerkung: Später werden wir sehen:

$$\forall x \in \mathbb{R} : \lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n = \sum_{k=0}^{\infty} \frac{x^k}{k!}.$$

7.8. Umordnungen: Es sei (a_n) eine Folge und $\varphi : \mathbb{N} \rightarrow \mathbb{N}$ eine bijektive Abbildung. Setze $b_n := a_{\varphi(n)}$ ($n \in \mathbb{N}$). 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: Die Folge $(a_2, a_4, a_1, a_3, a_6, a_8, a_5, a_7, \dots)$ ist eine Umordnung von (a_n) (aber **keine** Teilfolge von (a_n)).

Satz: Es sei (b_n) eine Umordnung von (a_n) . Dann gilt:

(1) Ist (a_n) konvergent, so konvergiert auch (b_n) und

$$\lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} b_n.$$

(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 $a := \lim_{n \rightarrow \infty} a_n$. Sei $\varepsilon > 0$. Dann gilt $|a_n - a| < \varepsilon$ ffa $n \in \mathbb{N}$, also auch $|a_{\varphi(n)} - a| < \varepsilon$ ffa $n \in \mathbb{N}$ (hier wurde nur die Injektivität von φ verwendet).

(2) Wegen der Bemerkungen in 7.3 reicht es, den Fall $a_n \in \mathbb{R}$ und $a_n \geq 0$ ($n \in \mathbb{N}$) zu betrachten. Für jedes $N \in \mathbb{N}$ gilt

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

Daraus folgt die Behauptung. □

Riemannscher Umordnungssatz (ohne Beweis): Die reelle Reihe $\sum_{n=1}^{\infty} a_n$ sei konvergent, aber **nicht** absolut konvergent (man sagt dazu auch *bedingt 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.

7.9. Das Cauchyprodukt: Es seien $\sum_{n=0}^{\infty} a_n$ und $\sum_{n=0}^{\infty} b_n$ Reihen. Für jedes $n \in \mathbb{N}_0$ sei

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

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 $a_n \geq 0, b_n \geq 0$ ($n \in \mathbb{N}_0$) zu betrachten. Es gilt dann für jedes $N \in \mathbb{N}_0$:

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

woraus die Behauptung folgt. $\square$

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

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

Nach obigem Satz konvergiert diese Reihe absolut und es gilt

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

7.10. Die Exponentialfunktion: Da die Exponentialreihe nach 7.7 für jedes $z \in \mathbb{C}$ (absolut) konvergiert, können wir durch

$$\exp : \mathbb{C} \rightarrow \mathbb{C}, \quad z \mapsto \exp(z) := \sum_{n=0}^{\infty} \frac{z^n}{n!}$$

eine Funktion definieren; $\exp$ heißt die *komplexe Exponentialfunktion*.

Eigenschaften der Exponentialfunktion:

- (1) Es gilt $\exp(0) = 1$ und $\exp(1) = e$ (vgl. 7.7).
- (2) Für alle $z, w \in \mathbb{C}$ gilt die *Funktionalgleichung* der Exponentialfunktion:

$$\exp(z) \exp(w) = \exp(z + w).$$

Beweis. Nach 7.9 gilt

$$\exp(z) \exp(w) = \left(\sum_{n=0}^{\infty} \frac{z^n}{n!} \right) \left(\sum_{n=0}^{\infty} \frac{w^n}{n!} \right) = \sum_{n=0}^{\infty} c_n$$

mit

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

Damit folgt die Behauptung. $\square$

Folgerung: Es gilt

$$\exp(z_1) \exp(z_2) \cdots \exp(z_m) = \exp(z_1 + z_2 + \cdots + z_m) \quad (z_1, \dots, z_m \in \mathbb{C}).$$

(3) Für alle $z \in \mathbb{C}$ und $n \in \mathbb{Z}$ gilt

$$\exp(z) \neq 0, \quad \exp(-z) = \frac{1}{\exp(z)}, \quad \exp(z)^n = \exp(nz).$$

Beweis. Es ist $1 = \exp(0) = \exp(z + (-z)) = \exp(z) \exp(-z)$, woraus $\exp(z) \neq 0$ und $\exp(-z) = \exp(z)^{-1}$ folgt. Der Rest folgt aus (2). $\square$

(4) Es gilt $\exp(x) \in \mathbb{R}$, $\exp(x) > 0$ ($x \in \mathbb{R}$), und $\exp(x) > 1$ ($x > 0$).

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

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

Also ist für $x < 0$ nach (3): $\exp(x) = \exp(-x)^{-1} \in (0, 1)$. $\square$

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

Beweis. Ist $x < y$, so gilt nach (2) und (4):

$$\exp(y) = \exp(y - x) \exp(x) > 1 \cdot \exp(x).$$

$\square$

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

Beweis. Für jedes $x > 0$ gilt

$$\exp(x) \geq 1 + x > x \quad \wedge \quad \exp(-x) = \frac{1}{\exp(x)} < \frac{1}{x},$$

woraus die Behauptungen folgen. $\square$

(7) Für alle $z \in \mathbb{C}$ gilt $\exp(\bar{z}) = \overline{\exp(z)}$.

Beweis. Sei $w_N := \sum_{n=0}^N \frac{z^n}{n!}$. Dann gilt $w_N \rightarrow \exp(z)$, also $\overline{w_N} \rightarrow \overline{\exp(z)}$, sowie

$$\overline{w_N} = \sum_{n=0}^N \frac{\bar{z}^n}{n!} \rightarrow \exp(\bar{z}).$$

$\square$

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

Beweis. Die erste Gleichung folgt aus (2). Mit (7), (2) und (1) folgt:

$$|\exp(iy)|^2 = \exp(iy) \overline{\exp(iy)} = \exp(iy) \exp(-iy) = \exp(iy - iy) = 1.$$

□

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

Beweis. Nach (4) ist $\exp(x) > 0$, $\exp(\frac{x}{n}) > 0$, und es gilt nach (2):

$$\left(\exp\left(\frac{x}{n}\right)\right)^n = \exp\left(n \frac{x}{n}\right) = \exp(x).$$

□

(10) Es gilt $|\exp(z)| = \exp(\operatorname{Re} z)$ ($z \in \mathbb{C}$).

Beweis. Nach (7), (2) und (9) ist:

$$|\exp(z)| = \sqrt{\exp(z) \overline{\exp(z)}} = \sqrt{\exp(z) \exp(\bar{z})} = \sqrt{\exp(2 \operatorname{Re} z)} = \exp(\operatorname{Re} z).$$

□

Bemerkung und Definition: Wegen (1), (2), (3) und (9) gilt für alle $m \in \mathbb{Z}$ und alle $n \in \mathbb{N}$:

$$e^m = \exp(m), \quad \sqrt[n]{e} = \exp\left(\frac{1}{n}\right), \quad \sqrt[n]{e^m} = \exp\left(\frac{m}{n}\right) = (\sqrt[n]{e})^m.$$

Man schreibt deshalb auch $e^z := \exp(z)$ ($z \in \mathbb{C}$).

Wir zeigen noch die folgenden **Ungleichungen**:

(11) Es gilt

$$|\exp(z) - 1| \leq |z| \exp(|z|) \quad (z \in \mathbb{C}).$$

(12) Es gilt

$$\left| \frac{\exp(z) - 1}{z} - 1 \right| \leq |z| \exp(|z|) \quad (z \in \mathbb{C} \setminus \{0\}).$$

Beweis. Es ist $\exp(z) = 1 + z + \frac{z^2}{2!} + \frac{z^3}{3!} + \dots$. Also

$$|\exp(z) - 1| = \left| \sum_{n=1}^{\infty} \frac{z^n}{n!} \right| \leq \sum_{n=1}^{\infty} \frac{|z|^n}{n!} \leq |z| \sum_{n=1}^{\infty} \frac{|z|^{n-1}}{(n-1)!} = |z| \exp(|z|).$$

Somit gilt (11). Für $z \neq 0$ ist

$$\begin{aligned} \left| \frac{\exp(z) - 1}{z} - 1 \right| &= \left| \sum_{n=1}^{\infty} \frac{z^{n-1}}{n!} - 1 \right| \leq \sum_{n=2}^{\infty} \frac{|z|^{n-1}}{n!} \\ &= |z| \sum_{n=2}^{\infty} \frac{|z|^{n-2}}{n!} \leq |z| \sum_{n=2}^{\infty} \frac{|z|^{n-2}}{(n-2)!} = |z| \exp(|z|). \end{aligned}$$

□

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

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

(1) Es gilt $\sin 0 = 0$ und $\cos 0 = 1$. [folgt aus 7.10(1)]

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

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

Das folgt aus

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

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

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

Beweis. Aus (2) 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 7.10(8).

□

(4) **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 aus 7.10(2) (Übung).

Wir zeigen noch die folgenden **Ungleichungen**:

- (5) Für alle $z \in \mathbb{C}$ gilt: $|\sin z| \leq |z| \exp(|z|)$.
- (6) Für alle $z \in \mathbb{C}$ gilt: $|\cos z - 1| \leq |z| \exp(|z|)$.
- (7) Für alle $z \in \mathbb{C} \setminus \{0\}$ gilt: $|\frac{\sin z}{z} - 1| \leq |z| \exp(|z|)$.
- (8) Für alle $z \in \mathbb{C} \setminus \{0\}$ gilt: $|\frac{\cos z - 1}{z}| \leq |z| \exp(|z|)$.

Beweis. Diese Ungleichungen folgen aus (11) und (12) in 7.10, wenn man beachtet:

$$\begin{aligned}\sin z &= \frac{1}{2i}(\exp(iz) - 1 - (\exp(-iz) - 1)), \\ \cos z - 1 &= \frac{1}{2}(\exp(iz) - 1 + (\exp(-iz) - 1)), \\ \frac{\sin z}{z} - 1 &= \frac{\exp(iz) - \exp(-iz)}{2iz} - 1 = \frac{1}{2} \left(\frac{\exp(iz) - 1}{iz} - 1 + \frac{\exp(-iz) - 1}{-iz} - 1 \right), \\ \frac{\cos z - 1}{z} &= \frac{i \exp(iz) - 1 + \exp(-iz) - 1}{2iz} \\ &= \frac{i}{2} \left(\frac{\exp(iz) - 1}{iz} - 1 - \left(\frac{\exp(-iz) - 1}{-iz} - 1 \right) \right).\end{aligned}$$

□

7.12. Potenzreihen: Sei $z_0 \in \mathbb{C}$. Eine *Potenzreihe (PR)* um z_0 ist eine Reihe der 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

$$\sum_{n=0}^{\infty} a_n (x - x_0)^n.$$

Frage: Für welche $z \in \mathbb{C}$ konvergiert eine gegebene Potenzreihe (absolut)? Klar: Eine Potenzreihe konvergiert absolut in $z = z_0$.

Beispiele: (1) Die Potenzreihen für $\exp$, $\sin$ und $\cos$ konvergieren absolut für jedes $z \in \mathbb{C}$. Hier: $z_0 = 0$.

(2) Die geometrische Reihe $\sum_{n=0}^{\infty} z^n$ ist für $|z| < 1$ absolut konvergent und für $|z| \geq 1$ divergent. Hier: $z_0 = 0$ und $a_n = 1$ ($n \in \mathbb{N}_0$).

(3) Die Potenzreihe

$$\sum_{n=0}^{\infty} n^n z^n$$

konvergiert nur in $z = z_0 = 0$: Für $z \neq 0$ gilt

$$\sqrt[n]{n^n |z|^n} = n|z| \rightarrow \infty \quad (n \rightarrow \infty).$$

Nach 7.5 ist die Reihe dann divergent.

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

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

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 $|z - z_0| > R$.

R heißt *Konvergenzradius* (KR) der Potenzreihe.

Beweis: Für $z \in \mathbb{C} \setminus \{z_0\}$ gilt:

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

und die Behauptungen des Satzes folgen aus dem Wurzelkriterium 7.5. □

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.

¹Beachte, dass $\alpha := \limsup \sqrt[n]{|a_n|} \in [0, \infty]$ und dass also diese Konventionen für $1/\alpha \in [0, \infty]$ sinnvoll sind, vgl. 6.9.

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.

(c) Die Potenzreihe $\sum_{n=1}^{\infty} \frac{z^n}{n}$ hat Konvergenzradius $R = 1$, für jedes $z \in \mathbb{C}$ mit $|z| = 1$, $z \neq 1$ ist sie konvergent (ohne Beweis, bekannt für $z = -1$), für $z = 1$ ist sie divergent.

Folgerungen: (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 ∞ . Damit folgt u.a.: $\limsup_{n \rightarrow \infty} \sqrt[n]{1/n!} = 0$ und mit 6.10

$$\lim_{n \rightarrow \infty} \frac{1}{\sqrt[n]{n!}} = 0.$$

(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 $R = 0$.

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!

7.14. Satz (Konvergenzradius über Quotienten): Es sei

$$\sum_{n=0}^{\infty} a_n(z - z_0)^n$$

eine Potenzreihe mit $a_n \neq 0$ ffa $n \in \mathbb{N}$. Gilt

$$\left| \frac{a_{n+1}}{a_n} \right| \rightarrow \alpha \in [0, \infty],$$

so ist der Konvergenzradius $R = \frac{1}{\alpha}$.

Beweis. Die Behauptung folgt aus dem Quotientenkriterium (Bemerkung in 7.6). □

Beispiel: Für die Exponentialreihe ist $a_n = 1/n!$, also $\left| \frac{a_{n+1}}{a_n} \right| = \frac{1}{n+1} \rightarrow 0$ und $R = \frac{1}{0} = \infty$. Auf die Potenzreihen für $\sin$ und $\cos$ aus 7.11 lässt sich der Satz nicht anwenden.

Bemerkung: Hier erlauben $\limsup_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right|$ und $\liminf_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right|$ i.a. keine Entscheidung!

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

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

also $\limsup_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| = \infty$ und $\liminf_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| = 0$, aber der Konvergenzradius ist $R = 1$, denn

$$\limsup_{n \rightarrow \infty} \sqrt[n]{|a_n|} = \limsup_{n \rightarrow \infty} \frac{2 + (-1)^n}{3} = 1.$$

7.15. g -adische Entwicklung reeller Zahlen: Sei $g \in \mathbb{N}$ mit $g \geq 2$. Ist $(a_n)_{n=p}^\infty$ eine Folge in $\{0, 1, \dots, g-1\}$ mit $p \leq 0$, so definiert man

$$a_p a_{p+1} \dots a_0, a_1 a_2 a_3 \dots := \sum_{n=p}^{\infty} \frac{a_n}{g^n}$$

und nennt dies einen g -adischen Bruch, wobei aus dem Zusammenhang klar sein muss, was g hier sein soll. Jede reelle Zahl $x \geq 0$ kann als g -adischer Bruch geschrieben werden, dieser ist aber nicht eindeutig: Z.B. ist $g = 10$, so gilt (Übung):

$$1 = 1,00000\dots = 0,99999\dots$$

Man kann Eindeutigkeit herstellen indem man “ $a_n = g-1$ ffa $n \geq p$ ” verbietet. Wir betrachten den Fall $x \in [0, 1)$:

Ohne Beweis: Für jedes $x \in [0, 1)$ gibt es eine eindeutig bestimmte Folge $(a_n)_{n \in \mathbb{N}}$ in $\{0, 1, \dots, g-1\}$ so, dass für jedes $N \in \mathbb{N}$ gilt:

$$\sum_{n=1}^N \frac{a_n}{g^n} \leq x < \left(\sum_{n=1}^N \frac{a_n}{g^n} \right) + \frac{1}{g^N}. \quad (1)$$

Man hat dann

$$x = \sum_{n=1}^{\infty} \frac{a_n}{g^n} \text{ also } x = 0, a_1 a_2 a_3 \dots, \quad (2)$$

mit $a_n \neq g-1$ für unendlich viele $n \in \mathbb{N}$.

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

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.

8.1. Vektorraumaxiome:

Im folgenden sei stets $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$,

Definition: Es sei $V \neq \emptyset$ eine Menge mit zwei Verknüpfungen

$$+ : V \times V \rightarrow V, \quad \cdot : \mathbb{K} \times V \rightarrow V.$$

Die Menge V (genauer $(V, +, \cdot)$) heißt ein *Vektorraum über $\mathbb{K}$* oder ein *$\mathbb{K}$ -Vektorraum* ($\mathbb{K}$ -VR), wenn die folgenden Eigenschaften (V1) – (V7) erfüllt sind:

- (V1) $\forall x, y, z \in V : (x + y) + z = x + (y + z)$ (Assoziativität der Addition),
- (V2) $\forall x, y \in V : x + y = y + x$ (Kommutativität der Addition),
- (V3) $\exists 0 \in V \forall x \in V : x + 0 = x$ (Existenz der Null),
- (V4) $\forall x \in V \exists -x \in V : x + (-x) = 0$ (Existenz des Negativen),
- (V5) $\forall \alpha, \beta \in \mathbb{K} \forall x \in V : (\alpha\beta) \cdot x = \alpha \cdot (\beta \cdot x)$ (Assoziativität der Multiplikationen),
- (V6) $\forall \alpha, \beta \in \mathbb{K} \forall x, y \in V : \alpha \cdot (x + y) = \alpha \cdot x + \alpha \cdot y$ und $(\alpha + \beta) \cdot x = \alpha \cdot x + \beta \cdot x$ (Distributivität),
- (V7) $\forall x \in V : 1 \cdot x = x$ (Kompatibilität).

Die Elemente eines Vektorraumes heißen *Vektoren*. Der Punkt $\cdot$ für die Multiplikation mit Skalaren wird gewöhnlich weggelassen, und $x - y := x + (-y)$.

Satz: (Beweis als Übung) Es sei V ein $\mathbb{K}$ -Vektorraum. Dann gilt:

- (1) Es existiert genau ein $0 \in V$ mit $x + 0 = x$ ($x \in V$),
- (2) zu jedem $x \in V$ existiert genau ein $-x \in V$ mit $x + (-x) = 0$,
- (3) $\forall x \in V : -x = (-1) \cdot x$,
- (4) $\forall x \in V : 0 \cdot x = 0$.

Definition: Es seien A, B nichtleere Mengen.

$$B^A := \text{Menge alle Abbildungen } f : A \rightarrow B.$$

Im Fall $A = \{1, \dots, n\}$, $n \in \mathbb{N}$ schreiben wir B^n für $B^{\{1, \dots, n\}}$.

Beispiele: (1) Sei $n \in \mathbb{N}$. Dann ist

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

Hier sind Abbildungen $x : \{1, \dots, n\} \rightarrow \mathbb{K}$ als $(x_1, x_2, \dots, x_n)$ geschrieben, d.h. $x(k) = x_k$. Wir erklären Verknüpfungen

$$+ : \mathbb{K}^n \times \mathbb{K}^n \rightarrow \mathbb{K}^n, \quad \cdot : \mathbb{K} \times \mathbb{K}^n \rightarrow \mathbb{K}^n$$

durch

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

Z.B. in $\mathbb{R}^2$:

$$(1, 1) + (0, 1) = (1, 2), \quad 2 \cdot (1, -1) = (2, -2),$$

oder in $\mathbb{C}^3$:

$$(1, i, 0) + i(1, 2, i) = (1 + i, 3i, -1).$$

Mit obigen Verknüpfungen ist $\mathbb{K}^n$ ein $\mathbb{K}$ -Vektorraum mit Nullelement

$$(0, 0, \dots, 0)$$

und

$$-(x_1, \dots, x_n) = (-x_1, \dots, -x_n) \quad ((x_1, \dots, x_n) \in \mathbb{K}^n).$$

Bemerkung: $\mathbb{C}^n$ ist ein $\mathbb{C}$ -Vektorraum, aber $\mathbb{C}^n$ ist auch ein $\mathbb{R}$ -Vektorraum, wenn die Skalarmultiplikation auf $\mathbb{R}$ eingeschränkt wird. Insbesondere ist $\mathbb{C}$ ein $\mathbb{R}$ -Vektorraum. Allgemein ist jeder $\mathbb{C}$ -Vektorraum auch ein $\mathbb{R}$ -Vektorraum.

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

8.2. Wichtiges Beispiel: Es sei $I \neq \emptyset$ eine nichtleere Menge (häufig ist z.B. $I \subseteq \mathbb{R}$ ein Intervall). Setzt man für $f, g \in \mathbb{K}^I$ und $\alpha \in \mathbb{K}$:

$$\begin{aligned} f + g : I &\rightarrow \mathbb{K}, & t &\mapsto f(t) + g(t), \\ \alpha \cdot f : I &\rightarrow \mathbb{K}, & t &\mapsto \alpha f(t), \end{aligned}$$

so ist $\mathbb{K}^I$ bzgl. dieser Verknüpfungen $+$: $\mathbb{K}^I \times \mathbb{K}^I \rightarrow \mathbb{K}^I$ und $\cdot$: $\mathbb{K} \times \mathbb{K}^I \rightarrow \mathbb{K}^I$ ein $\mathbb{K}$ -Vektorraum. Für $f, g \in \mathbb{K}^I$ und $\alpha \in \mathbb{K}$ sind $f + g$ und αf also *punktweise* erklärt, d.h. 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 $f : [a, b] \rightarrow \mathbb{K}$ ist ein $\mathbb{K}$ -Vektorraum.

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

$$\mathbb{K}^{\mathbb{N}} = \{(x_1, x_2, \dots) : x_j \in \mathbb{K} \ (j \in \mathbb{N})\}$$

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

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

8.3. Untervektorräume: Es sei V ein $\mathbb{K}$ -Vektorraum.

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

Satz: (Beweis als Übung) Eine Teilmenge $U \subseteq V$ ist UVR von V genau dann, wenn gilt:

- (i) $U \neq \emptyset$ und
- (ii) $\forall x, y \in U \forall \alpha \in \mathbb{K} : x + y \in U$ und $\alpha x \in U$.

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

Beispiele: (1) Jeder Vektorraum V hat die trivialen Untervektorräume $\{0\}$ und V . Sind $U_1, U_2 \subseteq V$ UVR von V , so ist auch $U_1 \cap U_2$ ein UVR von V .

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

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

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

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

(6) $U_1 = \{(x, 0) : x \in \mathbb{R}\}$ und $U_2 = \{(0, y) : y \in \mathbb{R}\}$ sind Untervektorräume von $\mathbb{R}^2$, aber $U_1 \cup U_2$ ist kein UVR von $\mathbb{R}^2$, denn

$$(1, 0), (0, 1) \in U_1 \cup U_2, \text{ aber } (1, 1) = (1, 0) + (0, 1) \notin U_1 \cup U_2.$$

Aus den Rechenregeln für konvergente Folgen und Reihen ergeben sich mit obigem Satz folgende Beispiele für Untervektorräume von $\mathbb{K}^{\mathbb{N}}$.

(7) Der Raum der konvergenten Folgen

$$c := \{(x_n) \in \mathbb{K}^{\mathbb{N}} : (x_n) \text{ ist konvergent}\}$$

ist ein UVR von $\mathbb{K}^{\mathbb{N}}$. Die Menge der Nullfolgen

$$c_0 := \{(x_n) \in c : \lim_{n \rightarrow \infty} x_n = 0\}$$

ist ein UVR von c .

(8) Der Raum der absolut summierbaren Folgen

$$l^1 := \{(x_n) \in \mathbb{K}^{\mathbb{N}} : \sum_{n=1}^{\infty} |x_n| < \infty\}$$

ist ein UVR von $\mathbb{K}^{\mathbb{N}}$ oder auch von c_0 .

(9) Der Raum der beschränkten Folgen

$$l^\infty := \{(x_n) \in \mathbb{K}^\mathbb{N} : \sup\{|x_n| : n \in \mathbb{N}\} < \infty\}$$

ist ein UVR von $\mathbb{K}^\mathbb{N}$. Der Raum c ist ein UVR von l^∞ .

Allgemeiner ist für jede nichtleere Menge X der Raum der beschränkten Funktionen

$$B(X) := \{f \in \mathbb{K}^X : \|f\|_\infty < \infty\}$$

ein UVR von $\mathbb{K}^X$. Hier ist $\|f\|_\infty := \sup\{|f(x)| : x \in X\}$ die *Supremumsnorm* von f auf X . Man überzeuge sich von folgenden Eigenschaften: Für alle $f, g \in B(X)$ und $\alpha \in \mathbb{K}$ gilt:

- (a) $f = 0 \Leftrightarrow \|f\|_\infty = 0$,
- (b) $\|\alpha f\|_\infty = |\alpha| \|f\|_\infty$,
- (c) $\|f + g\|_\infty \leq \|f\|_\infty + \|g\|_\infty$.

8.4. Lineare Abbildungen: Es seien V, W $\mathbb{K}$ -Vektorräume.

Definition: Eine Abbildung $\phi : V \rightarrow W$ heißt *linear* (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

$$\text{Kern } \phi := \{x \in V : \phi(x) = 0\} = \phi^{-1}(\{0\})$$

der *Kern* von ϕ und

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

heißt *Bild* von ϕ .

Ist $\phi : V \rightarrow W$ linear, so gilt immer $\phi(0) = 0$. Mithilfe von Satz 8.3 sieht man leicht, dass Kern ϕ ein UVR von V ist und dass Bild ϕ ein UVR 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 $\alpha_0 \in \mathbb{K}$, so ist die Abbildung $V \rightarrow V$, $x \mapsto \alpha_0 x$ linear. Ist $y_0 \in V$, so ist die Abbildung $V \rightarrow V$, $x \mapsto x + y_0$ (Verschiebung um y_0) nur im Fall $y_0 = 0$ linear.

(2) Ist $n \in \mathbb{N}$, so ist für jedes $j \in \{1, 2, \dots, n\}$ die Abbildung auf die j -te Komponente

$$\phi_j : \mathbb{K}^n \rightarrow \mathbb{K}, \quad (x_1, \dots, x_n) \mapsto x_j$$

linear.

(3) Die Abbildung

$$\phi : c \rightarrow \mathbb{K}, \quad (x_n) \mapsto \lim_{n \rightarrow \infty} x_n,$$

die jede konvergente Folge auf ihren Grenzwert abbildet, ist nach 6.3(4) linear. Hier ist Kern $\phi = c_0$ (der Raum der Nullfolgen).

(4) Die Abbildung

$$\phi : l^1 \rightarrow \mathbb{K}, \quad (x_n) \mapsto \sum_{n=1}^{\infty} x_n$$

ist linear nach 7.2(4).

(5) Die Abbildung

$$\phi : l^{\infty} \rightarrow c_0, \quad (x_n) \mapsto \left(\frac{x_n}{n}\right)$$

ist linear und Bild ϕ ist ein UVR von c_0 . Aber Bild $\phi \neq c_0$: Z.B. ist $(\frac{1}{\sqrt{n}}) \notin \text{Bild } \phi$, da $(\sqrt{n}) \notin l^{\infty}$.

(6) Die Abbildung

$$\phi : l^1 \rightarrow \mathbb{K}, \quad (x_n) \mapsto \sum_{n=1}^{\infty} x_n^2$$

ist **nicht** linear.

9 Stetigkeit

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

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

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

9.2. Beispiele: (1) Ist $p \in \mathbb{R}[x]$ ein Polynom, d.h.

$$p(x) = a_m x^m + a_{m-1} x^{m-1} + \dots + a_1 x + a_0,$$

so ist die Funktion $p : \mathbb{R} \rightarrow \mathbb{R}$, $x \mapsto p(x)$ auf $\mathbb{R}$ stetig. Sei $x_0 \in \mathbb{R}$ und (x_n) eine Folge mit $x_n \rightarrow x_0$. Aus 6.3 folgt $x_n^k \rightarrow x_0^k$ ($k = 1, \dots, m$) und somit $p(x_n) \rightarrow p(x_0)$. Also gilt $p \in C(\mathbb{R})$.

(2) Betrachte $p \in \mathbb{N}$, $D := [0, \infty)$ und $f : D \rightarrow \mathbb{R}$, $f(x) = \sqrt[p]{x}$. Sei $x_0 \in D$ und (x_n) eine Folge in D mit $x_n \rightarrow x_0$. Nach 6.5 folgt $f(x_n) \rightarrow f(x_0)$. Also gilt $f \in C([0, \infty))$.

(3) Sei $D = [0, 1] \cup \{2\}$ und $f : D \rightarrow \mathbb{R}$ definiert durch

$$f(x) = \begin{cases} x^2, & x \in [0, 1) \\ \frac{1}{2}, & x = 1 \\ 1, & x = 2 \end{cases}$$

Offensichtlich ist f in jedem $x \in [0, 1)$ stetig.

Betrachte $x_0 = 1$: Die Folge $(x_n) = (1 - 1/n)$ ist eine Folge in D mit $x_n \rightarrow 1$ und $f(x_n) = x_n^2 \rightarrow 1 \neq 1/2 = f(1)$. Also ist f in $x_0 = 1$ nicht stetig.

Betrachte $x_0 = 2$: Sei (x_n) eine Folge in D mit $x_n \rightarrow 2$. Dann gilt $x_n = 2$ ffa $n \in \mathbb{N}$, also $f(x_n) = 1$ ffa $n \in \mathbb{N}$. Es folgt $f(x_n) \rightarrow 1 = f(2)$. Somit ist f in $x_0 = 2$ stetig.

Mit den Rechenregeln für konvergente Folgen in 6.3 folgt:

9.3. Satz: Seien $f, g : D \rightarrow \mathbb{R}$ in $x_0 \in D$ stetig und $\alpha, \beta \in \mathbb{R}$. Dann sind auch

$$\alpha f + \beta g, f \cdot g, |f| : D \rightarrow \mathbb{R}$$

in x_0 stetig. Ist $g(x_0) \neq 0$, so ist auch

$$\frac{f}{g} : \{x \in D : g(x) \neq 0\} \rightarrow \mathbb{R}$$

in x_0 stetig.

Folgerung: Ist $\emptyset \neq D \subseteq \mathbb{R}$, so ist die Menge $C(D)$ der stetigen Funktionen $f : D \rightarrow \mathbb{R}$ ein reeller Vektorraum (und ist ein UVR des Vektorraumes $\mathbb{R}^D$ aller Funktionen $f : D \rightarrow \mathbb{R}$).

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

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

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

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

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

$$\frac{p}{q} : \{x \in \mathbb{R} : q(x) \neq 0\} \rightarrow \mathbb{R}, \quad x \mapsto \frac{p(x)}{q(x)}$$

stetig.

9.4. ε - δ -Kriterium: Sei $\emptyset \neq D \subseteq \mathbb{R}$, $f : D \rightarrow \mathbb{R}$ eine Funktion und $x_0 \in D$. Dann ist f in x_0 stetig genau dann, wenn gilt:

$$\forall \varepsilon > 0 \exists \delta > 0 \forall x \in D : |x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \varepsilon.$$

Beweis. Es gelte “ ε - δ ”. Es sei (x_n) eine Folge in D mit $x_n \rightarrow x_0$. Sei $\varepsilon > 0$. Dann finden wir ein $\delta > 0$ mit $|f(x) - f(x_0)| < \varepsilon$ für alle $x \in D$ mit $|x - x_0| < \delta$. Wegen $x_n \rightarrow x_0$ gilt $|x_n - x_0| < \delta$ ffa $n \in \mathbb{N}$. Es folgt $|f(x_n) - f(x_0)| < \varepsilon$ ffa $n \in \mathbb{N}$. Also gilt $f(x_n) \rightarrow f(x_0)$.

