Höhere Mathematik I für die Fachrichtungen Elektroingenieurwesen, Physik und Geodäsie Wintersemester 2008/09

Peer Christian Kunstmann
Institut für Analysis, Universität Karlsruhe
Kaiserstr. 89, D – 76128 Karlsruhe, Germany
e-mail: peer.kunstmann@math.uni-karlsruhe.de

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

1 Logische Grundlagen

1.1 Aussagen

Definition: Eine *Aussage* ist ein Satz, der entweder **wahr (w)** oder **falsch (f)** sein kann.

Beispiele: i) Der Mond ist ein grüner Käse. (f)
ii) 2 ist eine Primzahl. (w)
iii) Gehen sie geradeaus und dann hinten links! (keine Aussage)

Bemerkung: Wir werden uns auf **mathematische Aussagen** konzentrieren.

Aussagen bezeichnen wir im folgenden mit A, B, C, ...

1.2 Verknüpfung von Aussagen

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

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

A	w	w	f	f
B	w	f	w	f
$A \wedge B$	w	f	f	f

$A \vee B$ (*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, dh es ist nicht ausgeschlossen, dass beide Aussagen A und B wahr sind.

Negation $\neg A$

A	w	f
$\neg A$	f	w

Man sagt: "non A" oder "nicht A".

Implikation $A \Rightarrow B$

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

Man sagt: "wenn A, dann B", "A impliziert B", "aus A folgt B".
Bemerkung: Aus Falschem folgt Beliebiges (*ex falso quodlibet*).
Beispiel: Aus $1 = 2$ folgt: ich bin der Papst. (w)

Äquivalenz $A \Leftrightarrow B$

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

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

1.3 Regeln

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

$\neg(\neg A) \quad \Leftrightarrow A$; 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 ich nicht der Papst bin, dann ist $1 \neq 2$. (w)

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

Tertium non datur: $A \vee \neg A$.

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 in eine Aussage übergeht.

Beispiel: x ist eine Primzahl.

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

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

Negation:
$\neg(\forall x : A(x)) \Leftrightarrow (\exists x : \neg A(x))$ und $\neg(\exists x : A(x)) \Leftrightarrow (\forall x : \neg A(x))$.

In den allermeisten Fällen werden Quantoren *eingeschränkt* und beziehen sich dann nur auf gewisse Objekte, z.B. $\forall x$ mit $A(x)$: $B(x)$. Negation ist dann $\exists x$ mit $A(x)$: $\neg B(x)$.

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

2 Mengen

2.1 Der Begriff der Menge

“Definition”: “Eine *Menge* ist die Zusammenfassung wohlbestimmter, wohlunterschiedener Objekte der Anschauung oder des Denkens zu einem neuen Ganzen.”

Diese Objekte heißen *Elemente* der Menge.
$x \in M$ bedeutet: x ist Element der Menge M, dh x gehört zu M.
$x \notin M$ bedeutet. x gehört nicht zu M, dh $x \notin M :\Leftrightarrow \neg(x \in M)$.

Für jede Menge M und jedes x gilt entweder $x \in M$ oder $x \notin M$.

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

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

2.2 Beziehungen zwischen Mengen

Seien M_1, M_2 Mengen.

Definition: $M_1 \subset M_2 :\Leftrightarrow \forall x : x \in M_1 \Rightarrow x \in M_2$ (bzw. $\forall x \in M_1 : x \in M_2$).
“M_1 ist Teilmenge von M_2”. Gelegentlich schreibe ich auch $\subseteq$ statt $\subset$. Gleichheit ist bei “$\subset$” nicht ausgeschlossen, dh es gilt $M \subset M$ für jede Menge M.

Für $M_1 \subset M_2$ und $M_1 \neq M_2$ schreibe ich der Deutlichkeit halber $M_1 \overset{\subset}{\neq} M_2$.

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

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

2.3 Operationen mit Mengen

Seien M_1, M_2, M_3 und Q Mengen.

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

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

Regeln: Wegen 1.3 gelten:
Kommutativität $M_1 \cap M_2 = M_2 \cap M_1$, $M_1 \cup M_2 = M_2 \cup M_1$.
Assoziativität $M_1 \cap (M_2 \cap M_3) = (M_1 \cap M_2) \cap M_3$, $M_1 \cup (M_2 \cup M_3) = (M_1 \cup M_2) \cup M_3$.
Distributivität

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

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

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

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

de Morgansche Regeln: Wegen 1.3 gelten auch

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

2.4 Die leere Menge

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

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

2.5 Die Potenzmenge

Ist M eine Menge, so heißt die Menge aller Teilmengen von M

$$\mathrm{Pot}(M) := \{N : N \subset M\}$$

die *Potenzmenge von* M.

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

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

2.6 Das kartesische Produkt

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

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

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

Beispiele:
i) $M_1 = \{0, 1\}$, $M_2 = \{1, 2, 3\}$, $M_1 \times M_2 = \{(0, 1), (0, 2), (0, 3), (1, 1), (1, 2), (1, 3)\}$.
ii) $M = \{0, 1\}$, $M^3 = \{(0, 0, 0), (0, 0, 1), (0, 1, 0), (0, 1, 1), (1, 0, 0), (1, 0, 1), (1, 1, 0), (1, 1, 1)\}$.

3 Funktionen

3.1 Zum Begriff der Funktion

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

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

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

Graph von f $\{(x, f(x)) : x \in X\}$.

Eine Funktion von X nach Y ist also eine Teilmenge M von $X \times Y$, die folgende Eigenschaften hat:

$$\begin{array}{l} \forall x \in X \exists y \in Y : (x, y) \in M \\ \forall x \in X \forall y_1, y_2 \in Y : (x, y_1) \in M \wedge (x, y_2) \in M \;\Rightarrow\; y_1 = y_2. \end{array}$$

Für eine Funktion $f : X \to Y$ heißt X *Definitionsbereich* und Y *Wertebereich* von f. Für $A \subset X$ heißt $f(A) := \{f(x) : x \in A\}$ *Bild von A unter f*, und für $B \subset Y$ heißt $f^{-1}(B) := \{x \in X : f(x) \in B\}$ *Urbild von B unter f*. Insbesondere heißt $f(X) = \{f(x) : x \in X\}$ *Bild von f* (Menge aller $y \in Y$, die von f getroffen werden).

Im Beispiel oben ist $f(\mathbb{N}) = \{x \in \mathbb{N} : x \text{ ist ungerade}\}$ und etwa $f^{-1}(\{1, 2, 3, 4, 5\}) = f^{-1}(\{1, 3, 5\}) = \{1, 2, 3\}$.

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

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

(b) f heißt *injektiv*, falls gilt $\forall x_1, x_2 \in X : x_1 \neq x_2 \Rightarrow f(x_1) \neq f(x_2)$.

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

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

3.3 Komposition

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

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

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

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

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

dh die Hintereinanderausführung von Funktionen ist assoziativ.

3.4 Die Umkehrabbildung

Ist $f : X \to Y$ eine bijektive Funktion, so definiert

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

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

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

Ende Woche 1

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

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

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

Beweis: Die Funktion $g \circ f$ ist injektiv, also ist f injektiv. Die Funktion $f \circ g$ ist surjektiv, also ist f surjektiv. Gezeigt: f ist bijektiv. Nach 3.4 existiert die Umkehrabbildung $f^{-1} : Y \to X$. 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.

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

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

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

[Das ist klar. Die Formel heißt “Hemd-Jacken-Regel”.]

4 Die reellen Zahlen

Grundmenge der Analysis ist die Menge $\mathbb{R}$ der *reellen Zahlen*. Wir führen diese Menge durch *Axiome* ein dh durch grundlegende Eigenschaften, aus denen sich alle **alle** weiteren Rechenregeln herleiten lassen.

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

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

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 "·" herleiten. Diese werden von nun an als bekannt vorausgesetzt.

Beispiele: (1) Die Null in $(A2)$ ist eindeutig, ebenso die Eins in $(A6)$. [Ist $\tilde{0} \in \mathbb{R}$ mit $\forall a \in \mathbb{R} : a + 0 = a$, so folgt wegen $(A2)$ und $(A4)$: $0 = 0 + \tilde{0} = \tilde{0} + 0 = \tilde{0}$.]

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

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

(4) $\forall a \in \mathbb{R}$: $-a = (-1) \cdot a$. [Es gilt (nach $(A6)$, $(A9)$, und (3)): $a + a \cdot (-1) = a \cdot 1 + a \cdot (-1) = a \cdot (1 + (-1)) = a \cdot 0 = 0$, also $-a = a \cdot (-1) = (-1) \cdot a$.]

(5) $\forall a \in \mathbb{R}$: $a^2 = (-a)^2$, wobei $a^2 := a \cdot a$. [Es ist, nach (4), $(A5)$, $(A8)$ und (2): $(-a)^2 = (-a) \cdot (-1) \cdot a = (-(-a)) \cdot a = a^2$.]

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

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

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) $\forall a \in \mathbb{R}$: $a^2 \geq 0$.
Beweis: Sei $a \in \mathbb{R}$. Fall 1, $a \geq 0$: Dann gilt $a \cdot a \geq 0 \cdot a$ nach $(A14)$ und somit $a^2 \geq 0$ nach 4.1(3).
Fall 2, $a < 0$: 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 $ac \geq bc$. [$c \leq 0 \Rightarrow -c \geq 0$ (siehe Beweis von (1)). Nach $(A14)$ ist dann $a(-c) \leq b(-c)$. Beachtet man $a(-c) = -ac$ und $b(-c) = -bc$ und addiert $ac + bc$ zu der Ungleichung ($(A13)$!), so erhält man $bc \leq ac$.]

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

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

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

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

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

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

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

Regeln: Seien $a, b, c \in \mathbb{R}$. Dann gilt:
(1) $|a| \geq 0$;
(2) $|a| = 0 \Leftrightarrow a = 0$;
(3) $|ab| = |a| \cdot |b|$;
(4) $\pm a \leq |a|$ und $[|a| \leq c \Leftrightarrow a \leq c$ und $-a \leq c]$;
(5) $|a + b| \leq |a| + |b|$ *Dreiecksungleichung*;
(6) $||a| - |b|| \leq |a - b|$ *umgekehrte Dreiecksungleichung.*

[(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 $||a| -| b|| \leq |a - b|$.]

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

4.4 Supremum und Infimum

Definition: Sei $M \subset \mathbb{R}$ mit $M \neq \emptyset$.

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

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

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

Wegen $(A11)$ sind $\max M$ und $\min M$ im Falle der Existenz *eindeutig bestimmt*.

Beispiele: $[1, 2]$ ist nach oben und nach unten beschränkt, es ist $1 = \min M$ und $2 = \max M$.
$(1, \infty)$ ist nach unten, aber nicht nach oben beschränkt und hat kein Minimum.

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

Nach $(A11)$ sind Supremum und Infimum im Falle der Existenz eindeutig bestimmt.

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

Beispiele: (1) $M = [1, 2)$. M ist nach unten und nach oben beschränkt. Es ist $1 = \min M = \inf M$, M hat kein Maximum, und es ist $\sup M = 2$.
Beweis: 2 ist OS von M. Zeige: es gibt keine echt kleinere obere Schranke. Sei $\gamma < 2$. Zeige: γ ist nicht OS von M. Falls $\gamma \leq 1$ ist, so gilt $\gamma < \frac{3}{2} \in M$, also ist γ keine OS von M. Falls $\gamma > 1$ ist, so ist $\gamma < \frac{\gamma+2}{2} \in M$ und γ ist keine OS von M.

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

4.5 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.

[Sei $M \subset \mathbb{R}$ mit $M \neq \emptyset$. Setze $-M := \{-x : x \in M\}$.
Vorbemerkung: Dann gilt γ ist US von $M \Leftrightarrow -\gamma$ ist OS von $-M$.
Da $M \neq \emptyset$ untere Schranken hat, ist $-M \neq \emptyset$ nach oben beschränkt, hat also ein Supremum $s := \sup(-M)$. Nach der Vorbemerkung ist $-s$ US von M, und für jede US γ von M ist $-\gamma$ eine OS von $-M$, also $s \leq -\gamma$, dh $\gamma \leq -s$. Somit ist $-s$ größte US von M, dh $-s = \inf M$.]

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

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

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

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

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

(3) Sei A nach oben [unten] beschränkt und γ eine OS [US] von A. Dann gilt:

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

Beweis: (1) und (2) sind leicht.
Zu (3): "$\Rightarrow$": Sei $\gamma = \sup A$ und $\varepsilon > 0$. Dann ist $\gamma - \varepsilon$ keine OS von A.
"$\Leftarrow$" (Kontraposition): Sei $\gamma \neq \sup A =: \tilde{\gamma}$. Dann $\gamma > \tilde{\gamma}$, da γ OS von A ist, und $\varepsilon := \gamma - \tilde{\gamma} > 0$. Für jedes $x \in A$ gilt dann $x \leq \tilde{\gamma} = \gamma - \varepsilon$. Wir haben gezeigt: $\exists \varepsilon > 0 \forall x \in A : x \leq \gamma - \varepsilon$, dh $\neg(\forall \varepsilon > 0 \exists x \in A : x > \gamma - \varepsilon)$.]

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

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

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

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

Einschub: Sind $\emptyset \neq X$ eine Menge und $\mathcal{A} \subset \mathrm{Pot}(X)$, so setzt man

$$\begin{array}{rcl} \bigcup \mathcal{A} & := & \{x \in X : \exists A \in \mathcal{A} : x \in A\} \quad \text{und, falls } \mathcal{A} \neq \emptyset, \\ \bigcap \mathcal{A} & := & \{x \in X : \forall A \in \mathcal{A} : x \in A\}. \end{array}$$

Man schreibt auch $\bigcup \mathcal{A} = \bigcup\limits_{A \in \mathcal{A}} A$, $\bigcap \mathcal{A} = \bigcap\limits_{A \in \mathcal{A}} A$.

Bemerkung: Es ist somit $\mathbb{N} = \bigcap \{A \subset \mathbb{R} : \ A \text{ ist IM } \}$.

Satz: (1) $\mathbb{N}$ ist eine IM [$\mathbb{N}$ ist die kleinste IM].

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

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

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

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

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

(3) folgt sofort aus (2).

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

4.8 Vollständige Induktion

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

Beweis: Nach Voraussetzung gilt $A \subset \mathbb{N}$. Da A IM ist, gilt $\mathbb{N} \subset A$.

Beweisverfahren durch Induktion

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

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

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

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

Beispiel: Für alle $n \in \mathbb{N}$ gilt $\underbrace{n \geq 1}_{=:A(n)}$.
Beweis durch Induktion nach n:
Induktionsanfang (IA): Es gilt $1 \geq 1$, dh $A(1)$ ist wahr.
Induktionsschluss (IS): Sei $n \in \mathbb{N}$. Es gelte $A(n)$, dh es gelte $n \geq 1$ (Induktionsvoraussetzung (IV)). Dann ist $n + 1 \geq 1 + 1$ nach (IV) und $1 + 1 \geq 1 + 0 = 1$ nach 4.2 und 4.1, also $n + 1 \geq 1$ und $A(n + 1)$ ist wahr.

Wir setzen die natürlichen Zahlen ab jetzt als bekannt voraus.

Definition durch Rekursion (bzw. durch Induktion)

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

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

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

(2) Summenzeichen $\sum$: Seien $a_1, a_2, \ldots \in \mathbb{R}$. Setze $\sum_{j=1}^{1} a_j := a_1$ und für jedes $n \in \mathbb{N}$: $\sum_{j=1}^{n+1} a_j := (\sum_{j=1}^{n} a_j) + a_{n+1}$. Dann ist $\sum_{j=1}^{n} a_j = a_1 + a_2 + \ldots + a_n$ für jedes $n \in \mathbb{N}$. Die leere Summe ist $\sum_{j=1}^{0} a_j := 0$.

(3) Produktzeichen $\prod$: Seien $a_1, a_2, \ldots \in \mathbb{R}$. Setze $\prod_{j=1}^{1} a_j := a_1$ und für jedes $n \in \mathbb{N}$: $\prod_{j=1}^{n+1} a_j := (\prod_{j=1}^{n} a_j) \cdot a_{n+1}$. Dann ist $\prod_{j=1}^{n} a_j = a_1 \cdot a_2 \cdot \ldots \cdot a_n$ für jedes $n \in \mathbb{N}$. Das leere Produkt ist $\prod_{j=1}^{0} a_j := 1$.

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

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.* Ende Woche 2

Bemerkung: Jede nichtleere nach unten beschränkte Teilmenge von $\mathbb{Z}$ hat ein Minimum.

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

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

[Beweis des Zusatzes: Setze $p := \min\{m \in \mathbb{Z} : m > x\}$. Dann gilt $x < p$ und $p - 1 \leq x$, also $p \leq x + 1 < y$.

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

4.10 Binomialkoeffizienten

Für $n, k \in \mathbb{N}_0$ mit $k \leq n$ setzt man $\binom{n}{k} := \frac{n!}{k!(n-k)!}$ ("n über k").

Etwa: $\binom{n}{0} = 1 = \binom{n}{n}$ für alle $n \in \mathbb{N}_0$, $\binom{4}{2} = \frac{4 \cdot 3}{2} = 6$.

Lemma: Für $1 \leq k \leq n$ gilt

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

Beweis: Es gilt

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

4.11 Potenzen

Für $a \in \mathbb{R}$ setzt man $a^0 := 1$ und $a^{n+1} := a^n \cdot a$ für alle $n \in \mathbb{N}_0$, also $a^n = \prod_{=1}^{n} a = \underbrace{a \cdot a \cdot \ldots \cdot a}_{}$, und für $a \neq 0$ setzt man $a^{-n} := \frac{1}{a^n}$.
n Faktoren
Es gelten die bekannten Rechenregeln, also etwa $(a^m)^n = a^{m \cdot n}$ und $a^m \cdot a^n = a^{m+n}$.

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

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

[Übungsaufgabe]

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

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

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

Beweis durch Induktion nach n: IA $n = 0$ ist klar.
IS Sei $n \in \mathbb{N}$ mit $(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}$. Dann gilt

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

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

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

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

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

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

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

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

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

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

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

[Übungsaufgabe]

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

Satz: Zu jedem $a \in \mathbb{R}$ mit $a \geq 0$ gibt es genau ein $b \in \mathbb{R}$ mit $b \geq 0$ und $b^n = a$.
Bezeichnung: $b = \sqrt[n]{a}$, “n-te Wurzel aus a”.

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

Folgerung: Für alle $a, b \geq 0$ gilt:

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

Beweis des Satzes: Der Fall $a = 0$ ist klar. Sei also $a > 0$. Setze $M := \{x > 0 : x^n < a\}$. Dann ist $1 + a$ OS von M und nach 4.11(4) ist M nichtleer. Also existiert $b := \sup M$.

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

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

[Es ist

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

und die andere Ungleichung zeigt man ähnlich.]

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

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

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

5 Die komplexen Zahlen

5.1 Konstruktion

Auf $\mathbb{R}^2$ erklären wir zwei Verknüpfungen "+" und "$*$" durch

$$\begin{aligned} (x_1, y_1) + (x_2, y_2) &:= (x_1 + x_2, y_1 + y_2) \qquad \text{und} \\ (x_1, y_1) * (x_2, y_2) &:= (x_1x_2 - y_1y_2, x_2y_1 + x_1y_2). \end{aligned}$$

Man kann nun nachrechnen, dass die Körperaxiome $(A1) - (A9)$ für diese Verknüpfungen gelten (wobei "$*$" die Rolle der Multiplikation übernimmt, vergleiche die Übungsaufgabe).

Es gilt $(x_1, 0) * (x_2, 0) = (x_1x_2, 0)$ und man kann das Paar $(x, 0)$ mit der reellen Zahl x identifizieren. Weiter ist $(0, 1) * (0, 1) = (-1, 0)$, dh $(0, 1)$ ist eine "Zahl", deren Quadrat $= -1$ ist. Man setzt nun $i := (0, 1)$ und schreibt $x + iy$ statt (x, y). Es ist dann

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

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

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

$$(x_1 + iy_1)(x_2 + iy_2) = x_1x_2 + i(y_1x_2 + x_1y_2) + i^2y_1y_2 = (x_1x_2 - y_1y_2) + i(x_2y_1 + x_1y_2),$$

vergleiche die Definition von "$*$" oben.

5.2 Konjugation und Betrag

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

Rechenregeln: Für alle $w, z \in \mathbb{C}$ gilt:
$\overline{(\overline{z})} = z$,
$|\overline{z}| = |z|$,
$\overline{z + w} = \overline{z} + \overline{w}$,
$\overline{z \cdot w} = \overline{z} \cdot \overline{w}$,
$Re\, z = \frac{1}{2}(z + \overline{z})$ und $Im\, z = \frac{1}{2i}(z - \overline{z})$,
$\max\{|Re\, z|, |Im\, z|\} \leq |z| \leq |Re\, z| + |Im\, z|$ [denn $\max\{|x|, |y|\} \leq \sqrt{x^2 + y^2} \leq |x| + |y|$ für $x, y \in \mathbb{R}$],
$|z \cdot w| = |z| \cdot |w|$ [denn $|zw|^2 = zw\overline{zw} = z\overline{z}w\overline{w} = |z|^2|w|^2$],
$|z + w| \leq |z| + |w|$ (Dreiecksungleichung) und $||z| -| w|| \leq |z - w|$.

[Es ist nämlich

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

woraus die Dreiecksungleichung durch Wurzelziehen folgt.]

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

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

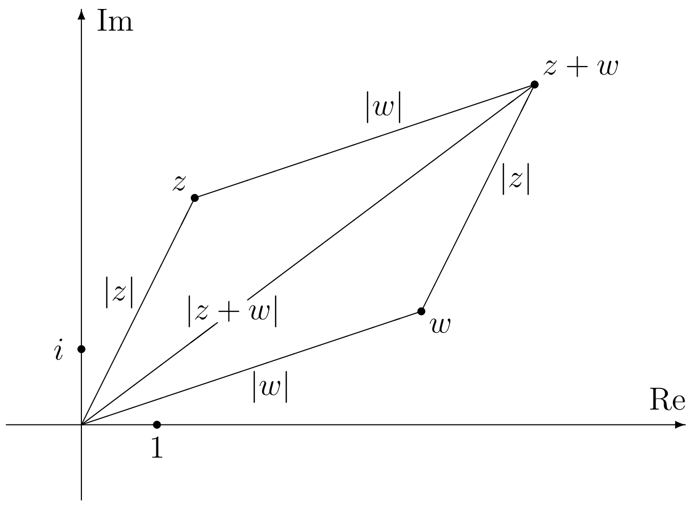
5.3 Zur anschaulichen Vorstellung

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

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

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

Dreiecksungleichung:



5.4 Polynome

Ein Polynom $p(z)$ mit komplexen Koeffizienten ist ein Ausdruck der Form $p(z)=a_nz^n+a_{n-1}z^{n-1}+\ldots+a_2z^2+a_1z+a_0$ mit $n\in\mathbb{N}_0$ und $a_j\in\mathbb{C}$ für alle $j\in\{0,1,\ldots,n\}$. Das Polynom heißt *reell*, wenn alle Koeffizienten a_j reell sind.

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

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

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

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

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

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

Satz: Ist $p(z)$ Polynom vom Grad $n \geq 1$ und ist z_0 Nullstelle von p, so gibt es ein Polynom $q(z)$ vom Grad $n - 1$ mit

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

Dabei heißt $z - z_0$ *Linearfaktor*.

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

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

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

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

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

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

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

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

6 Folgen und Konvergenz

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

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

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

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

Der Begriff der Konvergenz ist für die Analysis von zentraler Bedeutung.

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

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

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

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

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

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

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

Wir haben $\lim_{n\to\infty} \frac{1}{n} = 0$ nachgewiesen. Ende Woche 3

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

(2) Für jede Zahlenfolge (a_n) und jedes a gilt: $\lim a_n = a \Leftrightarrow \lim |a_n - a| = 0$. Insbesondere ist für $a = 0$: $\lim a_n = 0 \Leftrightarrow \lim |a_n| = 0$.

(3) Die Folge $(a_n)_{n\in\mathbb{N}} := ((-1)^n)_{n\in\mathbb{N}}$ ist divergent: Sei $a \in \mathbb{R}$. Setze $\varepsilon := \begin{cases} 1/2, & \text{falls } a \in \{1, -1\} \\ 1/2 \cdot \min\{|1-a|, |-1-a|\}, & \text{falls } a \notin \{1, -1\} \end{cases}$. Dann gilt $|a_n - a| \geq \varepsilon$ für unendlich viele $n \in \mathbb{N}$. Folglich konvergiert (a_n) nicht gegen a. Da a beliebig war, ist (a_n) divergent.

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

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

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

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

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

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

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

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

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

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

Bei Konvergenzfragen kommt also auf endlich viele Folgenglieder nicht an. Für $p \in \mathbb{Z}$ bezeichnet man auch $(a_n)_{n=p}^\infty$ als Folge.

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

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

(c) Ist (z_n) eine komplexe Zahlenfolge, so ist (z_n) genau dann konvergent, wenn die reellen Zahlenfolgen $(Re\, z_n)$ und $(Im\, z_n)$ konvergent sind. Das liegt an der Abschätzung (vgl. 5.2)

$$|Re\, z|, |Im\, z| \le |z| \le |Re\, z| + |Im\, z|.$$

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

Wir beschränken uns deshalb weitgehend auf reelle Zahlenfolgen.

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

(1) $|a_n - a| \le \alpha_n$ ffa $n \in \mathbb{N}$ und $\alpha_n \to 0 \Rightarrow a_n \to a$.

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

(3) $a_n \to a$ und $b_n \to b$ und $a_n \le b_n$ ffa $n \Rightarrow a \le b$.

(4) $a_n \to a$ und $b_n \to a$ und $a_n \le c_n \le b_n$ ffa $n \Rightarrow c_n \to a$.

