

Skript zur Vorlesung  
Höhere Mathematik für  
Bachelorstudiengänge

Prof. Dr. R. Herzog

gehalten im SS2011  
Technische Universität Chemnitz



TECHNISCHE UNIVERSITÄT  
CHEMNITZ

Auszug aus den Studienordnungen zu den Ausbildungszielen der mit dieser Vorlesung angesprochenen Studiengänge:

- **Wirtschaftsingenieurwesen (B):** „Ziel des Studienganges ist eine wirtschafts- und ingenieurwissenschaftliche Ausbildung, die zum einen für qualifizierte Tätigkeiten in **Schnittstellenbereichen zwischen Wirtschaft und Technik** und zum anderen für die Teilnahme an weiterführenden Masterstudiengängen befähigt.“

*mathematikintensive weiterführende Module:* Maschinenbau oder Elektrotechnik

- **Sports Engineering (B):** „Ziel des Studienganges ist es, die Studierenden unter Beachtung fachdidaktischer Gesichtspunkte zur selbstständigen und verantwortungsbewussten **wissenschaftlich-technischen Arbeit auf dem Gebiet der Sportwissenschaft und Sportgerätetechnik** zu qualifizieren.“

*mathematikintensive weiterführende Module:* Technische Mechanik

- **Automobilproduktion (B):** „Ziel des Studienganges ist es, **exzellente und nachgefragte ingenieurwissenschaftliche Fachkräfte für die Automobilindustrie** heranzubilden.“

*mathematikintensive weiterführende Module:* Technische Mechanik

- **Chemie (B):** „Die Ziele des Studienganges sind, die **chemischen Grundlagen inklusive des notwendigen mathematisch-naturwissenschaftlichen Fachwissens** in hinreichender Breite und Tiefe zu vermitteln.“

*mathematikintensive weiterführende Module:* Physik, physikalische Chemie

- **Media Production (B):** „Ziele des Studienganges sind die Berufsbefähigung der Absolventen für den Bereich **Print- und Medientechnik** einerseits und die Vorbereitung auf einen möglichen späteren Masterstudiengang zur Vertiefung oder fachübergreifenden Erweiterung andererseits.“

*mathematikintensive weiterführende Module:* Grundlagen Elektrotechnik; Mechanik und Werkstoffe

- **Technikkommunikation (B):** „Technikkommunikation, verstanden als Kommunikation über Technik und Technikgebrauch, basiert auf einer Mehrfachkompetenz, die Fachwissen als Technikwissen und dessen professionelle Vermittlung durch Kommunikationswissen umfasst. In dieser Sichtweise besteht das Ziel des Bachelorstudienganges Technikkommunikation in der Qualifizierung von Studierenden auf Universitätsniveau für Berufsfelder der **Experten-Nichtexperten-Kommunikation im Spezialbereich der Wissensvermittlung zu Technik und Technikgebrauch**.“

*mathematikintensive weiterführende Module:* Vertiefungsmodul Elektrotechnik oder Maschinenbau

- **Sensorik und kognitive Psychologie (B):** „Im Studium werden Grundkenntnisse auf wichtigen Teilgebieten der Psychologie, der Physik, aber auch der Mathematik und Informatik vermittelt. Die Studierenden erwerben Erfahrungen im **Umgang mit typischen Methoden der experimentellen und der theoretischen Arbeit in den Fachgebieten Physik und Psychologie**. Ein wesentliches Anliegen der Ausbildung ist es, die Fähigkeit zur möglichst selbständigen Einarbeitung in wechselnde Aufgaben zu fördern.“

*mathematikintensive weiterführende Module:* Methodenlehre und Statistik, Simulation naturwissenschaftlicher Prozesse, elektrische Messtechnik, digitale Signalverarbeitung, Computerphysik, nichtlineare Dynamik

- **Informatik für Journalisten (M):** „Ein erfolgreicher Journalist bedarf aber zunehmend nicht nur einer auf Kommunikationsaspekte fixierten Ausbildung, sondern benötigt zunehmend Kenntnisse aus dem Bereich der Informatik, um effizient die unterschiedlichen Technologien und Medien einsetzen zu können.“

*mathematikintensive weiterführende Module:* Informatik I und II, Datenbanken, Grundlagen der Computergeometrie

**Achtung:** Die Einteilung in Kapitel 1 (Höhere Mathematik I.1) und 2 (Höhere Mathematik I.2) muss nicht genau der tatsächlichen Einteilung des Stoffes in der Vorlesung entsprechen.

Dieses Vorlesungsskript orientiert sich zum Teil an früheren Vorlesungen von HSD Dr. Sybille Handrock und Prof. Dr. Hans Josef Pesch.

Fehler und Kommentare bitte an: [roland.herzog@mathematik.tu-chemnitz.de](mailto:roland.herzog@mathematik.tu-chemnitz.de)

Stand: 3. Juli 2011



## Inhaltsverzeichnis

Kapitel 1. Höhere Mathematik I.1	7
1 Elementare Grundlagen	7
1.1 Mathematische Logik	7
1.2 Mengenlehre	12
1.3 Umrechnung von Einheiten	14
1.4 Zahlbereiche	16
2 Lineare Algebra und analytische Geometrie	26
2.1 Vektoren im $\mathbb{R}^n$	26
2.2 Das Skalarprodukt	32
2.3 Matrizen	34
2.4 Lineare Gleichungssysteme	40
2.5 Inverse Matrizen und Determinanten	49
2.6 Analytische Geometrie in der Ebene	55
2.7 Analytische Geometrie im Raum	57
3 Folgen und Reihen	63
3.1 Folgen	63
3.2 Reihen	66
3.3 Folgen und Reihen in der Finanzmathematik	67
4 Funktionen einer Variablen	75
4.1 Polynominterpolation	78
4.2 Grenzwerte und Stetigkeit	79
4.3 Differentialrechnung	83
4.4 Anwendungen der Differentialrechnung	90
4.5 Optimierung (Kurvendiskussion) differenzierbarer Funktionen	93
4.6 Taylorpolynome	98
4.7 Integralrechnung	101
4.8 Vektorwertige Funktionen $\mathbb{R} \rightarrow \mathbb{R}^n$	108
Kapitel 2. Höhere Mathematik I.2	113
5 Differentialgleichungen und Dgl.-Systeme	113
5.1 Die trennbare Differentialgleichung $y'(x) = f(x)g(y)$	114
5.2 Die lineare Differentialgleichung $y'(x) = ay(x) + f(x)$	116
5.3 Eigenwerte und Eigenvektoren	119
5.4 Das lineare Differentialgleichungssystem $\vec{y}'(x) = A\vec{y}(x) + \vec{f}(x)$	122
6 Lineare Optimierung	130
6.1 Einführung	130
6.2 Grafische Lösung	131
6.3 Aufgaben in Normalform	133
6.4 Das Simplex-Verfahren	136

6.5	Einige Besonderheiten	146
7	Funktionen mehrerer Variabler	148
7.1	Funktionen $\mathbb{R}^m \rightarrow \mathbb{R}$	148
7.2	Anwendung: Lineare Fehlerrechnung	157
7.3	Optimierung differenzierbarer Funktionen	158
7.4	Anwendung: Ausgleichsrechnung	161
	Literaturverzeichnis	165
	Index	167
	Abbildungsnachweis	173

## KAPITEL 1

### Höhere Mathematik I.1

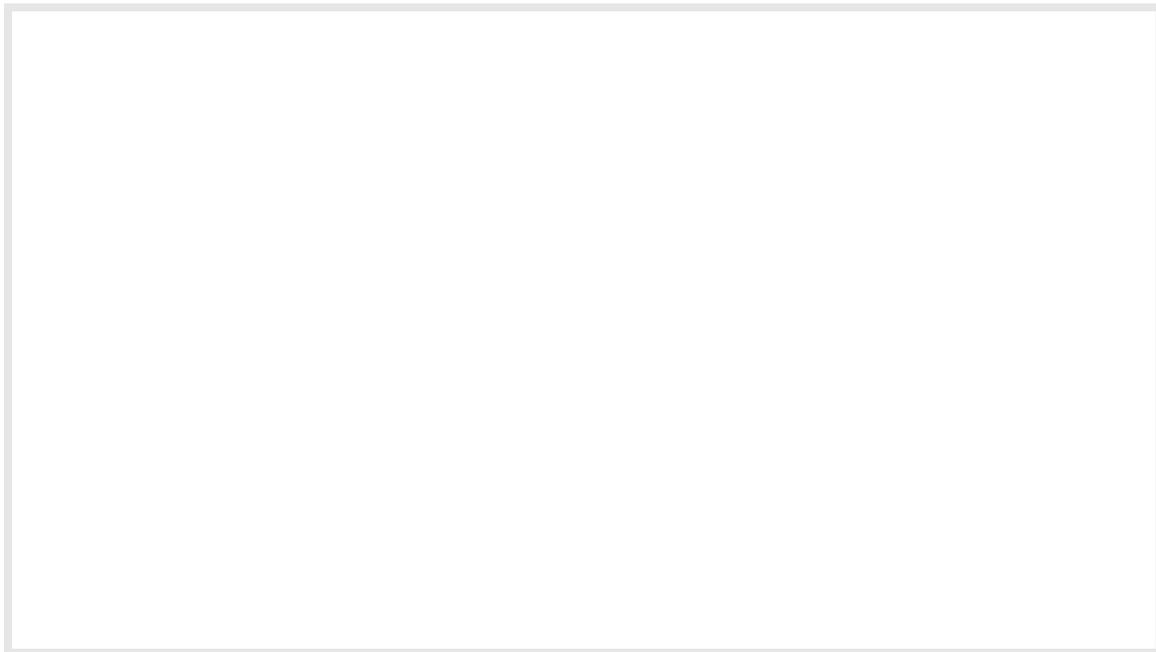
#### § 1 Elementare Grundlagen

##### § 1.1 Mathematische Logik

###### Definition 1.1 (Aussage)

Unter einer **Aussage** verstehen wir einen Satz, der entweder wahr (W) oder falsch (F) ist (Prinzip der Zweiwertigkeit). W (auch:  $\top$ , true oder 1) und F (auch:  $\perp$ , false oder 0) heißen **Wahrheitswerte**.  $\diamond$

###### Beispiel 1.2 (für Aussagen und Nicht-Aussagen)



Aus Aussagen (oft mit  $p, q$  etc. bezeichnet) kann man durch Verknüpfungen (Junktoren) neue Aussagen konstruieren. Diese sind, abhängig von den Wahrheitswerten von  $p$  und  $q$ , wiederum entweder wahr oder falsch.

###### Definition 1.3 (Junktoren)

Wir definieren folgende Junktoren:

(a) **Negation**  $\neg$

Die Operation  $\neg p$  (nicht  $p$ ) heißt **Negation**.  $\neg p$  ist wahr, wenn  $p$  falsch ist, und falsch, wenn  $p$  wahr ist.

$p$	$\neg p$
W	F
F	W

(b) **Alternative**  $\vee$  (**oder, nicht ausschließend**)

Die Aussage  $p \vee q$  ist wahr, wenn mindestens eine der Aussagen  $p$  und  $q$  wahr sind, ansonsten falsch.

$p$	$q$	$p \vee q$
W	W	W
W	F	W
F	W	W
F	F	F

(c) **Konjunktion**  $\wedge$  (**und**)

Die Aussage  $p \wedge q$  ist dann wahr, wenn  $p$  und  $q$  beide wahr sind, ansonsten falsch.

$p$	$q$	$p \wedge q$
W	W	W
W	F	F
F	W	F
F	F	F

(d) **Implikation**  $\Rightarrow$  (**führt zu, wenn ... dann**)

$p \Rightarrow q$  ist definiert als folgende Aussage: „Wann immer  $p$  (die **Prämisse**) wahr ist, ist auch  $q$  (die **Konklusion**) wahr.“ Man sagt auch: „ $p$  ist eine **hinreichende Bedingung** für  $q$ “ und „ $q$  ist eine **notwendige Bedingung** für  $p$ “.

$p$	$q$	$p \Rightarrow q$
W	W	W
W	F	F
F	W	W
F	F	W

(e) **Äquivalenz**  $\Leftrightarrow$  (**dann und nur dann, genau dann wenn**)

Die Aussage  $p \Leftrightarrow q$  ist wahr, wenn entweder  $p$  und  $q$  beide wahr oder beide falsch sind, ansonsten falsch.

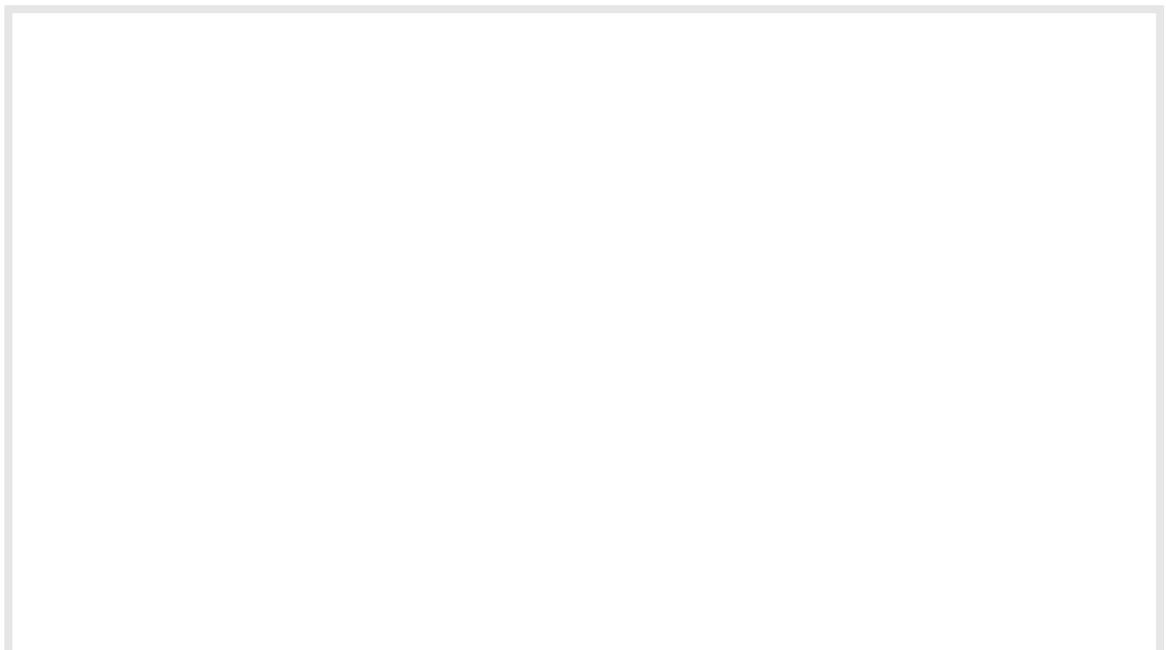
$p$	$q$	$p \Leftrightarrow q$
W	W	W
W	F	F
F	W	F
F	F	W

◇

**Beispiel 1.4** (zur Konstruktion neuer Aussagen)

$p$ : In einem Land  $A$  besitzt das Unternehmen  $U$  einen Marktanteil von über 30%.

$q$ : In einem Land  $B$  besitzt das Unternehmen  $U$  einen Marktanteil von über 30%.



**Bemerkung 1.5** (zur Implikation)

Umgangssprachlich wird „wenn . . . dann“ häufig kausal oder temporal gebraucht. Die oben definierte Implikation in der Logik „ $p \Rightarrow q$ “ drückt aber keinen tatsächlichen Zusammenhang aus, sondern bildet aus den Aussagen  $p$  und  $q$  eine neue Aussage, die wahr oder falsch sein kann.

Folgende Aussagen sind wahr (vgl. Wahrheitstabelle):

- |   |   |                    |
|---|---|--------------------|
| • Wenn Wien eine Stadt ist, ist Regen nass.<br>( $W \Rightarrow W$ )  | } | ex quodlibet verum |
| • Wenn Wien ein Dorf ist, ist Regen nass.<br>( $F \Rightarrow W$ )    |   |                    |
| • Wenn Wien ein Dorf ist, ist Regen nass.<br>( $F \Rightarrow W$ )    | } | ex falso quodlibet |
| • Wenn Wien ein Dorf ist, ist Regen trocken.<br>( $F \Rightarrow F$ ) |   |                    |

◇

Beginnend mit einer falschen Aussage kann man also alles Mögliche schlussfolgern. Dies ist natürlich dann ohne praktische Bedeutung. Darauf muss man beim Beweisen achten:

**Beispiel 1.6** (ein unzulässiger Beweis)

Junktoren unterliegen einer Rangfolge:

- (a)  $\neg$  bindet stärker als
- (b)  $\wedge$  bindet stärker als
- (c)  $\vee$  bindet stärker als
- (d)  $\Rightarrow$  bindet stärker als
- (e)  $\Leftrightarrow$

Durch Klammerung kann man andere Rangfolgen erreichen.

**Beispiel 1.7** (zur Rangfolge)

- (a)  $p \vee q \wedge r$  ist dasselbe wie  $p \vee (q \wedge r)$ , aber nicht dasselbe wie  $(p \vee q) \wedge r$ .
- (b)  $p \wedge q \Rightarrow r$  ist dasselbe wie  $(p \wedge q) \Rightarrow r$ , aber nicht dasselbe wie  $p \wedge (q \Rightarrow r)$ .

◇

**Satz 1.8**

Für beliebige Aussagen  $p$ ,  $q$  und  $r$  sind (unabhängig von ihren Wahrheitswerten!) folgende Aussagen immer wahr:

- (a)  $\neg(\neg p) \Leftrightarrow p$  (doppelte Negation)
- (b)  $(p \Rightarrow q) \Leftrightarrow \neg p \vee q$
- (c)  $((p \Rightarrow q) \wedge (q \Rightarrow r)) \Rightarrow (p \Rightarrow r)$  (Transitivität der Implikation)
- (d)  $(p \Rightarrow q) \Leftrightarrow (\neg q \Rightarrow \neg p)$  (Kontraposition)
- (e)  $\neg(p \wedge q) \Leftrightarrow \neg p \vee \neg q$   
 $\neg(p \vee q) \Leftrightarrow \neg p \wedge \neg q$  (De Morgansche Regeln)
- (f)  $p \wedge (q \vee r) \Leftrightarrow (p \wedge q) \vee (p \wedge r)$   
 $p \vee (q \wedge r) \Leftrightarrow (p \vee q) \wedge (p \vee r)$  (Distributivgesetze)

◇

*Beweis:* durch Wahrheits(wert)tabelle(n), z.B. für die erste De Morgansche Regel:

$p$	$q$	$p \wedge q$	$\neg(p \wedge q)$	$\neg p$	$\neg q$	$\neg p \vee \neg q$
W	W	W	F	F	F	F
W	F	F	W	F	W	W
F	W	F	W	W	F	W
F	F	F	W	W	W	W

□

Die Kontraposition wird beim indirekten Beweis ausgenutzt: Statt  $p \Rightarrow q$  direkt nachzuweisen, zeigt man die äquivalente Aussage  $\neg q \Rightarrow \neg p$ .

**Beispiel 1.9 (logisches Schließen)**

Bei einer Havarie kommen drei Bauteile  $A$ ,  $B$ ,  $C$  als Ursache (einzeln oder mehrere) in Frage. Die Gutachter kommen zu folgenden Aussagen:

- $p$ : Mindestens eines der Bauteile  $A$ ,  $B$ ,  $C$  verursachte die Havarie.
- $q$ : Falls  $A$  und  $B$  nicht beide unter den Verursachern waren, dann trifft  $C$  keine Schuld.
- $r$ : Ist  $A$  ein Verursacher oder ist  $C$  kein Verursacher, dann ist  $B$  kein Verursacher.

$A$	$B$	$C$	$A \vee B \vee C$	$(A \wedge B) \vee \neg C$	$(\neg A \wedge C) \vee \neg B$
W	W	W			
W	W	F			
W	F	W			
W	F	F			
F	W	W			
F	W	F			
F	F	W			
F	F	F			

Ende 1. V

14.10.2010

## § 1.2 Mengenlehre

**Definition 1.10** (Cantor (1895))

„Unter einer Menge verstehen wir jede Zusammenfassung von bestimmten wohlunterschiedenen Objekten unserer Anschauung oder unseres Denkens zu einem Ganzen.“  $\diamond$

Wir bezeichnen Mengen mit Großbuchstaben und ihre Elemente (Objekte) mit Kleinbuchstaben. Man schreibt  $a \in A$  (bzw.  $a \notin A$ ), falls  $a$  ein (kein) Element der Menge  $A$  ist. Die Beschreibung von Mengen erfolgt durch

- Aufzählen (endlicher Mengen):

$$A = \{1, 5, 99\}$$

- Angabe einer charakteristischen Eigenschaft der Elemente:

$$G = \{2n : n \in \mathbb{Z}\}$$

**Lies:**  $G$  ist die Menge aller Zweifachen der ganzen Zahlen (also die Menge der geraden ganzen Zahlen).

$$U = \{2n + 1 : n \in \mathbb{Z}\}$$

**Lies:**  $U$  ist die Menge aller ungeraden ganzen Zahlen.

$$A = \{\frac{1}{n} : n \in \mathbb{N}\}$$

**Lies:**  $A$  ist die Menge aller Kehrwerte von natürlichen Zahlen.

$$P = \{p \in \mathbb{N} : p \text{ ist Primzahl}\}$$

**Lies:**  $P$  ist die Menge aller Primzahlen.

$$X = \{x \in \mathbb{R} : x^2 \leq 4\}$$

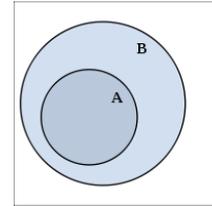
**Lies:**  $X$  ist die Menge aller reellen Zahlen  $x$ , für die gilt:  $x^2 \leq 4$ .

Die **leere Menge**  $\emptyset$  enthält kein Element. (Daher gilt die Aussage: „Wenn  $x$  ein Element der leeren Menge ist, dann ist es rot gestreift.“, kurz: „Alle Elemente der leeren Menge sind rot gestreift.“)

**Definition 1.11** (Relationen zwischen Mengen)

- (a) Eine Menge  $A$  heißt eine **Teilmenge** von  $B$  genau dann, wenn jedes Element von  $A$  auch ein Element von  $B$  ist, wenn also gilt:  $x \in A \Rightarrow x \in B$ . In dem Fall heißt  $B$  auch eine **Obermenge** von  $A$ .

$$A \subseteq B, \quad B \supseteq A, \quad \text{manchmal auch } A \subset B, \quad B \supset A.$$



- (b) Falls  $A \subseteq B$  gilt, aber nicht  $A = B$ , so heißt  $A$  eine **echte Teilmenge** von  $B$  und  $B$  eine **echte Obermenge** von  $A$ .

$$A \subset B, \quad B \supset A, \quad \text{manchmal auch } A \subsetneq B, \quad B \supsetneq A.$$

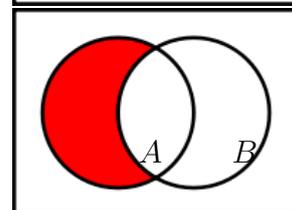
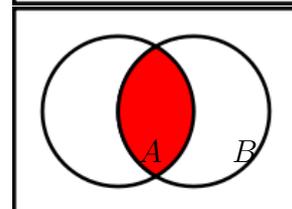
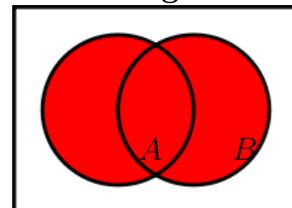
- (c) Für Mengen  $A$  und  $B$  definiert man

die **Vereinigung**  $A \cup B := \{x : x \in A \vee x \in B\}$ ,

den **Durchschnitt**  $A \cap B := \{x : x \in A \wedge x \in B\}$ ,

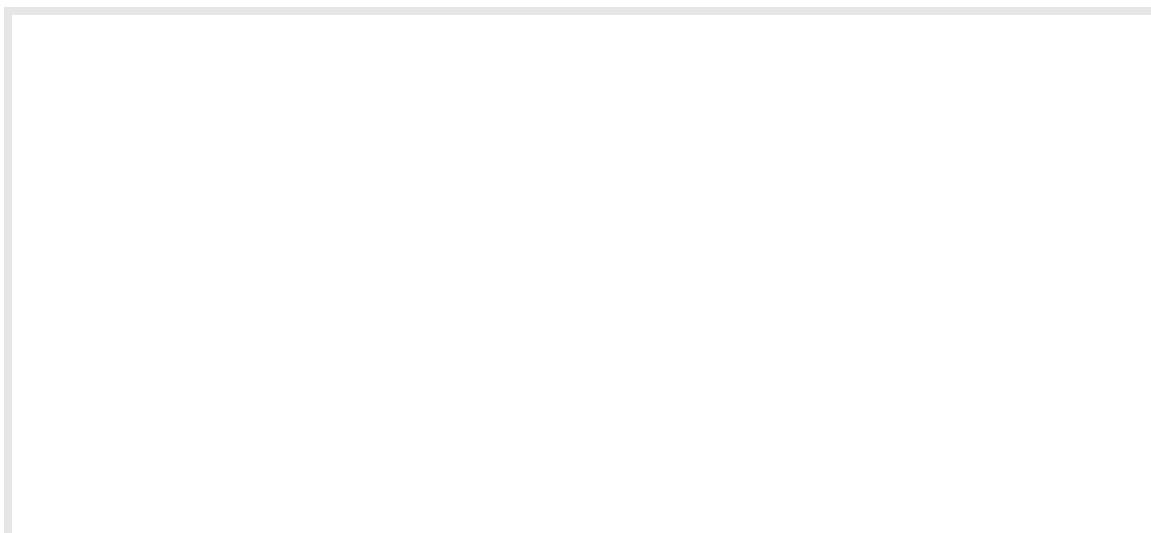
das **Komplement**  $A \setminus B := \{x : x \in A \wedge x \notin B\}$ ,  
auch „A ohne B“.

**Venn-Diagramme:**



- (d) Falls  $A \cap B = \emptyset$  gilt, so heißen die Mengen  $A$  und  $B$  **disjunkt**. ◇

**Beispiel 1.12** (Mengenrelationen)



**Satz 1.13**

Für Mengen  $A$ ,  $B$  und  $C$  gelten folgende Aussagen:

- (a)  $A \cup B = B \cup A$   
 $A \cap B = B \cap A$  (Kommutativität)
- (b)  $(A \cup B) \cup C = A \cup (B \cup C)$   
 $(A \cap B) \cap C = A \cap (B \cap C)$  (Assoziativität)
- (c)  $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$   
 $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$  (Distributivgesetz)
- (d)  $(A \subseteq B \wedge B \subseteq C) \Rightarrow A \subseteq C$
- (e)  $\emptyset \subseteq A$ ,  $A \cup \emptyset = A$ ,  $A \cap \emptyset = \emptyset$ ,  $A \setminus A = \emptyset$ ,  $\emptyset \setminus A = \emptyset$ ,  $A \setminus \emptyset = A$
- (f)  $(A \cup B) \setminus C = (A \setminus C) \cup (B \setminus C)$   
 $(A \cap B) \setminus C = (A \setminus C) \cap (B \setminus C)$  (Distributivgesetz)  $\diamond$

*Beweis:* durch Wahrheitstabellen für die Aussagen  $x \in A$ ,  $x \in B$  etc., durch **Venn-Diagramme** oder durch äquivalente Umformungen, z.B. im Fall (f<sub>1</sub>):

$$\begin{aligned} x \in (A \cup B) \setminus C &\Leftrightarrow (x \in A \vee x \in B) \wedge x \notin C \\ &\Leftrightarrow (x \in A \wedge x \notin C) \vee (x \in B \wedge x \notin C) \quad \text{nach Satz 1.8 (f)} \\ &\Leftrightarrow x \in (A \setminus C) \cup (B \setminus C). \end{aligned}$$

□

**Definition 1.14 (Kartesisches Produkt)**

Für zwei Mengen  $A$  und  $B$  heißt die Menge der (geordneten) Paare

$$A \times B = \{(a, b) : a \in A, b \in B\}$$

das **kartesische Produkt** von  $A$  und  $B$ .  $\diamond$

**Beispiel 1.15**

Sei  $A := \{\text{Kreuz, Pik, Herz, Karo}\}$  und  $B := \{7, 8, 9, 10, \text{Bube, Dame, König, As}\}$ . Dann hat

$$A \times B = \{(\text{Kreuz}, 7), (\text{Kreuz}, 8), \dots, (\text{Karo}, \text{As})\}$$

32 Elemente (wie ein Skatblatt).  $\diamond$

**§ 1.3 Umrechnung von Einheiten**

Viele physikalische Größen bestehen aus einer Maßzahl und einer Einheit, z.B. 3 m (Meter) oder 10 kg (Kilogramm). Es gibt im **SI-System**<sup>1</sup> (internationales Größensystem) sieben Basiseinheiten, nämlich

<sup>1</sup>Système International d'Unités

Dimension	Einheit	Einheitenzeichen
Länge	Meter	m
Masse	Kilogramm	kg
Zeit	Sekunde	s
Stromstärke	Ampere	A
Temperatur	Kelvin	K
Stoffmenge	Mol	mol
Lichtstärke	Candela	cd

Alle anderen Einheiten werden daraus abgeleitet, z.B.

$$\begin{aligned} \text{Newton (für Kräfte)} & \quad N = \frac{\text{kg m}}{\text{s}^2}, \\ \text{Joule (für Energie bzw. Arbeit)} & \quad J = N \text{ m} = \frac{\text{kg m}^2}{\text{s}^2}, \\ \text{Watt (für Leistung)} & \quad W = \frac{J}{\text{s}} = \frac{\text{kg m}^2}{\text{s}^3}, \\ \text{Volt (für elektrische Spannung)} & \quad V = \frac{W}{A} = \frac{\text{kg m}^2}{\text{A s}^3}. \end{aligned}$$

Für große und kleine Maßzahlen verwendet u.a. man die Vorsilben

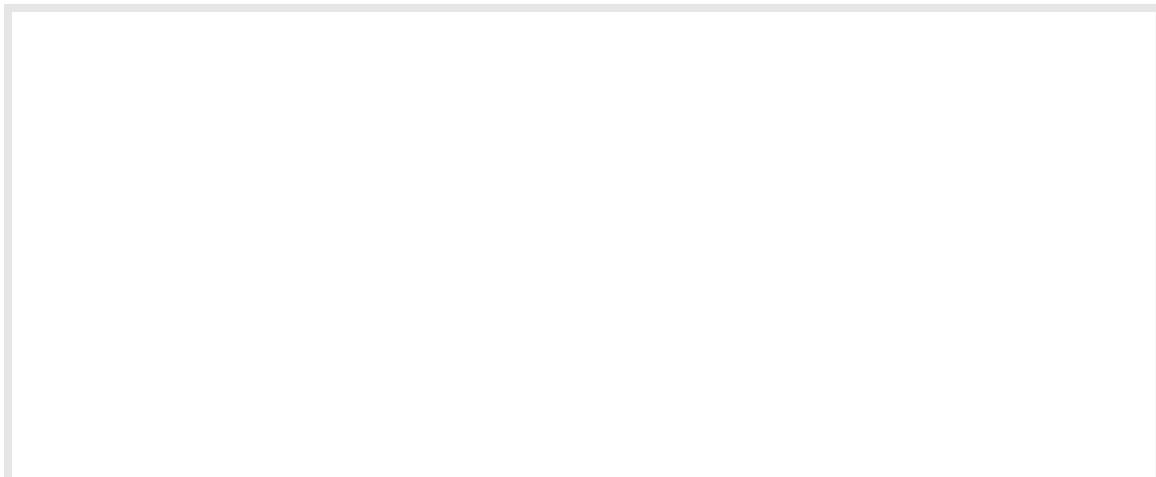
Deka	$10^1$	Dezi	$10^{-1}$
Hekto	$10^2$	Zenti	$10^{-2}$
Kilo	$10^3$	Milli	$10^{-3}$
Mega	$10^6$	Mikro	$10^{-6}$
Giga	$10^9$	Nano	$10^{-9}$ ,

z.B. sind 3,5 Megawatt =  $3,5 \cdot 10^6$  W. Unter Ausnutzung von Beziehungen wie

$$1 \text{ Meile} = 1,609 \text{ km}, \quad 3600 \text{ s} = 1 \text{ h (Stunde)},$$

kann man Einheiten ineinander umrechnen:

### Beispiel 1.16 (Umrechnung von Einheiten)



Für die Umrechnung von Grad- und Bogenmaß legen wir fest:

$$\pi = 180^\circ, \quad \text{also } 1^\circ = \frac{\pi}{180} = 0,0175$$

(Grad ist keine Einheit, sondern eine Zahl!) Daraus ergeben sich folgende wichtige Werte:

$0^\circ$	$45^\circ$	$90^\circ$	$135^\circ$	$180^\circ$	$225^\circ$	$270^\circ$	$315^\circ$	$360^\circ$
0	$\frac{\pi}{4}$	$\frac{\pi}{2}$	$\frac{3\pi}{4}$	$\pi$	$\frac{5\pi}{4}$	$\frac{3\pi}{2}$	$\frac{7\pi}{4}$	$2\pi$

## § 1.4 Zahlbereiche

Wir arbeiten in der Regel mit der Menge der **reellen Zahlen**  $\mathbb{R}$ . Diese kann man verstehen als Zahlen mit einer (möglicherweise unendlich langen) Dezimalbruchentwicklung, z.B.  $1,5563$  oder  $\pi = 3,141\,592\,653\dots$ . Wichtige Teilmengen von  $\mathbb{R}$  sind

- (a) die **natürlichen Zahlen**  $\mathbb{N}_0 := \{0, 1, 2, 3, \dots\}$  bzw.  $\mathbb{N} := \{1, 2, 3, \dots\}$ ,
- (b) die **ganzen Zahlen**  $\mathbb{Z} := \{0, \pm 1, \pm 2, \pm 3, \dots\}$ ,
- (c) die **rationalen Zahlen**  $\mathbb{Q} := \{\frac{m}{n} : m, n \in \mathbb{Z}, n \neq 0\}$ , z.B.  $\frac{3}{4}$  oder  $\frac{-11}{2008}$ .

Die rationalen Zahlen besitzen immer eine endliche oder periodisch unendliche Dezimalbruchentwicklung, z.B.  $3/4 = 0,75$ ,  $1/3 = 0,\bar{3}$  und  $1/22 = 0,04\bar{5}$ . Es gilt

$$\mathbb{N} \subsetneq \mathbb{N}_0 \subsetneq \mathbb{Z} \subsetneq \mathbb{Q} \subsetneq \mathbb{R}.$$

Alle diese Zahlbereiche können auf dem Zahlenstrahl dargestellt werden, der durch  $\mathbb{R}$  lückenlos ausgefüllt ist.

### § 1.4.1 Die reellen Zahlen $\mathbb{R}$

In  $\mathbb{R}$  sind die vier Grundrechenarten definiert, d.h., für  $a, b \in \mathbb{R}$  sind

$$a + b \in \mathbb{R}, \quad a - b \in \mathbb{R}, \quad ab \in \mathbb{R}, \quad \text{und, falls } b \neq 0, \quad \frac{a}{b} \in \mathbb{R}.$$

Reelle Zahlen sind stets vergleichbar, d.h., für  $a, b \in \mathbb{R}$  gilt

$$\text{entweder } a < b \quad \text{oder } a = b \quad \text{oder } a > b.$$

Die Schreibweise  $a \leq b$  bedeutet „ $a$  ist kleiner oder gleich  $b$ “, gleichbedeutend ist  $b \geq a$  oder „ $b$  ist größer oder gleich  $a$ “.

**Definition 1.17 (Intervalle)**

Für  $a, b \in \mathbb{R}$  und  $a < b$  definiert man folgende **Intervalle**:

(a) **endliche Intervalle**

$$[a, b] = \{x \in \mathbb{R} : a \leq x \leq b\} \quad (\text{abgeschlossenes Intervall, Endpunkte dabei})$$

$$(a, b) = \{x \in \mathbb{R} : a < x < b\} \quad (\text{offenes Intervall, Endpunkte nicht dabei})$$

$$[a, b) = \{x \in \mathbb{R} : a \leq x < b\} \quad (\text{rechtsoffenes Intervall})$$

$$(a, b] = \{x \in \mathbb{R} : a < x \leq b\} \quad (\text{links offenes Intervall})$$

(b) **einseitig unendliche Intervalle**

$$(-\infty, b] = \{x \in \mathbb{R} : x \leq b\}$$

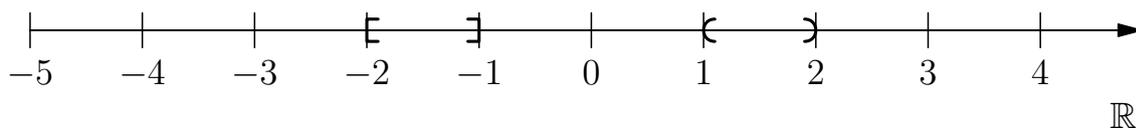
$$(-\infty, b) = \{x \in \mathbb{R} : x < b\}$$

$$[a, \infty) = \{x \in \mathbb{R} : a \leq x\}$$

$$(a, \infty) = \{x \in \mathbb{R} : a < x\}$$

(c)  $(-\infty, \infty) = \mathbb{R}$ .

Veranschaulichung der Intervalle  $[-2, -1]$  und  $(1, 2)$  am **Zahlenstrahl**:



◇

Ende 2. V

21.10.2010

Ungleichungen spielen im täglichen Leben eine wichtige Rolle: „Ich komme spätestens um 17 Uhr zu Dir.“ oder „Ich verkaufe meine Aktien, wenn der Kurs über €32 steigt.“

**Satz 1.18 (Rechenregeln für Ungleichungen)**

Für reelle Zahlen  $a, b, x, y$  gelten folgende Rechenregeln:

$$(a) \quad x \leq y \text{ und } a \leq b \quad \Rightarrow \quad x + a \leq y + b$$

$$x < y \text{ und } a < b \quad \Rightarrow \quad x + a < y + b$$

(Gleichgerichtete Ungleichungen darf man addieren.)

$$(b) \quad x + a \leq y \quad \Leftrightarrow \quad x \leq y - a$$

$$x + a < y \quad \Leftrightarrow \quad x < y - a$$

$$(c) \quad x \leq y \text{ und } 0 \leq a \quad \Rightarrow \quad ax \leq ay$$

$$x < y \text{ und } 0 < a \quad \Rightarrow \quad ax < ay$$

(Ungleichungen dürfen mit nicht-negativen Zahlen multipliziert werden.)

$$(d) \quad x \leq y \quad \Leftrightarrow \quad -y \leq -x$$

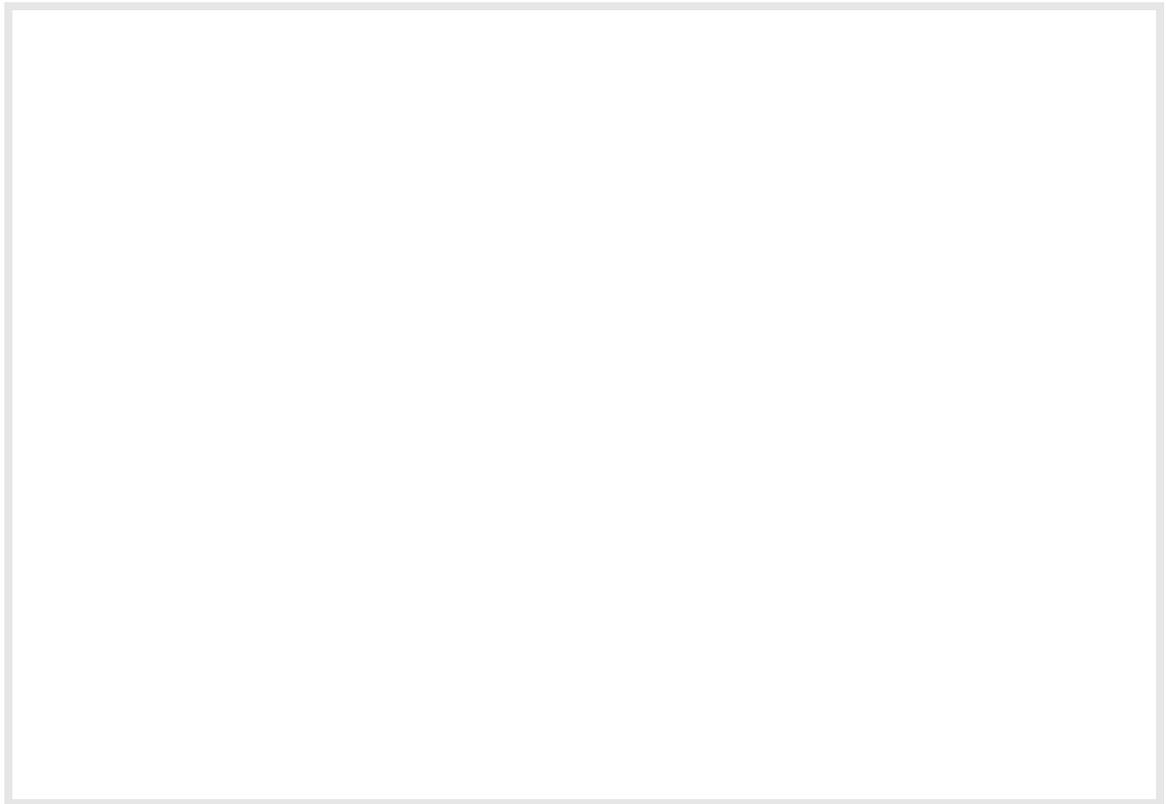
$$x < y \quad \Leftrightarrow \quad -y < -x$$

(Bei Multiplikation mit  $(-1)$  kehrt sich das Relationszeichen um.)

$$(e) \quad 0 < x \leq y \quad \Leftrightarrow \quad 0 < \frac{1}{y} \leq \frac{1}{x}$$

◇

**Beispiel 1.19 (Manipulation von Ungleichungen)**


**Definition 1.20 (Betrag)**

Für Zahlen  $x \in \mathbb{R}$  definieren wir den **Betrag** von  $x$  als

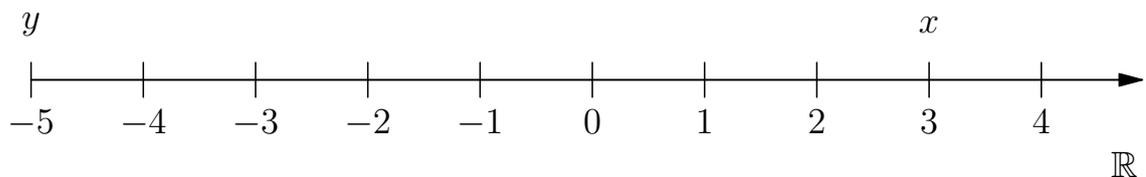
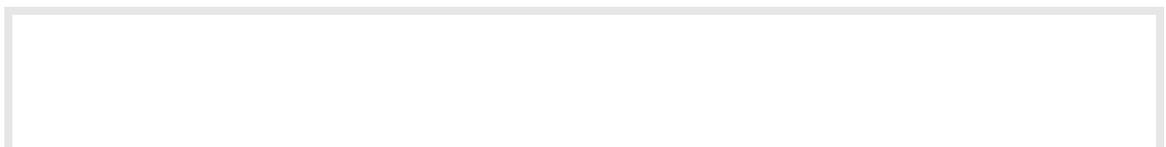
$$|x| := \begin{cases} x, & \text{falls } x \geq 0 \\ -x, & \text{falls } x < 0. \end{cases} \quad \diamond$$

Damit gilt:  $|x| \geq 0$  für alle  $x \in \mathbb{R}$ , z.B. ist  $|-4| = 4$  und  $|4| = 4$ .

**Beachte:**  $|x - y|$  bezeichnet den **Abstand** zwischen  $x$  und  $y$  auf dem Zahlenstrahl. Es gilt natürlich  $|x - y| = |y - x|$ , d.h.,  $x$  hat denselben Abstand zu  $y$  wie  $y$  zu  $x$ . Anhand des Abstandes lässt sich deshalb nicht entscheiden, welche der beiden Zahlen größer ist.

**Frage:** Wie groß ist der Abstand der Zahlen  $x = 3$  und  $y = -5$ ?

**Antwort:**  $|x - y| = |3 - (-5)| = |8| = 8$ :


**Beispiel 1.21 (Rechnen mit Beträgen)**


**Satz 1.22 (Rechenregeln für den Betrag)**

Für  $x, y \in \mathbb{R}$  gilt:

- (a)  $|x| = 0 \Leftrightarrow x = 0$
- (b)  $|xy| = |x| |y|$ , insbesondere  $|y| = |-y|$
- (c)  $|x + y| \leq |x| + |y|$  (Dreiecksungleichung)
- (d)  $-|x| \leq x \leq |x|$
- (e)  $\left|\frac{x}{y}\right| = \frac{|x|}{|y|}$ , falls  $y \neq 0$
- (f)  $|x| \leq y \Leftrightarrow -y \leq x \leq y$
- (g)  $|x| = \sqrt{x^2}$

◇

**§ 1.4.2 Die komplexen Zahlen**

Für manche Rechnungen reichen die reellen Zahlen nicht aus. Zum Beispiel hat die Gleichung

$$x^2 + 1 = 0$$

keine Lösung in  $\mathbb{R}$ , weil Quadrate reeller Zahlen immer  $\geq 0$  sind. Wir führen daher komplexe Zahlen ein. Ein weiterer Grund ist, dass sich manche Rechnungen und Zusammenhänge mit Hilfe komplexer Zahlen („im Komplexen“) **leichter ausdrücken**

lassen, z.B. in der Elektrotechnik und bei der Lösung von Differentialgleichungen (siehe § 5).

**Definition 1.23** (Komplexe Zahlen)

Ein Objekt der Form  $a + bi$  mit  $a, b \in \mathbb{R}$  heißt **komplexe Zahl**. Das Symbol  $i$  bezeichnet die **imaginäre Einheit**. Die Menge aller komplexen Zahlen ist

$$\mathbb{C} := \{z = a + bi : a, b \in \mathbb{R}\}. \quad \diamond$$

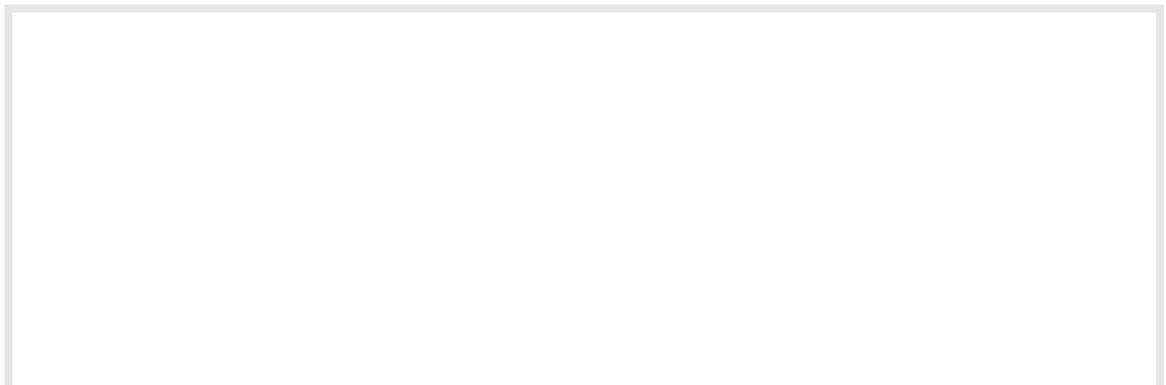
Für die Grundrechenarten mit komplexen Zahlen  $\mathbb{C}$  gelten dieselben Regeln wie für die reellen Zahlen  $\mathbb{R}$ . Wir vereinbaren jedoch zusätzlich die Regel

$$i^2 = -1.$$

**Beachte:** Es gilt  $\mathbb{R} \subsetneq \mathbb{C}$ , denn jedes  $a \in \mathbb{R}$  lässt sich als  $a + 0i$  schreiben.

**Achtung:** Komplexe Zahlen lassen sich i.A. nicht miteinander vergleichen. Für zwei verschiedene  $z, w \in \mathbb{C}$  kann man also weder von „ $z < w$ “ noch von „ $z > w$ “ sprechen.

**Beispiel 1.24** (Grundrechenarten mit komplexen Zahlen)



Binomische Formeln gelten auch für komplexe Zahlen, z.B.

$$(z + w)^2 = z^2 + 2zw + w^2$$

$$(z + w)^3 = z^3 + 3z^2w + 3zw^2 + w^3$$

$$(z + w)^4 = z^4 + 4z^3w + 6z^2w^2 + 4zw^3 + w^4$$

für alle  $z, w \in \mathbb{C}$ .

### Bemerkung 1.25

Mit komplexen Zahlen lassen sich jetzt alle quadratischen Gleichungen der Form

$$ax^2 + bx + c = 0 \quad \text{mit } a, b, c \in \mathbb{R} \text{ und } a \neq 0$$

lösen! Die bekannte  $(a, b, c)$ -Formel gilt weiter: Die Lösungen der obigen Gleichung sind

$$x_{1,2} = -\frac{b}{2a} \pm \frac{1}{2a}\sqrt{b^2 - 4ac},$$

wobei die Zahl unter der Wurzel negativ sein kann. Man setzt dann „ $\sqrt{-1} = i$ “. Analog gilt die  $(p, q)$ -Formel zur Lösung von

$$x^2 + px + q = 0 \quad \text{mit } p, q \in \mathbb{R}.$$

Die Lösungen sind

$$x_{1,2} = -\frac{p}{2} \pm \sqrt{\frac{p^2}{4} - q}. \quad \diamond$$

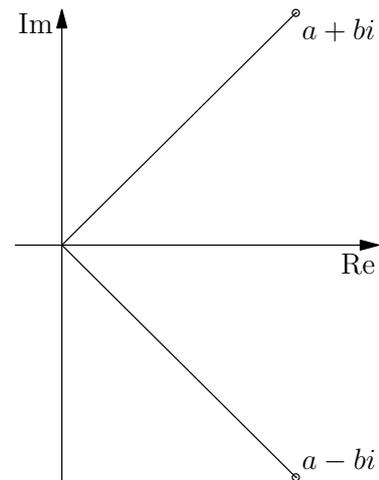
### Beispiel 1.26 (Lösungen quadrat. Gleichungen mit reellen Koeffizienten)

Eine Zahl  $z = a + bi$  mit  $a, b \in \mathbb{R}$  kann in der **komplexen Zahlenebene** (anstelle des Zahlenstrahles) dargestellt werden. Man bezeichnet die Achsen dieser Zahlenebene als **reellen Achse** bzw. **imaginäre Achse**. Man setzt:

$$\operatorname{Re}(z) := a \quad \text{der Realteil von } z$$

$$\operatorname{Im}(z) := b \quad \text{der Imaginärteil von } z.$$

Die Zahl  $\bar{z} = a - bi$  heißt die **konjugiert komplexe Zahl** zu  $z$ . Sie ergibt sich grafisch aus  $z$  durch Spiegelung an der Re-Achse.



**Satz 1.27** (Rechenregeln für konjugiert komplexe Zahlen)

Für  $z, w \in \mathbb{C}$  gilt:

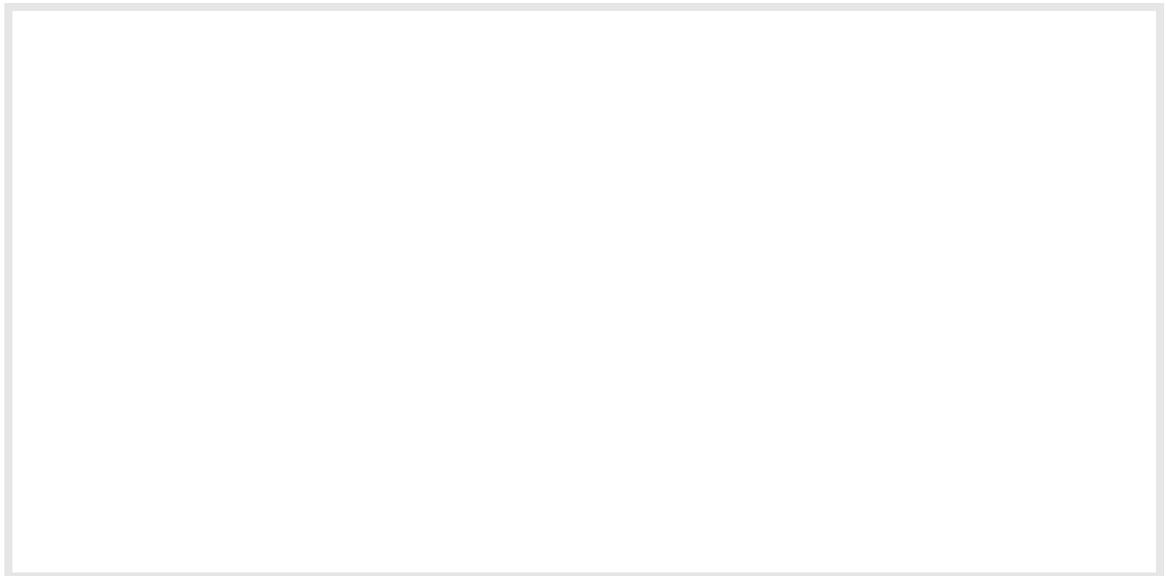
(a)  $\overline{z + w} = \bar{z} + \bar{w}$  und  $\overline{z - w} = \bar{z} - \bar{w}$

(b)  $\overline{z\bar{w}} = \bar{z} w$

(c)  $\overline{\left(\frac{z}{w}\right)} = \frac{\bar{z}}{\bar{w}}$

◇

**Beispiel 1.28** (Real- und Imaginärteil, konjugiert komplexe Zahl)



**Definition 1.29** (Betrag komplexer Zahlen)

Für  $z = a + bi$  mit  $a, b \in \mathbb{R}$  heißt

$$|z| := \sqrt{z\bar{z}} = \sqrt{(a + bi)(a - bi)} = \sqrt{a^2 + b^2}$$

der **Betrag** von  $z$ . Diese Definition erweitert die Definition 1.20 des Betrages für reelle Zahlen, denn: Für  $a \in \mathbb{R}$  und  $b = 0$  gilt  $|z| = |a + 0i| = \sqrt{a^2}$ , vergleiche Satz 1.22. ◇

**Beachte:** Es gilt  $|z| \in \mathbb{R}$  und sogar  $|z| \geq 0$ .  $|z - w|$  bezeichnet wieder den **Abstand** zwischen  $z$  und  $w$  in der komplexen Zahlenebene.

Ende 3. V

28.10.2010

**Satz 1.30** (Rechenregeln für den Betrag, vgl. Satz 1.22)

Für den Betrag und  $z, w \in \mathbb{C}$  gilt:

- (a)  $|z| = 0 \Leftrightarrow z = 0$
- (b)  $|z w| = |z| |w|$
- (c)  $|z + w| \leq |z| + |w|$
- (d)  $\left| \frac{z}{w} \right| = \frac{|z|}{|w|}$ , falls  $w \neq 0$ .

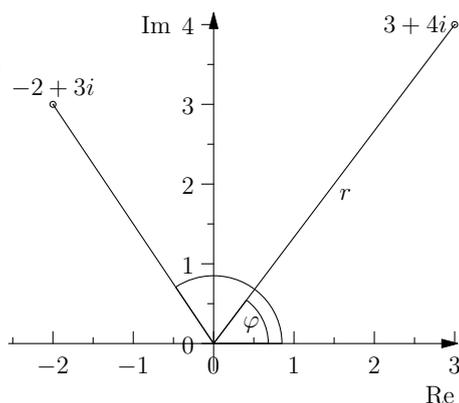
◇

In der komplexen Zahlenebene sieht man, dass man eine komplexe Zahl  $z = a + bi$  in der sogenannten **algebraischen Darstellung** auch durch Angabe von  $(r, \varphi)$  beschreiben kann, wobei

$r := |z| \geq 0$  der **Betrag** und

$\varphi := \arg(z) \in [0, 2\pi)$  das **Argument** von  $z$  ist.

Das Argument ist der Winkel gegen die positive reelle Achse.



Es gelten die folgenden Umrechnungsbeziehungen:

Umrechnung von  $(r, \varphi)$  in algebraische Darstellung  $z = a + bi$ :

$$a = r \cos \varphi, \quad b = r \sin \varphi.$$

Es gilt also:

$$\begin{aligned} z = a + bi &= r \cos \varphi + (r \sin \varphi) i \\ &= r (\cos \varphi + i \sin \varphi). \end{aligned}$$

Dies nennt man die **Polardarstellung** oder auch **trigonometrische Darstellung** der komplexen Zahl  $z$ .

Umrechnung von algebraischer Darstellung  $z = a + bi$  in Polardarstellung:

$$r = |z| = \sqrt{a^2 + b^2}$$

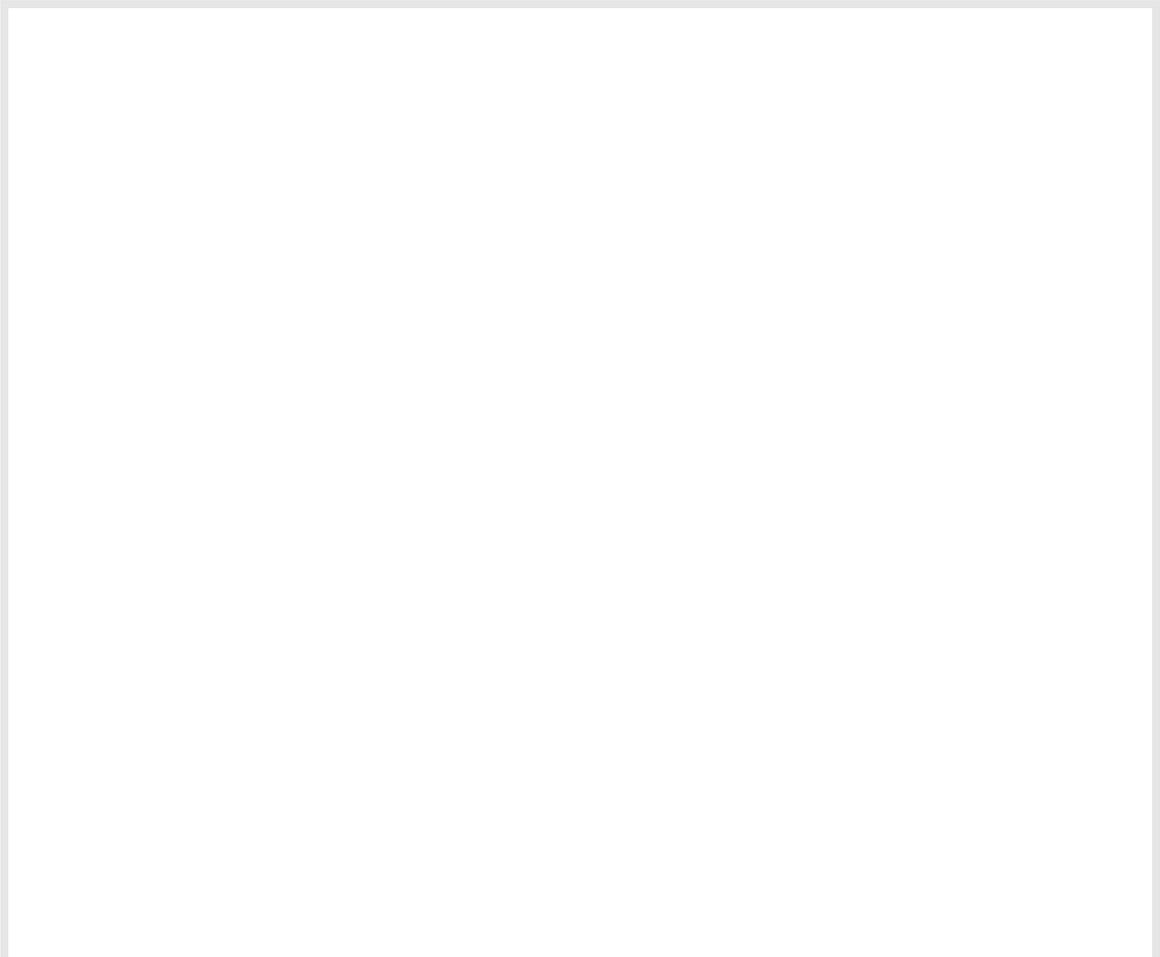
Falls  $a = 0$  ist (rein imaginäre Zahl):

$$\varphi = \begin{cases} \frac{\pi}{2}, & \text{falls } a = 0 \text{ und } b > 0 & \text{(obere Im-Achse)} \\ \text{beliebig,} & \text{falls } a = 0 \text{ und } b = 0 & \text{(Ursprung)} \\ \frac{3\pi}{2}, & \text{falls } a = 0 \text{ und } b < 0 & \text{(untere Im-Achse).} \end{cases}$$

Falls  $a \neq 0$  ist, nutzen wir die Beziehung  $\tan \varphi = \frac{b}{a}$ , die wir mit Hilfe der Umkehrfunktion (Arcustangens), siehe Beispiel 4.26, nach  $\varphi$  auflösen müssen. Am Vorzeichen von  $a$  und  $b$  können wir ablesen, in welchem der vier Quadranten (Winkelbereiche) die Zahl liegen muss:

$$\varphi = \begin{cases} \arctan(b/a), & \text{falls } a > 0 \text{ und } b \geq 0 & \text{(1. Quadrant)} \\ \pi + \arctan(b/a), & \text{falls } a < 0 \text{ und } b \geq 0 & \text{(2. Quadrant)} \\ \pi + \arctan(b/a), & \text{falls } a < 0 \text{ und } b < 0 & \text{(3. Quadrant)} \\ 2\pi + \arctan(b/a), & \text{falls } a > 0 \text{ und } b < 0 & \text{(4. Quadrant)} \end{cases}$$

### Beispiel 1.31 (Polardarstellung)



Wir führen jetzt die Abkürzung die **Eulersche Formel** ein:

$$e^{i\varphi} = \exp(i\varphi) := \cos \varphi + i \sin \varphi \quad \text{für } \varphi \in \mathbb{R}.$$

Diese Definition erweitert den Definitionsbereich der bekannten Exponentialfunktion  $e^x = \exp(x)$  auf komplexe Zahlen, die üblichen Rechenregeln gelten weiter. Mit Hilfe der Eulerschen Formel können wir die Polardarstellung auch schreiben als sogenannte **Exponentialdarstellung**:

$$z = r (\cos \varphi + i \sin \varphi) = r e^{i\varphi} \quad \text{oder} \quad z = r \exp(i\varphi).$$

Zwei komplexe Zahlen in Exponentialdarstellung kann man bequem multiplizieren, dividieren und potenzieren. Das Ergebnis liegt wieder in Exponentialdarstellung vor:

**Satz 1.32 (Rechnen in Exponentialdarstellung)**

Seien  $z = r_1 e^{i\varphi_1}$  und  $w = r_2 e^{i\varphi_2}$ . Dann ist

$\bar{z} = r_1 e^{-i\varphi_1}$	{	Betrag bleibt gleich Negation des Winkels
$z \cdot w = r_1 r_2 e^{i(\varphi_1 + \varphi_2)}$	{	Multiplikation der Beträge Addition der Winkel
$\frac{1}{w} = \frac{1}{r_2} e^{-i\varphi_2}$	{	Kehrwert des Betrags Negation des Winkels
$\frac{z}{w} = \frac{r_1}{r_2} e^{i(\varphi_1 - \varphi_2)}$	{	Division der Beträge Subtraktion der Winkel
$z^n = r_1^n e^{in\varphi_1} \quad \text{für } n \in \mathbb{N}$	{	Potenz des Betrags Vielfaches des Winkels <b>(Formel von Moivre)</b>

◇

Zusammenfassung der Vor- und Nachteile:

	algebraische Darstellung $z = a + bi$	Exponentialdarstellung $z = r e^{i\varphi}$
Addition $z + w$	✓	—
Subtraktion $z - w$	✓	—
Multiplikation $z \cdot w$	✓	✓
Division $\frac{z}{w}$	Erweitern mit $\bar{w}$	✓
Potenzieren $z^n$	mühsam	✓

## § 2 Lineare Algebra und analytische Geometrie

§ 2.1 Vektoren im  $\mathbb{R}^n$ **Definition 2.1** (Vektoren im  $\mathbb{R}^n$ )

(a) Ein geordnetes  $n$ -Tupel reeller Zahlen  $x_i$

$$\vec{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

heißt ein **Vektor** des **Vektorraumes**  $\mathbb{R}^n$ , wobei  $n \in \mathbb{N} = \{1, 2, 3, \dots\}$  gilt.

(b) Weiter heißt  $x_i$  die  $i$ -te **Komponente** des Vektors  $\vec{x}$ . Die Anzahl der Komponenten eines Vektors wird manchmal als **Länge** des Vektors  $\vec{x}$  bezeichnet. Hier besteht aber Verwechslungsgefahr mit der Länge nach Definition 2.15.

(c) Die Vektoren

$$\vec{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \quad \vec{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \dots, \quad \vec{e}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

heißen die **Einheitsvektoren** des  $\mathbb{R}^n$ . ◇

In der Anschauung verwendet man Vektoren des  $\mathbb{R}^2$  (Ebene, siehe § 2.6) bzw. des  $\mathbb{R}^3$  (Raum, siehe § 2.7). Allgemeine Vektoren im  $\mathbb{R}^n$  werden verwendet, um z.B. eine Anzahl von physikalischen Größen zu einer Größe zusammenzufassen. Man kann  $\mathbb{R}^n$  auffassen als kartesisches Produkt  $\mathbb{R} \times \mathbb{R} \times \dots \times \mathbb{R}$ . Es gilt  $\mathbb{R}^1 = \mathbb{R}$ .

