



Prof. Dr. Michael Herrmann  
Technische Universität Braunschweig  
Mathematik – Institut iPDE  
[michael.herrmann@tu-braunschweig.de](mailto:michael.herrmann@tu-braunschweig.de)

*Skript der Vorlesung*  
***Mathematik I***  
*im Wintersemester 2019/20*  
*für die Studiengänge ET, IST, WING der Fakultät 5*

© Michael Herrmann

Der Autor ist für Hinweise und Kommentare jederzeit dankbar.

Dieses Skript ist lizenziert unter **CC BY-SA 3.0 DE**.  
<http://creativecommons.org/licenses/by-sa/3.0/de/>



## Griechisches Alphabet

<i>klein</i>	<i>groß</i>	<i>Name</i>	<i>Laut</i>
$\alpha$	$A$	alpha	a
$\beta$	$B$	beta	b
$\gamma$	$\Gamma$	gamma	g
$\delta$	$\Delta$	delta	d
$\varepsilon, \epsilon$	$E$	epsilon	ě
$\zeta$	$Z$	zeta	z
$\eta$	$H$	eta	ē
$\theta, \vartheta$	$\Theta$	theta	th
$\iota$	$I$	iota	i
$\kappa$	$K$	kappa	k
$\lambda$	$\Lambda$	lambda	l
$\mu$	$M$	my	m

<i>klein</i>	<i>groß</i>	<i>Name</i>	<i>Laut</i>
$\nu$	$N$	ny	n
$\xi$	$\Xi$	xi	x
$o$	$O$	omikron	ö
$\pi$	$\Pi$	pi	p
$\varrho, \rho$	$P$	rho	r
$\sigma$	$\Sigma$	sigma	s
$\tau$	$T$	tau	t
$\upsilon$	$\Upsilon$	upsilon	y
$\varphi, \phi$	$\Phi$	phi	ph, f
$\chi$	$X$	chi	ch
$\psi$	$\Psi$	psi	ps
$\omega$	$\Omega$	omega	ō

## Literatur

Das Skript folgt in groben Zügen der Darstellung in

[AORS] R. ANSORGE, H.J. OBERLE, K. ROTHE, TH. SONAR:  
*Mathematik für Ingenieure I*. Wiley-VCH 2010, 4. Auflage.

Die wesentlichen Inhalte (Definitionen, Theoreme, Beweise, usw.) sowie weitere Beispiele können auch in vielen anderen Büchern gefunden werden, zum Beispiel in:

[ABHKLS] T. ARENS, R. BUSAM, FR. HETTLICH, CH. KARPFINGER,  
 KL. LICHTENEGGER, H. STACHEL:  
*Mathematik*. SpringerSpektrum 2015, 3. Auflage.

Zu empfehlen sind auch die Werke

[Pap1] L. PAPULA:  
*Mathematik für Ingenieure und Naturwissenschaftler: Band 1*.  
 SpringerVieweg 2018, 15. Auflage.

[Pap2] L. PAPULA:  
*Mathematik für Ingenieure und Naturwissenschaftler: Band 2*.  
 SpringerVieweg 2015, 14. Auflage.

[PapF] L. PAPULA:  
*Mathematische Formelsammlung*. SpringerVieweg 2018, 8. Auflage.

[PapA] L. PAPULA:  
*Mathematik für Ingenieure und Naturwissenschaftler: Anwendungsbeispiele*. SpringerVieweg 2017, 12. Auflage.

die zwar keine Beweise, aber sehr ausführliche Erklärungen, Übersichten, Beispiele und Anwendungen enthalten.

# Inhaltsverzeichnis

<b>1</b>	<b>Logik und Mengenlehre</b>	<b>5</b>
1.1	Aussagenlogik . . . . .	5
1.2	Mengenlehre . . . . .	9
1.3	*Exkurs über die Unendlichkeit . . . . .	14
<b>2</b>	<b>Zahlen und Zahlbereiche</b>	<b>15</b>
2.1	Natürliche Zahlen . . . . .	16
2.2	Reelle Zahlen . . . . .	20
2.3	Komplexe Zahlen . . . . .	29
<b>3</b>	<b>Vektorrechnung und Analytische Geometrie</b>	<b>41</b>
3.1	Vektoren im $\mathbb{R}^n$ . . . . .	41
3.2	(Affine) Geraden und Ebenen im $\mathbb{R}^n$ . . . . .	51
3.3	Allgemeine Vektorräume . . . . .	55
<b>4</b>	<b>Lineare Gleichungssysteme und Matrizen</b>	<b>63</b>
4.1	Matrizen . . . . .	63
4.2	Lineare Gleichungssysteme – Praxis . . . . .	67
4.3	Lineare Gleichungssysteme – Theorie . . . . .	75
4.4	Quadratische Matrizen . . . . .	81
4.5	*Elektrische Netzwerke und Matrizen . . . . .	92
<b>5</b>	<b>Lineare Abbildungen</b>	<b>103</b>
5.1	Basisdarstellungen und -wechsel . . . . .	103
5.2	Orthogonalität und Gram-Schmidt-Verfahren . . . . .	116
5.3	Orthogonale Matrizen . . . . .	122
<b>6</b>	<b>Spektraltheorie</b>	<b>129</b>
6.1	Eigenwerte und Eigenvektoren . . . . .	129
6.2	Normalform allgemeiner Matrizen . . . . .	136
6.3	Hauptachsentransformation . . . . .	143
<b>7</b>	<b>Folgen und Reihen</b>	<b>149</b>
7.1	Folgen . . . . .	149
7.2	Reihen . . . . .	157
<b>8</b>	<b>Stetigkeit und Differenzierbarkeit</b>	<b>165</b>
8.1	Stetigkeit . . . . .	165
8.2	Differenzierbarkeit . . . . .	175

---

<b>9</b>	<b>Differentialrechnung</b>	<b>187</b>
9.1	Extrema, Mittelwertsätze und l'Hospital . . . . .	187
9.2	Satz von Taylor . . . . .	192
9.3	Ableitungen in der Kurvendiskussion . . . . .	201
9.4	Exkurs über Potenzreihen . . . . .	205
<b>10</b>	<b>Integralrechnung</b>	<b>211</b>
10.1	Bestimmtes Riemann-Integral . . . . .	211
10.2	Der Hauptsatz der Differential- und Integralrechnung . . . . .	218
10.3	Weitere Aspekte der Integralrechnung . . . . .	228
<b>11</b>	<b>Fourier-Reihen</b>	<b>259</b>
11.1	Fourier-Reihen periodischer Funktionen . . . . .	259
11.2	Elemente der Theorie . . . . .	268
11.3	Varianten der Fourier-Reihe . . . . .	274
<b>A</b>	<b>Wiederholung: Lineare Abbildungen und Matrizen</b>	<b>279</b>

# Kapitel 1

## Logik und Mengenlehre

### 1.1 Aussagenlogik

Vorlesung 01, 23. Oktober 2019

---

**Motivation** Sauberes Schließen und fehlerfreies Argumentieren sind zentral in den Wissenschaften.

**Definition** Mathematische Aussagen sind Sätze, die wahr/richtig oder falsch sind (obwohl wir die Antwort vielleicht nicht kennen).

#### **Beispiele**

1. 2 ist eine Primzahl. (*richtig*)
2. Der Umfang eines Kreises von Radius  $r$  ist  $\pi r^2$ . (*falsch*)
3. Mein Zug hatte gestern Verspätung. (*je nach Kontext*)
4. Es gibt unendliche viele Primzahlzwillinge. (*nicht bekannt*)

#### **Gegenbeispiele**

1. Der Film war spannend.
2. Das Essen war gut.
3. Morgen wird es regnen.

**Notation** Aussagen werden meist mit Großbuchstaben ( $A, B, \dots$ ) bezeichnet; Wahrheitswert einer Aussage:

$$w(A) = 0 \text{ falls } A \text{ falsch ist,} \quad w(A) = 1 \text{ falls } A \text{ wahr ist,}$$

## Verknüpfungen von Aussagen

$\neg A$	nicht $A$	<u>Negation</u>
$A \wedge B$	$A$ und $B$	<u>Konjunktion</u> (von $A$ und $B$ )
$A \vee B$	$A$ oder $B$	<u>Disjunktion</u> bzw. <u>Alternative</u>
$A \Rightarrow B$	aus $A$ folgt $B$	<u>Implikation</u> (mit <u>Prämisse</u> $A$ und <u>Konklusion</u> $B$ )
$A \Leftrightarrow B$	$A$ genau dann, wenn $B$	<u>Äquivalenz</u>

$w(A)$	$w(B)$	$w(\neg A)$	$w(A \wedge B)$	$w(A \vee B)$	$w(A \Rightarrow B)$	$w(A \Leftrightarrow B)$
0	0	1	0	0	1	1
0	1	1	0	1	1	0
1	0	0	0	1	0	0
1	1	0	1	1	1	1

### Bemerkung

- Das mathematische ODER ist nicht exklusiv:  $A \vee B$  ist auch dann wahr, wenn sowohl  $A$  als auch  $B$  wahr sind.
- Eine Implikation mit falscher Prämisse ist immer richtig bzw. aus einer falschen Aussage folgt alles. *Beispiel:* Wenn 4 eine Primzahl ist, so ist auch 8 eine Primzahl.

### Beispiele

- Konjunktion: 3 ist eine Primzahl und 6 ist gerade.
- Alternative: 2 ist eine Primzahl oder 3 ist eine Primzahl.
- Implikation: Jede durch 4 teilbare Zahl ist gerade. ( $A \Rightarrow B$  mit  $A = „n$  ist durch 4 teilbar“ und  $B = „n$  ist gerade“)

**Definition** Tautologien sind Aussagen, die immer wahr sind.

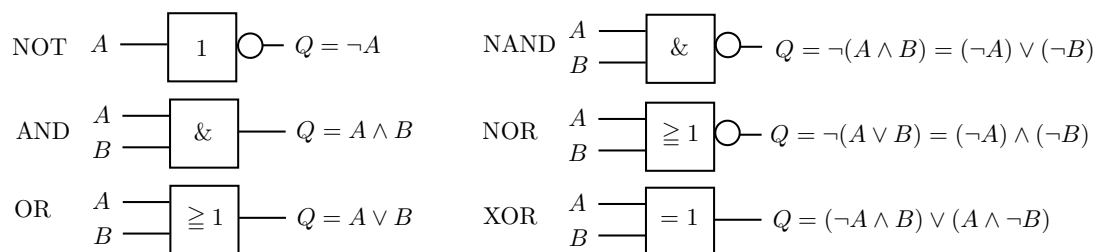
$A \vee \neg A$	(Satz vom ausgeschlossenen Dritten)
$\neg(A \wedge \neg A)$	(Satz vom Widerspruch)
$\neg(\neg A) \Leftrightarrow A$	(Negation der Negation)
$\neg(A \vee B) \Leftrightarrow (\neg A) \wedge (\neg B)$	(1. Regel von de Morgan)
$\neg(A \wedge B) \Leftrightarrow (\neg A) \vee (\neg B)$	(2. Regel von de Morgan)
$A \wedge (B \vee C) \Leftrightarrow (A \wedge B) \vee (A \wedge C)$	(Distributivgesetz)
$A \vee (B \wedge C) \Leftrightarrow (A \vee B) \wedge (A \vee C)$	(Distributivgesetz)
$(A \Rightarrow B) \Leftrightarrow (\neg B \Rightarrow \neg A)$	(Kontraposition)
$(A \Rightarrow B) \Leftrightarrow (\neg A \vee B)$	(...)
$(A \Rightarrow B) \wedge A \Rightarrow B$	(modus ponens)
$(A \Rightarrow B) \wedge \neg B \Rightarrow \neg A$	(modus tollens)
$(A \Rightarrow B) \wedge (B \Rightarrow C) \Rightarrow (A \Rightarrow C)$	(modus barbara)

**Beispiel** Überprüfung der *Kontraposition* mittels einer Wahrheitstabelle

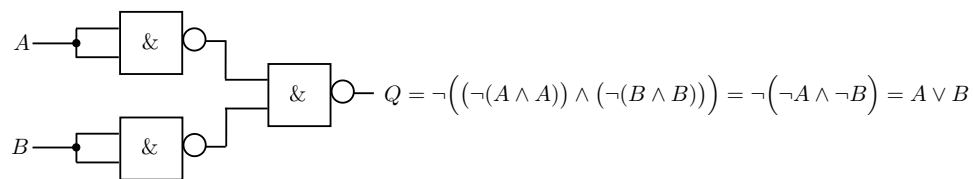
$A$	$B$	$A \Rightarrow B$	$\neg A$	$\neg B$	$\neg B \Rightarrow \neg A$	Kontraposition
0	0	1	1	1	1	1
0	1	1	1	0	1	1
1	0	0	0	1	0	1
1	1	1	0	0	1	1

**Bemerkung** Die Kontraposition ist die Grundlage des *indirekten Beweises* (siehe Große Übung II).

### Analogie in der Digitaltechnik = Logische Gatter



Die logischen Grundgatter.



Ein OR-Gatter, dass aus drei NAND-Gattern gebaut wurde.

**Definition** Aussageformen sind Aussagen, die von einer Variablen abhängen, zum Beispiel

$$A(x) = x \text{ ist gerade}$$

Der Wahrheitswert von  $A(x)$  hängt vom Wert der Variablen  $x$  ab.

### Aussagen mit Quantoren

Aussage	Bedeutung
$\forall x : A(x)$ ( <u>Allaussage</u> )	Für jedes $x$ ist $A(x)$ wahr.
$\exists x : A(x)$ ( <u>Existenzaussage</u> )	Es existiert (mind.) ein $x$ , so dass $A(x)$ wahr ist.

### Bemerkung

- Die Ausdrücke „ $\forall x : A(x)$ “ und „ $\forall y : A(y)$ “ beschreiben dieselbe mathematische Aussage. Analog für „ $\exists x : A(x)$ “ und „ $\exists y : A(y)$ “

2. Es gelten die folgenden zwei Tautologien der Negation

$$\neg(\forall x : A(x)) \Leftrightarrow \exists x : \neg A(x)$$

und

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

Insbesondere ist die Negation einer Allaussage *immer* eine Existenzaussage und die Negation einer Existenzaussage *immer* eine Allaussage.

### Beispiel

1. Um die Aussage „Alle Primzahlen sind ungerade“ zu widerlegen, brauchen wir nur ein Gegenbeispiel anzugeben (nämlich die 2).
2. Um die Aussage „Es gibt eine gerade Primzahl größer 2“ zu widerlegen, müssen wir zeigen, dass jede Primzahl größer 2 ungerade ist (was nicht sonderlich schwierig ist).

**Hinweis und Bitte** Benutzen Sie (z. Bsp. beim Aufschrieb der Hausaufgaben) möglichst keine (oder nur ganz wenige) Quantoren, sondern argumentieren Sie lieber verbal. Erfahrungsgemäß kommt sonst sehr schnell sehr viel Unsinn zustande.

Ganz allgemein gilt: Das Quintessenz der Mathematik besteht nicht darin, abstrakte Formeln zu benutzen, sondern logisch zu denken und seine Gedanken klar und möglichst einfach auszudrücken. Formeln können dabei äußerst hilfreich sein, aber sie können auch verwirren und vom Wesentlichen ablenken.

Vorlesung 02, 25. Oktober 2019

**Einschub** Für zwei ganze Zahlen  $k < l$  schreibt man

$$\sum_{j=k}^l a_j = a_k + a_{k+1} + a_{k+2} + \dots + a_{l-1} + a_l$$

sowie

$$\prod_{j=k}^l a_j = a_k \cdot a_{k+1} \cdot a_{k+2} \cdot \dots \cdot a_{l-1} \cdot a_l,$$

wobei

$$\sum_{j=k}^k a_j = a_k \quad \text{sowie} \quad \prod_{j=k}^k a_j = a_k$$

im Grenzfall  $k = l$  gilt. Insbesondere gilt

$$\sum_{j=1}^n j = 1 + 2 + \dots + n = \frac{1}{2}n(n+1) \quad (\text{Gauß-Formel})$$

sowie

$$n! = 1 \cdot 2 \cdot \dots \cdot n = \prod_{j=1}^n j \quad (\text{sprich: } n \text{ Fakultät}).$$



## 1.2 Mengenlehre

**Motivation** Die fundamentalen Konzepte der modernen Mathematik sind nicht – wie die meisten Leute meinen – Zahlen, sondern Mengen und Abbildungen. Diese Begriffe sind zum einen sehr einfach zu verstehen, aber zum anderen auch extrem nützlich.

**Definition** Eine Menge ist eine Zusammenfassung von unterscheidbaren Elementen der gleichen Art.

### Beispiele

1. Alle Studierenden, die heute an der Vorlesung teilnehmen.
2. Meine Freunde auf Facebook (zu einem Stichtag).
3. Zu jedem Text gibt es die Menge seiner Wort. Es gibt auch die Menge seiner Buchstaben.
4. Die Menge aller Dreiecke in der Ebene.
5. Die Menge der natürlichen/rationalen/reellen/komplexen Zahlen.
6.  $\{0, 1\}$ ,  $\{2, 3\}$ ,  $\{6, 7, 8, 9\}$  (drei Mengen von Zahlen).
7.  $\{\{0, 0\}, \{0, 1\}, \{1, 0\}, \{1, 1\}\}$  (eine Menge von Mengen von Zahlen).

**Notation** Es gibt verschiedene Möglichkeiten, eine Menge anzugeben:

<i>Art</i>	<i>Beispiel</i>
verbale Beschreibung	Die Menge $M$ , bestehend aus ...
Aufzistung der Elemente	$M = \{1, 6, 9\}$ $M = \{2, 4, 6, 8, \dots\}$
durch Aussageform	$M = \{x : A(x) \text{ ist wahr}\}$ $M = \{x : A(x)\}$ $M = \{x \mid A(x)\}$

Wir schreiben  $m \in M$  (bzw.  $m \notin M$ ), sofern  $m$  ein (bzw. kein) Element von  $M$  ist.

**Definition**  $\tilde{M}$  wird Teilmenge von  $M$  genannt (wir schreiben dann  $\tilde{M} \subset M$ ), wenn jedes Element von  $\tilde{M}$  auch Element von  $M$  ist (d.h., wenn die Implikation  $m \in \tilde{M} \Rightarrow m \in M$  gilt).

### Bemerkung

1. Die Elemente einer Menge müssen unterscheidbar sein, d.h. sie dürfen nicht doppelt vorkommen.  $\{1, 2, 2, 3\}$  beschreibt also keine Menge.
2. Zwei Mengen sind genau dann gleich, wenn sie dieselben Elementen enthalten. Insbesondere beschreiben  $\{1, 2, 3, 4\}$  und  $\{1, 3, 2, 4\}$  dieselbe Menge.
3. Es gibt die leere Menge  $\emptyset$ , die gar kein Element besitzt.
4. Für jede Menge  $M$  gilt  $\emptyset \subset M$  und  $M \subset M$ .

5. Die Anzahl der Elemente einer Menge  $M$  wird mit  $\#M$  bezeichnet und auch Mächtigkeit von  $M$  genannt. Sie kann den Wert 0 annehmen (für die leere Menge), eine natürliche Zahl sein (für endliche Mengen) oder den Wert  $\infty$  annehmen (für unendliche Mengen).
6. Es dürfen immer nur Elemente derselben Art zu einer Menge zusammengefasst werden, d.h. es gibt zum Beispiel keine Menge, die Personen, Dreiecke, und Zahlen als Elemente enthält.
7. Konstrukte wie „Die Menge aller Mengen“ führen zu logischen Widersprüchen, siehe Russels Paradox.

### Mengenoperationen

$$\begin{aligned} M_1 \cup M_2 &:= \{m : m \in M_1 \vee m \in M_2\} && \text{(Vereinigung)} \\ M_1 \cap M_2 &:= \{m : m \in M_1 \wedge m \in M_2\} && \text{(Durchschnitt)} \\ M_1 \setminus M_2 &:= \{m : m \in M_1 \wedge m \notin M_2\} && \text{(Differenz)} \end{aligned}$$

Zwei Mengen  $M_1$  und  $M_2$  mit  $M_1 \cap M_2 = \emptyset$  heißen disjunkt.

$$M_1 \times M_2 = \{(m_1, m_2) : m_1 \in M_1, m_2 \in M_2\} \quad \text{(Kreuzprodukt, Menge von Paaren)}$$

$$\mathcal{P}(M) = \{N : N \subset M\} \quad \text{(Potenzmenge, Menge aller Teilmengen von } M)$$

**Beispiel** Mit

$$M_1 = \{1, 3, 5\}, \quad M_2 = \{1, 2, 3\}$$

gilt

$$M_1 \cup M_2 = \{1, 2, 3, 5\}, \quad M_1 \cap M_2 = \{1, 3\}, \quad M_1 \setminus M_2 = \{5\}, \quad M_2 \setminus M_1 = \{2\}$$

sowie

$$M_1 \times M_2 = \{(1, 1), (1, 2), (1, 3), (3, 1), (3, 2), (3, 3), (5, 1), (5, 2), (5, 3)\}$$

und

$$\mathcal{P}(M_1) = \{\emptyset, \{1\}, \{3\}, \{5\}, \{3, 5\}, \{1, 5\}, \{1, 3\}, \{1, 3, 5\}\}.$$

**Beispiel** Es gilt

$$\mathbb{R}^2 = \mathbb{R} \times \mathbb{R},$$

wobei  $\mathbb{R}$  die Menge der reellen Zahlen bzw. die Menge der Punkte auf einer Geraden bezeichnet und  $\mathbb{R}^2 = \{(x_1, x_2) : x_1 \in \mathbb{R}, x_2 \in \mathbb{R}\}$  die Standardbezeichnung für die Zahlenebene ist.

**Bemerkung** Für jede endliche Menge  $M$  gilt

$$\#\mathcal{P}(M) = 2^{\#M}.$$

Man kann dies kombinatorisch (siehe Große Übung bzw. unten) oder mit vollständiger Induktion beweisen (siehe [AORS, Seite 18]).

**Definition** Eine Abbildung von einer Menge  $D$  in eine Menge  $Z$  ist eine Vorschrift, die jedem Element von  $D$  genau ein Element aus  $Z$  zuordnet. Je nach Kontext wird eine Abbildung auch Funktion genannt.

### Beispiele

1. Polynome sind Abbildungen der Menge der reellen Zahlen ( $\mathbb{R}$  genannt) in sich, d.h. es gilt  $D = Z = \mathbb{R}$ .
2. Ordnet man jedem Dreieck seinen Flächeninhalt zu, so entsteht eine Abbildung mit  $D = \{\text{Dreiecke}\}$  in  $Z = \mathbb{R}$ .
3. Man kann jede Menge von Personen in die Menge der Datumsangaben abbilden, indem man jeder Person ihren Geburtstag zuordnet.

**Notationen** Man schreibt meist  $f : D \rightarrow Z$  für eine Abbildungen  $f$  zwischen zwei gegebenen Mengen  $D$  und  $Z$ . Es gibt mehrere Möglichkeiten, eine solche zu definieren bzw. anzugeben, zum Beispiel:

1. Verbale Beschreibung.
2. Bei endlichen Mengen  $D$  kann man eine Tabelle benutzen, zum Beispiel

$x \in D$	1	2	3	4	5	6	7	8	9
$f(x) \in Z$	1	2	3	2	5	3	7	2	3

3. Oftmals wird eine Formel angegeben, zum Beispiel

$$f : \mathbb{R} \rightarrow \mathbb{R} \quad \text{mit} \quad f(x) = x^2 \quad \text{für} \quad x \in \mathbb{R}.$$

**Definition** Sei  $f : D \rightarrow Z$  eine beliebige Abbildung.

1.  $D$  ist Definitionsbereich von  $f$  und  $Z$  wird der Zielbereich von  $f$  (oder der Wertebereich von  $f$ ) genannt.
2.  $f$  heißt injektiv, wenn verschiedene Elemente aus  $D$  auf verschiedene Elemente aus  $Z$  abgebildet werden. Oder anders gesagt:  $f$  ist genau dann injektiv, wenn

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

3.  $f$  heißt surjektiv, wenn jedes  $y \in Z$  das Bild von einem  $x \in D$  unter  $f$  ist, d.h. wenn

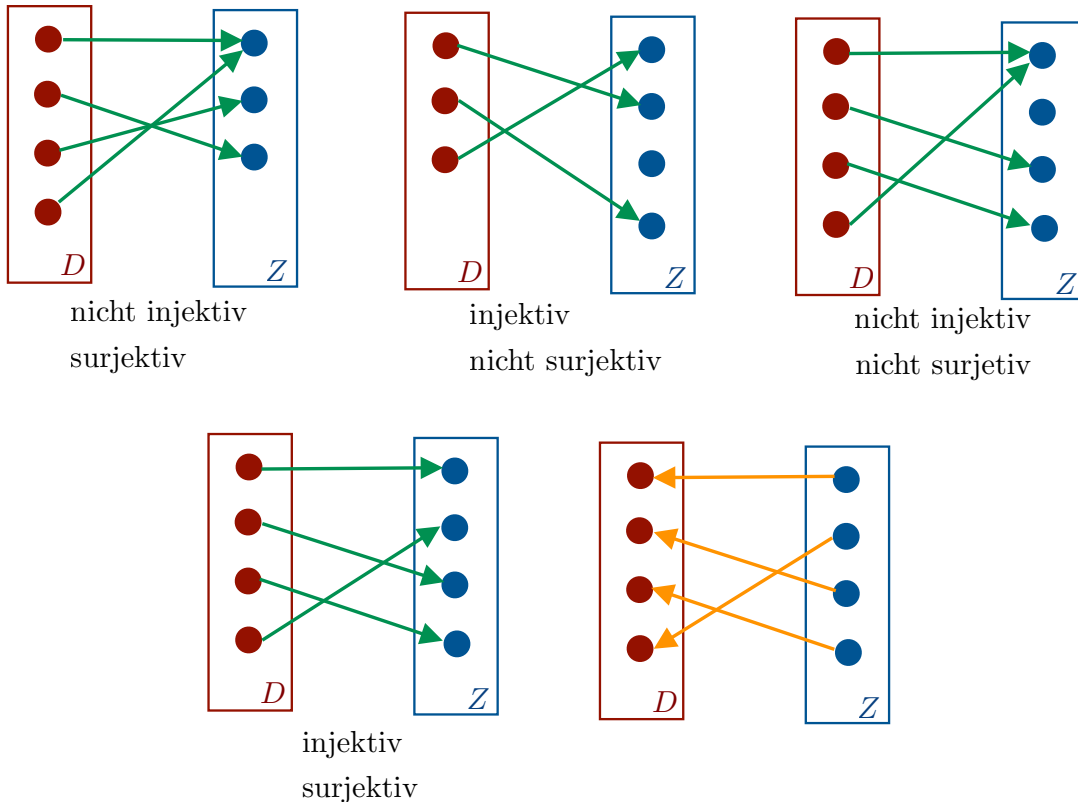
$$\forall y \in Z \exists x \in D : y = f(x)$$

4.  $f$  heißt bijektiv, falls  $f$  sowohl injektiv als auch surjektiv ist. In diesem Fall gibt es eine Abbildung  $g : Z \rightarrow D$  mit

$$g(f(x)) = x \quad \text{für alle} \quad x \in D, \quad (g(y)) = y \quad \text{für alle} \quad y \in Z.$$

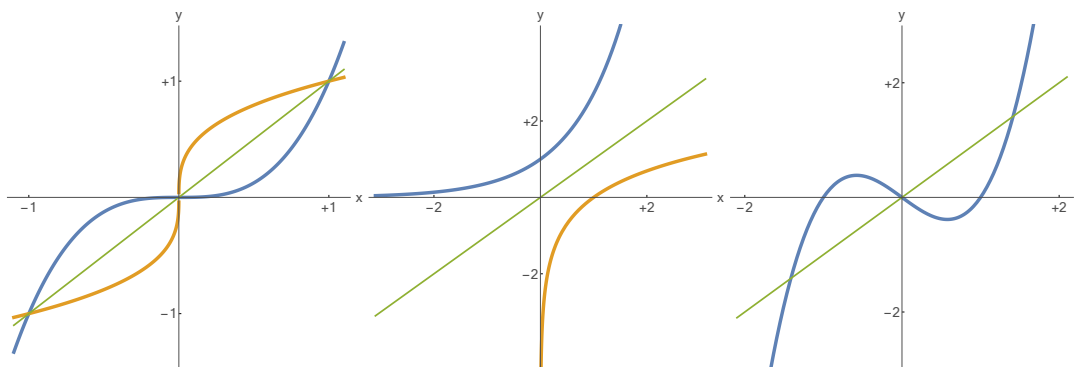
Die Abbildung  $g$  wird die Umkehrabbildung von  $f$  (oder die Inverse von  $f$ ) genannt und mit  $f^{-1}$  bezeichnet.

5. Bijektive Abbildungen heißen auch invertierbar.

**Beispiele**

**Beispiele** Sei  $D = Z = \mathbb{R}$  und  $f : \mathbb{R} \rightarrow \mathbb{R}$ .

1.  $f(x) = x^3$  definiert eine bijektive Abbildung mit  $f^{-1}(y) = \operatorname{sgn}(y) \cdot |y|^{-1/3}$ .
2. Mit  $f(x) = \exp(x)$  ist  $f$  injektiv, aber nicht surjektiv. Die inverse Abbildung (der sogenannte natürliche Logarithmus) existiert als Abbildung mit  $D = (0, \infty)$  und  $Z = \mathbb{R}$ , aber nicht als Abbildung von  $\mathbb{R}$  nach  $\mathbb{R}$ .
3.  $f(x) = x^3 - x$  liefert surjektive, aber nicht injektive Abbildung.



Die Graphen der drei Beispiel-Funktionen von  $\mathbb{R}$  nach  $\mathbb{R}$  (blau). Der Graph der inversen Funktion  $f^{-1}$  (orange) entsteht (sofern diese existiert) durch Spiegelung des Graphen von  $f$  an der Diagonalen  $y = x$  (grün).

**Definition** Sei  $f : D \rightarrow Z$  eine beliebige Abbildung.

## 1. Die Menge

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

wird Graph von  $f$  genannt.

2. Für jede Teilmenge  $A \subset D$  wird

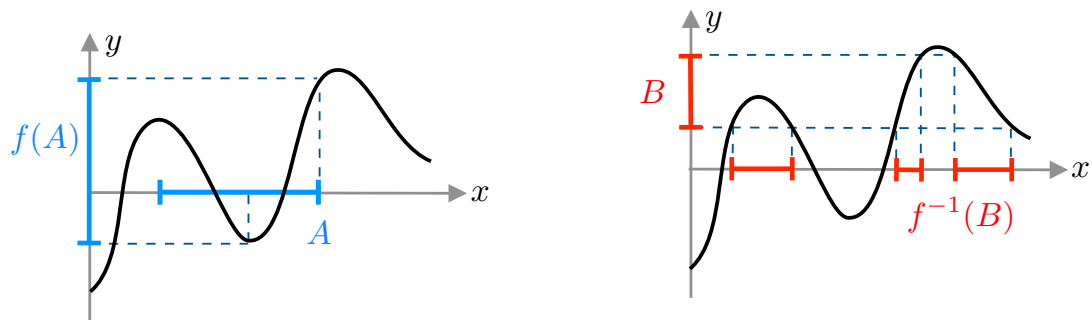
$$f(A) := \{y \in Z : y = f(x) \text{ für mindestens ein } x \in A\} = \{f(x) : x \in A\}$$

das Bild von  $A$  unter  $F$  genannt.

3. Für jede Teilmenge  $B \subset Z$  wird

$$f^{-1}(B) := \{x \in D : f(x) \in B\}$$

das Urbild von  $B$  unter  $f$  genannt. Diese Menge ist immer wohldefiniert, auch wenn die Umkehrabbildung  $f^{-1}$  nicht existiert.

**Beispiele**

**Definition** Für zwei Abbildungen  $f : D \rightarrow Z$  und  $g : Z \rightarrow P$  wird die Abbildung

$$g \circ f : D \rightarrow P \quad \text{mit} \quad (g \circ f)(x) = g(f(x))$$

als die Komposition von  $g$  und  $f$  bezeichnet.

**Bemerkung**

- $g \circ f$  ist nur definiert, wenn der Definitionsbereich von  $g$  eine Teilmenge des Zielbereichs von  $f$  ist.
- Für invertierbare Abbildungen  $f : D \rightarrow Z$  gilt immer

$$f^{-1} \circ f = \text{id}_D \quad f \circ f^{-1} = \text{id}_Z,$$

wobei für jede Menge  $M$  die Abbildung  $\text{id}_M : M \rightarrow M$  mit  $\text{id}_M(z) = z$  die sogenannte Identität ist.

- Sind  $f, g : D \rightarrow D$  zwei Funktionen, so ist die Aussage  $g \circ f = f \circ g$  im Allgemeinen **falsch**. Ein einfaches Gegenbeispiel ergibt sich mit  $D = \mathbb{R}$  sowie  $f(x) = x^2$  und  $g(x) = \sin(x)$ , denn es gilt in der Regel  $\sin(x^2) \neq (\sin(x))^2$ .
- Sind  $f, g, h : D \rightarrow D$  drei Funktionen, so gilt stets  $(h \circ g) \circ f = h \circ (g \circ f)$ .

### 1.3 \*Exkurs über die Unendlichkeit

**Definition** Zwei Mengen  $M$  und  $N$  heißen gleichmächtig, falls es eine bijektive Abbildung  $f : M \rightarrow N$  gibt (und damit auch eine bijektive Abbildung  $g = f^{-1} : N \rightarrow M$ ).

Eine Menge  $M$  heißt abzählbar, wenn sie gleichmächtig zur Menge  $\mathbb{N} = \{1, 2, 3, 4, \dots\}$  der natürlichen Zahlen ist.

**Beispiel** Die Menge der geraden Zahlen  $\mathbb{N}_g = \{2, 4, 6, \dots\}$  ist abzählbar, denn es gibt eine bijektive Abbildung  $f : \mathbb{N}_g \rightarrow \mathbb{N}$ , zum Beispiel:

$m \in \mathbb{N}_g$	2	4	6	8	10	12	14	16	18	20	22	...
$f(m) \in \mathbb{N}$	1	2	3	4	5	6	7	8	9	10	11	...

Oder anders – und etwas salopp – gesagt: Es gibt genauso viele gerade Zahlen wie natürliche Zahlen, eben weil man die geraden Zahlen *abzählen* bzw. *durchnummerieren* kann.

**Beispiel** Die Menge  $\mathbb{Z}$  der ganzen Zahlen ist auch abzählbar, wie die folgende Durchnummerierung zeigt:

$m \in \mathbb{Z}$	...	-4	-3	-2	-1	0	+1	+2	+3	+4	...
$f(m) \in \mathbb{N}$	...	9	7	5	3	1	2	4	6	8	...

**Beispiel** Die Menge  $\mathbb{Q}_+ = \{p/q : p, q \in \mathbb{N}\}$  ist abzählbar, wie das folgende quadratische Schema mit (roter) Durchnummerierung zeigt („-“ verweist auf Leerstellen, bei denen  $p$  und  $q$  nicht teilerfremd sind):

$p$	1	2	3	4	5	6	7
$q$	1	2	3	4	5	6	7
1	1	3	5	9	11	17	21
2	2	-	8	-	16	-	...
3	4	7	-	15	20	...	...
4	6	-	14	-	...	...	...
5	10	13	19	...	...	...	...
6	12	-	...	...	...	...	...
7	18	...	...	...	...	...	...

Mit einem modifizierten Argument kann man zeigen, dass auch  $\mathbb{Q}$  abzählbar ist, d.h. es gibt genauso viele rationale wie natürliche Zahlen.

**Bemerkung** Die Menge der reellen Zahlen ist überabzählbar, d.h. es gibt – salopp gesprochen – wesentlich mehr reelle als natürliche oder rationale Zahlen. Der klassische Beweis benutzt die Dezimaldarstellung reeller Zahlen sowie *Cantors Zweites Diagonalargument* (siehe WIKIPEDIA).

# Kapitel 2

## Zahlen und Zahlbereiche

Vorlesung 03, 25. Oktober 2019

### Zahlbereiche

natürliche Zahlen	$\mathbb{N} = \{1, 2, 3, 4, \dots\}$
	$\mathbb{N}_0 = \{0, 1, 2, 3, 4, \dots\}$
ganze Zahlen	$\mathbb{Z} = \{\dots, -3, -2, -1, 0, 1, 2, 3, \dots\}$
rationale Zahlen	$\mathbb{Q} = \{p/q : p \in \mathbb{Z}, q \in \mathbb{N}\}$
reelle Zahlen	$\mathbb{R}$
komplexe Zahlen	$\mathbb{C}$

**Notationen** Es gibt viele Arten, Zahlen zu schreiben:

1. Brüche für rationale Zahlen:

$$\frac{1}{2} \quad -\frac{7}{3} \quad 4/5$$

Zähler und Nenner sollten teilerfremd sein.

2. Dezimalzahlen für reelle Zahlen:

$$2,4567 \quad 1,3333 \dots \quad -67,34255476456 \dots$$

Rationale Zahlen haben endlich viele oder sich periodisch wiederholende Kommastellen. Irrationale Zahlen (Elemente von  $\mathbb{R} \setminus \mathbb{Q}$ ) haben unendlich viele, nicht-periodische Dezimalstellen.

*Achtung:* Im Englischen (und daher oft auch in der Wissenschaft) schreibt man Punkt statt Komma.

3. Symbolische Schreibweise:

$$\sqrt{2} \quad (\approx 1.414) \quad \text{Kreiszahl } \pi \quad (\approx 3,142) \quad \text{Eulerzahl } e \quad (\approx 2,718)$$

4. Gleitkommazahlen (mit Mantisse und Exponent):

$$1,345\text{E}+06 = 1,345 \cdot 10^6 = 1345000, \quad -4,02\text{E}-4 = -4,02 \cdot 10^{-4} = -0,000402$$

5. Dualzahlen:

$$3 = 11 = 1 \cdot 2^1 + 1 \cdot 2^0, \quad 17 = 1001 = 1 \cdot 2^4 + 1 \cdot 2^0$$

6. Hexadezimalzahlen,  $p$ -adischen Zahlen (siehe Große Übung II)

## 2.1 Natürliche Zahlen

### Beweisprinzip der vollständigen Induktion

**Prinzip** Eine Aussage der Form

$$\forall n \in \mathbb{N} : A(n) \quad (\text{„Für alle natürlichen Zahlen } n \text{ gilt die Aussage } A(n)\text{.“})$$

kann durch folgende zwei Schritte bewiesen werden:

1. Induktionsanfang ( $n = 1$ ): Man zeigt  $A(1)$ .
2. Induktionsschritt ( $n \rightsquigarrow n + 1$ ): Man zeigt die Implikation  $A(n) \Rightarrow A(n + 1)$ , d.h. man folgert aus der Gültigkeit von  $A(n)$  (Induktionsvoraussetzung) die Gültigkeit von  $A(n + 1)$  (Induktionsbehauptung)

$$\xrightarrow{\text{IV}} A(1) \xrightarrow{\text{IS}} A(2) \xrightarrow{\text{IS}} A(3) \xrightarrow{\text{IS}} A(4) \xrightarrow{\text{IS}} \dots$$

Schematische Darstellung des Induktionsprinzips.

**Beispiel** Wir wollen die Formel

$$\sum_{j=1}^n j^3 = 1^3 + 2^3 + \dots + n^3 = \frac{1}{4}n^2(n+1)^2$$

durch vollständige Induktion über  $n$  herleiten.

*Induktionsanfang* ( $n = 1$ ): Einfaches Rechnen liefert  $\sum_{j=1}^n j^3 = 1^3 = 1 = \frac{1}{4} \cdot 1^2 \cdot 2^2$  und damit die Gültigkeit von  $A(1)$ .

*Induktionsschritt* ( $n \rightsquigarrow n + 1$ ):

*Voraussetzung:*  $\sum_{j=1}^n j^3 = \frac{1}{4}n^2(n+1)^2$

*Behauptung:*  $\sum_{j=1}^{n+1} j^3 = \frac{1}{4}(n+1)^2(n+2)^2$

*Beweis:*

$$\begin{aligned} \sum_{j=1}^{n+1} j^3 &= \left( \sum_{j=1}^n j^3 \right) + (n+1)^3 && \text{(Aufspalten der Summe)} \\ &= \frac{1}{4}n^2(n+1)^2 + (n+1)^3 && \text{(nach Induktionsvoraussetzung)} \\ &= \frac{1}{4}(n+1)^2(n^2 + 4(n+1)) && \text{(Ausklammern)} \\ &= \frac{1}{4}(n+1)^2(n^2 + 4n + 4) \\ &= \frac{1}{4}(n+1)^2(n+2)^2 && \text{(binomische Formel)} \end{aligned}$$

□

**Beispiel** Wir wollen den mathematischen Satz

Die Zahl  $2^{3n} - 1$  ist für jedes  $n \in \mathbb{N}$  durch 7 teilbar.

durch vollständige Induktion über  $n$  beweisen.

*Induktionsanfang* ( $n = 1$ ):  $2^3 - 1 = 7$  ist offensichtlich durch 7 teilbar.



*Induktionsschritt* ( $n \rightsquigarrow n + 1$ ):

*Zu zeigen:* Wenn  $2^{3n} - 1$  durch 7 teilbar ist, so ist auch  $2^{3(n+1)} - 1$  durch 7 teilbar.

*Beweis:*

$$\begin{aligned}
 2^{3(n+1)} - 1 &= 8 \cdot 2^{3n} - 1 && \text{(Potenzgesetze)} \\
 &= 8 \cdot (2^{3n} - 1 + 1) - 1 && \text{ („nahrhafte Null“)} \\
 &= 8 \cdot (2^{3n} - 1) + 7 && \text{(Umformen)} \\
 &= 8 \cdot (7 \cdot m) + 7 && \text{(nach Ind.-Voraussetzung für ein } m \in \mathbb{N}\text{)} \\
 &= 7 \cdot (8 \cdot m + 1)
 \end{aligned}$$

□

**Variante des Induktionsprinzips** Zeigt man im Induktionsanfang die Aussage  $A(n_1)$  für ein gegebenes  $n_1 \in \mathbb{N}$  und beweist den Induktionsschritt wie oben, so folgt die Gültigkeit der Aussage  $A(n)$  für alle  $n \geq n_1$ .

## Kombinatorik — Lehre vom Zählen

**Definition** Die Größe

$$\binom{n}{k} := \frac{n!}{k!(n-k)!} = \binom{n}{n-k}$$

wird Binomialkoeffizient genannt.

**Binomischer Lehrsatz** Es gilt

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

wobei man diese Formeln entweder direkt (siehe Große Übung) oder mit vollständiger Induktion über  $n$  (Übungsausgabe) ableiten kann. Beispiele sind:

$$\begin{aligned}
 (a+b)^2 &= a^2 + 2ab + b^2, \\
 (a+b)^3 &= a^3 + 3a^2b + 3ab^2 + b^3, \\
 (a+b)^4 &= a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + b^4
 \end{aligned}$$

**Permutationen** Anordnung von  $n$  unterscheidbaren Objekten in einer Reihe:

$n!$  Möglichkeiten

*Beispiel:* Anordnungen von  $m_1, \dots, m_n$  (Zahlen, Buchstaben, Personen, oder ...)

$$\begin{aligned}
 n = 2 : & \quad m_1, m_2 \quad | \quad m_2, m_1 \\
 n = 3 : & \quad m_1, m_2, m_3 \quad | \quad m_1, m_3, m_2 \quad | \quad m_2, m_1, m_3 \\
 & \quad m_2, m_3, m_1 \quad | \quad m_3, m_1, m_2 \quad | \quad m_3, m_2, m_1
 \end{aligned}$$

**Kombinationen** Wählen von  $k$  Stück aus  $n$  Objekten, ohne Wiederholung und ohne Berücksichtigung der Reihenfolge:

$$\binom{n}{k} \text{ Möglichkeiten}$$

*Anwendung:*  $k$ -elementige Teilmengen der Menge  $\{m_1, \dots, m_n\}$

*Beispiel*

$$\begin{array}{ll} k = 0, n = 3 : & \{\} & \binom{3}{0} = \frac{3!}{0! \cdot 3!} = \frac{6}{1 \cdot 6} = 1 \\ k = 1, n = 3 : & \{m_1\}, \{m_2\}, \{m_3\} & \binom{3}{1} = \frac{6}{1 \cdot 2} = 3 \\ k = 2, n = 3 : & \{m_2, m_3\}, \{m_1, m_3\}, \{m_1, m_2\} & \binom{3}{2} = 3 \\ k = 3, n = 3 : & \{m_1, m_2, m_3\} & \binom{3}{3} = 1 \end{array}$$

**Variationen** Wählen von  $k$  Stück aus  $n$  Objekten, mit Wiederholung und mit Berücksichtigung der Reihenfolge:

$$n^k \text{ Möglichkeiten}$$

*Anwendung:* Worte der Länge  $k$ , die aus Alphabet mit  $n$  Buchstaben gebildet werden können

*Beispiel:*  $k = 3$ , Buchstaben  $\{A, B\}$ :

AAA, AAB, ABA, BAA, ABB, BAB, BBA, BBB

*Anwendung:* Eine Menge  $M = \{x_1, x_2, \dots, x_n\}$  mit  $n$  Elementen besitzt insgesamt  $2^n$  Teilmengen, denn wir können eine Teilmenge dadurch beschreiben, dass wir für jedes Element von  $M$  angeben, ob es dazu gehört oder nicht. Jede Teilmenge wird also eindeutig durch genau  $n$  binäre Entscheidungen charakterisiert.

## Teilbarkeit

**Definition**  $m \in \mathbb{N}$  heißt Teiler von  $n \in \mathbb{N}$  (oder  $n$  ist durch  $m$  teilbar oder  $n$  ist ein Vielfaches von  $m$ ), wenn es ein  $k \in \mathbb{N}$  gibt, so dass  $n = m \cdot k$ . Insbesondere ist jede natürliche Zahl  $n \geq 2$  durch 1 und durch sich selbst teilbar.

Gibt es keine weiteren Teiler, so heißt  $n$  Primzahl :

$$2, 3, 5, 7, 11, 13, 17, 19, 23, 29, 31, 37, 41, 43, \dots$$

Die Zahl 1 ist keine Primzahl.

**Theorem** Jede Zahl  $n$  lässt sich als das Produkt von Primzahlen schreiben. Insbesondere existieren  $k$  Primzahlen  $p_1, p_2, \dots, p_k$  sowie  $k$  Exponenten  $r_1, \dots, r_k \in \mathbb{N}$ , so dass

$$n = p_1^{r_1} \cdot p_2^{r_2} \cdot \dots \cdot p_k^{r_k}$$

Jede Zahl dieser Primzahlen  $p_i$  heißt Primfaktor von  $n$ . Die Primfaktorzerlegung von  $n$  ist bis auf die Vertauschung von Faktoren eindeutig.

*Beweis* siehe [AORS, Seite 22].

**Beispiele**

$$8 = 2^3, \quad 60 = 2^2 \cdot 3 \cdot 5 \quad 7 = 7, \quad 999 = 3^2 \cdot 11$$

**Theorem** Es gibt unendlich viele Primzahlen.

(Indirekter) Beweis: Wir nehmen an, es existieren nur endlich viele Primzahlen, sagen wir  $k$  Stück, und bezeichnen die Primzahlen mit  $q_1, \dots, q_k$ . Wir betrachten außerdem

$$n := q_1 \cdot q_2 \cdot \dots \cdot q_k + 1,$$

und bemerken, dass  $n$  bei Division durch jede der Zahlen  $q_1, \dots, q_k$  den Rest 1 lässt und damit kein Vielfaches dieser Zahl sein kann. Insbesondere kann keiner der Primfaktoren von  $n$  in  $\{q_1, \dots, q_k\}$  enthalten sein. Das ist aber ein Widerspruch zur Annahme, d.h. unsere Annahme muss falsch gewesen sein.  $\square$

**Definition** Für  $n, m \in \mathbb{N}$  nennt man

$$\text{kgV}(n, m) := \min\{k \in \mathbb{N} : k \text{ ist Vielfaches von } n \text{ und } m\}$$

das kleinste gemeinsame Vielfache von  $n$  und  $m$  und

$$\text{ggT}(n, m) := \max\{k \in \mathbb{N} : k \text{ teilt } n \text{ und } m\}$$

den größten gemeinsamen Teiler von  $n$  und  $m$ .

**Beispiel**

$$\text{kgV}(2, 8) = 8, \quad \text{kgV}(3, 13) = 39 \quad \text{kgV}(6, 9) = 18$$

$$\text{ggT}(5, 7) = 1, \quad \text{ggT}(3, 9) = 3 \quad \text{ggT}(6, 15) = 3$$

**Bemerkung** Sind die Primzahlzerlegungen von  $n$  und  $m$  bekannt, so können  $\text{kgV}(n, m)$  und  $\text{ggT}(n, m)$  leicht abgelesen werden. Für

$$525 = 3^1 \cdot 5^2 \cdot 7^1, \quad 180 = 2^2 \cdot 3^2 \cdot 5^1,$$

gilt zum Beispiel

$$\text{kgV}(525, 180) = 2^2 \cdot 3^2 \cdot 5^2 \cdot 7^1 = 6300 \quad \text{ggT}(525, 180) = 3^1 \cdot 5^1 = 15.$$

Insbesondere kann man mit Hilfe der Primfaktorzerlegungen von  $n$  und  $m$  die Formel

$$\text{kgV}(n, m) \cdot \text{ggT}(n, m) = m \cdot n$$

ableiten (siehe [AORS, Seite 23]). Bei großen Zahlen ist es allerdings nicht einfach, die Primfaktorzerlegung zu berechnen.

**Bemerkung**  $\text{ggT}(n, m)$  kann auch mittels des *Euklidischen Algorithmus* durch fortgesetzte Division mit Rest berechnet werden. Für  $n = 3054$  und  $m = 1002$  berechnet man zum Beispiel:

$$\begin{array}{ll} 3054 = 3 \cdot 1002 + 48 & \text{(Wie oft passt 1002 in 3054?)} \\ 1002 = 20 \cdot 48 + 42 & \text{(Wie oft passt 48 in 1002?)} \\ 48 = 1 \cdot 42 + 6 & \text{(Wie oft passt 42 in 48?)} \\ 42 = 7 \cdot 6 + 0 & \text{(Wie oft passt 6 in 42?)} \\ \text{Fertig wegen Rest } 0! & \end{array}$$

Es gilt nun  $\text{ggT}(3054, 1002) = 6$  und damit  $\text{kgV}(3054, 1002) = 3054 \cdot 1002 / 6 = 510018$ .

**\*lineare diophantische Gleichungen** Für gegebene natürliche Zahlen  $a, b, c$  besitzt die Gleichung

$$a \cdot x + b \cdot y = c$$

dann und nur dann eine ganzzahlige Lösungen  $x, y \in \mathbb{Z}$ , wenn  $g := \text{ggT}(a, b)$  ein Teiler von  $c$  ist (Diese Äquivalenzaussage kann mit zahlentheoretischen Argumenten bewiesen werden.). In diesem Fall kann man die Gleichung auch als

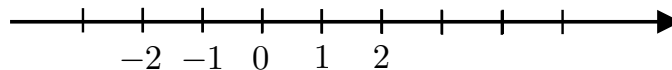
$$\tilde{a} \cdot x + \tilde{b} \cdot y = \tilde{c}$$

mit  $a = \tilde{a} \cdot g, b = \tilde{b} \cdot g, c = \tilde{c} \cdot g$  schreiben, wobei  $\tilde{a}$  und  $\tilde{b}$  teilerfremd sind. Man kann anschließend eine Lösung durch Raten oder durch den dem sogenannten *Erweiterten Euklidischen Algorithmus* (siehe zum Beispiel WIKIPEDIA) finden.

Vorlesung 04, 28. Oktober 2019 2019

## 2.2 Reelle Zahlen

Die reellen Zahlen können als Dezimalzahlen (mit vielleicht unendlich vielen Nachkommastellen) oder als die Punkte auf der Zahlengeraden betrachtet werden.



Die Zahlengerade. Es gilt  $x < y$  genau dann, wenn  $y$  rechts von  $x$  liegt.

### Wichtige Eigenschaften - Axiome

#### (I) Regeln der Addition

$$\begin{array}{ll} \text{(I.1)} & x + (y + z) = (x + y) + z & \text{(Assoziativgesetz)} \\ \text{(I.2)} & x + y = y + x & \text{(Kommutativgesetz)} \\ \text{(I.3)} & x + 0 = x & \text{(0 ist neutrales Element)} \\ \text{(I.4)} & x + (-x) = 0 & \text{(-x ist inverses Element zu x)} \end{array}$$

**(II) Regeln der Multiplikation**

- (II.1)  $x \cdot (y \cdot z) = (x \cdot y) \cdot z$  (Assoziativgesetz)  
 (II.2)  $x \cdot y = y \cdot x$  (Kommutativgesetz)  
 (II.3)  $x \cdot 1 = x$  (1 ist neutrales Element)  
 (II.4)  $x \cdot (1/x) = 1$  ( $1/x$  ist für  $x \neq 0$  inverses Element zu  $x$ )

**Achtung** Man darf **niemals** durch die Zahl 0 teilen!

**(III) Distributivgesetz**

$$(III.1) \quad x \cdot (y + z) = x \cdot y + x \cdot z$$

**(IV) Ordnungseigenschaften**

- (IV.1)  $x \leq x$   
 (IV.2)  $x \leq y \vee y \leq x$   
 (IV.3)  $x \leq y \wedge y \leq x \Rightarrow x = y$   
 (IV.4)  $x \leq y \wedge y \leq z \Rightarrow x \leq z$   
 (IV.5)  $x \leq y \Rightarrow x + z \leq y + z$   
 (IV.6)  $x \leq y \wedge z > 0 \Rightarrow x \cdot z \leq y \cdot z$  (Multipl. mit pos. Zahl)  
 (IV.7)  $x \leq y \wedge z < 0 \Rightarrow x \cdot z \geq y \cdot z$  (Multipl. mit neg. Zahl)

**Notationen**

1.  $x \leq y$  meint „ $x$  kleiner oder gleich  $y$ “
2.  $x < y$  meint „ $x$  kleiner  $y$ “
3.  $x \geq y$  meint „ $x$  größer oder gleich  $y$ “ bzw.  $y \leq x$
4.  $x > y$  meint  $y < x$ .
5.  $x$  heißt positiv bzw. nichtnegativ, falls  $x > 0$  bzw.  $x \geq 0$ .

**(V) Vollständigkeit**  $\mathbb{R}$  ist vollständig, d.h. es gibt keine „Lücken“ in  $\mathbb{R}$ .

Dieses Konzept ist sehr wichtig, aber nicht ganz so einfach zu verstehen. Wir wollen hier mit vorläufigen Erklärungen fortfahren; ein tiefes Verständnis von Vollständigkeit werden wir aber erst später – nämlich beim Studium von Cauchy-Folgen – erreichen.

1.  $\mathbb{Q}$  ist nicht vollständig, denn jede irrationale Zahl (zum Beispiel  $\sqrt{2}$  oder  $\pi$ ) beschreibt eine „Lücke“ in  $\mathbb{Q}$ .
2. Formale Beschreibung: Zerlegt man  $\mathbb{R}$  in zwei Mengen  $A$  und  $B$  mit

$$A \cup B = \mathbb{R}, \quad A \cap B = \emptyset \quad A, B \neq \emptyset, \quad x \leq y \quad \text{für alle } x \in A, y \in B,$$

so gibt es stets genau eine reelle Schnittzahl  $s \in \mathbb{R}$  mit

$$x \leq s \leq y \quad \text{für alle } x \in A, y \in B.$$

3. Ein *Standardbeispiel* ist  $s = \sqrt{2}$  für

$$A = \{x : x \leq 0 \text{ oder } x^2 \leq 2\}, \quad B = \{x : x > 0 \text{ und } x^2 > 2\}.$$

**\*Bemerkung** Sei  $M$  irgendeine Menge und  $+$  irgendeine binäre Verknüpfung — d.h. eine Abbildung  $M \times M \rightarrow M$  — so dass (I.1) – (I.4) gelten, so nennt man  $M$  eine Abelsche Gruppe mit Addition. Gibt es darüber hinaus noch eine Multiplikation, so dass auch noch (II.1) – (II.4) und (III.1) gelten, so wird  $M$  Körper genannt.

Insbesondere ist  $\mathbb{Q}$  ein Körper und erfüllt außerdem sogar die Regeln (IV.1) – (IV.7).  $\mathbb{Q}$  ist aber nicht vollständig.  $\mathbb{C}$  ist hingegen ein Körper und vollständig, besitzt aber keine sinnvolle Ordnungsrelation.

### Nützliche Tatsachen

1. Für jede reelle Zahl  $x > 0$  existieren zwei natürliche Zahlen  $n, m \in \mathbb{N}$ , so dass  $0 < 1/n < x < m$ .
2. Für je zwei reelle Zahlen  $a < b$  existieren unendlich viele rationale Zahlen  $q \in \mathbb{Q}$  mit  $a < q < b$  und unendlich viele irrationale Zahlen  $x$  mit  $a < x < b$ .

**Notation** Doppelungleichungen wie  $a < x < b$  oder  $b \geq x > a$  kodieren **immer zwei** Ungleichungen (nämlich  $a < x$  und  $x < b$  bzw.  $a < x$  und  $x \leq b$ ).

Wir schreiben **niemals**  $a < x > b$  oder  $b \geq x < a$ , da dies nur Verwirrung stiftet.

## Rechnen mit Ungleichungen und Beträgen

**Intervalle** Seien  $a, b$  zwei reelle Zahlen mit  $a < b$ .

1. offenes Intervall:  $(a, b) := \{x \in \mathbb{R} : a < x < b\}$
2. abgeschlossenes Intervall:  $[a, b] := \{x \in \mathbb{R} : a \leq x \leq b\}$
3. halboffene Intervalle gibt es in zwei Varianten:

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

4. Es gibt auch die unbeschränkten Intervalle  $(a, +\infty)$ ,  $[a, +\infty)$  sowie  $(-\infty, a)$ ,  $(-\infty, a]$ .
5. Man schreibt oftmals auch  $\mathbb{R} = (-\infty, +\infty)$ , aber niemals  $\mathbb{R} = [-\infty, +\infty]$  oder  $\mathbb{Q} = (-\infty, +\infty)$ .

**Weitere Ordnungseigenschaften** (die man alle aus den obigen Axiomen ableiten kann).

$$(IV.8) \quad x^2 \geq 0$$

$$(IV.9) \quad x_1 \leq x_2 \wedge y_1 \leq y_2 \Rightarrow x_1 + y_1 \leq x_2 + y_2$$

$$(IV.10) \quad 0 \leq x_1 \leq x_2 \wedge 0 \leq y_1 \leq y_2 \Rightarrow 0 \leq x_1 \cdot y_1 \leq x_2 \cdot y_2$$

### Beispiel

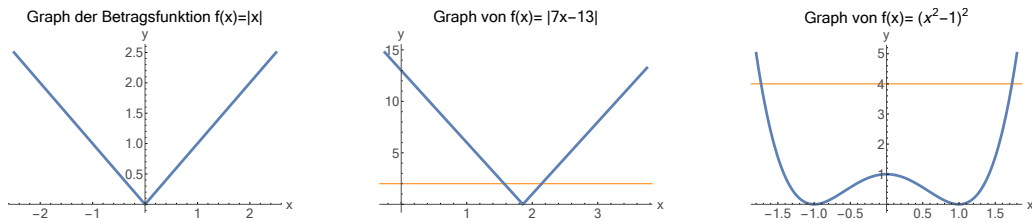
$$1 < 2, \quad 2 \cdot 1 < 2 \cdot 2, \quad (-1) \cdot 1 > (-1) \cdot 2 \quad 1 + 3 < 2 + 5, \quad 1 \cdot 3 < 2 \cdot 5$$

$$-2 < 1, \quad 2 \cdot (-2) < 3 \cdot 1, \quad (-1) \cdot (-2) > (-1) \cdot 1 \quad -2 + 3 < 1 + 5$$

**Definition** Der Betrag (oder Absolutbetrag) von  $x \in \mathbb{R}$  ist

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

und  $|x - y|$  ist der Abstand von  $x$  und  $y$  auf der Zahlengeraden.



Graphen einige Funktionen.

**Eigenschaften der Betragsfunktion** Durch Fallunterscheidung leitet man leicht die folgenden Aussagen für alle  $x \in \mathbb{R}$  ab:

1.  $|x| = |-x| > 0$  für  $x \neq 0$ .
2.  $|x| = \max\{x, -x\}$
3. Es gilt  $\sqrt{x^2} = |x|$ . Die Formel  $\sqrt{x^2} = x$  gilt nur für  $x \geq 0$ .
4. Es gilt  $|x| \leq c$  genau dann, wenn  $-c \leq x \leq c$
5.  $|x \cdot y| = |x| \cdot |y|$
6.  $|x - y| \leq |x - z| + |z - y|$  („Dreiecksungleichung“, Variante 1)
7.  $|x + y| \leq |x + z| + |y + z|$  („Dreiecksungleichung“, Variante 2)
8.  $||x| - |y|| \leq |x - y|$  („Dreiecksungleichung“, Variante 3)

**Bemerkung** Zum Beweis der 3. Variante bemerken wir, dass wegen

$$|x - 0| \leq |x - y| + |y - 0| \quad \text{und} \quad |y - 0| \leq |y - x| + |x - 0|$$

auch

$$|x| - |y| \leq |x - y| \quad \text{und} \quad -|x| + |y| \leq |x - y|$$

gelten. Insbesondere folgt

$$||x| - |z|| \leq \max\{|x| - |y|, |x| - |y|\} \leq |x - y|.$$

□

**Beispiel** Für welche  $x \in \mathbb{R}$  gilt  $|7x - 13| \leq 2$  ?

Lösung 1: Es müssen (siehe 4. Eigenschaft der Betragsfunktion) die beiden Ungleichungen

$$7x - 13 \leq 2 \quad \text{und} \quad 7x - 13 \geq -2$$

gelten und nach einfachen Umformungen beider Ungleichungen erhalten wir

$$|7x - 13| \leq 2 \quad \Leftrightarrow \quad \frac{11}{7} \leq x \quad \text{und} \quad x \leq \frac{15}{7} \quad \Leftrightarrow \quad x \in \left[\frac{11}{7}, \frac{15}{7}\right]$$

Lösung 2: Wir treffen eine Fallunterscheidung.

*Fall 1:* Wir betrachten  $x \in \mathbb{R}$  mit  $7x - 13 \geq 0$ , d.h.  $x \geq \frac{13}{7}$ , wobei auch

$$2 \geq |7x - 13| = 7x - 13 \quad \text{bzw.} \quad x \leq \frac{15}{7}$$

gelten soll. Die Menge der in diesem Fall zulässigen  $x$  ist also gerade das Intervall  $\left[\frac{13}{7}, \frac{15}{7}\right]$ .

*Fall 2:* Für  $7x - 13 \leq 0$  bzw.  $x \leq \frac{13}{7}$  soll

$$2 \geq |7x - 13| = -7x + 13 \quad \text{und damit} \quad x \geq \frac{11}{7}$$

gelten. Insgesamt ergibt sich  $x \in \left[\frac{11}{7}, \frac{15}{7}\right]$  für den zweiten Fall.

Durch Kombination der Teilresultate erhalten wir wieder dieselbe Antwort wie im ersten Lösungsweg.

Lösung 3: Man kann (in diesem einfachen Beispiel) die Lösung auch graphisch bestimmen, siehe dazu das Bild oben.

**Beispiel** Wie sieht die Menge  $M = \{x \in \mathbb{R} : (x^2 - 1)^2 < 4\}$  aus ?

Lösung 1:  $x$  ist genau dann Element von  $M$ , wenn

$$-2 < x^2 - 1 < 2 \quad \text{bzw.} \quad -1 < x^2 < 3.$$

Wegen  $x^2 \geq 0$  schließen wir, dass  $M$  gerade das offene Intervall  $(-\sqrt{3}, \sqrt{3})$  ist.

Lösung 2: Dasselbe Resultat mittels Fallunterscheidungen.

*Fall 1:* Unter der Annahme  $x^2 - 1 \geq 0$  soll  $x^2 - 1 < 2$  gelten, und beide Bedingungen liefern  $1 \leq x^2 < 3$ . Die Menge der zulässigen  $x$  ist also gerade die Vereinigung der Intervalle  $(-\sqrt{3}, -1]$  und  $[1, \sqrt{3})$

*Fall 2:* Für  $x^2 - 1 < 0$  soll  $1 - x^2 < 2$  gelten, und beide Bedingungen werden gerade durch  $x \in (-1, +1)$  erfüllt.

Insgesamt ergibt sich  $M = (-\sqrt{3}, -1] \cup (-1, +1) \cup [1, \sqrt{3}) = (-\sqrt{3}, \sqrt{3})$ .

Lösung 3: Es gibt auch wieder die Möglichkeit, die Lösung graphisch zu bestimmen.

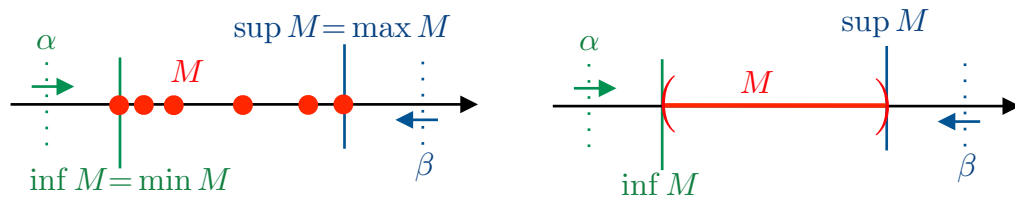
## Schranken

**Definition** Sei  $M \subset \mathbb{R}$  beliebig, aber nicht leer.

1.  $\alpha \in \mathbb{R}$  heißt (eine) untere Schranke von  $M$ , falls  $\alpha \leq x$  für alle  $x \in M$  gilt.  
 $\beta \in \mathbb{R}$  heißt (eine) obere Schranke von  $M$ , falls  $x \leq \beta$  für alle  $x \in M$  gilt.



2. Existieren sowohl eine untere als auch eine obere Schranke, so wird  $M$  beschränkt genannt.
3. Ist  $\alpha \in \mathbb{R}$  die *größte untere Schranke* von  $M$ , so heißt  $\alpha$  das Infimum von  $M$  und wird mit  $\alpha = \inf M$  abgekürzt.  
Ist  $\beta \in \mathbb{R}$  die *kleinste obere Schranke* von  $M$ , so heißt  $\beta$  das Supremum von  $M$  und wird mit  $\beta = \sup M$  abgekürzt.
4. Im Fall von  $\alpha = \inf M$  und  $\alpha \in M$ , nennen wir  $\alpha$  das Minimum von  $M$  und schreiben  $\alpha = \min M$ .  
Im Fall von  $\beta = \sup M$  und  $\beta \in M$ , nennen wir  $\beta$  das Maximum von  $M$  und schreiben  $\beta = \max M$ .



Zur Bedeutung von Schranken und von inf, sup, max, min.

**Beispiel** Für jedes offene/halboffene/abgeschlossene Intervall  $M$  mit Randpunkten  $a < b$  ist jedes  $\alpha < a$  eine untere Schranke und jedes  $\beta > b$  eine obere Schranke. Die größte untere Schranke ist dabei immer  $\inf M = a$ , die kleinste obere Schranke immer  $\sup M = b$ . Darüber hinaus gilt:

1. Für  $M = [a, b]$  existieren das Maximum und das Minimum,
2. Für  $M = [a, b)$  existiert nur das Minimum,
3. Für  $M = (a, b)$  existieren weder Minimum noch Maximum.

**Beispiel** Für  $M = \{1/n : n \in \mathbb{N}\}$  gilt

$$\sup M = \max M = 1, \quad \inf M = 0,$$

aber das Infimum gehört nicht zu  $M$  und ist daher kein Minimum.

**Bemerkung**

1. Für eine Menge mit endlich vielen Elementen ( $0 < \#M < \infty$ ) existieren stets  $\min M$  und  $\max M$ .
2. Gibt es eine untere (bzw. obere) Schranke, so ist  $\inf M$  bzw.  $\sup M$  wohldefiniert. Andernfalls schreiben wir  $\inf M = -\infty$  (bzw.  $\sup M = +\infty$ ).

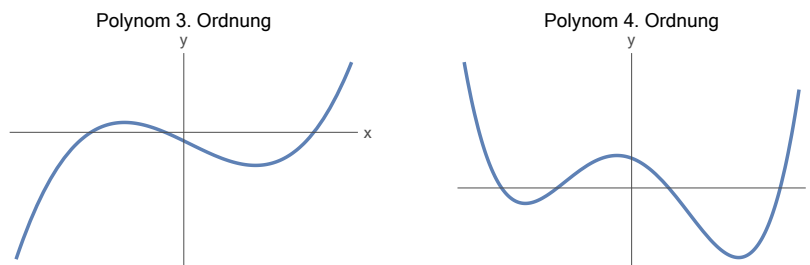
**Beispiel**  $\sup \mathbb{N} = \infty$ ,  $\min \mathbb{N}_0 = \inf \mathbb{N}_0 = 0$ ,  $\inf \mathbb{Z} = -\infty$ ,  $\sup \mathbb{Z} = +\infty$ .

## Wichtige Funktionen

**Definition** Ein (reelles) Polynom vom Grad  $n$  ist eine Funktion  $f : \mathbb{R} \rightarrow \mathbb{R}$  der Bauart

$$f(x) = \alpha_n \cdot x^n + \alpha_{n-1} \cdot x^{n-1} + \dots + \alpha_1 \cdot x^1 + \alpha_0 \cdot x^0,$$

wobei alle Koeffizienten  $\alpha_i$  reelle Zahlen sind und  $\alpha_n \neq 0$  gilt. Beachte, dass immer  $x^1 = x$  und  $x^0 = 1$  gilt und das Polynome immer auf ganz  $\mathbb{R}$  definiert sind. Ein Polynom vom Grad 1 oder 0 wird auch affine Funktion genannt.

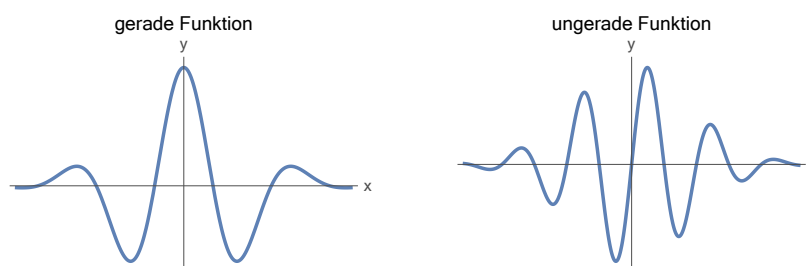


Zwei typische Polynome.

**Definition** Eine Funktion  $f : \mathbb{R} \rightarrow \mathbb{R}$  heißt gerade bzw. ungerade, falls

$$f(x) = f(-x) \quad \text{bzw.} \quad f(x) = -f(-x) \quad (2.1)$$

für alle  $x \in \mathbb{R}$  gilt.

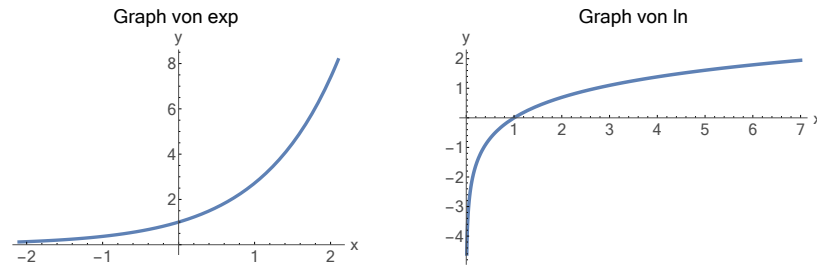


Zwei typische Beispiele für gerade und ungerade Funktionen.

**Exponentialfunktion** Die (reelle) Exponentialfunktion  $\exp : \mathbb{R} \rightarrow (0, \infty)$  ist durch Potenzen der Eulerzahl

$$\exp(x) = e^x, \quad e := \lim_{n \rightarrow \infty} (1 + 1/n)^n \approx 2,71828$$

definiert, wobei die ganze Definition später gegeben wird. Die Umkehrfunktion – die für alle positive Argumente definiert ist – wird natürlicher Logarithmus genannt, und mit  $\ln : (0, \infty) \rightarrow \mathbb{R}$  bezeichnet.



Die Exponentialfunktion und ihre Umkehrabbildung.

*Rechenregeln:*

1.  $\ln(\exp(x)) = x$
2.  $\exp(\ln(y)) = y$  (für  $y > 0$ )
3.  $\exp(x_1 + x_2) = \exp(x_1) \exp(x_2)$
4.  $(\exp(x))^r = \exp(r \cdot x)$  (für  $r > 0$ )
5.  $\ln(y_1 \cdot y_2) = \ln(y_1) + \ln(y_2)$
6.  $\ln(y^r) = r \cdot \ln(y)$
7.  $\exp(0) = 1, \quad \ln(1) = 0$

*Andere Logarithmen:* Für jedes  $a > 0$  wird die Umkehrfunktion von  $x \mapsto a^x$  der Logarithmus zur Basis  $a$  genannt und mit  $\log_a : (0, \infty) \rightarrow \mathbb{R}$  bezeichnet. Wegen

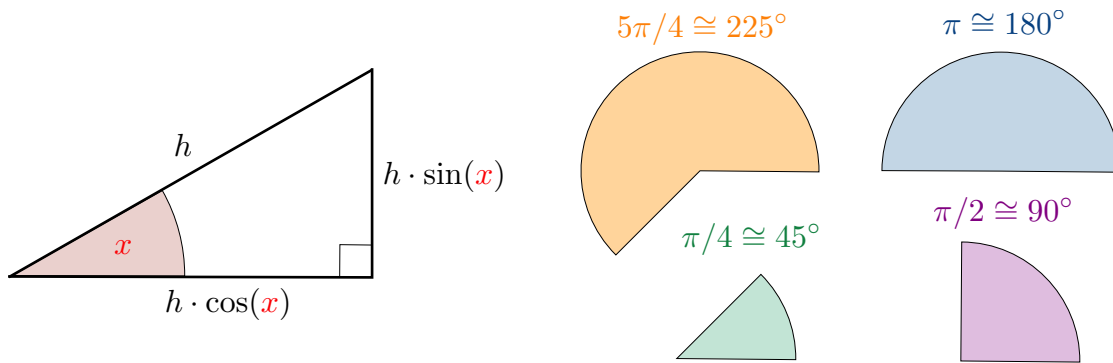
$$a^x = \exp(\ln(a) \cdot x)$$

gilt stets

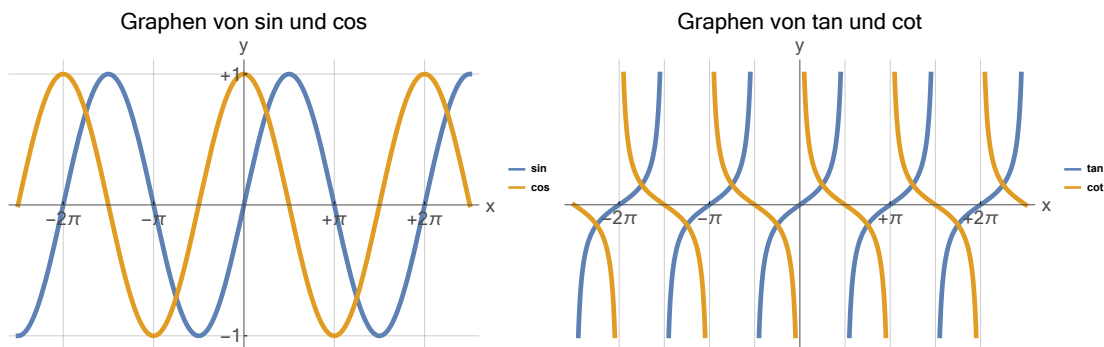
$$\log_a(x) = \frac{\ln(x)}{\ln(a)}.$$

**Trigonometrische Funktionen** Die Funktionen  $\sin$  (Sinus) und  $\cos$  (Kosinus) sind auf ganz  $\mathbb{R}$  definiert und dort  $2\pi$ -periodisch, d.h.  $\sin(x) = \sin(x + 2\pi)$  und  $\cos(x) = \cos(x + 2\pi)$ . Sie haben eine unmittelbare geometrische Bedeutung, sofern  $x$  als Winkel im Bogenmaß (oder Radian) interpretiert wird. Insbesondere entspricht  $2\pi$  dem Vollkreis bzw.  $360^\circ$ .

	$x = 0$ ( $0^\circ$ )	$x = \pi/6$ ( $30^\circ$ )	$x = \pi/4$ ( $45^\circ$ )	$x = \pi/3$ ( $60^\circ$ )	$x = \pi/2$ ( $90^\circ$ )
$\sin(x)$	$\frac{1}{2}\sqrt{0} = 0$	$\frac{1}{2}\sqrt{1} = 1/2$	$\frac{1}{2}\sqrt{2} = 1/\sqrt{2}$	$\frac{1}{2}\sqrt{3} = \sqrt{3}/2$	$\frac{1}{2}\sqrt{4} = 1$
$\cos(x)$	$\frac{1}{2}\sqrt{4} = 1$	$\frac{1}{2}\sqrt{3} = \sqrt{3}/2$	$\frac{1}{2}\sqrt{2} = 1/\sqrt{2}$	$\frac{1}{2}\sqrt{1} = 1/2$	$\frac{1}{2}\sqrt{0} = 0$



Zum Bogenmaß und der elementargeometrischen Bedeutung von sin und cos.



Die trigonometrischen Funktionen.

Darüber hinaus werden durch

$$\tan(x) = \frac{\sin(x)}{\cos(x)}, \quad \cot(x) = \frac{\cos(x)}{\sin(x)}$$

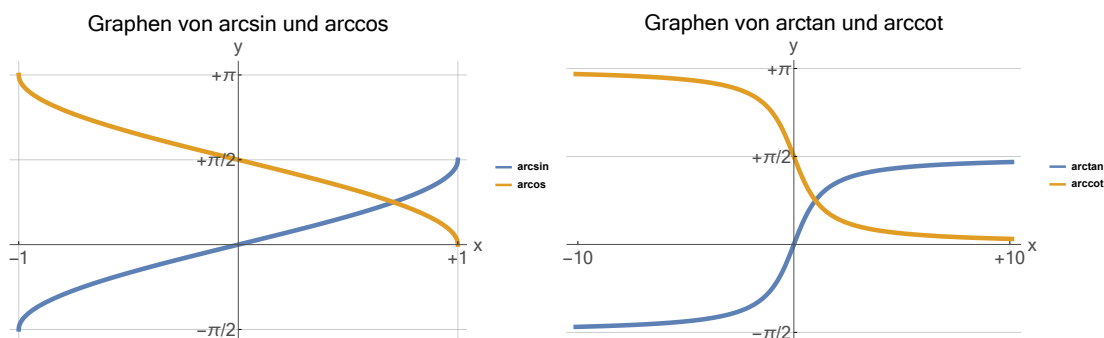
der Tangens und der Kotangens definiert. Diese Funktionen sind  $\pi$ -periodisch, aber nicht auf ganz  $\mathbb{R}$  definiert, sondern besitzen bei  $x = (\frac{1}{2} + m)\pi$  ( $m \in \mathbb{Z}$ , Nullstellen von cos) bzw.  $x = m\pi$  ( $m \in \mathbb{Z}$ , Nullstellen von sin) sogenannte Polstellen.

Die „Umkehrfunktionen“ der trigonometrischen Funktionen (Arkussinus usw. genannt) existieren nur in einem lokalen Sinne:

$$\arcsin : [-1, +1] \rightarrow [-\frac{1}{2}\pi, +\frac{1}{2}\pi], \quad \arccos : [-1, +1] \rightarrow [0, \pi],$$

und

$$\arctan : (-\infty, +\infty) \rightarrow (-\frac{1}{2}\pi, +\frac{1}{2}\pi), \quad \operatorname{arccot} : (-\infty, +\infty) \rightarrow (0, \pi).$$



Die inversen trigonometrischen Funktionen.

## 2.3 Komplexe Zahlen

**Motivation** Komplexe Zahlen sind extrem wichtig für das Studium der Elektrotechnik und der Mathematik.

### Definition und Rechenregeln

**Postulat** Es gibt eine Zahl  $i$  (die sogenannte imaginäre Einheit) mit

$$i^2 = i \cdot i = -1,$$

aber eine solche Zahl existiert nicht in  $\mathbb{R}$ .

**Anmerkung** In früheren Zeiten wurde viel über die philosophische und metaphysische Natur von Zahlen spekuliert und man dachte,  $i$  sei nur eine „imaginäre“, also „gedachte“ oder „virtuelle“ Größe, die zwar in unseren Gedanken, aber nicht in der Wirklichkeit existiert. Heute weiß man es besser: Komplexe und imaginäre Zahlen sind genauso „real“ wie die reellen Zahlen und eine Beschreibung der physikalischen oder technischen Welt ist ohne Verwendung komplexer Zahlen eigentlich unmöglich.

**Definition** Eine komplexe Zahl ist ein Ausdruck der Form

$$z = x + i \cdot y,$$

wobei  $x = \operatorname{Re}(z) \in \mathbb{R}$  der Realteil und  $y = \operatorname{Im}(z) \in \mathbb{R}$  der Imaginärteil von  $z$  sind. Die Menge aller komplexen Zahl wird mit  $\mathbb{C}$  bezeichnet.

**Addition und Multiplikation** Die Addition komplexer Zahlen wird ganz intuitiv eingeführt:

$$(x_1 + i \cdot y_1) + (x_2 + i \cdot y_2) = (x_1 + x_2) + i \cdot (y_1 + y_2)$$

Gleiches gilt für die Multiplikation, wobei allerdings schon  $i^2 = i \cdot i = -1$  benutzt wird:

$$\begin{aligned} (x_1 + i \cdot y_1) \cdot (x_2 + i \cdot y_2) &= x_1 \cdot x_2 + x_1 \cdot i \cdot y_2 + i \cdot x_2 \cdot y_1 + i \cdot y_1 \cdot i \cdot y_2 \\ &= x_1 \cdot x_2 + i \cdot x_1 \cdot y_2 + i \cdot y_1 \cdot x_2 + i^2 \cdot y_1 \cdot y_2 \\ &= (x_1 \cdot x_2 - y_1 \cdot y_2) + i \cdot (x_1 \cdot y_2 + x_2 \cdot y_1) \end{aligned}$$

**Definition** Für jedes  $z = x + y \cdot i \in \mathbb{C}$  heißt

$$\bar{z} := x - i \cdot y$$

die zu  $z$  konjugiert komplexe Zahl. Der Ausdruck

$$\sqrt{z \cdot \bar{z}} = \sqrt{x^2 - y^2 \cdot i^2 + i \cdot (x \cdot y - x \cdot y)} = \sqrt{x^2 + y^2} =: |z|$$

ist immer wohldefiniert (Wurzel einer nichtnegativen reellen Zahl) und wird Betrag von  $z$  genannt. Insbesondere gilt  $|z| > 0$  für jede komplexe Zahl  $z \neq 0 = 0 + 0 \cdot i$ .

**Subtraktion und Division** Zwei komplexe Zahlen können via

$$(x_1 + i \cdot y_1) - (x_2 + i \cdot y_2) = (x_1 - x_2) + i \cdot (y_1 - y_2)$$

ganz einfach subtrahiert werden. Die Division erfolgt durch Erweitern mit der konjugiert komplexen Zahl des Nenners via

$$\begin{aligned} \frac{x_1 + i \cdot y_1}{x_2 + i \cdot y_2} &= \frac{x_1 + i \cdot y_1}{x_2 + i \cdot y_2} \cdot \frac{x_2 - i \cdot y_2}{x_2 - i \cdot y_2} \\ &= \frac{(x_1 + i \cdot y_1) \cdot (x_2 - i \cdot y_2)}{x_2^2 + y_2^2} \\ &= \frac{x_1 \cdot x_2 + y_1 \cdot y_2}{x_2^2 + y_2^2} + i \cdot \frac{-x_1 \cdot y_2 + x_2 \cdot y_1}{x_2^2 + y_2^2}. \end{aligned}$$

Damit diese Formel Sinn hat, muss  $x_2 + i \cdot y_2 \neq 0$  gelten, d.h. wie dürfen auch im Komplexen nicht durch 0 teilen.

**Theorem** Es gelten die üblichen Rechenregeln der Addition und Multiplikation (siehe (I) – (III) aus Kapitel 2.2). Insbesondere ist  $\mathbb{C}$  ein sogenannter Körper.

*Beweis:* Alle Rechenregeln für reelle Zahlen können durch Nachrechnen auf die komplexen Zahlen übertragen werden.  $\square$

### **Bemerkung**

1. Jede reelle Zahl kann via  $x = x + i \cdot 0$  als komplexe Zahl aufgefasst werden. In diesem Sinne gilt  $\mathbb{R} \subset \mathbb{C}$ .
2. Eine komplexe Zahl der Bauart  $0 + i \cdot y$  heißt rein-imaginäre Zahl.
3. Die Zahl  $z = x + i \cdot y$  kann auch als  $x + y \cdot i$  oder  $i \cdot y + x$  oder  $y \cdot i + x$  geschrieben werden.
4. Die Formeln

$$\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}, \quad \overline{z_1 \cdot z_2} = \overline{z_1} \cdot \overline{z_2}$$

können leicht nachgerechnet werden (Übungsaufgabe).

5. Es gilt die komplexe Dreiecksungleichung

$$|z_1 + z_2| \leq |z_1| + |z_2|,$$

deren geometrische Bedeutung wir bald verstehen werden.

**Achtung** Es gibt keine sinnvolle Ordnung in  $\mathbb{C}$ , d.h. „ $\leq$ “ und „ $\geq$ “ sind nur für reelle, aber nicht für allgemeine komplexe Zahlen definiert. Insbesondere meint  $z \geq 0$  immer auch  $z \in \mathbb{R}$ .

**Beispiel** Man vereinfache die Formel

$$z = -2 \cdot i + \frac{4 + 5 \cdot i}{(1 + i)^2}$$

indem man den Realteil und den Imaginärteil von  $z$  explizit angibt.

Lösung: Wir betrachten zunächst einen Hilfsbruch und erweitern mit der konjugiert komplexen Zahl des Zählers:

$$\frac{1}{(1 + i)^2} = \frac{(1 - i)^2}{(1 + i)^2 \cdot (1 - i)^2} = \frac{(1 - i)^2}{(1 - i^2)^2} = \frac{(1 - i)^2}{4} = \frac{1 + i^2 - 2 \cdot i}{4} = -\frac{1}{2} \cdot i.$$

Durch Einsetzen erhalten wir

$$z = -2 \cdot i + (4 + 5 \cdot i) \cdot \left(-\frac{1}{2} \cdot i\right) = \frac{5}{2} - 4 \cdot i.$$

□

**quadratische Gleichungen mit reellen Koeffizienten** Jede quadratische Gleichung der Form

$$z^2 + p \cdot z + q = 0$$

mit reellen Koeffizienten  $p, q \in \mathbb{R}$  besitzt zwei komplexe Lösungen  $z_-, z_+ \in \mathbb{C}$ , für die die verallgemeinerte  $p$ - $q$ -Formel gilt:

$$z_{\pm} = \begin{cases} -\frac{p}{2} \pm \sqrt{\frac{p^2}{4} - q} & \text{falls } p^2 > 4 \cdot q, \\ -\frac{p}{2} & \text{falls } p^2 = 4 \cdot q, \\ -\frac{p}{2} \pm i \cdot \sqrt{q - \frac{p^2}{4}} & \text{falls } p^2 < 4 \cdot q. \end{cases}$$

Beweis: Die ersten beiden Fälle sind aus der Schule bekannt. Im dritten Fall rechnen wir:

$$\begin{aligned} (z_{\pm})^2 &= \left(-\frac{p}{2} \pm i \cdot \sqrt{q - \frac{p^2}{4}}\right)^2 \\ &= \left(\frac{p^2}{4} + i^2 \cdot \left(q - \frac{p^2}{4}\right)\right) - 2 \cdot \frac{p}{2} \cdot \left(\pm i \cdot \sqrt{q - \frac{p^2}{4}}\right) \\ &= \left(\frac{p^2}{2} - q\right) - p \cdot \left(z_{\pm} + \frac{p}{2}\right) \\ &= -p \cdot z_{\pm} - q \end{aligned}$$

□

**Achtung** Wir schreiben  $\sqrt{x}$  nur für Argumente  $x \in \mathbb{R}$  mit  $x \geq 0$ . Für komplexe Zahlen gibt es nämlich zwei denkbare Wurzeln (siehe auch den Fundamentalsatz der Algebra unten).

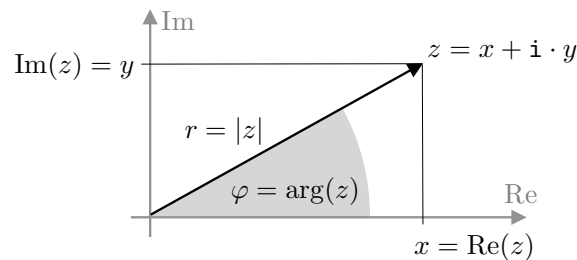
Allgemein gilt: Ganzzahlige Potenzen komplexer Zahlen sind (zumindest für  $z \neq 0$ ) harmlos und es gilt

$$\dots, \quad z^{-2} = \frac{1}{z \cdot z}, \quad z^{-1} = \frac{1}{z}, \quad z^0 = 1, \quad z^1 = z, \quad z^2 = z \cdot z, \quad z^3 = z \cdot z \cdot z, \quad \dots$$

Potenzen mit komplexer Basis und nichtganzzahligen Exponenten — wie zum Beispiel  $z^{1/2}$ ,  $z^{3/2}$ ,  $z^{-4/5}$ ,  $z^\pi$  oder gar  $z^i$  — sind mehrdeutig und werden daher **nicht** benutzt! Den Ursprung der Mehrdeutigkeit werden wir unten bzw. im dritten Semester beim Studium der *Komplexen Analysis* richtig verstehen.

## Geometrische Bedeutung

**Beobachtung** Zu jeder komplexen Zahl  $z = x + i \cdot y$  gehört genau ein Punkt  $(x, y)$  in der Zahlenebene  $\mathbb{R}^2$  und umgekehrt.



Komplexe Zahlen sind Punkte in der Ebene.

**Komplexe Zahlenebene** Jeden Punkt  $(x, y) \neq (0, 0)$  kann man als

$$x = r \cdot \cos(\varphi), \quad y = r \cdot \sin(\varphi)$$

schreiben, wobei die Länge  $r$  reell und positiv ist und der Winkel  $\varphi$  nur bis auf Vielfache von  $2\pi$  eindeutig bestimmt ist. Die Länge  $r$  ist gerade der Betrag der komplexen Zahl  $z = x + i \cdot y$  und jeder Winkel  $\varphi = \arg(z)$  heißt Argument von  $z$ .

Man nennt  $x, y$  die kartesischen Koordinaten von  $z = x + i \cdot y$  und  $r, \varphi$  die entsprechenden Polarkoordinaten. Die Berechnung von  $r$  und  $\varphi$  aus  $x$  und  $y$  gelingt mit den folgenden Formeln

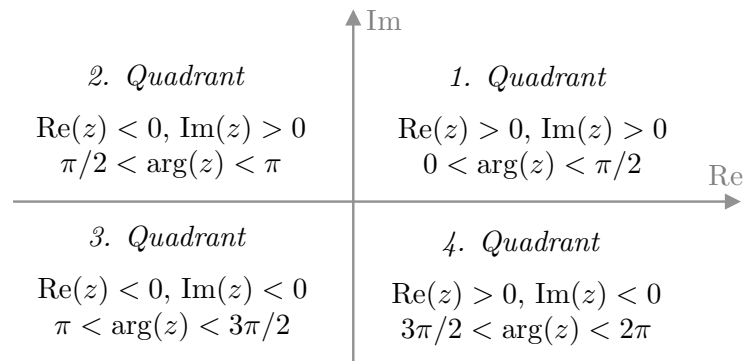
$$r = |z| = \sqrt{x^2 + y^2}$$

und

$$\varphi = \arg(z) = \begin{cases} 0 & \text{falls } x > 0, y = 0, \\ \arctan(y/x) & \text{falls } x > 0, y > 0, & (1. \text{ Quadrant}) \\ \pi/2 & \text{falls } x = 0, y > 0, \\ \pi + \arctan(y/x) & \text{falls } x < 0, y > 0, & (2. \text{ Quadrant}) \\ \pi & \text{falls } x < 0, y = 0, \\ \pi + \arctan(y/x) & \text{falls } x < 0, y < 0, & (3. \text{ Quadrant}) \\ 3\pi/2 & \text{falls } x = 0, y < 0, \\ 2\pi + \arctan(y/x) & \text{falls } x > 0, y < 0. & (4. \text{ Quadrant}) \end{cases}$$



wobei immer ein Winkel  $0 \leq \varphi < 2\pi$  geliefert wird und die Arkustangens-Funktion  $\arctan : \mathbb{R} \rightarrow (-\pi/2, \pi/2)$  in Kapitel 2.2 skizziert ist.

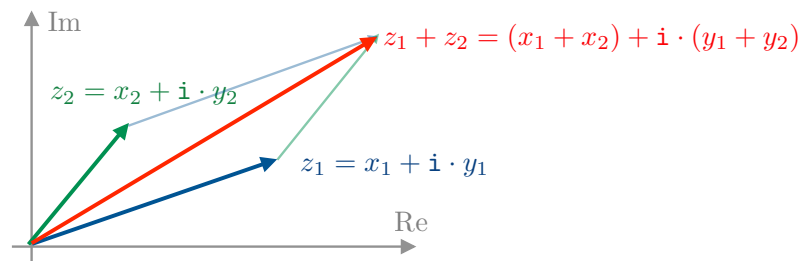


Einige Formeln für komplexe Zahlen hängen explizit davon ab, in welchen Quadranten die komplexe Zahl  $z \in \mathbb{C}$  liegt. Die vier Halbachsen beschreiben dabei gerade die Grenze zwischen je zwei Quadranten und alle Quadranten berühren sich im Punkte  $z = 0$ .

**Achtung** Für  $x = 0$  und  $y = 0$  (bzw. für  $z = 0$ ) gilt  $r = 0$ , aber  $\varphi$  ist nicht definiert.

**nochmal Addition** Die Summation zweier komplexer Zahlen ist gerade die übliche Vektoraddition in der Zahlenebene, denn die Gesetze der Multiplikation garantieren

$$\operatorname{Re}(z_1 + z_2) = \operatorname{Re}(z_1) + \operatorname{Re}(z_2), \quad \operatorname{Im}(z_1 + z_2) = \operatorname{Im}(z_1) + \operatorname{Im}(z_2).$$



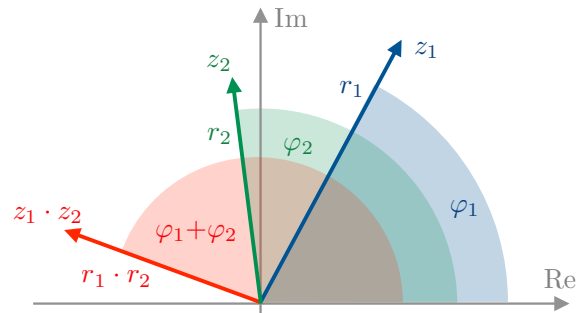
Komplexe Zahlen werden wie Vektoren in der Ebene addiert.

**nochmal Multiplikation** Für das Produkt zweier komplexer Zahlen erhalten wir in Polarkoordinaten die Formelkette

$$\begin{aligned} z_1 \cdot z_2 &= \left( r_1 \cdot \cos(\varphi_1) + i \cdot r_1 \cdot \sin(\varphi_1) \right) \cdot \left( r_2 \cdot \cos(\varphi_2) + i \cdot r_2 \cdot \sin(\varphi_2) \right) \\ &= r_1 \cdot r_2 \cdot \left( \cos(\varphi_1) \cdot \cos(\varphi_2) - \sin(\varphi_1) \cdot \sin(\varphi_2) \right) + \\ &\quad + i \cdot r_1 \cdot r_2 \cdot \left( \sin(\varphi_1) \cdot \cos(\varphi_2) + \cos(\varphi_1) \cdot \sin(\varphi_2) \right) \\ &= r_1 \cdot r_2 \cdot \left( \cos(\varphi_1 + \varphi_2) \right) + i \cdot r_1 \cdot r_2 \cdot \left( \sin(\varphi_1 + \varphi_2) \right), \end{aligned}$$

wobei wir zwei Additionstheoreme verwendet haben (siehe dazu weiter unten). Wir haben damit gezeigt:

$$|z_1 \cdot z_2| = |z_1| \cdot |z_2|, \quad \arg(z_1 \cdot z_2) = \arg(z_1) + \arg(z_2).$$



Die Multiplikation komplexer Zahlen kann als Summation der Argumente (Winkel) und Multiplikation der Beträge (Längen) interpretiert werden.

**Bemerkung** Mit  $z = r \cdot (\cos(\varphi) + i \cdot \sin(\varphi))$  gilt

$$\bar{z} = r \cdot (\cos(\varphi) - i \cdot \sin(\varphi)), \quad z^{-1} = r^{-1} \cdot (\cos(\varphi) - i \cdot \sin(\varphi)).$$

Vorlesung 06, 01. November 2019

## Potenzreihen und Euler-Formel

**Komplexe Exponentialfunktion** Wir werden im Laufe der Vorlesung sehen, dass die folgende Reihe (=unendliche Summe)

$$\exp(z) := \sum_{k=0}^{\infty} \frac{z^k}{k!} = 1 + z + \frac{z^2}{2} + \frac{z^3}{6} + \frac{z^4}{24} + \frac{z^5}{120} + \frac{z^6}{720} + \dots$$

für alle  $z \in \mathbb{C}$  wohldefiniert ist. Für reelle Argumente  $z = x + 0 \cdot i \in \mathbb{R}$  kann man außerdem zeigen, dass

$$\exp(x) = e^x$$

gilt, d.h. für  $x \in \mathbb{R}$  liefert die Reihe gerade die  $x$ -te Potenz der Eulerzahl  $e$ . Man schreibt auch für komplexe Argumente oftmals  $e^z$  statt  $\exp(z)$ .

**Bemerkung** Auch im Komplexen gilt

$$\exp(z_1 + z_2) = \exp(z_1) \cdot \exp(z_2).$$

(Wir werden später sehen, dass das ein Spezialfall der Cauchyschen Produktformel für Reihen ist).

**Komplexer Sinus und Kosinus** Man kann durch

$$\begin{aligned} \sin(z) &:= \sum_{m=0}^{\infty} (-1)^m \frac{z^{2m+1}}{(2m+1)!} &= z - \frac{z^3}{3!} + \frac{z^5}{5!} - \frac{z^7}{7!} + \dots \\ \cos(z) &:= \sum_{m=0}^{\infty} (-1)^m \frac{z^{2m}}{(2m)!} &= 1 - \frac{z^2}{2!} + \frac{z^4}{4!} - \frac{z^6}{6!} + \dots \end{aligned}$$

komplexe Varianten vom Sinus und Kosinus definieren, die für reelle Argumente mit den aus der Schule bekannten Funktionen (siehe auch Kapitel 2.2) übereinstimmen. (Wir werden später verstehen, warum das wirklich so ist).

**Bemerkung** Die obigen Reihen haben für jedes  $z \in \mathbb{C}$  Sinn, d.h. wir können  $\sin$  und  $\cos$  als Funktionen  $\mathbb{C} \rightarrow \mathbb{C}$  betrachten. Außerdem gilt immer

$$\sin(-z) = -\sin(z), \quad \cos(-z) = \cos(z), \quad (\sin(z))^2 + (\cos(z))^2 = 1.$$

**Theorem** Die Eulersche Formel

$$\exp(i \cdot z) = \cos(z) + i \cdot \sin(z)$$

gilt für alle  $z \in \mathbb{C}$ .

Formaler Beweis: Wir setzen voraus, dass wir die Reihen beliebig umsortieren dürfen (wir werden später diskutieren, warum das wirklich möglich ist) und rechnen:

$$\begin{aligned} \sum_{k=0}^{\infty} \frac{(i)^k \cdot z^k}{k!} &= \sum_{k \in \mathbb{N}_0 \text{ gerade}} \dots + \sum_{k \in \mathbb{N}_0 \text{ ungerade}} \dots \\ &= \sum_{k=2m, m \in \mathbb{N}_0} \dots + \sum_{k=2m+1, m \in \mathbb{N}_0} \dots \\ &= \sum_{m=0}^{\infty} \frac{(i^2)^m \cdot z^{2m}}{(2m)!} + \sum_{m=0}^{\infty} \frac{i \cdot (i^2)^m \cdot z^{2m+1}}{(2m+1)!} \\ &= \cos(x) + i \cdot \sin(z) \end{aligned}$$

□

**Bemerkung**

1. Ein Spezialfall ist die Eulersche Identität

$$e^{i \cdot \pi} + 1 = 0,$$

die die wichtigsten Zahlen der Mathematik in einer Formel vereint.

2. Die Euler-Formel impliziert (Nachrechnen!)

$$\sin(z) = \frac{e^{+i \cdot z} - e^{-i \cdot z}}{2 \cdot i}, \quad \cos(z) = \frac{e^{+i \cdot z} + e^{-i \cdot z}}{2}.$$

3. Wegen  $e^{i \cdot 2\pi} = 1$  gilt weiterhin immer

$$\sin(z) = \sin(z + 2\pi), \quad \cos(z) = \cos(z + 2\pi)$$

allerdings können Sinus und Kosinus im Komplexen auch nicht reelle Werte bzw. reelle Werte außerhalb von  $[-1, +1]$  annehmen.

4. Die Multiplikation von  $z = x + i \cdot y$  mit  $e^{i \cdot \alpha}$  ist für  $\alpha \in \mathbb{R}$  äquivalent zur Drehung des Vektors  $(x, y)$  um den Winkel  $\alpha$ .
5. Die Menge  $\{e^{i \cdot \alpha} : \alpha \in \mathbb{R}\}$  ist der Kreis vom Radius 1 um den Koordinatenursprung (der sogenannte Einheitskreis).

6. Potenzen der Bauart  $a^z$  mit positiver reeller Basis  $a > 0$  und Exponent  $z = x + i \cdot y$  sind via

$$\begin{aligned} a^z &= \exp(\ln(a) \cdot z) = \exp(\ln(a) \cdot x) \cdot \exp(i \cdot \ln(a) \cdot y) \\ &= a^x \cdot \left( \cos(\ln(a) \cdot y) + i \cdot \sin(\ln(a) \cdot y) \right) \end{aligned}$$

wohldefiniert.

7. Die Frage, ob es Komplexe Logarithmen gibt, werden wir erst im dritten Semester studieren können. Dasselbe gilt für allgemeine Potenzfunktionen auf  $\mathbb{C}$ .

**Additionstheoreme** Mit Hilfe der Euler-Formel können sehr leicht die Additionstheoreme abgeleitet werden. Es gilt zum Beispiel

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

Die Additionstheoreme für  $\cos$  bzw.  $\sin$  können nun durch Vergleich der Real- bzw. Imaginärteile auf der linken und der rechten Seite abgelesen werden. Analog kann man zu jedem  $m \in \mathbb{N}$  Formeln für  $\sin(m \cdot \alpha)$  und  $\cos(m \cdot \alpha)$  ableiten (siehe Hausaufgabe).

**Hyperbolischer Sinus und Kosinus** Man bezeichnet

$$\sinh(z) := \frac{e^{+z} - e^{-z}}{2}, \quad \cosh(z) := \frac{e^{+z} + e^{-z}}{2}$$

als den Sinus hyperbolicus und den Kosinus hyperbolicus. Aus diesen Definition folgt (Nachrechnen!)

$$\sinh(-z) = -\sinh(z), \quad \cosh(-z) = \cosh(z), \quad (\cosh(z))^2 - (\sinh(z))^2 = 1.$$

sowie

$$\sinh(z) := -i \cdot \sin(i \cdot z), \quad \cosh(z) = \cos(i \cdot z).$$

## Komplexe Polynome

**Definition** Ein (komplexes) Polynom vom Grad  $n$  ist eine Funktion  $f: \mathbb{C} \rightarrow \mathbb{C}$  der Bauart

$$f(z) = \alpha_n \cdot z^n + \alpha_{n-1} \cdot z^{n-1} + \dots + \alpha_1 \cdot z^1 + \alpha_0 \cdot z^0$$

mit  $\alpha_n \neq 0$ , wobei alle Koeffizienten  $\alpha_j$  komplexe Zahlen sein dürfen.

**Theorem (Fundamentalsatz der Algebra)** Für jedes komplexe Polynom  $f$  vom Grad  $n$  existieren  $n$  komplexe Zahlen  $z_1, \dots, z_n \in \mathbb{C}$ , so dass

$$f(z) = \alpha_n \cdot (z - z_n) \cdot (z - z_{n-1}) \cdot \dots \cdot (z - z_2) \cdot (z - z_1).$$

Diese Zahlen  $z_1, \dots, z_n$  sind gerade die Nullstellen von  $f$ . Sie sind bis auf Umnummerierung eindeutig, aber müssen nicht unbedingt verschieden sein, d.h. es kann Doppelnullstellen, Tripelnnullstellen usw. geben. (Ein Beweis des Fundamentalsatzes können wir erst im dritten Semester geben).

### Beispiele

1.  $z^2 - 1 = (z - 1) \cdot (z + 1)$ ,  $z^2 - 2z + 1 = (z - 1)^2$ ,  $z^2 + 1 = (z - i) \cdot (z + i)$
2.  $z^3 - 2z^2 + 2z = (z - 0) \cdot (z - 1 - i) \cdot (z - 1 + i)$
3.  $z^3 - (3 - i) \cdot z^2 + (2 - 3 \cdot i) \cdot z + 2 \cdot i = (z - 2) \cdot (z - 1) \cdot (z + i)$
4.  $z^4 + z^3 - z^2 + z - 2 = (z - i) \cdot (z + i) \cdot (z - 1) \cdot (z + 2)$

**Bemerkung** Die Bestimmung der Nullstellen eines gegebenen Polynoms ist keine leichte Aufgabe und es gibt leider *keinen* Lösungsalgorithmus, der für alle Polynome zum Ziel führt.

1. Für  $n = 2$  kann man immer mit quadratischer Ergänzung (siehe Große Übung I und die Diskussion zu Quadratwurzeln unten) bzw. die aus solchen Betrachtungen abgeleiteten Varianten der  $p$ - $q$ -Formel benutzen.
2. Für kubische ( $n = 3$ ) und quartische ( $n = 4$ ) Polynome gibt es zwar geschlossene Lösungsformeln (siehe zum Beispiel *Cardansche Formeln* unter WIKIPEDIA), aber diese sind recht unhandlich und nur selten wirklich nützlich.
3. Für  $n > 5$  gibt es keine geschlossenen Formeln (das kann man sogar beweisen: Satz von Abel-Ruffini).
4. Nullstellen können heutzutage näherungsweise mit dem Computer berechnet werden (siehe Große Übung II).

**Bemerkung** Selbst wenn die Koeffizienten  $\alpha_1, \dots, \alpha_n$  alle reell sind, können einige (oder auch alle) der Nullstellen  $z_1, \dots, z_n$  komplex sein. Allerdings treten die nicht-reellen Nullstellen dann als Paare konjugiert komplexer Zahlen auf (siehe Hausaufgabe).

**Polynomdivision** Hat man für ein gegebenes Polynom  $f$  vom Grad  $n$  eine Nullstelle  $z_n$  gefunden (zum Beispiel durch Raten), so kann man ein Polynom  $g$  vom Grad  $n - 1$  berechnen, so dass  $f(z) = (z - z_n) \cdot g(z)$  gilt. Anschließend kann man dann die Nullstellen des einfacheren Polynoms  $g$  untersuchen.

Wir wollen das allgemeine Verfahren am Beispiel des kubischen Polynoms

$$f(z) = z^3 + 5z^2 + 9z + 5 = 0$$

mit der (geratenen) Nullstelle  $z_3 = -1$  erklären, indem wir den Faktor  $(z + 1)$  sukzessive wie folgt abspalten:

$$\begin{aligned} z^3 + 5z^2 + 9z + 5 &= z^2 \cdot (z + 1) + 4z^2 + 9z + 5 && \text{(höchste Potenz } z^3) \\ 4z^2 + 9z + 5 &= 4z \cdot (z + 1) + 5z + 5 && \text{(höchste Potenz } z^2) \\ 5z + 5 &= 5 \cdot (z + 1) + 0 && \text{(höchste Potenz } z^1) \end{aligned}$$

Nach  $n$  (diesmal drei) Schritten enden wir mit dem Restterm 0 und erhalten durch Einsammeln der Teilresultate die Darstellung

$$f(z) = (z + 1) \cdot g(z) \quad \text{mit} \quad g(z) = z^2 + 4z + 5.$$

Im konkreten Fall ist  $g$  ein quadratisches Polynom (also vom Grad 2) und wir können die Nullstellen von  $g$  durch die  $p$ - $q$ -Formel von oben berechnen. Insbesondere gilt

$$g(z) = (z + 2 - \mathbf{i}) \cdot (z + 2 + \mathbf{i}).$$

**Einheitswurzeln** Für jedes  $n \in \mathbb{N}$  können die  $n$  komplexen Lösungen  $z \in \mathbb{C}$  der Gleichung

$$z^n = 1$$

explizit berechnet werden. Mit  $z = r \cdot e^{i \cdot \varphi}$  erhalten wir nämlich

$$1 = r^n \cdot \exp(i \cdot n \cdot \varphi) = r^n \cdot (\cos(n \cdot \varphi) + i \cdot \sin(n \cdot \varphi))$$

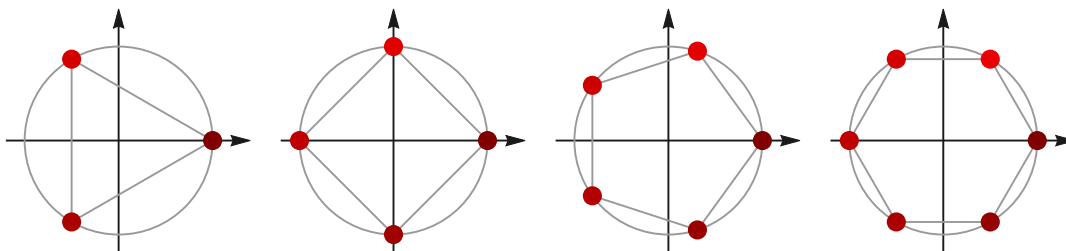
und schließend, dass  $r = 1$  gelten muss und dass  $n \cdot \varphi$  ein ganzzahliges Vielfaches von  $2\pi$  sein muss. Dies liefert die  $n$  verschiedenen Lösungen

$$z_1 = e^{i \cdot 2\pi \cdot 1/n}, \quad z_2 = e^{i \cdot 2\pi \cdot 2/n}, \quad \dots, \quad z_n = e^{i \cdot 2\pi \cdot n/n},$$

die eine direkte geometrische Bedeutung haben. Beachte, dass hier  $z_n = 1$  immer eine reelle Lösung ist und wir  $z_1 = -1$  und  $z_2 = +1$  im Spezialfall  $n = 2$  erhalten. Desweiteren gilt

$$z^n - 1 = (z - z_1) \cdot \dots \cdot (z - z_n)$$

nach dem Fundamentalsatz der Algebra.



Die  $n$ -ten Einheitswurzeln (rot,  $z_1$  jeweils am hellsten) liegen auf dem Einheitskreis und bilden die Ecken eines regelmäßigen  $n$ -Ecks (dargestellt für  $n \in \{3, 4, 5, 6\}$ ).

**\*Über das Wurzelziehen im Komplexen** Für gegebenes  $n \in \mathbb{N}$  und  $\zeta \in \mathbb{C}$  existieren  $n$  komplexe Lösungen der Gleichung

$$z^n = \zeta$$

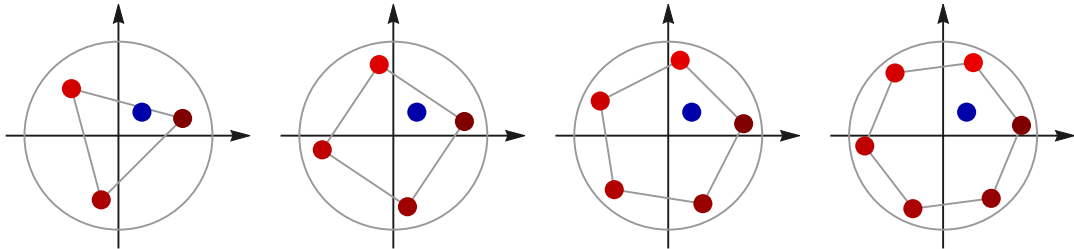
und die Rechenregeln der Multiplikation erzwingen (analog zur Berechnung der Einheitswurzeln), dass  $|z|^n = |\zeta|$  gelten muss und dass  $n \arg(z) - \arg(\zeta)$  ein ganzzahliges Vielfaches von  $2\pi$  sein muss. Insbesondere erhalten wir

$$z_k = |\zeta|^{1/n} \cdot \exp\left(\mathrm{i} \cdot \frac{\arg(\zeta) + 2\pi \cdot k}{n}\right) \quad \text{für} \quad k \in \{1, \dots, n\}.$$

Die Zahlen  $z_k$  werden die  $n$ -ten Wurzeln von  $\zeta$  genannt. Beachte, dass man auch

$$z_k = e^{\mathrm{i} \cdot 2\pi \cdot k/n} \cdot |\zeta|^{1/n} \cdot \exp\left(\mathrm{i} \cdot \frac{\arg(\zeta)}{n}\right)$$

schreiben kann, wobei nur der erste Faktor von  $k$  abhängt und gerade die  $k$ -te Einheitswurzel ist.



Die Lösungen (rot) der Gleichung  $z^n = \zeta$  für  $\zeta = \frac{1}{4}(1 + \mathrm{i})$  (blau) und  $n \in \{3, 4, 5, 6\}$ , wobei  $z_1$  jeweils am hellsten und  $z_n$  jeweils am dunkelsten erscheint. Der Kreis hat den Radius 1.

**\*Über quadratische Wurzeln** Im Fall von  $n = 2$  vereinfachen sich die obigen Formeln und die zwei komplexen Lösungen der Gleichung

$$z^2 = \zeta$$

können als

$$z_1 = -\sqrt{|\zeta|} \cdot \exp\left(\mathrm{i} \cdot \frac{\arg(\zeta)}{2}\right), \quad z_2 = +\sqrt{|\zeta|} \cdot \exp\left(\mathrm{i} \cdot \frac{\arg(\zeta)}{2}\right),$$

geschrieben werden (beachte, dass  $\exp(\mathrm{i} \cdot \pi) = -1$  und dass  $z_2 = -z_1$ ). Allerdings setzt dies die Kenntnis von  $\arg(\zeta)$ , also der Polarkoordinaten von  $\zeta$  voraus. Mit Hilfe von

$$\left(\operatorname{Re}(z) + \mathrm{i} \cdot \operatorname{Im}(z)\right)^2 = \operatorname{Re}(\zeta) + \mathrm{i} \cdot \operatorname{Im}(\zeta)$$

und elementarer Rechnungen (Übungsaufgabe) mit reellen Ausdrücken zeigt man, dass auch die Formeln

$$|\operatorname{Re}(z_1)| = |\operatorname{Re}(z_2)| = \sqrt{\frac{+\operatorname{Re}(\zeta) + \sqrt{\operatorname{Re}(\zeta)^2 + \operatorname{Im}(\zeta)^2}}{2}},$$

$$|\operatorname{Im}(z_1)| = |\operatorname{Im}(z_2)| = \sqrt{\frac{-\operatorname{Re}(\zeta) + \sqrt{\operatorname{Re}(\zeta)^2 + \operatorname{Im}(\zeta)^2}}{2}}$$

gelten, die man ohne die Polarkoordinaten von  $\zeta$  zu kennen auswerten kann. Die zu wählenden Vorzeichen hängen wie folgt von den Eigenschaften von  $\operatorname{Im}(\zeta)$  ab:

$$\operatorname{Im}(\zeta) > 0 \quad : \quad \operatorname{Re}(z_1) < 0, \operatorname{Im}(z_1) < 0, \quad \operatorname{Re}(z_2) > 0, \operatorname{Im}(z_2) > 0$$

sowie

$$\operatorname{Im}(\zeta) < 0 \quad : \quad \operatorname{Re}(z_1) > 0, \operatorname{Im}(z_1) < 0, \quad \operatorname{Re}(z_2) < 0, \operatorname{Im}(z_2) > 0.$$

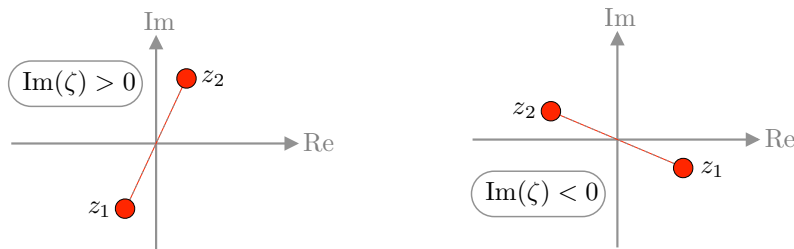
Für  $\operatorname{Im}(z) = 0$  ist  $\zeta$  reell und wir erhalten die vereinfachten Formeln

$$\zeta \in \mathbb{R} \text{ und } \zeta > 0 : \quad z_1 = -\sqrt{\zeta}, \quad z_2 = +\sqrt{\zeta}$$

bzw.

$$\zeta \in \mathbb{R} \text{ und } \zeta < 0 : \quad z_1 = -i \cdot \sqrt{-\zeta}, \quad z_2 = +i \cdot \sqrt{-\zeta}.$$

Wegen der Zweizahl der Lösungen und aufgrund der subtilen Vorzeichenbedingungen benutzen wir in dieser Vorlesung das Wurzelzeichen  $\sqrt{\quad}$  nur für nichtnegative reelle Zahlen, aber nicht für allgemeine komplexe Zahlen (obwohl einige Autoren dies tun).



