Höhere Mathematik I für Physiker

Karlsruher Institut für Technologie

Prof. Dr. Lamm

Wintersemester 2012/13

Inhaltsverzeichnis

1 Grundlagen

1.1 Logik

Wir behandeln mathematische Aussagen

Beispiel:
i) $4 > 2$ (wahr)
ii) $\forall n \in \mathbb{N} : n > 4 \Rightarrow \ n > 2$ (wahr)
iii) $5 < 3$ (falsch)

Satz vom ausgeschlossenen Dritten: Eine zulässige mathematische Aussage ist entweder wahr oder falsch, jedoch nie beides zugleich.

Bemerkung: nicht zulässig: „diese Aussage ist falsch“
$\rightarrow$ Axiome der Logik sind unvollständig

Logische Verknüpfungen: A, B mathematische Aussagen

$\neg A$	„Negation“
$A \wedge B$	„und“
$A \vee B$	„oder“
$A \rightarrow B$	„Implikation“
$A \leftrightarrow B$	„Äquivalenz“

sind definiert durch die Wahrheitstafel

A	B	$\neg A$	$A \wedge B$	$A \vee B$	$A \rightarrow B$	$A \leftrightarrow B$
w	w	f	w	w	w	w
w	f	f	f	w	f	f
f	w	w	f	w	w	f
f	f	w	f	f	w	w

Beispiel: $(n > 4) \rightarrow (n > 2)$

Bemerkung:
i) ist $A \rightarrow B$ wahr, so nennen wir dies Folgerung und schreiben $A \Rightarrow B$
ii) es gilt: $(A \rightarrow B) \wedge (B \rightarrow C) \Rightarrow (A \rightarrow C)$
Kette von Folgerungen: $A \Rightarrow B \Rightarrow \ldots \Rightarrow S$ (A: Annahme, S: Satz)
(Prinzip des direkten Beweises)

iii) ist die Äquivalenz $A \leftrightarrow B$ wahr, so schreiben wir $A \Leftrightarrow B$

Satz 1.1.1. *Es gilt:*

$$(A \to B) \Leftrightarrow (\neg B \to \neg A)$$

Beweis: Durch Wahrheitstafel:

A	B	$A \to B$	$\neg A$	$\neg B$	$\neg B \to \neg A)$
w	w	w	f	f	w
w	f	f	f	w	f
f	w	w	w	f	w
f	f	w	w	w	w

Vergleich der dritten und sechsten Spalte liefert die Behauptung. □

Damit folgt: Umkehrschluss
$(A \Rightarrow B) \Leftrightarrow (\neg B \Rightarrow \neg A)$

Indirekter Beweis
Um die Folgerung $A \Rightarrow B$ zu zeigen genügt es $\neg B \Rightarrow \neg A$ zu zeigen!

BEISPIEL:
A seien die Axiome in $\mathbb{N}$, d.h. $1 \in \mathbb{N}$ und $n \in \mathbb{N} \Rightarrow n+1 \in \mathbb{N}$
B: es gibt keine grösste natürliche Zahl

Behauptung: $A \Rightarrow B$

Beweis: Annahme: es existiert ein maximales $n_0 \in \mathbb{N}$ d.h. $n_0 \geq l \qquad \forall l \in \mathbb{N}$
aber $n_0 \in \mathbb{N} \Rightarrow n_0 < n_0 + 1 \in \mathbb{N}$, also $n_0 \notin \mathbb{N}$.

□

Satz 1.1.2. *Es gelten:*

i) $\neg(A \vee B) \Leftrightarrow (\neg A) \wedge (\neg B)$
ii) $\neg(A \wedge B) \Leftrightarrow (\neg A) \vee (\neg B)$

Beweis: i)

A	B	$A \vee B$	$\neg(A \vee B)$	$\neg A$	$\neg B$	$(\neg A) \wedge (\neg B)$
w	w	w	f	f	f	f
w	f	w	f	f	w	f
f	w	w	f	w	f	f
f	f	f	w	w	w	w

ii) folgt analog. □

1.2 Mengenlehre

Cantor: “Eine Menge ist die ungeordnete Zusammenfassung verschiedener Elemente zu einem Ganzen.“

Beispiele:

i)	$\mathbb{N} = \{1, 2, 3, \ldots\}$	$\mathbb{N}_0 = \{0, 1, 2, 3, \ldots\} = \{0\} \cup \mathbb{N}$
	$\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$	$\mathbb{R}_0^+ := \{x \in \mathbb{R} : x \geq 0\}$
		$\mathbb{N} = \mathbb{Z} \cap \mathbb{R}^+$
ii)	$\{\} = \emptyset$ leere Menge	
iii)	$\{n \in \mathbb{N} : n \text{ Primzahlen}\} = \{2, 3, 5, 7, \ldots\}$	

Achtung:
Die Menge M aller Mengen, die sich selbst als Element nicht enthält, existiert nicht:

$$\text{wäre} \qquad \begin{aligned} M \in M &\Rightarrow M \notin M \\ M \notin M &\Rightarrow M \in M \end{aligned}$$

Verknüpfungen von Mengen:

$X \cup Y := \{x : x \in X \vee x \in Y\}$	Vereinigung
$X \cap Y := \{x : x \in X \wedge x \in Y\}$	Durchschnitt
$X \backslash Y := \{x : x \in X \wedge x \notin Y\}$	Differenz
$X \subset Y :$	Teilmenge

Beispiel: $X = (X \backslash Y) \cup (X \cap Y)$

Bemerkung:

i)	$X = Y \Leftrightarrow (X \subset Y) \wedge (Y \subset X)$	
ii)	$X \cap (Y \cup Z) = (X \cap Y) \cup (X \cap Z)$	
iii)	$\mathcal{P}(X) := \{Y : Y \subset X\}$	Potenzmenge einer Menge X
iv)	$X \times Y := \{(x, y) : x \in X, y \in Y\}$	Produkt von Mengen X, Y

Quantoren

Für alle Elemente in Mengen benutzen wir die Quantoren:
$\forall$: für alle
$\exists$: es existiert
$\exists!$: es existiert genau ein

Beispiel:

i)	$\forall n \in \mathbb{N} : n > 0$	wahr
ii)	$\exists n_0 \in \mathbb{N}\ \forall k \in \mathbb{N} : k \leq n_0$	falsch
iii)	$\forall n_0 \in \mathbb{N}\ \exists k \in \mathbb{N} : k > 0$	wahr

Bemerkung:

$$\begin{aligned}
&\forall x \in M : A(x) \Leftrightarrow \{x \in M : A(x)\} = M \\
&\exists x \in M : A(x) \Leftrightarrow \{x \in M : A(x)\} \neq \emptyset
\end{aligned}$$

Satz 1.2.1. *Es gelten:*
i) $\neg(\forall x \in M : A(x)) \Leftrightarrow (\exists x \in M : \neg A(x))$
ii) $\neg(\exists x \in M : A(x)) \Leftrightarrow (\forall x \in M : \neg A(x))$

Beweis: i)

$M = \{x \in M : A(x)\} \cup \{x \in M : \neg A(x)\}$ Satz von ausgeschlossenen Dritten

$$\begin{aligned}
\Rightarrow \ \neg(\forall x \in M : A(x)) \quad &\Leftrightarrow \{x \in M : A(x)\} \neq M \\
&\Leftrightarrow \{x \in M : \neg A(x)\} \neq \emptyset \\
&\Leftrightarrow \exists x \in M : \neg A(x)
\end{aligned}$$

ii) folgt wieder analog. □

1.3 Funktionen

Aus der Schule bekannt: $y = f(x) = x - x^3, \ -1 \leq x \leq 1$

Definition 1.3.1. *Eine Funktion (Abbildung) $f : X \to Y$ (X, Y Mengen) ordnet jedem Element $x \in X$ genau ein „Bild" $y = f(x) \in Y$ zu. Jedes Element $z \in X$ mit $y = f(z)$ heisst „Urbild" von y.*

Bemerkung: Eine Funktion besteht also aus
i) dem Definitionsbereich (hier X)
ii) dem Bildbereich (hier Y)
iii) der Abbildungsvorschrift (hier $x \mapsto f(x)$)

Beispiel:

$$\begin{array}{llll}
\text{i)} & a) & f : \mathbb{R} \to \mathbb{R}, & x \mapsto x - x^3 \\
 & b) & f : [-1, 1] \to \mathbb{R}, & x \mapsto x - x^3 \\
\text{ii)} & & g : \mathbb{R} \to [-1, 1], & x \mapsto \cos(x) \\
\text{ii)} & & h : \mathbb{R} \to [0, \infty], & x \mapsto x^2
\end{array}$$

i) $a), b)$ sind verschiedene Funktionen, da die Definitionsbereiche verschieden sind.

Funktionen kann man durch ihren Graphen darstellen.

$G(f) := \{(x, f(x)) : x \in X\} \subset X \times Y$

Komposition von Abbildungen

$f : X \to Y,\ g : Y \to Z$

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

BEISPIEL:

$$\begin{aligned} h:&\ \mathbb{R} \to \mathbb{R} \quad x \mapsto e^{x^2} \\ f:&\ \mathbb{R} \to \mathbb{R} \quad x \mapsto x^2 \\ g:&\ \mathbb{R} \to \mathbb{R} \quad x \mapsto e^x \end{aligned}$$

Damit folgt $h = g \circ f$.

Satz 1.3.1. *Für Abbildungen $f : X \to Y,\ g : Y \to Z,\ h : Z \to W$ gilt:*

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

Beweis: Um zu zeigen das zwei Funktionen identisch sind, muss man zeigen das der Definitionsbereich, der Bildbereich und die Abbildungsvorschrift identisch sind.

Definitionsbereich und Bildbereich sind offensichtlich identisch. Weiter gilt für alle $x \in X$:

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

□

Definition 1.3.2. *Sei $f : X \to Y$ eine Abbildung.*

i) *f heisst surjektiv, falls $\forall y \in Y\ \exists x \in X$ mit $f(x) = y$*
ii) *f heisst injektiv, falls $\forall x_1, x_2 \in X :\ f(x_1) = f(x_2) \Rightarrow x_1 = x_2$*
oder: $\forall x_1, x_2 \in X : x_1 \neq x_2 \Rightarrow f(x_1) \neq f(x_2)$
iii) *f heisst bijektiv, falls f injektiv und surjektiv ist.*

BEISPIELE:

$f : \mathbb{R} \to [\,0, \infty)$,	$x \mapsto x^2$	surjektiv, nicht injektiv, z.B. $f(-1) = f(1) = 1$
$g : \mathbb{R} \to \mathbb{R}$,	$x \mapsto x^2$	nicht surjektiv, es ex. kein $x \in \mathbb{R}$ mit $g(x) = -1$
$h : [\,0, \infty) \to [\,0, \infty)$,	$x \mapsto x^2$	bijektiv
$i : \mathbb{R}\setminus\{0\} \to \mathbb{R}$,	$x \mapsto x^{-1}$	injektiv aber nicht surjektiv
$j : \mathbb{R}\setminus\{0\} \to \mathbb{R}\setminus\{0\}$,	$x \mapsto x^{-1}$	bijektiv
$k : \mathbb{R} \to \mathbb{R}$,	$x \mapsto x - x^3$	surjektiv aber nicht injektiv
$l : (-\frac{\pi}{2}, \frac{\pi}{2}) \to (-1, 1)$,	$x \mapsto \sin(x)$	bijektiv

BEMERKUNG: Ist f bijektiv, so können wir eine Abb $g : Y \to X$ definieren, welche jedem $y \in Y$ das eindeutig bestimmte Urbild $x \in X$ unter f zuordnet. Es gilt:

$$\begin{aligned} g \circ f &= id_x \qquad (id_x : X \to X, x \mapsto x) \\ f \circ g &= id_y \end{aligned}$$

g heisst Umkehrabbildung von f. Notation: $g = f^{-1}$

Weitere Erklärung: Sei $y \in Y$, f surjektiv $\Rightarrow \exists x \in X : f(x) = y$
Wir möchten definieren: $g(y) = x$ (Annahme: $\exists x_1, x_2 \in X : f(x_1) = f(x_2) = y$)
f surjektiv + injektiv $\Rightarrow \exists! x \in X : f(x) = y \Rightarrow g(y) = x$ wohldefiniert
$\Rightarrow g \circ f = id_x$ denn $g(f(x)) = x$

Satz 1.3.2. *Sei $f : X \to Y$ eine Abbildung. Dann gilt:*

i)	*f injektiv*	$\Rightarrow \exists g : Y \to X$ *mit* $g \circ f = id_x$
ii)	*f surjektiv*	$\Rightarrow \exists g : Y \to X$ *mit* $f \circ g = id_y$
iii)	*f bijektiv*	$\Rightarrow \exists g : Y \to X$ *mit* $g \circ f = id_x$ *und* $f \circ g = id_y$.

Beweis: i) $\forall y \in f(X) := \{f(x) : x \in X\} \subset Y \qquad \exists!$ Urbild $x \in X$ von y.

Definiere in diesem Fall $g(y) := x$

Sei weiter $x_0 \in X$ beliebig und setze $g(y) = x_0 \ \forall y \notin f(x)$

Damit ist $g : Y \to X$ wohldefiniert und es gilt: $g \circ f = id_x$. □

Urbild-Funktion
$f : X \to Y$ Abbildung, $A \subset X,\ B \subset Y$
$f(A) := \{f(x) : x \in A\} \subset Y$ heisst Bild von A

Definiere die Urbildfunktion $f^{-1} : \mathcal{P}(Y) \to \mathcal{P}(X)$ durch

$$f^{-1}(B) := \{x \in X : f(x) \in B\} \subset X.$$

Bemerkung: Für diese Defintion muss f nicht bijektiv sein.

Beispiel:

$$f : [-1,1] \to \mathbb{R},\ f(x) = x - x^3,\ B = \{0\} \qquad \Rightarrow f^{-1}(B) = \{-1, 0, 1\}$$
$$\tilde{B} = \left\{0, f\left(\frac{1}{\sqrt{3}}\right)\right\} \Rightarrow f^{-1}(\tilde{B}) = \left\{-1, 0, 1, \frac{1}{\sqrt{3}}\right\}$$

Satz 1.3.3. *Eine Abbildung $f : X \to Y$ ist genau dann bijektiv, wenn $f^{-1}(\{y\})$ für alle $y \in Y$ aus genau einem Element besteht.*

1.4 Äquivalenzrelationen

Sei X eine beliebige Menge.

Definition 1.4.1 Eine Beziehung „$\sim$“ auf X heisst Äquivalenzrelationen, falls gilt:

i)	Reflexivität:	$\forall x \in X:$	$x \sim x$
ii)	Symmetrie:	$\forall x, y \in X:$	$x \sim y \Rightarrow y \sim x$
iii)	Transitivität:	$\forall x, y, z \in X:$	$x \sim y,\ y \sim z \Rightarrow x \sim z$

Beispiel:

i) “=“ auf beliebigen Mengen
ii) Äquivalenz-Aussagen (Übung)
iii) Reste Modulo p:
Sei $p \in \mathbb{N}$ fest für $m, n \in \mathbb{Z}$ definiere $m \sim n$, falls $m = n + kp$ für ein $k \in \mathbb{Z}$

1)	$m \sim m$	wähle $k = 0$
2)	$m \sim n$	$\Rightarrow \exists k \in \mathbb{Z} : m = n + kp \Rightarrow \exists k \in \mathbb{Z} : n = m - kp$
		$\Rightarrow \exists k \in \mathbb{Z} : n = m + (-k)p$
		$\Rightarrow n \sim m$
3)	$m \sim n, n \sim l$	$\Rightarrow \exists k_1, k_2 \in \mathbb{Z} : m = n + k_1 p,\ n = l + k_2 p$
		$\Rightarrow m = l + (k_1 + k_2)p \ \Rightarrow m \sim l$

Sei „$\sim$“ Äquivalenzrelation auf X und sei $x \in X$. Die Menge $[x] := \{y \in X : x \sim y\}$ heisst Äquivalenzklasse von X. Es gilt:

1) $\forall y \in [x] : [y] = [x]$

Beweis: „$\subset$“ (es soll gezeigt werden, dass $[y] \subset [x]$)

Sei $z \in [y]$, also $y \sim z$. Da $x \sim y$ folgt aus der Transitivität von $\sim$ $x \sim z \ \Rightarrow \ z \in [x]$.

„$\supset$“ (es soll gezeigt werden, dass $[x] \subset [y]$)

$y \in [x] \Rightarrow x \sim y \Rightarrow y \sim x \Rightarrow x \in [y]$. Wie oben folgt damit $[x] \subset [y]$. □

2) $\forall y \in X : y \notin [x] \Rightarrow [y] \cap [x] = \emptyset$

Beweis: (indirekt)

Sei $y \notin [x]$ und $z \in [x] \cap [y]$

Aus 1) folgt: $[x] = [z] = [y] \ni y$ □

Satz 1.4.1. *Eine Äquivalenzrelation auf X definiert eine disjunkte Zerlegung von X in Äquivalenzklassen.*

2 Axiome der Reellen Zahlen und vollständige Induktion

2.1 Axiome der Reellen Zahlen

Gegeben seien die Strukturen

i)	$+$, welche $a, b \in \mathbb{R}$	$a + b \in \mathbb{R}$ zuordnet
ii)	$\cdot$, welche $a, b \in \mathbb{R}$	$a \cdot b \in \mathbb{R}$ zuordnet
iii)	eine Relation $a > b$ die für $a, b \in \mathbb{R}$ zutrifft oder nicht	

Körperaxiome

Assoziativgesetz:	$(a + b) + c = a + (b + c)$	$(a \cdot b) \cdot c = a \cdot (b \cdot c)$
Kommutativgesetz:	$a + b = b + a$	$a \cdot b = b \cdot a$
Distributivgesetz:	$a \cdot (b + c) = a \cdot b + a \cdot c$	

neutrales Element: es gibt $0 \in \mathbb{R},\ 1 \in \mathbb{R}$ mit $1 \neq 0$, so dass
0 : neutrales Element der Addition: $\forall a \in \mathbb{R}$ gilt: $a + 0 = a$
1 : neutrales Element der Multiplikation: $\forall a \in \mathbb{R}$ gilt: $a \cdot 1 = a$

Inverses Element: $\forall a \in \mathbb{R}\ \exists$ Lösungen $x, y \in \mathbb{R}$ von $a + x = 0,\ a \cdot y = 0\ \forall a \neq 0$

Bemerkung:
1) Eine Menge $\mathbb{K}$ mit Verknüpfungen $+, \cdot$ welche obige Axiome erfüllt, heisst Körper. Die Menge der reellen und rationalen Zahlen erfüllen diese Axiome.

2) Die neutralen Elemente 0 und 1 sind eindeutig bestimmt. Seien z.B. $0_1, 0_2$ neutrale Elemente bzgl. der Addition
$\Rightarrow 0_1 = 0_1 + 0_2 = 0_2 + 0_1 = 0_2$

3) Inverse Elemente sind auch eindeutig bestimmt. Seien dazu x_1, x_2 Lösungen von $a + x = 0$. Damit folgt aus den Körperaxiomen

$$x_1 = x_1 + 0 = x_1 + (a + x_2) = (x_1 + a) + x_2 = (a + x_1) + x_2 = 0 + x_2 = x_2 + 0 = x_2$$

4) Wir bezeichnen die Lösung x von $a + x = 0$ mit $-a$, sowie die Lösung y von $a \cdot y = 1$ mit $\frac{1}{a}$ bzw. a^{-1}

Definiere: $a - b := a + (-b),\ \frac{a}{b} := a \cdot \frac{1}{b}$

Satz 2.1.1. *Für $a, b \in \mathbb{R}$ gelten:*

$$-(-a) = a \qquad -(a+b) = (-a) + (-b)$$
$$\left(a^{-1}\right)^{-1} = a \qquad (a \cdot b)^{-1} = a^{-1} \cdot b^{-1} \quad \text{für } a, b \neq 0$$
$$a \cdot 0 = 0 \qquad a \cdot (-b) = -(a \cdot b)$$
$$(-a) \cdot (-b) = a \cdot b \qquad a \cdot (b - c) = a \cdot b - a \cdot c$$
$$a \cdot b = 0 \Rightarrow a = 0 \quad \text{oder } b = 0$$

Beweis:

$$(-a) + a = a + (-a) = 0 \quad \Rightarrow -(-a) = a$$
$$a \cdot 0 + a \cdot 0 = a \cdot (0 + 0) = a \cdot 0 \quad \Rightarrow a \cdot 0 = 0$$
$$\text{Sei } \; a \neq 0 \Rightarrow 0 = (a \cdot b) \cdot \frac{1}{a} = \left(a \cdot \frac{1}{a}\right) \cdot b = b \quad \Rightarrow a \cdot b = 0.$$

□

<u>Anordnungsaxiome</u>

i) Für jedes $a \in \mathbb{R}$ gilt genau eine der drei Aussagen: $a > 0,\ a = 0,\ a < 0$
ii) Aus $a, b > 0$ folgt: $a + b > 0$ und $a \cdot b > 0$

BEMERKUNG: Statt $-a > 0$ schreiben wir $a < 0$ und statt $a - b > 0$ schreiben wir $a > b$.

Satz 2.1.2. *Für $a, b, c, d \in \mathbb{R}$ gelten:*

1) Für $a, b \in \mathbb{R}$ gilt entweder $a > b,\ a = b$ oder $a < b$
2) $a > b,\ b > c \ \Rightarrow\ a > c$
3) $a > b \ \Rightarrow \begin{cases} a + c & > b + c \qquad \forall c \in \mathbb{R} \\ a \cdot c & > b \cdot c \qquad \forall c > 0 \\ a \cdot c & < b \cdot c \qquad \forall c < 0 \end{cases}$
4) $a > b,\ c > d \ \Rightarrow \begin{cases} a + c & > b + d \\ a \cdot c & > b \cdot d \qquad \text{falls } b, d > 0 \end{cases}$
5) $a \neq 0 \ \Rightarrow\ a^2 > 0$
6) $a > 0 \ \Rightarrow\ \frac{1}{a} > 0$
7) $a > b > 0 \ \Rightarrow\ \frac{1}{a} < \frac{1}{b}$

Beweis: 2) $a - c = (a - b) + (b - c) > 0$
3) $a \cdot c - b \cdot c = (a - b) \cdot c > 0 \quad \forall c > 0$
5) $a^2 = \begin{cases} a \cdot a & \text{falls } a > 0 \\ (-a) \cdot (-a) & \text{falls } a < 0 \end{cases} > 0$
6) $\frac{1}{a} = \left(\frac{1}{a^2}\right) \cdot a > 0.$ □

Definition 2.1.1. *Der Betrag von $a \in \mathbb{R}$ ist definiert durch*

$$|a| := \begin{cases} \quad a, & \text{falls } a \geq 0 \\ - \quad a, & \text{falls } a < 0 \end{cases}$$

Satz 2.1.3. *Für alle $a, b \in \mathbb{R}$ gilt:*

1) $|-a| = |a|, \ a \leq |a|$
2) $|a| \geq 0$ *und* $|a| = 0 \ \Rightarrow \ a = 0$
3) $|a \cdot b| = |a| \cdot |b|$
4) $|a + b| \leq |a| + |b|$ *(Dreiecks - Ungleichung)*
5) $|a - b| \geq ||a| - |b||$
6) $\forall \delta > 0 : \ 2|ab| \leq \delta a^2 + \frac{b^2}{\delta}$ *(Young'sche Ungleichung)*

Beweis: 1)$|-a| = \begin{cases} -a, & \text{falls } -a \geq 0 \\ -(-a), & \text{falls } -a < 0 \end{cases} = \begin{cases} -a, & \text{falls } a \leq 0 \\ a, & \text{falls } a > 0 \end{cases} = |a|$

Und $\ |a| - a = \begin{cases} 0, & \text{falls } a \geq 0 \\ (-a) + (-a), & \text{falls } a < 0 \end{cases} \geq 0$

2) und 3) Klar!

4) $|a + b| = \pm(a + b) = \pm a \pm b \leq |a| + |b|$

6) Ohne Einschränkung (o.E.) gelte: $a \geq 0, \ b \geq 0$
Sei $\varepsilon := \sqrt{\delta}$. Dann gilt:

$$0 \leq \left(\varepsilon a - \frac{b}{\varepsilon}\right)^2 = \epsilon^2 a^2 - 2ab + \frac{b^2}{\epsilon^2} = \delta a^2 - 2ab + \frac{b^2}{\delta}.$$

Damit erhalten wir wie gewünscht

$$2|ab| = 2ab \leq \delta a^2 + \frac{b^2}{\delta}.$$

Ist $a < 0$ oder $b < 0$, so wenden wir obiges Argument auf $-a$ bzw. $-b$ an. □

2.2 Vollständige Induktion

<u>Induktionsprinzip</u>
Sei $M \subset \mathbb{N}$ eine Menge mit den Eigenschaften
1) $\quad 1 \in M$
2) $\quad n \in M \ \Rightarrow \ n + 1 \in M$

dann gilt schon $M = \mathbb{N}$.

Satz 2.2.1. *Gegeben seien Aussagen $A(n)$ für alle $n \in \mathbb{N}$. Es gelte:*
i) $A(1)$ *ist wahr*
ii) $A(n)$ *ist wahr* $\Rightarrow$ $A(n+1)$ *ist wahr.*

Dann sind alle Aussagen $A(n)$ wahr.

Beweis: $M := \{n \in \mathbb{N} : A(n) \text{ ist wahr}\}$
i) $\Rightarrow 1 \in M$
ii) $n \in M \ \Rightarrow \ n+1 \in M$

Induktionsprinzip $\Rightarrow$ $M = \mathbb{N}$. □

Beispiel: Arithmetische Summe
Behauptung: $A(n) : 1+2+3+4+\ldots+n = \sum\limits_{k=1}^{n} k \ = \frac{n(n+1)}{2}$ ist wahr $\forall n \in \mathbb{N}$

<u>Beweis:</u>

1. Schritt: $A(1) : 1 = \frac{1\cdot(1+1)}{2}$ ist wahr
2. Schritt: angenommen $A(n)$ ist wahr. Dann gilt
$1+2+3+4+\ldots+n+(n+1) = \frac{n(n+1)}{2} + (n+1) = \frac{(n+1)(n+2)}{2}$,

d.h. $A(n+1)$ ist auch wahr $\Rightarrow$ $A(n)$ ist wahr $\forall n \in \mathbb{N}$!

Satz 2.2.2 (Bernoullische Ungleichung)**.** *Für $x \in \mathbb{R}$, $x \geq -1$ und für alle $n \in \mathbb{N}$ gilt:*

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

Beweis: 1. Schritt: $n = 1 :\ 1+x \geq 1+x \qquad \surd$
2. Schritt: $n \to n+1$:
$x \geq -1 \ \Rightarrow \ 1+x \geq 0$ und damit folgt
$(1+x)^{n+1} = (1+x)(1+x)^n \geq (1+x)(1+nx) = 1+(n+1)x+nx^2 \geq 1+(n+1)x.$ □

Definition 2.2.1. *Für $n \in \mathbb{N}_0$ definiere $n!$ induktiv durch:*
$0! = 1,\ (n+1)! = (n+1)n! \ \Rightarrow \ n! = 1 \cdot 2 \cdot 3 \cdot \ldots \cdot n.$
Für $x \in \mathbb{R}$, $k \in \mathbb{N}_0$ definiere den Binomialkoeffizient $\binom{x}{k}$ durch

$$\binom{x}{k} := \frac{x \cdot (x-1) \cdot \ldots \cdot (x-(k-1))}{k!} \qquad \text{für } k \in \mathbb{N}, \binom{x}{0} = 1$$

Bemerkung:

$$i) \binom{x}{k} = 0 \qquad \text{für } x \in \mathbb{N}_0 \text{ und } k > x$$

$$ii) \binom{n}{k} = \frac{n!}{k!(n-k)!} \qquad \text{für } n, k \in \mathbb{N}_0 \text{ und } k \leq n$$

Satz 2.2.3. *Für alle $k \in \mathbb{N}$ und $x \in \mathbb{R}$ gilt:*

$$\binom{x+1}{k} = \binom{x}{k-1} + \binom{x}{k}$$

Beweis: Siehe Übungen! □

Satz 2.2.4 (Binomischer Lehrsatz)**.** *Für alle $a, b \in \mathbb{R}$ und für alle $n \in \mathbb{N}$ gilt:*

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

Beweis: Induktionsanfang $n = 1$:

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

Induktionsschluss $n \to (n+1)$:

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

□

Beispiel:

$$\begin{aligned} n=2: \ (a+b)^2 &= \sum_{k=0}^{2} \binom{2}{k} a^k b^2 - k \\ &= \binom{2}{0} a^0 b^{2-0} + \binom{2}{1} a^1 b^{2-1} + \binom{2}{2} a^2 b^{2-2} \\ &= b^2 + 2ab + a^2 \end{aligned}$$

Satz 2.2.5. *Es gibt kein $r \in \mathbb{Q}$ mit $r^2 = 2$!*

Beweis: Annahme: $\exists r = \frac{p}{q} \in \mathbb{Q}$ mit $r^2 = 2$.
o.E. (ohne Einschränkung) gelte: p und q sind teilerfremd.
Aus $r^2 = 2$ folgt $p^2 = 2q^2$. Damit ist p^2 und also auch p gerade.
Schreibe z.B. $p = 2k$ mit $k \in \mathbb{N}$.
Setzt man dies in die obige Formel ein, so folgt $4k^2 = 2q^2$.
Damit ist also auch q gerade und dies ist ein Widerspruch zur Annahme. □

Definition 2.2.2. *Eine Menge $D \neq \emptyset$ heisst abzählbar, falls eine surjektive Abbildung $f : \mathbb{N} \to D$ existiert.*

Satz 2.2.6. *$\mathbb{Z}$ und $\mathbb{Q}$ sind abzählbar!*

Beweis für $\mathbb{Z}$: Definiere $f : \mathbb{N} \to \mathbb{Z}$ durch

$$f(n) = \begin{cases} \frac{n-1}{2}, & \text{für } n = 2k+1, \quad k \in \mathbb{N}_0 \\ -\frac{n}{2}, & \text{für } n = 2k, \quad k \in \mathbb{N} \end{cases}$$

Behauptung: f ist surjektiv.

Sei dazu $l \in \mathbb{Z}$ beliebig. Ist

$$\left.\begin{array}{lll} 1) & l \geq 0 \text{ dann gilt:} & f(2l+1) = l \\ 2) & l < 0 \text{ dann gilt:} & f(-2l) = l \end{array}\right\} \Rightarrow \text{Behauptung.}$$

□

3 $\mathbb{R}^n$, $\mathbb{C}$ und Polynome

3.1 $\mathbb{R}^n$

Definition 3.1.1. Vektorraum
Ein $(\mathbb{K}-)$ *Vektorraum* $(V,+,\cdot)$ *ist eine Menge* V *mit einer Addition* $+: V \times V \to V$ *und einer Skalarmultiplikation* $\cdot\ :\ \mathbb{K} \times V \to V$*, so dass gilt:*

- $x+(y+z) = (x+y)+z \ \forall x,y,z \in V$
- $x+y = y+x \ \forall x,y \in V$
- $\exists 0 \in V : x+0 = x \ \forall x \in V$
- $\forall x \in V \ \exists y \in V : x+y = 0$
- $a(x+y) = ax+ay \ \forall a \in \mathbb{K}\ , x,y \in V$
- $(a+b)x = ax+bx \ \forall a,b \in \mathbb{K}\ , x \in V$
- $a(bx) = (ab)x \ \forall a,b \in \mathbb{K},\ x \in V$
- $1 \cdot x = x \ \forall x \in V \qquad 1:$ *neutrales Element bezüglich* $\cdot$ *in* $\mathbb{K}$

Beispiel:

$$\begin{aligned} \mathbb{R}^n &= \{x = (x_1,\ldots,x_n) : x_i \in \mathbb{R} \ \forall\ 1 \le i \le n\} \text{ ist ein } \mathbb{R}-\text{Vektorraum} \\ x+y &= (x_1+y_1,\ldots,x_n+y_n) \qquad ax = (ax_1,\ldots,ax_n)\ \forall a \in \mathbb{R} \\ x &= (x_1,\ldots,x_n) \\ y &= (y_1,\ldots,y_n) \end{aligned}$$

Kanonische Basis $e_i = (0,\ldots,0,1,0,\ldots,0)\ 1 \le i \le n$
Jeder Vektor $x = (x_1,\ldots,x_n) \in \mathbb{R}^n$ lässt sich eindeutig als Linearkombination $x = (x_1,\ldots,x_n) = x_1e_1 + \ldots + x_ne_n = \sum\limits_{i=1}^n x_ie_i$ darstellen. (Standardbasis)

Skalarprodukt für $x = (x_1,\ldots,x_n) \in \mathbb{R}^n, y = (y_1,\ldots,y_n) \in \mathbb{R}^n$ definiere
$x \cdot y = < x,y > := x_1y_1 + \ldots + x_ny_n = \sum\limits_{i=1}^n x_iy_i \ \in \mathbb{R}$

Es gelten:

SP: i)	$x \cdot y = y \cdot x$	Symmetrie
SP: ii)	$x \cdot (y+z) = x \cdot y + x \cdot z$	Bi-Linearität
SP: iii)	$x \cdot (ay) = a(x \cdot y)$	$\forall a \in \mathbb{R}$

BEISPIEL: $x = (2,0,3),\ y = (-3,1,2) \Rightarrow x \cdot y = -6+0+6 = 0$ d.h. x und y sind orthogonal.
Dies gilt auch für $e_i \cdot e_j = 0\ \forall i \neq j$.

<u>Euklidische Norm</u>
$\|x\| := \sqrt{x \cdot x} = \sqrt{\sum_{i=1}^{n} x_i^2}$

BEISPIEL: i) $\|e_i\| = 1\ \forall 1 \leq i \leq n$, d.h. e_i sind paarweise orthogonal und auf Länge 1 normiert, sie sind orthonormal.

ii) Satz von Pythagoras: Seien $x, y \in \mathbb{R}^n$ orthogonal, d.h. $x \cdot y = 0$
$\Rightarrow\ \|x+y\|^2 = (x+y)\cdot(x+y) = x \cdot x + 2x \cdot y + y \cdot y = x \cdot x + y \cdot y = \|x\|^2 + \|y\|^2$

Satz 3.1.1. *Für die euklidische Norm gilt:*

i)	$\forall x \in \mathbb{R}^n :$	$\|x\| \geq 0$ *und* $\|x\| = 0 \ \Rightarrow\ x = 0$
ii)	$\forall x \in \mathbb{R}^n,\ a \in \mathbb{R} :$	$\|ax\| = \lvert a\rvert\, \|x\|$
iii)	$\forall x, y \in \mathbb{R}^n :$	$\|x+y\| \leq \|x\| + \|y\|$

Beweis:

$$
\begin{aligned}
i)\ \|x\| &\geq 0 \qquad \text{folgt direkt aus Definition der Norm} \\
\|x\| &= 0 \ \Rightarrow\ \sum_{i=1}^{n} x_i^2 = 0 \ \Rightarrow\ x_i = 0\ \forall\ 1 \leq i \leq n \ \Rightarrow\ x = 0 \\
ii)\ \|ax\| &= \sqrt{\sum_{i=1}^{n} (ax_i)^2} = \sqrt{a^2 \sum_{i=1}^{n} x_i^2} = |a|\, \|x\| \\
iii)\ \|x+y\|^2 &= (x+y)\cdot(x+y) = x \cdot x + 2x \cdot y + y \cdot y \\
&\leq \|x\|^2 + 2\,\|x\|\,\|y\| + \|y\|^2 = (\|x\| + \|y\|)^2 ,
\end{aligned}
$$

wobei im letzten Schritt Satz 3.1.2 benutzt wurde. □

Satz 3.1.2 (Cauchy-Schwarz Ungleichung)**.** *Für alle $x, y \in \mathbb{R}^n$ gilt:*

$$|x \cdot y| \leq \|x\| \cdot \|y\|$$

und Gleichheit tritt genau dann ein, wenn $y = \lambda x$ für ein $\lambda \in \mathbb{R}$.

Beweis: o.E.: $x \neq 0$ (sonst sind beide Seiten gleich 0)
Schreibe: $y = \frac{x}{\|x\|}\left(\frac{x}{\|x\|} \cdot y\right) + \left(y - \frac{x}{\|x\|}\left(\frac{x}{\|x\|} \cdot y\right)\right) =: y^{\|} + y^{\perp}$
Es gilt: $x \cdot y^{\perp} = x \cdot y - \frac{x \cdot x}{\|x\|}\left(\frac{x}{\|x\|} \cdot y\right) = x \cdot y - x \cdot y = 0.$
Damit folgt auch $y^{\|} \cdot y^{\perp} = (\frac{x}{\|x\|^2} \cdot y) x \cdot y^{\perp} = 0.$
Aus dem Satz von Pythagoras erhalten wir also

$$\frac{|x \cdot y|}{\|x\|} = \left\|y^{\|}\right\| \leq \sqrt{\|y^{\|}\|^2 + \|y^{\perp}\|^2} = \sqrt{\|y^{\|} + y^{\perp}\|^2} = \|y\|$$

Also gilt: $|x \cdot y| \leq \|x\| \, \|y\|$.

Gleichheit gilt, falls

$$\left\|y^{\|}\right\| = \sqrt{\|y^{\|}\|^2 + \|y^{\perp}\|^2} \Leftrightarrow y^{\perp} = 0,$$

also genau dann, wenn

$$y = (\frac{x}{\|x\|^2} \cdot y)x = \lambda x$$

mit $\lambda = \frac{x}{\|x\|^2} \cdot y$. □

Euklidische Metrik
Für $x, y \in \mathbb{R}^n$ definiere $d(x,y) := \|x - y\|$ („Abstand")
Es gilt für alle $x, y, z \in \mathbb{R}^n$

i) $d(x,y) \geq 0$, $d(x,y) = 0 \Leftrightarrow x = y$
ii) $d(x,y) = \|x - y\| = \|-(y - x)\| = |-1| \cdot \|y - x\| = d(y,x)$
iii) $d(x,z) = \|x - z\| = \|x - y + y - z\| \leq \|x - y\| + \|y - z\| = d(x,y) + d(y,z)$

3.2 $\mathbb{C}$

Auf dem $\mathbb{R}$-Vektorraum haben wir eine Addition
$\mathbb{R}^2 : + : (a,b), (c,d) \to (a + c, b + d) \in \mathbb{R}^2$
und wir definieren zusätzlich eine Multiplikation:
Komplexe Multiplikation $\cdot$: $(a,b), (c,d) \ \to (ac - bd, ad + bc) \in \mathbb{R}^2$

Ausage: $(\mathbb{R}^2, +, \cdot)$ ist ein Körper (Körper der komplexen Zahlen)
Neutrales Element bezüglich $\cdot$: $(1,0)$
$[(1,0) \cdot (a,b) = (a,b)]$
Inverses Element bezüglich $\cdot$: $(a,b) \neq (0,0)$
$(a,b) \ \left(\frac{a}{a^2+b^2}, \frac{-b}{a^2+b^2}\right) = \left(\frac{a^2}{a^2+b^2} + \frac{b^2}{a^2+b^2}, -\frac{ab}{a^2+b^2} + \frac{ab}{a^2+b^2}\right) = (1,0)$

Bemerkung: Wir können $\mathbb{R}$ in $\mathbb{C}$ „Einbetten".
$\mathbb{R} \ni x \to (x,0) \in \mathbb{C}$
Es gilt:
$x + y \to (x + y, 0) = (x,0) + (y,0)$
$x \cdot y \to (xy, 0) = (x,0) \cdot (y,0)$

Skalarmultiplikation in $\mathbb{R}^2$
$\alpha \in \mathbb{R} \ \ \alpha(x,y) = (\alpha x, \alpha y) = (\alpha, 0)(x,y)$

Jetzt identifizieren wir den Standardbasisvektor $e_1 = (1,0)$ mit der Zahl $1 \in \mathbb{R}$. Weiter setzen wir $i := e_2 = (0,1)$ („imaginäre Einheit")

$\Rightarrow \;\; i^2 = (0,1)\cdot(0,1) = (-1,0) \equiv -1$

Damit gilt für jedes $z = (x,y) \in \mathbb{C}: \;\; z = xe_1 + ye_2 = x + iy$
Realteil: $\Re(z) := x$
Imaginärteil: $\Im(z) := y$

Konjugation: Sei $z = x + iy$, dann ist $\overline{z} = x - iy \in \mathbb{C}$ die zu z konjugierte Zahl.

Es gilt:

$$\begin{aligned} i) \qquad z\cdot\overline{z} = (x+iy)(x-iy) &= x^2 - ixy + ixy - i^2y^2 \\ &= x^2 + y^2 \\ &= \|z\|^2 \end{aligned}$$

$$\begin{aligned} ii) \qquad \forall z_1 z_2 \in \mathbb{C} : \overline{z_1 + z_2} &= \overline{z_1} + \overline{z_2} \\ \overline{z_1 \cdot z_2} &= \overline{z_1}\cdot\overline{z_2}, denn \\ \overline{z_1 \cdot z_2} = \overline{(x_1+iy_1)(x_2+iy_2)} &= \overline{[x_1x_2 + i(x_2y_1 + x_1y_2) - y_1y_2]} \\ &= x_1x_2 - y_1y_2 - i(x_2y_1 + x_1y_2) \\ \overline{z_1}\cdot\overline{z_2} = (x_1 - iy_1)(x_2 - iy_2) &= x_1x_2 - y_1y_2 - i(x_2y_1 + x_1y_2) \end{aligned}$$

Folgerung:
i) $\forall z \in \mathbb{C}\setminus\{0\}$ gilt: $z^{-1} = \frac{\overline{z}}{\|z\|^2}$
ii) $\|zw\|^2 = (zw)\cdot(\overline{zw}) = zw\cdot(\overline{z}\;\overline{w})$
$= (z\cdot\overline{z})\cdot(w\cdot\overline{w}) = \|z\|^2\,\|w\|^2$
$\Rightarrow \;\; \|zw\| = \|z\|\cdot\|w\| \qquad \forall z,w \in \mathbb{C}$

Im Folgenden schreiben wir: $\|z\| = |z|$

Beispiel:

$$i)\; (2+i)^{-1} = \frac{2-i}{4+1} = \frac{2}{5} - \frac{i}{5}$$
$$ii)\; \frac{2+i}{2-1} := (2+i)(2-i)^{-1} = \frac{2+i}{2-i}\cdot\frac{2+i}{2+i} = \frac{1}{5}(4-1+4i) = \frac{3}{5} + \frac{4}{5}i$$

3.3 Polynome

In diesem Abschnitt sei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$.

Definition 3.3.1. *$p : \mathbb{K} \to \mathbb{K}$ heisst Polynom vom Grad $n \in \mathbb{N}_0$, falls $a_0, a_1, \ldots a_n \in \mathbb{K}$ existieren mit $a_n \neq 0$ und $p(z) = \sum\limits_{k=0}^{n} a_k z^k$ für alle $z \in \mathbb{K}$.*

Lemma 3.3.1. *Sei $p : \mathbb{K} \to \mathbb{K}$ ein Polynom vom Grad $n \in \mathbb{N}_0$ mit Koeffizienten $a_0, a_1, \ldots, a_n \in \mathbb{K}, \;\; a_n \neq 0$. Ist $p(\lambda) = 0$ für ein $\lambda \in \mathbb{K}$, so existiert ein Polynom*

$q : \mathbb{K} \to \mathbb{K}$ vom Grad $(n-1)$ mit Koeffizienten $b_0, b_1, \ldots, b_{n-1} = a_n$, so dass gilt:

$$p(z) = (z - \lambda)q(z).$$

Beweis: Es gilt für alle $k \in \mathbb{N}$ (siehe Übung):

$$z^k - \lambda^k = (z - \lambda) \sum_{j=1}^{k} \lambda^{k-j} z^{j-1}.$$

Damit erhalten wir

$$p(z) = p(z) - p(\lambda) = \sum_{k=0}^{n} a_k(z^k - \lambda^k) = (z - \lambda) \sum_{k=0}^{n} \sum_{j=1}^{k} a_k \lambda^{k-j} z^{j-1}.$$

Indem man nun q durch

$$q(z) := \sum_{k=1}^{n} \sum_{j=1}^{k} a_k \lambda^{k-j} z^{j-1} = b_0 + \ldots + b_{n-1} z^{n-1}$$

definiert, folgt

$$p(z) = (z - \lambda)q(z).$$

Die Potenz z^{n-1} entsteht in der Definition von q nur für $j = k = n \;\Rightarrow\; b_{n-1} = a_n \neq 0$. □

Lemma 3.3.2. *Sei $p : \mathbb{K} \to \mathbb{K}$ ein Polynom vom Grad $n \in \mathbb{N}_0$. Dann hat p höchstens n Nullstellen.*

Beweis: Induktion über $n \in \mathbb{N}_0$:

$n = 0 : \Rightarrow\; p(z) = a_0 \neq 0 \;\Rightarrow p$ hat keine Nullstellen $\quad\checkmark$

$n - 1 \to n :$ Sei p ein Polynom vom Grad n
Fall i): p hat keine Nullstelle $\Rightarrow$ fertig!
Fall ii): p hat eine Nullstelle $\lambda \in \mathbb{K} \;\Rightarrow\; p(z) = (z - \lambda)q(z)$ wegen Lemma 3.3.1, und $q : \mathbb{K} \to \mathbb{K}$ ist ein Polynom vom Grad $(n-1)$
Induktionsvoraussetzung (IV):
$\Rightarrow\; q$ hat höchstens $(n-1)$ Nullstellen
$\Rightarrow\; p$ hat höchstens n Nullstellen □

Satz 3.3.1. *Seien $p, q : \mathbb{K} \to \mathbb{K}$ Polynome vom Grad n bzw. m, d.h. es gilt mit $a_n, b_m \neq 0$*

$$p(z) = \sum_{i=0}^{n} a_i z^i, \qquad\qquad q(z) = \sum_{k=0}^{m} b_k z^k$$

Ist $p(\lambda_i) = q(\lambda_i)$ für paarweise verschiedene $\lambda_1, \ldots, \lambda_l \in \mathbb{K}$ mit $l > max\,\{n, m\}$, so folgt $m = n$ und $a_i = b_i \;\forall\; 1 \leq i \leq n$.

Beweis: Wäre $m \neq n$ oder $a_i \neq b_i$ für ein i, so wäre $p - q$ ein Polynom vom Grad höchstens $max\{n,m\}$ mit $l > max\{n,m\}$ Nullstellen $\lambda_1, \ldots, \lambda_l$. zu Lemma 3.3.2. □

Lemma 3.3.3. *Sei $p : \mathbb{K} \to \mathbb{K}$ ein Polynom vom Grad n. Dann besitzt p die eindeutige Darstellung*

$$p(z) = (z - \lambda_1)^{\nu_1} \cdot \ldots \cdot (z - \lambda_k)^{\nu_k} \cdot q(z) \qquad \forall z \in \mathbb{K},$$

wobei $\lambda_1, \ldots, \lambda_k$ die Nullstellen von p mit Vielfachheit $\nu_i \in \mathbb{N}$ sind, und q ein Polynom vom Grad $n - \sum_{i=1}^{k} \nu_i \in \{0, \ldots, n\}$ ist. Weiter gilt $q(z) \neq 0 \ \forall z \in \mathbb{K}$.

Beweis: Existenz: Folgt iterativ aus Lemma 3.3.1.
Eindeutigkeit: Sei für alle $z \in \mathbb{C}$

$$p(z) = (z - \lambda_1)^{\nu_i} \cdot \ldots \cdot (z - \lambda_k)^{\nu_k} \cdot q(z) = (z - \lambda_1)^{\mu_i} \cdot \ldots \cdot (z - \lambda_k)^{\mu_k} \cdot l(z)$$

Ohne Einschränkung gelte $\nu_1 \geq \mu_1$. Dann folgt für alle $z \in \mathbb{K} \setminus \{\lambda_1\}$

$$(z - \lambda_1)^{\nu_i - \mu_1} \cdot (z - \lambda_2)^{\nu_2} \cdot \ldots \cdot (z - \lambda_k)^{\nu_k} \cdot q(z) = (z - \lambda_2)^{\mu_2} \cdot \ldots \cdot (z - \lambda_k)^{\mu_k} \cdot l(z).$$

Beide Seiten sind Polynome und aus Satz 3.3.1 folgt damit, dass beide Seiten sogar für $z = \lambda_1$ übereinstimmen müssen. Dies impliziert sofort $\nu_1 = \mu_1$. Analog folgt $\nu_i = \mu_i \ \forall \ 1 \leq i \leq k$.
Nach Division erhalten wir damit

$$q(z) = l(z) \ \forall z \in \mathbb{K} \setminus \{\lambda_1, \ldots, \lambda_k\}.$$

Aus Satz 3.3.1) folgt dann wieder

$$q(z) = l(z) \ \forall z \in \mathbb{K}.$$

□

Satz 3.3.2 (Fundamentalsatz der Algebra)**.** *Jedes nichtkonstante Polynom $p : \mathbb{C} \to \mathbb{C}$ besitzt eine Nullstelle!*

Beweis: Folgt später. □

Lemma 3.3.4. *Jedes Polynom $p : \mathbb{C} \to \mathbb{C}$ vom Grad n mit Koeffizienten $a_i \in \mathbb{C}$ $0 \leq i \leq n$, zerfällt über $\mathbb{C}$ in Linearfaktoren, d.h. es existieren $\lambda_1, \ldots, \lambda_k \in \mathbb{C}$ und $\nu_1, \ldots, \nu_k \in \mathbb{N}$, so dass für alle $z \in \mathbb{C}$ gilt:*

$$p(z) = a_n (z - \lambda_1)^{\nu_1} \cdot \ldots \cdot (z - \lambda_k)^{\nu_k}.$$

Weiter gilt: $\sum_{i=1}^{k} \nu_i = n$.

Beweis: Folgt direkt aus Lemma 3.3.3 und Satz 3.3.2. □

4 Konvergenz und Vollständigkeit von $\mathbb{R}$

4.1 Vollständigkeit

Definition 4.1.1. *Sei $M \subset \mathbb{R}$ und $M \neq \emptyset$.*
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 obere (untere) Schranke von M.

Eine obere (untere) Schranke γ von M mit $\gamma \in M$ heisst Maximum (Minimum).
Notation: $\max M\ (\min M)$

Beispiel:
1) $M = [1,2]\,,\ maxM = 2,\ minM = 1$
2) $M = (1, \infty)$ ist nicht nach oben, aber nach unten beschränkt. (Minimum existiert jedoch nicht, da 1 nicht in M).

Definition 4.1.2. *Ist γ obere (untere) Schranke von M mit $\gamma \leq \widetilde{\gamma}\ (\gamma \geq \widetilde{\gamma})$ für jede andere obere (untere) Schranke $\widetilde{\gamma}$ von M, so heisst γ Supremum (Infimum) von M.*

Bemerkung: Ein Maximum (Minimum) ist immer auch Supremum (Infimum).
Notation: Supremum von M : $\sup M$ (Infimum von M : $\inf M$)
Es gilt: $\sup M = \max M \ \Leftrightarrow\ \sup M \in M$

Beispiel: 1) $M = [1,2)$
Beh.: $1 = \min M = \inf M$ und $2 = \sup M$. $\max M$ existiert nicht.