Vektoren der Länge eins, also Zahlen, nennt man oft auch **Skalare**. Wir definieren für Vektoren  $\vec{x}$  und  $\vec{y}$  des  $\mathbb{R}^n$  und Skalare  $\alpha \in \mathbb{R}$  folgende Operationen:

**Addition** von Vektoren  $\vec{x} + \vec{y} := \begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ \vdots \\ x_n + y_n \end{pmatrix}$

**Multiplikation** eines Vektors mit einem Skalar  $\alpha \vec{x} := \begin{pmatrix} \alpha x_1 \\ \alpha x_2 \\ \vdots \\ \alpha x_n \end{pmatrix}$ .

Die Addition des **Nullvektors**  $\vec{0} = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$  und die Multiplikation mit 1 lassen einen

Vektor  $\vec{x}$  unverändert. Unter  $-\vec{x}$  verstehen wir den Vektor  $(-1)\vec{x}$ . Damit definieren

wir die

**Subtraktion** von Vektoren  $\vec{x} - \vec{y} := \vec{x} + (-\vec{y}) = \begin{pmatrix} x_1 - y_1 \\ x_2 - y_2 \\ \vdots \\ x_n - y_n \end{pmatrix}.$

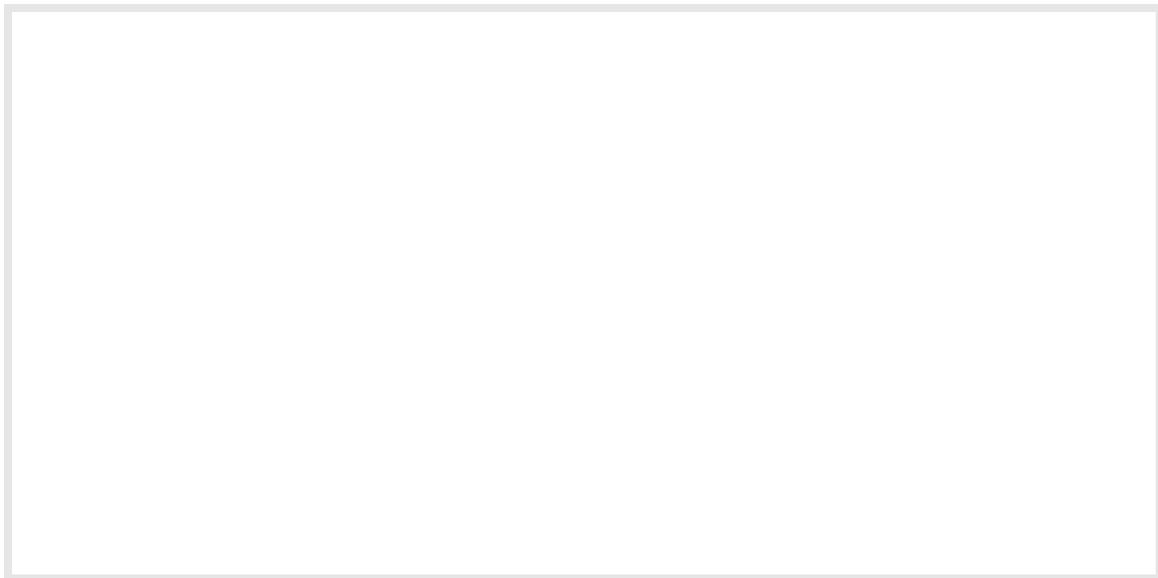
Es gilt also  $\vec{x} - \vec{x} = \vec{0}$ .

**Satz 2.2 (Eigenschaften der Vektoroperationen)**

Für Vektoren  $\vec{x}, \vec{y}, \vec{z} \in \mathbb{R}^n$  und Skalare  $\alpha, \beta \in \mathbb{R}$  gelten:

- (a)  $\vec{x} + \vec{y} = \vec{y} + \vec{x}$  (Kommutativität)
- (b)  $(\vec{x} + \vec{y}) + \vec{z} = \vec{x} + (\vec{y} + \vec{z})$  (Assoziativität)
- (c)  $\alpha(\beta \vec{x}) = (\alpha\beta) \vec{x}$  (Assoziativität der Multiplikation mit zwei Skalaren)
- (d)  $(\alpha + \beta) \vec{x} = \alpha \vec{x} + \beta \vec{x}$  (Distributivgesetz 1)
- (e)  $\alpha(\vec{x} + \vec{y}) = \alpha \vec{x} + \alpha \vec{y}$  (Distributivgesetz 2)
- (f)  $-(\vec{x} + \vec{y}) = -\vec{x} - \vec{y}$
- (g)  $-(-\vec{x}) = \vec{x}$  ◇

**Beispiel 2.3 (Vektorrechnung)**

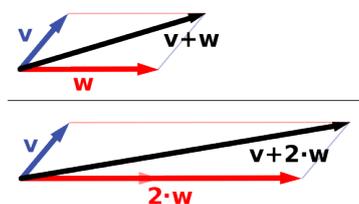


Oft kommen Summen von Vektoren und Multiplikation mit Skalaren gemeinsam vor:

$$1 \cdot \vec{v} + 1 \cdot \vec{w}$$

oder

$$1 \cdot \vec{v} + 2 \cdot \vec{w}.$$



Dies sind Linearkombinationen von  $\vec{v}$  und  $\vec{w}$ :

**Definition 2.4 (Linearkombination)**

Sei  $k \in \mathbb{N}$  und Vektoren  $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k \in \mathbb{R}^n$  gegeben.

(a) Eine Summe der Form

$$\sum_{i=1}^k \alpha_i \vec{v}_i := \alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2 + \dots + \alpha_k \vec{v}_k$$

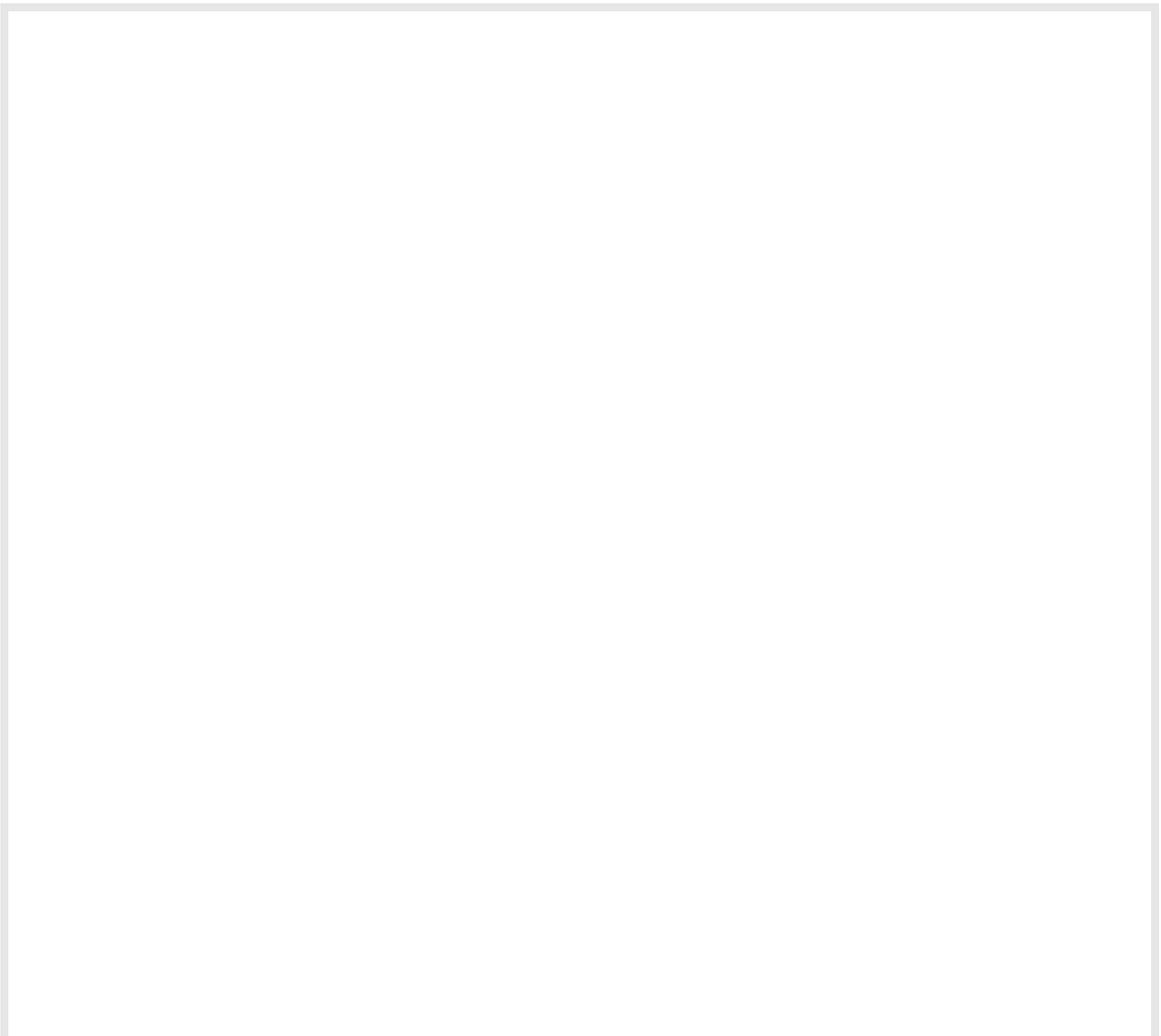
mit Skalaren (**Koeffizienten**)  $\alpha_1, \alpha_2, \dots, \alpha_k \in \mathbb{R}$  heißt eine **Linearkombination (LK)** der Vektoren  $\vec{v}_i, i = 1, \dots, k$ .

(b) Die LK heißt **trivial**, wenn alle  $\alpha_1 = \dots = \alpha_k = 0$  sind. (Das Ergebnis ist der Nullvektor.)

(c) Die Menge *aller* LK

$$\text{Lin}(\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k) := \left\{ \sum_{i=1}^k \alpha_i \vec{v}_i : \alpha_1, \dots, \alpha_k \in \mathbb{R} \right\}$$

heißt die **lineare Hülle** der Vektoren  $\vec{v}_i, i = 1, \dots, k$ . ◇

**Beispiel 2.5 (Linearkombination, lineare Hülle)**

**Definition 2.6 (Lineare Unabhängigkeit)** (a) Die Vektoren  $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k \in \mathbb{R}^n$  heißen **linear abhängig**, wenn es Zahlen  $\alpha_1, \dots, \alpha_k \in \mathbb{R}$  gibt, die nicht alle gleich null sind, sodass gilt:

$$\alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2 + \dots + \alpha_k \vec{v}_k = \vec{0}.$$

(Der Nullvektor lässt sich nicht-trivial aus den  $\vec{v}_i$  linear kombinieren.)

(b) Die Vektoren  $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k \in \mathbb{R}^n$  heißen hingegen **linear unabhängig**, wenn gilt:

$$\alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2 + \dots + \alpha_k \vec{v}_k = \vec{0} \quad \Rightarrow \quad \alpha_1 = \alpha_2 = \dots = \alpha_k = 0.$$

(Der Nullvektor lässt sich nur trivial aus den  $\vec{v}_i$  linear kombinieren.)  $\diamond$

**Beachte:** Man sagt *nicht*, ein Vektor sei linear (un)abhängig von anderen Vektoren.

Anschaulich bedeutet die lineare Unabhängigkeit der Vektoren  $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k$ , dass jeder von ihnen „in eine neue Richtung zeigt“.

**Satz 2.7 (Lineare Abhängigkeit)**

Die Vektoren  $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k \in \mathbb{R}^n$  sind genau dann linear abhängig, wenn mindestens einer der Vektoren als LK der anderen geschrieben werden kann.  $\diamond$

**Beispiel 2.8 (Lineare Unabhängigkeit)**



**Definition 2.9 (Unterraum)**

Eine Teilmenge  $V \subseteq \mathbb{R}^n$  heißt ein **Unterraum (UR)** des  $\mathbb{R}^n$ , wenn gilt:

- (a)  $\vec{x}, \vec{y} \in V \Rightarrow \vec{x} + \vec{y} \in V$  und
- (b)  $\vec{x} \in V$  und  $\alpha \in \mathbb{R} \Rightarrow \alpha \vec{x} \in V$ .

(Also: Summen und Vielfache von Elementen von  $V$  liegen wieder in  $V$ .)  $\diamond$

Einige Fakten:

- Der Raum  $V = \mathbb{R}^n$  selbst ist der größte, die Menge  $V = \{\vec{0}\}$  ist der kleinste UR von  $\mathbb{R}^n$ .
- Jeder UR enthält den Nullvektor  $\vec{0}$ .
- Die lineare Hülle  $\text{Lin}(\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k)$  von Vektoren  $\vec{v}_i \in \mathbb{R}^n$ ,  $i = 1, \dots, k$ , ist immer ein UR des  $\mathbb{R}^n$ .

**Definition 2.10 (Basis, Dimension)**

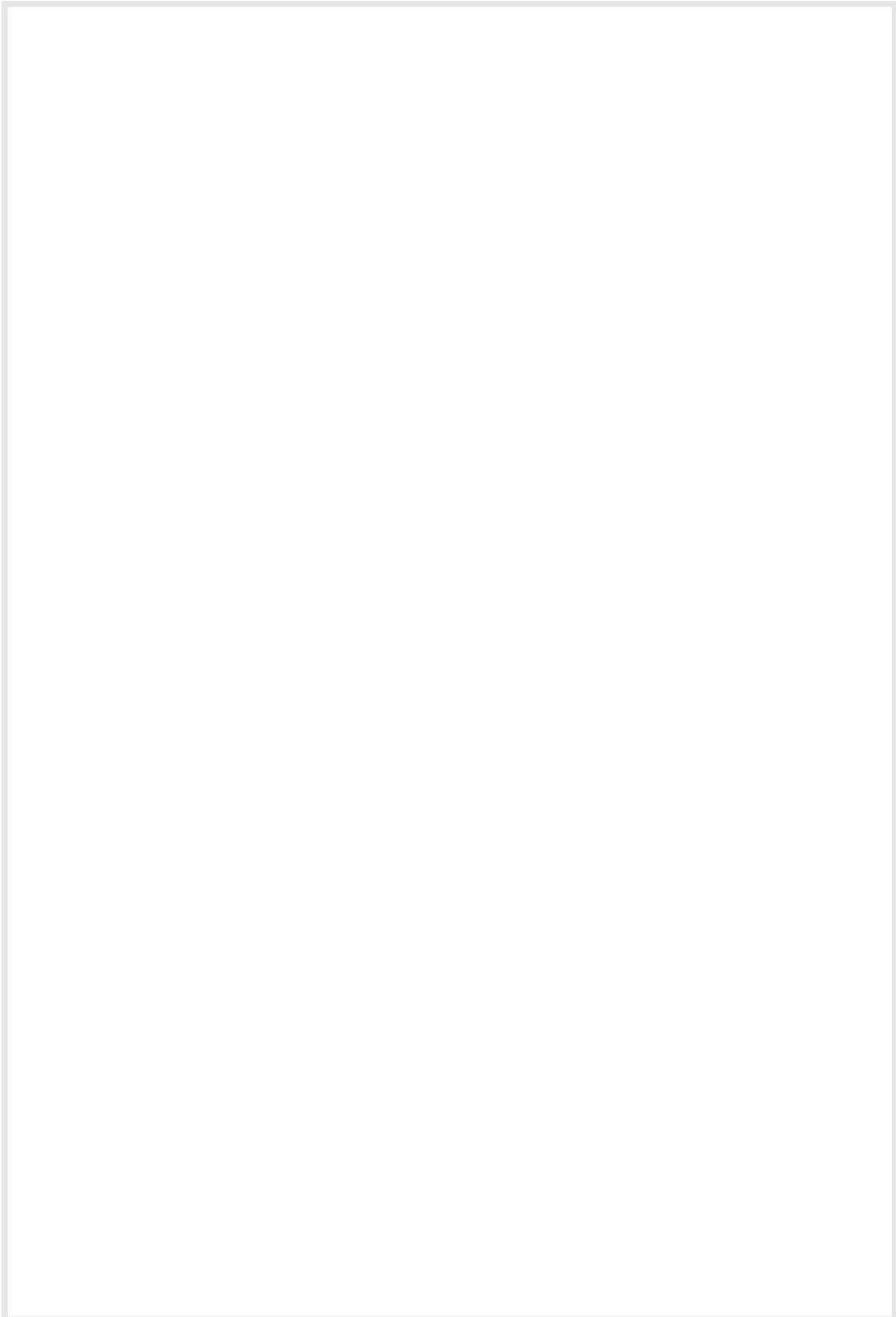
- (a) Die **Dimension** eines UR  $V \subseteq \mathbb{R}^n$  ( $\dim V$ ) ist die maximale Anzahl linear unabhängiger Vektoren in  $V$ . (Diese liegt zwischen 0 und  $n$ .)
- (b) Eine solche maximale Menge linear unabhängiger Vektoren heißt eine **Basis** des UR  $V$ .  $\diamond$

Unterräume des  $\mathbb{R}^n$  der Dimension 1 sind Geraden, die durch den Ursprung gehen. Unterräume des  $\mathbb{R}^n$  der Dimension 2 sind Ebenen, die durch den Ursprung gehen.

**Satz 2.11 (Bedeutung der Basis)**

Sei  $V \subseteq \mathbb{R}^n$  ein UR der Dimension  $m$ , und sei  $\{\vec{v}_1, \vec{v}_2, \dots, \vec{v}_m\}$  eine Basis von  $V$ . Dann ist jeder Vektor  $\vec{x} \in V$  in eindeutiger Art und Weise als LK der Basisvektoren darstellbar.  $\diamond$

**Beispiel 2.12** (Basis und Dimension)



Ende 5 V  
11.11.2010

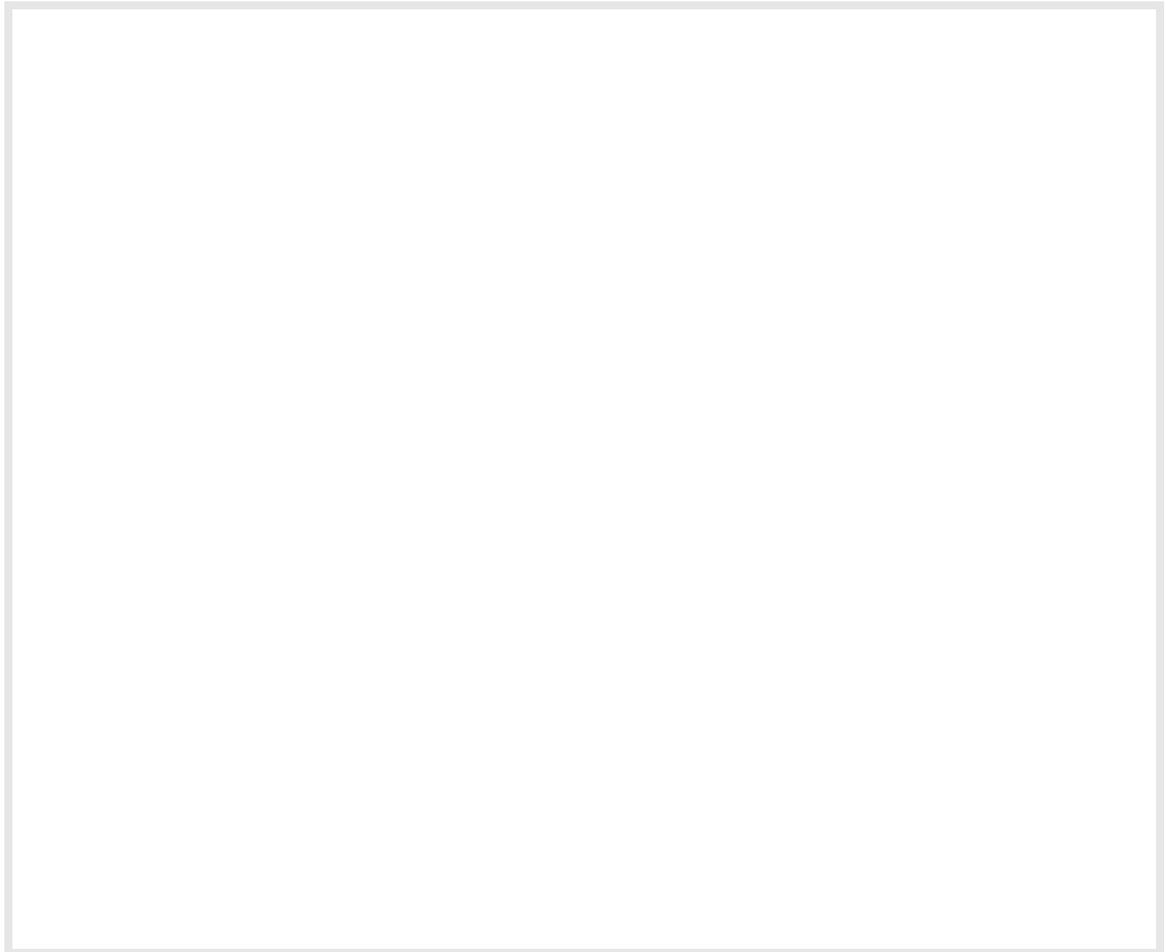
### Bemerkung 2.13

Alle Vektoren waren bisher **Spaltenvektoren**. Analog kann man auch **Zeilenvektoren** definieren, die den Vektorraum  $\mathbb{R}_n$  bilden:

$$\vec{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}_n.$$

Die Rechenregeln gelten analog weiter.  $\diamond$

### Beispiel 2.14 (Zeilen- und Spaltenvektoren)



## § 2.2 Das Skalarprodukt

### Definition 2.15 (Skalarprodukt)

Für Vektoren  $\vec{x}, \vec{y} \in \mathbb{R}^n$  definieren wir das **Skalarprodukt** (innere Produkt, Innenprodukt)

$$\vec{x} \cdot \vec{y} := \sum_{i=1}^n x_i y_i = x_1 y_1 + x_2 y_2 + \dots + x_n y_n.$$

Die Zahl

$$\|\vec{x}\| := \sqrt{\vec{x} \cdot \vec{x}} = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2} \geq 0$$

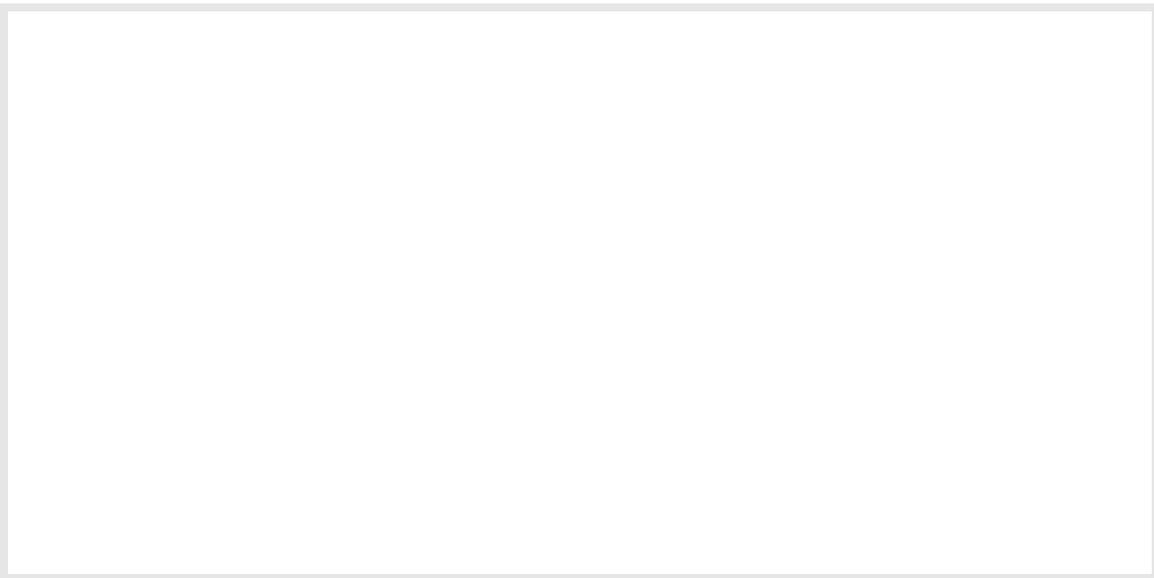
heißt **Betrag**, **Länge** oder (**euklidische**) **Norm** des Vektors  $\vec{x}$ . (Für  $\vec{x} \in \mathbb{R}^1 = \mathbb{R}$  stimmt dies mit dem Betrag aus Definition 1.20 überein.)  $\diamond$

**Satz 2.16** (Eigenschaften des Skalarproduktes und des Betrages)

Für Vektoren  $\vec{x}, \vec{y}, \vec{z} \in \mathbb{R}^n$  und Skalare  $\alpha \in \mathbb{R}$  gelten:

- (a)  $\vec{x} \cdot \vec{y} = \vec{y} \cdot \vec{x}$
- (b)  $\alpha (\vec{x} \cdot \vec{y}) = (\alpha \vec{x}) \cdot \vec{y} = \vec{x} \cdot (\alpha \vec{y})$
- (c)  $\vec{x} \cdot (\vec{y} + \vec{z}) = \vec{x} \cdot \vec{y} + \vec{x} \cdot \vec{z}$
- (d)  $\vec{x} \cdot \vec{x} = 0 \Leftrightarrow \|\vec{x}\| = 0 \Leftrightarrow \vec{x} = \vec{0}$
- (e)  $\|\alpha \vec{x}\| = |\alpha| \|\vec{x}\|$
- (f)  $-|\vec{x} \cdot \vec{y}| \leq \vec{x} \cdot \vec{y} \leq |\vec{x} \cdot \vec{y}| \leq \|\vec{x}\| \|\vec{y}\|$  (Cauchy-Schwarzsche Ungleichung)
- (g)  $\|\vec{x} + \vec{y}\| \leq \|\vec{x}\| + \|\vec{y}\|$  (Dreiecksungleichung)  $\diamond$

**Beispiel 2.17** (Skalarprodukt)



**Definition 2.18** (Winkel zwischen Vektoren) (a) Mit Hilfe des Skalarprodukts kann man einen **Winkel**  $\varphi \in [0, \pi]$  zwischen Vektoren  $\vec{x}$  und  $\vec{y}$  im  $\mathbb{R}^n$  definieren:

$$\varphi := \arccos \frac{\vec{x} \cdot \vec{y}}{\|\vec{x}\| \|\vec{y}\|}, \quad \text{falls } \vec{x}, \vec{y} \neq \vec{0}.$$

Es gilt also:

$$\vec{x} \cdot \vec{y} = \|\vec{x}\| \|\vec{y}\| \cos \varphi.$$

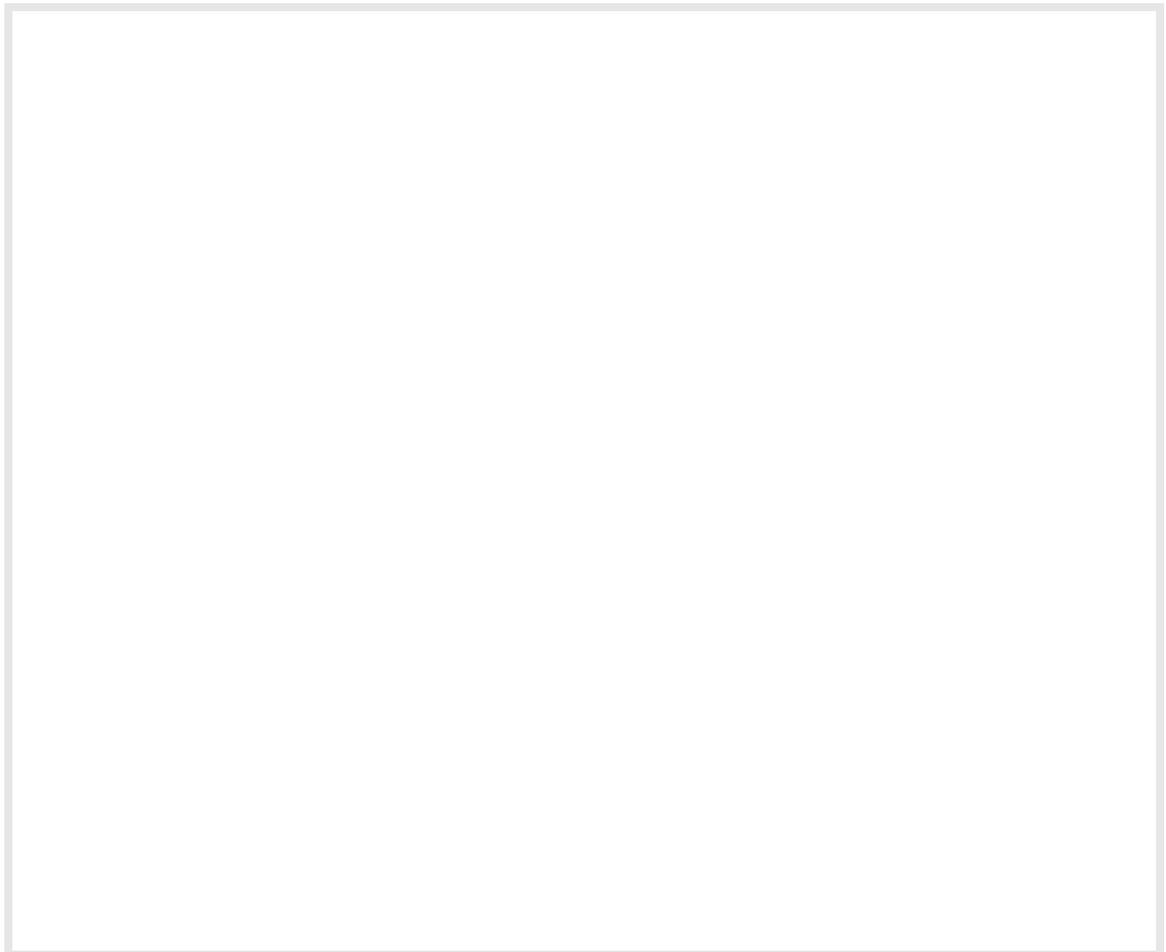
**Beachte:**

- Skalarprodukt  $\vec{x} \cdot \vec{y} > 0 \Leftrightarrow 0 < \varphi < 90^\circ = \pi/2$
- Skalarprodukt  $\vec{x} \cdot \vec{y} < 0 \Leftrightarrow \pi/2 = 90^\circ < \varphi \leq 180^\circ = \pi$

- (b) Die Vektoren  $\vec{x}$  und  $\vec{y}$  heißen **orthogonal** (**senkrecht**) zueinander oder kurz:  $\vec{x} \perp \vec{y}$ , wenn  $\vec{x} \cdot \vec{y} = 0$  gilt.
- (c) Die Vektoren  $\vec{x} \neq \vec{0}$  und  $\vec{y} \neq \vec{0}$  heißen **parallel**, wenn  $\vec{x} \cdot \vec{y} = \|\vec{x}\| \|\vec{y}\|$  gilt, also wenn  $\varphi = 0$  ist.
- (d) Die Vektoren  $\vec{x} \neq \vec{0}$  und  $\vec{y} \neq \vec{0}$  heißen **anti-parallel**, wenn  $\vec{x} \cdot \vec{y} = -\|\vec{x}\| \|\vec{y}\|$  gilt, also wenn  $\varphi = \pi$  ist.  $\diamond$

Diese Definitionen entsprechen den geometrischen Anschauungen im  $\mathbb{R}^2$  und  $\mathbb{R}^3$ .

### Beispiel 2.19 (Winkel zwischen Vektoren)



## § 2.3 Matrizen

**Definition 2.20** (**Matrix**) (a) Ein rechteckiges Zahlenschema der Form

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}$$

mit Einträgen  $a_{ij} \in \mathbb{R}$  heißt eine (reelle)  $m \times n$ -**Matrix** oder eine **Matrix vom Typ**  $(m, n)$ .  $A$  besitzt  $m$  Zeilen und  $n$  Spalten. Man schreibt kurz:

$A = (a_{ij}) \in \mathbb{R}^{m \times n}$ , wobei  $i = 1, \dots, m$  die Zeilen und  $j = 1, \dots, n$  die Spalten nummeriert.

- (b) Eine Matrix heißt **quadratisch**, wenn  $m = n$  gilt. Die Einträge  $a_{11}, a_{22}, \dots, a_{nn}$  bilden die **Hauptdiagonale**.
- (c) Eine quadratische Matrix heißt **obere Dreiecksmatrix**, wenn unterhalb der Hauptdiagonalen nur Nullen stehen:

$$\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & a_{22} & \cdots & a_{2n} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & a_{nn} \end{pmatrix}$$

Analog: **untere Dreiecksmatrix**.

- (d) Bei einer **Diagonalmatrix** stehen außerhalb der Hauptdiagonalen nur Nullen:

$$\begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & a_{22} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & a_{nn} \end{pmatrix} \quad \diamond$$

**Beachte:** Spaltenvektoren sind  $\mathbb{R}^{m \times 1}$ -Matrizen („schlank“), Zeilenvektoren sind  $\mathbb{R}^{1 \times n}$ -Matrizen („flach“).

Wir definieren für Matrizen  $A = (a_{ij})$  und  $B = (b_{ij})$  des  $\mathbb{R}^{m \times n}$  und Skalare  $\alpha \in \mathbb{R}$  folgende Operationen:

**Addition** von Matrizen  $A + B := \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \cdots & a_{1n} + b_{1n} \\ a_{21} + b_{21} & a_{22} + b_{22} & \cdots & a_{2n} + b_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} + b_{m1} & a_{m2} + b_{m2} & \cdots & a_{mn} + b_{mn} \end{pmatrix}$

**Multiplikation** mit einem Skalar  $\alpha A := \begin{pmatrix} \alpha a_{11} & \alpha a_{12} & \cdots & \alpha a_{1n} \\ \alpha a_{21} & \alpha a_{22} & \cdots & \alpha a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \alpha a_{m1} & \alpha a_{m2} & \cdots & \alpha a_{mn} \end{pmatrix}$ .

Die Addition der **Nullmatrix**  $\mathbf{0} = \begin{pmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{pmatrix} \in \mathbb{R}^{m \times n}$  und die Multiplikation mit

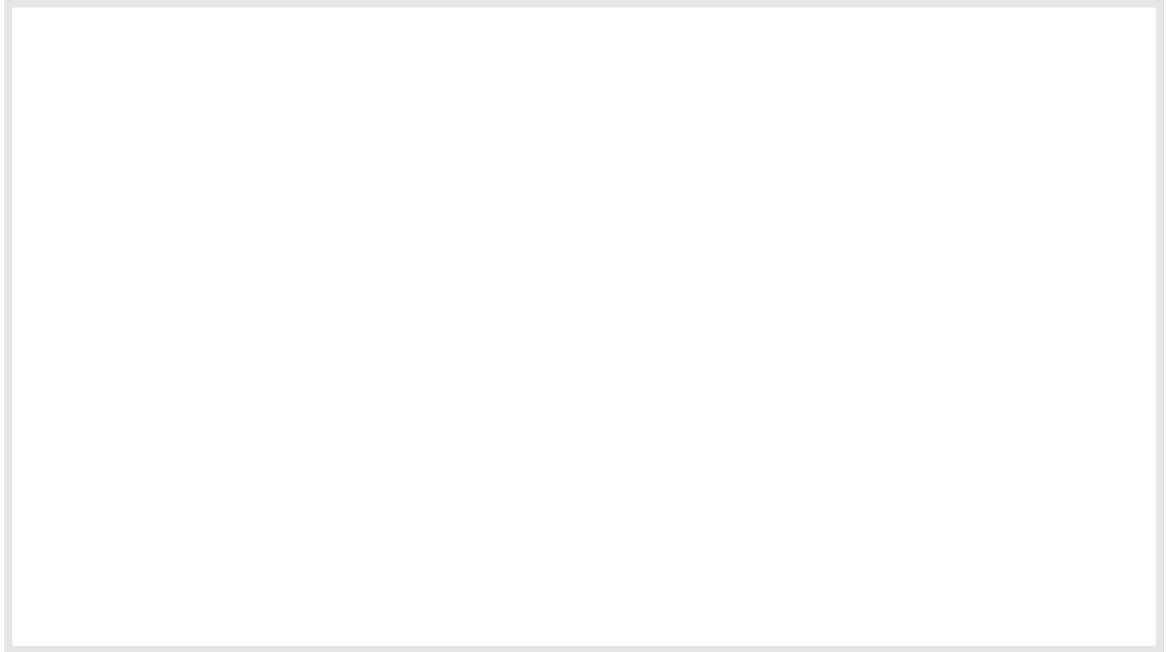
der Zahl 1 lassen eine Matrix  $A$  unverändert. Unter  $-A$  verstehen wir die Matrix  $(-1)A$ .

Die  $n \times n$ -**Einheitsmatrix**  $E$  (manchmal auch  $I$  oder **Identität**) besteht aus den Einheitsvektoren des  $\mathbb{R}^n$ :

$$E = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \quad (\text{Diagonalmatrix mit Hauptdiagonale aus Einsen}).$$

Es gelten die gleichen Eigenschaften wie in Satz 2.2 für Addition von Vektoren. Wie in Beispiel 2.14 für Vektoren können wir eine Matrix **transponieren**, indem wir Zeilen zu Spalten machen:  $A \in \mathbb{R}^{m \times n} \Leftrightarrow A^\top \in \mathbb{R}^{n \times m}$ . Eine (quadratische) Matrix heißt **symmetrisch**, wenn  $A = A^\top$  ist.

**Beispiel 2.21** (Addition, Multiplikation mit einem Skalar, Transposition)



Ende 6. V

18.11.2010

**Definition 2.22** (Matrix-Vektor-Multiplikation)

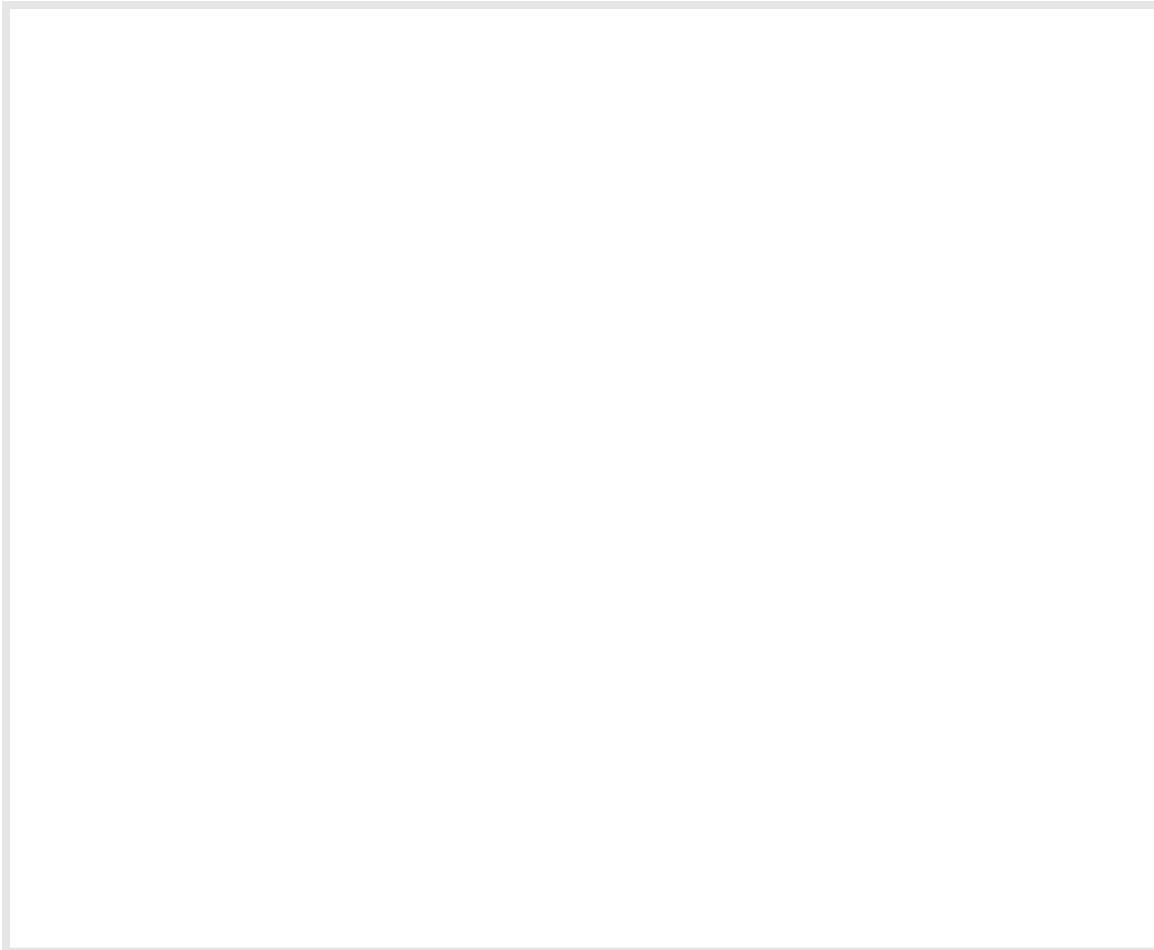
Für Matrizen  $A = (a_{ij}) \in \mathbb{R}^{m \times n}$  und Vektoren  $\vec{x} \in \mathbb{R}^n$  definieren wir das **Matrix-Vektor-Produkt**

$$\begin{aligned}
 A\vec{x} &= \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ \underbrace{a_{m1}}_{\vec{a}^1} & \underbrace{a_{m2}}_{\vec{a}^2} & \cdots & \underbrace{a_{mn}}_{\vec{a}^n} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} := x_1 \vec{a}^1 + x_2 \vec{a}^2 + \cdots + x_n \vec{a}^n \\
 &= \begin{pmatrix} a_{11} x_1 + a_{12} x_2 + \cdots + a_{1n} x_n \\ a_{21} x_1 + a_{22} x_2 + \cdots + a_{2n} x_n \\ \vdots \\ a_{m1} x_1 + a_{m2} x_2 + \cdots + a_{mn} x_n \end{pmatrix} \in \mathbb{R}^m.
 \end{aligned}$$

**Merkregel:** Man berechnet nacheinander die Skalarprodukte (Zeile von  $A$ )  $\cdot \vec{x}$  und schreibt die Ergebnisse untereinander.

Das Ergebnis ist eine LK der *Spaltenvektoren*  $\vec{a}^1, \dots, \vec{a}^n$  von  $A$  mit den Koeffizienten  $x_1, \dots, x_n$ .  $\diamond$

**Beachte:** Man kann das Matrix-Vektor-Produkt  $A\vec{x}$  nur dann bilden, wenn die Breite (Anzahl Spalten) von  $A$  übereinstimmt mit der Anzahl der Einträge des Spaltenvektors  $\vec{x}$ .

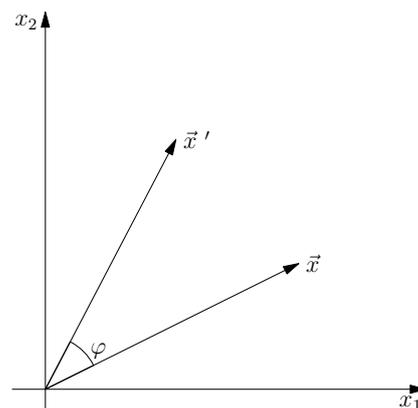
**Beispiel 2.23** (Matrix-Vektor-Multiplikation)

Matrizen werden auch benutzt, um Transformationen wie Drehungen, Spiegelungen, Skalierungen eines Ortsvektors in der Ebene (siehe § 2.6) und im Raum (§ 2.7) darzustellen, die etwa in der Computergrafik oft benötigt werden.

**Beispiel 2.24** (Geom. Transformationen als Matrix-Vektor-Produkte)

- (a) Sei  $\vec{x} = (x_1, x_2)^\top$  der Ortsvektor eines Punktes in der Ebene. Eine **Drehung** des Vektors  $\vec{x}$  um den Winkel  $\varphi$  mit dem Ursprung als Drehzentrum erreicht man durch

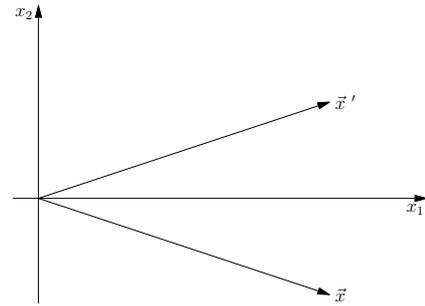
$$\vec{x}' := \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix} \vec{x}$$



(b) Durch

$$\vec{x}' := \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \vec{x}$$

erhält man eine **Spiegelung** des Punktes an der  $x_1$ -Achse.



(c) Die **Drehung** eines Punktes im Raum mit Ortsvektor  $\vec{x} = (x_1, x_2, x_3)^\top$  um den Winkel  $\varphi$  bzgl. der  $x_1$ -Achse erreicht man durch

$$\vec{x}' := \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \varphi & -\sin \varphi \\ 0 & \sin \varphi & \cos \varphi \end{pmatrix} \vec{x}.$$

◇

### Definition 2.25 (Matrix-Matrix-Multiplikation)

Seien  $A = (a_{ij}) \in \mathbb{R}^{m \times n}$  und  $B = (b_{ij}) \in \mathbb{R}^{n \times p}$  Matrizen.  $B$  bestehe aus den Spaltenvektoren  $\vec{b}^1, \dots, \vec{b}^p \in \mathbb{R}^n$ . Wir definieren das **Matrix-Matrix-Produkt**

$$AB = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1p} \\ b_{21} & b_{22} & \cdots & b_{2p} \\ \vdots & \vdots & & \vdots \\ \underbrace{b_{n1}}_{\vec{b}^1} & \underbrace{b_{n2}}_{\vec{b}^2} & \cdots & \underbrace{b_{np}}_{\vec{b}^p} \end{pmatrix}$$

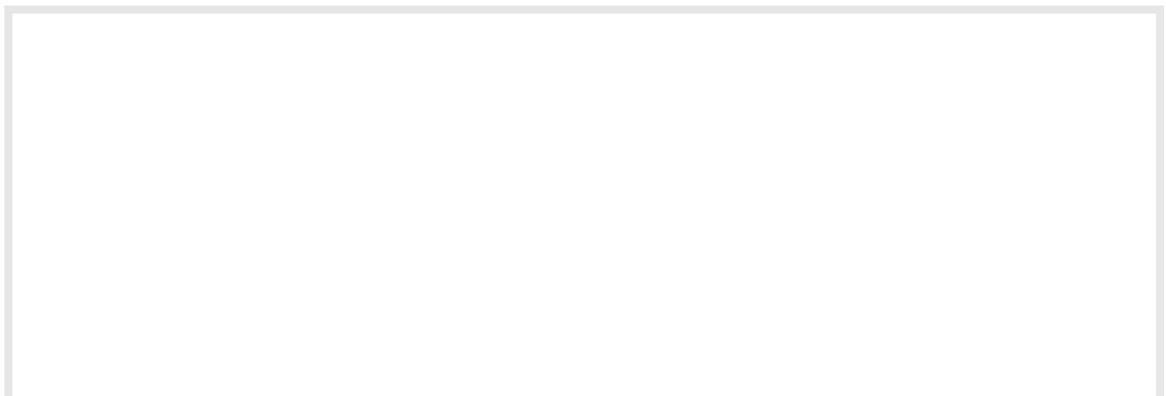
$$:= (A\vec{b}^1 \quad A\vec{b}^2 \quad \cdots \quad A\vec{b}^p) \in \mathbb{R}^{m \times p}.$$

**Merkregel:** Man berechnet nacheinander Matrix-Vektor-Produkte  $A \cdot$  (Spalte von  $B$ ) und schreibt die Ergebnisse nebeneinander. Jede Spalte des Ergebnisses ist eine LK der Spaltenvektoren von  $A$ . ◇

**Beachte:** Man kann das Matrix-Matrix-Produkt  $AB$  nur dann bilden, wenn die Breite (Anzahl Spalten) von  $A$  übereinstimmt mit der Höhe (Anzahl Zeilen) von  $B$ .

Die Multiplikation einer Matrix mit der Einheitsmatrix (passender Größe) von links oder rechts lässt diese unverändert:  $AE_{n \times n} = E_{m \times m}A = A$  für  $A \in \mathbb{R}^{m \times n}$ .

### Beispiel 2.26 (Matrix-Matrix-Produkt)



**Beispiel 2.27 (Anwendung in der Produktionsplanung)**

In einem Betrieb werden aus den Rohstoffen  $R_1, \dots, R_4$  fünf Zwischenprodukte  $Z_1, \dots, Z_5$  hergestellt. Aus diesen Zwischenprodukten werden schließlich drei Endprodukte  $E_1, E_2, E_3$  gefertigt. In den folgenden Tabellen (Matrizen) sind die Bedarfe (Verbrauchsnormen) für die jeweiligen Produktionsschritte angegeben:

	$Z_1$	$Z_2$	$Z_3$	$Z_4$	$Z_5$		$E_1$	$E_2$	$E_3$
$R_1$	0	1	1	1	2	$Z_1$	1	1	1
$R_2$	5	0	1	2	1	$Z_2$	1	2	0
$R_3$	1	1	1	1	0	$Z_3$	0	1	1
$R_4$	0	2	0	1	0	$Z_4$	4	1	1
						$Z_5$	3	1	1
$A_{Z \rightarrow R} \in \mathbb{R}^{4 \times 5}$						$A_{E \rightarrow Z} \in \mathbb{R}^{5 \times 3}$			

**Lies:** Für eine Einheit des Zwischenprodukts  $Z_5$  werden zwei Einheiten  $R_1$  und eine Einheit  $R_2$  benötigt.

- (a) Das Matrix-Matrix-Produkt  $A_{E \rightarrow R} := A_{Z \rightarrow R} A_{E \rightarrow Z} \in \mathbb{R}^{4 \times 3}$  gibt direkt die Bedarfe an Rohstoffen  $R_1, \dots, R_4$  für die Endprodukte  $E_1, E_2, E_3$  an:

$$A_{E \rightarrow R} := A_{Z \rightarrow R} A_{E \rightarrow Z} = \begin{pmatrix} 0 & 1 & 1 & 1 & 2 \\ 5 & 0 & 1 & 2 & 1 \\ 1 & 1 & 1 & 1 & 0 \\ 0 & 2 & 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 1 & 1 \\ 1 & 2 & 0 \\ 0 & 1 & 1 \\ 4 & 1 & 1 \\ 3 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 11 & 6 & 4 \\ 16 & 9 & 9 \\ 6 & 5 & 3 \\ 6 & 5 & 1 \end{pmatrix}.$$



für die Unbekannten (Variablen)  $x_1, \dots, x_n \in \mathbb{R}$  heißt ein **lineares Gleichungssystem (LGS)** mit  $m$  Gleichungen.

Nach Definition 2.22 können wir es kurz als  $A\vec{x} = \vec{b}$  schreiben. Dabei ist

$A = (a_{ij}) \in \mathbb{R}^{m \times n}$  die **Koeffizientenmatrix**

$\vec{x} = (x_1, \dots, x_n)^\top \in \mathbb{R}^n$  der **Vektor der Unbekannten** oder **Variablen**

$\vec{b} = (b_1, \dots, b_m)^\top \in \mathbb{R}^m$  der **Vektor der rechten Seite**.

Ein Vektor  $\vec{x} \in \mathbb{R}^n$ , der  $A\vec{x} = \vec{b}$  erfüllt, heißt eine **Lösung** des LGS.  $\diamond$

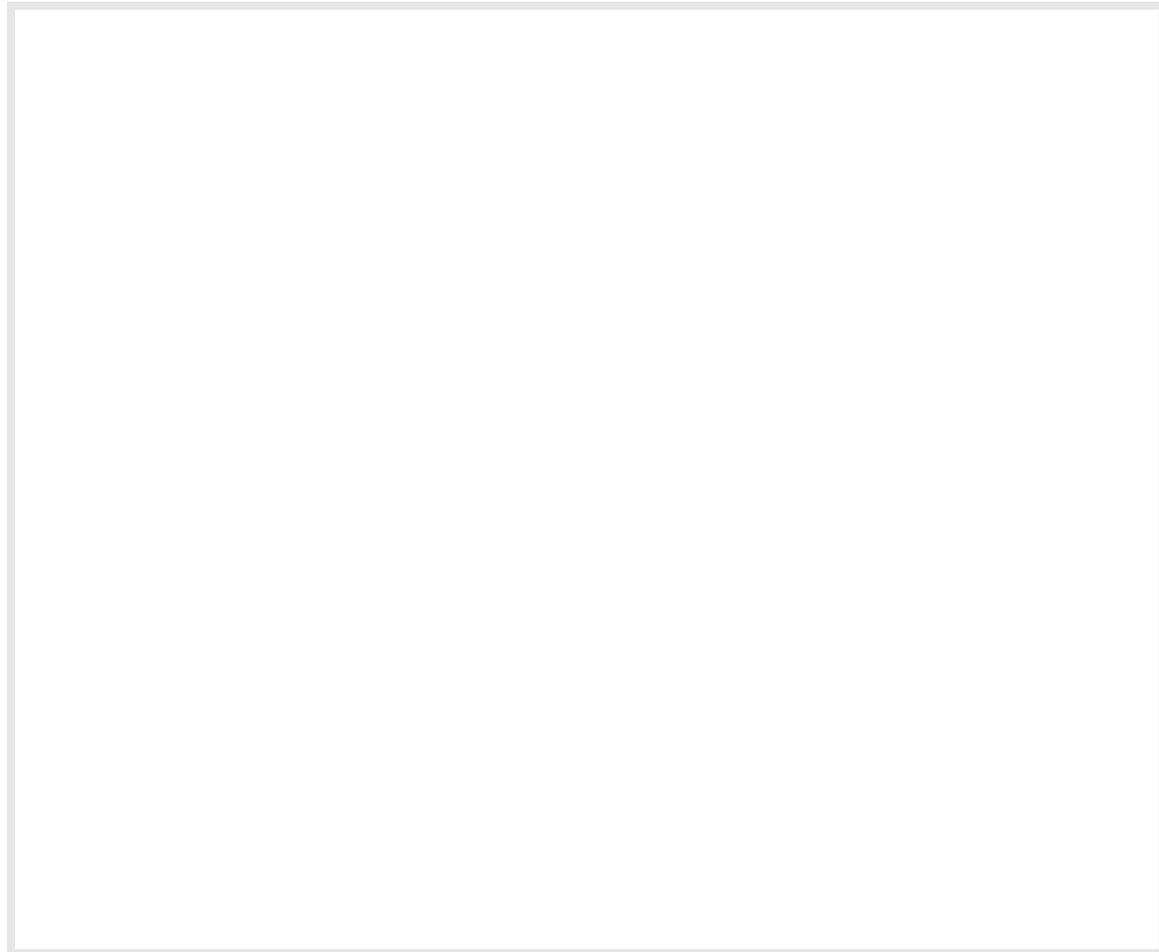
**Bemerkung 2.31** (a) Jede Zeile in  $A\vec{x} = \vec{b}$  entspricht einer Gleichung.

(b) Jede Spalte in  $A$  gehört zu einer der Variablen  $x_1, \dots, x_n$ .

Ende 7. V

25.11.2010

**Beispiel 2.32** (Lösungen linearer Gleichungssysteme)



**Frage:** Wie kann man *eine/alle* Lösungen eines LGS berechnen bzw. dessen Unlösbarkeit feststellen?

### § 2.4.1 Das homogene lineare Gleichungssystem

Für gegebenes  $A \in \mathbb{R}^{m \times n}$  untersuchen wir zunächst das **homogene LGS**  $A\vec{x} = \vec{0}$ .

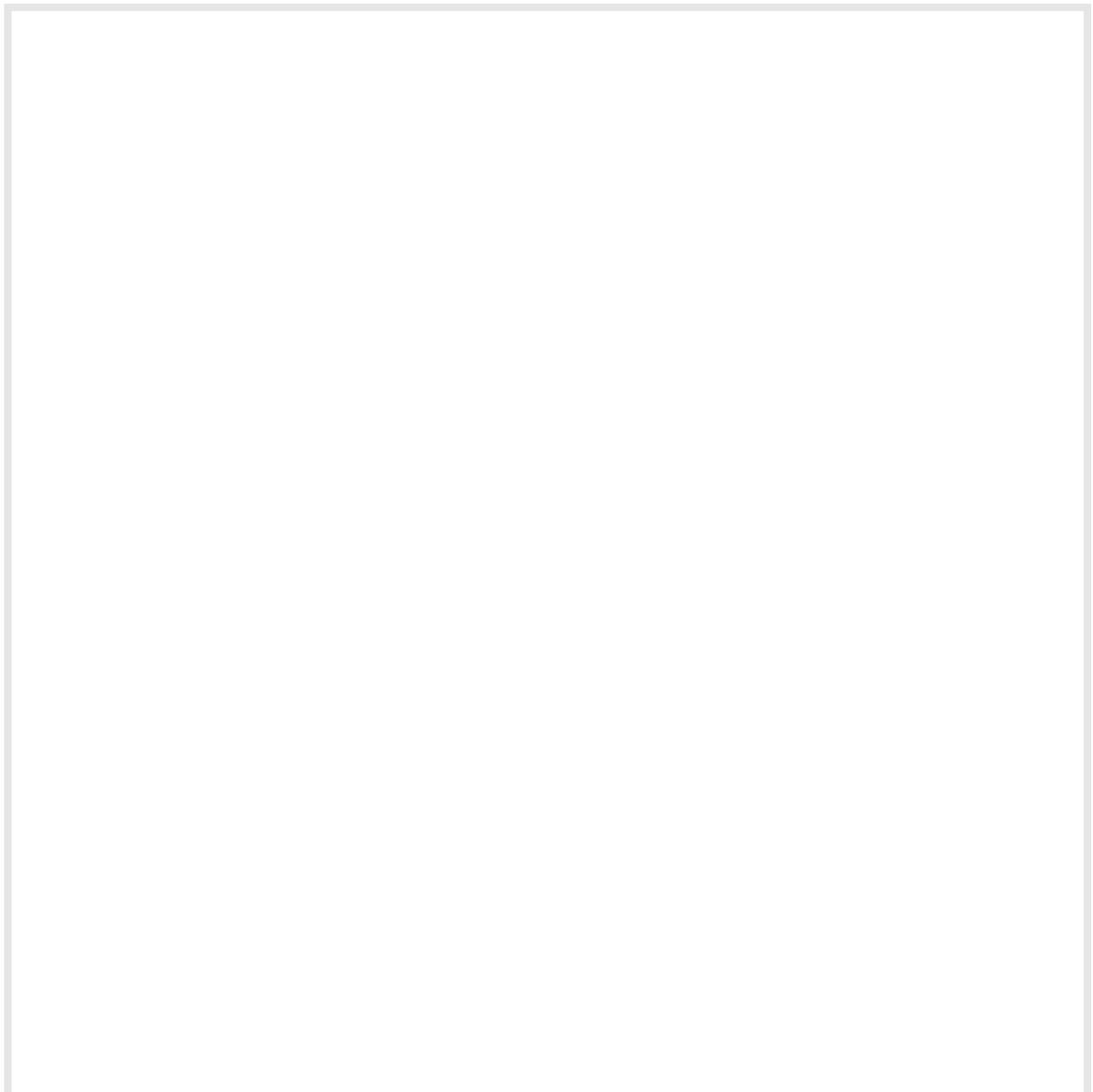
**Beachte:** Der Nullvektor  $\vec{0} \in \mathbb{R}^n$  ist immer eine Lösung. Die Lösungsmenge ändert sich nicht durch folgende Operationen (**Äquivalenzumformungen**):

- (a) Vertauschen zweier Gleichungen (Zeilen)
- (b) Multiplikation einer Gleichung (Zeile) mit einer Zahl  $\alpha \neq 0$
- (c) Addition des Vielfachen einer Gleichung (Zeile) zu einer anderen.

Diese Operationen sind Grundlage des **Gauß'schen Lösungsverfahrens**, in dem das LGS so umgeformt wird (**Vorwärtselimination**), dass man die Lösung(en) leicht ablesen kann (**Rückwärtssubstitution**).

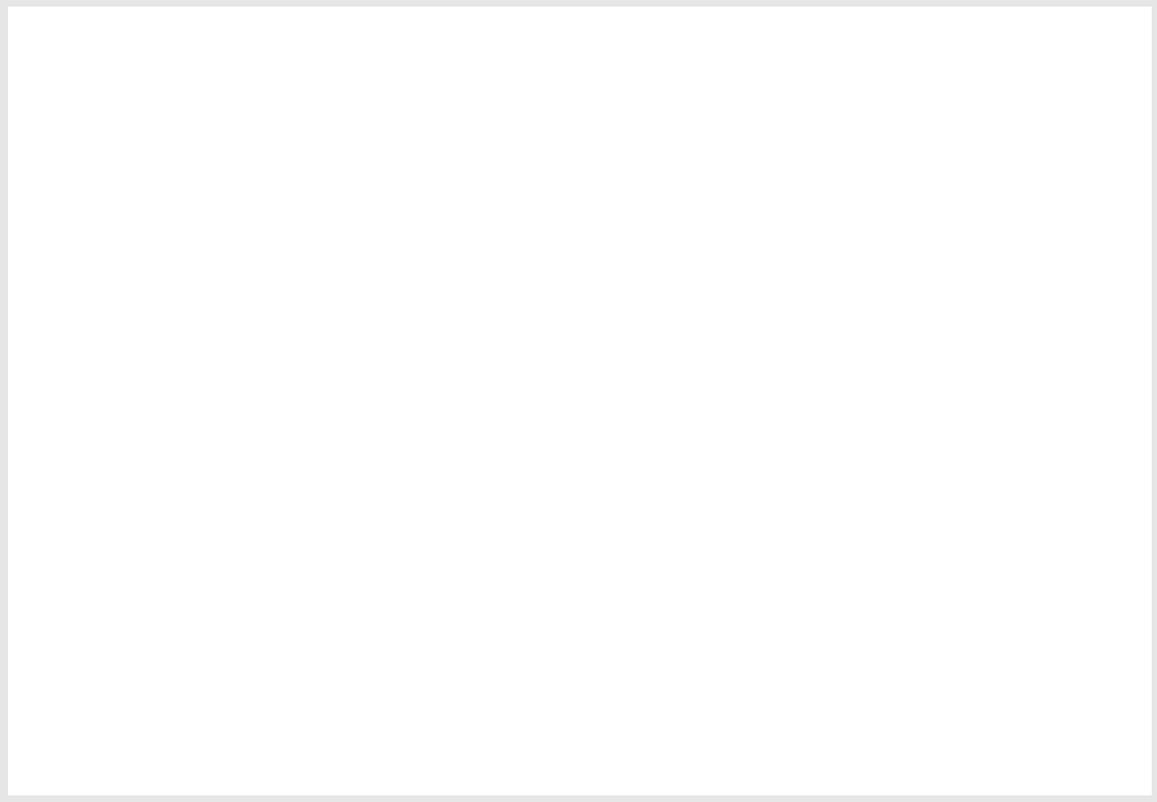
**Beispiel 2.33** (Vorwärtselimination)

- (a)

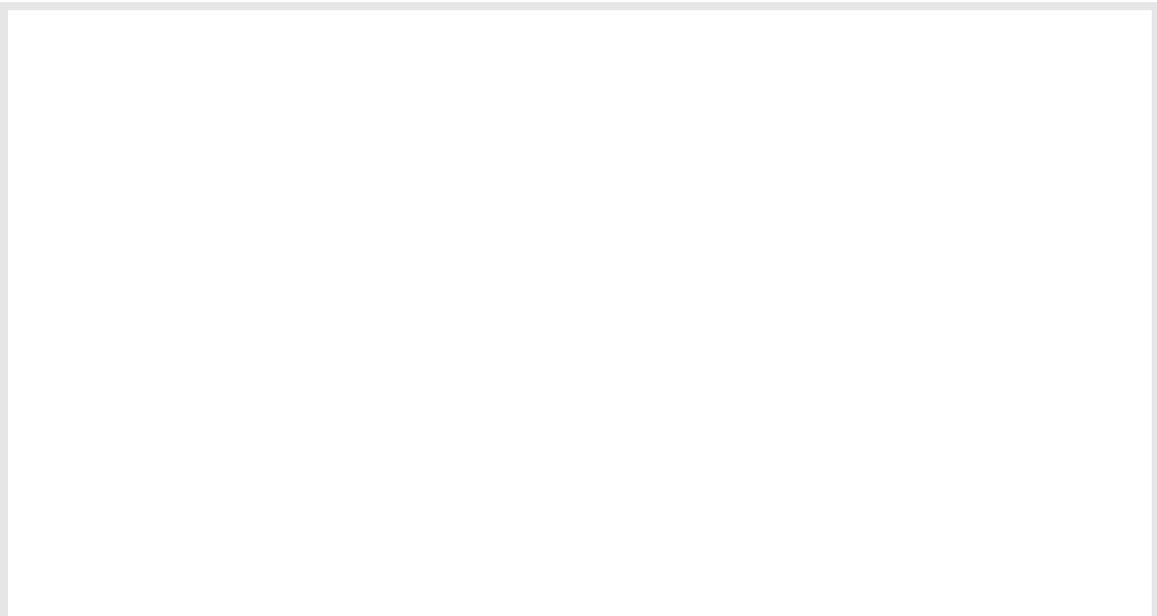




(b)



(c)



◇

Bei der Vorwärtselimination entsteht aus  $A \in \mathbb{R}^{m \times n}$  eine Matrix in **Zeilenstufenform (ZSF)**, z.B.

$$\begin{pmatrix} \bullet & * & * \\ 0 & \bullet & * \\ 0 & 0 & \bullet \end{pmatrix} \text{ im Fall (a) und (b) sowie } \begin{pmatrix} \bullet & * & * \\ 0 & 0 & \bullet \\ 0 & 0 & 0 \end{pmatrix} \text{ im Fall (c).}$$

Dabei bezeichnet einen  $\bullet$  Eintrag  $\neq 0$  und  $*$  irgendwelche Einträge. Weitere Beispiele für Matrizen in ZSF sind

$$\begin{pmatrix} 0 & \bullet & * & * & * \\ 0 & 0 & \bullet & * & * \\ 0 & 0 & 0 & 0 & \bullet \end{pmatrix}, \quad \begin{pmatrix} \bullet & * & * & * & * \\ 0 & \bullet & * & * & * \\ 0 & 0 & \bullet & * & * \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} \bullet & * \\ 0 & \bullet \\ 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

Kennzeichen der ZSF sind:

- (a) In jeder Zeile stehen links von  $\bullet$  nur Nullen.
- (b) Pro Zeile rückt  $\bullet$  mindestens um eine Stelle nach rechts.
- (c) Unterhalb eines  $\bullet$  stehen in derselben Spalte nur Nullen.

