PD. Dr. P.H. Lesky
Universität Stuttgart
Wintersemester 2018/19

Vorlesung Mathematik I für Studierende der Fächer inf, swt, msv

Warnung: Dies ist kein vollständiger Vorlesungsaufschrieb. Dieses Skript ist zur Erleichterung beim Mitschreiben gedacht, Ergänzungen sollen nachgetragen werden.

1 Grundlagen

1.1 Zur Aussagenlogik

1.1 Definition: Eine Aussage ist ein sprachlicher Satz, der entweder wahr oder falsch ist, unabhängig von Zeit und Ort, an dem er ausgesprochen wird.

1.2 Verknüpfung von Aussagen:

Verknüpfung	in Worten	Definition durch Wahrheitstabelle
$\neg A$	nicht A, Negation von A	

A	$\neg A$
w	
f	

Verknüpfung	in Worten
$A \vee B$	A oder B, logisches Oder
$A \wedge B$	A und B, logisches Und
$A \to B$	wenn A dann B
$A \leftrightarrow B$	A genau dann wenn B

A	B	$A \vee B$	$A \wedge B$	$A \to B$	$A \leftrightarrow B$
w	w				
w	f				
f	w				
f	f				

1.3 Implikation: $A \Rightarrow B$ bedeutet: Die Aussage $A \to B$ ist wahr.
Also: Wenn A wahr ist, dann ist auch B wahr. (Wenn A falsch ist, wird keine Aussage über B gemacht.)

Man sagt: A ist **hinreichende Bedingung** für B,
B ist **notwendige Bedingung** für A.

Z.B. $|x-4| < 1 \quad \Rightarrow \quad x < 5.$

1.4 Äquivalenz: $A \Leftrightarrow B$ bedeutet: Die Aussage $A \leftrightarrow B$ ist wahr.
Also: Entweder sind beide Aussagen wahr oder beide falsch.

Z.B. $(x-1)^2 \leq 4 \quad \Leftrightarrow \quad |x-1| \leq 2 \quad \Leftrightarrow \quad -2 \leq x-1 \leq 2 \quad \Leftrightarrow \quad -1 \leq x \leq 3.$

Z.B. $A \to B \quad \Leftrightarrow \quad \neg B \to \neg A.$

Beweis durch Wahrheitstafel:
(Beweis durch Fallunterscheidung,
Auflistung aller möglichen Fälle)

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

1.5 Wichtige Beweistechniken:

1) Direkter Beweis: Zeige $A \Rightarrow B$, indem aus der Gültigkeit von A durch Umformungen oder Folgerungen auf die Gültigkeit von B geschlossen wird.

2) Kontraposition: Zeige $A \Rightarrow B$ durch Nachweis von $\neg B \Rightarrow \neg A$.

3) Widerspruchsbeweis: Zeige $A \Rightarrow B$ durch Nachweis von $A \wedge \neg B \Rightarrow$ falsche Aussage (markiert durch $\lightning$ oder „Widerspruch“).

2) und 3) werden als **indirekte Beweise** bezeichnet.

Z.B. Beweise $\underbrace{|x-4| < 1}_{A} \quad \Rightarrow \quad \underbrace{x < 5}_{B}$:

1) Direkt: Wir gehen davon aus, dass $|x-4| < 1$ wahr ist.
Fall a) $x \geq 4 \quad \Rightarrow \quad x-4 = |x-4| < 1 \quad \Rightarrow \quad x < 5 \quad \Leftrightarrow \quad B$
Fall b) $x < 4 \quad \Rightarrow \quad x < 5 \quad \Leftrightarrow \quad B$

2) Kontraposition:

$$\neg B \quad \Leftrightarrow \quad x \geq 5 \quad \Rightarrow \quad x-4 \geq 1 \quad \Rightarrow \quad |x-4| = x-4 \geq 1 \quad \Leftrightarrow \quad \neg A.$$

3) Widerspruchsbeweis: Annahme $|x-4| < 1 \wedge x \geq 5$
$\Rightarrow \quad 1 \leq x-4 = |x-4| < 1 \quad \Rightarrow \quad 1 < 1 \quad \lightning$

Typisches Beispiel für einen Widerspruchsbeweis: $x^2 = 2 \quad \Rightarrow \quad x \notin \mathbb{Q}$.

Annahme: $x^2 = 2 \wedge x \in \mathbb{Q}$
$\Leftrightarrow \quad x^2 = 2 \wedge x = \frac{p}{q}$, wobei p, q teilerfremd
$\vdots$
$\Rightarrow \quad p, q$ haben den gemeinsamen Teiler 2 $\lightning$ (p, q sind teilerfremd)

1.2 Mengen

1.6 Definition (naiv)***:*** Eine **Menge** ist eine Zusammenfassung von Objekten zu einem Ganzen. Bezeichnung: $m \in M$ bedeutet: Das Objekt m liegt in M (ist in der Menge M enthalten), $m \notin M$ bedeutet: m liegt nicht in M.

Explizit: $M := \{1, 2, 3\}$, $\mathbb{N} = \{1, 2, 3, \ldots\}$.
Durch charakteristische Eigenschaft: $P := \{n \in \mathbb{N} : n \text{ ist Primzahl}\}$
$\emptyset$ (oder $\{\}$): Leere Menge, enthält keine Elemente.

1.7 Gleichheit: $A = B$, falls $x \in A \Leftrightarrow x \in B$.

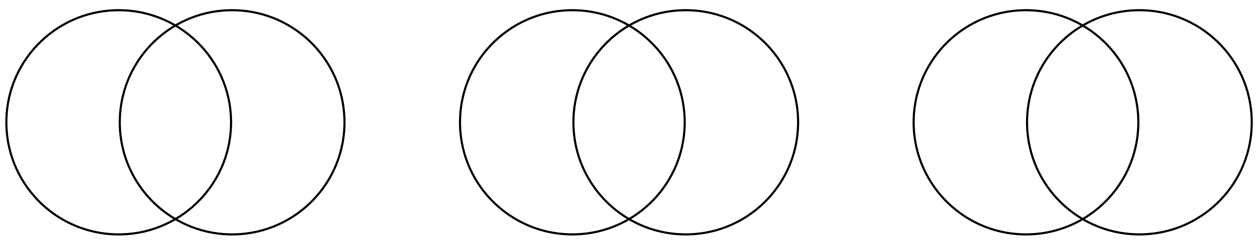
1.8 Teilmenge: $A \subseteq B$, falls: $x \in A \Rightarrow x \in B$.
Für jede Menge A gilt $\emptyset \subseteq A$.

Veranschaulichung im Venn-Diagramm:

1.9 Verknüpfung von Mengen:

$$\begin{aligned} A \cap B &:= \{x : x \in A \wedge x \in B\} \quad \text{(Schnittmenge)} \\ A \cup B &:= \{x : x \in A \vee x \in B\} \quad \text{(Vereinigungsmenge)} \\ A \setminus B &:= \{x : x \in A \wedge x \notin B\} \quad \text{(Differenzmenge)} \end{aligned}$$

Veranschaulichung im **Venn-Diagramm**:



Zwei Mengen heißen **disjunkt**, wenn ihre Schnittmenge leer ist.

1.10 Gesetze von De Morgan: Seien A, B, C Mengen. Dann gelten:

$$A \setminus (B \cup C) = (A \setminus B) \cap (A \setminus C), \qquad A \setminus (B \cap C) = (A \setminus B) \cup (A \setminus C).$$

1.11 Potenzmenge: $\mathcal{P}(A) := \{B : B \subseteq A\}$

Z.B.: $A = \{1, 2, 3\} \quad \Rightarrow \quad \mathcal{P}(A) =$

Z.B.: $A = \big\{\{1\}, \{2\}\big\} \quad \Rightarrow \quad \mathcal{P}(A) =$

1.12 Kartesisches Produkt: $A \times B \;:=\; \{(x, y) : x \in A \wedge y \in B\}$,
(x, y) heißt **geordnetes Paar** (geordnet: Die Reihenfolge ist wichtig)

Z.B.: $\{1, 2\} \times \{3, 4\} \;=$

Z.B.: $\mathbb{N} \times \mathbb{N} = \{\text{Gitterpunkte}\}$:

Geordnetes Paar definiert durch Mengenlehre: $(x, y) := \{x, \{y\}\}$.
Insbesondere: $(x, y) = (x', y') \;\Leftrightarrow\; x = x' \wedge y = y'$.

1.3 Quantoren

1.13 Definition: Sei M eine Menge, $A(x)$ eine von $x \in M$ abhängige Aussage.

$\forall x \in M : A(x)$
bedeutet: Für alle x aus der Menge M ist die Aussage $A(x)$ wahr.
Z.B.: $\forall n \in \mathbb{N} : n \leq n^2$.

$\exists x \in M : A(x)$
bedeutet: Es gibt (existiert) mindestens ein x in der Menge M, für das die Aussage $A(x)$ wahr ist.
Z.B.: $\exists n \in \mathbb{N} : n - 3 \notin \mathbb{N}$.

$\exists! x \in M : A(x)$
bedeutet: Es gibt genau ein x in der Menge M, für das die Aussage $A(x)$ wahr ist.
Z.B.: $\exists! x \in \mathbb{N} : x + 4 = 6$.

1.14 Negation von Aussagen mit Quantoren:
$\neg\big(\forall x \in M : A(x)\big) \;\Leftrightarrow\; \exists x \in M : \neg A(x)$,
$\neg\big(\exists x \in M : A(x)\big) \;\Leftrightarrow\; \forall x \in M : \neg A(x)$.

Z.B.: Lösbarkeit der Gleichung $x + a = b$ im Raum der ganzen Zahlen:

$$\forall a \in \mathbb{Z} \;\; \forall b \in \mathbb{Z} \;\; \exists x \in \mathbb{Z} : x + a = b$$

Negation:

$$\exists a \in \mathbb{Z} \;\; \exists b \in \mathbb{Z} \;\; \forall x \in \mathbb{Z} : x + a \neq b$$

1.4 Abbildungen

1.15 Definition: Seien A, B Mengen. Eine **Abbildung** f von A nach B ist eine Vorschrift, die jedem $x \in A$ ein eindeutig bestimmtes Element $f(x) \in B$ zuordnet. Manchmal heißt f auch **Funktion**.

Schreibe: $f : A \to B$

oder $f : A \to B : x \mapsto y = f(x)$

oder $f : A \ni x \mapsto y \in B$

oder $f : x \mapsto y = f(x)$.

A heißt **Definitionsbereich** (*engl. domain*) von f, B **Bildbereich** (*engl. codomain*) von f.

Falls $y = f(x)$, heißt y **Bild** von x (*engl. image*), x heißt **Urbild** (*engl. pre-image*) von y.

Die Menge

$$f(A) \ := \ \{f(x) \in B : x \in A\}$$

heißt **Bild von f** (*engl. Range*), für $C \subseteq B$ heißt

$$f^{-1}(C) \ := \ \{x \in A : f(x) \in C\}$$

Urbild von C.

Z.B.: $f : \mathbb{R} \to \mathbb{R} : x \mapsto x^2$, $y = 4$, $C = [0, 4]$.
Achtung: $f : \mathbb{R} \to \mathbb{R} : x \mapsto \frac{1}{x}$ ist keine Abbildung.

1.16 Definition: Zwei Abbildungen $f : A \to B$ und $f' : A' \to B'$ heißen **gleich**, falls $A' = A \wedge \forall x \in A : f(x) = f'(x)$.

1.17 Definition: $f : A \to B$ heißt

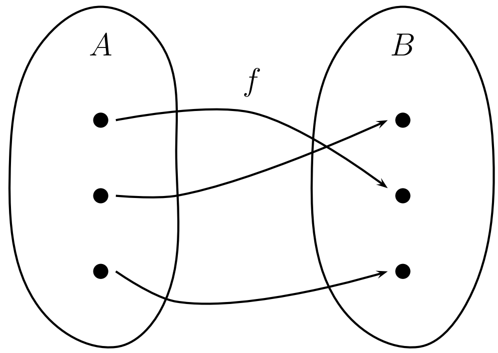
1) **injektiv**, falls für $x, x' \in A$ $\quad f(x) = f(x') \Rightarrow x = x'$
bzw. $\quad x' \neq x \Rightarrow f(x') \neq f(x)$

2) **surjektiv**, falls $f(A) = B$.

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

Injektiv bedeutet: In jedem $y \in B$ endet höchstens ein Pfeil
Surjektiv bedeutet: In jedem $y \in B$ endet mindestens ein Pfeil

Bijektiv:



1.18 Definition: Ist $f : A \to B$ bijektiv, so heißt $f^{-1} : B \to A : f(x) \mapsto x$ die **Umkehrabbildung** oder **inverse Abbildung** zu f.

Z.B.: $f_1 : \mathbb{R} \to \mathbb{R} : x \mapsto x^2$ weder injektiv noch surjektiv
$f_2 :]-\infty, 0] \to \mathbb{R} : x \mapsto x^2$ injektiv, nicht surjektiv
$f_3 : \mathbb{R} \to [0, \infty[: x \mapsto x^2$ nicht injektiv, aber surjektiv
$f_4 :]-\infty, 0] \to [0, \infty[: x \mapsto x^2$ injektiv und surjektiv, $f_4^{-1} : y \mapsto -\sqrt{y}$

1.19 Definition: Ist $f : A \to B$ und $A_0 \subseteq A$, so heißt $f_0 : A_0 \to B : x \mapsto f_0(x) := f(x)$ die **Einschränkung** von f auf A_0; f heißt **Fortsetzung** von f_0 auf A. Man schreibt $f_0 = f\big|_{A_0}$.

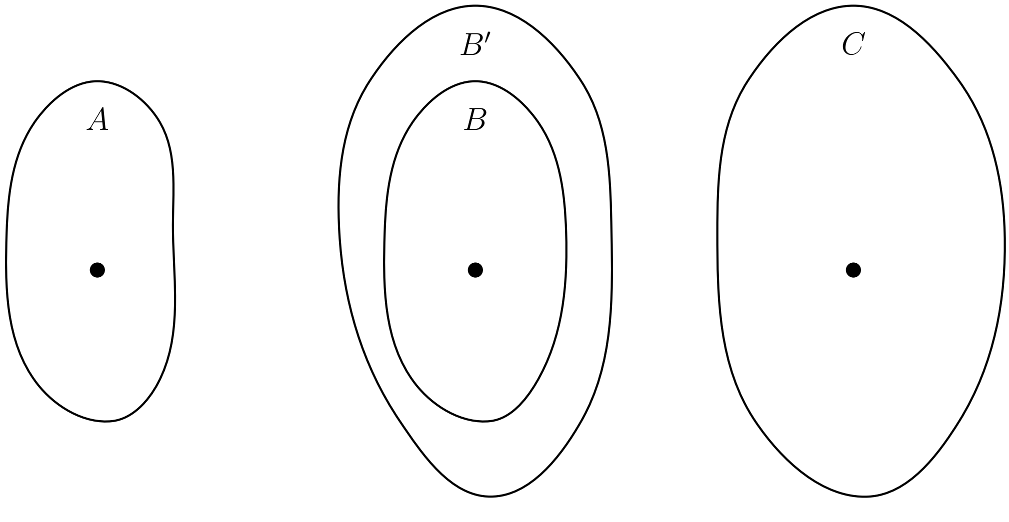
Die Abbildung $\mathrm{id}_A : A \to A : x \mapsto x$ heißt **identische Abbildung**.

Z.B.: Seien f_1, f_2 aus dem vorigen Beispiel. Dann ist f_2 die Einschränkung von f_1 auf $]-\infty, 0]$: $f_2 = f_1\big|_{]-\infty,0]}$. Umgekehrt ist f_1 eine Fortsetzung von f_2 auf $\mathbb{R}$.

1.20 Verknüpfung von Abbildungen: Sei $f : A \to B$, $g : B' \to C$, $B \subseteq B'$. Dann heißt

$$g \circ f : A \to C : x \mapsto (g \circ f)(x) := g(f(x))$$

die **Verknüpfung** oder **Hintereinanderausführung** g nach f.



1.21 Satz: Ist $f : A \to B$ bijektiv, so gilt

$$f \circ f^{-1} = \mathrm{id}_B, \qquad f^{-1} \circ f = \mathrm{id}_A$$

Beweis: **1)** Sei $x \in A$. $\Rightarrow \quad f^{-1}\big(f(x)\big) = x \quad \Rightarrow \quad f^{-1} \circ f = \mathrm{id}_A$.

2) $y \in B \;\Rightarrow\; \exists x \in A : y = f(x)$
$\Rightarrow\; x = f^{-1}(y) \;\Rightarrow\; f\big(f^{-1}(y)\big) = f(x) = y \;\Rightarrow\; f \circ f^{-1} = \mathrm{id}_B$ $\square$

1.22 Abbildungen und Mengenlehre: Sei $f : A \to B$. Dann heißt

$$G := \{(x, f(x)) : x \in A\} \subseteq A \times B$$

der **Graph** von f. Definiere eine Abbildung von A nach B als $G \subseteq A \times B$ mit der Eigenschaft

$$\forall x \in A \;\; \exists! y_x \in B : (x, y_x) \in G.$$

und setze $f(x) := y_x$.

1.5 Relationen

1.23 Definition: Seien A, B Mengen. Eine (zweistellige) **Relation** ist eine Teilmenge $R \subseteq A \times B$. Für $(a, b) \in R$ schreibe aRb: Zwischen a und b besteht die Relation R. R heißt **Relation auf** A, falls $R \subseteq A \times A$.

Z.B.: $\leq \; := \; \{(n, m) \in \mathbb{N} \times \mathbb{N} : n \leq m\}$.
Z.B.: $\mathbb{N}/2\mathbb{N} \; := \; \{(n, m) \in \mathbb{N} \times \mathbb{N} : |n - m| \text{ ist durch 2 teilbar}\}$.

1.24 Definition: Eine Relation auf A heißt

1) **reflexiv**, falls $\forall a \in A : (a, a) \in R$.

2) **symmetrisch**, falls $(a, b) \in R \Leftrightarrow (b, a) \in R$.

3) **antisymmetrisch**, falls $(a, b) \in R \;\wedge\; (b, a) \in R \;\Rightarrow\; a = b$.

4) **transitiv**, falls $(a, b) \in R \;\wedge\; (b, c) \in R \;\Rightarrow\; (a, c) \in R$.

1.25 Ordnungsrelation: Ist eine Relation auf A, die reflexiv, antisymmetrisch und transitiv ist. Für $(a,b) \in R$ schreibe $a \le b$. Es gilt also

$$\begin{aligned} &\forall a \in A : a \le a \\ &a \le b \ \wedge b \le a \ \Rightarrow \ a = b \\ &a \le b \ \wedge b \le c \ \Rightarrow \ a \le c \end{aligned}$$

1.26 Äquivalenzrelation: Ist eine Relation auf A, die reflexiv, symmetrisch und transitiv ist. Für $(a,b) \in R$ schreibe $a \sim b$. Es gilt also

$$\begin{aligned} &\forall a \in A : a \sim a \\ &a \sim b \ \Leftrightarrow \ b \sim a \\ &a \sim b \ \wedge b \sim c \ \Rightarrow \ a \sim c \end{aligned}$$

Z.B.: $\sim\, = \mathbb{N}/2\mathbb{N}$. Dadurch zerfällt $\mathbb{N}$ in zwei Klassen:

$$\begin{array}{ccccccc} \{n \in \mathbb{N} : n \sim 1\} & = & \{1,3,5,\dots\} & = & \{n \in \mathbb{N} : n \sim 3\} & = & \dots \\ \{n \in \mathbb{N} : n \sim 2\} & = & \{2,4,6,\dots\} & = & \{n \in \mathbb{N} : n \sim 4\} & = & \dots \end{array}$$

1.27 Äquivalenzrelation ist Klasseneinteilung: Sei $\sim$ eine Äquivalenzrelation auf A. Für $a \in A$ heißt

$$[a] \ := \ \{b \in A : b \sim a\}$$

die **Äquivalenzklasse** von a. Für Äquivalenzklassen gelten:

1) $\forall a \in A : a \in [a]$,

2) $[a] \cap [b] \ne \emptyset \ \Rightarrow \ [a] = [b]$.

Das bedeutet: Die Menge

$$\mathcal{A} \ := \ \{[a] : a \in A\} \ \subseteq \ \mathcal{P}(A)$$

besitzt die Eigenschaften

1) $\displaystyle\bigcup_{B \in \mathcal{A}} B \ = \ A$,

2) Sind $B_1, B_2 \in \mathcal{A}$, so gilt entweder $B_1 \cap B_2 = \emptyset$ oder $B_1 = B_2$.

Man nennt $\mathcal{A}$ eine **Klasseneinteilung** von A.

Beweis der Eigenschaften 1) und 2) der Äquivalenzklassen:
Für jedes $a \in A$ gilt $a \sim a \quad \Rightarrow \quad a \in [a]$.
Sei $[a] \cap [b] \ne \emptyset \quad \Rightarrow \quad \exists c \in [a] \cap [b] \quad \Rightarrow \quad a \sim c \sim b \quad \Rightarrow \quad a \sim b \quad \Rightarrow \quad [a] \subseteq [b] \subseteq [a]$.

Umgekehrt: Ist $\mathcal{A}$ eine Klasseneinteilung von A, so definiert

$$R \ := \ \{(a,b) \in A \times A \bigm| \exists B \in \mathcal{A} : a \in B \ \wedge \ b \in B\}$$

eine Äquivalenzrelation, und die Äquivalenzklassen sind genau die Elemente von $\mathcal{A}$.

Beweis, dass R eine Äquivalenzrelation ist:

Reflexivität: Sei $a \in A$. Wegen $A = \bigcup_{B \in \mathcal{A}} B$ existiert ein $B \in \mathcal{A}$ mit $a \in B \ \Rightarrow \ a\,R\,a$,

Symmetrie: $a\,R\,b \ \Rightarrow \ b\,R\,a$ klar,

Transitivität: $\underbrace{a\,R\,b}_{a,b \in B_1} \wedge \underbrace{b\,R\,c}_{b,c \in B_2} \ \Rightarrow \ B_1 = B_2 \ \Rightarrow \ a, c \in B_1 \ \Rightarrow \ a\,R\,c.$

1.28 Beispiele: **1)** $A = \mathbb{R} \times \mathbb{R}$, $(x,y) \sim (x',y') \ :\Leftrightarrow \ x^2 + y^2 = x'^2 + y'^2$
$[(x,y)] =$ Kreis um $(0,0)$ mit Radius $r = \sqrt{x^2 + y^2}$.

2) Brüche: Zwei Brüche $\frac{p}{q}, \frac{r}{s}$ sind gleich, wenn $p\,s = r\,q$. Definiere auf $\mathbb{Z} \times \mathbb{N}$ die Relation

$$(p,q) \sim (r,s) \ :\Leftrightarrow \ p\,s = r\,q$$

Ein Bruch ist dann die Äquivalenzklasse $[(p,q)] =: \frac{p}{q}$.

3) Größen von endlichen Mengen: Sei $\mathcal{M} := \{$Mengen M mit endlich vielen Elementen$\}$. Definiere auf $\mathcal{M}$

$$M_1 \sim M_2 \ :\Leftrightarrow \ \exists (f : M_1 \to M_2) : f \text{ ist bijektiv}$$

Die Äquivalenzklassen bestehen aus allen Mengen mit derselben Anzahl von Elementen.

1.6 Die natürlichen Zahlen

1.29 Peano (1889): Die natürlichen Zahlen bilden eine Menge $\mathbb{N}$, auf der eine Abbildung

$$\text{Nachfolger: } \mathbb{N} \to \mathbb{N}$$

erklärt ist mit folgenden Eigenschaften:

($\mathbb{N}$1) $\exists!\, x_0 \in \mathbb{N} : x_0 \notin \text{Nachfolger}(\mathbb{N})$. Bezeichnung $1 := x_0$.
(Es existiert genau eine Zahl, die nicht Nachfolger einer anderen Zahl ist).

($\mathbb{N}$2) Nachfolger ist injektiv
$(\text{Nachfolger}(n_1) = \text{Nachfolger}(n_2) \ \Rightarrow \ n_1 = n_2)$

($\mathbb{N}$3) (Induktionsaxiom) Ist $M \subseteq \mathbb{N}$ mit den beiden Eigenschaften

- $1 \in M$ und
- $n \in M \;\Rightarrow$ Nachfolger$(n) \in M$,

dann gilt $M = \mathbb{N}$.
(Enthält eine Teilmenge $M \subseteq \mathbb{N}$ die 1 und mit jeder Zahl auch ihren Nachfolger, dann ist $M = \mathbb{N}$).

Man schreibt: $2 :=$ Nachfolger(1), $3 :=$ Nachfolger(2), $\ldots$, $n+1 :=$ Nachfolger(n).

Aus diesen Axiomen lassen sich alle bekannten Rechenoperationen und Rechenregeln für die natürlichen Zahlen beweisen, z.B. $m+n = n+m$.

1.30 Vollständige Induktion: Für $n \in \mathbb{N}$ sei $A(n)$ eine Aussage, die von n abhängt, z.B.

$$1+2+\ldots+n \;=\; \frac{n(n+1)}{2},$$

Ziel: Beweise, dass $A(n)$ für alle $n \in \mathbb{N}$ wahr ist.

Verfahren: Beweise

1) $A(1)$ ist wahr (Induktionsanfang),

2) $A(n) \Rightarrow A(n+1)$ für beliebiges $n \in \mathbb{N}$ (Induktionsschritt)
(Falls $A(n)$ wahr („Induktionsannahme“), dann ist $A(n+1)$ wahr („Induktionsbehauptung“).

Induktionsschluss: Dann ist $A(n)$ wahr für alle $n \in \mathbb{N}$.

Beweis: Sei $M := \{n \in \mathbb{N} : A(n) \text{ ist wahr}\} \subseteq \mathbb{N}$. Dann $1 \in M$ und $n \in M \;\Rightarrow\; n+1 \in M$. Also $M = \mathbb{N}$ nach ($\mathbb{N}$3). $\square$

1.31 Beispiele: **1)** $A(n) : 1+2+\ldots+n = \frac{n(n+1)}{2}$ (**arithmetische Summe**).

- Induktionsanfang: $A(1): \quad 1 = \frac{1(1+1)}{2}$ ist wahr.
- Induktionsschritt: Falls $1+2+\ldots+n = \frac{n(n+1)}{2}$ ($A(n)$ wahr), dann

$$\begin{aligned} 1+2+\ldots+n+n+1 &= \frac{n(n+1)}{2}+n+1 = \frac{n(n+1)+2(n+1)}{2} \\ &= \frac{(n+2)(n+1)}{2}. \end{aligned}$$

  Also: $A(n+1)$ wahr.
- Induktionsschluss: $\forall n \in \mathbb{N} : 1+2+\ldots+n = \frac{n(n+1)}{2}$.

2) Sei $q \neq 1$. $A(n): \underbrace{q^0}_{:=1} + q^1 + \ldots + q^n = \frac{1-q^{n+1}}{1-q}$ (**geometrische Summe**).

3) $A(n)$: Ist $M \subseteq \mathbb{N}$ eine Menge mit n Elementen, so besitzt M ein Maximum:

$$k = \max(M) \quad :\Leftrightarrow \quad k \in M \ \wedge \ \forall l \in M : l \leq k.$$

4) Ist M eine Menge mit n Elementen, so besitzt die Potenzmenge $\mathcal{P}(M)$ 2^n Elemente.

1.32 Rekursive Definition: **1)** Fakultät $n!$:

$$\begin{aligned} 0! &:= 1 \\ (n+1)! &:= (n+1)\, n! \text{ für } n \in \mathbb{N}_0. \end{aligned}$$

2) Summen- und Produktzeichen:

$$\sum_{k=1}^{0} a_k := 0, \quad \sum_{k=1}^{n+1} a_k := \sum_{k=1}^{n} a_k + a_{n+1}, \qquad \text{salopp:} \quad \sum_{k=1}^{n} a_k = a_1 + a_2 + \ldots + a_n$$

$$\prod_{k=1}^{0} a_k := 1, \quad \prod_{k=1}^{n+1} a_k := \left(\prod_{k=1}^{n} a_k\right) \cdot a_{n+1}, \qquad \text{salopp:} \quad \prod_{k=1}^{n} a_k = a_1 \cdot a_2 \ldots a_n$$

3) Addition und Multiplikation aus Peano-Axiomen: Für $m \in \mathbb{N}$ setze

$$m + 1 := \text{Nachfolger}(m), \qquad m + (n+1) := \text{Nachfolger}(m+n).$$

Damit ist $m + n$ für alle $m, n \in \mathbb{N}$ definiert.
Analog: $m \cdot 1 := m,\ m \cdot (n+1) := m \cdot n + m$. Damit ist $m \cdot n$ für alle $n, m \in \mathbb{N}$ definiert.
Jetzt können alle Rechengesetze bewiesen werden.

1.33 Binomialkoeffizient:

$$\binom{n}{k} := \frac{n!}{(n-k)! \cdot k!} = \frac{n \cdot (n-1) \cdots (n-k+1)}{1 \cdot 2 \cdots k} \qquad \text{(sprich: „n über k“)}$$

für $n, k \in \mathbb{N}_0,\ k \leq n$. Dann gelten

1) $\binom{n}{k} = \binom{n}{n-k}$

2) $\binom{n}{k} + \binom{n}{k+1} = \binom{n+1}{k+1}$ für $k \leq n-1$.

3) $\binom{n}{0} = \binom{n}{n} = 1$

4) Pascalsches Dreieck:

$$\begin{array}{ccccccccc}
 & & & & \binom{0}{0} & & & & \\
 & & & \binom{1}{0} & & \binom{1}{1} & & & \\
 & & \binom{2}{0} & & \binom{2}{1} & & \binom{2}{2} & & \\
 & \binom{3}{0} & & \binom{3}{1} & & \binom{3}{2} & & \binom{3}{3} & \\
\binom{4}{0} & & \binom{4}{1} & & \binom{4}{2} & & \binom{4}{3} & & \binom{4}{4}
\end{array}
\qquad\qquad
\begin{array}{ccccccccc}
 & & & & 1 & & & & \\
 & & & 1 & & 1 & & & \\
 & & 1 & & 2 & & 1 & & \\
 & 1 & & 3 & & 3 & & 1 & \\
1 & & 4 & & 6 & & 4 & & 1
\end{array}$$

1.34 Binomischer Satz: Für beliebige Zahlen a, b und für $n \in \mathbb{N}$ gilt

$$(a+b)^n \;=\; \sum_{k=0}^{n} \binom{n}{k} a^{n-k}\, b^k \;=\; \binom{n}{0} a^n + \binom{n}{1} a^{n-1}\, b + \binom{n}{2} a^{n-2}\, b^2 + \ldots + \binom{n}{n} b^n.$$

Z.B.: $(a+b)^4 \;=\; a^4 + 4a^3\, b + 6a^2\, b^2 + 4a\, b^3 + b^4.$

Beweis: Vollständige Induktion ab $n = 1$:

Induktionsanfang: $\displaystyle\sum_{k=0}^{1} \binom{1}{k} a^{1-k}\, b^k \;=\; \binom{1}{0} a^1\, b^0 + \binom{1}{1} a^0\, b^1 \;=\; a+b \;=\; (a+b)^1$ stimmt.

Induktionsschritt: Induktionsannahme: Die Formel gilt für n.

Induktionsbehauptung: $\displaystyle (a+b)^{n+1} \;=\; \sum_{k=0}^{n+1} \binom{n+1}{k} a^{n+1-k}\, b^k.$

Beweis der Induktionsbehauptung:

$$\begin{aligned}
(a+b)^{n+1} &= (a+b)\cdot(a+b)^n \\
&= (a+b)\cdot \sum_{k=0}^{n} \binom{n}{k} a^{n-k}\, b^k \\
&= \sum_{k=0}^{n} \binom{n}{k} a^{n+1-k}\, b^k + \underbrace{\sum_{k=0}^{n} \binom{n}{k} a^{n-k}\, b^{k+1}}_{=\sum_{l=1}^{n+1} \binom{n}{l-1} a^{n-(l-1)}\, b^l} \\
&= \sum_{k=0}^{n} \binom{n}{k} a^{n+1-k}\, b^k + \sum_{k=1}^{n+1} \binom{n}{k-1} a^{n-(k-1)}\, b^k \\
&= \underbrace{\binom{n}{0}}_{=\binom{n+1}{0}} a^{n+1}\, b^0 + \sum_{k=1}^{n} \underbrace{\left(\binom{n}{k} + \binom{n}{k-1}\right)}_{=\binom{n+1}{k}} a^{n+1-k}\, b^k + \underbrace{\binom{n}{n}}_{\binom{n+1}{n+1}} a^0\, b^{n+1} \\
&= \sum_{k=0}^{n+1} \binom{n+1}{k} a^{n+1-k}\, b^k
\end{aligned}$$

Induktionsschluss: Die Formel gilt für alle $n \in \mathbb{N}$.

$\square$

1.35 Von $\mathbb{N}$ ***zu*** $\mathbb{Z}$***:*** Idee: $-1 = 1-2 = 2-3 = \ldots = m-(m+1) = \ldots$
Definiere auf $\mathbb{N}\times\mathbb{N}$ die Äquivalenzrelation

$$(m,n) \sim (m',n') \;:\Leftrightarrow\; m+n' = m'+n$$

1) Reflexivität: $(m,n)\sim(m,n)$.

2) Symmetrie: $(m,n)\sim(m',n') \;\Leftrightarrow\; (m',n')\sim(m,n)$

3) Transitivität: $(m,n)\sim(m',n') \;\wedge\; (m',n')\sim(m'',n'') \;\Rightarrow\; (m,n)\sim(m'',n'')$.

Sei $\mathbb{Z} := \{[(m,n)] : m,n\in\mathbb{N}\}$. Setze $0 = [(1,1)],\; n := [(n+1,1)],\; -n := [(1,n+1)]$.
Definiere $[(m,n)] + [(k,l)] := [(m+k, n+l)]$ und zeige, dass die Summe unabhängig von den Repräsentanten der Äquivalenzklassen ist.
$(\mathbb{Z},+)$ erfüllt alle bekannten Gesetze.

1.7 Teilbarkeit

1.36 Definition: Sei $n\in\mathbb{Z}$. Dann heißt $m\in\mathbb{N}$ **Teiler** von n (kurz $m \mid n$), falls

$$\exists k\in\mathbb{Z} : n = k\cdot m;$$

n heißt **teilbar** durch m. Insbesondere: 0 ist durch alle $m\in\mathbb{N}$ teilbar.

$$D(n) \;:=\; \{d\in\mathbb{N} : d \mid n\}$$

ist die Menge der Teiler von n. Für $n,m\in\mathbb{N}$ heißt

$$\mathrm{ggT}(m,n) \;:=\; \max\big(D(m)\cap D(n)\big)$$

der **größte gemeinsame Teiler** von m und n.

Z.B.: $\mathrm{ggT}(30,24)$:

1.37 Satz: Teiler ist Ordnungsrelation auf $\mathbb{N}$.

1.38 Hilfssatz: Seien $n,m,l,k\in\mathbb{Z},\; d\in\mathbb{N}$. Dann gilt

$$d \mid n \;\wedge\; d \mid m \;\Rightarrow\; d \mid (k\cdot n + l\cdot m).$$

Beweis: $d \mid n \ \wedge \ d \mid m \ \Rightarrow \ \exists k_1, k_2 \in \mathbb{Z} : n = k_1 \cdot d \ \wedge \ m = k_2 \cdot d.$
$\Rightarrow \ k \cdot n + l \cdot m = k \cdot k_1 \cdot d + l \cdot k_2 \cdot d = (k \cdot k_1 + l \cdot k_2) \cdot d$
$\Rightarrow d \mid (k \cdot n + l \cdot m).$ $\square$

1.39 Teilen mit Rest: Seien $n, m \in \mathbb{N}$ mit $n > m$. Dann gibt es eindeutig bestimmte Zahlen $q \in \mathbb{N}, r \in \mathbb{N}_0 := \mathbb{N} \cup \{0\}$ mit

$$n = q \cdot m + r \quad \wedge \quad r \leq m - 1.$$

Beweis: **1)** $q := \max\{k \in \mathbb{N} : k \cdot m \leq n\}, \ r := n - q \cdot m \in \mathbb{N}_0$
$\Rightarrow \ r \geq 0, \ r \leq m - 1, \ n = q \cdot m + r$
$\Rightarrow$ Existenz

2) Sei $q \cdot m + r = q' \cdot m + r'$

$$\Rightarrow \begin{cases} m \cdot (q - q') = r' - r & \text{falls } r' \geq r \\ m \cdot (q' - q) = r - r' & \text{falls } r' < r \end{cases}$$

$$\overset{|r'-r| \leq m-1}{\Rightarrow} \quad q = q' \ \wedge \ r = r'$$

$\Rightarrow$ Eindeutigkeit $\square$

1.40 Teilen mit Rest erhält den ggT: Ist $n = q \cdot m + r$ wie oben, so gilt

$$\mathrm{ggT}(n, m) = \mathrm{ggT}(m, r).$$

Beweis: Wir zeigen: $D(n) \cap D(m) = D(m) \cap D(r)$. Dann $\mathrm{ggT}(n, m) = \mathrm{ggT}(m.r)$.

$$\begin{aligned} \textbf{1)} \quad d \in D(n) \cap D(m) \ &\Rightarrow \ n = k_1 \cdot d \ \wedge \ m = k_2 \cdot d \\ &\Rightarrow \ r = n - q \cdot m = d \underbrace{(k_1 - q \cdot k_2)}_{\in \mathbb{Z}} \\ &\Rightarrow \ d \mid r \ \wedge \ d \mid m \\ &\Rightarrow \ d \in D(m) \cap D(r) \end{aligned}$$

2) $d \in D(m) \cap D(r) \overset{1.38}{\Rightarrow} d \mid n = q \cdot m + r \ \Rightarrow \ d \in D(n) \cap D(m)$

$\square$

Z.B.: $30 = 1 \cdot 24 + \underbrace{6}_{r = \mathrm{ggT}(30,24)}$

1.41 Euklidischer Algorithmus: Seien $m, n \in \mathbb{N}$ mit $m > n$. Dann existieren eindeutig bestimmte Zahlen $K \in \mathbb{N},\ r_1, \ldots, r_K, q_1, \ldots, q_{K+1} \in \mathbb{N}$ mit

$$\begin{aligned}
m &= q_1 \cdot n + r_1 \quad \text{mit } 0 < r_1 \le n-1 \\
n &= q_2 \cdot r_1 + r_2 \quad \text{mit } 0 < r_2 \le r_1 - 1 \\
r_1 &= q_3 \cdot r_2 + r_3 \quad \ldots \\
\vdots \quad & \quad \vdots \\
r_{K-2} &= q_K \cdot r_{K-1} + r_K \quad \ldots \\
r_{K-1} &= q_{K+1} \cdot r_K + 0,
\end{aligned}$$

und es gilt $r_K = \mathrm{ggT}(m, n)$.

Beweis: Aus Satz 1.39 folgt die Eindeutigkeit der q_j, r_j. Wegen $r_j \le r_{j-1} - 1 \le r_{j-2} - 2 \ldots$ bricht das Verfahren nach höchstens n Schritten ab.
Aus Satz 1.40 folgt, dass

$$\mathrm{ggT}(m,n) = \mathrm{ggT}(n, r_1) = \mathrm{ggT}(r_1, r_2) = \ldots = \mathrm{ggT}(r_{K-1}, r_K) = r_K.$$

□

1.42 Beispiele: **1)** $\mathrm{ggT}(210, 25)$:

2) $\mathrm{ggT}(132, 11)$:

Drei Folgerungen aus dem Euklidischen Algorithmus:

1.43 Hilfssatz: Seien $k, m, n \in \mathbb{N}$. Dann gilt

$$\mathrm{ggT}(km, kn) = k\,\mathrm{ggT}(m, n).$$

Beweis: Multipliziere den euklidischen Algorithmus für m, n mit k:

$$\begin{array}{rcl} m &=& q_1\, n + r_1 \\ n &=& q_2\, r_1 + r_2 \\ & \vdots & \\ r_{K-1} &=& q_{K+1}\, r_K + 0 \end{array} \quad \Rightarrow \quad \begin{array}{rcl} km &=& q_1\, kn + kr_1 \\ kn &=& q_2\, kr_1 + kr_2 \\ & \vdots & \\ k\, r_{K-1} &=& q_{K+1}\, kr_K + 0 \end{array}$$

$\Rightarrow\ \mathrm{ggT}(km, kn) = kr_K = k\,\mathrm{ggT}(m, n).$

□

1.44 Folgerung: $k \mid m \;\wedge\; k \mid n \;\Rightarrow\; k \mid \mathrm{ggT}(m,n)$.

Beweis: $m = l_1\,k \;\wedge\; n = l_2\,k \;\Rightarrow\; \mathrm{ggT}(m,n) = \mathrm{ggT}(k\,l_1, k\,l_2) = k\,\mathrm{ggT}(l_1,l_2)$. $\square$

1.45 Satz: Seien $m,n \in \mathbb{N}$ und $M(m,n) := \{mx + ny : x,y \in \mathbb{Z}\}$. Dann besteht $M(m,n)$ genau aus allen Vielfachen von $\mathrm{ggT}(m,n)$:

$$M(m,n) \;=\; \{z\,\mathrm{ggT}(m,n) : z \in \mathbb{Z}\} \;=:\; \mathrm{ggT}(m,n)\cdot\mathbb{Z}.$$

Beweis: **1)** $\mathrm{ggT}(m,n) \mid m \;\wedge\; \mathrm{ggT}(m,n) \mid n \;\Rightarrow\; \mathrm{ggT}(m,n) \mid (mx+ny) \;\Rightarrow\; M(m,n) \subseteq \mathrm{ggT}(m,n)\cdot\mathbb{Z}$.

2) Zeige $\mathrm{ggT}(m,n) \in M(m,n)$. Dann folgt $\mathrm{ggT}(m,n)\cdot\mathbb{Z} \subseteq M(m,n)$.
Für $l_1, l_2 \in \mathbb{Z}$ gilt: $k_1, k_2 \in M(m,n) \;\Rightarrow\; l_1\,k_1 + l_2\,k_2 \in M(m,n)$, denn:

$$\left.\begin{aligned} k_1 &= m\,x_1 + n\,y_1 \\ k_2 &= m\,x_2 + n\,y_2 \end{aligned}\right\} \;\Rightarrow\; l_1\,k_1 + l_2\,k_2 = m(l_1\,x_1 + l_2\,x_2) + n(l_1\,y_1 + l_2\,y_2) \in M(m,n)$$

Betrachte den euklidischen Algorithmus:

$$\begin{aligned} m &= q_1\,n + r_1 && \Rightarrow\; r_1 = m - q_1\,n \in M(m,n) \\ n &= q_2\,r_1 + r_2 && \Rightarrow\; r_2 = n - q_2\,r_1 \in M(m,n) \\ &\;\vdots \\ r_{K-2} &= q_K\,r_{K-1} + r_K && \Rightarrow\; \mathrm{ggT}(m,n) = r_K \in M(m,n) \end{aligned}$$

$\square$

1.8 Primzahlen

1.46 Definition: $p \in \mathbb{N}$ mit $p \geq 2$ heißt Primzahl, wenn $D(p) = \{1,p\}$. Insbesondere ist 1 keine Primzahl.

1.47 Euklidischer Hilfssatz: Seien $m,n \in \mathbb{N}$, p Primzahl. Dann:

$$p \mid (m\cdot n) \;\Rightarrow\; p \mid m \;\vee\; p \mid n.$$

Beweis: Fall $p \mid m$: okay
Fall $\neg(p \mid m)$: $\;p \mid mn \;\wedge\; p \mid pn \;\Rightarrow\; p \mid \mathrm{ggT}(mn,pn) = n\,\mathrm{ggT}(m,p) = n \;\Rightarrow\; p \mid n$. $\square$

Mit Induktion folgt: p Primzahl $\wedge\; p \mid n_1 \cdots n_k \;\Rightarrow\; \exists j : p \mid n_j$.

1.48 Fundamentalsatz der Arithmetik: Sei $n \in \mathbb{N}, n \geq 2$. Dann lässt sich n als Produkt von Primzahlen darstellen:

$$n \;=\; p_1 \cdot p_2 \cdots p_k, \qquad p_1, \ldots, p_k \text{ Primzahlen.}$$

Die Darstellung ist bis auf die Reihenfolge eindeutig.

Beweis: Wir beweisen: Für jedes $n \in \mathbb{N}$ gilt: Jede natürliche Zahl $j \in \{2, \ldots, n\}$ ist als Produkt von Primzahlen darstellbar.

1) Induktionsanfang $n = 2$: $\quad 2 = 2$ okay

2) Induktionsschritt:

$$n+1 \;=\; \begin{cases} \text{Primzahl } \Rightarrow \; n+1 = n+1, \text{ also Produkt aus der Primzahl } n+1 \\ \text{keine Primzahl: } n+1 = k \cdot l \\ \quad \overset{\text{Ind.Ann}}{\Rightarrow} \quad n+1 = \underbrace{p_1 \cdots p_j}_{=k} \cdot \underbrace{\tilde{p}_1 \cdots \tilde{p}_i}_{=l} = \text{ Produkt von Primzahlen} \end{cases}$$

3) Induktionsschluss: Die Aussage gilt für jedes $n \in \mathbb{N}$ mit $n \geq 2$.

Aus der bewiesenen Aussage folgt, dass sich jede Zahl $j \in \mathbb{N}$ mit $j \geq 2$ als Produkt von Primzahlen darstellen lässt. $\square$

1.49 Satz: Es gibt unendlich viele Primzahlen.

Beweis: Annahme: Es gibt nur endlich viele Primzahlen: $P = \{p_1, \ldots, p_n\}$.
Betrachte

$$m \;:=\; 1 + p_1 \cdot p_2 \cdots p_n \overset{1.48}{=} q_1 \cdots q_k, \quad q_j \text{ Primzahl}$$

Dann folgt $q_1 \notin P$: Andernfalls $q_1 = p_j \;\Rightarrow\; q_1 \mid 1 = q_1 \cdots q_k - p_1 \cdots p_n$ ↯
$\Rightarrow\; q_1$ Primzahl $\wedge\; q_1 \notin P$ ↯ $\square$

1.50 Bemerkung: Primzahlsatz (1896 Hadamard, Vallée Poussin)

$$\pi(n) \;:=\; \text{Anzahl der Primzahlen} \leq n \quad \Rightarrow \quad \lim_{n\to\infty} \frac{\pi(n)}{\frac{n}{\ln n}} \;=\; 1.$$

1.9 Kongruenzen

1.51 Definition: Sei $m \in \mathbb{N}$, $m \geq 2$. $a, b \in \mathbb{Z}$ heißen **kongruent modulo m** (geschrieben $a \equiv b \pmod{m}$ oder $a \equiv b\ (m)$), wenn $m \mid (a-b)$. D.h. a, b ergeben beim Teilen mit Rest durch m denselben Rest: Ist

$$\begin{aligned} a &= q_1 m + r_1, \quad 0 \leq r_1 \leq m-1 \\ b &= q_2 m + r_2, \quad 0 \leq r_2 \leq m-1 \end{aligned}$$

so sind $a \equiv b \pmod{m}$ und $r_1 = r_2$ äquivalent.

1.52 Satz: Sei $m \in \mathbb{N}$, $m \geq 2$ fest. Dann ist $\equiv\ (m)$ eine Äquivalenzrelation auf $\mathbb{Z}$.

Beweis: **1)** $a \equiv a \pmod{m}$ ist erfüllt,

2) $a \equiv b \pmod{m} \;\Rightarrow\; b \equiv a \pmod{m}$ ist erfüllt,

3) $a \equiv b \pmod{m} \;\wedge\; b \equiv c \pmod{m}$
$\Leftrightarrow\; m \mid (a-b) \;\wedge\; m \mid (b-c)$
$\Rightarrow\; m \mid (a-c) = (a-b) + (b-c)$
$\Rightarrow\; a \equiv c \pmod{m}$. $\square$

1.53 Satz: $+$ und $\cdot$ sind verträglich mit $\equiv\ (m)$, d.h.

$$a \equiv a'\ (m) \;\wedge\; b \equiv b'\ (m) \;\Rightarrow\; a + b \equiv a' + b'\ (m) \;\wedge\; a \cdot b \equiv a' \cdot b'\ (m).$$

1.54 Rechnen mit Äquivalenzklassen: Sei $m \in \mathbb{N}$, $m \geq 2$ fest. Dann besitzt die Äquivalenzrelation $\equiv\ (m)$ genau m Äquivalenzklassen $[0], [1], \ldots, [m-1]$, die **Restklassen modulo m**. Definiere

$$\begin{aligned} \mathbb{Z}/m\mathbb{Z} &:= \{[0], \ldots, [m-1]\} \\ [a] + [b] &:= [a+b], \qquad [a] \cdot [b] \;:=\; [a \cdot b] \end{aligned}$$

Dann gelten:

$$\begin{aligned} [a] + ([b] + [c]) &= ([a] + [b]) + [c] && \text{(Assoziativgesetz)} \\ [a] + [0] &= [a] && \text{(Neutrales Element)} \\ [a] + [-a] &= [0] && \text{(Inverses Element)} \\ [a] + [b] &= [b] + [a] && \text{(Kommutativgesetz)} \end{aligned}$$

D.h. $(Z/m\mathbb{Z}, +)$ bildet eine kommutative (abelsche) Gruppe (siehe später).

Z.B: Multiplikation in $\mathbb{Z}/4\mathbb{Z}$:

$\cdot$	[0]	[1]	[2]	[3]
[0]	[0]	[0]	[0]	[0]
[1]	[0]	[1]	[2]	[3]
[2]	[0]	[2]	[0]	[2]
[3]	[0]	[3]	[2]	[1]

Hier ist einiges ungewöhnlich:
$[3] \cdot [3] = [1]$, d.h. $[3] = \frac{[1]}{[3]}$
$[2] \cdot [2] = [0]$, obwohl $[2] \neq [0]$. Man nennt [2] einen Nullteiler.
Die Gleichung $[2] \cdot x = [2]$ hat zwei Lösungen $x = [1]$ und $x = [3]$.
Es gibt keine Zahl $[n]$, so dass $[2] \cdot [n] = [1]$.

1.55 Satz: Sei $m \in \mathbb{N}$, $m \geq 2$ fest. Dann:

1) $[a]$ besitzt ein inverses Element bezüglich $\cdot$ (d.h. eine Restklasse $[b]$ mit $[a] \cdot [b] = [1]$),
$\Leftrightarrow \quad \mathrm{ggT}(a, m) = 1$.

2) m ist Primzahl $\Leftrightarrow \forall a \in \{1, \ldots, m-1\} \;\; \exists b \in \{1, \ldots, m-1\} : [a] \cdot [b] = [1]$

Beweis: **1)**

$$\begin{aligned} [a] \cdot [b] = [1] &\Leftrightarrow [ab] = [1] \\ &\Leftrightarrow m \mid (ab - 1) \\ &\Leftrightarrow \exists y \in \mathbb{Z} : \underbrace{my = ab - 1}_{1 = ab - my} \\ &\Leftrightarrow 1 \in \{ax + my : x, y \in \mathbb{Z}\} = \mathrm{ggT}(a, m) \cdot \mathbb{Z} \text{ (vgl. Satz 1.45)} \\ &\Leftrightarrow \mathrm{ggT}(a, m) = 1 \end{aligned}$$

2) "$\Rightarrow$": Aus 1)
"$\Leftarrow$": Zeige: m keine Primzahl $\Rightarrow \exists k \in \mathbb{N} : [k]$ besitzt kein inverses Element.

Sei m keine Primzahl $\Rightarrow \exists k \in D(m) : k \neq 1 \;\wedge\; k \neq m$.
Behauptung: $[k]$ besitzt kein inverses Element.
Für $l \in \mathbb{Z}$ gilt $[k] \cdot [l] = [kl] = \{kl + ym : y \in \mathbb{Z}\}$
Nun gilt: $k \mid (kl + ym) \;\Rightarrow\; \forall y \in \mathbb{Z} : kl + ym \neq 1$.
$\Rightarrow\; 1 \notin [kl] \;\Rightarrow\; [kl] \neq [1]$.

$\square$

1.10 Darstellung natürlicher Zahlen

Zehnersystem: Stelle alle Zahlen mit den zehn Ziffern $0, 1, \ldots, 9$ dar:

$$\begin{array}{rrrcr l}
 & 1, & 2, & \ldots & 9 & \\
10, & 11, & 12, & \ldots & 19 & \text{z.B. } 13 = 1 \cdot 10 + 3 \cdot 1 \\
20, & 21, & 22, & \ldots & 29 & \\
\vdots & & & & & \\
100, & 101, & 102, & \ldots & 109 & \text{z.B. } 108 = 1 \cdot 10^2 + 0 \cdot 10^1 + 8 \cdot 10^0
\end{array}$$

Entsprechend geht das auch mit weniger oder mehr Ziffern:

1.56 Definition: Gegeben seien: die **Ziffernbasis** $g \in \mathbb{N},\ g \geq 2$,
die **Ziffern**: Menge mit g Elementen $Z = \{z_0, \ldots, z_{g-1}\}$.

Die Darstellung

$$n \;=\; (a_N a_{N-1} \ldots a_0)_g \;:=\; \sum_{j=0}^{N} a_j g^j,$$

wobei $N \in \mathbb{N},\ a_0, \ldots, a_N \in Z$, heißt ***g*-adische Entwicklung** von $n \in \mathbb{N}$.

Z.B: Zweier-System: $g = 2,\ Z = \{0, 1\}$:
$1011_2 = (1 \cdot 2^3 + 0 \cdot 2^2 + 1 \cdot 2^1 + 1 \cdot 2^0)_{10} = (8 + 0 + 2 + 1)_{10} = 11_{10}$.

Z.B: Hexadezimalsystem: $g = 16,\ z = \{0, 1, \ldots, 9, A, B, C, D, E, F\}$:
$D2A_{16} = D\,g^2 + 2\,g^1 + A\,g^0 = (13 \cdot 256 + 2 \cdot 16 + 10 \cdot 1)_{10} = 3370_{10}$.

1.57 Satz: Seien g, Z fest. Jede Zahl $n \in \mathbb{N}$ besitzt eine eindeutige g-adische Darstellung: Es existieren eindeutig bestimmte Zahlen $N \in \mathbb{N}$ und $a_0, \ldots, a_N \in Z$, so dass

$$n = (a_N a_{N-1} \ldots a_0)_g \;\wedge\; a_N \neq z_0.$$

Beweis: Existenz (konstruktiv):

Teilen mit Rest:

$$\begin{array}{rcl c rcl}
n & = & q_1 \cdot g + r_1 & \Rightarrow & a_0 & = & r_1 \\
q_1 & = & q_2 \cdot g + r_2 & \Rightarrow & a_1 & = & r_2 \\
 & \vdots & & & & & \\
q_{N-2} & = & q_{N-1} \cdot g + r_{N-1} & \Rightarrow & a_{N-2} & = & r_{N-1} \\
q_{N-1} & = & q_N \cdot g + r_N & \Rightarrow & a_{N-1} & = & r_N
\end{array}$$

Verfahren bricht ab, wenn $0 < q_N < g$. Dann setze noch $a_N := q_N$.

Rückwärts hochgehen:

$$\begin{array}{rclcl}
q_{N-1} & = & a_N \cdot g + a_{N-1} & & \\
q_{N-2} & = & q_{N-1} \cdot g + a_{N-2} & = & a_N \cdot g^2 + a_{N-1} \cdot g + a_{N-2} \\
q_{N-3} & = & q_{N-2} \cdot g + a_{N-3} & = & a_N \cdot g^3 + a_{N-1} \cdot g^2 + a_{N-2} \cdot g + a_{N-3} \\
 & \vdots & & & \\
n & = & \ldots & = & a_N \cdot g^N + \ldots + a_1 \cdot g + a_0
\end{array}$$

□

Z.B: 2007 im Fünfersystem:
$$\begin{aligned} 2007 : 5 &= 401 \quad \text{Rest } 2 \quad \Rightarrow \quad a_0 = 2 \\ 401 : 5 &= 80 \quad \text{Rest } 1 \quad \Rightarrow \quad a_1 = 1 \\ 80 : 5 &= 16 \quad \text{Rest } 0 \quad \Rightarrow \quad a_2 = 0 \\ 16 : 5 &= 3 \quad \text{Rest } 1 \quad \Rightarrow \quad a_3 = 1 \end{aligned}$$

Verfahren bricht ab, da $3 < 5$. Also $a_4 = 3$ und $(2007)_{10} = (31012)_5$.

1.11 Mächtigkeit von Mengen

1.58 Definition: Zwei Mengen A, B heißen **gleich groß** oder **gleich mächtig**, falls es eine bijektive Abbildung $f : A \to B$ gibt.

Z.B.: $f : \{1, 2, 3\} \to \{4, 5, 6\} : x \mapsto x + 3$ ist bijektiv

Z.B.: $f :]-\infty, -3] \to [-3, \infty[: x \mapsto -6 - x$ (Spiegelung an -3) ist bijektiv, also sind die Mengen gleich mächtig.

Z.B.: $G := \{2n : n \in \mathbb{N}\}$ und $\mathbb{N}$ sind gleich mächtig, obwohl G echte Teilmenge von $\mathbb{N}$.

Z.B.: $\mathbb{Q}$ ist gleich mächtig wie $\mathbb{N}$.

1.59 Definition: Sei A eine Menge.

1) A heißt **endlich**, falls jede injektive Abbildung $f : A \to A$ auch surjektiv ist. ($\emptyset$ ist automatisch endlich.)

2) A heißt **unendlich**, falls A nicht endlich. D.h. falls es eine injektive Abbildung $f : A \to A$ gibt, die nicht surjektiv ist, oder falls A gleich mächtig ist wie eine echte Teilmenge von A.

3) A heißt **abzählbar unendlich**, falls A gleich mächtig wie $\mathbb{N}$ ist.

4) A heißt **überabzählbar**, falls A unendlich und nicht abzählbar unendlich ist.

Z.B.: $\mathbb{Q}$ ist abzählbar. Später: $\mathbb{R}$ ist überabzählbar.

2 Zahlenkörper

2.1 Mengen und Verknüpfungen

Was ist $+$? Eine Abbildung: $+ : \mathbb{N} \times \mathbb{N} \to \mathbb{N} : (a, b) \mapsto a + b$.

2.1 Definition: Sei G Menge. Eine Abbildung $\circ : G \times G \to G : (g, h) \to g \circ h$ heißt **Verknüpfung** auf G.

1) Eine Verknüpfung heißt **assoziativ**, falls

$$\forall g, h, j \in G : g \circ (h \circ j) = (g \circ h) \circ j.$$

2) Eine Verknüpfung heißt **kommutativ**, falls

$$\forall g, h \in G : g \circ h = h \circ g.$$

3) Ein Element $e \in G$ heißt **neutrales Element**, falls

$$\forall g \in G : e \circ g = g = g \circ e.$$

4) Sei $g \in G$. Ein Element $h \in G$ heißt **inverses Element** zu g, falls $g \circ h = e = h \circ g$. Schreibweise: $g^{-1} := h$.

Z.B.: $(\mathbb{N}, +)$.

Z.B.: $(\mathbb{N}, \cdot)$.

Z.B.: $(\mathbb{Z}, +)$.

Z.B.: $(\mathbb{Q}, \cdot)$.

2.2 Definition: Sei $\circ$ eine Verknüpfung auf G.

1) $(G, \circ)$ heißt **Monoid**, falls $\circ$ assoziativ ist und ein neutrales Element existiert.

2) $(G, \circ)$ heißt **Gruppe**, falls $(G, \circ)$ ein Monoid ist und jedes Element von G ein inverses Element besitzt.

3) $(G, \circ)$ heißt **abelsche Gruppe**, falls $(G, \circ)$ eine Gruppe ist und $\circ$ kommutativ ist.

Z.B.: Monoide sind $(\mathbb{N}_0, +)$, $(\mathbb{N}, \cdot)$, $(\mathcal{P}(M), \cap)$.

Z.B.: Abelsche Gruppen sind $(\mathbb{Z}, +)$, $(\mathbb{Q}, +)$, $(\mathbb{Q} \setminus \{0\}, \cdot)$.

Wichtiges Beispiel: Die **Permutationsgruppe** $(S(M), \circ)$. Sei

$$\begin{aligned} &M \text{ endliche Menge,} \\ &S(M) \ := \ \{\pi : M \to M \ \big| \ \pi \text{ ist bijektive Abbildung}\}, \\ &\circ : (\pi_1, \pi_2) \mapsto \pi_1 \circ \pi_2 \text{ (Hintereinanderausführung).} \end{aligned}$$

Dann ist $(S(M), \circ)$ eine Gruppe. Falls $\sharp M \geq 3$ ist diese Gruppe nicht kommutativ.

2.3 Satz: In jeder Gruppe $(G, \circ)$ gelten:

1) e ist eindeutig.

2) g^{-1} ist eindeutig.

3) $\forall g, h \in G \ \ \exists! x \in G : g \circ x = h$ (nämlich $x = g^{-1} \circ h$).

4) $\forall g \in G : (g^{-1})^{-1} = g$.

5) $\forall g, h \in G : (g \circ h)^{-1} = h^{-1} \circ g^{-1}$.

Beweis: **1)** Seien e, e' zwei neutrale Elemente. Dann: $e = e' \circ e = e'$.

2) Sei $g \circ g^{-1} = e = g^{-1} \circ g$ und $g \circ h = e = h \circ g$.

$$\Rightarrow \quad h \ = \ e \circ h \ = \ (g^{-1} \circ g) \circ h \ = \ g^{-1} \circ (g \circ h) \ = \ g^{-1} \circ e \ = \ g^{-1}.$$

3) Existenz: $x := g^{-1} \circ h \ \Rightarrow \ g \circ x = g \circ (g^{-1} \circ h) = (g \circ g^{-1}) \circ h = e \circ h = h$.
Eindeutigkeit: $g \circ x = h \ \Rightarrow \ g^{-1} \circ g \circ x = g^{-1} \circ h \ \Rightarrow \ x = g^{-1} \circ h$.

4), **5)** als Übung. $\square$

Schreibweise: Für $g \in G$ und $n \in \mathbb{Z}$ setzt man

$$g^n \ := \ \begin{cases} \underbrace{g \circ \ldots \circ g}_{n \text{ Mal}} & \text{für } n \in \mathbb{N}, \\ e & \text{für } n = 0, \\ \underbrace{g^{-1} \circ \ldots \circ g^{-1}}_{-n \text{ Mal}} & \text{für } -n \in \mathbb{N} \end{cases}$$

Dann gilt $g^n \circ g^m = g^{n+m}$ für $n, m \in \mathbb{Z}$.

2.2 Zwei Verknüpfungen

2.4 Definition: Auf der Menge R seien zwei Verknüpfungen $+, \cdot$ definiert.

1) $(R, +, \cdot)$ heißt **Ring**, falls

- **a)** $(R, +)$ ist abelsche Gruppe und
- **b)** $(R, \cdot)$ ist Monoid und
- **c)** $0 \neq 1$ ($0 =$ neutrales Element bezüglich $+$ und $1 =$ neutrales Element bezüglich $\cdot$) und
- **d)** es gelten die **Distributivgesetze**:

$$\forall g, h, j \in R : (g + h) \cdot j = g \cdot j + h \cdot j \ \wedge g \cdot (h + j) = g \cdot h + g \cdot j.$$

2) $(R, +, \cdot)$ heißt **kommutativer Ring**, falls zusätzlich zu 1) $\cdot$ kommutitativ ist.

3) $(R, +, \cdot)$ heißt **Körper** (*englisch: field*), falls zusätzlich zu 1) $(R \setminus \{0\}, \cdot)$ eine abelsche Gruppe ist.

Z.B.: $(\mathbb{Z}, +, \cdot)$ ist kommutativer Ring.

Z.B.: $(\mathbb{Q}, +, \cdot)$ ist ein Körper.

Z.B.: p Primzahl. Dann ist $(\mathbb{Z}/p\mathbb{Z}, +, \cdot)$ ein Körper.

Z.B.: m keine Primzahl. Dann ist $(\mathbb{Z}/m\mathbb{Z}, +, \cdot)$ ein kommutativer Ring, aber kein Körper.

Z.B.: $R := \{f : \mathbb{Q} \to \mathbb{Q} \text{ Abbildung}\}$
$f + g : x \mapsto f(x) + g(x)$
$f \cdot g : x \mapsto f(x) \cdot g(x)$
Dann ist $(R, +, \cdot)$ ein kommutativer Ring.

2.5 Schreibweisen:

$$\begin{aligned}
-x &:= \text{Inverses bezüglich } + \\
x - y &:= x + (-y) \\
\frac{1}{x} := x^{-1} &:= \text{Inverses bezüglich } \cdot \\
\frac{x}{y} &:= x \cdot y^{-1} \text{ (oder auch } x : y) \\
x^n &:= \underbrace{x \cdot x \cdots x}_{n \text{ mal}}, \qquad x^{-n} := (x^{-1})^n
\end{aligned}$$

2.6 Folgerungen: Sei $(\mathbb{K}, +, \cdot)$ ein Körper. Für $a, b, c, d \in \mathbb{K}$ gelten:

1) $-(-x) = x$

2) $0 \cdot x = x \cdot 0 = 0$

3) $x \cdot y = 0 \ \Rightarrow \ x = 0 \vee y = 0$

4) $(-x) \cdot y = -(x \cdot y)$

5) $(-x) \cdot (-y) = x \cdot y$

6) $\dfrac{1}{\frac{1}{x}} = x$

7) $\dfrac{a}{b} + \dfrac{c}{d} = \dfrac{a \cdot d + b \cdot c}{b \cdot d}$

8) $\dfrac{a}{b} \cdot \dfrac{c}{d} = \dfrac{a \cdot c}{b \cdot d}$

9) $\dfrac{a}{b} : \dfrac{c}{d} = \dfrac{a}{b} \cdot \dfrac{d}{c} = \dfrac{a \cdot d}{b \cdot c}$

2.3 Die reellen Zahlen

2.7 Definition: Die **reellen Zahlen** $(\mathbb{R}, +, \cdot)$ bilden einen Körper mit folgenden Eigenschaften:

(O) $(\mathbb{R}, <)$ ist ein angeordneter Körper,

(A) $(\mathbb{R}, <)$ ist ein archimedischer Körper,

(V) $(\mathbb{R}, |.|)$ ist vollständig.

2.3.1 Die Anordnung in $\mathbb{R}$ und Folgerungen

2.8 Axiom* (O)*: $(\mathbb{R}, <)$ ist ein **angeordneter Körper**, d.h. auf $\mathbb{R}$ ist eine Relation $<$ (kleiner) definiert, so dass:

(O1) $\forall a, b \in \mathbb{R}$: Genau eine der Beziehungen $a < b$, $a = b$, $b < a$ ist wahr.

(O2) $\forall a, b, c \in \mathbb{R} : a < b \wedge b < c \Rightarrow a < c$ (Transitivität).

(O3) $\forall a, b, c \in \mathbb{R} : a < b \Rightarrow a + c < b + c$ (Verträglichkeit mit Addition).

(O4) $\forall a, b, c \in \mathbb{R} : a < b \wedge 0 < c \Rightarrow a\,c < b\,c$ (Verträglichkeit mit Multiplikation).

2.9 Bemerkungen: **1)** Für $a < b$ schreibe auch $b > a$ (größer).

2) $a \in \mathbb{R}$ heißt **positiv**, falls $a > 0$. Sei

$$\mathbb{R}^+ := \{a \in \mathbb{R} : a > 0\}, \qquad \mathbb{R}_0^+ := \mathbb{R}^+ \cup \{0\}.$$

3) Definiere

$$\begin{aligned} a \le b \quad &:\Leftrightarrow \quad a < b \vee a = b, \\ a \ge b \quad &:\Leftrightarrow \quad a > b \vee a = b; \end{aligned}$$

$\le$ und $\ge$ sind Ordnungsrelationen auf $\mathbb{R}$.

4) Ein Ausdruck der Form $a < b,\ a \le b,\ a > b,\ a \ge b$ heißt **Ungleichung**.

5) Endliche Körper können nicht angeordnet werden (siehe (U6) unten).

2.10 Folgerungen: **(U1)** $a < b \Leftrightarrow 0 < b - a \Leftrightarrow b - a > 0$

(U2) $a < b \wedge x < y \Rightarrow a + x < b + y$.
Insbesondere: $a < 0 \wedge x < 0 \Rightarrow a + x < 0; \quad 0 < b \wedge 0 < y \Rightarrow 0 < b + y$.

(U3) $a > 0 \;\Leftrightarrow\; -a < 0$ und $a < 0 \;\Leftrightarrow\; -a > 0$.

(U4) $0 < a < b \,\wedge\, 0 < x < y \;\Rightarrow\; a\,x < b\,y$.

(U5) $a < 0 \,\wedge\, x < y \;\Rightarrow\; a\,x > a\,y$.
bzw. $a < 0 \,\wedge\, x > y \;\Rightarrow\; a\,x < a\,y$,
Multiplikation von Ungleichungen mit negativen Zahlen vertauscht $<$ und $>$.

(U6) $a \neq 0 \;\Rightarrow\; a^2 > 0$. Insbesondere: $1 = 1^2 > 0$.
In endlichen Körpern gibt es keine Ordnung: $1 > 0 = 1 + 1 + \ldots + 1 > 1\, ↯$.

(U7) $a > 0 \;\Rightarrow\; \frac{1}{a} > 0$; $\quad a < 0 \;\Rightarrow\; \frac{1}{a} < 0$.

(U8) $0 < a < b \;\Rightarrow\; \frac{1}{a} > \frac{1}{b}$.

(U9) $0 < a < b \;\Rightarrow\; \frac{b}{a} > 1$.

2.11 Satz (Bernoulli-Ungleichung): Für $x \in \mathbb{R},\; x \geq -1$ und $n \in \mathbb{N}_0$ gilt

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

2.12 Intervalle: Sei $a < b$. Definiere

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

analog: $]a,b]$; $\quad]-\infty,b]$; $\quad]-\infty,b[$; $\quad]a,\infty[$; $\quad]-\infty,\infty[= \mathbb{R}$.

2.13 $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$ in $\mathbb{R}$: **1)** Identifiziere $\mathbb{N}$ als Teilmenge von $\mathbb{R}$: Definiere $f : \mathbb{N} \to \mathbb{R}$ durch

$$f(1_{\mathbb{N}}) := 1_{\mathbb{R}}, \quad f(n + 1_{\mathbb{N}}) := f(n) + 1_{\mathbb{R}}.$$

Nach dem Induktionsaxiom ist f für alle $n \in \mathbb{N}$ definiert.
Setze $\widetilde{\mathbb{N}} := f(\mathbb{N}) = \{f(n) : n \in \mathbb{N}\}$.

Die Abbildung f ist injektiv: Zeige $n \neq m \;\Rightarrow\; f(n) \neq f(m)$:
Sei $n \neq m$. O.B.d.A. $n < m$, also $m = n + k$.
$\Rightarrow\; f(n) < f(n) + 1_{\mathbb{R}} = f(n + 1_{\mathbb{N}}) < f(n + 2_{\mathbb{N}}) < \ldots < f(n+k) = f(m)$.
$\overset{\text{(O1)}}{\Rightarrow}\; f(n) \neq f(m)$.

Also: $f : \mathbb{N} \to \widetilde{\mathbb{N}}$ ist bijektiv.
Damit gelten die Peano-Axiome automatisch in $\widetilde{\mathbb{N}}$ mit Nachfolger$(n) = n + 1_{\mathbb{R}}$.
Identifiziere nun $\mathbb{N}$ mit $\widetilde{\mathbb{N}}$, schreibe im Folgenden $\mathbb{N}$ anstelle von $\widetilde{\mathbb{N}}$. (Typisches mathematisches Vorgehen.)

2) Setze $\mathbb{Z} := \{k \in \mathbb{R} \mid \exists n, m \in \mathbb{N} : k = n - m\}$.

3) Setze $\mathbb{Q} := \{x \in \mathbb{R} \mid \exists p \in \mathbb{Z},\ q \in \mathbb{N} : x = \frac{p}{q}\}$.

