

Oliver Deiser, Caroline Lasser

# Erste Hilfe in Linearer Algebra

*für Thalia und Larina*



---

# Inhalt

---

<b>Vorwort</b> .....	7
<b>0. Kapitel Mengentheoretisches Vorspiel</b> .....	9
0.1 Mengen .....	10
0.2 Endliche Mengen .....	12
0.3 Die Mengenkompensation .....	14
0.4 Algebraische Operationen mit Mengen .....	16
<b>1. Kapitel Relationen und Abbildungen</b> .....	19
1.1 Relationen .....	20
1.2 Äquivalenzrelationen .....	22
1.3 Ordnungen .....	24
1.4 Der Abbildungsbegriff .....	26
1.5 Konstruktion von Abbildungen .....	28
1.6 Notationen und Sprechweisen für Abbildungen .....	30
1.7 Umgang mit Funktionen .....	32
1.8 Operationen und Abgeschlossenheit .....	34
1.9 Abbildungseigenschaften .....	36
1.10 Mächtigkeitsvergleiche .....	38
1.11 Das Auswahlaxiom .....	40
1.12 Das Zornsche Lemma .....	42
<b>2. Kapitel Algebraische Strukturen</b> .....	45
2.1 Halbgruppen .....	46
2.2 Monoide .....	48
2.3 Gruppen .....	50
2.4 Rechenregeln in Gruppen .....	52
2.5 Kommutative Operationen .....	54
2.6 Untergruppen .....	56
2.7 Normalteiler und Faktorgruppen .....	58

2.8 Ringe .....	60
2.9 Körper .....	62
2.10 Angeordnete Körper .....	64
2.11 Polynomringe und Polynomfunktionen .....	66
2.12 Division und Nullstellen von Polynomen .....	68
<b>3. Kapitel Vektorräume .....</b>	<b>71</b>
3.1 Vektorräume .....	72
3.2 Unterräume .....	74
3.3 Produkte von Vektorräumen .....	76
3.4 Linearkombinationen und Erzeugendensysteme .....	78
3.5 Lineare Unabhängigkeit .....	80
3.6 Basen und Koordinatenvektoren .....	82
3.7 Austauschlemma und Austauschsatz .....	84
3.8 Die Dimension .....	86
3.9 Die Existenz von Basen .....	88
3.10 Summen von Vektorräumen .....	90
3.11 Quotientenräume .....	92
3.12 Affine Unterräume und Koordinaten .....	94
<b>4. Kapitel Lineare Abbildungen .....</b>	<b>97</b>
4.1 Gruppenhomomorphismen .....	98
4.2 Mono-, Epi-, Iso-, Endo- und Automorphismen .....	100
4.3 Kern und Bild .....	102
4.4 Der Homomorphiesatz .....	104
4.5 Lineare Abbildungen .....	106
4.6 Konstruktion linearer Abbildungen .....	108
4.7 Darstellung linearer Abbildungen .....	110
4.8 Fasern und lineare Gleichungssysteme .....	112
4.9 Isomorphie von Vektorräumen .....	114
4.10 Die Dimensionsformel .....	116
4.11 Lineare Abbildungen als Vektoren .....	118
4.12 Dualräume und duale Abbildungen .....	120
<b>5. Kapitel Matrizen .....</b>	<b>123</b>
5.1 Matrizen .....	124

5.2	Matrizen und lineare Abbildungen	126
5.3	Die Matrizenmultiplikation	128
5.4	Darstellende Matrizen für beliebige Basen	130
5.5	Invertierbare Matrizen	132
5.6	Die Elementarmatrizen	134
5.7	Die Permutationsmatrizen	136
5.8	Basiswechsel und Transformationsformel	138
5.9	Die Transposition	140
5.10	Der Rang	142
5.11	Die Zeilenstufenform	144
5.12	Eliminationsverfahren	146
<b>6.</b>	<b>Kapitel Euklidische und unitäre Vektorräume</b>	<b>149</b>
6.1	Das kanonische Skalarprodukt im $\mathbb{R}^n$	150
6.2	Das kanonische Skalarprodukt im $\mathbb{C}^n$	152
6.3	Allgemeine Skalarprodukte	154
6.4	Normierte Vektorräume	156
6.5	Normen im Endlich-Dimensionalen	158
6.6	Orthonormalbasen	160
6.7	Das Orthonormalisierungsverfahren	162
6.8	Orthogonale Komplemente und Projektionen	164
6.9	Orthogonale Homomorphismen und Matrizen	166
6.10	Der Rieszsche Darstellungssatz	168
6.11	Der adjungierte Endomorphismus	170
6.12	Sesquilinearformen	172
<b>7.</b>	<b>Kapitel Determinanten</b>	<b>175</b>
7.1	$2 \times 2$ -Determinanten	176
7.2	$n \times n$ -Determinanten	178
7.3	Das Vorzeichen einer Permutation	180
7.4	Die Leibniz-Formel	182
7.5	Multiplikation und Transposition	184
7.6	Der Entwicklungssatz von Laplace	186
7.7	Komplementärmatrizen und die Regel von Cramer	188
7.8	Die speziellen linearen Gruppen	190
7.9	Volumina von Parallelotopen	192

7.10 Das Kreuzprodukt .....	194
7.11 Positive Definitheit .....	196
7.12 Die Determinante eines Endomorphismus .....	198
<b>8. Kapitel Eigenwerte .....</b>	<b>201</b>
8.1 Eigenwerte und Eigenvektoren .....	202
8.2 Die Diagonalisierbarkeit .....	204
8.3 Das charakteristische Polynom .....	206
8.4 Das Diagonalisierbarkeitskriterium .....	208
8.5 Die Trigonalisierung .....	210
8.6 Der Spektralsatz .....	212
8.7 Hauptachsentransformation und Trägheitssatz .....	214
8.8 Die Singulärwertzerlegung .....	216
8.9 Lineare Abbildungen und Ellipsen .....	218
8.10 Minimalpolynome und der Satz von Cayley-Hamilton .....	220
8.11 Haupträume und Hauptraumzerlegung .....	222
8.12 Die Jordan-Normalform .....	224
<b>Überblick und Zusammenfassung .....</b>	<b>227</b>
1. Algebraische Grundstrukturen .....	228
2. Die Kongruenz modulo $m$ .....	229
3. Matrizen .....	230
4. Matrizen und lineare Abbildungen .....	232
5. Umformungen mit Elementarmatrizen .....	232
6. Matrixengruppen .....	234
7. Matrixzerlegungen .....	235
8. Die Sesquilinearformen $\langle \cdot, A \cdot \rangle$ und positive Definitheit .....	236
9. Quadriken in Normalform für $n = 2$ .....	237
10. Normalformen .....	238
11. Blockstrukturen .....	241
12. Berechnung und Bestimmung .....	242
<b>Ausblicke zu Eigenwerten .....</b>	<b>243</b>
1. Eigenwerte ohne Determinanten .....	244
2. Eigenwerte ohne Fundamentalsatz .....	245
3. Gershgorin-Kreise und die Lage der Eigenwerte .....	246
4. Matrixnormen .....	248
5. Matrixexponentiale .....	250
6. Lineare Systeme von Differentialgleichungen .....	252

<b>Anhänge</b> .....	255
1. Junktoren .....	256
2. Quantoren .....	258
3. Zum Funktionsbegriff .....	259
4. Zahlen .....	260
5. Geometrische Grundlagen .....	264
6. Die Axiome der Mengenlehre .....	266
<b>Literatur</b> .....	269
<b>Notationen</b> .....	271
<b>Index</b> .....	275
<b>Matrizensterne</b> .....	285





---

# Vorwort

---

Die Lineare Algebra gehört zum weltweit etablierten Kanon des mathematischen Grundwissens. In ihrem Zentrum stehen lineare Abbildungen zwischen Vektorräumen, die mit Hilfe von Matrizen analysiert werden. Traditionell findet innerhalb der Linearen Algebra auch eine erste Begegnung mit den algebraischen Grundstrukturen *Gruppen*, *Ringe*, *Körper* statt. In vielen Anfängervorlesungen werden zudem die für alle mathematischen Gebiete unentbehrlichen Begriffe aus dem Umfeld von Mengen, Relationen und Abbildungen behandelt. Das Buch folgt dieser Vorgehensweise. Nach einer kurzen Diskussion von Grundlagen (Kapitel 0 und 1) und algebraischen Strukturen (Kapitel 2) wenden wir uns sechs Kernthemen der Linearen Algebra zu (Kapitel 3–8). Diese sind:

Vektorräume, lineare Abbildungen, Matrizen,  
Skalarprodukte, Determinanten, Eigenwerte.

Der Umfang des Textes entspricht, je nach Vorwissen, Tempo und Lernzielen, einer bis zwei vierstündigen Vorlesungen. Es wurde ein systematischer und strukturierter Aufbau angestrebt, der sich sowohl zum Lernen und Wiederholen als auch für die Lehre eignet. Der Unterschied zu einem klassischen vorlesungsbegleitenden Lehrbuch lässt sich grob gesprochen durch das Fehlen einiger (aber nicht aller) Beweise und einer ausführlicheren Kommentierung zugunsten einer kompakten und übersichtlichen Darstellung unter Einschluss vieler Beispiele und Gegenbeispiele beschreiben. Dass Beweise fehlen, liegt an der Konzeption des Buches und nicht etwa an der Meinung der Autoren, dass man auf diese verzichten könne. Kurze Argumente sind ausgeführt, wenn sie den Einsatz mathematischer Eigenschaften aufzeigen und helfen, den Umgang mit mathematischen Objekten zu erlernen. Zahlreiche Abbildungen möchten die behandelten Begriffe, Ergebnisse und Methoden für die Anschauung greifbar und für die Erinnerung zugänglich machen. Wir möchten anregen, sich der Mathematik auch durch die eigenständige Anfertigung von Skizzen – die ganz einfach ausfallen können – zu nähern. Großer Wert wird auf exakte Definitionen, auf die oft unentbehrlichen Voraussetzungen der Sätze und auf die sorgfältige Verwendung der mathematischen Sprache gelegt.