### Beispiel 2.34 (Rückwärtssubstitution)

Wir fahren fort in Beispiel 2.33.

- (a) Aus dem in ZSF gebrachten System lesen wir ab:

$$\left( \begin{array}{ccc|c} 4 & 0 & -2 & 0 \\ 0 & -2 & 2 & 0 \\ 0 & 0 & \frac{19}{2} & 0 \end{array} \right)$$

lesen wir von unten nach oben ab:

$$\begin{aligned} \frac{19}{2}x_3 &= 0 \Rightarrow x_3 = 0 \\ -2x_2 + 2x_3 &= 0 \Rightarrow x_2 = 0 \\ 4x_1 - 2x_3 &= 0 \Rightarrow x_1 = 0. \end{aligned}$$

Also ist  $\vec{x} = (0, 0, 0)^\top$  die einzige Lösung des Systems  $A\vec{x} = \vec{0}$ .

- (b) Auch in diesem Fall ist  $\vec{x} = \vec{0}$  die einzige Lösung.

- (c) Das LGS in ZSF

$$\left( \begin{array}{ccc|c} 4 & -2 & -2 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right) \text{ mit dem Muster } \begin{pmatrix} \bullet & * & * \\ 0 & 0 & \bullet \\ 0 & 0 & 0 \end{pmatrix}$$

bedeutet ausgeschrieben

$$\begin{aligned}4x_1 - 2x_2 - 2x_3 &= 0 \\ 2x_3 &= 0.\end{aligned}$$

Diejenigen Spalten (Variablen), in denen *kein*  $\bullet$  vorkommt, sind die **freien Variablen**, hier  $x_2$ . Ihnen ordnen wir freie Parameter zu:

$$\boxed{x_2} = \lambda \quad \text{mit } \lambda \in \mathbb{R}.$$

Diejenigen Spalten (Variablen) mit  $\bullet$  sind die **abhängigen Variablen**, hier  $x_1$  und  $x_3$ . Sie ergeben sich eindeutig aus den Werten der freien Variablen. Wir berechnen sie von hinten nach vorne:

$$2\boxed{x_3} = 0 \quad \Rightarrow \quad x_3 = 0$$

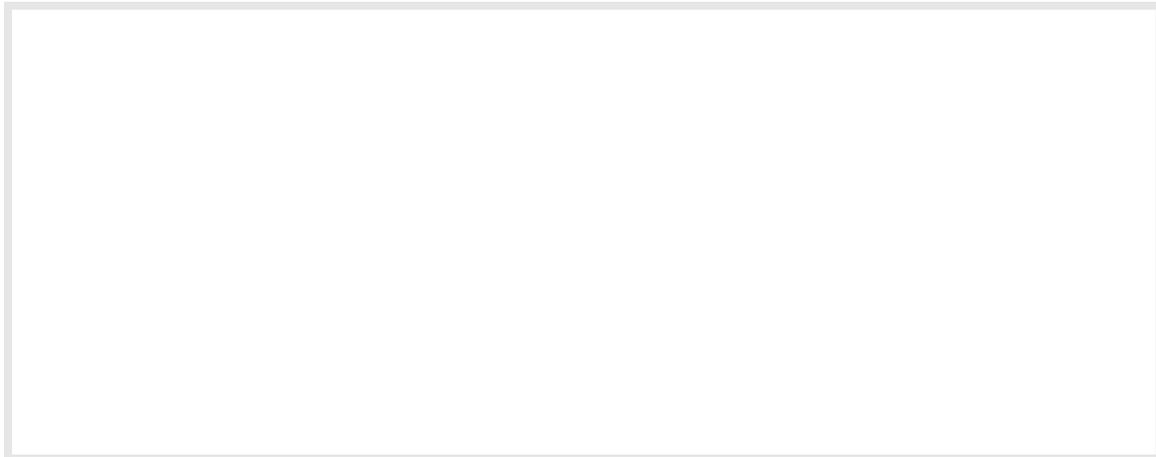
$$4\boxed{x_1} - 2x_2 - 2x_3 = 0 \quad \Rightarrow \quad 4x_1 = 2x_2 + 2x_3 = 2\lambda + 2 \cdot 0 \quad \Rightarrow \quad x_1 = \frac{1}{2}\lambda.$$

Die **allgemeine Lösung** des homogenen LGS aus Beispiel 2.33 (c) lautet also:

$$\left. \begin{array}{l} x_1 = 1/2 \lambda \\ x_2 = \lambda \\ x_3 = 0 \end{array} \right\} \quad \text{oder kurz} \quad \vec{x} = \lambda \begin{pmatrix} 1/2 \\ 1 \\ 0 \end{pmatrix}.$$

Mit anderen Worten: Die Lösungsmenge besteht aus allen Vielfachen des Vektors  $(1/2, 1, 0)^\top$ . Die Lösungsmenge ist damit ein Unterraum des  $\mathbb{R}^3$  der Dimension 1.

**Probe:**



◇

**Frage:** Wieviele freie Variablen hat man bei einem homogenen LGS?

**Definition 2.35 (Rang einer Matrix)**

Der **Rang** einer Matrix  $A \in \mathbb{R}^{m \times n}$  ist die Anzahl der  $\bullet$  in der zu  $A$  gehörigen Zeilenstufenform. ◇

**Satz 2.36 (zur Lösungsmenge eines homogenen LGS)**

Wir betrachten das homogene LGS  $A\vec{x} = \vec{0}$  mit  $A \in \mathbb{R}^{m \times n}$ .

- (a) Insgesamt hat das LGS  $n$  Variablen (Spalten). Jeder  $\bullet$  in der ZSF von  $A$  steht für eine abhängige Variable. Es bleiben also  $n - \text{Rang}(A)$  freie Variablen, die als freie Parameter in der allgemeinen Lösung erscheinen.
- (b) Die Lösungsmenge des homogenen LGS  $A\vec{x} = \vec{0}$  ist ein Unterraum des  $\mathbb{R}^n$  der Dimension  $n - \text{Rang}(A)$ .
- (c)  $\text{Rang}(A) = n \Leftrightarrow$  das homogene LGS  $A\vec{x} = \vec{0}$  hat nur die triviale Lösung  $\vec{x} = \vec{0}$ .
- (d)  $\text{Rang}(A)$  ist die Anzahl der für das LGS  $A\vec{x} = \vec{0}$  tatsächlich relevanten Gleichungen.

Ende 8. V

02.12.2010

**Bemerkung 2.37** (zum Rang einer Matrix)

Sei  $A \in \mathbb{R}^{m \times n}$ .

- (a) Es gilt  $\text{Rang}(A) \leq \min\{m, n\}$ .
- (b)  $\text{Rang}(A) =$  maximale Anzahl linear unabhängiger Zeilenvektoren von  $A$ .
- (c)  $\text{Rang}(A) =$  maximale Anzahl linear unabhängiger Spaltenvektoren von  $A$ .
- (d)  $\text{Rang}(A) = \text{Rang}(A^\top)$ .

**Beispiel 2.38** (Rang einer Matrix, vgl. Beispiel 2.33)

Wir bestimmen den Rang der Matrizen in Beispiel 2.33.

- (a)

- (b)

(c)

◇

### § 2.4.2 Das inhomogene lineare Gleichungssystem

Bei einem inhomogenen LGS  $A\vec{x} = \vec{b}$  gehen wir genauso wie oben vor. Während der Vorwärtselemination entsteht wieder eine ZSF, z.B.

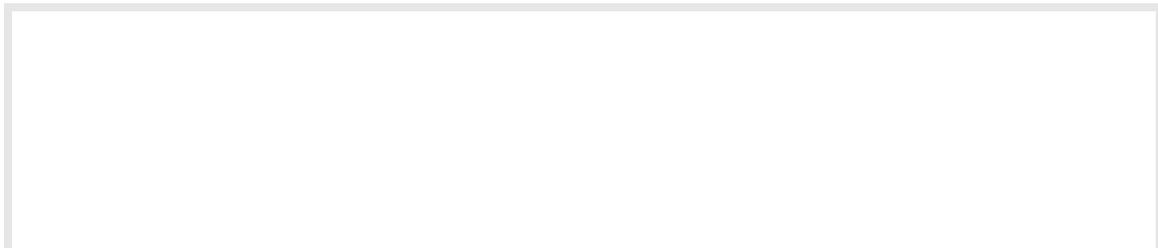
$$\left( \begin{array}{ccc|c} \bullet & * & * & * \\ 0 & 0 & \bullet & * \\ 0 & 0 & 0 & * \end{array} \right).$$

**Satz 2.39 (Lösbarkeit)**

Das LGS  $A\vec{x} = \vec{b}$  ist genau dann lösbar, wenn in jeder Nullzeile der zugehörigen ZSF auf der rechten Seite ebenfalls eine null steht. ◇

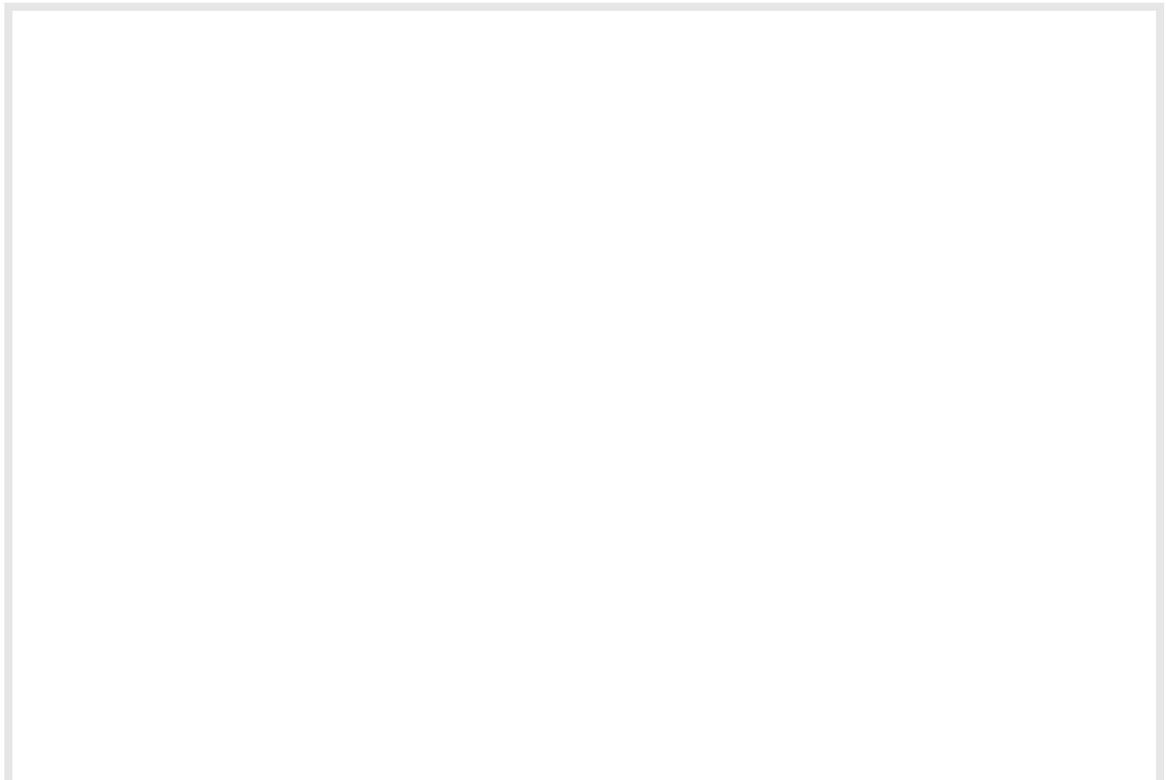
**Beispiel 2.40 (Lösung inhomogener LGS)**

(a)





(b)



**Bemerkung 2.41**

- (a) Die allgemeine Lösung setzt sich additiv zusammen
- aus irgendeiner **speziellen (partikulären) Lösung** von  $A\vec{x} = \vec{b}$ , im Beispiel (b):  $\vec{x} = (1, 0, -1)^\top$ ,
  - und der allgemeinen Lösung des zugehörigen homogenen Systems  $A\vec{x} = \vec{0}$ , im Beispiel (b):  $\lambda(1/2, 1, 0)^\top$ .

Dies ist das **Prinzip der Superposition**.

- (b) Die Lösungsmenge bildet also einen um den Vektor  $(1, 0, -1)^\top$  „verschobenen Unterraum“. Im Beispiel (b) ist  $n = 3$  und  $\text{Rang}(A) = 2$ . Daher gibt es  $n - \text{Rang}(A) = 3 - 2 = 1$  freie Parameter (hier  $\lambda$  genannt).
- (c) Lineare Gleichungssysteme  $A\vec{x} = \vec{b}$ , in denen die Matrix  $A$  oder die rechte Seite  $\vec{b}$  aus komplexen Zahlen bestehen, können genauso gelöst werden. Es ergibt sich dann i.A. ein Lösungsvektor  $\vec{x} \in \mathbb{C}^n$  mit komplexen Einträgen. ◊

**§ 2.5 Inverse Matrizen und Determinanten**

In diesem Abschnitt ist  $A \in \mathbb{R}^{n \times n}$  eine *quadratische* Matrix.

**Definition 2.42 (Inverse Matrix)** (a) Falls zu  $A \in \mathbb{R}^{n \times n}$  eine Matrix  $A^{-1} \in \mathbb{R}^{n \times n}$  existiert mit der Eigenschaft

$$A A^{-1} = A^{-1} A = E,$$

so nennt man  $A^{-1}$  die (eindeutige) **inverse Matrix** zu  $A$ .

- (b) Falls zu  $A \in \mathbb{R}^{n \times n}$  die inverse Matrix  $A^{-1}$  existiert, so heißt  $A$  **invertierbar** oder **regulär**, andernfalls **nicht invertierbar** oder **singulär**.  $\diamond$

**Beachte:** Für invertierbares  $A$  gilt  $(A^{-1})^{-1} = A$ .

**Satz 2.43 (Zusammenhang mit linearen Gleichungssystemen)**

Eine Matrix  $A \in \mathbb{R}^{n \times n}$  ist invertierbar genau dann, wenn das LGS  $A\vec{x} = \vec{b}$  für jede beliebige rechte Seite  $\vec{b} \in \mathbb{R}^n$  immer eindeutig lösbar ist. Die Lösung ist dann

$$A\vec{x} = \vec{b} \Leftrightarrow \underbrace{A^{-1}A}_{=E}\vec{x} = A^{-1}\vec{b} \Leftrightarrow \boxed{\vec{x} = A^{-1}\vec{b}}$$

**Bedeutung:** Wenn man  $A^{-1}$  kennt, so kann man *alle* LGS  $A\vec{x} = \vec{b}$  mit beliebiger rechter Seite  $\vec{b}$  einfach durch eine Matrix-Vektor-Multiplikation lösen.  $\diamond$

**Folgerung 2.44**

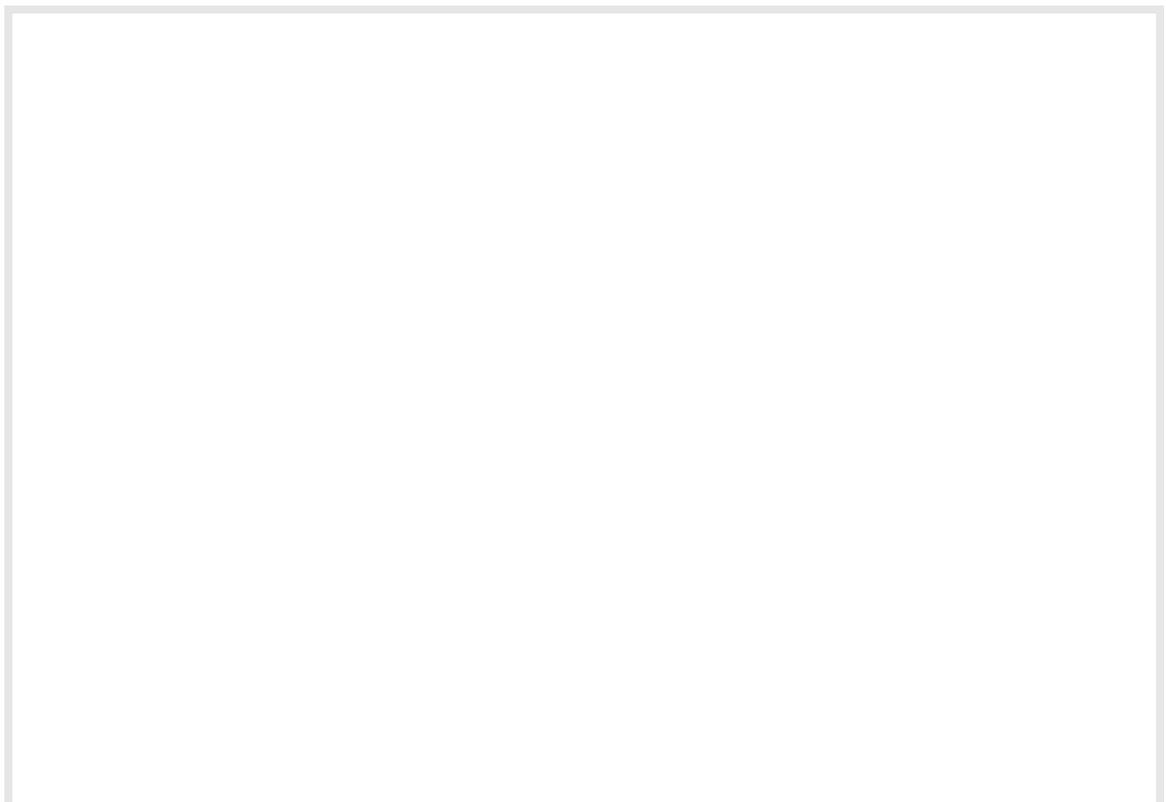
$A \in \mathbb{R}^{n \times n}$  ist invertierbar

$$\Leftrightarrow \text{die ZSF von } A \text{ hat das Muster } \begin{pmatrix} \bullet & * & \cdots & * \\ 0 & \bullet & \ddots & \vdots \\ \vdots & \ddots & \ddots & * \\ 0 & \cdots & 0 & \bullet \end{pmatrix}$$

$$\Leftrightarrow A \text{ hat vollen Rang, d.h. } \text{Rang}(A) = n. \quad \diamond$$

**Beispiel 2.45 (Berechnung der inversen Matrix)**

(a)





(b)

Ende 9. V

09.12.2010

**Definition 2.46 (Determinante)**Für  $A \in \mathbb{R}^{n \times n}$ ,  $A = (a_{ij})$ , definieren wir:(a)  $n = 1$ :

$$\det(A) = a_{11}$$

(b)  $n = 2$ :

$$\det(A) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11} a_{22} - a_{21} a_{12}$$

(c)  $n = 3$ :

$$\det(A) = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11} a_{22} a_{33} + a_{12} a_{23} a_{31} + a_{13} a_{21} a_{32} \\ - a_{31} a_{22} a_{13} - a_{32} a_{23} a_{11} - a_{33} a_{21} a_{12}.$$

**Merkregel (Regel von Sarrus)**, nur für  $3 \times 3$ -Matrizen:

$$\det(A) = \begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{11} & a_{12} \\ a_{21} & a_{22} & a_{23} & a_{21} & a_{22} \\ a_{31} & a_{32} & a_{33} & a_{31} & a_{32} \end{vmatrix}$$

(d) Für  $n \geq 4$  wird die Determinante mit Hilfe eines **Entwicklungssatzes** berechnet. Beispiel: Entwicklung nach der ersten Spalte:

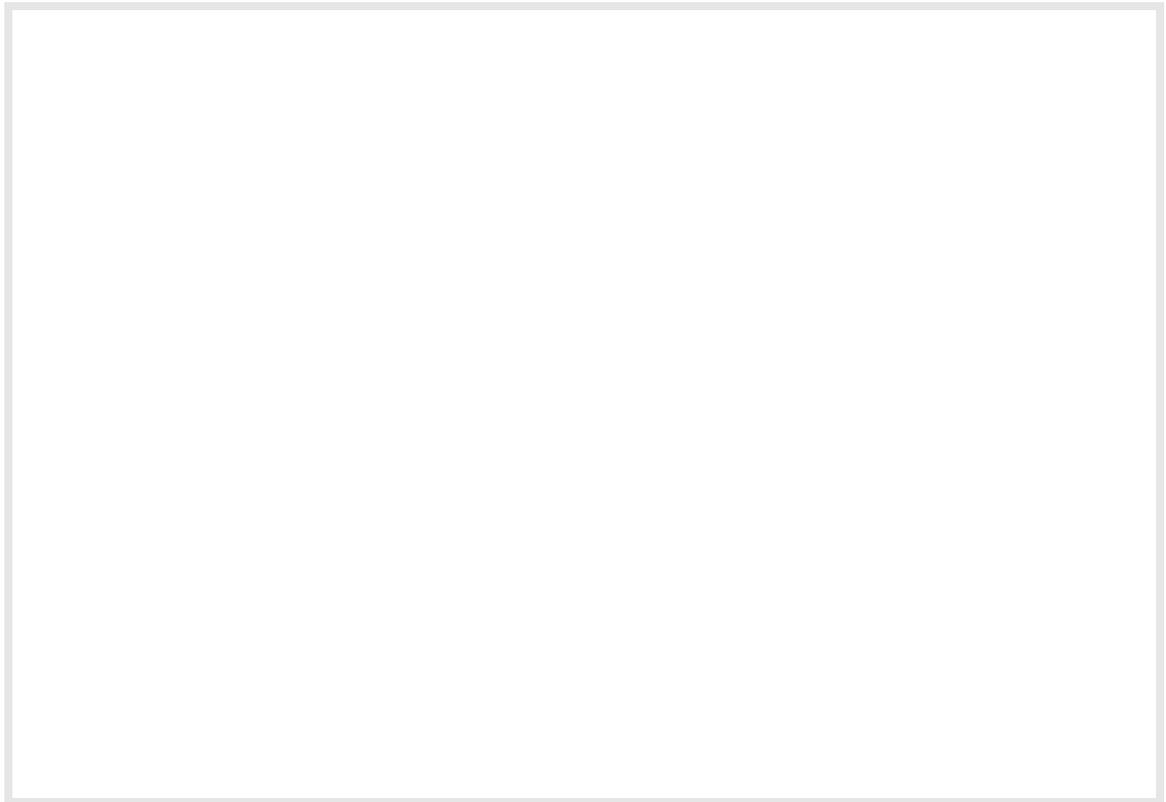
$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{vmatrix} = +a_{11} \begin{vmatrix} a_{22} & a_{23} & a_{24} \\ a_{32} & a_{33} & a_{34} \\ a_{42} & a_{43} & a_{44} \end{vmatrix} - a_{21} \begin{vmatrix} a_{12} & a_{13} & a_{14} \\ a_{32} & a_{33} & a_{34} \\ a_{42} & a_{43} & a_{44} \end{vmatrix} \\ + a_{31} \begin{vmatrix} a_{12} & a_{13} & a_{14} \\ a_{22} & a_{23} & a_{24} \\ a_{42} & a_{43} & a_{44} \end{vmatrix} - a_{41} \begin{vmatrix} a_{12} & a_{13} & a_{14} \\ a_{22} & a_{23} & a_{24} \\ a_{32} & a_{33} & a_{34} \end{vmatrix}.$$

Die Determinanten der verbleibenden  $3 \times 3$ -Matrizen kann man wieder mit der Regel von Sarrus berechnen.Analog kann man nach einer Zeile entwickeln. Bei Entwicklungen nach Spalten/Zeilen mit *ungerader* Nummer (hier: 1) beginnt man mit +, bei *gerader* Nummer mit - („Schachbrettmuster“). Bei Entwicklung nach der zweiten

Spalte ergibt sich also

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{vmatrix} = -a_{12} \begin{vmatrix} a_{21} & a_{23} & a_{24} \\ a_{31} & a_{33} & a_{34} \\ a_{41} & a_{43} & a_{44} \end{vmatrix} + a_{22} \begin{vmatrix} a_{11} & a_{13} & a_{14} \\ a_{31} & a_{33} & a_{34} \\ a_{41} & a_{43} & a_{44} \end{vmatrix} \\ - a_{32} \begin{vmatrix} a_{11} & a_{13} & a_{14} \\ a_{21} & a_{23} & a_{24} \\ a_{41} & a_{43} & a_{44} \end{vmatrix} + a_{42} \begin{vmatrix} a_{11} & a_{13} & a_{14} \\ a_{21} & a_{23} & a_{24} \\ a_{31} & a_{33} & a_{34} \end{vmatrix}. \quad \diamond$$

### Beispiel 2.47 (Determinanten)



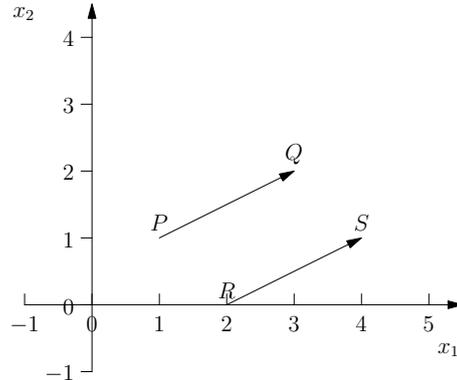
### Satz 2.48 (Bedeutung und Rechenregeln für die Determinante)

Für  $A, B \in \mathbb{R}^{n \times n}$  und  $\alpha \in \mathbb{R}$  gilt:

- (a)  $A$  ist invertierbar  $\Leftrightarrow \det(A) \neq 0$ .
- (b) Ist  $A$  invertierbar, dann gilt  $\det(A^{-1}) = \frac{1}{\det(A)}$ .
- (c)  $\det(A) = \det(A^\top)$
- (d)  $\det(AB) = \det(A) \det(B)$
- (e) Ist  $A = (a_{ij})$  eine obere oder untere Dreiecksmatrix (oder sogar eine Diagonalmatrix), so gilt  $\det(A) = a_{11} a_{22} \cdots a_{nn}$ , also das Produkt der Hauptdiagonal-Elemente.
- (f)  $\det(\alpha A) = \alpha^n \det(A)$  \(\diamond\)

§ 2.6 Analytische Geometrie in der Ebene

In der ebenen Geometrie verwenden wir ein rechtwinkliges Koordinatensystem mit Ursprung 0 und  $x_1$ - und  $x_2$ -Achse. Jeder Punkt  $P$  der Ebene kann in diesem Koordinatensystem dargestellt werden. Die **Verschiebung** eines Punktes  $P$  in einen Punkt  $Q$  wird mit  $\overrightarrow{PQ}$  bezeichnet und heißt **Vektor** von  $P$  nach  $Q$ . Zwei gleich lange, parallele und gleich gerichtete Pfeile (im Bild  $\overrightarrow{PQ}$  und  $\overrightarrow{RS}$ ) sind verschiedene Darstellungen desselben Vektors.



Der Vektor  $\vec{a} = \overrightarrow{0A}$  heißt **Ortsvektor** des Punktes  $A$ .

**Beispiel 2.49** (derselbe Vektor mit zwei Pfeilen)

Für die Punkte  $P = (1, 1)^\top$ ,  $Q = (3, 2)^\top$  sowie  $R = (2, 0)^\top$  und  $S = (4, 1)^\top$  gilt:

$$\begin{aligned} \overrightarrow{PQ} &= Q - P = \begin{pmatrix} 3 \\ 2 \end{pmatrix} - \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix} \\ \overrightarrow{RS} &= S - R = \begin{pmatrix} 4 \\ 1 \end{pmatrix} - \begin{pmatrix} 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}. \end{aligned}$$

Die Pfeile  $\overrightarrow{PQ}$  und  $\overrightarrow{RS}$  repräsentieren also denselben Vektor. Seine Länge (Betrag)

$$\|(2, 1)^\top\| = \sqrt{5}$$

entspricht dem **Abstand** der Punkte  $P$  und  $Q$  (und  $R$  und  $S$ ). ◇

Die Addition von Vektoren entspricht grafisch dem Aneinanderfügen von Pfeilen.

**Definition 2.50** (Projektion, orthogonale Zerlegung)

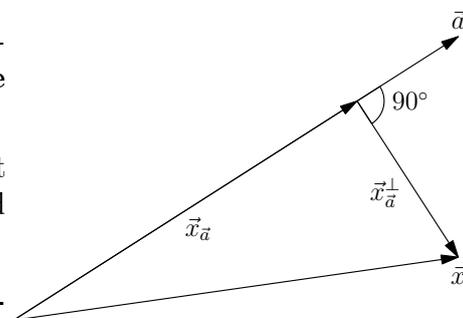
Es sei  $\vec{a} \in \mathbb{R}^2$  oder allgemein  $\vec{a} \in \mathbb{R}^n$  gegeben,  $\vec{a} \neq \vec{0}$ . Jeder Vektor  $\vec{x} \in \mathbb{R}^n$  kann eindeutig zerlegt werden in seine zu  $\vec{a}$  parallele (bzw. anti-parallele) Komponente  $\vec{x}_{\vec{a}}$  sowie die dazu senkrechte Komponente:

- (a) Für  $\vec{x} \in \mathbb{R}^n$  heißt

$$\vec{x}_{\vec{a}} := \frac{\vec{x} \cdot \vec{a}}{\|\vec{a}\|^2} \vec{a}$$

die **Projektion** von  $\vec{x}$  auf oder in Richtung von  $\vec{a}$ .  $\vec{x}_{\vec{a}}$  heißt die zu  $\vec{a}$  **parallele Komponente** von  $\vec{x}$ .

**Beachte:** Der Koeffizient  $(\vec{x} \cdot \vec{a})/\|\vec{a}\|^2$  ist positiv, wenn der Winkel zwischen  $\vec{x}$  und  $\vec{a}$  kleiner als  $90^\circ$  ist (wie in der Skizze).



- (b) Der Rest  $\vec{x}_{\vec{a}}^\perp := \vec{x} - \vec{x}_{\vec{a}}$  heißt die zu  $\vec{a}$  **senkrechte Komponente** von  $\vec{x}$ .
- (c) Die Darstellung  $\vec{x} = \vec{x}_{\vec{a}} + \vec{x}_{\vec{a}}^\perp$  heißt die **orthogonale Zerlegung** von  $\vec{x}$  bzgl.  $\vec{a}$ . ◇

**Satz 2.51** (a)  $\vec{x}_{\vec{a}}$  ist parallel zu  $\vec{a}$ .

(b) Der Rest  $\vec{x}_{\vec{a}}^{\perp} := \vec{x} - \vec{x}_{\vec{a}}$  steht senkrecht auf  $\vec{a}$ .

(c)  $\vec{a}_{\vec{a}} = \vec{a}$  (Projektion auf sich selbst). ◇

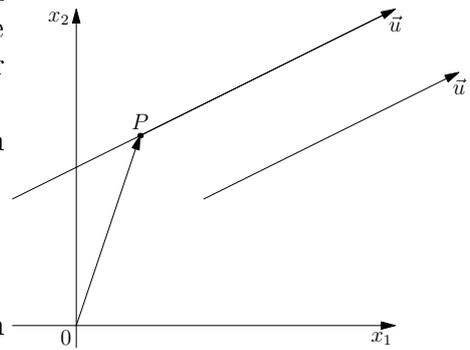
Zwei Punkte  $P \neq Q$  in der Ebene mit Ortsvektoren  $\vec{p}$  und  $\vec{q}$  legen eindeutig eine Gerade fest. Dasselbe gilt für einen Punkt  $P$  und einen Richtungsvektor  $\vec{u} \in \mathbb{R}^2$ ,  $\vec{u} \neq \vec{0}$ .

Ein Punkt  $X$  mit Ortsvektor  $\vec{x}$  liegt genau dann auf der Geraden, wenn  $\vec{x}$  die Darstellung

$$\vec{x} = \vec{p} + \lambda(\vec{q} - \vec{p}) \quad \textbf{Zwei-Punkte-Form} \quad \text{bzw.}$$

$$\vec{x} = \vec{p} + \lambda\vec{u} \quad \textbf{Punkt-Richtungs-Form}$$

mit einem Parameter  $\lambda \in \mathbb{R}$  hat. Beide Formen sind **Parameterdarstellungen** einer Geraden.



Wir bestimmen nun die Schnittpunkte zweier Geraden

$$g_1 : \vec{x} = \vec{p}_1 + \lambda_1 \vec{u}$$

$$g_2 : \vec{x} = \vec{p}_2 + \lambda_2 \vec{v}$$

in der Ebene, wobei  $\vec{p}_1, \vec{p}_2, \vec{u}, \vec{v} \in \mathbb{R}^2$  gegeben sind,  $\vec{u}, \vec{v} \neq \vec{0}$ .

**Ansatz:**

$$\begin{aligned} \vec{p}_1 + \lambda_1 \vec{u} &= \vec{p}_2 + \lambda_2 \vec{v} \\ \Leftrightarrow \lambda_1 \vec{u} - \lambda_2 \vec{v} &= \vec{p}_2 - \vec{p}_1 \\ \Leftrightarrow \begin{pmatrix} u_1 & -v_1 \\ u_2 & -v_2 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} &= \vec{p}_2 - \vec{p}_1 \end{aligned}$$

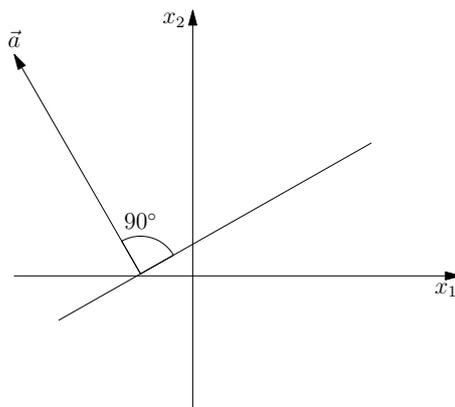
Dies ist ein LGS mit zwei Gleichungen für die beiden Unbekannten  $\lambda_1, \lambda_2$ . Folgende Fälle können auftreten:

- (a) Die Matrix hat Rang 2, d.h., die Richtungsvektoren  $\vec{u}$  und  $\vec{v}$  sind linear unabhängig (also nicht parallel oder anti-parallel). Es gibt eine eindeutige Lösung  $(\lambda_1, \lambda_2)$ , d.h., einen eindeutigen Schnittpunkt.
- (b) Die Matrix hat Rang 1, d.h., die Richtungsvektoren  $\vec{u}$  und  $\vec{v}$  sind linear abhängig (also parallel bzw. anti-parallel). Dann gibt es entweder unendlich viele Lösungen  $(\lambda_1, \lambda_2)$  mit einer freien Variablen (die Geraden sind identisch) oder keine Lösung (die Geraden sind parallel, aber nicht identisch).

Nach Satz 2.36 und Beispiel 2.40 wissen wir, dass die Lösungsmenge eines LGS in  $n$  Unbekannten ein verschobener Unterraum des  $\mathbb{R}^n$  der Dimension  $n - \text{Rang}(A)$  ist. Eine Gerade ist ein verschobener Unterraum der Dimension 1. Wir können deshalb eine Gerade in der Ebene ( $n = 2$ ) auch als Lösungsmenge eines LGS vom Rang 1 beschreiben, also durch eine Gleichung der Form

$$a_1 x_1 + a_2 x_2 = b \quad \text{bzw.} \quad \vec{a}^\top \vec{x} = b$$

mit  $\vec{a} = (a_1, a_2)^\top \neq \vec{0}$ . Dies ist eine **parameterfreie Darstellung** einer Geraden in der Ebene.



**Bemerkung 2.52**

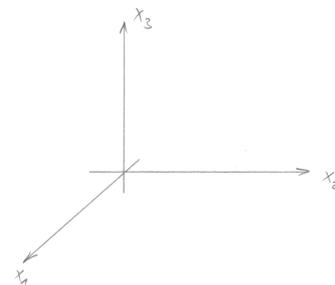
Falls  $a_2 \neq 0$  ist, so erhält man durch Division die bekannte **Normalform** der Geradengleichung in der Ebene:

$$x_2 = m x_1 + n \quad \text{bzw.} \quad y = m x + n.$$

Geraden, die parallel zur  $x_2$ -Achse sind, kann man so jedoch nicht darstellen.  $\diamond$

**§ 2.7 Analytische Geometrie im Raum**

In der räumlichen Geometrie verwenden wir ein rechtwinkliges Koordinatensystem mit Ursprung 0 und  $x_1$ -,  $x_2$ - und  $x_3$ -Achse, die ein **Rechtssystem** bilden. Jeder Punkt  $P \in \mathbb{R}^3$  kann in diesem Koordinatensystem dargestellt werden.



**Definition 2.53 (Kreuzprodukt)**

Für Vektoren  $\vec{a}, \vec{b} \in \mathbb{R}^3$  definieren wir das **Kreuzprodukt** oder **Vektorprodukt**

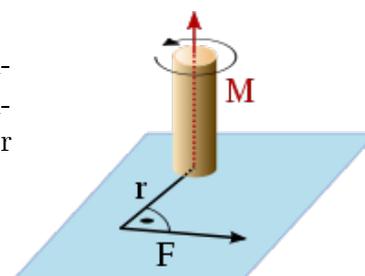
$$\vec{a} \times \vec{b} = \begin{pmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{pmatrix} \in \mathbb{R}^3.$$

**Merkregel:** Das Kreuzprodukt kann formal als Determinante geschrieben werden:

$$\vec{a} \times \vec{b} = \begin{vmatrix} \vec{e}_1 & a_1 & b_1 \\ \vec{e}_2 & a_2 & b_2 \\ \vec{e}_3 & a_3 & b_3 \end{vmatrix} = (a_2 b_3 - a_3 b_2) \vec{e}_1 + (a_3 b_1 - a_1 b_3) \vec{e}_2 + (a_1 b_2 - a_2 b_1) \vec{e}_3. \quad \diamond$$

Mit dem Kreuzprodukt kann man z.B. das Drehmoment  $\vec{M}$  einer Kraft  $\vec{F}$  bestimmen, das an einem Punkt angreift, der von der Drehachse über den Vektor  $\vec{r}$  zu erreichen ist:

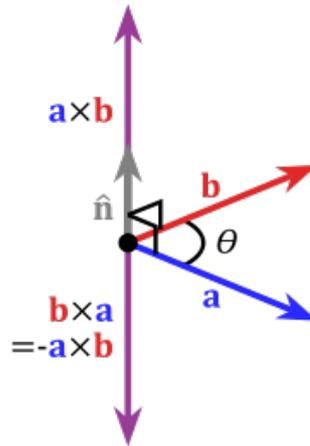
$$\vec{M} = \vec{r} \times \vec{F}.$$



**Satz 2.54 (Eigenschaften des Kreuzproduktes)**

Für Vektoren  $\vec{a}, \vec{b}, \vec{c} \in \mathbb{R}^3$  und  $\lambda \in \mathbb{R}$  gelten:

- (a)  $\vec{a} \times \vec{b}$  steht senkrecht auf  $\vec{a}$  und  $\vec{b}$
- (b)  $\vec{a}, \vec{b}$  und  $\vec{a} \times \vec{b}$  bilden in dieser Reihenfolge ein Rechtssystem
- (c)  $\|\vec{a} \times \vec{b}\| = \|\vec{a}\|\|\vec{b}\| \sin \varphi$ ,  
wobei  $\varphi$  der Winkel zwischen  $\vec{a}$  und  $\vec{b}$  ist (Definition 2.18)
- (d)  $\vec{a}$  steht senkrecht auf  $\vec{b} \iff \|\vec{a} \times \vec{b}\| = \|\vec{a}\|\|\vec{b}\|$   
(**Beachte:** Bequemer lässt sich das mit dem Skalarprodukt  $\vec{a} \cdot \vec{b} = 0$  testen.)
- (e)  $\|\vec{a} \times \vec{b}\|$  ist der Flächeninhalt des durch  $\vec{a}$  und  $\vec{b}$  aufgespannten Parallelogramms
- (f)  $\vec{a} \times \vec{b} = \vec{0} \iff \vec{a}$  und  $\vec{b}$  sind linear abhängig
- (g)  $\vec{a} \times \vec{b} = -\vec{b} \times \vec{a}$
- (h)  $(\vec{a} + \vec{b}) \times \vec{c} = \vec{a} \times \vec{c} + \vec{b} \times \vec{c}$
- (i)  $(\lambda \vec{a}) \times \vec{b} = \lambda(\vec{a} \times \vec{b}) = \vec{a} \times (\lambda \vec{b})$  ◇

**Beispiel 2.55 (zum Kreuzprodukt)**

Analog zum ebenen Fall legen zwei Punkte  $P \neq Q$  im Raum mit Ortsvektoren  $\vec{p}$  und  $\vec{q}$  oder ein Punkt  $P$  und ein Richtungsvektor  $\vec{u} \in \mathbb{R}^3$ ,  $\vec{u} \neq \vec{0}$  eindeutig eine Gerade fest. Ein Punkt  $X$  mit Ortsvektor  $\vec{x}$  liegt genau dann auf der Geraden, wenn  $\vec{x}$  die Parameterdarstellung

$$\vec{x} = \vec{p} + \lambda(\vec{q} - \vec{p})$$

**Zwei-Punkte-Form** bzw.

$$\vec{x} = \vec{p} + \lambda \vec{u}$$

**Punkt-Richtungs-Form**

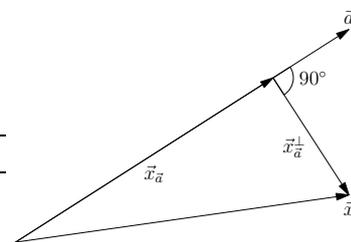
mit einem Parameter  $\lambda \in \mathbb{R}$  hat.

Wir erinnern an die schon in der Ebene benutzte **Projektion** eines Vektors  $\vec{x}$  auf  $\vec{a} \neq \vec{0}$ :

$$\vec{x}_a := \frac{\vec{x} \cdot \vec{a}}{\|\vec{a}\|^2} \vec{a},$$

die auch für Vektoren im  $\mathbb{R}^3$  gilt. Es gibt im  $\mathbb{R}^3$  eine explizite Möglichkeit, die zu  $\vec{a}$  senkrechte Komponente zu bestimmen:

$$\vec{x}_a^\perp = \vec{x} - \vec{x}_a = \frac{1}{\|\vec{a}\|^2} \vec{a} \times (\vec{x} \times \vec{a})$$



### Beispiel 2.56 (Abstand Punkt–Gerade)

**Frage:** Wie groß ist der Abstand eines Punktes  $A$  von der Geraden  $\vec{x} = \vec{p} + \lambda \vec{u}$ ?

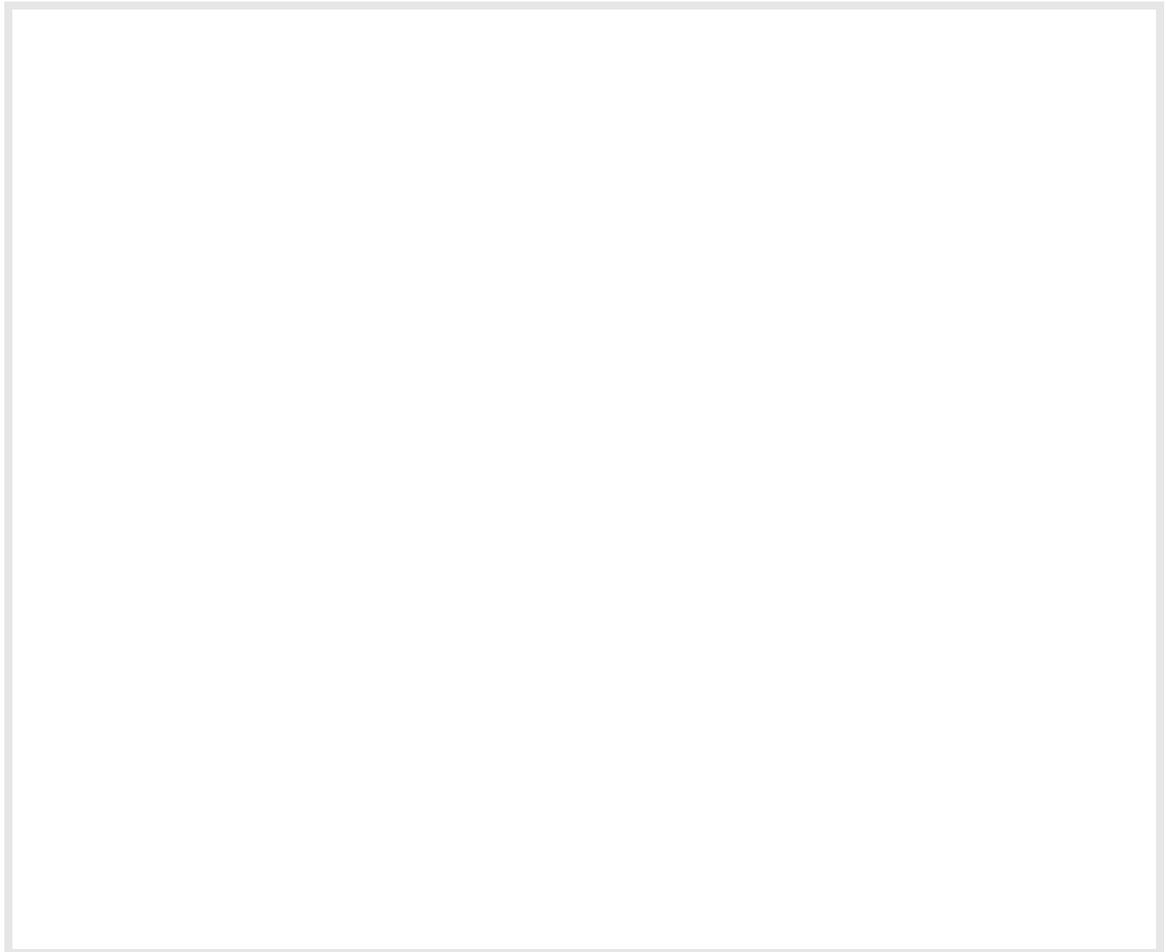
**Beispiel 2.57 (Abstand Gerade–Gerade)**

**Frage:** Wie groß ist der Abstand zweier Geraden

$$g_1 : \vec{x} = \vec{p}_1 + \lambda_1 \vec{u}$$

$$g_2 : \vec{x} = \vec{p}_2 + \lambda_2 \vec{v}$$

im Raum, wobei  $\vec{p}_1, \vec{p}_2, \vec{u}, \vec{v} \in \mathbb{R}^3$  gegeben sind,  $\vec{u}, \vec{v} \neq \vec{0}$ ?



Drei Punkte  $P$ ,  $Q$  und  $R$  im Raum, die nicht auf einer Geraden liegen, mit Ortsvektoren  $\vec{p}$ ,  $\vec{r}$  und  $\vec{q}$  legen eindeutig eine Ebene fest. Dasselbe gilt für einen Punkt  $P$  und zwei linear unabhängige Richtungsvektoren  $\vec{u}, \vec{v} \in \mathbb{R}^3$ .

Ein Punkt  $X$  mit Ortsvektor  $\vec{x}$  liegt genau dann in der Ebene, wenn  $\vec{x}$  die Parameterdarstellung

$$\vec{x} = \vec{p} + \lambda_1 (\vec{q} - \vec{p}) + \lambda_2 (\vec{r} - \vec{p})$$

**Drei-Punkte-Form** bzw.

$$\vec{x} = \vec{p} + \lambda_1 \vec{u} + \lambda_2 \vec{v}$$

**Punkt-Richtungs-Form**

mit Parametern  $\lambda_1, \lambda_2 \in \mathbb{R}$  hat. Der Vektor  $\vec{n} := \vec{u} \times \vec{v}$  heißt ein **Normalenvektor** der Ebene, da er senkrecht (normal) auf  $\vec{u}$  und  $\vec{v}$  steht. Durch Skalarmultiplikation der Gleichung mit  $\vec{n}$  ergibt sich die **parameterfreie Form** der Ebenengleichung:

$$\vec{x} \cdot \vec{n} = \vec{p} \cdot \vec{n} + \lambda_1 \underbrace{\vec{u} \cdot \vec{n}}_{=0} + \lambda_2 \underbrace{\vec{v} \cdot \vec{n}}_{=0}, \quad \text{also} \quad \vec{x} \cdot \vec{n} = \vec{p} \cdot \vec{n}.$$

Dies ist ein LGS für  $\vec{x} \in \mathbb{R}^3$  vom Rang 1, dessen Lösungsmenge gerade aus den Punkten der Ebene besteht (verschobener Unterraum des  $\mathbb{R}^3$  der Dimension 2).

Wir bestimmen nun die Schnittpunkte zwischen Ebene und Gerade

$$E : \vec{x} = \vec{p}_1 + \lambda_1 \vec{u} + \lambda_2 \vec{v}$$

$$g : \vec{x} = \vec{p}_3 + \lambda_3 \vec{w}$$

im Raum, wobei  $\vec{p}_1, \vec{p}_3, \vec{u}, \vec{v}, \vec{w} \in \mathbb{R}^3$  gegeben sind mit  $\vec{u}, \vec{v}$  linear unabhängig und  $\vec{w} \neq \vec{0}$ .

**Ansatz:**

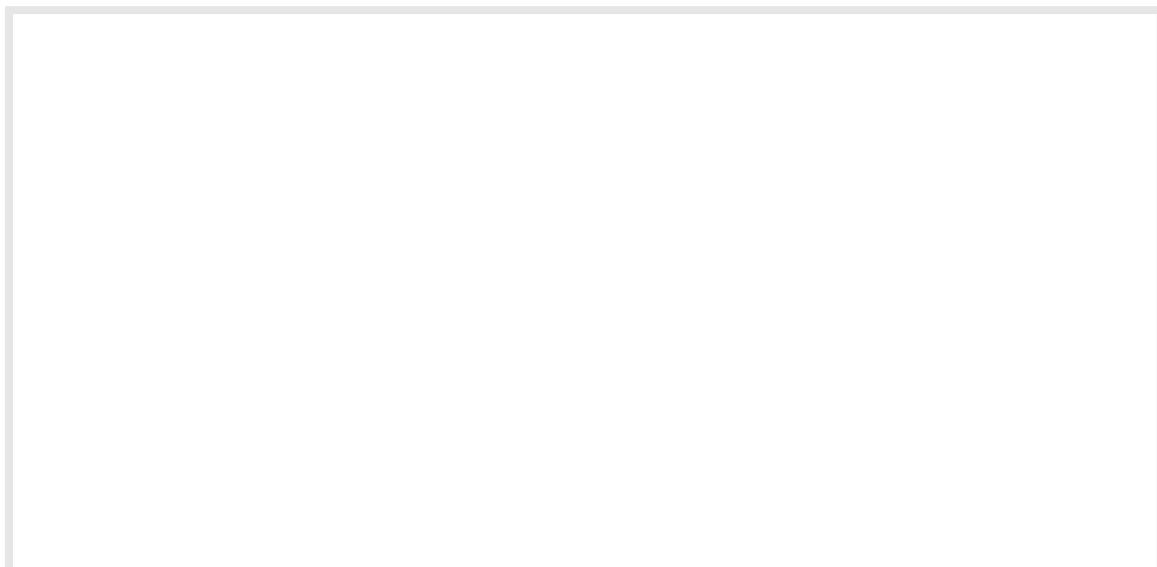
$$\begin{aligned} \vec{p}_1 + \lambda_1 \vec{u} + \lambda_2 \vec{v} &= \vec{p}_3 + \lambda_3 \vec{w} \\ \Leftrightarrow \begin{pmatrix} u_1 & v_1 & -w_1 \\ u_2 & v_2 & -w_2 \\ u_3 & v_3 & -w_3 \end{pmatrix} \begin{pmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \end{pmatrix} &= \vec{p}_3 - \vec{p}_1 \end{aligned}$$

Dies ist ein LGS mit drei Gleichungen für die drei Unbekannten  $\lambda_1, \lambda_2, \lambda_3$ . Folgende Fälle können auftreten:

- Die Matrix hat Rang 3, d.h., die Richtungsvektoren  $\vec{u}, \vec{v}, \vec{w}$  sind linear unabhängig. Es gibt eine eindeutige Lösung  $(\lambda_1, \lambda_2, \lambda_3)$ , d.h., einen eindeutigen Schnittpunkt.
- Die Matrix hat Rang 2, d.h., die Richtungsvektoren  $\vec{u}, \vec{v}, \vec{w}$  sind linear abhängig. Dann gibt es entweder unendlich viele Lösungen  $(\lambda_1, \lambda_2, \lambda_3)$  mit einer freien Variablen (die Gerade liegt in der Ebene) oder keine Lösung (die Gerade liegt parallel zur Ebene).
- Es kann nicht vorkommen, dass die Matrix nur Rang 1 oder 0 hat, da ja bereits die beiden Spalten  $\vec{u}$  und  $\vec{v}$  linear unabhängig sind.

### Beispiel 2.58 (Abstand Punkt–Ebene)

**Frage:** Wie groß ist der Abstand eines Punktes  $A$  von der Ebene  $\vec{x} = \vec{p} + \lambda_1 \vec{u} + \lambda_2 \vec{v}$ ?



Ende 11. V

06.01.2011



§ 3 Folgen und Reihen

§ 3.1 Folgen

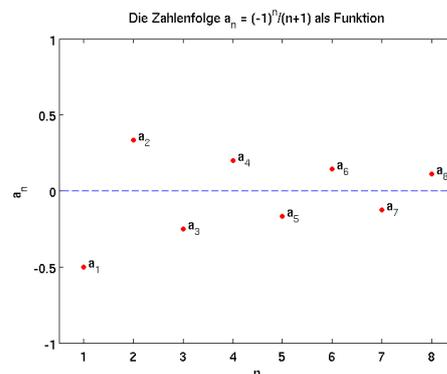
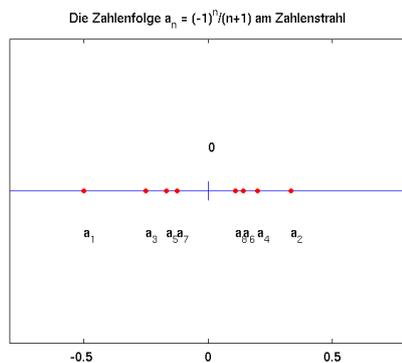
**Definition 3.1 (Folge)** (a) Eine Vorschrift, die jedem **Index**  $n \in \mathbb{N}_0$  eine Zahl  $a_n \in \mathbb{R}$  zuordnet, heißt eine **reelle Zahlenfolge** oder einfach **Folge**. Man schreibt  $(a_n)_{n \in \mathbb{N}_0}$  oder  $(a_n)_{n=0}^\infty$  oder  $a_0, a_1, \dots$ . Die Zahl  $a_n$  heißt das  **$n$ -te Glied** der Folge.

(b) Eine Folge braucht nicht bei  $a_0$  zu beginnen. Man schreibt dann z.B.  $(a_n)_{n \geq 3}$  oder  $(a_n)_{n=3}^\infty$ . ◇

Folgen werden definiert durch

- Angabe der Bildungsvorschrift, z.B.

$$a_n = \frac{(-1)^n}{n+1} \quad \text{für } n \geq 1, \quad \text{also } a_n = \left( \frac{-1}{2}, \frac{1}{3}, \frac{-1}{4}, \frac{1}{5}, \dots \right)$$



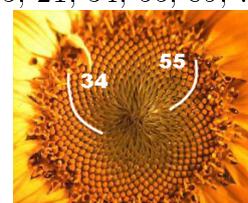
- rekursive Definition, z.B.

$$a_{n+1} = a_n + a_{n-1} \quad \text{für } n \geq 2$$

mit Startwerten  $a_1 = a_2 = 1$  (**Fibonacci-Folge**).

Die Fibonacci-Folge beginnt mit 1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, ...

Man trifft sie z.B. bei der spiralförmigen Anordnung von Samen in Blütenständen an, etwa bei **Sonnenblumen**.



**Beispiel 3.2 (Wichtige Beispiele von Folgen)**

Eine Folge mit Gliedern ...

- (a)  $a_n := c$  mit  $c \in \mathbb{R}$  heißt **konstante Folge**.
- (b)  $a_n := a_0 + nd$  mit  $a_0, d \in \mathbb{R}$  heißt **arithmetische Folge**. Die Differenz aufeinanderfolgender Glieder ist konstant (gleich  $d$ ).
- (c)  $a_n := a_0 q^n$  mit  $a_0, q \in \mathbb{R}$  heißt **geometrische Folge**. Der Quotient aufeinanderfolgender Glieder ist konstant (gleich  $q$ ). ◇

**Definition 3.3** (Beschränktheit und Konvergenz)

- (a) Eine Folge  $(a_n)$  heißt **beschränkt**, wenn es Zahlen  $c_1, c_2 \in \mathbb{R}$  gibt, sodass  $c_1 \leq a_n \leq c_2$  für alle  $n$  gilt. (Alle Folgenglieder liegen im Intervall  $[c_1, c_2]$ .)
- (b) Eine Folge  $(a_n)$  heißt **konvergent** gegen den **Grenzwert**  $a \in \mathbb{R}$ , wenn gilt: Für jede beliebig kleine Schranke  $\varepsilon > 0$  gibt es einen Index  $n_0 = n_0(\varepsilon) \in \mathbb{N}$ , sodass

$$|a_n - a| < \varepsilon \quad \text{für alle } n \geq n_0$$

gilt. (In jedem noch so kleinen Intervall mit Mittelpunkt  $a$  liegen alle hinreichend späten Folgenglieder.)

- (c) Man schreibt dann:

$$\lim_{n \rightarrow \infty} a_n = a \quad \text{oder} \quad a_n \xrightarrow{n \rightarrow \infty} a.$$

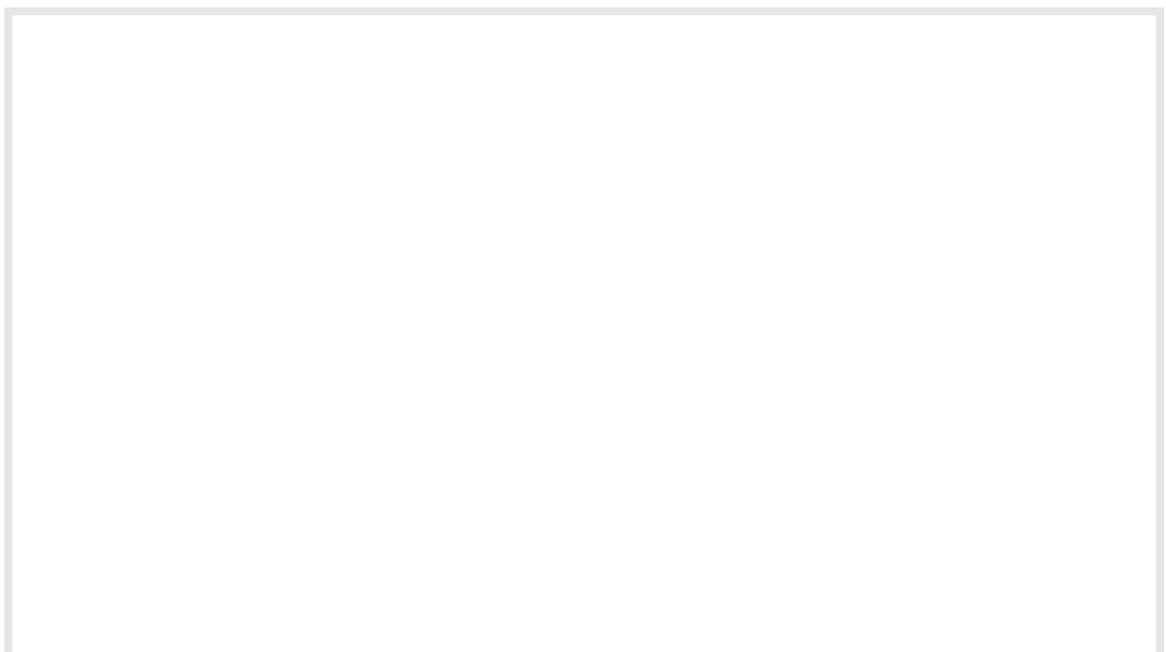
- (d) Eine nicht-konvergente Folge heißt **divergent**. ◇

**Bemerkung 3.4**

Es kommt bei der Frage nach der Beschränktheit und Konvergenz einer Folge nicht darauf an, ab welchem Index die Folge definiert ist. Deshalb brauchen wir den Indexbereich nicht anzugeben. ◇

**Satz 3.5** (Eindeutigkeit des Grenzwertes und Beschränktheit)

- (a) Der Grenzwert einer konvergenten Folge ist eindeutig.
- (b) Eine konvergente Folge ist immer beschränkt. ◇

**Beispiel 3.6** (Grenzwerte wichtiger Folgen)

**Satz 3.7 (Grenzwertsatz)**

Seien  $(a_n)$ ,  $(b_n)$  und  $(c_n)$  Folgen.

$$(a) \quad a_n \rightarrow a \text{ und } b_n \rightarrow b \quad \Rightarrow \quad a_n \pm b_n \rightarrow a \pm b.$$

$$(b) \quad a_n \rightarrow a \text{ und } b_n \rightarrow b \quad \Rightarrow \quad a_n b_n \rightarrow a b.$$

$$(c) \quad a_n \rightarrow a \text{ und } b_n \rightarrow b \neq 0 \quad \Rightarrow \quad \frac{a_n}{b_n} \rightarrow \frac{a}{b}.$$

**Beachte:** Die Folge  $\left(\frac{a_n}{b_n}\right)$  ist ab einem gewissen Index durchgängig definiert, weil  $b_n \neq 0$  ist.

$$(d) \quad a_n \leq b_n \text{ und } a_n \rightarrow a \text{ und } b_n \rightarrow b \quad \Rightarrow \quad a \leq b.$$

$$(e) \quad a_n \leq b_n \leq c_n \text{ und } a_n \rightarrow a \text{ sowie } c_n \rightarrow a \quad \Rightarrow \quad b_n \rightarrow a$$

**(Sandwich-Theorem).** ◇

**Definition 3.8 (Divergenz gegen  $\infty$ )** (a) Eine Folge  $(a_n)$  heißt **bestimmt divergent** gegen  $\infty$  [bzw.  $-\infty$ ], wenn gilt: Für jede beliebig große Zahl  $R > 0$  gibt es einen Index  $n_0 = n_0(R) \in \mathbb{N}$ , sodass

$$a_n > R \quad \text{für alle } n \geq n_0 \quad \text{bzw.} \quad a_n < -R \quad \text{für alle } n \geq n_0$$

gilt. Man schreibt dann:

$$\lim_{n \rightarrow \infty} a_n = \infty \quad \text{oder} \quad a_n \xrightarrow{n \rightarrow \infty} \infty \quad [\text{bzw. } -\infty].$$

(b) Eine divergente Folge, die nicht bestimmt divergiert, heißt **unbestimmt divergent**. ◇

Jede Folge  $(a_n)$  ist also entweder

- konvergent,
- bestimmt divergent gegen  $+\infty$ ,
- bestimmt divergent gegen  $-\infty$
- oder unbestimmt divergent.

**Beispiel 3.9 (Divergente Folgen)**

### § 3.2 Reihen

#### Definition 3.10 (Reihe)

Sei  $(a_n)_{n \in \mathbb{N}_0}$  eine Folge.

- (a) Wir bilden daraus eine neue Folge  $(s_n)_{n \in \mathbb{N}_0}$ :

$$s_n := \sum_{k=0}^n a_k = a_0 + a_1 + \dots + a_n.$$

Diese Folge  $(s_n)$  heißt die **Folge der Partialsummen** von  $(a_n)$ .