Schematische Darstellung der zwei Lösungen der Gleichung  $z^2 = \zeta$  für  $\zeta \notin \mathbb{R}$ .



# Kapitel 3

## Vektorrechnung und Analytische Geometrie

Vorlesung 07, 04. November 2019

**Motivation** Vektoren sind ein zentraler Baustein der höheren Mathematik. Insbesondere können mit ihrer Hilfe geometrische Objekte beschrieben werden.

### 3.1 Vektoren im $\mathbb{R}^n$

**Vorbemerkung** Wir arbeiten im Folgenden im  $n$ -dimensionalen „Raum“, wobei  $n$  beliebig aber fest ist. Bei der ersten Lektüre können Sie sich vorstellen, dass  $n = 2$  (2D, zweidimensionaler Raum = Ebene) oder  $n = 3$  (3D, der uns umgebende physikalische Raum) gilt. Der Großteil der Theorie kann aber problemlos für beliebiges  $n \in \mathbb{N}$  formuliert werden und wir werden später sehen, dass  $n$ -dimensionale Räume ein sehr sinnvolles Konzept sind.

**Beobachtung** Punkte in der zweidimensionalen Ebene bzw. im dreidimensionalen Raum bzw. im  $n$ -dimensionalen Raum können — sofern ein Achsenkreuz festgelegt ist — als

$$P = \begin{pmatrix} p_1 \\ p_2 \end{pmatrix} \quad \text{bzw.} \quad P = \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix} \quad \text{bzw.} \quad P = \begin{pmatrix} p_1 \\ \vdots \\ p_n \end{pmatrix}$$

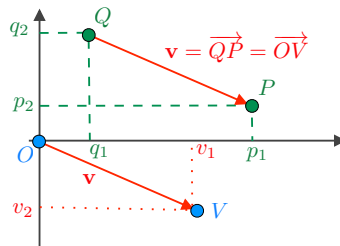
beschrieben werden, wobei die  $p_j$  die (kartesischen) Koordinaten sind und durch reelle Zahlen gegeben sind. Aus mathematischer Sicht handelt es sich also um Elemente des  $\mathbb{R}^2$  bzw.  $\mathbb{R}^3$  bzw.  $\mathbb{R}^n$ .

### Einführung

**Definition** Ein  $n$ -dimensionaler Vektor ist der gerichtete Verbindungspfeil zwischen zwei Punkten  $Q$  und  $P$ . Wir schreiben

$$\mathbf{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \overrightarrow{QP} = \begin{pmatrix} P_1 - Q_1 \\ \vdots \\ P_n - Q_n \end{pmatrix},$$

wobei  $v_j$  die  $j$ -te Komponente von  $\mathbf{v}$  genannt wird. Insbesondere wird auch jeder Vektor eindeutig durch ein Element aus  $\mathbb{R}^n$  beschrieben.



Über Punkte (grün und blau) und Vektoren (rot). Hier dargestellt in 2D, d.h. für  $n = 2$ .

### Klarstellung

1. Zu jedem Vektor  $\mathbf{v}$  gibt es genau einen Punkt  $V$ , so dass  $\mathbf{v} = \overrightarrow{OV}$ , wobei  $O$  der Koordinatenursprung ist, dessen Koordinaten sämtlich verschwinden. Insbesondere sind die Koordinaten von  $V$  gerade die Komponenten von  $\mathbf{v}$  und man nennt  $\mathbf{v}$  auch den Ortsvektor von  $V$ .
2. Obwohl sowohl Punkte als auch Vektoren jeweils durch ein  $n$ -Tupel reeller Zahlen beschrieben werden, gibt es einen wesentlichen konzeptionellen Unterschied: Vektoren können **frei verschoben** werden! (Mathematiker formulieren diese Tatsache mit anderen Worten: Vektoren bzw. Punkte gehören zum *linearen* bzw. zum *affinen* Raum).

### Bemerkung

1. Wir schreiben hier Vektoren immer als Spaltenvektoren und werden erst später das verwandte Konzept von Zeilenvektoren einführen.
2. Wir benutzen im Skript die Schreibweise  $\mathbf{v}$  für Vektoren. Manchmal wird in der Literatur statt des Fettdruckes die Pfeilschreibweise  $\vec{v}$  verwendet, aber diese ist sehr unhandlich. Mathematiker benutzen übrigens in aller Regel weder die Fett- noch die Pfeilschreibweise, sondern notieren einfach  $v \in \mathbb{R}^n$ .
3. An der Tafel gilt: Vektoren werden (in der Regel) mit  $\mathbf{u}, \mathbf{v}, \mathbf{w}$  bezeichnet und ihre Komponenten mit  $u_j, v_j, w_j$ . Andere reelle Zahlen (die wir auch Skalare nennen) werden mit kleinen griechischen Buchstaben ( $\lambda, \mu$  usw.) abgekürzt.

**Vektoraddition und Multiplikation mit Skalaren** Vektoren können komponentenweise addiert und mit einem Skalar  $\lambda \in \mathbb{R}$  multipliziert werden:

$$\mathbf{v} + \mathbf{w} = \begin{pmatrix} v_1 + w_1 \\ \vdots \\ v_n + w_n \end{pmatrix}, \quad \lambda \cdot \mathbf{v} = \begin{pmatrix} \lambda \cdot v_1 \\ \vdots \\ \lambda \cdot v_n \end{pmatrix}$$

**Rechenregeln** Mit den Notationen

$$\mathbf{0} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \quad (\text{Nullvektor}), \quad -\mathbf{v} = \begin{pmatrix} -v_1 \\ \vdots \\ -v_n \end{pmatrix} = (-1) \cdot \mathbf{v}$$

gilt offensichtlich

$$\begin{aligned} \text{(VR.I.1)} \quad \mathbf{u} + (\mathbf{v} + \mathbf{w}) &= (\mathbf{u} + \mathbf{v}) + \mathbf{w} \\ \text{(VR.I.2)} \quad \mathbf{v} + \mathbf{w} &= \mathbf{w} + \mathbf{v} \\ \text{(VR.I.3)} \quad \mathbf{v} + \mathbf{0} &= \mathbf{v} \\ \text{(VR.I.4)} \quad \mathbf{v} + (-\mathbf{v}) &= \mathbf{0} \end{aligned}$$

sowie

$$\begin{aligned} \text{(VR.II.1)} \quad 1 \cdot \mathbf{v} &= \mathbf{v} \\ \text{(VR.II.2)} \quad \lambda \cdot (\mu \cdot \mathbf{v}) &= (\lambda \cdot \mu) \cdot \mathbf{v} \\ \text{(VR.II.3)} \quad (\lambda + \mu) \cdot \mathbf{v} &= \lambda \cdot \mathbf{v} + \mu \cdot \mathbf{v} \\ \text{(VR.II.4)} \quad \lambda \cdot (\mathbf{v} + \mathbf{w}) &= \lambda \cdot \mathbf{v} + \lambda \cdot \mathbf{w} \end{aligned}$$

**Bemerkung** Vektoren können addiert und via

$$\mathbf{v} - \mathbf{w} = \mathbf{v} + (-1) \cdot \mathbf{w} = \begin{pmatrix} v_1 - w_1 \\ \vdots \\ v_n - w_n \end{pmatrix}$$

auch voneinander subtrahiert werden. Die Frage, ob bzw. in welchem Sinn Vektoren multipliziert werden können, diskutieren wir gleich, wollen aber hier schon festhalten: Vektoren dürfen **niemals** dividiert werden, d.h. Ausdrücke wie  $\mathbf{v}/\mathbf{w}$  haben keinen Sinn.

**Einheitsvektoren** In zwei bzw. drei Dimensionen nennt man

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

bzw.

$$\mathbf{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

die Einheitsvektoren. Analoge Formeln gelten in  $nD$ , wobei es dann  $n$  Einheitsvektoren gibt. Will man die Komponenten der Einheitsvektoren bezeichnen, so benutzt man Doppelindizes und schreibt in 3D zum Beispiel

$$\mathbf{e}_j = \begin{pmatrix} e_{j,1} \\ e_{j,2} \\ e_{j,3} \end{pmatrix}.$$

## Euklidische Norm und Skalarprodukt

**Definition** Die nichtnegative reelle Zahl

$$\|\mathbf{v}\| = \sqrt{v_1^2 + v_2^2 + \dots + v_n^2}$$

wird als euklidische Norm (oder Länge oder Betrag) des Vektors  $\mathbf{v}$  bezeichnet. Der normalisierte Vektor

$$\frac{1}{\|\mathbf{v}\|} \cdot \mathbf{v} \quad \left( \text{bzw. } \frac{\mathbf{v}}{\|\mathbf{v}\|} \right)$$

wird die Richtung von  $\mathbf{v} \neq \mathbf{0}$  genannt. Insbesondere ist die Richtung von  $\mathbf{v}$  ein Vektor der Länge 1, wobei der Nullvektor keine Richtung besitzt.

**Bemerkung** Jeder Vektor  $\mathbf{v} \neq \mathbf{0}$  ist eindeutig durch seine Länge und seine Richtung bestimmt.

**Notation** Zwei Vektoren  $\mathbf{v} \neq \mathbf{0}$  und  $\mathbf{w} \neq \mathbf{0}$  mit  $\mathbf{v}/\|\mathbf{v}\| = \pm\mathbf{w}/\|\mathbf{w}\|$  werden parallel genannt, wobei man dann oftmals  $\mathbf{v} \parallel \mathbf{w}$  schreibt.

### Eigenschaften der Norm

(VR.III.1)	$\ \mathbf{v}\  \geq 0$	(Nicht-Negativität)
(VR.III.2)	$\ \mathbf{v}\  = 0 \Leftrightarrow \mathbf{v} = \mathbf{0}$	(Nicht-Entartung)
(VR.III.3)	$\ \lambda \cdot \mathbf{v}\  =  \lambda  \cdot \ \mathbf{v}\ $	(Homogenität)
(VR.III.4)	$\ \mathbf{v} + \mathbf{w}\  \leq \ \mathbf{v}\  + \ \mathbf{w}\ $	(Dreiecksungleichung)

Die ersten drei Eigenschaften folgen dabei direkt aus der Definition der Norm; die vierte ist eine Konsequenz der Cauchy-Schwarz-Ungleichung, die wir weiter unten etablieren werden.

**Definition** Der Ausdruck

$$\langle \mathbf{v}, \mathbf{w} \rangle := v_1 \cdot w_1 + v_2 \cdot w_2 + \dots + v_n \cdot w_n$$

wird Skalarprodukt von  $\mathbf{v}$  und  $\mathbf{w}$  genannt. Insbesondere gilt  $\langle \text{Vektor}, \text{Vektor} \rangle = \text{Skalar}$ .

**Bemerkung** Das Skalarprodukt wird manchmal auch als  $\mathbf{v} \cdot \mathbf{w}$  geschrieben. Beachte: In der  $n$ -dimensionalen Vektorrechnung gibt es im Allgemeinen kein Produkt der Art „Vektor mal Vektor gleich Vektor“. Insbesondere meint „ $\cdot$ “ zwischen zwei Vektoren das Skalarprodukt und zwischen Vektor und Skalar die skalare Multiplikation von oben. Für  $n = 3$  werden wir allerdings unten das sogenannte Kreuzprodukt einführen, dass aber anders geschrieben wird.

**Eigenschaften des Skalarproduktes** Aus den Definitionen folgt unmittelbar:

(VR.IV.1)	$\langle \mathbf{v}, \mathbf{v} \rangle = \ \mathbf{v}\ ^2$	(SKP und Norm)
(VR.IV.2)	$\langle \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{w}, \mathbf{v} \rangle$	(Symmetrie)
(VR.IV.3)	$\langle \lambda \cdot \mathbf{v}, \mathbf{w} \rangle = \lambda \cdot \langle \mathbf{v}, \mathbf{w} \rangle$	(Bilinearität, Teil 1)
(VR.IV.4)	$\langle \mathbf{u} + \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{u}, \mathbf{w} \rangle + \langle \mathbf{v}, \mathbf{w} \rangle$	(Bilinearität, Teil 2)

**Bemerkung**

1. Das Skalarprodukt ist linear im ersten Argument und wegen der Symmetrie auch linear im zweiten Argument. d.h. es gilt

$$\begin{aligned}\langle \lambda \cdot \mathbf{u} + \mu \cdot \mathbf{v}, \mathbf{w} \rangle &= \lambda \cdot \langle \mathbf{u}, \mathbf{w} \rangle + \mu \cdot \langle \mathbf{v}, \mathbf{w} \rangle \\ &= \lambda \cdot \langle \mathbf{w}, \mathbf{u} \rangle + \mu \cdot \langle \mathbf{w}, \mathbf{v} \rangle \\ &= \langle \mathbf{w}, \lambda \cdot \mathbf{u} + \mu \cdot \mathbf{v} \rangle.\end{aligned}$$

Man sagt daher auch, das Skalarprodukt sei bilinear.

2. Es gilt

$$\langle \mathbf{e}_i, \mathbf{e}_j \rangle = \delta_{ij} := \begin{cases} 1 & \text{falls } i = j, \\ 0 & \text{falls } i \neq j, \end{cases}$$

wobei  $\delta_{ij}$  das sogenannte Kronecker-Delta ist.

3. Für jeden Vektor gilt die Darstellungsformel

$$\mathbf{v} = \sum_{j=1}^n v_j \mathbf{e}_j = \sum_{j=1}^n \langle \mathbf{v}, \mathbf{e}_j \rangle \cdot \mathbf{e}_j.$$

4. Unsere Definitionen implizieren die Formel

$$\|\mathbf{v}\|^2 = \sum_{j=1}^n v_j^2,$$

die als  $n$ -dimensionales Analogon zum Satz von Pythagoras betrachtet werden kann.

**Skalarprodukt und Winkel** Mit elementargeometrischen Argumenten in 2D (siehe Bild) zeigt man

$$\begin{aligned}v_1 \cdot w_1 + v_2 \cdot w_2 &= \sqrt{v_1^2 + v_2^2} \cdot \sqrt{w_1^2 + w_2^2} \cdot (\cos(\alpha) \cdot \cos(\beta) + \sin(\alpha) \cdot \sin(\beta)) \\ &= \sqrt{v_1^2 + v_2^2} \cdot \sqrt{w_1^2 + w_2^2} \cdot \cos(\beta - \alpha)\end{aligned}$$

bzw.

$$\langle \mathbf{v}, \mathbf{w} \rangle = \|\mathbf{v}\| \cdot \|\mathbf{w}\| \cdot \cos(\sphericalangle(\mathbf{v}, \mathbf{w})),$$

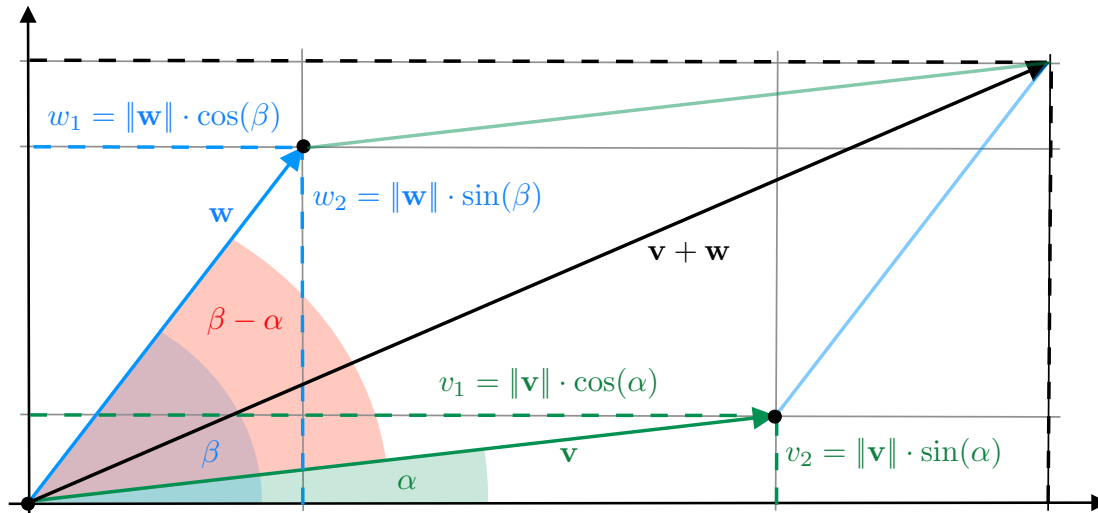
wobei Winkel in 2D üblicherweise in mathematisch positiver Orientierung (d.h. gegen den Uhrzeigersinn) gemessen und meist im Bogenmaß angegeben werden. Insbesondere gilt in 2D

$$\sphericalangle(\mathbf{w}, \mathbf{v}) = 2\pi - \sphericalangle(\mathbf{v}, \mathbf{w}),$$

wobei jeder der beiden Winkel im Intervall  $[0, 2\pi]$  liegt. Da drei nicht-entartete Punkte im  $\mathbb{R}^3$  — und analog im  $\mathbb{R}^n$  — immer eine Ebene aufspannen, kann man die obige Formel zwischen dem Skalarprodukt und dem Winkel auch in  $n$ D verwenden. Beachte aber, dass bei Winkelangaben in mehr als zwei Dimensionen meist keine Orientierung ausgezeichnet wird, d.h. es gilt

$$\sphericalangle(\mathbf{w}, \mathbf{v}) = \sphericalangle(\mathbf{v}, \mathbf{w})$$

und der Winkel wird als Wert im Intervall  $(0, \pi]$  angegeben.



Zum Winkel und zum Flächeninhalt in 2D.

**Weiteres zur Geometrie** Der Flächeninhalt des bei der Vektoraddition im Bild entstehenden Parallelogramms kann durch

$$(v_1 + w_1) \cdot (v_2 + w_2) - 2 \cdot v_2 \cdot w_1 - 2 \cdot \left(\frac{1}{2} \cdot v_1 \cdot v_2\right) - 2 \cdot \left(\frac{1}{2} \cdot w_1 \cdot w_2\right) = v_1 \cdot w_2 - v_2 \cdot w_1$$

berechnet werden (Parallelogramm = großes Rechteck – zwei kleine Rechtecke – zweimal zwei Dreiecke). Die Formel gilt immer, sofern ggf. noch das Vorzeichen korrigiert wird (Flächeninhalte sind bei uns immer positiv). Außerdem kann man analog zu oben im Bild die 2D-Formel

$$v_1 \cdot w_2 - v_2 \cdot w_1 = \sqrt{v_1^2 + v_2^2} \cdot \sqrt{w_1^2 + w_2^2} \cdot \sin(\beta - \alpha)$$

ablesen, die unten beim Kreuzprodukt verallgemeinert wird.

**\*Bemerkung** Für  $n = 2$  gilt folgende komplexe Variante

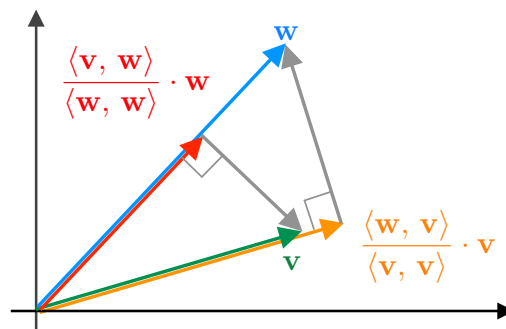
$$(v_1 - i \cdot v_2) \cdot (w_1 + i \cdot w_2) = (v_1 \cdot w_1 + v_2 \cdot w_2) + i \cdot (v_1 \cdot w_2 - v_2 \cdot w_1),$$

die aber im Rahmen dieser Vorlesung keine Rolle spielen wird.

**Bemerkung** Für zwei gegebene Vektoren  $\mathbf{v}$  und  $\mathbf{w} \neq \mathbf{0}$  heißt der Vektor

$$\frac{\langle \mathbf{v}, \mathbf{w} \rangle}{\langle \mathbf{w}, \mathbf{w} \rangle} \cdot \mathbf{w} = \left\langle \mathbf{v}, \frac{\mathbf{w}}{\|\mathbf{w}\|} \right\rangle \cdot \frac{\mathbf{w}}{\|\mathbf{w}\|}$$

die Projektion von  $\mathbf{v}$  auf  $\mathbf{w}$ . Sie ist das einzige skalare Vielfache  $\lambda \cdot \mathbf{w}$  von  $\mathbf{w}$ , so dass  $\lambda \neq 0$  und  $\langle \lambda \cdot \mathbf{w}, \mathbf{v} - \lambda \cdot \mathbf{w} \rangle = 0$ .



Zur Projektion von Vektoren. Die grauen Boxen stehen für rechte Winkel.

**Notation** Wir sagen, zwei Vektoren  $\mathbf{v} \neq \mathbf{0}$  und  $\mathbf{w} \neq \mathbf{0}$  mit  $\langle \mathbf{v}, \mathbf{w} \rangle = 0$  stehen senkrecht aufeinander und schreiben  $\mathbf{v} \perp \mathbf{w}$ . (Beachte, dass zwar  $\langle \mathbf{v}, \mathbf{0} \rangle = 0$  für alle  $\mathbf{v}$  gilt, wir aber in der Regel nicht  $\mathbf{0} \perp \mathbf{v}$  schreiben).

**Theorem** Die Cauchy-Schwarzsche-Ungleichung

$$|\langle \mathbf{v}, \mathbf{w} \rangle| \leq \|\mathbf{v}\| \cdot \|\mathbf{w}\|$$

gilt für alle  $v, w \in \mathbb{R}^n$ .

**Beweis** Wir können  $\mathbf{v} \neq \mathbf{0}$  und  $\mathbf{w} \neq \mathbf{0}$  annehmen, denn andernfalls ist Behauptung trivial (da dann sowohl die linke als auch die rechte Seite verschwinden). Nach den obigen Regeln gilt

$$0 \leq \langle \lambda \cdot \mathbf{v} \pm \mu \cdot \mathbf{w}, \lambda \cdot \mathbf{v} \pm \mu \cdot \mathbf{w} \rangle = \lambda^2 \cdot \langle \mathbf{v}, \mathbf{v} \rangle \pm 2 \cdot \lambda \cdot \mu \cdot \langle \mathbf{v}, \mathbf{w} \rangle + \mu^2 \cdot \langle \mathbf{w}, \mathbf{w} \rangle$$

für alle Vektoren  $\mathbf{v}, \mathbf{w} \in \mathbb{R}^n$  sowie alle Skalare  $\lambda, \mu \in \mathbb{R}$  (wegen  $\pm$  handelt es sich eigentlich um zwei Formeln). Mit der Wahl

$$\lambda := \|\mathbf{w}\| = \sqrt{\langle \mathbf{w}, \mathbf{w} \rangle} > 0, \quad \mu := \|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle} > 0$$

erhalten wir nach Umstellen

$$\mp 2 \cdot \lambda \cdot \mu \cdot \langle \mathbf{v}, \mathbf{w} \rangle \leq 2 \cdot \lambda^2 \cdot \mu^2$$

und nach Division durch  $2 \cdot \lambda \cdot \mu > 0$  schließlich

$$\mp \langle \mathbf{v}, \mathbf{w} \rangle \leq \lambda \cdot \mu = \|\mathbf{v}\| \cdot \|\mathbf{w}\| .$$

Es gilt die Behauptung wegen  $|\zeta| \leq \max\{-\zeta, +\zeta\}$  für alle  $\zeta \in \mathbb{R}$ . □

**Folgerung** Zum Beweis der Dreiecksungleichung bemerken wir

$$\begin{aligned} \|\mathbf{v} + \mathbf{w}\|^2 &= \langle \mathbf{v} + \mathbf{w}, \mathbf{v} + \mathbf{w} \rangle && \text{(Norm und SKP)} \\ &= \langle \mathbf{v}, \mathbf{v} \rangle + 2 \cdot \langle \mathbf{v}, \mathbf{w} \rangle + \langle \mathbf{w}, \mathbf{w} \rangle && \text{(Bilinearität des SKP)} \\ &\leq \langle \mathbf{v}, \mathbf{v} \rangle + 2 \cdot \|\mathbf{v}\| \cdot \|\mathbf{w}\| + \langle \mathbf{w}, \mathbf{w} \rangle && \text{(Cauchy-Schwarz)} \\ &= \|\mathbf{v}\|^2 + 2 \cdot \|\mathbf{v}\| \cdot \|\mathbf{w}\| + \|\mathbf{w}\|^2 && \text{(Norm und SKP)} \\ &= (\|\mathbf{v}\| + \|\mathbf{w}\|)^2 && \text{(Binomische Formel)} \end{aligned}$$

und die gesuchte Ungleichung ergibt sich durch Wurzelziehen auf beiden Seiten. □

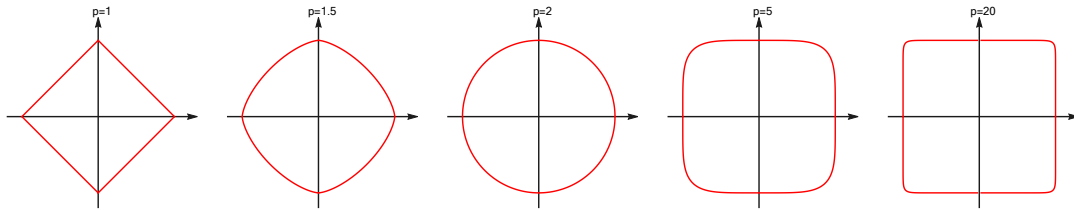
**\*Andere Normen** Neben der euklidischen Norm — bzw. dem euklidischen Abstands begriff — betrachtet man manchmal für jeden reellen Parameter  $p$  mit  $1 \leq p < \infty$  die sogenannte  $p$ -Norm, die für jedes  $\mathbf{v} \in \mathbb{R}^n$  durch

$$\|\mathbf{v}\|_p := \left( \sum_{i=1}^n |v_i|^p \right)^{1/p}$$

gegeben ist, wobei der Fall  $p = 2$  wieder die euklidische Norm liefert und

$$\|\mathbf{v}\|_\infty = \max \left\{ |v_i| : i \in \{1, \dots, n\} \right\}$$

im Grenzfall  $p \rightarrow \infty$  gilt. Die Formel (VR.IV.1) gilt aber nur für  $p = 2$ , d.h. nur die euklidische Norm ist mit dem Skalarprodukt verträglich.



Der „Einheitskreis“  $\{\mathbf{v} : \|\mathbf{v}\|_p = 1\}$  bzgl. der  $p$ -Norm in zwei Dimensionen und für verschiedene Wahlen von  $p$ .

Vorlesung 08, 06. November 2019

## Kreuzprodukt von Vektoren im $\mathbb{R}^3$

**Achtung** In diesem Abschnitt betrachten wir dreidimensionale Vektoren, d.h. es gilt  $n = 3$ . Den Grund für diese Einschränkung werden wir am Ende beleuchten.

**Definition** Für zwei 3D Vektoren  $\mathbf{v}$ ,  $\mathbf{w}$  definieren wir das Kreuzprodukt (manchmal auch Vektorprodukt genannt) durch

$$\mathbf{v} \times \mathbf{w} := \begin{pmatrix} v_2 \cdot w_3 - v_3 \cdot w_2 \\ v_3 \cdot w_1 - v_1 \cdot w_3 \\ v_1 \cdot w_2 - v_2 \cdot w_1 \end{pmatrix}.$$

Insbesondere gilt Vektor  $\times$  Vektor = Vektor.

### Beispiel

$$\mathbf{v} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad \mathbf{w} = \begin{pmatrix} 2 \\ 3 \\ -4 \end{pmatrix}, \quad \mathbf{v} \times \mathbf{w} = \begin{pmatrix} (-1) \cdot (-4) - 0 \cdot 3 \\ 0 \cdot 2 - 1 \cdot (-4) \\ 1 \cdot 3 - (-1) \cdot 2 \end{pmatrix} = \begin{pmatrix} 4 \\ 4 \\ 5 \end{pmatrix}$$

**Beispiel** Für die Einheitsvektoren ergibt sich durch Nachrechnen

$$\mathbf{e}_1 \times \mathbf{e}_2 = \mathbf{e}_3, \quad \mathbf{e}_2 \times \mathbf{e}_3 = \mathbf{e}_1, \quad \mathbf{e}_3 \times \mathbf{e}_1 = \mathbf{e}_2$$

sowie

$$\mathbf{e}_2 \times \mathbf{e}_1 = -\mathbf{e}_3, \quad \mathbf{e}_3 \times \mathbf{e}_2 = -\mathbf{e}_1, \quad \mathbf{e}_1 \times \mathbf{e}_3 = -\mathbf{e}_2.$$

### Bemerkung

1. Alle Formeln in diesem Abschnitt besitzen eine innere Symmetrie, die durch die zyklische Permutation

$$1 \rightsquigarrow 2, \quad 2 \rightsquigarrow 3, \quad 3 \rightsquigarrow 1$$

beschrieben werden kann, aber nicht mit der antizyklischen Permutation ( $1 \rightsquigarrow 3$ ,  $2 \rightsquigarrow 1$ ,  $3 \rightsquigarrow 2$ ) verträglich ist.

2. Es gilt stets  $\mathbf{0} \times \mathbf{v} = \mathbf{v} \times \mathbf{0} = \mathbf{0}$ .
3. Es gilt  $\mathbf{v} \times \mathbf{v} = \mathbf{0}$  für jeden Vektor  $\mathbf{v} \in \mathbb{R}^3$  und allgemeiner  $\mathbf{v} \times \mathbf{w} = \mathbf{0}$ , sofern  $\mathbf{v} \parallel \mathbf{w}$  (denn in diesem Fall existiert immer ein Skalar  $\lambda \in \mathbb{R}$ , so dass  $\mathbf{v} = \lambda \cdot \mathbf{w}$ ).
4. Sind  $\mathbf{v} \neq \mathbf{0}$  und  $\mathbf{w} \neq \mathbf{0}$  nicht parallel, so gilt  $\mathbf{v} \times \mathbf{w} \neq \mathbf{0}$  (siehe unten).



**Eigenschaften des Kreuzproduktes** Aus der Definition ergibt sich sofort:

$$\begin{aligned} \text{(VR.V.1)} \quad \mathbf{v} \times \mathbf{w} &= -\mathbf{w} \times \mathbf{v} && \text{(Antisymmetrie)} \\ \text{(VR.V.2)} \quad (\lambda \cdot \mathbf{v}) \times \mathbf{w} &= \lambda \cdot (\mathbf{v} \times \mathbf{w}) && \text{(Bilinearität, Teil 1)} \\ \text{(VR.V.3)} \quad (\mathbf{u} + \mathbf{v}) \times \mathbf{w} &= \mathbf{u} \times \mathbf{w} + \mathbf{v} \times \mathbf{w} && \text{(Bilinearität, Teil 2)} \end{aligned}$$

**Bemerkung** Die obigen Regeln implizieren außerdem

$$(\lambda \cdot \mathbf{u} + \mu \cdot \mathbf{v}) \times \mathbf{w} = \lambda \cdot \mathbf{u} \times \mathbf{w} + \mu \cdot \mathbf{v} \times \mathbf{w}$$

sowie

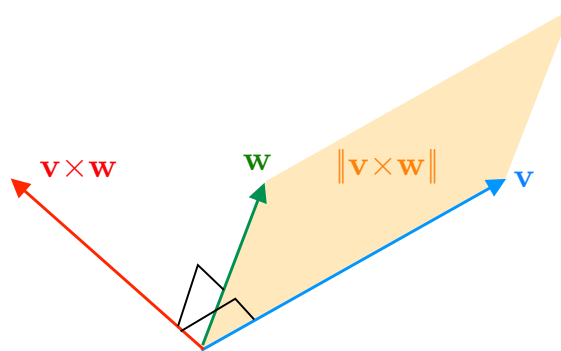
$$\mathbf{w} \times (\lambda \cdot \mathbf{u} + \mu \cdot \mathbf{v}) = \lambda \cdot \mathbf{w} \times \mathbf{u} + \mu \cdot \mathbf{w} \times \mathbf{v},$$

d.h. das Kreuzprodukt ist (trotz der Antisymmetrie) eine bilineare Operation.

**geometrische Interpretation** Es gilt (siehe auch die Bemerkungen zu Winkeln und zum Skalarprodukt)

$$\|\mathbf{v} \times \mathbf{w}\| = \|\mathbf{v}\| \cdot \|\mathbf{w}\| \cdot \sin(\angle(\mathbf{v}, \mathbf{w})),$$

und die Norm von  $\mathbf{v} \times \mathbf{w}$  ist gerade der Flächeninhalt des von  $\mathbf{v}$  und  $\mathbf{w}$  aufgespannten Parallelogramms.



Für nicht-parallele Vektoren  $\mathbf{v} \neq \mathbf{0}$  und  $\mathbf{w} \neq \mathbf{0}$  steht  $\mathbf{v} \times \mathbf{w}$  senkrecht auf  $\mathbf{v}$  und  $\mathbf{w}$ . Für  $\mathbf{v} \parallel \mathbf{w}$  gilt  $\mathbf{v} \times \mathbf{w} = \mathbf{0}$ .

Außerdem zeigt

$$\begin{aligned} \langle \mathbf{v}, \mathbf{v} \times \mathbf{w} \rangle &= v_1 \cdot (v_2 \cdot w_3 - v_3 \cdot w_2) + v_2 \cdot (v_3 \cdot w_1 - v_1 \cdot w_3) \\ &\quad + v_3 \cdot (v_1 \cdot w_2 - v_2 \cdot w_1) \\ &= 0 \end{aligned}$$

sowie eine analoge Rechnung für  $\langle \mathbf{w}, \mathbf{v} \times \mathbf{w} \rangle$  dass  $\mathbf{v} \times \mathbf{w}$  immer senkrecht auf sowohl  $\mathbf{v}$  und  $\mathbf{w}$  steht. Dabei gilt die Rechte-Hand-Regel:

$$\mathbf{v} \cong \text{Daumen}, \quad \mathbf{w} \cong \text{Zeigefinger}, \quad \mathbf{v} \times \mathbf{w} \cong \text{Mittelfinger}.$$

**Beispiel** Für gegebene  $\mathbf{p}, \mathbf{v} \in \mathbb{R}^n$  besteht die Menge

$$G = \left\{ \mathbf{x} \in \mathbb{R}^n : (\mathbf{x} - \mathbf{p}) \times \mathbf{v} = \mathbf{0} \right\}$$

aus den Ortsvektoren aller Punkte  $X$ , die auf der Geraden durch den Punkt  $P$  mit Richtungsvektor  $\mathbf{v}$ . Eine andere — aber geometrisch äquivalente — Schreibweise ist

$$G = \left\{ \mathbf{x} \in \mathbb{R}^n : \mathbf{x} = \mathbf{p} + \lambda \cdot \mathbf{v} \text{ für ein } \lambda \in \mathbb{R} \right\}.$$

**Beispiel** Der Flächeninhalt des durch die drei Punkte

$$U = \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix}, \quad V = \begin{pmatrix} 0 \\ -1 \\ 4 \end{pmatrix}, \quad W = \begin{pmatrix} 3 \\ 1 \\ 7 \end{pmatrix},$$

festgelegten Dreiecks in 3D ergibt sich mit

$$\mathbf{v} = \overrightarrow{UV} = \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}, \quad \mathbf{w} = \overrightarrow{UW} = \begin{pmatrix} 2 \\ 1 \\ 4 \end{pmatrix}$$

und

$$\mathbf{v} \times \mathbf{w} = \begin{pmatrix} (-1) \cdot 4 - 1 \cdot 1 \\ 1 \cdot 2 - (-1) \cdot 4 \\ (-1) \cdot 1 - (-1) \cdot 2 \end{pmatrix} = \begin{pmatrix} -5 \\ 6 \\ 1 \end{pmatrix}$$

aus

$$\frac{1}{2} \|\mathbf{v} \times \mathbf{w}\| = \frac{1}{2} \sqrt{5^2 + 6^2 + 1^2} = \frac{1}{2} \sqrt{62} = \sqrt{\frac{31}{2}} \approx 3,937.$$

### Anwendungen

1. Auf ein Teilchen mit Ladung  $q$  und Geschwindigkeit  $\mathbf{v}$  wirkt die elektromagnetische Lorentzkraft

$$\mathbf{F} = q \cdot (\mathbf{E} + \mathbf{v} \times \mathbf{B}),$$

wobei  $\mathbf{E}$  die elektrische Feldstärke ist und  $\mathbf{B}$  die magnetische Flussdichte bezeichnet.

2. Das Drehmoment der klassischen Mechanik ist durch

$$\mathbf{M} = \mathbf{r} \times \mathbf{F}$$

gegeben, wobei  $\mathbf{r}$  der Ortsvektor einer Punktmasse ist und  $\mathbf{F}$  die wirkende Kraft repräsentiert.

**Rechentrick** In der Literatur findet sich häufig die symbolische Formel

$$\begin{aligned} \mathbf{v} \times \mathbf{w} &= \det \begin{pmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{pmatrix} \\ &= (v_2 \cdot w_3 - v_3 \cdot w_2) \cdot \mathbf{e}_1 + (v_3 \cdot w_1 - v_1 \cdot w_3) \cdot \mathbf{e}_2 \\ &\quad + (v_1 \cdot w_2 - v_2 \cdot w_1) \cdot \mathbf{e}_3, \end{aligned}$$

die das Kreuzprodukt mit der Sarrus-Regel als Determinante einer formalen  $3 \times 3$ -Matrix darstellt. Matrizen und Determinanten werden wir später einführen.

**\*Spatprodukt** Der Ausdruck

$$[\mathbf{u}, \mathbf{v}, \mathbf{w}] = \langle \mathbf{u} \times \mathbf{v}, \mathbf{w} \rangle$$

beschreibt das Volumen des Parallelepipeds (auch *Spat* genannt), das durch die drei Vektoren aufgespannt wird (siehe zum Beispiel die Abbildungen in [AORS, Seite 60] oder unter WIKIPEDIA). Man kann durch Nachrechnen zeigen, dass

$$[\mathbf{u}, \mathbf{v}, \mathbf{w}] = \det \begin{pmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{pmatrix} = [\mathbf{v}, \mathbf{w}, \mathbf{u}] = [\mathbf{w}, \mathbf{u}, \mathbf{v}]$$

sowie  $[\mathbf{u}, \mathbf{v}, \mathbf{w}] = -[\mathbf{v}, \mathbf{u}, \mathbf{w}]$ .

**\*Verallgemeinerungen des Kreuzproduktes** Das Kreuzprodukt existiert in der eingeführten Form nur in 3D. Es gibt Verallgemeinerungen in 2D bzw. höheren Dimensionen, die man aus allen unabhängigen Varianten des Terms  $v_i \cdot w_j - v_j \cdot w_i$  zusammensetzt. Die Anzahl solcher Terme hängt aber von der Dimension  $n$  ab:

$$\begin{aligned} 2\text{D} : & \quad (i, j) \in \{(1, 2)\} && (1 \text{ Möglichkeit}) \\ 3\text{D} : & \quad (i, j) \in \{(1, 2), (2, 3), (3, 1)\} && (3 \text{ Möglichkeiten}) \\ 4\text{D} : & \quad (i, j) \in \{(1, 2), (2, 3), (3, 4), (4, 1), (2, 3), (3, 4)\} && (6 \text{ Möglichkeiten}) \end{aligned}$$

Insbesondere bemerken wir, dass nur in 3D die Bausteine zu einem dreidimensionalen Vektor vereint werden können. In 4D hat das Analogon zum Kreuzprodukt von zwei Vektoren 6 Komponenten und in 2D sogar nur eine Komponente.

## 3.2 (Affine) Geraden und Ebenen im $\mathbb{R}^n$

**Erinnerung** Ist eine kartesisches Achsenkreuz mit Ursprung  $O$  fixiert, so kann das  $n$ -Tupel

$$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$$

sowohl als Punkt  $A$  als auch als Ortsvektor  $\mathbf{a} = \overrightarrow{OA}$  interpretiert werden.

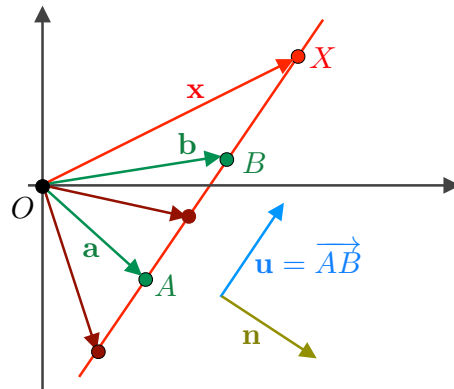
**Geraden im  $\mathbb{R}^n$**  sind eindimensionale Mengen von Punkten bzw. Ortsvektoren, die durch eine der folgenden zwei Gleichungen beschrieben werden können:

Zwei-Punkt-Gleichung mit gegebenen Punkten  $A, B \in \mathbb{R}^n$  und Parameter  $\lambda \in \mathbb{R}$ :

$$\mathbf{x} = \mathbf{a} + \lambda \cdot (\mathbf{b} - \mathbf{a})$$

Punkt-Richtungsgleichung mit  $A \in \mathbb{R}^n$ , Richtung  $\mathbf{u} \in \mathbb{R}^n$  und Parameter  $\lambda \in \mathbb{R}$ :

$$\mathbf{x} = \mathbf{a} + \lambda \cdot \mathbf{u}$$



Eine Gerade in 2D. Der Normalenvektor  $\mathbf{n}$  steht immer senkrecht auf dem Richtungsvektor  $\mathbf{u}$ .

### Bemerkung

1. Die Gerade  $G$  besteht aus allen Punkten  $X$  deren Ortsvektor  $\mathbf{x}$  die Gleichung für eine geeignete Wahl von  $\lambda$  erfüllen, d.h. zum Beispiel

$$G = \{X : \mathbf{x} = \mathbf{a} + \lambda \cdot (\mathbf{b} - \mathbf{a}) \text{ für ein } \lambda \in \mathbb{R}\}.$$

2. Damit die Gerade nicht entartet muss  $\mathbf{a} \neq \mathbf{b}$  bzw.  $\mathbf{u} \neq \mathbf{0}$  gelten. Andernfalls ist die Gerade eigentlich ein Punkt.
3. Beide Darstellungen sind äquivalent (sofern  $\mathbf{u}$  und  $\mathbf{b} - \mathbf{a}$  parallel sind, zum Beispiel  $\mathbf{u} = \mathbf{b} - \mathbf{a}$ ).
4. Die Daten (d.h.  $A$  und  $B$  bzw.  $A$  und  $\mathbf{u}$ ) sind nicht eindeutig.

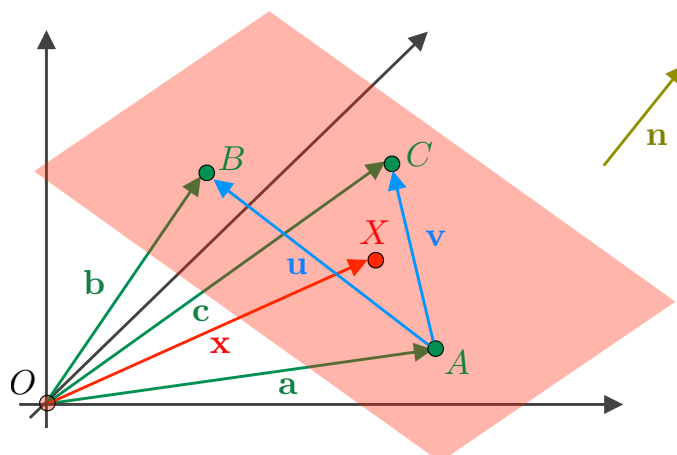
**Ebenen im  $\mathbb{R}^n$**  sind zweidimensional und können wie folgt angegeben werden:

Drei-Punkt-Gleichung mit Punkten  $A, B, C \in \mathbb{R}^n$  und zwei Parametern  $\lambda, \mu \in \mathbb{R}$ :

$$\mathbf{x} = \mathbf{a} + \lambda \cdot (\mathbf{b} - \mathbf{a}) + \mu \cdot (\mathbf{c} - \mathbf{a})$$

Punkt-Richtungsgleichung mit  $A \in \mathbb{R}^n$ , zwei Richtungen  $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$  sowie  $\lambda, \mu \in \mathbb{R}$ :

$$\mathbf{x} = \mathbf{a} + \lambda \cdot \mathbf{u} + \mu \cdot \mathbf{v}$$



Eine Ebene in 3D.

**Bemerkung**

1. Damit die Ebene nicht entartet, dürfen  $\mathbf{u}$  und  $\mathbf{v}$  (bzw.  $\mathbf{b} - \mathbf{a}$  und  $\mathbf{c} - \mathbf{a}$ ) nicht parallel sein und jeweils nicht verschwinden. Andernfalls ist die Ebene eine Gerade oder sogar ein Punkt.
2. Eine (nichtentartete) Ebene in 2D ist immer schon der gesamte  $\mathbb{R}^2$ .

**\* $k$ -dimensionale Unterräume im  $\mathbb{R}^n$**  werden durch  $k+1$  Punkte bzw. durch einen Punkt und  $k$  Richtungen via

$$\mathbf{x} = \left(1 - \sum_{m=1}^k \lambda_m\right) \cdot \mathbf{a} + \sum_{m=1}^k \lambda_m \cdot \mathbf{b}_m \quad \text{bzw.} \quad \mathbf{x} = \mathbf{a} + \sum_{m=1}^k \lambda_m \cdot \mathbf{u}_m$$

beschrieben. Die Nicht-Entartungsbedingung ist, dass alle Richtungsvektoren linear unabhängig sind (siehe dazu den nächsten Abschnitt).

**Geraden im  $\mathbb{R}^2$**  In 2D kann eine (nicht entartete) Gerade auch durch die Gleichung

$$n_1 \cdot x_1 + n_2 \cdot x_2 = \gamma \quad \text{bzw.} \quad \langle \mathbf{n}, \mathbf{u} \rangle = \gamma$$

beschrieben werden, wobei  $\mathbf{n} \neq \mathbf{0}$  ein sogenannter Normalenvektor an die Gerade ist und immer senkrecht auf dem Richtungsvektor  $\mathbf{u}$  steht. Diese Gleichung wird auch Normalform genannt und hat den Vorteil, dass kein Parameter  $\lambda$  mehr auftaucht.

Beweis In Komponenten wird eine Gerade durch die zwei Gleichungen

$$x_1 = a_1 + \lambda \cdot u_1, \quad x_2 = a_2 + \lambda \cdot u_2$$

charakterisiert, wobei  $u_1 \neq 0$  oder  $u_2 \neq 0$  gilt. Im ersten Fall (der zweite geht analog) können wir  $\lambda$  mittels der ersten Gleichung in der zweiten ersetzen. Mit anderen Worten: es gilt

$$\lambda = \frac{x_1 - a_1}{u_1}, \quad x_2 = a_2 + \frac{x_1 - a_1}{u_1} \cdot u_2$$

Nach Multiplikation der zweiten Gleichung mit  $u_1$  folgt die erste Behauptung mit  $n_1 = -u_2$ ,  $n_2 = u_1$ ,  $\gamma = a_2 \cdot u_1 - a_1 \cdot u_2$ .  $\square$

**Aufgabe** Man gebe die Punkt-Richtungsgleichung der Geraden an, die durch die Normalform

$$3 \cdot x_1 - 2 \cdot x_2 = 1$$

gegeben ist.

Lösung Wir suchen zunächst zwei Punkte  $X$  (bzw. ihre Ortvektoren  $\mathbf{x}$ ), die die Normalform genügen und nennen diese  $A$  und  $B$  (bzw.  $\mathbf{a}$  und  $\mathbf{b}$ ), zum Beispiel

$$\mathbf{a} = \begin{pmatrix} 0 \\ -\frac{1}{2} \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} \frac{1}{3} \\ 0 \end{pmatrix}.$$

Anschließend berechnen wir  $\mathbf{u} = \mathbf{b} - \mathbf{a}$ .

**Bemerkung**

1. Die Normalform gibt es so nur in 2D, da in nD eine Gerade nicht durch einen, sondern durch  $n - 1$  Normalenvektoren beschrieben werden muss.
2. Wird die Normalform (auf beiden Seiten) durch eine geeignete reelle Zahl dividiert, so kann man zusätzlich erreichen, dass

$$\|\mathbf{n}\| = 1, \quad \gamma \geq 0$$

gilt und erhält die Hessische Normalform der Geraden, wobei  $\gamma$  dann gerade der Abstand der Geraden vom Ursprung ist.

3. Die Literatur kennt weitere Varianten der Normalform, die aber alle ineinander umgeformt werden können:

$$\begin{aligned} x_2 - a_2 &= m \cdot (x_1 - a_1) && \text{(Punkt-Steigungsform, } m = u_2/u_1) \\ x_2 &= m \cdot x_1 + n && \text{(explizite Normalform, } n = a_2 - a_1) \\ \frac{x_1}{\alpha_1} + \frac{x_2}{\alpha_2} &= 1 && \text{(Achsenabschnittsform)} \end{aligned}$$

**Ebenen im  $\mathbb{R}^2$**  Eine Ebene in 3D kann durch die Normalform

$$n_1 \cdot x_1 + n_2 \cdot x_2 + n_3 \cdot x_3 = \gamma \quad \text{bzw.} \quad \langle \mathbf{n}, \mathbf{x} \rangle = \gamma$$

angegeben werden, wobei der Normalenvektor  $\mathbf{n} \in \mathbb{R}^3$  nun senkrecht auf beiden Richtungsvektoren  $\mathbf{u}$  und  $\mathbf{v}$  steht. Im Fall von  $\|\mathbf{n}\| = 1$  und  $\gamma \geq 0$  spricht man wieder von der Hessischen Normalform der Ebene.

Beweis Mit Hilfe zweier der drei Komponentengleichungen

$$x_1 = a_1 + \lambda \cdot u_1 + \mu \cdot v_1, \quad x_2 = a_2 + \lambda \cdot u_2 + \mu \cdot v_2, \quad x_3 = a_3 + \lambda \cdot u_3 + \mu \cdot v_3$$

können wir  $\lambda$  und  $\mu$  eliminieren. □

**Bemerkung**

1. Jeder Normalenvektor ist parallel zu  $\mathbf{n} = \mathbf{u} \times \mathbf{v}$ .
2. Zwei verschiedene Ebenen in 3D sind genau dann parallel, wenn sie parallele Normalenvektoren besitzen. Andernfalls schneiden Sie sich in einer Geraden.
3. Geraden in 3D können auch durch zwei Ebenen-Normalformen, d.h. durch ein Gleichungssystem der Bauart

$$\langle \mathbf{n}, \mathbf{x} \rangle = \gamma, \quad \langle \tilde{\mathbf{n}}, \mathbf{x} \rangle = \tilde{\gamma}$$

beschrieben werden, sofern  $\mathbf{n}$  und  $\tilde{\mathbf{n}}$  nicht parallel sind.

### 3.3 Allgemeine Vektorräume

**Definition** Jede Menge  $V$ , für die eine Addition (als Abbildung  $V \times V \rightarrow V$ ) und eine skalare Multiplikation (als Abbildung  $\mathbb{R} \times V \rightarrow V$ ) eingeführt werden können, so dass die Rechenregeln (VR.I) und (VR.II) gelten, nennt man reellen Vektorraum (oder auch linearen Raum).

#### Beispiel

1. Der Raum  $\mathbb{R}^n$ .
2. Die Menge der Balkendiagramme aus  $n$  Balken (siehe Bild), wobei Addition und skalare Multiplikation balkenweise definiert sind.
3. Die Menge aller Funktionen  $D \rightarrow \mathbb{R}$  auf der Menge  $D$  wird durch

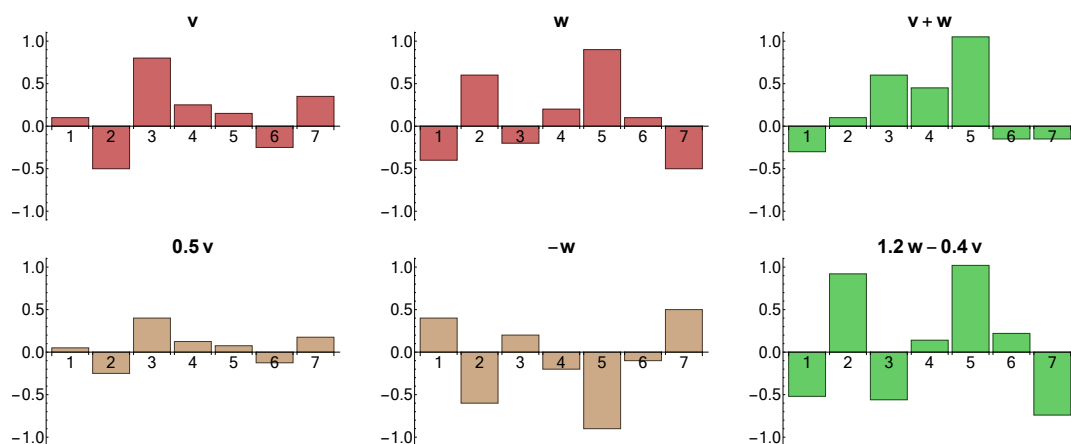
$$(f + g)(x) := f(x) + g(x) \quad \text{und} \quad (\lambda \cdot f)(x) = \lambda \cdot f(x)$$

zu einem Vektorraum ( $x \in D$ ,  $\lambda \in \mathbb{R}$ ,  $f, g : D \rightarrow \mathbb{R}$ ).

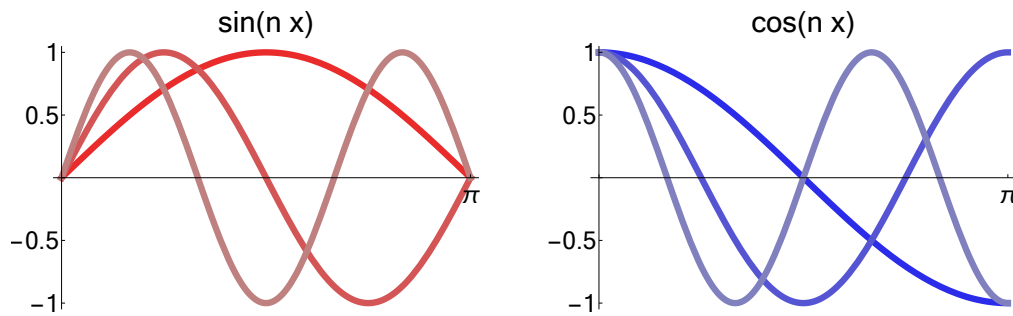
4. Die Menge aller stetigen (bzw. aller stetig differenzierbaren) Funktionen  $I \rightarrow \mathbb{R}$  auf einem Intervall  $I$  (siehe Bild).
5. Die Menge aller  $(m, n)$ -Matrizen (siehe nächstes Kapitel).
6. Der Menge aller Lösungen einer linearen Gleichung (wieder nächstes Kapitel).
7. Die Menge aller reellen Polynome in der Variablen  $x \in \mathbb{R}$  bzw. die Menge aller reellen Polynome vom Grad  $\leq k$ . Es gilt zum Beispiel

$$\begin{aligned} (3 \cdot x^4 + 4 \cdot x^2 - 5 \cdot x + 9) + (-x^3 - x^2 + 2 \cdot x) \\ = 3 \cdot x^4 - x^3 + 3 \cdot x^2 - 3 \cdot x + 9 \end{aligned}$$

und  $(-3) \cdot (x^3 - 2 \cdot x + 4) = -3 \cdot x^2 + 6 \cdot x - 12$ .



Die Menge aller Diagramme aus  $n$  Balken (hier  $n = 7$ ) ist in natürlicher Weise ein Vektorraum und isomorph zu  $\mathbb{R}^n$ .



Die Menge aller stetigen Funktionen auf einem Intervall  $I$  — hier  $I = (0, 2\pi)$  — ist (ein unendlich-dimensionaler) Vektorraum, den wir im Kapitel über Fourier-Reihen genauer studieren werden.

**Definition** Sei  $V$  ein Vektorraum und  $W \subset V$  eine nichtleere Teilmenge von  $V$ , die unter Addition und skalarer Multiplikation abgeschlossen ist, d.h.

$$\mathbf{v}, \mathbf{w} \in W \Rightarrow \mathbf{v} + \mathbf{w} \in W, \quad \mathbf{v} \in W, \lambda \in \mathbb{R} \Rightarrow \lambda \cdot \mathbf{v} \in W.$$

Dann nennt man  $W$  einen Untervektorraum bzw. einen linearen Unterraum von  $V$ .

### Beispiel

1. Jeder Vektorraum  $V$  ist auch Unterraum seiner selbst und jeder Unterraum ist auch wieder ein Vektorraum.
2. Der Raum aller Polynome vom Grad  $\leq 3$  ist ein Unterraum des Raumes aller Polynome.
3. Der Raum der stetig differenzierbaren Funktionen ist ein Unterraum des Raumes aller stetigen Funktionen.
4. Die Menge  $\{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x}\| = 1\}$  ist **kein** Unterraum von  $\mathbb{R}^n$ .
5. Die (affine) Gerade

$$G = \{X \in \mathbb{R}^n : \langle \mathbf{n}, \mathbf{x} \rangle = \gamma\}$$

ist genau dann ein linearer Unterraum des  $\mathbb{R}^n$ , wenn  $\gamma = 0$  bzw.  $O \in G$  gilt, d.h. wenn es sich um eine Ursprungsgerade handelt.

6. Sind  $\mathbf{v}_1, \dots, \mathbf{v}_m$  gegebene Vektoren in  $V$ , so ist

$$\text{span}(\mathbf{v}_1, \dots, \mathbf{v}_m) := \left\{ \sum_{j=1}^m \nu_j \cdot \mathbf{v}_j : \nu_j \in \mathbb{R} \right\}$$

ein Unterraum von  $V$  und wird der (von den  $\mathbf{v}_j$ ) aufgespannte Unterraum genannt. Die Elemente heißen Linearkombinationen der  $\mathbf{v}_j$ .

**Bemerkung** In jedem Vektorraum  $V$  gibt es einen Nullvektor  $\mathbf{0}$  mit  $0 \cdot \mathbf{v} = \mathbf{0}$  für alle  $\mathbf{v} \in V$ .

**Skalarprodukte bzw. Normen** können auch in allgemeinen Vektorräumen betrachtet werden, wobei es sich dann um Abbildungen  $V \rightarrow \mathbb{R}$  bzw.  $V \times V \rightarrow \mathbb{R}$  handelt, für die die Regeln (VR.I) bzw. (VR.II) gelten.



## Lineare Unabhängigkeit und Basen

**Definition** Die Vektoren  $\mathbf{v}_1, \dots, \mathbf{v}_m \in V$  heißen linear unabhängig, sofern die Implikation

$$\sum_{j=1}^m \nu_j \cdot \mathbf{v}_j = \mathbf{0} \quad \Rightarrow \quad \nu_1 = \dots = \nu_m = 0$$

gilt. Andernfalls werden sie linear abhängig genannt. Die Vektoren bilden außerdem eine Basis des Unterraumes  $W$ , wenn sie linear unabhängig sind und wenn  $\text{span}(\mathbf{v}_1, \dots, \mathbf{v}_m) = W$  gilt.

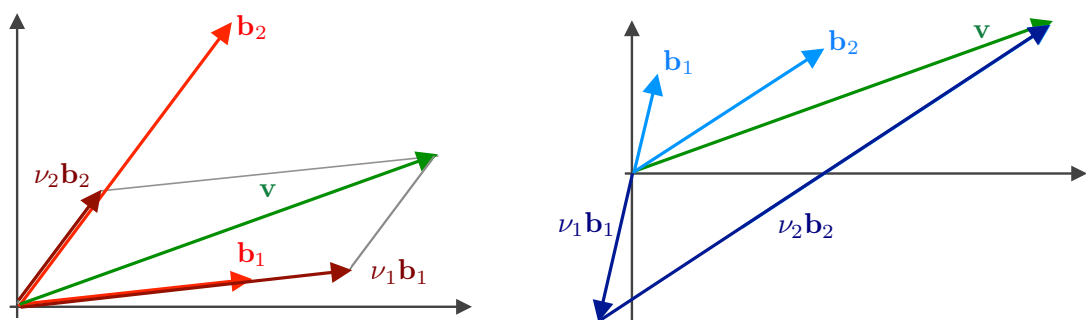
**Theorem** Sei  $\mathbf{b}_1, \dots, \mathbf{b}_m$  eine Basis von  $V$ . Dann existieren für jedes  $\mathbf{v} \in V$  eine eindeutige Wahl von skalaren Größen  $\nu_1, \dots, \nu_m \in \mathbb{R}$ , so dass die Basisdarstellung

$$\mathbf{v} = \sum_{j=1}^m \nu_j \cdot \mathbf{b}_j$$

gilt. Die  $\nu_j$  heißen die Komponenten (oder Koordinaten oder Koeffizienten) des Vektors  $\mathbf{v}$  bzgl. der gewählten Basis.

*Beweis:* Die Existenz der Komponenten folgt, weil  $V$  von den  $\mathbf{b}_j$  aufgespannt wird. Die Eindeutigkeit ergibt sich aus der linearen Unabhängigkeit wie folgt:

$$\begin{aligned} \sum_{j=1}^m \nu_j \cdot \mathbf{b}_j &= \sum_{j=1}^m \tilde{\nu}_j \cdot \mathbf{b}_j \quad \Rightarrow \quad \sum_{j=1}^m (\nu_j - \tilde{\nu}_j) \cdot \mathbf{b}_j = \mathbf{0} \\ &\Rightarrow \quad \nu_j - \tilde{\nu}_j = 0 \quad \text{für } j = 1..m. \end{aligned}$$



Basisdarstellung eines zweidimensionalen Vektors (grün) bzgl. zwei verschiedener Basen (blau und rot).

### Beispiel

1. Die Einheitsvektoren bilden eine Basis im  $\mathbb{R}^n$  — man spricht auch von der kanonischen Basis des  $\mathbb{R}^n$  — und es gilt offensichtlich

$$\begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} = \sum_{j=1}^m v_j \cdot \mathbf{e}_j,$$

d.h. die  $v_j$  sind gerade die Komponenten des Vektors bzgl. der kanonischen Basis.

2. Da

$$\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \frac{v_1 + v_2}{2} \cdot \mathbf{b}_1 + \frac{v_1 - v_2}{2} \cdot \mathbf{b}_2$$

mit

$$\mathbf{b}_1 := \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \mathbf{b}_2 := \begin{pmatrix} 1 \\ -1 \end{pmatrix},$$

für alle  $\mathbf{v} \in \mathbb{R}^2$  gilt, bilden  $\mathbf{b}_1$  und  $\mathbf{b}_2$  eine (nicht-kanonische) Basis des  $\mathbb{R}^2$  (die Formeln für die  $\nu_j$  können abgelesen werden).

3. Die Vektoren

$$\mathbf{b}_1 := \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}, \quad \mathbf{b}_2 := \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{b}_3 := \begin{pmatrix} -1 \\ 1 \\ 3 \end{pmatrix}$$

bilden eine Basis des  $\mathbb{R}^3$ , denn für jedes  $v \in \mathbb{R}^3$  gilt

$$\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \frac{3v_1 - 3v_2 + 2 \cdot v_3}{7} \cdot \mathbf{b}_1 + \frac{2v_1 + 5v_2 - v_3}{7} \cdot \mathbf{b}_2 + \frac{-2v_1 + 2v_2 + v_3}{7} \cdot \mathbf{b}_3.$$

*Bemerkung:* Die Gültigkeit der obigen Formel kann einfach nachgerechnet werden. Zur Herleitung dieser Formel muss ein lineares Gleichungssystem für die Vorfaktoren  $\nu_1, \nu_2, \nu_3$  gelöst werden und wir werden im nächsten Kapitel lernen, wie das geht.

4. Die Monome  $1, x, x^2$  usw. bilden eine Basis im unendlich-dimensionalen Raum aller Polynome in der Variablen  $x$ . (In welchem Sinne unendlich viele Vektoren einen Raum aufspannen, werden wir später mit Hilfe von Reihen genauer diskutieren).

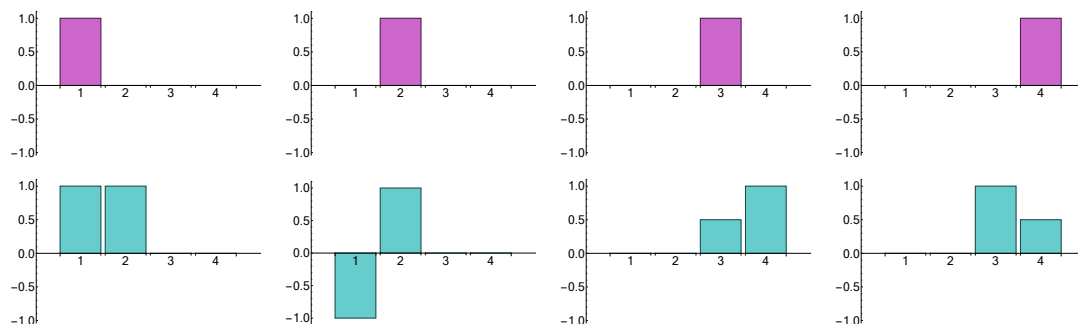
5. Die Funktionen

$$x \mapsto \sin(1 \cdot x), \quad x \mapsto \sin(2 \cdot x), \quad x \mapsto \sin(3 \cdot x), \quad \dots$$

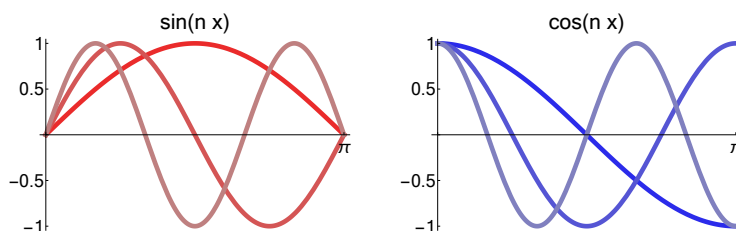
bzw.

$$x \mapsto \cos(0 \cdot x), \quad x \mapsto \cos(1 \cdot x), \quad x \mapsto \cos(2 \cdot x), \quad \dots$$

bilden Basen in gewissen Funktionenräumen auf dem Intervall  $[0, \pi]$ , die wir später genauer untersuchen werden.



Zwei Basen (lila und türkis) im Raum aller Balkendiagramme mit 4 Balken. Dieser Raum ist 4-dimensional.



Einige der Basisfunktionen für die Fourier-Sinus-Reihe (links) und die Fourier-Kosinus-Reihe (rechts) von Funktionen auf dem Intervall  $[0, \pi]$ .

**Definition** Ein Vektorraum  $V$  besitzt genau dann die Dimension  $n \in \mathbb{N}$  (wir schreiben  $\dim V = n$ ), wenn er eine Basis aus  $n$  Vektoren besitzt.

### Bemerkungen

1. Der  $\mathbb{R}^n$  besitzt die Dimension  $n$ .
2. Gilt  $\dim V = n$ , so besteht jede Basis aus  $n$ -Vektoren und je  $n + 1$  Vektoren sind immer linear abhängig (das wird sich aus den Resultaten des folgenden Kapitels ergeben).
3. Wird der Unterraum  $W$  von den Vektoren  $\mathbf{v}_1, \dots, \mathbf{v}_m \in V$  aufgespannt, so gilt  $\dim W \leq m$ . Sind diese Vektoren  $\mathbf{v}_j$  außerdem sogar linear unabhängig, so gilt sogar  $\dim W = m$ , denn die  $\mathbf{v}_j$  bilden ja schon eine Basis von  $W$ .
4. Ist  $W$  ein Unterraum von  $V$ , so gilt  $\dim W \leq \dim V$ .
5. Es gibt unendlich-dimensionale Räume, zum Beispiel Funktionenräume. Die entsprechende Theorie ist aber deutlich komplizierter als die Theorie der endlich-dimensionalen Räume.

**Zum Isomorphiebegriff** Die volle Tragweite des folgenden Konzeptes wird sich uns erst allmählich im Laufe der Vorlesung erschließen: Man nennt zwei Vektorräume  $V$  und  $W$  zueinander isomorph, falls es eine Abbildung  $T : V \rightarrow W$  gibt, so dass:

1.  $T$  ist bijektiv,
2.  $T$  ist linear, d.h. es gilt

$$T(\lambda \cdot \mathbf{v} + \tilde{\lambda} \cdot \tilde{\mathbf{v}}) = \lambda \cdot T(\mathbf{v}) + \tilde{\lambda} \cdot T(\tilde{\mathbf{v}})$$

für alle  $\mathbf{v}, \tilde{\mathbf{v}} \in V$  und  $\lambda, \tilde{\lambda} \in \mathbb{R}$ .

Bemerkung Salopp kann man sagen: Isomorphe Vektorräume sehen gleich aus und alles was man in dem einen verstanden oder berechnet hat, kann man 1 : 1 auf den anderen übertragen.

**Theorem** Jeder Vektorraum der Dimension  $n \in \mathbb{N}$  ist isomorph zum  $\mathbb{R}^n$ .

*Beweis* Wir wählen eine Basis  $\mathbf{b}_1, \dots, \mathbf{b}_n$  in  $V$  und können jeden gegebenen Vektor  $\mathbf{v} \in V$  in der entsprechenden Basisdarstellung schreiben. Wir können damit  $T : V \rightarrow \mathbb{R}^n$  durch

$$T\left(\sum_{j=1}^n \nu_j \cdot \mathbf{b}_j\right) := \begin{pmatrix} \nu_1 \\ \vdots \\ \nu_n \end{pmatrix}$$

definieren und bemerken, dass  $T$  offensichtlich bijektiv ist. Außerdem gilt

$$\begin{aligned} T\left(\lambda \cdot \sum_{j=1}^n \nu_j \cdot \mathbf{b}_j + \tilde{\lambda} \cdot \sum_{j=1}^n \tilde{\nu}_j \cdot \mathbf{b}_j\right) &= T\left(\sum_{j=1}^n (\lambda \cdot \nu_j + \tilde{\lambda} \cdot \tilde{\nu}_j) \cdot \mathbf{b}_j\right) \\ &= \begin{pmatrix} \lambda \cdot \nu_1 + \tilde{\lambda} \cdot \tilde{\nu}_1 \\ \vdots \\ \lambda \cdot \nu_n + \tilde{\lambda} \cdot \tilde{\nu}_n \end{pmatrix} \\ &= \lambda \cdot \begin{pmatrix} \nu_1 \\ \vdots \\ \nu_n \end{pmatrix} + \tilde{\lambda} \cdot \begin{pmatrix} \tilde{\nu}_1 \\ \vdots \\ \tilde{\nu}_n \end{pmatrix} \\ &= \lambda \cdot T\left(\sum_{j=1}^n \nu_j \cdot \mathbf{b}_j\right) + \tilde{\lambda} \cdot T\left(\sum_{j=1}^n \tilde{\nu}_j \cdot \mathbf{b}_j\right). \end{aligned}$$

□

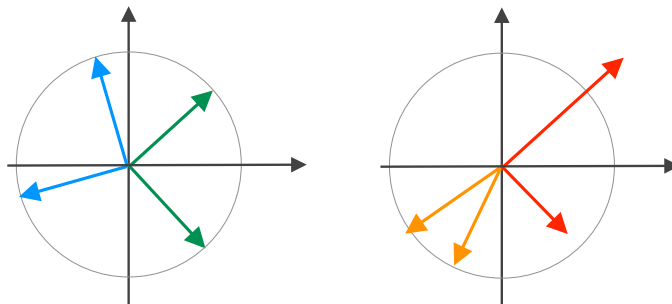
**O-Basen und ON-Basen** In einem  $n$ -dimensionalen Vektorraum mit Skalarprodukt spricht man von einer orthogonalen Basis, sofern

$$\langle \mathbf{b}_i, \mathbf{b}_j \rangle = 0 \quad \text{für alle } i, j = 1 \dots n \text{ mit } i \neq j, \quad \langle \mathbf{b}_i, \mathbf{b}_i \rangle > 0 \quad \text{für alle } i = 1 \dots n$$

Eine orthonormale Basis ist orthogonal und erfüllt außerdem noch die Normierungsbedingung  $\langle \mathbf{b}_i, \mathbf{b}_i \rangle = 1$  für alle  $i = 1 \dots n$ . Man schreibt dann meist

$$\langle \mathbf{b}_i, \mathbf{b}_j \rangle = \delta_{ij},$$

wobei die rechte Seite gerade das schon eingeführte Kronecker-Delta ist.



Zwei Beispiele (blau und grün) sowie zwei Gegenbeispiele (orange und rot) für eine ON-Basis in 2D. Der Kreis ist der Einheitskreis.

Bei einer ON-Basis können die Komponenten eines Vektors  $\mathbf{v}$  leicht mit Hilfe des Skalarproduktes berechnet werden, denn es gilt

$$\langle \mathbf{v}, \mathbf{b}_i \rangle = \left\langle \sum_{j=1}^n \nu_j \cdot \mathbf{b}_j, \mathbf{b}_i \right\rangle = \sum_{j=1}^n \nu_j \cdot \langle \mathbf{b}_j, \mathbf{b}_i \rangle = \sum_{j=1}^n \nu_j \cdot \delta_{ij} = \nu_i.$$

Für O-Basen gilt immerhin noch

$$\langle \mathbf{v}, \mathbf{b}_i \rangle = \nu_i \cdot \|\mathbf{b}_i\|^2,$$

aber bei nicht-orthogonalen Basen muss ein lineares Gleichungssystem gelöst werden, um für einen gegebenen Vektor  $\mathbf{v}$  seine Komponenten  $\nu_j$  zu berechnen.

**Komplexe Vektorräume** Neben den bisher betrachteten reellen Vektorräumen gibt es auch komplexe Vektorräume, für die die skalare Multiplikation mit  $\lambda \in \mathbb{C}$  definiert sein muss. Das  $n$ -dimensionale Standardbeispiel ist der  $\mathbb{C}^n$ , d.h. die Menge aller  $n$ -Tupel komplexer Zahlen.

Die Theorie komplexer Vektorräume kann ganz analog zu der Theorie der reellen Vektorräume entwickelt werden, wobei  $n$  komplexe Dimensionen immer — wegen  $\mathbb{C} \cong \mathbb{R}^2$  —  $2n$  reellen Dimensionen entsprechen. Der einzige Unterschied besteht darin, dass man das Skalarprodukt leicht anders definiert, nämlich durch

$$\langle \mathbf{v}, \mathbf{w} \rangle = \sum_{j=1}^n v_j \cdot \overline{w_j} \in \mathbb{C}$$

für beliebige

$$\mathbf{v} = \begin{pmatrix} v_1 \\ \vdots \\ v_n \end{pmatrix} \in \mathbb{C}^n, \quad \mathbf{w} = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} \in \mathbb{C}^n$$

Damit gilt weiterhin

$$\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle} = \sqrt{\sum_{j=1}^n v_j \cdot \overline{v_j}} = \sqrt{\sum_{j=1}^n |v_j|^2} \geq 0,$$

aber das Skalarprodukt ist nun nicht mehr bilinear sondern sesquilinear. Insbesondere gilt

$$\langle \mathbf{v}, \mathbf{w} \rangle = \overline{\langle \mathbf{w}, \mathbf{v} \rangle}$$

und damit zwar

$$\langle \lambda \cdot \mathbf{u} + \mu \cdot \mathbf{v}, \mathbf{w} \rangle = \lambda \cdot \langle \mathbf{u}, \mathbf{w} \rangle + \mu \cdot \langle \mathbf{v}, \mathbf{w} \rangle$$

aber auch

$$\langle \mathbf{w}, \lambda \cdot \mathbf{u} + \mu \cdot \mathbf{v} \rangle = \overline{\lambda} \cdot \langle \mathbf{w}, \mathbf{u} \rangle + \overline{\mu} \cdot \langle \mathbf{w}, \mathbf{v} \rangle.$$



# Kapitel 4

## Lineare Gleichungssysteme und Matrizen

Vorlesung 10, 11. November 2019

---

**Motivation** Matrizen bilden die Grundlage der Linearen Algebra und spielen eine zentrale Rolle in der Mathematik und allen Anwendungswissenschaften, da sie in natürlicher Weise beim Studium von Gleichungssystemen und von linearen Abbildungen auftauchen.

### 4.1 Matrizen

**Definition** Eine reelle (bzw. komplexe)  $(m, n)$ -Matrix ist

$$\mathbf{A} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$$

d.h. ein quadratisches Schema mit  $m$  Zeilen und  $n$  Spalten, deren Einträge  $a_{ij}$  reelle (oder komplexe) Zahlen sind. Wir schreiben  $\mathbb{R}^{(m,n)}$  (oder  $\mathbb{C}^{(m,n)}$ ) für die Menge aller Matrizen (andere Autoren benutzen  $\mathbb{R}^{m \times n}$  bzw.  $\mathbb{C}^{m \times n}$ ).

**Eiselsbrücke** In der Theorie von Matrizen gilt „Zeilen zuerst, Spalten später“.

**Addition und skalare Multiplikation** Zwei  $(m, n)$ -Matrizen  $\mathbf{A}$  und  $\mathbf{B}$  können komponentenweise addiert und mit einer reellen (oder komplexen) Zahl multipliziert werden. Insbesondere ist  $\mathbb{R}^{(m,n)}$  (oder  $\mathbb{C}^{(m,n)}$ ) ein  $n \cdot m$ -dimensionaler reeller (oder komplexer) Vektorraum. (Übungsaufgabe: Finden Sie Basen in  $\mathbb{R}^{(m,n)}$ .)

**Beispiel**

$$-2 \cdot \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} + 3 \cdot \begin{pmatrix} 1 & 3 & 5 \\ 2 & 4 & 6 \end{pmatrix} = \begin{pmatrix} 1 & 5 & 9 \\ -2 & 2 & 6 \end{pmatrix}$$

**Matrizenmultiplikation** Das Produkt einer  $(l, m)$ -Matrix  $\mathbf{A}$  und einer  $(m, n)$ -Matrix  $\mathbf{B}$  ist eine  $(l, n)$ -Matrix  $\mathbf{C} = \mathbf{A} \cdot \mathbf{B}$ , wobei der Eintrag  $c_{ij}$  gerade das Skalarprodukt des  $i$ -ten Zeilenvektors von  $\mathbf{A}$  und des  $j$ -ten Spaltenvektors von  $\mathbf{B}$  ist. Mit anderen Worten: Es gilt

$$c_{ij} = \sum_{k=1}^m a_{ik} b_{kj}$$

für alle  $i = 1 \dots l$  und alle  $j = 1 \dots n$ .

**Beispiel** Das Produkt einer  $(3, 2)$ -Matrix und einer  $(2, 4)$ -Matrix ist

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix} \cdot \begin{pmatrix} b_{11} & b_{12} & b_{13} & b_{14} \\ b_{21} & b_{22} & b_{23} & b_{24} \end{pmatrix} \\ = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} & a_{11}b_{13} + a_{12}b_{23} & a_{11}b_{14} + a_{12}b_{24} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} & a_{21}b_{13} + a_{22}b_{23} & a_{21}b_{14} + a_{22}b_{24} \\ a_{31}b_{11} + a_{32}b_{21} & a_{31}b_{12} + a_{32}b_{22} & a_{31}b_{13} + a_{32}b_{23} & a_{31}b_{14} + a_{32}b_{24} \end{pmatrix}$$

und damit eine  $(3, 4)$ -Matrix. Beachte, dass das Produkt einer  $(2, 4)$ -Matrix und einer  $(3, 2)$ -Matrix nicht definiert ist.

**Bemerkung** Wir werden im Laufe der Vorlesung sehen, dass diese auf den ersten Blick seltsam anmutende Matrizenmultiplikation sehr sinnvoll ist und in der Natur der Sache begründet liegt. Eine komponentenweise Multiplikation zweier  $(m, n)$ -Matrizen könnte zwar definiert werden, wäre aber ziemlich nutzlos.

**Beispiel** Das Produkt einer  $(2, 2)$ -Matrix und einer  $(2, 2)$ -Matrix ist

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \cdot \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} & a_{11}b_{12} + a_{12}b_{22} \\ a_{21}b_{11} + a_{22}b_{21} & a_{21}b_{12} + a_{22}b_{22} \end{pmatrix}.$$

Ganz allgemein können zwei quadratische  $(n, n)$ -Matrizen zu einer  $(n, n)$ -Matrix multipliziert werden. Beachte aber

$$\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \cdot \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} a_{11}b_{11} + a_{21}b_{12} & a_{12}b_{11} + a_{22}b_{12} \\ a_{11}b_{21} + a_{21}b_{22} & a_{12}b_{21} + a_{22}b_{22} \end{pmatrix}$$

d.h. es gilt im Allgemeinen  $\mathbf{A} \cdot \mathbf{B} \neq \mathbf{B} \cdot \mathbf{A}$ . Die Multiplikation quadratischer Matrizen ist also nicht kommutativ!

**Beispiel**

$$\begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix} \cdot \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix} = \begin{pmatrix} 2 & 2 \\ -2 & -2 \end{pmatrix}$$

Die rechte Seite im ersten Produkt ist übrigens die Null im Vektorraum aller  $(2, 2)$ -Matrizen (bzw. eine sogenannte Nullmatrix), die manchmal auch mit  $\mathbf{0}$  bezeichnet wird. Beachte, dass aus  $\mathbf{A} \cdot \mathbf{B} = \mathbf{0}$  weder  $\mathbf{B} \cdot \mathbf{A} = \mathbf{0}$  noch  $\mathbf{A} = \mathbf{0}$  oder  $\mathbf{B} = \mathbf{0}$  folgt.

**\*Bemerkung** Aus Sicht der Mathematischen Algebra ist die Menge aller  $(n, n)$ -Matrizen ein nichtkommutativer Ring, der nichttriviale Nullteiler besitzt.



**Beispiel** Elemente des  $\mathbb{R}^n$  (bzw. des  $\mathbb{C}^n$ ) können in natürlicher Weise als  $(n, 1)$ -Matrizen betrachtet werden, eben weil wir sie als Spaltenvektoren schreiben. Insbesondere kann eine  $(m, n)$ -Matrix mit einem Spaltenvektor der Dimension  $n$  multipliziert werden, wobei via

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \sum_{k=1}^n a_{1k}x_k \\ \vdots \\ \sum_{k=1}^n a_{mk}x_k \end{pmatrix}$$

ein Spaltenvektor der Dimension  $m$  entsteht.

**Beobachtung** Das lineare Gleichungssystem

$$\begin{aligned} a_{11}x_1 + \dots + a_{1n}x_n &= b_1 \\ &\vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n &= b_m \end{aligned}$$

kann als  $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$  geschrieben werden. Das ist eine erste — aber weder die letzte noch die beste — Motivation, Matrizen zu studieren.

**Rechenregeln der Multiplikation** Aus den obigen Definitionen kann man durch Nachrechnen die folgenden Regeln ableiten:

$$\begin{aligned} (\mathbf{A} + \mathbf{B}) \cdot \mathbf{C} &= \mathbf{A} \cdot \mathbf{C} + \mathbf{B} \cdot \mathbf{C} \\ \mathbf{A} \cdot (\mathbf{B} + \mathbf{C}) &= \mathbf{A} \cdot \mathbf{B} + \mathbf{A} \cdot \mathbf{C} \\ (\mathbf{A} \cdot \mathbf{B}) \cdot \mathbf{C} &= \mathbf{A} \cdot (\mathbf{B} \cdot \mathbf{C}) \\ \mathbf{A} \cdot (\lambda \mathbf{B}) &= (\lambda \cdot \mathbf{A}) \cdot \mathbf{B} = \lambda \cdot (\mathbf{A} \cdot \mathbf{B}) \end{aligned}$$

Hierbei müssen natürlich die Spalten- und Zeilenanzahlen der beteiligten Matrizen jeweils zueinander passen.

**Einheitsmatrix** Für jedes  $n \in \mathbb{N}$  wird die  $(n, n)$ -Matrix

$$\mathbf{I}_n = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix}$$

als Einheitsmatrix bezeichnet, wobei die Schreibweise so zu verstehen ist, dass jeder Diagonaleintrag 1 ist und alle anderen Einträge 0 sind. Es gilt also

$$\mathbf{I}_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathbf{I}_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{I}_4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \text{usw.}$$

Für jede  $(n, n)$ -Matrix rechnet man nach, dass  $\mathbf{A} \cdot \mathbf{I}_n = \mathbf{A} = \mathbf{I}_n \cdot \mathbf{A}$ .

**Die Transponierte einer Matrix** Für jede Matrix

$$\mathbf{A} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \in \mathbb{R}^{(m,n)} \quad \text{wird} \quad \mathbf{A}^T = \begin{pmatrix} a_{11} & \dots & a_{m1} \\ \vdots & & \vdots \\ a_{1n} & \dots & a_{mn} \end{pmatrix} \in \mathbb{R}^{(n,m)}$$

als die Transponierte zu  $\mathbf{A}$  bezeichnet. Analog wird die Transponierte von komplexen Matrizen eingeführt.

### Rechenregeln

$$\begin{aligned} (\mathbf{A} + \mathbf{B})^T &= \mathbf{A}^T + \mathbf{B}^T \\ (\lambda \cdot \mathbf{A})^T &= \lambda \cdot \mathbf{A}^T \\ (\mathbf{A}^T)^T &= \mathbf{A} \\ (\mathbf{A} \cdot \mathbf{B})^T &= \mathbf{B}^T \cdot \mathbf{A}^T \end{aligned}$$

### Beispiel

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}^T = \begin{pmatrix} 1 & 4 \\ 2 & 5 \\ 3 & 6 \end{pmatrix}, \quad \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}^T = \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix}$$

**Bemerkung** Die Transponierte eines Spaltenvektors ist dabei ein Zeilenvektor und umgekehrt, wobei das Produkt aus Zeilen- und Spaltenvektor gerade als Skalarprodukt interpretiert werden kann. Genauer gesagt, für  $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$  gilt

$$\mathbf{x}^T \cdot \mathbf{y} = (x_1 \ \dots \ x_n) \cdot \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \sum_{j=1}^n x_j y_j = \langle \mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{y}, \mathbf{x} \rangle = \mathbf{y}^T \cdot \mathbf{x} \in \mathbb{R}.$$

Beachte aber, dass

$$\mathbf{x} \cdot \mathbf{y}^T = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \cdot (y_1 \ \dots \ y_n) = \begin{pmatrix} x_1 \cdot y_1 & \dots & x_1 \cdot y_n \\ \vdots & & \vdots \\ x_n \cdot y_1 & \dots & x_n \cdot y_n \end{pmatrix} \in \mathbb{R}^{(n,n)}.$$

**Definition** Eine reelle quadratische Matrix  $\mathbf{A} \in \mathbb{R}^{(n,n)}$  heißt:

<u>symmetrisch</u>	falls $\mathbf{A} = \mathbf{A}^T$
<u>orthogonal</u>	falls $\mathbf{A} \cdot \mathbf{A}^T = \mathbf{A}^T \cdot \mathbf{A} = \mathbf{I}_n$
<u>positiv definit</u>	falls $\mathbf{A}$ symmetrisch ist und $\mathbf{x}^T \cdot \mathbf{A} \cdot \mathbf{x} > 0$ für alle $\mathbf{x} \neq \mathbf{0} \in \mathbb{R}^n$

Die Nützlichkeit dieser Begriffe wird sich uns aber erst später erschließen.

**Bemerkung** Bei einer orthogonalen  $(n, n)$ -Matrix bilden die Spaltenvektoren sowie die Zeilenvektoren jeweils eine ON-Basis des  $\mathbb{R}^n$ , zum Beispiel

$$\mathbf{A} = \begin{pmatrix} \cos(\theta) & -\sin(\theta) \\ \sin(\theta) & \cos(\theta) \end{pmatrix}, \quad \theta \in \mathbb{R}.$$

**Operationen mit komplexen Matrizen** Für  $\mathbf{A} \in \mathbb{C}^{(m,n)}$  ist

$$\overline{\mathbf{A}} := \begin{pmatrix} \overline{a_{11}} & \dots & \overline{a_{1n}} \\ \vdots & & \vdots \\ \overline{a_{m1}} & \dots & \overline{a_{mn}} \end{pmatrix} \in \mathbb{C}^{(m,n)}$$

die konjugierte Matrix und  $\mathbf{A}^* := \overline{\mathbf{A}}^T \in \mathbb{C}^{(m,n)}$  die adjungierte Matrix.

**Definition** Eine komplexe quadratische Matrix  $\mathbf{A} \in \mathbb{C}^{(n,n)}$  heißt:

<u>hermitesch</u>	falls $\mathbf{A} = \mathbf{A}^*$
<u>unitär</u>	falls $\mathbf{A} \cdot \mathbf{A}^* = \mathbf{A}^* \cdot \mathbf{A} = \mathbf{I}_n$
<u>positiv definit</u>	falls $\mathbf{A}$ hermitesch ist und $\mathbf{x}^* \cdot \mathbf{A} \cdot \mathbf{x} > 0$ für alle $\mathbf{x} \neq \mathbf{0} \in \mathbb{C}^n$

Vorlesung 11, 13. November 2019

## 4.2 Lineare Gleichungssysteme – Praxis

### Einfache Beispiele und Mächtigkeit der Lösungsmenge

#### Beispiel

$$\begin{array}{rcl} x_1 & + & x_2 = 1 \\ -x_1 & + & 2 \cdot x_2 = 5 \end{array}$$

Wir können die Lösungen – jede Lösung besteht aus je einem Wert für  $x_1$  und  $x_2$  – auf verschiedene Weisen ableiten

1. Wir raten eine Lösung, können uns dann aber nicht sicher sein, ob es noch andere gibt.
2. Wir bestimmen die Lösung graphisch als die Schnittpunkte zweier Geraden in der Ebene (siehe Bild).
3. Wir lösen die erste Gleichung nach  $x_2$  auf und setzen die entstehende Formel in die zweite Gleichung ein. Das ergibt

$$x_2 = 1 - x_1, \quad -x_1 + 2 \cdot (1 - x_1) = 5$$

und nach einfachen Umformungen erhalten wir

$$x_1 = -1, \quad x_2 = 2.$$

Insbesondere gibt es in diesem Beispiel genau eine Lösung.

*Bemerkung:* Wir hätten natürlich auch die erste Gleichung nach  $x_1$  oder die zweite Gleichung nach  $x_2$  auflösen können und letztlich dasselbe Ergebnis erhalten.

4. Wir betrachten geschickte *Linearkombinationen* der Gleichungen. Zweimal die erste Gleichung minus einmal die zweite Gleichung sowie die Summe beider Gleichungen ergeben

$$\begin{array}{l} 2 \cdot (x_1 + x_2) + (-1) \cdot (-x_1 + 2x_2) = 2 \cdot 1 + (-1) \cdot 5, \\ 1 \cdot (x_1 + x_2) + 1 \cdot (-x_1 + 2x_2) = 1 \cdot 1 + 1 \cdot 5 \end{array}$$

und damit auch

$$3 \cdot x_1 + 0 \cdot x_2 = -3, \quad 0 \cdot x_1 + 3 \cdot x_2 = 6.$$

Insbesondere gibt es im umgeformten Gleichungssystem viele Nullen und wir können die Lösung direkt ablesen.

Die ersten drei Lösungsmöglichkeiten sind nur bei kleinen Systemen praktikabel. Die vierte werden wir jedoch zu einem Algorithmus verallgemeinern können, der dann entweder manuell umgesetzt oder (mit einigen Modifikationen) auch im Computer implementiert werden kann.

**Beispiel** Ausgehend von

$$\begin{array}{rcl} x_1 & - & x_2 = 1 \\ -2x_1 & + & 2x_2 = -2 \end{array}$$

bemerken wir, dass zweimal die erste Gleichung plus die zweite Gleichung

$$0 \cdot x_1 + 0 \cdot x_2 = 0$$

liefert, d.h. die beiden Gleichungen sind linear abhängig. Oder anders gesagt: es handelt sich bei dem gegebenen System eigentlich gar nicht um zwei Gleichungen, sondern nur um eine, die nur zweimal leicht verschieden hingeschrieben wurde. Insbesondere gibt es keine eindeutige Lösung, sondern unendlich viele, nämlich

$$x_1 \in \mathbb{R} \text{ beliebig, } \quad x_2 = x_1 - 1.$$

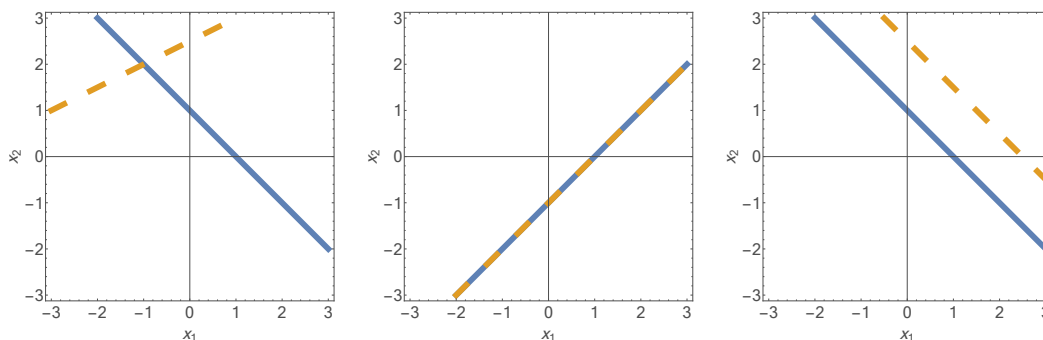
**Beispiel** Für das Gleichungssystem

$$\begin{array}{rcl} x_1 & + & x_2 = 1 \\ 2 \cdot x_1 & + & 2 \cdot x_2 = 5 \end{array}$$

multiplizieren wir die erste Gleichung mit 2 und subtrahieren die zweite Gleichung. Das liefert

$$0 \cdot x_1 + 0 \cdot x_2 = -3$$

und wir schließen, dass es in diesem Beispiel keine Lösung geben kann.



Die geometrische Lösungen der drei linearen Gleichungssysteme aus diesem Abschnitt. Lineare Gleichungssysteme aus  $m$  Gleichungen für  $n$  Unbekannte besitzen immer eine geometrische Interpretation im  $\mathbb{R}^n$  (als Schnittmengen von Geraden, Ebenen, Hyperebenen usw).

**Prinzip** Ganz allgemein kann man sagen: Ein lineares Gleichungssystem bestehend aus  $m$ -Gleichungen für  $n$  Unbekannte kann entweder keine, genau eine, oder unendlich viele (aber niemals genau zwei oder genau 17) Lösungen besitzen. Wir werden das weiter unten besser verstehen.

Gibt es keine bzw. unendlich viele Lösungen, so nennt man das Gleichungssystem überbestimmt bzw. unterbestimmt.

**Achtung** Man könnte meinen, dass ein lineares Gleichungssystem mit  $m < n$  (weniger Gleichungen als Unbekannte) oder  $m > n$  (mehr Gleichungen als Unbekannte) immer unter- bzw. immer überbestimmt ist. Das gilt aber in dieser strengen Form nicht immer, sondern nur, wenn die Gleichungen linear unabhängig sind und wenn sie sich nicht widersprechen. Genausowenig kann man sagen (siehe die Beispiele), dass ein System von  $n$  linearen Gleichungen für  $n$  Unbekannte immer eine eindeutige Lösung hat.

**Bemerkung** Wir entwickeln in diesem Abschnitt die Lösungstheorie linearer Gleichungen. Nichtlineare Gleichungssysteme wie zum Beispiel

$$x_1 - x_2^2 = 1, \quad x_1^2 + x_2^2 + 3 \cdot x_2 = 5$$

sind zwar auch sehr wichtig (in der Mathematik und in den Anwendungen), aber im Allgemeinen deutlich schwieriger zu verstehen.

**Bemerkung** Bei überbestimmten Gleichungssystemen kann mittels der Methode der *Kleinsten Fehlerquadrate* eine optimale approximative Lösung gefunden werden, siehe zum Beispiel [ABHKLS, Seite 508] für eine Einführung in dieses Thema.

## Gauß-Elimination und Stufenform

**Strategie** Durch elementare Operationen kann jedes Gleichungssystem schrittweise in eine Stufenform gebracht werden, aus der dann alle Lösungen durch direkte Rechnungen abgeleitet werden können.

**Vereinfachende Notation** Für das gegebene System

$$\begin{array}{ccccccc} a_{11}x_1 & + & \dots & + & a_{1n}x_n & = & b_1 \\ \vdots & & & & \vdots & & \vdots \\ a_{m1}x_1 & + & \dots & + & a_{mn}x_n & = & b_m \end{array}$$

werden

$$\begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \quad \text{bzw.} \quad \left( \begin{array}{ccc|c} a_{11} & \dots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \dots & a_{mn} & b_m \end{array} \right)$$

die Koeffizientenmatrix bzw. die erweiterte Koeffizientenmatrix genannt. Beachte, dass wir  $x_1, \dots, x_n$  in Abhängigkeit der Koeffizienten  $a_{ij}$  und  $b_i$  suchen.

**Klarstellung** Wir werden im Folgenden ausschließlich mit reellen Zahlen rechnen, d.h. es gilt  $a_{ij} \in \mathbb{R}$  und  $b_i \in \mathbb{R}$  sowie  $x_j \in \mathbb{R}$ . Alle wesentlichen Argumente und Algorithmen können aber leicht auf den Fall komplexer Zahlen übertragen werden.

**Beispiel** Wir untersuchen das Gleichungssystem

$$\left( \begin{array}{ccc|c} 1 & -2 & 1 & 3 \\ 3 & 2 & -1 & 1 \\ -2 & 2 & 1 & 0 \end{array} \right)$$

und erzeugen zunächst in der ersten Spalte Nullen, indem wir die erste Zeile mit  $-3$  bzw.  $+2$  multiplizieren und sie dann zur zweiten bzw. dritten Zeile addieren:

$$\left( \begin{array}{ccc|c} 1 & -2 & 1 & 3 \\ 0 & 8 & -4 & -8 \\ 0 & -2 & 3 & 6 \end{array} \right).$$

Anschließend multiplizieren wir die dritte Zeile mit 4 und addieren zu ihr die zweite Zeile:

$$\left( \begin{array}{ccc|c} 1 & -2 & 1 & 3 \\ 0 & 8 & -4 & -8 \\ 0 & 0 & 8 & 16 \end{array} \right).$$

Wir haben damit eine Stufenform erreicht und können die Lösung schrittweise durch sogenanntes Rückwärtseinsetzen bestimmen:

$$x_3 = \frac{16}{8} = 2, \quad x_2 = \frac{-8 + 4 \cdot x_3}{8} = 0 \quad x_1 = 3 + 2 \cdot x_2 - x_3 = 3 - x_3 = 1.$$

Insbesondere ist die Lösung in diesem Beispiel eindeutig und kann als

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}$$

geschrieben werden.

**Beispiel** Wir betrachten die erweiterte Koeffizientenmatrix

$$\left( \begin{array}{ccccc|c} 0 & 2 & -2 & -1 & 1 & 0 \\ 1 & 2 & 0 & 3 & 2 & -2 \\ -1 & 1 & -3 & 2 & 1 & 1 \end{array} \right)$$

mit 3 Gleichungen und 5 Unbekannten. Um in der ersten Spalte möglichst viele Nullen zu erzeugen, vertauschen wir zunächst die ersten beiden Zeilen und addieren anschließend die dann erste Zeile zur dritten Zeile:

$$\left( \begin{array}{ccccc|c} 1 & 2 & 0 & 3 & 2 & -2 \\ 0 & 2 & -2 & -1 & 1 & 0 \\ 0 & 3 & -3 & 5 & 3 & -1 \end{array} \right).$$

Im nächsten Schritt multiplizieren wir die dritte Zeile mit  $-2$  und addieren das Dreifache der zweiten Zeile:

$$\left( \begin{array}{ccccc|c} 1 & 2 & 0 & 3 & 2 & -2 \\ 0 & 2 & -2 & -1 & 1 & 0 \\ 0 & 0 & 0 & -13 & -3 & +2 \end{array} \right).$$

Damit haben wir auch in diesem Beispiel eine Stufenform erreicht und können die Lösung wie folgt ablesen: Die dritte Zeile der Stufenform impliziert

$$x_4 = -\frac{3}{13}x_5 - \frac{2}{13},$$

wobei  $x_5$  beliebig gewählt werden kann, und die anschließende Auswertung der zweiten Stufenzeile liefert

$$x_2 = x_3 + \frac{x_4 - x_5}{2} = x_3 - \frac{8}{13}x_5 - \frac{1}{13}.$$

Schließlich erzwingt die erste Stufe

$$x_1 = -2 - 2x_2 - 3x_4 - 2x_5 = -2x_3 - \frac{1}{13}x_5 - \frac{18}{13}$$

und wir erhalten insgesamt

$$\mathbf{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} = \frac{1}{13} \begin{pmatrix} -18 \\ -1 \\ 0 \\ -2 \\ 0 \end{pmatrix} + x_3 \begin{pmatrix} -2 \\ 1 \\ 1 \\ 0 \\ 0 \end{pmatrix} + \frac{x_5}{13} \begin{pmatrix} -1 \\ -8 \\ 0 \\ -3 \\ 13 \end{pmatrix}$$

als Formel für die allgemeine Lösung, wobei  $x_3$  und  $x_5$  jeweils beliebig gewählt werden dürfen.

**Beispiel** Die erweiterte Koeffizientenmatrix

$$\left( \begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & 1 & 1 \\ 2 & -2 & -4 & 2 \\ -1 & 1 & 2 & 1 \end{array} \right)$$

kann durch Addition geeigneter Vielfacher der ersten Zeile zur zweiten und dritten Zeile in

$$\left( \begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & -2 & -2 & 2 \\ 0 & 1 & 1 & 1 \end{array} \right)$$

umgeformt werden, und anschließend durch Addition geeigneter Vielfacher der zweiten Zeile zur dritten und vierten Zeile in die Stufenform

$$\left( \begin{array}{ccc|c} 1 & 0 & -1 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 4 \\ 0 & 0 & 0 & 0 \end{array} \right)$$

gebracht werden. Die vierte transformierte Gleichung formuliert dabei überhaupt keine Bedingung an die  $x_i$ , aber die dritte transformierte Gleichung zeigt, dass es in diesem Beispiel keine Lösung gibt. Es handelt sich also um ein überbestimmtes Gleichungssystem.

**Gauß-Elimination** Durch sukzessives Anwenden der Umformungsregeln

1. Vertausche zwei Zeilen.
2. Multipliziere eine Zeile mit einer von 0 verschiedenen Zahl.
3. Addiere ein geeignetes Vielfaches einer Zeile zu einer anderen Zeile.

kann jedes lineare Gleichungssystem in eine äquivalente Stufenform gebracht werden, wobei Stufenform meint, dass die Anzahl der Null-Koeffizienten in jeder Zeile zunimmt. Die allgemeine Lösung kann dann durch Rückwärtseinsetzen bestimmt werden.

**Bemerkung**

1. Bei der Anwendung der Gauß-Elimination hat man gewisse Freiheiten (zum Beispiel bei der Vertauschung der Zeilen). Das Endergebnis hängt aber nicht von den konkret gewählten Schritten ab.
2. Der erste von 0 verschiedene Eintrag in jeder Zeile wird auch Pivot-Element der Zeile genannt.
3. Varianten der Gauß-Elimination werden auch in der *Numerischen Mathematik* zur approximativen Lösung linearer Gleichungssysteme verwendet. Allerdings muss dann durch sorgfältiges Vertauschen der Zeilen sichergestellt werden, dass die Pivot-Elemente nicht zu klein werden, da andernfalls die unvermeidbaren Rundungsfehler zu groß werden können. Oder anders gesagt: Die Gauß-Elimination ist anfällig für Rundungsfehler und bei der numerischen Implementierung der Gauß-Eliminierung mit Gleitkommazahlen gibt es weitere Regeln zu beachten.
4. Manchmal wird in der Literatur bei der Gauß-Elimination das Vertauschen zweier Spalten zugelassen, was einer Umnummerierung der Unbekannten entspricht. Sie sollten dies nur dann tun, wenn Sie ausreichend Erfahrung besitzen und die wechselnde Nummerierung der Unbekannten immer konsequent und konsistent berücksichtigen.
5. Man kann die Gauß-Elimination als eine *LR-Zerlegung* der Koeffizientenmatrix interpretieren, d.h. man berechnet eigentlich eine Darstellung der Form

$$\mathbf{A} = \mathbf{L} \cdot \mathbf{R},$$

wobei  $R$  eine rechte obere Dreiecksmatrix ist (die Stufenform) und die Transformationsschritte in einer linken unteren Dreiecks-Matrix kodiert werden können. Für die Details sei auf WIKIPEDIA oder [AORS, Abschnitt 4.4] verwiesen.

**Notation** Es gibt keine Standardnotation, um die Ergebnisse der Einzelschritte bei einer Gauß-Elimination aufzuschreiben. Eine mögliche und praktische Notation benutzt das Äquivalenzzeichen  $\cong$ , d.h. es gilt zum Beispiel

$$\left( \begin{array}{cc|c} 2 & 1 & -1 \\ 1 & 3 & 1 \end{array} \right) \cong \left( \begin{array}{cc|c} 2 & 1 & -1 \\ 0 & 5 & 3 \end{array} \right) \quad \left[ \text{(I)}, -\text{(I)} + 2 \cdot \text{(II)} \right]$$

wobei die eckigen Klammern uns daran erinnern, dass die erste Zeile unverändert übernommen wurde und dass die zweite Zeile rechts als gewisse Linearkombination der ersten beiden Zeilen links berechnet wurde.



## Gauß-Jordan-Verfahren und Diagonalmatrizen

**Beobachtung** Man kann durch Zeilenoperationen weitere Nullen in der transformierten Koeffizientenmatrix erzeugen und im Fall von  $m = n$  die Koeffizientenmatrix ggf. sogar in eine Diagonalform bringen.

**Beispiel** Wir betrachten das Gleichungssystem

$$\left( \begin{array}{cccc|c} 0 & 1 & 1 & 0 & 1 \\ 1 & 1 & 0 & -1 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 1 & 0 \end{array} \right)$$

und benutzen die Gauß-Elimination um nacheinander die äquivalent umgeformten Systeme

$$\left( \begin{array}{cccc|c} 1 & 1 & 0 & -1 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & -1 & 0 & 1 & 0 \\ 0 & -1 & 1 & 1 & 0 \end{array} \right), \quad \left( \begin{array}{cccc|c} 1 & 1 & 0 & -1 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 2 & 1 & 1 \end{array} \right), \quad \left( \begin{array}{cccc|c} 1 & 1 & 0 & -1 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & -1 & -1 \end{array} \right)$$

abzuleiten, aus dem die Lösung durch Rückwärtsensetzen bestimmt werden könnte. Man kann alternativ aber auch – und das nennt man das Gauß-Jordan-Verfahren – weitere Umformungsschritte durchführen, indem man hier zunächst zu jeder Zeile ein Vielfaches der vierten Zeile addiert:

$$\left( \begin{array}{cccc|c} 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & -1 \end{array} \right).$$

Anschließend kann man mit Hilfe der dritten bzw. zweiten Zeile weiter umformen

$$\left( \begin{array}{cccc|c} 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & -1 \end{array} \right), \quad \left( \begin{array}{cccc|c} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & -1 \end{array} \right)$$

bis schließlich eine Diagonalmatrix entsteht. Im konkreten Beispiel kann man – da es sich um eine quadratische Matrix handelt und alle Diagonaleinträge nicht verschwinden – im letzten Schritt jede Zeile mit einer geeigneten Zahl multiplizieren. Dies liefert

$$\left( \begin{array}{cccc|c} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{array} \right)$$

und damit die eindeutige Lösung

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}.$$

**Beispiel** Man kann natürlich den Umweg über Stufenmatrizen vermeiden und gleich versuchen, möglichst viele Nullen in jeder Spalte zu produzieren (diese Variante nennt man auch Gauß-Jordan-Verfahren): Ausgehend von

$$\left( \begin{array}{cccc|c} 1 & 2 & 0 & 1 & 1 \\ 1 & 2 & -2 & 0 & 0 \\ 1 & 1 & -1 & 1 & 0 \end{array} \right)$$

gelingt dies zum Beispiel via

$$\left( \begin{array}{cccc|c} 1 & 2 & 0 & 1 & 1 \\ 0 & 0 & -2 & -1 & -1 \\ 0 & -1 & -1 & 0 & -1 \end{array} \right), \quad \left( \begin{array}{cccc|c} 1 & 2 & 0 & 1 & 1 \\ 0 & -1 & -1 & 0 & -1 \\ 0 & 0 & -2 & -1 & -1 \end{array} \right)$$

sowie

$$\left( \begin{array}{cccc|c} 1 & 0 & -2 & 1 & -1 \\ 0 & -1 & -1 & 0 & -1 \\ 0 & 0 & -2 & -1 & -1 \end{array} \right), \quad \left( \begin{array}{cccc|c} 1 & 0 & 0 & 2 & 2 \\ 0 & -1 & 0 & \frac{1}{2} & -\frac{1}{2} \\ 0 & 0 & -2 & -1 & -1 \end{array} \right)$$

und

$$\left( \begin{array}{cccc|c} 1 & 0 & 0 & 2 & 2 \\ 0 & 1 & 0 & -\frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 1 & \frac{1}{2} & \frac{1}{2} \end{array} \right).$$

Die allgemeine Lösung ist also:

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 4 \\ 1 \\ 1 \\ 0 \end{pmatrix} + \frac{x_4}{2} \begin{pmatrix} -4 \\ 1 \\ -1 \\ 2 \end{pmatrix}.$$

**Zusammenfassung** Die Grundidee der Gauß-Elimination ist es, eine Stufenform zu erzeugen, aus der alle Lösungen durch Rückwärtseinsetzen gewonnen werden können. Die Grundidee des Gauß-Jordan-Verfahrens ist es, durch Umformungen möglichst viele Nullen und pivotale Einsen auf der linken Seite der Mittellinie zu erzeugen, um aufwändige Rechnungen beim Rückwärtseinsetzen zu vermeiden. Beide Strategien basieren dabei auf denselben elementaren Zeilenoperationen und liefern bei fehlerfreier Umsetzung das gleiche Ergebnis.

**gleichzeitiges Lösen mehrerer Gleichungen** Gleichungssysteme mit gleicher Koeffizientenmatrix, aber verschiedener rechter Seite, können parallel gelöst werden, indem die Koeffizientenmatrix um mehrere Spalten erweitert wird.

**Beispiel** Um das Gleichungssystem

$$x_1 + 2x_2 = b_1, \quad x_1 + 4x_4 = b_2$$

für die zwei rechten Seiten

$$\mathbf{b} = \begin{pmatrix} 0 \\ -8 \end{pmatrix} \quad \text{bzw.} \quad \mathbf{b} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

zu lösen, betrachten wir die erweiterte Koeffizientenmatrix

$$\left( \begin{array}{cc|cc} 1 & 2 & 0 & 1 \\ 1 & 4 & -8 & 1 \end{array} \right)$$

und transformieren diese via

$$\left( \begin{array}{cc|cc} 1 & 2 & 0 & 1 \\ 0 & 2 & -8 & 0 \end{array} \right), \quad \left( \begin{array}{cc|cc} 1 & 0 & +8 & 1 \\ 0 & 2 & -8 & 0 \end{array} \right), \quad \left( \begin{array}{cc|cc} 1 & 0 & +8 & 1 \\ 0 & 1 & -4 & 0 \end{array} \right).$$

Insbesondere gilt

$$\mathbf{x} = \begin{pmatrix} 8 \\ -4 \end{pmatrix} \quad \text{bzw.} \quad \mathbf{x} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

**Beispiel** Wir betrachten die erweiterte Koeffizientenmatrix

$$\left( \begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & -1 & 0 & 1 & 0 \\ 1 & -1 & 2 & 0 & 0 & 1 \end{array} \right),$$

in der die Einheitsmatrix  $I_3$  rechts von der Mittellinie steht. Nach den Umformungen

$$\left( \begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & -2 & -1 & 1 & 0 \\ 0 & -1 & 1 & -1 & 0 & 1 \end{array} \right), \quad \left( \begin{array}{ccc|ccc} 1 & 0 & 1 & 1 & 0 & 0 \\ 0 & 1 & -2 & -1 & 1 & 0 \\ 0 & 0 & -1 & -2 & 1 & 1 \end{array} \right)$$

und

$$\left( \begin{array}{ccc|ccc} 1 & 0 & 0 & -1 & 1 & 1 \\ 0 & 1 & 0 & 3 & -1 & -2 \\ 0 & 0 & -1 & -2 & 1 & 1 \end{array} \right), \quad \left( \begin{array}{ccc|ccc} 1 & 0 & 0 & -1 & 1 & 1 \\ 0 & 1 & 0 & 3 & -1 & -2 \\ 0 & 0 & 1 & 2 & -1 & -1 \end{array} \right)$$

steht die Einheitsmatrix schließlich links von der Mittellinie. Insbesondere implizieren unsere Rechnungen — siehe den Abschnitt unten über inverse Matrizen — die Matrixgleichung

$$\mathbf{A} \cdot \mathbf{B} = \mathbf{I}_3 = \mathbf{B} \cdot \mathbf{A}, \quad \mathbf{A} = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & -1 \\ 1 & -1 & 2 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} -1 & 1 & 1 \\ 3 & -1 & -2 \\ 2 & -1 & -1 \end{pmatrix} = \mathbf{A}^{-1}$$

und wir schließen, dass das Gauß-Jordan-Verfahren zur *Berechnung inverser Matrizen* geeignet ist. Sollte die Inverse nicht existieren, so entsteht eine transformierte Koeffizientenmatrix, in der mindestens eine Zeile kein von Null verschiedenes Pivot-Element besitzt.

## 4.3 Lineare Gleichungssysteme – Theorie

**Klarstellung** Alle Resultate aus diesem Abschnitt gelten analog auch für komplexe Matrizen.

**Motivation** Wir betrachten in diesem Abschnitt das Gleichungssystem

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{b},$$

wobei wir in aller Regel die Lösung  $\mathbf{x} \in \mathbb{R}^n$  suchen und  $\mathbf{A} \in \mathbb{R}^{(m,n)}$  sowie  $\mathbf{b} \in \mathbb{R}^m$  gegeben sind. Dabei heißt das Gleichungssystem homogen, falls  $\mathbf{b} = \mathbf{0}$  gilt, andernfalls inhomogen. Beachte, dass es im homogenen Fall immer die triviale Lösung  $\mathbf{x} = \mathbf{0}$  gibt.

**Theorem (Lösungsraum linearer Gleichungen)**

1. Die Menge aller Lösungen der homogenen Gleichung ist ein linearer Unterraum von  $\mathbb{R}^n$ , den man den Kern der Matrix  $\mathbf{A}$  nennt – bzw. mit  $\ker(\mathbf{A})$  abkürzt – und dessen Dimension wir mit  $k \leq n$  bezeichnen. Insbesondere gibt es  $k$  Basisvektoren  $\mathbf{x}^{(1)}, \dots, \mathbf{x}^{(k)}$ , so dass die allgemeine homogene Lösung von  $\mathbf{A} \cdot \mathbf{x} = \mathbf{0}$  als

$$\mathbf{x} = \sum_{j=1}^k \alpha_j \cdot \mathbf{x}^{(j)}$$

mit  $\alpha_j \in \mathbb{R}$  geschrieben werden kann.

2. Wurde eine spezielle Lösung  $\mathbf{x}^{(*)}$  der inhomogenen Gleichung gefunden, so kann die allgemeine inhomogene Lösung als

$$\mathbf{x} = \mathbf{x}^{(*)} + \sum_{j=1}^k \alpha_j \cdot \mathbf{x}^{(j)}$$

dargestellt werden.

Beweis, Teil 1: Die Menge

$$\ker(\mathbf{A}) := \{\mathbf{x} \in \mathbb{R}^n : \mathbf{A} \cdot \mathbf{x} = \mathbf{0}\}$$

ist offensichtlich Unterraum, denn für alle  $\mathbf{x}, \tilde{\mathbf{x}} \in \ker(\mathbf{A})$  und alle  $\lambda, \tilde{\lambda} \in \mathbb{R}$  gilt wegen

$$\mathbf{A} \cdot (\lambda \cdot \mathbf{x} + \tilde{\lambda} \cdot \tilde{\mathbf{x}}) = \lambda \cdot \mathbf{A} \cdot \mathbf{x} + \tilde{\lambda} \cdot \mathbf{A} \cdot \tilde{\mathbf{x}} = \lambda \cdot \mathbf{0} + \tilde{\lambda} \cdot \mathbf{0} = \mathbf{0}$$

auch  $\lambda \cdot \mathbf{x} + \tilde{\lambda} \cdot \tilde{\mathbf{x}} \in \ker \mathbf{A}$ .

Beweis, Teil 2: Für jede Lösung  $\mathbf{x}$  der inhomogenen Gleichung gilt

$$\mathbf{A} \cdot (\mathbf{x} - \mathbf{x}^{(*)}) = \mathbf{A} \cdot \mathbf{x} - \mathbf{A} \cdot \mathbf{x}^{(*)} = \mathbf{b} - \mathbf{b} = \mathbf{0},$$

d.h. die Differenz zweier inhomogener Lösungen ist eine Lösung der homogenen Gleichung.  $\square$

**Bemerkung** Sofern es überhaupt Lösungen der inhomogenen Gleichung gibt, hängt die Anzahl der Lösungen der Gleichung  $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$  entscheidend von den Eigenschaften der Matrix  $\mathbf{A}$  ab.

**Beispiel** Für

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & -3 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 3 \\ 0 \end{pmatrix}$$

gilt

$$\ker(\mathbf{A}) = \text{span} \left\{ \begin{pmatrix} -2 \\ 3 \\ 1 \end{pmatrix} \right\}$$

und damit  $k = 1$ , denn die Gleichung  $\mathbf{A} \cdot \mathbf{x} = \mathbf{0}$  impliziert

$$x_1 + 2x_3 = 0, \quad x_2 - 3x_3 = 0$$

und damit

$$\mathbf{x} \in \ker \mathbf{A} \quad \Longrightarrow \quad \mathbf{x} = \begin{pmatrix} -2x_3 \\ 3x_3 \\ x_3 \end{pmatrix} = x_3 \begin{pmatrix} -2 \\ 3 \\ 1 \end{pmatrix}.$$

Außerdem ist

$$\mathbf{x}^{(*)} = \begin{pmatrix} 1 \\ 3 \\ 1 \end{pmatrix}$$

eine spezielle inhomogene Lösung (Nachrechnen!). Die allgemeine Lösung der inhomogenen Gleichung ist damit

$$\mathbf{x} = x_3 \begin{pmatrix} -2 \\ 3 \\ 1 \end{pmatrix} \quad \text{bzw.} \quad \mathbf{x} = \begin{pmatrix} 1 \\ 3 \\ 1 \end{pmatrix} + x_3 \begin{pmatrix} -2 \\ 3 \\ 1 \end{pmatrix},$$

wobei  $x_3 \in \mathbb{R}$  jeweils beliebig gewählt werden darf.