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

Beweis: (1) ist leicht. Für (2) verwendet man $||a_n| - |a|| \leq |a_n - a|$ (vgl. 4.3(6)) und (1). (3) und (4) hier ohne Beweis.

zu (5): $(a_n + b_n)$ ohne Beweis. Bei $(a_n \cdot b_n)$ verwendet man

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

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

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

und

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

woraus die Behauptung folgt.

6.4 Monotone Folgen

Definition: Eine reelle Folge (a_n) heißt

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

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

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

6.5 Beispiele (1) $a_n \geq 0$ für alle $n \in \mathbb{N}$, $a_n \to a$ und $p \in \mathbb{N} \Rightarrow \sqrt[p]{a_n} \to \sqrt[p]{a}$.
Beweis: Wir stellen zunächst fest, dass für alle $y > x \geq 0$ gilt $\sqrt[p]{y} - \sqrt[p]{x} \leq \sqrt[p]{y - x}$. [Es ist nämlich nach dem binomischen Satz $(\sqrt[p]{y-x} + \sqrt[p]{x})^p = y - x + \underbrace{\ldots}_{\geq 0} + x \geq y$.]

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

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

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

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

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

(3) Konvergiert die Folge (a_n) so gilt $a_{n+1} - a_n \to 0$. Es gilt aber auch z.B. $\sqrt{n+1} - \sqrt{n} \to 0$. Denn es gilt auch $a_{n+1} \to 0$, und $a_{n+1} - a_n \to 0$ folgt aus Satz 6.3(5).
Für jedes $n \in \mathbb{N}$ ist

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

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

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

6.6 Teilfolgen

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

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

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

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

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

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

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

Bemerkung: Der Satz gilt für reelle und komplexe Zahlenfolgen.

6.7 Rechnen mit ∞

Definition: Sei (a_n) eine reelle Folge.

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

6.8 Limes superior und Limes inferior

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

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

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

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

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

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

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

$$a_{k(n)} \to \limsup a_n \qquad\qquad a_{l(n)} \to \liminf a_n.$$

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

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

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

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

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

7 Reihen

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

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

Ist $\sum_{n=1}^\infty a_n$ konvergent, so heißt $\lim_{N\to\infty} s_N$ der *Reihenwert* von $\sum_{n=1}^\infty a_n$ und wird ebenfalls mit $\sum_{n=1}^\infty a_n$ bezeichnet. Ende Woche 4

Bemerkung: (a) Die Folge (a_n) kann hier reell oder komplex sein. Ist (a_n) komplex, so heißt die Reihe $\sum_{n=1}^\infty \mathit{Re}\, a_n$ *Realteil* der komplexen Reihe $\sum_{n=1}^\infty a_n$ und $\sum_{n=1}^\infty \mathit{Im}\, a_n$ heißt *Imaginärteil* von $\sum_{n=1}^\infty a_n$.
Eine komplexe Reihe konvergiert genau dann, wenn ihr Realteil **und** ihr Imaginärteil **beide** konvergieren. In diesem Fall ist

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

vgl. mit 6.2.

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

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

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

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

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

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

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

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

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

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

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

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

(3) Ist $\sum_{n=1}^\infty a_n$ konvergent und $r_N := \sum_{n=N+1}^\infty a_n$ für jedes $N \in \mathbb{N}$, so gilt $r_N \to 0$ $(N \to \infty)$.

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

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

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

Beweis: (1) folgt aus 6.4, angewandt auf die monton wachsende Folge $(s_N)_{N\in\mathbb{N}}$.
(2) Setzt man $\sigma_N := \sum_{n=p+1}^N a_n$ für $N \geq p+1$, so ist $\sigma_N = s_N - s_p \to \sum_{n=1}^\infty a_n - s_p$ $(N \to \infty)$.
(3) Wegen (2) gilt $r_N = \sum_{n=1}^\infty a_n - s_N \to \sum_{n=1}^\infty a_n - \sum_{n=1}^\infty a_n = 0$.
(4) Es ist $a_n = s_n - s_{n-1} \to 0$ wegen 6.6(1).
(5) folgt aus 6.3(5).

7.3 Absolute Konvergenz

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

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

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

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

(Dreiecksungleichung für Reihen). (Satz hier ohne Beweis.)

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

Beweis wie im Beispiel mit $b_n = 1/n$.

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

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

7.5 Majoranten- und Minorantenkriterium: 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: Für (1) verwende 7.2(1). (2) folgt aus (1).

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

7.6 Wurzelkriterium: Sei (a_n) eine Folge und $\alpha := \limsup \sqrt[n]{|a_n|}$.

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

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

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

Beweis: Ist $\alpha < 1$, so wählen wir $x \in (\alpha, 1)$. Dann gilt $\sqrt[n]{|a_n|} \leq x$ ffa $n \in \mathbb{N}$, also $|a_n| \leq x^n$ ffa $n \in \mathbb{N}$. Wegen $x \in (0,1)$ konvergiert $\sum_{n=1}^\infty x^n$, und die Behauptung folgt aus 7.5(1).

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

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

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

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

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

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

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

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

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

(2) Ist $\alpha := \limsup c_n < 1$, so wähle $x \in (\alpha, 1)$. Es gilt dann $c_n \leq x$ ffa $n \in \mathbb{N}$, dh wir finden n_0 mit $c_n \leq x$ für alle $n \geq n_0$. Es folgt $|a_n| \leq x|a_{n-1}| \leq \ldots \leq x^{n-n_0}|a_{n_0}| = x^n(|a_{n_0}|/x^{n_0})$. Wegen $x \in (0,1)$ konvergiert $\sum_{n=1}^\infty x^n$, und nach 7.5(1) ist $\sum_{n=1}^\infty a_n$ absolut konvergent. Ist $\beta := \liminf c_n > 1$, so folgt $c_n \geq 1$ ffa n und $\sum_{n=1}^\infty a_n$ divergiert nach (1).

7.8 Die Exponentialreihe

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

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

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

7.9 Umordnungen

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

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

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

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

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

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

Beispiele hierzu gibt es in den Übungsaufgaben.

7.10 Das Cauchyprodukt

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

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

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

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

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

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

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

Nach dem Satz konvergiert diese Reihe absolut und es gilt

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

7.11 Die Exponentialfunktion

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

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

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

(0) Es gilt $E(0) = 1$.

(1) Für alle $z, w \in \mathbb{R}$ gilt $E(z)E(w) = E(z+w)$.

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

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

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

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

[Es ist $1 = E(0) = E(z + (-z)) = E(z)E(-z)$, woraus $E(z) \neq 0$ und $E(z)^{-1} = E(-z)$ folgt. Der Rest folgt aus (1).] Ende Woche 5

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

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

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

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

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

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

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

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

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

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

[Dies folgt aus der Definition, sowie der Tatsache, dass $w_n \to w_0$ für eine komplexe Folge (w_n) impliziert: $\overline{w_n} \to \overline{w}$.]

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

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

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

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

[Es gilt nach (2): $E(\frac{x}{n})^n = E(n\frac{x}{n}) = E(x)$.]

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

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

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

Definition: Man setzt $e := E(1)$ *Eulersche Zahl.* Es ist also

$$e = \sum_{n=0}^{\infty} \frac{1}{n!} = 1 + 1 + \frac{1}{2} + \frac{1}{3!} + \frac{1}{4!} + \ldots \approx 2.718.$$

Bemerkung: Für alle $x \in \mathbb{R}$ gilt $(1 + \frac{x}{n})^n \to E(x)$ $(n \to \infty)$. Insbesondere ist $e = \lim_{n\to\infty}(1 + \frac{1}{n})^n$. (ohne Beweis.)

Bemerkung und Definition: Wegen (2) und (8) gilt $e^n = E(n)$ für alle $n \in \mathbb{Z}$ und $\sqrt[n]{e} = E(\frac{1}{n})$ für alle $n \in \mathbb{N}$. Man schreibt deshalb auch $e^z := E(z)$ für jedes $z \in \mathbb{C}$. Eine andere Bezeichnung ist $\exp(z) := E(z)$.

Wir zeigen noch die folgenden **Abschätzungen**:

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

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

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

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

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

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

7.12 Sinus und Cosinus

Definition: Wir definieren die Funktionen $\sin : \mathbb{C} \to \mathbb{C}$ und $\cos : \mathbb{C} \to \mathbb{C}$ durch

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

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

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

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

Das folgt aus

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

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

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

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

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

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

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

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

Das folgt sofort aus 7.11(1).

Wir zeigen noch die folgenden **Abschätzungen**:

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

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

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

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

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

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

7.13 Potenzreihen

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

Wir nennen eine Potenzreihe *reell*, falls z_0 und alle a_n reell sind.

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

Beispiele: Die Potenzreihen für exp, sin und cos konvergieren für jedes $z \in \mathbb{C}$.
Die geometrische Reihe konvergiert für $|z| < 1$.
Alle diese Reihen haben als Entwicklungspunkt $z_0 = 0$.

Bemerkung: Jede Potenzreihe $\sum_{n=0}^{\infty} a_n(z - z_0)^n$ konvergiert für $z = z_0$, dh im Entwicklungspunkt.

7.14 Der Konvergenzradius

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

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

und die Behauptungen des Satzes folgen aus dem Wurzelkriterium 7.6.

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

[Das folgt aus dem Quotientenkriterium (Bemerkung in 7.7).]

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

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

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

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

8 Stetigkeit

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

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

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

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

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

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

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

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

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

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

Bemerkung: Ist g stetig in x_0 und $g(x_0) \neq 0$, so gibt es ein $\delta > 0$ mit $(x_0 - \delta, x_0 + \delta) \cap \{x \in D : g(x) = 0\} = \emptyset$.

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

Beispiele: Sind p, q reelle Polynome mit $q \neq 0$, so ist die Funktion $\frac{p}{q} : \mathbb{R} \setminus \{x : q(x) = 0\} \to \mathbb{R}$, $x \mapsto \frac{p(x)}{q(x)}$ stetig. Ende Woche 6

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

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

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

Beispiele: $\lim_{x\to 0} E(x) = 1$, $\lim_{x\to\infty} E(x) = \infty$, $\lim_{x\to-\infty} E(x) = 0$, $\lim_{x\to 0} \frac{\sin x}{x} = 1$, $\lim_{x\to 0} \frac{E(x)-1}{x} = 1$, $\lim_{x\to 0} \frac{\cos x - 1}{x} = 0$. [Die Aussagen folgen aus 7.11 und 7.12.]

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

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

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

$$x_0 \text{ ist Häufungspunkt von } D \;\Leftrightarrow\; \forall \varepsilon > 0 \, U_\varepsilon(x_0) \cap (D \setminus \{x_0\}) \neq \emptyset.$$

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

Satz: Die Funktion $f : D \to \mathbb{R}$ ist stetig genau dann, wenn für jeden Häufungspunkt x_0 von D gilt $\lim_{x\to x_0} f(x) = f(x_0)$.

Bemerkung: Ist x_0 kein Häufungspunkt von D, so gibt es ein $\varepsilon > 0$ so, dass $U_\varepsilon(x_0) \cap (D \setminus \{x_0\}) = \emptyset$ gilt, dh dass $U_\varepsilon(x_0) \cap D = \{x_0\}$ ist. Ein solches x_0 heißt *isolierter Punkt von D*. Eine Funktion $f : D \to \mathbb{R}$ ist in jedem isolierten Punkt von D stetig.

Beispiel: Sei $D = [0,1] \cup \{2\}$ und $f : D \to \mathbb{R}$, $f(x) := \begin{cases} x^2 & , x \in [0,1) \\ 1/2 & , x = 1 \\ 0 & , x = 2 \end{cases}$.

(i) $x_0 \in [0,1)$: Ist (x_n) eine Folge in D mit $x_n \to x_0$, so gilt $x_n \in [0,1)$ ffa n, und somit $f(x_n) = x_n^2 \to x_0^2$. Also ist f stetig in x_0.

(ii) $x_0 = 1$: Ist (x_n) eine Folge in $D \setminus \{1\}$ mit $x_n \to 1$, so gilt ebenfalls $x_n \in [0,1)$ ffa n, also $f(x_n) = x_n^2 \to 1$. Somit gilt $\lim_{x\to 1} f(x) = 1 \neq \frac{1}{2} = f(1)$ und f ist in $x_0 = 1$ nicht stetig.

(iii) $x_0 = 2$: Ist (x_n) eine Folge in D mit $x_n \to 2$, so gilt $x_n = 2$ ffa n, also $f(x_n) = f(2) = 0$ ffa n und $f(x_n) \to 0 = f(2)$. Also ist f in $x_0 = 2$ stetig (beachte, dass 2 ein isolierter Punkt von D ist).

8.5 Satz (ε-δ-Kriterium): Sei $\emptyset \neq D \subset \mathbb{R}$, $f : D \to \mathbb{R}$ eine Funktion und $x_0 \in D$. Dann ist f in x_0 stetig genau dann, wenn zu jedem $\varepsilon > 0$ ein $\delta > 0$ existiert mit

$$|f(x) - f(x_0)| < \varepsilon \quad \text{für alle } x \in D \text{ mit } |x - x_0| < \delta.$$

Beweis: Es gelte "ε-δ". Sei (x_n) eine Folge in D mit $x_n \to x_0$ und $\varepsilon > 0$. Dann finden wir ein $\delta > 0$ mit $|f(x) - f(x_0)| < \varepsilon$ für alle $x \in D$ mit $|x - x_0| < \delta$. Wegen $x_n \to x_0$ gilt $|x_n - x_0| < \delta$ ffa n. Es folgt $|f(x_n) - f(x_0)| < \varepsilon$ ffa n. Damit ist $f(x_n) \to f(x_0)$ gezeigt.

Wenn "ε-δ" nicht gilt, finden wir ein $\varepsilon > 0$ und zu jedem $n \in \mathbb{N}$ ein $x_n \in D$ mit $|x_n - x_0| < 1/n[= \delta]$ und $|f(x_n) - f(x_0)| \geq \varepsilon$. Es gilt dann $x_n \to x_0$, aber $f(x_n) \not\to f(x_0)$. Somit ist f in x_0 nicht stetig.

8.6 Satz (Komposition stetiger Funktionen): Seien $D_1, D_2 \subset \mathbb{R}$, $f : D_1 \to \mathbb{R}$ und $g : D_2 \to \mathbb{R}$ Funktionen. Es sei $f(D_1) \subset D_2$, $x_0 \in D_1$ und $y_0 := f(x_0)$. Ist f stetig in x_0 und g stetig in y_0, so ist $g \circ f : D_1 \to \mathbb{R}$ stetig in x_0.

Beweis: Sei (x_n) eine Folge in D_1 mit $x_n \to x_0$. Da f in x_0 stetig ist, folgt $y_n := f(x_n) \to f(x_0) = y_0$. Da g stetig in y_0 ist, folgt

$$(g \circ f)(x_n) = g(f(x_n)) = g(y_n) \to g(y_0) = g(f(x_0)) = (g \circ f)(x_0).$$

Also ist $g \circ f$ in x_0 stetig.

Beispiel: Sei $p \in \mathbb{N}$. Die Funktion $g_p : [0, \infty) \to \mathbb{R}$, $x \mapsto \sqrt[p]{x}$ ist stetig (nach 6.5(1)). Da $f : x \mapsto x^2$ stetig ist, ist auch $x \mapsto g_2 \circ f(x) = \sqrt{x^2} = |x|$ auf $\mathbb{R}$ stetig.

8.7 Satz (Stetigkeit von Potenzreihen): Sei $\sum_{n=0}^{\infty} a_n(x - x_0)^n$ eine reelle Potenzreihe mit Konvergenzradius R, wobei $0 < R \leq \infty$. Dann ist die Funktion

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

auf $D := (x_0 - R, x_0 + R)$ stetig. (hier ohne Beweis!)

8.8 Zwischenwertsatz (ZWS): Seien $a, b \in \mathbb{R}$ mit $a < b$ und $f : [a, b] \to \mathbb{R}$ stetig. Es sei y_0 *zwischen* $f(a)$ *und* $f(b)$ (dh $y_0 \in [f(a), f(b)]$, falls $f(a) \leq f(b)$, und $y_0 \in [f(b), f(a)]$, falls $f(a) > f(b)$). Dann gibt es ein $x_0 \in [a, b]$ mit $f(x_0) = y_0$.

Folgerung: Sei I ein Intervall und $f : I \to \mathbb{R}$ eine stetige Funktion. Dann ist $f(I)$ ein Intervall.

Bemerkung: Eine Teilmenge $J \subset \mathbb{R}$ ist genau dann ein Intervall, wenn für alle $c, d \in J$ mit $c < d$ gilt: $[c, d] \subset J$.

Beweis der Folgerung: Setze $J := f(I)$. Seien $c, d \in J$ 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 $y_0 \in f(I) = J$.

Beispiel: Ist p ein normiertes Polynom von ungeradem Grad m, so gilt $p(\mathbb{R}) = \mathbb{R}$.

[Wegen des Zwischenwertsatzes reicht es zu zeigen, dass $p(x) \to \infty$ für $x \to \infty$ und dass $p(x) \to -\infty$ für $x \to -\infty$. Die Behauptungen sind klar für $m = 1$. Sei also $m \geq 3$ und

$p(x) = x^m + a_{m-1}x^{m-1} + \ldots + a_1 x + a_0$. Für $x \geq 1 + |a_{m-1}| + |a_{m-2}| + \ldots + |a_1| + |a_0|$ gilt dann

$$\begin{aligned}
p(x) &\geq x^m - |a_{m-1}|x^{m-1} - |a_{m-2}|x^{m-2} - \ldots - |a_1|x - |a_0| \\
&\geq x^m - |a_{m-1}|x^{m-1} - |a_{m-2}|x^{m-1} - \ldots - |a_1|x^{m-1} - |a_0|x^{m-1} \\
&\geq x^{m-1} \underbrace{(x - (|a_{m-1}| + |a_{m-2}| + \ldots + |a_1| + |a_0|))}_{\geq 1} \\
&\geq x^{m-1} \to \infty \quad (x \to \infty).
\end{aligned}$$

Die Aussage für $x \to -\infty$ zeigt man ähnlich.

8.9 Einseitige Grenzwerte

Sei $D \subset \mathbb{R}$ und $f : D \to \mathbb{R}$ eine Funktion. Sei $x_0 \in \mathbb{R}$ ein Häufungspunkt von $D \cap (x_0, \infty)$ [bzw. von $D \cap (-\infty, x_0)$]. Wir schreiben

$$\lim_{x \to x_0+} f(x) = \beta \in \mathbb{R} \cup \{\infty, -\infty\} \quad [\text{bzw. } \lim_{x \to x_0-} f(x) = \beta \in \mathbb{R} \cup \{\infty, -\infty\}],$$

falls für jede Folge (x_n) in $D \cap (x_0, \infty)$ [bzw. in $D \cap (-\infty, x_0)$] mit $x_n \to x_0$ gilt: $f(x_n) \to \beta$.

Im Falle der Existenz heißt $\lim_{x \to x_0+} f(x)$ *rechtsseitiger Grenzwert* und $\lim_{x \to x_0-} f(x)$ heißt *linksseitiger Grenzwert* von f in x_0.

Schreibweisen: $\lim_{x \to x_0+0} f(x)$, $\lim_{x \downarrow x_0} f(x)$ statt $\lim_{x \to x_0+} f(x)$ und $\lim_{x \to x_0-0} f(x)$, $\lim_{x \uparrow x_0} f(x)$ statt $\lim_{x \to x_0-} f(x)$.

Beispiel: Sei $D = \mathbb{R}$ und $f : \mathbb{R} \to \mathbb{R}$ gegeben durch $f(x) := \begin{cases} 1 & , x \in (0,1) \\ 0 & , x \notin (0,1) \end{cases}$. Dann gilt $\lim_{x \to 0+} f(x) = 1$ und $\lim_{x \to 0-} f(x) = 0$.

Satz: Sei $f : D \to \mathbb{R}$ eine Funktion und $x_0 \in D$ mit $U_\varepsilon(x_0) \subset D$. f ist stetig in x_0 genau dann, wenn $\lim_{x \to x_0+} f(x) = f(x_0) = \lim_{x \to x_0-}$ gilt.

8.10 Monotone Funktionen und Stetigkeit

Definition: Sei $\emptyset \neq D \subset \mathbb{R}$ und $f : D \to \mathbb{R}$ eine Funktion. f heißt *monoton wachsend* [bzw. *monoton fallend*], falls für alle $x_1, x_2 \in D$ gilt:

$$x_1 \leq x_2 \Longrightarrow f(x_1) \leq f(x_2) \quad [\text{bzw. } f(x_1) \geq f(x_2)].$$

f heißt *streng monoton wachsend* [bzw. *streng monoton fallend*], falls für alle $x_1, x_2 \in D$ gilt:

$$x_1 < x_2 \Longrightarrow f(x_1) < f(x_2) \quad [\text{bzw. } f(x_1) > f(x_2)].$$

f heißt *monoton*, falls f monoton wachsend oder monoton fallend ist, und f heißt *streng monoton*, falls f streng monoton wachsend oder streng monoton fallend ist.

Beispiele: $E : \mathbb{R} \to \mathbb{R}$ ist streng monoton wachsend.
Die Funktion $\mathbb{R} \to \mathbb{R}$, $x \mapsto x^2$ ist nicht monoton.

Die Funktion $f : \mathbb{R} \to \mathbb{R}$, $x \mapsto \begin{cases} 1 & , x \geq 0 \\ 0 & , x < 0 \end{cases}$ ist monoton wachsend, aber *nicht streng* monoton wachsend.

Satz: Sei $I \subset \mathbb{R}$ ein Intervall und $f : I \to \mathbb{R}$ streng monoton wachsend [bzw. fallend].

(a) Dann ist f injektiv und die Umkehrfunktion $f^{-1} : f(I) \to \mathbb{R}$ ist ebenfalls streng monoton wachsend [bzw. fallend].

(b) Ist f zusätzlich stetig, so ist $f(I)$ ein Intervall und $f^{-1} : f(I) \to \mathbb{R}$ ist stetig.

(hier ohne Beweis)

8.11 Der Logarithmus

Es gilt $E(\mathbb{R}) = (0, \infty)$. Die Abbildung $E : \mathbb{R} \to (0, \infty)$ ist bijektiv und stetig.

Definition: Die Umkehrfunktion der Exponentialfunktion $E : \mathbb{R} \to (0, \infty)$ heißt (natürlicher) *Logarithmus* $\ln : (0, \infty) \to \mathbb{R}$, dh also $\ln x := \log x := E^{-1}(x)$ für $x \in (0, \infty)$.

Somit ist $\ln(E(x)) = x$ für alle $x \in \mathbb{R}$ und $E(\ln y) = y$ für alle $y \in (0, \infty)$.

Eigenschaften: ln ist streng monoton wachsend und stetig. Es gilt $\ln((0, \infty)) = \mathbb{R}$, $\ln(1) = 0$ und $\ln(e) = 1$.

$\ln x \to \infty$ für $x \in \infty$ und $\ln x \to -\infty$ für $x \to 0+$.

Für alle $x, y > 0$ gilt

$$\ln(xy) = \ln x + \ln y \quad \text{und} \quad \ln(x/y) = \ln x - \ln y.$$

[Es gilt nämlich

$$\begin{aligned} E(\ln x + \ln y) &= E(\ln x)E(\ln y) = xy = E(\ln(xy)), \\ E(\ln x - \ln y) &= E(\ln x)E(-\ln y) = xE(\ln y)^{-1} = x/y = E(\ln(x/y)).] \end{aligned}$$

8.12 Die allgemeine Potenz

Wir definieren für $a > 0$ die allgemeine Potenz:

$$a^x := E(x \ln a) \quad \text{für jedes } x \in \mathbb{R}.$$

Bemerkung: Nach 7.11 stimmt dies für $x \in \mathbb{Z}$ mit der bisherigen Definition überein. Für $x = 1/n$ mit $n \in \mathbb{N}$ ist (ebenfalls nach 7.11) $a^{1/n} = \sqrt[n]{a}$. Für $x = p/q \in \mathbb{Q}$ mit $p \in \mathbb{Z}$, $q \in \mathbb{N}$ folgt aus 7.11:

$$a^{p/q} = \sqrt[q]{a^p} = (\sqrt[q]{a})^p.$$

Für $a = e$ erhalten wir $e^x = E(x)$ wie in 7.11. Wir schreiben in Zukunft in der Regel e^x statt $E(x)$.

Eigenschaften: Für $a > 0$, $x, y \in \mathbb{R}$ gilt:

(1) $a^x > 0$;

(2) die Funktion $x \mapsto a^x$ ist auf $\mathbb{R}$ stetig (wegen 8.6);

(3) $a^{x+y} = e^{(x+y)\ln a} = e^{x \ln a} e^{y \ln a} ? a^x a^y$;

(4) $a^{-x} = e^{-x \ln a} = (e^{x \ln a})^{-1} = (a^x)^{-1} = 1/a^x$;

(5) $\ln(a^x) = \ln(E(x \ln a)) = x \ln a$;

(6) $(a^y)^x = e^{x \ln(a^y)} \overset{(5)}{=} e^{xy \ln a} = a^{xy}$.

Der allgemeine Logarithmus: Sei $a > 1$. Dann ist die Funktion $\mathbb{R} \to (0, \infty)$, $x \mapsto a^x$, streng monoton wachsend, stetig und bijektiv.

Die Umkehrfunktion $\log_a : (0, \infty) \to \mathbb{R}$ heißt *Logarithmus zur Basis a*. Es ist also $\log_a(a^x) = x$ für alle $x \in \mathbb{R}$ und $a^{\log_a y} = y$ für alle $y \in (0, \infty)$.

Satz: Für alle $y \in (0, \infty)$ gilt $\log_a y = \frac{\ln y}{\ln a}$.

Denn $a^{\frac{\ln y}{\ln a}} = E(\frac{\ln y}{\ln a} \ln a) = E(\ln y) = y$.