- (b) Die Folge  $(s_n)_{n \in \mathbb{N}_0}$  heißt auch die **(unendliche) Reihe** mit der **Gliederfolge**  $(a_n)$ .  $\diamond$

#### Bemerkung 3.11

Eine Reihe ist also nichts anderes als eine Folge. Sie kann konvergieren oder (bestimmt bzw. unbestimmt) divergieren. Falls  $(s_n)$  gegen  $s$  konvergiert, so schreibt man

$$\lim_{n \rightarrow \infty} s_n = s = \sum_{n=0}^{\infty} a_n$$

und nennt  $s$  die **Summe** der Reihe.  $\diamond$

#### Beispiel 3.12 (Reihen)

- (a) Die **geometrische Reihe** ist die Reihe mit der Gliederfolge  $a_n = a_0 q^n$ . Die Folge der Partialsummen kann explizit angegeben werden:

$$\begin{cases} s_n = \sum_{k=0}^n a_0 q^k = a_0 \frac{1 - q^{n+1}}{1 - q}, & \text{falls } q \neq 1 \\ s_n = a_0 (n + 1), & \text{falls } q = 1. \end{cases}$$

Wir nehmen  $a_0 \neq 0$  an (sonst ist  $s_n \equiv 0$ ). Es gilt also:

$$s_n \rightarrow \frac{a_0}{1 - q} \quad \text{oder kurz} \quad \sum_{n=0}^{\infty} a_0 q^n = \frac{a_0}{1 - q} \quad \text{für } |q| < 1.$$

Für  $|q| \geq 1$  divergiert die Reihe. Sie divergiert bestimmt gegen  $\infty$  im Fall  $q \geq 1$  und  $a_0 > 0$ .

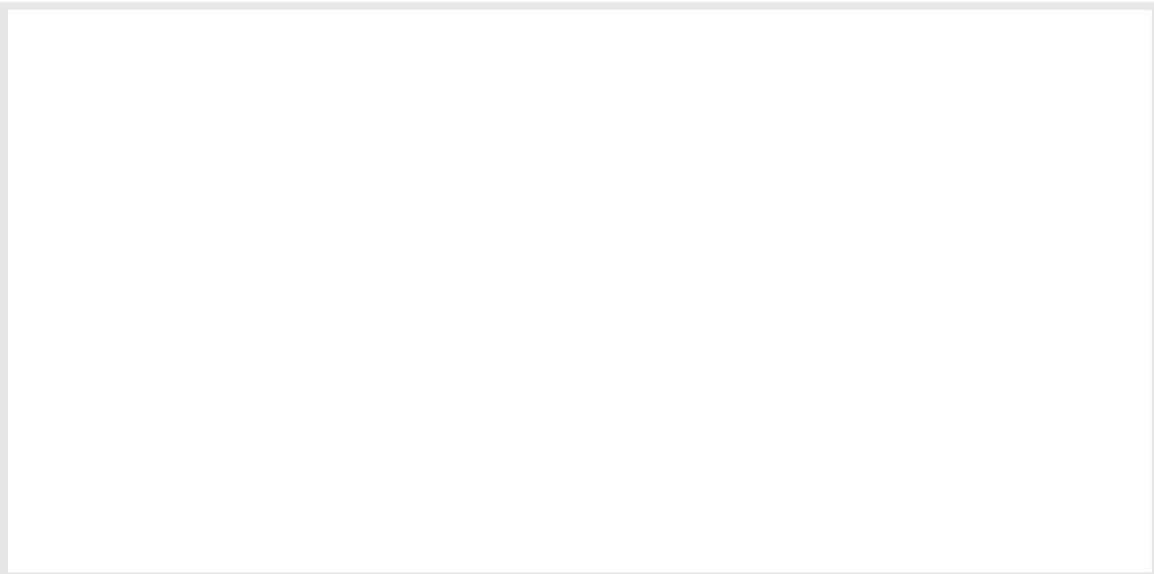
- (b) Die **harmonische Reihe** ist die Reihe mit der Gliederfolge  $a_n = \frac{1}{n}$ . Diese Reihe divergiert bestimmt gegen  $\infty$ .  $\diamond$

**Satz 3.13 (Konvergenzkriterien für Reihen)**

Es seien  $(a_n)_{n \in \mathbb{N}_0}$  und  $(b_n)_{n \in \mathbb{N}_0}$  die Gliederfolgen der Reihen  $(s_n)_{n \in \mathbb{N}_0}$  und  $(t_n)_{n \in \mathbb{N}_0}$ .

- (a) Falls  $(s_n)$  konvergiert, so gilt notwendig  $\lim_{n \rightarrow \infty} a_n = 0$ .
- (b) Falls  $0 \leq a_n \leq b_n$  erfüllt ist für alle  $n \geq n_0 \in \mathbb{N}_0$ , dann gilt:  
 $(t_n)$  konvergiert  $\Rightarrow$   $(s_n)$  konvergiert.
- (c) Falls  $0 \leq a_n \leq b_n$  erfüllt ist für alle  $n \geq n_0 \in \mathbb{N}_0$ , dann gilt:  
 $(s_n)$  divergiert bestimmt gegen  $\infty \Rightarrow$   $(t_n)$  divergiert bestimmt gegen  $\infty$ .  $\diamond$

**Beispiel 3.14 (Bestimmt divergente Reihe)**



Ende 12. V

13.01.2011

**§ 3.3 Folgen und Reihen in der Finanzmathematik**

**§ 3.3.1 Zinsrechnung**

Unter dem Begriff **Zinsen** versteht man die Vergütung für die Überlassung eines Geldbetrages in einer bestimmten Zeit (**Zinsperiode**). Die Höhe der Zinsen hängt von den folgenden drei Einflussgrößen ab:

- vom **Startkapital** (Geldbetrag),
- von der **Laufzeit** (Dauer der Überlassung) und
- vom **Zinssatz**.

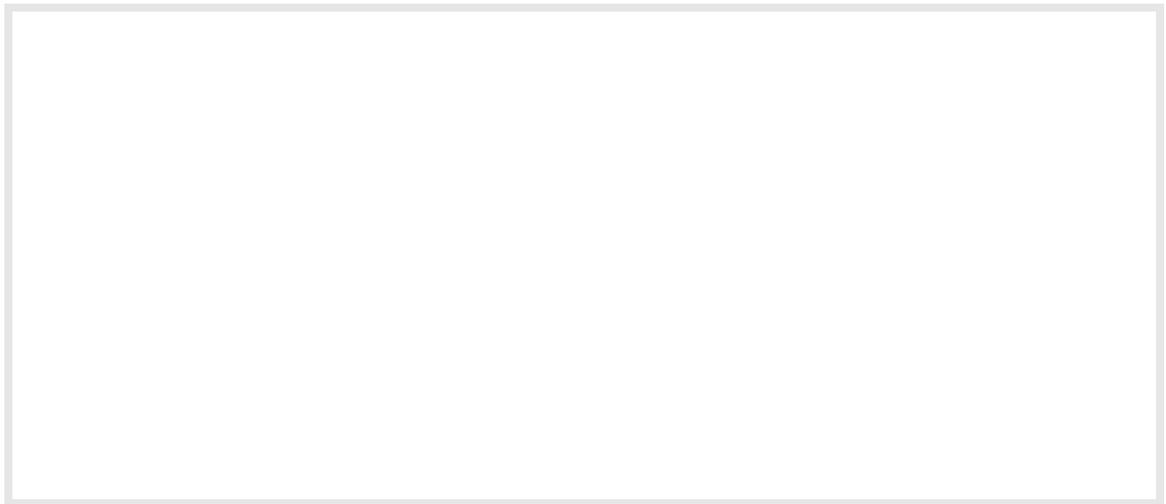
**Beachte:**  $1\% = 0,01$ .

**Beispiel 3.15 (Einfache Verzinsung)**

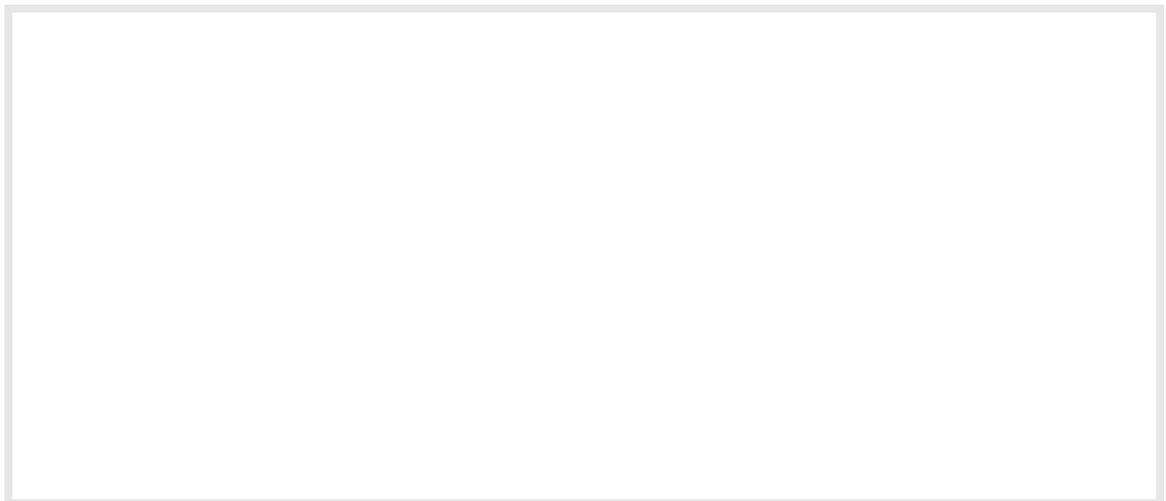
Am Ende der Zinsperiode werden die Zinsen ausgezahlt bzw. einem anderen Konto gutgeschrieben. Mit den Bezeichnungen:

- $K_0$  – Startkapital,
- $t$  – Teil der Zinsperiode ( $t = 1$  entspricht einer vollen Zinsperiode) und
- $p$  – Zinssatz für eine Zinsperiode

erhält man die Beziehungen



**Rechenbeispiel:** In Deutschland wird ein Jahr zu 360 und jeder volle Monat zu 30 Zinstagen angenommen. Ein am 11.03. eines Jahres eingezahlter Betrag von € 3 000 wird am 16.08. desselben Jahres wieder abgehoben. Wieviel Zinsen erbringt er bei einer jährlichen Verzinsung von 5%?



### **Beispiel 3.16** (Zinseszinsrechnung)

Am Ende einer Zinsperiode werden die Zinsen dem Kapital zugeschlagen und im Weiteren mit verzinst. Mit den Bezeichnungen

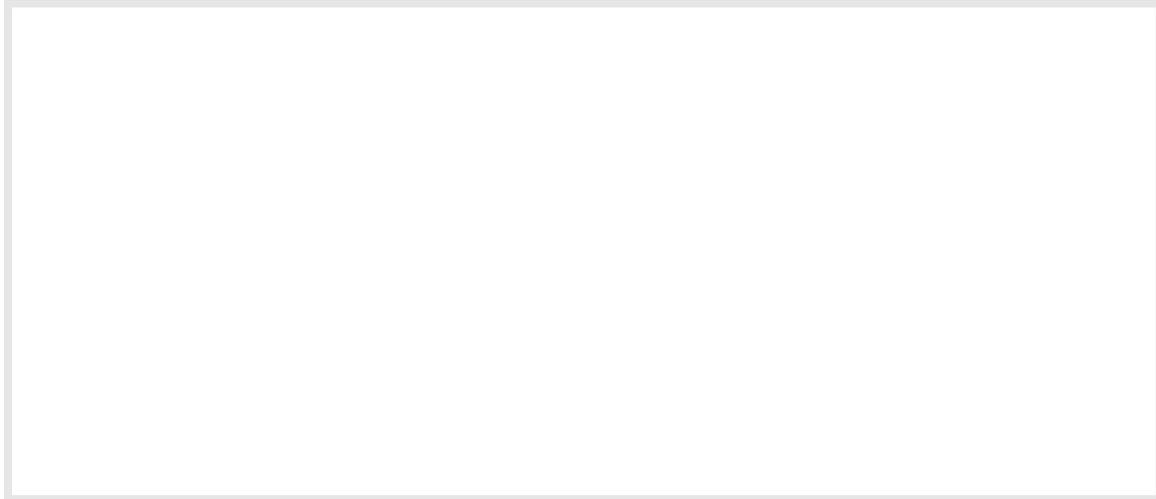
- $K_0$  – Startkapital,
- $n$  – Anzahl der Zinsperioden (z.B. Jahre oder Quartale) und

- $p$  – Zinssatz für eine Zinsperiode

und den Größen

- $K_n$  – Kapital am Ende der  $n$ -ten Zinsperiode (Zeitwert)
- $q = (1 + p)$  – **Aufzinsungsfaktor**

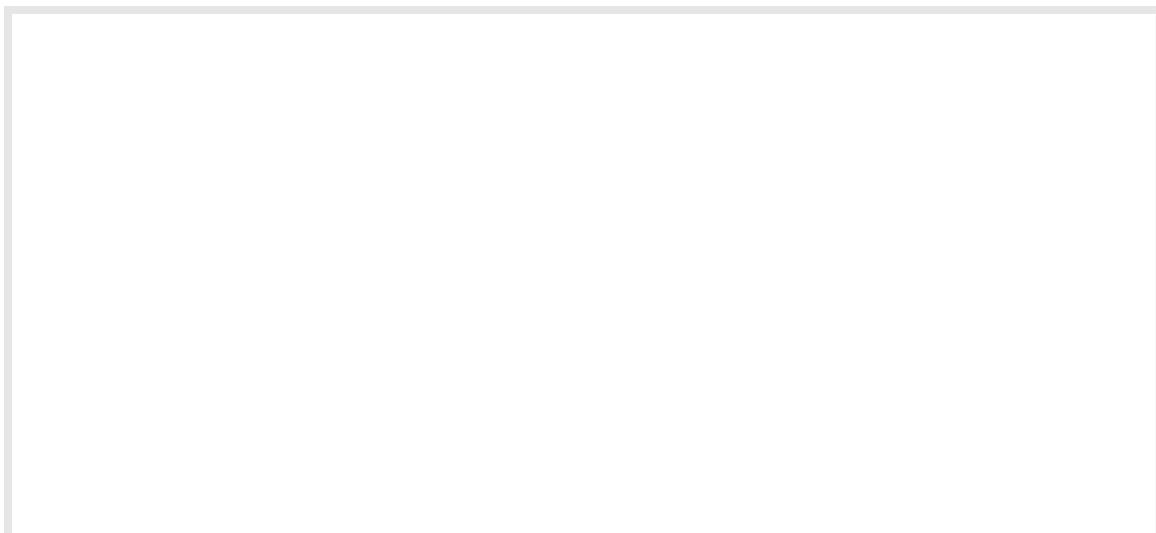
erhält man die Beziehungen



Man nennt  $q^n$  den **Aufzinsungsfaktor für  $n$  Zinsperioden** und  $q^{-n}$  den **Abzinsungsfaktor für  $n$  Zinsperioden**. Die Berechnung des Barwertes  $K_0$  aus dem (gewünschten) Endkapital  $K_n$  bezeichnet man auch als **Abzinsen** oder **Diskontieren**.

**Rechenbeispiel:**

- (a) Ein Bürger kauft Finanzierungsschätze des Bundes (Laufzeit: 2 Zinsperioden (Jahre)) im Nominalwert von €5 000 und muss dafür €4 441,60 bezahlen. Welcher Verzinsung pro Jahr entspricht dies?



- (b) Am 01.01.2002 verleiht A an B €10 000 zu 10% Zinsen pro Jahr. Welchen Betrag muss B am Rückzahlungstermin, dem 31.12.2008, zurückzahlen bei

**Beispiel 3.17 (Kontinuierliche Verzinsung)**

Eine Bank A bietet einen jährlichen Zinssatz von  $p$ . Die Zinsen werden jährlich gutgeschrieben. Eine weitere Bank B bietet den Zinssatz  $p/2$  für eine halbjährige Zinsperiode, eine dritte Bank C bietet  $p/4$  für ein Quartal. Welchem jährlichen Zinssatz entsprechen die Angebote?

### § 3.3.2 Rentenrechnung

Eine in gleichen Zeitabständen erfolgende Zahlung in bestimmter Höhe nennt man **Rente**. Diese Zahlungen können einem Guthaben entnommen werden, sodass dieses nach einer endlichen Anzahl von Zahlungen erlöschen kann. Die Zahlungen können aber auch dazu dienen, ein Guthaben anzusammeln.

Dabei bezeichnet  $r$  die Höhe der **Ratenzahlung** und  $n$  die Anzahl der Ratenzahlungen bzw. **Perioden**. Eine Rente heißt

- **vorschüssig**, wenn die Zahlungen zu Beginn, und
- **nachschüssig**, wenn die Zahlungen am Ende jeder Periode erfolgen.

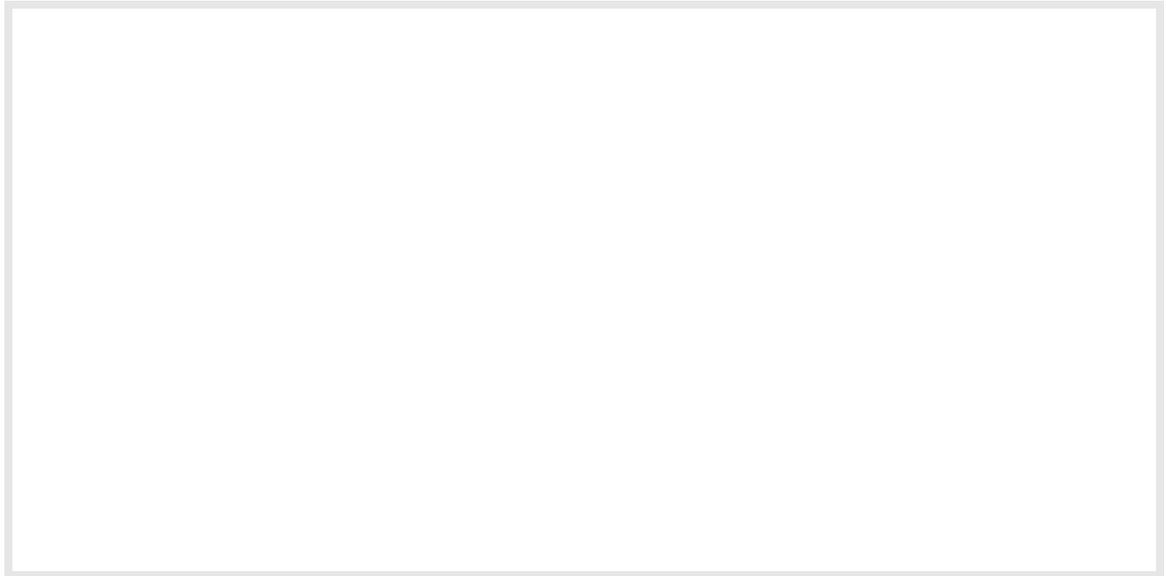
Zur Vereinfachung nehmen wir immer an, dass die *Ratenperiode gleich der Zinsperiode* (z.B. Jahr) ist.

Ferner unterscheidet man **Zeitrenten** (von begrenzter Dauer) und **ewige Renten** (von unbegrenzter Dauer). Die Rentenhöhe kann entweder **gleichbleibend** oder **dynamisch** (meist wachsend) sein. Wir betrachten hier nur *Renten konstanter Höhe*. Uns interessiert der **Barwert**  $B$  und der **Endwert**  $E$  aller Rentenzahlungen.

#### Beispiel 3.18 (Vorschüssige Zeitrenten)

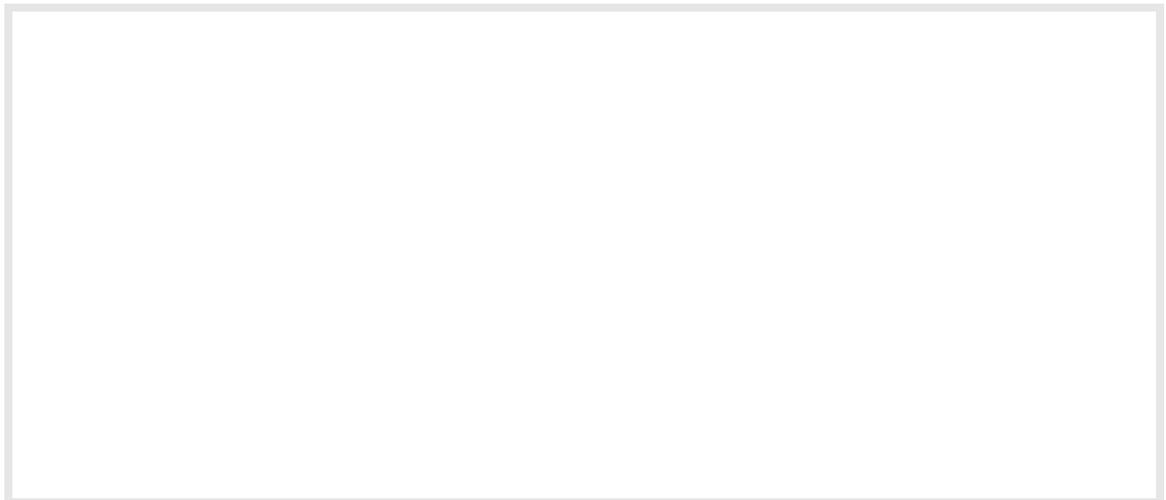
Der **Renten-Endwert**  $E_n^V$  ist derjenige Betrag, der zum Zeitpunkt  $n$  (nach  $n$  Raten-/Zinsperioden) ein Äquivalent für die  $n$  zu zahlenden Raten darstellt. Zur Berechnung von  $E_n^V$  bestimmen wir die Endwerte der einzelnen Zahlungen mit  $K_0 = r$ . Entsprechend den unterschiedlichen Zahlungszeitpunkten werden die Raten der Höhe  $r$  über eine unterschiedliche Anzahl von Perioden aufgezinst. Anschließend werden alle Endwerte aufsummiert:

Der **Renten-Barwert**  $B_n^V$  ist derjenige Betrag, der zum Zeitpunkt 0 einmalig angelegt werden müsste, um zum Zeitpunkt  $n$  den **Renten-Endwert**  $E_n^V$  zu erreichen. Man erhält ihn durch Abzinsen von  $E_n^V$  über  $n$  Zinsperioden:



**Beispiel 3.19** (Nachschüssige Zeitrenten)

Der **Renten-Endwert**  $E_n^N$  wird wieder durch Addition der  $n$  einzelnen Zahlungen errechnet. Da die Zahlungen hier am *Ende* der Periode erfolgen, erhält man



Der **Renten-Barwert**  $B_n^N$  ergibt sich wieder durch Abzinsen dieses Ausdrucks über  $n$  Zinsperioden:



Die nachstehende Tabelle zeigt die Zusammenhänge zwischen **Bar-** und **Endwerten** von **vorschüssigen** (Beispiel 3.18) und **nachschüssigen Renten**:

	Vorschüssige Rente	Nachschüssige Rente
<b>Renten-Endwert</b>	$E_n^V = q^n \cdot B_n^V = r \cdot \text{REF}^V$	$E_n^N = q^n \cdot B_n^N = r \cdot \text{REF}^N$
<b>Renten-Barwert</b>	$B_n^V = q^{-n} \cdot E_n^V = r \cdot \text{RBF}^V$	$B_n^N = q^{-n} \cdot E_n^N = r \cdot \text{RBF}^N$

**Rechenbeispiel:** Ein Großvater zahlt für seine Enkelin jeweils zu Jahresende € 1 200 bei einer Bank ein. Auf welchen Betrag sind die Einzahlungen nach 15 Jahren bei 6,5% jährlicher Verzinsung angewachsen, und welchem Barwert entspricht dieses Guthaben?

**Beispiel 3.20 (Ewige Renten)**

Wegen der sinnvollen Voraussetzung  $q = 1 + p > 1$  (positive Verzinsung) divergieren die Endwerte  $E_n^V$  und  $E_n^N$  für  $n \rightarrow \infty$  bestimmt gegen  $\infty$ , siehe Beispiel 3.12 (a). Man kann jedoch einen Renten-Barwert für ewige Renten berechnen:

$$B_\infty^V = \lim_{n \rightarrow \infty} B_n^V = \lim_{n \rightarrow \infty} \frac{r}{q^{n-1}} \frac{1 - q^n}{1 - q} = \lim_{n \rightarrow \infty} r q \frac{q^{-n} - 1}{1 - q} = \frac{r q}{q - 1} = r \frac{1 + p}{p}$$
$$B_\infty^N = \lim_{n \rightarrow \infty} B_n^N = \lim_{n \rightarrow \infty} \frac{r}{q^n} \frac{1 - q^n}{1 - q} = \lim_{n \rightarrow \infty} r \frac{q^{-n} - 1}{1 - q} = \frac{r}{q - 1} = \frac{r}{p},$$

denn  $\lim_{n \rightarrow \infty} q^{-n} = \lim_{n \rightarrow \infty} (1/q)^n = 0$ . Der **Barwert** ist z.B. bei Stiftungen von Interesse, bei denen nur die Zinsen ausgezahlt werden sollen und das eigentliche Kapital unangetastet bleiben soll.

## § 4 Funktionen einer Variablen

**Definition 4.1** (Funktion, Definitionsbereich, Wertemenge)

- (a) Eine Vorschrift  $f$ , die jedem Element  $x$  einer Menge  $X$  genau ein Element  $y$  einer Menge  $Y$  zuordnet, heißt **Funktion** mit **Definitionsmenge**  $D(f) = X$  und **Zielmenge**  $Y$ . Kurz:  $f : X \rightarrow Y$ .
- (b) Die **Wertemenge** oder **Bildmenge** einer Funktion  $f : X \rightarrow Y$  ist die Menge der tatsächlich angenommenen Funktionswerte  $W(f) = \{y \in Y : y = f(x) \text{ für ein } x \in X\} \subseteq Y$ .
- (c) Bei  $f : D(f) \subseteq \mathbb{R} \rightarrow \mathbb{R}$  spricht man von einer **reellen (reellwertigen) Funktion einer Variablen**. Solche betrachten wir in erster Linie in § 4. In § 4.8 treten dann auch Funktionen  $f : \mathbb{R} \rightarrow \mathbb{R}^n$  auf.  $\diamond$

**Beachte:** Eine Folge (Definition 3.1) ist eine Funktion  $a : \mathbb{N} \rightarrow \mathbb{R}$ . Man schreibt z.B.  $a_n = \frac{1}{n}$  statt  $a(n) = \frac{1}{n}$ .

Darstellungsmöglichkeiten einer Funktion:

- explizite Darstellung  $y = f(x)$
- Tabelle von Funktionswerten  $(x_i, f(x_i))$ ,  $i = 1, \dots, n$  (z.B. Messwerte)
- grafische Darstellung; die Menge  $\{(x, f(x)) : x \in D(f)\} \subseteq \mathbb{R}^2$  heißt der **Graph** von  $f$

**Definition 4.2** (Eigenschaften reeller Funktionen)

Sei  $f : D(f) \rightarrow \mathbb{R}$  und  $I \subseteq D(f)$  ein Intervall (Definition 1.17).  $f$  heißt ...

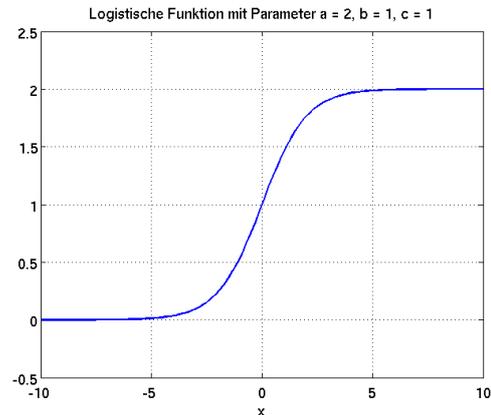
- (a) **konstant** auf  $I$ , wenn  $f(x) = a$  gilt für alle  $x \in I$ .
- (b) **beschränkt** auf  $I$ , wenn es Zahlen  $c_1, c_2 \in \mathbb{R}$  gibt, sodass  $c_1 \leq f(x) \leq c_2$  für alle  $x \in I$  gilt. (Alle Funktionswerte liegen im Intervall  $[c_1, c_2]$ .)
- (c) **monoton wachsend** bzw. **monoton fallend** auf  $I$ , wenn  $x_1, x_2 \in I$  und  $x_1 \leq x_2 \Rightarrow f(x_1) \leq f(x_2)$  bzw.  $f(x_1) \geq f(x_2)$ .
- (d) **streng monoton wachsend** bzw. **streng monoton fallend** auf  $I$ , wenn  $x_1, x_2 \in I$  und  $x_1 < x_2 \Rightarrow f(x_1) < f(x_2)$  bzw.  $f(x_1) > f(x_2)$ .
- (e) **periodisch** auf  $D(f)$  mit Periode  $p > 0$ , wenn  $x \in D(f) \Rightarrow x + p \in D(f)$  und gilt:  $f(x + p) = f(x)$  für alle  $x \in D(f)$ .  $\diamond$

**Beispiel 4.3 (Logistische Funktion)**

Die **logistische Funktion** ist definiert durch

$$f(x) = \frac{a}{1 + b e^{-cx}}$$

mit Parametern  $a, b, c > 0$ . Sie ist streng monoton wachsend und beschränkt (durch  $c_1 = 0$  und  $c_2 = a$ ) auf ihrem Definitionsbereich  $D(f) = \mathbb{R}$ .



◇

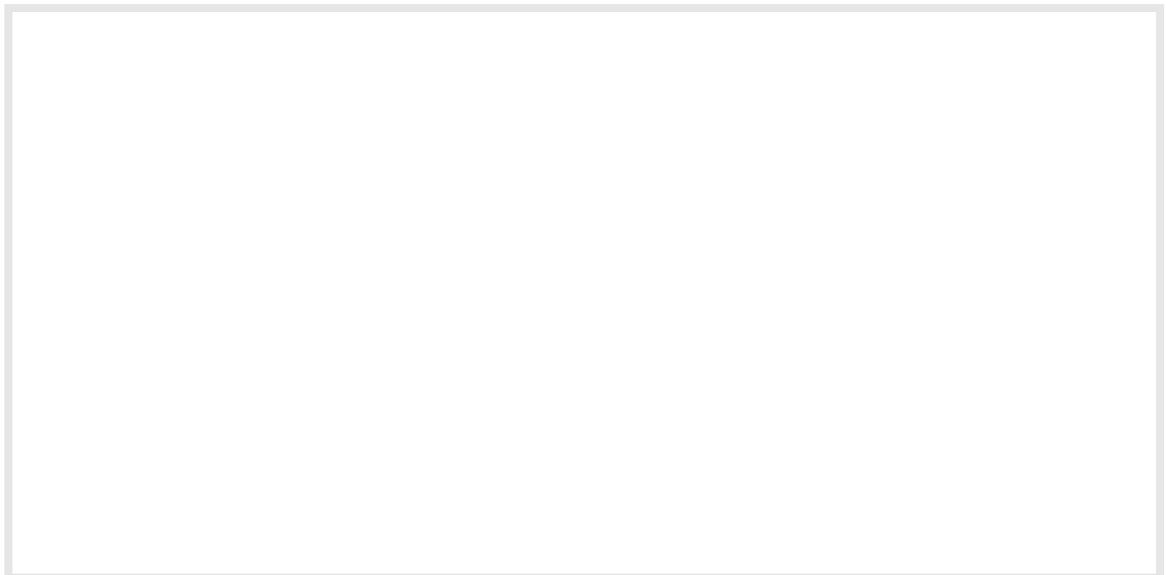
**Definition 4.4 (Verkettete Funktionen)**

Es seien  $g : \mathbb{R} \supseteq D(g) \rightarrow \mathbb{R}$  und  $f : \mathbb{R} \supseteq D(f) \rightarrow \mathbb{R}$  zwei Funktionen mit  $W(g) \subseteq D(f)$ . Dann heißt die Funktion  $f \circ g : D(g) \rightarrow \mathbb{R}$  definiert durch

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

$f$  **verkettet mit**  $g$  oder  $f$  **von**  $g$  oder die **Komposition** von  $f$  mit  $g$ .

◇

**Beispiel 4.5 (Verkettete Funktionen)****Definition 4.6 (Urbild, Umkehrfunktion)**

Sei  $f : \mathbb{R} \supseteq D(f) \rightarrow \mathbb{R}$ .

(a) Für  $y \in W(f)$  heißt die Menge

$$\{x \in D(f) : f(x) = y\}$$

das **Urbild** von  $y$ .

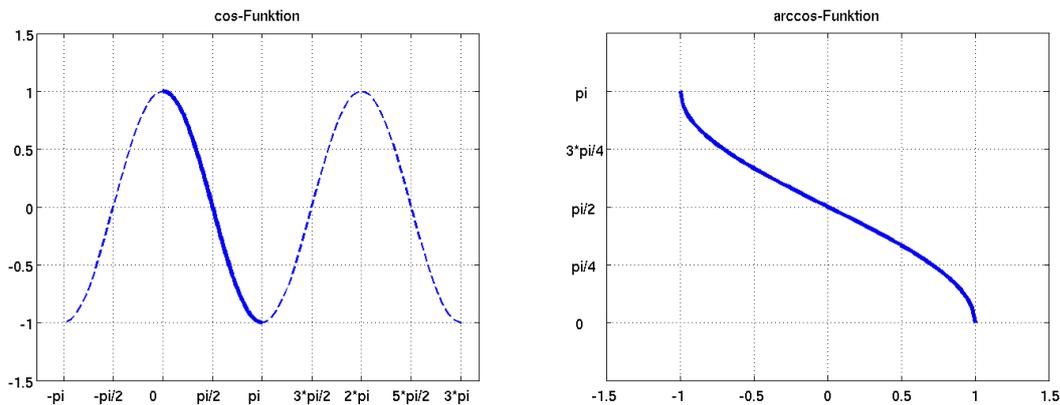
(b) Gilt für alle  $y \in W(f)$ , dass das Urbild von  $y$  aus nur einem Element  $x$  besteht, so ist dadurch die **Umkehrfunktion**

$$f^{-1} : W(f) \rightarrow D(f)$$

definiert. In dem Fall heißt  $f$  **umkehrbar** oder **invertierbar** oder **eindeutig**. ◇

**Beispiel 4.7 (Invertierbarmachen einer Funktion durch Einschränkung ihres Definitionsbereiches)**

Die Funktion  $f : \mathbb{R} \rightarrow \mathbb{R}$ ,  $f(x) = \cos(x)$  besitzt den Wertebereich  $[-1, 1]$ . Sie ist nicht umkehrbar. Bei Einschränkung des Definitionsbereiches auf  $D(f) = [0, \pi]$  wird die Funktion jedoch streng monoton fallend und damit invertierbar.



Die Umkehrfunktion  $f^{-1} : [-1, 1] \rightarrow [0, \pi]$  heißt **Arkuscosinus** (arccos). Sie ist ebenfalls streng monoton fallend. ◇

**Definition 4.8 (Polynom)**

Eine Funktion der Form

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \dots + a_1 x + a_0$$

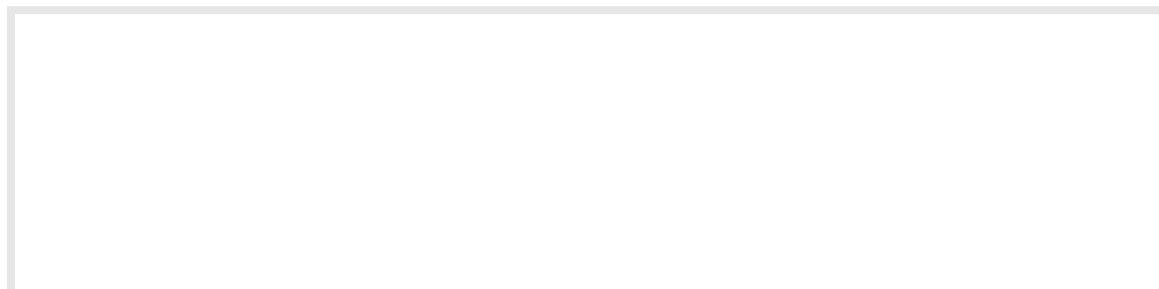
mit Zahlen (**Koeffizienten**)  $a_0, a_1, \dots, a_n \in \mathbb{R}$  und  $n \in \mathbb{N}_0$  heißt ein **Polynom**. Es heißt vom **Grad**  $n \in \mathbb{N}_0$ , wenn  $a_n \neq 0$  ist. ◇

Polynome sind besonders einfach zu handhabende Funktionen. Zum Beispiel werden sie in Form von **Bézierkurven**

- zur Modellierung von Karosseriefächen im Automobilbau und
- zur Beschreibung von Schriften (Fonts) in elektronischen Dokumenten

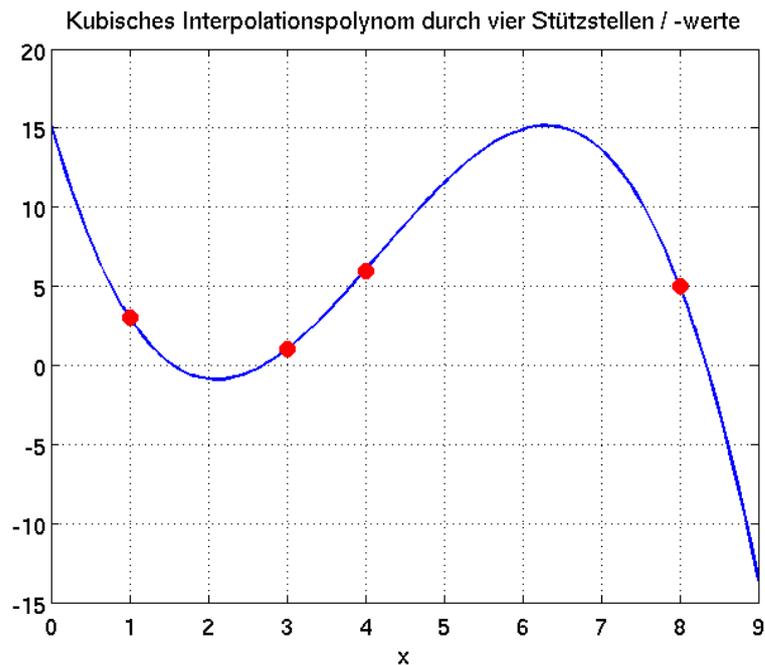
verwendet.

**Beispiel 4.9 (Polynome)**



### § 4.1 Polynominterpolation

**Aufgabe:** Zu einer Anzahl von Punkten  $(x_i, y_i) \in \mathbb{R}^2$ ,  $i = 1, \dots, n$  (z.B. Paaren von Messwerten) soll ein Polynom bestimmt werden, dessen Graph die gegebenen Punkte **interpoliert** („fittet“). Die Zahlen  $x_i \in \mathbb{R}$  heißen **Stützstellen** und die  $y_i \in \mathbb{R}$  **Stützwerte**.



Das Bild zeigt das eindeutige kubische Interpolationspolynom durch die  $n = 4$  Stützstellen und -werte

$x_i$	1	3	4	8
$y_i$	3	1	6	5

**Satz 4.10** (Polynominterpolation)

Gegeben seien  $n$  Paare  $(x_i, y_i)$ ,  $i = 1, \dots, n$  mit verschiedenen Stützstellen  $x_i \neq x_j$  für  $i \neq j$ . Dann gibt genau ein **Interpolationspolynom**  $p_{n-1}$  vom Grad  $\leq n - 1$ , das die Punkte interpoliert:  $p_{n-1}(x_i) = y_i$  für alle  $i = 1, \dots, n$ .  $\diamond$

**Beispiel 4.11 (Interpolation)**

Zu zwei Punkten  $(x_1, y_1)$  und  $(x_2, y_2)$  mit  $x_1 \neq x_2$  ist das eindeutige Interpolationspolynom vom Grad  $\leq 1$  gegeben durch

$$p_1(x) = y_1 + \frac{y_2 - y_1}{x_2 - x_1}(x - x_1). \quad \diamond$$

Im allgemeinen ergibt das Einsetzen der  $n$  Punkte  $(x_1, y_1), \dots, (x_n, y_n)$  in den Ansatz

$$p_{n-1}(x) = a_{n-1} x_i^{n-1} + \dots + a_1 x_i + a_0 = y_i$$

ein lineares Gleichungssystem mit  $n$  Gleichungen für die  $n$  Unbekannten  $a_0, a_1, \dots, a_{n-1}$ . Praktischer ist manchmal jedoch die Berechnung des Interpolationspolynoms in der **Lagrange-Form**:

$$p_{n-1}(x) = \sum_{i=1}^n y_i L_i(x).$$

Dabei ist

$$L_i(x) = \prod_{k=1, k \neq i}^n \frac{x - x_k}{x_i - x_k}$$

das  $i$ -te Lagrange-Polynom, das zu den Stützstellen  $x_1, \dots, x_n$  gehört. Es ist vom Grad  $n - 1$  und hat die Eigenschaft

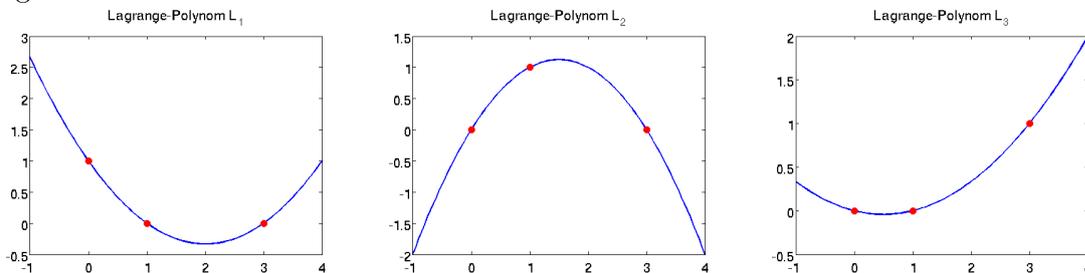
$$L_i(x_j) = \begin{cases} 1 & \text{für } i = j \\ 0 & \text{für } i \neq j. \end{cases}$$

**Beispiel 4.12 (Lagrange-Polynome bei  $n = 3$  Stützstellen)**

Für drei Stützstellen  $x_1 = 0, x_2 = 1, x_3 = 3$  haben die drei Lagrange-Polynome

$$\begin{aligned} L_1(x) &= \frac{x - x_2}{x_1 - x_2} \cdot \frac{x - x_3}{x_1 - x_3} \\ L_2(x) &= \frac{x - x_1}{x_2 - x_1} \cdot \frac{x - x_3}{x_2 - x_3} \\ L_3(x) &= \frac{x - x_1}{x_3 - x_1} \cdot \frac{x - x_2}{x_3 - x_2} \end{aligned}$$

folgenden Verlauf:



$\diamond$

§ 4.2 Grenzwerte und Stetigkeit

**Definition 4.13 (rechts- und linksseitige Grenzwerte)**

Gegeben sei ein Intervall  $I \subseteq \mathbb{R}$  und ein Punkt  $a \in I$  sowie  $f : I \setminus \{a\} \rightarrow \mathbb{R}$

- (a) Es sei  $a$  nicht der rechte Endpunkt von  $I$ . Falls für *jede* Folge  $(x_n)_{n \in \mathbb{N}}$  mit der Eigenschaft  $x_n > a$  und  $\lim_{n \rightarrow \infty} x_n = a$  gilt: die Folge  $(f(x_n))_{n \in \mathbb{N}}$  konvergiert gegen  $c \in \mathbb{R}$ , so heißt  $c$  der **rechtsseitige Grenzwert** von  $f$  für  $x$  gegen  $a$ . Man schreibt

$$\lim_{x \rightarrow a^+} f(x) = c \quad \text{oder} \quad \lim_{x \searrow a} f(x) = c.$$

- (b) Es sei  $a$  nicht der linke Endpunkt von  $I$ . Falls für *jede* Folge  $(x_n)_{n \in \mathbb{N}}$  mit der Eigenschaft  $x_n < a$  und  $\lim_{n \rightarrow \infty} x_n = a$  gilt: die Folge  $(f(x_n))_{n \in \mathbb{N}}$  konvergiert gegen  $d \in \mathbb{R}$ , so heißt  $d$  der **linksseitige Grenzwert** von  $f$  für  $x$  gegen  $a$ . Man schreibt

$$\lim_{x \rightarrow a^-} f(x) = d \quad \text{oder} \quad \lim_{x \nearrow a} f(x) = d.$$

- (c) Falls die rechts- und linksseitigen Grenzwerte von  $f$  bei  $a$  existieren und übereinstimmen, so nennt man dies den **Grenzwert** von  $f$  für  $x$  gegen  $a$  und schreibt

$$\lim_{x \rightarrow a} f(x) = \lim_{x \rightarrow a^+} f(x) = \lim_{x \rightarrow a^-} f(x)$$

**Beachte:** Der evtl. definierte Funktionswert  $f(a)$  spielt beim Grenzwert keine Rolle!

- (d) Der rechtsseitige Grenzwert aus (a) kann auf den Fall  $a = -\infty$  erweitert werden. Man schreibt dann:

$$\lim_{x \rightarrow -\infty} f(x) = c.$$

- (e) Der linksseitige Grenzwert aus (b) kann auf den Fall  $a = \infty$  erweitert werden. Man schreibt dann:

$$\lim_{x \rightarrow \infty} f(x) = d.$$

◇

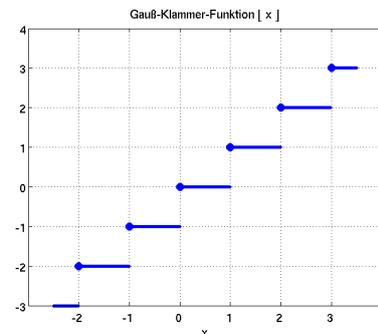
Ende 14. V

27.01.2011

### Beispiel 4.14 (Rechts- und linksseitige Grenzwerte)

Zu  $x \in \mathbb{R}$  bezeichnet  $\lfloor x \rfloor$  die größte ganze Zahl, die kleiner oder gleich  $x$  ist (**Gaußklammer**). Es gilt:

$$\begin{aligned} \lim_{x \rightarrow a^+} \lfloor x \rfloor &= a && \text{für } a \in \mathbb{Z} \\ \lim_{x \rightarrow a^-} \lfloor x \rfloor &= a - 1 && \text{für } a \in \mathbb{Z} \\ \lim_{x \rightarrow a^-} \lfloor x \rfloor &= \lim_{x \rightarrow a^+} \lfloor x \rfloor = \lfloor a \rfloor && \text{für } a \in \mathbb{R} \setminus \mathbb{Z}. \end{aligned}$$



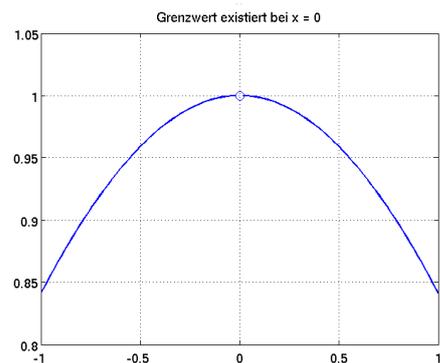
◇

### Beispiel 4.15 (Grenzwerte)

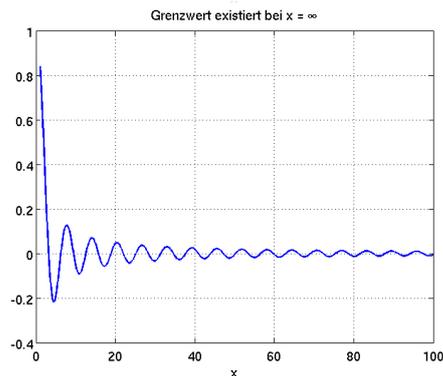
$$(a) \lim_{x \rightarrow 2} \frac{x^2 - 4}{(x - 2)(x + 1)} = \lim_{x \rightarrow 2} \frac{x + 2}{x + 1} = \frac{4}{3}$$



$$(b) \lim_{x \rightarrow 0} \frac{\sin x}{x} = 1$$



$$(c) \lim_{x \rightarrow \infty} \frac{\sin x}{x} = 0$$



◇

#### Definition 4.16 (Uneigentliche Grenzwerte)

Die Definition 4.13 lässt sich auf die Fälle  $c = \pm\infty$  oder  $d = \pm\infty$  erweitern, wenn die Folge der Funktionswerte  $(f(x_n))_{n \in \mathbb{N}}$  bestimmt gegen  $\infty$  oder  $-\infty$  divergiert. Man schreibt dann z.B.

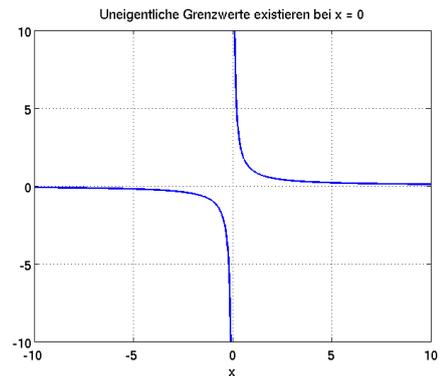
$$\lim_{x \rightarrow a^+} f(x) = \infty \quad \text{oder} \quad \lim_{x \rightarrow \infty} f(x) = -\infty.$$

und nennt dies **uneigentliche Grenzwerte**.

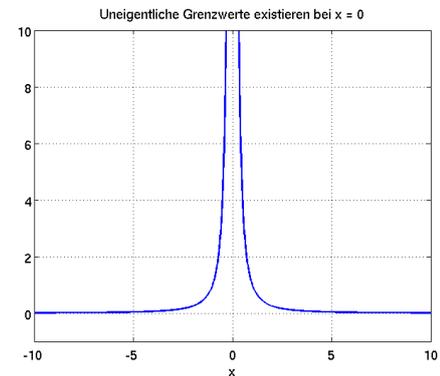
◇

#### Beispiel 4.17 (Uneigentliche Grenzwerte)

$$(a) \lim_{x \rightarrow 0^+} \frac{1}{x} = \infty \text{ und } \lim_{x \rightarrow 0^-} \frac{1}{x} = -\infty$$



$$(b) \lim_{x \rightarrow 0^+} \frac{1}{x^2} = \lim_{x \rightarrow 0^-} \frac{1}{x^2} = \lim_{x \rightarrow 0} \frac{1}{x^2} = \infty$$



### Definition 4.18 (Stetigkeit)

Gegeben sei ein offenes Intervall  $I \subseteq \mathbb{R}$  und ein Punkt  $a \in I$  sowie  $f : I \rightarrow \mathbb{R}$ .

- (a) Die Funktion  $f$  heißt **stetig** im Punkt  $a$ , wenn der Grenzwert  $\lim_{x \rightarrow a} f(x) \in \mathbb{R}$  existiert und gilt:

$$\lim_{x \rightarrow a} f(x) = f(a).$$

(Der Grenzwert an der Stelle  $x = a$  stimmt also mit dem Funktionswert überein.)

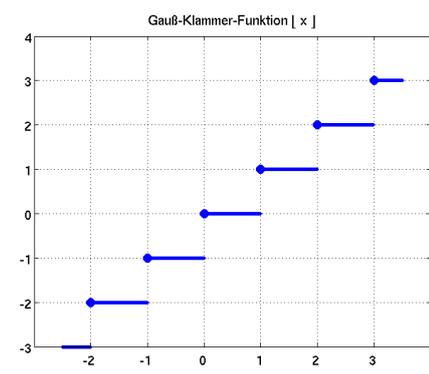
- (b) Andernfalls heißt  $f$  **unstetig** im Punkt  $a$ . (Entweder existiert der Grenzwert von  $f$  für  $x$  gegen  $a$  nicht, oder er weicht vom Funktionswert ab.)
- (c) Falls  $f(x)$  in jedem Punkt  $x \in I$  stetig ist, so heißt  $f$  **stetig** auf  $I$ .  $\diamond$

### Bemerkung 4.19

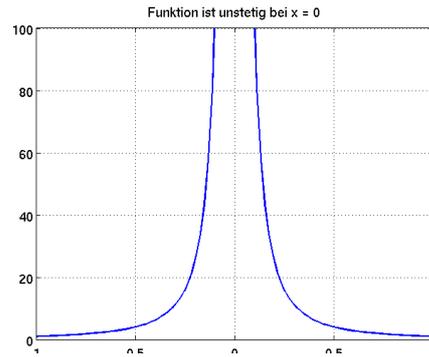
Eine Funktion ist stetig in einem Punkt  $a$ , wenn man sie in einer Umgebung von  $a$  zeichnen kann, ohne den Stift abzusetzen.  $\diamond$

### Beispiel 4.20 (Stetige und unstetige Funktionen)

- (a) Die Funktion  $f(x) = \lfloor x \rfloor$  (siehe Beispiel 4.14) ist unstetig in allen Punkten  $a \in \mathbb{Z}$ , da rechts- und linksseitige Grenzwerte verschieden sind, d.h., der Grenzwert  $\lim_{x \rightarrow a} f(x)$  existiert nicht. In allen anderen Punkten  $a \in \mathbb{R} \setminus \mathbb{Z}$  ist die Funktion stetig.

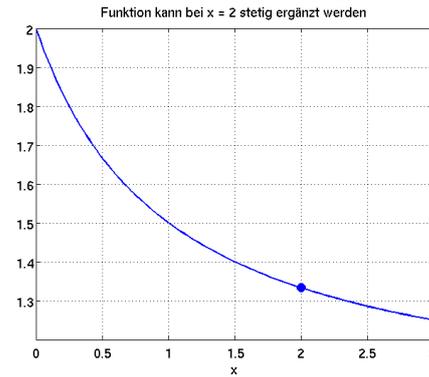


(b) Die Funktion  $f(x) = \frac{1}{x^2}$  mit Definitionsbereich  $D(f) = \mathbb{R} \setminus \{0\}$  ist unstetig im Punkt 0. Es existiert zwar (im uneigentlichen Sinne)  $\lim_{x \rightarrow 0} f(x) = \infty$  (siehe Beispiel 4.17 (b)), aber in Definition 4.18 ist  $\lim_{x \rightarrow a} f(x) \in \mathbb{R}$  gefordert.



(c) Die Funktion  $f(x) = \frac{x^2-4}{(x-2)(x+1)}$  ist an der Stelle  $x = 2$  nicht definiert (also unstetig), besitzt dort aber den Grenzwert  $\lim_{x \rightarrow 2} f(x) = \frac{4}{3}$  (siehe Beispiel 4.15 (a)).

Man kann eine **stetige Ersatzfunktion** für  $f$  bilden, indem man  $f(x)$  durch Setzen von  $f(2) := \frac{4}{3}$  stetig ergänzt.



◇

**Satz 4.21 (Rechenregeln für stetige Funktionen)**

Es seien  $f$  und  $g$  stetig im Punkt  $a$ . Dann sind auch folgende Funktionen stetig in diesem Punkt:

- (a)  $f(x) \pm g(x)$
- (b)  $c f(x)$  für alle  $c \in \mathbb{R}$
- (c)  $f(x) \cdot g(x)$
- (d)  $\frac{f(x)}{g(x)}$ , falls  $g(a) \neq 0$  ist.

◇

**§ 4.3 Differentialrechnung**

**Definition 4.22 (Differenzierbarkeit und Tangente)**

Gegeben sei ein offenes Intervall  $I \subseteq \mathbb{R}$  und ein Punkt  $a \in I$  sowie  $f : I \rightarrow \mathbb{R}$ .

(a) Der Ausdruck

$$\frac{f(a+h) - f(a)}{h}, \quad h \neq 0$$

heißt ein **Differenzenquotient** von  $f$  an der Stelle  $a$ . Er bezeichnet den mittleren Anstieg des Graphen von  $f$  im Intervall  $[a, a+h]$  bzw. (falls  $h < 0$ )  $[a+h, a]$ .

(b) Die Funktion (Gerade)

$$S(x; a, h) = f(a) + \frac{f(a+h) - f(a)}{h} (x - a)$$

heißt die **Sekante** von  $f$  durch die Punkte  $(a, f(a))$  und  $(a + h, f(a + h))$ , vergleiche Beispiel 4.11. Der Differenzenquotient gibt also die Steigung der Sekante an.

- (c) Die Funktion  $f$  heißt **differenzierbar (diffbar)** im Punkt  $a$ , wenn der Grenzwert der Differenzenquotienten

$$\lim_{h \rightarrow 0} \frac{f(a+h) - f(a)}{h} \quad \text{oder auch (setze } x = a+h) \quad \lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a}$$

in  $\mathbb{R}$  existiert. Diesen Grenzwert nennen wir die **Ableitung** von  $f$  an der Stelle  $a$ , kurz:  $f'(a)$  oder  $\frac{d}{dx}f(x)|_{x=a}$  oder  $\frac{df}{dx}(a)$ .

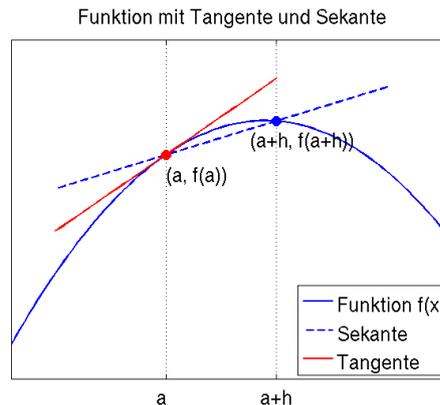
- (d) Die Funktion (Gerade)

$$T(x; a) = f(a) + f'(a)(x - a)$$

heißt die **Tangente** von  $f$  im Punkt  $a$ . Die Ableitung  $f'(a)$  gibt also die Steigung der Tangente an.

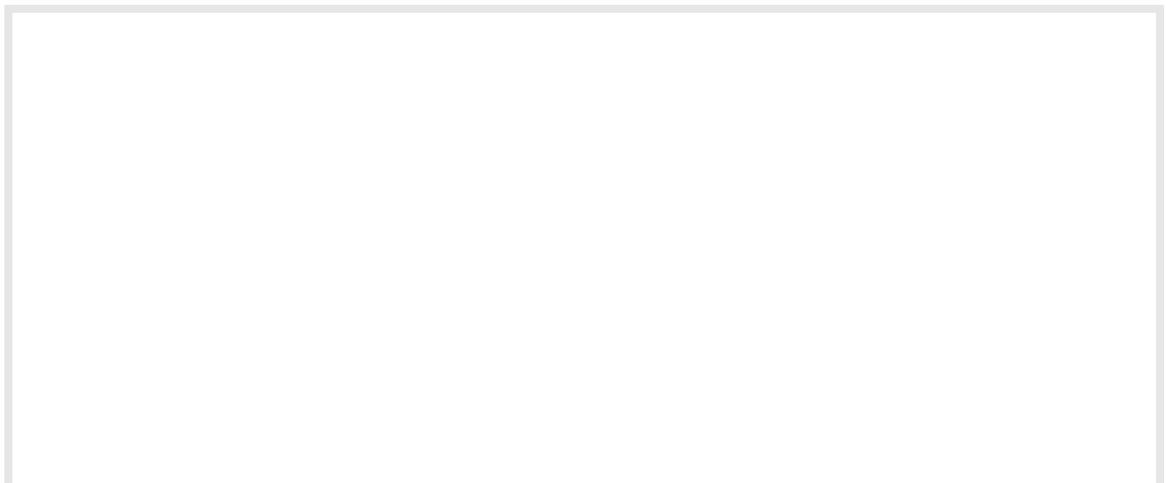
- (e) Falls die Ableitung  $f'(x)$  in jedem Punkt  $x \in I$  existiert, so heißt  $f$  **diffbar** auf  $I$ . Wenn dabei außerdem  $f'(x)$  eine stetige Funktion auf  $I$  ist, so heißt  $f$  **stetig diffbar** auf  $I$ .  $\diamond$

Für  $h \rightarrow 0$  gehen die Steigungen der Sekanten durch  $(a, f(a))$  und  $(a + h, f(a + h))$  über in die Steigung der Tangente  $f'(a)$ .



**Frage:** Welche „Einheit“ hat die Ableitung einer Funktion?

**Beispiel 4.23** (zur Einheit der Ableitung)



**Antwort:** Die Einheit der Ableitung einer Funktion ist gleich der Einheit der Funktionswerte geteilt durch die Einheit der unabhängigen Variablen.

**Bemerkung 4.24**

- (a) In den Wirtschaftswissenschaften bezeichnet man die Ableitung  $f'(x)$  auch als **Grenzfunktion** zur Funktion  $f(x)$ . Stellt z.B.  $f(x)$  eine Gewinnfunktion dar, so heißt  $f'(x)$  der **Grenzwinn** bzw. die **Grenzwinnfunktion**.
- (b) Wenn  $f'(a)$  existiert, dann ist  $f$  notwendig stetig an der Stelle  $a$ . Eine Funktion, die an einer Stelle  $a$  springt, kann dort also keine Ableitung besitzen.
- (c) Wenn  $f'(a)$  existiert, so bedeutet das:

$$\lim_{x \rightarrow a} \frac{f(x) - f(a)}{x - a} = f'(a) \quad \text{oder auch}$$

$$\lim_{x \rightarrow a} \frac{f(x) - \overbrace{(f(a) + f'(a)(x - a))}^{\text{Tangente}}}{x - a} = 0.$$

Mit anderen Worten: Die Funktion  $f(x)$  wird in der Nähe von  $a$  so gut durch ihre Tangente approximiert (angenähert), dass für das **Restglied**

$$r(x; a) := f(x) - (f(a) + f'(a)(x - a))$$

gilt:

$$\lim_{x \rightarrow a} \frac{r(x; a)}{x - a} = 0.$$

Man sagt: Das Restglied konvergiert für  $x \rightarrow a$  schneller gegen null als  $x - a$ .

- (d) Keine andere Gerade approximiert  $f(x)$  in der Nähe von  $a$  so gut wie die Tangente. Die Tangente  $T(x; a)$  ist also das beste **lineare Modell** für  $f(x)$  in der Nähe von  $a$ , das man finden kann!  $\diamond$

**Beispiel 4.25 (Approximation einer Funktion durch ihre Tangente)**

Die Tangente kann z.B. benutzt werden, um  $f(x)$  zu schätzen, wenn man  $f(a)$  und  $f'(a)$  für ein  $a$  nahe bei  $x$  kennt. Wir wollen  $\sqrt{2}$  näherungsweise berechnen und setzen  $f(x) = \sqrt{x}$ , also  $f'(x) = \frac{1}{2\sqrt{x}}$ . Wir wählen  $a = 1,96$ , denn wir kennen  $f(a) = \sqrt{1,96} = 1,4$  und  $f'(a) = \frac{1}{2 \cdot 1,4}$ . Es gilt:

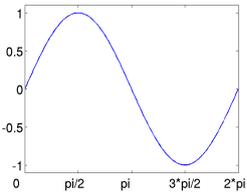
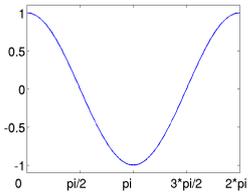
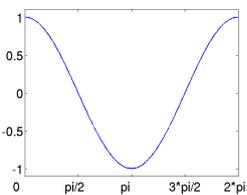
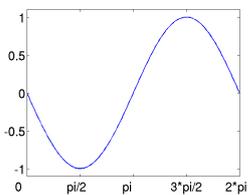
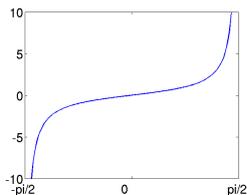
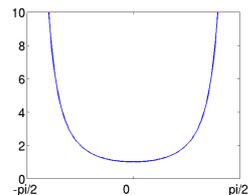
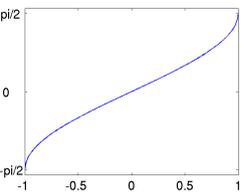
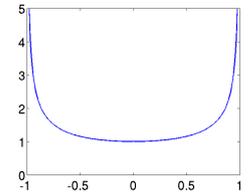
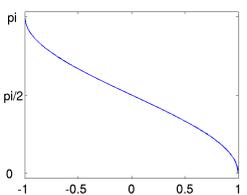
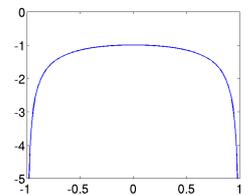
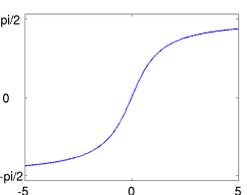
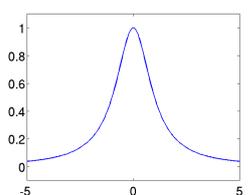
$$\begin{aligned} \sqrt{2} &= f(x) \approx f(a) + f'(a) \cdot (x - a) \\ &= 1,4 + \frac{1}{2,8} \cdot (2 - 1,96) \\ &= 1,4 + 0,3571 \cdot (2 - 1,96) \\ &= 1,4142857 \dots \end{aligned}$$

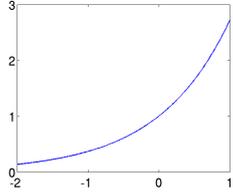
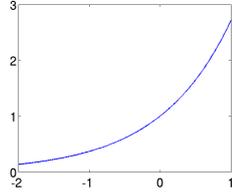
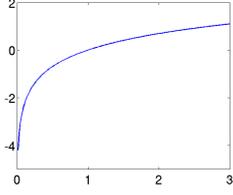
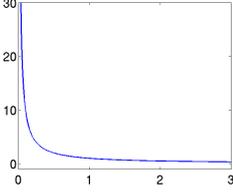
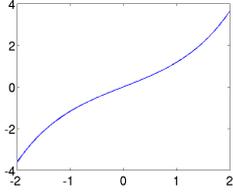
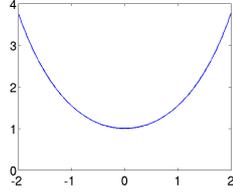
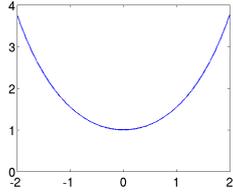
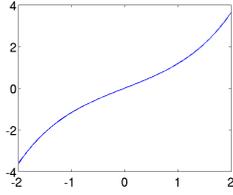
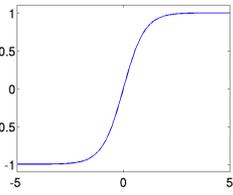
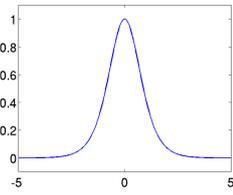
Für eine Abschätzung des Fehlers siehe Beispiel 4.48.  $\diamond$

Ende 15. V

03.02.2011

**Beispiel 4.26** (Elementare Funktionen und ihre Ableitungen)

$f(x)$	Abbildung	$f'(x)$	Abbildung	Bemerkung
$x^\alpha$		$\alpha x^{\alpha-1}$		Definitionsbereich hängt von $\alpha \in \mathbb{R}$ ab
$\sin(x)$		$\cos(x)$		$x \in \mathbb{R}$
$\cos(x)$		$-\sin(x)$		$x \in \mathbb{R}$
$\tan(x)$		$\frac{1}{\cos^2(x)}$		$x \neq (2k+1)\frac{\pi}{2}$ für $k \in \mathbb{Z}$
$\arcsin(x)$		$\frac{1}{\sqrt{1-x^2}}$		$ x  < 1$
$\arccos(x)$		$\frac{-1}{\sqrt{1-x^2}}$		$ x  < 1$
$\arctan(x)$		$\frac{1}{1+x^2}$		$x \in \mathbb{R}$

$f(x)$	Abbildung	$f'(x)$	Abbildung	Bemerkung
$e^x = \exp(x)$		$e^x$		$x \in \mathbb{R}$
$a^x$		$a^x \ln(a)$		$a > 0, \quad x \in \mathbb{R}$
$\ln(x)$		$\frac{1}{x}$		$x > 0$
$\log_a(x)$		$\frac{1}{x \ln(a)}$		$x > 0$
$\sinh(x)$		$\cosh(x)$		$x \in \mathbb{R}$
$\cosh(x)$		$\sinh(x)$		$x \in \mathbb{R}$
$\tanh(x)$		$\frac{1}{\cosh^2(x)}$		$x \in \mathbb{R}$

◇

**Satz 4.27 (Rechenregeln für die Ableitung)**

Es seien  $f$  und  $g$  diffbar im Punkt  $a$ . Dann existieren folgende Ableitungen:

(a)  $(f \pm g)'(a) = f'(a) \pm g'(a)$

(b)  $(cf)'(a) = cf'(a)$  für alle  $c \in \mathbb{R}$

(c)  $(f \cdot g)'(a) = f'(a) \cdot g(a) + f(a) \cdot g'(a)$

(d)  $\left(\frac{f}{g}\right)'(a) = \frac{f'(a) \cdot g(a) - f(a) \cdot g'(a)}{g^2(a)}$ , falls  $g(a) \neq 0$

(e) insbesondere  $\left(\frac{1}{g}\right)'(a) = \frac{-g'(a)}{g^2(a)}$ , falls  $g(a) \neq 0$  ◇

**Satz 4.28 (Kettenregel)**

Es sei  $g$  diffbar an der Stelle  $a$  und  $f$  diffbar an der Stelle  $g(a)$ . Dann ist  $f \circ g$  diffbar an der Stelle  $a$ , und es gilt

$$(f \circ g)'(a) = f'(g(a)) \cdot g'(a). \quad \diamond$$

**Beispiel 4.29 (Kettenregel)**

(a) Was ist die Ableitung von  $h(x) = (x^4 + 6x + 5)^3$ ?