Das Buch kann verwendet werden

- (1) als Begleittext für Hörer der Linearen Algebra I und (in Teilen) II; ein ausführlicheres und beweisvollständiges Lehrbuch kann und will es dabei nicht ersetzen,
- (2) zur Wiederholung und Prüfungsvorbereitung,
- (3) zur Selbstkontrolle (Kann ich die wichtigsten Begriffe und Ergebnisse genau wiedergeben? Kann ich kurze Argumentationen eigenständig und sicher führen? Kann ich Beispiele und Gegenbeispiele angeben? Kann ich abstrakte Begriffe veranschaulichen?),
- (4) zur Wissensauffrischung (insb. für Gymnasiallehrer, Physiker, Informatiker),
- (5) als Anregung für Dozenten im Sinne eines ausgearbeiteten Ansatzes, Lernprozesse von Studienanfängern zu unterstützen.

München, im März 2021

Oliver Deiser und Caroline Lasser

# 0. Kapitel

---

## Mengentheoretisches Vorspiel

---

## 0.1 Mengen

### Intuitiver Mengenbegriff

#### Mengen und ihre Elemente

Eine *Menge* ist eine Zusammenfassung von Objekten zu einem Objekt. Die Objekte, die eine Menge bilden, heißen ihre *Elemente*. Eine Menge ist durch ihre Elemente bestimmt.

#### Elementbeziehung

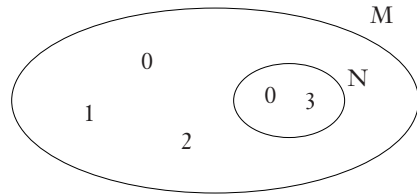
Ist ein Objekt  $x$  ein Element einer Menge  $M$ , so schreiben wir

$$x \in M, \quad (\text{Epsilon- oder Element-Beziehung})$$

gelesen: „ $x$  epsilon  $M$ “, „ $x$  Element  $M$ “, „ $x$  ist in  $M$  als Element enthalten“.

Ist  $x$  kein Element von  $M$ , so schreiben wir  $x \notin M$ .

„Menge“ ist ein nicht definierter Grundbegriff der Mathematik. Welche Mengen existieren, wird durch Axiome geregelt, die man nicht unbedingt kennen muss. Intuitiv ist eine (mathematische) Menge eine Zusammenfassung von (mathematischen) Objekten. Man kann sich diese Zusammenfassung als „Sack“ oder „umzäuntes Gebiet“ vorstellen, in dem sich die Objekte befinden.



Die Elemente einer Menge können beliebige Objekte und damit selbst Mengen sein. Die Menge  $M$  hat die Elemente 0, 1, 2 und  $N$ . Dabei ist  $N$  die aus 0 und 3 gebildete Menge. Es gilt  $3 \notin M$ .

### Beispiele

- (1) Ist  $M$  die aus den Zahlen 1, 2 und 3 gebildete Menge, so gilt  $1 \in M$ ,  $2 \in M$  und  $3 \in M$ . Für alle anderen  $x$  gilt  $x \notin M$ .
- (2)  $\mathbb{N}$  = „die Menge aller natürlichen Zahlen (einschließlich der Null)“,  
 $\mathbb{Z}$  = „die Menge aller ganzen Zahlen“,  
 $\mathbb{Q}$  = „die Menge aller rationalen Zahlen“,  
 $\mathbb{R}$  = „die Menge aller reellen Zahlen“.

Obwohl das Wort „Menge“ in der Umgangssprache eher „Vieles“ suggeriert, ist es in der Mathematik nützlich, auch den „leeren Sack“ als Menge zuzulassen:

### Die leere Menge $\emptyset$

Wir bezeichnen die Menge, die kein Element enthält, mit  $\emptyset$ .

Es gilt also  $x \notin \emptyset$  für alle  $x$ .

## Die Teilmengenbeziehung (Inklusion)

Wir definieren:

### Definition (Teilmenge, echte Teilmenge, Obermenge, echte Obermenge)

Eine Menge  $N$  heißt *Teilmenge* einer Menge  $M$ , falls jedes Element von  $N$  ein Element von  $M$  ist. In Zeichen schreiben wir  $N \subseteq M$ . Gilt  $N \subseteq M$  und  $N \neq M$ , so heißt  $N$  eine *echte Teilmenge* von  $M$ . In Zeichen schreiben wir  $N \subset M$ .

Gilt  $N \subseteq M$  bzw.  $N \subset M$ , so nennen wir  $M$  auch eine *Obermenge* von  $N$  bzw. eine *echte Obermenge* von  $N$ . Wir schreiben hierfür auch  $M \supseteq N$  bzw.  $M \supset N$ .

### Warnung

Viele Mathematiker (vor allem in der Analysis und Wahrscheinlichkeitstheorie) schreiben der Kürze halber  $M \subset N$  für  $M \subseteq N$  sowie  $M \subsetneq N$  für  $M \subset N$ . Die seit Felix Hausdorff betonte Analogie der Notation zu  $\leq$  und  $<$  geht dadurch verloren.

### Beispiele

- (1) Die Menge der geraden Zahlen ist eine echte Teilmenge der natürlichen Zahlen. Die Menge der von 2 verschiedenen Primzahlen ist eine echte Teilmenge der Menge der ungeraden Zahlen.
- (2) Es gilt  $\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q} \subset \mathbb{R}$ .
- (3) Gilt  $N \subset M$ , so gilt auch  $N \subseteq M$ .
- (4) Gilt  $M_1 \subseteq M_2$  und  $M_2 \subseteq M_3$ , so gilt auch  $M_1 \subseteq M_3$  (Transitivität der Inklusion). Das Gleiche gilt für die echte Inklusion  $\subset$ .
- (5) Für jede Menge  $M$  gilt  $\emptyset \subseteq M$  (da jedes Element von  $\emptyset$  auch Element von  $M$  ist) und  $M \subseteq M$  (da jedes Element von  $M$  ein Element von  $M$  ist).

## Das Extensionalitätsprinzip

Dass eine Menge durch ihre Elemente bestimmt ist, wird oft wie folgt zum Ausdruck gebracht:

### Extensionalitätsprinzip

Zwei Mengen sind genau dann gleich, wenn sie dieselben Elemente besitzen.

Es gibt also keine „gelbe“ von den Zahlen 1, 2 und 3 gebildete Menge, die von einer „roten“ von 1, 2 und 3 gebildeten Menge zu unterscheiden wäre. Mit Hilfe der Inklusion können wir das Prinzip auch so formulieren:

### Extensionalitätsprinzip, Umformulierung

Für alle Mengen  $M$  und  $N$  gilt  $M = N$  genau dann, wenn  $M \subseteq N$  und  $N \subseteq M$ .

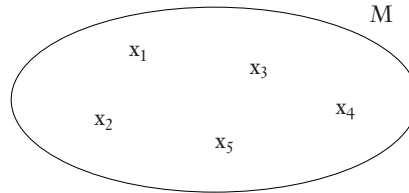
## 0.2 Endliche Mengen

### Definition (*endliche Komprehension, Auflistung der Elemente*)

Sind  $x_1, \dots, x_n$  Objekte, so schreiben wir

$$\{x_1, \dots, x_n\}$$

für die durch  $x_1, \dots, x_n$  gebildete Menge. Weiter schreiben wir auch  $\{\}$  für die leere Menge  $\emptyset$ .



Die endliche Menge

$$M = \{x_1, x_2, x_3, x_4, x_5\}.$$

Die Elemente  $x_1, \dots, x_5$  der Menge müssen nicht paarweise verschieden sein.

Die geschweiften Klammern  $\{$  und  $\}$  werden in der Mathematik fast ausschließlich zur Notation von Mengen verwendet. Auch außerhalb der Mathematik werden sie oft als „Mengenklammern“ bezeichnet.

Die Menge  $M = \{x_1, \dots, x_n\}$  hat genau die Elemente  $x_1, \dots, x_n$ , d. h., es gilt:

$$(a) \quad x_1 \in M, \dots, x_n \in M.$$

$$(b) \quad \text{Ist } x \in M, \text{ so ist } x = x_1 \text{ oder } \dots \text{ oder } x = x_n.$$

Ist umgekehrt  $N$  eine Menge mit den Eigenschaften (a) und (b), so gilt

$$N = \{x_1, \dots, x_n\}$$

nach dem Extensionalitätsprinzip.

Anhand der Mengenbildung durch Auflistung können viele wichtige allgemeine Eigenschaften des Mengenbegriffs deutlich gemacht werden. Wir diskutieren vier davon.