**Beispiel** Mit

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 0 & -1 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 2 \\ -3 \end{pmatrix}$$

gilt  $k = 0$ , d.h.  $\ker(\mathbf{A})$  ist der sogenannte Nullraum  $\{\mathbf{0}\}$ , der nur aus der trivialen Lösung besteht, und es gilt  $k = 0$ . Insbesondere gibt es genau eine Lösung des inhomogenen Problems, nämlich

$$\mathbf{x} = \begin{pmatrix} -4 \\ 3 \end{pmatrix}.$$

## Rang von Matrizen

**Definition** Die Anzahl der linear unabhängigen Zeilenvektoren der Matrix  $\mathbf{A}$  wird Zeilenrang von  $\mathbf{A}$  genannt und mit  $\text{Rang}(\mathbf{A})$  bezeichnet.

**Theorem (Dimensionsformel)** Für jede  $(m, n)$ -Matrix  $\mathbf{A}$  gilt

$$\dim(\ker(\mathbf{A})) + \text{Rang}(\mathbf{A}) = n$$

*Beweis* Wir bemerken zunächst, dass die Umstellungen der Gauß-Elimination weder den Rang noch den Kern der Matrix ändern (Übungsaufgabe). Mit  $\mathbf{b} = \mathbf{0}$  bricht die Gauß-Elimination nach  $r$  Schritten ab und liefert die erweiterte Koeffizientenmatrix

$$\left( \begin{array}{cccccccc|c} \tilde{a}_{11} & \tilde{a}_{12} & \dots & \tilde{a}_{1(r-1)} & \tilde{a}_{1r} & \tilde{a}_{1(r+1)} & \dots & \tilde{a}_{1n} & 0 \\ 0 & \tilde{a}_{22} & \dots & \tilde{a}_{2(r-1)} & \tilde{a}_{2r} & \tilde{a}_{2(r+1)} & \dots & \tilde{a}_{2n} & 0 \\ \vdots & & & & & & & & \vdots \\ 0 & 0 & \dots & 0 & \tilde{a}_{rr} & \tilde{a}_{r(r+1)} & \dots & \tilde{a}_{rn} & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \\ 0 & 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 \end{array} \right),$$

wobei wir hier der Einfachheit halber annehmen wollen, dass alle Pivot-Elemente  $\tilde{a}_{11}, \tilde{a}_{22}, \dots, \tilde{a}_{rr}$  jeweils nicht verschwinden (um dies in wirklich jedem Fall sicherzustellen, muss man ggf. die Unbekannten umnummerieren bzw. Spalten vertauschen). Man sieht nun leicht, dass in einer homogenen Lösung — d.h. in einem Vektor  $\mathbf{x} \in \mathbb{R}^n$  mit  $\mathbf{A} \cdot \mathbf{x} = \mathbf{0}$  — die Komponenten  $x_{r+1}, \dots, x_n$  beliebig gewählt werden können und dass dann alle anderen Komponenten  $x_1, \dots, x_r$  durch Rückwärtseinsetzen bestimmt werden können. Die Stufenmatrix zeigt also, dass der Kern von  $\mathbf{A}$  genau  $k = n - r$  Dimensionen (bzw. Freiheitsgrade) besitzt. Außerdem gibt  $r$  gerade die Anzahl der linear unabhängigen Zeilen an, d.h.  $r$  ist der Zeilenrang der Matrix  $\mathbf{A}$ .  $\square$

Klarstellung

**Klarstellung** Die obige Dimensionsformeln gilt sowohl für quadratische und als auch für nicht-quadratische Matrizen, wobei die Dimension des Kernes und der Zeilenrang sich immer zur Anzahl der Spalten addieren. Die folgenden Äquivalenzaussagen gelten in der angegebenen Form nur für quadratische Matrizen.

**Folgerung (Lösbarkeitskriterium bei quadratischen Matrizen)** Die folgenden vier Aussagen sind für jede quadratische  $(n, n)$ -Matrix zueinander äquivalent.

1. Es gilt  $\dim(\ker(\mathbf{A})) = 0$ .
2. Es gilt  $\text{Rang}(\mathbf{A}) = n$ .
3. Die homogene Gleichung  $\mathbf{A} \cdot \mathbf{x} = \mathbf{0}$  besitzt nur die triviale Lösung  $\mathbf{x} = \mathbf{0}$ .
4. Die inhomogene Gleichung  $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$  besitzt für jedes  $\mathbf{b} \in \mathbb{R}^n$  genau eine Lösung.

*Beweis* Das ist eine Übungsaufgabe zur Anwendung der Dimensionsformel und des Satzes über den Lösungsraum linearer Gleichungen. Beachte, dass sehr viele Implikationen zu zeigen sind, nämlich

$$\text{Aussage } i \quad \Rightarrow \quad \text{Aussage } j$$

für jede Wahl von  $i, j \in \{1, 2, 3, 4\}$  mit  $i \neq j$ .

**Beispiel** Für

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 0 \\ 0 & 3 & -4 \\ 1 & 1 & 2 \end{pmatrix}$$

gilt offensichtlich  $\text{Rang}(\mathbf{A}) = 3$  und damit  $\dim(\ker(\mathbf{A})) = 0$ .

**Bemerkung** Sei  $m < n$  und  $\mathbf{A}$  eine beliebige  $(m, n)$ -Matrix. Dann besitzt die homogene Gleichung  $\mathbf{A} \cdot \mathbf{x} = \mathbf{0}$  nach dem Dimensionssatz mindestens  $n - m$  linear unabhängige Lösungen (aber vielleicht noch mehr). Die inhomogene Gleichung muss aber trotzdem keine Lösung besitzen.

**Beispiel** Für

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 1 & 1 \\ -1 & 0 & -1 & -1 \end{pmatrix}$$

gilt  $\text{Rang}(\mathbf{A}) = 1$  und damit  $\dim(\ker(\mathbf{A})) = 3$ . Insbesondere beschreibt

$$\ker(\mathbf{A}) = \text{span} \left\{ \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \end{pmatrix} \right\}$$

alle Lösungen  $\mathbf{x} \in \mathbb{R}^4$  der homogenen Gleichung  $\mathbf{A} \cdot \mathbf{x} = \mathbf{0} \in \mathbb{R}^2$ , aber es gibt für

$$\mathbf{b} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

keine Lösung der inhomogenen Gleichung  $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$ .

**Lemma (Bedeutung der Spalten einer Matrix)** Sei  $\mathbf{A} \in \mathbb{R}^{(m, n)}$  eine beliebige Matrix und seien  $\mathbf{a}^{(1)}, \dots, \mathbf{a}^{(n)}$  die entsprechenden Spaltenvektoren (Elemente des  $\mathbb{R}^m$ ). Dann gilt

$$\mathbf{A} \cdot \mathbf{x} = x_1 \cdot \mathbf{a}^{(1)} + \dots + x_n \cdot \mathbf{a}^{(n)}.$$

*Beweis:* Nachrechnen analog zum folgenden Beispiel.

**Beispiel**

$$\begin{aligned} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} &= \begin{pmatrix} a_{11}x_1 + a_{12}x_2 + a_{13}x_3 \\ a_{21}x_1 + a_{22}x_2 + a_{23}x_3 \end{pmatrix} \\ &= x_1 \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix} + x_3 \begin{pmatrix} a_{13} \\ a_{23} \end{pmatrix} \end{aligned}$$

**Definition** Der Spaltenrang von  $\mathbf{A} \in \mathbb{R}^{(m, n)}$  ist die Anzahl der linear unabhängigen Spaltenvektoren.

**Theorem (Es gibt nur einen Rang.)** Für jede (quadratische oder nicht-quadratische) Matrix sind Zeilen- und Spaltenrang gleich.

*Beweis* Sei  $l$  der Spaltenrang von  $\mathbf{A}$ . Dann können wir o.B.d.A. — d.h. „ohne Beschränkung der Allgemeinheit“ — annehmen, dass die  $n$  Spaltenvektoren  $\mathbf{a}^{(1)}, \dots, \mathbf{a}^{(l)} \in \mathbb{R}^m$  linear unabhängig sind und finden Koeffizienten  $\gamma_{ki} \in \mathbb{R}$  so dass

$$\mathbf{a}^{(k)} = \sum_{i=1}^l \gamma_{ki} \mathbf{a}^{(i)} \quad \text{für alle } k = l+1, \dots, n$$

Damit gilt (beachte den Wechsel des Laufindex von  $k$  zu  $i$  in der ersten Summe der zweiten Zeile)

$$\begin{aligned} \mathbf{A} \cdot \mathbf{x} &= \sum_{k=1}^n x_k \mathbf{a}^{(k)} = \sum_{k=1}^l x_k \mathbf{a}^{(k)} + \sum_{k=l+1}^n \sum_{i=1}^l x_k \gamma_{ki} \mathbf{a}^{(i)} \\ &= \sum_{i=1}^l x_i \mathbf{a}^{(i)} + \sum_{i=1}^l \left( \sum_{k=l+1}^n x_k \gamma_{ki} \right) \mathbf{a}^{(i)} \\ &= \sum_{i=1}^l \left( x_i + \sum_{k=l+1}^n x_k \gamma_{ki} \right) \mathbf{a}^{(i)} \end{aligned}$$

und wir erhalten die Implikation

$$\mathbf{0} = \mathbf{A} \cdot \mathbf{x} \quad \Longleftrightarrow \quad x_i = - \sum_{k=l+1}^n x_k \gamma_{ki} \quad \text{für alle } i = 1 \dots l.$$

Hieraus schließen wir, dass die Dimension von  $\ker(\mathbf{A})$  gerade  $n - l$  ist, denn für jeden Vektor  $\mathbf{x} \in \ker \mathbf{A}$  können die Komponenten  $x_{l+1}, \dots, x_n$  beliebig gewählt werden und legen dann die Werte von  $x_1, \dots, x_l$  eindeutig fest (siehe auch das nächste Beispiel). Die Behauptung folgt nun mit der Dimensionsformel.  $\square$

**Beispiel** Um die Argumente im Beweis zu illustrieren, betrachten wir

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 1 \\ -1 & -2 & -1 \\ 0 & 1 & 2 \end{pmatrix}.$$

Diese Matrix besitzt die Spaltenvektoren

$$\mathbf{a}^{(1)} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad \mathbf{a}^{(2)} = \begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix}, \quad \mathbf{a}^{(3)} = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix},$$

wobei die zwei Vektoren  $\mathbf{a}^{(1)}, \mathbf{a}^{(2)}$  linear unabhängig sind und außerdem

$$\mathbf{a}^{(3)} = \gamma_{31} \cdot \mathbf{a}^{(1)} + \gamma_{32} \cdot \mathbf{a}^{(2)} \quad \text{mit} \quad \gamma_{31} = -3, \quad \gamma_{32} = 2$$

gilt. Insbesondere besitzt  $\mathbf{A}$  den Spaltenrang 2. Andererseits hatten wir im Beweis gesehen, dass

$$\mathbf{A} \cdot \mathbf{x} = (x_1 + \gamma_{31}x_3) \cdot \mathbf{a}^{(1)} + (x_2 + \gamma_{32}x_3) \cdot \mathbf{a}^{(2)}$$



für jeden Vektor  $\mathbf{x} \in \mathbb{R}^3$  gilt und hieraus der Kern von  $\mathbf{A}$  abgelesen werden kann: Es gilt  $\mathbf{A} \cdot \mathbf{x} = \mathbf{0}$  genau dann, wenn

$$x_1 = -\gamma_{31}x_3, \quad x_2 = -\gamma_{32}x_3 \quad \text{bzw.} \quad \mathbf{x} = x_3 \begin{pmatrix} 3 \\ -2 \\ 1 \end{pmatrix},$$

und es folgt

$$\ker \mathbf{A} = \text{span} \left\{ \begin{pmatrix} 3 \\ -2 \\ 1 \end{pmatrix} \right\}.$$

Die Dimensionsformel besagt, dass der Zeilenrang von  $\mathbf{A}$  auch  $2 = 3 - \dim \ker \mathbf{A}$  ist, und dies ist offensichtlich auch die Anzahl der linear unabhängigen Zeilenvektoren von  $\mathbf{A}$ .

**Beispiel** Die Matrix

$$\begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \end{pmatrix}$$

besitzt offensichtlich den Zeilenrang 2 und damit nach dem Theorem auch den Spaltenrang 2. Insbesondere kann der Nullvektor in  $\mathbb{R}^2$  durch eine nichttriviale Linearkombination der drei Spaltenvektoren der Matrix dargestellt werden, nämlich als

$$\begin{pmatrix} 0 \\ 0 \end{pmatrix} = 1 \cdot \begin{pmatrix} 2 \\ 1 \end{pmatrix} - 2 \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix} + 3 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

aber je zwei der drei Spaltenvektoren sind aber linear unabhängig. Nach Dimensionsformel gilt auch  $\dim(\ker \mathbf{A}) = 3 - 2 = 1$  und mit direkte Rechnungen ergibt sich

$$\ker \mathbf{A} = \text{span} \left\{ \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix} \right\}.$$

## 4.4 Quadratische Matrizen

**Klarstellung** Alles gilt wieder auch für komplexe Matrizen.

### Invertierbare und Inverse Matrizen

**Definition** Eine quadratische Matrix  $\mathbf{A} \in \mathbb{R}^{(n,n)}$  mit  $\text{Rang}(\mathbf{A}) = n$  bzw.  $\text{Rang}(\mathbf{A}) < n$  heißt regulär bzw. singulär.

**Definition** Eine quadratische Matrix  $\mathbf{A} \in \mathbb{R}^{(n,n)}$  heißt invertierbar, falls es eine weitere quadratische Matrix  $\mathbf{A}^{-1} \in \mathbb{R}^{(n,n)}$  gibt, so dass

$$\mathbf{A} \cdot \mathbf{A}^{-1} = \mathbf{I}_n = \mathbf{A}^{-1} \cdot \mathbf{A},$$

wobei  $\mathbf{A}^{-1}$  die sogenannte Inverse Matrix von  $\mathbf{A}$  ist.

**Theorem (Invertierbarkeit gleich Regularität)** Eine Matrix  $\mathbf{A} \in \mathbb{R}^{(n,n)}$  ist genau dann invertierbar, wenn sie regulär ist.

*Beweis, Hinrichtung:* Wir zeigen, dass jede invertierbare Matrix  $\mathbf{A}$  auch regulär ist. Dazu betrachten wir  $\mathbf{x} \in \mathbb{R}^n$  mit  $\mathbf{A} \cdot \mathbf{x} = \mathbf{0}$  und bemerken

$$\mathbf{x} = \mathbf{I}_n \cdot \mathbf{x} = \mathbf{A}^{-1} \cdot \mathbf{A} \cdot \mathbf{x} = \mathbf{A}^{-1} \cdot \mathbf{0} = \mathbf{0}.$$

Insbesondere gilt also  $\ker(\mathbf{A}) = \{0\}$  bzw.  $\dim(\ker(\mathbf{A})) = 0$  und die Dimensionsformel liefert  $\text{Rang}(\mathbf{A}) = n$ .

*Beweis, Rückrichtung:* Wir wollen nun zeigen, dass jede reguläre Matrix  $\mathbf{A}$  auch invertierbar ist. Dazu bemerken wir, dass es für jedes  $i = 1, \dots, n$  genau einen Vektor  $\mathbf{x}^{(i)} \in \mathbb{R}^n$  gibt, so dass

$$\mathbf{A} \cdot \mathbf{x}^{(i)} = \mathbf{e}_i,$$

wobei  $\mathbf{e}_i$  gerade der  $i$ -te Einheitsvektor (bzw. der  $i$ -te kanonische Basisvektor) ist. Diese Vektoren sind wegen

$$\begin{aligned} \mathbf{0} = \sum_{i=1}^n \nu_i \cdot \mathbf{x}^{(i)} &\implies \mathbf{0} = \mathbf{A} \cdot \left( \sum_{i=1}^n \nu_i \cdot \mathbf{x}^{(i)} \right) = \sum_{i=1}^n \nu_i \cdot \mathbf{A} \cdot \mathbf{x}^{(i)} = \sum_{i=1}^n \nu_i \cdot \mathbf{e}_i \\ &\implies \nu_1 = \dots = \nu_n = 0 \end{aligned}$$

linear unabhängig. Wir bezeichnen mit  $\mathbf{X}$  die Matrix, deren Spaltenvektoren gerade die  $\mathbf{x}^{(i)}$  sind, und bemerken, dass nach Konstruktion

$$\mathbf{A} \cdot \mathbf{X} = \mathbf{I}_n$$

gilt (Nachrechnen analog zum vorherigen Lemma!). Außerdem haben wir soeben gezeigt, dass der Spaltenrang von  $\mathbf{X}$  gleich  $n$  ist, d.h. es gilt  $\text{Rang}(\mathbf{X}) = n$  und  $\mathbf{X}$  ist also regulär. Es bleibt zu zeigen, dass auch  $\mathbf{X} \cdot \mathbf{A} = \mathbf{I}_n$  gilt (das folgt wegen der Nicht-Kommutativität der Matrizenmultiplikation nicht unmittelbar aus  $\mathbf{A} \cdot \mathbf{X} = \mathbf{I}_n$ ). Wir können alle unsere Argumente mit  $\mathbf{X}$  statt  $\mathbf{A}$  wiederholen und erhalten eine reguläre Matrix  $\mathbf{Y}$  mit  $\mathbf{X} \cdot \mathbf{Y} = \mathbf{I}_n$ . Dann gilt

$$\begin{aligned} \mathbf{X} \cdot \mathbf{A} &= (\mathbf{X} \cdot \mathbf{A}) \cdot \mathbf{I}_n = (\mathbf{X} \cdot \mathbf{A}) \cdot (\mathbf{X} \cdot \mathbf{Y}) \\ &= \mathbf{X} \cdot (\mathbf{A} \cdot \mathbf{X}) \cdot \mathbf{Y} = \mathbf{X} \cdot \mathbf{I}_n \cdot \mathbf{Y} = \mathbf{X} \cdot \mathbf{Y} \\ &= \mathbf{I}_n \end{aligned}$$

und der Beweis ist beendet, da  $\mathbf{X}$  alle gewünschten Eigenschaften besitzt.  $\square$

### Bemerkung

1. Zu jeder regulären Matrix  $\mathbf{A}$  kann es nur eine Inverse geben, denn  $\mathbf{A} \cdot \mathbf{X} = \mathbf{I}_n$  impliziert, dass der  $i$ -te Spaltenvektor  $\mathbf{x}^{(i)}$  von  $\mathbf{X}$  gerade die eindeutige Lösung der inhomogenen Gleichung  $\mathbf{A} \cdot \mathbf{x}^{(i)} = \mathbf{e}_i$  ist.
2. Für jedes reguläre  $\mathbf{A}$  kann  $\mathbf{A}^{-1}$  durch das Gauß-Jordan-Verfahren berechnet werden (siehe oben).
3. Ist  $\mathbf{A}$  regulär, so ist für jedes  $\mathbf{b} \in \mathbb{R}^n$  die eindeutige Lösung der Gleichung  $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$  durch  $\mathbf{x} = \mathbf{A}^{-1} \cdot \mathbf{b}$  gegeben.
4. Nicht-quadratische Matrizen können niemals invertierbar sein. Es gibt aber das allgemeinere Konzept von *Pseudo-Inversen*, siehe WIKIPEDIA.

**Inverse von (2, 2)-Matrizen** Für eine (2, 2)-Matrix

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \quad \text{gilt} \quad \mathbf{A}^{-1} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix},$$

sofern  $\det \mathbf{A} = a_{11}a_{22} - a_{12}a_{21} \neq 0$ .

**Regeln für Inverse Matrizen**

1.  $\mathbf{A} \cdot \mathbf{B} = \mathbf{I}$  impliziert schon (siehe zum Beispiel die Determinantenformel weiter unten), dass  $\mathbf{A}$  und  $\mathbf{B}$  beide regulär und damit auch invertierbar sind. Insbesondere gilt dann schon  $\mathbf{A} = \mathbf{B}^{-1}$  und  $\mathbf{B} = \mathbf{A}^{-1}$  und  $\mathbf{B} \cdot \mathbf{A} = \mathbf{I}$ .
2. Sind  $\mathbf{A}$  und  $\mathbf{B}$  beide invertierbar, so ist auch  $\mathbf{A} \cdot \mathbf{B}$  invertierbar und es gilt

$$(\mathbf{A} \cdot \mathbf{B})^{-1} = \mathbf{A}^{-1} \cdot \mathbf{B}^{-1}$$

3. Es gilt  $(\mathbf{A}^{-1})^{-1} = \mathbf{A}$ .

Vorlesung 13, 18. November 2019

## Exkurs über Permutationen

**Definition** Eine  $n$ -stellige Permutation  $\sigma$  ist eine *bijektive Abbildung* der Menge  $\{1, 2, \dots, n\}$  in sich. Wir schreiben

$$\sigma : i \in \{1, \dots, n\} \mapsto \sigma_i \in \{1, \dots, n\} \quad \text{bzw.} \quad \sigma = (\sigma_1, \dots, \sigma_n)$$

Die Menge aller  $n$ -stelligen Permutationen wollen wir mit  $S_n$  bezeichnen und bemerken, dass  $\#S_n = n!$ .

### Beispiele

$$S_2 = \{(1, 2), (2, 1)\}$$

$$S_3 = \{(1, 2, 3), (1, 3, 2), (2, 1, 3), (2, 3, 1), (3, 1, 2), (3, 2, 1)\}$$

**Operationen mit Permutationen** Permutationen können — da sie Abbildungen sind — in natürlicher Weise komponiert werden. So gilt zum Beispiel

$$(1, 3, 2) \circ (2, 1, 3) = (3, 1, 2),$$

$$(2, 1, 3) \circ (1, 3, 2) = (2, 3, 1),$$

$$(1, 3, 2) \circ (1, 3, 2) = (1, 2, 3),$$

sowie

$$(2, 1, 4, 3) \circ (3, 4, 1, 2) = (4, 3, 2, 1).$$

Da Permutationen aber auch bijektive Abbildungen sind, können sie sogar invertiert werden. So gilt zum Beispiel

$$\sigma^{-1} = (1, 3, 2) \quad \text{für} \quad \sigma = (1, 3, 2) \quad \sigma^{-1} = (3, 1, 2) \quad \text{für} \quad \sigma = (2, 3, 1)$$

sowie

$$\sigma^{-1} = (4, 1, 3, 2) \quad \text{für} \quad \sigma = (2, 4, 3, 1)$$

**\*Bemerkung** Die Menge  $S_n$  bildet mit der Komposition eine nicht-kommutative Gruppe, wobei das neutrale Element gerade die Identitätsabbildung  $(1, 2, \dots, n-1, n)$  (bzw. die triviale Permutation) ist.

**Einfache Permutationen** Transpositionen sind Permutationen, bei denen nur zwei Elemente vertauscht werden, zum Beispiel

$$(1, 3, 2), \quad (4, 2, 3, 1), \quad (2, 1, 3, 4)$$

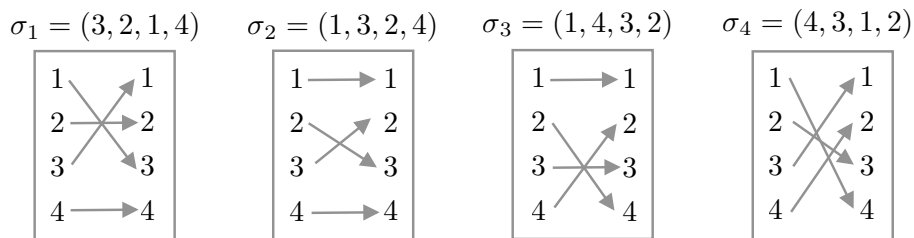
sowie

$$(1, 6, 3, 4, 5, 2), \quad (1, 2, 4, 3, 5, 6).$$

Insbesondere ist jede Transposition zu sich selbst invers und wir benutzen für die  $n$ -stellige Transposition, die die Zahlen  $i$  und  $j$  miteinander vertauscht, auch die Kurzschreibweise  $[i, j]$ .

**Theorem (Zerlegungssatz)** Jede Permutation kann als Komposition endlich vieler Transpositionen geschrieben werden. Diese Zerlegung ist zwar im Allgemeinen nicht eindeutig, aber man benötigt für eine gegebene Permutation entweder immer eine gerade oder immer eine ungerade Anzahl von Transpositionen.

*Beweis:* siehe [AORS, Seite 98].



$$\sigma_4 = \sigma_3 \circ \sigma_2 \circ \sigma_1$$

Beispiele für 4-stellige Permutationen.

### Beispiele

$$\begin{aligned} (3, 2, 1) &= [1, 2] \circ [2, 3] \circ [1, 2] = (2, 1, 3) \circ (1, 3, 2) \circ (2, 1, 3) \\ &= [2, 3] \circ [1, 2] \circ [2, 3] = (1, 3, 2) \circ (2, 1, 3) \circ (1, 3, 2) \end{aligned}$$

$$(4, 3, 1, 2) = [2, 3] \circ [1, 3] \circ [1, 4] = (1, 3, 2, 4) \circ (3, 2, 1, 4) \circ (4, 2, 3, 1)$$

**Definition** Das Vorzeichen  $\text{sgn}(\sigma)$  einer Permutation  $\sigma$  ist  $+1$  bzw.  $-1$ , sofern  $\sigma$  in eine gerade bzw. eine ungerade Anzahl von Transpositionen zerlegt werden kann.

### Beispiele

$$\text{sgn}((1, 2)) = +1, \quad \text{sgn}((2, 1)) = -1$$

$$\text{sgn}((1, 2, 3)) = +1$$

$$\text{sgn}((1, 3, 2)) = \text{sgn}((3, 2, 1)) = \text{sgn}((2, 1, 3)) = -1$$

$$\text{sgn}((2, 3, 1)) = \text{sgn}((3, 1, 2)) = +1,$$

**Bemerkung** Da Permutationen bijektive Abbildungen sind und weil  $[i, j]^{-1} = [i, j]$  gilt, verifiziert man leicht die logische Äquivalenz

$$\sigma = [i_k, j_k] \circ \dots \circ [i_1, j_1] \iff \sigma^{-1} = [i_1, j_1] \circ \dots \circ [i_k, j_k].$$

Insbesondere gilt  $\text{sgn}(\sigma) = \text{sgn}(\sigma^{-1})$  für jede Permutation  $\sigma$ .

## Determinanten von Matrizen

**Motivation** Wir führen nun eine zunächst sehr abstrakt wirkende Größe ein, die aber eine sehr bedeutende Rolle in der Theorie und den Anwendungen von Matrizen spielt.

**Definition** Die Determinante einer quadratischen Matrix  $\mathbf{A} \in \mathbb{R}^{(n, n)}$  ist

$$\det \mathbf{A} := \begin{vmatrix} a_{11} & \dots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \dots & a_{nn} \end{vmatrix} := \sum_{\sigma \in S_n} \text{sgn}(\sigma) \cdot a_{1\sigma_1} \cdot \dots \cdot a_{n\sigma_n}$$

und damit eine alternierende Summe von  $n!$  Summanden, die jeweils das Produkt von  $n$  verschiedenen Matrixeinträgen sind. Beachte, dass jeder Spalten- und jeder Zeilenindex in jedem Summanden genau einmal vorkommt.

**Beispiel** Für  $n = 2$  gilt

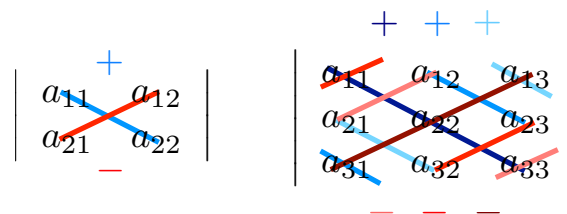
$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11} \cdot a_{22} - a_{12} \cdot a_{21}$$

und  $\det \mathbf{A}$  ist gerade der vorzeichenbehaftete Flächeninhalt des durch die zwei Spaltenvektoren von  $\mathbf{A}$  aufgespannten Parallelogramms.

Für  $n = 3$  ergibt sich

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{matrix} + & a_{11} \cdot a_{22} \cdot a_{33} & - & a_{11} \cdot a_{23} \cdot a_{32} \\ + & a_{12} \cdot a_{23} \cdot a_{31} & - & a_{12} \cdot a_{21} \cdot a_{33} \\ + & a_{13} \cdot a_{21} \cdot a_{32} & - & a_{13} \cdot a_{22} \cdot a_{31} \end{matrix},$$

und wir können  $\det A$  als das vorzeichenbehaftete Volumen des Spates interpretieren, der aus den drei Spaltenvektoren von  $\mathbf{A}$  gebildet werden kann. Beachte auch die Sarrussche Regel für  $n = 2$  und  $n = 3$  (siehe Bild).



Die schematische Sarrus-Regel zur Berechnung der Determinante von (2, 2)- bzw. (3, 3)-Matrizen. Es gibt kein direktes Analogon für  $(n, n)$ -Matrizen mit  $n \geq 4$ .

**Rechenregeln**

$$\det \mathbf{A} = \det \mathbf{A}^T$$

$$\det (\mathbf{A} \cdot \tilde{\mathbf{A}}) = \det \mathbf{A} \cdot \det \tilde{\mathbf{A}}$$

Beweis: Es gilt

$$\det \mathbf{A}^T = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) \cdot a_{\sigma_1 1} \cdot \dots \cdot a_{\sigma_n n},$$

aber durch Umsortierung der Summanden und nach Substitution  $\tau = \sigma^{-1}$  kann man diese Formel auch als

$$\det \mathbf{A}^T = \sum_{\tau \in S_n} \operatorname{sgn}(\tau^{-1}) \cdot a_{1\tau_1} \cdot \dots \cdot a_{n\tau_n}$$

schreiben. Die erste Behauptung folgt nun wegen  $\operatorname{sgn}(\tau) = \operatorname{sgn}(\tau^{-1})$ . Die zweite Behauptung kann ebenfalls direkt nachgerechnet werden, aber die Rechnungen sind sehr lang und mühsam.

**Entwicklungssatz von Laplace** Es gilt

$$\det \mathbf{A} = \sum_{k=1}^n (-1)^{j+k} a_{jk} \det \mathbf{A}_{jk} \quad (\text{Entwicklung nach der } j\text{-ten Zeile})$$

sowie

$$\det \mathbf{A} = \sum_{k=1}^n (-1)^{j+k} a_{kj} \det \mathbf{A}_{kj} \quad (\text{Entwicklung nach der } j\text{-ten Spalte}),$$

wobei  $\mathbf{A}_{ij}$  die reduzierte  $(n-1, n-1)$ -Matrix bezeichnet, die aus  $\mathbf{A}$  durch Löschung der  $i$ -ten Zeile und der  $j$ -ten Spalte entsteht. Beide Behauptungen können mit Hilfe der Definition einfach nachgerechnet werden, aber das erfordert wieder etwas Geschick und Geduld im Umgang mit Permutationen. Die Details können zum Beispiel in [AORS, Seite 105] nachgeschlagen werden.

$$\mathbf{A}_{11} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \quad \mathbf{A}_{21} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \quad \mathbf{A}_{13} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

Beispiele für reduzierte Matrizen wie im Entwicklungssatz für  $n = 3$ . Die Streichungen sind grau markiert.

**Beispiel** Entwicklung nach der ersten Zeile

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = +a_{11} \cdot \det \begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} - a_{12} \cdot \det \begin{pmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{pmatrix} \\ + a_{13} \cdot \det \begin{pmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix}$$

**Beispiel** Für

$$\mathbf{A} = \begin{pmatrix} 1 & 3 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{pmatrix}$$

entwickeln wir nach der vierten Spalte (dort gibt es viele Nullen)

$$\det(\mathbf{A}) = (-1)^{4+4} \cdot 1 \cdot \det \begin{pmatrix} 1 & 3 & 0 \\ 0 & 0 & 2 \\ -1 & 0 & 1 \end{pmatrix},$$

wobei der Vorfaktor gerade 1 ist. Anschließend können wir nach der zweiten Zeile entwickeln und erhalten

$$\det(\mathbf{A}) = (-1)^{2+3} \cdot 2 \cdot \det \begin{pmatrix} 1 & 3 \\ -1 & 0 \end{pmatrix}.$$

Im letzten Schritt können wir verbleibende (2, 2) mit der Sarrusschen Regel berechnen:

$$\det(\mathbf{A}) = -2 \cdot (1 \cdot 0 - 3 \cdot (-1)) = -6.$$

**Determinante von Dreiecksmatrizen** Man kann — zum Beispiel direkt oder mit dem Entwicklungssatz oder durch vollständige Induktion über  $n$  — die Formel

$$\det \begin{pmatrix} a_{11} & \dots & a_{1n} \\ & \ddots & \vdots \\ 0 & & a_{nn} \end{pmatrix} = a_{11} \cdot \dots \cdot a_{nn}$$

ableiten (siehe Hausaufgabe), wobei bei einer oberen Dreiecksmatrix  $a_{ij} = 0$  für alle  $1 \leq j < i \leq n$  gilt.

**Spalten- und Zeilenweise Linearität der Determinante** Identifizieren wir jede quadratische Matrix  $\mathbf{A}$  via

$$\mathbf{A} = \left( \begin{array}{c|c|c} | & & | \\ \mathbf{a}^{(1)} & \dots & \mathbf{a}^{(n)} \\ | & & | \end{array} \right)$$

mit dem entsprechenden  $n$ -Tupel aus  $n$ -dimensionalen Spaltenvektoren  $\mathbf{a}^{(i)}$ , so gilt

$$\begin{aligned} \det \left( \begin{array}{c|c|c|c|c|c} | & & | & & | & \\ \mathbf{a}^{(1)} & \dots & \mathbf{a}^{(k-1)} & \check{\lambda} \cdot \check{\mathbf{a}}^{(k)} + \tilde{\lambda} \cdot \tilde{\mathbf{a}}^{(k)} & \mathbf{a}^{(k+1)} & \dots & \mathbf{a}^{(n)} \\ | & & | & & | & & | \end{array} \right) = \\ \check{\lambda} \cdot \det \left( \begin{array}{c|c|c|c|c|c} | & & | & & | & \\ \mathbf{a}^{(1)} & \dots & \mathbf{a}^{(k-1)} & \check{\mathbf{a}}^{(k)} & \mathbf{a}^{(k+1)} & \dots & \mathbf{a}^{(n)} \\ | & & | & & | & & | \end{array} \right) + \\ \tilde{\lambda} \cdot \det \left( \begin{array}{c|c|c|c|c|c} | & & | & & | & \\ \mathbf{a}^{(1)} & \dots & \mathbf{a}^{(k-1)} & \tilde{\mathbf{a}}^{(k)} & \mathbf{a}^{(k+1)} & \dots & \mathbf{a}^{(n)} \\ | & & | & & | & & | \end{array} \right) \end{aligned}$$

für jedes  $k = 1, \dots, n$  sowie eine analoge Formel für jede Zeile der Matrix  $\mathbf{A}$ . Auch dies kann direkt mit Hilfe der Definition nachgerechnet werden, siehe zum Beispiel [AORS, Seite 102].

**Beispiel** Die Linearität bzgl. der dritten Spalte impliziert zum Beispiel

$$\det \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = a_{13} \cdot \det \begin{pmatrix} a_{11} & a_{12} & 1 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & 0 \end{pmatrix} + a_{23} \cdot \det \begin{pmatrix} a_{11} & a_{12} & 0 \\ a_{21} & a_{22} & 1 \\ a_{31} & a_{32} & 0 \end{pmatrix} + a_{33} \cdot \det \begin{pmatrix} a_{11} & a_{12} & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & 1 \end{pmatrix}.$$

Für  $(2, 2)$ -Matrizen ergibt sich aus der Linearität bzgl. der ersten und der zweiten Spalte die Formel

$$\begin{aligned} \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} &= a_{11} \cdot \det \begin{pmatrix} 1 & a_{12} \\ 0 & a_{22} \end{pmatrix} + a_{21} \cdot \det \begin{pmatrix} 0 & a_{12} \\ 1 & a_{22} \end{pmatrix} \\ &= a_{11} \cdot a_{12} \cdot \det \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} + a_{11} \cdot a_{22} \cdot \det \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \\ &\quad a_{21} \cdot a_{12} \cdot \det \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + a_{21} \cdot a_{22} \cdot \det \begin{pmatrix} 0 & 0 \\ 1 & 1 \end{pmatrix} \\ &= a_{11} \cdot a_{12} \cdot 0 + a_{11} \cdot a_{22} \cdot (+1) + a_{21} \cdot a_{12} \cdot (-1) + a_{21} \cdot a_{22} \cdot 0 \\ &= a_{11} \cdot a_{22} - a_{12} \cdot a_{21}, \end{aligned}$$

die wir aber schon kennen.

**Achtung** Es gilt im Allgemeinen  $\det(\lambda \cdot \mathbf{A} + \tilde{\lambda} \cdot \tilde{\mathbf{A}}) \neq \lambda \cdot \det \mathbf{A} + \tilde{\lambda} \cdot \det \tilde{\mathbf{A}}$ , d.h. die Determinante ist **keine** lineare Abbildung auf dem Raum der Matrizen. Es ist eine sogenannte *multi-lineare* Abbildung. Außerdem gilt

$$\det(\lambda \cdot \mathbf{A}) = \lambda^n \cdot \det \mathbf{A}$$

für jede  $(n, n)$ -Matrix  $\mathbf{A}$  und jeden Skalar  $\lambda$ .

**\*Bemerkung** Die Determinante ist im Wesentlichen — d.h. bis auf einen universellen Vorfaktor — die einzige multi-lineare Abbildung  $\mathbb{R}^{(n,n)} \rightarrow \mathbb{R}$ .

**Beispiel** Es gilt zwar

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} = \lambda \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

aber auch

$$\begin{aligned} 1 &= \det \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \neq \det \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + \det \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = 0 + 0, \\ \lambda^2 &= \det \begin{pmatrix} \lambda & 0 \\ 0 & \lambda \end{pmatrix} \neq \lambda \det \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \lambda. \end{aligned}$$

**Folgerung (Determinante und Rang-Defizit)** Besitzt  $\mathbf{A}$  linear abhängige Spalten oder linear abhängige Zeilen, so gilt  $\det \mathbf{A} = 0$ .



**Beispiel** Es gilt

$$\det \begin{pmatrix} -1 & 0 & 2 \\ 0 & 3 & 0 \\ 1 & 0 & -2 \end{pmatrix} = 0.$$

Man kann dies entweder direkt nachrechnen oder mit der linearen Abhängigkeit der ersten und dritten Spalte begründen.

**Determinante und Permutationen von Zeilen oder Spalten** Durch Rechnungen mit Permutationen kann man ebenfalls zeigen (siehe etwa [AORS, Seite 101]), dass

$$\det \begin{pmatrix} | & & | \\ \mathbf{a}^{(\sigma_1)} & \dots & \mathbf{a}^{(\sigma_n)} \\ | & & | \end{pmatrix} = \operatorname{sgn}(\sigma) \cdot \det \begin{pmatrix} | & & | \\ \mathbf{a}^{(1)} & \dots & \mathbf{a}^{(n)} \\ | & & | \end{pmatrix}$$

für jede Permutation  $\sigma$  der Spalten und analog auch für jede Permutation der Zeilen gilt. Insbesondere ändert jede Vertauschung von zwei Spalten bzw. jede Vertauschung von zwei Zeilen das Vorzeichen der Determinante.

**Beispiel**

$$\det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = -\det \begin{pmatrix} a_{12} & a_{11} \\ a_{22} & a_{21} \end{pmatrix}, \quad \det \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = -\det \begin{pmatrix} a_{21} & a_{22} \\ a_{11} & a_{12} \end{pmatrix}$$

**Folgerung (Determinante und Gauß-Verfahren)** Bei der Gauß-Elimination und dem Gauß-Jordan-Verfahren gilt für die Determinante der Koeffizientenmatrix:

1. Sie ändert sich nicht, wenn das Vielfache einer Zeile zu einer anderen Zeile addiert wird.
2. Sie ändert sich jedesmal das Vorzeichen, wenn zwei Zeilen vertauscht werden.
3. Wird eine Zeile mit einer Zahl multipliziert, so wird sie auch mit dieser Zahl multipliziert.

Insbesondere kann in jedem Schritt die Änderung die Determinante leicht protokolliert werden. Da außerdem die Determinante der resultierenden Dreiecksmatrix leicht berechnet werden kann (siehe oben), können Determinanten auch durch Gauß-Verfahren berechnet werden.

**Bemerkung**

1. Bei großen Matrizen ist die Berechnung der Determinante mittels der Gauß-Elimination und der Diagonalförmel effektiver und deutlich schneller als die Verwendung der Definition.
2. Gibt es jedoch (sehr) viele Nulleinträge, so ist die sukzessive Anwendung der Laplaceschen Entwicklungsformel am besten geeignet.

**Folgerung (Hauptsatz über quadratische Matrizen)** Für jede quadratische Matrix  $\mathbf{A}$  sind die folgenden Aussagen paarweise äquivalent:

1.  $\det(\mathbf{A}) \neq 0$ .
2.  $\mathbf{A}$  ist invertierbar.
3.  $\mathbf{A}$  ist regulär.
4.  $\dim(\ker(\mathbf{A})) = 0$ .
5. Die  $n$  Zeilenvektoren von  $\mathbf{A}$  sind linear unabhängig.
6. Die  $n$  Spaltenvektoren von  $\mathbf{A}$  sind linear unabhängig.
7. Die Gleichung  $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$  besitzt für jedes  $\mathbf{b} \in \mathbb{R}^n$  genau eine Lösung  $\mathbf{x} \in \mathbb{R}^n$ .

**\*Kramersche Regel** Man kann viele der bisher vorgestellten Resultate über Matrizen mit der Hilfe von Determinanten ausdrücken. Zum Beispiel kann man für alle regulären Matrizen  $\mathbf{A}$  und jeden Vektor  $\mathbf{b} \in \mathbb{R}^n$  zeigen, dass die  $k$ -te Komponente  $x_k$  der eindeutigen Lösung zur inhomogenen Gleichung  $\mathbf{A} \cdot \mathbf{x} = \mathbf{b}$  auch als

$$x_k = \frac{\det \begin{pmatrix} | & & | & & | \\ \mathbf{a}^{(1)} & \dots & \mathbf{b} & \dots & \mathbf{a}^{(n)} \\ | & & | & & | \end{pmatrix}}{\det \begin{pmatrix} | & & | & & | \\ \mathbf{a}^{(1)} & \dots & \mathbf{a}^{(k)} & \dots & \mathbf{a}^{(n)} \\ | & & | & & | \end{pmatrix}}$$

geschrieben werden kann, wobei im Nenner gerade  $\det \mathbf{A}$  steht und im Zähler die Determinante der Matrix, die aus  $\mathbf{A}$  dadurch entsteht, dass man den  $k$ -ten Spaltenvektor durch den Vektor  $\mathbf{b}$  ersetzt.

**Beispiel** Für

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}, \quad \mathbf{b} = \begin{pmatrix} 5 \\ 6 \end{pmatrix}$$

ist die Lösung der inhomogenen Gleichung gerade

$$\mathbf{x} = \frac{1}{2} \begin{pmatrix} -8 \\ 9 \end{pmatrix}.$$

Man kann dies mit Hilfe des Gauß-Jordan-Verfahrens

$$\left( \begin{array}{cc|c} 1 & 2 & 5 \\ 3 & 4 & 6 \end{array} \right) \cong \left( \begin{array}{cc|c} 1 & 2 & 5 \\ 0 & -2 & -9 \end{array} \right) \cong \left( \begin{array}{cc|c} 1 & 0 & -4 \\ 0 & -2 & -9 \end{array} \right) \cong \left( \begin{array}{cc|c} 1 & 0 & -4 \\ 0 & 1 & 9/2 \end{array} \right)$$

oder mit Hilfe der Inversen von  $\mathbf{A}$  via

$$\mathbf{x} = \mathbf{A}^{-1} \cdot \mathbf{b} = \frac{1}{-2} \begin{pmatrix} 4 & -2 \\ -3 & 1 \end{pmatrix} \cdot \begin{pmatrix} 5 \\ 6 \end{pmatrix}$$

zeigen. Alternativ kann man aber auch Kramers Formel benutzen:

$$x_1 = \frac{\det \begin{pmatrix} 5 & 2 \\ 6 & 4 \end{pmatrix}}{\det \mathbf{A}} = \frac{8}{-2} = -4 \quad x_2 = \frac{\det \begin{pmatrix} 1 & 5 \\ 3 & 6 \end{pmatrix}}{\det \mathbf{A}} = \frac{-9}{-2} = \frac{9}{2}$$

**\*Adjunkte einer Matrix und abstrakte Inversionsformel** Die Matrix

$$\tilde{\mathbf{A}} = \begin{pmatrix} \tilde{a}_{11} & \dots & \tilde{a}_{1n} \\ \vdots & & \vdots \\ \tilde{a}_{n1} & & \tilde{a}_{nn} \end{pmatrix} \quad \text{mit} \quad \tilde{a}_{ij} := (-1)^{i+j} \det \mathbf{A}_{ji}$$

wird die zu  $\mathbf{A}$  komplementäre Matrix genannt, wobei jeder Eintrag  $\tilde{a}_{ij}$  ein sogenanntes Adjunkt (oder ein Kofaktor) von  $\mathbf{A}$  ist und mittels der Determinante einer reduzierten Matrix definiert ist (beachte die Indexvertauschung:  $\tilde{a}_{ij}$  wird aus  $\mathbf{A}_{ji}$  berechnet). Aus der Kramers-Formel folgt

$$\mathbf{A} \cdot \tilde{\mathbf{A}} = \tilde{\mathbf{A}} \cdot \mathbf{A} = \det \mathbf{A} \cdot \mathbf{I}_n$$

bzw.

$$\mathbf{A}^{-1} = \frac{1}{\det \mathbf{A}} \tilde{\mathbf{A}} \quad \text{sofern} \quad \det \mathbf{A} \neq 0.$$

Allerdings liefert dies für große  $n$  keinen effizienten Algorithmus zur Berechnung von Inversen.

**Beispiel** Für eine invertierbare  $(2, 2)$ -Matrix ergibt sich

$$\tilde{a}_{11} = a_{22}, \quad \tilde{a}_{12} = -a_{12}, \quad \tilde{a}_{21} = -a_{21}, \quad \tilde{a}_{22} = a_{11}$$

und wir erhalten wieder die schon bekannte Formel für  $\mathbf{A}^{-1}$ . Hierbei haben wir benutzt, dass  $\det(a_{11}) = a_{11}$  für  $(1, 1)$ -Matrizen gilt.

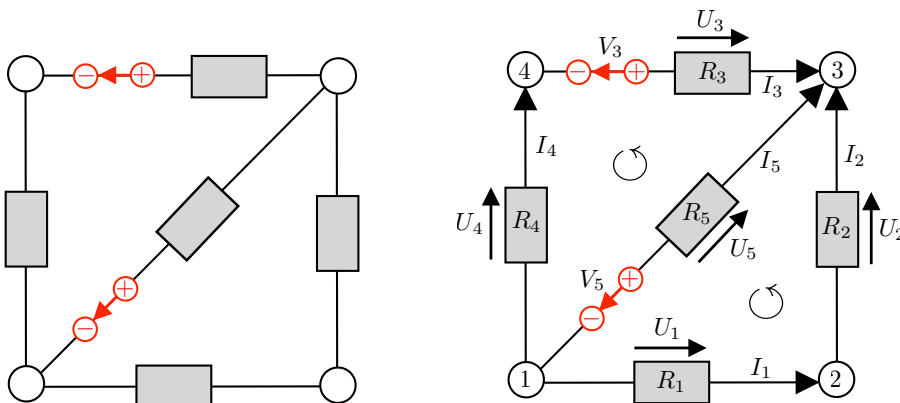
## 4.5 \*Elektrische Netzwerke und Matrizen

In diesem Abschnitt beschreiben wir drei unterschiedliche Verfahren, um die Ströme in einem gegebenen elektrischen Netzwerk mittels linearer Gleichungssysteme bzw. mit Hilfe geeigneter Matrizen berechnen zu können. Um die wesentlichen Ideen herauszuarbeiten, betrachten wir zunächst immer das im linken Bild dargestellte Beispiel für die folgende Netzwerkklasse:

1. Das Netzwerk (bzw. der *Graph*) besteht aus  $n$  Knoten,  $m$  Kanten und insgesamt  $l$  Maschen (graphentheoretische „Kreise“).
2. Auf jeder Kante gibt es einen gegebenen Widerstand und ggf. eine externe Stromquelle mit einer bekannten Spannung.

Als Vorbereitung für unsere Untersuchungen

1. nummerieren wir die Knoten, Kanten und Maschen jeweils beliebig,
2. wählen wir entlang jeder Kante eine beliebige Orientierung (also einen Richtungspfeil), wobei wir o.B.d.A. vereinbaren wollen, dass auf jeder Kante mit Stromquelle der Kantenpfeil entgegengesetzt zum Spannungspfeil der Stromquelle festgelegt wird,
3. wählen wir für jede Masche eine Orientierung, d.h. einen Umlaufsinn (bei planaren Netzwerken können wir zum Beispiel immer gegen den Uhrzeigersinn laufen).



Erstes Beispiel für ein elektrisches Netzwerk. Es gilt  $n = 4$ ,  $m = 5$  und  $l = 3$  mit den drei Maschen  $1 - 2 - 3 - 1$ ,  $1 - 3 - 4 - 1$  sowie  $1 - 2 - 3 - 4 - 1$ .

Sobald geeignete Nummerierungen und Kantenorientierungen gewählt sind, können wir für jede Kante  $i \in \{1, \dots, m\}$  die folgenden Größen betrachten:

1. einen Widerstandswert  $R_i$ , wobei wir immer  $0 < R_i < \infty$  voraussetzen wollen,
2. die Spannung  $V_i$  einer Stromquelle, wobei wir ggf.  $V_i = 0$  setzen,
3. einen Spannungsabfall  $U_i$  am Widerstand  $R_i$ ,
4. einen Strom  $I_i$  durch die Kante (auch *Zweigstrom* genannt).

In der Regel sind die Werte  $R_i$  und  $V_i$  bekannt, aber wir müssen die Ströme  $I_i$  sowie die Spannungen  $U_i$  bestimmen. Die erhaltenen Kantenströme werden am Ende nicht von der Wahl der Kanten- und Maschenorientierungen abhängen, da wir für jeden Kantenstrom entweder einen *positiven* bzw. einen *negativen* Wert erhalten werden, je nachdem ob der Strom *mit* bzw. *gegen* den willkürlich gewählten Kantenpfeil fließt.

## Zweigstrom-Analyse — Herleitung

Wir diskutieren zunächst die grundlegenden physikalischen Prinzipien und identifizieren ein vollständiges Gleichungssystem für die Kantenströme  $I_i$ .

### $n$ Kirchhoffsche Knotenregeln (Ladungserhaltung)

*In jedem Knoten des Netzwerkes ist die Summe der einlaufenden Ströme gleich der Summe der auslaufenden Ströme.*

Dieses Gesetz kann in jedem Knoten in eine Gleichung für die Ströme übersetzt werden, wobei die ein- bzw. auslaufenden Ströme mit  $+1$  bzw.  $-1$  gewichtet werden. Für das konkrete Netzwerkbeispiel erhalten wir:

$$\begin{array}{rcl} - I_1 & - I_4 - I_5 & = 0 & \text{(Knoten 1)} \\ + I_1 - I_2 & & = 0 & \text{(Knoten 2)} \\ & + I_2 + I_3 & + I_5 & = 0 & \text{(Knoten 3)} \\ & - I_3 + I_4 & & = 0 & \text{(Knoten 4)} \end{array}$$

Beachte, dass diese Gleichungen linear abhängig sind, da die Summe aller Gleichungen aus graphentheoretischen Erwägungen (jede Kante hat genau einen Anfangs- und genau einen Endknoten und erscheint daher genau zweimal mit verschiedenen Vorzeichen) die Tautologie  $0 = 0$  ergeben muss; je  $n - 1$  Knotengleichungen sollten aber linear unabhängig sein (sonst kann das Netzwerk nicht zusammenhängend sein).

### $l$ Kirchhoffsche Maschenregeln (Energieerhaltung)

*In jeder Masche des Netzwerkes heben sich alle vorzeichenbehafteten Teilspannungen insgesamt auf.*

Dieses Gesetz liefert in jeder Masche eine lineare Gleichung für die Spannungen, wobei alle Kanten der Masche gemäß der Maschenorientierung durchlaufen werden und jede richtig oder falsch orientierte Teilspannung mit dem Vorfaktor  $+1$  bzw.  $-1$  berücksichtigt wird. Im konkreten Netzwerk erhalten wir:

$$\begin{array}{rcl} + U_1 + U_2 & - U_5 + V_5 & = 0 & \text{(1. Masche: 1-2-3-1)} \\ & - U_3 + V_3 - U_4 + U_5 - V_5 & = 0 & \text{(2. Masche: 1-3-4-1)} \\ + U_1 + U_2 - U_3 + V_3 - U_4 & & = 0 & \text{(3. Masche: 1-2-3-4-1)} \end{array}$$

Auch hier sind die Gleichungen nicht unabhängig voneinander (siehe dazu die Bemerkung unten).

**$m$  Ohmsche Widerstände (Materialgesetze)** Schließlich wissen wir, dass es an jedem Widerstand einen einfachen Zusammenhang zwischen der Spannung und dem Strom gibt, nämlich

$$U_i = R_i \cdot I_i \quad \text{für jede Kante } i \in \{1, \dots, m\}.$$

**Gleichungssystem für die Ströme** Ersetzen wir die Spannung in den Maschenregeln mit Hilfe der Ohmschen Gesetze durch die Ströme  $I_i$ , so erhalten wir insgesamt  $n+l$  Gleichungen für die  $m$  unbekanntenen Ströme  $I_1, \dots, I_m$ . Die entsprechende erweiterte Koeffizientenmatrix ergibt sich im konkreten Beispiel zu

$$\left( \begin{array}{ccccc|c} -1 & 0 & 0 & -1 & -1 & 0 \\ 1 & -1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & -1 & 1 & 0 & 0 \\ R_1 & R_2 & 0 & 0 & -R_5 & -V_5 \\ 0 & 0 & -R_3 & -R_4 & R_5 & -V_3 + V_5 \\ R_1 & R_2 & -R_3 & -R_4 & 0 & -V_3 \end{array} \right)$$

und enthält nur die (bekannten) Widerstände  $R_i$  und die (bekannten) Beiträge  $V_i$  der Stromquellen, wobei wir letztere auf die rechte Seite der Gleichungen verschoben haben.

Wir können nun dieses Gleichungssystem für konkrete Werte der Parameter  $R_i$  und  $V_i$  lösen, wobei die Physik sicherstellt, dass es immer eine eindeutige Lösung für die Kantenströme  $I_i$  geben wird (oder wir uns irgendwo verrechnet haben). Anschließend können wir die Spannungen  $U_i$  mittels des Ohmschen Gesetzes aus den  $I_i$  direkt berechnen.

Für das konkrete Netzwerk ergibt sich zum Beispiel mit den Zahlenwerten

$$R_1 = R_2 = R_3 = R_4 = R_5 = 1$$

und

$$V_3 = 4, \quad V_5 = 2 \quad \text{bzw.} \quad V_3 = 8, \quad V_5 = 8$$

der Stromvektor

$$\mathbf{I} = \begin{pmatrix} -1 \\ -1 \\ 1 \\ 1 \\ 0 \end{pmatrix} \quad \text{bzw.} \quad \mathbf{I} = \begin{pmatrix} -3 \\ -3 \\ 1 \\ 1 \\ 2 \end{pmatrix}.$$

Die Negativität von  $I_1$  und  $I_2$  bedeutet jeweils, dass der Strom entgegen des von uns willkürlich gewählten Pfeiles fließt, wohingegen die Positivität von  $I_3$  und  $I_4$  zeigt, dass die gewählten Kantenpfeile mit der technischen Stromrichtung in der berechneten Lösung übereinstimmen.

## Zweigstrom-Verfahren — Umformulierung

Wir können die Gleichungen für ein elektrisches Netzwerk auch etwas systematischer studieren. Dazu betrachten wir zunächst die Inzidenzmatrix  $\mathbf{IM}$  des Netzwerkes, die  $m$  Zeilen sowie  $n$  Spalten besitzt und deren Einträge nur Werte in  $\{-1, 0, +1\}$  annehmen. Genauer gesagt: In der  $i$ -ten Zeile und der  $j$ -ten Spalte der Inzidenzmatrix steht genau dann  $-1$  bzw.  $+1$ , wenn der  $j$ -te Knoten des Netzwerkes der Anfangs- bzw. Endknoten der  $i$ -ten Kante ist. Für das obige Beispiel ergibt sich

$$\mathbf{IM} = \begin{pmatrix} -1 & +1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & 1 & -1 \\ -1 & 0 & 0 & 1 \\ -1 & 0 & 1 & 0 \end{pmatrix}, \quad \mathbf{IM}^T = \begin{pmatrix} -1 & 0 & 0 & -1 & -1 \\ 1 & -1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & -1 & 1 & 0 \end{pmatrix}.$$

Die Kirchhoffschen Knotenregeln lassen sich mit Hilfe der Inzidenzmatrix in kompakter Form als

$$\mathbf{IM}^T \cdot \mathbf{I} = \mathbf{0} \quad (n \text{ Gleichungen für die } m \text{ Ströme } I_i)$$

schreiben, wobei auf beiden Seiten des Gleichheitszeichens ein  $n$ -dimensionaler Vektor steht und  $\mathbf{I}$  den  $m$ -dimensionalen Spaltenvektor mit Komponenten  $I_i$  bezeichnet.

Desweiteren betrachten wir die Maschenmatrix  $\mathbf{MM}$  mit  $l$  Zeilen und  $m$ -Spalten, wobei das Element in der  $k$ -ten Zeile und der  $i$ -ten Spalte nur dann von 0 verschieden ist, wenn die Kante  $i$  zur Masche  $k$  gehört; in diesem Fall wählen wir den Wert  $+1$  bzw.  $-1$ , je nachdem ob die Orientierung der Kante  $i$  (der gewählte Pfeil) und die Orientierung der Masche  $k$  (meist gegen den Uhrzeigersinn) kompatibel bzw. inkompatibel sind. Für das konkrete Netzwerk ergibt sich damit:

$$\mathbf{MM} = \begin{pmatrix} 1 & 1 & 0 & 0 & -1 \\ 0 & 0 & -1 & -1 & +1 \\ 1 & 1 & -1 & -1 & 0 \end{pmatrix}, \quad \mathbf{MM}^T = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 0 & 1 \\ 0 & -1 & -1 \\ 0 & -1 & -1 \\ -1 & 1 & 0 \end{pmatrix}.$$

Die Kirchhoffschen Maschenregeln können nun als

$$\mathbf{MM} \cdot (\mathbf{U} - \mathbf{V}) = \mathbf{0} \quad (l \text{ Gleichungen für die } m \text{ Spannungen } U_i)$$

kodiert werden, wobei wir schon jeder Kante  $i$  ohne Stromquelle den Wert  $V_i = 0$  zugeordnet hatten.

Schließlich führen wir noch die quadratische Widerstandsmatrix  $\mathbf{WM}$  ein, indem wir aus allen Werten  $R_i$  eine  $(m, m)$ -Diagonalmatrix erzeugen. Im Beispiel meint dies

$$\mathbf{WM} = \begin{pmatrix} R_1 & 0 & 0 & 0 & 0 \\ 0 & R_2 & 0 & 0 & 0 \\ 0 & 0 & R_3 & 0 & 0 \\ 0 & 0 & 0 & R_4 & 0 \\ 0 & 0 & 0 & 0 & R_5 \end{pmatrix}.$$

Insgesamt können wir das Gleichungssystem für die Kantenströme in Form der zwei matrixwertigen Gleichungen

$$\mathbf{IM}^T \cdot \mathbf{I} = \mathbf{0}, \quad \mathbf{MM} \cdot (\mathbf{WM} \cdot \mathbf{I} - \mathbf{V}) = \mathbf{0}$$

schreiben, die jeweils die  $n$  Knotenregeln und die  $l$  Maschenregeln repräsentieren. (Die erweiterte Koeffizientenmatrix von oben umfasst alle  $n + l$  Gleichungen).

**Bemerkung** Es ist bei mittleren und großen Netzwerken nicht einfach zu erkennen, wieviele Maschengleichungen linear unabhängig sind bzw. wieviele man eigentlich ignorieren kann. Die Graphentheorie liefert für dieses Problem effiziente Algorithmen (zum Beispiel mit Hilfe geeignet gewählter oder berechneter Spannbäume), und man kann zeigen, dass es genau  $m - n + 1$  sogenannte Fundamentalmaschen gibt. Man kann aber auch immer gefahrlos die überflüssigen Gleichungen mitnehmen.

## Knotenpotential-Verfahren

Wir besprechen nun eine zweite Netzwerkmethod, bei der man die elektrischen Potentiale  $\Phi_j$  in den Knoten betrachtet. Der Schlüssel für dieses Verfahren sind die Kantengleichungen

$$U_i - V_i = \Phi_{a(i)} - \Phi_{e(i)},$$

die entlang jeder Kante  $i$  die Potentialdifferenz in den Eckknoten mit den Spannungen in Verbindung setzt. Hierbei meint

$$a(i) \in \{1, \dots, n\} \quad \text{bzw.} \quad e(i) \in \{1, \dots, n\}$$

den Knotenindex des Anfangs- bzw. des Endknoten der Kante  $i$ . Insbesondere sind diese Werte gerade die Spaltenindizes, für die in der  $i$ -ten Zeile der Inzidenzmatrix gerade  $-1$  bzw.  $+1$  steht. Im konkreten Fall gilt:

$$\begin{aligned} U_1 - V_1 &= \Phi_1 - \Phi_2 \\ U_2 - V_2 &= \Phi_2 - \Phi_3 \\ U_3 - V_3 &= \Phi_4 - \Phi_3 \\ U_4 - V_4 &= \Phi_1 - \Phi_4 \\ U_5 - V_5 &= \Phi_1 - \Phi_3 \end{aligned}$$

Ganz allgemein kann dies als

$$\mathbf{U} - \mathbf{V} = -\mathbf{IM} \cdot \Phi$$

geschrieben werden, wobei das Minus-Zeichen auf der rechten Seite die technische Stromrichtung repräsentiert (die sich von der physikalischen im Vorzeichen unterscheidet).

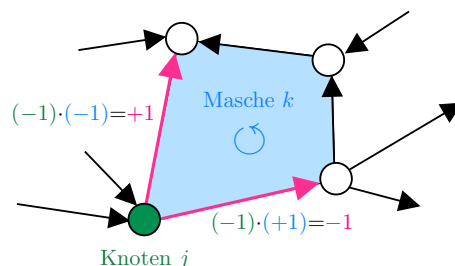
Die erste wesentliche Beobachtung ist nun, dass diese Gleichung – bzw. der Übergang zu den Knotenpotentialen – nach Multiplikation mit  $\mathbf{MM}$  schon die Gültigkeit aller Maschenregeln garantiert.

**Theorem** In jedem Netzwerk gilt

$$\mathbf{MM} \cdot \mathbf{IM} = \mathbf{0} \quad \text{bzw.} \quad \mathbf{IM}^T \cdot \mathbf{MM}^T = \mathbf{0}$$

im Sinne von  $(l, n)$ -Matrizen bzw.  $(n, l)$ -Matrizen. Beachte, dass die  $(l, m)$ -Matrix  $\mathbf{MM}$  und die  $(m, n)$ -Matrix  $\mathbf{IM}$  die *Geometrie des Netzwerkes* beschreiben, d.h. sie hängen weder von den Quellspannungen  $V_i$  noch von den Widerständen  $R_i$  oder den Strömen  $I_i$  ab.

*Beweis:* Die Regeln der Matrizenmultiplikation implizieren, dass der Eintrag in der  $k$ -ten Zeile und der  $j$ -ten Spalte des Matrixproduktes  $\mathbf{MM} \cdot \mathbf{IM}$  eine Summe über Kantenindizes ist, wobei die Nullen in der Inzidenz- und in der Maschenmatrix garantieren, dass nur zwei von Null verschiedene Summanden auftauchen, die jeweils ein Produkt von zwei Zahlen mit Betrag 1 sind. Insbesondere sind nur die beiden Kanten relevant, die sowohl zum Knoten  $j$  als auch zur Masche  $k$  gehören.





Mit Hilfe des Bildes und elementarer Vertauschungen (Wechsel der Kanten- oder/oder Maschenorientierung) zeigt man nun leicht, dass für jede Wahl von  $k \in \{1, \dots, l\}$  und  $j \in \{1, \dots, n\}$  die zwei von Null verschiedenen Summanden sich im Vorzeichen unterscheiden, und dies impliziert die erste Behauptung. Die zweite folgt dann nach Transposition aller Matrizen.  $\square$

**Bemerkung** Insbesondere gilt: Zu jeder Masche, d.h. zu jeder Zeile von  $\mathbf{MM}$ , gibt es einen Spaltenvektor im Kern von  $\mathbf{IM}^T$ , dessen Einträge nur Werte in  $\{-1, 0, +1\}$  annehmen.

Die zweite wesentliche Beobachtung ist, dass die Kirchhoffschen Knotenregeln ein Gleichungssystem für die Knotenpotentiale liefern. In der Tat, ersetzen wir  $\mathbf{U}$  durch  $\Phi$  und  $\mathbf{V}$ , so impliziert das Ohmsche Gesetz die Gleichung

$$\mathbf{I} = \mathbf{LM} \cdot \mathbf{U} = \mathbf{LM} \cdot (-\mathbf{IM} \cdot \Phi + \mathbf{V}),$$

wobei die Matrix  $\mathbf{LM}$  gerade die Inverse zu  $\mathbf{WM}$  ist. Insbesondere ist  $\mathbf{LM}$  wieder eine Diagonalmatrix, wobei der  $(i, i)$ -Eintrag gerade der  $i$ -te Leitwert  $L_i = 1/R_i$  ist. Im konkreten Netzwerk gilt insbesondere

$$\mathbf{LM} = \mathbf{WM}^{-1} = \begin{pmatrix} L_1 & 0 & 0 & 0 & 0 \\ 0 & L_2 & 0 & 0 & \\ 0 & 0 & L_3 & 0 & 0 \\ 0 & 0 & 0 & L_4 & 0 \\ 0 & 0 & 0 & 0 & L_5 \end{pmatrix}.$$

Die  $n$  Kirchhoffschen Knotenregeln können damit als

$$(\mathbf{IM}^T \cdot \mathbf{LM} \cdot \mathbf{IM}) \cdot \Phi = \mathbf{IM}^T \cdot \mathbf{LM} \cdot \mathbf{V}$$

geschrieben werden, wobei alle Terme auf der rechten Seite bekannt sind und auf der linken Seite vor dem unbekanntem Vektor  $\Phi \in \mathbb{R}^n$  eine  $(n, n)$ -Matrix steht. Für das Beispiel ergibt sich

$$(\mathbf{IM}^T \cdot \mathbf{LM} \cdot \mathbf{IM}) = \begin{pmatrix} L_1+L_4+L_5 & -L_1 & -L_5 & -L_4 \\ -L_1 & L_1+L_2 & -L_2 & 0 \\ -L_5 & -L_2 & L_2+L_3+L_5 & -L_3 \\ -L_4 & 0 & -L_3 & L_3+L_4 \end{pmatrix}$$

sowie

$$\mathbf{IM}^T \cdot \mathbf{LM} \cdot \mathbf{V} = \begin{pmatrix} -L_5 \cdot V_5 \\ 0 \\ L_3 \cdot V_3 + L_5 \cdot V_5 \\ -L_3 \cdot V_3 \end{pmatrix}.$$

Sobald alle Knotenpotentiale durch Lösen des Gleichungssystem bestimmt wurden, kann anschließend für jede Kante  $i$  die Spannung  $U_i$  sowie der Strom  $I_i$  mit Hilfe von  $R_i$ ,  $V_i$  und  $\Phi_{a(i)}$ ,  $\Phi_{e(i)}$  direkt berechnet werden.

Ein Vorteil des Knoten-Potentialverfahrens besteht darin, dass man die Maschenmatrix weder aufschreiben noch auf linear abhängige Zeilen untersuchen muss. Außerdem besitzt das zu lösende Gleichungssystem für die  $\Phi_j$  eine quadratische und

symmetrische Koeffizientenmatrix, und dies erleichtert bei großen Systemen die numerische Berechnung von Lösungen erheblich. Die gewonnenen Resultate sind aber immer äquivalent zu den Kantenströmen, die mit Hilfe von Rechtecksmatrizen in der Zweigstrom-Analyse ermittelt werden. Beachte auch, dass die Potentiale in den *Knoten* leben, wohingegen die Ströme durch *Kanten* fließen.

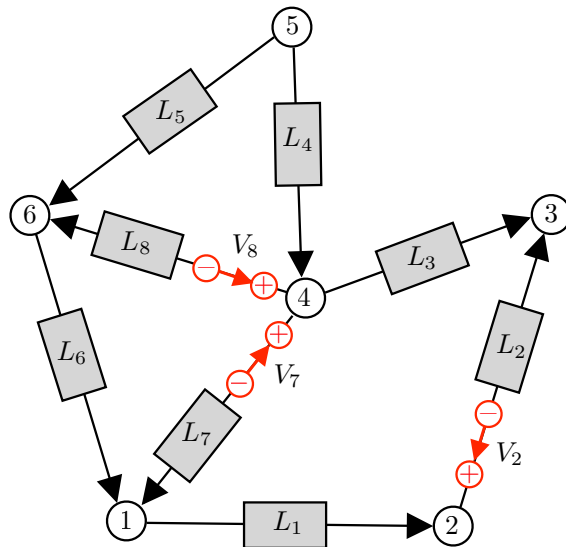
Für das konkrete Netzwerk ergibt sich mit den oben spezifizierten Zahlenwerten

$$\Phi = \begin{pmatrix} 1 \\ 2 \\ 3 \\ 0 \end{pmatrix} + \gamma \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} \quad \text{bzw.} \quad \Phi = \begin{pmatrix} 1 \\ 4 \\ 7 \\ 0 \end{pmatrix} + \gamma \cdot \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix},$$

als Lösung der entsprechenden Potentialgleichung, wobei der Parameter  $\gamma \in \mathbb{R}$  beliebig gewählt werden kann. Diese Mehrdeutigkeit ist nicht überraschend, da nur Potentialdifferenzen eine physikalische Bedeutung besitzen. Insbesondere kann man  $\gamma$  immer so wählen, dass das Potential in einem frei gewählten Knoten verschwindet (*Erdung* des Netzwerkes).

**Die Struktur der Potentialgleichung** Die Rechnungen für das erste Netzwerkbeispiel illustrieren die folgenden allgemeinen Prinzipien:

1. Für je zwei Knotenindizes  $j_1 \neq j_2$  verschwinden die Matrixeinträge  $(j_1, j_2)$  und  $(j_2, j_1)$  genau dann, wenn die Knoten  $j_1$  und  $j_2$  nicht durch eine Kante verbunden sind. Andernfalls steht in beiden Einträgen der negative Leitwert der Verbindungskante.
2. Alle Diagonaleinträge sind positiv, so dass alle Zeilen- und Spaltensummen in der Matrix verschwinden. Insbesondere enthält der Diagonaleintrag  $(j, j)$  die Summe der Leitwerte aller Kanten, für die der Knoten  $j$  entweder Anfangs- oder Endknoten ist.
3. Die  $j$ -te Zeile der rechten Seite ist eine gewichtete Summe von Beiträgen der Bauart  $L_i \cdot V_i$  über alle Kanten  $i$ , die den Knoten  $j$  enthalten, wobei ein- bzw. auslaufenden Kanten mit  $+1$  bzw.  $-1$  gewichtet sind. (Beachte, dass die Pfeile der Kanten und der Spannungsquellen entgegengesetzt gewählt wurden).



Zweites Beispiel für elektrisches Netzwerk mit  $n = 6$ ,  $m = 8$  und Leitwerten  $L_i = 1/R_i$ . Es gibt insgesamt  $l = 6$  Maschen, aber nur  $m - n + 1 = 3$  Fundamentalmaschen, aus denen die anderen zusammengesetzt werden können.

Mit diesen Regeln ergeben sich für das zweite, etwas kompliziertere Netzwerkbeispiel die Koeffizientenmatrix

$$\begin{pmatrix} L_1+L_7+L_6 & -L_1 & 0 & -L_7 & 0 & -L_6 \\ -L_1 & L_1+L_2 & -L_2 & 0 & 0 & 0 \\ 0 & -L_2 & L_2+L_3 & -L_3 & 0 & 0 \\ -L_7 & 0 & -L_3 & L_7+L_3+L_4+L_8 & -L_4 & -L_8 \\ 0 & 0 & 0 & -L_4 & L_4+L_5 & -L_5 \\ -L_6 & 0 & 0 & -L_8 & -L_5 & +L_6+L_8+L_5 \end{pmatrix}$$

sowie die rechte Seite

$$\begin{pmatrix} L_7 \cdot V_7 \\ -L_2 \cdot V_2 \\ L_2 \cdot V_2 \\ -L_7 \cdot V_7 - L_8 \cdot V_8 \\ 0 \\ L_8 \cdot V_8 \end{pmatrix},$$

aber beide Formeln können alternativ auch schrittweise aus den Kirchhoffschen Regeln abgeleitet werden.

Die im Knotenpotential-Verfahren entstehende Koeffizientenmatrix ist ein Beispiel für einen gewichteten Laplace-Operator auf Graphen. Solche und ähnliche Operatoren sind von zentraler Bedeutung in der Mathematik und den Anwendungswissenschaften. Sie tauchen in sehr vielen und sehr unterschiedlichen Kontexten auf.

## Maschenstrom-Verfahren

Es gibt noch ein drittes Verfahren, bei dem jeder Masche  $k$  ein sogenannter *Maschenstrom*  $J_k$  zugeordnet wird, wobei die Kantenströme sich als Linearkombinationen der Maschenströme darstellen lassen. Für das erste Netzwerkbeispiel ergeben sich unter sorgfältiger Berücksichtigung aller Orientierungen die Gleichungen

$$\begin{aligned} I_1 &= + J_1 && + J_3 \\ I_2 &= + J_1 && + J_3 \\ I_3 &= && - J_2 - J_3 \\ I_4 &= && - J_2 - J_3 \\ I_4 &= - J_1 &+ J_2 \end{aligned}$$

und ganz allgemein gilt immer

$$\mathbf{I} = \mathbf{M}\mathbf{M}^T \cdot \mathbf{J},$$

wobei  $\mathbf{J}$  gerade der  $l$ -dimensionale Vektor der Maschenströme ist. Der wesentliche Punkt bei dieser Methode ist, dass die Einführung der Maschenströme automatisch die Gültigkeit aller Kirchhoffschen Knotenregeln garantiert. In der Tat, es gilt

$$\mathbf{I}\mathbf{M}^T \cdot \mathbf{I} = \mathbf{I}\mathbf{M}^T \cdot \mathbf{M}\mathbf{M}^T \cdot \mathbf{J} = \mathbf{0}^T \cdot \mathbf{J} = \mathbf{0},$$

wobei wir das weiter oben bewiesene Geometrie-Theorem benutzt haben. Außerdem kann nun mit Hilfe der Kirchhoffschen Maschenregeln und der Ohmschen Gesetze das Gleichungssystem

$$\mathbf{M}\mathbf{M} \cdot \mathbf{R}\mathbf{M} \cdot \mathbf{M}\mathbf{M}^T \cdot \mathbf{J} = \mathbf{M}\mathbf{M} \cdot \mathbf{V}$$

für  $\mathbf{J}$  abgeleitet werden, wobei die Koeffizientenmatrix diesmal aus  $l$  Zeilen und  $l$  Spalten besteht. Für das erste Netzwerkbeispiel ergibt sich die rechte Seite

$$\mathbf{MM} \cdot \mathbf{V} = \begin{pmatrix} -V_5 \\ -V_3 + V_5 \\ -V_3 \end{pmatrix}$$

sowie die Koeffizientenmatrix

$$\mathbf{MM} \cdot \mathbf{RM} \cdot \mathbf{MM}^T = \begin{pmatrix} R_1 + R_2 + R_5 & -R_5 & R_1 + R_2 \\ -R_5 & R_3 + R_4 + R_5 & R_3 + R_4 \\ R_1 + R_2 & R_3 + R_4 & R_1 + R_2 + R_3 + R_4 \end{pmatrix}.$$

Beachte, dass die quadratische Koeffizientenmatrix im Maschenstrom-Verfahren wieder symmetrisch ist, aber diesmal werden die Einträge aus den Widerständen gebildet und maschenweise, nicht knotenweise zusammengesetzt. Außerdem wird die Koeffizientenmatrix im Allgemeinen einen Kern besitzen, wobei dessen Dimension gerade die Anzahl der linear abhängigen Maschengleichungen ist.

Mit den konkreten Zahlenwerten von oben berechnen wir

$$\mathbf{J} = \begin{pmatrix} -1 \\ -1 \\ 0 \end{pmatrix} + \delta \cdot \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix} \quad \text{bzw.} \quad \mathbf{J} = \begin{pmatrix} -3 \\ -1 \\ 0 \end{pmatrix} + \delta \cdot \begin{pmatrix} 1 \\ 1 \\ -1 \end{pmatrix}$$

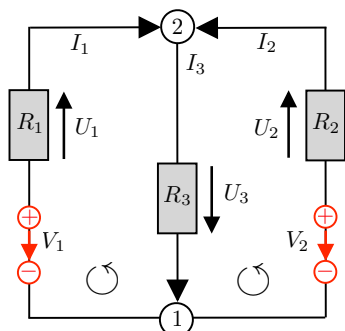
als allgemeine Lösung für die Maschenströme, wobei der Parameter  $\delta \in \mathbb{R}$  beliebig gewählt werden darf. Anschließend erhalten wir nach Multiplikation mit  $\mathbf{M}^T$  dieselben Kantenströme wie in der Zweigstrom-Analyse und im Knotenpotential-Verfahren, und zwar für jede Wahl von  $\delta$ .

## Diskussion

1. Alle drei Verfahren führen auf unterschiedlichen Wegen zum gleichen Ergebnis. Insbesondere werden einmal die Kantenströme direkt, die anderen beiden Male über den „Umweg“ Knotenpotentiale bzw. Maschenströme berechnet.
2. Jedes Verfahren führt zu linearen Gleichungssystemen, aber die entsprechenden Koeffizientenmatrizen sind sehr verschieden.
3. Jedes Verfahren hat seine Vor- und Nachteile. Bei großen Netzwerken wird in der Praxis entweder das Knotenpotential- oder das Maschenstrom-Verfahren verwendet (je nach Anwendung und Netzwerk).
4. Beim Maschenstrom-Verfahren muss man explizit alle Maschen kennen bzw. ermitteln. Das Knotenpotential-Verfahren beruht hingegen nur auf der Kenntnis von Kanten und Knoten.
5. Knotenpotential- und Maschenstrom-Verfahren sind aus mathematischer Sicht zueinander *konjugierte* oder *duale* Probleme.
6. Alle Verfahren erlauben es, durch geschickte Tricks (Zusammenfassung von Knoten und/oder Kanten, Einschränkung auf fundamentale Maschen usw.) die Anzahl der Gleichungen zu reduzieren. Im ersten Netzwerk könnte man zum Beispiel den Knoten 2 entfernen und die Kanten 1 und 2 zu einer einzigen Kante mit Widerstand  $R_1 + R_2$  zusammenfassen.

7. Wir haben in diesem Abschnitt die mathematisch-geprägte Sicht des Dozenten auf die drei Verfahren kennengelernt, die von der ingenieurwissenschaftlichen Darstellung in Notation und Terminologie abweichen kann (je nach Quelle und Autor).

Zum Abschluss geben wir noch ein kleines Übungsnetzwerk an, mit dem man sowohl die finale Äquivalenz als auch die algorithmischen Unterschiede des Zweigstrom-, des Knotenpotential- und des Maschenstrom-Verfahren sehr gut studieren kann.



Einfachster Fall:

$$V_1 = V_2 = V, \quad R_1 = R_2 = R$$

↓

$$I_1 = I_2 = \frac{V}{3R}, \quad I_3 = \frac{2V}{3R}$$

$$U_1 = U_2 = \frac{V}{3}, \quad U_3 = \frac{2V}{3}$$

Einfaches Netzwerk mit  $n = 2$ ,  $m = 2$  und  $l = 2 + 1$  als Testobjekt. Was passiert jeweils bei Änderung der Kanten- oder Maschenorientierungen?



# Kapitel 5

## Lineare Abbildungen

Vorlesung 14, 20. November 2019

### 5.1 Basisdarstellungen und -wechsel

**Klarstellung** Alle nachfolgenden Betrachtungen gelten wieder im Reellen und ganz analog im Komplexen.

**Definition** Eine Abbildung  $T : V \rightarrow W$  zwischen zwei Vektorräumen heißt linear, falls

$$T(\lambda \cdot \mathbf{v} + \tilde{\lambda} \cdot \tilde{\mathbf{v}}) = \lambda \cdot T(\mathbf{v}) + \tilde{\lambda} \cdot T(\tilde{\mathbf{v}})$$

für alle Vektoren  $\mathbf{v}, \tilde{\mathbf{v}} \in V$  und alle Skalare  $\lambda, \tilde{\lambda}$  gilt.

#### Beispiele

1. Jede  $(m, n)$ - Matrix  $\mathbf{A}$  definiert via

$$T(\mathbf{x}) = \mathbf{A} \cdot \mathbf{x}$$

in natürlicher Weise eine lineare Abbildung  $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ .

2. Die Vorschrift

$$f \mapsto T(f) := \int_0^1 f(s) \, ds$$

definiert eine lineare Abbildung vom Raum aller stetigen Funktionen auf dem Intervall  $[0, 1]$  in die Menge der reellen Zahlen.

3. Die Ableitungsoperation  $p \mapsto p'$  mit

$$p'(x) = \sum_{k=0}^{n-1} (k+1) \cdot \alpha_{k+1} \cdot x^k \quad \text{für} \quad p(x) = \sum_{k=0}^n \alpha_k \cdot x^k$$

ist eine lineare Abbildung des Raumes aller Polynome vom Grad  $\leq n$  in sich.

4. Die Umkehrabbildung einer linearen und bijektiven Abbildung ist auch linear.

**Bemerkung** Eine lineare Abbildung  $T : V \rightarrow W$  ist bereits eindeutig durch die Angabe von

$$T(\mathbf{v}_1), \dots, T(\mathbf{v}_n)$$

charakterisiert, sofern die Vektoren  $\mathbf{v}_1, \dots, \mathbf{v}_n$  eine Basis des Definitionsbereiches  $V$  bilden (Übungsaufgabe).

**Theorem (Basisdarstellung linearer Abbildung)** Seien  $V$  bzw.  $W$  endlich-dimensionale Räume der Dimension  $n$  bzw.  $m$  und seien

$$(\mathbf{v}_1, \dots, \mathbf{v}_n), \quad (\mathbf{w}_1, \dots, \mathbf{w}_m)$$

zwei gegebene Basen. Dann existiert für jede lineare Abbildung  $T : V \rightarrow W$  genau eine  $(m, n)$ -Matrix  $\mathbf{A}$ , so dass die folgenden Aussagen erfüllt sind:

1. Es gilt

$$T(\mathbf{v}_j) = \sum_{i=1}^m a_{ij} \cdot \mathbf{w}_i$$

für alle  $i = 1, \dots, m$  und alle  $j = 1, \dots, n$ .

2. Für je zwei Vektoren  $\mathbf{v} \in V$  und  $\mathbf{w} \in W$  mit den Basisdarstellungen

$$\mathbf{v} = \sum_{j=1}^n x_j \cdot \mathbf{v}_j, \quad \mathbf{w} = \sum_{i=1}^m y_i \cdot \mathbf{w}_i$$

gilt die logische Implikation

$$\mathbf{w} = T(\mathbf{v}) \quad \Leftrightarrow \quad \mathbf{y} = \mathbf{A} \cdot \mathbf{x},$$

wobei  $\mathbf{x} \in \mathbb{R}^n$  bzw.  $\mathbf{y} \in \mathbb{R}^m$  aus den Komponenten  $x_j$  bzw.  $y_i$  bestehen.

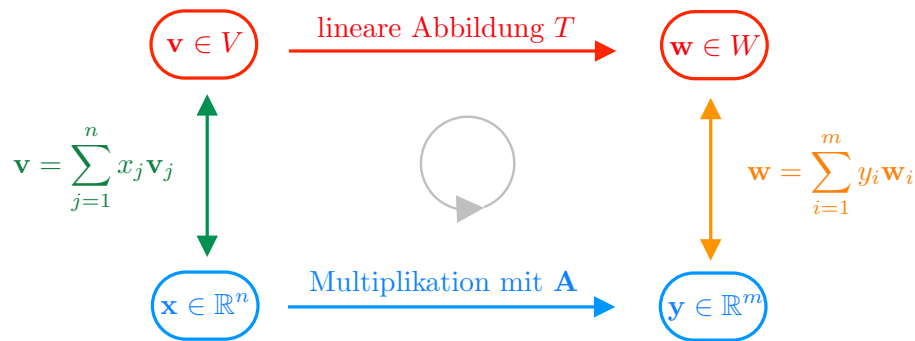
Die Matrix  $\mathbf{A}$  wird die Basisdarstellung (oder die Matrixdarstellung) der linearen Abbildung  $T$  bzgl. der gewählten Basen genannt.

Beweis: Für jedes feste  $j$  besitzt  $T(\mathbf{v}_j) \in W$  eine eindeutige Darstellung bzgl. der Basis in  $W$ . Die entsprechenden Komponenten nennen wir  $a_{1j}, \dots, a_{mj}$  und schreiben Sie als Einträge in die  $j$ -te Spalte einer  $(m, n)$ -Matrix  $\mathbf{A}$ . Insbesondere gilt damit die erste Behauptung nach Konstruktion und wir erhalten wegen der Linearität von  $T$  die Formel

$$\begin{aligned} T(\mathbf{v}) &= T\left(\sum_{j=1}^n x_j \cdot \mathbf{v}_j\right) = \sum_{j=1}^n x_j \cdot T(\mathbf{v}_j) = \sum_{j=1}^n x_j \cdot \sum_{i=1}^m a_{ij} \cdot \mathbf{w}_i \\ &= \sum_{i=1}^m \left(\sum_{j=1}^n a_{ij} \cdot x_j\right) \cdot \mathbf{w}_i \end{aligned}$$

für jedes  $\mathbf{v} \in V$ . Da die Vektoren  $\mathbf{w}_i$  aber eine Basis in  $W$  bilden, kann die Formel  $T(\mathbf{v}) = \sum_{i=1}^m y_i \cdot \mathbf{w}_i$  nur dann gelten, wenn die Koeffizienten in beiden Formeln übereinstimmen. Dies ist aber gerade die komponentenweise Formulierung der zweiten Behauptung.  $\square$





Flussdiagramm zur Basisdarstellung linearer Abbildungen.

**Beispiel** Wir betrachten  $V = W = \mathbb{R}^2$  sowie die lineare Abbildung  $T : V \rightarrow W$  mit

$$\begin{pmatrix} \varrho \cdot \cos(\theta) \\ \varrho \cdot \sin(\theta) \end{pmatrix} \xrightarrow{T} \begin{pmatrix} \varrho \cdot \cos(\theta + \alpha) \\ \varrho \cdot \sin(\theta + \alpha) \end{pmatrix},$$

die mittels Polarkoordinaten eingeführt wird und eine Drehung um den Winkel  $\alpha$  beschreibt. Wir wählen nun die kanonische Basis in  $V$  und  $W$ , d.h.

$$\mathbf{v}_1 = \mathbf{w}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \mathbf{v}_2 = \mathbf{w}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

und erhalten

$$T(\mathbf{v}_1) = \begin{pmatrix} \cos(\alpha) \\ \sin(\alpha) \end{pmatrix} = \cos(\alpha) \cdot \mathbf{w}_1 + \sin(\alpha) \cdot \mathbf{w}_2$$

sowie

$$T(\mathbf{v}_2) = \begin{pmatrix} \cos(\alpha + \pi/2) \\ \sin(\alpha + \pi/2) \end{pmatrix} = \begin{pmatrix} -\sin(\alpha) \\ \cos(\alpha) \end{pmatrix} = -\sin(\alpha) \cdot \mathbf{w}_1 + \cos(\alpha) \cdot \mathbf{w}_2,$$

wobei wir die Additionstheoreme benutzt haben. Wir sehen nun, dass die Drehung um den Winkel  $\alpha$  bzgl. der kanonischen Basis durch die Drehmatrix

$$\mathbf{A} = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}$$

beschrieben wird.

**Beispiel** Wir betrachten die Räume und die Basis aus dem vorangegangenen Beispiel sowie die lineare Abbildung

$$\begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \xrightarrow{T} \begin{pmatrix} v_2 \\ v_1 \end{pmatrix} = T\left(\begin{pmatrix} v_1 \\ v_2 \end{pmatrix}\right),$$

die einer Indexvertauschung entspricht. Diesmal finden wir

$$T(\mathbf{v}_1) = \mathbf{v}_2, \quad T(\mathbf{v}_2) = \mathbf{v}_1$$

und damit die einfache Permutationsmatrix

$$\mathbf{A} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

**Bemerkung** Statten wir die Räume  $V = \mathbb{R}^n$  und  $W = \mathbb{R}^m$  mit ihren jeweiligen kanonischen Basen aus, so gibt es keinen Unterschied zwischen linearen Abbildungen und ihren Basisdarstellungen durch Matrizen. Genauer gesagt: Jede  $(m, n)$ -Matrix definiert in natürlicher Weise eine lineare Abbildung  $\mathbb{R}^n \rightarrow \mathbb{R}^m$  und die durch das Theorem bereitgestellte Matrix ist gerade wieder die Matrix, mit der wir die Abbildung definiert haben. Für  $m = 3$  und  $n = 2$  folgt dies aus der ersten Formel im Theorem kombiniert mit

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} = a_{11} \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + a_{21} \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + a_{31} \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix} \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} = a_{12} \cdot \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + a_{22} \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + a_{32} \cdot \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

aber die Argumente gelten ganz analog auch für andere Werte von  $m$  und  $n$ .

Achtung: Bei allgemeinen Vektorräumen bzw. nicht-kanonischen Basen muss man aber zwischen Abbildungen und ihren Basisdarstellungen sauber trennen.

**Beispiel** Für  $V = \mathbb{R}^3$  und  $W = \mathbb{R}^2$  betrachten wir die lineare Abbildung

$$\mathbf{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \xrightarrow{T} \mathbf{w} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} 1 \cdot v_1 + 2 \cdot v_2 + 3 \cdot v_3 \\ 4 \cdot v_1 + 5 \cdot v_2 + 6 \cdot v_3 \end{pmatrix},$$

die durch Multiplikation mit einer konkreten  $(2, 3)$ -Matrix definiert ist, und wir hatten uns gerade überlegt, dass diese Matrix auch die Basisdarstellung der Abbildung bzgl. der kanonischen Basen in  $V$  und  $W$  ist.

Wir wollen nun die Basisdarstellung der Abbildung  $T$  bzgl. der Basen

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$$

und

$$\mathbf{w}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \mathbf{w}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

herleiten. Dazu berechnen wir zunächst das Bild jedes Basisvektors  $\mathbf{v}_j$  in  $V$  und stellen dieses dann bzgl. der Basis in  $W$  dar. Durch einfache Rechnungen verifizieren wir

$$T(\mathbf{v}_1) = \begin{pmatrix} 3 \\ 9 \end{pmatrix} = 6 \cdot \mathbf{w}_1 + 3 \cdot \mathbf{w}_2$$

$$T(\mathbf{v}_2) = \begin{pmatrix} 5 \\ 11 \end{pmatrix} = 8 \cdot \mathbf{w}_1 + 3 \cdot \mathbf{w}_2$$

$$T(\mathbf{v}_3) = \begin{pmatrix} 4 \\ 10 \end{pmatrix} = 7 \cdot \mathbf{w}_1 + 3 \cdot \mathbf{w}_2$$

und können nun (nach dem ersten Teil des Theorems) die Spalten der Matrix  $\mathbf{A}$  nacheinander ablesen. Wir erhalten

$$\mathbf{A} = \begin{pmatrix} 6 & 8 & 7 \\ 3 & 3 & 3 \end{pmatrix}$$

als Basisdarstellung von  $T$  bzgl. der gewählten Basen in  $V$  und  $W$ .

Interpretation: Die Matrizen

$$\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} 6 & 8 & 7 \\ 3 & 3 & 3 \end{pmatrix}$$

beschreiben dieselbe lineare Abbildung  $\mathbb{R}^3 \rightarrow \mathbb{R}^2$  nur bzgl. unterschiedlicher Basiswahlen. Die erste Matrix entspricht dabei den kanonischen Basen, die zweite den oben angegebenen.

Bemerkung: Wir wollen zur Vollständigkeit festhalten, dass wir eigentlich drei inhomogene Gleichungssysteme gelöst haben. In der Tat, für jedes  $j = 1, \dots, 3$  soll

$$T(\mathbf{v}_j) = a_{1j} \cdot \mathbf{w}_1 + a_{2j} \cdot \mathbf{w}_2 = a_{1j} \cdot \begin{pmatrix} w_{1,1} \\ w_{1,2} \end{pmatrix} + a_{2j} \cdot \begin{pmatrix} w_{2,1} \\ w_{2,2} \end{pmatrix} = \begin{pmatrix} w_{1,1} & w_{2,1} \\ w_{1,2} & w_{2,2} \end{pmatrix} \cdot \begin{pmatrix} a_{1j} \\ a_{2j} \end{pmatrix}$$

gelten und wir erhalten daher insgesamt die drei Gleichungssysteme

$$\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} a_{11} \\ a_{21} \end{pmatrix} = \begin{pmatrix} 3 \\ 9 \end{pmatrix}, \quad \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix} = \begin{pmatrix} 5 \\ 11 \end{pmatrix}, \quad \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} a_{13} \\ a_{23} \end{pmatrix} = \begin{pmatrix} 4 \\ 10 \end{pmatrix}$$

für die drei zweidimensionalen Spaltenvektoren der Matrix  $\mathbf{A}$ . Beachte, dass alle diese Gleichungssysteme die gleiche Koeffizientenmatrix besitzen, deren  $i$ -te Spalte gerade der Vektor  $\mathbf{w}_i$  ist. Wir hatten die Lösung oben schon im Kopf berechnet, aber wir hätten natürlich alternativ auch das Gauß-Jordan-Verfahren verwenden können:

$$\left( \begin{array}{cc|ccc} 1 & -1 & 3 & 5 & 4 \\ 1 & 1 & 9 & 11 & 10 \end{array} \right) \cong \left( \begin{array}{cc|ccc} 1 & -1 & 3 & 5 & 4 \\ 0 & 1 & 3 & 3 & 3 \end{array} \right) \cong \left( \begin{array}{cc|ccc} 1 & 0 & 6 & 8 & 7 \\ 0 & 1 & 3 & 3 & 3 \end{array} \right)$$

**Kochrezept** Unsere Beispiele illustrieren das folgende, allgemeine Prinzip für den Fall  $V = \mathbb{R}^n$  und  $W = \mathbb{R}^m$ : Wir können die Basisdarstellung  $\mathbf{A}$  der gegebenen linearen Abbildung  $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$  dadurch bestimmen, dass wir die erweiterte Koeffizientenmatrix

$$\left( \begin{array}{ccc|ccc} | & & | & | & & | \\ \mathbf{w}_1 & \dots & \mathbf{w}_m & T(\mathbf{v}_1) & \dots & T(\mathbf{v}_n) \\ | & & | & | & & | \end{array} \right)$$

solange mit dem Gauß-Jordan-Verfahren bearbeiten, bis links vom Mittelstrich die Einheitsmatrix  $I_m$  steht und wir die gesuchte Matrix  $\mathbf{A}$  rechts vom Mittelstrich ablesen können.

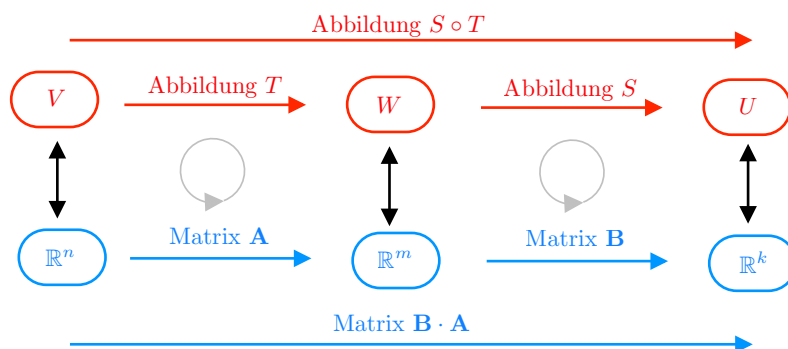
**Matrizenmultiplikation und Komposition von Abbildungen** Seien  $V, W, U$  nun drei Vektorräume der Dimension  $n, m, k$  und seien

$$T : V \rightarrow W, \quad S : W \rightarrow U$$

zwei lineare Abbildungen. Sind nun Basen  $(\mathbf{v}_1, \dots, \mathbf{v}_n)$  bzw.  $(\mathbf{w}_1, \dots, \mathbf{w}_m)$  bzw.  $(\mathbf{u}_1, \dots, \mathbf{u}_k)$  von  $V$  bzw.  $W$  bzw.  $U$  gegeben, und sind  $\mathbf{A}$  bzw.  $\mathbf{B}$  die Basisdarstellungen von  $T$  bzw.  $S$ , so ist

$$\mathbf{B} \cdot \mathbf{A}$$

die Basisdarstellung von  $S \circ T$ . Diese wichtige Tatsache ergibt sich unmittelbar aus der Definition von Basisdarstellung und den Regeln der Matrizenmultiplikation (oder anders gesagt: man kann dies einfach nachrechnen). Insbesondere zeigt sich hier, dass die Matrizenmultiplikation in sehr sinnvoller Weise definiert wurde: *Die Matrizenmultiplikation entspricht der Komposition linearer Abbildungen.*



Flussdiagramm zur Basisdarstellung linearer Abbildungen.

**Folgerung** Eine lineare Abbildung  $T : V \rightarrow W$  ist dann genau bijektiv, wenn die Basisdarstellung  $\mathbf{A}$  eine invertierbare Matrix ist und in diesem Fall ist die Basisdarstellung der Umkehrabbildung  $T^{-1}$  gerade die zu  $\mathbf{A}$  inverse Matrix  $\mathbf{A}^{-1}$ . Insbesondere kann im Fall  $\dim V \neq \dim W$  die lineare Abbildung  $T$  niemals bijektiv sein.

**Beispiel** Mit den Additionstheoremen ergibt sich

$$\begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} \cdot \begin{pmatrix} \cos(\beta) & -\sin(\beta) \\ \sin(\beta) & \cos(\beta) \end{pmatrix} = \begin{pmatrix} \cos(\alpha + \beta) & -\sin(\alpha + \beta) \\ \sin(\alpha + \beta) & \cos(\alpha + \beta) \end{pmatrix},$$

d.h. die Komposition zweier Teildrehungen in der Ebene ist ein Gesamtdrehung um die Summe der Teilwinkel. Insbesondere gilt damit auch

$$\begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}^{-1} = \begin{pmatrix} \cos(-\alpha) & -\sin(-\alpha) \\ \sin(-\alpha) & \cos(-\alpha) \end{pmatrix} = \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & \cos(\alpha) \end{pmatrix},$$

d.h. die Inverse der (2, 2)-Drehmatrix mit Winkel  $\alpha$  ist gerade die (2, 2)-Drehmatrix zum Winkel  $-\alpha$ .

**Beispiel** Mit den Additionstheoremen und der Matrizen-Multiplikation berechnen wir

$$\begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \cdot \begin{pmatrix} \cos(\alpha) & \sin(\alpha) \\ -\sin(\alpha) & \cos(\alpha) \end{pmatrix} = \begin{pmatrix} \cos(2\alpha) & \sin(2\alpha) \\ \sin(2\alpha) & -\cos(2\alpha) \end{pmatrix}.$$

Die linke Seite beschreibt die Komposition (von rechts nach links) der Drehung um den Winkel  $-\alpha$ , der Spiegelung an der  $x$ -Achse, und der Drehung um den Winkel  $+\alpha$  und wir können mit unserem geometrischen Wissen schließen, dass auf der rechten Seite eine Spiegelungsmatrix steht. Es handelt sich dabei um die Spiegelung an der Ursprungsgeraden

$$\text{span} \left\{ \begin{pmatrix} \cos(\alpha) \\ \sin(\alpha) \end{pmatrix} \right\},$$

denn diese Vektoren werden unter der Abbildung nicht bewegt. Mit anderen Worten: Es gilt

$$\begin{pmatrix} \cos(2\alpha) & \sin(2\alpha) \\ \sin(2\alpha) & -\cos(2\alpha) \end{pmatrix} \cdot \begin{pmatrix} \cos(\alpha) \\ \sin(\alpha) \end{pmatrix} = \begin{pmatrix} \cos(2\alpha)\cos(\alpha) + \sin(2\alpha)\sin(\alpha) \\ \sin(2\alpha)\cos(\alpha) - \cos(2\alpha)\sin(\alpha) \end{pmatrix} = \begin{pmatrix} \cos(\alpha) \\ \sin(\alpha) \end{pmatrix},$$

wobei die letzte Gleichheit wieder aufgrund von Additionstheoremen gilt.

**Basiswechsel und Übergangsmatrizen** Sei  $V$  ein gegebener Vektorraum der Dimension  $n$  und seien

$$(\mathbf{v}_1, \dots, \mathbf{v}_n) \quad \text{bzw.} \quad (\tilde{\mathbf{v}}_1, \dots, \tilde{\mathbf{v}}_n)$$

zwei verschiedene Basen in  $V$ . Dann existieren für jeden Vektor  $\mathbf{v} \in V$  via

$$\sum_{j=1}^n x_j \cdot \mathbf{v}_j = \mathbf{v} = \sum_{j=1}^n \tilde{x}_j \cdot \tilde{\mathbf{v}}_j$$

zwei Komponentensätze, und es stellt sich die Frage, ob bzw. wie man für gegebenes  $\mathbf{v} \in V$  die entsprechenden Basisdarstellungen  $\mathbf{x} \in \mathbb{R}^n$  und  $\tilde{\mathbf{x}} \in \mathbb{R}^n$  ineinander umrechnen kann.

**Theorem (Übergangsmatrizen und Basiswechsel)** Mit den obigen Notationen gibt es genau zwei  $(n, n)$ -Matrizen  $\mathbf{S}$  und  $\tilde{\mathbf{S}}$  mit Einträgen  $s_{ij}$  und  $\tilde{s}_{ij}$ , so dass die folgenden vier Aussagen erfüllt sind:

1. Es gilt

$$\tilde{\mathbf{v}}_k = \sum_{j=1}^n s_{jk} \cdot \mathbf{v}_j \quad \text{sowie} \quad \mathbf{v}_k = \sum_{j=1}^n \tilde{s}_{jk} \cdot \tilde{\mathbf{v}}_j$$

für alle  $k = 1, \dots, n$ .

2. Es gilt  $\mathbf{S} \cdot \tilde{\mathbf{S}} = \tilde{\mathbf{S}} \cdot \mathbf{S} = \mathbf{I}_n$ , d.h. beide Matrizen sind zueinander invers.
3. Es gilt immer  $\mathbf{x} = \mathbf{S} \cdot \tilde{\mathbf{x}}$ , d.h.  $\mathbf{S}$  beschreibt den Übergang von  $\tilde{\mathbf{x}}$  nach  $\mathbf{x}$ .
4. Es gilt immer  $\tilde{\mathbf{x}} = \tilde{\mathbf{S}} \cdot \mathbf{x}$ , d.h.  $\tilde{\mathbf{S}}$  beschreibt den Übergang von  $\mathbf{x}$  nach  $\tilde{\mathbf{x}}$ .

*Beweis:* Wir können die erste Behauptung als Definition der Einträge von  $\mathbf{S}$  und  $\tilde{\mathbf{S}}$  betrachten und müssen nur noch zeigen, dass dann alle anderen Behauptungen erfüllt sind. Zunächst gilt

$$\sum_{j=1}^n \delta_{jk} \cdot \tilde{\mathbf{v}}_j = \tilde{\mathbf{v}}_k = \sum_{l=1}^n s_{lk} \cdot \mathbf{v}_l = \sum_{l=1}^n s_{lk} \cdot \left( \sum_{j=1}^n \tilde{s}_{jl} \cdot \tilde{\mathbf{v}}_j \right) = \sum_{j=1}^n \left( \sum_{l=1}^n \tilde{s}_{jl} \cdot s_{lk} \right) \cdot \tilde{\mathbf{v}}_j,$$

und ein Koeffizientenvergleich (die  $\tilde{\mathbf{v}}_j$  bilden eine Basis) liefert

$$\sum_{l=1}^n \tilde{s}_{jl} \cdot s_{lk} = \delta_{jk} = \begin{cases} 1 & \text{sofern } j = k \\ 0 & \text{sonst} \end{cases} \quad \text{für alle } j, k = 1, \dots, n,$$

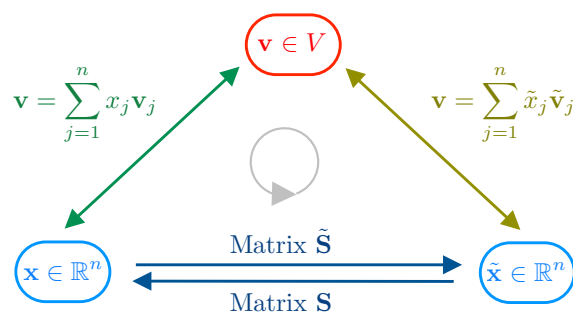
aber dies ist gerade die Komponentenschreibweise von  $\tilde{\mathbf{S}} \cdot \mathbf{S} = \mathbf{I}_n$ . Die Gleichung  $\mathbf{S} \cdot \tilde{\mathbf{S}} = \mathbf{I}_n$  kann vollkommen analog bewiesen werden. Andererseits gilt auch stets

$$\sum_{j=1}^n x_j \cdot \mathbf{v}_j = \mathbf{v} = \sum_{k=1}^n \tilde{x}_k \cdot \tilde{\mathbf{v}}_k = \sum_{k=1}^n \tilde{x}_k \cdot \left( \sum_{j=1}^n s_{jk} \cdot \mathbf{v}_j \right) = \sum_{j=1}^n \left( \sum_{k=1}^n s_{jk} \cdot \tilde{x}_k \right) \cdot \mathbf{v}_j$$

und ein Koeffizientenvergleich zeigt

$$x_j = \sum_{k=1}^n s_{jk} \cdot \tilde{x}_k \quad \text{für alle } j = 1, \dots, n,$$

aber dies ist gerade die Komponentenschreibweise von  $\mathbf{x} = \mathbf{S} \cdot \tilde{\mathbf{x}}$ . Die Herleitung der vierten Behauptung erfolgt analog.  $\square$



Flussdiagramm zu Basiswechseln und Übergangsmatrizen.

**Beispiel** Für  $V = \mathbb{R}^2$  und

$$\mathbf{v}_1 = \begin{pmatrix} 7 \\ -3 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} -19 \\ 8 \end{pmatrix}, \quad \tilde{\mathbf{v}}_1 = \begin{pmatrix} 2 \\ -1 \end{pmatrix}, \quad \tilde{\mathbf{v}}_2 = \begin{pmatrix} -3 \\ 1 \end{pmatrix}$$

verifizieren wir — durch scharfes Hinsehen oder kleine Rechnungen — die Formeln

$$\tilde{\mathbf{v}}_1 = 3 \cdot \mathbf{v}_1 + 1 \cdot \mathbf{v}_2, \quad \tilde{\mathbf{v}}_2 = 5 \cdot \mathbf{v}_1 + 2 \cdot \mathbf{v}_2.$$

sowie

$$\mathbf{v}_1 = 2 \cdot \tilde{\mathbf{v}}_1 - 1 \cdot \tilde{\mathbf{v}}_2, \quad \mathbf{v}_2 = -5 \cdot \tilde{\mathbf{v}}_1 + 3 \cdot \tilde{\mathbf{v}}_2$$

Wir können nun spaltenweise die Formeln

$$\mathbf{S} = \begin{pmatrix} 3 & 5 \\ 1 & 2 \end{pmatrix}, \quad \tilde{\mathbf{S}} = \begin{pmatrix} 2 & -5 \\ -1 & 3 \end{pmatrix}$$

ablesen und außerdem leicht per Hand nachprüfen, dass diese Matrizen wirklich invers zueinander sind.

Komponenten von Vektoren Zur Vollständigkeit wollen wir auch die Aussagen des Theorems zur Basisdarstellung von ausgewählten Vektoren  $\mathbf{v} \in V$  nachprüfen. Für den ersten Einheitsvektor gilt zum Beispiel

$$\mathbf{v} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = -8 \cdot \mathbf{v}_1 - 3 \cdot \mathbf{v}_2 = -1 \cdot \tilde{\mathbf{v}}_1 - 1 \cdot \tilde{\mathbf{v}}_2,$$

und dies impliziert die Basisdarstellungen

$$\mathbf{x} = \begin{pmatrix} -8 \\ -3 \end{pmatrix}, \quad \tilde{\mathbf{x}} = \begin{pmatrix} -1 \\ -1 \end{pmatrix}.$$

Wir können nun leicht nachrechnen, dass mit diesen konkreten Werten wirklich  $\mathbf{x} = \mathbf{S} \cdot \tilde{\mathbf{x}}$  und  $\tilde{\mathbf{x}} = \tilde{\mathbf{S}} \cdot \mathbf{x}$  gilt. Analoges gilt für den zweiten Einheitsvektor

$$\mathbf{v} = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = -19 \cdot \mathbf{v}_1 - 7 \cdot \mathbf{v}_2 = -3 \cdot \tilde{\mathbf{v}}_1 - 2 \cdot \tilde{\mathbf{v}}_2,$$

mit

$$\mathbf{x} = \begin{pmatrix} -19 \\ -7 \end{pmatrix}, \quad \tilde{\mathbf{x}} = \begin{pmatrix} -3 \\ -2 \end{pmatrix}.$$

Beachte, dass in beiden Fällen  $\mathbf{x} \neq \tilde{\mathbf{S}} \cdot \tilde{\mathbf{x}}$  und  $\tilde{\mathbf{x}} \neq \mathbf{S} \cdot \mathbf{x}$  gilt, aber diese Kombinationen tauchen auch weder im Theorem noch im Bild auf.

**Beispiel** Wir betrachten  $V = \mathbb{R}^3$  sowie

$$\mathbf{v}_1 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}$$

und

$$\tilde{\mathbf{v}}_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad \tilde{\mathbf{v}}_2 = \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}, \quad \tilde{\mathbf{v}}_3 = \begin{pmatrix} 0 \\ -1 \\ 0 \end{pmatrix}$$

Um die Koeffizienten der Matrix  $\mathbf{S}$  zu bestimmen, bemerken wir zunächst für jedes  $k$ , dass die Darstellungsformel für  $\tilde{\mathbf{v}}_k$  als lineares Gleichungssystem für die Komponenten der  $k$ -ten Spalte von  $\mathbf{S}$  geschrieben werden kann:

$$\begin{aligned} \begin{pmatrix} \tilde{v}_{k,1} \\ \tilde{v}_{k,2} \\ \tilde{v}_{k,3} \end{pmatrix} &= \tilde{\mathbf{v}}_k = s_{1k} \cdot \mathbf{v}_1 + s_{2k} \cdot \mathbf{v}_2 + s_{3k} \cdot \mathbf{v}_3 \\ &= \begin{pmatrix} | & | & | \\ \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \\ | & | & | \end{pmatrix} \cdot \begin{pmatrix} s_{1k} \\ s_{2k} \\ s_{3k} \end{pmatrix} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} s_{1k} \\ s_{2k} \\ s_{3k} \end{pmatrix}. \end{aligned}$$

Insbesondere können wir die Matrix  $\mathbf{S}$  durch das Gauß-Jordan-Verfahren bestimmen, und erhalten via

$$\begin{aligned} \left( \begin{array}{ccc|ccc} 0 & 1 & 1 & 0 & 2 & 0 \\ 1 & 0 & 1 & 0 & 0 & -1 \\ 1 & 1 & 0 & 1 & 0 & 0 \end{array} \right) &\cong \left( \begin{array}{ccc|ccc} 1 & 0 & 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 & 2 & 0 \\ 0 & 1 & -1 & 1 & 0 & 1 \end{array} \right) \cong \left( \begin{array}{ccc|ccc} 1 & 0 & 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 & 2 & 0 \\ 0 & 0 & -2 & 1 & -2 & 1 \end{array} \right) \\ &\cong \left( \begin{array}{ccc|ccc} 1 & 0 & 1 & 0 & 0 & -1 \\ 0 & 1 & 1 & 0 & 2 & 0 \\ 0 & 0 & 1 & -\frac{1}{2} & 1 & -\frac{1}{2} \end{array} \right) \cong \left( \begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{2} & -1 & -\frac{1}{2} \\ 0 & 1 & 0 & \frac{1}{2} & 1 & \frac{1}{2} \\ 0 & 0 & 1 & -\frac{1}{2} & 1 & -\frac{1}{2} \end{array} \right) \end{aligned}$$

die Formel

$$\mathbf{S} = \frac{1}{2} \begin{pmatrix} 1 & -2 & -1 \\ 1 & 2 & 1 \\ -1 & 2 & -1 \end{pmatrix}.$$

Anschließend können wir  $\tilde{\mathbf{S}}$  entweder als Inverse von  $\mathbf{S}$  oder mittels der Gauß-Jordan-Rechnung

$$\left( \begin{array}{ccc|ccc} 0 & 2 & 0 & 0 & 1 & 1 \\ 0 & 0 & -1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 1 & 0 \end{array} \right) \cong \left( \begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 2 & 0 & 0 & 1 & 1 \\ 0 & 0 & -1 & 1 & 0 & 1 \end{array} \right) \cong \left( \begin{array}{ccc|ccc} 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & 1 & -1 & 0 & -1 \end{array} \right)$$

bestimmen und erhalten

$$\tilde{\mathbf{S}} = \mathbf{S}^{-1} = \frac{1}{2} \begin{pmatrix} 2 & 2 & 0 \\ 0 & 1 & 1 \\ -2 & 0 & -2 \end{pmatrix}.$$

*Bemerkung:* Wir können aus den Spaltenvektoren von  $\mathbf{S}$  die Identitäten

$$\begin{aligned} \tilde{\mathbf{v}}_1 &= +\frac{1}{2} \cdot \mathbf{v}_1 + \frac{1}{2} \cdot \mathbf{v}_2 - \frac{1}{2} \cdot \mathbf{v}_3 \\ \tilde{\mathbf{v}}_2 &= -1 \cdot \mathbf{v}_1 + 1 \cdot \mathbf{v}_2 + 1 \cdot \mathbf{v}_3 \\ \tilde{\mathbf{v}}_3 &= -\frac{1}{2} \cdot \mathbf{v}_1 + \frac{1}{2} \cdot \mathbf{v}_2 + \frac{1}{2} \cdot \mathbf{v}_3 \end{aligned}$$

ablesen, wohingegen die Spalten von  $\tilde{\mathbf{S}}$  uns

$$\begin{aligned} \mathbf{v}_1 &= +1 \cdot \tilde{\mathbf{v}}_1 + 0 \cdot \tilde{\mathbf{v}}_2 - 1 \cdot \tilde{\mathbf{v}}_3 \\ \mathbf{v}_2 &= +1 \cdot \tilde{\mathbf{v}}_1 + \frac{1}{2} \cdot \tilde{\mathbf{v}}_2 + 0 \cdot \tilde{\mathbf{v}}_3 \\ \mathbf{v}_3 &= +0 \cdot \tilde{\mathbf{v}}_1 + \frac{1}{2} \cdot \tilde{\mathbf{v}}_2 - 1 \cdot \tilde{\mathbf{v}}_3 \end{aligned}$$

liefern.

**Kochrezept** Unsere Beispiele illustrieren das folgende, allgemeine Prinzip für  $V = \mathbb{R}^n$ : Wir können die Übergangsmatrizen  $\mathbf{S}$  und  $\tilde{\mathbf{S}}$  zum Beispiel dadurch bestimmen, dass wir das Gauß-Jordan-Verfahren auf die erweiterten Koeffizienten-Matrizen

$$\left( \begin{array}{ccc|ccc} | & | & | & | & | & | \\ \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 & \tilde{\mathbf{v}}_1 & \tilde{\mathbf{v}}_2 & \tilde{\mathbf{v}}_3 \\ | & | & | & | & | & | \end{array} \right) \quad \text{und} \quad \left( \begin{array}{ccc|ccc} | & | & | & | & | & | \\ \tilde{\mathbf{v}}_1 & \tilde{\mathbf{v}}_2 & \tilde{\mathbf{v}}_3 & \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \\ | & | & | & | & | & | \end{array} \right)$$

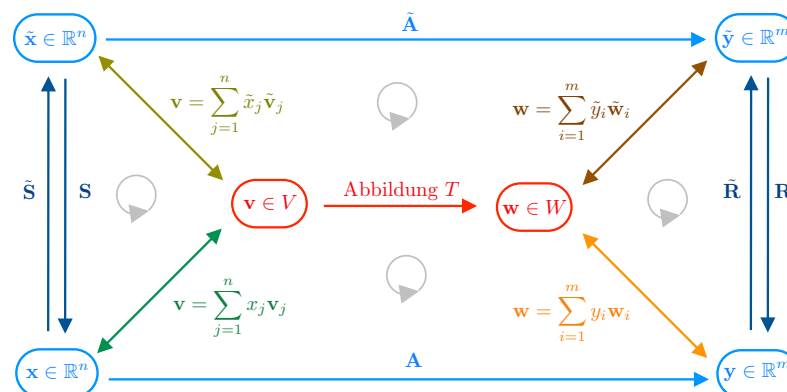
anwenden. Wenn wir  $\mathbf{S}$  oder  $\tilde{\mathbf{S}}$  bereits kennen, so können wir die jeweils andere auch als inverse Matrix (zum Beispiel wieder mit einer Variante des Gauß-Jordan-Verfahrens) bestimmen. Ein wichtiger Spezialfall entsteht, wenn jeder Basisvektor  $\mathbf{v}_j$  der  $j$ -te Einheitsvektor ist: In diesem Fall kann die Matrix  $\mathbf{S}$  direkt aus den Vektoren  $\tilde{\mathbf{v}}_j$  zusammengesetzt werden.