(b) Was ist die Ableitung von  $h(x) = [\sin(x^4 + 2x)^2]^5$ ?

## § 4.4 Anwendungen der Differentialrechnung

Es folgen drei Anwendungen der Differentialrechnung.

### § 4.4.1 Das Newton-Verfahren

**Frage:** Wie kann man Nullstellen einer diffbaren Funktion  $f(x) = 0$  bestimmen?

Wir legen im Punkt  $x_0$  (Schätzung der Nullstelle von  $f$ ) die Tangente an die Funktion und bestimmen *deren* Nullstelle. Dies geht sehr einfach, da die Tangente ja eine lineare Funktion von  $x$  ist:

$$f(x_0) + f'(x_0)(x - x_0) = 0 \quad \Leftrightarrow \quad x = x_0 - \frac{f(x_0)}{f'(x_0)}.$$

Dies gilt zumindest dann, wenn die Tangente im Punkt  $x_0$  nicht horizontal verläuft, falls also  $f'(x_0) \neq 0$  ist. Im **Newton-Verfahren** dient die Nullstelle der Tangente als nächste Schätzung  $x_1$  der Nullstelle von  $f$  usw.:

```

1: Setze  $n := 0$ 
2: Wähle einen Startwert  $x_0$ 
3: while  $|f(x_n)| > \varepsilon$  do
4:   Berechne  $x_{n+1} := x_n - \frac{f(x_n)}{f'(x_n)}$ 
5:   Setze  $n := n + 1$ 
6: end while

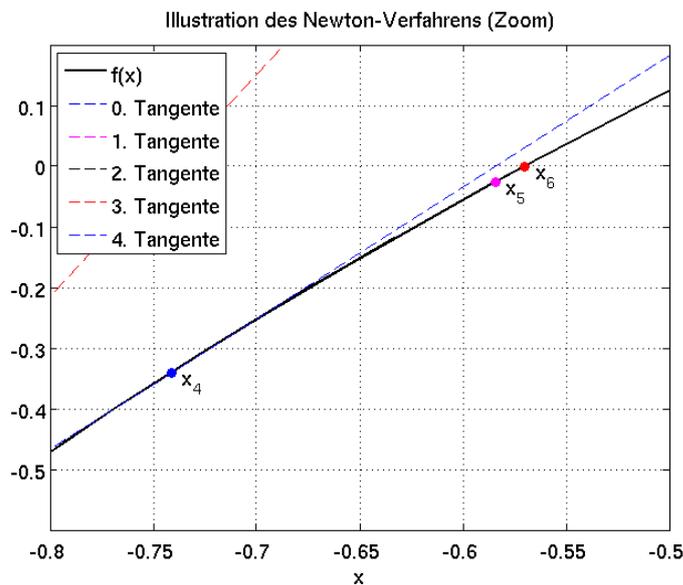
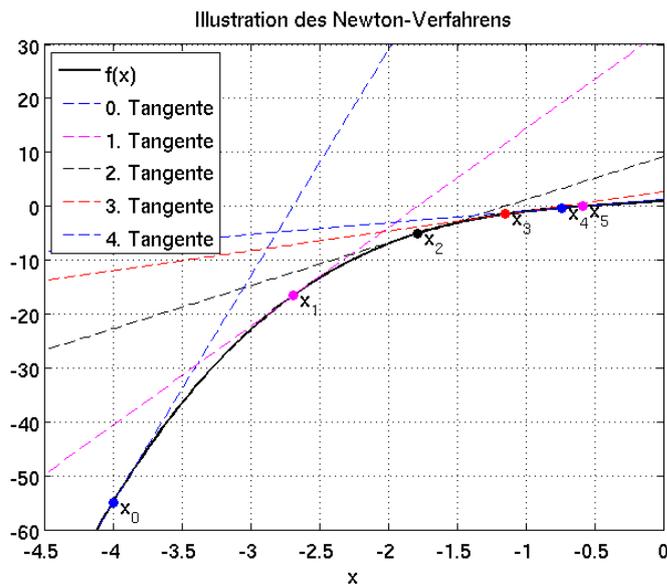
```

**Beachte:** Das Newton-Verfahren konvergiert in der obigen einfachen Form nicht immer. Man profitiert aber davon, den Startwert  $x_0$  möglichst gut zu wählen.

#### Beispiel 4.30

Bestimmung der Nullstelle von  $f(x) = x^3 + x^2 + 2x + 1$ : Ausgehend vom Startwert  $x_0 = -4$  erzeugt das Newton-Verfahren folgende Iterierte (korrekte Nachkommastellen sind unterstrichen):

$n$	$x_n$	$f(x_n)$
0	-4.000000000000000	-55.000000000000000
1	-2.69047619047619	-16.61773836518735
2	-1.78413817371242	-5.06430500889450
3	-1.14960657606113	-1.49693249686015
4	-0.74123041640946	-0.34028699346109
5	-0. <u>58411252385808</u>	-0.02632946418370
6	-0. <u>56992132285406</u>	-0.00014865543482
7	-0. <u>56984029353807</u>	-0.00000000465960
8	-0. <u>56984029099805</u>	-0.00000000000000



### § 4.4.2 Lineare Fehlerrechnung

**Frage:** Wie wirken sich Messfehler (Ungenauigkeiten)  $\Delta x$  in einer Größe  $x$  auf ein berechnetes Ergebnis  $f(x)$  aus?

Mit Hilfe der linearen Approximation (Tangente) erhält man

$$\underbrace{\Delta f}_{\text{Fehler im Ergebnis}} := f(x + \Delta x) - f(x) \approx f'(x) \underbrace{\Delta x}_{\text{Fehler in } x}.$$

Der Fehler in  $x$  (z.B. Messdaten) wird also in etwa mit dem Faktor  $f'(x)$  verstärkt (bzw. abgeschwächt). Die Ableitung  $f'(x)$  gibt damit die Empfindlichkeit (**Sensitivität**) von  $f$  bzgl. Störungen in  $x$  an. Die **relative Empfindlichkeit** von  $f$  beträgt damit

$$\underbrace{\frac{\Delta f}{f(x)}}_{\text{relativer Fehler im Ergebnis}} \approx \frac{f'(x)}{f(x)} \Delta x = \frac{f'(x)}{f(x)} x \underbrace{\frac{\Delta x}{x}}_{\text{relativer Fehler in } x}.$$

In den Wirtschaftswissenschaften bezeichnet man den Faktor bei der relativen Empfindlichkeit

$$\frac{f'(x)}{f(x)} x$$

auch als **Elastizität** der Funktion  $f(x)$ . Sie ist eine dimensionslose Größe, vergleiche Beispiel 4.23, denn  $f'(x) \cdot x$  hat wieder dieselbe Einheit wie  $f(x)$ .

**Bemerkung 4.31 (Differential)**

Die Größe  $df := f'(x) \Delta x$  heißt auch das **Differential** der Funktion  $f$  an der Stelle  $x$  zum Inkrement  $\Delta x$ .

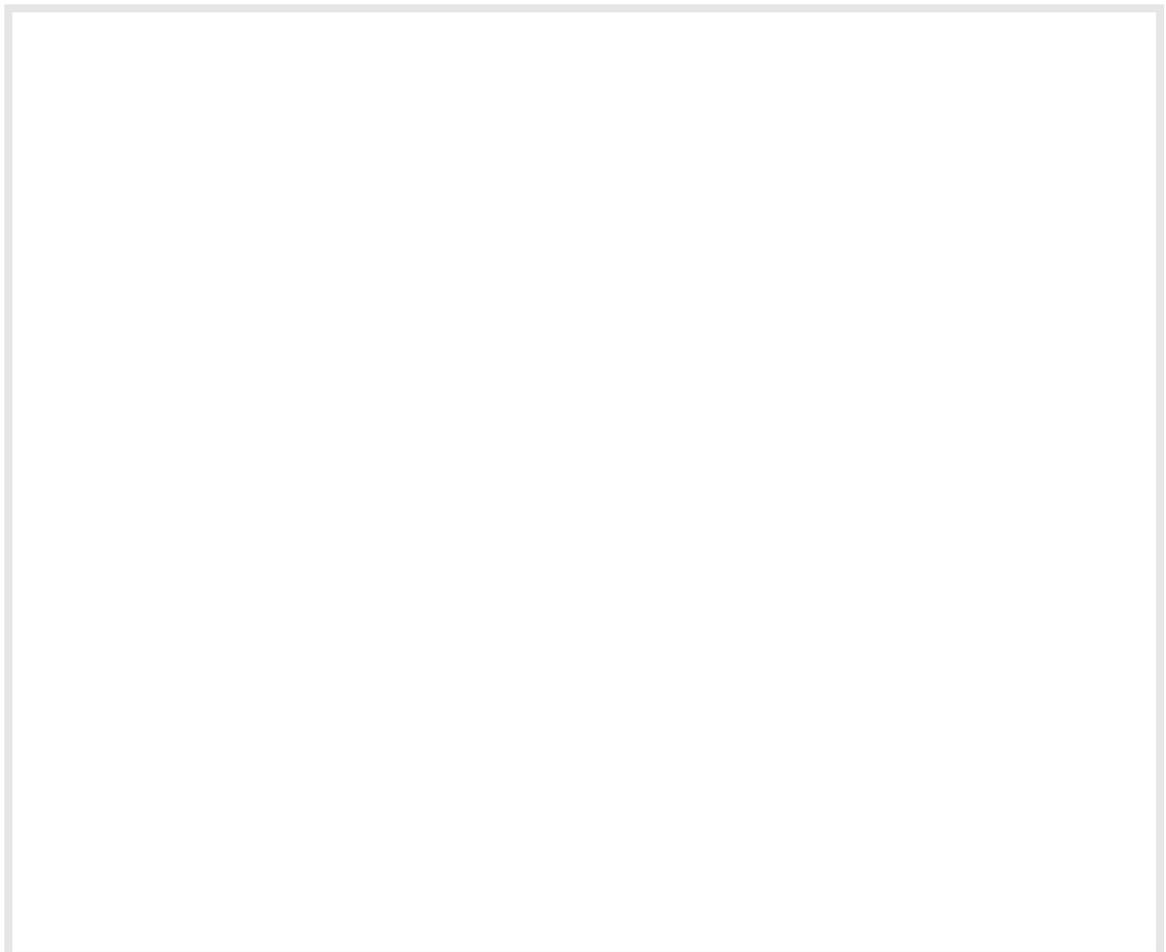
**Beispiel 4.32 (Lineare Fehlerrechnung)**

Auf einer Messstrecke der Länge  $s$  wird die Durchfahrzeit  $t$  eines Autos (z.B. mit Lichtschranken) gestoppt und daraus die Geschwindigkeit ermittelt:

$$v = \frac{s}{t}.$$

Wie wirkt sich ein Fehler  $\Delta t$  in der Zeitmessung auf das Ergebnis aus?

$$\Delta v \approx v'(t) \Delta t = \frac{-s}{t^2} \Delta t.$$



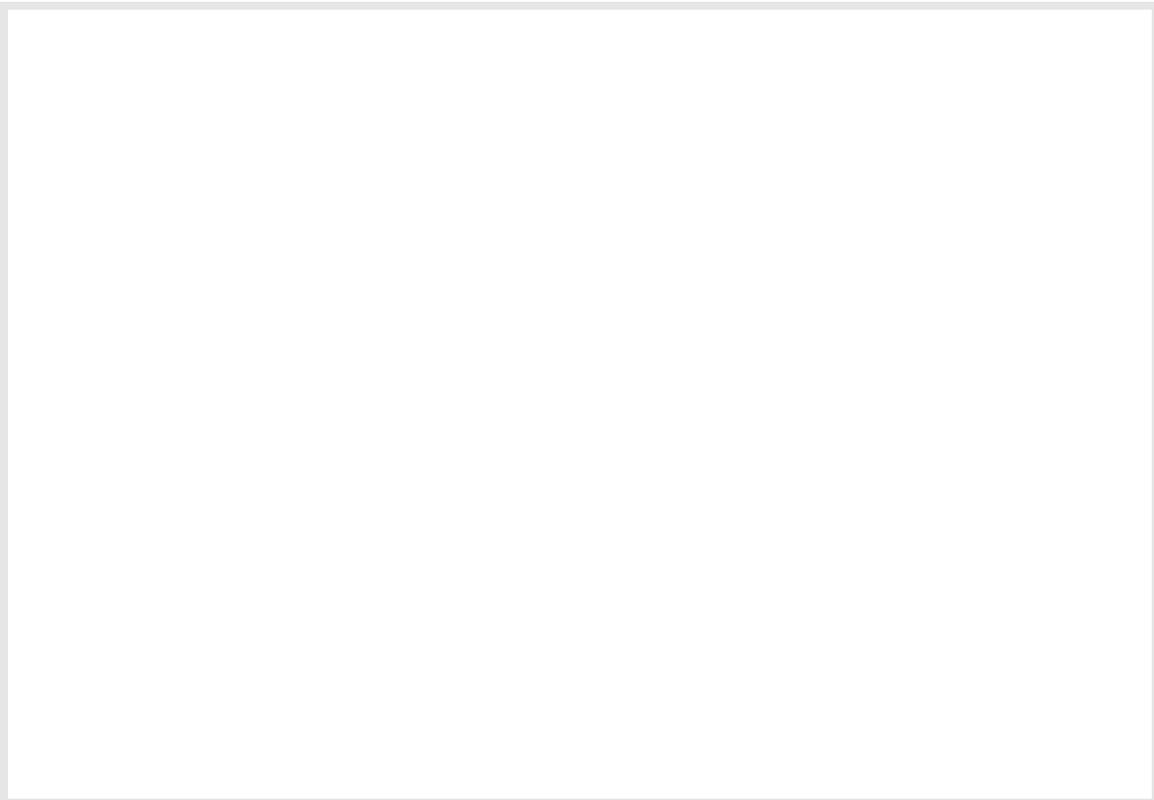
## § 4.4.3 Regel von L'Hospital

**Satz 4.33 (Regel von L'Hospital (1696))** (a) Seien  $f$  und  $g$  auf dem Intervall  $I = (a, b)$  diffbar und  $x_0 \in I$ . Falls  $\lim_{x \rightarrow x_0} f(x) = \lim_{x \rightarrow x_0} g(x) = 0$  oder  $\lim_{x \rightarrow x_0} f(x) = \lim_{x \rightarrow x_0} g(x) = \infty$  gilt, so ist

$$\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} = \lim_{x \rightarrow x_0} \frac{f'(x)}{g'(x)},$$

falls der rechte Grenzwert als eigentlicher Grenzwert (Definition 4.13 (c)) existiert.

- (b) Dasselbe gilt analog für einseitige Grenzwerte  $x \searrow a$  und  $x \nearrow b$  (Definition 4.13 (a) und (b)) sowie für Grenzwerte  $x \rightarrow \infty$  und  $x \rightarrow -\infty$  (Definition 4.13 (d) und (e)).  $\diamond$

**Beispiel 4.34 (Regel von L'Hospital)**

Ende 1. V

05.04.2011

## § 4.5 Optimierung (Kurvendiskussion) differenzierbarer Funktionen

**Definition 4.35 (Höhere Ableitungen)** (a) Die Ableitung der Ableitung von  $f(x)$  an der Stelle  $x$  bezeichnen wir mit  $f''(x)$  oder  $\frac{d}{dx} \left( \frac{d}{dx} f(x) \right)$  oder  $\frac{d^2}{dx^2} f(x)$ .

- (b) Wir definieren allgemein

$$f^{(0)}(x) := f(x), \quad f^{(1)}(x) := f'(x), \quad f^{(2)}(x) := f''(x) \quad \text{usw.}$$

- (c) Ist die  $k$ -te Ableitung  $f^{(k)}$  von  $f$  auf einem offenen Intervall  $I$  stetig, so nennen wir  $f$  auf  $I$   $k$ -mal **stetig diffbar** auf  $I$  oder kurz  $f \in C^k(I)$ .  $\diamond$

**Beispiel 4.36** (a) Wir differenzieren das (auf ganz  $\mathbb{R}$  definierte) Polynom

$$\begin{aligned} f(x) &= f^{(0)}(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0 \\ f'(x) &= f^{(1)}(x) = n a_n x^{n-1} + (n-1) a_{n-1} x^{n-2} + \cdots + a_1 \\ f''(x) &= f^{(2)}(x) = n(n-1) a_n x^{n-2} + (n-1)(n-2) a_{n-1} x^{n-3} + \cdots + 2 a_2 \\ &\vdots = \vdots \quad \vdots \\ f^{(n-1)}(x) &= n(n-1) \cdots 2 a_n x + (n-1)(n-2) \cdots 2 a_{n-1} \\ f^{(n)}(x) &= n(n-1) \cdots 2 \cdot 1 \cdot a_n = n! \cdot a_n \equiv \text{const} \\ f^{(n+1)}(x) &= 0. \end{aligned}$$

Hierbei ist die **Fakultät** einer natürlichen Zahl definiert als

$$0! := 1, \quad 1! := 1, \quad n! := 1 \cdot 2 \cdot \dots \cdot n \quad \text{für } n \in \mathbb{N}.$$

- (b) Ein konkretes Beispiel:

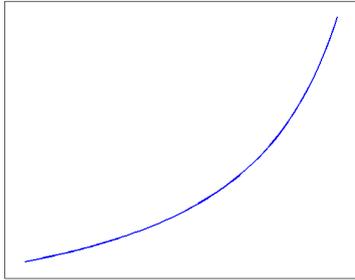
$$\begin{aligned} f(x) &= f^{(0)}(x) = -x^4 + 2x^3 + 7x^2 - 5x + 14 \\ f'(x) &= f^{(1)}(x) = -4x^3 + 6x^2 + 14x - 5 \\ f''(x) &= f^{(2)}(x) = -12x^2 + 12x + 14 \\ f^{(3)}(x) &= -24x + 12 \\ f^{(4)}(x) &= -24 \\ f^{(5)}(x) &= 0. \end{aligned}$$

$\diamond$

**Satz 4.37** (**Bedeutung der 1. und 2. Ableitung**)

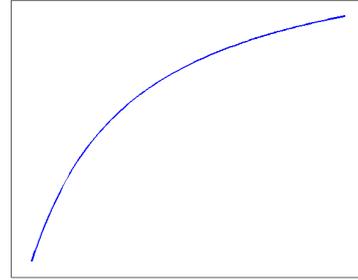
Sei  $f(x)$  auf einem offenen Intervall  $I$  einmal bzw. zweimal stetig diffbar.

- (a)  $f'(x) \geq 0$  auf  $I$   $\Leftrightarrow$   $f$  ist monoton wachsend auf  $I$   
 (b)  $f'(x) > 0$  auf  $I$   $\Rightarrow$   $f$  ist streng monoton wachsend auf  $I$   
 (c)  $f'(x) \leq 0$  auf  $I$   $\Leftrightarrow$   $f$  ist monoton fallend auf  $I$   
 (d)  $f'(x) < 0$  auf  $I$   $\Rightarrow$   $f$  ist streng monoton fallend auf  $I$   
 (e)  $f''(x) \geq 0$  auf  $I$   $\Leftrightarrow$   $f'$  ist monoton wachsend auf  $I$   
 $\Leftrightarrow$   $f$  ist **konvex** (linksgekrümmt) auf  $I$   
 (f)  $f''(x) \leq 0$  auf  $I$   $\Leftrightarrow$   $f'$  ist monoton fallend auf  $I$   
 $\Leftrightarrow$   $f$  ist **konkav** (rechtsgekrümmt) auf  $I$   $\diamond$

**Beispiel 4.38** (Wachsende und fallende, konvexe und konkave Fkt.en)

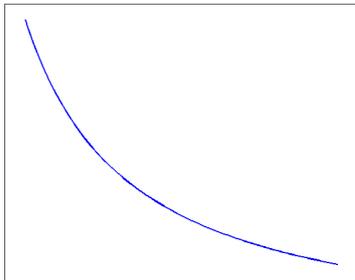
monoton wachsend und konvex

Beispiel: Erdbevölkerung über Zeit



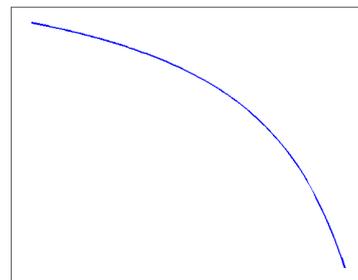
monoton wachsend und konkav

Beispiel: Gesamtpreis über Abnahmemenge



monoton fallend und konvex

Beispiel: Temperatur eines sich abkühlenden Körpers über Zeit



monoton fallend und konkav

Beispiel: Höhe über Zeit beim freien Fall

Wir wollen das Wissen aus Satz 4.37 zur *Optimierung* verwenden.

**Definition 4.39** (Extremstellen)

Sei  $f : \mathbb{R} \supseteq D(f) \rightarrow \mathbb{R}$ .

- (a)  $f$  hat im Punkt  $a \in D(f)$  ein **globales Minimum**, wenn gilt:

$$f(a) \leq f(x) \quad \text{für alle } x \in D(f).$$

**Also:** Es gibt keine Punkte im Definitionsbereich, an denen  $f(x)$  einen kleineren Funktionswert annimmt.

- (b)  $f$  hat im Punkt  $a \in D(f)$  ein **lokales Minimum**, wenn gilt

$$f(a) \leq f(x) \quad \text{für alle } x \in D(f) \cap I,$$

wobei  $I$  ein (kleines) Intervall mit Mittelpunkt  $a$  ist.

**Also:** In der Nähe von  $a$  gibt es keine Punkte, an denen  $f(x)$  einen kleineren Funktionswert annimmt.

- (c) Entsprechend sind **globales** und **lokales Maximum** definiert.

- (d) Allgemein spricht man von (lokalen oder globalen) **Extremstellen**.  $\diamond$

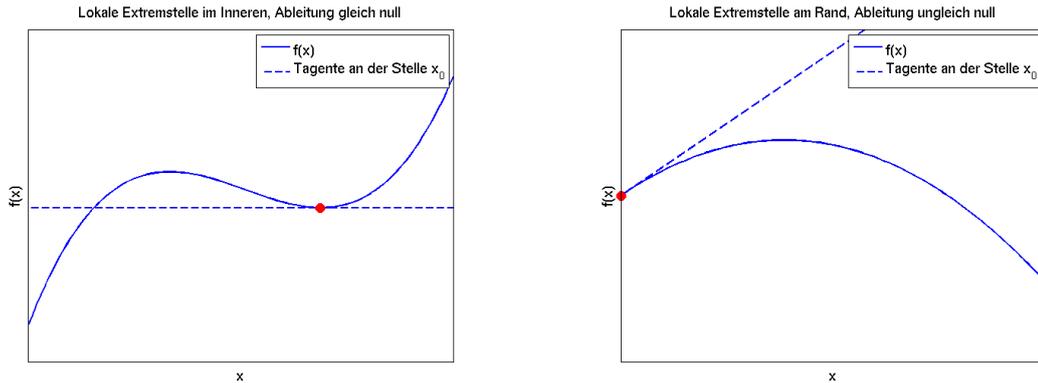
**Satz 4.40 (Notwendige Optimalitätsbedingung erster Ordnung)**

Sei  $I$  ein Intervall und  $f : I \rightarrow \mathbb{R}$  diffbar. Der Punkt  $x_0 \in I$  sei kein Randpunkt von  $I$ , d.h., er liegt im Inneren von  $I$ . Dann gilt:

$$x_0 \in I \text{ ist lokale Extremstelle} \Rightarrow f'(x_0) = 0,$$

d.h.,  $f$  hat an der Stelle  $x_0$  notwendigerweise eine waagerechte Tangente.  $\diamond$

**Beachte:** Für die Aussage des Satzes ist es wichtig, dass  $x_0$  kein Randpunkt von  $I$  ist:



An einer lokalen Extremstelle  $x_0$  im Inneren des Intervalles  $I$  ist  $f'(x_0) = 0$  (horizontale Tangente).

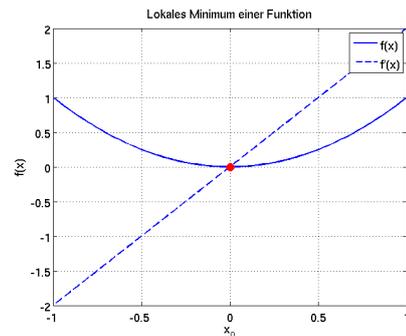
An einer lokalen Extremstelle  $x_0$ , die am Rand des Intervalles  $I$  liegt, kann  $f'(x_0) \neq 0$  sein.

**Beachte:**  $f'(x_0) = 0$  ist nur *notwendig* für das Vorliegen einer Extremstelle. Daher: Wenn möglich, zweite Ableitungen testen.

**Satz 4.41**

Sei  $I$  ein offenes Intervall,  $f : I \rightarrow \mathbb{R}$  diffbar und  $x_0 \in I$ . Wenn  $f'(x_0) = 0$  gilt und  $f'(x)$  unmittelbar links von  $x_0$  negativ und unmittelbar rechts von  $x_0$  positiv ist, dann ist  $x_0$  ein lokales Minimum.

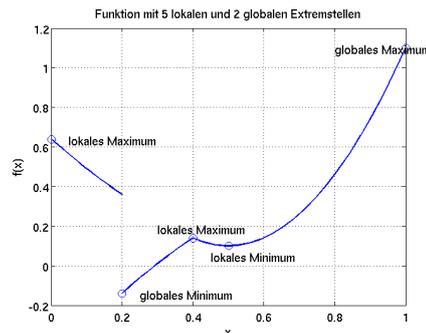
Das ist zum Beispiel dann der Fall, wenn  $f''(x)$  existiert und stetig ist und  $f''(x_0) > 0$  gilt (d.h., die erste Ableitung  $f'(x)$  wächst in der Nähe von  $x_0$ , siehe Satz 4.37).



**Bemerkung 4.42 (Kandidaten für Extremstellen)**

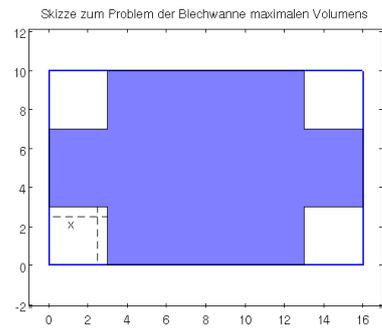
Kandidaten für Extremstellen sind:

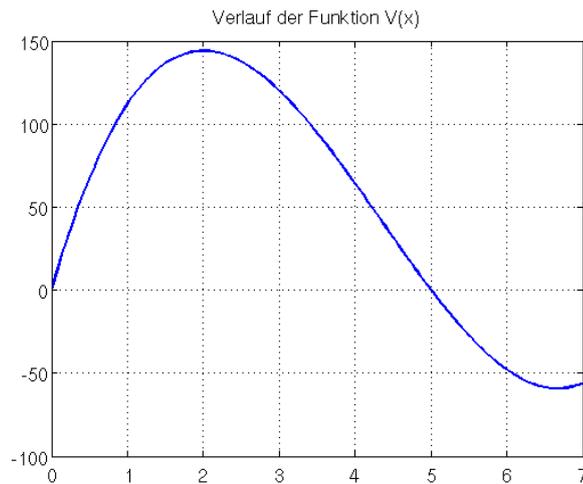
- (a) die Randpunkte des Definitionsbereiches  $D(f)$ ,
- (b) die Punkte in  $D(f)$ , in denen  $f$  nicht diffbar ist,
- (c) die Punkte im Inneren des Definitionsbereiches  $D(f)$ , an denen  $f'(x) = 0$  gilt. (Solche Punkte heißen **stationäre Punkte**.)



**Beispiel 4.43 (Blechzuschnitt)**

Aus einer Blechplatte mit Seitenlängen 16 cm und 10 cm soll eine oben offene Wanne mit möglichst großem Volumen geformt werden. Dazu muss man in den Ecken Quadrate der Seitenlänge  $x$  entfernen ( $0 \leq x \leq 5$ ).





### § 4.6 Taylorpolynome

**Frage:** In Beispiel 4.25 hatten wir eine Funktion durch ihre Tangente (also ein lineares Polynom) approximiert (angenähert). Kann man die Approximation verbessern, indem man Polynome höheren Grades nimmt? Gibt es eine Fehlerabschätzung?

#### Definition 4.44 (Taylorpolynom)

Das Polynom  $n$ -ten Grades

$$\begin{aligned} T_n(x; a) &= \sum_{k=0}^n \frac{f^{(k)}(a)}{k!} (x-a)^k \\ &= f(a) + \frac{f'(a)}{1!} (x-a) + \frac{f''(a)}{2!} (x-a)^2 + \frac{f^{(3)}(a)}{3!} (x-a)^3 + \dots + \frac{f^{(n)}(a)}{n!} (x-a)^n \end{aligned}$$

heißt das  $n$ -te **Taylorpolynom** oder das **Taylorpolynom  $n$ -ter Ordnung** der Funktion  $f$  mit der Entwicklungsstelle  $a$ .  $\diamond$

**Beachte:** Das *erste Taylorpolynom* einer Funktion  $f$  mit der Entwicklungsstelle  $a$  ist ihre *Tangente* an der Stelle  $a$ , die wir bisher mit  $T(x; a)$  bezeichnet haben. Sie stimmt an der Stelle  $a$  in Funktionswert und erster Ableitung mit  $f$  überein:

$$T_1(a; a) = f(a) \quad \text{und} \quad T_1'(a; a) = f'(a).$$

Bei Taylorpolynomen höheren Grades stimmen entsprechend mehr Ableitungen mit denen von  $f$  an der Entwicklungsstelle  $a$  überein.

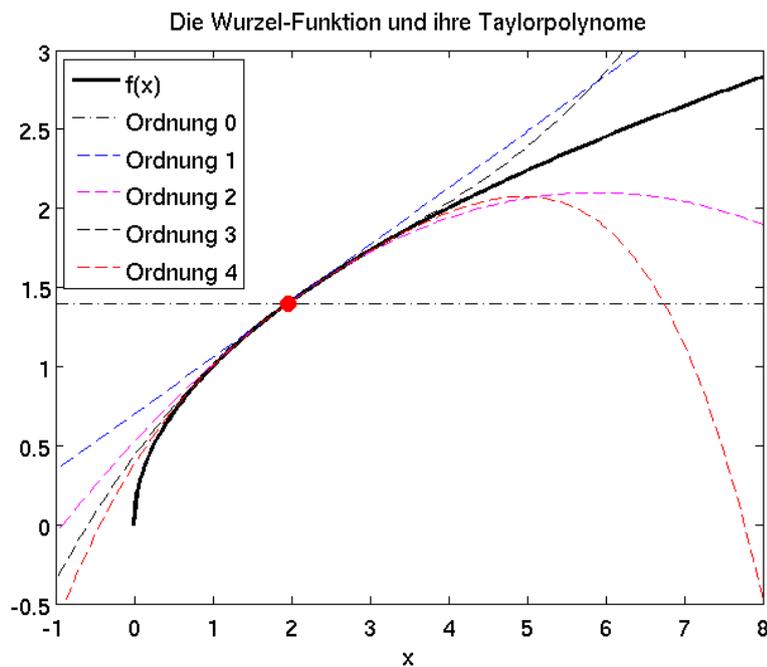
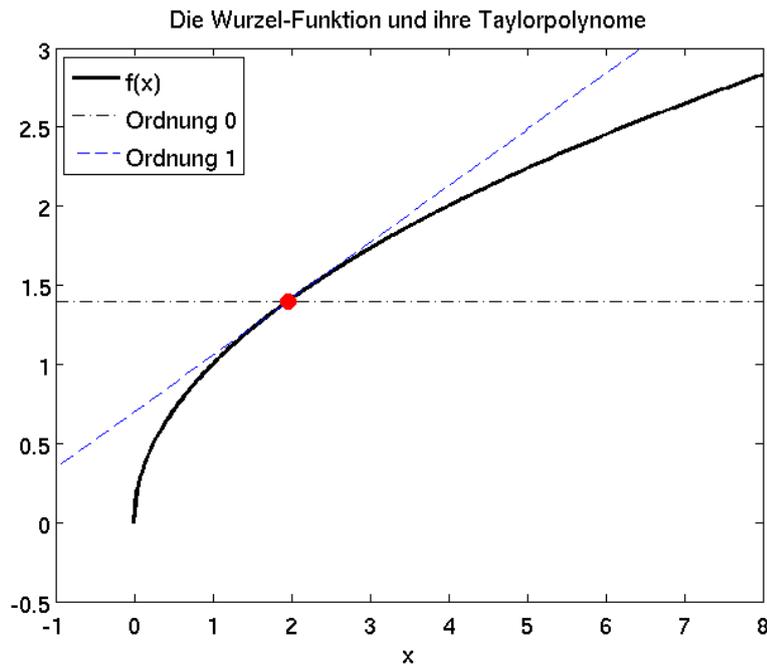
#### Beispiel 4.45

Die Wurzelfunktion hat folgende Ableitungen an der Stelle  $a = 1$ , 96:

$f(x) = \sqrt{x} = x^{1/2}$	$f(a) = \sqrt{a} = 1,4000$
$f'(x) = \frac{1}{2} x^{-1/2}$	$f'(a) = \frac{1}{2} a^{-1/2} = 0,3571$
$f''(x) = -\frac{1}{4} x^{-3/2}$	$f''(a) = -\frac{1}{4} a^{-3/2} = -0,0911$
$f^{(3)}(x) = \frac{3}{8} x^{-5/2}$	$f^{(3)}(a) = \frac{3}{8} a^{-5/2} = 0,0697$
$f^{(4)}(x) = -\frac{15}{16} x^{-7/2}$	$f^{(4)}(a) = -\frac{15}{16} a^{-7/2} = -0,0889$

Das vierte Taylorpolynom von  $f(x)$  mit der Entwicklungsstelle  $a = 1,96$  ist also

$$T_4(x; 1,96) = \underbrace{1,4000 + \frac{0,3571}{1}(x - 1,96)}_{\text{Tangente, vergleiche Beispiel 4.25}} - \frac{0,0911}{2}(x - 1,96)^2 + \frac{0,0697}{6}(x - 1,96)^3 - \frac{0,0889}{24}(x - 1,96)^4.$$



◇

Ende 2. V

12.04.2011

**Satz 4.46** (Satz von Taylor)

Es sei  $I$  ein offenes Intervall,  $f : I \rightarrow \mathbb{R}$  eine auf  $I$   $(n+1)$ -mal stetig diffbare Funktion

und  $a \in I$ . Dann ist

$$f(x) = T_n(x; a) + R_{n+1}(x; a) \quad \text{für alle } x \in I,$$

wobei für das **Restglied** (in **Lagrange-Form**) gilt:

$$R_{n+1}(x; a) = \frac{f^{(n+1)}(\xi)}{(n+1)!} (x-a)^{n+1}$$

mit einem (unbekannten)  $\xi$  zwischen  $a$  und  $x$ . ◇

**Folgerung 4.47 (Fehlerabschätzung)**

Obwohl man  $\xi$  nicht kennt, kann man oft eine Abschätzung für  $f^{(n+1)}(\xi)$  finden. Damit kann man den Fehler zwischen  $f(x)$  und  $T_n(x; a)$  abschätzen:

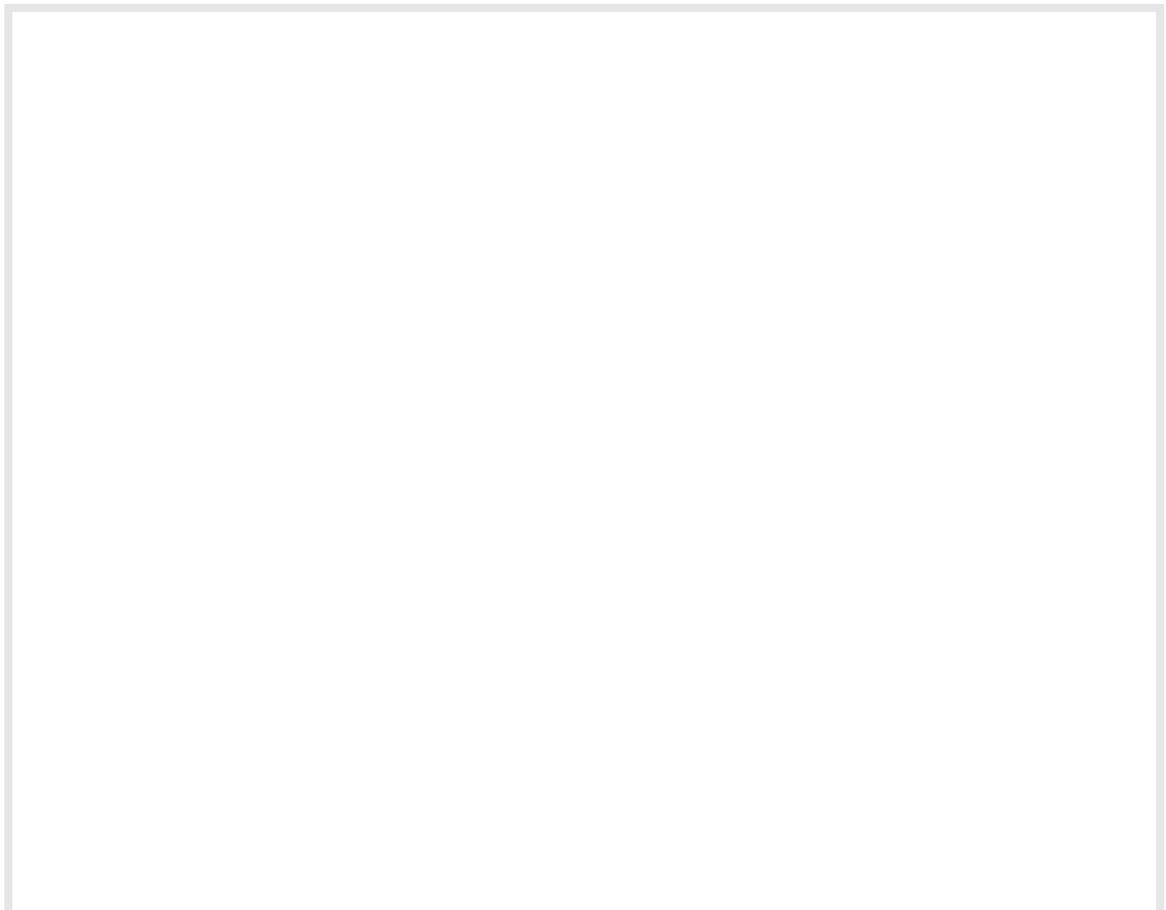
$$|f(x) - T_n(x; a)| = |R_{n+1}(x; a)| \leq \max_{\xi \text{ zwischen } x \text{ und } a} |f^{(n+1)}(\xi)| \frac{|x-a|^{n+1}}{(n+1)!}.$$

Somit kann man z.B. entscheiden, welchen Grad  $n$  des Taylorpolynoms man mindestens benötigt, um einen bestimmten Fehler nicht zu überschreiten. ◇

**Beispiel 4.48 (Fehlerabschätzung)**

Wie groß war der Fehler in Beispiel 4.25, als wir  $\sqrt{2}$  mit Hilfe des ersten Taylorpolynoms (Tangente) an der Entwicklungsstelle  $a = 1,96$  bei  $x = 2$  näherungsweise berechnet haben? Wir hatten dort

$$\begin{aligned} \sqrt{2} = f(x) &\approx T_1(x; a) = f(a) + f'(a) \cdot (x - a) \\ &= 1,4 + 0,3571 \cdot (2 - 1,96) = 1,4142857 \dots \end{aligned}$$



## § 4.7 Integralrechnung

Die Integration ist die Umkehrung der Differentiation. Man versucht dabei, eine Funktion aus den Werten ihrer Ableitung zu rekonstruieren.

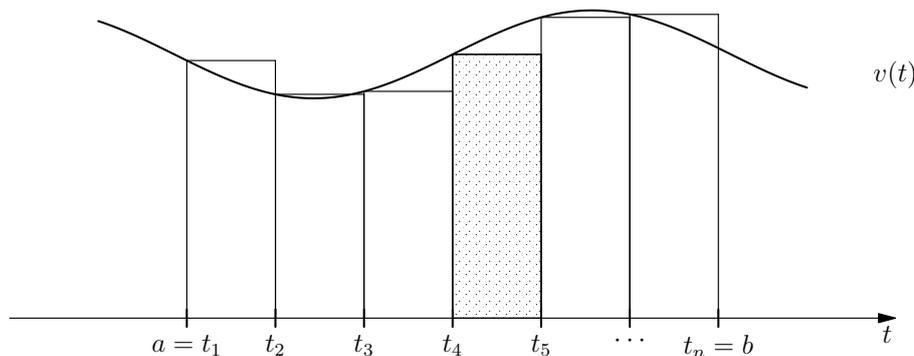
## § 4.7.1 Das bestimmte Integral

**Beispiel 4.49** (Fahrtenschreiber)

Ein Fahrtenschreiber in einem LKW zeichnet die Momentangeschwindigkeit  $v(t)$  des Fahrzeugs über die Zeit auf. Wie können wir daraus die zurückgelegte Strecke  $s(t)$  berechnen?

**Beachte:**  $s'(t) = v(t)$ . (Wir versuchen also, eine Funktion  $s(t)$  aus den Werten ihrer Ableitung  $v(t)$  zu rekonstruieren.)

Dazu betrachten wir den zeitlichen Verlauf (Graph) der Geschwindigkeitsfunktion  $v(t)$



und unterteilen das Zeitintervall  $[a, b]$ . Wir denken uns die Geschwindigkeit in jedem kleinen Teilintervall  $[t_i, t_{i+1}]$  konstant, z.B.  $v(t_i)$  (linker Randpunkt). Die in jedem Teilintervall zurückgelegte Strecke ist dann etwa gleich  $v(t_i)(t_{i+1} - t_i)$ . Die Summation aller Teilstrecken ergibt eine Näherung der gesuchten Gesamtstrecke:

$$\underbrace{s(b)}_{\text{Ort zum Zeitpunkt } t = b} - \underbrace{s(a)}_{\text{Ort zum Zeitpunkt } t = a} \approx \sum_{i=1}^{n-1} v(t_i)(t_{i+1} - t_i).$$

◇

**Definition 4.50** (Riemann-Summe und bestimmtes Integral)

Es sei  $f : [a, b] \rightarrow \mathbb{R}$  eine beschränkte und **stückweise stetige** Funktion (nur endlich viele Sprünge).

(a) Wir zerlegen das Intervall durch Einfügen von Teilpunkten

$$a = x_1 < x_2 < \dots < x_n = b, \quad n \geq 1$$

und wählen in jedem Teilintervall einen Zwischenpunkt  $x_i \leq \xi_i \leq x_{i+1}$  für alle  $i = 1, \dots, n - 1$ . Dann heißt

$$\sum_{i=1}^{n-1} f(\xi_i)(x_{i+1} - x_i)$$

eine **Riemann-Summe** von  $f$  im Intervall  $[a, b]$ .

- (b) Für feiner werdende Zerlegungen konvergieren die Riemann-Summen gegen eine Zahl, die unabhängig von der Zerlegung und der Wahl der Zwischenpunkte  $\xi$  ist. Diese Zahl nennen wir das **bestimmte Integral** von  $f$  über  $[a, b]$  und schreiben

$$\int_a^b f(x) dx \quad \text{oder} \quad \int_a^b f(t) dt \quad \text{etc.}$$

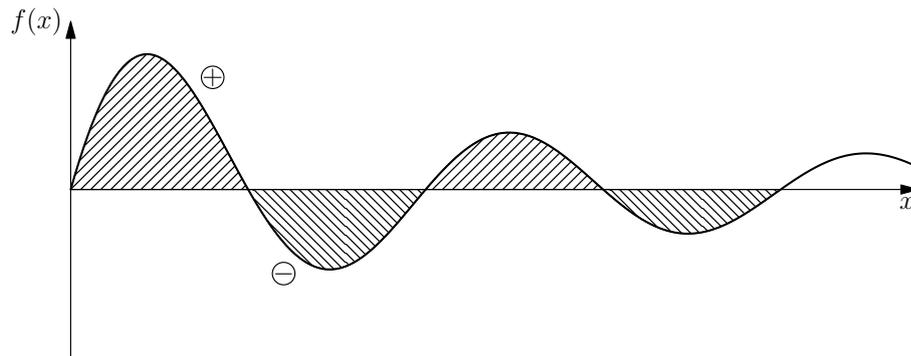
- (c) Wir definieren außerdem:

$$\int_a^a f(x) dx := 0, \quad \int_b^a f(x) dx := - \int_a^b f(x) dx. \quad \diamond$$

**Beachte:** Riemann-Summen dienen sowohl zur Definition des bestimmten Integrals als auch zu dessen numerischer Approximation.

**Bemerkung 4.51 (Geometrische Deutung des Integrals)**

Das Integral  $\int_a^b f(x) dx$  ist eine Maßzahl für den Flächeninhalt unter dem Graphen von  $f(x)$  im Intervall  $[a, b]$ , wobei Flächenstücke unterhalb der  $x$ -Achse negativ gezählt werden.



Will man auch die Flächenstücke unterhalb der  $x$ -Achse positiv zählen, so geht man einfach zu  $\int_a^b |f(x)| dx$  über.  $\diamond$

**Satz 4.52 (Rechenregeln für bestimmte Integrale)**

Es seien  $f, g : [a, b] \rightarrow \mathbb{R}$  beschränkt und stückweise stetig,  $b' \in [a, b]$  und  $c \in \mathbb{R}$ .

(a)  $\int_a^b f(x) dx + \int_a^b g(x) dx = \int_a^b [f(x) + g(x)] dx$

(b)  $\int_a^b c f(x) dx = c \int_a^b f(x) dx$

(c)  $\int_a^{b'} f(x) dx + \int_{b'}^b f(x) dx = \int_a^b f(x) dx$

(d)  $f(x) \leq g(x)$  für alle  $x \in [a, b] \Rightarrow \int_a^b f(x) dx \leq \int_a^b g(x) dx$

(e)  $\left| \int_a^b f(x) dx \right| \leq \int_a^b |f(x)| dx$

- (f)  $m \leq f(x) \leq M$  für alle  $x \in [a, b]$   
 $\Rightarrow m(b-a) \leq \int_a^b f(x) dx \leq M(b-a)$   $\diamond$

**Frage:** Integrale über Riemann-Summen auszurechnen oder auch nur zu approximieren ist (zumindest von Hand) mühsam. Gibt es eine einfachere Möglichkeit?

### § 4.7.2 Stammfunktionen

#### Definition 4.53 (Stammfunktion)

Sei  $I$  ein offenes Intervall und  $F : I \rightarrow \mathbb{R}$  diffbar.  $F$  heißt eine **Stammfunktion** von  $f : I \rightarrow \mathbb{R}$ , falls  $F'(x) = f(x)$  für alle  $x \in I$  gilt.  $\diamond$

**Frage:** Gibt es Stammfunktionen? Wie können wir sie zur Berechnung eines bestimmten Integrals  $\int_a^b f(x) dx$  verwenden?

#### Satz 4.54 (Hauptsatz der Differential- und Integralrechnung)

Sei  $f : [a, b] \rightarrow \mathbb{R}$  stetig.

- (a) Die durch

$$F_a(x) := \int_a^x f(t) dt, \quad x \in [a, b]$$

definierte Funktion (jeder Funktionswert ist ein bestimmtes Integral) *ist eine* Stammfunktion von  $f$  auf  $(a, b)$ . Das heißt,  $F_a$  ist differenzierbar in  $(a, b)$ , und es gilt

$$F'_a(x) = \frac{d}{dx} \left( \int_a^x f(t) dt \right) = f(x), \quad x \in (a, b).$$

- (b) *Jede andere* Stammfunktion von  $f$  hat die Form  $F(x) = F_a(x) + c$  für ein  $c \in \mathbb{R}$ . Die Menge aller Stammfunktionen heißt **unbestimmtes Integral** von  $f$ . Wir schreiben

$$\int f(x) dx = F(x) + c,$$

wenn  $F$  irgendeine Stammfunktion von  $f$  ist.

- (c) Mit einer beliebigen Stammfunktion  $F$  von  $f$  gilt

$$\int_a^b f(x) dx = F(x)|_a^b := F(b) - F(a). \quad \diamond$$

#### Bemerkung 4.55 (Berechnung bestimmter Integrale)

Zur Berechnung des bestimmten Integrals

$$\int_a^b f(x) dx$$

können wir also wie folgt vorgehen:

1. eine Stammfunktion  $F$  von  $f$  bestimmen (nachschaugen oder mit Hilfe eines Computeralgebraprogrammes),
2. die Probe machen ( $F'(x) = f(x)$  prüfen),
3. das gesuchte Integral als  $F(b) - F(a)$  berechnen.

**Beachte:** Die Differentiationstabelle aus Beispiel 4.26 dient auch zur Bestimmung von Stammfunktionen!  $\diamond$

**Beispiel 4.56** (a) Eine gleichmäßig beschleunigte Masse bewege sich mit der Geschwindigkeit  $v(t) = v_0 + gt$ . Welchen Weg legt sie im Zeitintervall  $[0, 2]$  zurück?

(b) Bei einer chemischen Reaktion wird ein Stoff (Produkt) mit der zeitabhängigen Rate

$$r(t) = k \exp(-at)$$

produziert ( $a, k > 0$ ). Wie hängt die Stoffmenge  $S(t)$  von der Zeit  $t$  ab?

**Beachte:** Die Einheit der Stammfunktion einer Funktion ist gleich der Einheit der Funktionswerte multipliziert mit der Einheit der unabhängigen Variablen, vergleiche Beispiel 4.23.

Im Beispiel (a) oben hatte z.B.  $v(t)$  die Einheit m/s, die unabhängige Variable  $t$  die Einheit s, also die Stammfunktion  $s(t)$  die Einheit m. Analog hatte  $r(t)$  die Einheit mol/s, die unabhängige Variable  $t$  die Einheit s, also die Stammfunktion  $S(t)$  die Einheit mol.

Ende 3. V

19.04.2011

### § 4.7.3 Berechnung von Stammfunktionen

Aus den Differentiationsregeln (Satz 4.27 und 4.28) erhalten wir Regeln zur Berechnung von Stammfunktionen:

#### **Satz 4.57 (Rechenregeln für Stammfunktionen)**

Seien  $f$  und  $g$  stetige bzw. stetig diffbare Funktionen,  $c \in \mathbb{R}$  und  $F$  eine Stammfunktion von  $f$ .

$$(a) \int [f(x) \pm g(x)] dx = \int f(x) dx \pm \int g(x) dx$$

$$(b) \int c f(x) dx = c \int f(x) dx$$

$$(c) \int f'(x) g(x) dx = f(x) g(x) - \int f(x) g'(x) dx \quad (\text{partielle Integration})$$

$$(d) \int f(g(x)) g'(x) dx = F(g(x)) + c \quad (\text{Substitutionsregel})$$

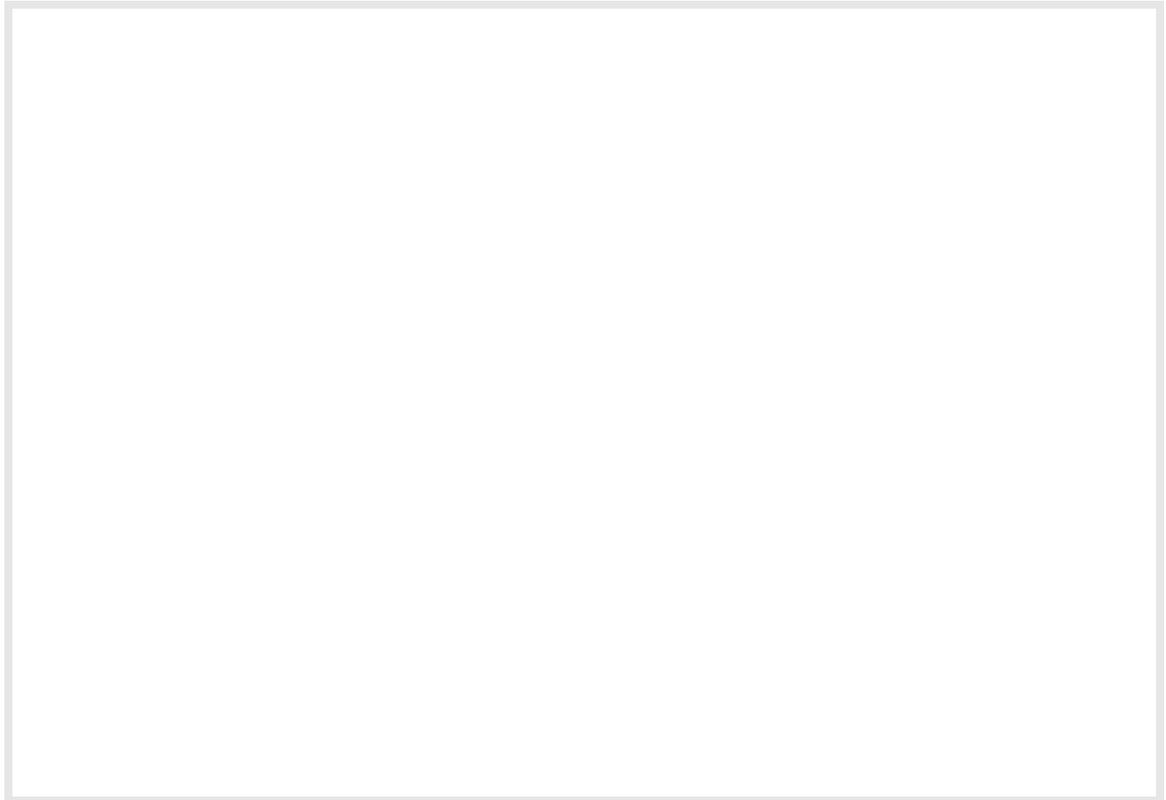
$$(e) \int \frac{g'(x)}{g(x)} dx = \ln |g(x)| + c \quad (\text{logarithmische Integration})$$

$$(f) \int f(ax + b) dx = \frac{1}{a} F(ax + b) + c \quad (\text{lineare Substitution}) \quad \diamond$$

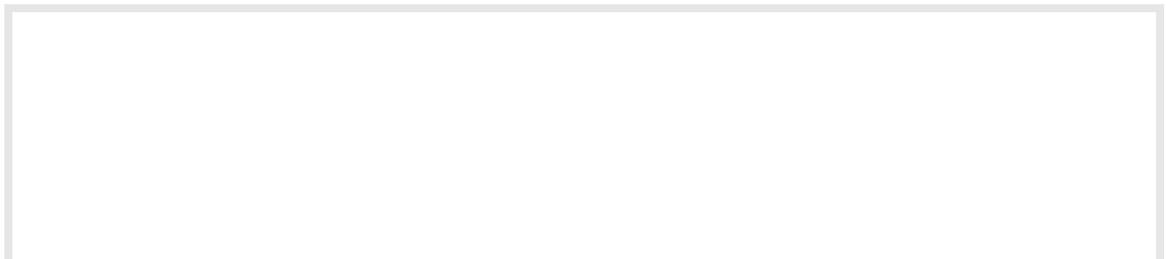
**Beachte:** (e) ist ein Spezialfall von (d) mit  $f(z) = \frac{1}{z}$ , und (f) ist ein Spezialfall von (d) mit  $g(x) = ax + b$ .

**Beispiel 4.58** (Bestimmung von Stammfunktionen)

- (a) Partielle Integration ist gut anwendbar, wenn sich ein Faktor  $[g(x)]$  durch Differentiation vereinfacht, der andere Faktor  $[f'(x)]$  sich durch Integration aber nicht wesentlich verkompliziert.

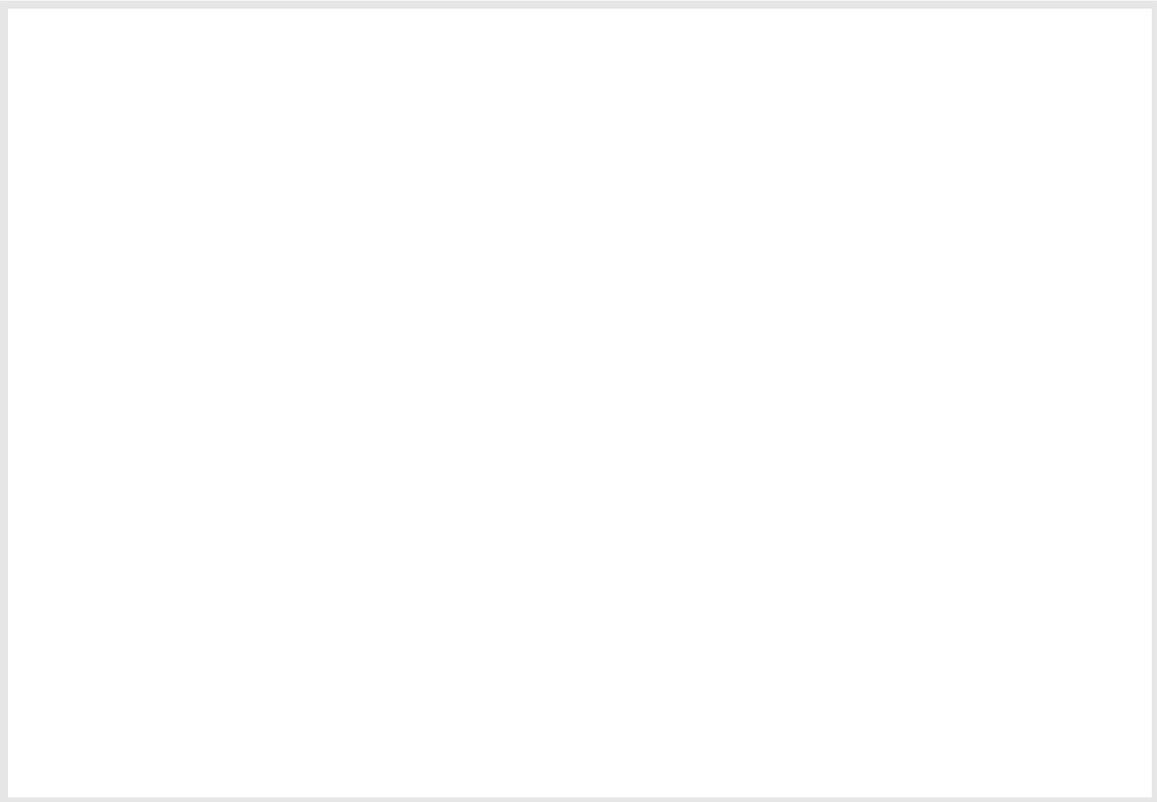


- (b)





(c)



#### § 4.7.4 Uneigentliche Integrale

**Frage:** Kann man eine Funktion „bis  $\infty$ “ integrieren?

**Definition 4.59** (Uneigentliches Integral)

Für geeignete Funktionen  $f$  definiert man folgende **uneigentliche Integrale**:

(a)

$$\int_a^{\infty} f(x) \, dx := \lim_{b \rightarrow \infty} \int_a^b f(x) \, dx$$

(b)

$$\int_{-\infty}^b f(x) dx := \lim_{a \rightarrow -\infty} \int_a^b f(x) dx$$

(c) Falls  $\lim_{x \nearrow b} f(x) = \pm\infty$  ist:

$$\int_a^b f(x) dx := \lim_{c \nearrow b} \int_a^c f(x) dx$$

(d) Falls  $\lim_{x \searrow a} f(x) = \pm\infty$  ist:

$$\int_a^b f(x) dx := \lim_{c \searrow a} \int_c^b f(x) dx$$

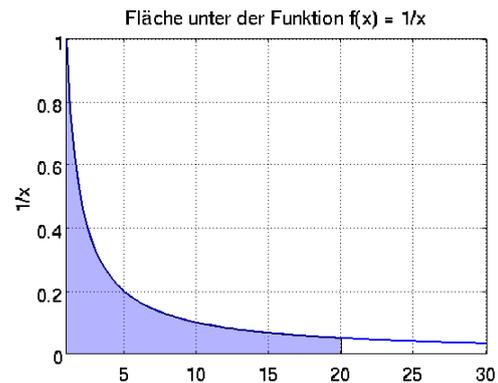
Je nachdem, ob der Grenzwert existiert, heißen die Integrale **konvergent** oder (bestimmt bzw. unbestimmt) **divergent**.  $\diamond$

### Beispiel 4.60 (Uneigentliche Integrale)

(a)

$$\begin{aligned} \int_1^{\infty} \frac{1}{x} dx &= \lim_{b \rightarrow \infty} \int_1^b \frac{1}{x} dx \\ &= \lim_{b \rightarrow \infty} \ln x \Big|_1^b = \lim_{b \rightarrow \infty} \ln b = \infty \end{aligned}$$

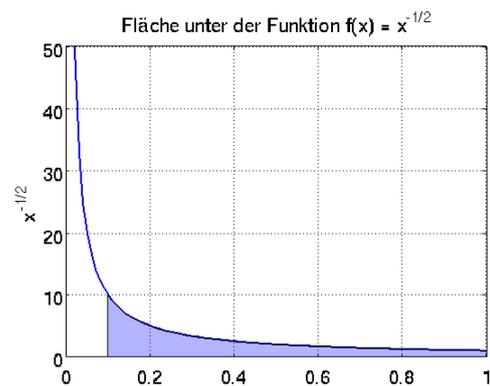
Das Integral ist divergent.



(b)

$$\begin{aligned} \int_0^1 \frac{1}{\sqrt{x}} dx &= \lim_{a \searrow 0} \int_a^1 \frac{1}{\sqrt{x}} dx \\ &= \lim_{a \searrow 0} 2\sqrt{x} \Big|_a^1 = \lim_{a \searrow 0} 2 - 2\sqrt{a} = 2. \end{aligned}$$

Das Integral ist konvergent.