Beweis: Man sieht sofort, dass 2 eine obere Schranke von M ist.
Sei jetzt $\widetilde{\gamma} < 2$. Im Fall $\widetilde{\gamma} < 1$ kann $\widetilde{\gamma}$ keine obere Schranke sein, denn $1 \in M$.

Ist $\widetilde{\gamma} \geq 1$ so ist $\widetilde{\gamma} < \frac{\widetilde{\gamma}+2}{2} \in M$. In diesem Fall kann $\widetilde{\gamma}$ also auch keine obere Schranke sein. Insofern ist 2 tatsächlich die kleinste ober Schranke (also das Supremum) von M. $\square$

2) $M = \left\{1 - \frac{1}{n} : n \in \mathbb{N}\right\}$
Beh: $\inf M = \min M = 0$ und $\sup M = 1$. $\max M$ existiert nicht, denn $1 \notin M$.

Es ist einfach zu sehen, dass 1 eine ober Schranke von M ist. Für alle $\varepsilon > 0$ ist die Zahl $1 - \varepsilon$ keine obere Schranke von M, denn mit Hilfe von Satz 4.1.1 erhält man die Existenz eines $n \in \mathbb{N}$ mit der Eigenschaft

$$1 - \frac{1}{n} > 1 - \varepsilon \;\Leftrightarrow\; \varepsilon > \frac{1}{n}.$$

Insofern ist 1 die kleinste obere Schranke von M.

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

Lemma 4.1.1. *Äquivalent zum Vollständigkeitsaxiom ist: Jeder nichtleere nach unten beschränkte Teilmenge von $\mathbb{R}$ hat ein Infimum.*

Satz 4.1.1. *(Archimedische Eigenschaft von $\mathbb{R}$) Es gelten*
i) $\forall r \in \mathbb{R}^+ \; \exists n \in \mathbb{N} : n > r$
ii) $\forall \varepsilon > 0 \; \exists n \in \mathbb{N} : \frac{1}{n} < \varepsilon.$

Beweis: Zuerst zeigen wir die Implikation $i) \Rightarrow ii)$. Dazu sei $\varepsilon > 0$. Nach $i)$ existiert ein $n \in \mathbb{N}$ mit

$$n > \frac{1}{\varepsilon},$$

oder

$$\frac{1}{n} < \varepsilon.$$

Wir müssen also nur noch $i)$ beweisen. Dazu nehmen wir an, dass ein $r \in \mathbb{R}^+$ existiert, so dass für alle $n \in \mathbb{N}$ gilt:

$$n \leq r,$$

d.h. $\mathbb{N}$ ist beschränkt. Nach dem Supremumsaxiom existiert also ein $s = \sup \mathbb{N} \in \mathbb{R}$ mit der Eigenschaft

$$n + 1 \leq s \quad \forall \; n \in \mathbb{N}\,.$$

Dies impliziert aber

$$n \leq s - 1 \quad \forall \; n \in \mathbb{N}$$

und damit ist auch $s - 1$ eine ober Schranke, im Widerspruch zu $s = \sup \mathbb{N}$. Also war unsere Annahme falsch, und damit ist $i)$ bewiesen. $\square$

4.2 Folgen

Beispiele:
i) Fibonacci-Folge: $1, 1, 2, 3, 5, 8, 13, \ldots$
Diese Folge ist für alle $n \in \mathbb{N}$ rekursiv definiert durch

$$a_0 = 1, \; a_1 = 1, \; a_{n+1} = a_n + a_{n-1}.$$

ii) Zinsfaktoren bei $\frac{1}{n}$-tel jährlicher Verzinsung.

$$a_n = (1 + \frac{1}{n})^n$$

Wir werden zeigen, dass a_n konvergiert, und den Grenzwert bezeichnen wir mit e, der Eulerschen Zahl.

iii) Geometrische Reihe: Für alle $n \in \mathbb{N}$ betrachte

$$s_n = 1 + q + q^2 + \ldots + q^n = \sum_{k=0}^{n} q^k.$$

<u>Grenzwert einer Folge</u>
Sei $(a_n)_{n\in\mathbb{N}} = (a_1, a_2, \ldots)$ Folge in $\mathbb{R}$, $a \in \mathbb{R}$

Definition 4.2.1. *i)* $(a_n)_{n\in\mathbb{N}}$ *<u>konvergiert</u> gegen* a *für* $n \to \infty$*, falls gilt:*

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

Wir schreiben $a = \lim\limits_{n\to\infty} a_n$ *oder* $a_n \to a$*.* a *heisst Grenzwert oder Limes der Folge.*

ii) (a_n) *heisst konvergent, falls die Folge einen Grenzwert besitzt. Ansonsten heisst sie divergent.*

Beispiele:

i) a) $a_n = \frac{1}{n}$ $\forall n \in \mathbb{N}$. Behauptung: $a_n \to 0$.

Beweis: Sei $\varepsilon > 0$. Nach Satz 4.1.1 existiert ein $n_0 \in \mathbb{N}$, so dass gilt:

$$\frac{1}{n_0} < \varepsilon.$$

Für alle $n \geq n_0$ gilt damit

$$-\varepsilon < 0 < \frac{1}{n} \leq \frac{1}{n_0} < \varepsilon,$$

also

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

$\square$

b) Analog argumentiert man für $a_n = \sqrt{\frac{1}{n}} \to 0$:
Wiederum nach Satz 4.1.1 existiert für alle $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$ mit

$$n_0 > \frac{1}{\varepsilon^2} \;\Leftrightarrow\; \sqrt{n_0} > \frac{1}{\varepsilon}.$$

Damit folgt für alle $n \geq n_0$

$$0 < \frac{1}{\sqrt{n}} \leq \frac{1}{\sqrt{n_0}} < \varepsilon.$$

ii) Sei $q \in \mathbb{R}$ mit $0 < q < 1$. Dann gilt: $q^n \to 0$.

Beweis: Sei $\varepsilon > 0$ und schreibe $\frac{1}{q} = 1 + \delta$ mit $\delta > 0$.

Mit Hilfe der Bernoullischen Ungleichung, Satz 2.2.2, schätzen wir für alle $n \in \mathbb{N}$ ab:

$$\frac{1}{q^n} = \left(\frac{1}{q}\right)^n = (1+\delta)^n \geq 1 + n\delta \geq n\delta.$$

Damit gilt für alle $n \in \mathbb{N}$

$$0 < q^n \leq \frac{1}{n\delta}.$$

Nach Satz 4.1.1 existiert ein $n_0 \in \mathbb{N}$ mit

$$n_0 > \frac{1}{\varepsilon\delta}.$$

Insgesamt folgt damit für alle $n \geq n_0$

$$0 < q^n \leq \frac{1}{n\delta} \leq \frac{1}{n_0\delta} < \varepsilon.$$

$\square$

iii) $a_n = (-1)^n$ ist divergent.
Für alle $a \in \mathbb{R},\ n \in \mathbb{N}$ gilt:

$$|a_n - a| + |a_{n+1} - a| \geq |a_n - a + a - a_{n-1}| = |a_n - a_{n+1}| = 2.$$

Es kann also insbesondere kein $a \in \mathbb{R}$ existieren, welches die Abschätzung $|a_n - a| < \frac{1}{2}$ für alle $n \geq n_0$ erfüllt. Also ist die Folge divergent.

iv) $a_n = n$ ist divergent, denn für alle $a \in \mathbb{R}$ existiert nach Satz 4.1.1 ein $n_0 \in \mathbb{N}$ mit

$$a + 1 < n_0.$$

Für alle $n \geq n_0$ gilt dann

$$|a_n - a| = n - a \geq n_0 - a > 1$$

und damit ist a_n divergent.

Satz 4.2.1. *Die Folge (a_n) konvergiere sowohl gegen $a \in \mathbb{R}$, als auch gegen $b \in \mathbb{R}$. Dann gilt $a = b$.*

Beweis: Wir nehmen an das $a \neq b$ ist, und wir definieren $\varepsilon := \frac{|a-b|}{2} > 0$.

Weiter wählen wir $n_0 \in \mathbb{N}$ so, dass für alle $n \geq n_0$ gilt

$$|a_n - a| < \varepsilon, \ |a_n - b| < \varepsilon.$$

Damit erhalten wir mit Hilfe der Dreiecksungleichung den Widerspruch

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

□

Satz 4.2.2. *Seien (a_n) und (b_n) konvergente Folgen mit Grenzwerten a bzw. b. Dann konvergieren auch die Folgen $(a_n + b_n)$, $(a_n b_n)$ und es gilt:*

i) $a_n + b_n \to a + b$

ii) $a_n b_n \to ab$

iii) Gilt zusätzlich $b \neq 0 \neq b_n$ für alle $n \in \mathbb{N}$, dann gilt auch $\frac{a_n}{b_n} \to \frac{a}{b}$.

iv) Ist $a_n \leq b_n \ \forall n \in \mathbb{N}$ so gilt $a \leq b$.

Beweis: Zu $\varepsilon > 0$ wähle $n_0 \in \mathbb{N}$, so dass für alle $n \geq n_0$ gilt:

$$|a_n - a| < \varepsilon, \quad |b_n - b| < \varepsilon.$$

zu i) Für alle $n \geq n_0$ schätzen wir mit der Dreiecksungleichung ab:

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

Da $\varepsilon > 0$ beliebig war, folgt damit die Konvergenz $a_n + b_n \to a + b$.

zu ii) Ohne Einschränkung sei $\varepsilon < 1$. Dann gilt für alle $n \geq n_0$:

$$|b_n| \leq |b_n - b| + |b| < 1 + |b|.$$

Damit erhalten wir wieder aus der Dreiecksungleichung für alle $n \geq n_0$

$$\begin{aligned} |a_n b_n - ab| =& |(a_n - a)b_n + a(b_n - b)| \\ \leq& |b_n||a_n - a| + |a||b_n - b| \\ \leq& (1 + |b| + |a|)\varepsilon \end{aligned}$$

und da $\varepsilon > 0$ beliebig war, folgt damit wieder die gewünschte Aussage.

iii) Wir betrachten zuerst den Spezialfall $a_n = a = 1$.

Ohne Einschränkung sei diesmal $0 < \varepsilon < \frac{|b|}{2}$. Damit gilt für alle $n \geq n_0$

$$|b_n| = |b_n - b + b| \geq |b| - |b_n - b| \geq |b| - \varepsilon \geq \frac{|b|}{2}$$

und

$$|\frac{1}{b_n} - \frac{1}{b}| = \frac{|b - b_n|}{|b||b_n|} \leq \frac{2}{|b|^2}\varepsilon.$$

Da ε wieder beliebig war folgt $b_n^{-1} \to b^{-1}$.

Im allgemeinen Fall benutzen wir die gerade bewiesene Aussage und *ii*) und erhalten

$$\frac{a_n}{b_n} = a_n \frac{1}{b_n} \to a\frac{1}{b} = \frac{a}{b}.$$

iv) Ist $a > b$ so folgt mit $2\varepsilon := a - b > 0$ für alle $n \geq n_0$:

$$a < a_n + \varepsilon \leq b_n + \varepsilon < b + 2\varepsilon = a.$$

Dies ist ein Widerspruch und damit muss $a \leq b$ sein. □

Bemerkung: Aus $a_n < b_n \ \forall n \in \mathbb{N}$ folgt nicht $a < b$! Betrachte dazu die Folgen $a_n = 0$ und $b_n = \frac{1}{n}$. Es gilt: $a_n \leq b_n$ für alle $n \in \mathbb{N}$, aber $a_n \to 0$ und $b_n \to 0$.

Beispiel:
1) $\lim\limits_{n\to\infty} \sqrt[n]{n} = 1$

Beweis: Mit Hilfe der binomischen Formel, Satz 2.2.4, schliessen wir für alle $x \geq 0$ und $n \geq 2$:

$$(1+x)^n = \sum_{k=0}^{n} \binom{n}{k} x^k \geq 1 + nx + \binom{n}{2} x^2 \geq \frac{n(n-1)}{2} x^2 \geq \frac{n^2}{4} x^2.$$

Hierbei haben wir benutzt, dass $n(n-1) \geq \frac{n}{2}$ für alle $n \geq 2$ gilt.

Als nächstes wählen wir $x = \frac{2}{\sqrt{n}}$ und erhalten

$$\left(1 + \frac{2}{\sqrt{n}}\right)^n \geq n \geq 1,$$

oder

$$1 + \frac{2}{\sqrt{n}} \geq \sqrt[n]{n} \geq 1.$$

Aus dem obigen Satz und den vorherigen Beispielen folgt somit die Behauptung. □

2) $\lim\limits_{n\to\infty} n^p q^n = 0$ für alle $0 < q < 1$ und für alle $p \in \mathbb{N}$

Beweis: Wir definieren $0 < s := \sqrt[p]{q} < 1$ und wir schreiben $s = \frac{1}{(1+\varepsilon)^2}$, also $\varepsilon = \frac{1}{\sqrt{s}} - 1 > 0$. Nach dem vorherigen Beispiel existiert ein $n_0 \in \mathbb{N}$, so dass für alle

$n \geq n_0$ gilt:

$$\sqrt[n]{n} < 1 + \varepsilon.$$

Damit folgt

$$0 \leq n^p q^n = (s \sqrt[n]{n})^{np} \leq \left(\frac{1}{1+\varepsilon}\right)^{np} = \left(\frac{1}{(1+\varepsilon)^p}\right)^n \to 0,$$

denn $\frac{1}{(1+\varepsilon)^p} < 1$. $\square$

Definition 4.2.2. *Eine Folge (a_n) heisst nach oben (unten) beschränkt, falls gilt:*

$$\exists b \in \mathbb{R} \quad \forall n \in \mathbb{N} : a_n \leq b \quad (b \leq a_n),$$

d.h. falls die Menge $A = \{a_n : n \in \mathbb{N}\}$ nach oben (unten) beschränkt ist.

Satz 4.2.3. *Jede konvergente Folge (a_n) ist beschränkt.*

Beweis: Wir setzen $a = \lim_{n\to\infty} a_n$ und zu $\varepsilon = 1$ sei $n_0 \in \mathbb{N}$ so gewählt, dass für alle $n \geq n_0$ gilt:

$$|a_n - a| < 1.$$

Damit folgt für alle $n \geq n_0$

$$|a_n| = |a_n - a + a| \leq |a| + |a_n - a| \leq |a| + 1.$$

Dies impliziert für alle $n \in \mathbb{N}$ die Abschätzung

$$|a_n| \leq \max\{|a| + 1, |a_1|, \ldots, |a_{n_0-1}|\} =: b.$$

$\square$

Bemerkung: Die Umkehrung der obigen Aussage ist nicht richtig, wie das Beispiel $a_n = (-1)^n$ zeigt. Diese Folge ist beschränkt, aber nicht konvergent.

Es gilt aber der folgende Satz;

Satz 4.2.4. *Sei (a_n) eine nach oben beschränkte und <u>monoton wachsende</u> Folge, d.h. für ein $b \in \mathbb{R}$ und für alle $n \in \mathbb{N}$ gilt:*

$$a_1 \leq a_2 \leq \ldots \leq a_n \leq a_{n+1} \leq \ldots \leq b,$$

dann ist (a_n) konvergent.

Die analoge Aussage gilt auch für eine nach unten beschränkte und monoton fallende Folge.

Beweis: Wir definieren die Menge $A := \{a_n : n \in \mathbb{N}\} \neq \emptyset$. Nach Voraussetzungist A nach oben beschränkt (durch b). Aus dem Vollständigkeitsaxiom folgt die Existenz

von
$$a = \sup A = \sup_{n \in \mathbb{N}} a_n.$$
Wir behaupten es gilt
$$a = \lim_{n \to \infty} a_n.$$
Sei dazu $\varepsilon > 0$. Nach Aufgabe 4a) auf dem Übungsblatt 3 existiert ein $n_0 = n_0(\varepsilon) \in \mathbb{N}$ mit der Eigenschaft, dass
$$a_{n_0} > a - \varepsilon.$$
Aus der Monotonie der Folge folgt damit für alle $n \geq n_0$
$$a - \varepsilon < a_{n_0} \leq a_n \leq \sup_{l \in \mathbb{N}} a_l = a < a + \varepsilon,$$
oder äquivalent dazu:
$$|a_n - a| < \varepsilon.$$
Dies impliziert die Behauptung und beendet damit den Beweis. □

Eine wichtige Anwendung des gerade bewiesenen Satzes ist die Folgende:

Eine Folge von Intervallen $I_n = [a_n, b_n] \subset \mathbb{R}$ heisst Intervallschachtelung, falls gilt:
i) $I_{n+1} \subset I_n \quad \forall n \in \mathbb{N}$
ii) $|I_n| := b_n - a_n \to 0$

Satz 4.2.5. *Jede Intervallschachtelung $I_n = [a_n, b_n]$ enthält genau einen Punkt $a = \lim_{n \to \infty} a_n = \lim_{n \to \infty} b_n$.*

Beweis: Nach Vorraussetzung gilt für alle $n \in \mathbb{N}$
$$a_1 \leq a_2 \leq \ldots \leq a_n \leq a_{n+1} \leq b_{n+1} \leq b_n \leq \ldots \leq b_2 \leq b_1.$$
Also ist die Folge (a_n) monoton wachsend und nach oben beschränkt (durch b_1), und die Folge (b_n) ist monoton fallend und nach unten beschränkt (durch a_1).

Aus Satz 4.2.4 folgt die Existenz von $a = \lim_{n \to \infty} a_n$ und $b = \lim_{n \to \infty} b_n$.

Mit Hilfe der Eigenschaft *ii)* einer Intervallschachtelung erhalten wir schliesslich $b_n - a_n \to 0$, also $a = b$. □

Beispiel:
1) Eulersche Zahl e
Wir betrachten die Folgen $a_n = \left(1 + \frac{1}{n}\right)^n < b_n = \left(1 + \frac{1}{n}\right)^{n+1}$ und wir behaupten das für alle $n \geq 2$ gilt:
$$2 = a_1 \leq \ldots \leq a_{n-1} \leq a_n \leq b_n \leq b_{n-1} \leq \ldots \leq b_1 = 4.$$

Dazu zeigen wir erstmal die Abschätzung

$$\begin{aligned}
\frac{a_n}{a_{n-1}} &= \frac{\left(1+\frac{1}{n}\right)^n}{\left(1+\frac{1}{n-1}\right)^{n-1}} = \frac{\left(1+\frac{1}{n}\right)^n}{\left(1+\frac{1}{n-1}\right)^n}\left(1+\frac{1}{n-1}\right) \\
&= \frac{\left(\frac{n+1}{n}\right)^n}{\left(\frac{n}{n-1}\right)^n}\left(1+\frac{1}{n-1}\right) = \left(\frac{(n+1)(n-1)}{n^2}\right)^n\left(1+\frac{1}{n-1}\right) \\
&= \left(1-\frac{1}{n^2}\right)^n\left(1+\frac{1}{n-1}\right) \\
&\geq \left(1-\frac{1}{n}\right)\left(1+\frac{1}{n-1}\right) = 1,
\end{aligned}$$

wobei wir im Schritt von der vorletzten zur letzten Zeite die Bernoulli'sche Ungleichung benutzt haben. Insgesamt zeigt diese Abschätzung das (a_n) monoton wachsend ist.

Ähnlich argumentiert man um zu zeigen das (b_n) monoton fallend ist.

Aus Satz 4.2.4 folgt dann die Existenz von $a, b \in \mathbb{R}$ mit

$$2 \leq a = \lim a_n \leq b = \lim b_n \leq 4.$$

Weiter gilt:

$$1 \leq \frac{b}{a} = \frac{\lim b_n}{\lim a_n} = \lim\left(1+\frac{1}{n}\right) = 1$$

und damit definieren wir $a = b =: e$, die Eulersche Zahl.

2) Für $c > 1$ definieren wir für alle $n \in \mathbb{N}$ die rekursiv definierte Folge $a_1 = c$ und

$$a_{n+1} = \frac{1}{2}\left(a_n + \frac{c}{a_n}\right) = a_n + \frac{c-a_n^2}{2a_n}.$$

Behauptung: Der Grenzwert $a = \lim a_n$ existiert und es gilt: $a^2 = c$.

Um die Behauptung zu zeigen gehen wir Schrittweise vor:

i) Es gilt $a_n > 0$ fü alle $n \in \mathbb{N}$.

Dies folgt durch Induktion aus der Tatsache das $a_1 = c > 1 > 0$ und $a_{n+1} = \frac{1}{2}\left(a_n + \frac{c}{a_n}\right) > 0$, falls $a_n > 0$.

ii) Es gilt $a_n^2 \geq c$ für alle $n \in \mathbb{N}$.

Wiederum benutzen wir Induktion: $a_1^2 = c^2 > c$, da $c > 1$, und weiter

$$a_{n+1}^2 = \left(a_n + \frac{c-a_n^2}{2a_n}\right)^2 = a_n^2 + (c-a_n^2) + \left(\frac{c-a_n^2}{2a_n}\right)^2 \geq c.$$

iii) Aus i) und ii) folgt direkt: $a_n \geq 1$ für alle $n \in \mathbb{N}$

iv) Es gilt $a_{n+1} \leq a_n$ für alle $n \in \mathbb{N}$, denn aus ii) folgt

$$a_{n+1} = a_n + \frac{c - a_n^2}{2a_n} \leq a_n.$$

Aus iii) und iv) folgt die Abschätzung

$$1 \leq a_{n+1} \leq a_n \leq \ldots \leq a_1 = c.$$

Also ist (a_n) monoton fallend und nach unten beschränkt. Aus Satz 4.2.4 folgt somit die Existenz von $a = \lim a_n$.

Weiter folgt aus $a_{n+1} = \frac{1}{2}\left(a_n + \frac{c}{a_n}\right)$ die Gleichung

$$a = \frac{1}{2}\left(a + \frac{c}{a}\right),$$

oder

$$a^2 = c.$$

Bemerkung: Die für $c > 1$ rekursiv definierte Folge $a_1 = c$ und

$$a_{n+1} = \frac{1}{k}\left((k-1) \cdot a_n + \frac{c}{a_n^{k-1}}\right)$$

konvergiert für $n \to \infty$ gegen $\sqrt[k]{c}$.

Im Folgenden sei (a_k) eine Folge in $\mathbb{R}$.

Definition 4.2.3. *Sei $\Lambda \subset \mathbb{N}$ eine unendliche Teilmenge und sei $\mathbb{N} \ni n \to l(n) \in \Lambda$ eine monotone Abzählung von Λ. Dann heisst die Folge $(a_l)_{l\in\Lambda} = (a_{l(n)})_{n\in\mathbb{N}}$ Teilfolge von (a_n).*

Beispiel: i) Die Folge (a_n) mit $a_n = (-1)^n$ hat die konstanten Teilfolgen (a_{2n}) bzw. (a_{2n-1}).
ii) Die Folge $b_n = 2^n$ ist eine Teilfolge der Folge $a_n = n$.

Definition 4.2.4. *$a \in \mathbb{R}$ heisst Häufungspunkt von (a_n), falls (a_n) eine gegen a konvergente Teilfolge besitzt, d.h. falls $\Lambda \subset \mathbb{N}$ existiert mit $a_l \to a$ $(l \to \infty, l \in \Lambda)$.*

Beispiel: Die Folge $a_n = (-1)^n$ hat die Häufungspunkte ± 1.

Satz 4.2.6. *Eine Zahl $a \in \mathbb{R}$ ist genau dann Häufungspunkt der Folge (a_n), falls gilt:*

$$\forall \varepsilon > 0 \ \forall n_0 \in \mathbb{N} \ \exists l > n_0 : |a - a_l| < \varepsilon.$$

Beweis: i) Falls $a_l \to a$ mit $l \to \infty$ und $l \in \Lambda$, so existiert offenbar zu jedem $\varepsilon > 0$ ein $l_0 \in \Lambda$ mit
$$\forall l \geq l_0,\ \ l \in \Lambda: \ \ |a - a_l| < \varepsilon.$$
Ist nun zusätzlich $n_0 \in \mathbb{N}$ vorgegeben, so wähle $l \in \Lambda$ beliebig mit $l \geq \max\{n_0, l_0\}$. Es folgt $|a - a_l| < \varepsilon$, und $l \geq n_0$.

ii) Umgekehrt wähle $l(1) = 1$ und definiere $l(n)$ $(n \in \mathbb{N})$ induktiv, wie folgt. Sei $n \in \mathbb{N}$, und seien $l(1) < l(2) < \ldots < l(n)$ bereits bestimmt. Es gilt dann automatisch $l(j) \geq j$, $1 \leq j \leq n$. Zu $\varepsilon = \frac{1}{n} > 0$ und $n_0 = l(n) + 1$ existiert nach Annahme ein Index $l(n+1) := l \geq n_0 > l(n)$ mit $|a - a_l| < \frac{1}{n}$. Die so konstruierte Teilfolge $a_{l(n)}$ konvergiert offenbar gegen a. □

Limes Superior, Limes Inferior

Sei (a_n) eine beschränkte Folge, d.h. es existiert ein $M \in \mathbb{R}$, so dass für alle $n \in \mathbb{N}$ gilt:
$$|a_n| \leq M.$$
Wir definieren für alle $k \in \mathbb{N}$ die Mengen $A_k = \{a_n : n \geq k\}$. Wegen dem Vollständigkeitsaxiom existieren
$$c_k = \inf_{n \geq k} a_n = \inf A_k \leq \sup A_k = \sup_{n > k} a_n = b_k.$$

Für beliebige Mengen $A \subset B$ gilt $\inf A \geq \inf B$ und $\sup A \leq \sup B$ und daraus schliessen wir wegen $A_{k+1} \subset A_k$ für alle $k \in \mathbb{N}$:
$$-M \leq c_1 \leq c_2 \leq \ldots \leq c_k \leq c_{k+1} \leq b_{k+1} \leq b_k \leq \ldots \leq b_2 \leq b_1 \leq M.$$

Also ist die Folge (b_n) monoton fallend und nach unten beschränkt und die Folge (c_n) ist monoton wachsend und nach oben beschränkt. Nach Satz 4.2.4 konvergieren also beide Folgen und damit existieren
$$b = \lim_{k\to\infty} b_k \ \ \text{und} \ \ c = \lim_{k\to\infty} c_k.$$
Aufgrund der obigen Abschätzung folgt weiter $b \geq c$. Wir definieren
$$b = \lim_{k\to\infty} \sup_{n\geq k} a_n =: \limsup_{n\to\infty} a_n$$
und
$$c = \lim_{k\to\infty} \inf_{n\geq k} a_n =: \liminf_{n\to\infty} a_n.$$
Ist die Folge (a_n) nicht nach oben beschränkt, so definieren wir
$$\limsup_{n\to\infty} a_n := \infty$$

und falls (a_n) nicht nach unten beschränkt ist, so setzen wir

$$\liminf_{n\to\infty} a_n := -\infty.$$

Lemma 4.2.1. *b und c sind Häufungspunkte von (a_n).*

Beweis. Wir beweisen die Aussage für b.

Sei $\varepsilon > 0$ und $n_0 \in \mathbb{N}$. Aus $b_k = \sup_{n\geq k} a_n \to b$ folgt die Existenz einer Zahl $k_0 = k_0(\varepsilon)$, so dass für alle $k \geq k_0$ gilt

$$|b_k - b| < \varepsilon,$$

d.h. für alle $k \geq k_0$ gilt:

$$b - \varepsilon < b_k = \sup_{n\geq k} a_n < b + \varepsilon.$$

Ohne Einschränkung sei $k_0 \geq n_0$ (sonst ersetze n_0 durch k_0). Für $k = k_0$ folgt für alle $l \geq k_0$:

$$a_l \leq \sup_{n\geq k_0} a_n < b + \varepsilon < b + 2\varepsilon.$$

Weiter existiert wegen Blatt 3, Aufgabe 4 a) ein $l \geq k_0$ mit

$$a_l \geq \sup_{n\geq k_0} a_n - \varepsilon > b - 2\varepsilon.$$

Also existiert ein $l \geq k_0 \geq n_0$ mit:

$$|a_l - b| < 2\varepsilon.$$

Aus Satz 4.2.6 folgt damit die Behauptung. □

Eine direkte Konsequenz des Lemmas ist der Satz von Bolzano-Weierstrass.

Satz 4.2.7 (Bolzano-Weierstrass)**.** *Jede beschränkte Folge (a_n) besitzt eine konvergente Teilfolge.*

Bemerkung: i) Sei die Folge (a_n) beschränkt. Zu $\varepsilon > 0$ wähle $k_0 = k_0(\varepsilon)$, so dass für alle $k \geq k_0$ gilt

$$c - \varepsilon < c_k = \inf_{n\geq k} a_n \leq \sup_{n\geq k} a_n = b_k < b + \varepsilon.$$

Für $k = k_0$ folgt damit für alle $n \geq k_0$:

$$c - \varepsilon < a_n < b + \varepsilon,$$

d.h. für alle bis auf endlich viele $n \in \mathbb{N}$ gilt

$$a_n \in (c - \varepsilon, b + \varepsilon).$$

ii) Aus i) folgt direkt das b der grösste und c der kleinste Häufungspunkt von $a(_n)$ ist.

iii) Ist $b = c$ so konvergiert (a_n) konvergiert gegen $b = c$. iv) Gilt $a_n \to a$ so konvergiert auch jede Teilfolge von (a_n) gegen a.

Satz 4.2.8. *Sei (a_n) eine beschränkte Folge. Es sind äquivalent:*
i) $a_n \to a$

ii) $\liminf a_n = \limsup a_n = a$

iii) Jede Teilfolge von (a_n) besitzt eine Teilfolge $(a_l)_{l\in\Lambda}$ mit $a_l \to a$

Beweis: Der Beweis folgt aus obiger Bemerkung. □

Definition 4.2.5. *Eine Folge (a_n) heisst Cauchy-Folge, falls gilt:*

$$\forall \varepsilon > 0 \ \exists n_0 \in \mathbb{N}, \ \forall n, l \geq n_0 : |a_n - a_l| < \varepsilon.$$

Satz 4.2.9. *Jede konvergente Folge (a_n) ist eine Cauchy-Folge.*

Beweis: Es gelte $a_n \to a$ und sei $\varepsilon > 0$. Wir wählen $n_0 \in \mathbb{N}$ so, dass für alle $n \geq n_0$ gilt

$$|a_n - a| < \varepsilon.$$

Mit Hilfe der Dreiecksungleichung schätzen wir damit für alle $n, l \geq n_0$ ab:

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

also ist (a_n) eine Cauchy-Folge. □

Die Umkehrung des Satzes ist auch richtig.

Satz 4.2.10. *Jede Cauchy-Folge (a_n) ist konvergent.*

Beweis: Zuerst zeigen wir das jede Cauchy-Folge beschränkt ist. Dazu wählen wir für $\varepsilon = 1$ ein $n_0 \in \mathbb{N}$ so, dass für alle $n, l \geq n_0$ gilt:

$$|a_n - a_l| < 1.$$

Setzen wir $n = n_0$, so erhalten wir für alle $l \geq n_0$

$$|a_l| = |a_l - a_{n_0} + a_{n_0}| \leq |a_l - a_{n_0}| + |a_{n_0}| < |a_{n_0}| + 1.$$

Insgesamt gilt damit für alle $l \in \mathbb{N}$

$$|a_l| \leq \max\left\{|a_1|, \ldots, |a_{n_0-1}, |a_{n_0}| + 1|\right\},$$

also ist die Folge beschränkt.

Nach dem Satz von Bolzano-Weierstrass, Satz 4.2.7, existiert damit $\Lambda \subset \mathbb{N}$ und $a \in \mathbb{R}$ mit
$$a_l \to a \ (l \to \infty, \ l \in \Lambda).$$
Zu $\varepsilon > 0$ wählen wir $l_0 \in \mathbb{N}$ mit der Eigenschaft das für alle $l \geq l_0, l \in \Lambda$ gilt
$$|a_l - a| < \varepsilon.$$
Weiter wählen wir $n_0 \in \mathbb{N}$ so, dass für alle $l, n \geq n_0$ gilt
$$|a_l - a_n| < \varepsilon.$$
Für ein $l \in \Lambda$ mit $l \geq max\{l_0, n_0\}$ gilt also für alle $n \geq n_0$
$$|a_n - a| \leq |a_n - a_l| + |a_l - a| < 2\varepsilon$$
und dies zeigt die Konvergenz der Folge (a_n). □

Beispiel: Wir definieren rekursiv $a_1 = 1$ und $a_{n+1} = a_n + \frac{1}{n+1}$ für alle $n \in \mathbb{N}$. Damit erhalten wir in geschlossener Form die harmonische Reihe
$$a_n = \frac{1}{n} + \frac{1}{n-1} + \ldots + 1 = \sum_{k=1}^{n} \frac{1}{k}.$$
Für alle $n \in \mathbb{N}$ gilt die Abschätzung
$$a_{2n} - a_n = \frac{1}{n+1} + \frac{1}{n+2} + \ldots + \frac{1}{2n} \geq n\frac{1}{2n} = \frac{1}{2}$$
und dies zeigt das die Folge (a_n) keine Cauchy-Folge sein kann. Insofern kann (a_n) nach obigem Satz auch nicht konvergent sein.

Bemerkung: Für alle $\alpha > 1$ konvergiert aber die Folge (siehe Blatt 5, Aufgabe 4)
$$b_n := \sum_{k=1}^{n} \frac{1}{k^\alpha}.$$

4.3 Folgen in $\mathbb{R}$ und $\mathbb{C}$

Sei (a_n) eine Folge in $\mathbb{R}^d$ mit $a_n = (a_n^1, \ldots, a_n^d) \in \mathbb{R}^d$ und sei $a = (a^1, \ldots, a^d) \in \mathbb{R}^d$.

Satz 4.3.1. *Es sind äquivalent: i)* $a_n \to a$, *d.h.* $\|a_n - a\| = d(a_n, a) \to 0$ *und*

ii) Für alle $i \in \{1, \ldots, d\}$ *gilt* $a_n^i \to a^i$.

Beweis: Zuerst benöetigen wir eine wichtige Abschätzung: Für alle $x = (x^1, \ldots, x^d) \in$

$\mathbb{R}^d$ gilt

$$\max_{1\le i\le d}|x^i| \le \sqrt{\sum_{i=1}^{d}|x^i|^2} = ||x|| \le \sqrt{d}\max_{1\le i\le d}|x^i|.$$

Gilt jetzt $a_n \to a$, so folgt mit der obigen Abschätzung für alle $1 \le i \le d$

$$|a_n^i - a^i| \le \|a_n - a\| \to 0.$$

Also folgt ii) aus i).

Umgekehrt folgt aus $a_n^i \to a^i$ für alle $1 \le i \le d$, wiederum aus obiger Abschätzung

$$\|a_n - a\| \le \sqrt{d}\max_{1\le i\le d}|a_n^i - a^i| \to 0.$$

Insofern impliziert ii) auch i). □

Aus diesem Satz und den Sätzen 4.2.9 und 4.2.10 folgt direkt der

Satz 4.3.2. *Es sind äquivalent:*
i) (a_n) *ist Cauchy-Folge (d.h.* $\forall\varepsilon > 0\ \exists n_0 \in \mathbb{N}\ \ \forall n, l \ge n_0 : \|a_n - a_l\| < \varepsilon$*)*
ii) (a_n) *ist konvergent*

Definition 4.3.1. *Die Folge* (a_n) *ist beschränkt falls gilt:*

$$\exists c \in \mathbb{R}\ \ \forall n \in \mathbb{N} : \|a_n\| \le c.$$

Wir können auch den satz von Bolzano-Weierstrass in $\mathbb{R}^d$ zeigen.

Satz 4.3.3. *(Bolzano-Weierstrass) Jede beschränkte Folge* (a_n) *besitzt eine konvergente Teilfolge.*

Beweis: Für alle $1 \le i \le d$ und für alle $n \in \mathbb{N}$ gilt:

$$|a_n^i| \le \|a_n\| \le c.$$

Nach Satz 4.2.7 erhalten wir somit Teilfolgen $\mathbb{N} \supset \Lambda_1 \supset \ldots \supset \Lambda_d =: \Lambda$ und Häufungspunkte $a^i \in \mathbb{R}$, $1 \le i \le d$, mit

$$a_n^i \to a^i\ (n \to \infty, n \in \Lambda \subset \Lambda_i)\ 1 \le i \le d.$$

Aus Satz 4.3.1 folgt damit die Konvergenz

$$a_n \to a\ (n \to \infty, n \in \Lambda).$$

□

4.4 Reihen

In diesem Abschnitt sei (a_k) ein Folge in $\mathbb{R}$ oder $\mathbb{C}$.

Die n-te Partialsumme s_n von (a_k) ist definiert durch

$$s_n := a_1 + a_2 + a_3 + \ldots + a_n = \sum_{k=1}^{n} a_k.$$

Definition 4.4.1. *Wir sagen die Reihe* $\sum\limits_{k=1}^{\infty} a_k$ *ist konvergent, falls*

$$\lim_{n\to\infty} s_n = \lim_{n\to\infty} \sum_{k=1}^{n} a_k =: \sum_{k=1}^{\infty} a_k.$$

Beispiele:

1) Sei $q \in \mathbb{R}$ mit $|q| < 1$. Aus der geometrische Summenformel (Aufgabe 4, Blatt 2) folgt:

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

Im Abschnitt 4.2 haben gezeigt, dass $q^n \to 0$ für alle $q \in \mathbb{R}$ mit $|q| < 1$ gilt. Also konvergiert die Reihe $\sum\limits_{k=0}^{n} q^k$ und es gilt

$$\sum_{k=0}^{\infty} q^k = \frac{1}{1 - q}.$$

Die ist die sogenannte geometrische Reihe.

2) Die harmonische Reihe:

Im Abschnitt 4.2 haben wir bereits gezeigt, dass die harmonische Reihe

$$\sum_{k=1}^{\infty} \frac{1}{k}$$

divergiert, denn die Folge $s_n = \sum\limits_{k=1}^{n} \frac{1}{k}$ ist keine Cauchy-Folge.

Satz 4.4.1. *Es gelten*

i) Die Reihe $\sum\limits_{k=1}^{\infty} a_k$ *ist konvergent, falls gilt:*

$$\forall \varepsilon > 0 \; \exists n_0 \in \mathbb{N} \;\; \forall n \geq l \geq n_0 : \;\; |s_n - s_l| = \Big|\sum_{k=l+1}^{n} a_k\Big| < \varepsilon,$$

d.h. $|\sum\limits_{k=l}^{n} a_k| \to 0$ *für* $n \geq l \to \infty$.

ii) Sei $a_k \geq 0$ *für alle* $k \in \mathbb{N}$. *Dann konvergiert die Reihe* $\sum\limits_{k=1}^{\infty} a_k$ *genau dann, wenn die Folge* $s_n = \sum_{k=1}^{n} a_k$ *nach oben beschränkt ist.*

Beweis: Aus Satz 4.2.9 und Satz 4.2.10 folgt:

$$(s_n) \text{ konvergent} \quad \Leftrightarrow \quad (s_n) \text{ Cauchy-Folge.}$$

Dies beweist i).

ii) folgt aus Satz 4.2.3 und Satz 4.2.4, denn nach Vorraussetzung ist (s_n) monoton wachsend. □

Bemerkung: 1) Ist $\sum\limits_{k=1}^{\infty} a_k$ konvergent, so folgt durch Wahl von $l = n-1$ im Punkt i) des obigen Satzes, das $a_n \to 0$ gelten muss.

2) Wie man am Beispiel der harmonischen Reihe sieht, folgt aus $a_k \to 0$ nicht das die Reihe $\sum\limits_{k=1}^{\infty} a_k$ konvergieren muss.

Satz 4.4.2 (Quotientenkriterium)**.** *Sei* (a_k) *eine Folge mit* $a_k \neq 0$ *für alle* $k \in \mathbb{N}$. *Dann gilt*

i) Ist $\limsup\limits_{k\to\infty} |\frac{a_{k+1}}{a_k}| < 1$, *so ist die Reihe* $\sum\limits_{k=1}^{\infty} a_k$ *konvergent.*

ii) Ist $\liminf\limits_{k\to\infty} |\frac{a_{k+1}}{a_k}| > 1$, *so ist die Reihe* $\sum\limits_{k=1}^{\infty} a_k$ *divergent.*

Beweis: i) Wir setzen

$$q_0 := \limsup_{k\to\infty} |\frac{a_{k+1}}{a_k}| = \lim_{n\to\infty} \sup_{k\geq n} |\frac{a_{k+1}}{a_k}| < 1$$

und wir wählen $q \in \mathbb{R}$ mit $q_0 < q < 1$. Damit existiert ein $n_0 \in \mathbb{N}$ so, dass für alle $n \geq n_0$ gilt:

$$\sup_{k\geq n} |\frac{a_{k+1}}{a_k}| \leq q.$$

Indem wir $n = n_0$ wählen, erhalten wir speziell für alle $k \geq n_0$:

$$|\frac{a_{k+1}}{a_k}| \leq q.$$

Damit bekommen wir für alle $k \geq n_0$

$$|a_k| = |\frac{a_k}{a_{k-1}} \cdot \frac{a_{k-1}}{a_{k-2}} \cdot \ldots \cdot \frac{a_{n_0+1}}{a_{n_0}} \cdot a_{n_0}| \leq q^{k-n_0}\ |a_{n_0}|.$$

Für alle $n > l \geq n_0$ gilt also

$$|\sum_{k=l}^{n} a_k| \leq \sum_{k=l}^{n} |a_k| \leq \sum_{k=l}^{n} q^{k-n_0}|a_{n_0}| = |a_{n_0}|q^{l-n_0} \sum_{k=l}^{n} q^{k-l} \leq \frac{|a_{n_0}|q^{l-n_0}}{1-q}, \tag{4.1}$$

denn nach obigem Beispiel zur geometrischen Reihe gilt für alle $n > l \geq n_0$

$$\sum_{k=l}^{n} q^{k-l} \leq \sum_{k=0}^{\infty} q^k = \frac{1}{1-q}.$$

Lassen wir $n > l \to \infty$, so folgt also

$$|\sum_{k=l}^{n} a_k| \to 0,$$

und aus Satz 4.4.1 folgt damit die Konvergenz der Reihe $\sum_{k=1}^{\infty} a_k$.

ii) Nach Voraussetzung gilt

$$\liminf_{n\to\infty} |\frac{a_{k+1}}{a_k}| = \lim_{n\to\infty} \inf_{k\geq n} |\frac{a_{k+1}}{a_k}| > 1.$$

Also existiert ein $n_0 \in \mathbb{N}$ mit der Eigenschaft, dass für alle $k \geq n_0$ gilt

$$|\frac{a_{k+1}}{a_k}| \geq \inf_{l\geq n_0} |\frac{a_{l+1}}{a_l}| \geq 1.$$

Als Konsequenz erhalten wir für alle $k \geq n_0$ die Abschätzung

$$|a_k| = |\frac{a_k}{a_{k-1}} \cdot \ldots \cdot \frac{a_{n_0+1}}{a_{n_0}} \cdot a_{n_0}| \geq |a_{n_0}| > 0$$

und damit folgt aus der obigen Bemerkung das die Reihe $\sum\limits_{k=1}^{\infty} a_k$ nicht konvergieren kann. □

Beispiele:

1) Die Exponentialreihe: Für $z \in \mathbb{C}$ definieren wir

$$\exp(z) := \sum_{k=0}^{\infty} \frac{z^k}{k!}.$$

Es gilt $\exp(0) = 1$ und für alle $z \in \mathbb{C} \setminus \{0\}$ folgt mit $a_k := \frac{z^k}{k!}$:

$$\left|\frac{a_{k+1}}{a_k}\right| = \left|\frac{z^{k+1}}{(k+1)!}\frac{k!}{z^k}\right| = \frac{|z|}{k+1} \to 0 \ \text{ für } \ k \to \infty.$$

Nach dem Quotientenkriterium konvergiert die Exponentialreihe also für alle $z \in \mathbb{C}$.

Weiter erhalten wir aus der obigen Abschätzung für alle $z \in \mathbb{C}$ mit $|z| \leq \frac{k+1}{2}$ und alle $l \geq k$

$$\left|\frac{a_{l+1}}{a_l}\right| = \frac{|z|}{l+1} \leq \frac{|z|}{k+1} \leq \frac{1}{2}.$$

Aus (4.1) folgt dann mit $k = l = n_0$ und $q = \frac{1}{2}$

$$\left|\exp(z) - \sum_{j=0}^{k-1} \frac{z^j}{j!}\right| = \left|\sum_{j=k}^{\infty} \frac{z^j}{j!}\right| \leq \frac{2|z|^k}{k!}. \tag{4.2}$$

2) Frage: Für welche $z \in \mathbb{C}$ konvergiert die Reihe

$$f(z) = \sum_{k=0}^{\infty} \frac{z^k k!}{k^k}?$$

Für $z = 0$ gilt $f(0) = 1$ und für $z \neq 0$ definieren wir $a_k = \frac{z^k k!}{k^k}$. Dann gilt:

$$\left|\frac{a_{k+1}}{a_k}\right| = \left|\frac{z^{k+1}(k+1)!}{(k+1)^{k+1}} \cdot \frac{k^k}{z^k k!}\right| = |z| \cdot \frac{k^k}{(k+1)^k} = |z| \cdot \frac{1}{\left(1 + \frac{1}{k}\right)^k} \to \frac{|z|}{e}$$

für $k \to \infty$. Insgesamt folgt also aus dem Quotientenkriterium das die reihe für alle $z \in \mathbb{C}$ mit $|z| < e$ knovergiert und für alle $z \in \mathbb{C}$ mit $|z| > e$ divergiert.

Im nächsten Satz beweisen wir ein weiteres Kriterium zur Konvergenz bzw. Divergenz von reihen.

Satz 4.4.3 (Wurzelkriterium)**.** *Sei (a_k) eine Folge in $\mathbb{R}$ oder $\mathbb{C}$. Dann gilt:*

i) Ist $\limsup\limits_{k\to\infty} \sqrt[k]{|a_k|} < 1$, so ist die Reihe $\sum\limits_{k=1}^{\infty} a_k$ konvergent.

ii) Ist $\limsup\limits_{k\to\infty} \sqrt[k]{|a_k|} > 1$, so ist die Reihe $\sum\limits_{k=1}^{\infty} a_k$ divergent.

Beweis: i) Wir wählen ein $q \in \mathbb{R}$ mit $\limsup \sqrt[k]{|a_k|} < q < 1$ und bemerken das dann ein $n_0 \in \mathbb{N}$ existiert mit der Eigenschaft das für alle $k \geq n_0$ gilt

$$\sqrt[k]{|a_k|} \leq q, \quad \text{oder äquivalent} \quad |a_k| \leq q^k.$$

Damit gilt für alle $n > l \to \infty$

$$|\sum_{k=l}^{n} a_k| \leq \sum_{k=l}^{n} q^k \leq \frac{q^l}{1-q} \to 0$$

und damit konvergiert die Reihe $\sum\limits_{k=1}^{\infty} a_k$ nach Satz 4.4.1.

ii) Wir wählen ε so, dass

$$\limsup \sqrt[k]{|a_k|} =: 1 + \varepsilon > 1.$$

Also gilt für alle $k_0 \in \mathbb{N}$

$$\sup_{k \geq k_0} \sqrt[k]{|a_k|} \geq 1 + \frac{\varepsilon}{2}$$

Wegen Blatt 3, Aufgabe 4a, existiert dann für alle $k_0 \in \mathbb{N}$ ein $k \geq k_0$ so, das

$$\sqrt[k]{|a_k|} \geq 1 \quad \text{oder} \quad |a_k| \geq 1.$$

Damit konvergiert (a_k) nicht gegen Null, d.h. $\sum\limits_{k=1}^{\infty} a_k$ kann nicht konvergieren. $\square$

Potenzreihen

Sei (c_k) eine Folge in $\mathbb{R}$ oder $\mathbb{C}$. Die Reihe

$$p(z) = c_0 + c_1 z + \ldots = \sum_{k=0}^{\infty} c_k z^k$$

heisst Potenzreihe. Wir definieren $a_k := c_k z^k$ und erhalten

$$\sqrt[k]{|a_k|} = |z| \sqrt[k]{|c_k|}.$$

Aus dem Wurzelkriterium folgt damit direkt der folgende Satz.

Satz 4.4.4. *Die Potenzreihe $p(z) = \sum\limits_{k=0}^{\infty} c_k z^k$ konvergiert für alle $z \in \mathbb{C}$ mit*

$$|z| < \rho := \frac{1}{\limsup \sqrt[k]{|c_k|}} \in [0, \infty].$$

Die Reihe ist divergent für alle $z \in \mathbb{C}$ mit $|z| > \rho$.

Die Zahl ρ heisst Konvergenzradius von p.

Beispiel:
Wir betrachten die Potenzreihe $p(z) = \sum_{k=0}^{\infty} c_k z^k$ mit

$$c_k = \begin{cases} 1 & , \text{ für } k \text{ ungerade} \\ \frac{1}{k} & , \text{ für } k \text{ gerade} \end{cases}$$

Wir prüfen zuerst ob wir das Quotientenkriterium anwenden können um den Konvergenzradius von p zu bestimmen. Dazu berechnen wir

$$\frac{|c_{k+1}|}{|c_k|} = \begin{cases} k & , \ k \text{ gerade} \\ \frac{1}{k+1} & , \ k \text{ ungerade} \end{cases}$$

Also gilt

$$\limsup \frac{|c_{k+1}|}{|c_k|} = \infty$$

und

$$\liminf \frac{|c_{k+1}|}{|c_k|} = 0.$$

Das Quotientenkriterium liefert also keine Aussage über die Konvergenz bzw. Divergenz von p.

Andererseits folgt aber

$$\sqrt[k]{|c_k|} = \begin{cases} 1 & , \ k \text{ ungerade} \\ \frac{1}{\sqrt[k]{k}} & , \ k \text{ gerade} \end{cases} \quad \Rightarrow \quad \limsup \sqrt[k]{|c_k|} = 1$$

Also folgt aus dem Wurzelkriterium:

$$\sum_{k=0}^{\infty} c_k z^k \begin{cases} \text{ist konvergent für alle } z \in \mathbb{C} \text{ mit } |z| < 1 \\ \text{ist divergent für alle } z \in \mathbb{C} \text{ mit } |z| > 1 \end{cases}$$

Satz 4.4.5. *(Leibniz-Kriterium)*
Sei (a_k) eine monoton fallende reelle Nullfolge ($\Rightarrow \quad a_k \geq 0 \ \forall k \in \mathbb{N}$). Dann ist die alternierende Reihe $S = \sum\limits_{k=0}^{\infty} (-1)^k a_k$ konvergent und für alle $n \in \mathbb{N}_0$ gilt die Abschätzung

$$0 \leq \left| S - \sum_{k=0}^{n} (-1)^k a_k \right| \leq a_{n+1}. \tag{4.3}$$

Beweis: Zuerst bemerken wir, dass für jede monoton fallende relle Nullfolge (a_k) gilt: $a_k \geq 0$ für alle $k \in \mathbb{N}$.