Wenn “ ε - δ ” nicht gilt, finden wir ein $\varepsilon > 0$ so, dass zu jedem $n \in \mathbb{N}$ (zu $\delta = 1/n$) ein $x_n \in D$ existiert mit

$$|x_n - x_0| < \frac{1}{n} \quad \text{und} \quad |f(x_n) - f(x_0)| \geq \varepsilon.$$

Es gilt dann $x_n \rightarrow x_0$, aber $f(x_n) \not\rightarrow f(x_0)$. Somit ist f in x_0 nicht stetig. $\square$

Bemerkung: Die ε - δ -Bedingung findet sich in der Literatur auch oft als Definition der Stetigkeit.

Definition: Die Funktion f heißt *gleichmäßig stetig in/auf* D , falls gilt: Sind $(x_n), (y_n)$ Folgen in D mit $x_n - y_n \rightarrow 0$, so gilt $f(x_n) - f(y_n) \rightarrow 0$. Äquivalent ist:

$$\forall \varepsilon > 0 \exists \delta > 0 \forall x, y \in D : |x - y| < \delta \Rightarrow |f(x) - f(y)| < \varepsilon.$$

Bemerkung: Ist f gleichmäßig stetig auf D , so gilt $f \in C(D)$. Die Umkehrung gilt nicht. Z.B. ist $f : \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = x^2$ stetig auf $\mathbb{R}$, aber nicht gleichmäßig stetig auf $\mathbb{R}$ (Übung). Im Fall $D = [a, b]$ ist “Stetigkeit auf D ” und “gleichmäßige Stetigkeit auf D ” dasselbe (Satz von Heine).

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

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

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

Beispiele: (1) Aus den Ungleichungen in 7.10 und 7.11 folgt etwa:

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

(2) Sei $m \in \mathbb{N}_0$. Aus der Ungleichung

$$\exp(x) = \sum_{n=0}^{\infty} \frac{x^n}{n!} \geq \frac{x^{m+1}}{(m+1)!} \quad (x > 0)$$

folgen

$$\frac{\exp(x)}{x^m} \geq \frac{x}{(m+1)!} \quad (x > 0)$$

sowie

$$0 \leq x^m \exp(-x) \leq \frac{(m+1)!}{x} \quad (x > 0).$$

Also $\frac{\exp(x)}{x^m} \rightarrow \infty$ ($x \rightarrow \infty$) und $x^m \exp(-x) \rightarrow 0$ ($x \rightarrow \infty$). Insbesondere gilt (mit $m = 0$):

$$\lim_{x \rightarrow \infty} \exp(x) = \infty, \quad \lim_{x \rightarrow -\infty} \exp(x) = 0.$$

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

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

(3) Die Rechenregeln für den Folgenlimes aus 6.3 und 6.9 übertragen sich auf den Funktionenlimes. Gilt z.B.

$$\lim_{x \rightarrow \alpha} f(x) = \beta_1, \quad \lim_{x \rightarrow \alpha} g(x) = \beta_2$$

und ist $\beta_1 + \beta_2$ definiert, so gilt

$$\lim_{x \rightarrow \alpha} (f(x) + g(x)) = \beta_1 + \beta_2.$$

Definition: Ist $\alpha = x_0 \in \mathbb{R}$ und gibt es eine Folge in $D \setminus \{x_0\}$ mit $x_n \rightarrow x_0$, 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 $(0, 1)$, 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

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

Ein solches x_0 heißt *isolierter Punkt von D* . So ist etwa 0 ein isolierter Punkt von $\mathbb{Z}$.

Eine Funktion $f : D \rightarrow \mathbb{R}$ ist in jedem isolierten Punkt x_0 von D stetig, vgl. Beispiel (3) in 9.2.

Satz: Die Funktion $f : D \rightarrow \mathbb{R}$ ist stetig auf D genau dann, wenn für jedes $x_0 \in D$, welches Häufungspunkt von D ist, gilt:

$$\lim_{x \rightarrow x_0} f(x) = f(x_0).$$

9.6. Einseitige Grenzwerte: Es sei $\emptyset \neq D \subseteq \mathbb{R}$, $f : D \rightarrow \mathbb{R}$ eine Funktion, und $x_0 \in \mathbb{R}$ sei ein Häufungspunkt von $D \cap (x_0, \infty)$ [bzw. von $D \cap (-\infty, x_0)$]. Wir schreiben

$$\lim_{x \rightarrow x_0+} f(x) = \beta \in \mathbb{R} \cup \{\infty, -\infty\} \quad [\text{bzw.} \quad \lim_{x \rightarrow x_0-} f(x) = \beta \in \mathbb{R} \cup \{\infty, -\infty\}],$$

falls für jede Folge (x_n) in $D \cap (x_0, \infty)$ [bzw. in $D \cap (-\infty, x_0)$] mit $x_n \rightarrow x_0$ gilt: $f(x_n) \rightarrow \beta$.

Im Falle der Existenz heißt $\lim_{x \rightarrow x_0+} f(x)$ *rechtsseitiger Limes* und $\lim_{x \rightarrow x_0-} f(x)$ heißt *linksseitiger Limes* von f in x_0 .

Bemerkung: Man kann sich für den rechtsseitigen Limes auf monoton fallende Folgen und für den linksseitigen Limes auf monoton wachsende Folgen (x_n) beschränken!

Beispiel: Sei $D = \mathbb{R}$ und $f : \mathbb{R} \rightarrow \mathbb{R}$ gegeben durch

$$f(x) := \begin{cases} 1, & x > 0 \\ -7, & x = 0 \\ \frac{1}{x}, & x < 0 \end{cases}.$$

Dann gilt $\lim_{x \rightarrow 0+} f(x) = 1$ und $\lim_{x \rightarrow 0-} f(x) = -\infty$.

Satz: Sei $f : D \rightarrow \mathbb{R}$ eine Funktion und $x_0 \in D$ Häufungspunkt sowohl von $D \cap (-\infty, x_0)$ als auch von $D \cap (x_0, \infty)$. Dann ist f stetig in x_0 genau dann, wenn gilt:

$$\lim_{x \rightarrow x_0+} f(x) = f(x_0) = \lim_{x \rightarrow x_0-} f(x).$$

9.7. Satz zur Komposition stetiger Funktionen: Es seien $D_1, D_2 \subseteq \mathbb{R}$ nichtleer, $f : D_1 \rightarrow \mathbb{R}$ und $g : D_2 \rightarrow \mathbb{R}$ 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 $g \circ f : D_1 \rightarrow \mathbb{R}$ stetig in x_0 .

Beweis. Sei (x_n) eine Folge in D_1 mit $x_n \rightarrow x_0$. Da f in x_0 stetig ist, folgt $y_n := f(x_n) \rightarrow f(x_0) = y_0$. Da g stetig in y_0 ist, folgt

$$(g \circ f)(x_n) = g(f(x_n)) = g(y_n) \rightarrow g(y_0) = g(f(x_0)) = (g \circ f)(x_0).$$

Also ist $g \circ f$ in x_0 stetig. □

Beispiel: Sei $p \in \mathbb{N}$. Die Funktion $g_p : [0, \infty) \rightarrow \mathbb{R}$, $g_p(x) = \sqrt[p]{x}$ ist auf $[0, \infty)$ stetig (nach 9.2(2)). Die Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = x^2 + 1$ ist auf $\mathbb{R}$ stetig und $f(x) \in [0, \infty)$ ($x \in \mathbb{R}$). Also ist auch $g_p \circ f : \mathbb{R} \rightarrow \mathbb{R}$, $(g_p \circ f)(x) = \sqrt[p]{x^2 + 1}$ auf $\mathbb{R}$ stetig.

9.8. Stetigkeit von Potenzreihen: Es sei

$$\sum_{n=0}^{\infty} a_n (x - x_0)^n$$

eine reelle Potenzreihe mit Konvergenzradius $R \in (0, \infty]$. Dann ist die Funktion

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

auf $D := (x_0 - R, x_0 + R)$ stetig.

Einen Beweis dieses Satzes führen wir in Abschnitt 9.9 unten.

Wichtige Beispiele: Es gilt

$$\exp, \sin, \cos \in C(\mathbb{R}).$$

Obiger Satz hat die folgende Ergänzung (ohne Beweis):

Abelscher Grenzwertsatz: Ist in der Situation des Satzes $R < \infty$ und konvergiert die Potenzreihe für $x = x_0 + R$, so gilt

$$\lim_{x \rightarrow x_0 + R} f(x) = \sum_{n=0}^{\infty} a_n R^n.$$

Man kann also f stetig fortsetzen auf $(x_0 - R, x_0 + R]$. Entsprechendes gilt bei Konvergenz in $x = x_0 - R$.

9.9. Funktionenfolgen und Funktionenreihen: Es sei $\emptyset \neq D \subseteq \mathbb{R}$ und $(f_n)_{n \in \mathbb{N}}$ eine Folge von Funktionen $f_n : D \rightarrow \mathbb{R}$. Setzt man

$$s_N := \sum_{n=1}^N f_n \quad (N \in \mathbb{N}),$$

so heißt (s_N) eine *Funktionenreihe* und wird mit

$$\sum_{n=1}^{\infty} f_n$$

bezeichnet.

Definition: (1) Eine Funktionenfolge (f_n) heißt *punktweise konvergent auf D* , falls

$$\forall x \in D : (f_n(x)) \text{ ist konvergent.}$$

In diesem Fall heißt $f : D \rightarrow \mathbb{R}$, $f(x) := \lim_{n \rightarrow \infty} f_n(x)$ die *Grenzfunktion* der Folge (f_n) .

(2) Eine Funktionenreihe $\sum_{n=1}^{\infty} f_n$ heißt *punktweise konvergent auf D* , falls

$$\forall x \in D : (s_N(x)) \text{ ist konvergent.}$$

In diesem Fall heißt $f : D \rightarrow \mathbb{R}$, $f(x) := \sum_{n=1}^{\infty} f_n(x)$ die *Summenfunktion* der Reihe $\sum_{n=1}^{\infty} f_n$.

Beispiele: (1) $f_n : [0, 1] \rightarrow \mathbb{R}$, $f_n(x) = x^n$ ($n \in \mathbb{N}$). Hier gilt:

$$f_n(1) = 1 \rightarrow 1 \quad (n \rightarrow \infty), \quad \forall x \in [0, 1) : f_n(x) = x^n \rightarrow 0 \quad (n \rightarrow \infty).$$

Die Funktionenfolge (f_n) konvergiert also punktweise auf $[0, 1]$ gegen die Grenzfunktion

$$f : [0, 1] \rightarrow \mathbb{R}, \quad f(x) = \begin{cases} 0, & x \in [0, 1) \\ 1, & x = 1 \end{cases}.$$

(2) Sei

$$\sum_{n=0}^{\infty} a_n(x - x_0)^n$$

eine reelle Potenzreihe mit Konvergenzradius $R \in (0, \infty]$ und $D := (x_0 - R, x_0 + R)$. Für $n \in \mathbb{N}_0$ sei $f_n : D \rightarrow \mathbb{R}$ definiert durch $f_n(x) = a_n(x - x_0)^n$. Nach 7.12 konvergiert die Funktionenreihe $\sum_{n=0}^{\infty} f_n$ auf D punktweise gegen die Summenfunktion

$$f : D \rightarrow \mathbb{R}, \quad f(x) = \sum_{n=0}^{\infty} f_n(x) = \sum_{n=0}^{\infty} a_n(x - x_0)^n.$$

Beachte: Punktweise Konvergenz von (f_n) auf D gegen f bedeutet:

$$\forall x \in D \quad \forall \varepsilon > 0 \quad \exists n_0 = n_0(x, \varepsilon) \in \mathbb{N} \quad \forall n \geq n_0 : |f_n(x) - f(x)| < \varepsilon.$$

Definition:

Die Funktionenfolge [Funktionenreihe] (f_n) $[\sum_{n=1}^{\infty} f_n]$ konvergiert auf D *gleichmäßig* gegen $f : D \rightarrow \mathbb{R} \Leftrightarrow$

$$\forall \varepsilon > 0 \quad \exists n_0 = n_0(\varepsilon) \in \mathbb{N} \quad \forall n \geq n_0 \quad \forall x \in D : |f_n(x) - f(x)| < \varepsilon$$

$$[\forall \varepsilon > 0 \exists N_0 = N_0(\varepsilon) \in \mathbb{N} \forall N \geq N_0 \forall x \in D : |s_N(x) - f(x)| < \varepsilon]$$

Bemerkung: Aus gleichmäßiger Konvergenz folgt punktweise Konvergenz aber nicht umgekehrt.

Beispiele: (1) Die Funktionenfolge (f_n) mit $f_n(x) = x^n$ ($x \in [0, 1]$) konvergiert auf $[0, 1]$ nicht gleichmäßig gegen ihre Grenzfunktion f : Andernfalls existiert zu $\varepsilon = 1/3$ ein $n_0 \in \mathbb{N}$ mit

$$\forall n \geq n_0 \forall x \in [0, 1] : |x^n - f(x)| < \frac{1}{3}.$$

Sei $n \geq n_0$ und $x = 1/\sqrt[n]{2} \in [0, 1)$. Dann gilt $f(x) = 0$, also

$$|x^n - f(x)| = \frac{1}{2} < \frac{1}{3},$$

ein Widerspruch.

(2) Sei $D = (-1, 1)$. Die Funktionenreihe $\sum_{n=0}^{\infty} x^n$ konvergiert auf D punktweise gegen $f : D \rightarrow \mathbb{R}$, $f(x) = 1/(1-x)$, aber nicht gleichmäßig: Hier ist

$$s_N(x) = \sum_{n=0}^N x^n = \frac{1 - x^{N+1}}{1 - x}.$$

Annahme: Es liegt gleichmäßige Konvergenz vor. Dann existiert zu $\varepsilon = 1$ ein $N_0 \in \mathbb{N}$ mit

$$\forall N \geq N_0 \forall x \in (-1, 1) : |s_N(x) - f(x)| = \frac{|x|^{N+1}}{1 - x} < 1.$$

Sei $N \geq N_0$ fest. Es gilt

$$\frac{|x|^{N+1}}{1 - x} \rightarrow \infty \quad (x \rightarrow 1-),$$

ein Widerspruch.

Wie zeigt man gleichmäßige Konvergenz?

(a) Ist (α_n) eine Nullfolge und gilt

$$\exists m \in \mathbb{N} \forall n \geq m \forall x \in D : |f_n(x) - f(x)| \leq \alpha_n,$$

so ist (f_n) auf D gleichmäßig konvergent gegen f .

(b) (Kriterium von Weierstraß): Ist (c_n) eine Folge in $[0, \infty)$ mit $\sum_{n=1}^{\infty} c_n < \infty$ und gilt

$$\exists m \in \mathbb{N} \forall n \geq m \forall x \in D : |f_n(x)| \leq c_n,$$

so gilt:

(i) $\sum_{n=1}^{\infty} f_n(x)$ ist für jedes $x \in D$ absolut konvergent, und

(ii) $\sum_{n=1}^{\infty} f_n$ konvergiert auf D gleichmäßig gegen $f : D \rightarrow \mathbb{R}$, $f(x) := \sum_{n=1}^{\infty} f_n(x)$.

Beweis. (a) Sei $\varepsilon > 0$. Dann existiert ein $n_0 \geq m$ mit $\alpha_n < \varepsilon$ ($n \geq n_0$), also

$$\forall n \geq n_0 \forall x \in D : |f_n(x) - f(x)| \leq \alpha_n < \varepsilon.$$

(b) (i) folgt aus 7.4. (ii) Für $s_N = f_1 + \cdots + f_N$ gilt für alle $N \geq m$ und $x \in D$:

$$|s_N(x) - f(x)| = \left| \sum_{n=N+1}^{\infty} f_n(x) \right| \leq \sum_{n=N+1}^{\infty} c_n =: r_N.$$

Nach 7.2(2) gilt $r_N \rightarrow 0$ und die Behauptung folgt aus (a). $\square$

Satz: Es sei $\sum_{n=0}^{\infty} a_n(x - x_0)^n$ eine reelle Potenzreihe mit Konvergenzradius $R \in (0, \infty]$, $D := (x_0 - R, x_0 + R)$ und

$$f(x) := \sum_{n=0}^{\infty} a_n(x - x_0)^n \quad (x \in D).$$

Weiter sei $0 \leq r < R$. Dann konvergiert die Potenzreihe auf $[x_0 - r, x_0 + r]$ gleichmäßig gegen f .

Beweis. Für $n \in \mathbb{N}_0$ sei $f_n(x) := a_n(x - x_0)^n$ ($x \in [x_0 - r, x_0 + r]$). Es gilt

$$|f_n(x)| \leq |a_n|r^n =: c_n \quad (n \in \mathbb{N}_0, x \in [x_0 - r, x_0 + r]).$$

Wegen $0 \leq r < R$ ist $\sum_{n=0}^{\infty} c_n$ konvergent und die Behauptung folgt aus dem Kriterium von Weierstraß. $\square$

Wir zeigen nun, dass Stetigkeit bei gleichmäßiger Konvergenz erhalten bleibt.

Satz: Es sei $\emptyset \neq D \subseteq \mathbb{R}$ und $(f_n)_{n \in \mathbb{N}}$ eine Folge in $C(D)$, die auf D gleichmäßig gegen eine Funktion $f : D \rightarrow \mathbb{R}$ konvergiere. Dann ist $f \in C(D)$.

Beweis. Sei $x_0 \in D$ und $\varepsilon > 0$. Für jedes $x \in D$ und $n \in \mathbb{N}$ ist dann

$$|f(x) - f(x_0)| \leq |f(x) - f_n(x)| + |f_n(x) - f_n(x_0)| + |f_n(x_0) - f(x_0)|.$$

Zunächst finden wir $m \in \mathbb{N}$ so, dass für alle $x \in D$ gilt: $|f_m(x) - f(x)| < \varepsilon/3$. Da f_m in x_0 stetig ist, finden wir ein $\delta > 0$ so, dass für alle $x \in D$ mit $|x - x_0| < \delta$ gilt: $|f_m(x) - f_m(x_0)| < \varepsilon/3$. Für jedes $x \in D$ mit $|x - x_0| < \delta$ folgt dann:

$$|f(x) - f(x_0)| \leq |f(x) - f_m(x)| + |f_m(x) - f_m(x_0)| + |f_m(x_0) - f(x_0)| < \varepsilon,$$

womit Stetigkeit von f in x_0 gezeigt ist. $\square$

Beispiele: (1) Die Funktionenreihe $\sum_{n=1}^{\infty} \frac{\cos(nx)}{n^2}$ konvergiert gleichmäßig auf $\mathbb{R}$ (wegen (b) mit $c_n = \frac{1}{n^2}$) und die Summenfunktion

$$f(x) = \sum_{n=1}^{\infty} \frac{\cos(nx)}{n^2} \quad (x \in \mathbb{R})$$

ist stetig (da alle Funktionen $x \mapsto \cos(nx)/n^2$ stetig sind).

(2) Bekannt: Die durch $f_n(x) := x^n$ ($x \in [0, 1]$) definierte Funktionenfolge konvergiert auf $[0, 1]$ punktweise aber nicht gleichmäßig gegen ihre Grenzfunktion f . Zum Nachweis, dass (f_n) nicht gleichmäßig konvergiert kann am auch obigen Satz benutzen: Es gilt $f_n \in C([0, 1])$ ($n \in \mathbb{N}$) aber $f \notin C([0, 1])$.

Aus obigem Satz folgt nun auch 9.8 (die Stetigkeit von Potenzreihen): Es sei

$$\sum_{n=0}^{\infty} a_n(x - x_0)^n$$

eine reelle Potenzreihe mit Konvergenzradius $R \in (0, \infty]$, $D := (x_0 - R, x_0 + R)$ und

$$f : D \rightarrow \mathbb{R}, \quad f(x) := \sum_{n=0}^{\infty} a_n(x - x_0)^n.$$

Sei $x \in D$. Wähle ein $r \in (0, R)$ mit $x \in (x_0 - r, x_0 + r)$. Da die Potenzreihe auf $[x_0 - r, x_0 + r]$ gleichmäßig konvergiert und alle Funktionen $x \mapsto a_n(x - x_0)^n$ auf $[x_0 - r, x_0 + r]$ stetig sind, ist auch f stetig auf $[x_0 - r, x_0 + r]$, also stetig in x . Da $x \in D$ beliebig war folgt $f \in C(D)$.

9.10. Zwischenwertsatz (ZWS): Seien $a, b \in \mathbb{R}$ mit $a < b$ und $f : [a, b] \rightarrow \mathbb{R}$ stetig. Es sei y_0 zwischen $f(a)$ und $f(b)$, d.h.

$$\min\{f(a), f(b)\} \leq y_0 \leq \max\{f(a), f(b)\}.$$

Dann existiert ein $x_0 \in [a, b]$ mit $f(x_0) = y_0$.

Beweis. O.B.d.A. sei $f(a) \leq f(b)$, also $y_0 \in [f(a), f(b)]$ (andernfalls betrachte $-f$ und $-y_0$). Wir konstruieren rekursiv eine Intervallschachtelung $(I_n)_{n=0}^{\infty}$ mit $I_n = [a_n, b_n]$ so, dass $b_n - a_n = (b - a)/2^n$ und $y_0 \in [f(a_n), f(b_n)]$. Sei $I_0 := [a, b]$, also $a_0 := a$, $b_0 := b$. Ist $I_n = [a_n, b_n]$ definiert, so sei $c_n = \frac{a_n + b_n}{2}$ und

$$I_{n+1} := \begin{cases} [a_n, c_n], & \text{falls } f(c_n) \geq y_0 \\ [c_n, b_n], & \text{falls } f(c_n) < y_0 \end{cases}.$$

Setze

$$x_0 = \lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} b_n.$$

Da f stetig ist, folgt

$$f(x_0) = \lim_{n \rightarrow \infty} f(a_n) = \lim_{n \rightarrow \infty} f(b_n)$$

und wegen $y_0 \in [f(a_n), f(b_n)]$ ($n \in \mathbb{N}$) dann auch

$$y_0 = \lim_{n \rightarrow \infty} f(a_n) = \lim_{n \rightarrow \infty} f(b_n).$$

□

Folgerung: Ist I ein Intervall und $f : I \rightarrow \mathbb{R}$ stetig, so ist $f(I)$ ein Intervall.

Bemerkung: Eine Teilmenge $\emptyset \neq J \subseteq \mathbb{R}$ 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 $J := \exp(\mathbb{R}) = (0, \infty)$. Nach 7.10(4) ist $\exp(x) > 0$ ($x \in \mathbb{R}$), also gilt $J \subseteq (0, \infty)$. Nach der Folgerung ist J ein Intervall und nach 7.10(6) ist $\sup J = \infty$, $\inf J = 0$. Folglich ist $(0, \infty) \subseteq J$.

(2) Ist $p \in \mathbb{R}[X]$ ein Polynom von ungeradem Grad m , so gilt $p(\mathbb{R}) = \mathbb{R}$.

Beweis. O.B.d.A. sei p normiert. Wegen des ZWSes reicht es zu zeigen, dass $p(x) \rightarrow \infty$ ($x \rightarrow \infty$) und dass $p(x) \rightarrow -\infty$ ($x \rightarrow -\infty$). Dies ist klar für $m = 1$. Sei also $m \geq 3$ und $p(x) = x^m + a_{m-1}x^{m-1} + \dots + a_1x + a_0$. Für $x \geq 1 + |a_{m-1}| + |a_{m-2}| + \dots + |a_1| + |a_0|$ gilt dann

$$\begin{aligned} p(x) &\geq x^m - |a_{m-1}|x^{m-1} - |a_{m-2}|x^{m-2} - \dots - |a_1|x - |a_0| \\ &\geq x^m - |a_{m-1}|x^{m-1} - |a_{m-2}|x^{m-1} - \dots - |a_1|x^{m-1} - |a_0|x^{m-1} \\ &\geq x^{m-1} \underbrace{(x - (|a_{m-1}| + |a_{m-2}| + \dots + |a_1| + |a_0|))}_{\geq 1} \\ &\geq x^{m-1} \rightarrow \infty \quad (x \rightarrow \infty). \end{aligned}$$

Die Aussage $p(x) \rightarrow -\infty$ ($x \rightarrow -\infty$) zeigt man ähnlich. □

9.11. Monotone Funktionen und Stetigkeit: Es sei $\emptyset \neq D \subseteq \mathbb{R}$ und $f : D \rightarrow \mathbb{R}$ eine Funktion. f heißt (*monoton*) *wachsend* [bzw. (*monoton*) *fallend*], falls für alle $x_1, x_2 \in D$ gilt:

$$x_1 \leq x_2 \implies f(x_1) \leq f(x_2) \quad [\text{bzw. } f(x_1) \geq f(x_2)].$$

f heißt *streng (monoton) wachsend* [bzw. *streng (monoton) fallend*], falls für alle $x_1, x_2 \in D$ gilt:

$$x_1 < x_2 \implies f(x_1) < f(x_2) \quad [\text{bzw. } f(x_1) > f(x_2)].$$

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: (1) Nach 7.10(5) ist $\exp : \mathbb{R} \rightarrow \mathbb{R}$ streng monoton wachsend.

(2) Die Funktion $\mathbb{R} \rightarrow \mathbb{R}, x \mapsto x^2$ ist nicht monoton.

(3) Die Funktion

$$f : \mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto \begin{cases} 1, & x \geq 0 \\ 0, & x < 0 \end{cases}$$

ist monoton wachsend, aber nicht streng monoton wachsend.

Satz: Sei $I \subseteq \mathbb{R}$ ein Intervall und $f : I \rightarrow \mathbb{R}$ streng monoton wachsend [bzw. fallend].

(a) Dann ist f injektiv, also $f : I \rightarrow f(I)$ bijektiv, und die Umkehrfunktion $f^{-1} : f(I) \rightarrow \mathbb{R}$ ist ebenfalls streng monoton wachsend [bzw. fallend].

(b) Ist f zusätzlich stetig, so ist $f(I)$ ein Intervall und $f^{-1} : f(I) \rightarrow \mathbb{R}$ ist stetig.

Beweis. (a) Injektivität von f ist klar. Ist f z.B. streng monoton wachsend, so erhalten wir (durch Kontraposition):

$$\forall y_1, y_2 \in f(I) : y_1 \leq y_2 \implies f^{-1}(y_1) \leq f^{-1}(y_2).$$

Da f^{-1} injektiv ist, ist somit f^{-1} streng monoton wachsend.

(b) Wir verwenden 9.6 zum Nachweis der Stetigkeit von f^{-1} : Sei $y_0 = f(x_0) \in f(I)$ und $(y_n) = (f(x_n))$ eine monotone Folge in $f(I)$ mit $y_n \rightarrow y_0$. 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 $y_n = f(x_n) \rightarrow f(\alpha)$. Somit ist $f(\alpha) = y_0 = f(x_0)$. Es folgt $\alpha = x_0$, und damit

$$f^{-1}(y_n) = x_n \rightarrow x_0 = f^{-1}(y_0).$$

□

Beispiel: Ist $D \subseteq \mathbb{R}$ **kein** Intervall und $f \in C(D)$ streng monoton, so kann es vorkommen, dass $f^{-1} : f(D) \rightarrow \mathbb{R}$ nicht auf $f(D)$ stetig ist: Betrachte $f : [0, 1] \cup (2, 3] \rightarrow \mathbb{R}, f(x) = x$ ($x \in [0, 1]$), $f(x) = x - 1$ ($x \in (2, 3]$). Übung: $f^{-1} : [0, 2] \rightarrow \mathbb{R}$ ist in 1 unstetig.

9.12. Der Logarithmus: Nach 9.10 gilt $\exp(\mathbb{R}) = (0, \infty)$. Die Abbildung $\exp : \mathbb{R} \rightarrow (0, \infty)$ ist bijektiv und stetig.

Definition: Die Umkehrfunktion der Exponentialfunktion $\exp : \mathbb{R} \rightarrow (0, \infty)$ heißt (*natürlicher*) *Logarithmus* $\log := \ln : (0, \infty) \rightarrow \mathbb{R}$, d.h. also $\log x := \exp^{-1}(x)$ ($x \in (0, \infty)$).

Somit ist

$$\log(\exp(x)) = x \quad (x \in \mathbb{R}), \quad \exp(\log y) = y \quad (y \in (0, \infty)).$$

Eigenschaften des Logarithmus: $\log$ ist streng monoton wachsend und stetig. Es gilt $\log((0, \infty)) = \mathbb{R}$, $\log(1) = 0$ und $\log(e) = 1$, sowie

$$\log x \rightarrow \infty \quad (x \rightarrow \infty), \quad \log x \rightarrow -\infty \quad (x \rightarrow 0+).$$

Für alle $x, y > 0$ gilt

$$\log(xy) = \log x + \log y \quad \text{und} \quad \log(x/y) = \log x - \log y.$$

Es gilt nämlich

$$\begin{aligned} \exp(\log x + \log y) &= \exp(\log x) \exp(\log y) = xy = \exp(\log(xy)), \\ \exp(\log x - \log y) &= \exp(\log x) \exp(-\log y) = x \exp(\log y)^{-1} = x/y = \exp(\log(x/y)). \end{aligned}$$

9.13. Die allgemeine Potenz: Wir definieren für $a > 0$ die allgemeine Potenz:

$$\forall x \in \mathbb{R} : a^x := \exp(x \log a).$$

Weiter setzen wir $0^x := 0$ ($x > 0$) (Erinnerung: $0^0 := 1$).

Bemerkung: Nach 7.10 stimmt dies für $x \in \mathbb{Z}$ mit der bisherigen Definition (von ganzzahligen Potenzen) überein. Aus 7.10(9) folgt:

$$a^{1/n} = \exp\left(\frac{1}{n} \log(a)\right) = \sqrt[n]{\exp(\log(a))} = \sqrt[n]{a} \quad (n \in \mathbb{N}).$$

Für $a = e$ erhalten wir $e^x = \exp(x)$ wie in 7.10. Wir schreiben in Zukunft oft 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 9.7);
- (3) $a^{x+y} = e^{(x+y) \log a} = e^{x \log a} e^{y \log a} = a^x a^y$;
- (4) $a^{-x} = e^{-x \log a} = (e^{x \log a})^{-1} = (a^x)^{-1} = 1/a^x$;
- (5) $\log(a^x) = \log(e^{x \log a}) = x \log a$;
- (6) $(a^y)^x = e^{x \log(a^y)} \stackrel{(5)}{=} e^{xy \log a} = a^{xy}$, insbesondere $\sqrt[q]{a^p} = (\sqrt[q]{a})^p = a^{p/q}$ ($p \in \mathbb{Z}, q \in \mathbb{N}$).
- (7) $(ab)^x = e^{x \log(ab)} \stackrel{9.12}{=} e^{x \log a} e^{x \log b} = a^x b^x$.
- (8) $a^{b^x} := a^{(b^x)}$; i.a. ist $a^{b^x} \neq (a^b)^x$.