Wir übertragen Arithmetik und Ordnung von $\mathbb{R}$ auf diese Teilmengen, soweit das geht.

2.14 Definition: Der **Betrag** oder **Absolutbetrag** von $x \in \mathbb{R}$ ist definiert durch

$$|x| := \begin{cases} x & \text{für } x \geq 0 \\ -x & \text{für } x < 0\ . \end{cases}$$

Insbesondere gelten

$$|x| \ \geq \ x, \quad |x| \ = \ |-x| \ \geq \ -x.$$

2.15 Eigenschaften von |.| : Für $x, y \in \mathbb{R}$ gelten:

(B1) $|x| \geq 0 \quad \wedge \quad (|x| = 0 \ \Leftrightarrow \ x = 0)$

(B2) $|x \cdot y| = |x| \cdot |y|$

(B3) $|x + y| \ \leq \ |x| + |y|$ (**Dreiecksungleichung**, **$\boldsymbol{\Delta}$-Ungleichung**)

Beweis: (B1)

$$\begin{aligned} x > 0 \ &\Rightarrow \ |x| = x > 0 \\ x = 0 \ &\Rightarrow \ |x| = x = 0 \\ x < 0 \ &\Rightarrow \ |x| = -x > 0 \quad \text{(nach (U3))} \end{aligned}$$

(B2) Unterscheide 4 mögliche Fälle $x \geq 0 \ \wedge y \geq 0, \ldots$, getrennt nachrechnen.

(B3)

$$\begin{aligned} x + y \geq 0 \ &\Rightarrow \ |x + y| = x + y \leq |x| + |y| \\ x + y < 0 \ &\Rightarrow \ |x + y| = -(x + y) = (-x) + (-y) \leq |x| + |y| \end{aligned}$$

□

2.16 Definition: Ein Körper $(\mathbb{K}, +, \cdot)$ mit einem Betrag, d.h. einer Funktion $|.| : \mathbb{K} \to \mathbb{R}_0^+$ mit (B1), (B2), (B3), heißt **bewerteter Körper** (z.B. auch $\mathbb{C}$).

2.17 Satz: In jedem bewerteten Körper gelten:

1) $|1| \ = \ |-1| \ = \ 1$, insbesondere $|-x| \ = \ |x|$.

2) $\left|\dfrac{x}{y}\right| \ = \ \dfrac{|x|}{|y|}$ falls $y \neq 0$.

3) $|x-y| \;\geq\; \big||x|-|y|\big| \quad (\geq\; |x|-|y|),$
$|x+y| \;\geq\; \big||x|-|y|\big|$
(**Δ-Ungleichung nach unten**).

Beweis: **2)** $x = \dfrac{x}{y}\cdot y \;\Rightarrow\; |x| = \left|\dfrac{x}{y}\right|\cdot|y| \;\Rightarrow$ Behauptung.

1) $|1| = \left|\dfrac{1}{1}\right| \overset{2)}{=} \dfrac{|1|}{|1|} = 1.$

$$|-1|\cdot|-1| \overset{(B2)}{=} \big|(-1)\big|^2 = |1| = 1.$$

$$\begin{aligned} b := |-1| \quad &\Rightarrow \quad b^2 = 1 \;\wedge\; b \geq 0 \\ &\Rightarrow \quad b \geq 0 \;\wedge\; 0 = b^2-1 = (b+1)(b-1) \\ &\underset{\text{Nullprodukt}}{\overset{\text{Satz vom}}{\Rightarrow}} \quad b \geq 0 \;\wedge\; (b=-1 \;\vee\; b=1) \\ &\Rightarrow \quad b = -1 \end{aligned}$$

3) Als Übung

□

2.18 Definition: In einem bewerteten Körper ist der **Abstand** zweier Elemente $x, y \in \mathbb{K}$ definiert durch

$$d(x,y) \;:=\; |x-y|.$$

Die Abstandsfunktion $d : \mathbb{K}\times\mathbb{K} \to \mathbb{R}_0^+$ besitzt die Eigenschaften

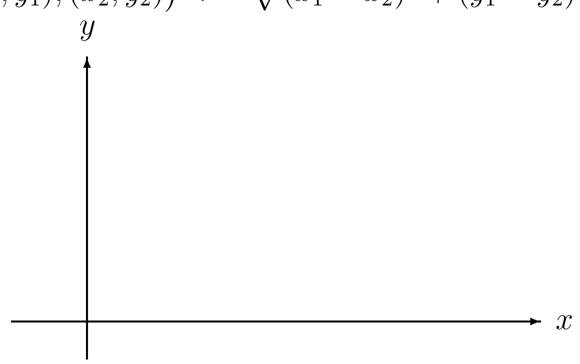
(M1) $d(x,y) \;\geq\; 0 \quad\wedge\quad (d(x,y) \;=\; 0 \;\Leftrightarrow\; x=y)$ (**Positivität**, **Definitheit**).

(M2) $d(x,y) \;=\; d(y,x)$ (**Symmetrie**).

(M3) $d(x,y) \;\leq\; d(x,z)+d(z,y)$ (**Δ-Ungleichung**).

Diese drei Eigenschaften werden später dazu benützt, abstrakt den Begriff eines Abstandes (Metrik) zu definieren.

Z.B.: $\mathbb{R}^2 = \mathbb{R}\times\mathbb{R},\; d\big((x_1,y_1),(x_2,y_2)\big) \;:=\; \sqrt{(x_1-x_2)^2+(y_1-y_2)^2}$ (euklidischer Abstand).



2.19 Definition: Sei M Menge. Eine **Folge in M** ist eine Abbildung $\mathbb{N} \ni n \mapsto x_n \in M$. Schreibweise: (x_n) oder $(x_n)_{n\in\mathbb{N}}$.

2.20 Konvergenz: Sei $(\mathbb{K}, |.|)$ bewerteter Körper. Die Folge (x_n) in $\mathbb{K}$ heißt **konvergent gegen $x \in \mathbb{K}$**, falls

$$\forall \varepsilon > 0 \ \ \exists N_\varepsilon \in \mathbb{N} \ \ \forall n > N_\varepsilon : |x_n - x| \ < \ \varepsilon.$$

D.h. für jedes noch so kleine $\varepsilon > 0$ ist der Abstand aller x_n zu x kleiner als ε, sobald $n > N_\varepsilon$ gilt. In diesem Fall schreiben wir

$$x \ = \ \lim_{n\to\infty} x_n, \qquad x_n \overset{n\to\infty}{\longrightarrow} x, \qquad x_n \to x \text{ für } n \to \infty, \qquad x_n \to x;$$

x heißt **Grenzwert** oder **Limes** der Folge. Eine Folge (x_n) heißt **konvergent**, falls

$$\exists x \in \mathbb{K} : x = \lim_{n\to\infty} x_n.$$

Ist eine Folge nicht konvergent, so heißt sie **divergent**.

2.21 Eindeutigkeit des Grenzwertes: Sei $(\mathbb{K}, |.|)$ bewerteter Körper. Ist $\lim_{n\to\infty} x_n = x$, so besitzt (x_n) keinen weiteren Grenzwert.

Beweis: Widerspruchsbeweis, Annahme: $x = \lim_{n\to\infty} x_n \ \wedge \ y = \lim_{n\to\infty} x_n \ \wedge \ x \neq y$.

Wähle $\varepsilon := \dfrac{|x-y|}{2}$. Dann: $\exists N_\varepsilon \in \mathbb{N} \ \ \forall n > N_\varepsilon : |x_n - x| \ < \ \varepsilon$

$\exists \widetilde{N}_\varepsilon \in \mathbb{N} \ \ \forall n > \widetilde{N}_\varepsilon : |x_n - y| \ < \ \varepsilon$

Für $n > \max\{N_\varepsilon, \widetilde{N}_\varepsilon\}$ folgt

$$|x-y| = |x - x_n + x_n - y| \le |x - x_n| + |x_n - y| = |x - x_n| + |y - x_n| < 2\varepsilon = |x - y| \quad ↯$$

$\square$

2.22 Satz: Sei (x_n) konvergente Folge in $\mathbb{R}$ und $K \in \mathbb{R}$, $N \in \mathbb{N}$. Dann

$$(\forall n \ge N : x_n \le K) \ \ \Rightarrow \ \ \lim_{n\to\infty} x_n \le K,$$
$$(\forall n \ge N : x_n \ge K) \ \ \Rightarrow \ \ \lim_{n\to\infty} x_n \ge K.$$

2.23 Achtung: $x_n := \dfrac{1}{n} > 0$, aber $\underbrace{\lim_{n\to\infty} x_n = 0}_{\text{Beweis später}}$.

2.3.2 Die archimedische Anordnung

***2.24 Axiom* (A):** $(\mathbb{R},<)$ ist ein **archimedisch geordneter** Körper, d.h. es gilt

$$\forall x \in \mathbb{R} \quad \exists n \in \mathbb{N} : n > x.$$

2.25 Folgerungen: Aus (A) folgt

1) $\forall y \in \mathbb{R} \quad \forall x \in \mathbb{R}^+ \quad \exists n \in \mathbb{N} : n\,x > y$

2) $\forall x \in \mathbb{R}_0^+ \quad \exists! n \in \mathbb{N}_0 : n \leq x < n+1$
Definiere die **untere Gaußklammer**: $\lfloor x \rfloor := \max\{n \in \mathbb{N}_0 : n \leq x\}$.
Dann $\lfloor x \rfloor \in \mathbb{N}$ und $\lfloor x \rfloor \leq x < \lfloor x \rfloor + 1$.

3) $\forall \varepsilon > 0 \quad \exists n \in \mathbb{N} : \frac{1}{n} < \varepsilon$

4) $x \in \mathbb{R}_0^+ \ \wedge \ (\forall n \in \mathbb{N} : x < \frac{1}{n}) \ \Rightarrow \ x = 0$

5) $\forall b > 1 \quad \forall x > 0 \quad \exists n \in \mathbb{N} : b^n > x$

6) $\forall b \in \,]0,1[\quad \forall \varepsilon > 0 \quad \exists n \in \mathbb{N} : b^n < \varepsilon$

2.26 Anwendung auf Folgen: **1)** $\dfrac{1}{n} \to 0.$

2) Sei $|q| < 1$. Dann $q^n \to 0$.

3) Sei $|q| < 1$ und

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

Dann $s_n \to \dfrac{1}{1-q}$.

4) Die Folge (x_n) mit $x_n = (-1)^n$ ist nicht konvergent (Beweis später).

2.3.3 Die Vollständigkeit

2.27 Verdichtungsprinzip: Sei (x_n) konvergent. Dann gilt:

$$\forall \varepsilon > 0 \quad \exists N_\varepsilon \in \mathbb{N} \quad \forall n,m > N_\varepsilon : |x_n - x_m| \ < \ \varepsilon. \qquad (*)$$

Beweis: Sei $\varepsilon > 0$. Aus $x_n \to x$: $\exists N_\varepsilon \in \mathbb{N} \;\; \forall n > N_\varepsilon : |x_n - x| \; < \; \frac{\varepsilon}{2}$.

Für $n, m > N_\varepsilon$ folgt

$$|x_n - x_m| = |x_n - x + x - x_m| \le |x_n - x| + |x - x_m| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2}.$$

□

2.28 Definition: Eine Folge (x_n) in einem bewerteten Körper heißt **Cauchy-Folge**, falls sie das Verdichtungsprinzip $(*)$ erfüllt.

Also: (x_n) konvergent $\Rightarrow$ (x_n) Cauchy-Folge

(x_n) keine Cauchy-Folge $\Rightarrow$ (x_n) divergent

2.29 Dezimalbrüche: Sei $a_0 \in \mathbb{N}_0$, und für $j \in \mathbb{N}$ sei $d_j \in \{0, 1, \ldots, 9\}$. Setze

$$a_n \;:=\; a_0 + \frac{d_1}{10} + \frac{d_2}{100} + \ldots + \frac{d_n}{10^n} \in \mathbb{Q} \subseteq \mathbb{R}.$$

Schreibe zur Abkürzung $a_n = a_0, d_1 d_2 \ldots d_n$. Man nennt die Folge (a_n) einen **Dezimalbruch**. Behauptung: (a_n) ist eine Cauchy-Folge.

Beweis: Sei $m > n$, also $m = n + k$. Dann gilt

$$\begin{aligned}
|a_m - a_n| = a_m - a_n &= \frac{d_{n+1}}{10^{n+1}} + \ldots + \frac{d_{n+k}}{10^{n+k}} \\
&\le \frac{9}{10^{n+1}} + \ldots \frac{9}{10^{n+k}} \\
&= \frac{9}{10^{n+1}} \left(1 + \frac{1}{10} + \ldots \frac{1}{10^{k-1}}\right) \\
&= \frac{9}{10^{n+1}} \frac{1 - \left(\frac{1}{10}\right)^k}{1 - \frac{1}{10}} \\
&= \frac{9}{10^{n+1}} \frac{10}{9} \left(1 - \left(\frac{1}{10}\right)^k\right) \\
&\le \frac{1}{10^n} < \varepsilon \text{ für } n > N_\varepsilon, \text{ da } \left(\frac{1}{10}\right)^n \to 0
\end{aligned}$$

□

2.30 Axiom* (V)*: $(\mathbb{R}, |.|)$ ist **vollständig**, d.h. in $\mathbb{R}$ konvergiert jede Cauchy-Folge.

Also: In $\mathbb{R}$ gilt $\quad (x_n)$ konvergent $\Leftrightarrow$ (x_n) ist Cauchy-Folge

2.31 Satz: **1)** Jeder Dezimalbruch konvergiert gegen eine reelle Zahl.

2) Für jede relle Zahl $x \in \mathbb{R}$ existiert ein Dezimalbruch, der gegen x konvergiert.

Meist identifiziert man x mit dem Dezimalbruch, z.B.

$$\sqrt{2} \;=\; 1,41421\ldots$$

Beweis: **1)** Folgt aus (V) und 2.29.

2) Sei $x \in \mathbb{R}_0^+$.
Setze $a_0 := \lfloor x \rfloor$. Dann $0 \leq x - a_0 < 1$ und $0 \leq 10(x - a_0) < 10$.

Setze $d_1 := \lfloor 10(x - a_0) \rfloor \;\Rightarrow \begin{cases} d_1 \in \{0, \ldots, 9\} \\ \text{Für } a_1 := a_0 + \frac{d_1}{10} \text{ gilt } 0 \leq x - a_1 < \frac{1}{10} \end{cases}$

Setze $d_2 := \lfloor 100(x - a_1) \rfloor \;\Rightarrow \begin{cases} d_2 \in \{0, \ldots, 9\} \\ \text{Für } a_2 := a_0 + \frac{d_1}{10} + \frac{d_2}{100} \text{ gilt } 0 \leq x - a_2 < \frac{1}{100} \end{cases}$

$\vdots$

Dann ist (a_n) Dezimalbruch und $|x - a_n| < \frac{1}{10^n}$, also $a_n \to x$.

Für $x \in \mathbb{R}^-$ führe Verfahren für $-x$ durch. $\square$

Mit diesem Verfahren erhält man zu x einen eindeutigen Dezimalbruch. Alle Dezimalbrüche kommen vor bis auf diejenigen, für die gilt:

$$\exists n_0 \in \mathbb{N} \;\; \forall n \geq n_0 : d_n = 9.$$

Denn sei z.B. $a_n = 3,19\ldots9. \quad \Rightarrow \; |3,2 - a_n| = 10^{-n} \to 0 \; \Rightarrow \; a_n \to 3,2.$
Für $x = 3,2$ liefert das obige Verfahren aber $x = 3,200\ldots$.

2.32 Bemerkung: Entsprechend kann eine g-adische Entwicklung von $x \in \mathbb{R}$ konstruiert werden:

$$x \;=\; (a_N a_{N-1} \ldots a_0, a_{-1}, a_{-2} \ldots)_g \;:=\; \lim_{n\to\infty} \sum_{k=0}^{n} a_{N-k}\, g^{N-k}$$

mit $N \in \mathbb{N},\; a_j \in \{0, 1, \ldots, g-1\}$.

2.33 Satz ($\mathbb{R}$ ***und*** $\mathbb{Q}$***):*** **1)** $\mathbb{Q}$ ist abzählbar ($\mathbb{Q} \subseteq \mathbb{Z} \times \mathbb{N}$, Cantorsches Diagonalverfahren).

2) $\mathbb{R}$ ist überabzählbar. Insbesondere $\mathbb{Q} \subsetneqq \mathbb{R}$.

3) Für jedes $x \in \mathbb{R}$ existiert eine Folge (q_n) in $\mathbb{Q}$ mit $x = \lim\limits_{n\to\infty} q_n$.
Man sagt: $\mathbb{Q}$ liegt **dicht** in $\mathbb{R}$.

Beweis: **1)** Schon gemacht

2) Zeige: Ist $f : \mathbb{N} \to \mathbb{R}$, so ist f nicht surjektiv.
Sei also eine Abbildung $f : \mathbb{N} \to \mathbb{R}$ gegeben. Definiere $x := 0, d_1\, d_2 \ldots$, wobei

$$d_j \; := \; \begin{cases} 0 \quad \text{falls } f(j) = \tilde{a_0}, \tilde{d}_1\, \tilde{d}_2 \ldots \; \text{ und } \tilde{d}_j \neq 0 \\ 1 \quad \text{falls } f(j) = \tilde{a_0}, \tilde{d}_1\, \tilde{d}_2 \ldots \; \text{ und } \tilde{d}_j = 0 \end{cases}$$

$\Rightarrow \; x \in \mathbb{R} \; \wedge \; x \neq f(j)$ für $j \in \mathbb{N}$
$\Rightarrow \; x \in \mathbb{R} \setminus f(\mathbb{N})$.

3) Wähle (q_n) = zu x konstruierter Dezimalbruch. $\square$

Also: Auf dem Zahlenstrahl hat die Menge $\mathbb{Q}$ Lücken. Die Menge der Lücken ist größer als $\mathbb{Q}$.

2.34 Definition: **1)** Eine Folge (a_n) heißt **monoton wachsend / monoton fallend**, falls $a_{n+1} \geq a_n \; / \; a_{n+1} \leq a_n$ für $n \in \mathbb{N}$.

2) Seien $(a_n), (b_n)$ Folgen in $\mathbb{R}$. Die Folge von Intervallen $([a_n, b_n])$ heißt **Intervallschachtelung**, falls

- (a_n) monoton wachsend, (b_n) monoton fallend,
- $\forall n \in \mathbb{N} : a_n \leq b_n$,
- $\lim\limits_{n\to\infty} (b_n - a_n) = 0$.

2.35 Satz: Ist $([a_n, b_n])_{n\in\mathbb{N}}$ eine Intervallschachtelung, so konvergieren die Folgen $(a_n), (b_n)$ gegen denselben Grenzwert. Außerdem existiert genau ein $x \in \mathbb{R}$ mit $x \in \bigcap\limits_{n=1}^{\infty} [a_n, b_n]$, nämlich $x = \lim\limits_{n\to\infty} a_n = \lim\limits_{n\to\infty} b_n$.

Beweis: 1) (a_n) konvergiert: Sei $\varepsilon > 0$.
$b_n - a_n \to 0 \; \Rightarrow \; \exists N_\varepsilon \in \mathbb{N} \;\; \forall n > N_\varepsilon : |b_n - a_n| < \varepsilon$.
Für $m > n > N_\varepsilon$ folgt $|a_m - a_n| = a_m - a_n \leq b_m - a_n \leq b_n - a_n < \varepsilon$
$\Rightarrow \; (a_n)$ ist Cauchy-Folge
$\Rightarrow \; a_n \to a \in \mathbb{R}$.

2) $b_n \to a$: Sei $\varepsilon > 0$.

$$b_n - a_n \to 0 \;\; \Rightarrow \;\; \exists N_\varepsilon \in \mathbb{N} \;\; \forall n > N_\varepsilon |b_n - a_n| < \frac{\varepsilon}{2}$$
$$a_n \to a \;\; \Rightarrow \;\; \exists \tilde{N}_\varepsilon \in \mathbb{N} \;\; \forall n > \tilde{N}_\varepsilon |a_n - a| < \frac{\varepsilon}{2}$$

$\Rightarrow$ für $n > \max\{N_\varepsilon, \tilde{N}_\varepsilon\} : |b_n - a| = |b_n - a_n + a_n - a| \leq |b_n - a_n| + |a_n - a| < \varepsilon$.

3) $a \in [a_n, b_n]$ für alle $n \in \mathbb{N}$: Sei $n \in \mathbb{N}$ fest.

$$\left.\begin{array}{lll} \text{Für } m > n: & a_m \leq b_m \leq b_n & \Rightarrow \quad a = \lim_{m\to\infty} a_m \leq b_n \\ \text{Für } m > n: & a_n \leq a_m & \Rightarrow \quad a_n \leq \lim_{m\to\infty} a_m = a \end{array}\right\} \Rightarrow \; a_n \leq a \leq b_n.$$

4) Eindeutigkeit:

$$\begin{array}{lllll} y < a & \Rightarrow & \exists n \in \mathbb{N} : a_n > y & \Rightarrow & y \notin \bigcap\limits_{n\in\mathbb{N}} [a_n, b_n] \\ y > a & \Rightarrow & \exists n \in \mathbb{N} : b_n < y & \Rightarrow & y \notin \bigcap\limits_{n\in\mathbb{N}} [a_n, b_n] \end{array}$$

□

2.36 Definition: Sei $M \subseteq \mathbb{R}$, $M \neq \emptyset$.

1) $a \in \mathbb{R}$ heißt eine **obere** (**untere**) **Schranke** von M, falls

$$\forall x \in M : x \leq a \quad (x \geq a)$$

2) M heißt **nach oben beschränkt**, falls M eine (und damit unendlich viele) obere Schranke besitzt. Analog: Nach unten beschränkt.
M heißt **beschränkt**, falls M eine obere und eine untere Schranke besitzt.

3) $a \in \mathbb{R}$ heißt **maximales** (**minimales**) Element von M oder **Maximum** (**Minimum**) von M, falls

$$a \in M \;\wedge\; \forall x \in M : x \leq a \quad (x \geq a)$$

Schreibe $a =: \max M$ bzw. $a =: \min M$.

4) $a \in \mathbb{R}$ heißt **Supremum** von M, falls a kleinste obere Schranke ist:

$$\big(\forall x \in M : x \leq a\big) \;\wedge\; \big((\forall x \in M : x \leq b) \;\Rightarrow\; b \geq a\big).$$

Analog: **Infimum** von M = größte untere Schranke. Schreibe $a =: \sup M$ bzw. $a =: \inf M$.

5) Falls M nicht nach oben beschränkt ist, besitzt M kein Supremum. Schreibe dann: $\sup M = \infty$. Genauso $\inf M = -\infty$, falls M nicht nach unten beschränkt.

Z.B.:
$$\begin{array}{lll} M_1 & = &]-\infty, 1] \\ M_2 & = &]2, \infty[\\ M_3 & = & [0, 3[\\ M_4 & = & \{\frac{1}{n} : n \in \mathbb{N}\} \;=\; \{1, \frac{1}{2}, \frac{1}{3}, \ldots\} \end{array}$$

2.37 Satz(Existenz Infimum/Supremum): Jede nach unten (oben) beschränkte Menge $\emptyset \neq M \subseteq \mathbb{R}$ besitzt ein Infimum (Supremum).

Beweis: Intervallschachtelung.

Für Infimum: Wähle a_0 untere Schranke, $b_0 \in M$. Setze $x_1 := \dfrac{a_0 + b_0}{2}$.

Entweder: x_1 ist untere Schranke. Dann $a_1 := x_1,\ b_1 := b_0$.

Oder: x_1 ist keine untere Schranke: $\exists y \in M : a_0 \le y < x_1$. Dann: $a_1 := a_0,\ b_1 := y$

Setze $x_2 := \dfrac{a_1 + b_1}{2}$.

$\vdots$

Dann ist $[a_n, b_n]$ Intervallschachtelung mit

$$\forall n \in \mathbb{N} : a_n \text{ ist untere Schranke } \wedge\ b_n \in M\ \wedge\ 0 \le b_n - a_n \le \frac{1}{2^n}(b_0 - a_0)$$

$\Rightarrow\ a := \bigcap\limits_{n=1}^{\infty} [a_n, b_n]$ ist Infimum. $\square$

2.38 Hauptsatz über monotone Folgen: Ist (a_n) monoton wachsend und nach oben beschränkt ($a_n \le K$ für $n \in \mathbb{N}$) oder monoton fallend und nach unten beschränkt, so ist (a_n) konvergent.

Beweis: Sei (a_n) monoton wachsend, $a_n \le K$. Setze $a := \sup\{a_n : n \in \mathbb{N}\}$.
Behauptung: $a = \lim\limits_{n\to\infty} a_n$.
Sei $\varepsilon > 0$ fest. $a - \varepsilon$ ist keine obere Schranke.
$\Rightarrow\ \exists N_\varepsilon \in \mathbb{N} : a_{N_\varepsilon} > a - \varepsilon$.

$$\Rightarrow\ \forall n > N_\varepsilon : \left.\begin{array}{l} a_n \ge a_{N_\varepsilon} > a - \varepsilon \\ a_n \le a \end{array}\right\} \Rightarrow\ |a - a_n| < \varepsilon.$$

$\square$

2.39 Existenz der Wurzel: Seien $a \in \mathbb{R}^+$ und $x_0 \in \mathbb{R}^+$. Definiere (x_n) rekursiv durch

$$x_{n+1} \ := \ \frac{1}{2}\left(x_n + \frac{a}{x_n}\right) \quad (n \in \mathbb{N}_0).$$

Dann existiert

$$b \ := \ \lim_{n\to\infty} x_n,$$

und es gilt $b^2 = a$ und $b \ge 0$, d.h. $b =: \sqrt{a}$ (Babylonisches Wurzelziehen, Heron-Verfahren).

Beweis: **1)** $x_n > 0$: Induktionsanfang: $x_0 > 0$

Induktionsschritt: $x_n > 0\ \Rightarrow\ x_{n+1} = \dfrac{1}{2}\left(x_n + \dfrac{a}{x_n}\right) > 0$

Induktionsschluss: $x_n > 0$ für alle $n \in \mathbb{N}_0$.

2) $\forall n \in \mathbb{N} : x_n^2 \geq a$:

$$\begin{aligned} x_n^2 - a &= \frac{1}{4}\left(x_{n-1} + \frac{a}{x_{n-1}}\right)^2 - a \\ &= \frac{1}{4}\left(x_{n-1}^2 + 2a + \left(\frac{a}{x_{n-1}}\right)^2 - 4a\right) \\ &= \frac{1}{4}\left(x_{n-1} - \frac{a}{x_{n-1}}\right)^2 \\ &\geq 0 \end{aligned}$$

3) (x_n) ist monoton fallend ab $n = 1$:

$$x_n - x_{n+1} = x_n - \frac{1}{2}\left(x_n + \frac{a}{x_n}\right) = \frac{1}{2}\left(x_n - \frac{a}{x_n}\right) = \frac{1}{2x_n}\left(x_n^2 - a\right) \overset{1),\ 2)}{\geq} 0$$

4) Wegen $x_1 \geq x_n > 0$ und der Monotonie konvergiert (x_n).

$$b := \lim_{n\to\infty} x_n \geq 0.$$

Es gilt $b > 0$: $\quad b^2 \overset{\text{siehe später}}{=} \lim_{n\to\infty} x_n^2 \overset{2)}{\geq} a > 0.$

5) $b^2 = a$:

$$\begin{aligned} b &= \lim_{n\to\infty} x_n = \lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} \frac{1}{2}\left(x_n + \frac{a}{x_n}\right) \overset{\text{siehe später}}{=} \frac{1}{2}\left(b + \frac{a}{b}\right) \\ \Rightarrow\quad 2b &= b + \frac{a}{b} \\ \Leftrightarrow\quad b - \frac{a}{b} &= 0 \\ \Leftrightarrow\quad b^2 - a &= 0 \end{aligned}$$

$\square$

2.40 Bemerkungen: **1)** Definiert man den **relativen Fehler** $d_n := \frac{x_n - \sqrt{a}}{x_n}$, dann zeigt eine kurze Rechnung

$$d_{n+1} \leq d_n^2.$$

Insbesondere verdoppelt sich die Zahl der richtigen Nachkommastellen bei jedem Schritt.

2) Wir wissen: $\sqrt{2} \in \mathbb{R} \setminus \mathbb{Q}$.

Analog: Sei $k \in \mathbb{N},\ k \geq 2$. Setze

$$x_{n+1} := \frac{1}{k}\left((k-1)x_n + \frac{a}{x_n^{k-1}}\right), \quad x_0 > 0.$$

Dann existiert $b := \lim_{n\to\infty} x_n$, und es gilt $b^k = a$ und $b \geq 0$. Schreibe $b =: \sqrt[k]{a}$.

2.41 Die eulersche Zahl* e*: Sei $a_n := \left(1+\frac{1}{n}\right)^n$.
Dann ist (a_n) monoton und beschränkt, also konvergent. Der Grenzwert heißt **eulersche Zahl**:

$$\mathrm{e} := \lim_{n\to\infty}\left(1+\frac{1}{n}\right)^n = 2,7182818\ldots$$

Man kann zeigen: e ist irrational. Neben π ist e eine der wichtigsten Zahlen in der Mathematik.

Monotonie:

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

Also: $a_n \geq a_{n-1}$.

Beschränktheit:

$$\begin{aligned}
a_n &\overset{\text{Binomi}}{=} \sum_{k=0}^{n}\binom{n}{k}\frac{1}{n^k} \\
&= \sum_{k=0}^{n}\frac{n(n-1)\cdots(n-k+1)}{n\cdot n\cdots n}\cdot\frac{1}{k!} \\
&\leq \sum_{k=0}^{n}\frac{1}{k!} \\
&= 1+\sum_{k=1}^{n}\frac{1}{k!} \\
&\overset{k!\geq 2^{k-1}}{\leq} 1+\sum_{k=1}^{n}\frac{1}{2^{k-1}} \\
&\overset{l=k-1}{=} 1+\sum_{l=0}^{n-1}\left(\frac{1}{2}\right)^n \\
&= 1+\frac{1-(\frac{1}{2})^n}{1-\frac{1}{2}} \\
&\leq 1+\frac{1}{\frac{1}{2}} \\
&= 3
\end{aligned}$$

Insbesondere folgt $\mathrm{e} \leq 3$.

2.4 Die komplexen Zahlen

2.42 Überblick:

$3+x=2$ hat keine Lösung in $\mathbb{N}$. $\quad \mathbb{N}\cup\{-1\} \xrightarrow{\text{Gruppe}} \mathbb{Z}$.

$3\cdot x=1$ hat keine Lösung in $\mathbb{Z}$. $\quad \mathbb{Z}\cup\{\frac{1}{3}\} \xrightarrow{\text{Körper}} \mathbb{Q}$

$x^2=2$ hat keine Lösung in $\mathbb{Q}$. $\quad \mathbb{Q}\cup\{\sqrt{2}\} \xrightarrow{\text{vollst. Körper}} \mathbb{R}$

$x^2=-1$ hat keine Lösung in $\mathbb{R}$. $\quad \mathbb{R}\cup\{\sqrt{-1}\} \xrightarrow{\text{Körper}} \mathbb{C}$

Aber: In $\mathbb{C}$ gibt es keine Anordnung.

Vieles in der Mathematik versteht man leichter in $\mathbb{C}$, z.B. Nullstellen von Polynomen, Matrizen, Potenzreihen, Lösungen von Differentialgleichungen.

2.4.1 Der Körper der komplexen Zahlen

2.43 Idee: Suche Körper $(\mathbb{C},+,\cdot)$ mit $\mathbb{R}\subseteq\mathbb{C}$ und einem Element $\mathrm{i}\in\mathbb{C}$ mit $\mathrm{i}^2=-1$.

Dann $x+\mathrm{i}\cdot y\in\mathbb{C}$ für $x,y\in\mathbb{R}$. Aus den Körperrechengesetzen folgen

$$\begin{aligned}(x_1+\mathrm{i}\,y_1)+(x_2+\mathrm{i}\,y_2) &= x_1+\mathrm{i}\,y_1+x_2+\mathrm{i}\,y_2\\ &= x_1+x_2+\mathrm{i}\,y_1+\mathrm{i}\,y_2\\ &= (x_1+x_2)+\mathrm{i}\,(y_1+y_2)\end{aligned}$$

und

$$\begin{aligned}(x_1+\mathrm{i}\,y_1)\cdot(x_2+\mathrm{i}\,y_2) &= x_1\,x_2+\mathrm{i}\,y_1\,x_2+x_1\mathrm{i}\,y_2+\mathrm{i}\,y_1\,\mathrm{i}\,y_2\\ &= x_1\,x_2+(\mathrm{i}\cdot\mathrm{i})\,y_1\,y_2+\mathrm{i}\,y_1\,x_2+\mathrm{i}\,x_1\,y_2\\ &= (x_1\,x_2-y_1\,y_2)+\mathrm{i}\,(y_1\,x_2+x_1\,y_2)\end{aligned}$$

2.44 Satz: $\mathbb{C}:=\mathbb{R}\times\mathbb{R}$ mit den Verknüpfungen

$$\begin{aligned}(x_1,y_1)+(x_2,y_2) &:= (x_1+x_2,y_1+y_2)\\ (x_1,y_1)\cdot(x_2,y_2) &:= (x_1x_2-y_1y_2,x_1y_2+y_1x_2)\end{aligned}$$

ist ein Körper mit

$$\begin{array}{lrcl}\text{Nullelement} & 0 &=& (0,0)\\ \text{Einselement} & 1 &=& (1,0)\\ & -(x,y) &=& (-x,-y)\\ & (x,y)^{-1} &=& \left(\dfrac{x}{x^2+y^2}\,,\,\dfrac{-y}{x^2+y^2}\right) \qquad \text{falls } (x,y)\neq(0,0).\end{array}$$

$(\mathbb{C},+,\cdot)$ heißt Körper der **komplexen Zahlen**. In $\mathbb{C}$ kann man also rechnen wie in jedem Körper.

Beweis: Durch Nachrechnen, z.B.

$$(x,y)\cdot\left(\frac{x}{x^2+y^2}\,,\,\frac{-y}{x^2+y^2}\right) \;=\; \left(\frac{x^2}{x^2+y^2}-\frac{-y^2}{x^2+y^2}\,,\,\frac{x\cdot(-y)}{x^2+y^2}+\frac{y\cdot x}{x^2+y^2}\right) \;=\; (1,0)$$

□

2.45 $\mathbb{R}$ ***als Teilmenge von*** $\mathbb{C}$***:*** Die Abbildung $f:\mathbb{R}\to\mathbb{C}: x\mapsto(x,0)$ ist injektiv, und es gilt

$$\begin{aligned} f(x)+f(y) &= (x,0)+(y,0) &&= (x+y,0) &&= f(x+y)\\ f(x)\cdot f(y) &= (x,0)\cdot(y,0) &&= (x\cdot y,0) &&= f(x\cdot y)\end{aligned}$$

Es ist egal, ob ich in $\mathbb{R}$ rechne und dann abbilde, oder erst abbilde und dann in $\mathbb{C}$ rechne.
Also: Identifiziere $x\in\mathbb{R}$ mit $(x,0)\in\mathbb{C}$ bzw. $\mathbb{R}=f(\mathbb{R})$. Dann sind die Verknüpfungen $+,\cdot:\mathbb{C}\times\mathbb{C}\to\mathbb{C}$ Fortsetzungen von $+,\cdot:\mathbb{R}\times\mathbb{R}\to\mathbb{R}$. Weiter gilt für $\alpha\in\mathbb{R},\ (x,y)\in\mathbb{C}$

$$\alpha\cdot(x,y) \;=\; (\alpha,0)\cdot(x,y) \;=\; (\alpha x,\alpha y) \;=\; (x,y)\cdot\alpha.$$

2.46 Vereinfachung: Sei $\mathrm{i}:=(0,1)$. Dann gilt

$$\mathrm{i}^2 \;=\; -1\;\in\mathbb{R}.$$

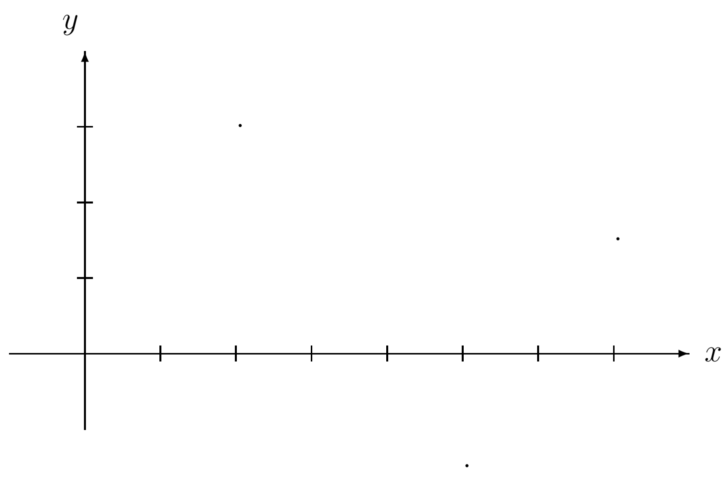
Die komplexe Zahl i heißt **imaginäre Einheit**. Für $(x,y)\in\mathbb{C}$ gilt

$$(x,y) \;=\; (x,0)+(0,y) \;=\; (x,0)+(0,1)\cdot(y,0) \;=\; x+\mathrm{i}\cdot y.$$

Fazit: Schreibe statt (x,y) auch $x+\mathrm{i}\,y$ (**Normalform** einer komplexen Zahl), rechne wie in $\mathbb{R}$, beachte $\mathrm{i}^2=-1$.

Z.B.: Berechne $\dfrac{4+3\mathrm{i}}{2-\mathrm{i}}=1+2\mathrm{i}$.

Veranschaulichung in der **Gauß'schen Zahlenebene**.



2.47 Definition: Sei $z = x + \mathrm{i}\, y \in \mathbb{C}$ mit $x, y \in \mathbb{R}$.

1) x heißt **Realteil** von z : $\quad x = \operatorname{Re} z$.

2) y heißt **Imaginärteil** von z : $\quad y = \operatorname{Im} z$.

3) $\overline{z} := x - \mathrm{i}\, y$ heißt die zu z **konjugierte** Zahl.

2.48 Satz: Für $z, z_1, z_2 \in \mathbb{C}$ gelten:

1) $z \mapsto \overline{z} \mathrel{\widehat{=}}$ Spiegelung an reeller Achse in der Gauß'schen Zahlenebene

2) $\overline{\overline{z}} \;=\; z$

3) $z \in \mathbb{R} \quad \Leftrightarrow \quad \overline{z} \;=\; z$

4) $\operatorname{Re} z \;=\; \dfrac{z + \overline{z}}{2}, \quad \operatorname{Im} z \;=\; \dfrac{z - \overline{z}}{2\mathrm{i}}$

5) $\overline{z_1 + z_2} \;=\; \overline{z}_1 + \overline{z}_2$

6) $\overline{z_1 \cdot z_2} \;=\; \overline{z}_1 \cdot \overline{z}_2, \quad \overline{\left(\dfrac{1}{z}\right)} \;=\; \dfrac{1}{\overline{z}}$

7) $z \;=\; x + \mathrm{i}\, y \;\Rightarrow\; z \cdot \overline{z} \;=\; x^2 + y^2 \;\in \mathbb{R}_0^+$

Z.B.: $z = 3 + 5\mathrm{i} \;\Rightarrow\; z \cdot \overline{z} = 34$.

2.49 Definition: Sei $z = x + \mathrm{i}\, y \in \mathbb{C}$, $x, y \in \mathbb{R}$. Dann heißt

$$|z| := \sqrt{z \cdot \overline{z}} = \sqrt{x^2 + y^2} \qquad \text{(siehe Satz 2.48)}$$

der **Betrag** von z. (Geometrisch: $|z|$ = Länge von 0 bis z in Gauß'scher Ebene.)

2.50 Hilfssatz: **1)** $|z| \;=\; |\overline{z}|$

2) $|\operatorname{Re} z| \;\le\; |z|, \quad |\operatorname{Im} z| \;\le\; |z|$

3) $z \;=\; x \in \mathbb{R} \;\Rightarrow\; |z|_{\mathbb{C}} \;=\; \sqrt{x^2+0} \;=\; |x|_{\mathbb{R}}.$
Also: Der Betrag in $\mathbb{C}$ verallgemeinert den Betrag in $\mathbb{R}$.

2.51 Satz: $(\mathbb{C}, |.|)$ ist ein bewerteter Körper, d.h. es gelten:

(B1) $|z| \geq 0; \qquad |z| = 0 \;\Leftrightarrow\; z = 0.$

(B2) $|z_1 \cdot z_2| \;=\; |z_1| \cdot |z_2|.$

(B3) $|z_1 + z_2| \;\leq\; |z_1| + |z_2| \quad$ (Δ-Ungleichung).

Insbesondere: $\left|\dfrac{z_1}{z_2}\right| \;=\; \dfrac{|z_1|}{|z_2|},\; |z_1 \pm z_2| \;\geq\; \big||z_1| - |z_2|\big|$ (vgl. Folgerungen 2.17).

Beweis: **(B1)** $|z| \geq 0$ okay

$$|z| \;=\; 0 \;\Leftrightarrow\; x^2 + y^2 \;=\; 0 \;\Leftrightarrow\; x \;=\; 0 \;\wedge\; y \;=\; 0 \;\Leftrightarrow\; z \;=\; 0$$

(B2) $|z_1\, z_2|^2 \;=\; (z_1\, z_2) \cdot \overline{(z_1\, z_2)} \;=\; z_1\, z_2\, \overline{z_1}\, \overline{z_2} \;=\; z_1\, \overline{z_1}\, z_2\, \overline{z_2} \;=\; |z_1|^2\, |z_2|^2$

(B3)
$$\begin{aligned}
|z_1 + z_2|^2 \;&=\; (z_1 + z_2)\, \overline{(z_1 + z_2)} \\
&=\; (z_1 + z_2)\, (\overline{z_1} + \overline{z_2}) \\
&=\; z_1\, \overline{z_1} + z_1\, \overline{z_2} + z_2\, \overline{z_1} + z_2\, \overline{z_2} \\
&=\; |z_1|^2 + |z_2|^2 + \underbrace{2\mathrm{Re}\, z_1\, \overline{z_2}}_{\leq 2|z_1\, \overline{z_2}| = 2|z_1|\, |\overline{z_2}| = 2|z_1|\, |z_2|} \\
&\leq\; \big(|z_1| + |z_2|\big)^2
\end{aligned}$$
$\square$

2.52 Bemerkung: Wegen $\mathrm{i}^2 = -1^2$ kann $\mathbb{C}$ nicht angeordnet werden.

2.4.2 Folgen in $\mathbb{C}$

Definition von Konvergenz: Siehe Definition 2.20

2.53 Hilfssatz Vergleich von Beträgen: Sei $z = x + \mathrm{i}\, y$. Dann gilt

$$|z| \;\leq\; |x| + |y| \;\leq\; \sqrt{2}\; |z|.$$

D.h., wenn $|z|$ klein ist, sind auch $|x|, |y|$ klein und umgekehrt.

Beweis: 1) $|z|^2 \;=\; x^2 + y^2 \;=\; |x|^2 + |y|^2 \;\leq\; |x|^2 + |y|^2 + 2|x|\,|y| \;=\; (|x| + |y|)^2$
$\Rightarrow\; |z| \;\leq\; |x| + |y|.$

2) Es gilt $0 \leq (|x|-|y|)^2 = |x|^2 - 2|x|\,|y| + |y|^2 \;\Rightarrow\; 2|x|\,|y| \;\leq\; |x|^2+|y|^2$.

3) $(|x|+|y|)^2 \;=\; |x|^2 + 2|x|\,|y| + |y|^2 \;\overset{\text{(ii)}}{\leq}\; 2(|x|^2+|y|^2) \;=\; 2|z|^2 \;\Rightarrow\; |x|+|y| \;\leq\; \sqrt{2}\,|z|.$ □

2.54 Satz: Seien $x_n, y_n, x, y \in \mathbb{R}$ und $z_n = x_n + \mathrm{i}\,y_n,\;\; z = x + \mathrm{i}\,y$. Dann gelten

1) $\lim\limits_{n\to\infty} z_n = z \;\Leftrightarrow\; \lim\limits_{n\to\infty} x_n = x \;\wedge\; \lim\limits_{n\to\infty} y_n = y.$

2) (z_n) ist Cauchy-Folge in $\mathbb{C} \;\Leftrightarrow\; (x_n), (y_n)$ sind Cauchy-Folgen in $\mathbb{R}$.

3) $(\mathbb{C}, |.|)$ ist vollständig.

Z.B.: $z_n \;=\; \underbrace{\frac{1}{n}}_{\to 0} + \mathrm{i}\,\underbrace{(2-\frac{1}{n^2})}_{\to 2} \;\to\; 2\mathrm{i}$

Z.B.: $z_n \;=\; \frac{1}{n} + \mathrm{i}\,n:$ $\quad (\operatorname{Im} z_n)_{n\in\mathbb{N}}$ ist divergent $\;\Rightarrow\; (z_n)$ ist divergent.

Z.B.: $z_n = (-1)^n - \frac{\mathrm{i}}{n}:$ $\quad (\operatorname{Re} z_n)_{n\in\mathbb{N}}$ ist divergent $\;\Rightarrow\; (z_n)$ ist divergent.

2.55 Rechenregeln für konvergente Folgen: Seien $(a_n), (b_n)$ konvergente Folgen in $\mathbb{C}$ oder in $\mathbb{R}$. Dann gelten:

1) $\exists M \in \mathbb{R}^+ \;\; \forall n \in \mathbb{N} : |a_n| \leq M$ $\qquad$ (Jede konvergente Folge ist beschränkt).

2) $(a_n + b_n)_{n\in\mathbb{N}}$ ist konvergent mit $\lim\limits_{n\to\infty}(a_n+b_n) \;=\; \lim\limits_{n\to\infty} a_n + \lim\limits_{n\to\infty} b_n.$
(endliche Summe und Grenzwert sind vertauschbar)

3) Für $\lambda \in \mathbb{C}$ ist $(\lambda\, a_n)$ konvergent mit $\lim\limits_{n\to\infty} \lambda\, a_n \;=\; \lambda \lim\limits_{n\to\infty} a_n.$

4) $(|a_n|)$ ist konvergent mit $\lim\limits_{n\to\infty} |a_n| \;=\; \left|\lim\limits_{n\to\infty} a_n\right|.$

5) $(a_n\, b_n)$ ist konvergent und $\lim\limits_{n\to\infty} a_n\, b_n \;=\; \left(\lim\limits_{n\to\infty} a_n\right) \cdot \lim\limits_{n\to\infty} b_n.$

6) Sei $\lim\limits_{n\to\infty} a_n \neq 0$. Dann existiert ein $N \in \mathbb{N}$ mit $a_n \neq 0$ für $n > N$. Setze

$$c_n \;:=\; \begin{cases} \dfrac{b_n}{a_n} & \text{für } n \geq N+1 \\ 0 & \text{für } n \leq N. \end{cases}$$

Dann ist (c_n) konvergent mit $\lim\limits_{n\to\infty} c_n \;=\; \dfrac{\lim\limits_{n\to\infty} b_n}{\lim\limits_{n\to\infty} a_n}.$

7) Sind $(a_n), (b_n)$ Folgen in $\mathbb{R}$ und $N \in \mathbb{N}$, so gelten weiter

a) $(\forall n \geq N : a_n \leq b_n) \ \Rightarrow\ \lim\limits_{n\to\infty} a_n \ \leq\ \lim\limits_{n\to\infty} b_n$.

b) $\lim\limits_{n\to\infty} a_n = \lim\limits_{n\to\infty} b_n \wedge (\forall n \geq N : a_n \leq c_n \leq b_n) \ \Rightarrow\ (c_n)$ konvergent $\wedge \lim\limits_{n\to\infty} c_n = \lim\limits_{n\to\infty} a_n$ (Sandwich-Satz, Prinzip der Polizisten).

Beweis: **Zu 1)** Sei $\varepsilon := 1 \ \Rightarrow\ \exists n_1 \in \mathbb{N} \ \ \forall n > n_1 : |a_n - a| < 1$

$\Rightarrow\ \forall n > n_1 : |a_n| = |a_n - a + a| \leq |a_n - a| + |a| < 1 + |a|$

Wähle

$$M \ := \ \max\big\{|a_1|, |a_2|, \ldots, |a_{n_1}|, 1 + |a|\big\}.$$

Zu 5) Vorüberlegung:

$$\begin{aligned} |a_n\, b_n - a\, b| &= |a_n\, b_n - a_n\, b + a_n\, b - a\, b| \\ &\leq |a_n\, b_n - a_n\, b| + |a_n\, b - a\, b| \\ &= |a_n|\, |b_n - b| + |b|\, |a_n - a| \end{aligned}$$

Also: Sei $M > 0$ mit $\quad \forall n \in \mathbb{N} : |a_n| \leq M, \quad$ sei $\varepsilon > 0$ fest.

Wähle n_1 mit $\forall n > n_1 : |b_n - b| < \dfrac{\varepsilon}{2M}$

Wähle n_2 mit $\forall n > n_2 : |a_n - a| < \dfrac{\varepsilon}{2|b|}$ (falls $b \neq 0$, sonst n_2 beliebig)

Wähle $n_\varepsilon := \max\{n_1, n_2\}$. Für $n > n_\varepsilon$ gilt

$$|a_n\, b_n - a\, b| \ \leq\ M\, \frac{\varepsilon}{2M} + |b|\, \frac{\varepsilon}{2|b|} \ = \ \varepsilon$$

Zu 7) $a_n \leq c_n \leq b_n \ \Rightarrow\ 0 \leq b_n - c_n \leq b_n - a_n$, also $|b_n - c_n| \leq |b_n - a_n|$

$$\begin{aligned} \Rightarrow\ |c_n - a| &= |c_n - b_n + b_n - a| \\ &\leq |c_n - b_n| + |b_n - a| \\ &\leq |b_n - a_n| + |b_n - a| \\ &\leq |b_n - a| + |a - a_n| + |b_n - a| \\ &\to 0 \end{aligned}$$

$\square$

Z.B.: $a_n \ = \ \dfrac{2^n + \mathrm{i}\, 7^n}{\mathrm{i}\, n - 2 \cdot 7^n}$

Z.B.: $a_n \ = \ \dfrac{n^2 + \mathrm{i}\, n^3}{1 + \mathrm{i}\, n^2}$

Z.B.: $\left(1+\frac{1}{n}\right)^n \;\le\; \sum_{k=0}^{n} \frac{1}{k!} \;\le\; 3$ (vgl. Beweis von 2.41)

$a_n \;:=\; \left(1+\frac{1}{n}\right)^n \;\to\; \mathrm{e}$

$b_n \;:=\; \sum_{k=0}^{n} \frac{1}{k!}$ ist monoton wachsend und beschränkt, also in $\mathbb{R}$ konvergent

$\Rightarrow \quad \mathrm{e} \;\le\; \lim\limits_{n\to\infty} b_n \;\le\; 3.$

Man kann zeigen: $\sum_{k=0}^{\infty} \frac{1}{k!} \;:=\; \lim\limits_{n\to\infty} b_n \;=\; \mathrm{e}.$

2.4.3 Polardarstellung komplexer Zahlen

2.56 Satz: Für jedes $z \in \mathbb{C}\setminus\{0\}$ existieren eindeutige Zahlen $r \in \mathbb{R}^+$, $\varphi \in [0, 2\pi[$, so dass

$$z \;=\; r(\cos\varphi + \mathrm{i}\sin\varphi). \qquad (*)$$

Man nennt $(*)$ die **Polardarstellung** von z. Es gilt

$$r = |z|,\; \cos\varphi = \frac{\operatorname{Re} z}{|z|},\; \sin\varphi = \frac{\operatorname{Im} z}{|z|}.$$

Der Winkel φ heißt **Argument** von z: $\quad \varphi = \arg(z)$.

2.57 Tabelle wichtiger Werte:

φ	0	$30^\circ \hat{=} \frac{\pi}{6}$	$45^\circ \hat{=} \frac{\pi}{4}$	$60^\circ \hat{=} \frac{\pi}{3}$	$90^\circ \hat{=} \frac{\pi}{2}$
$\sin\varphi$	$\frac{\sqrt{0}}{2}$	$\frac{\sqrt{1}}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{3}}{2}$	$\frac{\sqrt{4}}{2}$
$\cos\varphi$	$\frac{\sqrt{4}}{2}$	$\frac{\sqrt{3}}{2}$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{1}}{2}$	$\frac{\sqrt{0}}{2}$

Z.B.: $z_1 = 1 + \mathrm{i}$, $z_2 = \mathrm{i}$, $z_3 = -1$, $z_4 = -1 - \sqrt{3}\mathrm{i}$.

2.58 Multiplikation: $z_1 = r_1(\cos\varphi_1 + \mathrm{i}\,\sin\varphi_1)$, $\;z_2 = r_2(\cos\varphi_2 + \mathrm{i}\,\sin\varphi_2)$

$$\begin{aligned} \Rightarrow\; z_1\, z_2 \;&=\; r_1\, r_2\big(\cos\varphi_1\, \cos\varphi_2 - \sin\varphi_1\, \sin\varphi_2 + \mathrm{i}\,(\cos\varphi_1\, \sin\varphi_2 + \sin\varphi_1\, \cos\varphi_2)\big) \\ &=\; r_1\, r_2(\cos(\varphi_1+\varphi_2) + \mathrm{i}\, \sin(\varphi_1+\varphi_2)) \end{aligned}$$

wegen der Additionstheoreme

$$\begin{aligned}\cos(\varphi_1+\varphi_2) &= \cos\varphi_1\,\cos\varphi_2 - \sin\varphi_1\,\sin\varphi_2,\\ \sin(\varphi_1+\varphi_2) &= \cos\varphi_1\,\sin\varphi_2 + \sin\varphi_1\,\cos\varphi_2.\end{aligned}$$

Also: Komplexe Zahlen werden multipliziert, in dem man ihre Beträge multipliziert und die Winkel addiert.

2.59 Vereinfachung: Wir schreiben $\mathrm{e}^{\mathrm{i}\varphi} := \cos\varphi + \mathrm{i}\sin\varphi$ (φ im Bogenmaß). Damit wird die Polardarstellung zu

$$z = r\,\mathrm{e}^{\mathrm{i}\varphi}.$$

Additionstheoreme $\Rightarrow$ $\mathrm{e}^{\mathrm{i}\,\varphi_1}\cdot\mathrm{e}^{\mathrm{i}\,\varphi_2} = \mathrm{e}^{\mathrm{i}(\varphi_1+\varphi_2)}$.
Multiplikation komplexer Zahlen: $r_1\,\mathrm{e}^{\mathrm{i}\varphi_1}\cdot r_2\,\mathrm{e}^{\mathrm{i}\varphi_2} = r_1 r_2\,\mathrm{e}^{\mathrm{i}(\varphi_1+\varphi_2)}$.
Wichtig: Für $\varphi\in\mathbb{R}$ gilt $\left|\mathrm{e}^{\mathrm{i}\varphi}\right| = \sqrt{\cos^2\varphi+\sin^2\varphi} = 1$.

Z.B.: $(1+\mathrm{i})^n$.

Z.B.: Bestimme alle Lösungen von $z^3 = -8$.

2.60 Die n-te Wurzel: Es seien $r\in\,]0,\infty[$, $\varphi\in[0,2\pi[$ und $n\in\mathbb{N}$ gegeben. Die Gleichung

$$z^n = r\,\mathrm{e}^{\mathrm{i}\varphi} = r(\cos\varphi + \mathrm{i}\,\sin\varphi)$$

hat genau n verschiedene Lösungen

$$z_k = r^{1/n}\,\mathrm{e}^{\mathrm{i}\,\frac{\varphi+2k\pi}{n}} = r^{1/n}\left(\cos\frac{\varphi+2k\pi}{n} + \mathrm{i}\,\sin\frac{\varphi+2k\pi}{n}\right), \qquad k = 0,1,\ldots,n-1.$$

D.h. die Gleichung $z^n = a$ hat für $a\neq 0$ genau n verschiedene Lösungen.
Spezialfall $n=2$: Für $a\in\mathbb{C}\setminus[0,\infty[$ bezeichnen wir mit $\sqrt{a}$ eine (egal welche) Lösung z von $z^2 = a$.

Z.B.: $\sqrt{8\mathrm{i}}$ bezeichnet $(2+2\mathrm{i})$ oder $-2-2\mathrm{i}$.

2.61 Quadratische Gleichungen: Die Gleichung $az^2 + bz + c \;=\; 0 \quad (a, b, c \in \mathbb{C},\; a \neq 0)$ hat die Lösungen

$$z_{1/2} \;=\; \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}\,.$$

Beweis:

$$\begin{aligned}
& a\,z^2 + b\,z + c \;=\; 0 \\
\Leftrightarrow\quad & z^2 + \frac{b}{a}\,z + \frac{c}{a} \;=\; 0 \\
\Leftrightarrow\quad & z^2 + 2\frac{b}{2a}\,z + \left(\frac{b}{2a}\right)^2 \;=\; \left(\frac{b}{2a}\right)^2 - \frac{c}{a} \\
\Leftrightarrow\quad & \left(z + \frac{b}{2a}\right)^2 \;=\; \frac{1}{4a^2}\,\left(b^2 - 4ac\right) \\
\Leftrightarrow\quad & z + \frac{b}{2a} \;=\; \pm\frac{1}{2a}\,\sqrt{b^2 - 4ac} \\
\Leftrightarrow\quad & z \;=\; -\frac{b}{2a} \pm \frac{1}{2a}\,\sqrt{b^2 - 4ac}
\end{aligned}$$

$\square$

2.4.4 Polynome

2.62 Definition: **1)** Ein **Polynom** ist eine Funktion

$$P : \mathbb{C} \to \mathbb{C} : z \mapsto \sum_{k=0}^{n} a_k\, z^k \;=\; a_0 + a_1\, z + \ldots + a_n\, z^n$$

mit den **Koeffizienten** $a_k \in \mathbb{C}$. Falls $a_0, a_1, \ldots, a_n \in \mathbb{R}$, heißt das Polynom **reell** (dann betrachtet man auch $P : \mathbb{R} \to \mathbb{R}$). Ist $a_n \neq 0$, so heißt n der **Grad** des Polynoms: $n = \operatorname{Grad} P$. Für das Nullpolynom: $\operatorname{Grad} 0 := -1$.

2) Eine **rationale Funktion** ist eine Funktion $z \mapsto \dfrac{P(z)}{Q(z)}$ mit geeignetem Definitionsbereich, wobei P, Q Polynome sind.

3) Die Menge der komplexen/reellen Polynome: $\mathbb{C}[x],\; \mathbb{R}[x]$.

2.63 Berechnung: $P(z) = (\ldots((a_n\, z + a_{n-1})\, z + a_{n-2})\, z + \ldots + a_1)\, z + a_0$. Hierfür werden nur n Multiplikationen benötigt statt $\frac{n(n+1)}{2}$ (**Hornerschema**).

Z.B.: $P(z) \;=\; 4z^3 - 3z^2 + 2z - 1 \;=\; \big((4z-3)z + 2\big)z - 1$

Berechnung von $P(2)$:

a_k	4	-3	2	-1
$z = 2$		8	10	24
	4	5	12	23 $= P(2)$

Klar ist: Sind $P_1, P_2 \neq 0$ Polynome, dann ist $P_1 \cdot P_2$ Polynom mit

$$\operatorname{Grad}(P_1 \cdot P_2) \;=\; (\operatorname{Grad} P_1) + \operatorname{Grad} P_2.$$

Philosophie: Rechnen mit Polynomen analog Rechnen mit ganzen Zahlen. Rechnen mit rationalen Funktionen analog Rechnen mit rationalen Zahlen.

2.64 Division mit Rest: Seien P, Q Polynome, $1 \leq \operatorname{Grad} Q \leq \operatorname{Grad} P$. Dann gibt es eindeutig bestimmte Polynome P_1 und R, so dass

$$P \;=\; Q \cdot P_1 + R \;\wedge\; \operatorname{Grad} R < \operatorname{Grad} Q.$$

Z.B.:
$$\begin{array}{rrrrl}
(4z^4 & -14z^3 & +6z^2 & +3z-5) & : \underbrace{(2z^3+3z+1)}_{Q} \;=\; \underbrace{2z-7}_{P_1} + \dfrac{\overbrace{22z+2}^{R}}{2z^3+3z+1} \\
4z^4 & & +6z^2 & +2z & \\ \hline
 & -14z^3 & & +z-5 & \\
 & -14z^3 & & -21z-7 & \\ \hline
 & & & 22z+2 &
\end{array}$$

2.65 Satz: Sei $P \in \mathbb{C}[x]$, Grad $P \geq 1$, $\lambda \in \mathbb{C}$. Dann gilt:

$$P(\lambda) = 0 \;\Leftrightarrow\; (z-\lambda) \text{ ist Teiler von } P, \text{ d.h. } P(z) = (z-\lambda)P_1(z).$$

Beweis: “$\Leftarrow$” klar.

“$\Rightarrow$”: Es gilt $P(z) \;=\; (z-\lambda)P_1(z) + R(z)$, $\quad \operatorname{Grad} R < 1$, also $R =$ konstant.
$0 \;=\; P(\lambda) \;=\; 0 \cdot P_1(\lambda) + R(\lambda) \;\Rightarrow\; R \;=\; 0$ $\square$

2.66 Folgerung: Ein Polynom vom Grad $n \geq 1$ hat höchstens n verschiedene Nullstellen.

Beweis: Annahme: $P(\lambda_j) = 0$, $j = 0, \ldots, n+1$, alle $\lambda_1, \ldots, \lambda_{n+1}$ verschieden.
$\Rightarrow\; P(z) \;=\; (z-\lambda_1)\cdots(z-\lambda_{n+1}) \cdot Q(z) \;=$ Polynom vom Grad $> n$ $\lightning$ $\square$

2.67 Identitätssatz: Sind

$$P(z) \;=\; \sum_{k=0}^{n} a_k z^k, \quad Q(z) \;=\sum_{k=0}^{n} b_k z^k$$

Polynome vom Grad $\leq n$, die an mindestens $n+1$ verschiedenen Stellen übereinstimmen, so gilt $P = Q$, d. h. $a_k = b_k$ für $k = 0, 1, \ldots, n$.

Beweis: $P - Q$ hat mindestens $n+1$ Nullstellen und $\operatorname{Grad}(P-Q) \leq n \;\Rightarrow\; P - Q = 0$. $\square$

2.68 Definition: $\lambda \in \mathbb{C}$ heißt **k-fache Nullstelle** von P oder Nullstelle mit **Vielfachheit k**, falls $(z-\lambda)^k$, aber nicht $(z-\lambda)^{k+1}$ Teiler von P ist. D. h. $P(z) = (z-\lambda)^k P_1(z)$ mit $P_1(\lambda) \neq 0$.

2.69 Fundamentalsatz der Algebra (C. F. Gauß 1799): Sei $P \in \mathbb{C}[x]$ Polynom vom Grad $n \geq 1$:

$$P(z) = a_n z^n + a_{n-1} z^{n-1} + \ldots + a_0 \qquad \text{mit } a_n \neq 0.$$

Dann gelten

1) P hat in $\mathbb{C}$ mindestens eine Nullstelle.

2) Es gibt komplexe Zahlen $\lambda_1, \ldots, \lambda_n$ (nicht notwendig verschieden), so dass

$$P(z) \;=\; a_n(z-\lambda_1)\cdots(z-\lambda_n).$$

Die λ_j sind bis auf Nummerierung eindeutig. Zählt man jede Nullstelle mit ihrer Vielfachheit, so hat P genau $\operatorname{Grad} P$ Nullstellen.

2.70 Hornerschema: Seien $P(z) = a_n z^n + a_{n-1} z^{n-1} + \ldots + a_0$ und $z_0 \in \mathbb{C}$ gegeben.

a_k	a_n	a_{n-1}	a_{n-2}	$\ldots$	a_0
$z = z_0$		$z_0 a_n$	$z_0 b_{n-1}$	$\ldots$	$z_0 b_1$
	a_n	$a_{n-1} + z_0 a_n$ $=: b_{n-1}$	$a_{n-2} + z_0 b_{n-1}$ $=: b_{n-2}$	$\ldots$	$a_0 + z_0 b_1$ $=: b_0 = P(z_0)$

Dann kann in der letzten Zeile das Ergebnis der Polynomdivision $P(z) : (z - z_0)$ abgelesen werden:

$$P(z) : (z - z_0) \;=\; a_n\, z^{n-1} + b_{n-1}\, z^{n-2} + \ldots + b_2\, z + b_1 + \frac{b_0}{z - z_0}$$

Z.B.: $P(z) \;=\; z^4 - 5z^2 - 2z, \qquad z_0 = -2$

a_k	1	0	-5	-2	0
$z_0 = -2$		-2	4	2	0
	1	-2	-1	0	0

$\Rightarrow\; P(z_0) = 0, \qquad P(z) : (z+2) = z^3 - 2z^2 - z.$

2.71 Reelle Polynome: Ist P reelles Polynom und $\lambda \in \mathbb{C}$ Nullstelle, dann ist auch $\overline{\lambda}$ Nullstelle.

Also: Eine Nullstelle ist entweder reell oder ist $\lambda \in \mathbb{C} \setminus \mathbb{R}$ Nullstelle, dann auch $\overline{\lambda}$.

Fall 1: $\lambda \in \mathbb{R}$: $\quad P(z) = (z - \lambda) P_1(z) \quad$ (P_1 reelles Polynom)

Fall 2: $\lambda \in \mathbb{C} \setminus \mathbb{R}$:
$$\begin{aligned} P(z) &= (z-\lambda)(z-\overline{\lambda})P_1(z) \\ &= (z^2 - (2\operatorname{Re}\lambda)z + |\lambda|^2)P_1(z) \qquad (P_1 \text{ reelles Polynom}) \end{aligned}$$

2.72 Folgerung: Ein reelles Polynom ist darstellbar als Produkt von linearen und quadratischen reellen Polynomen, wobei die quadratischen Polynome keine reellen Nullstellen besitzen (d.h. sie sind in $\mathbb{R}$ irreduzibel).

Z.B.: $P(z) = z^4 + 4z^3 + 2z^2 + 4z + 1$ hat Nullstelle $z = \mathrm{i}$

$\Rightarrow$ weitere Nullstelle $z = -\mathrm{i}$

$(z - \mathrm{i})(z + \mathrm{i}) = z^2 + 1$

$(z^4 + 4z^3 + 2z^2 + 4z + 1) : (z^2 + 1) = z^2 + 4z + 1$

$z^2 + 4z + 1 = 0 \;\Leftrightarrow\; z_{1/2} = \dfrac{-4 \pm \sqrt{16 - 4}}{2} = -2 \pm \sqrt{3}$

$\Rightarrow\; P(z) = (z^2 + 1) \cdot (z - (-2 + \sqrt{3}))(z - (-2 - \sqrt{3}))$

2.73 Rationale Nullstellen: Sei $P(x) = \sum\limits_{k=0}^{n} a_k\, x^k$ mit $a_0, \ldots, a_n \in \mathbb{Z}$. Ist $x = \dfrac{p}{q}$ Nullstelle von P, wobei $p, q \in \mathbb{Z}$ teilerfremd sind, dann ist p Teiler von a_0 und q Teiler von a_n.

Beweis: $0 \;=\; a_n \left(\dfrac{p}{q}\right)^n + a_{n-1} \left(\dfrac{p}{q}\right)^{n-1} + \ldots + a_0$

$\Leftrightarrow\;\; 0 \;=\; a_n\, p^n + a_{n-1}\, p^{n-1}\, q + \ldots + a_0\, q^n$

$\Leftrightarrow\;\; a_n\, p^n \;=\; -q(\ldots)$

$\Rightarrow\;\; q \mid a_n\, p^n \overset{p,q \text{ teilerfremd}}{\Rightarrow} q \mid a_n$

Genauso: $q^n\, a_0 = -p(\ldots) \;\Rightarrow\; p \mid q^n\, a_0 \;\Rightarrow\; p \mid a_0$ $\square$

Z.B.: $P(x) \;=\; 1 \cdot x^3 - 2x^2 - 6x + 4.$

Wegen $a_n = a_3 = 1$ sind alle rationalen Nullstellen ganze Zahlen.

Wegen $a_0 = 4$ kommen nur $x = \pm 1, \pm 2, \pm 4$ in Frage.

Probieren: $P(-2) = 0$.

Z.B.: $P(x) \;=\; 12x^4 - 4x^3 + 6x^2 + x - 1.$

Teiler von $a_n = a_4 = 12 :\quad 1, 2, 3, 4, 6, 12.$

Teiler von $a_0 = -1 :\quad 1$

$\Rightarrow$ Mögliche rationale Nullstellen $x = \pm 1, \pm\frac{1}{2}, \pm\frac{1}{3}, \pm\frac{1}{4}, \pm\frac{1}{6}, \pm\frac{1}{12}$.

Probieren: $P(\frac{1}{3}) = 0$.

3 Lineare Algebra

3.1 Linearität

3.1 Definition: Ein **Vektorraum** oder **linearer Raum** über einem Körper $\mathbb{K}$ (z.B. $\mathbb{K} = \mathbb{R}$, $\mathbb{K} = \mathbb{C}$, $\mathbb{K} = \mathbb{F}_p$) ist eine abelsche (d.h. kommutative) Gruppe $(V, +)$, versehen mit einer **Skalarenmultiplikation** $\cdot : \mathbb{K} \times V \to V : (\lambda, v) \mapsto \lambda \cdot v$, so dass

(S1) $(\lambda + \mu) \cdot v = \lambda \cdot v + \mu \cdot v$

(S2) $\lambda \cdot (v + w) = \lambda \cdot v + \lambda \cdot w$

(S3) $\lambda \cdot (\mu \cdot v) = (\lambda \, \mu) \cdot v$

(S4) $1 \cdot v = v$

Die Elemente des Vektorraumes heißen **Vektoren**. Das neutrale Element der Gruppe $(V, +)$ heißt auch **Nullvektor** 0.

Beispiele: **1)** $V_1 = \mathbb{R}^n = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} : x_1, \ldots, x_n \in \mathbb{R} \right\}$ ist ein Vektorraum über $\mathbb{R}$ mit

$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}, \qquad \lambda \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \lambda \, x_1 \\ \vdots \\ \lambda \, x_n \end{pmatrix}$$

2) $V_1 = \mathbb{C}^n = \left\{ \begin{pmatrix} z_1 \\ \vdots \\ z_n \end{pmatrix} : z_1, \ldots, z_n \in \mathbb{C} \right\}$ ist ein Vektorraum über $\mathbb{C}$ oder über $\mathbb{R}$ mit Definition von $+, \cdot$ wie in 1).

3) Vektorraum der Folgen:

$$V_3 := \{(a_n)_{n \in \mathbb{N}} \mid \forall n \in \mathbb{N} : a_n \in \mathbb{C}\}$$

ist Vektorraum über $\mathbb{C}$ (oder über $\mathbb{R}$) mit

$$(a_n) + (b_n) := (a_n + b_n), \quad \lambda \cdot (a_n) := (\lambda \, a_n).$$

4) Vektorraum der konvergenten Folgen:

$$V_4 := \mathcal{C} := \{(a_n)_{n \in \mathbb{N}} \mid \forall n \in \mathbb{N} : a_n \in \mathbb{C} \ \wedge \ \exists a \in \mathbb{C} : a_n \to a\}$$

ist Vektorraum über $\mathbb{C}$ (oder über $\mathbb{R}$) mit Definition von $+, \cdot$ wie in 3).

5) $V_5 := \mathcal{F} := \{f : \mathbb{C} \to \mathbb{C}\}$: Vektorraum der komplexen Funktionen mit

$$f + g : x \mapsto f(x) + g(x), \quad \lambda \cdot f : z \mapsto \lambda \, f(z)$$

6) $V_6 := \mathcal{P} := \{P : \mathbb{C} \to \mathbb{C} \mid P \text{ ist Polynom}\}$ mit $+, \cdot$ wie in 5) ist Vektorraum der Polynome, Untervektorraum von $\mathcal{F}$.