**Basiswechsel für lineare Abbildungen** Sei  $T : V \rightarrow W$  eine lineare Abbildung und sei  $\mathbf{A}$  ihre Basisdarstellung bzgl. zweier Basen  $(\mathbf{v}_1, \dots, \mathbf{v}_n)$  und  $(\mathbf{w}_1, \dots, \mathbf{w}_m)$  von  $V$  und  $W$ . Sind nun  $(\tilde{\mathbf{v}}_1, \dots, \tilde{\mathbf{v}}_n)$  und  $(\tilde{\mathbf{w}}_1, \dots, \tilde{\mathbf{w}}_m)$  zwei andere Basen in  $V$  und  $W$  und  $\tilde{\mathbf{A}}$  die entsprechende Basisdarstellung von  $T$ , so gilt

$$\tilde{\mathbf{A}} = \tilde{\mathbf{R}} \cdot \mathbf{A} \cdot \mathbf{S} \quad \text{sowie} \quad \mathbf{A} = \mathbf{R} \cdot \tilde{\mathbf{A}} \cdot \tilde{\mathbf{S}},$$

wobei  $\mathbf{S}$  und  $\tilde{\mathbf{S}}$  bzw.  $\mathbf{R}$  und  $\tilde{\mathbf{R}}$  die Basiswechsel in  $V$  und  $W$  beschreiben (siehe Bild). Diese Formeln ergeben sich unmittelbar aus den bisher abgeleiteten Resultaten in diesem Abschnitt.



Flussdiagramm zu Basiswechsel und Basisdarstellung bei linearen Abbildungen. Beachte, dass weder  $T$  noch  $\mathbf{A}$  oder  $\tilde{\mathbf{A}}$  invertierbar sein müssen, aber dass immer  $\tilde{\mathbf{S}} = \mathbf{S}^{-1}$  und  $\tilde{\mathbf{R}} = \mathbf{R}^{-1}$  gilt.

**Beispiel** Wir betrachten den einfachen Fall  $V = W = \mathbb{R}^2$  sowie die lineare Abbildung

$$\mathbf{w} = T(\mathbf{v}) \quad \text{mit} \quad \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \begin{pmatrix} a_{11} \cdot v_1 + a_{12} \cdot v_2 \\ a_{21} \cdot v_1 + a_{22} \cdot v_2 \end{pmatrix}$$

mit gegebenen Koeffizienten  $a_{ij} \in \mathbb{R}$ . Wir wollen der Einfachheit halber auch jeweils eine Basis in  $V$  und  $W$  kanonisch wählen, d.h. wir setzen

$$\mathbf{v}_1 = \mathbf{w}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \mathbf{v}_2 = \mathbf{w}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

Damit kann die entsprechende Basisdarstellung  $\mathbf{A}$  von  $T$  (unterer Pfeil im Bild) direkt als

$$\mathbf{A} = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

abgelesen werden und wir wollen die Matrix  $\tilde{\mathbf{A}}$  (oberer Pfeil im Bild) für zwei Wahlen der jeweils anderen Basen berechnen.

Basiswechsel 1: Wir betrachten sowohl im Bild- als auch im Urbildraum jeweils die Basis

$$\tilde{\mathbf{v}}_1 := \tilde{\mathbf{w}}_1 := \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad \tilde{\mathbf{v}}_2 := \tilde{\mathbf{w}}_2 := \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

und bemerken zunächst, dass wegen

$$\tilde{\mathbf{v}}_1 = 1 \cdot \mathbf{v}_1 + 1 \cdot \mathbf{v}_2, \quad \tilde{\mathbf{v}}_2 = -1 \cdot \mathbf{v}_1 + 1 \cdot \mathbf{v}_2$$

bzw.

$$\mathbf{v}_1 = \frac{1}{2} \cdot \tilde{\mathbf{v}}_1 - \frac{1}{2} \cdot \tilde{\mathbf{v}}_2, \quad \mathbf{v}_2 = \frac{1}{2} \cdot \tilde{\mathbf{v}}_1 + \frac{1}{2} \cdot \tilde{\mathbf{v}}_2$$

die entsprechenden Übergangsmatrizen und ihre Inverse spaltenweise als

$$\mathbf{S} = \mathbf{R} = \begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}, \quad \tilde{\mathbf{S}} = \tilde{\mathbf{R}} = \mathbf{S}^{-1} = \mathbf{R}^{-1} = \frac{1}{2} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$$

identifiziert werden können. Mit den Formeln für die Basiswechsel-Matrizen ergibt sich insbesondere

$$\begin{aligned} \tilde{\mathbf{A}} &= \tilde{\mathbf{S}} \cdot \mathbf{A} \cdot \mathbf{S} = \tilde{\mathbf{S}} \cdot \begin{pmatrix} a_{11} + a_{12} & -a_{11} + a_{12} \\ a_{21} + a_{22} & -a_{21} + a_{22} \end{pmatrix} \\ &= \frac{1}{2} \begin{pmatrix} a_{11} + a_{12} + a_{21} + a_{22} & -a_{11} + a_{12} - a_{21} + a_{22} \\ -a_{11} - a_{12} + a_{21} + a_{22} & a_{11} - a_{12} - a_{21} + a_{22} \end{pmatrix}. \end{aligned}$$

Wir wollen nun abschließend eben hergeleitete Formel alternativ begründen. Dazu berechnen wir

$$T(\tilde{\mathbf{v}}_1) = \begin{pmatrix} a_{11} + a_{12} \\ a_{21} + a_{22} \end{pmatrix} = \frac{a_{11} + a_{12} + a_{21} + a_{22}}{2} \cdot \tilde{\mathbf{w}}_1 + \frac{-a_{11} - a_{12} + a_{21} + a_{22}}{2} \cdot \tilde{\mathbf{w}}_2$$

sowie

$$T(\tilde{\mathbf{v}}_2) = \begin{pmatrix} -a_{11} + a_{12} \\ -a_{21} + a_{22} \end{pmatrix} = \frac{-a_{11} + a_{12} - a_{21} + a_{22}}{2} \cdot \tilde{\mathbf{w}}_1 + \frac{a_{11} - a_{12} - a_{21} + a_{22}}{2} \cdot \tilde{\mathbf{w}}_2,$$

und können aus die Formeln die Matrix  $\tilde{\mathbf{A}}$  spaltenweise und analog zum ersten Theorem in diesem Abschnitts berechnen. Wir erhalten dabei wieder das gleiche Ergebnis. Man kann  $\tilde{\mathbf{A}}$  schließlich auch auf eine dritte Art, nämlich durch das oben beschriebene Kochrezept berechnen. Dazu wendet man das Gauß-Jordan-Verfahrens auf die folgende erweiterte Koeffizientenmatrix an:

$$\left( \begin{array}{cc|cc} 1 & -1 & a_{11} + a_{12} & -a_{11} + a_{12} \\ 1 & 1 & a_{21} + a_{22} & -a_{21} + a_{22} \end{array} \right)$$

Basiswechsel 2: Mit der alternativen Wahl

$$\tilde{\mathbf{v}}_1 = \tilde{\mathbf{w}}_1 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \tilde{\mathbf{v}}_2 = \tilde{\mathbf{w}}_2 = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$$

können wir auch wieder die Matrix  $\mathbf{A}$  auf zwei verschiedene Weisen berechnen. Wir beginnen diesmal mit dem direkten Weg ohne Basiswechsel und lesen aus

$$T(\tilde{\mathbf{v}}_1) = \begin{pmatrix} a_{12} \\ a_{22} \end{pmatrix} = a_{22} \cdot \tilde{\mathbf{w}}_1 - a_{12} \cdot \tilde{\mathbf{w}}_2$$

sowie

$$T(\tilde{\mathbf{v}}_2) = \begin{pmatrix} -a_{11} \\ -a_{21} \end{pmatrix} = -a_{21} \cdot \tilde{\mathbf{w}}_1 + a_{11} \tilde{\mathbf{w}}_2$$

spaltenweise die Formel

$$\tilde{\mathbf{A}} = \begin{pmatrix} a_{22} & -a_{21} \\ a_{-12} & a_{11} \end{pmatrix}$$

ablesen. Andererseits gilt

$$\tilde{\mathbf{v}}_1 = 0 \cdot \mathbf{v}_1 + 1 \cdot \mathbf{v}_2, \quad \tilde{\mathbf{v}}_2 = -1 \cdot \mathbf{v}_1 + 0 \cdot \mathbf{v}_2$$

bzw.

$$\mathbf{v}_1 = 0 \cdot \tilde{\mathbf{v}}_1 - 1 \cdot \tilde{\mathbf{v}}_2, \quad \mathbf{v}_2 = 1 \cdot \tilde{\mathbf{v}}_1 - 0 \cdot \tilde{\mathbf{v}}_2$$

und damit auch

$$\mathbf{S} = \mathbf{R} = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \quad \text{bzw.} \quad \tilde{\mathbf{S}} = \tilde{\mathbf{R}} = \mathbf{S}^{-1} = \tilde{\mathbf{R}}^{-1} \tilde{\mathbf{S}}^{-1} = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}.$$

Wir können nun leicht nachrechnen, dass auch wieder  $\tilde{\mathbf{A}} = \tilde{\mathbf{S}} \cdot \mathbf{A} \cdot \mathbf{S}$  gilt. Dasselbe Ergebnis kann man drittens auch mittels der erweiterten Koeffizientenmatrix und dem

$$\left( \begin{array}{cc|cc} 0 & -1 & a_{12} & -a_{11} \\ 1 & 0 & a_{22} & -a_{21} \end{array} \right)$$

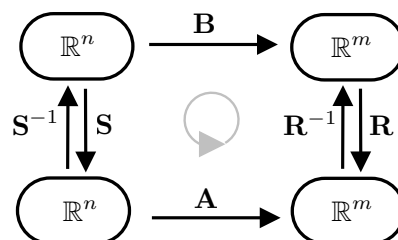
Gauß-Jordan-Verfahren erreichen.

*Bemerkung:* Analog zeigt man: Bei Vertauschungen von Basisvektoren müssen die Einträge in der Basisdarstellung  $\mathbf{A}$  entsprechend vertauscht werden. Wir werden das später beim Studium von Permutationsmatrizen besser verstehen.

**Definition** Zwei  $(m, n)$ -Matrizen  $\mathbf{A}$  und  $\mathbf{B}$  heißen äquivalent, falls es eine invertierbare  $(n, n)$ -Matrix  $\mathbf{S}$  sowie eine invertierbare  $(m, m)$ -Matrix  $\mathbf{R}$  gibt, so dass

$$\mathbf{B} = \mathbf{R}^{-1} \cdot \mathbf{A} \cdot \mathbf{S} \quad \text{bzw.} \quad \mathbf{A} = \mathbf{R} \cdot \mathbf{B} \cdot \mathbf{S}^{-1}$$

gilt, wobei die jeweils eine Gleichung die andere impliziert (Übungsaufgabe). Im Fall von  $m = n$  und  $\mathbf{R} = \mathbf{S}$  nennt man die (dann quadratischen) Matrizen  $\mathbf{A}$  und  $\mathbf{B}$  auch ähnlich.



Zur Äquivalenz und Ähnlichkeit von Matrizen.

**Bemerkung** Bei ähnlichen Matrizen gilt

$$\det \mathbf{B} = \det (\mathbf{S}^{-1} \cdot \mathbf{A} \cdot \mathbf{S}) = (\det \mathbf{S}^{-1}) \cdot (\det \mathbf{A}) \cdot (\det \mathbf{S}) = \det \mathbf{A},$$

weil mit  $\mathbf{S}^{-1} \cdot \mathbf{S}$  auch  $(\det \mathbf{S}^{-1}) \cdot \det \mathbf{S} = 1$  erfüllt ist

**Beispiele** Mit

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 0 \\ 1 & 0 & 2 \\ 0 & 2 & 1 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 3 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

und

$$\mathbf{S}^{-1} = \mathbf{R}^{-1} = \frac{1}{15} \begin{pmatrix} 3 & 6 & 6 \\ 1 & -3 & 2 \\ -5 & 0 & 5 \end{pmatrix} \quad \mathbf{S} = \mathbf{R} = \begin{pmatrix} 1 & 2 & -2 \\ 1 & -3 & 0 \\ 1 & 2 & 1 \end{pmatrix}$$

gilt  $\mathbf{A} = \mathbf{R} \cdot \mathbf{B} \cdot \mathbf{R}^{-1}$ . Insbesondere sind die 3,  $-2$ , und 1 die sogenannten *Eigenwerte* (siehe nächstes Kapitel) der zu einander ähnlichen Matrizen  $\mathbf{A}$  und  $\mathbf{B}$ .

**Bemerkung** Ähnliche bzw. äquivalente Matrizen beschreiben letztlich dieselbe lineare Abbildung, nur bzgl. jeweils anderer Basen. Eine naheliegende und wichtige Frage ist nun, welche Eigenschaften einer Matrix sich nicht beim Wechsel von Basen ändern (*Invariantentheorie*) bzw. ob man immer geschickt Basen so wählen kann, dass die entsprechende Basisdarstellung möglichst einfach wird (*Theorie der Normalformen*). Wir werden beide Aspekte weiter untersuchen und insbesondere die Klasse der zugelassenen Übergangsmatrizen  $\mathbf{R}$  und  $\mathbf{S}$  weiter einschränken.

## 5.2 Orthogonalität und Gram-Schmidt-Verfahren

**Setting für diesen Abschnitt**

1.  $V$  ist reeller Vektorraum der Dimension  $n \in \mathbb{N}$  und besitzt ein Skalarprodukt  $\langle \cdot, \cdot \rangle$  sowie eine entsprechende Norm  $\|\cdot\|$ .
2.  $\mathbf{v}_1, \dots, \mathbf{v}_n$  ist gegebene Basis von  $V$ .
3.  $W$  ist  $m$ -dimensionaler Unterraum von  $V$ , der von den linear unabhängigen Vektoren  $\mathbf{w}_1, \dots, \mathbf{w}_m$  aufgespannt wird.

**Klarstellung** Alle Aussagen in diesem Abschnitt gelten zunächst nur für reelle Matrizen. Im Komplexen gelten analoge Formeln und Ergebnisse, aber man muss berücksichtigen, dass Skalarprodukte dann nicht mehr bilinear, sondern sesquilinear sind und deshalb an einigen Stellen konjugiert komplexe Zahlen auftreten werden.

**Erinnerung** Die Basis  $(\mathbf{w}_1, \dots, \mathbf{w}_m)$  heißt Orthogonalbasis (O-Basis) in  $W$ , falls

$$\langle \mathbf{w}_l, \mathbf{w}_i \rangle = 0 \quad \text{für alle } l, i = 1, \dots, m \text{ mit } l \neq i$$

bzw. Orthonormalbasis (ON-Basis), falls zusätzlich

$$\langle \mathbf{w}_i, \mathbf{w}_i \rangle = 1 \quad \text{für alle } i = 1, \dots, m$$

gilt. Man spricht alternativ auch von einem O-System bzw. einem ON-System in  $V$  (denn dort sind die  $\mathbf{w}_i$  immer noch linear unabhängig und orthogonal, aber für  $m < n$  keine Basis mehr).

**Lemma (Fourier-Entwicklung und Satz von Pythagoras)** Bilden die  $\mathbf{w}_i$  eine ON-Basis in  $W$ , so gilt

$$\mathbf{w} = \sum_{i=1}^m \langle \mathbf{w}, \mathbf{w}_i \rangle \cdot \mathbf{w}_i, \quad \|\mathbf{w}\|^2 = \sum_{i=1}^m \langle \mathbf{w}, \mathbf{w}_i \rangle^2$$

für jedes  $\mathbf{w} \in W$ .

*Beweis:* Der Ansatz

$$\mathbf{w} = \sum_{l=1}^m y_l \cdot \mathbf{w}_l$$

impliziert

$$\langle \mathbf{w}_i, \mathbf{w} \rangle = \left\langle \mathbf{w}_i, \sum_{l=1}^m y_l \cdot \mathbf{w}_l \right\rangle = \sum_{l=1}^m y_l \cdot \langle \mathbf{w}_i, \mathbf{w}_l \rangle = \sum_{l=1}^m y_l \cdot \delta_{il} = y_i,$$

wobei  $\delta_{il}$  wieder das Kronecker-Delta ist und nur für  $i = l$  eine 1, sonst immer eine 0 liefert. Analog folgt

$$\langle \mathbf{w}, \mathbf{w} \rangle = \left\langle \sum_{l=1}^m y_l \cdot \mathbf{w}_l, \sum_{i=1}^m y_i \cdot \mathbf{w}_i \right\rangle = \sum_{i=1}^m \sum_{j=1}^m y_i \cdot y_j \cdot \delta_{ij} = \sum_{i=1}^m y_i^2.$$

□

## Gram-Schmidtsches Orthogonalisierungs-Verfahren

**Idee** Mit einfachen Algorithmen kann aus einer gegebenen Basis  $\mathbf{w}_1, \dots, \mathbf{w}_m$  von  $W$  rekursiv eine orthogonale bzw. eine orthonormale Basis  $\tilde{\mathbf{w}}_1, \dots, \tilde{\mathbf{w}}_n$  erzeugt werden.

### O-Variante

1. Setze im ersten Schritt  $\tilde{\mathbf{w}}_1 := \mathbf{w}_1$ .
2. Berechne im  $i$ -ten Schritt ( $i = 2, \dots, m$ ) den Vektor  $\tilde{\mathbf{w}}_i$  via

$$\tilde{\mathbf{w}}_i := \mathbf{w}_i - \sum_{l=1}^{i-1} c_{il} \cdot \tilde{\mathbf{w}}_l, \quad c_{il} := \frac{\langle \mathbf{w}_i, \tilde{\mathbf{w}}_l \rangle}{\langle \tilde{\mathbf{w}}_l, \tilde{\mathbf{w}}_l \rangle}$$

aus  $\mathbf{w}_i$  sowie den Ergebnissen der vorangegangenen Schritte.

**Eigenschaften** Dieses Verfahren erzeugt aus den folgenden Gründen sukzessive eine ON-Basis:

1. Der Algorithmus impliziert

$$\tilde{\mathbf{w}}_2 = \mathbf{w}_2 - c_{21} \cdot \tilde{\mathbf{w}}_1 \quad \text{bzw.} \quad \tilde{\mathbf{w}}_3 = \mathbf{w}_3 - c_{31} \cdot \tilde{\mathbf{w}}_1 - c_{32} \cdot \tilde{\mathbf{w}}_2$$

für den zweiten bzw. den dritten Schritt, wobei die „Korrekturkoeffizienten“

$$c_{21} = \frac{\langle \mathbf{w}_2, \tilde{\mathbf{w}}_1 \rangle}{\langle \tilde{\mathbf{w}}_1, \tilde{\mathbf{w}}_1 \rangle} \quad \text{bzw.} \quad c_{31} = \frac{\langle \mathbf{w}_3, \tilde{\mathbf{w}}_1 \rangle}{\langle \tilde{\mathbf{w}}_1, \tilde{\mathbf{w}}_1 \rangle}, \quad c_{32} = \frac{\langle \mathbf{w}_3, \tilde{\mathbf{w}}_2 \rangle}{\langle \tilde{\mathbf{w}}_2, \tilde{\mathbf{w}}_2 \rangle}$$

gerade so gewählt sind, dass  $\tilde{\mathbf{w}}_2$  aus  $\tilde{\mathbf{w}}_1$  senkrecht steht bzw. das  $\tilde{\mathbf{w}}_3$  auf  $\tilde{\mathbf{w}}_1$  und  $\tilde{\mathbf{w}}_2$  senkrecht steht, d.h.

$$\langle \tilde{\mathbf{w}}_2, \tilde{\mathbf{w}}_1 \rangle = 0 \quad \text{bzw.} \quad \langle \tilde{\mathbf{w}}_3, \tilde{\mathbf{w}}_1 \rangle = 0 \quad \text{und} \quad \langle \tilde{\mathbf{w}}_3, \tilde{\mathbf{w}}_2 \rangle = 0.$$

2. Ganz allgemein kann man die Implikation

$$\tilde{\mathbf{w}}_1, \dots, \tilde{\mathbf{w}}_{i-1} \text{ sind O-System} \quad \implies \quad \tilde{\mathbf{w}}_1, \dots, \tilde{\mathbf{w}}_i \text{ sind O-System}$$

zeigen. In der Tat, mit  $1 \leq k \leq i-1$  gilt nach Konstruktion

$$\begin{aligned} \langle \tilde{\mathbf{w}}_k, \tilde{\mathbf{w}}_i \rangle &= \langle \tilde{\mathbf{w}}_k, \mathbf{w}_i \rangle - \left\langle \tilde{\mathbf{w}}_k, \sum_{l=1}^{i-1} \frac{\langle \mathbf{w}_i, \tilde{\mathbf{w}}_l \rangle}{\langle \tilde{\mathbf{w}}_l, \tilde{\mathbf{w}}_l \rangle} \cdot \tilde{\mathbf{w}}_l \right\rangle \\ &= \langle \tilde{\mathbf{w}}_k, \mathbf{w}_i \rangle - \sum_{l=1}^{i-1} \frac{\langle \mathbf{w}_i, \tilde{\mathbf{w}}_l \rangle}{\langle \tilde{\mathbf{w}}_l, \tilde{\mathbf{w}}_l \rangle} \cdot \langle \tilde{\mathbf{w}}_k, \tilde{\mathbf{w}}_l \rangle \\ &= \langle \tilde{\mathbf{w}}_k, \mathbf{w}_i \rangle - \frac{\langle \mathbf{w}_i, \tilde{\mathbf{w}}_k \rangle}{\langle \tilde{\mathbf{w}}_k, \tilde{\mathbf{w}}_k \rangle} \cdot \langle \tilde{\mathbf{w}}_k, \tilde{\mathbf{w}}_k \rangle \\ &= 0, \end{aligned}$$

wobei wir benutzt haben, dass

$$\langle \tilde{\mathbf{w}}_k, \tilde{\mathbf{w}}_l \rangle = \delta_{kl} \cdot \langle \tilde{\mathbf{w}}_k, \tilde{\mathbf{w}}_k \rangle$$

nach Induktionsvoraussetzung gilt. Wir schließen, dass der im  $i$ -ten Schritt erzeugte neue Basisvektor  $\tilde{\mathbf{w}}_i$  in der Tat auf allen vorher konstruierten senkrecht steht (eben weil wir die richtigen Korrektorkoeffizienten verwenden).

3. Nach Konstruktion gilt

$$\tilde{\mathbf{w}}_i \in \text{span}\{\mathbf{w}_1, \dots, \mathbf{w}_i\}$$

für alle  $i = 1, \dots, m$  und rekursiv über  $i$  zeigt man

$$\tilde{\mathbf{w}}_i \notin \text{span}\{\tilde{\mathbf{w}}_1, \dots, \tilde{\mathbf{w}}_{i-1}\}$$

mit Hilfe der Basis-Eigenschaft der  $\mathbf{w}_i$  (Übungsaufgabe für Fortgeschrittene). Insbesondere kann  $\tilde{\mathbf{w}}_i$  niemals der Nullvektor sein und das Verfahren ist wohldefiniert und liefert nach  $m$  Schritten wieder eine Basis von  $W$ .

**Bemerkung** Der Ausdruck

$$\frac{\langle \mathbf{w}, \tilde{\mathbf{w}} \rangle}{\langle \tilde{\mathbf{w}}, \tilde{\mathbf{w}} \rangle} \cdot \tilde{\mathbf{w}} = \left\langle \mathbf{w}, \frac{\tilde{\mathbf{w}}}{\|\tilde{\mathbf{w}}\|} \right\rangle \cdot \frac{\tilde{\mathbf{w}}}{\|\tilde{\mathbf{w}}\|}$$

ist gerade der Anteil von  $\mathbf{w}$  in Richtung von  $\tilde{\mathbf{w}}$  (auch Projektion von  $\mathbf{w}$  auf  $\text{span}\{\tilde{\mathbf{w}}\}$  genannt).

**Beispiel** Für  $n = 4$  und  $m = 3$  sei

$$\mathbf{w}_1 = \begin{pmatrix} 2 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{w}_2 = \begin{pmatrix} 3 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \quad \mathbf{w}_3 = \begin{pmatrix} 0 \\ 2 \\ 4 \\ 2 \end{pmatrix}$$

die gegebene Basis von  $W \subset V = \mathbb{R}^4$ . Die ersten beiden Schritte im Algorithmus liefern

$$\tilde{\mathbf{w}}_1 = \begin{pmatrix} 2 \\ 1 \\ 1 \\ 0 \end{pmatrix}, \quad \tilde{\mathbf{w}}_2 = \begin{pmatrix} 3 \\ 0 \\ 0 \\ 1 \end{pmatrix} - c_{21} \cdot \begin{pmatrix} 2 \\ 1 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix}$$

mit dem Zwischenergebnis

$$c_{21} = \frac{3 \cdot 2 + 0 \cdot 1 + 0 \cdot 1 + 1 \cdot 0}{2^2 + 1^2 + 1^2 + 0^2} = 1.$$

Im dritten und letzten Schritt erhalten wir

$$\tilde{\mathbf{w}}_3 = \begin{pmatrix} 0 \\ 2 \\ 4 \\ 2 \end{pmatrix} - c_{31} \cdot \begin{pmatrix} 2 \\ 1 \\ 1 \\ 0 \end{pmatrix} - c_{32} \cdot \begin{pmatrix} 1 \\ -1 \\ -1 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ 2 \\ 3 \end{pmatrix},$$

wobei

$$c_{31} = \frac{0 \cdot 2 + 2 \cdot 1 + 4 \cdot 1 + 2 \cdot 0}{2^2 + 1^2 + 1^2 + 0^2} = 1, \quad c_{32} = \frac{0 \cdot 1 - 2 \cdot 1 - 4 \cdot 1 + 2 \cdot 1}{1^2 + 1^2 + 1^2 + 1^2} = -1.$$

**Bemerkung** Hat man eine O-Basis  $\tilde{\mathbf{w}}_1, \dots, \tilde{\mathbf{w}}_m$  berechnet, so kann man durch abschließende Normierung aller Vektoren via

$$\hat{\mathbf{w}}_i := \frac{\tilde{\mathbf{w}}_i}{\|\tilde{\mathbf{w}}_i\|}, \quad i = 1, \dots, m$$

auch eine entsprechende ON-Basis  $\hat{\mathbf{w}}_1, \dots, \hat{\mathbf{w}}_m$  erzeugen. Man kann die Normierung aber auch gleich in jedem Zwischenschritt ausführen und erhält die folgende Variante des Algorithmus.

### ON-Variante

1. Setze im ersten Schritt  $\hat{\mathbf{w}}_1 := \frac{1}{\|\mathbf{w}_1\|} \cdot \mathbf{w}_1$ .

2. Berechne im  $i$ -ten Schritt

$$\hat{\mathbf{w}}_i := \frac{1}{\|\tilde{\mathbf{w}}_i\|} \cdot \tilde{\mathbf{w}}_i, \quad \tilde{\mathbf{w}}_i := \mathbf{w}_i - \sum_{l=1}^{i-1} d_{il} \cdot \hat{\mathbf{w}}_l, \quad d_{il} := \langle \mathbf{w}_i, \hat{\mathbf{w}}_l \rangle$$

**Beispiel** Wir betrachten  $W = \mathbb{R}^3$  mit der gegebenen Basis

$$\mathbf{w}_1 = \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix}, \quad \mathbf{w}_2 = \begin{pmatrix} 4 \\ -3 \\ 4 \end{pmatrix}, \quad \mathbf{w}_3 = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}.$$

Im ersten Schritt ergibt sich

$$\hat{\mathbf{w}}_1 = \frac{1}{\sqrt{0^2 + 2^2 + 0^2}} \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

und anschließend erhalten wir

$$\hat{\mathbf{w}}_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad \hat{\mathbf{w}}_3 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$$

aus den Zwischenresultaten

$$\tilde{\mathbf{w}}_2 = \begin{pmatrix} 4 \\ -3 \\ 4 \end{pmatrix} - d_{21} \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 4 \\ 0 \\ 4 \end{pmatrix}, \quad d_{21} = 4 \cdot 0 - 3 \cdot 1 + 4 \cdot 0 = -3$$

und

$$\tilde{\mathbf{w}}_3 = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix} - d_{31} \cdot \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} - \frac{d_{32}}{\sqrt{2}} \cdot \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}$$

mit

$$d_{31} = 3 \cdot 0 + 2 \cdot 1 + 1 \cdot 0 = 2, \quad d_{32} = \frac{3 \cdot 1 + 2 \cdot 0 + 1 \cdot 1}{\sqrt{2}} = 2\sqrt{2}.$$

**Bemerkung** Wir können die Verfahren auch auf die Basis  $\mathbf{v}_1, \dots, \mathbf{v}_n$  anwenden, um in  $n$  Schritten eine O-Basis oder ON-Basis  $\tilde{\mathbf{v}}_1, \dots, \tilde{\mathbf{v}}_n$  zu erzeugen. Außerdem überlegt man sich leicht die Implikation

$$\mathbf{w}_1 = \mathbf{v}_1, \quad \dots, \quad \mathbf{w}_m = \mathbf{v}_m \quad \Longrightarrow \quad \tilde{\mathbf{w}}_1 = \tilde{\mathbf{v}}_1 \quad \dots, \quad \tilde{\mathbf{w}}_m = \tilde{\mathbf{v}}_m$$

Vorlesung 16, 25. November 2019

## Orthogonalprojektionen

**Komplemente von Unterräumen** Die Menge

$$W^\perp := \{v \in V : \langle v, w \rangle = 0 \text{ für alle } w \in W\}$$

ist auch ein linearer Unterraum von  $V$  (Übungsaufgabe), besitzt die Dimension  $n - m$  und wird das orthogonale Komplement von  $W$  genannt. Außerdem zeigt man leicht

$$(W^\perp)^\perp = W.$$

**Beispiel** Für  $V = \mathbb{R}^3$  und  $W = \text{span}\{\mathbf{w}_1, \mathbf{w}_2\}$  gilt

$$W^\perp = \text{span}\{\mathbf{w}_1 \times \mathbf{w}_2\},$$

sofern  $\mathbf{w}_1$  und  $\mathbf{w}_2$  linear unabhängig sind.



**Theorem (Projektionssatz)** Für jedes  $\mathbf{v} \in V$  existiert eine eindeutige Zerlegung

$$\mathbf{v} = \mathbf{v}^{\parallel} + \mathbf{v}^{\perp} \quad \text{mit} \quad \mathbf{v}^{\parallel} \in W, \quad \mathbf{v}^{\perp} \in W^{\perp}.$$

Die Vektoren  $\mathbf{v}^{\parallel}$  und  $\mathbf{v}^{\perp}$  werden die Orthogonalprojektionen von  $\mathbf{v}$  auf  $W$  und  $W^{\perp}$  genannt.

Beweis: Nach einem geeigneten Basiswechsel können wir (Übungsaufgabe)

$$W = \text{span}\{\mathbf{v}_1, \dots, \mathbf{v}_m\}$$

annehmen und durch eine Anwendung des Gram-Schmidtschen Verfahrens außerdem sicherstellen, dass  $\mathbf{v}_1, \dots, \mathbf{v}_n$  sogar eine ON-Basis in  $V$  ist. Dies impliziert dann automatisch (wieder Übungsaufgabe)

$$W^{\perp} = \text{span}\{\mathbf{v}_{m+1}, \dots, \mathbf{v}_n\}.$$

Mit Hilfe der Fourier-Entwicklung

$$\mathbf{v} = \sum_{j=1}^n \langle \mathbf{v}, \mathbf{v}_j \rangle \cdot \mathbf{v}_j$$

sehen wir nun leicht, dass

$$\mathbf{v}^{\parallel} := \sum_{j=1}^m \langle \mathbf{v}, \mathbf{v}_j \rangle \cdot \mathbf{v}_j, \quad \mathbf{v}^{\perp} := \sum_{j=m+1}^n \langle \mathbf{v}, \mathbf{v}_j \rangle \cdot \mathbf{v}_j$$

die gewünschten Eigenschaften besitzen. Um auch noch die Eindeutigkeit dieser Formeln zu zeigen, nehmen wir an, dass zusätzlich

$$\mathbf{v} = \mathbf{u}^{\parallel} + \mathbf{u}^{\perp}, \quad \mathbf{u}^{\parallel} \in W, \quad \mathbf{u}^{\perp} \in W^{\perp}$$

erfüllt ist und erhalten

$$0 = \langle \mathbf{v}_j, \mathbf{v} - \mathbf{v} \rangle = \langle \mathbf{v}_j, \mathbf{v}^{\parallel} + \mathbf{v}^{\perp} - \mathbf{u}^{\parallel} - \mathbf{u}^{\perp} \rangle = \langle \mathbf{v}_j, \mathbf{v}^{\parallel} \rangle - \langle \mathbf{v}_j, \mathbf{u}^{\parallel} \rangle$$

für jedes  $j = 1, \dots, m$ . Insbesondere gilt

$$\mathbf{v}^{\parallel} = \sum_{j=1}^m \langle \mathbf{v}_j, \mathbf{v}^{\parallel} \rangle \cdot \mathbf{v}_j = \sum_{j=1}^m \langle \mathbf{v}_j, \mathbf{u}^{\parallel} \rangle \cdot \mathbf{u}_j = \mathbf{u}^{\parallel},$$

und damit auch  $\mathbf{v}^{\perp} = \mathbf{v} - \mathbf{v}^{\parallel} = \mathbf{v} - \mathbf{u}^{\parallel} = \mathbf{u}^{\perp}$ . □

**Bemerkung** Die Projektionen  $\mathbf{v}^{\parallel}$  und  $\mathbf{v}^{\perp}$  existieren unabhängig von der gewählten Basis und können im Prinzip mit jeder Basis berechnet werden. Bei nicht-orthonormalen Basen sind die entsprechenden Formeln aber deutlich komplizierter als die im Beweis angegebenen.

**Theorem (Approximationssatz)** Es gilt

$$\|\mathbf{v} - \mathbf{v}^\perp\| \leq \|\mathbf{v} - \mathbf{w}\| \quad \text{für alle } \mathbf{w} \in W$$

sowie

$$\|\mathbf{v} - \mathbf{v}^\perp\| \leq \|\mathbf{v} - \mathbf{u}\| \quad \text{für alle } \mathbf{u} \in W^\perp.$$

*Beweis* Wir betrachten wieder eine ON-Basis von  $V$  wie im Beweis des Projektionssatzes und bemerken, dass

$$\|\mathbf{v} - \mathbf{v}^\perp\|^2 = \|\mathbf{v}^\perp\|^2 = \sum_{j=m+1}^n \langle \mathbf{v}, \mathbf{v}_j \rangle^2, \quad \|\mathbf{v} - \mathbf{v}^\perp\|^2 = \|\mathbf{v}^\perp\|^2 = \sum_{j=1}^m \langle \mathbf{v}, \mathbf{v}_j \rangle^2.$$

Andererseits gilt wegen  $\mathbf{w} \in W$  und mit der Fourier-Entwicklung

$$\begin{aligned} \|\mathbf{v} - \mathbf{w}\|^2 &= \sum_{j=1}^m \langle \mathbf{v} - \mathbf{w}, \mathbf{v}_j \rangle^2 + \sum_{j=m+1}^n (\langle \mathbf{v}, \mathbf{v}_j \rangle - \langle \mathbf{w}, \mathbf{v}_j \rangle)^2 \\ &= \sum_{j=1}^m \langle \mathbf{v} - \mathbf{w}, \mathbf{v}_j \rangle^2 + \sum_{j=m+1}^n (\langle \mathbf{v}, \mathbf{v}_j \rangle - 0)^2 \\ &= \sum_{j=1}^m \langle \mathbf{v} - \mathbf{w}, \mathbf{v}_j \rangle^2 + \|\mathbf{v} - \mathbf{v}^\perp\|^2 \\ &\geq \|\mathbf{v} - \mathbf{v}^\perp\|^2. \end{aligned}$$

Analog ergibt sich

$$\begin{aligned} \|\mathbf{v} - \mathbf{u}\|^2 &= \sum_{j=1}^m (\langle \mathbf{v}, \mathbf{v}_j \rangle - 0)^2 + \sum_{j=m+1}^n \langle \mathbf{v} - \mathbf{u}, \mathbf{v}_j \rangle^2 \\ &= \|\mathbf{v} - \mathbf{v}^\perp\|^2 + \sum_{j=m+1}^n \langle \mathbf{v} - \mathbf{u}, \mathbf{v}_j \rangle^2 \\ &\geq \|\mathbf{v} - \mathbf{v}^\perp\|^2 \end{aligned}$$

und die beiden Behauptungen folgen nach Wurzelziehen.

## 5.3 Orthogonale Matrizen

*Erinnerung*

1. Eine reelle Matrix  $\mathbf{A} \in \mathbb{R}^{(m,m)}$  heißt orthogonal, falls  $\mathbf{A} \cdot \mathbf{A}^T = \mathbf{A}^T \cdot \mathbf{A} = \mathbf{I}_n$  gilt, d.h. wenn  $\mathbf{A}^T = \mathbf{A}^{-1}$ .
2. Das komplexe Analogon von orthogonal ist unitär, d.h.  $\mathbf{A} \cdot \mathbf{A}^* = \mathbf{A}^* \cdot \mathbf{A} = \mathbf{I}_n$  mit adjungierter Matrix  $\mathbf{A}^* = \overline{\mathbf{A}}^T$ .

**Nützliche Formeln**

1. Für alle  $\mathbf{x}, \tilde{\mathbf{x}} \in \mathbb{R}^n$  gilt

$$\mathbf{x}^T \cdot \tilde{\mathbf{x}} = \langle \mathbf{x}, \tilde{\mathbf{x}} \rangle = \langle \tilde{\mathbf{x}}, \mathbf{x} \rangle = \tilde{\mathbf{x}}^T \cdot \mathbf{x}$$

sowie

$$\langle \mathbf{A} \cdot \mathbf{x}, \tilde{\mathbf{x}} \rangle = (\mathbf{A} \cdot \mathbf{x})^T \cdot \tilde{\mathbf{x}} = \mathbf{x}^T \cdot (\mathbf{A}^T \cdot \tilde{\mathbf{x}}) = \langle \mathbf{x}, \mathbf{A}^T \cdot \tilde{\mathbf{x}} \rangle,$$

wobei  $\mathbf{A} \in \mathbb{R}^{(n,n)}$  beliebig ist.

2. Es gilt auch stets die Parallelogramm-Identität

$$4 \cdot \langle \mathbf{x}, \tilde{\mathbf{x}} \rangle = \langle \mathbf{x} + \tilde{\mathbf{x}}, \mathbf{x} + \tilde{\mathbf{x}} \rangle - \langle \mathbf{x} + \tilde{\mathbf{x}}, \mathbf{x} - \tilde{\mathbf{x}} \rangle = \|\mathbf{x} + \tilde{\mathbf{x}}\|^2 - \|\mathbf{x} - \tilde{\mathbf{x}}\|^2.$$

Insbesondere kann jedes Skalarprodukt durch Normen ausgedrückt werden.

**Theorem (Hauptsatz über orthogonale Matrizen)** Für eine quadratische Matrix  $\mathbf{A} \in \mathbb{R}^{(n,n)}$  sind die folgenden Aussagen äquivalent

1.  $\mathbf{A}$  ist längentreu, d.h. es gilt  $\|\mathbf{A} \cdot \mathbf{x}\| = \|\mathbf{x}\|$  für alle  $\mathbf{x} \in \mathbb{R}^n$ .
2.  $\mathbf{A}$  erhält das Skalarprodukt, d.h. es gilt

$$\langle \mathbf{A} \cdot \mathbf{x}, \mathbf{A} \cdot \tilde{\mathbf{x}} \rangle = \langle \mathbf{x}, \tilde{\mathbf{x}} \rangle$$

für alle  $\mathbf{x}, \tilde{\mathbf{x}} \in \mathbb{R}^n$

3.  $\mathbf{A}$  ist orthogonale Matrix.
4. Die Zeilen von  $\mathbf{A}$  bilden eine ON-Basis im  $\mathbb{R}^n$ .
5. Die Spalten von  $\mathbf{A}$  bilden eine ON-Basis im  $\mathbb{R}^n$ .

Beweis, 1.  $\Rightarrow$  2.: Die Parallelogramm-Identität und die Voraussetzung liefern

$$\begin{aligned} 4 \cdot \langle \mathbf{A} \cdot \mathbf{x}, \mathbf{A} \cdot \tilde{\mathbf{x}} \rangle &= \|\mathbf{A} \cdot (\mathbf{x} + \tilde{\mathbf{x}})\|^2 - \|\mathbf{A} \cdot (\mathbf{x} - \tilde{\mathbf{x}})\|^2 \\ &= \|\mathbf{x} + \tilde{\mathbf{x}}\|^2 - \|\mathbf{x} - \tilde{\mathbf{x}}\|^2 \\ &= 4 \cdot \langle \mathbf{x}, \tilde{\mathbf{x}} \rangle. \end{aligned}$$

Beweis, 2.  $\Rightarrow$  3.: Für die  $(i, j)$ -te Komponente des Matrixproduktes  $\mathbf{A}^T \cdot \mathbf{A}$  gilt

$$(\mathbf{A}^T \cdot \mathbf{A})_{ij} = \mathbf{e}_i^T \cdot \mathbf{A}^T \cdot \mathbf{A} \cdot \mathbf{e}_j = \langle \mathbf{A} \cdot \mathbf{e}_i, \mathbf{A} \cdot \mathbf{e}_j \rangle = \langle \mathbf{e}_i, \mathbf{e}_j \rangle = \delta_{ij} = (\mathbf{I}_n)_{ij},$$

und wir haben damit komponentenweise die Gleichung

$$\mathbf{A}^T \cdot \mathbf{A} = \mathbf{I}_n$$

gezeigt. Die Rechenregeln für inverse Matrizen garantieren, dass  $\mathbf{A}^T$  die Inverse von  $\mathbf{A}$  ist und damit auch  $\mathbf{A} \cdot \mathbf{A}^T = \mathbf{I}_n$  erfüllt.

Beweis, 3.  $\Rightarrow$  1.: Es gilt

$$\|\mathbf{A} \cdot \mathbf{x}\|^2 = \langle \mathbf{A} \cdot \mathbf{x}, \mathbf{A} \cdot \mathbf{x} \rangle = \langle \mathbf{x}, \mathbf{A}^T \cdot \mathbf{A} \cdot \mathbf{x} \rangle = \langle \mathbf{x}, \mathbf{x} \rangle = \|\mathbf{x}\|^2.$$

Beweis, 3.  $\Leftrightarrow$  4.  $\Leftrightarrow$  5.: Dies wird in den Großen Übungen gezeigt. □

**Folgerung** Das Produkt zweier orthogonaler Matrizen ist wieder orthogonal, wie die folgende Rechnung zeigt:

$$(\mathbf{A} \cdot \tilde{\mathbf{A}})^T = \tilde{\mathbf{A}}^T \cdot \mathbf{A}^T = \tilde{\mathbf{A}}^{-1} \cdot \mathbf{A}^{-1} \cdot (\mathbf{A} \cdot \tilde{\mathbf{A}})^{-1}$$

Außerdem sind mit  $\mathbf{A}$  auch immer  $\mathbf{A}^T$  und  $\mathbf{A}^{-1}$  orthogonal.

**Bemerkung**

1. Für jede orthogonale Matrix gilt  $\det \mathbf{A} = \pm 1$  (siehe Große Übung).
2. Die Menge aller orthogonalen Matrizen aus  $\mathbb{R}^{(n,n)}$  ist die orthogonale Gruppe  $O(n)$  und alle orthogonalen Matrizen mit Determinante 1 bilden die spezielle orthogonale Gruppe  $SO(n)$ .
3. Für  $n = 2$  werden alle orthogonalen Matrizen in der Großen Übung charakterisiert. Insbesondere gibt es nur Drehungen und Spiegelungen.

**Permutationsmatrizen** Jede  $(n, n)$ -Matrix der Form

$$\begin{pmatrix} | & | & | \\ \mathbf{e}_{\sigma(1)} & \cdots & \mathbf{e}_{\sigma(n)} \\ | & | & | \end{pmatrix}$$

mit Permutation  $\sigma \in S(n)$  ist orthogonal und heißt Permutationsmatrix zu  $\sigma$ , wobei  $\mathbf{e}_i \in \mathbb{R}^n$  der  $i$ -te Einheitsvektor ist. Die Determinante einer Permutationsmatrix ist gerade das Vorzeichen der Permutation und die Transponierte einer Permutationsmatrix ist auch eine Permutationsmatrix.

**Beispiel** Für  $n = 2$  gibt es die zwei Permutationsmatrizen

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

wohingegen

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

und

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

alle Permutationsmatrizen für  $n = 3$  sind.

## Orthogonale Matrizen für $n = 3$

**Spiegelungen an einer Ursprungsebene** Jede Ursprungsebene in  $\mathbb{R}^3$  wird durch zwei linear unabhängige Vektoren  $\mathbf{w}_1$  und  $\mathbf{w}_2$  aufgespannt und besitzt immer einen Normalenvektor

$$\mathbf{n} := \pm \frac{\mathbf{w}_1 \times \mathbf{w}_2}{\|\mathbf{w}_1 \times \mathbf{w}_2\|}$$

der Länge 1 (die Wahl des Vorzeichen ist hier egal). Die Spiegelung an dieser Ebene wird durch die Householder-Matrix

$$\mathbf{A} = \mathbf{I}_3 - 2 \cdot \mathbf{n} \cdot \mathbf{n}^T$$

beschrieben, denn wegen

$$\mathbf{w}_i^T \cdot \mathbf{n} = 0, \quad \mathbf{n}^T \cdot \mathbf{n} = 1$$

gilt mit

$$\mathbf{x} = \lambda_1 \cdot \mathbf{w}_1 + \lambda_2 \cdot \mathbf{w}_2 + \mu \cdot \mathbf{n}$$

auch

$$\begin{aligned} \mathbf{A} \cdot \mathbf{x} &= (\lambda_1 \cdot \mathbf{w}_1 + \lambda_2 \cdot \mathbf{w}_2 + \mu \cdot \mathbf{n}) - 2 \cdot (0 + 0 + \mu) \cdot \mathbf{n} \\ &= \lambda_1 \cdot \mathbf{w}_1 + \lambda_2 \cdot \mathbf{w}_2 - \mu \cdot \mathbf{n}. \end{aligned}$$

Insbesondere folgt  $\mathbf{A} \cdot \mathbf{A} \cdot \mathbf{x} = \mathbf{x}$ , d.h. es gilt  $\mathbf{A} \cdot \mathbf{A} = \mathbf{I}_3$ , und weil auch  $\mathbf{A}^T = \mathbf{A}$  gilt (Nachrechnen!) schließen wir, dass jede Householder-Matrix wirklich orthogonal ist.

**Beispiele** Die Matrizen

$$\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

beschreiben Spiegelungen an Ebenen, die jeweils senkrecht auf einem Einheitsvektor stehen. Die Matrix

$$\begin{aligned} \mathbf{A} &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} - 2 \begin{pmatrix} \cos \alpha \\ \sin \alpha \\ 0 \end{pmatrix} \cdot (\cos \alpha \quad \sin \alpha \quad 0) \\ &= \begin{pmatrix} \sin^2 \alpha - \cos^2 \alpha & -2 \sin \alpha \cos \alpha & 0 \\ -2 \sin \alpha \cos \alpha & -\sin^2 \alpha + \cos^2 \alpha & 0 \\ 0 & 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} -\cos(2\alpha) & -\sin(2\alpha) & 0 \\ -\sin(2\alpha) & \cos(2\alpha) & 0 \\ 0 & 0 & 1 \end{pmatrix} \end{aligned}$$

beschreibt auch eine Spiegelung.

Bemerkung Jeder Vektor  $\mathbf{n} \in \mathbb{R}^3$  mit  $\|\mathbf{n}\| = 1$  kann als

$$\mathbf{n} = \begin{pmatrix} \cos \alpha \cdot \cos \beta \\ \sin \alpha \cdot \cos \beta \\ \sin \beta \end{pmatrix}$$

geschrieben werden, wobei  $\alpha$  und  $\beta$  die sogenannten Euler-Winkel sind.

**Bemerkung** Householder-Matrizen existieren auch für  $n \geq 4$  und haben analoge Eigenschaften.

**Drehungen um eine Achse** führen auch auf orthogonale Matrizen.

**Beispiele** Die Matrizen

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\alpha) & -\sin(\alpha) \\ 0 & \sin(\alpha) & \cos(\alpha) \end{pmatrix}, \quad \begin{pmatrix} \cos(\alpha) & 0 & -\sin(\alpha) \\ 0 & 1 & 0 \\ \sin(\alpha) & 0 & \cos(\alpha) \end{pmatrix}, \quad \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) & 0 \\ \sin(\alpha) & \cos(\alpha) & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

beschreiben Drehungen um den Winkel  $\alpha$  bzgl. einer zu einem Einheitsvektor parallelen Drehachse.

**Theorem (Darstellungssatz für  $O(3)$ )** Sei  $\mathbf{A} \in \mathbb{R}^{(3,3)}$  orthogonal. Dann ist  $\mathbf{A}$

1. eine Drehung, sofern  $\det \mathbf{A} = 1$ ,
2. eine Komposition einer Drehung und einer Spiegelung an einer Ebene, falls  $\det \mathbf{A} = -1$ .

*Beweis, Teil 1a* : Wir zeigen zunächst, dass es einen Achsenvektor  $\mathbf{v}_* \in \mathbb{R}^3$  mit  $\mathbf{A} \cdot \mathbf{v}_* = \mathbf{v}_*$  gibt. Dazu bemerken wir, dass wegen

$$\begin{aligned} \det(\mathbf{A} - \mathbf{I}_3) &= \det(\mathbf{A} \cdot (\mathbf{I}_3 - \mathbf{A}^T)) = \det \mathbf{A} \cdot \det((\mathbf{I}_3 - \mathbf{A})^T) = 1 \cdot \det(\mathbf{I}_3 - \mathbf{A}) \\ &= \det((-1) \cdot (\mathbf{A} - \mathbf{I}_3)) = (-1)^3 \cdot \det(\mathbf{A} - \mathbf{I}_3) = -\det(\mathbf{A} - \mathbf{I}_3) \end{aligned}$$

schon  $\det(\mathbf{A} - \mathbf{I}_3) = 0$  gelten muss und wir daher  $\mathbf{v}_* \neq 0$  im Kern von  $\mathbf{A} - \mathbf{I}_3$  wählen können, wobei durch Multiplikation mit einem geeigneten Skalar zusätzlich  $\|\mathbf{v}_*\| = 1$  sichergestellt werden kann.

*Beweis, Teil 1b* : Wir wählen nun (zum Beispiel mit Gram-Schmidt) eine ON-Basis  $\tilde{\mathbf{v}}_1, \tilde{\mathbf{v}}_2, \tilde{\mathbf{v}}_3$  mit  $\tilde{\mathbf{v}}_1 = \mathbf{v}_*$  und betrachten die Matrix

$$\tilde{\mathbf{A}} = \mathbf{S}^{-1} \cdot \mathbf{A} \cdot \mathbf{S}, \quad \mathbf{S} = \begin{pmatrix} | & | & | \\ \tilde{\mathbf{v}}_1 & \tilde{\mathbf{v}}_2 & \tilde{\mathbf{v}}_3 \\ | & | & | \end{pmatrix}.$$

Da  $\mathbf{S}$  und  $\mathbf{S}^{-1} = \mathbf{S}^T$  orthogonal sind (die Spalten bzw. Zeilen sind gerade die ON Vektoren  $\tilde{\mathbf{v}}_j$ ), ist  $\tilde{\mathbf{A}}$  als Produkt orthogonaler Matrizen auch wieder orthogonal. Nach Konstruktion gilt außerdem

$$\tilde{\mathbf{v}}_1 = \mathbf{S} \cdot \mathbf{e}_1, \quad \mathbf{A} \cdot \tilde{\mathbf{v}}_1 = \tilde{\mathbf{v}}_1, \quad \mathbf{S}^{-1} \cdot \tilde{\mathbf{v}}_1 = \mathbf{e}_1$$

und daher auch

$$\tilde{\mathbf{A}} \cdot \mathbf{e}_1 = \mathbf{S}^{-1} \cdot \mathbf{A} \cdot \mathbf{S} \cdot \mathbf{e}_1 = \mathbf{S}^{-1} \cdot \mathbf{A} \cdot \tilde{\mathbf{v}}_1 = \mathbf{S}^{-1} \cdot \tilde{\mathbf{v}}_1 = \mathbf{e}_1$$

und wir schließen, dass die erste Spalte von  $\tilde{\mathbf{A}}$  gerade der erste Einheitsvektor  $\mathbf{e}_1$  ist. Da die Spalten von  $\tilde{\mathbf{A}}$  aber ein ON-Basis bilden, muss damit schon

$$\tilde{\mathbf{A}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & b_{11} & b_{12} \\ 0 & b_{21} & b_{22} \end{pmatrix}$$

gelten, wobei

$$\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix}$$

eine orthogonale (2, 2)-Matrix mit Determinante +1 ist. Der Darstellungssatz für  $O(2)$  aus der Großen Übung garantiert nun

$$\tilde{\mathbf{A}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos(\alpha) & -\sin(\alpha) \\ 0 & \sin(\alpha) & \cos(\alpha) \end{pmatrix}$$

für ein  $\alpha \in \mathbb{R}$ . Diese Matrix beschreibt die Drehung mit Winkel  $\alpha$  um die zu  $\mathbf{e}_1$  parallele Drehachse.

Beweis, Teil 1c: Die Matrizen  $\mathbf{A}$  und  $\tilde{\mathbf{A}}$  beschreiben dieselbe lineare Abbildung  $T$  bzgl. zweier verschiedener Basen —  $\mathbf{A}$  bzgl. der kanonischen Basisvektoren  $\mathbf{e}_i$  und  $\tilde{\mathbf{A}}$  bzgl. der geschickt gewählten Basisvektoren  $\tilde{\mathbf{v}}_i$  — und wir schließen, dass auch  $\mathbf{A}$  die Drehung mit Winkel  $\alpha$  und Drehachse  $\mathbf{v}_*$  beschreibt.

Beweis, Teil 2: Im Fall von  $\det \mathbf{A} = -1$  betrachten wir die Matrix

$$\check{\mathbf{A}} := \mathbf{A} \cdot \hat{\mathbf{A}}, \quad \hat{\mathbf{A}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix},$$

die als Produkt orthogonaler Matrizen selbst orthogonal ist und für die wir wegen

$$\det \check{\mathbf{A}} = \det \mathbf{A} \cdot (\det \hat{\mathbf{A}}) = (-1) \cdot (-1) = 1$$

die Ergebnisse des ersten Teils anwenden können. Insbesondere ist  $\check{\mathbf{A}}$  eine Drehung und weil die Householder-Matrix  $\hat{\mathbf{A}} = \hat{\mathbf{A}}^T = \hat{\mathbf{A}}^{-1}$  eine Spiegelung beschreibt, folgt die Behauptung aus  $\mathbf{A} = \check{\mathbf{A}} \cdot \hat{\mathbf{A}}^{-1}$ .  $\square$

Bemerkung: Im Beweis beschreibt die Matrix  $\mathbf{S}$  mit den Spalten  $\tilde{\mathbf{v}}_j$  gerade den Basiswechsel von  $(\tilde{\mathbf{v}}_1, \dots, \tilde{\mathbf{v}}_3)$  nach  $(\mathbf{e}_1, \dots, \mathbf{e}_3)$ . Dies kann entweder mit den abstrakten Resultaten von oben begründet oder wie folgt direkt verifiziert werden: Wenn wir die  $i$ -te Komponente des Spaltenvektors  $\tilde{\mathbf{v}}_j$  mit  $\tilde{v}_{j,i}$  und den  $(i, j)$ -Eintrag von  $\mathbf{S}$  wie üblich mit  $s_{ij}$  bezeichnen, so gilt  $\tilde{v}_{j,i} = s_{ij}$  und aus

$$\begin{aligned} \sum_{i=1}^3 x_i \cdot \mathbf{e}_i &= \sum_{j=1}^3 \tilde{x}_j \cdot \tilde{\mathbf{v}}_j = \sum_{j=1}^3 \tilde{x}_j \cdot \left( \sum_{i=1}^3 \tilde{v}_{j,i} \mathbf{e}_i \right) = \sum_{j=1}^3 \tilde{x}_j \cdot \left( \sum_{i=1}^3 s_{ij} \mathbf{e}_i \right) \\ &= \sum_{i=1}^3 \left( \sum_{j=1}^3 s_{ij} \cdot \tilde{x}_j \right) \cdot \mathbf{e}_i \end{aligned}$$

ergibt sich  $x_i = \sum_{j=1}^3 s_{ij} \cdot \tilde{x}_j$  bzw.  $\mathbf{x} = \mathbf{S} \cdot \tilde{\mathbf{x}}$  nach Koeffizientenvergleich.

### Folgerung

1. Die Komposition zweier beliebiger Drehungen im  $\mathbb{R}^3$  ist eine einzelne Drehung!  
Die Komposition zwei Spiegelungen ist auch eine Drehung.

2. Jede (3, 3)-Permutationsmatrix ist entweder Drehung oder Drehspiegelung.
3. Jede Permutationsmatrix ist entweder Drehung oder Komposition von Drehung und Spiegelung. Das kann man auch direkt sehen!

**Bemerkung** Für  $n \geq 4$  sind die Darstellungssätze von  $O(n)$  deutlich komplizierter.

**Beispiel** Die Matrix

$$\mathbf{A} = \frac{1}{3} \begin{pmatrix} 1 & 2 & 2 \\ 2 & 1 & -2 \\ -2 & 2 & -1 \end{pmatrix}$$

ist orthogonal mit Determinante +1 (Nachrechnen!) und beschreibt daher eine Drehung. Für die Drehachse suchen wir Lösungen  $\mathbf{v}_*$  der Gleichung  $(\mathbf{A} - \mathbf{I}_3) \cdot \mathbf{v}_* = \mathbf{0}$  und finden

$$\mathbf{v}_* = \lambda \cdot \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix},$$

wobei  $\lambda \in \mathbb{R}$  freier Parameter ist (und etwa als  $\lambda = 1/\sqrt{2}$  gewählt werden könnte, um  $\|\mathbf{v}_*\| = 1$  zu erreichen). Um auch noch den Drehwinkel  $\alpha$  zu bestimmen, können wir die Transformationen aus dem Beweis benutzen oder alternativ mit

$$\mathbf{w} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}, \quad \mathbf{A} \cdot \mathbf{w} = \frac{1}{3} \begin{pmatrix} 2 \\ -2 \\ -1 \end{pmatrix}$$

wie folgt argumentieren:  $\mathbf{w}$  und  $\mathbf{A} \cdot \mathbf{w}$  stehen beide senkrecht auf der durch  $\mathbf{v}_*$  aufgespannten Drehachse. Also ist der Drehwinkel gerade der Winkel zwischen diesen beiden Vektoren und wir erhalten

$$\cos(\alpha) = \frac{\langle \mathbf{w}, \mathbf{A} \cdot \mathbf{w} \rangle}{\|\mathbf{w}\| \cdot \|\mathbf{A} \cdot \mathbf{w}\|} = -\frac{1}{3}$$

bzw.  $\alpha = \arccos\left(-\frac{1}{3}\right) \approx 1,9 \cong 109,5^\circ$ .



# Kapitel 6

## Spektraltheorie

Vorlesung 17, 27. November 2019

**Vorbemerkung** In diesem Kapitel studieren wir ausschließlich quadratische  $(n, n)$ -Matrizen und führen die fundamentalen Konzepte der sogenannten *Eigenwerttheorie* oder *Spektraltheorie* ein. Um die Darstellung möglichst einfach zu halten, wollen wir jede quadratische Matrix standardmäßig immer als Element von  $\mathbb{C}^{(n,n)}$  betrachten, selbst dann, wenn alle ihre Einträge reelle Zahlen sind. Der Grund ist, dass auch in der Spektraltheorie reeller Matrizen in natürlicher Weise komplexe Zahlen auftauchen können, nämlich als Nullstellen eines Polynoms. Alternativ kann man zwischen einer reellen und einer komplexen Theorie unterscheiden. In diesem Fall schreibt man oftmals  $\mathbf{A} \in \mathbb{K}^{(n,n)}$ , sofern eine Aussage sowohl für reelle als auch für komplexe Matrizen gilt und  $\mathbb{K}$  entweder durch  $\mathbb{R}$  oder durch  $\mathbb{C}$  ersetzt werden kann.

### 6.1 Eigenwerte und Eigenvektoren

**Definition** Eine Matrix  $\mathbf{A} \in \mathbb{C}^{(n,n)}$  heißt diagonalisierbar, falls

$$\mathbf{\Lambda} = \mathbf{S}^{-1} \cdot \mathbf{A} \cdot \mathbf{S}$$

für eine Diagonalmatrix  $\mathbf{\Lambda} \in \mathbb{C}^{(n,n)}$  (mit Diagonaleinträgen  $\lambda_1, \dots, \lambda_n$ ) sowie eine invertierbare Matrix  $\mathbf{S} \in \mathbb{C}^{(n,n)}$  gilt.

**Beobachtung** Ist  $\mathbf{A}$  diagonalisierbar, so gilt  $\lambda_j \cdot \mathbf{e}_j = \mathbf{S}^{-1} \cdot \mathbf{A} \cdot \mathbf{S} \cdot \mathbf{e}_j$  für jeden Einheitsvektor  $\mathbf{e}_j$  und damit

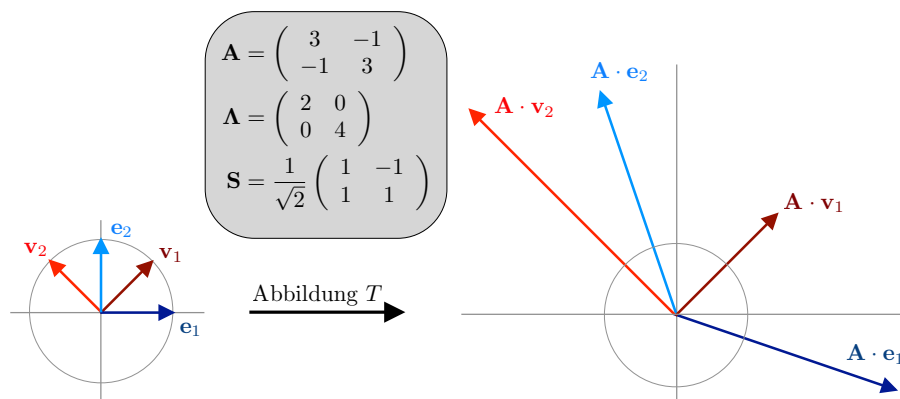
$$\lambda_j \cdot \mathbf{v}_j = \mathbf{A} \cdot \mathbf{v}_j,$$

wobei  $\mathbf{v}_j := \mathbf{S} \cdot \mathbf{e}_j$  der  $j$ -te Spaltenvektor von  $\mathbf{S}$  ist.

**Interpretation** Wenn  $\mathbf{A}$  diagonalisierbar ist, dann existiert eine Basis  $\mathbf{v}_1, \dots, \mathbf{v}_n$  (die Spaltenvektoren von  $\mathbf{S}$ ), so dass die durch  $\mathbf{A}$  kodierte lineare Abbildung  $T(\mathbf{v}) = \mathbf{A} \cdot \mathbf{v}$  bzgl. der neuen Basis sehr einfach aussieht:

$$\mathbf{v} = \sum_{j=1}^n \nu_j \cdot \mathbf{v}_j \quad \Longrightarrow \quad T(\mathbf{v}) = \sum_{j=1}^n \lambda_j \cdot \nu_j \cdot \mathbf{v}_j$$

Insbesondere ist  $T(\mathbf{v}_j) = \lambda_j \cdot \mathbf{v}_j$ , d.h. die Abbildung  $T$  streckt jeden Vektor  $\mathbf{v}_j$  um den Faktor  $\lambda_j$  (siehe Bild).



Einfaches Beispiel für eine diagonalisierbare Matrix, wobei hier alle auftretenden Matrizen reell sind.

**Definition** Gilt

$$\lambda \cdot \mathbf{v} = \mathbf{A} \cdot \mathbf{v}$$

für ein  $\lambda \in \mathbb{C}$  und ein  $\mathbf{v} \in \mathbb{C}^n$  mit  $\mathbf{v} \neq \mathbf{0}$ , so heißt  $\lambda$  Eigenwert von  $\mathbf{A} \in \mathbb{C}^{(n,n)}$  und  $\mathbf{v}$  wird Eigenvektor zum Eigenwert  $\lambda$  genannt.

**Beispiel** Es gilt  $\mathbf{S}^{-1} \cdot \mathbf{A} \cdot \mathbf{S} = \mathbf{\Lambda}$  für

$$\mathbf{A} = \begin{pmatrix} 2 & 1 \\ 6 & 3 \end{pmatrix}, \quad \mathbf{S} = \begin{pmatrix} -1 & 1 \\ 2 & 3 \end{pmatrix}, \quad \mathbf{S}^{-1} = \frac{1}{5} \begin{pmatrix} -3 & 1 \\ 2 & 1 \end{pmatrix}, \quad \mathbf{\Lambda} = \begin{pmatrix} 0 & 0 \\ 0 & 5 \end{pmatrix}$$

d.h.  $\mathbf{A}$  ist diagonalisierbar (wobei in diesem Beispiel nur reelle Zahlen auftauchen). Insbesondere sind

$$\lambda_1 = 0, \quad \lambda_2 = 5, \quad \mathbf{v}_1 = \begin{pmatrix} -1 \\ 2 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 1 \\ 3 \end{pmatrix}$$

die Eigenwerte und Eigenvektoren von  $\mathbf{A}$ .  $\lambda_1$  und  $\lambda_2$  sind auch die Eigenwerte der Diagonalmatrix  $\mathbf{\Lambda}$ , aber die entsprechenden Eigenvektoren sind die Einheitsvektoren  $\mathbf{e}_1$  und  $\mathbf{e}_2$ .

**Beispiel** Die reelle Matrix

$$\mathbf{A} = \begin{pmatrix} a & -b \\ b & a \end{pmatrix}, \quad a, b \in \mathbb{R}$$

ist diagonalisierbar mit komplexen Matrizen

$$\mathbf{S} = \begin{pmatrix} -i & i \\ 1 & 1 \end{pmatrix}, \quad \mathbf{S}^{-1} = \frac{1}{2} \begin{pmatrix} i & 1 \\ -i & 1 \end{pmatrix}, \quad \mathbf{\Lambda} = \begin{pmatrix} a - i \cdot b & 0 \\ 0 & a + i \cdot b \end{pmatrix}.$$

Insbesondere liefern die Diagonaleinträge von  $\mathbf{\Lambda}$  bzw. die Spalten von  $\mathbf{S}$  die Eigenwerte bzw. Eigenvektoren von  $\mathbf{A}$ , denn einfache Rechnungen zeigen

$$\mathbf{A} \cdot \mathbf{v}_1 = \mathbf{A} \cdot \begin{pmatrix} -i \\ 1 \end{pmatrix} = \begin{pmatrix} -b - i \cdot a \\ a - i \cdot b \end{pmatrix} = (a - i \cdot b) \begin{pmatrix} -i \\ 1 \end{pmatrix} = \lambda_1 \cdot \mathbf{v}_1$$

sowie

$$\mathbf{A} \cdot \mathbf{v}_2 = \mathbf{A} \cdot \begin{pmatrix} i \\ 1 \end{pmatrix} = \begin{pmatrix} -b + i \cdot a \\ a + i \cdot b \end{pmatrix} = (a + i \cdot b) \begin{pmatrix} i \\ 1 \end{pmatrix} = \lambda_2 \cdot \mathbf{v}_2$$

Ein Spezialfall sind zweidimensionale Drehmatrizen mit  $a = \cos(\theta)$  und  $b = \sin(\theta)$ .

**Achtung**

1. Die Eigenwerte und -vektoren einer reellen Matrix können nicht-verschwindende Imaginärteile aufweisen! Insbesondere gilt: In der Eigenwert-Theorie reeller Matrizen sind komplexe Zahlen im Allgemeinen nicht zu vermeiden!
2. Nicht jede Matrix  $\mathbf{A}$  ist diagonalisierbar, und eine zentrale Frage ist, welche Matrizen dies sind bzw. nicht sind. Standard-Beispiele für Nicht-Diagonalisierbarkeit sind die weiter unten betrachteten Jordan-Matrizen.
3. Ist  $\mathbf{A}$  reelle Matrix und  $\lambda$  ein reeller Eigenwert von  $\mathbf{A}$ , so gibt es auch immer einen entsprechenden reellen Eigenvektor  $\mathbf{v}$ . In der Tat, unter diesen Annahmen kann leicht gezeigt werden, dass aus jedem Eigenvektor  $\mathbf{v}$  mit komplexen Einträgen durch komponentenweisen Übergang zum Realteil und zum Imaginärteil reelle Eigenvektoren konstruiert werden können (Übungsaufgabe).

**Bemerkung**

1. Die Menge aller Eigenwerte von  $\mathbf{A}$  wird Spektrum von  $\mathbf{A}$  genannt und mit  $\text{spec } \mathbf{A}$  abgekürzt. Das Spektrum von  $\mathbf{A}$  ist also eine Teilmenge von  $\mathbb{C}$ .
2. Die Menge

$$E_\lambda := \{\mathbf{v} \in \mathbb{C}^n : \lambda \cdot \mathbf{v} = \mathbf{A} \cdot \mathbf{v}\} = \ker(\mathbf{A} - \lambda \cdot \mathbf{I})$$

ist für jedes  $\lambda \in \mathbb{C}$  ein Unterraum von  $\mathbb{C}^n$  und wird für Eigenwerte  $\lambda \in \text{spec } \mathbf{A}$  der entsprechende Eigenraum genannt. Insbesondere liefert der Hauptsatz über quadratische Matrizen (angewendet auf  $\mathbf{A} - \lambda \cdot \mathbf{I}$ ) die logische Äquivalenz

$$\begin{aligned} \lambda \in \text{spec } \mathbf{A} &\Leftrightarrow E_\lambda \neq \{\mathbf{0}\} \Leftrightarrow \dim(\ker(\mathbf{A} - \lambda \cdot \mathbf{I})) > 0 \\ &\Leftrightarrow \det(\mathbf{A} - \lambda \cdot \mathbf{I}) = 0 \end{aligned}$$

bzw.  $\lambda \notin \text{spec } \mathbf{A} \Leftrightarrow E_\lambda = \{\mathbf{0}\} \Leftrightarrow \det(\mathbf{A} - \lambda \cdot \mathbf{I}) \neq 0$ . Die Determinantenbedingung führt in sehr natürlicher Weise zum weiter unten diskutierten *charakteristischen Polynom* der Matrix  $\mathbf{A}$ .

3. Das Spektrum kodiert alle wichtigen Eigenschaften einer Matrix und spielt eine zentrale Rolle in vielen Bereichen der Mathematik und den Anwendungswissenschaften. Zum Beispiel müssen bei *Stabilitätsuntersuchungen* in der Regel Eigenwerte quadratischer Matrizen berechnet und ausgewertet werden.
4. Für jede Diagonalmatrix zeigt man durch einfache Rechnungen: Die Diagonaleinträge sind die Eigenwerte und die kanonischen Einheitsvektoren sind die entsprechenden Eigenvektoren.
5. Das Spektrum ähnlicher Matrizen ist identisch. Genauer gesagt: Aus

$$\mathbf{B} = \mathbf{R}^{-1} \cdot \mathbf{A} \cdot \mathbf{R} \quad \text{und} \quad \lambda \cdot \mathbf{v} = \mathbf{A} \cdot \mathbf{v}$$

folgt (Nachrechnen!)

$$\lambda \cdot \mathbf{w} = \mathbf{B} \cdot \mathbf{w} \quad \text{für} \quad \mathbf{w} = \mathbf{R} \cdot \mathbf{v}.$$

Insbesondere ist jeder Eigenwert von  $\mathbf{A}$  auch Eigenwert von  $\mathbf{B}$ , und wegen  $\mathbf{A} = \mathbf{R} \cdot \mathbf{B} \cdot \mathbf{R}^{-1}$  und  $\mathbf{v} = \mathbf{R}^{-1} \cdot \mathbf{w}$  gilt analog auch die Umkehrung. Die entsprechenden Eigenvektoren  $\mathbf{v}$  und  $\mathbf{w}$  sind zwar verschieden, können aber mittels der Übergangsmatrizen  $\mathbf{R}$  bzw.  $\mathbf{R}^{-1}$  ineinander umgerechnet werden.

**Lemma (Unabhängigkeit von Eigenvektoren)** Seien  $\lambda_1, \dots, \lambda_k$  paarweise verschiedene Eigenwerte der Matrix  $\mathbf{A}$  und seien  $\mathbf{v}_1, \dots, \mathbf{v}_k$  entsprechende Eigenvektoren. Dann sind diese Eigenvektoren linear unabhängig.

*Beweis:* Wir benutzen vollständige Induktion über  $k$  und bemerken, dass der Induktionsanfang  $k = 1$  trivial ist. Für den Induktionsschritt  $k - 1 \rightsquigarrow k$  betrachten wir eine Linearkombination der  $\mathbf{v}_j$  mit Gewichten  $\nu_j \in \mathbb{C}$ , so dass

$$0 = \sum_{j=1}^k \nu_j \cdot \mathbf{v}_j.$$

Durch Multiplizieren beider Seiten mit der Matrix  $\mathbf{A} - \lambda_k \cdot \mathbf{I}$  ergibt sich

$$\begin{aligned} \mathbf{0} &= (\mathbf{A} - \lambda_k \cdot \mathbf{I}) \cdot \left( \sum_{j=1}^k \nu_j \cdot \mathbf{v}_j \right) = \sum_{j=1}^k \nu_j \cdot \mathbf{A} \cdot \mathbf{v}_j - \sum_{j=1}^k \nu_j \cdot \lambda_k \cdot \mathbf{v}_j \\ &= \sum_{j=1}^k \nu_j \cdot (\lambda_j - \lambda_k) \cdot \mathbf{v}_j = \left( \sum_{j=1}^{k-1} \nu_j \cdot (\lambda_j - \lambda_k) \cdot \mathbf{v}_j \right) + \mathbf{0}. \end{aligned}$$

Die Induktionsvoraussetzung impliziert nun  $\nu_j \cdot (\lambda_j - \lambda_k) = 0$  für alle  $j = 1, \dots, k - 1$  und wegen  $\lambda_j \neq \lambda_k$  erhalten wir

$$\nu_1 = \dots = \nu_{k-1} = 0.$$

Außerdem muss aber auch  $\nu_k = 0$  gelten, weil ja insgesamt die betrachtete Linearkombination verschwinden soll, und wir haben den Induktionsschritt etabliert.  $\square$

**Theorem (1. Diagonalisierbarkeitssatz)** Jede Matrix  $\mathbf{A} \in \mathbb{C}^{(n,n)}$ , die  $n$  paarweise verschiedene Eigenwerte besitzt, ist diagonalisierbar.

*Beweis:* Wir wählen für jeden Eigenwert  $\lambda_j$  einen Eigenvektor  $\mathbf{v}_j$  und bilden aus diesen spaltenweise eine Matrix  $\mathbf{S}$ , d.h.

$$\mathbf{S} := \begin{pmatrix} | & & | \\ \mathbf{v}_1 & \dots & \mathbf{v}_n \\ | & & | \end{pmatrix}.$$

Nach dem vorangegangenen Lemma sind die Vektoren  $\mathbf{v}_1, \dots, \mathbf{v}_n$  linear unabhängig, d.h. die Matrix  $\mathbf{S}$  besitzt maximalen Rang und ist damit invertierbar (Hauptsatz über quadratische Matrizen). Außerdem ist  $\mathbf{\Lambda} := \mathbf{S}^{-1} \cdot \mathbf{A} \cdot \mathbf{S}$  gerade die Diagonalmatrix der Eigenwerte  $\lambda_j$ , denn es gilt

$$\mathbf{S} \cdot \mathbf{e}_j = \mathbf{v}_j, \quad \mathbf{A} \cdot \mathbf{S} \cdot \mathbf{e}_j = \mathbf{A} \cdot \mathbf{v}_j = \lambda_j \cdot \mathbf{v}_j, \quad \mathbf{S}^{-1} \cdot \mathbf{A} \cdot \mathbf{S} \cdot \mathbf{e}_j = \lambda_j \cdot \mathbf{S}^{-1} \cdot \mathbf{v}_j = \lambda_j \cdot \mathbf{e}_j$$

für alle  $j = 1, \dots, n$ , d.h. die Matrix-Multiplikation mit  $\mathbf{S}^{-1} \cdot \mathbf{A} \cdot \mathbf{S}$  bildet jeden Einheitsvektor  $\mathbf{e}_j$  auf  $\lambda_j \cdot \mathbf{e}_j$  ab.  $\square$

**Beispiel** Die Matrix

$$\mathbf{A} = \begin{pmatrix} 1 & 0 & 8 \\ 0 & 3 & 0 \\ 2 & 0 & 1 \end{pmatrix}$$

besitzt die drei verschiedenen Eigenwerte

$$\lambda_1 = -3, \quad \lambda_2 = 3, \quad \lambda_3 = 5,$$

wobei wir gleich sehen werden, dass diese als Nullstellen des Polynoms

$$p_{\mathbf{A}}(z) = -z^3 + 5z^2 + 9z - 45$$

ermittelt werden können. Für jeden Eigenwert  $\lambda_j$  bestimmen wir nun durch Lösen des entsprechenden Gleichungssystems einen Eigenvektor  $\mathbf{v}_i$  mit  $(\mathbf{A} - \lambda_i \cdot \mathbf{I})$  und schreiben diese als Spalten in die Matrix  $\mathbf{S}$ . Dies liefert zum Beispiel

$$\mathbf{S} = \begin{pmatrix} -2 & 0 & 2 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

und durch direkte Rechnungen verifizieren wir

$$\mathbf{S}^{-1} = \frac{1}{4} \begin{pmatrix} -1 & 0 & 2 \\ 0 & 4 & 0 \\ 1 & 0 & 2 \end{pmatrix}, \quad \mathbf{S}^{-1} \cdot \mathbf{A} \cdot \mathbf{S} = \begin{pmatrix} -3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 5 \end{pmatrix},$$

wobei die Diagonalmatrix wieder aus den Eigenwerten besteht.

*Bemerkung:* Die Eigenvektoren sind nur bis auf einen Vorfaktor eindeutig bestimmt. Zum Beispiel: Um den Raum  $E_{\lambda_1}$  bzw. den Kern von  $\mathbf{A} - \lambda_1 \cdot \mathbf{I}$  zu bestimmen, müssen wir das lineare und homogene Gleichungssystem

$$(\mathbf{A} - \lambda_1 \cdot \mathbf{I}) \cdot \mathbf{v}_1 = \mathbf{0} \quad \text{bzw.} \quad \begin{pmatrix} 4 & 0 & 8 \\ 0 & 6 & 0 \\ 2 & 0 & 4 \end{pmatrix} \cdot \begin{pmatrix} v_{1,1} \\ v_{1,2} \\ v_{1,3} \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

lösen und erhalten

$$\mathbf{v}_1 = \begin{pmatrix} v_{1,1} \\ v_{1,2} \\ v_{1,3} \end{pmatrix} = v_{1,3} \begin{pmatrix} -2 \\ 0 \\ 1 \end{pmatrix}$$

als allgemeine Lösung, wobei  $v_{1,3} \neq 0$  beliebig gewählt werden kann. Die Wahl der Matrix  $\mathbf{S}$  ist also nicht eindeutig bzw. wir werden weiter unten sehen, dass man weitere Bedingungen an die Eigenvektoren stellen kann.

**Polynome und Spektrum** Der Hauptsatz über quadratische Matrizen — bzw. die entsprechende Bemerkung weiter oben — zeigt, dass  $\lambda \in \mathbb{C}$  genau dann Eigenwert von  $\mathbf{A}$  ist, wenn  $\lambda$  komplexe Nullstelle von  $p_{\mathbf{A}}$  ist, wobei

$$p_{\mathbf{A}}(z) := \det(\mathbf{A} - z \cdot \mathbf{I})$$

das charakteristische Polynom von  $\mathbf{A}$  ist, das immer den Grad  $n$  besitzt. Der Fundamentalsatz der Algebra impliziert daher, dass jede quadratische  $(n, n)$ -Matrix (reell oder komplex) genau  $n$  komplexe Eigenwerte besitzt, wobei Mehrfachnullstellen möglich sind und einige (oder sogar alle) komplexen Nullstellen reell sein dürfen.

*\*Bemerkung* Der Fundamentalsatz der Algebra gilt nur im Komplexen, nicht im Reellen. Das ist der Grund, warum komplexe Spektraltheorie viel klarer und einfacher ist als die reelle Variante der Theorie.

**Beispiele**

1. Mit

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix}$$

erhalten wir

$$p_{\mathbf{A}}(z) = \begin{vmatrix} 1-z & 2 \\ 4 & 3-z \end{vmatrix} = (1-z) \cdot (3-z) - 2 \cdot 4 = z^2 - 4z - 5$$

und die  $pq$ -Formel liefert  $\text{spec } \mathbf{A} = \{-1, 5\}$ .

2. Für

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{pmatrix}$$

ergibt sich mit der Sarrus-Regel

$$\begin{aligned} p_{\mathbf{A}}(z) &= \begin{vmatrix} 1-z & 2 & 3 \\ 4 & 5-z & 6 \\ 7 & 8 & 9-z \end{vmatrix} \\ &= (1-z) \cdot (5-z) \cdot (9-z) + 2 \cdot 6 \cdot 7 + 3 \cdot 4 \cdot 8 \\ &\quad - 3 \cdot (5-z) \cdot 7 - 6 \cdot 8 \cdot (1-z) - (z-9) \cdot 2 \cdot 4 \\ &= -z \cdot (z^2 - 15 \cdot z - 18) \end{aligned}$$

und anschließend  $\text{spec } \mathbf{A} = \{0, \frac{3}{2}(5 - \sqrt{33}), \frac{3}{2}(5 + \sqrt{33})\}$  nach Berechnung der Nullstellen von  $p_{\mathbf{A}}$ .

3. Für Diagonalmatrizen ist wieder alles ganz einfach: Mit

$$\mathbf{\Lambda} = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

ergibt sich

$$p_{\mathbf{\Lambda}}(z) = (\lambda_1 - z) \cdot \dots \cdot (\lambda_n - z), \quad \text{spec } \mathbf{\Lambda} = \{\lambda_1, \dots, \lambda_n\}.$$

Beachte, dass wir hier nicht vorausgesetzt haben, dass alle  $\lambda_i$  paarweise verschieden sind.

**Vielfachheiten von Eigenwerten** Eine weitere Konsequenz aus dem Fundamentalsatz der Algebra kann wie folgt formuliert werden: Bezeichnen wir mit  $m$  die Anzahl der paarweise verschiedene Eigenwerte von  $\mathbf{A}$  (es gilt  $1 \leq m \leq n$ ), d.h. gilt

$$\text{spec } \mathbf{A} = \{\lambda_1, \dots, \lambda_m\}, \quad \lambda_{j_1} \neq \lambda_{j_2},$$

so existieren  $m$  natürliche Zahlen  $a_1, \dots, a_m$  mit  $a_1 + \dots + a_m = n$ , so dass

$$p_{\mathbf{A}}(z) = (-1)^n \cdot (z - \lambda_1)^{a_1} \cdot \dots \cdot (z - \lambda_m)^{a_m}.$$

Die Zahl  $a_j = a(\lambda_j)$  heißt dabei die algebraische Vielfachheit des Eigenwerts  $\lambda_j$ , wohingegen

$$g(\lambda_j) = \dim \ker (\mathbf{A} - \lambda_j \cdot \mathbf{I})$$

die geometrische Vielfachheit genannt wird. Beide Vielfachheiten müssen nicht identisch sein.

**Bemerkung**

1. Für zwei ähnliche Matrizen  $\mathbf{A}$  und  $\mathbf{B}$  gilt  $p_{\mathbf{A}}(z) = p_{\mathbf{B}}(z)$  für jedes  $z \in \mathbb{C}$  sowie  $\dim \ker(\mathbf{A} - \lambda \cdot \mathbf{I}) = \dim \ker(\mathbf{B} - \lambda \cdot \mathbf{I})$  für jedes  $\lambda \in \text{spec } \mathbf{A} = \text{spec } \mathbf{B}$ . Insbesondere können die algebraischen und die geometrischen Vielfachheiten der Eigenwerte sowohl mit  $\mathbf{A}$  als auch mit  $\mathbf{B}$  berechnet werden.
2. Für jeden Eigenwert  $\lambda$  gilt  $1 \leq g(\lambda) \leq a(\lambda) \leq n$ . Dies folgt aus einfachen Dimensionsbetrachtungen und geschickter Umformulierung mit ähnlichen Matrizen, siehe zum Beispiel [AORS, Seite 158].

**Beispiel** Die Matrix

$$\mathbf{A} = \begin{pmatrix} 6 & -1 & 2 \\ 0 & 4 & 0 \\ 2 & -1 & 6 \end{pmatrix}, \quad p_{\mathbf{A}}(z) = -z^3 + 16z^2 - 80z + 128 = -(z-4)^2 \cdot (z-8)$$

besitzt den Eigenwert 4 mit algebraischer Vielfachheit 2 sowie den Eigenwert 8 mit algebraischer Vielfachheit 1. Durch eine Gauß-Elimination finden wir

$$\ker(\mathbf{A} - 4 \cdot \mathbf{I}) = \text{span}\{\mathbf{v}_1, \mathbf{v}_2\}, \quad \ker(\mathbf{A} - 8 \cdot \mathbf{I}) = \text{span}\{\mathbf{v}_3\}$$

mit

$$\mathbf{S} = \begin{pmatrix} | & | & | \\ \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \\ | & | & | \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 2 & 0 \\ -1 & 0 & 1 \end{pmatrix},$$

wobei die Eigenvektoren  $\mathbf{v}_1$ ,  $\mathbf{v}_2$  und  $\mathbf{v}_3$  linear unabhängig sind und spaltenweise zur Matrix  $\mathbf{S}$  vereint wurden. Wir können nun leicht nachrechnen, dass  $\mathbf{A}$  via

$$\mathbf{S}^{-1} = \frac{1}{4} \begin{pmatrix} 2 & -1 & -2 \\ 0 & 2 & 0 \\ 2 & -1 & 2 \end{pmatrix}, \quad \mathbf{S}^{-1} \cdot \mathbf{A} \cdot \mathbf{S} = \begin{pmatrix} 4 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 8 \end{pmatrix} = \mathbf{\Lambda}$$

diagonalisierbar ist, wobei die Diagonalelemente von  $\mathbf{\Lambda}$  gerade die Eigenwerte von  $\mathbf{A}$  sind.

**Theorem (2. Diagonalisierbarkeitssatz)** Eine Matrix  $\mathbf{A} \in \mathbb{C}^{(n,n)}$  ist genau dann diagonalisierbar, wenn  $a(\lambda) = g(\lambda)$  für jeden Eigenwert  $\lambda$  von  $\mathbf{A}$  gilt.

*Beweis, Hinrichtung:* Die Behauptung folgt aus den Eigenschaften von Diagonalmatrizen und der Tatsache, dass ähnliche Matrizen dieselben Eigenwerte und Vielfachheiten besitzen.

*Beweis, Rückrichtung:* Da sich die geometrischen Vielfachheiten aller Eigenwerte nach Voraussetzung zu  $n$  addieren (im Allgemeinen gilt dies nach dem Fundamentalsatz der Algebra nur für die algebraischen Vielfachheiten), können wir eine Basis  $\mathbf{v}_1, \dots, \mathbf{v}_n$  von Eigenvektoren wählen und anschließend analog zum Beweis des 1. Diagonalisierbarkeitssatzes argumentieren. Siehe dazu auch das vorherige Beispiel.

**\*Bemerkung** Der erste Diagonalisierbarkeitssatz gilt analog auch in der reellen Spektraltheorie, der zweite aber nicht, denn er basiert auf dem Fundamentalsatz der Algebra.

Vorlesung 18, 29. November 2019

**Berechnungsvorschrift** Für eine gegebene Matrix  $\mathbf{A} \in \mathbb{C}^{(n,n)}$  müssen wir im Allgemeinen folgende Schritte ausführen, um alle Größen aus diesem Abschnitt zu bestimmen:

1. Bestimme das charakteristische Polynom  $p_{\mathbf{A}}$  durch Determinantenberechnung mit Parameter  $z$ .
2. Finde alle Eigenwerte  $\lambda$  von  $\mathbf{A}$  als komplexe Nullstellen von  $p_{\mathbf{A}}$  sowie deren algebraischen Vielfachheiten  $a(\lambda)$ . Dies kann für große Werte von  $n$  sehr schwierig bzw. aufwändig sein und gelingt vielleicht nur approximativ auf dem Computer.
3. Berechne für jeden Eigenwert  $\lambda \in \text{spec } \mathbf{A}$  den Eigenraum  $E_{\lambda}$  durch Lösen des homogenen Gleichungssystems  $(\mathbf{A} - \lambda \cdot \mathbf{I}) \cdot \mathbf{v} = \mathbf{0}$  mit unbekanntem Vektor  $\mathbf{v}$ .  $E_{\lambda}$  ist gerade der entsprechende Lösungsraum (d.h. die Menge aller Lösungen  $\mathbf{v} \in \mathbb{C}^n$ ) und seine Dimension die geometrische Vielfachheit  $g(\lambda)$ .

## 6.2 Normalform allgemeiner Matrizen

Nicht alle Matrizen sind diagonalisierbar, denn es kann Eigenwerte  $\lambda$  geben, für die  $1 \leq g(\lambda) < a(\lambda)$  gilt. Wir wollen zunächst anhand einfacher Beispiele verstehen, was dies bedeutet.

**Vorbemerkung** Die einfachste (3,3)-Diagonalmatrix ist

$$\mathbf{D} = \begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda \end{pmatrix} \quad \text{mit} \quad p_{\mathbf{D}}(z) = (\lambda - z)^3,$$

d.h.  $\lambda \in \mathbb{C}$  ist algebraisch dreifacher Eigenwert mit  $a(\lambda) = 3$ . Wie bemerken außerdem, dass wegen

$$\ker(\mathbf{D} - \lambda \cdot \mathbf{I}) = \ker \mathbf{0} = \mathbb{C}^3 = \text{span}\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$$

die geometrische Vielfachheit von  $\lambda$  ebenfalls 3 ist und es einen dreidimensionalen komplexen Eigenraum gibt (nämlich ganz  $\mathbb{C}^3$ ). Insbesondere gilt für die kanonische Basis

$$\begin{aligned} \mathbf{e}_1 & \xrightarrow{\mathbf{D} - \lambda \mathbf{I}} \mathbf{0}, \\ \mathbf{e}_2 & \xrightarrow{\mathbf{D} - \lambda \mathbf{I}} \mathbf{0}, \\ \mathbf{e}_3 & \xrightarrow{\mathbf{D} - \lambda \mathbf{I}} \mathbf{0}, \end{aligned}$$

d.h. jeder der Vektoren  $\mathbf{e}_i$  ist Eigenvektor von  $\mathbf{D}$  und im Kern von  $\mathbf{D} - \lambda \cdot \mathbf{I}$ . Wir werden nun sehen, dass diese Beobachtung uns besser verstehen läßt, was eigentlich bei nicht-diagonalisierbaren Matrizen passiert.



**Jordan-Matrizen** Wir betrachten nun zum Vergleich

$$\mathbf{J} = \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix} \quad \text{mit} \quad p_{\mathbf{J}}(\lambda) = \begin{vmatrix} \lambda - z & 1 & 0 \\ 0 & \lambda - z & 1 \\ 0 & 0 & \lambda - z \end{vmatrix} = (\lambda - z)^3$$

als prototypisches Beispiel einer nicht-diagonalisierbaren  $(3, 3)$ -Matrix. In der Tat,  $\lambda$  ist wieder der einzige Eigenwert mit algebraischer Vielfachheit  $a(\lambda) = 3$ . Wegen

$$\mathbf{J} - \lambda \cdot \mathbf{I} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

ist aber diesmal die geometrische Vielfachheit  $g(\lambda) = 1$ , d.h. es gibt nur einen eindimensionalen komplexen Eigenraum (aufgespannt durch  $\mathbf{e}_1$ ). Durch Nachrechnen bzw. scharfes Hinsehen verifizieren wir nun die sogenannte Kettenbedingung

$$\begin{aligned} \mathbf{e}_1 & \xrightarrow{\mathbf{J} - \lambda \cdot \mathbf{I}} \mathbf{0}, \\ \mathbf{e}_2 & \xrightarrow{\mathbf{J} - \lambda \cdot \mathbf{I}} \mathbf{e}_1, \\ \mathbf{e}_3 & \xrightarrow{\mathbf{J} - \lambda \cdot \mathbf{I}} \mathbf{e}_2, \end{aligned}$$

bzw.

$$\begin{aligned} \ker(\mathbf{J} - \lambda \cdot \mathbf{I})^1 &= \text{span}\{\mathbf{e}_1\}, \\ \ker(\mathbf{J} - \lambda \cdot \mathbf{I})^2 &= \text{span}\{\mathbf{e}_1, \mathbf{e}_2\}, \\ \ker(\mathbf{J} - \lambda \cdot \mathbf{I})^3 &= \text{span}\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}. \end{aligned}$$

Mit anderen Worten:  $\mathbf{e}_2$  und  $\mathbf{e}_3$  sind zwar keine Eigenvektoren von  $\mathbf{J}$  mehr, aber immer noch verallgemeinerte Eigenvektoren (oder sogenannte Hauptvektoren), die zwar nicht im Kern von  $\mathbf{J} - \lambda \cdot \mathbf{I}$ , aber im Kern gewisser Potenzen von  $\mathbf{J} - \lambda \cdot \mathbf{I}$  enthalten sind.

**Verallgemeinerung** Ist  $\lambda$  ein Eigenwert von  $\mathbf{A}$  mit  $g(\lambda) < a(\lambda)$ , so muss man nicht nur den Kern von  $\mathbf{A} - \lambda \cdot \mathbf{I}$  berechnen, um die Eigenvektoren zu erhalten, sondern man muss zusätzlich den Kern der Potenzen von  $\mathbf{A} - \lambda \cdot \mathbf{I}$  studieren, da dort verallgemeinerte Eigenvektoren zu finden sind.

**Definition** Die Matrizen

$$\begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}, \quad \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix}, \quad \begin{pmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & \lambda \end{pmatrix}, \quad \begin{pmatrix} \lambda & 1 & 0 & 0 & 0 \\ 0 & \lambda & 1 & 0 & 0 \\ 0 & 0 & \lambda & 1 & 0 \\ 0 & 0 & 0 & \lambda & 1 \\ 0 & 0 & 0 & 0 & \lambda \end{pmatrix}$$

werden die 2-, 3-, 4- und 5-dimensionale Jordan-Matrix zum Eigenwert  $\lambda$  genannt. Analoge Formeln gibt es für die  $n$ -dimensionale Jordan-Matrix zum Eigenwert  $\lambda$  und das 1-dimensionale Analogon ist die  $(1, 1)$ -Matrix  $(\lambda)$ .

Jordan-Matrizen liefern nicht nur typische Beispiele für nicht-diagonalisierbare Matrizen, sondern — wie wir gleich sehen werden — auch die fundamentalen Bausteine, aus denen man jede andere Matrix in gewisser Weise zusammensetzen kann.

**Bemerkung** Für eine  $n$ -dimensionale Jordan-Matrix mit  $n > 1$  gilt (Übungsaufgabe) die Kettenbedingung

$$\mathbf{e}_n \xrightarrow{\mathbf{J}-\lambda\mathbf{I}} \mathbf{e}_{n-1} \xrightarrow{\mathbf{J}-\lambda\mathbf{I}} \dots \xrightarrow{\mathbf{J}-\lambda\mathbf{I}} \mathbf{e}_2 \xrightarrow{\mathbf{J}-\lambda\mathbf{I}} \mathbf{e}_1 \xrightarrow{\mathbf{J}-\lambda\mathbf{I}} \mathbf{0}$$

und damit auch

$$\ker(\mathbf{A} - \lambda \cdot \mathbf{I})^j = \text{span}\{\mathbf{e}_1, \dots, \mathbf{e}_j\}$$

für alle  $j = 1, \dots, n$ .

**Bemerkung** Ist  $\mathbf{J}$   $n$ -dimensionale Jordan-Matrix zum Eigenwert  $\lambda$ , so ist  $\mathbf{J} - \lambda \cdot \mathbf{I}$  Jordan-Matrix zum Eigenwert 0 und es gilt

$$(\mathbf{J} - \lambda \cdot \mathbf{I})^n = \mathbf{0}$$

im Sinne von  $(n, n)$ -Matrizen. Man sagt deswegen auch, die  $n$ -dimensionale Jordan-Matrix zum Eigenwert 0 ist nilpotent, eben weil die  $n$ -te Potenz die Nullmatrix ergibt.

## Jordansche Normalform

Wir wollen nun mit Hilfe der Jordan-Matrizen einen ganz zentralen Satz der Matrizen-theorie formulieren, der für jede beliebige quadratische Matrix  $\mathbf{A} \in \mathbb{C}^{(n,n)}$  gilt. Dazu bezeichnen wir mit

$$\lambda_1, \dots, \lambda_m \in \mathbb{C},$$

die paarweise verschiedenen Eigenwerte ( $1 \leq m \leq n$ ) von  $\mathbf{A}$  und mit

$$a(\lambda_k) \quad \text{bzw.} \quad g(\lambda_k)$$

die algebraische bzw. geometrische Vielfachheit des Eigenwerts  $\lambda_k$  ( $k = 1, \dots, m$ ). Insbesondere gilt stets

$$1 \leq g(\lambda_k) \leq a(\lambda_k) \leq n, \quad \sum_{k=1}^m a(\lambda_k) = n.$$

**Theorem (Jordansche Normalform)** Für jedes  $\mathbf{A}$  existieren Matrizen  $\mathbf{J}$  und  $\mathbf{S}$ , so dass die folgenden Aussagen erfüllt sind (siehe Bild):

1. Es gilt  $\mathbf{J} = \mathbf{S}^{-1} \cdot \mathbf{A} \cdot \mathbf{S}$ .
2.  $\mathbf{J}$  ist eine Blockmatrix

$$\mathbf{J} = \begin{pmatrix} \mathbf{J}^{(1)} & & \\ & \ddots & \\ & & \mathbf{J}^{(m)} \end{pmatrix}$$

aus  $m$  Blöcken, wobei jede Matrix  $\mathbf{J}^{(k)}$  der Jordan-Block zum Eigenwert  $\lambda_k$  genannt wird und die folgenden Eigenschaften aufweist:

- (a) Sie ist quadratisch mit  $a(\lambda_k)$  vielen Zeilen und Spalten.

(b) Sie ist selbst wieder Blockmatrix

$$\mathbf{J}^{(k)} = \begin{pmatrix} \mathbf{J}^{(k,1)} & & \\ & \ddots & \\ & & \mathbf{J}^{(k,l_k)} \end{pmatrix}.$$

(c) Jeder Unterblock  $\mathbf{J}^{(k,j)}$

(i) ist eine Jordan-Matrix (im Sinne der obigen Definition) der Dimension  $d^{(k,i)}$ , wobei der Entartungsfall  $d^{(k,i)} = 1$  zugelassen ist,

(ii) wird Jordan-Kästchen zum Eigenwert  $\lambda_k$  genannt.

(d) Es gilt  $l_k = g(\lambda_k)$ , d.h. die Anzahl der Jordan-Kästchen zum Eigenwert  $\lambda_k$  ist gerade die geometrische Vielfachheit von  $\lambda_k$ ,

(e) Es gilt  $\sum_{j=1}^{l_k} d^{(k,j)} = a(\lambda_k)$ , d.h. die Dimensionen aller Jordan-Kästchen zum Eigenwert  $\lambda_k$  summieren sich zur algebraischen Vielfachheit von  $\lambda_k$ .

3. Für jedes Jordan-Kästchen  $\mathbf{J}^{(k,j)}$  erfüllen die entsprechenden Spaltenvektoren  $\mathbf{v}_1^{(k,j)}, \dots, \mathbf{v}_{d^{(k,j)}}^{(k,j)}$  von  $\mathbf{S}$  die Kettenbedingungen

$$(\mathbf{A} - \lambda \cdot \mathbf{I}) \mathbf{v}_1^{(k,j)} = 0$$

sowie

$$(\mathbf{A} - \lambda \cdot \mathbf{I}) \cdot \mathbf{v}_2^{(k,j)} = \mathbf{v}_1^{(k,j)}, \quad \dots, \quad (\mathbf{A} - \lambda \cdot \mathbf{I}) \cdot \mathbf{v}_i^{(k,j)} = \mathbf{v}_{i-1}^{(k,j)}.$$

für alle  $i = 2, \dots, d^{(k,j)}$ .

Die Matrix  $\mathbf{J}$  wird Jordansche Normalform von  $\mathbf{A}$  genannt und jede Spalten von  $\mathbf{S}$  ist entweder Eigenvektor oder Hauptvektor zu einem Eigenwert von  $\mathbf{A}$ .

$$\mathbf{J} = \begin{pmatrix} \lambda_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \lambda_1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \lambda_1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \lambda_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \lambda_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \lambda_2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_2 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_2 & 1 \end{pmatrix}$$

Beispiel für die Jordansche Normalform einer (15,15)-Matrix mit zwei Jordanblöcken, wobei der erste Block (rot) zum Eigenwert  $\lambda_1$  aus 2 Jordankästchen der Länge 1 bzw. 3 besteht. Der zweite Block (blau) gehört zum Eigenwert  $\lambda_2$  und umfasst 5 Jordankästchen der Längen 1, 1, 2, 2 und 5. Insbesondere gilt  $a(\lambda_1) = 4$ ,  $g(\lambda_1) = 2$  sowie  $a(\lambda_2) = 11$ ,

$g(\lambda_1) = 5$ . Beachte, dass die Jordankästchen der Länge 1 in manchen Formulierungen des Theorems zu

Diagonalblöcken zusammengefasst werden (lila).

$$\mathbf{J} = \begin{pmatrix} \lambda_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \lambda_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \lambda_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \lambda_2 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \lambda_2 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \lambda_3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \lambda_4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_5 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_5 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \lambda_5 \end{pmatrix}$$

Zweites Beispiel für eine Jordansche Normalform mit einem Jordanblock bestehend aus 3 Kästchen und vier Jordanblöcken mit jeweils einem Kästchen.

**Merkregel** Jede quadratische komplexe Matrix kann nach einem geeigneten Basiswechsel vollständig in Jordankästchen zerlegt werden. Dabei ist die Matrix genau dann diagonalisierbar, wenn alle Jordankästchen die Dimension 1 besitzen.

### Bemerkung

1. Es gibt verschiedene Formulierungen des Theorems, die aber letztlich alle äquivalent sind.
2. Die Matrix  $\mathbf{J}$  ist bis auf Permutationen der Eigenwerte und der Jordankästchen eindeutig durch  $\mathbf{A}$  bestimmt. Die Matrix  $\mathbf{S}$  ist ebenfalls nicht eindeutig, aber eine vollständige Beschreibung aller Mehrdeutigkeiten ist recht aufwändig.
3. Wir können den Beweis des Satzes im Rahmen dieser Vorlesung nicht führen. Er beruht auf dem *Satz von Hamilton-Caley*. Dieses sehr tiefe Resultat besagt, dass mit

$$p_{\mathbf{A}}(z) = (-1)^n \cdot (z - \lambda_1)^{a(\lambda_1)} \cdot \dots \cdot (z - \lambda_m)^{a(\lambda_m)}$$

auch schon die Gleichheit

$$(\mathbf{A} - \lambda_1 \cdot \mathbf{I})^{a(\lambda_1)} \cdot \dots \cdot (\mathbf{A} - \lambda_m \cdot \mathbf{I})^{a(\lambda_m)} = \mathbf{0}$$

im Sinne quadratischer Matrizen gilt. Oder anders formuliert: Jede quadratische Matrix  $\mathbf{A}$  annulliert ihr eigenes charakteristisches Polynom!

**\*Beispiel** Für

$$\mathbf{A} = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$

gilt

$$p_{\mathbf{A}}(z) = (1 - z) \cdot (4 - z) - 2 \cdot 3 = 1 \cdot z \cdot z - 5 \cdot z - 2 = 1 \cdot z^2 - 5 \cdot z - 2 \cdot z^0$$

und die Erkenntnis von Hamilton-Caley kann im konkreten Fall einfach nachgerechnet werden (wobei wir  $\mathbf{A}^0 = \mathbf{I}$  verwenden):

$$\begin{aligned} 1 \cdot \mathbf{A}^2 - 5 \cdot \mathbf{A} - 2 \cdot \mathbf{A}^0 &= 1 \cdot \mathbf{A} \cdot \mathbf{A} - 5 \cdot \mathbf{A} - 2 \cdot \mathbf{I} \\ &= 1 \cdot \begin{pmatrix} 7 & 10 \\ 15 & 22 \end{pmatrix} - 5 \cdot \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} - 2 \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \end{aligned}$$

**Beispiel** Wir wollen die Jordansche Normalform der Matrix

$$\mathbf{A} = \begin{pmatrix} -3 & 15 & -11 \\ -9 & 21 & -9 \\ -9 & 15 & 0 \end{pmatrix}$$

bestimmen und berechnen zunächst die Nullstellen des charakteristischen Polynoms via

$$p_{\mathbf{A}}(z) = \begin{vmatrix} -3-z & 15 & -11 \\ -9 & 21-z & -9 \\ -9 & 15 & -z \end{vmatrix} = -z^3 + 18z^2 - 108z + 216 = -(z-6)^3.$$

Insbesondere gibt es in diesem (sehr einfachen) Beispiel nur den Eigenwert  $\lambda_1 = 6$  mit algebraischer Vielfachheit  $a(\lambda_1) = 3$  und wir schließen, dass es auch nur einen Jordanblock geben wird, der aber vielleicht aus mehreren Kästchen besteht. Um dies zu klären, bestimmen wir nun die geometrische Vielfachheit von  $\lambda_1$ , indem wir  $\ker \mathbf{A} - 6 \cdot \mathbf{I}$  durch Gauß-Elimination berechnen. Mit

$$\left( \begin{array}{ccc|c} -9 & 15 & -11 & 0 \\ -9 & 15 & -9 & 0 \\ -9 & 15 & -6 & 0 \end{array} \right) \cong \left( \begin{array}{ccc|c} -9 & 15 & -11 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 5 & 0 \end{array} \right) \cong \left( \begin{array}{ccc|c} -9 & 15 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right)$$

erhalten wir

$$\ker(\mathbf{A} - \lambda_1 \cdot \mathbf{I}) = \text{span} \left\{ \begin{pmatrix} 5 \\ 3 \\ 0 \end{pmatrix} \right\}$$

und damit die geometrische Vielfachheit  $g(\lambda_1) = 1$ . Insbesondere wissen wir nun schon, dass der Jordanblock nur aus einem Jordan-Kästchen der Dimension 3 bestehen kann. Es gilt also

$$\mathbf{J} = \begin{pmatrix} 6 & 1 & 0 \\ 0 & 6 & 1 \\ 0 & 0 & 6 \end{pmatrix}.$$

Wir müssen nun noch einen zulässigen Basiswechsel finden, wobei wir wegen der Kettenbedingung schon wissen, dass die erste Spalte von  $\mathbf{S}$  im Kern von  $\mathbf{A} - 6 \cdot \mathbf{I}$  liegen wird. Wir haben nun prinzipiell mehrere Möglichkeiten, eine geeignete Matrix  $\mathbf{S}$  zu finden:

1. Es muss  $\mathbf{S} \cdot \mathbf{J} - \mathbf{A} \cdot \mathbf{S} = \mathbf{0}$  gelten und dies liefert 9 lineare und homogene Gleichungen für die 9 Komponenten von  $\mathbf{S}$ , deren (nicht eindeutige Lösung) man zum Beispiel mit Gauß-Elimination berechnen kann. Diese Variante ist aber im Allgemeinen recht aufwändig und wenig elegant, da wir zunächst eine (9,9)-Koeffizientenmatrix berechnen und anschließend das entsprechende Gleichungssystem lösen müssen.
2. Unsere Rechnungen oben implizieren, dass

$$\mathbf{v}_1 = \begin{pmatrix} 5 \\ 3 \\ 0 \end{pmatrix}$$

eine mögliche Wahl des ersten Spaltenvektors von  $\mathbf{S}$  ist, denn es gilt

$$(\mathbf{A} - 6 \cdot \mathbf{I}) \cdot \mathbf{v}_1 = \mathbf{0}.$$

Wir können nun mittels der Kettenbedingung  $\mathbf{v}_2$  als Lösung der Gleichung

$$(\mathbf{A} - 6 \cdot \mathbf{I}) \cdot \mathbf{v}_2 = \mathbf{v}_1$$

und anschließend  $\mathbf{v}_3$  als Lösung von

$$(\mathbf{A} - 6 \cdot \mathbf{I}) \cdot \mathbf{v}_3 = \mathbf{v}_2$$

bestimmen (es wird aber keine eindeutige Lösung geben). In dieser Variante müssen wir neben der bereits gelösten homogenen Gleichung zusätzlich noch zwei inhomogene Gleichungssysteme lösen, aber die  $(3, 3)$ -Koeffizientenmatrix ist jedesmal dieselbe.

3. Wir können schließlich alternativ auch die Kerne

$$\ker(\mathbf{A} - 6 \cdot \mathbf{I})^j \quad \text{für} \quad j = 1, 2, 3$$

bestimmen, wobei wir schon wissen, dass der erste bzw. zweite bzw. dritte Kern die Dimension 1 bzw. 2 bzw. 3 besitzt. Insbesondere können wir mittels der allgemeinen Lösungen drei verschiedener homogener Gleichungssysteme sukzessive drei linear unabhängige Vektoren  $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$  identifizieren, so dass

$$\begin{aligned} \ker(\mathbf{A} - 6 \cdot \mathbf{I})^1 &= \text{span}\{\mathbf{v}_1\}, \\ \ker(\mathbf{A} - 6 \cdot \mathbf{I})^2 &= \text{span}\{\mathbf{v}_1, \mathbf{v}_2\}, \\ \ker(\mathbf{A} - 6 \cdot \mathbf{I})^3 &= \text{span}\{\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3\}. \end{aligned}$$

Mit den letzten beiden Methoden können geeignete Vektoren  $\mathbf{v}_1, \mathbf{v}_2, \mathbf{v}_3$  als Spalten von  $\mathbf{S}$  konstruiert werden. Im konkreten Fall ergibt sich zum Beispiel

$$\mathbf{S} = \begin{pmatrix} 15 & 2 & 1 \\ 9 & 0 & 0 \\ 0 & -3 & -1 \end{pmatrix}, \quad \mathbf{S}^{-1} = \frac{1}{9} \begin{pmatrix} 0 & 1 & 0 \\ -9 & 15 & -9 \\ 27 & -45 & 18 \end{pmatrix},$$

aber das ist nicht die einzige mögliche Wahl.

**Bemerkung** Im Allgemeinen ist die Berechnung der Jordanschen Normalform recht aufwändig und erfordert die Lösung großer und/oder vieler linearer Gleichungssysteme. Heutzutage übernehmen die Aufgabe natürlich meist Computer.

**Beispiel** Die MATHEMATICA-Befehle `JordanDecomposition` und `Inverse` liefern für

$$\mathbf{A} = \begin{pmatrix} 0 & 0 & -2 & -1 & 0 & 3 \\ 0 & 1 & 0 & -1 & 0 & 0 \\ -1 & 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 2 & 0 & 0 \\ -1 & -2 & -1 & -2 & 3 & 2 \\ -2 & 0 & -2 & 0 & 0 & 5 \end{pmatrix}$$

zum Beispiel

$$\mathbf{J} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & 1 \\ 0 & 0 & 0 & 0 & 0 & 3 \end{pmatrix}$$

sowie

$$\mathbf{S} = \begin{pmatrix} 1 & 0 & -1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}, \quad \mathbf{S}^{-1} = \begin{pmatrix} 1 & 0 & 1 & 0 & 0 & -1 \\ 0 & 1 & 0 & 1 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & -1 & 1 & 0 \\ -1 & 0 & -1 & 0 & 0 & 2 \end{pmatrix}.$$

Insbesondere gibt es in diesem Beispiel 3 Jordan-Blöcke (da drei verschiedene Eigenwerte) mit insgesamt 4 Jordan-Kästchen. Die Befehle `CharacteristicPolynomial` und `FactorList` bzw. `Solve` sind auch nützlich und liefern in diesem Beispiel

$$p_{\mathbf{A}}(z) = (z - 1)^2 \cdot (z - 2)^2 \cdot (z - 3)^2.$$

**Bemerkung** Weitere Beispiele finden sich in [AORS, Seite 164 ff] und in anderen Büchern.

Vorlesung 19, 02. Dezember 2019

## 6.3 Hauptachsentransformation

**Erinnerung** (Skalarprodukt in  $\mathbb{C}^n$ )

1.  $\langle \mathbf{v}, \mathbf{w} \rangle = \sum_{j=1}^n v_j \cdot \bar{w}_j = \overline{\langle \mathbf{w}, \mathbf{v} \rangle}, \quad \langle \mathbf{v}, \mathbf{w} \rangle = \mathbf{v}^T \cdot \bar{\mathbf{w}} = \mathbf{w}^* \cdot \mathbf{v}$
2. Skalarprodukte sind im Komplexen sesquilinear, d.h.

$$\langle \lambda \cdot \mathbf{v} + \tilde{\lambda} \cdot \tilde{\mathbf{v}}, \mathbf{w} \rangle = \lambda \cdot \langle \mathbf{v}, \mathbf{w} \rangle + \tilde{\lambda} \cdot \langle \tilde{\mathbf{v}}, \mathbf{w} \rangle$$

aber

$$\langle \mathbf{v}, \mu \cdot \mathbf{w} + \tilde{\mu} \cdot \tilde{\mathbf{w}} \rangle = \bar{\mu} \cdot \langle \mathbf{v}, \mathbf{w} \rangle + \bar{\tilde{\mu}} \cdot \langle \mathbf{v}, \tilde{\mathbf{w}} \rangle.$$

3.  $\|\mathbf{v}\| = \sqrt{\langle \mathbf{v}, \mathbf{v} \rangle} = \sqrt{\sum v_i \cdot \bar{v}_i}$
4.  $\mathbf{A}^* = \bar{\mathbf{A}}^T$
5.  $\langle \mathbf{A} \cdot \mathbf{v}, \mathbf{w} \rangle = (\mathbf{A} \cdot \mathbf{v})^T \cdot \bar{\mathbf{w}} = \mathbf{v}^T \cdot \mathbf{A}^T \cdot \bar{\mathbf{w}} = \mathbf{v}^T \cdot \overline{(\mathbf{A}^* \cdot \mathbf{w})} = \langle \mathbf{v}, \mathbf{A}^* \cdot \mathbf{w} \rangle$

**Bemerkung** Die genaue Definition von sesquilinear variiert in der Literatur. In einigen Büchern benutzt man die Formel  $\langle \mathbf{v}, \mathbf{w} \rangle = \sum_{j=1}^n \bar{v}_j \cdot w_j$ , so dass das Skalarprodukt linear im zweiten Argument  $\mathbf{w}$  und anti-linear im ersten Argument  $\mathbf{v}$  ist.

**Erinnerung**

1. Eine komplexe Matrix  $\mathbf{A}$  heißt unitär, falls  $\mathbf{A} \cdot \mathbf{A}^* = \mathbf{A}^* \cdot \mathbf{A} = \mathbf{I}$  bzw.  $\mathbf{A}^* = \mathbf{A}^{-1}$ , und hermitesch, falls  $\mathbf{A}^* = \mathbf{A}$ . Insbesondere ist eine reelle Matrix genau dann unitär bzw. hermitesch, falls sie orthogonal bzw. symmetrisch ist.
2. Analog zum Reellen zeigt man:  $\mathbf{A}$  ist genau dann unitär, wenn die Spaltenvektoren (oder die Zeilenvektoren) von  $\mathbf{A}$  eine ON-Basis in  $\mathbb{C}^n$  bilden.

**Beispiele** Die Matrizen

$$\frac{1}{\sqrt{3}} \begin{pmatrix} \mathbf{i} & 1 - \mathbf{i} \\ 1 + \mathbf{i} & \mathbf{i} \end{pmatrix} \quad \text{und} \quad \frac{1}{2} \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & \mathbf{i} & -1 & -\mathbf{i} \\ 1 & -1 & 1 & -1 \\ 1 & -\mathbf{i} & -1 & \mathbf{i} \end{pmatrix}$$

sind unitär; die Matrizen

$$\begin{pmatrix} 3 & 2 + \mathbf{i} \\ 2 - \mathbf{i} & 4 \end{pmatrix} \quad \text{und} \quad \begin{pmatrix} -2 & 1 + \mathbf{i} & 1 + 4\mathbf{i} & 2 \\ 1 - \mathbf{i} & 2 & 1 & -2\mathbf{i} \\ 1 - 4\mathbf{i} & 1 & 4 & 1 - \mathbf{i} \\ 2 & 2\mathbf{i} & 1 + \mathbf{i} & -2 \end{pmatrix}$$

sind hermitesch.