8.13 Abgeschlossene Mengen

Definition: Sei $D \subset \mathbb{R}$. D heißt *abgeschlossen*, falls für jede Folge (x_n) in D mit $x_n \to x_0 \in \mathbb{R}$ gilt: $x_0 \in D$.

Beispiele: (1) Folgende Mengen sind abgeschlossen: $\mathbb{R}$, $\emptyset$, $[a, b]$, $(-\infty, a]$, $[a, \infty)$, $\{1/n : n \in \mathbb{N}\} \cup \{0\}$.

(2) Folgende Mengen sind nicht abgeschlossen: (a, b), $(a, b]$, $[a, b)$, (a, ∞), $(-\infty, a)$, $\{1/n : n \in \mathbb{N}\}$.

8.14 Satz: Sei $\emptyset \neq D \subset \mathbb{R}$, D abgeschlossen und beschränkt, sowie $f : D \to \mathbb{R}$ eine stetige Funktion. Dann ist $f(D)$ abgeschlossen und beschränkt und es gibt $x_1, x_2 \in D$ mit

$$\text{für alle } x \in D \text{ gilt: } f(x_1) \leq f(x) \leq f(x_2)$$

(dh "f nimmt auf D Maximum und Minimum an").

Folgerung: Sind $a, b \in \mathbb{R}$ mit $a < b$ und ist $f : [a, b] \to \mathbb{R}$ stetig, so gibt es $c, d \in \mathbb{R}$ mit $f([a, b]) = [c, d]$.

Beispiel: $D = [1, 53]$ und $f(x) := x^3 \sin(e^x + \ln x)$. Ende Woche 7

9 Trigonometrische Funktionen und Hyperbelfunktionen

9.1 Die Zahl π

Vorbereitungen: Wir interessieren uns für Nullstellen der Cosinus-Funktion. Nach 7.12 gilt für jedes $x \in \mathbb{R}$:

$$\cos x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!} = 1 - \frac{x^2}{2} + \frac{x^4}{4!} - \frac{x^6}{6!} + \dots$$

Die Reihe ist alternierend. Wie beim Leibnizkriterium 7.4 können wir durch Abbrechen der Reihe bei $n = n_0$ Abschätzungen nach oben oder nach unten angeben, wenn nur die Folge $(x^{2n}/(2n)!)_{n \geq n_0}$ monoton fallend ist. Wegen

$$\frac{x^{2n}}{(2n)!} \geq \frac{x^{2(n+1)}}{(2(n+1))!} \Leftrightarrow (2n+2)(2n+1) \geq x^2$$

gilt dies für $n_0 = 1$, wenn $x^2 \leq 12$, dh $|x| \leq 2\sqrt{3}$ ist.
Also gilt für alle $x \in [0, 2\sqrt{3}]$:

$$1 - \frac{x^2}{2} + \frac{x^4}{24} \geq \cos x \geq 1 - \frac{x^2}{2}.$$

Insbesondere ist $\cos \sqrt{2} \geq 0$ und $\cos(\sqrt{2}\sqrt{3 - \sqrt{3}}) \leq 0$. Nach dem Zwischenwertsatz 8.8 gibt es mindestens ein $x_0 \in [\sqrt{2}, \sqrt{2}\sqrt{3 - \sqrt{3}}]$ mit $\cos x_0 = 0$.

Definition: Wir definieren $\pi/2$ als die kleinste Nullstelle des Cosinus im Intervall $[0, 2]$, also

$$\frac{\pi}{2} := \inf\{x_0 \in [0, 2] : \cos x_0 = 0\}.$$

Beachte: Es gilt

$$\sqrt{2} \leq \pi/2 \leq \sqrt{2}\sqrt{3 - \sqrt{3}} < 2 < 2\sqrt{3}$$

und $\sqrt{2} \approx 1.41421$, sowie $\sqrt{2}\sqrt{3 - \sqrt{3}} \approx 1.59245$.

Bemerkung: Nach 7.12 gilt

$$\sin x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}, \quad \text{für alle } x \in \mathbb{R}.$$

Die Argumente in der Vorbetrachtung zeigen dann

$$\sin x \geq x - \frac{x^3}{3!} \geq 0 \quad \text{für alle } x \in [0, \sqrt{6}].$$

Wegen $\sin^2 x + \cos^2 x = 1$ folgt also $\sin \frac{\pi}{2} = 1$.

9.2 Eigenschaften

Aus den Additionstheoremen in 7.12 erhalten wir:

(1) $\sin \pi = 0$, $\cos \pi = -1$;

(2) $\forall x \in \mathbb{R}$: $\sin(x+\pi) = -\sin x$, $\cos(x+\pi) = -\cos x$;

(3) $\sin 2\pi = 0$, $\cos 2\pi = 1$;

(4) $\forall x \in \mathbb{R}$: $\sin(x+2\pi) = \sin x$, $\cos(x+2\pi) = \cos x$ (dh sin und cos sind 2π-*periodisch*).

(5) $\forall x \in \mathbb{R}$: $\sin(\pi + x) = -\sin(\pi - x)$, $\cos(\pi + x) = \cos(\pi - x)$;

(6) $\forall x \in \mathbb{R}$: $\sin(\frac{\pi}{2} + x) = \sin(\frac{\pi}{2} - x)$, $\cos(\frac{\pi}{2} + x) = -\cos(\frac{\pi}{2} - x)$;

(7) $\forall x \in \mathbb{R}$: $\sin(x + \frac{\pi}{2}) = \cos x$.

Hieraus können wir alle Nullstellen von sin und cos bestimmen:

(8) Für $x \in \mathbb{R}$ gilt: $\cos x = 0 \Leftrightarrow$ es gibt $k \in \mathbb{Z}$ mit $x = (2k+1)\frac{\pi}{2}$.

(9) Für $x \in \mathbb{R}$ gilt: $\sin x = 0 \Leftrightarrow$ es gibt $k \in \mathbb{Z}$ mit $x = k\pi$.

(10) Für $k \in \mathbb{Z}$ gilt: $\cos(k\pi) = (-1)^k$ und $\sin((2k+1)\pi/2) = (-1)^k$.

Schließlich geben wir noch an:

(11) $\sin \frac{\pi}{4} = \cos \frac{\pi}{4} = \frac{\sqrt{2}}{2}$.

Außerdem erinnern wir an:

(12) Für alle $x \in \mathbb{R}$ gilt $\sin(-x) = -\sin x$ und $\cos(-x) = \cos x$.

9.3 Monotonie für sin **und** cos

Es reicht, die Funktionen auf dem Intervall $[0, \pi/2]$ zu betrachten. Dabei ist $\sin x > 0$ für $x \in (0, \pi/2]$ und $\cos x > 0$ für $x \in [0, \pi/2)$. Für $x, x+h \in [0, \pi/2]$ mit $h > 0$ gilt also

$$\cos(x+h) = \cos x \cos h - \underbrace{\sin x \sin h}_{>0} > \cos x,$$

dh cos **ist auf** $[0, \pi/2]$ **streng monoton fallend**.

Folglich ist $\sin x$ auf $[0, \pi/2]$ streng monoton wachsend (wegen $\sin x = \sqrt{1 - \cos^2 x}$ für $x \in [0, \pi/2]$).

9.4 Arcussinus und Arcuscosinus

Wegen 9.3 und 9.2(6) ist $\cos : [0, \pi] \to [-1, 1]$ streng monoton fallend, nach 8.2(2) ist diese Abbildung stetig und wegen 8.10 ist sie bijektiv.

Die Umkehrabbildung $\arccos : [-1, 1] \to [0, \pi]$ heißt *Arcuscosinus*.

Eigenschaften: $\arccos : [-1,1] \to [0,\pi]$ ist streng monoton fallend, stetig und bijektiv. Es gilt $\arccos(-1) = \pi$, $\arccos 0 = \pi/2$, $\arccos 1 = 0$.

Wegen 9.3 und 9.2(12) ist $\sin : [-\pi/2, \pi/2] \to [-1,1]$ streng monoton wachsend. Auch diese Funktion ist stetig und bijektiv.

Die Umkehrfunktion $\arcsin : [-1,1] \to [-\pi/2, \pi/2]$ heißt *Arcussinus*.

Eigenschaften: Die Funktion $\arcsin : [-1,1] \to [-\pi/2, \pi/2]$ ist streng monoton wachsend, stetig und bijektiv. Es gilt $\arcsin(1) = \pi/2$, $\arcsin 0 = 0$ und $\arcsin(-x) = -\arcsin x$ für alle $x \in [-1,1]$.

Achtung: Wegen 9.2 sind für jedes $k \in \mathbb{Z}$ auch die Abbildungen

$$\cos : [k\pi, (k+1)\pi] \to [-1,1], \qquad \sin : [(2k-1)\pi/2, (2k+1)\pi/2] \to [-1,1]$$

streng monoton, stetig und bijektiv. Für $k \neq 0$ sind ihre Umkehrabbildungen von den eben definierten Funktionen arccos und arcsin verschieden!

9.5 Der Tangens

Definition: Die Funktion

$$\tan : \mathbb{R} \setminus (\frac{\pi}{2} + \pi\mathbb{Z}) \to \mathbb{R}, \quad x \mapsto \tan x := \frac{\sin x}{\cos x}$$

heißt *Tangens*. (Beachte, dass die Menge $\frac{\pi}{2} + \pi\mathbb{Z}$ gerade die Nullstellen des Cosinus enthält.)

Eigenschaften: Es gilt $\tan 0 = 0$, $\tan \pi/4 = 1$, sowie für alle x im Definitionsbereich, dh für alle $x \in \mathbb{R} \setminus (\frac{\pi}{2} + \pi\mathbb{Z})$:

$$\tan(-x) = -\tan x \qquad \tan(x+\pi) = \tan x.$$

Somit ist der Tangens eine π-periodische Funktion.

Außerdem ist tan auf $(-\pi/2, \pi/2)$ streng monoton wachsend mit $\tan x \to \infty$ für $x \to \pi/2-$ und $\tan x \to -\infty$ für $x \to -\pi/2+$.

Folglich ist $\tan : (-\pi/2, \pi/2) \to \mathbb{R}$ streng monoton wachsend, stetig und bijektiv. Die Umkehrfunktion $\arctan : \mathbb{R} \to (-\pi/2, \pi/2)$ heißt *Arcustangens*.

Eigenschaften: Es gilt $\arctan 0 = 0$, $\arctan 1 = \pi/4$ und $\arctan(-x) = -\arctan x$ für alle $x \in \mathbb{R}$.

Bemerkung: Sei $k \in \mathbb{Z}$. Man sieht leicht, dass die Umkehrabbildung von $\tan : (k\pi - \pi/2, k\pi + \pi/2) \to \mathbb{R}$ gegeben ist durch

$$\mathbb{R} \to (k\pi - \pi/2, k\pi + \pi/2), x \mapsto k\pi + \arctan x.$$

9.6 Anwendung: Polarkoordinaten

Definition und Satz: Für jede komplexe Zahl $z \in \mathbb{C} \setminus \{0\}$ gibt es genau ein $r \in (0, \infty)$ und genau einen Winkel $\varphi \in (-\pi, \pi]$ mit $z = re^{i\varphi}$. Dabei heißt $r = |z|$ *Länge von z* und $\varphi =: \arg z$ heißt das *Argument von z*.

Satz: Für jedes Paar $(a, b) \in \mathbb{R}^2$ mit $a^2 + b^2 = 1$ gibt es genau ein $\varphi \in (-\pi, \pi]$ mit $\cos \varphi = a$ und $\sin \varphi = b$.

Bemerkung: Dabei ist $\varphi = \arcsin b$, falls $a \geq 0$ ist, und $\varphi = \arccos a$, falls $b \geq 0$ ist. Sind $a, b < 0$, so ist $\varphi = \arccos |a| - \pi = \arcsin |b| - \pi$.

Alternativ hat man $\varphi = \arctan(b/a)$ für $a > 0$, $\varphi = b\pi/2$ für $a = 0$ (dann $b \in \{-1, 1\}$), sowie $\varphi = \pi + \arctan(b/a)$ für $a < 0$, $b \geq 0$, und $\varphi = -\pi + \arctan(b/a)$ für $a, b < 0$.

Bemerkung: Es gilt $e^{i\psi} = 1$ genau dann, wenn $\psi \in 2\pi\mathbb{Z}$ ist. Insbesondere gibt es zu jedem $z \in \mathbb{C} \setminus \{0\}$ unendlich viele $\psi \in \mathbb{R}$ mit $z = |z|e^{i\psi}$.

9.7 Hyperbelfunktionen

Definiere für $x \in \mathbb{R}$:

$$\cosh x := \frac{1}{2}(e^x + e^{-x}), \qquad \sinh x := \frac{1}{2}(e^x - e^{-x})$$

(*Cosinus hyperbolicus* und *Sinus hyperbolicus*).

Dann gilt $\cosh 0 = 1$ und $\sinh 0 = 0$, sowie

$$\forall x \in \mathbb{R} : \cosh(-x) = \cosh x, \quad \sinh(-x) = -\sinh x.$$

Außerdem gilt

$$\begin{aligned} \cosh x \to \infty, \quad \sinh x &\to \infty \quad \text{für} \quad x \to \infty \\ \cosh x \to \infty, \quad \sinh x &\to -\infty \quad \text{für} \quad x \to -\infty, \end{aligned}$$

sowie

$$\cosh^2 x - \sinh^2 x = 1 \quad \text{für alle } x \in \mathbb{R}.$$

Reihendarstellungen: Für alle $x \in \mathbb{R}$ gilt

$$\cosh x = \sum_{n=0}^{\infty} \frac{x^{2n}}{(2n)!}, \qquad \sinh x = \sum_{n=0}^{\infty} \frac{x^{2n+1}}{(2n+1)!}.$$

Folgerung: Die Funktionen

$$\cosh : [0, \infty) \to [1, \infty) \quad \text{und} \quad \sinh : \mathbb{R} \to \mathbb{R}$$

sind streng monoton wachsend, stetig und bijektiv.

Additionstheoreme: Für alle $x, y \in \mathbb{R}$ gilt

$$\begin{aligned}\cosh(x+y) &= \cosh x \cosh y + \sinh x \sinh y,\\ \sinh(x+y) &= \sinh x \cosh y + \cosh x \sinh y.\end{aligned}$$

Definition: Die Funktion

$$\tanh : \mathbb{R} \to \mathbb{R}, x \mapsto \tanh x := \frac{\sinh x}{\cosh x},$$

heißt *Tangens hyperbolicus*.

Es ist $\tanh(-x) = -\tanh x$ für alle $x \in \mathbb{R}$ und $\tanh 0 = 0$. Außerdem gilt $\tanh x \to 1$ für $x \to \infty$ und $\tanh x \to -1$ für $x \to -\infty$.

Die Funktion $\tanh : \mathbb{R} \to (-1, 1)$ ist streng monoton wachsend, stetig und bijektiv.

9.8 Areafunktionen

Definition: Die Umkehrfunktionen von $\sinh : \mathbb{R} \to \mathbb{R}$, $\cosh : [0, \infty) \to [1, \infty)$ und $\tanh : \mathbb{R} \to (-1, 1)$ heißen $\operatorname{Arsinh} : \mathbb{R} \to \mathbb{R}$ (*Areasinus*), $\operatorname{Arcosh} : [1, \infty) \to [0, \infty)$ (*Areacosinus*) und $\operatorname{Artanh} : (-1, 1) \to \mathbb{R}$ (*Areatangens*). Diese Funktionen sind jeweils streng monoton wachsend, stetig und bijektiv.

9.9 Weitere Funktionen

Es gibt noch weitere Funktionen, auf die wir hier nicht näher eingehen, z.B. $\cot x = 1/\tan x$, $\sec x = 1/\sin x$, $\operatorname{sech} x = 1/\sinh x$ etc.

10 Integration

10.1 Ober- und Untersummen, oberes und unteres Integral

Seien $a, b \in \mathbb{R}$ mit $a < b$ und sei $f : [a, b] \to \mathbb{R}$ *beschränkt*, dh $f([a, b]) = \{f(x) : x \in [a, b]\}$ beschränkt.

Definition: $Z := \{x_0, x_1, \ldots, x_n\}$ heißt eine *Zerlegung* von $[a, b]$, falls $a = x_0 < x_1 < \ldots < x_n = b$.

Sei $m := \inf f([a, b])$ und $M := \sup f([a, b])$. Sei $Z = \{x_0, x_1, \ldots, x_n\}$ ein Zerlegung von $[a, b]$. Für $j = 1, 2, \ldots, n$ setze

$$\begin{aligned} I_j &:= [x_{j-1}, x_j], \quad |I_j| := x_j - x_{j-1}, \quad m_j := \inf f(I_j), \quad M_j := \sup f(I_j), \text{ sowie} \\ s_f(Z) &:= \sum_{j=1}^{n} m_j |I_j| \qquad \textit{Untersumme} \text{ von } f \text{ bzgl. } Z, \\ S_f(Z) &:= \sum_{j=1}^{n} M_j |I_j| \qquad \textit{Obersumme} \text{ von } f \text{ bzgl. } Z. \end{aligned}$$

Es ist $m \leq m_j \leq M_j \leq M$, also wegen $|I_j| > 0$:

$$\sum_{j=1}^{n} m|I_j| \leq \sum_{j=1}^{n} m_j|I_j| \leq \sum_{j=1}^{n} M_j|I_j| \leq \sum_{j=1}^{n} M|I_j|.$$

Somit gilt für jede Zerlegung Z von $[a, b]$:

$$(*) \qquad m(b - a) \leq s_f(Z) \leq S_f(Z) \leq M(b - a).$$

Satz: Seien Z_1, Z_2 Zerlegungen.

(1) Ist $Z_1 \subset Z_2$, 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 $Z_1 \subset Z$ und $Z_2 \subset Z$ 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:

$$s_f := \underline{\int_a^b} f(x)\, dx := \sup\{s_f(Z) : Z \text{ ist Zerlegung von } [a, b]\},$$

das *untere Integral von* f über $[a, b]$ und

$$S_f := \overline{\int_a^b} f(x)\, dx := \inf\{S_f(Z) : Z \text{ ist Zerlegung von } [a, b]\}$$

das *obere Integral von* f über $[a, b]$.

Wegen $(*)$ und (2) gilt dann

$$(**) \qquad m(b-a) \le s_f \le S_f \le M(b-a).$$

[Zunächst folgt aus (2) durch Supremumsbildung über Z_1: $s_f \le S_f(Z_2)$ für jede Zerlegung Z_2. Dann bilde man das Infimum über alle Z_2.]

10.2 Definition: Seien $a, b \in \mathbb{R}$ mit $a < b$ und sei $f : [a, b] \to \mathbb{R}$ *beschränkt.* Dann heißt f (Riemann-)*integrierbar* (*ib*), falls $s_f = S_f$ gilt.

In diesen Falle heißt

$$\int_a^b f\, dx := \int_a^b f(x)\, dx := S_f (= s_f)$$

das (Riemann-)*Integral* von f über $[a, b]$.

Beispiele: (1) Sei $c \in \mathbb{R}$ und $f(x) = c$, $x \in [a, b]$. Dann ist $m = M = c$ und aus $(**)$ folgt $s_f = S_f = c(b-a)$. Also ist f integrierbar und

$$\int_a^b c\, dx = c(b-a).$$

(2) Sei $a = 0$, $b = 1$ und $f(x) := \begin{cases} 1 & , x \in [0,1] \cap \mathbb{Q} \\ 0 & , x \in [0,1] \setminus \mathbb{Q} \end{cases}$. Ist $Z = \{x_0, x_1, \ldots, x_n\}$ Zerlegung von $[0, 1]$ und sind m_j, M_j, I_j wie oben, so haben wir $m_j = 0$, $M_j = 1$ für alle j, also $s_f(Z) = 0$ und $S_f(Z) = 1$. Folglich ist

$$s_f = 0 \ne 1 = S_f,$$

und f ist nicht integrierbar über $[0, 1]$. Ende Woche 8

Definition: Wir setzen

$$R[a, b] := \{f : [a, b] \to \mathbb{R} : \ f \text{ ist beschränkt und über } [a, b] \text{ integrierbar}\}$$

(Menge der über $[a, b]$ integrierbaren Funktionen).

10.3 Satz: Seien $f, g \in R[a, b]$.

(1) Gilt $f \le g$ auf $[a, b]$, so ist $\int_a^b f\, dx \le \int_a^b g\, dx$.

(2) Für $\alpha, \beta \in \mathbb{R}$ ist $\alpha f + \beta g \in R[a,b]$ und

$$\int_a^b (\alpha f + \beta g)\, dx = \alpha \int_a^b f\, dx + \beta \int_a^b g\, dx.$$

[(1) ist klar, (2) hier ohne Beweis.]

10.4 Satz: Sei $f : [a,b] \to \mathbb{R}$ eine beschränkte Funktion. Es gilt $f \in R[a,b]$ genau dann, wenn es zu jedem $\varepsilon > 0$ eine Zerlegung Z gibt mit $S_f(Z) - s_f(Z) < \varepsilon$.

[Das ergibt sich aus der Definition.]

10.5 Beispiel: Sei $m \in \mathbb{N}$ und $y > 0$. Wir zeigen, dass $f : x \mapsto x^m$ über $[0,y]$ integrierbar ist und dass gilt

$$\int_0^y x^m\, dx = \frac{y^{m+1}}{m+1}.$$

Dazu sei $n \in \mathbb{N}$, $x_j := jy/n$ für $j = 0, 1, \ldots, n$ und Z die Zerlegung $\{x_0, x_1, \ldots, x_n\}$. Sind I_j, m_j, M_j wie in 10.1, so gilt $|I_j| = y/n$, $m_j = f(x_{j-1}) = ((j-1)y/n)^m$, $M_j = f(x_j) = (jy/n)^m$ und

$$\begin{aligned} S_f(Z) &= \sum_{j=1}^n M_j |I_j| = \frac{y}{n} \sum_{j=1}^n \left(\frac{jy}{n}\right)^m = \frac{y^{m+1}}{n^{m+1}} \sum_{j=1}^n j^m \\ s_f(Z) &= \sum_{j=1}^n m_j |I_j| = \frac{y}{n} \sum_{j=1}^n \left(\frac{(j-1)y}{n}\right)^m = \frac{y^{m+1}}{n^{m+1}} \sum_{j=1}^{n-1} j^m. \end{aligned}$$

Insbesondere ist $S_f(Z) - s_f(Z) = \frac{y^{m+1}}{n^{m+1}}\, n^m = y^{m+1}/n \to 0$ für $n \to \infty$, so dass $f \in R[0,y]$ aus 10.4 folgt. Außerdem sieht man

$$\int_0^y x^m\, dx = \lim_{n\to\infty} \frac{y^{m+1}}{n^{m+1}} \sum_{j=1}^n j^m,$$

so dass wir nur noch $n^{-(m+1)} \sum_{j=1}^n j^m \to (m+1)^{-1}$ für $n \to \infty$ zeigen müssen.

Dazu verwenden wir folgende Formel:

10.6 Die Summen m-ter Potenzen

Für alle $n, m \in \mathbb{N}_0$ sei $s_m(n) := \sum_{j=1}^n j^m$. Dann gilt für alle $n, m \in \mathbb{N}_0$:

$$n^{m+1} + \sum_{k=1}^{m+1} \binom{m+1}{k} (-1)^k s_{m+1-k}(n) = 0.$$

Beweis durch Induktion nach n, wobei $m \in \mathbb{N}_0$ fest ist. Der Induktionsanfang $n = 0$ ist klar, da jeder Summand auf der linken Seite $= 0$ ist.

Für den Induktionsschritt $n \to n+1$ beachte man

$$n^{m+1} = ((n+1)-1)^{m+1} = (n+1)^{m+1} + \sum_{k=1}^{m+1} \binom{m+1}{k} (-1)^k (n+1)^{m+1-k}.$$

Fortsetzung von 10.5: Dividiert man die Formel in 10.6 durch n^{m+1} und beachtet, dass $s_{m+1-k}(n) \le n^{m+2-k}$ für $k \ge 2$ gilt, so erhält man

$$\lim_{n\to\infty} \left(1 - (m+1)\frac{1}{n^{m+1}}\, s_m(n)\right) = 0, \quad \text{also} \quad \lim_{n\to\infty} \frac{1}{n^{m+1}}\, s_m(n) = \frac{1}{m+1},$$

wie gewünscht.

10.7 Satz: Ist $f : [a,b] \to \mathbb{R}$ monoton, so ist $f \in R[a,b]$.

Beweis: Sei f monoton wachsend (sonst betrachte $-f$ und verwende 10.3(2)). Sei $n \in \mathbb{N}$ und $x_j := a + j(b-a)/n$ für $j = 0, 1, \ldots, n$ und $Z := \{x_0, x_1, \ldots, x_n\}$. Sind I_j, M_j und m_j wie in 10.1, so gilt $|I_j| = (b-a)/n$, $M_j = f(x_j)$ und $m_j = f(x_{j-1})$. Somit haben wir

$$S_f(Z) - s_f(Z) = \sum_{j=1}^{n} (M_j - m_j)|I_j| = \frac{b-a}{n} \sum_{j=1}^{n} (f(x_j) - f(x_{j-1})) = \frac{b-a}{n}(f(b) - f(a)) \to 0$$

für $n \to \infty$, und die Behauptung folgt aus 10.4.