Der allgemeine Logarithmus: Sei $a > 1$. Dann ist die Funktion $\mathbb{R} \rightarrow (0, \infty)$, $x \mapsto a^x$, streng monoton wachsend, stetig und bijektiv.

Die Umkehrfunktion $\log_a : (0, \infty) \rightarrow \mathbb{R}$ heißt *Logarithmus zur Basis a* . Es ist also

$$\log_a(a^x) = x \quad (x \in \mathbb{R}), \quad a^{\log_a y} = y \quad (y \in (0, \infty)),$$

und es gilt

$$\log_a y = \frac{\log y}{\log a} \quad (y \in (0, \infty)),$$

denn

$$a^{\frac{\log y}{\log a}} = \exp\left(\frac{\log y}{\log a} \log a\right) = \exp(\log y) = y \quad (y \in (0, \infty)).$$

9.14. Abgeschlossene und kompakte Mengen: Es sei $D \subseteq \mathbb{R}$.

Definition: D heißt *abgeschlossen*, wenn für jede Folge (x_n) in D mit $x_n \rightarrow x_0 \in \mathbb{R}$ gilt: $x_0 \in D$.

Beispiele: (1) Folgende Mengen sind abgeschlossen ($a, b \in \mathbb{R}$, $a \leq b$):

$$\mathbb{R}, \emptyset, [a, b], (-\infty, a], [a, \infty), \{1/n : n \in \mathbb{N}\} \cup \{0\}.$$

(2) Folgende Mengen sind nicht abgeschlossen ($a, b \in \mathbb{R}$, $a < b$):

$$(a, b), (a, b], [a, b), (a, \infty), (-\infty, a), \{1/n : n \in \mathbb{N}\}.$$

Bemerkung: D ist abgeschlossen genau dann, wenn jeder Häufungspunkt von D zu D gehört.

Satz: Ist $\emptyset \neq D \subseteq \mathbb{R}$ abgeschlossen und nach oben [bzw. nach unten] beschränkt, so gilt $\sup D \in D$ [bzw. $\inf D \in D$].

Beweis. Sei $\gamma := \sup D$. Es ist $\gamma \in \mathbb{R}$. Nach 4.6 finden wir zu jedem $n \in \mathbb{N}$ ein $x_n \in D$ mit $\gamma - \frac{1}{n} < x_n \leq \gamma$. Dann gilt $x_n \rightarrow \gamma$, also $\gamma \in D$, da D abgeschlossen ist. $\square$

Definition: D heißt *kompakt*, wenn jede Folge (x_n) in D eine konvergente Teilfolge $(x_{k(n)})$ hat, mit

$$\lim_{n \rightarrow \infty} x_{k(n)} \in D.$$

Satz: Es gilt: D ist kompakt $\Leftrightarrow D$ ist abgeschlossen und beschränkt.

Beweis. “ $\Leftarrow$ ”: Da D beschränkt ist, hat eine Folge in D nach Bolzano-Weierstraß (siehe 6.8) eine Teilfolge, die in $\mathbb{R}$ konvergiert. Da D abgeschlossen ist, liegt deren Grenzwert in D .

“ $\Rightarrow$ ”: Sei (x_n) eine Folge in D , die gegen $x_0 \in \mathbb{R}$ konvergiert. Diese hat eine konvergente Teilfolge mit Grenzwert in D . Jede Teilfolge konvergiert gegen x_0 . Also gilt $x_0 \in D$. Damit ist D abgeschlossen. Wäre D nicht beschränkt, so könnten wir eine Folge (x_n) in D finden mit $|x_n| \rightarrow \infty$. Diese hätte keine konvergente Teilfolge. $\square$

9.15. Satz: Sei $\emptyset \neq D \subseteq \mathbb{R}$ kompakt, sowie $f : D \rightarrow \mathbb{R}$ eine stetige Funktion. Dann ist $f(D)$ kompakt, und es gibt $x_1, x_2 \in D$ mit

$$\forall x \in D : f(x_1) \leq f(x) \leq f(x_2)$$

(d.h. “ f nimmt auf D Maximum und Minimum an”).

Beweis. Sei $(y_n) = (f(x_n))$ eine Folge in $f(D)$. Dann ist (x_n) eine Folge in D , und wir finden ein $x_0 \in D$ und eine Teilfolge $(x_{k(n)})$ von (x_n) mit $x_{k(n)} \rightarrow x_0$. Da f stetig ist, folgt $y_{k(n)} = f(x_{k(n)}) \rightarrow f(x_0) \in f(D)$. Somit ist $f(D)$ kompakt. Die zweite Aussage folgt aus den Sätzen in 9.14. $\square$

Folgerung: Sind $a, b \in \mathbb{R}$ mit $a < b$ und ist $f : [a, b] \rightarrow \mathbb{R}$ stetig, so gibt es $c, d \in \mathbb{R}$ mit $c \leq d$ und $f([a, b]) = [c, d]$.

Beweis. Nach 9.10 ist $f([a, b])$ ein Intervall. Da $[a, b]$ kompakt ist existieren

$$c := \min f([a, b]), \quad d := \max f([a, b]).$$

Also gilt $f([a, b]) = [c, d]$. $\square$

10 Trigonometrische und Hyperbel-Funktionen

10.1. Die Zahl π : Wir beginnen mit einer **Vorbetrachtung**: Nach 7.11 gilt

$$\sin x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} \quad (x \in \mathbb{R}).$$

Für $x \in (0, 2)$ und $n \in \mathbb{N}$ ist

$$\frac{x^{2n-1}}{(2n-1)!} - \frac{x^{2n+1}}{(2n+1)!} = \frac{x^{2n-1}}{(2n-1)!} \left(1 - \frac{x^2}{2n(2n+1)}\right) \geq \frac{x^{2n-1}}{(2n-1)!} \left(1 - \frac{x^2}{6}\right) > 0,$$

also

$$\sin x = \underbrace{\left(x - \frac{x^3}{3!}\right)}_{>0} + \underbrace{\left(\frac{x^5}{5!} - \frac{x^7}{7!}\right)}_{>0} + \underbrace{\left(\frac{x^9}{9!} - \frac{x^{11}}{11!}\right)}_{>0} + \dots > x - \frac{x^3}{3!} > 0.$$

Speziell: $\sin 1 > 1 - \frac{1}{6} = \frac{5}{6}$. Weiter gilt $\cos 0 = 1 > 0$ und

$$\begin{aligned} \cos 2 &= \cos(1+1) = \cos^2 1 - \sin^2 1 = \cos^2 1 + \sin^2 1 - 2\sin^2 1 \\ &= 1 - 2\sin^2 1 \leq 1 - 2 \cdot \frac{25}{36} < 0. \end{aligned}$$

Nach dem Zwischenwertsatz 9.10 gibt es mindestens ein $x_0 \in (0, 2)$ mit $\cos x_0 = 0$.

Definition: Wir definieren π als das zweifache der kleinsten Nullstelle des Cosinus im Intervall $[0, 2]$, also

$$\pi := 2 \inf\{x_0 \in [0, 2] : \cos x_0 = 0\} = 2 \min\{x_0 \in [0, 2] : \cos x_0 = 0\}.$$

Damit gilt $\pi \in (0, 4)$ und $\cos(x) > 0$ ($x \in [0, \pi/2)$). Wegen $\cos(\pi/2) = 0$ ist

$$\sin^2 \frac{\pi}{2} = 1 - \cos^2 \frac{\pi}{2} = 1, \text{ also } |\sin \frac{\pi}{2}| = 1,$$

und wegen $\sin(x) > 0$ in $(0, 2)$ folgt

$$\sin \frac{\pi}{2} = 1.$$

Bemerkung: Tatsächlich gibt es nur eine Nullstelle des Cosinus im Intervall $[0, 2]$.

10.2. Eigenschaften von Sinus und Cosinus: Aus den Additionstheoremen in 7.11 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) $\forall x \in \mathbb{R}: \sin(x + 2\pi) = \sin x, \cos(x + 2\pi) = \cos x$ (d.h. $\sin$ 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: $\cos x = 0 \Leftrightarrow \exists k \in \mathbb{Z} : x = k\pi + \frac{\pi}{2}$.
 (9) Für $x \in \mathbb{R}$ gilt: $\sin x = 0 \Leftrightarrow \exists k \in \mathbb{Z} : x = k\pi$.
 (10) $\forall k \in \mathbb{Z} : \cos(k\pi) = (-1)^k$ und $\sin(k\pi + \frac{\pi}{2}) = (-1)^k$.

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 $[0, \frac{\pi}{2}]$ zu betrachten. Dabei ist $\sin x > 0$ für $x \in (0, \frac{\pi}{2}]$ und $\cos x > 0$ für $x \in [0, \frac{\pi}{2})$. Es folgt $\cos x = \sqrt{1 - \sin^2 x} < 1 = \cos 0$ für $x \in (0, \frac{\pi}{2}]$, und für $x, h \in (0, \frac{\pi}{2}]$ mit $x + h \in (0, \frac{\pi}{2}]$ gilt also

$$\cos(x + h) = \cos x \underbrace{\cos h}_{\leq 1} - \underbrace{\sin x \sin h}_{> 0} < \cos x,$$

d.h. $\cos$ ist auf $[0, \frac{\pi}{2}]$ streng monoton fallend.

Folglich ist $\sin x$ auf $[0, \frac{\pi}{2}]$ streng monoton wachsend, wg. $\sin x = \sqrt{1 - \cos^2 x}$ ($x \in [0, \frac{\pi}{2}]$).

10.4. Arcussinus und Arcuscosinus: Wegen 10.3 und 10.2(6) ist $\cos : [0, \pi] \rightarrow [-1, 1]$ streng monoton fallend, nach 9.8 ist diese Abbildung stetig und wegen 9.11 ist sie bijektiv. Die Umkehrabbildung $\arccos : [-1, 1] \rightarrow [0, \pi]$ heißt *Arcuscosinus*.

Eigenschaften: $\arccos : [-1, 1] \rightarrow [0, \pi]$ ist streng monoton fallend, stetig und bijektiv. Es gilt $\arccos(-1) = \pi, \arccos 0 = \frac{\pi}{2}, \arccos 1 = 0$.

Wegen 10.3 und 10.2(12) ist $\sin : [-\frac{\pi}{2}, \frac{\pi}{2}] \rightarrow [-1, 1]$ streng monoton wachsend. Auch diese Funktion ist stetig und bijektiv.

Die Umkehrfunktion $\arcsin : [-1, 1] \rightarrow [-\frac{\pi}{2}, \frac{\pi}{2}]$ heißt *Arcussinus*.

Eigenschaften: Die Funktion $\arcsin : [-1, 1] \rightarrow [-\frac{\pi}{2}, \frac{\pi}{2}]$ ist streng monoton wachsend, stetig und bijektiv. Es gilt $\arcsin 1 = \frac{\pi}{2}$, $\arcsin 0 = 0$ 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

$$\cos : [k\pi, (k+1)\pi] \rightarrow [-1, 1], \quad \sin : [k\pi - \frac{\pi}{2}, k\pi + \frac{\pi}{2}] \rightarrow [-1, 1]$$

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

$$\tan : \mathbb{R} \setminus (\frac{\pi}{2} + \pi\mathbb{Z}) \rightarrow \mathbb{R}, \quad x \mapsto \tan x := \frac{\sin x}{\cos x}$$

heißt *Tangens*. (Beachte, dass die Menge $\frac{\pi}{2} + \pi\mathbb{Z}$ genau die Nullstellen des Cosinus enthält.)

Eigenschaften: Es gilt $\tan 0 = 0$, $\tan \frac{\pi}{4} = 1$, sowie

$$\forall x \in \mathbb{R} \setminus (\frac{\pi}{2} + \pi\mathbb{Z}) : \tan(-x) = -\tan x, \quad \tan(x + \pi) = \tan x.$$

Somit ist der Tangens eine π -periodische Funktion.

Außerdem ist $\tan$ auf $(-\frac{\pi}{2}, \frac{\pi}{2})$ streng monoton wachsend mit $\tan x \rightarrow \infty$ für $x \rightarrow \frac{\pi}{2}-$ und $\tan x \rightarrow -\infty$ für $x \rightarrow -\frac{\pi}{2}+$.

Folglich ist $\tan : (-\frac{\pi}{2}, \frac{\pi}{2}) \rightarrow \mathbb{R}$ streng monoton wachsend, stetig und bijektiv. Die Umkehrfunktion $\arctan : \mathbb{R} \rightarrow (-\frac{\pi}{2}, \frac{\pi}{2})$ heißt *Arcustangens*.

Eigenschaften: Es gilt $\arctan 0 = 0$, $\arctan 1 = \frac{\pi}{4}$ und $\arctan(-x) = -\arctan x$ ($x \in \mathbb{R}$), sowie

$$\arctan x \rightarrow \frac{\pi}{2} \quad (x \rightarrow \infty), \quad \arctan x \rightarrow -\frac{\pi}{2} \quad (x \rightarrow -\infty).$$

Bemerkung: Sei $k \in \mathbb{Z}$. Man sieht leicht, dass die Umkehrabbildung von

$$\tan : (k\pi - \frac{\pi}{2}, k\pi + \frac{\pi}{2}) \rightarrow \mathbb{R}$$

gegeben ist durch $x \mapsto k\pi + \arctan x$.

10.6. Anwendung (Polarkoordinaten):

Vorbemerkung: Sei $t \in \mathbb{R}$. Dann gilt

$$e^{it} = 1 \Leftrightarrow \cos(t) + i\sin(t) = 1 \Leftrightarrow \cos(t) = 1 \wedge \sin(t) = 0 \Leftrightarrow \exists k \in \mathbb{Z} : t = 2k\pi.$$

Satz: Jede komplexe Zahl $z \in \mathbb{C} \setminus \{0\}$ besitzt eine Darstellung

$$z = re^{i\varphi} \text{ mit } r = |z| \text{ und } \varphi \in \mathbb{R},$$

wobei φ bis auf Addition von $2k\pi$ mit $k \in \mathbb{Z}$ bestimmt ist. Insbesondere gibt es genau einen Winkel $\varphi \in (-\pi, \pi]$ mit $z = re^{i\varphi}$. Dieser heißt das *Argument von z* . Wir schreiben $\arg z := \varphi$.

Beweis. 1. Fall: Es sei $\operatorname{Im} z \geq 0$. Wir setzen $x + iy := z/|z|$ mit $x, y \in \mathbb{R}$; dann ist $x^2 + y^2 = 1$ und $y \geq 0$. Wir setzen $\varphi := \arccos x$. Dann ist $\varphi \in [0, \pi]$, und es gilt $\cos \varphi = x$ und $\sin \varphi \geq 0$. Wegen $y^2 = 1 - (\cos \varphi)^2 = (\sin \varphi)^2$ und $y \geq 0$ folgt $\sin \varphi = y$. Also ist

$$x + iy = \cos \varphi + i \sin \varphi, \text{ somit } z = |z|e^{i\varphi}.$$

2. Fall: Es sei $\operatorname{Im} z < 0$. Nach dem 1. Fall hat $\bar{z}$ eine Darstellung $\bar{z} = |z|e^{-i\varphi}$, $-\varphi \in [0, \pi]$, also $z = |z|e^{i\varphi}$ mit $\varphi \in (-\pi, 0]$, denn $\varphi = -\pi$ ist nicht möglich wegen $\operatorname{Im} z < 0$.

Ist nun $z = |z|e^{i\psi}$ eine weitere Darstellung, so ist $e^{i(\varphi-\psi)} = 1$ und aus der Vorbemerkung folgt $\varphi - \psi = 2k\pi$ für ein $k \in \mathbb{Z}$. $\square$

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

$$zw = rse^{i(\varphi+\psi)},$$

d.h. 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 $\arg z = \arg w = \frac{3\pi}{4}$ und $\arg z + \arg w = \frac{3\pi}{2} \neq -\frac{\pi}{2} = \arg(-2i) = \arg(wz)$ 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 $w_1, w_2, \dots, w_n \in \mathbb{C} \setminus \{0\}$ mit $w_k^n = z$ ($k = 1, 2, \dots, n$), d.h. es gibt genau n verschiedene n -te Wurzeln von z . Schreibe dazu $z = re^{i\varphi}$ mit $r > 0$ und $\varphi \in (-\pi, \pi]$ und setze

$$w_{k+1} := \sqrt[n]{r} e^{i\frac{\varphi+2k\pi}{n}} \quad (k = 0, 1, \dots, n-1).$$

Diese Zahlen sind paarweise verschieden: Es seien $j, k \in \{0, \dots, n-1\}$, $k \geq j$. Ist

$$\sqrt[n]{r} e^{i\frac{\varphi+2k\pi}{n}} = \sqrt[n]{r} e^{i\frac{\varphi+2j\pi}{n}}, \text{ also } e^{i\frac{2(k-j)\pi}{n}} = 1,$$

so existiert ein $l \in \mathbb{Z}$ mit

$$\frac{2(k-j)\pi}{n} = 2l\pi, \text{ also } \frac{k-j}{n} = l.$$

Somit ist

$$0 \leq l = \frac{k-j}{n} \leq \frac{k}{n} \leq \frac{n-1}{n} = 1 - \frac{1}{n} < 1.$$

Wegen $l \in \mathbb{Z}$ folgt damit $l = 0$, also $k = j$.

Weiter gilt $w_k^n = re^{i\varphi} = z$ und nach dem Fundamentalsatz der Algebra 5.5 hat das Polynom $p(w) = w^n - z$ höchstens n verschiedene Nullstellen. Also gilt $w^n - z = \prod_{k=1}^n (w - w_k)$.

Beispiele: (1) $n = 2$, $z \in \mathbb{C} \setminus \{0\}$: Die beiden Quadratwurzeln von z sind

$$w_1 = \sqrt{|z|}e^{i\arg(z)/2}, \quad w_2 = \sqrt{|z|}e^{i(\arg(z)+2\pi)/2} = \sqrt{|z|}e^{i\arg(z)/2} \underbrace{e^{i\pi}}_{=-1} = -w_1.$$

Eine weitere Möglichkeit Quadratwurzeln zu bestimmen ist folgende: Ist $z \in (-\infty, 0]$, so sind $w_{1,2} = \pm i\sqrt{-z}$ die Wurzeln von z .

Behauptung: Ist $z \in \mathbb{C} \setminus (-\infty, 0]$, so sind

$$w_{1,2} = \pm \sqrt{|z|} \frac{z + |z|}{|z + |z||}$$

die Wurzeln von z .

Beweis. Es gilt:

$$\begin{aligned} \left(\pm \sqrt{|z|} \frac{z + |z|}{|z + |z||} \right)^2 &= |z| \frac{(z + |z|)(z + |z|)}{(z + |z|)(\bar{z} + |z|)} = |z| \frac{z + |z|}{\bar{z} + |z|} \\ &= \frac{|z|z + z\bar{z}}{\bar{z} + |z|} = z \frac{|z| + \bar{z}}{\bar{z} + |z|} = z. \end{aligned}$$

□

Z.B. sind die Wurzeln von $-3 + 4i$:

$$\pm \sqrt{5} \frac{-3 + 4i + 5}{|-3 + 4i + 5|} = \pm \sqrt{5} \frac{2 + 4i}{\sqrt{20}} = \pm(1 + 2i).$$

(2) $n = 3$, $z = 1$: Die drei dritten Wurzeln von 1 sind

$$w_1 = 1, \quad w_2 = e^{i\frac{2\pi}{3}}, \quad w_3 = e^{i\frac{4\pi}{3}} = e^{-i\frac{2\pi}{3}} = \bar{w}_2.$$

(3) $n = 4$, $z = 1$: Die vier vierten Wurzeln von 1 sind

$$w_1 = 1, \quad w_2 = e^{i\frac{\pi}{2}} = i, \quad w_3 = e^{i\pi} = -1, \quad w_4 = e^{i\frac{3\pi}{2}} = -i.$$

(4) $n \in \mathbb{N}$, $z = 1$: Dann gilt $w^n - 1 = \prod_{k=1}^n (w - w_k)$, also (mit $w = 0$):

$$\prod_{k=1}^n w_k = (-1)^{n+1}.$$

Die n -ten Wurzeln von 1 heißen auch die n -ten *Einheitswurzeln*.

10.7. Hyperbelfunktionen: Die Funktionen

$$\cosh : \mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto \cosh x := \frac{1}{2}(e^x + e^{-x}),$$

$$\sinh : \mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto \sinh x := \frac{1}{2}(e^x - e^{-x})$$

heißen *Cosinus hyperbolicus* bzw. *Sinus hyperbolicus*.

Es 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

$$\begin{aligned} \cosh x &\rightarrow \infty, \quad \sinh x \rightarrow \infty \quad \text{für } x \rightarrow \infty \\ \cosh x &\rightarrow \infty, \quad \sinh x \rightarrow -\infty \quad \text{für } x \rightarrow -\infty, \end{aligned}$$

sowie

$$\forall x \in \mathbb{R} : \cosh^2 x - \sinh^2 x = 1.$$

Reihendarstellungen: Für alle $x \in \mathbb{R}$ gilt

$$\cosh x = \sum_{n=0}^{\infty} \frac{x^{2n}}{(2n)!}, \quad \sinh x = \sum_{n=0}^{\infty} \frac{x^{2n+1}}{(2n+1)!}.$$

Folgerung: Die Funktionen

$$\cosh : [0, \infty) \rightarrow [1, \infty) \quad \text{und} \quad \sinh : \mathbb{R} \rightarrow \mathbb{R}$$

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

$$\tanh : \mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto \tanh x := \frac{\sinh x}{\cosh x} = \frac{e^x - e^{-x}}{e^x + e^{-x}},$$

heißt *Tangens hyperbolicus*.

Es ist $\tanh(-x) = -\tanh x$ ($x \in \mathbb{R}$) und $\tanh 0 = 0$. Außerdem gilt $\tanh x \rightarrow 1$ ($x \rightarrow \infty$) und $\tanh x \rightarrow -1$ ($x \rightarrow -\infty$).

Die Funktion $\tanh : \mathbb{R} \rightarrow (-1, 1)$ ist streng monoton wachsend, stetig und bijektiv.

10.8. Areafunktionen: Die Umkehrfunktionen von $\sinh : \mathbb{R} \rightarrow \mathbb{R}$, $\cosh : [0, \infty) \rightarrow [1, \infty)$ und $\tanh : \mathbb{R} \rightarrow (-1, 1)$ heißen $\operatorname{Arsinh} : \mathbb{R} \rightarrow \mathbb{R}$ (*Areasinus*), $\operatorname{Arcosh} : [1, \infty) \rightarrow [0, \infty)$ (*Areacosinus*) und $\operatorname{Artanh} : (-1, 1) \rightarrow \mathbb{R}$ (*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 $f : I \rightarrow \mathbb{R}$ 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* (db), falls der Limes

$$\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} \in \mathbb{R}$$

existiert. Dieser Grenzwert heißt *die Ableitung von f in x_0* . Bezeichnung: $f'(x_0)$, $\frac{d}{dx}f(x_0)$.

Die Funktion f heißt *auf I differenzierbar*, falls f in jedem $x \in I$ differenzierbar ist. In diesem Fall heißt $f' : I \rightarrow \mathbb{R}$, $x \mapsto f'(x)$, die *Ableitung von f auf I* .

Bemerkung: Falls f in x_0 differenzierbar ist gilt

$$\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h}.$$

Mit 9.4 sieht man: $\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} = a$ ist äquivalent zu

$$\forall \varepsilon > 0 \exists \delta > 0 \forall x \in I : |x - x_0| < \delta \Rightarrow |f(x) - f(x_0) - a(x - x_0)| \leq \varepsilon |x - x_0|.$$

Beispiele: (1) $f : I \rightarrow \mathbb{R}$, $f(x) = c \in \mathbb{R}$, ist auf I differenzierbar mit $f'(x) = 0$ ($x \in I$).

(2) $f : \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = |x|$ 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 $\lim_{h \rightarrow 0} \frac{f(h) - f(0)}{h}$ nicht.

(3) $I = \mathbb{R}$, $n \in \mathbb{N}$, $f(x) = x^n$. f ist auf $\mathbb{R}$ differenzierbar mit $f'(x_0) = nx_0^{n-1}$ für jedes $x_0 \in \mathbb{R}$: Für $x \neq x_0$ ist nach 4.11(1)

$$\frac{x^n - x_0^n}{x - x_0} = \sum_{k=0}^{n-1} x^{n-1-k} x_0^k \rightarrow nx_0^{n-1} \quad (x \rightarrow x_0).$$

(4) $\exp, \sin, \cos, \sinh, \cosh$ sind auf $\mathbb{R}$ differenzierbar mit

$$\exp' = \exp, \quad \sin' = \cos, \quad \cos' = -\sin, \quad \sinh' = \cosh, \quad \cosh' = \sinh.$$

Für $x, h \in \mathbb{R}$ mit $h \neq 0$ gilt nämlich nach 7.10(2) und 9.5:

$$\frac{e^{x+h} - e^x}{h} = e^x \underbrace{\frac{e^h - 1}{h}}_{\rightarrow 1} \rightarrow e^x \quad (h \rightarrow 0),$$

sowie nach 7.11(4) (Additionstheoreme) und 9.5:

$$\begin{aligned} \frac{\sin(x+h) - \sin x}{h} &= \frac{\sin x \cos h + \cos x \sin h - \sin x}{h} \\ &= \sin x \underbrace{\frac{\cos h - 1}{h}}_{\rightarrow 0} + \cos x \underbrace{\frac{\sin h}{h}}_{\rightarrow 1} \rightarrow \cos x \quad (h \rightarrow 0). \end{aligned}$$

Die Beweise für $\cos, \sinh$ und $\cosh$ sind analog. Später werden wir sehen, dass Potenzreihen stets differenzierbar sind.

Satz: Ist $f : I \rightarrow \mathbb{R}$ in $x_0 \in I$ differenzierbar, so ist f in x_0 stetig.

Beweis. Da I ein Intervall ist, ist $x_0 \in I$ ein Häufungspunkt von I . Für $x \in I \setminus \{x_0\}$ gilt dann

$$f(x) - f(x_0) = \frac{f(x) - f(x_0)}{x - x_0} (x - x_0) \rightarrow f'(x_0) \cdot 0 = 0 \quad (x \rightarrow x_0),$$

d.h. $\lim_{x \rightarrow x_0} f(x) = f(x_0)$. Also ist f in x_0 stetig. $\square$

11.2. Ableitungsregeln: Seien $f, g : I \rightarrow \mathbb{R}$ 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 und

$$(\alpha f + \beta g)'(x_0) = \alpha f'(x_0) + \beta g'(x_0).$$

(2) fg ist differenzierbar in x_0 und

$$(fg)'(x_0) = f'(x_0)g(x_0) + f(x_0)g'(x_0) \quad \text{Produktregel.}$$

(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 $\frac{f}{g} : J \rightarrow \mathbb{R}$ ist in x_0 differenzierbar und

$$\left(\frac{f}{g}\right)'(x_0) = \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g(x_0)^2} \quad \text{Quotientenregel.}$$

Bemerkung: (1) besagt, dass

$$V := \{f : I \rightarrow \mathbb{R} : f \text{ ist differenzierbar in } x_0\}$$

ein $\mathbb{R}$ -Vektorraum ist und dass die Abbildung $V \rightarrow \mathbb{R}, f \mapsto f'(x_0)$ linear ist.

Beweis. (1) ist klar wegen 6.3 bzw. 9.5.

(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 $x \rightarrow x_0$ die Behauptung folgt (beachte, dass $g(x) \rightarrow g(x_0)$ für $x \rightarrow x_0$, 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

$$\frac{1/g(x) - 1/g(x_0)}{x - x_0} = -\frac{\frac{g(x)-g(x_0)}{x-x_0}}{g(x)g(x_0)} \rightarrow -\frac{g'(x_0)}{g(x_0)^2} \quad (x \rightarrow x_0).$$

Wieder beachte man $g(x) \rightarrow g(x_0)$ ($x \rightarrow x_0$). □

Beispiele: $\tan = \frac{\sin}{\cos}$ und $\tanh = \frac{\sinh}{\cosh}$ sind auf $(-\frac{\pi}{2}, \frac{\pi}{2})$ bzw. auf $\mathbb{R}$ differenzierbar, und es gilt:

$$\tan' = \frac{1}{\cos^2} = 1 + \tan^2 \quad \text{auf } (-\frac{\pi}{2}, \frac{\pi}{2}), \quad \tanh' = \frac{1}{\cosh^2} = 1 - \tanh^2 \quad \text{auf } \mathbb{R}.$$

Es ist nämlich

$$\tan' = \left(\frac{\sin}{\cos}\right)' = \frac{\sin' \cos - \sin \cos'}{\cos^2} = \frac{\cos^2 + \sin^2}{\cos^2} = \frac{1}{\cos^2} = 1 + \tan^2.$$

Der Beweis für $\tanh$ ist analog.

11.3. Kettenregel: Sei $f : I \rightarrow \mathbb{R}$ in $x_0 \in I$ differenzierbar und $J \subseteq \mathbb{R}$ ein Intervall mit $f(I) \subseteq J$. Sei $g : J \rightarrow \mathbb{R}$ in $y_0 := f(x_0)$ differenzierbar. Dann ist $g \circ f : I \rightarrow \mathbb{R}$ in x_0 differenzierbar und

$$(g \circ f)'(x_0) = g'(f(x_0)) \cdot f'(x_0)$$

(“äußere Ableitung mal innere Ableitung”).

Beweis. Wir setzen

$$q : J \rightarrow \mathbb{R}, \quad q(y) := \begin{cases} \frac{g(y)-g(y_0)}{y-y_0}, & y \neq y_0 \\ g'(y_0), & y = y_0 \end{cases}.$$

Dann gilt $q(y) \rightarrow g'(y_0)$ ($y \rightarrow y_0$), also auch $q(f(x)) \rightarrow g'(f(x_0))$ ($x \rightarrow x_0$), 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

$$\frac{g(f(x)) - g(f(x_0))}{x - x_0} = q(f(x)) \cdot \frac{f(x) - f(x_0)}{x - x_0} \rightarrow g'(f(x_0)) \cdot f'(x_0) \quad (x \rightarrow x_0).$$

□

Beispiel: Sei $a > 0$ und $h(x) = a^x = e^{x \log a}$ ($x \in \mathbb{R}$). Setze $g(y) = e^y$ und $f(x) = x \log a$. Also ist $h = g \circ f$ auf $\mathbb{R}$ differenzierbar und es gilt

$$h'(x) = (a^x)' = g'(x \log a) f'(x) = a^x \log a \quad (x \in \mathbb{R}).$$

Beachte die unpräzise aber übliche Notation “ $(a^x)'$ ”.