Wir definieren $S_n = \sum\limits_{k=0}^{n} (-1)^k a_k$ und bemerken, dass für alle $n \in \mathbb{N}_0$ gilt

$$\begin{aligned} S_{2n+2} - S_{2n} &= a_{2n+2} - a_{2n+1} \leq 0, \\ S_{2n+3} - S_{2n+1} &= a_{2n+2} - a_{2n+3} \geq 0 \quad \text{und} \\ S_{2n+1} - S_{2n} &= -a_{2n+1} \leq 0, \end{aligned}$$

da die Folge a_k monoton fallend ist. Damit erhalten wir für alle $n \in \mathbb{N}$

$$S_1 \leq S_3 \leq \ldots \leq S_{2n+1} \leq S_{2n} \leq \ldots \leq S_2 \leq S_0.$$

Also ist die Folge (S_{2n}) monoton fallend und nach unten beschränkt (durch S_1) und die Folge S_{2n+1} ist monoton wachsend und nach oben beschränkt (durch s_0).

Wegen Satz 4.2.4 konvergieren also die beiden Folgen $S_{2n} \to a$ bzw. $S_{2n+1} \to b$. Weiter folgt

$$a - b = \lim_{n\to\infty}(S_{2n} - S_{2n+1}) = \lim_{n\to\infty} a_{2n+1} = 0.$$

Damit konvergiert die gesamte Folge $S_n \to a = b =: S$.

Zusätzlich erhalten wir für alle $n \in \mathbb{N}_0$

$$S_{2n+1} \leq S \leq S_{2n+2} \leq S_{2n}$$

und damit

$$\begin{aligned} 0 \leq& S - S_{2n+1} \leq S_{2n+2} - S_{2n+1} = a_{2n+2} \quad \text{und} \\ 0 \leq& S_{2n} - S \leq S_{2n} - S_{2n+1} = a_{2n+1}. \end{aligned}$$

Zusammen folgt also für alle $n \in \mathbb{N}_0$

$$0 \leq |S - S_n| \leq a_{n+1}.$$

$\square$

Beispiel: $S = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} \mp \ldots = \sum\limits_{k=0}^{\infty}(-1)^k \frac{1}{k+1}$. Nach dem obigen Satz konvergiert diese Reihe, da die Folge $a_k = \frac{1}{k+1}$ eine monoton fallende Nullfolge ist. Später zeigen wir das $S = \log 2$ gilt.

Wir können allerdings die Reihenglieder auch umordnen und erhalten

$$\begin{aligned} S &= 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} \mp \ldots \\ &= \left(1 - \frac{1}{2}\right) - \frac{1}{4} + \left(\frac{1}{3} - \frac{1}{6}\right) - \frac{1}{8} + \left(\frac{1}{5} - \frac{1}{10}\right) - \frac{1}{12} \pm \ldots \\ &= \frac{1}{2} - \frac{1}{4} + \frac{1}{6} - \frac{1}{8} + \frac{1}{10} - \frac{1}{12} \pm \ldots \\ &= \frac{1}{2}\left(1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} \pm \ldots\right) \\ &= \frac{S}{2}. \end{aligned}$$

Dies würde aber $S = 0$ implizieren. Also durften wir die Reihe insgesamt nicht umordnen.

Absolute Konvergenz

Sei (a_n) eine Folge in $\mathbb{R}$ oder $\mathbb{C}$.

Definition 4.4.2. *Die Reihe* $\sum_{k=1}^{\infty} a_k$ *konvergiert absolut, falls die Reihe* $\sum_{k=1}^{\infty} |a_k|$ *konvergiert.*

Beispiele:

1) Die Reihe $\sum_{k=0}^{\infty} (-1)^k \frac{1}{k+1}$ konvergiert nicht absolut, denn die Reihe $\sum_{k=0}^{\infty} \frac{1}{k+1}$ divergiert.

2) Die Potenzreihe $p(z) = \sum_{k=0}^{\infty} c_k z^k$ konvergiert absolut für alle $z \in \mathbb{C}$ mit $|z| < \rho$ (ρ ist der Konvergenzradius der Reihe).

Satz 4.4.6 (Umordnungssatz)**.** *Die Reihe* $\sum_{k=1}^{\infty} a_k$ *sei absolut konvergent und sei* $\varphi : \mathbb{N} \to \mathbb{N}$ *bijektiv. Dann ist auch die Reihe* $\sum_{k=1}^{\infty} a_{\varphi(k)}$ *absolut konvergent und es gilt:*

$$\sum_{k=1}^{\infty} a_{\varphi(k)} = \sum_{k=1}^{\infty} a_k.$$

Beweis: Sei $S = \sum_{k=1}^{\infty} a_k$. Nach Voraussetzung gilt für $n \to \infty$

$$\sum_{k=n+1}^{\infty} |a_k| = \sum_{k=0}^{\infty} |a_k| - \sum_{k=0}^{n} |a_k| \to 0.$$

Also existiert für alle $\varepsilon > 0$ ein $n \in \mathbb{N}$ mit

$$\sum_{k=n+1}^{\infty} |a_k| < \frac{\varepsilon}{2}.$$

Für m hinreichend gross gilt

$$\varphi^{-1}(\{1, \ldots, n\}) \subset \{1, \ldots, m\},$$

bzw. $\{1, \ldots, n\} \subset \varphi(\{1, \ldots, m\})$, und es folgt

$$\begin{aligned}
\Big| S - \sum_{j=1}^{m} a_{\varphi(j)} \Big| &\leq \Big| S - \sum_{j \in \{1,\ldots,m\}, \varphi(j) \leq n} a_{\varphi(j)} \Big| + \underbrace{\Big| \sum_{j \in \{1,\ldots,m\}, \varphi(j) > n} a_{\varphi(j)} \Big|}_{< \frac{\varepsilon}{2}} \\
&< \Big| S - \sum_{k=1}^{n} a_k \Big| + \frac{\varepsilon}{2} \\
&< \varepsilon.
\end{aligned}$$

$\square$

Satz 4.4.7. *Für alle $z \in \mathbb{C}$ gilt:*

$$\exp z = \sum_{k=0}^{\infty} \frac{z^k}{k!} = \lim_{n\to\infty} \left(1 + \frac{z}{n}\right)^n.$$

Beweis: Sei $z \in \mathbb{C}$. Zu $\varepsilon > 0$ wähle $m \in \mathbb{N}_0$ mit

$$\sum_{k=m+1}^{\infty} \frac{|z|^k}{k!} < \frac{\varepsilon}{3}.$$

Für alle $z \in \mathbb{C}$ folgt aus Aufgabe 3, Blatt 5

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

Damit gilt für $n > m$:

$$\begin{aligned}
\left|\left(1 + \frac{z}{n}\right)^n - \exp z\right| =& \left|\sum_{k=0}^{n} \frac{n(n-1)\cdot\ldots\cdot(n-k+1)}{n^k}\frac{z^k}{k!} - \sum_{k=0}^{\infty}\frac{z^k}{k!}\right| \\
\leq& \left|\sum_{k=0}^{m} \left(\frac{n(n-1)\cdot\ldots\cdot(n-k+1)}{n^k} - 1\right)\frac{z^k}{k!}\right| \\
&+ \sum_{k=m+1}^{n} \frac{n(n-1)\cdot\ldots\cdot(n-k+1)}{n^k}\frac{|z|^k}{k!} + \sum_{k=m+1}^{\infty}\frac{|z|^k}{k!} \\
\leq& \frac{2\varepsilon}{3} + \sum_{k=0}^{m}\left|\frac{n(n-1)\cdot\ldots\cdot(n-k+1)}{n^k} - 1\right|\frac{|z|^k}{k!}.
\end{aligned}$$

Für $n \to \infty$ gilt:

$$\left|\frac{n(n-1)\cdot\ldots\cdot(n-k+1)}{n^k} - 1\right| \to 0$$

und damit existiert ein $n_0 \in \mathbb{N}$ so, dass für alle $n \geq n_0$ gilt:

$$\left|\sum_{k=0}^{m}\left|\frac{n(n-1)\cdot\ldots\cdot(n-k+1)}{n^k} - 1\right|\frac{z^k}{k!}\right| < \frac{\epsilon}{3}.$$

Insgesamt folgt damit

$$\left|\left(1 + \frac{z}{n}\right)^n - \exp z\right| < \varepsilon.$$

$\square$

Satz 4.4.8 (Cauchy–Produkt)**.** *Die Reihen $\sum\limits_{k=0}^{\infty} a_k$ und $\sum\limits_{k=0}^{\infty} b_k$ seien absolut konvergent. Dann ist auch die Reihe $\sum\limits_{n=0}^{\infty} c_n$ mit $c_n = \sum\limits_{k+l=n} a_k b_l = \sum\limits_{k=0}^{n} a_k b_{n-k}$ absolut* *konvergent, und es gilt:*

$$\sum_{n=0}^{\infty} c_n = \sum_{n=0}^{\infty} \Big(\sum_{k+l=n} a_k b_\ell \Big) = \Big(\sum_{k=0}^{\infty} a_k \Big) \cdot \Big(\sum_{k=0}^{\infty} b_k \Big)$$

Beweis. Für $N \in \mathbb{N}$ definieren wir

$$A_N := \sum_{k=0}^{N} |a_k| \to \sum_{k=0}^{\infty} |a_k| =: A \text{ und } B_N := \sum_{k=0}^{N} |b_k| \to \sum_{k=0}^{\infty} |b_k| =: B,$$

wobei $A, B < \infty$. Es folgt

$$\sum_{n=0}^{N} |c_n| \leq \sum_{k+l\leq N} |a_k|\ |b_l| \leq \sum_{k,l\leq N} |a_k|\ |b_l| = A_N B_N \leq A\ B < \infty.$$

Nach Satz 4.4.1 ist $\sum\limits_{n=0}^{\infty} c_n$ absolut konvergent und wir berechnen

$$\begin{aligned}
\Big|\Big(\sum_{k=0}^{2N} a_k\Big) \Big(\sum_{l=0}^{2N} b_l\Big) - \sum_{n=0}^{2N} c_n\Big| &= \Big| \sum_{k,l\leq 2N} a_k b_l - \sum_{k+l\leq 2N} a_k b_l \Big| \\
&\leq \sum_{\substack{k,l\leq 2n \\ \max\{k,l\}>N}} |a_k|\ |b_l| \\
&= A_{2N} B_{2N} - A_N B_N \to 0
\end{aligned}$$

□

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

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

Beweis: Wir benutzen Satz 4.4.8 mit $a_k = \frac{z^k}{k!}$ und $b_k = \frac{w^k}{k!}$. Damit erhalten wir

$$\begin{aligned}
\exp z \cdot \exp w &= \left(\sum_{k=0}^{\infty} \frac{z^k}{k!}\right) \left(\sum_{k=0}^{\infty} \frac{w^k}{k!}\right) \\
&= \sum_{n=0}^{\infty} \sum_{k=0}^{n} \frac{z^k w^{n-k}}{k!(n-k)!} \\
&= \sum_{n=0}^{\infty} \frac{1}{n!} \sum_{k=0}^{n} \binom{n}{k} z^k w^{n-k} \\
&= \sum_{n=0}^{\infty} \frac{(z+w)^n}{n!} \\
&= \exp(z+w).
\end{aligned}$$

□

Lemma 4.4.1. *Für $a_k \in \mathbb{C}$ sei die Reihe $S = \sum\limits_{k=0}^{\infty} a_k$ konvergent. Dann ist auch die komplex konjugierte Reihe $\sum\limits_{k=0}^{\infty} \overline{a}_k$ konvergent mit Grenzwert $\overline{S}$.*

Beweis: Mit $S_n = \sum\limits_{k=0}^{n} a_k$ gilt $\overline{S}_n = \sum\limits_{k=0}^{n} \overline{a}_k$ und damit

$$|\overline{S} - \overline{S}_n| = |\overline{S - S_n}| = |S - S_n| \to 0$$

für $n \to \infty$. □

Satz 4.4.9. *Es gelten die folgenden Eigenschaften:*

$$\begin{aligned}
&1) \exp z \cdot \exp(-z) = 1 \quad \forall\, z \in \mathbb{C}, \quad \textit{insbesondere ist} \quad \exp z \neq 0,\\
&2) \exp x > 0 \quad \forall\, x \in \mathbb{R}\\
&3) |\exp(iy)| = 1 \quad \forall\, y \in \mathbb{R}\\
&4) |\exp(x+iy)| = \exp x \quad \forall\, x, y \in \mathbb{R}\\
&5) \exp(pz) = (\exp z)^p \quad \forall\, z \in \mathbb{C},\ p \in \mathbb{Z}
\end{aligned}$$

Beweis: zu 1) Aus dem obigen Beispiel zum Cauchy-Produkt von Reihen folgt

$$\exp(z)\exp(-z) = \exp(z - z) = \exp 0 = 1.$$

zu 2) Aus der Definition von exp folgt für alle $x \geq 0$

$$\exp x = 1 + x + \frac{x^2}{2} + \ldots \in [1, \infty)\,.$$

Für $x < 0$ erhalten wird

$$\exp x = \frac{1}{\exp(-x)} \in (0, 1).$$

zu 3) Wir berechnen

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

zu 4) Es gilt mit Hilfe von 2) und 3)

$$|\exp(x+iy)| = |\exp x|\ |\exp(iy)| = \exp x.$$

zu 5) Für alle $p \in \mathbb{N}$ schliessen wir

$$\exp(pz) = \exp(z + \ldots + z) = (\exp z)^p.$$

Für $p \in \mathbb{Z} \setminus \mathbb{N}$ folgt damit

$$1 = \exp(pz)\exp((-p)z) = \exp(pz)(\exp z)^{-p},$$

also die Behauptung. □

Winkelfunktionen

Sei $S^1 := \{z \in C : |z| = 1\}$. Die Abbildung $c : \mathbb{R} \to S^1$ mit $c(t) = \exp(it)$ ist nach Satz 4.4.9, 4), wohldefiniert.

Definition 4.4.3. *Wir definieren für alle* $t \in \mathbb{R}$

$$\cos t := \Re(\exp(it)) = \frac{\exp(it) + \exp(-it)}{2} \quad \textit{und}$$
$$\sin t := \Im(\exp(it)) = \frac{\exp(it) - \exp(-it)}{2i}$$

BEMERKUNG: Aus der Definition folgt sofort die Eulersche Formel

$$\exp(it) = \cos t + i \sin t.$$

Lemma 4.4.2. *Es gelten die folgenden Gleichungen:*

1) $\cos^2(t) + \sin^2(t) = 1 \quad \forall\, t \in \mathbb{R}$
2) $\cos(-t) = \cos t,\ \sin(-t) = -\sin t \quad \forall\, t \in \mathbb{R}$
3) $\cos(\alpha + \beta) = \cos\alpha\cos\beta - \sin\alpha\sin\beta \quad \forall\, \alpha, \beta \in \mathbb{R}$
4) $\sin(\alpha + \beta) = \sin\alpha\cos\beta + \cos\alpha\sin\beta \quad \forall\, \alpha, \beta \in \mathbb{R}.$

Beweis: zu 1) Dies folgt direkt aus

$$1 = |\exp(it)|^2 = \cos^2 t + \sin^2 t.$$

2) folgt durch Einsetzen direkt aus der Definition der beiden Funktionen.

3) und 4) Es gilt

$$\begin{aligned}\cos(\alpha+\beta) + i\sin(\alpha+\beta) &= \exp(i(\alpha+\beta)) = \exp(i\alpha)\exp(i\beta)\\ &= (\cos\alpha + i\sin\alpha)(\cos\beta + i\sin\beta)\\ &= \cos\alpha\cos\beta - \sin\alpha\sin\beta + i(\sin\alpha\cos\beta - \cos\alpha\sin\beta).\end{aligned}$$

Vergleich der Real- bzw. Imaginärteile liefert die Behauptung. □

Satz 4.4.10. *Die beiden Reihen*

$$\sum_{k=0}^{\infty} (-1)^k \frac{t^{2k}}{(2k)!}$$

und

$$\sum_{k=0}^{\infty} (-1)^k \frac{t^{2k+1}}{(2k+1)!}$$

sind für alle $t \in \mathbb{R}$ absolut konvergent und es gilt:

$$\cos t = \sum_{k=0}^{\infty} (-1)^k \frac{t^{2k}}{(2k)!} = 1 - \frac{t^2}{2} \pm \dots$$
$$\sin t = \sum_{k=0}^{\infty} (-1)^k \frac{t^{2k+1}}{(2k+1)!} = t - \frac{t^3}{6} \pm \dots$$

Beweis: Die absolute Konvergenz der beiden Reihen folgt aus der Abschätzung

$$\sum_{k=0}^{N} \frac{|t|^{2k}}{(2k)!} + \sum_{k=0}^{N} \frac{|t|^{2k+1}}{(2k+1)!} = \sum_{k=0}^{2N+1} \frac{|t|^k}{k!} \leq \sum_{k=0}^{\infty} \frac{|t|^k}{k!} = \exp(|t|) < \infty \quad \forall\, t \in \mathbb{R}, \;\; N \in \mathbb{N}\,.$$

Weiter schliessen wir

$$\begin{aligned} \cos t + i \sin t = \exp(it) &= \lim_{N \to \infty} \sum_{k=0}^{2N+1} \frac{(it)^k}{k!} \\ &= \lim_{N \to \infty} \left(\sum_{k=0}^{N} (-1)^k \frac{t^{2k}}{(2k)!} + i \sum_{k=0}^{N} (-1)^k \frac{t^{2k+1}}{(2k+1)!} \right) \\ &= \sum_{k=0}^{\infty} (-1)^k \frac{t^{2k}}{(2k)!} + i \sum_{k=0}^{\infty} (-1)^k \frac{t^{2k+1}}{(2k+1)!}. \end{aligned}$$

Erneuter Vergleich der Real- und Imaginärteile liefert die Reihenentwicklungen für sin und cos. □

Lemma 4.4.3. *Für alle $t \in \mathbb{R}$ mit $|t| \leq 1$ gilt:*

$$|\sin t - t| \leq t^2$$

Beweis: Es gilt:

$$|\sin t - t| = |\Im\left[\exp(it) - (1 + it)\right]| \leq |\exp(it) - (1 + it)|.$$

Die Abschätzung (4.2) impliziert für alle $|t| \leq 1$

$$|\exp(it) - (1 + it)| \leq \frac{2|it|^2}{2} = t^2.$$

Zusammen liefern diese beiden Abschätzungen die Behauptung. □

5 Topologie des $\mathbb{R}^n$

5.1 Offene und abgeschlossene Mengen

Definition 5.1.1. *Die offene Kugel um $x_0 \in \mathbb{R}^n$ mit Radius $r > 0$ ist die Menge*

$$B_r(x_0) := \{x \in \mathbb{R} : d(x, x_0) < r\} = \{x \in \mathbb{R}^n : \|x - x_0\| < r\}.$$

Definition 5.1.2. *Eine Menge $\Omega \subset \mathbb{R}^n$ heisst offen, falls für alle $x \in \Omega$ ein $\varepsilon > 0$ existiert mit $B_\epsilon(x) \subset \Omega$.*

Beispiel: $B_r(x_0)$ ist offen für alle $r > 0$ und $x_0 \in \mathbb{R}^n$.

Beweis: Sei $x \in B_r(x_0)$ und definiere $\varepsilon = r - d(x, x_0)$. Dann gilt $B_\varepsilon(x) \subset B_r(x_0)$, denn für alle $y \in B_\varepsilon(x)$ gilt nach der Dreiecksungleichung

$$d(y, x_0) \leq d(y, x) + d(x, x_0) < \varepsilon + d(x, x_0) = r,$$

also $y \in B_r(x_0)$. □

Satz 5.1.1. *Es gelten die Eigenschaften*

1) $\emptyset$ und $\mathbb{R}^n$ sind offen
2) Der Durchschnitt $\bigcap_{i=1}^{k} U_i$ von endlich vielen offenen Mengen U_i ist offen.
3) Die Vereinigung $\bigcup_{i=I} U_i$ von beliebig vielen offenen Mengen U_i ist offen.

Man sagt: Die offenen Mengen bilden eine Topologie.

Beweis: 1) Für die leere Menge ist nichts nachzuprüfen und für jedes $x \in \mathbb{R}^n$ gilt: $B_1(x) \subset \mathbb{R}^n$. Also ist auch $\mathbb{R}^n$ offen.
2) Sei $x \in \bigcap_{i=1}^{k} U_i$. Dann ist $x \in U_i$ für alle $1 \leq i \leq k$. Da alle U_i offen sind existieren für alle $1 \leq i \leq k$ $\varepsilon_i > 0$ mit $B_{\varepsilon_i}(x) \subset U_i$. Durch Wahl von $\varepsilon := \min_{1 \leq i \leq k} \varepsilon_i > 0$ folgt für alle $1 \leq i \leq k$

$$B_\varepsilon(x) \subset B_{\varepsilon_i}(x) \subset U_i$$

und damit

$$B_\varepsilon(x) \subset \bigcap_{i=1}^{k} U_i.$$

3) Sei $x \in \bigcup_{i\in I} U_i$. Also existiert ein $i_0 \in I$ so, dass $x \in U_{i_0}$. Da U_{i_0} offen ist, existiert ein $\varepsilon > 0$ mit $B_\varepsilon(x) \subset U_{i_0}$. Damit gilt auch

$$B_\varepsilon(x) \subset \bigcup_{i\in I} U_i.$$

□

BEMERKUNG: Es gilt $\bigcap_{i\in\mathbb{N}} B_{\frac{1}{i}}(x) = \{x\} \quad \forall\, x \in \mathbb{R}^n$, also sind Durchschnitte unendlich vieler offener Mengen nicht notwendigerweise wieder offen.

Definition 5.1.3. *Eine Teilmenge $A \subset \mathbb{R}^n$ heisst abgeschlossen, falls folgende Implikation stets gilt:*

$$x_k \in A,\ x_k \to x \ \Rightarrow\ x \in A.$$

Satz 5.1.2. *Für alle Mengen $M \subset \mathbb{R}^n$ gilt die Äquivalenz*

$$M \quad \textit{offen} \quad \Leftrightarrow \quad \mathbb{R}^n \backslash M \quad \textit{abgeschlossen.}$$

Beweis: $\underline{\Rightarrow}$: Sei M offen und sei $x_k \in \mathbb{R}^n\backslash M$ eine Folge mit $x_k \to x$.

Wäre $x \in M$, so würde wegen der Offenheit von M ein $\varepsilon > 0$ existieren mit $B_\varepsilon(x) \subset M$. Für grosse Werte von k folgt dann aus der Konvergenz von x_k gegen x auch $x_k \in B_\varepsilon(x) \subset M$ und dies ist ein Widerspruch.

$\underline{\Leftarrow}$: Wäre M nicht offen, so gäbe es ein $x \in M$ mit $B_\varepsilon(x)\backslash M \neq \emptyset$ für alle $\varepsilon > 0$. Induktiv können wir dann $x_k \in B_{\frac{1}{k}}(x)$ mit $x_k \in \mathbb{R}^\Bbbk\backslash\mathbb{M}$ finden.

Damit folgt $x_k \to x \in M$ und dies widerspricht der Tatsache das $\mathbb{R}^n \backslash M$ abgeschlossen ist. □

BEISPIEL:
$\overline{B_r(x_0)} := \{x \in \mathbb{R}^n : \quad \|x - x_0\| \leq r\}$ ist abgeschlossen.

Beweis: Wir zeigen das $\mathbb{R}^n\backslash\overline{B_r(x_0)}$ offen ist. Dazu sei $x \in \mathbb{R}^n$ mit $\|x - x_0\| > r$. Setze $\varepsilon = \|x-x_0\|-r > 0$. Für alle $y \in B_\varepsilon(x)$ gilt dann mit der Dreiecksungleichung

$$\|y - x_0\| \geq \|x - x_0\| - \|y - x\| > \|x - x_0\| - \varepsilon = r.$$

Dies impliziert $B_\varepsilon(x) \subset \mathbb{R}^n\backslash\overline{B_r(x_0)}$. □

Als direkte Konsequenz von Satz 5.1.1 und Satz 5.1.2 erhalten wird das

Lemma 5.1.1. *Es gelten:*

1. *$\emptyset$ und $\mathbb{R}^n$ sind abgeschlossen.*

2. Die Vereinigung endlich vieler abgeschlossener Mengen ist abgeschlossen.

3. Der Durchschnitt von beliebig vielen abgeschlossenen Mengen ist abgeschlossen

Relativtopologie

Auf jeder Teilmenge $M \subset \mathbb{R}^n$ existiert eine Funktion $d_M : \; M \times M \to \mathbb{R}$ mit

1. $d_M(x,y) \geq 0 \quad$ "$=$"$\Leftrightarrow x = y$

2. $d_M(x,y) = d_M(y,x)$

3. $d_M(x,y) \leq d_M(x,z) + d_M(z,y) \qquad\qquad \forall\, x,y,z \in M$

Genauer definieren wir $d_M(x,y) = d(x,y) \quad \forall x,y \in M$.
Beispiel:
Sei $M = S^{n-1} = \{x \in \mathbb{R}^n : \|x\| = 1\}$

Dann ist $d_{S^{n-1}}(x,y) := d(x,y)$ für alle $x,y \in S^{n-1} \subset \mathbb{R}^n$.

Für Kugeln bezüglich der Abstandsfunktion d_M gilt:

$$B_r^M(x_0) = \{y \in M : \; d_M(y,x_0) < r\} = B_r(x_0) \cap M.$$

Bemerkung:
$\Omega \subset M$ offen $\Leftrightarrow \forall x \in \Omega \; \exists \varepsilon > 0 : \quad B_\varepsilon^M(x) \subset \Omega$.

5.2 Das Innere, der Rand und der Abschluss einer Menge

Definition 5.2.1. *Sei $M \subset \mathbb{R}^n$, dann definieren wir:*

$\operatorname{int} M := \{x \in M : \exists \varepsilon > 0 \, mit \, B_\varepsilon(x) \subset M\}$ *(das Innere von M)*

$\overline{M} := \{x \in \mathbb{R}^n : \;\; \forall \varepsilon > 0 \; ist \, B_\varepsilon(x) \cap M \neq \emptyset\}$ *(der Abschluss von M)*

$\partial M := \{x \in \mathbb{R}^n : \;\; \forall \varepsilon > 0 \; sind \, B_\varepsilon(x) \cap M, \, B_\varepsilon(x) \cap (\mathbb{R}^n \backslash M) \neq \emptyset\}$ *(der Rand von M)*

Bemerkung:

Für alle Mengen $M \subset \mathbb{R}^n$ gilt

$$\operatorname{int} \;\; M \subset M \subset \overline{M}.$$

Weiter gilt: Ist M offen, so ist $\operatorname{int} \; M = M$ und ist M abgeschlossen, so ist $M = \overline{M}$.

Satz 5.2.1. *Für $M \subset \mathbb{R}^n$ gilt:*

1. $\operatorname{int} M$ *ist offen und es gilt die Implikation:*

$$\Omega \quad \textit{offen}, \quad \Omega \subset M \Rightarrow \Omega \subset \operatorname{int} M$$

2. $\overline{M}$ *ist abgeschlossen, und es gilt:*

$$A \quad \textit{abgeschlossen}, \quad A \supset M \Rightarrow A \supset \overline{M}$$

3. ∂M *ist abgeschlossen und es gilt:*

$$\partial M = \overline{M} \backslash \operatorname{int} M.$$

Beweis: Sei $x \in \operatorname{int} M$, also existiert ein $\varepsilon > 0$ mit $B_\varepsilon(x) \subset M$. Für $y \in B_\varepsilon(x)$ gilt $B_r(y) \subset B_\varepsilon(x) \subset M$ mit $r = \varepsilon - d(y,x)$. Damit ist

$$B_\varepsilon(x) \subset \operatorname{int} M,$$

also ist $\operatorname{int} M$ offen.

Sei weiter $\Omega \subset M$ offen. Zu $x \in \Omega$ existiert ein $\varepsilon > 0$ mit

$$B_\varepsilon(x) \subset \Omega \subset M.$$

Also ist $x \in \operatorname{int} M$ und damit $\Omega \subset \operatorname{int} M$.

Für (b) verwenden wir (a) und Satz 5.1.2. Nach Definition ist $\mathbb{R}^n \backslash \overline{M} = \operatorname{int}(\mathbb{R}^n \backslash M)$, also ist $\overline{M} = \mathbb{R}^n \backslash \operatorname{int}(\mathbb{R}^n \backslash M)$ abgeschlossen. Ist nun $A \subset \mathbb{R}^n$ eine beliebige abgeschlossene Menge mit $A \supset M$, so ist $\mathbb{R}^n \backslash A$ offen sowie $\mathbb{R}^n \backslash A \subset \mathbb{R}^n \backslash M$, also $\mathbb{R}^n \backslash A \subset \operatorname{int}(\mathbb{R}^n \backslash M)$ nach (a), und somit $A \supset \overline{M}$. Dies beweist (b).

Nach Definition gilt weiter $\partial M = \overline{M} \cap \overline{(\mathbb{R}^n \backslash M)}$, also ist ∂M abgeschlossen nach (b) und Lemma 5.1.1. Ferner ist ebenfalls nach Definition $\mathbb{R}^n \backslash \operatorname{int} M = \overline{\mathbb{R}^n \backslash M}$, folglich

$$\partial M = \overline{M} \cap (\mathbb{R}^n \backslash \operatorname{int} M) = \overline{M} \backslash \operatorname{int} M.$$

□

Satz 5.2.2. *Die Mengen* $\mathbb{Q}$ *und* $\mathbb{R}\backslash\mathbb{Q}$ *sind dicht in* $\mathbb{R}$*, d.h.* $\forall\, x, y \in \mathbb{R}$ *mit* $x < y$ *existiert ein* $q \in \mathbb{Q}$ *bzw. ein* $r \in \mathbb{R}\backslash\mathbb{Q}$ *mit*

$$x < q < y \quad \textit{bzw.} \quad x < r < y.$$

Beweis: 1) $\mathbb{Q}$

Ohne Einschränkung sei $0 \le x < y$. Ansonsten wählen wir $q = 0$ im Fall $0 < x < y$ und im Fall $x < y \le 0$ erhalten wir $0 \le (-y) < (-x)$ und wir sind zurück im Fall zweier nichtnegativer Zahlen.

Im Fall $0 \leq x < y$ existiert nach Satz 4.1.1 ein $n \in \mathbb{N}$ mit $\frac{1}{n} < y - x$.

Wir definieren die Menge

$$A := \{m \in \mathbb{N} : \frac{m}{n} > x\} \subset \mathbb{N}$$

Die Menge A ist nichtleer und nach unten beschränkt, also existiert $\inf A$. Da $A \subset \mathbb{Z}$ ist das Infimum sogar ein Minimimum, d.h.

$$\exists\ m_0 \in \mathbb{N} : \quad m_0 = \min A.$$

Es gilt:

$$x < \frac{m_0}{n} = \frac{m_0 - 1}{n} + \frac{1}{n} < x + (y - x) = y.$$

Für $q = \frac{m_0}{n}$ gilt damit

$$x < q < y.$$

2) $\mathbb{R} \backslash \mathbb{Q}$

Nach 1) existiert für alle $x < y$ ein $q \in \mathbb{Q}$ mit

$$x - \sqrt{2} < q < y - \sqrt{2}.$$

Damit folgt

$$x < q + \sqrt{2} < y$$

und $q + \sqrt{2} \in \mathbb{R} \setminus \mathbb{Q}$. $\square$

Beispiel: 1)
Sei $M = B_r(x_0) \subset \mathbb{R}^n$. Da M offen ist gilt

$$\operatorname{int} M = M = B_r(x_0).$$

Wir behaupten

$$\partial M = \{x \in \mathbb{R}^n : \ \|x - x_0\| = r\}.$$

"$\supset$":

Sei $y \in \{x \in \mathbb{R}^n : \ \|x - x_0\| = r\}$ und sei $\varepsilon > 0$. Dann existieren Punkte x_1, x_2 mit

$$x_1 \in B_\varepsilon(y) \cap B_r(x_0) \quad \text{und} \quad x_2 \in B_\varepsilon(y) \cap (\mathbb{R}^n \setminus B_r(x_0)),$$

z.B. kann man $x_1 = x_0 + (1 - \frac{\min(\varepsilon, 2r)}{2r})(y - x_0)$, $x_2 = x_0 + (1 + \frac{\varepsilon}{2r})(y - x_0)$ wählen.

Damit folgt aber aus der Definition von $\partial B_r(x_0)$:

$$y \in \partial B_r(x_0).$$

"$\subset$":

Wir zeigen:

$$\partial M \cap B_r(x_0) = \emptyset \quad \text{und} \quad \partial M \cap (\mathbb{R}^n \setminus B_r(x_0)) = \emptyset.$$

Sei zuerst $x \in \partial M \cap B_r(x_0)$. Dann ist $x \in B_r(x_0)$ und aus der Offenheit erhalten wir die Existenz eines $\varepsilon > 0$ mit

$$B_\varepsilon(x) \subset B_r(x_0).$$

Damit erhalten wir den Widerspruch $x \notin \partial B_r(x_0)$.

Da die Menge $\{x \in \mathbb{R}^n : \ \|x - x_0\| \leq r\}$ abgeschlossen ist, muss das Komplement $M' := \{x \in \mathbb{R}^n : \ \|x - x_0\| > r\}$ offen sein. Gäbe es also ein $x \in \partial M \cap M'$ so würde wieder ein $\varepsilon > 0$ mit

$$B_\varepsilon(x) \subset M'$$

existieren. Damit wäre wieder $x \notin \partial M$.

Aus Satz 5.2.1 folgt schliesslich $\partial M = \overline{M} \setminus \operatorname{int} M$ und damit

$$\overline{B_r(x_0)} = \partial B_r(x_0) \cup B_r(x_0) = \{x \in \mathbb{R}^n : \ \|x - x_0\| \leq r\}.$$

2)

i) Wir behaupten

$$\operatorname{int}(\mathbb{R} \setminus \mathbb{Q}) = \emptyset.$$

Aus der Annahme der Existenz eines $x \in \operatorname{int}(\mathbb{R} \setminus \mathbb{Q})$ folgt die Existenz eines $\varepsilon > 0$ mit

$$B_\varepsilon(x) = (x - \varepsilon, x + \varepsilon) \subset \mathbb{R} \setminus \mathbb{Q}.$$

Nach Satz 5.2.2 existiert aber für alle $\varepsilon > 0$ ein $q \in \mathbb{Q}$ mit

$$x - \varepsilon < q < x + \varepsilon$$

und dies ist ein Widerspruch.

ii) Analog zu *i)* zeigt man

$$\operatorname{int} \mathbb{Q} = \emptyset.$$

iii) Wir behaupten weiter:

$$\overline{\mathbb{Q}} = \overline{(\mathbb{R} \setminus \mathbb{Q})} = \mathbb{R}.$$

Wiederum nach Satz 5.2.2 existiert für alle $\varepsilon > 0$ ein $q \in \mathbb{Q}$ bzw. ein $r \in \mathbb{R} \setminus \mathbb{Q}$ mit

$$q, r \in B_\varepsilon(x).$$

Damit folgt die Behauptung.

iv) Aus Satz 5.2.1, *ii)* und *iii)* folgt:

$$\partial \mathbb{Q} = \overline{Q} \setminus \operatorname{int} \mathbb{Q} = \mathbb{R},$$

also kann der Rand einer Menge sehr gross sein.

Definition 5.2.2. *Eine Menge $M \subset \mathbb{R}^n$ heisst kompakt, falls gilt:*

Jede Folge $x_k \in M$ hat eine Teilfolge welche gegen ein $x \in M$ konvergiert.

Satz 5.2.3. *Eine Menge $M \subset \mathbb{R}^n$ ist genau dann kompakt, wenn M abgeschlossen und beschränkt ist.*

Beweis: "$\Rightarrow$"

Sei $x_k \in M$ eine Folge mit $x_k \to x_0$. Wir zeigen $x_0 \in M$ und damit ist M dann abgeschlossen. Da M kompakt ist, existiert eine Teilfolge (x_{k_j}) von x_k mit: $x_{k_j} \to x_1 \in M$. Aus der Eindeutigkeit des Grenzwertes folgt

$$x_1 = \lim_{j \to \infty} x_{k_j} = x_0 \in M.$$

Wäre M nicht beschränkt, so gäbe es zu $k \in \mathbb{N}$ ein $x_k \in M$ mit $\|x_k\| \geq k$. Die Folge (x_k) besitzt keine konvergente Teilfolge und damit wäre M nicht kompakt, ein Widerspruch.

"$\Leftarrow$"

Sei $x_k \in M$ eine Folge. Da M beschränkt ist, ist auch $\|x_k\|$ beschränkt. Nach dem Satz von Bolzano-Weierstrass) existiert also eine konvergente Teilfolge x_{k_j} mit $x_{k_j} \to x_0$. Da M zusätzlich abgeschlossen ist, gilt auch $x_0 \in M$. $\square$

Beispiele:

- $\overline{B_r(x_0)}$ ist kompakt.
- $[a, b]$ ist kompakt.
- $M = \mathbb{R}$ ist nicht kompakt, denn die Folge $x_k = k$ besitzt keine konvergente Teilfolge.
- $M = (0, 1)$ ist nicht kompakt, denn die Folge $x_k = \frac{1}{k+1} \in (0, 1)$ konvergiert gegen 0, aber $0 \notin (0, 1)$.

6 Stetigkeit und Grenzwerte von Funktionen

6.1 Stetigkeit

Definition 6.1.1. *Sei $x_0 \in D \subset \mathbb{R}^n$. Die Funktion $f : D \to \mathbb{R}^m$ heisst stetig im Punkt x_0, falls für jede Folge $x_k \in D$ mit $x_k \to x_0$ gilt: $f(x_k) \to f(x_0)$.*

f heisst stetig auf D, falls f stetig in allen $x_0 \in D$ ist.

Beispiele:

- $f(x) = c \in \mathbb{R}^m$ stetig auf $D = \mathbb{R}^n$
- $f(x) = x \in \mathbb{R}^n$ stetig auf $D = \mathbb{R}^n$

  Sei dazu $x_0 \in \mathbb{R}^n$ und $x_k \to x_0$ Damit folgt

$$\|f(x_k) - f(x_0)\| = \|x_k - x_0\| \to 0$$

- $f(x) = [x] := \max\{k = \mathbb{Z} : k \leq x\}$

  Für $D = [1, \infty)$ ist f ist stetig in $x_0 = 1$

  Sei dazu $x_k \in D$ eine Folge mit $x_k \to x_0 = 1$. Für k gross genug gilt dann $x_k \in [1, 2)$ und damit $f(x_k) = [x_k] = 1 = f(1)$.
  Für $D = (-\infty, 1]$ ist f nicht stetig in $x_0 = 1$, denn für $x_k = 1 - \frac{1}{k} \in D$ gilt $x_k \to 1$, aber $0 = f(x_k) \not\to 1 = f(1)$.

- $\exp : \mathbb{C} \to \mathbb{C}$ ist stetig auf $\mathbb{C}$

  i) $z_0 = 0$. Sei $z_k \in \mathbb{C}$ mit $z_k \to 0$. Für k gross genug gilt: $|z_k| \leq 1$ und wegen (4.2) gilt dann

$$|\exp(z_k) - \exp(0)| = |\exp(z_k) - 1| \leq \frac{|z_k|}{1!} \to 0.$$

  ii) z_0 beliebig. Sei $z_k \in \mathbb{C}$ eine Folge mit $z_k \to z_0$. Dann gilt

$$\exp(z_k) = \exp((z_k - z_0) + z_0) = \exp(z_k - z_0)\exp(z_0) \to \exp(z_0)$$

BEMERKUNG: $\exp : \mathbb{R} \to \mathbb{R}$ ist stetig auf $\mathbb{R}$.

Allgemein gilt: Sei $E \subset D \subset \mathbb{R}^n$ und $f : D \to \mathbb{R}^n$ stetig in $x_0 \in E \subset D$. Dann ist auch $f : E \to \mathbb{R}^n$ stetig in $x_0 \in E$, denn in der Defintion der Stetigkeit sind weniger Folgen zulässig.

- Die Funktion $f(x) = \|x\|$ ist stetig auf $D = \mathbb{R}^n$. Sei $x_k \in \mathbb{R}^n$ eine Folge mit $x_k \to x_0 \in \mathbb{R}^n$, dann gilt mit Hilfe der Dreiecksungleichung

$$|f(x_k) - f(x_0)| = \|\|x_k\| - \|x_0\|\| \leq \|x_k - x_0\| \to 0.$$

- Sei $f : \mathbb{R}^n \to \mathbb{R}^n$ linear, d.h.

  $f(\alpha x + \beta y) = \alpha f(x) + \beta f(y) \; \forall \alpha,\, \beta \in \mathbb{R},\; x, y \in \mathbb{R}^n$

  Beh. f ist stetig auf $\mathbb{R}^n$:

  i) Wir zeigen zuerst die Existenz eines $K > 0$ mit der Eigenschaft das für alle $x \in \mathbb{R}^n$ gilt

$$\|f(x)\| \leq K\|x\|.$$

  Sei dazu $x = \sum_{i=1}^{n} x_1 e_1$. Aus der Linearität von f folgt

$$\|f(x)\| = \|f(\sum_{i=1}^{n} x_i e_i)\| = \|\sum_{i=1}^{n} x_i f(e_i)\| \leq \sum_{i=1}^{n} |x_i|\|f(e_i)\| =: v \cdot w$$

  mit $v := (|x_1|, .., |x_n|) \in \mathbb{R}^n$ und $w := (\|f(e_1)\|, .., \|f(e_n)\|) \in \mathbb{R}^n$. Aus der Cauchy-Schwarz Ungleichung folgt jetzt

$$\|f(x)\| \leq \|v\|\|w\| = \sqrt{\sum_{i=1}^{n} |x_i|^2} \cdot \sqrt{\sum_{i=1}^{n} \|f(e_i)\|^2} =: K\|x\|$$

  mit $K = \sqrt{\sum_{i=1}^{n} \|f(e_i)\|^2}$.

  ii) Sei jetzt $x_0 \in \mathbb{R}^n$ und $x_k \to x_0$ eine Folge. Dann gilt

$$\|f(x_k) - f(x_0)\| = \|f(x_k - x_0)\| \leq K\|x_k - x_0\| \to 0.$$

Definition 6.1.2. *Sei $D \subset \mathbb{R}^n$ eine Menge. Eine Funktion $f : D \to \mathbb{R}^m$ heisst **Lipschitz**-stetig mit Konstante L, falls gilt:*

$$\|f(x) - f(y)\| \leq L\|x - y\| \quad \forall x, y \in D.$$

Satz 6.1.1. *Sei $D \subset \mathbb{R}^n$ und sei $f : D \to \mathbb{R}^m$ Lipschitzstetig. Dann ist die Funktion f stetig in D.*

Beweis: Sei $x_0 \in D$ und $x_k \in D$ eine Folge mit $x_k \to x_0$. Dann gilt:

$$\|f(x_k) - f(x_0)\| \leq L\|x_k - x_0\| \to 0.$$

$\square$

Satz 6.1.2. *Sei $x_0 \in D \subset \mathbb{R}^n$ und sei $f : D \to \mathbb{R}^m$ eine Funktion. Es sind äquivalent:*

i) f ist stetig in x_0

ii) Für alle $\varepsilon > 0$ existiert ein $\delta > 0$ so, dass für alle $x \in D$ mit $\|x - x_0\| < \delta$ gilt:

$$\|f(x) - f(x_0)\| < \varepsilon.$$

Äquivalent ist: $f(B_\delta(x_0)) \subset B_\varepsilon(f(x_0))$.

iii) Für alle offenen Mengen $V \subset \mathbb{R}^m$ ist $f^{-1}(V) \subset \mathbb{R}^n$ offen.

Beweis: i) $\Rightarrow$ ii)

Angenommen die Aussage ist falsch. Dann existiert ein $\varepsilon > 0$, so dass zu $\delta = \frac{1}{k}$ ein $x_k \in D$ existiert mit $\|x_k - x_0\| < \frac{1}{k}$ aber $\|f(x_k) - f(x_0)\| \geq \varepsilon$. Damit gilt dann $x_k \to x_0$, aber $f(x_k) \not\to f(x_0)$ und dies ist ein Widerspruch zur Stetigkeit von f in x_0.

ii) $\Rightarrow$ iii)

Sei $V \subset \mathbb{R}^n$ offen und sei $x_0 \in f^{-1}(V)$ (im Fall $f^{-1}(V) = \emptyset$ sind wir fertig). Dann ist $y_0 := f(x_0) \in V$ und damit existiert ein $\varepsilon > 0$ mit $B_\varepsilon(y_0) \subset V$. Nach Voraussetzung existiert ein $\delta > 0$ mit $f(B_\delta(x_0) \cap D) \subset B_\varepsilon(y_0) \subset V$. Ohne Einschränkung sei δ so klein, dass $B_\delta(x_0) \subset D$, Dann folgt $B_\delta(x_0) \subset f^{-1}(V)$ und damit ist $f^{-1}(V)$ offen.

iii) $\Rightarrow$ i)

Sei $x_k \in D$ eine Folge mit $x_k \to x_0$ und sei $\varepsilon > 0$ mit $V = B_\varepsilon(f(x_0)) \subset \mathbb{R}^m$. Nach Voraussetzung ist $U = f^{-1}(V)$ offen und $x_0 \in U$. Damit ist also auch $x_k \in U$ für grosse k. Somit ist $f(x_k) \in f(U) = V$ für k gross, und da ε beliebig war, folgt $f(x_k) \to f(x_0)$. Also ist f stetig in x_0. $\square$

Beispiel: Sei $D = \mathbb{R}$ und $f(x) = x^2$. Zu $\varepsilon > 0$ und $x_0 \in \mathbb{R}$ definieren wir $\delta = min\{1, \frac{\epsilon}{1+2|x_0|}\}$. Aus $|x - x_0| < \delta$ folgt

$$\begin{aligned}|x^2 - x_0^2| =& |x + x_0||x - x_0| \leq (|x| + |x_0|)|x - x_0| \\ \leq& (|x - x_0| + 2|x_0|)|x - x_0| \leq (1 + 2|x_0|)|x - x_0| \\ <& (1 + 2|x_0|)\delta \leq \varepsilon\end{aligned}$$

Also ist $f(x) = x^2$ stetig auf $\mathbb{R}$.

Satz 6.1.3. *Sei $D \subset \mathbb{R}^n$ und seien $f, g\ D \to \mathbb{R}$ stetig in $x_0 \in D$. Dann gilt:*

1. *$\forall \alpha, \beta \in \mathbb{R}$ ist $\alpha f + \beta f$ stetig in x_0*

2. *fg ist stetig in x_0*

3. *Ist $g(x_0) \neq 0$, dann ist $\frac{f}{g}$ stetig in x_0.*

Beweis: Sei $x_k \in D$ eine Folge mit $x_k \to x_0$, also gilt auch $f(x_k) \to f(x_0)$ und $g(x_k) \to g(x_0)$. Aus Satz 4.2.2 folgt damit

1. $\alpha f(x_k) + \beta g(x_k) \to \alpha f(x_0) + \beta g(x_0)$

2. $f(x_k)g(x_k) \to f(x_0)g(x_0)$

3. $g(x_0) \neq 0 \;\Rightarrow\; g(x_k) \neq 0$ für k gross und damit $\frac{f(x_k)}{f(x_k)} \to \frac{f(x_0)}{f(x_0)}$

$\square$

Bemerkung: Die Menge $C^0(D, \mathbb{R}^m) = \{f : D \to \mathbb{R}^m : \ f \text{ ist stetig auf D}\}$ ist ein $\mathbb{R}$-Vektorraum.

Satz 6.1.4. *Sei $D \subset \mathbb{R}^n$, $E \subset \mathbb{R}^m$ und seien $f : D \to \mathbb{R}^m$ und $g : E \to \mathbb{R}^k$ mit $f(D) \subset E$. Ist f stetig in $x_0 \in D$ und g stetig in $y_0 = f(x_0)$, dann ist $g \circ f : D \to \mathbb{R}^k$ stetig in x_0.*

Beweis: Sei $x_k \in D$ eine Folge mit $x_k \to x_0$. Dann gilt $f(x_k) \to f(x_0)$ und somit auch $g(f(x_k)) \to g(f(x_0))$. $\square$

Beispiele:

1. Sei $D \subset \mathbb{R}$ und $f, g : D \to \mathbb{R}$ seien stetig auf D.

   $\Rightarrow\ |f|\ \max\{f, g\}, \min\{f, g\} : D \to \mathbb{R}$ sind stetig auf D.

2. Sei $D \subset \mathbb{C}$ und $f : D \to \mathbb{C}$ sei stetig auf D. Dann sind auch $Re(f) = \frac{1}{2}(f + \overline{f})$, $Im(f) = \frac{1}{2i}(f + \overline{f}) : D \to \mathbb{R}$ stetig auf D.

Satz 6.1.5. *Die Funktionen* $\cos, \sin : \mathbb{R} \to \mathbb{R}$ *sind stetig auf* $\mathbb{R}(\exp(i\cdot) : \ \mathbb{R} \to \mathbb{C} \quad t \to$ *explizit*)

$[t \to itm\ z \to \exp(z)]$

Beweis: Es gilt:

$$\cos x = Re(\exp(ix)), \quad \sin x = Im(\exp(ix))$$

und damit sind die beiden Funktionen nach obigen Beispielen stetig. □

6.2 Grenzwerte von Funktionen

Definition 6.2.1. *Sei $D \subset \mathbb{R}^n$ und sei $x_0 \in D$ ein Häufungspunkt von D. Die Funktion $f : D \to \mathbb{R}^m$ hat in x_0 den Grenzwert $a \in \mathbb{R}$, falls für alle Folgen $x_k \to x_0$ gilt:*

$$f(x_k) \to a.$$

Notation: $\lim\limits_{x \to x_0} f(x) = a$.

Direkt aus der Definitiom folgt das Lemma

Lemma 6.2.1. *Sei $x_0 \in D$ ein Häufungspunkt von $D \subset \mathbb{R}^n$. Für $f : D \to \mathbb{R}^m$ sind äquivalent:*

- $\lim\limits_{x \to x_0} f(x) = f(x_0)$
- *f ist stetig in x_0.*

Definition 6.2.2. *Sei $f : (a, \infty) \to \mathbb{R}^n$ (bzw. $f : (-\infty, b) \to \mathbb{R}^m$). Es gilt $f(x) \to c$ mit $x \to \infty$ $\quad$ ($f(x) \to c$ mit $x \to -\infty$), falls für jede Folge $x_k \to \infty$ ($x_k \to -\infty$) gilt*

$$f(x_k) \to c.$$

Definition 6.2.3. *Sei $x_0 \in D$ ein Häufungspunkt von D. Die Funktion $f : D \to \mathbb{R}$ konvergiert für $x \to x_0$ gegen $+\infty$ ($-\infty$), falls für jede Folge $x_k \to x_0$ gilt:*

$$f(x_k) \to \infty \quad (-\infty).$$

Beispiele:

- $f(x) = \frac{1}{x} \to \begin{cases} 0 & \text{für } x \to +\infty \\ +\infty & \text{für } x \downarrow 0 \\ -\infty & \text{für } x \uparrow 0 \\ 0 & \text{für } x \to -\infty \end{cases}$
- $\lim\limits_{x \to \infty} (\sqrt{x+1} - \sqrt{x}) = \lim\limits_{x \to \infty} \left(\dfrac{1}{\sqrt{x+1} + \sqrt{x}} \right) = 0$
- $\lim\limits_{x \to 0} \dfrac{\sin x}{x} = 1 + \lim\limits_{x \to 0} \dfrac{\sin x - x}{x} = 1,$

  denn nach Lemma 4.4.3 gilt für alle $|x| \leq 1$: $|\sin x - x| \leq |x|^2$.