10.8 Satz: Ist $f : [a,b] \to \mathbb{R}$ stetig, so ist $f \in R[a,b]$.

Beweis: Sei $n \in \mathbb{N}$ und $x_j := a + j(b-a)/n$ für $j = 0, 1, \ldots, n$ und $Z := \{x_0, x_1, \ldots, x_n\}$. Sind I_j, M_j und m_j wie in 10.1, so gilt $|I_j| = (b-a)/n$ und

$$S_f(Z) - s_f(Z) = \sum_{j=1}^{n} (M_j - m_j)|I_j| = \frac{b-a}{n} \sum_{j=1}^{n} (M_j - m_j) \le (b-a) \max\{M_j - m_j : j = 1, 2, \ldots, n\}.$$

Sei nun $\varepsilon > 0$. Hier ohne Beweis (das liegt daran, dass f auf $[a,b]$ *gleichmäßig stetig* ist): Es gibt ein $\delta > 0$ so, dass für alle $x, \tilde{x} \in [a,b]$ gilt:

$$|x - \tilde{x}| < \delta \;\Rightarrow\; |f(x) - f(\tilde{x})| < \frac{\varepsilon}{b-a}.$$

Wir wählen n so groß, dass $\frac{b-a}{n} < \delta$ ist. Sind dann $x, \tilde{x} \in I_j$ für ein $j \in \{1, 2, \ldots, n\}$, so folgt $|f(x) - f(\tilde{x})| < \varepsilon/(b-a)$. Wir haben also für jedes $j \in \{1, 2, \ldots, n\}$:

$$M_j - m_j = \sup\{f(x) : x \in I_j\} - \inf\{f(\tilde{x}) : \tilde{x} \in I_j\} = \sup\{f(x) - f(\tilde{x}) : x, \tilde{x} \in I_j\} \le \frac{\varepsilon}{b-a}.$$

Zusammen erhalten wir

$$S_f(Z) - s_f(Z) \leq (b-a)\max\{M_j - m_j : j = 1,\dots,n\} \leq (b-a)\frac{\varepsilon}{b-a} = \varepsilon,$$

woraus nach 10.4 die Behauptung folgt.

Beispiele: (1) exp, sin, cos sind über jedem Intervall $[a,b] \subset \mathbb{R}$ integrierbar.

(2) Eine reelle Potenzreihe mit Konvergenzradius $R \in (0,\infty]$ und Entwicklungspunkt $x_0 \in \mathbb{R}$ ist über jedem $[a,b] \subset (x_0 - R, x_0 + R)$ integrierbar.

Zur Berechnung von Integralen:

10.9 Satz: Sei I ein Intervall und $f : I \to \mathbb{R}$ eine Funktion.

(1) Seien $a, b, c \in I$ mit $a < c < b$. Dann gilt $f \in R[a,b]$ genau dann, wenn $f \in R[a,c]$ und $f \in R[c,b]$. In diesem Fall ist

$$\int_a^b f\,dx = \int_a^c f\,dx + \int_c^b f\,dx.$$

(2) Sei $f \in R[a,b]$ für alle $[a,b] \subset I$, und sei $c \in I$. Setzt man

$$F : I \to \mathbb{R}, y \mapsto F(y) := \int_c^y f(x)\,dx,$$

wobei $\int_c^y f(x)\,dx := -\int_y^c f(x)\,dx$ für $y < c$ und $\int_c^c f(x)\,dx := 0$ gesetzt ist, so gilt für alle $a, b \in I$:

$$\int_a^b f(x)\,dx = F(b) - F(a).$$

[(2) folgt aus (1). Zum Beweis von (1) sei zunächst $f \in R[a,b]$ und $\varepsilon > 0$. Wir finden eine Zerlegung Z von $[a,b]$ mit $S_f(Z) - s_f(Z) < \varepsilon$. Wegen Satz 10.1(1) können wir davon ausgehen, dass $c \in Z$ gilt. Setzen wir dann $Z_1 := Z \cap [a,c]$ und $Z_2 := Z \cap [c,b]$, so gilt für $k = 1, 2$:

$$S_f(Z_k) - s_f(Z_k) \leq S_f(Z) - s_f(Z) < \varepsilon.$$

Ist umgekehrt $f \in R[a,c]$ und $f \in R[c,b]$, sowie $\varepsilon > 0$, so finden wir Zerlegungen Z_1 von $[a,c]$ und Z_2 von $[c,b]$ mit $S_f(Z_k) - s_f(Z_k) < \varepsilon/2$ für $k = 1, 2$. Wir setzen $Z := Z_1 \cup Z_2$ und erhalten

$$S_f(Z) - s_f(Z) = S_f(Z_1) + S_f(Z_2) - s_f(Z_1) - s_f(Z_2) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon.$$

Die Behauptung folgt also wieder mithilfe von 10.4.]

Beispiel: Für alle $a, b \in \mathbb{R}$ mit $a < b$ und alle $n \in \mathbb{N}_0$ gilt

$$\int_a^b x^n\,dx = \frac{b^{n+1} - a^{n+1}}{n+1}.$$

10.10 Satz: Sei $f : [a,b] \to \mathbb{R}$ eine beschränkte Funktion. Existiert zu jedem $\varepsilon > 0$ eine Funktion $g \in R[a,b]$ mit
$$\forall x \in [a,b] : |f(x) - g(x)| < \varepsilon,$$
so gilt $f \in R[a,b]$.

Bew: Verwende 10.4. Sei $\varepsilon > 0$. Wähle $g \in R[a,b]$ mit
$$\forall x \in [a,b] : |f(x) - g(x)| < \frac{\varepsilon}{3(b-a)} =: \delta$$
und eine Zerlegung $Z = \{x_0, x_1, \ldots, x_n\}$ mit $S_g(Z) - s_g(Z) < \varepsilon/3$. Dann gilt
$$S_f(Z) - s_f(Z) \leq |S_f(Z) - S_g(Z)| + \underbrace{S_g(Z) - s_g(Z)}_{<\varepsilon/3} + |s_g(Z) - s_f(Z)|.$$

Sei nun I_j wie in 10.1 und für $h \in \{f, g\}$:
$$M_j(h) := \sup h(I_j) \quad \text{und} \quad m_j(h) := \inf h(I_j).$$
Dann gilt für jedes $j = 1, \ldots, n$ (Übung!):
$$|M_j(f) - M_j(g)| \leq \delta \quad \text{und} \quad |m_j(f) - m_j(g)| \leq \delta.$$
Somit erhält man
$$|S_f(Z) - S_g(Z)| \leq \varepsilon/3 \quad \text{und} \quad |s_f(Z) - s_g(Z)| \leq \varepsilon/3,$$
und es folgt $S_f(Z) - s_f(Z) < \varepsilon$.

Bemerkung: Sind $f, g \in R[a,b]$ mit
$$\forall x \in [a,b] : |f(x) - g(x)| < \varepsilon,$$
so gilt (vergleiche auch 10.13(2))
$$|\int_a^b f\,dx - \int_a^b g\,dx| \leq \varepsilon(b-a).$$

10.11 Anwendung (Integration von Potenzreihen): Sei $\sum_{n=0}^{\infty} a_n(x-x_0)^n$ eine reelle Potenzreihe mit Konvergenzradius $R \in (0, \infty]$ und $I := (x_0 - R, x_0 + R)$. Dann gilt für jedes $y \in I$:
$$\int_{x_0}^{y} \left(\sum_{n=0}^{\infty} a_n(x-x_0)^n \right) dx = \sum_{n=0}^{\infty} \frac{a_n}{n+1}(y-x_0)^{n+1}.$$

Beweis für $x_0 = 0$ (sonst verschieben!): Setze $f(x) := \sum_{n=0}^\infty a_n x^n$ für $x \in I$. Sei $\varepsilon > 0$ und $y \in I$. Wir finden $n_0 \in \mathbb{N}$ mit $\sum_{n=n_0+1}^\infty |a_n||y|^n \le \varepsilon$ (da die Reihe für $x = y$ absolut konvergent ist). Somit gilt für alle $N \ge n_0$:

$$\forall x \in [-|y|, |y|] : \sum_{n=N+1}^{\infty} |a_n||x|^n \le \varepsilon.$$

Für $N \ge n_0$ setzen wir $g_N(x) := \sum_{n=0}^N a_n x^n$. Dann gilt $|f(x) - g_N(x)| \le \varepsilon$ für alle $x \in [-|y|, |y|]$. Da $g_N \in R[-|y|, |y|]$ gilt, ist $f \in R[-|y|, |y|]$. Außerdem gilt

$$|\int_0^y f(x)\,dx - \underbrace{\int_0^y g_N(x)\,dx}_{=\sum_{n=0}^N a_n y^{n+1}/(n+1)} | \le \varepsilon |y|$$

für jedes $N \ge n_0$. Damit ist auch die Formel für das Integral gezeigt.

Beispiele: (1) Für jedes $y \in \mathbb{R}$ gilt $\int_0^y \cos x\,dx = \sin y$, $\int_0^y \sin x\,dx = 1 - \cos y$, $\int_0^y e^x\,dx = e^y - 1$. Insbesondere haben wir $\int_0^\pi \sin x\,dx = 2$, $\int_0^{\pi/2} \cos x\,dx = 1$ und $\int_0^1 e^x\,dx = e - 1$.
(2) Ist $a \ne 0$, so erhalten wir

$$\int_0^y \sum_{n=0}^\infty a_n (ax)^n\,dx = \sum_{n=0}^\infty a_n a^n \frac{y^{n+1}}{n+1} = \frac{1}{a} \sum_{n=0}^\infty a_n \frac{(ay)^{n+1}}{n+1}$$

für y mit $ay \in I$.
Es ist also $\int_0^y \cos(ax)\,dx = a^{-1}\sin(ay)$, $\int_0^y \sin(ax)\,dx = \frac{1-\cos(ay)}{a}$, $\int_0^y e^{ax}\,dx = \frac{e^{ay}-1}{a}$.
Insbesondere ist für $a > 0$: $\int_0^y e^{-ax}\,dx = \frac{1}{a}(1 - e^{-ay})$ und $\lim_{y\to\infty} \int_0^y e^{-ax}\,dx = 1/a$.
Als weitere Anwendung notieren wir

$$\int_0^y \sin x \cos x\,dx = \frac{1}{2}\int_0^y \sin(2x)\,dx = \frac{1}{4}(1 - \cos(2y)).$$

Damit ist insbesondere $\int_0^\pi \sin x \cos x\,dx = 0 = \int_0^{2\pi} \sin x \cos x\,dx$ und $\int_0^{\pi/2} \sin x \cos x\,dx = 1/2$.

10.12 Riemann-Summen

Sei $f : [a, b] \to \mathbb{R}$ beschränkt und $Z = \{x_0, x_1, \ldots, x_n\}$ eine Zerlegung, sowie I_j, M_j, m_j wie in 10.1. Dann heißt

$$|Z| := \max\{|I_j| : j = 1, 2, \ldots, n\}$$

die *Feinheit* von Z.

Ein n-Tupel $\xi = (\xi_1, \ldots, \xi_n)$ heißt *passender Zwischenvektor*, wenn $\xi_j \in I_j$ für jedes $j = 1, \ldots, n$ gilt. Für einen solchen heißt

$$\sigma_f(Z, \xi) := \sum_{j=1}^{n} f(\xi_j)|I_j|$$

eine *Riemannsche Summe*. Wegen $m_j \leq f(\xi_j) \leq M_j$ gilt dabei $s_f(Z) \leq \sigma_f(Z, \xi) \leq S_f(Z)$.

Satz: Sei $f \in R[a, b]$ und $(Z_m)_{m\in\mathbb{N}}$ eine Folge von Zerlegungen mit $|Z_m| \to 0$, sowie $(\xi^{(m)})$ eine Folge von passenden Zwischenvektoren. Dann gilt

$$\lim_{m\to\infty} \sigma_f(Z_m, \xi^{(m)}) = \int_a^b f\, dx.$$

(ohne Beweis)

Bemerkung: Es gilt auch die Umkehrung: Haben alle Folgen Riemannscher Summen, deren Feinheit gegen Null geht, denselben Limes, so gilt $f \in R[a, b]$ und der Limes ist gleich dem Integral $\int_a^b f(x)\, dx$.

10.13 Satz: Seien $f, g \in R[a, b]$ und $D := f([a, b])$.

(1) Sei $L \geq 0$ und $h : D \to \mathbb{R}$ eine Funktion mit

$$\forall s, t \in D : |h(t) - h(s)| \leq L|t - s|. \tag{L}$$

Dann gilt $h \circ f \in R[a, b]$.

(2) Es ist $|f| \in R[a, b]$ und

$$|\int_a^b f\, dx| \leq \int_a^b |f|\, dx \qquad \text{(Dreiecksungleichung für Integrale).}$$

(3) Es ist $fg \in R[a, b]$.

(4) Ist $c > 0$ und gilt $|f(x)| \geq c$ für alle $x \in [a, b]$, so ist $1/f \in R[a, b]$.

Bemerkung: Eine Funktion h mit (L) heißt *Lipschitz-stetig* mit *Lipschitz-Konstante* L. Eine solche Funktion ist insbesondere stetig (sogar gleichmäßig stetig).

Beweis: (1) Sei $Z = \{x_0, x_1, \ldots, x_n\}$ eine Zerlegung und I_j, M_j, m_j wie in 10.1, sowie

$$\tilde{M}_j := \sup h \circ f(I_j) \quad \text{und} \quad \tilde{m}_j := \inf h \circ f(I_j).$$

Wie im Beweis von 10.8 zeigt man

$$\tilde{M}_j - \tilde{m}_j = \sup\{\underbrace{h(f(x)) - h(f(\tilde{x}))}_{\leq L|f(x)-f(\tilde{x})|} : x, \tilde{x} \in I_j\} \leq L(M_j - m_j).$$

Man erhält $S_{h\circ f}(Z) - s_{h\circ f}(Z) \leq L(S_f(Z) - s_f(Z))$, und $h \circ f \in R[a,b]$ folgt mithilfe von 10.4.

(2) Wende (1) an auf $h(t) = |t|$. Es gilt $||t| -| s|| \leq |t - s|$ für alle $t, s \in \mathbb{R}$, dh (L) gilt mit $L = 1$. Also ist $|f| \in R[a,b]$. Weiter ist

$$|\int_a^b f\,dx| = \max\{\int_a^b \underbrace{f}_{\leq |f|}\,dx, \int_a^b \underbrace{(-f)}_{\leq |f|}\,dx\} \leq \int_a^b |f|\,dx.$$

(3) Wegen $fg = \frac{1}{4}((f+g)^2 - (f-g)^2)$ reicht es, $f^2 \in R[a,b]$ zu zeigen. Da f beschränkt ist, gibt es $\gamma > 0$ mit $|f(x)| \leq \gamma$ für alle $x \in [a,b]$. Also gilt $D \subseteq [-\gamma, \gamma]$. Für $s, t \in D$ gilt somit:

$$|t^2 - s^2| = |t + s||t - s| \leq (|t| + |s|)|t - s| \leq 2\gamma|t - s|.$$

Mit $h(t) = t^2$ in (1) folgt $f^2 \in R[a,b]$.

(4) Nach Voraussetzung gilt $D \subset (-\infty, -c] \cup [c, \infty)$. Somit ist für $t, s \in D$:

$$\left|\frac{1}{t} - \frac{1}{s}\right| = \frac{|s-t|}{|s||t|} \leq \frac{1}{c^2}|t - s|.$$

Mit $h(t) := 1/t$ in (1) folgt die Behauptung.

Ende Woche 9

10.14 Satz: Sei $f \in R[a,b]$ und $x_0 \in [a,b]$.

(1) Es gilt

$$\int_a^{x_0+h} f(x)\,dx \to \int_a^{x_0} f(x)\,dx \quad (h \to 0).$$

(2) Ist f stetig in x_0, so gilt

$$\frac{1}{h}\int_{x_0}^{x_0+h} f(x)\,dx \to f(x_0) \quad (h \to 0).$$

Hierbei betrachten wir für $x_0 = a$ nur $h \to 0+$ und für $x_0 = b$ nur $h \to 0-$.

Beweis: (1) Wir wählen $\gamma > 0$ mit $|f(x)| \leq \gamma$ für alle $x \in [a,b]$. Es gilt dann für $h \neq 0$ mit $x_0, x_0 + h \in [a,b]$:

$$\begin{aligned} &|\int_a^{x_0+h} f(x)\,dx - \int_a^{x_0} f(x)\,dx| = |\int_{x_0}^{x_0+h} f(x)\,dx| \\ \leq &\int_{\min\{x_0,x_0+h\}}^{\max\{x_0,x_0+h\}} \underbrace{|f(x)|}_{\leq\gamma}\,dx \leq \gamma|h| \to 0\ (h \to 0). \end{aligned}$$

(2) Sei $\varepsilon > 0$. Da f in x_0 stetig ist, finden wir $\delta > 0$ mit $|f(x) - f(x_0)| \leq \varepsilon$ für $x \in [a,b]$ mit $|x - x_0| < \delta$. Für $|h| < \delta$ mit $x_0, x_0 + h \in [a,b]$ gilt wegen $h^{-1} \int_{x_0}^{x_0+h} 1 \, dx = 1$:

$$
\begin{aligned}
& |\frac{1}{h} \int_{x_0}^{x_0+h} f(x) \, dx - f(x_0)| = |\frac{1}{h} \int_{x_0}^{x_0+h} (f(x) - f(x_0)) \, dx| \\
\leq \quad & \frac{1}{|h|} \int_{\min\{x_0, x_0+h\}}^{\max\{x_0, x_0+h\}} \underbrace{|f(x) - f(x_0)|}_{\leq \varepsilon} \, dx \leq \varepsilon.
\end{aligned}
$$

Daraus folgt die Behauptung.

11 Differentialrechnung

In diesem Abschnitt sei $I \subset \mathbb{R}$ stets ein Intervall.

11.1 Differenzierbarkeit

Sei $f : I \to \mathbb{R}$ eine Funktion und $x_0 \in I$.

Idee: Approximation von f "in der Nähe von x_0" durch eine *lineare* Funktion (da lineare Funktionen leichter zu behandeln sind):

$$f(x) = f(x_0) + \frac{f(x) - f(x_0)}{x - x_0}(x - x_0) \sim f(x_0) + a(x - x_0).$$

Definition: f heißt in $x_0 \in I$ *differenzierbar* (*dbar*), falls der Limes $\lim_{x \to x_0} \frac{f(x)-f(x_0)}{x-x_0} \in \mathbb{R}$ existiert. Dieser Grenzwert heißt *die Ableitung von f in x_0*, Bezeichnung: $f'(x_0)$.

Die Funktion f heißt *auf I differenzierbar*, falls f in jedem $x \in I$ differenzierbar ist. In diesem Fall heißt $f' : I \to \mathbb{R}$, $x \mapsto f'(x)$, die *Ableitung von f auf I*.

Bemerkung: Es ist $\lim_{x \to x_0} \frac{f(x)-f(x_0)}{x-x_0} = \lim_{h \to 0} \frac{f(x_0+h)-f(x_0)}{h}$.

Beispiele: (1) $f : I \to \mathbb{R}$, $f(x) = c \in \mathbb{R}$, ist auf I differenzierbar mit $f' = 0$ auf I.

(2) $f : \mathbb{R} \to \mathbb{R}$, $f(x) = |x|$ ist in $x_0 = 0$ **nicht** differenzierbar, denn für $h \neq 0$ ist

$$\frac{f(h) - f(0)}{h} = \frac{|h|}{h} = \begin{cases} 1 & , h > 0 \\ -1 & h < 0 \end{cases}$$

Also existiert $\lim_{h \to 0} \frac{f(h)-f(0)}{h}$ nicht.

(3) $I = \mathbb{R}$, $n \in \mathbb{N}$, $f(x) = x^n$. f ist auf $\mathbb{R}$ differenzierbar mit $f'(x_0) = nx_0^{n-1}$ für jedes $x_0 \in \mathbb{R}$. Für $x \neq x_0$ ist nämlich nach 4.11(1)

$$\frac{x^n - x_0^n}{x - x_0} = \sum_{k=0}^{n-1} x^{n-1-k} x_0^k \to nx_0^{n-1} \ (x \to x_0).$$

(4) exp, sin, cos, sinh, cosh sind auf $\mathbb{R}$ differenzierbar mit $\exp' = \exp$, $\sin' = \cos$, $\cos' = -\sin$, $\sinh' = \cosh$ und $\cosh' = \sinh$. Für $x, h \in \mathbb{R}$ mit $h \neq 0$ gilt nämlich nach 7.11(1) und 8.4:

$$\frac{e^{x+h} - e^x}{h} = e^x \underbrace{\frac{e^h - 1}{h}}_{\to 1} \to e^x \ (h \to 0),$$

sowie nach 7.12(3) und 8.4:

$$\frac{\sin(x+h) - \sin x}{h} = \frac{\sin x \cos h + \cos x \sin h - \sin x}{h} = \sin x \underbrace{\frac{\cosh -1}{h}}_{\to 0} + \cos x \underbrace{\frac{\sin h}{h}}_{\to 0} \to \cos x \ (h \to 0).$$

Die Beweise für cos, sinh und cosh sind analog.

(5) Ist $f : I \to \mathbb{R}$ stetig, $a \in I$, und $F : I \to \mathbb{R}$ definiert durch $F(x) := \int_a^x f(y)\,dy$, so ist F auf I differenzierbar mit $F' = f$ auf I (vgl. 10.14, es ist

$$\frac{1}{h}\left(F(x_0+h) - F(x_0)\right) = \frac{1}{h}\int_{x_0}^{x_0+h} f(y)\,dy.$$

Satz: Ist f in x_0 differenzierbar, so ist f in x_0 stetig.

Beweis: Da I ein Intervall ist, ist $x_0 \in I$ Häufungspunkt von I. Für $x \in I \setminus \{x_0\}$ ist dann

$$f(x) - f(x_0) = \frac{f(x) - f(x_0)}{x - x_0}(x - x_0) \to f'(x_0) \cdot 0 = 0 \ (x \to x_0).$$

Also ist f in x_0 stetig.

11.2 Ableitungsregeln: Seien $f, g : I \to \mathbb{R}$ Funktionen, die in $x_0 \in I$ differenzierbar sind, und $\alpha, \beta \in \mathbb{R}$. Dann gilt:

(1) $\alpha f + \beta g$ in differenzierbar in $x_0 \in I$ und

$$(\alpha f + \beta g)'(x_0) = \alpha f'(x_0) + \beta g'(x_0).$$

(2) fg ist differenzierbar in x_0 und

$$(fg)'(x_0) = f'(x_0)g(x_0) + f(x_0)g'(x_0) \qquad \textit{Produktregel.}$$

(3) Ist $g(x_0) \neq 0$, so gibt es ein $\delta > 0$ mit $g(x) \neq 0$ für alle $x \in I \cap U_\delta(x_0) =: J$. Die Funktion $\frac{f}{g} : J \to \mathbb{R}$ ist in x_0 differenzierbar und

$$\left(\frac{f}{g}\right)'(x_0) = \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g(x_0)^2} \qquad \textit{Quotientenregel.}$$

Beweis: (1) ist klar wegen 6.3 und 8.4.

(2) Für $x \in I \setminus \{x_0\}$ gilt

$$\frac{(fg)(x) - (fg)(x_0)}{x - x_0} = \frac{f(x) - f(x_0)}{x - x_0} g(x) + f(x_0)\frac{g(x) - g(x_0)}{x - x_0},$$

woraus für $x \to x_0$ die Behauptung folgt (beachte, dass $g(x) \to g(x_0)$, da g nach 11.1 in x_0 stetig ist.

(3) Wegen (2) reicht $f = 1$. Die Existenz von δ erhalten wir aus 8.3, da g in x_0 stetig ist. Beachte, dass J ein Interball ist. Für $x \in J \setminus \{x_0\}$ ist

$$\frac{1/g(x) - 1/g(x_0)}{x - x_0} = \frac{\frac{g(x_0) - g(x)}{x - x_0}}{g(x)g(x_0)},$$

woraus für $x \to x_0$ die Behauptung folgt. Wieder beachte man $g(x) \to g(x_0)$ für $x \to x_0$.

Beispiele: $\tan = \frac{\sin}{\cos}$ und $\tanh = \frac{\sinh}{\cosh}$ sind auf $\mathbb{R} \setminus (\pi\mathbb{Z} + \pi/2)$ bzw. auf $\mathbb{R}$ differenzierbar, und es gilt:

$$\tan' = \frac{1}{\cos^2} = 1 + \tan^2 \text{ auf } \mathbb{R} \setminus (\pi\mathbb{Z} + \pi/2), \quad \tanh' = \frac{1}{\cosh^2} = 1 - \tanh^2 \text{ auf } \mathbb{R}.$$

[Es ist nämlich

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

Der Beweis für tanh ist analog.]

11.3 Kettenregel: Sei $f : I \to \mathbb{R}$ in $x_0 \in I$ differenzierbar und $J \subset \mathbb{R}$ ein Intervall mit $f(I) \subset J$. Sei $g : J \to \mathbb{R}$ in $y_0 := f(x_0)$ differenzierbar. Dann ist $g \circ f : I \to \mathbb{R}$ in x_0 differenzierbar und

$$(g \circ f)'(x_0) = g'(f(x_0)) \cdot f'(x_0)$$

("äußere Ableitung mal innere Ableitung").

Beweis: Die Idee ist zu schreiben

$$\frac{g(f(x)) - g(f(x_0))}{x - x_0} = \frac{g(f(x)) - g(f(x_0))}{f(x) - f(x_0)} \frac{f(x) - f(x_0)}{x - x_0}.$$

Da man dabei eventuell durch Null dividiert, setzen wir

$$d : J \to \mathbb{R}, d(y) := \begin{cases} \frac{g(y)-g(y_0)}{y-y_0} & , y \neq y_0 \\ g'(y_0) & , y = y_0 \end{cases} .$$

Dann gilt $d(y) \to g'(y_0)$ für $y \to y_0$, also auch $d(f(x)) \to g'(f(x_0))$ für $x \to x_0$, da f in x_0 stetig ist. Außerdem gilt $g(y) - g(y_0) = d(y)(y - y_0)$ für **alle** $y \in J$, also

$$\frac{g(f(x)) - g(f(x_0))}{x - x_0} = d(f(x)) \frac{f(x) - f(x_0)}{x - x_0} \to g'(f(x_0)) \cdot f'(x_0) \quad (x \to x_0).$$

Beispiel: Sei $a > 0$ und $h(x) = a^x$ für $x \in \mathbb{R}$. Dann ist h auf $\mathbb{R}$ differenzierbar und für jedes $x \in \mathbb{R}$ gilt:

$$h'(x) = (a^x)' = a^x \ln a.$$

11.4 Satz über die Umkehrfunktion: Sei $f : I \to \mathbb{R}$ stetig und streng monoton auf I. Ist f in $x_0 \in I$ differenzierbar und $f'(x_0) \neq 0$, so ist die Umkehrfunktion $f^{-1} : f(I) \to \mathbb{R}$ in $y_0 := f(x_0)$ differenzierbar und

$$(f^{-1})'(y_0) = \frac{1}{f'(x_0)} = \frac{1}{f'(f^{-1}(y_0))}.$$

Beweis: Nach 8.8 ist $f(I)$ ein Intervall. Sei $y \in f(I)$ und $x := f{-}1(y)$. Dann gilt

$$\frac{f^{-1}(y) - f^{-1}(y_0)}{y - y_0} = \frac{x - x_0}{f(x) - f(x_0)} \to \frac{1}{f'(x_0)} \quad (y \to y_0),$$

da wegen der Stetigkeit von f^{-1} aus $y \to y_0$ folgt $x \to x_0$.