11.4. Ableitung der Umkehrfunktion: Sei $f : I \rightarrow \mathbb{R}$ stetig und streng monoton auf I . Ist f in $x_0 \in I$ differenzierbar und $f'(x_0) \neq 0$, so ist die Umkehrfunktion $f^{-1} : f(I) \rightarrow \mathbb{R}$ 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 9.10 ist $f(I)$ ein Intervall. Sei $y \in f(I) \setminus \{y_0\}$ und $x := f^{-1}(y)$. Dann gilt

$$\frac{f^{-1}(y) - f^{-1}(y_0)}{y - y_0} = \frac{x - x_0}{f(x) - f(x_0)} \rightarrow \frac{1}{f'(x_0)} \quad (y \rightarrow y_0),$$

da wegen der Stetigkeit von f^{-1} aus $y \rightarrow y_0$ folgt $x \rightarrow x_0$. □

Beispiele: Die Funktionen $\log : (0, \infty) \rightarrow \mathbb{R}$, $\arctan : \mathbb{R} \rightarrow (-\frac{\pi}{2}, \frac{\pi}{2})$, $\operatorname{Arsinh} : \mathbb{R} \rightarrow \mathbb{R}$, $\operatorname{Artanh} : (-1, 1) \rightarrow \mathbb{R}$ sind differenzierbar. Dabei ist $\log$ die Umkehrfunktion von $\exp$ und $\exp' = \exp$, also

$$\log' y = \frac{1}{e^{\log y}} = \frac{1}{y} \quad (y > 0).$$

Weiter gilt

$$\arctan'(y) = \frac{1}{1 + y^2} \quad (y \in \mathbb{R}), \quad \operatorname{Arsinh}'(y) = \frac{1}{\sqrt{1 + y^2}} \quad (y \in \mathbb{R})$$

$$\operatorname{Artanh}'(y) = \frac{1}{1 - y^2} \quad (y \in (-1, 1)).$$

Die Funktionen $\arcsin : (-1, 1) \rightarrow (-\frac{\pi}{2}, \frac{\pi}{2})$, $\arccos : (-1, 1) \rightarrow (0, \pi)$ und $\operatorname{Arcosh} : (1, \infty) \rightarrow (0, \infty)$ sind differenzierbar mit

$$\arcsin'(y) = \frac{1}{\sqrt{1 - y^2}} \quad (y \in (-1, 1)), \quad \arccos'(y) = -\frac{1}{\sqrt{1 - y^2}} \quad (y \in (-1, 1)),$$

$$\operatorname{Arcosh}'(y) = \frac{1}{\sqrt{y^2 - 1}} \quad (y \in (1, \infty)).$$

Anwendung: Für $\alpha \in \mathbb{R}$ betrachten wir $x \mapsto x^\alpha$ auf $(0, \infty)$. Es gilt:

$$(x^\alpha)' = (e^{\alpha \log x})' = e^{\alpha \log x} \cdot \frac{\alpha}{x} = \alpha x^{\alpha-1} \quad (x \in (0, \infty)).$$

Ein wichtiger Grenzwert: Für jedes $x \in \mathbb{R}$ gilt:

$$\lim_{n \rightarrow \infty} \left(1 + \frac{x}{n}\right)^n = e^x.$$

Beweis. Für $n > -x$ ist $n \log(1 + \frac{x}{n})$ definiert. Wir zeigen zuerst

$$n \log(1 + \frac{x}{n}) \rightarrow x \quad (n \rightarrow \infty).$$

Das ist klar für $x = 0$. Für $x \neq 0$ gilt

$$\lim_{n \rightarrow \infty} \frac{\log(1 + \frac{x}{n})}{\frac{1}{n}} = x \lim_{n \rightarrow \infty} \frac{\log(1 + \frac{x}{n})}{\frac{x}{n}} = x \lim_{h \rightarrow 0} \frac{\log(1 + h) - \log 1}{h} = x \log'(1) = x.$$

Wegen $\exp \in C(\mathbb{R})$ folgt

$$\left(1 + \frac{x}{n}\right)^n = \exp\left(n \log\left(1 + \frac{x}{n}\right)\right) \rightarrow e^x \quad (n \rightarrow \infty).$$

□

11.5. Lokale Extremstellen: Sei $D \subseteq \mathbb{R}$ mit $D \neq \emptyset$, $g : D \rightarrow \mathbb{R}$ eine Funktion und $x_0 \in D$.

Definition: (a) g hat in x_0 ein *lokales Maximum* [*Minimum*], falls es ein $\delta > 0$ gibt mit

$$\forall x \in D \cap U_\delta(x_0) : g(x) \leq g(x_0) \quad [g(x) \geq g(x_0)],$$

d.h.

$$\forall x \in D : |x - x_0| < \delta \Rightarrow g(x) \leq g(x_0) \quad [g(x) \geq g(x_0)].$$

Ein lokales Maximum/Minimum wird auch als *relatives Maximum/Minimum* bezeichnet.

Ein *relatives* oder *lokales Extremum* ist ein lokales Maximum oder ein lokales Minimum.

(b) g hat in x_0 ein *globales Maximum* [*Minimum*], falls

$$\forall x \in D : g(x) \leq g(x_0) \quad [g(x) \geq g(x_0)].$$

Ein *globales Extremum* ist ein globales Maximum oder ein globales Minimum.

(c) Globale oder lokale Extrema heißen *strikt*, wenn für $x \neq x_0$ jeweils “<” bzw. “>” gilt.

(d) Ein Punkt $x_0 \in D$ heißt *innerer Punkte von D*, falls es ein $\delta > 0$ mit $U_\delta(x_0) \subseteq D$ gibt.

Bemerkung: Ist $D = I$ ein Intervall, so ist $x_0 \in I$ genau dann ein innerer Punkt von I , wenn $x_0 \notin \{\sup I, \inf I\}$ gilt.

11.6. Satz: Die Funktion $f : I \rightarrow \mathbb{R}$ habe in einem inneren Punkt x_0 von I ein lokales Extremum und sei in x_0 differenzierbar. Dann gilt $f'(x_0) = 0$.

Beweis: Wir nehmen an, dass f in x_0 ein lokales Maximum hat (sonst betrachte $-f$). Dann existiert ein $\delta > 0$ mit $U_\delta(x_0) \subseteq I$ und $f(x) \leq f(x_0)$ ($x \in U_\delta(x_0)$). Also gilt

$$\frac{f(x) - f(x_0)}{x - x_0} \begin{cases} \leq 0, & x \in (x_0, x_0 + \delta) \\ \geq 0, & x \in (x_0 - \delta, x_0) \end{cases}$$

und damit

$$f'(x_0) = \begin{cases} \lim_{x \rightarrow x_0+} \frac{f(x) - f(x_0)}{x - x_0} \leq 0 \\ \lim_{x \rightarrow x_0-} \frac{f(x) - f(x_0)}{x - x_0} \geq 0 \end{cases}.$$

Somit ist $f'(x_0) = 0$. □

11.7. Mittelwertsatz der Differentialrechnung (MWS): Es sei $f : [a, b] \rightarrow \mathbb{R}$ stetig und auf (a, b) differenzierbar. Dann gilt:

$$\exists \xi \in (a, b) : f'(\xi) = \frac{f(b) - f(a)}{b - a}.$$

Beweis. Wir betrachten zunächst den Fall $f(a) = f(b)$. Nach 9.15 existieren $x_1, x_2 \in [a, b]$ mit

$$f(x_1) \leq f(x) \leq f(x_2) \quad (x \in [a, b]).$$

Somit hat f in x_1 ein globales Minimum und in x_2 ein globales Maximum.

Fall 1, $x_1 \in (a, b)$ oder $x_2 \in (a, b)$: Nach 11.6 ist dann $f'(x_1) = 0$ oder $f'(x_2) = 0$.

Fall 2, $x_1, x_2 \in \{a, b\}$: Dann ist f konstant und $f'(x) = 0$ für jedes $x \in (a, b)$.

Im allgemeinen Fall setzen wir

$$g(x) := f(x) - \frac{f(b) - f(a)}{b - a}(x - a) \quad (x \in [a, b]).$$

Dann ist $g(a) = f(a) = g(b)$ und wir erhalten ein $\xi \in (a, b)$ mit

$$0 = g'(\xi) = f'(\xi) - \frac{f(b) - f(a)}{b - a}.$$

□

11.8. Folgerungen: Sei $I \subseteq \mathbb{R}$ ein Intervall und $f, g : I \rightarrow \mathbb{R}$ 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 $f' \geq 0$ [bzw. $f' \leq 0$, $f' > 0$, $f' < 0$] auf I , so ist f auf I monoton wachsend [bzw. monoton fallend, streng monoton wachsend, streng monoton fallend].

Beweis. (1) “ $\Leftarrow$ ”: Nach dem MWS ist $f(y) = f(x)$ für alle $x, y \in I$.

(2) Wende (1) auf $f - g$ an.

(3) Ist $f' \geq 0$ auf I , so gilt für $x, y \in I$ mit $x < y$:

$$\exists \xi \in (x, y) : \frac{f(y) - f(x)}{y - x} = f'(\xi) \geq 0.$$

Es folgt $f(y) \geq f(x)$. Die anderen Aussagen beweist man analog. $\square$

Eine erste Differentialgleichung: Sei $a \in \mathbb{R}$ und $f : \mathbb{R} \rightarrow \mathbb{R}$ eine differenzierbare Funktion mit

$$f'(x) = af(x) \quad (x \in \mathbb{R}),$$

sowie $x_0 \in \mathbb{R}$. Dann gilt $f(x) = f(x_0)e^{a(x-x_0)}$ ($x \in \mathbb{R}$): Setzt man $g(x) := f(x)e^{-ax}$ ($x \in \mathbb{R}$), so ist g differenzierbar auf $\mathbb{R}$ mit $g' = 0$ auf $\mathbb{R}$, und somit

$$f(x)e^{-ax} = g(x) = g(x_0) = f(x_0)e^{-ax_0} \quad (x \in \mathbb{R}).$$

Somit hat für feste $x_0, c \in \mathbb{R}$ das *Anfangswertproblem*

$$y'(x) = ay(x), \quad y(x_0) = c$$

auf $\mathbb{R}$ genau eine Lösung, gegeben durch $y(x) = ce^{a(x-x_0)}$ ($x \in \mathbb{R}$).

11.9. Die Regeln von de l'Hospital: Es seien $a \in \mathbb{R} \cup \{-\infty\}$, $b \in \mathbb{R} \cup \{\infty\}$ und $f, g : (a, b) \rightarrow \mathbb{R}$ auf (a, b) differenzierbar mit $g'(x) \neq 0$ ($x \in (a, b)$). Weiter sei $c = a$ oder $c = b$ und

$$\lim_{x \rightarrow c} \frac{f'(x)}{g'(x)} = L \in \mathbb{R} \cup \{-\infty, \infty\}.$$

(a) Ist $\lim_{x \rightarrow c} f(x) = \lim_{x \rightarrow c} g(x) = 0$, so gilt

$$\frac{f(x)}{g(x)} \rightarrow L \quad (x \rightarrow c).$$

(b) Ist $\lim_{x \rightarrow c} g(x) = \pm\infty$, so gilt

$$\frac{f(x)}{g(x)} \rightarrow L \quad (x \rightarrow c).$$

Beweis. Wir beweisen nur (a) mit $c = b \in \mathbb{R}$: Setze $f(b) = g(b) = 0$. Dann sind f, g stetig auf $(a, b]$. Nach dem MWS existiert zu jedem $x \in (a, b)$ ein $\eta \in (x, b)$ mit

$$g(x) = g(x) - g(b) = g'(\eta)(x - b) \neq 0.$$

Also ist $g(x) \neq 0$ ($x \in (a, b)$).

Sei nun (x_n) eine Folge in (a, b) mit $x_n \rightarrow b$. Sei $n \in \mathbb{N}$. Betrachte

$$h(x) := g(x)f(x_n) - f(x)g(x_n) \quad (x \in [x_n, b]).$$

Dann ist $h(x_n) = 0 = h(b)$. Nach dem MWS existiert $\xi_n \in (x_n, b)$ mit $h'(\xi_n) = 0$, d.h. mit

$$\frac{f(x_n)}{g(x_n)} = \frac{f'(\xi_n)}{g'(\xi_n)}.$$

Die Behauptung folgt nun für $n \rightarrow \infty$. □

Beispiele: (1) Für $a, b > 0$ gilt

$$\lim_{x \rightarrow 0+} \frac{a^x - b^x}{x} = \lim_{x \rightarrow 0+} \frac{a^x \log a - b^x \log b}{1} = \log a - \log b.$$

(2) Es ist

$$\lim_{x \rightarrow \infty} \frac{\log x}{x} = \lim_{x \rightarrow \infty} \frac{1/x}{1} = 0.$$

(3) Es ist

$$\lim_{x \rightarrow 0+} x \log x = \lim_{x \rightarrow 0+} \frac{\log x}{1/x} = \lim_{x \rightarrow 0+} \frac{1/x}{-1/x^2} = \lim_{x \rightarrow 0+} (-x) = 0.$$

(4) Aus (3) folgt mit der Stetigkeit der Exponentialfunktion:

$$\lim_{x \rightarrow 0+} x^x = \lim_{x \rightarrow 0+} e^{x \log x} = e^0 = 1.$$

Bemerkung: Hier ist es jeweils so, dass erst die Existenz des letzten Limes die Existenz des ersten Limes garantiert (vgl. die Regeln oben).

(5) Vorsicht:

$$0 = \lim_{x \rightarrow 1+} \frac{\log x}{x} \neq \lim_{x \rightarrow 1+} \frac{1/x}{1} = 1$$

Der folgende Satz kann in Anwendungen nützlich sein.

11.10. Satz: Sei $f : [a, b] \rightarrow \mathbb{R}$ eine Funktion und $x_0 \in (a, b)$. Weiter sei f auf (a, x_0) und (x_0, b) differenzierbar. Existiert der Grenzwert $\alpha := \lim_{x \rightarrow x_0} f'(x) \in \mathbb{R}$ 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 $x \rightarrow x_0$ hat man $\xi(x) \rightarrow x_0$ und also $f'(\xi(x)) \rightarrow \alpha$ nach Voraussetzung. $\square$

11.11. Höhere Ableitungen: Sei $I \subseteq \mathbb{R}$ ein Intervall und $f : I \rightarrow \mathbb{R}$ sei auf I differenzierbar.

Definition: (a) f heißt in $x_0 \in I$ *zweimal differenzierbar*, falls f' in x_0 differenzierbar ist. Dann heißt

$$f''(x_0) := (f')'(x_0)$$

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

$$f'''(x_0), f^{(4)}(x_0), f^{(5)}(x_0), \dots \text{ bzw. } f''', f^{(4)}, f^{(5)} \dots$$

(b) Sei $n \in \mathbb{N}$. f heißt auf I *n-mal stetig differenzierbar*, falls f auf I n -mal differenzierbar ist und $f^{(n)} \in C(I)$ ist. Dafür schreiben wir $f \in C^n(I)$. Für $n = 0$ ist $f^{(0)} := f$ und $C^0(I) := C(I)$. Wir setzen

$$C^\infty(I) := \bigcap_{n \in \mathbb{N}_0} C^n(I),$$

d.h. $f \in C^\infty(I)$ wenn f auf I beliebig oft differenzierbar ist.

Beispiele: (1) Es gilt $\exp \in C^\infty(\mathbb{R})$. Hier ist $\exp^{(n)} = \exp$ ($n \in \mathbb{N}_0$).

(2) Es gilt $\sin, \cos \in C^\infty(\mathbb{R})$. Hier ist

$$\sin' = \cos, \quad \sin'' = \cos' = -\sin, \quad \sin''' = -\sin' = -\cos, \quad \sin^{(4)} = -\cos' = \sin, \quad \text{etc..}$$

Genauso sind $\sinh, \cosh \in C^\infty(\mathbb{R})$.

(3) Die Funktion $f(x) = \frac{1}{1-x}$ ist auf $I = (-\infty, 1)$ bzw. auf $I = (1, \infty)$ beliebig oft differenzierbar mit

$$f^{(n)}(x) = \frac{n!}{(1-x)^{n+1}} \quad (n \in \mathbb{N}_0).$$

(4) Jedes Polynom ist beliebig oft differenzierbar auf $\mathbb{R}$.

(5) Die Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$,

$$f(x) := \begin{cases} x^2 \sin(1/x), & x \neq 0 \\ 0, & x = 0 \end{cases}$$

ist auf $\mathbb{R}$ differenzierbar: Es gilt

$$f'(x) = 2x \sin(1/x) - \cos(1/x) \quad (x \neq 0).$$

Weiter ist

$$\left| \frac{f(x) - f(0)}{x - 0} \right| = |x| |\sin(1/x)| \leq |x| \rightarrow 0 \quad (x \rightarrow 0),$$

also ist f in 0 differenzierbar und $f'(0) = 0$. Man beachte, dass der Limes $\lim_{x \rightarrow 0} f'(x)$ nicht existiert: Z.B. für $(x_n) := (\frac{1}{n\pi})$ gilt $x_n \rightarrow 0$ und

$$f'(x_n) = \frac{2}{n\pi} \sin(n\pi) - \cos(n\pi) = (-1)^{n+1} \quad (n \in \mathbb{N}).$$

Damit ist f' in 0 nicht stetig, also $f \notin C^1(\mathbb{R})$.

(6) Für die Funktion $f : \mathbb{R} \rightarrow \mathbb{R}$,

$$f(x) := \begin{cases} e^{-1/x}, & x > 0 \\ 0, & x \leq 0 \end{cases}$$

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 $p_1(\frac{1}{x}) = \frac{1}{x^2}$, 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 ist

$$f^{(n+1)}(x) = p'_n(1/x) \cdot (-x^{-2})e^{-1/x} + p_n(1/x)e^{-1/x} \cdot x^{-2} \quad (x > 0),$$

also $p_{n+1}(s) = -s^2 p'_n(s) + s^2 p_n(s)$. Wir wenden den Satz 11.10 wiederholt auf $f, f', f'', \dots$ an: Für jedes $n \in \mathbb{N}$ gilt (vgl. Bsp.(2) in 9.5):

$$\lim_{x \rightarrow 0+} f^{(n)}(x) = \lim_{x \rightarrow 0+} p_n\left(\frac{1}{x}\right)e^{-1/x} = \lim_{s \rightarrow \infty} p_n(s)e^{-s} = 0.$$

Wir erhalten so, dass $f, f', f'', \dots$ auf $\mathbb{R}$ differenzierbar sind, mit

$$\forall n \in \mathbb{N}_0 : f^{(n)}(0) = 0.$$

11.12. Satz von Taylor: Sei $n \in \mathbb{N}_0$ und $f : I \rightarrow \mathbb{R}$ sei $n + 1$ -mal auf I differenzierbar. Seien $x, x_0 \in I$ mit $x \neq x_0$. Dann gibt es ein ξ zwischen x und x_0 mit

$$\begin{aligned} f(x) &= f(x_0) + \frac{f'(x_0)}{1!}(x - x_0) + \dots + \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n + \frac{f^{(n+1)}(\xi)}{(n+1)!}(x - x_0)^{n+1} \\ &= \underbrace{\sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k}_{n\text{-tes Taylorpolynom}} + \underbrace{\frac{f^{(n+1)}(\xi)}{(n+1)!}(x - x_0)^{n+1}}_{\text{Restglied}}. \end{aligned}$$

Bemerkung: Für $n = 0$ ist das der Mittelwertsatz, den wir auch im Beweis verwenden.

Definition: Für $n \in \mathbb{N}_0$ schreiben wir

$$T_n(f; x_0)(x) := \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k,$$

für das n -te Taylorpolynom, und

$$R_{n+1}(f; x_0)(x) := f(x) - T_n(f; x_0)(x)$$

für das n -te Restglied von f bei Entwicklung um x_0 .

Beweis des Satzes. O.B.d.A. sei $x_0 = 0$ und $x > 0$. Wir wählen $c \in \mathbb{R}$ so, dass gilt

$$f(x) = \sum_{k=0}^n \frac{f^{(k)}(0)}{k!}x^k + c \frac{x^{n+1}}{(n+1)!},$$

und definieren $g : [0, x] \rightarrow \mathbb{R}$ durch

$$g(y) := f(x) - \sum_{k=0}^n \frac{f^{(k)}(y)}{k!}(x - y)^k - c \frac{(x - y)^{n+1}}{(n+1)!}.$$

Dann haben wir $g(x) = 0 = g(0)$. Nach Voraussetzung ist g in $[0, x]$ differenzierbar. Nach dem MWS existiert ein $\xi \in (0, x)$ mit $g'(\xi) = 0$. Wegen (nachrechnen!)

$$g'(y) = (c - f^{(n+1)}(y)) \frac{(x - y)^n}{n!} \quad (y \in [0, x])$$

folgt dann $c = f^{(n+1)}(\xi)$. □

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:

$$T_n(f; x_0)(x) \rightarrow f(x) \quad (n \rightarrow \infty).$$

Gleichbedeutend damit ist $R_{n+1}(f; x_0)(x) \rightarrow 0$ ($n \rightarrow \infty$) ($x \in I$).

Beispiel: Wir betrachten $f : (-\infty, 1) \rightarrow \mathbb{R}$, $f(x) := \log(1 - x)$ und $x_0 = 0$. Hier ist $f'(x) = -\frac{1}{1-x}$ und also nach 11.11(3)

$$f^{(k)}(x) = -\frac{(k-1)!}{(1-x)^k}, \quad f^{(k)}(0) = -(k-1)! \quad (k \in \mathbb{N}).$$

Wir erhalten somit für jedes $n \in \mathbb{N}$ und $x < 1$:

$$T_n(f; 0)(x) = -\sum_{k=1}^n \frac{x^k}{k}, \quad R_{n+1}(f; 0)(x) = -\frac{1}{(1-\xi)^{n+1}} \frac{x^{n+1}}{n+1}.$$

Da ξ zwischen 0 und x liegt, erhalten wir

$$0 < \frac{x}{1-\xi} \leq \frac{x}{1-x} \leq 1 \quad (x \in (0, \frac{1}{2}]), \quad 0 \leq \left| \frac{x}{1-\xi} \right| = \frac{|x|}{1-\xi} \leq |x| \leq 1 \quad (x \in [-1, 0]).$$

Also gilt $R_{n+1}(f; 0)(x) \rightarrow 0$ ($n \rightarrow \infty$) ($x \in [-1, \frac{1}{2}]$), und damit

$$\log(1-x) = -\sum_{k=1}^{\infty} \frac{x^k}{k} \quad (x \in [-1, \frac{1}{2}]).$$

Tatsächlich gilt dies auch für $x \in (1/2, 1)$ (vgl. 11.14 unten). Wir notieren als Spezialfall ($x = -1$)

$$\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{k} = \log 2.$$

Warnung: Nicht jede Funktion $f \in C^\infty(I)$ ist durch ihre Taylorreihe darstellbar!

Beispiel: Wir betrachten,

$$f : \mathbb{R} \rightarrow \mathbb{R}, \quad f(x) = \begin{cases} e^{-1/x}, & x > 0 \\ 0, & x \leq 0 \end{cases},$$

und $x_0 = 0$. Es gilt $f^{(n)}(0) = 0$ ($n \in \mathbb{N}_0$), also $T_n(f; 0)(x) = 0$ ($n \in \mathbb{N}_0, x \in \mathbb{R}$). Für **kein** $x > 0$ gilt somit $T_n(f; 0)(x) \rightarrow f(x)$.

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 . Weiter sei

$$f'(x_0) = f''(x_0) = \dots = f^{(n-1)}(x_0) = 0 \quad \textbf{und} \quad f^{(n)}(x_0) \neq 0.$$

(a) Ist n gerade und $f^{(n)}(x_0) > 0$ [bzw. $f^{(n)}(x_0) < 0$], so hat f in x_0 ein striktes lokales Minimum [bzw. Maximum].

(b) Ist n ungerade, so hat f in x_0 kein lokales Extremum.

Bemerkung: In Anwendungen ist meist $n = 2$ (vgl. Schule). Ist $f'(x_0) = 0$, so gilt:

für $f''(x_0) > 0$ hat f in x_0 ein striktes lokales Minimum,

für $f''(x_0) < 0$ hat f in x_0 ein striktes lokales Maximum,

für $f''(x_0) = 0$ erhält man keine Entscheidung.

Beweis. Es gilt $f^{(n)} \in C(I)$ mit $f^{(n)}(x_0) \neq 0$. Somit existiert ein $\delta > 0$ 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

$$f(x) = T_{n-1}(f; x_0)(x) + R_n(f; x_0)(x) = f(x_0) + \frac{f^{(n)}(\xi)}{n!}(x - x_0)^n.$$

Nun betrachte man das Vorzeichen des Restgliedes. □

Für die Suche nach globalen Extremstellen ist folgende Situation besonders geeignet:

Konvexe und konkave Funktionen: $f : I \rightarrow \mathbb{R}$ heißt *konvex* [*konkav*], wenn gilt:

$$(*) \quad \forall x, y \in I \quad \forall \lambda \in (0, 1) : f((1 - \lambda)x + \lambda y) \leq [\geq](1 - \lambda)f(x) + \lambda f(y),$$

d.h. wenn die Funktion auf jedem Teilintervall $[x, y]$ unterhalb [oberhalb] der Geraden durch $(x, f(x))$ und $(y, f(y))$ liegt. Ohne Beweis: Ist $f \in C^2(I)$, so ist dies äquivalent zu $f'' \geq 0$ auf I [$f'' \leq 0$ auf I].

Satz: Es sei $f \in C^2(I)$ mit $f'' \geq 0$ auf I [bzw. mit $f'' \leq 0$ auf I]. Ist $x_0 \in I$ mit $f'(x_0) = 0$, so hat f in x_0 ein globales Minimum [bzw. ein globales Maximum].

Beweis. Wir betrachten den Fall $f'' \geq 0$ auf I . Dann ist f' auf I nach 11.8 monoton wachsend, also $f' \leq 0$ links von x_0 und $f' \geq 0$ rechts von x_0 . Wieder nach 11.8 ist also f monoton fallend links von x_0 und monoton wachsend rechts von x_0 . Wir erhalten so $f(x) \geq f(x_0)$ ($x \in I$). □

Beispiele: (1) Es sei $p > 1$, $\alpha > 0$. Bestimme das Maximum von

$$f : (0, \infty) \rightarrow \mathbb{R}, \quad f(x) = \alpha x - x^p/p.$$

Für $x > 0$ ist

$$f'(x) = \alpha - x^{p-1}, \quad f''(x) = -(p-1)x^{p-2} < 0.$$

Also ist f konkav. Es gilt $f'(x_0) = 0 \Leftrightarrow x_0 = \alpha^{1/(p-1)}$. Also hat f in x_0 ein globales Maximum. 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.

(2) $f : \mathbb{R} \rightarrow \mathbb{R}$, $f(x) = e^x - x$ ist in $C^\infty(\mathbb{R})$ 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 Konvergenzradius $R \in (0, \infty]$, $I := (x_0 - R, x_0 + R)$ und

$$f : I \rightarrow \mathbb{R}, \quad x \mapsto f(x) := \sum_{n=0}^{\infty} a_n(x - x_0)^n.$$

Dann ist f auf I differenzierbar, und für jedes $x \in I$ gilt

$$f'(x) = \sum_{n=1}^{\infty} n a_n(x - x_0)^{n-1}.$$

Da die Potenzreihe auf jedem Intervall $[x_0 - r, x_0 + r]$, $0 \leq r < R$ gleichmäßig konvergiert (vgl. 9.9) folgt dies aus dem allgemeineren

Satz: Sei (f_n) eine Folge differenzierbarer Funktionen $f_n : I \rightarrow \mathbb{R}$, die auf einem Intervall I punktweise gegen die Funktion $f : I \rightarrow \mathbb{R}$ konvergiert. Die Folge (f'_n) der Ableitungen konvergiere auf I gleichmäßig gegen eine Funktion g und g sei stetig in $x_0 \in I$. Dann ist f in x_0 differenzierbar und $f'(x_0) = g(x_0)$.

Beweis. Nach dem MWS gilt für $x \in I \setminus \{x_0\}$ und $n \in \mathbb{N}$:

$$\frac{f_n(x) - f_n(x_0)}{x - x_0} - g(x_0) = f'_n(\xi_n(x)) - g(\xi_n(x)) + g(\xi_n(x)) - g(x_0),$$

für ein $\xi_n(x)$ zwischen x und x_0 . Zu $\varepsilon > 0$ existiert ein $n_0 \in \mathbb{N}$ so, dass

$$\forall n \geq n_0 \quad \forall x \in I : |f'_n(x) - g(x)| < \frac{\varepsilon}{2}.$$

Da g in x_0 stetig ist, existiert ein $\delta > 0$ so, dass

$$\forall x \in I : |x - x_0| < \delta \Rightarrow |g(x) - g(x_0)| < \frac{\varepsilon}{2}.$$

Für $x \in I$ mit $|x - x_0| < \delta$ und $n \geq n_0$ gilt wegen $|\xi_n(x) - x_0| < \delta$ nun

$$\left| \frac{f_n(x) - f_n(x_0)}{x - x_0} - g(x_0) \right| \leq |f'_n(\xi_n(x)) - g(\xi_n(x))| + |g(\xi_n(x)) - g(x_0)| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Für $n \rightarrow \infty$ erhalten wir

$$\forall x \in I : |x - x_0| < \delta \Rightarrow \left| \frac{f(x) - f(x_0)}{x - x_0} - g(x_0) \right| \leq \varepsilon.$$

□

Folgerung: Sei (f_n) ein Funktionenfolge in $C^1([a, b])$, die punktweise gegen eine Funktion $f : [a, b] \rightarrow \mathbb{R}$ konvergiert und deren Ableitungen (f'_n) gleichmäßig gegen eine Funktion $g : [a, b] \rightarrow \mathbb{R}$ konvergieren. Dann ist $f \in C^1([a, b])$ und $f' = g$ auf $[a, b]$.