### 1. Die Reihenfolge der Elemente spielt keine Rolle.

#### Beispiele

$$(1) \quad \{1, 2\} = \{2, 1\},$$

$$(2) \quad \{-1, 0, 1\} = \{1, 0, -1\} = \{1, -1, 0\} \text{ usw.}$$

### 2. Wiederholungen der Elemente spielen keine Rolle.

#### Beispiele

$$(1) \quad \{1, 3, 1, 2, 2, 2, 1\} = \{1, 2, 3\},$$

$$(2) \quad \{1\} = \{1, 1\} = \{1, 1, 1\} = \{1, 1, 1, 1\} \text{ usw.}$$

Die Anzahl der Elemente einer Menge  $\{x_1, \dots, x_n\}$  kann also kleiner als  $n$  sein.

### 3. Die Elemente einer Menge können selbst Mengen sein.

#### Beispiele

- (1) Die Menge  $M = \{1, 2, \{1, 2, 3\}\}$  hat genau drei verschiedene Elemente, nämlich die Zahlen 1 und 2 sowie die Zahlenmenge  $\{1, 2, 3\}$ . Es gilt  $\{1, 2, 3\} \in M$ . Auf der linken Seite der  $\in$ -Beziehung können also auch Mengen vorkommen. Sind alle Elemente von  $M$  Mengen, so heißt  $M$  auch ein *Mengensystem*.
- (2) Es gilt  $\{\} \neq \{\{\}\} = \{\emptyset\}$ , denn die leere Menge  $\{\}$  hat kein Element, während die Menge  $\{\emptyset\}$  die leere Menge als Element besitzt.

### 4. Elemente und Teilmengen sind zu unterscheiden.

#### Beispiele

- (1) Sei  $M = \{1, 2, 3, \{1, 2, 3\}\}$ . Dann gilt
 
$$\{1, 2\} \subseteq M, \{1, 2\} \notin M, \{1, 2, 3\} \subseteq M, \{1, 2, 3\} \in M.$$
- (2)  $\emptyset \subseteq \{1, 2, 3\}, \emptyset \notin \{1, 2, 3\}, \{1\} \subseteq \{1, 2, 3\}, 1 \in \{1, 2, 3\}.$

### Geordnete Tupel

Dass Reihenfolge und Wiederholungen keine Rolle spielen, ist manchmal unerwünscht, man denke etwa an die  $x$ - $y$ -Koordinaten eines Punktes der Ebene oder an die  $x$ - $y$ - $z$ -Koordinaten eines Punktes des dreidimensionalen Raums. Will man Reihenfolge und Wiederholungen respektieren, so verwendet man runde Klammern:

$$(x, y), (x, y, z), \dots, (x_1, \dots, x_n), \dots \quad (\text{geordnetes Paar, Tripel, \dots, } n\text{-Tupel, \dots})$$

Für alle  $x_1, \dots, x_n, y_1, \dots, y_m$  gilt dann

$$(+ \quad (x_1, \dots, x_n) = (y_1, \dots, y_m) \quad \text{genau dann, wenn} \quad n = m \text{ und } x_i = y_i \text{ für alle } 1 \leq i \leq n.$$

Um keine weiteren undefinierten Grundbegriffe zulassen zu müssen, kann man  $n$ -Tupel als Mengen einführen, sodass (+) gilt. Man definiert hierzu

$$(x, y) = \{\{x\}, \{x, y\}\}, \quad (\text{Kuratowski-Paar})$$

$$(x, y, z) = ((x, y), z), \quad (x_1, x_2, x_3, x_4) = ((x_1, x_2, x_3), x_4) \text{ usw.}$$

Diese Definitionen spielen im mathematischen Alltag zwar keine Rolle, aber sie illustrieren eine Stärke des Mengenbegriffs, die die moderne Mathematik geprägt hat und die im mathematischen Alltag überall spürbar ist:

*Mengen eignen sich zur Definition aller anderen mathematischen Objekte.*

Das bringt nicht nur Philosophen zum Schwärmen. Das Kuratowski-Paar ist ein schönes Beispiel, um mit diesem Gedanken vertraut zu werden. Es will nicht erklären, was  $(x, y)$  „ist“ (ontologisch), es will  $(x, y)$  lediglich so definieren („als Menge interpretieren“), dass alle erwünschten Eigenschaften erfüllt sind.

### 0.3 Die Mengenkomprension

**Definition** (*Mengenkomprension über Eigenschaften*)

Sei  $\mathcal{E}(x)$  eine Eigenschaft und  $M$  eine Menge. Wir schreiben

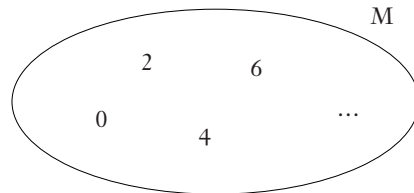
$M = \{ x \mid \mathcal{E}(x) \}$ , (gelesen: „ $M$  ist die Menge aller  $x$  mit der Eigenschaft  $\mathcal{E}(x)$ .“)

falls für alle  $x$  gilt:

$x \in M$  genau dann, wenn  $\mathcal{E}(x)$ .

Gilt  $M = \{ x \mid \mathcal{E}(x) \}$ , so ist nach Definition  $y \in \{ x \mid \mathcal{E}(x) \}$  gleichwertig mit  $\mathcal{E}(y)$ . Insbesondere gilt  $M = \{ x \mid x \in M \}$ . Oft wird aber

$M = \{ x \mid \mathcal{E}(x) \}$



$M = \{ n \mid n \text{ ist eine gerade natürliche Zahl} \}$

zur Definition einer Menge  $M$  verwendet. Intuitiv sammelt man dabei im mathematischen Universum alle  $x$  mit  $\mathcal{E}(x)$  auf und bildet aus diesen  $x$  die Menge  $M$ . Diese freizügigen Reisen durch das mathematische Weltall sind alles andere als unproblematisch, da sie widersprüchliche Zusammenfassungen wie

$R = \{ x \mid x \notin x \} =$  „die Menge aller  $x$ , die sich selbst nicht als Element enthalten“  
(*Russell-Komprension*)

zulassen, die durch die Frage, ob  $R \in R$  oder  $R \notin R$  gilt, die Mathematik ins Wanken bringen: Gilt nämlich  $R \notin R$ , so ist  $R$  ein  $x$  mit  $x \notin x$ , sodass doch  $R \in R$ . Gilt aber  $R \in R$ , so gilt Eigenschaft  $\mathcal{E}(x)$  für  $x = R$ , also doch wieder  $R \notin R$  (da  $\mathcal{E}(x) = „x \notin x“$ ). Bertrand Russell hat diese Paradoxie wie folgt anschaulich gemacht:

**Der Dorfbarbier**

Ein Dorfbarbier behauptet, dass er genau jenen Dorfbewohnern die Haare schneidet, die sich die Haare nicht selbst schneiden. Stimmt diese Aussage, so muss er sich selbst die Haare genau dann schneiden, wenn er sie sich selbst nicht schneidet.

Es gibt also keinen solchen Dorfbarbier, und ebenso gibt es die Menge  $R = \{ x \mid x \notin x \}$  nicht. Man schließt heute derartige Mengenbildungen aus, indem man sich axiomatisch von der leeren Menge  $\emptyset$  zu immer größeren Mengen hochhangelt und hofft, dass dadurch keine Widersprüche entstehen. Viel mehr kann man nicht tun, denn Gödel hat in seinen Unvollständigkeitssätzen bewiesen, dass die Mengenlehre – und auch jedes vergleichbare System – ihre eigene Widerspruchsfreiheit nicht beweisen kann.

Die erfreuliche Nachricht für den Anfänger lautet: Die Hintergrundaxiomatik der Mengenlehre, in der zumindest bis zur Veröffentlichung dieses Textes kein Widerspruch gefunden wurde, ist so stark, dass die Definition von  $M = \{ x \mid \mathcal{E}(x) \}$  in der „mathematischen Praxis“ immer erlaubt ist. Der Stachel der Russell-Paradoxie und der fast mythische



Status der Gödelschen Sätze bleiben bestehen. Das gehört zur Mathematik dazu.

### Pünktchen-Notationen

In Mengenkompensationen wird die definierende Eigenschaft oft gar nicht genannt, sondern muss durch ein angedeutetes Bildungsgesetz „erraten“ werden. Das ist suggestiv, wenn auch nicht ungefährlich.