**Theorem (Spektrum hermitescher Matrizen)** Für jede hermitesche Matrix gilt:

1. Alle Eigenwerte sind reell.
2. Eigenvektoren zu verschiedenen Eigenwerten stehen senkrecht aufeinander.

Beweis, Teil 1: Aus den obigen Rechenregeln und wegen  $\mathbf{A} = \mathbf{A}^*$  folgen für jeden Eigenvektor  $\mathbf{v} \neq \mathbf{0}$  zum Eigenwert  $\lambda$  die Formeln

$$\lambda \cdot \|\mathbf{v}\|^2 = \langle \lambda \cdot \mathbf{v}, \mathbf{v} \rangle = \langle \mathbf{A} \cdot \mathbf{v}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{A}^* \cdot \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{A} \cdot \mathbf{v} \rangle = \langle \mathbf{v}, \lambda \cdot \mathbf{v} \rangle = \bar{\lambda} \cdot \|\mathbf{v}\|^2,$$

und wir erhalten  $\lambda = \bar{\lambda}$  und damit  $\lambda \in \mathbb{R}$  wegen  $\|\mathbf{v}\| \neq 0$ .

Beweis, Teil 2: Mit

$$\mathbf{A} \cdot \mathbf{v} = \lambda \cdot \mathbf{v}, \quad \mathbf{A} \cdot \mathbf{w} = \mu \cdot \mathbf{w}, \quad \bar{\mu} = \mu \neq \lambda = \bar{\lambda}, \quad \mathbf{v} \neq \mathbf{0} \neq \mathbf{w}$$

ergibt sich

$$\lambda \cdot \langle \mathbf{v}, \mathbf{w} \rangle = \langle \lambda \cdot \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{A} \cdot \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{v}, \mathbf{A} \cdot \mathbf{w} \rangle = \langle \mathbf{v}, \mu \cdot \mathbf{w} \rangle = \bar{\mu} \cdot \langle \mathbf{v}, \mathbf{w} \rangle,$$

und deshalb muss  $\langle \mathbf{v}, \mathbf{w} \rangle = 0$  gelten. □



**Theorem (Hauptachsen-Transformation hermitescher Matrizen)** Jede hermitesche Matrix  $\mathbf{A}$  ist diagonalisierbar. Insbesondere existieren eine reelle Diagonalmatrix  $\mathbf{\Lambda}$  sowie eine unitäre Matrix  $\mathbf{S}$ , so dass

$$\mathbf{\Lambda} = \mathbf{S}^{-1} \cdot \mathbf{A} \cdot \mathbf{S}.$$

Dabei sind die Diagonaleinträge von  $\mathbf{\Lambda}$  gerade die (reellen) Eigenwerte von  $\mathbf{A}$  und die Spalten von  $\mathbf{S}$  bestehen aus Eigenvektoren.

*Beweis, Teil 1:* Angenommen,  $\mathbf{A}$  wäre nicht diagonalisierbar. Dann gibt es nach dem Theorem über die Jordansche Normalform mindestens einen (reellen) Eigenwert  $\lambda$  mit  $1 \leq g(\lambda) < a(\lambda)$  sowie einen verallgemeinerten Eigenvektor  $\mathbf{v}$ , so dass

$$(\mathbf{A} - \lambda \cdot \mathbf{I})^2 \cdot \mathbf{v} = \mathbf{0}, \quad (\mathbf{A} - \lambda \cdot \mathbf{I}) \cdot \mathbf{v} \neq \mathbf{0}.$$

Da  $\mathbf{A} - \lambda \cdot \mathbf{I}$  auch hermitesch ist (Nachrechnen!) gilt nun

$$\begin{aligned} \mathbf{0} &= \langle (\mathbf{A} - \lambda \cdot \mathbf{I}) \cdot (\mathbf{A} - \lambda \cdot \mathbf{I}) \cdot \mathbf{v}, \mathbf{v} \rangle \\ &= \langle (\mathbf{A} - \lambda \cdot \mathbf{I}) \cdot \mathbf{v}, (\mathbf{A} - \lambda \cdot \mathbf{I}) \cdot \mathbf{v} \rangle \\ &= \|(\mathbf{A} - \lambda \cdot \mathbf{I}) \cdot \mathbf{v}\|^2 \end{aligned}$$

und wir erhalten mit  $(\mathbf{A} - \lambda \cdot \mathbf{I}) \cdot \mathbf{v} = \mathbf{0}$  den gewünschten Widerspruch. Insbesondere ist  $\mathbf{A}$  diagonalisierbar und es gilt (siehe den 2. Diagonalisierbarkeitssatz)  $g(\lambda) = a(\lambda)$  für jeden Eigenwert  $\lambda$  von  $\mathbf{A}$ .

*Beweis, Teil 2:* Die Matrix  $\mathbf{A}$  besitzt  $m$  verschiedene Eigenwerte  $\lambda_1, \dots, \lambda_m \in \mathbb{R}$ , wobei

$$\sum_{j=1}^m g(\lambda_j) = \sum_{j=1}^m a(\lambda_j) = n$$

gilt. Wir wählen nun für jedes  $j = 1, \dots, m$  genau  $g(\lambda_j)$  linear unabhängige Eigenvektoren

$$\mathbf{v}_1^{(j)}, \dots, \mathbf{v}_{g(\lambda_j)}^{(j)},$$

wobei wir O.B.d.A. annehmen dürfen, dass diese eine ON-Basis im Eigenraum  $E_{\lambda_j}$  bilden, denn wir können aus jeder anderen Basis mit Hilfe des Gram-Schmidt-Verfahrens eine solche konstruieren. Der zweite Teil des Satzes über das Spektrum hermitescher Matrizen garantiert, dass die  $n$  Vektoren

$$\mathbf{v}_1^{(1)}, \dots, \mathbf{v}_{g(\lambda_1)}^{(1)}, \quad \dots, \quad \mathbf{v}_1^{(m)}, \dots, \mathbf{v}_{g(\lambda_m)}^{(m)}$$

eine komplexe ON-Basis von  $\mathbb{C}^n$  bilden bzw. dass die spaltenweise aus diesen Vektoren gebildete Matrix  $\mathbf{S}$  hermitesch ist. Die behauptete Diagonalisierbarkeit folgt nun analog zum Beweis des ersten Diagonalisierungssatzes, sofern die  $\lambda_j$  der Reihe nach und gemäß ihrer Vielfachheit in die Diagonaleinträge von  $\mathbf{\Lambda}$  geschrieben werden.  $\square$

### Bemerkung

1. Die Hauptachsentransformation ist ein Spezialfall des Theorems über die Jordansche Normalform (im Sinne von  $\mathbf{J} = \mathbf{\Lambda}$ ).

2. Die Matrix  $\mathbf{A}$  ist wieder bis auf Permutationen der Eigenwerte bzw. Diagonalelemente eindeutig bestimmt. Die Matrix  $\mathbf{S}$  besitzt weitere Mehrdeutigkeiten, da es viele ON-Basen im Eigenraum  $E_{\lambda_j}$  gibt (jedenfalls sofern  $\dim E_{\lambda_j} = g(\lambda_j) > 1$ ).

3. Beachte, dass nicht jede diagonalisierbare Matrix hermitesch ist. Ein einfaches Gegenbeispiel ist

$$\mathbf{A} = \begin{pmatrix} 1 & 0 \\ 1 & -1 \end{pmatrix}, \quad \mathbf{\Lambda} = \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \mathbf{S} = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 2 \\ 1 & 1 \end{pmatrix}, \quad \mathbf{S}^{-1} = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 & 2 \\ 1 & 0 \end{pmatrix},$$

wobei aber  $\mathbf{A}$  nicht hermitesch und  $\mathbf{S}$  nicht unitär ist.

4. Es gilt die folgende partielle Umkehrung (Übungsaufgabe): Gibt es für  $\mathbf{A}$  eine unitäre Matrix  $\mathbf{S}$ , so dass  $\mathbf{S}^{-1} \cdot \mathbf{A} \cdot \mathbf{S}$  eine reelle Diagonalmatrix ist, so ist  $\mathbf{A}$  hermitesch.

5. Ist  $\mathbf{A}$  reellwertig und symmetrisch, so kann  $\mathbf{S}$  immer als reelle und orthogonale Matrix gewählt werden, eben weil es dann immer auch reelle Eigenvektoren gibt.

**Beispiel** Die hermitesche Matrix

$$\mathbf{A} = \begin{pmatrix} 2 & 1 + \mathbf{i} \\ 1 - \mathbf{i} & 1 \end{pmatrix}$$

besitzt (Nachrechnen!) die jeweils einfachen Eigenwerte

$$\lambda_1 = 0, \quad \lambda_2 = 3$$

sowie die entsprechenden Eigenräume

$$E_0 = \text{span} \left\{ \begin{pmatrix} -1 - \mathbf{i} \\ 2 \end{pmatrix} \right\}, \quad E_3 = \text{span} \left\{ \begin{pmatrix} 1 + \mathbf{i} \\ 1 \end{pmatrix} \right\}.$$

Insbesondere gilt das Theorem über die Hauptachsen-Transformation mit

$$\mathbf{A} = \begin{pmatrix} 0 & 0 \\ 0 & 3 \end{pmatrix}, \quad \mathbf{S} = \frac{1}{\sqrt{6}} \begin{pmatrix} -1 - \mathbf{i} & (1 + \mathbf{i})\sqrt{2} \\ 2 & \sqrt{2} \end{pmatrix}, \quad \mathbf{S}^{-1} = \frac{1}{\sqrt{6}} \begin{pmatrix} -1 + \mathbf{i} & 2 \\ (1 - \mathbf{i})\sqrt{2} & \sqrt{2} \end{pmatrix},$$

wobei  $\mathbf{S}^{-1} = \mathbf{S}^*$ .

**Beispiel** Wir wollen die Hauptachsentransformation für die symmetrische Matrix

$$\mathbf{A} = \begin{pmatrix} 1 & 3 & 0 \\ 3 & -2 & -1 \\ 0 & -1 & 1 \end{pmatrix}$$

durchführen und berechnen zunächst ihr charakteristisches Polynom zu

$$p_{\mathbf{A}}(z) = (1 - z) \cdot (-2 - z) \cdot (1 - z) - 1 - 9 = -z^3 + 13 \cdot z - 12.$$

Durch scharfes Hinsehen bemerken wir  $p_{\mathbf{A}}(1) = 0$  und erhalten nach Polynomdivision sowie quadratischer Ergänzung

$$p_{\mathbf{A}}(z) = -(z - 1) \cdot (z - 3) \cdot (z + 4)$$

und damit die drei Eigenwerte

$$\lambda_1 = 1, \quad \lambda_2 = 3, \quad \lambda_3 = -4,$$

deren algebraische und geometrische Vielfachheit jeweils 1 ist. Wir berechnen nun für  $\lambda_1$  mittels der Gauß-Elimination

$$\left( \begin{array}{ccc|c} 0 & 3 & 0 & 0 \\ 3 & -3 & -1 & 0 \\ 0 & -1 & 0 & 0 \end{array} \right) \cong \left( \begin{array}{ccc|c} 3 & -3 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 3 & 0 & 0 \end{array} \right) \cong \left( \begin{array}{ccc|c} 3 & 0 & -1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right)$$

den Eigenvektor

$$\mathbf{v}_1 = \frac{1}{\sqrt{10}} \begin{pmatrix} 1 \\ 0 \\ 3 \end{pmatrix},$$

wobei wir  $\|\mathbf{v}_1\| = 1$  durch geeignete Wahl des Vorfaktors sichergestellt haben. Anschließend berechnen wir den zweiten Eigenvektor via

$$\begin{aligned} \left( \begin{array}{ccc|c} -2 & 3 & 0 & 0 \\ 3 & -5 & -1 & 0 \\ 0 & -1 & -2 & 0 \end{array} \right) &\cong \left( \begin{array}{ccc|c} 1 & -2 & -1 & 0 \\ 3 & -5 & -1 & 0 \\ 0 & 1 & 2 & 0 \end{array} \right) \\ &\cong \left( \begin{array}{ccc|c} 1 & -2 & -1 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 1 & 2 & 0 \end{array} \right) \cong \left( \begin{array}{ccc|c} 1 & 0 & 3 & 0 \\ 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right) \end{aligned}$$

und geeigneter Normierung zu

$$\mathbf{v}_2 = \frac{1}{\sqrt{14}} \begin{pmatrix} 3 \\ 2 \\ -1 \end{pmatrix},$$

Wir könnten nun analog  $\mathbf{v}_3$  konstruieren, aber alternativ können wir auch die Formel

$$\mathbf{v}_3 = \mathbf{v}_1 \times \mathbf{v}_2 = \frac{1}{\sqrt{10} \cdot \sqrt{14}} \begin{pmatrix} 0 \cdot (-1) - 3 \cdot 2 \\ 3 \cdot 3 - 1 \cdot (-1) \\ 1 \cdot 2 - 0 \cdot 3 \end{pmatrix} = \frac{1}{2 \cdot \sqrt{35}} \begin{pmatrix} -6 \\ 10 \\ 2 \end{pmatrix} = \frac{1}{\sqrt{35}} \begin{pmatrix} -3 \\ 5 \\ 1 \end{pmatrix}$$

verwenden, weil wir ja schon wissen, dass  $\mathbf{v}_3$  senkrecht auf  $\mathbf{v}_1$  und  $\mathbf{v}_2$  stehen muss. Insgesamt ergibt sich

$$\lambda = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -4 \end{pmatrix}, \quad \mathbf{S} = \begin{pmatrix} | & | & | \\ \mathbf{v}_1 & \mathbf{v}_2 & \mathbf{v}_3 \\ | & | & | \end{pmatrix} \quad \mathbf{S}^{-1} = \mathbf{S}^T = \begin{pmatrix} - & \mathbf{v}_1^T & - \\ - & \mathbf{v}_2^T & - \\ - & \mathbf{v}_3^T & - \end{pmatrix}.$$

**\*Reelle Quadriken** Eine Funktion  $q: \mathbb{R}^n \rightarrow \mathbb{R}$  der Bauart

$$q(\mathbf{x}) = \mathbf{x}^T \cdot \mathbf{A} \cdot \mathbf{x} + \mathbf{b}^T \cdot \mathbf{x} + c$$

heißt quadratisches Polynom in den Variablen  $x_1, \dots, x_n$ , wobei die symmetrische Matrix  $\mathbf{A} = \mathbf{A}^T \in \mathbb{R}^{(n,n)}$ , der Vektor  $\mathbf{b} \in \mathbb{R}$  und der Skalar  $c \in \mathbb{R}$  gegeben sind.

Die Nullstellenmenge von  $q$  wird die entsprechende Quadrik genannt, siehe die Große Übung bzw. [AORS, Seite 173ff] für Beispiele.

Durch die Hauptachsen-Transformation

$$\mathbf{S}^{-1} \cdot \mathbf{A} \cdot \mathbf{S} = \mathbf{\Lambda} =: \tilde{\mathbf{A}} \quad \text{sowie} \quad \tilde{\mathbf{b}} := \mathbf{S}^{-1} \cdot \mathbf{b} = \mathbf{S}^T \cdot \mathbf{b}$$

wird ein neues quadratisches Polynom

$$\tilde{q}(\tilde{\mathbf{x}}) := \tilde{\mathbf{x}}^T \cdot \tilde{\mathbf{A}} \cdot \tilde{\mathbf{x}} + \tilde{\mathbf{b}}^T \cdot \tilde{\mathbf{x}} + c$$

definiert, und mittels der Koordinatentransformationen

$$\mathbf{x} = \mathbf{S} \cdot \tilde{\mathbf{x}} \quad \tilde{\mathbf{x}} = \mathbf{S}^{-1} \cdot \mathbf{x}$$

kann leicht gezeigt werden, dass

$$q(\mathbf{x}) = 0 \quad \iff \quad \tilde{q}(\tilde{\mathbf{x}}) = 0.$$

Oder anders gesagt: Die Quadriken von  $q$  und  $\tilde{q}$  unterscheiden sich nur durch eine orthogonale Abbildung (zum Beispiel Drehung oder Spiegelung). Diese Beobachtung ist sehr nützlich und erlaubt es, Quadriken anhand von Normalformen zu klassifizieren, siehe wieder [AORS, Seite 173ff], wobei dann meist noch eine geeignete Verschiebung eingebaut wird.

# Kapitel 7

## Folgen und Reihen

Vorlesung 20, 06. Dezember 2019

### 7.1 Folgen

**Definition** Eine (komplexe Zahlen-) Folge ist eine Abbildung  $a : \mathbb{N} \rightarrow \mathbb{C}$ , aber wir schreiben meist

$$a_n \quad \text{statt} \quad a(n)$$

für die Glieder oder Elemente der Folge. Die Folge selbst wird mit  $a$  oder  $(a_n)_{n \in \mathbb{N}}$  bezeichnet.

Gehören alle Folgenglieder zu einer Menge  $M \subset \mathbb{C}$ , so schreiben wir auch  $(a_n)_{n \in \mathbb{N}} \subset M$  (andere Autoren benutzen die Abkürzung  $a \in M^{\mathbb{N}}$ ). Zum Beispiel gilt  $(a_n)_{n \in \mathbb{N}} \subset \mathbb{R}$  für jede reelle Zahlenfolge.

**Bemerkung** Es gibt mehrere Arten, eine Folge zu beschreiben:

1. verbal, zum Beispiel „die Folge der Quadratzahlen“,
2. durch eine explizite Formel, wie zum Beispiel

$$a_n = (1 + i \cdot n)^3 \quad \text{für alle } n \in \mathbb{N}. \quad (7.1)$$

3. rekursiv, via

$$a_1 = 0, \quad a_{n+1} = f(a_n) \quad \text{für alle } n \geq 1,$$

wobei  $f : \mathbb{C} \rightarrow \mathbb{C}$  die sogenannte Iterationsvorschrift ist.

4. salopp, wie zum Beispiel

$$a = (2, 4, 6, 8, \dots),$$

für die Folge der geraden Zahlen, obwohl diese Notation streng genommen missverständlich ist.

5. Manchmal wählt man einen anderen Startindex, d.h. man betrachtet Folgen  $(a_n)_{n \geq n_1} = (a_{n_1}, a_{n_1+1}, a_{n_1+2} \dots)$ .

**Definition** Ist  $(a_n)_{n \in \mathbb{N}}$  eine Folge und  $(n_j)_{j \in \mathbb{N}} \subset \mathbb{N}$  eine Folge natürlicher Zahlen mit

$$n_1 < n_2 < n_3 < \dots,$$

so wird die Folge  $(a_{n_j})_{j \in \mathbb{N}}$  auch Teilfolge von  $(a_n)_{n \in \mathbb{N}}$  genannt.

### Beispiel

1. Die Folge der Quadratzahlen  $(1, 4, 9, 16, \dots)$  ist Teilfolge der Folge der natürlichen Zahlen  $(1, 2, 3, 4, \dots)$ , denn mit  $a_n = n$  und  $n_j = j^2$  gilt  $a_{n_j} = j^2$ .
2. Die Folge der Primzahlen  $(2, 3, 5, 7, \dots)$  ist auch Teilfolge der Folge aller natürlichen Zahlen, obwohl es keine geschlossene Formel für  $n_j$  gibt.
3. Die *geometrische* Folge  $(1, \frac{1}{2}, \frac{1}{4}, \frac{1}{8}, \frac{1}{16}, \dots)$  ist Teilfolge der *harmonischen* Folge  $(1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \dots)$ .

**Definition** Eine Folge  $a : \mathbb{N} \rightarrow \mathbb{C}$  heißt

1. beschränkt, falls

$$\exists C > 0 \quad \forall n \in \mathbb{N} \quad : \quad |a_n| \leq C.$$

2. konvergent mit Grenzwert  $a_\infty \in \mathbb{C}$ , falls

$$\forall \varepsilon > 0 \quad \exists N \in \mathbb{N} \quad \forall n \geq N \quad : \quad |a_n - a_\infty| < \varepsilon,$$

wobei  $N$  von  $\varepsilon$  abhängen darf (und meist wird). Wir schreiben

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

oder ganz kurz  $a_n \rightarrow a_\infty$ .

3. konvergent, falls sie gegen einen Grenzwert konvergiert, und divergent, falls sie nicht konvergiert,

4. Cauchy-Folge, falls

$$\forall \varepsilon > 0 \quad \exists N \in \mathbb{N} \quad \forall n_1, n_2 \geq N \quad : \quad |a_{n_1} - a_{n_2}| < \varepsilon,$$

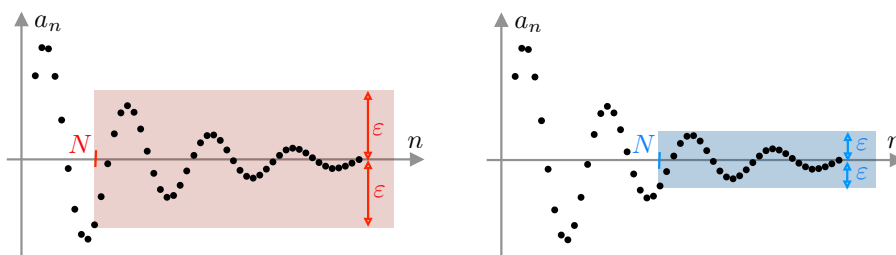
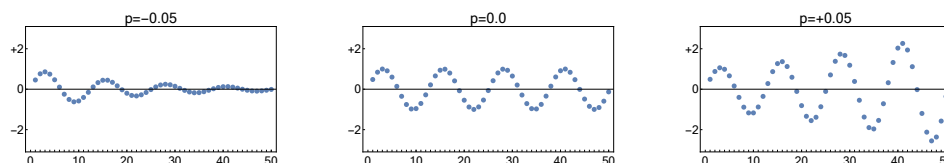


Illustration der Konvergenz von Folgen am Beispiel einer reellen Nullfolge  $a_n \rightarrow a_\infty = 0$ . Je kleiner  $\varepsilon > 0$  ist, desto größer muss  $N$  gewählt werden.



Die Folge  $a_n = \exp(p \cdot n) \cdot \sin(n/2)$  für verschiedene Werte des Parameters  $p$ . *Links*: konvergent und damit beschränkt. *Mitte*: beschränkt, aber nicht konvergent. *Rechts*: weder konvergent noch beschränkt.

**Beispiele**

1. Die harmonische Folge  $a_n = 1/n$  konvergiert gegen  $a_\infty = 0$ , denn für jedes  $\varepsilon > 0$  können wir  $N$  mit  $N > 1/\varepsilon$  wählen und erhalten

$$|a_n - a_\infty| = a_n = 1/n \leq 1/N < \varepsilon$$

für alle  $n \geq N$ .

2. Die komplexe Folge  $a_n = 1/(\ln n) + i \cdot (1 + (-1)^n/n)$  konvergiert gegen  $a_\infty = i$ , denn wegen  $\ln n \leq n$  gilt

$$|a_n - a_\infty| \leq \sqrt{\frac{1}{(\ln n)^2} + \frac{1}{n^2}} \leq \sqrt{\frac{2}{(\ln n)^2}}$$

für alle  $n \in \mathbb{N}$  und damit auch

$$|a_n - a_\infty| \leq \frac{\sqrt{2}}{\ln N} < \varepsilon$$

für alle  $n \geq N$ , sofern  $N > \exp(\sqrt{2}/\varepsilon)$ .

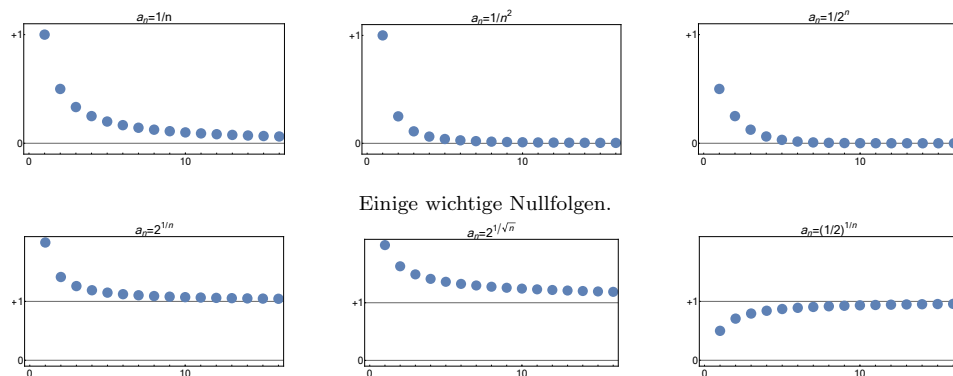
3. Für jeden reellen Parameter  $p > 0$  sind die vier Folgen

$$a_n = \frac{1}{n^p}, \quad a_n = \frac{1}{(\ln n)^p}, \quad a_n = \frac{1}{(\ln \ln n)^p}, \quad a_n = \frac{1}{(1+p)^n}$$

alle Nullfolgen, d.h. es gilt stets  $a_n \rightarrow a_\infty = 0$ . Die vier Folgen

$$a_n = p^{1/n^2}, \quad a_n = p^{1/n}, \quad a_n = p^{1/\sqrt{n}}, \quad a_n = p^{1/\ln n}$$

konvergieren hingegen alle gegen  $a_\infty = 1$ .



Einige Folgen, die gegen 1 konvergieren.

**Achtung** Es ist in aller Regel alles andere als einfach, die Konvergenz einer gegebenen Folge mit Hilfe der Definition nachzuweisen. Wir werden daher meist versuchen, eine oder mehrere der folgenden Regeln anzuwenden und die Grenzwerte bekannter Folgen einzusetzen.

**Rechenregeln für Grenzwerte** Für zwei konvergente Folgen  $a_n \rightarrow a_\infty$  und  $b_n \rightarrow b_\infty$  gilt:

1.  $\lambda \cdot a_n + \mu \cdot b_n \rightarrow \lambda \cdot a_\infty + \mu \cdot b_\infty$  für alle  $\lambda, \mu \in \mathbb{C}$
2.  $a_n \cdot b_n \rightarrow a_\infty \cdot b_\infty$  und  $a_n/b_n \rightarrow a_\infty/b_\infty$ , sofern  $b_\infty \neq 0$
3.  $f(a_n) \rightarrow f(a_\infty)$  für jede *stetige* Funktion  $f$  (siehe dazu auch das nächste Kapitel)

Außerdem implizieren unsere Definitionen die folgenden Aussagen:

1. Es gilt  $a_n \rightarrow a_\infty$  genau dann, wenn  $|a_n - a_\infty| \rightarrow 0$ .
2. Für eine komplexe Folge gilt  $a_n \rightarrow a_\infty$  genau dann, wenn  $\operatorname{Re}(a_n) \rightarrow \operatorname{Re}(a_\infty)$  und  $\operatorname{Im}(a_n) \rightarrow \operatorname{Im}(a_\infty)$  (Übungsaufgabe).
3. Jede Teilfolge einer konvergenten Folge konvergiert auch gegen den Grenzwert der Folge.

### Beispiele

1.  $\lim_{n \rightarrow \infty} (-2/n + 3/n^2) = -2 \cdot \lim_{n \rightarrow \infty} 1/n + 3 \cdot \lim_{n \rightarrow \infty} 1/n^2 = -2 \cdot 0 + 3 \cdot 0 = 0$
2.  $\lim_{n \rightarrow \infty} (2 + 1/n) / (5 - 1/n^2) = (2 + \lim_{n \rightarrow \infty} 1/n) / (5 + \lim_{n \rightarrow \infty} 1/n^2) = 2/5$ .
3.  $\lim_{n \rightarrow \infty} \sin(2 + 1/n) = \sin(2 + \lim_{n \rightarrow \infty} 1/n) = \sin(2)$

**Beispiel** Es gilt

$$a_n = \sqrt{n^2 + n} - n \rightarrow \frac{1}{2}$$

denn wir können  $a_n$  wie folgt umformen:

$$\begin{aligned} a_n &= \frac{\sqrt{n^2 + n} - n}{1} \cdot \frac{\sqrt{n^2 + n} + n}{\sqrt{n^2 + n} + n} = \frac{n}{\sqrt{n^2 + n} + n} \\ &= \frac{n}{\sqrt{n^2 + n} + n} \cdot \frac{1/n}{1/n} = \frac{1}{\sqrt{1 + 1/n} + 1} \rightarrow \frac{1}{2} \end{aligned}$$

**Theorem (Eigenschaften konvergenter Folgen)** Jede konvergente Folge

1. ist beschränkt,
2. ist Cauchy-Folge,
3. besitzt genau einen Grenzwert.

Beweis, Teil 1: Es existiert ein  $N \in \mathbb{N}$ , so dass

$$|a_n - a_\infty| \leq 1 \quad \text{und damit} \quad |a_n| \leq |a_\infty| + |a_n - a_\infty| \leq |a_\infty| + 1$$

für alle  $n \geq N$ . Die Behauptung folgt mit  $C = \max\{|a_1|, \dots, |a_N|, |a_\infty| + 1\}$ .

Beweis, Teil 2: Für jedes feste  $\varepsilon > 0$  wählen wir  $N$ , so dass

$$|a_n - a_\infty| < \varepsilon/2 \quad \text{für alle} \quad n \geq N$$



und erhalten mit der Dreiecksungleichung

$$|a_n - a_m| \leq |a_n - a_\infty| + |a_m - a_\infty| < 2\varepsilon/2 \quad \text{für alle } n, m \geq N.$$

*Beweis, Teil 3:* Angenommen, es gelte  $a_n \rightarrow a_\infty$  sowie  $a \rightarrow \tilde{a}_\infty$ . Dann gibt es für jedes  $\varepsilon > 0$  Indizes  $N, \tilde{N} \in \mathbb{N}$  mit

$$|a_n - a_\infty| < \varepsilon/2 \quad \text{für alle } n \geq N, \quad |a_n - \tilde{a}_\infty| < \varepsilon/2 \quad \text{für alle } n \geq \tilde{N},$$

und wir erhalten für  $n \geq \max\{N, \tilde{N}\}$  die Abschätzung

$$0 \leq |a_\infty - \tilde{a}_\infty| \leq |a_n - a_\infty| + |a_n - \tilde{a}_\infty| < \varepsilon.$$

Da  $\varepsilon$  beliebig klein gewählt werden kann, folgt  $a_\infty = \tilde{a}_\infty$ .  $\square$

**Theorem (Bedeutung der Cauchy-Eigenschaft)** Jede Cauchy-Folge ist konvergent.

*Beweis:* Das ist eine sehr tiefe Erkenntnis, die wir hier nicht beweisen können. Sie gilt letztlich, weil  $\mathbb{R}$  und  $\mathbb{C}$  *vollständig* sind, d.h. weil es keine *Lücken* in den reellen oder komplexen Zahlen gibt (siehe dazu Kapitel 2).  $\square$

**\*Bemerkung** Da  $\mathbb{Q}$  nicht vollständig ist, gibt es Folgen rationaler Zahlen, die der Cauchy-Eigenschaft genügen und gegen einen Grenzwert konvergieren, der zwar zu  $\mathbb{R}$ , aber nicht zu  $\mathbb{Q}$  gehört. Ein einfaches Beispiel sind die endlichen Dezimalapproximationen von  $\pi \notin \mathbb{Q}$ , d.h.

$$a_1 = 3.1, \quad a_2 = 3.14, \quad a_3 = 3.141, \quad a_4 = 3.1415, \quad a_5 = 3.14159, \quad \dots$$

Es handelt sich hierbei um eine Cauchy-Folge rationaler Zahlen, denn für  $n_2 \geq n_1$  gilt

$$|a_{n_2} - a_{n_1}| \leq \sum_{n=n_1+1}^{n_2} 9 \cdot 10^{-n} < 10^{-n_1},$$

da sich  $a_{n_1}$  und  $a_{n_2}$  nur in der  $n_1+1$ -ten bis zur  $n_2$ -ten Nachkommastelle unterscheiden. Oder anders gesagt: Es gilt  $|a_{n_2} - a_{n_1}| < \varepsilon$  für alle  $n_2 \geq n_1 \geq N$ , sofern bei gegebenem  $\varepsilon > 0$  der Index  $N$  so groß gewählt wird, dass  $10^{-N} \leq \varepsilon$  gilt.

Vorlesung 21a, 09. Dezember 2019

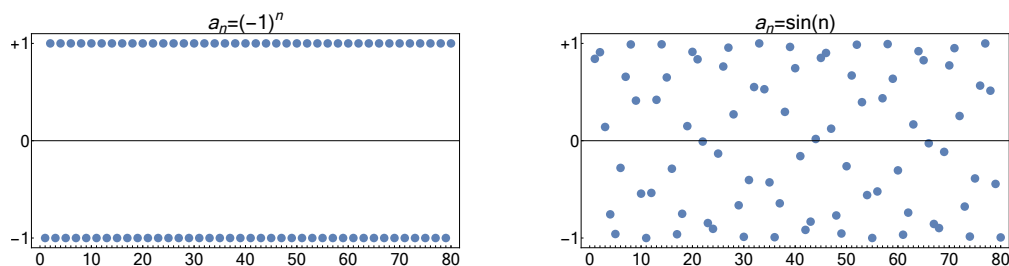
**Theorem (Folgenkompaktheit bzw. Satz von Bolzano-Weierstraß)** Jede beschränkte Folge besitzt (mindestens) eine konvergente Teilfolge, wobei der Grenzwert einer solchen Teilfolge Häufungspunkt der Folge heißt.

*Beweis:* Auch diese Aussage können wir hier nicht beweisen.  $\square$

### Beispiele

1. Die Folge  $a_n = (-1)^n$  besitzt die beiden Häufungspunkte  $-1$  und  $+1$ , denn die beiden Teilfolgen  $(a_{2j})_{j \in \mathbb{Z}}$  und  $(a_{2j+1})_{j \in \mathbb{Z}}$  sind konstant und daher auch konvergent.

2. Die Folge  $(a_n)_{n \in \mathbb{N}}$  mit  $a_n := \sin(n)$  besitzt sehr viele Häufungspunkte, nämlich jeden Wert in  $[0, 1]$ . (Das ist allerdings gar nicht so einfach zu beweisen).



Zwei Beispiele für beschränkte, aber nicht konvergente Folgen. Es gibt immer konvergente Teilfolgen und Häufungspunkte.

**Folgen in Vektorräumen** Alle bisherigen Definitionen und Aussagen lassen sich leicht auf vektorwertige Folgen  $a : \mathbb{N} \rightarrow V$  übertragen, sofern der Betrag  $|\cdot|$  durch eine Norm  $\|\cdot\|$  ersetzt wird. Hierbei ist  $V$  ein Vektorraum, zum Beispiel  $V = \mathbb{R}^n$  oder  $V = \mathbb{C}^{(n,m)}$ .

**Achtung** Bei unendlich-dimensionalen Räumen — zum Beispiel Funktionenräume — kann man aber nicht unbedingt davon ausgehen, dass sie *vollständig* sind, d.h. dass jede Cauchy-Folge auch konvergiert. Außerdem sind in einem unendlich-dimensionalen Vektorraum nicht alle Normen äquivalent, d.h. ob eine gegebene Folge konvergiert oder nicht, hängt von der gewählten Norm ab. In endlich-dimensionalen Räumen treten diese Subtilitäten hingegen nicht auf.

**Bemerkung** Für jede Folge mit Werten in  $\mathbb{C}^n$  oder  $\mathbb{C}^{(n,m)}$  sind die folgenden Aussagen äquivalent:

1. sie konvergiert,
2. sie konvergiert komponentenweise,
3. sie ist Cauchy-Folge.

### Beispiele

$$\begin{pmatrix} 2 + \sin(1/n) \\ 1 \\ n^2 + \cos(1/n) \end{pmatrix} \xrightarrow{n \rightarrow \infty} \begin{pmatrix} 2 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 1 + 1/n & 2 + i \\ -i/n^2 & \cos(1/n) \end{pmatrix} \xrightarrow{n \rightarrow \infty} \begin{pmatrix} 1 & 2 + i \\ 0 & 1 \end{pmatrix}.$$

## Konvergenzkriterien für reelle Zahlenfolgen

**Definition** Wir sagen, eine reelle Folge  $(a_n)_{n \in \mathbb{N}} \subset \mathbb{R}$  konvergiert uneigentlich gegen  $+\infty$  (bzw. gegen  $-\infty$ ), falls

$$\forall C > 0 \quad \exists N \in \mathbb{N} \quad \forall n \geq N \quad : \quad a_n > C \quad (\text{bzw. } a_n < -C).$$

### Beispiele

$$\lim_{n \rightarrow \infty} \sqrt{n} = +\infty, \quad \lim_{n \rightarrow \infty} \ln(1/n) = -\infty.$$

**Definition** Eine reelle Folge  $(a_n)_{n \in \mathbb{N}} \subset \mathbb{R}$  heißt

1. monoton wachsend, falls  $a_n \leq a_{n+1}$  für alle  $n \in \mathbb{N}$ ,
2. strikt (oder streng) monoton wachsend, falls  $a_n < a_{n+1}$  für alle  $n \in \mathbb{N}$ ,
3. nach oben beschränkt, falls es eine Konstante  $C \in \mathbb{R}$  gibt, so dass  $a_n \leq C$  für alle  $n \in \mathbb{N}$ .

Durch analoge Formeln werden die Begriffe monoton fallend, strikt monoton fallend, und nach unten beschränkt eingeführt. Insbesondere ist eine reelle Folge genau dann beschränkt, falls sie nach oben und nach unten beschränkt ist.

**Theorem (Monotonie + Beschränktheit = Konvergenz)** Für die reelle Folge  $(a_n)_{n \in \mathbb{N}}$  gelte

$$a_n \leq a_{n+1} \leq C \quad (\text{bzw. } C \leq a_{n+1} \leq a_n)$$

für eine geeignete Konstante  $C \in \mathbb{R}$  und alle  $n \in \mathbb{N}$  (siehe Bild). Dann gilt

$$a_n \rightarrow a_\infty$$

mit  $a_\infty := \sup\{a_n : n \in \mathbb{N}\}$  (bzw.  $a_\infty := \inf\{a_n : n \in \mathbb{N}\}$ ).

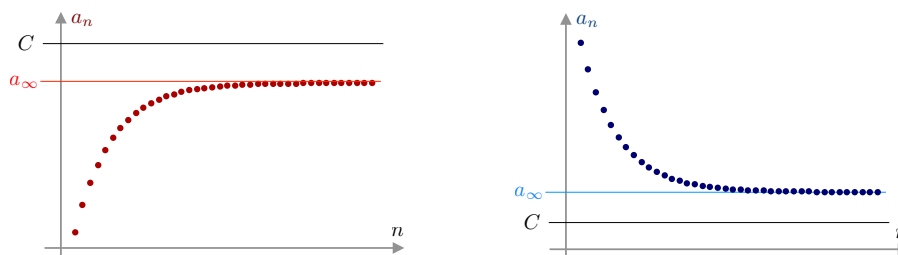
Beweis: Wir betrachten nur den ersten Fall, da wir im zweiten analog argumentieren können. Die Eigenschaften des Supremums implizieren, dass es für jedes  $\varepsilon > 0$  ein  $N \in \mathbb{N}$  gibt, so dass

$$a_\infty - \varepsilon < a_N,$$

denn andernfalls könnte  $a_\infty$  nicht die kleinste obere Schranke sein. Andererseits gilt nach Konstruktion von  $a_\infty$  und der Monotonie von  $a_n$  auch

$$a_N \leq a_n \leq a_\infty \quad \text{für alle } n \geq N.$$

Die Kombination beider Aussagen liefert  $|a_n - a_\infty| = a_\infty - a_n \leq a_\infty - a_N < \varepsilon$ .  $\square$



Zur monotonen Konvergenz reeller Folgen.

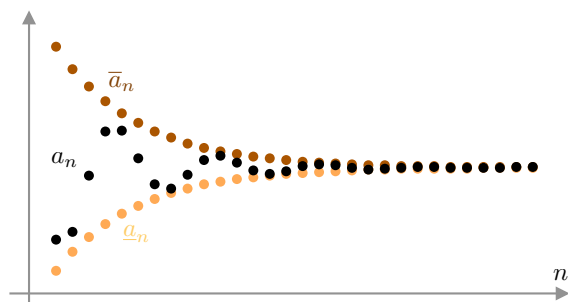
**Sandwich-Prinzip** Seien  $(a_n)_{n \in \mathbb{N}}$ ,  $(\underline{a}_n)_{n \in \mathbb{N}}$ ,  $(\bar{a}_n)_{n \in \mathbb{N}}$  drei reelle Folgen mit

$$\underline{a}_n \leq a_n \leq \bar{a}_n, \quad \underline{a}_n \leq \underline{a}_{n+1}, \quad \bar{a}_{n+1} \leq \bar{a}_n$$

für alle  $n \in \mathbb{N}$  und sei außerdem  $\lim_{n \rightarrow \infty} \bar{a}_n = \lim_{n \rightarrow \infty} \underline{a}_n =: \alpha$ . Dann gilt

$$|\underline{a}_n - \alpha| \leq \max\{a_\infty - \underline{a}_n, \bar{a}_n - a_\infty\} \xrightarrow{n \rightarrow \infty} 0$$

und damit insbesondere  $a_n \rightarrow \alpha$ .



Das Sandwich-Prinzip ist ein sehr mächtiges und robustes Werkzeug, um die Konvergenz reeller Folgen nachzuweisen.

**Beispiel** Die Konvergenzaussage

$$a_n = \frac{\sin^2(n^2)}{n(2 + \cos(n))} \xrightarrow{n \rightarrow \infty} 0,$$

folgt aus dem Sandwich-Prinzip mit  $\underline{a}_n = 0$  und  $\bar{a}_n = 1/n$ .

**Achtung** Durch Grenzübergang können aus strikten Ungleichungen Gleichungen werden. Zum Beispiel gilt  $x_n = 1/n > 0$  für alle  $n \in \mathbb{N}$ , aber  $x_\infty = 0$ .

**Theorem (Eulersche Zahl und reelle Exponentialfunktion)** Für jeden Parameter  $p \in \mathbb{R}$  gilt

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

Insbesondere erhalten wir für  $p = 1$  die mathematische Definition der Eulerschen Zahl  $e = 2.7182\dots$

Beweis, Vorbemerkung: Die Bernoullische Ungleichung

$$(1 + x)^n \geq 1 + nx \quad \text{für alle } x \geq -1 \text{ und alle } n \in \mathbb{N}$$

kann für festes  $x$  leicht mittels vollständiger Induktion über  $n$  abgeleitet werden. Wir werden außerdem nur den Spezialfall  $p = 1$  mit

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

betrachten und verweisen auf [AORS, Seite 204f] für eine analoge Diskussion im allgemeinen Fall.

Beweis, Monotonie der Folge: Mit elementaren Umformungen sowie der Bernoulli-Ungleichung für  $x = -1/(n+1)^2$  berechnen wir

$$\begin{aligned} \frac{a_{n+1}}{a_n} &= \left(1 + \frac{1}{n+1}\right) \cdot \left(\frac{1 + \frac{1}{n+1}}{1 + \frac{1}{n}}\right)^n = \frac{n+2}{n+1} \cdot \left(\frac{\frac{n+2}{n+1}}{\frac{n+1}{n}}\right)^n \\ &= \frac{n+2}{n+1} \cdot \left(\frac{n(n+2)}{(n+1)^2}\right)^n = \frac{n+2}{n+1} \cdot \left(1 - \frac{1}{(n+1)^2}\right)^n \\ &\geq \frac{n+2}{n+1} \cdot \left(1 - \frac{n}{(n+1)^2}\right) = \frac{n+2}{n+1} \cdot \frac{n^2 + n + 1}{(n+1)^2} \\ &= \frac{n^3 + 3n^2 + 3n + 2}{n^3 + 3n^2 + 3n + 1} \geq 1 \end{aligned}$$

und schließen, dass die Folge  $(a_n)_{n \in \mathbb{N}}$  in der Tat monoton wachsend ist.

Beweis, obere Beschränktheit und Konvergenz der Folge: Desweiteren gilt

$$a_n \leq a_{2n} = \left( \frac{1}{\left(1 - \frac{1}{2n+1}\right)^n} \right)^2 \leq \left( \frac{1}{1 - \frac{n}{2n+1}} \right)^2 = \left( \frac{2n+1}{n+1} \right)^2 \leq 4,$$

wobei wir diesmal die Bernoullische Ungleichung für  $x = 1/(2n+1)$  benutzt haben. Aus der Monotonie und der Beschränktheit der Folge  $(a_n)_{n \in \mathbb{N}}$  folgt nun die Konvergenz, wobei der Grenzwert  $e$  genannt wird.  $\square$

## 7.2 Reihen

**Partialsommenfolgen** Zu einer gegebenen Zahlenfolge  $(a_n)_{n \in \mathbb{N}}$  definiert man die Partialsommenfolge  $(s_m)_{m \in \mathbb{N}}$  durch

$$s_m := \sum_{n=1}^m a_n,$$

und kann nun fragen, ob es einen entsprechenden Grenzwert  $s_\infty$  gibt.

**Definition** Konvergiert die Partialsommenfolge, so schreiben wir

$$s_\infty = \lim_{m \rightarrow \infty} s_m = \sum_{n=1}^{\infty} a_n$$

und sagen, die Reihe zur Folge  $(a_n)_{n \in \mathbb{N}}$  konvergiert. Wir reden darüber hinaus von absoluter Konvergenz, wenn sogar die Reihe  $\sum_{n=1}^{\infty} |a_n|$  konvergiert, d.h. wenn der Limes der Folge

$$\lim_{m \rightarrow \infty} \sum_{n=1}^m |a_n|$$

existiert.

### Beispiele

1. Für jedes  $z \in \mathbb{C}$  mit  $|z| < 1$  konvergiert die Reihe der entsprechenden geometrischen Folge  $(a_n)_{n \in \mathbb{N}}$  mit  $a_n = z^n$ , wobei

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

In der Tat, mit  $a_n = z^n$  gilt

$$z \cdot s_m - s_m = \sum_{n=1}^m z^{n+1} - \sum_{n=1}^m z^n = \sum_{n=2}^{m+1} z^n - \sum_{n=1}^m z^n = z^{m+1} - z^1$$

und damit

$$s_m = \frac{z - z^{m+1}}{1 - z} \xrightarrow{m \rightarrow \infty} \frac{z}{1 - z}.$$

Die Konvergenz ist sogar absolut, denn unsere Resultate gelten ja auch für  $0 < |z| < 1$  und implizieren  $\sum_{n=1}^{\infty} |z| = |z|/(1 - |z|) < \infty$ .

2. Für jeden Parameter  $s > 1$  ist die Riemannsche Zeta-Funktion

$$\zeta(s) = \sum_{n=1}^{\infty} \frac{1}{n^s}$$

wohldefiniert. Insbesondere kann man zeigen, dass

$$\zeta(2) = 1 + \frac{1}{4} + \frac{1}{9} + \frac{1}{16} + \dots = \frac{\pi^2}{6}, \quad \zeta(4) = \frac{\pi^4}{90}, \quad \zeta(6) = \frac{\pi^6}{945}.$$

*Bemerkung:* Die Frage, welche Eigenschaften die Zeta-Funktion für komplexe Parameter  $s$  besitzt, führt zu einem der sogenannten Millenniums-Probleme der Mathematik sowie zur Frage nach der asymptotischen Verteilung der Primzahlen.

3. Wegen

$$s_m = \sum_{n=1}^m \frac{1}{n(n+1)} = \sum_{n=1}^m \left( \frac{1}{n} - \frac{1}{n+1} \right) = \sum_{n=1}^m \frac{1}{n} - \sum_{n=2}^{m+1} \frac{1}{n} = 1 - \frac{1}{m+1}$$

gilt

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

Die Konvergenz ist sogar absolut, da die Folgenglieder sämtlich positive reelle Zahlen sind.

Vorlesung 21b, 11. Dezember 2019

### **Bemerkung**

1. Das Cauchy-Kriterium für die Konvergenz einer Reihe lautet

$$\forall \varepsilon > 0 \quad \exists M \in \mathbb{N} \quad \forall m_2 > m_1 \geq M \quad : \quad |s_{m_2} - s_{m_1}| = \left| \sum_{n=m_1+1}^{m_2} a_n \right| < \varepsilon,$$

und ein analoges Kriterium mit  $|a_n|$  statt  $a_n$  charakterisiert die absolute Konvergenz.

2. Das Cauchy-Kriterium impliziert: Konvergiert die Reihe zu  $(a_n)_{n \in \mathbb{N}}$  (absolut oder nicht), so gilt  $a_n \rightarrow 0$ . Insbesondere können nur die Reihen zu Nullfolgen konvergieren.

3. Absolute Konvergenz untersucht die Konvergenz der *reellen* und *monoton wachsenden* Partialsummenfolge  $(t_m)_{m \in \mathbb{N}}$  mit

$$t_m := \sum_{n=1}^m |a_n| ,$$

wobei der Grenzwert entweder im eigentlichen ( $t_\infty < \infty$ ) oder im uneigentlichen Sinne existiert ( $t_\infty = \infty$ , siehe oben). Insbesondere konvergiert die Reihe der Folge  $(a_n)_{n \in \mathbb{N}}$  genau dann absolut, wenn  $(t_m)_{m \in \mathbb{N}}$  nach oben beschränkt ist, d.h. wenn  $\sum_{n=1}^{\infty} |a_n| = \lim_{m \rightarrow \infty} t_m < \infty$ .

4. Absolute Konvergenz impliziert Konvergenz, denn mit  $m_2 \geq m_1$  gilt nach Dreiecksungleichung

$$|s_{m_2} - s_{m_1}| = \left| \sum_{n=1}^{m_2} a_n - \sum_{n=1}^{m_1} a_n \right| \leq \sum_{n=m_1+1}^{m_2} |a_n| = |t_{m_2} - t_{m_1}| ,$$

d.h.  $(s_m)_{m \in \mathbb{N}}$  ist Cauchy-Folge, sofern  $(t_m)_{m \in \mathbb{N}}$  Cauchy-Folge ist.

5. Konvergenz impliziert im Allgemeinen nicht absolute Konvergenz. Das Standard-Gegenbeispiel (siehe auch Große Übung II bzw. den Abschnitt über Taylor-Reihen) ist

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \dots = \ln 2 , \quad \sum_{n=1}^{\infty} \frac{1}{n} = \infty .$$

Beachte aber, dass für reelle Folgen mit nur positiven oder nur negativen Folgengliedern Konvergenz und absolute Konvergenz gleichbedeutend sind.

6. Manchmal ist es sinnvoll, eine Reihe bei dem Index 0 oder  $n_0 \in \mathbb{N}$  zu starten:

$$\sum_{n=0}^{\infty} a_n = a_0 + \sum_{n=1}^{\infty} a_n = a_0 + \lim_{m \rightarrow \infty} \sum_{n=1}^m a_n , \quad \sum_{n=n_0}^{\infty} a_n = \sum_{j=1}^{\infty} a_{n_0-1+j} ,$$

und alle Resultate in diesem Abschnitt gelten sinngemäß auch für solche Reihen. Insbesondere gilt  $\sum_{n=0}^{\infty} z^n = 1/(1-z)$  für alle  $z \in \mathbb{C}$  mit  $|z| < 1$  im Sinne absolut konvergenter Reihen.

## Konvergenzkriterien

**Theorem** Die folgenden Bedingungen implizieren die absolute Konvergenz der Reihe zur Folge  $(a_n)_{n \in \mathbb{N}}$ :

1. (Majorantenkriterium) Es gilt  $|a_n| \leq b_n$  für alle  $n \in \mathbb{N}$  mit  $\sum_{n=1}^{\infty} b_n < \infty$ , wobei die Folge  $(b_n)_{n \in \mathbb{N}}$  die sogenannte Majorante ist.
2. (Quotientenkriterium) Es existiert  $0 \leq q < 1$ , so dass  $|a_{n+1}/a_n| \leq q < 1$  für alle  $n \in \mathbb{N}$ .
3. (Wurzelkriterium) Es existiert  $0 \leq q < 1$ , so dass  $\sqrt[n]{|a_n|} \leq q < 1$  für alle  $n \in \mathbb{N}$ .

Beweis:

1. Die Folge  $(t_m)_{m \in \mathbb{N}}$  mit

$$t_m = \sum_{n=1}^m |a_n|$$

ist monoton wachsend (wegen  $t_{m+1} = t_m + |a_{m+1}| \geq t_m$  für alle  $m \in \mathbb{N}$ ) und nach Voraussetzung auch nach oben beschränkt via

$$t_m \leq \sum_{n=1}^m b_n \leq \sum_{n=1}^{\infty} b_n < \infty.$$

Hieraus folgt die Existenz des Grenzwertes  $\lim_{m \rightarrow \infty} t_m = \sum_{n=1}^{\infty} |a_n|$  und damit die absolute Konvergenz der Reihe  $\sum_{n=1}^{\infty} a_n$ .

2. Es gilt

$$|a_2| \leq q \cdot |a_1|, \quad |a_3| \leq q \cdot |a_2| \leq q^2 \cdot |a_1|, \quad |a_4| \leq q \cdot |a_3| \leq q^3 \cdot |a_1|$$

und mit vollständiger Induktion über  $n$  folgt

$$|a_n| \leq q^n \cdot \frac{|a_1|}{q} =: b_n.$$

Die behauptete absolute Konvergenz folgt aus dem Majorantenkriterium wegen

$$\sum_{n=1}^{\infty} b_n = \frac{|a_1|}{q} \sum_{n=1}^{\infty} q^n = \frac{|a_1|}{q} \cdot \frac{q}{1-q} = \frac{|a_1|}{1-q} < \infty.$$

3. Nach Voraussetzung gilt diesmal

$$|a_n| \leq q^n =: b_n$$

und wir können wieder das Majorantenkriterium verwenden. □

### **Bemerkung**

1. Man kann das Quotienten- bzw. das Wurzelkriterium wie folgt abschwächen: Die Bedingungen

$$\lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| < 1 \quad \text{bzw.} \quad \lim_{n \rightarrow \infty} \sqrt[n]{|a_n|} < 1$$

implizieren (mit leicht abgewandeltem Beweis) auch schon die absolute Konvergenz der Reihe zu  $(a_n)_{n \in \mathbb{N}}$ .

2. Aus den Bedingungen

$$\lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| > 1 \quad \text{bzw.} \quad \lim_{n \rightarrow \infty} \sqrt[n]{|a_n|} > 1$$

folgt jedoch  $\sum_{n=1}^{\infty} |a_n| = \infty$ , d.h. die absolute Nicht-Konvergenz, den man kann für  $|a|_n$  eine divergente Minorante finden.



## 3. In den Grenzfällen

$$\lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right| = 1 \quad \text{bzw.} \quad \lim_{n \rightarrow \infty} \sqrt[n]{|a_n|} = 1$$

ist keine einfache Aussage möglich und das Verhalten von  $|a_n|$  muss genauer untersucht werden.

4. Analoge Aussagen gelten für Reihen  $\sum_{n=n_0}^{\infty} a_n$  mit  $n_0 \in \mathbb{Z}$  (zum Beispiel  $n_0 = 0$ ).

**Beispiele**

## 1. Die komplexe Reihe

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

konvergiert wegen

$$\left| \frac{a_{n+1}}{a_n} \right| = \frac{|a_{n+1}|}{|a_n|} = \frac{|z|}{n+1} \xrightarrow{n \rightarrow \infty} 0$$

für jedes  $z \in \mathbb{C}$  absolut nach dem (verallgemeinerten) Quotientenkriterium. Analoge Aussagen gelten für

$$\sum_{n=0}^{\infty} (-1)^n \frac{z^{2n+1}}{(2n+1)!} = z - \frac{z^3}{3!} + \frac{z^5}{5!} - \frac{z^7}{7!} + \dots = \sin(z)$$

und

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

## 2. Die Reihe

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

konvergiert wegen

$$\left| \frac{a_{n+1}}{a_n} \right| = |z|^2 \cdot \frac{2n+1}{2n+3} \xrightarrow{n \rightarrow \infty} |z|^2$$

für  $|z| < 1$  nach dem verallgemeinerten Quotientenkriterium absolut.

**Theorem (Leibnitz-Kriterium für alternierende Reihen)** Für jede monoton fallende Nullfolge  $(\tilde{a}_n)_{n \in \mathbb{N}}$  — d.h.  $0 \leq \tilde{a}_{n+1} \leq \tilde{a}_n$  für alle  $n \in \mathbb{N}$  und  $\tilde{a}_n \rightarrow 0$  — konvergiert die alternierende Reihe

$$\sum_{n=1}^{\infty} (-1)^n \tilde{a}_n,$$

aber die Konvergenz dieser Reihe ist im Allgemeinen nicht absolut (sondern nur, falls außerdem  $\sum_{n=1}^{\infty} \tilde{a}_n < \infty$  gilt).

*Beweis:* Wir definieren die Folgen  $(u_m)_{m \in \mathbb{N}}$  und  $(v_m)_{m \in \mathbb{N}}$  durch

$$u_m := \sum_{n=1}^{2m-1} (-1)^n \tilde{a}_n = -\tilde{a}_1 + \tilde{a}_2 - \tilde{a}_3 \dots - \tilde{a}_{2m-1},$$

$$v_m := \sum_{n=1}^{2m} (-1)^n \tilde{a}_n = -\tilde{a}_1 + \tilde{a}_2 - \tilde{a}_3 \dots - \tilde{a}_{2m-1} + \tilde{a}_{2m}$$

und bemerken, dass

$$0 < v_m - u_m = \tilde{a}_{2m} \xrightarrow{m \rightarrow \infty} 0.$$

Andererseits gilt

$$u_{m+1} = -\tilde{a}_1 + \tilde{a}_2 \dots - \tilde{a}_{2m-1} + \tilde{a}_{2m} - \tilde{a}_{2m+1} = u_m + \tilde{a}_{2m} - \tilde{a}_{2m+1} \geq u_m$$

sowie

$$v_{m+1} = -\tilde{a}_1 + \tilde{a}_2 \dots - \tilde{a}_{2m-1} + \tilde{a}_{2m} - \tilde{a}_{2m+1} + \tilde{a}_{2m+2} = v_m - \tilde{a}_{2m+1} + \tilde{a}_{2m+2} \leq v_m$$

d.h.  $(u_m)_{m \in \mathbb{N}}$  ist monoton wachsend und  $(v_m)_{m \in \mathbb{N}}$  ist monoton fallend. Wir schließen nun via

$$u_m \leq v_m \leq v_1, \quad u_1 \leq u_m \leq v_m$$

dass beide Folgen nach oben bzw. nach unten beschränkt sind und damit konvergieren, wobei die Grenzwerte wegen der ersten Identität sogar gleich sein müssen. Dies impliziert schließlich mit  $u_m = s_{2m-1}$  und  $v_m = s_{2m}$  die Konvergenz der Partialsummenfolge  $(s_m)_{m \in \mathbb{N}}$ .  $\square$

### Bemerkung

1. Analoge Resultate gelten für  $\sum_{n=n_0}^{\infty} (-1)^n \tilde{a}_n$  oder  $\sum_{n=n_0}^{\infty} (-1)^{n+1} \tilde{a}_n$ . Wichtig beim Leibnitz-Kriterium ist nur, dass die Vorzeichen alternieren (d.h.  $\operatorname{sgn}(a_{n+1}) = -\operatorname{sgn}(a_n)$  für alle  $n$ ) und dass  $(|a_n|)_{n \in \mathbb{N}}$  eine monoton fallende Nullfolge ist.
2. Die Konvergenz der alternierenden geometrischen Reihe

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} = \frac{1}{1} - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \dots$$

kann mit dem Leibnitz-Kriterium begründet werden, aber den genauen Zahlenwert der Reihe (nämlich  $\ln 2$ ) erhält man so nicht.

3. Weitere Beispiele sind

$$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n^2} = \frac{1}{1} - \frac{1}{4} + \frac{1}{9} - \frac{1}{16} + \dots$$

oder

$$\sum_{n=2}^{\infty} \frac{(-1)^n}{\ln n} = -\frac{1}{\ln 2} + \frac{1}{\ln 3} - \frac{1}{\ln 4} + \frac{1}{\ln 5} - \frac{1}{\ln 6} + \dots,$$

wobei die erste Reihe sogar absolut konvergiert, das zweite aber nicht.

**Theorem (Umordnungssatz)** Absolut konvergente Reihen können beliebig umgeordnet werden, d.h. es gilt

$$\sum_{n=1}^{\infty} |a_n| < \infty \quad \Longrightarrow \quad \sum_{n=1}^{\infty} a_n = \sum_{n=1}^{\infty} a_{\sigma_n}$$

für jede bijektive Abbildung  $\sigma : \mathbb{N} \rightarrow \mathbb{N}$ .

*Beweis:* Es gilt

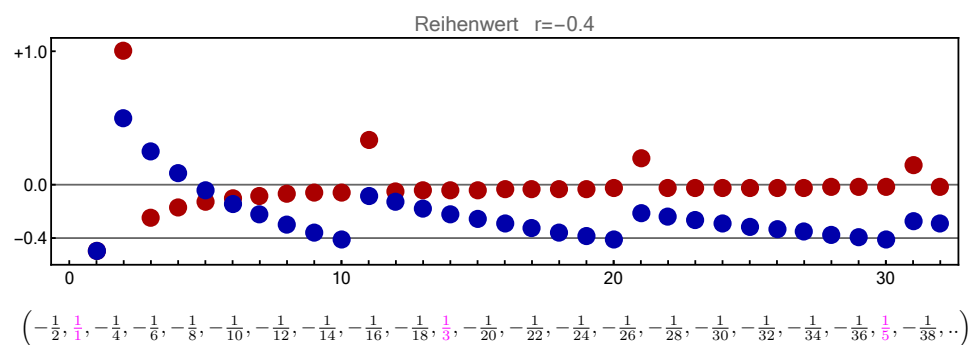
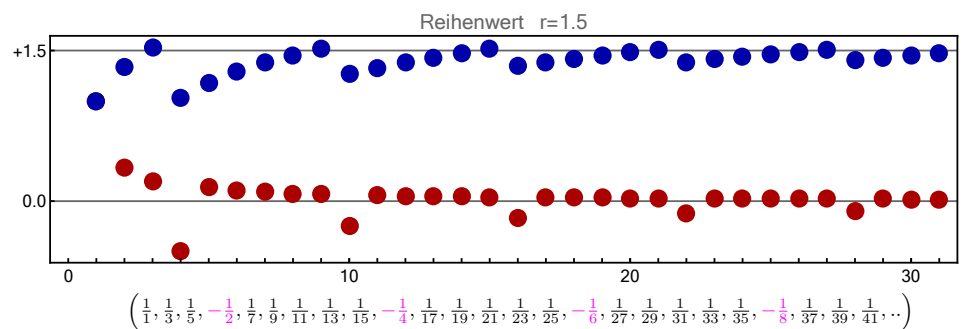
$$\sum_{n=1}^m |a_{\sigma_n}| \leq \sum_{n=1}^{\infty} |a_n| < \infty$$

und wegen der Monotonie der linken Seite bzgl.  $m$  folgt sowohl die absolute Konvergenz der Reihe zu  $(a_{\sigma_n})_{n \in \mathbb{N}}$  als auch  $\sum_{n=1}^{\infty} |a_{\sigma_n}| \leq \sum_{n=1}^{\infty} |a_n|$ . Die umgekehrte Ungleichung kann analog abgeleitet werden.  $\square$

**\*Paradox der Umordnung** Die absolute Konvergenz der Reihe ist wesentlich im Umordnungssatz. Für die alternierende harmonische Folge  $a_n = (-1)^{n+1}/n$  kann zum Beispiel gezeigt werden (\*\*-Übungsaufgabe), dass für jedes  $r \in \mathbb{R}$  eine bijektive Abbildung  $\sigma : \mathbb{N} \rightarrow \mathbb{N}$  existiert, so dass

$$\sum_{n=1}^{\infty} a_{\sigma_n} = r.$$

Man kann also so zu jedem vorgeschriebenem Reihenwert  $r$  eine entsprechende Umordnung finden (siehe Bild).



Zwei verschiedene Umordnungen der alternierenden harmonischen Folge (rote Punkt sowie die Zahlen) sowie die entsprechende Partialsummenfolge (blaue Punkte).

**Cauchy-Produkt von Reihen** Mit absolut konvergenten Reihen kann man so rechnen, wie man es von endlichen Summen gewöhnt ist. Zum Beispiel gilt nach dem Umordnungssatz die Formel

$$\begin{aligned} \left( \sum_{n=0}^{\infty} a_n \right) \cdot \left( \sum_{k=0}^{\infty} b_k \right) &= (a_0 + a_1 + a_2 + a_3 + \dots) \cdot (b_0 + b_1 + b_2 + b_3 + \dots) \\ &= a_0 \cdot b_0 + (a_0 \cdot b_1 + a_1 \cdot b_0) + (a_0 \cdot b_2 + a_1 \cdot b_1 + a_2 \cdot b_0) + \dots \\ &= \sum_{m=0}^{\infty} \sum_{l=0}^m a_l \cdot b_{m-l}. \end{aligned}$$

Als wichtigen Spezialfall dieser Produktformel erhalten wir

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

für je zwei komplexe Zahlen  $y, z$ .

**Vektor- und matrisenwertige Reihen** Neben Reihen zu Zahlenfolgen können auch unendliche Summen für Folgen  $a : \mathbb{N} \rightarrow \mathbb{C}^n$  oder  $a : \mathbb{N} \rightarrow \mathbb{C}^{(n,n)}$  betrachtet werden, wobei die Konvergenz wieder ganz natürlich als Grenzwert der Partialsummen  $s_m = \sum_{n=1}^m a_n$  definiert wird und absolute Konvergenz  $\sum_{n=1}^{\infty} \|a_n\| < \infty$  meint. Ein wichtiges Beispiel für eine absolut konvergente (und damit auch konvergente) Reihe im Raum der quadratischen Matrizen ist das Matrix-Exponential

$$\exp(\mathbf{A}) = \sum_{n=0}^{\infty} \frac{\mathbf{A}^n}{n!} = \mathbf{I} + \mathbf{A} + \frac{1}{2} \cdot \mathbf{A} \cdot \mathbf{A} + \frac{1}{6} \cdot \mathbf{A} \cdot \mathbf{A} \cdot \mathbf{A} + \frac{1}{24} \cdot \mathbf{A} \cdot \mathbf{A} \cdot \mathbf{A} \cdot \mathbf{A} + \dots,$$

das eine sehr wichtige Rolle in der Theorie linearer Differentialgleichungen spielt (siehe Große Übung II). Insbesondere kann für jedes  $\mathbf{A} \in \mathbb{C}^{(n,n)}$  die Allgemeine Lösung des Differentialgleichungssystems

$$\frac{d}{dt} \mathbf{x}(t) = \mathbf{A} \cdot \mathbf{x}(t)$$

als

$$\mathbf{x}(t) = \exp(t \cdot \mathbf{A}) \cdot \mathbf{x}(0)$$

geschrieben bzw. berechnet werden.

# Kapitel 8

## Stetigkeit und Differenzierbarkeit

Vorlesung 22, 13. Dezember 2019

**Vorbemerkung** Wir betrachten Funktionen  $f : D \rightarrow \mathbb{R}$  auf einem Definitionsbereich  $D \subset \mathbb{R}$  und führen die zentralen Begriffe Stetigkeit und Differenzierbarkeit ein. Wir werden auch immer diskutieren, welche Konzepte und Resultate auf Abbildungen  $F : D \rightarrow V$  verallgemeinert werden können, sofern  $V$  ein (endlich-dimensionaler) Vektorraum ist. Die Analysis von  $f : \Omega \rightarrow \mathbb{R}^m$  mit  $\Omega \subset \mathbb{R}^n$  ist komplizierter und wird erst im zweiten Kurs dieser Reihe behandelt.

### 8.1 Stetigkeit

**Definition** Sei  $f : D \rightarrow \mathbb{R}$  eine Funktion auf der Menge  $D \subset \mathbb{R}$  und sei  $x_* \in D$  gegeben. Wir sagen,  $f$  konvergiert für  $x \rightarrow x_*$  gegen  $y_* \in \mathbb{R}$ , falls für jede Folge  $(x_n)_{n \in \mathbb{N}}$  mit

$$x_n \in D \quad \text{und} \quad x_n \neq x_* \quad \text{für alle } n$$

die Implikation

$$x_n \rightarrow x_* \quad \implies \quad f(x_n) \rightarrow y_*$$

gilt. Wir schreiben dann

$$y_* = \lim_{x \rightarrow x_*} f(x) \quad \text{oder} \quad f(x) \xrightarrow{x \rightarrow x_*} y_*$$

oder auch  $f(x) \rightarrow y_*$  für  $x \rightarrow x_*$ .

#### **Bemerkung**

1. Es gelten analoge Rechenregeln zur Konvergenz von Folgen, d.h. zum Beispiel

$$\begin{aligned} \lim_{x \rightarrow x_*} (\lambda \cdot f + \mu \cdot g)(x) &= \lim_{x \rightarrow x_*} (\lambda \cdot f(x) + \mu \cdot g(x)) \\ &= \lambda \cdot \left( \lim_{x \rightarrow x_*} f(x) \right) + \mu \cdot \left( \lim_{x \rightarrow x_*} g(x) \right) \end{aligned}$$

oder

$$\lim_{x \rightarrow x_*} (f \cdot g)(x) = \lim_{x \rightarrow x_*} (f(x) \cdot g(x)) = \left( \lim_{x \rightarrow x_*} f(x) \right) \cdot \left( \lim_{x \rightarrow x_*} g(x) \right).$$

- Mit  $y_n = f(x_n)$  kann man die Konvergenz  $f(x_n) \rightarrow y_*$  als eine Konvergenz von Punkten auf dem Graphen von  $f$ , d.h. auf der Menge

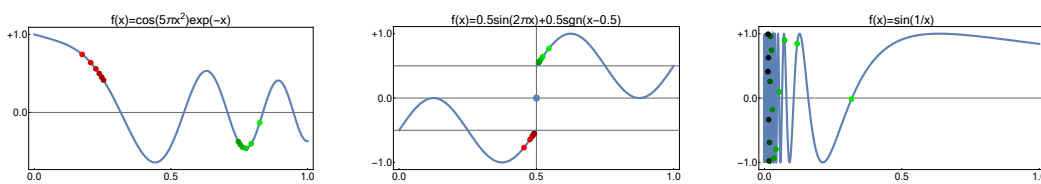
$$\text{graph}(f) = \{(x, y) : x \in D, y = f(x)\} \subset \mathbb{R}^2$$

verstehen: Mit  $x_n \rightarrow x_*$  gilt  $y_n \rightarrow y_*$  genau dann, wenn  $(x_n, y_n) \rightarrow (x_*, y_*)$ .

- Es gibt das allgemeinere Konzept der einseitigen Grenzwerte

$$\lim_{x \nearrow x_*} f(x) = \lim_{x \rightarrow x_*, x < x_*} f(x), \quad \lim_{x \searrow x_*} f(x) = \lim_{x \rightarrow x_*, x > x_*} f(x).$$

Dabei gilt  $\lim_{x \nearrow x_*} f(x) = \lim_{x \searrow x_*} f(x) = y_*$  (Gleichheit der einseitigen Grenzwerte) genau dann, wenn  $y_* = \lim_{x \rightarrow x_*} f(x)$  (Konvergenz).

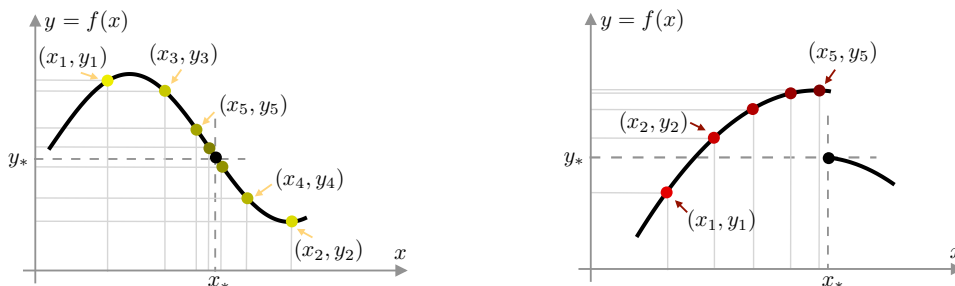


Drei Beispiele für Funktionen  $f : (0, 1) \rightarrow \mathbb{R}$  (blau) sowie Folgen  $(x_n, y_n) \subset \text{graph } f$  (rote bzw. grüne Punkte). Im ersten Beispiel folgt aus  $x_n \rightarrow x_*$  stets  $f(x_n) \rightarrow y_* = f(x_*)$ , da diese Funktion auf  $(0, 1)$  stetig ist (siehe unten). Im zweiten Beispiel gilt dies nur für  $x_* \neq 1/2$ , da die Funktion an der Stelle  $x_* = 1/2$  unstetig ist. Es existieren aber noch die einseitigen Grenzwerte  $\lim_{x \nearrow 1/2} f(x)$  und  $\lim_{x \searrow 1/2} f(x)$ , die allerdings verschieden sind. Das rechte Beispiel zeigt für  $x_n \rightarrow 0$  ein recht seltsames Verhalten, da es sehr viele verschiedene Häufungspunkte von  $f(x)$  für  $x \rightarrow 0$  gibt.

**Definition** Eine Funktion  $f : D \rightarrow \mathbb{R}$  heißt

- stetig an der Stelle  $x_* \in D$ , falls  $f(x_*) = \lim_{x \rightarrow x_*} f(x)$ ,
- unstetig an der Stelle  $x_* \in D$ , falls  $f$  nicht stetig in  $x_*$  ist,
- stetig auf der Menge  $\tilde{D} \subseteq D$ , falls  $f$  stetig in jeder Stelle  $x_* \in \tilde{D}$  ist,
- stetig, falls sie stetig auf dem Definitionsbereich  $D$  ist.

Statt „an der Stelle  $x_*$ “ sagt man oft auch „im Punkt  $x_*$ “ oder einfach „in  $x_*$ “.



Zur Definition von Stetigkeit durch Grenzwerte:  $f : D \rightarrow \mathbb{R}$  ist genau dann stetig im Punkt  $x_*$ , wenn mit  $x_n \rightarrow x_*$  schon  $y_n = f(x_n) \rightarrow y_* = f(x_*)$  und damit auch  $(x_n, y_n) \rightarrow (x_*, y_*)$  gilt.

**Interpretation** Salopp kann man sagen: Eine Funktion ist genau dann stetig, wenn man ihren Graphen zeichnen kann, ohne den Stift neu ansetzen zu müssen. Die mathematische Definition ist allerdings viel genauer, da sie Stetigkeit zunächst als punktweise Eigenschaft definiert.

**Beispiele**

1. Jedes reelle Polynom  $f : \mathbb{R} \rightarrow \mathbb{R}$  vom Grad  $m$  mit

$$f(x) = \alpha_m \cdot x^m + \dots + \alpha_1 x + \alpha_0 \quad \alpha_n \neq 0$$

ist auf ganz  $\mathbb{R}$  stetig, denn aus  $x_n \rightarrow x_*$  folgt  $x_n^i \rightarrow x_*^i$  für jedes  $i = 0, \dots, m$  und damit auch  $f(x_n) \rightarrow f(x_*)$ .

2. Man zeigt leicht, dass die Signum-Funktion

$$f(x) = \operatorname{sgn}(x) = \begin{cases} -1 & \text{für } x < 0 \\ 0 & \text{für } x = 0 \\ +1 & \text{für } x > 0 \end{cases}$$

unstetig an der Stelle  $x_* = 0$ , aber stetig in jeder anderen Stelle  $x_* \neq 0$  ist.

In der Tat, die Unstetigkeit bei  $x_* = 0$  folgt zum Beispiel mit  $x_n = 1/n \rightarrow 0$  und  $1 = f(x_n) \not\rightarrow f(x_*) = 0$  und die Stetigkeit in  $x_* < 0$  bzw.  $x_* > 0$  kann wie folgt begründet werden: Aus  $x_n \rightarrow x_*$  folgt  $x_n < 0$  bzw.  $x_n > 0$  für fast alle (d.h. bis auf vielleicht endlich viele)  $n \in \mathbb{N}$ , d.h. die Folge  $(f(x_n))_{n \in \mathbb{N}}$  ist fast konstant und nimmt ab einem gewissen Index nur den Wert  $-1$  bzw.  $+1$  an.

3. Die Funktion  $f : \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R}$  mit  $f(x) = 1/x$  ist stetig auf ihrem gesamten Definitionsbereich. Die Funktion besitzt zwar eine Singularität bei  $x_* = 0$ , aber dieser Punkt ist nicht Teil des Definitionsbereiches.
4. Die Exponentialfunktion  $\exp : \mathbb{R} \rightarrow \mathbb{R}$  ist stetig auf ganz  $\mathbb{R}$ . Dies folgt zum Beispiel aus der Tatsache, dass die Exponentialfunktion als *Potenzreihe* geschrieben werden kann (siehe weiter unten).
5. Die trigonometrischen Funktionen  $\sin$ ,  $\cos$ ,  $\tan$ ,  $\cot$  usw. sind stetig auf ihrem jeweiligen Definitionsbereich.
6. Da  $\exp : \mathbb{R} \rightarrow \mathbb{R}_+$  strikt monoton wachsend und stetig ist, ist auch die Umkehrfunktion  $\ln : \mathbb{R}_+ \rightarrow \mathbb{R}$  stetig (siehe das Theorem über die Stetigkeit der Umkehrabbildung).
7. Die Funktion  $f : \mathbb{R}^+ \rightarrow \mathbb{R}$  mit  $f(x) = x^p$  ist für jeden Parameter  $p > 0$  stetig, denn es gilt  $f(x) = \exp(\ln p \cdot x)$ .

**Komposition stetiger Abbildungen** Sind  $f_1 : D_1 \rightarrow D_2$  und  $f_2 : D_2 \rightarrow \mathbb{R}$  zwei stetige Funktionen, so ist auch die Komposition  $f_2 \circ f_1 : D_1 \rightarrow \mathbb{R}$  stetig. Zum Beispiel wird durch

$$f(x) = \sin(x^3 - 2x) \quad \text{bzw.} \quad f(x) = \ln(1 + \cos^2(x))$$

jeweils eine stetige Funktion  $f : \mathbb{R} \rightarrow \mathbb{R}$  definiert.

**\*Raum der stetigen Funktionen** Die Menge

$$C(D) := \{f : D \rightarrow \mathbb{R} : f \text{ ist stetig}\}$$

bildet in natürlicher Weise einen reellen Vektorraum, dessen Dimension allerdings nicht endlich ist. Außerdem handelt es sich sogar um ein Modul im Sinne der Algebra, da das (punktweise) Produkt zweier Elemente wieder eine stetige Funktion auf  $D$  ist. Ist  $D = [a, b]$  ein kompaktes Intervall, so wird durch

$$\|f\|_\infty := \max \{|f(x)| : x \in D\}$$

eine Norm auf  $C([a, b])$  definiert, wobei es aber kein entsprechendes Skalarprodukt gibt.

**Bemerkung** Nach Betrachtung der Beispiele könnte man meinen, dass jede Funktion in fast allen Punkten stetig ist und gegebenenfalls nur einige wenige Unstetigkeitsstellen besitzt. Dies ist aber so nicht richtig: Die Formel

$$f(x) = \begin{cases} 1 & \text{für } x \in \mathbb{Q} \\ 0 & \text{für } x \in \mathbb{R} \setminus \mathbb{Q} \end{cases}$$

definiert zum Beispiel eine Funktion, die in jedem Punkt  $x_* \in \mathbb{R}$  unstetig ist. Es geht aber noch verrückter: Mit

$$f(x) = \begin{cases} 0 & \text{für } x \in \mathbb{R} \setminus \mathbb{Q} \\ 1/q & \text{für } x = p/q \text{ mit teilerfremden Zahlen } p \in \mathbb{Z} \text{ und } q \in \mathbb{N} \end{cases}$$

erhält man eine Funktion, die in allen rationalen Punkten  $x_* \in \mathbb{Q}$  unstetig, in allen irrationalen Punkten  $x_* \in \mathbb{R} \setminus \mathbb{Q}$  aber stetig ist (\*\*-Übungsaufgabe)

**Theorem (Äquivalente Charakterisierung von Stetigkeit)** Eine Funktion  $f : D \rightarrow \mathbb{R}$  ist genau dann stetig in  $x_* \in D$ , falls

$$\forall \varepsilon > 0 \quad \exists \delta > 0 \quad \forall x \in D : |x - x_*| < \delta \implies |f(x) - f(x_*)| < \varepsilon.$$

Beachte, dass  $\delta$  von  $x_*$  und  $\varepsilon$  (aber nicht von  $x$ ) abhängen darf.

**Beweis:** Siehe zum Beispiel [AORS, Seite 222].

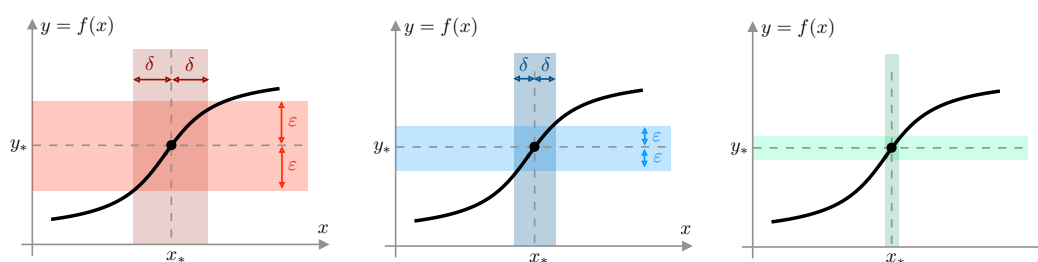


Illustration des  $\varepsilon$ - $\delta$ -Kriteriums der Stetigkeit in einem Stetigkeitspunkt: Für jedes  $\varepsilon > 0$  gibt es entsprechendes  $\delta > 0$ , und je kleiner  $\varepsilon$  ist, umso kleiner muss auch  $\delta$  gewählt werden.

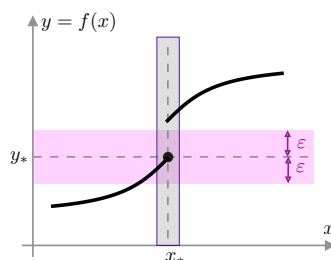


Illustration des  $\varepsilon$ - $\delta$ -Kriteriums der Stetigkeit in einem Unstetigkeitspunkt: Für gewisse Werte von  $\varepsilon > 0$  kann kein  $\delta > 0$  mit den gewünschten Eigenschaften gefunden werden.



**Verallgemeinerung** Alle bisher eingeführten Konzepte können auf Funktionen  $f : D \rightarrow V$  mit  $D \subset \mathbb{R}$  übertragen werden, sofern  $V$  ein beliebiger endlich-dimensionaler Vektorraum ist und der Betrag  $|\cdot|$  ggf. durch eine Norm  $\|\cdot\|$  ersetzt wird. Insbesondere gilt: Eine Abbildung  $f : D \rightarrow \mathbb{C}$  ist genau dann stetig, wenn  $\operatorname{Re}(f) : D \rightarrow \mathbb{R}$  und  $\operatorname{Im}(f) : D \rightarrow \mathbb{R}$  stetig sind, und eine Abbildung  $f : D \rightarrow \mathbb{C}^n$  oder  $f : D \rightarrow \mathbb{C}^{(n,m)}$  ist genau dann stetig, wenn jede Komponentenabbildung stetig ist.

**Beispiel** Die Formeln  $z(x) = \exp(-x^2) + i \cdot (x^3 - 2)$  sowie

$$\mathbf{v}(x) = \begin{pmatrix} 3x \\ -x^2 \\ 4x^3 \end{pmatrix}, \quad \mathbf{A}(x) = \begin{pmatrix} \cos(x) & \sin(x) \\ -\sin(x) & \cos(x) \end{pmatrix} + \begin{pmatrix} 0 & i \cdot x \\ -i \cdot x & 0 \end{pmatrix}$$

definieren stetige Abbildungen  $z : \mathbb{R} \rightarrow \mathbb{C}$ ,  $\mathbf{v} : \mathbb{R} \rightarrow \mathbb{R}^3$ ,  $\mathbf{A} : \mathbb{R} \rightarrow \mathbb{C}^{(2,2)}$ , eben weil jede Komponente stetig von  $x$  abhängt.

## Stetige Funktionen auf kompakten Intervallen

**Bemerkung** Ein Intervall  $[a, b]$  mit  $-\infty < a < b < \infty$  wird kompakt genannt. Es handelt sich insbesondere um eine beschränkte Menge, die darüber hinaus auch abgeschlossen unter Konvergenz ist. Letzteres meint, dass für jede Folge  $(x_n)_{n \in \mathbb{N}}$ , deren Glieder sämtlich in  $[a, b]$  liegen und die konvergiert, der Grenzwert  $x_\infty$  auch im Intervall  $[a, b]$  liegt. Beachte, dass ein offenes Intervall nicht abgeschlossen unter Konvergenz ist; die Glieder der Folge  $(1/n)_{n \in \mathbb{N}}$  liegen zum Beispiel alle im Intervall  $(0, 2)$ , aber der Grenzwert 0 liegt außerhalb.

**Theorem (Zwischenwertsatz)** Sei  $f : [a, b] \rightarrow \mathbb{R}$  eine stetige Funktion auf dem kompakten Intervall  $[a, b]$ . Gilt  $f(a) < f(b)$ , so existiert für jedes  $y_*$  mit  $f(a) < y_* < f(b)$  mindestens ein  $x_* \in [a, b]$ , so dass  $y_* = f(x_*)$ . Eine analoge Aussage gilt für  $f(a) > f(b)$ .

Beweis, Teil 1: Wir definieren rekursiv zwei Folgen  $(a_n)_{n \in \mathbb{N}_0}$  und  $(b_n)_{n \in \mathbb{N}_0}$  durch die Anfangsbedingung

$$a_0 := a, \quad b_0 := b$$

sowie den folgenden Iterationsschritt  $n \rightsquigarrow n + 1$ : Wir betrachten

$$x_n := \frac{a_n + b_n}{2}$$

und nutzen eine der folgenden Alternativen:

1. Gilt  $f(x_n) = y_*$ , so sind wir fertig und können mit  $x_* := x_n$  sofort abbrechen.
2. Gilt  $f(x_n) < y_*$ , so setzen wir  $a_{n+1} := x_n$  und  $b_{n+1} := b_n$ .
3. Gilt  $f(x_n) > y_*$ , so setzen wir  $a_{n+1} := a_n$  und  $b_{n+1} := x_n$ .

Beweis, Teil 2: Es kann nun sein, dass unser Algorithmus nach endlich vielen Schritten ein geeignetes  $x$  liefert. Andernfalls zeigt man leicht durch Induktion über  $n$ , dass

$$f(a_n) < y_* < f(b_n), \quad a_n \leq a_{n+1}, \quad b_{n+1} \leq b_n, \quad a_n \leq x_n \leq b_n$$

sowie

$$b_{n+1} - a_{n+1} = \frac{b_n - a_n}{2} = \frac{b_0 - a_0}{2^n}$$

für alle  $n \in \mathbb{N}_0$  gilt. Mit anderen Worten,  $a_n$  bzw.  $b_n$  wachsen bzw. fallen mit  $n$  und die Differenz halbiert sich in jedem Schritt. Insbesondere existiert nach dem Sandwich-Prinzip ein  $x_\infty \in [a, b]$  mit

$$x_\infty = \lim_{n \rightarrow \infty} x_n = \lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} b_n,$$

und nach Konstruktion sowie der Stetigkeit von  $f$  folgt

$$\lim_{n \rightarrow \infty} f(a_n) = f(x_\infty) \leq y_* \leq f(x_\infty) = \lim_{n \rightarrow \infty} f(b_n)$$

und damit  $f(x_\infty) = y_*$ . Das ist gerade die erste Behauptung mit  $x_* := x_\infty$ . Der Beweis der zweiten Behauptung kann analog geführt werden (Übungsaufgabe), sofern die Ordnungsrelationen angepasst werden.  $\square$

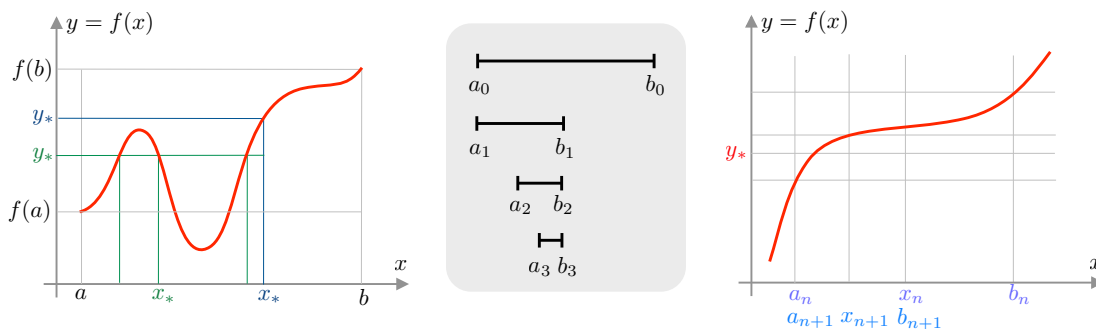


Illustration der Aussage (links) und des Beweises (rechts) des Zwischenwertsatzes sowie der entsprechenden Intervallschachtelung (Mitte), die durch sukzessives Halbieren entsteht.

**Bemerkung** Die im Beweis des Zwischenwertsatzes verwendete Methode ist ein Beispiel für eine Intervallschachtelung. Dies ist ein sehr robustes Konzept, das in vielen Bereichen angewendet werden kann.

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

$$f(a) < 0 < f(b) \quad \text{oder} \quad f(a) > 0 > f(b).$$

Dann besitzt  $f$  mindestens eine reelle Nullstelle im Intervall  $[a, b]$ .

**Folgerung** Jedes reelle Polynom ungerader Ordnung besitzt mindestens eine reelle Nullstelle.

**Beweis:** Je nach Vorzeichen des Koeffizienten vor der höchsten  $x$  Potenz gilt entweder

$$\lim_{x \searrow -\infty} f(x) = -\infty \quad \text{und} \quad \lim_{x \nearrow +\infty} f(x) = +\infty$$

oder

$$\lim_{x \searrow -\infty} f(x) = +\infty \quad \text{und} \quad \lim_{x \nearrow +\infty} f(x) = -\infty.$$

Insbesondere können immer  $a, b \in \mathbb{R}$  gefunden werden, so dass die vorherige Folgerung angewendet werden kann.  $\square$

**Theorem (Minima und Maxima)** Sei  $f : [a, b] \rightarrow \mathbb{R}$  stetig auf dem kompakten Intervall  $[a, b]$ . Dann existieren  $\bar{x}, \underline{x} \in [a, b]$ , so dass

$$f(\bar{x}) = \max \{f(x) : x \in [a, b]\}, \quad f(\underline{x}) = \min \{f(x) : x \in [a, b]\}.$$

Man schreibt

$$f(\bar{x}) = \max_{a \leq x \leq b} f(x) = \max f, \quad \bar{x} = \operatorname{argmax}_{a \leq x \leq b} f(x) = \operatorname{argmax} f$$

und nennt  $f(\bar{x})$  bzw.  $\bar{x}$  das Maximum bzw. einen Maximierer (oder eine Maximalstelle) von  $f$ . Analoge Notationen und Bezeichnungen verwendet man für das Minimum  $f(\underline{x})$  und den Minimierer  $\underline{x}$ .

Beweis: Im ersten Teil betrachten wir eine *maximierende Folge*  $(x_n)_{n \in \mathbb{N}} \subset [a, b]$  mit

$$f(x_n) \xrightarrow{n \rightarrow \infty} \sup f := \sup \{f(x) : x \in [a, b]\}.$$

Ein solche Folge existiert immer, denn wir können nach Definition des Supremums zum Beispiel für jedes  $n \in \mathbb{N}$  eine Zahl  $x_n \in [a, b]$  wählen, so dass

$$\sup f - \frac{1}{n} < f(x_n) \leq \sup f.$$

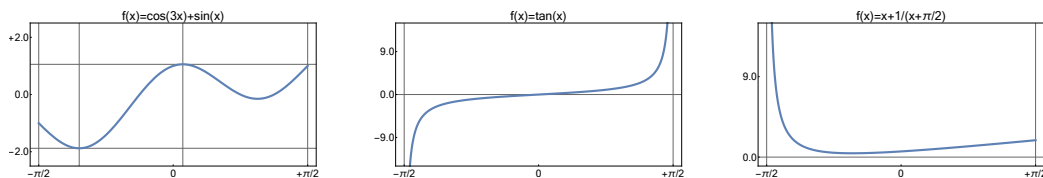
Da das Intervall  $[a, b]$  sowohl beschränkt als auch abgeschlossen ist, gibt es (siehe den Satz von Bolzano-Weierstraß) eine konvergente Teilfolge  $(x_{n_j})_{j \in \mathbb{N}}$  mit

$$x_{n_j} \xrightarrow{j \rightarrow \infty} x_\infty \in [a, b].$$

Die Stetigkeit von  $f$  impliziert

$$f(x_\infty) = \sup f$$

und damit auch die Behauptung via  $\bar{x} = x_\infty$  und  $\sup f = \max f$ . Der Beweis des zweiten Teiles folgt analog mit Hilfe einer minimierenden Folge.  $\square$



*Links*: Beispiel für eine stetige Funktion auf dem Intervall  $[-\pi/2, \pi/2]$ , die dort auch ihr Minimum und ihr Maximum annimmt. *Mitte* und *Rechts*: Stetige Funktionen auf den nicht kompakten Intervallen  $(-\pi/2, \pi/2)$  oder  $(-\pi/2, \pi/2]$  können, aber müssen kein Minimum oder Maximum besitzen.

**Bemerkung** Das soeben bewiesene Theorem ist ein Spezialfall des allgemeinen topologischen Prinzips

*Jede stetige Funktion  $f : K \rightarrow \mathbb{R}$  auf einer kompakten Menge  $K$  nimmt ihr Minimum und ihr Maximum an.*

Diese sehr bedeutende Erkenntnis gilt nicht nur auf Teilmengen  $K = D \subset \mathbb{R}$ , sondern auf einer sehr viel größeren Klasse von Mengen  $K$ , zum Beispiel auf einer Vereinigung kompakter Intervalle sowie auf jeder beschränkten und abgeschlossenen Teilmenge eines endlich-dimensionalen Vektorraumes. Die von uns verwendete Beweismethode kann ebenfalls sehr viel allgemeiner angewendet werden und wird auch die Direkte Methode der Optimierung genannt.

**Theorem (Existenz und Stetigkeit einer Umkehrfunktion)** Die Funktion  $f : [a, b] \rightarrow \mathbb{R}$  sei stetig und strikt monoton (wachsend oder fallend). Dann existiert die Umkehrabbildung  $f^{-1} : [f(a), f(b)] \rightarrow \mathbb{R}$  und ist auch strikt monoton und stetig.

*Beweis, Teil 1:* Wir betrachten den Fall wachsender Funktionen (der andere Fall kann analog behandelt werden) und bemerken zunächst, dass  $f$  das Intervall  $I := [a, b]$  bijektiv und monoton auf das Intervall  $J := [f(a), f(b)]$  abbildet (siehe Bild und den Zwischenwertsatz). Insbesondere ist  $f^{-1} : J \rightarrow I$  wohldefiniert und auch strikt monoton, denn mit  $y_i = f(x_i)$  bzw.  $x_i = f^{-1}(y_i)$  gilt  $x_1 < x_2$  dann und nur dann, wenn  $y_1 < y_2$ .

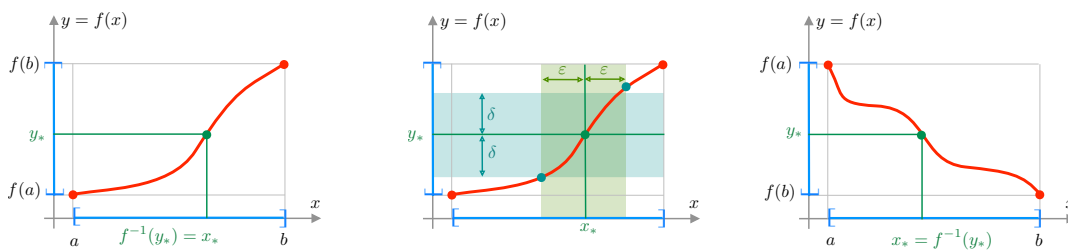
*Beweis, Teil 2:* Um die Stetigkeit von  $f^{-1}$  zu zeigen, fixieren wir  $y_* \in J$  sowie ein beliebiges  $\varepsilon > 0$ . Wir setzen nun (siehe Bild)

$$\delta := \min \{ y_* - f(x_* - \varepsilon), f(x_* + \varepsilon) - y_* \}$$

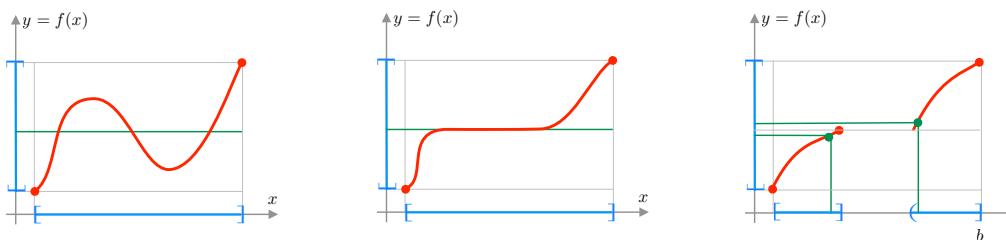
mit  $x_* = f^{-1}(y_*)$ , so dass nach Konstruktion und wegen der strikten Monotonie von  $f$  die Implikation

$$y \in [y_* - \delta, y_* + \delta] \implies x = f^{-1}(y) \in [x_* - \varepsilon, x_* + \varepsilon]$$

gilt. Da  $\varepsilon > 0$  beliebig war, folgt die Stetigkeit von  $f^{-1}$ . □



*Links und Rechts:* Jede strikt monotone und stetige Funktionen auf einem Intervall besitzt eine strikt monotone und stetige Umkehrfunktion. *Mitte:* Zum Beweis der Stetigkeit der Umkehrfunktion  $f^{-1}$ . Beachte, dass das Bild den Graph von  $f$  zeigt. Der Graph von  $f^{-1}$  kann aber hieraus leicht rekonstruiert werden (siehe unten).



*Links und Mitte:* Stetige, aber nicht monotone oder nicht strikt-monotone Funktionen auf Intervallen sind nicht invertierbar. *Rechts:* Stetige und strikt-monotone Funktionen auf Nicht-Intervallen besitzen zwar immer eine Umkehrfunktion, aber diese muss nicht stetig sein.

**Bemerkung**

1. Dieses Resultat gilt analog auf jedem Intervall und auf ganz  $\mathbb{R}$ , aber nicht auf jeder Teilmenge von  $\mathbb{R}$ .
2. Eine nicht-strikt monotone Funktion auf einem Intervall besitzt in der Regel keine Umkehrfunktion und ist  $D$  kein Intervall, so gibt es stetige und invertierbare Funktionen  $f : D \rightarrow \mathbb{R}$ , deren Umkehrabbildung zwar wohldefiniert, aber nicht stetig ist (siehe Bild).

3. Existiert die Umkehrfunktion  $f^{-1}$  von  $f$ , so können die Graphen von  $f$  und  $f^{-1}$  durch Spiegelung an der Geraden  $y = x$  ineinander überführt werden.

**Theorem (gleichmäßige Stetigkeit)** Jede stetige Funktion  $f : [a, b] \rightarrow \mathbb{R}$  ist gleichmäßig stetig, d.h.

$$\forall \varepsilon > 0 \exists \delta > 0 \forall x, x_* \in [a, b] : |x - x_*| < \delta \implies |f(x) - f(x_*)| < \varepsilon.$$

Oder anders gesagt: Beim  $\varepsilon$ - $\delta$ -Kriterium kann man bei kompakten Intervallen  $\delta$  zwar nicht unabhängig von  $\varepsilon$ , aber doch unabhängig von  $x_*$  wählen.

*Beweis:* Dieses technische — aber doch wichtige und nützliche Resultat — wird zum Beispiel in [AORS, Seite 229] hergeleitet.

## Weitere Aspekte

**Klassifikation von Unstetigkeitsstellen** In der Praxis treten oftmals isolierte Unstetigkeitsstellen auf, d.h. Punkte  $x_* \in D$ , so dass

1.  $f$  ist unstetig in  $x_*$ .
2. Es existiert  $\eta > 0$ , so dass  $f$  stetig auf den Intervallen  $(x_* - \eta, x_*)$  und  $(x_*, x_* + \eta)$  ist.

In diesem Fall kann man  $x_*$  wie folgt klassifizieren:

1.  $x_*$  heißt hebbare Unstetigkeit, wenn  $y_* = \lim_{x \rightarrow x_*} f(x)$  existiert.
2.  $x_*$  wird Sprungstelle genannt, wenn die einseitigen Grenzwerte  $\lim_{x \nearrow x_*} f(x)$  und  $\lim_{x \searrow x_*} f(x)$  (im eigentlichen Sinne) existieren und verschieden sind.
3.  $x_*$  heißt Polstelle, wenn die einseitigen Grenzwerte  $\lim_{x \nearrow x_*} f(x)$  und  $\lim_{x \searrow x_*} f(x)$  im uneigentlichen Sinne existieren (also den Wert  $+\infty$  oder  $-\infty$  annehmen).

Im ersten Fall gilt dabei  $y_* \neq f(x_*)$ , da andernfalls  $f$  in  $x_*$  nicht unstetig, sondern stetig wäre. Der wesentliche Punkt ist aber, dass die im Punkt  $x_*$  abgeänderte Funktion  $\tilde{f} : D \rightarrow \mathbb{R}$  mit

$$\tilde{f}(x) = \begin{cases} y_* & \text{für } x = x_* \\ f(x) & \text{für } x \neq x_* \end{cases}$$

in  $x_*$  stetig ist. Bei Sprungstellen kann man nicht durch eine punktweise Änderung aus  $f$  eine stetige Funktion machen, aber es existieren sowohl der linksseitige als auch der rechtsseitige Grenzwert, wobei

$$\mathbb{R} \ni \llbracket f \rrbracket_{x_*} := \lim_{\varepsilon \searrow 0} f(x_* + \varepsilon) - \lim_{\varepsilon \searrow 0} f(x_* - \varepsilon) = \lim_{x \searrow x_*} f(x) - \lim_{x \nearrow x_*} f(x)$$

der sogenannte Sprung von  $f$  an  $x_*$  ist. Bei einem Pol explodiert die Funktion auf beiden Seite zu  $\pm\infty$ , wobei es möglich ist, dass auf beiden Seiten verschiedene oder gleiche uneigentliche Grenzwerte angenommen werden (etwa  $f(x) = 1/x$  für  $x \neq x_* = 0$  und  $f(x_*) = 0$  bzw.  $f(x) = 1/x^2$  für  $x \neq x_* = 0$  und  $f(x_*) = 0$ ). Darüber hinaus gibt es aber auch isolierte Nullstellen, die weder hebbare Unstetigkeit noch Sprungstelle oder Pol sind.

**Beispiel** Wir betrachten die Funktionen  $f_i : \mathbb{R} \rightarrow \mathbb{R}$  mit

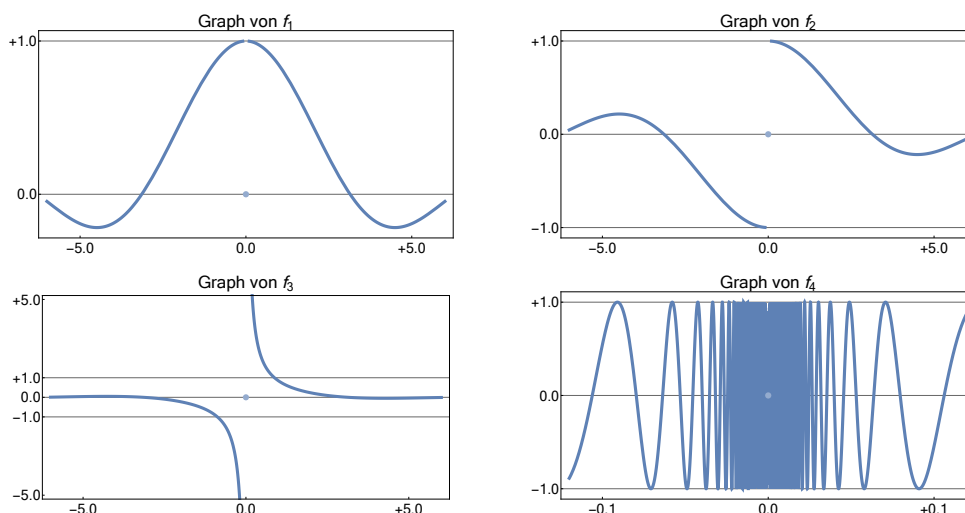
$$f_1(x) = \begin{cases} 0 & \text{für } x = 0 \\ \frac{\sin(x)}{x} & \text{für } x \neq 0 \end{cases}, \quad f_2(x) = \begin{cases} 0 & \text{für } x = 0 \\ \frac{\sin(x)}{\sqrt{x^2}} & \text{für } x \neq 0 \end{cases}$$

sowie

$$f_3(x) = \begin{cases} 0 & \text{für } x = 0 \\ \frac{\sin(x)}{x^2} & \text{für } x \neq 0 \end{cases}, \quad f_4(x) = \begin{cases} 0 & \text{für } x = 0 \\ \sin\left(\frac{1}{x}\right) & \text{für } x \neq 0 \end{cases}.$$

In jedem Fall ist  $x_* = 0$  eine isolierter Unstetigkeitsstelle, aber es handelt sich immer um einen anderen Typ (siehe Bild). Beachte dabei, dass

$$\lim_{x \searrow 0} \frac{\sin x}{x} = 1 = \lim_{x \nearrow 0} \frac{\sin x}{x}.$$



Die vier im Text angegebenen Beispiele für eine isolierte Unstetigkeitsstelle bei  $x_* = 0$ : Es handelt sich um eine hebbare Unstetigkeit, eine Sprungstelle, einen Pol und einen nicht-klassifizierten Typ.

**Lücken im Definitionsbereich** Neben isolierten Unstetigkeitsstellen kann es auch isolierte Lücken im Definitionsbereich  $D$  geben, d.h. Punkte  $x_*$  so dass

1.  $x_* \notin D$ ,
2. es existiert  $\eta > 0$ , so dass  $f$  auf den Intervalle  $(x_* - \eta, x_*)$  und  $(x_*, x_* + \eta)$  definiert und stetig ist.

Man kann nun die Lücke  $x_*$  ganz analog zu oben klassifizieren, wobei man im ersten Fall auch von einer hebbaren Singularität redet.

**Beispiele** Die obigen Formeln definieren für  $x \neq 0$  stetige Funktionen  $f_i : \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R}$ , wobei dann  $x_*$  keine isolierte Unstetigkeitsstelle, sondern eine isolierte Lücke im Definitionsbereich ist.

**Vertauschbarkeit von Grenzübergängen** In vielen praktischen Fällen hängt ein Ausdruck nicht nur von einer reellen Variablen  $x$ , sondern auch von einer zweiten Variablen (oder einem Parameter)  $\xi$  ab. Man kann dann die Grenzprozesse  $x \rightarrow x_*$  und  $\xi \rightarrow \xi_*$  auf zwei verschiedene Arten hintereinander ausführen.

**Beispiele**

1. Wir betrachten  $f : \mathbb{R}_+ \times \mathbb{R}_+ \rightarrow \mathbb{R}$  mit

$$f(x, \xi) = \exp(-x/\xi)$$

und berechnen

$$\lim_{\xi \searrow 0} \lim_{x \searrow 0} f(x, \xi) = \lim_{\xi \searrow 0} 1 = 1$$

sowie

$$\lim_{x \searrow 0} \lim_{\xi \searrow 0} f(x, \xi) = \lim_{x \searrow 0} 0 = 0.$$

Insbesondere hängt hier das Ergebnis von der Reihenfolge der Grenzübergänge ab.

2. Wir betrachten  $f : \mathbb{R}_+ \times \mathbb{R}_+ \rightarrow \mathbb{R}$  mit

$$f(x, \xi) = \xi \exp(-x)$$

und berechnen

$$\lim_{\xi \searrow 0} \lim_{x \searrow 0} f(x, \xi) = \lim_{\xi \searrow 0} \xi = 0$$

sowie

$$\lim_{x \searrow 0} \lim_{\xi \searrow 0} f(x, \xi) = \lim_{x \searrow 0} 0 = 0.$$

Dieses Beispiel ist regulärer, denn hier können beide Grenzübergänge vertauscht werden.

**Achtung** Im Allgemeinen ist die Aussage

$$\lim_{\xi \rightarrow \xi_*} \lim_{x \rightarrow x_*} f(x, \xi) = \lim_{x \rightarrow x_*} \lim_{\xi \rightarrow \xi_*} f(x, \xi),$$

*falsch*, d.h. verschiedene Grenzprozesse können nicht unbedingt miteinander vertauscht werden. Wir können diese Problematik hier nicht erschöpfend behandeln, wollen aber ein wichtiges Resultat vorwegnehmen bzw. informell formulieren: Man kann, vereinfacht gesprochen, nur dann die Gültigkeit der obigen Formel erwarten, wenn die Konvergenz für  $x \rightarrow x_*$  gleichmäßig bzgl.  $\xi$  oder die Konvergenz für  $\xi \rightarrow \xi_*$  gleichmäßig bzgl.  $x$  erfolgt.

**8.2 Differenzierbarkeit**

**Definition**  $f : D \rightarrow \mathbb{R}$  heißt differenzierbar an der Stelle  $x_* \in D$ , falls der Grenzwert

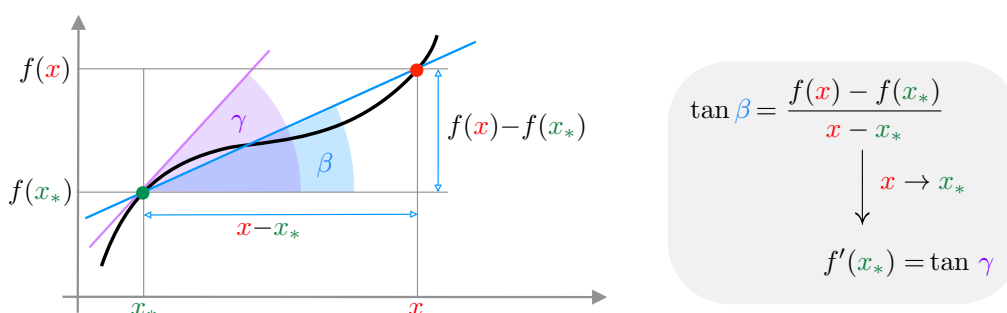
$$\lim_{x \rightarrow x_*} \frac{f(x) - f(x_*)}{x - x_*} =: f'(x_*) =: \frac{d}{dx} f(x_*)$$

existiert, wobei er dann die Ableitung von  $f$  in  $x_*$  genannt wird. Existiert  $f'(x)$  für jedes  $x_* \in D$  (bzw. für jedes  $x_* \in \tilde{D} \subset D$ ), so nennt man  $f$  differenzierbar (bzw. differenzierbar auf der Menge  $\tilde{D}$ ).

**Interpretation** Der Ausdruck

$$\frac{f(x) - f(x_*)}{x - x_*}$$

ist der Anstieg derjenigen Sekante, die die Punkte  $(x, f(x))$  und  $(x_*, f(x_*))$  auf dem Graphen von  $f$  miteinander verbindet. Im Limes  $x \rightarrow x_*$  wird daraus der Anstieg der Tangente an den Graphen von  $f$  im Punkt  $(x_*, f(x_*))$  (siehe Bild).



Die Ableitung  $f'(x_*)$  ist der Anstieg der Tangente (lila) an den Graphen bzw. der Grenzwert des Anstiegs der Sekanten (blau).

**Bemerkung** Bei der Berechnung von Differenzenquotienten sowie den entsprechenden Konvergenzuntersuchungen wird immer stillschweigend  $x \neq x_*$  vorausgesetzt.

Vorlesung 24, 18. Dezember 2019

### Beispiele

1. Für jedes Monom  $f(x) = x^m$  mit  $m \in \mathbb{N}_0$  sowie der Abkürzung  $h = x - x_*$  gilt

$$\begin{aligned} \frac{f(x) - f(x_*)}{x - x_*} &= \frac{f(x_* + h) - f(x_*)}{h} \\ &= \frac{\left( \sum_{j=0}^m \binom{m}{j} \cdot h^j \cdot x_*^{m-j} \right) - x_*^m}{h} \\ &= \frac{(x_*^m + m \cdot h \cdot x_*^{m-1} + \dots + m \cdot h^{m-1} \cdot x_* + h^m) - x_*^m}{h} \\ &= m \cdot x_*^{m-1} + \sum_{j=2}^m \binom{m}{j} \cdot h^{j-1} \cdot x_*^{m-j} . \end{aligned}$$

Der Limes  $x \rightarrow x_*$  bzw.  $h \rightarrow 0$  liefert nun

$$f'(x_*) = m \cdot x_*^{m-1}$$

für alle  $x_* \in \mathbb{R}$ , wobei diese Formel insbesondere auch für  $m = 0$  bzw.  $m = 1$  im Sinne von  $f'(x_*) = 0$  und  $f'(x_*) = 1$  gilt. Da  $x_*$  beliebig ist, können wir am Ende auf beiden Seiten der Formel  $x_*$  durch  $x$  ersetzen, d.h. es gilt  $f'(x) = m \cdot x^{m-1}$  für alle  $x \in \mathbb{R}$ .



2. Analog berechnet man für jedes Polynom

$$f(x) = \alpha_m \cdot x^m + \alpha_{m-1} \cdot x^{m-1} + \dots + \alpha_2 \cdot x^2 + \alpha_1 \cdot x + \alpha_0$$

die Formel

$$f'(x) = m \cdot \alpha_m \cdot x^{m-1} + (m-1) \cdot \alpha_{m-1} \cdot x^{m-2} + \dots + 2 \cdot \alpha_2 \cdot x + \alpha_1.$$

3. Mit  $f(x) = 1/x$  für  $x \neq 0$  berechnen wir für jedes feste  $x_*$  den Differenzenquotienten

$$\begin{aligned} \frac{f(x) - f(x_*)}{x - x_*} &= \frac{f(x_* + h) - f(x_*)}{h} = \frac{1}{h \cdot (x_* + h)} - \frac{1}{h \cdot x_*} \\ &= \frac{x_* - (x_* + h)}{h \cdot (x_* + h) \cdot x_*} = -\frac{1}{(x_* + h) \cdot x_*} \end{aligned}$$

und erhalten  $f'(x_*) = -1/x_*^2$  nach Grenzübergang  $h \rightarrow 0$ . Alternativ können wir auch

$$\frac{d}{dx} \frac{1}{x} = -\frac{1}{x^2}$$

schreiben.

4. Analog zeigt man für alle  $m \in \mathbb{N}$  und jedes  $x \neq 0$  die Formel

$$\frac{d}{dx} \left( \frac{1}{x^m} \right) = -\frac{m}{x^{m+1}}.$$

Ganz allgemein gilt  $\frac{d}{dx} x^k = k \cdot x^{k-1}$  für alle  $k \in \mathbb{Z}$  und  $x \neq 0$ .

5. Die Theorie der Potenzreihen (siehe nächstes Kapitel) impliziert

$$\exp'(x) = \exp(x), \quad \sin'(x) = \cos(x), \quad \cos'(x) = -\sin(x).$$

6. Die Differentiationsregeln für Umkehrfunktionen (siehe unten) garantieren

$$\ln'(x) = \frac{1}{x}$$

für alle  $x > 0$ .

7. Die Betragsfunktion  $f: \mathbb{R} \rightarrow \mathbb{R}$  mit  $f(x) = |x|$  ist an jeder Stelle  $x_* < 0$  bzw.  $x_* > 0$  differenzierbar mit  $f'(x_*) = -1$  bzw.  $f'(x_*) = +1$ , denn es gilt

$$\frac{|x_* + h| - |x_*|}{h} = \frac{-h}{h} = -1 \quad \text{bzw.} \quad \frac{|x_* + h| - |x_*|}{h} = \frac{+h}{h} = +1$$

für alle hinreichend kleinen  $h$  (es sollte  $|h| < |x_*|$  gelten). Die Betragsfunktion ist aber nicht in  $x_* = 0$  differenzierbar, denn es gilt

$$\lim_{h \searrow 0} \frac{|h|}{h} = 1 \neq -1 = \lim_{h \nearrow 0} \frac{|h|}{h},$$

d.h. der rechtsseitige und der linksseitige Grenzwert des Differenzenquotienten sind verschieden. Das ist nicht verwunderlich, da ja die Betragsfunktion einen *Knick* bei  $x = 0$  aufweist.

8. Die Vorzeichenfunktion  $\text{sgn}: \mathbb{R} \rightarrow \mathbb{R}$  ist in jedem Punkt  $x \neq 0$  differenzierbar mit  $\text{sgn}'(x) = 0$ , aber die Ableitung existiert nicht an der Stelle  $x = 0$ .

**Lemma** Wenn  $f : D \rightarrow \mathbb{R}$  an der Stelle  $x_*$  differenzierbar ist, so ist  $f$  auch stetig an der Stelle  $x_*$ . Die Umkehrung ist aber nicht richtig (zum Beispiel Knickstellen!).