Beispiele: Die Funktionen $\ln : (0, \infty) \to \mathbb{R}$, $\arctan : \mathbb{R} \to (-\pi/2, \pi/2)$, $\operatorname{Arsinh} : \mathbb{R} \to \mathbb{R}$, $\operatorname{Artanh} : (-1, 1) \to \mathbb{R}$ sind differenzierbar. Dabei ist ln die Umkehrfunktion von $f(x) = e^x$, also

$$\ln' y = (e^{\ln y})^{-1} = \frac{1}{y} \quad \text{für jedes } y > 0.$$

Die Funktionen $\arcsin : (-1, 1) \to (-\pi/2, \pi/2)$, $\arccos : (-1, 1) \to (0, \pi)$ und $\operatorname{Arcosh} : (1, \infty) \to (0, \infty)$ sind differenzierbar.

11.5 Lokale Extremstellen

Sei $D \subset \mathbb{R}$ mit $D \neq \emptyset$, $f : D \to \mathbb{R}$ eine Funktion und $x_0 \in D$.

Definition: g hat in x_0 ein *lokales Maximum* [*Minimum*], falls es ein $\delta > 0$ gibt mit

$$\forall x \in D \cap U_\delta(x_0) : g(x) \leq g(x_0) \quad [\text{ bzw. } g(x) \geq g(x_0)],$$

dh $\forall x \in D$: $|x - x_0| < \delta \Rightarrow g(x) \leq g(x_0)$ [bzw. $g(x) \geq g(x_0)$].

Ein lokales Maximum/Minimum wird auch als *relatives Maximum/Minimum* bezeichnet.

Ein *relatives* oder *lokales Extremum* ist ein lokales Maximum oder ein lokales Minimum.

11.6 Satz: Die Funktion $f : I \to \mathbb{R}$ habe in $x_0 \in I$ ein lokales Extremum und sei in x_0 differenzierbar. Gibt es ein $\delta > 0$ mit $U_\delta(x_0) \subset I$, so ist $f'(x_0) = 0$.

Bemerkung: Ist $M \subset \mathbb{R}$ eine Menge, so heißen Punkte $x_0 \in M$, für die es ein $\delta > 0$ mit $U_\delta(x_0) \subset M$ gibt, *innere Punkte von* M.

Da I ein Intervall ist, ist $x_0 \in I$ genau dann ein innerer Punkt von I, wenn $x_0 \notin \{\sup I, \inf I\}$ gilt.

Beweis des Satzes: Wir nehmen an, dass f in x_0 einlokales Maximum hat (sonst betrachte $-f$). Wir können weiter annehmen, dass das δ aus 11.5 mit dem δ aus dem Satz übereinstimmt. Für $x \in (x_0, x_0 + \delta)$ gilt dann $f(x) \leq f(x_0)$, also $\frac{f(x)-f(x_0)}{x-x_0} \leq 0$ und damit $f'(x_0) = \lim_{x \to x_0+} \frac{f(x)-f(x_0)}{x-x_0} \leq 0$. Für $x \in (x_0 - \delta, x_0)$ gilt $f(x) \leq f(x_0)$ und $\frac{f(x)-f(x_0)}{x-x_0} \geq 0$, woraus $f'(x_0) \geq 0$ folgt.

11.7 Mittelwertsatz der Differentialrechnung (MWS): Sei $f : [a, b] \to \mathbb{R}$ stetig und auf (a, b) differenzierbar. Dann gibt es ein $\xi \in (a, b)$ mit

$$f'(\xi) = \frac{f(b) - f(a)}{b - a}.$$

Beweis: Wir betrachten zunächst den Fall $f(a) = f(b)$. Nach 8.14 finden wir $x_1, x_2 \in [a, b]$ mit $f(x_1) \leq f(x) \leq f(x_2)$ für alle $x \in [a, b]$. Somit hat f in x_1 ein lokales Minimum und in x_2 ein lokales Maximum.

Fall 1, $x_1 \in (a, b)$: Dann wenden wir 11.6 an und erhalten $f'(x_1) = 0$.

Fall 2, $x_1 \in \{a, b\}$, $x_2 \in (a, b)$: Wir wenden 11.6 an und erhalten $f'(x_2) = 0$.

Fall 3, $x_1, x_2 \in \{a, b\}$: Dann ist f konstant und $f'(x) = 0$ für jedes $x \in (a, b)$.

Im allgemeinen Fall setzen wir $g(x) := f(x) - \frac{f(b)-f(a)}{b-a}(x - a)$ für $x \in [a, b]$ und wenden das Gezeigte auf g an (es ist $g(a) = f(a) = g(b)$). Wir erhalten $\xi \in (a, b)$ mit

$$0 = g'(\xi) = f'(\xi) - \frac{f(b) - f(a)}{b - a}$$

wie gewünscht.

11.8 Folgerungen: Seien $f, g : I \to \mathbb{R}$ auf I differenzierbar.

(1) f ist auf I konstant $\Leftrightarrow$ $f' = 0$ 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 MWS ist $f(b) = f(a)$ für alle $a, b \in I$.

(2) Wende (1) an auf $f - g$.

(3) Ist $f' \geq 0$ auf I, so gilt für $x, y \in I$ mit $x < y$:

$$\frac{f(y) - f(x)}{y - x} = f'(\xi) \geq 0$$

für ein $\xi \in (x, y)$. Es folgt $f(y) \geq f(x)$. Die anderen Aussagen beweist man analog.

11.9 Hauptsatz der Differential- und Integralrechnung

Sei $f : [a, b] \to \mathbb{R}$ stetig.

(1) Ist $G : [a, b] \to \mathbb{R}$ eine differenzierbare Funktion mit $G' = f$ auf $[a, b]$, so gilt

$$\int_a^b f(x)\, dx = G(b) - G(a) \ \left(=: G(x)|_a^b =: [G(x)]_a^b\right).$$

(2) Setzt man $F : [a, b] \to \mathbb{R}$, $x \mapsto F(x) := \int_a^x f(\xi)\, d\xi$, so ist F auf $[a, b]$ differenzierbar mit $F' = f$ auf $[a, b]$.

Beweis: (2) haben wir in 10.14 gezeigt (vgl. Beispiel 11.1(5)).

(1) folgt aus (2) und 11.8(2): Es ist $G = F + c$, also $G(b) - G(a) = F(b) - F(a)$. Ende Woche 10

11.10 Stammfunktionen

Definition: Sind $f, F : I \to \mathbb{R}$ Funktionen, wobei F auf I differenzierbar ist mit $F' = f$ auf I, so heißt F eine *Stammfunktion von f auf I*.

Bemerkung: Der Hauptsatz 11.9 besagt, dass man zur Berechnung von Integralen Stammfunktionen verwenden kann. Für eine Stammfunktion von $f : I \to \mathbb{R}$ schreibt man deshalb auch $\int f(x)\,dx$ (*unbestimmtes Integral*).

Zwei Stammfunktionen von $f : I \to \mathbb{R}$ unterscheiden sich nach 11.8(2) nur durch eine additive Konstante.

Beispiel: $\int \cos x\,dx = \sin x + c$, wobei $c \in \mathbb{R}$ eine feste Konstante ist.

Definition: Eine Funktion $f : I \to \mathbb{R}$ heißt *stetig differenzierbar auf I*, falls f auf I differenzierbar ist und $f' : I \to \mathbb{R}$ stetig ist. Wir schreiben dafür auch $f \in C^1(I)$.

Bemerkung: Ist $f \in C^1(I)$, so gilt $\int f'\,dx = f$ auf I.

Beispiel: Sei $f : [0,1] \to \mathbb{R}$ gegeben durch $f(x) := \begin{cases} x^{3/2}\sin(\frac{1}{x}) & , x \in (0,1] \\ 0 & , x = 0 \end{cases}$. Für $x \in (0,1]$ ist f in x differenzierbar mit

$$f'(x) = \frac{3}{2}\sqrt{x}\sin(\frac{1}{x}) - \frac{1}{\sqrt{x}}\cos(\frac{1}{x}).$$

Beachte, dass diese Funktion auf $(0,1]$ nicht beschränkt ist.

Die Funktion f ist in 0 differenzierbar mit $f'(0) = 0$, denn für jedes $h \in (0,1]$ gilt

$$\left|\frac{f(h) - f(0)}{h}\right| = \sqrt{h}|\sin(\frac{1}{h})| \leq \sqrt{h} \to 0 \quad (h \to 0+),$$

dh $\lim_{h\to 0+} \frac{f(h)-f(0)}{h} = 0$.

Die Funktion f ist auf $[0,1]$ differenzierbar, aber $f \notin C^1([0,1])$.

Die beiden folgenden Integrationsregeln ergeben sich über den Hauptsatz aus Differentiationsregeln.

11.11 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

$$\int_a^b f'(x)g(x)\,dx = f(x)g(x)|_a^b - \int_a^b f(x)g'(x)\,dx.$$

Beweis: Es gilt $(fg)' = f'g + fg'$ auf I. Somit hat $f'g + fg'$ die Stammfunktion fg auf I, woraus die Behauptungen folgen (für die zweite Formel verwenden wir den Hauptsatz 11.9).

Beispiele: (1) $\int \underbrace{x}_{g} \underbrace{e^x}_{f'} \, dx = \underbrace{x}_{g} \underbrace{e^x}_{f} - \int \underbrace{1}_{g'} \cdot \underbrace{e^x}_{f} \, dx = xe^x - e^x$.

(2) $\int \ln x \, dx = \int \underbrace{1}_{f'} \cdot \underbrace{\ln x}_{g} \, dx = \underbrace{x}_{f} \underbrace{\ln x}_{g} - \int \underbrace{x}_{f} \underbrace{\frac{1}{x}}_{g'} \, dx = x \ln x - x$.

11.12 Integration durch Substitution: Seien $I, J \subset \mathbb{R}$ Intervalle, $f : I \to \mathbb{R}$ stetig und $g \in C^1(J)$ mit $g(J) \subseteq I$. Dann ist

$$\int f(g(t))g'(t) \, dt = \int f(x) \, dx|_{x=g(t)} \quad \text{auf } J.$$

Ist $g'(t) \neq 0$ für alle $t \in J$, so ist g auf J streng monoton und

$$\int f(x) \, dx = \int f(g(t))g'(t) \, dt|_{t=g^{-1}(x)} \quad \text{auf } g(J).$$

Ist $I = [a, b]$ und $g(\alpha) = a$, $g(\beta) = b$ sowie $J = \begin{cases} [\alpha, \beta] & , \alpha \leq \beta \\ [\beta, \alpha] & , \alpha > \beta \end{cases}$, so gilt

$$\int_a^b f(x) \, dx = \int_\alpha^\beta f(g(t))g'(t) \, dt$$

Beweis: Wähle eine Stammfunktion F von f auf I. Dann ist $G := F \circ g$ eine Stammfunktion von $h := (f \circ g) \cdot g'$ auf J. Also ist

$$\int h(t) \, dt = G(t) = F(g(t)) = \int f(x) \, dx|_{x=g(t)}$$

auf J. Ist $g'(t) \neq 0$ für alle $t \in J$, so ist nach dem Zwischenwertsatz entweder $g' > 0$ auf J oder es ist $g' < 0$ auf J. In jedem Fall ist g streng monoton auf J und besitzt also eine Umkehrfunktion $g^{-1} : g(J) \to J$ (beachte auch, dass $g(J)$ ein Intervall ist). Dann ist

$$\int h(t) \, dt|_{t=g^{-1}(x)} = G(g^{-1}(x)) = F(g(g^{-1}(x)) = F(x) = \int f(x) \, dx$$

auf $g(J)$. Schließlich verwenden wir den Hauptsatz.

Merkregel: Ist $y = y(x)$ eine differenzierbare Funktion, so schreibt man für die Ableitung y' auch $\frac{dy}{dx}$. Substituiere nun in $\int f(x) \, dx$: $x = g(t)$, dh fasse x als Funktion von t auf. Dann ist $\frac{dx}{dt} = g'(t)$ und man erhält (formal!) "$dx = g'(t)\,dt$" (dies ist nur eine Schreibweise, da "dx" oder "dt" hier **keine mathematische Bedeutung** tragen).

Beispiele: (1) $\int_0^1 \frac{e^{2x}+1}{e^x}\,dx$, substituiere $t = e^x$, also $x = \ln t$. Dann ist $dx = dt/t$ und aus $x : 0 \to 1$ ergibt sich $t : 1 \to e$. Wir erhalten:

$$\int_0^1 \frac{e^{2x}+1}{e^x}\,dx = \int_1^e (1+t^{-2})\,dt = (t - t^{-1})|_1^e = e - 1/e.$$

(2) $\int_0^1 \sqrt{1-x^2}\,dx$, substituiere $x = \sin t$. Dann ist $dx = \cos t\,dt$ und aus $x : 0 \to 1$ ergibt sich $t : 0 \to \pi/2$. Wir erhalten

$$\int_0^1 \sqrt{1-x^2}\,dx = \int_0^{\pi/2} \underbrace{\sqrt{1-\sin^2 t}}_{=\sqrt{\cos^2 t}} \cos t\,dt = \int_0^{\pi/2} \cos^2 t\,dt,$$

da $\cos \geq 0$ auf $[0, \pi/2]$ ist. Nun schreiben wir

$$\cos(2t) = \cos^2 t - \sin^2 t = \cos^2 t - (1 - \cos^2 t) = 2\cos^2 t - 1,$$

also $\cos^2 t = \frac{1}{2}(1 + \cos(2t))$, und erhalten

$$\int_0^{\pi/2} \cos^2 t\,dt = \frac{1}{2}\left(\int_0^{\pi/2} 1\,dt + \underbrace{\int_0^{\pi/2} \cos(2t)\,dt}_{=0}\right) = \frac{\pi}{4}.$$

Das ist der Flächeninhalt eines Viertelkreises mit Radius 1.

In 11.1 bei der Definition der Differenzierbarkeit haben wir gesagt, dass sich die lineare Approximation einer Funktion oft leichter behandeln lässt als die Funktion selber. Das folgende ist dafür ein Beispiel.

11.13 Das Newton-Verfahren: Gegeben sei eine differenzierbare Funktion $f : I \to \mathbb{R}$, für die man eine *Nullstelle* berechnen möchte, dh also ein $x^* \in I$ mit $f(x^*) = 0$ (wir nehmen an, dass f in I genau eine Nullstelle hat).

Zu einem gewählten *Startwert* $x_0 \in I$ betrachtet man statt f die linearisierte Funktion $x \mapsto f(x_0) + f'(x_0)(x - x_0)$ und bestimmt für diese eine Nullstelle x_1. Dies ist möglich, wenn $f'(x_0) \neq 0$ ist; dann ist $x_1 = x_0 - f'(x_0)^{-1} f(x_0)$.

Man hofft, dass x_1 eine Näherung für x^* ist, und bestimmt rekursiv

$$x_{k+1} = x_k - f'(x_k)^{-1} f(x_k), \quad k \in \mathbb{N}_0,$$

(unter der Voraussetzung, dass immer $f'(x_k) \neq 0$ ist), in der Hoffnung, dass die Folge (x_k) gegen x^* konvergiert.

Bemerkung: (1) Das Newton-Verfahren konvergiert unter den Voraussetzungen $f \in C^1(I)$ und $f'(x^*) \neq 0$ zwar "schnell", jedoch im allgemeinen nur, wenn x_0 schon genügend nahe bei x^* liegt.

(2) Wenn $f \in C^1(I)$ ist mit $f'(x^*) \neq 0$ und (x_k) konvergiert, so ist $\lim x_k$ eine Nullstelle von f.

Satz: Sei $f : I \to \mathbb{R}$ differenzierbar, $x^* \in I$ mit $f(x^*) = 0$ und $f'(x^*) > 0$. Ist $x_0 > x^*$ und f' auf $[x^*, x_0]$ monoton wachsend, so konvergiert das Newton-Verfahren gegen x^*.

Beweis: Wie man leicht sieht, gilt $x_k \geq x^*$ für jedes k. Für jedes $k \in \mathbb{N}_0$ gilt:

$$x_{k+1} - x^* = x_k - x^* - f'(x_k)^{-1} f(x_k)$$

und nach dem MWS

$$f(x_k) = f(x_k) - f(x^*) = f'(\xi)(x_k - x^*)$$

für ein $\xi \in (x^*, x_k)$, also

$$x_{k+1} - x^* = (1 - f'(\xi)/f'(x_k))(x_k - x^*) \leq \underbrace{(1 - f'(x^*)/f'(x_0))}_{=:\alpha \in [0,1)}(x_k - x^*).$$

Wir erhalten $x_k - x^* \leq \alpha^k (x_0 - x^*) \to 0$ für $k \to \infty$.

Beispiel: $m \in \mathbb{N}$, $a > 0$ und $f(x) = x^m - a$. Für jeden Startwert $x_0 > \sqrt[m]{a}$ konvergiert das Newton-Verfahren gegen $\sqrt[m]{a}$.

11.14 Verallgemeinerter Mittelwertsatz: Seien $f, g : [a, b] \to \mathbb{R}$ stetig und auf (a, b) differenzierbar mit $g'(x) \neq 0$ für alle $x \in (a, b)$. Dann gilt $g(a) \neq g(b)$ und es gibt ein $\xi \in (a, b)$ mit

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(\xi)}{g'(\xi)}.$$

Beweis: Wäre $g(b) = g(a)$, so gäbe es nach dem MWS ein $\eta \in (a, b)$ mit $g'(\eta) = 0$ im Widerspruch zur Voraussetzung. Definiere $h : [a, b] \to \mathbb{R}$ durch

$$h(x) := (f(b) - f(a))g(x) - (g(b) - g(a))f(x).$$

Dann ist h stetig und auf (a, b) differenzierbar mit

$$h(a) = f(b)g(a) - g(b)f(a) = h(b)$$

und

$$h'(x) = (f(b) - f(a))g'(x) - (g(b) - g(a))f'(x).$$

Die Behauptung folgt aus dem MWS, angewandt auf h.

11.15 Die Regeln von de l'Hospital: Seien $a \in \mathbb{R} \cup \{-\infty\}$, $b \in \mathbb{R} \cup \{\infty\}$ und $L \in \mathbb{R} \cup \{-\infty, \infty\}$. Seien $f, g : (a, b) \to \mathbb{R}$ auf (a, b) differenzierbar mit $g'(x) \neq 0$ für alle $x \in (a, b)$. Weiter sei $\lim_{x \to a} \frac{f'(x)}{g'(x)} = L$.

(a) Ist $\lim_{x \to a} f(x) = \lim_{x \to a} g(x) = 0$, so gilt

$$\frac{f(x)}{g(x)} \to L \quad (x \to a).$$

(b) Ist $\lim_{x \to a} g(x) = \pm\infty$, so gilt

$$\frac{f(x)}{g(x)} \to L \quad (x \to a).$$

(hier ohne Beweis, ginge mit 11.14.)

Beispiele: (1) Für $a, b > 0$ gilt

$$\lim_{x \to 0+} \frac{a^x - b^x}{x} = \lim_{x \to 0+} \frac{a^x \ln a - b^x \ln b}{1} = \ln a - \ln b.$$

(2) Es ist

$$\lim_{x \to \infty} \frac{\ln x}{x} = \lim_{x \to \infty} \frac{1/x}{1} = 0.$$

(3) Es ist

$$\lim_{x \to 0+} x \ln x = \lim_{x \to 0+} \frac{\ln x}{1/x} = \lim_{x \to 0+} \frac{1/x}{-1/x^2} = \lim_{x \to 0+} (-x) = 0.$$

(4) Aus (3) folgt mit der Stetigkeit der Exponentialfunktion:

$$\lim_{x \to 0+} x^x = \lim_{x \to 0+} e^{x \ln x} = e^0 = 1.$$

Bemerkung: Hier ist es jeweils so, dass erst die Existenz des letzten Limes die Existenz des ersten Limes garantiert (vgl. die Regeln oben).

12 Uneigentliche Integrale

Die folgenden Vereinbarungen sollen für das gesamte Kapitel gelten:

> Ist $I \subseteq \mathbb{R}$ ein Intervall und $f : I \to \mathbb{R}$ eine Funktion, so sei f über jedes Intervall $[a, b] \subset I$ integrierbar.
>
> Es seien stets $a, b \in \mathbb{R}$, $\alpha \in \mathbb{R} \cup \{-\infty\}$ und $\beta \in \mathbb{R} \cup \{\infty\}$ mit $a < \beta$ und $\alpha < b$.

12.1 Konvergenz uneigentlicher Integrale

Definition: Sei $f : [a, \beta) \to \mathbb{R}$ (bzw. $(\alpha, b] \to \mathbb{R}$) eine Funktion. Das *uneigentliche Integral* $\int_a^\beta f(x)\,dx$ (bzw. $\int_\alpha^b f(x)\,dx$) heißt *konvergent*, falls der Limes $\lim_{r\to\beta-} \int_a^r f(x)\,dx$ (bzw. $\lim_{r\to\alpha+} \int_r^b f(x)\,dx$) existiert und reell ist. In diesem Fall setzt man

$$\int_a^\beta f(x)\,dx := \lim_{r\to\beta-} \int_a^r f(x)\,dx \quad \Big(\text{bzw.} \quad \int_\alpha^b f(x)\,dx := \lim_{r\to\alpha+} \int_r^b f(x)\,dx\Big).$$

Ein nicht konvergentes uneigentliches Integral heißt *divergent*.

Beispiele: (1) Sei $\gamma > 0$. Dann gilt für jedes $r > 1$:

$$\int_1^r \frac{1}{x^\gamma}\,dx = \begin{cases} \ln r & ,\gamma = 1 \\ \frac{r^{1-\gamma}-1}{1-\gamma} & ,\gamma \neq 1 \end{cases},$$

also ist $\int_1^\infty x^{-\gamma}\,dx$ konvergent genau dann, wenn $\gamma > 1$ ist. In diesem Fall ist das Integral $= 1/(\gamma - 1)$.

(2) Für $r > 0$ gilt

$$\int_0^r \frac{1}{1+x^2}\,dx = \arctan r \to \frac{\pi}{2} \quad (r \to \infty).$$

Also ist $\int_0^\infty \frac{1}{1+x^2}\,dx$ konvergent und $= \pi/2$.

(3) Sei $\gamma > 0$. Dann gilt für $r \in (0, 1)$:

$$\int_r^1 \frac{1}{x^\gamma}\,dx = \begin{cases} -\ln r & ,\gamma = 1 \\ \frac{1-r^{1-\gamma}}{1-\gamma} & ,\gamma \neq 1 \end{cases},$$

also ist $\int_0^1 x^{-\gamma}\,dx$ konvergent genau dann, wenn $\gamma < 1$ ist. In diesem Fall ist das Integral $= 1/(1 - \gamma)$.

(4) Analog zu (2): $\int_{-\infty}^0 \frac{1}{1+x^2}\,dx$ ist konvergent und $= \pi/2$.

Definition: Sei $f : (\alpha, \beta) \to \mathbb{R}$ eine Funktion. Das *uneigentliche Integral* $\int_\alpha^\beta f(x)\,dx$ heißt *konvergent*, falls es ein $c \in (\alpha, \beta)$ so gibt, dass die uneigentlichen Integrale $\int_\alpha^c f(x)\,dx$ und $\int_c^\beta f(x)\,dx$ **beide** konvergent sind. In diesem Fall setzt man

$$\int_\alpha^\beta f(x)\,dx := \int_\alpha^c f(x)\,dx + \int_c^\beta f(x)\,dx,$$

und die Definition ist unabhängig von $c \in (\alpha, \beta)$.

Das Integral $\int_\alpha^\beta f(x)\,dx$ heißt *divergent*, falls $\int_\alpha^\beta f(x)\,dx$ nicht konvergent ist.

Beispiele: (5) Sei $\gamma > 0$. Nach den Beispielen (1) und (3) ist $\int_0^\infty x^{-\gamma}\,dx$ divergent.

(6) Nach den Beispielen (2) und (4) ist $\int_{-\infty}^\infty \frac{1}{1+x^2}\,dx$ konvergent und $= \pi$.

Bemerkung: Wir betrachten im folgenden nur den Fall von Funktionen $f : [a, \beta) \to \mathbb{R}$. Entsprechendes gilt jeweils auch für Funktionen $f : (\alpha, b] \to \mathbb{R}$ bzw. $f : (\alpha, \beta) \to \mathbb{R}$.

12.2 Cauchykriterium

Satz (ohne Beweis): $\int_a^\beta f(x)\,dx$ konvergiert genau dann, wenn es zu jedem $\varepsilon > 0$ ein $c = c(\varepsilon) \in (a, \beta)$ gibt mit

$$\forall u, v \in (c, \beta) : \left| \int_u^v f(x)\,dx \right| < \varepsilon.$$

Beispiel: $\int_0^\infty \frac{\sin x}{x}\,dx$ ist konvergent.