#### Beispiele

- (1)  $\{0, 2, 4, \dots, 96, 98, 100\} = \{n \in \mathbb{N} \mid n \text{ ist gerade und } n \leq 100\},$
- (2)  $\{0, 2, 4, \dots\} = \{n \in \mathbb{N} \mid n \text{ ist gerade}\},$
- (3)  $\{0, 3, -3, 6, -6, 9, -9, \dots\} = \{a \in \mathbb{Z} \mid \text{es gibt ein } b \in \mathbb{Z} \text{ mit } a = 3b\},$
- (4)  $\{1, 2, 4, 8, 16, 32, \dots\} = \{n \in \mathbb{N} \mid \text{es gibt ein } k \in \mathbb{N} \text{ mit } n = 2^k\}.$

### Wichtige Kompensationen

Menge	Definition	Name
$\emptyset, \{\}$	$\{x \mid x \neq x\}$	leere Menge
$\{a\}$	$\{x \mid x = a\}$	Einermenge, Singleton
$\{a, b\}$	$\{x \mid x = a \text{ oder } x = b\}$	Paarmenge
$\{a_1, \dots, a_n\}$	$\{x \mid x = a_1 \text{ oder } \dots \text{ oder } x = a_n\}$	Auflistung
$(a, b)$	$\{\{a\}, \{a, b\}\}$	geordnetes Paar
$(a, b, c)$	$((a, b), c)$	Tripel
$A \times B$ $A \times B \times C$ $A^2, A^3, A^4, \dots$	$\{(a, b) \mid a \in A \text{ und } b \in B\}$ $\{(a, b, c) \mid a \in A \text{ und } b \in B \text{ und } c \in C\}$ $A \times A, A^2 \times A, A^3 \times A, \dots$	Kreuzprodukt
$\{x \in A \mid \mathcal{C}(x)\}$	$\{x \mid x \in A \text{ und } \mathcal{C}(x)\}$	Aussonderung
$\mathcal{P}(M)$	$\{A \mid A \subseteq M\}$	Potenzmenge

#### Beispiele

- (1)  $\{0, 1\} \times \{1, 2\} = \{(0, 1), (0, 2), (1, 1), (1, 2)\},$
- (2)  $\mathbb{R}^2 = \{(x, y) \mid x, y \in \mathbb{R}\}, \quad \mathbb{R}^n = \{(x_1, \dots, x_n) \mid x_k \in \mathbb{R} \text{ für alle } 1 \leq k \leq n\},$
- (3)  $\mathcal{P}(\emptyset) = \{\emptyset\}, \quad \mathcal{P}(\{0\}) = \{\emptyset, \{0\}\}, \quad \mathcal{P}(\{0, 1\}) = \{\emptyset, \{0\}, \{1\}, \{0, 1\}\}.$

## 0.4 Algebraische Operationen mit Mengen

### Definition (Durchschnitt, Vereinigung, Komplement)

#### Boolesche Mengenoperationen

Sei  $M$  eine Menge, und seien  $A, B \subseteq M$ . Dann definieren wir

$$A \cap B = \{x \mid x \in A \text{ und } x \in B\}, \quad (\text{Durchschnitt})$$

$$A \cup B = \{x \mid x \in A \text{ oder } x \in B\}, \quad (\text{Vereinigung})$$

$$A - B = A \setminus B = \{x \mid x \in A \text{ und } x \notin B\}, \quad (\text{Differenz})$$

$$A^c = M - A, \quad (\text{Komplement in } M)$$

$$A \Delta B = (A - B) \cup (B - A). \quad (\text{symmetrische Differenz})$$

Wir nennen  $A$  und  $B$  *disjunkt*, falls  $A \cap B = \emptyset$ , und *komplementär*, falls  $B = A^c$ .

#### Boolesche Operationen für Mengensysteme

Sei  $M$  eine Menge, und seien  $\mathcal{A}, \mathcal{B} \subseteq \mathcal{P}(M)$ , d. h., jedes Element von  $\mathcal{A}$  und von  $\mathcal{B}$  ist eine Teilmenge von  $M$ . Dann definieren wir

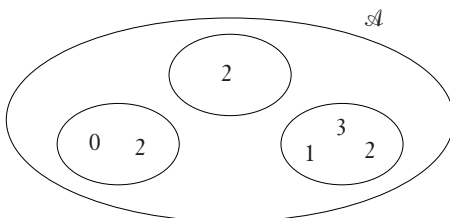
$$\bigcap \mathcal{A} = \bigcap_{A \in \mathcal{A}} A = \{x \in M \mid \text{für alle } A \in \mathcal{A} \text{ gilt } x \in A\}, \quad (\text{Durchschnitt})$$

$$\bigcup \mathcal{A} = \bigcup_{A \in \mathcal{A}} A = \{x \in M \mid \text{es gibt ein } A \in \mathcal{A} \text{ mit } x \in A\}, \quad (\text{Vereinigung})$$

$$\mathcal{A}^c = \{A^c \mid A \in \mathcal{A}\}. \quad (\text{Komplementsystem})$$

Weiter nennen wir das Mengensystem  $\mathcal{A}$

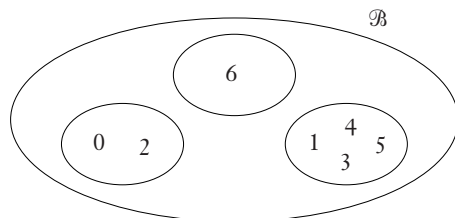
- (a) *paarweise disjunkt*, falls  $A \cap B = \emptyset$  für alle  $A, B \in \mathcal{A}$  mit  $A \neq B$ ,
- (b) eine *Überdeckung (von  $M$ )*, falls  $\bigcup \mathcal{A} = M$ ,
- (c) eine *Zerlegung, Partition oder Klasseneinteilung (von  $M$ )*, falls  $\mathcal{A}$  eine paarweise disjunkte Überdeckung mit  $\emptyset \notin \mathcal{A}$  ist.



$$\mathcal{A} = \{\{0, 2\}, \{2\}, \{1, 2, 3\}\}$$

$$\bigcap \mathcal{A} = \{2\}$$

$$\bigcup \mathcal{A} = \{0, 1, 2, 3\}$$



$$\mathcal{B} = \{\{0, 2\}, \{6\}, \{1, 3, 4, 5\}\}$$

Das System  $\mathcal{B}$  ist eine Zerlegung von

$$M = \{0, 1, 2, 3, 4, 5, 6\}.$$

Rechenregeln für die booleschen Operationen		
$A \cap (B \cap C) = (A \cap B) \cap C$	$A \cup (B \cup C) = (A \cup B) \cup C$	Assoziativität
$A \cap B = B \cap A$	$A \cup B = B \cup A$	Kommutativität
$A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$	$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$	Distributivität
$A \cap A^c = \emptyset$	$A \cup A^c = M$	Komplementierung
$(A \cap B)^c = A^c \cup B^c$	$(A \cup B)^c = A^c \cap B^c$	De-Morgan-Regeln
$A - B = A \cap B^c$ $(A \cap B) - C = A \cap (B - C)$	$(A - B) - C = A - (B \cup C)$ $(A - B) \cup B = A \cup B$	Differenzenregeln
$A \Delta (B \Delta C) = (A \Delta B) \Delta C$	$A \Delta B = (A \cup B) - (A \cap B)$	Regeln für $\Delta$

Für  $\cap, \cup$  und  $\Delta$  können wir Klammern weglassen und  $A \cap B \cap C$  statt  $(A \cap B) \cap C$  oder  $A \cap (B \cap C)$  usw. schreiben. Dagegen ist  $A - (B - C)$  von  $(A - B) - C$  zu unterscheiden.

**Beispiele**

Für  $M = \{ 1, 2, 3 \}$  gilt

$\{ 1 \} \cap \{ 2, 3 \} = \emptyset, \{ 1, 2 \}^c = \{ 3 \}, \{ 1, 2 \} - \{ 2, 3 \} = \{ 1 \},$

$\{ 1, 2 \} \Delta \{ 2, 3 \} = \{ 1, 3 \}, \{ 1, 2 \} \Delta \{ 2, 3 \} \Delta \{ 1 \} = \{ 3 \}.$

Für den Durchschnitt und die Vereinigung von Mengensystemen gilt:

$\cap \mathcal{A} =$  „was alle haben“,  $\cup \mathcal{A} =$  „was mindestens einer hat“.

**Beispiele**

- (1)  $\cap \{ A \} = A, \cap \{ A, B \} = A \cap B, \cup \{ A, B, C \} = A \cup B \cup C,$   
 $\cap \{ A, \emptyset, B, C \} = \emptyset, \cup \{ A, B, M, C \} = M,$   
 $\cap \emptyset = M, \cap \{ \emptyset \} = \emptyset, \cup \emptyset = \cup \{ \emptyset \} = \emptyset.$
- (2)  $\{ \{ 1, 2 \}, \{ 2, 3, 4 \}, \{ 1, 4 \}, \emptyset \}$  ist eine Überdeckung von  $\{ 1, 2, 3, 4 \},$
- (3)  $\{ \{ (n, m) \mid n \in \mathbb{N} \} \mid m \in \mathbb{N} \}$  ist die Zerlegung von  $\mathbb{N}^2$  in „Zeilen“  $\mathbb{N} \times \{ m \}.$

Für Systeme gilt:

$(\cap \mathcal{A})^c = \cup \mathcal{A}^c, (\cup \mathcal{A})^c = \cap \mathcal{A}^c. \qquad \qquad \qquad (allgemeine\ De-Morgan-Regeln)$



# 1. Kapitel

---

## Relationen und Abbildungen

---

## 1.1 Relationen

### Definition (Relation)

#### Relationen

Eine Menge  $R$  heißt eine (zweistellige) *Relation*, falls jedes Element von  $R$  ein geordnetes Paar ist. Gilt  $R \subseteq A \times A$  für eine Menge  $A$ , so heißt  $R$  eine Relation auf  $A$ . Anstelle von  $(a, b) \in R$  schreiben wir auch  $a R b$ .