Beweis: Für jedes  $x \neq x_*$  gilt

$$f(x) - f(x_*) = \frac{f(x) - f(x_*)}{x - x_*} \cdot (x - x_*)$$

und da die rechte Seite nach Voraussetzung wegen der Rechenregeln für Grenzwerte für  $x \rightarrow x_*$  gegen  $f'(x_*) \cdot 0 = 0$  konvergiert, gilt also  $\lim_{x \rightarrow x_*} (f(x) - f(x_*)) = 0$  und damit  $\lim_{x \rightarrow x_*} f(x) = f(x_*)$ .  $\square$

### Bemerkung

1. Differenzierbarkeit ist (wie Stetigkeit) zunächst eine punktweise Eigenschaft, die an jeder Stelle aus  $D$  überprüft werden muss. Es gibt seltsame Funktionen, die zwar überall stetig sind, aber an keiner oder nur an sehr wenigen Stellen differenzierbar sind.
2. Ist  $f : D \rightarrow \mathbb{R}$  an jeder Stelle  $x \in D$  differenzierbar, so existiert die Ableitung als Funktion  $f' : D \rightarrow \mathbb{R}$ .
3. Besonders wichtig sind stetig differenzierbare Funktionen, bei denen  $f' : D \rightarrow \mathbb{R}$  nicht nur in jeder Stelle wohldefiniert, sondern auch stetig ist.
4. Es gibt das allgemeinere Konzept der einseitigen Ableitungen

$$\lim_{x \searrow x_*} \frac{f(x) - f(x_*)}{x - x_*} \quad \lim_{x \nearrow x_*} \frac{f(x) - f(x_*)}{x - x_*},$$

das vor allem bei Knickstellen sinnvoll ist.

5. Ableitungen kann man für Funktionen  $f : \mathbb{C} \rightarrow \mathbb{C}$  ganz analog via

$$f'(z_*) = \lim_{z \rightarrow z_*} \frac{f(z) - f(z_*)}{z - z_*}$$

für  $z, z_* \in \mathbb{C}$  einführen, aber die entsprechende Theorie ist ganz anders als im Reellen. Wir werden komplexe Ableitungen erst im dritten Teil dieses Vorlesungszyklus behandeln.

**Elementare Rechenregeln für Ableitungen** Für zwei an der Stelle  $x_* \in D$  differenzierbare Funktionen  $f, g : D \rightarrow \mathbb{R}$  gelten die folgenden Aussagen:

1. (Linearität der Ableitung) Die obigen Definitionen implizieren

$$(\lambda \cdot f + \mu \cdot g)'(x_*) = \lambda \cdot f'(x_*) + \mu \cdot g'(x_*)$$

für beliebige reelle Zahlen  $\lambda, \mu \in \mathbb{R}$ .

2. (Produktregel) Es gilt auch

$$(f \cdot g)'(x_*) = f'(x_*) \cdot g(x_*) + f(x_*) \cdot g'(x_*).$$

In der Tat, für  $h \neq 0$  gilt

$$\begin{aligned} & \frac{f(x_* + h) \cdot g(x_* + h) - f(x_*) \cdot g(x_*)}{h} \\ &= \frac{f(x_* + h) - f(x_*)}{h} \cdot g(x_* + h) + f(x_*) \cdot \frac{g(x_* + h) - g(x_*)}{h} \end{aligned}$$

und der Limes  $h \rightarrow 0$  liefert die gewünschte Formel, da  $f$  und  $g$  in  $x_*$  differenzierbar und damit auch stetig sind.

3. (Quotientenregel) Unter der Zusatzannahme  $g(x_*) \neq 0$  gilt

$$\left(\frac{f}{g}\right)'(x_*) = \frac{f'(x_*) \cdot g(x_*) - f(x_*) \cdot g'(x_*)}{g^2(x_*)},$$

wobei diese Regel aus der Produkt- und der Kettenregel abgeleitet werden kann (Übungsaufgabe).

**Beispiele** Für alle  $x$  im Definitionsbereich der jeweiligen Funktionen gilt

$$\frac{d}{dx} (3x^2 + 4 \sin(x)) = 3 \frac{d}{dx} x^2 + 4 \frac{d}{dx} \sin(x) = 6x + 4 \cos(x)$$

sowie

$$\frac{d}{dx} (\ln(x) \cdot \sin(x)) = \ln'(x) \cdot \sin(x) + \ln(x) \cdot \sin'(x) = \frac{\sin(x)}{x} + \ln(x) \cdot \cos(x)$$

und

$$\begin{aligned} \frac{d}{dx} \tan(x) &= \frac{d}{dx} \left( \frac{\sin(x)}{\cos(x)} \right) = \frac{\sin'(x) \cdot \cos(x) - \sin(x) \cdot \cos'(x)}{\cos^2(x)} \\ &= \frac{\cos^2(x) + \sin^2(x)}{\cos^2(x)} = \frac{1}{\cos^2(x)} = \sec^2(x), \end{aligned}$$

wobei  $\sec(x) = 1/\cos(x)$  der Sekans von  $x$  ist. Der Kosekans ist übrigens  $\csc(x) = 1/\sin(x)$ .

**Kettenregel für Ableitungen** Ist  $f : D \rightarrow E$  in  $x_*$  differenzierbar und ist ausserdem  $g : E \rightarrow \mathbb{R}$  in  $f(x_*)$  differenzierbar, so ist auch  $g \circ f$  in  $x_*$  differenzierbar und es gilt

$$(g \circ f)'(x_*) = \frac{d}{dx} g(f(x)) = g'(f(x_*)) \cdot f'(x_*),$$

wobei die Terme  $g'(f(x_*))$  bzw.  $f'(x_*)$  manchmal die äußere bzw. die innere Ableitung genannt werden.

Ein rigoroser Beweis findet sich in [AORS, Seite 235], aber informell kann die Kettenregel wie folgt begründet werden: Für  $h \neq 0$  betrachten wir

$$\eta(h) := f(x_* + h) - f(x_*) \xrightarrow{h \rightarrow 0} 0$$

und schreiben

$$\begin{aligned}\frac{g(f(x_* + h)) - g(f(x_*))}{h} &= \frac{g(f(x_* + h)) - g(f(x_*))}{f(x_* + h) - f(x_*)} \cdot \frac{f(x_* + h) - f(x_*)}{h} \\ &= \frac{g(f(x_*) + \eta(h)) - g(f(x_*))}{\eta(h)} \cdot \frac{f(x_* + h) - f(x_*)}{h},\end{aligned}$$

wobei wir stillschweigend vorausgesetzt haben (und das ist die kleine Lücke in unserem „Beweis“), dass  $\eta(h) \neq 0$  für  $h \neq 0$  gilt. Die behauptete Formel folgt nun durch Grenzübergang  $h \rightarrow 0$ .

**Beispiele** Durch direkte Rechnungen verifizieren wir

$$\frac{d}{dx} (\ln(1 + x^2)) = \ln'(1 + x^2) \cdot (2 \cdot x) = \frac{2 \cdot x}{1 + x^2}$$

sowie

$$\begin{aligned}\frac{d}{dx} \exp(\sin(x^3)) &= \exp'(\sin(x^3)) \cdot \sin'(x^3) \cdot (3 \cdot x^2) \\ &= 3 \cdot x^2 \cdot \cos(x^3) \cdot \exp(\sin(x^3)).\end{aligned}$$

**Ableitung der Umkehrfunktion** Besitzt die Funktion  $f : D \rightarrow E$  die Umkehrfunktion  $f^{-1} : E \rightarrow D$  und ist  $f$  in  $x_*$  differenzierbar mit  $f'(x_*) \neq 0$ , so ist  $f^{-1}$  in  $y_* = f(x_*)$  differenzierbar mit

$$(f^{-1})'(y_*) = \frac{1}{f'(x_*)}.$$

Auch hier verweisen wir für einen rigorosen Beweis auf die Literatur (etwa [AORS, Seite 235]), wollen aber erwähnen, dass die Formel mittels der Kettenregel durch symbolische Differentiation von

$$x = f^{-1}(f(x))$$

nach  $x$  abgeleitet werden kann. In der Tat, wir erhalten

$$1 = \frac{d}{dx} f^{-1}(f(x)) = (f^{-1})'(f(x)) \cdot f'(x) \quad \text{bzw.} \quad (f^{-1})'(f(x)) = \frac{1}{f'(x)}$$

für alle  $x$  und damit auch für  $x = x_*$ . Alternativ hätten wir in

$$y = f(f^{-1}(y))$$

beide Seiten nach  $y$  differenzieren können, um  $(f^{-1})'(y) = 1/f'(f^{-1}(y))$  für alle  $y$  zu erhalten.

**Beispiele**

1. Für alle  $x \in \mathbb{R}$  gilt

$$x = \exp(\ln(x))$$

und durch Differentiation nach  $x$  erhalten wir

$$1 = \exp'(\ln(x)) \cdot \ln'(x) = \exp(\ln(x)) \cdot \ln'(x) = x \cdot \ln'(x)$$

und damit  $\ln'(x) = 1/x$  für alle  $x > 0$ .

2. Aus

$$x = \sin(\arcsin(x)) \quad \text{für alle } x \in (-1, +1)$$

ergibt sich

$$\begin{aligned} 1 &= \sin'(\arcsin(x)) \cdot \arcsin'(x) \\ &= \cos(\arcsin(x)) \cdot \arcsin'(x) \\ &= \sqrt{1 - \sin^2(\arcsin(x))} \cdot \arcsin'(x) \\ &= \sqrt{1 - x^2} \cdot \arcsin'(x), \end{aligned}$$

und damit  $\arcsin'(x) = 1/\sqrt{1-x^2}$ . Hierbei haben wir benutzt, dass  $\arcsin$  Werte in  $(-\pi/2, +\pi/2)$  annimmt und dass daher  $\cos(\arcsin(x))$  immer positiv ist. Beachte, dass  $f'(x) \rightarrow \infty$  für  $x \rightarrow \pm 1$ .

Vorlesung 25, 20. Dezember 2019

**Höhere Ableitungen** Ist  $f$  differenzierbar, so stellt sich die natürliche Frage, ob denn auch die Funktion  $f' : D \rightarrow \mathbb{R}$  an einzelnen oder gar an allen Stellen  $x_* \in D$  differenzierbar ist. Ist die Antwort positiv, so spricht man von der zweiten Ableitung von  $f$  und schreibt

$$f''(x_*) = \lim_{x \rightarrow x_*} \frac{f'(x) - f'(x_*)}{x - x_*}$$

an einer ausgezeichneten Stelle  $x_*$  bzw.

$$\frac{d^2}{dx^2} f(x) = \frac{d}{dx} f'(x),$$

sofern die zweite Ableitung im gesamten Definitionsbereich existiert. Analog kann man dritte, vierte usw. Ableitungen einführen, wobei man dann die  $n$ -te Ableitung üblicherweise als  $f^{(n)}$  und nicht als  $f''''$  schreibt. Insbesondere gilt

$$f^{(n)}(x_*) = \lim_{x \rightarrow x_*} \frac{f^{(n-1)}(x) - f^{(n-1)}(x_*)}{x - x_*}$$

bzw.

$$f^{(n)}(x) = \frac{d^n}{dx^n} f(x) = \frac{d}{dx} f^{(n-1)}(x).$$

Aus Konsistenzgründen schreibt man auch  $f^{(0)}(x) = f(x)$ , d.h. jede Funktion gleicht ihrer nullten Ableitung.

### Beispiele

1. Für das Monom vom Grad  $m \in \mathbb{N}$  gilt

$$f^{(0)}(x) = x^m, \quad f^{(1)}(x) = m \cdot x^{m-1}, \quad f^{(2)}(x) = m \cdot (m-1) \cdot x^{m-2}$$

usw. bis

$$f^{(m-1)}(x) = m \cdot (m-1) \cdot \dots \cdot 2 \cdot x^1, \quad f^{(m)}(x) = m!$$

und dann

$$f^{(m+n)}(x) = 0$$

für alle  $x$  und alle  $n \in \mathbb{N}$ .

2. Für Monome mit negativem Exponent gilt

$$f^{(0)}(x) = \frac{1}{x^m}, \quad f^{(1)}(x) = -\frac{m}{x^{m+1}}, \quad f^{(2)}(x) = -\frac{m(m+1)}{x^{m+2}}, \quad \dots$$

3. Es gilt  $\sin' = \cos$ ,  $\sin'' = -\sin$ ,  $\sin''' = -\cos$ ,  $\sin'''' = \sin$  und analog für den Kosinus.

**\*Räume differenzierbarer Funktionen** Die Menge

$$C^k(D) := \{f : D \rightarrow \mathbb{R} : f \text{ ist } k\text{-mal stetig differenzierbar}\}$$

ist ein Vektorraum unendlicher Dimension. Ist  $D$  ein abgeschlossenes Intervall, so kann man durch

$$\|f\|_{k,\infty} := \|f^{(0)}\|_{\infty} + \dots + \|f^{(k)}\|_{\infty} = \sum_{j=0}^k \max_{x \in D} |f^{(j)}(x)|$$

in sinnvoller Weise eine Norm einführen, wobei  $\|\cdot\|_{\infty}$  gerade die Norm in  $C(D)$ , also im Raum der stetigen Funktionen ist.

## Kurven in Vektorräumen

**Verallgemeinerung** Die Konzepte *Differenzierbarkeit* und *Ableitung* können auch auf vektorwertige Abbildungen  $\mathbf{f} : D \subset \mathbb{R} \rightarrow V$  verallgemeinert werden, sofern  $V$  ein endlich-dimensionaler Vektorraum mit Norm  $\|\cdot\|$  ist, wobei man dann oftmals die Elemente aus  $D$  mit  $t$  und nicht mit  $x$  bezeichnet. Insbesondere gilt

$$\mathbf{f}'(t) = \frac{d}{dt}\mathbf{f}(t) = \lim_{h \rightarrow 0} h^{-1}(\mathbf{f}(t+h) - \mathbf{f}(t)) \in V.$$

Hierbei sind  $h \neq 0$  und  $t$  reelle Zahlen, wohingegen  $\mathbf{f}(t)$  und  $\mathbf{f}'(t)$  immer Elemente des Vektorraumes  $V$  sind.

### Bemerkung

1. Zu einer Abbildung  $\mathbf{x} : D \rightarrow \mathbb{R}^n$  gibt es in natürlicher Weise Komponentenabbildungen  $x_i : D \rightarrow \mathbb{R}$ , so dass

$$\mathbf{x}(t) = \begin{pmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{pmatrix}.$$

Man kann nun leicht zeigen, dass  $\mathbf{x}$  genau dann differenzierbar ist, wenn alle Komponenten dies sind, wobei

$$\mathbf{x}'(t) = \begin{pmatrix} x_1'(t) \\ \vdots \\ x_n'(t) \end{pmatrix}.$$

2. Eine Abbildung  $z : D \subset \mathbb{R} \rightarrow \mathbb{C}$  entspricht in natürlicher Weise zwei Abbildungen  $x : D \rightarrow \mathbb{R}$  und  $y : D \rightarrow \mathbb{R}$ , so dass

$$x(t) = \operatorname{Re}(z(t)), \quad y(t) = \operatorname{Im}(z(t))$$

und damit

$$z(t) = x(t) + i \cdot y(t), \quad z'(t) = x'(t) + i \cdot y'(t).$$

Beachte, dass bei uns  $t$  immer noch reell ist; der Fall komplexer Argumente  $t$  wird erst in *Mathe III* diskutiert.

3. Eine Abbildung  $\mathbf{A} : D \subset \mathbb{R} \rightarrow \mathbb{C}^{(m,n)}$  und ihre Ableitung (sofern existent) können als

$$\mathbf{A}(t) = \begin{pmatrix} a_{11}(t) & \dots & a_{1n}(t) \\ \vdots & & \vdots \\ a_{m1}(t) & \dots & a_{mn}(t) \end{pmatrix}, \quad \mathbf{A}'(t) = \begin{pmatrix} a'_{11}(t) & \dots & a'_{1n}(t) \\ \vdots & & \vdots \\ a'_{m1}(t) & \dots & a'_{mn}(t) \end{pmatrix}$$

geschrieben werden, wobei  $a_{ij} : D \rightarrow \mathbb{R}$  die  $ij$ -te Komponentenfunktion ist.

4. Komponentenabbildungen gibt es auch für allgemeine Vektorräume, aber dann muss zunächst eine Basis gewählt werden. Es gilt dann

$$\frac{d}{dt} \left( \sum_{j=1}^n \nu_j(t) \mathbf{v}_j \right) = \sum_{j=1}^n \nu'_j(t) \mathbf{v}_j$$

bzw.

$$\frac{d}{dt} \left( \sum_{j=1}^n \nu_j(t) \mathbf{v}_j(t) \right) = \sum_{j=1}^n \left( \nu'_j(t) \mathbf{v}_j(t) + \nu_j(t) \mathbf{v}'_j(t) \right)$$

für den Fall, dass die Basisvektoren  $\mathbf{v}_j$  selbst von  $t$  abhängen.

### Beispiele

1. Die Position eines zur Zeit  $t = 0$  aus dem Koordinatenursprung abgefeuerten Projektils kann (stark vereinfacht) durch

$$\mathbf{x}(t) = \begin{pmatrix} v_1 \cdot t \\ v_2 \cdot t \\ v_3 \cdot t - \frac{1}{2}gt^2 \end{pmatrix}$$

beschrieben werden, wobei  $\mathbf{x}(t)$  die momentane Position bezeichnet,  $v_j$  die  $j$ -te Komponente der Anfangsgeschwindigkeit  $\mathbf{v} \in \mathbb{R}^3$  ist und  $g$  für die Erdbeschleunigung steht. Die Vektoren  $\mathbf{x}'(t)$  und  $\mathbf{x}''(t)$  geben für jedes  $t$  gerade die Geschwindigkeit und die Beschleunigung an. Die Formel macht allerdings nur Sinn, solange  $x_3(t) > 0$  gilt, d.h. solange das Projektil nicht wieder den Erdboden erreicht hat, und sofern  $\|\mathbf{v}\|$  nicht zu groß ist.

2. Ein vertikal in einer Ebene aufgehängtes Pendel kann näherungsweise durch

$$\mathbf{x}(t) = \begin{pmatrix} \sin(\omega t) \\ -\cos(\omega t) \end{pmatrix}$$

beschrieben werden, wobei  $\omega t$  gerade der Winkel zur vertikalen Lotgeraden ist. Man könnte dies auch durch die komplexe Funktion

$$z(t) = \sin(\omega t) - i \cos(\omega t) = \exp(i\omega t - i\pi/2) = -i \exp(i\omega t)$$

modellieren.

3. Ein einfaches matrizenwertiges Beispiel ist

$$\mathbf{A}(t) = \begin{pmatrix} \cos(t) & \sin(t) \\ -\sin(t) & \cos(t) \end{pmatrix}, \quad \mathbf{A}'(t) = -\begin{pmatrix} \sin(t) & -\cos(t) \\ \cos(t) & \sin(t) \end{pmatrix}.$$

Insbesondere gilt  $\mathbf{A}''(t) + \mathbf{A}(t) = 0$  für alle  $t$ , d.h.  $\mathbf{A}$  löst eine sehr einfache *Differentialgleichung* zweiter Ordnung.

**Bemerkung** Eine Abbildung  $\mathbf{f} : D \rightarrow V$  nennt man auch parametrisierte Kurve in  $V$ , wobei  $t$  oftmals als Zeit interpretiert werden kann. In diesem Sinne beschreibt eine Kurve die zeitliche Änderung eines Vektors und  $f(t)$  kann als momentane Position im Vektorraum  $V$  verstanden werden. Der Ableitungsvektor  $\mathbf{f}'(t)$  ist dann gerade der momentane Geschwindigkeitsvektor bzw. der sogenannte Tangentialvektor an die Kurve. In der Physik und der Mathematischen Dynamik schreibt man auch gerne

$$\dot{\mathbf{f}}(t) \quad \text{und} \quad \ddot{\mathbf{f}}(t) \quad \text{statt} \quad \mathbf{f}'(t) \quad \text{und} \quad \mathbf{f}''(t).$$

**Geometrie von Kurven im  $\mathbb{R}^2$**  Eine planare Kurve wird durch eine (zweimal stetig differenzierbare) Abbildung  $\mathbf{x} : D \rightarrow \mathbb{R}^2$  mit

$$D \ni t \quad \mapsto \quad \mathbf{x}(t) = \begin{pmatrix} x_1(t) \\ x_2(t) \end{pmatrix} \in \mathbb{R}^2$$

beschrieben, wobei  $D$  üblicherweise ein Intervall ist und die Abbildung  $\mathbf{x}$  auch Parametrisierung der Kurve genannt wird. Man kann nun mit Hilfe von nullten, ersten und zweiten Ableitungen die folgenden geometrischen Konzepte einführen bzw. berechnen:

$$\ell(s) = \|\mathbf{x}'(t)\| = \sqrt{(x_1'(t))^2 + (x_2'(t))^2} \quad (\text{infinitesimales Längenelement})$$

$$\mathbf{b}_1(t) = \frac{\mathbf{x}'(t)}{\|\mathbf{x}'(t)\|} = \frac{1}{\ell(t)} \begin{pmatrix} x_1'(t) \\ x_2'(t) \end{pmatrix} \quad (\text{normierter Tangentialvektor})$$

$$\mathbf{b}_2(t) = \mathbf{b}_1(t)^\perp = \frac{1}{\ell(t)} \begin{pmatrix} -x_2'(t) \\ x_1'(t) \end{pmatrix} \quad (\text{normierter Normalenvektor})$$

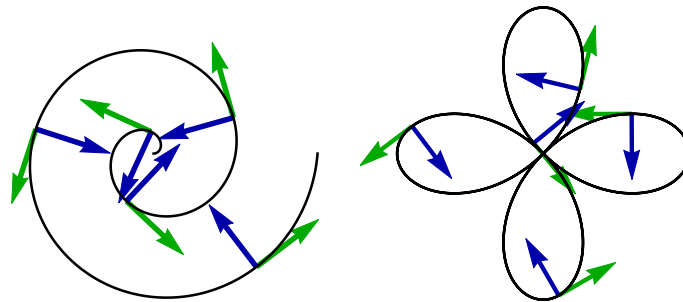
$$\kappa(t) = \frac{x_1''(t) \cdot x_2'(t) - x_2''(t) \cdot x_1'(t)}{\ell(t)^3} \quad (\text{Krümmung der Kurve})$$



**Beispiele** Die Formeln

$$\mathbf{x}(t) = t \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix} \quad \text{bzw.} \quad \mathbf{x}(t) = 2 \cos(2t) \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix}$$

beschreiben eine Spirale bzw. ein vierblättriges Kleeblatt (siehe Bild und Film), wobei die zweite Kurve sogar geschlossen ist .



Die zwei im Text beschriebenen Kurven sowie die mitbewegte ON-Basis  $\mathbf{b}_1(t)$ ,  $\mathbf{b}_2(t)$  (Frenetsches Zweibein) für ausgewählte Werte von  $t$ .



# Kapitel 9

## Differentialrechnung

Vorlesung 26, 06. Januar 2020

### 9.1 Extrema, Mittelwertsätze und l'Hospital

**Annahme** In diesem Abschnitt betrachten wir Funktionen  $f : [a, b] \rightarrow \mathbb{R}$ , die stetig in  $[a, b]$  und differenzierbar in  $(a, b)$  sind.

**Lokale und globale Extrema** Wir sagen, die Funktion  $f$  nimmt an der Stelle (oder im Punkt)  $x_* \in [a, b]$

1. ein globales Maximum bzw. ein striktes globales Maximum an, falls  $f(x_*) \geq f(x)$  bzw.  $f(x_*) > f(x)$  für alle  $x \in [a, b]$  mit  $x \neq x_*$  gilt,
2. ein lokales Maximum bzw. ein striktes lokales Maximum an, sofern ein  $\varepsilon > 0$  existiert, so dass die entsprechende Ungleichung an jeder Stelle  $x \in [a, b]$  mit  $0 < |x - x_*| < \varepsilon$  gilt.

Analog werden die verschiedenen Arten von Minima definiert und ein Extremum ist entweder ein Minimum oder ein Maximum, wobei  $x_*$  dann Extremalstelle heißt.

**Lemma (notwendige Bedingung für Extrema in inneren Punkten)** Nimmt  $f$  an der Stelle  $x_* \in (a, b)$  ein lokales Extremum (strikt oder nicht) an, so gilt  $f'(x_*) = 0$ .

**Beweis:** Für alle  $x \in [a, b]$ , die hinreichend nahe bei einer Maximalstelle  $x_* \in (a, b)$  liegen, gilt

$$\frac{f(x) - f(x_*)}{x - x_*} \geq 0 \quad \text{für } x < x_*, \quad \frac{f(x) - f(x_*)}{x - x_*} \leq 0 \quad \text{für } x > x_*.$$

Durch die einseitigen Grenzübergänge  $x \nearrow x_*$  bzw.  $x \searrow x_*$  erhalten wir

$$f'(x_*) \geq 0 \quad \text{bzw.} \quad f'(x_*) \leq 0$$

und damit die Behauptung. Minimalstellen können analog behandelt werden.

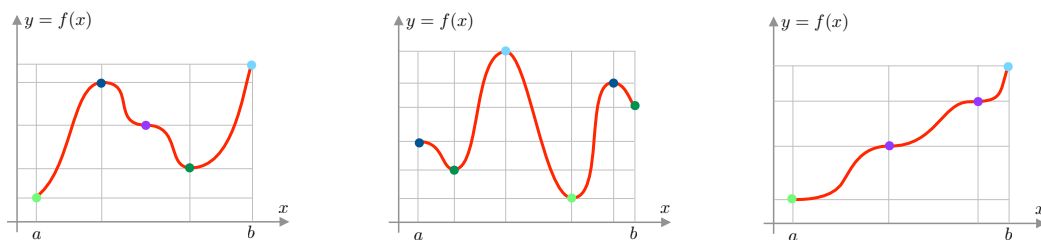
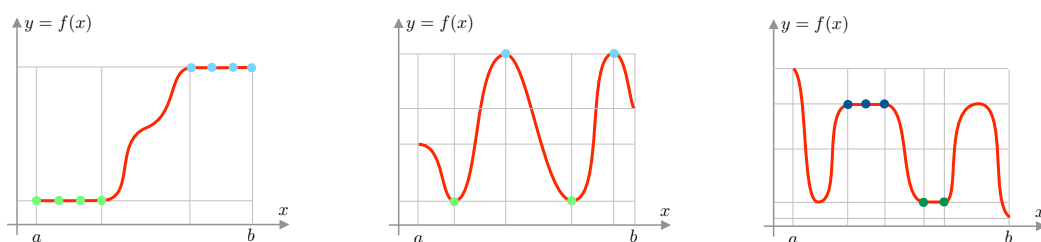


Illustration der verschiedenen Konzepte von Extremum / Extremalstelle. Globale Minima (hellgrün), nur-lokale Minima (dunkelgrün), globale Maxima (hellblau), nur-lokale Maxima (dunkelblau) sowie Sattelpunkte (lila).



Beispiele für nicht-strikte Extrema (Farbkodierung wie oben).

### Bemerkung

1. Da  $[a, b]$  kompakt ist und  $f$  stetig ist, existiert immer (mindestens) ein globales Minimum sowie (mindestens) ein globales Maximum, die aber jeweils nicht strikt sein müssen.
2. Lokale Extrema können, aber müssen nicht existieren.
3. Jedes globale Extremum ist nach unserer Definition auch ein lokales Extremum.
4. Nicht-strikte Extremalstellen können isoliert sein oder nicht, strikte sind immer isoliert (siehe Bilder).
5. Jeder innere Punkt  $x_* \in (a, b)$  mit  $f'(x_*) = 0$  heißt stationärer Punkt oder kritischer Punkt von  $f$ . Nicht jeder kritische Punkt ist aber lokale Extremalstelle, denn es gibt auch Sattelpunkte (siehe Bilder).
6.  $f$  kann im linken Randpunkt  $x_* = a$  ein lokales oder globales Extremum annehmen, aber dann gilt  $f'(a) \geq 0$  für ein Minimum bzw.  $f'(a) \leq 0$  für ein Maximum (sofern  $f'(a)$  als Grenzwert einseitiger Differenzenquotienten existiert).
7. Für Minima bzw. Maxima, die am rechten Randpunkt  $x_* = b$  angenommen werden, gilt  $f'(b) \leq 0$  bzw.  $f'(b) \geq 0$ .

**Theorem (Mittelwertsätze)** Ist  $f$  eine Funktion wie oben, so gilt:

1. (1. Mittelwertsatz) Es existiert  $x_* \in (a, b)$ , so dass  $f'(x_*) = \frac{f(b) - f(a)}{b - a}$ .
2. (Satz von Rolle) Im Fall von  $f(a) = f(b)$  gibt es ein  $x_* \in (a, b)$  mit  $f'(x_*) = 0$ .
3. (2. Mittelwertsatz) Ist  $g$  eine weitere Funktion wie oben mit  $g'(x) \neq 0$  für jedes  $x \in (a, b)$ , so gilt

$$\frac{f'(x_*)}{g'(x_*)} = \frac{f(b) - f(a)}{g(b) - g(a)}$$

für ein geeignetes  $x_* \in (a, b)$ .

*Beweis, Teil A:* Obwohl der Satz von Rolle ein Spezialfall sowohl des 1. und des 2. Mittelwertsatzes ist, beweisen wir ihn zuerst. Wir wählen  $\underline{x} \in [a, b]$  und  $\bar{x} \in [a, b]$ , so dass  $f$  in  $\underline{x}$  bzw. in  $\bar{x}$  ein globales Minimum bzw. Maximum annimmt. Im Entartungsfall  $f(\underline{x}) = f(\bar{x})$  ist  $f$  konstant, d.h. es gilt  $f'(x) = 0$  für jedes  $x \in (a, b)$  und wir können  $x_*$  beliebig wählen. Andernfalls liegt  $\underline{x}$  oder  $\bar{x}$  im offenen Intervall  $(a, b)$  (oder beide) und wir können  $x_* = \underline{x}$  oder  $x_* = \bar{x}$  wählen.

*Beweis, Teil B:* Um den zweiten Mittelwertsatz zu beweisen, bemerken wir, dass nach Voraussetzung an  $g$  auch  $g(b) \neq g(a)$  gelten muss, da andernfalls  $g'$  nach dem Satz von Rolle ein Nullstelle besäße. Wir können daher neben  $f$  auch die Funktion  $\tilde{f} : [a, b] \rightarrow \mathbb{R}$  mit

$$\tilde{f}(x) := f(x) - (g(x) - g(a)) \cdot \frac{f(b) - f(a)}{g(b) - g(a)}$$

betrachten, wobei  $\tilde{f}(a) = \tilde{f}(b) = f(a)$  und

$$\tilde{f}'(x) = f'(x) - g'(x) \cdot \frac{f(b) - f(a)}{g(b) - g(a)}$$

nach Konstruktion gilt. Der zweite Mittelwertsatz ergibt sich nun, indem wir den Satz von Rolle auf  $\tilde{f}$  anwenden. Der erste Mittelwertsatz ist dann gerade der Spezialfall  $g(x) = x$ .  $\square$

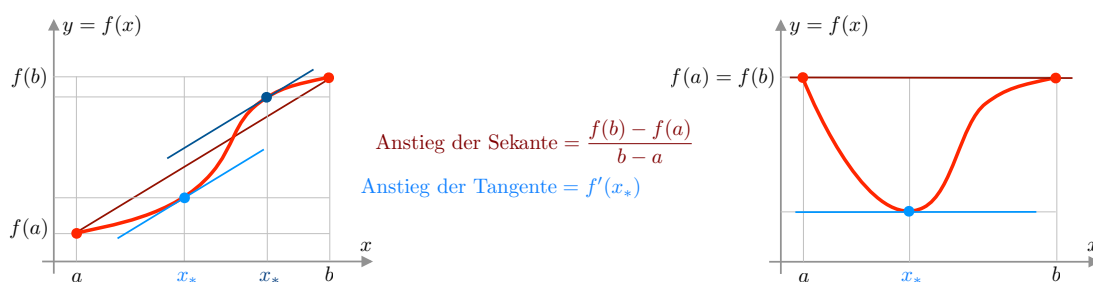


Illustration des 1. Mittelwertsatzes (links) sowie des Satzes von Rolle (rechts) als Spezialfall. Der zweite Mittelwertsatz ist eine Verallgemeinerung des ersten.

**Folgerung** Gilt  $f'(x) = 0$  für alle  $x \in (a, b)$ , so ist  $f$  konstant auf  $[a, b]$ .

*Beweis:* Angenommen, es existieren  $\tilde{a}, \tilde{b} \in [a, b]$  mit  $f(\tilde{a}) \neq f(\tilde{b})$  und  $\tilde{a} < \tilde{b}$ . Dann existiert nach dem Mittelwertsatz, angewendet auf die Einschränkung von  $f$  auf das Intervall  $[\tilde{a}, \tilde{b}]$ , ein  $x_* \in (\tilde{a}, \tilde{b}) \subset (a, b)$  mit

$$f'(x_*) = \frac{f(\tilde{b}) - f(\tilde{a})}{\tilde{b} - \tilde{a}} \neq 0,$$

aber dies ist ein Widerspruch zur Annahme.  $\square$

**Folgerung (Monotonie und Vorzeichen der ersten Ableitung)**

1. Die Funktion  $f$  ist genau dann monoton wachsend (bzw. fallend) auf  $[a, b]$ , wenn  $f'(x) \geq 0$  (bzw.  $f'(x) \leq 0$ ) für alle  $x \in (a, b)$  gilt.

2. Gilt sogar  $f'(x) > 0$  (bzw.  $f'(x) < 0$ ) für alle  $x \in (a, b)$ , so ist  $f$  strikt monoton wachsend (bzw. strikt monoton fallend).

*Beweis:* Alle Implikationen, die aus einer Vorzeichenbedingung an  $f'$  die Monotonie von  $f$  folgern, können direkt aus dem Mittelwertsatz bzw. analog zum Beweis des Satzes von Rolle abgeleitet werden (Übungsaufgabe). Die Umkehrungen im ersten Teil können mit Vorzeichenuntersuchungen für Differenzenquotienten und anschließendem Grenzübergang begründet werden.  $\square$

**Bemerkung** Die erste Aussage der Folgerung ist eine Äquivalenz, die zweite nicht. Zum Beispiel ist das kubische Monom  $f(x) = x^3$  strikt monoton wachsend auf  $[-1, 1]$  mit  $f'(0) = 0$ .

**Theorem (Regel von l'Hospital für Grenzwerte „Null durch Null“)** Für je zwei Funktionen  $f$  und  $g$  wie oben mit

$$f(x_*) = g(x_*) = 0, \quad g'(x) \neq 0 \quad \text{für } x \in (a, b)$$

gilt

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

sofern der Grenzwert auf der rechten Seite existiert. Insbesondere erhalten wir

$$\lim_{x \rightarrow x_*} \frac{f(x)}{g(x)} = \frac{f'(x_*)}{g'(x_*)}$$

für den Fall, dass der Quotient auf der rechten Seite wohldefiniert ist.

*Beweis:* Für jedes  $x \in [a, x_*)$  existiert nach dem 2. Mittelwertsatz — angewendet auf die Einschränkung von  $f$  auf das Intervall  $[x, x_*)$  — ein  $\hat{x} \in (x, x_*)$  mit

$$\frac{f(x)}{g(x)} = \frac{f(x) - f(x_*)}{g(x) - g(x_*)} = \frac{f'(\hat{x})}{g'(\hat{x})},$$

wobei  $\hat{x}$  im Allgemeinen von  $x$  und  $x_*$  abhängen wird. Analoges gilt für alle  $x \in (x_*, b]$ . Die Behauptung folgt nun nach Grenzübergang, wobei wir benutzen, dass  $x \rightarrow x_*$  auch  $\hat{x} \rightarrow x_*$  impliziert (denn es gilt  $|\hat{x} - x_*| \leq |x - x_*|$ ).

**Bemerkung**

1. Die Formel von l'Hospital gilt immer dann, wenn  $f$  und  $g$  in einem (kleinen) Intervall, das  $x_*$  als inneren Punkt enthält, differenzierbar sind und  $g'$  dort keine Nullstelle hat.
2. Manchmal ist der Quotient  $f'(x_*)/g'(x_*)$  selbst „Null durch Null“ und dann gilt

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

sofern  $f$  und  $g$  zweimal differenzierbar sind und der Limes auf der rechten Seite wohldefiniert ist.

3. In der Regel von l'Hospital kann  $x \rightarrow x_*$  durch die einseitigen Grenzwerte  $x \searrow x_*$  bzw.  $x \nearrow x_*$  oder die uneigentlichen Grenzwerte  $x \rightarrow -\infty$  bzw.  $x \rightarrow +\infty$  ersetzt werden.
4. Man kann die l'Hospital-Formel auch dann verwenden, wenn  $x_*$  eine isolierte Lücke im Definitionsbereich von  $f$  und  $g$  ist. Dies gilt insbesondere, wenn  $\lim_{x \rightarrow x_*} f(x) = \pm\infty$  und  $\lim_{x \rightarrow x_*} g(x) = \pm\infty$ , d.h. wenn der Quotient  $f(x_*)/g(x_*)$  „Unendlich durch Unendlich“ ist. Ein rigoroser Beweis findet sich in [AORS, Seite 255], aber formal kann dies mit

$$\begin{aligned} \lim_{x \rightarrow x_*} \frac{f(x)}{g(x)} &= \lim_{x \rightarrow x_*} \frac{\frac{1}{g(x)}}{\frac{1}{f(x)}} = \lim_{x \rightarrow x_*} \frac{-\frac{g'(x)}{g^2(x)}}{\frac{-f'(x)}{f^2(x)}} \\ &= \lim_{x \rightarrow x_*} \frac{\left(\frac{f(x)}{g(x)}\right)^2}{\frac{f'(x)}{g'(x)}} = \frac{\left(\lim_{x \rightarrow x_*} \frac{f(x)}{g(x)}\right)^2}{\lim_{x \rightarrow x_*} \frac{f'(x)}{g'(x)}} \end{aligned}$$

und nach einer elementaren Umstellung der Terme begründet werden.

### Beispiele

1. Direkte Anwendungen sind

$$\lim_{x \rightarrow 0} \frac{\sin(x)}{x} = \lim_{x \rightarrow 0} \frac{\cos(x)}{1} = 1, \quad \lim_{x \rightarrow 0} \frac{\cos(x) - 1}{x} = \lim_{x \rightarrow 0} \frac{-\sin(x)}{1} = 0.$$

2. Eine zweifache Anwendung ist

$$\lim_{x \rightarrow 0} \frac{1 - \cos(x)}{x^2} = \lim_{x \rightarrow 0} \frac{\sin(x)}{2 \cdot x} = \lim_{x \rightarrow 0} \frac{\cos(x)}{2} = \frac{1}{2}.$$

3. Manchmal muss man die Terme erst geeignet umformen. Ein Beispiel ist

$$\begin{aligned} \lim_{x \rightarrow 0} \left( \frac{1}{x} - \frac{1}{\sin(x)} \right) &= \lim_{x \rightarrow 0} \frac{\sin(x) - x}{x \cdot \sin(x)} = \lim_{x \rightarrow 0} \frac{\cos(x) - 1}{\sin(x) + x \cdot \cos(x)} \\ &= \lim_{x \rightarrow 0} \frac{-\sin(x)}{2 \cdot \cos(x) - x \cdot \sin(x)} = \frac{0}{2 \cdot 1 - 0 \cdot 0} = 0, \end{aligned}$$

wobei wir die Regel von l'Hospital wieder zweimal angewendet haben.

4. Ein einfaches Beispiel mit Polstellen ist

$$\lim_{x \searrow 0} x \ln(x) = \lim_{x \searrow 0} \frac{\ln(x)}{\frac{1}{x}} = \lim_{x \searrow 0} \frac{\frac{1}{x}}{-\frac{1}{x^2}} = -\lim_{x \searrow 0} x = 0.$$

Insbesondere erhalten wir damit

$$\lim_{x \searrow 0} x^x = \lim_{x \searrow 0} \exp(x \ln(x)) = \exp\left(\lim_{x \searrow 0} x \ln(x)\right) = \exp(0) = 1,$$

wobei wir die Stetigkeit der Exponentialfunktion  $\exp$  benutzt haben.

5. Manchmal ist auch eine Substitution sinnvoll. Zum Beispiel gilt

$$\lim_{x \rightarrow \infty} \left(1 + \frac{a}{x}\right)^x = \exp \left( \lim_{x \rightarrow \infty} \frac{\ln \left(1 + \frac{a}{x}\right)}{\frac{1}{x}} \right)$$

sowie

$$\lim_{x \rightarrow \infty} \frac{\ln \left(1 + \frac{a}{x}\right)}{\frac{1}{x}} = \lim_{s \searrow 0} \frac{\ln(1 + a \cdot s)}{s} = \lim_{s \searrow 0} \frac{\frac{a}{1 + a \cdot s}}{1} = a,$$

wobei wir  $x$  durch  $s = 1/x$  ersetzt haben. Insgesamt erhalten wir

$$\left(1 + \frac{a}{x}\right)^x \xrightarrow{x \rightarrow \infty} e^a$$

und damit eine Verallgemeinerung der Definition der Eulerschen Zahl mittels Folgen.

6. Ein weiteres Beispiel mit  $x \rightarrow \infty$  ist

$$\begin{aligned} \lim_{x \rightarrow \infty} \left( x \cdot \ln \left( \frac{x+1}{x-1} \right) \right) &= \lim_{x \rightarrow \infty} \frac{\ln(x+1) - \ln(x-1)}{\frac{1}{x}} = \lim_{x \rightarrow \infty} \frac{\frac{1}{x+1} - \frac{1}{x-1}}{-\frac{1}{x^2}} \\ &= \lim_{x \rightarrow \infty} \frac{2 \cdot x^2}{(x+1) \cdot (x-1)} = \lim_{x \rightarrow \infty} \frac{2}{\left(1 + \frac{1}{x}\right) \cdot \left(1 - \frac{1}{x}\right)} = 2. \end{aligned}$$

Das vorletzte Gleichheitszeichen hätten wir via

$$\lim_{x \rightarrow \infty} \frac{2 \cdot x^2}{(x+1) \cdot (x-1)} = \lim_{x \rightarrow \infty} \frac{4 \cdot x}{2 \cdot x} = 2.$$

auch aus der Regeln von l'Hospital ableiten können.

Vorlesung 27, 08. Januar 2020

## 9.2 Satz von Taylor

**Landau-Symbole** Ist  $\varepsilon$  eine positive reelle Zahl,  $\delta(\varepsilon)$  eine reelle Größe, die von  $\varepsilon$  abhängt, und  $r > 0$  ein positiver Exponent, so schreibt man:

1.  $\delta(\varepsilon) = o(\varepsilon^r)$ , falls

$$\frac{|\delta(\varepsilon)|}{\varepsilon^r} \xrightarrow{\varepsilon \rightarrow 0} 0,$$

d.h. falls  $|\delta(\varepsilon)|$  *schneller als*  $\varepsilon^r$  gegen 0 konvergiert.

2.  $\delta(\varepsilon) = O(\varepsilon)$ , falls  $|\delta(\varepsilon)|$  *nicht langsamer* als  $\varepsilon^r$  gegen 0 konvergiert, d.h. falls es Konstanten  $\varepsilon_* > 0$  und  $C_* > 0$  gibt, so dass

$$\frac{|\delta(\varepsilon)|}{\varepsilon^r} \leq C_*,$$

für alle  $\varepsilon$  mit  $0 < \varepsilon \leq \varepsilon_*$  (d.h. für alle hinreichend kleinen  $\varepsilon$ ) gilt.



**Beispiele**

1. Es gilt

$$\varepsilon^2 = o(\varepsilon), \quad \varepsilon^3 = o(\varepsilon), \quad \varepsilon^3 = o(\varepsilon^2)$$

und ganz allgemein  $\varepsilon^s = o(\varepsilon^r)$  für jedes  $s > r$ .

2. Es gilt  $C\varepsilon^r = O(\varepsilon^r)$  sowie  $\varepsilon^r + C\varepsilon^{r+1} = O(\varepsilon^r)$  für jedes  $c > 0$  und alle  $r > 0$ .

3. Für  $\delta(\varepsilon) = \exp(-1/\varepsilon)$  gilt  $\delta(\varepsilon) = o(\varepsilon^r)$  für jedes  $r > 0$ , da  $\delta(\varepsilon)$  exponentiell mit  $\varepsilon$  und damit schneller als jede Potenz von  $\varepsilon$  gegen 0 konvergiert.

4. Mit  $\delta(\varepsilon) = \varepsilon^r / \ln(1/\varepsilon)$  gilt  $\delta(\varepsilon) = o(\varepsilon^r)$  aber  $\varepsilon^r / \ln(\varepsilon) \neq O(\varepsilon^s)$  für jedes  $r \geq 0$  und alle  $s > r$ . Insbesondere gilt  $1/\ln(1/\varepsilon) = o(1)$ .

**Bemerkung** Die Landau-Symbole  $o$  und  $O$  sind extrem nützlich, aber es gelten einige auf den ersten Blick seltsam anmutende Regeln:

$$5 \cdot O(\varepsilon^2) = O(\varepsilon^2), \quad O(\varepsilon^2) + O(\varepsilon^3) = O(\varepsilon^2), \quad O(\varepsilon^2) \cdot O(\varepsilon^3) = O(\varepsilon^5).$$

Der Grund ist, dass  $o(\varepsilon^r)$  bzw.  $O(\varepsilon^r)$  keine konkrete Zahl oder Größe bezeichnet, sondern eigentlich eine Klasse von Zahlen (und die Notation  $\varepsilon^3 \in O(\varepsilon^2)$  eigentlich klarer und intuitiver wäre). Nach etwas Übung und Gewöhnung kann man sehr gut und effektiv mit Landau-Symbolen rechnen.

**Annahme** Wir betrachten nun wieder eine beliebige stetige Funktion  $f : [a, b] \rightarrow \mathbb{R}$ , wollen aber diesmal annehmen, dass sie  $n$ -mal differenzierbar im offenen Intervall  $(a, b)$  ist. Außerdem bezeichnet  $x_*$  einen beliebig fixierten Punkt in  $(a, b)$ .

**Taylor-Polynome** Das reelle Polynom

$$T_n(x) := \sum_{k=0}^n \frac{f^{(k)}}{k!} \cdot (x - x_*)^k$$

wird das Taylor-Polynom vom Grad  $n$  zu  $f$  im Entwicklungspunkt  $x_*$  genannt und ist für alle  $x \in \mathbb{R}$  definiert. Manchmal schreibt man auch  $T_{n,f,x_*}(x)$  oder  $T_n(x; x_*)$  um die Abhängigkeit von  $f$  und/oder  $x_*$  deutlich zu machen.

Per Definition gilt

$$\begin{aligned} T_0(x) &= f(x_*), \\ T_1(x) &= f(x_*) + f'(x_*) \cdot (x - x_*), \\ T_2(x) &= f(x_*) + f'(x_*) \cdot (x - x_*) + \frac{f''(x_*)}{2} \cdot (x - x_*)^2 \\ T_3(x) &= f(x_*) + f'(x_*) \cdot (x - x_*) + \frac{f''(x_*)}{2} \cdot (x - x_*)^2 + \frac{f'''(x_*)}{6} \cdot (x - x_*)^3 \end{aligned}$$

und durch direkte Rechnungen zeigt man

$$T_0^{(0)}(x_*) = f(x_*)$$

sowie

$$T_1^{(0)}(x_*) = f^{(0)}(x_*) = f(x_*), \quad T_1^{(1)}(x_*) = f^{(1)}(x_*) = f'(x_*)$$

und

$$T_2^{(0)}(x_*) = f(x_*), \quad T_2^{(1)}(x_*) = f'(x_*), \quad T_2^{(2)}(x_*) = f''(x_*).$$

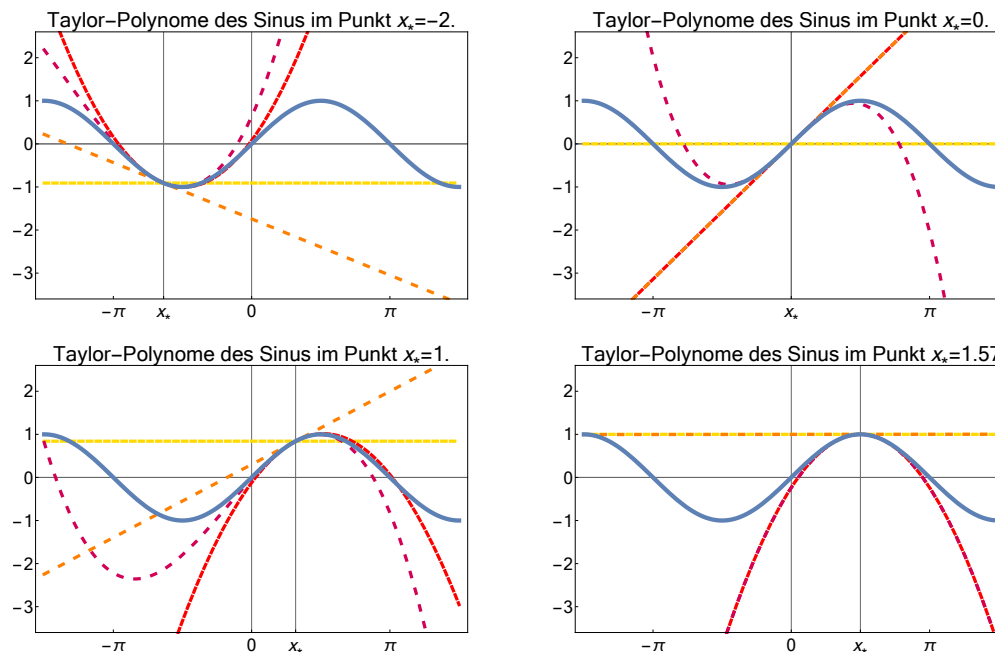
Ganz allgemein gilt

$$T_n^{(k)}(x_*) = f^{(k)}(x_*) \quad \text{für jedes } n \text{ und alle } k = 0, \dots, n.$$

Geometrisch bedeutet dies, dass sich der Graph von  $T_n$  im Punkt  $(x_*, f(x_*))$  bis zur  $n$ -ten Ordnung an den Graphen von  $f$  anschmiegt (siehe Bild). Der Term

$$R_n(x) = f(x) - T_n(x)$$

wird dabei das  $n$ -te Taylor-Restglied genannt.



Die Taylor-Polynome des nullten (gelb, konstant), ersten (orange, linear), zweiten (rot, quadratisch) und dritten (dunkelrot, kubisch) Grades in vier verschiedenen Entwicklungspunkten  $x_*$ . Beachte, dass im zweiten Bild ( $x_* = 0$ )

$T_2 = T_1$  wegen  $f''(x_*) = 0$  gilt und dass im vierten Bild ( $x_* = \pi/2$ ) aufgrund von  $f'(x_*) = f'''(x_*) = 0$  die

Taylor-Polynome  $T_0$  und  $T_1$  sowie  $T_2$  und  $T_3$  zusammenfallen.

## Beispiele

1. Für  $f = \sin$  erhalten wir

$$T_0(x) = \sin(x_*),$$

$$T_1(x) = \sin(x_*) + \cos(x_*)(x - x_*),$$

$$T_2(x) = \sin(x_*) + \cos(x_*)(x - x_*) - \frac{1}{2} \sin(x_*)(x - x_*)^2$$

$$T_3(x) = \sin(x_*) + \cos(x_*)(x - x_*) - \frac{1}{2} \sin(x_*)(x - x_*)^2 - \frac{1}{6} \cos(x_*)(x - x_*)^3$$

und mit  $x_* = 0$  ergibt sich

$$T_0(x) = 0, \quad T_1(x) = T_2(x) = x, \quad T_3(x) = T_4(x) = x - \frac{1}{6}x^3.$$

Für andere Entwicklungspunkte siehe das Bild.

2. Für das kubische Polynom  $f(x) = x^3 - x$  ergibt sich  $T_0(x) = x^3 - x_*$  sowie

$$T_1(x) = (x_*^3 - x_*) + (3x_*^2 - 1)(x - x_*) = (3x_*^2 - 1)x - 2x_*^3.$$

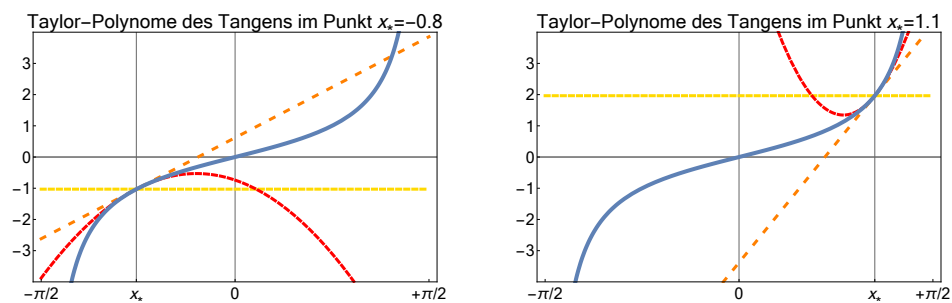
Außerdem berechnen wir

$$\begin{aligned} T_2(x) &= T_1(x) + \frac{1}{2}(6x_*) (x - x_*)^2 = T_1(x) + 3x_*x^2 - 6x_*^2x + 3x_*^3 \\ &= (3x_*)x^2 + (-3x_*^2 - 1)x + x_*^3 \end{aligned}$$

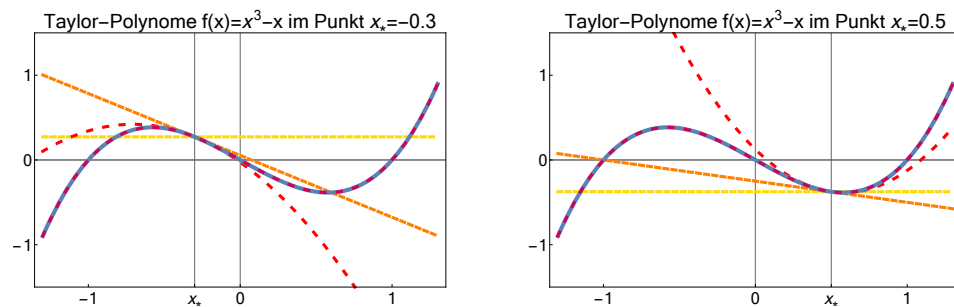
und

$$T_3(x) = T_2(x) + \frac{1}{6}6(x - x_*)^3 = T_2(x) + x^3 - 3x_*x^2 + 3x_*^2x - x_*^3 = x^3 - x.$$

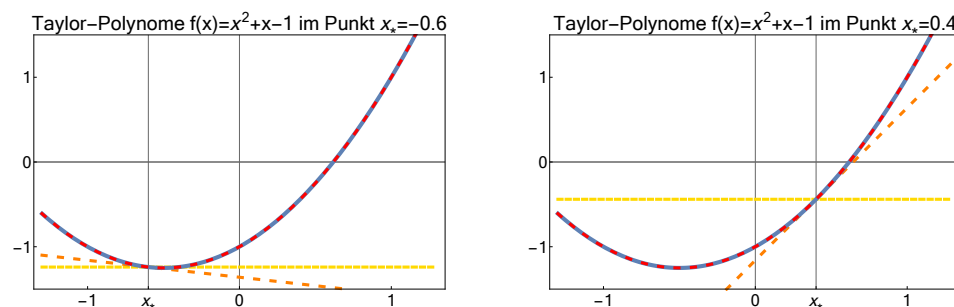
Insbesondere gilt hier  $T_3(x) = f(x)$  für jedes  $x_*$  und alle  $x$ .



Die ersten drei Taylor-Polynome des Tangens in zwei Entwicklungspunkten



Die Taylor-Polynome  $T_0, T_1, T_2, T_3$  für das Polynom  $f(x) = x^3 - x$  in zwei verschiedenen Entwicklungspunkten.



Die Taylor-Polynome  $T_0, T_1, T_2$  für das Polynom  $f(x) = x^2 + x - 1$  und zwei verschiedene Entwicklungspunkte.

**Lemma (Taylor-Polynome von Polynomen)** Ist  $f$  ein Polynom vom Grad  $n$ , so gilt  $T_{n+m}(x) = f(x)$  für alle  $x$  und alle  $m \in \mathbb{N}_0$ .

*Beweis:* Induktion über  $n$ . Alternativ kann die Behauptung auch direkt aus der Lagrange-Formel für das Taylor-Restglied (siehe unten) abgeleitet werden.  $\square$

**Theorem (Satz von Taylor)** Es gilt

$$R_n(x) = o(|x - x_*|^n) \quad \text{bzw.} \quad \lim_{x \rightarrow x_*} \left| \frac{R_n(x)}{(x - x_*)^n} \right| = 0.$$

*Beweis, Vorbemerkung:* Die Behauptung kann mit unserem bisherigen Wissen abgeleitet werden, siehe zum Beispiel [AORS, Seite 244f], aber der Beweis ist sehr technisch. Wir werden daher eine alternative Herleitung mittels Induktion über  $n$  vorstellen, die allerdings schon Integration sowie den Hauptsatz der Differential- und Integralrechnung

$$f(x) = f(x_*) + \int_{x_*}^x f'(y) \, dy$$

benutzt. Wir werden außerdem voraussetzen, dass  $f$  nicht nur  $n$ -mal differenzierbar ist, sondern dass sogar  $f^{(n+1)}$  existiert und stetig ist, und unter dieser Zusatzannahme die verbesserte Abschätzung

$$|R_n(x)| \leq C_n \cdot |x - x_*|^{n+1} = O(|x - x_*|^{n+1})$$

ableiten, wobei die Konstante  $C_n$  von  $f$  und allen ihren Ableitungen, aber nicht von  $x$  abhängt. Der Induktionsanfang  $n = 0$  ergibt sich dabei wegen

$$R_0(x) = f(x) - T_0(x) = f(x) - f(x_*) = \int_{x_*}^x f'(y) \, dy$$

aus der Abschätzung

$$|R_0(x)| = |f(x) - T_0(x)| \leq \int_{x_*}^x |f'(y)| \, dy \leq C_0 |x - x_*|,$$

wobei  $C_0 := \max_{y \in [a, b]} |f'(y)|$ .

*Beweis, Induktionsschritt  $n-1 \rightsquigarrow n$ :* Wir können die Induktionsvoraussetzung auf  $f'$  anwenden und schreiben

$$\begin{aligned} f'(x) &= \sum_{l=0}^{n-1} \frac{f^{(l+1)}(x_*)}{l!} (x - x_*)^l + S_{n-1}(x) \\ &= \sum_{k=1}^n \frac{f^{(k)}(x_*)}{(k-1)!} (x - x_*)^{k-1} + S_{n-1}(x). \end{aligned}$$

Hier ist  $S_{n-1}$  das Taylor-Restglied in der Entwicklung von  $f'$  und erfüllt  $|S_n(x)| \leq D_{n-1} |x - x_*|^n$  für eine geeignete Konstante  $D_{n-1}$  und alle  $x \in [a, b]$ . Der Hauptsatz der Differential- und Integralrechnung liefert

$$f(x) = f(x_*) + \int_{x_*}^x f'(y) \, dy$$

und mittels

$$R_n(x) := \int_{x_*}^x S_{n-1}(y) \, dy$$

ergibt sich via

$$\begin{aligned} f(x) &= f(x_*) + \int_{x_*}^x \sum_{k=1}^n \frac{f^{(k)}(x_*)}{(k-1)!} \cdot (y-x_*)^{k-1} \, dy + R_n(x) \\ &= f(x_*) + \sum_{k=1}^n \frac{f^{(k)}(x_*)}{(k-1)!} \cdot \int_{x_*}^x (y-x_*)^{k-1} \, dy + R_n(x) \\ &= f(x_*) + \sum_{k=1}^n \frac{f^{(k)}(x_*)}{(k-1)!} \cdot \left[ \frac{(y-x_*)^k}{k} \right]_{y=x_*}^{y=x} + R_n(x) \\ &= f(x_*) + \sum_{k=1}^n \frac{f^{(k)}(x_*)}{k!} \cdot (x-x_*)^k + R_n(x) \\ &= T_n(x) + R_n(x) \end{aligned}$$

die gewünschte Taylor-Formel. Für das Restglied gilt dabei die Abschätzung

$$|R_n(x)| \leq \int_{x_*}^x |S_{n-1}(y)| \, dy \leq D_{n-1} \int_{x_*}^x |y-x_*|^n \, dy = C_n |x-x_*|^{n+1},$$

wobei  $C_n := D_{n-1}/n$ . □

**Interpretation** Der Satz von Taylor ist ausgesprochen nützlich und wird sehr häufig in der Mathematik und den Ingenieurwissenschaften angewendet. Er garantiert, dass jede hinreichend oft differenzierbare Funktion  $f$  in der Nähe eines jeden Punktes  $x_*$  durch die Taylor-Polynome *approximiert* werden kann, wobei die Approximationsgüte mit aufsteigendem Grad immer besser wird, denn für  $|x-x_*| \ll 1$  gilt

$$|x-x_*|^n \ll |x-x_*|^{n-1} \ll \dots \ll |x-x_*|^3 \ll |x-x_*|^2 \ll |x-x_*|^1$$

Für große Abstände  $|x-x_*|$  kann man jedoch nicht erwarten, dass die Taylor-Approximation sinnvolle Näherungsformeln liefert. Das sieht man zum Beispiel sehr gut an der Sinus-Funktion: Diese ist eine beschränkte Funktion auf  $\mathbb{R}$ , aber alle Taylor-Polynome  $T_n$  sind für  $n > 0$  unbeschränkt.

**Restglied-Darstellungen** Es gibt verschiedene Möglichkeiten, das Restglied in der Taylor-Entwicklung genauer anzugeben, wobei wieder die Existenz und Stetigkeit der  $n+1$ -ten Ableitung von  $f$  vorausgesetzt sei:

1. Es gilt die Lagrange-Darstellung

$$R_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} (x-x_*)^{n+1}$$

für eine Zwischenstelle  $\xi$ , die zwischen  $x$  und  $x_*$  liegt (aber im allgemeinen nicht explizit berechnet werden kann). Diese Formel impliziert zum Beispiel, dass

$$C_n = \frac{\sup_{y \in [a, b]} |f^{(n+1)}(y)|}{(n+1)!}$$

die optimale Konstante in unserem Beweis des Satzes von Taylor ist.

2. Die Schlömilch-Darstellung ist

$$R_n(x) = \frac{f^{(n+1)}((1-\theta)x_* + \theta x)}{p \cdot (n+1)!} (x - x_*)^{n+1} (1-\theta)^{n+1-p},$$

wobei  $p \in \{1, 2, \dots, n+1\}$  beliebig ist und  $\theta \in [0, 1]$  in Abhängigkeit von  $x$ ,  $x_*$  und  $p$  zu wählen ist. Im Spezialfall  $p = 1$  wird diese Formel auch die Cauchy-Darstellung genannt.

3. Die Integraldarstellung

$$R_n(x) = \frac{1}{n} \int_{x_*}^x (x-y)^{n+1} f^{(n+1)}(y) dy$$

kann direkt aus dem Hauptsatz der Differential- und Integralrechnung abgeleitet werden und ist besonders für Beweise sehr gut geeignet.

**Bemerkung** Jede der oberen Darstellungsformeln impliziert

$$R_n(x) = O(|x - x_*|^{n+1}).$$

Beachte aber, dass der Satz von Taylor auch dann gilt, wenn nur die  $n$ -te Ableitung existiert. In diesem Fall gibt es aber keine einfache Darstellungsformel für das Restglied und man kann nur die schlechtere Abschätzung  $R_n(x) = o(|x - x_*|^n)$  beweisen.

**Alternativer Beweis von l'Hospital** Mit  $f(x_*) = g(x_*) = 0$  ergibt sich

$$\begin{aligned} \frac{f(x)}{g(x)} &= \frac{f(x_*) + f'(x_*)(x - x_*) + O(|x - x_*|)^2}{g(x_*) + g'(x_*)(x - x_*) + O(|x - x_*|)^2} \\ &= \frac{f'(x_*) + O(|x - x_*|)}{g'(x_*) + O(|x - x_*|)} \xrightarrow{x \rightarrow x_*} \frac{f'(x_*)}{g'(x_*)} \end{aligned}$$

aus Satzes von Taylor, wobei wir hier vorausgesetzt haben, dass  $f$  und  $g$  zweimal stetig differenzierbar sind mit  $g'(x_*) \neq 0$ .

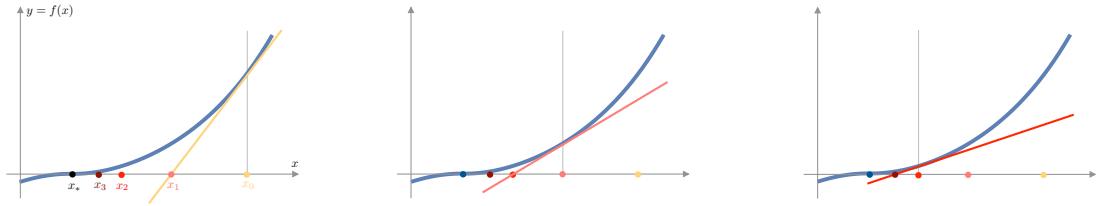
**Zur Interpretation und Konvergenz des Newton-Verfahrens** Das Newton-Verfahren zu approximativen Berechnung von Nullstellen ist durch die Rekursion

$$x_{n+1} = x_n - \frac{f'(x_n)}{f(x_n)}$$

gegeben, die auch als

$$f(x_n) + f'(x_n) \cdot (x_{n+1} - x_n) = T_{1, f, x_n}(x_{n+1}) = 0$$

geschrieben werden kann. Oder anders gesagt: Man ersetzt im  $n+1$ -ten Schritt  $f$  durch das lineare Taylor-Polynom an der Stützstelle  $x_n$  und bestimmt  $x_{n+1}$  als die entsprechende Nullstelle des Taylor-Polynoms.



Bei Newton-Verfahren wird im Schritt  $x_n \rightsquigarrow x_{n+1}$  die Funktion  $f$  im Punkt  $x_n$  linearisiert — d.h. durch das entsprechende Taylor-Polynom 1. Grades im Entwicklungspunkt  $x_n$  ersetzt — und  $x_{n+1}$  wird als Nullstelle der linearen Approximation bestimmt.

Um die Konvergenz in der Nähe einer exakten Nullstelle  $x_*$  zu berechnen, schreiben wir

$$0 = f(x_*) = f(x_n) + f'(x_n) \cdot (x_* - x_n) + \frac{1}{2} \cdot f''(\xi_n) \cdot (x_* - x_n)^2,$$

wobei  $\xi_n$  die Lagrange-Zwischenstelle ist und zwischen  $x_*$  und  $x_n$  liegt. Durch Einsetzen erhalten wir

$$(x_{n+1} - x_*) = \frac{f''(\xi_n)}{2 \cdot f'(x_n)} \cdot (x_n - x_*)^2.$$

Ist nun der Startwert  $x_0$  hinreichend nahe bei  $x_*$ , so gilt

$$|x_{n+1} - x_*| \approx C_* \cdot (x_n - x_*)^2 \quad \text{mit} \quad C_* = \left| \frac{f''(x_*)}{2 \cdot f'(x_*)} \right|,$$

und damit auch  $|x_{n+1} - x_*| \leq |x_n - x_*|$ . Insbesondere konvergiert die Newton-Folge  $(x_n)_{n \in \mathbb{N}}$  sehr schnell (genauer gesagt: quadratisch) gegen die exakte Nullstelle  $x_*$ . Allerdings gilt dies nur, wenn  $x_0$  schon hinreichend nahe bei  $x_*$  liegt.

Vorlesung 28, 10. Januar 2020

**Über Taylor-Reihen** Besitzt  $f$  unendlich viele Ableitungen, so kann man formal die Taylor-Reihe

$$T_\infty(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_*)}{k!} (x - x_*)^k$$

definieren und die folgende zwei Fragen untersuchen:

1. Für welche  $x \in \mathbb{R}$  ist  $T_\infty(x)$  im Sinne einer konvergenten oder gar absolut konvergenten Reihe wohldefiniert?
2. Für welche  $x \in [a, b]$  gilt  $f(x) = T_\infty(x)$ ?

Wir werden diese Frage am Ende dieses Kapitels bzw. in *Mathe II* genauer studieren, wollen aber hier schon festhalten, dass für allgemeine Funktionen  $f$  diese Fragen nicht ganz einfach zu beantworten sind.

**Beispiel** Wir betrachten  $f(x) = \exp(x)$  auf ganz  $\mathbb{R}$  (bzw. auf einem beliebig gewählten Intervall  $[a, b]$ ). Dann gilt (beachte, dass  $\exp^{(k)} = \exp$ )

$$T_n(x) = \sum_{k=0}^n \frac{\exp(x_*)}{k!} (x - x_*)^k \xrightarrow{n \rightarrow \infty} T_\infty(x) = \sum_{k=0}^{\infty} \frac{\exp(x_*)}{k!} (x - x_*)^k,$$

für jeden Entwicklungspunkt  $x_*$  und jedes  $x \in \mathbb{R}$ , denn das Quotientenkriterium garantiert via

$$\frac{\frac{\exp(x_*)}{(k+1)!} (x - x_*)^{k+1}}{\frac{\exp(x_*)}{k!} (x - x_*)^k} = \frac{x - x_*}{k+1} \xrightarrow{k \rightarrow \infty} 0$$

dass die Partialsummenfolge  $(T_n(x))_{n \in \mathbb{N}}$  absolut konvergiert. Außerdem verifizieren wir

$$T_\infty(x) = \exp(x_*) \sum_{k=0}^{\infty} \frac{(x - x_*)^k}{k!} = \exp(x_*) \exp(x - x_*) = \exp(x) = f(x)$$

für alle  $x \in \mathbb{R}$ . Man sagt auch, die Exponentialfunktion ist reell-analytisch, eben weil  $f = T_\infty$  für jede Entwicklungsstelle  $x_* \in \mathbb{R}$  gilt.

**Gegenbeispiel** Wir betrachten

$$f(x) = \begin{cases} \exp(1/x) & \text{für } x < 0, \\ 0 & \text{für } x = 0, \\ \exp(-1/x) & \text{für } x > 0. \end{cases}$$

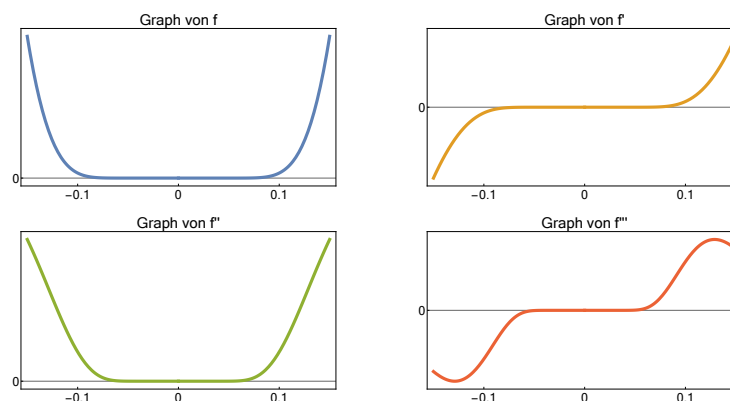
Man kann zeigen (\*\*-Übungsaufgabe), dass  $f$  an jeder Stelle  $x_*$  unendlich oft differenzierbar ist, wobei

$$f^{(k)}(0) = 0$$

für alle  $k \in \mathbb{N}$  gilt (es gilt aber  $f^{(k)}(x) \neq 0$  für alle  $x \neq 0$  und alle  $k \in \mathbb{N}$ ). Im Entwicklungspunkt  $x_* = 0$  erhalten wir deshalb

$$T_n(x) = 0$$

für alle  $n \in \mathbb{N}$  und jedes  $x \in \mathbb{R}$  und damit auch  $T_\infty(x) = 0$ . Insbesondere existiert die entsprechende Taylor-Reihe, aber es gilt offensichtlich  $T_\infty \neq f$ . Der Grund ist, dass  $\exp(-1/x)$  für kleine  $x$  sehr viel kleiner als jede Potenz von  $x$  ist, d.h.  $\exp(-1/x) = o(x^r)$  für alle Exponenten  $r > 0$ .



Die ersten Ableitungen der Funktion aus dem oben eingeführten Gegenbeispiel. Beachte, dass alle Ableitungen an der Stelle  $x_* = 0$  verschwinden und daher die Taylor-Reihe trivial ist.



## 9.3 Ableitungen in der Kurvendiskussion

**Hinreichende Kriterien für lokale Extrema** Ist  $f$  dreimal stetig differenzierbar, so gilt in der Nähe eines kritischen Punktes

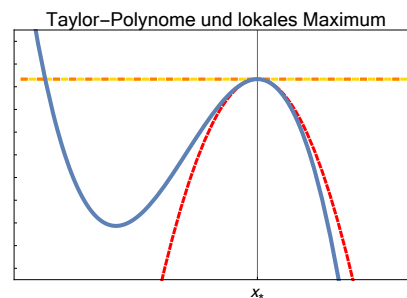
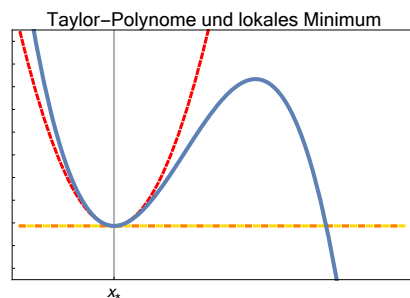
$$f(x) = f(x_*) + \frac{1}{2} f''(x_*) (x - x_*)^2 + O(|x - x_*|^3),$$

wegen  $f'(x_*) = 0$  und nach dem Satz von Taylor. Wir schließen nun, dass das Vorzeichen von  $f''(x_*)$  darüber entscheidet, ob es sich um ein lokales Minimum oder Maximum handelt.

**Theorem (Zweite Ableitung und lokale Extrema)** Ist  $f$  zweimal stetig differenzierbar, so gelten für jeden inneren Punkt  $x_* \in (a, b)$  die folgenden beiden Implikationen

1. Gilt  $f'(x_*) = 0$  und  $f''(x_*) > 0$ , so nimmt  $f$  in  $x_*$  ein striktes lokales Minimum an.
2. Gilt  $f'(x_*) = 0$  und  $f''(x_*) < 0$ , so nimmt  $f$  in  $x_*$  ein striktes lokales Maximum an.

*Beweis:* Ein rigoroser Beweis findet sich in [AORS, Seite 250] und benötigt nur zweite Ableitungen von  $f$ . Unter der leicht stärkeren Voraussetzung, dass  $f'''$  existiert und stetig ist, kann die Behauptung aber direkt aus der obigen Taylor-Formel abgelesen werden, denn für  $|x - x_*| \ll 1$  dominiert der zweite Taylor-Term  $\frac{1}{2} f''(x_*) (x - x_*)^2$  den Fehlerterm  $O(|x - x_*|^3)$ .



Gilt  $f'(x_*) = 0$  sowie  $f''(x_*) > 0$  bzw.  $f''(x_*) < 0$ , so nimmt  $f$  in  $x_*$  ein striktes Minimum bzw. ein striktes Maximum an. Gilt jedoch  $f'(x_*) = 0$  sowie  $f''(x_*) = 0$ , so kann  $x_*$  nicht so einfach klassifiziert werden.

**Bemerkung** Gilt  $f'(x_*) = 0$  sowie  $f''(x_*) = 0$ , so ist keine einfache Aussage über das Verhalten von  $f$  in der Nähe von  $x_*$  möglich (bzw. erfordert die Auswertung von höheren Ableitungen). Einfache Beispiele sind

$$f_1(x) = x^3, \quad f_2(x) = +x^4, \quad f_3(x) = -x^4.$$

Insbesondere gilt stets  $f'_j(0) = 0 = f''(0)$  im Entwicklungspunkt  $x_* = 0$ , aber  $f_2$  bzw.  $f_3$  nimmt in 0 ein striktes Minimum bzw. Maximum an, wohingegen 0 ein Sattelpunkt für  $f_1$  ist.

**Konvexe und konkave Funktionen** Die Funktion  $f$  heißt konvex bzw. konkav, wenn

$$f((1 - \lambda)x + \lambda\tilde{x}) \leq (1 - \lambda)f(x) + \lambda f(\tilde{x})$$

bzw.

$$f((1 - \lambda)x + \lambda\tilde{x}) \geq (1 - \lambda)f(x) + \lambda f(\tilde{x})$$

für alle  $x, \tilde{x} \in [a, b]$  mit  $x \neq \tilde{x}$  und alle  $0 < \lambda < 1$  gilt. Gilt sogar immer  $<$  bzw.  $>$  statt  $\leq$  bzw.  $\geq$ , so nennt man  $f$  strikt konvex bzw. strikt konkav. Beachte, dass diese Definitionen keine Ableitungen von  $f$  beinhalten und dass es eine direkt geometrische Interpretation dieser Ungleichungen gibt (siehe Bild).

**Lemma (Erste Ableitung und Konvexität/Konkavität)** Für differenzierbare Funktionen  $f$  gilt:

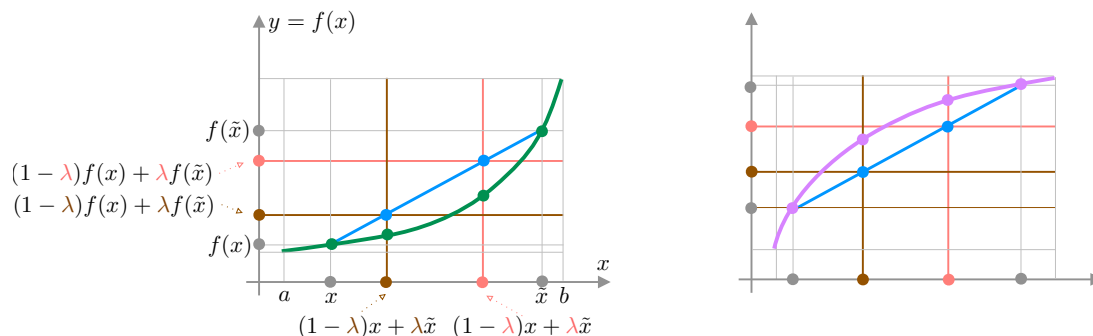
1.  $f$  ist genau dann konvex bzw. konkav, wenn  $f'$  monoton wachsend bzw. monoton fallend ist.
2. Ist  $f'$  strikt monoton wachsend bzw. strikt monoton fallend, so ist  $f$  strikt konvex bzw. strikt konkav.

*Beweis:* Wir wollen hier auf einen rigorosen Beweis verzichten, da alle Behauptungen auch aus der geometrischen Anschauung abgeleitet werden können.  $\square$

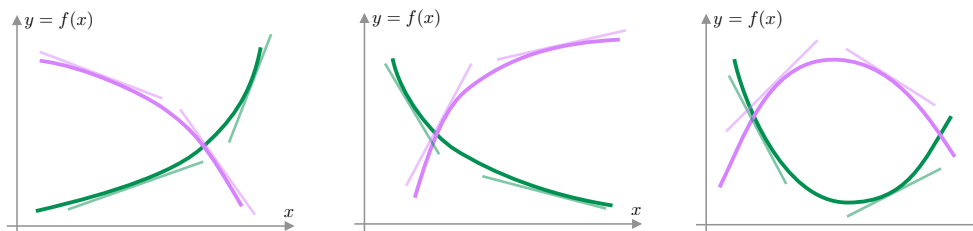
**Lemma (Zweite Ableitung und Konvexität/Konkavität)** Ist  $f$  zweimal differenzierbar, so gilt:

1.  $f$  ist genau dann konvex bzw. konkav, wenn  $f''(x) \geq 0$  bzw.  $f''(x) \leq 0$  für alle  $x \in [a, b]$  gilt.
2. Gilt  $f''(x) > 0$  bzw.  $f''(x) < 0$  für alle  $x \in [a, b]$ , so ist  $f$  strikt konvex bzw. strikt konkav.

*Beweis:* Wir hatten uns schon überlegt, dass die Monotonie einer Funktion äquivalent zu gewissen Ungleichungsbedingungen an die Ableitung ist, und wir können diese Ergebnisse nun auf die erste Ableitung  $f'$  anwenden.  $\square$



Bei einer strikt konvexen (links, grün) bzw. konkaven (rechts, lila) Funktion liegt der Graph unterhalb bzw. oberhalb jeder Sekante (blau) an den Graphen. Affine Funktionen (d.h. Polynome von Grad 0 oder 1) sind gleichzeitig konvex und konkav (aber jeweils nicht strikt), da dann Graph und Sekante zusammenfallen.



Typische Beispiele für strikt konvexe (grün) bzw. strikt konkave (lila) Funktionen auf kompakten Intervallen. Die dünnen geraden Linien repräsentieren Tangenten an den Graphen von  $f$ , deren Anstiege gerade durch die monoton wachsende bzw. fallende Funktion  $f'$  beschrieben werden.

**Lemma (Konvexität/Konkavität und Extrema)** Ist  $f$  konvex bzw. konkav auf  $[a, b]$  und  $x_* \in (a, b)$  ein kritischer Punkt, so nimmt  $f$  in  $x_*$  ein globales Minimum bzw. globales Maximum an. Ist  $f$  darüber hinaus sogar strikt konvex bzw. konkav, so existiert höchstens ein kritischer Punkt.

Beweis, Teil 1: Die Taylor-Formel mit Lagrange-Restglied impliziert

$$f(x) = f(x_*) + f'(x_*) (x - x_*) + \frac{1}{2} f''(\xi) (x - x_*)^2$$

für eine Zwischenstelle  $\xi$ . Mit  $f'(x_*) = 0$  sowie  $f''(\xi) \geq 0$  bzw.  $f''(\xi) \leq 0$  erhalten wir  $f(x) \geq f(x_*)$  bzw.  $f(x) \leq f(x_*)$  für alle  $x$  und damit  $f(x_*) = \min f$  bzw.  $f(x_*) = \max f$ .

Beweis, Teil 2: Sei  $f$  strikt konvex und nehmen wir an, es gäbe zwei verschiedene kritische Punkte  $x_*, \tilde{x}_* \in (a, b)$ . Dann gilt nach dem ersten Teil  $f(x_*) = f(\tilde{x}_*) = \min f$  und die Definition von Konvexität liefert via

$$\min f \leq f\left(\frac{1}{2}x_* + \frac{1}{2}\tilde{x}_*\right) < \frac{1}{2}f(x_*) + \frac{1}{2}f(\tilde{x}_*) = \min f$$

einen Widerspruch. Für konkave Funktionen kann analog argumentiert werden.  $\square$

**Hinreichende Kriterien für Wendepunkte** Viele Funktionen sind weder konvex noch konkav auf dem Intervall  $[a, b]$ , sondern Konvexitäts- und Konkavitätsintervalle wechseln sich ab. An einer Übergangsstelle  $x_*$  muss dann  $f''(x_*) = 0$  gelten und damit auch

$$\begin{aligned} f(x) &= f(x_*) + f'(x_*) (x - x_*) + \frac{1}{6} f'''(\xi) (x - x_*)^3 \\ &= f(x_*) + f'(x_*) (x - x_*) + \frac{1}{6} f'''(x_*) (x - x_*)^3 + O(|x - x_*|^4), \end{aligned}$$

wobei  $\xi$  wieder eine Lagrange-Zwischenstelle bezeichnet. Außerdem kann der Satz von Taylor auch auf  $f''$  angewendet werden und liefert

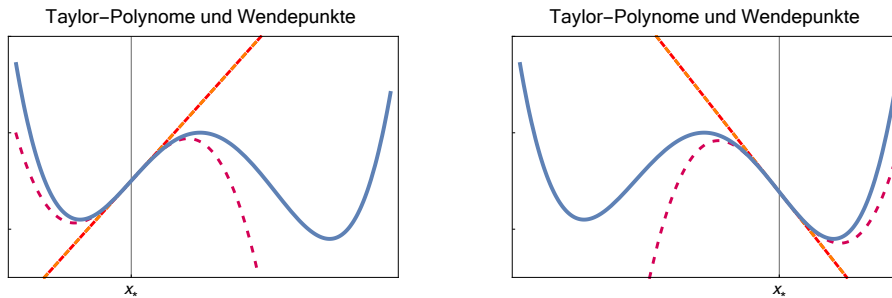
$$f''(x) = 0 + f'''(x_*) (x - x_*) + O(|x - x_*|^2),$$

Aus jeder dieser Formeln kann das folgende Resultat abgeleitet werden.

**Theorem (Dritte Ableitung und Wendepunkte)** Ist  $f$  dreimal stetig differenzierbar, so gelten für jeden inneren Punkt  $x_* \in (a, b)$  die folgenden beiden Implikationen

1. Gilt  $f''(x_*) = 0$  und  $f'''(x_*) < 0$ , so ist  $f$  in der Nähe von  $x_*$  konvex-konkav.

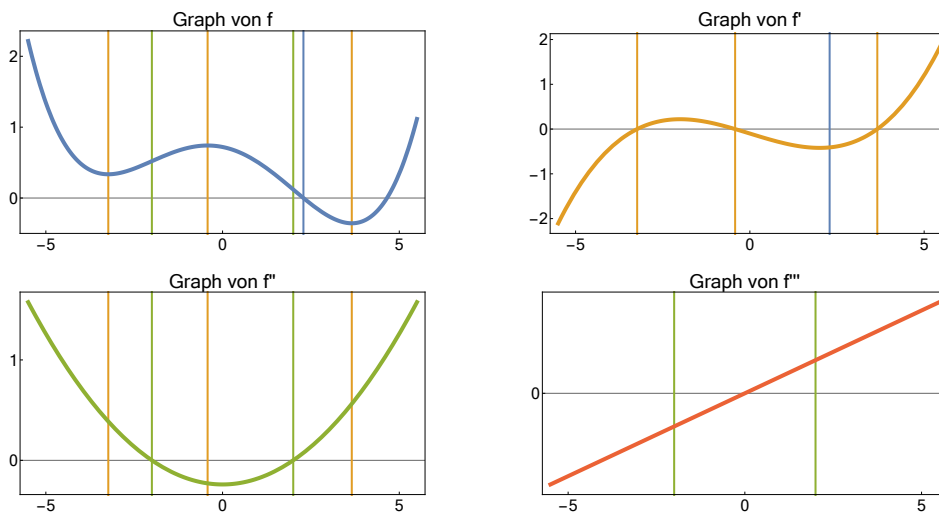
2. Gilt  $f''(x_*) = 0$  und  $f'''(x_*) > 0$ , so ist  $f$  in der Nähe von  $x_*$  konkav-konvex.



Die Taylor-Polynome  $T_1 = T_2$  und  $T_3$  in Entwicklungspunkten  $x_*$ , die zu Wendepunkten von  $f$  gehören (d.h. es gilt  $f''(x_*) = 0 \neq f'''(x_*)$ ). Das Vorzeichen von  $f'''(x_*)$  gibt dabei an, ob die Funktion von konvex zu konkav oder umgekehrt wechselt.

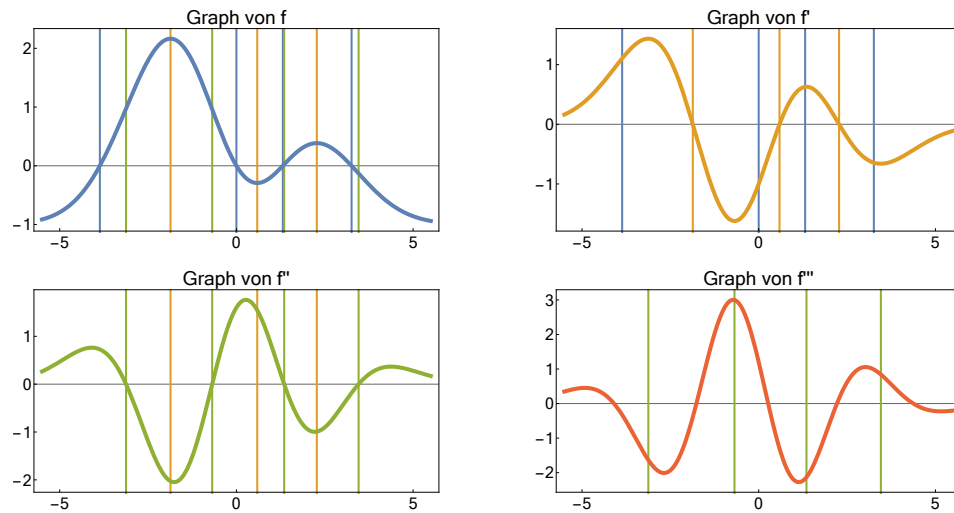
**Zur Kurvendiskussion** Um das qualitative Verhalten einer gegebenen Funktion zu studieren, sollte man immer versuchen, die Nullstellen der nullten, ersten, und zweiten Ableitung zu bestimmen:

1. Punkte  $x_*$  mit  $f(x_*) = 0$  sind die Nullstellen von  $f$ . Ist dabei die Nicht-Entartungsbedingung  $f'(x_*) \neq 0$  erfüllt, so handelt es sich um einen Schnittpunkt des Graphen mit der  $x$ -Achse. Andernfalls kann es sich um einen (entarteten) Schnittpunkt oder einen echten Berührungspunkt handeln, aber um dies zu entscheiden, muss man höhere Ableitungen auswerten.
2. Punkte  $x_*$  mit  $f'(x_*) = 0$  entsprechen den lokalen Extrema, zumindest sofern die Nicht-Entartungsbedingung  $f''(x_*) \neq 0$  gilt.
3. Punkte  $x_*$  mit  $f''(x_*) = 0$  liefern unter der Nicht-Entartungsvoraussetzung  $f'''(x_*) \neq 0$  die Wendepunkte.



Erstes Beispiel zur Bedeutung der Nullstellen von  $f$ ,  $f'$  und  $f''$ . Beachte, dass zu  $f^{(m)}(x_*) = 0$  die Nicht-Entartungsbedingung  $f^{(m+1)}(x_*) \neq 0$  gehört.

Außerdem ist es meist sinnvoll, auch das Verhalten der Funktion in den Randpunkten und isolierten Lücken des Definitionsbereiches sowie das asymptotische Verhalten für  $x \rightarrow \pm\infty$  genauer zu untersuchen. Im praktischen Leben kann man natürlich eine gegebene Funktion mit Hilfe eines geeigneten Computerprogramms plotten und dadurch viele wertvolle Informationen ohne aufwändige Rechnungen erhalten.

Zweites Beispiel zur Bedeutung der Nullstellen von  $f$ ,  $f'$  und  $f''$ .

Vorlesung 29, 13. Januar 2020

## 9.4 Exkurs über Potenzreihen

**Potenzreihen** Für eine gegebene Koeffizientenfolge  $(\beta_k)_{k \in \mathbb{N}_0} \subset \mathbb{C}$  heißt

$$\phi(z) = \sum_{k=0}^{\infty} \beta_k \cdot |z - z_*|^k$$

die entsprechende komplexe Potenzreihe im Entwicklungspunkt  $z_* \in \mathbb{C}$  und in der Variablen  $z \in \mathbb{C}$ . Eine reelle Potenzreihe ist analog durch

$$f(x) = \sum_{k=0}^{\infty} \alpha_k \cdot |x - x_*|^k$$

gegeben, wobei nun  $x, x_* \in \mathbb{R}$  sowie  $(\alpha_k)_{k \in \mathbb{N}_0} \subset \mathbb{R}$ . Wir wollen nun für eine reelle Potenzreihe die Fragen

1. Für welche  $x$  ist  $f(x)$  im Sinne einer konvergenten Reihe wohldefiniert?
2. Ist  $f$  differenzierbar?
3. Wie sehen die Taylor-Polynome aus?
4. Was kann man über die Taylor-Reihe von  $f$  sagen?

untersuchen, beginnen aber mit einem Resultat über komplexe Potenzreihen, das wir aber erst in der Vorlesung *Mathe III* beweisen werden.

**Theorem (Hauptsatz über die Konvergenz komplexer Potenzreihen)** Für jede Potenzreihe — d.h. für jede gegebene Koeffizientenfolge  $(\beta_k)_{k \in \mathbb{N}}$  und jeden gegebenen Entwicklungspunkt  $z_*$  — gelten die folgenden Aussagen:

1. Es existiert ein Konvergenzradius  $r$  mit  $0 \leq r \leq \infty$ , so dass:
  - (a) Für jedes  $z \in \mathbb{C}$  mit  $|z - z_*| < r$  konvergiert die Reihe zu  $\phi(z)$  absolut.

(b) Für jedes  $z \in \mathbb{C}$  mit  $|z - z_*| > r$  ist  $\phi(z)$  nicht definiert, da die entsprechende Reihe divergiert.

2. Es gilt stets die Formel von Cauchy-Hadamard

$$r = \frac{1}{\limsup_{k \rightarrow \infty} \sqrt[k]{|\beta_k|}},$$

wobei  $\limsup_{k \rightarrow \infty} \sqrt[k]{|\beta_k|}$  der größte Häufungspunkt der Folge  $(\sqrt[k]{|\beta_k|})_{k \in \mathbb{N}_0}$  ist und  $1/0 = \infty$  und  $1/\infty = 0$  vereinbart sei. Insbesondere gilt

$$r = \lim_{k \rightarrow \infty} \left| \frac{\beta_k}{\beta_{k+1}} \right| \quad \text{bzw.} \quad r = \lim_{k \rightarrow \infty} \frac{1}{\sqrt[k]{|\beta_k|}},$$

sofern der jeweilige Grenzwert existiert.

3. Es gilt die lokal gleichmäßige Konvergenz

$$\sup_{z: |z - z_*| \leq \tilde{r}} \left| \sum_{k=0}^{\infty} \dots - \sum_{k=0}^n \dots \right| = \sup_{z: |z - z_*| \leq \tilde{r}} \left| \sum_{k=n+1}^{\infty} \beta_k \cdot (z - z_*)^k \right| \xrightarrow{n \rightarrow \infty} 0$$

für jeden Radius  $\tilde{r}$  mit  $0 < \tilde{r} < r$ .

### Bemerkungen

1. Die Cauchy-Hadamard-Formel liefert immer einen eindeutigen Wert von  $r$ , auch wenn die Grenzwertformeln nicht verwendet werden dürfen.
2. Der Entartungsfall  $r = 0$  kann eintreten, zum Beispiel für  $\beta(k) = k!$ .
3. Im Fall von  $r = \infty$  ist die Potenzreihe auf ganz  $\mathbb{C}$  definiert.
4. Der Hauptsatz macht keine Aussage über die Konvergenz der Reihe für  $z \in \mathbb{C}$  mit  $|z - z_*| = r$ , d.h. auf dem Rand der Kreisscheibe vom Radius  $r$  um  $z_*$ .
5. Die Cauchy-Hadamard-Formel impliziert (Nachrechnen!), dass die Potenzreihen

$$\sum_{k=0}^{\infty} \beta_k \cdot (z - z_*)^k \quad \text{und} \quad \sum_{k=1}^{\infty} k \cdot \beta_k \cdot (z - z_*)^k$$

denselben Konvergenzradius besitzen, wobei die zweite aus der ersten durch gliedweise Differentiation hervorgeht.

Ergänzung

**Achtung** Das Quotienten- bzw. Wurzelkriterium für Potenzreihen ergibt sich aus dem entsprechenden Kriterium aus dem Abschnitt über Reihen von Zahlen, aber die Details in den Formeln sind leicht anders. So garantiert zum Beispiel

$$\lim_{k \rightarrow \infty} \frac{|\beta_{k+1}| |z - z_*|^{k+1}}{|\beta_k| |z - z_*|^k} < 1$$

die absolute Konvergenz von  $\phi(z)$  für fixiertes  $z \in \mathbb{C}$ , und diese Bedingung kann als

$$|z - z_*| < \frac{1}{\lim_{k \rightarrow \infty} \frac{|\beta_{k+1}|}{|\beta_k|}} = \lim_{k \rightarrow \infty} \frac{|\beta_k|}{|\beta_{k+1}|}$$

geschrieben werden (sofern die Grenzwerte existieren).

**Beispiele**

1. Für  $\beta_k = 1/k!$  und  $z_* = 0$  gilt

$$\phi(z) = \exp(z) = \sum_{k=0}^{\infty} \frac{z^k}{k!} = 1 + z + \frac{1}{2} \cdot z^2 + \frac{1}{6} \cdot z^3 + \frac{1}{24} \cdot z^4 + \dots$$

und

$$r = \lim_{k \rightarrow \infty} \frac{(k+1)!}{k!} = \lim_{k \rightarrow \infty} k = \infty$$

ist der entsprechende Konvergenzradius.

2. Für die beiden trigonometrischen Funktionen

$$\sin(z) = \sum_{m=0}^{\infty} (-1)^{2m} \cdot \frac{z^{2m+1}}{(2m+1)!} = z - \frac{1}{6} \cdot z^3 + \frac{1}{120} \cdot z^5 - \frac{1}{5040} \cdot z^7 + \dots$$

bzw.

$$\cos(z) = \sum_{m=0}^{\infty} (-1)^{2m} \cdot \frac{z^{2m}}{(2m)!} = 1 - \frac{1}{2} \cdot z^2 + \frac{1}{24} \cdot z^4 - \frac{1}{720} \cdot z^6 + \dots$$

ist der Konvergenzradius ebenfalls  $\infty$ .

3. Für  $\beta_k = (-1)^k/k + 1$  ergibt sich

$$r = \lim_{k \rightarrow \infty} \frac{k+1}{k+2} = 1$$

und wir erhalten  $r = 1$  als Konvergenzradius von  $\phi(z) = 1 - \frac{1}{2} \cdot z + \frac{1}{3} \cdot z^3 - \frac{1}{4} \cdot z^4 + \dots$

4. Mit  $\beta_k = 1$  und  $z_* = 0$  erhalten wir  $r = 1$  als Konvergenzradius der geometrischen Reihe und es gilt

$$\phi(z) = \sum_{k=0}^{\infty} z^k = 1 + z + z^2 + z^3 + z^4 + \dots = \frac{1}{1-z}$$

für alle  $|z| < 1$ .

**Reelle Potenzreihen (zum Beispiel reelle Taylor-Reihen)** Der komplexe Hauptsatz gilt mit

$$\beta_k = \alpha_k, \quad z = x, \quad z_* = x_*$$

auch für eine reelle Potenzreihe und impliziert, dass  $f$  mit

$$f(x) = \alpha_0 + \alpha_1 \cdot (x - x_*) + \alpha_2 \cdot (x - x_*)^2 + \alpha_3 \cdot (x - x_*)^3 + \alpha_4 \cdot (x - x_*)^4 + \dots$$

auf dem offenen Intervall  $(x_* - r, x_* + r)$  im Sinne absolut konvergenter Reihen wohldefiniert ist.

**Theorem (Hauptsatz über reelle Potenzreihen)** Eine reelle Potenzreihe ist in ihrem reellen Definitionsbereich unendlich oft differenzierbar, wobei alle Ableitungen gliedweise berechnet werden können. Insbesondere gilt

$$\begin{aligned} f'(x) &= \alpha_1 + 2 \cdot \alpha_2 \cdot (x - x_*) + 3 \cdot \alpha_3 \cdot (x - x_*)^2 + 4 \cdot \alpha_4 \cdot (x - x_*)^3 + \dots \\ &= \sum_{k=1}^{\infty} k \cdot \alpha_k \cdot (x - x_*)^{k-1} = \sum_{l=0}^{\infty} (l+1) \cdot \alpha_{l+1} \cdot (x - x_*)^l \end{aligned}$$

sowie

$$\begin{aligned} f''(x) &= 2 \cdot \alpha_2 + 2 \cdot 3 \cdot \alpha_3 \cdot (x - x_*)^1 + 3 \cdot 4 \cdot \alpha_4 \cdot (x - x_*)^2 + \dots \\ &= \sum_{k=2}^{\infty} (k-1) \cdot k \cdot \alpha_k \cdot (x - x_*)^{k-2} \end{aligned}$$

und ganz allgemein

$$f^{(n)}(x) = \sum_{k=n}^{\infty} (k-n+1) \cdot \dots \cdot (k-1) \cdot k \cdot \alpha_k \cdot (x - x_*)^{k-n}$$

Beweis: Wir verweisen wieder auf *Mathe III*, bemerken aber, dass wegen

$$\limsup_{l \rightarrow \infty} \sqrt[l]{(l+1) \cdot |\alpha_{l+1}|} = \limsup_{l \rightarrow \infty} \sqrt[l]{l+1} \cdot \limsup_{l \rightarrow \infty} \sqrt[l]{|\alpha_{l+1}|} = 1 \cdot \limsup_{k \rightarrow \infty} \sqrt[k]{|\alpha_k|}$$

die Konvergenzradien von  $f$  und  $f'$  gleich sind. Induktiv folgt dann, dass die Konvergenzradien aller Ableitungen übereinstimmen.  $\square$

**Folgerung** Für eine reelle Potenzreihe gilt

$$f^{(n)}(x_*) = (n-n+1) \cdot \dots \cdot (n-1) \cdot n \cdot \alpha_n = n! \cdot \alpha_n$$

und damit

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(x_*)}{k!} \cdot (x - x_*)^k$$

für alle  $x$  mit  $|x - x_*| < r$ . Insbesondere ist jede reelle Potenzreihe ihre eigene Taylor-Reihe.

### Beispiele

1. Für alle  $|x| < 1$  gilt

$$\frac{1}{1-x} = \sum_{k=0}^{\infty} x^k = 1 + x + x^2 + x^3 + \dots$$

und durch Differenziation nach  $x$  erhalten wir

$$\frac{1}{(1-x)^2} = \sum_{k=1}^{\infty} k \cdot x^{k-1} = 1 + 2 \cdot x + 3 \cdot x^2 + 4 \cdot x^3 + \dots$$



sowie

$$\frac{2}{(1-x)^3} = \sum_{k=2}^{\infty} k \cdot (k-1) \cdot x^{k-2} = 2 + 6 \cdot x + 12 \cdot x^2 + 20 \cdot x^3 \dots,$$

wobei die rechte Seite jeweils die Taylor-Reihe der linken Seite ist und für  $|x| < 1$  bzw.  $|x| > 1$  konvergiert bzw. divergiert. Diese Formeln gelten sinngemäß für  $z \in \mathbb{C}$  mit  $|z| < 1$ , aber dies werden wir erst in *Mathe III* richtig verstehen.

2. Für alle  $|x| < 1$  gilt außerdem die Taylor-Reihendarstellung

$$\ln(1+x) = \sum_{k=0}^{\infty} (-1)^k \cdot \frac{x^{k+1}}{k+1} = x - \frac{1}{2} \cdot x^2 + \frac{1}{3} \cdot x^3 - \frac{1}{4} \cdot x^4 + \dots,$$

wobei der entsprechende Konvergenzradius gerade 1 ist und der Grenzübergang  $x \nearrow 1$  die schon bekannte Formel

$$\ln 2 = 1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \dots$$

für die alternierende harmonische Reihe liefert. Beachte außerdem, dass wir nach Differentiation bzgl.  $x$  die Formeln

$$\frac{1}{1+x} = 1 - x + x^2 - x^3 + \dots = \sum_{k=0}^{\infty} (-1)^k \cdot x^k$$

und damit einen Spezialfall der geometrischen Reihe erhalten.

3. Wegen  $\arctan 0 = 0$  und

$$\frac{d}{dx} \arctan(x) = \frac{1}{1+x^2} = \sum_{k=0}^{\infty} (-1)^k \cdot x^{2k} = 1 - x^2 + x^4 - x^6 + \dots$$

gilt auch

$$\arctan(x) = x - \frac{1}{3} \cdot x^3 + \frac{1}{5} \cdot x^5 - \frac{1}{7} \cdot x^7 + \dots = \sum_{k=0}^{\infty} (-1)^k \cdot \frac{x^{2k+1}}{2k+1}$$

für alle  $|x| < 1$ . Insbesondere gilt  $\arctan^{(n)}(0) = 0$  für alle geraden  $n$  sowie  $\arctan^{(1)}(0) = 1! = 1$ ,  $\arctan^{(3)}(0) = -2! = -2$  und  $\arctan^{(5)}(0) = 4! = 24$  usw. (allgemein gilt  $\arctan^{(2k+1)}(0) = (2k+1)! \cdot (-1)^k \cdot 1/(2k+1) = (-1)^k \cdot (2k)!$ ).

4. Der Tangens und der Kotangens können via

$$\tan(x) = x + \frac{1}{3} \cdot x^3 + \frac{2}{15} \cdot x^5 + \frac{17}{315} \cdot x^7 + \dots \quad \text{für } x \in (-\pi/2, \pi/2)$$

und

$$\cot(x) = -(x - \frac{1}{2}\pi) - \frac{1}{3} \cdot (x - \frac{1}{2}\pi)^3 - \frac{2}{15} \cdot (x - \frac{1}{2}\pi)^5 + \dots \quad \text{für } x \in (0, \pi)$$

auch als Potenz- bzw. Taylor-Reihe geschrieben werden, aber die Formeln für die Koeffizienten sind recht kompliziert.

## 5. Die Funktion

$$f(x) = \frac{\cos(x)}{1-x}$$

kann für  $|x| < 1$  als Potenzreihe und damit auch als Taylor-Reihe geschrieben werden. Insbesondere gilt

$$\begin{aligned} f(x) &= \left( \sum_{k=0}^{\infty} (-1)^k \cdot \frac{x^k}{(2k)!} \right) \cdot \left( \sum_{l=0}^{\infty} x^l \right) \\ &= \left( 1 - \frac{1}{2} \cdot x^2 + \frac{1}{24} x^4 + \dots \right) \cdot \left( 1 + x + x^2 + x^3 + \dots \right) \\ &= 1 + x + \frac{1}{2} \cdot x^2 + \frac{1}{2} \cdot x^3 + \frac{13}{24} \cdot x^4 + \frac{13}{24} \cdot x^5 + \frac{389}{720} \cdot x^6 + \frac{389}{720} \cdot x^7 + \dots \\ &= \sum_{m=0}^{\infty} \left( \sum_{k=0}^m \frac{(-1)^k}{(2k)!} \right) \cdot x^m, \end{aligned}$$

wobei wir die Cauchysche Produktformel für absolut konvergente Reihen (siehe Kapitel 7) verwendet haben.

# Kapitel 10

## Integralrechnung

**Motivation** Integrale sind sehr wichtig, weil sie

1. die Fläche unter einer Kurve liefern,
2. invers zum Ableitungsbegriff sind, und
3. in vielen Anwendungsproblemen auftauchen.

### 10.1 Bestimmtes Riemann-Integral

**Vorbemerkung** Im Folgenden sei  $f : [a, b] \rightarrow \mathbb{R}$  immer eine beschränkte Funktion, d.h. es gelte stets  $\sup_{x \in [a, b]} |f(x)| < \infty$ .

**Zerlegung des Intervalles in Teilintervalle** Eine Zerlegung des Intervalles  $[a, b]$  besteht aus  $n$  geordneten Punkten  $x_j$  mit

$$a = x_0 < x_1 < x_2 < \dots < x_{n-1} < x_n = b,$$

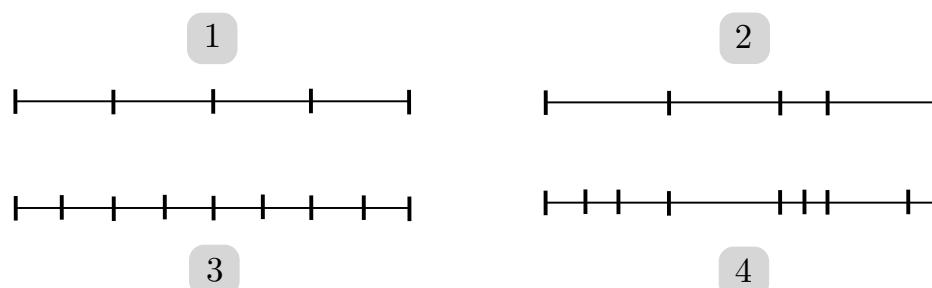
wobei  $n \in \mathbb{N}$  beliebig gewählt werden kann und

$$0 < \|Z\| := \max_{j=1 \dots n} |x_j - x_{j-1}| < \infty$$

die Feinheit von  $Z$  genannt wird. Eine weitere Zerlegung  $\tilde{Z}$  mit Punkten

$$a = \tilde{x}_0 < \tilde{x}_1 < \dots < \tilde{x}_{\tilde{n}} = b$$

heißt Verfeinerung von  $Z$ , wenn jeder Punkt in  $Z$  auch zu  $\tilde{Z}$  gehört, d.h. wenn für jedes  $j \in 1 \dots n$  ein  $\tilde{j} = 1 \dots \tilde{n}$  existiert, so dass  $x_j = \tilde{x}_{\tilde{j}}$  (siehe Bild).



Vier Zerlegungen eines kompakten Intervalles. Die erste und die dritte Zerlegung sind äquidistant (d.h. alle Teilintervalle  $[x_{j-1}, x_j]$  haben dieselbe Länge, nämlich  $(b - a)/n$ ), wohingegen die zweite und die vierte Zerlegung diese Eigenschaft nicht besitzen. Außerdem ist die dritte bzw. vierte Zerlegung eine Verfeinerung der ersten bzw. zweiten.

**Allgemeine und spezielle Riemann-Summen** Für jede Wahl von  $n$  Stützstellen  $\xi_j \in [x_{j-1}, x_j]$  wird

$$R(Z) = \sum_{j=1}^n f(\xi_j)(x_j - x_{j-1})$$

die Riemann-Summe bzgl. der Zerlegung und der Stützstellen genannt (wobei wir die Abhängigkeit von  $f$  und den  $\xi_j$  nicht explizit schreiben). Von besonderer Bedeutung sind die folgenden zwei Sonderfälle: Die Untersumme bzw. die Obersumme ist durch

$$U(Z) = \sum_{j=1}^n \inf f|_{[x_{j-1}, x_j]} \cdot (x_j - x_{j-1})$$

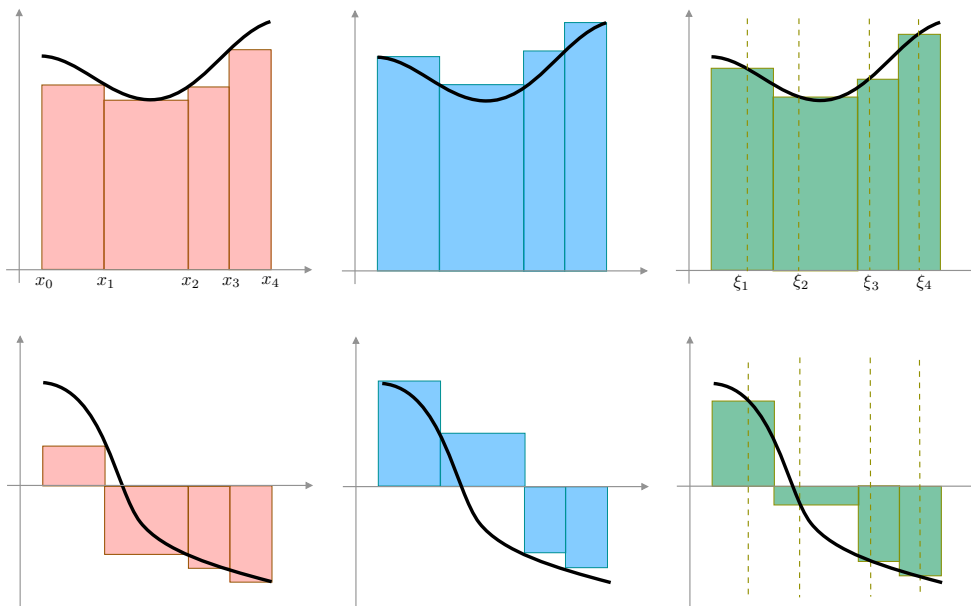
bzw.

$$O(Z) = \sum_{j=1}^n \sup f|_{[x_{j-1}, x_j]} \cdot (x_j - x_{j-1})$$

gegeben, wobei für eine stetige Funktion  $f$  die Formel

$$\begin{aligned} \inf f|_{[x_{j-1}, x_j]} &= \inf \{f(x) : x_{j-1} \leq x \leq x_j\} \\ &= \min \{f(x) : x_{j-1} \leq x \leq x_j\} = \min f|_{[x_{j-1}, x_j]} \end{aligned}$$

und bzw. analog  $\sup f|_{[x_{j-1}, x_j]} = \max f|_{[x_{j-1}, x_j]}$  gilt.



Untersumme (rot), Obersumme (blau) und Riemann-Summe (grün) für eine gegebene, nicht-äquidistante Zerlegung von  $[a, b]$  in vier Teilintervalle und zwei verschiedene Funktionen. Alle Summen approximieren das bestimmte Integral, d.h. die vorzeichenbehaftete Fläche unter dem Graphen der Funktion, wobei sich nach oben bzw. unten öffnende Balken einen positiven bzw. negativen Beitrag liefern.

### Beobachtung

1. Es gilt

$$U(Z) \leq R(Z) \leq O(Z)$$

für jede Zerlegung  $Z$ .

2. Ist  $\tilde{Z}$  eine beliebige Verfeinerung von  $Z$ , so folgt

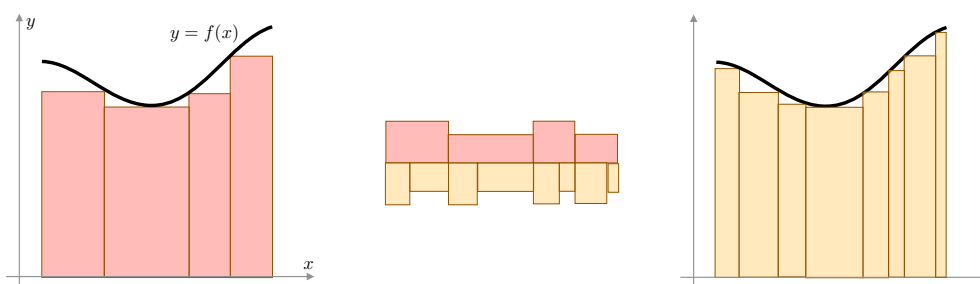
$$U(Z) \leq U(\tilde{Z}) \leq O(\tilde{Z}) \leq O(Z).$$

3. Sind  $Z^{(1)}$  und  $Z^{(2)}$  zwei beliebige Zerlegungen, so gilt

$$U(Z^{(1)}) \leq O(Z^{(2)}).$$

In der Tat, es gibt immer eine Zerlegung  $Z$ , die Verfeinerung von sowohl  $Z^{(1)}$  als auch  $Z^{(2)}$  ist, und für diese gilt

$$U(Z^{(1)}) \leq U(Z) \leq O(Z) \leq O(Z^{(2)})$$



Untersumme bzgl. einer groben (rot) und einer verfeinerten (gelb) Zerlegung.

Vorlesung 29, 15. Januar 2020

**Folgerung** Die reellen Zahlen

$$\bar{U} := \sup\{U(Z) : Z \text{ ist Zerlegung von } [a, b]\}$$

bzw.

$$\underline{O} := \inf\{O(Z) : Z \text{ ist Zerlegung von } [a, b]\}$$

sind wohldefiniert und werden Unterintegral bzw. Oberintegral von  $f$  genannt. Insbesondere gilt

$$U(Z) \leq \bar{U} \leq \underline{O} \leq O(Z),$$

für jede Zerlegung  $Z$ , d.h.  $\text{Untersumme} \leq \text{Unterintegral} \leq \text{Oberintegral} \leq \text{Obersumme}$ .

**Definition** Gilt

$$\bar{U} = \underline{O} =: \int_a^b f(x) dx$$

so heißt die Funktion  $f$  (Riemann-)integrierbar und das Symbol auf der rechten Seite wird das (Riemann-)Integral von  $f$  genannt.

**Notation** Das  $x$  im bestimmten Integral ist keine freie, sondern eine gebundene Variable. Insbesondere gilt

$$\int_a^b f(x) dx = \int_a^b f(\tilde{x}) d\tilde{x} = \int_a^b f(y) dy = \int_a^b f(t) dt.$$

**Integrierbarkeit und Folgen von Zerlegungen** Sei  $(Z^{(m)})_{m \in \mathbb{N}}$  eine beliebige Folge von Zerlegungen, wobei die  $m$ -te Zerlegung  $Z^{(m)}$  aus den  $n^{(m)}$  Punkten

$$a = x_0^{(m)} < x_1^{(m)} < x_2^{(m)} < \dots < x_{n-1}^{(m)} < x_{n^{(m)}}^{(m)} = b,$$

besteht. Gilt nun

$$\lim_{m \rightarrow \infty} U(Z^{(m)}) = \lim_{m \rightarrow \infty} O(Z^{(m)}),$$

so ist die Funktion  $f$  integrierbar (wobei die Grenzwerte gerade das Integral liefern).

In der Tat, nach den Beobachtungen von oben gilt

$$U(Z^{(m)}) \leq \bar{U} \leq \underline{O} \leq O(Z^{(m)})$$

und die Behauptung ergibt sich nach Grenzübergang  $m \rightarrow \infty$  aus dem Sandwich-Prinzip. Beachte, dass in aller Regel

$$\lim_{m \rightarrow \infty} \|Z^{(m)}\| = 0$$

gelten wird, dass heißt man kann nur dann die asymptotische Gleichheit von Unter- und Obersumme erwarten, wenn die Feinheit von  $Z^{(m)}$  mit  $m \rightarrow \infty$  gegen 0 konvergiert.

**Beispiel** Für eine affine Funktion  $f(x) = c_1 \cdot x + c_0$  gilt

$$\int_a^b f(x) dx = \frac{1}{2} \cdot c_1 \cdot (b-a)^2 + c_0 \cdot (b-a),$$

wobei wir den Beweis nur im vereinfachten Fall  $c_1 > 0$  mit  $a = 0$  und  $b = 1$  führen wollen (der allgemeine Fall geht analog mit modifizierten Formeln). Wir definieren durch

$$Z^{(m)} = \left( x_0^{(m)}, \dots, x_m^{(m)} \right), \quad x_j^{(m)} = \frac{j}{m}$$

eine Folge von jeweils äquidistanten Zerlegungen des Intervalles  $[0, 1]$  und berechnen

$$\begin{aligned} U(Z_m) &= \sum_{j=1}^m \frac{f(x_{j-1})}{m} = c_1 \cdot \sum_{j=1}^m \frac{j-1}{m^2} + c_0 \cdot \sum_{j=1}^m \frac{1}{m} \\ &= \frac{c_1}{m^2} \cdot \frac{m(m-1)}{2} + c_0 = \frac{c_1}{2} \left( 1 - \frac{1}{m} \right) + c_0 \end{aligned}$$

sowie

$$O(Z_m) = \sum_{j=1}^m \frac{f(x_j)}{m} = c_1 \cdot \sum_{j=1}^m \frac{j}{m^2} + c_0 \cdot \sum_{j=1}^m \frac{1}{m} = \frac{c_1}{2} \left( 1 + \frac{1}{m} \right) + c_0.$$

Insbesondere gilt  $\lim_{m \rightarrow \infty} U(Z_m) = \lim_{m \rightarrow \infty} O(Z_m) = \frac{1}{2} \cdot c_1 + c_0$ .