7) $V_7 := \mathcal{P}_n := \{P : \mathbb{C} \to \mathbb{C} \mid P \text{ ist Polynom vom Grad } \leq n\}$ mit $+, \cdot$ wie in 5) ist Untervektorraum von $\mathcal{F}$ und von $\mathcal{P}$.

8) $V_8 := \{f : \mathbb{C} \to \mathbb{C} : z \mapsto \lambda z^n \mid \lambda \in \mathbb{C}\}$ ist Untervektorraum von $\mathcal{F}$, $\mathcal{P}$ und von $\mathcal{P}_n$. „Gerade im Raum der Polynome“

9) $(\mathbb{K}, +, \cdot)$ ist ein Vektorraum über $\mathbb{K}$.

3.2 Rechenregeln: Sei $(V, +, \cdot)$ ein Vektorraum. Dann gelten:

1) $\lambda \cdot v = 0 \;\Leftrightarrow\; \lambda = 0 \;\vee\; v = 0$.

2) $(-1) \cdot v \;=\; -v$ (inverses Element zu v).

3) Seien $u, v \in V$ gegeben. Die Gleichung $u + x = v$ besitzt die eindeutige Lösung $x = v + (-u) =: v - u$.

3.3 Definition: Sei $(V, +, \cdot)$ ein Vektorraum über $\mathbb{K}$. Eine Teilmenge $U \subseteq V$ heißt **Untervektorraum** oder **linearer Teilraum** von V, falls $(U, +, \cdot)$ ein Vektorraum über $\mathbb{K}$ ist (hier bezeichnen $+, \cdot$ die Einschränkungen der V-Operationen auf $U \times U$ bzw. $\mathbb{K} \times U$).

3.4 Untervektorraumkriterium: Es seien $(V, +, \cdot)$ ein Vektorraum über $\mathbb{K}$ und $U \subseteq V$. Dann sind äquivalent:

(i) U ist Untervektorraum von V.

(ii) $U \neq \emptyset \;\wedge\; (\forall v, w \in U \;\; \forall \lambda, \mu \in \mathbb{K} : \lambda \cdot v + \mu \cdot w \in U)$.

(iii) $U \neq \emptyset \;\wedge\; (\forall v, w \in U : v + w \in U) \;\wedge\; (\forall v \in U \;\; \forall \lambda \in \mathbb{K} : \lambda \cdot v \in U)$.

Beispiele: **1)** $V = \mathbb{R}^3$:

$$U_1 := \left\{ \begin{pmatrix} x_1 \\ x_2 \\ 0 \end{pmatrix} : x_1, x_2 \in \mathbb{R} \right\}$$

$$U_2 := \left\{ \begin{pmatrix} x_1 \\ x_2 \\ 1 \end{pmatrix} : x_1, x_2 \in \mathbb{R} \right\}$$

$$U_3 := \left\{ \begin{pmatrix} x_1 \\ 0 \\ 0 \end{pmatrix} : x_1 \in \mathbb{R} \right\} = \left\{ \lambda \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} : \lambda \in \mathbb{R} \right\}$$

$$U_4 := \left\{ \begin{pmatrix} x_1 \\ 0 \\ 2x_1 \end{pmatrix} : x_1 \in \mathbb{R} \right\} = \left\{ \lambda \cdot \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} : \lambda \in \mathbb{R} \right\}$$

2) $V = \{(a_n)_{n\in\mathbb{N}} \mid \forall n \in \mathbb{N} : a_n \in \mathbb{C}\}$.

$U_1 := \{(a_n)_{n\in\mathbb{N}} \mid a_n \to 0\}$.

$U_2 := \{(a_n)_{n\in\mathbb{N}} \mid a_n \to 1\}$.

3) $\mathcal{P}_n$ ist Untervektorraum von $\mathcal{P}$.
Sei $P_j : \mathbb{C} \to \mathbb{C} : z \mapsto z^j \ (j = 0, \dots, n)$

$$\Rightarrow \mathcal{P}_n = \left\{\sum_{j=0}^{n} \lambda_j \cdot P_j : \lambda_j \in \mathbb{C}\right\}.$$

3.5 Definition: **1)** Seien $n \in \mathbb{N}$, $\lambda_1, \dots, \lambda_n \in \mathbb{K}$, $v_1, \dots, v_n \in V$. Dann heißt

$$v := \sum_{j=1}^{n} \lambda_j \cdot v_j$$

Linearkombination der Vektoren $v_1, \dots, v_n$.
Achtung: Eine Linearkombination ist immer eine endliche Summe.

2) Sei $M \subseteq V$. Dann heißt

$$\begin{aligned} \mathrm{LH}(M) &:= \{\text{Linearkombinationen von Elementen aus } M\} \\ &= \left\{\sum_{j=1}^{n} \lambda_j \cdot v_j : n \in \mathbb{N},\ \lambda_j \in \mathbb{K},\ v_j \in M\right\} \end{aligned}$$

die **lineare Hülle** von M. Offensichtlich gelten:

- $\mathrm{LH}(M)$ ist ein Untervektorraum von V.
- $\mathrm{LH}(M)$ ist der kleinste Unterraum von V, der M enthält:

$$M \subseteq W \ \wedge \ W \text{ Unterraum von } V \ \Rightarrow \ \mathrm{LH}(M) \subseteq W.$$

Man sagt: M **spannt** den Untervektorraum $\mathrm{LH}(M)$ auf.

Beispiele: **1)** $V = \mathbb{R}^3$

$M_1 := \left\{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}\right\}$

$M_2 := \left\{\lambda \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} : \lambda \in \mathbb{R}\right\}$

$M_3 := \left\{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 2 \end{pmatrix}\right\}$

2) $V = \{(a_n) \text{ Folge in } \mathbb{C}\}$, $M := \{e_j \in V \mid j \in \mathbb{N} \ \wedge \ e_j = (0, 0, \ldots, 0, 1, 0, 0, \ldots)\}$
$\uparrow$ j-tes Folgenglied

3) $V = \mathcal{P}, \ M := \{P_j : z \mapsto z^j \mid j = 0, 1, 2, \ldots\}$.

3.6 Minimale aufspannende Mengen: Sei $M \subseteq V$ mit $\mathrm{LH}(M) = V$. Dann sind äquivalent:

(i) $\forall m \in M : \mathrm{LH}\big(M \setminus \{m\}\big) \neq V$.

(ii) Für jede endliche Teilmenge $\{m_1, \ldots, m_n\} \subseteq M$ gilt

$$\forall \lambda_1, \ldots \lambda_n \in \mathbb{K} : \left(\sum_{j=1}^{n} \lambda_j \cdot m_j = 0 \ \Rightarrow \ \lambda_1 = \ldots = \lambda_n = 0 \right).$$

(iii) Ist $v = \sum\limits_{j=1}^{n} \lambda_j \cdot m_j$ mit $m_1, \ldots, m_n \in M$, so sind die Koeffizienten (Koordinaten) $\lambda_1, \ldots, \lambda_n$ eindeutig.

3.7 Definition: Seien V, W Vektorräume über $\mathbb{K}$. Eine Abbildung $L : V \to W$ heißt **linear** oder **Vektorraum-Homomorphismus**, falls

(i) $\big(\forall u, v \in V : L(u + v) = L(u) + L(v)\big) \ \wedge \ \big(\forall u \in V \ \ \forall \lambda \in \mathbb{K} : L(\lambda \cdot v) = \lambda \cdot L(v)\big)$

oder äquivalent

(ii) $\forall u, v \in V \ \ \forall \lambda, \mu \in \mathbb{K} : L(\lambda \cdot u + \mu \cdot v) = \lambda \cdot L(u) + \mu \cdot L(v)$.

Beispiele: **1)** $L : \mathbb{R}^3 \to \mathbb{R} : \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mapsto x_1$.

2) $L : \mathbb{R}^2 \to \mathbb{R}^2 : \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} a_1 x_1 + a_2 x_2 \\ b_1 x_1 + b_2 x_2 \end{pmatrix}$.

3) $L : \mathcal{F} \to \mathbb{C} : f \mapsto f(1)$.

4) $L : \mathcal{P} \to \mathcal{P} : P \mapsto P'$.

5) $V = \mathcal{C} = \{(a_n) \text{ Folge in } \mathbb{C} \mid (a_n) \text{ konvergent}\}$, $L : \mathcal{C} \to \mathbb{C} : (a_n) \mapsto \lim\limits_{n \to \infty} a_n$.

3.8 Satz: Seien U, V, W Vektorräume über $\mathbb{K}$. Sind $K : U \to V$ und $L : V \to W$ linear, dann ist auch $L \circ K : U \to W$ linear.

3.9 Definition: Sei $L : V \to W$ linear. Dann heißt

$$\mathrm{Bild}\,(L) \;:=\; \{L(x) : x \in V\}$$

das **Bild** von L und das Urbild der Menge $\{0\} \subseteq W$

$$\mathrm{Kern}\,(L) \;:=\; \{x \in V : L(x) = 0\} \;=\; L^{-1}(\{0\})$$

der **Kern** von L.

Beispiele: **1)** $L : \mathbb{R}^2 \mapsto \mathbb{R}^2 : \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} x_1 \\ 0 \end{pmatrix}$.

2) $L : \mathcal{F} \to \mathbb{C} : f \mapsto f(1)$.

3) $L : \mathcal{P}_n \to \mathcal{P}_n : P \mapsto P'$.

4) $L : \mathcal{C} \to \mathbb{C} : (a_n) \mapsto \lim\limits_{n\to\infty} a_n$.

3.10 Eigenschaften linearer Abbildungen: Sei $L : V \to W$ linear. Dann gelten:

1) $L(0) = 0$.

2) $\mathrm{Bild}\,(L)$ ist ein Untervektorraum von W.

3) $\mathrm{Kern}\,(L)$ ist ein Untervektorraum von V.

4) Ist zusätzlich L bijektiv, so ist $L^{-1} : W \to V$ linear.

3.11 Definition: Sei $L : V \to W$ linear und bijektiv. Dann heißt L **Isomorphismus**, die Vektorräume V und W heißen **isomorph**. Man kann dann V und W identifizieren, sie sind nur verschiedene Realisierungen desselben Vektorraumes.

Beispiele: **1)** $L : \mathbb{C}^{n+1} \to \mathcal{P}_n : \begin{pmatrix} a_1 \\ \vdots \\ a_{n+1} \end{pmatrix} \mapsto \big(P : z \mapsto a_1 + a_2\, z + \ldots + a_{n+1}\, z^n\big)$.

2) $L : \mathcal{P} \to \{\text{abbrechende Folgen}\}$:

$$\big(P : z \mapsto a_0 + a_1\, z + \ldots + a_n\, z^n\big) \mapsto (a_0, a_1, \ldots, a_n, 0, 0, \ldots).$$

Anwendung: In einem endlichen Körper $\mathbb{K}$ mit n Elementen gibt es nur n^n verschiedene Abbildungen $f : \mathbb{K} \to \mathbb{K}$. Der Raum der Polynome soll aber unendlich viele Elemente enthalten. Deshalb definiert man den Vektorraum der Polynome $\mathcal{P}$ nicht als Raum von Abbildungen sondern durch

$$\mathcal{P} \;:=\; \{\text{abbrechende Folgen in } \mathbb{K}\}.$$

Im Fall von $\mathbb{K} = \mathbb{R}$ ist diese Definition äquivalent zu unserer ursprünglichen Definition.

3.12 Kern und Injektivität: Für $L : V \to W$ linear sind äquivalent:

(i) $\mathrm{Kern}\,(L) = \{0\}$.

(ii) L ist injektiv.

Beweis: **(i) $\Rightarrow$ (ii):**
$$\begin{aligned} Lu = Lv \quad &\Leftrightarrow \quad Lu - Lv = 0 \\ &\Leftrightarrow \quad L(u-v) = 0 \\ &\Leftrightarrow \quad u - v \in \mathrm{Kern}\,(L) \\ &\Leftrightarrow \quad u - v = 0 \end{aligned}$$

(ii) $\Rightarrow$ (i): $L(0) = 0 \ \wedge\ L$ injektiv $\ \Rightarrow\ L^{-1}(\{0\}) = \{0\}$. □

3.2 Lineare Gleichungssysteme - Vorläufiges

3.13 Definition: Seien $m, n \in \mathbb{N}$ und $a_{jk}, b_j \in \mathbb{K}$ für $j = 1, \ldots, m,\ k = 1, \ldots, n$ gegeben. Das System

$$\left.\begin{array}{ccccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \ldots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & \ldots & & & + & a_{2n}x_n & = & b_2 \\ \vdots & & & & & & & & \\ a_{m1}x_1 & + & \ldots & & & + & a_{mn}x_n & = & b_m \end{array}\right\} \qquad (*)$$

für die n Unbekannten $x_1, \ldots, x_n \in \mathbb{K}$ heißt **lineares Gleichungssystem** (**LGS**). Falls die rechte Seite verschwindet (d.h. $b_1 = \ldots = b_m = 0$) heißt das LGS **homogen**, sonst **inhomogen**. Ist $(*)$ inhomogen, so heißt das LGS mit denselben Koeffizienten a_{jk}, aber $b_1 = \ldots = b_m = 0$, das **zugehörige homogene** LGS.

Kürzere Schreibweise für $(*)$:

$$\sum_{k=1}^{n} a_{jk}\, x_k \;=\; b_j, \quad j = 1, \ldots, m$$

Praktische Schreibweise für kleinere LGS:

$$\left(\begin{array}{cccc|c} a_{11} & a_{12} & \ldots & a_{1n} & b_1 \\ \vdots & & & \vdots & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} & b_m \end{array}\right)$$

3.14 Bemerkung: Die Abbildung

$$L : \mathbb{K}^n \to \mathbb{K}^n : \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \begin{pmatrix} \sum_{k=1}^{n} a_{1k}\, x_k \\ \vdots \\ \sum_{k=1}^{n} a_{mk}\, x_k \end{pmatrix}$$

ist linear.
Also: $(*)$ lösen bedeutet: $L^{-1}\left(\left\{\begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}\right\}\right)$ bestimmen.

3.15 Das Gauß-Verfahren: **1)** Vertausche die Zeilen, so dass keine Zeile weiter links beginnt als die erste Zeile:

$$\forall j \geq 2 : \min\{k : a_{jk} \neq 0\} \geq \min\{k : a_{1k} \neq 0\}.$$

2) Durch Addition von Vielfachen der 1. Zeile zu den anderen Zeilen erreiche „führende Nullen“ ab der 2. Zeile:

$$\forall j \geq 2 : \min\{k : a_{jk} \neq 0\} \geq \min\{k : a_{1k} \neq 0\} + 1.$$

3) Ab jetzt wird die erste Zeile immer beibehalten. Betrachte die zweite Zeile als erste und wiederhole den Vorgang.

Abbruchbedingung: Führe das Gauß-Verfahren so lange durch, bis das LGS **Zeilenstufenform** hat, d.h. dass jede Zeile entweder vom linken Rand her mehr aufeinanderfolgende Nullkoeffizienten als die vorhergehende Zeile hat oder links vom Gleichheitszeichen aus lauter Nullen besteht:

$$\begin{array}{cccccccccccc}
a_{11}\,x_1 & + & a_{12}\,x_2 & + & \dots & & & & + & a_{1n}\,x_n & = & b_1 \\
0\,x_1 & + & a_{22}\,x_2 & + & \dots & & & & + & a_{2n}\,x_n & = & b_2 \\
0\,x_1 & + & 0\,x_2 & + & a_{33}\,x_3 & + & \dots & & + & a_{3n}\,x_n & = & b_3 \\
 & & & & & \ddots & & & & & & \\
0\,x_1 & + & \dots & & & & a_{lk}\,x_k & + \dots & + & a_{ln}\,x_n & = & b_l \qquad (k \geq l) \\
 & & & & & & & & & 0 & = & b_{l+1} \\
 & & & & & & & & & & \vdots & \\
 & & & & & & & & & 0 & = & b_m
\end{array}$$

Klar ist: Falls $b_{l+1} \neq 0$ oder $b_{l+2} \neq 0$ $\dots$ oder $b_m \neq 0$, dann gibt es keine Lösung. Pech! Andernfalls kann man das LGS „von unten nach oben“ auflösen.

3.16 Begründung: Ist $(x_1, \dots, x_n)$ Lösung von $(*)$, dann auch von dem umgeformten System nach einem Gauß-Schritt (Wir haben ja nur Gleichungen addiert).
Ist $(x_1, \dots, x_n)$ Lösung des umgeformten Systems, dann können wir die ursprünglichen Gleichungen durch Gauß-Umformungen rekonstruieren.
Also: Ein Gauß-Schritt ändert die Lösungsgesamtheit nicht.

3.17 Folgerung: Ein homogenes LGS besitzt immer die Lösung $x_1 = \dots = x_n = 0$. Im Fall $m \leq n - 1$ (weniger Zeilen als Variablen) besitzt ein homogenes LGS immer nichttriviale Lösungen.

3.3 Basis und Dimension

3.18 Definition: Sei $(V, +, \cdot)$ ein Vektorraum über $\mathbb{K}$.

1) Eine Teilmenge $M \subseteq V$ heißt **linear unabhängig**, falls folgende äquivalente Bedingungen erfüllt sind:

(i) Für jede endliche Teilmenge $\{m_1, \ldots, m_n\} \subseteq M$ gilt

$$\forall \lambda_1, \ldots \lambda_n \in \mathbb{K} : \left(\sum_{j=1}^{n} \lambda_j \cdot m_j = 0 \;\Rightarrow\; \lambda_1 = \ldots = \lambda_n = 0 \right).$$

(ii) Ist $v = \sum\limits_{j=1}^{n} \lambda_j \cdot m_j$ mit $m_1, \ldots, m_n \in M$, so sind die Koeffizienten $\lambda_1, \ldots, \lambda_n$ eindeutig.

(iii) $\forall m \in M : \mathrm{LH}\big(M \setminus \{m\}\big) \neq \mathrm{LH}(M)$.

Falls M nicht linear unabhängig ist, heißt M **linear abhängig**.

2) Eine Teilmenge $B \subseteq V$ heißt **Basis**, falls B linear unabhängig ist und $\mathrm{LH}(B) = V$.
Insbesondere besitzt dann jeder Vektor $v \in V$ eine eindeutige Darstellung

$$v \;=\; \sum_{j=1}^{n} \lambda_j \cdot b_j \qquad \text{mit } n \in \mathbb{N},\ \lambda_j \in \mathbb{K},\ b_j \in B.$$

Die Koeffizienten $\lambda_1, \ldots, \lambda_n$ heißen **Koordinaten** von v bezüglich B. Wir schreiben

$$v \;=\; \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix}_B$$

Beispiele: **1)** $V = \mathbb{K}^n$, $e_1 := \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \ldots e_n := \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$.
$B := \{e_1, \ldots, e_n\}$ ist die kanonische Basis von $\mathbb{K}^n$.

2) $V = \mathbb{K}^n$, $a_1, \ldots a_n \in \mathbb{K}^n \setminus \{0\}$, so dass a_j für $j > k$ mehr aufeinanderfolgende Nullen „von oben her" hat als a_k. Dann ist $\{a_1, \ldots, a_n\}$ Basis von $\mathbb{K}^n$.

3) $V = \{(a_n) \text{ Folge in } \mathbb{C}\}$, $e_j = (0, 0, \ldots, 0, 1, 0, 0, \ldots)$
$\uparrow$ j-tes Folgenglied

$M = \{e_1, e_2, \ldots\}$ ist linear unabhängig, aber keine Basis.

4) Seien $P_j : \mathbb{C} \to \mathbb{C} : z \mapsto z^j$, $j \in \mathbb{N}_0$.
$\{P_0, \ldots, P_n\}$ ist Basis von $\mathcal{P}_n$,
$\{P_0, P_1, \ldots\}$ ist Basis von $\mathcal{P}$.

5) $V = \mathcal{F} = \{f : \mathbb{R} \to \mathbb{R} \text{ Abbildung}\}$.
$M := \{x \mapsto \mathrm{e}^{ax} \mid a \in \mathbb{R}\}$ ist linear unabhängig, aber keine Basis.

3.19 Die Idee von Descartes: Sei $\{v_1, \ldots, v_n\} \subseteq V$. Definiere

$$T : \mathbb{K}^n \to V : \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \sum_{j=1}^{n} x_j \cdot v_j$$

Falls T umkehrbar: Jeder Vektor $v \in V$ besitzt Koordinaten $x_1, \ldots, x_n$. Mit Koordinaten kann man rechnen.

3.20 Satz: **1)** T ist linear

2) Äquivalent sind: **(i)** $\mathrm{LH}\,\{v_1, \ldots, v_n\} = V$.
(ii) T ist surjektiv.

3) Äquivalent sind: **(i)** $\{v_1, \ldots, v_n\}$ ist linear unabhängig.
(ii) T ist injektiv.

4) Äquivalent sind: **(i)** $\{v_1, \ldots, v_n\}$ ist Basis von V.
(ii) T ist bijektiv.

3.21 Folgerung: Sei V Vektorraum über $\mathbb{K}$ mit einer endlichen Basis B und $n := \sharp B$. Dann ist V isomorph zu $\mathbb{K}^n$.

3.22 Satz: Der Vektorraum V besitze eine Basis mit $n \in \mathbb{N}$ Elementen. Dann gelten:

1) Jede Menge mit mehr als n Elementen ist linear abhängig.

2) Jede Basis hat genau n Elemente.

Beweis: **1)** Sei $B = \{b_1, \ldots, b_n\}$ die Basis. Dann ist $T : \mathbb{K}^n \to V$ Isomorphismus. Seien nun $w_1, \ldots, w_{n+1} \in V$. Zu zeigen:

$$\begin{aligned}
&\lambda_1 \cdot w_1 + \ldots + \lambda_{n+1} \cdot w_{n+1} = 0 \text{ hat nichttriviale Lösungen} \\
&\quad \Leftrightarrow \quad T^{-1}(\lambda_1 \cdot w_1 + \ldots + \lambda_{n+1} \cdot w_{n+1}) = 0 \text{ hat nichttriviale Lösungen} \\
&\quad \Leftrightarrow \quad \lambda_1 \cdot T^{-1}(w_1) + \ldots + \lambda_{n+1} \cdot T^{-1}(w_{n+1}) = 0 \text{ hat nichttriviale Lösungen} \\
&\quad \Leftrightarrow \quad \text{wahr,}
\end{aligned}$$

da hier ein homogenes LGS mit $n + 1$ Unbekannten und n Zeilen vorliegt.

2) Seien B, B' Basen von V.

$$\left.\begin{array}{l} 1) \;\Rightarrow\; \sharp B' \leq \sharp B \quad \text{(da } B' \text{ linear unabhängig)} \\ 1) \;\Rightarrow\; \sharp B \leq \sharp B' \quad \text{(da } B \text{ linear unabhängig)} \end{array}\right\} \;\Rightarrow\; \sharp B' = \sharp B.$$

□

3.23 Definition: Besitzt V eine endliche Basis B, so heißt $\sharp B$ die **Dimension** von V :

$$\dim V := \sharp B.$$

Besitzt V keine endliche Basis, so heißt V **unendlichdimensional**.

Beispiele: **1)** $\dim(\mathbb{K}^n) = n$,

2) $\mathbb{C}^n$ als Vektorraum über $\mathbb{R}$: $\dim(\mathbb{C}^n) = 2n$,

3) $V = \mathrm{LH}\left\{ \begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\} \Rightarrow \dim(V) = 2$,

4) $\dim(\mathcal{P}_n) = n + 1$,

5) $\mathcal{P}$ ist unendlichdimensional.

3.24 Folgerung: Falls $\dim V = n$ und $\{v_1, \ldots, v_n\} \subseteq V$ linear unabhängig, ist $\{v_1, \ldots, v_n\}$ eine Basis.

3.25 Basisergänzungssatz: Sei V endlichdimensional, $m \in \mathbb{N}$, $m < n := \dim V$ und $\{v_1, \ldots, v_m\} \subseteq V$ linear unabhängig. Dann existieren Vektoren $v_{m+1}, \ldots, v_n \in V$, so dass $\{v_1, \ldots, v_n\}$ eine Basis von V ist.

Beweis: Sei $\{b_1, \ldots, b_n\}$ Basis von V.

$$v_1 = \sum_{j=1}^{n} \lambda_j \cdot b_j \;\wedge\; v_1 \neq 0 \;\Rightarrow\; \exists \lambda_1 : \lambda_{j_1} \neq 0 \;\Rightarrow\; b_{j_1} = \frac{1}{\lambda_{j_1}} \Big(-\sum_{j \neq j_1} \lambda_j \cdot b_j + v_1 \Big)$$

Dann kann b_{j_1} durch v_1 ersetzt werden, d.h.

$$\big(\{b_1, \ldots, b_n\} \setminus \{b_{j_1}\}\big) \cup \{v_1\} \text{ ist Basis von } V.$$

Genauso folgt, dass es ein b_{j_2} gibt, das durch v_2 ersetzt werden kann, und die enstehende Menge ist immer noch Basis von V.

So fortfahrend erhalten wir

$$\big(\{b_1, \ldots, b_n\}\big) \setminus \{b_{j_1}, \ldots, b_{j_m}\} \cup \{v_1, \ldots, v_m\}$$

als gewünschte Basis von V.

□

3.26 Bemerkung: Man kann beweisen: Jeder Vektorraum besitzt eine Basis.

3.27 Dimensionsformel: Sei $L : V \to W$ linear, $\dim V < \infty$. Dann gilt

$$\dim V \; = \; \dim \operatorname{Kern}(L) + \dim \operatorname{Bild}(L)$$

(wobei $\dim\big(\{0\}\big) := 0$).

Beweis: Sei $k := \dim \operatorname{Kern}(L),\; n := \dim(V)$.

Fall $k = n$**:** $\Rightarrow\; V = \operatorname{Kern}(L) \;\wedge\; \operatorname{Bild}(L) = \{0\} \;\Rightarrow\; n = 0 + n$ stimmt.

Fall $1 \le k \le n-1$**:** Sei $\{b_1, \dots b_k\}$ Basis von Kern (L).
Ergänze zu einer Basis $\{b_1, \dots, b_n\}$ von V

$$\Rightarrow \; \operatorname{Bild}(L) = L(V) = L\big(\mathrm{LH}\{b_1, \dots, b_n\}\big) = \mathrm{LH}\{Lb_1, \dots, Lb_n\} = \mathrm{LH}\{Lb_{k+1}, \dots, Lb_n\}.$$

Behauptung: $\{Lb_{k+1}, \dots, Lb_n\}$ ist linear unabhängig.
Dann: $n = k + (n-k)$ stimmt.

$$\begin{aligned} 0 \; &= \; \lambda_{k+1} \cdot Lb_{k+1} + \ldots + \lambda_n \cdot Lb_n \; = \; L(\lambda_{k+1} \cdot b_{k+1} + \ldots + \lambda_n \cdot b_n) \\ &\Rightarrow \; \lambda_{k+1} \cdot b_{k+1} + \ldots + \lambda_n \cdot b_n \in \operatorname{Kern}(L) \\ &\Rightarrow \; \lambda_{k+1} \cdot b_{k+1} + \ldots + \lambda_n \cdot b_n = \lambda_1 \cdot b_1 + \ldots + \lambda_k \cdot b_k \\ &\Rightarrow \; \lambda_{k+1} = \ldots = \lambda_n = 0. \end{aligned}$$

Fall $k = 0$**:** Sei $\{b_1, \dots, b_n\}$ Basis von V. Behauptung: $\{Lb_1, \dots, Lb_n\}$ ist linear unabhängig.
Dann: $n = 0 + n$ stimmt.

$$\begin{aligned} 0 \; &= \; \lambda_1 \cdot Lb_1 + \ldots + \lambda_n \cdot Lb_n \; = \; L(\lambda_1 \cdot b_1 + \ldots + \lambda_n \cdot b_n) \\ &\Rightarrow \; \lambda_1 \cdot b_1 + \ldots + \lambda_n \cdot b_n \in \operatorname{Kern}(L) = \{0\} \\ &\Rightarrow \; \lambda_1 \cdot b_1 + \ldots + \lambda_n \cdot b_n = 0 \\ &\Rightarrow \; \lambda_1 = \ldots = \lambda_n = 0. \end{aligned}$$

$\square$

Beispiele: **1)** $L : \mathbb{R}^4 \to \mathbb{R}^2 : \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \mapsto \begin{pmatrix} x_2 \\ 0 \end{pmatrix}$ ist linear.

$$\begin{aligned} \operatorname{Kern}(L) \; &= \; \left\{ \begin{pmatrix} x_1 \\ 0 \\ x_3 \\ x_4 \end{pmatrix} : x_1, x_3, x_4 \in \mathbb{R} \right\} & \Rightarrow \; \dim(\operatorname{Kern}(L)) = 3 \\ \operatorname{Bild}(L) \; &= \; \left\{ \begin{pmatrix} x_2 \\ 0 \end{pmatrix} : x_2 \in \mathbb{R} \right\} & \Rightarrow \; \dim(\operatorname{Bild}(L)) = 1 \\ 3 + 1 \; &= \; 4 \; = \; \dim(\mathbb{R}^4) \end{aligned}$$

2) $L : \mathcal{P}_n \to \mathcal{P} : P \mapsto P'$:

$$\left.\begin{array}{rcl} \dim(\mathrm{Kern}\,(L)) & = & 1 \\ \dim(\mathcal{P}_n) & = & n+1 \end{array}\right\} \Rightarrow \ \dim(\mathrm{Bild}\,(L)) \ = \ n+1-1 \ = \ n$$

3.28 Definition: Sei $L : V \to W$ linear. Dann heißt $\mathrm{Rang}(L) := \dim\big(\mathrm{Bild}\,(L)\big)$ der **Rang** der Abbildung L.

3.4 Lineare Abbildungen und Matrizen

3.29 Lineare Abbildung und Basis: Sei $\{b_1, \ldots, b_n\}$ Basis von V, $\{w_1, \ldots, w_n\} \subseteq W$. Dann existiert genau eine lineare Abbildung $L : V \to W$ mit

$$L(b_1) = w_1, \ \ldots, \ L(b_n) = w_n.$$

Beweis: **1)** Annahme: $L : V \to W$ ist eine lineare Abbildung mit $L(b_j) = w_j, \ j = 1, \ldots, n$.

$$\text{Sei } v \in V \ \Rightarrow \ v = \sum_{j=1}^{n} \lambda_j \cdot b_j, \ \lambda_j \text{ eindeutig}$$

$$\Rightarrow \ L(v) \ = \ \sum_{j=1}^{n} \lambda_j \cdot L(b_j) \ = \ \sum_{j=1}^{n} \lambda_j \cdot w_j \ \Rightarrow \ L(v) \text{ ist eindeutig}$$

$\Rightarrow$ Eindeutigkeit von L.

2) Definiere $L : V \to W : \sum_{j=1}^{n} \lambda_j \cdot b_j \mapsto \sum_{j=1}^{n} \lambda_j \cdot w_j$ Dann ist L linear (durch Nachrechnen)

$\Rightarrow$ Existenz von L. $\square$

Beispiele: **1)** $V = \mathbb{K}^n, \ b_j = e_j$ (kanonische Basis)

$$\Rightarrow \ L : \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \sum_{j=1}^{n} x_j \cdot w_j$$

2) $L : \mathbb{R}^2 \to \mathbb{R}^3$ definiert durch $L\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 2 \\ 2 \end{pmatrix}$, $L\begin{pmatrix} 0 \\ 1 \end{pmatrix} \mapsto \begin{pmatrix} 0 \\ -1 \\ -1 \end{pmatrix}$

$$\Rightarrow \ L\begin{pmatrix} x \\ y \end{pmatrix} = L\left(x \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + y \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix}\right) = x \cdot \begin{pmatrix} 0 \\ 2 \\ 2 \end{pmatrix} + y \cdot \begin{pmatrix} 0 \\ -1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 2x-y \\ 2x-y \end{pmatrix}$$

3.30 Satz: **1)** Äquivalent sind: **(i)** $\mathrm{LH}\left(\{w_1, \ldots, w_n\}\right) = W$.
(ii) L ist surjektiv.

2) Äquivalent sind: **(i)** $\{w_1, \ldots, w_n\}$ ist linear unabhängig.
(ii) L ist injektiv.

3) Äquivalent sind: **(i)** $\{w_1, \ldots, w_n\}$ ist Basis von W.
(ii) L ist bijektiv.

Beweis: Siehe 3.20 □

3.31 Folgerung: Seien $B = \{b_1, \ldots, b_n\}, C = \{c_1, \ldots, c_m\}$ Basen von V bzw. W, und

$$w_k = \sum_{j=1}^{m} \alpha_{jk} \cdot c_j \qquad (k = 1, \ldots, n)$$

die eindeutigen Darstellungen von $w_1, \ldots, w_n$ bezüglich der Basis C. Dann ist die durch 3.29 gegebene Abbildung eindeutig bestimmt durch Angabe der α_{jk}. Vereinbarung: Schreibe die α_{jk} in **Matrixform**:

$$M = M_L^{C,B} = \begin{pmatrix} \alpha_{11} & \ldots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{m1} & \ldots & \alpha_{mn} \end{pmatrix} = (\alpha_{jk}) = (\alpha_{jk})_{\substack{j=1,\ldots,m \\ k=1,\ldots,n}} .$$

3.32 Definition: $M = M_L^{C,B}$ heißt die $\boldsymbol{m \times n}$ **- Matrix** der linearen Abbildung L bezüglich der Basen B und C. Es gelten:

1) Zeilenzahl = dim W (Bildraum),

2) Spaltenzahl = dim V (Urbildraum),

3) in den Spalten stehen die Koordinaten bezüglich C der Bilder der Basisvektoren:

$$M_L^{C,B} = \begin{pmatrix} \cdot & \alpha_{12} & \ldots & \ldots \\ \cdot & \alpha_{22} & \ldots & \ldots \\ \vdots & \vdots & \vdots & \vdots \\ \cdot & \alpha_{m2} & \ldots & \ldots \end{pmatrix} \Rightarrow L(b_2) = \alpha_{12} \cdot c_1 + \ldots + \alpha_{m2} \cdot c_m =: \begin{pmatrix} \alpha_{12} \\ \alpha_{22} \\ \vdots \\ \alpha_{m2} \end{pmatrix}_C$$

Beispiele: **1)** $L : \mathbb{R}^2 \to \mathbb{R}^2 : \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} -x_2 \\ x_1 \end{pmatrix}$

$$B = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\} = C \Rightarrow M_L^{C,B} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}$$

Aber:

$$C' = \left\{ \begin{pmatrix} 0 \\ 2 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \end{pmatrix} \right\} \Rightarrow M_L^{C',B} = \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{pmatrix}$$

2) Zu $\mathbb{0} : V \to W : v \mapsto 0$ gehört immer die **Nullmatrix**

$$M_{\mathbb{0}}^{C,B} = \begin{pmatrix} 0 & \dots & 0 \\ \vdots & & \vdots \\ 0 & \dots & 0 \end{pmatrix} =: \mathbb{0}$$

3) Zu $\mathrm{Id}_V : V \to V : x \mapsto x$ gehört bezüglich **einer** Basis B die Einheitsmatrix:

$$M_{\mathrm{Id}_V}^{B,B} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix} =: E_n \text{ } (n = \text{Anzahl Spalten/Zeilen})$$

Aber: $M_{\mathrm{Id}_V}^{C,B}$ sieht anders aus, wenn $C \neq B$ (Transformationsmatrizen).

4) Sei $L : V \to W$ bijektiv, $B = \{b_1, \dots, b_n\}$ Basis von V, $C := \{Lb_1, \dots, Lb_n\}$ (nach 3.30 ist C Basis von W).

$$\Rightarrow M_L^{C,B} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & 1 \end{pmatrix} = E_n$$

Diese Matrix sagt über die Abbildung nicht viel aus.

5) $V = \mathbb{R}^3,\ B = \left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\}, W = \mathbb{R}^2,\ C = \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right\}.$

$$M_L^{C,B} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \Rightarrow \begin{cases} L\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 4 \end{pmatrix} \\ L\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 5 \end{pmatrix} \\ L\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 6 \end{pmatrix} \end{cases}$$

$$\Rightarrow L\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = L\left(x_1 \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + x_2 \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + x_3 \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right) = \begin{pmatrix} 1x_1 + 2x_2 + 3x_3 \\ 4x_1 + 5x_2 + 6x_3 \end{pmatrix}$$

3.33 Bild eines beliebigen Vektors: Sei $L : V \to W$ gegeben durch $M_L^{C,B} = (\alpha_{jk})$, und sei $B = \{b_1, \dots, b_n\}$, $C = \{c_1, \dots, c_m\}$.

Für $v \in V$ gilt $v = \sum_{k=1}^{n} v_k \cdot b_k$ und

$$\begin{aligned} L(v) &= \sum_{k=1}^{n} v_k \cdot L(b_k) \\ &= \sum_{k=1}^{n} v_k \cdot \left(\sum_{j=1}^{m} \alpha_{jk} \cdot c_j \right) \\ &= \sum_{j=1}^{m} \underbrace{\left(\sum_{k=1}^{n} \alpha_{jk}\, v_k \right)}_{= \text{Koordinaten von } Lv \text{ bezüglich C}} \cdot c_j \end{aligned}$$

Also $L : v \mapsto Lv$:

$$\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_B \mapsto \begin{pmatrix} \sum\limits_{k=1}^{n} \alpha_{1k}\, v_k \\ \vdots \\ \sum\limits_{k=1}^{n} \alpha_{mk}\, v_k \end{pmatrix}_C =: \begin{pmatrix} \alpha_{11} & \alpha_{12} & \dots & \alpha_{1n} \\ \vdots & & & \vdots \\ \alpha_{m1} & \alpha_{m2} & \dots & \alpha_{mn} \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}_B$$

D.h.: 1. Koordinate des Bildvektors = 1. Zeile der Matrix „Mal“ Vektor

$\vdots$

m-te Koordinate des Bildvektors = m-te Zeile der Matrix „Mal“ Vektor

3.34 Satz: Seien $B = \{b_1, \dots, b_n\}, C = \{c_1, \dots, c_m\}$ Basen von V bzw. W und $M_L^{C,B}$ die Matrix der Abbildung $L : V \to W$. Dann berechnen sich die Koordinaten von Lv bezüglich der Basis C aus den Koordinaten von v bezüglich B durch

$$\big(Lv\big)_C = \begin{pmatrix} (Lv)_1 \\ \vdots \\ (Lv)_m \end{pmatrix}_C = \begin{pmatrix} \alpha_{11} & \dots & \alpha_{1n} \\ \vdots & & \vdots \\ \alpha_{m1} & \dots & \alpha_{mn} \end{pmatrix} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}_B = M_L^{C,B} \cdot v_B$$

3.35 Satz: Sei $L : \mathbb{K}^n \to \mathbb{K}^m$ linear, E die kanonische Basis von $\mathbb{K}^n$, E' die kanonische Basis von $\mathbb{K}^m$ und $M_L^{E',E} = (a_{jk})$ mit den **Spaltenvektoren**

$$a_k := \begin{pmatrix} a_{1k} \\ a_{2k} \\ \vdots \\ a_{mk} \end{pmatrix} \qquad (k = 1, \dots, n).$$

Dann gelten

$$\begin{aligned} \text{Bild}\,(L) &= \text{LH}\{a_1, \dots, a_n\}, \\ \text{Rang}\,(L) &= \dim\Big(\text{LH}\{a_1, \dots, a_n\}\Big). \end{aligned}$$

3.36 Satz: Sei $L : V \to W$ linear, $\dim(V) = n$, $\dim(W) = m$. Dann existieren Basen B, C von V bzw. W so, dass

$$M_L^{C,B} = \begin{array}{c} \begin{pmatrix} 1 & 0 & 0 & \dots & & & 0 \\ 0 & 1 & 0 & \dots & & & 0 \\ \vdots & \ddots & \ddots & \ddots & & & \vdots \\ 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ 0 & \dots & & & & \dots & 0 \\ \vdots & & & & & & \vdots \\ 0 & \dots & & & & \dots & 0 \end{pmatrix} \begin{array}{l} \\ \\ \\ \leftarrow k\text{-te Zeile} \\ \left.\begin{array}{l} \\ \\ \\ \end{array}\right\} m-k \text{ Nullzeilen} \end{array} \\ \begin{array}{ccccccc} & & & \uparrow & \underbrace{\qquad\qquad}_{n-k \text{ Nullspalten}} \\ & & & k\text{-te Spalte} & \end{array} \end{array}$$

und es gelten

$$\begin{array}{rcl} \text{Rang}\,(L) &=& \text{Anzahl der Spalten ungleich Null} &=& k \\ \dim \text{Kern}\,(L) &=& \text{Anzahl der Nullspalten} &=& n-k &=& \dim V - k \end{array}$$

3.37 Verknüpfung linearer Abbildungen: Seien U, V, W Vektorräume mit endlichen Basen B_U, B_V, B_W und

$K : U \to V$ linear mit Matrix $\quad M_K^{B_V,B_U} = (\beta_{jk})_{\substack{j=1,\ldots,m \\ k=1,\ldots,n}}$

$L : V \to W$ linear mit Matrix $\quad M_L^{B_W,B_V} = (\alpha_{ij})_{\substack{i=1,\ldots,l \\ j=1,\ldots,m}}$

Dann ist die Matrix von $L \circ K$ gegeben durch

$$(\gamma_{ik})_{\substack{i=1,\ldots,l \\ k=1,\ldots,n}} \;=\; M_{L\circ K}^{B_W,B_U} \;=:\; M_L^{B_W,B_V} \cdot M_K^{B_V,B_U} \;=\; (\alpha_{ij}) \cdot (\beta_{jk})$$

wobei das **Matrizenprodukt** definiert ist durch

$$\gamma_{ik} \;:=\; \sum_{j=1}^{m} \alpha_{ij}\,\beta_{jk}$$

($\gamma_{ik} = i$-te Zeile der Matrix (α_{ij}) „Mal“ k-te Spalte der Matrix (β_{jk}).

Beweis: $B_u = \{a_1, \ldots, a_n\},\ B_V = \{b_1, \ldots, b_m\},\ B_W = \{c_1, \ldots, c_l\}$

$$\begin{aligned} \Rightarrow\ (L \circ K)(a_k) &= L\big(K(a_k)\big) \\ &= L\left(\sum_{j=1}^{m} \beta_{jk} \cdot b_j\right) \\ &= \sum_{j=1}^{m} \beta_{jk} \cdot L(b_j) \\ &= \sum_{j=1}^{m} \beta_{jk} \cdot \left(\sum_{i=1}^{l} \alpha_{ij} \cdot c_i\right) \\ &= \sum_{i=1}^{l} \underbrace{\left(\sum_{j=1}^{m} \alpha_{ij}\,\beta_{jk}\right)}_{=\gamma_{ik}} \cdot c_i \end{aligned}$$

$\square$

3.38 Definition: Die Matrix $C = (\gamma)_{ik}$ heißt **Matrizenprodukt** von A und B: $\quad C = A \cdot B$. Das Matrizenprodukt $A \cdot B$ kann nur gebildet werden, wenn Spaltenzahl von A = Zeilenzahl von B.

Beispiele: **1)** $\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}.$

2) $\begin{pmatrix} 1 & 2 & -1 & 1 \\ 0 & 1 & 2 & -2 \\ 2 & -1 & 0 & -3 \end{pmatrix} \cdot \begin{pmatrix} 1 & -2 \\ 0 & -1 \\ 1 & 1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & -5 \\ 0 & 1 \\ -1 & -3 \end{pmatrix}.$

3) $\begin{pmatrix} 1 & 2 & -1 & 1 \\ 0 & 1 & 2 & -2 \\ 2 & -1 & 0 & -3 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}.$

4) $\begin{pmatrix} 1 & 2 & -1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix} = (1).$

5) $\begin{pmatrix} 1 \\ 0 \\ 1 \\ 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & -1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & -1 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 2 & -1 & 1 \\ 1 & 2 & -1 & 1 \end{pmatrix}$

(manchmal auch “dyadisches Produkt”). Insbesondere $A \cdot B \neq B \cdot A$ (vergleiche mit dem vorigen Beispiel).

3.39 Matrizen und lineare Abbildungen: Zu einer $m \times n$-Matrix

$$M = (a_{jk})_{\substack{j=1,\dots,m \\ k=1,\dots,n}} \quad \text{mit Spaltenvektoren } a_k = \begin{pmatrix} a_{1k} \\ \vdots \\ a_{mk} \end{pmatrix} \in \mathbb{K}^m$$

definieren wir die lineare Abbildung

$$L_M : \mathbb{K}^n \to \mathbb{K}^m : \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \sum_{k=1}^{n} x_k \cdot a_k = M \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Dann gilt $M_{L_M}^{E',E} = M$, wenn E, E' die kanonischen Basen von $\mathbb{K}^n$ bzw. $\mathbb{K}^m$ bezeichnen. Die Abbildung

$$\Phi : M \mapsto L_M \text{ mit Umkehrabbildung } \Phi^{-1} : L \mapsto M_L^{E',E}$$

ist eine bijektive Abbildung von der Menge aller $m \times n$-Matrizen auf die Menge aller linearen Abbildungen $L : \mathbb{K}^n \to \mathbb{K}^m$

3.40 Matrizen und Verkettung linearer Abbildungen: Sei M eine $m \times n$-Matrix und K eine $l \times m$-Matrix. Dann: $\Phi(M) : \mathbb{K}^n \to \mathbb{K}^m$, $\Phi(K) : \mathbb{K}^m \to \mathbb{K}^l$, und

$$\begin{array}{ccccl} (K, & M) & \longmapsto & K \cdot M & \text{Matrizenmultiplikation} \\ \Phi \updownarrow & \updownarrow \Phi & & \updownarrow \Phi & \\ (\Phi(K), & \Phi(M)) & \longmapsto & \Phi(K) \circ \Phi(M) & \text{Hintereinanderausführung} \end{array}$$

Also: $\Phi(K \cdot M) = \Phi(K) \circ \Phi(M)$.

So wurde die Matrizenmultiplikation definiert! 1

Da die Verkettung von Abbildungen assoziativ ist, ist auch die Matrizenmultipliktion assoziativ: Sind M, K, L Matrizen, so dass $M \cdot (K \cdot L)$ definiert ist, so ist auch $(M \cdot K) \cdot L$ definiert und ergibt die selbe Matrix.

3.41 Umkehrabbildungen: **1)** Seien $L : V \to V$ linear und bijektiv und B eine Basis von V mit n Elementen. Dann gilt

$$M_L^{B,B} \cdot M_{L^{-1}}^{B,B} \;=\; M_{L^{-1}}^{B,B} \cdot M_L^{B,B} \;=\; M_{\mathrm{Id}}^{B,B} \;=\; E_n \;:=\; \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$$

Die $n \times n$-Matrix E_n heißt **Einheitsmatrix**.

2) Sei M eine $n \times n$-Matrix (über $\mathbb{K}$). Falls eine $n \times n$-Matrix M^{-1} existiert mit $M \cdot M^{-1} = E_n$, so gilt auch $M^{-1} \cdot M = E_n$. Außerdem ist die Matrix M^{-1} eindeutig, d.h. die Matrizengleichung

$$M \cdot X \;=\; E_n$$

besitzt genau eine Lösung $X = M^{-1}$.

3.42 Definition: Eine $n \times n$-Matrix M über $\mathbb{K}$ heißt **invertierbar**, falls eine $n \times n$-Matrix M^{-1} existiert mit $M \cdot M^{-1} = E_n$. Dann gilt auch $M^{-1} \cdot M = E_n$, und die Matrix M^{-1} heißt **inverse Matrix** zu M.

3.43 Folgerung: Es sei $L : V \to V$ linear mit Matrix $M = M_L^{B,B}$. Dann sind äquivalent:

(i) L ist bijektiv.

(ii) M ist invertierbar.

Beweis: (i) $\Rightarrow$ (ii): (i) $\Rightarrow$ $M_{L^{-1}}^{B,B} = M^{-1}$ $\Rightarrow$ (ii)

(ii) $\Rightarrow$ (i): Setze $K : V \to V : x \mapsto K(x)$ mit $\big(K(x)\big)_B := M^{-1} \cdot x_B$.

$\Rightarrow$ $\big((L \circ K)(x)\big)_B = L\big(K(x)\big)_B = M \cdot \big(K(x)\big)_B = M \cdot (M^{-1} \cdot x_B) = (M^{-1} \cdot M) \cdot x_B = x_B$

$\Rightarrow$ $L \circ K = \mathrm{Id}_V$, also $K = L^{-1}$

$\Rightarrow$ (i) $\square$

3.44 Folgerung: Sind A, B invertierbare Matrizen, dann ist auch $A \cdot B$ invertierbar und es gilt

$$(A \cdot B)^{-1} \;=\; B^{-1} \cdot A^{-1}.$$

3.45 Satz: Die Menge

$$\left\{L : \mathbb{K}^n \to \mathbb{K}^n \;\middle|\; L \text{ ist linear und bijektiv}\right\}$$

mit der Hintereinanderausführung $\circ$ als Verknüpfung bildet eine Gruppe, die **lineare Gruppe $\mathbf{GL}_n(\mathbb{K})$**.

3.46 Folgerung: Die Menge

$$\left\{M \;\middle|\; M \text{ ist invertierbare } n \times n\text{-Matrix über } \mathbb{K}\right\}$$

mit der Matrizenmultiplikation als Verknüpfung bildet eine Gruppe. Sie wird ebenfalls als **lineare Gruppe $\mathbf{GL}_n(\mathbb{K})$** bezeichnet.

3.5 Lineare Gleichungssysteme II

3.47 LGS und Matrizen: Sei ein LGS

$$\left.\begin{array}{lclclclcl} a_{11}x_1 & + & a_{12}x_2 & + & \dots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & \dots & & & + & a_{2n}x_n & = & b_2 \\ \vdots & & & & & & & & \\ a_{m1}x_1 & + & \dots & & & + & a_{mn}x_n & = & b_m \end{array}\right\} \qquad (*)$$

mit der **Koeffizientenmatrix**

$$M := \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$$

gegeben. Dann ist $x_1, \dots, x_n$ genau dann Lösung von $(*)$, wenn der Vektor $x := \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ Lösung von

$$L_M(x) = b := \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \qquad (**)$$

ist, oder anders geschrieben:

$$M \cdot x = b \qquad (***)$$

3.48 Folgerung: Das LGS $M \cdot x = b$ ist genau dann lösbar, wenn $b \in \operatorname{Bild}(L_M)$.

3.49 Lösungsstruktur: Sei $L : \mathbb{K}^n \to \mathbb{K}^m$ linear und das inhomogene LGS $L\,x = b$ gegeben. Ist x_0 eine Lösung, dann sind alle Lösungen gegeben durch

$$x \;\in\; x_0 + \mathrm{Kern}\,(L) \;:=\; \{x_0 + y : y \in \mathrm{Kern}\,(L)\} \;=\; \{x_0 + y : Ly = 0\}.$$

Man nennt x_0 **partikuläre Lösung**, obwohl x_0 nur irgendeine Lösung bezeichnet (falls es mehrerere Lösungen gibt).

3.50 Folgerung: Sei $L : \mathbb{K}^n \to \mathbb{K}^n$ linear und das LGS $L\,x = b$ gegeben. Dann sind äquivalent:

(i) Das LGS besitzt für jeden Vektor $b \in \mathbb{K}^n$ genau eine Lösung.

(ii) $\mathrm{Kern}\,(L) = \{0\}$

(iii) $\mathrm{Rang}\,(L) = n$

Beweis: (i) $\Rightarrow$ $\mathrm{Kern}\,(L) = \{0\}$ $\Leftrightarrow$ (ii)

(ii) $\Leftrightarrow$ (iii): Dimensionsformel $n = \dim(\mathrm{Kern}\,(L)) + \underbrace{\dim(\mathrm{Bild}\,(L))}_{=\mathrm{Rang}(L)}$

(ii) $\Rightarrow$ (i): (ii) $\Rightarrow$ (iii) $\Rightarrow$ $\mathrm{Bild}\,(L) = \mathbb{K}^n$ $\Rightarrow$ $b \in \mathrm{Bild}\,(L)$ $\Rightarrow$ mindestens eine Lösung existiert.
(ii) $\Rightarrow$ Lösung ist eindeutig
$\Rightarrow$ (i) $\square$

3.51 Definition: Sei $M = (a_{jk})$ eine $m \times n$-Matrix mit den Spaltenvektoren $a_k = \begin{pmatrix} a_{1k} \\ \vdots \\ a_{mk} \end{pmatrix}$ $(k = 1, \ldots, n)$. Der **Rang** von M ist definiert durch

$$\mathrm{Rang}\,(M) \;:=\; \mathrm{Rang}\,(L_M) \;=\; \dim(\mathrm{Bild}\,(L_M)) \;=\; \dim \mathrm{LH}\big(\{a_1, \ldots, a_n\}\big).$$

3.52 Folgerung: Eine $n \times n$-Matrix M ist genau dann invertierbar, wenn $\mathrm{Rang}\,(M) = n$.

Beweis: M invertierbar

$$\begin{aligned}
&\overset{3.43}{\Leftrightarrow} && L_M \text{ bijektiv} \\
&\Leftrightarrow && \begin{cases} \mathrm{Bild}\,(L_M) = \mathbb{K}^n & (L_M \text{ surjektiv}) \\ \mathrm{Kern}\,(L_M) = \{0\} & (\Leftrightarrow\; L_M \text{ injektiv}) \end{cases} \\
&\Leftrightarrow && \dim(\mathrm{Bild}\,(L_M)) = n \\
&\overset{\text{Dimensionsformel}}{\Leftrightarrow} && \mathrm{Rang}(L_M) = n
\end{aligned}$$

$\square$

3.53 Satz: Sei das LGS $A \cdot x = b$ gegeben, $A' := (A\,b)$ die **erweiterte Koeffizientenmatrix**. Dann sind äquivalent:

(i) $A \cdot x = b$ besitzt mindestens eine Lösung.

(ii) $\text{Rang}\,(A) = \text{Rang}\,(A')$.

Beweis: (i)
$$\begin{aligned} &\Leftrightarrow\ b \in \text{LH}\{a_1, \ldots, a_n\} \\ &\Leftrightarrow\ \text{LH}\{a_1, \ldots, a_n\} = \text{LH}\{a_1, \ldots, a_n, b\} \\ &\Leftrightarrow\ \text{Rang}(A) = \text{Rang}(A') \end{aligned}$$
□

3.54 Berechnung inverser Matrizen: Sei M eine $n \times n$-Matrix. Schreibe $\big(M\big|E_n\big)$. Wende Gaußalgorithmus an, bis links vom Trennstrich die Einheitsmatrix steht. Dann steht rechts die inverse Matrix:

$$\big(M\big|E_n\big) \Leftrightarrow \ldots \Leftrightarrow \big(E_n\big|M^{-1}\big).$$

Tritt in einem Gauß-Schritt links vom Trennstrich eine Nullzeile auf, so ist M nicht invertierbar.

3.6 Basiswechsel

3.55 Definition: Sei V ein Vektorraum mit Basen B, B'. Dann heißt

$$S \;:=\; M_{\text{Id}_V}^{B',B}$$

die **Basiswechselmatrix** von B auf B'.

3.56 Eigenschaften: **1)** Ist $B = \{b_1, \ldots, b_n\}$ und $B' = \{c_1, \ldots, c_n\}$ und

$$v \;=\; \sum_{j=1}^{n} x_j \cdot b_j \;=\; \sum_{j=1}^{n} y_j \cdot c_j,$$

so können die Koordinaten $x_1, \ldots, x_n$ und $y_1, \ldots, y_n$ folgendermaßen umgerechnet werden:

$$\begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}_{B'} = M_{\text{Id}_V}^{B',B} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}_{B} \quad \text{bzw.} \quad \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}_{B} = M_{\text{Id}_V}^{B,B'} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}_{B'}$$

2) S ist invertierbar: $S^{-1} = M_{\text{Id}_V}^{B,B'}$.

3) Ist $L : V \to V$ linear, dann ist

$$M_L^{B',B'} \;=\; M_{\mathrm{Id}_V}^{B',B} \cdot M_L^{B,B} \cdot M_{\mathrm{Id}_V}^{B,B'}$$

4) Allgemeiner: Seien $L : U \to V$, B, B' Basen von U, C, C' Basen von V, $M := M_L^{C,B}$, $R := M_{\mathrm{Id}_U}^{B',B}$, $S := M_{\mathrm{Id}_V}^{C',C}$. Dann gelten

$$\begin{aligned} M_L^{C',B} &= M_{\mathrm{Id}_V}^{C',C} \cdot M_L^{C,B} &&= S \cdot M \\ M_L^{C,B'} &= M_L^{C,B} \cdot M_{\mathrm{Id}_U}^{B,B'} &&= M \cdot R^{-1} \\ M_L^{C',B'} &= M_{\mathrm{Id}_V}^{C',C} \cdot M_L^{C,B} \cdot M_{\mathrm{Id}_U}^{B,B'} &&= S \cdot M \cdot R^{-1} \end{aligned}$$

3.57 Definition: **1)** Zwei $m \times n$-Matrizen A, B heißen **äquivalent**, wenn es invertierbare Matrizen S $(m \times m)$ und R $(n \times n)$ gibt mit

$$B \;=\; S \cdot A \cdot R^{-1} \qquad (\text{vgl. oben } M_L^{C',B'} = S \cdot M \cdot R^{-1}).$$

D.h. A und B beschreiben bezüglich geeigneter Basen dieselbe Abbildung.
Insbesondere: $\mathrm{Rang}\,(A) = \mathrm{Rang}\,(B)$.
Diese Relation definiert auf der Menge aller $m \times n$-Matrizen eine Äquivalenzrelation.

2) Zwei $n \times n$-Matrizen A, B heißen **ähnlich**, wenn es eine invertierbare $n \times n$-Matrix S gibt mit

$$B \;=\; S \cdot A \cdot S^{-1}.$$

3.58 Satz: Es sei M eine $m \times n$-Matrix. Die Anwendung des Gauß-Algorithmus (Vertauschen von Zeilen und Addition des Vielfachen einer Zeile zu einer anderen Zeile) ändert den Rang von M nicht. Insbesondere gilt

$$\begin{aligned} \mathrm{Rang}\,(M) &= \text{maximale Anzahl linear unabhängiger Spalten} \\ &= \text{maximale Anzahl linear unabhängiger Zeilen} \end{aligned}$$

D.h. Zeilenrang = Spaltenrang.

Beweis: **1)** Vertauschung zweier Zeilen, z.B. 1. und 2. Zeile:

$$M' = \begin{pmatrix} 0 & 1 & 0 & \dots & & 0 \\ 1 & 0 & 0 & \dots & & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & 0 & \ddots & & \vdots \\ & & \vdots & & \ddots & 0 \\ 0 & 0 & 0 & \dots & 0 & 1 \end{pmatrix} \cdot M \cdot E_n$$

$\Rightarrow$ M', M sind äquivalent, haben denselben Rang.

2) Addition $\lambda \cdot$ (1. Zeile) zur 2. Zeile:

$$M' = \begin{pmatrix} 1 & 0 & \dots & & 0 \\ \lambda & 1 & 0 & \dots & 0 \\ 0 & 0 & \ddots & \vdots & \\ \vdots & \vdots & & & 0 \\ 0 & 0 & \dots & 0 & 1 \end{pmatrix} \cdot M \cdot E_n$$

$\Rightarrow$ M', M sind äquivalent, haben denselben Rang.

3) Forme M mit Gaußschritten um zu Zeilenstufenmatrix M'

$$\Rightarrow \ \text{Rang}(M) = \text{Rang}(M') = \text{ Anzahl der nicht-Null Zeilen in } M'$$

Gauß-Schritte ändern nicht die maximale Anzahl der linear unabhängigen Zeilen. $\square$

3.7 Länge von Vektoren

3.59 Definition: Seien $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$ und V ein $\mathbb{K}$-Vektorraum. Eine Abbildung $\|.\| : V \to \mathbb{R} : x \mapsto \|x\|$ heißt **Norm**, falls für $x, y \in V$ und $\lambda \in \mathbb{K}$ gelten:

(N1) $\|x\| \geq 0 \quad \wedge \quad (\|x\| = 0 \Leftrightarrow x = 0)$ (**Positivität, Definitheit**)

(N2) $\|\lambda \cdot x\| = |\lambda| \, \|x\|$ (**Homogenität**)

(N3) $\|x + y\| \leq \|x\| + \|y\|$ (**Δ-Ungleichung**)

(vgl. Eigenschaften des Betrages in Satz 2.15).

3.60 Definition: Ist $\|.\|$ eine Norm auf V, so heißt $K_1(0) := \{x \in V : \|x\| \leq 1\}$ **Einheitskugel**. Ist $\|x\| = 1$, so heißt x **Einheitsvektor**.

3.61 Definition: Sei V Vektorraum über $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$ mit Norm $\|.\|$. Dann heißt

$$d(x, y) := \|x - y\| \qquad \text{für } x, y \in V$$

der **Abstand** von x zu y. Der Abstand besitzt folgende Eigenschaften:

(M1) $d(x, y) \geq 0$ sowie $d(x, y) = 0 \Leftrightarrow x = y$ (**Positivität**, **Definitheit**).

(M2) $d(x, y) = d(y, x)$ (**Symmetrie**).

(M3) $d(x, y) \leq d(x, z) + d(z, y)$ (**Δ-Ungleichung**).

(vgl. Abstand in Körpern, Definition 2.18).

3.8 Winkel im Vektorraum

3.62 Definition: Sei V Vektorraum über $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$. Eine Abbildung

$$V \times V \to \mathbb{K} : (x, y) \mapsto \langle x, y \rangle$$

heißt **Skalarprodukt** (*engl. inner product*) auf V, falls

(SP1) $\langle x, x \rangle \geq 0 \quad \wedge \quad (\langle x, x \rangle = 0 \Leftrightarrow x = 0)$

(SP2) $\langle y, x \rangle = \overline{\langle x, y \rangle}$ (insbesondere $\langle x, x \rangle \in \mathbb{R}$),

(SP3) $\langle \lambda \cdot x + \mu \cdot y, z \rangle = \lambda \langle x, z \rangle + \mu \langle y, z \rangle$

Ein Vektorraum mit Skalarprodukt $(V, \langle ., . \rangle)$ heißt auch **euklidischer** ($\mathbb{K} = \mathbb{R}$) oder **unitärer** ($\mathbb{K} = \mathbb{C}$) Vektorraum.

Beispiele: **1)** Standard-Skalarprodukt im $\mathbb{R}^n$:

$$\left\langle \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right\rangle = \sum_{j=1}^{n} x_j \, y_j.$$

2) Standard-Skalarprodukt im $\mathbb{C}^n$:

$$\left\langle \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right\rangle = \sum_{j=1}^{n} x_j \, \overline{y_j}.$$

3) Im $\mathbb{R}^n$ ist auch

$$\left\langle \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right\rangle^{\sim} = \left\langle \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \right\rangle = \sum_{j=1}^{n} \lambda_j \, x_j \, y_j$$

ein Skalarprodukt, falls $\lambda_1, \ldots, \lambda_n > 0$.

3.63 Eigenschaften: **1)** Für festes $y \in V$ ist die Abbildung

$$x \mapsto \langle x, y \rangle$$

linear nach (SP3).

2) Nach (SP2) und (SP3)

$$\langle x, \lambda \cdot y + \mu \cdot z \rangle = \overline{\lambda} \, \langle x, y \rangle + \overline{\mu} \, \langle x, z \rangle .$$

3) $\forall x \in V : \langle x, 0 \rangle = 0 = \langle 0, x \rangle$.

4) $\big(\forall y \in V : \langle x, y \rangle = 0\big) \;\Rightarrow\; x = 0$ (Setze $y := x$, siehe (SP1)).

3.64 Satz (Cauchy-Schwarz-Bunjakowski Ungleichung (CSB)): Für $x, y \in V$ gilt

$$|\langle x, y \rangle| \;\leq\; \sqrt{\langle x, x \rangle} \cdot \sqrt{\langle y, y \rangle}.$$

Beweis: Fall $y = 0 : \Rightarrow\; 0 \leq 0$ stimmt.

Fall $y \neq 0$: Setze $\lambda := -\dfrac{\langle x, y \rangle}{\langle y, y \rangle}$.

$$\begin{aligned}
\text{(SP1)} \Rightarrow \quad 0 \;&\leq\; \langle y, y \rangle \langle x + \lambda y, x + \lambda y \rangle \\
&=\; \langle y, y \rangle \big(\langle x, x \rangle + \lambda \langle y, x \rangle + \overline{\lambda} \langle x, y \rangle + |\lambda|^2 \langle y, y \rangle \big) \\
&=\; \langle y, y \rangle \langle x, x \rangle - |\langle x, y \rangle|^2 - |\langle x, y \rangle|^2 + |\langle x, y \rangle|^2 \\
&=\; \langle y, y \rangle \langle x, x \rangle - |\langle x, y \rangle|^2 \\
\Rightarrow \quad & |\langle x, y \rangle|^2 \leq \langle y, y \rangle \langle x, x \rangle
\end{aligned}$$

□

3.65 Satz: Ist $(V, \langle ., . \rangle)$ ein Vektorraum mit Skalarprodukt, so definiert

$$\|x\|_2 \;:=\; \sqrt{\langle x, x \rangle}$$

eine Norm auf V. Die Cauchy-Schwarz-Bunjakowski-Ungleichung wird zu

$$|\langle x, y \rangle| \;\leq\; \|x\|_2 \cdot \|y\|_2.$$

Beweis: **(N1)** $\|x\|_2 \geq 0$ stimmt, $\|x\|_2 = 0 \;\Leftrightarrow\; \langle x, x \rangle = 0 \;\Leftrightarrow\; x = 0$.

(N2) $\|\lambda \cdot x\|_2 = \sqrt{\langle \lambda \cdot x, \lambda \cdot x \rangle} = \sqrt{|\lambda|^2 \langle x, x \rangle} = |\lambda|\, \|x\|_2$.

(N3)
$$\begin{aligned}
\|x + y\|_2^2 \;&=\; \langle x + y, x + y \rangle \;=\; \langle x, x \rangle + \langle x, y \rangle + \langle y, x \rangle + \langle y, y \rangle \\
&=\; \|x\|_2^2 + \|y\|_2^2 + 2\mathrm{Re}\, \langle x, y \rangle \\
&\leq\; \|x\|_2^2 + \|y\|_2^2 + 2 |\langle x, y \rangle| \\
&\leq\; \|x\|_2^2 + \|y\|_2^2 + 2 \|x\|_2 \, \|y\|_2 \\
&=\; \big(\|x\|_2 + \|y\|_2\big)^2
\end{aligned}$$

□

3.9 Orthogonalität

3.66 Definition: Sei V ein Vektorraum mit Skalarprodukt $\langle .,. \rangle$.

1) $x, y \in V$ heißen **orthogonal** ($x \perp y$), falls $\langle x, y \rangle = 0$.

2) Eine Familie $(v_i)_{i \in I} \subseteq V$ (I bezeichnet irgendeine Menge von Indices) mit $v_i \neq 0$ für $i \in I$ heißt **Orthogonalsystem**, falls $v_i \perp v_j$ für $i \neq j$, und **Orthonormalsystem** (ONS), falls

$$\langle v_i, v_j \rangle = \delta_{ij} = \begin{cases} 0 & \text{für } i \neq j \\ 1 & \text{für } i = j. \end{cases}$$

Ist ein ONS gleichzeitig Basis von V, so heißt es **Orthonormalbasis** (ONB).

3.67 Satz des Pythagoras: Sei $x \perp y$. Dann gilt

$$\|x + y\|_2^2 = \|x\|_2^2 + \|y\|_2^2 .$$

3.68 Satz: Ist $(v_i)_{i \in I}$ ein Orthogonalsystem, so ist $\{v_i : i \in I\}$ linear unabhängig.

Beweis: Sei $J \subseteq I$, J endlich und $\sum\limits_{j \in J} \lambda_j \cdot v_j = 0$. Für $k \in J$ folgt

$$0 = \langle 0, v_k \rangle = \Big\langle \sum_{j \in J} \lambda_j \cdot v_j, v_k \Big\rangle = \sum_{j \in J} \lambda_j \underbrace{\langle v_j, v_k \rangle}_{=0 \text{ für } j \neq k} = \lambda_k \|v_k\|_2^2 \Rightarrow \lambda_k = 0.$$

$\square$

3.69 Satz: Sei $(V, \langle .,. \rangle)$ Vektorraum mit Skalarprodukt, $\{v_1, v_2, \ldots\} \subseteq V$ endliche oder abzählbare linear unabhängige Menge. Dann gibt es ein ONS $\{e_1, e_2, \ldots\}$ mit der Eigenschaft

$$\forall k \in \mathbb{N} \text{ mit } k \leq \sharp\{v_1, v_2, \ldots\} : \mathrm{LH}\{v_1, \ldots, v_k\} = \mathrm{LH}\{e_1, \ldots, e_k\}$$

Insbesondere gilt: Ist $\dim V < \infty$, so besitzt V eine ONB, und jedes ONS lässt sich zu einer ONB ergänzen.

Beweis: Konstruktiv, **Gram-Schmidtsches Orthogonalisierungsverfahren**:

Schritt 1: $e_1 := \dfrac{1}{\|v_1\|_2} \cdot v_1$

Schritt 2: $f_2 := v_2 - \langle v_2, e_1 \rangle \cdot e_1 \neq 0$, da $v_2 \notin \mathrm{LH}\{v_1\} = \mathrm{LH}\{e_1\}$

Es gilt $f_2 \perp e_1 : \langle f_2, e_1 \rangle = \langle v_2, e_1 \rangle - \langle v_2, e_1 \rangle \underbrace{\langle e_1, e_1 \rangle}_{=1} = 0$

$$\text{Setze } e_2 := \frac{1}{\|f_2\|_2} \cdot f_2 \;\Rightarrow\; \begin{cases} \{e_1, e_2\} \text{ ist ONS} \\ \mathrm{LH}(\{e_1, e_2\} = \mathrm{LH}(\{v_1, v_2\} \end{cases}$$

$\vdots$

Schritt k: Sei $\{e_1, \ldots, e_{k-1}\}$ bereits konstruiert.

Dann gilt $v_k \notin \mathrm{LH}\{v_1, \ldots, v_{k-1}\} = \mathrm{LH}\{e_1, \ldots, e_{k-1}\}$

$\Rightarrow\; f_k := v_k - \sum\limits_{j=1}^{k-1} \langle v_k, e_j \rangle \cdot e_j \neq 0$

Es gilt $f_k \perp e_j$ für $j = 1, \ldots, k-1$:

$$\langle f_k, e_j \rangle = \langle v_k, e_j \rangle - \sum_{i=1}^{k-1} \langle v_k, e_i \rangle \underbrace{\langle e_i, e_j \rangle}_{=\delta_{ij}} = \langle v_k, e_j \rangle - \langle v_k, e_j \rangle = 0.$$

$$\text{Setze } e_k := \frac{1}{\|f_k\|_2} \cdot f_k \;\Rightarrow\; \begin{cases} \{e_1, \ldots, e_k\} \text{ ist ONS} \\ \mathrm{LH}(\{e_1, \ldots, e_k\} = \mathrm{LH}(\{v_1, \ldots, v_k\} \end{cases}$$

□

3.70 Koordinaten bezüglich ONB: Sei $B = \{e_1, \ldots, e_n\}$ ONB von V. Dann:

1) Für alle $v \in V$ gilt

$$v \;=\; \sum_{j=1}^{n} \langle v, e_j \rangle \cdot e_j, \quad \text{d.h. } v_B = \begin{pmatrix} \langle v, e_1 \rangle \\ \vdots \\ \langle v, e_n \rangle \end{pmatrix}$$

d.h. $\langle v, e_1 \rangle, \ldots, \langle v, e_n \rangle$ sind die Koordinaten von v bezüglich $\{e_1, \ldots, e_n\}$.

2) $\forall u, v \in V : \langle u, v \rangle \;=\; \langle u_B, v_B \rangle_{\mathbb{K}^n} \;=\; \sum\limits_{j=1}^{n} \langle u, e_j \rangle \overline{\langle v, e_j \rangle} \;=\; \sum\limits_{j=1}^{n} \langle u, e_j \rangle \langle e_j, v \rangle$

Insbesondere gilt

$$\|u\|_2^2 \;=\; \langle u, u \rangle \;=\; \sum_{j=1}^{n} |\lambda_j|^2 \;=\; \sum_{j=1}^{n} \big| \langle u, e_j \rangle \big|^2$$

(**Parsevalsche Gleichung**).