Zunächst beachten wir $\frac{\sin x}{x} \to 1$ für $x \to 0+$, weshalb das Integral bei 0 nicht uneigentlich ist. Für $v > u > 0$ gilt:

$$\begin{aligned}
\left|\int_u^v \frac{\sin x}{x}\,dx\right| &= \left|\int_u^v \underbrace{\frac{1}{x}}_{g} \underbrace{\sin x}_{f'}\,dx\right| \\
&= \left|-\frac{\cos x}{x}\Big|_u^v - \int_u^v \frac{\cos x}{x^2}\,dx\right| \\
&\le \frac{|\cos u|}{u} + \frac{|\cos v|}{v} + \int_u^v \frac{|\cos x|}{x^2}\,dx \\
&\le \frac{1}{u} + \frac{1}{v} + \left[-\frac{1}{x}\right]_u^v = \frac{2}{u}.
\end{aligned}$$

Sei $\varepsilon > 0$. Für $v > u > 2/\varepsilon$ gilt dann $|\int_u^v \frac{\sin x}{x}\,dx| \le 2/u < \varepsilon$.

12.3 Absolute Konvergenz uneigentlicher Integrale

Definition: $\int_a^\beta f(x)\,dx$ heißt *absolut konvergent*, falls $\int_a^\beta |f(x)|\,dx$ konvergent ist.

Beispiel: $\int_1^\infty \frac{\sin x}{x}\,dx$ ist **nicht** absolut konvergent. Für $k \in \mathbb{N}_0$ gilt nämlich

$$\int_{k\pi}^{(k+1)\pi} \frac{|\sin x|}{x}\,dx \geq \underbrace{\int_0^\pi |\sin x|\,dx}_{=2} \frac{1}{(k+1)\pi} = \frac{2}{(k+1)\pi},$$

also für $n \in \mathbb{N}$:

$$\int_0^{(n+1)\pi} \frac{|\sin x|}{x}\,dx \geq \sum_{k=0}^n \frac{2}{(k+1)\pi} \to \infty \quad (n \to \infty).$$

Mithilfe von 12.2 kann man zeigen (ähnlich wie bei Reihen, vgl. 7.3 und 7.5): Ende Woche 11

12.4 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) $\int_1^\infty \frac{x}{\sqrt{1+x^5}}\,dx$; setze $f(x) := \frac{x}{\sqrt{1+x^5}}$ und $g(x) := x^{-3/2}$. Dann gilt $|f(x)| = f(x) \leq g(x)$ für jedes $x \geq 1$. Da $\int_1^\infty x^{-3/2}\,dx$ konvergiert, konvergiert auch $\int_1^\infty f(x)\,dx$ und es gilt:

$$\int_1^\infty f(x)\,dx \leq \int_1^\infty g(x)\,dx = 2.$$

(2) $\int_1^\infty \underbrace{\frac{x}{x^2+7x}}_{=:f(x)}\,dx$; setze $g(x) = 1/x$. Dann gilt $f(x)/g(x) = \frac{x^2}{x^2+7x} \to 1$ $(x \to \infty)$ und wir finden $c \geq 1$ mit $f/g \geq 1/2$ auf $[c,\infty)$, dh es gilt $f(x) \geq g(x)/2 = (2x)^{-1}$ für alle $x \geq c$ (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 $\int_1^\infty f(x)\,dx$.

Warnung: Sind $\int_a^\beta f\,dx$ und $\int_a^\beta g\,dx$ konvergent, so muss $\int_a^\beta fg\,dx$ **nicht** konvergieren!

13 Fortsetzung der Differentialrechnung

Der folgende Satz kann in Anwendungen nützlich sein und folgt aus der Regel von de l'Hospital.

13.1 Satz: Sei $f : [a,b] \to \mathbb{R}$ eine Funktion und $x_0 \in (a,b)$. Weiter sei f auf (a,x_0) und (x_0,b) differenzierbar. Existiert der Grenzwert $\alpha := \lim_{x\to x_0} f'(x) \in \mathbb{R}$ und ist f in x_0 stetig, so ist f in x_0 differenzierbar mit $f'(x_0) = \alpha$ und f' ist in x_0 stetig.

Beispiel: Die Funktion $f : \mathbb{R} \to \mathbb{R}$, $f(x) := \left\{ \begin{array}{cl} e^{-1/x} & , x > 0 \\ 0 & , x \leq 0 \end{array} \right.$ ist auf $\mathbb{R}$ stetig differenzierbar. Es gilt

$$f'(x) = \left\{ \begin{array}{cl} x^{-2}e^{-1/x} & , x > 0 \\ 0 & , x < 0 \end{array} \right. .$$

Wegen $\lim_{x\to 0+} f(x) = \lim_{x\to 0+} e^{-1/x} = 0$ ist f in 0 stetig. Wegen $\lim_{x\to 0+} f'(x) = \lim_{x\to 0+} x^{-2}e^{-1/x} = \lim_{s\to\infty} s^2 e^{-s} = 0$ ist f nach dem Satz in 0 differenzierbar mit $f'(0) = 0$ und f' ist stetig auf $\mathbb{R}$.

13.2 Höhere Ableitungen

Definition: Sei $I \subset \mathbb{R}$ ein Intervall und $f : I \to \mathbb{R}$ sei auf I differenzierbar.

(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 $f'''(x_0)$, $f^{(4)}(x_0)$ etc. bzw. f''', $f^{(4)}$, ...

(b) Sei $n \in \mathbb{N}$. f heißt *auf I n-mal stetig differenzierbar*, falls f auf I n-mal differenzierbar ist und $f, f', f'', \ldots, f^{(n)} : I \to \mathbb{R}$ stetig sind. Dafür schreiben wir $f \in C^n(I)$.

Außerdem: $f \in C^0(I) = C(I)$, falls f auf I stetig ist, und $f \in C^\infty(I)$, falls $f \in C^n(I)$ für jedes $n \in \mathbb{N}$ gilt, dh wenn f auf I beliebig oft differenzierbar ist.

Beispiele: (1) Für $f : \mathbb{R} \to \mathbb{R}$, $f(x) = e^x$, gilt $f \in C^\infty(\mathbb{R})$. Hier ist $f^{(n)}(x) = e^x$ für jedes $n \in \mathbb{N}$, $x \in \mathbb{R}$.

(2) Es gilt $\sin, \cos \in C^\infty(\mathbb{R})$. Hier ist $\sin' = \cos$, $\sin'' = \cos' = -\sin$, $\sin''' = -\sin' = -\cos$ und $\sin^{(4)} = -\cos' = \sin$, etc.

(3) Für die Funktion f aus dem Beispiel in 13.1 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:

$$f^{(n)}(x) = \left\{ \begin{array}{cl} p_n(\frac{1}{x})e^{-1/x} & , x > 0 \\ 0 & , x < 0 \end{array} \right. ,$$

wobei p_n ein Polynom ist (für $n = 1$ ist $p_1(s) = s^2$ und also $p_1(\frac{1}{x}) = \frac{1}{x^2}$, vgl. Beispiel in 13.1 oben; die allgemeine Aussage zeigt man durch Induktion nach n). Wir wenden den Satz 13.1 jetzt wiederholt an (dh sukzessive auf f, f', f'', etc.), wobei wir beachten, dass für jedes $n \in \mathbb{N}$ gilt:

$$\lim_{x\to 0+} f^{(n)}(x) = \lim_{x\to 0+} p_n(\frac{1}{x})e^{-1/x} = \lim_{s\to\infty} p_n(s)e^{-s} = 0.$$

Wir erhalten so, dass f, f', f'', etc. auf $\mathbb{R}$ stetig differenzierbar sind.

13.3 Ableitung von Potenzreihen

Sei $\sum_{n=0}^{\infty} a_n(x - x_0)^n$ eine reelle Potenzreihe mit Konvergenradius $R \in (0, \infty]$, $I := (x_0 - R, x_0 + R)$ und

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

Dann ist f auf I differenzierbar, und für jedes $x \in I$ gilt

$$f'(x) = \sum_{n=1}^{\infty} na_n(x - x_0)^{n-1}.$$

Bemerkung: Der Satz besagt, dass man Potenzreihen im Inneren des Konvergenzintervalls gliedweise differenzieren kann. Da f' wieder eine Potenzreihe mit demselben Konvergenzradius ist, kann man den Satz wiederholt anwenden und erhält damit: $f \in C^\infty(I)$ und für alle $k \in \mathbb{N}$ gilt:

$$f^{(k)}(x) = \sum_{n=k}^{\infty} n(n-1)(n-2)\cdot\ldots\cdot(n-(k-1))a_n(x - x_0)^{n-k}, \quad x \in I.$$

Für $x = x_0$ erhält man

$$f^{(k)}(x_0) = k(k-1)(k-2)\cdot\ldots\cdot(k-(k-1))a_k = k!a_k.$$

Somit gilt

$$a_k = \frac{f^{(k)}(x_0)}{k!} \quad \text{für alle } k \in \mathbb{N}_0,$$

und wir erhalten die Darstellung

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k, \quad x \in I.$$

Beweis des Satzes: Setze $g(x) := \sum_{n=1}^{\infty} na_n(x - x_0)^{n-1}$. Diese Potenzreihe hat auch den Konvergenzradius R und ist somit auf I stetig. Nach 10.11 gilt für $x \in I$:

$$a_0 + \int_{x_0}^{x} g(x)\,dx = \sum_{n=0}^{\infty} a_n(x - x_0)^n = f(x).$$

Nach dem Hauptsatz ist f auf I differenzierbar und $f' = g$ auf I.

Beispiel: Für $x \in (-1,1)$ gilt

$$\arctan'(x) = \frac{1}{1+x^2} = \frac{1}{1-(-x^2)} = \sum_{k=0}^{\infty}(-x^2)^k = \sum_{k=0}^{\infty}(-1)^k x^{2k}.$$

Wegen $\arctan 0 = 0$ ist also

$$\arctan(x) = \sum_{k=0}^{\infty}(-1)^k \frac{x^{2k+1}}{2k+1}, \quad x \in (-1,1).$$

13.4 Satz von Taylor: Sei $n \in \mathbb{N}_0$, $f \in C^n(I)$ und $f^{(n)}$ sei auf I differenzierbar. Seien $x, x_0 \in I$. Dann gibt es ein ξ zwischen x und x_0 mit

$$\begin{aligned} f(x) &= f(x_0) + \frac{f'(x_0)}{1!}(x-x_0) + \ldots + \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n + \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-x_0)^{n+1} \\ &= \underbrace{\sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k}_{n\text{-tes Taylorpolynom}} + \underbrace{\frac{f^{(n+1)}(\xi)}{(n+1)!}(x-x_0)^{n+1}}_{\text{Restglied}}. \end{aligned}$$

Bemerkung: Für $n = 0$ ist das der Mittelwertsatz.

Definition: Für $n \in \mathbb{N}_0$ schreiben wir

$$T_n(f;x_0)(x) := \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k, \quad x \in I,$$

für das n-te Taylorpolynom von f bei Entwicklung um x_0.

Bemerkung: Falls $f \in C^\infty(I)$, so ist f um x_0 in eine auf I konvergente Potenzreihe entwickelbar, falls für jedes $x \in I$ gilt:

$$T_n(f;x_0)(x) \to f(x) \quad (n \to \infty).$$

Gleichbedeutend damit ist, dass das entsprechende Restglied für jedes $x \in I$ für $n \to \infty$ gegen Null geht.

Warnung: Das ist **nicht** für jedes $f \in C^\infty(I)$ der Fall!

Beispiel: Wir betrachten $f : \mathbb{R} \to \mathbb{R}$, $f(x) = \begin{cases} e^{-1/x} & , x > 0 \\ 0 & , x < 0 \end{cases}$, in $x_0 = 0$. Es gilt $f^{(n)}(0) = 0$ für jedes $n \in \mathbb{N}_0$, also auch $T_n(f;0)(x) = 0$ für alle $n \in \mathbb{N}$, $x \in \mathbb{R}$. Für **kein** $x > 0$ gilt somit $T_n(f;0)(x) \to f(x)$!

13.5 Lokale Extrema: Sei $n \geq 2$, $I \subset \mathbb{R}$ ein Intervall, $f \in C^n(I)$ und $x_0 \in I$ ein innerer Punkt von I (dh kein Randpunkt). Weiter sei

$$f'(x_0) = f''(x_0) = \ldots = f^{(n-1)}(x_0) = 0 \quad \textbf{und} \quad f^{(n)}(x_0) \neq 0.$$

(a) Ist n gerade und $f^{(n)}(x_0) > 0$ [bzw. $f^{(n)}(x_0) < 0$], so hat f in x_0 ein lokales Minimum [bzw. ein lokales Maximum].

(b) Ist n ungerade, so hat f in x_0 kein lokales Extremum.

Bemerkung: In der Anwendung ist meist $n = 2$. Ist $f'(x_0) = 0$, so gilt:

für $f''(x_0) > 0$ hat f in ein lokales Minimum,

für $f''(x_0) < 0$ hat f in ein lokales Maximum,

für $f''(x_0) = 0$ erhält man keine Entscheidung.

Beweis: $f^{(n)}$ ist stetig auf I mit $f^{(n)}(x_0) \neq 0$, somit gibt es $\delta > 0$ mit

$$f^{(n)}(\xi) f^{(n)}(x_0) > 0 \quad \text{für alle } \xi \in U_\delta(x_0) \subset I.$$

Nach dem Satz von Taylor und der Voraussetzung gibt es für jedes $x \in U_\delta(x_0)$ ein $\xi \in U_\delta(x_0)$ mit

$$f(x) - f(x_0) = \frac{f^{(n)}(x_0)}{n!}(x - x_0)^n.$$

Nun betrachte man das Vorzeichen der rechten Seite.

Beispiel: Sei $p > 1$, $\alpha > 0$. Bestimme das Maximum von $f(x) := \alpha x - x^p/p$ über $x \geq 0$.

Es gilt $f(0) = 0$ und $f(x) \to -\infty$ für $x \to \infty$. Für $x > 0$ ist $f'(x) = \alpha - x^{p-1}$ und $f''(x) = -(p-1)x^{p-2} < 0$. Weiter ist $f'(x_0) = 0$ genau dann, wenn $x_0 = \alpha^{1/(p-1)}$ ist. Der Funktionswert an dieser Stelle ist $f(\alpha^{1/(p-1)}) = (1 - \frac{1}{p})\alpha^{p/(p-1)} > 0$. Dies ist das gesuchte Maximum.

13.6 Identitätssatz für Potenzreihen: Sei $r > 0$ und $\sum_{n=0}^\infty a_n(x-x_0)^n$, $\sum_{n=0}^\infty b_n(x-x_0)^n$ reelle Potenzreihen, die auf $I := (x_0 - r, x_0 + r)$ konvergieren. Seien $f, g : I \to \mathbb{R}$ gegeben durch $f(x) := \sum_{n=0}^\infty a_n(x - x_0)^n$ und $g(x) := \sum_{n=0}^\infty b_n(x - x_0)^n$, $x \in I$. Es gebe eine streng monotone Folge $(x_m)_{m\in\mathbb{N}}$ in I mit $x_m \to x_0$ für $m \to \infty$ und $f(x_m) = g(x_m)$ für jedes $m \in \mathbb{N}$. Dann gilt $a_n = b_n$ für alle $n \in \mathbb{N}_0$.

Bemerkung: Insbesondere gilt also:

$$f = g \text{ auf } I \quad \Rightarrow \quad \forall n \in \mathbb{N}_0 : a_n = b_n,$$

aber die Voraussetzung lässt sich sehr abschwächen.

Beweis: Wir dürfen $g = 0$ und $b_n = 0$ annehmen (durch Betrachtung von $f - g$ und $a_n - b_n$).

Annahme: es gibt $n_0 := \min\{n \in \mathbb{N}_0 : a_n \neq 0\}$. Dann gilt $f(x) = \sum_{n=n_0}^{\infty} a_n(x - x_0)^n$ und

$$\frac{f(x)}{(x - x_0)^{n_0}} = \sum_{n=n_0}^{\infty} a_n(x - x_0)^{n-n_0} \to a_{n_0} \quad (x \to x_0).$$

Also ist nach Voraussetzung $a_{n_0} = \lim_{m\to\infty} \frac{f(x_m)}{(x_m - x_0)^{n_0}} = 0$ im Widerspruch zur Annahme.

Beispiel: Wir wollen eine Funktion y bestimmen mit $y'(x) = xy(x)$ und $y(0) = 1$ und machen einen *Potenzreihenansatz* $y(x) = \sum_{n=0}^{\infty} a_n x^n$. Wegen $y(0) = 0$ ist $a_0 = 1$.

Nach 13.3 ist $y'(x) = \sum_{n=1}^{\infty} n a_n x^{n-1}$. Außerdem ist

$$xy(x) = \sum_{n=0}^{\infty} a_n x^{n+1} = \sum_{n=2}^{\infty} a_{n-2} x^{n-1}.$$

Durch *Koeffizientenvergleich* (dh nach 13.6) erhalten wir $a_1 = 0$ und

$$na_n = a_{n-2}, \quad n \geq 2.$$

Durch Induktion erhält man hieraus $a_3 = a_5 = a_7 = \ldots = 0$ und $a_{2k} = (2^k k!)^{-1}$ für $k = 0, 1, 2, \ldots$. Somit

$$y(x) = \sum_{k=0}^{\infty} \frac{x^{2k}}{2^k k!} = \sum_{k=0}^{\infty} \frac{1}{k!}\left(\frac{x^2}{2}\right)^k = e^{x^2/2}, \quad x \in \mathbb{R},$$

und dies genügt tatsächlich den geforderten Bedingungen.

13.7 Zur Exponentialfunktion: Für jedes $x \in \mathbb{R}$ gilt:

$$\lim_{n\to\infty} \left(1 + \frac{x}{n}\right)^n = e^x.$$

Beweis: Für $n > -x$ ist die linke Seite $= \exp(n \ln(1 + \frac{x}{n}))$. Also reicht zu zeigen

$$n \ln(1 + \frac{x}{n}) \to x \quad (n \to \infty).$$

Das ist klar für $x = 0$. Für $x \neq 0$ erhalten wir

$$\lim_{n\to\infty} \frac{\ln(1 + \frac{x}{n})}{\frac{x}{n}} = \lim_{h\to 0} \frac{\ln(1 + h) - \ln 1}{h} = \ln'(1) = \frac{1}{t}\Big|_{t=1} = 1,$$

woraus die Behauptung folgt.

Ende
Woche 12

14 Grundzüge der linearen Algebra

14.1 Verknüpfungen in $\mathbb{R}^n$ und $\mathbb{C}^n$

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

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

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

Wir erklären Verknüpfungen

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

durch

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

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

Für die **Anschauung** besonders wichtig sind $\mathbb{R}^2$ und $\mathbb{R}^3$.

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

14.2 Vektorraumaxiome

Setzen wir $V = \mathbb{K}^n$, wobei $n \in \mathbb{N}$, so haben die in 14.1 definierten Verknüpfungen $+ : V \times V \to V$ und $\cdot : \mathbb{K} \times V \to V$ die folgenden Eigenschaften:

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

Definition: Ist $V \neq 0$ eine Menge mit Verknüpfungen $+ : V \times V \to V$ und $\cdot : \mathbb{K} \times V \to V$, für welche die Eigenschaften $(V1)$–$(V7)$ gelten, so heißt V ein *Vektorraum über* $\mathbb{K}$ oder ein $\mathbb{K}$-*Vektorraum* ($\mathbb{K}$-*VR*). Die Elemente eines Vektorraumes heißen *Vektoren*.

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

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

14.3 Ein wichtiges Beispiel

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

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

Allgemeiner sei nun $M \neq \emptyset$ eine Menge und

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

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

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

so ist $\mathbb{K}^M$ bzgl. dieser Verknüpfungen $+ : \mathbb{K}^M \times \mathbb{K}^M \to \mathbb{K}^M$ und $\cdot : \mathbb{K} \times \mathbb{K}^M \to \mathbb{K}^M$ ein $\mathbb{K}$-Vektorraum.

Beispiele: (1) $M = \mathbb{N}$: $\mathbb{K}^\mathbb{N}$ ist die Menge aller Folgen in $\mathbb{K}$, dh

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

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

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

(2) $M = [a, b]$ mit $a < b$ reell: Die Menge $\mathbb{K}^{[a,b]}$ aller Funktionen $f : [a, b] \to \mathbb{K}$ ist ein $\mathbb{K}$-Vektorraum. Für $f, g \in \mathbb{K}^{[a,b]}$ und $\alpha \in \mathbb{K}$ sind $f + g$ und αf *punktweise* erklärt (wie sonst ja auch schon!).

Bemerkung: Ist V ein $\mathbb{K}$-Vektorraum, so gilt für alle $x \in V$: $0 \cdot x = 0$ und $-x = (-1) \cdot x$.

Beweis: Es ist $0 \cdot x = (0 + 0) \cdot x = (0 \cdot x) + (0 \cdot x)$, also

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

Weiter ist

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

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

14.4 Untervektorräume

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

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

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

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

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

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

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

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

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

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

14.5 Der lineare Aufspann

Definition: Ist V ein $\mathbb{K}$-Vektorraum, $n \in \mathbb{N}$ und sind $v_1, v_2, \ldots, v_n$ Vektoren aus V, so heißt jeder Vektor der Form

$$\sum_{j=1}^{n} \alpha_j v_j \quad \text{mit } \alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{K}$$

eine *Linearkombination* (*LK*) der Vektoren $v_1, v_2, \ldots, v_n$.

Satz: Ist V ein $\mathbb{K}$-Vektorraum und $M \subset V$, so ist

$$\operatorname{lin}(M) := \{\sum_{j=1}^{n} \alpha_j v_j : n \in \mathbb{N}_0, \alpha_j \in \mathbb{K}, v_j \in M\}$$

ein Untervektorraum von V, genannt der *lineare Aufspann* (oder das *lineare Erzeugnis*) von M (erinnert sei an die Konvention $\sum_{j=1}^{0} \alpha_j v_j = 0$).

Bemerkung: (1) $\operatorname{lin}(M)$ besteht aus allen Linearkombinationen von Vektoren aus M (und ev. Null, denn $\operatorname{lin}(\emptyset) = \{0\}$).

(2) Offenbar ist $\operatorname{lin}(M)$ der *kleinste* Untervektorraum von V, der M enthält. Insbesondere ist U ein Untervektorraum von V genau dann, wenn $\operatorname{lin}(U) = U$ gilt.

Beispiele: (1) In $V = \mathbb{R}^2$ gilt $\operatorname{lin}(\{(1, 0)\}) = \mathbb{R} \times \{0\}$.

(2) In $V = \mathbb{R}^2$ gilt $\operatorname{lin}(\{(1, 0), (0, 1)\}) = \mathbb{R}^2 = \operatorname{lin}(\{(1, 0), (0, 1), (1, 1)\}$.

(3) Für K aus Beispiel 14.4(3) gilt $\operatorname{lin}(K) = \mathbb{R}^2$.

(4) Für C aus Beispiel 14.4(4) gilt $\operatorname{lin}(C) = \mathbb{R}^2$.

14.6 Lineare Unabhängigkeit

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

Definition: Eine Teilmenge $M \subset V$ heißt *linear unabhängig*, falls für jedes $v \in M$ gilt, dass $\mathrm{lin}(M \setminus \{v\}) \neq \mathrm{lin}(M)$, das heißt also falls man keinen Vektor aus M entfernen kann, ohne den linearen Aufspann zu verkleinern. Ist M nicht linear unabhängig, so heißt M *linear abhängig.*

Bemerkung: Eine Teilmenge $M \subset V$ ist linear abhängig genau dann, wenn es einen Vektor $v_0 \in M$ gibt, der Linearkombination von Vektoren aus $M \setminus \{v_0\}$ ist. Dies ist genau dann der Fall, wenn es $n \in \mathbb{N}$ und $v_1, v_2, \ldots, v_n \in M$ und $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{K}$ gibt mit

$$\sum_{j=1}^{n} \alpha_j v_j = 0, \quad \text{wobei nicht alle } \alpha_j = 0 \text{ sind.}$$

Definition: Man nennt n Vektoren $v_1, v_2, \ldots, v_n \in V$ *linear unabhängig*, falls für alle $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{K}$ gilt:

$$\sum_{j=1}^{n} \alpha_j v_j = 0 \quad \Longrightarrow \quad \alpha_1 = \alpha_2 = \ldots = \alpha_n = 0.$$

Man sagt dazu, dass man den Nullvektor nur als *triviale Linearkombination* der $v_1, v_2, \ldots, v_n$ erhalten kann.

Sind $v_1, v_2, \ldots, v_n$ nicht linear unabhängig, so heißen sie *linear abhängig.*

Beispiele: (1) $M = \emptyset$ ist linear unabhängig.

(2) $M = \{v\}$ ist linear unabhängig genau dann, wenn $v \neq 0$ gilt.

(3) In $\mathbb{R}^2$ sind $(1,0), (0,1)$ linear unabhängig, aber $(1,0), (0,1), (1,1)$ sind linear abhängig.

(4) In $\mathbb{K}^n$ sind die *Einheitsvektoren* $e_1, e_2, \ldots, e_n$ linear unabhängig. Hierbei ist

$$e_j := (0, \ldots, 0, \underbrace{1}_{j\text{-te Stelle}}, 0, \ldots, 0) \quad \text{für } j = 1, 2, \ldots, n.$$

Mithilfe des *Kroneckersymbols* $\delta_{jk} = \begin{cases} 1 & , j = k \\ 0 & , j \neq k \end{cases}$ ist also $e_j = (\delta_{jk})_{k=1}^n$.