Bemerkung: Man kann also Potenzreihen im Inneren des Konvergenzintervalls gliedweise differenzieren. Da f' wieder durch eine Potenzreihe mit Konvergenzradius R gegeben ist, kann man diese Aussage wiederholt anwenden und erhält damit: $f \in C^\infty(I)$ und für alle $k \in \mathbb{N}$ gilt:

$$f^{(k)}(x) = \sum_{n=k}^{\infty} n(n-1)(n-2) \cdot \dots \cdot (n-(k-1)) a_n (x-x_0)^{n-k} \quad (x \in I).$$

Für $x = x_0$ erhält man

$$f^{(k)}(x_0) = k(k-1)(k-2) \cdot \dots \cdot (k-(k-1)) a_k = k! a_k.$$

Somit gilt

$$a_k = \frac{f^{(k)}(x_0)}{k!} \quad (k \in \mathbb{N}_0),$$

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 dieser 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

$$\arctan(x) = g(x) = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{2k+1} \quad (x \in (-1, 1)).$$

Wir stellen fest, dass die Arcustangens-Reihe auch noch für $x = \pm 1$ konvergiert (Leibnizkriterium), und würden gerne $x = 1$ einsetzen. Das geht nach dem Abelschen Grenzwertsatz (siehe 9.8) und wir erhalten

$$\begin{aligned}\frac{\pi}{4} &= \arctan 1 = \lim_{x \rightarrow 1-} \arctan(x) = \lim_{x \rightarrow 1-} \sum_{k=0}^{\infty} \frac{(-1)^k x^{2k+1}}{2k+1} \\ &= \sum_{k=0}^{\infty} \frac{(-1)^k}{2k+1} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \dots\end{aligned}$$

Eine ähnliche Argumentation zeigt

$$\log(1-x) = -\sum_{n=1}^{\infty} \frac{x^n}{n}, \quad x \in (-1, 1).$$

Da diese Reihe auch für $x = -1$ konvergiert, erhalten wir aus dem Abelschen Grenzwertsatz

$$\log 2 = \sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n}.$$

11.15. Identitätssatz für Potenzreihen: Sei $r > 0$ und $I := (x_0 - r, x_0 + r)$. Seien $f, g : I \rightarrow \mathbb{R}$ gegeben durch

$$f(x) := \sum_{n=0}^{\infty} a_n (x - x_0)^n, \quad g(x) := \sum_{n=0}^{\infty} b_n (x - x_0)^n,$$

wobei beide Potenzreihen auf I konvergieren. Gilt $f = g$ auf I , so folgt mit 11.14

$$\forall n \in \mathbb{N}_0 : a_n = \frac{f^{(n)}(x_0)}{n!} = \frac{g^{(n)}(x_0)}{n!} = b_n.$$

Bemerkung: Die Voraussetzung lässt sich sehr abschwächen. Gibt es z.B. eine streng monotone Folge $(x_m)_{m \in \mathbb{N}}$ in I mit $x_m \rightarrow x_0$ für $m \rightarrow \infty$ und $f(x_m) = g(x_m)$ ($m \in \mathbb{N}$). Dann gilt $a_n = b_n$ ($n \in \mathbb{N}_0$) (ohne Beweis).

Beispiel: Wir wollen ein Intervall I mit $0 \in I$ und eine differenzierbare Funktion $y : I \rightarrow \mathbb{R}$ bestimmen mit $y'(x) = xy(x)$ ($x \in I$) und $y(0) = 1$. Wir machen einen *Potenzreihenansatz*

$$y(x) = \sum_{n=0}^{\infty} a_n x^n \quad (x \in I = (-R, R)),$$

wobei R der Konvergenzradius der Potenzreihe sei. Wegen $y(0) = 1$ ist $a_0 = 1$. Ist $R > 0$, so ist nach 11.14:

$$y'(x) = \sum_{n=1}^{\infty} n a_n x^{n-1} \quad (x \in I).$$

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* (d.h. nach 11.15) folgt

$$a_1 = 0, \quad a_n = \frac{a_{n-2}}{n} \quad (n \geq 2).$$

Durch Induktion erhält man hieraus $0 = a_1 = a_3 = a_5 = a_7 = \dots$ und

$$a_{2k} = \frac{1}{2^k k!} \quad (k \in \mathbb{N}_0).$$

Wir erhalten

$$y(x) = \sum_{k=0}^{\infty} \frac{x^{2k}}{2^k k!} = \sum_{k=0}^{\infty} \frac{1}{k!} \left(\frac{x^2}{2}\right)^k = e^{x^2/2} \quad (x \in \mathbb{R}),$$

und diese Funktion genügt tatsächlich den geforderten Bedingungen (einsetzen). Wir haben also durch Potenzreihenansatz eine Lösung des Anfangswertproblems

$$y'(x) = xy(x), \quad y(0) = 1$$

bestimmt.

12 Integration

12.1. Ober- und Untersummen, oberes und unteres Integral: Es seien $a, b \in \mathbb{R}$ mit $a < b$ und es sei $f : [a, b] \rightarrow \mathbb{R}$ *beschränkt*, d.h. $f([a, b]) = \{f(x) : x \in [a, b]\}$ ist beschränkt.

Definition: $Z := \{x_0, x_1, \dots, x_n\}$ heißt eine *Zerlegung* von $[a, b]$, falls

$$a = x_0 < x_1 < \dots < x_n = b.$$

Es bezeichne $\mathcal{Z} = \mathcal{Z}([a, b])$ die Menge aller Zerlegungen (von $[a, b]$).

Wir setzen $m := \inf f([a, b])$ und $M := \sup f([a, b])$. Es sei $Z = \{x_0, x_1, \dots, x_n\} \in \mathcal{Z}$. Für $j = 1, 2, \dots, n$ sei

$$\begin{aligned} I_j &:= [x_{j-1}, x_j], & |I_j| &:= x_j - x_{j-1}, & m_j &:= \inf f(I_j), & M_j &:= \sup f(I_j), & \text{ sowie} \\ s_f(Z) &:= \sum_{j=1}^n m_j |I_j| & & \text{Untersumme von } f \text{ bzgl. } Z, \\ S_f(Z) &:= \sum_{j=1}^n M_j |I_j| & & \text{Obersumme von } f \text{ bzgl. } Z. \end{aligned}$$

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 $[a, b]$:

$$(*) \quad m(b-a) \leq s_f(Z) \leq S_f(Z) \leq M(b-a).$$

Satz: Seien $Z_1, Z_2 \in \mathcal{Z}$. Dann gilt:

- (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:

$$s_f := \sup\{s_f(Z) : Z \in \mathcal{Z}\},$$

das *untere Integral von f* über $[a, b]$ und

$$S_f := \inf\{S_f(Z) : Z \in \mathcal{Z}\}$$

das *obere Integral von f* über $[a, b]$.

Wegen (*) und (2) gilt dann

$$(**) \quad m(b-a) \leq s_f \leq S_f \leq M(b-a).$$

[Aus (2) folgt durch Supremumbildung 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 $f : [a, b] \rightarrow \mathbb{R}$ *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

$$R([a, b]) := \{f : [a, b] \rightarrow \mathbb{R} : f \text{ ist über } [a, b] \text{ integrierbar}\}$$

(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 $Z = \{x_0, x_1, \dots, x_n\} \in \mathcal{Z}([0, 1])$ und sind m_j, M_j, I_j wie oben, so gilt $m_j = 0, M_j = 1$ für alle j ², also $s_f(Z) = 0$ und $S_f(Z) = 1$. Damit ist

$$s_f = 0 \neq 1 = S_f,$$

und f ist nicht integrierbar über $[0, 1]$.

²Hier wird benutzt, dass jedes Intervall (mit mindestens zwei Punkten) rationale und irrationale Zahlen enthält.

12.3. Satz: Es seien $f, g \in R([a, b])$.

(1) Monotonie des Integrals: Gilt $f \leq g$ auf $[a, b]$, so ist $\int_a^b f \, dx \leq \int_a^b g \, dx$.

(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 $R([a, b]) \rightarrow \mathbb{R}, f \mapsto \int_a^b f \, dx$, ist linear.

(3) Setzt man $\gamma := \sup\{|f(x)| : x \in [a, b]\}$, so gilt

$$\left| \int_a^b f(x) \, dx \right| \leq \gamma(b-a).$$

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

$$s_f(Z) + s_g(Z) \leq s_{f+g}(Z) \leq S_{f+g}(Z) \leq S_f(Z) + S_g(Z) \quad (Z \in \mathbb{Z}),$$

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 (**):

$$\left| \int_a^b f(x) \, dx \right| \leq \max\{|M|, |m|\}(b-a).$$

Beachte nun $\gamma = \max\{|M|, |m|\}$. □

12.4. Riemann-Kriterium: Sei $f : [a, b] \rightarrow \mathbb{R}$ eine beschränkte Funktion. Dann gilt:

$$f \in R([a, b]) \Leftrightarrow \forall \varepsilon > 0 \exists Z \in \mathbb{Z} : S_f(Z) - s_f(Z) < \varepsilon.$$

Äquivalent ist: $f \in R([a, b])$ genau dann, wenn es eine Folge $(Z_l)_{l \in \mathbb{N}}$ in $\mathbb{Z}$ gibt mit $\lim_{l \rightarrow \infty} S_f(Z_l) - s_f(Z_l) = 0$. In diesem Fall ist (wegen $s_f(Z_l) \leq s_f = S_f \leq S_f(Z_l)$ ($l \in \mathbb{N}$))

$$\int_a^b f(x) \, dx = \lim_{l \rightarrow \infty} S_f(Z_l) = \lim_{l \rightarrow \infty} s_f(Z_l).$$

Beweis. “ $\Leftarrow$ ”: Sei $\varepsilon > 0$. Dann existiert also ein $Z \in \mathbb{Z}$ mit $S_f(Z) - s_f(Z) < \varepsilon$. Somit gilt

$$S_f \leq S_f(Z) \leq s_f(Z) + \varepsilon \leq s_f + \varepsilon, \text{ also } 0 \leq S_f - s_f \leq \varepsilon.$$

Da $\varepsilon > 0$ beliebig war folgt $S_f = s_f$.

“ $\Rightarrow$ ”: Dies folgt aus der Definition des Integrals mithilfe von 12.1 Satz (1). (Übung.) □

Wir betrachten eine *äquidistante* Zerlegung $Z_n := \{x_0, x_1, \dots, x_n\}$ für $n \in \mathbb{N}$, d.h.

$$x_j := a + \frac{j}{n}(b-a) \quad (j = 0, 1, \dots, n).$$

Sind I_j , M_j und m_j wie in 12.1, so gilt $|I_j| = (b-a)/n$ ($j = 1, \dots, n$) und

$$(***) \quad S_f(Z_n) - s_f(Z_n) = \frac{b-a}{n} \sum_{j=1}^n (M_j - m_j) \leq (b-a) \max_{j=1, \dots, n} (M_j - m_j).$$

Das werden wir in 12.5 und 12.6 ausnutzen.

12.5. Satz: Es gilt $C([a, b]) \subseteq R([a, b])$.

Beweis. Wir zeigen

$$\max_{j=1, \dots, n} (M_j - m_j) \rightarrow 0 \quad (n \rightarrow \infty),$$

dann folgt die Behauptung aus 12.4 und (***). Nach 9.15 existieren jeweils $y_j, \tilde{y}_j \in I_j$ mit $M_j - m_j = f(y_j) - f(\tilde{y}_j)$. Wegen $|y_j - \tilde{y}_j| \leq |I_j| = \frac{b-a}{n}$ reicht es somit zu zeigen, dass f gleichmäßig stetig ist (vgl. Satz von Heine, 9.4), also

$$\forall \varepsilon > 0 \exists \delta > 0 \forall y, \tilde{y} \in [a, b] : |y - \tilde{y}| < \delta \Rightarrow |f(y) - f(\tilde{y})| < \varepsilon.$$

Ist dies falsch, so existiert ein $\varepsilon_0 > 0$ so, dass zu jedem $m \in \mathbb{N}$ Punkte $y_m, \tilde{y}_m \in [a, b]$ existieren mit $|y_m - \tilde{y}_m| < 1/m$ und $|f(y_m) - f(\tilde{y}_m)| \geq \varepsilon_0$. Nach Bolzano-Weierstraß existiert eine konvergente Teilfolge $(y_{k(m)})$ von (y_m) . Sei

$$y_0 := \lim_{m \rightarrow \infty} y_{k(m)}; \text{ beachte } y_0 \in [a, b].$$

Dann gilt auch $\tilde{y}_{k(m)} \rightarrow y_0$ ($m \rightarrow \infty$). Da f in y_0 stetig ist, folgt

$$\varepsilon_0 \leq |f(y_{k(m)}) - f(\tilde{y}_{k(m)})| \rightarrow |f(y_0) - f(y_0)| = 0 \quad (m \rightarrow \infty),$$

ein Widerspruch. □

12.6. Satz: Ist $f : [a, b] \rightarrow \mathbb{R}$ monoton, so ist $f \in R([a, b])$.

Beweis. Sei f monoton wachsend (den Fall “ f fallend” beweist man analog). Wir verwenden wieder (***). Es gilt $M_j = f(x_j)$ und $m_j = f(x_{j-1})$, also

$$S_f(Z_n) - s_f(Z_n) = \frac{b-a}{n} \sum_{j=1}^n (f(x_j) - f(x_{j-1})) = \frac{b-a}{n} (f(b) - f(a)) \rightarrow 0 \quad (n \rightarrow \infty),$$

und die Behauptung folgt aus 12.4. □

Beispiele: (1) Sei $b > 0$ und $f : [0, b] \rightarrow \mathbb{R}$, $f(x) = x$. Nach 12.6 ist $f \in R([0, b])$. Wir berechnen das Integral: Sei $Z_n = \{x_0, \dots, x_n\}$ mit $x_j = jb/n$ ($j = 0, 1, \dots, n$). Dann gilt $M_j = f(x_j) = x_j = jb/n$, $m_j = (j-1)b/n$ und $|I_j| = b/n$, also

$$\begin{aligned} S_f(Z_n) &= \sum_{j=1}^n f(x_j)|I_j| = \sum_{j=1}^n j \frac{b^2}{n^2} = \frac{b^2}{n^2} \frac{n(n+1)}{2} \rightarrow \frac{b^2}{2} \quad (n \rightarrow \infty), \\ s_f(Z_n) &= \sum_{j=1}^n f(x_{j-1})|I_j| = \sum_{j=1}^n (j-1) \frac{b^2}{n^2} = \frac{b^2}{n^2} \frac{n(n-1)}{2} \rightarrow \frac{b^2}{2} \quad (n \rightarrow \infty), \end{aligned}$$

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 $x \mapsto \log x$ über $[a, b]$ integrierbar.

12.7. Riemann-Summen: Sei $f : [a, b] \rightarrow \mathbb{R}$ beschränkt und $Z = \{x_0, x_1, \dots, x_n\}$ eine Zerlegung, sowie I_j , M_j , m_j wie in 12.1. Dann heißt

$$\|Z\| := \max\{|I_j| : j = 1, 2, \dots, n\}$$

die *Feinheit* der Zerlegung Z .

Ein n -Tupel $\xi = (\xi_1, \dots, \xi_n)$ heißt *passender Zwischenvektor*, wenn $\xi_j \in I_j$ für jedes $j = 1, \dots, n$ gilt. Für einen solchen heißt

$$\sigma_f(Z, \xi) := \sum_{j=1}^n f(\xi_j)|I_j|$$

eine *Riemannsche Summe*. Wegen $m_j \leq f(\xi_j) \leq M_j$ gilt dabei $s_f(Z) \leq \sigma_f(Z, \xi) \leq S_f(Z)$.

Satz: Sei $f \in R([a, b])$ und $(Z_l)_{l \in \mathbb{N}}$ eine Folge von Zerlegungen mit $\|Z_l\| \rightarrow 0$, sowie $(\xi^{(l)})$ eine Folge von passenden Zwischenvektoren. Dann gilt

$$\lim_{l \rightarrow \infty} \sigma_f(Z_l, \xi^{(l)}) = \int_a^b f(x) \, dx.$$

Beweisidee. Wegen 12.4 reicht zu zeigen: $S_f(Z_l) - s_f(Z_l) \rightarrow 0$. □

12.8. Satz: Sei I ein Intervall und $f : I \rightarrow \mathbb{R}$ eine Funktion.

(1) Seien $a, b, c \in I$ mit $a < c < b$. Dann gilt

$$f \in R([a, b]) \Leftrightarrow f \in R([a, c]) \wedge f \in R([c, b]).$$

In diesem Fall ist

$$\int_a^b f dx = \int_a^c f dx + \int_c^b f dx.$$

(2) Sei $f \in R([a, b])$ ($[a, b] \subseteq I$), und sei $c \in I$. Setzt man

$$F : I \rightarrow \mathbb{R}, \quad F(y) := \int_c^y f(x) dx,$$

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 $\varepsilon > 0$. Dann existiert ein $Z \in \mathcal{Z}([a, b])$ mit $S_f(Z) - s_f(Z) < \varepsilon$. Wegen Satz 12.1(1) können wir davon ausgehen, dass $c \in Z$ gilt. Für $Z_1 := Z \cap [a, c]$ und $Z_2 := Z \cap [c, b]$ gilt dann :

$$S_f(Z_k) - s_f(Z_k) \leq S_f(Z) - s_f(Z) < \varepsilon \quad (k = 1, 2).$$

Nach 12.4 ist also $f \in R([a, c])$ und $f \in R([c, b])$.

Sei umgekehrt $f \in R([a, c])$ und $f \in R([c, b])$. Zu $\varepsilon > 0$ finden wir $Z_1 \in \mathcal{Z}([a, c])$ und $Z_2 \in \mathcal{Z}([c, b])$ 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 wieder mit 12.4. Weiter gilt:

$$\begin{aligned} \int_a^b f dx &= \sup\{s_f(Z) : Z \in \mathcal{Z}([a, b])\} = \sup\{s_f(Z) : Z \in \mathcal{Z}([a, b]) \wedge c \in Z\} \\ &= \sup\{s_f(Z_1) + s_f(Z_2) : Z_1 \in \mathcal{Z}([a, c]), Z_2 \in \mathcal{Z}([c, b])\} \\ &= \sup\{s_f(Z_1) : Z_1 \in \mathcal{Z}([a, c])\} + \sup\{s_f(Z_2) : Z_2 \in \mathcal{Z}([c, b])\} \\ &= \int_a^c f dx + \int_c^b f dx, \end{aligned}$$

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. $\square$

Bemerkung: Ist $f \in R([a, b])$ und $g : [a, b] \rightarrow \mathbb{R}$ so, dass die Menge

$$\{x \in [a, b] : f(x) \neq g(x)\}$$

endlich ist, so 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. [Übung: 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.]

Folgerung: Sei $g : [a, b] \rightarrow \mathbb{R}$ so, dass es eine Zerlegung $Z = \{x_0, x_1, \dots, x_n\} \in \mathcal{Z}([a, b])$ gibt, mit

1. $g : (x_{j-1}, x_j) \rightarrow \mathbb{R}$ ist stetig für $j = 1, 2, \dots, n$,
2. $\lim_{x \rightarrow x_j-} g(x)$ und $\lim_{x \rightarrow x_j+} g(x)$ existieren in $\mathbb{R}$ ($j = 1, \dots, n$) bzw. ($j = 0, \dots, n-1$).

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 $g : [x_{j-1}, x_j] \rightarrow \mathbb{R}$ durch ändern in den Randpunkten zu einer stetigen Funktion auf $[x_{j-1}, x_j]$ machen kann. Dann verwende 12.8.

Solche Funktionen $g : [a, b] \rightarrow \mathbb{R}$ heißen *stückweise stetig*. Sie spielen in der Regelungstechnik z.B. als “Steuerfunktionen” eine Rolle.

12.9. 1. Hauptsatz der Differential- und Integralrechnung:

Definition Sei $I \subseteq \mathbb{R}$ ein Intervall. Sind $f, F : I \rightarrow \mathbb{R}$ Funktionen, wobei F auf I differenzierbar ist mit $F' = f$ auf I , so heißt F eine *Stammfunktion von f auf I* .

Bemerkungen:

- (1) Sind F_1, F_2 Stammfunktionen von f auf I , so gilt $F_1' = f = F_2'$ auf I . Damit folgt

$$\exists c \in \mathbb{R} \forall x \in I : F_1(x) = F_2(x) + c.$$

- (2) Für eine Stammfunktion von f schreibt man auch

$$\int f(x) dx \quad (\text{unbestimmtes Integral}),$$

also z.B.

$$\int x^2 dx = \frac{1}{3} x^3 \text{ auf } \mathbb{R}, \quad \int x^2 dx = \frac{1}{3} x^3 + 17 \text{ auf } \mathbb{R}, \quad \int \frac{1}{x} dx = \log x \text{ auf } (0, \infty).$$

(3) Es gibt integrierbare Funktionen die keine Stammfunktion besitzen: Übung: Es sei $f : [0, 1] \rightarrow \mathbb{R}$ definiert durch

$$f(x) = \begin{cases} 0, & x = 0 \\ 1, & x \in (0, 1] \end{cases}$$

Dann ist $f \in R([0, 1])$ aber f hat keine Stammfunktion auf $[0, 1]$.

(4) Es gibt Funktionen die eine Stammfunktion besitzen aber nicht integrierbar sind: Übung: Die Funktion $F : [0, 1] \rightarrow \mathbb{R}$,

$$F(x) := \begin{cases} x^{\frac{3}{2}} \sin(1/x), & x \in (0, 1] \\ 0, & x = 0 \end{cases}$$

ist differenzierbar auf $[0, 1]$, also eine Stammfunktion ihrer Ableitung $f := F'$. Aber f ist unbeschränkt auf $[0, 1]$. Also ist $f \notin R([0, 1])$. D.h.

$$\int f(x)dx = F(x) \text{ auf } [0, 1], \text{ aber } \int_0^1 f(x)dx \text{ existiert nicht!}$$

1. Hauptsatz: Es sei $f \in R([a, b])$ und f besitze auf $[a, b]$ eine Stammfunktion F . Dann gilt

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

Beweis. Es sei $Z = \{x_0, \dots, x_n\} \in \mathcal{Z}$ und I_j, m_j, M_j wie in 12.1. Mit dem Mittelwertsatz folgt für jedes $j = 1, \dots, n$:

$$\exists \xi_j \in I_j : F(x_j) - F(x_{j-1}) = f(\xi_j)(x_j - x_{j-1}) = f(\xi_j)|I_j|.$$

Summation über j liefert (beachte $m_j \leq f(\xi_j) \leq M_j$):

$$s_f(Z) \leq \sum_{j=1}^n f(\xi_j)|I_j| = \sum_{j=1}^n (F(x_j) - F(x_{j-1})) = F(b) - F(a) \leq S_f(Z).$$

Also gilt

$$\forall Z \in \mathcal{Z} : s_f(Z) \leq F(b) - F(a) \leq S_f(Z).$$

Wegen $f \in R([a, b])$ folgt

$$\int_a^b f(x)dx = s_f \leq F(b) - F(a) \leq S_f = \int_a^b f(x)dx.$$

□

Beispiele: Es gilt

$$\int_0^\pi \cos x dx = [\sin x]_0^\pi = 0, \quad \int_a^b x^n dx = \left[\frac{x^{n+1}}{n+1} \right]_a^b = \frac{b^{n+1} - a^{n+1}}{n+1}.$$

12.10. 2. Hauptsatz der Differential- und Integralrechnung: Es sei $f \in R([a, b])$ und $F : [a, b] \rightarrow \mathbb{R}$ definiert durch

$$F(x) := \int_a^x f(t)dt.$$

Dann gilt:

(1) F ist *Lipschitz-stetig*, d.h. es gibt ein $L \geq 0$ (eine sog. *Lipschitz-Konstante*) mit

$$\forall x, y \in [a, b] : |F(y) - F(x)| \leq L|y - x|.$$

Insbesondere ist $F \in C([a, b])$.

(2) Ist $f \in C([a, b])$, so ist F eine Stammfunktion von f auf $[a, b]$. Insbesondere ist dann $F \in C^1([a, b])$.

Beweis. (1) Es existiert ein $L \geq 0$ mit $|f(x)| \leq L$ ($x \in [a, b]$). Seien $x, y \in [a, b]$. Mit 12.8 und 12.13(2) folgt

$$|F(y) - F(x)| = \left| \int_x^y f(t)dt \right| \leq \left| \int_x^y |f(t)|dt \right| \leq \left| \int_x^y Ldt \right| = L|y - x|.$$

Insbesondere gilt $F(y) \rightarrow F(x)$ ($y \rightarrow x$).

(2) Wir zeigen für $x \in [a, b]$:

$$\lim_{h \rightarrow 0+} \frac{F(x+h) - F(x)}{h} = f(x),$$

(analog zeigt man für $x \in (a, b]$: $\lim_{h \rightarrow 0-} \frac{F(x+h) - F(x)}{h} = f(x)$).

Sei also $x \in [a, b]$ fest, $h > 0$ und $x+h \in [a, b]$. Es ist

$$\frac{1}{h} \int_x^{x+h} f(t)dt = f(x), \quad \frac{F(x+h) - F(x)}{h} = \frac{1}{h} \int_x^{x+h} f(t)dt.$$

Somit gilt:

$$\begin{aligned} D(h) &:= \left| \frac{F(x+h) - F(x)}{h} - f(x) \right| = \frac{1}{h} \left| \int_x^{x+h} (f(t) - f(x))dt \right| \\ &\leq \frac{1}{h} \int_x^{x+h} |f(t) - f(x)|dt. \end{aligned}$$

Die Funktion $t \mapsto |f(t) - f(x)|$ ist stetig auf $[a, b]$. Nach 9.15 gilt:

$$\exists \xi_h \in [x, x+h] \quad \forall t \in [x, x+h] : |f(t) - f(x)| \leq |f(\xi_h) - f(x)|.$$

Also gilt:

$$D(h) \leq \frac{1}{h} \int_x^{x+h} |f(\xi_h) - f(x)|dt = |f(\xi_h) - f(x)|.$$

Für $h \rightarrow 0+$ gilt $\xi_h \rightarrow x$. Da f in x stetig ist folgt $f(\xi_h) \rightarrow f(x)$ ($h \rightarrow 0+$), also $D(h) \rightarrow 0$ ($h \rightarrow 0+$). $\square$

Folgerung: Ist $I \subseteq \mathbb{R}$ ein Intervall, $f \in C(I)$, $c \in I$ und

$$F : I \rightarrow \mathbb{R}, \quad F(x) := \int_c^x f(t) dt,$$

so ist F eine Stammfunktion von f auf I .

12.11. 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 1. Hauptsatz 12.9). $\square$

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)

$$\int \log x \, dx = \int \underbrace{1}_{f'} \cdot \underbrace{\log x}_g \, dx = \underbrace{x}_f \underbrace{\log x}_g - \int \underbrace{x}_f \underbrace{\frac{1}{x}}_{g'} \, dx = x \log x - x.$$

12.12. Integration durch Substitution: Es seien $I, J \subseteq \mathbb{R}$ Intervalle, $f \in C(I)$ 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 $g'(t) \neq 0$ ($t \in J$), 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 $I = [a, b]$, $J = [\min\{\alpha, \beta\}, \max\{\alpha, \beta\}]$ und $g(\alpha) = a$, $g(\beta) = b$, 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 $G' = (F \circ g)' = (F' \circ g) \cdot g' = (f \circ g) \cdot g'$). 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 $g^{-1} : g(J) \rightarrow J$ (beachte, 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 1. Hauptsatz:

$$\int_{\alpha}^{\beta} f(g(t))g'(t) dt = G(\beta) - G(\alpha) = F(b) - F(a) = \int_a^b f(x) dx.$$

□

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 $x = g(t)$, d.h. 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 $x = \log t$. Dann ist $dx = dt/t$ und aus “ x von 0 bis 1” ergibt sich “ t von 1 bis e ”. Wir erhalten:

$$\int_0^1 \frac{e^{2x}+1}{e^x} dx = \int_1^e (1+t^{-2}) dt = (t - t^{-1})|_1^e = e - \frac{1}{e}.$$

(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 von 0 bis 1” ergibt sich “ t von 0 bis $\frac{\pi}{2}$ ”. Wir erhalten

$$\int_0^1 \sqrt{1-x^2} dx = \int_0^{\frac{\pi}{2}} \sqrt{1-\sin^2 t} dt = \int_0^{\frac{\pi}{2}} \sqrt{\cos^2 t} \cos t dt = \int_0^{\frac{\pi}{2}} \cos^2 t dt,$$

da $\cos \geq 0$ auf $[0, \frac{\pi}{2}]$ ist. Wegen

$$\cos(2t) = \cos^2 t - \sin^2 t = \cos^2 t - (1 - \cos^2 t) = 2 \cos^2 t - 1$$

gilt $\cos^2 t = \frac{1}{2}(1 + \cos(2t))$, also

$$\int_0^{\frac{\pi}{2}} \cos^2 t dt = \frac{1}{2} \left(\int_0^{\frac{\pi}{2}} 1 dt + \underbrace{\int_0^{\frac{\pi}{2}} \cos(2t) dt}_{=0} \right) = \frac{\pi}{4}.$$

Der Flächeninhalt eines Kreises mit Radius 1 ist also π .

(3) $\int x e^{-x^2} dx$, substituiere $u = x^2$, also $2x dx = du$ bzw. $x dx = \frac{du}{2}$. Dann ist

$$\int x e^{-x^2} dx = \frac{1}{2} \int e^{-u} du \Big|_{u=x^2} = -\frac{1}{2} e^{-u} \Big|_{u=x^2} = -\frac{1}{2} e^{-x^2}.$$

Mittels partieller Integration folgt daraus z.B.:

$$\int x^3 e^{-x^2} dx = \int \underbrace{x^2}_g \cdot \underbrace{x e^{-x^2}}_{f'} dx = -\frac{x^2}{2} e^{-x^2} + \int x e^{-x^2} dx = -\frac{x^2}{2} e^{-x^2} - \frac{1}{2} e^{-x^2}.$$

Man kann auf diese Weise Stammfunktionen von $x^n e^{-x^2}$ bestimmen, falls $n \in \mathbb{N}$ ungerade ist. 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 $x = \log y$ 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 sogenannten *Partialbruchzerlegung*. Wir haben also

$$\int \frac{1}{(1+y)y} dy \Big|_{y=e^x} = \int \frac{1}{y} - \frac{1}{1+y} dy \Big|_{y=e^x} = (\log y - \log(1+y)) \Big|_{y=e^x} = x - \log(1+e^x).$$

12.13. Satz: Seien $f, g \in R([a, b])$ und $D := f([a, b])$.

(1) Sei $L \geq 0$ und $h : D \rightarrow \mathbb{R}$ eine Funktion mit

$$\forall s, t \in D : |h(t) - h(s)| \leq L|t - s| \quad (\text{L})$$

(h ist also eine Lipschitz-stetige Funktion). Dann gilt $h \circ f \in R([a, b])$.

(2) Es gilt $|f| \in R([a, b])$ und

$$\left| \int_a^b f dx \right| \leq \int_a^b |f| dx \quad (\text{Dreiecksungleichung für Integrale}).$$

(3) Es ist $f \cdot g \in R([a, b])$.

(4) Ist $c > 0$ und gilt $|f(x)| \geq c$ ($x \in [a, b]$), so ist $\frac{1}{f} \in R([a, b])$.

Bemerkungen:

(1) Vorsicht: Bei Komposition integrierbarer Funktionen können nicht integrierbare Funktionen entstehen.