Beweis: **1)** B Basis $\Rightarrow$ $v = \sum\limits_{j=1}^{n} \lambda_j \cdot e_j$.

$$\Rightarrow \text{ für } k = 1, \ldots, n : \; \langle v, e_k \rangle = \sum_{j=1}^{n} \lambda_j \langle e_j, e_k \rangle = \lambda_k.$$

2) $$\langle u, v\rangle = \left\langle \sum_{j=1}^{n} \lambda_j \cdot e_j, \sum_{k=1}^{n} \mu_k \cdot e_k \right\rangle = \sum_{j=1}^{n}\sum_{k=1}^{n} \lambda_j \overline{\mu_k} \langle e_j, e_k\rangle = \sum_{j=1}^{n} \lambda_j \overline{\mu_j}$$ □

3.71 Folgerung: Der Isomorphismus

$$L : \mathbb{K}^n \to V : \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mapsto \sum_{j=1}^{n} x_j \cdot e_j$$

erhält das Skalarprodukt:

$$\langle x, y\rangle_{\mathbb{K}^n} = \langle L\,x, L\,y\rangle_V ,$$

und damit auch die zugehörige Norm. $\mathbb{K}^n$ und V können als identisch angesehen werden.

Beispiele: **1)** $V = \mathbb{R}^2$, $B = \{e_1, e_2\} = \left\{ \begin{pmatrix} \cos\varphi \\ \sin\varphi \end{pmatrix}, \begin{pmatrix} -\sin\varphi \\ \cos\varphi \end{pmatrix} \right\}$ mit festem $\varphi \in \mathbb{R}$.

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}_E \quad \Rightarrow \quad x = \begin{pmatrix} x_1 \cos\varphi + x_2 \sin\varphi \\ -x_1 \sin\varphi + x_2 \cos\varphi \end{pmatrix}_B$$

2) $V = C([0, 2\pi]) := \{f : [0, 2\pi] \to \mathbb{C} \mid f \text{ ist stetig}\}$.

$$\langle f, g\rangle := \int_0^{2\pi} f(x)\,\overline{g(x)}\,\mathrm{d}x \quad \text{ist Skalarprodukt}$$
$$\|f\|_2 := \left(\int_0^{2\pi} |f(x)|^2\,\mathrm{d}x \right)^{1/2} \quad \text{zugehörige Norm}$$

Setze

$$\begin{aligned} e_0 : &\quad x \mapsto \frac{1}{\sqrt{2\pi}} \\ e_{2n-1} : &\quad x \mapsto \frac{1}{\sqrt{\pi}} \sin(nx) \\ e_{2n} : &\quad x \mapsto \frac{1}{\sqrt{\pi}} \cos(nx) \end{aligned}$$

Nachrechnen: $\{e_0, e_1, \ldots\}$ ist ONS. Für $f \in \mathrm{LH}\{e_0, e_1, \ldots\}$ gilt

$$f = \sum_n \langle f, e_n\rangle \cdot e_n = b_0 + \sum_{n=1}^{\infty} a_n \sin(nx) + \sum_{n=1}^{\infty} b_n \cos(nx)$$
$$b_0 = \frac{1}{2\pi} \int_0^{2\pi} f(x)\,\mathrm{d}x$$
$$b_n = \frac{1}{\pi} \int_0^{2\pi} f(x) \cos(nx)\,\mathrm{d}x$$
$$a_n = \frac{1}{\pi} \int_0^{2\pi} f(x) \sin(nx)\,\mathrm{d}x$$

Lässt man auch „unendliche Summen“ zu, so kann man praktisch jede Funktion f so entwickeln: Theorie der Fourierreihen.

3.72 Satz: Seien V ein Vektorraum mit Skalarprodukt, U ein endlichdimensionaler Unterraum, $\{e_1,\dots,e_k\}$ ONB von U und $x\in V$ fest. Für $y\in U$ sind äquivalent

(i) $y=\sum_{j=1}^{k}\langle x,e_j\rangle\cdot e_j$

(ii) $x=y+z$ mit $z\perp U$ (d.h. $\forall\tilde y\in U: z\perp\tilde y$)

(iii) $\forall\tilde y\in U: \|x-y\|_2\le\|x-\tilde y\|_2$.

D.h.: y ist das Element von U, das zu x den kleinsten Abstand hat.

Beweis: **(i) $\Rightarrow$ (ii):** Sei $z:=x-y$ und $\widetilde y\in U,\ \widetilde y=\sum_{j=1}^{k}\lambda_j\cdot e_j$.

$$\begin{aligned}\Rightarrow\ \langle z,\widetilde y\rangle &= \langle x-y,\widetilde y\rangle = \langle x,\widetilde y\rangle-\langle y,\widetilde y\rangle\\ &= \sum_{j=1}^{k}\overline{\lambda_j}\,\langle x,e_j\rangle-\sum_{j=1}^{k}\langle x,e_j\rangle\sum_{l=1}^{k}\langle e_j,\lambda_l\cdot e_l\rangle\\ &= \sum_{j=1}^{k}\overline{\lambda_j}\,\langle x,e_j\rangle-\sum_{j=1}^{k}\langle x,e_j\rangle\,\overline{\lambda_j}\cdot 1\\ &= 0\end{aligned}$$

(ii) $\Rightarrow$ (iii): Sei $z:=x-y\perp U,\ \widetilde y\in U$.

$$\begin{aligned}\Rightarrow\ \|x-\widetilde y\|_2^2 &= \|\underbrace{x-y}_{\perp U}+\underbrace{y-\widetilde y}_{\in U}\|_2^2\\ &= \|x-y\|_2^2+\|y-\widetilde y\|_2^2\\ &\ge \|x-y\|_2^2\end{aligned}$$

(iii) $\Rightarrow$ (i): Fall $x\in U$: Dann $y=x=\sum_{j=1}^{k}\langle x,e_j\rangle\cdot e_j$

Fall $x\notin U$: Sei $\{e_1,\dots,e_k,e_{k+1}\}$ ONB von $\mathrm{LH}\{e_1,\dots,e_k,x\}$. Es gilt

$$x=\sum_{j=1}^{k+1}\langle x,e_j\rangle\cdot e_j,\quad y=\sum_{j=1}^{k+1}\lambda_j\cdot e_j \text{ mit } \lambda_{k+1}=0.$$

$$\begin{aligned}\Rightarrow\ \|x-y\|_2^2 &= \left\|\sum_{j=1}^{k+1}(\langle x,e_j\rangle-\lambda_j)\cdot e_j\right\|_2^2\\ &= \sum_{j=1}^{k+1}\left|\,\langle x,e_j\rangle-\lambda_j\right|^2\\ &= \text{Minimum, falls } \lambda_j=\langle x,e_j\rangle \text{ für } 1\le j\le k\end{aligned}$$

$\square$

3.73 Bemerkung: Aus dem Satz: Für $x \in V\ \exists!\ y \in U : \|x - y\|_2$ ist minimal. Es gilt

$$y = \sum_{j=1}^{k} \langle x, e_j \rangle \cdot e_j.$$

Die Abbildung $P : V \to U : x \mapsto \sum_{j=1}^{k} \langle x, e_j \rangle \cdot e_j$ ist linear und heißt **orthogonale Projektion** von V auf U.

3.10 Die adjungierte Matrix

3.74 Definition: Sei

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$$

eine $m \times n$-Matrix. Dann heißt die $n \times m$-Matrix

$$A^* := \begin{pmatrix} \overline{a_{11}} & \dots & \overline{a_{m1}} \\ \vdots & & \vdots \\ \overline{a_{1n}} & \dots & \overline{a_{mn}} \end{pmatrix}$$

die **adjungierte Matrix** zu A und die $n \times m$-Matrix

$$A^T := \begin{pmatrix} a_{11} & \dots & a_{m1} \\ \vdots & & \vdots \\ a_{1n} & \dots & a_{mn} \end{pmatrix}$$

die **transponierter Matrix** von A. Im Fall $\mathbb{K} = \mathbb{R}$ gilt $A^* = A^T$.

3.75 Definition: Eine $n \times n$-Matrix A heißt

1) **selbstadjungiert**, falls $A^* = A$,

2) **symmetrisch**, falls $A^T = A$,

3) **unitär**, falls $A^* = A^{-1}$. Im Fall $\mathbb{K} = \mathbb{R}$ heißt A dann auch **orthogonal** $(A^T = A^{-1})$.

3.76 Satz: Seien B, C zwei ONB's von V. Dann ist $M_{\mathrm{Id}}^{C,B}$ unitär, d.h. die inverse Matrix ist besonders einfach zu berechnen: $M_{\mathrm{Id}}^{B,C} = \left(M_{\mathrm{Id}}^{C,B}\right)^*$.

Beweis: Für $V = \mathbb{C}^n$ und $C = E =$ kanonische Basis. Sei $B = \{b_1, \ldots, b_n\}$.

$$M_{\mathrm{Id}}^{E,B} = \begin{pmatrix} (b_1)_1 & \ldots & (b_n)_1 \\ \vdots & & \vdots \\ \underbrace{(b_1)_n}_{\substack{\text{Koord.}\\ \text{von } b_1\\ \text{bzgl. } E}} & & \underbrace{(b_n)_n}_{\substack{\text{Koord.}\\ \text{von } b_n\\ \text{bzgl. } E}} \end{pmatrix} = \big((b_1)_E, \ldots, (b_n)_E\big)$$

$$\begin{aligned} \Rightarrow \left(M_{\mathrm{Id}}^{E,B}\right)^* \cdot M_{\mathrm{Id}}^{E,B} &= \begin{pmatrix} (b_1)_E^* \\ \vdots \\ (b_n)_E^* \end{pmatrix} \cdot \begin{pmatrix} (b_1)_E & \ldots & (b_n)_E \end{pmatrix} \\ &= \begin{pmatrix} \langle b_1, b_1\rangle & \langle b_1, b_2\rangle & \ldots & \langle b_1, b_n\rangle \\ \langle b_2, b_1\rangle & & & \\ \vdots & & & \\ \langle b_n, b_1\rangle & \ldots & \ldots & \langle b_n, b_n\rangle \end{pmatrix} \\ &= E \quad \text{(Einheitsmatrix)} \end{aligned}$$

$\square$

3.77 Satz: Seien $\langle .\,,.\rangle_{\mathbb{K}^n}$, $\langle .\,,.\rangle_{\mathbb{K}^m}$ die Standard-Skalarprodukte in $\mathbb{K}^n$ bzw. $\mathbb{K}^m$, und sei A eine $m \times n$-Matrix. Für $x \in \mathbb{K}^n$, $y \in \mathbb{K}^m$ gilt

$$\langle A \cdot x, y\rangle_{\mathbb{K}^m} = \langle x, A^* \cdot y\rangle_{\mathbb{K}^n}\,.$$

Beweis: j-te Koordinate von Ax: $(A \cdot x)_j = \sum_{k=1}^{n} a_{jk}\, x_k$

$$\Rightarrow \langle Ax, y\rangle_{\mathbb{K}^m} = \sum_{j=1}^{m} \left(\sum_{k=1}^{n} a_{jk}\, x_k\right) \overline{y_j}.$$

Genauso

$$\begin{aligned} \langle x, A^* \cdot y\rangle_{\mathbb{K}^n} &= \sum_{k=1}^{n} x_k \overline{(A^* \cdot y)_k} \\ &= \sum_{k=1}^{n} x_k \overline{\left(\sum_{j=1}^{m} \overline{a_{jk}}\, y_j\right)} \\ &= \sum_{j=1}^{m} \sum_{k=1}^{n} \overline{\overline{a_{jk}}}\, x_k\, \overline{y_j} \\ &= \langle A \cdot x, y\rangle_{\mathbb{K}^m} \end{aligned}$$

$\square$

3.78 Folgerungen: **1)** Ist A selbstadjungierte $n \times n$-Matrix, so gilt für $x, y \in \mathbb{K}^n$:

$$\langle A \cdot x, y \rangle \; = \; \langle x, A \cdot y \rangle \, .$$

2) Ist A unitär, so gilt für $x, y \in \mathbb{K}^n$

$$\langle A \cdot x, A \cdot y \rangle \; = \; \langle x, y \rangle \, .$$

Insbesondere gilt $\|Ax\|_2 = \sqrt{\langle A \cdot x, A \cdot x \rangle} = \|x\|_2$. Eine unitäre Matrix erhält Längen und Winkel.

3.11 Die adjungierte Abbildung

3.79 Satz und Definition: Sei $L : V \to W$. Dann existiert genau eine lineare Abbildung $L^* : W \to V$, so dass

$$\forall v \in V \quad \forall w \in W : \langle L\, v, w \rangle_W = \langle v, L^*\, w \rangle_V \, .$$

L^* heißt die **adjungierte Abbildung** zu L.
Ist $B = \{b_1, \ldots, b_n\}$ ONB von V und $C = \{c_1, \ldots, c_m\}$ ONB von W, so gilt

$$M_{L^*}^{B,C} \; = \; \left(M_L^{C,B} \right)^* ,$$

d.h. zur adjungierten Abbildung gehört die adjungierte Matrix.

Beweis: **Existenz:** Sei $(Lv)_C = M_L^{C,B} \cdot v_B$. Offensichtlich ist $M_L^{C,B}$ eine $m \times n$-Matrix und $\left(M_L^{C,B} \right)^*$ eine $n \times m$-Matrix. Definiere

$$L^* : W \to V : (L^* w)_B := \left(M_L^{C,B} \right)^* (w)_C .$$

Dann gilt $M_{L^*}^{B,C} = \left(M_L^{C,B} \right)^*$ und

$$\begin{aligned}
\langle v, L^* w \rangle &= \langle v_B, (L^* w)_B \rangle_{\mathbb{K}^n} \\
&= \left\langle v_B, \left(M_L^{C,B} \right)^* \cdot w_C \right\rangle_{\mathbb{K}^n} \\
&= \left\langle M_L^{C,B} \cdot v_B, w_C \right\rangle_{\mathbb{K}^m} \\
&= \langle (Lv)_C, w_C \rangle_{\mathbb{K}^m} \\
&= \langle Lv, w \rangle_W
\end{aligned}$$

Eindeutigkeit: Sei $K : W \to V$ mit $\langle Lv, w \rangle = \langle v, Kw \rangle_V$

$$\begin{aligned}
&\Rightarrow \; \langle Kw - L^* w, v \rangle \; = \; \langle w, Lv \rangle - \langle w, Lv \rangle \; = \; 0 \quad \forall v \in V \\
&\Rightarrow \; Kw - L^* w = 0.
\end{aligned}$$

□

3.80 Eigenschaften der Adjungierten: Es sei $L : V \to W,\ K : W \to U$. Dann gelten

1) $(L^*)^* = L$.

2) $(K \circ L)^* = L^* \circ K^*$ (Reihenfolge dreht sich um!).

3) $\mathrm{Kern}\,(L) = (\mathrm{Bild}\,(L^*))^\perp := \{x \in V \mid \forall y \in \mathrm{Bild}\,(L^*) : \langle x, y\rangle = 0\}$.

4) Falls $L : V \to V$, gilt $\mathrm{Rang}\,(L) = \mathrm{Rang}\,(L^*)$.

Beweis: **1)** Für $x \in V,\ y \in W$ gilt $\langle x, (L^*)^* y\rangle_V = \langle L^* x, y\rangle_W = \langle x, Ly\rangle_V$

$\Rightarrow\ \forall x \in V : \langle x, Ly - (L^*)^* y\rangle = 0$

$\overset{3.63}{\Rightarrow}\ (L^*)^* y = Ly$ für beliebiges $y \in W$.

2) $\langle (K \circ L)x, y\rangle_U = \langle K(L(x)), y\rangle_U = \langle Lx, K^* y\rangle_W = \langle x, L^*(K^* y)\rangle_V = \langle x, L^* \circ K^* y\rangle_V$.

Andererseits: $\langle (K \circ L)x, y\rangle_U = \langle x, (K \circ L)^* y\rangle_V$

Wie oben folgt dann die Behauptung.

3)
$$\begin{aligned} x \in \mathrm{Kern}\,(L) &\Leftrightarrow Lx = 0 \\ &\Leftrightarrow \forall y \in W : \langle Lx, y\rangle = 0 \\ &\Leftrightarrow \forall y \in W : \langle x, L^* y\rangle = 0 \\ &\Leftrightarrow x \in (\mathrm{Bild}\,(L^*))^\perp \end{aligned}$$

4) Dimensionsformel: $\dim V = \dim\ \mathrm{Kern}\,(L) + \dim\ \mathrm{Bild}\,(L)$.

$$\begin{aligned} \Rightarrow\ \mathrm{Rang}\,(L) &= \dim\ \mathrm{Bild}\,(L) = \dim(V) - \dim\ \mathrm{Kern}\,(L) \\ &= \dim(V) - \dim\big(\mathrm{Bild}\,(L^*)^\perp\big) \\ &= \dim(V) - \big(\dim(V) - \dim\ \mathrm{Bild}\,(L^*)\big) \\ &= \dim\ \mathrm{Bild}\,(L^*) = \mathrm{Rang}\,(L^*). \end{aligned}$$

□

3.81 Definition: Eine Abbildung $L : V \to V$ heißt

1) **normal**, falls $L \circ L^* = L^* \circ L$.

2) **unitär**, falls $L^* = L^{-1}$. Ist L unitär und gilt $\mathbb{K} = \mathbb{R}$, so heißt L auch **orthogonal**.

3) **selbstadjungiert**, falls $L^* = L$. Ist L selbstadjungiert und gilt $\mathbb{K} = \mathbb{R}$, so heißt L auch **symmetrisch**.

3.82 Eigenschaften unitärer Abbildungen: Für $L : V \to V$ linear sind äquivalent:

(i) L ist unitär.

(ii) $\forall u, v \in V : \langle L\,u, L\,v\rangle = \langle u, v\rangle$. Insbesondere ist L längen- und winkelerhaltend.

(iii) L ist **isometrisch**, d.h $\forall u \in V : \|L\,u\|_2 = \|u\|_2$.

(iv) L transformiert ONBen in ONBen.

(v) Es existiert eine ONB B, so dass die Spalten von $M_L^{B,B}$ in $\mathbb{K}^n$ eine ONB bilden ($n = \dim(V)$).

(vi) Für jede ONB B bilden die Spalten von $M_L^{B,B}$ eine ONB in $\mathbb{K}^n$.

3.12 Determinanten

3.83 Definition: Seien $x, y \in \mathbb{R}^2$. Die **orientierte Fläche** des von x, y aufgespannten Parallelogramms ist definiert durch

$$F(x, y) \ := \ x_1\,y_2 - x_2\,y_1 \ = \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix}$$

3.84 Folgerung: Die orientierte Fläche besitzt folgende Eigenschaften:

1) $F(x, y) = 0 \ \Leftrightarrow \ \{x, y\}$ ist linear abhängig.

2) $F(e_1, e_2) = 1$.

3) $F(\lambda \cdot x, y) = \lambda\,F(x, y)$, $F(x, \lambda \cdot y) = \lambda\,F(x, y)$, $F(\lambda \cdot x, \lambda \cdot y) = \lambda^2\,F(x, y)$.

4) $F(x + z, y) = F(x, y) + F(z, y)$, $F(x, y + z) = F(x, y) + F(x, z)$.

5) $F(y, x) = -F(x, y)$.

3.85 Definition: Seien $x, y, z \in \mathbb{R}^3$. Dann heißt

$$\begin{aligned} V(x, y, z) \ &:= \ x_1\,y_2\,z_3 + y_1\,z_2\,x_3 + z_1\,x_2\,y_3 - z_1\,y_2\,x_3 - x_1\,z_z\,y_3 - y_1\,x_2\,z_3 \\ &= \left|\begin{array}{ccc|cc} x_1 & y_1 & z_1 & x_1 & y_1 \\ x_2 & y_2 & z_2 & x_2 & y_2 \\ x_3 & y_3 & z_3 & x_3 & y_3 \end{array}\right. \end{aligned}$$

das **orientierte Volumen** des von x, y, z aufgespannten Parallelepipeds.

3.86 Definition: Es sei $M_{n,n}(\mathbb{K})$ die Menge der $n \times n$-Matrizen mit Elementen in $\mathbb{K}$. Eine Abbildung $\det : M_{n,n}(\mathbb{K}) \to \mathbb{K} : A \mapsto \det(A) =: |A|$ heißt **Determinante**, falls sie folgende Eigenschaften besitzt: Ist

$$A \;=\; \begin{pmatrix} a_1 & \dots & a_n \end{pmatrix} \qquad \text{mit Spaltenvektoren } a_j \in \mathbb{K}^n,$$

so gelten:

(D1) $\det(E_n) = 1$.

(D2) Für jedes $k \in \{1, \dots, n\}$ gilt:

$$\begin{aligned} &\det\big(a_1 \dots a_{k-1}\ (\alpha \cdot a_k + \beta \cdot b)\ a_{k+1} \dots a_n\big) = \\ &\quad = \alpha \det(a_1 \dots a_{k-1}\ a_k\ a_{k+1} \dots a_n\big) + \beta \det(a_1 \dots a_{k-1}\ b\ a_{k+1} \dots a_n\big). \end{aligned}$$

D.h. die Abbildung det ist in jedem Argument linear.

(D3) Für alle Paare (k, j) mit $k \neq j$ gilt:

$$\det(a_1 \dots a_k \dots a_j \dots a_n\big) = (-1) \det(a_1 \dots a_j \dots a_k \dots a_n\big),$$

d.h. Vertauschung zweier Spalten ändert das Vorzeichen.

3.87 Bemerkung: Man kann beweisen, dass es genau eine Determinante $\det : M_{n,n}(\mathbb{K}) \to \mathbb{K}$ mit diesen Eigenschaften gibt.

3.88 Satz: Es sei A eine $n \times n$-Matrix. Für die Determinante gelten:

1) Verschwindet eine Spalte: $a_k = 0$ für ein k, so gilt $\det(A) = 0$.

2) Enthält A zwei gleiche Spalten: $a_k = a_j$ für ein Paar (k, j) mit $k \neq j$, so gilt $\det(A) = 0$.

3) $\det \begin{pmatrix} a_{11} & 0 & \dots & 0 \\ 0 & a_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & a_{nn} \end{pmatrix} = a_{11}\, a_{22} \cdots a_{nn}$.

4) Addition des Vielfachen einer Spalte zu einer anderen ändert nicht den Wert der Determinante.

3.89 Satz: Es sei $A = (a_1 \dots a_n)$ eine $n \times n$-Matrix mit den Spaltenvektoren a_j. Dann gilt:

$$\{a_1, \dots, a_n\} \text{ ist linear abhängig } \Leftrightarrow \det(A) = 0.$$

Oder anders ausgedrückt: Es gilt

$$\mathrm{Rang}\,(A) < n \;\Leftrightarrow\; \det(A) = 0 \quad \text{bzw.} \quad \mathrm{Rang}\,(A) = n \;\Leftrightarrow\; \det(A) \neq 0.$$

Beweis: „$\Rightarrow$“: Es sei $\lambda_1 \cdot a_1 + \dots + \lambda_n \cdot a_n = 0$. OBdA $\lambda_1 \neq 0 \;\Rightarrow\; a_1 + \frac{\lambda_2}{\lambda_1} \cdot a_2 + \dots + \frac{\lambda_n}{\lambda_1} \cdot a_n = 0$

$$\Rightarrow \det(A) = \det\Bigg(\underbrace{\left(a_1 + \frac{\lambda_2}{\lambda_1} \cdot a_2 + \dots + \frac{\lambda_n}{\lambda_1} \cdot a_n\right)}_{=0} \; a_2 \; \dots \; a_n\Bigg) = 0$$

„$\Leftarrow$“: Kontraposition: Zeige $\{a_1, \dots, a_n\}$ linear unabhängig $\Rightarrow \det(A) \neq 0$.

Sei $\{a_1, \dots, a_n\}$ linear unabhängig

$$\begin{aligned}
\Rightarrow \quad e_1 \quad &= \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} = \lambda_1 \cdot a_1 + \dots + \lambda_n \cdot a_n, \quad \text{o.B.d.A. } \lambda_1 \neq 0 \\
\Rightarrow \quad \det(A) &\overset{(D2)}{=} \frac{1}{\lambda_1} \det\,(\lambda_1 \cdot a_1 \; a_2 \; a_3 \; \dots \; a_n) \\
&\overset{3.88,\,4)}{=} \frac{1}{\lambda_1} \det\,(\lambda_1 \cdot a_1 + \dots + \lambda_n \cdot a_n \; a_2 \; a_3 \; \dots \; a_n) \\
&= \frac{1}{\lambda_1} \det\,(e_1 \; a_2 \; a_3 \; \dots \; a_n) \\
&\;\vdots \\
&= \frac{1}{\lambda_1 \cdot \lambda_2 \cdots \lambda_n} \det(e_1 \; e_2 \; \dots \; e_n) \\
&\neq 0
\end{aligned}$$

$\square$

3.90 Lineare Gleichungssysteme: Es sei A eine $n \times n$-Matrix. Dann gelten:

1) Das LGS $A \cdot x = 0$ besitzt nichttriviale Lösungen genau dann, wenn $\det(A) = 0$.

2) Das LGS $A \cdot x = b$ ist für jedes $b \in \mathbb{K}^n$ eindeutig lösbar genau dann, wenn $\det(A) \neq 0$.

Beweis: **1)** $\det(A) = 0 \;\Leftrightarrow\; \mathrm{Rang}\,(A) < n \;\Leftrightarrow\;$ LGS $Ax = 0$ besitzt mehr als eine Lösung.

2) $\det(A) \neq 0 \;\Leftrightarrow\; \mathrm{Rang}\,(A) = n \;\Leftrightarrow\;$ LGS $Ax = b$ ist für jedes $b \in \mathbb{K}^n$ eindeutig lösbar. $\square$

3.91 Rechenregeln: **1)** $\det(A \cdot B) = \big(\det(A)\big) \det(B)$.

2) Falls A invertierbar ist: $\det(A^{-1}) = \dfrac{1}{\det(A)}$.

3) $\det(A^T) = \det(A)$.
Insbesondere: Addition des Vielfachen einer Zeile zu einer anderen Zeile ändert nicht den Wert der Determinante.

3.92 Laplace-Entwicklung: Für $A = (\alpha_{ij}) \in M_{n,n}$ setze

$$A_{ij} \;:=\; \begin{pmatrix} & \alpha_{1j} & \\ \alpha_{i1} & & \alpha_{in} \\ & \alpha_{nj} & \end{pmatrix} \;\in\; M_{n-1,n-1}$$

(Streichen der i-ten Zeile und j-ten Spalte in A).

Für jedes $i \in \{1, \dots, n\}$ gilt

$$\det A \;=\; \sum_{j=1}^{n} (-1)^{i+j} \alpha_{ij} \det A_{ij}$$

(Entwicklung nach i-ter Zeile), und für jedes $j \in \{1, \dots, n\}$ gilt

$$\det A \;=\; \sum_{i=1}^{n} (-1)^{i+j} \alpha_{ij} \det A_{ij}$$

(Entwicklung nach j-ter Spalte).

3.93 Geometrische Bedeutung: **1)** Für $v_1, \dots, v_n \in \mathbb{R}^n$ ist

$$V(v_1, \dots, v_n) \;:=\; \det \big(v_1 \;\dots\; v_n\big)$$

das orientierte Volumen des von den Vektoren $v_1, \dots, v_n$ aufgespannten Parallelepipeds.

2) Sei A eine reelle $n \times n$-Matrix. Betrachte $L : \mathbb{R}^n \to \mathbb{R}^n : x \mapsto A \cdot x$. Sind $v_1, \dots, v_n \in \mathbb{R}^n$, so gilt

$$V(L\,v_1, \dots, L\,v_n) \;=\; \det(A)\, V(v_1, \dots, v_n),$$

d.h. $\det(A)$ gibt an, wie sich das Volumen unter der Abbildung L verändert.

3.94 Definition: Sei $L : V \to V$ linear und B Basis von V. Dann ist die **Determinante** von L definiert durch

$$|L| \;:=\; \det(L) \;:=\; \left| M_L^{B,B} \right|.$$

3.95 Bemerkung: Sind B, B' Basen von V und ist $L : V \to V$ linear, so gilt

$$\left|M_L^{B,B}\right| = \left|M_L^{B',B'}\right|.$$

Beweis: Sei $M_{\mathrm{Id}}^{B',B}$ die Basiswechselmatrix. Dann gilt:

$$\begin{aligned}
\left|M_L^{B',B'}\right| &= \left|M_{\mathrm{Id}}^{B',B} \cdot M_L^{B,B} \cdot M_{\mathrm{Id}}^{B,B'}\right| \\
&= \left|\underbrace{M_{\mathrm{Id}}^{B',B}}_{=\left(M_{\mathrm{Id}}^{B,B'}\right)^{-1}}\right| \cdot \left|M_L^{B,B}\right| \cdot \left|M_{\mathrm{Id}}^{B,B'}\right| \\
&= \frac{1}{\left|M_{\mathrm{Id}}^{B,B'}\right|} \cdot \left|M_L^{B,B}\right| \cdot \left|M_{\mathrm{Id}}^{B,B'}\right|
\end{aligned}$$

□

3.13 Diagonalisierung

3.96 Definition: Sei $L : V \to V$ linear.

1) $\lambda \in \mathbb{K}$ heißt **Eigenwert** (EW) von L, falls ein $v \in V$ mit $v \neq 0$ existiert, so dass $L\,v = \lambda \cdot v$. Der Vektor v heißt dann **Eigenvektor** (EV) zum Eigenwert λ (*englisch: eigenvalue, eigenvector*).

2) Die Menge $\sigma(L) := \big\{\lambda \in \mathbb{K} : \lambda \text{ ist EW von } L\big\}$ heißt **Spektrum** von L.

3) Für $\lambda \in \sigma(L)$ heißte der lineare Unterraum $E(\lambda) := \{v \in V : L(v) = \lambda \cdot v\}$ **Eigenraum** zum Eigenwert λ (*englisch: eigenspace*).

3.97 Satz: Sei $L : V \to V$ linear, $B = \{b_1, \ldots, b_n\}$ Basis von V. Äquivalent sind:

(i) B besteht nur aus Eigenvektoren.

(ii) Die Matrix $M_L^{B,B}$ hat **Diagonalgestalt**, d.h.

$$M_L^{B,B} = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

Sind (i) und (ii) erfüllt, so ist b_j Eigenvektor zum Eigenwert λ_j: $L(b_j) = \lambda_j \cdot b_j$, und es gilt

$$\begin{aligned} M_{L^k}^{B,B} &= \begin{pmatrix} \lambda_1^k & & 0 \\ & \ddots & \\ 0 & & \lambda_n^k \end{pmatrix} \\ M_{L^{-1}}^{B,B} &= \begin{pmatrix} \lambda_1^{-1} & & 0 \\ & \ddots & \\ 0 & & \lambda_n^{-1} \end{pmatrix} \quad \text{(falls } L \text{ invertierbar)} \\ \det(L) &= \lambda_1 \cdots \lambda_n \end{aligned}$$

Beweis: (i) $\Leftrightarrow\ \forall j \in \{1, \dots, n\}\ \exists \lambda_j \in \mathbb{K} : Lb_j = \lambda_j \cdot b_j \ \Leftrightarrow\ M_L^{B,B} = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$ □

3.98 Definition: $L : V \to V$ linear heißt **diagonalisierbar**, falls es eine Basis von V aus Eigenvektoren von L gibt.

3.99 Satz: Es sei $L : V \to V$ linear. Dann sind für $\lambda \in \mathbb{K}$ äquivalent:

(i) λ ist Eigenwert von L,

(ii) $\det(L - \lambda \cdot \mathrm{Id}) = 0$.

Beweis: Sei B Basis von V, $\dim(V) = n$. Dann:

$$\begin{aligned} \text{(i)} &\Leftrightarrow \exists v \in V \setminus \{0\} : Lv = \lambda \cdot v \\ &\Leftrightarrow \exists v \in V \setminus \{0\} : M_L^{B,B} \cdot v_B = \lambda \cdot v_B \\ &\Leftrightarrow \text{Das LGS } \left(M_L^{B,B} - \lambda \cdot E_n\right) v_B = 0 \text{ hat nichttriviale Lösungen} \\ &\Leftrightarrow \det\left(M_L^{B,B} - \lambda \cdot E_n\right) = 0 \\ &\Leftrightarrow \det(L - \lambda \cdot \mathrm{Id}) = 0 \end{aligned}$$

□

3.100 Satz und Definition: Sei $L : V \to V$ linear, $\dim(V) = n$.

1) $p_L : \mathbb{K} \ni \lambda \mapsto p_L(\lambda) := \det(L - \lambda \cdot \mathrm{Id})$ ist Polynom n-ten Grades in λ. p_L heißt **charakteristisches Polynom** von L.

2) Ist $A = M_L^{B,B}$ bezüglich einer Basis B von V, so ist

$$p_L(\lambda) = \det(A - \lambda \cdot E_n)$$

und hängt nicht von B ab.

3) Es gilt $p_L(\lambda) = (-1)^n\lambda^n + (-1)^{n-1}(\alpha_{11} + \ldots + \alpha_{nn})\lambda^{n-1} + \ldots + \det A$.
Die Summe der Diagonalelemente von A hängt also nicht von der Basis B ab und heißt **Spur** von L (oder auch Spur von A).
Schreibweise: $\mathrm{Sp}(L)$ oder $\mathrm{tr}(L)$ (trace) bzw. $\mathrm{Sp}(A)$ oder $\mathrm{tr}(A)$.

3.101 Folgerung: Für eine lineare Abbildung $L : V \to V$ gelten:

1) $\lambda \in \mathbb{K}$ ist Eigenwert von $L \quad \Leftrightarrow \quad p_L(\lambda) = 0$.

2) Falls $\mathbb{K} = \mathbb{C}$ besitzt L mindestens einen Eigenwert.

3) L hat höchstens n verschiedene Eigenwerte.

3.102 Definition: Sei $\lambda \in \mathbb{K}$ Eigenwert von L. Dann heißt
$n_g(\lambda) := \dim\{v \in V : L\,v = \lambda \cdot v\}$ die **geometrische Vielfachheit** von λ.
$n_a(\lambda) :=$ Ordnung der Nullstelle λ von p_L die **algebraische Vielfachheit** von λ.

3.103 Satz: Sei λ Eigenwert von L. Dann gilt $n_g(\lambda) \le n_a(\lambda)$.

Beweis: Wir bezeichnen den Eigenwert mit λ_0. Sei $\{b_1, \ldots, b_k\}$ eine Basis des Eigenraumes $\{v \in V : Lv = \lambda_0 \cdot v\}$ (also $k = n_g(\lambda_0)$). Ergänze zu einer Basis $B = \{b_1, \ldots, b_n\}$ von V.

$$\Rightarrow \quad M_L^{B,B} \quad = \quad \begin{pmatrix} \lambda_0 & 0 & \ldots & 0 & * \\ 0 & \lambda_0 & \ddots & \vdots & \\ \vdots & \ddots & \ddots & 0 & \\ 0 & \ldots & 0 & \lambda_0 & * \\ 0 & \ldots & 0 & 0 & \\ \vdots & & \vdots & \vdots & A \\ 0 & \ldots & 0 & 0 & \end{pmatrix}$$

$$\begin{aligned} \Rightarrow \quad p_L(\lambda) \quad &= \quad \det(M_L^{B,B} - \lambda \cdot E) \quad = \quad \det \begin{pmatrix} \lambda_0 - \lambda & & 0 & * \\ & \ddots & & \\ 0 & & \lambda_0 - \lambda & * \\ 0 & \ldots & 0 & A - \lambda \cdot E \end{pmatrix} \\ &= \quad (\lambda_0 - \lambda)^k \det(A - \lambda \cdot E) \end{aligned}$$

$\Rightarrow \lambda = \lambda_0$ ist mindestens k-fache Nullstelle von p_l, also $n_a(\lambda_0) \ge k = n_g(\lambda_0)$. $\square$

3.104 Satz: Sind $\lambda_1, \ldots, \lambda_k \in \mathbb{K}$ paarweise verschiedene Eigenwerte von L mit Eigenvektoren $v_1, \ldots, v_k$, so ist $\{v_1, \ldots, v_k\}$ linear unabhängig (Eigenvektoren zu verschiedenen Eigenwerten sind linear unabhängig).

Beweis: Induktion nach k:

Induktionsanfang $k = 1$: $\{v_1\}$ ist linear unabhängig, da $v_1 \neq 0$.

Induktionsschritt: Sei $\alpha_1 \cdot v_1 + \ldots + \alpha_{k+1} \cdot v_{k+1} = 0$.

$$\begin{aligned}
\Rightarrow \quad 0 &= (L - \lambda_{k+1} \cdot \mathrm{Id})(\alpha_1 \cdot v_1 + \ldots + \alpha_{k+1} \cdot v_{k+1}) \\
&= \alpha_1(\lambda_1 - \lambda_{k+1}) \cdot v_1 + \ldots + \alpha_k(\lambda_k - \lambda_{k+1}) \cdot v_k + 0 \\
\Rightarrow \quad & \alpha_1 = \ldots = \alpha_k = 0 \\
\Rightarrow \quad & \alpha_{k+1} = 0
\end{aligned}$$

$\square$

3.105 Hauptsatz: Für eine lineare Abbildung $L : V \to V$ sind äquivalent

(i) L ist diagonalisierbar.

(ii) p_L zerfällt über $\mathbb{K}$ in Linearfaktoren (hat also n nicht notwendig verschiedene Nullstellen = Eigenwerte in $\mathbb{K}$):

$$p(\lambda) = (-1)^n(\lambda - \lambda_1)(\lambda - \lambda_2) \cdots (\lambda - \lambda_n),$$

und für jede Nullstelle von P_L gilt geometrische = algebraische Vielfachheit.

Beweis: Sei $n := \dim V$.

(i) ⇒ (ii):

$$\begin{aligned}
\text{(i)} \quad &\Leftrightarrow \quad V \text{ besitzt Basis aus EVen } B = \{b_1, \ldots, b_n\} \\
&\Leftrightarrow \quad M_L^{B,B} = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \\
&\Rightarrow \quad p_L(\lambda) = \det(M_L^{B,B} - \lambda \cdot E_n) = (\lambda_1 - \lambda) \cdots (\lambda_n - \lambda)
\end{aligned}$$

Da zu jedem λ_j ein Eigenvektor b_j aus der Basis gehört, gilt $n_g(\lambda_j) \geq n_a(\lambda_j)$. Mit Satz 3.103 folgt $n_g(\lambda_j) = n_a(\lambda_j)$.

(ii) ⇒ (i):

Seien $\lambda_1, \ldots, \lambda_k$ die paarweise verschiedenen Nullstellen von p_L. Da das Polynom in Linearfaktoren zerfällt, gilt

$$\sum_{j=1}^{k} n_a(\lambda_j) = n.$$

Wegen der Bedingung $n_a(\lambda_j) = n_g(\lambda_j)$ folgt

$$\sum_{j=1}^{k} n_g(\lambda_j) = n.$$

Sind $B_1, \ldots, B_k$ Basen der Eigenräume zu $\lambda_1, \ldots, \lambda_k$, so folgt aus Satz 3.104, dass $B := B_1 \cup \ldots \cup B_k$ linear unabhängig ist. Wegen

$$\sharp B = \sum_{j=1}^{k} n_g(\lambda_j) = n = \dim V$$

ist B dann eine Basis von V.

$\square$

3.14 Sesquilinearformen

3.106 Definition: Sei V ein Vektorraum über $\mathbb{K}$. Eine Abbildung $b : V \times V \to \mathbb{K}$ mit

$$\begin{aligned} b(\lambda_1 \cdot v_1 + \lambda_2 \cdot v_2, w) &= \lambda_1 \, b(v_1, w) + \lambda_2 \, b(v_2, w), \\ b(w, v) &= \overline{b(v, w)} \\ \Big(\text{Dann gilt auch } b(v, \lambda_1 \cdot w_1 + \lambda_2 \cdot w_2) &= \overline{\lambda_1} \, b(v, w_1) + \overline{\lambda_2} \, b(v, w_2) \\ \text{und } b(v, 0) &= 0 \; = \; b(0, w)\Big) \end{aligned}$$

heißt **Sesquilinearform** auf V. Falls $\mathbb{K} = \mathbb{R}$, heißt b auch **Bilinearform**.

3.107 Satz: Sei V Vektorraum mit Skalarprodukt, E ONB von V, b eine Sesquilinearform auf V. Dann existiert eine eindeutig bestimmte selbstadjungierte Matrix $M_b^E \in M_{n,n}(\mathbb{K})$, so dass

$$b(v, w) \;=\; \left\langle M_b^E \cdot v_E, w_E \right\rangle_{\mathbb{K}^n}.$$

Falls $\mathbb{K} = \mathbb{R}$, ist M_b^E symmetrisch.

Beweis: Sei $E = \{e_1, \ldots, e_n\}$ und $M_b^E := (\alpha_{jk})$ mit $\alpha_{jk} := b(e_k, e_j)$.

$$\Rightarrow \;\; \alpha_{kj} = b(e_j, e_k) = \overline{b(e_k, e_j)} = \overline{\alpha_{jk}}, \quad \text{also } \left(M_b^E\right)^* = M_b^E$$

Außerdem:

$$\begin{aligned} b(v, w) &= b\left(\sum_{j=1}^{n} \underbrace{\langle v, e_j \rangle}_{=v_k} \cdot e_j, \sum_{k=1}^{n} \underbrace{\langle w, e_k \rangle}_{=w_k \text{ (k-te Koordinate von w_E)}} \cdot e_k\right) \\ &= \sum_{j=1}^{n} \sum_{k=1}^{n} v_j \overline{w_k} \, b(e_j, e_k) \\ &= \sum_{k=1}^{n} \left(\sum_{j=1}^{n} \alpha_{kj} \, v_j\right) \overline{w_k} \\ &= \left\langle M_b^E \cdot v_E, w_E \right\rangle_{\mathbb{K}^n} \end{aligned}$$

$\square$

3.108 Bemerkung: Durch $b \mapsto M_b^E$ wird eine bijektive Abbildung von der Menge aller Sesquilinearformen auf V auf die Menge aller selbstadjungierten $n \times n$-Matrizen in $\mathbb{K}$ definiert. Dies entspricht dem Vorgehen bei linearen Abbildungen: Durch $L \mapsto M_L^{B',B}$ wird eine bijektive Abbildung von der Menge alle linearen Abbildungen $L : V \to W$ auf die Menge aller $m \times n$-Matrizen definiert.

3.109 Satz: Zusätzlich zum letzten Satz sei E' eine weitere Orthonormalbasis von V und $S := M_{\mathrm{Id}}^{E,E'}$ die Basiswechselmatrix von E' zu E. Dann gilt

$$M_b^{E'} \;=\; S^* \cdot M_b^E \cdot S \;=\; S^{-1} \cdot M_b^E \cdot S\,.$$

Beweis:

$$\begin{aligned} b(v,w) &= \left\langle M_b^E \cdot v_E, w_E \right\rangle \\ &= \left\langle M_b^E \cdot M_{\mathrm{Id}}^{E,E'} \cdot v_{E'}, M_{\mathrm{Id}}^{E,E'} \cdot w_{E'} \right\rangle \\ &= \left\langle S^* \cdot M_b^E \cdot S \cdot v_{E'}, w_{E'} \right\rangle, \end{aligned}$$

und $S^* = S^{-1}$, da S unitär ist. $\square$

3.110 Definition: Sei b eine Sesquilinearform auf V.

1) Die Abbildung

$$q : V \to \mathbb{R} : v \mapsto b(v,v)$$

heißt die zu b gehörige **quadratische Form**.

2) Eine quadratische Form heißt **positiv semidefinit**, falls

$$\forall v \in V : q(v) \;\geq\; 0$$

b heißt **positiv semidefinit**, falls die zugehörige quadratische Form positiv semidefinit ist.
Eine selbstadjungierte $n \times n$-Matrix A heißt **positiv semidefinit**, falls

$$\forall x \in \mathbb{K}^n : \langle A \cdot x, x \rangle_{\mathbb{K}^n} \;\geq 0.$$

3) Eine Sesquilinearform bzw. die zugehörige quadratische Form heißt **positiv definit**, falls

$$\forall v \in V \setminus \{0\} : b(v,v) \;=\; q(v) \;>\; 0.$$

Eine selbstadjungierte $n \times n$-Matrix A heißt **positiv definit**, falls

$$\forall x \in \mathbb{K}^n \setminus \{0\} : \langle A \cdot x, x \rangle_{\mathbb{K}^n} \;> 0.$$

Jede positiv definite Sesquilinearform definiert ein Skalarprodukt.

3.15 Diagonalisierung mit Orthonormalbasen

3.111 Satz: Sei V Vektorraum mit Skalarprodukt. Ist $L : V \to V$ normal (d.h. $L^* \circ L = L \circ L^*$), so gilt:

$$v \text{ ist Eigenvektor von } L \text{ zum Eigenwert } \lambda \;\Rightarrow\; v \text{ ist Eigenvektor von } L^* \text{ zum Eigenwert } \overline{\lambda}$$

Beweis:
$$\begin{aligned}
\left\|\left(L^* - \overline{\lambda} \cdot \mathrm{Id}\right) v\right\|_2^2 &= \left\langle \left(L^* - \overline{\lambda} \cdot \mathrm{Id}\right) v, \left(L^* - \overline{\lambda} \cdot \mathrm{Id}\right) v \right\rangle \\
&= \underbrace{\langle L^*v, L^*v \rangle}_{=\langle Lv, Lv \rangle} - \underbrace{\langle L^*v, \overline{\lambda} \cdot v \rangle}_{=\langle \lambda v, Lv \rangle} - \underbrace{\langle \overline{\lambda} \cdot v, L^*v \rangle}_{=\langle Lv, \lambda \cdot v \rangle} + \underbrace{\langle \overline{\lambda} \cdot v, \overline{\lambda} \cdot v \rangle}_{=\langle \lambda \cdot v, \lambda \cdot v \rangle} \\
&= \langle (L - \lambda \cdot \mathrm{Id})\, v, (L - \lambda \cdot \mathrm{Id})\, v \rangle \\
&= 0
\end{aligned}$$

□

3.112 Charakterisierung normaler Abbildungen: Sei V Vektorraum über $\mathbb{C}$ mit Skalarprodukt. Für eine lineare Abbildung $L : V \to V$ sind äquivalent:

(i) L besitzt eine ONB aus Eigenvektoren.

(ii) $L^* \circ L = L \circ L^*$ (d.h. L ist normal).

3.113 Folgerung: Ist L normal, dann gelten:

1) Eigenvektoren zu verschiedenen Eigenwerten sind orthogonal.

2) L ist diagonalisierbar

Beweis: **(i) ⇒ (ii):** Sei B die ONB aus EVen.

$$\begin{aligned}
\text{(i)} \;\Rightarrow\; M_L^{B,B} &= \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}, \quad M_{L^*}^{B,B} = \begin{pmatrix} \overline{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \overline{\lambda_n} \end{pmatrix} \\
\Rightarrow M_{L \circ L^*}^{B,B} &= M_L^{B,B} \circ M_{L^*}^{B,B} = M_{L^*}^{B,B} \circ M_L^{B,B} \circ = M_{L^* \circ L}^{B,B} \\
\Rightarrow L \circ L^* &= L^* \circ L
\end{aligned}$$

(ii) ⇒ (i): Vollständige Induktion nach $n = \dim V$:

Induktionsanfang $n = 1$: Klar, jede lineare Abbildung besitzt eine ONB $\{b_1\}$ aus EVen.

Induktionsschritt: Sei $\dim V = n + 1$. Da $\mathbb{K} = \mathbb{C}$, besitzt p_L mindestens eine Nullstelle $\lambda_1 \;\Rightarrow\; \lambda_1$ ist EW. Sei v_1 EV zu λ_1.
Betrachte $W := \{v_1\}^\perp = \{v \in V : \langle v, v_1 \rangle = 0\}$. W ist Unterraum von V mit $\dim W = n$.

(α) Es gilt $L(W) \subseteq W$:

$$w \in M \;\Rightarrow\; \langle Lw, v_1\rangle = \langle w, L^* v_1\rangle = \left\langle w, \overline{\lambda_1} v_1\right\rangle = 0 \;\Rightarrow\; Lw \in W.$$

(β) Es gilt $L^*(W) \subseteq W$: Beweis genauso.

Betrachte nun $\widetilde{L} := L\big|_W : W \to W$. Dann gilt $\widetilde{L}^* = L^*\big|_W$.

$$\Rightarrow\; \widetilde{L}^* \circ \widetilde{L} = L^*\big|_W \circ L\big|_W = L\big|_W \circ L^*\big|_W = \widetilde{L} \circ \widetilde{L}^*.$$

Also ist $\widetilde{L}$ normal.

Induktionsvoraussetzung: Es existiert eine ONB $\{e_1, \ldots, e_n\}$ von W aus EVen von $\widetilde{L}$

$$\Rightarrow\; \{e_1, \ldots, e_n, \frac{v_1}{\|v_1\|}\} \text{ ist ONB aus EVen von } L.$$

$\square$

3.114 Charakterisierung unitärer Abbildungen: Sei V Vektorraum über $\mathbb{C}$ und $L : V \to V$ linear. Dann sind äquivalent:

(i) L ist unitär,

(ii) L ist normal und $\sigma(L) \subseteq \{\lambda \in \mathbb{C} : |\lambda| = 1\}$.

Insbesondere ist jede unitäre Abbildung über $\mathbb{C}$ diagonalisierbar, und alle Eigenwerte liegen auf dem komplexen Kreis um 0 mit Radius 1.

Beweis: **(i) $\Rightarrow$ (ii):** $L^* = L^{-1} \;\Rightarrow\; L$ ist normal, also diagonalisierbar mit ONB.

Sei B ONB aus EVen $\Rightarrow\; M_L^{B,B} = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$, $M_{L^*}^{B,B} = \begin{pmatrix} \overline{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \overline{\lambda_n} \end{pmatrix}$.

$$L \circ L^* = \mathrm{Id} \;\Rightarrow\; \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \cdot \begin{pmatrix} \overline{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \overline{\lambda_n} \end{pmatrix} = \mathrm{E}_n.$$

$\Rightarrow\; |\lambda_1|^2 = \ldots = |\lambda_n|^2 = 1.$

(ii) $\Rightarrow$ (i): L ist normal, also diagonalisierbar mit ONB. Sei B ONB aus EVen

$$\Rightarrow\; M_L^{B,B} \cdot M_{L^*}^{B,B} = \begin{pmatrix} |\lambda_1|^2 & & 0 \\ & \ddots & \\ 0 & & |\lambda_n|^2 \end{pmatrix} = \mathrm{E}_n = M_{\mathrm{Id}}^{B,B}.$$

$\square$

3.115 Satz und Definition: **1)** Die unitäre Gruppe

$$U(n) \ := \ \{L : \mathbb{C}^n \to \mathbb{C}^n \mid L \text{ ist unitär}\}$$

ist eine **Untergruppe** der linearen Gruppe $(\mathrm{GL}_n(\mathbb{C}), \circ)$, d.h. dass $U(n) \subseteq \mathrm{GL}_n(\mathbb{C})$ gilt und $(U(n), \circ)$ eine Gruppe ist.

2) Die lineare Gruppe $\mathrm{GL}_n(\mathbb{R})$ enthält folgende Untergruppen:

- Die orthogonale Gruppe $\mathrm{O}(n) \ := \ \{L : \mathbb{R}^n \to \mathbb{R}^n \mid L \text{ ist orthogonal}\}$,
- Die spezielle orthogonale Gruppe $\mathrm{SO}(n) \ := \ \{L \in \mathrm{O}(n) \mid \det(L) = 1\}$.

Beweis: Zur Erinnerung: $(\mathrm{GL}_n(\mathbb{K}), \circ) = \{L : \mathbb{K}^n \to \mathbb{K}^n \mid L \text{ bijektiv}\}$, $(\mathrm{GL}_n(\mathbb{K}), \circ)$ ist eine nicht abelsche Gruppe.

1) Sei $U : \mathbb{C}^n \to \mathbb{C}^n$ unitär.

$$\Rightarrow \ |\det(U)| = \left|\prod_{j=1}^{n} \lambda_j\right| = 1 \neq 0 \ \Rightarrow \ U \text{ ist bijektiv.}$$

Also: $U(n) \subseteq \mathrm{GL}(\mathbb{K})$.

Seien U, V unitär.

$$\Rightarrow \begin{cases} U \circ V \text{unitär:} & (U \circ V)^* = V^* \circ U^* = V^{-1} \circ U^{-1} = (U \circ V)^{-1} \\ U^{-1} \text{ unitär:} & (U^{-1})^* = (U^*)^* = U = U^{-1})^{-1} \end{cases}$$

(Insbesondere $\mathrm{Id} = U \circ U^{-1} \in U(n)$)
Damit ist $U(n)$ Untergruppe von $(\mathrm{GL}_n(\mathbb{K}), \circ)$.

2) Wie in 1) folgt, dass $O(n)$ Untergruppe von $(\mathrm{GL}_n(\mathbb{K}), \circ)$ ist.
Es gilt: $U \in O(n) \ \Rightarrow \ \det U = \pm 1$.

Seien $U, V \in SO(n) \ \Rightarrow \begin{cases} U \circ V \in SO(n) : \det(U \circ V) = (\det U) \cdot \det V = 1 \\ U^{-1} \in SO(n) : \det(U^{-1}) = \dfrac{1}{\det U} = 1 \end{cases}$

$\square$

3.116 Charakterisierung selbstadjungierter Abbildungen: Sei V Vektorraum über $\mathbb{C}$ mit Skalarprodukt und $L : V \to V$ linear. Dann sind äquivalent:

(i) L ist selbstadjungiert,

(ii) L ist normal und $\sigma(L) \subseteq \mathbb{R}$,

(iii) $\forall v \in V : \langle L(v), v\rangle \in \mathbb{R}$.

Insbesondere besitzt L eine ONB aus EVen, ist also diagonalisierbar.

Beweis: **(i) $\Rightarrow$ (ii):** $L^* = L \;\Rightarrow\; L$ ist normal.

Sei B ONB aus EVen. Dann:

$$M_L^{B,B} = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \overset{L=L^*}{=} M_{L^*}^{B,B} = \begin{pmatrix} \overline{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \overline{\lambda_n} \end{pmatrix}$$

$\Rightarrow\; \forall j \in \{1, \ldots, n\} : \lambda_j \in \mathbb{R}.$

(ii) $\Rightarrow$ (iii): Sei B ONB aus EVen.

$$\Rightarrow\; \langle Lv, v\rangle = \left\langle M_L^{B,B} \cdot v_B, v_B \right\rangle = \lambda_1^2|v_1|^2 + \ldots + \lambda_n|v_n|^2 \in \mathbb{R}.$$

(iii) $\Rightarrow$ (i): $\langle Lv, v\rangle = \langle v, Lv\rangle = \langle L^*v, v\rangle$.

$$\Rightarrow\; \forall v \in V : \langle (L - L^*)v, v\rangle = 0.$$

Nun gilt: $L - L^*$ ist normal (Nachrechnen!)

Sei B ONB aus EVen von $L - L^*$, $B = \{b_1, \ldots, b_n\}$.

$$\Rightarrow\; 0 = \langle (L - L^*)b_j, b_j\rangle = \lambda_j \|b_j\|^2 \;\Rightarrow\; \lambda_j = 0$$

$\Rightarrow\; L = L^*$, L ist selbstadjungiert. $\square$

3.117 Satz: Es sei M eine selbstadjungierte $n \times n$-Matrix in $\mathbb{K}$ und $L_M : x \mapsto M \cdot x$.

1) Äquivalent sind: **(i)** M ist positiv semidefinit,
(ii) Für alle Eigenwerte λ von L_M gilt $\lambda \geq 0$.

2) Äquivalent sind: **(i)** M ist positiv definit,
(ii) Für alle Eigenwerte λ von L_M gilt $\lambda > 0$.

Beweis: Sei E die kanonische Basis, B eine ONB aus EVen (existiert nach 3.116) und $S := M_{\mathrm{Id}}^{E,E'}$.

$$\begin{aligned} \Rightarrow \underbrace{S^*}_{=S^{-1}} \cdot M \cdot S &= M_{\mathrm{Id}}^{E',E} \cdot M_{L_M}^{E,E} \cdot M_{\mathrm{Id}}^{E,E'} = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \\ \Rightarrow q(x) &= \langle Mx, x\rangle \\ &= \langle S^* \cdot M \cdot S(x)_{E'}, (x)_{E'}\rangle \\ &= \lambda_1(x_1')^2 + \ldots \lambda_n(x_n')^2. \end{aligned}$$

$\square$

3.118 Charakterisierung symmetrischer Abbildungen: Sei $L : \mathbb{R}^n \to \mathbb{R}^n : x \mapsto A \cdot x$. Dann sind äquivalent:

(i) A bzw. L ist symmetrisch,

(ii) L besitzt eine Orthonormalbasis aus Eigenvektoren.

Insbesondere zerfällt das charakteristische Polynom p_L in reelle Linearfaktoren, und für jeden Eigenwert λ von L gilt: $n_a(\lambda) = n_g(\lambda)$.

Beweisskizze:

(ii) $\Rightarrow$ (i): Sei B ONB aus EVen

$$\begin{aligned} \Rightarrow \qquad & S = M_{\mathrm{Id}}^{B,E} \text{ ist orthogonal,} \quad S \cdot A \cdot S^{-1} = D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix} \\ \Rightarrow \qquad & A^T = (S^{-1} \cdot D \cdot S)^T = (S^T \cdot D \cdot S)^T = S^T \cdot D \cdot (S^T)^T = S^T \cdot D \cdot S = A. \end{aligned}$$

(i) $\Rightarrow$ (ii): Idee: Betrachte $\widetilde{L} : \mathbb{C}^n \to \mathbb{C}^n : x \mapsto Ax$.

$\widetilde{L}$ ist selbstadjungiert $\Rightarrow \begin{cases} \widetilde{B} \text{ ONB aus EVen existiert} \\ \text{Alle Eigenwerte sind reell} \end{cases}$

Ist nun $Ab_j = \lambda_j b_j$, so folgt für $c_j := \operatorname{Re} b_j = \frac{1}{2}(b_j + \overline{b_j}) \in \mathbb{R}^n$ und $\widetilde{c_j} := \operatorname{Im} b_j = \frac{1}{2}(b_j - \overline{b_j}) \in \mathbb{R}^n$:

$$Ac_j = \lambda_j \cdot c_j \ \wedge \ A\widetilde{c_j} = \lambda_j \cdot \widetilde{c_j}.$$

3.119 Anwendung auf quadratische Formen: Auf $\mathbb{R}^n$ sei die quadratische Form

$$q(x) \ := \ \langle A \cdot x, x \rangle$$

mit beliebiger quadratischer Matrix $A \in M_{n,n}(\mathbb{R})$ gegeben. Dann ist $(A + A^T)$ symmetrisch, und q ist die zur Bilinearform

$$b(x, y) \ = \ \left\langle \frac{1}{2}(A + A^T)x, y \right\rangle$$

gehörende quadratische Form. Sei nun $B = \{b_1, \ldots, b_n\}$ eine Orthonormalbasis aus Eigenvektoren: $\frac{1}{2}(A + A^T) \cdot b_j = \lambda_j \cdot b_j$. Dann gilt

$$q(x) \ = \ \lambda_1 \widetilde{x}_1^2 + \ldots + \lambda_n \widetilde{x}_n^2 \qquad \text{mit } \widetilde{x} = x_B.$$

3.16 Die Jordansche Normalform

3.120 Definition: Für $\lambda_0 \in \mathbb{K}$ und $k \in \mathbb{N}$ heißt die $k \times k$-Matrix

$$J_{\lambda_0}^{(k)} := \begin{pmatrix} \lambda_0 & 1 & 0 & \dots & 0 \\ 0 & \lambda_0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ & & & \ddots & 1 \\ 0 & \dots & & 0 & \lambda_0 \end{pmatrix} \in M_{k,k}(\mathbb{K})$$

Jordan-Matrix oder **Jordan-Block**.

3.121 Satz: Für eine Jordan-Matrix $J_{\lambda_0}^{(k)}$ gelten:

1) $\lambda = \lambda_0$ ist einziger Eigenwert von $J_{\lambda_0}^{(k)}$,

2) $n_a(\lambda_0) = k$ und $n_g(\lambda_0) = 1$. Insbesondere ist $J_{\lambda_0}^{(k)}$ nicht diagonalisierbar für $k \geq 2$.

3) $\big(J_{\lambda_0}^{(k)} - \lambda_0 \cdot E\big)^k = 0$, d.h. die Matrix $J_{\lambda_0}^{(k)} - \lambda_0 \cdot E$ bzw. die zugehörige lineare Abbildung ist **nilpotent**.

Beweis: **1)** Es gilt $\det(J_{\lambda_0}^{(k)} - \lambda \cdot E_k) = (\lambda_0 - \lambda)^k$.

2) Folgt aus

$$\big(J_{\lambda_0}^{(k)} - \lambda_0 \cdot \mathrm{Id}\big)v = 0 \Leftrightarrow \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ & & & \ddots & 1 \\ 0 & \dots & & 0 & 0 \end{pmatrix} v = 0 \Leftrightarrow v = t \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

3) Die Anwendung von $\big(J_{\lambda_0}^{(k)} - \lambda_0 \cdot \mathrm{Id}\big)$ auf einen Vektor v verschiebt alle Koordinateneinträge um eine Stelle nach oben und setzt die letzte Koordinate auf 0. Da der Vektor genau k Koordinaten hat, sind nach k Anwendungen alle Koordinaten 0, also $\big(J_{\lambda_0}^{(k)} - \lambda_0 \cdot \mathrm{Id}\big)^k \cdot v = 0$. □

3.122 Satz: Es sei $k \geq 2$ und $L : \mathbb{K}^k \to \mathbb{K}^k$ eine lineare Abbildung, die nur einen Eigenwert λ_0 besitzt, und für die gilt: $n_a(\lambda_0) = k$, $n_g(\lambda_0) = 1$. Für eine Basis B von $\mathbb{K}^k$ sind äquivalent:

(i) $M_L^{B,B} = J_{\lambda_0}^{(k)}$,

(ii) Die Basis $B = \{b_1, \ldots, b_k\}$ besteht aus einer **Vektorkette** zum Eigenwert λ_0 von L, d.h. es gilt

$$\begin{aligned}(L - \lambda_0 \cdot \mathrm{Id}) \cdot b_1 &= 0 \quad (b_1 \text{ ist Eigenvektor}),\\ (L - \lambda_0 \cdot \mathrm{Id}) \cdot b_{j+1} &= b_j \quad \text{für } j = 1, 2, \ldots, k-1.\end{aligned}$$

Das bedeutet, man kommt von einem Element der Vektorkette zum vorherigen durch Anwendung der Abbildung $L - \lambda_0 \cdot \mathrm{Id}$.

Beweis: **(i) $\Rightarrow$ (ii):** Durch Nachrechnen, z.B.

$$L\, b_2 = J_{\lambda_0}^{(k)} \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}_B = \begin{pmatrix} 1 \\ \lambda_0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}_B = b_1 + \lambda_0 \cdot b_2$$

(ii) $\Rightarrow$ (i):

$$L\, b_1 = \lambda_0 \cdot b_1 \quad \Rightarrow \quad M_L^{B,B} = \begin{pmatrix} \lambda_0 & * & \ldots & * \\ 0 & * & \ldots & * \\ \vdots & \vdots & & \vdots \\ 0 & * & \ldots & * \end{pmatrix}$$

$$L\, b_2 = \lambda_0 \cdot b_2 + b_1 \quad \Rightarrow \quad M_L^{B,B} = \begin{pmatrix} \lambda_0 & 1 & * & \ldots & * \\ 0 & \lambda_0 & * & \ldots & * \\ 0 & 0 & * & \ldots & * \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & * & \ldots & * \end{pmatrix}$$

usw. $\square$

3.123 Jordansche Normalform: Sei $L : V \to V$ linear. Falls p_L über $\mathbb{K}$ in Linearfaktoren zerfällt:

$$p_L(\lambda) = (\lambda_1 - \lambda)^{n_1} \cdots (\lambda_k - \lambda)^{n_k},$$

so gibt es eine Basis B von V, so dass

$$M_L^{B,B} = \begin{pmatrix} \boxed{J_1} & & & 0 \\ & \boxed{J_2} & & \\ & & \ddots & \\ 0 & & & \boxed{J_r} \end{pmatrix},$$

wobei $J_i = J_{\lambda_i}^{(k_i)}$ Jordan-Blöcke sind. L ist genau dann diagonalisierbar, wenn alle Jordan-Blöcke 1-dimensional sind.

Vorsicht: Es kann mehrere Jordan-Blöcke zu einem Eigenwert geben. Sind z.B. $J_1, \ldots, J_l$ alle Jordan-Blöcke zum Eigenwert λ_1, so gilt $k_1 + \ldots + k_l = n_a(\lambda_1)$ und $n_g(\lambda_1) = l$.

Ergänzung: Beweis des Satzes 3.123

3.124 Grundvoraussetzung: Für den Beweis von Satz 3.123 sei im folgenden vorausgesetzt:

- $L : V \to V$ ist linear, und es gilt $\dim(V) = n$.
- Für das charakteristische Polynom von L gilt

$$p_L(\lambda) \;=\; (\lambda_1 - \lambda)^{n_1} \cdots (\lambda_k - \lambda)^{n_k},$$

  wobei die Eigenwerte $\lambda_1, \dots, \lambda_k$ paarweise verschieden sind. D.h. die algebraische Vielfachheit von λ_j ist $n_a(\lambda_j) = n_j$. Insbesondere gilt $n_1 + \ldots + n_k = n$.
- Wir betrachten die Untervektorräume

$$V_j := \text{Kern}\left((L - \lambda_j \cdot \text{Id})^{n_j}\right) \quad \text{mit Basis} \quad B_j \qquad (j = 1, \dots, k).$$

  Zunächst sei B_j ganz allgemein eine Basis von V_j, später werden diese Basen speziell gewählt.

3.125 Satz: **1)** Für jedes $j \in \{1, \dots, k\}$ ist V_j **invariant** unter L, d.h. es gilt

$$L(V_j) \;\subseteq V_j.$$

2) Für $i \neq j$ ist die Abbildung $(L - \lambda_i \cdot \text{Id})\big|_{V_j} : V_j \to V_j$ bijektiv.

Beweis: **1)** Sei j fest. Für $v \in V_j$ gilt $(L - \lambda_j \cdot \text{Id})^{n_j} v = 0$. Daraus folgt

$$(L - \lambda_j \cdot \text{Id})^{n_j}(Lv) \;=\; L(L - \lambda_j \cdot \text{Id})^{n_j} v \;=\; 0,$$

also $Lv \in V_j$.

2) Zunächst erhalten wir aus 1), dass für $v \in V_j$ auch $(L - \lambda_i \cdot \text{Id})v = L\,v - \lambda_i \cdot v \in V_j$.
Sei $v \in \text{Kern}\left((L - \lambda_i \cdot \text{Id})\big|_{V_j}\right)$. Das bedeutet einerseits $v \in V_j$, also $(L - \lambda_j \cdot \text{Id})^{n_j} v = 0$, und andererseits $L\,v = \lambda_i \cdot v$, also $(L - \lambda_j \cdot \text{Id})^{n_j} v = (\lambda_i - \lambda_j)^{n_j} \cdot v$. Wegen $\lambda_i - \lambda_j \neq 0$ folgt $v = 0$. Also gilt $\text{Kern}\left((L - \lambda_i \cdot \text{Id})\big|_{V_j}\right) = \{0\}$, und daraus folgt die Bijektivität. $\square$

3.126 Satz: $B := \bigcup_{j=1}^{k} B_j$ ist linear unabhängig (B_j = Basis von V_j, vgl. 3.124). Insbesondere gilt

$$\sum_{j=1}^{k} \dim(V_j) \;\leq\; \dim(V) \;=\; n.$$

Beweis: Wir bezeichnen die Elemente von B_j mit $b_1^{(j)}, \ldots, b_{d_j}^{(j)}$, wobei $d_j := \dim(V_j)$. Wir zeigen, wie aus

$$0 = \sum_{j=1}^{k} \sum_{l=1}^{d_j} c_{jl} \cdot b_l^{(j)}$$

$c_{1\,1} = c_{1\,2} = \ldots = c_{1\,d_1} = 0$ folgt. Genauso folgt dann $c_{j\,l} = 0$ für alle j, l, also die lineare Unabhängigkeit von B.

Wegen $(L - \lambda_j \cdot \mathrm{Id})^{n_j} \cdot b_l^{(j)} = 0$ gilt

$$0 = \left(\prod_{j=2}^{k} (L - \lambda_j \cdot \mathrm{Id})^{n_j} \right) \left(\sum_{j=1}^{k} \sum_{l=1}^{n_j} c_{jl} \cdot b_l^{(j)} \right) = \sum_{l=1}^{n_1} c_{1l} \cdot \left(\prod_{j=2}^{k} (L - \lambda_j \cdot \mathrm{Id})^{n_j} \right) b_l^{(1)}$$

Für $j \neq 1$ ist die Abbildung $(L - \lambda_j \cdot \mathrm{Id})\Big|_{V_1} : V_1 \to V_1$ bijektiv. Das bedeutet, dass linear unabhängige Mengen auf linear unabhängige Mengen abgebildet werden. Somit ist die Menge

$$\left\{ \prod_{j=2}^{k} (L - \lambda_j \cdot \mathrm{Id})^{n_j} \cdot b_l^{(1)} : l = 1, \ldots, d_1 \right\}$$

linear unabhängig, und es folgt $c_{1\,1} = c_{1\,2} = \ldots = c_{1\,d_1} = 0$. □

3.127 Satz: Es gilt $\dim(V_j) = n_j$ für $j = 1, \ldots, k$.

Beweis: Es genügt, $\dim(V_j) \geq n_j$ zu beweisen. Mit $\dim(V_1) + \ldots + \dim(V_k) \leq n_1 + \ldots n_k$ (letzter Satz) folgt dann die Behauptung.

Wir führen den Beweis für $j = 1$. Dazu sei $\tilde{B}$ eine Basis von V, die durch Ergänzung von B_1 entstanden ist. Da V_1 invariant unter L ist, folgt

$$M_L^{\tilde{B},\tilde{B}} = \begin{pmatrix} A_1 & * \\ 0 & A \end{pmatrix},$$

mit einer $d_1 \times d_1$-Matrix A_1 und einer $(n - d_1) \times (n - d_1)$-Matrix A ($d_1 = \dim(V_1)$).

Aus dem Entwicklungssatz folgt

$$p_L(\lambda) = \det(L - \lambda \cdot \mathrm{Id}) = \big(\det(A_1 - \lambda \cdot E_{d_1})\big) \cdot \det(A - \lambda \cdot E_{n-d_1})$$

($E_l =$ $l \times l$-Einheitsmatrix). Falls λ_1 keine Nullstelle von $\det(A - \lambda \cdot E_{n-d_1})$ ist, muss $(\lambda_1 - \lambda)^{n_1}$ Teiler von $\det(A_1 - \lambda \cdot E_{d_1})$ sein und somit $d_1 \geq n_1$ gelten.

Annahme: λ_1 ist Nullstelle von $\det(A - \lambda \cdot E_{n-d_1})$. Dann besitzt A einen Eigenvektor zum Eigenwert λ_1. Hieraus folgt, dass L einen Eigenvektor zum Eigenwert λ_1 besitzt, der nicht in V_1 (von B_1 aufgespannt) liegt. Dies ist ein Widerspruch zu $V_1 = \mathrm{Kern}\big((L - \lambda_1 \cdot \mathrm{Id})^{n_1}\big)$. □

3.128 Satz: $B := \bigcup_{j=1}^{k} B_j$ ist Basis von V, und es gilt

$$M_L^{B,B} = \begin{pmatrix} A_1 & 0 & \dots & 0 \\ 0 & A_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & A_k \end{pmatrix},$$

wobei für $j = 1, \dots, k$ die Matrix A_j eine $n_j \times n_j$-Matrix ist.