14.7 Zeilenumformungen

Seien n Vektoren $v_1, v_2, \ldots, v_n$ aus dem $\mathbb{K}^m$ gegeben, deren lineare Unabhängigkeit wir untersuchen wollen. Dabei sei $v_j = (v_{j1}, v_{j2}, \ldots, v_{jm})$ für $j = 1, 2, \ldots, n$. Wir schreiben die Vektoren $v_1, v_2, \ldots, v_n$ als **Zeilen** in eine Martix A, dh

$$A = \begin{pmatrix} v_{11} & v_{12} & \ldots & v_{1m} \\ v_{21} & v_{22} & \ldots & v_{2m} \\ \ldots & \ldots & \ldots & \ldots \\ v_{n1} & v_{n2} & \ldots & v_{nm} \end{pmatrix}.$$

Die Matrix A ist eine $n \times m$-Matrix mit n Zeilen und m Spalten. Wir schreiben dafür $A \in \mathbb{K}^{n \times m}$.

Die Matrix A bringen wir in eine Form, an der wir die lineare Unabhängigkeit der Zeilen ablesen können, mittels der folgenden *Zeilenumformungen* für eine gegebene $n \times m$-Matrix B mit Zeilen $w_1, w_2, \ldots, w_n \in \mathbb{K}^m$:

(Z1) Ersetze eine Zeile w_j durch αw_j, wobei $\alpha \in \mathbb{K} \setminus \{0\}$;

(Z2) Ersetze eine Zeile w_j durch $w_j + \beta w_k$, wobei $\beta \in \mathbb{K}$ und $k \neq j$;

(Z3) Vertausche die Zeilen w_j und w_k, wobei $j \neq k$.

Bemerkung: Die Zeilen von B sind genau dann linear unabhängig, wenn die Zeilen der umgeformten Matrix linear unabhängig sind.

Satz: Jede Matrix $A \in \mathbb{K}^{n \times m}$ kann man durch endlich viele Zeilenumformungen in eine Matrix $C \in \mathbb{K}^{m \times n}$ überführen, die in *Zeilenstufenform* (*ZSF*) ist. Dabei heißt

$$C = \begin{pmatrix} c_{11} & c_{12} & \ldots & c_{1m} \\ c_{21} & c_{22} & \ldots & c_{2m} \\ \ldots & \ldots & \ldots & \ldots \\ c_{n1} & c_{n2} & \ldots & c_{nm} \end{pmatrix}$$

in *Zeilenstufenform*, falls es ein $r \in \{0, 1, \ldots, n\}$ und $1 \leq k_1 < k_2 < \ldots < k_r \leq m$ gibt mit

für $j = 1, 2, \ldots, r$ gilt $c_{jk} = 0$ für $k = 1, 2, \ldots, k_j - 1$ und $\gamma_j := c_{jk_j} \neq 0$;

für $j = r + 1, \ldots, n$ gilt $c_{jk} = 0$ für alle $k = 1, 2, \ldots, m$.

Man braucht dafür sogar nur Umformungen der Art (Z2) und (Z3).

Beispiel 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:

$$\begin{pmatrix} 0 & \gamma_1 & * & * & * & * & * & * \\ 0 & 0 & 0 & \gamma_2 & * & * & * & * \\ 0 & 0 & 0 & 0 & \gamma_3 & * & * & * \\ 0 & 0 & 0 & 0 & 0 & 0 & \gamma_4 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Folgerung: Die n Zeilen der Matrix A sind genau dann linear unabhängig, wenn für eine zugehörige Matrix in Zeilenstufenform $r = n$ gilt. Dies kann wegen $r \leq m$ höchstens dann sein, wenn $n \leq m$ ist.

Beispiel: Wir untersuchen die Vektoren $v_1 = (1,3,-2,4)$, $v_2 = (-1,-1,5,-9)$, $v_3 = (2,0,-13,23)$, $v_4 = (1,5,1,-2) \in \mathbb{R}^4$ und erhalten nacheinander die Matrizen

$$\begin{pmatrix} 1 & 3 & -2 & 4 \\ -1 & -1 & 5 & -9 \\ 2 & 0 & -13 & 23 \\ 1 & 5 & 1 & -2 \end{pmatrix} \quad \begin{pmatrix} 1 & 3 & -2 & 4 \\ 0 & 2 & 3 & -5 \\ 0 & -6 & -9 & 15 \\ 0 & 2 & 3 & -6 \end{pmatrix} \quad \begin{pmatrix} 1 & 3 & -2 & 4 \\ 0 & 2 & 3 & -5 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \quad \begin{pmatrix} 1 & 3 & -2 & 4 \\ 0 & 2 & 3 & -5 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Die Vektoren v_1, v_2, v_3, v_4 sind nicht linear unabhängig.

Bemerkung: Man sieht aber auch, dass Zeilenumformungen den linearen Aufspann der Zeilen einer Matrix nicht ändern. Ein Blick auf die dritte Matrix zeigt, dass v_3 eine Linearkombination der Vektoren v_1, v_2, v_4 ist (es wurden bis dahin nur Umformungen der Art (Z2) vorgenommen, dh es wurden keine Zeilen vertauscht!): es gilt also

$$\mathrm{lin}(v_1, v_2, v_3, v_4) = \mathrm{lin}(v_1, v_2, v_4).$$

Ausserdem sieht man, dass v_1, v_2, v_4 linear unabhängig sind.

14.8 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 14.7) gilt:

> für $j = 1, 2, \ldots, r$ gilt: $\gamma_j = c_{jk_j} = 1$ und $c_{lk_j} = 0$ für $l = 1, 2, \ldots, j-1$ (dh oberhalb von c_{jk_j} stehen nur Nullen).

Beispiel mit $n = 5$ und $m = 8$, $r = 4$ aus 14.7:

$$\begin{pmatrix} 0 & 1 & * & 0 & 0 & * & 0 & * \\ 0 & 0 & 0 & 1 & 0 & * & 0 & * \\ 0 & 0 & 0 & 0 & 1 & * & 0 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & * \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Beispiel: Für die konkrete Matrix im Beispiel in 14.7 formen wir weiter um:

$$\begin{pmatrix} 1 & 3 & -2 & 4 \\ 0 & 2 & 3 & -5 \\ 0 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 1 & 3 & -2 & 4 \\ 0 & 1 & 3/2 & -5/2 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 & -13/2 & 23/2 \\ 0 & 1 & 3/2 & -5/2 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \quad \begin{pmatrix} 1 & 0 & -13/2 & 0 \\ 0 & 1 & 3/2 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

14.9 Basen und Dimension

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

Definition: Eine Teilmenge $B \subset V$ heißt *Basis von V*, falls B linear unabhängig ist und $\mathrm{lin}(B) = V$ gilt. V heißt *endlich-dimensional*, falls V eine endliche Basis hat, und V heißt *unendlich-dimensional*, wenn es keine endliche Basis von V gibt.

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 $\dim V$.

Beispiel: Es gilt $\dim \mathbb{K}^n = n$ und eine Basis ist gegeben durch die *Einheitsvektoren* $e_1, e_2, \ldots, e_n$.

Ende Woche 13

Bemerkung: V ist unendlich-dimensional genau dann, wenn V eine unendliche linear unabhängige Teilmenge enthält. Eine unendliche Teilmenge M von V ist linear unabhängig genau dann, wenn je endlich viele Vektoren aus M linear unabhängig sind.

Beispiel: Die Menge $P[a,b]$ der Polynomfunktionen $p : [a,b] \to \mathbb{R}$ ist ein unendlich-dimensionaler $\mathbb{R}$-Vektorraum: Da Summen und Produkte von Polynomen wieder Polynome sind, ist $P[a,b]$ ein Untervektorraum von $\mathbb{R}^{[a,b]}$. Für jedes $k \in \mathbb{N}_0$ sei p_k gegeben durch $p_k(x) := x^k$. Dann ist $\{p_k : k \in \mathbb{N}_0\}$ linear unabhängig, denn für $m \in \mathbb{N}$ und $\alpha_0, \alpha_1, \ldots, \alpha_m \in \mathbb{R}$ mit

$$\alpha_0 p_0 + \alpha_1 p_1 + \ldots + \alpha_m p_m = 0 \quad \text{auf } I$$

folgt $\alpha_0 = \alpha_1 = \ldots = \alpha_m = 0$, da nur das Nullpolynom unendlich viele Nullstellen hat (vgl. 5.4).

Bemerkung: $B \subset V$ ist Basis von V genau dann, wenn sich jeder Vektor $v \in V$ **auf genau eine Weise** als Linearkombination von Vektoren aus B schreiben lässt. Insbesondere gilt:

Ist $n \in \mathbb{N}$ und $b_1, b_2, \ldots, b_n$ eine Basis von V, so gibt es zu jedem $v \in V$ eindeutig bestimmte $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{K}$ mit

$$v = \alpha_1 b_1 + \alpha_2 b_2 + \ldots + \alpha_n b_n = \sum_{j=1}^{n} \alpha_j b_j,$$

die $\alpha_1, \alpha_2, \ldots, \alpha_n$ heißen *Koordinaten* von v bzgl. der Basis $b_1, b_2, \ldots, b_n$.

Beweis: Existenz folgt aus $v \in \mathrm{lin}\{b_1, b_2, \ldots, b_n\}$. Sind $\alpha_j, \tilde{\alpha}_j \in \mathbb{K}$ für $j = 1, \ldots, n$ mit $\sum_{j=1}^{n} \alpha_j b_j = v = \sum_{j=1}^{n} \tilde{\alpha}_j b_j$, so folgt

$$0 = v - v = \sum_{j=1}^{n} (\alpha_j - \tilde{\alpha}_j) b_j$$

und weiter $\alpha_j = \tilde{\alpha}_j$ für alle $j = 1, \ldots, n$, da $b_1, b_2, \ldots, b_n$ linear unabhängig sind.

Satz: Sei V ein $\mathbb{K}$-Vektorraum und seien $v_1, v_2, \ldots, v_n \in V$. Dann ist die Dimension $\dim \mathrm{lin}(v_1, v_2, \ldots, v_n)$ des linearen Aufspanns von $v_1, v_2, \ldots, v_n$ gerade die Maximalzahl linear unabhängiger Vektoren unter den $v_1, v_2, \ldots, v_n$.

Beispiel: Im $\mathbb{R}^3$ sei $v_1 := (1,1,0)$, $v_2 := (1,0,0)$ und $v_3 := (0,1,0)$. Dann gilt $v_1 \in \operatorname{lin}\{v_2, v_3\}$ und v_2, v_3 ist Basis von $\operatorname{lin}\{v_1, v_2, v_3\} = \operatorname{lin}\{v_2, v_3\}$. Es ist $\dim \operatorname{lin}\{v_1, v_2, v_3\} = 2$.

Warnung: Für einen $\mathbb{C}$-Vektorraum V ist es für die Dimensionsbestimmung wichtig, ob man ihn als $\mathbb{C}$-Vektorraum oder als $\mathbb{R}$-Vektorraum betrachtet, dh ob man Linearkombinationen mit komplexen oder nur mit reellen Koeffizienten zulässt. So gilt $\mathbb{C}\text{-}\dim \mathbb{C} = 1$, aber $\mathbb{R}\text{-}\dim \mathbb{C} = 2$ (wie schon an der Veranschaulichung von $\mathbb{C}$ als komplexe Zahlen*ebene* zu sehen ist). Eine reelle Basis von $\mathbb{C}$ ist gegeben durch $1, i$.

14.9 Lineare Gleichungssysteme

Seien $n, m \in \mathbb{N}$. Wir schreiben ein *lineares Gleichungssystem* (*LGS*)

$$\begin{array}{ccccc} a_{11}x_1+ & a_{12}x_2+ & \dots & +a_{1m}x_m & = b_1 \\ a_{21}x_1+ & a_{22}x_2+ & \dots & +a_{2m}x_m & = b_2 \\ \vdots & \vdots & \dots & \vdots & \vdots \\ a_{n1}x_1+ & a_{n2}x_2+ & \dots & +a_{nm}x_m & = b_n \end{array}$$

mit gegebenen $a_{jk}, b_j \in \mathbb{K}$ und gesuchten $x_1, x_2, \dots, x_m \in \mathbb{K}$ in Matrixschreibweise als

$$\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & \vdots & \dots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}$$

oder als $Ax = b$, wobei $A \in \mathbb{K}^{n\times m}$ und $b \in \mathbb{K}^n$ gegeben und $x \in \mathbb{K}^m$ gesucht ist. Hierbei verwenden wir das *Matrix-Vektor-Produkt*:

Für $A = (a_{jk})_{j=1}^{n}{}_{k=1}^{m} \in \mathbb{K}^{n\times m}$ und $x = (x_k)_{k=1}^m \in \mathbb{K}^m$ ist

$$Ax = \Big(\sum_{k=1}^m a_{jk}x_k\Big)_{j=1}^n \in \mathbb{K}^n.$$

Beachte: Wir schreiben Vektoren $x = (x_1, x_2, \dots, x_m) \in \mathbb{K}^m$, die wir bisher als m-Tupel geschrieben haben, ab jetzt ggf. als **Spaltenvektoren**! Dies ist die übliche Konvention.

Beispiel: Sei $l \in \{1, 2, \dots, m\}$. Für den l-ten Einheitsvektor $e_l = (\delta_{lk})_{k=1}^m \in \mathbb{K}^m$ erhalten wir

$$Ae_l = \Big(\sum_{k=1}^m a_{jk}\delta_{kl}\Big)_{j=1}^n = (a_{jl})_{j=1}^n \in \mathbb{K}^n,$$

dh Ae_l **ist die l-te Spalte der Matrix** A. Für $n = m = 3$ und $l = 2$ ist etwa

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \end{pmatrix}.$$

14.10 Eigenschaften des Matrix-Vektor-Produktes

Für alle $A, B \in \mathbb{K}^{n\times m}$ und $x, y \in \mathbb{K}^m$, $\alpha \in \mathbb{K}$ gilt:

$$A(x+y) = Ax + Ay, \qquad (A+B)x = Ax + Bx, \qquad A(\alpha x) = \alpha(Ax) = (\alpha A)x,$$

wobei für $A = (a_{jk})$, $B = (b_{jk})$ die Matrizen $A + B, \alpha A \in \mathbb{K}^{n\times m}$ gegeben sind durch

$$A + B = (a_{jk} + b_{jk})_{j=1}^{n}{}_{k=1}^{m}, \qquad \alpha A = (\alpha a_{jk})_{j=1}^{n}{}_{k=1}^{m},$$

dh an jeder Stelle (j, k) werden die Einträge addiert bzw. mit α multipliziert.

Satz: Mit diesen Verknüpfungen ist $\mathbb{K}^{n\times m}$ ein $\mathbb{K}$-Vektorraum der Dimension $n \cdot m$. Eine Basis ist gegeben durch die Matrizen

$$B_{pq} := (\delta_{jp}\delta_{kq})_{j=1}^{n}{}_{k=1}^{m}, \qquad p = 1, \dots, n,\ q = 1, \dots, m.$$

Die Matrix B_{pq} hat gerade an der Stelle (p, q) eine Eins und sonst Nullen.

14.11 Die Lösungsmenge eines linearen Gleichungssystems

Seien $A \in \mathbb{K}^{n\times m}$ und $b \in \mathbb{K}^n$ gegeben. Wir setzen

$$\operatorname{Kern} A := \{x \in \mathbb{K}^m : Ax = 0\} \quad \text{und} \quad \operatorname{Bild} A := \{Ax : x \in \mathbb{K}^m\}.$$

$\operatorname{Kern} A$ ist die Lösungsmenge der *homogenen Gleichung* $Ax = 0$. Die Gleichung $Ax = b$ heißt *inhomogen* und b heißt *Inhomogenität* der Gleichung.

Satz: (1) $\operatorname{Kern} A$ ist ein Untervektorraum von $\mathbb{K}^m$ und $\operatorname{Bild} A$ ist ein Untervektorraum von $\mathbb{K}^n$.

(2) $\operatorname{Bild} A$ ist der lineare Aufspann der Spalten von A.

(3) Sind $x, y \in \mathbb{K}^m$ mit $Ax = b = Ay$, so ist $x - y \in \operatorname{Kern} A$. Ist $z \in \operatorname{Kern} A$, so ist $A(x + z) = b$.

Beweis: (2) Für jeden Vektor $x = (x_1, x_2, \dots, x_m) \in \mathbb{K}^m$ gilt

$$x = \sum_{l=1}^{m} x_l e_l, \quad \text{also} \quad Ax = A(\sum_{l=1}^{m} x_l e_l) = \sum_{l=1}^{m} x_l(\underbrace{Ae_l}_{l\text{-te Spalte von } A})$$

(vgl. Beispiel aus 14.9). (1) und (3) folgen aus 14.10.

Folgerung: (1) Wir erhalten die Gesamtheit der Lösungen der inhomogenen Gleichung (falls es überhaupt welche gibt), indem zu einer **speziellen** Lösung der inhomogenen Gleichung alle Lösungen der zugehörigen homogenen Gleichung addieren.

(2) Es gilt

$$Ax = b \text{ ist lösbar} \;\Leftrightarrow\; b \in \operatorname{Bild} A \;\Leftrightarrow\; b \text{ ist LK der Spalten von } A.$$

Ist $Ax = b$ lösbar, so ist die Lösung eindeutig genau dann, wenn $\operatorname{Kern} A = \{0\}$ gilt.

Beispiele: (1) $A = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$, $b = \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right)$. Dann hat $Ax = b$ keine Lösung.

(2) Sei A wie eben und $b = \left(\begin{smallmatrix} 1 \\ 0 \end{smallmatrix}\right)$. Dann ist $Ax = b$ lösbar und die Menge aller Lösungen ist

$$\left\{ \begin{pmatrix} \alpha \\ 1 \end{pmatrix} : \alpha \in \mathbb{K} \right\} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \underbrace{\operatorname{lin}\begin{pmatrix} 1 \\ 0 \end{pmatrix}}_{= \text{Kern } A} .$$

(3) Sei $A = \left(\begin{smallmatrix} 2 & 1 \\ 0 & 1 \end{smallmatrix}\right)$ und $b = \left(\begin{smallmatrix} 0 \\ 1 \end{smallmatrix}\right)$. Dann ist $\left(\begin{smallmatrix} -1/2 \\ 1 \end{smallmatrix}\right)$ die eindeutige Lösung von $Ax = b$.

Bemerkung: Ein lineares Gleichungssystem kann also keine, genau eine oder unendlich viele Lösungen haben.

14.12 Lösungsalgorithmus

Seien $A \in \mathbb{K}^{n\times m}$ und $b \in \mathbb{K}^n$ gegeben. Zur Lösung erweitern wir die Matrix A um b als $(m+1)$-te Spalte, betrachten also $(A\,|\,b) \in \mathbb{K}^{n\times(m+1)}$. Für $A = (a_{jk})$ und $b = (b_j)$ ist

$$(A\,|\,b) = \left(\begin{array}{cccc|c} a_{11} & a_{12} & \dots & a_{1m} & b_1 \\ a_{21} & a_{22} & \dots & a_{2m} & b_2 \\ \vdots & \vdots & \dots & \vdots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nm} & b_n \end{array}\right)$$

Wir werden Zeilenumformungen verwenden (vgl. 14.7).

Beobachtung: Geht die Matrix $(\tilde{B}\,|\,\tilde{c})$ aus der Matrix $(B\,|\,c)$ durch eine Zeilenumformung hervor, so gilt

$$\{x \in \mathbb{K}^m : \tilde{B}x = \tilde{c}\} = \{x \in \mathbb{K}^m : Bx = c\},$$

dh die Lösungsmenge der entsprechenden linearen Gleichungssysteme ändert sich nicht.

Allgemein gilt: Geht $\tilde{B}$ aus B durch eine Zeilenumformung hervor und ist die k-te Spalte von B Linearkombination anderer Spalten von B, so ist die k-te Spalte von $\tilde{B}$ Linearkombination der entsprechenden Spalten von $\tilde{B}$ und zwar *mit denselben Koeffizienten*.

Algorithmus (Eliminationsverfahren nach Gauß):

(1) Matrix A um b als letzte Spalte erweitern.
(2) Die erweiterte Matrix $(A\,|\,b)$ auf ZNF (oder ZSF) bringen.
(3) Lösbarkeit und Lösung ablesen.

Zu Schritt (2) vergleiche 14.7 und 14.8.

Algorithmus Schritt (3), Ablesen der Lösbarkeit: Das lineare Gleichungssystem $Ax = b$ ist **nicht lösbar** genau dann, wenn eine ZSF von $(A\,|\,b)$ 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 14.7 ist dies genau dann der Fall, wenn $k_r = m+1$ gilt.

Beispiel: siehe Beispiel 14.11(1). Dort ist $n = m = 2$, $r = 2$, $k_1 = 2$, $k_2 = 3$.

Algorithmus Schritt (3), Ablesen der Lösung: 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 $(A\,|\,b)$ die folgende Gestalt hat (hier ist $n = 4$, $m = 5$):

$$\left(\begin{array}{ccccc|c} 1 & 0 & \alpha_1 & 0 & \beta_1 & c_1 \\ 0 & 1 & \alpha_2 & 0 & \beta_2 & c_2 \\ 0 & 0 & 0 & 1 & \beta_3 & c_3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right).$$

Man nehme die Variablen zu den Spalten, die “hinter” den Stufen stehen (im Beispiel die dritte und die fünfte Spalte), als *freie Parameter* (im Beispiel also etwa $s = x_3$ und $t = x_5$). Schreibt man die Gleichungen wieder aus, so erhält man

$$x_1 = c_1 - s\alpha_1 - t\beta_1, \quad x_2 = c_2 - s\alpha_2 - t\beta_2, \quad x_4 = c_3 - t\beta_3.$$

Also ist die Lösungsmenge von $Ax = b$ gegeben durch

$$\left\{\begin{pmatrix} c_1 \\ c_2 \\ 0 \\ c_3 \\ 0 \end{pmatrix} + s\begin{pmatrix} -\alpha_1 \\ -\alpha_2 \\ 1 \\ 0 \\ 0 \end{pmatrix} + t\begin{pmatrix} -\beta_1 \\ -\beta_2 \\ 0 \\ -\beta_3 \\ 1 \end{pmatrix} : s,t \in \mathbb{K}\right\}.$$

Hier gilt $\operatorname{Kern} A = \operatorname{lin}\{\begin{pmatrix} -\alpha_1 \\ -\alpha_2 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -\beta_1 \\ -\beta_2 \\ 0 \\ -\beta_3 \\ 1 \end{pmatrix}\}$ und $\dim(\operatorname{Kern} A) = 2$. Die Lösungsmenge ist eine Ebene durch den Punkt $\begin{pmatrix} c_1 \\ c_2 \\ 0 \\ c_3 \\ 0 \end{pmatrix}$.

Zum selben Ergebnis gelangt man mit dem (-1)**-Ergänzungstrick:**

> Ausgehend von der Zeilennormalform lasse man zunächst die Nullzeilen weg. Dann ergänze man unter den Zeilen mit “längeren” Stufen eine Zeile mit -1

und sonst Nullen, so dass auf der Diagonalen nur ± 1 steht. Im Beispiel

$$\left(\begin{array}{ccccc|c} 1 & 0 & \alpha_1 & 0 & \beta_1 & c_1 \\ 0 & 1 & \alpha_2 & 0 & \beta_2 & c_2 \\ 0 & 0 & 0 & 1 & \beta_3 & c_3 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right) \quad \left(\begin{array}{ccccc|c} 1 & 0 & \alpha_1 & 0 & \beta_1 & c_1 \\ 0 & 1 & \alpha_2 & 0 & \beta_2 & c_2 \\ 0 & 0 & 0 & 1 & \beta_3 & c_3 \end{array}\right) \quad \left(\begin{array}{ccccc|c} 1 & 0 & \alpha_1 & 0 & \beta_1 & c_1 \\ 0 & 1 & \alpha_2 & 0 & \beta_2 & c_2 \\ 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & \beta_3 & c_3 \\ 0 & 0 & 0 & 0 & -1 & 0 \end{array}\right).$$

Dann nimmt man für jede (-1)-Spalte einen freien Parameter und kann die Lösungsmenge hinschreiben: letzte Spalte plus jeweils freier Parameter mal entsprechender Spalte, also

$$\{\begin{pmatrix} c_1 \\ c_2 \\ 0 \\ c_3 \\ 0 \end{pmatrix} + s\begin{pmatrix} \alpha_1 \\ \alpha_2 \\ -1 \\ 0 \\ 0 \end{pmatrix} + t\begin{pmatrix} \beta_1 \\ \beta_2 \\ 0 \\ \beta_3 \\ -1 \end{pmatrix} : s, t \in \mathbb{K}\}.$$

Man vergleiche dies mit der Darstellung oben.

Bemerkung: Enthält die ZNF links Nullspalten, so ergänze man für diese die (-1)-Zeilen *oben*. Im Beispiel 14.11(2) also

$$\left(\begin{array}{cc|c} 0 & 1 & 1 \\ 0 & 0 & 0 \end{array}\right) \quad (0\,1\,|\,1) \quad \left(\begin{array}{cc|c} -1 & 0 & 0 \\ 0 & 1 & 1 \end{array}\right)$$

mit der abgelesenen Lösungsmenge

$$\{\begin{pmatrix} 0 \\ 1 \end{pmatrix} + s\begin{pmatrix} -1 \\ 0 \end{pmatrix} : s \in \mathbb{K}\}.$$

14.13 Der Rang einer Matrix

Satz und Definition: Für eine Matrix $A \in \mathbb{K}^{n\times m}$ ist die Maximalzahl r linear unabhängiger Zeilen gleich der Maximalzahl linear unabhängiger Spalten (man sagt auch "Zeilenrang = Spaltenrang"). Die Zahl r heißt *Rang der Matrix A*, geschrieben $\operatorname{rg} A$.

Beweis: Man betrachte die ZNF von A.

Bemerkung: Es gilt $\operatorname{rg} A = \dim \operatorname{Bild} A$.