Satz 6.2.1. *Es gilt:*

1. $a, b \in \mathbb{R},\ x_0 \in \mathbb{R} \cup \{\pm\infty\},\ f(x) \to a,\ g(x) \to b$ *für* $x \to x_0$
$$\Rightarrow \begin{cases} \alpha f(x) + \beta g(x) \to \alpha a + \beta b \quad (\alpha, \beta \in \mathbb{R}) \\ f(x)g(x) \to ab \\ \frac{f(x)}{g(x)} \to \frac{a}{b}, \textit{falls } b \neq 0 \end{cases}$$

2. $f : D \to \mathbb{R},\ g : E \to \mathbb{R}$ *mit* $f(D) \subset E$ *und es gelte* $f(x) \to y_0$ *mit* $x \to x_0$ *und* g *sei stetig in* y_0. *Dann gilt* $(g \circ f)(x) \to g(y_0)$ *mit* $x \to x_0$.

3. $f(x) \to a,\ g(x) \to b$ *für* $x \to x_0$ *und ist* $f(x) \leq g(x)\ \forall\ \|x - x_0\| < \delta$, *so ist* $a \leq b$.

4. *Ist* $f > 0$ *auf* D, *so ist* $\lim\limits_{x \to x_0} f(x) = 0$ *äquivalent zu* $\lim\limits_{x \to x_0} \dfrac{1}{f(x)} = \infty$.

7 Der Zwischenwertsatz, der Satz vom Maximum und Anwendungen

7.1 Der Zwischenwertsatz

Lemma 7.1.1. *Sei $f : [a, b] \to \mathbb{R}$ stetig und es gelte $f(a) < 0 < f(b)$. Dann existiert ein $x_0 \in (a, b)$ mit $f(x_0) = 0$.*

Beweis: Wir konstruieren eine Intervallschachtelung wie folgt:

Sei $I_0 = [a, b]$ und sei $I_k = [a_k, b_k]$ schon definiert mit $f(a_k) \leq 0 \leq f(b_k)$. Wir setzen $m_k = \frac{a_k+b_k}{2}$ und definieren

$$I_{k+1} = \begin{cases} [a_k, m_k] & \text{falls } f(m_k) \geq 0 \\ [m_k, b_k] & \text{falls } f(m_k) < 0. \end{cases}$$

Damit folgt $I_{k+1} \subset I_k$ und $|I_{k+1}| = \frac{1}{2} |I_k| = \left(\frac{1}{2}\right)^{k+1} |I_0| \to 0$. Wegen Satz 4.2.5 exitiert dann genau ein $x_0 \in \mathbb{R}$ mit $x_0 \in I_k \; \forall \, k$ und $x_0 = \lim a_k = \lim b_k$.

Da f stetig ist erhalten wir

$$f(x_0) = \lim f(a_k) \leq 0 \leq \lim f(b_k) = f(x_0)$$

und damit

$$f(x_0) = 0.$$

$\square$

Satz 7.1.1 (Zwischenwertsatz). *Sei $f : [a, b] \to \mathbb{R}$ stetig mit $f(a) \leq y_0 \leq f(b)$ bzw. $f(a) \geq y_0 \geq f(b)$. Dann existiert ein $x_0 \in [a, b]$ mit*

$$f(x_0) = y_0.$$

Beweis: Wir definieren die stetige Funktion $g(x) := f(x) - y_0$ (bzw. $g(x) := y_0 - f(x)$) und erhalten

$$g(a) \leq 0 \leq g(b).$$

Nach dem obigen Lemma exitiert ein $x_0 \in [a, b]$ mit

$$g(x_0) = 0 \quad \Leftrightarrow \quad f(x_0) = y_0.$$

□

Beispiel: Das Polynom $p(x) = x^n - a$ hat für alle $n \in \mathbb{N}$ und alle $a > 0$ eine Nullstelle, denn $p(0) = -a < 0$ und $p(x) > 0$ für x sehr gross.

Lemma 7.1.2. *Sei $I \subset \mathbb{R}$ ein Intervall und sei $f : I \to \mathbb{R}$ stetig. Dann ist $f(I) \subset \mathbb{R}$ ein Intervall mit Endpunkten $\alpha = \inf\limits_{x \in I} f(x)$ und $\beta = \sup\limits_{x \in I} f(x)$*

Beweis: Zu $y \in (\alpha, \beta)$ exitieren wegen Blatt 3, Aufgabe 4a), Punkte $x_1, x_2 \in I$ mit $f(x_1) < y < f(x_2)$. Aus dem Zwischenwertsatz folgt dann die Existenz eines $x \in [x_1, x_2]$ (oder $x \in [x_2, x_1]$) mit $f(x) = y$. Damit ist

$$(\alpha, \beta) \subset f(I) \subset [\alpha, \beta]\,.$$

□

Satz 7.1.2. *Sei $f : [a, b] \to \mathbb{R}$ stetig mit $f([a, b]) \subset [a, b]$. Dann existiert ein $x \in [a, b]$ mit $f(x) = x$.*

Beweis: Wir definieren die Funktion $g : [a, b] \to \mathbb{R}$, $g(x) = f(x) - x$ und berechnen $g(a) = f(a) - a \geq 0$ und $g(b) = f(b) - b \leq 0$. Wegen dem Zwischenwertsatz existstiert also ein $x \in [a, b]$ mit $g(x) = 0 \Leftrightarrow f(x) = x$. □

7.2 Satz vom Maximum

Satz 7.2.1 (Satz vom Maximum)**.** *Sei $D \subset \mathbb{R}^n$, $D \neq \emptyset$ kompakt und sei $f : D \to \mathbb{R}$ stetig. Dann ist f beschränkt und f nimmt sein Supremum und Infimum an, d.h. es existieren $x_0, x_1 \in D$ mit*

$$f(x_0) = \inf_{x \in D} f(x), \quad f(x_1) = \sup_{x \in D} f(x).$$

Beweis: Sei $\alpha = \inf\limits_{x \in D} f(x)$. Nach Blatt 3, Aufgabe 4a, existiert für alle $\varepsilon > 0$ ein $y \in D$ mit $\alpha \leq f(y) < \alpha + \varepsilon$.

Damit können wir eine Folge $x_k \in D$ konstruieren, welche $f(x_k) \to \alpha$ erfüllt. Da D kompakt ist, existiert eine Teilfolge x_{k_j}, welche gegen ein $x_0 \in D$ konvergiert.

Da f stetig ist gilt

$$f(x_0) = \lim f(x_{k_j}) = \alpha.$$

Der Beweis für das Supremum folgt analog. □

Beispiel: Sei $D = (0, \infty)$, $f(x) = \frac{1}{1+x}$ weder $\sup f = 1$ noch $\inf f = 0$ werden angenommen, da D nicht kompakt ist.

Lemma 7.2.1. *Sei $D \subset \mathbb{R}^n$ kompakt und sei $f : D \to \mathbb{R}$ stetig mit $f(x) \neq 0$ für alle $x \in D$. Dann existiert ein $c > 0$ mit $|f(x)| \geq c > 0$ für alle $x \in D$.*

Beweis: Mit f ist auch die Funktion $|f| : D \to \mathbb{R}$ stetig. Wegen dem Satz vom Maximum existiert ein $x_0 \in D$ mit $|f(x_0)| = \inf\limits_{x \in D} |f(x)|$. Damit folgt

$$|f(x)| \geq |f(x_0)| =: c > 0 \quad \forall x \in D.$$

□

Lemma 7.2.2. *Sei $f : [a,b] \to \mathbb{R}$ stetig. Dann gilt*

$$f([a,b]) = [\alpha, \beta]$$

mit

$\alpha = \inf\limits_{x \in [a,b]} f(x)$ *und* $\beta = \sup\limits_{x \in [a,b]} f(x)$.

Beweis: Folgt direkt aus Lemma 7.1.2 und Satz 7.2.1. □

7.3 Umkehrfunktionen und Anwendungen

Definition 7.3.1. *Eine Funktion $f : (a,b) \to \mathbb{R}$ heisst monoton wachsend (fallend), falls für alle $x, y \in (a,b)$ mit $x \leq y$ gilt:*

$$f(x) \underset{(\geq)}{\leq} f(y).$$

f heisst streng monoton wachsend (fallend), falls für alle $x < y$ gilt

$$f(x) \underset{(>)}{<} f(y).$$

Bemerkung:
Ist f streng monoton wachsend oder fallend, so ist f injektiv.

Lemma 7.3.1. *Sei $f : I = (a,b) \to \mathbb{R}$ monoton wachsend. Dann gilt:*

$$\lim_{x \searrow a} f(x) = \inf\{f(x) : x \in I\}, \quad \lim_{x \nearrow b} = \sup\{f(x) : x \in I\}$$

Beweis: Mit $\alpha := \inf\{f(x) : x \in I\} \in [-\infty, \infty)$ gilt $f(x) \geq \alpha$ für alle $x \in I$. Andererseits existiert zu jedem $\alpha' > \alpha$ ein $x' \in I$ mit $f(x') < \alpha'$ und dann ist $f(x) < \alpha'$ für alle $a < x \leq x'$. Damit folgt

$$\lim_{x \searrow a} f(x) = \alpha.$$

Analog argumentiert man in der anderen Situation. $\square$

Satz 7.3.1. *Sei I Teilmenge $I \in \mathbb{R}$ ein Intervall mit Endpunkten $a < b$, und sei $f : I \to \mathbb{R}$ stetig und streng monoton wachsend. Dann gilt:*

1. *Die Umkehrfunktion $g : f(I) \to \mathbb{R}$ ist stetig und streng monoton wachsend.*
2. *$I^* = f(I)$ ist ein Intervall mit Endpunkten $\alpha = \lim_{x \searrow a} f(x) < \lim x \nearrow bf(x) = \beta$*
3. *$\lim_{y \searrow \alpha} g(y) = a$ und $\lim_{y \nearrow \beta} g(y) = b$*

Analoge Aussagen gelten wenn f streng monoton fallend ist.

Beweis: Die zweite und dritte Aussage folgen aus Lemma 7.2.2 und Lemma 7.3.1 angewandt auf f und g.
1) g ist streng monoton wachsend:
Wäre g nicht streng monoton wachsend, so gäbe es y_1, $y_2 \in f(I)$ mit $y_1 < y_2$, aber $g(y_2) \leq g(y_1)$. Damit erhalten wir aber den Widerspruch

$$y_2 = f(g(y_2)) \leq f(g(y_1)) = y_1.$$

2) g ist stetig:
Sei $x_0 \in I$ und $y_0 = f(x_0) \in f(I)$. Weiter sei $\varepsilon > 0$. Wegen Lemma 7.1.2 ist $f((x_0-\varepsilon, x_0+\varepsilon) \cap I)$ ein Intervall mit Grenzen $\alpha' = \inf\{f(x) : x \in (x_0-\varepsilon, x_0+\varepsilon) \cap I\}$ und $\beta' = \sup\{f(x) : x \in (x_0 - \varepsilon, x_0 + \varepsilon) \cap I\}$

i) Im Fall $a < x_0 < b$ gilt $\alpha' < f(x_0) = y_0 < \beta'$ und damit existiert ein $\delta > 0$ mit

$$(y_0 - \delta, y_0 + \delta) \subset f\left((x_0 - \varepsilon, x_0 + \varepsilon) \cap I\right).$$

Nach Anwendung von g folgt also

$$g\left((y_0 - \delta, y_0 + \delta)\right) \subset (g \circ f)\left((x_0 - \varepsilon, x_0 + \varepsilon) \cap I\right) = (x_0 - \varepsilon, x_0 + \varepsilon) \cap I.$$

Also gilt für alle y mit $|y - y_0| < \delta$: $|g(y) - g(y_0)| < \varepsilon$ und damit ist g stetig in y_0.

ii) Im Fall $a = x_0$ ist $\alpha = f(x_0) = y_0 < \beta'$. Also gilt für δ hinreichend klein:

$$\begin{aligned}
&\implies g((y_0 - \delta, y_0 + \delta) \cap f(I)) \subset [y_0, y_0 + \delta) \subset f((x_0 - \varepsilon, x_0 + \varepsilon) \cap I) \\
&\implies g((y_0 - \delta, y_0 + \delta) \cap f(I)) \subset (x_0 - \varepsilon, x_0 + \varepsilon)
\end{aligned}$$

Wie vorher folgt damit die Stetigkeit von g in y_0.

Analog argumentiert man für $x_0 = b$. $\square$

Bemerkung:
$a \in I \Leftrightarrow \alpha \in f(I) = I^*$ und $b \in I \Leftrightarrow \beta \in f(I) = I^*$

Beweis:
- $a \in I \Rightarrow f(a) = \inf\{f(x) : x \in I\} = \alpha \Rightarrow \alpha \in f(I)$
- Sei $\alpha \in f(I)$, d.h. $\exists x \in I$ mit $\alpha = f(x)$. Wäre $x > a$, so gäbe es ein $x' \in (a, x)$ mit $f(x') < f(x) = \alpha$ und dies ist ein Widerspruch. Damit ist$x = a \in I$.

□

Satz 7.3.2. *Die Funktion* $\exp : \mathbb{R} \to (0, \infty)$ *ist streng monoton wachsend, stetig und bijektiv. Es gilt:*

$$\lim_{x \to -\infty} \exp(x) = 0, \ \lim_{x \to \infty} \exp(x) = \infty.$$

Die Umkehrfunktion $\log : (0, \infty) \to \mathbb{R}$ *heisst natürlicher Logarithmus.* $\log$ *ist streng monoton wachsend, stetig und bijektiv und es gilt:*

$$\lim_{y \searrow 0} \log(y) = -\infty, \ \lim_{y \to \infty} \log(y) = \infty$$

Weiter ist

$$\log(1) = 0, \log(\mathrm{e}) = 1$$

und für alle $y_1, y_2 \in (0, \infty)$ *gilt*

$$\log(y_1 y_2) = \log(y_1) + \log(y_2).$$

Beweis: Aus der Definition von exp folgt für alle $x > 0$

$$\exp(x) = \sum_{k=0}^{\infty} \frac{x^k}{k!} > 1 + x.$$

Also gilt $\lim\limits_{x \to \infty} \exp(x) = \infty$. Weiter gilt $\exp(-x) = \frac{1}{\exp(x)} \to 0$ für $x \to \infty$ und damit

$$\lim_{x \to -\infty} \exp(x) = 0.$$

Für $x_2 > x_1$ gilt:

$$\frac{\exp(x_2)}{\exp(x_1)} = \exp(x_2 - x_1) > 1,$$

und dies zeigt, dass exp streng monoton wachsend ist.

Nach Satz 7.3.1 ist die Umkehrfunktion $\log : (0, \infty) \to \mathbb{R}$ also stetig und streng monoton wachsend. Weiter gilt:

$$\lim_{y \searrow 0} \log(y) = -\infty, \ \lim_{y \nearrow \infty} \log(y) = \infty$$

Ausserdem haben wir

$$\log(1) = \log(\exp(0)) = 0 \quad \text{und} \quad \log(\mathrm{e}) = \log(\exp(1)) = 1.$$

Es gilt $\exp(x_1 + x_2) = \exp(x_1)\exp(x_2)$ und nach Wahl von $x_i = \log(y_i)$, $i = 1, 2$ erhalten wir

$$\exp(\log(y_1) + \log(y_2)) = \exp(\log(y_1))\exp(\log(y_2)) = y_1 y_2,$$

oder

$$\log(y_1) + \log(y_2) = \log(\exp(\log(y_1) + \log(y_2))) = \log(y_1 y_2).$$

□

Definition 7.3.2. *Für $a > 0$ definieren wir die Funktion $f : \mathbb{R} \to \mathbb{R}$,*

$$f(x) = \exp(x \log(a)) =: a^x.$$

Insbesondere ist $\mathrm{e}^x = \exp(x)$
Für $\alpha \in \mathbb{R}$ definieren wir weiter die Funktion $g : (0, \infty) \to \mathbb{R}$,

$$g(x) = \exp(\alpha \log(x)) =: x^\alpha.$$

Lemma 7.3.2. *Die Funktion $f : \mathbb{R} \to \mathbb{R}$ habe die Eigenschaften*

- $f(x + y) = f(x)f(y) \quad \forall x, y \in \mathbb{R}$
- $\lim\limits_{x \to 0} f(x) = f(0)$

Dann ist entweder $f \equiv 0$ oder $a := f(1) > 0$ und $f(x) = a^x$ $\forall x \in \mathbb{R}$.

Beweis: Es gilt

$$f(1) = f(\frac{1}{2})^2 \geq 0$$

und wir unterscheiden jetzt zwei Fälle:

a) $f(1) = 0$. Dann folgt für alle $x \in \mathbb{R}$: $f(x) = f(x-1)f(1) = 0$.

b) $f(1) =: a > 0$

   In diesem Fall ist $f(0) = 1$, denn $f(1) = f(1 + 0) = f(1)f(0)$. Weiter ist f stetig auf ganz $\mathbb{R}$, denn für alle $x_0 \in \mathbb{R}$ und $x_k \to x_0$ gilt

$$f(x_k) = f(x_k - x_0)f(x_0) \to f(x_0).$$

   Sei jetzt $p \in \mathbb{N}$. Wir berechnen induktiv

$$f(p) = f(1)f(p-1) = f(1)^2 f(p-2) = ... = f(1)^p = a^p.$$

   Ausserdem ist $f(-p) = f(p-p)(f(p))^{-1} = a^{-p}$ für alle $p \in \mathbb{N}$ und damit gilt für alle $p \in \mathbb{Z}$

$$f(p) = a^p.$$

Sei nun $r = \dfrac{p}{q}$ mit $p \in \mathbb{Z}, q \in \mathbb{N}$. Dann erhalten wir

$$f(r)^q = f(qr) = f(p) = a^p,$$

also $f(r) = a^r$ für alle $r \in \mathbb{Q}$. Da f stetig auf ganz $\mathbb{R}$ ist, und $f(x) = a^x$ für alle $x \in \mathbb{Q}$ gilt, schliessen wir mit Blatt 7, dass für alle $x \in \mathbb{R}$ gilt:

$$f(x) = a^x.$$

$\square$

7.4 Polarkoordinaten und die Zahl π

Lemma 7.4.1. *Für alle $x \in [0, 2]$ gilt*

$$\begin{aligned} 1 - \frac{x^2}{2} &\leq \cos x \leq 1 - \frac{x^2}{2} + \frac{x^4}{24} \quad \textit{und} \\ x - \frac{x^3}{6} &\leq \sin x \leq x. \end{aligned}$$

Insbesondere gilt: $\cos 0 = 1 > 0$, $\cos 2 \leq -\dfrac{1}{3} < 0$ *und* $\sin x > 0$ *für alle* $x \in (0, 2]$.

Beweis: Aus der Reihendarstellung des Cosinus, $\cos x = \sum_{k=0}^{\infty} (-1)^k \dfrac{x^{2k}}{(2k)!}$, folgt

$$\cos x - 1 = -\frac{x^2}{2} + \frac{x^4}{24} - \frac{x^6}{720} \pm \ldots =: -\frac{x^2}{2} \sum_{k=0}^{\infty} (-1)^k a_k,$$

wobei $a_0 = 1$ und $a_k = \dfrac{x^{2k}}{3 \cdot 4 \cdot 5 \cdot ...(2k+2)}$ für alle $k \in \mathbb{N}$ ist.

Für $0 \leq x \leq 2$ ist a_k eine monoton fallende Nullfolge, denn

- $0 \leq a_k \leq \dfrac{\overbrace{2 \cdot 2 \cdot ... \cdot 2}^{2k-\text{mal}}}{3 \cdot 4 \cdot 5 \cdot (2k+2)} \leq \dfrac{1}{k+1} \to 0$ und
- $0 \leq \dfrac{a_{k+1}}{a_k} = \dfrac{x^2}{(2k+3)(2k+4)} \leq \dfrac{4}{12} = \dfrac{1}{3} < 1.$

Aus dem Leibniz-Kritierium, Satz 4.4.5, folgt damit

$$0 \le a_0 - \sum_{k=0}^{\infty} (-1)^k a_k = 1 - \sum_{k=0}^{\infty} (-1)^k a_k \le a_1 = \frac{x^2}{12}.$$

Insgesamt erhalten wir

$$0 \le -\frac{x^2}{2} \left(\sum_{k=0}^{\infty} (-1)^k a_k - 1 \right) = \cos x - 1 + \frac{x^2}{2}$$

und

$$\cos x - 1 + \frac{x^2}{2} = \left(-\frac{x^2}{2} \right) \left(\sum_{k=0}^{\infty} (-1)^k a_k - 1 \right) \le \frac{x^4}{24}.$$

Aus den beiden Ungleichungen folgt

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

Für den Sinus gehen wir ähnlich vor. Genauer setzen wir

$$\sin x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!} =: \sum_{k=0}^{\infty} (-1)^k b_k.$$

Die Folge b_k ist eine monoton fallende Nullfolge, denn

$$\frac{b_{k+1}}{b_k} = \frac{x^2}{(2k+2)(2k+3)} \le \frac{4}{6} = \frac{2}{3} < 1.$$

Aus dem Leibniz-Kriterium folgt diesmal

$$0 \le b_0 - \sin x = x - \sin x \le b_1 = \frac{x^3}{6},$$

also

$$x - \frac{x^3}{6} \le \sin x \le x.$$

□

Lemma 7.4.2. *Für alle* $\alpha, \beta \in \mathbb{R}$ *gilt*

$$\cos\alpha - \cos\beta = -2\sin\frac{\alpha+\beta}{2}\sin\frac{\alpha-\beta}{2} \quad \textit{und}$$
$$\sin\alpha - \sin\beta = 2\cos\frac{\alpha+\beta}{2}\sin\frac{\alpha-\beta}{2}.$$

Beweis: Wir beweisen nur die Aussage für den Kosinus. Aus den Additionstheoremen, Lemma 4.4.2, folgt

$$\cos\alpha = \cos\left(\frac{\alpha+\beta}{2} + \frac{\alpha-\beta}{2}\right) = \cos\frac{\alpha+\beta}{2}\cos\frac{\alpha-\beta}{2} - \sin\frac{\alpha+\beta}{2}\sin\frac{\alpha-\beta}{2}$$

und

$$\begin{aligned}\cos\beta &= \cos\left(\frac{\alpha+\beta}{2}+\frac{\beta-\alpha}{2}\right) = \cos\frac{\alpha+\beta}{2}\cos\frac{\beta-\alpha}{2} - \sin\frac{\alpha+\beta}{2}\sin\frac{\beta-\alpha}{2}\\ &= \cos\frac{\alpha+\beta}{2}\cos\frac{\beta-\alpha}{2} + \sin\frac{\alpha+\beta}{2}\sin\frac{\alpha-\beta}{2}.\end{aligned}$$

Nach Subtraktion der beiden Gleichungen erhält man die Behauptung. $\square$

Satz 7.4.1. *Es existiert eine eindeutig bestimmte Zahl $x_0 \in (0,2)$ mit $\cos x_0 = 0$. Wir definieren $\pi := 2x_0$. Auf $\left[0, \frac{\pi}{2}\right)$ ist* cos *streng monoton fallend und* sin *streng monoton wachsend. Weiter gilt*

$$\sin\frac{\pi}{2} = 1 \quad \textit{und} \quad \exp\left(i\frac{\pi}{2}\right) = i.$$

Beweis: Aus Lemma 7.4.1 folgt $\cos 0 = 1$ und $\cos 2 \leq -\frac{1}{3} < 0$. Nach dem Zwischenwertsatz, Satz 7.1.1, existiert also ein $x_0 \in (0,2)$ mit $\cos x_0 = 0$.

Für $0 \leq t_1 < t_2 \leq 2$ gilt weiter

$$0 < \frac{t_1+t_2}{2}, \ \frac{t_2-t_1}{2} \leq 2$$

und aus Lemma 7.4.1, 7.4.2 folgt

$$\cos t_2 - \cos t_1 = -2\sin\frac{t_1+t_2}{2}\sin\frac{t_2-t_1}{2} < 0.$$

Also ist cos streng monoton fallend auf $[0,2]$ und damit ist x_0 die einzige Nullstelle von cos im Intervall $[0,2]$.

Für $0 \leq t_1 < t_2 < \frac{\pi}{2}$ gilt wieder

$$0 < \frac{t_1+t_2}{2}, \frac{t_2-t_1}{2} < \frac{\pi}{2}$$

und damit

$$\sin t_2 - \sin t_1 = 2\cos\frac{t_1+t_2}{2}\sin\frac{t_2-t_1}{2} > 0.$$

Also ist sin streng monoton wachsend auf $\left[0, \frac{\pi}{2}\right)$.

Zuletzt schliessen wir aus $\cos\frac{\pi}{2} = 0$, $\sin x \geq 0$ für alle $x \in [0,2]$ und Lemma 4.4.2, dass

$$1 = \sin\frac{\pi}{2}.$$

Daraus folgt auch

$$\exp\left(i\frac{\pi}{2}\right) = \cos\frac{\pi}{2} + i\sin\frac{\pi}{2} = i.$$

$\square$

Lemma 7.4.3. *Für alle $t \in \mathbb{R}$ gilt*

$$\cos\left(t + \frac{\pi}{2}\right) = -\sin t \quad \textit{und} \quad \sin(t + \frac{\pi}{2}) = \cos t.$$

Beweis: Aus der Definition von cos und sin folgt

$$\begin{aligned} \cos(t + \frac{\pi}{2}) + i\sin(t + \frac{\pi}{2}) =& \exp\left(i(t + \frac{\pi}{2})\right) \\ =& \exp\left(i\frac{\pi}{2}\right)\exp(it) \\ =& i(\cos t + i\sin t) = -\sin t + i\cos t \end{aligned}$$

und nach Vergleich von Real- bzw. Imaginärteil folgt die Behauptung. $\square$

Aus den obigen Resultaten erhalten wir die folgende Wertetabelle

t	0		$\frac{\pi}{2}$		π		$\frac{3\pi}{2}$		2π
$\exp(it)$	1	I	i	II	-1	III	$-$i	IV	1
$\cos t$	1	$\searrow$	0	$\searrow$	-1	$\nearrow$	0	$\nearrow$	1
$\sin t$	0	$\nearrow$	1	$\searrow$	0	$\searrow$	-1	$\nearrow$	0

dabei bezeichnen die römischen Zahlen I-IV die Quadranten in der Ebene.

Bemerkung:
1) $\exp(it),\ \cos t,\ \sin t$ sind 2π-periodisch, denn

$$\exp(2\pi i) = 1 \;\Rightarrow\; \exp(i(t + 2\pi)) = \exp(2\pi i)\exp(it) = \exp(it)$$

und $\exp(i\cdot)$ kann nach obiger Wertetabelle auch keine kleinere Periode haben.

2)

$$\begin{aligned} \exp(it) =& 1 \quad \Leftrightarrow \quad t = 2\pi k \quad \text{mit} \quad k \in \mathbb{Z} \\ \cos t =& 0 \quad \Leftrightarrow \quad t = \frac{\pi}{2} + k\pi \quad \text{mit} \quad k \in \mathbb{Z} \\ \sin t =& 0 \quad \Leftrightarrow \quad t = k\pi \quad \text{mit} \quad k \in \mathbb{Z}\,. \end{aligned}$$

Satz 7.4.2. *Die Funktionen* $\cos : [0, \pi] \to [-1, 1]$ *und* $\sin : \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] \to [-1, 1]$ *sind stetig, streng monoton fallend bzw. wachsend und bijektiv. Ihre Umkehrfunktionen* $\arccos : [-1, 1] \to [0, \pi]$ *bzw.* $\arcsin : [-1, 1] \to \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$ *sind stetig, streng monoton fallend bzw. wachsend und es gilt:*

$$\begin{aligned} \lim_{x \searrow -1} \arccos x =& \pi, \quad \lim_{x \nearrow 1} \arccos x = 0 \quad \textit{und} \\ \lim_{y \searrow -1} \arcsin y =& -\frac{\pi}{2}, \quad \lim_{y \nearrow 1} \arcsin y = \frac{\pi}{2}. \end{aligned}$$

Beweis: Folgt aus Satz 7.3.1 und obiger Wertetabelle. $\square$

Satz 7.4.3. *Zu jedem $z \in \mathbb{C} \setminus \{0\}$ gibt es eindeutig bestimmte $r > 0$, $\varphi \in [0, 2\pi)$ mit $z = r \exp(i\varphi)$*

Beweis: <u>Schritt 1:</u> Die Abbildung $c : [0, 2\pi) \to S^1$, $c(t) = \exp(it)$ ist bijektiv.

Aus $\exp(it_1) = \exp(it_2)$ folgt $\exp(i(t_1 - t_2)) = 1$, also $t_1 - t_2 = 2\pi k$ mit $k \in \mathbb{Z}$. Für $t_1, t_2 \in [0, 2\pi)$ folgt damit $k = 0$, also $t_1 = t_2$. Damit ist c injektiv.

Für die Surjektivität definieren wir zu $x + iy \in S^1$ (also $x^2 + y^2 = 1$)

$$\varphi = \begin{cases} \arccos x & \text{für } y \geq 0 \\ 2\pi - \arccos x & \text{für } y < 0 \end{cases}$$

Wegen $\arccos x \in [0, \pi]$ und $\sin \varphi = \sqrt{1 - \cos^2 \varphi}$ für $\varphi \in [0, \pi]$ folgt für $y \geq 0$:

$$\begin{aligned} \exp(i\varphi) &= \cos \varphi + i \sin \varphi = \cos(\arccos x) + i \sin(\arccos x) \\ &= x + i\sqrt{1 - \cos^2(\arccos x)} = x + i\sqrt{1 - x^2} \\ &= x + iy. \end{aligned}$$

Aus $\cos(2\pi - x) = \cos x$ bzw. $\sin(2\pi - x) = -\sin x$ folgt für $y < 0$:

$$\begin{aligned} \exp(i\varphi) &= \cos(2\pi - \arccos x) + i \sin(2\pi - \arccos x) \\ &= x - i\sqrt{1 - \cos^2(\arccos x)} = x - i\sqrt{1 - x^2} = x + iy. \end{aligned}$$

Insgesamt haben wir $\exp(i\varphi) = x + iy$ gezeigt und damit ist c auch surjektiv.

<u>Schritt 2:</u> Sei $z \in \mathbb{C} \setminus \{0\}$ beliebig.
Wir setzen $r = |z|$ und nach Schritt 1 wählen wir $\varphi \in [0, 2\pi)$ mit $\frac{z}{|z|} = \exp(i\varphi)$. Damit gilt dann

$$z = r \exp(i\varphi).$$

$\square$

Satz 7.4.4. *Sei $n \in \mathbb{N}$. Die Gleichung $z^n = 1$ hat genau n Lösungen. Sie sind durch die n-ten Einheitswurzeln*

$$\omega_k = \exp(i\frac{2\pi k}{n}); \quad k = 0, 1, \ldots, n-1$$

gegeben.

Beweis: Es gilt $z^n = 1$ und damit $|z|^n = 1$, also $r = |z| = 1$.
Sei also $z = \exp(i\varphi)$ mit $\varphi \in [0, 2\pi)$. Aus $z^n = 1$ folgt $\exp(in\varphi) = 1$. Also muss $n\varphi = 2\pi k$ für ein $k \in \mathbb{Z}$ sein. Dies impliziert $\varphi = \frac{2\pi k}{n}$. Die ω_k sind also paarweise verschiedene Lösungen von $z^n = 1$.

Da $p(z) := z^n - 1$ ein Polynom n-ten Grades ist, kann p nach Lemma 3.3.2 höchstens n-Nullstellen haben. Damit haben wir mit den ω_k schon alle möglichen Nullstellen gefunden. $\square$

8 Differentialrechnung

In diesem Kapitel ist $I \subset \mathbb{R}$ stets ein offenes Intervall.

8.1 Differentialrechung

Definition 8.1.1. *Die Funktion $f : I \to \mathbb{R}^n$ hat in $x_0 \in I$ die Ableitung $a \in \mathbb{R}^n$ (Notation: $f'(x_0) = a$), falls gilt:*

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h} = a$$

Wir nennen f differenzierbar in x_0, falls f in x_0 eine Ableitung $a \in \mathbb{R}^n$ besitzt.

Lemma 8.1.1. *Die Funktion $f = (f_1, \ldots, f_n) : I \to \mathbb{R}^n$ ist genau dann in x_0 differenzierbar, wenn für alle $1 \leq i \leq n$ die Funktionen $f_i : I \subset \mathbb{R}$ differenzierbar in x_0 sind.*

Ausserdem haben wir

$$f'(x_0) = (f_1'(x_0), \ldots, f_n'(x_0)).$$

Definition 8.1.2. *Die Funktion $f : I \to \mathbb{R}^n$ heisst differenzierbar, falls f in jedem Punkt $x_0 \in I$ differenzierbar ist. Die Funktion $f' : I \to \mathbb{R}^n$*

$$x \to f'(x) = \lim_{h \to 0} \frac{f(x + h) - f(x)}{h}$$

heisst Ableitung von f.

Beispiele:

1) $f(x) = c \in \mathbb{R}^n \quad \forall\, x \in I \Longrightarrow \dfrac{f(x) - f(x_0)}{x - x_0} = 0 \quad \forall\, x \neq x_0$
   $\Longrightarrow \; f'(x_0) = 0 \quad \forall\, x_0 \in I$

2) $f : \mathbb{R} \to \mathbb{R}$, $f(x) = |x|$ ist nicht differenzierbar in $x_0 = 0$, denn der Grenzwert

$$\lim_{s \to 0} \frac{f(x) - f(0)}{x - 0} = \lim_{x \to 0} \frac{|x|}{x} = \lim_{x \to 0} \frac{x}{|x|} \quad \text{existiert nicht}$$

3) $f : \mathbb{R}_0^+ \to \mathbb{R}$ $f(x) = \sqrt{x}$. Sei $x_0 \geq 0$.

$$\implies \frac{f(x_0+h) - f(x_0)}{h} = \frac{\sqrt{x_0+h} - \sqrt{x_0}}{h} = \frac{1}{\sqrt{x_0+h} + \sqrt{x_0}}$$

$$\implies \lim_{h=0} \frac{f(x_0+h) - f(x_0)}{h} \begin{cases} = \dfrac{1}{2\sqrt{x_0}} & \text{für } x_0 \neq 0 \\ \text{existiert nicht für } x_0 = 0 \end{cases}$$

$\implies$ $f(x) = \sqrt{x}$ nicht differenzierbar in $x_0 = 0$, aber differenzierbar für alle $x_0 \in \mathbb{R}^+$ mit Ableitung

$$f'(x) = \frac{1}{2\sqrt{x}}.$$

Satz 8.1.1. *Sei $a > 0$. Die Funktion $f : \mathbb{R} \to \mathbb{R}$, $f(x) = a^x$ ist differenzierbar mit Ableitung $f'(x) = (\log a)a^x$. (Spezialfall: $(e^x)' = e^x$)*

Beweis. Für $x_0 \in \mathbb{R}$ gilt mit Hilfe von Blatt 7, Aufgabe 1

$$\lim_{x \to x_0} \frac{a^x - a^{x_0}}{x - x_0} = a^{x_0} \lim_{x \to x_0} \frac{a^{x-x_0} - 1}{x - x_0} = a^{x_0} \lim_{y \to 0} \frac{a^y - 1}{y} = a^{x_0}(\log a).$$

□

Satz 8.1.2. *Sei $n \in \mathbb{N}$. Die Funktion $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^n$ ist differenzierbar mit $f'(x) = nx^{n-1}$.*

Beweis. Sei $x \in \mathbb{R}$. Wiederum mit Hilfe von Blatt 7, Aufgabe 1, folgt

$$\lim_{h \to 0} \frac{f(x+h) - f(x)}{h} = \lim_{h \to 0} \frac{(x+h)^n - x^n}{h} = nx^{n-1}.$$

□

Satz 8.1.3. *Die Funktion $f : I \to \mathbb{R}^n$ hat in $x_0 \in I$ genau dann die Ableitung $a \in \mathbb{R}^n$, falls für $r(h) := f(x_0+h) - (f(x_0) + ah)$ gilt:*

$$\lim_{h \to 0} \frac{r(h)}{h} = 0.$$

Beweis. Für $h \neq 0$ gilt:

$$\frac{r(h)}{h} = \frac{f(x_0+h) - f(x_0) - ah}{h} = \frac{f(x_0+h) - f(x_0)}{h} - a$$

□

Satz 8.1.4. *Ist $f : I \to \mathbb{R}^n$ differenzierbar in $x_0 \in I$, so ist f auch stetig in x_0.*

Beweis. Wir müssen zeigen, dass $\lim\limits_{h\to 0} f(x_0+h) = f(x_0)$ gilt.

Für $h \neq 0$ gilt mit Satz 8.1.3

$$f(x_0+h) = f(x_0) + ah + h\,\frac{r(h)}{h} \to f(x_0) \quad \text{mit } h \to 0.$$

$\square$

Satz 8.1.5. *Seien $f, g : I \to \mathbb{R}$ differenzierbar in $x_0 \in I$. Dann sind auch die Funktionen $\alpha f + \beta g (\alpha, \beta \in \mathbb{R})$, fg und $\frac{f}{g}$ (im Fall $g(x_0) \neq 0$) in x_0 differenzierbar mit den Ableitungen*

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

2) $(fg)'(x_0) = f'(x_0)g(x_0) + f(x_0)g'(x_0)$ *„Produktregel“*

3) $\left(\frac{f}{g}\right)'(x_0) = \dfrac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g^2(x_0)}$ *„Quotientenregel“*

Beweis. 1)

$$\begin{aligned}\frac{\alpha f(x) + \beta g(x) - \alpha f(x_0) - \beta g(x_0)}{x - x_0} &= \alpha\frac{f(x) - f(x_0)}{x - x_0} + \beta\frac{g(x) - g(x_0)}{x - x_0} \\ &\to \alpha f'(x_0) + \beta g'(x_0) \quad \text{für } \ x \to x_0.\end{aligned}$$

2)

$$\begin{aligned}\frac{f(x)g(x) - f(x_0)g(x_0)}{x - x_0} &= \frac{f(x) - f(x_0)}{x - x_0}\, g(x) + f(x_0)\,\frac{g(x) - g(x_0)}{x - x_0} \\ &\to \ f'(x_0)g(x_0) + f(x_0)g'(x_0) \\ &\quad \text{da g stetig in } x_0 \text{ nach Satz 8.1.4}\end{aligned}$$

3) a) $f(x) = 1 \ \forall\, x \in I$

$$\frac{1}{x - x_0}\left(\frac{1}{g(x)} - \frac{1}{g(x_0)}\right) = -\frac{1}{g(x)g(x_0)}\,\frac{g(x) - g(x_0)}{x - x_0} \to \frac{-g'(x_0)}{g^2(x_0)}$$

b) f beliebig

$$\implies \left(\frac{f}{g}\right)'(x_0) = \left(f \cdot \frac{1}{g}\right)'(x_0) \overset{2+3a)}{=} f'(x_0)\frac{1}{g(x_0)} - f(x_0)\frac{g'(x_0)}{g^2(x_0)}.$$

$\square$

Beispiele:

1) Sei $f(x) = x^{-n} = \frac{1}{x^n}$ für $n \in \mathbb{N}$ und $x > 0$. Aus der Quotientenregel folgt

$$f'(x) = -\frac{nx^{n-1}}{x^{2n}} = -nx^{-n-1}.$$

2) Für $f(t) = \exp(it)$ gilt

$$f'(t) = \lim_{h \to 0} \exp(it) \frac{\exp(ih) - 1}{h} = i \exp(it).$$

Also folgt

$$(\cos + i \sin)'(t) = i(\cos + i \sin)(t) = -\sin t + i \cos t$$

und damit

$$\cos'(t) = -\sin t \quad \text{und} \quad \sin'(t) = \cos t.$$

Satz 8.1.6 (Kettenregel)**.** *Seien $f : I \to \mathbb{R}$, $g : J \to \mathbb{R}^n$ mit $f(I) \subset J$. Ist f in $x_0 \in I$ differenzierbar und g in $y_0 = f(x_0) \in J$ differenzierbar, so ist $g \circ f : I \to \mathbb{R}$ in x_0 differenzierbar mit Ableitung*

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

Beweis. Betrachte für $x \neq x_0$

$$\frac{g(f(x)) - g(f(x_0))}{x - x_0} = \begin{cases} \dfrac{g(f(x)) - g(f(x_0))}{f(x) - f(x_0)} \cdot \dfrac{f(x) - f(x_0)}{x - x_0} & \text{für} f(x) \neq f(x_0) \\ 0 & \text{für} f(x) = f(x_0) \end{cases}$$

Weiter definieren wir

$$a : J \to \mathbb{R}, a(y) = \begin{cases} \dfrac{g(y) - g(f(x_0))}{y - f(x_0)} & \text{für } y \neq f(x_0) \\ g'(f(x_0)) & \text{für } y = f(x_0) \end{cases}$$

a ist stetig in $f(x_0)$.

Für $x \to x_0$, $x \neq x_0$, folgt $f(x) \to f(x_0)$ und damit

$$\begin{aligned} \frac{g(f(x)) - g(f(x_0))}{x - x_0} &= a(f(x)) \frac{f(x) - f(x_0)}{x - x_0} \to a(f(x_0))f'(x_0) \\ &= g'(f(x_0))f'(x_0). \end{aligned}$$

□

Satz 8.1.7 (Ableitung der Umkehrfunktion)**.** *Sei $f : I \to \mathbb{R}$ streng monoton und stetig auf I und differenzierbar in $x_0 \in I$ mit $f'(x_0) \neq 0$. Dann ist $I^* = f(I)$ ein offenes Intervall und die Umkehrfunktion $g : I^* \to \mathbb{R}$ ist differenzierbar in $y_0 = f(x_0)$ mit Ableitung*

$$g'(y_0) = \frac{1}{f'(g(y_0))}.$$

Beweis. Nach Satz 7.3.1 und der anschliessenden Bemerkung ist $I^\star$ ein offenes Intervall und $g : I^\star \to \mathbb{R}$ ist streng monoton und stetig, also gilt $g(y) \to g(y_0) = x_0$ mit $y \to y_0$. Für $y \neq y_0$ folgt:

$$\frac{g(y) - g(y_0)}{y - y_0} = \frac{g(y) - g(y_0)}{f(g(y)) - f(g(y_0))} = \frac{1}{\frac{f(g(y)) - f(g(y_0))}{g(y) - g(y_0)}} \to \frac{1}{f'(x_0)}.$$

□

Satz 8.1.8. *Die Funktion* $\log : (0, \infty) \to \mathbb{R}$ *ist differenzierbar mit Ableitung*

$$(\log)'(y) = \frac{1}{y}.$$

Beweis. Aus Satz 8.1.1 und Satz 8.1.7 folgt:

$$\log'(y) = \frac{1}{\exp'(\log y)} = \frac{1}{\exp(\log y)} = \frac{1}{y}.$$

□

BEISPIELE:

1) $f : (0, \infty) \to \mathbb{R}$, $f(x) = x^\alpha = exp(\alpha \log(x))$ mit $\alpha \in \mathbb{R}$
$\implies f'(x) = (\exp(\alpha \log x))' = \exp(\alpha \log x) \cdot (\alpha \log x)' = x^\alpha \cdot \dfrac{\alpha}{x} = \alpha x^{\alpha - 1}.$

2) $\arccos : (-1, 1) \to (0, \pi)$
$\arcsin : (-1, 1) \to \left(-\dfrac{\pi}{2}, \dfrac{\pi}{2}\right)$

Für $x \in (-1, 1)$ gilt:

$$\begin{aligned}
\arccos'(x) =& \frac{1}{\cos'(\arccos x)} = -\frac{1}{\sin(\arccos x)} = -\frac{1}{\sqrt{1 - \cos^2(\arccos)}} \\
=& -\frac{1}{\sqrt{1 - x^2}} \\
\arcsin'(x) =& \frac{1}{\sin'(\arcsin x)} = \frac{1}{\cos(\arcsin x)} = \frac{1}{\sqrt{1 - x^2}}
\end{aligned}$$

8.2 Der Mittelwertsatz und Anwendungen

Definition 8.2.1. *Die Funktion* $f : (a, b) \to \mathbb{R}$ *hat in* $x_0 \in (a, b)$ *ein lokales Minimum (Maximum), falls es ein* $\delta > 0$ *gibt, so dass gilt:*

$$f(x_0) \overset{(\geq)}{\leq} f(x) \qquad \forall\, x \in (x_0 - \delta, x_0 + \delta).$$

Ist sogar $f(x_0) \overset{(>)}{<} f(x)$ für $x \in (x_0-\delta, x_0+\delta)\setminus\{x_0\}$, so heisst das lokale Minimum (Maximum) isoliert.

Satz 8.2.1. *Die Funktion $f : (a,b) \to \mathbb{R}$ habe in $x_0 \in (a,b)$ ein lokales Extremum. Ist f in x_0 differenzierbar, so gilt $f'(x_0) = 0$.*

Beweis. Sei z. B. x_0 ein lokales Minimum von f. Dann gibt es ein $\delta > 0$ mit $f(x) \geq f(x_0) \quad \forall\, x \in (x_0-\delta, x_0+\delta)$, also gilt

$$\frac{f(x)-f(x_0)}{x-x_0} \begin{cases} \geq 0 & \text{für } x \in (x_0, x_0+\delta) \\ \leq 0 & \text{für } x \in (x_0-\delta, x_0) \end{cases}$$

Mit $x \searrow x_0$ folgt $f'(x_0) \geq 0$, mit $x \nearrow x_0$ folgt $f'(x_0) \leq 0$, also insgesamt $f'(x_0) = 0$. □

Beispiel: $f : \mathbb{R} \to \mathbb{R}$, $f(x) = x^3$ erfüllt $f'(0) = 0$, aber für $x_0 = 0$ liegt kein lokales Extremum vor.

Satz 8.2.2 (Mittelwertsatz der Differentialrechung)**.** *Sei $f : [a,b] \to \mathbb{R}$ stetig und differenzierbar auf (a,b). Dann gibt es ein $\xi \in (a,b)$ mit*

$$f'(\xi) = \frac{f(b)-f(a)}{b-a}$$

Beweis. **Schritt 1**: (Satz von Rolle) $f(a) = f(b) = 0$

Aus Satz 7.2.1 folgt die Existenz von $\xi_1, \xi_2 \in [a,b]$ mit

$$f(\xi_1) = \inf_{x\in[a,b]} f(x), \;\; f(\xi_2) = \sup_{x\in[a,b]} f(x)$$

Ist $\xi_1 \in (a,b)$ so folgt aus Satz 8.2.1

$$f'(\xi_1) = 0 = \frac{f(b)-f(a)}{b-a}.$$

Ebenso folgt dies falls $\xi_2 \in (a,b)$.

Für den verbleibenden Fall $\xi_1, \xi_2 \in \{a,b\}$ folgt $\inf f = \sup f = 0$, bzw.

$$f(x) = 0 \qquad \forall\, x \in [a,b].$$

Damit ist auch $f'(x) = 0 \qquad \forall\, x \in (a,b)$

Schritt 2: $f(a), f(b)$ beliebig

Definiere $h : [a,b] \to \mathbb{R}$ durch Abziehen der Sekante:

$$h(x) = f(x) - \Big(f(a) + \frac{f(b)-f(a)}{b-a}(x-a)\Big)$$

Es gilt $h(a) = 0 = h(b)$. Nach Schritt 1 existiert ein $\xi \in (a, b)$ mit

$$0 = h'(\xi) = f'(\xi) - \frac{f(b) - f(a)}{b - a}.$$

□

Satz 8.2.3 (Monotoniekriterium)**.** *Sei $f : [a, b] \to \mathbb{R}$ stetig und auf (a, b) differenzierbar. Dann gilt:*

$$\begin{array}{llll} f'(x) \geq 0 & \forall\, x \in (a,b) & \Longrightarrow & f \quad \textit{ist wachsend auf } [a,b] \\ f'(x) \leq 0 & \forall\, x \in (a,b) & \Longrightarrow & f \quad \textit{ist fallend auf } [a,b] \\ f'(x) = 0 & \forall\, x \in (a,b) & \Longrightarrow & f \quad \textit{ist konstant.} \end{array}$$

Bei strikter Ungleichung folgt strenge Monotonie.

Beweis. Sei $a \leq x_1 < x_2 \leq b$. Nach Satz 8.2.2 existiert ein $\xi \in (x_1, x_2)$, so dass gilt:

$$f(x_2) - f(x_1) = f'(\xi) \underbrace{(x_2 - x_1)}_{>0} \begin{cases} \geq 0, & \text{wenn } f'(\xi) \geq 0 \\ > 0, & \text{wenn } f'(\xi) > 0 \\ \leq 0, & \text{wenn } f'(\xi) \leq 0 \\ < 0, & \text{wenn } f'(\xi) < 0 \\ = 0, & \text{wenn } f'(\xi) = 0 \end{cases}$$

□

Lemma 8.2.1. *Sei $a \in \mathbb{R}$. Jede differenzierbare Funktion $y : \mathbb{R} \to \mathbb{R}$, welche die Differentialgleichung*

$$y'(x) = ay(x) \quad \forall x \in \mathbb{R}$$

erfüllt, hat die Gestalt $y(x) = Ce^{ax}$, wobei C eine Konstante ist.

Insbesondere ist $y : \mathbb{R} \to \mathbb{R}$, $y(x) = e^x$ die einzige differenzierbare Lösung der Gleichung

$$y'(x) = y(x) \qquad \textit{mit } y(0) = 1.$$

Beweis. Die Funktion $g(x) = y(x)e^{-ax}$ erfüllt

$$g'(x) = (y'(x) - ay(x))e^{-ax} = 0 \qquad \forall\, x \in \mathbb{R}$$

Nach Satz 8.2.3 existiert damit ein $C \in \mathbb{R}$ mit $g(x) = C$ für alle $x \in \mathbb{R}$. □

Satz 8.2.4 (Schrankensatz)**.** *Ist $f : [a, b] \to \mathbb{R}$ stetig und auf (a, b) differenzierbar, so gilt für $a \leq x_1 < x_2 \leq b$:*

$$f'(x) \geq m \quad \forall\, x \in (a,b) \implies \frac{f(x_2) - f(x_1)}{x_2 - x_1} \geq m,$$
$$f'(x) \leq M \quad \forall\, x \in (a,b) \implies \frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq M.$$

Beweis. Wir zeigen die erste Behauptung. Die zweite folgt analog.