Beweis: Nach dem letzten Satz spannt B einen Raum der Dimension $n_1 + \dots + n_k = n = \dim(V)$ auf, ist also Basis von V.

Die Gestalt der Matrix $M_L^{B,B}$ folgt aus der Tatsache, dass die von den Basen B_j aufgespannten Räume V_j unter L invariant sind. □

Mit dem nächsten Satz bzw. der Verallgemeinerung für B_j folgt dann die Behauptung von Satz 3.123.

3.129 Satz: Die Basis B_1 von V_1 kann so gewählt werden, dass sie als Vereinigung von Vektorketten dargestellt werden kann, d.h. es gibt Vektorketten $K_1, \dots, K_m$, so dass $B_1 = K_1 \cup K_2 \cup \dots \cup K_m$. Für diese Wahl von B_1 hat die Matrix A_1 im letzten Satz die Form

$$A_1 = \begin{pmatrix} \boxed{J_1} & & & 0 \\ & \boxed{J_2} & & \\ & & \ddots & \\ 0 & & & \boxed{J_m} \end{pmatrix},$$

wobei $J_i = J_{\lambda_1}^{(k_i)}$ Jordan-Blöcke zum Eigenwert λ_1 mit $k_i = \sharp K_i$ sind $(i = 1, \dots, m)$.

Zum Beweis reicht es, eine Basis von V_1 zu konstruieren, die nur aus Vektorketten zum Eigenwert λ_1 von L besteht. Zu jeder Vektorkette (die auch nur aus einem Vektor bestehen kann, das ist dann ein Eigenvektor) gehört dann in A ein Jordan-Block (vgl. 3.122). Wir untersuchen zunächst die lineare Unabhängigkeit von Vektorketten.

3.130 Satz: E seien l Vektorketten $K_1 = \{v_1^{(1)}, \dots\}, \; \dots, \; K_l = \{v_1^{(l)}, \dots\}$ zum Eigenwert λ_1 von L gegeben. Dann sind äquivalent:

(i) $K_1 \cup \dots \cup K_l$ ist linear unabhängig.

(ii) $\{v_1^{(1)}, \ldots, v_1^{(l)}\}$ ist linear unabhängig (d.h. die Menge aus allen ersten Vektoren der Vektorketten).

Beweis: (i) $\Rightarrow$ (ii): Folgt direkt aus $\{v_1^{(1)}, \ldots, v_1^{(l)}\} \subseteq K_1 \cup \ldots \cup K_l$.

(ii) $\Rightarrow$ (i): Wir beweisen folgende einfachere Fassung: Es seien $K_1 = \{v_1, \ldots, v_i\}$ und $K_2 = \{w_1, \ldots, w_j\}$ zwei Vektorketten zum Eigenwert λ_1. Ist $\{v_1, w_1\}$ linear unabhängig, dann auch $K_1 \cup K_2$. Der allgemeine Beweis geht genauso.

Sei
$$0 \;=\; \sum_{l=1}^{i} c_l \cdot v_l + \sum_{l=1}^{j} d_l \cdot w_l.$$
Annahme: Mindestens einer der Koeffizienten c_l, d_l ist von Null verschieden. Dann ist
$$M \;:=\; \max\{l : c_l \neq 0 \;\vee\; d_l \neq 0\} \;\geq\; 1.$$
Da alle c_l, d_l mit $l \geq 0$ verschwinden, folgt aus der Definition von Vektorketten
$$\begin{aligned} 0 \;&=\; (L - \lambda_1 \cdot \mathrm{Id})^{M-1} \left(\sum_{l=1}^{M} c_l \cdot v_l + \sum_{l=1}^{M} d_l \cdot w_l \right) \\ &=\; c_M \cdot v_1 + d_M \cdot w_1 \end{aligned}$$
und aus der linearen Unabhängigkeit von $\{v_1, w_1\}$ schließlich $c_M = d_M = 0$ im Widerspruch zur Definition von M. Also müssen alle Koeffizienten c_l, d_l verschwinden. $\square$

Beweis: (von Satz 3.129, konstruktiv, kann so programmiert werden)

Wir ändern die Basis B_1 von V_1 sukzessive ab, bis sie nur noch aus Vektorketten besteht. Dazu betrachten wir für jedes Element $v \in B_1$ die zugehörige Vektorkette
$$\begin{aligned} K(v) \;&:=\; \left\{(L - \lambda_1 \cdot \mathrm{Id})^l \cdot v, (L - \lambda_1 \cdot \mathrm{Id})^{l-1} \cdot v, \ldots, v\right\} \\ &\qquad \text{mit } (L - \lambda_1 \cdot \mathrm{Id})^l \cdot v \neq 0 \;\wedge\; (L - \lambda_1 \cdot \mathrm{Id})^{l+1} \cdot v = 0. \end{aligned}$$
Aus der Definition von V_1 folgt $l \leq n_1 - 1$.

Schritt 1: Wir setzen
$$L(B_1) \;:=\; \max\{\sharp K(v) : v \in B_1\} \qquad \text{(maximale Kettenlänge)},$$
Wir stellen fest, wie viele der Vektorketten maximaler Länge benötigt werden:
$$d(B_1) \;:=\; \dim \mathrm{LH}\left\{(L - \lambda_1 \cdot \mathrm{Id})^{L(B_1)-1} \cdot v : v \in B_1\right\}.$$

Nun seien $v_1, \ldots, v_{d(B_1)} \in B_1$ so ausgewählt, dass gilt:

$$\big\{(L - \lambda_1 \cdot \mathrm{Id})^{L(B_1)-1} \cdot v_l : l = 1, \ldots, d(B_1)\big\}$$
$$\text{ist eine Basis von } \mathrm{LH}\big\{(L - \lambda_1 \cdot \mathrm{Id})^{L(B_1)-1} \cdot v : v \in B_1\big\}.$$

Nun reduzieren wir die Länge der anderen Ketten maximaler Länge. Dazu wird jedes $v \in B_1 \setminus \{v_1, \ldots, v_{d(B_1)}\}$, das eine Kette $K(v)$ mit maximaler Länge erzeugt ($\sharp K(v) = L(B_1)$), ersetzt durch

$$\tilde{v} \;:=\; v - \sum_{j=1}^{d_1} c_j \cdot v_j, \quad \text{so dass } (L - \lambda_1 \cdot \mathrm{Id})^{L(B_1)-1} \cdot \tilde{v} = 0.$$

Dadurch entsteht eine Basis $\tilde{B}_1$, in der nur die Ketten $K(v_1), \ldots, K(v_{d(B_1)})$ die Länge $\sharp K(v) = L(B_1)$ haben. Für alle anderen Ketten gilt $\sharp K(v) < L(B_1)$.

Nun verwenden wir den letzten Satz. Aus der Definition der Vektoren $v_1, \ldots, v_{d(B_1)}$ folgt, dass die Menge $K(v_1) \cup \ldots \cup K(v_{d(B_1)})$ linear unabhängig ist. Nun tauschen wir diese Vektoren mit ebenso vielen passend gewählten Vektoren aus der Basis $\tilde{B}_1$ aus (Basisaustauschsatz). Dann haben wir eine Basis $B_1^{(1)}$ von V_1 mit

$$K(v_1) \cup \ldots \cup K(v_{d(B_1)}) \;\subseteq\; B_1^{(1)}.$$

Schritt 2: Falls $C_1 := B_1^{(1)} \setminus \big(K(v_1) \cup \ldots \cup K(v_{d(B_1)})\big) \neq 0$, führen wir dieselbe Prozedur für diese Menge durch: Wir bestimmen $L(C_1)$ und wählen $d(C_1)$ Vektoren aus C_1 aus, so dass gilt:

$$\big\{(L - \lambda_1 \cdot \mathrm{Id})^{L(C_1)-1} \cdot v_l : l = d(B_1) + 1, \ldots, d(B_1) + d(C_1)\big\}$$
$$\text{ist eine Basis von } \mathrm{LH}\big\{(L - \lambda_1 \cdot \mathrm{Id})^{L(C_1)-1} \cdot v : v \in C_1\big\}.$$

Entsprechend der obigen Vorgehensweise entsteht eine Basis $B_1^{(2)}$, so dass

$$K(v_1) \cup \ldots \cup K(v_{d(B_1)+d(B_2)}) \;\subseteq\; B_1^{(2)}.$$

Dieses Verfahren kann so lange fortgesetzt werden, bis die entstandene Basis $B_1^{(l)}$ nur noch aus Vektorketten (eventuell der Länge 1, dann enthält die Vektorkette nur einen Eigenvektor und keine anderen Vektoren) besteht.

$\square$

3.131 Bemerkungen: **1)** Aus dem letzten Beweis folgt, dass $V_1 = \mathrm{Kern}\,\big((L - \lambda_1 \cdot \mathrm{Id})^{L(B_1)}\big)$ gilt, wobei $L(B_1) \leq n_1$.

2) Sei p_L das charakteristische Polynom von L. Man kann nun in die Variable auch lineare Abbildungen (oder Matrizen) einsetzen, da Potenzen und Linearkombinationen von linearen Abbildungen definiert sind. Dann ist z.B. $p_L(L)$ wieder eine lineare Abbildung. Aus der Jordanschen Normalform von L folgt, dass

$$p_L(L) \;:=\; (L - \lambda_1 \cdot \mathrm{Id})^{n_1} \cdots (L - \lambda_k \cdot \mathrm{Id})^{n_k} \;=\; \mathbb{0}.$$

Diese Gleichung heißt **Gleichung von Cayley-Hamilton**. Mit der vorigen Bemerkung folgt sogar

$$(L - \lambda_1 \cdot \mathrm{Id})^{L(B_1)} \cdots (L - \lambda_k \cdot \mathrm{Id})^{L(B_k)} \;=\; \mathbb{0}.$$

4 Konvergenz

4.1 Abstände

4.1 Definition: Sei M nichtleere Menge. Eine **Metrik** oder ein **Abstand** auf M ist eine Abbildung $d : M \times M \to \mathbb{R}$ mit folgenden Eigenschaften:

(M1) $\forall x, y \in M : d(x,y) \geq 0 \wedge \forall x, y \in M : \big(d(x,y) = 0 \Leftrightarrow x = y\big)$ (Positivität, Definitheit).

(M2) $\forall x, y \in M : d(x,y) = d(y,x)$ (Symmetrie).

(M3) $\forall x, y, z \in M : d(x,y) \leq d(x,z) + d(z,y)$ (Dreiecks-Ungleichung, Δ-Ungleichung).

(M, d) heißt **metrischer Raum**.

4.2 Satz: Ist $(V, \|.\|)$ ein normierter Vektorraum, dann ist

$$d(x,y) \ := \ \|x - y\|$$

eine Metrik auf V.

4.3 Beispiele: **1)** Ist $(\mathbb{K}, |.|)$ ein bewerteter Körper (z.B. $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$), dann ist

$$d(x,y) \ := \ |x - y|$$

eine Metrik auf $\mathbb{K}$.

2) Sei $V = \mathbb{R}^n$ oder $V = \mathbb{C}^n$. Aus Satz 4.2 und Beispielen zur Definition von Norm in 3.59

$$\begin{aligned} d_2(x,y) \ &:= \ \sqrt{\sum_{j=1}^{n} |x_j - y_j|^2}, \\ d_1(x,y) \ &:= \ \sum_{j=1}^{n} |x_j - y_j|, \\ d_\infty(x,y) \ &:= \ \max_{1 \leq j \leq n} |x_j - y_j|, \\ d_p(x,y) \ &:= \ \left(\sum_{j=1}^{n} |x_j - y_j|^p\right)^{1/p} \qquad \text{mit } 1 \leq p < \infty \end{aligned}$$

sind Metriken auf V. Je nach Anwendung kann eine andere Metrik passend sein.

3) Der Raum der $n \times n$-Matrizen $V = \mathcal{M}_{n,n} = \left\{ \big(a_{ij}\big)_{\substack{1 \le i \le n \\ 1 \le j \le n}} : a_{ij} \in \mathbb{C} \right\}$

$$\begin{aligned}
A + B &= \big(a_{ij}\big) + \big(b_{ij}\big) := \big(a_{ij} + b_{ij}\big) \\
\lambda \cdot A &= \lambda \cdot \big(a_{ij}\big) := \big(\lambda a_{ij}\big) \\
A \cdot B &= \big(a_{ij}\big) \cdot \big(b_{ij}\big) := \left(\sum_{l=1}^{n} a_{il}\, b_{lj} \right)_{i,j} \qquad \text{(Matrizenmultiplikation)} \\
\|A\| &= \big\| \big(a_{ij}\big) \big\| := n \max |a_{ij}| \quad \text{ist eine Norm} \\
\Rightarrow\ d(A, B) &= \big\| \big(a_{ij}\big) - \big(b_{ij}\big) \big\| \quad \text{ist eine Metrik auf } \mathcal{M}_{n,n}
\end{aligned}$$

Spezialität: $$\begin{aligned}
\|A \cdot B\| &\le n \max \sum_{l=1}^{n} |a_{il}\, b_{lj}| \\
&\le n^2 \max |a_{ij}| \max |b_{ij}| \\
&= \|A\| \, \|B\|
\end{aligned}$$

Insbesondere gilt: $\big\|A^k\big\| \le \|A\|^k$ für $k \in \mathbb{N}$.

4) $V = C([a, b] \to \mathbb{R})$: Raum der auf $[a, b]$ stetigen reellwertigen Funktionen.

$$\begin{aligned}
\|f\|_\infty &:= \max_{a \le x \le b} |f(x)| \quad \text{ist Norm} \\
d(f, g) &:= \max_{a \le x \le b} |f(x) - g(x)| \quad \text{ist Metrik auf } C([a, b] \to \mathbb{R})
\end{aligned}$$

5) Metrik auf einer beliebigen Menge: Sei $M \neq \emptyset$.

$$d(x, y) := \begin{cases} 0 & \text{falls } x = y \\ 1 & \text{falls } x \neq y \end{cases}$$

ist eine Metrik auf M.

4.4 Dreiecks-Ungleichung nach unten: Sei (M, d) ein metrischer Raum. Für $x, y, z \in M$ gilt

$$d(x, y) \ge \big| d(x, z) - d(z, y) \big|.$$

4.5 Definition: Sei (M, d) ein metrischer Raum, $x_0 \in M$ und $R \in \,]0, \infty[$. Dann heißt

$$B_R(x_0) := \big\{ x \in M : d(x, x_0) < R \big\}$$

offene Kugel um x_0 mit Radius R.

4.2 Folgen

4.6 Konvergenz: Sei (M, d) ein metrischer Raum, (x_n) eine Folge in M und $x \in M$.

1) (x_n) heißt **konvergent gegen $\boldsymbol{x}$**, falls

$$\forall \varepsilon > 0 \;\; \exists N_\varepsilon \in \mathbb{N} \;\; \forall n > N_\varepsilon : d(x_n, x) < \varepsilon.$$

In diesem Fall schreiben wir

$$x \;=\; \lim_{n\to\infty} x_n, \qquad x_n \overset{n\to\infty}{\longrightarrow} x, \qquad x_n \to x \text{ für } n \to \infty, \qquad x_n \to x;$$

x heißt **Grenzwert** oder **Limes** der Folge (x_n).

2) (x_n) heißt **konvergent**, falls

$$\exists x \in M : x = \lim_{n\to\infty} x_n.$$

Ist eine Folge nicht konvergent, so heißt sie **divergent**.

3) (x_n) heißt **Cauchy-Folge**, falls

$$\forall \varepsilon > 0 \;\; \exists N_\varepsilon \in \mathbb{N} \;\; \forall n, m > N_\varepsilon : d(x_n, x_m) < \varepsilon.$$

4.7 Satz und Definition: Sei (M, d) ein metrischer Raum.

1) Jede konvergente Folge ist eine Cauchy-Folge.

2) Gilt umgekehrt: Jede Cauchy-Folge ist in M konvergent, dann heißt (M, d) **vollständig**.

4.8 Eindeutigkeit des Grenzwertes: Sei (M, d) ein metrischer Raum. Ist $\lim\limits_{n\to\infty} x_n = x$, so besitzt (x_n) keinen weiteren Grenzwert.

4.9 Beispiele: **1)** $M = \mathbb{Q}$, $d(x, y) = |x - y|$:

$$x_n \to x \;\Leftrightarrow\; \forall \varepsilon > 0 \;\; \exists N_\varepsilon \in \mathbb{N} \;\; \forall n > N_\varepsilon : |x_n - x| < \varepsilon.$$

Dies ist der Konvergenzbegriff, den wir bereits kennen.
$(\mathbb{Q}, d)$ ist **nicht vollständig**.

2) $M = \mathbb{R}$ mit $d(x, y) = |x - y|$ ist vollständig (vgl. 2.30).

3) $M = \mathbb{C}$, $d(z, u) = |z - u|$. Es gelten (vgl. 2.54)

$$z_n \to z \;\;\Leftrightarrow\;\; \operatorname{Re} z_n \to \operatorname{Re} z \;\wedge\; \operatorname{Im} z_n \to \operatorname{Im} z$$

$$(z_n) \text{ ist Cauchy-Folge} \;\;\Leftrightarrow\;\; (\operatorname{Re} z_n), (\operatorname{Im} z_n) \text{ sind Cauchy-Folgen.}$$

Da $\mathbb{R}$ vollständig ist, ist auch $\mathbb{C}$ vollständig.

4) $M = \mathbb{C}^n$, $d(z,u) := d_\infty(z,u) = \max\limits_{1\le j\le n} |z_j - u_j|$ ist vollständig:

Sei (z_m) Cauchy-Folge, $z_m = \begin{pmatrix} (z_m)_1 \\ \vdots \\ (z_m)_n \end{pmatrix}$.

$$\begin{aligned}
&\Rightarrow\ \ \forall m,k > N_\varepsilon \ \ \forall j = 1,\dots,n : |(z_m)_j - (z_k)_j| \le d_\infty(z_m, z_k) < \varepsilon \\
&\Rightarrow\ \ \forall j = 1,\dots,n : ((z_m)_j)_{m\in\mathbb{N}} \text{ ist Cauchy-Folge in } \mathbb{C} \\
&\Rightarrow\ \ \forall j = 1,\dots,n : (z_m)_j \to u_j \in \mathbb{C}
\end{aligned}$$

Setze $u := \begin{pmatrix} u_1 \\ \vdots \\ u_n \end{pmatrix} \ \Rightarrow\ d(z_m,u) = \max\limits_{1\le j\le n} \left|(z_m)_j - u_j\right| \to 0 \ \Rightarrow\ z_m \to u$ in $(\mathbb{C}^n, d_\infty)$.

Entsprechend kann bewiesen werden: $(\mathbb{C}^n, d_p)$ ist vollständig (vgl. 4.3)

5) $V = \mathcal{M}_{n,n},\ d(A,B) = n \max |a_{ij} - b_{ij}|$:
Genauso wie 4): (V,d) ist vollständig.
Sei $A \in \mathcal{M}_{n,n}$ mit $\|A\| < 1$. Dann gilt $A^n \to 0$:

$$d(A^n, 0) = \|A^n - 0\| = \|A^n\| \le \|A\|^n \to 0.$$

6) $V = C([a,b] \to \mathbb{R}),\ d_\infty(f,g) = \max\limits_{a\le x\le b} |f(x) - g(x)|$: Später: (V, d_∞) ist vollständig.

4.10 Definition: Gilt für die **relle** Folge (a_n):

$$\forall M > 0 \ \ \exists N \in \mathbb{N} \ \ \forall n > N : a_n > M$$

oder

$$\forall M > 0 \ \ \exists N \in \mathbb{N} \ \ \forall n > N : a_n < -M,$$

so heißt die Folge (a_n) **bestimmt divergent**. Man schreibt

$$\lim_{n\to\infty} a_n = \infty \text{ oder } a_n \to \infty \quad \text{bzw.} \quad \lim_{n\to\infty} a_n = -\infty \text{ oder } a_n \to -\infty.$$

Das bedeutet aber nicht, dass für solche Folgen ein Kovergenzbegriff definiert wird.

4.11 Beispiele: $\lim\limits_{n\to\infty} n = \infty$, $\lim\limits_{n\to\infty} \left(-n^{1/1000}\right) = -\infty$.
Die Folge $a_n = (-1)^n n$ ist nicht bestimmt divergent.

4.3 Reihen

4.12 Definition: Sei (a_k) eine Folge in einem normierten Vektorraum (z.B. $\mathbb{R}$, $\mathbb{R}^n$, $\mathbb{C}$, $\mathbb{C}^n$).

1) Die (unendliche) **Reihe** $\sum\limits_{k=0}^{\infty} a_k$ bezeichnet die Folge der **Partialsummen** $(s_n)_{n\in\mathbb{N}_0}$ mit $s_n := \sum\limits_{k=0}^{n} a_k$, d.h.
$$\sum_{k=0}^{\infty} a_k \;:=\; (s_n)_{n\in\mathbb{N}_0} \;=\; \left(\sum_{k=0}^{n} a_k\right)_{n\in\mathbb{N}_0}.$$
Die Folgenglieder a_k heißen auch die **Summanden** der Reihe.

2) Die Reihe $\sum\limits_{k=0}^{\infty} a_k$ heißt **konvergent**, falls (s_n) konvergiert, sonst **divergent**.

3) Falls die Reihe konvergiert, schreibt man
$$\sum_{k=0}^{\infty} a_k \;:=\; \lim_{n\to\infty} s_n \;=\; \lim_{n\to\infty} \sum_{k=0}^{n} a_k\,.$$

4) Eine relle Reihe (d.h. die Summanden sind reell) heißt **bestimmt divergent**, falls
$$\forall K \in \mathbb{R} \;\; \exists N_K \in \mathbb{N} \;\; \forall n > N_K : \sum_{k=0}^{n} a_k > K$$
(d.h. $s_n \to \infty$) bzw.
$$\forall K \in \mathbb{R} \;\; \exists N_K \in \mathbb{N} \;\; \forall n > N_K : \sum_{k=0}^{n} a_k < -K$$
(d.h. $s_n \to -\infty$). Schreibweise:
$$\sum_{k=0}^{\infty} a_k \;=\; \infty \quad \text{bzw.} \quad \sum_{k=0}^{\infty} a_k \;=\; -\infty.$$

5) Genauso $\sum\limits_{k=k_0}^{\infty} a_k$ mit $k_0 \in \mathbb{Z}$.

Achtung: $\sum\limits_{k=0}^{\infty} a_k$ bzeichnet zwei verschiedene Dinge: **(i)** Die Teilsummenfolge (s_n),
(ii) den Grenzwert der Teilsummenfolge.

4.13 Bemerkung: Jede Folge kann als Reihe dargestellt werden:
$$x_n = \underbrace{x_1}_{=:a_1} + \underbrace{x_2 - x_1}_{=:a_2} + \underbrace{x_3 - x_2}_{=:a_3} + \ldots + \underbrace{x_n - x_{n-1}}_{=:a_n} \quad \Rightarrow \quad \sum_{k=1}^{\infty} a_k = (x_n).$$

4.14 Beispiele: **1)** $a_k = 1: \ s_n = n+1, \ \sum_{k=0}^{\infty} 1 = \infty$ (bestimmt divergent).

2) Die Geometrische Reihe: $a_k = q^k, \ q \in \mathbb{C} \setminus \{1\}$

$$\begin{aligned} s_n \ &= \ \sum_{k=0}^{n} q^k \ = \ \frac{1-q^{n+1}}{1-q} \\ |q| < 1 \ \Rightarrow \ \sum_{k=0}^{\infty} q^k \ &= \ \frac{1}{1-q} \\ |q| > 1 \ \Rightarrow \ \sum_{k=0}^{\infty} q^k & \text{ ist divergent} \\ q \in \mathbb{R} \ \wedge \ q > 1 \ \Rightarrow \ \sum_{k=0}^{\infty} q^k \ &= \ \infty \end{aligned}$$

3) Wichtigste divergente Reihe: **Harmonische Reihe** $\sum_{k=1}^{\infty} \frac{1}{k}$:

Aus $\frac{1}{m+1} + \frac{1}{m+2} + \ldots + \frac{1}{2m} \geq \frac{1}{2m} + \frac{1}{2m} + \ldots + \frac{1}{2m} = \frac{1}{2}$ folgt

$$\begin{aligned} s_{2^N} \ = \ \sum_{k=1}^{2^N} \frac{1}{k} \ &= \ 1 + \frac{1}{2} + \underbrace{\frac{1}{3} + \frac{1}{4}}_{\geq \frac{1}{2}} + \underbrace{\ldots}_{\geq \frac{1}{2}} + \ldots + \underbrace{\frac{1}{2^{N-1}+1} + \ldots + \frac{1}{2^{N-1}+2^{N-1}}}_{\geq \frac{1}{2}} \\ &\geq \ 1 + \frac{N}{2} \end{aligned}$$

$\Rightarrow \ s_n \geq 1 + \frac{N}{2}$ für $n \geq 2^N$ für beliebiges $N \in \mathbb{N}$.

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

Sehr langsame Divergenz: $s_{1000} = 7,4\ldots; s_{10\,000} = 9,7\ldots$.

Auf jedem Taschenrechner konvergiert die harmonische Reihe.

4) Dezimalbrüche: $\pi = 3,14159\ldots$ bedeutet

$$\pi \ = \ \sum_{k=0}^{\infty} \frac{a_k}{10^k} \qquad \text{mit } a_0 = 3, \ a_1 = 1, \ a_2 = 4, \ \ldots .$$

Die Teilsummenfolge ist monoton: $s_{n+1} - s_n = \frac{a_{n+1}}{10^{n+1}} \geq 0$

und beschränkt: $s_n \leq \sum_{k=0}^{n} \frac{9}{10^k} \leq 9 \sum_{k=0}^{\infty} \left(\frac{1}{10}\right)^k = 9 \, \frac{1}{1-\frac{1}{10}}$.

$\Rightarrow$ Konvergenz (vgl. Hauptsatz über monotone Folgen 2.38).

4.15 Elementare Eigenschaften: **1)** Sind $\sum_{k=0}^{\infty} a_k$, $\sum_{k=0}^{\infty} b_k$ konvergent und $\lambda \in \mathbb{K}$, dann sind auch $\sum_{k=0}^{\infty}(a_k + b_k)$ und $\sum_{k=0}^{\infty} \lambda\, a_k$ konvergent mit

$$\begin{aligned} \sum_{k=0}^{\infty}(a_k + b_k) &= \sum_{k=0}^{\infty} a_k + \sum_{k=0}^{\infty} b_k \\ \sum_{k=0}^{\infty} \lambda\, a_k &= \lambda \cdot \sum_{k=0}^{\infty} a_k \end{aligned}$$

2) Cauchy-Kriterium: Ist $\sum_{k=0}^{\infty} a_k$ eine Reihe in einem vollständigen Vektorraum, dann ist sie genau dann konvergent wenn

$$\forall \varepsilon > 0 \;\; \exists N_\varepsilon \in \mathbb{N} \;\; \forall n, m > N_\varepsilon : \quad \underbrace{\left\| \sum_{k=n+1}^{m} a_k \right\|}_{=\|s_m - s_n\| = d(s_m, s_n)} < \varepsilon$$

3) Sind (a_k) und (b_k) Folgen, die sich nur an endlich vielen Stellen unterscheiden, so ist $\sum_{k=0}^{\infty} a_k$ genau dann konvergent, wenn $\sum_{k=0}^{\infty} b_k$ konvergiert. (Die Grenzwerte können verschieden sein.)

4) Ist $\sum_{k=0}^{\infty} a_k$ konvergent, dann ist $\lim_{k\to\infty} a_k = 0$. Die Umkehrung gilt nicht: $\sum_{k=1}^{\infty} \frac{1}{k}$.
Nullfolge-Kriterium: $\neg(a_n \to 0) \;\Rightarrow\; \sum a_k$ ist divergent.

5) Ist $a_k \in \mathbb{R}$, $a_k \geq 0$ für $k \in \mathbb{N}$, und sind die Partialsummen beschränkt, so ist $\sum_{k=0}^{\infty} a_k$ konvergent.

4.4 Konvergenzkriterien für Reihen

4.16 Leibniz-Kriterium: Sei (a_k) in $\mathbb{R}$ eine positive, monoton fallende Nullfolge, d.h.

$$a_k \geq 0, \;\; a_k \geq a_{k+1} \text{ und } \lim_{k\to\infty} a_k = 0.$$

Dann ist die alternierende Reihe $\sum_{k=0}^{\infty}(-1)^k\, a_k$ konvergent. und es gilt

$$\left| \sum_{k=0}^{\infty}(-1)^k\, a_k - \sum_{k=0}^{n}(-1)^k\, a_k \right| \;\leq\; a_{n+1}.$$

Beweis: Setze $s_n := \sum_{k=0}^{n} (-1)^k\, a_k$.

1) $(s_{2k})_{k\in\mathbb{N}}$ ist monoton fallend:

$$s_{2(k+1)} - s_{2k} \;=\; a_{2k+2} - a_{2k+1} \;\leq\; 0$$

2) $(s_{2k+1})_{k\in\mathbb{N}}$ ist monoton wachsend: Genauso

3) Es gilt: $s_1 \;\leq\; s_{2k+1} \;=\; s_{2k} - a_{2k+1} \;\leq\; s_{2k} \;\leq\; s_0$.

$\Rightarrow$ $(s_{2k}), (s_{2k+1})$ sind monoton und beschränkt, also konvergent.
Wegen $s_{2k+1} - s_{2k} = -a_{k+1} \to 0$ gilt

$$\lim_{k\to\infty} s_{2k} \;=\; \lim_{k\to\infty} s_{2k+1} \;=:\; s.$$

$\Rightarrow$ (s_n) ist konvergent, $\lim\limits_{n\to\infty} s_n = s$.

4) Fehlerabschätzung:

$$\begin{aligned} s_{2k-1} \leq s \leq s_{2k} \quad &\Rightarrow \quad |s - s_{2k-1}| = s - s_{2k-1} \leq s_{2k} - s_{2k-1} = a_{2k},\\ s_{2k+1} \leq s \leq s_{2k} \quad &\Rightarrow \quad |s - s_{2k}| = s_{2k} - s \leq s_{2k} - s_{2k+1} = a_{2k+1}. \end{aligned}$$

$\square$

4.17 Beispiel: $\displaystyle\sum_{k=1}^{\infty} \frac{(-1)^{k-1}}{k}$ ist konvergent (gegen $\ln 2$).

4.18 Definition: Die Reihe $\sum a_n$ heißt **absolut konvergent**, falls $\sum \|a_n\|$ konvergiert bzw. falls $\sum |a_n|$ konvergiert (wenn (a_n) reelle oder komplexe Folge). Eine konvergente Reihe, die nicht absolut konvergiert, heißt **bedingt konvergent**.

4.19 Beispiel: $\displaystyle\sum_{k=1}^{\infty} \frac{(-1)^{k}}{k}$ ist bedingt konvergent, denn $\displaystyle\sum_{k=1}^{\infty} \frac{1}{k} = \infty$.

4.20 Satz: In einem vollständigen Vektorraum ist jede absolut konvergente Reihe konvergent.

Beweis: Cauchy-Kriterium:

$$\|s_m - s_n\| \;=\; \left\| \sum_{k=n+1}^{m} a_k \right\| \;\leq\; \sum_{k=n+1}^{m} \|a_k\| \;<\; \varepsilon \qquad \text{für } m \geq n > N_\varepsilon.$$

$\square$

4.21 Majorantenkriterium: Es sei (a_n) eine Folge in einem vollständigen Vektorraum und (c_n) eine reelle Folge. Falls

$$\exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : \|a_n\| \leq c_n \ \wedge \ \sum_{n=0}^{\infty} c_n \text{ konvergent},$$

dann konvergiert auch $\sum a_n$ absolut.

Die Reihe $\sum c_n$ heißt **Majorante** für $\sum a_n$.

Beweis: Cauchy-Kriterium:

$$\|s_m - s_n\| \;=\; \left\| \sum_{k=n+1}^{m} a_k \right\| \;\leq\; \sum_{k=n+1}^{m} \|a_k\| \;\leq\; \sum_{k=n+1}^{m} c_k \;<\; \varepsilon \qquad \text{für } m \geq n > N_\varepsilon.$$

□

4.22 Minorantenkriterium: Sind $(a_n), (c_n)$ **relle** Folgen, und gilt

$$\exists n_0 \in \mathbb{N} \quad \forall n \geq n_0 : a_n \geq c_n \geq 0 \ \wedge \ \sum_{n=0}^{\infty} c_n = \infty,$$

dann ist $\sum a_n$ bestimmt divergent: $\displaystyle\sum_{n=0}^{\infty} a_n = \infty$.

4.23 Beispiele: **1)** $\displaystyle\sum_{n=0}^{\infty} \frac{(-1)^n}{2^n + 3^n}$ konvergiert absolut wegen

$$\frac{1}{2^n + 3^n} \leq \frac{1}{3^n}, \qquad \sum_{n=0}^{\infty} \left(\frac{1}{3}\right)^n = \frac{1}{1 - \frac{1}{3}} \text{ (konvergent).}$$

2) $\displaystyle\sum_{n=1}^{\infty} \frac{1}{\sqrt{n}} = \infty$ wegen $\displaystyle\frac{1}{\sqrt{n}} \geq \frac{1}{n}$ für $n \in \mathbb{N}$.

3) $\displaystyle\sum_{n=0}^{\infty} \frac{n}{4^n}$ konvergiert absolut wegen

$$\frac{n}{4^n} \;=\; \frac{n}{2^n (1+1)^n} \;\overset{\text{Bernoullie Ungl.}}{\leq}\; \frac{n}{2^n\, n} \;=\; \frac{1}{2^n} \text{ und } \sum_{n=0}^{\infty} \frac{1}{2^n} = \frac{1}{1 - \frac{1}{2}}.$$

4) $\displaystyle\sum_{k=1}^{\infty} \frac{1}{k^2}$ konvergiert wegen $\displaystyle\frac{1}{k^2} \leq \frac{1}{k(k-1)}$ für $k \geq 2$ und

$$\begin{aligned} \frac{1}{k(k-1)} \;&=\; \frac{1}{k-1} - \frac{1}{k} \\ \Rightarrow \sum_{k=2}^{n} \frac{1}{k(k-1)} \;&=\; \sum_{k=2}^{n} \frac{1}{k-1} - \sum_{k=2}^{n} \frac{1}{k} \;=\; 1 - \frac{1}{n} \;\to\; 1. \end{aligned}$$

4.24 Wurzelkriterium: Es sei (a_n) eine Folge in einem vollständigen Vektorraum.

1) Falls es ein q mit $0 \leq q < 1$ gibt, so dass

$$\forall n \geq n_0 : \sqrt[n]{\|a_n\|} \leq q$$

(für irgendein $n_0 \in \mathbb{N}$), dann ist $\sum_{n=0}^{\infty} a_n$ absolut konvergent.

Aus $\sqrt[n]{\|a_n\|} \geq 1$ für $n \geq n_0$ folgt Divergenz.

2) Falls die Folge $\left(\sqrt[n]{\|a_n\|}\right)_{n\in\mathbb{N}}$ konvergiert, setze

$$q \ := \ \lim_{n\to\infty} \sqrt[n]{\|a_n\|}.$$

Dann gilt:

$$\begin{aligned}
q < 1 \quad &\Rightarrow \quad \sum a_n \text{ konvergiert absolut,}\\
q > 1 \quad &\Rightarrow \quad \sum a_n \text{ divergiert,}\\
q = 1 \quad &\Rightarrow \quad \text{Mit Wurzelkriterium keine Aussage möglich.}
\end{aligned}$$

Beweis: **1)** **a)** $\left.\begin{array}{l} \forall n \geq n_0 : \|a_n\| \leq q^n \\ \sum q^n \text{ konvergent} \end{array}\right\} \Rightarrow \sum a_n$ absolut konvergent

b) $\|a_n\| \geq 1^n = 1$ für $n \geq n_0 \ \Rightarrow \ \neg(a_n \to 0) \Rightarrow \ \sum a_n$ divergent.

2) Sei $q < 1$. Setze $\varepsilon := \dfrac{1-q}{2} > 0. \Rightarrow$ Für $n > N_\varepsilon$ gilt $\sqrt[n]{\|a_n\|} < q + \varepsilon = \dfrac{q+1}{2} < 1.$ $\square$

4.25 Beispiel: $\sum_{n=0}^{\infty} \dfrac{n}{4^n}$ konvergiert absolut wegen

$$\lim_{n\to\infty} \sqrt[n]{\frac{n}{4^n}} \ = \ \lim_{n\to\infty} \frac{n^{1/n}}{4} \ = \ \frac{1}{4} \lim_{n\to\infty} (n^{1/n}) \ \overset{\text{siehe unten}}{=} \ \frac{1}{4} \ < \ 1.$$

4.26 Satz: Es gilt $\lim_{n\to\infty} n^{1/n} = 1$.

Beweis: $n > 1 \ \Rightarrow \ n^{1/n} > 1$.

Setze $\delta_n := n^{1/n} - 1 \ \Rightarrow \ n^{1/n} = 1 + \delta_n, \ \delta_n > 0$

$$\begin{aligned}
&\Rightarrow \ n \ = \ (1+\delta_n)^n \ = \ \sum_{j=0}^{n} \underbrace{\binom{n}{j} \delta_n^j \, 1^{n-j}}_{\geq 0} \ \geq \ \binom{n}{2} \delta_n^2 \ = \ \frac{n(n-1)}{2} \delta_n^2\\
&\Rightarrow \ \delta_n^2 \ \leq \ \frac{2n}{n(n-1)} \ \to \ 0\\
&\Rightarrow \ \delta_n \ \to \ 0\\
&\Rightarrow \ n^{1/n} \ = \ 1 + \delta_n \ \to \ 1
\end{aligned}$$

$\square$

4.27 Quotientenkriterium: Sei (a_n) eine relle oder komplexe Folge mit $a_n \neq 0$ für $n \geq n_0$.

1) Falls es ein q mit $0 \leq q < 1$ gibt mit

$$\forall n \geq n_0 : \left|\frac{a_{n+1}}{a_n}\right| \leq q,$$

dann ist $\sum_{n=0}^{\infty} a_n$ absolut konvergent.

Aus $\left|\frac{a_{n+1}}{a_n}\right| \geq 1$ für $n \geq n_0$ folgt Divergenz.

2) Falls die Folge $\left(\left|\frac{a_{n+1}}{a_n}\right|\right)$ konvergiert sei $q := \lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right|$. Dann gilt:

$$\begin{aligned} q < 1 \quad &\Rightarrow \quad \sum a_n \text{ konvergiert absolut,} \\ q > 1 \quad &\Rightarrow \quad \sum a_n \text{ divergiert,} \\ q = 1 \quad &\Rightarrow \quad \text{Mit Quotientenkriterium keine Aussage möglich.} \end{aligned}$$

4.28 Beispiele: **1)** $\sum_{n=0}^{\infty} n^2 z^n$ mit $z \in \mathbb{C}$ ist absolut konvergent für $|z| < 1$ und divergent für $|z| \geq 1$:

$$\begin{aligned} \text{Wurzelkriterium:} \quad & \sqrt[n]{|n^2 z^n|} = \sqrt[n]{n^2}\,|z| = \left(n^{1/n}\right)^2 |z| \to |z|, \\ \text{Quotientenkriterium:} \quad & \left|\frac{a_{n+1}}{a_n}\right| = \left|\frac{(n+1)^2 |z|^{n+1}}{n^2 z^n}\right| = \left(\frac{n+1}{n}\right)^2 |z| \to |z| \end{aligned}$$

Für $|z| = 1$: $|n^2 z^n| = n^2 \to \infty \;\Rightarrow\; \neg(n^2 z^n \to 0) \;\Rightarrow\; \sum n^2 z^n$ ist divergent.

2) $\sum_{n=0}^{\infty} \frac{z^n}{n!}$ konvergiert absolut für alle $z \in \mathbb{C}$:

$$\text{Quotientenkriterium:} \quad \left|\frac{\frac{z^{n+1}}{(n+1)!}}{\frac{z^n}{n!}}\right| = |z|\frac{1}{n+1} \to 0.$$

3) $\sum_{n=1}^{\infty} \frac{1}{n^2}$ ist konvergent, und $\sqrt[n]{\frac{1}{n^2}} \to 1$ bzw. $\frac{\frac{1}{(n+1)^2}}{\frac{1}{n^2}} \to 1$

$\sum_{n=1}^{\infty} \frac{1}{n}$ ist divergent, und $\sqrt[n]{\frac{1}{n}} \to 1$ bzw. $\frac{\frac{1}{n+1}}{\frac{1}{n}} \to 1$

4) $\sum_{n=0}^{\infty} \frac{(2+(-1)^n)^n}{4^n}$ ist absolut konvergent:

$$\text{Wurzelkriterium:} \quad \sqrt[n]{|a_n|} = \begin{cases} \sqrt[n]{\frac{3^n}{4^n}} = \frac{3}{4} & \text{für gerades } n, \\ \sqrt[n]{\frac{1^n}{4^n}} = \frac{1}{4} & \text{für ungerades } n. \end{cases} \quad \Rightarrow \text{ absolute Konv.}$$

Aber Quotientenkriterium: $$\left|\frac{a_{n+1}}{a_n}\right| = \begin{cases} \frac{\left(\frac{1}{4}\right)^{n+1}}{\left(\frac{3}{4}\right)^n} = \frac{1}{4\cdot 3^{n+1}} & \text{für gerades } n \\ \frac{\left(\frac{3}{4}\right)^{n+1}}{\left(\frac{1}{4}\right)^n} = \frac{3^{n+1}}{4} & \text{für ungerades } n \end{cases}$$

Mit dem Quotientenkriterium ist keine Aussage möglich!

Das Wurzelkriterium ist stärker als das Quotientenkrtiterium (ohne Beweis).

4.29 Produkt von Reihen: Es seien $\sum a_n$ und $\sum b_n$ absolut konvergent und

$$c_n := \sum_{j=0}^{n} a_j \cdot b_{n-j}.$$

Dann konvergiert das **Cauchy-Produkt** $\sum c_n$ absolut, und für die Grenzwerte gilt

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

4.30 Beispiele: **1)** $\frac{1}{(1-q)^2} = \left(\sum_{n=0}^{\infty} q^n\right)\left(\sum_{n=0}^{\infty} q^n\right)$, Reihe ist für $|q| < 1$ absolut konvergent.

$$c_n = \sum_{j=0}^{n} q^j\, q^{n-j} = \sum_{j=0}^{n} q^n = (n+1)q^n$$

$$\Rightarrow \frac{1}{(1-q)^2} = \sum_{n=0}^{\infty}(n+1)q^n = \sum_{n=1}^{\infty} n\, q^{n-1}.$$

2) $\sum_{n=0}^{\infty} \frac{(-1)^n}{\sqrt{n+1}}$ ist bedingt konvergent.

Cauchy-Produkt mit sich selber:

$$c_n = \sum_{j=0}^{n} \frac{(-1)^j}{\sqrt{j+1}} \frac{(-1)^{n-j}}{\sqrt{n-j+1}} = (-1)^n \sum_{j=0}^{n} \frac{1}{\sqrt{j+1}\sqrt{n-j+1}}.$$

Die Produktreihe $\sum c_n$ ist divergent:

Für festes $n \in \mathbb{N}$ betrachte

$$\begin{aligned} f(x) &= x(n+2-x) \\ &= -\underbrace{\left(x^2 - (n+2)x + \left(\frac{n+2}{2}\right)^2\right)}_{=(x-\frac{n+2}{2})^2 \leq 0} + \left(\frac{n+2}{2}\right)^2 \leq \left(\frac{n+2}{2}\right)^2. \end{aligned}$$

$\Rightarrow$ für $0 < x < n+2$ gilt $\frac{1}{f(x)} \geq \left(\frac{2}{n+2}\right)^2$ bzw. $\frac{1}{\sqrt{f(x)}} \geq \frac{2}{n+2}$

$\Rightarrow \frac{1}{\sqrt{j+1}\sqrt{n-j+1}} = \frac{1}{\sqrt{f(j+1)}} \geq \frac{2}{n+2}$

$\Rightarrow \; |c_n| \geq \sum\limits_{j=0}^{n} \dfrac{2}{n+2} = \dfrac{2(n+1)}{n+2}$
$\Rightarrow \; \neg(c_n \to 0)$
$\Rightarrow \; \sum c_n$ ist divergent.

4.5 Potenzreihen

4.31 Definition: Sei $z_0 \in \mathbb{C}$ und (a_n) eine komplexe Folge. Dann heißt

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

Potenzreihe um $\boldsymbol{z_0}$ mit den Koeffizienten a_n. Falls $z, z_0 \in \mathbb{R}$ und (a_n) reelle Folge, so heißt die Potenzreihe **reell**.

4.32 Beispiel:

$$\frac{1}{2-z} \;=\; \frac{1}{2}\,\frac{1}{1-\frac{z}{2}} \;=\; \frac{1}{2}\sum_{n=0}^{\infty}\left(\frac{z}{2}\right)^n \;=\; \sum_{n=0}^{\infty}\left(\frac{1}{2}\right)^{n+1} z^n \qquad \text{für } \left|\frac{z}{2}\right| < 1 \text{ bzw. } |z| < 2$$

oder

$$\frac{1}{2-z} \;=\; \frac{1}{1-(z-1)} \;=\; \sum_{n=0}^{\infty}(z-1)^n \qquad \text{für } |z-1| < 1$$

4.33 Konvergenzradius: Sei $\sum a_n\,(z-z_0)^n$ Potenzreihe.

1) Es gibt eine eindeutig bestimmte „Zahl“ $R \in [0, \infty]$, so dass

$$\sum_{n=0}^{\infty} a_n\,(z-z_0)^n \text{ ist } \begin{cases} \text{absolut konvergent} & \text{für } |z-z_0| < R, \\ \text{divergent} & \text{für } |z-z_0| > R. \end{cases}$$

(Für $|z - z_0| = R$ kann alles passieren.) Die Größe R heißt **Konvergenzradius** der Potenzreihe.

2) Falls

$$r := \lim_{n\to\infty}\left|\frac{a_{n+1}}{a_n}\right| \qquad \text{oder} \qquad r := \lim_{n\to\infty}\sqrt[n]{|a_n|}$$

existiert oder $r = \infty$ (bestimmte Divergenz), so gilt

$$R \;=\; \begin{cases} \dfrac{1}{r} & \text{falls } r > 0 \\ \infty & \text{falls } r = 0 \\ 0 & \text{falls } r = \infty \end{cases}$$

Im Allgemeinen: $r = \limsup\limits_{n\to\infty}\sqrt[n]{|a_n|}$ (Limes superior = größter durch Teilfolge erreichbarer Grenzwert), R wie oben.

Beweis: Sei $z_0 = 0$. Zum Beweis von 2) wende das Wurzelkriterium an:

$$\begin{aligned} \sqrt[n]{|a_n\, z^n|} \;&=\; \sqrt[n]{|a_n|}\,|z| \overset{n\to\infty}{\to} \left(\lim_{n\to\infty} \sqrt[n]{|a_n|}\right)|z| \overset{\text{soll sein}}{<} 1 \\ \Leftrightarrow\; &\begin{cases} |z| < \dfrac{1}{\lim\limits_{n\to\infty} \sqrt[n]{|a_n|}} & \text{falls } \lim\limits_{n\to\infty} \sqrt[n]{|a_n|} > 0 \\ z \text{ beliebig} & \text{falls } \lim\limits_{n\to\infty} \sqrt[n]{|a_n|} = 0 \end{cases} \end{aligned}$$

Genauso: Anwendung des Quotientenkriteriums liefert die andere Formel für den Konvergenzradius.

□

4.34 Beispiele: **1)** $\sum \frac{z^n}{n!}$: $a_n = \frac{1}{n!} \;\Rightarrow\; \left|\frac{a_{n+1}}{a_n}\right| = \frac{1}{n+1} \to 0 \;\Rightarrow\; R = \infty.$

2) $\sum n^2 z^n$: $R = 1$, Reihe ist konvergent für $|z| < 1$, divergent für $|z| \geq 1$ (vgl. Beispiel auf Seite 116).

3) $\sum \frac{2^n}{n^2}(z-1)^n$: $a_n = \frac{2^n}{n^2} \;\Rightarrow\; \left|\frac{a_{n+1}}{a_n}\right| = \frac{n^2}{(n+1)^2}\,2 \to 2$

$$\Rightarrow\; R = \frac{1}{2}, \text{ also: } \quad \begin{array}{l} \text{absolute Konvergenz für } |z-1| < \frac{1}{2} \\ \text{Divergenz für } |z-1| > \frac{1}{2} \end{array}$$

Für $|z-1| = \frac{1}{2}$: $\left|\frac{2^n}{n^2}\,(z-1)^n\right| = \frac{1}{n^2} \;\Rightarrow$ Konvergenz (Majo-Kriterium 4.21).

4) Geometrische Reihe:

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

Also $R = 1$. Die Funktion $f(z) = \frac{1}{1+z^2}$ hat in $z = \pm\mathrm{i}$ eine Singularität (Nennernullstelle). Deshalb kann der Konvergenzradius nicht größer sein.

5) $\sum (3 + (-1)^n)^n\, z^n$: $\sqrt[n]{|a_n|} = \begin{cases} 2, & n \text{ ungerade} \\ 4, & n \text{ gerade} \end{cases} \quad \Rightarrow\; R = \frac{1}{4}.$

4.35 Identitätssatz: Seien $f(z) = \sum_{n=0}^{\infty} a_n\,(z-z_0)^n$ und $g(z) = \sum_{n=0}^{\infty} b_n\,(z-z_0)^n$ Potenzreihen mit positiven Konvergenzradien. Existiert eine Folge (z_k) mit $z_k \to z_0$, $z_k \neq z_0$ und $f(z_k) = g(z_k)$ für $k \in \mathbb{N}$, so folgt $a_n = b_n$, also auch $f = g$.

Insbesondere ist jede Funktion auf höchstens eine Weise als Potenzreihe um z_0 darstellbar.

4.36 Multiplikation: Seien

$$f(z) = \sum_{n=0}^{\infty} a_n (z - z_0)^n \text{ mit Konvergenzradius } R_1 > 0 \text{ und } g(z) = \sum_{n=0}^{\infty} b_n (z - z_0)^n \text{ mit } R_2 > 0.$$

Dann gilt wenigstens für $|z - z_0| < \min\{R_1, R_2\}$:

$$f(z) \cdot g(z) \;=\; \sum_{n=0}^{\infty} c_n (z - z_0)^n \qquad \text{mit } c_n = \sum_{k=0}^{n} a_k \, b_{n-k}.$$

Beweis: $f(z), g(z)$ sind absolut konvergent für $|z| < \min\{R_1, R_2\}$. Aus Satz 4.29 (Cauchy-Produkt von Reihen): Für jedes feste $z \in \mathbb{C}$ mit $|z| < \min\{R_1, R_2\}$ gilt:

$$\begin{aligned} f(z)\, g(z) &= \sum_{n=0}^{\infty} \widetilde{c}_n, \\ \widetilde{c}_n &= \sum_{k=0}^{n} a_k (z - z_0)^k \, b_{n-k} (z - z_0)^{n-k} \;=\; (z - z_0)^n \sum_{k=0}^{n} a_k \, b_{n-k}. \end{aligned}$$

□

4.6 Spezielle Funktionen

4.37 Definition: Wir setzen:

$$\begin{aligned} \mathrm{e}^z &:= \sum_{n=0}^{\infty} \frac{z^n}{n!} && \text{für } z \in \mathbb{C}, \\ \ln z &:= \sum_{n=0}^{\infty} \frac{(-1)^n}{n+1} (z-1)^{n+1} && \text{für } z \in \mathbb{C} \text{ mit } |z - 1| < 1, \\ \sin z &:= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} \, z^{2n+1} && \text{für } z \in \mathbb{C}, \\ \cos z &:= \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} \, z^{2n} && \text{für } z \in \mathbb{C}. \end{aligned}$$

4.38 Eigenschaften: **1)** $\mathrm{e}^{z+w} = \mathrm{e}^z \cdot \mathrm{e}^w$ für $z, w \in \mathbb{C}$, insbesondere gilt $\mathrm{e}^{-z} = \dfrac{1}{\mathrm{e}^z}$.

Weiter gilt $\mathrm{e}^0 = 1$, $\mathrm{e}^1 = \sum_{n=0}^{\infty} \frac{1}{n!} = \mathrm{e}$, $\quad \mathrm{e}^{1/2} \cdot \mathrm{e}^{1/2} = \mathrm{e}^1 = \mathrm{e} \wedge \mathrm{e}^{1/2} > 0 \Rightarrow \mathrm{e}^{1/2} = \sqrt{\mathrm{e}}$.
Mit Induktion folgt $\mathrm{e}^q = q$-te Potenz von e für $q \in \mathbb{Q}$.

2) Später: Die reelle Exponentialfunktion $\mathrm{e}^{(\cdot)} : \mathbb{R} \to]0, \infty[: x \mapsto \mathrm{e}^x$ ist bijektiv. Die Umkehrabbildung stimmt für $|x - 1| < 1$ mit der oben definierten Logarithmusfunktion überein, d.h. es gilt $\ln(\mathrm{e}^x) = x = \mathrm{e}^{\ln x}$.

3) $\sin(-z) = -\sin z$, $\cos(-z) = \cos z$.

4) $\mathrm{e}^{\mathrm{i}z} = \cos z + \mathrm{i}\sin z$ (Eulersche Formel) bzw.

$$\frac{\mathrm{e}^{\mathrm{i}\,z} + \mathrm{e}^{-\mathrm{i}\,z}}{2} = \cos z, \qquad \frac{\mathrm{e}^{\mathrm{i}\,z} - \mathrm{e}^{-\mathrm{i}\,z}}{2\mathrm{i}} = \sin z.$$

Insbesondere gilt die Formel von Moivre:

$$(\cos z + \mathrm{i}\sin z)^n = \left(\mathrm{e}^{\mathrm{i}\,z}\right)^n = \mathrm{e}^{\mathrm{i}\,nz} = \cos(nz) + \mathrm{i}\sin(nz) \qquad \text{für } n \in \mathbb{N}.$$

5) $\sin^2 z + \cos^2 z = 1$. Insbesondere $\left|\mathrm{e}^{\mathrm{i}\,t}\right| = 1$ für $t \in \mathbb{R}$.

6) Additionstheoreme: $\sin(z_1 + z_2) = \sin z_1 \cdot \cos z_2 + \cos z_1 \cdot \sin z_2$,
$\cos(z_1 + z_2) = \cos z_1 \cdot \cos z_2 - \sin z_1 \cdot \sin z_2$

7) Später: Für $z = x \in \mathbb{R}$ stimmen $\sin x$, $\cos x$ mit der rellen (bereits bekannten) Sinus- bzw. Cosinusfunktion überein. Daraus folgt:

a) Die Abbildung $[0, 2\pi[\, \ni t \mapsto \mathrm{e}^{\mathrm{i}t}$ parametrisiert den Einheitskreis in $\mathbb{C}$. Insbesondere gilt $\mathrm{e}^{\mathrm{i}\pi/2} = \mathrm{i}$, $\mathrm{e}^{\mathrm{i}\pi} = -1$, $\mathrm{e}^{\mathrm{i}3\pi/2} = -\mathrm{i}$, $\mathrm{e}^{2\pi\mathrm{i}} = 1$.

b) Die komplexe Exponentialfunktion $\mathrm{e}^{(\cdot)} : \mathbb{C} \to \mathbb{C} : z \mapsto \mathrm{e}^z$ hat die Periode $2\pi\mathrm{i}$:

$$\mathrm{e}^{z+2\pi\mathrm{i}} \;=\; \mathrm{e}^z \qquad \text{für } z \in \mathbb{C}.$$

c) Die komplexe Sinus- und Cosinusfunktion haben die Periode 2π:

$$\sin(z + 2\pi) = \sin z, \quad \cos(z + 2\pi) = \cos z \qquad \text{für } z \in \mathbb{C}, \text{ insbesondere für } z \in \mathbb{R}.$$

4.39 Definition: Sei $V = \mathcal{M}_{n,n}$ der vollständige Vektorraum der $n \times n$-Matrizen mit Norm $\|A\| = \|(a_{ij})\| = n \max\limits_{1 \le i,j \le n} |a_{ij}|$. Wir setzen

$$\mathrm{e}^A \;:=\; \sum_{k=0}^{\infty} \frac{A^k}{k!} \quad \text{für } A \in \mathcal{M}_{n,n} \quad \text{mit } A^0 := E_n \text{ (Einheitsmatrix)}$$

(Exponentialfunktion für Matrizen). Dann gilt

$$\left\|\mathrm{e}^A\right\| \;\le\; \mathrm{e}^{\|A\|}.$$

Weiter sind die Sinus- und Cosinusfunktionen definiert durch

$$\sin(A) \;:=\; \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n+1)!} A^{2n+1} \quad \text{für } A \in \mathcal{M}_{n,n}$$
$$\cos(A) \;:=\; \sum_{n=0}^{\infty} \frac{(-1)^n}{(2n)!} A^{2n} \quad \text{für } A \in \mathcal{M}_{n,n}$$

4.40 Eigenschaften: **1)** $\mathrm{e}^{A+B} = \mathrm{e}^A \cdot \mathrm{e}^B$ für $A, B \in \mathcal{M}_{n,n}$, insbesondere gilt $\mathrm{e}^{-A} = \left(\mathrm{e}^A\right)^{-1}$.

2) $\sin(-A) = -\sin A$, $\cos(-A) = \cos A$.

3) $\mathrm{e}^{\mathrm{i}A} = \cos A + \mathrm{i}\sin A$

4.41 Beispiele: **1)** $A = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \Rightarrow A^n = \begin{pmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{pmatrix}$

$$\Rightarrow \mathrm{e}^A = \sum_{n=0}^{\infty} \frac{1}{n!} \begin{pmatrix} \lambda_1^n & 0 \\ 0 & \lambda_2^n \end{pmatrix} = \begin{pmatrix} \sum \frac{\lambda_1^n}{n!} & 0 \\ 0 & \sum \frac{\lambda_2^n}{n!} \end{pmatrix} = \begin{pmatrix} \mathrm{e}^{\lambda_1} & 0 \\ 0 & \mathrm{e}^{\lambda_2} \end{pmatrix}$$

2) $A = \begin{pmatrix} 1 & \lambda \\ 0 & 1 \end{pmatrix} \Rightarrow \mathrm{e}^A = \begin{pmatrix} \mathrm{e} & \lambda\,\mathrm{e} \\ 0 & \mathrm{e} \end{pmatrix}$ (Übungen).

4.42 Die geometrische Reihe für Matrizen: Für $A \in \mathcal{M}_{n,n}$ mit $\|A\| < 1$ gilt

$$\sum_{n=0}^{\infty} A^n = (E - A)^{-1} \qquad \text{(Neumannsche Reihe).}$$

5 Stetigkeit

5.1 Um was gehts?

5.1 Definition und Satz: Seien (E, d_E) und (F, d_F) metrische Räume und $f : E \supseteq D \to F$, $x_0 \in D$ (D = Definitionsbereich von f). Dann heißt f **stetig in $\boldsymbol{x_0}$**, falls

(i) $\forall \varepsilon > 0 \quad \exists \delta > 0 \quad \forall x \in D : d_E(x, x_0) < \delta \Rightarrow d_F(f(x), f(x_0)) < \varepsilon$

(ε-δ-Kriterium für Stetigkeit)

oder äquivalent

(ii) Für jede Folge (x_n) in D mit $x_n \to x_0$ gilt $f(x_n) \to f(x_0)$, d.h. $\lim\limits_{n\to\infty} f(x_n) = f\Big(\lim\limits_{n\to\infty} x_n\Big)$

(Folgenkriterium für Stetigkeit).

f heißt **stetig**, falls f in jedem Punkt $x_0 \in D$ stetig ist.

Zur Äquivalenz: (i) $\Rightarrow$ (ii): Sei (x_n) Folge in D mit $x_n \to x_0$.
Zu $\varepsilon > 0$ wähle $\delta > 0$, so dass gilt: $\quad d_E(x, x_0) < \delta \Rightarrow d_F(f(x), f(x_0)) < \varepsilon$.
Wähle N_δ mit $d_E(x_n, x_0) < \delta$ für $n > N_\delta$. Für $n > N_\delta$ folgt $d_F(f(x_n), f(x_0)) < \varepsilon$
$\Rightarrow \; f(x_n) \to f(x_0)$.
(ii) $\Rightarrow$ (i): Zeige $\neg$(i) $\Rightarrow$ $\neg$(ii):
$\neg$(i) $\Leftrightarrow \; \exists \varepsilon > 0 \;\; \forall \delta > 0 \;\; \exists x \in D : d_E(x, x_0) < \delta \;\wedge\; d_F(f(x), f(x_0)) \geq \varepsilon$
Für $n \in \mathbb{N}$ wähle $\delta_n := \frac{1}{n}$: $\;\exists x_n \in D : d_E(x_n, x_0) < \frac{1}{n} \;\wedge\; d_F(f(x_n), f(x_0)) \geq \varepsilon$
$\Rightarrow \; x_n \to x_0 \;\wedge\; \neg\big(f(x_n) \to f(x_0)\big)$
$\Rightarrow \; \neg$(ii),

5.2 Diskussion: **1)** Das Bild:

2) Konstante Funktionen sind stetig und $\mathrm{id} : E \to E : x \mapsto x$ ist stetig.
(Das sind die wichtigsten stetigen Funktionen.)

3) Heaviside Funktion, "Einschaltfunktion"

$$h : \mathbb{R} \to \mathbb{R} : x \mapsto \begin{cases} 0, & x < 0 \\ 1, & x \geq 0 \end{cases}$$

ist unstetig in $x_0 = 0$, sonst stetig.

4) $f : \mathbb{R} \to \mathbb{R} : x \mapsto \begin{cases} 1, & x \in \mathbb{Q} \\ 0, & \text{sonst} \end{cases}$

ist nirgends stetig.

5) Unglaublich, aber wahr: $D := \mathbb{Z} \subseteq \mathbb{R}$, $f : D \to \mathbb{R} : n \mapsto \begin{cases} 1 & n \text{ gerade} \\ 0 & n \text{ ungerade} \end{cases}$ ist stetig.

$$f : \mathbb{R} \setminus \{0\} \to \mathbb{R} : x \mapsto \frac{1}{x} \text{ ist stetig.}$$

6) $f : \mathbb{R}^2 \to \mathbb{R} : (x, y) \mapsto x \cdot y$ ist stetig.

5.3 Rechenregeln für stetige Funktionen in Vektorräumen: Seien E, F Vektorräume über $\mathbb{R}$ oder $\mathbb{C}$ und $f, g : E \supseteq D \to F$ stetig in $x_0 \in D$. Dann gelten:

1) $f + g$, λf, $\|f\| : x \mapsto \|f(x)\|$ sind stetig in x_0.

2) Ist $F = \mathbb{R}$ oder $F = \mathbb{C}$, so ist $f \cdot g : x \mapsto f(x)\, g(x)$ stetig in x_0.
Ist zusätzlich $g(x_0) \neq 0$, so sind auch

$$\frac{1}{g} : x \mapsto \frac{1}{g(x)}, \quad \frac{f}{g} : x \mapsto \frac{f(x)}{g(x)}$$

stetig in x_0.

3) Ist $F = \mathbb{C}$, so sind $\mathrm{Re} f$ und $\mathrm{Im} f$ stetig in x_0.

5.4 Anwendung: Aus 1): $C(\mathbb{R} \to \mathbb{R}) := \{f : \mathbb{R} \to \mathbb{R} \mid f \text{ ist stetig}\}$ ist ein Vektorraum.

5.5 Komposition stetiger Funktionen: Sei f stetig in x_0, g stetig in $f(x_0)$, dann ist auch $g \circ f : x \mapsto g(f(x))$ stetig in x_0.

5.6 Beispiele: **1)** Polynom $P : \mathbb{C} \to \mathbb{C} : z \mapsto a_0 + a_1 z + \ldots + a_n z^n$ ist stetig auf ganz $\mathbb{C}$.

2) P, Q Polynome, $D := \{z \in \mathbb{C} : Q(z) \neq 0\}$. Dann ist $\dfrac{P}{Q}$ stetig.

3) Seien $p, q \in \mathbb{N}$ fest. Später: $f : [0, \infty[\to [0, \infty[: x \mapsto x^{1/q}$ ist stetig.
Setze $g(y) := y^p \;\Rightarrow\; g \circ f : x \mapsto x^{p/q}$ ist stetig.

5.2 Stetige Funktionen sind gute Funktionen

5.7 Satz: Sei $f : E \supseteq D \to F$ stetig in $x_0 \in D$ und $f(x_0) \neq y_0$. Dann

$$\exists \delta > 0 \;\; \forall x \in D : d_E(x, x_0) < \delta \Rightarrow f(x) \neq y_0.$$

5.8 Zwischenwertsatz von Bolzano: Sei $f : \mathbb{R} \supseteq [a, b] \to \mathbb{R}$ stetig. Dann nimmt f auch jeden Wert zwischen $f(a)$ und $f(b)$ wenigstens einmal an. Insbesondere:

$$f(a) < 0 \;\wedge\; f(b) > 0 \;\Rightarrow\; \exists x \in \,]a, b[\, : f(x) = 0.$$

Beweisidee: Sei $f(a) < 0 < f(b)$.
Intervallschachtelung: Setze $a_0 := a,\; b_0 := b,\; x_0 := \dfrac{a_0 + b_0}{2}$
Falls $f(x_0) = 0$, sind wir fertig

$$\left.\begin{array}{l} \text{Falls } f(x_0) > 0\text{: Setze } a_1 := a_0,\; b_1 := x_0 \\ \text{Falls } f(x_0) < 0\text{: Setze } a_1 := x_0,\; b_1 := b_0 \end{array}\right\} \Rightarrow \begin{array}{l} f(a_1) < 0 < f(b_1) \\ b_1 - a_1 = \frac{b_0 - a_0}{2}. \end{array}$$

Führe Verfahren fort. Dann Nullstelle in endlich vielen Schritten oder $a_n \to x$ mit $f(x) = 0$.

5.9 Diskussion: **1)** Dazu wichtig: $[a, b]$ hat keine Löcher, also Vollständigkeit.

2) Beweis gibt iteratives Verfahren zur Bestimmung der Nullstelle.

3) Es folgt: Jedes Polynom ungerader Ordnung hat mindestens eine reelle Nullstelle:

$$\frac{P(x)}{a_n} = \frac{a_0 + a_1 x + \ldots + a_n x^n}{a_n} = \underbrace{x^n}_{\substack{> 0 \text{ für } x > 0 \\ < 0 \text{ für } x < 0}} \underbrace{\left(1 + \frac{a_{n-1}}{a_n x} + \ldots + \frac{a_0}{a_n x^n}\right)}_{\to 1 \text{ für } x \to \pm\infty}$$

$$\Rightarrow \;\; \exists x_1, x_2 \in \mathbb{R} : f(x_1) < 0 < f(x_2).$$

Der Zwischenwertsatz garantiert nun die Existenz mindestens eines $x \in \mathbb{R}$ mit $\dfrac{P(x)}{a_n} = 0$, also $P(x) = 0$.

5.10 Satz: Sei $f : \mathbb{R} \supset [a,b] \to \mathbb{R}$ stetig. Dann nimmt f auf $[a,b]$ das Maximum und Minimum an, d.h.

$$\exists x_1, x_2 \in [a,b] \quad \forall x \in [a,b] : f(x_1) \le f(x) \le f(x_2)$$

Verallgemeinerung: Ist $f : D \to F$ stetig und $D \subset \mathbb{R}^n$ beschränkt und abgeschlossen, dann nimmt f auf D das Maximum und Minimum an.

Beweisidee: Sei $M := \sup\{f(x) : a \le x \le b\}$ (eventuell $M = \infty$). Wähle Folge (x_n) in $[a,b]$ mit $f(x_n) \to M$.
Diese Folge besitzt eine Teilfolge $(x_{n_k})_{k \in \mathbb{N}}$, die in $[a,b]$ konvergiert: $x_{n_k} \to x_0 \in [a,b]$ für $k \to \infty$ (ohne Beweis).

$$\Rightarrow \begin{cases} f(x_{n_k}) \to M & \text{für } k \to \infty \\ f(x_{n_k}) \to f(x_0) & \text{für } k \to \infty \end{cases} \quad \Rightarrow \quad M = f(x_0), \text{ insbesondere } M < \infty.$$

5.11 Beispiele: **1)** $f : [-1,2] \to \mathbb{R} : x \mapsto (x-1)(x-2)$: $f(3/2) \le f(x) \le f(-1)$ für $x \in [-1,2]$.

2) $f :]0,1[\to \mathbb{R} : x \mapsto \dfrac{1}{x}$: Weder Minimum noch Maximum wird angenommen.

3) $f : [1,4] \to \mathbb{R} : x \mapsto \begin{cases} 3 & \text{für } x = 1 \\ x & \text{für } 1 < x < 4 \\ 2 & \text{für } x = 4 \end{cases}$: Weder Minimum noch Maximum wird angenommen.

5.12 Folgerung: Sei F normierter Vektorraum und $f : \mathbb{R} \supset [a,b] \to F$ stetig. Dann ist f **beschränkt**, das heißt

$$\exists K \in \mathbb{R} \quad \forall x \in [a,b] : \|f(x)\| \le K.$$

Beweis: Setze $g(x) := \|f(x)\| \;\Rightarrow\; g : [a,b] \to \mathbb{R}$ ist stetig.
$\Rightarrow\; \exists x_1, x_2 \in [a,b] : g(x_1) \le g(x) = \|f(x)\| \le g(x_2)$, d.h. $g(x_2)$ ist obere Schranke für $\|f(x)\|$. $\square$

5.3 Umkehrfunktionen

5.13 Definition: $f : \mathbb{R} \supseteq D \to \mathbb{R}$ heißt **monoton wachsend** (streng monoton wachsend, monoton fallend, streng monoton fallend), falls

$$\forall x, y \in D : x < y \;\Rightarrow\; f(x) \le f(y) \quad (f(x) < f(y),\; f(x) \ge f(y),\; f(x) > f(y)).$$

5.14 Beispiele: **1)** $f : \mathbb{R} \to \mathbb{R} : x \mapsto 1$ ist monoton wachsend und monoton fallend.

2) $f : \mathbb{R} \to \mathbb{R} : x \mapsto x$ ist streng monoton wachsend.

3) $f : \mathbb{R} \to \mathbb{R} : x \mapsto \begin{cases} x \text{ für } x < 1, \\ 1 \text{ für } 1 \leq x \leq 2, \\ x \text{ für } x > 2, \end{cases}$
ist monoton wachsend.

5.15 Satz: Es sei $f : \mathbb{R} \supseteq D \to \mathbb{R}$ streng monoton wachsend oder streng monoton fallend. Dann ist f injektiv.

Beweis: Sei f streng monoton wachsend.

$$x_1 \neq x_2 \;\Rightarrow\; \left\{ \begin{array}{lll} \text{Entweder} & x_1 < x_2 \;\Rightarrow\; & f(x_1) < f(x_2) \\ \text{oder} & x_1 > x_2 \;\Rightarrow\; & f(x_1) > f(x_2) \end{array} \right\} \;\Rightarrow\; f(x_1) \neq f(x_2)$$

□

5.16 Satz: Sei $f : [a, b] \to \mathbb{R}$ stetig und streng monoton wachsend. Dann bildet f das Intervall $[a, b]$ bijektiv auf das Intervall $[f(a), f(b)]$ ab. Die Umkehrfunktion $f^{-1} : [f(a), f(b)] \to [a, b]$ ist stetig und streng monoton wachsend.
Analog für streng monoton fallende Funktionen $f : [a, b] \to [f(b), f(a)]$ und auch für offene Intervalle.

Beweis: **1)** f streng monoton wachsend $\Rightarrow$ f injektiv $\wedge$ $f([a, b]) \subseteq [f(a), f(b)]$.
Da f stetig, folgt mit dem Zwischenwertsatz, dass f surjektiv ist.
$\Rightarrow$ f bijektiv.