#### Definitions- und Wertebereich

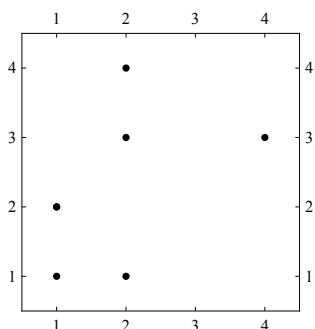
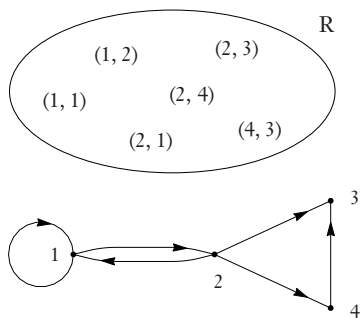
Für eine Relation  $R$  setzen wir (mit  $\text{dom}$  und  $\text{rng}$  für engl. domain bzw. range):

$$\text{Def}(R) = \text{dom}(R) = \{a \mid \text{es gibt ein } b \text{ mit } a R b\}, \quad (\text{Definitionsbereich})$$

$$\text{Bild}(R) = \text{rng}(R) = \{b \mid \text{es gibt ein } a \text{ mit } a R b\}, \quad (\text{Bild oder Wertebereich})$$

#### Eigenschaften einer Relation $R$ bzgl. einer Menge $A$

$R$ heißt ... auf $A$	falls für alle $a, b, c \in A$ gilt:
<i>reflexiv</i>	$a R a$
<i>irreflexiv</i>	nicht( $a R a$ )
<i>symmetrisch</i>	$a R b$ impliziert $b R a$
<i>antisymmetrisch</i>	$(a R b \text{ und } b R a)$ impliziert $a = b$
<i>transitiv</i>	$(a R b \text{ und } b R c)$ impliziert $a R c$



Drei Darstellungen einer Relation  $R$  auf  $\{1, 2, 3, 4\}$ . Es gilt  
 $1 R 1, 1 R 2, 2 R 1,$   
 $2 R 4, 2 R 3, 4 R 3,$   
 $\text{Def}(R) = \{1, 2, 4\},$   
 $\text{Bild}(R) = \{1, 2, 3, 4\}.$

In einer Relation  $R$  sind alle Paare  $(a, b)$ , die in einer „bestimmten Beziehung“ stehen, versammelt. Statt  $(a, b) \in R$  wird meistens  $a R b$  geschrieben, wie man es etwa von  $a \leq b$  oder  $a = b$  gewohnt ist. Wir vereinbaren zudem:

$a R b R c$  bedeutet  $a R b$  und  $b R c$ .

Man vergleiche hierzu wieder  $a \leq b \leq c$  und  $a = b = c$ .

## Beispiele

- (1) Die Kleinerleich-Relation auf  $\mathbb{N}$  kann definiert werden durch

$$\leq = \{ (n, m) \in \mathbb{N}^2 \mid \text{es gibt ein } k \in \mathbb{N} \text{ mit } n + k = m \}, \quad (\text{Kleinerleich auf } \mathbb{N})$$

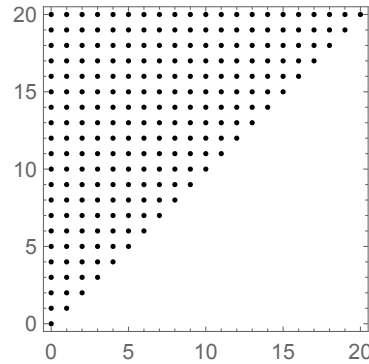
oder gleichwertig – und besser lesbar – durch die Setzung

$$n \leq m, \text{ falls es gibt ein } k \in \mathbb{N} \text{ mit } n + k = m$$

für alle  $n, m \in \mathbb{N}$ . Es gilt

$$\text{Def}(\leq) = \text{Bild}(\leq) = \mathbb{N}.$$

Die  $\leq$ -Relation ist reflexiv, antisymmetrisch und transitiv.



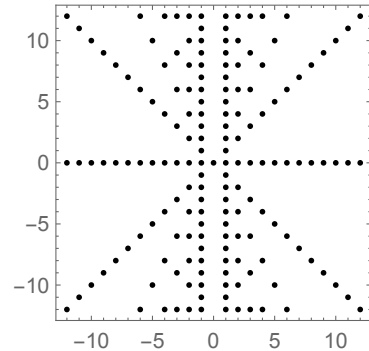
- (2) Für alle  $d, a \in \mathbb{Z}$  setzen wir

$$d \mid a, \text{ falls es gibt ein } k \in \mathbb{Z} \text{ mit } kd = a. \quad (\text{Teilbarkeit auf } \mathbb{Z})$$

Gilt  $d \mid a$ , so heißt  $d$  ein *Teiler* oder *Divisor* von  $a$  und  $a$  ein (*ganzzahliges*) *Vielfaches* von  $d$ . Es gilt

$$\text{Def}(\mid) = \text{Bild}(\mid) = \mathbb{Z}.$$

Die  $\mid$ -Relation ist reflexiv und transitiv. Sie ist nicht antisymmetrisch, da  $-2 \mid 2$  und  $2 \mid -2$ , aber  $2 \neq -2$ .



- (3) Sei  $m \in \mathbb{N} - \{0\}$ . Dann setzen wir für alle  $a, b \in \mathbb{Z}$

$$a \equiv_m b, \text{ falls } m \mid (a - b). \quad (\text{Kongruenz modulo } m)$$

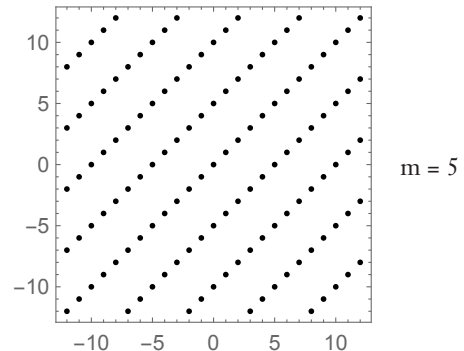
Gilt  $a \equiv_m b$ , so sagen wir, dass die Zahlen  $a$  und  $b$  *kongruent modulo*  $m$  sind.

Die Relation  $\equiv_m$  ist reflexiv, symmetrisch und transitiv.

Wir schreiben oftmals auch  $a \equiv b \pmod{m}$  anstelle von  $a \equiv_m b$ . So gilt zum Beispiel

$$0 \equiv 5 \equiv -25 \pmod{5},$$

$$-5 \equiv 2 \equiv 16 \pmod{7}.$$



## 1.2 Äquivalenzrelationen

**Definition** (*Äquivalenzrelation, Äquivalenzklasse, Repräsentantensystem*)

### Äquivalenzrelationen

Eine Relation  $\sim$  auf  $A$  heißt eine *Äquivalenzrelation* oder kurz eine *Äquivalenz*, falls  $\sim$  reflexiv, symmetrisch und transitiv ist. Gilt  $a \sim b$  für  $a, b \in A$ , so sagen wir, dass  $a$  und  $b$  *äquivalent* (bzgl.  $\sim$ ) sind.

### Äquivalenzklassen und Faktorisierung

Wir setzen

$$a/\sim = \{b \in A \mid a \sim b\} \quad \text{für alle } a \in A, \quad (\text{Äquivalenzklasse von } a, a \text{ modulo } \sim)$$

$$A/\sim = \{a/\sim \mid a \in A\}. \quad (\text{Faktorisierung, } A \text{ modulo } \sim)$$

### Repräsentanten und Repräsentantensysteme

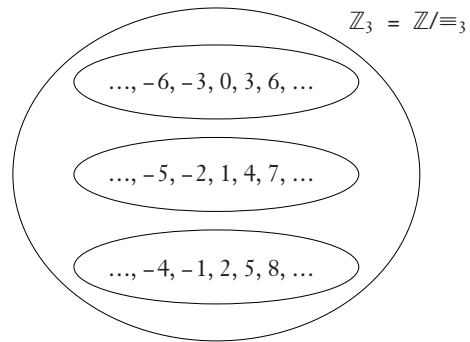
Gilt  $b \sim a$ , so heißt  $b$  ein *Repräsentant* der Äquivalenzklasse  $a/\sim$ . Eine Menge  $B \subseteq A$  heißt ein (*vollständiges*) *Repräsentantensystem* für die Äquivalenz  $\sim$ , falls es für alle  $a \in A$  genau ein  $b \in B$  mit  $a \sim b$  gibt.

Eine Äquivalenzrelation bringt eine „Ähnlichkeit“, „Gleichwertigkeit“, „Gleichheit in bestimmter Hinsicht“ zum Ausdruck. Sie beschreibt das Absehen von als unwesentlich erachteten Eigenschaften und damit das Abstrahieren. Die Begriffsbildung ist die „Abstraktion der Abstraktion“.

Das Trio „reflexiv, symmetrisch, transitiv“ lässt sich durch die Eigenschaften der Gleichheit motivieren. Denn für alle  $a, b, c$  gilt

$$a = a, \quad a = b \text{ impliziert } b = a,$$

$$a = b \text{ und } b = c \text{ impliziert } b = c.$$



Die Kongruenz modulo 3 auf  $\mathbb{Z}$  besitzt drei Äquivalenzklassen. Die Mengen  $\{0, 1, 2\}$  und  $\{0, -2, 8\}$  sind zwei Beispiele für vollständige Repräsentantensysteme.

### Notationen

- (1) Statt  $a/\sim$  schreibt man auch  $[a]_\sim$  oder auch nur  $[a]$ , wenn  $\sim$  aus dem Kontext heraus klar ist. Daneben ist auch  $\bar{a}$  für  $[a]$  üblich.
- (2) Äquivalenzrelationen können auch mit  $R, S, \dots$  bezeichnet werden. Meistens werden jedoch Zeichen wie

$$\sim, \sim^*, \approx, \equiv, \simeq, \cong$$

verwendet, die an das Gleichheitssymbol  $=$  erinnern.