Mit $g(x) = mx$ gilt $(f-g)' = f' - g' \geq m - m = 0$ für alle $x \in (a,b)$. Nach Satz 8.2.3 ist $(f-g)$ wachsend auf $[a,b]$, d. h.

$$f(x_2) - mx_2 \geq f(x_1) - mx_1 \qquad \forall\, a \leq x_1 < x_2 \leq b$$

□

BEMERKUNG: Für Funktionen $f : [a,b] \to \mathbb{R}^n$ gilt der Mittelwertsatz nicht. Sei z. B. $c : [0, 2\pi] \to \mathbb{C}$ mit $c(t) = e^{it}$. Dann gilt

$$\frac{c(2\pi) - c(0)}{2\pi - 0} = 0, \quad \text{aber } |c'(t)| = |ie^{it}| = 1 \quad \forall\, t \in [0, 2\pi].$$

Definition 8.2.2. *Die k–te Ableitung ($k \in \mathbb{N}_0$) von $f : (a,b) \to \mathbb{R}^n$ in x_0 ist induktiv definiert durch*

$$f^{(k)}(x_0) = (f^{(k-1)})'(x_0) \quad \textit{und} \quad f^{(0)}(x_0) = f(x_0).$$

Damit $f^{(k)}(x_0)$ definiert ist, müssen also alle Ableitungen bis zur Ordnung $k-1$ in einer Umgebung von x_0 definiert sein, und $f^{(k-1)}$ muss in x_0 differenzierbar sein.

Definition 8.2.3. *Sei $I \subset \mathbb{R}$ ein offenes Intervall und $k \in \mathbb{N}_0$. Wir bezeichnen mit $C^k(I, \mathbb{R}^n)$ den $\mathbb{R}$-Vektorraum der k-mal stetig differenzierbaren Funktionen $f : I \to \mathbb{R}^n$, das heisst*

$$C^k(I, \mathbb{R}^n) = \{f : I \to \mathbb{R}^n : \ f^{(i)} : I \to \mathbb{R}^n \ \textit{sind definiert und stetig für } 0 \leq i \leq k\}.$$

Weiter definieren wir $C^\infty(I, \mathbb{R}^n)$ als den $\mathbb{R}$-Vektorraum der unendlich oft differenzierbaren Funktionen, also ist

$$C^\infty(I, \mathbb{R}^n) = \cap_{k \in \mathbb{N}_0} C^k(I, \mathbb{R}^n).$$

Einfachheitshalber definieren wir noch im Fall $n = 1$ und $k \in \mathbb{N}_0 \cup \{\infty\}$

$$C^k(I) := C^k(I, \mathbb{R}).$$

BEMERKUNG: Nicht jede differenzierbare Funktion ist stetig differenzierbar (also in C^1). Betrachte dazu $f : \mathbb{R} \to \mathbb{R}$,

$$f(x) = \begin{cases} x^2 \cos \dfrac{1}{x}, & x \neq 0 \\ 0, & x = 0. \end{cases}$$

f ist stetig, da $\lim\limits_{x \to 0} x^2 \cos \dfrac{1}{x} = 0$.

Für $x \neq 0$ ist f differenzierbar mit Ableitung

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

Weiter gilt

$$f'(0) = \lim_{x \to 0} \frac{f(x) - f(0)}{x - 0} = \lim_{x \to 0} x \cos \frac{1}{x} = 0,$$

also ist f auf ganz $\mathbb{R}$ differenzierbar, aber f' ist in $x = 0$ nicht stetig, denn es gilt

$$f'\left(\frac{1}{k\pi}\right) = \frac{2}{k\pi} \cos k\pi + 0 \to 0$$

für $k \to \infty$, aber

$$f'\left(\frac{1}{\frac{\pi}{2} + k2\pi}\right) = \frac{2}{\frac{\pi}{2} + k2\pi} \cos\left(\frac{\pi}{2} + 2\pi k\right) + 1 = 1 \quad \forall\, k \in \mathbb{Z}.$$

Satz 8.2.5. *Sei $f : (a,b) \to \mathbb{R}$ differenzierbar. In $x_0 \in (a,b)$ gelte $f'(x_0) = 0$ und $f''(x_0)$ sei definiert. Dann gilt:*

1) Ist $f''(x_0) > 0$, so hat f in x_0 ein isoliertes, lokales Minimum.

2) Hat f in x_0 ein lokales Minimum, so folgt $f''(x_0) \geq 0$.

Analoge Aussagen gelten mit umgekehrten Ungleichungen für Maxima.

Beweis. Zu 1): Da $f'(x_0) = 0$, gilt

$$f''(x_0) = \lim_{x \to x_0} \frac{f'(x) - f'(x_0)}{x - x_0} = \lim_{x \to x_0} \frac{f'(x)}{x - x_0}$$

Also existiert ein $\delta > 0$ mit $f'(x) < 0$ für $x \in (x_0 - \delta, x_0)$ und $f'(x) > 0$ für $x \in (x_0, x_0 + \delta)$. Damit folgt aus Satz 8.2.3, dass f auf $(x_0 - \delta, x_0)$ streng monoton fallend und auf $(x_0, x_0 + \delta)$ streng monoton wachsend ist.
Somit besitzt f in x_0 ein isoliertes lokales Minimum.

Zu 2): Wäre $f''(x_0) < 0$ in 2), so hätte f nach 1) in x_0 ein isoliertes lokales Maximum im Widerspruch zur Voraussetzung. □

Satz 8.2.6 (Verallgemeinerter Mittelwertsatz)**.** *Seien $f, g : [a,b] \to \mathbb{R}$ stetig und auf (a,b) differenzierbar. Weiter sei $g'(x) \neq 0 \quad \forall\, x \in (a,b)$. Dann ist $g(b) \neq g(a)$ 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 $g(a) = g(b)$, so würde nach Satz 8.2.2 ein $\xi \in (a,b)$ mit $g'(\xi) = 0$ existieren. Dies widerspricht der Voraussetzung.

Definiere jetzt $F : [a,b] \to \mathbb{R}$ durch

$$F(x) = f(x) - \frac{f(b) - f(a)}{g(b) - g(a)}\,(g(x) - g(a))$$

Damit gilt $F(a) = F(b) = f(a)$ und nach Satz 8.2.2 existiert ein $\xi \in (a,b)$ mit

$$0 = F'(\xi) = f'(\xi) - \frac{f(b) - f(a)}{b - a} g'(\xi).$$

□

Satz 8.2.7 (L'Hospitalsche Regeln). *Seien $f, g : (a,b) \to \mathbb{R}$ differenzierbar, und es sei $g'(x) \neq 0$ für alle $x \in (a,b)$. In jeder der beiden Situationen*

1) $f(x) \to 0$ und $g(x) \to 0$ für $x \to a$

2) $f(x) \to \infty$ und $g(x) \to \infty$ für $x \to a$

gilt:

Existiert $\lim\limits_{x \searrow a} \dfrac{f'(x)}{g'(x)}$, so existiert auch $\lim\limits_{x \searrow a} \dfrac{f(x)}{g(x)}$ und es gilt:

$$\lim_{x \searrow a} \frac{f(x)}{g(x)} = \lim_{x \searrow a} \frac{f'(x)}{g'(x)}.$$

Beweis. 1) Wir fassen f und g als in a stetige Funktionen auf mit

$$f(a) = g(a) = 0.$$

Nach Satz 8.2.6 existiert für alle $x \in (a,b)$ ein $\xi \in (a,x)$ mit

$$\frac{f(x)}{g(x)} = \frac{f'(\xi)}{g'(\xi)}.$$

Die Bedingung $x \to a$ impliziert $\xi \to a$ und damit folgt die Behauptung.

2) Sei $\lim\limits_{x \searrow a} \dfrac{f'(x)}{g'(x)} =: A$. Zu $\varepsilon > 0$ wähle man $\delta > 0$ so, dass

$$\Big|\frac{f'(t)}{g'(t)} - A\Big| < \varepsilon \qquad \forall\, t \in (a, a+\delta)$$

Nach Satz 8.2.6 gilt dann für $x, y \in (a, a+\delta)$ mit $x \neq y$ für ein $z \in (x,y)$ bzw. $z \in (y,x)$

$$\Big|\frac{f(x) - f(y)}{g(x) - g(y)} - A\Big| = \Big|\frac{f'(z)}{g'(z)} - A\Big| < \varepsilon$$

Nun ist

$$\frac{f(x)}{g(x)} = \frac{f(x)-f(y)}{g(x)-g(y)} \cdot \frac{1-\frac{g(y)}{g(x)}}{1-\frac{f(y)}{f(x)}}.$$

Jetzt fixieren wir $y = y_0 \in (a, a+\delta)$ und für $x \to a$ gilt dann

$$\frac{1-\frac{g(y_0)}{g(x)}}{1-\frac{f(y_0)}{f(x)}} \to 1$$

Also existiert ein $\delta^* > 0$, so dass $\forall\ x \in (a, a+\delta^*)$ gilt:

$$\Big|\frac{f(x)}{g(x)} - \frac{f(x)-f(y_0)}{g(x)-g(y_0)}\Big| < \varepsilon.$$

Für x mit $a < x < a + \min\{\delta, \delta^*\}$ gilt dann

$$\begin{aligned}\Big|\frac{f(x)}{g(x)} - A\Big| &\le \Big|\frac{f(x)}{g(x)} - \frac{f(x)-f(x_0)}{g(x)-g(x_0)}\Big| \\ &\quad + \Big|\frac{f(x)-f(x_0)}{g(x)-g(x_0)} - A\Big| \\ &< 2\varepsilon\end{aligned}$$

$$\Longrightarrow \lim_{x\searrow a} \frac{f(x)}{g(x)} = A = \lim_{x\searrow a} \frac{f(x)}{g'(x)}$$

□

Beispiele:

1) Sei $s > 0$. Dann gilt

$$\lim_{x\searrow 0} x^s \log x = -\lim_{x\searrow 0} \frac{-\log x}{x^{-s}} = -\lim_{x\searrow 0} \frac{-\frac{1}{x}}{-sx^{-s-1}} = \lim_{x\searrow 0}(-sx^s) = 0$$

Speziell erhält man damit:

$$\lim_{x\searrow 0} x^x = \lim_{x\searrow 0} \exp(x \log x) = e^0 = 1.$$

2) Als nächstes betrachten wir

$$\lim_{x\to 0}\Big(\frac{1}{\sin x} - \frac{1}{x}\Big) = \lim_{x\to 0} \frac{x - \sin x}{x \sin x}.$$

Es gilt $\lim_{x\to 0}(x - \sin x) = 0 = \lim_{x\to 0} x \sin x$. Jetzt folgt mit Hilfe der L'Hospitalschen Regel

$$\lim_{x\to 0} \frac{x-\sin x}{x \sin x} = \lim_{x\to 0} \frac{1-\cos x}{\sin x + x\cos x}.$$

Da $\lim_{x\to 0}(1-\cos x) = 0 = \lim_{x\to 0}(\sin x + x\cos x)$, können wir die L'Hospitalsche Regel nochmals anwenden und erhalten

$$\lim_{x\to 0}\frac{1-\cos x}{\sin x + x\cos x} = \lim_{x\to 0}\frac{\sin x}{2\cos x - x\sin x} = \frac{0}{2} = 2.$$

Insgesamt ergibt sich

$$\lim_{x\to 0}(\frac{1}{\sin x} - \frac{1}{x}) = 0.$$

9 Konvergenz von Funktionenfolgen

9.1 Konvergenz von Funktionenfollgen

Definition 9.1.1. *Es sei $D \subset \mathbb{R}^m$ eine Menge und für jedes $n \in \mathbb{N}$ sei $f_n : D \to \mathbb{C}$ eine Funktion.*

1) *Die Folge (f_n) konvergiert* punktweise *gegen eine Funktion $f : D \to \mathbb{C}$, falls $\forall\ x \in D$ die Folge $(f_n(x))$ gegen $f(x)$ konvergiert., d. h. $\forall\ x \in D \quad \forall\ \varepsilon > 0 \ \ \exists\ N = N(x, \varepsilon) > 0$ mit*

$$|f_n(x) - f(x)| < \varepsilon \quad \forall\ n \geq N.$$

2) *Die Folge (f_n) konvergiert* gleichmässig *gegen eine Funktion $f : D \to \mathbb{C}$, falls $\forall\ \varepsilon > 0 \ \ \exists\ N = N(\varepsilon) > 0$ mit*

$$|f_n(x) - f(x)| < \varepsilon \quad \forall\ x \in D,\ \forall\ n \geq N.$$

Beispiele:

1) Sei $f_n : [0, 1] \to \mathbb{R}$, $f_n(x) = x^n$.

   Für $0 \leq x < 1$ gilt $f_n(x) \to 0$ mit $n \to \infty$ und es gilt $f_n(1) = 1 \ \ \forall\ n \in \mathbb{N}$.

   Damit konvergiert (f_n) punktweise gegen die Funktion $f : [0, 1] \to \mathbb{R}$,

$$f(x) = \begin{cases} 0 & \text{für } \ 0 \leq x < 1 \\ 1 & \text{für } \ x = 1. \end{cases}$$

   (f_n) konvergiert nicht gleichmässig gegen f, denn sei z. B. $\varepsilon = \dfrac{1}{4}$.

   Da $f_n\ \forall\ n \in \mathbb{N}$ stetig ist, existiert $\forall\ n \in \mathbb{N}$ nach dem Zwischenwertsatz, Satz 7.1.1, ein $x_n \in (0, 1)$ mit $f_n(x_n) = \dfrac{1}{2}$.

   Damit folgt $\dfrac{1}{2} = |f_n(x_n) - f(x_n)| > \dfrac{1}{4}$.

2) $f_n : [0, 1] \to \mathbb{R}$, $f_n(x) = \dfrac{1}{n} \sin(\pi n x)$.

Wegen $|\sin y| \leq 1$ für alle $y \in \mathbb{R}$ gilt

$$f_n(x) \to 0 \qquad \forall\, x \in [0,1]$$

$\implies$ (f_n) konvergiert punktweise gegen $f : [0,1] \to \mathbb{R}$, $f(x) = 0$.

Zu $\varepsilon > 0$ $\exists$ $N \in N$ mit $\dfrac{1}{N} < \varepsilon$ (siehe Satz 4.1.1). Damit folgt $\forall\, n \geq N$ und $\forall\, x \in [0,1]$:

$$|f_n(x)| \leq \frac{1}{N} < \varepsilon$$

$\implies$ (f_n) konvergiert auch gleichmässig gegen f.

Bemerkung:

1) Konvergiert $f_n : D \to \mathbb{C}$ gleichmässig gegen $f : D \to \mathbb{C}$, so konvergiert $f_n)$ auch punktweise gegen f.

   Die Umkehrung gilt nicht (siehe Beispiel 9.1.2).

2) Ist $D \subset \mathbb{R}$, $f_n : D \to \mathbb{R}$ gleichmässig konvergent gegen $f : D \to \mathbb{R}$, dann liegen für $\varepsilon > 0$ beliebig die Graphen von f_n für $n \geq N$ in einem „ε–Schlauch" um den Graphen von f.

Satz 9.1.1. *Es sei $D \subset \mathbb{R}^m$ und $f_n : D \to \mathbb{C}$ stetig in $x_0 \in D$ $\forall\, n \in \mathbb{N}$. Konvergiert (f_n) gleichmässig gegen $f : D \to \mathbb{C}$, so ist f ebenfalls setig in $x_0 \in D$.*

Beweis. Sei $\varepsilon > 0$. Da (f_n) gleichmässig gegen f konvergiert, $\exists$ $N \in \mathbb{N}$ mit

$$|f_N(x) - f(x)| < \frac{\varepsilon}{3} \quad \forall\, x \in D.$$

Da f_N stetig in x_0 ist, existiert ein $\delta > 0$ mit

$$|f_N(x) - f_N(x_0)| < \frac{\varepsilon}{3} \qquad \forall\, x \in D \cap B_\delta(x_0)$$

Damit folgt für alle $x \in D \cap B_\delta(x_0)$

$$|f(x) - f(x_0)| \leq \; |f(x) - f_N(x)| + |f_N(x) - f_N(x_0)| + |f_N(x_0) - f(x_0)| < \varepsilon.$$

□

Definition 9.1.2. *Sei $D \subset \mathbb{R}^m$ und $f : D \to \mathbb{R}^n$. Wir definieren die Supremumsnorm*

$$\|f\|_\infty := \sup\{\|f(x)\| : x \in D\}.$$

Die Menge der beschränkten Funktionen auf D, also $\|f\|_\infty < \infty$ bezeichnen wir mit $B(D)$.

Satz 9.1.2. *Sei $D \subset \mathbb{R}^m$ und $f_n : D \to \mathbb{C}$ $\forall\, n \in \mathbb{N}$. Dann gilt: (f_n) konvergiert* *genau dann gleichmässig gegen $f : D \to \mathbb{C}$, falls*

$$\|f_n - f\|_\infty \to 0 \quad \textit{für} \quad n \to \infty.$$

Beweis. „$\Longrightarrow$": Sei $\varepsilon > 0$. Dann existiert ein $N \in \mathbb{N}$ mit $|f_n(x) - f(x)| < \frac{\varepsilon}{2}$ für alle $x \in D$ und für alle $n \geq N$. Damit folgt für alle $n \geq N$

$$\|f_n - f\|_\infty \leq \frac{\varepsilon}{2} < \varepsilon.$$

„$\Longleftarrow$": Sei $\varepsilon > 0$. Es existiert ein $N \in \mathbb{N}$ mit $\|f_n - f\|_\infty < \frac{\varepsilon}{2}$ für alle $n \geq N$.

Also gilt für alle $n \geq N$ und $x \in D$

$$|f_n(x) - f(x)| \leq \sup_{x \in D} |f_n(x) - f(x)| = \|f_n - f\|_\infty < \varepsilon.$$

$\square$

Satz 9.1.3. *Sei $p(z) = \sum\limits_{k=0}^{\infty} a_k z^k$ eine Potenzreihe mit Konvergenzradius $R > 0$. Dann ist P stetig auf $B_R(0) = \{z \in \mathbb{C} : |z| < R\}$.*

Beweis. Sei $w \in B_R(0)$ und sei $\Theta \in (0, 1)$ so gewählt, dass $|w| \leq \Theta R$. Wir zeigen, dass p stetig in w ist.

Für alle $n \in \mathbb{N}$ ist $p_n(z) = \sum\limits_{k=0}^{n} a_n z^n$ stetig auf $\mathbb{C}$. Sei $z_0 = \sqrt{\Theta} R$. Da p in z_0 absolut konvergiert, existiert ein $M \in \mathbb{R}$ so, dass für alle $k \in \mathbb{N}$ gilt

$$|a_k||z_0|^k \leq M.$$

Damit gilt für alle $z \in \mathbb{C}$ mit $|z| \leq \Theta R$

$$\begin{aligned}
|p(z) - p_n(z)| &\leq \sum_{k=n+1}^{\infty} |a_k||z|^k \leq M \sum_{k=n+1}^{\infty} \left(\frac{|z|}{|z_0|}\right)^k \\
&\leq \frac{M}{1 - \frac{|z|}{|z_0|}} \left(\frac{|z|}{|z_0|}\right)^{n+1} \leq \frac{M}{1 - \frac{\Theta R}{\sqrt{\Theta} R}} \left(\frac{\Theta R}{\sqrt{\Theta} R}\right)^{n+1} \\
&= \frac{M}{1 - \sqrt{\Theta}} \Theta^{\frac{n+1}{2}} \\
&\to 0 \quad \text{für } n \to \infty
\end{aligned}$$

Damit konvergiert die Folge $p_n : \{z \in \mathbb{C} : |z| \leq \Theta R\} \to \mathbb{C}$ gleichmässig gegen $p : \{z \in \mathbb{C} : |z| \leq \Theta R\} \to \mathbb{C}$ und nach Satz 9.1.1 ist p stetig auf $\{z \in \mathbb{C} : |z| \leq \Theta R\}$. Insbesondere ist p stetig in w. $\square$

10 Integration in $\mathbb{R}$

10.1 Stammfunktionen

Seien $-\infty < a < b < \infty$ und sei $f \in C^0((a,b))$.

Definition 10.1.1. *Eine Funktion $F \in C^1((a,b))$ heisst Stammfunktion zu f, falls gilt:*

$$F' = f \quad \text{in} \quad (a,b).$$

Beispiel: Die Funktion $f(x) = \dfrac{1}{x}$ besitzt die Stammfunktion $F(x) = \log x$, aber auch die Stammfunktion $F(x) = \log x + 1500$.
Insbesondere ist mit jeder Stammfunktion F auch $F+c$, $c \in \mathbb{R}$ beliebig, eine Stammfunktion.

Satz 10.1.1. *Sind $F_1, F_2 \in C^1((a,b))$ Stammfunktionen zu $f \in C^0((a,b))$, dann existiert ein $c \in \mathbb{R}$ mit*

$$F_1 = F_2 + c.$$

Beweis: Auf (a,b) gilt

$$(F_1 - F_2)' = F_1' - F_2' = f - f = 0.$$

Nach Satz 8.2.3 existiert dann ein $c \in \mathbb{R}$ mit $F_1 - F_2 = c$. □

Definition 10.1.2. *Sei $f \in C^0((a,b))$ und sei $F \in C^1((a,b))$ eine Stammfunktion von f.*

1) Für $a < x_0 < x < b$ heisst

$$\int_{x_0}^{x} f(\xi)\, d\xi := F(x) - F(x_0)$$

das Integral von f über $[x_0, x]$.

2) $\int f(x)\, dx$ heisst unbestimmtes Integral und ist eine alternative Notation für die Stammfunktion zu f.

Bemerkung: Wegen Satz 10.1.1 ist die Defintion des Integrals von f unabhängig von der Wahl der Stammfunktion.

Beispiele:

f	$\int f(x)\, dx$
$x^\alpha, \alpha \neq -1$	$\frac{1}{\alpha+1}x^{\alpha+1}$
x^{-1}	$\log x$
e^x	e^x
$\cos x$	$\sin x$
$\sin x$	$-\cos x$
$\frac{1}{\sqrt{1-x^2}}$	$\arcsin x$
$-\frac{1}{\sqrt{1-x^2}}$	$\arccos x$

Satz 10.1.2. *1. Seien $f, g \in C^0((a,b))$ mit Stammfunktionen $F, G \in C^1((a,b))$ und seien $\alpha, \beta \in \mathbb{R}$. Dann ist $\alpha F + \beta G \in C^1((a,b))$ eine Stammfunktion von $\alpha f + \beta g$, d.h*

$$\int (\alpha f + \beta g)(\xi)\, d\xi = \alpha \int f(\xi)\, d\xi + \beta \int g(\xi)\, d\xi.$$

2. (Partielle Integration) Seien $u, v \in C^1((a,b))$ und es existiere eine Stammfunktion $F \in C^1((a,b))$ von $f = uv' \in C^0((a,b))$. Dann besitzt die Funktion $u'v \in C^0((a,b))$ die Stammfunktion:

$$\int (u'v)(\xi)\, d\xi = uv - \int (uv')(\xi)\, d\xi.$$

Beweis: 1. Die Summenregel, Satz 8.1.5, liefert

$$(\alpha F + \beta G)' = \alpha F' + \beta G' = \alpha f + \beta g$$

und damit die Behauptung.

2. Mit $F' = f = uv'$ folgt aus der Produktregel, Satz 8.1.5,

$$(uv - F)' = u'v + uv' - F' = u'v$$

und damit ist $uv - F$ eine Stammfunktion von $u'v$.

□

Beispiele:

- Sei $p(x) = a_n x^n + ... + a_1 x + a_0$. Mit $\int x^k\, dx = \frac{x^{k+1}}{k+1}$ für alle $k \in \mathbb{N}_0$ folgt:

$$\int p(x)\, dx = \frac{a_n}{n+1} x^{n+1} + ... + \frac{a_1}{2} x^2 + a_0 x.$$

- Seien $u(x) = x$ und $v(x) = \log x$ mit $u'(x) = 1$ bzw. $v'(x)\frac{1}{x}$. Dann folgt

$$\int \log x \, dx = x \log x - \int 1 \, dx = x \log x - x.$$

Satz 10.1.3. *Seien $f, g \in C^0((a,b))$ mit Stammfunktionen $F, G \in C^1((a,b))$ und sei $f(x) \leq g(x)$ für alle $x \in (a,b)$. Dann gilt für $a < x_0 < x_1 < b$*

$$\int\limits_{x_0}^{x_1} f(x) \, dx \leq \int\limits_{x_0}^{x_1} g(x) \, dx.$$

Beweis: Definiere $h \in C^0((a,b))$ durch $h(x) = g(x) - f(x) \geq 0$. Nach Satz 10.1.2 hat h eine Stammfunktion $H \in C^1((a,b))$ mit $H' = h \geq 0$. Damit ist H nach Satz 8.2.3 monoton wachsend auf dem Intervall (a,b). Speziell gilt

$$\int\limits_{x_0}^{x_1} g(x) \, dx - \int\limits_{x_0}^{x_1} f(x) \, dx = \int\limits_{x_0}^{x_1} h(x) \, dx = H(x_1) - H(x_0) \geq 0.$$

□

Satz 10.1.4. *Sei $f \in C^0((a,b))$ mit Stammfunktion $F \in C^1((a,b))$ und seien $a < x_0 < x_1 < x_2 < b$. Dann gilt*

$$\int\limits_{x_0}^{x_1} f(x) \, dx + \int\limits_{x_1}^{x_2} f(x) \, dx = \int\limits_{x_0}^{x_2} f(x) \, dx.$$

Beweis: Wir berechnen

$$\begin{aligned}\int\limits_{x_0}^{x_1} f(x) \, dx + \int\limits_{x_1}^{x_2} f(x) \, dx &= (F(x_1) - F(x_0)) + (F(x_2) - F(x_1)) \\ &= F(x_2) - F(x_0) = \int\limits_{x_0}^{x_2} f(x) \, dx.\end{aligned}$$

□

Satz 10.1.5. *(Substitutionsregel) Seien $f, g \in C^1((a,b))$. Dann gilt für $a < x_0 < x_1 < b$:*

$$\int\limits_{x_0}^{x_1} f'(g(x))g'(x) \, dx = f(g(x_1)) - f(g(x_0)) = \int\limits_{g(x_0)}^{g(x_1)} f'(y) \, dy$$

Beweis: Aus der Kettenregel, Satz 8.1.6, folgt

$$(f \circ g)'(x) = f'(g(x))g'(x)$$

und damit ist $(f \circ g)$ eine Stammfunktion von $(f' \circ g)g'$. Da auch f eine Stammfunktion von f' ist folgt damit die Behauptung. □

BEMERKUNG: Man kann also formal die Variable $y = g(x)$ substituieren mit "$dy = g'(x)\,dx$".

BEISPIELE:

- $$\int_0^2 \frac{x}{\sqrt{1+x^2}}\,dx \overset{y=1+x^2}{=} \frac{1}{2}\int_1^5 \frac{1}{\sqrt{y}}\,dy = \sqrt{y}\,\Big|_{y=1}^{y=5} = \sqrt{5}-1$$

- $$\int_0^1 \frac{x}{1+x^2}\,dx \overset{y=1+x^2}{=} \int_1^2 \frac{1}{2y}dy = \frac{\log 2}{2}$$

- **i)** $$\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^2\phi\,d\phi \overset{x=\sin\phi}{=} \int_{-1}^{1} \sqrt{1-x^2}\,dx$$

  ii) Mit $u(\phi) = \sin\phi,\ v(\phi) = \cos\phi$ folgt $u'(\phi) = \cos\phi, v'(\phi) = -\sin\phi$ und mit partieller Integration erhält man

  $$\begin{aligned}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^2\phi\,d\phi &= \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} u'(\phi)v(\phi)\,d\phi \\ &= [\sin\phi\cos\phi]_{-\frac{\pi}{2}}^{\frac{\pi}{2}} + \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sin^2\phi\,d\phi \\ &= \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} (1-\cos^2\phi)\,d\phi \\ &= \pi - \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^2\phi\,d\phi.\end{aligned}$$

  Damit folgt

  $$\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos^2\phi\,d\phi = \frac{\pi}{2} = \int_{-1}^{1} \sqrt{1-x^2}\,dx.$$

PARTIALBRUCHZERLEGUNG:

Wir wollen $\int \frac{dx}{1-x^2}$ berechnen und wir ignorieren das $\arctan'(x) = \frac{1}{1-x^2}$.
Es gilt $1 - x^2 = (1-x)(1+x)$ und damit

$$\frac{1}{1-x^2} = \frac{1}{(1-x)(1+x)} \stackrel{!}{=} \frac{a}{1-x} + \frac{b}{1+x}.$$

Damit das letzte Gleichheitszeichen gilt müssen a und b wie folgt gewählt werden:

$$a(1+x) + b(1-x) = 1 \quad \forall\, x \in \mathbb{R}\,.$$

Also ist $a + b + x(a-b) = 1 \quad \forall\, x$ und damit folgt nach Lemma 3.3.2

$$\begin{cases} a = b \\ a + b = 1, \end{cases}$$

also $a = b = \frac{1}{2}$. Insgesamt erhalten wir somit

$$\int \frac{dx}{1-x^2} = \frac{1}{2}\int \frac{dx}{1-x} + \frac{1}{2}\int \frac{dx}{1+x} = -\frac{\log(1-x)}{2} + \frac{\log(1+x)}{2}.$$

10.2 Riemannsches Integral

Das Ziel der Einführung des Riemannschen Integrals ist eine Erweiterung des Integralbegriffes. Unter anderem soll auch gezeigt werden, dass jede stetige Funktion eine Stammfunktion besitzt.

Beispiele:

- Sei $f(x) = c$ für ein $c > 0$ und für alle $x \in [a,b]$. Dann ist $F(x) = cx$ eine Stammfunktion von f und das Integral von f berechnet sich zu

  $$\int\limits_a^b f(x)\, dx = c(b-a).$$

  Für $c < 0$ erhalten wir auch den Flächeninhalt unter dem Graphen von f (falls richtig orientiert!)

- Sei $f(x) = mx,\ m \in \mathbb{R}$ mit Stammfunktion $F(x) = \frac{1}{2}mx^2$. Auch in diesem Fall entspricht das Integral

  $$\int\limits_a^b (mx) dx = \frac{1}{2}m(b^2 - a^2)$$

  dem Flächeninhalt unter dem Graphen von f.

Frage: Für welche Funktionen $f : [a,b] \to \mathbb{R}$ kann man den Flächeninhalt zwischen dem Intervall $[a,b]$ und Graph(f) messen?

Definition 10.2.1. *1. Eine Funktion $f: [a,b] \to \mathbb{R}$ heisst Treppenfunktion, falls für eine Zerlegung von $I = [a,b]$ in disjunkte Teilintervalle (offen, abgeschlossen oder halboffen) $I_1, ..., I_k$ mit Konstanten $c_k \in \mathbb{R}$ gilt:*

$$f(x) = c_k \quad \forall\, x \in I_k,\ 1 \le k \le K$$

d.h.

$$f = \sum_{k=1}^{K} c_k \chi_{I_k},$$

wobei

$$\chi_{I_k}(x) = \begin{cases} 1, & x \in I_k \\ 0 & \text{sonst} \end{cases}$$

die charakteristische Funktion von I_k ist.

2. Das Integral einer Treppenfunktion $f = \sum_{k=1}^{K} c_k \chi_{I_k} : [a,b] \to \mathbb{R}$ ist

$$\int_a^b \left(\sum_{k=1}^{K} c_k \chi_{I_k}(x) \right) dx := \sum_{k=1}^{K} c_k |I_k|,$$

wobei $|I_k|$ die Länge von I_k bezeichnet.

Bemerkung: Die konstante Funbktion $f(x) = c,\ c \in \mathbb{R}$ kann wie folgt als Treppenfunktion aufgefasst werden:

$$f = \sum_{k=1}^{K} c_k \chi_{I_k}$$

mit $c_k = c \quad \forall\ 1 \le k \le K$, wobei $I_k,\ 1 \le k \le K$, eine disjunkte Zerlegung von $I = [a,b]$ ist. Es gilt:

$$\int_a^b f(x)dx = c(b-a) = \sum_{k=1}^{K} c_k |I_k|.$$

Lemma 10.2.1. *Sind $e, g : [a,b] \to \mathbb{R}$ Treppenfunktionen mit $e \le g$ (d.h. $e(x) \le g(x) \quad \forall\, x \in [a,b]$), dann gilt:*

$$\int_a^b e(x)\, dx \le \int_a^b g(x)\, dx.$$

Beweis: Seien $e = \sum_{k=1}^{K} c_k \chi_{I_k}$ und $g = \sum_{l=1}^{L} d_l \chi_{I_l}$ Treppenfunktionen mit disjunkten

Intervallen $I_1, .., I_k$, bzw. $J_1, ..., J_l$, wobei

$$I = [a, b] = \bigcup_{k=1}^{K} I_k = \bigcup_{l=1}^{L} J_l.$$

Betrachte die Intervalle

$$I_{kl} = I_k \cap J_l, \ 1 \leq k \leq K, \ 1 \leq l \leq L.$$

Die I_{kl} sind auch disjunkt mit

$$I_k = \bigcup_{l=1}^{L} I_{kl} = \bigcup_{l=1}^{L} I_k \cap J_l, \quad \text{und} \quad J_l = \bigcup_{k=1}^{K} I_{kl}.$$

Weiter gilt:

$$\bigcup_{\substack{l=1 \\ k=1}}^{K,L} I_{kl} = \bigcup_{k=1}^{K} \left(\bigcup_{l=1}^{L} I_{kl} \right) = \bigcup_{k=1}^{K} I_k = I.$$

Für beliebige $1 \leq k \leq K, \ 1 \leq l \leq L$ und für alle $x \in I_{kl}$ haben wir die Abschätzung

$$c_k = e(x) \leq g(x) = d_l \quad \text{also} \quad c_k \leq d_l.$$

Damit gilt:

$$\begin{aligned} \int_a^b e(x)\ dx &= \sum_{k=1}^{K} c_k |I_k| = \sum_{k=1}^{K} \sum_{l=1}^{L} c_k |I_{kl}| \\ &\leq \sum_{k=1}^{K} \sum_{l=1}^{L} d_l |I_{kl} = \sum_{l=1}^{L} d_l |J_l| = \int_a^b g(x)\ dx. \end{aligned}$$

□

Sei jetzt $f : [a, b] \to \mathbb{R}$ beschränkt, d.h. es existiert ein $c \in \mathbb{R}$ so, dass für alle $x \in [a, b]$ gilt

$$|f(x)| \leq c.$$

Dann existieren Treppenfunktionen $e = -c\chi_{[a,b]}$ und $g = c\chi_{[a,b]}$ welche

$$e(x) \leq f(x) \leq g(x) \quad \forall\, x \in [a, b]$$

erfüllen.

Definition 10.2.2. *1. Für beschränktes $f : [a, b] \to \mathbb{R}$ definiere das untere, bzw. obere Riemann-Integral (R-Integral) durch*

$$\underline{\int_a^b} f(x)\ dx := \sup\{\int_a^b e(x)\ dx : e \text{ Treppenfunktion}, \ e \leq f\}$$

und
$$\overline{\int_a^b} f(x)\,dx := \inf\{\int_a^b g(x)\,dx : g \text{ Treppenfunktion},\ f \le g\}.$$

2. *f heisst über $[a,b]$ Riemann-integrabel (R-integrabel), falls*
$$\underline{\int_a^b} f(x)\,dx = \overline{\int_a^b} f(x)\,dx = A.$$

   In diesem Fall heisst $A =: \int_a^b f(x)dx$ das ***Riemann-Integral von f.***

BEMERKUNGEN:

- Für alle beschränkten Funktionen f gilt die Abschätzung
$$\underline{\int_a^b} f(x)\,dx \le \overline{\int_a^b} f(x)\,dx.$$

- Eine Funktion f ist genau dann R-integrabel, wenn für alle $\varepsilon > 0$ Treppenfunktionen $e, g : [a,b] \to \mathbb{R}$ existieren mit $e \le f \le g$ und
$$\int_a^b g(x)\,dx - \int_a^b e(x)\,dx < \varepsilon.$$

Beweis:

- Seien e, g Treppenfunktionen mit $e \le f \le g$. Aus Lemma 10.2.1 folgt
$$\int_a^b e(x)\,dx \le \int_a^b g(x)\,dx$$
und damit
$$\underline{\int_a^b} f(x)\,dx = \sup\{\int_a^b e(x)\,dx : e \text{ Treppenfunktion},\ e \le f\} \le \int_a^b g(x)\,dx.$$
Indem man das Infimum über alle zulässigen Treppenfunktionen g bildet, erhält man die Behauptung.

- Mit Hilfe des gerade bewiesenen können wir für alle Treppenfunktionen e, g

abschätzen

$$\int_a^b e(x)\,dx \le \underline{\int_a^b} f(x)\,dx \le \overline{\int_a^b} f(x)\,dx \le \int_a^b g(x)\,dx.$$

Damit folgt direkt die Behauptung.

□

Satz 10.2.1. *Sei $D \subset \mathbb{R}$ kompakt und sei $f : D \to \mathbb{R}^m$ stetig. Dann ist f gleichmässig stetig, d.h. zu $\varepsilon > 0$ gibt es ein $\delta > 0$ mit:*

$$x, x' \in D,\ \|x - x'\| < \delta \quad \Rightarrow \quad \|f(x) - f(x')\| < \varepsilon.$$

Beweis: Anderenfalls gibt es für ein $\varepsilon > 0$ Punkte $x_n, x'_n \in D$ mit

$$\|x_n - x'_n\| \to 0, \quad \text{aber} \quad \|f(x_n) - f(x'_n)\| \ge \varepsilon.$$

Da D kompakt ist, existieren Teilfolgen x_{n_k}, x'_{n_k} mit $x_{n_k} \to x_0 \in D$ und $x'_{n_k} \to x_0$. Da f stetig ist erhalten wir den Widerspruch

$$\varepsilon \le \lim_{k\to\infty} \|f(x_{n_k}) - f(x'_{n_k})\| = \|f(x_0) - f(x_0)\| = 0.$$

□

Beispiel:
Ist D nicht kompakt, so muss eine stetige Funktion $f : D \to \mathbb{R}$ nicht gleichmässig stetig sein. Sei z.B. $f : \left(0, \frac{1}{\pi}\right) \to \mathbb{R},\ f(x) = \sin\frac{1}{x}$.
Mit $x_n = \frac{1}{n\pi},\ x'_n = \frac{1}{n\pi + \frac{\pi}{2}}$ gilt $x_n \to 0,\ x'_n \to 0$, aber

$$\left|f(x_n) - f(x'_n)\right| = 1 \quad \forall\, n \in \mathbb{N}.$$

Satz 10.2.2. *Sei $f : [a,b] \to \mathbb{R}$ stetig. Dann ist f über $[a,b]$ R-integrabel.*

Beweis: Da $[a,b]$ kompakt, ist beschränkt und gleichmässig stetig. Zu $\varepsilon > 0$ existiert damit ein $\delta > 0$ mit

$$\forall\, x, y \in [a,b] : |x - y| < \delta \quad \Rightarrow \quad |f(x) - f(y)| < \varepsilon. \tag{10.1}$$

Für $k \in \mathbb{N}$ mit $\frac{b-a}{K} < \delta$ unterteile man $[a,b]$ in disjunkte Teilintervalle I_k mit Endpunkten

$$a_k = a + (k-1)\frac{b-a}{K},\ b_k = a_k + \frac{b-a}{K}$$

und Länge

$$|I_k| = \frac{b-a}{K},\ 1 \le k \le K.$$

Weiter setze man

$$c_k = \inf_{I_k} f \le \sup_{I_k} f = d_k.$$

Dann sind

$$e = \sum_{k=1}^{K} c_k \chi_{I_k},\ g = \sum_{k=1}^{K} d_k \chi_{I_k}$$

Treppenfunktionen mit $e \le f \le g$.
Da für $1 \le k \le K$ gilt nach nach Konstruktion

$$\sup_{x,y \in I_k} |x-y| = |I_k| < \delta,$$

folgt mit (10.1) auch

$$d_k - c_k = \sup_{x,y \in I_k} (f(x) - f(y)) \le \sup_{x,y \in I_k} |f(x) - f(y)| \le \varepsilon.$$

Damit können wir abschätzen

$$\int_a^b g\ dx - \int_a^b e\ dx = \sum_{k=1}^{K} (d_k - c_k)\, |I_k| \le \varepsilon \sum_{k=1}^{K} |I_k| = (b-a)\varepsilon.$$

Die Behauptung folgt aus der obigen Bemerkung. □

Das Integral einer stetigen Funktion $f \in C^0([a,b])$ kann man numerisch approximieren:

Satz 10.2.3. *Sei $f \in C^0([a,b])$. Dann gilt für eine beliebige Folge von Zerlegungen*

$$I = [a,b] = \bigcup_{k=1}^{K_n} I_k^n$$

von I in disjunkte Teilintervalle I_k^n, $1 \le k \le K_n$, mit Feinheit

$$\delta_n = \sup_{1 \le k \le K_n} |I_k^n| \to 0\, (n \to \infty)$$

und eine beliebige Auswahl von Punkten $x_k^n \in I_k^n$, $1 \le k \le K_n$, stets

$$\int_a^b \left(\sum_{k=1}^{K_n} f(x_k^n) \chi_{I_k^n} \right) dx = \sum_{k=1}^{K_n} f(x_k^n)\, |I_K^n| \to \int_a^b f\ dx \quad (n \to \infty).$$

Beweis: Zu $\varepsilon > 0$ wähle $\delta = \delta(\varepsilon) > 0$ mit (10.1), dazu $n_0 = n_0(\varepsilon) \in \mathbb{N}$ mit

$$\forall\, n \ge n_0 :\ \delta_n < \delta.$$

Für $n \in \mathbb{N}$ setze weiter

$$c_k^n = \inf_{I_k^n} f \leq f(x_k^n) \leq \sup_{I_k^n} f = d_k^n, \ 1 \leq k \leq K_n.$$

Wie in Satz 10.2.2 erhalten wir für $n \geq n_0$ die Abschätzung

$$d_k^n - c_k^n \leq \sup_{x,y \in I_k^n} |f(x) - f(y)| \leq \epsilon, \ 1 \leq k \leq K_n.$$

Definiere die Treppenfunktionen

$$e_n = \sum_{k=1}^{K_n} c_k^n \chi_{I_k^n} \leq f_n = \sum_{k=1}^{N} f\left(x_k^n\right) \chi_{I_k^n} \leq g_n = \sum_{k=1}^{K_n} d_k^n \chi_{I_k^n}.$$

Da f nach Satz 10.2.2 R-integrabel ist, können wir für $n \leq n_0(\epsilon)$ abschätzen

$$\begin{aligned}
R_n := \int_a^b f \, dx - \int_a^b \left(\sum_{k=1}^{K-n} f\left(x_k^n\right) \chi_{I_k^n} \right) dx &= \overline{\int_a^b} f_n \, dx - \int_a^b f_n \, dx \\
\leq \int_a^b g_n \, dx - \int_a^b e_n \, dx = \sum_{k=1}^{K_n} \left(d_k^n - c_k^n\right) |I_k^n| &\leq \varepsilon \sum_{k=1}^{K_n} |I_k^n| = (b-a)\varepsilon
\end{aligned}$$

Analog erhalten wir $R_n \geq -(b-a)\varepsilon$ und damit $|R_n| \leq (b-a)\varepsilon$. daraus folgt die Behauptung. □

10.3 Integrationsregeln und Hauptsatz

Satz 10.3.1 (Monotonie des R-Integrals)**.** *Seien $f_1, f_2 : [a,b] \to \mathbb{R}$ beschränkt und R-integrabel mit $f_1 \leq f_2$. Dann gilt:*

$$\int_a^b f_1 \, dx \leq \int_a^b f_2 \, dx.$$

Beweis: Jede Treppenfunktion $g : [a,b] \to \mathbb{R}$ mit $f_2 \leq g$ erfüllt auch $f_1 \leq g$, also gilt:

$$\int_a^b f_1 \, dx = \overline{\int_a^b} f_1 \, dx \leq \int_a^b g \, dx.$$

Nach Übergang zum Infimum bzgl. derartiger Treppenfunktionen $g \geq f_2$ erhalten wir

$$\int_a^b f_1 \, dx \leq \overline{\int_a^b} f_2 \, dx = \int_a^b f_2 \, dx.$$

□

Satz 10.3.2 (Linearität des R-Integrals). *Seien $f, f_1, f_2 : [a,b] \to \mathbb{R}$ beschränkt und R-integrabel und sei $\alpha \in \mathbb{R}$. Dann sind die Funktionen $\alpha f, f_1 + f_2$ über $[a,b]$ R-integrabel, und*

$$\begin{aligned}\int_a^b (\alpha f)\, dx =& \alpha \int_a^b f\, dx \\ \int_a^b (f_1 + f_2)\, dx =& \int_a^b f_1\, dx + \int_a^b f_2\, dx.\end{aligned}$$

Beweis: Für Treppenfunktionen lassen sich die beiden Behauptungen leicht nachrechnen.

Im Folgenden seien jetzt f, f_1, f_2 wie im Satz und $\alpha \in \mathbb{R}$.

i) Für $\alpha \geq 0$ und Treppenfunktionen e, g mit $e \leq f \leq g$ gilt $\alpha e \leq \alpha f \leq \alpha g$ und damit

$$\begin{aligned}\overline{\int_a^b} (\alpha f)\, dx \leq& \inf \left\{ \int_a^b (\alpha g)\, dx : \ g \text{ Treppenfkt.}, \ g \geq f \right\} \\ =& \alpha \inf \left\{ \int_a^b g\, dx : \ g \text{ Treppenfkt.}, \ g \geq f \right\} = \alpha \overline{\int_a^b} f\, dx \\ =& \alpha \underline{\int_a^b} f\, dx = \alpha \sup \left\{ \int_a^b e\, dx : \ e \text{ Treppenfkt.}, \ e \leq f \right\} \\ =& \sup \left\{ \int_a^b (\alpha e)\, dx : \ e \text{ Treppenfkt.}, \ e \leq f \right\} \leq \underline{\int_a^b} f\, dx \leq \overline{\int_a^b} (\alpha f)\, dx,\end{aligned}$$

wobei in der dritten Zeite benutzt wurde das f R-integrabel ist. Aus dieser Ungleichungskette folgt das (αf) auch R-integrabel ist mit

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

Analog argumentiert man für $\alpha < 0$.

ii) Seien e_i, g_i Treppenfunktionen mit $e_i \leq f_i \leq g_i$ für $1 \leq i \leq 2$. Die Funktionen $e_1 + e_2$, $g_1 + g_2$ sind wieder Treppenfunktionen mit

$$e_1 + e_2 \leq f_1 + f_2 \leq g_1 + g_2.$$

Damit erhalten wir die Abschätzung

$$\begin{aligned}
\overline{\int_a^b}(f_1+f_2)\,dx &\leq \inf\{\int_a^b (g_1+g_2)\,dx:\ g_1,g_2 \text{ Treppenfkt., } g_i\geq f_i,\ 1\leq i\leq 2\}\\
&= \inf\{\int_a^b g_1\,dx:\ g_1 \text{ TF., } g_1\geq f_1\} + \inf\{\int_a^b g_2\,dx:\ g_2 \text{ TF, } g_2\geq f_2\}\\
&= \overline{\int_a^b} f_1\,dx + \overline{\int_a^b} f_2\,dx = \underline{\int_a^b} f_1\,dx + \underline{\int_a^b} f_2\,dx\\
&= \sup\{\int_a^b e_1\,dx:\ e_1 \text{ TF}, e_1\leq f_1\} + \sup\{\int_a^b e_2\,dx:\ e_2 \text{ TF}, e_2\leq f_2\}\\
&= \sup\{\int_a^b e_1\,dx + \int_a^b e_2\,dx:\ e_1,e_2 \text{ Treppenfkt., } e_i\leq f_i,\ 1\leq i\leq 2\}\\
&= \sup\{\int_a^b (e_1+e_2)\,dx:\ e_1,e_2 \text{ Treppenfkt. , } e_i\leq f_i,\ 1\leq i\leq 2\}\\
&\leq \underline{\int_a^b}(f_1+f_2)\,dx \leq \overline{\int_a^b}(f_1+f_2)\,dx.
\end{aligned}$$

Damit folgt wie in i) das f_1+f_2 R-integrabel ist mit

$$\int_a^b (f_1+f_2)\,dx = \int_a^b f_1\,dx + \int_a^b f_2\,dx.$$

□

Korollar 10.3.1. *Für $f\in C^0([a,b])$ gilt:*

$$\left|\int_a^b f\,dx\right| \leq \int_a^b |f|\ dx \leq ||f||_\infty (b-a).$$

Beweis: Für allle $x\in[a,b]$ gilt

$$\pm f(x)\leq |f(x)|\leq ||f||_\infty .$$

Der Beweis folgt aus Satz 10.3.1 und Satz 10.3.2. □

Korollar 10.3.2. *Seien die Funktionen $f, f_k\in C^0([a,b])$ und f_k konvergiere gleichmässig gegen f. Dann gilt:*

$$\left|\int_a^b f_k\,dx - \int_a^b f\,dx\right| \leq \int_a^b |f_k-f|\ dx \leq ||f_k-f||_\infty (b-a)\to 0\quad (\text{für } k\to\infty)$$

Beispiel: Das Korollar ist nicht richtig wenn die Funktionen f_k nur punktweise gegen f konvergieren. Dazu betrachte man für $n \geq 2$ die Funktionen $f_n : [0,1] \to \mathbb{R}$ definiert durch

$$f_n(x) := \max(n - n^2|x - \frac{1}{n}|, 0).$$

Es gilt $f_n \in C^0([0,1])$ und $f_k(x) \to 0$ für $n \to \infty$. Damit konvergiert f_n punktweise gegen die Nullfunktion. Allerdings konvergiert f_n nicht gleichmässig, denn für kein $n \geq 2$ gilt

$$||f_n - 0||_\infty < 1,$$

da $f_n(\frac{1}{n}) = n \geq 2$. Für die Integrale berechnen wir für alle $n \geq 2$

$$\int_0^1 f_n \, dx = 1 \neq 0 = \int_0^1 0 \, dx.$$

Bemerkung:
Sei $p(x) = \sum_{k=0}^{\infty} c_k x^k$ eine Potenzreihe mit Konvergenzradius $\rho := \frac{1}{\limsup_{k\to\infty} \sqrt[k]{|c_k|}}$.
Für alle $0 \leq r < p$ ist p nach Satz 9.1.3 auf $B_r(0)$ gleichmässig konvergent. Damit impliziert das obige Korollar für alle $-\rho < a < b < \rho$:

$$\begin{aligned}
\int_a^b p(x) \, dx &= \int_a^b \left(\lim_{n\to\infty} \sum_{k=0}^{n} c_k x^k \right) dx \\
&= \lim_{n\to\infty} \int_a^b \left(\sum_{k=0}^{n} c_k x^k \right) dx = \lim_{n\to\infty} \left(\sum_{k=0}^{n} c_k \int_a^b x^k \, dx \right) \\
&= \lim_{n\to\infty} \sum_{k=0}^{n} \frac{c_k}{k+1} \left(b^{k+1} - a^{k+1} \right) = \sum_{k=0}^{\infty} \frac{c_k}{k+1} \left(b^{k+1} - a^{k+1} \right).
\end{aligned}$$

Damit erhalten wir

Korollar 10.3.3. *Potenzreihen dürfen innerhalb ihres Konvergenzkreises (d.h. $B_\rho(0)$) gliedweise integriert werden.*

Beispiel:
Für $0 \leq b < 1$ erhalten wir:

$$\begin{aligned}
\log(1+b) &= \int_1^{1+b} \frac{dx}{x} = \int_1^{1+b} \frac{dx}{1-(1-x)} \\
&= \int_1^{1+b} \left(\sum_{k=0}^{\infty} (1-x)^k \right) dx = \sum_{k=0}^{\infty} \left[\frac{-(1-x)^{k+1}}{k+1} \right]_{x=1}^{x=1+b} \\
&= \sum_{k=0}^{\infty} (-1)^k \frac{b^{k+1}}{k+1}.
\end{aligned}$$

Da die Folge $\frac{b^{k+1}}{k+1}$ für alle $0 \le b \le 1$ eine monoton fallende Nullfolge ist (es gilt

$$\frac{b^{k+1}}{k+1}\frac{k}{b^k} = b\frac{k}{k+1} < 1)$$

konvergiert die Reihe $\sum_{k=0}^{\infty}(-1)^k\frac{b^{k+1}}{k+1}$ nach dem Leibniz-Kriterium, Satz 4.4.5. Die Fehlerabschätzung beim Leibniz-Kriterium liefert weiter

$$\left|\log(1+b) - \sum_{k=0}^{n-1}(-1)^k\frac{b^{k+1}}{k+1}\right| \le \frac{b^{n+1}}{n+1} \quad \forall\, 0 \le b \le 1 \quad \forall\, n \in \mathbb{N}.$$

Für $b \to 1$ folgt damit

$$\left|\log(2) - \sum_{k=0}^{n-1}(-1)^k\frac{1}{k+1}\right| \le \frac{1}{n+1} \quad \forall\, n \in \mathbb{N}$$

und für $n \to \infty$ gilt schliesslich

$$\log 2 = \sum_{k=0}^{\infty}(-1)^k\frac{1}{k+1}.$$

Satz 10.3.3. *Sei $f : [a,b] \to \mathbb{R}$ beschränkt und R-integrabel über $[a,b]$ und sei $x_0 \in [a,b]$. Dann ist f eingeschränkt auf $[a,x_0]$ und $[x_0,b]$ auch R-integrabel mit*

$$\int_a^b f\;dx = \int_a^{x_0} f\;dx + \int_{x_0}^b f\;dx.$$

Beweis: Ist f eine Treppenfunktion, $f = \sum_{k=1}^{K} c_k\chi_{I_k}$, so sind auch $f_1 := \sum_{k=1}^{K} c_k\chi_{I_k\cap[a,x_0]}$, $f_2 := \sum_{k=1}^{K} c_k\chi_{I_k\cap[x_0,b]}$ Treppenfunktionen und es gilt $f = f_1 + f_2$. Damit folgt direkt die Behauptung für Treppenfunktionen.