2) f^{-1} ist streng monoton wachsend: Sei $y_1 < y_2$ mit $y_j = f(x_j)$.
Annahme: $f^{-1}(y_1) \geq f^{-1}(y_2)$
$\Rightarrow$ $x_1 = f^{-1}(y_1) \geq f^{-1}(y_2) = x_2$ $\Rightarrow$ $y_1 = f(x_1) \geq f(x_2) = y_2$ $\lightning$

3) f^{-1} ist stetig: Seien $y_0 = f(x_0)$ und $\varepsilon > 0$ fest. Wegen der strengen Monotonie gilt $f(x_0 - \varepsilon) < f(x_0) < f(x_0 + \varepsilon)$. Wähle nun $\delta > 0$ so klein, dass $f(x_0 - \varepsilon) < f(x_0) - \delta$ und $f(x_0) + \delta < f(x_0 + \varepsilon)$. Für $|y - y_0| < \delta$ gilt nun

$$\begin{aligned} y \in \,]f(x_0) - \delta, f(x_0) + \delta[\; &\subseteq\; [f(x_0 - \varepsilon), f(x_0 + \varepsilon)] = f([x_0 - \varepsilon, x_0 + \varepsilon]) \\ \Rightarrow\; f^{-1}(y) &\in [x_0 - \varepsilon, x_0 + \varepsilon] \;\Rightarrow\; |f^{-1}(y) - x_0| < \varepsilon. \end{aligned}$$

□

5.17 Bemerkungen: **1)** $f^{-1} \neq \dfrac{1}{f}$.

2) Graph von $x = f^{-1}(y)$: derselbe wie der von $y = f(x)$.
Graph von $y = f^{-1}(x)$: Spiegelung des Graphen von $y = f(x)$ an $y = x$.

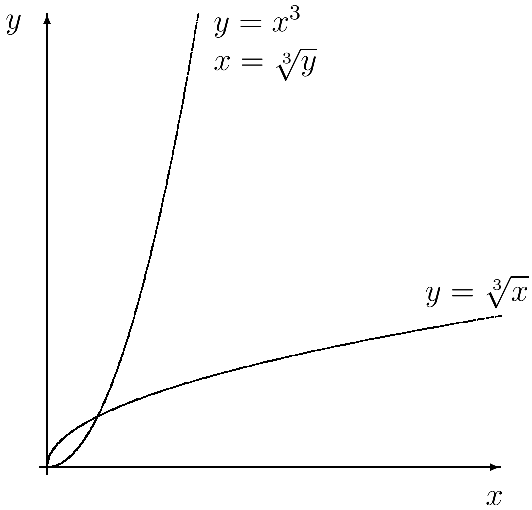
5.18 Wurzelfunktionen: Für festes $n \in \mathbb{N}$ sei $f : \mathbb{R}_0^+ \to \mathbb{R}_0^+ : x \mapsto x^n$.

1) f ist stetig.

2) f ist streng monoton wachsend.

3) f ist surjektiv.

Also existiert die Umkehrfunktion $f^{-1} : \mathbb{R}_0^+ \to \mathbb{R}_0^+ : x \mapsto \sqrt[n]{x}$. Wie oben: f^{-1} ist stetig und streng monoton wachsend.



Allgemeiner folgt:

Für jedes $q \in \mathbb{Q}$ mit $q \geq 0$ ist $f : [0, \infty[\to [0, \infty[: x \mapsto x^q$ stetig.
Für jedes $q \in \mathbb{Q}$ mit $q < 0$ ist $f :]0, \infty[\to]0, \infty[: x \mapsto x^q$ stetig.

5.4 Stetigkeit von Grenzfunktionen

5.19 Beispiele für Funktionenfolgen: **1)** Sei $f_n : \mathbb{R} \to \mathbb{R} : x \mapsto x^{2n}$.

Die Folge $\big(f_n(x)\big)_n$ konvergiert nicht für alle $x \in D = \mathbb{R}$.

$$\begin{aligned} |x| < 1 &\;\Rightarrow\; f_n(x) \to 0 \\ |x| = 1 &\;\Rightarrow\; f_n(x) \to 1 \\ |x| > 1 &\;\Rightarrow\; f_n(x) \to \infty \end{aligned}$$

2) Sei $f_n : \mathbb{R} \to \mathbb{R} : x \mapsto \dfrac{1}{1 + x^{2n}}$.
$\big(f_n(x)\big)_n$ konvergiert für $n \to \infty$ in jedem Punkt $x \in D = \mathbb{R}$.
Wir setzen $f(x) := \lim\limits_{n\to\infty} f_n(x)$ für $x \in D$.

$$\Rightarrow \ f(x) \ = \begin{cases} 1 & |x| < 1 \\ \frac{1}{2} & |x| = 1 \\ 0 & |x| > 1 \end{cases}$$

Alle f_n sind stetig, aber f nicht.

3) Sei $f_n : [-2,2] \to \mathbb{R} : x \mapsto 1 + \dfrac{1}{n}\, x$.

Für jedes $x \in D = [-2,2]$ gilt $f_n(x) \to 1$.

Alle f_n sind stetig, und die Grenzfunktion $f : x \mapsto 1$ auch.

5.20 Definition: Seien $f, f_n : E \supseteq D \to F$ Funktionen.

1) Die Folge (f_n) heißt **punktweise konvergent gegen $\boldsymbol{f}$**, falls $\lim\limits_{n\to\infty} f_n(x) = f(x)$ für alle $x \in D$, d.h.

$$\forall x \in D \ \ \forall \varepsilon > 0 \ \ \exists \underbrace{N_\varepsilon \in \mathbb{N}}_{N_\varepsilon \text{ darf von } x \text{ abhängen}} \ \ \forall n > N_\varepsilon : d_F(f_n(x), f(x)) < \varepsilon.$$

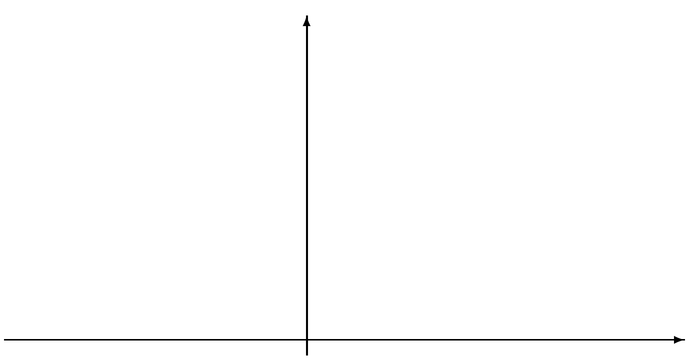
2) Die Folge (f_n) heißt **gleichmäßig konvergent gegen $\boldsymbol{f}$** (auf D), falls

$$\forall \varepsilon > 0 \ \ \exists N_\varepsilon \in \mathbb{N} \ \ \forall n > N_\varepsilon \ \ \forall x \in D : \underbrace{d_F(f_n(x), f(x)) < \varepsilon}_{\text{gilt unabhängig von } x}.$$

Schreibweise: $\lim\limits_{n\to\infty} f_n = f$ punktweise/gleichmäßig; f heißt **Grenzfunktion** der Folge (f_n).

5.21 Diskussion: **1)** $f_n \to f$ gleichmäßig bedeutet:

$$\forall \varepsilon > 0 \ \ \exists N_\varepsilon \in \mathbb{N} \ \ \forall n > N_\varepsilon : \text{Der Graph von } f_n \text{ liegt im } \varepsilon\text{-Schlauch um } f$$



2) Beispiel 1: nicht punktweise konvergent, nicht gleichmäßig konvergent.

Beispiel 2: punktweise konvergent, aber nicht gleichmäßig konvergent.

Beispiel 3: punktweise und gleichmäßig konvergent.

3) gleichmäßig konvergent $\Rightarrow$ punktweise konvergent.
Aber „gleichmäßig konvergent“ ist wesentlich mehr als „punktweise konvergent“.

4) Rechenregeln für konvergente Folgen übertragen sich auf Funktionenfolgen.

5) Reihen: Eine Reihe (bzw. die Folge der Partialsummen) kann punktweise oder gleichmäßig konvergieren, bedingt oder absolut.

5.22 Satz: Der gleichmäßige Grenzwert stetiger Funktionen ist stetig:

$$f_n \text{ stetig für } n \in \mathbb{N} \ \wedge \ f_n \to f \text{ gleichmäßig} \ \Rightarrow \ f \text{ ist stetig.}$$

Beweis: Seien $x_0 \in D$ und $\varepsilon > 0$ fest. Wähle N_ε, so dass

$$d_F\big(f_n(x), f(x)\big) < \frac{\varepsilon}{3} \qquad \text{für } n > N_\varepsilon,\ x \in D.$$

Da $f_{N_\varepsilon+1}$ stetig ist, kann $\delta > 0$ so gewählt werden, dass

$$d_F\big(f_{N_\varepsilon+1}(x), f_{N_\varepsilon+1}(x_0)\big) < \frac{\varepsilon}{3} \qquad \text{für } d_E(x, x_0) < \delta.$$

Für $d_E(x, x_0) < \delta$ folgt

$$\begin{aligned} d_F(f(x), f(x_0)) &\leq d_F\big(f(x), f_{N_\varepsilon+1}(x)\big) + d_F\big(f_{N_\varepsilon+1}(x), f_{N_\varepsilon+1}(x_0)\big) + d_F\big(f_{N_\varepsilon+1}(x_0), f(x_0)\big) \\ &< \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon. \end{aligned}$$

$\square$

5.23 Kriterien für gleichmäßige Konvergenz: **1)** Cauchy-Kriterium: Sei (F, d_F) ein vollständiger metrischer Raum. Falls

$$\forall \varepsilon > 0 \ \ \exists N_\varepsilon \in \mathbb{N} \ \ \forall n, m > N_\varepsilon \ \ \forall x \in D : d_F(f_n(x), f_m(x)) < \varepsilon,$$

dann existiert eine Funktion $f : D \to F$ mit $f = \lim f_n$ gleichmäßig.

2) **Weierstraß-Kriterium:** Falls F vollständiger Vektorraum und (c_n) reelle Folge mit $c_n \geq 0$ und

$$\sum c_n < \infty \text{ und } \|f_n(x)\| \leq c_n \text{ für } x \in D,\ n \geq n_0,$$

dann konvergieren $\sum f_n$ und $\sum \|f_n\|$ gleichmäßig auf D.

Beweis: **1)** Wir betrachten nur den Fall $F = \mathbb{R},\ d(x,y) = |x-y|$. Der allgemeine Fall wird entsprechend bewiesen.

Für festes $x \in D$ ist $\big(f_n(x)\big)_{n\in\mathbb{N}}$ eine Cauchy-Folge. also konvergent.

Definiere $f(x) := \lim\limits_{n\to\infty} f_n(x)$ für $x \in D$.

Aus $\big|f_n(x) - f_m(x)\big| < \varepsilon$ für $m, n > N_\varepsilon$ und $x \in D$ folgt (mit Satz 2.55)

$$|f_n(x) - f(x)| \;=\; \lim_{m\to\infty} |f_n(x) - f_m(x)| \;\leq\; \varepsilon \qquad \text{für } n > N_\varepsilon,\ x \in D.$$

2) Wir zeigen, dass die Teilsummenfolge (s_n) das Cauchy-Kriterium aus 1) erfüllt:

Mit $s_n(x) := \sum\limits_{k=0}^{n} f_k(x)$ gilt

$$\begin{aligned} \|s_n(x) - s_m(x)\| \;&=\; \left\| \sum_{k=m+1}^{n} f(x) \right\| \;\leq\; \sum_{k=m+1}^{n} \|f(x)\| \\ &\leq\; \sum_{k=m+1}^{n} c_k \;<\; \varepsilon \qquad \text{für } n > m > N_\varepsilon. \end{aligned}$$

Also konvergiert (s_n) gleichmäßig gegen f mit $f(x) = \sum\limits_{n=0}^{\infty} f_n(x)$.

$\square$

5.24 Beispiele: **1)** Die geometrische Reihe:

Bekannt: $\sum\limits_{n=0}^{\infty} z^n = \dfrac{1}{1-z}$ punktweise für $|z| < 1$.

a) Sei $0 < q < 1$, q fest. Dann ist $\sum z^n$ gleichmäßig konvergent für $|z| \leq q$.

b) Aber: Auf $\{z \in \mathbb{C} : |z| < 1\}$ ist $\sum z^n$ nicht gleichmäßig konvergent.

2) Sei $f_n :]-1, 1] \to \mathbb{R} : x \mapsto x^n$. f_n ist stetig und es gilt

$$f_n(x) \to f(x) = \begin{cases} 0, & |x| < 1 \\ 1, & x = 1 \end{cases} \quad \text{punktweise.}$$

Die Grenzfunktion ist unstetig, also konvergiert (f_n) nicht gleichmäßig auf $]-1, 1]$.

5.25 Potenzreihen: Es sei $\sum a_n\,(z - z_0)^n$ eine Potenzreihe mit Konvergenzradius $R \in]0, \infty]$. Dann gilt:

1) Ist $\rho \in]0, R[$, so konvergiert die Potenzreihe in $\{z : |z - z_0| \leq \rho\}$ gleichmäßig.

2) Die Grenzfunktion $f(z) = \sum\limits_{n=0}^{\infty} a_n\,(z - z_0)^n$ für $|z - z_0| < R$ ist stetig.

Beweis: **1)** Weierstraß-Kriterium: $|a_n(z-z_0)^n| \le |a_n|\rho^n$, $\sum |a_n\,\rho^n|$ ist konvergent.

2) Sei $|z_1 - z_0| < R \;\Rightarrow\; \exists \rho \in]0, R[: |z_1 - z_0| < \rho$.
Aus 1) und Satz 5.22:
f ist stetig in der Menge $\{z \in \mathbb{C} : |z - z_0| < \rho\}$, also insbesondere in $z = z_1$. $\square$

5.26 Anwendung: **1)** Die Funktionen $\mathrm{e}^{(\cdot)}$, sin, cos sind stetig auf $\mathbb{C}$.

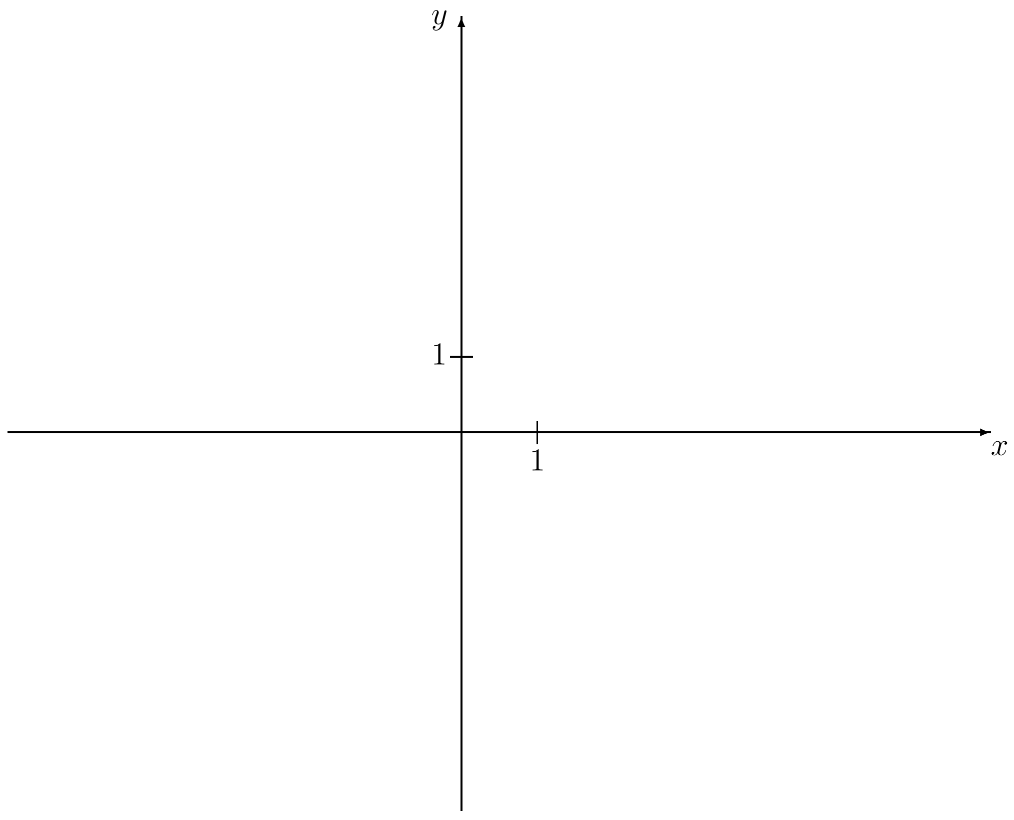
2) Die Funktion ln ist stetig auf $\{z \in \mathbb{C} : |z-1| < 1\}$.

3) Für $A \in \mathcal{M}_{n,n}$ ist die Funktion $\mathbb{R} \ni t \mapsto \mathrm{e}^{t\cdot A} \in \mathcal{M}_{n,n}$ stetig auf $\mathbb{R}$.

5.27 Der relle Logarithmus: Für die Funktion $\mathrm{e}^{(\cdot)} :]-\infty, \infty[\to]0, \infty[$ gelten:

(i) $\mathrm{e}^{(\cdot)}$ ist stetig.
(ii) $\mathrm{e}^{(\cdot)}$ ist streng monoton wachsend, insbesondere injektiv.
(iii) $\mathrm{e}^{(\cdot)}$ ist surjektiv.

Also existiert eine Umkehrfunktion $\ln :]0, \infty[\to]-\infty, \infty[$. Sie ist stetig und streng monoton wachsend.



5.5 Grenzwerte von Funktionen

5.28 Beispiele: **1)** Heaviside-Funktion:

$$h(x) = \begin{cases} 0 & x < 0 \\ 1 & x \geq 0 \end{cases}$$

Für x von oben gegen 0 „strebt“ $h(x)$ gegen 1

Für x von unten gegen 0 „strebt“ $h(x)$ gegen 0

2) Sei $f : \mathbb{R} \supset \mathbb{Z} : n \mapsto n$. Grenzwert von $f(x)$ für $x \to x_0 \in \mathbb{R}$ ist nicht sinnvoll.

5.29 Häufungspunkte: Sei (E, d) ein metrischer Raum und $M \subseteq E$, $x_0 \in E$. Dann heißt x_0 **Häufungspunkt** von M, falls

$$\forall \varepsilon > 0 : \big(B_\varepsilon(x_0) \setminus \{x_0\}\big) \cap M \neq \emptyset$$

oder äquivalent

$$\exists (x_n) : x_n \in M \wedge x_n \neq x_0 \wedge x_n \to x_0.$$

$x_0 \in M$ heißt **isolierter Punkt** von M, falls x_0 kein Häufungspunkt von M ist.

5.30 Definition: Sei $f : E \supseteq D \to F$ und $x_0 \in E$.

1) Schreibe

$$\lim_{x \to x_0} f(x) = a,$$

falls x_0 Häufungspunkt von D ist und für jede Folge (x_n) in D mit $x_n \to x_0$ gilt:

$$\lim_{n \to \infty} f(x_n) = a. \tag{$*$}$$

2) Im Fall $f : \mathbb{R} \supseteq D \to F$ schreibe

$$\lim_{x \downarrow x_0} f(x) = a \quad \Big(\text{oder } \lim_{x \to x_0+} f(x) = a\Big),$$

falls x_0 Häufungspunkt von $D \cap]x_0, \infty[$ ist und $(*)$ für jede Folge (x_n) in D mit $x_n \to x_0$, $x_n > x_0$ gilt.

Analog $\lim\limits_{x \uparrow x_0} f(x) = a$.

3) Falls x_0 Häufungspunkt von D ist und für jede Folge (x_n) in D mit $x_n \to x_0$ die Folge $(f(x_n))$ bestimmt gegen $+\infty$ $(-\infty)$ divergiert, schreibe

$$\lim_{x \to x_0} f(x) \;=\; +\infty \; (-\infty) \qquad (\textbf{uneigentlicher Grenzwert}).$$

4) Alles analog für $x \to +\infty$, $x \to -\infty$.

5.31 Beispiele: $\lim\limits_{x\downarrow 0}\dfrac{1}{x}=\infty,\quad \lim\limits_{x\uparrow 0}\dfrac{1}{x}=-\infty,\quad \lim\limits_{x\to\infty}\dfrac{1}{x}=0,\quad \lim\limits_{x\to 1}\dfrac{x^2-1}{x-1}=2.$

5.32 Bemerkungen: **1)** Die Rechenregeln für konvergente Folgen übertragen sich sinngemäß auf das Rechnen mit Grenzwerten.

2) Gilt $x_0\notin D$ und $\lim\limits_{x\to x_0}f(x)=a$, so ist

$$\tilde f: D\cup\{x_0\}\to F: x\mapsto\begin{cases} f(x) & \text{für } x\in D,\\ a & \text{für } x=x_0\end{cases}$$

stetig in x_0. Man sagt: f ist **stetig ergänzbar** in x_0 oder **stetig fortsetzbar** nach x_0.

z.B.: $f:\mathbb{R}\setminus\{-1\}\to\mathbb{R}: x\mapsto\dfrac{x^2-1}{x+1}=\dfrac{(x-1)(x+1)}{x+1}$

$\Rightarrow\ \widetilde{f}:\mathbb{R}\to\mathbb{R}: x\mapsto x-1$

3) Sei $f:\mathbb{R}\supseteq D\to F$ und $x_0\in D$ Häufungspunkt von $D\cap\,]x_0,\infty[$ und von $D\cap\,]-\infty,x_0[$. Dann sind äquivalent:

(i) f ist stetig in x_0.

(ii) $f(x_0)=\lim\limits_{x\downarrow x_0}f(x)=\lim\limits_{x\uparrow x_0}f(x)$.

5.33 Grenzwerte mit Potenzreihen: Zum Beispiel $\lim\limits_{x\to 0}\dfrac{\sin x-x\cos x}{x^2(\mathrm{e}^x-1)}=?.$

Nenner: $x^2\,(\mathrm{e}^x-1)\ =\ x^2\sum\limits_{n=1}^{\infty}\dfrac{x^n}{n!}\ =\ x^3\underbrace{\sum\limits_{n=1}^{\infty}\dfrac{x^{n-1}}{n!}}_{=:f(x)}$

Da f durch eine Potenzreihe dargestellt wird, ist f stetig. Außerdem gilt $f(0)=\dfrac{1}{1!}=1$.

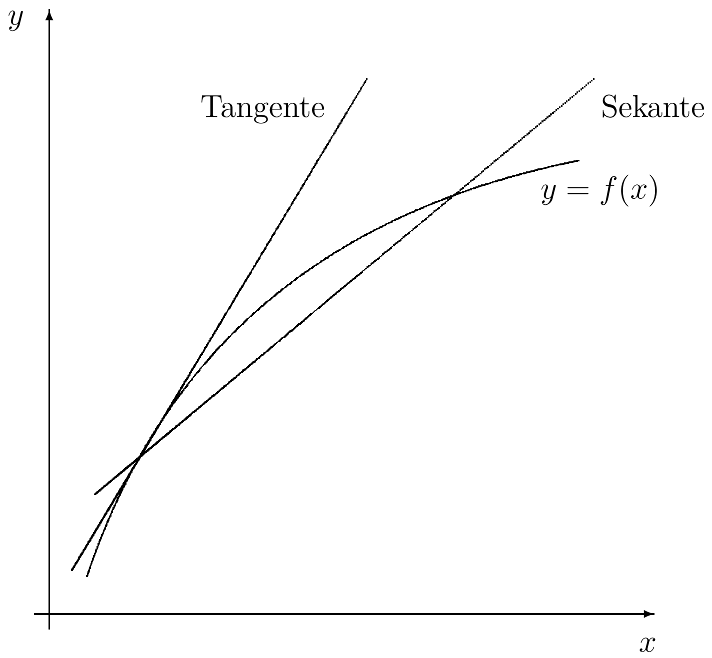
$$\begin{aligned}\text{Zähler: } \sin x-x\cos x\ &=\ \sum_{n=0}^{\infty}\frac{(-1)^n}{(2n+1)!}\,x^{2n+1}-x\sum_{n=0}^{\infty}\frac{(-1)^n}{(2n)!}\,x^{2n}\\ &=\ x^3\underbrace{\left(\sum_{n=1}^{\infty}\frac{(-1)^n}{(2n+1)!}\,x^{2n-2}-\sum_{n=1}^{\infty}\frac{(-1)^n}{(2n)!}\,x^{2n-2}\right)}_{=:g(x)}\end{aligned}$$

g ist stetig mit $g(0)=\dfrac{-1}{3!}-\dfrac{-1}{2!}=\dfrac{1}{2}-\dfrac{1}{6}=\dfrac{1}{3}$.

$$\Rightarrow\ \lim_{x\to 0}\frac{\sin x-x\cos x}{x^2(\mathrm{e}^x-1)}\ =\ \lim_{x\to 0}\frac{x^3\,g(x)}{x^3\,f(x)}\ =\ \lim_{x\to 0}\frac{g(x)}{f(x)}\ =\ \frac{g(0)}{f(0)}\ =\ \frac{1}{3}.$$

6 Differentialrechnung in einer Variablen

6.1 Die Ableitung



Gegeben: $y = f(x)$.
Frage: Tangente in $(x_0, f(x_0))$?

Für $x_1 \to x_0$ “nähert” sich Sekante an Tangente

Sekantengleichung:
$$y = f(x_0) + (x - x_0) \underbrace{\frac{f(x_1) - f(x_0)}{x_1 - x_0}}_{\text{Steigung der Sekante}}$$

6.1 Definition: Sei $\mathbb{K} = \mathbb{R}$ oder $\mathbb{K} = \mathbb{C}$ und $f : \mathbb{K} \supseteq D \to \mathbb{K}$, $x_0 \in D$, x_0 Häufungspunkt von D.

1) Die Abbildung $\Delta f : D \setminus \{x_0\} \to \mathbb{K} : x \mapsto \dfrac{f(x) - f(x_0)}{x - x_0}$ heißt **Differenzenquotient**.

2) f heißt **differenzierbar in x_0**, falls Δf in x_0 stetig ergänzbar ist, d.h. falls

$$\lim_{\substack{x \to x_0 \\ x \in D,\ x \neq x_0}} \frac{f(x) - f(x_0)}{x - x_0} =: f'(x_0) =: \frac{df}{dx}(x_0) =: \left.\frac{df(x)}{dx}\right|_{x=x_0}$$

existiert. In diesem Fall heißt der Grenzwert die **Ableitung von f in x_0** (oder **Differentialquotient**).

3) f heißt **differenzierbar** (in D), falls f in jedem Punkt $x_0 \in D$ differenzierbar ist. Die Funktion f' heißt **Ableitung** von f.

6.2 Beispiele: **1)** Konstante Funktion $f : \mathbb{C} \to \mathbb{C} : z \mapsto c$ ist differenzierbar: $f'(z) = 0$.

2) Identische Funktion id: $\mathbb{C} \to \mathbb{C} : z \mapsto z$ ist differenzierbar: $\mathrm{id}'(z) = 1$.

3) Die Exponentialfunktion $f : \mathbb{C} \to \mathbb{C} : z \mapsto \mathrm{e}^z$ ist differenzierbar: $(\mathrm{e}^z)' = \mathrm{e}^z$.

4) $f : \mathbb{R} \to \mathbb{R} : x \mapsto |x|$ ist in $x_0 = 0$ nicht differenzierbar.

6.3 Satz: $f : \mathbb{K} \supseteq D \to \mathbb{K}$ differenzierbar in $x_0 \;\Rightarrow\; f$ stetig in x_0.

6.4 Bemerkung: Komplexe differenzierbare Funktionen haben erstaunliche Eigenschaften und heißen **holomorphe** oder **analytische** Funktionen $\rightsquigarrow$ Funktionentheorie

6.5 Rechenregeln: Seien $f, g : \mathbb{K} \supseteq D \to \mathbb{K}$ differenzierbar (in $x_0 \in D$). Dann gilt

1) $\lambda \cdot f$ ($\lambda \in \mathbb{K}$ fest) ist differenzierbar (in x_0), und es gilt $(\lambda \cdot f)'(x_0) = \lambda\, f'(x_0)$.

2) $f + g$ ist differenzierbar (in x_0), und es gilt $(f+g)'(x_0) = f'(x_0) + g'(x_0)$.

3) $f \cdot g$ ist differenzierbar (in x_0) mit

$$(f \cdot g)'(x_0) \;=\; f'(x_0)\, g(x_0) + f(x_0)\, g'(x_0) \quad (\textbf{Produktregel}).$$

4) Falls $g(x_0) \neq 0$, ist $\dfrac{f}{g}$ differenzierbar in x_0 mit

$$\left(\frac{f}{g}\right)'(x_0) \;=\; \frac{f'(x_0)\, g(x_0) - f(x_0)\, g'(x_0)}{g^2(x_0)} \quad (\textbf{Quotientenregel}).$$

5) Ableitung der Umkehrfunktion: Sei $D = [a, b]$ und $f : D \to \mathbb{R}$ stetig und streng monoton, $f^{-1} : f(D) \to D$ sei die Umkehrfunktion. Ist f in x_0 differenzierbar und $f'(x_0) \neq 0$, dann ist f^{-1} in $y_0 := f(x_0)$ differenzierbar und

$$\left(f^{-1}\right)'(y_0) \;=\; \frac{1}{f'(x_0)} \quad \left(= \frac{1}{f'\big(f^{-1}(y_0)\big)}\right).$$

6) Seien $f : D \to \mathbb{K}$, $g : D' \to \mathbb{K}$ mit $f(D) \subseteq D'$. Sei f in $x_0 \in D$ differenzierbar, g in $f(x_0) \in D'$ differenzierbar. Dann ist $g \circ f : D \to \mathbb{K}$ in x_0 differenzierbar und

$$(g \circ f)'(x_0) \;=\; \underbrace{g'(f(x_0))}_{\text{äußere Ableitung}} \cdot \underbrace{f'(x_0)}_{\text{innere Ableitung}} \quad (\textbf{Kettenregel}).$$

6.6 Beispiele: **1)** Für festes $n \in \mathbb{Z}$ und $f : \mathbb{C} \setminus \{0\} \to \mathbb{C} : z \mapsto z^n$ gilt $f'(z) = n\, z^{n-1}$.

2) Für festes $q \in \mathbb{Q} \setminus \{0\}$ sei $f :]0, \infty[\to]0, \infty[: x \mapsto x^q$. Dann gilt $f'(x) = q\, x^{q-1}$.

3) $f(x) = \sqrt{x^2 + 1} = \left(x^2 + 1\right)^{1/2}$.

4) Für den reellen Logarithmus $\ln :]0, \infty[\to]0, \infty[: x \mapsto \ln x$ gilt $\dfrac{\mathrm{d} \ln x}{\mathrm{d}x} = \dfrac{1}{x}$.

5) Aus der Schule: $f : [-\frac{\pi}{2}, \frac{\pi}{2}] \to [-1, 1] : x \mapsto \sin x$ ist bijektiv mit $f'(x) = \cos x$. Für die Umkehrabbildung arcsin gilt $\dfrac{\mathrm{d} \arcsin x}{\mathrm{d}x} \;=\; \dfrac{1}{\sqrt{1 - x^2}}$.

6.2 Höhere Ableitungen

6.7 Offene Mengen: Sei (M, d) ein metrischer Raum, $D \subseteq M$.

1) $x_0 \in D$ heißt **innerer Punkt** von D, falls

$$\exists \varepsilon > 0 : B_\varepsilon(x_0) \subseteq D.$$

Falls $M = \mathbb{R}^n$ oder $M = \mathbb{C}^n$ mit der üblichen Metrik, dann ist jeder innere Punkt automatisch Häufungspunkt von D.

2) D heißt **offen**, falls alle Elemente von D innere Punkte sind.

6.8 Beispiele: $]a, b[\subset \mathbb{R}$, $\{z \in \mathbb{C} : |z| < 1\}$ sind offen;
$[a, b[\subset \mathbb{R}$, $\{(x, y) \in \mathbb{R}^2 : \sqrt{x^2 + y^2} \leq 2\} \subset \mathbb{R}^2$ sind nicht offen.

6.9 Definition: Sei $f : \mathbb{K} \supseteq D \to \mathbb{K}$ und $k \in \mathbb{N}$.

1) f heißt im inneren Punkt $x_0 \in D$ ***k*-mal differenzierbar**, falls f auf einer Kugel $B_\varepsilon(x_0) \subseteq D$ $(k-1)$-mal differenzierbar ist und die $(k-1)$-te Ableitung von f in x_0 differenzierbar ist. Schreibe

$$f^{(k)}(x_0) \;:=\; \frac{d^k f}{dx^k}(x_0) \;:=\; \frac{d}{dx}\big(f^{(k-1)}\big)(x_0) \qquad (f^{(2)} =: f'',\ f^{(3)} =: f''',\ f^{(0)} =: f).$$

2) Sei nun D offen. Dann heißt f ***k*-mal differenzierbar auf *D***, falls f in jedem $x_0 \in D$ k-mal differenzierbar ist; $x \mapsto f^{(k)}(x)$ heißt die ***k*-te Ableitung** von f auf D; f heißt manchmal 0-te Ableitung von f.

3) f heißt auf D ***k*-mal stetig differenzierbar**, falls f k-mal differenzierbar auf D und $f^{(k)}$ stetig auf D ist. Die Menge der auf D k-mal stetig differenzierbaren Funktionen bildet einen Vektorraum, bezeichnet mit $\boldsymbol{C^k(D \to \mathbb{K})}$, $C^\infty(D \to \mathbb{K})$ ist der Vektorraum der beliebig oft differenzierbaren Funktionen auf D.

6.10 Bemerkung: Die Rechenregeln gelten analog für höhere Ableitungen. Insbesondere:

$$(f \cdot g)^{(n)} \;=\; \sum_{k=0}^{n} \binom{n}{k} f^{(k)} \cdot g^{(n-k)} \qquad (\text{“Leibnizregel”})$$

(Beweis durch vollständige Induktion.)

6.11 Beispiel: $(x \cdot \mathrm{e}^x)^{(1000)} \;=\; (x + 1000)\,\mathrm{e}^x$.

6.3 Ableitung von Potenzreihen

6.12 Satz: Sei

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

eine (komplexe) Potenzreihe mit Konvergenzradius $R > 0$. Dann hat die Potenzreihe

$$g(z) = \sum_{n=1}^{\infty} n\, a_n (z - z_0)^{n-1}$$

denselben Konvergenzradius, und es gilt $f'(z) = g(z)$ auf $B_R(z_0)$. Das bedeutet, eine Potenzreihe kann auf dem ganzen Konvergenzkreis gliedweise differenziert werden.

Veranschaulichung: Seien R der Konvergenzradius von f, R' der Konvergenzradius von g:

$$R = \frac{1}{\overline{\lim\limits_{n\to\infty}} \sqrt[n]{|a_n|}} \quad \Rightarrow \quad R' = \frac{1}{\overline{\lim\limits_{n\to\infty}} \sqrt[n]{|n a_n|}} = \frac{1}{\overline{\lim\limits_{n\to\infty}} \sqrt[n]{n}\, \sqrt[n]{|a_n|}} = R$$

6.13 Folgerung: Jede Potenzreihe mit positivem Konvergenzradius ist beliebig oft differenzierbar. Eine Funktion, die als Potenzreihe darstellbar ist, heißt **analytische Funktion**.

6.14 Beispiele: **1)** $\dfrac{1}{1-z} = \sum\limits_{n=0}^{\infty} z^n \;\Rightarrow\; \dfrac{1}{(1-z)^2} = \sum\limits_{n=1}^{\infty} n\, z^{n-1}$

$\Rightarrow\; \dfrac{2}{(1-z)^3} = \sum\limits_{n=2}^{\infty} n(n-1)\, z^{n-2}.$

2) $(\mathrm{e}^z)' = \mathrm{e}^z$ für $z \in \mathbb{C}$,
insbesondere gilt für die reelle Exponentialfunktion: $\dfrac{\mathrm{d}\mathrm{e}^x}{\mathrm{d}x} = \mathrm{e}^x$ für $x \in \mathbb{R}$.

3) $\ln(z)' = \dfrac{1}{z}$ für $z \in \mathbb{C}$ mit $|z-1| < 1$. Früheres Beispiel: $(\ln x)' = \dfrac{1}{x}$ für $x \in \,]0, \infty[\,$.

4) $\cos(z)' = -\sin z$ für $z \in \mathbb{C}$, genauso für den reellen Cosinus.

5) $\sin(z)' = \cos z$ für $z \in \mathbb{C}$, genauso für den reellen Sinus.

6) $\arctan(x)' = \dfrac{1}{1+x^2}$ für $x \in \mathbb{R}$.

7) $\arcsin(x)' = \dfrac{1}{\sqrt{1-x^2}}$ für $x \in \,]-1, 1[\,$.

8) $\arccos(x)' = -\dfrac{1}{\sqrt{1-x^2}}$ für $x \in \,]-1, 1[\,$.

6.4 Extrema

6.15 Definition: Sei (M, d) metrischer Raum. Die Funktion $f : M \supseteq D \to \mathbb{R}$ hat in $x_0 \in D$ ein **lokales Maximum** (bzw. **lokales Minimum**), falls

$$\exists \varepsilon > 0 \;\; \forall x \in D : d(x, x_0) < \varepsilon \Rightarrow f(x) \le f(x_0) \quad (\text{bzw. } f(x) \ge f(x_0)).$$

Ein lokales Maximum oder Minimum heißt auch **lokales Extremum**. Falls $f(x) = f(x_0)$ nur für $x = x_0$, so heißt das lokale Extremum **strikt** oder **isoliert**.

Falls

$$\forall x \in D : f(x) \le f(x_0) \; / \; f(x) \ge f(x_0),$$

hat f in x_0 ein **globales Maximum/globales Minimum**.

6.16 Notwendiges Kriterium für Extremum: Sei $f : \mathbb{R} \supseteq D \to \mathbb{R}$ differenzierbar in $x_0 \in D$ und x_0 innerer Punkt von D. Hat f in x_0 ein lokales Extremum, so gilt $f'(x_0) = 0$:

$$f \text{ hat in } x_0 \text{ ein lokales Extremum } \Rightarrow \; f'(x_0) = 0.$$

Ist f differenzierbar in x_0 und $f'(x_0) = 0$, so heißt x_0 **stationärer** oder **kritischer** Punkt von f.

Beweis: f habe ein lokales Maximum in x_0. Dann folgt

$$\left.\begin{array}{rcll} f'(x_0) & = & \displaystyle\lim_{h \downarrow 0} \frac{f(x_0 + h) - f(x_0)}{h} & \le \; 0 \\ f'(x_0) & = & \displaystyle\lim_{h \uparrow 0} \frac{f(x_0 + h) - f(x_0)}{h} & \ge \; 0 \end{array}\right\} \Rightarrow \; f'(x_0) \; = \; 0$$

$\square$

6.17 Beispiele: **1)** $f : \mathbb{R} \to \mathbb{R} : x \mapsto x^4 - x^3 = x^3(x - 1)$:

Kritische Punkte: $x = 0 \;\vee\; x = \dfrac{3}{4}$.

In $x_0 = \dfrac{3}{4}$ hat f ein lokales und globales Minimum, in $x_0 = 0$ kein Extremum.

2) $f : [0, 1] \to \mathbb{R} : x \mapsto x$:

Keine kritischen Punkte, aber bei $x_0 = 1$ hat f ein lokales und globales Maximum, bei $x_0 = 0$ ein lokales und globales Minimum.

6.5 Mittelwertsätze und Anwendungen

6.18 Satz von Rolle: Sei $f : [a, b] \to \mathbb{R}$ stetig, differenzierbar auf $]a, b[$. Dann gilt:

$$f(a) = f(b) \;\Rightarrow\; \exists \xi \in\,]a, b[: f'(\xi) = 0.$$

Beweis: 5.10 $\Rightarrow$ f nimmt auf $[a, b]$ Minimum und Maximum an.

Fall 1) Minimum und Maximum liegen in a und b.

$$\Rightarrow\; f(x) = \text{const} = f(a) \text{ für } x \in [a, b] \;\Rightarrow\; f'(x) = 0 \text{ in } [a, b]$$

Fall 2) Maximum oder Minimum in $x_0 \in\,]a, b[$.

$$\stackrel{6.16}{\Rightarrow}\; f'(x_0) = 0.$$

□

6.19 Mittelwertsatz der Differentialrechnung: Sei $f : [a, b] \to \mathbb{R}$ stetig, differenzierbar auf $]a, b[$. Dann gilt:

$$\exists \xi \in\,]a, b[: f'(\xi) = \frac{f(b) - f(a)}{b - a}.$$

Beweis: Setze $F(x) := f(x) - \dfrac{f(b) - f(a)}{b - a}(x - a)$.

$\Rightarrow$ F erfüllt die Voraussetzungen des Satzes von Rolle ($F(a) = f(a) = F(b)$)

$\Rightarrow$ $\exists \xi \in\,]a, b[: 0 = F'(\xi) = f'(\xi) - \dfrac{f(b) - f(a)}{b - a}$.

□

6.20 Verallgemeinerter Mittelwertsatz: Seien $f, g : [a, b] \to \mathbb{R}$ stetig, differenzierbar auf $]a, b[$. Dann gilt: Ist $g'(x) \neq 0 \;\forall x \in]a, b[$, so gilt

$$\exists \xi \in\,]a, b[: \frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(\xi)}{g'(\xi)}.$$

6.21 Nullableitung: Sei $f : [a, b] \to \mathbb{R}$ stetig, differenzierbar auf $]a, b[$. Dann gilt:

$$\big(\forall x \in\,]a, b[: f'(x) = 0\big) \;\Rightarrow\; f = \text{konstant auf } [a, b].$$

Beweis: Sei $y \in\,]a, b]$. f erfüllt die Voraussetzungen des Mittelwertsatzes im Intervall $[a, y]$.

$$\Rightarrow\; \exists \xi \in\,]a, y[: \underbrace{f'(\xi)}_{=0} = \frac{f(y) - f(a)}{y - a} \;\Rightarrow\; f(y) = f(a)$$

□

6.22 Monotonie: Sei $f : [a, b] \to \mathbb{R}$ stetig, differenzierbar auf $]a, b[$.

1) Falls

$$\forall x \in]a, b[: f'(x) \geq 0 \qquad (f'(x) > 0 \quad / \quad f'(x) \leq 0 \quad / \quad f'(x) < 0),$$

so ist f auf $[a, b]$ monoton wachsend (streng monoton wachsend/monoton fallend/ streng monoton fallend).

2) Ist f auf $[a, b]$ monoton wachsend/fallend, so gilt: $f'(x) \geq 0/f'(x) \leq 0$ auf $]a, b[$.

Beweis: **1)** Folgt aus Mittelwertsatz und $f(x_1) - f(x_2) = f'(\xi)(x_2 - x_1)$

2) $f'(x) = \lim\limits_{x \to x_0} \dfrac{f(x) - f(x_0)}{x - x_0}$. Vorzeichenüberlegung! □

6.23 Lokale Extrema: Sei $D = [a, b] \subseteq \mathbb{R}$, $f : D \to \mathbb{R}$ stetig auf $[a, b]$ und differenzierbar auf $]a, b[$, $x_0 \in]a, b[$.

1) Gibt es ein $\varepsilon > 0$, so dass

$$\begin{aligned} f'(x) \geq 0 \; (f'(x) > 0) \quad &\text{für} \quad x_0 - \varepsilon < x < x_0, \\ f'(x) \leq 0 \; (f'(x) < 0) \quad &\text{für} \quad x_0 < x < x_0 + \varepsilon, \end{aligned}$$

so hat f in x_0 ein lokales (striktes) Maximum: $\forall x \in B_\varepsilon(x_0) : f(x) \leq f(x_0)$ $(f(x) < f(x_0))$.

Entsprechend für Minimum.

2) Ist f in x_0 zweimal differenzierbar und $f'(x_0) = 0$, $f''(x_0) < 0$ $(f''(x_0) > 0)$, so besitzt f in x_0 ein striktes lokales Maximum/Minimum.

3) Ist f in x_0 zweimal differenzierbar und besitzt f in x_0 ein lokales Maximum/Minimum, so ist $f'(x_0) = 0$ und $f''(x_0) \leq 0/ \geq 0$.

Beweis: **1)** Aus letztem Satz:

$$f \text{ ist monoton wachsend für } x_0 - \varepsilon \leq x \leq x_0$$
$$f \text{ ist monoton fallend für } x_0 \leq x \leq x_0 + \varepsilon$$

2) $f'(x_0) = 0 \;\wedge\; \lim\limits_{x \to x_0} \dfrac{f'(x) - 0}{x - x_0} = f''(x_0) < 0$

$$\Rightarrow \quad \exists \varepsilon > 0 : \frac{f'(x)}{x - x_0} < 0 \text{ für } x_0 - \varepsilon < x < x_0 \;\vee\; x_0 < x < x_0 + \varepsilon$$

$$\Rightarrow \quad f'(x) \begin{cases} < 0 & \text{für } x_0 < x < x_0 + \varepsilon \\ > 0 & \text{für } x_0 - \varepsilon < x < x_0 \end{cases}$$

3) 6.16 $\Rightarrow$ $f'(x_0) = 0$. Wäre $f''(x_0) > 0$ $\Rightarrow$ Minimum $\lightning$ $\square$

6.24 Achtung: $f'(x_0) = 0$, $f''(x_0) = 0$ bedeutet gar nichts! Z.B. $f : x \mapsto x^4$ oder $f : x \mapsto x^5$ bei $x_0 = 0$. In diesem Fall muss man höhere Ableitungen betrachten.

6.25 Regeln von de l'Hospital: Seien $f, g\ :]a, b[\to \mathbb{R}$ differenzierbar, $\lim\limits_{x\downarrow a} f(x) = 0 \ \wedge \ \lim\limits_{x\downarrow a} g(x) = 0 \ \wedge \ g'(x) \neq 0$ auf $]a, b[$. Dann gilt:

$$\lim_{x\downarrow a} \frac{f'(x)}{g'(x)} = c \quad \Rightarrow \quad \lim_{x\downarrow a} \frac{f(x)}{g(x)} = c.$$

Beweis: Setze $f(a) := 0$, $g(a) := 0$. Dann gilt für $x \in]a, b[:$ $f, g : [a, x] \to \mathbb{R}$ stetig, differenzierbar auf $]a, x[$. Verallgemeinerter Mittelwertsatz auf $[a, x]$:

$$\exists \xi_x \in]a, x[: \underbrace{\frac{f(x) - f(a)}{g(x) - g(a)}}_{=\frac{f(x)}{g(x)}} = \underbrace{\frac{f'(\xi_x)}{g'(\xi_x)}}_{\to c \text{ für } x\downarrow a}$$

$\square$

6.26 Varianten: **1)** Typ $\left[\frac{\infty}{\infty}\right]$: $\lim\limits_{x\downarrow a} f(x) = \infty = \lim\limits_{x\downarrow a} g(x)$. Dann

$$\lim_{x\downarrow a} \frac{f(x)}{g(x)} = \lim_{x\downarrow a} \frac{f'(x)}{g'(x)}$$

2) Dasselbe für $x \uparrow b$ oder $x \to +\infty$ oder $x \to -\infty$.

3) Typ "$0 \cdot \infty$": $\lim f(x) = 0$, $\lim g(x) = \infty$: $\lim f(x) \cdot g(x) = \lim \dfrac{f(x)}{\frac{1}{g(x)}}$.

4) Typ "1^∞": $\lim f(x) = 1$, $\lim g(x) = \infty$: $f(x)^{g(x)} = \mathrm{e}^{g(x)\cdot \ln f(x)}$ geht mit vorigem Fall.

5) Alle Varianten gelten auch bei bestimmter Divergenz:

$$\lim \frac{f'}{g'} = +\infty\ (-\infty) \ \Rightarrow \ \lim \frac{f}{g} = +\infty\ (-\infty).$$

6.27 Beispiele: **1)** $\lim\limits_{x\to 0} \dfrac{1 - \cos x}{x^2} = \dfrac{1}{2}$.

2) $\lim\limits_{x\to\infty} \dfrac{\ln x}{x} = 0$.

3) $\lim\limits_{x\to\infty} x^{1/x} = 1$.

4) $\lim\limits_{x\downarrow 0} x \ln x = 0$.

5) $\lim\limits_{x\to-\infty} x^2\,\mathrm{e}^x = 0$.

6.6 Der Satz von Taylor

6.28 Vorbemerkung: Sei $f :]a,b[\to \mathbb{R}$ n-mal stetig differenzierbar auf $]a,b[$, $x_0 \in]a,b[$. Es gibt genau ein Polynom T_n n-ten Grades, so dass

$$T_n^{(k)}(x_0) \;=\; f^{(k)}(x_0) \qquad \text{für } 0 \le k \le n.$$

Für dieses Polynom gilt

$$T_n(x) \;=\; \sum_{j=0}^{n} \frac{f^{(j)}(x_0)}{j!}\,(x-x_0)^j.$$

T_n heißt das ***n*-te Taylorpolynom** für f, $R_n(x) := f(x) - T_n(x)$ das ***n*-te Restglied**.

6.29 Beispiel: $f : x \mapsto \mathrm{e}^x$, $x_0 = 0$: $f^{(k)}(0) = \mathrm{e}^0 = 1$ für $k \in \mathbb{N}_0$.

$$\Rightarrow\; T_n(x) = \sum_{j=0}^{n} \frac{1}{j!}\,x^j, \qquad R_n(x) = \mathrm{e}^x - \sum_{j=0}^{n} \frac{x^j}{j!}.$$

$x_0 = 1$: $f^{(k)}(1) = \mathrm{e} \;\Rightarrow\; T_n(x) = \displaystyle\sum_{j=0}^{n} \frac{\mathrm{e}}{j!}(x-1)^j$.

6.30 Satz (Taylor, Lagrange): Sei f $(n+1)$-mal stetig differenzierbar auf $]a,b[$, $x_0 \in]a,b[$ und $x \in]x_0,b[$. Dann existiert $\xi \in]x_0,x[$, so dass

$$f(x) \;=\; T_n(x) + \underbrace{\frac{f^{(n+1)}(\xi)}{(n+1)!}\,(x-x_0)^{n+1}}_{=R_n(x)}$$

(gilt genauso für $x \in]a,x_0[$, nur dann $\xi \in]x,x_0[$).

6.31 Bemerkungen: **1)** Für $n = 0$ ist dies wieder der Mittelwertsatz.

2) T_n approximiert f an einer Stelle x_0 bis zur n-ten Ableitung. Erst durch Diskussion des Restgliedes erkennt man, wie gut die Approximation durch T_n in einer Umgebung von x_0 ist.

6.32 Beispiele: **1)** $f : x \mapsto \mathrm{e}^x$, $x_0 = 0$, $x = 1$: e^1 soll bis auf Genauigkeit 10^{-5} durch $T_n(1)$ bestimmt werden.

$$|R_n(1)| \;=\; \left|\frac{\mathrm{e}^\xi}{(n+1)!}\,(1-0)^{n+1}\right| \overset{0<\xi<1}{\le} \frac{\mathrm{e}^1}{(n+1)!} \;\le\; \frac{3}{(n+1)!} \;\overset{!}{\le}\; 10^{-5}$$

$$\Leftrightarrow\; (n+1)! \;\ge\; 3\cdot 10^5 \;\Rightarrow\; n \ge 8 \quad (9! = 362880)$$

Das bedeutet: $\left|\mathrm{e} - T_8(1)\right| < 10^{-5}$. Mit Rechner:

$$\begin{array}{rclcrcl} T_6(1) & = & 2.7180556, & & R_6 & = & 2.262 \cdot 10^{-4} \\ T_7(1) & = & 2.7182540, & & R_7 & = & 2.78 \cdot 10^{-5} \\ T_8(1) & = & 2.7182788, & & R_8 & = & 3.0 \cdot 10^{-6} \\ \mathrm{e}^1 & = & 2.7182818\ldots & & & & \end{array}$$

2) $f : \mathbb{R} \to \mathbb{R} : x \mapsto \mathrm{e}^x,\ x_0 = 0$. Was passiert für $n \to \infty$?

$$|R_n(x)| = \left|\frac{\mathrm{e}^\xi}{(n+1)!}(x-0)^{n+1}\right| \leq \max\{\mathrm{e}^x, 1\} \frac{|x|^{n+1}}{(n+1)!} \to 0$$

für $n \to \infty$ bei jedem festen $x \in \mathbb{R}$.

$$\Rightarrow \mathrm{e}^x = \sum_{j=0}^{\infty} \frac{x^j}{j!} \qquad \text{für } x \in \mathbb{R}$$

(wie bereits bekannt).

3) $f(x) = \begin{cases} \mathrm{e}^{-1/x^2}, & x \neq 0, \\ 0, & x = 0. \end{cases}$, $x_0 = 0$.

Es gilt $f^{(n)}(0) = 0$ für $n \in \mathbb{N}_0$, also auch $T_n(x) = 0$ für $n \in \mathbb{N},\ x \in \mathbb{R}$. Diese Funktion f ist um $x_0 = 0$ nicht durch das Taylorpolynom approximierbar.

4) Die **Binomialreihe**: Für $r \in \mathbb{R} \setminus \mathbb{N}_0$ gilt

$$(1+x)^r = \sum_{j=0}^{\infty} \binom{r}{j} x^j \qquad \text{für } x \in]-1, 1[\qquad \left(\binom{r}{j} := \frac{r \cdot (r-1) \cdots (r-j+1)}{j!}\right).$$

6.7 Extrema: Hinreichende Bedingungen

6.33 Satz: Sei $D \subseteq \mathbb{R}$ offen, $f \in C^m(D \to \mathbb{R})$, $x_0 \in D$ und $f'(x_0) = \ldots = f^{m-1}(x_0) = 0$, $f^{(m)}(x_0) \neq 0$.

1) Ist m ungerade, so hat f in x_0 einen **Sattelpunkt**, d.h. Wendepunkt mit waagrechter Tangente.

2) Falls m gerade ist:

- **a)** Ist $f^{(m)}(x_0) > 0$, so hat f in x_0 ein striktes lokales Minimum.
- **b)** Ist $f^{(m)}(x_0) < 0$, so hat f in x_0 ein striktes lokales Maximum.

Beweis: Aus Satz von Taylor: $f(x) = f(x_0) + \sum_{j=1}^{m-1} 0 + \frac{f^{(m)}(\xi)}{m!}(x-x_0)^m$,

und $f^{(m)}(\xi)$ hat konstantes Vorzeichen für $|\xi - x_0| < \varepsilon$. $\square$

6.34 Beispiele: $y = x^6$, $y = x^7$.

6.8 Ableitung II

6.35 Definition: Sei $\gamma : [a,b] \to \mathbb{R}^n$ stetig. Die Menge

$$K \;:=\; \big\{\gamma(t) : t \in [a,b]\big\} \subseteq \mathbb{R}^n$$

heißt **Kurve** im $\mathbb{R}^n$, $(\gamma, [a,b])$ heißt **Parameterdarstellung** von K. Ist $\gamma(a) = \gamma(b)$, so heißt K **geschlossen**. Ist $\gamma\big|_{[a,b[}$ injektiv, so heißt K **Jordan-Kurve**.

6.36 Beispiele: **1)** $\gamma : [0, 2\pi] \to \mathbb{R}^2 : t \mapsto \begin{pmatrix} 2\cos t \\ \sin t \end{pmatrix}$.

2) $\gamma : [0, \pi] \to \mathbb{R}^2 : t \mapsto \begin{pmatrix} 2\cos 2t \\ -\sin 2t \end{pmatrix}$: Dieselbe Kurve wie in 1).

3) $\gamma : [0, 2] \to \mathbb{R}^3 : t \mapsto \begin{pmatrix} \cos(2\pi t) \\ \sin(2\pi t) \\ t \end{pmatrix}$ (Schraubenline).

4) $\gamma : [0, 2] \to \mathbb{R}^2 : t \mapsto \begin{pmatrix} t\cos(2\pi t) \\ t\sin(2\pi t) \end{pmatrix}$ (Archimedische Schneckenlinie).

6.37 Definition und Satz: Sei $f : \mathbb{R} \supseteq D \to \mathbb{R}^n$ und x_0 Häufungspunkt von D. Dann heißt f **differenzierbar in x_0**, falls

$$f'(x_0) \;:=\; \lim_{\substack{x \to x_0 \\ x \in D,\ x \neq x_0}} \frac{1}{x - x_0} \cdot \big(f(x) - f(x_0)\big)$$

existiert. Dies ist genau dann der Fall, wenn jede Koordinate f_j von f in x_0 differenzierbar ist $(j = 1, 2, \ldots, n)$.

Ist f in x_0 differenzierbar, so gilt

$$f'(x_0) \;=\; \begin{pmatrix} f_1'(x_0) \\ \vdots \\ f_n'(x_0) \end{pmatrix}$$

6.38 Beispiele: **1)** $f(x) = \begin{pmatrix} x^2 \\ e^{x^2 - x} \end{pmatrix} \;\Rightarrow\; f'(x) = \begin{pmatrix} 2x \\ (2x-1)\, e^{x^2 - x} \end{pmatrix}$

2) Seien $A \in \mathcal{M}_{n,n}$, $v \in \mathbb{R}^n$ und $g : \mathbb{R} \to \mathbb{R}^n : t \mapsto e^{t \cdot A}\, v$. Dann gilt $g'(t) = A \cdot e^{t \cdot A}\, v$.

3) Man kann die obige Definition der Ableitung auf $f : \mathbb{R} \supseteq D \to V$ erweitern, wobei V ein normierter Vektorraum ist. Mit $V = \mathcal{M}_{n,n}$, $A \in \mathcal{M}_{n,n}$ und $f : \mathbb{R} \to \mathcal{M}_{n,n} : t \mapsto e^{t \cdot A}$ folgt dann $f'(t) = A \cdot e^{t \cdot A}$.

6.39 Geometrische Bedeutung der Ableitung: Ist $\gamma : [a,b] \to \mathbb{R}^n$ in t_0 differenzierbar mit $\gamma'(t_0) \neq 0$, so ist $\gamma'(t_0)$ ein Tangentenvektor (Richtungsvektor der Tangente an K) im Punkt $\gamma'(t_0)$. $\|\gamma'(t_0)\|$ gibt die „Momentangeschwindigkeit“ an, mit der K durchlaufen wird.

6.40 Beispiele: **1)** $\gamma : [0, 2\pi] \to \mathbb{R}^2 : t \mapsto \begin{pmatrix} 2\cos t \\ \sin t \end{pmatrix}$:

$\gamma'(t_0) = \begin{pmatrix} -2\sin t_0 \\ \cos t_0 \end{pmatrix}$ ist Tangentenvektor im Punkt $(2\cos t_0, \sin t_0)$.

2) $\gamma : [0, \pi] \to \mathbb{R}^2 : t \mapsto \begin{pmatrix} 2\cos 2t \\ \sin 2t \end{pmatrix}$:

$\gamma'(\frac{t_0}{2}) = \begin{pmatrix} -4\sin t_0 \\ 2\cos t_0 \end{pmatrix}$ ist Tangentenvektor im Punkt $(2\cos t_0, \sin t_0)$. Die Ellipse wird mit doppelter Geschwindigkeit durchlaufen.

6.41 Achtung bei $\gamma'(t_0) = 0$: Sei z.B. $\gamma(t) := \begin{pmatrix} t^2 \\ 0 \end{pmatrix}$ für $-1 \leq t < 0$, $\gamma(t) := \begin{pmatrix} 0 \\ t^2 \end{pmatrix}$ für $0 \leq t \leq 1$. Dann gilt $\gamma \in C^1([a,b] \to \mathbb{R}^2)$, aber K hat eine Ecke in $(0,0)$.

6.42 Definition: Sei $\gamma : I = [a,b] \to \mathbb{R}$ stetig.

1) Ist γ in $t_0 \in I$ differenzierbar mit $\gamma'(t_0) \neq 0$, so heißt $T(t_0) := \dfrac{\gamma'(t_0)}{\|\gamma'(t_0)\|}$ der **Tangenteneinheitsvektor** in t_0.

2) γ heißt **regulär**, falls γ differenzierbar auf I ist mit $\gamma'(t) \neq 0$ auf I.

3) γ heißt **singulär in $t_0 \in I$**, falls $\gamma'(t_0) = 0$.

6.43 Bemerkung: Besitzt die Kurve K eine reguläre Parameterdarstellung $(\gamma, [a,b])$, so hat K keine Ecken.

7 Differentialrechnung im $\mathbb{R}^n$

7.1 Ableitung III

7.1 Der Graph von Funktionen mehrerer Veränderlicher: Ist $f : \mathbb{R}^2 \supseteq D \to \mathbb{R}$ und D offen, so ist der Graph von f

$$G = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ f(x_1, x_2) \end{pmatrix} \in \mathbb{R}^3 : \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in D \right\}$$

eine (gekrümmte) Fläche im $\mathbb{R}^3$. Veranschaulichung durch Höhenlinien ($f(x) = \text{konst}$) oder durch Schnitte mit Ebenen.

Allgemeiner: Ist $f : \mathbb{R}^n \supseteq D \to \mathbb{R}$ und D offen, so ist der Graph von f

$$G = \left\{ \begin{pmatrix} x \\ f(x) \end{pmatrix} \in \mathbb{R}^{n+1} : x \in D \right\}$$

eine „Hyperfläche“ im $\mathbb{R}^{n+1}$.

7.2 Beispiele: **1)** $f : \mathbb{R}^2 \to \mathbb{R} : \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto 4 + 5x_1 - 2x_2$:

Graph $G = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ 4 + 5x_1 - 2x_2 \end{pmatrix} : x_1, x_2 \in \mathbb{R} \right\}$ (Ebene).

2) Allgemeiner: $f : \mathbb{R}^n \to \mathbb{R} : x \mapsto a_1x_1 + a_2x_2 + \ldots a_nx_n$: Graph ist Hyperebene im $\mathbb{R}^{n+1}$.

3) $f : D = \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} : x_1^2 + x_2^2 \leq 1 \right\} \to \mathbb{R} : (x_1, x_2) \mapsto \sqrt{1 - x_1^2 - x_2^2}$:

Graph von f: $G = \left\{ \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} : x_1^2 + x_2^2 < 1 \ \wedge \ x_3 = \sqrt{1 - x_1^2 - x_2^2} \right\}$ (obere Halbkugel).

4) $f : \mathbb{R}^2 \to \mathbb{R} : \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto x_1^2 + 4x_2^2$ (elliptisches Paraboloid).

5) $f : \mathbb{R}^2 \to \mathbb{R} : \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto x_1x_2$ (Der Graph ist eine Sattelfläche).

7.3 Äquivalente Definition der Ableitung: Seien $f : \mathbb{R} \supseteq D \to \mathbb{R}$ und x_0 Häufungspunkt von D. Dann sind äquivalent:

(i) f ist differenzierbar in x_0.

(ii) Es existiert eine Zahl $a \in \mathbb{R}$, so dass

$$\lim_{\substack{x \to x_0 \\ x \in D,\ x \neq x_0}} \frac{1}{|x - x_0|} \big(f(x) - f(x_0) - a(x - x_0) \big) = 0$$

Sind diese Bedingungen erfüllt, so gilt $a = f'(x_0)$.

7.4 Definition: Sei $f : \mathbb{R}^n \supseteq D \to \mathbb{R}$, D offen und $x_0 \in D$. Dann heißt f **differenzierbar** in x_0, falls eine $1 \times n$-Matrix $A = (a_1, \ldots, a_n)$ $(a_j \in \mathbb{R})$ existiert mit

$$\lim_{\substack{x \to x_0 \\ x \in D,\ x \neq x_0}} \frac{1}{\|x - x_0\|}\big(f(x) - f(x_0) - A \cdot (x - x_0)\big) \;=\; 0$$

oder äquivalent:

$$f(x) \;=\; f(x_0) + A \cdot (x - x_0) + \|x - x_0\|\, r(x) \qquad \text{mit } \lim_{x \to x_0} r(x) = 0.$$

Die Matrix A ist durch diese Bedingung eindeutig bestimmt und heißt **Ableitung** von f in x_0. Wir schreiben $f'(x_0) := A = (a_1, \ldots, a_n)$. Der affine Raum

$$T \;:=\; \left\{ \begin{pmatrix} x \\ x_{n+1} \end{pmatrix} : x \in \mathbb{R}^n \wedge x_{n+1} = f(x_0) + f'(x_0) \cdot (x - x_0) \right\}$$

heißt **Tangentialebene** (falls $n = 2$) bzw. **Tangentialhyperebene** an den Graphen von f im Punkt $\begin{pmatrix} x_0 \\ f(x_0) \end{pmatrix}$.

7.5 Die Idee:: Die Funktion $g(x) := f(x_0) + f'(x_0)(x - x_0)$ ist die optimale „lineare“ Approximation von f bei x_0.

7.6 Beispiele: **1)** $f : \mathbb{R}^2 \to \mathbb{R} : \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto x_1\, x_2$.

2) $f : \left\{ \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} : x_1^2 + x_2^2 < 1 \right\} \to \mathbb{R} : \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \sqrt{1 - x_1^2 - x_2^2}$.

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

7.8 Wichtige Idee: Zur Untersuchung oder Veranschaulichung des Graphen in der Umgebung eines Punktes $\big(x_0, f(x_0)\big)$ kann man das Verhalten von f längs der Geraden $x = x_0 + t \cdot v$ $(t \in \mathbb{R})$ mit festem Richtungsvektor $v \in \mathbb{R}^n$ betrachten. Anders ausgedrückt: Man schneidet G mit geeigneten Ebenen:

$$\begin{aligned} E \;&:=\; \big\{(x_0 + t \cdot v, s) \in \mathbb{R}^{n+1} : t, s \in \mathbb{R}\big\} \\ \Rightarrow\; E \cap G \;&=\; \big\{\big(x_0 + t \cdot v, f(x_0 + t \cdot v)\big) : t \in D'\big\} \qquad (D' \subseteq \mathbb{R} \text{ geeignet}). \end{aligned}$$

Diese Schnittkurve $E \cap G$ kann man als Kurve im $\mathbb{R}^2$ auffassen:

$$K \;:=\; \big\{\big(t, f(x_0 + t \cdot v)\big) : t \in D'\big\}.$$

Vorteil: Man untersucht die Funktion $t \mapsto f(x_0 + t \cdot v)$, die nur von **einer reellen** Variablen abhängt.

Noch spezieller: Wähle als v die Koordinatenvektoren $e_1, \ldots, e_n$:

$$g_j(t) \;:=\; \big(t, f(x_0 + t \cdot e_j)\big).$$

und betrachte deren Tangentialvektoren im Punkt $t = 0$ (falls existent):

$$g_j'(0) \;=\; \Big(1, \lim_{h\to 0} \frac{f(x_0 + h \cdot e_j) - f(x_0)}{h}\Big).$$

7.9 Partielle Ableitungen: Sei $f : \mathbb{R}^n \supseteq D \to \mathbb{R}$, D offen und f differenzierbar in $x_0 \in D$. Dann existiert für jedes $j \in \{1, \ldots, n\}$ der Grenzwert

$$\lim_{\substack{h \to 0 \\ h \neq 0}} \frac{f(x_0 + h \cdot e_j) - f(x_0)}{h} \;=:\; \frac{\partial f}{\partial x_j}(x_0) \;=:\; \partial_{x_j} f(x_0) \;=:\; \partial_j f(x_0).$$

Dieser Grenzwert heißt **partielle Ableitung** von f nach x_j. Es gilt

$$f'(x_0) \;=\; \big(\partial_{x_1} f(x_0),\ \partial_{x_2} f(x_0),\ \ldots,\ \partial_{x_n} f(x_0)\big).$$

Der Vektor

$$\nabla f(x_0) \;:=\; \begin{pmatrix} \partial_1 f(x_0) \\ \vdots \\ \partial_n f(x_0) \end{pmatrix}$$

heißt **Gradient** von f.

Beweis: $$0 \;=\; \lim_{\substack{x \to x_0 \\ x \neq x_0}} \frac{1}{\|x - x_0\|}\big(f(x) - f(x_0) - f'(x_0)(x - x_0)\big)$$

$$\Rightarrow\ 0 \;=\; \lim_{\substack{h \to 0 \\ h \neq 0}} \underbrace{\frac{1}{\|x_0 + h \cdot e_j - x_0\|}}_{=\frac{1}{|h|}} \Big(f(x_0 + h \cdot e_j) - f(x_0) - \underbrace{f'(x_0)(x_0 + h \cdot e_j - x_0)}_{=h\left(f'(x_0)\right)_j}\Big)$$

$$=\; \lim_{\substack{h \to 0 \\ h \neq 0}} \operatorname{sign}(h) \left(\frac{f(x_0 + h \cdot e_j) - f(x_0)}{h} - \big(f'(x_0)\big)_j\right)$$

$$\Rightarrow\ \partial_{x_j} f(x_0) \;=\; \lim_{\substack{h \to 0 \\ h \neq 0}} \frac{f(x_0 + h \cdot e_j) - f(x_0)}{h} \quad (j = 1, \ldots, n) \text{ existieren und}$$

$$f'(x_0) \;=\; \big(\partial_{x_1} f(x_0),\ \partial_{x_2} f(x_0),\ \ldots,\ \partial_{x_n} f(x_0)\big).$$

□

7.10 Bemerkung: $f'(x_0) \cdot (x - x_0) \;=\; \langle \nabla f(x_0), x - x_0 \rangle$,
$\nabla f(x_0) = f'(x_0)^T$ (transponierte Matrix).

7.11 Geometrische Bedeutung der partiellen Ableitung: Schneide den Graphen von f mit der Ebene

$$E := \big\{(x_0 + t \cdot e_j, s) \in \mathbb{R}^{n+1} : t, s \in \mathbb{R}\big\}.$$

Die Schnittkurve ist der Graph der Funktion

$$t \mapsto f(x_0 + t \cdot e_j).$$

Für die Ableitung dieser Funktion im Punkt $t = 0$ gilt

$$\frac{\mathrm{d}}{\mathrm{d}t} f(x_0 + t \cdot e_j)\big|_{t=0} = \partial_{x_j} f(x_0).$$

7.12 Geometrische Bedeutung des Gradienten: Für festes $v \in \mathbb{R}^n$ mit $\|v\| = 1$ schneide den Graph von f mit der Ebene

$$E := \big\{(x_0 + t \cdot v, s) : t, s \in \mathbb{R}\big\}.$$

Die Steigung der Schnittfunktion $t \mapsto f(x_0 + t \cdot v)$ im Punkt $t = 0$ ergibt sich zu

$$\frac{\mathrm{d}}{\mathrm{d}t} f(x_0 + t \cdot v)\big|_{t=0} = f'(x_0)\, v = \langle \nabla f(x_0), v \rangle .$$

Die Steigung (in Abhängigkeit von v) ist am größten, wenn v parallel zu $\nabla f(x_0)$ ist. Dann hat die Steigung den Wert $\|\nabla f(x_0)\|$.

Also: Der Gradientenvektor zeigt in die Richtung der größten Zunahme von $f(x)$.

7.13 Ein Kriterium für Differenzierbarkeit: Sei $f : \mathbb{R}^n \supseteq D \to \mathbb{R}$ stetig, D offen, alle partiellen Ableitungen $\partial_j f$ existieren in D und seien stetig in $x_0 \in D$. Dann ist f in x_0 differenzierbar, und es gilt $f'(x_0) = \big(\partial_{x_1} f(x_0), \ldots, \partial_{x_n} f(x_0)\big)$. Also:

Für Differenzierbarkeit in x_0 ist notwendig:	f ist stetig und $\nabla f(x_0)$ existiert.
Für Differenzierbarkeit in x_0 ist hinreichend:	f ist stetig, ∇f existiert in D, und ∇f ist stetig in x_0.

Für $k \in \mathbb{N}_0$ ist $C^k(D \to \mathbb{R})$ der Raum aller Funktionen, deren partielle Ableitungen bis zur k-ten Ordnung auf D existieren und stetig sind.

7.14 Ergänzung: Die Existenz aller partiellen Ableitungen in einem Punkt x_0 garantiert nicht die Differenzierbarkeit in x_0. Sei z.B.

$$f : \mathbb{R}^2 \to \mathbb{R} \quad \text{mit} \quad f\begin{pmatrix} x \\ y \end{pmatrix} := \begin{cases} 0 & \text{falls } (x,y) = (0,0) \ \vee \ y \neq x \\ 1 & \text{falls } y = x \ \wedge \ x \neq 0 \end{cases}$$

Dann gilt $\partial_x f(0,0) = 0$, $\partial_y f(0,0) = 0$, aber die Funktion f ist in $(0,0)$ unstetig, also sicher nicht differenzierbar.