Die Faktorisierung  $A/\sim$  ist ein Mengensystem. Jedes Element  $a/\sim$  von  $A/\sim$  ist eine Teilmenge von  $A$  und damit gilt  $A/\sim \subseteq \mathcal{P}(A)$ . Es gilt

$$(\#) \quad a/\sim \neq \emptyset; \quad a/\sim \cap b/\sim = \emptyset \quad \text{genau dann, wenn} \quad \text{non}(a \sim b); \quad \bigcup A/\sim = A.$$

Wie die Menge der Schüler einer Schule in Schulklassen zerfällt, so zerfällt  $A$  in Äquivalenzklassen. Weitere Alltagsbeispiele sind die Einteilung von Kleidungsstücken in die Größenklassen XS, S, M, L, XL, die Zustandsbeschreibungen „neu, wie neu, gebraucht, akzeptabel“, die Einteilung der Welt in Länder und jede Form der Teambildung im Sport.

Die drei Eigenschaften in (#) besagen, dass  $\mathcal{A} = A/\sim$  eine Zerlegung der Menge  $A$  bildet (vgl. 0.4). Ist umgekehrt  $\mathcal{A}$  eine Zerlegung von  $A$ , so definiert

$$a \sim b, \quad \text{falls} \quad \text{es gibt ein } A \in \mathcal{A} \text{ mit } a, b \in A \quad \text{für alle } a, b \in A$$

eine Äquivalenzrelation auf  $A$  mit  $A/\sim = \mathcal{A}$ . Damit gilt:

*Äquivalenzrelationen und Zerlegungen entsprechen einander.*

Wählen wir aus jeder Äquivalenzklasse  $a/\sim$  genau ein Element aus und fassen wir die ausgewählten Elemente zu einer Menge  $B \subseteq A$  zusammen, so erhalten wir ein Repräsentantensystem (vgl. 1.11 zu „wählen“). Im Schulbeispiel: Klassensprecherversammlung.

## Beispiele

- (1) Für alle  $m \geq 1$  ist die Kongruenz  $\equiv_m$  eine Äquivalenz auf  $\mathbb{Z}$  (vgl. 1.1). Wir schreiben kurz  $[a]_m$  oder  $[a]$  statt  $a/\equiv_m$  und  $\mathbb{Z}_m$  statt  $\mathbb{Z}/\equiv_m$ . Für  $m = 3$  gilt

$$[0] = [3] = [-3] = \dots = \{\dots, -6, -3, 0, 3, 6, \dots\},$$

$$[1] = [4] = [-2] = \dots = \{\dots, -5, -2, 1, 4, 7, \dots\},$$

$$[2] = [5] = [-1] = \dots = \{\dots, -4, -1, 2, 5, 8, \dots\}.$$

Die Menge  $\{0, 1, 2\}$  ist ein Repräsentantensystem. Man nennt es „kanonisches“ oder „Standard-Repräsentantensystem“, da sich 0, 1, 2 bei Division durch 3 als Reste anbieten. Aber auch  $\{0, -1, -2\}$  und  $\{3, 7, -7\}$  sind prinzipiell gleichwertige Repräsentantensysteme. Es gilt

$$\mathbb{Z}_3 = \mathbb{Z}/\equiv_3 = \{[a] \mid a \in \mathbb{Z}\} = \{[0], [1], [2]\} = \{[0], [-1], [-2]\} \text{ usw.}$$

- (2) Die geometrische Kongruenz (Deckungsgleichheit) zweier Teilmengen  $A, B$  der Ebene  $\mathbb{R}^2$  ist eine Äquivalenz auf  $\mathcal{P}(\mathbb{R}^2) = \{A \mid A \subseteq \mathbb{R}^2\}$ . Ebenso ist die Ähnlichkeit von Dreiecken eine Äquivalenz auf der Menge aller Dreiecke.
- (3) Die Relation  $\sim = \{(a, a) \mid a \in A\}$  ist eine Äquivalenzrelation auf  $A$  (Motto: „Jeder ist anders.“, „Einzelunterricht“). Es gilt  $a/\sim = \{a\}$  für alle  $a \in A$  und  $A/\sim = \{\{a\} \mid a \in A\}$ . Die Menge  $A$  ist das einzige Repräsentantensystem.
- (4) Die Relation  $\sim = \{(a, b) \mid a, b \in A\} = A^2$  ist eine Äquivalenz auf  $A$  (Motto: „Alle sind gleich.“, „Dorfschule mit einer Klasse“). Es gilt  $a/\sim = A$  für alle  $a \in A$  und  $A/\sim = \{A\}$ . Für alle  $a \in A$  ist  $\{a\}$  ein Repräsentantensystem.

### 1.3 Ordnungen

#### Definition (partielle und lineare Ordnungen)

##### Partielle und lineare Ordnung

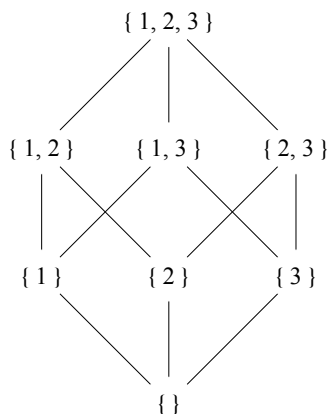
Eine Relation  $\leq$  auf  $A$  heißt eine (*partielle*) *Ordnung* auf  $A$ , falls  $\leq$  reflexiv, antisymmetrisch und transitiv auf  $A$  ist. Für alle  $a, b \in A$  setzen wir  $a < b$ , falls  $a \leq b$  und  $a \neq b$ .

Sind  $a, b \in A$  mit  $a \leq b$  oder  $b \leq a$ , so heißen  $a$  und  $b$  *vergleichbar*. Sind je zwei Elemente vergleichbar, so heißt die Ordnung  $\leq$  *linear* oder *total*.

##### Ordnungsbegriffe

Seien  $\leq$  eine partielle Ordnung auf  $A$ ,  $a \in A$  und  $X \subseteq A$ .

a heißt ...	in Zeichen	falls ...
obere Schranke von $X$	$X \leq a, a \geq X$	für alle $x \in X$ gilt $x \leq a$
untere Schranke von $X$	$a \leq X, X \geq a$	für alle $x \in X$ gilt $a \leq x$
Maximum von $X$	$a = \max(X)$	$a \in X$ und $X \leq a$
Minimum von $X$	$a = \min(X)$	$a \in X$ und $a \leq X$
Supremum von $X$	$a = \sup(X)$	$X \leq a$ und für alle $b \geq X$ gilt $a \leq b$
Infimum von $X$	$a = \inf(X)$	$a \leq X$ und für alle $b \leq X$ gilt $b \leq a$
maximal in $X$	–	$a \in X$ und es gibt kein $x \in X$ mit $a < x$
minimal in $X$	–	$a \in X$ und es gibt kein $x \in X$ mit $x < a$

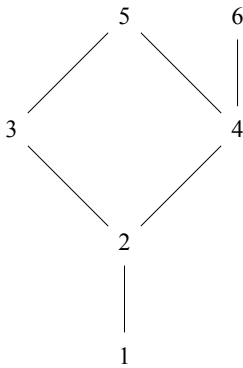


Die Inklusion  $\subseteq$  ist eine partielle Ordnung auf

$$A = \mathcal{P}(\{1, 2, 3\}) =$$

$$\{\{\}, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{1, 3\}, \{2, 3\}, \{1, 2, 3\}\}.$$

Sie lässt sich durch ein sog. Hasse-Diagramm darstellen: Die Ordnung wird durch Linien angezeigt, wobei größere Elemente über kleineren stehen. In der Ordnung ist  $\{1\}$  kleiner als  $\{1, 2, 3\}$ , während  $\{1\}$  und  $\{2, 3\}$  unvergleichbar sind. Auch viele Ordnungen auf unendlichen Mengen kann man in verwandter Weise visualisieren, man denke etwa an Zahlenstrahldarstellungen von  $\mathbb{N}$ ,  $\mathbb{Z}$ ,  $\mathbb{Q}$  oder  $\mathbb{R}$ .



Wir betrachten die durch das Hasse-Diagramm links dargestellte partielle Ordnung  $\leq$  auf

$$A = \{1, 2, 3, 4, 5, 6\}$$

und die Teilmenge  $X = \{2, 3, 4\}$  von  $A$ . Es gilt:

- (a) 5 ist eine obere Schranke von  $X$ , 6 ist keine obere Schranke von  $X$ ,
- (b) 1 und 2 sind untere Schranken von  $X$ ,
- (c)  $2 = \min(X)$ ,  $\max(X)$  existiert nicht,
- (d) 3 und 4 sind maximal in  $X$ , 2 ist minimal in  $X$ .

Im Gegensatz zu einer Äquivalenzrelation, die eine Menge  $A$  in disjunkte Äquivalenzklassen unterteilt, bringt eine partielle Ordnung die Elemente von  $A$  in eine netzartige Struktur. Ist die Ordnung linear (total), so wird  $A$  in die Form einer „Kette“ oder abstrakten „Linie“ gebracht.

Ist  $\leq$  eine partielle Ordnung auf  $A$ , so ist die zugehörige Relation  $<$  irreflexiv und transitiv (und damit antisymmetrisch). Eine irreflexive und transitive Relation auf  $A$  nennt man auch eine *strikte partielle Ordnung* auf  $A$ . Ist  $<$  eine strikte partielle Ordnung auf  $A$ , so erhält man eine partielle Ordnung  $\leq$  auf  $A$  durch

$$a \leq b, \quad \text{falls} \quad a < b \text{ oder } a = b \quad \text{für alle } a, b \in A.$$

Es ist also Geschmackssache, ob man  $\leq$  oder  $<$  bevorzugt. Man hat immer beides.