(2) Im folgenden Beweis benutzen wir: Ist $A \subseteq \mathbb{R}$ beschränkt und nicht leer, so gilt

$$\sup A - \inf A = \sup\{a - \tilde{a} : a, \tilde{a} \in A\} = \sup\{|a - \tilde{a}| : a, \tilde{a} \in A\}.$$

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

$$\widetilde{M}_j := \sup(h \circ f)(I_j) \quad \text{und} \quad \widetilde{m}_j := \inf(h \circ f)(I_j).$$

Dann gilt

$$\widetilde{M}_j - \widetilde{m}_j = \sup\{\underbrace{|h(f(x)) - h(f(\tilde{x}))|}_{\leq L|f(x) - f(\tilde{x})|} : x, \tilde{x} \in I_j\} \leq L(M_j - m_j).$$

Wir erhalten

$$\forall Z \in \mathcal{Z} : S_{h \circ f}(Z) - s_{h \circ f}(Z) \leq L(S_f(Z) - s_f(Z)),$$

und $h \circ f \in R([a, b])$ folgt aus 12.4.

(2) Wende (1) an auf $h(t) = |t|$. Es gilt $||t| - |s|| \leq |t - s|$ ($t, s \in \mathbb{R}$), d.h. (L) gilt mit $L = 1$. Also ist $|f| \in R([a, b])$. Weiter ist

$$\left| \int_a^b f \, dx \right| = \max \left\{ \int_a^b \underbrace{f}_{\leq |f|} \, dx, \int_a^b \underbrace{(-f)}_{\leq |f|} \, dx \right\} \leq \int_a^b |f| \, dx.$$

(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 $|f(x)| \leq \gamma$ ($x \in [a, b]$). Also gilt $D \subseteq [-\gamma, \gamma]$. Für $s, t \in D$ gilt nun:

$$|t^2 - s^2| = |t + s||t - s| \leq (|t| + |s|)|t - s| \leq 2\gamma|t - s|.$$

Mit $h(t) = t^2$ in (1) folgt $f^2 \in R([a, b])$.

(4) Nach Voraussetzung gilt $D \subseteq (-\infty, -c] \cup [c, \infty)$. Für $t, s \in D$ gilt somit:

$$\left| \frac{1}{t} - \frac{1}{s} \right| = \frac{|s - t|}{|s||t|} \leq \frac{1}{c^2}|t - s|.$$

Mit $h(t) := 1/t$ in (1) folgt die Behauptung. □

Wir erinnern an den Begriff der gleichmäßigen Konvergenz aus 9.9: Eine Folge (f_n) von Funktionen $f_n : [a, b] \rightarrow \mathbb{R}$ ist auf $[a, b]$ gleichmäßig konvergent gegen eine Funktion $f : [a, b] \rightarrow \mathbb{R}$, falls es eine Nullfolge (α_n) gibt, mit:

$$\forall n \in \mathbb{N} \, \forall x \in [a, b] : |f_n(x) - f(x)| \leq \alpha_n.$$

12.14. Konvergenzsatz für Integrale: Sei $f : [a, b] \rightarrow \mathbb{R}$ eine Funktion und (f_n) eine Folge in $R([a, b])$, die auf $[a, b]$ gleichmäßig gegen f konvergiert. Dann gilt $f \in R([a, b])$ und

$$\lim_{n \rightarrow \infty} \int_a^b f_n(x) dx = \int_a^b f(x) dx \quad \left[= \int_a^b \lim_{n \rightarrow \infty} f_n(x) dx \right].$$

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 gilt für jedes $j \in \{1, \dots, n\}$:

$$M_j - m_j = \sup\{|f(x) - f(\tilde{x})| : x, \tilde{x} \in I_j\}.$$

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

$$M_j - m_j \leq 2\delta + \sup\{|f_n(x) - f_n(\tilde{x})| : x, \tilde{x} \in I_j\},$$

also

$$S_f(Z) - s_f(Z) \leq 2\delta(b-a) + S_{f_n}(Z) - s_{f_n}(Z) = \frac{2\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.$$

Damit folgt $f \in R([a, b])$ aus 12.4. Wegen 12.13(2) gilt

$$\left| \int_a^b f_n(x) dx - \int_a^b f(x) dx \right| \leq \int_a^b \underbrace{|f_n(x) - f(x)|}_{\leq \alpha_n} dx \leq (b-a)\alpha_n \rightarrow 0 \quad (n \rightarrow \infty).$$

□

Beispiele: (1) Sei $b > 0$ und (f_n) gegeben durch $f_n(x) := e^{-x^2/n}$ ($x \in [-b, b]$). Dann konvergiert (f_n) auf $[-b, b]$ punktweise gegen $f : [-b, b] \rightarrow \mathbb{R}$, $f(x) = 1$. Für jedes $x \in [-b, b]$ und $n \in \mathbb{N}$ gilt

$$|f_n(x) - f(x)| = 1 - e^{-x^2/n} \leq 1 - e^{-b^2/n} =: \alpha_n.$$

Wegen $\alpha_n \rightarrow 0$ konvergiert (f_n) auf $[-b, b]$ gleichmäßig gegen f , also

$$\int_{-b}^b e^{-x^2/n} dx \longrightarrow \int_{-b}^b 1 dx = 2b \quad (n \rightarrow \infty).$$

(2) Sei $f_n : [0, 2] \rightarrow \mathbb{R}$ 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

$$\int_0^2 f_n(x) dx = 1 \quad (n \in \mathbb{N}),$$

und für jedes $x \in [0, 2]$ konvergiert die Folge $(f_n(x))$ gegen 0. Setzt man $f : [0, 2] \rightarrow \mathbb{R}$, $f(x) = 0$, so konvergiert die Folge (f_n) auf $[0, 2]$ *punktweise* gegen f . Also gilt

$$\int_0^2 f_n(x) dx = 1 \quad \underbrace{\rightarrow}_{n \rightarrow \infty} \quad 1 \neq 0 = \int_0^2 f(x) dx.$$

Insbesondere konvergiert (f_n) also auf $[0, 2]$ *nicht gleichmäßig* gegen f (das kann man natürlich auch direkt einsehen). Dieses Beispiel zeigt, dass Satz 12.14 bei punktweiser Konvergenz im allgemeinen falsch ist.

(3) Für $n \in \mathbb{N}$ sei $f_n : [0, \pi] \rightarrow \mathbb{R}$, $f_n(x) = \sin(nx)$. Dann gilt

$$(f_n(\frac{\pi}{2})) = (1, 0, -1, 0, 1, 0, -1, 0, \dots),$$

insbesondere ist (f_n) auf $[0, \pi]$ nicht punktweise konvergent. Trotzdem gilt

$$\int_0^\pi f_n(x) dx = \int_0^\pi \sin(nx) dx = \frac{1}{n} \left[-\cos(nx) \right]_0^\pi = \frac{1}{n} (1 - (-1)^n) \rightarrow 0 \quad (n \rightarrow \infty).$$

Aus Konvergenz der Integrale kann also nicht auf punktweise Konvergenz der Integranden geschlossen werden.

13 Differentialgleichungen

13.1. Differentialgleichungen erster Ordnung: Sei $D \subseteq \mathbb{R} \times \mathbb{R}$ und $f : D \rightarrow \mathbb{R}$ eine Funktion. Wir betrachten Differentialgleichungen der Form

$$y'(x) = f(x, y(x)) \quad (\text{kurz } y' = f(x, y)), \quad (\text{DGL})$$

und für gegebene $(x_0, y_0) \in D$ das **Anfangswertproblem:**

$$\begin{aligned} y'(x) &= f(x, y(x)), \\ y(x_0) &= y_0. \end{aligned} \quad (\text{AWP})$$

Eine *Lösung von* (DGL) ist eine differenzierbare Funktion $y : I \rightarrow \mathbb{R}$, wobei $I \subseteq \mathbb{R}$ ein Intervall ist, mit $(x, y(x)) \in D$ ($x \in I$) und

$$y'(x) = f(x, y(x)) \quad (x \in I).$$

Eine *Lösung von* (AWP) ist eine Lösung $y : I \rightarrow \mathbb{R}$ von (DGL) mit $x_0 \in I$ und $y(x_0) = y_0$.

Ein Anfangswertproblem (AWP), welches Lösungen besitzt, heißt *eindeutig lösbar*, wenn für je zwei Lösungen $y_j : I_j \rightarrow \mathbb{R}$, $j = 1, 2$, von (AWP) gilt

$$y_1(x) = y_2(x) \quad (x \in I_1 \cap I_2).$$

Ist (AWP) eindeutig lösbar, so gibt es genau eine Lösung $y_{\max} : I_{\max} \rightarrow \mathbb{R}$ mit maximalem Existenzintervall $I_{\max}$. Diese heißt die *maximale Lösung* von (AWP) (jede Lösung von (AWP) erhält man dann als Einschränkung von $y_{\max}$ auf ein Teilintervall $I \subseteq I_{\max}$ mit $x_0 \in I$). Insbesondere kann, falls (AWP) für jedes $(x_0, y_0) \in D$ eindeutig lösbar ist, jede Lösung von (DGL) eindeutig auf ein maximales Existenzintervall fortgesetzt werden.

Beispiel: $y' = \lambda y$, $y(x_0) = y_0$ ist für alle $x_0, y_0 \in \mathbb{R}$ eindeutig lösbar und die maximale Lösung ist gegeben durch $y(x) = y_0 e^{\lambda(x-x_0)}$ ($x \in \mathbb{R}$) (vgl. 11.8).

Bedeutung von (DGL) und (AWP): Häufig ist x hier ein Zeitparameter und $y = y(x)$ beschreibt einen Zustand zur Zeit x . (DGL) besagt, dass die Zustandsänderung (zum Zeitpunkt x) abhängt vom gegebenen Zeitpunkt x und dem gegebenen Zustand $y(x)$. Ist (AWP) eindeutig lösbar, so legen Anfangswert $y(x_0) = y_0$ und Systembeschreibung (DGL) den zeitlichen Zustandsverlauf in Vergangenheit und Zukunft fest.

13.2. Lineare Differentialgleichungen erster Ordnung: Sei $J \subseteq \mathbb{R}$ ein Intervall und $a, b \in C(J)$. Dann heißt

$$y' = a(x)y + b(x) \quad (\text{L1})$$

eine *Lineare Differentialgleichung (1.Ordnung)* und

$$y' = a(x)y. \quad (\text{H1})$$

heißt die zugehörige *homogene Gleichung*.

Es sei

$$\begin{aligned}\mathcal{L} &:= \{y : J \rightarrow \mathbb{R} : y \text{ ist Lösung von (L1)}\}, \\ \mathcal{L}_h &:= \{y : J \rightarrow \mathbb{R} : y \text{ ist Lösung von (H1)}\}.\end{aligned}$$

Satz: Es sei A eine Stammfunktion von a auf J . Dann gilt:

(1)

$$y \in \mathcal{L}_h \Leftrightarrow \exists c \in \mathbb{R} : y(x) = ce^{A(x)}.$$

(2) Sei $y_p \in \mathcal{L}$ (eine spezielle Lösung). Dann gilt:

$$y \in \mathcal{L} \Leftrightarrow \exists y_h \in \mathcal{L}_h : y = y_p + y_h.$$

(3) Es sei $x_0 \in J$ und $y_0 \in \mathbb{R}$. Dann ist das Anfangswertproblem

$$y' = a(x)y + b(x), \quad y(x_0) = y_0$$

eindeutig lösbar und $I_{\max} = J$.

Bemerkung: *Variation der Konstanten:* Der Ansatz

$$y_p(x) = c(x)e^{A(x)}$$

mit einer noch unbekannten Funktion $x \mapsto c(x)$ führt stets auf eine spezielle Lösung y_p (vgl. Beispiele).

Beweis. (1) Ist $c \in \mathbb{R}$ und $y(x) = ce^{A(x)}$, so gilt:

$$y'(x) = cA'(x)e^{A(x)} = a(x)ce^{A(x)} = a(x)y(x) \quad (x \in J),$$

also $y \in \mathcal{L}_h$. Ist umgekehrt $y \in \mathcal{L}_h$, so gilt für

$$\phi(x) := e^{-A(x)}y(x) \quad (x \in J) :$$

$$\phi'(x) = -A'(x)e^{-A(x)}y(x) + e^{-A(x)}y'(x) = -a(x)e^{-A(x)}y(x) + e^{-A(x)}a(x)y(x) = 0.$$

Somit gilt:

$$\exists c \in \mathbb{R} \forall x \in J : \phi(x) = c,$$

also

$$y(x) = ce^{A(x)} \quad (x \in J).$$

(2) Ist $y = y_p + y_h$ mit $y_h \in \mathcal{L}_h$, so gilt:

$$\begin{aligned}y'(x) &= y_p'(x) + y_h'(x) = a(x)y_p(x) + b(x) + a(x)y_h(x) \\ &= a(x)(y_p(x) + y_h(x)) + b(x) = a(x)y(x) + b(x).\end{aligned}$$

Ist umgekehrt $y \in \mathcal{L}$, so gilt für $y_h(x) := y(x) - y_p(x)$:

$$\begin{aligned} y_h'(x) &= y'(x) - y_p'(x) = (a(x)y(x) + b(x)) - (a(x)y_p(x) + b(x)) \\ &= a(x)(y(x) - y_p(x)) = a(x)y_h(x). \end{aligned}$$

Also ist $y_h \in \mathcal{L}_h$.

(3) Die allgemeine Lösung von (L1) ist $y(x) = y_p(x) + ce^{A(x)}$.

$$y_0 = y(x_0) = y_p(x_0) + ce^{A(x_0)} \iff c = (y_0 - y_p(x_0))e^{-A(x_0)}.$$

Das Anfangswertproblem hat also auf J genau eine Lösung. Betrachtet man (L1) und (H1) auf einem Teilintervall $I \subseteq J$ mit $x_0 \in I$, so ergibt sich als eindeutige Lösung des Anfangswertproblems auf I die Einschränkung dieser Lösung auf I . Damit ist das AWP eindeutig lösbar und $I_{\max} = J$. $\square$

Beispiel: $y' = -(\sin x)y + \sin^3 x$. Hier ist $J = \mathbb{R}$. Wähle $A(x) = \cos x$. Die Lösungen der homogenen Gleichung sind gegeben durch $y(x) = ce^{\cos x}$, $c \in \mathbb{R}$. Eine spezielle Lösung der inhomogenen Gleichung erhalten wir durch **Variation der Konstanten**. Ansatz: $y_p(x) = c(x)e^{\cos x}$. Es gilt $y_p \in \mathcal{L}$ genau dann wenn gilt

$$\begin{aligned} y_p'(x) &= c'(x)e^{\cos x} - c(x)e^{\cos x} \sin x = -\sin x(c(x)e^{\cos x}) + \sin^3 x \\ &\iff c'(x) = \sin^3 x e^{-\cos x}. \end{aligned}$$

Z.B. mit partieller Integration erhält man die Stammfunktion

$$c(x) = -e^{-\cos x}(1 + \cos x)^2.$$

Die allgemeine Lösung der Differentialgleichung ist also gegeben durch

$$y(x) = -(1 + \cos x)^2 + ce^{\cos x} \quad (x \in \mathbb{R}),$$

wobei $c \in \mathbb{R}$ eine Konstante ist.

Will man z.B. das Anfangswertproblem $y' = -(\sin x)y + \sin^3 x$, $y(0) = 1$ lösen, so führt die Anfangsbedingung auf $1 = -(1 + 1)^2 + ce$, also $c = 5e^{-1}$. Die (maximale) Lösung dieses Anfangswertproblems ist also

$$y(x) = -(1 + \cos x)^2 + 5e^{-1+\cos x} \quad (x \in \mathbb{R}).$$

13.3. Trennung der Variablen: Seien $I, J \subseteq \mathbb{R}$ Intervalle und $f : I \rightarrow \mathbb{R}$, $g : J \rightarrow \mathbb{R}$ stetig. Eine Differentialgleichung der Form

$$y' = f(x)g(y) \tag{TdV-1}$$

heißt Differentialgleichung *mit getrennten Variablen* (oder *Veränderlichen*). Das Anfangswertproblem

$$\begin{aligned} y' &= f(x)g(y), \\ y(x_0) &= y_0. \end{aligned} \quad (\text{TdV-2})$$

mit $x_0 \in I$, $y_0 \in J$ behandelt man wie folgt:

- 1. Fall** $g(y_0) = 0$: Eine Lösung ist gegeben durch $y(x) = y_0$ ($x \in I$).
- 2. Fall** $g(y_0) \neq 0$: Sei $\tilde{J} \subseteq J$ das größte Teilintervall von J mit $y_0 \in \tilde{J}$ und $g \neq 0$ auf $\tilde{J}$. Ist $y : \tilde{I} \rightarrow \mathbb{R}$ eine Lösung von (TdV-2) mit $y(\tilde{I}) \subseteq \tilde{J}$, also mit $g(y(x)) \neq 0$ ($x \in \tilde{I}$), so gilt

$$\frac{y'(x)}{g(y(x))} = f(x) \quad (x \in \tilde{I}),$$

und mittels Substitution $\eta = y(t)$, $d\eta = y'(t) dt$ folgt:

$$\int_{y_0}^{y(x)} \frac{1}{g(\eta)} d\eta = \int_{x_0}^x \frac{y'(t)}{g(y(t))} dt = \int_{x_0}^x f(t) dt \quad (x \in \tilde{I}).$$

Sei F Stammfunktion von f auf I und G Stammfunktion von $1/g$ auf $\tilde{J}$. Dann gilt für $x \in \tilde{I}$:

$$G(y(x)) - G(y_0) = F(x) - F(x_0), \quad \text{also} \quad G(y(x)) = F(x) - F(x_0) + G(y_0).$$

Wegen $G' = 1/g \neq 0$ auf $\tilde{J}$ und $1/g \in C(\tilde{J})$ ist G auf $\tilde{J}$ streng monoton und besitzt also auf $G(\tilde{J})$ eine Umkehrfunktion G^{-1} . Wir erhalten

$$y(x) = G^{-1}(F(x) - F(x_0) + G(y_0)) \quad (x \in \tilde{I}). \quad (\text{TdV-3})$$

Insbesondere ist die Lösung im Falle der Existenz auf $\tilde{I}$ eindeutig.

Umgekehrt definiert (TdV-3) auch eine Lösung von (TdV-2), wenn man $\tilde{I}$ als ein Teilintervall von I nimmt mit $x_0 \in \tilde{I}$ und $F(x) - F(x_0) + G(y_0) \in G(\tilde{J})$ ($x \in \tilde{I}$): Dann ist

$$y(x_0) = G^{-1}(F(x_0) - F(x_0) + G(y_0)) = G^{-1}(G(y_0)) = y_0$$

und für $x \in \tilde{I}$ gilt

$$\begin{aligned} y'(x) &= (G^{-1})'(F(x) - F(x_0) + G(y_0)) \cdot F'(x) \\ &= \frac{F'(x)}{G'(G^{-1}(F(x) - F(x_0) + G(y_0)))} = \frac{F'(x)}{G'(y(x))} = f(x)g(y(x)). \end{aligned}$$

Beispiele: 1.) $y' = \sqrt{|y|}$, $y(x_0) = y_0$, mit $x_0, y_0 \in \mathbb{R}$. Hier ist $f(x) = 1$ auf $I = \mathbb{R}$ und $g(y) = \sqrt{|y|}$ auf $J = \mathbb{R}$.

Für $y_0 = 0$ ist $y(x) = 0$ ($x \in \mathbb{R}$) eine Lösung.

Wir betrachten nun $y_0 > 0$. Dann ist $\tilde{J} = (0, \infty)$. Auf $\tilde{J}$ gilt $G(\eta) = \int 1/\sqrt{|\eta|} d\eta = 2\sqrt{\eta}$, also $G(\tilde{J}) = (0, \infty)$. Somit haben wir

$$F(x) - F(x_0) = x - x_0, \quad G(y) - G(y_0) = 2\sqrt{y} - 2\sqrt{y_0}$$

und

$$F(x) - F(x_0) + G(y_0) = x - x_0 + 2\sqrt{y_0} > 0 \iff x > x_0 - 2\sqrt{y_0}.$$

Folglich nehmen wir $\tilde{I} = (x_0 - 2\sqrt{y_0}, \infty)$ und erhalten als eindeutige Lösung auf diesem Intervall

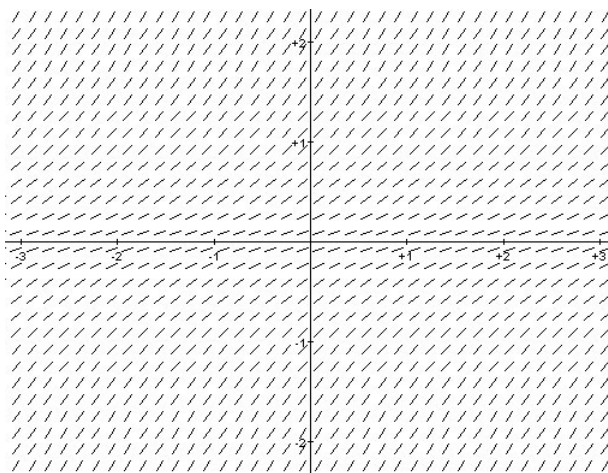
$$y(x) = G^{-1}(x - x_0 + 2\sqrt{y_0}) = \frac{(x - x_0 + 2\sqrt{y_0})^2}{4} \quad (x > x_0 - 2\sqrt{y_0}).$$

Für $x \leq c := x_0 - 2\sqrt{y_0}$ können wir die Lösung fortsetzen durch $y(x) = 0$ oder durch $y(x) = -\frac{(x-c)^2}{4}$ oder für ein gewähltes $b < c$ durch

$$y(x) = \begin{cases} 0, & x \in [b, c] \\ -\frac{(x-b)^2}{4}, & x < b \end{cases}.$$

Das Anfangswertproblem $y' = \sqrt{|y|}$, $y(x_0) = y_0$ hat also Lösungen auf $\mathbb{R}$. Diese sind nur auf $[x_0 - 2\sqrt{y_0}, \infty)$ eindeutig, nicht jedoch auf $\mathbb{R}$. Sie verzweigen sich im Punkt $(x_0 - 2\sqrt{y_0}, 0)$. Insbesondere sehen wir, dass das Anfangswertproblem mit $y(x_0) = 0$ auf keinem Intervall eindeutig lösbar ist.

Bemerkungen: (1) Man kann sich Differentialgleichungen der Form $y' = f(x, y)$ auch durch das **Richtungsfeld** veranschaulichen: In jedem Punkt (x, y) gibt $f(x, y)$ die durch die Gleichung gegebene **Steigung** in diesem Punkt an.



Richtungsfeld von $y' = \sqrt{|y|}$.

(2) In HM III werden wir sehen: Ist $f \in C(I)$ und $g \in C^1(J)$, so ist (TdV-2) für jedes $(x_0, y_0) \in I \times J$ eindeutig lösbar.

2.) $y' = e^y \sin x$ (zunächst ohne Anfangswerte). Hier ist $I = J = \mathbb{R}$, $f(x) = \sin x$ und $g(y) = e^y \neq 0$ ($y \in \mathbb{R}$), also $\tilde{J} = \mathbb{R}$ in jedem Fall. Trennung der Veränderlichen führt auf

$$-e^{-y} = \int e^{-y} dy = \int \sin x dx = -\cos x - c,$$

d.h. $e^{-y} = \cos x + c$ und

$$y(x) = -\log(\cos x + c), \quad \text{wobei } \cos x + c > 0.$$

Berücksichtigen wir den Anfangswert $y(x_0) = y_0$ und setzen $F(x) = -\cos x$, $G(y) = -e^{-y}$, so ist $G(\tilde{J}) = G(\mathbb{R}) = (-\infty, 0)$ und $G^{-1} : (-\infty, 0) \rightarrow \mathbb{R}$, $s \mapsto -\log(-s)$. Wir erhalten

$$y(x) = -\log(\cos x - \cos x_0 + e^{-y_0}) \quad (x \in \tilde{I}),$$

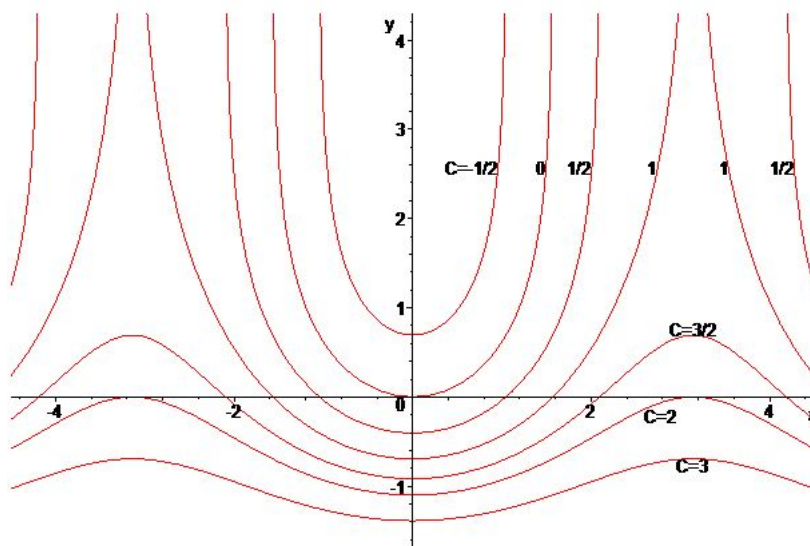
wobei $\tilde{I}$ ein Intervall ist mit $x_0 \in \tilde{I}$ und $-\cos x + \cos x_0 - e^{y_0} \in G(\mathbb{R})$, also mit

$$\cos x - \cos x_0 + e^{-y_0} > 0 \quad (x \in \tilde{I}).$$

Durch Vergleich sehen wir $c = e^{-y_0} - \cos x_0$, also insbesondere $c > -1$.

Für $c > 1$ kann man $\tilde{I} = \mathbb{R}$ setzen. Für $c \in (-1, 1]$ hingegen ist das Existenzintervall der Lösung beschränkt. Für $x_0 = 0$ ist etwa $c \leq 1$ genau dann, wenn $y_0 \geq -\log 2$. Für $y_0 \geq -\log 2$ ist dann das maximale Existenzintervall gegeben durch

$$\tilde{I} = \{x \in (-\pi, \pi) : |x| < \arccos(1 - e^{-y_0})\}.$$



$$y(x) = -\log(\cos x + c) \quad \text{für } c + \cos x > 0.$$

13.4. Linearen Differentialgleichungen zweiter Ordnung: Sei $J \subseteq \mathbb{R}$ ein Intervall und $a_1, a_0, b \in C(J)$.

Dann heit

$$y'' + a_1(x)y' + a_0(x)y = b(x) \quad (\text{L2})$$

eine *Lineare Differentialgleichung (2.Ordnung)* und

$$y'' + a_1(x)y' + a_0(x)y = 0 \quad (\text{H2})$$

heit die zugehrige *homogene Gleichung*.

Es sei

$$\mathcal{L} := \{y : J \rightarrow \mathbb{R} : y \text{ ist Lsung von (L2)}\},$$

$$\mathcal{L}_h := \{y : J \rightarrow \mathbb{R} : y \text{ ist Lsung von (H2)}\}.$$

Beachte: Lsungen sind nun Funktionen in $y \in C^2(J)$, welche der jeweiligen Differentialgleichung gengen.

Satz: Es gilt:

(1) Sind $u, v : J \rightarrow \mathbb{R}$ Lsungen von (H2) mit der Eigenschaft

$$\alpha u(x) + \beta v(x) = 0 \quad (x \in J) \Rightarrow \alpha = \beta = 0,$$

so ist

$$\mathcal{L}_h = \{c_1 u + c_2 v : c_1, c_2 \in \mathbb{R}\}.$$

Derartige Lsungen u, v existieren (und heien ein *Fundamentalsystem* von (H2)).

(2) Sei $y_p \in \mathcal{L}$ (eine spezielle Lsung). Dann gilt:

$$y \in \mathcal{L} \Leftrightarrow \exists y_h \in \mathcal{L}_h : y = y_p + y_h.$$

(3) Es sei $x_0 \in J$ und $y_0, y_1 \in \mathbb{R}$. Dann ist das Anfangswertproblem

$$y'' + a_1(x)y' + a_0(x)y = b(x), \quad y(x_0) = y_0, \quad y'(x_0) = y_1$$

eindeutig lsbar und $I_{\max} = J$.

Beweis. (2) Nachrechnen (wie in 13.2).

(1) und (3): Spter in HM III. □

Wir betrachten nun den Fall, dass a_0 und a_1 konstant sind (mit $J = \mathbb{R}$). Dann heit

$$p(\lambda) := \lambda^2 + a_1\lambda + a_0$$

das zu (H2) gehrende *charakteristische Polynom*.

Satz: Es seien λ_1, λ_2 die Nullstellen von p , also $p(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2)$.

Im Fall $\lambda_1, \lambda_2 \in \mathbb{R}, \lambda_1 \neq \lambda_2$ setze

$$u(x) = e^{\lambda_1 x}, \quad v(x) = e^{\lambda_2 x}.$$

Im Fall $\lambda_1, \lambda_2 \in \mathbb{R}, \lambda_1 = \lambda_2$ setze

$$u(x) = e^{\lambda_1 x}, \quad v(x) = x e^{\lambda_1 x}.$$

Im Fall $\lambda_1 = \alpha + i\beta$ mit $\alpha \in \mathbb{R}$ und $\beta \in \mathbb{R} \setminus \{0\}$ (dann ist $\lambda_2 = \alpha - i\beta$) setze

$$u(x) = e^{\alpha x} \cos(\beta x), \quad v(x) = e^{\alpha x} \sin(\beta x).$$

Dann ist

$$\mathcal{L}_h = \{c_1 u + c_2 v : c_1, c_2 \in \mathbb{R}\}.$$

Beweis. Nachrechnen: $u, v \in \mathcal{L}_h$. Betrachte z.B. den Fall $\lambda_1, \lambda_2 \in \mathbb{R}, \lambda_1 \neq \lambda_2$. Sind $\alpha, \beta \in \mathbb{R}$ mit

$$\alpha e^{\lambda_1 x} + \beta e^{\lambda_2 x} = 0 \quad (x \in \mathbb{R}),$$

so folgt mit $x = 0$ bzw. $x = 1$:

$$\alpha + \beta = 0 \quad \wedge \quad \alpha e^{\lambda_1} + \beta e^{\lambda_2} = 0,$$

also $\alpha = e^{\lambda_1 - \lambda_2} \alpha$ und damit $\alpha = \beta = 0$. Die anderen Fälle behandelt man analog. $\square$

Beispiele:

(1) $y'' + 2y' + y = 0$, $p(\lambda) = \lambda^2 + 2\lambda + 1 = (\lambda + 1)^2$, also $\lambda_1 = \lambda_2 = -1$. Damit ist $u(x) = e^{-x}$, $v(x) = x e^{-x}$. Die allgemeine Lösung der Differentialgleichung ist also

$$y(x) = c_1 e^{-x} + c_2 x e^{-x} \quad (x \in \mathbb{R}), \quad c_1, c_2 \in \mathbb{R}.$$

Wir wollen das Anfangswertproblem

$$y'' + 2y' + y = 0, \quad y(0) = 1, \quad y'(0) = -2$$

lösen. Die Ableitung der allgemeine Lösung ist

$$y'(x) = -c_1 e^{-x} + c_2 (e^{-x} - x e^{-x}).$$

Wie erhalten die Gleichungen

$$1 = y(0) = c_1, \quad -2 = y'(0) = -c_1 + c_2,$$

also $c_2 = -1$. Die Lösung des Anfangswertproblems ist also

$$y(x) = e^{-x} - x e^{-x} \quad (x \in \mathbb{R})$$

(2) Wir wollen das Anfangswertproblem

$$y'' + y = 0, \quad y(\pi) = 1, \quad y'(\pi) = -3$$

lösen. $p(\lambda) = \lambda^2 + 1 = (\lambda - i)(\lambda + i)$, also $\lambda_1 = i = 0 + i \cdot 1$, $\lambda_2 = -i$. Damit ist $u(x) = \cos x$, $v(x) = \sin x$. Die allgemeine Lösung der Differentialgleichung ist also

$$y(x) = c_1 \cos x + c_2 \sin x \quad (x \in \mathbb{R}), \quad c_1, c_2 \in \mathbb{R},$$

mit Ableitung

$$y'(x) = -c_1 \sin x + c_2 \cos x.$$

Wie erhalten die Gleichungen

$$1 = y(\pi) = -c_1, \quad -3 = y'(\pi) = -c_2.$$

Die Lösung des Anfangswertproblems ist also

$$y(x) = -\cos x + 3 \sin x \quad (x \in \mathbb{R}).$$

Wie man zu einer speziellen Lösung von (L2) kommt behandeln wir nur für rechte Seiten $b(x)$ der Form

$$b(x) = q_m(x)e^{\alpha x} \begin{cases} \cos(\beta x) \\ \sin(\beta x) \end{cases},$$

mit $\alpha, \beta \in \mathbb{R}$ und q_m ein Polynom vom Grad $m \in \mathbb{N}_0$. Ist $\alpha + i\beta$ eine ν -fache Nullstelle des charakteristischen Polynoms p , so führt der **Ansatz**

$$y_p(x) = \left[r_m(x) \cos(\beta x) + \tilde{r}_m(x) \sin(\beta x) \right] e^{\alpha x} x^\nu$$

zum Ziel, wobei r_m und $\tilde{r}_m$ unbekannte Polynome vom Grad $\leq m$ sind (beachte, dass die Spezialfälle $\alpha = 0$, $\beta = 0$, $m = 0$, $\nu = 0$ möglich sind).