Ist f beliebig, so wählen wir zu $\varepsilon > 0$ Treppenfunktionen e, g mit $e \le f \le g$ und $\int_a^b g\;dx - \int_a^b e\;dx < \varepsilon$. Dann gilt

$$0 \le \left(\int_a^{x_0} g\;dx - \int_a^{x_0} e\;dx\right) + \left(\int_{x_0}^b g\;dx - \int_{x_0}^b e\;dx\right) = \int_a^b g\;dx - \int_a^b e\;dx < \varepsilon,$$

also ist f R-integrabel auf $[a, x_0]$ bzw. $[x_0, b]$. Weiter haben wir

$$A := \int_a^b f\ dx - \left(\int_a^{x_0} f\ dx + \int_{x_0}^b f\ dx \right) \leq \int_a^b g\ dx - \left(\int_a^{x_0} e\ dx + \int_{x_0}^b e\ dx \right) < \varepsilon.$$

Analog erhält man die Abschätzung $A > -\varepsilon$ und damit ist $A = 0$. □

Satz 10.3.4 (Hauptsatz der Differential- und Integralrechnung)**.** *Sei $f \in C^0([a,b])$. Setze*

$$F : x \mapsto \int_a^x f(\xi)\ d\xi,\ x \in [a,b].$$

Dann gilt:

$$F \in C^1((a,b)) \quad \text{und } F' = f.$$

Beweis: Sei $x_0 \in (a,b)$ und zu $\varepsilon > 0$ wählen wir ein $\delta > 0$ so, dass für alle $x \in [a,b]$ gilt:

$$|x - x_0| < \delta \quad \Rightarrow \quad |f(x) - f(x_0)| < \varepsilon.$$

Sei $x_0 < x < x_0 + \delta$. Aus Satz 10.3.3 folgt

$$F(x) - F(x_0) = \int_a^x f(\xi)\ d\xi - \int_a^{x_0} f(\xi)\ d\xi = \int_{x_0}^x f(\xi)\ d\xi.$$

Korollar 10.3.1 impliziert

$$\left| \int_{x_0}^x f(\xi)\ d\xi - f(x_0)(x - x_0) \right| \leq |x - x_0| \sup_{|y - x_0| < \delta} |f(y) - f(x_0)| \leq \varepsilon(x - x_0),$$

wobei wir die Identität $f(x_0)(x - x_0) = \int_{x_0}^x f(x_0)\ d\xi$ benutzt haben. Kombiniert man die letzten beiden Abschätzungen, so erhält man

$$\left| \frac{F(x) - F(x_0)}{x - x_0} - f(x_0) \right| \leq \varepsilon.$$

Analog folgt die Abschätzung auch für $x - \delta < x < x_0$. Insgesamt sieht man damit

$$\sup_{|x - x_0| < \delta} \left| \frac{F(x) - F(x_0)}{x - x_0} - f(x_0) \right| \leq \varepsilon$$

und dies impliziert

$$F \in C^1(a,b) \quad \text{und} \quad F' = f.$$

□

Satz 10.3.5 (Mittelwertsatz der Integralrechnung)**.** *Seien $f, g : I = [a,b] \to \mathbb{R}$ stetig* *und* $g(x) \geq 0 \quad \forall\, x \in I$. *Dann existiert ein* $\xi \in [a,b]$ *mit*

$$\int\limits_a^b fg\ dx = f(\xi) \int\limits_a^b g\ dx.$$

Im Spezialfall $g \equiv 1$ *erhalten wir*

$$\int\limits_a^b f\ dx = f(\xi)(b-a).$$

Beweis: Da $g \geq 0$ impliziert Satz 10.3.1 das $\int\limits_a^b g\ dx \geq 0$. Jetzt unterscheiden wir zwei Fälle.

1. Fall: $\int\limits_a^b g\ dx = 0$

In diesem Fall behaupten wir das $g \equiv 0$ sein muss. Dazu nehmen wir an das ein $x_0 \in [a,b]$ existiert mit

$$g(x_0) =: \varepsilon > 0.$$

Da g stetig ist existiert ein $\delta > 0$ so, dass für alle $x \in B_\delta(x_0) \cap I$ gilt

$$g(x) \geq \frac{\epsilon}{2}.$$

Definiere jetzt eine Treppenfunktion $h:\ I \to \mathbb{R}$ durch

$$h(x) = \begin{cases} \frac{\varepsilon}{2}, & x \in B_\delta(x_0) \cap I \\ 0, & \text{sonst} \end{cases}.$$

Damit gilt $g(x) \geq h(x)$ für alle $x \in I$, aber aus satz 10.3.1 folgt

$$\int\limits_a^b g\ dx \geq \int\limits_a^b h\ dx \geq \frac{\epsilon}{2\delta} > 0.$$

Dies ist ein Widerspruch und deshalb ist $g \equiv 0$.

2. Fall: $\int\limits_a^b g\ dx > 0$

Wir setzen $m = \inf\limits_{x \in I} f(x)$ und $M = \sup\limits_{x \in I} f(x)$. Wegen Satz 7.2.1 sind m und M wohldefiniert. Für alle $x \in I$ gilt

$$mg(x) \leq f(x)g(x) \leq Mg(x)$$

und damit folgt aus Satz 10.3.1

$$m \int_a^b g \, dx \leq \int_a^b fg \, dx \leq M \int_a^b g \, dx$$

oder

$$m \leq \frac{\int_a^b fg \, dx}{\int_a^b g \, dx} \leq M.$$

Aus dem Zwischenwertsatz, Satz 7.1.1 folgt die Existenz eines $\xi \in I$ mit

$$f(\xi) = \frac{\int_a^b fg \, dx}{\int_a^b g \, dx}.$$

□

Satz 10.3.6. *Sei $f_n \in C^1(I)$ eine Folge von Funktionen auf $I = (a, b)$, die punktweise gegen f konvergiert. Falls die Folge f_n' gleichmässig gegen $g : I \to \mathbb{R}$ konvergiert, so ist $f \in C^1(I)$ mit $f' = g$.*

Beweis: Für $x, x_0 \in I$ gilt

$$f_n(x) = f_n(x_0) + \int_{x_0}^x f_n' \, dx.$$

Aus Korollar 10.3.2 folgt

$$\int_{x_0}^x f_n' \, dx \to \int_{x_0}^x g \, dx$$

und da $f_n(x) \to f(x)$ für alle $x \in I$ gilt, folgt insgesamt

$$f(x) = f(x_0) + \int_a^b g \, dx.$$

Die Behauptung folgt nun aus Satz 10.3.4. □

Satz 10.3.7. *Die Potenzreihe $p(x) = \sum_{k=0}^{\infty} c_k x^k$ habe den Konvergenzradius $\rho > 0$.*

Dann ist $p: (-\rho, \rho) \to \mathbb{R}$ differenzierbar mit Ableitung

$$p'(x) = \sum_{k=1}^{\infty} k c_k x^{k-1}.$$

Insbesondere ist $p \in C^\infty(-\rho, \rho)$.

Beweis: Sei $x_0 \in (-\rho, \rho)$, $|x_0| < R < \rho$ und sei $p_n(x) = \sum_{k=0}^{n} c_k x^k$. Für alle $x \in (-\rho, \rho)$ konvergiert $p_n(x)$ gegen $p(x)$. Wir definieren

$$q(x) := \sum_{k=1}^{\infty} \underbrace{k c_k}_{=d_k} x^{k-1}$$

und da

$$\rho = \limsup_{k \to \infty} \frac{1}{\sqrt[k]{|c_k|}} = \limsup_{k \to \infty} \frac{1}{\sqrt[k-1]{k\,|c_k|}}$$

hat q auch den Konvergenzradius ρ. Aus Satz 9.1.3 folgt damit das $p_n' \to q$ gleichmässig auf dem Intervall $[-R, R]$ konvergiert. Die Behauptung folgt aus Satz 10.3.6. $\square$

Beispiel:
Für alle $x \in (-1, 1)$ gilt:

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

Wir definieren die Potenzreihe $p(x) := \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{2n+1}$. Aus der Geleichung

$$\limsup_{n \to \infty} \frac{1}{\sqrt[n]{\left|\frac{1}{2n+1}\right|}} = 1$$

folgt das p den Konvergenzradius 1 hat. Satz 10.3.7 impliziert $p \in C^1((-1, 1))$ und

$$p'(x) = \sum_{n=0}^{\infty} (-1)^n x^{2n} = \arctan'(x) \quad \forall\, x \in (-1, 1).$$

Also existiert ein $c \in \mathbb{R}$ mit

$$p(x) = c + \arctan x \quad \forall\, x \in (-1, 1).$$

Es gilt $p(0) = 0 = \arctan 0$ und damit $c = 0$. Insgesamt erhalten wir damit eine Reihendarstellung für den arctan auf dem Intervall $(-1, 1)$:

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

Aus der Fehlerabschätzung des Leibniz-Kriteriums, Satz 4.4.5, folgt

$$\left|\arctan x - \sum_{n=0}^{N-1} (-1)^n \frac{x^{2n+1}}{2n+1}\right| \leq \frac{|x|^{2N+1}}{2N+1} \quad \forall\, N \in \mathbb{N}, \quad \forall\, x \in (-1,1).$$

Mit $x \nearrow 1$ und dann $N \to \infty$ folgt

$$\arctan 1 = \frac{\pi}{4} = \sum_{n=0}^{\infty} (-1)^n \frac{1}{2n+1} = 1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} \pm ...$$

10.4 Das R-Integral vektorwertiger Funktionen

Definition 10.4.1. *Eine Funktion $f = (f_1, ..., f_n) : [a,b] \to \mathbb{R}^n$ ist R-integrabel über $[a,b]$ genau dann, wenn $f_i : [a,b] \to \mathbb{R}$, $1 \leq i \leq n$ R-integrabel ist, und*

$$\int_a^b f\ dx := \left(\int_a^b f_1\ dx, ..., \int_a^b f_n\ dx \right).$$

Satz 10.4.1. *Sei $f \in C^0\left([a,b], \mathbb{R}^n\right)$. Dann gilt:*

$$\| \int_a^b f\ dx \| \leq \int_a^b \|f\|\ dx \leq (b-a)\|f\|_\infty.$$

Beweis: Setze

$$P := \int_a^b f\ dx \in \mathbb{R}^n.$$

Dann gilt mit Hilfe der Cauchy-Schwarz Ungleichung, Satz 3.1.2

$$\|P\|^2 = \int_a^b f\ dx \cdot P = \int_a^b (f \cdot P)\ dx \leq \|P\| \int_a^b \|f\|\ dx,$$

also folgt mit Hilfe von Korollar 10.3.1

$$\|P\| = \| \int_a^b f\ dx \| \leq \int_a^b \|f\|\ dx \leq (b-a)\|f\|_\infty.$$

□

10.5 Uneigentliches R-Integral

Sei $f : (a,b) \to \mathbb{R}$ über jedes kompakte Intervall $[c,d] \subset (a,b)$ R-integrabel.

Definition 10.5.1. *f heisst über (a, b) uneigentlich R-integrabel, falls*

$$\int_a^b f\ dx := \lim_{c \searrow a, d \nearrow b} \int_c^d f\ dx$$

existiert.

Beispiel:

1) Für $\alpha < -1$ existiert

$$\int_1^\infty x^\alpha = \lim_{d\to\infty} \int_1^d x^\alpha\ dx = \lim_{d\to\infty} \left[\frac{x^{\alpha+1}}{\alpha+1}\right]_{x=1}^{x=d} = \lim_{d\to\infty} \left(\frac{d^{\alpha+1}}{\alpha+1} - \frac{1}{\alpha+1}\right) = -\frac{1}{\alpha+1}.$$

2) Für $\alpha > -1$ existiert

$$\int_0^1 x^\alpha\ dx = \lim_{c\searrow 0} \int_c^1 x^\alpha\ dx = \lim_{c\searrow 0} \left[\frac{x^{\alpha+1}}{\alpha+1}\right]_{x=c}^{x=1} = \lim_{c\searrow 0} \left(\frac{1}{\alpha+1} - \frac{c^{\alpha+1}}{\alpha+1}\right) = \frac{1}{\alpha+1}.$$

3) $\displaystyle\int_0^1 \frac{dx}{x} = \lim_{c\to 0} \int_c^1 \frac{dx}{x} = \lim_{c\to 0}(-\log c)$ existiert nicht.

4) $\displaystyle\int_0^\infty e^{-t}\ dt = \lim_{d\to\infty} \int_0^d e^{-t}\ dt = \lim_{d\to\infty} \left[-e^{-t}\right]_{t=0}^{t=1} = \lim_{d\to\infty} (1 - e^{-d}) = 1.$

Satz 10.5.1. *Sei $f : [1, \infty) \to \mathbb{R}^+ = \{x \in \mathbb{R} : x > 0\}$ monoton fallend. Dann konvergiert die Reihe $\sum_{k=1}^\infty f(k)$ genau dann, wenn das Integral $\int_1^\infty f\ dx$ konvergiert und es gilt*

$$0 \leq \sum_{k=1}^\infty f(k) - \int_1^\infty f\ dx \leq f(1).$$

Beweis: Wir definieren die Funktionen $e = \sum_{k=1}^\infty f(k+1)\chi_{[k,k+1)}$ und $g = \sum_{k=1}^\infty f(k)\chi_{[k,k+1]}$. Auf dem Intervall $[1, \infty)$ gilt $e \leq f \leq g$ und auf dem Intervall $[k, k+1)$ haben wir weiter

$$e = f(k+1) \leq f(x) \leq f(k) = g.$$

Dies impliziert die Abschätzung

$$\begin{aligned}
\int\limits_1^n e\;dx &= \sum_{k=1}^{n-1} f(k+1) = \sum_{k=1}^{n} f(k) - f(1) \\
&\leq \int\limits_1^n f\;dx \leq \int\limits_1^n g\;dx \\
&= \sum_{k=1}^{n-1} f(k) = \sum_{k=1}^{n} f(k) - f(n)
\end{aligned}$$

und damit folgt

$$0 < f(n) \leq \sum_{k=1}^{n} f(k) - \int\limits_1^n f\;dx \leq f(1).$$

Mit $n \to \infty$ folgt daraus die Behauptung. $\square$

Beispiel:

1) Die Riemannsche Zeta-Funktion ist definiert durch $\zeta(s) = \sum\limits_{k=1}^{\infty} k^{-s},\ \zeta : (1,\infty) \to \mathbb{R}$.
Die Funktion $f(x) := x^{-s},\ f : [1,\infty) \to \mathbb{R}^+$ ist monoton fallend, denn $f' < 0$. Durch eine einfache Rechnung erhalten wir für alle $1 < s < \infty$

$$\int\limits_1^n x^{-s}\;dx = \frac{1}{1-s}\left[x^{1-s}\right]_{x=1}^{x=n} = \frac{n^{1-s}}{1-s} - \frac{1}{1-s}$$

und mit $n \to \infty$ folgt also

$$\lim_{n\to\infty} \int\limits_1^n x^{-s}\;dx = -\frac{1}{1-s} = \frac{1}{s-1}.$$

Aus Satz 10.5.1 folgt nun

$$0 \leq \zeta(s) - \int\limits_1^\infty x^{-s}\;dx = \zeta(s) - \frac{1}{s-1} \leq f(1) = 1.$$

2) Wir wollen die Reihe $\sum\limits_{k=2}^{\infty} \frac{1}{k(\log k)^s},\ \ s > 1$ auf Konvergenz untersuchen. Die Funktion $f(x) = \frac{1}{x(\log x)^s},\ f\,[2,\infty) \to \mathbb{R}^+$ ist monoton fallend und wir berechnen

$$\lim_{n\to\infty} \int\limits_2^n f(x)\;dx = \lim_{n\to\infty} \int\limits_2^n \frac{1}{x(\log x)^s}\;dx = \lim_{n\to\infty} \int\limits_{\log 2}^{\log n} y^{-s}\;dy < \infty.$$

Wegen Satz 10.5.1 konvergiert die Reihe also.

11 Taylor–Reihen

11.1 Taylor-Polynome

Definition 11.1.1. *Sei $I \subset \mathbb{R}$ ein Intervall und sei $f \in C^{(n)}(I)$ $(n \in \mathbb{N}_0)$. Das Taylor–Polynom n–ter Ordnung von f im Punkt $x_0 \in I$ ist gegeben durch*

$$(T^n_{x_0} f)(x) := \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}\,(x-x_0)^k.$$

Bemerkung: $T^n_{x_0} f$ ist das eindeutig bestimmte Polynom, welches

$$f^{(k)}(x_0) = (T^n_{x_0})^{(k)}(x_0) \quad \forall\ 0 \le k \le n$$

erfüllt.

Satz 11.1.1 (Taylor)**.** *Sei $f \in C^{(n+1)}(I)$. Dann gilt für alle $x_0, x \in I$:*

$$f(x) = (T^n_{x_0} f)(x) + R_{n+1}(x),$$

wobei

$$R_{n+1}(x) = \frac{1}{n!} \int_{x_0}^{x} (x-t)^n f^{(n+1)}(t)dt.$$

Beweis. Induktion nach n

$\underline{n=0}$: f ist Stammfunktion zu f' und damit folgt

$$f(x) = f(x_0) + \int_{x_0}^{x} f'(t)dt = (T^0_{x_0} f)(x) + R_1(x).$$

$\underline{n-1 \to n}$:

Nach Induktionsvoraussetzung ist $f(x) = (T^{n-1}_{x_0} f)(x) + R_n(x)$ mit

$$R_n(x) = \frac{1}{(n-1)!} \int_{x_0}^{x} (x-t)^{n-1} f^{(n)}(t)dt.$$

Jetzt gilt mit Hilfe von partieller Integration

$$\begin{aligned}R_n(x) &= -\frac{1}{n!}[(x-t)^n f^{(n)}(t)]_{t=x_0}^{t=x} + \frac{1}{n!}\int_{x_0}^{x_0}(x-t)^n f^{(n+1)}(t)dt \\ &= \frac{1}{n!}(x-x_0)^n f^{(n)}(x_0) + R_{n+1}(x)\end{aligned}$$

und damit $f(x) = (T_{x_0}^n f)(x) + R_{n+1}(x)$. □

Satz 11.1.2 (Lagrange)**.** *Sei $f \in C^{n+1}(I)$ $(n \in \mathbb{N}_0)$ und $x_0, x \in I$. Dann existiert ein ξ zwischen x_0 und x mit*

$$f(x) = (T_{x_0}^n f)(x) + \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-x_0)^{n+1}.$$

Beweis. Sei ohne Einschränkung $x_0 < x$. Aus Satz 10.3.5 folgt die Existenz eines $\xi \in [x_0, x]$ mit

$$\begin{aligned}R_{n+1} &= \frac{1}{n!}\int_{x_0}^{x}(x-t)^n f^{(n+1)}(t)dt = f^{(n+1)}(\xi)\int_{x_0}^{x}\frac{(x-t)^n}{n!}dt \\ &= -f^{(n+1)}(\xi)\Big[\frac{(x-t)^{n+1}}{(n+1)!}\Big]_{t=x_0}^{t=x} \\ &= f^{(n+1)}(\xi)\ \frac{(x-x_0)^{n+1}}{(n+1)!}.\end{aligned}$$

□

Lemma 11.1.1. *Sei $f \in C^{(n+1)}(I)$ mit $f^{(n+1)}(x) = 0$ für alle $x \in I$. Dann gilt*

$$f(x) = (T_{x_0}^n f)(x) \qquad \forall\ x \in I.$$

Beweis. Folgt aus Satz 11.1.2. □

Lemma 11.1.2. *Sei $f \in C^{(n)}(I)$ $(n \in \mathbb{N})$ und sei $x_0 \in I$. Dann gilt:*

$$f(x) = (T_{x_0}^n f)(x) + \eta(x)(x-x_0)^n \qquad \forall\ x \in I,$$

wobei $\eta : I \to \mathbb{R}$ eine Funktion mit $\lim\limits_{x \to x_0} \eta(x) = 0$ ist.

Beweis. Sei $x \in I$. Aus Satz 11.1.2 folgt die Existenz von ξ zwischen x_0 und x mit

$$\begin{aligned}f(x) &= (T_{x_0}^{n-1} f)(x) + \frac{f^{(n)}(\xi)}{n!}\ (x-x_0)^n \\ &= (T_{x_0}^n f)(x) + \frac{f^{(n)}(\xi) - f^{(n)}(x_0)}{n!}(x-x_0)^n.\end{aligned}$$

Definiere nun $\eta(x) := \dfrac{f^{(n)}(\xi) - f^{(n)}(x_0)}{n!}$ (ξ hängt von x ab!)

Aus der Stetigkeit von $f^{(n)}$ und aus der Tatsache das mit $x \to x_0$ auch $\xi \to x_0$ gilt, folgt:

$$\lim_{x \to x_0} \eta(x) = 0.$$

$\square$

Beispiele:

1) $f(x) = e^x$, $x_0 = 0$. Es gilt

$$f^{(k)}(x) = e^x \quad \forall\, k \in \mathbb{N}_0 \implies f^{(k)}(0) = 1 \quad \forall\, k \in \mathbb{N}_0\,.$$

Damit erhalten wir $(T_0^n f)(x) = \sum\limits_{k=0}^{n} \dfrac{x^k}{k!} \quad \forall\, n \in \mathbb{N}_0$.

2) Sei $f : (-1, 1) \to \mathbb{R}$ definiert durch $f(x) = \sqrt{1+x}$ und sei $x_0 = 0$. Wir berechnen $f(0) = 1$, $f'(0) = \dfrac{1}{2}$. Aus dem obigen Lemma folgt damit

$$f(x) = \sqrt{1+x} = 1 + \frac{x}{2} + \eta(x)x$$

wobei $\eta(x) \to 0$ für $x \to 0$.

Also gilt für alle $n > 1$

$$\begin{aligned} \sqrt{n + \sqrt{n}} &= \sqrt{n\Big(1 + \frac{1}{\sqrt{n}}\Big)} = \sqrt{n}\,\sqrt{1 + \frac{1}{\sqrt{n}}} \\ &= \sqrt{n}\Big(1 + \frac{1}{2\sqrt{n}} + \eta\Big(\frac{1}{\sqrt{n}}\Big)\frac{1}{\sqrt{n}}\Big) \\ &= \sqrt{n} + \frac{1}{2} + \eta\Big(\frac{1}{\sqrt{n}}\Big) \end{aligned}$$

und somit schliessen wir

$$\lim_{n \to \infty} (\sqrt{n + \sqrt{n}} - \sqrt{n}) = \frac{1}{2}.$$

3) Sei $f : (-1, \infty) \to \mathbb{R}$, $f(x) = \log(1+x)$ und sei $x_0 = 0$. Induktiv berechnen wir $f^{(k)}(x) = (-1)^{k-1}(k-1)!(1+x)^{-k} \quad \forall\, k \in \mathbb{N}$, also $f(0) = 0$, $f^{(k)}(0) = (-1)^{k-1}(k-1)! \quad \forall\, k \in \mathbb{N}$. Damit folgt für alle $n \in \mathbb{N}$

$$(T_0^n f)(x) = \sum_{k=1}^{n} (-1)^{k-1} \frac{x^k}{k}.$$

Satz 11.1.3. *Sei $n \in \mathbb{N}$, $f : I \to \mathbb{R}$, $f \in C^{(n)}(I)$ und für $x_0 \in I$ gelte*

$$f'(x_0) = f''(x_0) = \ldots = f^{(n-1)}(x_0) = 0 \quad \textit{und} \quad f^{(n)}(x_0) \neq 0.$$

Ist n ungerade, so ist x_0 keine Extremalstelle.

Ist n gerade, so ist x_0 ein Minimum falls $f^{(n)}(x_0) > 0$ und ein Maximum, falls $f^{(n)}(x_0) < 0$.

Beweis. Sei $x \in I$. Aus Satz 11.1.2 folgt die Existenz von ξ zwischen x_0 und x mit

$$f(x) = (T_{x_0}^{n-1} f)(x) + \frac{f^{(n)}(\xi)}{n!}(x - x_0)^n = f(x_0) + \frac{f^{(n)}(\xi)}{n!}(x - x_0)^n$$

$f^{(n)}$ ist stetig und $f^{(n)}(x_0) \neq 0 \implies \exists\, \delta > 0$ mit $f^{(n)}(x) \neq 0\ \forall\, x \in B_\delta(x_0) \cap I$.

1) n **ungerade** und $f^{(n)}(x_0) > 0$

$$\implies f(x) - f(x_0) = f^n(\xi)(x - x_0)^n \begin{cases} > 0 & \text{für } x \in B_\delta(x_0) \cap I, x > x_0 \\ < 0 & \text{für } x \in B_\delta(x_0) \cap I, x < x_0 \end{cases}$$

(Im Fall $f^{(n)}(x_0) < 0$ erhält man die umgekehrten Ungleichungen!)

$\implies f$ hat kein Extremum in x_0.

2) n **gerade**

$$\implies f(x) - f(x_0) = f^{(n)}(\xi)(x - x_0)^n \begin{cases} > 0 & \text{falls } f^{(n)}(\xi) > 0, x \in B_\delta(x_0) \cap I \\ < 0 & \text{falls } f^{(n)}(\xi) < 0, x \in B_\delta(x_0) \cap I \end{cases}$$

Damit hat f ein Minimum (Maximum) in x_0 falls $f^{(n)}(x_0) > 0$ ($f^{(n)}(x_0) < 0$). □

11.2 Taylor-Reihen

Definition 11.2.1. *Sei $f \in C^\infty(I)$ und $x_0 \in I$. Dann heisst*

$$(T_{x_0} f)(x) := \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!}\ (x - x_0)^k$$

die Taylor–Reihe von f in x_0.

Bemerkung: Für $f \in C^\infty(I)$ und $x_0 \in I$ gilt:

$$f = T_{x_0} f \quad \Leftrightarrow \quad \lim_{n\to\infty} (f(x) - (T_{x_0}^n f)(x)) = \lim_{n\to\infty} R_{n+1}(x) = 0 \quad \forall x \in I.$$

Satz 11.2.1. *Sei $f \in C^0(I)$ und sei $x_0 \in I$. Dann sind äquivalent:*

1) f ist durch eine konvergente Potenzreihe auf I mit Entwicklungspunkt x_0 darstellbar:

$$f(x) = \sum_{k=0}^{\infty} a_k (x - x_0)^k \qquad \forall\ x \in I.$$

2) $f \in C^\infty(I)$ und es gilt für alle $x \in I$:

$$f(x) - (T_{x_0}^n f)(x) \to 0 \qquad \textit{für} \ \ n \to \infty.$$

Beweis. **1) $\Longrightarrow$ 2)**:

Aus Satz 10.3.7 folgt direkt $f \in C^\infty(I)$ und $f^{(k)}(x_0) = k! a_k$. Damit gilt

$$\begin{aligned} R_{n+1}(x) &= f(x) - \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k \\ &= f(x) - \sum_{k=0}^{n} a_k (x - x_0)^k \to 0 \ \text{ für } n \to \infty. \end{aligned}$$

2) $\Longrightarrow$ 1):

Es gilt $f(x) = \sum_{k=0}^{n} \dfrac{f^{(k)}(x_0)}{k!} (x - x_0)^k + \underbrace{R_{n+1}(x)}_{\to 0 \text{ mit } n \to \infty}$ und damit

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k.$$

□

Beispiel: Die Taylor–Reihen von e^x, $\sin x$, $\cos x$, $\arctan x$, $\log(1 + x)$ sind auf ihren jeweiligen Definitionsbereichen gegeben durch

$$\begin{aligned} &e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}, \ \cos x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!}, \ \sin x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!}, \\ &\arctan x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{2k+1}, \\ &\log(1 + x) = \sum_{k=1}^{\infty} (-1)^{k-1} \frac{x^k}{k}. \end{aligned}$$

Frage: Sei $f \in C^\infty(I)$, $x_0 \in I$ und gelte $(T_{x_0}^n f)(x) \to (T_{x_0} f)(x) \quad \forall\ x \in I$. Gilt dann $f = T_{x_0} f$?

Beispiel: Die Antwort ist Nein! Dazu sei $f : \mathbb{R} \to \mathbb{R}$ gegeben durch

$$f(x) = \begin{cases} e^{-\frac{1}{x}} & \text{für } x > 0 \\ 0 & \text{für } x \le 0 \end{cases}$$

In den Aufgaben wurde gezeigt:

$$f \in C^\infty(\mathbb{R}) \quad \text{und} \quad f^{(k)}(0) = 0 \quad \forall\, k \in \mathbb{N}\,.$$

Damit folgt $(T_0^n f)(x) = 0$ für alle $n \in \mathbb{N}_0$ und alle $x \in \mathbb{R}$. Also gilt $(T_0 f)(x) = \lim\limits_{n\to\infty} (T_0^n f)(x) = 0$ für alle $x \in \mathbb{R}$, aber $f(x) > 0$ für alle $x > 0$.

12 Untervektorräume

12.1 Untervektorräume

Erinnerung Vektorraum (s. Kapitel 3)

Sei $\mathbb{K}$ ein Körper. Ein $\mathbb{K}$–Vektorraum $(V, +, \cdot)$ ist eine Menge V auf der eine Addition $+$ und skalare Multiplikation $\cdot$ definiert sind, so dass gilt:

- $x + (y + z) = (x + y) + z \qquad \forall\ x, y, z \in V$
- $x + y = y + x \qquad \forall\ x, y \in V$
- $\exists\ 0 \text{mit } x + 0 = x \qquad \forall\ x \in V$
- $\forall\ x \in V \quad \exists\ y \in V$ ist $x + y = 0$
- $a(x + y) = ax + ay \qquad \forall\ x, y \in V;\ a \in \mathbb{K}$
- $(a + b)x = ax + bx \qquad \forall\ x \in V; a, b \in \mathbb{K}$
- $(ab)x = a(bx)$
- $1x = x$, wobei 1 das neutrale Element der Multiplikation in $\mathbb{K}$ ist.

Beispiele:

1) $\mathbb{R}^n$, $\mathbb{C}$ sind $\mathbb{R}$–Vektorräume

2) $C^k(I, \mathbb{R}^n), C^\infty(I, \mathbb{R})$, $R(I)$, $B(I)$ sind $\mathbb{R}$–Vektorräume

3) Sind $V_1, \ldots, V_n$ $(n \in \mathbb{N})$ $\mathbb{K}$-Vektorräume, so ist auch

$$V := V_1 \times \ldots \times V_n = \{v_1, \ldots, v_n) : v_i \in V_i \quad \forall\ 1 \leq i \leq n\}$$

ein $\mathbb{K}$-Vektorraum. V heisst das direkte Produkt der $V_1, \ldots, V_n$.

4) $\mathbb{K}^n = \{(\alpha_1, \ldots, \alpha_n) : \alpha_i \in \mathbb{K} \qquad \forall\ 1 \leq i \leq n\}$ (folgt aus 3)

Definition 12.1.1. *Sei V ein $\mathbb{K}$–Vektorraum. Eine Teilmenge $U \subset V$ heisst ein $\mathbb{K}$–Untervektorraum von V, falls gilt:*

1) $U \neq \emptyset$

2) $a, b \in U \implies a + b \in U$

3) $\alpha \in \mathbb{K}$, $a \in U \implies \alpha\, a \in U$

Lemma 12.1.1. *Eine Teilmenge U eines $\mathbb{K}$–Vektorraums V ist genau dann ein $\mathbb{K}$–Untervektorraum, wenn er mit der aus V induzierten Addition und skalaren Multiplikation selbst ein Vektorraum ist.*

Beweis. „$\Longrightarrow$“: 2) + 3) $\implies$ U abgeschlossen unter der Addition und skalaren Multiplikation

1) $\implies$ $U \neq \emptyset \implies \exists\, a \in U \implies -a = (-1)a \in U \implies 0_V = a - a \in U$

Die anderen Eigenschaften sind leicht nachzurechnen.

„$\Longleftarrow$“: Klar! □

Beispiele:

1) $\{0\}$, V sind Untervektorräume von V.

2) Sei $V = \{f : I \to \mathbb{R}^n\}$, I Intervall.
$\implies R(I), C^k(I), C^\infty(I)$ sind Untervektorräume von V

Lemma 12.1.2. *Sei V ein $\mathbb{K}$–Vektorraum und sei $(U_i)_{i\in I}$ eine Familie von Untervektorräumen von V. Dann ist $U := \bigcap_{i\in I} U_i$ ebenfalls ein Untervektorraum von V.*

Beweis.

- $0 \in U_i \quad \forall\, i \in I \implies 0 \in U \quad \implies U \neq \emptyset$
- $a, b \in U \implies a, b \in U_i \quad \forall\, i \in I \implies a + b \in U_i \quad \forall i \in I \implies a + b \in U$
- $\alpha \in \mathbb{K}$, $a \in U \implies a \in U_i \quad \forall\, i \in I \implies \alpha\, a \in U_i \quad \forall\, i \in I \implies \alpha\, a \in U$

□

Definition 12.1.2. *Sei V ein $\mathbb{K}$–Vektorraum und seien $v_1, \ldots, v_m \in V$ $(m \in \mathbb{N})$. Ein Vektor $v \in V$ heisst Linearkombination der $v_i, 1 \leq i \leq m$, falls $\alpha_1, \ldots, \alpha_m \in \mathbb{K}$ existieren, mit*

$$v = \sum_{i=1}^{m} \alpha_i v_i.$$

Die Menge

$$\mathrm{Span}_{\mathbb{K}}(v_1, \ldots, v_m) := \{\alpha_1 v_1 + \ldots + \alpha_m v_m : \alpha_i \in \mathbb{K} \quad \forall\, 1 \leq i \leq m\}$$

heisst der von $v_1, \ldots, v_m$ aufgespannte Raum.

Wir setzen $\mathrm{Span}_{\mathbb{K}}(\emptyset) = \{0\}$.

Bemerkung: Sei V ein $\mathbb{K}$–Vektorraum, $U \subset V$ ein Untervektorraum und seien $v_1, \ldots, v_m \in U$ $(m \in \mathbb{N})$. Dann gilt für alle $\alpha_i \in \mathbb{K}$, $1 \leq i \leq m$:

$$\alpha_1 v_1 + \ldots + \alpha_m v_m \in U.$$

Satz 12.1.1. *In einem $\mathbb{K}$–Vektorraum V seien $v_1, \ldots, v_m \in V$ $(m \in \mathbb{N}_0)$ gegeben. Dann gilt*

1) $W := \mathrm{Span}_{\mathbb{K}}(v_1, \ldots, v_m) \subset V$ ist ein Untervektorraum

2) W ist der kleinste Untervektorraum von V, der $v_1, \ldots, v_m$ enthält.

Beweis. 1) $0 = \sum\limits_{i=1}^{m} 0 v_i \in W \quad \Longrightarrow W \neq \emptyset$.

Für $v = \sum\limits_{i=1}^{m} \alpha_i v_i \in W$ und $w = \sum\limits_{i=1}^{m} \beta_i v_i \in W$ gilt: $v + w = \sum\limits_{i=1}^{m} (\alpha_i + \beta_i) v_i \in W$. Ausserdem gilt für $\alpha \in \mathbb{K}$: $\alpha\, v = \sum\limits_{i=1} (\alpha\, \alpha_i) v_i \in W$.

2) Es gilt $v_i = 0v_1 + \ldots + 0v_{i-1} + 1v_i + 0v_{i+1} + \ldots + 0v_m \in W \qquad \forall\ 1 \leq i \leq m$

Sei nun $W' \subset V$ ein Untervektorraum mit $v_1, \ldots, v_m \in W'$

$\Longrightarrow \sum\limits_{i=1}^{m} \alpha_i v_i \in W' \qquad \forall\ \alpha_i \in \mathbb{K}, 1 \leq i \leq m \Longrightarrow\ W \subset W'$.

$\square$

Definition 12.1.3. *Sei V ein $\mathbb{K}$-Vektorraum und seien $(v_1, \ldots, v_n) \in V^n$.*

$(v_1, \ldots, v_n)$ heissen linear unabhängig, falls aus

$$0 = \alpha_1 v_1 + \alpha_n v_n \ \textit{ folgt, dass } \alpha_i = 0 \quad \forall\ 1 \leq i \leq n.$$

Ansonsten heissen die $(v_1, \ldots, v_n)$ linear abhängig.

Satz 12.1.2. *Sei V ein $\mathbb{K}$–Vektorraum. $(v_1, \ldots, v_n) \in V^n$ sind linear unabhängig genau dann, wenn aus*

$$v = \sum_{i=1}^{n} \alpha_i v_i = \sum_{i=1}^{n} \beta_i v_i \ \textit{ folgt}: \ \alpha_i = \beta_i \quad \forall\ 1 \leq i \leq n.$$

Beweis. „$\Longrightarrow$“: Aus $v = \sum\limits_{i=1}^{n} \alpha_i v_i = \sum\limits_{i=1}^{n} \beta_i v_i$ folgt $\sum\limits_{i=1}^{n} (\alpha_i - \beta_i) v_i = v - v = 0$. Da $(v_1, \ldots, v_n)$ linear unabhängig sind, impliziert dies $\alpha_i = \beta_i$ für alle $1 \leq i \leq n$.

„$\Longleftarrow$“: Sei $\alpha_1 v_1 + \ldots + \alpha_n v_n = 0 = 0v_1 + \ldots + 0v_n$. Aus der Eindeutigkeit der Darstellung als Linearkombination folgt:

$$\alpha_i = 0 \qquad \forall\ 1 \le i \le n.$$

$\square$

Beispiele:

1) Sei $V = \mathbb{R}[X] = \{\text{Polynome mit reellen Koeffizienten}\}$. Wir definieren

$$p_0(x) = 1, p_1(x) = x, \ldots, p_n(x) = x^n \ \text{ für alle } \ x.$$

Dann sind $(p_0, \ldots, p_n)$ linear unabhängig für alle $n \in \mathbb{N}_0$.

Denn aus

$$0 = \alpha_0 p_0 + \ldots + \alpha_n p_n$$

folgt mit Lemma 3.3.2:

$$\alpha_i = 0 \qquad \forall\ 0 \le i \le n.$$

2) $\mathbb{R}$ ist ein $\mathbb{Q}$–Vektorraum und $1, \sqrt{2}$ sind linear unabhängig, denn aus

$$\begin{aligned} &\lambda_1 1 + \lambda_2 \sqrt{2} = 0_{\mathbb{R}} \ \text{ mit } \lambda_1, \lambda_2 \in 0 \ \text{ folgt für } \lambda_2 \neq 0 \\ &\sqrt{2} = -\frac{\lambda_1}{\lambda_2} \in \mathbb{Q} \quad \lightning \quad \Longrightarrow \quad \lambda_2 = 0 \Longrightarrow \lambda_1 = 0 \end{aligned}$$

Lemma 12.1.3. *Sei V ein $\mathbb{K}$–Vektorraum. $(v_1, \ldots, v_n) \in V^n$ sind genau dann linear abhängig, wenn ein $i \in \{1, \ldots, n\}$ existiert mit*

$$v_i \in \operatorname{Span}_{\mathbb{K}}(v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n).$$

Beweis. „$\Longrightarrow$“: $\exists\ i \in \{1, \ldots, n\}$ und $\alpha_i \in \mathbb{K}$, $\alpha_i \neq 0$ mit $\sum\limits_{j=1}^{n} \alpha_j v_j = 0$, also

$$\begin{aligned} \Longrightarrow \ v_i &= -\frac{\alpha_1}{\alpha_i} v_1 - \ldots - \frac{\alpha_{i-1}}{\alpha_i} v_{i-1} - \frac{\alpha_{i+1}}{\alpha_i} v_{i+1} - \ldots - \frac{\alpha_n}{\alpha_i} v_n \\ &\in \operatorname{Span}_{\mathbb{K}}(v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n) \end{aligned}$$

„$\Longleftarrow$“: $\exists\ \alpha_1, \ldots, \alpha_{i-1}, \alpha_{i-1}, \ldots, \alpha_n \in \mathbb{K}$ mit

$$\begin{aligned} &v_i = \alpha_1 v_1 + \ldots + \alpha_{i-1} v_{i-1} + \alpha_{i+1} v_{i+1} + \ldots + \alpha_n v_n \\ &\Longrightarrow \quad \text{für } \alpha_i := -1 \neq 0 \text{ gilt: } \quad \sum_{i=1}^{n} \alpha_i v_i = 0 \end{aligned}$$

$\square$

12.2 Basis und Dimension

Definition 12.2.1. *Sei V ein $\mathbb{K}$–Vektorraum. $(v_1, \ldots, v_n) \in V^n$ heisst endliches Erzeugendensystem von V, falls $\mathrm{Span}_{\mathbb{K}}(v_1, \ldots, v_n) = V$.*

V heisst endlich erzeugt, falls ein endliches Erzeugendensystem von V existiert.

Ein endliches Erzeugendensystem $(v_1, \ldots, v_n)$ heisst Basis, falls $(v_1, \ldots, v_n)$ linear unabhängig sind.

Beispiele:

1) $\mathbb{R}^n = \mathrm{Span}_{\mathbb{R}}(e_1, \ldots, e_n)$ und $(e_1, \ldots, e_n)$ linear unabhängig $\Longrightarrow$ $(e_1, \ldots, e_n)$ ist eine Basis von $\mathbb{R}^n$.

2) $V = \mathbb{R}[X]$ ist nicht endlich erzeugt, denn seien $p_1, \ldots, p_n \in \mathbb{R}[x]$ Polynome vom Grad d_i, $1 \leq i \leq n$.

   $\Longrightarrow$ $\alpha_1 p_1 + \ldots + \alpha_n p_n, \alpha_i \in \mathbb{R}$, ist ein Polynom vom Grad $\leq d = \max\limits_{1 \leq i \leq n} d_i$.

   $\Longrightarrow$ Polynome vom Grad $> d$ sind nicht in $\mathrm{Span}_{\mathbb{R}}(p_1, \ldots, p_n)$.

Satz 12.2.1 (Basis-Auswahlsatz)**.** *Sei V ein $\mathbb{K}$–Vektorraum und es gelte*

$$V = \mathrm{Span}_{\mathbb{K}}(v_1, \ldots, v_n) \quad \text{mit } v_i \in V, \quad \forall\, 1 \leq i \leq n.$$

Dann gibt es ein $m \in \mathbb{N}$ mit $0 \leq m \leq n$ und Indizes $i_1, \ldots, i_m \in \{1, \ldots, n\}$, so dass

$$B := (v_{i_1}, \ldots, v_{i_m})$$

eine Basis von V ist.

Speziell besitzt jeder endlich erzeugte Vektorraum eine endliche Basis.

Beweis. Ist $(v_1, \ldots, v_n)$ linear unabhängig, so wählen wir $m = n$. Ist $(v_1, \ldots, v_n)$

linear abhängig, so existiert nach Lemma 12.1.3 ein $i \in \{1, \ldots, n\}$ mit

$v_i \in \mathrm{Span}_{\mathbb{K}}(v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n)$, also

$$\mathrm{Span}_{\mathbb{K}}(v_1, \ldots, v_n) = \mathrm{Span}_{\mathbb{K}}(v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n).$$

Ist $(v_1, \ldots, v_{i-1}, v_{i+1}, \ldots, v_n)$ linear unabhängig, so sind wir fertig.

Andernfalls wiederholen wir den vorherigen Schritt.

Diese Iteration führt nach höchstens n–Schritten zum Ziel.

Denn falls wir bei $V = \mathrm{Span}_{\mathbb{K}}(v_j)$, $1 \leq j \leq n$, angekommen sind, dann gilt entweder

- $v_j \neq 0 \implies m = 1$, oder
- $v_j = 0 \implies m = 0$ und $\mathrm{Span}_{\mathbb{K}}(\emptyset) = \{0\}$.

$\square$

Lemma 12.2.1. *Sei V ein $\mathbb{K}$–Vektorraum und sei $B := (v_1, \ldots, v_n)$ eine Basis von V. Sei weiter $w \in V$ mit $w \neq 0$ und es gelte*

$$w = \beta_1 v_1 + \ldots + \beta_n v_n.$$

Ist $\beta_i \neq 0$ für ein $1 \leq i \leq n$, so ist

$$B' := (v_1, \ldots, v_{i-1}, w, v_{i+1}, \ldots, v_n)$$

eine Basis von V.

Beweis. Ohne Einschränkung sei $i = 1$

Schritt 1: $V = \mathrm{Span}_{\mathbb{K}} B'$

„ $\subset$ “: Ist $v \in V$, so gilt $v = \alpha_1 v_1 + \ldots + \alpha_n v_n$ mit $\alpha_i \in \mathbb{K}$. Aus

$$\begin{aligned}
&v_1 = \frac{1}{\beta_1} w - \frac{\beta_2}{\beta_1} v_2 - \ldots - \frac{\beta_n}{\beta_1} v_n \quad \text{folgt damit} \\
&v = \frac{\alpha_1}{\beta_1} w + \Big(\alpha_2 - \frac{\alpha_2 \beta_2}{\beta_1}\Big) v_2 + \ldots + \Big(\alpha_n - \frac{\alpha_1 \beta_n}{\beta_1}\Big) v_n \\
&\implies v \in \mathrm{Span}_{\mathbb{K}} B'
\end{aligned}$$

„ $\supset$ “:

$$\begin{aligned}
v \in \mathrm{Span}_{\mathbb{K}} B' \implies v &= \alpha_1 w + \alpha_2 v_2 + \ldots + \alpha_n v_n \\
&= \alpha_1 \beta_1 v_1 + (\alpha_2 + \alpha_1 \beta_2) v_2 + \ldots + (\alpha_n + \alpha_1 \beta_n) v_n \\
&\in \mathrm{Span}_{\mathbb{K}} B = V
\end{aligned}$$

Schritt 2: B' linear unabhängig

Dazu sei $\alpha_1 w + \alpha_2 v_2 + \ldots + \alpha_n v_n = 0$
$\implies \alpha_1 \beta_1 v_1 + (\alpha_2 + \alpha_1 \beta_2) v_2 + \ldots + (\alpha_n + \alpha_1 \beta_n) v_n = 0$

B linear unabhängig
$\implies \alpha_1 \beta_1 = \alpha_2 + \alpha_1 \beta_2 = \ldots = \alpha_n + \alpha_1 \beta_n = 0$

Aus $\beta_1 \neq 0$ folgt $\alpha_1 = 0 \implies \alpha_2 = 0, \ldots, \alpha_n = 0$

$\implies B'$ linear unabhängig $\square$

Satz 12.2.2 (Basis–Austauschsatz)**.** *Sei V ein $\mathbb{K}$–Vektorraum und sei $B := (v_1, \ldots, v_n)$ eine Basis von V. Weiter seien $(w_1, \ldots, w_m) \in V^m$ linear unabhängig. Dann muss $m \leq n$ sein und man kann $i_1, \ldots, i_m \in \{1, \ldots, n\}$ finden, so dass in der Basis die Vektoren $v_{i_1}, \ldots, v_{i_m}$ durch $w_1, \ldots, w_m$ ersetzt werden können.*

Ist $i_1 = 1, \ldots, i_m = m$ gewählt, so ist

$$B' := (w_1, \ldots, w_m, v_{m+1}, \ldots, v_n)$$

wieder eine Basis von V.

Beweis. Induktion über m

$\boldsymbol{m = 0}$: In diesem Fall ist nichts auszutauschen.

$\boldsymbol{m = 1}$: $w_1 \neq 0 \implies V \neq \{0\}$, also $n \geq 1$ und die Behauptung folgt aus Lemma 12.2.1.

$\boldsymbol{m - 1 \to m}$: Ist $(w_1, \ldots, w_m)$ linear unabhängig, so ist auch $(w_1, \ldots, w_{m-1})$ linear unabhängig

Induktionsannahme $\implies m - 1 \leq n$, also $m \leq n + 1$ und

$B' := (w_1, \ldots, w_{m-1}, v_m, \ldots, v_n)$ ist Basis von V.

Um $m \leq n$ zu zeigen, müssen wir den Fall $m - 1 = n$ ausschliessen.

Dann wäre aber $(w_1, \ldots, w_m)$ linear unabhängig und $(w_1, \ldots, w_{m-1})$ Basis.

Dies würde $w_m \in \mathrm{Span}_{\mathbb{K}}(w_1, \ldots, w_{m-1})$ implizieren und damit wäre

$w_m = \alpha_1 w_1 + \ldots + \alpha_{m-1} w_{m-1}$, also $(w_1, \ldots, w_m)$ linear abhängig.

Insgesamt folgt so $m \leq n$.

Jetzt müssen wir noch w_m eintauschen.