Für partielle Ordnungen werden meistens Zeichen wie  $\leq$ ,  $\leqslant$ ,  $\leq^*$  mit Unterstrich und die zugehörigen strikten Versionen  $<$ ,  $<$ ,  $<^*$  verwendet.

## Beispiele

(1) Die üblichen  $\leq$ -Relationen auf  $\mathbb{N}$ ,  $\mathbb{Z}$ ,  $\mathbb{Q}$  und  $\mathbb{R}$  sind lineare Ordnungen.

(2) Ist  $\leq$  eine lineare Ordnung auf  $A$ , so definiert

$$(a, b) \leq_{\text{lex}} (c, d), \quad \text{falls} \quad a \leq c \text{ oder } (a = c \text{ und } b \leq d)$$

eine lineare Ordnung auf  $A^2$ , die sog. *lexikographische Ordnung* auf  $A^2$ . Speziell kann die Menge  $\mathbb{R}^2 = \mathbb{C}$  in dieser Weise linear geordnet werden.

(3) Für jede Menge  $M$  ist die Inklusion  $\subseteq$  eine partielle Ordnung auf  $A = \mathcal{P}(M)$ . Sind  $a, b \in M$  verschieden, so sind  $\{a\}$  und  $\{b\}$  nicht vergleichbar. Für alle  $\mathcal{A} \subseteq \mathcal{P}(M)$  gilt  $\sup(\mathcal{A}) = \bigcup \mathcal{A}$  und  $\inf(\mathcal{A}) = \bigcap \mathcal{A}$ .

(4) Die Anfangsstückrelation auf der Menge der endlichen 01-Folgen ist eine baumartige partielle Ordnung. Es gilt zum Beispiel

$$010 \leq 01011, \quad \text{nicht } (010 \leq 11101).$$

Die Ordnung hat keine maximalen Elemente, da  $s \leq s0$  (und  $s \leq s1$ ) für alle  $s$  gilt.

## 1.4 Der Abbildungsbegriff

### Definition (Abbildung, Funktion)

Eine Relation  $f$  heißt eine *Abbildung* oder *Funktion*, falls für alle  $a, b, c$  gilt:

$a f b$  und  $a f c$  impliziert  $b = c$ . (Rechtseindeutigkeit)

Statt  $a f b$  (d.h.  $(a, b) \in f$ ) schreiben wir auch

$f(a) = b$  oder  $fa = b$  oder  $f : a \mapsto b$  oder  $a \xrightarrow{f} b$ .

Gilt  $\text{Def}(f) = A$ , so heißt  $f$  eine Funktion *auf*  $A$ .

Abbildungen sind also Mengen von geordneten Paaren  $(a, b)$ , bei denen der erste Eintrag den zweiten bestimmt. Bei Relationen können einem  $a$  viele  $b$  entsprechen:

$$1 < 3, 1 < 5, 2 \equiv_4 6, 2 \equiv_4 -2.$$

Für eine Abbildung  $f$  auf  $A$  gibt es zu jedem  $a \in A$  genau ein  $b$  mit  $a f b$ .

Ob man „Abbildung“ oder „Funktion“ sagt, ist einerlei. Wir wechseln ab.

f	
a	b
b	u
c	v
d	b
e	e

Eine Funktion  $f$  als Zuordnungstabelle. Es gilt

$$f(a) = b, f(b) = u, f(c) = v, \dots$$

Die Einträge in der rechten Spalte können mehrfach auftreten, den Einträgen in der linken Spalte entspricht immer ein eindeutiger rechter Eintrag.

### Beispiele

- (1)  $f = \{ (1, 2), (2, 1), (3, 1) \}$  ist eine Abbildung. Es gilt  $f(1) = 2, f(2) = 1, f(3) = 1$ . Dagegen ist  $R = \{ (1, 2), (1, 3), (2, 1) \}$  eine Relation, aber keine Abbildung. Die Rechtseindeutigkeit ist verletzt, da  $1 R 2$  und  $1 R 3$ . Die leere Menge  $\emptyset$  ist dagegen eine Abbildung.
- (2) Für jede Menge  $A$  heißt  $\text{id}_A = \{ (a, a) \mid a \in A \}$  die *Identität* auf  $A$ . Es gilt  $\text{id}_A(a) = a$  für alle  $a \in A$ ,  $\text{Def}(\text{id}_A) = \text{Bild}(\text{id}_A) = A$ .
- (3) Für jede Menge  $A$  und jedes  $c$  heißt  $\text{const}_c^A = \{ (a, c) \mid a \in A \}$  die *konstante Abbildung* auf  $A$  mit *Wert*  $c$ . Es gilt  $\text{const}_c^A(a) = c$  für alle  $a \in A$ ,  $\text{Def}(\text{const}_c^A) = A$ ,  $\text{Bild}(\text{const}_c^A) = \{ c \}$  für  $A \neq \emptyset$  und  $\text{Bild}(\text{const}_c^A) = \emptyset$  für  $A = \emptyset$ .
- (4) Ist  $A$  eine Menge und  $B \subseteq A$ , so heißt

$$\chi_B^A = 1_B^A = \{ (a, 1) \mid a \in B \} \cup \{ (a, 0) \mid a \in A - B \}$$

die *charakteristische Funktion* oder *Indikatorfunktion* von  $B$  bzgl.  $A$ . Für  $\chi = \chi_B^A$  gilt:  $\chi(a) = 1$ , falls  $a \in B$ ;  $\chi(a) = 0$ , falls  $a \in A - B$ ;  $\text{Def}(\chi) = A$ ;  $\text{Bild}(\chi) = \{ 0, 1 \}$ , falls  $B \neq \emptyset$ ;  $\text{Bild}(\chi) = \{ 0 \}$ , falls  $A \neq \emptyset$  und  $B = \emptyset$ ;  $\text{Bild}(\chi) = \emptyset$ , falls  $A = B = \emptyset$ .

Die mengentheoretische Notation dieser Beispiele illustriert die Definition. Andere und oft suggestivere Definitionsformen werden wir gleich kennenlernen.

„Die“ beste Interpretation, Anschauung oder Visualisierung einer Funktion existiert nicht. Wir stellen ohne Anspruch auf Vollständigkeit fünf Interpretationen vor.

### A Die Tabelleninterpretation

Abbildungen sind zweiseitige Tabellen. Die Bildung von  $b = f(a)$  entspricht dem Nachschlagen („table look-up“): Finde  $a$  in der linken und das zugehörige  $b = f(a)$  in der rechten Spalte. Entscheidend ist, dass der Übergang von  $a$  zu  $b$  eindeutig ist, d. h.,  $a$  darf links nur einmal auftauchen. Dagegen darf  $b$  in der rechten Spalte mehrfach erscheinen.

### B Die Zuordnungsinterpretation

Eine Abbildung „bildet ab“ (wie der Name schon sagt) oder „ordnet zu“. Jedem Element  $a$  einer Menge  $A$  wird ein eindeutiges  $b = f(a)$  einer Menge  $B$  zugeordnet. Die Mengen  $A$  und  $B$  visualisiert man oft als „getrennte Welten“ (obwohl  $A \cap B \neq \emptyset$  gelten kann), und Pfeile deuten an, auf welches  $b \in B$  ein  $a \in A$  abgebildet wird. Oft haben  $A$  und  $B$  algebraische oder geometrische Strukturen, die durch  $f$  respektiert werden oder auch nicht, etwa „Liegt  $a$  nahe bei  $b$ , so liegt  $f(a)$  nahe bei  $f(b)$ .“

### C Die kartesische Interpretation

Ist  $f \subseteq \mathbb{R}^2$ , so kann man  $f$  „plotten“, indem man jeden Punkt  $(a, f(a))$  für  $a \in \text{Def}(f)$  schwarz einfärbt. Jeder kennt die reelle Einheitsparabel und den reellen Sinus in dieser Form. Ist allgemein  $f \subseteq A \times B$  für beliebige Mengen  $A$  und  $B$ , so ist diese Interpretation noch abstrakt möglich, indem man sich  $A$  als Waagrechte und  $B$  als Senkrechte vorstellt.

### D Die Pfeilinterpretation

Das Feld  $\text{Def}(f) \cup \text{Bild}(f)$  von  $f$  sieht man als eine Menge von isolierten Punkten an. Von einem Punkt  $a$  des Feldes führt genau dann ein Pfeil zu einem Punkt  $b$  des Feldes, wenn  $b = f(a)$ . Die Funktion  $f$  „schickt“  $a$  nach  $b$  oder „zeigt“ von  $a$  nach  $b$ . Es entsteht ein Pfeilsystem, das zum Beispiel die Bildung von Bahnen  $a, f(a), f(f(a))$  (falls  $f(a) \in \text{Def}(f)$ ), ... suggeriert. Besonders geeignet ist diese Interpretation für endliche Felder.

### E Die Dienstboten-Interpretation

Eine Funktion  $f$  ist ein Dienstbote, dem man  $a$  übergeben kann und der einem irgendwann  $b = f(a)$  zurückbringt. Ersetzt man hier „Dienstbote“ durch „Computerprogramm“ und schränkt  $a$  auf Daten ein, so erhält man die berechenbaren Funktionen, die in der Informatik, Logik und in allen rechnerischen Anwendungen der Mathematik eine herausragende Rolle spielen. Der Funktionsbegriff der Mathematik ist wesentlich allgemeiner, die „Dienstboten“ sind in der Regel keine Computerprogramme, sondern abstrakte und ideale Gegenstände der Mathematik. Wie man Funktionen in der Mathematik konstruieren kann, beschreiben wir im folgenden Abschnitt.