## § 4.8 Vektorwertige Funktionen $\mathbb{R} \rightarrow \mathbb{R}^n$

Wir erinnern uns an Definition 4.1: Eine Vorschrift  $f$ , die jedem Element  $x$  einer Menge  $X$  genau ein Element  $y$  einer Menge  $Y$  zuordnet, heißt **Funktion** mit **Definitionsmenge**  $D(f) = X$  und **Zielfmenge**  $Y$ . Bisher war stets  $D(f) \subseteq \mathbb{R}$  und  $Y = \mathbb{R}$  (reelle Funktion einer reellen Variablen).

Nun ist weiterhin  $D(\vec{f}) \subseteq \mathbb{R}$ , aber  $Y = \mathbb{R}^n$ . Jeder Zahl  $x \in D(\vec{f})$  wird ein *Vektor*

$$\vec{f}(x) = \begin{pmatrix} f_1(x) \\ f_2(x) \\ \vdots \\ f_n(x) \end{pmatrix} \in \mathbb{R}^n, \quad n \geq 1$$

zugeordnet.  $\vec{f}$  heißt eine **vektorwertige Funktion**. Die einzelnen  $f_i(x)$  heißen **Komponentenfunktionen**.

**Beispiel 4.61 (Vektorwertige Funktionen)**

(a) Bezeichnet  $x$  die Geschwindigkeit eines Autos, so sind z.B.

$$\begin{aligned} f_1(x) &= \text{Benzinverbrauch } (x), \\ f_2(x) &= \text{Luftwiderstand } (x), \\ f_3(x) &= \text{Rollwiderstand } (x) \end{aligned}$$

Funktionen von  $x$ , die wir mittels

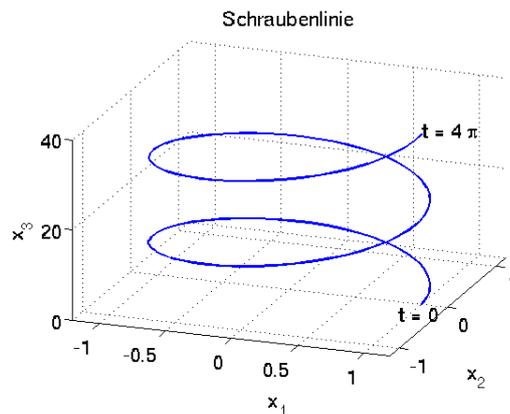
$$\vec{f}(x) = \begin{pmatrix} f_1(x) \\ f_2(x) \\ f_3(x) \end{pmatrix}$$

zu einer vektorwertigen Funktion zusammenfassen können.

(b) Mit  $\vec{s}(t) \in \mathbb{R}^3$  können wir die Bahnkurve eines Punktes im Raum zum Zeitpunkt  $t \in \mathbb{R}$  beschreiben, z.B. durch die **Schraubenlinie**

$$\vec{s}(t) = \begin{pmatrix} \cos(t) \\ \sin(t) \\ 3t \end{pmatrix}.$$

Die Komponentenfunktionen  $s_1(t)$ ,  $s_2(t)$  und  $s_3(t)$  beschreiben jeweils die  $x_1$ -,  $x_2$ - und  $x_3$ -Koordinaten.



◇

Die Begriffe **Grenzwert** (Definition 4.13), **Stetigkeit** und **Differenzierbarkeit** (Definition 4.22) übertragen sich auf Funktionen  $\vec{f} : \mathbb{R} \rightarrow \mathbb{R}^n$ , indem wir sie auf alle Komponentenfunktionen anwenden:

**Definition 4.62 (Stetigkeit und Diffbarkeit)**

Gegeben sei ein offenes Intervall  $I \subseteq \mathbb{R}$  und ein Punkt  $a \in I$  sowie  $\vec{f} : I \rightarrow \mathbb{R}^n$ .

(a) Die Funktion  $\vec{f}$  heißt **stetig** (Definition 4.18) im Punkt  $a$ , wenn alle Komponentenfunktionen  $f_i : I \rightarrow \mathbb{R}$ ,  $i = 1, \dots, n$ , im Punkt  $a$  stetig sind.

- (b) Die Funktion  $\vec{f}$  heißt **diffbar** im Punkt  $a$ , wenn alle Komponentenfunktionen  $f_i : I \rightarrow \mathbb{R}$ ,  $i = 1, \dots, n$ , im Punkt  $a$  diffbar sind. Die **Ableitung** ist dann

$$\vec{f}'(a) = \frac{d}{dx} \vec{f}(a) = \begin{pmatrix} f_1'(x) \\ f_2'(x) \\ \vdots \\ f_n'(x) \end{pmatrix} \in \mathbb{R}^n.$$

◇

Analog zu Definition 4.22 (e) können wir die **Tangente** an die Funktion  $f : \mathbb{R} \rightarrow \mathbb{R}^n$  im Punkt  $a \in \mathbb{R}$  definieren:

$$\vec{T}_1(x; a) = \underbrace{\vec{f}(a)}_{\in \mathbb{R}^n} + \underbrace{\vec{f}'(a)}_{\in \mathbb{R}^n} \underbrace{(x - a)}_{\in \mathbb{R}}.$$

Sie ist eine Gerade im  $\mathbb{R}^n$ .

### Beispiel 4.63 (Ableitungen vektorwertiger Funktionen)

(a)

$$\vec{f}(x) = \begin{pmatrix} x^2 \\ -x^4 \end{pmatrix} \Rightarrow \vec{f}'(x) = \begin{pmatrix} 2x \\ -4x^3 \end{pmatrix}.$$

- (b) Die Funktion  $\vec{s}(t)$  aus Beispiel 4.61 (b) ist diffbar, und die Ableitung ist

$$\vec{v}(t) = \vec{s}'(t) = \begin{pmatrix} -\sin(t) \\ \cos(t) \\ 3 \end{pmatrix}.$$

Die Ableitung an einer Stelle  $a$  ist der **Tangentialvektor (Geschwindigkeitsvektor)** an die Bahnkurve  $\vec{s}(t)$  an dieser Stelle.

Die Gleichung der Tangente an die Funktion  $\vec{s}(t)$  im Zeitpunkt Punkt  $a$  ist

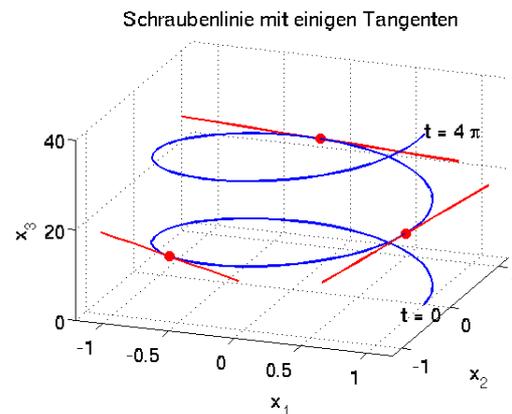
$$\vec{s}(a) + \vec{s}'(a)(t - a).$$

Der Tangentialvektor  $\vec{s}'(a)$  ist der Richtungsvektor der Tangente im Punkt  $a$ . ◇

### Satz 4.64 (Rechenregeln für die Ableitung, vgl. Satz 4.27)

Es seien  $\vec{f} : \mathbb{R} \rightarrow \mathbb{R}^n$ ,  $\vec{g} : \mathbb{R} \rightarrow \mathbb{R}^n$  und  $\alpha : \mathbb{R} \rightarrow \mathbb{R}$  diffbar im Punkt  $a \in \mathbb{R}$ . Dann existieren auch folgende Ableitungen:

- (a)  $(\vec{f} \pm \vec{g})'(a) = \vec{f}'(a) \pm \vec{g}'(a)$   
 (b)  $(c\vec{f})'(a) = c\vec{f}'(a)$  für alle  $c \in \mathbb{R}$



$$(c) (\vec{f} \cdot \vec{g})'(a) = \vec{f}'(a) \cdot \vec{g}(a) + \vec{f}(a) \cdot \vec{g}'(a)$$

$$\text{bzw. } (\vec{f}^\top \vec{g})'(a) = \vec{f}'(a)^\top \vec{g}(a) + \vec{f}(a)^\top \cdot \vec{g}'(a)$$

$$(d) (\alpha \cdot \vec{f})'(a) = \alpha'(a) \cdot \vec{f}(a) + \alpha(a) \cdot \vec{f}'(a)$$

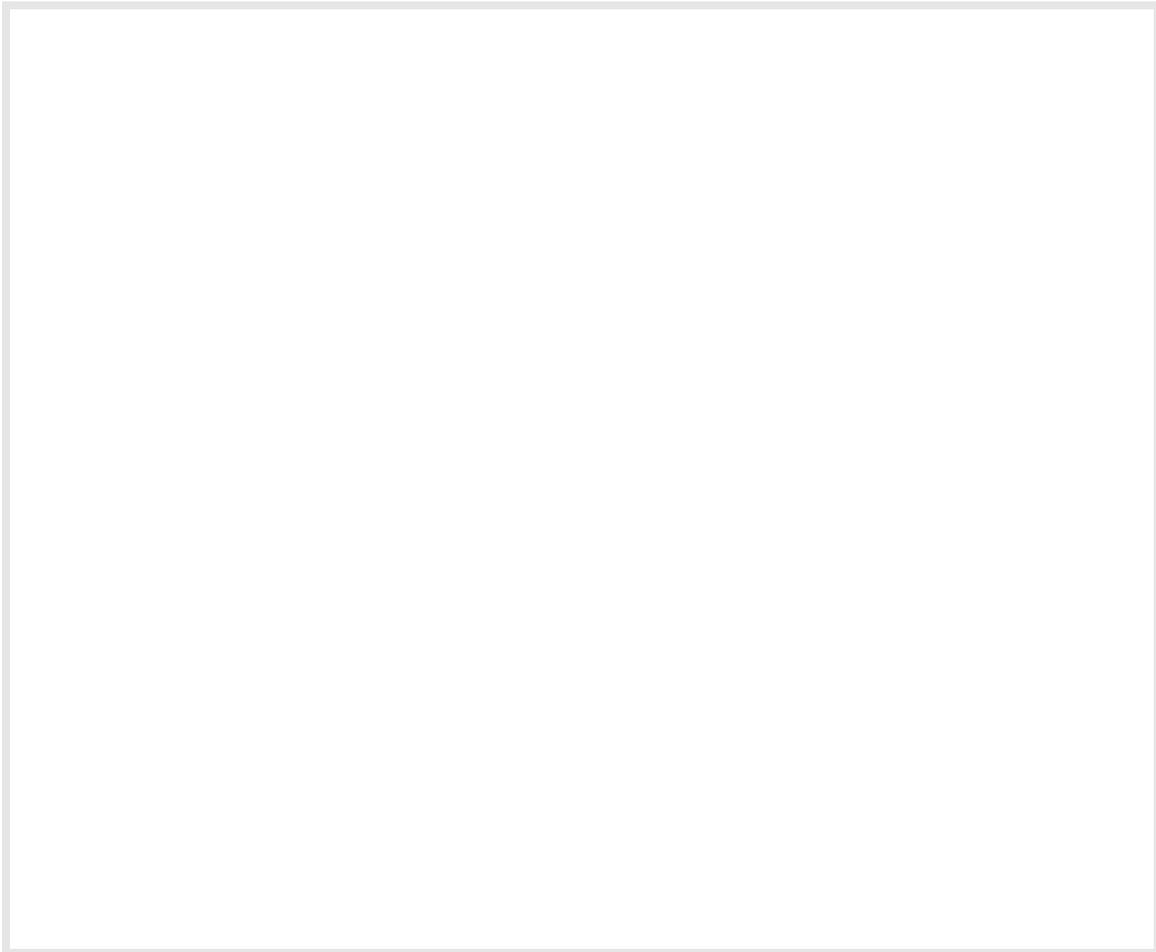
$$(e) (\vec{f} \times \vec{g})'(a) = \vec{f}'(a) \times \vec{g}(a) + \vec{f}(a) \times \vec{g}'(a) \quad (\text{für } n = 3)$$

◇

Ende 4. V

**Beispiel 4.65** (Anwendungen dieser Rechenregeln)

26.04.2011





## Höhere Mathematik I.2

### § 5 Differentialgleichungen und Differentialgleichungssysteme

Differentialgleichungen (Dgl) beschreiben einen Zusammenhang zwischen einer unbekannt (gesuchten) Funktion  $y(x)$  und ihrer Ableitung  $y'(x)$ , evtl. auch höheren Ableitungen  $y''(x)$  usw. Viele Naturgesetze und Beobachtungen in der Physik, Chemie, Biologie, Mechanik etc. werden durch Dgl ausgedrückt.

**Beachte:** Die Lösung einer Dgl ist keine Zahl oder ein Vektor, sondern eine Funktion.

#### Beispiel 5.1 (Exponentielles und logistisches Wachstum)

In diesem Beispiel bezeichnet  $x$  die unabhängige Variable (Zeit) und  $y(x)$  die zur Zeit  $x$  vorhandene Menge einer Bakterienpopulation.

- (a) In einem einfachen Modell wird angenommen, dass die Wachstumsrate  $y'(x)$  proportional zur bereits vorhandenen Bakterienanzahl (aufgrund von Zellteilung) ist:

$$y'(x) = a y(x) \quad \text{mit einem Proportionalitätsfaktor } a > 0.$$

Man rechnet leicht nach, dass die Funktion

$$y(x) = c \exp(ax)$$

für jedes  $c \in \mathbb{R}$  eine Lösung dieser Dgl ist, denn:

$$y'(x) = c a \exp(ax) = a y(x).$$

Allerdings ist  $y(x)$  für  $c \neq 0$  unbeschränkt für  $x \rightarrow \infty$ , sodass das Dgl-Modell nur bedingt sinnvoll ist. Wir sprechen von der Dgl des **exponentiellen Wachstums**.

Die Dgl können wir auch so lesen: Die Steigung  $y'(x)$  der (Tangente der) Lösung am Punkt  $(x, y(x)) \in \mathbb{R}^2$  ist gleich  $a y(x)$ . Damit können wir in der  $(x, y)$ -Ebene ein Richtungsfeld zeichnen, hier  $(1, a y)$ , dem die Lösungen folgen.

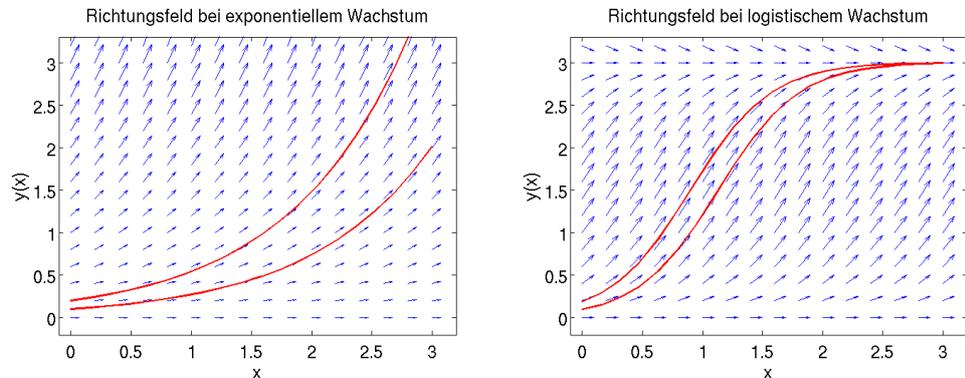
- (b) Wir verfeinern das Modell und nehmen an, dass die Wachstumsrate  $y'(x)$  gleichzeitig proportional zu  $y(x)$  und zu  $b - y(x)$  ist:

$$y'(x) = a y(x) (b - y(x))$$

Damit sollte sich eine Sättigung bei  $y(x) = b$  ergeben. Man rechnet wieder leicht nach, dass die Funktion

$$y(x) = \frac{b}{1 + c e^{-abx}}$$

für jedes  $c \in \mathbb{R}$  eine Lösung dieser Dgl ist. Da dies eine logistische Funktion ist, vgl. Beispiel 4.3<sup>1</sup>, sprechen wir von der Dgl des **logistischen Wachstums**.



◇

### § 5.1 Die trennbare Differentialgleichung $y'(x) = f(x)g(y)$

#### Definition 5.2 (Trennbare Differentialgleichung)

(a) Eine Gleichung der Form

$$y'(x) = g(x)h(y(x)) \quad \text{oder kurz} \quad y' = g(x)h(y)$$

heißt **Dgl 1. Ordnung mit trennbaren Veränderlichen (trennbare Dgl)**. Dabei sind  $g(x)$  und  $h(y)$  als stetige Funktionen vorausgesetzt.

(b) Eine Funktion  $y(x)$  heißt **Lösung** der Dgl auf einem Intervall  $I$ , wenn  $y : I \rightarrow \mathbb{R}$  diffbar ist und die Dgl für alle  $x \in I$  erfüllt. ◇

Um eine trennbare Dgl zu lösen, gehen wir formal wie folgt vor: Wir schreiben (vergleiche Definition 4.22)

$$\begin{aligned} y' &= \frac{dy}{dx} = g(x)h(y) \\ \Rightarrow \frac{dy}{h(y)} &= g(x) dx \end{aligned}$$

Jetzt wird für beide Seiten eine Stammfunktion gesucht (unbestimmt integriert), siehe § 4.7.2:

$$\int \frac{dy}{h(y)} = \int g(x) dx + c$$

Wenn möglich, wird diese Gleichung nach  $y$  aufgelöst.

#### Beispiel 5.3 (Trennbare Differentialgleichung) (a) Wir lösen die Dgl

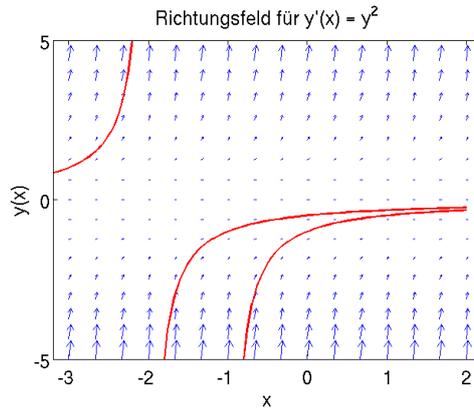
$$\frac{dy}{dx} = y' = 1 \cdot y^2 =: g(x)h(y),$$

die mit  $g(x) = 1$  und  $h(y) = y^2$  von trennbarer Form ist. Offenbar ist  $y(x) \equiv 0$  eine Lösung, deswegen können wir ab sofort von  $y < 0$  oder  $y > 0$

<sup>1</sup>Die Konstanten sind dort jedoch anders benannt.

ausgehen. Wir schreiben um:

$$\begin{aligned} & \int \frac{dy}{y^2} = \int 1 \, dx \\ \Rightarrow & -\frac{1}{y} = x + c \\ \Rightarrow & y = \frac{-1}{x + c}, \quad c \in \mathbb{R}. \end{aligned}$$



Diese Funktionen sind für jedes  $c \in \mathbb{R}$  Lösung jeweils auf den Intervallen  $(-\infty, -c)$  und  $(-c, \infty)$ . Eine weitere Lösung ist  $y \equiv 0$ .

- (b) Die logistische Differentialgleichung aus Beispiel 5.1 ist ebenfalls von trennbarer Form:

$$y' = a y (b - y) = g(x) h(y)$$

mit

$$g(x) = a, \quad h(y) = y (b - y).$$

Wir schreiben (mittels Partialbruchzerlegung) um:

$$\int \frac{dy}{y(b-y)} = \int \frac{1}{by} dy + \int \frac{1}{b(b-y)} dy = \int a \, dx + C$$

und berechnen die Stammfunktionen:

$$\frac{1}{b} \ln|y| - \frac{1}{b} \ln|b-y| = ax + C.$$

Nun lösen wir nach  $y$  auf:

$$\begin{aligned} & \ln|y| - \ln|b-y| = b(ax + C) \\ \Rightarrow & \ln \left| \frac{y}{b-y} \right| = b(ax + C) \\ \Rightarrow & \left| \frac{y}{b-y} \right| = \exp(b(ax + C)) \\ \Rightarrow & \left| \frac{b-y}{y} \right| = \left| \frac{b}{y} - 1 \right| = \exp(-b(ax + C)) \\ \Rightarrow & \frac{b}{y} = \begin{cases} \exp(-b(ax + C)) + 1 = \exp(-abx) \exp(-bC) + 1 \\ -\exp(-b(ax + C)) + 1 = -\exp(-abx) \exp(-bC) + 1. \end{cases} \end{aligned}$$

Da  $C \in \mathbb{R}$  beliebig war, ist  $\exp(-bC)$  eine beliebige *positive* reelle Zahl. Wenn wir eine neue Konstante  $c := \pm \exp(-bC)$  einführen, können wir beide Fälle wieder zusammenführen:

$$\begin{aligned} \Rightarrow & \frac{b}{y} = c \exp(-abx) + 1 \\ \Rightarrow & y = \frac{b}{c \exp(-abx) + 1} \quad \text{mit} \quad c \in \mathbb{R}. \end{aligned}$$

Die Lösung ist für jedes  $c \in \mathbb{R}$  auf ganz  $I = \mathbb{R}$  definiert.

**Beachte:** Auch für  $c = 0$  ergibt sich mit  $y(x) = b$  eine Lösung der Dgl.  $\diamond$

## § 5.2 Die lineare Differentialgleichung $y'(x) = ay(x) + f(x)$

### Definition 5.4 (Lineare Differentialgleichung)

(a) Eine Gleichung der Form

$$y'(x) = ay(x) + f(x)$$

heißt **lineare Dgl 1. Ordnung** mit konstantem Koeffizienten  $a \in \mathbb{R}$ . (Es besteht ein linearer Zusammenhang zwischen  $y(x)$  und der Ableitung, und es tritt nur die erste Ableitung auf.)

(b) Die Dgl heißt **homogen**, wenn die **rechte Seite**  $f(x) \equiv 0$  ist, ansonsten **inhomogen**.

(c) Eine Funktion  $y(x)$  heißt **Lösung** der Dgl auf einem Intervall  $I$ , wenn  $y : I \rightarrow \mathbb{R}$  diffbar ist und die Dgl auf  $I$  erfüllt.  $\diamond$

Die rechte Seite  $f(x)$  entspricht je nach Aufgabenstellung einer äußeren Anregung, z.B. einer äußeren Kraft in der Mechanik, einem Zu- bzw. Abfluss bei chemischen Reaktionen etc.

### § 5.2.1 Die homogene Dgl

Wir untersuchen zuerst die homogene Dgl:

#### Satz 5.5 (Lösungen der homogenen linearen Dgl)

Die Lösungen der homogenen linearen Dgl (auf beliebigen Intervallen  $I \subseteq \mathbb{R}$ )

$$y'(x) = ay(x), \quad a \in \mathbb{R}$$

sind genau die Funktionen

$$y(x) = ce^{ax}$$

mit einer beliebigen Konstanten  $c \in \mathbb{R}$ . Man nennt die Gesamtheit der Lösungen die **allgemeine Lösung** der homogenen Dgl.  $\diamond$

**Beachte:**

$$\left\{ \begin{array}{ll} a > 0 & \Rightarrow e^{ax} \text{ nimmt mit } x \text{ zu} \\ a = 0 & \Rightarrow e^{ax} \text{ ist konstant in } x \\ a < 0 & \Rightarrow e^{ax} \text{ nimmt mit } x \text{ ab.} \end{array} \right.$$

### § 5.2.2 Die inhomogene Dgl

Nehmen wir an, wir kennen irgendeine spezielle (partikuläre) Lösung  $y_p(x)$  der inhomogenen Dgl

$$y'(x) = ay(x) + f(x), \quad a \in \mathbb{R}.$$

Dann können wir jede Lösung schreiben als

$$y(x) = y_h(x) + y_p(x),$$

wobei  $y_h(x)$  irgendeine Lösung der homogenen Dgl ist, denn:

$$y'(x) = y'_h(x) + y'_p(x) = ay_h(x) + ay_p(x) + f(x) = ay(x) + f(x).$$

**Beachte:** Die Lösungsstruktur „partikuläre + allgemeine homogene“ Lösung ist genau wie bei inhomogenen LGS, vgl. Beispiel 2.40. Dies ist das **Prinzip der Superposition**.

**Frage:** Woher bekommen wir eine partikuläre Lösung?

**Ansatz:** Wir versuchen es mit  $y_p(x) = c(x) e^{ax}$ . Diese Technik heißt **Variation der Konstanten**. Wir erhalten die Ableitung

$$y_p'(x) = c'(x) e^{ax} + c(x) a e^{ax}$$

und setzen in die inhomogene Dgl ein:

$$\begin{aligned} c'(x) e^{ax} + c(x) a e^{ax} &\stackrel{!}{=} a c(x) e^{ax} + f(x) \\ \Rightarrow c'(x) e^{ax} &= f(x) \\ \Rightarrow c'(x) &= e^{-ax} f(x) \\ \Rightarrow c(x) &= \int_{x_0}^x e^{-at} f(t) dt \end{aligned}$$

mit irgendeinem  $x_0 \in \mathbb{R}$ . (Wir nehmen immer  $x_0 = 0$ .) Die Bestimmung einer partikulären Lösung

$$y_p(x) = e^{ax} \int_0^x e^{-at} f(t) dt$$

läuft also darauf hinaus, eine Stammfunktion für  $e^{-at} f(t)$  zu finden.

Ende 5. V

**Satz 5.6 (Lösungen der inhomogenen linearen Dgl)**

03.05.2011

Die Lösungen der inhomogenen linearen Dgl (auf beliebigen Intervallen  $I \subseteq \mathbb{R}$ )

$$y'(x) = a y(x) + f(x), \quad a \in \mathbb{R}$$

sind genau die Funktionen

$$\begin{aligned} y(x) &= c e^{ax} + e^{ax} \int_0^x e^{-at} f(t) dt \\ &= e^{ax} \left( c + \int_0^x e^{-at} f(t) dt \right) \end{aligned}$$

mit einer beliebigen Konstanten  $c \in \mathbb{R}$ . Man nennt die Gesamtheit der Lösungen die **allgemeine Lösung** der inhomogenen Dgl.  $\diamond$

**Frage:** Wie wählt man aus den vielen Lösungen die „richtige“ aus?

Oft ist neben der Dgl noch eine **Anfangsbedingung** (AB)  $y(x_0) = y_0$  gegeben. Wir sprechen dann von einem **Anfangswertproblem** (AWP). Wir nehmen immer  $x_0 = 0$  an. Die AB  $y(0) = y_0$  wählt aus der Lösungsschar genau eine Lösung aus:

$$y(0) = e^{a \cdot 0} \left( c + \int_0^0 e^{-at} f(t) dt \right) = 1 \cdot (c + 0) \stackrel{!}{=} y_0 \quad \Rightarrow \quad c = y_0.$$

Die eindeutige Lösung des AWP mit  $y(0) = y_0$  lautet also

$$y(x) = e^{ax} \left( y_0 + \int_0^x e^{-at} f(t) dt \right).$$

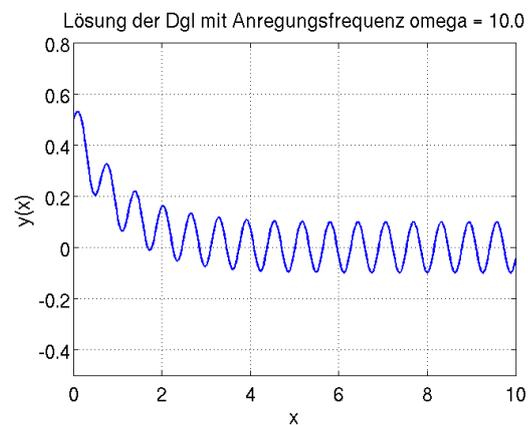
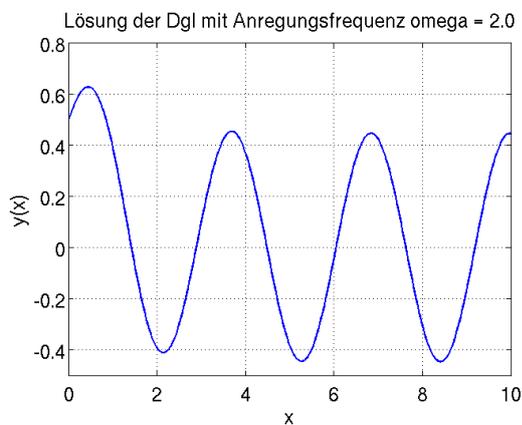
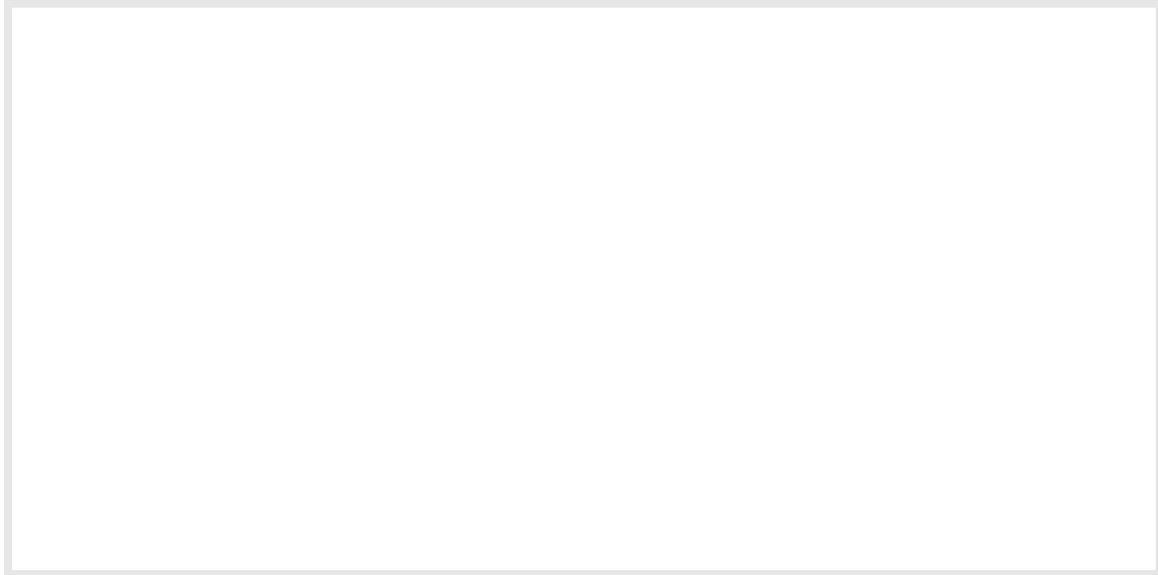
**Beachte:** Diese Formel muss man nicht auswendig können. Man kann sie sich immer durch Variation der Konstanten herleiten.



zur Bestimmung von  $c_1$  und  $c_2$ . Das System ist immer eindeutig lösbar, denn die Determinante der Matrix (vgl. Definition 2.46 und Satz 2.48) ist

$$(-a)^2 - \omega(-\omega) = a^2 + \omega^2 \neq 0.$$

**Beispiel 5.8** (Ansatz vom Typ der rechten Seite)



◇

### § 5.3 Eigenwerte und Eigenvektoren

Für die Lösung des Dgl-Systems

$$\vec{y}'(x) = A\vec{y}(x) + \vec{f}(x)$$

im nächsten Abschnitt stellen wir folgende

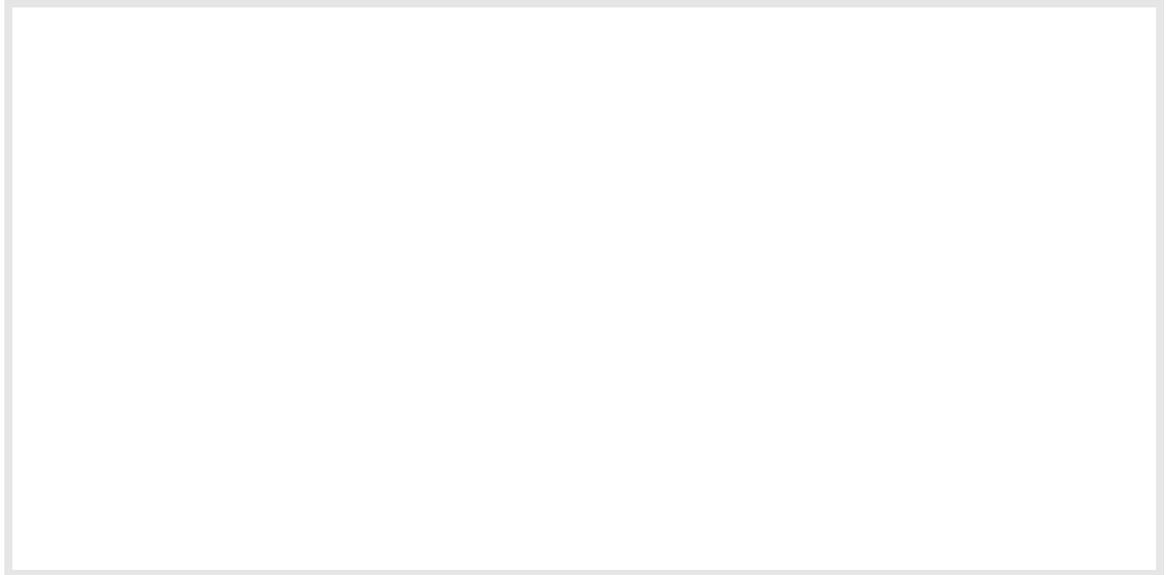
**Frage:** Gegeben sei eine quadratische Matrix  $A \in \mathbb{R}^{n \times n}$ . Gibt es Vektoren  $\vec{x} \in \mathbb{R}^n$ ,  $\vec{x} \neq \vec{0}$ , sodass  $A\vec{x}$  parallel (oder anti-parallel) zu  $\vec{x}$  liegt?

$$A\vec{x} = \lambda\vec{x} \Leftrightarrow A\vec{x} = \lambda E\vec{x} \Leftrightarrow (A - \lambda E)\vec{x} = \vec{0}.$$

Wir suchen also nach Zahlen  $\lambda$ , sodass das homogene LGS mit Koeffizientenmatrix  $A - \lambda E$  noch andere Lösungen außer  $\vec{x} = \vec{0}$  hat.

$$\begin{aligned} (A - \lambda E) \vec{x} = \vec{0} & \text{ hat noch andere Lösungen als } \vec{x} = \vec{0} \\ \Leftrightarrow A - \lambda E & \text{ ist nicht invertierbar (Satz 2.43)} \\ \Leftrightarrow \det(A - \lambda E) = 0 & \text{ (Satz 2.48)}. \end{aligned}$$

### Beispiel 5.9 (Charakteristisches Polynom)



### Definition 5.10 (Charakteristisches Polynom, Eigenwert, Eigenvektor)

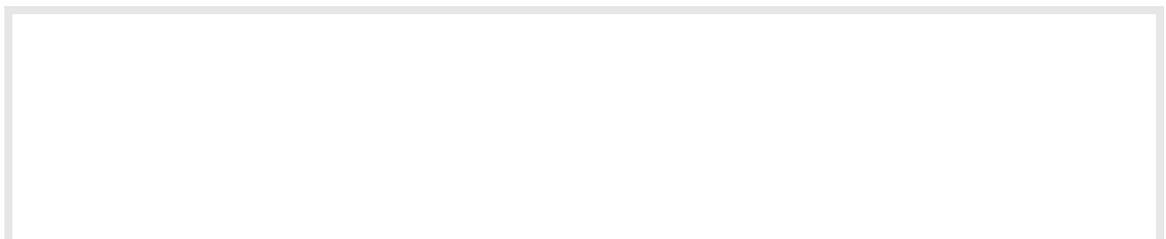
Sei  $A \in \mathbb{R}^{n \times n}$ .

- Der Ausdruck  $\det(A - \lambda E)$  ist ein Polynom in  $\lambda$  vom Grad  $n$ , genannt das **charakteristische Polynom** der Matrix  $A$ .
- Die Nullstellen  $\lambda_i \in \mathbb{C}$ ,  $i = 1, \dots, n$  (möglicherweise sind nicht alle verschieden) heißen die **Eigenwerte (EW)** der Matrix  $A$ .
- Ist  $\lambda \in \mathbb{C}$  ein Eigenwert, dann heißt jede Lösung  $\vec{x} \in \mathbb{C}^n$  des homogenen LGS

$$(A - \lambda E) \vec{x} = \vec{0} \quad \text{oder} \quad A \vec{x} = \lambda \vec{x}$$

(mit Ausnahme von  $\vec{x} = \vec{0}$ ) ein **Eigenvektor (EV)** zum Eigenwert  $\lambda$ . Man sagt auch:  $(\vec{x}, \lambda)$  bilden ein **Eigenpaar**.  $\diamond$

### Beispiel 5.11 (Eigenwerte und Eigenvektoren)



Ende 6. V

10.05.2011

**Satz 5.12 (Eigenwerte, Diagonalisierbarkeit)**

Die Matrix  $A \in \mathbb{R}^{n \times n}$  habe die EW  $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{C}$  (möglicherweise sind nicht alle verschieden).

- (a) Einige oder alle der EW können komplexe Zahlen sein. Sie treten immer in (komplex-konjugierten) Paaren  $a \pm bi$  auf. In dem Fall werden die EV auch komplexe Vektoren  $\vec{x} \in \mathbb{C}^n$  sein.
- (b) EV zu verschiedenen EW sind linear unabhängig.<sup>2</sup> Das heißt: Es seien  $\lambda_1, \lambda_2, \dots, \lambda_k$  paarweise verschiedene EW von  $A$  und  $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_k$  zugehörige EV. Dann gilt (vgl. Definition 2.6):

$$\alpha_1 \vec{v}_1 + \alpha_2 \vec{v}_2 + \dots + \alpha_k \vec{v}_k = \vec{0} \quad \text{mit irgendwelchen } \alpha_1, \dots, \alpha_k \in \mathbb{C}$$

$$\Rightarrow \alpha_1 = \alpha_2 = \dots = \alpha_k = 0.$$

- (c) Falls es  $n$  linear unabhängige EV  $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$  von  $A$  zu den EW  $\lambda_1, \lambda_2, \dots, \lambda_n$  gibt, die man spaltenweise in eine Matrix  $V$  einträgt, dann gilt:

$$V^{-1}AV = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \dots & \\ & & & \lambda_n \end{pmatrix} = D \quad \text{Diagonalmatrix}$$

oder auch  $AV = VD$ , also

$$A \begin{pmatrix} \vdots & \vdots & \dots & \vdots \\ \vec{v}_1 & \vec{v}_2 & \dots & \vec{v}_n \\ \vdots & \vdots & & \vdots \end{pmatrix} = \begin{pmatrix} \vdots & \vdots & \dots & \vdots \\ \vec{v}_1 & \vec{v}_2 & \dots & \vec{v}_n \\ \vdots & \vdots & & \vdots \end{pmatrix} \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \dots & \\ & & & \lambda_n \end{pmatrix}.$$

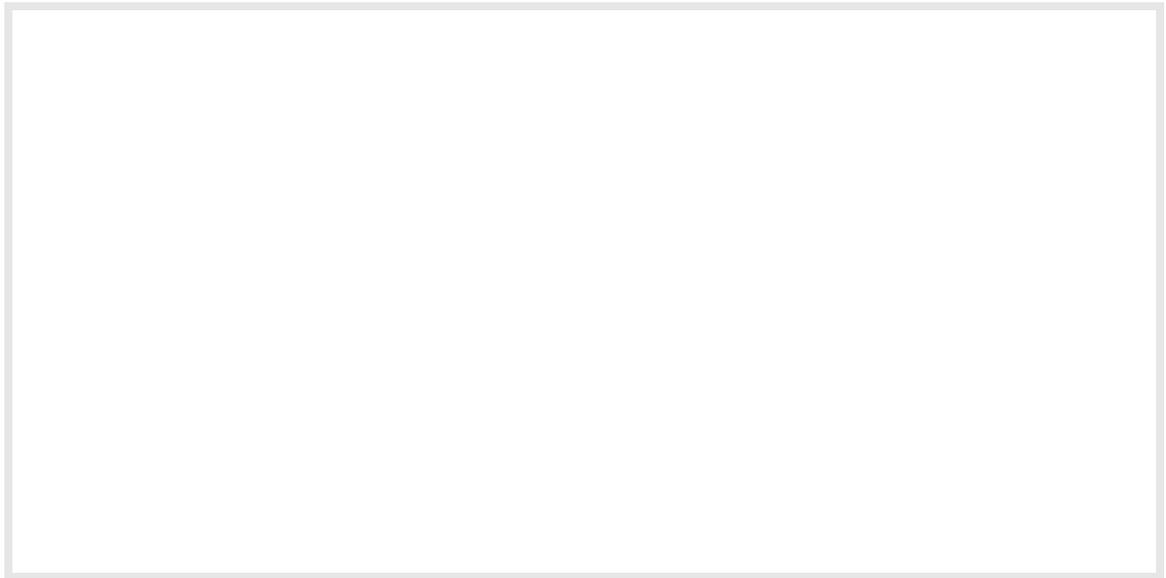
Man sagt dann:  $A$  ist **diagonalisierbar**.  $\diamond$

**Bemerkung 5.13**

Eine Matrix ist nur dann *nicht* diagonalisierbar, wenn ein EW mehrfach (als Nullstelle des charakteristischen Polynoms) vorkommt und es nicht genügend viele linear unabhängige EV dazu gibt.  $\diamond$

**Beispiel 5.14 (Diagonalisierung)**

<sup>2</sup>über den komplexen Zahlen  $\mathbb{C}$



**Satz 5.15 (Eigenwerte, Diagonalisierbarkeit symmetrischer Matrizen)**

Die Matrix  $A \in \mathbb{R}^{n \times n}$  sei symmetrisch, also  $A = A^\top$ .

- (a) Alle EW  $\lambda_1, \lambda_2, \dots, \lambda_n$  sind reelle Zahlen. Also können die EV immer als reelle Vektoren gewählt werden.
- (b) EV zu verschiedenen EW sind nicht nur linear unabhängig, sondern sogar orthogonal (senkrecht) zueinander.
- (c) Es gibt  $n$  jeweils zueinander orthogonale EV  $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n$  von  $A$  zu den EW  $\lambda_1, \lambda_2, \dots, \lambda_n$ , d.h.,  $A$  ist diagonalisierbar. Die daraus gebildete Matrix  $V$  erfüllt wieder

$$V^{-1}AV = \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix} = D \quad \text{Diagonalmatrix.}$$

- (d) Wenn man die EV aus (c) zusätzlich **normiert** (sodass  $\|\vec{v}_i\| = 1$  für  $i = 1, \dots, n$  gilt), dann ist

$$V^{-1} = V^\top.$$

(Eine solche Matrix  $V$  heißt **orthogonal**.) ◇

**§ 5.4 Das lineare Differentialgleichungssystem  $\vec{y}'(x) = A\vec{y}(x) + \vec{f}(x)$**

Wir betrachten jetzt nicht mehr eine einzelne Dgl, sondern mehrere Dgl, die gekoppelt sind, also Systeme von Dgl.

**Beispiel 5.16 (Chemische Reaktion)**

Eine einfache chemische Reaktion laufe nach dem Schema



mit Reaktionsraten  $k_1, k_2 > 0$  ab. Wir bezeichnen mit  $y_1(x)$  die Stoffkonzentration (bzw. Stoffmenge) von  $C_1$  zum Zeitpunkt  $x$ , analog  $y_2(x)$  und  $y_3(x)$ . Für sie gilt

$$y_1'(x) = -k_1 y_1(x),$$

d.h., die Abnahmerate der Stoffmenge ist proportional zur vorhandenen Menge. Die Konzentration von  $C_2$  nimmt aufgrund der ersten Reaktion zu und aufgrund der zweiten ab:

$$y_2'(x) = +k_1 y_1(x) - k_2 y_2(x).$$

Schließlich ändert sich  $C_3$  wie folgt:

$$y_3'(x) = +k_2 y_2(x).$$

Die Reaktionsgleichung lässt sich also kurz schreiben als

$$\begin{pmatrix} y_1'(x) \\ y_2'(x) \\ y_3'(x) \end{pmatrix} = \begin{pmatrix} -k_1 & 0 & 0 \\ k_1 & -k_2 & 0 \\ 0 & k_2 & 0 \end{pmatrix} \begin{pmatrix} y_1(x) \\ y_2(x) \\ y_3(x) \end{pmatrix} \quad \text{oder} \quad \vec{y}'(x) = A \vec{y}(x). \quad \diamond$$

**Definition 5.17 (Lineares Dgl-System)** (a) Ein Dgl-System der Form

$$\vec{y}'(x) = A \vec{y}(x) + \vec{f}(x)$$

heißt **lineares Dgl-System 1. Ordnung** mit konstanter Koeffizientenmatrix  $A \in \mathbb{R}^{n \times n}$ .

- (b) Das Dgl-System heißt **homogen**, wenn  $\vec{f}(x) \equiv \vec{0}$  ist, ansonsten **inhomogen**.
- (c) Eine Funktion  $\vec{y}(x)$  heißt **Lösung** des Dgl-Systems auf einem Intervall  $I$ , wenn  $\vec{y}: I \rightarrow \mathbb{R}^n$  diffbar ist und jede der Dgl (jede Zeile des Dgl-Systems) auf  $I$  erfüllt ist.  $\diamond$

**Bemerkung 5.18**

Solche Dgl-Systeme können viele verschiedene Prozesse beschreiben, z.B.

- elektrische Netzwerke
- mechanische Schwingungssysteme
- Klimamodelle
- die Bewegung von Fahrzeugen
- biologische Systeme (z.B. Populationsgrößen)
- die Ausbreitung von Krankheiten.  $\diamond$

### § 5.4.1 Das homogene Dgl-System

Wir untersuchen zuerst wieder den homogenen Fall und machen den Ansatz

$$\vec{y}(x) = e^{\lambda x} \vec{v}$$

mit einer noch zu bestimmenden Zahl  $\lambda$  und einem Vektor  $\vec{v}$ . Wir berechnen

$$\vec{y}'(x) = \lambda e^{\lambda x} \vec{v}$$

und setzen den Ansatz in das Dgl-System ein:

$$\begin{aligned} \vec{y}'(x) = A \vec{y}(x) &\Leftrightarrow \lambda e^{\lambda x} \vec{v} = A e^{\lambda x} \vec{v} \quad | \cdot e^{-\lambda x} \\ &\Leftrightarrow A \vec{v} = \lambda \vec{v}. \end{aligned}$$

**Fazit:** Die Funktion  $e^{\lambda x} \vec{v}$  ist genau dann eine Lösung von  $\vec{y}'(x) = A\vec{y}(x)$ , wenn  $\lambda$  ein EW von  $A$  und  $\vec{v}$  ein dazugehöriger EV ist.

**Beachte:** Dann ist auch  $c e^{\lambda x} \vec{v}$  eine Lösung des homogenen Dgl-Systems für jedes  $c \in \mathbb{R}$  (und sogar jedes  $c \in \mathbb{C}$ ).

**Satz 5.19 (Lösungen des homogenen linearen Dgl-Systems)**

Die Matrix  $A \in \mathbb{R}^{n \times n}$  habe die EW  $\lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{C}$  (möglicherweise sind nicht alle verschieden) mit zugehörigen EV  $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n \in \mathbb{C}^n$ .

(a) Dann ist

$$\vec{y}(x) = \sum_{i=1}^n c_i e^{\lambda_i x} \vec{v}_i$$

für jede Wahl von  $c_1, c_2, \dots, c_n \in \mathbb{C}$  eine Lösung von  $\vec{y}'(x) = A\vec{y}(x)$  (auf beliebigen Intervallen  $I \subseteq \mathbb{R}$ ).

(b) Wenn die EV  $\vec{v}_1, \vec{v}_2, \dots, \vec{v}_n \in \mathbb{C}^n$  linear unabhängig sind (also  $A$  diagonalisierbar ist), dann sind dies *alle* Lösungen des Dgl-Systems.<sup>3</sup> Man spricht dann wieder von der **allgemeinen Lösung** des homogenen Dgl-Systems.  $\diamond$

**Bemerkung 5.20**

Die allgemeine Lösung enthält  $n$  unbestimmte Konstanten  $c_1, c_2, \dots, c_n \in \mathbb{C}$ . Diese können durch die **Anfangsbedingung** (AB)  $\vec{y}(x_0) = \vec{y}_0 \in \mathbb{C}^n$  eindeutig festgelegt werden.  $\diamond$

**Beispiel 5.21 (Chemische Reaktion)**

Wir lösen das Dgl-System aus Beispiel 5.16 für den Fall  $k_1 = 1$ ,  $k_2 = 2$ , also mit

$$A = \begin{pmatrix} -1 & 0 & 0 \\ 1 & -2 & 0 \\ 0 & 2 & 0 \end{pmatrix}.$$

Die Eigenwerte und zugehörigen Eigenvektoren (Eigenpaare) von  $A$  sind

$$\begin{aligned} \lambda_1 &= 0, & \vec{v}_1 &= \alpha \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \\ \lambda_2 &= -1, & \vec{v}_2 &= \alpha \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix} \\ \lambda_3 &= -2, & \vec{v}_3 &= \alpha \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \end{aligned}$$

mit  $\alpha \in \mathbb{C}$ ,  $\alpha \neq 0$ . Damit ergibt sich nach Satz 5.19 folgende allgemeine Lösung des homogenen Dgl-Systems  $\vec{y}'(x) = A\vec{y}(x)$ :

$$\vec{y}(x) = \sum_{i=1}^3 c_i e^{\lambda_i x} \vec{v}_i = c_1 e^{0x} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + c_2 e^{-x} \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix} + c_3 e^{-2x} \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$$

<sup>3</sup>Wenn  $A$  nicht diagonalisierbar ist, dann „fehlen“ Lösungen. Um diese zu bestimmen, benötigt man neben den Eigenvektoren noch sogenannte Hauptvektoren. Diese Thematik wird aber hier nicht weiter vertieft.

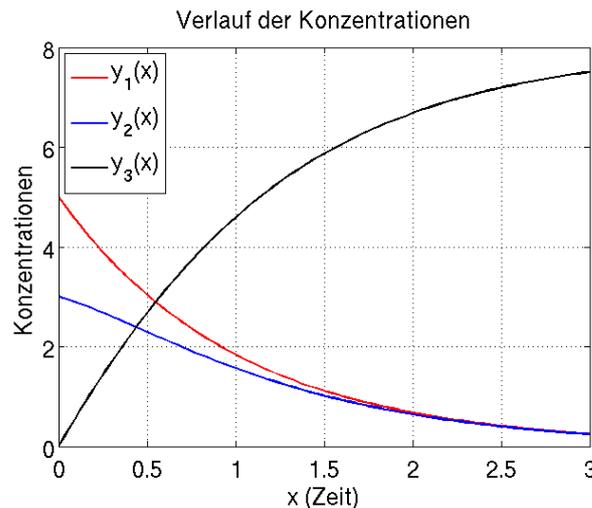
mit  $c_1, c_2, c_3 \in \mathbb{C}$ .

**Beachte:** Wenn wir reellwertige Konstanten  $c_1, c_2, c_3 \in \mathbb{R}$  verwenden, ist die Lösung  $\vec{y}(x)$  ebenfalls reellwertig. Das liegt daran, dass die EW  $\lambda_1 = 0, \lambda_2 = -1, \lambda_3 = -2$  ebenfalls alle reell waren.

Die AB  $\vec{y}(0) = (5, 3, 0)^\top$  wählt aus diesen Lösungen genau eine aus. Einsetzen ergibt:

$$\vec{y}(0) = c_1 \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} + c_2 \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix} + c_3 \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 5 \\ 3 \\ 0 \end{pmatrix},$$

also ein LGS für  $(c_1, c_2, c_3)^\top$ ! Dessen eindeutige Lösung ist hier  $(c_1, c_2, c_3) = (8, 5, -2)$ .



◇

Ende 7. V

17.05.2011

**Frage:** Was passiert bei komplexen Eigenwerten?

**Beispiel 5.22 (Mathematisches Pendel)**

Wir betrachten die Dgl des mathematischen Pendels (diese gilt nur für kleine Winkel):

$$\varphi''(x) = -\frac{g}{L} \varphi(x)$$

wobei

- $\varphi(x)$  Ausschlagswinkel zur Zeit  $x$
- $g$  Gravitationskonstante
- $L$  Pendellänge

bezeichnen. Wir formen diese Dgl zweiter Ordnung in ein System mit zwei Dgl erster Ordnung um und setzen dafür:

$$\begin{aligned} y_1(x) = \varphi(x) & \quad \text{Winkel} & \Rightarrow & \quad y_1'(x) = \varphi'(x) = y_2(x) \\ y_2(x) = \varphi'(x) & \quad \text{Winkelgeschwindigkeit} & \Rightarrow & \quad y_2'(x) = \varphi''(x) = -\frac{g}{L} y_1(x). \end{aligned}$$

Dann ergibt sich ein Dgl-System in gewohnter Form

$$\begin{pmatrix} y_1'(x) \\ y_2'(x) \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ -g/L & 0 \end{pmatrix} \begin{pmatrix} y_1(x) \\ y_2(x) \end{pmatrix}.$$

Das charakteristische Polynom der Matrix ist

$$\det(A - \lambda E) = \det \begin{pmatrix} -\lambda & 1 \\ -g/L & -\lambda \end{pmatrix} = \lambda^2 + \frac{g}{L}.$$

Die beiden Eigenwerte der Matrix sind  $\lambda_{1,2} = \pm i \sqrt{g/L} \in \mathbb{C}$ , also komplexwertig.

Wir setzen zur Abkürzung  $\omega_0 = \sqrt{g/L}$  und berechnen die (komplexen) Eigenvektoren. Das lineare Gleichungssystem

$$\begin{pmatrix} -i\omega_0 & 1 & | & 0 \\ -\omega_0^2 & -i\omega_0 & | & 0 \end{pmatrix} \quad \text{hat die allgemeine Lösung} \quad \vec{x} = \alpha \begin{pmatrix} -i \\ \omega_0 \end{pmatrix}$$

$$\begin{pmatrix} i\omega_0 & 1 & | & 0 \\ -\omega_0^2 & i\omega_0 & | & 0 \end{pmatrix} \quad \text{hat die allgemeine Lösung} \quad \vec{x} = \alpha \begin{pmatrix} i \\ \omega_0 \end{pmatrix}$$

jeweils mit  $\alpha \in \mathbb{C}$ . Nach Satz 5.19 besteht die allgemeine Lösung des Dgl-Systems aus

$$\begin{aligned} \vec{y}(x) &= c_1 e^{i\omega_0 x} \begin{pmatrix} -i \\ \omega_0 \end{pmatrix} + c_2 e^{-i\omega_0 x} \begin{pmatrix} i \\ \omega_0 \end{pmatrix} \\ &= c_1 \left( \cos(\omega_0 x) + i \sin(\omega_0 x) \right) \begin{pmatrix} -i \\ \omega_0 \end{pmatrix} + c_2 \left( \cos(-\omega_0 x) + i \sin(-\omega_0 x) \right) \begin{pmatrix} i \\ \omega_0 \end{pmatrix} \end{aligned}$$

mit  $c_1, c_2 \in \mathbb{C}$ , vgl. Eulersche Formel (§ 1.4.2):

**Frage:** Diese komplexwertigen Lösungen für  $y_1(x)$  (Winkel) und  $y_2(x)$  (Winkelgeschwindigkeit) sind für uns nicht sinnvoll. Wie erhalten wir reellwertige Lösungen?

Wir nehmen dazu *eine* der beiden Lösungen (die zu einem Paar komplex-konjugierter Eigenwerte gehören), bilden die Real- und Imaginärteile und versehen diese mit neuen Konstanten  $\tilde{c}_1, \tilde{c}_2 \in \mathbb{R}$ :

$$\begin{aligned} \operatorname{Re} \left( \cos(\omega_0 x) + i \sin(\omega_0 x) \right) \begin{pmatrix} -i \\ \omega_0 \end{pmatrix} &= \begin{pmatrix} \sin(\omega_0 x) \\ \omega_0 \cos(\omega_0 x) \end{pmatrix} \\ \operatorname{Im} \left( \cos(\omega_0 x) + i \sin(\omega_0 x) \right) \begin{pmatrix} -i \\ \omega_0 \end{pmatrix} &= \begin{pmatrix} -\cos(\omega_0 x) \\ \omega_0 \sin(\omega_0 x) \end{pmatrix}. \end{aligned}$$

Dies führt auf die endgültige allgemeine (reellwertige) Lösung

$$\vec{y}(x) = \begin{pmatrix} \varphi(x) \\ \varphi'(x) \end{pmatrix} = \tilde{c}_1 \begin{pmatrix} \sin(\omega_0 x) \\ \omega_0 \cos(\omega_0 x) \end{pmatrix} + \tilde{c}_2 \begin{pmatrix} -\cos(\omega_0 x) \\ \omega_0 \sin(\omega_0 x) \end{pmatrix}$$

mit  $\tilde{c}_1, \tilde{c}_2 \in \mathbb{R}$ . Die Lösungen entsprechen periodischen Schwingungsbewegungen. Die Zahl  $\omega_0 = \sqrt{g/L}$  heißt die **Eigen-Kreisfrequenz** des Pendels (Einheit: 1/s). Die **Periodendauer** beträgt  $2\pi/\omega_0$ .

Durch Einsetzen der AB  $\vec{y}(0) = (2, 0)^\top$  ergibt sich das LGS

$$\vec{y}(0) = \tilde{c}_1 \begin{pmatrix} 0 \\ \omega_0 \end{pmatrix} + \tilde{c}_2 \begin{pmatrix} -1 \\ 0 \end{pmatrix} \stackrel{!}{=} \begin{pmatrix} 2 \\ 0 \end{pmatrix}$$

mit der eindeutigen Lösung  $(\tilde{c}_1, \tilde{c}_2) = (0, -2)$ , also

$$\vec{y}(x) = -2 \begin{pmatrix} -\cos(\omega_0 x) \\ \omega_0 \sin(\omega_0 x) \end{pmatrix}.$$



**Beispiel 5.23 (Mathematisches Pendel mit periodischer Anregung)**

Wir kommen zurück zum Dgl-System des mathematischen Pendels, siehe Beispiel 5.22. Wir geben zusätzlich eine harmonische Anregung einer Form eines Moments vor, das periodisch auf das Pendel wirkt:

$$\vec{f}(x) = \begin{pmatrix} 0 \\ 50 \sin(2\omega_0 x) \end{pmatrix}.$$

Die Anregungsfrequenz ist also die doppelte Eigenfrequenz ( $\omega = 2\omega_0$ ). Wir erhalten das inhomogene Dgl-System

$$\begin{pmatrix} y_1'(x) \\ y_2'(x) \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & 1 \\ -\omega_0^2 & 0 \end{pmatrix}}_A \begin{pmatrix} y_1(x) \\ y_2(x) \end{pmatrix} + \begin{pmatrix} 0 \\ 50 \sin(2\omega_0 x) \end{pmatrix}$$

mit der Abkürzung  $\omega_0 = \sqrt{g/L}$ . Zur Bestimmung einer partikulären Lösung machen wir den Ansatz vom Typ der rechten Seite.

**Beachte:** Obwohl in  $\vec{f}(x)$  nur sin-Terme vorkommen, müssen wir trotzdem im Ansatz sin- und cos-Terme vorsehen:

$$\begin{aligned} \vec{y}_p(x) &= \vec{c}_1 \sin(2\omega_0 x) + \vec{c}_2 \cos(2\omega_0 x) \\ \Rightarrow \vec{y}_p'(x) &= 2\omega_0 \vec{c}_1 \cos(2\omega_0 x) - 2\omega_0 \vec{c}_2 \sin(2\omega_0 x) \end{aligned}$$

mit Unbekannten  $\vec{c}_1 = (c_{11}, c_{12})^\top$  und  $\vec{c}_2 = (c_{21}, c_{22})^\top$ . Zum späteren Einsetzen in die Dgl  $\vec{y}_p'(x) = A\vec{y}_p(x) + \vec{f}(x)$  berechnen wir auch

$$\begin{aligned} A\vec{y}_p(x) &= \begin{pmatrix} 0 & 1 \\ -\omega_0^2 & 0 \end{pmatrix} [\vec{c}_1 \sin(2\omega_0 x) + \vec{c}_2 \cos(2\omega_0 x)] \\ &= \begin{pmatrix} 0 \cdot c_{11} \sin(2\omega_0 x) + 1 \cdot c_{12} \sin(2\omega_0 x) + 0 \cdot c_{21} \cos(2\omega_0 x) + 1 \cdot c_{22} \cos(2\omega_0 x) \\ -\omega_0^2 \cdot c_{11} \sin(2\omega_0 x) + 0 \cdot c_{12} \sin(2\omega_0 x) - \omega_0^2 \cdot c_{21} \cos(2\omega_0 x) + 0 \cdot c_{22} \cos(2\omega_0 x) \end{pmatrix} \\ &= \begin{pmatrix} c_{12} \sin(2\omega_0 x) + c_{22} \cos(2\omega_0 x) \\ -\omega_0^2 \cdot c_{11} \sin(2\omega_0 x) - \omega_0^2 \cdot c_{21} \cos(2\omega_0 x) \end{pmatrix}. \end{aligned}$$

Koeffizientenvergleich in  $\vec{y}_p'(x) = A\vec{y}_p(x) + \vec{f}(x)$  ergibt das LGS (vgl. (\*))

$$\begin{pmatrix} 0 & -1 & -2\omega_0 & 0 \\ \omega_0^2 & 0 & 0 & -2\omega_0 \\ 2\omega_0 & 0 & 0 & -1 \\ 0 & 2\omega_0 & \omega_0^2 & 0 \end{pmatrix} \begin{pmatrix} c_{11} \\ c_{12} \\ c_{21} \\ c_{22} \end{pmatrix} = \begin{pmatrix} 0 \\ 50 \\ 0 \\ 0 \end{pmatrix}$$

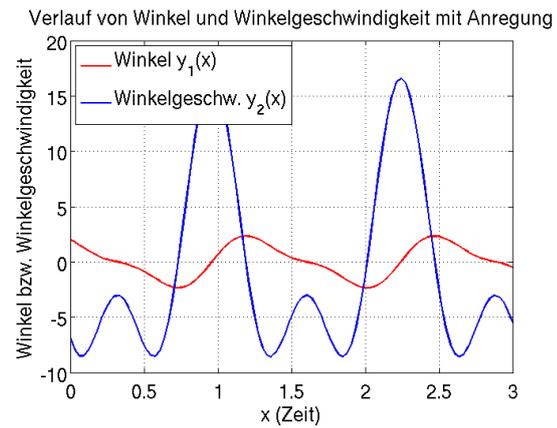
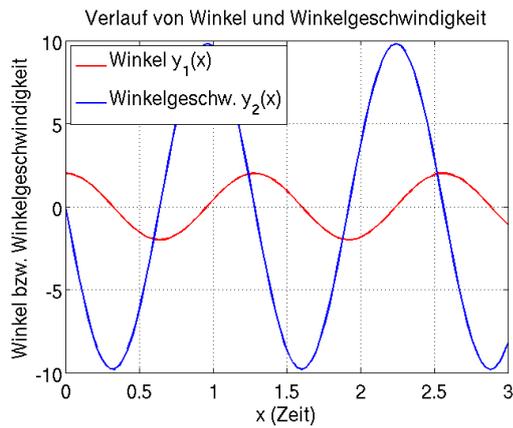
mit der eindeutigen Lösung

$$\vec{c}_1 = \left( \frac{-50}{3\omega_0^2}, 0 \right)^\top, \quad \vec{c}_2 = \left( 0, \frac{-100}{3\omega_0} \right)^\top.$$

Zusammen mit der allgemeinen Lösung des *homogenen* Systems (siehe Beispiel 5.22) bekommen wir

$$\vec{y}(x) = \tilde{c}_1 \begin{pmatrix} \sin(\omega_0 x) \\ \omega_0 \cos(\omega_0 x) \end{pmatrix} + \tilde{c}_2 \begin{pmatrix} -\cos(\omega_0 x) \\ \omega_0 \sin(\omega_0 x) \end{pmatrix} - \frac{50}{3\omega_0^2} \begin{pmatrix} \sin(2\omega_0 x) \\ 0 \end{pmatrix} - \frac{100}{3\omega_0} \begin{pmatrix} 0 \\ \cos(2\omega_0 x) \end{pmatrix}$$

mit  $\tilde{c}_1, \tilde{c}_2 \in \mathbb{R}$ . Durch Vorgabe einer AB  $\vec{y}(0) = \vec{y}_0$  wird diese Lösung wieder eindeutig.