B' Basis $\implies w_m = \beta_1 w_1 + \ldots + \beta_{m-1} w_{m-1} + \beta_m v_m + \ldots + \beta_n v_n$

$(w_1, \ldots, w_m)$ linear unabhängig $\implies \exists\ m \leq i \leq n$ mit $\beta_i \neq 0$

$\overset{\text{Lemma 12.2.1}}{\implies}$ w_m kann für v_i $(m \leq i \leq n)$ eingetauscht werden. □

Satz 12.2.3. *Je zwei endliche Basen eines $\mathbb{K}$–Vektorraums haben gleich Länge.*

Beweis. Sind $B = (v_1, \ldots, v_n)$, $B' = (w_1, \ldots, w_m)$ Basen von V

$\overset{\text{Satz 12.2.2}}{\implies}$ $n \leq m$ und $m \leq n$

$\Longrightarrow \;\; n = m.$ $\square$

Definition 12.2.2. *Sei V ein $\mathbb{K}$–Vektorraum. Ist V endlich erzeugt, so existiert nach Satz 11.2.3 eine Basis $B = (v_1, \ldots, v_n)$ von V. Wir definieren die Dimension von V über $\mathbb{K}$ als*

$$\dim_{\mathbb{K}} V := n.$$

Dieser Ausdruck ist wegen Satz 12.2.3 wohldefiniert.

Ist V nicht endlich erzeugt, so definieren wir

$$\dim_{\mathbb{K}} V := \infty$$

Bemerkung: Gibt es für alle $m \in \mathbb{N}$ linear unabhängige Vektoren $w_1, \ldots, w_m \in W$, so ist $\dim V = \infty$.

Beispiele:

1) $\dim_{\mathbb{R}} \mathbb{R}^n = n$

2) Der $\mathbb{R}$–Vektorraum $\mathbb{C}$ hat die Basis $(1, i) \Longrightarrow \dim_{\mathbb{R}} \mathbb{C} = 2$

3) $\dim_{\mathbb{R}} \mathbb{R}[x] = \infty$

Satz 12.2.4 (Basis–Ergänzungssatz)**.** *Sei V ein endlich erzeugter $\mathbb{K}$–Vektorraum und sei $(w_1, \ldots, w_m) \in V^m$ linear unabhängig.*

Dann existieren $v_{m+1}, \ldots, v_n \in V$, so dass

$$B := (w_1, \ldots, w_m, v_{m+1}, \ldots, v_m) \quad \textit{eine Basis von } V \quad \textit{ist.}$$

Beweis. Da V endlich erzeugt ist existiert ein Erzeugendensystem $(v_1, \ldots, v_N)$.

Aus Satz 12.2.1 folgt die existenz eines $n \leq N$, so dass $(v_1, \ldots, v_n) =: B$ eine Basis von V ist.

Die Behauptung folgt jetzt aus Satz 12.2.2. $\square$

Satz 12.2.5. *Sei V ein $\mathbb{K}$–Vektorraum und sei $U \subset V$ ein Untervektorraum. Ist V endlich erzeugt, so ist auch U endlich erzeugt und es gilt:*

$$\dim_{\mathbb{K}} U \leq \dim_{\mathbb{K}} V.$$

Aus $\dim_{\mathbb{K}} U = \dim_{\mathbb{K}} V$ folgt $U = V$.

Beweis. **Annahme**: $U \neq \{0_V\}$ ist nicht endlich erzeugt.

Behauptung: $\forall\, m \in \mathbb{N}$ existieren linear unabhängige Vektoren $w_1, \ldots, w_m \in U$.

Angenommen es wurden schon $(w_1, \ldots, w_{m-1})$ linear unabhängig gefunden. Da U nicht endlich erzeugt ist, muss ein Vektor $w_m \in U$ existieren, so dass

$$\mathrm{Span}_{\mathbb{K}}(w_1, \ldots, w_{m-1}) \overset{\subset}{\neq} \mathrm{Span}_{\mathbb{K}}(w_1, \ldots, w_{m-1}, w_m)$$

$\Longrightarrow (w_1, \ldots, w_m) \in V^m$ ist linear unabhängig

Aus Satz 12.2.4 folgt $\dim_{\mathbb{K}} V \geq m \quad \forall\, m \in \mathbb{N}$. Dies ist ein Widerspruch.

$\Longrightarrow$ U ist endlich erzeugt.

Sei nun $(w_1, \ldots, w_m) \in V^m$ eine Basis von U und $(v_1, \ldots, v_n)$ eine Basis von V. In dieser Situation implizert Satz 12.2.2 $m \leq n$.

Ist $m = n$, so ist $(w_1, \ldots, w_m)$ auch eine Basis von $V \implies U = V$. □

12.3 Summen und direkte Summen

Definition 12.3.1. *Sei V ein $\mathbb{K}$–Vektorraum und seien $W_1, \ldots, W_k \subset V$ Untervektorräume. Wir definieren die Summe der W_i, $1 \leq i \leq k$, durch*

$$W_1 + \ldots + W_k := \{v \in V :\ \exists\, w_i \in W_i, 1 \leq i \leq k \ \ \textit{mit}\ v = w_1 + \ldots + w_k\}$$

Bemerkung 12.3.2. *$W_1 + \ldots + W_k$ ist ein Untervektorraum von V und da die Basen der W_i zusammengenommen die Summe erzeugen, ist*

$$\dim_{\mathbb{K}}(W_1 + \ldots W_k) \leq \dim_{\mathbb{K}} W_1 + \ldots + \dim_{\mathbb{K}} W_k.$$

Satz 12.3.1. *Sei V ein $\mathbb{K}$–Vektorraum und seien $W_1, W_2 \subset V$ endlich dimensionale Untervektorräume. Dann gilt:*

$$\dim_{\mathbb{K}}(W_1 + W_2) = \dim_{\mathbb{K}} W_1 + \dim_{\mathbb{K}} W_2 - \dim_{\mathbb{K}}(W_1 \cap W_2)$$

Beweis. Sei $\beta_0 := (v_1, \ldots, v_m)$ eine Basis von $W_1 \cap W_2$.

$$\overset{\text{Satz 12.2.4}}{\Longrightarrow} \quad \begin{array}{l} \exists\, B_1 := (v_1, \ldots, v_m, w_1, \ldots, w_k) \ \text{ Basis von } W_1 \\ \exists\, B_2 := (v_1, \ldots, v_m, u_1, \ldots, u_\ell) \ \text{ Basis von } W_2. \end{array}$$

Behauptung: $B := (v_1, \ldots, v_m, w_1, \ldots, w_k, u_1, \ldots, u_\ell)$ Basis von $W_1 + W_2$

$\Longrightarrow \dim_{\mathbb{K}}(W_1 + W_2) = m + k + \ell = (m + k) + (m + \ell) - m = \dim_{\mathbb{K}} W_1 + \dim_{\mathbb{K}} W_2 - \dim_{\mathbb{K}}(W_1 \cap W_2)$

Beweis der Behauptung:

1. $\mathrm{Span}_{\mathbb{K}}(B) = W_1 + W_2$

   - Sei $v \in \mathrm{Span}_{\mathbb{K}}(B) \implies v = \lambda v_1 + \ldots + \lambda_m v_m + \mu_1 w_1 + \ldots + \mu_k w_k$
     $+ \rho_1 u_1 + \ldots + \rho_\ell u_\ell$
     $\in W_1 + W_2$

   - Sei $v \in W_1 + W_2 \implies v = (\lambda_1 v_1 + \ldots + \lambda_m v_m + \mu_1 w_1 + \ldots + \mu_k w_k)$
     $+ (\lambda_1' v_1 + \ldots + \lambda_m' v_m + \rho_1 u_1 + \ldots + \rho_\ell u_\ell)$
     $= (\lambda_1 + \lambda_1')v_1 + \ldots + (\lambda_m + \lambda_m')v_m + \mu_1 w_1$
     $+ \ldots + \mu_k w_k + \rho_1 u_1 + \ldots + \rho_\ell u_\ell$
     $\in \mathrm{Span}_{\mathbb{K}} B$

2. B ist linear unabhängig.

   Sei $\lambda_1 v_1 + \ldots + \lambda_m v_m + \mu_1 w_1 + \ldots + \mu_k w_k + \rho_1 u_1 + \ldots + \rho_\ell u_\ell = 0 \quad (*)$

   Definiere $v := \lambda_1 v_1 + \ldots + \lambda_m v_m + \mu_1 w_1 + \ldots + \mu_k w_k \in W_1$

   $\implies v = -\rho_1 u_1 - \ldots - \rho_\ell u_\ell \in W_2 \implies v \in W_1 \cap W_2$

   $\implies \exists\, \tilde{\lambda}_i, 1 \leq i \leq m$, mit $v = \tilde{\lambda}_1 v_1 + \ldots + \tilde{\lambda}_m v_m = \lambda_1 v_1 + \ldots + \lambda_m v_m + \mu_1 w_1 + \ldots + \mu_k w_k$

   B_1 Basis$\implies \lambda_i = \tilde{\lambda}_i \ \forall\ 1 \leq i \leq m$ und $\mu_j = 0 \ \ \forall\ 1 \leq j \leq k$

   $\overset{(*)}{\implies} \lambda_1 v_1 + \ldots + \lambda_m v_m + \rho_1 u_1 + \ldots \rho_\ell u_\ell = 0$

   B_2 Basis $\implies \lambda_i = 0 \ \forall\ 1 \leq i \leq m$ und $\rho_j = 0 \ \ \forall\ 1 \leq j \leq \ell$ (s. Satz 11.1.12)

□

Definition 12.3.3. *Sei V ein $\mathbb{K}$–Vektorraum und seien $W_1, W_2 \subset V$ Untervektorräume. V ist die direkte Summe von W_1 und W_2, $V = W_1 \oplus W_2$, falls*

$$V = W_1 + W_2 \quad \textit{und} \quad W_1 \cap W_2 = \{0\}.$$

Lemma 12.3.1. *Sei V ein $\mathbb{K}$–Vektorraum mit $\dim_{\mathbb{K}} V < \infty$ und seien $W_1, W_2 \subset V$ Untervektorräume. Dann sind äquivalent:*

1) $V = W_1 \oplus W_2$

2) Sei $B_1 = (w_1, \ldots, w_k)$ Basis von W_1 und $B_2 = (u_1, \ldots, u_\ell)$ Basis von W_2, dann ist $B = (w_1, \ldots, w_k, u_1, \ldots, u_\ell)$ Basis von V.

3) $V = W_1 + W_2$ *und* $\dim_{\mathbb{K}} V = \dim_{\mathbb{K}} W_1 + \dim_{\mathbb{K}} W_2$

Beweis. 1) $\implies$ 2): Folgt aus Beweis von vorherigem Satz, da $\emptyset$ eine Basis von $\{0\} = W_1 \cap W_2$ ist.

2) $\Longrightarrow$ 3): B Basis von $V \implies V = W_1 + W_2$ und $\dim_{\mathbb{K}} V = k + \ell = \dim_{\mathbb{K}} W_1 + \dim_{\mathbb{K}} W_2$

3) $\Longrightarrow$ 1): Satz 11.4.3 $\implies \dim_{\mathbb{K}}(W_1 \cap W_2) = 0 \Longrightarrow W_1 \cap W_2 = \{0\}$. □

Lemma 12.3.2. *Sei V ein $\mathbb{K}$–Vektorraum mit $\dim_{\mathbb{K}} V < \infty$ und seien W_1, W_2 Untervektorräume von V. Dann sind äquivalent:*

1) $V = W_1 \oplus W_2$

2) Jedes $v \in V$ ist eindeutig darstellbar als $v = w_1 + w_2$ mit $w_1 \in W_1, w_2 \in W_2$

3) $V = W_1 + W_2$ und aus $w_1 + w_2 = 0$ mit $w_1 \in W_1, w_2 \in W_2$ folgt $w_1 = w_2 = 0$.

Beweis. 1) $\Longrightarrow$ 2): Ist $v = w_1 + w_2 = w_1' + w_2' \implies w_1 - w_1' = w_2' - w_2 \in W_1 \cap W_2 = \{0\}$
$\implies w_1 = w_1'$ und $w_2 = w_2'$

2) $\Longrightarrow$ 3): $w_1 + w_2 = 0 = 0 + 0 \implies w_1 = w_2 = 0$

3) $\Longrightarrow$ 1): **Annahme**: $\exists\, v \in W_1 \cap W_2, v \neq 0$
$\implies v + (-v) = 0$, Widerspruch zu 3). □

13 Lineare Gleichungssysteme

13.1 LGS

Gegeben sei ein System aus n Gleichungen mit m Unbekannten, d. h. für $x = (x_1, \dots, x_m) \in \mathbb{R}^m$ gelte

$$\begin{array}{ccccccccc} a_{11} & x_1 & + & \dots & + & a_{1m} & x_m & = & b_1 \\ \vdots & & & & & \vdots & & & \vdots \\ a_{n1} & x_1 & + & \dots & + & a_{nm} & x_m & = & b_n \end{array} \qquad (*)$$

Dabei sind die *Koeffizienten* a_{ij} und b_i gegebene reelle Zahlen. Der Index $i \in \{1, \dots, n\}$ bezeichnet die Nummer der Gleichung, $j \in \{1, \dots, m\}$ ist die Nummer der zugehörigen Unbekannten x_j.

Die Lösungsmenge ist definiert durch

$$L = \{x \in \mathbb{R}^m : \ x \ \text{ löst } (*)\}.$$

Zur Vereinfachung definieren wir die Matrix

$$A := \begin{pmatrix} a_{11} & \dots & a_{1m} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nm} \end{pmatrix} \quad \text{und eine Spalte } b := \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

A hat n Zeilen und m Spalten.

Definition 13.1.1. *Für ein festes $1 \leq i \leq n$ heissen die (a_{ij}), $1 \leq j \leq m$ Zeilenvektoren von A und für festes $1 \leq j \leq m$ sind (a_{ij}), $1 \leq i \leq n$ die Spaltenvektoren von A.*

Weiter wird

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \quad \text{als Spalte geschrieben.}$$

Wir definieren:

$$A \cdot x := \begin{pmatrix} a_{11} & x_1 & + & \dots & + & a_{1m} & x_m \\ & & & \vdots & & & \\ a_{n1} & x_1 & + & \dots & + & a_{nm} & x_m \end{pmatrix}, \quad \text{d.,h.}$$

$L =:$ LÖS $(A, b) = \{x \in \mathbb{R}^m : A \cdot x = b\} \subset \mathbb{R}^m$. Zuletzt definieren wir die Koeffizientenmatrix

$$\begin{pmatrix} a_{11} & \dots & a_{1m} & | & b_1 \\ \vdots & & \vdots & | & \vdots \\ a_{n1} & \dots & a_{nm} & | & b_n \end{pmatrix}$$

(A, B) ist eine $n \times (m+1)$–Matrix

Bemerkung: Für eine gegebene Matrix $A \in \mathbb{R}^{n,m}$ erhält man durch die folgende Zuordnung eine lineare Abbildung $\mathbb{R}^m \to \mathbb{R}^n$:

$$x \to A \cdot x.$$

Umgekehrt lässt sich jede lineare Abbildung $B : \mathbb{R}^m \to \mathbb{R}^n$ durch genau eine $n \times m$-Matrix $A = (a_{ij})$ darstellen: $B(x) = A \cdot x \ \forall x \in \mathbb{R}^m$. Dabei sind die Spaltenvektoren $a_j = (a_{ij})$, $1 \leq i \leq n$ durch $a_j = B(e_j)$ (e_j ist der j-te Basisvektor) gegeben. Aufgrund der Linearität von B erhält man damit eine wohldefinierte Abbildung auf ganz $\mathbb{R}^m$.

Beispiel: Für die lineare Abbildung

$$B : \ \mathbb{R}^2 \to \mathbb{R}^2 \quad B(x) = \begin{pmatrix} x_1 + x_2 \\ x_2 \end{pmatrix}$$

gilt

$$B(e_1) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \ B(e_2) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

und damit gilt

$$B(x) = A \cdot x \quad \text{mit} \quad A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$$

Man sagt, (A, b) hat *Zeilenstufenform*, falls sie von der Form

$$\left(\begin{array}{ccccc|c} \underline{a_{1j_1}} & & & & & b_1 \\ & \underline{a_{2j_2}} & & & & \vdots \\ 0 & & \ddots & & & \vdots \\ & & & \underline{a_{rj_r}} & \dots & b_r \\ & & & & & \vdots \\ & & & & & b_n \end{array}\right)$$

Die $a_{1j_1}, \dots, a_{rj_r}$ heissen *Angelpunkte*. Sie müssen $\neq 0$ sein und es gilt $1 \leq j_1 < j_2 < \dots < j_r \leq m$.

Die Einträge a_{ij} unter der Stufenlinie müssen $= 0$ sein.

Das Gaussche Eliminationsverfahren liefert, dass jede Matrix in Zeilenstufenform gebracht werden kann, ohne die Lösungsmenge des dazugehörigen linearen Gleichungssystems zu ändern.

Definition 13.1.2. *Für eine Koeffizientenmatrix (A, B) eines linearen Gleichungssystems ist eine Zeilenumformung entweder*

1) *man vertauscht die Zeilen i und k $(i \neq k)$ oder*

2) *man zählt zur Zeile k das λ–fache der Zeile i dazu, wobei $i \neq k$ und $\lambda \in \mathbb{R}$.*

Satz 13.1.1. 1) *Ist $(\tilde{A}, \tilde{b})$ aus (A, B) durch Zeilenumformungen entstanden, so gilt:*

$$LÖS\ (\tilde{A}, \tilde{b}) = LÖS\ (A, b).$$

2) *Jede Koeffizientenmatrix kann durch endlich viele Zeilenumformungen auf Zeilenstufenform gebracht werden.*

Beweis. 1) Eine Zeilenumformung vom Typ 1) ändert nur die Reihenfolge der Bedingungen
$\Longrightarrow$ Lösungsmenge bleibt gleich

Für eine Zeilenumformung vom Typ 2) definieren wir $\forall\ 1 \leq i \leq n$

$$s_i := (a_{i1}, \ldots, a_{in}) \in \mathbb{R}^m\,.$$

Ist $x = (x_1, \ldots, x_m) \in \mathbb{R}^m$, so ist die i–te Gleichung des Systems $\langle s_i, x\rangle = \sum\limits_{j=1}^{m} a_{ij}x_j = b_i$.

Zu zeigen bleibt das für alle $x \in \mathbb{R}^m$ und $\lambda \in \mathbb{R}$ gilt:

$$\begin{cases} \langle s_i, x\rangle = b_i \quad \text{und} \\ \langle s_k, x\rangle = b_k \end{cases} \quad \Longleftrightarrow \quad \begin{cases} \langle s_i, x\rangle = b_i \quad \text{und} \\ \langle s_k + \lambda s_i, x\rangle = b_k + \lambda b_i \end{cases}$$

„ $\Longrightarrow$ “: $\langle s_k + \lambda s_i, x\rangle = \langle s_k, x\rangle + \lambda\langle s_i, x\rangle = b_k + \lambda b_i$

„ $\Longleftarrow$ “: $\langle s_k, x\rangle = \langle s_k + \lambda s_i, x\rangle - \lambda\langle s_i, x\rangle = b_k + \lambda b_i - \lambda b_i = b_k$.

2) i) $a_{ij} = 0 \quad \forall\ i, j \implies$ Zeilenstufenform mit $r = 0$.

ii) Sei j_1 die erste (von links gezählte) Spalte mit einem Eintrag $\neq 0$. Ist $a_{1j_1} \neq 0$ so ist dies der erste Angelpunkt.

Ist $a_{1j_1} = 0$ so nehme man eine Zeile mit $a_{ij_1} \neq 0$ und vertausche sie mit der ersten.

Definiere den neuen Angelpunkt $\tilde{a}_{1j_1} = a_{ij_1}$.

Damit mache man durch Umformungen vom Typ 2) alle darunter stehenden Einträge $a_{2j_1}, \ldots, a_{mj_1}$ zu Null.

Ist etwa $a_{2j_1} \neq 0$, so soll $a_{2j_1} + \lambda \tilde{a}_{1j_1} = 0$ sein, also muss $\lambda = -\dfrac{a_{2j_1}}{\tilde{a}_{1j_1}}$ für die Umformung der 2. Zeile gewählt werden.

Im nächsten Schritt betrachtet man die nach den Umformungen entstandene Teilmatrix mit den Zeilen $i \geq 2$ und den Spalten $j \geq j_1 + 1$. Mit dieser Teilmatrix verfährt man wie zuvor:

Man sucht die erste Spalte $j_2 > j_1$ mit einem Eintrag $\neq 0 \ldots$

Hat man schliesslich einen Angelpunkt $a_{rj_r} \neq 0$ gefunden, so dass nach allen Zeilenumformungen in den Zeilen $i \geq r + 1$ und den Spalten $j \geq j_r$ nur noch Nullen stehen, so ist die Zeilenstufenform erreicht.

$\square$

Beispiel:

$$(A, B) = \left(\begin{array}{cccccc|c} 0 & 0 & 3 & 0 & 1 & 2 & 3 \\ 0 & 1 & 2 & 4 & 0 & 1 & 2 \\ 0 & -1 & 4 & -4 & 4 & 3 & 8 \\ 0 & 2 & 7 & 8 & 2 & 4 & 9 \end{array}\right) j_1 = 2$$

Vertausche Zeile 1) und 2)

$$\left(\begin{array}{cccccc|c} 0 & 1 & 2 & 4 & 0 & 1 & 2 \\ 0 & 0 & 3 & 0 & 1 & 2 & 3 \\ 0 & -1 & 4 & -4 & 4 & 3 & 8 \\ 0 & 2 & 7 & 8 & 2 & 4 & 9 \end{array}\right)$$

Addiere 1. Zeile zur 3. Zeile und (-2)·1. Zeile zur 4. Zeile

$$\left(\begin{array}{cccccc|c} 0 & 1 & 2 & 4 & 0 & 1 & 2 \\ 0 & 0 & 3 & 0 & 1 & 2 & 3 \\ 0 & 0 & 6 & 0 & 4 & 4 & 10 \\ 0 & 0 & 3 & 0 & 2 & 2 & 5 \end{array}\right) j_2 = 3$$

Addiere (-2) 2. Zeile zur dritten und (-1) 2. Zeile zur vierten

$$\left(\begin{array}{cccccc|c} 0 & 1 & 2 & 4 & 0 & 1 & 2 \\ 0 & 0 & 3 & 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 & 2 & 0 & 4 \\ 0 & 0 & 0 & 0 & 1 & 0 & 2 \end{array}\right) j_3 = 5$$

Addiere $(-\frac{1}{2})$ 3. Zeile zur 4.

$$\left(\begin{array}{cccccc|c} 0 & 1 & 2 & 4 & 0 & 1 & 2 \\ 0 & 0 & 3 & 0 & 1 & 2 & 3 \\ 0 & 0 & 0 & 0 & 2 & 0 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right) =: (\tilde{A}, \tilde{b})$$

$\Longrightarrow r = 3.$

Lemma 13.1.1. *Ist (A, b) in Zeilenstufenform und existiert ein $r+1 \leq i \leq n$ mit $b_i \neq 0$, so ist LÖS $(A, b) = \emptyset$.*

Beweis. Die i–te Gleichung des Systems lautet:

$$0 = 0_{x_1} + \ldots + 0_{x_m} = b_i \neq 0$$

□

Im Folgenden nehmen wir ohne Einschränkung $j_1 = 1, j_2 = 2, \ldots, j_r = r$ (sonst vertauschen wir Spalten und benennen die x_i um) und $b_{r+1} = \ldots = b_n = 0$ an, d. h.

$$\left(\begin{array}{ccccc|c} \underline{a_{11}} & & & & & b_1 \\ & \underline{a_{2j_2}} & & & & \vdots \\ 0 & & \ddots & & & \vdots \\ & & & \underline{a_{rr}} & \cdots & b_r \\ & & & & & 0 \\ & & & & & \vdots \\ & & & & & 0 \end{array}\right)$$

Setze $x_{r+1} = \lambda_1, \ldots, x_m = \lambda_k$, wobei $k := m - r \geq 0$.

Die r-te Gleichung lautet dann

$$a_{rr}x_r + a_{r,r+1}x_{r+1} + \ldots + a_{rm}x_m = b_r$$
$$\overset{a_{rr}\neq 0}{\Longrightarrow} x_r = \frac{1}{a_{rr}}(b_r - a_{r,r+1}\lambda_1 - \ldots - a_{rm}\lambda_k)$$

Im nächsten Schritt setzen wir $x_r, x_{r+1}, \ldots, x_m$ in die $(r-1)$–te Gleichung ein, usw. Insgesamt folgt damit

Satz 13.1.2. *Gegeben sei ein lineares Gleichungssystem $Ax = b$ in Zeilenstufenform mit $j_1 = 1, \ldots, j_r = r$. Dann gilt:*

$$\textit{LÖS}\ (A, b) \neq \emptyset \iff b_{r+1} = \ldots = b_n = 0.$$

Beispiel: Sei

$$(A|b) := \begin{pmatrix} 2 & 1 & 3 & | & 4 \\ 0 & 3 & 1 & | & 6 \end{pmatrix}$$

Setze $x_3 = \lambda_1$. Aus $3x_2+x_3 = 6$ folgt $x_2 = 2-\frac{\lambda_1}{3}$ und damit folgt aus $2x_1+x_2+3x_3 = 4$: $x_1 = 2 - 1 - \frac{3}{2}\lambda_1 + \frac{\lambda_1}{6} = 1 - \frac{4}{3}\lambda_1$. Also ist

$$x = (1 - \frac{4}{3}\lambda_1, 2 - \frac{\lambda_1}{3}, \lambda_1)$$

für jedes $\lambda_1 \in \mathbb{R}$ eine Lösung des LGS.

Definition 13.1.3. *Ist $b = 0_{\mathbb{R}^n}$, dann heisst das Gleichungssystem homogen. Ansonsten heisst es inhomogen.*

Lemma 13.1.2. *Es gelten:*

1) LÖS $(A,0)$ ist ein Untervektorraum von $\mathbb{R}^m$.

2) Ist $v \in \mathbb{R}^m$ eine Lösung von $Ax = b$, dann gilt:

$$LÖS\ (A,b) = v + \quad LÖS\ (A,0).$$

Beweis: 1) i) $0 \in$ LÖS $(A,0)$

ii) $u, v \in$ LÖS $(A,0) \implies Au = 0 = Av$
$\implies A(u+v) = A\,u + A\,v = 0$, also $u + v \in$ LÖS$(A,0)$

iii) $\lambda \in \mathbb{R}, u \in$ LÖS $(A,0) \implies A(\lambda u) = \lambda(Au) = 0$, also $\lambda u \in$ LÖS$(A,0)$

2) „ $\subset$ “: Sei $u \in$ LÖS $(A,b) \implies Au = b = Av$
$\implies A(u-v) = Au - A\,v = 0$
$\implies u - v \in$ LÖS $(A,0)$
$\implies u \in v +$ LÖS $(A,0)$

„ $\supset$ “: Sei $u \in v +$ LÖS $(A,0)$

$\implies u - v \in$ LÖS $(A,0) \implies A(u-v) = 0$
$\implies Au = A(u-v) + A\,v = b + 0 = b$
$\implies u \in$ LÖS (A,b).

□

Bemerkung: Ist A in Zeilenstufenform mit $j_1 = 1, ..., j_r = r$ und $k = m - r$, dann ist $\dim_{\mathbb{K}} \text{LÖS}(A,0) \geq k$, denn es existieren k linear unabhängige Vektoren

$$\begin{aligned} x^1 =& (x_1^1, \ldots, x_r^1, 1, 0, \ldots, 0) \\ &\vdots \\ x^k =& (x_1^k, \ldots, x_r^k, 0, \ldots, 0, 1), \end{aligned}$$

welche die Gleichung $Ax = 0$ lösen.

Die Vektoren sind linear unabhängig, denn aus $\lambda_1 x^1 + \ldots + \lambda_k x^k = 0$ folgt

$$(\sum_{l=1}^{k} \lambda_l x_1^l, \ldots, \sum_{l=1}^{k} \lambda_l x_r^l, \lambda_1, \ldots, \lambda_k) = 0$$

und damit $\lambda_1 = \ldots = \lambda_k = 0$.

Des Weiteren können wir wie vor dem Satz 13.1.2 argumentieren um zu sehen, dass die ersten r-Komponenten der Vektoren $x^1, \ldots, x^k$ tatsächlich so gewählt werden können, dass wir Lösungen des LGS erhalten.

Lemma 13.1.3. *Gegeben sei ein lineares Gleichungssystem $Ax = b$ in Zeilenstufenform.*

1) *Ist $r = m = n$ und $b \in \mathbb{R}^n$ beliebig, so besitzt das System genau eine Lösung. (Ist $b = 0$, so gilt $x = 0$).*

2) *Ist $m > n$ und $b = 0$, so existiert eine Lösung $x \neq 0$ des Systems.*

Beweis. 1) In diesem Fall ist

$$(A, b) = \left(\begin{array}{cccc|c} a_{11} & & & & b_1 \\ & a_{22} & & & \vdots \\ 0 & & \ddots & & \vdots \\ & & & a_{nn} & b_n \end{array}\right) \text{ mit } a_{ii} \neq 0 \quad \forall\, 1 \leq i \leq m$$

Aus $a_{nn} x_n = b_n$ folgt $x_n = \dfrac{b_n}{a_{nn}}$ aus $a_{n-1,n-1} x_{n-1} + a_{n-1,n} x_n = b_{n-1}$ folgt $x_{n-1} = \dfrac{1}{a_{n-1,n-1}} \Big(b_{n-1} - \dfrac{a_{n-1,n}}{a_{nn}} b_n \Big)$ und durch weiteres Einsetzen erhält man eine eindeutige Lösung $x \in \mathbb{R}^n$.

2) folgt aus der vorherigen Bemerkung, da $r \leq n$ und damit $k = m - r > 0$.

$\square$

Bezeichnen wir die Spaltenvektoren der Matrix

$$A = \begin{pmatrix} a_{11} & \dots & a_{1m} \\ a_{21} & \dots & a_{2m} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nm} \end{pmatrix} \quad \text{mit} \quad a_j := \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{pmatrix},$$

so kann das Gleichungssystem auch in der übersichtlichen Form

$$\sum_{j=1}^{m} x_j \, a_j \;=\; b$$

geschrieben werden. Daraus ergibt sich:

Satz 13.1.3. *a) Das System $A\,x = b$ ist genau dann lösbar, wenn*

$b \in Span\{a_1, \dots, a_m\}$ ist.

b) Angenommen, das System $A\,x = b$ besitze eine Lösung $x \in \mathbb{R}^m$ zu einem gegebenen $b \in \mathbb{R}^n$. Diese Lösung ist genau dann eindeutig, wenn die Spaltenvektoren $a_1, \dots, a_m$ von A linear unabhängig sind.

c) Das System $A\,x = b$ ist zu einem $b \in \mathbb{R}^n$ genau dann eindeutig lösbar, wenn $a_1, \dots, a_m \in \mathbb{R}^n$ linear unabhängig und $b \in Span\{a_1, \dots, a_m\}$ sind.

d) Das System ist genau dann für alle $b \in \mathbb{R}^n$ eindeutig lösbar, wenn $n = m$ gilt und $\{a_1, \dots, a_n\}$ eine Basis in $\mathbb{R}^n$ bildet.

Beweis: (a) ist offensichtlich auf Grund der obigen Schreibweise des Systems.

(b) Wenn das inhomogene Gleichungssystem $\sum_{j=1}^{m} x_j \, a_j \;=\; b$ zu einem $b \in \mathbb{R}^n$ genau eine Lösung besitzt, dann folgt aus Lemma 13.1.2: Das zugehörige homogene System $\sum_{j=1}^{m} x_j \, a_j \;=\; 0$ hat nur die triviale Lösung, d.h. $\text{LÖS}(A, 0) = \{0\}$. Insbesondere ist die Gleichung

$$\lambda_1 a_1 + \ldots + \lambda_m a_m = 0$$

nur für $\lambda_1 = \ldots = \lambda_m = 0$ erfüllt. Also sind $a_1, \dots, a_m$ linear unabhängig in $\mathbb{R}^n$.

Die andere Richtung folgt aus Satz 12.1.2.

(c) folgt unmittelbar aus den Teilaussagen (a) und (b).

(d) Wenn das System $Ax = b$ für jedes $b \in \mathbb{R}^n$ eindeutig lösbar ist, dann kann es insbesondere nur eine Darstellung der 0 als Linearkombination

$$\sum_{j=1}^{m} x_j \, a_j \;=\; A\,x \;=\; 0$$

geben, nämlich die triviale Lösung $x = 0 \in \mathbb{R}^m$, d.h. $a_1, \ldots, a_m$ sind linear unabhängig in $\mathbb{R}^n$ Außerdem gilt dabei $\mathbb{R}^n = \text{Span}\{a_1, \ldots, a_m\}$, d.h. nach Definition, $\{a_1, \ldots, a_m\}$ ist eine Basis von des $\mathbb{R}$-Vektorraumes $\mathbb{R}^n$.
Schließlich garantiert Satz 12.2.3, dass die Anzahl m der Elemente von $\{a_1, \ldots, a_m\}$ gleich $\dim \mathbb{R}^n = n$ ist.
Der umgekehrte Schluss folgt aus a) und b). $\square$

Korollar 13.1.1. *Das homogene Gleichungssystem* $A\,x = 0$ *mit* $A \in \mathbb{R}^{n,n}$ *hat genau dann nichttriviale Lösungen in* $\mathbb{R}^n$*, wenn die Spaltenvektoren von* A *linear abhängig sind.* $\square$

13.2 Dimension der Teilräume bei linearen Abbildungen

Definition 13.2.1. *V und W seien Vektorräume über $\mathbb{R}$ mit $m := \dim V < \infty$, und $A : V \longrightarrow W$ bezeichne eine lineare Abbildung. Wir definieren*

$$\begin{aligned} \text{Kern}\,(A) &= \underline{\textit{Nullraum}} \text{ von } A := \{v \in V \mid A\,v = 0\} \subset V, \\ \text{Bild}\,(A) &= \underline{\textit{Wertebereich}} \text{ von } A := \{A\,v \mid v \in V\} \subset W. \end{aligned}$$

Falls zusätzlich W endlichdimensional ist, so lässt sich A durch eine Matrix bzgl. Basen von V und W darstellen, und es heißen

$$\begin{aligned} \underline{\textit{Zeilenrang}} \text{ von } A &:= \dim \textit{Span}\{\text{ Zeilenvektoren von } A\}, \\ \underline{\textit{Spaltenrang}} \text{ von } A &:= \dim \textit{Span}\{\text{ Spaltenvektoren von } A\}. \end{aligned}$$

Bemerkung Zeilen- und Spaltenrang hängen nicht von der Wahl der verwendeten Basen von V bzw. W ab. Ferner folgt direkt aus der Definition

$$\text{Spaltenrang von } A = \dim \text{ Bild}\,(A).$$

Satz 13.2.1. *Mit den Bezeichnungen aus Definition 13.2.1 sind* Kern (A) *und* Bild (A) $\mathbb{R}$-*Untervektorräume von V bzw. W, und es gilt*

$$\dim \text{ Kern}\,(A) + \dim \text{ Bild}\,(A) = \dim V.$$

Beweis: Es lässt sich recht leicht nachprüfen, dass Kern (A) und Bild (A) abgeschlossen bzgl. Vektoraddition und skalarer Multiplikation sind.

Mit Blick auf die behauptete Dimensionsformel sei $\{v_1, \ldots, v_k\}$ eine Basis von Kern (A), $k := \dim$ Kern (A). Dann gibt es $v_{k+1}, \ldots, v_m$, so dass $\{v_1, \ldots, v_m\}$ eine Basis von V bildet (siehe Satz 12.2.4). Jedes Element von Bild $(A) \subset W$ hat die Gestalt

$$A\,x = A\left(\sum_{j=1}^{m} c_j\, v_j\right) = \sum_{j=k+1}^{m} c_j \cdot A\,v_j,$$

d.h. es gilt die Ungleichung $\dim \text{ Bild}\,(A) \leq m - k$.

Andererseits impliziert

$$\sum_{j=k+1}^{m} \lambda_j \cdot A\, v_j \;=\; 0$$

immer, dass $\lambda_{k+1}\, v_{k+1} + \ldots + \lambda_m\, v_m$ in $\mathrm{Kern}\,(A) = \mathrm{Span}\{v_1,\, \ldots\, , v_k\}$ liegt. Dies ist nur mit $0 = \lambda_{k+1} = \,\ldots\, \lambda_m$ möglich, weil $\{v_1,\, \ldots\, , v_m\}$ als Basis linear unabhängig sind. Daraus folgt $\dim\ \mathrm{Bild}\,(A) \geq m - k$. $\square$

Korollar 13.2.1. *V und W seien $\mathbb{R}$-Vektorräume mit gleicher Dimension $< \infty$. Eine lineare Abbildung $A : V \longrightarrow W$ besitzt genau dann eine Umkehrfunktion $A^{-1} : W \longrightarrow V$, wenn* $\mathrm{Kern}\,(A) = \{0\}$ *ist.*
Gleichbedeutend damit gilt: A bijektiv $\iff$ A injektiv. $\square$

Satz 13.2.2. *V_0 sei ein k–dimensionaler Untervektorraum von $\mathbb{R}^m$, und $V_0^{\perp}$ bezeichne das (bezüglich des üblichen Skalarprodukts) orthogonale Komplement von V_0:*

$$V_0^{\perp} \;:=\; \Big\{ x \in \mathbb{R}^m \mid \langle x,\, v\rangle = 0 \;\; \forall\, v \in V_0 \Big\}.$$

Dann erfüllt die Dimension stets $\dim V_0^{\perp} = m - k$.

Beweis: Zunächst sind die Untervektorräume V_0 und $V_0^{\perp}$ von $\mathbb{R}^m$ "direkt", d.h. $V_0 \cap V_0^{\perp} = \{0\}$, denn für jedes $x \in V_0 \cap V_0^{\perp}$ gilt $\|x\|^2 = \langle x, x\rangle = 0$. Also ist nach Lemma 12.3.1

$$\dim V_0 + \dim V_0^{\perp} \;=\; \dim(V_0 + V_0^{\perp}) \;\leq\; m$$

und somit $\dim V_0^{\perp} \leq m - k$.

Außerdem löst jedes Elemente x von $V_0^{\perp}$ das homogene Gleichungssystem

$$\langle x, v_j\rangle = 0 \qquad \text{für } j = 1,\, \ldots\, , k,$$

wobei $\{v_1,\, \ldots\, , v_k\}$ eine Basis von V_0 ist. Es besteht aus k Gleichungen für m Unbekannte. Die Bemerkung nach Lemma 13.1.2 besagt nun das der Lösungsraum mindestens $(m - k)$–dimensional, d.h. es gilt $\dim V_0^{\perp} \geq m - k$. $\square$

Satz 13.2.3 (Rang einer Matrix)**.** *Für jede Matrix $A \in \mathbb{R}^{n,m}$ gilt*

$$\textit{Zeilenrang von } A \;=\; \textit{Spaltenrang von } A \;=:\; \mathrm{rang}(A).$$

Beweis: Wie in obiger Bemerkung erwähnt, ist der Spaltenrang gleich der Dimension des Bildraums der durch A erzeugten linearen Abbildung $\mathbb{R}^m \longrightarrow \mathbb{R}^n$.
Der Kern ist gleich dem orthogonalen Komplement des durch die Zeilenvektoren $a_1,\, \ldots\, , a_n$ aufgespannten Untervektorraums von $\mathbb{R}^m$:

$$\begin{aligned} u \in \mathrm{Kern}\,(A) \subset \mathbb{R}^m \;&\iff\; A\, u \;=\; 0 \\ &\iff\; \text{In der } j\text{-ten Zeile: } \langle a_j,\, u\rangle \;=\; 0 \quad (j = 1,\, \ldots\, , n), \end{aligned}$$

und Satz 13.2.2 ergibt

$$\dim\ \mathrm{Kern}\,(A) \;=\; m - \text{Zeilenrang } (A)$$

Die Dimensionsformal aus Satz 13.2.1 besagt schließlich

$$\begin{aligned} m \;&=\; \dim\ \mathrm{Kern}\,(A) \qquad\quad + \; \dim\ \mathrm{Bild}\,(A) \\ &=\; m - \text{Zeilenrang } (A) \;+\; \text{Spaltenrang } (A). \end{aligned}$$ $\square$

14 Determinante einer Matrix und Eigenwerte

14.1 Die Determinante

Definition 14.1.1. $\mathbb{K}^{n,n}$ *(bzw.* $\mathbb{R}^{n,n}$ *oder* $\mathbb{C}^{n,n}$*) bezeichne die Menge der* $n \times n$*–Matrizen mit Einträgen aus* $\mathbb{K}$ *(bzw.* $\mathbb{R}$ *oder* $\mathbb{C}$*).*
Eine Funktion $\det : \mathbb{K}^{n,n} \longrightarrow \mathbb{K}$ *heißt eine* <u>Determinantenfunktion</u> *(oder kurz* <u>Determinante</u>*), wenn sie folgende Eigenschaften besitzt:*

1.) $\det(\cdot)$ *ist "multilinear":* $\det(\cdot)$ *ist linear in jeder Spalte, d.h. wenn* a^j *bzw.* b^j *jeweils für eine entsprechende Spalte steht,*

$$\begin{aligned} &\det\left(a^1, \dots, a^{j-1}, \lambda a^j + \mu\, b^j, a^{j+1}, \dots, a^n\right) \\ &= \lambda \cdot \det(a^1, \dots, a^n) + \mu \cdot \det\left(a^1, \dots, a^{j-1}, b^j, a^{j+1}, \dots, a^n\right) \\ &\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad (j = 1, \dots, n) \end{aligned}$$

2.) $\det(\cdot)$ *ist "alternierend": Bei Vertauschung zweier Spalten einer Matrix geht deren Determinante in ihr Negatives über.*

3.) $\det(\cdot)$ *ist "normiert":* $\det\begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} = 1$

Es wird weiter unten zu zeigen sein, dass eine solche Funktion existiert und dass sie sogar eindeutig in $\mathbb{K}^{n,n}$ ist. Zunächst wollen wir aus der Definition einige wichtige Eigenschaften ableiten:

Satz 14.1.1. *Eine Determinantenfunktion hat die folgenden Eigenschaften:*

(a) *Es gilt* $\det A = 0$*, falls*
  (i) *eine Spalte gleich 0 ist oder*
  (ii) *zwei Spalten gleich sind.*

(b) $\det A = 0$ *gilt genau dann, wenn die Spalten linear abhängig sind. Äquivalent dazu ist:* $\det A \neq 0 \Longleftrightarrow$ *die Spalten von* A *sind linear unabhängig.*

(c) *Addition eines Vielfachen einer Spalte zu einer anderen Spalte ändert die Determinante nicht, d.h. für* $\lambda \in \mathbb{K}$ *und* $k \neq j$ *ist*

$$\det\Big(a^1,\,\ldots\,,a^{j-1},\;a^j+\lambda a^k,\;a^{j+1},\,\ldots\,,a^n\Big)\;=\;\det\left(a^1,\,\ldots\,,a^n\right).$$

Beweis: (a) (i) folgt aus der Linearität bezüglich der jeweiligen Spalte: Ist z. B. die j–te Spalte gleich 0, so gilt

$$\begin{aligned}\det\Big(a^1,\,\ldots\,,a^n\Big) &= \det\Big(a^1,\,\ldots\,,a^{j-1},\;0\cdot a^j,\;a^{j+1},\,\ldots\,,a^n\Big)\\ &= 0\cdot\det\Big(a^1,\,\ldots\,,a^n\Big)\;=\;0.\end{aligned}$$

(ii) Vertauschung der beiden gleichen Spalten ändert einerseits nichts, führt aber andererseits zum negativen Wert.

(b) “$\Longleftarrow$” Gilt z. B. $a^k=\sum\limits_{j\neq k}\alpha_j a^j$, so folgt aus der Linearität bezüglich der k–ten Spalte

$$\det A\;=\;\sum_{j\neq k}\alpha_j\cdot\det\Big(a^1,\,\ldots\,,a^{k-1},a^j,a^{k+1},\,\ldots\,,a_n\Big)\;=\;0,$$

da in den Determinanten der Summe jeweils die Spalten k,j gleich sind.

(b) “$\Longrightarrow$” $a^1,\,\ldots\,,a^n\in\mathbb{K}^n$ seien linear unabhängig, also eine Basis von $\mathbb{K}^n$. Dann lässt sich insbesondere jeder Vektor e^k der kanonischen Basis (d.h. $e^k=(0,\,\ldots\,,0,1,0,\,\ldots\,,0)$ mit 1 an der k–ten Stelle) als Linearkombination der a^j darstellen:

$$e^k=\sum_{j=1}^n x_{kj}\,a^j\qquad\text{mit}\qquad x_{kj}\in\mathbb{K}\,.$$

Damit folgt

$$\begin{aligned}1\;&=\;\det(e^1,\,\ldots\,,e^n)\;=\;\det\left(\sum_{j_1=1}^n x_{1j_1}\,a^{j_1},\;\sum_{j_2=1}^n x_{2j_2}\,a^{j_2},\;\ldots\;,\;\sum_{j_n=1}^n x_{nj_n}\,a^{j_n}\right)\\ &=\;\sum_{j_1=1}^n\cdots\sum_{j_n=1}^n x_{1j_1}\,x_{2j_2}\,\ldots\,x_{nj_n}\,\cdot\det(a^{j_1},\,\ldots\,,a^{j_n})\\ &\qquad\text{(jede dieser Determinanten ist }\;\pm\det(a^1,\,\ldots\,,a^n),\;\;\text{wenn sie aus}\\ &\qquad\det(a^1,\,\ldots\,,a^n)\;\text{ durch endlich viele Vertauschungen hervorgeht,}\\ &\qquad\text{oder }\;=0,\text{wenn mindestens 2 Spalten gleich sind,}\\ &=\;\alpha\cdot\det(a^1,\,\ldots\,,a^n)\qquad\text{mit einem }\alpha\in\mathbb{K}\,.\end{aligned}$$

Also ist $\det(a^1,\,\ldots\,,a^n)\neq 0$.

(c) Die Linearität in jeder Spalte und Teilaussage (a) (ii) ergeben für alle $k\neq j$ und $\lambda\in\mathbb{K}$

$$\begin{aligned}&\det\Big(a^1,\,\ldots\,,a^{j-1},\;a^j+\lambda\,a^k,\;a^{j+1},\,\ldots\,,a^n\Big)\\ &\quad=\;\det\left(a^1,\,\ldots\,,a^n\right)\;+\;\lambda\cdot\det\Big(a^1,\,\ldots\,,a^{j-1},\;a^k,\;a^{j+1},\,\ldots\,,a^n\Big)\\ &\quad=\;\det\left(a^1,\,\ldots\,,a^n\right)\;+\;\lambda\cdot 0\;=\;\det\left(a^1,\,\ldots\,,a^n\right).\end{aligned}$$

□

Definition 14.1.2. *Eine injektive Abbildung* $\{1, \dots, n\} \longrightarrow \{1, \dots, n\}$ *heißt* Permutation. *Die Menge aller dieser Permutationen, versehen mit der Komposition als Verknüpfung, bildet die sog.* symmetrische Gruppe S_n.

Bemerkung: (1.) Es ist recht einfach nachzuweisen, dass die Komposition von Abbildungen eine Gruppenstruktur auf den Permutationen induziert. Dabei ist die Identität

$$\{1, \dots, n\} \longrightarrow \{1, \dots, n\}, \quad j \longmapsto j$$

das *neutrale Element*, und das *inverse Element* einer Permutation erhalten wir jeweils durch Umkehrung der Zuordnung.

(2.) S_n enthält $n!$ Elemente (dem ersten können n Elemente zugeordnet werden, dem zweiten $n-1$, dem dritten $n-2$ ...). Die Elemente σ von S_n werden in der Form notiert

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & \dots & n \\ \sigma(1) & \sigma(2) & \sigma(3) & \dots & \sigma(n) \end{pmatrix}.$$

Definition 14.1.3 (Signum)**.** *Sei* $\sigma \in S_n$. *Wir sagen ein Fehlstand liegt vor, falls* $i < j$ *aber* $\sigma(i) > \sigma(j)$ *und wir bezeichnen mit* $r(\sigma)$ *die Zahl aller Fehlstände von* σ.

Das Signum *einer Permutation* $\sigma : \{1, \dots, n\} \longrightarrow \{1, \dots, n\}$ *wird damit definiert als*

$$\operatorname{sign} \sigma := (-1)^{r(\sigma)}.$$

Beispiel 14.1.4. *Die Permutation*

$$\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$$

hat zwei Fehlstände: $1 < 2$, *aber* $3 > 1$ *und* $1 < 3$ *aber* $3 > 2$. *Also gilt* $\operatorname{sign} \sigma = 1$.

Bemerkung: Für zwei Permutationen σ und τ gilt immer

$$\operatorname{sign}(\sigma \circ \tau) = \operatorname{sign} \sigma \cdot \operatorname{sign} \tau.$$

Eine Permutation $\tau \in S_n$ heisst Transposition, wenn sie zwei verschiedene Zahlen vertauscht und alle anderen festlässt. Genauer gesagt, wenn es $1 \leq i < j \leq n$ gibt mit $\tau(i) = j$, $\tau(j) = i$ und $\tau(k) = k$ für alle $k \in \{1, \dots, n\} \setminus \{i, j\}$.

Offensichtlich gilt für jede Transposition $\tau \in S_n$: $\tau^{-1} = \tau$ und $\operatorname{sign} \tau = -1$.