Beispiel: $y'' + y = \cos(\beta x)$ mit $\beta \geq 0$. Bekannt: $p(\lambda) = (\lambda - i)(\lambda + i)$, und die Lösungen der homogenen Gleichung $y'' + y = 0$ sind

$$y(x) = c_1 \cos x + c_2 \sin x \quad (x \in \mathbb{R}), \quad c_1, c_2 \in \mathbb{R}.$$

1. Fall: $\beta \neq 1$: Dann ist $i\beta$ *keine* Nullstelle von p (d.h. $\nu = 0$) und wir machen den Ansatz

$$y_p(x) = \gamma \cos(\beta x) + \delta \sin(\beta x).$$

Dann ist

$$y'_p(x) = -\gamma\beta \sin(\beta x) + \delta\beta \cos(\beta x), \quad y''_p(x) = -\gamma\beta^2 \cos(\beta x) - \delta\beta^2 \sin(\beta x),$$

und

$$y_p''(x) + y_p(x) = \gamma(1 - \beta^2) \cos(\beta x) + \delta(1 - \beta^2) \sin(\beta x) \stackrel{!}{=} \cos(\beta x)$$

führt auf $\delta = 0$, $\gamma = (1 - \beta^2)^{-1}$ (beachte $\beta^2 \neq 1$). Wir erhalten die Lösung

$$y_p(x) = (1 - \beta^2)^{-1} \cos(\beta x) \quad (x \in \mathbb{R}).$$

Die allgemeine Lösung

$$y(x) = c_1 \cos x + c_2 \sin x + (1 - \beta^2)^{-1} \cos(\beta x) \quad (x \in \mathbb{R}),$$

ist also für jede Wahl der Konstanten $c_1, c_2 \in \mathbb{R}$ beschränkt.

2. Fall $\beta = 1$: Hier ist $i\beta = i$ *einfache* Nullstelle des charakteristischen Polynoms p (d.h. $\nu = 1$). Der Ansatz

$$y_p(x) = \gamma x \cos x + \delta x \sin x$$

gibt

$$y_p'(x) = (\gamma + \delta x) \cos x + (\delta - \gamma x) \sin x, \quad y_p''(x) = (2\delta - \gamma x) \cos x + (-2\gamma - \delta x) \sin x,$$

und

$$y_p''(x) + y_p(x) = 2\delta \cos x - 2\gamma \sin x \stackrel{!}{=} \cos x$$

führt auf $\delta = 1/2$, $\gamma = 0$. Wir erhalten die Lösung

$$y_p(x) = \frac{x}{2} \sin x \quad (x \in \mathbb{R}).$$

Die allgemeine Lösung

$$y(x) = c_1 \cos x + c_2 \sin x + \frac{x}{2} \sin x \quad (x \in \mathbb{R}),$$

ist also für jede Wahl der Konstanten $c_1, c_2 \in \mathbb{R}$ **unbeschränkt**: Die periodische Anregung mit der Eigenfrequenz $\omega = 1$ des ungestörten Systems führt zur *Resonanzkatastrophe*.

14 Uneigentliche Integrale

Die folgenden Vereinbarungen sollen für den Rest dieses Kapitels gelten:

1. Ist $I \subseteq \mathbb{R}$ ein Intervall und $f : I \rightarrow \mathbb{R}$ eine Funktion, so sei f über jedes Intervall $[a, b] \subseteq I$ integrierbar.
2. 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 $f : [a, \beta) \rightarrow \mathbb{R}$ [bzw. $(\alpha, b] \rightarrow \mathbb{R}$] eine Funktion. Das *uneigentliche Integral* $\int_a^\beta f(x) dx$ [bzw. $\int_\alpha^b f(x) dx$] heißt *konvergent*, falls der Limes $\lim_{r \rightarrow \beta-} \int_a^r f(x) dx$ [bzw. $\lim_{r \rightarrow \alpha+} \int_r^b f(x) dx$] existiert und reell ist. In diesem Fall setzt man

$$\int_a^\beta f(x) dx := \lim_{r \rightarrow \beta-} \int_a^r f(x) dx \quad \left[\text{bzw.} \quad \int_\alpha^b f(x) dx := \lim_{r \rightarrow \alpha+} \int_r^b f(x) dx \right].$$

Ein nicht konvergentes uneigentliches Integral heißt *divergent*.

Beispiele: (1) Sei $\gamma > 0$. Dann gilt für jedes $r > 1$:

$$\int_1^r \frac{1}{x^\gamma} dx = \begin{cases} \log r, & \gamma = 1 \\ \frac{r^{1-\gamma}-1}{1-\gamma}, & \gamma \neq 1 \end{cases},$$

also ist $\int_1^\infty x^{-\gamma} dx$ konvergent genau dann, wenn $\gamma > 1$ ist. In diesem Fall ist

$$\int_1^\infty \frac{1}{x^\gamma} dx = \frac{1}{\gamma - 1}.$$

(2) Für $r > 0$ gilt

$$\int_0^r \frac{1}{1+x^2} dx = \arctan r \rightarrow \frac{\pi}{2} \quad (r \rightarrow \infty).$$

Also ist $\int_0^\infty \frac{1}{1+x^2} dx$ konvergent und $= \frac{\pi}{2}$.

(3) Sei $\gamma > 0$. Dann gilt für $r \in (0, 1)$:

$$\int_r^1 \frac{1}{x^\gamma} dx = \begin{cases} -\log r, & \gamma = 1 \\ \frac{1-r^{1-\gamma}}{1-\gamma}, & \gamma \neq 1 \end{cases},$$

also ist $\int_0^1 x^{-\gamma} dx$ konvergent genau dann, wenn $\gamma < 1$ ist. In diesem Fall ist

$$\int_0^1 \frac{1}{x^\gamma} dx = \frac{1}{1-\gamma}.$$

(4) Analog zu (2): $\int_{-\infty}^0 \frac{1}{1+x^2} dx$ ist konvergent und $= \frac{\pi}{2}$.

Definition: Sei $f : (\alpha, \beta) \rightarrow \mathbb{R}$ 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

$$\int_{\alpha}^{\beta} f(x) dx := \int_{\alpha}^c f(x) dx + \int_c^{\beta} f(x) dx.$$

Diese 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 $\int_{-r}^r x dx = 0$ für $r \rightarrow \infty$ gegen 0.

Bemerkung: Wir betrachten im folgenden nur den Fall von Funktionen $f : [a, \beta) \rightarrow \mathbb{R}$. Entsprechendes gilt jeweils auch für Funktionen $f : (\alpha, b] \rightarrow \mathbb{R}$ bzw. $f : (\alpha, \beta) \rightarrow \mathbb{R}$.

14.2. Satz (Cauchy Kriterium): $\int_a^{\beta} f(x) dx$ konvergiert genau dann, wenn gilt:

$$\forall \varepsilon > 0 \exists c = c(\varepsilon) \in (a, \beta) \forall u, v \in (c, \beta) : \left| \int_u^v f(x) dx \right| < \varepsilon.$$

(ohne Beweis)

Beispiel: $\int_0^{\infty} \frac{\sin x}{x} dx$ ist konvergent.

Es gilt $\frac{\sin x}{x} \rightarrow 1$ ($x \rightarrow 0+$), weshalb das Integral bei 0 nicht uneigentlich ist. Für $v > u > 0$ gilt:

$$\begin{aligned} \left| \int_u^v \frac{\sin x}{x} dx \right| &= \left| \int_u^v \underbrace{\frac{1}{x}}_g \underbrace{\sin x}_{f'} dx \right| \\ &= \left| \left[-\frac{\cos x}{x} \right]_u^v - \int_u^v \frac{\cos x}{x^2} dx \right| \\ &\leq \frac{|\cos u|}{u} + \frac{|\cos v|}{v} + \int_u^v \frac{|\cos x|}{x^2} dx \\ &\leq \frac{1}{u} + \frac{1}{v} + \left[-\frac{1}{x} \right]_u^v = \frac{2}{u}. \end{aligned}$$

Sei $\varepsilon > 0$. Für $v > u > c(\varepsilon) := 2/\varepsilon$ gilt dann

$$\left| \int_u^v \frac{\sin x}{x} dx \right| \leq 2/u < \varepsilon.$$

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

$$\int_{k\pi}^{(k+1)\pi} \frac{|\sin x|}{x} dx \geq \underbrace{\int_0^\pi |\sin x| dx}_{=2} \frac{1}{(k+1)\pi} = \frac{2}{(k+1)\pi},$$

also gilt für $n \in \mathbb{N}$:

$$\int_0^{(n+1)\pi} \frac{|\sin x|}{x} dx \geq \sum_{k=0}^n \frac{2}{(k+1)\pi} \rightarrow \infty \quad (n \rightarrow \infty).$$

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

$$\left| \int_a^\beta f(x) dx \right| \leq \int_a^\beta |f(x)| dx.$$

(2) **Majorantenkriterium:**

Ist $|f| \leq g$ auf $[a, \beta)$ und $\int_a^\beta g(x) dx$ konvergent, so ist $\int_a^\beta f(x) dx$ absolut konvergent.

(3) **Minorantenkriterium:**

Ist $f \geq g \geq 0$ auf $[a, \beta)$ und $\int_a^\beta g(x) dx$ divergent, so ist $\int_a^\beta f(x) dx$ divergent.

Beispiele: (1) $\int_1^\infty \frac{x}{\sqrt{1+x^5}} dx$ ist konvergent: Es gilt

$$f(x) := \frac{x}{\sqrt{1+x^5}} \leq \frac{1}{x^{3/2}} =: g(x) \quad (x \geq 1).$$

Da $\int_1^\infty g(x) dx$ konvergiert, konvergiert auch $\int_1^\infty f(x) dx$ und es gilt:

$$0 \leq \int_1^\infty f(x) dx \leq \int_1^\infty g(x) dx = 2.$$

(2) $\int_1^\infty \underbrace{\frac{x}{x^2 + 7x}}_{=:f(x)} dx$ ist divergent: Setze $g(x) := 1/x$. Dann gilt

$$\frac{f(x)}{g(x)} = \frac{x^2}{x^2 + 7x} \rightarrow 1 \quad (x \rightarrow \infty).$$

Also existiert ein $c \geq 1$ mit $f/g \geq 1/2$ auf $[c, \infty)$, d.h. es gilt

$$f(x) \geq \frac{1}{2}g(x) = \frac{1}{2x} \quad (x \geq c)$$

(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 $\int_1^\infty f(x) dx$.

Warnungen:

(1) Sind $\int_a^\beta f dx$ und $\int_a^\beta g dx$ konvergent, so muss $\int_a^\beta fg dx$ **nicht** konvergieren!

Betrachte z.B. $f(x) = g(x) = x^{-1/2}$. Dann ist $\int_0^1 f(x) dx$ konvergent, aber $\int_0^1 f^2(x) dx$ ist divergent.

(2) Im Unterschied zu Reihen (vgl. 7.2 (3)) folgt aus der Konvergenz von

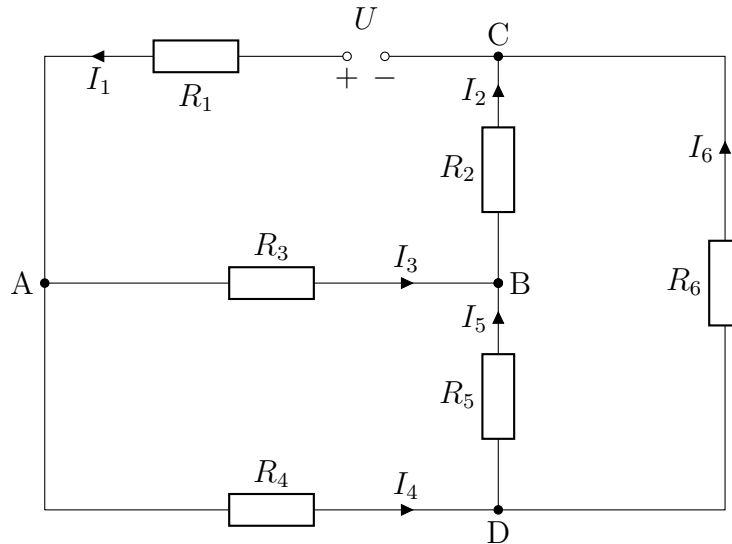
$$\int_a^\infty f(x) dx$$

i.a. **nicht** $f(x) \rightarrow 0$ ($x \rightarrow \infty$). Betrachte z.B. (mit $a = 0$) die Funktion

$$f : [0, \infty) \rightarrow \mathbb{R}, \quad f(x) = \begin{cases} 0, & x \in [0, \infty) \setminus \mathbb{N} \\ x, & x \in \mathbb{N} \end{cases}.$$

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 $R_1, R_2, \dots, R_6$ bekannt sind, und wollen die Ströme $I_1, I_2, \dots, I_6$ bestimmen. Durch Betrachtung der Knoten A, B, C, D bzw. der Maschen ABC, ABD, BCD, ADC erhalten wir die Gleichungen

$$\begin{array}{rcccccccl}
 I_1 & & -I_3 & -I_4 & & & = & 0 \\
 & -I_2 & +I_3 & & +I_5 & & = & 0 \\
 -I_1 & +I_2 & & & & +I_6 & = & 0 \\
 & & & I_4 & -I_5 & -I_6 & = & 0 \\
 R_1 I_1 & +R_2 I_2 & +R_3 I_3 & & & & = & U \\
 & & R_3 I_3 & -R_4 I_4 & -R_5 I_5 & & = & 0 \\
 & R_2 I_2 & & & +R_5 I_5 & -R_6 I_6 & = & 0 \\
 R_1 I_1 & & & +R_4 I_4 & & +R_6 I_6 & = & U.
 \end{array}$$

Das ist ein *lineares Gleichungssystem (LGS)* mit acht Gleichungen für die sechs Unbekannten $I_1, \dots, I_6$. Die natürlichen Fragen sind zunächst die nach **Existenz** und **Eindeutigkeit** der Lösung und ggf. die nach deren **Berechnung**.

Bemerkung: Man kann hier sehen, dass z.B. Addition der ersten drei Gleichungen (bis aufs Vorzeichen) die vierte Gleichung ergibt und Addition der letzten drei Gleichungen die fünfte 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 $\vec{x} = (I_1, I_2, \dots, I_6) \in \mathbb{R}^6$ 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: Im folgenden sei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$. Es seien $n, m \in \mathbb{N}$. Wir schreiben ein *lineares Gleichungssystem (LGS)*

$$\begin{array}{cccccc} a_{11}x_1 + & a_{12}x_2 + & \dots & + a_{1m}x_m & = & b_1 \\ a_{21}x_1 + & a_{22}x_2 + & \dots & + a_{2m}x_m & = & b_2 \\ \vdots & \vdots & \ddots & \vdots & & \vdots \\ a_{n1}x_1 + & a_{n2}x_2 + & \dots & + a_{nm}x_m & = & b_n \end{array}$$

mit gegebenen $a_{jk}, b_j \in \mathbb{K}$ und gesuchten $x_1, x_2, \dots, x_m \in \mathbb{K}$ in Matrixschreibweise als

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

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 $A = (a_{jk})_{j=1, k=1}^n = (a_{jk}) \in \mathbb{K}^{n \times m}$ und $\vec{x} = (x_k)_{k=1}^m = (x_k) \in \mathbb{K}^m$ ist

$$A\vec{x} = \left(\sum_{k=1}^m a_{jk} x_k \right)_{j=1}^n \in \mathbb{K}^n.$$

Beachte: Wir schreiben Vektoren $\vec{x} = (x_1, x_2, \dots, x_m) \in \mathbb{K}^m$, die wir bisher als m -Tupel geschrieben haben, ab jetzt ggf. als **Spaltenvektoren!** Dies ist die übliche Konvention.

Beispiel: Sei $l \in \{1, 2, \dots, m\}$. 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. Es ist also

$$\vec{e}_l := (0, \dots, 0, \underbrace{1}_{l\text{-te Stelle}}, 0, \dots, 0) \quad (l = 1, 2, \dots, m).$$

Wir erhalten dann

$$A\vec{e}_l = \left(\sum_{k=1}^m a_{jk} \delta_{kl} \right)_{j=1}^n = (a_{jl})_{j=1}^n \in \mathbb{K}^n,$$

d.h. $A\vec{e}_l$ 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:

$$A(\vec{x} + \vec{y}) = A\vec{x} + A\vec{y}, \quad (A + B)\vec{x} = A\vec{x} + B\vec{x}, \quad A(\alpha\vec{x}) = \alpha(A\vec{x}) = (\alpha A)\vec{x},$$

wobei für $A = (a_{jk})$, $B = (b_{jk})$ die Matrizen $A + B, \alpha A \in \mathbb{K}^{n \times m}$ gegeben sind durch

$$A + B = (a_{jk} + b_{jk}), \quad \alpha A = (\alpha a_{jk}),$$

d.h. an jeder Stelle (j, k) werden die Einträge addiert bzw. mit α multipliziert.

Diese Rechenregeln kann man folgendermaßen formulieren:

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 $\mathbb{K}^m \rightarrow \mathbb{K}^n$, $\vec{x} \mapsto A\vec{x}$ linear.

3) Für festes $\vec{x} \in \mathbb{K}^m$ ist die Abbildung $\mathbb{K}^{n \times m} \rightarrow \mathbb{K}^n$, $A \mapsto A\vec{x}$ linear.

Bemerkung: Das bedeutet, dass das Lineare Gleichungssystem

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

(mit $A \in \mathbb{K}^{n \times m}$ und $\vec{b} \in \mathbb{K}^n$ gegeben und $\vec{x} \in \mathbb{K}^m$ gesucht) ein Spezialfall der folgenden Situation ist:

15.4. Lineare Gleichungen: Seien V und W zwei $\mathbb{K}$ -Vektorräume, $\phi : V \rightarrow W$ sei eine lineare Abbildung und $w \in W$ gegeben. Man sucht $v \in V$ mit

$$\phi(v) = w. \quad (\text{LG})$$

Ist $w = 0$, so heißt die lineare Gleichung (LG) *homogen*, sonst *inhomogen*, wobei $w \neq 0$ als *Inhomogenität* bezeichnet wird.

Beispiele: Sei $I \subseteq \mathbb{R}$ ein Intervall (mit mindestens zwei Punkten).

1) Lineare Differentialgleichung erster Ordnung (siehe 13.2): Seien $a, b \in C(I)$, $V := C^1(I)$ und $W := C(I)$, sowie

$$L : C^1(I) \rightarrow C(I), \quad L(y) = y' - ay.$$

Dann ist L linear und wir können (L1) in 13.2 schreiben als $L(y) = b$.

2) Lineare Differentialgleichung zweiter Ordnung (siehe 13.4): Seien $a_0, a_1, b \in C(I)$, $V := C^2(I)$ und $W := C(I)$, sowie

$$L : C^2(I) \rightarrow C(I), \quad L(y) = y'' + a_1 y' + a_0 y.$$

Dann ist L linear und wir können (L2) in 13.4 schreiben als $L(y) = b$.

Erinnerung: Wir hatten in 8.4 für eine lineare Abbildung $\phi : V \rightarrow W$ die Notationen

$$\text{Kern } \phi = \{v \in V : \phi(v) = 0\}, \quad \text{Bild } \phi = \{\phi(v) : v \in V\}$$

eingeführt und gezeigt, dass Kern ϕ ein UVR von V und Bild ϕ ein UVR von W ist. Beachte, dass Kern ϕ die Lösungsmenge der homogenen Gleichung $\phi(v) = 0$ ist.

Satz: 1) Die Gleichung (LG) ist genau dann lösbar, wenn $w \in \text{Bild } \phi$ ist.

2) Ist (LG) lösbar und $v_0 \in V$ eine Lösung, so gilt für die Lösungsmenge

$$\{v \in V : \phi(v) = w\} = v_0 + \text{Kern } \phi := \{v_0 + v_1 : v_1 \in \text{Kern } \phi\}.$$

Insbesondere ist v_0 genau dann die einzige Lösung, wenn Kern $\phi = \{0\}$ ist.

3) ϕ ist injektiv genau dann, wenn 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 $v - v_0 \in \text{Kern } \phi$, d.h. $v = v_0 + (v - v_0) \in v_0 + \text{Kern } \phi$. Ist umgekehrt $v = v_0 + v_1$ mit $v_1 \in \text{Kern } \phi$, so gilt

$$\phi(v) = \phi(v_0) + \phi(v_1) = w + 0 = w.$$

Somit ist 2) gezeigt. 3) folgt aus 2). □

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 UVR 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 gilt

$$U = \{v - v_0 : v \in F\} = \{v_1 - v_2 : v_1, v_2 \in F\}.$$

Ist umgekehrt $\emptyset \neq F \subseteq V$, $v_1 \in F$ und ist

$$U := \{v_1 - v_2 : v_2 \in F\}$$

ein UVR von V , so ist F ein affiner Teilraum von V mit $F = v + U$ für jedes $v \in F$.

Beispiel: In $\mathbb{R}^2$ sei $(a, b) \neq (0, 0)$. Dann ist $U := \{(x, y) \in \mathbb{R}^2 : ax + by = 0\}$ ein UVR (Gerade durch $(0, 0)$) und $F := \{(x, y) \in \mathbb{R}^2 : ax + by = 1\}$ ein affiner 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: $A\vec{x} = \vec{b}$, mit $A \in \mathbb{K}^{n \times m}$ und $\vec{b} \in \mathbb{K}^n$: Dann sind (da $\vec{x} \mapsto A\vec{x}$ linear ist)

$$\text{Kern } A := \{\vec{x} \in \mathbb{K}^m : A\vec{x} = \vec{0}\} \subseteq \mathbb{K}^m \quad \text{und} \quad \text{Bild } A := \{A\vec{x} : \vec{x} \in \mathbb{K}^m\} \subseteq \mathbb{K}^n$$

Untervektorräume von $\mathbb{K}^m$ bzw. $\mathbb{K}^n$. 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 $\vec{b} \in \text{Bild } A$ ist. Ist $\vec{z} \in \mathbb{K}^m$ eine Lösung, so ist die Menge aller Lösungen $\vec{z} + \text{Kern } A$ ein affiner Teilraum von $\mathbb{K}^m$.

15.5. Die lineare Hülle: Ist V ein $\mathbb{K}$ -Vektorraum, $n \in \mathbb{N}$ und sind $v_1, v_2, \dots, v_n$ Vektoren aus V , so heißt jeder Vektor der Form

$$\sum_{j=1}^n \alpha_j v_j \quad \text{mit } \alpha_1, \alpha_2, \dots, \alpha_n \in \mathbb{K}$$

eine *Linearkombination* (LK) der Vektoren $v_1, v_2, \dots, v_n$.

Satz: Ist V ein $\mathbb{K}$ -Vektorraum und $\emptyset \neq M \subseteq V$, so ist

$$\text{lin}M := \text{lin}(M) := \left\{ \sum_{j=1}^n \alpha_j v_j : n \in \mathbb{N}, \alpha_j \in \mathbb{K}, v_j \in M \right\}$$

ein UVR von V und heit die *lineare Hlle von M* (oder *der von M erzeugte UVR*).

Bemerkungen: (1) Im Fall $M \neq \emptyset$ ist also $\text{lin}(M)$ die Menge aller Linearkombinationen von Vektoren aus M . Im Fall $M = \emptyset$ setzt man $\text{lin}M := \text{lin}(M) := \{0\}$.

(2) Offenbar ist $\text{lin}(M)$ der *kleinste* UVR von V , der M enthlt. Insbesondere ist $U \subseteq V$ ein UVR von V genau dann, wenn $\text{lin}(U) = U$ gilt.

(3) Ist $M_1 \subseteq M_2 \subseteq V$, so gilt $\text{lin}(M_1) \subseteq \text{lin}(M_2)$.

Beispiele: (1) In $V = \mathbb{R}^2$ gilt

$$\text{lin}\{(1, 0)\} = \{(\alpha, 0) : \alpha \in \mathbb{R}\} = \mathbb{R} \times \{0\}.$$

(2) In $V = \mathbb{R}^2$ gilt $\text{lin}\{(1, 0), (0, 1)\} = \mathbb{R}^2 = \text{lin}\{(1, 0), (0, 1), (1, 1)\}$.

(3) Fr $K = \{(x_1, x_2) \in \mathbb{R}^2 : x_1^2 + x_2^2 \leq 1\}$ gilt $\text{lin}(K) = \mathbb{R}^2$.

(4) Fr $C = [0, \infty) \times [0, \infty)$ gilt $\text{lin}(C) = \mathbb{R}^2$.

(5) In $V = \mathbb{K}^m$ gilt $\mathbb{K}^m = \text{lin}\{\vec{e}_1, \vec{e}_2, \dots, \vec{e}_m\}$, denn

$$\forall \vec{x} = (x_1, x_2, \dots, x_m) \in \mathbb{K}^m : \vec{x} = \sum_{j=1}^m x_j \vec{e}_j.$$

(6) Sei $A \in \mathbb{K}^{n \times m}$. Fr $\vec{x} \in \mathbb{K}^m$ gilt

$$A\vec{x} = A\left(\sum_{j=1}^m x_j \vec{e}_j\right) = \sum_{j=1}^m x_j A\vec{e}_j,$$

also

$$\text{Bild } A = \text{lin}\{A\vec{e}_1, A\vec{e}_2, \dots, A\vec{e}_m\}.$$

Da $A\vec{e}_j$ nach dem Beispiel in 15.2 die j -te Spalte von A ist, haben wir somit:

Bild A ist die lineare Hlle der Spalten von A .

(7) Sei $\mathcal{L}_h := \{u \in C^2(\mathbb{R}) : u'' + u = 0\}$. Das charakteristische Polynom der Differentialgleichung ist $\lambda^2 + 1 = (\lambda + i)(\lambda - i)$. Nach 13.4 gilt

$$\mathcal{L}_h = \text{lin}\{\sin, \cos\}.$$

15.6. Lineare Unabhängigkeit: Sei V ein $\mathbb{K}$ -Vektorraum.

Definition: Man nennt n Vektoren $v_1, v_2, \dots, v_n \in V$ *linear unabhängig*, falls für alle $\alpha_1, \alpha_2, \dots, \alpha_n \in \mathbb{K}$ gilt:

$$\sum_{j=1}^n \alpha_j v_j = 0 \implies \alpha_1 = \alpha_2 = \dots = \alpha_n = 0.$$

Man kann in diesem Fall den Nullvektor also **nur** als *triviale Linearkombination* der $v_1, v_2, \dots, v_n$ erhalten, d.h. nur wenn alle $\alpha_j = 0$ sind.

Sind $v_1, v_2, \dots, v_n$ nicht linear unabhängig, so heißen sie *linear abhängig*. Die Vektoren $v_1, v_2, \dots, v_n$ sind also genau dann linear abhängig, wenn man $\alpha_1, \alpha_2, \dots, \alpha_n \in \mathbb{K}$ 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* $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n$ linear unabhängig.

(4) In $C^2(\mathbb{R})$ sind $\sin, \cos$ linear unabhängig: Sind $\alpha, \beta \in \mathbb{R}$ mit

$$\alpha \sin + \beta \cos = 0, \quad \text{d.h.} \quad \alpha \sin x + \beta \cos x = 0 \quad (x \in \mathbb{R}),$$

so folgt $\alpha = 0$ 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 $\phi : V \rightarrow W$ linear und **injektiv** und sind $v_1, v_2, \dots, v_n \in V$ linear unabhängig, so sind $\phi(v_1), \phi(v_2), \dots, \phi(v_n)$ linear unabhängig: Für $\alpha_1, \alpha_2, \dots, \alpha_n \in \mathbb{K}$ gilt

$$\sum_{j=1}^n \alpha_j \phi(v_j) = 0 \implies \phi\left(\sum_{j=1}^n \alpha_j v_j\right) = 0 \implies \sum_{j=1}^n \alpha_j v_j = 0 \implies \alpha_1 = \alpha_2 = \dots = \alpha_n = 0.$$

15.7. Zeilenumformungen: Seien n Vektoren $v_1, v_2, \dots, v_n$ aus dem $\mathbb{K}^m$ gegeben, deren lineare Unabhängigkeit wir untersuchen wollen. Dabei sei $v_j = (v_{j1}, v_{j2}, \dots, v_{jm})$ für $j = 1, 2, \dots, n$. Wir schreiben die Vektoren $v_1, v_2, \dots, v_n$ als **Zeilen** in eine Matrix A , d.h.

$$A = \begin{pmatrix} v_{11} & v_{12} & \dots & v_{1m} \\ v_{21} & v_{22} & \dots & v_{2m} \\ \dots & \dots & \dots & \dots \\ v_{n1} & v_{n2} & \dots & v_{nm} \end{pmatrix} \in \mathbb{K}^{n \times m}.$$

Die Matrix A bringen wir in eine Form, an der wir die lineare Unabhängigkeit der Zeilen ablesen können, mittels der folgenden *Zeilenumformungen*:

- (Z1) Ersetze eine Zeile v_j durch αv_j , wobei $\alpha \in \mathbb{K} \setminus \{0\}$;
- (Z2) Ersetze eine Zeile v_j durch $v_j + \beta v_k$, wobei $\beta \in \mathbb{K}$ und $k \neq j$;
- (Z3) Vertausche die Zeilen v_j und v_k , wobei $j \neq k$.