**Beachte:** Im Beispiel haben wir eine periodische Anregung  $\vec{f}(x)$  mit Kreisfrequenz  $2\omega_0$  (doppelte Eigen-Kreisfrequenz) gewählt. Bei Anregungen mit Frequenzen nahe der Eigenfrequenz  $\omega_0$  kommt es zur **Resonanzkatastrophe**. ◇

Ende 8. V

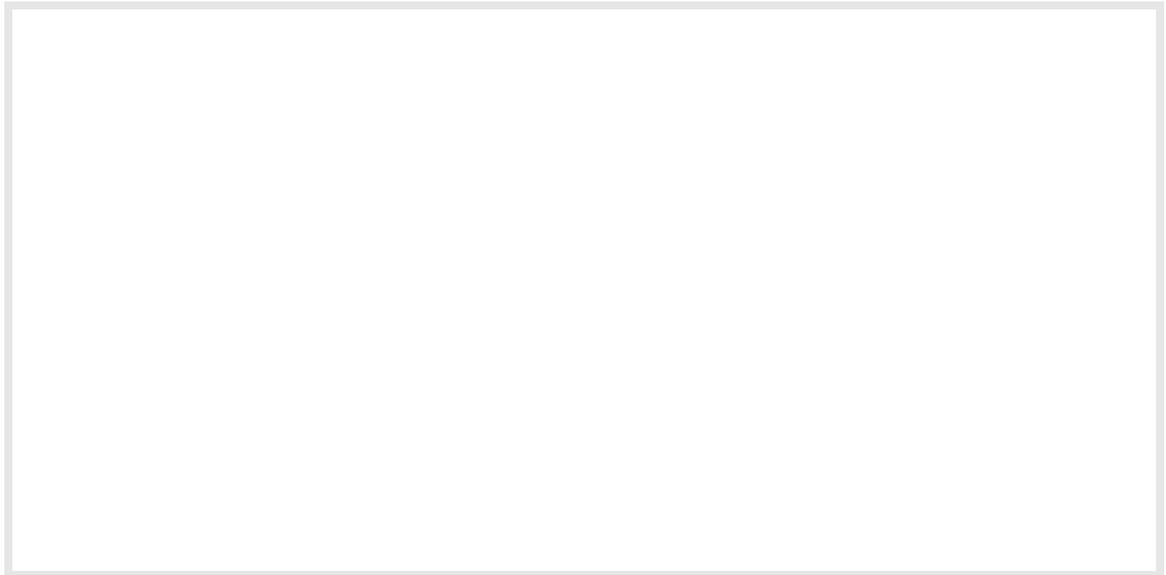
24.05.2011

## § 6 Lineare Optimierung

### § 6.1 Einführung

Lineare Optimierungsaufgaben spielen insbesondere in den Wirtschaftswissenschaften eine große Rolle.

#### Beispiel 6.1 (Mozartproblem)



#### Definition 6.2 (Lineare Optimierungsaufgabe)

(a) Eine Aufgabe der Form

$$\text{Minimiere/Maximiere } c_1 x_1 + c_2 x_2 + \dots + c_n x_n$$

$$\text{sodass } a_{11} x_1 + a_{12} x_2 + \dots + a_{1n} x_n \quad \square \quad b_1$$

$$\text{und } a_{21} x_1 + a_{22} x_2 + \dots + a_{2n} x_n \quad \square \quad b_2$$

$$\vdots \quad \vdots \quad \vdots \quad \vdots$$

$$\text{und } a_{m1} x_1 + a_{m2} x_2 + \dots + a_{mn} x_n \quad \square \quad b_m$$

$$\text{sowie } x_1 \geq 0, x_2 \geq 0, \dots, x_n \geq 0$$

gelten, heißt **lineare Optimierungsaufgabe** (LOA). Dabei steht „ $\square$ “ jeweils für  $\leq$ ,  $=$  oder  $\geq$ . Die Unbekannten  $x_1, x_2, \dots, x_n$  heißen die **Variablen** der LOA.

(b) Die **Zielfunktion** kann auch als Skalarprodukt (siehe Definition 2.15)  $\vec{c} \cdot \vec{x}$  oder  $\vec{c}^\top \vec{x}$  mit dem **Kostenvektor**  $\vec{c} \in \mathbb{R}^n$  geschrieben werden. (Er kann statt Kosten auch einen Nutzen beschreiben.)

(c) Die Menge aller  $\vec{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$ , die alle Gleichungen und Ungleichungen (**Beschränkungen**) erfüllen, heißt **zulässige Menge** oder **zulässiger Bereich** der LOA.

- (d) Ein Punkt  $\vec{x}^* \in \mathbb{R}^n$  heißt **optimale Lösung** der LOA, wenn er zulässig ist und kein anderer zulässiger Punkt einen größeren (bei Maximierung) bzw. kleineren (bei Minimierung) Zielfunktionswert hat.  $\diamond$

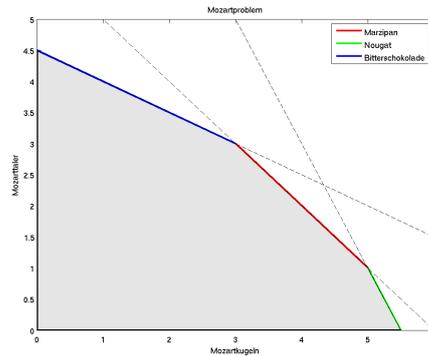
§ 6.2 Grafische Lösung

Wenn nur zwei Optimierungsvariablen  $x_1, x_2$  vorliegen, kann man eine LOA grafisch lösen.

**Beispiel 6.3 (Mozartproblem)**

- (1.) Darstellung der zulässigen Menge

$$\begin{aligned} x_1 + x_2 &\leq 6 && \text{Marzipan} \\ 2x_1 + x_2 &\leq 11 && \text{Nougat} \\ x_1 + 2x_2 &\leq 9 && \text{Bitterschokolade} \\ x_1 &\geq 0 \\ x_2 &\geq 0 \end{aligned}$$

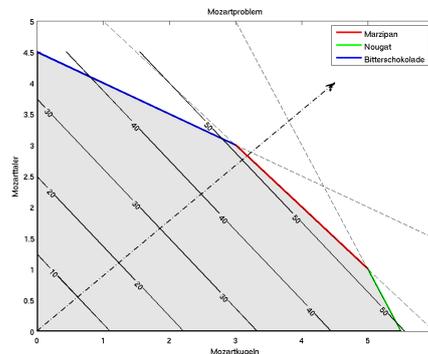


- (2.) Darstellung einiger Höhenlinien der Zielfunktion

Die Zielfunktion ist  $9x_1 + 8x_2$ , also ist

$$\vec{c} = \begin{pmatrix} 9 \\ 8 \end{pmatrix}$$

der Kostenvektor der Aufgabe. Die Höhenlinien der Zielfunktion verlaufen senkrecht zu  $\vec{c}$ . Die Funktion steigt in Richtung von  $\vec{c}$  und fällt in Richtung von  $-\vec{c}$ .



- (3.) Verschiebe die Höhenlinien soweit wie möglich in Maximierungsrichtung  $\vec{c}$  bzw. Minimierungsrichtung  $-\vec{c}$ , sodass sie den zulässigen Bereich gerade noch schneiden.

- (4.) Bestimmung der Lösung

- (A) Die Lösung liegt in einer Ecke des zulässigen Bereiches: Lies die Koordinaten aus der Zeichnung ab, oder bestimme sie als Schnittpunkt zweier Geraden.

Im Beispiel:  $\vec{x}^* = (5, 1)^\top$  ablesen oder als Schnittpunkt der Geraden  $2x_1 + x_2 = 11$  und  $x_1 + x_2 = 6$  bestimmen (LGS).

- (B) Die Lösung ist eine ganze Kante des zulässigen Bereiches: Gib die Geradengleichung der Kante und die Endpunkte an.

Setze die gefundene Lösung in die Zielfunktion ein, um den Maximalwert bzw. Minimalwert zu ermitteln.

Im Beispiel:  $9 \cdot 5 + 8 \cdot 1 = 53$ .

**Lösung:** Der maximale Umsatz von 53 ergibt sich bei Produktion von  $x_1^* = 5$  Einheiten an Mozartkugeln und  $x_2^* = 1$  Einheit an Mozarttalern.  $\diamond$

**Frage:** Welche Fälle können vorkommen?

- Die LOA ist **lösbar**, d.h., es existiert mindestens eine Lösung der Aufgabe. Diese liegt dann in einer Ecke, oder aber eine ganze Kante (allgemeiner: Seitenfläche) der zulässigen Menge ist Lösung.
- Es existiert keine Lösung der LOA, weil der zulässige Bereich leer ist. Die LOA heißt dann **unzulässig**.
- Es existiert keine Lösung der LOA, weil die Zielfunktion auf dem zulässigen Bereich beliebig groß (bei Maximierung) oder beliebig klein (bei Minimierung) werden kann. Die LOA heißt dann **unbeschränkt**.  $\diamond$

**Beispiel 6.4 (Mögliche Fälle bei LOA)** (a) Im obigen Beispiel 6.1 ist die eindeutige Lösung  $\vec{x}^* = (5, 1)^\top$ . Ändert man aber den Kostenvektor z.B. in  $\vec{c} = (8, 8)^\top$ , so besteht die Lösungsmenge aus einer ganzen Kante (die rote „Marzipan“-Kante).

(b) Fügt man dem Beispiel 6.1 die Beschränkung  $x_1 + x_2 \geq 8$  hinzu, wird der zulässige Bereich leer. Das Problem ist dann unzulässig.

(c) Folgende Aufgabe ist unbeschränkt:

$$\text{Maximiere } 4x_1 + 3x_2$$

$$\text{sodass } -\frac{1}{2}x_1 + x_2 \geq -1$$

$$-x_1 + x_2 \leq 1$$

$$\text{und } x_1 \geq 0$$

$$x_2 \geq 0.$$

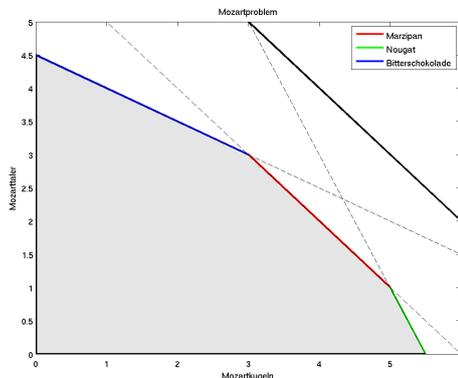


Illustration des Falles (b): unzulässige LOA

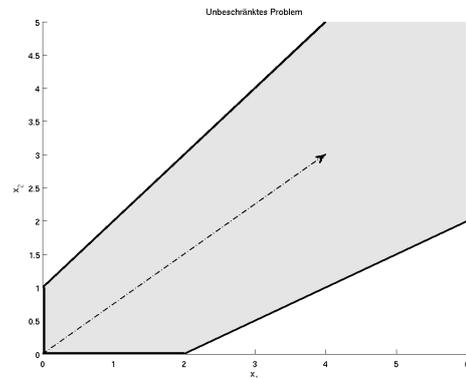


Illustration des Falles (c): unbeschränkte LOA

$\diamond$



**Beachte:** Es kommen wieder neue Variablen hinzu. Es gilt jedoch weiterhin: Alle Variablen unterliegen einer Nicht-Negativitätsbeschränkung. Alle weiteren Bedingungen sind Gleichungen.

- (4.) Falls in einer oder mehreren Gleichungen  $b_i < 0$  gilt, so werden diese mit  $(-1)$  multipliziert.
- (5.) Liegt ein Minimierungsproblem vor, so wird stattdessen die negative Zielfunktion maximiert:

$$\text{Minimiere } \vec{c}^\top \vec{x} \rightsquigarrow \text{Maximiere } -\vec{c}^\top \vec{x}.$$

Additive Konstanten in der Zielfunktion können ignoriert werden.

**Beachte:** Nach Lösung der Aufgabe müssen diese Transformationen rückgängig gemacht werden, um die Lösung interpretieren zu können.

### Beispiel 6.7 (Transformation in Normalform<sup>4</sup>)

Wir transformieren die Aufgabe

$\begin{aligned} &\text{Minimiere} && -x_1 + x_2 - x_3 - 2x_4 \\ &\text{sodass} && 3x_1 + x_2 - 3x_3 - x_4 \leq -12 \\ &&& 2x_1 + x_2 + x_3 = 2 \\ &\text{und} && x_1 \geq 0, \quad x_3 \geq 2, \quad x_4 \leq 5 \end{aligned}$
---

in Normalform. Wir ersetzen in Schritten (1.) und (2.)

$$\begin{aligned} x_4 &\rightsquigarrow \tilde{x}_4 = 5 - x_4 && \Rightarrow \tilde{x}_4 \geq 0 \\ x_3 &\rightsquigarrow \tilde{x}_3 = x_3 - 2 && \Rightarrow \tilde{x}_3 \geq 0 \\ x_2 &\rightsquigarrow x_2 = x_2^+ - x_2^- && \Rightarrow x_2^+ \geq 0, \quad x_2^- \geq 0 \end{aligned}$$

Die neuen Variablen sind  $x_1, x_2^+, x_2^-, \tilde{x}_3, \tilde{x}_4$ . Wir ersetzen in der Zielfunktion und den restlichen Beschränkungen die alten durch die neuen Variablen:

$\begin{aligned} &\text{Minimiere} && -x_1 + (x_2^+ - x_2^-) - (\tilde{x}_3 + 2) - 2(5 - \tilde{x}_4) \\ &\text{sodass} && 3x_1 + (x_2^+ - x_2^-) - 3(\tilde{x}_3 + 2) - (5 - \tilde{x}_4) \leq -12 \\ &&& 2x_1 + (x_2^+ - x_2^-) + (\tilde{x}_3 + 2) = 2 \\ &\text{und} && x_1 \geq 0, \quad x_2^+ \geq 0, \quad x_2^- \geq 0, \quad \tilde{x}_3 \geq 0, \quad \tilde{x}_4 \geq 0 \end{aligned}$
---

Durch Vereinfachen:

$\begin{aligned} &\text{Minimiere} && -x_1 + x_2^+ - x_2^- - \tilde{x}_3 + 2\tilde{x}_4 - 12 \\ &\text{sodass} && 3x_1 + x_2^+ - x_2^- - 3\tilde{x}_3 + \tilde{x}_4 \leq -1 \\ &&& 2x_1 + x_2^+ - x_2^- + \tilde{x}_3 = 0 \\ &\text{und} && x_1 \geq 0, \quad x_2^+ \geq 0, \quad x_2^- \geq 0, \quad \tilde{x}_3 \geq 0, \quad \tilde{x}_4 \geq 0 \end{aligned}$
---

In Schritt (3.) führen wir eine Schlupfvariable  $u_1$  ein und erhalten

<sup>4</sup>aus [Luderer and Würker, 1997, Abschnitt 5.3]

Minimiere	$-x_1 + x_2^+ - x_2^- - \tilde{x}_3 + 2\tilde{x}_4 - 12$
sodass	$3x_1 + x_2^+ - x_2^- - 3\tilde{x}_3 + \tilde{x}_4 + u_1 = -1$
	$2x_1 + x_2^+ - x_2^- + \tilde{x}_3 = 0$
und	$x_1 \geq 0, \quad x_2^+ \geq 0, \quad x_2^- \geq 0, \quad \tilde{x}_3 \geq 0, \quad \tilde{x}_4 \geq 0, \quad u_1 \geq 0$

In Schritt (4.) multiplizieren wir die erste Gleichung mit  $(-1)$ , um eine nicht-negative rechte Seite zu erhalten:

Minimiere	$-x_1 + x_2^+ - x_2^- - \tilde{x}_3 + 2\tilde{x}_4 - 12$
sodass	$-3x_1 - x_2^+ + x_2^- + 3\tilde{x}_3 - \tilde{x}_4 - u_1 = 1$
	$2x_1 + x_2^+ - x_2^- + \tilde{x}_3 = 0$
und	$x_1 \geq 0, \quad x_2^+ \geq 0, \quad x_2^- \geq 0, \quad \tilde{x}_3 \geq 0, \quad \tilde{x}_4 \geq 0, \quad u_1 \geq 0$

Schließlich wird in Schritt (5.) die Zielfunktion auf Maximierung umgeschrieben:

Maximiere	$x_1 - x_2^+ + x_2^- + \tilde{x}_3 - 2\tilde{x}_4 + 12$
sodass	$-3x_1 - x_2^+ + x_2^- + 3\tilde{x}_3 - \tilde{x}_4 - u_1 = 1$
	$2x_1 + x_2^+ - x_2^- + \tilde{x}_3 = 0$
und	$x_1 \geq 0, \quad x_2^+ \geq 0, \quad x_2^- \geq 0, \quad \tilde{x}_3 \geq 0, \quad \tilde{x}_4 \geq 0, \quad u_1 \geq 0$

oder in Kurzform

$$A = \begin{pmatrix} -3 & -1 & 1 & 3 & -1 & -1 \\ 2 & 1 & -1 & 1 & 0 & 0 \end{pmatrix}, \quad \vec{b} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \text{und } \vec{c} = (1, -1, 1, 1, -2, 0)^\top.$$

Die additive Konstante 12 in der Zielfunktion können wir ignorieren. Die Lösung (z.B. mit dem Simplex-Verfahren, siehe § 6.4) lautet

$$(1, 0, 4, 2, 0, 6)^\top$$

oder in ursprünglichen Variablen

$$x_1^* = 1, \quad x_2^* = x_2^+ - x_2^- = -4, \quad x_3^* = \tilde{x}_3 + 2 = 4, \quad x_4^* = 5 - \tilde{x}_4 = 5.$$

Die Schlupfvariable  $u_1 = 6$  sagt uns, dass in der ursprünglichen Ungleichungsbeschränkung

$$\underbrace{3x_1^* + x_2^* - 3x_3^* - x_4^*}_{=-18} \leq -12$$

noch 6 Einheiten „Luft ist“.

Der optimale Wert der Zielfunktion ist

$$-x_1^* + x_2^* - x_3^* - 2x_4^* = -1 - 4 - 4 - 10 = -19.$$

◇

### § 6.4 Das Simplex-Verfahren

Wir nehmen an, dass die LOA in Normalform vorliegt:

$$\text{Maximiere } \vec{c}^\top \vec{x}, \quad \text{wobei } A\vec{x} = \vec{b} \quad \text{sowie} \quad \vec{x} \geq \vec{0}$$

mit  $A \in \mathbb{R}^{m \times n}$ ,  $\vec{b} \in \mathbb{R}^m$  und  $\vec{c} \in \mathbb{R}^n$ .

**Zur Erinnerung:**  $n$  = Anzahl der Variablen,  $m$  = Anzahl der Gleichungsbeschränkungen.

#### Voraussetzung 6.8

Neben  $\vec{b} \geq \vec{0}$  nehmen wir noch an, dass  $m \leq n$  gilt (weniger Gleichungen als Variable) und dass  $A$  keine linear abhängigen (überflüssigen/widersprüchlichen) Zeilen enthält, also  $\text{Rang}(A) = m$  ist.  $\diamond$

**Vorbereitungsschritt (Erzeugung einer zulässigen Basislösung):** Forme das LGS  $A\vec{x} = \vec{b}$  ähnlich wie beim Gauß-Verfahren so um, dass einige Spalten zusammen die Einheitsmatrix  $E \in \mathbb{R}^{m \times m}$  ergeben:

Beispiel:  $A \in \mathbb{R}^{4 \times 6}$ , also  $n = 6$  (Variablen) und  $m = 4$  (Gleichungen)

$$\underbrace{\begin{array}{cccccc|c} * & * & * & * & * & * & * \\ * & * & * & * & * & * & * \\ * & * & * & * & * & * & * \\ * & * & * & * & * & * & * \end{array}}_A \underbrace{\quad}_\vec{b} \rightsquigarrow \underbrace{\begin{array}{cccccc|c} 1 & 0 & * & 0 & * & 0 & \tilde{b}_1 \\ 0 & 1 & * & 0 & * & 0 & \tilde{b}_2 \\ 0 & 0 & * & 1 & * & 0 & \tilde{b}_4 \\ 0 & 0 & * & 0 & * & 1 & \tilde{b}_6 \end{array}}_{\tilde{A}} \quad \text{oder} \quad \begin{array}{cccccc|c} 0 & 0 & * & 1 & * & 0 & \tilde{b}_4 \\ 0 & 0 & * & 0 & * & 1 & \tilde{b}_6 \\ 1 & 0 & * & 0 & * & 0 & \tilde{b}_1 \\ 0 & 1 & * & 0 & * & 0 & \tilde{b}_2 \end{array}$$

Diejenigen Spalten (Variablen), die zur Einheitsmatrix beitragen, heißen **Basisvariablen**<sup>5</sup>, hier  $x_1, x_2, x_4$  und  $x_6$ . Die anderen Spalten (Variablen) sind die **Nichtbasisvariablen**, hier  $x_3$  und  $x_5$ .

#### Definition 6.9 (Basis, Basislösung)

- Eine Menge  $B \subseteq \{1, 2, \dots, n\}$  mit  $m$  Elementen heißt **Basis**, wenn die zugehörigen Spalten von  $A$  linear unabhängig sind. Die Restmenge  $N$  heißt **Nichtbasis**.  
(Im Beispiel:  $B = \{1, 2, 4, 6\}$ ,  $N = \{3, 5\}$ .)
- Jeder Vektor  $\vec{x} \in \mathbb{R}^n$  kann in  $x_B$  und  $x_N$  zerlegt werden.  
(Im Beispiel:  $x_B = (x_1, x_2, x_4, x_6)^\top$  und  $x_N = (x_3, x_5)^\top$ .)
- Ein Vektor  $\vec{x}$  heißt **Basislösung** oder **Basisvektor**, wenn gilt:

$$A\vec{x} = \vec{b}, \quad \vec{x}_N = \vec{0}.$$

Er heißt **zulässige Basislösung** oder **zulässiger Basisvektor**, wenn zusätzlich  $\vec{x}_B \geq \vec{0}$  gilt.  $\diamond$

<sup>5</sup>Die Bezeichnung kommt daher, dass die zugehörigen Spalten 1,2,4,6 von  $A$  linear unabhängig sind, also eine Basis des  $\mathbb{R}^4$  bilden, vgl. Definition 2.10.

Im Beispiel: Nach dem Vorbereitungsschritt lässt sich eine Basislösung leicht ablesen.<sup>6</sup>

$$\vec{x} = \begin{pmatrix} \tilde{b}_1 \\ \tilde{b}_2 \\ 0 \\ \tilde{b}_4 \\ 0 \\ \tilde{b}_6 \end{pmatrix} \quad \text{oder kurz: } \vec{x}_B = \tilde{\vec{b}}, \quad \vec{x}_N = \vec{0}.$$

Falls die entstandene rechte Seite  $\tilde{\vec{b}} \geq \vec{0}$  erfüllt, so ist diese Basislösung auch zulässig.

**Bemerkung 6.10**

Die zulässigen Basislösungen entsprechen den Ecken der zulässigen Menge. ◇

**Idee des Simplex-Verfahrens:** Gehe von einer Ecke der zulässigen Menge zu einer benachbarten Ecke (d.h., tausche einen Index  $B \leftrightarrow N$ ), wobei sich der Zielfunktionswert verbessern (vergrößern) soll.

Wir wählen einen Index  $k \in N$  und lassen  $x_k \geq 0$  werden.<sup>7</sup> Die restlichen Einträge in  $\vec{x}_N$  behalten wir bei 0. Um weiterhin  $A\vec{x} = \vec{b}$  zu garantieren, muss sich auch  $\vec{x}_B$  ändern:

$$A\vec{x} = \vec{b} \Leftrightarrow E\vec{x}_B + \tilde{A}_N \underbrace{\vec{x}_N}_{=0 \text{ bis auf } x_k} = \vec{b} \Leftrightarrow \vec{x}_B = \tilde{\vec{b}} - \tilde{a}_k x_k,$$

wobei  $\tilde{a}_k$  die  $k$ -te Spalte von  $\tilde{A}$  ist.

$\vec{x}$ alt	$\vec{x}$ neu
$x_j = 0, \quad j \in N \setminus \{k\}$	$x_j = 0, \quad j \in N \setminus \{k\}$
$x_k = 0$	$x_k \geq 0$
$x_j = \tilde{b}_j, \quad j \in B$	$x_j = \tilde{b}_j - \tilde{a}_{jk} x_k, \quad j \in B$

Wir vergleichen die Funktionswerte der Zielfunktion:

$$\begin{aligned} \vec{x} \text{ alt : } & \sum_{j=1}^n c_j x_j = \sum_{j \in B} c_j x_j = \sum_{j \in B} c_j \tilde{b}_j \\ \vec{x} \text{ neu : } & \sum_{j=1}^n c_j x_j = \sum_{j \in B} c_j x_j + c_k x_k = \sum_{j \in B} c_j (\tilde{b}_j - \tilde{a}_{jk} x_k) + c_k x_k \\ & = \sum_{j \in B} c_j \tilde{b}_j - \left( \sum_{j \in B} c_j \tilde{a}_{jk} - c_k \right) x_k. \end{aligned}$$

Die Funktionswerte unterscheiden sich also um

$$-\Delta_k x_k \quad \text{mit} \quad \Delta_k := \sum_{j \in B} c_j \tilde{a}_{jk} - c_k.$$

Man nennt  $\Delta_k$  einen **Optimalitätsindikator**.

**Frage:** Welches  $k \in N$  wählen wir, um den Funktionswert zu verbessern?

<sup>6</sup>Dies gilt unabhängig von der Reihenfolge in  $B$ !

<sup>7</sup>Das heißt, der Index  $k$  wechselt von  $N$  nach  $B$ .

Wir wählen  $k \in N$  so, dass  $\Delta_k < 0$  ist! Damit wird der Funktionswert beim Übergang von „ $\vec{x}$  alt“ zu „ $\vec{x}$  neu“ größer, da ja  $x_k \geq 0$  gelten muss.

**Satz 6.11 (Optimalitätstest)**

Sei  $\vec{x} \in \mathbb{R}^n$  eine zulässige Basislösung. Wenn  $\Delta_k \geq 0$  für alle  $k \in N$  sind, dann ist  $\vec{x}$  eine Lösung der LOA.  $\diamond$

**Frage:** Auf welchen Wert setzen wir  $x_k \geq 0$ ?

So groß wie möglich, denn der Zugewinn in der Zielfunktion ist  $-\Delta_k x_k$ , also proportional zu  $x_k$ ! Dabei ist aber zu beachten, dass kein  $x_j$  für  $j \in B$  negativ werden darf:

$$x_j = \tilde{b}_j - \tilde{a}_{jk} x_k \stackrel{!}{\geq} 0 \quad \text{für alle } j \in B.$$

Dazu machen wir den **Quotiententest**: Bilde die Quotienten

$$Q_j = \frac{\tilde{b}_j}{\tilde{a}_{jk}} \quad \text{für diejenigen } j \in B \text{ mit } \tilde{a}_{jk} > 0$$

und wähle den kleinsten davon. Wir merken uns den Index  $r \in B$ , bei dem das kleinste  $Q_j$  auftritt. Dieser Index wechselt von  $B$  nach  $N$ :

$$\begin{aligned} N &\rightsquigarrow N \cup \{r\} \setminus \{k\} \\ B &\rightsquigarrow B \cup \{k\} \setminus \{r\}. \end{aligned}$$

Im neuen  $\vec{x}$  wird  $x_r = 0$  sein.

**Satz 6.12 (Unbeschränkheitstest)**

Wenn beim Quotiententest auffällt, dass  $\tilde{a}_{jk} \leq 0$  sind für alle  $j \in B$ , dann ist die LOA unbeschränkt, also nicht lösbar.  $\diamond$

**Frage:** Was ist noch zu tun?

Es muss wieder die Situation wie am Anfang hergestellt werden, d.h., die Matrix  $\tilde{A}$  und rechte Seite  $\tilde{b}$  müssen so umgeformt werden wie beim Vorbereitungsschritt. Dabei müssen nun die zur *neuen* Basis  $B$  gehörenden Spalten die Einheitsmatrix bilden. Damit ist *ein Schritt* des Simplex-Verfahrens beschrieben. Das gesamte Verfahren kann bequem mit Hilfe eines sogenannten **Simplextableaus** ausgeführt werden.

**Beispiel 6.13**

In unserem Beispiel ist

$$A = \begin{pmatrix} 8 & -3 & 1 & 1 & 0 \\ 3 & 2 & 0 & 1 & 1 \\ 1 & 4 & 0 & 0 & 1 \end{pmatrix}, \quad \vec{b} = \begin{pmatrix} 33 \\ 29 \\ 21 \end{pmatrix} \quad \text{und} \quad \vec{c} = (3, 2, -3, -4, -1)^\top.$$

Wir wählen also ein Simplextableau für  $n = 5$  Variablen und  $m = 3$  Gleichungen.

**Vorbereitung des Simplextableaus**

- (1.) Trage in die erste Zeile die **Namen der Variablen** (z.B.  $x_1, x_2, \dots, x_n$ ) ein und darunter die **Koeffizienten der Zielfunktion** ( $c_1, c_2, \dots, c_n$ ).

Nr.	Var. →							$x_B \downarrow$	$Q \downarrow$
	$B \downarrow$	$c_B \downarrow$							
1									
2									
3									
		$\Delta \rightarrow$							$\leftarrow z$

(2.) Trage darunter die Matrix  $A$  und rechts davon die rechte Seite  $\vec{b}$  in das Tableau ein.

Nr.	Var. →							$x_B \downarrow$	$Q \downarrow$
	$B \downarrow$	$c_B \downarrow$							
1									
2									
3									
		$\Delta \rightarrow$							$\leftarrow z$

(3.) Forme das LGS  $A\vec{x} = \vec{b}$  wie beim Gauß-Verfahren so um, dass einige Spalten zusammen die **Einheitsmatrix**  $E \in \mathbb{R}^{m \times m}$  ergeben. Diese Spalten in der richtigen Reihenfolge bilden die **Basis**  $B$ . Im Beispiel bietet sich  $B = \{3, 4, 5\}$  an, da die zugehörigen Spalten von  $A$  schon nahe an der Einheitsmatrix sind.

Nr.	Var. →							$x_B \downarrow$	$Q \downarrow$
	$B \downarrow$	$c_B \downarrow$							
1									
2									
3									
		$\Delta \rightarrow$							$\leftarrow z$

Nr.	Var. →							$x_B \downarrow$	$Q \downarrow$
	$B \downarrow$	$c_B \downarrow$							
1									
2									
3									
		$\Delta \rightarrow$							$\leftarrow z$

**Achtung:** Hier muss kontrolliert werden, dass alle Einträge in der Spalte  $x_B$  wirklich  $\geq 0$  sind! Ansonsten wäre der Basisvektor nicht zulässig, und wir müssten es entweder mit einer anderen Basis  $B$  versuchen oder (systematischer) die sogenannte Phase I vorschalten (siehe § 6.5.3).

- (4.) Schreibe die **Namen der Basisvariablen** in die Spalte  $B$  und daneben die **zugehörigen Koeffizienten**  $c_B$  der Zielfunktion.

Nr.	Var. →							$x_B \downarrow$	$Q \downarrow$
	$B \downarrow$	$c_B \downarrow$							
1									
2									
3									
		$\Delta \rightarrow$							$\leftarrow z$

- (5.) Berechne die **Optimalitätsindikatoren**  $\Delta_k := \sum_{j \in B} c_j \tilde{a}_{jk} - c_k$  und trage sie in die mit  $\Delta$  bezeichnete Zeile ein.

Nr.	Var. →							$x_B \downarrow$	$Q \downarrow$
	$B \downarrow$	$c_B \downarrow$							
1									
2									
3									
		$\Delta \rightarrow$							$\leftarrow z$

**Kontrolle:** Bei den Basisvariablen  $k \in B = \{3, 4, 5\}$  muss  $\Delta_k = 0$  stehen.

- (6.) Berechne die **Zielfunktion**  $z = \sum_{j \in B} c_j x_j$  und trage den Wert in die mit  $z$  bezeichnete Zelle ein. Das Simplextableau ist nun für den ersten Simplexschritt vorbereitet.

Nr.	Var. →							$x_B \downarrow$	$Q \downarrow$
	$B \downarrow$	$c_B \downarrow$							
1									
2									
3									
		$\Delta \rightarrow$							$\leftarrow z$

Ende 10. V

07.06.2011

### Durchführung eines Simplexschrittes

- (7.) Führe den **Optimalitätstest** durch: Sind alle  $\Delta_k \geq 0$ ?

- ▷ Wenn ja, dann ist das aktuelle  $\vec{x}_B$  mit  $\vec{x}_N = \vec{0}$  eine optimale Lösung der LOA. **ENDE**
- ▷ Wenn nein, dann wähle (eine Spalte)  $k$  mit  $\Delta_k < 0$ . Im Beispiel kommt nur  $k = 1$  in Frage.

(8.) Führe den **Quotiententest** durch: Berechne dazu

$$Q_j = \frac{\tilde{b}_j}{\tilde{a}_{jk}} \quad \text{für diejenigen } j \in B \text{ mit positiven Matrixeinträgen } \tilde{a}_{jk} > 0.$$

Nr.	Var. →						$x_B \downarrow$	$Q \downarrow$
	$B \downarrow$	$c_B \downarrow$						
1								
2								
3								
		$\Delta \rightarrow$						$\leftarrow z$

- ▷ Wenn in der gewählten Spalte  $k$  alle  $\tilde{a}_{jk} \leq 0$  sind, ist die LOA unbeschränkt und damit nicht lösbar. **ENDE**
- ▷ Ansonsten merke den Index  $r \in B$ , bei dem  $Q_j$  am kleinsten ist. Im Beispiel:  $r = 4$ , da der kleinste Quotient in der zu  $x_4$  gehörenden Zeile auftritt.

(9.) Bereite das Tableau für den nächsten Schritt vor. Die bisherige Basisvariable  $x_k$  wird durch  $x_r$  ersetzt (Spalte  $B$ ). Der daneben stehende Koeffizient aus der Zielfunktion  $c_B$  wird ebenfalls ersetzt. Im Beispiel ist  $B = \{3, 1, 5\}$  die neue Basis. Die Spalte  $Q$  wird gelöscht.

Nr.	Var. →						$x_B \downarrow$	$Q \downarrow$
	$B \downarrow$	$c_B \downarrow$						
1								
2								
3								
		$\Delta \rightarrow$						$\leftarrow z$

(10.) Die Matrix und rechte Seite müssen wie im Vorbereitungsschritt (3.) auf die richtige Form gebracht werden, und zwar passend zur neuen Basis. Im Beispiel:  $B = \{3, 1, 5\}$ .

- ▷ Die 3. Spalte der Matrix muss  $(1, 0, 0)^T$  sein (stimmt noch).
- ▷ Die 1. Spalte der Matrix muss  $(0, 1, 0)^T$  werden.
- ▷ Die 5. Spalte der Matrix muss  $(0, 0, 1)^T$  sein (stimmt noch).

**Beachte:** Es gibt immer nur eine Spalte, die nicht stimmt, nämlich die zum Index  $\boxed{k}$  der neuen Basisvariablen, im Beispiel:  $k = 1$ .

Zunächst bringen wir eine 1 auf die entsprechende Zeile und erzeugen anschließend darüber und darunter Nullen:

	Var. $\rightarrow$								
Nr.	$B \downarrow$	$c_B \downarrow$						$x_B \downarrow$	$Q \downarrow$
1									
2									
3									
		$\Delta \rightarrow$							$\leftarrow z$

	Var. $\rightarrow$								
Nr.	$B \downarrow$	$c_B \downarrow$						$x_B \downarrow$	$Q \downarrow$
1									
2									
3									
		$\Delta \rightarrow$							$\leftarrow z$

	Var. $\rightarrow$								
Nr.	$B \downarrow$	$c_B \downarrow$						$x_B \downarrow$	$Q \downarrow$
1									
2									
3									
		$\Delta \rightarrow$							$\leftarrow z$

**Kontrolle:** Die Basisvariablen (Spalte  $x_B$ ) dürfen keine negativen Werte annehmen.

- (11.) Abschließend müssen wir noch die Optimalitätsindikatoren  $\Delta$  und den Wert der Zielfunktion  $z$  aktualisieren. Dazu können wir diese entweder wie bei den Vorbereitungsschritten (5.) und (6.) neu berechnen, oder wir können die  $\Delta$ -Zeile wie eine zusätzliche Zeile der Matrix behandeln und eine Null in Spalte  $k = 1$  erzeugen:

Nr.	Var. →							$x_B \downarrow$	$Q \downarrow$
	$B \downarrow$	$c_B \downarrow$							
1									
2									
3									
		$\Delta \rightarrow$							$\leftarrow z$

Nr.	Var. →							$x_B \downarrow$	$Q \downarrow$
	$B \downarrow$	$c_B \downarrow$							
1									
2									
3									
		$\Delta \rightarrow$							$\leftarrow z$

**Kontrolle:** Bei den Basisvariablen  $k \in B = \{3, 1, 5\}$  muss wieder  $\Delta_k = 0$  stehen.

**Kontrolle:** Der neue Wert der Zielfunktion (hier  $z = -8$ ) darf nicht kleiner sein als der alte Wert ( $z = -128$ ).

**Durchführung des nächsten Simplexschrittes**

(7'.) Beim **Optimalitätstest** sind wieder nicht alle  $\Delta_k \geq 0$ . Wir müssen einen weiteren Simplexschritt durchführen. Wieder kommt nur eine Spalte  $k = 2$  in Frage.

(8'.) Für den **Quotiententest** berechnen wir wieder

$$Q_j = \frac{\tilde{b}_j}{\tilde{a}_{jk}} \quad \text{für diejenigen } j \in B \text{ mit positiven Matrixeinträgen } \tilde{a}_{jk} > 0.$$

Nr.	Var. →							$x_B \downarrow$	$Q \downarrow$
	$B \downarrow$	$c_B \downarrow$							
1									
2									
3									
		$\Delta \rightarrow$							$\leftarrow z$

Der kleinste Quotient tritt in der Zeile mit  $x_3$  auf, also setzen wir  $r = 3$ .

(9'.) Die bisherige Basisvariable  $x_k$  wird durch  $x_r$  ersetzt (Spalte  $B$ ). Der daneben stehende Koeffizient aus der Zielfunktion  $c_B$  wird ebenfalls ersetzt. Die Spalte  $Q$  wird gelöscht. Im Beispiel ist  $B = \{2, 1, 5\}$  die neue Basis.

Nr.	Var. $\rightarrow$							$x_B \downarrow$	$Q \downarrow$
	$B \downarrow$	$c_B \downarrow$							
1									
2									
3									
		$\Delta \rightarrow$							$\leftarrow z$

(10'.) Die Matrix und rechte Seite müssen wieder auf die richtige Form gebracht werden, passend zur neuen Basis  $B = \{2, 1, 5\}$ .

▷ Die 2. Spalte der Matrix muss  $(1, 0, 0)^\top$  werden.

▷ Die 1. Spalte der Matrix muss  $(0, 1, 0)^\top$  sein (stimmt noch).

▷ Die 5. Spalte der Matrix muss  $(0, 0, 1)^\top$  sein (stimmt noch).

Zunächst bringen wir eine 1 auf die entsprechende Zeile und erzeugen anschließend darunter Nullen:

Nr.	Var. $\rightarrow$							$x_B \downarrow$	$Q \downarrow$
	$B \downarrow$	$c_B \downarrow$							
1									
2									
3									
		$\Delta \rightarrow$							$\leftarrow z$

Nr.	Var. $\rightarrow$							$x_B \downarrow$	$Q \downarrow$
	$B \downarrow$	$c_B \downarrow$							
1									
2									
3									
		$\Delta \rightarrow$							$\leftarrow z$

Nr.	Var. $\rightarrow$							$x_B \downarrow$	$Q \downarrow$
	$B \downarrow$	$c_B \downarrow$							
1									
2									
3									
		$\Delta \rightarrow$							$\leftarrow z$

**Kontrolle:** Die Basisvariablen (Spalte  $x_B$ ) dürfen keine negativen Werte annehmen.

(11'.) Abschließend müssen wir noch die Optimalitätsindikatoren  $\Delta$  und den Wert der Zielfunktion  $z$  aktualisieren.

Nr.	Var. $\rightarrow$						$x_B \downarrow$	$Q \downarrow$
	$B \downarrow$	$c_B \downarrow$						
1								
2								
3								
		$\Delta \rightarrow$						$\leftarrow z$

Nr.	Var. $\rightarrow$						$x_B \downarrow$	$Q \downarrow$
	$B \downarrow$	$c_B \downarrow$						
1								
2								
3								
		$\Delta \rightarrow$						$\leftarrow z$

**Kontrolle:** Bei den Basisvariablen  $k \in B = \{2, 1, 5\}$  muss wieder  $\Delta_k = 0$  stehen.

**Kontrolle:** Der neue Wert der Zielfunktion (hier  $z = -3$ ) darf nicht kleiner sein als der alte Wert ( $z = -8$ ).

(7'.) Dieses Mal sind beim **Optimalitätstest** alle  $\Delta_k \geq 0$ !

Der aktuelle Vektor  $\vec{x}_B$  mit  $\vec{x}_N = \vec{0}$  ist *eine* optimale Lösung (Ecke) der LOA:

$$\vec{x}^* = \begin{pmatrix} 21/5 \\ 1/5 \\ 0 \\ 0 \\ 16 \end{pmatrix}.$$

Der optimale Wert der Zielfunktion ist  $z^* = -3$ .

**Beachte:** Falls die ursprüngliche Aufgabe erst in Normalform gebracht werden musste, müssen wir die Lösung noch zurücktransformieren. Zum Beispiel werden Variablen ohne Vorzeichenbeschränkung wieder als  $x_i = x_i^+ - x_i^-$  zusammengefasst, Schlupfvariablen gestrichen etc., siehe Beispiel 6.7 am Ende. Im gerade gerechneten Beispiel ist das aber nicht notwendig, da die Aufgabe schon in Normalform gestellt war. ◇

## § 6.5 Einige Besonderheiten

### § 6.5.1 Nichteindeutige Lösungen

Wir hatten in Beispiel 6.4 (a) schon gesehen, dass die Lösung einer LOA nicht eindeutig sein muss. Die Lösungsmenge kann auch aus einer ganzen Kante (allgemeiner: Seitenfläche) der zulässigen Menge bestehen.

**Frage:** Wie erkennt man das im Simplex-Verfahren?

Wenn beim Optimalitätstest  $\Delta_k \geq 0$  für alle  $k \in N$  gilt<sup>8</sup>, haben wir eine Lösung gefunden, siehe Schritt (7.) im Beispiel 6.13. Sind alle  $\Delta_k > 0$  für  $k \in N$ , so ist diese Lösung auch eindeutig. Ist jedoch  $\Delta_k = 0$  für ein oder mehrere  $k \in N$ , so könnte man das Verfahren noch einen Schritt weiterführen, zu einer anderen Basis übergehen und evtl. eine neue Lösung erhalten (natürlich mit gleichem Zielfunktionswert).<sup>9</sup>

Das war bei Beispiel 6.13 sogar der Fall! Dort ist am Ende  $B = \{2, 1, 5\}$  und damit  $\Delta_2 = \Delta_1 = \Delta_5 = 0$ , aber auch  $\Delta_4 = 0$ . Führt man das Verfahren noch einen Schritt weiter, kommt man zu einer anderen optimalen Lösung (Ecke):

$$\vec{x}^* = \begin{pmatrix} 121/25 \\ 101/25 \\ 0 \\ 32/5 \\ 0 \end{pmatrix}.$$

Alle Vektoren auf der Verbindungsstrecke, also

$$\vec{x}^* = \alpha \begin{pmatrix} 21/5 \\ 1/5 \\ 0 \\ 0 \\ 16 \end{pmatrix} + (1 - \alpha) \begin{pmatrix} 121/25 \\ 101/25 \\ 0 \\ 32/5 \\ 0 \end{pmatrix} \quad \text{mit } \alpha \in [0, 1]$$

sind damit ebenfalls Lösungen der Aufgabe.

### § 6.5.2 Entartete Basislösungen

Im Simplex-Verfahren sind die Nichtbasisvariablen  $\vec{x}_N$  immer gleich  $\vec{0}$ , die Basisvariablen  $\vec{x}_B$  stets  $\geq \vec{0}$ . Ist jedoch mindestens ein  $x_j = 0$  für ein  $j \in B$ , so nennt man die Basislösung  $\vec{x}$  **entartet**. Unter sehr ungünstigen Umständen kann es dann vorkommen, dass man einige Schritte weiterrechnet, ohne dass sich  $\vec{x}$  ändert. Es ändern sich nur  $B$  und  $N$ , und es kann ein Zyklus entstehen, wobei sich die Folgen von  $B$  und  $N$  ständig wiederholen.<sup>10</sup>

**Abhilfe:** Sind im Optimalitätstest mehrere  $\Delta_k < 0$ , so wähle den kleinstmöglichen Index  $k$  (nicht: das kleinstmögliche  $\Delta_k$ ). Ist im Quotiententest das kleinste  $Q_j$  nicht eindeutig, so wähle  $r$  als den kleinsten der möglichen Indizes. Diese Strategie heißt die **Regel von Bland**. Sie vermeidet Zyklen.

<sup>8</sup>Die zu Basisvariablen gehörenden  $\Delta_k$  sind sowieso null.

<sup>9</sup>Bei gleichzeitiger Entartung der Basislösung kann es auch passieren, dass wir dieselbe Lösung mit anderem  $B$  wieder erhalten.

<sup>10</sup>Ein Beispiel dafür findet sich in [Luderer and Würker, 1997, Abschnitt 5.4].

## § 6.5.3 Phase I

Bei der Vorbereitung des Simplextableaus muss man darauf achten, dass man überhaupt eine *zulässige* Basislösung erhält, also  $\vec{x}_B \geq \vec{0}$ , siehe Beispiel 6.13, Schritt (3.) Man kann verschiedene Basen  $B$  ausprobieren oder aber (systematischer) die sogenannte **Phase I** vorschalten.

Wir gehen wieder von einer LOA in Normalform aus:

$$\text{Maximiere } \vec{c}^\top \vec{x}, \quad \text{wobei } A\vec{x} = \vec{b} \quad \text{sowie} \quad \vec{x} \geq \vec{0}$$

mit  $A \in \mathbb{R}^{m \times n}$ ,  $\vec{b} \in \mathbb{R}^m$  und  $\vec{c} \in \mathbb{R}^n$ ,  $\vec{b} \geq \vec{0}$ . Wir führen vorübergehend künstliche Variablen  $v_1, v_2, \dots, v_m$  ein und untersuchen die folgende Hilfsaufgabe (HA) mit  $n + m$  Variablen:

$$\text{Maximiere } -v_1 - v_2 - \dots - v_m, \quad \text{wobei } A\vec{x} + E\vec{v} = \vec{b} \quad \text{sowie} \quad \vec{x} \geq \vec{0}, \quad \vec{v} \geq \vec{0}.$$

Für die HA steht sofort eine zulässige Basislösung fest, nämlich  $\vec{x} = \vec{0}$  und  $\vec{v} = \vec{b}$ , wobei die  $\vec{v}$  gerade die Basisvariablen sind und die  $\vec{x}$  die Nichtbasisvariablen. Das Simplex-Verfahren für die HA kann also gestartet werden (Phase I).

**Satz 6.14 (Bedeutung der Hilfsaufgabe)**

Die HA besitzt immer eine Lösung, ist also weder unbeschränkt noch unzulässig.

- (a) Ist der Zielfunktionswert der optimalen Lösung der HA  $z^* < 0$ , dann ist die ursprüngliche Aufgabe unzulässig, d.h., die zulässige Menge ist leer.
- (b) Ist der optimale Zielfunktionswert dagegen  $z^* = 0$ , dann gibt es eine zulässige Basislösung der ursprünglichen Aufgabe. Diese erhalten wir aus den  $\vec{x}$ -Komponenten der optimalen Lösung der HA.<sup>11</sup> Die Komponenten  $\vec{v}^*$  sind  $\vec{0}$ . ◇

Ende 11. V

14.06.2011

<sup>11</sup>Evtl. ist noch Basistausch notwendig.

## § 7 Funktionen mehrerer Variabler

### § 7.1 Funktionen $\mathbb{R}^m \rightarrow \mathbb{R}$

Oft hängen Größen nicht nur von einer, sondern von mehreren unabhängigen Variablen ab: Diesen Sachverhalt beschreiben wir durch Funktionen  $f : D(f) \subseteq \mathbb{R}^m \rightarrow \mathbb{R}$ .

**Beispiel 7.1** (a) Die Temperaturverteilung in einem Raum wird durch eine Funktion  $T : \mathbb{R}^3 \rightarrow \mathbb{R}$  beschrieben.  $T(x, y, z)$  bezeichnet die Temperatur am Punkt  $(x, y, z)$ .

(b) Die kinetische Energie (Bewegungsenergie) eines Körpers ist

$$E(m, v) = \frac{m v^2}{2},$$

wobei  $m$  die Masse und  $v$  die Geschwindigkeit bezeichnen.

(c) Die logistische Funktion aus Beispiel 4.3 kann auch als Funktion

$$f(x, a, b, c) = \frac{a}{1 + b e^{-cx}},$$

also  $f : \mathbb{R}^4 \rightarrow \mathbb{R}$  verstanden werden.  $\diamond$

#### § 7.1.1 Darstellungsmöglichkeiten

Funktionen  $f : D(f) \subset \mathbb{R}^2 \rightarrow \mathbb{R}$  können wie folgt grafisch dargestellt werden:

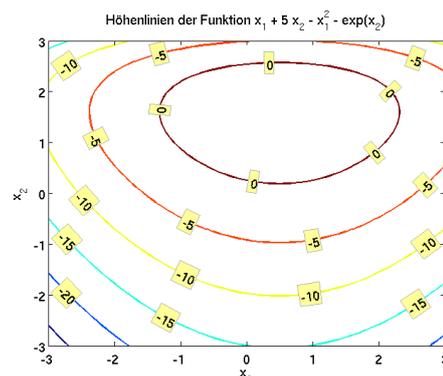
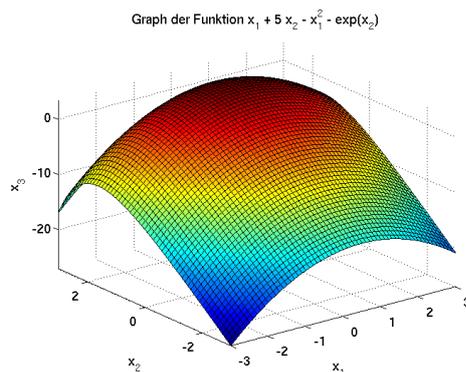
- Die Menge  $\{(x_1, x_2, f(x_1, x_2)) : (x_1, x_2) \in D(f)\} \subseteq \mathbb{R}^3$  heißt der **Graph** von  $f$ .
- Die Menge  $\{(x_1, x_2) \in D(f) : f(x_1, x_2) = c\} \subseteq \mathbb{R}^2$  heißt die **Höhenlinie** oder Niveaulinie von  $f$  zum Niveau  $c \in \mathbb{R}$ .

Liegt der Definitionsbereich von  $f$  im  $\mathbb{R}^m$  mit  $m \geq 2$ , so kann man sich bei der Darstellung mit Schnitten behelfen, bei denen alle bis auf zwei Variablen festgehalten werden.

**Beispiel 7.2** (a) Wir betrachten die Funktion

$$f(x_1, x_2) = x_1 + 5x_2 - x_1^2 - \exp(x_2)$$

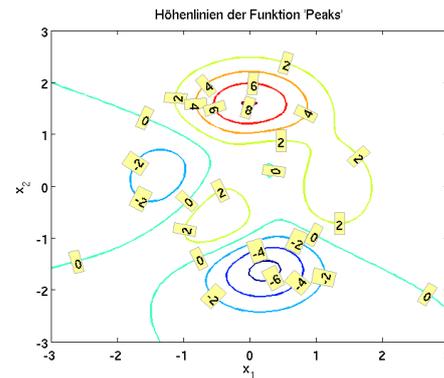
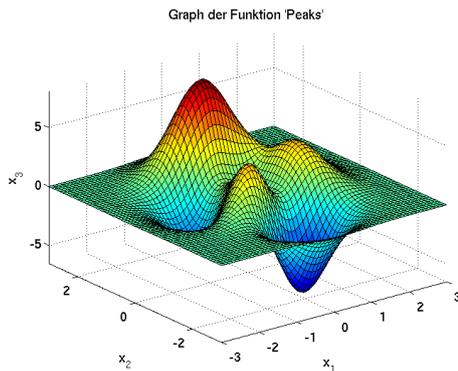
auf der Menge  $D(f) = [-3, 3] \times [-3, 3] \subset \mathbb{R}^2$ .



(b) Wir betrachten die Funktion<sup>12</sup>

$$f_{\text{peaks}}(x_1, x_2) = 3(1 - x_1)^2 \exp(-x_1^2 - (x_2 + 1)^2) - 10\left(\frac{x_1}{5} - x_1^3 - x_2^5\right) \exp(-x_1^2 - x_2^2) - \frac{1}{3} \exp(-(x_1 + 1)^2 - x_2^2)$$

auf der Menge  $D(f) = [-3, 3] \times [-3, 3] \subset \mathbb{R}^2$ .



◇

### § 7.1.2 Grenzwerte und Stetigkeit

**Definition 7.3** (Folgen im  $\mathbb{R}^m$ , vgl. Definition 3.1) (a) Eine Vorschrift, die jedem **Index**  $n \in \mathbb{N}_0$  einen Vektor  $\vec{a}_n \in \mathbb{R}^m$  zuordnet, heißt eine **Folge** in  $\mathbb{R}^m$ .

(b) Eine Folge  $(\vec{a}_n)$  heißt **konvergent** gegen den **Grenzwert**  $\vec{a} \in \mathbb{R}^m$ , wenn gilt: Für jede beliebig kleine Schranke  $\varepsilon > 0$  gibt es einen Index  $n_0 = n_0(\varepsilon) \in \mathbb{N}$ , sodass

$$\|\vec{a}_n - \vec{a}\| < \varepsilon \quad \text{für alle } n \geq n_0$$

gilt. (In jeder noch so kleinen Kugel mit Mittelpunkt  $\vec{a}$  liegen alle hinreichend späten Folgenglieder.)

(c) Man schreibt dann:

$$\lim_{n \rightarrow \infty} \vec{a}_n = \vec{a} \quad \text{oder} \quad \vec{a}_n \xrightarrow{n \rightarrow \infty} \vec{a}.$$

◇

#### Bemerkung 7.4

Konvergenz einer Folge in  $\mathbb{R}^m$  ist äquivalent zur Konvergenz in jeder Koordinate. ◇

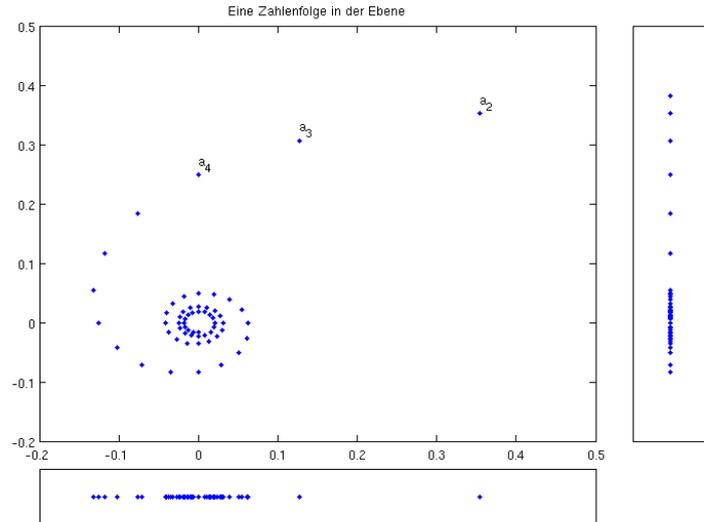
<sup>12</sup>Dies ist die peaks-Funktion in MATLAB.

**Beispiel 7.5**

Die Folge

$$\vec{a}_n = \frac{1}{n} \begin{pmatrix} \cos\left(\frac{8\pi}{n}\right) \\ \sin\left(\frac{8\pi}{n}\right) \end{pmatrix}$$

konvergiert gegen  $(0, 0)^\top$ . In der Darstellung erkennt man, dass auch die Folge der ersten und zweiten Koordinaten von  $(\vec{a}_n)$  konvergieren.



◇

**Definition 7.6** (Grenzwerte und Stetigkeit, vgl. Definition 4.13 und 4.18)

Sei  $f : D(f) \subseteq \mathbb{R}^m \rightarrow \mathbb{R}$  eine Funktion und  $\vec{a} = (a_1, a_2, \dots, a_m) \in D(f)$  ein Punkt im Inneren von  $D(f)$ .<sup>13</sup>

- (a) Falls für *jede* Folge  $(\vec{x}_n)_{n \in \mathbb{N}} \subseteq D(f)$  mit der Eigenschaft  $x_n \neq a$  und  $\lim_{n \rightarrow \infty} \vec{x}_n = \vec{a}$  gilt: die Folge  $(f(\vec{x}_n))_{n \in \mathbb{N}}$  konvergiert gegen  $c \in \mathbb{R}$ , so heißt  $c$  der **Grenzwert** von  $f$  für  $\vec{x}$  gegen  $\vec{a}$ . Man schreibt

$$\lim_{\vec{x} \rightarrow \vec{a}} f(\vec{x}) = c \quad \text{oder} \quad f(\vec{x}) \rightarrow c \quad \text{für} \quad \vec{x} \rightarrow \vec{a}.$$

- (b) Die Funktion  $f$  heißt **stetig** im Punkt  $\vec{a}$ , wenn der Grenzwert  $\lim_{\vec{x} \rightarrow \vec{a}} f(\vec{x}) \in \mathbb{R}$  existiert und gilt:

$$\lim_{\vec{x} \rightarrow \vec{a}} f(\vec{x}) = f(\vec{a}).$$

◇

**Beispiel 7.7** (Polynome im  $\mathbb{R}^m$ )

Analog zu Beispiel 4.9 nennen wir folgende Funktionen  $f : \mathbb{R}^m \rightarrow \mathbb{R}$  **Polynome** im  $\mathbb{R}^m$  vom Grad ...

- 0: konstante Funktionen  $f(\vec{x}) = c$ .
- 1: (affin-)lineare Funktionen

$$f(\vec{x}) = \vec{b} \cdot \vec{x} + c = \vec{b}^\top \vec{x} + c = \sum_{i=1}^m b_i x_i + c.$$

- 2: quadratische Funktionen

$$f(\vec{x}) = \vec{x}^\top A \vec{x} + \vec{b}^\top \vec{x} + c = \sum_{i,j=1}^m a_{ij} x_i x_j + \sum_{i=1}^m b_i x_i + c.$$

◇

<sup>13</sup>Das heißt, dass eine ganze Kugel mit Mittelpunkt  $\vec{a}$ , also  $\{\vec{x} \in \mathbb{R}^m : \|\vec{x} - \vec{a}\| < \varepsilon\}$ , in  $D(f)$  enthalten ist.

Dabei sind  $c \in \mathbb{R}$ ,  $\vec{b} = (b_1, \dots, b_m)^\top \in \mathbb{R}^m$  und  $A = (a_{ij})_{i,j=1}^m \in \mathbb{R}^{m \times m}$ .

**Beachte:** Es kann ohne Einschränkung der Allgemeinheit davon ausgegangen werden, dass  $A = A^\top$  (symmetrisch) ist. Warum?

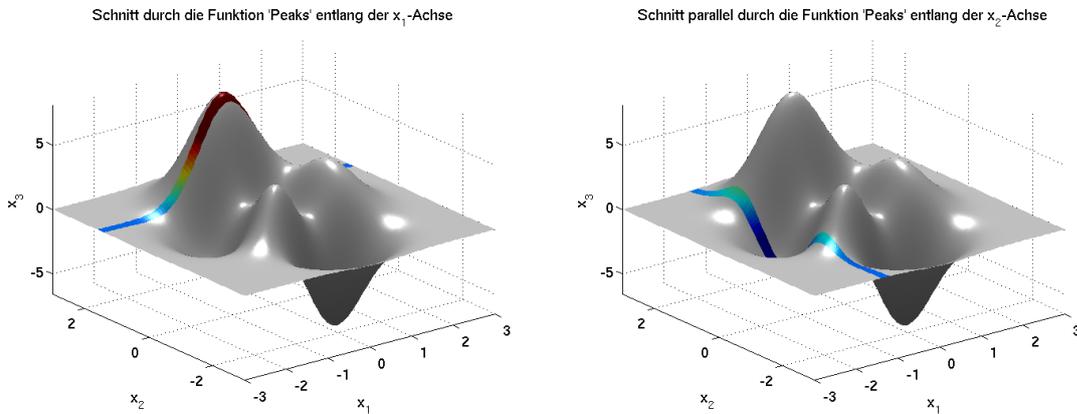
Die o.g. Funktionen sind in jedem Punkt von  $\mathbb{R}^m$  stetig. ◇

### § 7.1.3 Partielle Differenzierbarkeit

Sei  $f : D(f) \subseteq \mathbb{R}^m \rightarrow \mathbb{R}$  eine Funktion. Wir fixieren einen Punkt  $\vec{a} = (a_1, a_2, \dots, a_m) \in D(f)$  im Inneren von  $D(f)$  und betrachten die **partielle Funktion**

$$\mathbb{R} \ni x_1 \mapsto f(x_1, a_2, \dots, a_m) \in \mathbb{R},$$

d.h., den **Schnitt** durch die Funktion  $f$  entlang der  $x_1$ -Achse durch den Punkt  $\vec{a}$ , wobei alle anderen Argumente konstant gehalten werden.



### Definition 7.8 (Partielle Differenzierbarkeit und Richtungsableitung)

- (a) Falls die Funktion  $\mathbb{R} \ni x_1 \mapsto f(x_1, a_2, \dots, a_m) \in \mathbb{R}$  im Punkt  $a_1 \in \mathbb{R}$  diffbar ist, dann heißt die Ableitung

$$\begin{aligned} \frac{\partial f(\vec{a})}{\partial x_1} &= \lim_{x_1 \rightarrow a_1} \frac{f(x_1, a_2, \dots, a_m) - f(a_1, a_2, \dots, a_m)}{x_1 - a_1} \\ &= \lim_{h \rightarrow 0} \frac{f(\vec{a} + h \vec{e}_1) - f(\vec{a})}{h} \end{aligned}$$

die **partielle Ableitung** von  $f$  nach  $x_1$  im Punkt  $\vec{a}$ , und  $f$  heißt **partiell diffbar** nach  $x_1$  im Punkt  $\vec{a}$ . (Hier ist  $\vec{e}_i$  der  $i$ -te Einheitsvektor im  $\mathbb{R}^m$ , siehe Beispiel 2.8.)

- (b) Analog definieren wir die partiellen Ableitungen im Punkt  $\vec{a}$  auch in die anderen Koordinatenrichtungen für  $i = 1, \dots, m$ :

$$\begin{aligned} \frac{\partial f(\vec{a})}{\partial x_i} &= \lim_{x_i \rightarrow a_i} \frac{f(a_1, a_2, \dots, a_{i-1}, x_i, a_{i+1}, \dots, a_m) - f(a_1, a_2, \dots, a_m)}{x_i - a_i} \\ &= \lim_{h \rightarrow 0} \frac{f(\vec{a} + h \vec{e}_i) - f(\vec{a})}{h}. \end{aligned}$$

Die partiellen Ableitungen an der Stelle  $\vec{a}$  werden auch mit den Symbolen

$$\frac{\partial}{\partial x_i} f(\vec{a}) \quad \text{und} \quad f_{x_i}(\vec{a}), \quad i = 1, \dots, m$$

bezeichnet.

- (c) Der (Spalten-)Vektor aus *allen*  $m$  partiellen Ableitungen an einem Punkt  $\vec{a}$  heißt der **Gradient** von  $f$  im Punkt  $\vec{a}$ :

$$\text{grad } f(\vec{a}) = \nabla f(\vec{a}) = \begin{pmatrix} \frac{\partial f(\vec{a})}{\partial x_1} \\ \frac{\partial f(\vec{a})}{\partial x_2} \\ \vdots \\ \frac{\partial f(\vec{a})}{\partial x_m} \end{pmatrix} \quad \text{sprich: „nabla“ für } \nabla$$

- (d) Die Funktion

$$T_1(\vec{x}; \vec{a}) = f(\vec{a}) + \text{grad } f(\vec{a}) \cdot (\vec{x} - \vec{a})$$

heißt die **Tangentialebene** von  $f$  im Punkt  $\vec{a}$ . Sie ist, ebenso wie  $f$ , eine Funktion  $\mathbb{R}^m \rightarrow \mathbb{R}$ . Die Tangentialebene wird auch das **Taylorpolynom 1. Ordnung** (siehe Definition 4.44 für den Fall  $f : \mathbb{R} \rightarrow \mathbb{R}$ ) an  $f$  im Punkt  $\vec{a}$  genannt. Sie ist das beste **lineare Modell** für  $f(\vec{x})$  in der Nähe von  $\vec{a}$ , das man finden kann!