7.2 Extrema

Sei $f : \mathbb{R}^n \supseteq D \to \mathbb{R}$, D offen und f differenzierbar in $x_0 \in D$. Wir wollen feststellen, ob f in x_0 ein Extremum besitzt. Wenn f in x_0 ein Maximum besitzt, dann haben auch alle Funktionen

$$g_v : t \mapsto f(x_0 + t \cdot v)$$

mit beliebigen Vektoren $v \in \mathbb{R}^n$ für $t = 0$ ein Maximum (genauso mit Minimum). Wir wissen:

- Notwendig für ein lokales Extremum ist $g_v'(0) = 0$.
- Hinreichend für ein striktes lokales Extremum ist $g_v'(0) = 0$ und $g_v''(0) > 0$ oder $g_v''(0) < 0$.

Weiter ist bekannt:

$$g_v'(t) \;=\; f'(x_0 + t \cdot v)\, v \;=\; \sum_{j=1}^{n} \partial_{x_j} f(x_0 + t \cdot v) v_j.$$

Daraus folgt sofort:

7.15 Notwendige Bedingung: Sei $f : \mathbb{R}^n \supseteq D \to \mathbb{R}$, D offen und f differenzierbar in $x_0 \in D$. Dann gilt:

$$f \text{ hat in } x_0 \text{ ein Extremum} \;\Rightarrow\; f'(x_0) = \big(\partial_{x_1} f(x_0),\ \ldots,\ \partial_{x_n} f(x_0)\big) = 0.$$

Nun nehmen wir an, dass die partiellen Ableitungen $\mathbb{R}^n \supseteq D \ni x \mapsto \partial_{x_j} f$ wieder differenzierbar sind. Dann gilt

$$\begin{aligned}
\frac{\mathrm{d}\partial_{x_j} f(x_0 + t \cdot v)}{\mathrm{d}t} &= \sum_{k=1}^{n} \partial_{x_k}(\partial_{x_j} f)(x_0 + t \cdot v)\, v_k \\
\Rightarrow\; g_v''(t) &= \sum_{j,k=1}^{n} \partial_{x_k}(\partial_{x_j} f)(x_0 + t \cdot v)\, v_k\, v_j \\
\Rightarrow\; g_v''(0) &= \sum_{j,k=1}^{n} \partial_{x_k}(\partial_{x_j} f)(x_0)\, v_k\, v_j \\
&= \left\langle \underbrace{\begin{pmatrix} \partial_{x_1}\partial_{x_1} f(x_0) & \partial_{x_1}\partial_{x_2} f(x_0) & \ldots & \partial_{x_1}\partial_{x_n} f(x_0) \\ \vdots & & & \vdots \\ \partial_{x_n}\partial_{x_1} f(x_0) & \partial_{x_n}\partial_{x_2} f(x_0) & \ldots & \partial_{x_n}\partial_{x_n} f(x_0) \end{pmatrix}}_{=: H_f(x_0) \text{ (Hessematrix)}} \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix}, \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \right\rangle
\end{aligned}$$

7.16 Hinreichende Bedingung: Sei $D \subseteq \mathbb{R}^n$ offen, $f \in C^2(D \to \mathbb{R})$, $x_0 \in D$ und $H_f(x_0)$ die Hessematrix von f in x_0. Dann gelten:

1) Falls $f'(x_0) = 0$ und $H_f(x_0)$ positiv definit ist, besitzt f in x_0 ein striktes lokales Minimum.

2) Falls $f'(x_0) = 0$ und $H_f(x_0)$ negativ definit ist (d.h. $-H_f(x_0)$ ist positiv definit), besitzt f in x_0 ein striktes lokales Maximum.

Beweis: Wenn 1) erfüllt ist, zeigt die obige Rechnung, dass alle Funktionen g_v in $t = 0$ ein Minimum besitzen. Dies veranschaulicht, dass die Bedingung 1) für das Vorliegen eines Minimums sinnvoll ist. Den strikten Beweis kann man erst mit dem Satz von Taylor (Satz 7.32) führen.

$\square$

Wir werden später sehen: $H_f(x_0)$ ist symmetrisch. Das bedeutet, dass $H_f(x_0)$ diagonalisierbar ist. Deshalb kann man den letzten Satz auch anders formulieren:

7.17 Satz: Sei $D \subseteq \mathbb{R}^n$ offen und $f \in C^2(D \to \mathbb{R})$, $x_0 \in D$ mit $f'(x_0) = 0$. Dann gilt:

1) Hat die Hesse-Matrix $H_f(x_0)$ nur positive Eigenwerte, so hat f in x_0 ein striktes lokales Minimum.

2) Hat die Hesse-Matrix $H_f(x_0)$ nur negative Eigenwerte, so hat f in x_0 ein striktes lokales Maximum.

3) Ohne Beweis: Hat die Hesse-Matrix $H_f(x_0)$ positive und negative Eigenwerte, so hat f in x_0 einen Sattelpunkt: In jeder Umgebung von x_0 gibt es Punkte x mit $f(x) > f(x_0)$ und Punkte mit $f(x) < f(x_0)$.

7.18 Beispiele: **1)** $f\begin{pmatrix} x \\ y \end{pmatrix} = (x^2 - y^2)\ln x$ in $D := \{(x, y) \in \mathbb{R}^2 : x > 0\}$.

2) $f\begin{pmatrix} x \\ y \end{pmatrix} = 4xy - x^3y - xy^3$

3) $f : D = \left\{\begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 : x^2 + y^2 \le 1\right\} \to \mathbb{R} : \begin{pmatrix} x \\ y \end{pmatrix} \mapsto x\,y$:
Einziger kritischer Punkt: $(x, y) = (0, 0)$ mit $f(0, 0) = 0$. Hier liegt ein Sattelpunkt vor. Also hat f im Inneren von D kein Extremum.

Da D beschränkt und abgeschlossen ist, nimmt f auf D das Maximum und Minimum an. Also muss f das Minimum und Maximum am Rand von D annehmen.

Untersuche f am Rand: $\Rightarrow$ Maximum $f\begin{pmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} = \frac{1}{2}$, Minimum $f\begin{pmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{pmatrix} = -\frac{1}{2}$

7.3 Ableitung IV

7.19 Definition: Sei $f : \mathbb{R}^n \supseteq D \to \mathbb{R}^m$, D offen und $x_0 \in D$. Dann heißt f **differenzierbar** in x_0, falls eine $m \times n$-Matrix A existiert mit

$$\lim_{\substack{x \to x_0 \\ x \in D,\ x \neq x_0}} \frac{1}{\|x - x_0\|}\big(f(x) - f(x_0) - A \cdot (x - x_0)\big) \;=\; 0$$

oder äquivalent:

$$f(x) \;=\; f(x_0) + A \cdot (x - x_0) + \|x - x_0\|\, r(x) \qquad \text{mit } \lim_{x \to x_0} r(x) = 0.$$

Die Matrix A ist durch diese Bedingung eindeutig bestimmt und heißt **Ableitung** von f in x_0. Wir schreiben $f'(x_0) := A$.

7.20 Bemerkungen: **1)** Im Fall $m = 1$ ist dies genau die Definition 7.4.

2) Ist f in x_0 differenzierbar, so gilt

$$f'(x_0) \;=\; \begin{pmatrix} \partial_{x_1} f_1 & \partial_{x_2} f_1 & \dots & \partial_{x_n} f_1 \\ \partial_{x_1} f_2 & \partial_{x_2} f_2 & \dots & \partial_{x_n} f_2 \\ \vdots & & & \vdots \\ \partial_{x_1} f_m & \partial_{x_2} f_m & \dots & \partial_{x_n} f_m \end{pmatrix}$$

Die $m \times n$-Matrix $J_f(y) = \big(\partial_i f_j(y)\big)_{i,j}$ heißt auch **Jacobi-Matrix** von f.

7.21 Beispiel: $f(x, y) = \begin{pmatrix} x\,\mathrm{e}^y \\ x\,y \end{pmatrix} \;\Rightarrow\; J_f(x, y) = f'(x, y) = \begin{pmatrix} \mathrm{e}^y & x\,\mathrm{e}^y \\ y & x \end{pmatrix}.$

7.22 Ableitung des Gradienten: Sei $f : \mathbb{R}^n \supseteq D \to \mathbb{R}$ differenzierbar in D, D offen. Dann gilt $\nabla f : D \to \mathbb{R}^n$. Ist ∇f in $x_0 \in D$ differenzierbar, so ist die Jacobi-Matrix von ∇f:

$$(\nabla f)'(x_0) \;=\; \begin{pmatrix} \partial_1^2 f(x_0) & \partial_2 \partial_1 f(x_0) & \dots & \partial_n \partial_1 f(x_0) \\ \vdots & & & \vdots \\ \partial_1 \partial_n f(x_0) & \partial_2 \partial_n f(x_0) & \dots & \partial_n^2 f(x_0) \end{pmatrix}$$

die **Hesse-Matrix** von f.

7.23 Beispiel: $f(x, y) = x\,\mathrm{e}^y \;\Rightarrow\; \nabla f(x, y) = \begin{pmatrix} \mathrm{e}^y \\ x\,\mathrm{e}^y \end{pmatrix}, \quad (\nabla f)'(x, y) = \begin{pmatrix} 0 & \mathrm{e}^y \\ \mathrm{e}^y & x\,\mathrm{e}^y \end{pmatrix}.$

7.4 Mittelwertsätze und mehr

7.24 Mittelwertsatz bei mehreren Variablen: Seien $D \subseteq \mathbb{R}^n$ offen, $f \in C^1(D \to \mathbb{R})$ und $x_0, y_0 \in D$, so dass die Verbindungsstrecke in D liegt: $\{x_0 + t \cdot (y_0 - x_0) : 0 \leq t \leq 1\} \subset D$. Dann

$$\exists \tau \in]0,1[: f(y_0) - f(x_0) \ = \ \big\langle \nabla f\big(x_0 + \tau \cdot (y_0 - x_0)\big), y_0 - x_0 \big\rangle .$$

Beweis: Setze $g(t) := f\big(x_0 + t \cdot (y_0 - x_0)\big)$ für $0 \leq t \leq 1$. Da f differenzierbar ist, gilt

$$g'(t) \ = \ f'\big(x_0 + t \cdot (y_0 - x_0)\big)(y_0 - x_0) \ = \ \big\langle \nabla f\big(x_0 + t \cdot (y_0 - x_0)\big), y_0 - x_0 \big\rangle .$$

Der Mittelwertsatz 6.19 liefert

$$f(y_0) - f(x_0) \ = \ g(1) - g(0) \ = \ g'(\tau) \cdot 1 \ = \ \big\langle \nabla f\big(x_0 + \tau \cdot (y_0 - x_0)\big), y_0 - x_0 \big\rangle$$

□

7.25 Satz von Schwarz: Für $f \in C^2(D \to \mathbb{R})$ gilt $\partial_j \partial_k f = \partial_k \partial_j f$ in D. Insbesondere ist die Hesse-Matrix von f symmetrisch.

Entsprechend: Sei $f \in C^m(D \to \mathbb{R})$, $\alpha_1, \alpha_2 \in \mathbb{N}$ mit $\alpha_1 + \alpha_2 \leq m$. Dann gilt

$$\partial_j^{\alpha_1} \partial_k^{\alpha_2} f \ = \partial_k^{\alpha_2} \partial_j^{\alpha_1} f \qquad \text{in } D.$$

7.26 Fehlerfortpflanzung: Bei Berechnungen mit Maschinenzahlen treten im Allgemeinen mindestens Rundungsfehler auf, d.h. die Ergebnisse weichen vom exakten Wert ab. Führt man mit einem fehlerbehafteten Wert eine neue Rechnung durch, so wird sich die ursprüngliche Abweichung auf das Ergebnis auswirken. Dies nennt man **Fehlerfortpflanzung**.

Sei nun $x \in \mathbb{R}^n$ der exakte Wert und $x + \Delta x \in \mathbb{R}^n$ der fehlerbehaftete Wert. Wir schätzen für eine Funktion $f \in C^1(\mathbb{R}^n \to \mathbb{R})$ ab, wie groß die Abweichung $f(x + \Delta x) - f(x)$ ist. Mit dem Mittelwertsatz 7.24:

$$f(x + \Delta x) - f(x) \ = \ \langle \nabla f(x + \tau \Delta x), \Delta x \rangle \approx \langle \nabla f(x), \Delta x \rangle .$$

Z.B. folgt für die Funktion $f(x,y) = xy$ bei $xy \neq 0$

$$f(x + \Delta x, y + \Delta y) - f(x,y) \ \approx \ y\Delta x + x \Delta y$$

und

$$\frac{f(x + \Delta x, y + \Delta y) - f(x,y)}{f(x,y)} \ \approx \ \frac{\Delta x}{x} + \frac{\Delta y}{y}.$$

bzw. für den relativen Fehler

$$\left| \frac{f(x + \Delta x, y + \Delta y) - f(x,y)}{f(x,y)} \right| \ \lessapprox \ \left| \frac{\Delta x}{x} \right| + \left| \frac{\Delta y}{y} \right|.$$

Also: Bei Produktbildung addieren sich die relativen Fehler

7.27 Kettenregel: Sei $f : \mathbb{R}^n \supseteq D \to \mathbb{R}^m,\ g : \mathbb{R}^m \supseteq D' \to \mathbb{R}^k$ mit $f(D) \subseteq D'$. Ist f in $x_0 \in D$ und g in $f(x_0)$ differenzierbar, so ist $h := g \circ f$ in x_0 differenzierbar mit

$$h'(x_0) \;=\; g'\big(f(x_0)\big) \cdot f'(x_0).$$

7.28 Beispiele: **1)** $f : \mathbb{R} \to \mathbb{R}^2 : x \mapsto \begin{pmatrix} x^2 \\ \mathrm{e}^x \end{pmatrix} : f'(x) = \begin{pmatrix} 2x \\ \mathrm{e}^x \end{pmatrix}$

$$g : \mathbb{R}^2 \to \mathbb{R}^2 : y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \mapsto \begin{pmatrix} y_1\, y_2 \\ y_1 - y_2 \end{pmatrix} : g'(y) = \begin{pmatrix} y_2 & y_1 \\ 1 & -1 \end{pmatrix}$$

$$\text{Kettenregel: } (g \circ f)'(x) \;=\; \begin{pmatrix} \mathrm{e}^x & x^2 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 2x \\ \mathrm{e}^x \end{pmatrix} \;=\; \begin{pmatrix} \mathrm{e}^x(x^2 + 2x) \\ 2x - \mathrm{e}^x \end{pmatrix}$$

$$\text{Direkt: } g \circ f(x) = \begin{pmatrix} x^2\, \mathrm{e}^x \\ x^2 - \mathrm{e}^x \end{pmatrix} \;\Rightarrow\; (g \circ f)'(x) = \begin{pmatrix} \mathrm{e}^x(x^2 + 2x) \\ 2x - \mathrm{e}^x \end{pmatrix}$$

2) $g(x_1, x_2, x_3) = \begin{pmatrix} x_1\, x_2\, x_3 \\ x_1^2 + x_2^2 + x_3^2 \end{pmatrix},\ f(y_1, y_2) = \mathrm{e}^{y_1^2\, y_2}.$

3) Wichtiges Beispiel: $g : \mathbb{R} \to \mathbb{R}^n,\ f : \mathbb{R}^n \to \mathbb{R}$.

7.29 Multiindizes: Für partielle Ableitungen höherer Ordnung ist folgende Notation geschickt:

$$\nabla^\alpha f \;:=\; \partial_{x_1}^{\alpha_1} \cdots \partial_{x_n}^{\alpha_n} f \qquad \text{für } \alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{N}_0^n$$

(∇ sprich „Nabla“). Zum Rechnen mit sogenannten **Multiindizes** $\alpha, \beta \in \mathbb{N}_0^n$ vereinbart man:

$$\begin{aligned} |\alpha| \;&:=\; \alpha_1 + \ldots + \alpha_n \\ \alpha \le \beta \;&\Leftrightarrow\; \alpha_1 \le \beta_1 \;\wedge\; \ldots \;\wedge\; \alpha_n \le \beta_n \\ \alpha! \;&:=\; (\alpha_1!) \cdots (\alpha_n!) \\ \binom{\alpha}{\beta} \;&:=\; \frac{\alpha!}{\beta!\,(\alpha - \beta)!} \;=\; \binom{\alpha_1}{\beta_1} \cdot \binom{\alpha_2}{\beta_2} \cdots \binom{\alpha_n}{\beta_n} \end{aligned}$$

7.30 Leibniz-Formel: Seien $D \subseteq \mathbb{R}^n$ offen und $f, g \in C^m(D \to \mathbb{R})$. Für $\alpha \in \mathbb{N}_0^n$ mit $|\alpha| \le m$ gilt:

$$\nabla^\alpha (g \cdot f) \;=\; \sum_{\beta \in \mathbb{N}_0^n,\ \beta \le \alpha} \binom{\alpha}{\beta} \big(\nabla^\beta f\big) \cdot \big(\nabla^{\alpha - \beta} g\big) \qquad \text{in } D.$$

7.31 Ableitungen längs einer Geraden: Für $f : \mathbb{R}^n \supseteq D \to \mathbb{R}$, $x_0 \in D$, $v \in \mathbb{R}^n$ und $\{x_0 + t \cdot v : 0 \le t \le 1\} \subset D$ setze $g(t) := f(x_0 + t \cdot v)$. Falls $f \in C^m(D \to \mathbb{R})$, so ist g m-Mal differenzierbar mit

$$g^{(m)}(t) \;=\; \sum_{|\alpha|=m} \frac{m!}{\alpha!} \left(\nabla^\alpha f\right)(x_0 + t \cdot v) \cdot v^\alpha,$$

wobei $v^\alpha := v_1^{\alpha_1} \cdots v_n^{\alpha_n}$.

Beweis:

$$\begin{aligned}
g'(t) &= \sum_{j_1=1}^{n} \partial_{j_1} f(x_0 + t \cdot v)\, v_{j_1} \qquad \left(= \sum_{|\alpha|=1} \nabla^\alpha f(x_0 + t \cdot v)\, v^\alpha \right) \\
g''(t) &= \sum_{j_2=1}^{n} \sum_{j_1=1}^{n} \partial_{j_2} \partial_{j_1} f(x_0 + t \cdot v)\, v_{j_1}\, v_{j_2} \\
&\;\;\vdots \\
g^{(m)}(t) &= \sum_{j_m=1}^{n} \sum_{j_{m-1}=1}^{n} \cdots \sum_{j_1=1}^{n} \partial_{j_m} \cdots \partial_{j_1} f(x_0 + t \cdot v)\, v_{j_1}\, v_{j_2} \cdots v_{j_m}.
\end{aligned}$$

In dieser m-fachen Summe kommen genau alle $\nabla^\alpha f$ mit $|\alpha| = m$ vor, manche mehrfach.

Wie oft kommt ein $\alpha \in \mathbb{N}_0^n$ mit $|\alpha| = m$ vor? Kombinatorik: Verteile $m = \alpha_1 + \ldots + \alpha_n$ Einträge auf die Plätze $j_1, \ldots, j_m$: Das sind $m!$ Möglichkeiten. Davon sind aber α_1 Möglichkeiten gleich, $\alpha_2, \ldots, \alpha_n$ Möglichkeiten gleich.

$$\Rightarrow \; g^{(m)}(t) \;=\; \sum_{|\alpha|=m} \frac{m!}{\alpha_1! \cdots \alpha_n!} \nabla^\alpha f(x_0 + t \cdot v).$$

□

7.32 Satz von Taylor II: Sei $D \subseteq \mathbb{R}^n$ offen, $f \in C^{m+1}(D \to \mathbb{R})$, $x, x_0 \in D$ so dass die Verbindungsstrecke ganz in D liegt: $\{x_0 + t \cdot (x - x_0) : 0 \le t \le 1\} \subset D$. Dann existiert ein $\tau \in \,]0,1[$, so dass

$$f(x) \;=\; \sum_{k=0}^{m} \sum_{|\alpha|=k} \frac{1}{\alpha!} \big(\nabla^\alpha f\big)(x_0) \cdot (x - x_0)^\alpha + \sum_{|\alpha|=m+1} \frac{1}{\alpha!} \big(\nabla^\alpha f\big)\big(x_0 + \tau \cdot (x - x_0)\big) \cdot (x - x_0)^\alpha.$$

Beweis: Wende den Satz von Taylor 6.30 auf die Funktion $g(t) := f\big(x_0 + t \cdot (x - x_0)\big)$ an, verwende den letzten Satz für $g^{(m)}(t)$.

□

7.33 Beispiel: $f(x, y) = \mathrm{e}^{x^2 + xy + y^2}$ bei $(x_0, y_0) = (0, 0)$.

8 Integration

8.1 Treppen- und Regelfunktionen

Gegeben: Funktion $f : [a,b] \to \mathbb{R}$.
Gesucht: Fläche zwischen $y = f(x)$ und x-Achse.

Typisch Mathematiker:
Existiert diese Fläche überhaupt?
Idee: Alles auf Rechteckflächen aufbauen.

8.1 Definition: **1)** Für $I \subseteq \mathbb{R}$, $I \neq \emptyset$ heißt

$$\chi_I : \mathbb{R} \to \mathbb{R} : x \mapsto \begin{cases} 1 & \text{falls } x \in I, \\ 0 & \text{sonst} \end{cases}$$

charakteristische Funktion von I (χ sprich „chi“).

2) $f : [a,b] \to \mathbb{R}$ heißt **Treppenfunktion** auf $[a,b]$, falls es $a = x_0 < x_1 < \ldots < x_n = b$ gibt, so dass f jeweils auf $]x_{i-1}, x_i[$ konstant ist ($1 \leq i \leq n$).

3) Zwei Treppenfunktionen f, g heißen **gleich fast überall** ($f = g$ f.ü.), falls $f(x) = g(x)$ für $x \in [a,b]$ mit höchstens endlich vielen Ausnahmen.

Die Relation „gleich fast überall“ ist eine Äquivalenzrelation auf der Menge alle Funktionen auf $[a,b]$.

8.2 Beispiel:

$$f(x) = \begin{cases} 3, & 0 \leq x < 1 \\ 1, & x = 1 \\ 4, & 1 < x \leq 2 \\ 2, & 2 < x \leq 3 \end{cases}$$

$$\Rightarrow \; f = 3 \cdot \chi_{[0,1]} + 4 \cdot \chi_{[1,2]} + 2 \cdot \chi_{[2,3]} \text{ f.ü.}$$

8.3 Satz: **1)** $f : [a,b] \to \mathbb{R}$ ist Treppenfunktion

$\Leftrightarrow$ Es gibt Intervalle $I_k \subseteq [a,b]$ und Konstanten $c_k \in \mathbb{R}$ mit $f = \sum\limits_{k=1}^{n} c_k \, \chi_{I_k}$ f.ü.

2) Die Menge der Treppenfunktionen auf $[a,b]$ mit der üblichen Addition und Multiplikation mit Konstanten bildet einen Vektorraum. Außerdem sind Produkt und Betrag von Treppenfunktionen wieder Treppenfunktionen.

8.4 Definition: Ist f Treppenfunktion auf $[a,b]$ und $f = \sum_{k=1}^{n} c_k \, \chi_{]x_{k-1},x_k[}$ f.ü., so heißt

$$\int_a^b f(x)\,\mathrm{d}x \;:=\; \sum_{k=1}^{n} c_k(x_k - x_{k-1})$$

das (bestimmte) **Integral** von f. Schreibe auch $\displaystyle\int_a^b f$.

Insbesondere: $f = g$ f.ü. $\Rightarrow \displaystyle\int_a^b f = \int_a^b g$.

Falls $a \le c \le d \le b$ setze $\displaystyle\int_c^d f := \int_a^b f \cdot \chi_{[c,d]}$.

8.5 Definition: $f : [a,b] \to \mathbb{R}$ heißt **Regelfunktion**, falls es eine Folge (t_n) von Treppenfunktionen auf $[a,b]$ gibt mit $f = \lim t_n$ gleichmäßig auf $[a,b]$. Insbesondere ist jede Regelfunktion beschränkt.

Beweis der Beschränktheit: Wähle $\varepsilon = 1$. Dann:

$$\exists N_1 \in \mathbb{N} \quad \forall n > N_1 \quad \forall x \in [a,b] : |f(x) - t_n(x)| < 1.$$

t_{N_1+1} ist Treppenfunktion $\Rightarrow \max\limits_{a \le x \le b} |t_{N_1+1}(x)|$ existiert.

$$\begin{aligned} \Rightarrow \; |f(x)| \;&\le\; |f(x) - t_{N_1+1}(x)| + |t_{N_1+1}(x)| \;<\; 1 + \max_{a \le x \le b} |t_{N_1+1}(x)| \;=:\; M \\ \Rightarrow \; |f(x)| \;&\le\; M \qquad \text{für } x \in [a,b]. \end{aligned}$$

□

8.6 Welche Funktionen sind Regelfunktionen: Sei $f : [a,b] \to \mathbb{R}$.

1) f ist stetig oder monoton $\Rightarrow f$ ist Regelfunktion.

Genauer: f ist Regelfunktion $\Leftrightarrow$ f besitzt an jeder Stelle $x \in \,]a,b]$ einen linksseitigen Grenzwert und an jeder Stelle $x \in [a,b[$ einen rechtsseitigen Grenzwert

2) $\mathcal{R}([a,b]) :=$ Menge der Regelfunktionen ist ein Vektorraum (übliche Addition von Funktionen, Multiplikation von Funktionen mit Skalaren).
Außerdem: Produkt und Betrag von Regelfunktionen sind Regelfunktionen.

$\|f\|_\infty := \sup\limits_{[a,b]} |f(x)|$ ist eine Norm auf $\mathcal{R}([a,b])$,

$\big(\mathcal{R}([a,b]), \|.\|_\infty\big)$ ist ein vollständiger normierter Raum (**Banachraum**).

8.7 Satz und Definition: Sei $f : [a, b] \to \mathbb{R}$ Regelfunktion, (t_n) eine Folge von Treppenfunktionen mit $f = \lim t_n$ gleichmäßig. Dann existiert der Grenzwert

$$\int_a^b f \;:=\; \int_a^b f(x)\,\mathrm{d}x \;:=\; \lim_{n\to\infty} \int_a^b t_n(x)\,\mathrm{d}x$$

und ist unabhängig von der gewählten Folge (t_n).

$\int_a^b f(x)\,\mathrm{d}x$ heißt das (bestimmte) **Regel-** oder **Cauchy-Integral** von f, f heißt auch (Regel-) **integrierbar**.

Es ist $\int_a^a f = 0$. Wir setzen $\int_b^a f := -\int_a^b f$ für $a < b$.

Beweis: Sei $I_n := \int_a^b t_n(x)\,\mathrm{d}x$.

Schritt 1: Beweise, dass (I_n) eine Cauchy-Folge ist. Dann konvergiert (I_n) in $\mathbb{R}$.

Sei $\varepsilon > 0$ fest. Es existiert ein $N \in \mathbb{N}$, so dass für $n > N$

$$|t_n(x) - f(x)| \;<\; \frac{\varepsilon}{2(b-a)} \qquad \text{für } x \in [a, b].$$

Für $m, n > N$ folgt

$$|t_n(x) - t_m(x)| \;\le\; |t_n(x) - f(x)| + |f(x) - t_m(x)| \;<\; \frac{\varepsilon}{b-a} \qquad \text{für } x \in [a, b].$$

Wähle $a = x_0 < x_1 < \ldots < x_K = b$, so dass

$$t_n = \sum_{k=1}^{K} c_k\,\chi_{[x_{k-1},x_k]} \text{ f.ü.} \quad \text{und} \quad t_m = \sum_{k=1}^{K} d_k\,\chi_{[x_{k-1},x_k]} \text{ f.ü.}$$

$$\Rightarrow\ |c_k - d_k| \;<\; \frac{\varepsilon}{b-a}$$

$$\Rightarrow\ |I_n - I_m| \;\le\; \sum |c_k - d_k|(x_k - x_{k-1}) \;<\; \frac{\varepsilon}{b-a} \sum (x_k - x_{k-1}) \;=\; \varepsilon.$$

Also konvergiert (I_n).

Schritt 2: Eindeutigkeit des Grenzwerts: Ist $(\widetilde{t}_n)$ eine weitere Folge von Treppenfunktionen, die gleichmäßig gegen f konvergiert und $\widetilde{I}_n := \int_a^b \widetilde{t}_n(x)\,\mathrm{d}x$, so betrachte die Folge $t_1, \widetilde{t}_1, t_2, \widetilde{t}_2, \ldots$. Diese konvergiert ebenfalls gleichmäßig gegen f. Somit konvergiert auch die Folge $I_1, \widetilde{I}_1, I_2, \widetilde{I}_2, \ldots$. Dann müssen die Teilfolgen (I_n) und $(\widetilde{I}_n)$ gegen denselben Grenzwert konvergieren. $\square$

8.8 Eigenschaften des Integrals: Seien $a < b$ und $f, g \in \mathcal{R}([a, b])$. Dann gilt

1) Linearität: $\int_a^b (\alpha f + \beta g) = \alpha \int_a^b f + \beta \int_a^b g \qquad$ für $\alpha, \beta \in \mathbb{R}$.

2) Monotonie: $f \geq 0$ auf $[a, b] \quad \Rightarrow \quad \int_a^b f \geq 0$,

$$f \leq g \text{ auf } [a, b] \quad \Rightarrow \quad \int_a^b f \leq \int_a^b g.$$

Spezieller und besser:

Ist $f : [a, b] \to \mathbb{R}$ stetig, $f \geq 0$ auf $[a, b]$, $f(x_0) > 0$ für ein $x_0 \in [a, b]$, dann ist $\int_a^b f > 0$.

3) $f = g$ f.ü. $\Rightarrow \int_a^b f = \int_a^b g$.

4) $\left| \int_a^b f \right| \leq \int_a^b |f| \leq (b-a) \cdot \|f\|_\infty$.

Insbesondere: $\left| \int_a^b f - \int_a^b g \right| \leq (b-a) \cdot \|f - g\|_\infty$,

d.h. die Abbildung $\mathcal{R}([a, b]) \ni f \mapsto \int_a^b f \in \mathbb{R}$ ist stetig.

5) Ist f stetig auf $[a, b]$, dann existiert ein $\xi \in [a, b]$ mit $\int_a^b f = (b-a)\, f(\xi)$.

Allgemeiner: Ist $g \geq 0$ und f stetig, dann existiert $\xi \in [a, b]$ mit $\int_a^b f \cdot g = f(\xi) \cdot \int_a^b g$ (Mittelwertsatz der Integralrechnung).

6) Für $a \leq c \leq d \leq b$ setze

$$\int_c^d f := \int_a^b f \cdot \chi_{[c,d]}.$$

Für $a \leq c \leq b$ gilt $\int_a^b f = \int_a^c f + \int_c^b f$. Diese Formel gilt auch für $a \leq b \leq c$, falls f integrierbar über $[a, c]$.

Beweis: Zu 1), 3) und 4): Klar für Treppenfunktionen.
Grenzübergang $\Rightarrow$ Aussage gilt auch für Regelfunktionen.

Zu 2): Sei $f \geq 0$ und $t_n \to f$ gleichmäßig auf $[a, b]$. Setze $\widetilde{t}_n(x) := \max\{0, t_n(x)\}$. Dann gelten

- $\widetilde{t}_n$ ist Treppenfunktion,
- $\widetilde{t}_n \geq 0$, also $\int_a^b \widetilde{t}_n \geq 0$,
- $\widetilde{t}_n \to f$ gleichmäßig auf $[a, b]$, da $|f(x) - \widetilde{t}_n(x)| \overset{f \text{ positiv}}{\leq} |f(x) - t_n(x)|$.

$$\Rightarrow \int_a^b f = \lim_{n\to\infty} \int_a^b \widetilde{t}_n \geq 0.$$

Sei nun f stetig auf $[a,b]$, $f(x) \geq 0$ auf $[a,b]$ und $f(x_0) > 0$. Wähle $\varepsilon = \dfrac{f(x_0)}{2}$. Aus der Stetigkeit von f folgt für $x_0 - \delta < x < x_0 + \delta$:

$$f(x) \;=\; f(x_0) + f(x) - f(x_0) \;\geq\; f(x_0) - |f(x) - f(x_0)| \;>\; f(x_0) - \varepsilon \;=\; \frac{f(x_0)}{2}.$$

$$\begin{aligned} &\Rightarrow \quad f \geq g \text{ mit } g(x) := \begin{cases} 0 & |x - x_0| \geq \delta \\ \dfrac{f(x_0)}{2} & |x - x_0| < \delta \end{cases} \\ &\overset{4)}{\Rightarrow} \quad \int_a^b f \;\geq\; \int_a^b g \;=\; \frac{f(x_0)}{2} 2\delta \;>\; 0. \end{aligned}$$

Achtung: Diese Aussage gilt nicht, wenn die Stetigkeit von f weggelassen wird: Betrachte f auf $[0,2]$ mit $f(x) = 0$ für $x \neq 1$ und $f(x) = 1$ für $x = x_0 = 1$.

Zu 5): Sei $m := \min\limits_{[a,b]} f(x),\;\; M := \max\limits_{[a,b]} f(x)$.

$$\begin{aligned} &\overset{2)}{\Rightarrow} \quad m \int_a^b g \;\leq\; \int_a^b gf \;\leq\; M \int_a^b g \\ &\Rightarrow \quad \exists \mu \in [m, M] : \mu \int_a^b g = \int_a^b fg \end{aligned}$$

Zwischenwertsatz (siehe 5.8): $\exists \xi \in [a,b] : \mu = f(\xi)$.

Zu 6): Folgt aus $f = f \cdot \chi_{[a,c]} + f \cdot \chi_{[c,b]}$ f.ü. $\square$

8.2 Stammfunktionen

8.9 Wichtige Idee: Ist $f : [a,b] \to \mathbb{R}$ integrierbar, dann heißen die Abbildungen

$$F : [a,b] \to \mathbb{R} : x \mapsto \int_a^x f(t)\,\mathrm{d}t, \quad G : [a,b] \to \mathbb{R} : x \mapsto \int_x^b f(t)\,\mathrm{d}t$$

Flächeninhaltsfunktionen.

8.10 Satz: Ist $f : [a,b] \to \mathbb{R}$ integrierbar, dann sind die Flächeninhaltsfunktionen stetig.

Beweis: $$\begin{aligned} |F(x) - F(x_0)| \;&=\; \left| \int_a^x f - \int_a^{x_0} f \right| \;=\; \left| \int_{x_0}^x f \right| \\ &\overset{8.8,\text{ Teil4}}{\leq}\; |x - x_0|\, \|f\|_\infty \;<\; \varepsilon \quad \text{falls } |x - x_0| < \frac{\varepsilon}{\|f\|_\infty} \end{aligned}$$

Also ist F stetig. Die Stetigkeit von G folgt aus $G(x) = \displaystyle\int_a^b f - F(x)$. $\square$

8.11 Hauptsatz der Integral- und Differentialrechnung(1667): Sei $f : [a,b] \to \mathbb{R}$ stetig und $F : [a,b] \to \mathbb{R} : x \mapsto \int_a^x f(t)\,\mathrm{d}t$. Dann ist F differenzierbar auf $]a,b[$, und es gilt $F' = f$.

Also: $\underbrace{\text{Integration}}_{\text{Flächenproblem}}$ ist Umkehrung der $\underbrace{\text{Differentiation}}_{\text{Tangentenproblem}}$.

Beweis: Sei $x_0 \in \,]a,b[\,$, $\Delta(x) := \dfrac{F(x)-F(x_0)}{x-x_0}$ für $x \neq x_0$. Zeige $\lim\limits_{\substack{x \to x_0 \\ x \neq x_0}} \Delta(x) = f(x_0)$.

$$\begin{aligned}
|\Delta(x) - f(x_0)| &= \left| \frac{1}{x-x_0} \left(\int_a^x f(t)\,\mathrm{d}t - \int_a^{x_0} f(t)\,\mathrm{d}t - \int_{x_0}^x f(x_0)\,\mathrm{d}t \right) \right| \\
&= \frac{1}{|x-x_0|} \left| \int_{x_0}^x (f(t) - f(x_0))\,\mathrm{d}t \right| \\
&\overset{\text{8.8, Teil 4}}{\leq} \frac{1}{|x-x_0|} |x-x_0| \max_{t \in [x,x_0]} |f(t) - f(x_0)|.
\end{aligned}$$

Da f stetig ist, gibt es ein $\delta > 0$, so dass

$$\begin{aligned}
|f(t) - f(x_0)| &< \varepsilon \qquad \text{für } |t - x_0| < \delta \\
\Rightarrow |\Delta(x) - f(x_0)| &\leq \varepsilon \qquad \text{für } |x - x_0| < \delta.
\end{aligned}$$

□

8.12 Definition: Sei $f : [a,b] \to \mathbb{R}$. Eine Funktion $F : [a,b] \to \mathbb{R}$ heißt **Stammfunktion** von f, falls F stetig auf $[a,b]$, differenzierbar auf $]a,b[$ und $F' = f$ auf $]a,b[$ ist.

Falls f stetig ist, besitzt f eine Stammfunktion (Hauptsatz 8.11).

Alle Stammfunktionen zu f:

(i) F Stammfunktion $\Rightarrow$ $F + c$ ist Stammfunktion für beliebige Konstante c.

(ii) F, G Stammfunktionen $\Rightarrow$ $(F-G)' = f - f = 0 \overset{6.21}{\Rightarrow} F - G = \text{const}$.

8.13 Flächenberechnung: Ist $f : [a,b] \to \mathbb{R}$ stetig und F eine Stammfunktion von f, so gilt für $a \leq c \leq d \leq b$:

$$\int_c^d f(t)\,\mathrm{d}t \;=\; F(d) - F(c) \;=:\; F(x)\Big|_{x=c}^{d} \;=:\; \Big[F(x)\Big]_{x=c}^{d}.$$

Beweis: Sei $G(x) := \int_a^x f \overset{\text{Hauptsatz}}{\Rightarrow} G$ ist Stammfunktion

$$\begin{aligned}
\Rightarrow \quad F &= G + \gamma \\
\Rightarrow \quad \int_c^d f &= \int_a^d f - \int_a^c f \;=\; G(d) - G(c) \;=\; F(d) - \gamma - (F(c) - \gamma) \;=\; F(d) - F(c)
\end{aligned}$$

□

8.14 Beispiele: **1)** $\displaystyle\int_0^2 |x^2-1|\,\mathrm{d}x \;=\; \int_0^1 (1-x^2)\,\mathrm{d}x + \int_1^2 (x^2-1)\,\mathrm{d}x \;=\; \ldots \;=\; 2.$

2) Sei $f(x) := \begin{cases} 1 & 0 \le x < 1 \\ 2-x & 1 \le x < 2 \\ 1 & 2 \le x \le 3 \end{cases}$

$$\Rightarrow\ F(x) \;:=\; \int_0^x f(t)\,\mathrm{d}t \;=\; \begin{cases} x & 0 \le x < 1 \\ -\frac{1}{2} + 2x - \frac{x^2}{2} & 1 \le x < 2 \\ x - \frac{1}{2} & 2 \le x < 3 \end{cases}$$

F ist in $x=1$ differenzierbar, in $x=2$ nicht.

8.15 Definition: Die Menge aller Stammfunktionen

$$\int f(x)\,\mathrm{d}x \;:=\; \big\{F : [a,b] \to \mathbb{R} \;\big|\; F \text{ ist Stammfunktion von } f\big\}$$

heißt das **unbestimmte Integral** von f.

Ist F eine Stammfunktion von f, so gilt $\displaystyle\int f(x)\,\mathrm{d}x \;=\; \{F+c : c \in \mathbb{R}\}$, und wir schreiben kurz

$$\int f(x)\,\mathrm{d}x \;=\; F(x) + c.$$

8.3 Wie findet man Stammfunktionen?

Wichtigste Methode: Raten. Z.B.

$$\int x^\alpha\,\mathrm{d}x \;=\; \frac{x^{\alpha+1}}{\alpha+1} + c \quad (\alpha \ne -1), \qquad \int \sin x\,\mathrm{d}x \;=\; -\cos x + c,$$

$$\int \mathrm{e}^x\,\mathrm{d}x \;=\; \mathrm{e}^x + c, \qquad \int \frac{1}{1+x^2}\,\mathrm{d}x \;=\; \arctan x + c,$$

$$\int \frac{1}{x}\,\mathrm{d}x \;=\; \begin{cases} \ln x + c, & \text{für } x > 0, \\ \ln(-x) + c, & \text{für } x < 0, \end{cases} \qquad \text{kurz: } \int \frac{1}{x}\,\mathrm{d}x \;=\; \ln|x| + c \text{ für } x \ne 0.$$

8.16 Ratehilfen: **1)** Seien $u, v : [a,b] \to \mathbb{R}$ differenzierbar in $]a,b[$. Dann gilt

$$\int u' \cdot v \;=\; u \cdot v - \int u \cdot v' \qquad (\textbf{Partielle Integration}).$$

2) Sei $f : [a,b] \to \mathbb{R}$ stetig, $\varphi : [\alpha,\beta] \to [a,b]$ stetig differenzierbar. Dann gilt

$$\int f(\varphi(t)) \cdot \varphi'(t)\,\mathrm{d}t \;=\; \left[\int f(x)\,\mathrm{d}x\right]_{x=\varphi(t)} \qquad (\textbf{Integration durch Substitution}).$$

Wichtig: Diese Formel kann von links nach rechts **und** von rechts nach links benützt werden: Ist φ invertierbar, so gilt

$$\int f(x)\,\mathrm{d}x \;=\; \int f(\varphi(t)) \cdot \varphi'(t)\,\mathrm{d}t\,\Big|_{t=\varphi^{-1}(x)} \qquad (\textbf{Integration durch Substitution})$$

8.17 Beispiele: **1)** $\displaystyle\int x\sin x\,\mathrm{d}x \;=\; -x\cos x+\sin x+c$

2) $\displaystyle\int \ln x\,\mathrm{d}x \;=\; x\cdot\ln|x|-x+c$ für $x\neq 0$.

3) $\displaystyle\int \cos^2 x\,\mathrm{d}x \;=\; \frac{1}{2}(\cos x\cdot\sin x+x)+c.$

4) $\displaystyle\int 2t\sin(t^2+1)\,\mathrm{d}t \;=\; -\cos(t^2+1)+c.$ Dies hätte man auch erraten können!

5) Für $x\in\,]-1,1[$: $\displaystyle\int\sqrt{1-x^2}\,\mathrm{d}x=\frac{1}{2}\left(x\sqrt{1-x^2}+\arcsin x+c\right)$ (Substitution $x=\sin t$).

6) Für $x>-\sqrt[3]{4}$: $\displaystyle\int\frac{1}{4+x^{1/3}}\,\mathrm{d}x=\frac{3}{2}x^{2/3}-12x^{1/3}+48\ln(x^{1/3}+4)+c$ (Substitution $x=t^3$).

8.4 Integration rationaler Funktionen

Ziel: Darstellung einer rationalen Funktion $f(x)=\dfrac{P(x)}{Q(x)}$ (P,Q Polynome) als Summe „einfacher“ Brüche, die dann integriert werden können

8.18 Hilfssatz: Seien P,Q teilerfremde Polynome mit $\operatorname{Grad}P<\operatorname{Grad}Q$, und λ sei n-fache Nullstelle von Q:

$$Q(z) \;=\; (z-\lambda)^n\,Q_1(z) \quad\text{mit}\quad Q_1(\lambda)\neq 0.$$

Dann gibt es ein $a_1\in\mathbb{C}$ und ein Polynom P_1, so dass

$$\frac{P(z)}{Q(z)} \;=\; \frac{P(z)}{(z-\lambda)^nQ_1(z)} \;=\; \frac{a_1}{(z-\lambda)^n}+\frac{P_1(z)}{(z-\lambda)^{n-1}Q_1(z)};$$

a_1 und P_1 sind eindeutig, und es gilt $\operatorname{Grad}P_1<(\operatorname{Grad}Q)-1$.

Fortsetzung dieses Verfahrens:

$$\frac{P(z)}{Q(z)} \;=\; \frac{a_1}{(z-\lambda)^n}+\frac{a_2}{(z-\lambda)^{n-1}}+\ldots+\frac{a_n}{(z-\lambda)^1}+\frac{P_n(z)}{Q_1(z)}$$

mit $\operatorname{Grad}P_n<(\operatorname{Grad}Q)-n=\operatorname{Grad}Q_1$.

8.19 Beispiel: $\dfrac{4z^6-5z+8}{(z-\mathrm{i})^4(z+1+\mathrm{i})(z-4)^2}$

$$=\;\frac{a_1}{z-\mathrm{i}}+\frac{a_2}{(z-\mathrm{i})^2}+\frac{a_3}{(z-\mathrm{i})^3}+\frac{a_4}{(z-\mathrm{i})^4}+\frac{b}{z+1+\mathrm{i}}+\frac{c_1}{z-4}+\frac{c_2}{(z-4)^2}.$$

8.20 Komplexe Partialbruchzerlegung: Seien P, Q teilerfremd mit Grad $P <$ Grad Q,

$$Q(z) \;=\; a_n(z-\lambda_1)^{n_1}\cdots(z-\lambda_k)^{n_k},$$

$\lambda_1, \ldots, \lambda_k$ paarweise verschieden (d.h $n_1, \ldots, n_k$ sind die (algebraischen) Vielfachheiten der Nullstellen $\lambda_1, \ldots, \lambda_k$). Dann existieren eindeutig bestimmte Zahlen $a_{j,i} \in \mathbb{C}$, so dass

$$\frac{P(z)}{Q(z)} \;=\; \sum_{i=1}^{k} \left(\frac{a_{1,i}}{\lambda-\lambda_i} + \frac{a_{2,i}}{(\lambda-\lambda_i)^2} + \ldots + \frac{a_{n_i,i}}{(\lambda-\lambda_i)^{n_i}} \right).$$

8.21 Bemerkungen: **1)** Falls Grad $P \geq$ Grad Q: Erst abdividieren:

$$\frac{P}{Q} \;=\; P_1 + \frac{R}{Q} \qquad \text{mit } \operatorname{Grad} R < \operatorname{Grad} Q,\; P_1 \text{ Polynom},$$

dann 8.20 auf $\dfrac{R}{Q}$ anwenden.

2) Falls P und Q gemeinsame Nullstellen haben: Durch größtes gemeinsames Teilerpolynom kürzen.

8.22 Beispiel: $f(x) \;=\; \dfrac{P(x)}{Q(x)} \;=\; \dfrac{4x^2+9x-4}{(x-1)(x+2)^2} \overset{\text{Theorie}}{=} \dfrac{a}{x-1} + \dfrac{b}{x+2} + \dfrac{c}{(x+2)^2}.$

Zuhaltemethode: Multipliziere mit $(x-1)$:

$$\frac{4x^2+9x-4}{(x+2)^2} \;=\; a + (x-1)(\ldots) \overset{x=1}{\Rightarrow} a = \frac{9}{3^2} = 1.$$

Zuhaltemethode: Multipliziere mit $(x+2)^2$:

$$\frac{4x^2+9x-4}{x-1} \;=\; c + (x+2)(\ldots) \overset{x=-2}{\Rightarrow} c = \frac{-6}{-3} = 2.$$

b kann nicht durch die Zuhaltemethode bestimmt werden. Einen x-Wert einsetzen:

$$\frac{4x^2+9x-4}{(x-1)(x+2)^2} \;=\; \frac{1}{x-1} + \frac{b}{x+2} + \frac{2}{(x+2)^2} \overset{x=0}{\Rightarrow} \frac{-4}{-4} = -1 + \frac{b}{2} + \frac{1}{2} \;\Rightarrow\; b = 3.$$

$$\Rightarrow \int f(x)\,\mathrm{d}x \;=\; \int \left(\frac{1}{x-1} + \frac{3}{x+2} + \frac{2}{(x+2)^2} \right) \mathrm{d}x \;=\; \ln|x-1| + 3\ln|x+2| - \frac{2}{x+2} + c.$$

8.23 Wiederholung reelle Polynome: Sei Q ein reelles Polynom, d.h. die Koeffizienten von Q sind reell. Dann gilt

$$Q(\lambda) = 0 \;\Rightarrow\; Q(\overline{\lambda}) = 0.$$

Ist $\lambda \in \mathbb{C} \setminus \mathbb{R}$ Nullstelle der Vielfachheit n, so folgt

$$Q(x) \;=\; \left(x-\lambda\right)^n \left(x-\overline{\lambda}\right)^n Q_1(x),$$

wobei Q_1 wieder ein reelles Polynom ist und $Q_1(\lambda_j) \neq 0$ gilt. Mit

$$\big(x-\lambda\big)\,\big(x-\overline{\lambda}\big) \;=\; x^2 - 2\mathrm{Re}(\lambda)\,x + |\lambda|^2 \;=:\; x^2 + \beta\,x + \gamma$$

ergibt sich als reelle Faktorisierung

$$Q(x) \;=\; a_n(x-\lambda_1)^{n_1}\cdots(x-\lambda_k)^{n_k}(x^2+\beta_{k+1}x+\gamma_{k+1})^{n_{k+1}}\cdots(x^2+\beta_l x+\gamma_l)^{n_l}$$

($\lambda_1,\dots,\lambda_k$: reelle Nullstellen;
$\lambda_{k+1},\overline{\lambda}_{k+1},\dots,\lambda_l,\overline{\lambda}_l$ nichtrelle Nullstellen, $\beta_j := -2\mathrm{Re}(\lambda_j)$, $\gamma_j := |\lambda_j|^2$).

8.24 Relle Partialbruchzerlegung: Sind P, Q teilerfremde reelle Polynome mit $\mathrm{Grad}\,P < \mathrm{Grad}\,Q$, und hat Q die oben stehende relle Zerlegung, dann kann $\dfrac{P(x)}{Q(x)}$ als Summe von Termen der Form

$$\frac{a}{(x-\lambda_j)^i} \quad (j=1,\dots,k;\, 1\le i\le n_j), \qquad \frac{bx+c}{(x^2+\beta_j x+\gamma_j)^i} \quad (j=k+1,\dots,l;\, 1\le i\le n_j)$$

dargestellt werden.

Berechnung durch

1) Zuhaltemethode,

2) Einsetzen verschiedener x-Werte führt auf lineares Gleichungssystem.

3) Hauptnenner und Koeffizientenvergleich,

8.25 Beispiele: **1)** $Q(x) = (x-1)^4(x+2)(x-(1+\mathrm{i}))^3(x-(1-\mathrm{i}))^3$, P Polynom mit $\mathrm{Grad}\,P \le 10$, $P(1)\neq 0$, $P(-2)\neq 0$, $P(1\pm\mathrm{i})\neq 0$:

$(x-(1+\mathrm{i}))\,(x-(1-\mathrm{i})) = x^2+2x+4$. Also gibt es Konstanten, so dass

$$\begin{aligned}\frac{P(x)}{Q(x)} \;&=\; \frac{a_1}{x-1}+\frac{a_2}{(x-1)^2}+\frac{a_3}{(x-1)^3}+\frac{a_4}{(x-1)^4}+\frac{b}{x+2}\\ &\quad+\frac{c_1\,x+d_1}{x^2-2x+4}+\frac{c_2\,x+d_2}{(x^2-2x+4)^2}+\frac{c_3\,x+d_3}{(x^2-2x+4)^3}\end{aligned}$$

2)
$$\begin{aligned}\frac{x^4+x^3+x}{x^3+1} \;&=\; x+1-\frac{\frac{1}{3}}{x+1}-\frac{-\frac{1}{3}\,x+\frac{2}{3}}{x^2-x+1}\\ &=\; x+1-\frac{1}{3}\,\frac{1}{x+1}+\frac{1}{6}\,\frac{2x-1}{x^2-x+1}-\frac{1}{2}\,\frac{1}{x^2-x+1}.\end{aligned}$$

$$\Rightarrow \int \frac{x^4+x^3+x}{x^3+1}\,\mathrm{d}x$$
$$= \frac{x^2}{2}+x-\frac{1}{3}\ln|x+1|+\frac{1}{6}\ln(x^2-x+1)-\frac{1}{2}\sqrt{\frac{4}{3}}\arctan\left(\sqrt{\frac{4}{3}}\left(x-\frac{1}{2}\right)\right)+c$$

8.5 Uneigentliche Integrale

Ziel: Erweiterung des Integralbegriffs auf unbeschränkte Intervalle und unbeschränkte Funktionen.

8.26 Definition: Sei $I \subseteq \mathbb{R}$ beliebiges Intervall; $f : I \to \mathbb{R}$ heißt **lokal integrierbar**, falls die Einschränkung von f auf jedes Intervall $[a, b] \subseteq I$ integrierbar ist.

8.27 Beispiele: **1)** $f : I = [0, \infty[\to \mathbb{R} : x \mapsto \dfrac{1}{x+1}$ ist lokal intergrierbar.

2) $f : I = \,]0, 1] \to \mathbb{R} : x \mapsto \dfrac{1}{x}$ ist lokal integrierbar.

3) Ist $f : I = [a, b] \to \mathbb{R}$ integrierbar, dann ist f auf I lokal integrierbar.

8.28 Definition: Sei $I = [a, b[\,,\ -\infty < a < b \leq \infty$, $f : I \to \mathbb{R}$ lokal integrierbar. Dann heißt f **uneigentlich integrierbar** über I, falls

$$\lim_{\beta \uparrow b} \int_a^\beta f(x)\,\mathrm{d}x \ =: \int_a^b f(x)\,\mathrm{d}x$$

existiert. Wir sagen auch: $\displaystyle\int_a^b f(x)\,\mathrm{d}x$ **konvergiert**.

Falls der Grenzwert für $\beta \uparrow b$ nicht existiert: $\displaystyle\int_a^b f(x)\,\mathrm{d}x$ **divergiert**.

Genauso: $I = \,]a, b]$, $-\infty \leq a < b$:

$$\int_a^b f(x)\,\mathrm{d}x := \lim_{\alpha \downarrow a} \int_\alpha^b f(x)\,\mathrm{d}x,$$

$I = \,]a, b[\,,\ -\infty \leq a < b \leq \infty$:

$$\int_a^b f(x)\,\mathrm{d}x \ := \ \lim_{\alpha \downarrow a} \int_\alpha^c f(x)\,\mathrm{d}x + \lim_{\beta \uparrow b} \int_c^\beta f(x)\,\mathrm{d}x, \qquad c \in \,]a, b[\ \text{ beliebig.}$$

8.29 Beispiele: **1)** $\displaystyle\int_0^\infty \mathrm{e}^{-x}\,\mathrm{d}x \ = \ 1.$

2) $\displaystyle\int_{-\infty}^\infty \frac{1}{1+x^2}\,\mathrm{d}x \ = \ \pi.$

3) $\displaystyle\int_1^\infty \frac{1}{x}\,\mathrm{d}x$ ist divergent

4) $$\int_1^\infty \frac{1}{x^s}\,\mathrm{d}x \ = \ \begin{cases} \dfrac{1}{s-1} & \text{für } s > 1, \\ \text{divergent} & \text{für } s \leq 1. \end{cases}$$

5) $\displaystyle\int_0^1 \frac{1}{x^s}\,\mathrm{d}x \;=\; \begin{cases} \dfrac{1}{1-s} & \text{für } s<1, \\ \text{divergent} & \text{für } s \geq 1. \end{cases}$

8.30 Vergleichskriterium für uneigentliche Integrale: Sei $f : [a,b[\to \mathbb{R}$ lokal integrierbar. Falls

$$\big(\forall x \in [a,b[: |f(x)| \leq g(x)\big) \;\wedge\; g \text{ über } [a,b[\text{ uneigentlich integrierbar,}$$

dann ist auch f uneigentlich integrierbar über $[a,b[$.
Entsprechend für uneigentliche Integrale über $]a,b]$ oder $]a,b[$.

Beweis: Sei (x_n) Folge in $[a,b[$, $x_n \uparrow b$ und $y_n := \displaystyle\int_a^{x_n} f(x)\,\mathrm{d}x$. Für $n > m$ gilt

$$\begin{aligned} |y_n - y_m| \;&=\; \left|\int_{x_m}^{x_n} f(x)\,\mathrm{d}x\right| \;\leq\; \int_{x_m}^{x_n} |f(x)|\,\mathrm{d}x \\ &\leq\; \int_{x_m}^{x_n} g(x)\,\mathrm{d}x \;=\; \int_a^{x_n} g(x)\,\mathrm{d}x - \int_a^{x_m} g(x)\,\mathrm{d}x \\ &<\; \varepsilon \qquad \text{für } n > m > N_\varepsilon. \end{aligned}$$

$\Rightarrow$ (y_n) ist Cauchy-Folge in $\mathbb{R}$, also konvergent.

Genauso zeigt man: Ist $(\widetilde{x}_n)$ eine weitere Folge in $[a,b[$ mit $x_n \uparrow b$, so gilt

$$\lim_{n\to\infty} \int_a^{\widetilde{x}_n} f(x)\,\mathrm{d}x \;=\; \lim_{n\to\infty} y_n.$$

□

8.31 Beispiele: **1)** $\displaystyle\int_1^\infty \frac{x^s}{1+x^2}\,\mathrm{d}x$ konvergiert für $s<1$, divergiert für $s \geq 1$.

2) $\displaystyle\int_1^\infty \frac{1}{x^2+3x+2}\,\mathrm{d}x.$

8.32 Integralkriterium für Reihen: Sei $f : [1,\infty[\to [0,\infty[$ monoton fallend. Dann gilt:

$$\int_1^\infty f(x)\,\mathrm{d}x \text{ konvergiert} \;\Leftrightarrow\; \sum_{n=1}^\infty f(n) \text{ konvergiert.}$$

Beweis: (i) "$\Rightarrow$":

$$\begin{aligned} \sum_{n=2}^m f(n) \;&\leq\; \int_1^m f(x)\,\mathrm{d}x \;\leq\; \int_1^\infty f(x)\,\mathrm{d}x \\ &\Rightarrow\; \sum_{n=1}^\infty f(n) \text{ ist beschränkt} \\ &\overset{f(n)\geq 0}{\Rightarrow}\; \sum_{n=1}^\infty f(n) \text{ ist konvergent} \end{aligned}$$

(ii) "$\Leftarrow$": Definiere $g(x) := f(n)$ für $n \le x < n+1$.

$$\Rightarrow \left\{ \begin{array}{l} 0 \le f(x) \le g(x) \\ \displaystyle\int_1^\infty g(x)\,\mathrm{d}x \;=\; \sum_{n=1}^\infty f(n) \text{ ist konvergent} \end{array} \right. \quad \Rightarrow \int_1^\infty f(x)\,\mathrm{d}x \text{ ist konvergent.}$$

$\square$

8.33 Beispiel: $\sum \frac{1}{n^s}$ ist konvergent für $s > 1$ und divergent für $s \le 1$.

8.6 Parameterabhängige Integrale

8.34 Beispiele: **1)** $\displaystyle\int_1^2 \sin(tx)\,\mathrm{d}x \;=\; \begin{cases} \frac{1}{t}(\cos t - \cos 2t) & t \neq 0 \\ 0 & t = 0 \end{cases}$

2) Für $t \neq 0$: $\displaystyle\int_t^{1/t} \mathrm{e}^x\,\mathrm{d}x \;=\; \mathrm{e}^{1/t} - \mathrm{e}^t.$

8.35 Satz: Sei $f \in C^1(\mathbb{R}^2 \to \mathbb{R})$, $I(t) := \displaystyle\int_a^b f(t,x)\,\mathrm{d}x$. Dann gilt

$$I'(t) \;=\; \int_a^b \partial_1 f(t,x)\,\mathrm{d}x.$$

(Ableitung und Integration sind vertauschbar)

8.36 Folgerung: Seien $\alpha, \beta \in C^1(\mathbb{R} \to \mathbb{R})$, $f \in C^1(\mathbb{R}^2 \to \mathbb{R})$, $I(t) := \displaystyle\int_{\alpha(t)}^{\beta(t)} f(t,x)\,\mathrm{d}x$. Dann gilt

$$I'(t) \;=\; \int_{\alpha(t)}^{\beta(t)} \partial_1 f(t,x)\,\mathrm{d}x + f(t,\beta(t))\,\beta'(t) - f(t,\alpha(t))\,\alpha'(t).$$

Beweis: Setze $F(u_1,u_2,u_3) := \displaystyle\int_{u_1}^{u_2} f(u_3,x)\,\mathrm{d}x$

$$\Rightarrow \left\{ \begin{array}{lcl} \partial_1 F(u_1,u_2,u_3) & = & -f(u_3,u_1) \\ \partial_2 F(u_1,u_2,u_3) & = & f(u_3,u_2) \\ \partial_3 F(u_1,u_2,u_3) & = & \int_{u_1}^{u_2} \partial_1 f(u_3,x)\,\mathrm{d}x \end{array} \right.$$

Kettenregel 7.27 anwenden auf $I(t) = F(\alpha(t),\beta(t),t)$:

$$I'(t) \;=\; (\partial_1 F)\,\alpha'(t) + (\partial_2 F)\,\beta'(t) + (\partial_3 F)\,1.$$

$\square$

8.37 Beispiel: $I(t) = \displaystyle\int_{\sqrt{t}}^{t^2} \mathrm{e}^{-tx^2}\,\mathrm{d}x \;\Rightarrow\; I'(t) = \int_{\sqrt{t}}^{t^2} (-x^2)\mathrm{e}^{-tx^2}\,\mathrm{d}x + \mathrm{e}^{-t^5}\,2t - \mathrm{e}^{-t^2}\,\frac{1}{2\sqrt{t}}$

8.7 Einführung in die numerische Integration

Problem: $\int\limits_a^b f$ nicht explizit berechenbar (z.B. $\int\limits_0^1 \mathrm{e}^{-x^2}\,\mathrm{d}x$, $\int_0^1 \frac{\sin x}{x}\,\mathrm{d}x$).

1. Idee (Interpolation): Wähle **Stützstellen** $a = x_0 < x_1 < \ldots < x_n = b$, berechne Interpolationspolynom P_n durch $(x_j, f(x_j))$ und hoffe $\int_a^b f \approx \int_a^b P_n$.

8.38 Definition: Sei $f : [a,b] \to \mathbb{R}$, $a = x_0 < x_1 < \ldots < x_n = b$. Das Polynom P_n n-ten Grades mit

$$P(x_j) \;=\; f(x_j), \qquad \text{für } j = 0, \ldots, n$$

heißt **Interpolationspolynom** zu f und $\{x_0, \ldots, x_n\}$.

Existenz: Sei $P(x) = a_0 + a_1 x + \ldots + a_n x^n$. Bestimme die Koeffizienten a_j als Lösung von

$$P(x_j) = f(x_j) \;\Leftrightarrow\; a_0 + a_1 x_j + \ldots + a_n x_j^n = f(x_j) \qquad (j = 0, \ldots, n).$$

Dies ist ein LGS mit $n+1$ Gleichungen für $n+1$ Unbekannte.

Aus dem Identitätssatz 2.67 wissen wir: Es gibt höchstens eine Lösung. Für unser LGS bedeutet das: Die Koeffizientenmatrix hat Höchstrang. Oder anders

$$\begin{vmatrix} 1 & x_0 & x_0^2 & \ldots & x_0^n \\ 1 & x_1 & x_1^2 & \ldots & x_1^n \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_n & x_n^2 & \ldots & x_n^n \end{vmatrix} \;\neq\; 0.$$

$\Rightarrow$ Die Koeffizientenmatrix ist invertierbar

$\Rightarrow$ Es existiert eine eindeutige Lösung

8.39 Satz: Sei $f \in C^{n+1}([a,b] \to \mathbb{R})$, P_n das Interpolationspolynom zu den Stützstellen $x_0, \ldots, x_n$ und $R_n := f - P_n$. Dann gibt es ein $\xi \in \,]a,b[\,$, so dass

$$R_n(x) \;=\; \frac{1}{(n+1)!}\, f^{(n+1)}(\xi) \prod_{j=0}^{n} (x - x_j).$$

Insbesonder kann der Fehler R_n abgeschätzt werden durch

$$|R_n(x)| \;\leq\; \frac{1}{(n+1)!} \max_{t \in [a,b]} \left| f^{(n+1)}(t) \right| \prod_{j=0}^{n} |x - x_j|.$$

Beweis: Sei $\omega(x) := \prod_{j=0}^{n}(x - x_j) = x^{n+1} + c_n x^n + \ldots + c_0 \;\Rightarrow\; \omega^{(n+1)}(x) = (n+1)!$.

Für festes $x \in [a,b],\; x \neq x_0, \ldots, x_n$ betrachte

$$\varphi(t) \;:=\; R_n(t) - \frac{R_n(x)}{\omega(x)}\,\omega(t).$$

Wegen $R_n(x_j) = 0,\; \omega(x_j) = 0$ für $j = 0, \ldots, n$ gilt

$$\varphi(x_0) = 0,\; \ldots,\; \varphi(x_n) = 0,\; \varphi(x) = 0.$$

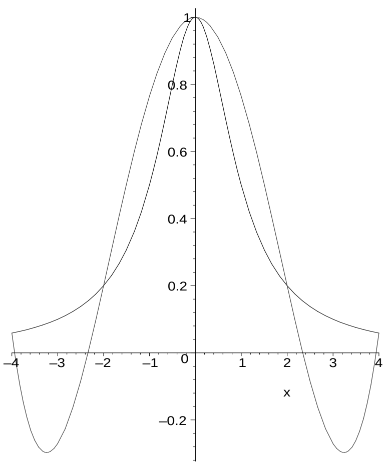
Also hat φ mindestens $n+2$ verschiedene Nullstellen in $[a,b]$.

$$\begin{aligned}
&\overset{\text{Satz von Rolle}}{\Rightarrow} && \varphi' \text{ hat mindestens } n+1 \text{ verschiedene Nullstellen in } \;]a,b[\\
&\Rightarrow && \varphi'' \text{ hat mindestens } n \text{ verschiedene Nullstellen in } \;]a,b[\\
& && \vdots\\
&\Rightarrow && \varphi^{(n+1)} \text{ hat mindestens eine Nullstelle in } \;]a,b[\\
& && \text{d.h. } \exists \xi \in\,]a,b[: 0 = \varphi^{(n+1)}(\xi) = \underbrace{R_n^{(n+1)}(\xi)}_{=(f-P_n)^{(n+1)} = f^{(n+1)}} - \frac{R_n(x)}{\omega(x)}\,(n+1)!\\
&\Rightarrow && f^{(n+1)}(\xi) = \frac{R_n(x)}{\omega(x)}\,(n+1)!
\end{aligned}$$

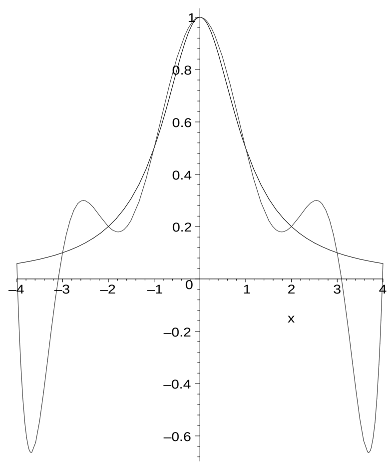
□

8.40 Die große Enttäuschung: $f : [-4,4] \to \mathbb{R} : x \mapsto \dfrac{1}{1+x^2}$:

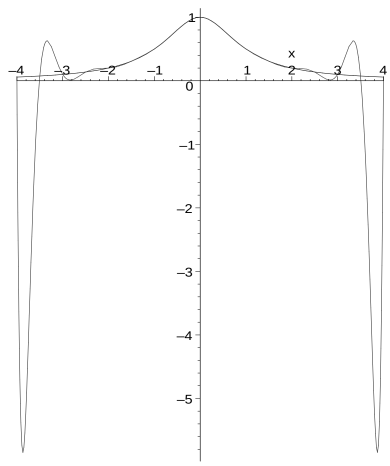
5 Stützstellen



9 Stützstellen



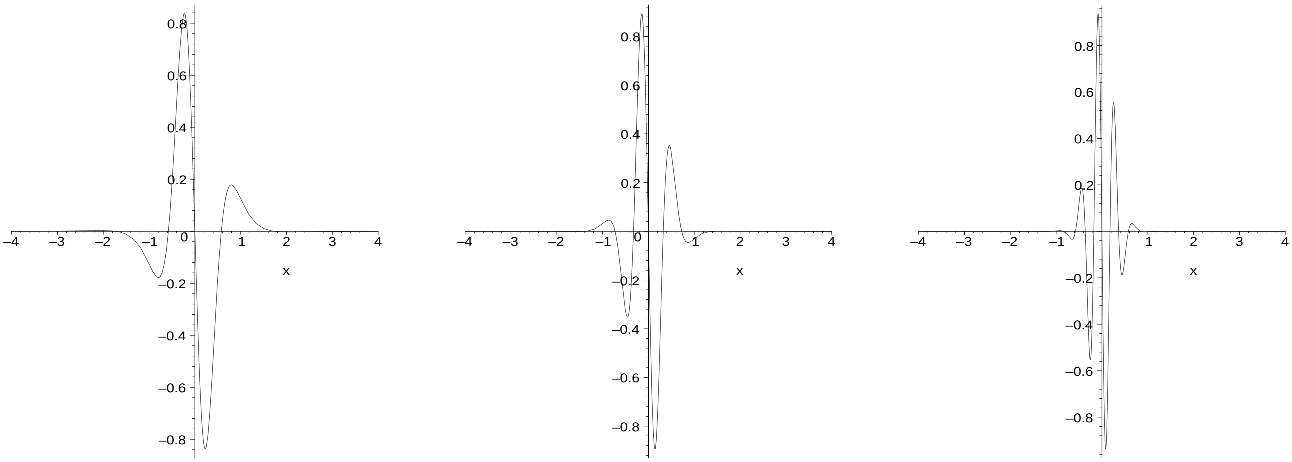
17 Stützstellen



Mit wachsender Anzahl der Stützstellen wird die Approximation schlechter. Warum?

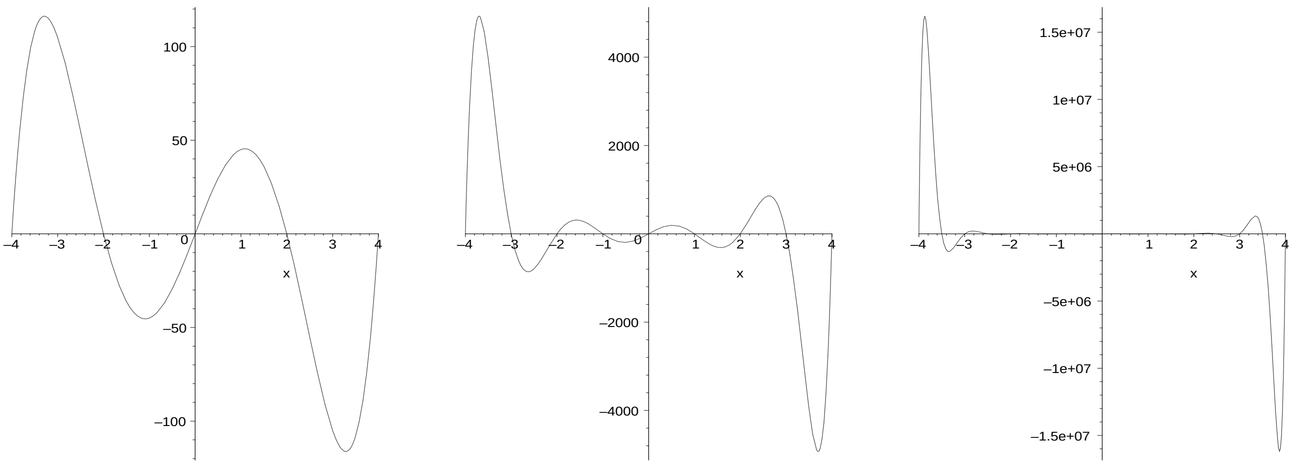
Bestimme Maximum von $\dfrac{1}{(n+1)!}\,\left|f^{(n+1)}(x)\right|$:

$n = 4$ $\qquad$ $n = 8$ $\qquad$ $n = 16$



Bestimme $\displaystyle\prod_{j=0}^{n} |x - x_j|$:

$n = 4$ $\qquad$ $n = 8$ $\qquad$ $n = 16$



8.41 Bemerkung: Man kann die Stützstellen geschickter wählen (nicht äquidistant), damit das Polynom $\displaystyle\prod_{j=0}^{n} |x - x_j|$ nicht so groß wird (Nullstellen von Tschebyscheff-Polynomen).