**Gegenbeispiel** Die Funktion

$$f(x) = \begin{cases} 1 & \text{falls } x \text{ rational} \\ 0 & \text{falls } x \text{ irrational} \end{cases} \quad \text{für } x \in [a, b] = [0, 1]$$

ist nicht Riemann-integrierbar, denn es gilt

$$U(Z) = \sum_{j=1}^n 0 \cdot (x_j - x_{j-1}) = 0$$

sowie

$$\begin{aligned} O(Z) &= \sum_{j=1}^n 1 \cdot (x_j - x_{j-1}) \\ &= (x_1 - x_0) + (x_2 - x_1) + \dots + (x_{n-2} - x_{n-1}) + (x_{n-1} - x_n) \\ &= x_n - x_0 = b - a = 1 \end{aligned}$$

für jede Zerlegung  $Z$  und damit  $\bar{U} = 0 < 1 = \underline{O}$ . Die Formel  $\sum_{j=1}^n x_j - x_{j-1} = x_n - x_0$  heißt übrigens Teleskopsummen-Prinzip.

**\*Lebesgue-Integral** Statt des Riemann-Integrals wird in der Mathematik meist das allgemeinere Konzept des Lebesgue-Integrals verwendet. Dabei gilt:

1. Jede Riemann-integrierbare Funktion ist Lebesgue-integrierbar und beide Integrale sind gleich.
2. Es gibt Funktionen, die Lebesgue-, aber nicht Riemann-integrierbar sind. Ein Standardbeispiel ist das angegebene Gegenbeispiel zur Riemann-Integrierbarkeit.
3. Das Lebesgue-Integral besitzt bessere Eigenschaften als das Riemann-Integral, aber die mathematische Theorie ist wesentlich anspruchsvoller.

Im Rahmen dieser Vorlesung betrachten wir nur das Riemann-Integral und meinen mit „integrierbar“ immer Riemann-integrierbar.



Beim Riemann-Integral (links) wird der Urbildbereich der Funktion in Teilintervalle zerlegt, wohingegen beim Lebesgue-Integral (rechts) der Bildbereich unterteilt wird. Für reguläre (zum Beispiel stetige) Funktionen sind beide Ansätze äquivalent, aber für irreguläre Funktionen führt die Lebesguesche Idee zu einem wesentlich robusteren Integralbegriff.

**Theorem (hinreichende Bedingungen für Integrierbarkeit)**

1. Jede monotone Funktion ist integrierbar.
2. Jede stetige Funktion ist integrierbar.

*Beweis, Teil 1:* Wir nehmen O.B.d.A. an, dass  $f$  monoton wachsend ist (der Beweis für monoton fallende Funktionen geht analog) und definieren durch

$$x_j^{(m)} = a + \frac{j}{m} \cdot (b - a), \quad n^{(m)} = m, \quad j = 0 \dots m$$

eine Folge äquidistanter Zerlegungen des Intervalls  $[a, b]$ . Dann gilt

$$\begin{aligned} O(Z^{(m)}) - U(Z^{(m)}) &= \sum_{j=1}^m f(x_j) \cdot (x_j - x_{j-1}) - \sum_{j=1}^m f(x_{j-1}) \cdot (x_j - x_{j-1}) \\ &= \frac{b-a}{m} \cdot \sum_{j=1}^m (f(x_j) - f(x_{j-1})) \\ &= \frac{b-a}{m} \cdot (f(b) - f(a)) \end{aligned}$$

nach der Teleskopsummenformel und es folgt  $\lim_{m \rightarrow \infty} U(Z^{(m)}) = \lim_{m \rightarrow \infty} O(Z^{(m)})$ .

*Beweis, Teil 2:* Wir wählen  $\varepsilon > 0$  beliebig und bemerken, dass nach dem Theorem über die Gleichmäßige Stetigkeit ein  $\delta > 0$  existiert, so dass

$$|x - \tilde{x}| < \delta \quad \implies \quad |f(x) - f(\tilde{x})| < \varepsilon$$

Ist nun  $Z$  eine Zerlegung (äquidistant oder nicht) mit  $\|Z\| < \delta$ , so folgt

$$\begin{aligned} 0 \leq O(Z) - U(Z) &= \sum_{j=1}^m (\max_{f|_{[x_{j-1}, x_j]}} - \min_{f|_{[x_{j-1}, x_j]}}) (x_{j-1} - x_j) \\ &< \varepsilon \sum_{j=1}^m (x_j - x_{j-1}) = \varepsilon \cdot (x_n - x_1) = \varepsilon \cdot (b - a), \end{aligned}$$

wobei wir benutzt haben, dass die stetige Funktion  $f$  auf dem Intervall  $[x_{j-1}, x_j]$  ihr Minimum und ihr Maximum annimmt, und dass der Abstand eines entsprechenden Minimierers und eines Maximierers durch  $x_j - x_{j-1} \leq \|Z\| < \delta$  beschränkt ist. Insgesamt erhalten wir

$$0 \leq \underline{Q} - \bar{U} \leq O(Z) - U(Z) \leq \varepsilon \cdot (b - a),$$

und weil  $\varepsilon$  beliebig klein gewählt werden kann, muss  $\bar{U} = \underline{Q}$  gelten.  $\square$

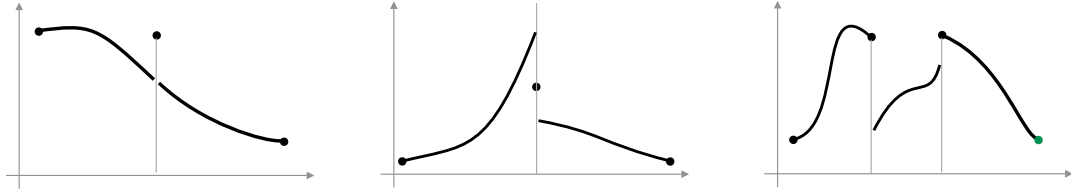
### Bemerkung

1. Das Theorem garantiert die Existenz des Integrals, aber liefert erstmal keine einfache Berechnungsformel.
2. Mit wenig mehr Aufwand kann man zeigen, dass auch jede stückweise stetige Funktion integrierbar ist, d.h. jede Funktion, die in *endlich* vielen Stellen unstetig, aber in allen anderen Stellen ihres Definitionsbereiches stetig ist.
3. Nicht ganz so einfach zu zeigen ist, dass auch eine Funktion mit nur abzählbar vielen Unstetigkeitsstellen immer noch Riemann-integrierbar ist. Ein Beispiel mit  $a = 0$  und  $b = 1$  ist die Funktion

$$f(x) = \begin{cases} 0 & \text{für } x \in [0, 1] \setminus \mathbb{Q} \\ 1/q & \text{für } x = p/q \text{ mit teilerfremden natürlichen Zahlen } 1 \leq p \leq q \end{cases}$$

deren Riemann-Integral übrigens 0 ist.





Typische Beispiele für beschränkte, aber unstetige Funktionen auf einem kompakten Intervall (hebbare Unstetigkeit, ein bzw. zwei Sprungstellen). Die Kreise stellen die die Funktionswerte am Rand und in den Unstetigkeitsstellen dar, aber diese isolierten Werte tragen nichts zum Integral bei.

**Eigenschaften von bestimmten Riemann-Integralen** Unsere bisherigen Definitionen und Resultate implizieren (Übungsaufgabe) die folgenden Aussagen:

1. (Linearität) Es gilt

$$\int_a^b \lambda \cdot f(x) + \tilde{\lambda} \cdot \tilde{f}(x) dx = \lambda \cdot \int_a^b f(x) dx + \tilde{\lambda} \cdot \int_a^b \tilde{f}(x) dx$$

für alle integrierbaren Funktionen  $f, \tilde{f}$  und alle Zahlen  $\lambda, \tilde{\lambda} \in \mathbb{R}$ .

2. (Gebietsadditivität) Für jedes  $\eta$  mit  $a \leq \eta \leq b$  gilt

$$\int_a^b f(x) dx = \int_a^{\eta} f(x) dx + \int_{\eta}^b f(x) dx.$$

Aus Konsistenzgründen definiert man daher auch

$$\int_{\eta}^{\eta} f(x) dx := 0, \quad \int_b^a f(x) dx := - \int_a^b f(x) dx.$$

3. (Monotonie und Positivität) Es gelten die Implikationen

$$f(x) \geq \tilde{f}(x) \quad \text{für alle } x \in [a, b] \quad \Longrightarrow \quad \int_a^b f(x) dx \geq \int_a^b \tilde{f}(x) dx$$

sowie

$$f(x) \geq 0 \quad \text{für alle } x \in [a, b] \quad \Longrightarrow \quad \int_a^b f(x) dx \geq 0.$$

4. (Maximum- und Minimumprinzip) Es gilt

$$(\inf f) \cdot (b - a) \leq \int_a^b f(x) dx \leq (\sup f) \cdot (b - a)$$

für jede integrierbare Funktion auf  $[a, b]$ .

5. („Dreiecksungleichung“ für Integrale)

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

bzw.

$$\left| \int_a^b (f(x) - \tilde{f}(x)) \, dx \right| \leq \int_a^b |f(x) - \tilde{f}(x)| \, dx.$$

**Funktionen mit überall verschwindendem Integral** Gilt

$$\int_{\tilde{a}}^{\tilde{b}} f(x) \, dx = 0$$

für alle  $\tilde{a}, \tilde{b}$  mit  $a \leq \tilde{a} \leq \tilde{b} \leq b$ , so gilt  $f(x) = 0$  in allen Stetigkeitsstellen. Beachte hierbei, dass dies nicht in Unstetigkeitspunkten gelten muss. Ein Standardgegenbeispiel mit  $a = -1$  und  $b = 1$  ist

$$f(x) = \begin{cases} 1 & \text{falls } x = 0 \\ 0 & \text{sonst} \end{cases},$$

denn man kann für jedes Teilintervall  $[\tilde{a}, \tilde{b}]$  mit Hilfe von Ober- und Untersummen leicht zeigen, dass das entsprechende Integral von  $f$  verschwindet.

## 10.2 Der Hauptsatz der Differential- und Integralrechnung

Vorlesung 30, 17. Januar 2020

**Theorem (HDI, Fundamentalsatz der Analysis)** Ist  $f : [a, b] \rightarrow \mathbb{R}$  stetig und ist  $F : [a, b] \rightarrow \mathbb{R}$  Stammfunktion von  $f$ , d.h.  $F'(\xi) = f(\xi)$  für alle  $\xi \in [a, b]$ , so gilt

$$\int_{\alpha}^{\beta} f(x) \, dx = F(\beta) - F(\alpha)$$

für alle  $\alpha, \beta \in [a, b]$ . Insbesondere werden durch

$$F_1(\xi) := + \int_a^{\xi} f(x) \, dx \quad \text{und} \quad F_2(\xi) := - \int_{\xi}^b f(x) \, dx$$

zwei spezielle Stammfunktionen  $F_1$  und  $F_2$  von  $f$  definiert.

Beweis, Teil 1: Wir beweisen zunächst den zweiten Teil der Behauptung. Für alle  $\xi \in$

$[a, b]$  und alle  $h > 0$  mit  $\xi + h \in [a, b]$  ergibt sich aus den Eigenschaften bestimmter Integrale die Formel

$$F_1(\xi + h) - F_1(\xi) := \int_a^{\xi+h} f(x) dx - \int_a^{\xi} f(x) dx = \int_{\xi}^{\xi+h} f(x) dx.$$

Da außerdem auch

$$\int_{\xi}^{\xi+h} f(\xi) dx = f(\xi) \int_{\xi}^{\xi+h} 1 dx = f(\xi) \cdot h$$

gilt (wir integrieren über  $x$ , aber der Integrand hängt nur von  $\xi$  ab), erhalten wir

$$\begin{aligned} \left| \frac{F_1(\xi + h) - F_1(\xi)}{h} - f(\xi) \right| &= \left| \frac{1}{h} \int_{\xi}^{\xi+h} (f(x) - f(\xi)) dx \right| \\ &\leq \frac{1}{h} \int_{\xi}^{\xi+h} |f(x) - f(\xi)| dx \leq C(\xi, h) \frac{h}{h} \end{aligned}$$

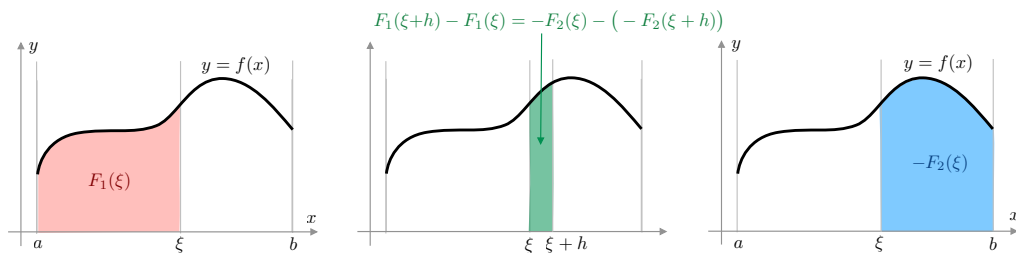
mit  $C(\xi, h) := \sup_{x \in [\xi, \xi+h]} |f(x) - f(\xi)|$ . Die Stetigkeit von  $f$  impliziert  $C(\xi, h) \rightarrow 0$  für  $h \searrow 0$  und damit

$$\lim_{h \searrow 0} \frac{F_1(\xi + h) - F_1(\xi)}{h} = f(\xi).$$

Eine analoge Rechnung gilt für  $h < 0$  und wir erhalten insgesamt das gewünschte Ergebnis für  $F_1$ . Desweiteren gilt

$$F_1(\xi + h) - F_1(\xi) = -F_2(\xi) + F_2(\xi + h)$$

und dies impliziert  $F_1'(\xi) = F_2'(\xi)$  nach Division durch  $h$  und Grenzübergang  $h \rightarrow 0$ .



Zur Aussage und zum Beweis des Hauptsatzes der Differential und Integralrechnung.

*Beweis, Teil 2:* Wir beweisen nun den ersten Teil der Behauptung. Nach Voraussetzung an  $F$  und aufgrund des ersten Beweisteils gilt  $F'(\xi) = F_1'(\xi)$  für alle  $\xi \in [a, b]$ , d.h. die Ableitung der Differenzfunktion  $F - F_1$  überall. Der Mittelwertsatz impliziert (siehe Kapitel 9.1) die Existenz einer Konstanten  $c$ , so dass

$$F(\xi) - F_1(\xi) = c$$

für alle  $\xi \in [a, b]$ . Insbesondere gilt  $F(\beta) - F_1(\beta) = F(\alpha) - F_1(\alpha)$  und nach Umstellung der Terme sowie der Eigenschaften bestimmter Integrale ergibt sich via

$$F(\beta) - F(\alpha) = F_1(\beta) - F_1(\alpha) = \int_a^{\beta} f(x) dx - \int_a^{\alpha} f(x) dx = \int_{\alpha}^{\beta} f(x) dx,$$

die gewünschte Formel. □

**Beispiel** Für  $f(x) = x^2$  definiert  $F(x) = \frac{1}{3}x^3 + C$  eine Stammfunktion, wobei  $C$  die sogenannte Integrationskonstante ist. Damit gilt

$$\int_{\alpha}^{\beta} f(x) = F(\beta) - F(\alpha) = \frac{1}{3}(\beta^3 - \alpha^3)$$

sowie

$$F_1(\xi) = \frac{1}{3}\xi^3 - \frac{1}{3}a^3, \quad F_2(\xi) = \frac{1}{3}\xi^3 - \frac{1}{3}b^3$$

für alle  $\xi$  nach dem Hauptsatz der Differential- und Integral-Gleichung. Letzteres kann auch als

$$F_1(x) = \frac{1}{3}x^3 - \frac{1}{3}a^3, \quad F_2(x) = \frac{1}{3}x^3 - \frac{1}{3}b^3$$

geschrieben werden. Insbesondere gilt  $F = F_1$  bzw.  $F = F_2$ , sofern  $C = -\frac{1}{3}a^3$  bzw.  $C = -\frac{1}{3}b^3$  gewählt wird.

### Bemerkung

1. Ganz allgemein gilt

$$\frac{d}{d\beta} \int_{\alpha}^{\beta} f(x) dx = +f(\beta) \quad \text{sowie} \quad \frac{d}{d\alpha} \int_{\alpha}^{\beta} f(x) dx = -f(\alpha),$$

d.h. die Ableitung des bestimmten Integrals von  $f$  nach der oberen bzw. unteren Integralgrenze ist  $+f$  bzw.  $-f$ , ausgewertet an der oberen bzw. unteren Integralgrenze.

2. *Achtung:* Ausdrücke wie  $\int_a^x f(x) dx$  oder  $\int_x^b f(x) dx$  haben keinen Sinn, da  $x$  nicht gleichzeitig Integrationsvariable und obere/untere Integrationsgrenze sein kann.
3. Die Aussagen des Hauptsatzes gelten auch, wenn  $f$  nur stückweise stetig ist, wobei dann jede Stammfunktion  $F$  in den Unstetigkeitsstellen von  $f$  nicht differenzierbar ist, sondern Knicke aufweist.
4. Eine Stammfunktion von  $f$  wird auch als

$$\int f(x) dx$$

geschrieben und unbestimmtes Integral (also ohne Integrationsgrenzen) genannt. Das unbestimmte Integral ist dabei nur bis auf eine Integrationskonstante (oftmals  $C$  genannt) bestimmt, und dieser Freiheitsgrad sollte immer mit angegeben werden.

*Achtung:* Das unbestimmte Integral ist immer noch eine Funktion in  $x$ , d.h. die Ausdrücke  $\int f(x) dx$  und  $\int f(y) dy$  sind nicht identisch (bei bestimmten Integralen war das anders). Insbesondere gilt

$$\int x^2 dx = \frac{1}{3}x^3 + C, \quad \int y^2 dy = \frac{1}{3}y^3 + C, \quad \int_0^1 x^2 dx = \frac{1}{3} = \int_0^1 y^2 dy.$$

5. Man kann den ersten Teil des Hauptsatzes auch wie folgt formulieren:

$$\int_{\alpha}^{\beta} \frac{dF}{dx}(x) dx = \int_{\alpha}^{\beta} F'(x) dx = F(\beta) - F(\alpha).$$

**Wichtige Stammfunktionen** Die folgenden Formeln können mittels Differentiation nach  $x$  nachgerechnet werden, wobei die (!)-Einträge besonders wichtig sind und  $p$  einen reellen Parameter bezeichnet:

$$\begin{aligned}
 (!) \int x^p dx &= \frac{x^{p+1}}{p+1} + C && \text{für } p \neq -1 \\
 (!) \int \frac{dx}{x} &= \ln|x| + C && \text{für } x \neq 0 \\
 (!) \int \sin(x) dx &= -\cos(x) + C \\
 (!) \int \cos(x) dx &= \sin(x) + C \\
 \int \tan(x) dx &= -\ln|\cos(x)| + C && \text{für } \cos(x) \neq 0 \\
 \int \cot(x) dx &= \ln|\sin(x)| + C && \text{für } \sin(x) \neq 0 \\
 \int \frac{dx}{\cos^2(x)} &= \tan(x) + C && \text{für } x \neq k\pi + \pi/2, k \in \mathbb{Z} \\
 \int \frac{dx}{\sin^2(x)} &= -\cot(x) + C && \text{für } x \neq k\pi, k \in \mathbb{Z} \\
 \int \frac{dx}{\sqrt{1-x^2}} &= \arcsin(x) + C && \text{für } |x| < 1 \\
 \int \frac{dx}{\sqrt{x^2-1}} &= \ln|x + \sqrt{x^2-1}| + C && \text{für } |x| > 1 \\
 \int \frac{dx}{\sqrt{1+x^2}} &= \ln(x + \sqrt{1+x^2}) + C \\
 \int \frac{dx}{1+x^2} &= \arctan(x) + C \\
 \int \frac{dx}{1-x^2} &= \frac{1}{2} \ln \left| \frac{1+x}{1-x} \right| + C && \text{für } |x| \neq 1 \\
 (!) \int \exp(px) dx &= \frac{\exp(px)}{p} + C && \text{für } p \neq 0 \\
 (!) \int p^x dx &= \frac{p^x}{\ln p} + C && \text{für } p > 0 \text{ mit } p \neq 1 \\
 (!) \int \ln|x| dx &= x \cdot \ln|x| - x + C && \text{für } x \neq 0 \\
 (!) \int \sinh(x) dx &= \cosh(x) + C \\
 (!) \int \cosh(x) dx &= \sinh(x) + C \\
 \int \tanh(x) dx &= \ln(\cosh(x)) + C \\
 \int \coth(x) dx &= \ln|\sinh(x)| + C && \text{für } x \neq 0
 \end{aligned}$$

Diese Liste der bekannten Integrale ist bei weitem nicht vollständig. Es fehlen zum Beispiel die Umkehrfunktionen der trigonometrischen und der hyperbolischen Funktionen (arcsin und arsinh usw.), deren Integrale man etwa unter WIKIPEDIA in der *Tabelle von Ableitungs- und Stammfunktionen* oder im Anhang von [PapF] findet.

**Bemerkung**

1. Das Bilden einer Stammfunktion ist linear im Sinne von

$$\int (\lambda \cdot f(x) + \tilde{\lambda} \cdot \tilde{f}(x)) dx = \lambda \cdot \int f(x) dx + \tilde{\lambda} \cdot \int \tilde{f}(x) dx.$$

2. Differentiation und Integration sind (in gewisser Weise) zueinander *inverse* Operationen. Beachte aber, dass in der Praxis das Differenzieren oftmals vergleichsweise einfach ist (da es die Ketten- und die Produktregel gibt), aber dass die exakte Berechnung von Integralen sehr schnell sehr kompliziert wird.
3. Einige nicht explizit berechenbare Integrale haben spezielle Namen bekommen, zum Beispiel:

$$\text{Si}(\xi) := \int_0^{\xi} \frac{\sin(x)}{x} dx \quad (\text{Integralsinus-Funktion})$$

$$\text{Erf}(\xi) := \frac{2}{\sqrt{\pi}} \int_0^{\xi} \exp(-x^2) dx \quad (\text{Fehlerfunktion})$$

**Partielle Integration** Die Produktregel der Differentiation impliziert

$$\int u(x) \cdot v'(x) dx = u(x) \cdot v(x) - \int u'(x) \cdot v(x) dx$$

für unbestimmte Integrale bzw.

$$\int_a^b u(x) \cdot v'(x) dx = u(b) \cdot v(b) - u(a) \cdot v(a) - \int_a^b u'(x) \cdot v(x) dx$$

für bestimmte Integrale. Dabei wird oftmals

$$u(b) \cdot v(b) - u(a) \cdot v(a) = \left[ u(x) \cdot v(x) \right]_{x=a}^{x=b}$$

geschrieben.

**Beispiele**

1. Wegen  $\exp' = \exp$  erhalten wir

$$\int x \cdot \exp(x) dx = x \cdot \exp(x) - \int 1 \cdot \exp(x) dx = (x - 1) \cdot \exp(x) + C$$

für das unbestimmte Integral und dies impliziert

$$\int_a^b x \cdot \exp(x) dx = (b - 1) \cdot \exp(b) - (a - 1) \cdot \exp(a)$$

für das bestimmte Integral über dem Intervall  $[a, b]$ .

2. Für das unbestimmte Integral gilt

$$\int \ln(x) \, dx = \int \ln(x) \cdot 1 \, dx = x \cdot \ln(x) - \int \frac{1}{x} \cdot x \, dx = x \ln(x) - x + C$$

und für bestimmte Integrale erhalten wir analog

$$\int_a^b \ln(x) \, dx = \left[ x \cdot \ln(x) \right]_{x=a}^{x=b} - \int_a^b 1 \, dx = b \cdot \ln b - a \cdot \ln a - b + a.$$

*Bemerkung:*  $\ln(x)$  ist nur für  $x > 0$  definiert und daher ist auch die Stammfunktion  $\int \ln(x) \, dx$  nur auf dem offenen Intervall  $(0, \infty)$  definiert.

3. Die Ableitungsformeln der trigonometrischen Funktionen implizieren

$$\begin{aligned} \int \sin(x) \cdot \sin(x) \, dx &= -\sin(x) \cdot \cos(x) + \int \cos(x) \cdot \cos(x) \, dx \\ &= -\sin(x) \cdot \cos(x) + \int (1 - \sin(x) \cdot \sin(x)) \, dx \end{aligned}$$

und damit

$$\int \sin^2(x) \, dx = \frac{1}{2} (x - \sin(x) \cdot \cos(x)) + C$$

bzw.

$$\int \cos^2(x) \, dx = \frac{1}{2} (x + \sin(x) \cdot \cos(x)) + C,$$

wobei die Integrationskonstante  $C$  in jeder der beiden Formeln beliebig gewählt werden darf. Analoge Rechnungen kann man für die bestimmten Integrale durchführen.

Vorlesung 31, 20. Januar 2020

**Substitutionsformel** Ist  $h : [c, d] \rightarrow [a, b]$  stetig differenzierbar, so gilt

$$\int_{h(c)}^{h(d)} f(x) \, dx = \int_c^d f(h(s)) \cdot h'(s) \, ds.$$

Man kann diese sehr wichtige und nützliche Formel als die Integral-Variante der Kettenregel betrachten. In der Tat, ist  $F$  eine Stammfunktion zu  $f$ , so gilt

$$\int_c^d f(h(s)) \cdot h'(s) \, ds = \int_c^d \frac{d}{ds} F(h(s)) \, ds = F(h(d)) - F(h(c)) = \int_{h(c)}^{h(d)} f(x) \, dx,$$

wobei wir einmal die Kettenregel und zweimal den Hauptsatz angewendet haben. Alternativ kann man die Substitutionsregel auch wie folgt verstehen: Wir setzen

$$x = h(s) \quad \text{und damit} \quad f(x) = f(h(s)).$$

Außerdem schreiben wir die Formel

$$\frac{dx}{ds} = \frac{d}{ds}h(s) = h'(s)$$

symbolisch als

$$dx = h'(s) ds$$

und bemerken, dass  $s = c$  bzw.  $s = d$  genau dann gilt, wenn  $x = h(c)$  bzw.  $x = h(d)$  gilt. Die rechte Seite in der Substitutionsformeln erhält man nun, indem alle Terme auf der linken Seite *substituiert* werden.

Das Analogon für unbestimmte Integrale ist

$$\int f(x) dx = \int f(h(s)) \cdot h'(s) ds,$$

wobei die linke Seite eine Funktion in  $x$  und die rechte Seite eine Funktion in  $s$  ist, d.h. man muss nach Berechnung der Stammfunktion entweder links  $x$  durch  $s$  ersetzen oder rechts  $s$  durch  $x$  ersetzen (siehe Beispiel).

## Beispiele

- Um das bestimmte Integral

$$I := \int_{-c}^{+c} \sqrt{1 - \frac{x^2}{c^2}} dx$$

zu berechnen, substituieren wir

$$x = c \sin(s) \quad \text{mit} \quad \frac{dx}{ds} = c \cos(s) \quad \text{bzw.} \quad dx = c \cos(s) ds$$

und bemerken, dass

$$x \in [-c, +c] \quad \Leftrightarrow \quad s \in [-\pi/2, +\pi/2], \quad x = \pm c \quad \Leftrightarrow \quad s = \pm\pi/2.$$

Deshalb und wegen  $\cos(s) \geq 0$  erhalten wir mit

$$\begin{aligned} I &= c \cdot \int_{-\pi/2}^{+\pi/2} \sqrt{1 - \sin^2(s)} \cdot \cos(s) ds = c \cdot \int_{-\pi/2}^{+\pi/2} \cos^2(s) ds \\ &= \frac{1}{2} \cdot c \cdot \left[ s + \sin(s) \cdot \cos(s) \right]_{s=-\pi/2}^{s=+\pi/2} = \frac{1}{2} \cdot \pi \cdot c \end{aligned}$$

eine explizite Formel für das gesuchte Integral, wobei die Stammfunktion von  $\cos^2$  bereits in den Beispielen zu partieller Integration berechnet hatten.

- Wir wollen das bestimmte Integral

$$I := \int_1^e \frac{dx}{x(1 + \ln x)}$$



berechnen und substituieren

$$s = \ln(x), \quad ds = \frac{dx}{x}.$$

Wegen  $0 = \ln 1$  und  $1 = \ln e$  erhalten wir

$$I = \int_0^1 \frac{ds}{1+s} = \left[ \ln(1+s) \right]_{s=0}^{s=1} = \ln 2$$

und damit die gesuchte Zahl.

*Bemerkung:* Wir hätten dieses Ergebnis wegen

$$\frac{d}{dx} \ln(1 + \ln x) = \frac{1}{x(1 + \ln(x))}$$

alternativ auch direkt mit dem Hauptsatz berechnen können:

$$I = \ln(1 + \ln e) - \ln(1 + \ln 1) = \ln 2.$$

3. Für  $x > 0$  gilt

$$\int 2x \exp(-x^2) dx = \int 2\sqrt{s} \exp(-s) \frac{ds}{2\sqrt{s}} = \int \exp(-s) ds = -\exp(-s) + C$$

wobei wir die Substitution

$$x = \sqrt{s}, \quad dx = \frac{ds}{2\sqrt{s}}$$

verwendet haben. Die Formel hat aber nur dann Sinn, wenn wir die rechte Seite durch Rücksubstitution als Funktion von  $x$  schreiben. Dies liefert

$$\int 2x \exp(-x^2) dx = -\exp(-x^2) + C.$$

*Bemerkung:* Alternativ hätte wir auch diese Formel via

$$\int 2x \exp(-x^2) dx = \int \frac{d}{dx} \exp(-x^2) dx$$

direkt aus dem Hauptsatz der Differential- und Integralrechnung ableiten können.

4. Mit der Substitution

$$\sqrt{x} = s, \quad ds = \frac{dx}{2\sqrt{x}}, \quad dx = 2s ds$$

ergibt sich

$$\begin{aligned} \int \exp(\sqrt{x}) dx &= \int 2s \exp(s) ds \\ &= 2(s-1) \exp(s) + C = 2(\sqrt{x}-1) \exp(\sqrt{x}) + C, \end{aligned}$$

wobei wir das Integral bzgl.  $s$  mit partieller Integration berechnet haben und im letzten Schritt wieder die Rücksubstitution durchgeführt haben.

*Bemerkung:* Wir können abschließend und bei Bedarf noch die Probe machen: Die Ableitung der rechten Seite nach  $x$  liefert gerade den Integranden auf der linken Seite (Nachrechnen!).

5. Durch Substitution von

$$s = e^x, \quad dx = \frac{ds}{\frac{ds}{dx}} = \frac{ds}{e^x} = \frac{ds}{s}$$

berechnen wir

$$\begin{aligned} \int \frac{1}{e^x + e^{-x}} dx &= \int \frac{1}{s + \frac{1}{s}} \frac{ds}{s} \\ &= \int \frac{ds}{s^2 + 1} = \arctan(s) + C = \arctan(e^x) + C. \end{aligned}$$

6. Die Formeln

$$\int \frac{g'(x)}{g(x)} dx = \ln |g(x)| + C, \quad \int g'(x) \cdot g(x) dx = \frac{1}{2} g(x)^2 + C$$

können via  $s = g(x)$  und  $ds = g'(x) ds$  abgeleitet werden.

### **Bemerkungen**

1. Wird das unbestimmte Integral  $\int f(x) dx$  durch Substitution der Variablen  $x$  berechnet, so muss die Rücksubstitution durchgeführt werden, da wir am Ende ja eine Stammfunktion in der Variablen  $x$  erhalten müssen. Bei der Substitution im bestimmten Integral  $\int_a^b f(x) dx$  gibt es keine Rücksubstitution, da das Ergebnis nicht von  $x$  abhängt (sondern im einfachsten Fall eine Zahl ist).
2. Sowohl für bestimmte als auch für unbestimmte Integrale gilt: Die Substitutionsregel ist viel einfacher und robuster als sie auf den ersten Blick erscheinen mag. Nach etwas Übung kann man mit ihr sehr gut komplizierte Integrale berechnen.
3. Für viele Klassen von Integranden ist bekannt, welche Substitution zum Ziel, d.h. zu expliziten oder vereinfachten Formeln, führt (siehe die Große Übung II). Im Allgemeinen gibt es aber leider keinen Algorithmus und kein Kochrezept zum Auffinden einer geeigneten Substitution.
4. Für die Gültigkeit der Substitutionsformel bestimmter Integrale muss man nicht voraussetzen, dass  $h$  bijektiv ist. In der Praxis wird  $h$  aber meist strikt monoton wachsend mit  $a = h(c)$  und  $b = h(d)$  oder strikt monoton fallend mit  $a = h(d)$  und  $b = h(c)$  sein.

**Zwei wichtige Integralsätze** Seien  $f, p : [a, b] \rightarrow \mathbb{R}$  zwei stetige Funktionen und gelte  $p(x) \geq 0$  für alle  $x \in [a, b]$ .

**Lemma (Mittelwertsatz für Integrale)** Es existiert (mindestens) ein  $\xi \in [a, b]$ , so dass

$$f(\xi) = \frac{\int_a^b f(x) p(x) dx}{\int_a^b p(x) dx},$$

wobei die rechte Seite der Formel das Integralmittel von  $f$  bzgl.  $p$  ist und  $p$  auch Gewichtsfunktion genannt wird.

Beweis. Da  $f$  stetig ist, existieren  $\min f$  und  $\max f$  und es gilt

$$\min f \leq f(x) \leq \max f$$

für alle  $x \in [a, b]$ . Nach Multiplikation mit  $p(x) \geq 0$  und anschließender Integration erhalten wir

$$(\min f) \int_a^b p(x) \, dx \leq \int_a^b f(x) p(x) \, dx \leq \int_a^b (\max f) p(x) \, dx,$$

d.h. die rechte Seite in der Behauptung liegt zwischen  $\min f$  und  $\max f$ . Der Zwischenwertsatz für stetige Funktionen liefert nun die Existenz von  $\xi$ .  $\square$

**Bemerkung** Gilt  $p(x) = 1$  für alle  $x \in [a, b]$ , so gilt

$$F'(\xi) = f(\xi) = \frac{\int_a^b f(x) \, dx}{b-a} = \frac{F(b) - F(a)}{b-a}$$

für jede Stammfunktion  $F$  von  $f$  und der Mittelwertsatz für Integrale ist gerade der Erste Mittelwertsatz für  $F$ .

**Lemma (Jensensche Ungleichung)** Es gilt

$$\psi \left( \frac{\int_a^b f(x) p(x) \, dx}{\int_a^b p(x) \, dx} \right) \leq \frac{\int_a^b \psi(f(x)) p(x) \, dx}{\int_a^b p(x) \, dx}$$

für jede konvexe Funktion  $\psi : \mathbb{R} \rightarrow \mathbb{R}$ .

Beweis. Es sei auf die Literatur verwiesen. Die Strategie ist, zunächst eine analoge Ungleichung für Riemann-Summen zu etablieren und anschließend einen Grenzübergang durchzuführen.

**Beispiele**

1. Mit  $p(x) = 1$  ergibt sich

$$\psi \left( \frac{1}{b-a} \int_a^b f(x) \, dx \right) \leq \frac{1}{b-a} \int_a^b \psi(f(x)) \, dx$$

für jedes konvexe  $\psi$ .

2. Mit  $\psi(s) = s^2$  erhalten wir

$$\left( \int_a^b f(x) p(x) \, dx \right)^2 \leq \left( \int_a^b f^2(x) p(x) \, dx \right) \cdot \left( \int_a^b p(x) \, dx \right).$$

## 10.3 Weitere Aspekte der Integralrechnung

Vorlesung 32, 22. Januar 2020

### Stammfunktionen rationaler Funktionen

**Vorbemerkung** Für eine rationale Funktion

$$f(x) = \frac{p(x)}{q(x)} \quad \text{mit reellen Polynomen } p \text{ und } q$$

können unbestimmte Integrale explizit berechnet werden, sofern die Nullstellen des Nennerpolynoms  $q$  bekannt sind. Ein wesentlicher Teil der Berechnung der Integrale ist die sogenannte *Partialbruchzerlegung*, die  $f$  als Summe einfacher Bausteine (die auch Partialbrüche genannt werden) darstellt. Wir berechnen nun zunächst die Stammfunktionen der Bausteine und diskutieren anschließend sowohl Beispiele als auch das allgemeine Konzept.

**Baustein 1 – Polynome** Ist das Nennerpolynom trivial ( $q \equiv 1$ ), so ist die rationale Funktion  $f$  ein Polynom und die Stammfunktion kann via

$$\int (\alpha_m \cdot x^m + \dots + \alpha_1 \cdot x + \alpha_0) dx = \alpha_m \cdot \frac{x^{m+1}}{m+1} + \dots + \alpha_1 \cdot \frac{x^2}{2} + \alpha_0 \cdot x + C$$

explizit berechnet werden (mit freier Integrationskonstante  $C$ ).

**Baustein 2 – Inverse Monome mit reellen Nullstellen** Für festes  $x_*$  gilt

$$\int \frac{dx}{(x - x_*)} = \ln |x - x_*| + C$$

bzw.

$$\int \frac{dx}{(x - x_*)^k} dx = \frac{1}{(1 - k) \cdot (x - x_*)^{k-1}} + C$$

für alle  $k \in \mathbb{N}$  mit  $k > 1$ , wobei wir immer stillschweigend  $x \neq x_*$  voraussetzen.

**Baustein 3 – Inverse Monome mit nicht-reellen Nullstellen** Wir wollen nun die Stammfunktionen

$$I_k(x) = \int \frac{1}{((x - a)^2 + b^2)^k} dx, \quad J_k(x) = \int \frac{x - a}{((x - a)^2 + b^2)^k} dx$$

berechnen. Für  $k = 1$  ergeben sich durch Hinsehen (bzw. Nachrechnen) die Formeln

$$I_1(x) = \frac{1}{b} \cdot \arctan \left( \frac{x - a}{b} \right) + C, \quad J_1(x) = \frac{1}{2} \cdot \ln \left( (x - a)^2 + b^2 \right) + C$$

und für  $k > 1$  gilt

$$\begin{aligned} J_k(x) &= -\frac{1}{2 \cdot (k-1)} \int \frac{d}{dx} \left( \frac{1}{((x - a)^2 + b^2)^{k-1}} \right) dx \\ &= -\frac{1}{2 \cdot (k-1) \cdot ((x - a)^2 + b^2)^{k-1}} + C. \end{aligned}$$

Die Funktion  $I_k$  kann leider nicht direkt, sondern nur rekursiv berechnet werden. Dazu bemerken wir

$$\begin{aligned}
 I_{k-1}(x) &= \int \frac{b^2}{((x-a)^2 + b^2)^k} dx + \int \frac{(x-a)^2}{((x-a)^2 + b^2)^k} dx \\
 &= b^2 \cdot I_k(x) + \int \frac{x-a}{2 \cdot (k-1)} \cdot \frac{2 \cdot (k-1) \cdot (x-a)}{((x-a)^2 + b^2)^k} dx \\
 &= b^2 \cdot I_k(x) - \int \frac{x-a}{2 \cdot (k-1)} \cdot \frac{d}{dx} \frac{1}{((x-a)^2 + b^2)^{k-1}} dx \\
 &= b^2 \cdot I_k(x) - \frac{x-a}{2 \cdot (k-1) \cdot ((x-a)^2 + b^2)^{k-1}} \\
 &\quad + \int \frac{dx}{2 \cdot (k-1) \cdot ((x-a)^2 + b^2)^{k-1}} \\
 &= b^2 \cdot I_k(x) + \frac{I_{k-1}(x)}{2 \cdot (k-1)} - \frac{x-a}{2 \cdot (k-1) \cdot ((x-a)^2 + b^2)^{k-1}}
 \end{aligned}$$

und erhalten nach Umstellung die Rekursionsformel

$$I_k(x) = \frac{2 \cdot k - 3}{b^2 \cdot 2 \cdot (k-1)} \cdot I_{k-1}(x) + \frac{x-a}{b^2 \cdot 2 \cdot (k-1) \cdot ((x-a)^2 + b^2)^{k-1}}.$$

Insbesondere können wir  $I_k$  aus  $I_{k-1}$  berechnen, zum Beispiel

$$I_2(x) = \frac{1}{2 \cdot b^3} \cdot \arctan\left(\frac{x-a}{b}\right) + \frac{x-a}{2 \cdot b^2 \cdot ((x-a)^2 + b^2)}$$

aus der Formel für  $I_1$ .

### Einfache Partialbruchzerlegungen

1. Ein sehr einfaches Beispiel ist

$$f(x) = \frac{2 \cdot x - 1}{x + 2} = \frac{2 \cdot (x + 2) - 5}{x + 2} = 2 - \frac{5}{x + 2}$$

denn die Partialbruchzerlegung — alle Summanden auf rechten Seite sind Bausteine — kann hier sehr einfach abgeleitet werden und liefert

$$\int f(x) dx = 2 \cdot x - 5 \cdot \ln|x+2| + C,$$

wobei die Stammfunktion nur für  $x \neq -2$  definiert ist und  $C$  die freie Integrationskonstante bezeichnet.

2. Ein komplizierteres, aber immer noch einfaches Beispiel ist die Funktion

$$\begin{aligned}
 f(x) &= \frac{x^2 + 2 \cdot x - 4}{x - 1} = \frac{((x-1) + 1)^2 + 2 \cdot ((x-1) + 1) - 4}{x - 1} \\
 &= \frac{(x-1)^2 + 4 \cdot (x-1) - 1}{x - 1} = (x-1) + 4 - \frac{1}{x-1},
 \end{aligned}$$

für die wir wieder die Partialbruch- (bzw. Baustein-) Zerlegung zu Fuß hergeleitet haben. Insgesamt erhalten wir

$$\int f(x) dx = \frac{1}{2} \cdot x^2 + 3 \cdot x - \ln|x-1| + C,$$

wobei diese Stammfunktion nur für  $x \neq 1$  definiert ist.

### 3. Für die rationale Funktion

$$f(x) = \frac{x-5}{x^2-4} = \frac{x-5}{(x-2) \cdot (x+2)}$$

beginnen wir mit dem Ansatz

$$f(x) = \frac{c_1}{x-2} + \frac{c_2}{x+2}$$

und wollen nun die beiden Koeffizienten  $c_1, c_2$  bestimmen, denn wenn diese existieren und bekannt sind, so haben wir mit

$$\int f(x) dx = c_1 \cdot \ln|x-2| + c_2 \cdot \ln|x+2|$$

die gesuchte Stammfunktion für  $x \neq \pm 2$  gefunden.

Koeffizientenvergleich: Wir können die Terme in der angesetzten Partialbruchzerlegung für  $f$  auf einen gemeinsamen Nenner bringen. Dies liefert

$$\frac{c_1}{x-2} + \frac{c_2}{x+2} = \frac{c_1 \cdot (x+2) + c_2 \cdot (x-2)}{x^2-4} = \frac{(c_1+c_2) \cdot x + 2 \cdot (c_1-c_2)}{x^2-4}$$

und nach einem Koeffizientenvergleich mit dem Zählerpolynom von  $f$  ergibt sich das lineare Gleichungssystem

$$c_1 + c_2 = 1, \quad c_1 - c_2 = -\frac{5}{2}$$

für die gesuchten Koeffizienten, wobei

$$c_1 = -\frac{3}{4}, \quad c_2 = \frac{7}{4}.$$

offensichtlich die einzige Lösung ist.

Einsetzungsmethode: Man kann Gleichungen für die Koeffizienten  $c_j$  auch dadurch gewinnen, dass man den Ansatz und die Funktion  $f$  jeweils an verschiedenen Stellen  $x$  auswertet. Im konkreten Fall muss zum Beispiel

$$-\frac{1}{2} \cdot c_1 + \frac{1}{2} \cdot c_2 = f(0) = \frac{5}{4}, \quad -c_1 + \frac{1}{3} \cdot c_2 = f(1) = \frac{4}{3}$$

gelten, wobei wir der Reihe nach  $x = 0$  und  $x = 1$  gesetzt haben. Wir hätten auch andere Werte wählen können, aber werden am Ende immer dieselben Koeffizienten erhalten.

### 4. Für die rationale Funktion

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

können wir sukzessive eine Partialbruchzerlegung ableiten, in dem wir zunächst solange wie möglich Polynome in additiver Weise abspalten (bzw. eine Polynomdivision mit Rest ausführen):

$$\begin{aligned} f(x) &= \frac{x^4 - x^2}{x^2 - 1} + \frac{x^3 + x^2 + 1}{x^2 - 1} = x^2 + \frac{x^3 - x}{x^2 - 1} + \frac{x^2 + x + 1}{x^2 - 1} \\ &= x^2 + x + \frac{x^2 - 1}{x^2 - 1} + \frac{x + 2}{x^2 - 1} = x^2 + x + 1 + \frac{x + 2}{x^2 - 1}. \end{aligned}$$

Der letzte (d.h. der nicht-polynomiale) Summand auf der rechten Seite ist eine sogenannte echt-gebrochene rationale Funktion, deren Zählerpolynom einen kleineren Grad besitzt als das Nennerpolynom. Dieser Term kann nun analog zu oben weiter aufgespalten werden und wir erhalten

$$f(x) = x^2 + x + 1 + \frac{3}{2} \cdot \frac{1}{x - 1} - \frac{1}{2} \cdot \frac{1}{x + 1}$$

als Partialbruchzerlegung von  $f$ , in der alle Summanden in der obigen Liste der Bausteine vorkommen. Insbesondere erhalten wir

$$\int f(x) dx = \frac{1}{3} \cdot x^3 + \frac{1}{2} \cdot x^2 + x + \frac{3}{2} \cdot \ln|x - 1| - \frac{1}{2} \cdot \ln|x + 1| + C$$

und damit wieder eine explizite Formel für die Stammfunktion von  $f$ .

Bemerkung: Es empfiehlt sich, die Probe zu machen, um Rechenfehler auszuschließen.

### Allgemeines Prinzip der Partialbruchzerlegung (im Reellen)

1. Abspalten eines Polynoms: Jede rationale Funktion  $f(x) = p(x)/q(x)$  kann in eindeutiger Weise als

$$\frac{p(x)}{q(x)} = p_1(x) + \frac{p_2(x)}{q(x)}$$

mit Polynomen  $p_1$  und  $p_2$  geschrieben werden, wobei der Grad von  $p_2$  echt kleiner als der Grad von  $q$  ist. Insbesondere gilt  $p_1 \equiv 0$  (bzw.  $p_1 \equiv c$ ), sofern der Grad von  $p$  kleiner als der (bzw. gleich dem) Grad von  $q$  ist.

2. Zerlegung des Nennerpolynoms: Das Polynom  $q$  kann als Produkt

$$q(x) = d \cdot \prod_{j=1}^m (x - x_j)^{k_j} \cdot \prod_{j=m+1}^{m+n} ((x - a_j)^2 + b_j^2)^{k_j}$$

geschrieben werden, wobei  $d \in \mathbb{R}$  der führende Koeffizient von  $q$  ist. Desweiteren ist  $x_j$  die  $j$ -te reelle Nullstelle mit Vielfachheit  $k_j$  (für  $j = 1 \dots m$ ) und jedes Paar  $(a_j, b_j)$  entspricht dem konjugiert-komplexen Nullstellenpaar  $a_j \pm i \cdot b_j$  mit Vielfachheit  $k_j$  (für  $j = m + 1 \dots n$ ).

3. Zerlegung des echt-gebrochenen rationalen Anteils: Es gibt eindeutige reelle Koeffizienten  $c_{jl}$  sowie  $\gamma_{jl}$  und  $\delta_{jl}$ , so dass

$$\frac{p_2(x)}{q(x)} = \sum_{j=1}^m \sum_{l=1}^{k_j} \frac{c_{jl}}{(x - x_j)^l} + \sum_{j=m+1}^n \sum_{l=1}^{k_j} \frac{\gamma_{jl} \cdot x + \delta_{jl}}{((x - a_j)^2 + b_j^2)^l}.$$

Diese Koeffizienten können als eindeutige Lösung eines linearen Gleichungssystems ermittelt werden.

4. Zusammensetzen der Bausteine: Die Stammfunktion von  $f$  kann nun mit Hilfe der Bausteine einfach angegeben werden. Insbesondere können wir immer explizite Formeln für die Stammfunktion herleiten, obgleich diese sehr kompliziert sein können.

### Bemerkung

1. Wir werden später (Mathe III) eine komplexes Analogon der Partialbruchzerlegung kennen lernen. Dieses wird sogar einfacher als die reelle Variante sein, da man nach dem Hauptsatz der Algebra im Komplexen jeden Baustein vom Typ 3 in Bausteine vom Typ 2 zerlegen kann. Mit anderen Worten:

$$(x - a_j)^2 + b_j^2 = (x - a_j - i \cdot b_j) \cdot (x - a_j + i \cdot b_j)$$

2. Die Aussage des Zerlegungsschrittes für  $q$  kann als reelle Variante des Hauptsatzes der Algebra betrachtet werden. Beachte dabei, dass für jedes Polynom mit reellen Koeffizienten die echt-komplexen Nullstellen (also die mit nicht-verschwindendem Imaginärteil) in Paaren konjugiert-komplexer Zahlen auftreten (das war eine Übungsaufgabe und kann leicht gezeigt werden).
3. Wichtig ist: In der Zerlegung des echt gebrochen-rationalen Anteils  $p_2(x)/q(x)$  ist die Anzahl der Bausteine bzw. die Anzahl der zu bestimmen Konstanten gerade der Polynomgrad von  $q$ . Nicht wichtig ist, wie diese Konstanten bezeichnet werden. Oben haben wir die Symbole  $c_{jl}$ ,  $\gamma_{jl}$  und  $\delta_{jl}$  verwendet, aber in jedem konkreten Fall wird man die Konstanten oftmals einfach durchnummerieren ( $c_1$ ,  $c_2$  usw.).

**Beispiel** Die rationale Funktion

$$f(x) = \frac{1-x}{x^4+x^2} = \frac{1-x}{x^2 \cdot (x^2+1)}$$

ist echt-gebrochen — d.h. der Grad des Zählerpolynoms (nämlich 1) ist kleiner als der Grad des Nennerpolynoms (nämlich 4) — und deshalb existieren Koeffizienten  $c_1, c_2, \gamma, \delta \in \mathbb{R}$ , so dass

$$f(x) = \frac{c_1}{x} + \frac{c_2}{x^2} + \frac{\gamma \cdot x + \delta}{x^2 + 1}$$

gilt. Durch Ausmultiplizieren und Koeffizientenvergleich (oder mittels der Einsetzungsmethode) erhalten wir

$$c_1 = -1, \quad c_2 = 1, \quad \gamma = 1, \quad \delta = -1$$

bzw.

$$f(x) = -\frac{1}{x} + \frac{1}{x^2} + \frac{x}{x^2+1} - \frac{1}{x^2+1}$$

als Partialbruchzerlegung von  $f$ . Insgesamt ergibt sich

$$\begin{aligned} \int f(x) dx &= -\int \frac{1}{x} dx + \int \frac{1}{x^2} dx + \int \frac{x}{x^2+1} dx - \int \frac{1}{x^2+1} dx \\ &= -\ln|x| - \frac{1}{x} + \frac{1}{2} \cdot \ln(x^2+1) - \arctan(x) + C \end{aligned}$$

als Formel für die Stammfunktion.

Bemerkung: Machen Sie bei komplizierten Partialbruchzerlegungen immer die Probe! Entweder per Hand oder mit MATHEMATICA.



**Beispiele**

1. Entsteht
- $f$
- aus

$$p(x) = -x^2 + 20 \cdot x + 149, \quad q(x) = x^3 + 4 \cdot x^2 - 11 \cdot x - 30,$$

so gibt es wieder keinen polynomiellen Anteil, d.h. wir haben  $p_1 = 0$ ,  $p_2 = p$  und der erste Schritt entfällt. Außerdem gilt

$$q(x) = (x - 3) \cdot (x + 2) \cdot (x + 5)$$

und deshalb reduziert sich die Partialbruchzerlegung von  $f$  auf

$$f(x) = \frac{c_1}{x - 3} + \frac{c_2}{x + 2} + \frac{c_3}{x + 5},$$

wobei wir nun die entsprechenden Koeffizienten  $c_1, c_2, c_3 \in \mathbb{R}$  bestimmen müssen. Durch Koeffizientenvergleich (oder alternativ mit Hilfe der Einsetzungsmethode) berechnen wir

$$c_1 = 5, \quad c_2 = -7, \quad c_3 = 1$$

und erhalten

$$\begin{aligned} \int f(x) dx &= 5 \cdot \ln|x - 3| - 7 \cdot \ln|x + 2| + \ln|x + 5| + C \\ &= \ln \left| \frac{(x - 3)^5 \cdot (x + 5)}{(x + 2)^7} \right| + C. \end{aligned}$$

2. Zu den Polynomen

$$p(x) = 5 \cdot x^2 - 37 \cdot x + 54, \quad q(x) = x^3 - 6 \cdot x^2 + 9 \cdot x = x \cdot (x - 3)^2$$

gehört die Partialbruchzerlegung

$$f(x) = \frac{p(x)}{q(x)} = \frac{6}{x} - \frac{1}{x - 3} - \frac{4}{(x - 3)^2}$$

sowie die Stammfunktion

$$\int f(x) dx = 6 \cdot \ln|x| - \ln|x - 3| + \frac{4}{x - 3} + C.$$

3. Im Fall von

$$p(x) = x^4 - 2 \cdot x^3 + 5 \cdot x^2 - 2 \cdot x + 2$$

und

$$q(x) = x^3 - x^2 + x - 1 = (x - 1) \cdot (x^2 + 1)$$

erhalten wir mittels Polynomdivision mit Rest im ersten Schritt zunächst

$$f(x) = x - 1 + \frac{3 \cdot x^2 + 1}{(x - 1) \cdot (x^2 + 1)}$$

und anschließend die Partialbruchzerlegung

$$f(x) = x - 1 + \frac{2}{x - 1} + \frac{x + 1}{x^2 + 1}.$$

Insgesamt ergibt sich damit

$$\int f(x) dx = \frac{1}{2} \cdot x^2 - x + 2 \cdot \ln|x - 1| + \arctan(x) + \frac{1}{2} \cdot \ln(x^2 + 1) + C,$$

wobei natürlich  $x \neq 1$  gelten muss.

## Uneigentliche Riemann-Integrale

Vorlesung 33, 24. Januar 2020

**Vorbemerkung** Bisher hatten wir bei der Berechnung von bestimmten Integralen (d.h. des Flächeninhaltes unter dem Graphen) vorausgesetzt, dass der Integrand  $f$  aus einem kompakten Intervall  $[a, b]$  definiert ist. Wir wollen nun die folgenden, allgemeineren Fälle studieren:

1. halb-offene Intervalle wie  $[a, b]$  und  $(a, b]$  bzw.  $[a, +\infty)$  und  $(-\infty, b]$ ,
2. offene Intervalle wie  $(a, b)$ ,  $(-\infty, +\infty)$  oder auch  $(-\infty, b)$ ,  $(a, +\infty)$ ,
3. isolierte Lücken im Definitionsbereich.

### Uneigentliche Integrale auf halboffenen und halbunendlichen Intervallen

Ist  $f$  auf  $[a, \infty)$  bzw.  $(-\infty, b]$  definiert und (stückweise) stetig, so nennt man

$$\int_a^{+\infty} f(x) \, dx := \lim_{\beta \nearrow +\infty} \int_a^{\beta} f(x) \, dx \quad \text{bzw.} \quad \int_{-\infty}^b f(x) \, dx := \lim_{\alpha \searrow -\infty} \int_{\alpha}^b f(x) \, dx$$

das uneigentliche Riemann-Integral von  $f$ , sofern der entsprechende Grenzwert im eigentlichen Sinne existiert (d.h. wenn er wohldefiniert und reell ist).

### Beispiele

1. Die Funktion  $f(x) = x^p$  ist für  $p < -1$  und jedes  $a > 0$  auf  $[a, +\infty)$  integrierbar mit

$$\int_a^{+\infty} x^p \, dx = -\frac{a^{p+1}}{p+1}$$

denn es gilt

$$\int_a^{\beta} x^p \, dx = \left[ \frac{x^{p+1}}{p+1} \right]_{x=a}^{x=\beta} = \frac{\beta^{p+1} - a^{p+1}}{p+1} \xrightarrow{\beta \nearrow +\infty} -\frac{a^{p+1}}{p+1},$$

wobei der Grenzwert wohldefiniert ist und einen Wert in  $(0, \infty)$  annimmt.

Achtung: Für  $p \geq -1$  ist die Funktion aber nicht integrierbar. Für  $p > -1$  folgt dies bereits aus obigen Formeln, da  $\lim_{\beta \nearrow +\infty} \int_a^{\beta} x^p \, dx = +\infty$  gilt, und für  $p = -1$  berechnen wir

$$\int_a^{\beta} x^{-1} \, dx = \left[ \ln(x) \right]_{x=a}^{x=\beta} = \ln(\beta) - \ln(a) \xrightarrow{\beta \nearrow +\infty} +\infty.$$

Heuristisch können diese Ergebnisse wie folgt verstanden werden: Nur für  $p < -1$  klingt die Funktion im Unendlichen so schnell ab, dass der Flächeninhalt unter dem Graphen immer noch endlich ist. Für  $-1 \leq p < 0$  ist das Abklingen jedoch

vergleichsweise schwach und für  $p \geq 0$  klingt die Funktion gar nicht ab.

*Bemerkung:* Da  $f$  keine Vorzeichenwechsel aufweist, schreibt man oftmals auch  $\int_a^\infty x^p dx = +\infty$  für  $p \geq -1$ , aber dies ist streng genommen nur eine Abkürzung und gilt nicht im Sinne der obigen Definition uneigentlicher Integrale.

2. Analog zu oben zeigt man

$$\lim_{\beta \nearrow +\infty} \int_a^\beta \frac{dx}{x \cdot \ln(x)} = \lim_{\beta \nearrow +\infty} \left( \ln(\ln \beta) - \ln(\ln 2) \right) = +\infty,$$

d.h.  $1/(x \cdot \ln x)$  klingt für  $x \rightarrow \infty$  zwar schon schneller ab als  $1/x$ , aber dieses schnellere Abklingen reicht noch nicht, um die Integrierbarkeit sicherzustellen.

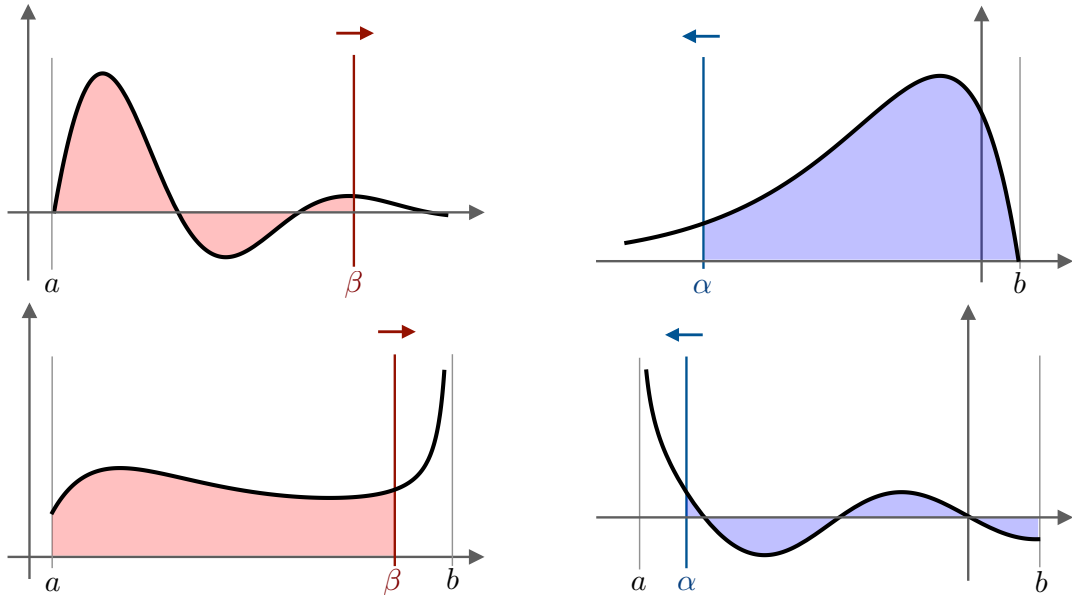
3. Für jedes  $a \in \mathbb{R}$  gilt

$$\int_a^{+\infty} x \cdot \exp(-x^2) dx = \frac{1}{2} \cdot e^{-a^2},$$

denn mit der Substitution  $s = x^2$  ergibt sich  $ds = 2 \cdot x \cdot dx$  sowie

$$\int_a^\beta x \cdot e^{-x^2} dx = \int_{a^2}^{\beta^2} \frac{1}{2} \cdot e^{-s} ds = \left[ -\frac{1}{2} \cdot e^{-s} \right]_{s=a^2}^{s=\beta^2} = \frac{1}{2} \cdot (e^{-a^2} - e^{-\beta^2})$$

und der Grenzübergang  $\beta \nearrow +\infty$  liefert die Behauptung.



Zum uneigentlichen Integralbegriff auf den halboffenen Intervallen  $[a, +\infty)$  und  $(-\infty, b]$  (oben) bzw.  $[a, b)$  und  $(a, b]$  (unten)

**Uneigentliche Integrale auf halboffenen beschränkten Intervallen** Ist  $f$  stetig in  $(a, b]$  bzw.  $[a, b)$  (aber vielleicht nicht definiert in  $a$  bzw.  $b$ ), so setzt man

$$\int_a^b f(x) dx := \int_a^{\beta \nearrow b} f(x) dx \quad \text{bzw.} \quad \int_a^b f(x) dx := \int_{\alpha \searrow a}^b f(x) dx$$

sofern der jeweilige Grenzwert existiert und reell ist. Beachte, dass die bereits studierten halbunendlichen Intervalle den Grenzfällen  $a = -\infty$  oder  $b = +\infty$  entsprechen.

### Beispiele

1. Für jedes  $b > 0$  gilt

$$\int_{\alpha}^b x^p dx = \left[ \frac{b^{p+1} - \alpha^{p+1}}{p+1} \right] \xrightarrow{\alpha \searrow 0} \begin{cases} \frac{b^{p+1}}{p+1} & \text{für } p > -1 \\ \infty & \text{für } p \leq -1 \end{cases},$$

d.h. die Funktion  $f(x) = x^p$  ist nur für  $p > -1$  auf dem Intervall  $(0, b)$  uneigentlich Riemann-integrierbar.

2. Die Funktion  $f(x) = x/(1-x^2)$  ist auf dem Intervall  $[0, +1)$  nicht uneigentlich integrierbar, denn es gilt

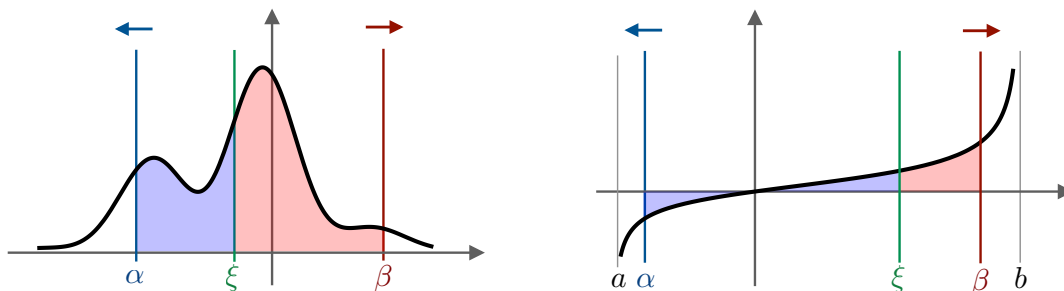
$$\int_0^{\beta} \frac{x \cdot dx}{1-x^2} = \left[ -\frac{1}{2} \cdot \ln(1-x^2) \right]_{x=0}^{x=\beta} - \frac{1}{2} \cdot \ln(1-x^2) \xrightarrow{\beta \nearrow 1} +\infty,$$

d.h. der Grenzwert existiert nur im uneigentlichen Sinne.

**Uneigentliche Integrale auf offenen Intervallen** Ist  $f$  stetig in  $(a, b)$ , so definiert man

$$\int_a^b f(x) dx := \lim_{\alpha \searrow a} \int_{\alpha}^{\xi} f(x) dx + \lim_{\beta \nearrow b} \int_{\xi}^{\beta} f(x) dx$$

sofern jeder der beiden Grenzwerte im eigentlichen Sinne existiert und  $\xi$  eine beliebige Zwischenstelle mit  $a < \xi < b$  ist. Diese Formel kann sowohl für endliche als auch für unendliche Werte von  $a \in [-\infty, +\infty)$  und  $b \in (a, +\infty)$  verwendet werden. Beachte außerdem, dass sowohl die Existenz beider Grenzwerte als auch der Wert ihrer Summe nicht von der Wahl von  $\xi$  abhängen (Übungsaufgabe).



Uneigentliche Integrale auf den offenen Intervallen  $(-\infty, \infty)$  und  $(a, b)$  erfordern die Berechnung zweier unabhängiger Grenzwerte (rot und blau), wobei die Zwischenstelle  $\xi$  (an der das Intervall zerteilt wird) beliebig gewählt werden darf.

Analoges gilt für  $(-\infty, b)$  und  $(a, +\infty)$ .

**Beispiele**

1. Die Funktion  $f(x) = x^p$  ist für keinen Wert des Parameters  $p$  auf dem Intervall  $(0, \infty)$  definiert, denn für jede Wahl von  $\xi$  gilt

$$\lim_{\alpha \searrow 0} \int_{\alpha}^{\xi} x^p dx = \infty \quad \text{oder} \quad \lim_{\beta \nearrow \infty} \int_{\xi}^{\beta} x^p dx = \infty,$$

je nachdem ob  $p \geq -1$  oder  $p \leq -1$  (siehe auch das Bild unten). Insbesondere existieren für keinen Wert des Parameters  $p$  beide Integrale im eigentlichen Sinne.

2. Es gilt

$$\begin{aligned} \int_{-\infty}^{\infty} e^{-x^2} dx &= \lim_{\alpha \searrow -\infty} \int_{\alpha}^0 e^{-x^2} dx + \lim_{\beta \nearrow +\infty} \int_0^{\beta} e^{-x^2} dx \\ &= \lim_{\beta \nearrow \infty} \frac{1}{2} \cdot \sqrt{\pi} \cdot \operatorname{Erf}(\beta) - \lim_{\alpha \searrow -\infty} \frac{1}{2} \cdot \sqrt{\pi} \cdot \operatorname{Erf}(\alpha) \\ &= \frac{1}{2} \cdot \sqrt{\pi} \cdot (\operatorname{Erf}(+\infty) - \operatorname{Erf}(-\infty)) = \sqrt{\pi}, \end{aligned}$$

wobei wir die verwendeten asymptotischen Eigenschaften der Fehlerfunktion  $\operatorname{Erf}$  erst später verstehen werden. Einfacher zu Rechnen ist der Fall

$$\begin{aligned} \int_{-\infty}^{+\infty} x \cdot e^{-x^2} dx &= \lim_{\alpha \searrow -\infty} \int_{\alpha}^0 x \cdot e^{-x^2} dx + \lim_{\beta \nearrow +\infty} \int_0^{\beta} x \cdot e^{-x^2} dx \\ &= \lim_{\beta \nearrow +\infty} \frac{1}{2} \cdot (1 - e^{-\beta^2}) - \lim_{\alpha \searrow -\infty} \frac{1}{2} \cdot (1 - e^{-\alpha^2}) = 0. \end{aligned}$$

3. Es gilt

$$\begin{aligned} \int_{-1}^{+1} \frac{dx}{\sqrt{1-x^2}} &= \lim_{\alpha \nearrow -1} \int_{\alpha}^0 \frac{dx}{\sqrt{1-x^2}} + \lim_{\beta \nearrow +1} \int_0^{\beta} \frac{dx}{\sqrt{1-x^2}} \\ &= \lim_{\alpha \nearrow -1} (\arcsin(0) - \arcsin(\alpha)) + \lim_{\beta \nearrow +1} (\arcsin(\beta) - \arcsin(0)) \\ &= \arcsin(+1) - \arcsin(-1) = \pi \end{aligned}$$

im Sinne uneigentlicher Integrale. Das ist nicht überraschend, denn ein Taylor-Argument (siehe unten für die Details) offenbart

$$\frac{1}{\sqrt{1-x^2}} = \frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{1-x}} (1 + O(1-x)) \quad \text{für } x \lesssim +1$$

sowie

$$\frac{1}{\sqrt{1-x^2}} = \frac{1}{\sqrt{2}} \cdot \frac{1}{\sqrt{1+x}} (1 + O(1+x)) \quad \text{für } x \gtrsim -1,$$

d.h. der Integrand besitzt an den Intervallrändern  $x = -1$  und  $x = +1$  integrierbare Singularitäten. Siehe dazu auch das allgemeine Prinzip weiter unten.

4. Für die Funktion  $f : \mathbb{R} \rightarrow \mathbb{R}$  mit

$$f(x) = \frac{x}{1+x^2}$$

berechnen wir

$$\int_{\alpha}^{\beta} f(x) dx = \left[ \frac{1}{2} \cdot \ln(1+x^2) \right]_{x=\alpha}^{x=\beta} = \frac{1}{2} \cdot \ln \left( \frac{1+\beta^2}{1+\alpha^2} \right).$$

Dies impliziert

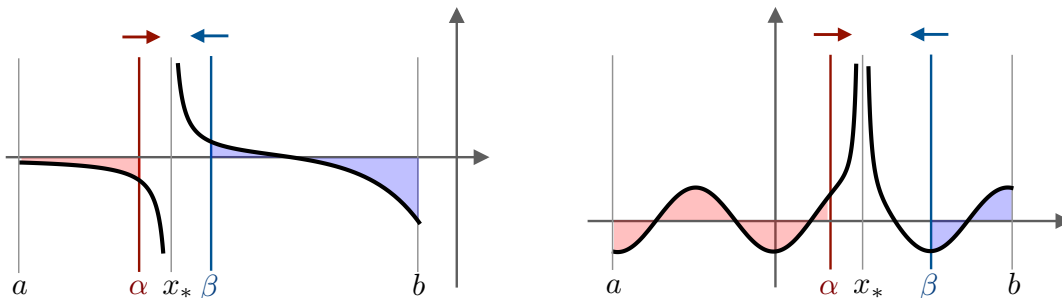
$$\int_{\alpha}^{\xi} f(x) dx \xrightarrow{\alpha \searrow -\infty} -\infty, \quad \int_{\xi}^{\beta} f(x) dx \xrightarrow{\alpha \nearrow \infty} \infty$$

für jedes feste  $\xi \in \mathbb{R}$  und wir schließen, dass  $f$  nicht uneigentlich integrierbar auf  $\mathbb{R}$  ist.

**Uneigentliche Integrale bei isolierten Lücken im Definitionsbereich** Ist  $f$  wohldefiniert auf  $[a, x_*) \cup (x_*, b]$ , so bezeichnet man

$$\int_a^b f(x) dx := \lim_{\alpha \nearrow x_*} \int_a^{\alpha} f(x) dx + \lim_{\beta \searrow x_*} \int_{\beta}^b f(x) dx$$

als das entsprechende uneigentliche Integral, sofern die Grenzwerte auf der rechten Seite beide wohldefiniert sind.



Auch bei isolierten Lücken im Definitionsbereich wird das uneigentliche Integral mit Hilfe zweier unabhängiger Grenzübergänge (wieder rot und blau) definiert.

### Beispiele

1. Für jedes  $p \neq -1$  betrachten wir die Funktion  $f(x) = (x - x_*)^p$  und wegen

$$\int_a^{\alpha} (x - x_*)^p dx = \frac{(\alpha - x_*)^{p+1} - (a - x_*)^{p+1}}{p+1}$$

und

$$\int_{\beta}^b (x - x_*)^p dx = \frac{(b - x_*)^{p+1} - (\beta - x_*)^{p+1}}{p+1}$$

schließen wir, dass  $f$  für  $p > -1$  uneigentlich integrierbar ist mit

$$\int_a^b f(x) dx = \frac{(b - x_*)^{p+1} - (a - x_*)^{p+1}}{p + 1}.$$

Für  $p \leq -1$  gilt dies aber nicht, da die zu betrachtenden Grenzwerte nur im uneigentlichen, aber nicht im eigentlichen Sinn existieren. Beachte, dass der Fall  $p = -1$  eine leicht andere Rechnungen erfordert.

2. Die Funktion  $|f|(x) = |x - x_*|^p$  ist auch nur für  $p > -1$  uneigentlich integrierbar, wobei diesmal

$$\int_a^b |f|(x) dx = \frac{(b - x_*)^{p+1} + (a - x_*)^{p+1}}{p + 1}$$

gilt.

**\*Allgemeines Prinzip für Lücken und Randpunkte** Kann eine stetige Funktion  $f$  in der Nähe einer isolierten Lücke  $x_*$  ihres Definitionsbereichs durch

$$f(x) = \begin{cases} c_- \cdot (x_* - x)^{p_-} (1 + O(x_* - x)) & \text{für } x < x_* \\ c_+ \cdot (x - x_*)^{p_+} (1 + O(x - x_*)) & \text{für } x > x_* \end{cases}$$

approximiert werden, so ist  $f$  in der Nähe von  $x_*$  genau dann uneigentlich integrierbar, wenn die beiden Exponenten den Bedingungen  $p_- > -1$  und  $p_+ > -1$  genügen. Die einseitigen Analoga gelten auch, wenn  $x_*$  äußerer Randpunkt des Definitionsbereichs ist.

### Beispiel

1. Die Funktion

$$f(x) = \frac{1}{\sqrt{1 - x^2}}$$

ist wohldefiniert auf  $(-1, +1)$  und wir betrachten nun das Verhalten für positive aber kleine  $\varepsilon = 1 - x$ . Dies liefert

$$\begin{aligned} f(x) = f(1 - \varepsilon) &= \frac{1}{\sqrt{1 - (1 - \varepsilon)^2}} = \frac{1}{\sqrt{2 \cdot \varepsilon - \varepsilon^2}} \\ &= \frac{1}{\sqrt{2} \cdot \sqrt{\varepsilon}} \cdot \frac{1}{\sqrt{1 - \frac{1}{2} \cdot \varepsilon}} = \frac{1}{\sqrt{2} \cdot \sqrt{\varepsilon}} \cdot \left(1 + \frac{1}{4} \cdot \varepsilon + O(\varepsilon^2)\right) \\ &= \frac{1}{\sqrt{2} \cdot \sqrt{1 - x}} \left(1 + \frac{1}{4} \cdot (1 - x) + O(1 - x)^2\right). \end{aligned}$$

Wir haben hier benutzt, dass

$$g(\varepsilon) := \frac{1}{\sqrt{1 - \frac{1}{2}\varepsilon}} = g(0) + g'(0) \cdot \varepsilon + \dots$$

nach dem Satz von Taylor gilt, wobei  $g(0) = 1$  und  $g'(0) = \frac{1}{4}$ .

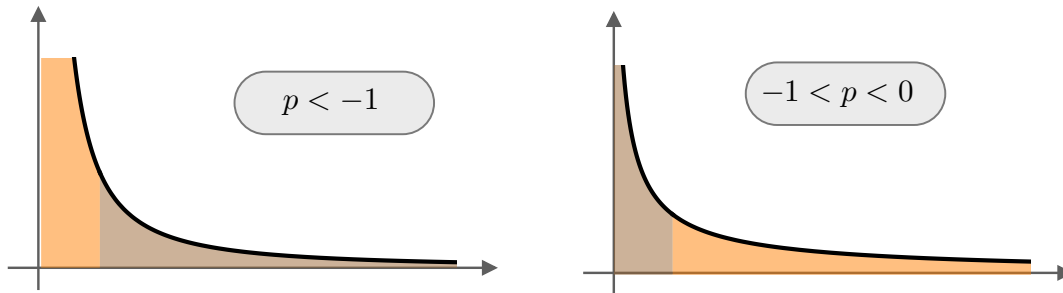
## 2. Die rationale Funktion

$$f(x) = \frac{x^3 + 3}{x - 1}$$

besitzt einen Pol bei  $x_* = 1$  und ist auf keinem Intervall  $[a, b]$  mit  $a < 1 < b$  uneigentlich integrierbar. Mit  $x = x_* + \varepsilon$  gilt nämlich

$$f(x) = f(1 + \varepsilon) = \frac{1}{\varepsilon} \cdot \left( (1 + \varepsilon)^3 + 3 \right) = \frac{4}{\varepsilon} \cdot \left( 1 + O(\varepsilon) \right),$$

wobei  $\varepsilon$  hier sowohl positiv als auch negativ sein kann.



Zur uneigentlichen Integrierbarkeit von  $f(x) = x^p$  mit  $p < 0$  auf  $(0, \infty)$ , wobei die braunen bzw. gelben Flächen endlich bzw. unendlich sind. Für  $p = -1$  sind beide Flächen gelb.

**\*Allgemeines Prinzip für das Abkling-Verhalten** Kann das Abklingverhalten einer stetigen Funktion auf  $\mathbb{R}$  durch

$$f(x) = \begin{cases} c_- \cdot |x|^{p_-} \left( 1 + O(|1/x|) \right) & \text{für } x \ll -1 \\ c_+ \cdot |x|^{p_+} \left( 1 + O(|1/x|) \right) & \text{für } x \gg +1 \end{cases}$$

beschrieben werden, so ist  $f$  genau dann uneigentlich integrierbar auf  $\mathbb{R}$ , wenn  $p_- < -1$  und  $p_+ < -1$  gilt.

**\*Cauchysche Hauptwerte** Es gibt ein schwächeres Konzept als Uneigentliche Riemann-Integrale, nämlich die sogenannten Hauptwerte. Wir wollen dieses Konzept hier nur streifen, indem wir zwei typische Beispiele studieren.

Die stetige Funktion  $f : \mathbb{R} \rightarrow \mathbb{R}$  mit  $f(x) = (x + 2)/(x^2 + 1)$  ist nicht uneigentlich integrierbar auf  $(-\infty, \infty)$ , eben weil die Funktion im Unendlichen zu langsam abklingt und daher die Grenzwerte

$$\lim_{\alpha \searrow -\infty} \int_{\alpha}^{\xi} f(x) dx = -\infty, \quad \lim_{\beta \nearrow +\infty} \int_{\xi}^{\beta} f(x) dx = +\infty$$

nicht im eigentlichen, sondern nur im uneigentlichen Sinne existieren. Es gilt aber

$$\begin{aligned} \lim_{\zeta \nearrow +\infty} \int_{-\zeta}^{+\zeta} f(x) dx &= \lim_{\zeta \nearrow +\infty} \left( \int_{-\zeta}^{+\zeta} \frac{x}{x^2 + 1} dx + \int_{-\zeta}^{+\zeta} \frac{2}{x^2 + 1} dx \right) \\ &= \lim_{\zeta \nearrow +\infty} \left( \left[ \frac{1}{2} \cdot \ln(1 + x^2) \right]_{x=-\zeta}^{x=+\zeta} + \left[ 2 \cdot \arctan(x) \right]_{x=-\zeta}^{x=+\zeta} \right) \\ &= 2 \cdot \lim_{\zeta \nearrow +\infty} \left( \arctan(+\zeta) - \arctan(-\zeta) \right) = 2 \cdot \pi, \end{aligned}$$



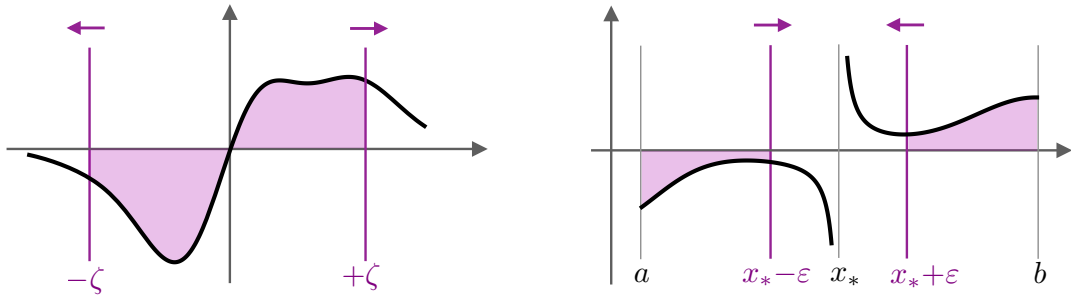
wobei der Grenzwert auf der linken Seite der Cauchysche Hauptwert von  $f$  genannt wird.

Analog ist die Funktion

$$f(x) = \frac{x+1}{x} = 1 + \frac{1}{x}$$

nicht uneigentlich integrierbar auf  $[-1, 0) \cup (0, 1]$ , aber es existiert wieder ein Hauptwert im Sinne von

$$\begin{aligned} \lim_{\varepsilon \searrow 0} \left( \int_{-1}^{-\varepsilon} f(x) dx + \int_{+\varepsilon}^{+1} f(x) dx \right) &= \lim_{\varepsilon \searrow 0} \left( \left[ x + \ln |x| \right]_{x=-1}^{x=-\varepsilon} + \left[ x + \ln |x| \right]_{x=+\varepsilon}^{x=+1} \right) \\ &= \lim_{\varepsilon \searrow 0} \left( (-\varepsilon + \ln \varepsilon + 1) + (1 - \ln \varepsilon - \varepsilon) \right) \\ &= \lim_{\varepsilon \searrow 0} (2 - 2 \cdot \varepsilon) = 2. \end{aligned}$$



Beim Cauchyschen Hauptwert wird statt zweier separater Grenzwerte nur ein Limes betrachtet. Dieses Konzept ist schwächer als das uneigentliche Integral, da es Auslöschungseffekte geben kann, bei denen sich positive und negative Beiträge gegenseitig aufheben.

## Parametrisierte Integrale

Vorlesung 34, 27. Januar 2020

**Motivation** In vielen Anwendungen tauchen bestimmte Integrale auf, bei denen die Integrationsgrenzen und/oder der Integrand von einem reellen Parameter  $p \in \mathbb{R}$  abhängen. Will man nun verstehen, wie sich der Wert des Integrals  $I(p)$  unter kleinen Störungen  $\Delta p$  des Parameters  $p$  verändert, so muss man die Ableitung von  $I$  nach  $p$  studieren, denn es gilt

$$I(p + \Delta p) = I(p) + \frac{dI}{dp}(p) \cdot \Delta p + O((\Delta p)^2) \approx I(p) + \frac{dI}{dp}(p) \cdot \Delta p$$

nach dem Satz von Taylor.

**Parameterabhängige Integrationsgrenzen** Es gilt

$$\frac{d}{dp} \int_{a(p)}^{b(p)} f(x) dx = f(b(p)) \cdot \frac{db}{dp}(p) - f(a(p)) \cdot \frac{da}{dp}(p),$$

sofern die Integralgrenzen in differenzierbarer Weise von  $p$  abhängen. Dieses Gesetz kann auf verschiedene Weisen abgeleitet werden.

*Herleitung 1:* Ist  $F$  eine Stammfunktion zu  $f$  (d.h.  $F' = f$ ), so impliziert der Hauptsatz der Differential- und Integralrechnung

$$I(p) = F(b(p)) - F(a(p)).$$

Die Formel für  $\frac{d}{dp}I(p)$  ergibt sich nun durch Differentiation nach  $p$  und Anwendung der Kettenregel.

*Herleitung 2:* Man kann auch Differenzenquotienten benutzen: Die Eigenschaften des Integrals implizieren für jedes  $\Delta p$  die Formel

$$\begin{aligned} \frac{I(p + \Delta p) - I(p)}{\Delta p} &= \frac{1}{\Delta p} \int_{b(p)}^{b(p+\Delta p)} f(x) dx - \frac{1}{\Delta p} \int_{a(p)}^{a(p+\Delta p)} f(x) dx \\ &= \frac{\Delta b}{\Delta p} \cdot \frac{1}{\Delta b} \int_{b(p)}^{b(p)+\Delta b} f(x) dx - \frac{\Delta a}{\Delta p} \cdot \frac{1}{\Delta a} \int_{a(p)}^{a(p)+\Delta a} f(x) dx, \end{aligned}$$

wobei

$$\Delta a := a(p + \Delta p) - a(p) \xrightarrow{\Delta p \rightarrow 0} 0, \quad \Delta b := b(p + \Delta p) - b(p) \xrightarrow{\Delta p \rightarrow 0} 0.$$

Die gewünschte Formel für  $\frac{d}{dp}I(p)$  ergibt sich nun nach Grenzübergang  $\Delta p \rightarrow 0$ , wobei die Differenzierbarkeit der Funktionen  $a$  und  $b$ , die Stetigkeit von  $f$  und die Mittelwert-eigenschaft der Integration in Form von  $\lim_{h \rightarrow 0} (h^{-1} \int_{x_*}^{x_*+h} f(x) dx) = f(x_*)$  verwendet werden.

*Herleitung 3:* Man kann das höherdimensionale Analogon zur Kettenregel verwenden (siehe *Mathe II*).

### Beispiele

1. Die obige Formel liefert für das parametrisierte Integral

$$I(p) = \int_0^{\sin(p)} x^2 dx$$

die Ableitung

$$\frac{d}{dp}I(p) = \sin^2(p) \cdot \cos(p).$$

Alternativ können wir aber auch zunächst  $I(p)$  direkt berechnen und anschließend nach  $p$  differenzieren. Dies liefert aber wegen

$$I(p) = \left[ \frac{1}{3} \cdot x^3 \right]_{x=0}^{x=\sin(p)} = \frac{1}{3} \cdot (\sin(p))^3$$

das gleiche Endergebnis und wir haben die abstrakte Formel für den konkreten Fall nachgewiesen.

2. Für

$$I(p) = \int_{\exp(-p)}^1 x^4 dx$$

erhalten wir mit bzw. ohne Verwendung der Formel

$$\frac{d}{dp} I(p) = -(\exp(-p))^4 \cdot (-\exp(-p)) = \exp(-5 \cdot p),$$

bzw.

$$\begin{aligned} \frac{d}{dp} I(p) &= \frac{d}{dp} \left( \left[ \frac{1}{5} \cdot x^5 \right]_{x=\exp(-p)}^{x=1} \right) \\ &= \frac{d}{dp} \left( \frac{1}{5} \cdot \exp(-1) - \frac{1}{5} \cdot \exp(-5 \cdot p) \right) = \exp(-5 \cdot p). \end{aligned}$$

Insbesondere liefern wieder beide Rechenwege dasselbe Ergebnis.

3. Die Formel impliziert

$$\frac{d}{dp} \int_{p^2}^{p^3} \sin(-x^3) dx = \sin(-p^9) \cdot (3 \cdot p^6) - \sin(-p^4) \cdot (2 \cdot p),$$

aber in diesem Beispiel gibt es keine echte Alternative, da wir die Stammfunktion des Integranden nicht kennen.

**Parameterabhängige Integranden** Hängt jedoch der Integrand, aber nicht die Integrationsgrenzen von  $p$  ab, so gilt

$$\frac{d}{dp} \int_a^b g(x, p) dx = \int_a^b \frac{\partial}{\partial p} g(x, p) dx,$$

wobei

$$\frac{\partial}{\partial p} g(x, p) = \lim_{h \rightarrow 0} \frac{g(x, p+h) - g(x, p)}{h}$$

die sogenannte partielle Ableitung von  $g$  nach  $p$  ist (die wir eigentlich erst in *Mathe II* studieren) und wir annehmen, dass die Funktion  $g$  in hinreichend regulärer Weise von den beiden Argumenten  $x$  und  $p$  abhängt (stetige Differenzierbarkeit ist zum Beispiel ausreichend). Diese Formel kann zum Beispiel wieder mit Differenzenquotienten begründet werden.

### Beispiele

1. Für das parametrisierte Integral

$$I(p) := \int_0^1 \exp(p \cdot x) dx$$

folgt aus der Formel sowie anschließender partieller Integration

$$\begin{aligned}\frac{d}{dp}I(p) &= \int_0^1 x \cdot \exp(p \cdot x) dx = \left[ x \cdot \frac{\exp(p \cdot x)}{p} \right]_{x=0}^{x=1} - \int_0^1 1 \cdot \frac{\exp(p \cdot x)}{p} dx \\ &= \left[ x \cdot \frac{\exp(p \cdot x)}{p} \right]_{x=0}^{x=1} - \left[ \frac{\exp(p \cdot x)}{p^2} \right]_{x=0}^{x=1} = \left( \frac{1}{p} - \frac{1}{p^2} \right) \cdot \exp(p) + \frac{1}{p^2}.\end{aligned}$$

Alternativ können wir via

$$\frac{d}{dp}I(p) = \frac{d}{dp} \left( \left[ \frac{\exp(p \cdot x)}{p} \right]_{x=0}^{x=1} \right) = \frac{d}{dp} \left( \frac{\exp(p)}{p} - \frac{1}{p} \right)$$

dasselbe Ergebnis direkt herleiten (und damit die Formel in einem Spezialfall beweisen).

2. Ein weiteres einfaches Beispiel ist

$$\frac{d}{dp} \int_1^\pi \frac{\sin(px)}{x} dx = \int_1^\pi \cos(px) dx = \left[ \frac{-\sin(px)}{p} \right]_{x=1}^{x=\pi} = \frac{\sin(p) - \sin(p\pi)}{p}.$$

3. Für die sogenannte  $n$ -te Besselfunktion (die später wichtig werden wird)

$$J_n(x) := \frac{1}{\pi} \int_0^\pi \cos(x \cdot \sin(t) - n \cdot t) dt$$

erhalten wir

$$J'_n(x) = \frac{d}{dx} J_n(x) = -\frac{1}{\pi} \int_0^\pi \sin(x \cdot \sin(t) - n \cdot t) \cdot \sin(t) dt$$

sowie

$$J''_n(x) = \frac{d}{dx} J'_n(x) = -\frac{1}{\pi} \int_0^\pi \cos(x \cdot \sin(t) - n \cdot t) \cdot \sin^2(t) dt.$$

Beachte, dass hier  $t$  die Integrationsvariable und  $x$  der Parameter im Integranden ist.

**Bemerkung** Bestimmte Integrale können nicht nach der Integrationsvariablen differenziert werden, d.h. die Ausdrücke

$$\frac{d}{dx} \int_a^b f(x) dx, \quad \frac{d}{dx} \int_a^b g(x, p) dx$$

haben **keinen Sinn**, eben weil das Integral für jede Funktion  $f$  bzw.  $g$  eine reelle Zahl ist, die im zweiten Fall zwar von  $p$ , aber in keinen Fall von  $x$  abhängt. Bei unbestimmten

Integralen ist das anders. Insbesondere impliziert der Hauptsatz der Differential- und Integralrechnung die Formel

$$\frac{d}{dx} \int f(x) dx = f(x), \quad \frac{\partial}{\partial x} \int g(x, p) dx = g(x, p),$$

sofern  $f$  und  $g$  hinreichend regulär sin. Außerdem gilt

$$\frac{\partial}{\partial p} \int g(x, p) dx = \int \frac{\partial}{\partial p} g(x, p) dx,$$

eben weil  $p$  nicht die Integrationsvariable ist.

**Allgemeiner Fall** Oftmals hängen sowohl die Integralgrenzen als auch der Integrand von einem Parameter ab. Dann gilt

$$\frac{d}{dp} \int_{a(p)}^{b(p)} g(x, p) dx = \int_{a(p)}^{b(p)} \frac{\partial}{\partial p} g(x, p) dx + g(b(p), p) \cdot \frac{d}{dp} b(p) - g(a(p), p) \cdot \frac{d}{dp} a(p),$$

sofern alle involvierten Funktionen hinreichend regulär sind.

**Beispiel** Für das parameterabhängige Integral

$$I(p) := \int_p^{e^p} (p \cdot x)^2 dx \quad \text{für } p > 0$$

ergibt sich unter Verwendung der Formel

$$\begin{aligned} \frac{d}{dp} I(p) &= \int_p^{e^p} 2 \cdot p \cdot x^2 dx + (p \cdot e^p)^2 \cdot e^p - (p \cdot p)^2 \cdot 1 \\ &= \left[ \frac{2}{3} \cdot p \cdot x^3 \right]_{x=p}^{x=e^p} + p^2 \cdot e^{3 \cdot p} - p^4 = \left( \frac{2}{3} \cdot p + p^2 \right) \cdot e^{3 \cdot p} - \frac{5}{3} \cdot p^4, \end{aligned}$$

wobei die Terme weiter vereinfacht werden können. Alternativ kann man auch hier zunächst  $I(p)$  durch die Stammfunktion des Integranden (als Funktion von  $x$ ) ausdrücken und anschließend das Ergebnis nach  $p$  differenzieren. Mit anderen Worten, es gilt

$$I(p) = \left[ \frac{1}{3} \cdot p^2 \cdot x^3 \right]_{x=p}^{x=e^p} = \frac{1}{3} \cdot p^2 \cdot e^{3 \cdot p} - \frac{1}{3} \cdot p^5$$

und wir können nun direkt nach  $p$  differenzieren.

**Gamma-Funktion** Es gibt natürlich auch uneigentliche parametrisierte Integrale und wir wollen hier ein wichtiges Beispiel besprechen. Die sogenannte  $\Gamma$ -Funktion ist definiert durch

$$\Gamma(x) = \int_0^{\infty} t^{x-1} \cdot e^{-t} dt$$

wobei hier wieder  $x$  der Parameter und  $t$  die Integrationsvariable ist. Diese Funktion ist wohldefiniert (weil der Integrand für jedes festes  $x$  und  $t \rightarrow \infty$  sehr schnell abklingt) sowie differenzierbar auf  $(0, \infty)$  mit Ableitung

$$\Gamma'(x) := \frac{d}{dx} \int_0^{\infty} t^{x-1} \cdot e^{-t} dt = \int_0^{\infty} \ln(t) \cdot t^{x-1} \cdot e^{-t} dt.$$

Außerdem gilt

$$\Gamma(1) = \int_0^{\infty} e^{-t} dt = \left[ -e^{-t} \right]_{t=0}^{t=\infty} = 1$$

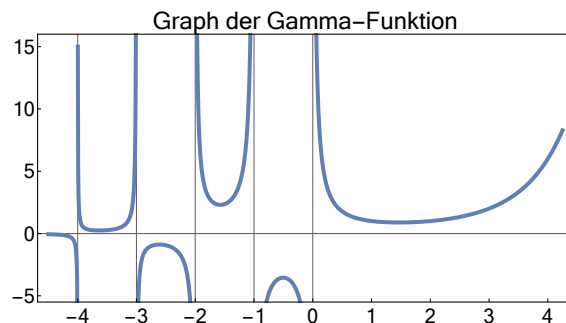
und die partielle Integration liefert die Funktionalgleichung

$$\begin{aligned} \Gamma(x+1) &= \int_0^{\infty} t^x \cdot e^{-t} dt = \left[ -t^x \cdot e^{-t} \right]_{t=0}^{t=\infty} + \int_0^{\infty} x \cdot t^{x-1} \cdot e^{-t} dt \\ &= 0 + x \cdot \Gamma(x), \end{aligned}$$

wobei wir  $0^x = 0$  und  $\lim_{t \rightarrow \infty} t^x \cdot e^{-t} = 0$  benutzt haben. Man zeigt nun leicht durch vollständige Induktion über  $n$ , dass

$$\Gamma(n) = n! \quad \text{für alle } n \in \mathbb{N}$$

gilt, d.h. die Funktion  $\Gamma$  kann als Verallgemeinerung der Fakultät betrachtet werden. Beachte aber, dass  $\lim_{x \searrow 0} \Gamma(x) = \infty \neq 0!$  (siehe Bild). Außerdem kann die Funktionalgleichung in der Form  $\Gamma(x-1) = \Gamma(x)/(x-1)$  benutzt werden, um  $\Gamma(x)$  schrittweise für alle negativen, aber nicht ganzzahligen Argumente  $x$  zu definieren.



Die Gamma-Funktion  $\Gamma$  besitzt Polstellen an den nichtpositiven ganzen Zahlen. Für  $x > 0$  ist  $\Gamma(x)$  durch ein uneigentliches Integral definiert und es gilt  $\Gamma(n) = n!$  für jedes  $n \in \mathbb{N}$ .

## Integrale für Rotationskörper

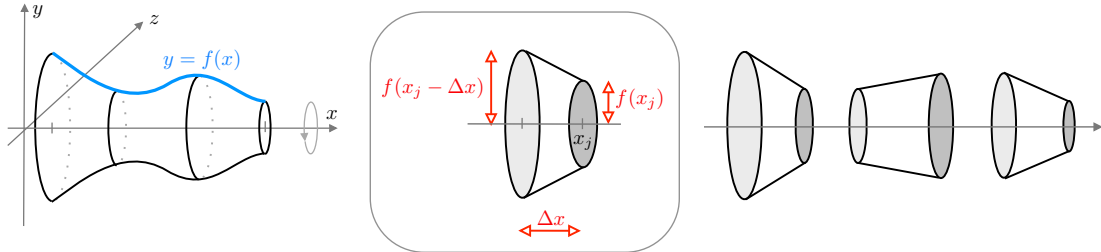
Wir betrachten in diesem Abschnitt geometrische Objekte, die durch Rotation des Graphen einer Funktion  $f : (a, b) \rightarrow (0, \infty)$  um die  $x$ -Achse entstehen (siehe Bild). Die Punktmenge des entsprechenden Rotationskörpers kann als

$$C := \{(x, y, z) : a < x < b, \sqrt{y^2 + z^2} \leq f(x)\}$$

geschrieben werden, wobei die Menge

$$\partial_M C := \{(x, y, z) : a < x < b, \sqrt{y^2 + z^2} = f(x)\}$$

gerade die Punkte der Mantelfläche von  $C$  enthält. Die Einschränkung auf offene Intervalle  $(a, b)$  und nicht-negative Funktionen  $f$  ist dabei nicht wirklich wesentlich. Desweiteren können mit analogen Argumenten auch Rotationskörper mit beliebiger Drehachse studiert werden.



*Links:* Ein Rotationskörper  $C$ , der durch Rotation des Graphen der Funktion  $f$  (blau) um die  $x$ -Achse entsteht.

*Rechts:* Das Prinzip von Cavalieri besagt, dass jeder Rotationskörper durch eine Vereinigung von endlich vielen Kreiskegelscheiben approximiert werden kann, wobei eine gute Approximation aus einer großen Anzahl von sehr dünnen Scheiben besteht. Im Bild wurden die insgesamt drei Scheiben zur besseren Darstellung versetzt gezeichnet.

*Mitte:* Eine typische Kreiskegelscheibe  $S_j$  mit kleiner Dicke  $\Delta x$  und Fußpunkt  $x_j$ .

**Volumen von  $C$**  Nach dem Prinzip von Cavalieri kann der Rotationskörper in viele kleine Kegelstümpfe zerlegt werden (siehe das Bild für eine sehr grobe Zerlegung in nur 3 Teile). Genauer gesagt, mit einer äquidistanten Intervall-Zerlegung

$$\Delta x := \frac{b-a}{n}, \quad x_j := a + (\Delta x) \cdot j, \quad j = 0 \dots n$$

in  $n$  Teilintervalle gilt  $a = x_0 < x_1 < \dots < x_n = b$  und wir berechnen mit unserem schulmathematischen Wissen (siehe auch die nachfolgenden Beispiele)

$$\begin{aligned} \text{vol}(S_j) &= \frac{\pi \cdot \Delta x}{3} \cdot \left( (f(x_j - \Delta x))^2 + f(x_j - \Delta x) \cdot f(x_j) + (f(x_j))^2 \right) \\ &= \pi \cdot (f(x_j))^2 \cdot \Delta x + O((\Delta x)^2) \end{aligned}$$

wobei  $S_j$  die  $j$ -te Scheibe meint und wir die Taylor-Formel

$$f(x + \Delta x) = f(x) + O(\Delta x)$$

verwendet haben. Durch Summation über alle Teile erhalten wir

$$\pi^{-1} \cdot \sum_{j=1}^n \text{vol}(S_j) = \Delta x \cdot \sum_{j=1}^n (f(x_j))^2 + O(1/n)$$

wegen  $\Delta x = O(1/n)$ . Der führende Term auf der rechten Seite ist nun gerade die Riemann-Summe eines Integrales und der Limes  $n \rightarrow \infty$  (bzw.  $\Delta x \rightarrow 0$ ) liefert daher die Formel

$$\text{vol}(C) = \pi \cdot \int_a^b (f(x))^2 dx$$

für das Volumen von  $C$ .

**Mantelfläche von  $C$**  Die Mantelfläche des Kreiskegelstumpfes  $S_j$  ist durch

$$\text{area}_M(S_j) = \pi \cdot (f(x_j - \Delta x) + f(x_j)) \cdot \sqrt{(f(x_j) - f(x_j - \Delta x))^2 + (\Delta x)^2}$$

gegeben und mit dem Satz von Taylor erhalten wir

$$\begin{aligned} \text{area}_M(S_j) &= \pi \cdot (2 \cdot f(x_j) + O(\Delta x)) \cdot \sqrt{(\Delta x)^2 \cdot ((f'(x_j) + O(\Delta x))^2 + 1)} \\ &= 2 \cdot \pi \cdot f(x_j) \cdot \sqrt{1 + (f'(x_j))^2} \cdot \Delta x + O((\Delta x)^2). \end{aligned}$$

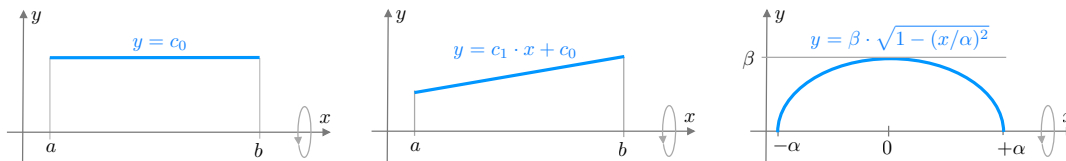
Das Prinzip von Cavalieri die Formel liefert daher die Formel

$$\text{area}_M(C) = 2 \cdot \pi \cdot \int_a^b f(x) \cdot \sqrt{1 + (f'(x))^2} dx$$

für die Mantelfläche von  $C$  bzw.,

$$\text{area}(C) = \text{area}_M(C) + \pi \cdot ((f(a))^2 + (f(b))^2),$$

für die Gesamtoberfläche von  $C$ , wobei die zusätzlichen Terme die Grund- und die Deckfläche quantifizieren.



Die Funktionen für einen Kreiszyylinder, einen Kreiskegelstumpf und einen Ellipsoidsen.

### Beispiele

1. Für eine konstante Funktion  $f(x) = c_0$  erhalten wir

$$\text{vol}(C) = \pi \cdot \int_a^b c_0^2 dx = \pi \cdot (b - a) \cdot c_0^2 = \pi \cdot h \cdot r^2,$$

$$\text{area}_M(C) = 2 \cdot \pi \int_a^b c_0 \cdot \sqrt{1 + 0} dx = 2 \cdot \pi \cdot c_0 \cdot (b - a) = 2 \cdot \pi \cdot h \cdot r$$

und damit die elementargeometrischen Formeln für den Kreiszyylinder. Siehe *Zylinder (Geometrie)* unter WIKIPEDIA und beachte, dass  $r = c_0$  der Radius und  $h = b - a$  die Höhe des Zylinders sind.

2. Mit  $f(x) = c_1 \cdot x + c_0$  ergibt sich

$$\begin{aligned} \text{vol}(C) &= \pi \cdot \int_a^b (c_1^2 \cdot x^2 + 2 \cdot c_1 \cdot c_0 \cdot x + c_0^2) dx \\ &= \pi \cdot \frac{1}{3} \cdot c_1^2 \cdot (b^3 - a^3) + \pi \cdot c_1 \cdot c_0 \cdot (b^2 - a^2) + c_0^2 \cdot (b - a) \\ &= \frac{\pi \cdot h}{3} \cdot (r^2 + r \cdot R + R^2), \end{aligned}$$



sowie

$$\begin{aligned} \text{area}_M(C) &= 2 \cdot \pi \cdot \int_a^b (c_1 \cdot x + c_0) \cdot \sqrt{1 + c_1^2} \, dx \\ &= 2 \cdot \pi \cdot \sqrt{1 + c_1^2} \cdot \left( \frac{1}{2} \cdot c_1 \cdot (b^2 - a^2) + c_0 \cdot (b - a) \right) \\ &= \pi \cdot m \cdot (r + R). \end{aligned}$$

Dies sind die bekannten Formeln für den Kreiskegelstumpf mit Höhe  $h = b - a$  sowie den Radien  $r = f(a) = c_1 \cdot a + c_0$  und  $R = f(b) = c_1 \cdot b + c_0$ , wobei  $m = \sqrt{(R - r)^2 + h^2}$  die *Mantellinie* ist, siehe *Kegelstumpf* bei WIKIPEDIA.

**Beispiel** Durch

$$f(x) = \beta \cdot \sqrt{1 - \left(\frac{x}{\alpha}\right)^2} = \frac{\beta}{\alpha} \cdot \sqrt{\alpha^2 - x^2} \quad -\alpha < x < +\alpha$$

wird ein Rotationsellipsoid  $C_{\alpha,\beta}$  beschrieben. Es gilt

$$\begin{aligned} \text{vol}(C_{\alpha,\beta}) &= \pi \cdot \beta^2 \cdot \int_{-\alpha}^{\alpha} \left(1 - \left(\frac{x}{\alpha}\right)^2\right) \, dx = \pi \cdot \alpha \cdot \beta^2 \cdot \int_{-1}^1 (1 - s^2) \, ds \\ &= \pi \cdot \alpha \cdot \beta^2 \cdot \left[ s - \frac{1}{3}s^3 \right]_{s=-1}^{s=+1} = \frac{4}{3} \cdot \pi \cdot \alpha \cdot \beta^2, \end{aligned}$$

wobei wir  $x = \alpha \cdot s$  substituiert haben. Im Spezialfall  $\alpha = \beta = \varrho$  ergibt sich mit

$$\text{vol}(C_{\varrho,\varrho}) = \frac{4}{3} \cdot \pi \cdot \varrho^3$$

die berühmte Formel für das Volumen der Kugel vom Radius  $\varrho$ . Wegen

$$f'(x) = \frac{\beta}{\alpha} \cdot \frac{1}{2} \cdot \frac{2 \cdot x}{\sqrt{\alpha^2 - x^2}} = \frac{\beta \cdot x}{\alpha \cdot \sqrt{\alpha^2 - x^2}}$$

und

$$\sqrt{1 + (f'(x))^2} = \frac{\sqrt{\alpha^4 + (\beta^2 - \alpha^2) \cdot x^2}}{\alpha \cdot \sqrt{\alpha^2 - x^2}}$$

erhalten wir außerdem

$$\text{area}(C_{\alpha,\beta}) = \text{area}_M(C_{\alpha,\beta}) = \frac{2 \cdot \pi \cdot \beta}{\alpha^2} \cdot \int_{-\alpha}^{+\alpha} \sqrt{\alpha^4 + (\beta^2 - \alpha^2) \cdot x^2} \, dx$$

als vorläufige Formeln für die Mantelfläche, die wir nun weiter vereinfachen wollen. Für  $\alpha = \beta = \varrho$  hängt der Integrand gar nicht von  $x$  ab und wir berechnen die Kugeloberfläche zu

$$\text{area}(C_{\varrho,\varrho}) = \frac{2 \cdot \pi}{\varrho} \cdot 2 \cdot \varrho \cdot \sqrt{\varrho^4} = 4 \cdot \pi \cdot \varrho^2.$$

Im Fall von  $\beta < \alpha$  substituieren wir  $x = \alpha^2 / \sqrt{\alpha^2 - \beta^2} \cdot s$  und erhalten

$$\begin{aligned} \text{area}(C_{\alpha,\beta}) &= \frac{2 \cdot \pi \cdot \beta \cdot \alpha^2}{\sqrt{\alpha^2 - \beta^2}} \cdot \int_{-\sqrt{\alpha^2 - \beta^2}/\alpha}^{+\sqrt{\alpha^2 - \beta^2}/\alpha} \sqrt{1 - s^2} \, ds \\ &= \frac{2 \cdot \pi \cdot \beta \cdot \alpha^2}{\sqrt{\alpha^2 - \beta^2}} \cdot \left[ s \cdot \sqrt{1 - s^2} + \arcsin(s) \right]_{s=-\sqrt{\alpha^2 - \beta^2}/\alpha}^{s=+\sqrt{\alpha^2 - \beta^2}/\alpha} \\ &= \frac{2 \cdot \pi \cdot \beta \cdot \alpha^2}{\sqrt{\alpha^2 - \beta^2}} \cdot \left( \frac{\sqrt{\alpha^2 - \beta^2}}{\alpha} \cdot \frac{\beta}{\alpha} + \arcsin\left(\frac{\sqrt{\alpha^2 - \beta^2}}{\alpha}\right) \right) \\ &= 2 \cdot \pi \cdot \beta^2 + \frac{2 \cdot \pi \cdot \beta \cdot \alpha^2}{\sqrt{\alpha^2 - \beta^2}} \cdot \arcsin\left(\frac{\sqrt{\alpha^2 - \beta^2}}{\alpha}\right), \end{aligned}$$

wobei wir die unbestimmte Integralformel  $\int \sqrt{1 - s^2} \, ds = \frac{1}{2} \cdot s \cdot \sqrt{1 - s^2} + \frac{1}{2} \cdot \arcsin(s) + C$  verwendet haben. Im Fall von  $\beta > \alpha$  ergibt sich hingegen

$$\text{area}(C_{\alpha,\beta}) = 2 \cdot \pi \cdot \beta^2 + \frac{2 \cdot \pi \cdot \beta \cdot \alpha^2}{\sqrt{\beta^2 - \alpha^2}} \cdot \operatorname{arsinh}\left(\frac{\sqrt{\beta^2 - \alpha^2}}{\alpha}\right),$$

durch analoge Rechnungen.

## Kurvenintegrale

**Erinnerung** Eine (stückweise) stetig differenzierbare Abbildung  $\mathbf{c} : [a, b] \rightarrow \mathbb{R}^m$  mit

$$\mathbf{c}(t) = \begin{pmatrix} c_1(t) \\ \vdots \\ c_m(t) \end{pmatrix}$$

wird Parametrisierung einer Kurve genannt, wobei die geometrische Kurve gerade der Graph der Parametrisierung, d.h. die Punktmenge

$$C = \{\mathbf{c}(t) : t \in [a, b]\} \subset \mathbb{R}^m,$$

ist und in den Anwendungen in der Regel  $m = 2$  oder  $m = 3$  gilt. Dabei kann der Kurvenparameter  $t \in [a, b]$  meist als Zeit interpretiert werden und

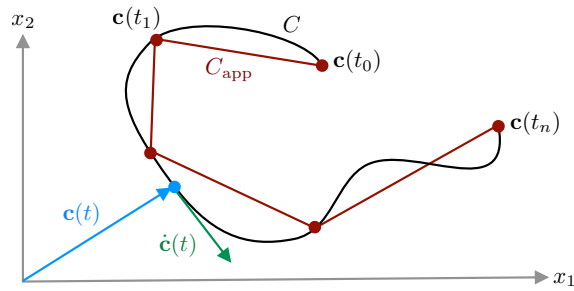
$$\dot{\mathbf{c}}(t) = \frac{d}{dt} \mathbf{c}(t) = \begin{pmatrix} \dot{c}_1(t) \\ \vdots \\ \dot{c}_m(t) \end{pmatrix}$$

ist der momentane Tangential- oder Geschwindigkeitsvektor, der die vektorielle Geschwindigkeit eines entlang der parametrisierten Kurve bewegten Teilchens beschreibt. Beachte, dass bei Ableitungen nach der Zeit in Mathematik und Physik oftmals die Punkt- statt der Strichnotation verwendet wird, d.h. wir schreiben  $\dot{\mathbf{c}}$  statt  $\mathbf{c}'$ .

Die Kurve ist geschlossen, wenn

$$\mathbf{c}(a) = \mathbf{c}(b)$$

gilt, d.h. wenn Anfangs- und Endpunkt zusammenfallen.



Eine parametrisierte Kurve und ein approximierender Polygonzug bestehend aus 4 Strecken (die einer Zerlegung von  $[a, b]$  in vier Teilintervalle entspricht). Mittels sehr feiner Zerlegungen kann die Kurve beliebig gut durch einen Polygonzug angenähert werden.

### Bogenlänge einer Kurve Das Integral

$$L(\mathbf{c}) := \int_a^b \|\dot{\mathbf{c}}(t)\| dt$$

liefert die Länge einer Kurve. Um dies zu verstehen, betrachten wir eine Zerlegung

$$a = t_0 < t_1 < t_2 < \dots < t_n = b$$

des Intervalles  $[a, b]$  und definieren (siehe Bild) eine stückweise affine Approximation  $\mathbf{c}_{\text{app}} : [a, b] \rightarrow \mathbb{R}^m$  durch

$$\mathbf{c}_{\text{app}}(t) = \frac{t_j - t}{t_j - t_{j-1}} \cdot \mathbf{c}(t_{j-1}) + \frac{t - t_{j-1}}{t_j - t_{j-1}} \cdot \mathbf{c}(t_j) \quad \text{für } t \in [t_{j-1}, t_j],$$

d.h. in jedem Teilintervall  $[t_{j-1}, t_j]$  verwenden wir eine leicht andere Formel. Wegen

$$\dot{\mathbf{c}}_{\text{app}}(t) = \frac{1}{t_j - t_{j-1}} \cdot (\mathbf{c}(t_j) - \mathbf{c}(t_{j-1})) \quad \text{für } t \in [t_{j-1}, t_j]$$

gilt damit

$$\begin{aligned} L(\mathbf{c}_{\text{app}}) &= \int_a^b \|\dot{\mathbf{c}}_{\text{app}}(t)\| dt = \sum_{j=1}^n \int_{t_{j-1}}^{t_j} \|\dot{\mathbf{c}}_{\text{app}}(t)\| dt \\ &= \sum_{j=1}^n \int_{t_{j-1}}^{t_j} \left\| \frac{\mathbf{c}(t_j) - \mathbf{c}(t_{j-1})}{t_j - t_{j-1}} \right\| dt = \sum_{j=1}^n \frac{\|\mathbf{c}(t_j) - \mathbf{c}(t_{j-1})\|}{t_j - t_{j-1}} \int_{t_{j-1}}^{t_j} 1 dt \\ &= \sum_{j=1}^n \|\mathbf{c}(t_j) - \mathbf{c}(t_{j-1})\|. \end{aligned}$$

Wir sehen nun für den Polygonzug  $\mathbf{c}_{\text{app}}$ , dass die obige Definition von Länge mittels eines Integrals gerade die Summe der euklidischen Längen aller Teilstrecken ist und daher mit dem intuitiven Längenbegriff übereinstimmt. Da jede (stückweise) differenzierbare Kurve beliebig gut durch Polygonzüge approximiert werden kann, schließen wir, dass die obige Integralformel in der Tat die Länge einer Kurve liefert.

**Beispiele**

1. Die Standardparametrisierung der Kreislinie

$$C = \{(x_1, x_2) \in \mathbb{R}^2 : x_1^2 + x_2^2 = \varrho^2\}$$

vom Radius  $\varrho$  ist

$$\mathbf{c}(t) = \varrho \cdot \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix}, \quad \dot{\mathbf{c}}(t) = \varrho \cdot \begin{pmatrix} -\sin(t) \\ \cos(t) \end{pmatrix}$$

mit  $t \in [0, 2\pi]$  und wir erhalten

$$L(\mathbf{c}) = \int_0^{2\pi} \varrho dt = 2 \cdot \pi \cdot \varrho.$$

Es gibt andere Parametrisierungen der Kreislinie, aber diese werden dieselbe Länge liefern (siehe unten).

2. Die
- Kardioide
- ist die planare Kurve

$$\mathbf{c}(t) = (1 + \cos(t)) \cdot \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix}, \quad t \in [0, 2\pi]$$

die wegen  $\mathbf{c}(0) = \mathbf{c}(2\pi)$  eine geschlossene Kurve ist. Nach Produkt- und Kettenregel gilt

$$\dot{\mathbf{c}}(t) = (1 + \cos(t)) \cdot \begin{pmatrix} -\sin(t) \\ \cos(t) \end{pmatrix} - \sin(t) \cdot \begin{pmatrix} \cos(t) \\ \sin(t) \end{pmatrix},$$

und wir erhalten

$$\|\dot{\mathbf{c}}(t)\| = \sqrt{(1 + \cos(t))^2 + \sin^2(t)} = \sqrt{2 \cdot (1 + \cos(t))} = 2 \cdot \left| \cos\left(\frac{1}{2} \cdot t\right) \right|,$$

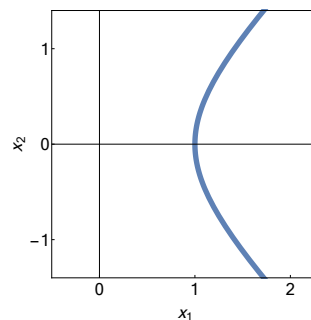
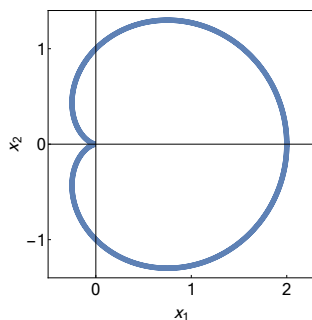
wobei wir zunächst  $\cos^2(1) + \sin^2(1)$  und anschließend

$$\cos(t) = \cos(2 \cdot t/2) = \cos^2(t/2) - \sin^2(t/2) = 2 \cdot \cos^2(t/2) - 1$$

verwendet haben. Insgesamt ergibt sich

$$\begin{aligned} L(C) &= 2 \cdot \int_0^{2\pi} \left| \cos\left(\frac{1}{2} \cdot t\right) \right| dt = 2 \cdot \int_0^{\pi} \cos\left(\frac{1}{2} \cdot t\right) dt - 2 \cdot \int_{\pi}^{2\pi} \cos\left(\frac{1}{2} \cdot t\right) dt \\ &= 2 \cdot \left[ 2 \cdot \sin\left(\frac{1}{2} \cdot t\right) \right]_{t=0}^{t=\pi} - 2 \cdot \left[ 2 \cdot \sin\left(\frac{1}{2} \cdot t\right) \right]_{t=\pi}^{t=2\pi} = 8 \end{aligned}$$

und damit eine ganzzahlige Länge der Kardioide.