- (e) Sei  $\vec{v} \in \mathbb{R}^m$  mit  $\|\vec{v}\| = 1$  gegeben.<sup>14</sup> Der Grenzwert

$$\frac{\partial f(\vec{a})}{\partial \vec{v}} = \lim_{h \rightarrow 0} \frac{f(\vec{a} + h\vec{v}) - f(\vec{a})}{h}$$

heißt die **Richtungsableitung** von  $f$  im Punkt  $\vec{a}$  in Richtung  $\vec{v}$ .  $\diamond$

Ende 12. V

21.06.2011

**Bemerkung 7.9 (zur Richtungsableitung)** (a) Die partiellen Ableitungen sind gerade die Richtungsableitungen in Richtung der Einheitsvektoren:

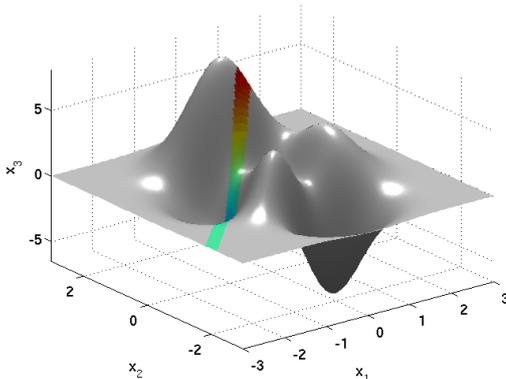
$$\frac{\partial f(\vec{a})}{\partial x_i} = \frac{\partial f(\vec{a})}{\partial \vec{e}_i}.$$

- (b) Die Richtungsableitung  $\frac{\partial f(\vec{a})}{\partial \vec{v}}$  ist die Ableitung der eindimensionalen „Schnittfunktion“

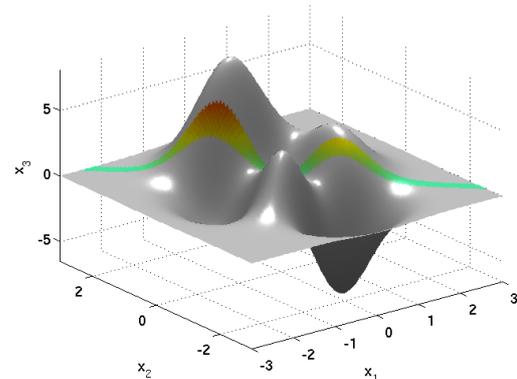
$$\mathbb{R} \ni t \mapsto f(\vec{a} + t\vec{v}) \in \mathbb{R}$$

an der Stelle  $t = 0$ .  $\diamond$

Schnitt parallel durch die Funktion 'Peaks' entlang eines Vektors



Schnitt parallel durch die Funktion 'Peaks' entlang eines Vektors



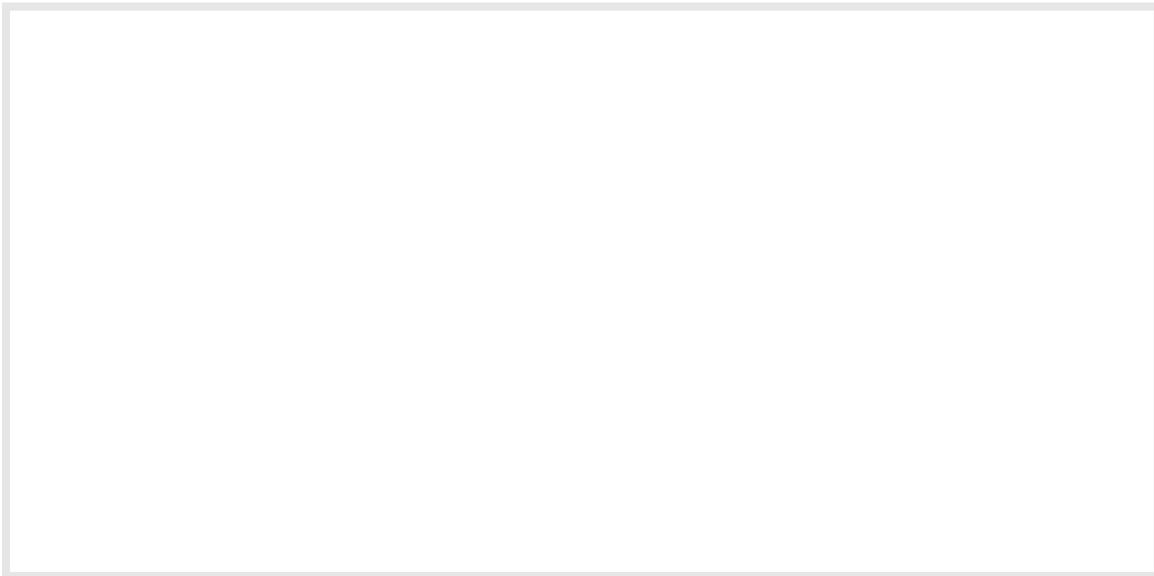
<sup>14</sup>Die Bedingung  $\|\vec{v}\| = 1$  wird auch oft fallengelassen.

**Beispiel 7.10 (Partielle Ableitungen)**

Was sind die partiellen Ableitungen der logistischen Funktion

$$f(x, a, b, c) = \frac{a}{1 + b e^{-cx}}$$

aus Beispiel 7.1? Für die partielle Ableitung  $\partial f(x, a, b, c)/\partial x$  denken wir uns  $(a, b, c)$  konstant und leiten nach  $x$  ab, analog für die anderen partiellen Ableitungen:

**Beispiel 7.11 (Partielle Ableitungen)**

Wir berechnen die partiellen Ableitungen der Funktion aus Beispiel 7.2:

$$f(x_1, x_2) = 3(1 - x_1)^2 \exp(-x_1^2 - (x_2 + 1)^2) - 10\left(\frac{x_1}{5} - x_1^3 - x_2^5\right) \exp(-x_1^2 - x_2^2) - \frac{1}{3} \exp(-(x_1 + 1)^2 - x_2^2).$$

Für die partielle Ableitung  $\partial f(x_1, x_2)/\partial x_1$  denken wir uns  $x_2$  konstant und leiten nach  $x_1$  ab:

$$\begin{aligned} \frac{\partial f(x_1, x_2)}{\partial x_1} &= -6(1 - x_1) \exp(-x_1^2 - (x_2 + 1)^2) + 3(1 - x_1)^2 \exp(-x_1^2 - (x_2 + 1)^2)(-2x_1) \\ &\quad - 10\left(\frac{1}{5} - 3x_1^2\right) \exp(-x_1^2 - x_2^2) - 10\left(\frac{x_1}{5} - x_1^3 - x_2^5\right) \exp(-x_1^2 - x_2^2)(-2x_1) \\ &\quad - \frac{1}{3} \exp(-(x_1 + 1)^2 - x_2^2)(-2(x_1 + 1)). \end{aligned}$$

Für die partielle Ableitung  $\partial f(x_1, x_2)/\partial x_2$  denken wir uns  $x_1$  konstant und leiten nach  $x_2$  ab:

$$\begin{aligned} \frac{\partial f(x_1, x_2)}{\partial x_2} &= 3(1 - x_1)^2 \exp(-x_1^2 - (x_2 + 1)^2)(-2(x_2 + 1)) \\ &\quad - 10(-5x_2^4) \exp(-x_1^2 - x_2^2) - 10\left(\frac{x_1}{5} - x_1^3 - x_2^5\right) \exp(-x_1^2 - x_2^2)(-2x_2) \\ &\quad - \frac{1}{3} \exp(-(x_1 + 1)^2 - x_2^2)(-2x_2). \end{aligned}$$

◇

**Beispiel 7.12 (Tangentialebene)**

Wir berechnen Funktionswert und Gradienten der Funktion aus Beispiel 7.11 an der Stelle  $\vec{a} = (-1, 0; 1, 4)$ :

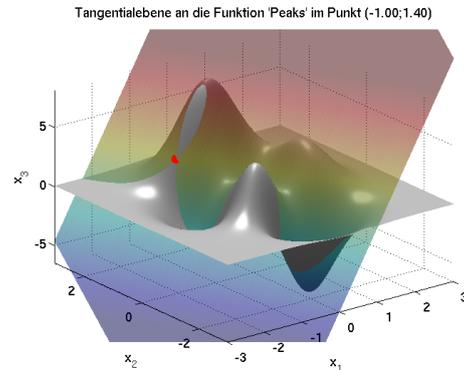
$$f(\vec{a}) = \dots = 2,3394,$$

$$\text{grad } f(\vec{a}) = \begin{pmatrix} \frac{\partial f(\vec{a})}{\partial x_1} \\ \frac{\partial f(\vec{a})}{\partial x_2} \end{pmatrix} = \dots = \begin{pmatrix} 6,2096 \\ 3,3754 \end{pmatrix}.$$

Mit diesen Daten können wir die Tangentialebene

$$T_1(\vec{x}; \vec{a}) = f(\vec{a}) + \text{grad } f(\vec{a}) \cdot (\vec{x} - \vec{a})$$

angeben.



◇

**§ 7.1.4 Differenzierbarkeit****Definition 7.13**

Sei  $f : D(f) \subseteq \mathbb{R}^m \rightarrow \mathbb{R}$  eine Funktion und  $\vec{a} = (a_1, a_2, \dots, a_m) \in D(f)$  ein Punkt im Inneren von  $D(f)$ . Die Funktion  $f$  heißt **total diffbar** oder einfach **diffbar** im Punkt  $\vec{a}$ , wenn alle partiellen Ableitungen im Punkt  $\vec{a}$  existieren und wenn die Tangentialebene die Funktion  $f(\vec{x})$  in der Nähe von  $\vec{a}$  so gut approximiert, dass für das Restglied

$$r(\vec{x}; \vec{a}) := f(\vec{x}) - \underbrace{(f(\vec{a}) + \text{grad } f(\vec{a}) (\vec{x} - \vec{a}))}_{T_1(\vec{x}; \vec{a})}$$

gilt:

$$\lim_{\vec{x} \rightarrow \vec{a}} \frac{r(\vec{x}; \vec{a})}{\|\vec{x} - \vec{a}\|} = 0.$$

In diesem Fall heißt  $\text{grad } f(\vec{a})$  die **Ableitung** von  $f$  an der Stelle  $\vec{a}$ .

◇

**Bemerkung 7.14**

Diese Definition erweitert die Definition für den Fall  $f : D(f) \subseteq \mathbb{R} \rightarrow \mathbb{R}$ , vgl. Bemerkung nach Definition 4.22.

◇

**Frage:** Wie kann man Diffbarkeit einer Funktion feststellen, ohne die Definition zu überprüfen?

**Satz 7.15 (Ein hinreichendes und ein notwendiges Kriterium für Diffbarkeit)**

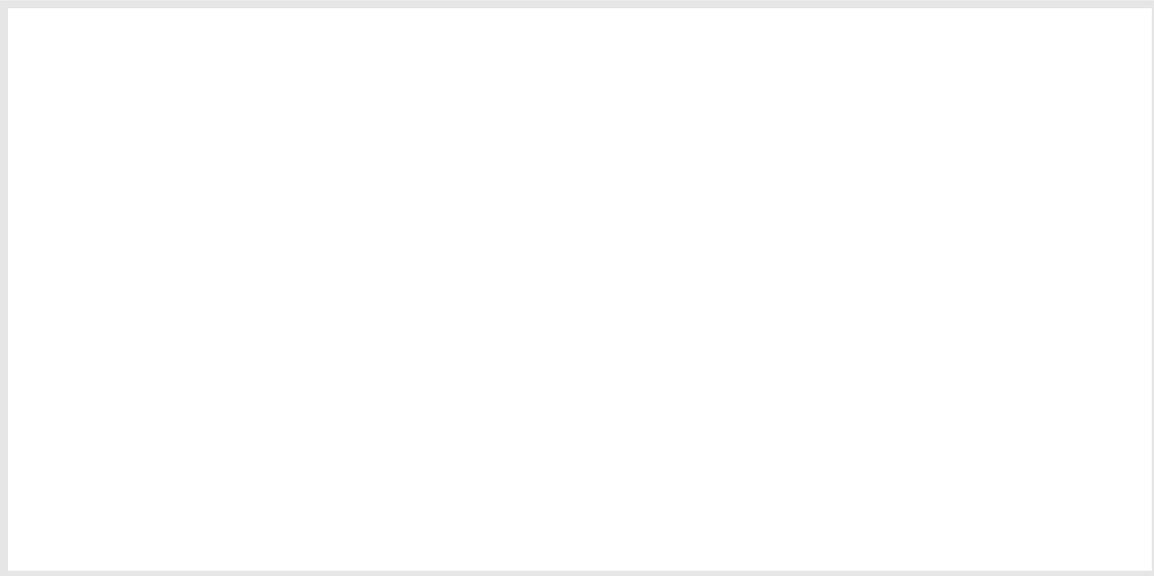
Seien  $f$  und  $\vec{a}$  wie in Definition 7.13.

(a) Falls die partiellen Ableitungen  $\frac{\partial f}{\partial x_i}$  alle in einer Umgebung von  $\vec{a}$  existieren und dort stetig sind, dann ist  $f$  total diffbar im Punkt  $\vec{a}$ .

(b) Falls  $f$  in  $\vec{a}$  total diffbar ist, dann ist  $f$  dort auch stetig.

◇

**Beispiel 7.16 (Differenzierbare Funktionen)**



**Satz 7.17 (Berechnung der Richtungsableitung)**

Seien  $f$  und  $\vec{a}$  wie in Definition 7.13. Falls  $f$  im Punkt  $\vec{a}$  diffbar ist, dann existieren die Richtungsableitungen an der Stelle  $\vec{a}$  in alle Richtungen  $\vec{v}$  mit  $\|\vec{v}\| = 1$ , und es gilt

$$\frac{\partial f(\vec{a})}{\partial \vec{v}} = \text{grad } f(\vec{a}) \cdot \vec{v}. \quad \diamond$$

**Beachte:** Der Gradient  $\text{grad } f(\vec{a})$  erlaubt also die Berechnung *aller* Richtungsableitungen, ohne dass man die Definition 7.8 (e) bemühen muss.

**Frage:** In welche Richtung  $\vec{v}$  auf der Einheitskugel ( $\|\vec{v}\| = 1$ ) ist die Richtungsableitung am größten?

**Folgerung 7.18 (Bedeutung des Gradienten)**

Die Richtungsableitung ist das Skalarprodukt zwischen dem Gradienten  $\text{grad } f(\vec{a})$  und der Richtung  $\vec{v}$ , daher gilt (Definition 2.18)

$$\frac{\partial f(\vec{a})}{\partial \vec{v}} = \text{grad } f(\vec{a}) \cdot \vec{v} = \|\text{grad } f(\vec{a})\| \underbrace{\|\vec{v}\|}_{=1} \cos \varphi.$$

Die Richtungsableitung ist demnach am größten, wenn  $\varphi = 0$ , also  $\vec{v}$  parallel zu  $\text{grad } f(\vec{a})$  ist. Daher sagt man auch: Der Gradient ist die **Richtung des größten** bzw. **steilsten Anstiegs** einer Funktion. Entsprechend ist  $-\text{grad } f(\vec{a})$  die Richtung des **steilsten Abstiegs**.  $\diamond$

**Satz 7.19 (Rechenregeln für die Ableitung, vgl. Satz 4.27)**

Es seien  $f : \mathbb{R}^m \rightarrow \mathbb{R}$  und  $g : \mathbb{R}^m \rightarrow \mathbb{R}$  und  $\vec{s} : \mathbb{R} \rightarrow \mathbb{R}^m$  diffbar. Es sei  $\vec{a} \in \mathbb{R}^m$  und  $t \in \mathbb{R}$ . Dann existieren folgende Ableitungen:

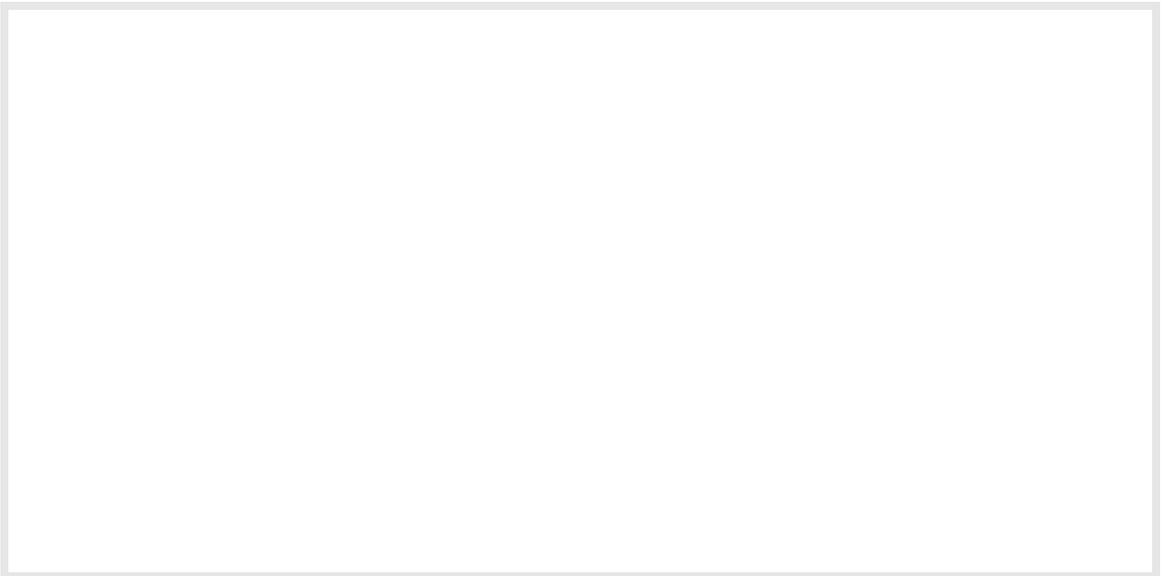
- (a)  $\text{grad}(f \pm g)(\vec{a}) = \text{grad } f(\vec{a}) \pm \text{grad } g(\vec{a})$
- (b)  $\text{grad}(cf)(\vec{a}) = c \text{grad } f(\vec{a})$  für alle  $c \in \mathbb{R}$
- (c)  $\text{grad}(fg)(\vec{a}) = \text{grad } f(\vec{a}) g(\vec{a}) + f(\vec{a}) \text{grad } g(\vec{a})$
- (d)  $(f \circ \vec{s})'(t) = \text{grad } f(\vec{s}(t)) \cdot \vec{s}'(t)$  (Kettenregel)  $\diamond$

Ende 13. V

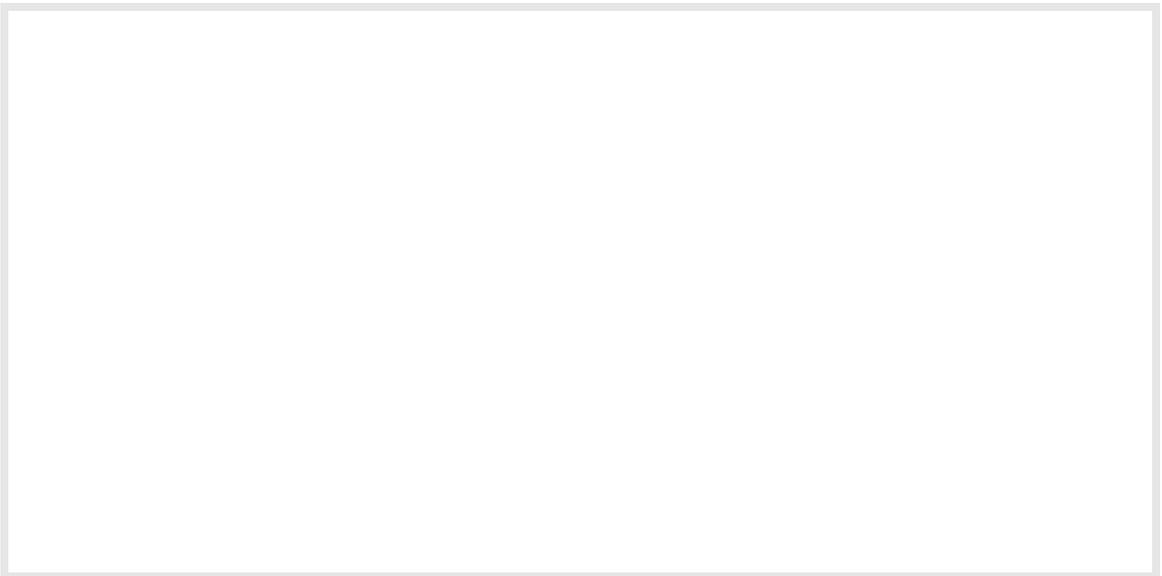
28.06.2011

**Beispiel 7.20** (Anwendungen dieser Rechenregeln)

(a)



(b)



(c) Durch  $x(r, \varphi) = r \cos \varphi$  und  $y(r, \varphi) = r \sin \varphi$  wird die sogenannte **Polarkoordinaten-Transformation** definiert. Dadurch wird eine Funktion  $f(x, y)$  auf Polarkoordinaten transformiert:

$$F(r, \varphi) := f(x(r, \varphi), y(r, \varphi)).$$

---

### § 7.2 Anwendung: Lineare Fehlerrechnung

Wir verallgemeinern die Betrachtungen zur linearen Fehlerrechnung (§ 4.4.2) für Funktionen mehrerer Variabler,  $f: \mathbb{R}^m \rightarrow \mathbb{R}$ . Mit Hilfe der linearen Approximation (Tangentialebene) erhält man

$$\underbrace{\Delta f}_{\text{Fehler im Ergebnis}} := f(\vec{x} + \Delta \vec{x}) - f(\vec{x}) \approx \text{grad } f(\vec{x}) \cdot \underbrace{\Delta \vec{x}}_{\text{Fehler in } \vec{x}} = \sum_{i=1}^m \frac{\partial}{\partial x_i} f(\vec{x}) \Delta x_i.$$

Der Fehler in  $\vec{x}$  (z.B. Messdaten) wird also mit dem Faktor  $\text{grad } f(\vec{x})$  skalar multipliziert. Der Gradient  $\text{grad } f(\vec{x})$  gibt damit die Empfindlichkeit (**Sensitivität**) von  $f$  bzgl. Störungen in  $\vec{x}$  an. Jeder Eintrag  $\frac{\partial}{\partial x_i} f(\vec{x})$  in diesem Vektor steht für die Empfindlichkeit von  $f$  bzgl. Störungen in der  $i$ -ten Komponente von  $\vec{x}$ .

In der Regel wird man Toleranzen  $|\Delta x_i|$  für jedes  $x_i$  einzeln kennen. Dann hilft die Abschätzung (Dreiecksungleichung, Satz 1.22)

$$|\Delta f| \approx \left| \sum_{i=1}^m \frac{\partial}{\partial x_i} f(\vec{x}) \Delta x_i \right| \leq \sum_{i=1}^m \left| \frac{\partial}{\partial x_i} f(\vec{x}) \right| |\Delta x_i|,$$

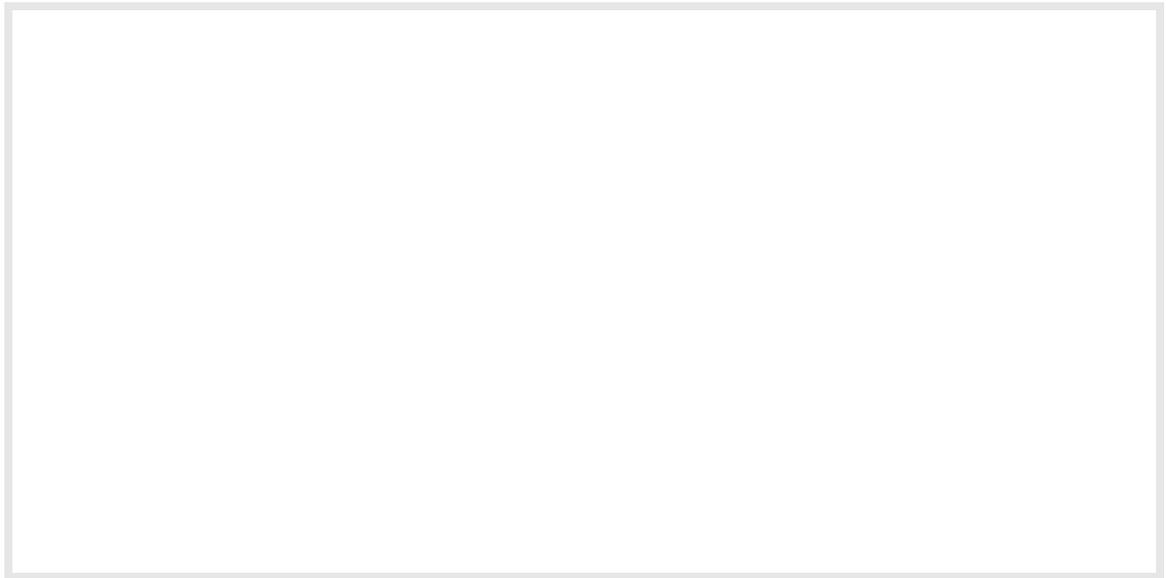
um den Gesamtfehler im Ergebnis einzugrenzen.

#### Beispiel 7.21 (Lineare Fehlerrechnung)

Wir verwenden wieder das Beispiel 4.32, bei dem auf einer Messstrecke der Länge  $s$  aus der Durchfahrzeit  $t$  eines Autos die Geschwindigkeit

$$v = \frac{s}{t}$$

ermittelt wird. Dieses Mal betrachten wir sowohl die Zeitmessung als unsicher (Fehler  $\Delta t$ ) als auch die Länge der Messstrecke (Fehler  $\Delta s$ ). Wie wirken sich die Fehler auf das Ergebnis aus?



### § 7.3 Optimierung differenzierbarer Funktionen

Ähnlich wie Definition 4.39 können wir für Funktionen  $f : \mathbb{R}^m \rightarrow \mathbb{R}$  lokale und globale Maxima und Minima definieren. Das Analogon zu Satz 4.40 ist:

#### Satz 7.22

Sei  $f : \mathbb{R}^m \rightarrow \mathbb{R}$  diffbar, und die partiellen Ableitungen seien wieder stetige Funktionen. Dann gilt:

$$\vec{x}_0 \in \mathbb{R}^m \text{ ist lokale Extremstelle} \quad \Rightarrow \quad \text{grad } f(\vec{x}_0) = \vec{0},$$

d.h.,  $f$  hat an der Stelle  $\vec{x}_0$  eine waagerechte Tangentialebene. ◇

**Beachte:**  $\text{grad } f(\vec{x}_0) = \vec{0}$  ist wie schon im Fall  $f : \mathbb{R} \rightarrow \mathbb{R}$  (Satz 4.40) wieder nur *notwendig* für das Vorliegen einer Extremstelle. Man nennt einen Punkt  $\vec{x}_0$ , an dem  $\text{grad } f(\vec{x}_0) = \vec{0}$  gilt, einen **stationären Punkt** der Funktion  $f(\vec{x})$ . Die Entscheidung, ob bei  $\vec{x}_0$  tatsächlich ein lokales Minimum oder Maximum vorliegt, kann man oft mit Hilfe der zweiten Ableitungen fällen.

#### Definition 7.23 (Zweite Ableitungen)

Sei  $f : \mathbb{R}^m \rightarrow \mathbb{R}$  eine Funktion.

- (a) Wenn die partielle Ableitung  $\frac{\partial f}{\partial x_i}$  in der Umgebung eines Punktes  $\vec{a} \in \mathbb{R}^m$  existiert und diese Funktion im Punkt  $\vec{a}$  wiederum partiell diffbar nach  $x_j$  ist, so bezeichnet man diese **partielle Ableitung zweiter Ordnung** mit

$$f_{x_i x_j}(\vec{a}) = \frac{\partial}{\partial x_j} \left( \frac{\partial f(\vec{a})}{\partial x_i} \right) = \frac{\partial^2}{\partial x_j \partial x_i} f(\vec{a}).$$

**Lies:** Ableitung erst nach  $x_i$ , dann nach  $x_j$ .

- (b) Insgesamt gibt es  $m^2$  verschiedene partielle Ableitungen zweiter Ordnung, die man in der **Hesse-Matrix** der Funktion  $f$  anordnet:

$$H(\vec{a}) = \nabla^2 f(\vec{a}) = \begin{pmatrix} \frac{\partial^2 f(\vec{a})}{\partial x_1^2} & \frac{\partial^2 f(\vec{a})}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f(\vec{a})}{\partial x_1 \partial x_m} \\ \frac{\partial^2 f(\vec{a})}{\partial x_2 \partial x_1} & \frac{\partial^2 f(\vec{a})}{\partial x_2^2} & \cdots & \frac{\partial^2 f(\vec{a})}{\partial x_2 \partial x_m} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f(\vec{a})}{\partial x_m \partial x_1} & \frac{\partial^2 f(\vec{a})}{\partial x_m \partial x_2} & \cdots & \frac{\partial^2 f(\vec{a})}{\partial x_m^2} \end{pmatrix}. \quad \diamond$$

**Achtung:** Manchmal wird in der Definition die Reihenfolge von  $x_i$  und  $x_j$  vertauscht. Dies spielt jedoch in der Regel keine Rolle, denn es gilt:

**Satz 7.24 (Satz von Schwarz)**

Wenn alle zweiten partiellen Ableitungen einer Funktion  $f$  in der Umgebung eines Punktes  $\vec{a}$  existieren und diese Funktionen dort alle wieder stetig sind, dann gilt:

$$\frac{\partial^2 f(\vec{a})}{\partial x_j \partial x_i} = \frac{\partial^2 f(\vec{a})}{\partial x_i \partial x_j},$$

d.h., die Hesse-Matrix  $H(\vec{a})$  ist symmetrisch. ◇

Mit Hilfe der Hesse-Matrix kann man das **Taylorpolynom 2. Ordnung** (vgl. Definition 4.44 für den Fall  $f : \mathbb{R} \rightarrow \mathbb{R}$ ) an die Funktion  $f$  im Punkt  $\vec{a}$  definieren:

$$T_2(\vec{x}; \vec{a}) = \underbrace{f(\vec{a}) + \text{grad } f(\vec{a}) \cdot (\vec{x} - \vec{a})}_{=T_1(\vec{x}; \vec{a})} + \frac{1}{2}(\vec{x} - \vec{a})^\top H(\vec{a})(\vec{x} - \vec{a}).$$

Dies ist das beste **quadratische Modell** für  $f(\vec{x})$  in der Nähe von  $\vec{a}$ , das man finden kann.

**Definition 7.25**

Sei  $A \in \mathbb{R}^{n \times n}$  eine symmetrische Matrix.  $A$  heißt ...

- (a) **positiv (semi-)definit**, wenn  $A$  ausschließlich positive (nicht-negative) Eigenwerte besitzt.
- (b) **negativ (semi-)definit**., wenn  $A$  ausschließlich negative (nicht-positive) Eigenwerte besitzt.
- (c) **indefinit**, wenn  $A$  mindestens einen positiven und mindestens einen negativen Eigenwert besitzt. ◇

**Satz 7.26 (Extremwerttest)**

Die Funktion  $f : \mathbb{R}^m \rightarrow \mathbb{R}$  erfülle die Voraussetzungen von Satz 7.24, und es sei  $\text{grad } f(\vec{a}) = \vec{0}$  (also  $\vec{a}$  ein stationärer Punkt).

- (a) Wenn die Hesse-Matrix  $H(\vec{a})$  nur positive Eigenwerte  $\lambda_i > 0$  hat (also positiv definit ist), dann ist  $\vec{a}$  ein lokales Minimum.
- (b) Wenn die Hesse-Matrix  $H(\vec{a})$  nur negative Eigenwerte  $\lambda_i < 0$  hat (also negativ definit ist), dann ist  $\vec{a}$  ein lokales Maximum.

- (c) Wenn die Hesse-Matrix  $H(\vec{a})$  sowohl positive als auch negative Eigenwerte  $\lambda_i < 0$  hat (also indefinit ist), dann ist  $\vec{a}$  ein sogenannter **Sattelpunkt** (keine Extremstelle!).

In allen anderen Fällen ist auf diesem Weg keine Entscheidung möglich.  $\diamond$

**Folgerung 7.27** (Spezialfall für Funktionen  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ )

Im Fall  $m = 2$  kann man sich die Berechnung der Eigenwerte ersparen, denn man kann ihre Vorzeichen auch einfacher ermitteln:

$$(a) \det(H(\vec{a})) > 0 \text{ und } f_{x_1x_1}(\vec{a}) > 0 \Rightarrow \begin{aligned} &H(\vec{a}) \text{ ist positiv definit} \\ &\Rightarrow \vec{a} \text{ ist lokales Minimum.} \end{aligned}$$

$$(b) \det(H(\vec{a})) > 0 \text{ und } f_{x_1x_1}(\vec{a}) < 0 \Rightarrow \begin{aligned} &H(\vec{a}) \text{ ist negativ definit} \\ &\Rightarrow \vec{a} \text{ ist lokales Maximum.} \end{aligned}$$

$$(c) \det(H(\vec{a})) < 0 \Rightarrow \vec{a} \text{ ist Sattelpunkt.} \quad \diamond$$

**Beachte:**  $f_{x_1x_1}(\vec{a})$  ist der Eintrag oben links in der Hesse-Matrix. Für die Determinante siehe Definition 2.46.

**Beispiel 7.28** (Extremwerte von Funktionen  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ )

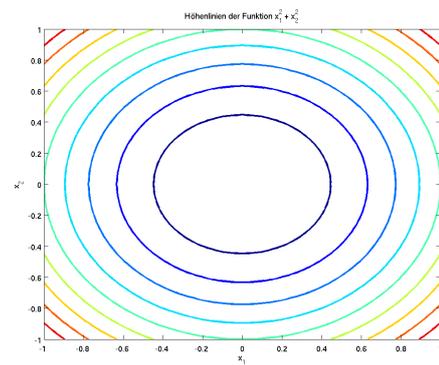
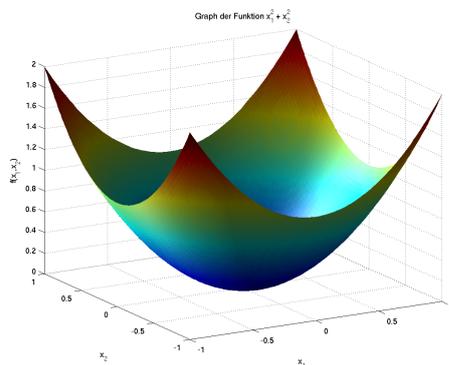
- (a) Gegeben ist die Funktion

$$f(x_1, x_2) = x_1^2 + x_2^2 \Rightarrow \text{grad } f(x_1, x_2) = 2 \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}.$$

Einzigster stationärer Punkt ist  $\vec{a} = (0, 0)^\top$ . Die Hesse-Matrix dort ist

$$H(\vec{a}) = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} \quad \text{mit} \quad \det(H(\vec{a})) = 4.$$

Nach Folgerung 7.27 ist  $H(\vec{a})$  positiv definit, also ist  $\vec{a} = (0, 0)^\top$  ein lokales Minimum.



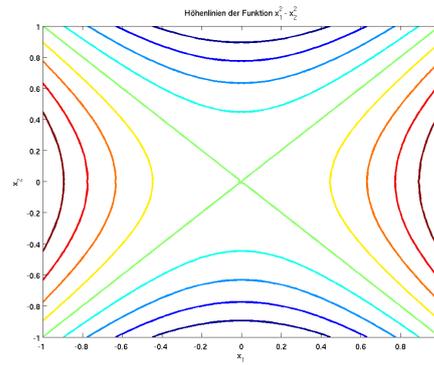
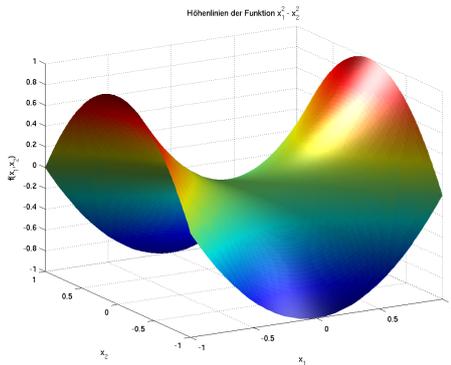
- (b) Gegeben ist die Funktion

$$f(x_1, x_2) = x_1^2 - x_2^2 \Rightarrow \text{grad } f(x_1, x_2) = 2 \begin{pmatrix} x_1 \\ -x_2 \end{pmatrix}.$$

Einzigster stationärer Punkt ist wieder  $\vec{a} = (0, 0)^\top$ . Die Hesse-Matrix dort ist

$$H(\vec{a}) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix} \quad \text{mit} \quad \det(H(\vec{a})) = -4.$$

Nach Folgerung 7.27 ist  $\vec{a} = (0, 0)^T$  ein Sattelpunkt, also keine Extremstelle.



Ende 14. V  
05.07.2011

§ 7.4 Anwendung: Ausgleichsrechnung

**Frage:** Wie kann man eine Modellfunktion mit unbekanntem Parametern an ermittelte Messdaten anpassen?

**Beispiel 7.29 (Bremsverzögerung: Aufgabenstellung<sup>15</sup>)**

Ein Fahrzeug wird von einer Geschwindigkeit  $v$  bis zum Stillstand abgebremst. Bei gleichmäßiger Verzögerung kann für den Bremsweg  $y$  eine quadratische Abhängigkeit

$$y = a v^2 + b v + c.$$

angenommen werden. Zur Bestimmung der unbekanntem Modellparameter  $(a, b, c)$  werden Messungen durchgeführt:

$v$	9	17	17	25	35
$y$	3	9	5	14	23

Da mehr Messungen durchgeführt wurden (5), als Parameter zu identifizieren sind (3), erhalten wir aus dem Modellansatz ein **überbestimmtes System** linearer Gleichungen für die unbekanntem Parameter:

$$\begin{aligned}
 3 &= 9^2 a + 9 b + c \\
 9 &= 17^2 a + 17 b + c \\
 5 &= 17^2 a + 17 b + c \\
 14 &= 25^2 a + 25 b + c \\
 23 &= 35^2 a + 35 b + c
 \end{aligned}
 \quad \text{oder kurz:} \quad
 \underbrace{\begin{pmatrix} 81 & 9 & 1 \\ 289 & 17 & 1 \\ 289 & 17 & 1 \\ 625 & 25 & 1 \\ 1225 & 35 & 1 \end{pmatrix}}_A
 \underbrace{\begin{pmatrix} a \\ b \\ c \end{pmatrix}}_{\vec{x}}
 =
 \underbrace{\begin{pmatrix} 3 \\ 9 \\ 5 \\ 14 \\ 23 \end{pmatrix}}_{\vec{b}}.$$

Dieses LGS ist natürlich unlösbar. Stattdessen suchen wir nach Werten  $(a, b, c)$  für die Parameter, sodass die Gleichungen in einem gewissen Sinne bestmöglich erfüllt sind. ◇

Gegeben sei ein überbestimmtes LGS

$$A\vec{x} = \vec{b}, \quad \text{mit } A \in \mathbb{R}^{m \times n}, \quad \vec{b} \in \mathbb{R}^m, \quad m > n.$$

<sup>15</sup>aus [Meyberg and Vachenaer, 2001, Kapitel 7, § 3.5]

Dabei steht  $m$  für die Anzahl der Messwerte und  $n$  für die Anzahl der zu identifizierenden Parameter. Da dieses LGS i.A. unlösbar ist, suchen wir nach einem Vektor  $\vec{x} \in \mathbb{R}^n$ , der die Gleichungen **im quadratischen Mittel** bestmöglich erfüllt:

$$\text{Minimiere } \frac{1}{2} \|A\vec{x} - \vec{b}\|^2, \quad \vec{x} \in \mathbb{R}^n.$$

Eine solche Aufgabe nennt man ein **lineares Ausgleichsproblem**. Der Faktor  $1/2$  dient allein der Bequemlichkeit; er kürzt sich beim Ableiten weg.

Zur Lösung der Optimierungsaufgabe berechnen wir zunächst die erste Ableitung (den Gradienten) der zu minimierenden Funktion

$$f(\vec{x}) := \frac{1}{2} \|A\vec{x} - \vec{b}\|^2 = \frac{1}{2} \sum_{i=1}^m \underbrace{(\vec{a}^i \cdot \vec{x} - b_i)}_{r_i(\vec{x})}^2,$$

wobei  $\vec{a}^1, \dots, \vec{a}^m$  die Zeilenvektoren von  $A$  sind. Die Zahlen  $r_i(\vec{x})$  heißen die **Residuen** der Aufgabe. Jedes Residuum gehört zu einem der  $m$  Messwerte. Es gibt an, wie weit die Vorhersage durch die Modellfunktion (bei Modellparametern  $\vec{x}$ ) den jeweiligen Messwert verfehlt.

Der Gradient von  $f(\vec{x})$  lässt sich mit Satz 7.19 (a) und (c) bestimmen:

$$\begin{aligned} \text{grad } f(\vec{x}) &= \frac{1}{2} \sum_{i=1}^m \vec{a}^i (\vec{a}^i \cdot \vec{x} - b_i) + (\vec{a}^i \cdot \vec{x} - b_i) \vec{a}^i \\ &= \sum_{i=1}^m \vec{a}^i (\vec{a}^i \cdot \vec{x} - b_i) \\ &= A^\top (A\vec{x} - \vec{b}). \end{aligned}$$

Stationäre Punkte erfüllen also

$$A^\top (A\vec{x} - \vec{b}) = \vec{0} \quad \text{oder} \quad A^\top A\vec{x} = A^\top \vec{b}.$$

Dieses ist die sogenannte **Normalengleichung** des linearen Ausgleichsproblems, sie ist ein lineares Gleichungssystem der Dimension  $n \times n$ .

### Beispiel 7.30 (Bremsverzögerung: Normalengleichung)

Im Beispiel 7.29 erhalten wir als Normalengleichung das LGS

$$\underbrace{\begin{pmatrix} 81 & 289 & 289 & 625 & 1225 \\ 9 & 17 & 17 & 25 & 35 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix}}_{A^\top} \underbrace{\begin{pmatrix} 81 & 9 & 1 \\ 289 & 17 & 1 \\ 289 & 17 & 1 \\ 625 & 25 & 1 \\ 1225 & 35 & 1 \end{pmatrix}}_A \underbrace{\begin{pmatrix} a \\ b \\ c \end{pmatrix}}_{\vec{x}} = \underbrace{\begin{pmatrix} 81 & 289 & 289 & 625 & 1225 \\ 9 & 17 & 17 & 25 & 35 \\ 1 & 1 & 1 & 1 & 1 \end{pmatrix}}_{A^\top} \underbrace{\begin{pmatrix} 3 \\ 9 \\ 5 \\ 14 \\ 23 \end{pmatrix}}_{\vec{b}}$$

oder zusammengefasst

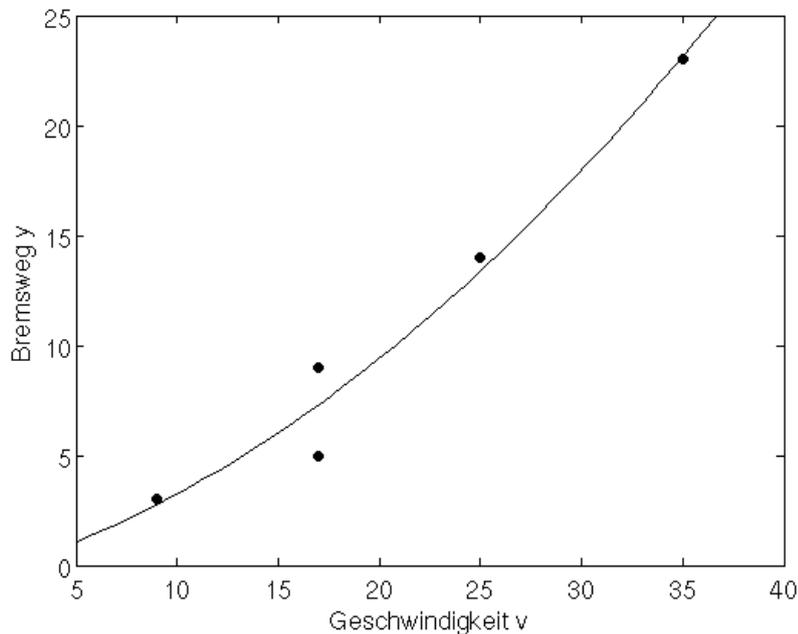
$$\begin{pmatrix} 2064853 & 69055 & 2509 \\ 69055 & 2509 & 103 \\ 2509 & 103 & 5 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 41214 \\ 1420 \\ 54 \end{pmatrix}$$

mit der eindeutigen Lösung

$$\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} 0,0119 \\ 0,2590 \\ -0,5248 \end{pmatrix}.$$

Die dazugehörige (optimal an die Messdaten angepasste) Modellfunktion  $y(v) = av^2 + bv + c$  ist zusammen mit den Messdaten hier abgebildet:

Messdaten und angepasste Modellfunktion (Lösung des Ausgleichsproblems)



◇

Wir können sicher sein, dass der oben berechnete stationäre Punkt tatsächlich ein globales Minimum von  $f(\vec{x})$  ist, denn es gilt folgender Satz.

**Satz 7.31 (zum linearen Ausgleichsproblem)**

Seien  $A \in \mathbb{R}^{m \times n}$  und  $\vec{b} \in \mathbb{R}^m$  gegeben,  $m > n$ .

- Das lineare Ausgleichsproblem „Minimiere  $\frac{1}{2} \|A\vec{x} - \vec{b}\|^2$ “ ist immer lösbar.
- Jeder stationäre Punkt, also jede Lösung der Normalengleichung  $A^\top A \vec{x} = A^\top \vec{b}$ , ist eine Lösung (globales Minimum) des Ausgleichsproblems.
- Falls die Spalten von  $A$  linear unabhängig sind, also  $\text{Rang}(A) = n$  gilt, dann ist die Lösung des Ausgleichsproblems eindeutig. ◇

**Beachte:** Die Hesse-Matrix von  $f(\vec{x})$  ist

$$H(\vec{x}) = A^\top A.$$

Diese Matrix ist unabhängig von der Stelle  $\vec{x}$  (!) und stets positiv semi-definit. Falls  $\text{Rang}(A) = n$  gilt, dann ist sie sogar positiv definit und damit invertierbar. In diesem Fall gibt es also genau einen stationären Punkt.

**Bemerkung 7.32 (zu Ausgleichsproblemen)** (a) Das obige Beispiel führte auf ein *lineares* Ausgleichsproblem, da die zu bestimmenden Modellparameter  $\vec{x} = (a, b, c)$  linear in die Modellfunktion  $y = av^2 + bv + c$  eingehen.

(b) Falls die Modellparameter  $\vec{x}$  nichtlinear in das Modell eingehen, so erhält man ein *nichtlineares* Ausgleichsproblem der Form

$$\text{Minimiere } f(\vec{x}) := \frac{1}{2} \sum_{i=1}^m r_i(\vec{x})^2,$$

was auch ein **nichtlineares Kleinste-Quadrate-Problem** genannt wird. Wieder geben die Residuen  $r_i(\vec{x})$  an, wie weit die Vorhersage durch die Modellfunktion (bei den Modellparametern  $\vec{x}$ ) die Messwerte verfehlt.

(c) Zur Lösung (linearer und nichtlinearer) Ausgleichsprobleme gibt es spezielle Software (z.B. `lsqcurvefit` in MATLAB), die meist das **Gauß-Newton-** oder das **Levenberg-Marquardt-Verfahren** implementiert.

(d) Die Verwendung der Kleinste-Quadrate-Technik kann stochastisch begründet werden: Dazu nimmt man an, dass die Messdaten mit Messfehlern behaftet sind, die unabhängig und identisch normal-verteilt sind mit Mittelwerten null. Dann ergibt die Lösung  $\vec{x}$  der Kleinste-Quadrate-Aufgabe gerade diejenigen Parameter, die — bei den gegebenen Messdaten — am wahrscheinlichsten sind.  $\diamond$

Ende 15. V

12.07.2011

## Literaturverzeichnis

- B. Luderer and U. Würker. *Einstieg in die Wirtschaftsmathematik*. Teubner, 1997.  
K. Meyberg and P. Vachenauer. *Höhere Mathematik 1*. Springer, Berlin, 2001.



## Index

- $k$ -mal stetig differenzierbare Funktion, 94
- Äquivalenz, 8
- Äquivalenzumformung, 42
  
- AB, *siehe* Anfangsbedingung
- abhängige Variablen, 45
- Ableitung, 154
  - Einheit, 84
- Ableitung einer Funktion, 84, 110
- Abstand, 18, 23, 55
- Abzinsen, 69
- Abzinsungsfaktor, 69
- affine lineare Funktion, 150
- algebraische Darstellung, 23
- allgemeine Lösung einer homogenen Differentialgleichung, 116
- allgemeine Lösung einer inhomogenen Differentialgleichung, 117
- allgemeine Lösung eines homogenen Differentialgleichungssystems, 124
- Alternative, 8
- Anfangsbedingung, 117, 124
- Anfangswertproblem, 117
- Ansatz vom Typ der rechten Seite, 118, 127
- anti-parallel, 34
- Argument, 23
- arithmetische Folge, 63
- Aufzinsungsfaktor, 69
- Ausgleichsproblem, 162
  - lineares, 162
  - Normalengleichung, 162
- Aussage, 7
- AWP, *siehe* Anfangswertproblem
  
- Barwert, 71
- Basis, 30, 136
- Basislösung, 136
- Basisvariable, 136
- Basisvektor, 136
  - zulässiger, 136
- Bedingung
  - hinreichende, 8
  - notwendige, 8
- beschränkte Folge, 64
- beschränkte Funktion, 75
  
- bestimmt divergente Folge, 65
- bestimmtes Integral, 102
- Betrag
  - einer komplexen Zahl, 22, 23
  - einer reellen Zahl, 18
  - eines Vektors, 33
- Bildmenge, 75
  
- charakteristisches Polynom, 120
  
- Definitionsmenge, 75, 108
- Determinante, 53
  - Entwicklungssatz, 53
- Dgl, *siehe* Differentialgleichung
- diagonalisierbare Matrix, 121
- Diagonalmatrix, 35
- Differential einer Funktion, 92
- Differentialgleichung
  - allgemeine Lösung, 116, 117
  - Anfangsbedingung, 117
  - Anfangswertproblem, 117
  - Ansatz vom Typ der rechten Seite, 118
  - harmonische Anregung, 118, 127
  - homogene, 116
  - inhomogene, 116
  - Lösung, 114, 116
  - lineare, 116
  - rechte Seite, 116
  - trennbare, 114
  - Variation der Konstanten, 117
- Differentialgleichungssystem
  - allgemeine Lösung, 124
  - Anfangsbedingung, 124
  - Ansatz vom Typ der rechten Seite, 127
  - Eigenkreisfrequenz, 126
  - homogenes, 123
  - inhomogenes, 123
  - Lösung, 123
  - lineares, 123
  - Periodendauer, 126
- Differenzenquotient, 83
- differenzierbare Funktion, 84, 110, 154
- Dimension, 30
- disjunkt, 13
- Diskontieren, 69

- divergente Folge, 64
- divergentes uneigentliches Integral, 108
- Drehung
  - im Raum, 38
  - in der Ebene, 37
- Dreiecksmatrix
  - obere, 35
  - untere, 35
- Durchschnitt, 13
- Ebene
  - Drei-Punkte-Form, 60
  - Normalenvektor, 60
  - parameterfreie Darstellung, 60
  - Punkt-Richtungs-Form, 60
- Eigenkreisfrequenz, 126
- Eigenpaar, 120
- eigentlicher Grenzwert einer Funktion, 81
- Eigenvektor, 120
  - normiert, 122
- Eigenwert, 120
- eineindeutige Funktion, 77
- Einheit der Ableitung, 84
- Einheit der Stammfunktion, 105
- Einheitsmatrix, 35
- Einheitsvektor, 26
- Elastizität einer Funktion, 92
- Empfindlichkeit, 91, 157
  - relative, 91
- Endwert, 71
- Entwicklungssatz, 53
- Ersatzfunktion
  - stetige, 83
- Eulersche Formel, 25
- EV, *siehe* Eigenvektor
- EW, *siehe* Eigenwert
- ewige Rente, 71
- Exponentialdarstellung, 25
- exponentielles Wachstum, 113
- Extremstelle einer Funktion, 95
  - unbestimmt divergente, 65
- Formel von Moivre, 25
- freie Variablen, 45
- Funktion, 75, 108
  - $k$ -mal stetig differenzierbare, 94
  - (total) differenzierbare, 154
  - Ableitung, 84, 110
  - affine lineare, 150
  - beschränkte, 75
  - Bildmenge, 75
  - Definitionsmenge, 75, 108
  - Differential, 92
  - differenzierbare, 84, 110
  - eineindeutig, 77
  - Extremstelle, 95
  - Gradient, 152
  - Graph, 75, 148
  - Höhenlinie, 148
  - Hesse-Matrix, 159
  - invertierbare, 77
  - Komposition, 76
  - konkave, 94
  - konstante, 75
  - konvexe, 94
  - logistische, 76
  - monoton fallende, 75
  - monoton wachsende, 75
  - Niveaulinie, 148
  - partiell differenzierbare, 151
  - partielle, 151
  - partielle Ableitung, 151
  - partielle Ableitung zweiter Ordnung, 158
  - periodische, 75
  - quadratische, 150
  - Richtungsableitung, 152
  - Sattelpunkt, 160
  - stückweise stetige, 101
  - stationärer Punkt, 96, 158
  - stetig diffbare, 84
  - stetige, 82, 109, 150
  - streng monoton fallende, 75
  - streng monoton wachsende, 75
  - Tangente, 84, 110
  - Tangentialebene, 152
  - umkehrbare, 77
  - unstetige, 82
  - Urbild, 76
  - vektorwertige, 109
  - verkettete, 76
  - Wertemenge, 75
  - Zielmenge, 75, 108
- ganze Zahlen, 16
- Gauß-Verfahren, 42
- Gaußklammer, 80
- geometrische Folge, 63
- geometrische Reihe, 66

- Gerade
  - Normalform, 57
  - Parameterdarstellung, 56
  - parameterfreie Darstellung, 57
  - Punkt-Richtungs-Form, 56, 59
  - Zwei-Punkte-Form, 56, 59
- Geschwindigkeitsvektor, 110
- Gleichungssystem
  - lineares, 41
- Gliederfolge, 66
- globales Maximum, 95
- globales Minimum, 95
- Grad eines Polynoms, 77
- Gradient, 152
- Graph einer Funktion, 75, 148
- Grenzfunktion, 85
- Grenzwert, 85
- Grenzwert einer Funktion, 80, 109, 150
  - eigentlicher, 81
  - linksseitiger, 80
  - rechtsseitiger, 80
- Höhenlinie einer Funktion, 148
- harmonische Anregung, 118, 127
- harmonische Reihe, 67
- Hauptdiagonale, 35
- Hesse-Matrix, 159
- hinreichende Bedingung, 8
- homogene Differentialgleichung, 116
- homogenes Differentialgleichungssystem, 123
- Identität, *siehe* Einheitsmatrix
- imaginäre Achse, 22
- imaginäre Einheit, 20
- Imaginärteil, 22
- Implikation, 8
- inhomogene Differentialgleichung, 116
- inhomogenes Differentialgleichungssystem, 123
- Innenprodukt, *siehe* Skalarprodukt
- inneres Produkt, 32
- Integral
  - bestimmtes, 102
  - divergentes uneigentliches, 108
  - konvergentes uneigentliches, 108
  - unbestimmtes, 103
  - uneigentliches, 107
- Interpolation, 78
  - Stützstelle, 78
  - Stützwert, 78
- Interpolationspolynom
  - Lagrange-Form, 79
- Interpolationspolynom, 78
- Intervall, 17
- inverse Matrix, 49
- invertierbare Funktion, 77
- invertierbare Matrix, 50
- kartesisches Produkt, 14
- Koeffizient eines Polynoms, 77
- Koeffizientenmatrix, 41
- Komplement, 13
- komplexe Zahlen, 20
  - Argument, 23
  - Betrag, 22, 23
  - Exponentialdarstellung, 25
  - Polardarstellung, 23
  - trigonometrische Darstellung, 23
- komplexe Zahlenebene, 22
- komplexe Zahlen
  - algebraische Darstellung, 23
- Komponentenfunktion, 109
- Komposition von Funktionen, 76
- konjugiert komplex, 22
- Konjunktion, 8
- konkave Funktion, 94
- Konklusion, 8
- konstante Folge, 63
- konstante Funktion, 75
- konvergente Folge, 64, 149
- konvergentes uneigentliches Integral, 108
- konvexe Funktion, 94
- Kostenvektor, 130
- Kreuzprodukt, 57
- Länge
  - eines Vektors, 33
- lösbare lineare Optimierungsaufgabe, 132
- Lösung einer Differentialgleichung, 114, 116
- Lösung eines Differentialgleichungssystems, 123
- Lagrange-Form des Restgliedes, 100
- Lagrange-Form eines
  - Interpolationspolynoms, 79
- Laufzeit, 67
- leere Menge, 12
- LGS, *siehe* lineares Gleichungssystem
- linear abhängig, 29
- linear unabhängig, 29
- lineare Differentialgleichung, 116
- lineare Fehlerrechnung, 92, 157
- lineare Hülle, 28
- Lineare Optimierungsaufgabe, 130
  - Basis, 136
  - Basislösung, 136
  - Basisvariable, 136
  - Basisvektor, 136
    - zulässiger, 136
  - Beschränkung, 130
  - entartete Basislösung, 146

- entarteter Basisvektor, 146
- Kostenvektor, 130
- Nichtbasis, 136
- Normalform, 133
- optimale Lösung, 131
- Regel von Bland, 146
- Schlupfvariable, 133
- unbeschränkte, 132
- unzulässige, 132
- Variablen, 130
- Zielfunktion, 130
- zulässige Basislösung, 136
- zulässige Menge, 130
- zulässiger Bereich, 130
- lineares Differentialgleichungssystem, 123
- lineares Gleichungssystem, 41
  - abhängige Variablen, 45
  - allgemeine Lösung, 45
  - freie Variablen, 45
  - homogenes, 41
  - Lösung, 41
  - partikuläre Lösung, 49
  - rechte Seite, 41
- Linearkombination, 28
  - triviale, 28
- linksseitiger Grenzwert, 80
- LK, *siehe* Linearkombination
- LOA, *siehe* Lineare Optimierungsaufgabe
- logistische Funktion, 76
- logistisches Wachstum, 114
- lokales Maximum, 95
- lokales Minimum, 95
- Matrix, 34
  - Addition, 35
  - charakteristisches Polynom, 120
  - Determinante, 53
  - diagonalisierbare, 121
  - Eigenvektor, 120
  - Eigenwert, 120
  - inverse, 49
  - invertierbare, 50
  - Multiplikation mit einem Skalar, 35
  - negativ definite, 159
  - nicht invertierbare, 50
  - orthogonale, 122
  - positiv definite, 159
  - quadratische, 35
  - Rang, 45
  - reguläre, 50
  - singuläre, 50
  - symmetrisch, 36
- Matrix–Matrix-Produkt, 38
- Matrix–Vektor-Produkt, 36
- Maximum
  - globales, 95
  - lokales, 95
- Menge
  - leere, 12
- Minimum
  - globales, 95
  - lokales, 95
- monoton fallende Funktion, 75
- monoton wachsende Funktion, 75
- nachschüssige Rente, 71
- natürliche Zahlen, 16
- Negation, 7
- negativ definite Matrix, 159
- Newton-Verfahren, 90
- nicht invertierbare Matrix, 50
- Nichtbasis, 136
- Niveaulinie, 148
- Norm
  - eines Vektors, 33
- Normalengleichung, 162
- Normalform einer linearen
  - Optimierungsaufgabe, 133
- notwendige Bedingung, 8
- Nullmatrix, 35
- Nullvektor, 26
- obere Dreiecksmatrix, 35
- Obermenge, 13
  - echte, 13
- optimale Lösung, 131
- Optimalitätsindikator, 137
- orthogonal, 34
- orthogonale Matrix, 122
- orthogonale Zerlegung, 55
- Ortsvektor, 55
- parallel, 34
- Partialsummenfolge, 66
- partiell differenzierbare Funktion, 151
- partielle Ableitung, 151
- partielle Ableitung zweiter Ordnung, 158
- partielle Funktion, 151
- Periodendauer, 126
- periodische Funktion, 75
- Pfeil, 55
- Phase I, 147
- Polardarstellung, 23
- Polarkoordinaten, 156
- Polynom, 77
  - Grad, 77
  - Interpolation, 78
- Polynominterpolation, 78
- positiv definite Matrix, 159
- Prämisse, 8
- Produkt
  - Matrix–Matrix, 38
  - Matrix–Vektor, 36
- Projektion, 55, 59

- quadratische Funktion, 150
- quadratische Matrix, 35
- Quotiententest, 138
  
- Rückwärtssubstitution, 42
- Rang, 45
- Ratenzahlung, 71
- rationale Zahlen, 16
- Realteil, 22
- rechte Seite
  - lineares Gleichungssystem, 41
- rechte Seite einer Differentialgleichung, 116
- rechtsseitiger Grenzwert, 80
- Rechtssystem, 57
- reelle Achse, 22
- reelle Zahlen, 16
  - Betrag, 18
- Regel von Bland, 146
- Regel von Sarrus, 53
- reguläre Matrix, 50
- Reihe, 66
  - geometrische, 66
  - Gliederfolge, 66
  - harmonische, 67
  - Summe, 66
- relative Empfindlichkeit, 91
- relative Sensitivität, 91
- Rente, 71
  - Barwert, 71
  - Endwert, 71
  - nachschüssige, 71
  - vorschüssige, 71
- Renten-Barwert, 72
- Renten-Endwert, 71, 72
- Resonanzkatastrophe, 129
- Restglied, 85
- Restglied eines Taylorpolynoms, 100
- Richtungs des größten Anstiegs, 155
- Richtungsableitung, 152
- Riemann-Summe, 101
  
- Sandwich-Theorem, 65
- Sattelpunkt, 160
- Satz
  - von Schwarz, 159
  - von Taylor, 99
- Schlupfvariable, 133
- Schraubenlinie, 109
- Sekante, 84
- senkrecht, 34
- Sensitivität, 91, 157
  - relative, 91
- Simplex-Verfahren
  - Optimalitätsindikator, 137
  - Phase I, 147
  - Quotiententest, 138
  - Simplextableau, 138
- Simplextableau, 138
- sinuläre Matrix, 50
- Skalar, 26
- Skalarprodukt, 32
- Spaltenvektor, 32
- Spiegelung
  - in der Ebene, 38
- stückweise stetige Funktion, 101
- Stützstelle, 78
- Stützwert, 78
- Stammfunktion, 103
  - Einheit, 105
- Startkapitel, 67
- stationärer Punkt, 96, 158
- steilster Abstieg, 155
- steilster Anstieg, 155
- stetig diffbare Funktion, 84
- stetige Ersatzfunktion, 83
- stetige Funktion, 82, 109, 150
- streng monoton fallende Funktion, 75
- streng monoton wachsende Funktion, 75
- Superposition, 49, 117
- symmetrische Matrix, 36
  
- Tangente, 84, 110
- Tangentialebene, 152
- Tangentialvektor, 110
- Taylorpolynom, 98, 152, 159
  - Restglied, 100
- Teilmenge, 13
  - echte, 13
- transponieren
  - Matrix, 36
- trennbare Differentialgleichung, 114
- trigonometrische Darstellung, 23
  
- umkehrbare Funktion, 77
- Umkehrfunktion, 77
- unbeschränkte lineare Optimierungsaufgabe, 132
- unbestimmt divergente Folge, 65
- unbestimmtes Integral, 103
- uneigentliches Integral, 107
- unstetige Funktion, 82
- untere Dreiecksmatrix, 35
- Unterraum, 30
- unzulässige lineare Optimierungsaufgabe, 132
- UR, *siehe* Unterraum
- Urbild, 76
  
- Variation der Konstanten, 117
- Vektor, 26
  - Addition, 26
  - Betrag, 33
  - Komponente, 26
  - Kreuzprodukt, 57

- Länge, 26, 33
- Multiplikation, 26
- Norm, 33
- orthogonale Zerlegung, 55
- Subtraktion, 27
- Vektorprodukt, 57
- Vektorprodukt, 57
- Vektorraum  $\mathbb{R}^n$ , 26
- vektorielle Funktion, 109
- Venn-Diagramm, 13
- Vereinigung, 13
- verkettete Funktion, 76
- vorschüssige Rente, 71
- Vorwärtselimination, 42
  
- Wahrheitswert, 7
- Wertemenge, 75
- Winkel, 33
  
- Zahlen
  - ganze, 16
  - komplexe, 20
  - natürliche, 16
  - rationale, 16
  - reelle, 16
- Zahlenebene
  - komplexe, 22
- Zahlenstrahl, 17
- Zeilenstufenform, 44
- Zeilenvektor, 32
- Zeitrente, 71
- Zielfunktion, 130
- Zielmenge, 75, 108
- Zinsen, 67
- Zinsperiode, 67
- Zinssatz, 67
- ZSF, *siehe* Zeilenstufenform
- zulässige Basislösung, 136
- zulässige Menge, 130
- zulässiger Bereich, 130

## Abbildungsnachweis

Teilmenge		
	<a href="http://commons.wikimedia.org/wiki/File:Set_subsetAofB.svg">http://commons.wikimedia.org/wiki/File:Set_subsetAofB.svg</a>	13
Vereinigungsmenge, Schnittmenge, Restmenge		
	<a href="http://commons.wikimedia.org/wiki/File:Venn0001.svg">http://commons.wikimedia.org/wiki/File:Venn0001.svg</a>	13
Linearkombinationen zweier Vektoren		
	<a href="http://commons.wikimedia.org/wiki/File:Vector_addition_and_scaling.svg">http://commons.wikimedia.org/wiki/File:Vector_addition_and_scaling.svg</a>	27
Drehmoment		
	<a href="http://commons.wikimedia.org/wiki/File:Drehmoment.svg">http://commons.wikimedia.org/wiki/File:Drehmoment.svg</a>	57
Kreuzprodukt		
	<a href="http://commons.wikimedia.org/wiki/File:Cross_product_vector.svg">http://commons.wikimedia.org/wiki/File:Cross_product_vector.svg</a>	58
Fibonacci-Folge am Blütenstand der Sonnenblume		
	<a href="http://commons.wikimedia.org/wiki/File:Goldener_Schnitt_Bluetenstand_Sonnenblume.jpg">http://commons.wikimedia.org/wiki/File:Goldener_Schnitt_Bluetenstand_Sonnenblume.jpg</a>	63