Jede Permutation $\sigma \in S_n$ als Verknüpfung endlich vieler Transpositionen schreiben lässt. Genauer existieren ein $k \in \mathbb{N}$ und Transpositionen $\tau_1, \dots, \tau_k \in S_n$ mit

$$\sigma = \tau_1 \circ \dots \circ \tau_k.$$

Satz 14.1.2 (Existenz und Eindeutigkeit der Determinante)**.**
Die Determinantenfunktion $\det : \mathbb{K}^{n,n} \longrightarrow \mathbb{K}$ *existiert und ist eindeutig bestimmt durch*

$$\det\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} = \sum_{\sigma\in S_n} a_{\sigma(1),1}\cdots a_{\sigma(n),n}\cdot \operatorname{sign}\sigma$$
$$= \sum_{\sigma\in S_n} a_{1,\sigma(1)}\cdots a_{n,\sigma(n)}\cdot \operatorname{sign}\sigma\,.$$

Beweis: Auf Grund der obigen Rechenregeln gilt (zunächst nur formal)

$$\det\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} = \det\left(\sum_{k_1=1}^{n} a_{k_1,1}\, e^{k_1}, \dots, \sum_{k_n=1}^{n} a_{k_n,n}\, e^{k_n}\right)$$
$$= \sum_{k_1=1}^{n}\cdots\sum_{k_n=1}^{n} a_{k_1 1}\,\dots\, a_{k_n n}\cdot\det\left(e^{k_1},\dots,e^{k_n}\right).$$

Falls in der rechts stehenden Summe zwei der kanonischen Basisvektoren $e^{k_1}, \dots, e^{k_n}$ gleich sind, so ist die zugehörige Determinante gleich 0 nach Satz 14.1.1 (a)(ii). Also brauchen wir nur die Indextupel $(k_1, \dots, k_n)$ gemäß einer Permutation $\sigma : \{1, \dots, n\} \longrightarrow \{1, \dots, n\}$ zu berücksichtigen:

$$\det\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} = \sum_{\sigma\in S_n} a_{\sigma(1),1}\,\dots\, a_{\sigma(n),n}\cdot\det\left(e^{\sigma(1)},\dots,e^{\sigma(n)}\right)$$
$$= \sum_{\sigma\in S_n} a_{\sigma(1),1}\cdots a_{\sigma(n),n}\cdot \operatorname{sign}\sigma.$$

Wenn wir zu jeder Permutation $\sigma \in S_n$ ihre Umkehrung σ^{-1} betrachten, so ergibt die obige Bemerkung $\operatorname{sign}(\sigma^{-1}) = \operatorname{sign}\sigma$, und wir erhalten schließlich

$$\det\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} = \sum_{\sigma\in S_n} a_{1,\sigma^{-1}(1)}\,\dots\, a_{n,\sigma^{-1}(n)}\cdot \operatorname{sign}(\sigma^{-1})$$
$$= \sum_{\tilde\sigma\in S_n} a_{1,\tilde\sigma(1)}\,\dots\, a_{n,\tilde\sigma(n)}\cdot \operatorname{sign}\tilde\sigma.$$

Dies zeigt die Eindeutigkeit der Determinante. Um die Existenz zu beweisen, muss man nachrechen das die obige Formel die Definition einer Determinante erfüllt. □

Korollar 14.1.1. (*i*) *$\det A^t = \det A$ für $A \in \mathbb{R}^{n,n}$ bzw. $\det(A^*) = \overline{\det A}$ für $A \in \mathbb{C}^{n,n}$, wobei für $A = (a_{ij})$ $(1 \le i, j \le n)$ die Matrizen A^t und $A^\star$ durch $A^t = (a_{ji})$ bzw. $A^\star = (\overline{a}_{ji})$ definiert sind.*

(*ii*) *Alle bisherigen Aussagen bzgl. der Spalten gelten auch bzgl. der Zeilen.*

(*iii*) *Für jedes $A = \begin{pmatrix} a_{11} & 0 \\ c & B \end{pmatrix}$ mit $a_{11} \in \mathbb{K}$, $c \in \mathbb{K}^{n-1}$ und $B \in \mathbb{K}^{n-1,n-1}$ ist*

$$\det A = \sum_{\substack{\sigma \in S_n: \\ \sigma(1)=1}} a_{1,\,\sigma(1)} \quad \cdots \quad a_{n,\,\sigma(n)} \cdot \operatorname{sign} \sigma$$
$$= \sum_{\tau \in S_{n-1}} a_{11} \; a_{2,\,1+\tau(1)} \cdots a_{n,\,1+\tau(n-1)} \cdot \operatorname{sign} \tau \; = \; a_{11} \, \cdot \, \det B.$$

Eine weitere wichtige Konsequenz ist der

Satz 14.1.3 (LAPLACE'sche Entwicklungssatz für Determinanten)**.**
Für alle $i, j \in \{1, \ldots, n\}$ bezeichne $A_{ij} \in \mathbb{K}^{(n-1),(n-1)}$ die sog. i, j–te Minorante *von $A \in \mathbb{K}^{n,n}$, die durch Streichung der i–ten Zeile und j–ten Spalte entsteht. Dann gilt für jedes $i \in \{1, \ldots, n\}$ bzw. $j \in \{1, \ldots, n\}$*

$$\det \begin{pmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \ldots & a_{nn} \end{pmatrix} = \sum_{k=1}^{n} (-1)^{i+k} \quad a_{ik} \quad \det(A_{ik})$$
$$= \sum_{k=1}^{n} (-1)^{k+j} \quad a_{kj} \quad \det(A_{kj}),$$

die sog. Entwicklung nach der i–ten Zeile *bzw.* nach der j–ten Spalte*.*

Beweis: Wir beschränken uns auf die Entwicklung nach der i–ten Zeile:
Durch $i-1$ Zeilenvertauschungen und später $j-1$ Spaltenvertauschungen erhält

man

$$\det\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{pmatrix} = (-1)^{i-1}\cdot\det\begin{pmatrix} a_{i1} & \dots & a_{in} \\ a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{i-1,1} & \dots & a_{i-1,n} \\ a_{i+1,1} & \dots & a_{i+1,n} \\ \vdots & & \vdots \\ a_{n1} & & a_{nn} \end{pmatrix}$$

$$= (-1)^{i-1}\cdot\sum_{j=1}^{n}\det\begin{pmatrix} 0 & \dots & a_{ij} & \dots & 0 \\ a_{11} & \dots & a_{1j} & \dots & a_{1n} \\ \vdots & & \vdots & & \vdots \\ a_{i-1,1} & \dots & a_{i-1,j} & \dots & a_{i-1,n} \\ a_{i+1,1} & \dots & a_{i+1,j} & \dots & a_{i+1,n} \\ \vdots & & \vdots & & \vdots \\ a_{n,1} & \dots & a_{n,j} & \dots & a_{n,n} \end{pmatrix}$$

$$= \sum_{j=1}^{n}(-1)^{i+j-2}\cdot\det\begin{pmatrix} a_{ij} & 0 & \dots & 0 & 0 & \dots & 0 \\ a_{1j} & a_{11} & \dots & a_{1,j-1} & a_{1,j+1} & \dots & a_{1n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ a_{i-1,j} & a_{i-1,1} & \dots & a_{i-1,j-1} & a_{i-1,j+1} & \dots & a_{i-1,n} \\ a_{i+1,j} & a_{i+1,1} & \dots & a_{i+1,j-1} & a_{i+1,j+1} & \dots & a_{i+1,n} \\ \vdots & \vdots & & \vdots & \vdots & & \vdots \\ a_{n,j} & a_{n1} & \dots & a_{n,j-1} & a_{n,j+1} & \dots & a_{n,n} \end{pmatrix}$$

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

□

Beispiel: (1.) $n = 2$

$$\det\begin{pmatrix} a & b \\ c & d \end{pmatrix} = a\,d - b\,c.$$

(2.) Als (noch relativ überschaubaren) Spezialfall erhält man für $n = 3$ die Sarrus**'sche Regel**

$$\det\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = a_{11}\,a_{22}\,a_{33} + a_{12}\,a_{23}\,a_{31} + a_{13}\,a_{21}\,a_{32}$$
$$- a_{13}\,a_{22}\,a_{31} - a_{11}\,a_{23}\,a_{32} - a_{12}\,a_{21}\,a_{33}.$$

Bemerkung: Matrixprodukt

Jede $n\times m$–Matrix $A = (a_{ij})_{\substack{1\le i\le n\\ 1\le j\le m}}$ definiert in naheliegender Weise eine lineare Abbildung $A:\mathbb{K}^m \longrightarrow \mathbb{K}^n$ und umgekehrt.
Ist nun A eine $n\times m$–Matrix und B eine $\ell\times n$–Matrix, so wird die Komposition $BA:\mathbb{K}^m \longrightarrow \mathbb{K}^\ell$ der zugehörigen linearen Abbildungen $A:\mathbb{K}^m \longrightarrow \mathbb{K}^n$,

$B : \mathbb{K}^n \longrightarrow \mathbb{K}^\ell$ durch das sog. *Matrixprodukt* beschrieben:

$$B\,A \;=\; \begin{pmatrix} \sum\limits_{j=n}^{n} b_{1j}\,a_{j1} & \dots & \sum\limits_{j=1}^{n} b_{1j}\,a_{jn} \\ \vdots & & \\ \sum\limits_{j=1}^{n} b_{\ell j}\,a_{j1} & \dots & \sum\limits_{j=1}^{n} b_{\ell j}\,a_{jn} \end{pmatrix} \;\in\; \mathbb{K}^{\ell,m}\,.$$

Diese Darstellung lässt sich recht einfach direkt nachrechnen:

$$\begin{aligned} (B\,A)\,x \;=\; B\,(A\,x) \;&=\; \begin{pmatrix} b_{11} & \dots & b_{1n} \\ \vdots & & \\ b_{\ell 1} & \dots & b_{\ell n} \end{pmatrix} \begin{pmatrix} \sum\limits_{j=1}^{m} a_{1j}\,x_j \\ \vdots \\ \sum\limits_{j=1}^{n} a_{nj}\,x_j \end{pmatrix} \\ &=\; \begin{pmatrix} \sum\limits_{i=1}^{n} b_{1i} \sum\limits_{j=1}^{m} a_{ij}\,x_j \\ \vdots \\ \sum\limits_{i=1}^{n} b_{\ell i} \sum\limits_{j=1}^{m} a_{ij}\,x_j \end{pmatrix} \;=\; \begin{pmatrix} \sum\limits_{j=1}^{m} \Big(\sum\limits_{i=1}^{n} b_{1i}\,a_{ij} \Big)\,x_j \\ \vdots \\ \sum\limits_{j=1}^{m} \Big(\sum\limits_{i=1}^{n} b_{\ell i}\,a_{ij} \Big)\,x_j \end{pmatrix}. \end{aligned}$$

Dieses Rechenschema wird überschaubarer, wenn man den Blick auf die Spaltenvektoren von A richtet: Der k–te Spaltenvektor von A wird mit B durch ein Matrix-Vektor-Produkt verknüpft, und das Resultat wird zum k–ten Spaltenvektor der Produktmatrix BA.

Das Matrixprodukt ist assoziativ und distributiv. Bereits einfache Beispiele zeigen aber, dass das Matrixprodukt im Allgemeinen *nicht* kommutativ ist.

Speziell für quadratische Matrizen ergibt sich damit der

Satz 14.1.4 (Determinantenmultiplikationssatz)**.** *Für $A, B \in \mathbb{K}^{n,n}$ gilt*

$$\det(BA) \;=\; \det A \cdot \det B \;=\; \det(AB).$$

Beweis:

<u>1. Fall</u>: $\det B = 0$.

Dann sind die Spaltenvektoren von B linear abhängig. Das Bild der entsprechenden linearen Abbildung B (dies ist die lineare Hülle der Spaltenvektoren) ist also $\neq \mathbb{K}^n$. Dann ist aber auch das Bild des Produktes BA ungleich $\mathbb{K}^n$, d.h. die Spaltenvektoren von BA sind auch linear abhängig. Also ist

$$0 \;=\; \det(BA) \;=\; 0 \cdot \det A \;=\; \det B \cdot \det A.$$

<u>2. Fall</u>: $\det B \neq 0$.

Mit Blick auf die Identität

$$\det(BA) \;=\; \frac{\det(BA)}{\det B}\;\det B$$

folgern wir aus den Eigenschaften des Matrixprodukts: $A \longmapsto \frac{\det(BA)}{\det B}$ ist

(i) linear in jeder Spalte von A (denn das Matrixprodukt BA hängt immer linear von jeder Spalte von A ab),

(ii) alternierend bzgl. Spaltenvektoren (denn das Vertauschen zweier Spalten von A bedingt das Vertauschen der entsprechenden Spalten von BA),

(iii) $= 1$, falls $A = \mathbb{E}_n = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$ ist.

Also ist $\mathbb{K}^{n,n} \longrightarrow \mathbb{K}$, $A \longmapsto \frac{\det(BA)}{\det B}$ eine normierte alternierende Multilinearform, und die Eindeutigkeit aus Satz 14.1.2 ergibt $\frac{\det(BA)}{\det B} \;=\; \det A$.
Die zweite Aussage ist nun offensichtlich. □

Definition 14.1.5. *Eine Matrix $A \in \mathbb{K}^{n,n}$ heißt* <u>invertierbar</u>*, falls es eine Matrix $A^{-1} \in \mathbb{K}^{n,n}$ mit $A^{-1}A \;=\; A\,A^{-1} \;=\; \mathbb{E}_n$ gibt. A^{-1} heißt* Inverse *zu A.*

Bemerkung: (1.) Allgemein kann eine $n \times m$–Matrix A (dann und) nur dann eine "inverse" Matrix $B \in \mathbb{K}^{m,n}$ mit $B\,A \;=\; \mathbb{E}_m$, $A\,B \;=\; \mathbb{E}_n$ besitzen, wenn die zugehörige lineare Abbildung $\mathbb{K}^m \longrightarrow \mathbb{K}^n$, $x \longmapsto A\,x$ bijektiv ist. Also muss notwendigerweise gelten: $m = n$.

(2.) $A \in \mathbb{K}^{n,n}$ ist invertierbar $\overset{13.2.1}{\Longleftrightarrow}$ Spaltenvektoren sind linear unabhängig
$\overset{13.2.3}{\Longleftrightarrow}$ Zeilenvektoren sind linear unabhängig
$\overset{14.1.1}{\Longleftrightarrow}$ $\det A \neq 0$

Korollar 14.1.2. *Für jede invertierbare Matrix $A \in \mathbb{K}^{n,n}$ gilt*

$$\det(A^{-1}) \;=\; (\det A)^{-1}.$$

□

Wir können nun die Lösung von linearen Gleichungssystemen mit n Gleichungen und n Unbekannten im Fall der eindeutigen Lösbarkeit geschlossen angeben.

Satz 14.1.5 (Cramer'sche Regel für lineare Gleichungssysteme)**.** *$A \in \mathbb{K}^{n,n}$ erfülle $\det A \neq 0$. Dann ist die eindeutig bestimmte Lösung*

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{K}^n \quad \textit{von } Ax = b \quad \textit{mit } b = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \in \mathbb{K}^n$$

gegeben durch

$$x_i = \frac{1}{\det A} \cdot \det \begin{pmatrix} a_{11} & \dots & a_{1,i-1} & b_1 & a_{1,i+1} & \dots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & & a_{n,i-1} & b_n & a_{n,i+1} & \dots & a_{nn} \end{pmatrix} \quad \textit{für } i = 1, \dots, n.$$

Beweis: Das System $A\,x = b$ können wir in der Form

$$x_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{n1} \end{pmatrix} + \dots + x_n \begin{pmatrix} a_{1n} \\ \vdots \\ a_{nn} \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$$

darstellen bzw. für jedes $i = 1, \dots, n$

$$x_1 \begin{pmatrix} a_{11} \\ \vdots \\ a_{n1} \end{pmatrix} + \dots + x_{i-1} \begin{pmatrix} a_{1,i-1} \\ \vdots \\ a_{n,i-1} \end{pmatrix} + \begin{pmatrix} x_i\, a_{1i} - b_1 \\ \vdots \\ x_i\, a_{ni} - b_n \end{pmatrix} + x_{i+1} \begin{pmatrix} a_{1,i+1} \\ \vdots \\ a_{n,i+1} \end{pmatrix} + \dots + x_n \begin{pmatrix} a_{1n} \\ \vdots \\ a_{nn} \end{pmatrix} = 0.$$

Die Spalten der Matrix

$$\begin{pmatrix} a_{11} & \dots & a_{1,i-1} & x_i\, a_{1i} - b_1 & a_{1,i+1} & \dots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \dots & a_{n,i-1} & x_i\, a_{ni} - b_n & a_{n,i+1} & \dots & a_{nn} \end{pmatrix}$$

sind also linear abhängig. Für jeden Index i folgt also wegen der Linearität in der i–ten Spalte

$$x_i \cdot \det A - \det \begin{pmatrix} a_{11} & \dots & a_{1,i-1} & b_1 & a_{1,i+1} & \dots & a_{1n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & \dots & a_{n,i-1} & b_n & a_{n,i+1} & \dots & a_{nn} \end{pmatrix} = 0$$

Auflösen nach x_i ergibt die CRAMER'sche Regel. □

Satz 14.1.6 (Matrixinversion). *$A \in \mathbb{K}^{n,n}$ sei invertierbar (d.h. $\det A \neq 0$). Dann ist*

$$A^{-1} = \frac{1}{\det A}\, B$$

mit den Minoranten $A_{ij} \in \mathbb{K}^{n-1,n-1}$ aus Satz 14.1.3 und

$$B := \left((-1)^{i+j} \cdot \det A_{ji}\right) = \begin{pmatrix} \det A_{11} & -\det A_{21} & \det A_{31} & \dots \\ -\det A_{12} & \det A_{22} & -\det A_{32} & \dots \\ \det A_{13} & -\det A_{23} & \det A_{33} & \dots \\ \vdots & \vdots & \vdots & \end{pmatrix}$$

Speziell ist im Fall $n = 2$: $\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \dfrac{1}{ad - bc} \cdot \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$

Beweis: Wegen $AA^{-1} = \mathbb{E}_n$ ist der j–te Spaltenvektor $\left(a^-_{ij}\right)_{1 \le i \le n} \in \mathbb{K}^n$ der inversen Matrix $A^{-1} = \left(a^-_{ik}\right)_{1 \le i,k \le n} \in \mathbb{K}^{n,n}$ die Lösung des Gleichungssystems

$$A\,x \;=\; e^j \;=\; \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \qquad \text{(mit 1 an } j\text{-ter Stelle).}$$

Also folgt aus der CRAMER'schen Regel für $i = 1, \ldots, n$

$$a_{ij}^{-} \;=\; \tfrac{1}{\det A} \cdot \det \begin{pmatrix} a_{11} & \ldots & a_{1,i-1} & 0 & a_{1,i+1} & \ldots & a_{1n} \\ \vdots & & \vdots & 1 & \vdots & & \vdots \\ a_{n1} & & a_{n,i-1} & 0 & a_{n,i+1} & \ldots & a_{nn} \end{pmatrix},$$

wobei die i–te Spalte der letzten Matrix nur in der j–ten Zeile eine 1 und in den anderen Zeilen 0 hat. (Dabei ist zu beachten, dass i, j hier ihre Rolle als Zeilen- bzw. Spaltenindex gewechselt haben.)
Durch $(i-1)$–faches Vertauschen benachbarter Spalten und anschließendes $(j-1)$–faches Vertauschen benachbarter Zeilen erhalten wir

$$\begin{aligned} a_{ij}^{-} &= \tfrac{1}{\det A} \cdot (-1)^{(i-1)+(j-1)} \det \begin{pmatrix} 1 & 0 \\ 0 & A_{ji} \end{pmatrix} \\ &= \tfrac{1}{\det A} \cdot (-1)^{i+j} \qquad\qquad \det A_{ji} \end{aligned}$$

□

14.2 Eigenwerte und Eigenvektoren linearer Abbildungen

Definition 14.2.1. *V sei ein Vektorrraum über $\mathbb{K}$ und $A : V \longrightarrow V$ eine lineare Abbildung. Ein Skalar $\lambda \in \mathbb{K}$ heißt ein* Eigenwert (EW) *von A, wenn es einen Vektor $x \in V \setminus \{0\}$ gibt mit $A\,x = \lambda\,x$. Solch ein Vektor $x \neq 0$ heißt* Eigenvektor (EV) *oder auch* Eigenelement (EE) *von A zum Eigenwert λ.*

BEMERKUNG: Jede Linearkombination $\alpha_1\,x_1 + \alpha_2\,x_2$ zweier Eigenvektor $x_1, x_2 \in V$ von A zum selben Eigenwert $\lambda \in \mathbb{K}$ ist wiederum ein Eigenvektor von A zum Eigenwert λ, denn es gilt

$$A\,(\alpha_1\,x_1 + \alpha_2\,x_2) \;=\; \alpha_1\,A\,x_1 + \alpha_2\,A\,x_2 \;=\; \lambda\;(\alpha_1\,x_1 + \alpha_2\,x_2)$$

Daher bilden alle Eigenvektoren von A zu einem Eigenwert $\lambda \in \mathbb{K}$ (gemeinsam mit dem Nullvektor) einen Untervektorraum von V.

BEMERKUNG: V sei ein n–dimensionaler $\mathbb{K}$–Vektorraum und $\{v^1, \ldots, v^n\}$ eine Basis von V. Dann beschreibt die lineare Abbildung

$$U : V \;\longrightarrow\; \mathbb{K}^n, \quad \sum_{j=1}^{n} c_j\,v^j \;\longmapsto\; (c_1, \ldots, c_n)$$

wie die Vektoren aus V jeweils durch ein Komponententupel in $\mathbb{K}^n$ bzgl. dieser Basis dargestellt werden. Die Transformation U ist immer bijektiv und hat die Umkehrabbildung

$$U^{-1} : \mathbb{K}^n \longrightarrow V, \quad (c_1, \ldots, c_n) \longmapsto \sum_{j=1}^{n} c_j\, v^j.$$

Zu jeder linearen Abbildung $A : V \longrightarrow V$ und gegebenen Basis $\{v^1, \ldots, v^n\}$ lassen sich die Koeffizienten $a_{ij} \in \mathbb{K}$ $(i, j = 1, \ldots, n)$ nun so wählen, dass für jedes $j = 1, \ldots, n$ gilt: $\quad A\, v^j = \sum\limits_{i=1}^{n} a_{ij}\, v^i.$

Dann ist $(a_{ij})_{1 \le i,j \le n}$ eine Matrix in $\mathbb{K}^{n,n}$, und wir finden sie wieder bei der Komposition $U\, A\, U^{-1} : \mathbb{K}^n \longrightarrow \mathbb{K}^n$ in Form des Matrix-Vektor-Produkts:

$$\begin{aligned}
\Big(U\, A\, U^{-1}\Big)(c_1, \ldots, c_n) &= U\, A \sum_{j=1}^{n} c_j\, v^j = U \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij}\, c_j\, v^i \\
&= \left(\sum_{j=1}^{n} a_{1j}\, c_j, \ \ldots, \sum_{j=1}^{n} a_{nj}\, c_j \right),
\end{aligned}$$

d.h. $UAU^{-1} : \mathbb{K}^n \longrightarrow \mathbb{K}^n$ ist die durch die Matrix (a_{ij}) in $\mathbb{K}^{n,n}$ erzeugte lineare Abbildung. Man sagt: Die lineare Abbildung $A : V \longrightarrow V$ wird bezüglich der Basis $\{v^1, \ldots, v^n\}$ durch die Matrix (a_{ij}) erzeugt.

Offenbar ist λ genau dann ein Eigenwert von A mit einem Eigenvektor x, wenn λ Eigenvektor von UAU^{-1} ist mit Eigenvektor Ux. Dies folgt aus der Gleichung

$$(UAU^{-1})(Ux) = U(Ax) = \lambda Ux.$$

Wir können uns deshalb im endlichdimensionalen Fall auf die Untersuchung von linearen Abbildungen $\mathbb{K}^n \longrightarrow \mathbb{K}^n$ (und ihre erzeugenden $n \times n$–Matrizen) konzentrieren.

14.3 Diagonalisierbare lineare Abbildungen

Definition 14.3.1. *V sei ein endlichdimensionaler $\mathbb{K}$–Vektorraum.*
Eine lineare Abbildung $A : V \longrightarrow V$ heißt diagonalisierbar, *wenn eine Basis von V existiert, bezüglich der A durch eine Diagonalmatrix $(a_{ij})_{1 \le i,j \le n} \in \mathbb{K}^{n,n}$ (d.h. $a_{ij} = 0$ für alle $i \neq j$) dargestellt wird.*

Angesichts der Bemerkung im vorherigen Abschnitt ist eine direkte Konsequenz der Definitionen:

Satz 14.3.1. *Eine lineare Abbildung A in einem endlichdimensionalen Vektorraum V ist genau dann diagonalisierbar, wenn V eine Basis aus Eigenvektoren von A besitzt. A ist dann bezüglich dieser Basis durch die Diagonalmatrix erzeugt, auf deren Diagonalen die Eigenwerte von A stehen.* □

Offenbar sind diagonalisierbare Abbildungen besonders einfach zu untersuchen. Es ist deshalb eine wichtige Frage, woran man die Diagonalisierbarkeit erkennt.

Beispiel: Die Spiegelung S an der e^1–Achse in $\mathbb{R}^2$ wird durch die Matrix $\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ dargestellt, hat also die Eigenwerte 1 und -1 mit (den) zugehörigen Eigenvektoren e^1 und e^2. (Es gilt entsprechend in $\mathbb{C}^2$ – wenn auch eine geometrische Interpretation als Spiegelung dann "deutlich schwieriger" ist.)

Beispiel: Die Drehung D um $90° \sim \pi/2$ in $\mathbb{R}^2$ (in math. positiver Richtung) wird bekanntlich durch die Matrix

$$D := \begin{pmatrix} \cos\frac{\pi}{2} & -\sin\frac{\pi}{2} \\ \sin\frac{\pi}{2} & \cos\frac{\pi}{2} \end{pmatrix} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

dargestellt. Sie hat keinen reellen Eigenwert. Dies erscheint geometrisch plausibel, denn kein Vektor $\neq 0$ kann seine Richtung beibehalten. Rechnerisch werden Eigenwert und -vektor charakterisiert durch die Bedingung:

$$D\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \lambda\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \iff \begin{cases} -x_2 = \lambda\, x_1 \\ \quad x_1 = \lambda\, x_2 \end{cases} \iff \begin{cases} -x_2 = \lambda^2\, x_2 \\ \quad x_1 = -\lambda^2\, x_1 \end{cases}$$

und letzteres besitzt keine Lösung $\lambda \in \mathbb{R}$ mit $(x_1, x_2) \neq (0, 0)$.

In $\mathbb{C}^2$ hat dieselbe Matrix offenbar die Eigenwerte $\lambda = \pm i$ mit den Eigenvektoren $u_+ = \begin{pmatrix} i \\ 1 \end{pmatrix}$ für $\lambda = i$ und $u_- = \begin{pmatrix} 1 \\ i \end{pmatrix}$ für $\lambda = -i$.

14.4 Das charakteristische Polynom

Wie findet man Eigenwerte und Eigenvektoren einer Matrix $A = (a_{ij})_{1 \le i,j \le n}$?

Damit $\lambda \in \mathbb{K}$ Eigenwert ist, muss offenbar die homogene Gleichung

$$\begin{pmatrix} a_{11}-\lambda & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22}-\lambda & \dots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn}-\lambda \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = 0$$

eine nichttriviale Lösung in $\mathbb{K}^n$ haben. Unter Verwendung von Korollar 13.1.1 und Satz 14.1.1 (b) ist das gleichbedeutend mit

$$\begin{aligned} & \text{Kern } (A - \lambda\,\mathbb{E}_n) \neq \{0\} \\ \iff \quad & \det(A - \lambda\,\mathbb{E}_n) = 0. \end{aligned}$$

Definition 14.4.1. *A sei eine $n \times n$–Matrix über $\mathbb{K}$.*
$\lambda \longmapsto \det(A - \lambda\,\mathbb{E}_n)$ ist ein Polynom vom Grad n und heißt charakteristisches Polynom von A. (Seine Nullstellen sind genau die Eigenwerte von A.)

Der Fundamentalsatz der Algebra garantiert für jedes komplexe Polynom n-ten Grades, dass mindestens eine und höchstens n Nullstellen in $\mathbb{C}$ existieren. Daraus folgt hier direkt:

Satz 14.4.1. *Jede Matrix $A \in \mathbb{C}^{n,n}$ hat mindestens einen komplexen Eigenwert und höchstens n verschiedene Eigenwerte in $\mathbb{C}$.*

Zugehörige Eigenvektoren lassen sich dann mit Hilfe des Gauß'schen Eliminationsverfahren bestimmen.

Beispiel:

$$A = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{pmatrix}, \quad A - \lambda\,\mathbb{E}_3 = \begin{pmatrix} 2-\lambda & 0 & 0 \\ 0 & -\lambda & 1 \\ 0 & -1 & -\lambda \end{pmatrix}.$$

Das charakteristische Polynom ist $p_A(\lambda) := (2-\lambda)\,(\lambda^2+1)$ mit den Nullstellen $\lambda = 2, i, -i$. Zugehörige Eigenvektoren sind

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ i \end{pmatrix}, \begin{pmatrix} 0 \\ i \\ 1 \end{pmatrix}.$$

Satz 14.4.2. *Die Matrix $A \in \mathbb{K}^{n,n}$ besitze (paarweise) verschiedene Eigenwerte $\lambda_1, \ldots, \lambda_r \in \mathbb{K}$ und die zugehörigen Eigenvektoren $x^1, \ldots, x^r \in \mathbb{K}^n$.*

Dann sind $\{x^1, \ldots, x^r\}$ linear unabhängig.

(Die entsprechende lineare Unabhängigkeit gilt auch für beliebig viele verschiedene Eigenwerte einer linearen Abbildung $V \longrightarrow V$ mit einem endlichdimensionalen $\mathbb{K}$-Vektorraum V.)

Beweis: Induktionsanfang $r = 1$: Die Definition garantiert $x^1 \neq 0$, und damit ist die lineare Unabhängigkeit trivial.

Induktionsschritt $r-1 \longrightarrow r$: $x^1, \ldots, x^{r-1}$ seien bereits linear unabhängig. Nun werden $c_1, \ldots, c_r \in \mathbb{K}$ gewählt mit $\sum\limits_{i=1}^{r} c_i\, x^i = 0$. Dann gilt

$$\begin{aligned} 0 &= A\left(\sum_{i=1}^{r} c_i\, x^i\right) &&= \sum_{i=1}^{r} \lambda_i\, c_i\, x^i \\ \Longrightarrow \quad 0 &= \lambda_r \sum_{i=1}^{r} c_i\, x^i - \sum_{i=1}^{r} \lambda_i\, c_i\, x^i &&= \sum_{i=1}^{\mathbf{r-1}} (\lambda_r - \lambda_i)\, c_i\, x^i. \end{aligned}$$

Die lineare Unabhägigkeit von $x^1, \ldots, x^{r-1}$ ergibt $(\lambda_r - \lambda_i)\, c_i = 0$ und schließlich $c_i = 0$ für $i = 1, \ldots, r-1$, da die Eigenwerte (paarweise) verschieden sind. Wegen $\sum\limits_{i=1}^{r} c_i\, x^i = 0$ ist dann auch $c_r = 0$. $\square$

15 Euklidische und unitäre Vektorräume

15.1 Skalarprodukte

Definition 15.1.1. *V sei ein Vektorraum über $\mathbb{K}$.*
Ein Skalarprodukt *auf V ist eine Funktion $\langle \cdot, \cdot \rangle : V \times V \longrightarrow \mathbb{K}$ mit folgenden Eigenschaften:*

$$
\begin{array}{llcll}
(S1) & \langle x,\ x\rangle & \geq & 0 & \text{für alle } x \in V, \\
 & \langle x,\ x\rangle & = & 0 \iff x = 0, & \\
(S2) & \langle x,\ y\rangle & = & \overline{\langle y,\ x\rangle} & \text{für alle } x, y \in V, \\
(S3) & \langle x,\ \lambda\, y\rangle & = & \lambda\, \langle x,\ y\rangle & \text{für alle } x, y \in V,\ \lambda \in \mathbb{K}, \\
(S4) & \langle x,\ y + z\rangle & = & \langle x,\ y\rangle + \langle x,\ z\rangle & \text{für alle } x, y, z \in V.
\end{array}
$$

Ein reeller Vektorraum mit einem Skalarprodukt heißt euklidischer Vektorraum*, und ein komplexer Vektorraum mit einem Skalarprodukt heißt* unitärer Vektorraum*.*

Bemerkung: Jedes Skalarprodukt $\langle \cdot, \cdot \rangle$ auf einem $\mathbb{K}$–Vektorraum V induziert eine *Norm* auf V gemäß $\|x\|_V := \sqrt{\langle x,\ x\rangle}$ für $x \in V$.

Beispiel: Das sog. Standard-Skalarprodukt auf $\mathbb{R}^n$ bzw. $\mathbb{C}^n$ ist

$$
(x, y)_{\mathbb{K}^n} := \begin{cases} \sum\limits_{j=1}^{n} x_j\, y_j & \text{in } \mathbb{R}^n, \\ \sum\limits_{j=1}^{n} \overline{x}_j\, y_j & \text{in } \mathbb{C}^n, \end{cases}
$$

für $x = (x_1, \ldots, x_n),\ y = (y_1, \ldots, y_n)$.

Lemma 15.1.1. *V sei ein $\mathbb{K}$-Vektorraum und $\{v^1, \ldots, v^n\}$ eine Basis von V. Zu jedem Skalarprodukt $\langle \cdot, \cdot \rangle$ auf V gibt es genau eine Matrix $A = (a_{ij}) \in \mathbb{K}^{n,n}$, so dass für alle Vektoren $v = \sum_{j=1}^{n} \alpha_j\, v^j$ und $w = \sum_{j=1}^{n} \beta_j\, v^j$ in V gilt:*

$$
\begin{aligned}
\langle v,\ w\rangle &= \Big((\alpha_i)_{1 \leq i \leq n},\ A\, (\beta_k)_{1 \leq k \leq n} \Big)_{\mathbb{C}^n} \\
&= \left((\alpha_i)_{1 \leq i \leq n},\ \left(\sum_{k=1}^{n} a_{jk}\, \beta_k \right)_{1 \leq j \leq n} \right)_{\mathbb{C}^n} \\
&= \sum_{j,k=1}^{n} \overline{\alpha}_j\ a_{jk}\ \beta_k\, .
\end{aligned}
$$

Dabei ist $a_{jk} = \langle v^j,\ v^k\rangle$ für alle $j, k = 1, \ldots, n$. □

Definition 15.1.2. *V sei ein $\mathbb{K}$–Vektorraum mit einem Skalarprodukt $\langle \cdot, \cdot \rangle$. Zwei Vektoren $v, w \in V$ heißen* <u>orthogonal</u> *(bzgl. $\langle \cdot, \cdot \rangle$), falls $\langle v,\, w \rangle \;=\; 0$ ist.*

Eine Basis $\{v^1, \ldots, v^m\}$ von V heißt <u>Orthogonalbasis</u>*, falls je zwei ihrer Vektoren orthogonal sind, d.h. wenn $\langle v^i,\, v^j \rangle \;=\; 0$ für alle Indizes $i \neq j$ gilt.*

Eine Basis $\{v^1, \ldots, v^m\}$ von V heißt <u>Orthonormalbasis</u> (ONB)*, falls gilt:*

$$\langle v^i,\, v^j \rangle \;=\; \begin{cases} 0, & \textit{falls } i \neq j, \\ 1, & \textit{falls } i = j. \end{cases}$$

Lemma 15.1.2. *V sei ein $\mathbb{K}$–Vektorraum mit einem Skalarprodukt. Wenn jeweils zwei der Vektoren $\{w_1, \ldots, w_n\}$ in $V \setminus \{0\}$ orthogonal sind, dann sind $\{w_1, \ldots, w_n\}$ linear unabhängig.*

Beweis: Aus $0 \;=\; \sum\limits_{j=1}^{n} \lambda_j\, w_j$ folgt für jedes $k = 1, \ldots, n$

$$0 \;=\; \left\langle w_k,\, \sum_{j=1}^{n} \lambda_j\, w_j \right\rangle \;=\; \sum_{j=1}^{n} \lambda_j\, \langle w_k,\, w_j \rangle \;=\; \lambda_k\, \langle w_k,\, w_k \rangle$$

und wegen der Voraussetzung $w_k \neq 0$ muss $\lambda_k = 0$ sein. □

Beispiel: In $\mathbb{R}^n$ bzw. $\mathbb{C}^n$ bilden die kanonischen Einheitsvektoren $\{e^1, \ldots, e^n\}$ eine Orthonormalbasis, wobei jeweils $e^j = (0, \ldots, 1, \ldots, 0) \in \mathbb{K}^n$ die Zahl 1 an der j–ten Stelle besitze.
Doch die Existenz einer Orthonormalbasis ist keine Besonderheit von $\mathbb{K}^n$:

Satz 15.1.1. *Jeder endlichdimensionale Vektorraum V mit Skalarprodukt besitzt eine Orthonormalbasis.*

Beweis: $\{f^1, \ldots, f^n\}$ sei eine beliebige Basis von V.

Das folgende Gram–Schmidt'sche Orthonormalisierungverfahren führt dann sukzessive zu einer Orthonormalbasis von V:

$$\begin{aligned}
v^1 &:= \frac{1}{\|f^1\|}\, f^1, \\
v^2 &:= \frac{1}{\|f^2 - \langle v^1, f^2 \rangle\, v^1\|}\, \left(f^2 - \langle v^1, f^2 \rangle\, v^1\right), \quad \ldots \\
v^{j+1} &:= \frac{1}{\left\| f^{j+1} - \sum\limits_{i=1}^{j} \langle v^i, f^{j+1} \rangle\, v^i \right\|} \left(f^{j+1} - \sum_{i=1}^{j} \langle v^i, f^{j+1} \rangle\, v^i \right)
\end{aligned}$$

Offenbar sind die v^j normiert und paarweise orthogonal. Also sind sie nach Lemma 15.1.2 linear unabhängig. Da sie außerdem den gleichen Span haben wie $\{f^1, \ldots, f^n\}$, bilden sie eine Basis von V. □

15.2 Orthogonale und unitäre lineare Abbildungen

Definition 15.2.1. *V_1, V_2 seien $\mathbb{K}$–Vektorräume mit Skalarprodukten $\langle\cdot,\cdot\rangle_{V_j}$.*
Eine Abbildung $U : V_1 \longrightarrow V_2$ heißt orthogonal (bei $\mathbb{K} = \mathbb{R}$) bzw. unitär (bei $\mathbb{K} = \mathbb{C}$), wenn für alle $x, y \in V_1$ gilt:

$$\langle Ux,\, Uy\rangle_{V_2} \;=\; \langle x,\, y\rangle_{V_1}.$$

Bemerkung: Jede orthogonale Abbildung $U : V_1 \longrightarrow V_2$ ist *isometrisch* in folgendem Sinne: $\quad \|Ux\|_{V_2} \;=\; \|x\|_{V_1} \quad$ für alle $x \in V_1$.

Satz 15.2.1. *V_1 und V_2 seien n–dimensionale Vektorräume mit Skalarprodukt (beide über $\mathbb{R}$ oder beide über $\mathbb{C}$). $\{e^1, \dots, e^n\}$ sei eine ONB von V_1.*

Eine lineare Abbildung $U : V_1 \longrightarrow V_2$ ist genau dann orthogonal bzw. unitär, wenn $\{f^1, \dots, f^n\}$ mit $f^j := Ue^j$ eine ONB in V_2 ist.

Beweis: "$\Longrightarrow$" Wenn U orthogonal bzw. unitär ist, so folgt aus der Orthonormalität von $\{e^1, \dots, e^n\}$:

$$\langle f^i,\, f^j\rangle_{V_2} \;=\; \langle Ue^i,\, Ue^j\rangle_{V_2} \;=\; \langle e^i,\, e^j\rangle_{V_1} \;=\; \delta_{ij}$$

für alle i, j. Gemäß Lemma 15.1.2 ist $\{f^1, \dots, f^n\}$ linear unabhängig und somit eine ONB.

"$\Longleftarrow$" Unter Verwendung der ONB $\{e^1, \dots, e^n\}$ haben beliebige Elemente $x, y \in V_1$ die Form

$$x \;=\; \sum_{j=1}^{n} c_j\, e^j \quad \text{mit} \quad \sum_{j=1}^{n} |c_j|^2 \;=\; \|x\|^2,$$
$$y \;=\; \sum_{j=1}^{n} d_j\, e^j \quad \text{mit} \quad \sum_{j=1}^{n} |d_j|^2 \;=\; \|y\|^2.$$

Wenn $\{f^1, \dots, f^n\}$ eine ONB in V_2 ist, so folgt daraus

$$\begin{aligned}
\langle Ux,\, Uy\rangle_{V_2} \;&=\; \left\langle U\sum_{j=1}^{n} c_j\, e^j,\; U\sum_{j=1}^{n} d_j\, e^j\right\rangle_{V_2} \;=\; \left\langle \sum_{j=1}^{n} c_j\, U\, e^j,\; \sum_{j=1}^{n} d_j\, U\, e^j\right\rangle_{V_2} \\
&=\; \left\langle \sum_{j=1}^{n} c_j\, f^j,\quad \sum_{j=1}^{n} d_j\, f^j\right\rangle_{V_2} \;=\; \sum_{j=1}^{n} \overline{c_j}\, d_j \\
&=\; \left\langle \sum_{j=1}^{n} c_j\, e^j,\quad \sum_{j=1}^{n} d_j\, e^j\right\rangle_{V_1} \;=\; \langle x, y\rangle_{V_1}.
\end{aligned}$$

□

Korollar 15.2.1. (*1.*) *Unter den Annahmen von Proposition 15.2.1 ist jede orthogonale (bzw. unitäre) Abbildung $U : V_1 \longrightarrow V_2$ bijektiv, und ihre Umkehrung ist ebenfalls orthogonal (bzw. unitär).*

(*2.*) *Für eine Matrix $A = (a_{ij}) \in \mathbb{C}^{n,n}$ ist die lineare Abbildung $\mathbb{C}^n \longrightarrow \mathbb{C}^n$, $x \longmapsto A\,x$ genau dann unitär bzgl. $(\cdot,\cdot)_{\mathbb{C}^n}$, wenn die Spaltenvektoren von A eine Orthonormalbasis von $\mathbb{C}^n$ sind. Dann ist $\overline{A}^t\, A \;=\; \mathbb{E}_n \;=\; A\,\overline{A}^t$.*

15.3 Selbstadjungierte lineare Abbildungen

Satz 15.3.1. *V sei ein endlichdimensionaler Vektorraum über $\mathbb{K}$ mit einem Skalarprodukt $\langle\cdot,\cdot\rangle$.*

Zu jeder linearen Abbildung $A : V \longrightarrow V$ gibt es genau eine lineare Abbildung $A^* : V \longrightarrow V$, die sog. zu A adjungierte Abbildung, *mit*

$$\langle x,\, A\,y\rangle \;=\; \langle A^*\,x,\, y\rangle \qquad\qquad \textit{für alle } x, y \in V.$$

(Das Entsprechende gilt für jede lineare Abbildung A von einem Vektorraum $(V_1, \langle\cdot,\cdot\rangle_{V_1})$ nach $(V_2, \langle\cdot,\cdot\rangle_{V_2})$, A^* ist dann eine Abbildung von V_2 nach V_1.)

Beweis: Eindeutigkeit: B_1 und B_2 seien lineare Abbildungen $V \longrightarrow V$ mit der von A^* geforderten Eigenschaft. Dann gilt für alle $x, y \in V$

$$\langle (B_1 - B_2)\,x,\, y\rangle \;=\; \langle x,\, A\,y\rangle \;-\; \langle x,\, A\,y\rangle \;=\; 0,$$

also ist $(B_1 - B_2)\,x \;=\; 0$ für alle $x \in V$, d.h. $B_1 = B_2$.

Existenz: *Spezialfall* $V = \mathbb{K}^n$ mit dem Standard-Skalarprodukt $(\cdot,\cdot)_{\mathbb{K}^n}$:

Wird $A : \mathbb{K}^n \longrightarrow \mathbb{K}^n$ durch die Matrix $(a_{ij}) \in \mathbb{K}^{n,n}$ dargestellt, so wird A^* durch die *adjungierte Matrix* $\left(\overline{a_{ij}}\right)^t \;=\; \left(\overline{a_{ji}}\right)$ erzeugt:

$$\begin{aligned}(x,\, Ay)_{\mathbb{K}^n} \;&=\; \sum_{i=1}^{n}\left(\overline{x_i}\,\sum_{j=1}^{n} a_{ij}\,y_j\right)\\ &=\; \sum_{j=1}^{n}\left(y_j\,\sum_{i=1}^{n} a_{ij}\,\overline{x_i}\right)\\ &=\; \sum_{j=1}^{n}\left(y_j\,\overline{\left(\sum_{i=1}^{n}\overline{a_{ij}}\,x_i\right)}\right),\end{aligned}$$

also ist $\quad A^*x \;=\; \left(\sum\limits_{i=1}^{n}\overline{a_{i1}}\,x_i,\;\ldots\;,\sum\limits_{i=1}^{n}\overline{a_{in}}\,x_i\right).$

Allgemeiner Fall $(V, \langle\cdot,\cdot\rangle_V)$: V besitzt eine Orthonormalbasis $\{v_1, \ldots, v_n\}$ (bzgl. $\langle\cdot,\cdot\rangle_V$) nach Satz 15.1.1. Für $j = 1, \ldots, n$ bezeichne $e^j \in \mathbb{K}^n$ den j–ten kanonischen Einheitsvektor.

Dann gibt es genau ein lineare Abbildung $U : \mathbb{K}^n \longrightarrow V$ mit $U\,e^j = v_j$ für jedes $j = 1, \ldots, n$, wie in der Bemerkung in Abschnitt 14.2 erwähnt. Insbesondere ist U nun sogar orthonormal bzw. unitär laut Satz 15.2.1 (bzgl. des Standardskalarproduktes auf $\mathbb{K}^n$): $\langle U\,x,\, U\,y\rangle_V \;=\; (x,\, y)_{\mathbb{K}^n}$ für alle $x, y \in \mathbb{K}^n$. Gemäß Korollar 15.2.1 (1.) ist auch $U^{-1} : V \longrightarrow \mathbb{K}^n$ orthogonal bzw. unitär.

Dann ist $B := U^{-1}\,A\,U : \mathbb{K}^n \longrightarrow \mathbb{K}^n$ eine lineare Abbildung in $\mathbb{K}^n$ mit einer Adjungierten $B^*: \mathbb{K}^n \longrightarrow \mathbb{K}^n$. Für alle $v, w \in V$ gilt

$$\begin{aligned}\langle v, Aw\rangle_V =&(U^{-1}v, U^{-1}AUU^{-1}w)_{\mathbb{K}^n} = (U^{-1}v, BU^{-1}w)_{\mathbb{K}^n}\\ =&(B^*U^{-1}v, U^{-1}w)_{\mathbb{K}^n} = \langle UB^*U^{-1}v, w\rangle_V,\end{aligned}$$

d.h. $U\,B^*\,U^{-1}: V \longrightarrow V$ ist die Adjungierte von $A: V \longrightarrow V$. □

Definition 15.3.1. *V sei ein $\mathbb{K}$–Vektorraum mit einem Skalarprodukt $\langle\cdot,\cdot\rangle$. Eine lineare Abbildung $A: V \longrightarrow V$ heißt* selbstadjungiert *(d.h. zu sich selbst adjungiert), wenn $A = A^*$ ist, d.h. für alle $x, y \in V$ gilt:*

$$\langle x,\ A\,y\rangle \ = \ \langle A\,x,\ y\rangle.$$

Bemerkung: Aus dem Beweis von Satz 15.3.1 geht hervor: Die durch eine Matrix $A = (a_{ij}) \in \mathbb{K}^{n,n}$ erzeugte Abbildung $\big(\mathbb{K}^n, (\cdot,\cdot)_{\mathbb{K}^n}\big) \longrightarrow \big(\mathbb{K}^n, (\cdot,\cdot)_{\mathbb{K}^n}\big)$, $x \longmapsto Ax$ ist genau dann selbstadjungiert, wenn alle ihre Komponenten $a_{ij} = \overline{a_{ji}}$ $(i, j \in \{1, \ldots, n\})$ erfüllen (kurz notiert als $A = \overline{A}^t = A^*$).

Definition 15.3.2. *Eine Matrix $(a_{ij}) \in \mathbb{C}^{n,n}$ mit $a_{ij} = \overline{a_{ji}}$ für alle i, j heißt* hermitesch. *Analog heißt $(a_{ij}) \in \mathbb{R}^{n,n}$ mit $a_{ij} = a_{ji}$ für alle i, j* symmetrisch.

Satz 15.3.2. *V sei ein endlichdimensionaler $\mathbb{K}$–Vektorraum mit Skalarprodukt. Dann gilt für jede selbstadjungierte lineare Abbildung $A: V \longrightarrow V$:*

(*a*) *Alle Eigenwerte von A sind reell.*

(*b*) *A hat mindestens einen Eigenwert.*

(*c*) *Eigenvektoren zu verschiedenen Eigenwerten sind orthogonal.*

Beweis: (a) Für jeden Eigenwert $\lambda \in \mathbb{C}$ von A und einen zugehörigen Eigenvektor $x \in V$ gilt

$$\begin{aligned}\overline{\lambda}\,\|x\|^2 \ &= \ \overline{\lambda}\,\langle x,\ x\rangle \ = \ \langle\lambda\,x,\ x\rangle \ = \ \langle A\,x,\ x\rangle \ = \ \langle x,\ A\,x\rangle \ = \ \langle x,\ \lambda\,x\rangle\\ &= \ \lambda\,\langle x,\ x\rangle \ = \ \lambda\,\|x\|^2,\end{aligned}$$

d.h. λ ist reell.

(b) Es gibt eine orthogonale bzw. unitäre lineare Abbildung $U: \mathbb{K}^m \longrightarrow V$, wie bereits im Beweis von Satz 15.3.1 erläutert wurde. Sie besitzt stets eine orthogonale bzw. unitäre Umkehrung $U^{-1}: V \longrightarrow \mathbb{K}^m$.

Dann ist $B := U^{-1}\,A\,U: \mathbb{K}^m \longrightarrow \mathbb{K}^m$ selbstadjungiert bzgl. des Standardskalarprodukts, denn es gilt

$$\begin{aligned}(B\,x,\ y)_{\mathbb{K}^m} \ &= \ (U^{-1}\,A\,U\,x,\ \ y)_{\mathbb{K}^m} &&= \ \langle U\,U^{-1}\,A\,U\,x,\ \ U\,y\rangle_V\\ &= \ \langle A\,U\,x,\ \ U\,y\rangle_V &&= \ \langle U\,x,\ \ A\,U\,y\rangle_V\\ &= \ (U^{-1}\,U\,x,\ \ U^{-1}\,A\,U\,y)_{\mathbb{K}^m} &&= \ (x,\ \ B\,y)_{\mathbb{K}^m}.\end{aligned}$$

Also wird B durch eine hermitesche (bzw. symmetrische) Matrix erzeugt. Im komplexen Fall garantiert Satz 14.4.1 für B mindestens einen Eigenwert, der nach Teil (a) reell ist.

Im reellen Fall erzeugt die reelle Matrix von B eine selbstadjungierte Abbildung $\mathbb{C}^m \longrightarrow \mathbb{C}^m$, die (wie eben gezeigt) mindestens einen Eigenwert $\lambda \in \mathbb{R}$ hat, d.h. es gilt $p_B(\lambda) = 0$ für das charakteristische Polynom von B. Also ist $B - \lambda\,\mathbb{E}_m$ nicht injektiv in $\mathbb{R}^m$, und somit ist λ auch ein Eigenwert von $B : \mathbb{R}^m \longrightarrow \mathbb{R}^m$.

Ist $v \in \mathbb{K}^m$ der zu λ gehörige Eigenvektor von $B = U^{-1}\,A\,U : \mathbb{K}^m \longrightarrow \mathbb{K}^m$, so erhalten wir mit $U\,v \in V$ einen Eigenvektor von A zum Eigenwert λ:

$$A\,(U\,v) \;=\; U\,(U^{-1}\,A\,U)\,v \;=\; U\,B\,v \;=\; U\,(\lambda\,v) \;=\; \lambda U\,v.$$

(c) $\lambda_1 \neq \lambda_2$ seien Eigenwerte von A und $v_1, v_2 \in V \setminus \{0\}$ zugehörige Eigenvektoren. Gemäß Teil (a) sind λ_1, λ_2 reell, und daraus folgt

$$\begin{aligned}(\lambda_2 - \lambda_1)\,\langle v_1,\, v_2\rangle \;&=\; \langle v_1,\, \lambda_2\,v_2\rangle \;-\; \langle \lambda_1\,v_1,\, v_2\rangle\\ &=\; \langle v_1,\, A\,v_2\rangle \;-\; \langle\, A\,v_1,\, v_2\rangle \;=\; 0.\end{aligned}$$

Also sind v_1 und v_2 orthogonal: $\langle v_1, v_2\rangle = 0$. $\square$