Links: Die Kardioide. Rechts: Der Hyperbelast  $x_1^2 = x_2^2 + 1$  mit  $x_1 > 0$ .

**\*Wichtige Eigenschaft** Die Länge ist invariant unter Reparametrisierung in folgendem Sinne: Ist  $h : [\tilde{a}, \tilde{b}] \rightarrow [a, b]$  eine bijektive, streng monoton wachsende und differenzierbare Funktion, so beschreiben die parametrisierten Kurven  $\mathbf{c} : [a, b] \rightarrow \mathbb{R}^m$  und  $\tilde{\mathbf{c}} : [\tilde{a}, \tilde{b}] \rightarrow \mathbb{R}^m$  mit

$$\tilde{\mathbf{c}}(\tilde{t}) := \mathbf{c}(h(\tilde{t}))$$

dasselbe geometrische Objekt (es gilt  $\tilde{C} = C$ ), aber mit Hilfe unterschiedlicher Parametern, nämlich  $\tilde{t}$  bzw.  $t = h(\tilde{t})$ . Mit der Kettenregel für Ableitungen und der Substitutionsformel für Integrale ergibt sich dann

$$L(\mathbf{c}) = \int_a^b \left\| \frac{d\mathbf{c}}{dt}(t) \right\| dt = \int_{\tilde{a}}^{\tilde{b}} \left\| \frac{d\mathbf{c}}{dt}(h(\tilde{t})) \right\| \frac{dh}{d\tilde{t}}(\tilde{t}) d\tilde{t} = \int_{\tilde{a}}^{\tilde{b}} \left\| \frac{d\tilde{\mathbf{c}}}{d\tilde{t}}(\tilde{t}) \right\| d\tilde{t} = L(\tilde{\mathbf{c}}),$$

wobei wir benutzt haben, dass nach Annahme  $\frac{dh}{d\tilde{t}}(\tilde{t}) > 0$  gilt. Insbesondere ist die Länge eine geometrische Eigenschaft von  $C$ .

**\*Bemerkung** Durch

$$s(t) := \int_a^t \|\dot{\mathbf{c}}(\tau)\| d\tau$$

wird die Bogenlängenfunktion von  $\mathbf{c}$  definiert, die eine monoton wachsende Funktion auf  $[a, b]$  ist. Gilt

$$\frac{d}{dt}s(t) = \|\dot{\mathbf{c}}(t)\| = 1,$$

für alle  $t \in [a, b]$ , so nennt man  $\mathbf{c}$  auch eine Bogenlängenparametrisierung von  $C$ . Für jede geometrische Kurve  $C$  kann immer eine Parametrisierung  $\mathbf{c}$  gewählt werden, so dass diese Bedingung erfüllt ist.

**Überstrichene Fläche einer planaren Kurve** Im Fall einer planaren Kurve (d.h.  $m = 2$ ) beschreibt das Integral

$$F(\mathbf{c}) = \frac{1}{2} \cdot \int_a^b (c_1(t) \cdot \dot{c}_2(t) - \dot{c}_1(t) \cdot c_2(t)) dt$$

die vorzeichenbehaftete Fläche, die vom Ortsvektor eines entlang der Kurve bewegten Punktes überstrichen wird. Für einen approximativen Polygonzug reduziert sich diese Formel auf

$$F(\mathbf{c}_{\text{app}}) = \frac{1}{2} \cdot \sum_{j=1}^n c_1(t_{j-1}) c_2(t_j) - c_1(t_j) c_2(t_{j-1})$$

und kann als Summe von positiv und negativ gezählten Dreiecksflächen interpretiert werden (siehe Bild).

Herleitung: Die reduzierte Formel kann mittels folgender Rechnungen hergeleitet werden. Die Formel für den approximierenden Polygonzug impliziert

$$\begin{aligned} F_{12,j} &:= \int_{t_{j-1}}^{t_j} c_1(t) \dot{c}_2(t) dt \\ &= \int_{t_{j-1}}^{t_j} \left( \frac{t_j - t}{t_j - t_{j-1}} c_1(t_{j-1}) + \frac{t - t_{j-1}}{t_j - t_{j-1}} c_1(t_j) \right) \cdot \frac{c_2(t_j) - c_2(t_{j-1})}{t_j - t_{j-1}} dt \end{aligned}$$

und wir berechnen unter Verwendung von

$$\int_{t_{j-1}}^{t_j} \frac{t_j - t}{t_j - t_{j-1}} dt = \int_{t_{j-1}}^{t_j} \frac{t - t_{j-1}}{t_j - t_{j-1}} dt = \frac{t_j - t_{j-1}}{2}$$

die vereinfachte Formel

$$F_{12,j} = \frac{1}{2} \cdot (c_1(t_j) + c_1(t_{j-1})) \cdot (c_2(t_j) - c_2(t_{j-1})).$$

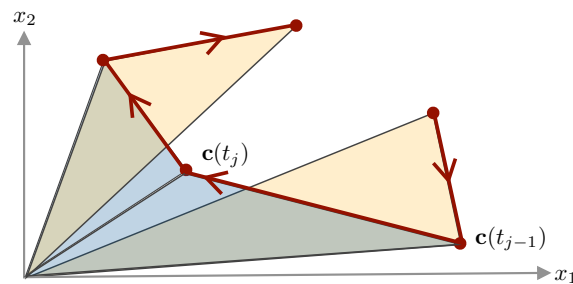
Durch Vertauschung der Koordinatenindizes 1  $\leftrightarrow$  2 erhalten wir den analogen Ausdruck

$$F_{21,j} = \frac{1}{2} \cdot (c_2(t_j) + c_2(t_{j-1})) \cdot (c_1(t_j) - c_1(t_{j-1})),$$

und mit

$$F(\mathbf{c}_{\text{app}}) = \frac{1}{2} \cdot \sum_{j=1}^n (F_{12,j} - F_{21,j})$$

folgt schließlich die behauptete Formel  $F(\mathbf{c}_{\text{app}})$  nach einfachen Umformungen.



Die durch eine planare Kurve überstrichene Fläche kann auch durch ein Integral berechnet werden. Für planare Polygonzüge ist die entsprechende Formel äquivalent zu elementargeometrischen Argumenten, wobei die blauen bzw. gelben Dreiecksflächen einen positiven bzw. negativen Beitrag liefern.

## Beispiele

1. Für die Standardparametrisierung des Kreises vom Radius  $\rho$  (siehe oben) erhalten wir

$$F(C) = \frac{1}{2} \cdot \rho^2 \cdot \int_0^{2\pi} (\cos(t) \cdot \cos(t) + \sin(t) \cdot \sin(t)) dt = \pi \cdot \rho^2.$$

Durchlaufen wir den Kreis jedoch via

$$\mathbf{c}(t) = \varrho \cdot \begin{pmatrix} \cos(-t) \\ \sin(-t) \end{pmatrix} = \begin{pmatrix} \cos(t) \\ -\sin(t) \end{pmatrix}, \quad \dot{\mathbf{c}}(t) = \varrho \cdot \begin{pmatrix} -\sin(t) \\ -\cos(t) \end{pmatrix}$$

in mathematisch negativer Richtung (also gegen den Uhrzeigersinn), so gilt

$$F(\mathbf{c}) = \frac{1}{2} \cdot \int_0^{2\pi} (-\cos(t) \cdot \cos(t) - \sin(t) \cdot \sin(t)) dt = -\pi \cdot \varrho^2,$$

d.h. wir erhalten den negativen Flächeninhalt.

2. Durch

$$\mathbf{c}(t) = \begin{pmatrix} \cosh(t) \\ \sinh(t) \end{pmatrix}, \quad \dot{\mathbf{c}}(t) = \begin{pmatrix} \sinh(t) \\ \cosh(t) \end{pmatrix}$$

ist eine Parametrisierung eines Hyperbel-Astes (siehe Bild) gegeben, denn es gilt

$$c_1^2(t) - c_2^2(t) = 1 \quad c_1(t) > 0$$

für alle  $t \in \mathbb{R}$ . Wir erhalten

$$\int_{\alpha}^{\beta} (c_1(t) \dot{c}_2(t) - \dot{c}_1(t) c_2(t)) dt = \int_{\alpha}^{\beta} (\cosh^2(t) - \sinh^2(t)) dt = \int_{\alpha}^{\beta} 1 dt = \beta - \alpha.$$

Insbesondere werden beim Durchlaufen dieser Kurve in gleichen Zeitintervallen immer gleiche Flächen überstrichen.

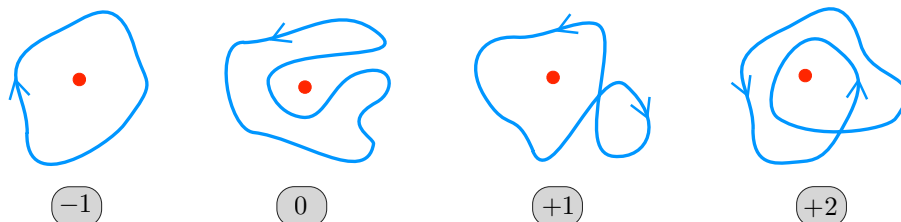
3. Für die Kardioide von oben ergibt sich

$$F(\mathbf{c}) = \frac{1}{2} \cdot \int_0^{2\pi} (1 + \cos(t))^2 dt = \frac{1}{2} \cdot \int_0^{2\pi} (1 + 2 \cos(t) + \cos^2(t)) dt = \frac{3}{2} \cdot \pi.$$

**\*Umlaufzahl** Beschreibt  $\mathbf{c}$  eine geschlossene planare Kurve und ist  $\xi \notin C$  ein beliebiger Punkt außerhalb der Kurve, so ist

$$\text{wnd}(\mathbf{c}, \xi) := \frac{1}{2\pi} \cdot \int_a^b \frac{(c_1(t) - \xi_1) \cdot \dot{c}_2(t) - \dot{c}_1(t) \cdot (c_2(t) - \xi_2)}{(c_1(t) - \xi_1)^2 + (c_2(t) - \xi_2)^2} dt$$

die (ganzzahlige) Windungs- oder Umlaufzahl der Kurve  $C$  um den Punkt  $\xi$  (siehe Bild).



Beispiele für die Umlaufzahl (schwarz) einer Kurve  $\mathbf{c}$  (blau) bzgl. eines Punktes  $\xi$  (rot).

**Erstes (oder skalares) Kurvenintegral** Ist  $f : D \subset \mathbb{R}^m \rightarrow \mathbb{R}$  eine stetige Funktion und  $\mathbf{c} : [a, b] \rightarrow D$  eine parametrisierte Kurve in  $D$ , so wird

$$\int_{\mathbf{c}} f(\mathbf{x}) \, ds := \int_a^b f(\mathbf{c}(t)) \|\dot{\mathbf{c}}(t)\| \, dt \in \mathbb{R}$$

das Kurvenintegral der 1. Art von  $f$  bzgl.  $C$  genannt.

**Anwendung** Wir nehmen an,  $\mathbf{c}$  beschreibt einen (dreidimensionalen, aber sehr dünnen) Draht mit konstanter Massendichte  $\varrho$ . Dann ist der Schwerpunkt des Drahtes durch

$$\mathbf{X} = \begin{pmatrix} X_1 \\ X_2 \\ X_3 \end{pmatrix}, \quad X_i = \int_{\mathbf{c}} x_i \, ds \quad / \quad \int_{\mathbf{c}} 1 \, ds$$

gegeben, wobei das Nennerintegral gerade die Länge des Drahtes ist. Mit analogen Formel in zwei Dimensionen erhalten wir für die Kardioide die Schwerpunktkoordinaten  $X_1 = 4/5$ ,  $X_2 = 0$  (Nachrechnen!).

**Zweites (oder vektoriell) Kurvenintegral** Ist  $\mathbf{f} : D \subset \mathbb{R}^m \rightarrow \mathbb{R}^m$  eine stetige Funktion und  $\mathbf{c} : [a, b] \rightarrow D$  eine parametrisierte Kurve in  $D$ , so wird

$$\int_{\mathbf{c}} \mathbf{f}(\mathbf{x}) \cdot d\mathbf{x} := \int_a^b \mathbf{f}(\mathbf{c}(t)) \cdot \dot{\mathbf{c}}(t) \, dt \in \mathbb{R}$$

das Kurvenintegral der 2. Art von  $\mathbf{f}$  bzgl.  $C$  genannt, wobei  $\cdot$  auf der linken und der rechten Seite das Skalarprodukt meint.

**Anwendung** Beschreibt  $\mathbf{c}$  die Bahn eines geladenen Teilchens mit Ladung  $q$  und Masse  $m$  und ist  $\mathbf{f}$  das elektrische Feldstärke, so ist

$$q \int_{\mathbf{c}} \mathbf{f}(\mathbf{x}) \cdot d\mathbf{x}$$

gerade die vom elektrischen Feld am Teilchen verrichtete Arbeit. Analog ist

$$m \int_{\mathbf{c}} \mathbf{g}(\mathbf{x}) \cdot d\mathbf{x}$$

die vom Beschleunigungsfeld  $\mathbf{g}$  (etwa Erdbeschleunigung) geleistete Arbeit.

**\*Wichtige Eigenschaft** Beide Kurvenintegrale sind invariant unter Reparametrisierung, denn mit den obigen Notationen gilt

$$\begin{aligned} \int_{\mathbf{c}} \mathbf{f}(\mathbf{x}) \cdot d\mathbf{x} &= \int_a^b \mathbf{f}(\mathbf{c}(t)) \cdot \frac{d\mathbf{c}}{dt}(t) \, dt = \int_{\tilde{a}}^{\tilde{b}} \mathbf{f}(\mathbf{c}(h(\tilde{t}))) \cdot \frac{d\mathbf{c}}{dt}(h(\tilde{t})) \frac{dh}{d\tilde{t}}(\tilde{t}) \, d\tilde{t} \\ &= \int_{\tilde{a}}^{\tilde{b}} \mathbf{f}(\tilde{\mathbf{c}}(\tilde{t})) \cdot \frac{d\tilde{\mathbf{c}}}{d\tilde{t}}(\tilde{t}) \, d\tilde{t} = \int_{\tilde{\mathbf{c}}} \mathbf{f}(\mathbf{x}) \cdot d\mathbf{x} \end{aligned}$$



sowie eine analoge Formel gilt für die Integrale der 1. Art. Insbesondere hängt der Wert beider Kurvenintegrale nur von den Eigenschaften der geometrischen Kurve  $C$  und der Abbildung  $\mathbf{f}$  bzw.  $f$  ab.



# Kapitel 11

## Fourier-Reihen

**Motivation** Jedes periodische Signal kann in elementare Bausteine zerlegt werden bzw. aus diesen zusammengesetzt werden.

Vorlesung 36, 31. Januar 2019

### 11.1 Fourier-Reihen periodischer Funktionen

#### Über $T$ -periodische Funktionen

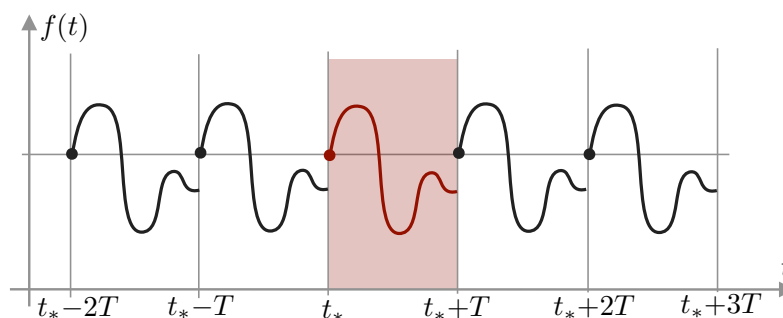
**Definition** Für eine gegebene Periode  $T > 0$  sagen wir, die Funktion  $f : \mathbb{R} \rightarrow \mathbb{R}$  ist  $T$ -periodisch, falls

$$f(t + T) = f(t)$$

für alle  $t \in \mathbb{R}$  gilt.

**Beobachtung** Jede  $T$ -periodische Funktion ist eindeutig durch ihre Werte auf einem halboffenen Intervall der Länge  $T$ , also auf einem Intervall der Bauart  $[t_*, t_* + T)$  oder  $(t_*, t_* + T]$ , definiert (siehe Bild). So gilt zum Beispiel:

$$f(t) = \begin{cases} f(t + 2T) & \text{für } t_* - 2T \leq t < t_* - T \\ f(t + T) & \text{für } t_* - T \leq t < t_* \\ f(t) & \text{für } t_* \leq t < t_* + T \\ f(t - T) & \text{für } t_* + T \leq t < t_* + 2T \\ f(t - 2T) & \text{für } t_* + 2T \leq t < t_* + 3T \\ \dots & \dots \end{cases}$$



Zur  $T$ -periodischen Fortsetzung einer auf dem Intervall  $[t_*, t_* + T)$  (rot) gegebenen Funktion. Beachte, dass durch die Fortsetzung Unstetigkeiten entstehen können.

**Lemma (Nützliche Integralformeln)** Für jede  $T$ -periodische Funktion gilt

$$\int_{t_*}^{t_*+T} f(t) dt = \int_0^T f(t) dt = \int_{-T/2}^{T/2} f(t) dt, \quad \int_0^{nT} f(t) dt = n \int_0^T f(t) dt,$$

sofern auch nur eines der Integrale wohldefiniert ist.

Beweis: Wir bemerken zunächst, dass wegen

$$\frac{d}{dt_*} \int_{t_*}^{t_*+T} f(t) dt = f(t_* + T) - f(t_*) = 0,$$

das Integral auf der linken Seite nicht von  $t_*$  abhängt. Außerdem gilt

$$\int_0^{nT} f(t) dt = \sum_{i=1}^n \int_{(i-1)T}^{iT} f(t) dt = \sum_{i=1}^n \int_0^T f(\tilde{t} + (i-1)T) d\tilde{t} = \sum_{i=1}^n \int_0^T f(\tilde{t}) d\tilde{t},$$

wobei wir im  $i$ -ten Teilintegral  $t = \tilde{t} + (i-1)T$  substituiert und die Periodizität von  $f$  eingesetzt haben.  $\square$

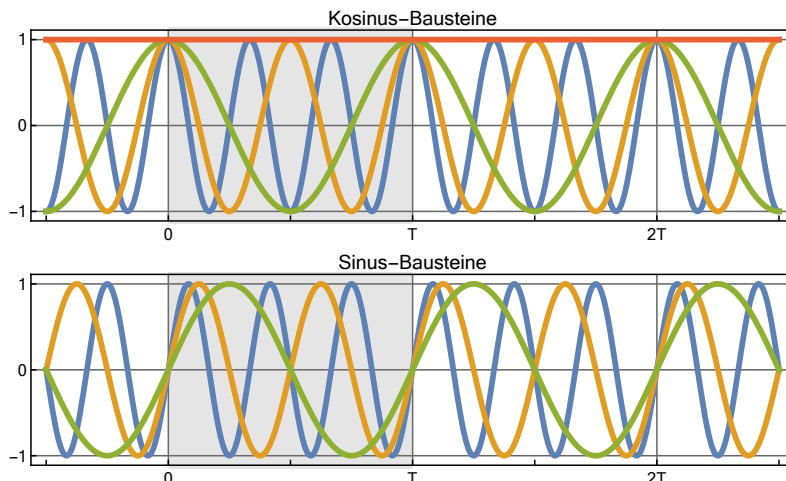
**Harmonische Funktionen als Bausteine periodischer Funktionen** Die Funktionen

$$1 = \cos(0\omega t), \quad \cos(m\omega t), \quad \sin(m\omega t)$$

mit  $m \in \mathbb{N} = \{1, 2, 3, \dots\}$  sind alle  $T$ -periodisch, sofern die Kreisfrequenz durch

$$\omega := \frac{2\pi}{T}$$

festgelegt ist (siehe Bild).



Die harmonischen Funktionen für  $m = 1$  (grün),  $m = 2$  (gelb),  $m = 3$  (blau) und die konstante Funktion (rot).

Durch Nachrechnen im Reellen (oder im Komplexen wie weiter unten) verifizieren wir die sogenannten Orthogonalitätsrelationen

$$\int_0^T 1 \cdot 1 \, dt = T, \quad \int_0^T 1 \cdot \cos(m\omega t) \, dt = 0, \quad \int_0^T 1 \cdot \sin(m\omega t) \, dt = 0$$

sowie

$$\int_0^T \sin(m\omega t) \cdot \sin(l\omega t) \, dt = \int_0^T \cos(m\omega t) \cdot \cos(l\omega t) \, dt = \frac{1}{2} \cdot T \cdot \delta_{ml}$$

und

$$\int_0^T \cos(k\omega t) \cdot \sin(l\omega t) \, dt = 0.$$

Hierbei sind  $m, l \in \mathbb{N}$  beliebig und

$$\delta_{ml} = \begin{cases} 1 & \text{für } m = l, \\ 0 & \text{für } m \neq l \end{cases}$$

ist wieder das Kronecker-Delta. Diese Relationen sind sehr wichtig und beschreiben, dass die harmonischen Funktionen paarweise senkrecht aufeinander stehen, wobei ein geeigneter Vektorraum (der sogenannte Lebesguesche Funktionenraum  $L^2$ ) sowie das Skalarprodukt

$$\langle u, v \rangle = \frac{1}{T} \int_0^T u(t)v(t) \, dt$$

zu Grunde liegen. Wir werden zwar die Formeln benutzen, aber von dem abstrakten Setting keinen weiteren Gebrauch machen.

**Grundkonzept** Die zentrale Idee der Fourier-Analyse ist, dass sich jede hinreichend gute  $T$ -periodische Funktion als Reihe

$$f(t) = \frac{a_0}{2} + \sum_{m=1}^{\infty} a_m \cos(m\omega t) + \sum_{m=1}^{\infty} b_m \sin(m\omega t)$$

mit geeigneten Koeffizienten  $a_m$  und  $b_m$  schreiben lässt.

**Fragen** Für welche Funktion  $f$  ist die Reihendarstellung möglich? Und wenn ja, wie berechnen sich die Koeffizienten?

**Notwendige Bedingungen** Die zweite Frage können wir vergleichsweise einfach beantworten: Wir multiplizieren beide Seiten der Reihendarstellung mit einer der harmonischen Funktionen, integrieren über eine Periodenlänge, vertauschen Integration und Reihenbildung (obwohl man streng genommen erst rechtfertigen muss, dass

man das wirklich darf), und setzen die Orthogonalitätsrelationen ein. Für den konstanten Baustein ergibt sich

$$\begin{aligned} \int_0^T f(t) \cdot 1 \, dt &= \frac{a_0}{2} \int_0^T dt + \sum_{m=1}^{\infty} a_m \int_0^T \cos(m\omega t) \, dt + \sum_{m=1}^{\infty} b_m \int_0^T \sin(m\omega t) \, dt \\ &= \frac{a_0}{2} T + \sum_{m=0}^{\infty} a_m \cdot 0 + \sum_{m=1}^{\infty} b_m \cdot 0 = \frac{1}{2} \cdot a_0 \cdot T \end{aligned}$$

und damit eine notwendige Bedingung für  $a_0$ . Für den  $l$ -ten Kosinus erhalten wir

$$\begin{aligned} \int_0^T f(t) \cos(l\omega t) \, dt &= \frac{a_0}{2} \int_0^T \cos(l\omega t) \, dt + \sum_{m=1}^{\infty} a_m \int_0^T \cos(m\omega t) \cdot \cos(l\omega t) \, dt \\ &\quad + \sum_{m=1}^{\infty} b_m \int_0^T \sin(m\omega t) \cdot \cos(l\omega t) \, dt \\ &= \frac{1}{2} \cdot a_0 \cdot 0 + \frac{1}{2} \cdot T \cdot \sum_{m=0}^{\infty} a_m \cdot \delta_{ml} + \sum_{m=1}^{\infty} b_m \cdot 0 \\ &= \frac{1}{2} \cdot T \cdot a_l \end{aligned}$$

und mit dem  $l$ -ten Sinus analog

$$\int_0^T f(t) \sin(l\omega t) \, dt = \frac{1}{2} \cdot T \cdot b_l.$$

**Fourier-Koeffizienten und Fourier-Polynome** Wir betrachten nun die notwendigen Bedingungen als Definition der Koeffizienten, d.h. wir setzen

$$a_0 := \frac{2}{T} \int_0^T f(t) \, dt, \quad a_m := \frac{2}{T} \int_0^T f(t) \cos(m\omega t) \, dt, \quad b_m := \frac{2}{T} \int_0^T f(t) \sin(m\omega t) \, dt$$

für jedes  $m \in \mathbb{N}$ , und definieren

$$F_N(t) := \frac{a_0}{2} + \sum_{m=1}^N a_m \cos(m\omega t) + \sum_{m=1}^N b_m \sin(m\omega t).$$

Beachte, dass  $F_N$  für jedes  $N \in \mathbb{N}$  unendlich oft stetig differenzierbar und  $T$ -periodisch ist. Außerdem kann man mit den Additionstheoremen

$$\sin(2\omega t) = \sin(\omega t) \cos(\omega t), \quad \cos(2\omega t) = \cos^2(\omega t) - \sin^2(\omega t), \quad \text{usw.}$$

zeigen, dass jede Funktion  $F_N$  sich als trigonometrisches Polynom schreiben lässt, d.h. als Linearkombination von Produkten/Potenzen von  $\sin(\omega t)$  und  $\cos(\omega t)$ . Die genauen Formeln interessieren uns hier aber nicht.

**Notation** Manchmal schreibt man  $F_{f,n}(t)$  und  $a_{f,m}$ ,  $b_{f,m}$  statt  $F_n(t)$  und  $a_m$ ,  $b_m$ , um die Abhängigkeit von  $f$  deutlich zu machen.

**Fragen** Wir können nun für jede feste Funktion  $f$  die Fragen von oben wie folgt erweitern:

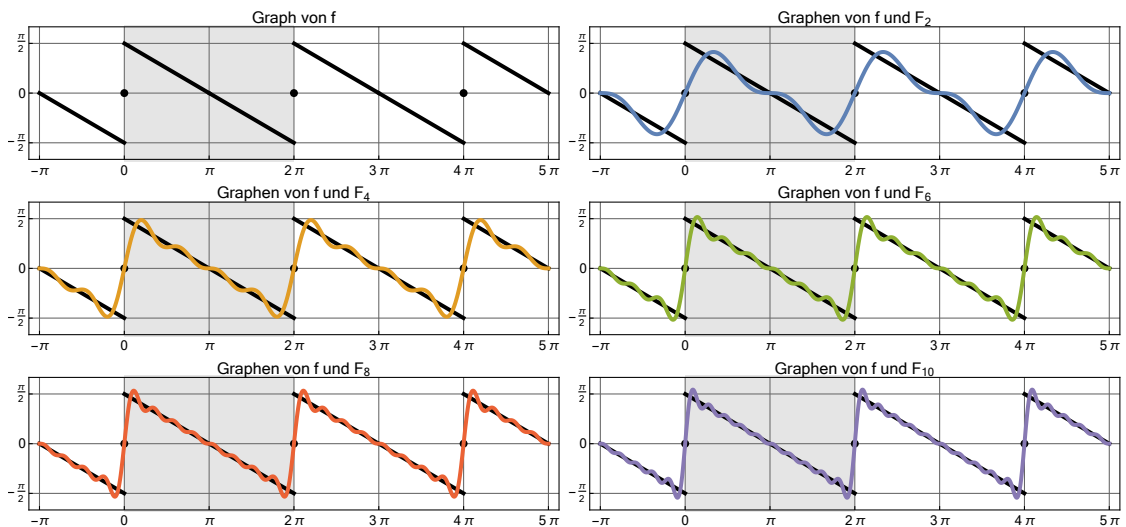
1. Für welche Werte von  $t$  existiert der Limes  $F_\infty(t) = \lim_{n \rightarrow \infty} F_n(t)$  und wie schnell/langsam ist die Konvergenz?
2. Für welche Werte von  $t$  gilt  $F_\infty(t) = f(t)$ ?

Wir wollen nun zunächst einige Beispiele rechnen, wobei in diesen Beispielen immer  $T = 2\pi$  und damit  $\omega = 1$  gilt.

**Die Sägezahnfunktion** Wir definieren  $f$  als  $2\pi$ -periodische Fortsetzung von

$$f(t) := \frac{1}{2}(\pi - t) \quad \text{für } 0 < t < 2\pi, \quad f(0) = 0,$$

wobei der Wert von  $f(0)$  nicht wirklich wichtig ist.



Beispiel 1: Die Sägezahnfunktion.

Da  $f$  ungerade ist (d.h.  $f(x) = -f(-x)$ ) ergibt sich (siehe auch das Lemma oben)

$$a_0 = \frac{1}{\pi} \int_{-\pi}^{+\pi} f(t) dt = 0$$

sowie

$$a_m = \frac{1}{\pi} \int_0^{2\pi} f(t) \cos(mt) dt = \frac{1}{\pi} \int_{-\pi}^{+\pi} f(t) \cos(mt) dt = 0,$$

weil der Integrand als Produkt einer geraden und einer ungeraden Funktion selbst ungerade ist. Außerdem gilt

$$\begin{aligned} b_m &= \frac{1}{\pi} \int_0^{2\pi} f(t) \sin(mt) dt = \frac{1}{\pi} \int_0^{2\pi} \frac{1}{2}(\pi - t) \sin(mt) dt \\ &= -\frac{1}{2\pi} \int_0^{2\pi} t \sin(mt) dt = \frac{1}{2\pi} \left[ t \frac{\cos(mt)}{m} \right]_{t=0}^{t=2\pi} - \frac{1}{2\pi} \int_0^{2\pi} \frac{\cos(mt)}{m} dt \\ &= \frac{1}{2\pi} \left[ t \frac{\cos(mt)}{m} \right]_{t=0}^{t=2\pi} = \frac{1}{m}, \end{aligned}$$

und wir erhalten

$$F_\infty(t) = \sin(t) + \frac{\sin(2t)}{2} + \frac{\sin(3t)}{3} + \frac{\sin(4t)}{4} + \dots$$

als explizite Formel für die Fourier-Reihe von  $f$  sowie die entsprechenden Ausdrücke für die Partialsummen  $F_N(t)$  (siehe Bild).

**Rechteckschwingung** Für die  $2\pi$ -periodische Funktion  $f$  mit

$$f(t) := +1 \text{ für } 0 < t < \pi, \quad f(t) := -1 \text{ für } \pi < t < 2\pi, \quad f(0) := f(\pi) := 0,$$

gilt wieder

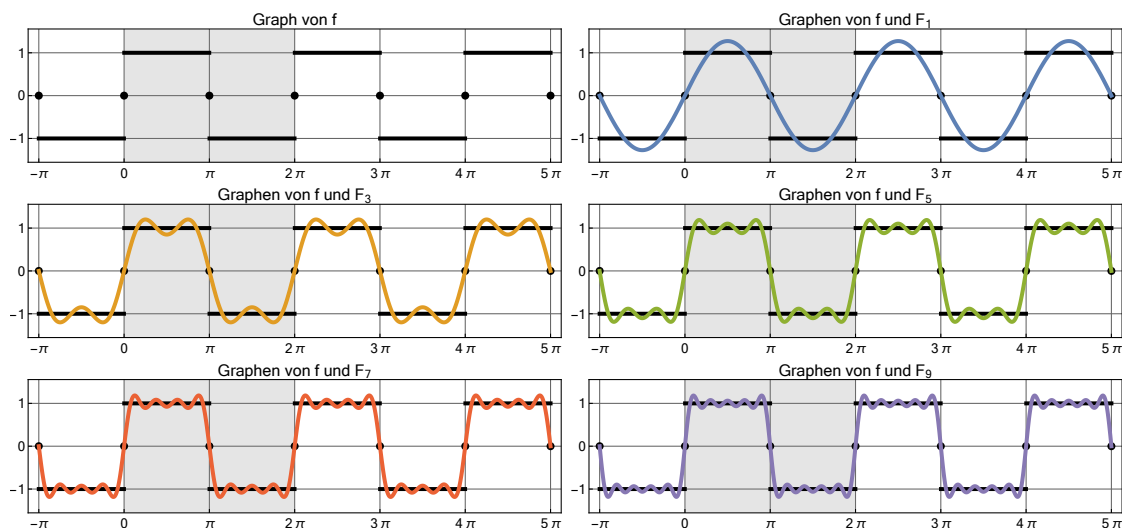
$$a_0 = 0, \quad a_m = 0$$

für alle  $m \in \mathbb{N}$ , da  $f$  ungerade ist (siehe Bild). Des Weiteren berechnen wir

$$\begin{aligned} b_m &= \frac{1}{\pi} \int_{-\pi}^{+\pi} f(t) \sin(mt) dt = -\frac{1}{\pi} \int_{-\pi}^0 \sin(mt) dt + \frac{1}{\pi} \int_0^{+\pi} \sin(mt) dt \\ &= \left[ \frac{\cos(mt)}{m\pi} \right]_{t=-\pi}^{t=0} - \left[ \frac{\cos(mt)}{m\pi} \right]_{t=0}^{t=+\pi} = \frac{2}{m\pi} - \frac{2\cos(m\pi)}{m\pi} \\ &= \frac{4}{m\pi} \begin{cases} 1 & \text{für } m \text{ ungerade} \\ 0 & \text{für } m \text{ gerade} \end{cases} \end{aligned}$$

und erhalten

$$F_\infty(t) = \frac{4 \sin(t)}{\pi} + \frac{4 \sin(3t)}{3\pi} + \frac{4 \sin(5t)}{5\pi} + \dots$$



Beispiel 2: Die Rechteckschwingung.

**Nebenresultat** Es gilt

$$F_\infty\left(\frac{\pi}{2}\right) = \frac{4}{\pi} \left( \frac{1}{1} - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \dots \right)$$



und, wie wir unten werden unten sehen, auch  $f(\pi/2) = F_\infty(\pi/2)$ . Durch Einsetzen und Umformen erhalten wir

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} = \frac{\pi}{4} f\left(\frac{\pi}{2}\right) = \frac{\pi}{4}.$$

Ganz allgemein gilt: Durch das Studium von Fourier-Reihen können viele exakte Summenformeln für Reihen abgeleitet werden, obwohl das weder die ursprüngliche Motivation noch die wichtigste Anwendung die Theorie ist.

**Periodisch fortgesetzte Parabel** Die  $2\pi$ -periodische Fortsetzung von

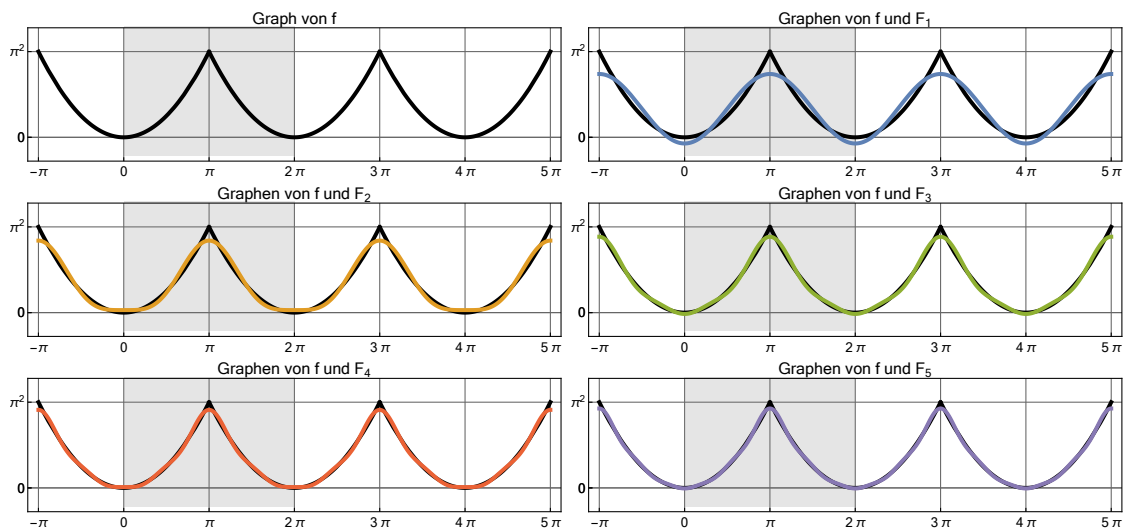
$$f(t) = t^2 \quad \text{für} \quad -\pi \leq t < \pi$$

ist gerade (siehe Bild), d.h., diesmal gilt

$$b_m = \frac{1}{\pi} \int_{-\pi}^{+\pi} f(t) \sin(mt) dt = 0,$$

und eine direkte Rechnung zeigt

$$a_0 = \frac{1}{\pi} \int_{-\pi}^{+\pi} t^2 dt = \frac{2}{\pi} \int_0^{+\pi} t^2 dt = \left[ \frac{2t^3}{3\pi} \right]_{t=0}^{t=\pi} = \frac{2}{3} \pi^2.$$



Beispiel 3: Die periodische Fortsetzung der Parabel.

Eine zweifache partielle Integration liefert außerdem

$$a_m = \frac{2}{\pi} \int_0^{+\pi} t^2 \cos(mt) dt = \frac{(-1)^m 4}{m^2},$$

und wir erhalten

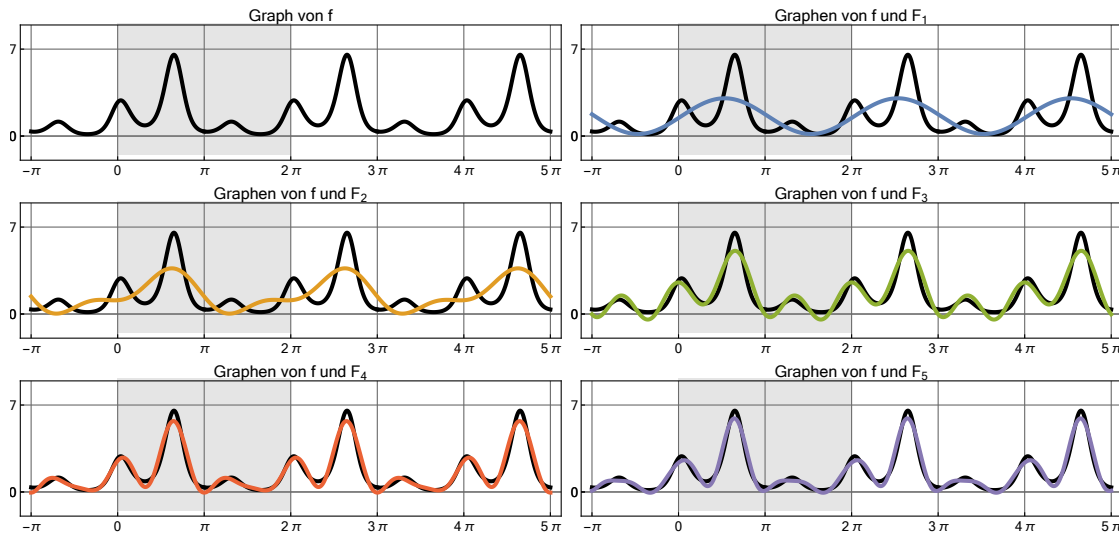
$$F_\infty(t) = \frac{\pi^2}{3} - \frac{4 \cos(t)}{1} + \frac{4 \cos(2t)}{4} - \frac{4 \cos(3t)}{9} + \dots$$

**Nebenresultat** Kombinieren wir

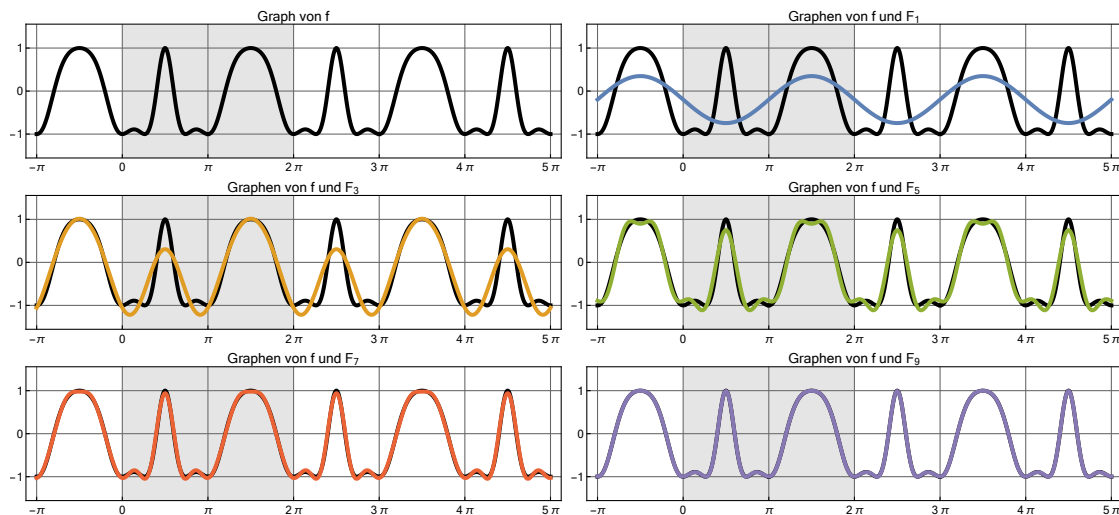
$$F_\infty(\pi) = \frac{\pi^2}{3} + \frac{4}{1} + \frac{4}{4} + \frac{4}{9} + \dots$$

mit  $F_\infty(\pi) = f(\pi)$  (siehe unten), so erhalten wir die klassische Formel

$$\sum_{m=1}^{\infty} \frac{1}{m^2} = \frac{f(\pi) - \frac{1}{3}\pi^2}{4} = \frac{\pi^2}{6}.$$



Beispiel 4:  $2\pi$ -periodische Funktion  $f(t) = \exp(\sin(t) + \cos(3t))$  mit numerisch berechneten Fourier-Koeffizienten.



Beispiel 5:  $2\pi$ -periodische Funktion  $f(t) = \cos(\sin(2t) + \pi \cos(t))$  mit numerisch berechneten Fourier-Koeffizienten.

Vorlesung 37, 03. Februar 2019

**Paritätseigenschaften** Ist  $f$  gerade bzw. ungerade, d.h. gilt

$$f(t) = +f(-t) \quad \text{bzw.} \quad f(t) = -f(-t) \quad \text{für alle} \quad t \in \mathbb{R},$$

so rechnet man leicht nach, dass

$$a_m = \frac{4}{T} \int_0^{T/2} \cos(m\omega t) f(t) dt \quad \text{und} \quad b_m = 0$$

bzw.

$$a_m = 0 \quad \text{und} \quad b_m = \frac{4}{T} \int_0^{T/2} \sin(m\omega t) f(t) dt$$

gilt, wobei  $m \in \mathbb{N}_0$  jeweils beliebig ist.

**Spektraldarstellung** Man kann auch mit dem Ansatz

$$a_m \cos(m\omega t) + b_m \sin(m\omega t) = \varrho_m \sin(m\omega t + \theta_m)$$

die Fourier-Approximationen der Funktion  $f$  in der Form

$$T_\infty(t) = \frac{1}{2} a_0 + \sum_{m=1}^{\infty} \varrho_m \sin(m\omega t + \theta_m)$$

schreiben, wobei einfache Rechnungen zeigen, dass

$$\begin{pmatrix} a_m \\ b_m \end{pmatrix} = \varrho_m \begin{pmatrix} \sin(\theta_m) \\ \cos(\theta_m) \end{pmatrix}$$

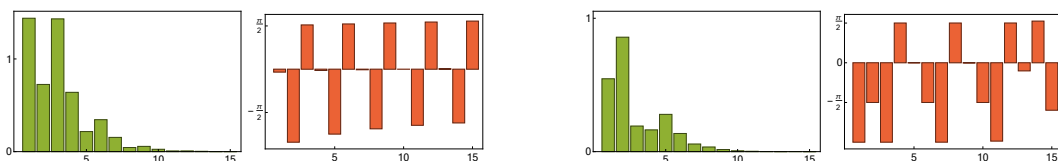
bzw.

$$\varrho_m = \sqrt{a_m^2 + b_m^2}, \quad \tan(\theta_m) = \frac{b_m}{a_m}$$

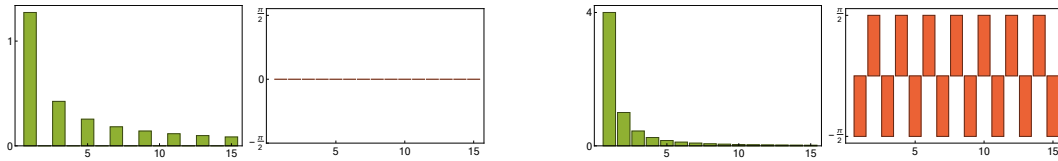
gelten muss. Man plottet nun oftmals die Größe  $\varrho_m$  bzw.  $\theta_m$  über  $m$  und spricht vom Amplitudenspektrum bzw. Phasenspektrum von  $f$ .

### Beispiele

1. Für die Sägezahnfunktion und die Rechtecksspannung ergibt sich unmittelbar  $\varrho_m = b_m$ ,  $\theta_m = 0$  und für die periodische Fortsetzung der Parabel gilt  $\varrho_m = m$ ,  $\theta_m = \pm\pi/2$ .
2. Für die Beispiele 4 und 5 können die Spektren numerisch berechnet werden (siehe Bild).



Das Amplitudenspektrum (grün) und das Phasenspektrum (rot) für die Beispiel 4 (links) und 5 (rechts). Beachte, dass in beiden Beispielen die (nichtnegative) Amplitude  $\varrho_m$  für  $m \rightarrow \infty$  sehr schnell (nämlich exponentiell) abklingt. Der Grund ist, dass die jeweilige Funktion  $f$  als Potenzreihe geschrieben werden kann und damit insbesondere unendlich oft differenzierbar ist. Die Phase  $\theta_m$  hat aber kein Vorzeichen und klingt auch nicht ab.



Die Spektren für die Beispiele 2 und 3, wobei die Amplitude aufgrund des Sprunges bzw. des Knickes der Funktion  $f$  nur vergleichsweise langsam (nämlich nur polynomiell in  $m$ ) abklingt.

## 11.2 Elemente der Theorie

**Die komplexe Variante** Mit elementaren Umformungen zeigen wir, dass die Fourier-Approximation  $F_N$  auch als

$$F_N(t) = \sum_{j=-N}^{+N} \gamma_j \exp(+i j \omega t)$$

geschrieben werden kann, wobei der Index  $j \in \mathbb{Z}$  nun eine ganze Zahl ist und der Koeffizient

$$\gamma_j = \frac{1}{T} \int_0^T f(t) \exp(-i j \omega t) dt$$

Werte in  $\mathbb{C}$  annimmt. In der Tat, die Formel für  $\gamma_j$  kombiniert mit der Euler-Formel und den Eigenschaften von Integralen liefert

$$\gamma_0 = \frac{\alpha_0}{2}, \quad \gamma_{+m} = \frac{a_m - i b_m}{2}, \quad \gamma_{-m} = \frac{a_m + i b_m}{2}$$

für alle  $m \in \mathbb{N}$  und insgesamt erhalten wir mit

$$\begin{aligned} \sum_{j=-N}^{+N} \gamma_j \exp(i j \omega t) &= \gamma_0 + \sum_{m=1}^N (\gamma_{-m} \exp(-i m \omega t) + \gamma_{+m} \exp(+i m \omega t)) \\ &= \gamma_0 + \sum_{m=1}^N (\gamma_{-m} + \gamma_{+m}) \cos(m \omega t) \\ &\quad + \sum_{m=1}^N (-\gamma_{-m} + \gamma_{+m}) i \sin(m \omega t) \\ &= \frac{1}{2} a_0 + \sum_{m=1}^N a_m \cos(m \omega t) + \sum_{m=1}^N b_m \sin(m \omega t) \end{aligned}$$

die behauptete Gleichheit der beiden Formeln für  $F_N$  (sowohl für  $N \in \mathbb{N}$  als auch im Grenzfall  $N = \infty$ ). Des Weiteren gelten die Orthogonalitätsrelationen

$$\begin{aligned} \frac{1}{T} \int_0^T \exp(i j \omega t) \overline{\exp(i n \omega t)} dt &= \frac{1}{T} \int_0^T \exp(i (j - n) \omega t) dt \\ &= \frac{1}{\omega T} \int_0^{2\pi} \exp(i (j - n) s) ds = \frac{2\pi}{\omega T} \delta_{jn} = \delta_{jn}, \end{aligned}$$

wobei wir in der Integralberechnung  $s = \omega t$  substituiert haben. Mit dieser Identität und der Euler-Formel können leicht die reellen Orthogonalitätsrelationen nachgeprüft werden. Außerdem können wir die Parsevalsche Gleichung

$$\frac{1}{T} \int_0^T |f(t)|^2 dt = \sum_{j=-\infty}^{\infty} |\gamma_j|^2$$

herleiten, die eine wesentliche Rolle in der Theorie der Fourier-Reihen spielt.

**Interpretation** Die komplexen Bausteine  $\exp(i j \omega t)$  bilden eine orthonormale Basis im unendlich-dimensionalen und komplexen Vektorraum der  $T$ -periodischen Funktionen, wobei das entsprechende Skalarprodukt zweier  $T$ -periodischer Funktionen  $u, v : \mathbb{R} \rightarrow \mathbb{C}$  durch

$$\langle u, v \rangle = \frac{1}{T} \int_0^T u(t) \overline{v(t)} dt$$

gegeben ist. Die Parsevalsche Ungleichung ist dann gerade die entsprechende Variante des Satzes von Pythagoras (vgl. dazu die analoge endlich-dimensionale Diskussion in Abschnitt §5.2).

### Bemerkung

1. Die komplexe Variante kann auch auf Funktionen  $f : \mathbb{R} \rightarrow \mathbb{C}$  angewendet werden.
2. In der Mathematik schreibt man statt  $\gamma_j$  auch  $\hat{f}_j$  und bezeichnet die doppelt-unendliche Folge  $(\hat{f}_j)_{j \in \mathbb{Z}}$  als die Fourier-Transformierte von  $f$ . Insbesondere entspricht jeder  $T$ -periodischen Funktion  $f : \mathbb{R} \rightarrow \mathbb{C}$  eine Funktion  $\hat{f} : \mathbb{Z} \rightarrow \mathbb{C}$  und umgekehrt. Dieses Konzept kann dann auch sehr viel allgemeiner gefasst werden. Schlagworte sind hier *diskrete oder kontinuierliche Fourier-Transformation, Fourier-Integrale, abstrakte harmonische Analysis*.
3. Es hat sich leider weder in der Mathematik noch in den Anwendungswissenschaften eine einheitliche Notation für Fourier-Reihen (und allgemeiner für die sogenannten Fourier-Methoden) herausgebildet, sondern Formeln, Sprechweisen, Normierungen usw. variieren in der Literatur zum Teil stark. Die zugrunde liegende Idee ist aber immer dieselbe: Eine Funktion/ein Signal wird in seine harmonischen Bausteine zerlegt.
4. Fourier-Methoden sind extrem wichtig, sowohl in der Mathematik als auch in den Natur- und Ingenieurwissenschaften. In der Quantenmechanik beschreibt die Fourier-Transformation zum Beispiel den Welle-Teilchen-Dualismus und die damit verbundene Heisenbergsche Unschärferelation.

**Konvergenz der Fourier-Reihe** Wir formulieren nun das wesentliche Konvergenzresultat, dessen Beweis allerdings im Rahmen dieser Vorlesungsreihe nicht gegeben werden kann.

**Theorem (Hauptsatz über Fourier-Reihen)** Sei  $f$  beschränkt und stückweise stetig differenzierbar, so dass in jeder Unstetigkeitsstelle  $t$  via

$$f(t \pm 0) := \lim_{\varepsilon \searrow 0} f(t \pm \varepsilon)$$

der linksseitige Grenzwert  $f(t - 0)$  und der rechtsseitige Grenzwert  $f(t + 0)$  existiert.

1. (Existenz des punktweisen Grenzwertes) Für jedes  $t \in \mathbb{R}$  gilt

$$F_N(t) \xrightarrow{N \rightarrow \infty} \frac{f(t - 0) + f(t + 0)}{2},$$

wobei  $f(t - 0) = f(t + 0) = f(t)$  in jeder Stetigkeitsstelle  $t \in \mathbb{R}$  gilt.

2. (Lokal gleichmäßige Konvergenz im Stetigkeitsbereich) Ist  $f$  stetig auf dem offenen Intervall  $I$ , so gilt

$$\sup_{t \in K} |F_N(t) - f(t)| \xrightarrow{N \rightarrow \infty} 0$$

für jedes kompakte Intervall  $K \subset I$ .

3. (Konvergenz im Integralsinn) Es gilt

$$\|F_n - f\|_p := \left( \int_0^T |F_n(t) - f(t)|^p dt \right)^{1/p} \xrightarrow{N \rightarrow \infty} 0$$

für jeden Exponenten  $p$  mit  $1 \leq p < \infty$ .

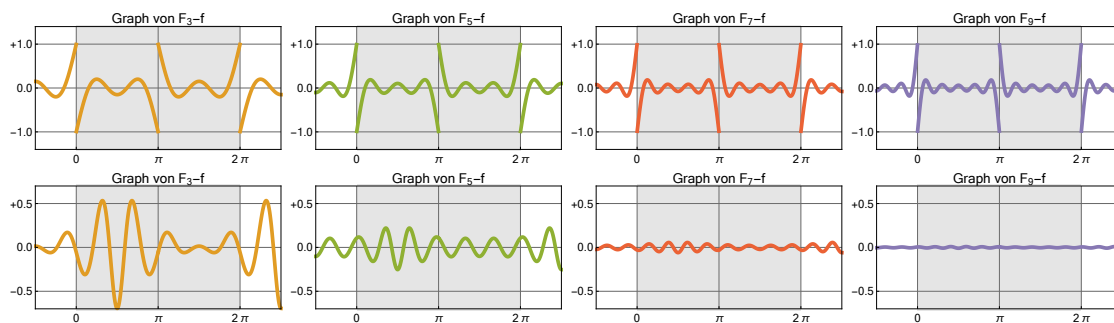
**Folgerung** Die Fourier-Reihe  $F_\infty$  ist an jeder Stelle  $t \in \mathbb{R}$  wohldefiniert mit

$$F_\infty(t) = \frac{f(t - 0) + f(t + 0)}{2}.$$

Außerdem gilt

$$\int_0^T |F_\infty(t) - f(t)| dt = 0$$

sowie  $F_\infty(t) = f(t)$  in jeder Stetigkeitsstelle  $t$ .



Der Approximationsfehler  $E_N = F_N - f$  als Funktion von  $t$  für die Beispiele 2 (oben) und 5 (unten). Beachte, dass die Konvergenz in Beispiel 2 deutlich schlechter als in Beispiel 5 ist, da die Funktion  $f$  in Beispiel 2 einen Sprung aufweist und in Beispiel 5 unendlich oft differenzierbar ist.

	$\ F_3 - f\ _1$	$\ F_5 - f\ _1$	$\ F_7 - f\ _1$	$\ F_9 - f\ _1$
Beispiel 2	1.371	1.026	0.829	0.700
Beispiel 5	1.147	0.575	0.145	0.024

**Bemerkung** Der Approximationsfehler  $E_N := F_N - f$  ist zunächst auch eine  $T$ -periodische Funktion und konvergiert für  $N \rightarrow \infty$  auf drei verschiedene Art und Weisen (die Konvergenz von Funktionenfolgen ist subtiler als die Konvergenz von Zahlenfolgen, weil es verschiedene, nicht-äquivalente Konvergenzbegriffe gibt).

**Differentiation** Ist  $f$  stetig differenzierbar, so gilt auch

$$\dot{f}(t) = \sum_{m=1}^{\infty} b_m m \omega \cos(m \omega t) - \sum_{m=1}^{\infty} a_m m \omega \sin(m \omega t),$$

d.h. Fourier-Reihen dürfen gliedweise differenziert werden und die Fourier-Koeffizienten der Ableitung  $\dot{f}$  können aus denen für  $f$  berechnet werden. Analog gilt

$$\ddot{f}(t) = - \sum_{m=1}^{\infty} a_m m^2 \omega^2 \cos(m \omega t) - \sum_{m=1}^{\infty} b_m m^2 \omega^2 \sin(m \omega t),$$

sofern die zweite Ableitung  $\ddot{f}$  überall existiert. Ganz allgemein kann man sagen: Aus den Fourier-Koeffizienten von  $f$  können sehr leicht die Fourier-Koeffizienten aller Ableitungen von  $f$  berechnet werden (sofern diese existieren).

*Bemerkung:* Auch Stammfunktionen können gliedweise berechnet werden.

**\*Konvergenzgeschwindigkeit** Es gilt die Faustregel: Je mehr Ableitungen von  $f$  existieren, umso besser ist die Konvergenz für  $N \rightarrow \infty$ . Quantitative Abschätzungen können dabei aus der Parsevalschen Gleichung mit Hilfe des komplexen Kalküls gewonnen werden. Ist zum Beispiel  $f$  insgesamt  $K$ -mal stetig differenzierbar, so gilt

$$\sum_{m=1}^{\infty} (\varrho_m \cdot m^K)^2 = C_K < \infty, \quad \varrho_m = \sqrt{|a_m|^2 + |b_m|^2}$$

für eine Konstante  $C_K$ , die von  $F$  und  $K$  abhängt. Insbesondere müssen  $\varrho_m$  — bzw. sowohl  $|a_m|$  als auch  $|b_m|$  — im Limes  $m \rightarrow \infty$  mindestens so schnell wie  $\sqrt{C_K}/m^K$  abklingen, und diese Schranke ist asymptotisch für große  $K$  besser als für kleine  $K$ .

**\*Gibbs-Phänomen** Wir wollen das problematische Konvergenzverhalten in der Nähe von Unstetigkeitsstellen noch genauer untersuchen und betrachten dazu das prototypische Beispiel der Sägezahnfunktion mit Fehler

$$E_N(t) = F_N(t) - f(t) = \frac{1}{2}(t - \pi) + \sum_{n=1}^N \frac{\sin(nt)}{n}.$$

Die Ableitung kann nun sowohl durch

$$\dot{E}_N(t) = \frac{1}{2} + \sum_{n=1}^N \cos(nt)$$

als auch durch

$$\dot{E}_N(t) = \frac{\sin\left(\left(N + \frac{1}{2}\right)t\right)}{2 \sin\left(\frac{1}{2}t\right)}$$

berechnet werden, wobei die Gleichheit der Formeln aus den Additionstheoremen hergeleitet werden kann. Insbesondere kann man zeigen, dass

$$t_N = \frac{\pi}{N + \frac{1}{2}}$$

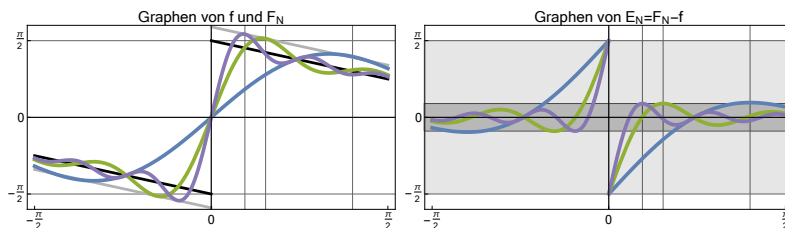
das erste lokale Extremum von  $F_N$  in Intervall  $(0, \infty)$  ist, und wir erhalten

$$E_N(t_N) = \int_0^{t_N} \dot{E}_N(t) dt - \frac{\pi}{2} = \int_0^{\pi} \frac{\sin(u)}{u} \frac{u/(2N+1)}{\sin(u/(2N+1))} du - \frac{\pi}{2},$$

wobei wir die Substitution  $u = (N + \frac{1}{2})t$  verwendet haben. Nach l'Hospital (für festes  $u$  angewendet auf die  $N$ -Abhängigkeit) gilt nun

$$E_N(t_N) \xrightarrow{N \rightarrow \infty} \int_0^{\pi} \frac{\sin(u)}{u} du - \frac{\pi}{2} = \text{Si}(\pi) - \frac{\pi}{2} \approx 0.1789 f(0+0).$$

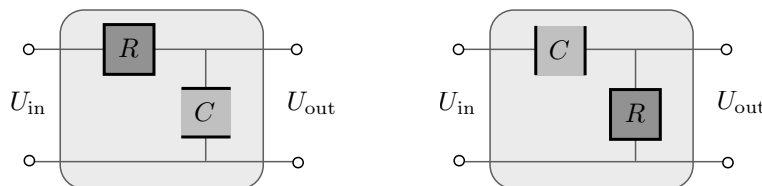
Oder anders gesagt: Die Approximationen  $F_N$  überschwingen für jedes  $N$  die Funktion  $f$ , wobei für große  $N$  die Stärke dieses Überschwingens ca. 18% beträgt. Diese Erkenntnis wird als Gibbs-Phänomen bezeichnet. Es tritt bei jedem  $N$  auf, und zwar nicht nur für die Sägezahnfunktion, sondern analog auch für jede andere Funktion  $f$  an jeder ihrer Unstetigkeitsstellen.



Das Gibbs-Phänomen in Beispiel 1. Die Höhe des dunklen Streifens ist 18% der Höhe des hellen Streifens.

Vorlesung 38, 05. Februar 2019

## Anwendung RC-Tiefpass



*Links:* Ein RC-Tiefpass oder Tiefpass 1. Ordnung. Die mathematischen Eigenschaften werden im Text, die elektrischen zum Beispiel unter *Tiefpass* auf WIKIPEDIA diskutiert. *Rechts:* Der Hochpass kann mit ähnlichen Formeln beschrieben werden.

**mathematisches Modell** Man kann ein idealisiertes Bauelement durch die lineare Differentialgleichung

$$\tau \dot{U}_{\text{out}}(t) + U_{\text{out}}(t) = U_{\text{in}}(t), \quad \tau = R \cdot C$$



beschreiben, wobei  $R$  und  $C$  gegebene Konstanten sind (Widerstand und Kapazität, zum Beispiel in SI-Einheiten). Außerdem ist die Eingangsspannung  $U_{\text{in}}(t)$  eine bekannte, die Ausgangsspannung  $U_{\text{out}}$  jedoch eine zu bestimmende Funktion der Zeit  $t$ . Wir lernen die Theorie der Differentialgleichungen erst in *Mathe II* kennen, können aber hier schon wesentliche Resultate ableiten.

**harmonischer Input** Der Ausgangspunkt ist die Beobachtung

$$U_{\text{in}}(t) = \exp(\mathrm{i} \Omega t) \quad \Longrightarrow \quad U_{\text{out}}(t) = \frac{1}{1 + \mathrm{i} \tau \Omega} \exp(\mathrm{i} \Omega t),$$

d.h. wenn der Input eine harmonische Funktion mit Frequenz  $\Omega$  ist, so ist auch der Output eine harmonische Funktion mit derselben Frequenz, aber wegen

$$\frac{1}{1 + \mathrm{i} \tau \Omega} = \frac{1}{\sqrt{1 + \tau^2 \Omega^2}} \exp(-\mathrm{i} \arctan(\tau \Omega)).$$

mit gedämpfter Amplitude und Phasenverschiebung. Dieses sehr einfache Gesetz über den Zusammenhang zwischen harmonischem Input und harmonischem Output kann zum Beispiel physikalisch begründet werden. Alternativ können wir einfach nachrechnen, dass die Formeln eine Lösung der Differentialgleichung liefern:

$$\tau \dot{U}_{\text{out}}(t) + U_{\text{out}}(t) = \tau \frac{\mathrm{i} \Omega}{1 + \mathrm{i} \tau \Omega} \exp(\mathrm{i} \Omega t) + \frac{1}{1 + \mathrm{i} \tau \Omega} \exp(\mathrm{i} \Omega t) = U_{\text{in}}(t).$$

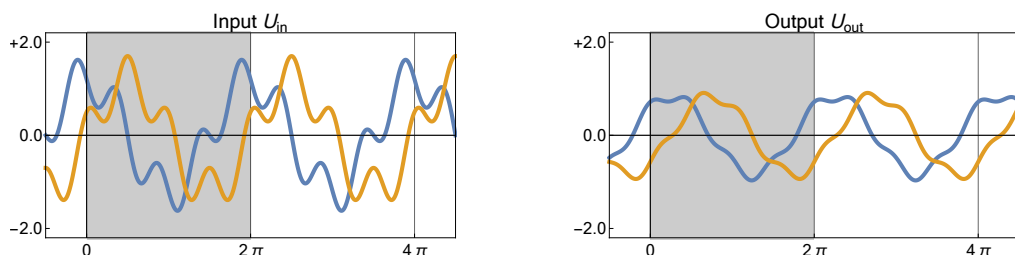
**$T$ -periodischer Input** Mit der obigen Beobachtung sowie der Theorie der Fourier-Reihen können wir nun alle  $T$ -periodischen Inputfunktionen behandeln, wobei wir gleich im Komplexen rechnen wollen. Ist die Eingangsspannung durch die Fourier-Reihe

$$U_{\text{in}}(t) = \sum_{j=-\infty}^{\infty} \hat{U}_{\text{in},j} \exp(\mathrm{i} j \omega t)$$

gegeben, so können wir durch

$$U_{\text{out}}(t) = \sum_{j=-\infty}^{\infty} \hat{U}_{\text{out},j} \exp(\mathrm{i} j \omega t), \quad \hat{U}_{\text{out},j} = \frac{\hat{U}_{\text{in},j}}{1 + \mathrm{i} j \tau \omega}$$

die entsprechende Ausgangsspannung erzeugen, wobei wir das obige Gesetz für  $\Omega = j \omega$  ausgewertet haben. Durch gliedweise Differentiation kann außerdem leicht die Probe gemacht werden, d.h. wir können wieder nachrechnen, dass  $U_{\text{out}}$  in der Tat die Differentialgleichung erfüllt.



Beispiel mit  $T = 2\pi$  und  $U_{\text{in}}(t) = 1.2 e^{\mathrm{i}t} + \mathrm{i} 0.5 e^{\mathrm{i}3t}$ , wobei Real- und Imaginärteil in Blau und Gelb erscheinen.

Lösungsstrategie: Die obigen Formeln können auch als Handlungsanweisung formuliert werden:

1. Berechne aus  $U_{\text{in}}$  die komplexen Fourier-Koeffizienten  $\hat{U}_{\text{in},j}$  mit Hilfe der Integralformel

$$\hat{U}_{\text{in},j} = \frac{1}{T} \int_0^T U_{\text{in}}(t) \exp(-i j \omega t) dt.$$

2. Berechne  $\hat{U}_{\text{out},j}$  für jedes  $j$  aus  $\hat{U}_{\text{in},j}$  durch Multiplikation mit  $1/(1 + i \tau \omega)$ .
3. Setze  $U_{\text{out}}$  durch Summation bzw. Reihenbildung aus seinen Fourier-Bausteinen zusammen.

Das reelle Analogon zu den komplexen Formeln ist übrigens

$$a_{\text{out},m} = \frac{a_{\text{in},m} - \tau m \omega b_{\text{in},m}}{1 + \tau^2 m^2 \omega^2}, \quad b_{\text{out},m} = \frac{\tau m \omega a_{\text{in},m} + b_{\text{in},m}}{1 + \tau^2 m^2 \omega^2}$$

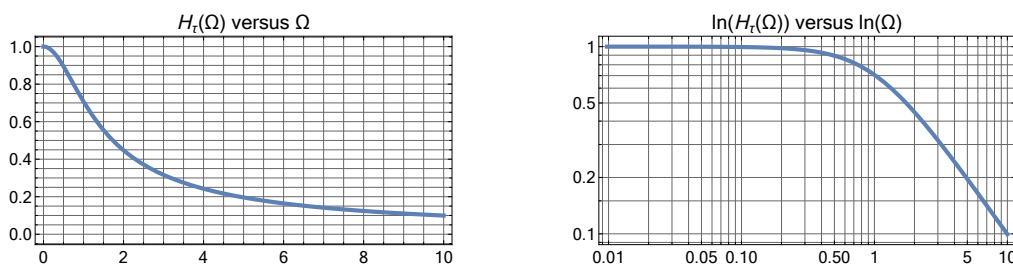
und kann auch als

$$\frac{\varrho_{\text{out},m}}{\varrho_{\text{in},m}} = \frac{1}{\sqrt{1 + \tau^2 m^2 \omega^2}}, \quad \theta_{\text{out},m} - \theta_{\text{in},m} = -\arctan(\tau m \omega),$$

geschrieben werden, wobei man diese Formeln den Amplitudengang bzw. Phasengang des Tiefpasses nennt. Der Graph der Funktion

$$H_\tau(\Omega) = \frac{1}{\sqrt{1 + \tau^2 \Omega^2}}$$

wird dabei üblicherweise in einem doppelt logarithmischen Plot gezeichnet.



Normaler and doppelt-logarithmischer Plot der Funktion  $H_\tau$  (für  $\tau = 1$ ), die den Amplitudengang beschreibt.

## 11.3 Varianten der Fourier-Reihe

**Vorbemerkung** Es gibt Varianten der Fourier-Reihe, die wichtig für viele Anwendungen sind und die wir abschließend kurz vorstellen wollen. Wir beschränken uns dabei auf den Fall reellwertiger Funktionen, obwohl analoge Aussagen auch für komplexwertige Funktionen getroffen werden können.

**Drei verschiedene Fortsetzungen** Eine gegebene Funktion  $f : [0, T] \rightarrow \mathbb{R}$  kann auf verschiedene Arten zu einer  $2T$ -periodischen Funktion auf ganz  $\mathbb{R}$  fortgesetzt werden, wobei wir akzeptieren, dass an den Vielfachen von  $T$  Unstetigkeiten bzw. Definitionsprobleme entstehen können.

1. Wir setzen  $f$  zu einer  $T$ -periodischen Funktion fort und können  $f$  durch die Fourier-Reihe aus dem letzten Abschnitt approximieren. Beachte, dass jede  $T$ -periodische Funktion auch  $2T$ -periodisch ist.
2. Wir setzen  $f$  im ersten Schritt zunächst zu einer ungeraden Funktion auf  $[-T, T]$  fort und anschließend zu einer  $2T$ -periodischen Funktion auf ganz  $\mathbb{R}$  fort. Die Fourier-Reihe dieser  $2T$ -periodischen Fortsetzung kann als Fourier-Sinus-Reihe

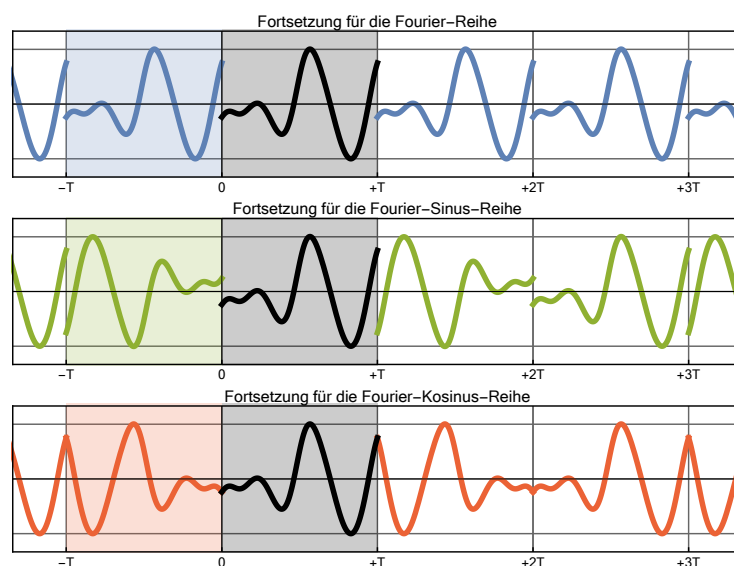
$$f(t) = \sum_{m=1}^{\infty} \beta_m \sin\left(\frac{1}{2} m \omega\right), \quad \beta_m = \frac{2}{T} \int_0^T f(x) \sin\left(\frac{1}{2} m \omega t\right) dt$$

geschrieben werden, da die Kosinus-Beiträge aus Paritätsgründen verschwinden. Beachte aber, dass wir hier nun auch halbzahlige Sinus-Bausteine auf dem Intervall  $[0, T]$  finden, da wir ja die Fourier-Reihe für  $2T$ -periodische Funktionen benutzen. Insbesondere liefert der erste ganze Sinus-Baustein auf dem Doppelintervall  $[-T, +T]$  einen halben Sinus auf  $[0, T]$ .

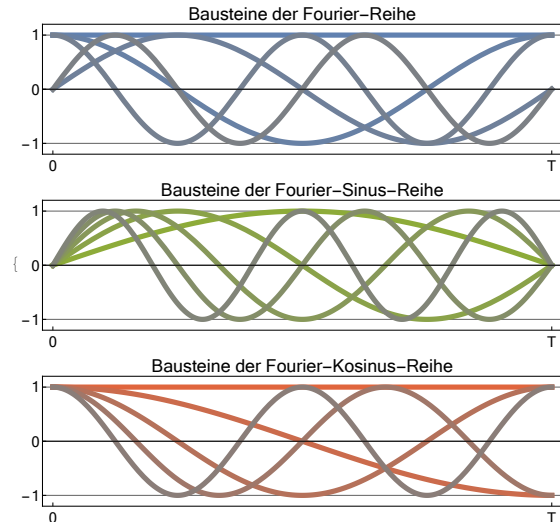
3. Setzen wir jedoch  $f$  zunächst zu einer geraden und anschließend zu einer  $2T$ -periodischen Funktion fort, so erhalten wir die Fourier-Kosinus-Reihe

$$f(t) = \frac{1}{2} \alpha_0 + \sum_{m=1}^{\infty} \alpha_m \cos\left(\frac{1}{2} m \omega\right), \quad \alpha_m = \frac{2}{T} \int_0^T f(x) \cos\left(\frac{1}{2} m \omega t\right) dt$$

als Darstellung mit halb- und ganzzahligen Kosinus-Bausteinen.

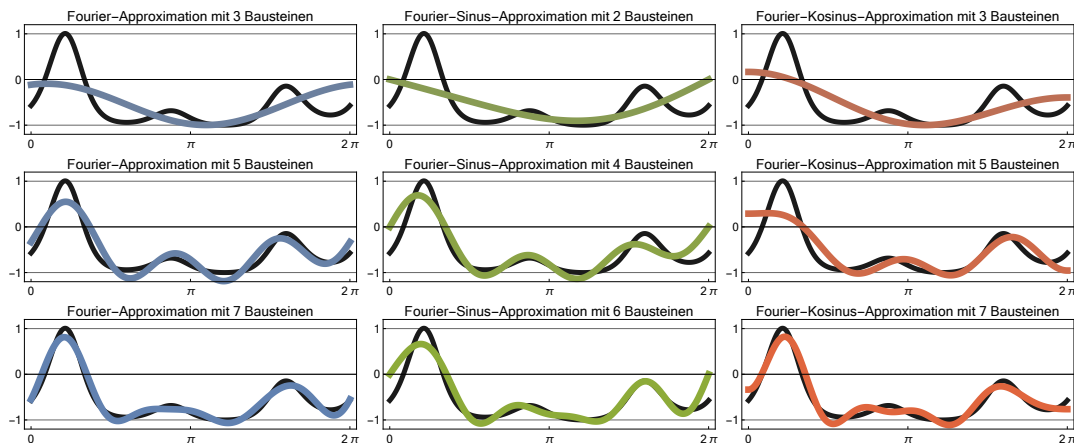


Eine Funktion (schwarz) auf dem Intervall  $[0, T]$  kann wie im Text beschrieben auf drei verschiedene Arten zu einer  $2T$ -periodischen Funktion fortgesetzt werden.

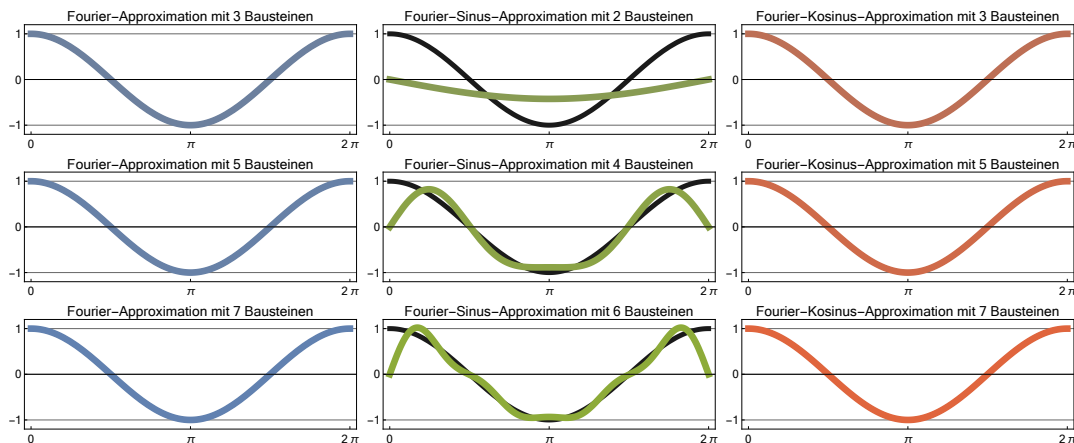


Die Bausteine (bzw. Basis-Funktionen) der drei Reihen auf dem Intervall  $[0, T]$ . Bei der Fourier-Sinus-Reihe bzw. der Fourier-Kosinus-Reihe treten sowohl halbzahlige als auch ganzzahlige Sinus- bzw. Kosinus-Funktionen auf, während bei der Fourier-Reihe nur ganzzahlige Funktionen, aber dafür aus beiden Familien auftreten.

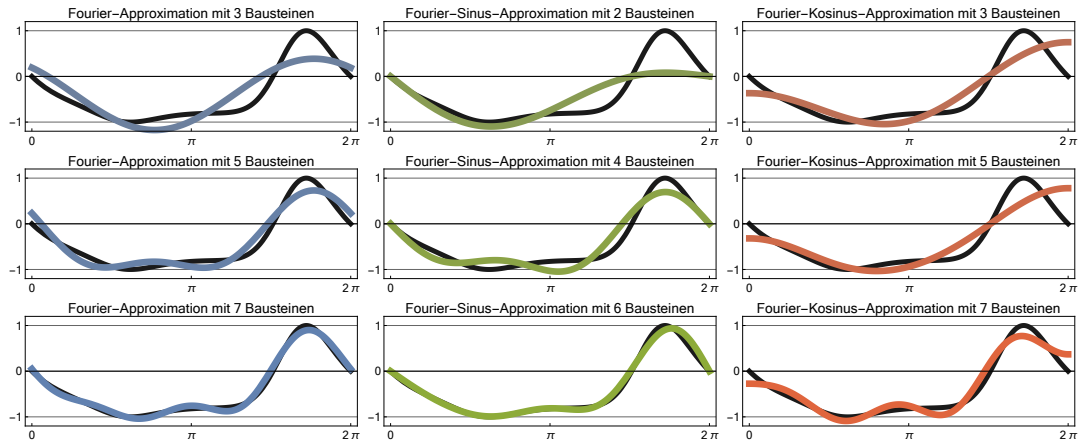
**Folgerung** Jede Funktion auf  $[0, T]$  kann auf drei verschiedene Weisen als konvergente trigonometrische Reihe geschrieben werden, wobei jeweils Gibbs-Phänomene an den Randpunkten  $t = 0$  und  $t = T$  auftreten können.



Beispiel A:  $T = 2\pi$  und  $f(t) = \exp(\cos(t)) + \sin(3t)$ .



Beispiel B:  $T = 2\pi$  und  $f(t) = \cos(t)$ . Beachte die Gibbs-Phänomene, die bei der Sinus-Reihe am Intervallrand auftreten.



Beispiel C: Es gelten die Dirichlet-Randbedingungen  $f(0) = f(T) = 0$ , die aber nur bei den Fourier-Sinus-Approximationen erfüllt werden.

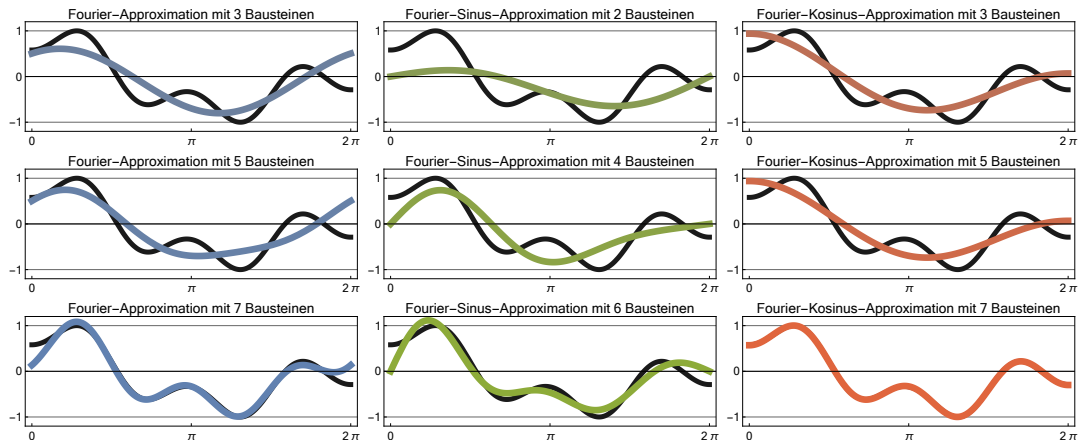
**Bemerkung** Die Kosinus-Funktion kann aus Sinus-Bausteinen zusammengesetzt werden. Mit  $T = 2\pi$  ergeben sich die Koeffizienten

$$\beta_m = -\frac{2m((-1)^m - 1)}{\pi(m^2 - 4)}$$

und damit die Darstellung

$$\cos(t) = -\frac{4 \sin\left(\frac{t}{2}\right)}{3\pi} + \frac{12 \sin\left(\frac{3t}{2}\right)}{5\pi} + \frac{20 \sin\left(\frac{5t}{2}\right)}{21\pi} + \dots$$

für jedes  $t$  mit  $0 < t < 2\pi$ . Für  $t = 0$  und  $t = 2\pi$  gilt diese Formel so nicht, da dort die fortgesetzte Funktion unstetig ist und die Fourier-Sinus-Reihe den Wert 0 annimmt (Gibbs-Phänomen). Siehe auch die Bilder für Beispiel B.



Beispiel D: Die Neumann-Randbedingungen  $\dot{f}(0) = \dot{f}(T) = 0$  werden nur von den Kosinus-Bausteinen respektiert.

### Anwendungsregel

1. Gilt  $f(0) = f(T) = 0$  (Dirichlet-Randbedingungen), so ist die Fourier-Sinus-Reihe am besten geeignet.
2. Gilt  $\dot{f}(0) = \dot{f}(T) = 0$  (Neumann-Randbedingungen), so bietet sich die Fourier-Kosinus-Reihe an.
3. In allen anderen Fällen wird meist die Fourier-Reihe verwendet.



# Anhang A

## Wiederholung: Lineare Abbildungen und Matrizen

Vorlesung 39, 07. Februar 2020

**Vorbemerkung** Wir rechnen in diesem Anhang reell, aber alles gilt analog auch im Komplexen.

### Basiswechsel

**Setting** In einem (reellen) Vektorraum  $V$  der Dimension  $n = \dim V$  seien zwei (verschiedene) Basen gegeben:

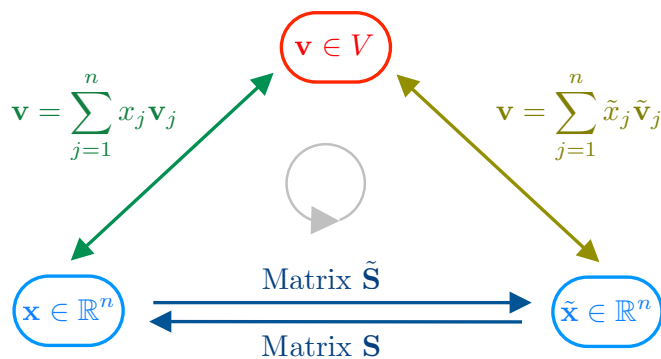
(Basis 1)  $\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n$

(Basis 2)  $\tilde{\mathbf{v}}_1, \tilde{\mathbf{v}}_2, \dots, \tilde{\mathbf{v}}_n$

Für jeden Vektor  $\mathbf{v} \in V$  existieren damit zwei eindeutige Koordinatenvektoren  $\mathbf{x} \in \mathbb{R}^n$  und  $\tilde{\mathbf{x}} \in \mathbb{R}^n$ , so dass

$$\mathbf{v} = \sum_{j=1}^n x_j \cdot \mathbf{v}_j = \sum_{j=1}^n \tilde{x}_j \cdot \tilde{\mathbf{v}}_j.$$

Beachte, dass für gegebenes  $\mathbf{v}$  jeder der entsprechenden Koordinaten  $x_j$  bzw.  $\tilde{x}_j$  im Allgemeinen durch ein lineares Gleichungssystem zu bestimmen ist.



Das Diagramm zum Basiswechsel aus Kapitel 5.

**Prinzip/Theorem** Mit den obigen Bezeichnungen gilt:

1. Es gibt zwei quadratische Basiswechsel-Matrizen (auch Übergangsmatrizen genannt) und  $\mathbf{S}$  und  $\tilde{\mathbf{S}}$ , so dass stets

$$\tilde{\mathbf{x}} = \tilde{\mathbf{S}} \cdot \mathbf{x}, \quad \mathbf{x} = \mathbf{S} \cdot \tilde{\mathbf{x}}$$

gilt.

2. Die  $(n, n)$ -Matrizen  $\mathbf{S}$  und  $\tilde{\mathbf{S}}$  sind invers zueinander, d.h. es gilt  $\mathbf{S} \cdot \tilde{\mathbf{S}} = \mathbf{I}_n = \tilde{\mathbf{S}} \cdot \mathbf{S}$ .
3. Die Einträge der Basiswechsel-Matrizen können durch die Formeln

$$\tilde{\mathbf{v}}_k = \sum_{j=1}^n s_{jk} \cdot \mathbf{v}_j, \quad \mathbf{v}_k = \sum_{j=1}^n \tilde{s}_{jk} \cdot \tilde{\mathbf{v}}_j$$

bestimmt werden.

4. Ist der zu Grunde liegende Vektorraum  $V$  der  $\mathbb{R}^n$ , so kann das folgende Kochrezept verwendet werden: Wird die erweiterte Koeffizientenmatrix

$$\left( \begin{array}{ccc|ccc} | & & | & | & & | \\ \mathbf{v}_1 & \dots & \mathbf{v}_n & \tilde{\mathbf{v}}_1 & \dots & \tilde{\mathbf{v}}_n \\ | & & | & | & & | \end{array} \right) \text{ bzw. } \left( \begin{array}{ccc|ccc} | & & | & | & & | \\ \tilde{\mathbf{v}}_1 & \dots & \tilde{\mathbf{v}}_n & \mathbf{v}_1 & \dots & \mathbf{v}_n \\ | & & | & | & & | \end{array} \right)$$

solange mit dem Gauß-Jordan-Verfahren transformiert, bis links vom Mittelstrich die Einheitsmatrix steht, so kann rechts die Matrix  $\mathbf{S}$  bzw.  $\tilde{\mathbf{S}}$  abgelesen werden.

*Bemerkung:* Man kann natürlich auch die erste Koeffizientenmatrix solange umformen bis rechts die Einheitsmatrix und links  $\tilde{\mathbf{S}}$  steht.

**Beispiel 1.1** Für  $n = 3$  und  $V = \mathbb{R}^3$  betrachten wir die kanonische Basis

$$\mathbf{v}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{v}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \mathbf{v}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

sowie die nicht-kanonische Basis

$$\tilde{\mathbf{v}}_1 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \quad \tilde{\mathbf{v}}_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad \tilde{\mathbf{v}}_3 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}.$$

Wir können nun die Basisvektoren  $\tilde{\mathbf{v}}_j$  als Linearkombination der  $\mathbf{v}_i$  darstellen und umgekehrt. Durch Scharfes Hinsehen (bzw. nach Lösen entsprechender Gleichungssysteme) erhalten wir

$$\begin{aligned} \tilde{\mathbf{v}}_1 &= \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = 0 \cdot \mathbf{v}_1 + 1 \cdot \mathbf{v}_2 + 1 \cdot \mathbf{v}_3 \\ \tilde{\mathbf{v}}_2 &= \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = 1 \cdot \mathbf{v}_1 + 0 \cdot \mathbf{v}_2 + 1 \cdot \mathbf{v}_3 \\ \tilde{\mathbf{v}}_3 &= \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} = 1 \cdot \mathbf{v}_1 + 1 \cdot \mathbf{v}_2 + 0 \cdot \mathbf{v}_3 \end{aligned}$$



bzw.

$$\begin{aligned}\mathbf{v}_1 &= \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = -\frac{1}{2} \cdot \tilde{\mathbf{v}}_1 + \frac{1}{2} \cdot \tilde{\mathbf{v}}_2 + \frac{1}{2} \cdot \tilde{\mathbf{v}}_3 \\ \mathbf{v}_2 &= \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = +\frac{1}{2} \cdot \tilde{\mathbf{v}}_1 - \frac{1}{2} \cdot \tilde{\mathbf{v}}_2 + \frac{1}{2} \cdot \tilde{\mathbf{v}}_3 \\ \mathbf{v}_3 &= \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = +\frac{1}{2} \cdot \tilde{\mathbf{v}}_1 + \frac{1}{2} \cdot \tilde{\mathbf{v}}_2 - \frac{1}{2} \cdot \tilde{\mathbf{v}}_3\end{aligned}$$

und können die Matrizen  $\mathbf{S}$  bzw.  $\tilde{\mathbf{S}}$  jeweils spaltenweise aus den einzelnen Gleichungen zusammensetzen. Im konkreten Fall ergibt sich

$$\mathbf{S} = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \quad \text{bzw.} \quad \tilde{\mathbf{S}} = \frac{1}{2} \begin{pmatrix} -1 & 1 & 1 \\ 1 & -1 & 1 \\ 1 & 1 & -1 \end{pmatrix}.$$

Wir können außerdem durch einfache und direkte Rechnungen nachprüfen, dass in der Tat  $\tilde{\mathbf{S}} \cdot \mathbf{S} = \mathbf{S} \cdot \tilde{\mathbf{S}} = \mathbf{I}_3$  gilt. Das Kochrezept liefert natürlich dieselben Ergebnisse, zum Beispiel via

$$\begin{aligned}\left( \begin{array}{ccc|ccc} 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 \end{array} \right) &\cong \left( \begin{array}{ccc|ccc} 0 & 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 \end{array} \right) \cong \left( \begin{array}{ccc|ccc} 0 & 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & -1 & 0 & -1 & 1 \end{array} \right) \\ &\cong \left( \begin{array}{ccc|ccc} -1 & 0 & 1 & 1 & 0 & -1 \\ 1 & 0 & 0 & 0 & 1 & 1 \\ 0 & 1 & -1 & 0 & -1 & 1 \end{array} \right) \cong \left( \begin{array}{ccc|ccc} -1 & 0 & 1 & 1 & 0 & -1 \\ 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & -1 & 0 & 0 & 2 \end{array} \right) \\ &\cong \left( \begin{array}{ccc|ccc} -1 & 0 & 1 & 1 & 0 & -1 \\ 1 & 0 & 0 & 0 & 1 & 1 \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & 0 & 0 & 1 \end{array} \right) \cong \left( \begin{array}{ccc|ccc} -1 & 0 & 1 & 1 & 0 & -1 \\ \frac{1}{2} & -\frac{1}{2} & +\frac{1}{2} & 0 & 1 & 0 \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & 0 & 0 & 1 \end{array} \right) \\ &\cong \left( \begin{array}{ccc|ccc} -\frac{1}{2} & \frac{1}{2} & \frac{1}{2} & 1 & 0 & 0 \\ \frac{1}{2} & -\frac{1}{2} & +\frac{1}{2} & 0 & 1 & 0 \\ \frac{1}{2} & \frac{1}{2} & -\frac{1}{2} & 0 & 0 & 1 \end{array} \right),\end{aligned}$$

wobei wir das Gauß-Jordan-Verfahren benutzt haben, um rechts die Einheitsmatrix zu erzeugen.

**Beispiel 1.2** Wie wollen noch ein weiteres, zweidimensionales Beispiel rechnen, aber schreiben für spätere Zwecke  $W$  statt  $V$ ,  $\mathbf{w}$  statt  $\mathbf{v}$ ,  $\mathbf{y}$  statt  $\mathbf{x}$ ,  $\mathbf{R}$  statt  $\mathbf{S}$  usw. In diesem Beispiel gilt  $W = \mathbb{R}^2$  und wir betrachten die zwei Basen

$$\mathbf{w}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \mathbf{w}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \tilde{\mathbf{w}}_1 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad \tilde{\mathbf{w}}_2 = \begin{pmatrix} -1 \\ 0 \end{pmatrix}.$$

Analog zu oben bemerken wir

$$\begin{aligned}\tilde{\mathbf{w}}_1 &= \begin{pmatrix} 1 \\ -1 \end{pmatrix} = +1 \cdot \mathbf{w}_1 - 1 \cdot \mathbf{w}_2 \\ \tilde{\mathbf{w}}_2 &= \begin{pmatrix} -1 \\ 0 \end{pmatrix} = -1 \cdot \mathbf{w}_1 + 0 \cdot \mathbf{w}_2\end{aligned}$$

bzw.

$$\begin{aligned}\mathbf{w}_1 &= \begin{pmatrix} 1 \\ 0 \end{pmatrix} = +0 \cdot \tilde{\mathbf{w}}_1 - 1 \cdot \tilde{\mathbf{w}}_2 \\ \mathbf{w}_2 &= \begin{pmatrix} 0 \\ 1 \end{pmatrix} = -1 \cdot \tilde{\mathbf{w}}_1 - 1 \cdot \tilde{\mathbf{w}}_2\end{aligned}$$

und können nun die Basiswechsel-Matrizen

$$\mathbf{R} = \begin{pmatrix} 1 & -1 \\ -1 & 0 \end{pmatrix}, \quad \tilde{\mathbf{R}} = \begin{pmatrix} 0 & -1 \\ -1 & -1 \end{pmatrix}$$

ablesen, wobei  $\mathbf{R}$  und  $\tilde{\mathbf{R}}$  wieder invers zueinander sind. Insbesondere gilt

$$\sum_{j=1}^n y_j \cdot \mathbf{w}_j = \sum_{j=1}^n \tilde{y}_j \cdot \tilde{\mathbf{w}}_j \quad \iff \quad \tilde{\mathbf{y}} = \tilde{\mathbf{R}} \cdot \mathbf{y}, \quad \mathbf{y} = \mathbf{R} \cdot \tilde{\mathbf{y}}.$$

**Beispiel 1.3** Wir betrachten schließlich den (abstrakten) Vektorraum

$$\Pi_3 := \{ \text{reelle Polynome in der Variable } t \text{ vom Grad } \leq 3 \}$$

für den es neben der monomialen Basis

$$p_1(t) = 1, \quad \tilde{p}_2(t) = t, \quad p_3(t) = t^2, \quad \tilde{p}_4(t) = t^3$$

viele weitere Basen gibt, zum Beispiel

$$\tilde{p}_1(t) = 1, \quad \tilde{p}_2(t) = t - 1, \quad \tilde{p}_3(t) = (t - 2)^2, \quad \tilde{p}_4(t) = (t - 3)^3.$$

Insbesondere gilt  $\dim \Pi_3 = 4$ , aber wir benutzen hier (wie üblich bei Funktionenräumen) nicht die Fettschreibweise, um Elemente aus  $\Pi_3$  zu kennzeichnen. Für jedes Element von  $\Pi_3$  – d.h. für jedes Polynom  $p$  vom Grad  $\leq 3$  – gibt es nun Koeffizientenvektoren  $\mathbf{c} \in \mathbb{R}^4$  und  $\tilde{\mathbf{c}} \in \mathbb{R}^4$ , so dass zum einen

$$p(t) = c_1 \cdot p_1(t) + c_2 \cdot p_2(t) + c_3 \cdot p_3(t) + c_4 \cdot p_4(t)$$

und zum anderen

$$p(t) = \tilde{c}_1 \cdot \tilde{p}_1(t) + \tilde{c}_2 \cdot \tilde{p}_2(t) + \tilde{c}_3 \cdot \tilde{p}_3(t) + \tilde{c}_4 \cdot \tilde{p}_4(t)$$

für alle  $t \in \mathbb{R}$  gilt. Wegen

$$\begin{aligned}\tilde{p}_1(t) &= 1 = +1 \cdot 1 + 0 \cdot t + 0 \cdot t^2 + 0 \cdot t^3 \\ \tilde{p}_2(t) &= t - 1 = -1 \cdot 1 + 1 \cdot t + 0 \cdot t^2 + 0 \cdot t^3 \\ \tilde{p}_3(t) &= (t - 2)^2 = +4 \cdot 1 - 4 \cdot t + 1 \cdot t^2 + 0 \cdot t^3 \\ \tilde{p}_4(t) &= (t - 3)^3 = -27 \cdot 1 + 27 \cdot t - 9 \cdot t^2 + 1 \cdot t^3\end{aligned}$$

und

$$\begin{aligned}p_1(t) &= 1 = +1 \cdot 1 + 0 \cdot (t - 1) + 0 \cdot (t - 2)^2 + 0 \cdot (t - 3)^3 \\ p_2(t) &= t = +1 \cdot 1 + 1 \cdot (t - 1) + 0 \cdot (t - 2)^2 + 0 \cdot (t - 3)^3 \\ p_3(t) &= t^2 = +0 \cdot 1 + 4 \cdot (t - 1) + 1 \cdot (t - 2)^2 + 0 \cdot (t - 3)^3 \\ p_4(t) &= t^3 = +0 \cdot 1 + 9 \cdot (t - 1) + 9 \cdot (t - 2)^2 + 1 \cdot (t - 3)^3\end{aligned}$$

können wir nun die Basiswechsel-Matrizen hinschreiben. Insbesondere gilt in diesem Beispiel immer

$$\begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 4 & -27 \\ 0 & 1 & -4 & 27 \\ 0 & 0 & 1 & -9 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} \tilde{c}_1 \\ \tilde{c}_2 \\ \tilde{c}_3 \\ \tilde{c}_4 \end{pmatrix}, \quad \begin{pmatrix} \tilde{c}_1 \\ \tilde{c}_2 \\ \tilde{c}_3 \\ \tilde{c}_4 \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 4 & 9 \\ 0 & 0 & 1 & 9 \\ 0 & 0 & 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} c_1 \\ c_2 \\ c_3 \\ c_4 \end{pmatrix}.$$

## Matrixdarstellung linearer Abbildungen

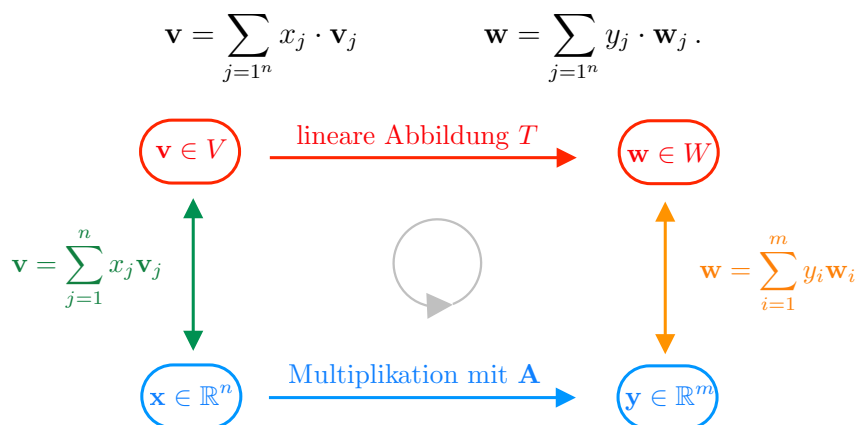
**Setting** Gegeben sei eine lineare Abbildung  $T$  zwischen zwei Vektorräumen  $V$  und  $W$  mit  $n = \dim V$  und  $m = \dim W$ , wobei  $T$  genau dann linear ist, wenn

$$T(\lambda \cdot \mathbf{v} + \tilde{\lambda} \cdot \tilde{\mathbf{v}}) = \lambda \cdot T(\mathbf{v}) + \tilde{\lambda} \cdot T(\tilde{\mathbf{v}})$$

für alle Vektoren  $\mathbf{v}, \tilde{\mathbf{v}} \in V$  und alle Skalare  $\lambda, \tilde{\lambda} \in \mathbb{R}$  gilt. Wir betrachten außerdem zwei Basen

$$\mathbf{v}_1, \mathbf{v}_2, \dots, \mathbf{v}_n \quad (\text{von } V), \quad \mathbf{w}_1, \mathbf{w}_2, \dots, \mathbf{w}_m \quad (\text{von } W).$$

Damit gibt es für jedes  $\mathbf{v} \in V$  genau ein  $\mathbf{x} \in \mathbb{R}^n$  sowie für jedes  $\mathbf{w} \in W$  genau ein  $\mathbf{y} \in \mathbb{R}^m$  mit



Das Diagramm zur Basisdarstellung aus Kapitel 5.

**Prinzip/Theorem** Mit den obigen Bezeichnungen gilt:

1. Es existiert eine eindeutige  $(m, n)$ -Matrix  $\mathbf{A}$ , so dass

$$T(\mathbf{v}) = \mathbf{w} \quad \iff \quad \mathbf{y} = \mathbf{A} \cdot \mathbf{x}.$$

Diese Matrix  $\mathbf{A}$  heißt Matrixdarstellung (oder auch Basisdarstellung) von  $T$  und repräsentiert die Abbildung  $T$  bzgl. der gewählten Basen.

2. Die Komponenten  $a_{ij}$  von  $\mathbf{A}$  sind eindeutig durch die  $n$  vektorwertigen Gleichungen

$$T(\mathbf{v}_j) = \sum_{i=1}^m a_{ij} \cdot \mathbf{w}_i,$$

bestimmt.

3. Im Fall von  $V = \mathbb{R}^n$  und  $W = \mathbb{R}^m$  gibt es wieder ein Kochrezept: Wir formen die erweiterte Koeffizientenmatrix

$$\left( \begin{array}{ccc|ccc} | & & | & | & | & | \\ \mathbf{w}_1 & \dots & \mathbf{w}_m & T(\mathbf{v}_1) & \dots & T(\mathbf{v}_n) \\ | & & | & | & | & | \end{array} \right) \text{ bzw.}$$

solange mit dem Gauß-Jordan-Verfahren um, bis links vom Mittelstrich die Einheitsmatrix steht. Die Matrix  $\mathbf{A}$  kann dann rechts abgelesen werden.

**Beispiel 2.1a** Wir betrachten  $V = \mathbb{R}^3$  und  $W = \mathbb{R}^2$  sowie die lineare Abbildung

$$\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} \xrightarrow{T} \begin{pmatrix} w_1 \\ w_2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix} \cdot \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix},$$

die selbst durch eine Rechtecksmatrix vermittelt wird. Bzgl. der kanonischen Basen in  $V$  und  $W$  gilt

$$T(\mathbf{v}_1) = \begin{pmatrix} 1 \\ 4 \end{pmatrix} = 1 \cdot \mathbf{w}_1 + 4 \cdot \mathbf{w}_2,$$

$$T(\mathbf{v}_2) = \begin{pmatrix} 2 \\ 5 \end{pmatrix} = 2 \cdot \mathbf{w}_1 + 5 \cdot \mathbf{w}_2,$$

$$T(\mathbf{v}_3) = \begin{pmatrix} 3 \\ 6 \end{pmatrix} = 3 \cdot \mathbf{w}_1 + 6 \cdot \mathbf{w}_2,$$

und durch spaltenweises Zusammensetzen erhalten wir

$$\mathbf{A} = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{pmatrix}$$

als Matrixdarstellung von  $T$  bzgl. der gewählten Basen in  $V$  und  $W$ . Da diese aber beide kanonisch waren, ist  $\mathbf{A}$  gerade die Matrix, mit der wir  $T$  definiert hatten. *Bemerkung:* Der Unterschied zwischen einer linearen Abbildung und ihren Matrixdarstellungen tritt erst dann deutlich zu Tage, wenn wir auch nicht-kanonischen Basen betrachten.

**Beispiel 2.1b** Wir betrachten nochmal  $V$ ,  $W$  und  $T$  wie in letzten Beispiel 2.1a, aber wählen nun die Basen  $\tilde{\mathbf{v}}_1, \tilde{\mathbf{v}}_2, \tilde{\mathbf{v}}_3$  bzw.  $\tilde{\mathbf{w}}_1, \tilde{\mathbf{w}}_2$  aus den Beispielen 1.1 und 1.3. Wir bezeichnen aus Konsistenzgründen die entsprechende Matrixdarstellung mit  $\tilde{\mathbf{A}}$  (und nicht mit  $\mathbf{A}$ ) und wegen

$$T(\tilde{\mathbf{v}}_1) = \begin{pmatrix} 5 \\ 11 \end{pmatrix} = -11 \cdot \tilde{\mathbf{w}}_1 - 16 \cdot \tilde{\mathbf{w}}_2$$

$$T(\tilde{\mathbf{v}}_2) = \begin{pmatrix} 4 \\ 10 \end{pmatrix} = -10 \cdot \tilde{\mathbf{w}}_1 - 14 \cdot \tilde{\mathbf{w}}_2$$

$$T(\tilde{\mathbf{v}}_3) = \begin{pmatrix} 3 \\ 9 \end{pmatrix} = -9 \cdot \tilde{\mathbf{w}}_1 - 12 \cdot \tilde{\mathbf{w}}_2$$

erhalten wir

$$\tilde{\mathbf{A}} = \begin{pmatrix} -11 & -10 & -9 \\ -16 & -14 & -12 \end{pmatrix},$$

wobei die Koeffizienten auf der rechten Seite der Gleichung für  $T(\mathbf{v}_j)$  gerade die  $j$ -te Spalte von  $\mathbf{A}$  liefert. Alternativ hätten wir auch das Kochrezept benutzen können:

$$\begin{aligned} \left( \begin{array}{cc|ccc} 1 & -1 & 5 & 4 & 3 \\ -1 & 0 & 11 & 10 & 9 \end{array} \right) &\cong \left( \begin{array}{cc|ccc} -1 & 0 & 11 & 10 & 9 \\ 1 & -1 & 5 & 4 & 3 \end{array} \right) \cong \left( \begin{array}{cc|ccc} -1 & 0 & 11 & 10 & 9 \\ 0 & -1 & 16 & 14 & 12 \end{array} \right) \\ &\cong \left( \begin{array}{cc|ccc} 1 & 0 & -11 & -10 & -9 \\ 0 & 1 & -16 & -14 & -12 \end{array} \right) \end{aligned}$$

Bedeutung der Matrix  $\tilde{\mathbf{A}}$ : Der Vektor

$$\mathbf{v} = 1 \cdot \tilde{\mathbf{v}}_1 - 1 \cdot \tilde{\mathbf{v}}_3 = \begin{pmatrix} -1 \\ 0 \\ +1 \end{pmatrix},$$

wird zum Beispiel der unter der Abbildung (Pfeil oben im Diagramm) auf

$$T(\mathbf{v}) = \begin{pmatrix} 2 \\ 2 \end{pmatrix} = -2 \cdot \tilde{\mathbf{w}}_1 - 4 \cdot \tilde{\mathbf{w}}_2$$

abgebildet. Wir wissen nun, dass  $\mathbf{v}$  bzw.  $T(\mathbf{v})$  den Koordinatenvektoren

$$\tilde{\mathbf{x}} = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} \quad \text{bzw.} \quad \tilde{\mathbf{y}} = \begin{pmatrix} -2 \\ -4 \end{pmatrix}$$

entspricht. Diese Formeln repräsentieren den linken und den rechten Pfeil im Diagramm, wobei alle Größen in der unteren Ebene des Diagramms aus Konsistenzgründen mit der Schlange  $\sim$  zu versehen sind. Der untere Pfeil im Diagramm meint nun gerade

$$\tilde{\mathbf{A}} \cdot \tilde{\mathbf{x}} = \begin{pmatrix} -11 & -10 & -9 \\ -16 & -14 & -12 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = \begin{pmatrix} -2 \\ -4 \end{pmatrix} = \tilde{\mathbf{y}}.$$

**Beispiel 2.2a** Diesmal betrachten wir  $V = W = \Pi_3$  sowie die lineare Abbildung  $T : \Pi_3 \rightarrow \Pi_3$ , die durch die Differentiation vermittelt wird. Mit anderen Worten, in diesem Beispiel gilt

$$T(p) = \dot{p}$$

für jedes Polynom  $p$  und mit  $\dot{p}(t) = \frac{d}{dt}p(t)$ . Wählen wir sowohl im Bild- als auch im Urbildraum jeweils die monomiale Basis, so erhalten wir wegen

$$\begin{aligned} \frac{d}{dt}1 &= 0 = 0 \cdot 1 + 0 \cdot t + 0 \cdot t^2 + 0 \cdot t^3 \\ \frac{d}{dt}t &= 1 = 1 \cdot 1 + 0 \cdot t + 0 \cdot t^2 + 0 \cdot t^3 \\ \frac{d}{dt}t^2 &= 2t = 0 \cdot 1 + 2 \cdot t + 0 \cdot t^2 + 0 \cdot t^3 \\ \frac{d}{dt}t^3 &= 3t^2 = 0 \cdot 1 + 0 \cdot t + 3 \cdot t^2 + 0 \cdot t^3 \end{aligned}$$

die Matrixdarstellung

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Oder anders gesagt: Ist  $\mathbf{c} \in \mathbb{R}^4$  der Koeffizientenvektor eines Polynoms  $p$  bzgl. der monomialen Basis, so besitzt das Polynom  $\dot{p}$  den Koeffizientenvektor  $\mathbf{A} \cdot \mathbf{c}$ .

**Beispiel 2.2b** Wir betrachten nochmal die Abbildung  $T$  von eben, wollen aber nun die Matrixdarstellung bzgl. der nicht-monomialen Basis aus Beispiel 1.3 betrachten. Wegen

$$\begin{aligned} \frac{d}{dt}\tilde{p}_1(t) &= 0 = +0 \cdot 1 + 0 \cdot (t-1) + 0 \cdot (t-2)^2 + 0 \cdot (t-3)^3 \\ \frac{d}{dt}\tilde{p}_2(t) &= 1 = +1 \cdot 1 + 0 \cdot (t-1) + 0 \cdot (t-2)^2 + 0 \cdot (t-3)^3 \\ \frac{d}{dt}\tilde{p}_3(t) &= 2(t-2) = -2 \cdot 1 + 2 \cdot (t-1) + 0 \cdot (t-2)^2 + 0 \cdot (t-3)^3 \\ \frac{d}{dt}\tilde{p}_4(t) &= 3(t-3)^2 = +9 \cdot 1 - 6 \cdot (t-1) + 3 \cdot (t-2)^2 + 0 \cdot (t-3)^3 \end{aligned}$$

ergibt sich diesmal die Matrix

$$\tilde{\mathbf{A}} = \begin{pmatrix} 0 & 1 & -2 & +9 \\ 0 & 0 & 2 & -6 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Insbesondere beschreiben die soeben berechnete Matrix  $\tilde{\mathbf{A}}$  sowie die Matrix  $\mathbf{A}$  aus dem vorherigen Beispiel dieselbe lineare Operation (nämlich das Differenzieren von Polynomen), aber bzgl. unterschiedlicher Basiswahlen.

**Beispiel 2.3.b** Wir betrachten  $V = W = \mathbb{R}^n$  sowie die lineare Abbildung  $T$ , die durch eine gegebene  $(n, n)$ -Matrix  $\mathbf{A}$  vermittelt wird. Mit anderen Worten, es gilt

$$\mathbf{v} \xrightarrow{T} \mathbf{w} = T(\mathbf{v}) = \mathbf{A} \cdot \mathbf{v},$$

wobei wir der Einfachheit halber voraussetzen wollen, dass  $\mathbf{A}$  symmetrisch ist (so dass es  $n$  reelle Eigenwerte  $\lambda_1, \dots, \lambda_n$  gibt). Wir wählen nun sowohl im Urbild- als auch im Bildraum von  $T$  die Eigenbasis

$$\tilde{\mathbf{v}}_1, \dots, \tilde{\mathbf{v}}_n \quad \text{mit} \quad \mathbf{A} \cdot \tilde{\mathbf{v}}_j = \lambda_j \cdot \tilde{\mathbf{v}}_j$$

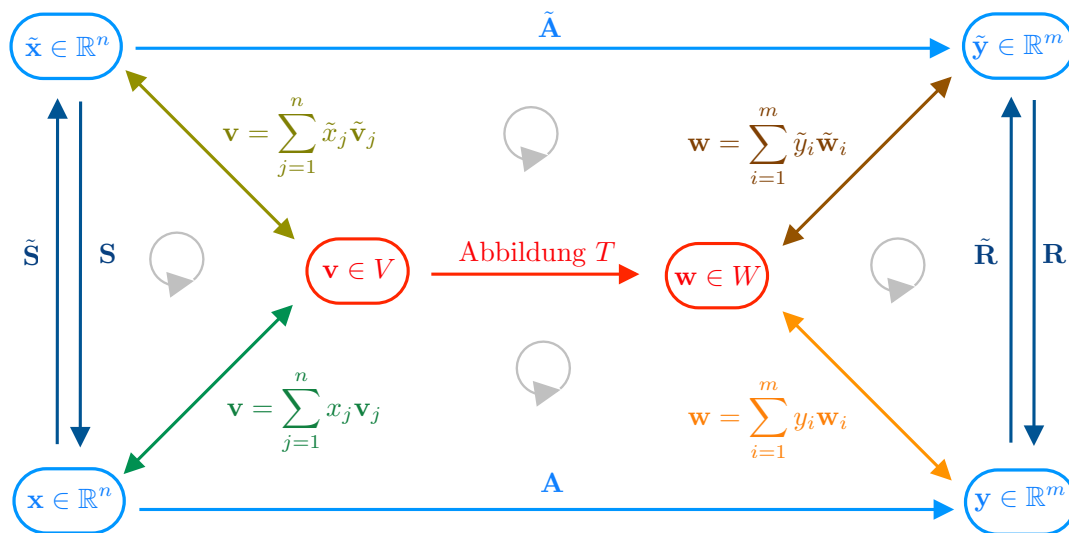
und bezeichnen die Matrixdarstellung von  $T$  bzgl. dieser Basis mit  $\tilde{\mathbf{A}}$  ( $\mathbf{A}$  ist hier auch wieder die Matrixdarstellung von  $T$  bzgl. der kanonischen Basis in Bild- und Urbildraum). Wegen

$$T(\tilde{\mathbf{v}}_1) = \mathbf{A} \cdot \tilde{\mathbf{v}}_1 = \lambda_1 \cdot \tilde{\mathbf{v}}_1 + 0 \cdot \tilde{\mathbf{v}}_2 + \dots + 0 \cdot \tilde{\mathbf{v}}_n$$

sehen wir, dass die erste Spalte von  $\tilde{\mathbf{A}}$  in der ersten Zeile gerade  $\lambda_1$ , in allen anderen Zeilen aber 0 enthält. Analoge Aussagen kann man leicht für alle anderen Spalten

von  $\tilde{\mathbf{A}}$  zeigen, d.h.  $\tilde{\mathbf{A}}$  ist gerade die Diagonalmatrix, die aus den Eigenwerten von  $\mathbf{A}$  besteht. Oder anders gesagt: Bei einer diagonalisierbaren Matrix bietet es sich an, den Basiswechsel auf die Eigenbasis zu vollziehen, eben weil man dann nur noch mit Diagonalmatrizen rechnen muss. Wir werden dies insbesondere bei den Linearen Differentialgleichungen in *Mathe II* so machen.

### Matrixdarstellung und Basiswechsel



Das kombinierte Diagramm aus Kapitel 5. Die Dreiecke beschreiben Basiswechsel, die Vierecke unterschiedliche Matrixdarstellung der Abbildung  $T$  im Zentrum.

**Prinzip/Theorem** Das obige Diagramm kommutiert, d.h. man kann die Pfeile beliebig kombinieren. Insbesondere gilt

$$\mathbf{R} \cdot \tilde{\mathbf{A}} = \mathbf{A} \cdot \mathbf{S}, \quad \tilde{\mathbf{R}} \cdot \mathbf{A} = \tilde{\mathbf{A}} \cdot \tilde{\mathbf{S}}$$

sowie

$$\tilde{\mathbf{S}} \cdot \mathbf{S} = \mathbf{S} \cdot \tilde{\mathbf{S}} = \mathbf{I}_n, \quad \tilde{\mathbf{R}} \cdot \mathbf{R} = \mathbf{R} \cdot \tilde{\mathbf{R}} = \mathbf{I}_m$$

und

$$\tilde{\mathbf{A}} = \tilde{\mathbf{R}} \cdot \mathbf{A} \cdot \mathbf{S}, \quad \mathbf{A} = \mathbf{R} \cdot \tilde{\mathbf{A}} \cdot \tilde{\mathbf{S}}$$

im Sinne der Matrizenmultiplikation.

**Beispiel 3.1** Wir betrachten noch einmal eine diagonalisierbare Matrix  $\mathbf{A}$  wie in Beispiel 2.3.b mit  $V = W = \mathbb{R}^n$ . Wählen wir im obigen Diagramm wieder  $\mathbf{v}_j = \mathbf{w}_j$  als  $j$ -ten kanonischen Einheitsvektor sowie  $\tilde{\mathbf{v}}_j = \tilde{\mathbf{w}}_j$  als den  $j$ -ten Eigenvektoren von  $\mathbf{A}$ , so gilt  $\mathbf{R} = \mathbf{S}$  und  $\tilde{\mathbf{R}} = \tilde{\mathbf{S}} = \mathbf{S}^{-1} = \mathbf{R}^{-1}$  sowie

$$\tilde{\mathbf{A}} = \tilde{\mathbf{S}} \cdot \mathbf{A} \cdot \mathbf{S},$$

wobei die Spalten von  $\mathbf{S}$  gerade die Eigenvektoren von  $\mathbf{A}$  sind. Oder anders gesagt: Die Spektraltheorie quadratischer Matrizen ist ein Spezialfall des allgemeinen Diagramms.