Satz: Sind $A \in \mathbb{K}^{n\times m}$ und $b \in \mathbb{K}^n$, so gilt:

$$Ax = b \text{ ist lösbar} \quad \Leftrightarrow \quad \operatorname{rg} A = \operatorname{rg}(A\,|\,b),$$

dh ein lineares Gleichungssystem ist lösbar genau dann, wenn der Rang der Matrix A gleich dem Rang der erweiterten Matrix $(A\,|\,b)$ ist.

14.14 Dimensionsformel: Sei $A \in \mathbb{K}^{n\times m}$. Dann gilt

$$m = \dim \operatorname{Bild} A + \dim \operatorname{Kern} A = \operatorname{rg} A + \dim \operatorname{Kern} A.$$

Man vergleiche mit 14.12 (insbesondere mit der Form der Matrix, die beim (-1)-Erweiterungstrick entsteht).

Beispiele: (1) $A = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 0 \end{pmatrix}$. Hier ist $n = m = 3$ und $\operatorname{rg} A = 3$, da die Zeilen (oder auch die Spalten) linear unabhängig sind. Somit ist $\dim(\operatorname{Kern} A) = 3 - 3 = 0$ und $\operatorname{Kern} A = \{0\}$.

(2) $A = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$. Hier ist $n = m = 3$ und $\operatorname{rg} A = 2$, da die ersten beiden Spalten linear unabhängig sind, die dritte Spalte jedoch Summe der ersten beiden. Somit ist $\dim(\operatorname{Kern} A) = 3 - 2 = 1$. Durch Probieren findet man $\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} \in \operatorname{Kern} A$. Folglich ist $\operatorname{Kern} A = \operatorname{lin}\{\begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}\}$.

14.15 Lineare Abbildungen

Seien V, W $\mathbb{K}$-Vektorräume.

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

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

Für ein solches ϕ heißt

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

der *Kern von* ϕ und

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

heißt *Bild von* ϕ.

Beispiele: (1) Ist $A \in \mathbb{K}^{n\times m}$, so ist die Abbildung $\phi_A : \mathbb{K}^m \to \mathbb{K}^n$, $x \mapsto Ax$, linear. Es ist $\operatorname{Kern} \phi_A = \operatorname{Kern} A$ und $\operatorname{Bild} \phi_A = \operatorname{Bild} A$.

(2) Die Abbildung $R[a, b] \to \mathbb{R}$, $f \mapsto \int_a^b f(x)\, dx$, ist linear.

(3) Die Abbildung $C^1([a, b]) \to C([a, b])$, $f \mapsto f'$, ist linear.

Satz: Sei $\phi : V \to W$ linear. Dann gilt:

(1) $\operatorname{Kern}\phi$ ein Untervektorraum von V.
(2) $\operatorname{Bild}\phi$ ist ein Untervektorraum von W.
(3) ϕ ist injektiv $\Leftrightarrow \operatorname{Kern}\phi = \{0\}$.
(4) Ist ϕ 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 in W.

Ende Woche 14

Beweis: (4) Seien $\alpha_1, \alpha_2, \dots, \alpha_n \in \mathbb{K}$ mit $\sum_{j=1}^n \alpha_j \phi(v_j) = 0$. Da ϕ linear ist, folgt $\phi(\sum_{j=1}^n \alpha_j v_j) = 0$. Da ϕ injektiv ist, folgt $\sum_{j=1}^n \alpha_j v_j = 0$, und aus linearen Unabhängigkeit von $v_1, v_2, \dots, v_n$ folgt dann $\alpha_1 = \alpha_2 = \dots = \alpha_n = 0$.

Bemerkung: Ist Z ein weiterer $\mathbb{K}$-Vektorraum und sind $\phi : V \to W$ und $\psi : W \to Z$ linear, so ist auch $\psi \circ \phi : V \to Z$ linear.

14.16 Lineare Abbildungen als Matrizen

Seien V, W endlich-dimensionale $\mathbb{K}$-Vektorräume und $\phi : V \to W$ eine lineare Abbildung. In V und W seien Basen $v_1, v_2, \dots, v_m$ bzw. $w_1, w_2, \dots, w_n$ gegeben. Dann gibt es genau eine Matrix $A \in \mathbb{K}^{n\times m}$ mit der Eigenschaft

$$\phi(\sum_{k=1}^m \alpha_k v_k) = \sum_{j=1}^n \beta_j w_j \;\Leftrightarrow\; A \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_m \end{pmatrix} = \begin{pmatrix} \beta_1 \\ \vdots \\ \beta_n \end{pmatrix}.$$

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 $v_1, v_2, \dots, v_m$ auf die Koordinatentupel $(\beta_1, \beta_2, \dots, \beta_n) \in \mathbb{K}^n$ von $\phi(v)$ bzgl. der Basis $w_1, w_2, \dots, w_n$ abgebildet.

A heißt dann *Darstellung von ϕ bzgl. der Basen $v_1, v_2 \dots, v_n$ und $w_1, w_2, \dots, w_n$*.

Man erhält die Matrix A wie folgt:

> Zu jedem Basisvektor v_k, $k = 1, \dots, m$ bestimme man die Koordinaten von $\phi(v_k) \in W$ bzgl. der Basis $w_1, w_2, \dots, w_n$. Diese bilden die k-te Spalte von A.

Beispiele: (1) Die Darstellungsmatrix von ϕ_A aus Beispiel 14.15(1) bzgl. der Standardbasen ist A.

(2) Die lineare Abbildung $\phi : \mathbb{K}^3 \to \mathbb{K}^3$ sei gegeben durch

$$\phi(x_1, x_2, x_3) = (2x_1 - x_2 + x_3, -x_1 + 2x_2 + x_3, x_1 - x_2 + 2x_3).$$

Wir bestimmen die Matrix A zu ϕ bzgl. der Basis v_1, v_2, v_3 im Argument- und Zielraum, wobei

$$v_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}.$$

Es ist

$$\begin{aligned}
\phi(v_1) &= \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = 1 \cdot v_1 + 0 \cdot v_2 + 0 \cdot v_3, \\
\phi(v_2) &= \begin{pmatrix} 3 \\ 0 \\ 3 \end{pmatrix} = 0 \cdot v_1 + 3 \cdot v_2 + 0 \cdot v_3, \\
\phi(v_3) &= \begin{pmatrix} 0 \\ 3 \\ 1 \end{pmatrix} = 1 \cdot v_1 + (-1) \cdot v_2 + 2 \cdot v_3,
\end{aligned}$$

also

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 3 & -1 \\ 0 & 0 & 2 \end{pmatrix}.$$

Verwenden wir hingegen v_1, v_2, v_3 nur "vorne" und die Standardbasis "hinten", so erhalten wir als Darstellungsmatrix

$$\begin{pmatrix} 1 & 3 & 0 \\ 1 & 0 & 3 \\ 0 & 3 & 1 \end{pmatrix}.$$

14.17 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. Nach 14.15 sind die Abbildungen $\phi_A : \mathbb{K}^m \to \mathbb{K}^n$ und $\phi_B : \mathbb{K}^q \to \mathbb{K}^m$ linear, und $\phi_A \circ \phi_B : \mathbb{K}^q \to \mathbb{K}^n$ ist auch linear. Die Darstellungsmatrix von $\phi_A \circ \phi_B$ bzgl. der Standardbasen ist gegeben durch das *Matrixprodukt* $AB \in \mathbb{K}^{n \times q}$, wobei für $A = (a_{jk})_{j=1}^{n}{}_{k=1}^{m}$ und $B = (b_{kl})_{k=1}^{m}{}_{l=1}^{q}$ die Matrix AB Einträge (c_{jl}) hat mit

$$c_{jl} = \sum_{k=1}^{n} a_{jk} b_{kl}, \qquad j = 1, \ldots, n, \; l = 1, \ldots, q.$$

Für $x = (x_1, x_2, \ldots, x_q) \in \mathbb{K}^q$ ist nämlich

$$\begin{aligned}
A(Bx) &= (a_{jk})_{j,k} \Big(\sum_{l=1}^{q} b_{kl} x_l \Big)_k = \Big(\sum_{k=1}^{m} a_{jk} \sum_{l=1}^{q} b_{kl} x_l \Big)_j \\
&= \Big(\sum_{l=1}^{q} (\sum_{k=1}^{m} a_{jk} b_{kl}) x_l \Big)_j = \Big(\sum_{k=1}^{m} a_{jk} b_{kl} \Big)_{j,l} (x_l)_l.
\end{aligned}$$

Man beachte, dass hierbei die Anzahl der Spalten von A gleich der Anzahl der Zeilen von B ist (nämlich m), andernfalls ist das Matrixprodukt **nicht definiert**.

Bemerkung: Man erhält die l-te Spalte von AB, indem man die Matrix A mit der l-ten Spalte von B multipliziert (im Sinne des Matrix-Vektor-Produktes aus 14.9).

Beispiel mit $n = m = 2$:
$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Eigenschaften des Produktes von Matrizen

Für alle $A_1, A_2 \in \mathbb{K}^{n\times m}$, $B_1, B_2 \in \mathbb{K}^{m\times q}$ und alle $\alpha \in \mathbb{K}$ gilt:

$$(A_1 + \alpha A_2)B_1 = A_1B_1 + \alpha A_2B_1, \qquad A_1(B_1 + \alpha B_2) = A_1B_1 + \alpha A_1B_2,$$

dh für festes $B \in \mathbb{K}^{m\times q}$ ist $\mathbb{K}^{n\times m} \to \mathbb{K}^{n\times q}$, $A \mapsto AB$ linear, und für festes $A \in \mathbb{K}^{n\times m}$ ist $\mathbb{K}^{m\times q} \to \mathbb{K}^{n\times q}$, $B \mapsto AB$ linear.

Außerdem ist das Produkt von Matrizen assoziativ.

Warnung: Im allgemeinen gilt $AB \neq BA$! Damit beide Produkte existieren, muss zunächst $n = q$ sein. Dann ist AB eine $n \times n$-Matrix und BA eine $m \times m$-Matrix. Gleichheit kann also höchstens für $n = m = q$ gelten. Für $n = m = q = 2$ ist aber z.B.

$$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 0 \end{pmatrix} \neq \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}.$$

14.18 Invertierbare Matrizen

Definition: Für $n \in \mathbb{N}$ heißt $I_n := (\delta_{jk}) \in \mathbb{K}^{n\times n}$ die *Einheitsmatrix* (wir schreiben auch kurz I, wenn n aus dem Zusammenhang klar ist).

Es gilt dann
$$IA = AI = A \quad \text{für alle } A \in \mathbb{K}^{n\times n}$$
und die zugehörige Abbildung $\phi_I : \mathbb{K}^n \to \mathbb{K}^n$ ist die Identität.

Definition: Eine Matrix $A \in \mathbb{K}^{n\times n}$ heißt *regulär*, falls $\operatorname{rg} A = n$ 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:

$$B = BI = B(A\tilde{B}) = (BA)\tilde{B} = I\tilde{B} = \tilde{B}. \tag{1}$$

Diese Matrix B heißt *Inverse von A* (oder *zu A inverse Matrix*) und wird mit A^{-1} bezeichnet.

Rechenregeln: Seien $A, B \in \mathbb{K}^{n\times n}$ invertierbar. Dann sind A^{-1} und AB invertierbar, und es gilt:
$$(A^{-1})^{-1} = A \qquad \text{und} \qquad (AB)^{-1} = B^{-1}A^{-1}.$$

Die zweite Regel ist die *Hemd-Jacken-Regel* (vgl. 3.6).

Satz: Sei $A \in \mathbb{K}^{n\times n}$. Dann gilt:

$$A \text{ regulär} \;\Leftrightarrow\; \operatorname{Bild} A = \mathbb{K}^n \;\Leftrightarrow\; \operatorname{Kern} A = \{0\} \;\Leftrightarrow\; A \text{ invertierbar}.$$

Für die induzierte lineare Abbildung $\phi_A : \mathbb{K}^n \to \mathbb{K}^n$ gilt:

$$\phi_A \text{ surjektiv} \;\Leftrightarrow\; \phi_A \text{ injektiv} \;\Leftrightarrow\; \phi_A \text{ bijektiv}.$$

Bemerkung: Sind $A, B \in \mathbb{K}^{n\times n}$ mit $AB = I$, so sind A und B invertierbar und es gilt $A^{-1} = B$. So ist etwa $\left(\begin{smallmatrix} 1 & 1 \\ 0 & 1 \end{smallmatrix}\right)^{-1} = \left(\begin{smallmatrix} 1 & -1 \\ 0 & 1 \end{smallmatrix}\right)$ (vgl. Beispiel in 14.17).

Beweis: Gilt $AB = I$, so ist $\operatorname{Kern} B = \{0\}$ und $\operatorname{Bild} A = \mathbb{K}^n$, und nach dem Satz sind A, B invertierbar. Für $A^{-1} = B$ verwenden wir ein Argument wie in (1).

Berechnung von A^{-1}: Man löse für jedes $l \in \{1, 2, \dots, n\}$ das Gleichungssytem $Ax = e_l$. Die Lösung ist die l-te Spalte von A^{-1}. Man kann diese Gleichungssyteme auch simultan lösen.

Beispiel: Sei $A = \left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right) \in \mathbb{K}^{2\times 2}$ invertierbar und $a \neq 0$. Wir setzen $\delta := ad - bc$ und sehen an der dritten Matrix, dass $\delta \neq 0$ ist.

$$\left(\begin{array}{cc|cc} a & b & 1 & 0 \\ c & d & 0 & 1 \end{array}\right) \quad \left(\begin{array}{cc|cc} 1 & b/a & 1/a & 0 \\ 0 & d - bc/a & -c/a & 1 \end{array}\right) \quad \left(\begin{array}{cc|cc} 1 & b/a & 1/a & 0 \\ 0 & \delta/a & -c/a & 1 \end{array}\right)$$
$$\left(\begin{array}{cc|cc} 1 & b/a & 1/a & 0 \\ 0 & 1 & -c/\delta & a/\delta \end{array}\right) \quad \left(\begin{array}{cc|cc} 1 & 0 & \frac{1}{a} + \frac{bc}{\delta a} & -b/\delta \\ 0 & 1 & -c/\delta & a/\delta \end{array}\right).$$

Wegen $\frac{1}{a} + \frac{bc}{\delta a} = d/\delta$ ist also

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{\delta}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Den Fall $c \neq 0$ behandelt man analog, er führt auf dieselbe Formel. Wir sehen auch, dass Invertierbarkeit äquivalent ist zu $\delta \neq 0$.

14.19 Skalarprodukt und Norm in $\mathbb{C}^n$

Definition: Wir definieren das *Skalarprodukt* in $\mathbb{C}^n$ durch

$$(x|y) := \sum_{j=1}^{n} x_j \overline{y_j} \quad \text{für } x = (x_1, x_2, \dots, x_n), y = (y_1, y_2, \dots, y_n) \in \mathbb{C}^n.$$

Die so definierte Abbildung $(\cdot|\cdot) : \mathbb{C}^n \times \mathbb{C}^n \to \mathbb{C}$ hat folgende Eigenschaften

(S1) $\forall x, y \in \mathbb{C}^n$: $(x|y) = \overline{(y|x)}$,
(S2) $\forall x, y, z \in \mathbb{C}^n$, $\alpha \in \mathbb{C}$: $(\alpha x + y|z) = \alpha(x|z) + (y|z)$,
(S3) $\forall x \in \mathbb{C}^n \setminus \{0\}$: $(x|x) > 0$.

Bemerkung: Ist V ein $\mathbb{C}$-Vektorraum, so heißt eine Abbildung $(\cdot|\cdot) : V \times V \to \mathbb{C}$ mit den Eigenschaften (S1)–(S3) (für V satt $\mathbb{C}^n$) ein *Skalarprodukt auf* V.

Eigenschaften eines Skalarproduktes auf einem $\mathbb{C}$-Vektorraum V sind

> $\forall x, y \in V, \alpha \in \mathbb{C}$: $(x|\alpha y + z) = \overline{\alpha}(x|y) + (x|z)$,
> $\forall x \in V$: $(x|0) = (0|x) = 0$,
> Cauchy-Schwarz-Ungleichung: $\forall x, y \in V : |(x|y)| \leq \sqrt{(x|x)}\sqrt{(y|y)}$, wobei Gleichheit nur gilt, wenn x, y linear abhängig sind.

Eine weitere wichtige Eigenschaft beschreibt der folgende Satz.

Satz: Ist V ein $\mathbb{C}$-Vektorraum und $(\cdot|\cdot) : V \times V \to \mathbb{C}$ ein Skalarprodukt, so hat die Abbildung $\|\cdot\| : V \to [0, \infty)$, $v \mapsto \sqrt{(v|v)}$ folgende Eigenschaften:

> (N1) $\forall v \in V$: $\|v\| = 0 \Rightarrow v = 0$,
> (N2) $\forall v \in V$, $\alpha \in \mathbb{C}$: $\|\alpha v\| = |\alpha|\|v\|$,
> (N3) $\forall u, v \in V$: $\|u + v\| \leq \|u\| + \|v\|$ (Dreiecks-Ungleichung)

Bemerkung: Eine Abbildung $\|\cdot\| : V \to [0, \infty)$ mit den Eigenschaften (N1)–(N3) heißt eine *Norm auf* V. Ist $\|\cdot\|$ eine Norm auf V, so wird für $u, v \in V$ die Zahl $\|u - v\| \geq 0$ als **Abstand** von u und v interpretiert, und es gilt $\|0\| = 0$, sowie

$$|\|u\| - \|v\|| \leq \|u - v\| \quad \text{(umgekehrte Dreiecksungleichung).}$$

Bemerkung: Die Aussagen über das Skalaprodukt und die Norm gelten auch in $\mathbb{R}^n$ und allgemeinen $\mathbb{R}$-Vektorraumen V. Auf die komplexe Konjugation kann dann verzichtet werden.

Anschauliche Bedeutung des Skalarproduktes im $\mathbb{R}^n$: Sind $u, v \in \mathbb{R}^n$, so ist

$$(u|v) = \|u\|\|v\| \cos\varphi,$$

wobei φ den von u, v eingeschlossenen **Winkel** bezeichne.

Andere gängige Bezeichnungen: Für das Skalarprodukt von $x, y \in \mathbb{R}^n$ schreibt man häufig auch $x \cdot y$ (und gelegentlich sogar nur xy).

14.20 Orthogonalität

Definition: Die Vektoren $v_1, v_2, \ldots, v_m \in \mathbb{K}^n$ heißen *orthogonal*, falls für alle $j, k \in \{1, 2 \ldots, m\}$ mit $j \neq k$ gilt $(v_j|v_k) = 0$.

Die Vektoren $v_1, v_2, \ldots, v_m$ heißen *orthonormal* oder ein *Orthonormalsystem* (*ONS*), falls sie orthogonal sind und zusätzlich $(v_j|v_j) = 1$ für alle $j = 1, \ldots, n$ gilt (bzw. falls für alle j, k gilt: $(v_j|v_k) = \delta_{jk}$).

Eine Basis in $\mathbb{K}^n$ heißt *Orthonormalbasis* (*ONB*), falls sie aus orthonormalen Vektoren besteht.

Beispiel: Die Standardbasis $e_1, e_2, \ldots, e_n$ ist eine Orthonormalbasis von $\mathbb{C}^n$, denn es gilt $(e_j|e_k) = \delta_{jk}$ für alle $j, k = 1, \ldots, n$.

Satz: Sei V ein $\mathbb{K}$-Vektorraum mit Skalarprodukt und seien $v_1, v_2, \ldots, v_n \in V \setminus \{0\}$ orthogonal. Dann sind $v_1, v_2, \ldots, v_n$ linear unabhängig.

Beweis: Seien $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{K}$ mit

$$\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n = 0.$$

Sei nun $k \in \{1, 2, \ldots, n\}$. Wir nehmen das Skalarprodukt der Gleichung mit v_k und erhalten

$$0 = (0|v_k) = (\alpha_1 v_1 + \alpha_2 v_2 + \ldots + \alpha_n v_n | v_k) = \alpha_k (v_k|v_k)$$

wegen der vorausgesetzten Orthogonalität. Wegen $v_k \neq 0$ ist auch $(v_k|v_k) \neq 0$, und wir erhalten $\alpha_k = 0$. Da k beliebig war, sind $v_1, v_2, \ldots, v_n$ linear unabhängig.

14.21 Das Gram-Schmidt-Verfahren

Gegeben sei ein $\mathbb{C}$-Vektorraum mit Skalarprodukt und $v_1, v_2, \ldots, v_m \in V$ seien linear unabhängig. Wir wollen ein Orthonormalsystem $b_1, b_2, \ldots, b_m$ finden mit

$$\operatorname{lin}\{v_1, v_2, \ldots, v_m\} = \operatorname{lin}\{b_1, b_2, \ldots, b_m\}.$$

Falls $v_1, v_2, \ldots, v_m$ eine Basis von V ist, so ist $b_1, b_2, \ldots, b_m$ eine Orthonormalbasis von V.

Die Vektoren $b_1, \ldots, b_m$ werden sukzessiv so konstruiert, dass gilt:

$$\operatorname{lin}\{v_1, v_2, \ldots, v_k\} = \operatorname{lin}\{b_1, b_2, \ldots, b_k\} \quad \text{für alle } k = 1, \ldots, m.$$

Wir setzen $b_1 := v_1/\|v_1\|$ und für $k = 2, \ldots, m$:

$$\tilde{b}_k := v_k - \sum_{j=1}^{k-1} (v_k|b_j) b_j, \quad b_k := \frac{\tilde{b}_k}{\|\tilde{b}_k\|},$$

bzw.

$$\tilde{b}_k := v_k - \sum_{j=1}^{k-1} \frac{(v_k|\tilde{b}_j)}{(\tilde{b}_j|\tilde{b}_j)} \tilde{b}_j, \quad b_k := \frac{\tilde{b}_k}{\|\tilde{b}_k\|}.$$

Beispiele: (1) Wir betrachten $n = 3$, $V = \mathbb{C}^3$ und

$$v_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \quad v_2 = \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}.$$

Dann ist

$$\begin{aligned}
b_1 &= \frac{v_1}{\|v_1\|} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \\
\tilde{b}_2 &= \begin{pmatrix} 2 \\ 1 \\ 0 \end{pmatrix} - \frac{1}{2} \cdot 2 \cdot \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \\
b_2 &= \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \\
\tilde{b}_3 &= \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} - \frac{1}{2} \cdot (-1) \cdot \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} - \frac{1}{3} \cdot 0 \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1/2 \\ -1 \\ 1/2 \end{pmatrix} \\
b_3 &= \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ -2 \\ 1 \end{pmatrix}.
\end{aligned}$$

(2) v_1, v_2 wie eben, aber $v_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$. Dann sind b_1, b_2 wie eben, aber

$$\begin{aligned}
\tilde{b}_3 &= \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} - \frac{1}{2} \cdot (-1) \cdot \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} - \frac{1}{3} \cdot 2 \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -1/6 \\ 1/3 \\ -1/6 \end{pmatrix} \\
b_3 &= \frac{1}{\sqrt{6}} \begin{pmatrix} -1 \\ 2 \\ -1 \end{pmatrix}.
\end{aligned}$$

Man beachte, dass sich dieses b_3 von dem in Beispiel (1) nur durch das Vorzeichen unterscheidet. Dies ist nicht erstaunlich, da es genau zwei Möglichkeiten gibt, b_1, b_2 zu einer Orthonormalbasis zu ergänzen.

14.22 Transponierte und adjungierte Matrizen

Definition: Für eine Matrix $A = (a_{jk}) \in \mathbb{K}^{n \times m}$ heißt die Matrix $\in \mathbb{K}^{m \times n}$, die durch Vertauschen von Zeilen und Spalten entsteht, die *transponierte Matrix zu A* und wird mit A^T bezeichnet. Für alle $k \in \{1, \ldots, m\}$, $j \in \{1, \ldots, n\}$ steht an der Stelle (k, j) in der Matrix A^T also der Eintrag a_{jk}, der in der Matrix A an der Stelle (j, k) steht. Setzen wir $B := A^T$ mit $B = (b_{kj})_{k=1}^{m}{}_{j=1}^{n} \in \mathbb{K}^{m \times n}$, so gilt also

$$b_{kj} = a_{jk}, \quad k \in \{1, \ldots, m\}, j \in \{1, \ldots, n\}.$$

Im Falle $\mathbb{K} = \mathbb{C}$ heißt die Matrix $\in \mathbb{C}^{m\times n}$, für die an jeder Stelle (k, j) der Eintrag $\overline{a_{jk}}$ steht, die *adjungierte Matrix zu A* und wird mit A^* bezeichnet.

Bemerkung: Setzt man $\overline{A} := (\overline{a_{jk}}) \in \mathbb{C}^{n\times m}$ (*konjugiert komplexe Matrix zu A*), so gilt

$$A^* = \overline{A^T} = \left(\overline{A}\right)^T.$$

Schreibweisen des Skalarprodukts:

Für $x, y \in \mathbb{R}^n = \mathbb{R}^{n\times 1}$ gilt $(x|y) = y^T x = x^T y$.
Für $x, y \in \mathbb{C}^n = \mathbb{C}^{n\times 1}$ gilt $(x|y) = y^* x = \overline{x^* y}$.

Ende Woche 15

Rechenregeln: Für Matrizen A, B, deren Produkt erklärt ist, gilt:

$$(AB)^T = B^T A^T \quad \text{und} \quad (AB)^* = B^* A^*.$$

Folgerung: Sei $A \in \mathbb{K}^{n\times m}$. Dann gilt:

(a) Im Fall $\mathbb{K} = \mathbb{R}$ ist $(Ax|y) = (x|A^T y)$ für alle $x \in \mathbb{R}^m$, $y \in \mathbb{R}^n$.

(b) Im Fall $\mathbb{K} = \mathbb{C}$ ist $(Ax|y) = (x|A^* y)$ für alle $x \in \mathbb{C}^m$, $y \in \mathbb{C}^n$.

Beweis: Man schreibe das Skalarprodukt wie oben angegeben und benutze die Rechenregeln.