Man beschränkt sich deshalb auf Polynome kleiner Ordnung.

Zunächst der Fall $n = 1$, d.h. f wird durch ein Polynom 1. Grades („Gerade“) approximiert:

8.42 Satz: Sei $f \in C^2([a,b] \to \mathbb{R})$. Dann gilt die **Trapezregel**

$$\int_a^b f(x)\,\mathrm{d}x \;=\; \frac{b-a}{2}\big(f(a)+f(b)\big) + R \quad \text{mit } |R| \;\le\; \frac{(b-a)^3}{12}\,\|f''\|_\infty\,.$$

Beweis: $|R| \overset{8.39}{\leq} \dfrac{1}{2!} \max |f''(t)| \underbrace{\int_a^b |x-a|\,|x-b| \,\mathrm{d}x}_{=\frac{(b-a)^3}{6}}.$ $\square$

Der Fall $n = 2$, dh. f wird durch ein Polynom 2. Grades („Parabel“) approximiert:

8.43 Satz: Es sei $f \in C^4([a,b] \to \mathbb{R})$. Dann gilt die **Simpson-Formel** (Keplersche Fassregel):

$$\int_a^b f(x)\,\mathrm{d}x \;=\; \frac{b-a}{6}\left(f(a) + 4f\left(\frac{a+b}{2}\right) + f(b)\right) + R \text{ mit } |R| \;\leq\; \frac{(b-a)^5}{2880}\,\left\|f^{(4)}\right\|_\infty.$$

2. Idee (Summation): Unterteile $[a,b]$ in Teilintervalle $[a, a+h], [a+h, a+2h], \ldots, [b-h, b]$, wende in jedem Teil die Trapezregel (oder Simpson-Formel) an.

8.44 Satz: Sei $f \in C^2([a,b])$, $h := \dfrac{b-a}{k}$ mit einem $k \in \mathbb{N}$. Dann gilt

$$\begin{aligned} \int_a^b f \;&=\; T(h) + R \;:=\; \frac{h}{2}\left(f(a) + 2f(a+h) + 2f(a+2h) + \ldots + 2f(b-h) + f(b)\right) + R, \\ |R| \;&\leq\; (b-a)\,\frac{h^2}{12}\,\|f''\|_\infty \end{aligned}$$

(**summierte Trapezregel**).

8.45 Diskussion: **1)** Der Fehler ist von Ordnung $h^2 = \dfrac{(b-a)^2}{k^2}$ (k = Anzahl der Teilintervalle).

2) Wählt man $h = \dfrac{b-a}{2^l}$, so kann man Rechenoperationen beim Übergang von l auf $l+1$ sparen, da die alten Stützstellen beibehalten werden.

3) Man könnte auch Simpson summieren.

3. Idee Romberg-Extrapolation (genial!): Beweise eine genauere Fehlerabschätzung

$$\int_a^b f(x)\,\mathrm{d}x \;=\; T(h) + c_1 h^2 + c_2 h^4 + \ldots + h^{2k} \cdot \rho(h)$$

mit beschränkten $\rho(h)$ für $h \downarrow 0$ (Taylorentwicklung).

$$\Rightarrow \int_a^b f \;=\; T\left(\frac{h}{2}\right) + c_1\left(\frac{h}{2}\right)^2 + c_2\left(\frac{h}{2}\right)^4 + \ldots + \left(\frac{h}{2}\right)^{2k} \cdot \rho\left(\frac{h}{2}\right).$$

Subtraktion der ersten Abschätzung von 4-mal der zweiten eliminiert den 1. Fehlerterm:

$$3\int_a^b f \;=\; 4\,T\left(\frac{h}{2}\right) - T(h) + c_2\left(\frac{4}{16} - 1\right) h^4 + \ldots.$$

Also Verfahren:

$$
\begin{array}{lllllllll}
T_0 & := & T(b-a) & & & & & & \\
T_1 & := & T\left(\dfrac{b-a}{2}\right) & T_1^{(1)} & := & \dfrac{4\,T_1 - T_0}{3} & & & \\
T_2 & := & T\left(\dfrac{b-a}{4}\right) & T_2^{(1)} & := & \dfrac{4\,T_2 - T_1}{3} & T_2^{(2)} & := & \dfrac{4^2\,T_2^{(1)} - T_1^{(1)}}{4^2-1} \\
 & \vdots & & & \vdots & & & \vdots & & \ddots
\end{array}
$$

Dieses Verfahren liefert gute Ergebnisse, wenn f genügend oft differenzierbar ist.

Für ein Beispiel siehe Meyberg/Vachenauer.

9 Gewöhnliche Differentialgleichungen

9.1 Beispiele

1) Heißer Tee mit Temperatur $y(t)$ (t = Zeit), die Außentemperatur $y_{\text{außen}}$ sei konstant.

Physik: Der Verlauf von $y(t)$ ist bestimmt durch:

a) Anfangstemperatur y_0: $y(0) = y_0$,

b) Abfließen der Wärme: $\underbrace{y'(t)}_{\text{Änderung der Temperatur}} = -\underbrace{\alpha\big(y(t) - y_{\text{außen}}\big)}_{\text{abfließende Wärme } (\alpha > 0)}$

Mathematik:

Zu b) Seien $\alpha, y_{\text{außen}} \in \mathbb{R}$ vorgegeben. Die Differentialgleichung

$$y'(t) = -\alpha\big(y(t) - y_{\text{außen}}\big)$$

besitzt die Lösungen $y(t) = c\,\mathrm{e}^{-\alpha t} + y_{\text{außen}}$ ($c \in \mathbb{R}$ beliebig):

Einsetzen:

$$\left.\begin{array}{ll} \text{linke Seite:} & y'(t) = -\alpha c\,\mathrm{e}^{-\alpha t} \\ \text{rechte Seite:} & -\alpha\big(y(t) - y_{\text{außen}}\big) = -\alpha c\,\mathrm{e}^{-\alpha t} \end{array}\right\} \begin{array}{l} \text{stimmt} \\ \text{überein} \end{array}$$

Zu a) $y_0 \overset{!}{=} y(0) = c\,\mathrm{e}^0 + y_{\text{außen}} = c + y_{\text{außen}} \;\Rightarrow\; c = y_0 - y_{\text{außen}}$

Mathematik bestätigt Physik:

Sind $\alpha, y_{\text{außen}}, y_0 \in \mathbb{R}$ vorgegeben, so ist der Temperaturverlauf eindeutig:

$$y(t) \;=\; (y_0 - y_{\text{außen}})\,\mathrm{e}^{-\alpha t} + y_{\text{außen}}.$$

9.1 Beobachtung: Die Differentialgleichung hat unendlich viele Lösungen. Die Anfangsbedingung $y(0) = y_{\text{außen}}$ „wählt“ die richtige aus.

Für eine eindeutige Lösung werden Differentialgleichung und Anfangsbedingung benötigt.

2) Senkrechter Wurf (eines Steins) nach oben: Sei $y(t)$ die Höhe des Steins über dem See. Physik: Beschreibung der Bewegung $y(t)$ durch

a) Startpunkt $\quad y(0) = h$

b) Startgeschwindigkeit $\quad y'(0) = v$

c) Einwirkung der Gewichtskraft: $\quad \underbrace{m\,y''(t)}_{\text{Trägheitskraft}} = \underbrace{-m\,g}_{\text{Gewichtskraft}}$

Mathematik:

Zu c) $y''(t) = -g \;\Leftrightarrow\; y'(t) = -g\,t + c_1 \;\Leftrightarrow\; y(t) = -\frac{1}{2}\,g\,t^2 + c_1\,t + c_2.$

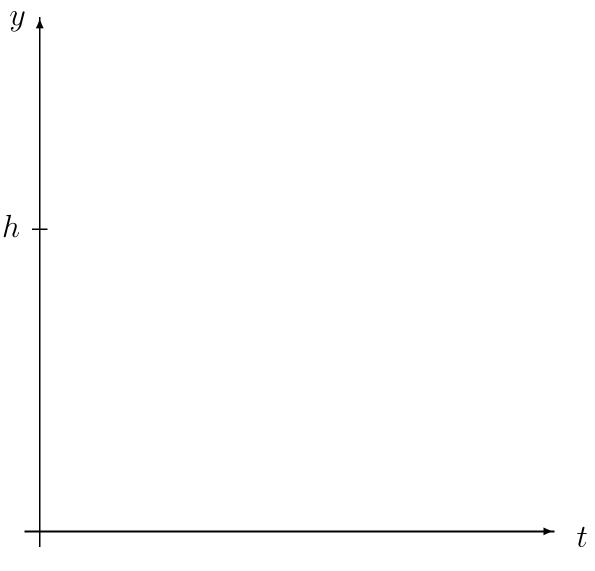
Die Differentialgleichung besitzt unendlich viele Lösungen

$$y(t) \;=\; -\frac{1}{2}\,g\,t^2 + c_1\,t + c_2 \qquad (c_1, c_2 \in \mathbb{R})$$

Zu a) $h \overset{!}{=} y(0) = -\frac{1}{2}\,g \cdot 0 + c_1 \cdot 0 + c_2 \;\Rightarrow\; c_2 = h$

Zu b) $v \overset{!}{=} y'(0) = -g \cdot 0 + c_1 \;\Rightarrow\; c_1 = v.$

$\Rightarrow$ Eindeutige Lösung $y(t) = -\frac{1}{2}\,g\,t^2 + v\,t + h.$



Höchster Punkt:

$$y'(t) = 0 \;\Leftrightarrow\; -g\,t + v = 0 \;\Leftrightarrow\; t = \frac{v}{g}$$

$$\Rightarrow\; y_{\max} \;=\; -\frac{g}{2}\left(\frac{v}{g}\right)^2 + v\,\frac{v}{g} + h \;=\; \frac{v^2}{2g} + h.$$

Eintauchzeitpunkt:

$$0 = y(t) = -\frac{g}{2}\,t^2 + v\,t + h \Rightarrow \ldots$$

In y' eingesetzt: Eintauchgeschwindigkeit.

9.2 Beobachtung: Die Differentialgleichung hat unendlich viele Lösungen. Die Anfangsbedingung $y(0) = h \;\wedge\; y'(0) = v$ „wählt“ die richtige aus.

Für eine eindeutige Lösung werden Differentialgleichung und Anfangsbedingung benötigt.

9.3 Definition: Sei $f : \mathbb{R}^n \to \mathbb{R}$ und $I \subseteq \mathbb{R}$ ein Intervall.

1) Die Gleichung

$$y^{(n)}(x) \;=\; \underbrace{f\big(x, y(x), y'(x), y''(x), \ldots, y^{(n-1)}(x)\big)}_{\text{Term, der von } x, y(x), y'(x), \ldots, y^{(n-1)}(x) \text{ abhängen darf}} \qquad \text{für } x \in I \qquad (*)$$

für die unbekannte Funktion y heißt (explizite) **gewöhnliche Differentialgleichung n-ter Ordnung**. Eine n-mal stetig differenzierbare Funktion y, die diese Gleichung erfüllt, heißt **Lösung**.

2) Seien $y_0, y_1, \ldots, y_{n-1} \in \mathbb{R}$ und $x_0 \in I$ gegeben. Die Bedingung

$$y(x_0) = y_0, \quad y'(x_0) = y_1, \; \ldots, \quad y^{(n-1)}(x_0) = y_{n-1} \qquad (**)$$

heißt **Anfangsbedingung** (n-ter Ordnung).

3) $(*) + (**)$ heißt **Anfangswertproblem** (n-ter Ordnung).

9.4 Bemerkungen: **1)** n-ter Ordnung: Die höchste auftretende Ableitung ist $y^{(n)}$.

2) Die Differentialgleichung heißt gewöhnlich, weil y nur von $x \in \mathbb{R}^1$ abhängt (y hängt von $x \in \mathbb{R}^n$ ab: partielle Differentialgleichung).

3) Abkürzende Schreibweise: $y^{(n)} = f(x, y, y', \ldots, y^{(n-1)})$.

4) Es gibt auch implizite Differentialgleichung, z.B. $(y'' + x)^2 - y = 0$.

5) Es gibt keine allgemeine Theorie zur Lösung von Differentialgleichungen. Im Folgenden: Verfahren für einige spezielle Typen von Differentialgleichungen.

9.5 Beispiele: **1)** $y^{(4)} = \underbrace{(x+y)^2 \cdot y' - y'''}_{=f(x,y,y',y'',y''')}$, $y(1) = 0$, $y'(1) = 1$, $y''(1) = 2$, $y'''(1) = 3$ (Anfangswertproblem 4. Ordnung). Lösung?

2) $y' = 2xy$, $y(0) = -1$: Lösung $y(x) = -\mathrm{e}^{x^2}$.

Beobachtungen:

a) Die Lösung ist auch „links von der Anfangsbedingung“ definiert. Oft wird kein Intervall für die Differentialgleichung angegeben. Bei Konstruktion der Lösung wird das Intervall mitgeliefert.

b) Die Lösung existiert für alle $x \in \mathbb{R}$: **globale Lösung**.

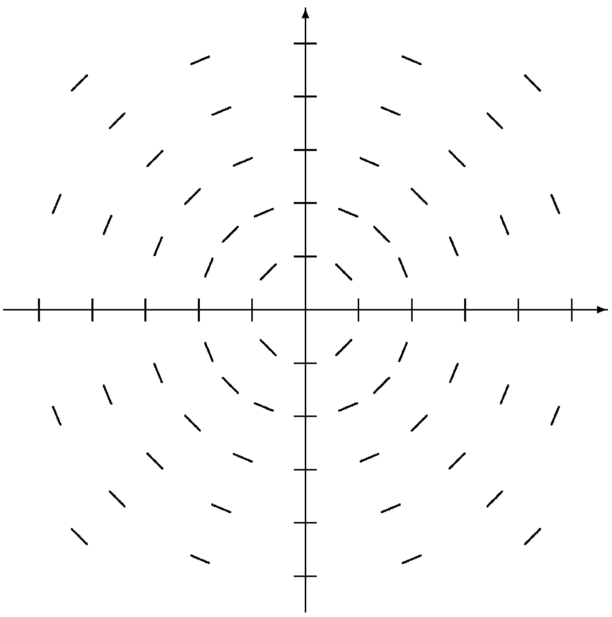
3) $y' = 1 + y^2$, $y\big(\frac{\pi}{4}\big) = 1$: Lösung $y(x) = \tan x$ für $-\dfrac{\pi}{2} < x < \dfrac{\pi}{2}$: **lokale Lösung**.

9.6 Geometrische Veranschaulichung: Eine Lösung der Differentialgleichung

$$y'(x) = f\big(x, y(x)\big)$$

hat im Punkt (x_0, y_0) die Steigung $y' = f(x_0, y_0)$.

Z.B. $y' = -\frac{x}{y}$: Man kann die Lösungssteigungen einzeichen (**Richtungsfeld**) und den Lösungsverlauf „erraten" oder numerisch berechnen ($\longrightarrow$ NumStoch, 3. Semester)



Vermutung:
Die Lösungen sind Halbkreise
$y(x) = \pm\sqrt{r^2 - x^2} = \pm(r^2 - x^2)^{1/2}$
für $-r < x < r$.
Nachrechnen:
$y'(x) = \pm\frac{1}{2}\,\frac{1}{(r^2 - x^2)^{1/2}}\,(-2x) = -\frac{x}{y}$.

9.2 Ein paar Lösungsmethoden

9.2.1 Nur integrieren

$$\begin{aligned} y'(x) &= f(x), \qquad x \in [a, b] \\ \overset{f \text{ stetig}}{\Leftrightarrow} \quad y(x) &= \underbrace{\int_a^x f(t)\,\mathrm{d}t + c}_{\text{dies sind alle Lösungen}} \end{aligned}$$

Durch Anfangsbedingung $y(x_0) = y_0$ ist c eindeutig bestimmt.

9.2.2 Trennung der Variablen

9.7 Definition: Seien $f, g : \mathbb{R} \to \mathbb{R}$. Die Differentialgleichung

$$y' \;=\; f(y) \cdot g(x)$$

heißt **separierbare Differentialgleichung**.

9.8 Beispiele: $y' = \underbrace{2x}_{g(x)} \cdot \underbrace{\cos y}_{f(y)}$ ist separierbar

$y' = 1 + y^2 = \underbrace{1}_{g(x)} \cdot \underbrace{(1+y^2)}_{f(y)}$ ist separierbar

$y' = 2x + \cos y$ ist nicht separierbar

9.9 Lösung durch Trennung der Variablen: **1)** Wo ist $f(y) = 0$?

$f(\eta) = 0 \;\Rightarrow\; y = \text{const} = \eta$ ist Lösung („spezielle Lösung").

2) Sei $y(x) \notin \{\eta : f(\eta) = 0\}$.

$$\begin{aligned} y' &= f(y)\,g(x) \;\Leftrightarrow\; \frac{y'}{f(y)} = g(x) \qquad \text{(Variablen getrennt!)} \\ &\Leftrightarrow\; \underbrace{\int \frac{y'(x)}{f(y(x))}\,\mathrm{d}x}_{} = \int g(x)\,\mathrm{d}x \\ &\overset{u=y(x)}{=} \left[\int \frac{1}{f(u)}\,\mathrm{d}u\right]_{u=y(x)} = [H(u)]_{u=y(x)} + c \quad (H' = \frac{1}{f}) \\ &\Leftrightarrow\; H(y(x)) = \int g(x)\,\mathrm{d}x = G(x) + c \quad (G' = g) \\ &\Leftrightarrow\; y(x) = H^{-1}\big(G(x) + c\big) \end{aligned}$$

(Da $H'(u) = \dfrac{1}{f(u)} \neq 0$, ist H lokal streng monoton, also invertierbar)

Also „Kochrezept" für die separierbare Differentialgleichung $y'(x) = f(y) \cdot g(x)$:

I) $f(\eta) = 0 \;\Rightarrow\; y(x) = \eta$ ist (spezielle) Lösung. **nicht vergessen!**

II) Für $y \notin \{\text{Nullstellen von } f\}$:

$$\begin{aligned} y' &= \frac{\mathrm{d}y}{\mathrm{d}x} = f(y) \cdot g(x) \\ \Leftrightarrow\; \int \frac{\mathrm{d}y}{f(y)} &= \int g(x)\,\mathrm{d}x \end{aligned}$$

Löse entstehende Gleichung nach y auf.

III) Allgemeine Lösung aus 1) und 2).

IV) Freie Konstante bestimmen: Allgemeine Lösung in Anfangsbedingung einsetzen.

9.10 Beispiele: **1)** $y' = \underbrace{-x}_{g(x)} \cdot \underbrace{y^2}_{f(y)}$:

I) Spezielle Lösung $y(x) = 0$.

II) $y \neq 0$: $y(x) = \dfrac{2}{x^2 + c}$, $c \in \mathbb{R}$.

III) Allgemeine Lösung: $y = \dfrac{2}{x^2+c}$, $c \in \mathbb{R}$ oder $y = 0$.

IV) Für die Anfangsbedingung $y(x_0) = y_0$ folgt:

Fall $\boldsymbol{y_0 = 0}$: Gesuchte Lösung ist $y(x) = 0$.

Fall $\boldsymbol{y_0 \neq 0}$: Gesuchte Lösung ist $y(x) = \dfrac{2}{x^2 + \frac{2}{y_0} - x_0^2}$.

Also: Durch jeden Punkt der x, y - Ebene geht genau eine Lösung, d.h. für beliebige $x_0, y_0 \in \mathbb{R}$ besitzt das Anfangswertproblem

$$y' \;=\; -x\,y^2 \quad \wedge \quad y(x_0) = y_0$$

genau eine Lösung.

Manche Lösungen sind global, manche lokal:

$$\begin{aligned} y(1) &= 1 \;\Rightarrow\; y(x) = \frac{2}{x^2+1} \text{ für } x \in \mathbb{R}: \text{ globale Lösung} \\ y(2) &= 1 \;\Rightarrow\; y(x) = \frac{2}{x^2-2} \text{ für } x > \sqrt{2}: \text{ lokale Lösung} \end{aligned}$$

2) $y' = |y-1|^{1/2}$: $f(y) = |y-1|^{1/2}$, $g(x) = 1$.

Allgemeine Lösung:
$$\begin{aligned} y(x) &= 1 + \frac{1}{4}(x+c)^2 \qquad \text{für } x > -c \; (c \in \mathbb{R}) \\ y(x) &= 1 - \frac{1}{4}(x+c)^2 \qquad \text{für } x < -c \; (c \in \mathbb{R}) \\ y(x) &= 1 \qquad \text{für } x \in \mathbb{R} \end{aligned}$$

Achtung: Das Anfangswertproblem $y' = |y-1|^{1/2}$, $y(1) = 2$ besitzt unendlich viele Lösungen: Jede der Funktionen

$$y(x) \;=\; \begin{cases} 1 + \frac{1}{4}\,(x+1)^2 & \text{für } x > -1, \\ 1 & \text{für } -c \le x \le -1, \\ 1 - \frac{1}{4}(x+c)^2 & \text{für } x < -c \end{cases}$$

($c \le 1$ beliebig) ist Lösung. Im Intervall $[1, \infty[$ ist der Lösungsverlauf eindeutig. Dies ist der Bereich, in dem $y(x) \ge 1$ ist.

9.2.3 Die Ähnlichkeitsdifferentialgleichung $\boldsymbol{y' = f\left(\frac{y}{x}\right)}$

Sei y Lösung von
$$y' \;=\; f\Big(\frac{y}{x}\Big).$$

Definiere $u : x \mapsto \dfrac{y(x)}{x}$. Welche Differentialgleichung erfüllt u?

$$\begin{aligned} u' &= \frac{y'}{x} - \frac{y}{x^2} = \frac{1}{x}\Big(y' - \frac{y}{x}\Big) = \frac{1}{x}\Big(f\Big(\frac{y}{x}\Big) - \frac{y}{x}\Big) \\ \Leftrightarrow\; u' &= \frac{1}{x}\,(f(u) - u) \end{aligned}$$

Hurra! Die entstehende Differentialgleichung ist separierbar.

Also Lösungsverfahren für $y' = f\left(\frac{y}{x}\right)$:
- Setze $u := \frac{y}{x}$,
- löse Differentialgleichung $u' = \frac{1}{x}\,(f(u) - u)$,
- Rücksubstitution $y(x) = x \cdot u(x)$.

9.11 Beispiel: $y' = \frac{1}{4} + \left(\frac{y}{x}\right)^2$ (nicht separierbar).

$u(x) := \frac{y}{x} \;\Rightarrow\; u'(x) = \frac{1}{x}\left(\frac{1}{4} + u^2 - u\right)$.

Allgemeine Lösung $u(x) = \frac{1}{2}$ oder $u(x) = \frac{1}{2} + \frac{1}{-\ln|x| + c}$ $(c \in \mathbb{R})$.

Rücksubstitution: $y(x) = \frac{1}{2}x$ oder $y(x) = x\left(\frac{1}{2} + \frac{1}{-\ln|x| + c}\right)$ $(c \in \mathbb{R})$

9.2.4 Die Differentialgleichung $\boldsymbol{y' = f\left(\frac{ax+by+c}{\alpha x + \beta y + \gamma}\right)}$

Fall 1: $\alpha = \beta = 0$, also $y' = f(Ax + By + C)$:

Die Substitution $u(x) := Ax + By(x) + C$ führt auf

$$u' \;=\; A + By' = A + Bf(u) \qquad \text{(separierbar).}$$

Fall 2: $\alpha \neq 0$ oder $\beta \neq 0$

Fall 2a: $\begin{vmatrix} a & b \\ \alpha & \beta \end{vmatrix} = 0 \quad\Rightarrow\quad \exists \lambda \in \mathbb{R} : \begin{pmatrix} a \\ b \end{pmatrix} = \lambda \begin{pmatrix} \alpha \\ \beta \end{pmatrix}$ (Zeilen sind linear abhängig)

Also: $y' = f\left(\frac{\lambda\alpha\, x + \lambda\beta\, y + c}{\alpha x + \beta y + \gamma}\right)$

Substitution $u(x) := \alpha x + \beta y \;\Rightarrow\; u' = \alpha + \beta y' = \alpha + \beta f\left(\frac{\lambda u + c}{u + \gamma}\right)$ (separierbar)

Fall 2b: $\begin{vmatrix} a & b \\ \alpha & \beta \end{vmatrix} \neq 0$: Transformation $\tilde{x} := x + \xi,\; \tilde{y} := y + \eta$

$$\begin{aligned}
\Rightarrow \frac{ax + by + c}{\alpha x + \beta y + \gamma} &= \frac{a\,(\tilde{x} - \xi) + b\,(\tilde{y} - \eta) + c}{\alpha\,(\tilde{x} - \xi) + \beta\,(\tilde{y} - \eta) + \gamma} \\
&= \frac{a\,\tilde{x} + b\,\tilde{y} + c - a\,\xi - b\,\eta}{\alpha\,\tilde{x} + \beta\,\tilde{y} + \gamma - \alpha\,\xi - \beta\,\eta} \\
&= \frac{a\,\tilde{x} + b\,\tilde{y}}{\alpha\,\tilde{x} + \beta\,\tilde{y}} \qquad \text{falls} \quad \begin{matrix} a\xi + b\eta = c \\ \alpha\xi + \beta\eta = \gamma \end{matrix}
\end{aligned}$$

Seien ξ, η Lösung dieses LGS (existiert wegen $\begin{vmatrix} a & b \\ \alpha & \beta \end{vmatrix} \neq 0$).

Neue Differentialgleichung: $\tilde{y}' := \frac{d\tilde{y}}{d\tilde{x}} = \frac{d\tilde{y}}{dx} \cdot \frac{dx}{d\tilde{x}} = y' \cdot 1$

$$\Rightarrow\; \tilde{y}' = f\left(\frac{a\tilde{x} + b\tilde{y}}{\alpha\tilde{x} + \beta\tilde{y}}\right) \overset{\text{falls } \tilde{x} \neq 0}{=} f\left(\frac{a + b\frac{\tilde{y}}{\tilde{x}}}{\alpha + \beta\frac{\tilde{y}}{\tilde{x}}}\right) \quad \text{(Ähnlichkeitsdifferentialgleichung)}$$

9.12 Beispiele: **1)** $y' = -\dfrac{4x+3y-1}{3x+4y+1}$: $\begin{vmatrix} 4 & 3 \\ 3 & 4 \end{vmatrix} = 16-9 = 7 \neq 0$

$$\left.\begin{array}{rcr} 4\,\xi+3\,\eta &=& -1 \\ 3\,\xi+4\,\eta &=& 1 \end{array}\right\} \Rightarrow \ \xi = -1,\ \eta = 1$$

Transformation $\tilde{x} = x-1,\ \tilde{y} = y+1$

$$\tilde{y}' \ = \ -\frac{4\,\tilde{x}+3\,\tilde{y}}{3\,\tilde{x}+4\,\tilde{y}} \overset{\text{falls } \tilde{x}\neq 0}{=} -\frac{4+3\dfrac{\tilde{y}}{\tilde{x}}}{3+4\dfrac{\tilde{y}}{\tilde{x}}}$$

Lösung als Übung.

2) $y' = \dfrac{1}{x+2y+3}$.

$$u := x+2y+3 \quad \Rightarrow \quad u' = 1+\frac{2}{u}.$$

Aber diese Differentialgleichung ist nicht explizit lösbar.

9.3 Theorie

Es stellen sich drei Fragen:

1) Existiert eine Lösung?

2) Ist die Lösung eindeutig?

3) Hängt die Lösung stetig von den Anfangsdaten ab?

9.13 Hauptsatz über lokale Existenz (Picard (1890), Lindelöf (1894)): Gegeben sei das Anfangswertproblem

$$y' \ = \ f(x,y), \qquad y(x_0) = y_0,$$

und es gelte

1) f ist stetig im Rechteck $R := [x_0-a, x_0+a] \times [y_0-b, y_0+b]$ und $M := \max\limits_R |f(x,y)| > 0$,

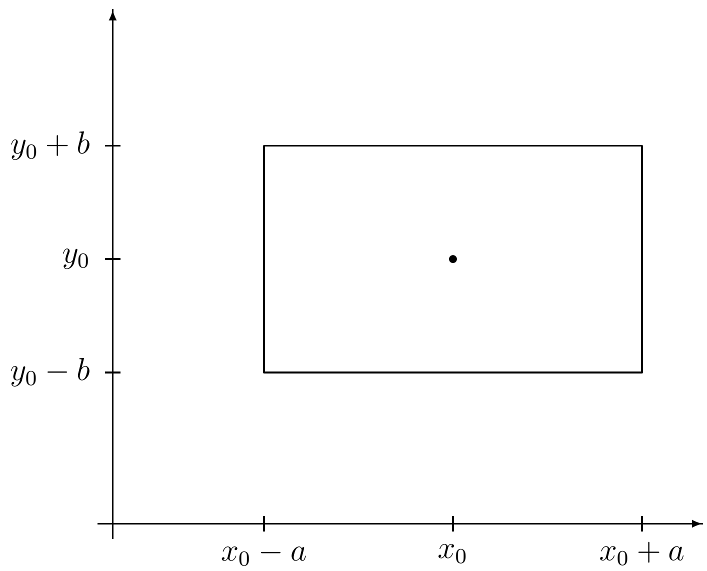
2) f genügt auf R einer **Lipschitz-Bedingung**:

$$\exists L > 0 \ \ \forall (x,y_1),(x,y_2) \in R : |f(x,y_1) - f(x,y_2)| \leq L|y_1 - y_2|.$$

Dann besitzt das Anfangswertproblem eine eindeutige lokale Lösung

$$y : [x_0-h, x_0+h] \to \mathbb{R} \text{ mit } h := \min\left\{a, \frac{b}{M}\right\}.$$

Veranschaulichung:



Wegen $|y'| \leq M$ befindet sich die Lösung in R.

Beweisskizze:

1) Äquivalente Integralgleichung:

$$y \in C\big([x_0 - h, x_0 + h] \to \mathbb{R}\big) \ \wedge \ y(x) = y_0 + \int_{x_0}^{x} f\big(t, y(t)\big)\,\mathrm{d}t \tag{$**$}$$

Behauptung: y Lösung von $(**)$ $\Leftrightarrow$

$$y \in C^1\big([x_0 - h, x_0 + h] \to \mathbb{R}\big) \ \wedge \ y' = f(x, y) \ \wedge \ y(x_0) = y_0 \tag{$*$}$$

"$\Leftarrow$": y Lösung von $(*)$ $\Rightarrow$

$$y(x) - y(x_0) \ = \ \int_{x_0}^{x} y'(t)\,\mathrm{d}t \ \overset{(*)}{=} \ \int_{x_0}^{x} f\big(t, y(t)\big)\,\mathrm{d}t.$$

"$\Rightarrow$": y Lösung von $(**)$ $\Rightarrow$

$$\begin{aligned} y'(x) &= f\big(x, y(x)\big), \quad \text{insbesondere } y \in C^1\big(\ldots\big), \\ y(x_0) &= y_0 + \int_{x_0}^{x_0} \ldots\,\mathrm{d}t \ = \ y_0. \end{aligned}$$

Philosophie: Integralgleichungen sind besser zu behandeln als Differentialgleichungen.

2) Approximation: Definiere Folge (η_n) in $C\big([x_0 - h, x_0 + h]\big)$ durch

$$\begin{aligned} \eta_0(x) &:= y_0, \\ \eta_n(x) &:= y_0 + \int_{x_0}^{x} f\big(t, \eta_{n-1}(t)\big)\,\mathrm{d}t, \qquad (n \in \mathbb{N}). \end{aligned}$$

Zeige: (η_n) ist Cauchy-Folge bezüglich $\|g\|_\infty := \max\limits_{x \in [x_0 - h, x_0 + h]} |f(t)|$.

$\Rightarrow$ gleichmäßige Konvergenz $\eta_n \to y : x \mapsto \lim\limits_{n \to \infty} \eta_n(x)$.

Zeige: Dies ist die gesuchte Lösung.

9.14 Bemerkung: Dies ist Anwendung einer allgemeinen Methode: Ist Lösung y von $y = F(y)$ gesucht (Fixpunkt), versuche ob die rekursiv definierte Folge $y_n := F(y_{n-1})$ mit geeignet gewähltem y_0 gegen den Fixpunkt konvergiert (siehe Banachscher Fixpunktsatz).

9.15 Beispiele: **1)** $y' = -x \cdot y^2 =: f(x,y)$:

$$\begin{aligned} |f(x,y_1) - f(x,y_2)| &= \left|x \cdot (y_1^2 - y_2^2)\right| = |x \cdot (y_1 + y_2) \cdot (y_1 - y_2)| \\ &\leq \underbrace{\left(\max_{x \in [x_0-a, x_0+a]} |x|\right) \left(\max_{y_1, y_2 \in [y_0-b, y_0+b]} |y_1 + y_2|\right)}_{=\max\{|x_0-a|,|x_0+a|\} \cdot 2 \cdot \max\{|y_0-b|,|y_0+b|\} =: L} \cdot |y_1 - y_2|. \end{aligned}$$

Aus Satz: Das Anfangswertproblem $y' = -x \cdot y^2$, $y(x_0) = y_0$ besitzt für jedes $(x_0, y_0) \in G$ genau eine lokale Lösung.

2) $y' = |y-1|^{1/2} =: f(x,y)$. Für $y_1 = 1$ gilt keine Lipschitz-Bedingung:

$$\frac{|f(x,1) - f(x,y_2)|}{|1 - y_2|} = \frac{|y_2 - 1|^{1/2}}{|1 - y_2|} = \frac{1}{|1 - y_2|^{1/2}} \to \infty \qquad \text{für } y_2 \to 1.$$

Dies erklärt, warum das Anfangswertproblem in Beispiel 9.10, Teil 2), unendlich viele Lösungen besitzt.

9.16 Bemerkung: Einerseits typisch Mathematiker: Existenz und Eindeutigkeit klar, aber keine allgemeine Lösungstheorie.

Andererseits: Wenn eine Lösung numerisch berechnet wird, muss man wissen, ob es eine gibt. Wenn sie nicht eindeutig ist, muss man überlegen, was der Computer berechnet.

9.17 Abhängigkeit vom Anfangswert: Es sei $f : \mathbb{R}^2 \to \mathbb{R}$ stetig und genüge der (globalen) Lipschitzbedingung

$$\exists L > 0 \;\; \forall (x,y_1),(x,y_2) \in \mathbb{R}^2 : |f(x,y_1) - f(x,y_2)| \leq L|y_1 - y_2|.$$

Sind $y, \tilde{y}$ zwei Lösungen von $y' = f(x,y)$, so gilt

$$|y(x) - \tilde{y}(x)| \;\leq\; |y(x_0) - \tilde{y}(x_0)| \cdot \mathrm{e}^{L|x-x_0|},$$

d.h. eine Änderung des Anfangswertes $y(x_0)$ pflanzt sich höchstens exponentiell wachsend fort.

9.18 Beispiele: **1)** $y' = y =: f(x,y)$:

$$|f(x,y_1) - f(x,y_2)| = |y_1 - y_2| \;\Rightarrow\; L = 1.$$

Die Lösung y mit $y(x_0) = y_0$ ist gegeben durch $y(x) = y_0 \, \mathrm{e}^{x-x_0}$.

Sind $y, \tilde{y}$ zwei Lösungen mit $y(x_0) = y_0$, $\tilde{y}(x_0) = \tilde{y}_0$, so folgt

$$|y(x) - \tilde{y}(x)| = |y_0 - \tilde{y}_0| \, \mathrm{e}^{x-x_0} = \mathrm{e}^{L(x-x_0)}.$$

Der Abstand wächst exponentiell, wie im Satz angegeben.

2) Es kann auch besser sein: $y' = \dfrac{y}{x} =: f(x,y)$ in $G = \{(x,y) \in \mathbb{R}^2 : x \geq 1\}$:

$$|f(x,y_1) - f(x,y_2)| \;=\; \left|\frac{y_1 - y_2}{x}\right| \;\leq\; \frac{|y_1 - y_2|}{1} \;\Rightarrow\; L = 1.$$

Lösung y mit $y(x_0) = y_0$: $y(x) = \dfrac{y_0}{x_0}\,x$.

$$\Rightarrow\; |y(x) - \tilde{y}(x)| \;=\; \frac{|y_0 - \tilde{y}_0|}{x_0}\,|x|.$$

Hier wächst der Abstand nur linear.

9.4 Systeme von Differentialgleichungen

9.19 Beispiel: Gegeben sei das zeitabhängige Vektorfeld

$$f(t,x_1,x_2) \;=\; \begin{pmatrix} t\,x_1 \\ x_2 \end{pmatrix}.$$

Gesucht: Kurve γ, die für $t = 0$ in $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ startet, und an die das Vektorfeld tangential ist.

Eindeutige Lösung $\gamma : t \mapsto \begin{pmatrix} \mathrm{e}^{t^2/2} \\ \mathrm{e}^{t} \end{pmatrix}$.

9.20 Definition: **1)** Eine Abbildung $f : \mathbb{R}^n \supseteq D \to \mathbb{R}^n$ heißt **Vektorfeld**.

2) Sei $f : \mathbb{R} \times \mathbb{R}^n \supseteq I \times D \to \mathbb{R}^n$ ein zeitabhängiges Vektorfeld. Die Gleichung

$$y'(t) \;=\; f(t, y(t))$$

für die Unbekannte $y : I \supseteq I' \to \mathbb{R}^n$ heißt **System von Differentialgleichungen 1. Ordnung** oder kurz **System 1. Ordnung**.

3) Sei $t_0 \in I$, $y_0 \in D$. Die Bedingung

$$y(t_0) \;=\; y_0$$

heißt **Anfangsbedingung**.

9.21 Picard-Lindelöf: Es sei $f : \mathbb{R} \times \mathbb{R}^n \supseteq I \times D \to \mathbb{R}^n$ ein zeitabhängiges Vektorfeld und $(t_0, y_0) \in I \times D$. Weiter gelte:

1) f ist stetig im Quader $R := [t_0 - a, t_0 + a] \times [y_0 - b, y_0 + b]$ und $M := \max\limits_{R} \|f(t,y)\| > 0$,

2) f genügt auf R einer **Lipschitz-Bedingung**:

$$\exists L > 0 \;\; \forall (t,y_1),(t,y_2) \in R : \|f(t,y_1) - f(t,y_2)\| \leq L\|y_1 - y_2\|.$$

Dann besitzt das Anfangswertproblem

$$y' \;=\; f(t,y), \qquad y(t_0) = y_0,$$

eine eindeutige lokale Lösung.

9.5 Lineare Systeme

Im Folgenden sei $I \subseteq \mathbb{R}$ ein beliebiges Intervall

9.22 Definition: Sei $A : I \to \mathbb{R}^{(n^2)}$ eine $n \times n$-Matrix mit zeitabhängigen Einträgen und $g : I \to \mathbb{R}^n$. Das System

$$y'(t) \;=\; A(t)\,y(t) + g(t) \qquad (\text{oder } y' = A(t)y + g(t)\ )$$

für $y : I \to \mathbb{R}^n$ heißt **lineares System** (von Differentialgleichungen 1. Ordnung).
Ist $g = 0$, so heißt das System **homogen**, sonst **inhomogen**.

9.23 Beispiel: $y'(t) \;=\; \begin{pmatrix} \frac{1}{t} & 1 \\ 0 & \frac{1}{t} \end{pmatrix} y(t) \quad \Leftrightarrow \quad \begin{cases} y_1' = \frac{1}{t}\,y_1 + y_2 \\ y_2' = \frac{1}{t}\,y_2 \end{cases}$

9.24 Satz: Sind A und g stetig auf I und $y_0 \in \mathbb{R}^n$, so besitzt das Anfangswertproblem

$$y' \;=\; A(t)\,y + g(t), \quad y(t_0) = y_0$$

eine eindeutige Lösung $y \in C^1(I \to \mathbb{R}^n)$. Insbesondere: Ist $I = \mathbb{R}$, so ist die Lösung global, d.h. es gilt $y \in C^1(\mathbb{R} \to \mathbb{R}^n)$.

9.5.1 Homogene Systeme

Für $A : I \to \mathbb{R}^{(n^2)}$ betrachte das homogene System

$$y' \;=\; A(t)\,y. \tag{$*$}$$

9.25 Satz: **1)** Sind $y_{[1]}, y_{[2]}$ Lösungen von $(*)$ und $\alpha, \beta \in \mathbb{R}$, so ist auch $y := \alpha \cdot y_{[1]} + \beta \cdot y_{[2]}$ eine Lösung von $(*)$.

2) Ist A stetig auf I, so bildet die Menge aller Lösungen

$$V \;:=\; \left\{ y \in C^1(I \to \mathbb{R}^n) : y' = A(t)\,y \right\}$$

einen Vektorraum der Dimension n.

Eine Basis $\{y_{[1]}, \ldots, y_{[n]}\}$ von V heißt **Fundamentalsystem** zum Differentialgleichungssystem $(*)$.

Beweis: **1)** $y' = \alpha y_{[1]}' + \beta y_{[2]}' = \alpha A\,y_{[1]} + \beta A\,y_{[2]} = A\left(\alpha y_{[1]} + \beta y_{[2]}\right) = A\,y.$

2) Nach 1) bildet V einen Vektorraum (mit der üblichen Addition von Funktionen und skalarer Multiplikation). Sei $t_0 \in I$ fest und

$$T_{t_0} : \mathbb{R}^n \to V : y_0 \mapsto y := \text{ Lösung von } y' = A\,y \;\wedge\; y(t_0) = y_0.$$

Dann ist T_{t_0} sinnvoll definiert, da die Lösung y existiert und eindeutig ist (Satz 9.24).

T_{t_0} ist offensichtlich linear,
injektiv: $y = \widetilde{y} \Rightarrow y_0 = y(t_0) = \widetilde{y}(t_0) = \widetilde{y}_0$
und surjektiv: Ist $y \in V$ gegeben, so setze $y_0 = y(t_0) \Rightarrow y = T_{t_0} y_0$.

Dimensionsformel (3.27):

$$n \;=\; \dim \mathbb{R}^n \;=\; \dim \underbrace{\text{Bild}(T_{t_0})}_{=V} + \dim \underbrace{\text{Kern}(T_{t_0})}_{=\{0\},\text{ da injektiv}} \;=\; \dim V.$$

$\square$

9.26 Beispiel: Das System $y' = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 2 & 1 \end{pmatrix} y$

besitzt die Lösungen $y_{[1]}(t) = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \mathrm{e}^t$, $y_{[2]}(t) = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \mathrm{e}^{-t}$, $y_{[3]}(t) = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \mathrm{e}^{3t}$,

$$\Rightarrow \; V = \Big\{ y = c_1\, y_{[1]} + c_2\, y_{[2]} + c_3\, y_{[3]} : c_j \in \mathbb{R} \Big\}.$$

Die Abbildung T_0 (mit $t_0 = 0$) ist gegeben durch

$$T_0 : \mathbb{R}^n \to V : \begin{pmatrix} \alpha \\ \beta \\ \gamma \end{pmatrix} \mapsto \alpha\, y_{[1]} + \frac{\beta - \gamma}{2}\, y_{[2]} + \frac{\beta + \gamma}{2}\, y_{[3]}.$$

9.27 Folgerung: Seien $y_{[1]}, \ldots, y_{[n]}$ Lösungen von $(*)$. Dann gilt:

$$\begin{aligned}
& \big\{ y_{[1]}, \ldots, y_{[n]} \big\} \text{ linear unabhängig in } V, \\
& \qquad \text{d.h. } \big(c_1 \cdot y_{[1]} + \ldots + c_n y_{[n]} = 0 \;\Rightarrow\; c_1 = \ldots = c_n = 0 \big) \\
\Leftrightarrow \quad & \big\{ y_{[1]}(t_0), \ldots, y_{[n]}(t_0) \big\} \text{ linear unabhängig im } \mathbb{R}^n \text{ für ein } t_0 \in I \\
\overset{t_0 \text{ war beliebig}}{\Leftrightarrow} \quad & \big\{ y_{[1]}(t), \ldots, y_{[n]}(t) \big\} \text{ linear unabhängig im } \mathbb{R}^n \text{ für alle } t \in I
\end{aligned}$$

Beweis: $T_{t_0} : \mathbb{R}^n \to V$ ist Vektorraumisomorphismus, bildet also linear unabhängige Vektoren auf linear unabhängige Vektoren ab.

$\square$

9.28 Satz: Seien $y_{[1]}, \ldots, y_{[n]}$ Lösungen von $(*)$. Dann sind äquivalent:

(i) $\{y_{[1]}, \ldots, y_{[n]}\}$ ist ein Fundamentalsystem.

(ii) $W(t) := \det\big(y_{[1]}(t) \ \ldots \ y_{[n]}(t)\big) \neq 0$ für ein $t \in I$.

(iii) $W(t) := \det\big(y_{[1]}(t) \ \ldots \ y_{[n]}(t)\big) \neq 0$ für alle $t \in I$.

$W(t)$ heißt die **Wronskideterminante**.

9.29 Beispiel: Auf $I :=]0, \infty[$ besitzt

$$y' = \begin{pmatrix} \frac{1}{t} & 1 \\ 0 & \frac{1}{t} \end{pmatrix} y$$

das Fundamentalsystem $\{y_{[1]}, y_{[2]}\} := \left\{ \begin{pmatrix} t^2 \\ t \end{pmatrix}, \begin{pmatrix} t \\ 0 \end{pmatrix} \right\}$, denn

$$W(t) = \begin{vmatrix} t^2 & t \\ t & 0 \end{vmatrix} = -t^2 \neq 0 \text{ auf } I$$

$\Rightarrow$ Allgemeine Lösung: $y(t) \ = \ c_1 \cdot \begin{pmatrix} t^2 \\ t \end{pmatrix} + c_2 \cdot \begin{pmatrix} t \\ 0 \end{pmatrix}$.

9.5.2 Inhomogene Systeme

Für $A : I \to \mathbb{R}^{(n^2)}$, $g : I \to \mathbb{R}^n$ betrachte das inhomogene System

$$y' = A(t)\, y + g(t). \qquad (**)$$

9.30 Satz: **1)** Sei $y_{[1]}$ eine Lösung von $(**)$. Für $y_{[2]} \in C^1(I \to \mathbb{R}^n)$ sind äquivalent:

(i) $y_{[2]}$ ist Lösung von $(**)$.

(ii) $y := y_{[1]} - y_{[2]}$ ist Lösung des zugehörigen homogenen Systems $y' = A(t)\, y$.

2) Ist $y_{[1]}$ eine Lösung von $(**)$ (eine **partikuläre Lösung**), so ist der affine Raum $\tilde{V}$ aller Lösungen von $(**)$ gegeben durch

$$\tilde{V} \ = \ y_{[1]} + V \ = \ \{y_{[1]} + y : y \in C^1(I \to \mathbb{R}^n) \ \wedge \ y' = A(t)\, y\}.$$

Beweis: **1)** (i) $\Rightarrow$ $y' = y'_{[1]} - y'_{[2]} = A\, y_{[1]} + g - (A\, y_{[2]} + g) = A\,\big(y_{[1]} - y_{[2]}\big) = A\, y \ \Rightarrow$ (ii).

(ii)) $\Rightarrow$ $y'_{[2]} = (y_{[1]} - y)' = A\, y_{[1]} - A\, y + g = A\,\big(y_{[1]} - y\big) + g = A\, y_{[2]} + g \ \Rightarrow$ (ii).

2) Folgt direkt aus 1). $\square$

9.31 Wie man eine partkuläre Lösung findet (Variation der Konstanten): Sei $\{y_{[1]}, \ldots, y_{[n]}\}$ ein Fundamentalsystem des homogenen Systems $(*)$ und

$$y := c_1(t) \cdot y_{[1]} + \ldots + c_n(t) \cdot y_{[n]}.$$

Dann sind äquivalent:

(i) y ist Lösung des inhomogenen Systems $(**)$.

(ii) $c' = \begin{pmatrix} c_1' \\ \vdots \\ c_n' \end{pmatrix}$ ist eindeutige Lösung des linearen Gleichungssystems

$$\underbrace{\begin{pmatrix} y_{[1]} & \ldots & y_{[n]} \end{pmatrix}}_{n \times n\text{-Matrix mit Höchstrang}} c'(t) \; = \; g(t).$$

Beweis: Durch Nachrechnen. $\square$

9.32 Bemerkung: Variation der Konstanten funktioniert auch im Fall $n = 1$ (kein System, eine lineare Differentialgleichung).

9.33 Beispiele: **1)** Auf $I =]0, \infty[$: $\quad y' = \begin{pmatrix} \frac{1}{t} & 1 \\ 0 & \frac{1}{t} \end{pmatrix} y + \begin{pmatrix} 2t^2 \\ 3t \end{pmatrix}$

Allgemeine Lösung: $y(t) = \begin{pmatrix} \frac{5}{2} t^3 \\ 3t^2 \end{pmatrix} + c_1 \cdot \begin{pmatrix} t^2 \\ t \end{pmatrix} + c_2 \cdot \begin{pmatrix} t \\ 0 \end{pmatrix}.$

2) $y' = (\sin x) y + x e^{-\cos x}$. Allgemeine Lösung: $y = \dfrac{x^2}{2} e^{-\cos x} + c\, e^{-\cos x}$.

9.6 Lineare Systeme mit konstanten Koeffizienten

9.34 Satz: Sei A eine relle konstante $n \times n$-Matrix. Für jeden Vektor $y_0 \in \mathbb{R}^n$ und jede Anfangszeit $t_0 \in \mathbb{R}$ besitzt das Anfangswertproblem

$$y' \; = \; A\, y, \qquad y(t_0) = y_0$$

die eindeutige globale Lösung

$$y(t) \; = \; e^{(t-t_0)A} y_0 \qquad \text{für } t \in \mathbb{R}.$$

Zur Existenz der Lösung:

$$\mathrm{e}^A = \sum_{j=0}^{\infty} \frac{A^j}{j!} \qquad \text{mit } A^0 = \mathrm{Id},\ \text{insbesondere } \mathrm{e}^0 = \mathrm{Id}$$

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\mathrm{e}^{(t-t_0)A} y_0\right) = \frac{\mathrm{d}}{\mathrm{d}t}\mathrm{e}^{tA}\left(\mathrm{e}^{-t_0 A} y_0\right) = A\,\mathrm{e}^{tA}\left(\mathrm{e}^{-t_0 A} y_0\right) = A\,y(t)$$

9.35 Fundamentalsystem: Sei A eine reelle konstante $n \times n$-Matrix. Ist $\{v_1, \ldots, v_n\} \subset \mathbb{R}^n$ linear unabhängig, so bildet

$$\left\{\mathrm{e}^{tA} \cdot v_1, \mathrm{e}^{tA} \cdot v_2, \ldots, \mathrm{e}^{tA} \cdot v_n\right\}$$

ein Fundamentalsystem für $y' = A\,y$.

Beweis: **1)** $y = \mathrm{e}^{tA} v_k \;\Rightarrow\; y' = A\,y$ siehe letzter Satz.

2) Wronski-Determinante und Satz 9.28

$$W(0) = \det(\mathrm{e}^{0A} v_1, \ldots, \mathrm{e}^{0A} v_n) = \det(v_1, \ldots, v_n) \neq 0.$$

□

9.36 Satz: Sei A eine relle konstante $n \times n$-Matrix.

1) Ist $v \in \mathbb{R}^n$ ein Eigenvektor von A zum Eigenwert λ, so gilt

$$\mathrm{e}^{tA} v = \mathrm{e}^{\lambda t} \cdot v.$$

2) Ist $\{v_1, \ldots, v_n\}$ eine Basis aus reellen Eigenvektoren zu den Eigenwerten $\lambda_1, \ldots, \lambda_n$, so ist

$$\left\{\mathrm{e}^{\lambda_1 t} \cdot v_1, \ldots, \mathrm{e}^{\lambda_n t} \cdot v_n\right\}$$

ein Fundamentalsystem für $y' = A\,y$.

9.37 Beispiel: $y' = \begin{pmatrix} 8 & -2 \\ -2 & 5 \end{pmatrix} y$. Fundamentalsystem: $\left\{\mathrm{e}^{9t} \cdot \begin{pmatrix} 2 \\ -1 \end{pmatrix}, \mathrm{e}^{4t} \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix}\right\}$.

9.38 Komplexe Eigenwerte: **1)** Sei A eine relle konstante $n \times n$-Matrix und

$\{v_1, \ldots, v_n\} \subset \mathbb{C}^n$ eine Basis aus Eigenvektoren zu den Eigenwerten $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$.

Dann ist

$$\left\{\mathrm{e}^{\lambda_1 t} \cdot v_1, \ldots, \mathrm{e}^{\lambda_n t} \cdot v_n\right\}$$

ein komplexes Fundamentalsystem für das relle lineare System $y' = A\,y$.

2) Ist $v \in \mathbb{C}^n$, so sind

$$y(t) := \mathrm{Re}\left(\mathrm{e}^{tA} v\right), \quad y(t) := \mathrm{Im}\left(\mathrm{e}^{tA} v\right)$$

zwei relle Lösungen des Differentialgleichungssystems $y' = Ay$.

9.39 Beispiel: $y' = \begin{pmatrix} 0 & 2 & 0 \\ 0 & 0 & 2 \\ -1 & 1 & 0 \end{pmatrix} y.$

Fundamentalsystem $\{y_{[1]}, y_{[2]}, y_{[3]}\}$:

$$\begin{aligned} y_{[1]}(t) &= \mathrm{e}^{-2t} \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix}, \\ y_{[2]}(t) &= \mathrm{e}^{t} \left(\cos t \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix} - \sin t \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix} \right), \\ y_{[3]}(t) &= \mathrm{e}^{t} \left(\sin t \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix} + \cos t \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix} \right). \end{aligned}$$

Also Lösungsmethode für $y' = A\,y + g(t)$ mit diagonalisierbarer Matrix A:

1) Eigenwerte $\lambda_1, \ldots, \lambda_n$ und Basis aus zugehörigen Eigenvektoren $b_1, \ldots, b_n$ bestimmen.

2) Für relle Eigenwerte λ_j wähle $y(t) := \mathrm{e}^{\lambda_j t} \cdot b_j$ als Element des Fundamentalsystems.

Falls $\lambda_j \in \mathbb{C} \setminus \mathbb{R}$ und $\lambda_k = \overline{\lambda_j}$: Wähle $y(t) = \operatorname{Re}\left(\mathrm{e}^{\lambda_j t} \cdot b_j\right)$ und $y(t) = \operatorname{Im}\left(\mathrm{e}^{\lambda_j t} \cdot b_j\right)$ als Elemente des Fundamentalsystems.

Dies liefert ein Fundamentalsystem $\left\{y_{[1]}, \ldots, y_{[n]}\right\}$ von $y' = A\,y$.

3) Bestimme eine partikuläre Lösung y_{part} von $y' = A\,y + g(t)$ mit Variation der Konstanten.

4) Allgemeine Lösung: $y = y_{\text{part}} + c_1\, y_{[1]} + \ldots + c_n\, y_{[n]}$ $(c_j \in \mathbb{R})$.

9.40 Satz: Sei A eine relle konstante $n \times n$-Matrix und $\{v_1, \ldots, v_k\}$ eine Vektorkette zum Eigenwert λ, d.h. es gilt

$$\begin{aligned} A\,v_1 &= \lambda \cdot v_1 \quad (v_1 \text{ ist Eigenvektor}) \\ A\,v_j &= \lambda \cdot v_j + v_{j-1} \text{ für } j = 2, \ldots, k. \end{aligned}$$

(vgl. Jordansche Normalform). Dann folgt

$$\begin{aligned} \mathrm{e}^{tA} v_1 &= \mathrm{e}^{\lambda t} \cdot v_1, \\ \mathrm{e}^{tA} v_2 &= \mathrm{e}^{\lambda t} \cdot (v_2 + t \cdot v_1), \\ \mathrm{e}^{tA} v_3 &= \mathrm{e}^{\lambda t} \cdot \left(v_3 + t \cdot v_2 + \frac{t^2}{2} \cdot v_1 \right), \\ &\vdots \\ \mathrm{e}^{tA} v_k &= \mathrm{e}^{\lambda t} \cdot \left(\sum_{j=0}^{k-1} \frac{t^j}{j!} \cdot v_{k-j} \right) \end{aligned}$$

9.41 Beispiel: $y' = \begin{pmatrix} 0 & 1 \\ -4 & -4 \end{pmatrix} y$

Fundamentalsystem: $\left\{ \mathrm{e}^{-2t} \cdot \begin{pmatrix} 1 \\ -2 \end{pmatrix}, \mathrm{e}^{-2t} \left(\begin{pmatrix} 1 \\ -1 \end{pmatrix} + t \cdot \begin{pmatrix} 1 \\ -2 \end{pmatrix} \right) \right\}$

9.7 Differentialgleichungen höherer Ordnung und Systeme

9.42 Beispiel:

$$y''' = 2y'' + y' - 2y \tag{$*$}$$

Geniale Idee: Setze $u = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} := \begin{pmatrix} y \\ y' \\ y'' \end{pmatrix}$. Dann erfüllt u die Dgl.

$$u' = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -2 & 1 & 2 \end{pmatrix} u \tag{$**$}$$

Allgemeine Lösung von $(**)$:

$$u(t) = c_1\,\mathrm{e}^t \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} + c_2\,\mathrm{e}^{-t} \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} + c_3\,\mathrm{e}^{2t} \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix}.$$

Die allgemeine Lösung von $(*)$ erhält man aus der ersten Koordinate von u::

$$y(t) = u_1(t) = c_1\,\mathrm{e}^t + c_2\,\mathrm{e}^{-t} + c_3\,\mathrm{e}^{2t}.$$

Allgemeines Vorgehen: Gegeben sei das Anfangswertproblem höherer Ordnung

$$\left.\begin{aligned} y^{(n)} &= f\big(x, y, y', \dots, y^{(n-1)}\big) \\ y(x_0) &= y_0, \quad y'(x_0) = y_1, \ \dots, \ y^{(n-1)}(x_0) = y_{n-1} \end{aligned}\right\} \tag{$*$}$$

mit $n \geq 2$. Setze

$$u(x) = \begin{pmatrix} u_1(x) \\ u_2(x) \\ \vdots \\ u_n(x) \end{pmatrix} := \begin{pmatrix} y \\ y' \\ \vdots \\ y^{(n-1)} \end{pmatrix}$$

$$\Rightarrow\ u' = F(x,u) := \begin{pmatrix} u_2 \\ \vdots \\ u_n \\ f(x, u_1, \dots, u_n) \end{pmatrix}, \qquad u(x_0) = u_0 := \begin{pmatrix} y_0 \\ y_1 \\ \vdots \\ y_{n-1} \end{pmatrix} \tag{$**$}$$

Umgekehrt: Ist u Lösung von $(**)$ und $y := u_1$, dann ist y Lösung von $(*)$.

9.43 Satz: $y \in C^n(I \to \mathbb{R})$ Lösung von $(*)$ $\Leftrightarrow$ $u \in C^1(I \to \mathbb{R}^n)$ Lösung von $(**)$.

Also: Satz von Picard-Lindelöf liefert Eindeutigkeit und lokale Existenz der Lösung von $(**)$ und damit auch der Lösung von $(*)$.

9.8 Lineare Differentialgleichungen höherer Ordnung

9.44 Definition: Seien $a_0, \dots, a_{n-1}, g \in C(I \to \mathbb{R})$. Die Differentialgleichung

$$y^{(n)} + a_{n-1}(t)\,y^{(n-1)} + \dots + a_0(t)\,y = g(t) \tag{1}$$

heißt **lineare Differentialgleichung** (n-ter Ordnung). Ist $g = 0$, so heißt sie **homogen**, sonst **inhomogen**.

9.45 Bemerkungen: **1)** Die Abbildung

$$y \mapsto y^{(n)} + a_{n-1}\, y^{(n-1)} + \ldots + a_0\, y$$

ist linear.

2) Definiert man wie oben $u_1 := y, \ldots, u_n := y^{(n-1)}$, so gilt

$$u' \;=\; \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & \ddots & 1 & \ddots & 0 \\ \vdots & & & \ddots & \vdots \\ 0 & \ldots & & 0 & 1 \\ -a_0(t) & -a_1(t) & \ldots & & -a_{n-1}(t) \end{pmatrix} u + \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ g(t) \end{pmatrix} \tag{2}$$

D.h. u ist Lösung eines linearen Systems.

Beachte: Ist u Lösung, so ist $y := u_1$ Lösung der ursprünglichen Differentialgleichung.

9.46 Homogene Differentialgleichung: **1)** Die Menge aller Lösungen

$$V \;:=\; \left\{ y \in C^n(I \to \mathbb{R}) : y^{(n)} + a_{n-1}(t)\, y^{(n-1)} + \ldots + a_0(t)\, y = 0 \right\}$$

bildet einen Vektorraum der Dimension n. Eine Basis von V heißt **Fundamentalsystem**.

2) n Lösungen $\left\{ y_{[1]}, \ldots, y_{[n]} \right\}$ der homogenen Differentialgleichung bilden genau dann ein Fundamentalsystem, wenn

$$\exists t \in I : W(t) \;:=\; \begin{vmatrix} y_{[1]}(t) & \ldots & y_{[n]}(t) \\ y'_{[1]}(t) & \ldots & y'_{[n]}(t) \\ \vdots & & \vdots \\ y_{[1]}^{(n-1)}(t) & \ldots & y_{[n]}^{(n-1)}(t) \end{vmatrix} \;\neq\; 0$$

($W(t)$ =: **Wronskideterminante**). In diesem Fall gilt $W(t) \neq 0$ für alle $t \in I$.

9.47 Inhomogene Differentialgleichung: **1)** Sei V der Lösungsraum der zugehörigen homogenen Differentialgleichung und y_{part} eine partikuläre Lösung der inhomogenen Differentialgleichung (1). Dann ist

$$\tilde{V} \;:=\; y_{\text{part}} + V$$

der affine Raum aller Lösungen von (1).

2) Variation der Konstanten: Ist $\left\{ y_{[1]}, \ldots, y_{[n]} \right\}$ ein Fundamentalsystem der zugehörigen homogenen Differentialgleichung, so ist

$$y \;:=\; c_1(t)\, y_{[1]} + \ldots + c_n(t)\, y_{[n]}$$

genau dann Lösung von (1), wenn

$$\begin{pmatrix} y_{[1]}(t) & \ldots & y_{[n]}(t) \\ y'_{[1]}(t) & \ldots & y'_{[n]}(t) \\ \vdots & & \vdots \\ y_{[1]}^{(n-1)}(t) & \ldots & y_{[n]}^{(n-1)}(t) \end{pmatrix} \begin{pmatrix} c'_1(t) \\ c'_2(t) \\ \vdots \\ c'_n(t) \end{pmatrix} \;=\; \begin{pmatrix} 0 \\ \vdots \\ 0 \\ g(t) \end{pmatrix}.$$

9.48 Beispiel: $y''' - 2y'' - y' + 2y = \mathrm{e}^t$:

1) Homogene Lösung: $y_{\text{hom}}(x) = c_1\,\mathrm{e}^x + c_2\,\mathrm{e}^{-x} + c_3\,\mathrm{e}^{2x}$ (vergleiche Beispiel am Anfang).

2) Partikuläre Lösung: $y_{\text{part}}(x) = -\dfrac{x}{2}\,\mathrm{e}^x$

3) Allgemeine Lösung: $y(x) = -\dfrac{x}{2}\,\mathrm{e}^x + c_1\,\mathrm{e}^x + c_2\,\mathrm{e}^{-x} + c_3\,\mathrm{e}^{2x}$

9.49 Bemerkung: Bei homogenen Differentialgleichungen der Ordnung $n \geq 2$ gibt es das Prinzip **Reduktion der Ordnung**: Ist $y_{[1]}$ Lösung der Differentialgleichung, so führt der Ansatz $y_{[2]} := c(x)\,y_{[1]}$ eingesetzt auf eine Differentialgleichung der Ordnung $n-1$ für c'. Damit kann aus der Lösung $y_{[1]}$ eine linear unabhängige Lösung $y_{[2]}$ gewonnen werden.

9.50 Beispiel: $y'' - \dfrac{5}{x}\,y' + \dfrac{9}{x^2}\,y = x^2$:

1) Homogen: $y'' - \dfrac{5}{x}\,y' + \dfrac{9}{x^2}\,y = 0$:

- **a)** Geraten: $y(x) = x^3$ ist Lösung.
- **b)** Für weitere Lösung Ansatz $y(x) = c(x)\,x^3 \;\Rightarrow\; y(x) = x^3 \ln|x|$.
- **c)** Also allgemeine Lösung: $y_{\text{hom}}(x) = c_1 \underbrace{x^3}_{=:y_{[1]}(x)} + c_2 \underbrace{x^3 \ln|x|}_{=:y_{[2]}(x)}$,

  $\{y_{[1]}, y_{[2]}\}$ ist Fundamentalsystem: $W(x) \neq 0$.

2) Partikuläre Lösung: Ansatz $y_{\text{part}} = c_1(x)\,y_{[1]} + c_2(x)\,y_{[2]}$

$$\Rightarrow\; y_{\text{part}} = -x^4 \ln|x| + x^4 + x^4 \ln|x| = x^4.$$

3) Allgemeine Lösung: $y(x) = x^4 + c_1\,x^3 + c_2\,x^3 \ln|x| \quad (c_1, c_2 \in \mathbb{R})$.

9.9 Lineare Differentialgleichungen mit konstanten Koeffizienten

Für $a_0, \ldots, a_{n-1} \in \mathbb{R}$ betrachte die homogene Differentialgleichung

$$y^{(n)} + a_{n-1}\,y^{(n-1)} + \ldots + a_0\,y \;=\; 0. \tag{$*$}$$

Für $u := \begin{pmatrix} y \\ y' \\ \vdots \\ y^{(n-1)} \end{pmatrix}$ gilt

$$u' = \underbrace{\begin{pmatrix} 0 & 1 & 0 & \dots & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ \vdots & & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & \ddots & \ddots & 0 \\ 0 & \dots & & & 0 & 1 \\ -a_0 & -a_1 & \dots & & & -a_{n-1} \end{pmatrix}}_{=:A} u \qquad (**)$$

Bekannt: $(**)$ hat Lösungen der Form $u(t) = \mathrm{e}^{\lambda t} \cdot v$. Also Ansatz für Lösung der Differentialgleichung $(*)$: $y(t) = \mathrm{e}^{\lambda t}$. Eingesetzt in $(*)$:

$$\mathrm{e}^{\lambda t}\big(\underbrace{\lambda^n + a_{n-1}\,\lambda^{n-1} + \ldots + a_0}_{=:p(\lambda)}\big) = 0.$$

Das Polynom p heißt **charakteristisches Polynom** und ist bis auf das Vorzeichen gleich dem charakteristischen Polynom der Matrix A. Besitzt p eine Nullstelle λ mit Vielfachheit größer als 1, so ist A nicht diagonalisierbar, und $(**)$ hat Lösungen der Form

$$\mathrm{e}^{\lambda t} \cdot \left(\sum_{j=0}^{k-1} \frac{t^j}{j!} \cdot v_{k-j} \right)$$

Dies liefert für $(*)$ Lösungen der Form $t^j\,\mathrm{e}^{\lambda t}$.

Also Kochrezept für die Differentialgleichung

$$y^{(n)} + a_{n-1}\,y^{(n-1)} + \ldots + a_0\,y = g(t)$$

mit $a_0, \ldots, a_{n-1} \in \mathbb{R}$ und $g \in C(I \to \mathbb{R})$.

1) Homogene DGl: Der Ansatz $y(t) = \mathrm{e}^{\lambda t}$ eingesetzt:

$$\mathrm{e}^{\lambda t}\big(\underbrace{\lambda^n + a_{n-1}\,\lambda^{n-1} + \ldots + a_0}_{=:p(\lambda)}\big) = 0.$$

Konstruktion des Fundamentalsystems $\{y_{[1]}, \ldots, y_{[n]}\}$: Sei $p(\lambda_j) = 0,\ 1 \le j \le n$.

a) Falls $\lambda_j \in \mathbb{R}$ einfache Nullstelle: Wähle $y_{[j]}(t) := \mathrm{e}^{\lambda_j t}$.

b) Falls $\lambda_j, \lambda_{j+1} = \overline{\lambda_j} \in \mathbb{C} \setminus \mathbb{R}$ einfache Nullstellen:

$$y_{[j]}(t) := \mathrm{e}^{\mathrm{Re}\,(\lambda_j)t} \cos\big(\mathrm{Im}\,(\lambda_j)t\big),\ y_{[j+1]}(t) := \mathrm{e}^{\mathrm{Re}\,(\lambda_j)t} \sin\big(\mathrm{Im}\,(\lambda_j)t\big).$$

c) Falls $\lambda_j = \lambda_{j+1} = \ldots = \lambda_{j+k-1} \in \mathbb{R}$ k-fache Nullstelle:

$$y_{[j]}(t) := \mathrm{e}^{\lambda_j t},\ y_{[j+1]}(t) := t\,\mathrm{e}^{\lambda_j t},\ \ldots,\ y_{[j+k-1]}(t) := t^{k-1}\,\mathrm{e}^{\lambda_j t}.$$

d) Analog für komplexe Nullstellen der Ordnung k:

$$t^m\,\mathrm{e}^{\mathrm{Re}\,(\lambda_j)t} \cos\big(\mathrm{Im}\,(\lambda_j)t\big),\ t^m\,\mathrm{e}^{\mathrm{Re}\,(\lambda_j)t} \sin\big(\mathrm{Im}\,(\lambda_j)t\big) \text{ für } 0 \le m \le k-1.$$

2) Partikuläre Lösung über Variation der Konstanten.

3) Allgemeine Lösung = partikuläre L. + allgemeine L. der homogenen Differentialgleichung

9.51 Beispiele: **1)** $y''' - y'' - y' + y = 0$: $p(\lambda) = \lambda^3 - \lambda^2 - \lambda + 1 = (\lambda - 1)^2(\lambda + 1)$

Fundamentalsystem: $y_{[1]}(t) = \mathrm{e}^t,\ y_{[2]}(t) = t\,\mathrm{e}^t,\ y_{[3]}(t) = \mathrm{e}^{-t}$.

Allgemeine Lösung: $y(t) = c_1\,\mathrm{e}^t + c_2 t\,\mathrm{e}^t + c_3\,\mathrm{e}^{-t}$

2) $y^{(4)} + 8y'' + 16y = 0$: $p(\lambda) = (\lambda^2 + 4)^2$:

Fundamentalsystem: $y_{[1]}(t) = \cos(2t),\ y_{[2]}(t) = \sin(2t),\ y_{[3]}(t) = t\cos(2t),\ y_{[4]}(t) = t\sin(2t)$.

3) $y'' - 4y = \mathrm{e}^{2t}$

a) Homogen: $y'' - 4y = 0$:
Fundamentalsystem $y_{[1]}(t) = \mathrm{e}^{2t},\ y_{[2]}(t) = \mathrm{e}^{-2t}$.

b) Partikuläre Lösung: $y = c_1(t)\,\mathrm{e}^{2t} + c_2(t)\,\mathrm{e}^{-2t}$ (Variation der Konstanten)

$$\Rightarrow\ y_{\text{part}} = \frac{1}{4}\,t\,\mathrm{e}^{2t} - \frac{1}{16}\mathrm{e}^{2t}$$

c) Allgemeine Lösung: $y(t) = \dfrac{1}{4}\,t\,\mathrm{e}^{2t} + c_1\,\mathrm{e}^{2t} - c_2\,\mathrm{e}^{-2t}$ $(c_1, c_2 \in \mathbb{R})$.

This is the end!

Inhaltsverzeichnis

Stichwortverzeichnis