Bemerkung: Die Zeilen von A sind genau dann linear unabhängig, wenn die Zeilen der umgeformten Matrix linear unabhängig sind.

Satz: Jede Matrix $A = (a_{jk}) \in \mathbb{K}^{n \times m}$ kann man durch endlich viele Zeilenumformungen in eine Matrix $C \in \mathbb{K}^{n \times m}$ überführen, die in *Zeilenstufenform* (ZSF) ist. Dabei heißt

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

in *Zeilenstufenform*, falls es ein $r \in \{0, 1, \dots, n\}$ und $1 \leq k_1 < k_2 < \dots < k_r \leq m$ gibt mit

- für $j = 1, 2, \dots, r$ gilt $c_{jk} = 0$ für $k = 1, 2, \dots, k_j - 1$ und $\gamma_j := c_{jk_j} \neq 0$;
- für $j = r + 1, \dots, n$ gilt $c_{jk} = 0$ für alle $k = 1, 2, \dots, m$.

Man braucht dafür sogar nur Umformungen der Art (Z2) und (Z3).

Beispiel für eine mögliche 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:

$$\begin{pmatrix} 0 & \gamma_1 & * & * & * & * & * & * \\ 0 & 0 & 0 & \gamma_2 & * & * & * & * \\ 0 & 0 & 0 & 0 & \gamma_3 & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & \gamma_4 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Hier ist also $k_1 = 2, k_2 = 4, k_3 = 5, k_4 = 7$.

Zum Beweis des Satzes betrachtet man zunächst die erste Spalte von A .

- 1) Ist $a_{11} \neq 0$, so setzt man $C_0 := A$, $\gamma_1 := a_{11}$ und zieht für $j = 2, \dots, n$ von der j -ten Zeile das a_{j1}/γ_1 -fache der ersten Zeile ab (d.h. man wendet (Z2) an). Man erhält so eine neue Matrix C_1 mit Nullen in der ersten Spalte unterhalb von γ_1 :

$$C_1 = \begin{pmatrix} \gamma_1 & a_{12} & \dots & a_{1m} \\ 0 & c_{22}^{(1)} & \dots & c_{2m}^{(1)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & c_{n2}^{(1)} & \dots & c_{nm}^{(1)} \end{pmatrix},$$

und macht mit der Matrix weiter, die man durch "Streichen" der ersten Zeile und ersten Spalte von C_1 erhält.

- 2) Ist $a_{11} = 0$ und $a_{j1} \neq 0$ für ein $j \in \{2, \dots, n\}$, so vertauscht man die erste und die j -te Zeile (d.h. man wendet (Z3) an), nennt die entstehende Matrix C_0 und mache dann wie in 1) weiter.
- 3) Ist $a_{j1} = 0$ für alle $j = 1, 2, \dots, n$, so wendet man das Verfahren an auf die Matrix, die durch "Streichen" der ersten Spalte von A entsteht.

Das Verfahren wird nun solange wiederholt, bis Zeilenstufenform erreicht ist. "Streichen" bedeutet hier nur, 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, d.h. 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

$$v_1 = (1, 3, -2, 4), \quad v_2 = (-1, -1, 5, -9), \quad v_3 = (2, 0, -13, 23), \quad v_4 = (1, 5, 1, -2) \in \mathbb{R}^4$$

und erhalten nacheinander die Matrizen

$$\begin{pmatrix} 1 & 3 & -2 & 4 \\ -1 & -1 & 5 & -9 \\ 2 & 0 & -13 & 23 \\ 1 & 5 & 1 & -2 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 3 & -2 & 4 \\ 0 & 2 & 3 & -5 \\ 0 & -6 & -9 & 15 \\ 0 & 2 & 3 & -6 \end{pmatrix} \rightarrow$$

$$\begin{pmatrix} 1 & 3 & -2 & 4 \\ 0 & 2 & 3 & -5 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 3 & -2 & 4 \\ 0 & 2 & 3 & -5 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Die Vektoren v_1, v_2, v_3, v_4 sind also linear abhängig.

Bemerkung: Man sieht auch, dass Zeilenumformungen die lineare Hülle 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, d.h. 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

$$\text{lin}\{v_1, v_2, v_3\} = \text{lin}\{v_1, v_2\},$$

und weiter also auch

$$\text{lin}\{v_1, v_2, v_3, v_4\} = \text{lin}\{v_1, v_2, v_4\}.$$

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:

$$\begin{aligned} & \begin{pmatrix} 1 & 0 & -1 & -1 & 0 & 0 \\ 0 & -1 & 1 & 0 & 1 & 0 \\ -1 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & -1 & -1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & -1 & -1 & 0 & 0 \\ 0 & -1 & 1 & 0 & 1 & 0 \\ 0 & 1 & -1 & -1 & 0 & 1 \\ 0 & 0 & 0 & 1 & -1 & -1 \end{pmatrix} \\ & \rightarrow \begin{pmatrix} 1 & 0 & -1 & -1 & 0 & 0 \\ 0 & -1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1 & 1 \\ 0 & 0 & 0 & 1 & -1 & -1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & -1 & -1 & 0 & 0 \\ 0 & -1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}. \end{aligned}$$

Es wurden hierbei keine Zeilen vertauscht. Die ersten drei Zeilen sind also linear unabhängig, alle vier Zeilen sind linear abhängig. Die vierte liegt somit in der linearen Hülle 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 $j = 1, 2, \dots, r$ gilt: $\gamma_j = c_{jk_j} = 1$ und $c_{lk_j} = 0$ für $l = 1, 2, \dots, j - 1$ (d.h. 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 alle γ_j auf 1 und gehen dann von rechts nach links vor:

$$\begin{aligned} & \begin{pmatrix} 1 & 3 & -2 & 4 \\ 0 & 2 & 3 & -5 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 3 & -2 & 4 \\ 0 & 1 & 3/2 & -5/2 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \rightarrow \\ & \begin{pmatrix} 1 & 3 & -2 & 0 \\ 0 & 1 & 3/2 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & -13/2 & 0 \\ 0 & 1 & 3/2 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}. \end{aligned}$$

15.9. Basen und Dimension: Es sei V ein $\mathbb{K}$ -Vektorraum.

Definition: Sei $n \in \mathbb{N}$. Die Vektoren $b_1, b_2, \dots, b_n \in V$ heißen eine (*geordnete*) *Basis* von V , falls sie linear unabhängig sind und

$$\text{lin}\{b_1, b_2, \dots, b_n\} = V$$

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

Der $\mathbb{K}$ -Vektorraum $\mathbb{K}^{n \times m}$ hat die Dimension $n \cdot m$. Eine Basis ist gegeben durch die Matrizen

$$B_{pq} := (\delta_{jp} \delta_{kq})_{j=1, k=1}^n \quad (p = 1, \dots, n, q = 1, \dots, m).$$

Die Matrix B_{pq} hat 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 $v_1, v_2, \dots, v_n \in V$ gibt. In diesem Fall setzt man $\dim V := \infty$.

Beispiel: Sei $a < b$. Dann gilt $\dim C([a, b]) = \infty$: Für jedes $k \in \mathbb{N}_0$ sei $p_k : [a, b] \rightarrow \mathbb{R}$, $p_k(x) := x^k$. Dann sind für jedes $n \in \mathbb{N}$ die Funktionen $p_0, p_1, \dots, p_n$ linear unabhängig, denn für $\alpha_0, \alpha_1, \dots, \alpha_n \in \mathbb{R}$ mit

$$\alpha_0 p_0(x) + \alpha_1 p_1(x) + \dots + \alpha_n p_n(x) = 0 \quad (x \in [a, b])$$

folgt $\alpha_0 = \alpha_1 = \dots = \alpha_n = 0$, da nur das Nullpolynom unendlich viele Nullstellen hat (vgl. 5.5).

Bemerkung: Ist $n \in \mathbb{N}$ und $b_1, b_2, \dots, b_n$ eine Basis von V , so gibt es zu jedem $v \in V$ eindeutig bestimmte $\alpha_1, \alpha_2, \dots, \alpha_n \in \mathbb{K}$ mit

$$v = \alpha_1 b_1 + \alpha_2 b_2 + \dots + \alpha_n b_n = \sum_{j=1}^n \alpha_j b_j.$$

Diese $\alpha_1, \alpha_2, \dots, \alpha_n$ heißen die *Koordinaten* von v bzgl. der Basis $b_1, b_2, \dots, b_n$.

Beweis. Die Existenz folgt aus $v \in V = \text{lin}\{b_1, b_2, \dots, b_n\}$. Sind $\alpha_j, \tilde{\alpha}_j \in \mathbb{K}$ ($j = 1, \dots, n$) 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 $\alpha_j = \tilde{\alpha}_j$ ($j = 1, \dots, n$), da $b_1, b_2, \dots, b_n$ linear unabhängig sind. □

Satz: Sei V ein $\mathbb{K}$ -Vektorraum und seien $v_1, v_2, \dots, v_n \in V$. Dann ist die Dimension $\dim \operatorname{lin}\{v_1, v_2, \dots, v_n\}$ der linearen Hülle von $v_1, v_2, \dots, v_n$ gerade die Maximalzahl linear unabhängiger Vektoren unter den $v_1, v_2, \dots, v_n$.

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 eine Basis von

$$\operatorname{lin}\{\vec{v}_1, \vec{v}_2, \vec{v}_3\} = \operatorname{lin}\{\vec{v}_1, \vec{v}_2\}.$$

Also 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 Vektorraum V ist es für die Dimensionsbestimmung wichtig, ob man ihn als $\mathbb{C}$ -Vektorraum oder als $\mathbb{R}$ -Vektorraum betrachtet, d.h. ob man Linearkombinationen mit komplexen oder nur mit reellen Koeffizienten zulässt. Z.B. gilt

$$\mathbb{C} - \dim \mathbb{C} = 1, \text{ aber } \mathbb{R} - \dim \mathbb{C} = 2$$

(wie schon an der Veranschaulichung von $\mathbb{C}$ als komplexe Zahlenebene zu sehen ist). Eine Basis des $\mathbb{R}$ -Vektorraums $\mathbb{C}$ ist gegeben durch $1, i$. Eine Basis des $\mathbb{C}$ -Vektorraums $\mathbb{C}$ ist gegeben durch 1 .

Ebenso ist die Dimension von $\mathbb{C}^n$ als $\mathbb{C}$ -Vektorraum $= n$ und eine Basis ist $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n$. Die Dimension von $\mathbb{C}^n$ als $\mathbb{R}$ -Vektorraum ist hingegen $= 2n$ und eine Basis ist z.B. $\vec{e}_1, \vec{e}_2, \dots, \vec{e}_n, i\vec{e}_1, i\vec{e}_2, \dots, i\vec{e}_n$.

15.10. Die Lösungsmenge eines linearen Gleichungssystems: Seien $A \in \mathbb{K}^{n \times m}$ und $\vec{b} \in \mathbb{K}^n$ gegeben. Bisher bekannt:

- (1) Kern A ist ein Untervektorraum von $\mathbb{K}^m$ und Bild A ist ein Untervektorraum von $\mathbb{K}^n$.
- (2) Bild A ist die lineare Hülle 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) $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$, $\vec{b} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Dann hat $A\vec{x} = \vec{b}$ keine Lösung.

(2) Sei A wie eben und $\vec{b} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Dann ist $A\vec{x} = \vec{b}$ lösbar und die Menge aller Lösungen ist der affine Teilraum

$$\left\{ \begin{pmatrix} \alpha \\ 1 \end{pmatrix} : \alpha \in \mathbb{K} \right\} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \underbrace{\operatorname{lin} \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\}}_{= \operatorname{Kern} A}.$$

(3) Sei $A = \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix}$ und $\vec{b} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$. Dann ist $\vec{x} = \begin{pmatrix} -1/2 \\ 1 \end{pmatrix}$ 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 $(\tilde{B} | \vec{c})$ aus der Matrix $(B | \vec{c})$ durch eine Zeilenumformung hervor, so gilt

$$\{\vec{x} \in \mathbb{K}^m : \tilde{B}\vec{x} = \vec{c}\} = \{\vec{x} \in \mathbb{K}^m : B\vec{x} = \vec{c}\},$$

d.h. 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:

$$R_1 = 5 \Omega, \quad R_2 = R_6 = 10 \Omega, \quad R_3 = R_4 = 20 \Omega, \quad R_5 = 40 \Omega, \quad U = 40 V.$$

Dabei lassen wir die vierte und die fünfte Zeile weg, da diese Linearkombination der ersten drei Zeilen bzw. der letzten drei Zeilen sind und also die Lösungsmenge nicht geändert

wird.

$$\left(\begin{array}{cccccc|c} 1 & 0 & -1 & -1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & -1 & -1 & 0 \\ R_1 & R_2 & R_3 & 0 & 0 & 0 & U \\ 0 & 0 & R_3 & -R_4 & -R_5 & 0 & 0 \\ 0 & R_2 & 0 & 0 & R_5 & -R_6 & 0 \\ R_1 & 0 & 0 & R_4 & 0 & R_6 & U \end{array} \right) \rightarrow \left(\begin{array}{cccccc|c} \underline{1} & 0 & -1 & -1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 20 & -20 & -40 & 0 & 0 \\ 0 & 10 & 0 & 0 & 40 & -10 & 0 \\ 5 & 0 & 0 & 20 & 0 & 10 & 40 \end{array} \right)$$

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:

$$I_6 = 1 A, \quad I_5 = 0 A, \quad I_4 = 1 A, \quad I_3 = 1 A, \quad I_2 = 1 A, \quad I_1 = 2 A.$$

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 $j \in \{1, \dots, n\}$ gibt mit $c_{jk} = 0$ für $k = 1, \dots, m$ 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 $A\vec{x} = \vec{b}$ 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 $\text{Kern } A = \text{lin}(\{ \begin{pmatrix} -\alpha_1 \\ -\alpha_2 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -\beta_1 \\ -\beta_2 \\ 0 \\ -\beta_3 \\ 1 \end{pmatrix} \})$ und $\dim \text{Kern } A = 2$. 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

$$\left(\begin{array}{ccccc|c} 1 & 0 & \alpha_1 & 0 & \beta_1 & c_1 \\ 0 & 1 & \alpha_2 & 0 & \beta_2 & c_2 \\ 0 & 0 & 0 & 1 & \beta_3 & c_3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right) \rightarrow \left(\begin{array}{ccccc|c} 1 & 0 & \alpha_1 & 0 & \beta_1 & c_1 \\ 0 & 1 & \alpha_2 & 0 & \beta_2 & c_2 \\ 0 & 0 & 0 & 1 & \beta_3 & c_3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array} \right) \xrightarrow{(-1)} \left(\begin{array}{ccccc|c} 1 & 0 & \alpha_1 & 0 & \beta_1 & c_1 \\ 0 & 1 & \alpha_2 & 0 & \beta_2 & c_2 \\ 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & \beta_3 & c_3 \\ 0 & 0 & 0 & 0 & -1 & 0 \end{array} \right).$$

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:

$$\left(\begin{array}{ccccc|c} 1 & 0 & \alpha_1 & \beta_1 & c_1 \\ 0 & 1 & \alpha_2 & \beta_2 & c_2 \\ 0 & 0 & 0 & 0 & 0 \end{array} \right) \rightarrow \left(\begin{array}{ccccc|c} 1 & 0 & \alpha_1 & \beta_1 & c_1 \\ 0 & 1 & \alpha_2 & \beta_2 & c_2 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \end{array} \right) \xrightarrow{(-1)} \left(\begin{array}{ccccc|c} 1 & 0 & \alpha_1 & \beta_1 & c_1 \\ 0 & 1 & \alpha_2 & \beta_2 & c_2 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \end{array} \right).$$

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 15.10(2) also

$$\left(\begin{array}{cc|c} 0 & 1 & 1 \\ 0 & 0 & 0 \end{array} \right) \rightarrow (01|1) \xrightarrow{(-1)} \left(\begin{array}{cc|c} -1 & 0 & 0 \\ 0 & 1 & 1 \end{array} \right)$$

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 $\text{Rang } A$.

Beweis. Man betrachte die ZNF von A . □

Bemerkung: Es gilt $\text{Rang } A = \dim \text{Bild } A$.

Satz: Sind $A \in \mathbb{K}^{n \times m}$ und $\vec{b} \in \mathbb{K}^n$, so gilt:

$$A\vec{x} = \vec{b} \text{ ist lösbar} \iff \text{Rang } A = \text{Rang } (A | \vec{b}),$$

d.h. 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

$$m = \dim \text{Bild } A + \dim \text{Kern } A = \text{Rang } A + \dim \text{Kern } A.$$

Man vergleiche mit 15.11 (insbesondere mit der Form der Matrix, die beim (-1) -Ergänzungsstrick entsteht).

Beispiele: (1) $A = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix}$. Hier ist $n = m = 3$ und $\text{Rang } A = 3$, da die Zeilen (oder auch die Spalten) linear unabhängig sind. Somit ist $\dim \text{Kern } A = 3 - 3 = 0$ und $\text{Kern } A = \{\vec{0}\}$. Ablesen lässt sich das natürlich auch an einer ZSF:

$$\begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 1 & 0 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix}.$$

(2) $A = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$. Hier ist $n = m = 3$ und $\text{Rang } A = 2$, da die ersten beiden Spalten linear unabhängig sind, die dritte Spalte jedoch Summe der ersten beiden. Somit ist $\dim \text{Kern } A = 3 - 2 = 1$. Durch Probieren findet man $\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} \in \text{Kern } A$. Folglich ist

$$\text{Kern } A = \text{lin}\left\{\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}\right\}.$$

Zum selben Ergebnis kommt man mit einer ZNF und anschließender (-1) -Ergänzung:

$$\begin{pmatrix} 1 & 1 & 2 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 1 \\ 0 & -1 & -1 \\ 0 & 1 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 1 & 2 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix} \\ \rightarrow \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix} \xrightarrow{(-1)} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & -1 \end{pmatrix}.$$

15.14. Lineare Abbildungen als Matrizen: Seien V, W endlich-dimensionale $\mathbb{K}$ -Vektorräume und $\phi : V \rightarrow W$ eine lineare Abbildung. In V und W seien Basen $b_1, b_2, \dots, b_m$ bzw. $c_1, c_2, \dots, c_n$ gegeben (insbesondere gilt $\dim V = m$ und $\dim W = n$). Dann gibt es genau eine Matrix $A \in \mathbb{K}^{n \times m}$ mit der Eigenschaft

$$\phi\left(\sum_{k=1}^m \alpha_k b_k\right) = \sum_{j=1}^n \beta_j c_j \iff A \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{pmatrix} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}.$$

Durch die Matrix A werden also die Koordinatentupel $(\alpha_1, \alpha_2, \dots, \alpha_m) \in \mathbb{K}^m$ von Vektoren $v \in V$ bzgl. der Basis $b_1, b_2, \dots, b_m$ auf die Koordinatentupel $(\beta_1, \beta_2, \dots, \beta_n) \in \mathbb{K}^n$ von $\phi(v)$ bzgl. der Basis $c_1, c_2, \dots, c_n$ abgebildet.

Diese Matrix A heißt *Darstellung von ϕ bzgl. der Basen $b_1, b_2, \dots, b_m$ und $c_1, c_2, \dots, c_n$* .

Man erhält die Matrix A wie folgt:

In der k -ten Spalte von A stehen die Koordinaten des Bildes des k -ten Basisvektors b_k , also die Koordinaten von $\phi(b_k)$ ($k = 1, \dots, m$).

Beispiele: (1) Sei $A \in \mathbb{K}^{n \times m}$. Die Darstellungsmatrix von $\phi_A : \mathbb{K}^m \rightarrow \mathbb{K}^n$, $\phi_A(\vec{x}) := A\vec{x}$, bzgl. der Standardbasen in $\mathbb{K}^m$ und $\mathbb{K}^n$ ist A .

(2) Die lineare Abbildung $\phi : \mathbb{K}^3 \rightarrow \mathbb{K}^3$ 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 im Argumentraum von ϕ und die Standardbasis 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 $\text{Id} : \mathbb{K}^n \rightarrow \mathbb{K}^n$, $\vec{x} \mapsto \vec{x}$, 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 $\vec{b}_k$, d.h. 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

$$C := \frac{1}{2} \begin{pmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \\ -1 & 1 & 1 \end{pmatrix}.$$

(4) Es sei $V = \text{lin}\{1, \sin, \cos\}$. Dann ist V ein UVR von $C(\mathbb{R})$ mit $\dim V = 3$. Betrachte die lineare Abbildung $\phi : V \rightarrow V$, $\phi(f) = f'$. Bzgl. der Basis $1, \sin, \cos$ im Argument- und Zielraum erhalten wir die Darstellungsmatrix

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

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 $\phi_A : \mathbb{K}^m \rightarrow \mathbb{K}^n$, $\phi_A(\vec{x}) = A\vec{x}$, und $\phi_B : \mathbb{K}^q \rightarrow \mathbb{K}^m$, $\phi_B(\vec{x}) = B\vec{x}$, linear, und $\phi_A \circ \phi_B : \mathbb{K}^q \rightarrow \mathbb{K}^n$ 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 $A = (a_{jk})_{j=1, k=1}^n$ und $B = (b_{kl})_{k=1, l=1}^m$ die Matrix AB Einträge $(c_{jl})_{j=1, l=1}^n$ hat mit

$$c_{jl} = \sum_{k=1}^m a_{jk} b_{kl} \quad (j = 1, \dots, n, \quad l = 1, \dots, q).$$

Beweis. Für $\vec{x} = (x_1, x_2, \dots, x_q) \in \mathbb{K}^q$ gilt

$$\begin{aligned} A(B\vec{x}) &= (a_{jk})_{j,k} \left(\sum_{l=1}^q b_{kl} x_l \right)_k = \left(\sum_{k=1}^m a_{jk} \sum_{l=1}^q b_{kl} x_l \right)_j \\ &= \left(\sum_{l=1}^q \left(\sum_{k=1}^m a_{jk} b_{kl} \right) x_l \right)_j = \left(\sum_{k=1}^m a_{jk} b_{kl} \right)_{j,l} (x_l)_l. \end{aligned}$$

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**. $\square$

Bemerkung: Man erhält den (j, l) -ten Eintrag von AB , indem man die j -te Zeile von 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:

$$(A_1 + \alpha A_2)B_1 = A_1 B_1 + \alpha A_2 B_1, \quad A_1(B_1 + \alpha B_2) = A_1 B_1 + \alpha A_1 B_2,$$

d.h. für festes $B \in \mathbb{K}^{m \times q}$ ist

$$\mathbb{K}^{n \times m} \rightarrow \mathbb{K}^{n \times q}, \quad A \mapsto AB$$

linear, und für festes $A \in \mathbb{K}^{n \times m}$ ist

$$\mathbb{K}^{m \times q} \rightarrow \mathbb{K}^{n \times q}, \quad B \mapsto AB$$

linear. Also ist die Abbildung $\mathbb{K}^{n \times m} \times \mathbb{K}^{m \times q} \rightarrow \mathbb{K}^{n \times q}$, $(A, B) \mapsto AB$ *bilinear* (d.h. linear in jedem Argument).

Außerdem ist das Produkt von Matrizen assoziativ: Sind A, B, C Matrizen und sind AB und BC **definiert**, so gilt

$$A(BC) = (AB)C \quad =: ABC.$$

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 $\phi : \mathbb{K}^3 \rightarrow \mathbb{K}^3$ 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 $\text{Id} : \mathbb{K}^3 \rightarrow \mathbb{K}^3$ 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). Es gilt nämlich für $\vec{x} \in \mathbb{K}^3$:

$$\vec{x} = \sum_{k=1}^3 \alpha_k \vec{b}_k = \sum_{k=1}^3 \beta_k \vec{e}_k \mapsto \phi(\vec{x}) = \sum_{k=1}^3 \gamma_k \vec{e}_k$$

mit

$$\vec{\beta} = B\vec{\alpha} \text{ und } \vec{\gamma} = F\vec{\beta}, \text{ also } \vec{\gamma} = FB\vec{\alpha}.$$

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 $(n \times n)$ -Einheitsmatrix (wir schreiben kurz I , wenn n aus dem Zusammenhang klar ist).

Es gilt dann

$$\forall A \in \mathbb{K}^{n \times n} : IA = AI = A$$

und die zugehörige Abbildung $\phi_I : \mathbb{K}^n \rightarrow \mathbb{K}^n$ ist die Identität $\vec{x} \mapsto \vec{x}$.

Definition: Eine Matrix $A \in \mathbb{K}^{n \times n}$ heißt *regulär*, falls $\text{Rang } A = n$ 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:

$$B = BI = B(A\tilde{B}) = (BA)\tilde{B} = I\tilde{B} = \tilde{B}. \quad (*)$$

Diese Matrix B heißt *Inverse von A* (oder *zu A inverse Matrix*) und wird mit A^{-1} bezeichnet. Die Abbildung $\mathbb{K}^n \rightarrow \mathbb{K}^n$, $\vec{x} \mapsto A^{-1}\vec{x}$ ist die Umkehrabbildung von $\mathbb{K}^n \rightarrow \mathbb{K}^n$, $\vec{x} \mapsto A\vec{x}$.

Rechenregeln: Seien $A, B \in \mathbb{K}^{n \times n}$ invertierbar. Dann sind A^{-1} und AB invertierbar, und es gilt:

$$(A^{-1})^{-1} = A \quad \text{und} \quad (AB)^{-1} = B^{-1}A^{-1}.$$

Die zweite Regel ist die *Hemd-Jacken-Regel* (vgl. 3.6).

Satz: Sei $A \in \mathbb{K}^{n \times n}$. Dann gilt:

$$A \text{ ist regulär} \iff \text{Bild } A = \mathbb{K}^n \iff \text{Kern } A = \{\vec{0}\} \iff A \text{ ist invertierbar.}$$

Für die induzierte lineare Abbildung $\phi_A : \mathbb{K}^n \rightarrow \mathbb{K}^n$ 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 $\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$ (vgl. Beispiel in 15.15).

Beweis. Gilt $AB = I$, so ist $\text{Kern } B = \{\vec{0}\}$ und $\text{Bild } A = \mathbb{K}^n$, und nach dem Satz sind A, B invertierbar. Somit ist dann auch

$$B = IB = A^{-1}AB = A^{-1}I = A^{-1}.$$

□

Bemerkung: Ist V ein unendlich-dimensionaler Vektorraum und sind $\phi, \psi : V \rightarrow V$ lineare Abbildungen mit $\phi \circ \psi = \text{Id}_V$, so sind ϕ und ψ nicht notwendig invertierbar. Betrachte zum Beispiel $V = C^\infty([0, 1])$ mit $\phi, \psi : V \rightarrow V$ definiert durch

$$\phi(f) = f', \quad \psi(f)(x) = \int_0^x f(t)dt \quad (x \in [0, 1]).$$

Dann gilt $\phi(\psi(f)) = f$ ($f \in V$). Aber ϕ ist nicht injektiv und ψ ist nicht surjektiv.

Berechnung von A^{-1} : Man löse für jedes $l \in \{1, 2, \dots, n\}$ das Gleichungssystem $A\vec{x} = \vec{e}_l$. Die Lösung ist die l -te Spalte von A^{-1} . Man kann diese Gleichungssysteme auch simultan lösen.

Beispiel: Sei $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \in \mathbb{K}^{2 \times 2}$ invertierbar und $a \neq 0$. Wir setzen $\delta := ad - bc$ und sehen an der dritten Matrix, dass $\delta \neq 0$ ist.

$$\begin{pmatrix} a & b & | & 1 & 0 \\ c & d & | & 0 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & b/a & | & 1/a & 0 \\ 0 & d - bc/a & | & -c/a & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & b/a & | & 1/a & 0 \\ 0 & \delta/a & | & -c/a & 1 \end{pmatrix} \\ \rightarrow \begin{pmatrix} 1 & b/a & | & 1/a & 0 \\ 0 & 1 & | & -c/\delta & a/\delta \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & | & \frac{1}{a} + \frac{bc}{\delta a} & -b/\delta \\ 0 & 1 & | & -c/\delta & a/\delta \end{pmatrix}.$$

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 $\text{Id} : \mathbb{K}^n \rightarrow \mathbb{K}^n$ 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 $\text{Id} : \mathbb{K}^n \rightarrow \mathbb{K}^n$ 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 $\phi : \mathbb{K}^n \rightarrow \mathbb{K}^n$ 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 $\phi : V \rightarrow V$ 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 $\text{Id} : V \rightarrow V$ 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:

$$\begin{pmatrix} -1 & -1 & 1 & | & 1 & 0 & 0 \\ 2i & -2i & 0 & | & 0 & 1 & 0 \\ 1 & 1 & 1 & | & 0 & 0 & 1 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 1 & -1 & | & -1 & 0 & 0 \\ 0 & -4i & 2i & | & 2i & 1 & 0 \\ 0 & 0 & 2 & | & 1 & 0 & 1 \end{pmatrix} \\ \rightarrow \begin{pmatrix} 1 & 1 & 0 & | & -1/2 & 0 & 1/2 \\ 0 & -4i & 0 & | & i & 1 & -i \\ 0 & 0 & 1 & | & 1/2 & 0 & 1/2 \end{pmatrix} \rightarrow \begin{pmatrix} 1 & 0 & 0 & | & -1/4 & -i/4 & 1/4 \\ 0 & 1 & 0 & | & -1/4 & i/4 & 1/4 \\ 0 & 0 & 1 & | & 1/2 & 0 & 1/2 \end{pmatrix}$$

(man überzeuge sich am Ende evtl. 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

$$b'_1 - b_1 = (-1-i)b_1 + ib_3, \quad b'_2 - b_2 = (-1+i)b_2 - ib_3, \quad b'_3 - b_3 = -\frac{i}{2}b_1 + \frac{i}{2}b_2 - b_3,$$

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).

Diagramm zum Basiswechsel in diesem Beispiel:

$$\begin{array}{ccccccc} V & \xrightarrow{\text{Id}_V} & V & \xrightarrow{\phi} & V & \xrightarrow{\text{Id}_V} & V \\ \downarrow b_j & & \downarrow e_j & & \downarrow e_j & & \downarrow b_j \\ \mathbb{K}^3 & \xrightarrow{B} & \mathbb{K}^3 & \xrightarrow{F} & \mathbb{K}^3 & \xrightarrow{B^{-1}} & \mathbb{K}^3 \end{array} \quad \vec{x} \mapsto B^{-1}FB\vec{x}$$