## 1.5 Konstruktion von Abbildungen

### Satz (Konstruktionsmöglichkeiten für Abbildungen)

#### Direkte Angabe

Seien  $a_1, \dots, a_n$  paarweise verschiedene Objekte, und seien  $b_1, \dots, b_n$  beliebige Objekte. Dann gibt es genau eine Abbildung  $f$  auf  $\{a_1, \dots, a_n\}$  mit

$$f(a_k) = b_k \quad \text{für alle } 1 \leq k \leq n.$$

#### Termindefinitionen

Ist  $t(x)$  ein Term, also ein aus der Variablen  $x$ , Konstanten, Funktionszeichen und Klammern aufgebauter formaler Ausdruck, so existiert für jede geeignete Menge  $A$  genau eine Abbildung  $f$  auf  $A$  mit

$$f(a) = t[a] \quad \text{für alle } a \in A,$$

wobei  $t[a]$  der Wert ist, den man erhält, wenn man  $a$  für die Variable  $x$  in  $t$  einsetzt. „Geeignet“ heißt, dass diese Termauswertung für alle Elemente  $a$  der Menge  $A$  erklärt ist.

#### Eindeutige Eigenschaften

Sei  $A$  eine Menge und sei  $\mathcal{C}(a, b)$  eine Eigenschaft mit:

(+) Für alle  $a \in A$  gibt es genau ein  $b$  mit  $\mathcal{C}(a, b)$ .

Dann gibt es genau eine Abbildung  $f$  auf  $A$  mit

$$f(a) = \text{„das eindeutige } b \text{ mit } \mathcal{C}(a, b)\text{“} \quad \text{für alle } a \in A.$$

Die drei (letztendlich axiomatisch postulierten) Möglichkeiten versammeln die meisten in der Mathematik auftretenden Konstruktionen von Funktionen. Eine vierte Möglichkeit diskutieren wir in 1.11.

### Möglichkeit 1: Direkte Angabe

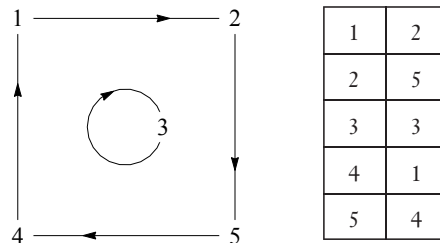
Hier gilt einfach  $f = \{(a_1, b_1), \dots, (a_n, b_n)\}$ . Die Rechtseindeutigkeit wird durch die Verschiedenheit der  $a_k$  sichergestellt. Die vielleicht wichtigsten derartigen Funktionen sind Permutationen:

#### Definition (Permutation)

Eine Funktion  $f$  auf  $\{1, \dots, n\}$  heißt *Permutation*, falls

$$\text{Bild}(f) = \{1, \dots, n\}.$$

Eine Permutation heißt *Transposition*, wenn es  $i \neq j$  gibt mit  $f(i) = j$ ,  $f(j) = i$  und  $f(k) = k$  für alle  $k \neq i, j$ .



Eine Permutation auf  $\{1, 2, 3, 4, 5\}$

Wir schreiben  $\begin{pmatrix} 1 & \cdots & n \\ b_1 & \cdots & b_n \end{pmatrix}$  oder kurz  $(b_1, \dots, b_n)$  für die Permutation  $f$  mit  $f(i) = b_i$  für alle  $i \in \{1, \dots, n\}$ .

### Beispiele

- (1)  $(1, 2, 3, 4), (1, 4, 3, 2), (4, 3, 2, 1)$  sind Permutationen auf  $\{1, 2, 3, 4\}$ .
- (2)  $(1, 3, 2), (2, 1, 3), (3, 2, 1)$  sind alle möglichen Transpositionen auf  $\{1, 2, 3\}$ .

### Möglichkeit 2: Termdefinitionen

Dies ist aus der Schule bekannt:  $f(x)$  = „rechnerischer Ausdruck in  $x$ “. Wir verzichten auf eine Präzisierung von „Term“ und „Termauswertung“ und begnügen uns mit:

### Beispiel

Sind  $\sin(x)$  und  $\cos(x)$  für alle  $x \in \mathbb{R}$  bereits definiert, so können wir eine Funktion  $f$  mit  $\text{Def}(f) = \mathbb{R}$  definieren durch

$$f(x) = \sin^2(x) + \cos^2(x) \quad \text{für alle } x \in \mathbb{R}.$$

Es stellt sich heraus, dass  $f(x) = 1$  für alle  $x \in \mathbb{R}$  gilt, sodass  $f = \text{const}_1^{\mathbb{R}}$ .

### Möglichkeit 3: „das (eindeutige) $y$ mit ...“

Dass Termdefinitionen nicht ausreichen, hat Leonhard Euler bereits im 18. Jahrhundert anhand von Fragestellungen wie in (1) bemerkt.

### Beispiele

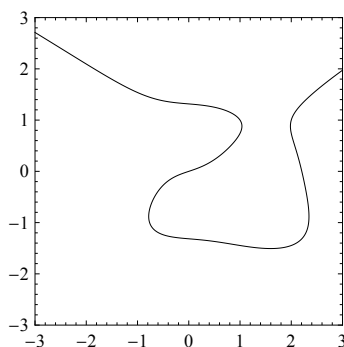
- (1) Wir betrachten für  $x, y \in \mathbb{R}$  die Eigenschaft

$$\mathcal{E}(x, y) = \text{„}y^5 - x^4 + 2x^3 - 3y + x = 0\text{“}$$

und definieren  $f$  mit  $\text{Def}(f) = [0, 2]$  durch

$$f(x) = \text{„das } y \in [-2, -1] \text{ mit } \mathcal{E}(x, y)\text{“}$$

für alle  $x \in [0, 2]$ . Die Eindeutigkeit von  $y$  muss man natürlich erst beweisen (die Abbildung rechts deutet an, dass sie gilt). Eine Termdefinition ist dagegen nicht ersichtlich. Gleiches gilt für:



Die Punkte  $(x, y) \in [-3, 3]^2$  mit  $\mathcal{E}(x, y)$

- (2) Sei  $g$  auf  $\{n \in \mathbb{N} \mid n \geq 1\}$  definiert durch

$$g(n) = \text{„das } k \geq 0 \text{ mit } 2^k \mid n \text{ und nicht } (2^{k+1} \mid n)\text{“} \quad \text{für alle } n \geq 1.$$

Die Abbildung  $g$  gibt für alle  $n \geq 1$  den Exponenten der 2 in der Primfaktorzerlegung von  $n$  an. Zum Beispiel gilt

$$g(8) = g(2^3) = 3, \quad g(120) = g(2^3 \cdot 3^1 \cdot 5^1) = 3, \quad g(1) = g(15) = 0.$$

## 1.6 Notationen und Sprechweisen für Abbildungen

**Definition** (*Stellen, Werte, Definiertheit,  $f: A \rightarrow B$ , Familien*)

Sei  $f$  eine Funktion auf  $A$ .

### Allgemeine Sprechweisen

Gilt  $f(a) = b$ , so heißt  $b$  der *Wert* von  $f$  an der *Stelle* oder für das *Argument*  $a$ .

Wir sagen, dass  $a$  durch  $f$  auf  $b$  *abgebildet* wird, dass die *Anwendung* von  $f$  auf  $a$  den Wert  $b$  *ergibt* oder dass  $f$  an der Stelle  $a$  den Wert  $b$  *annimmt*. Die Sprechweise „ $f(a)$  ist definiert“ ist gleichbedeutend mit  $a \in A$ .

### Die Notation $f: A \rightarrow B$

Ist  $B$  eine Menge mit  $\text{Bild}(f) = \{f(a) \mid a \in A\} \subseteq B$ , so schreiben wir

$$f: A \rightarrow B \quad \text{oder} \quad f: A \ni a \mapsto f(a) \in B.$$

Wir sagen dann, dass  $f$  eine Abbildung *von  $A$  nach  $B$*  oder *zwischen  $A$  und  $B$*  ist.

Die Menge  $B$  heißt ein *Wertevorrat* oder eine *Zielmenge* von  $f$ .

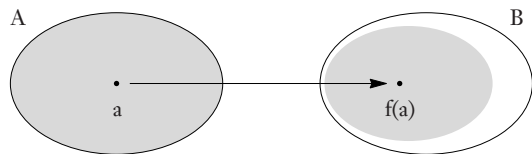
### Familien

Wir schreiben  $f$  auch in der *Familien-* oder *Folgennotation*

$$(b_a)_{a \in A} \quad \text{oder} \quad (b_a \mid a \in A), \quad \text{mit} \quad b_a = f(a) \quad \text{für alle} \quad a \in A.$$

Wir nennen eine in der Form  $(b_a)_{a \in A}$  notierte Funktion eine *Familie* mit *Indexmenge*  $A$  oder eine  *$A$ -Folge*. Gilt  $b_a \in B$  für alle  $a$ , so heißt sie eine Familie *in  $B$* .

Ein Wertevorrat  $B$  ist vom Wertebereich  $\text{Bild}(f)$  zu unterscheiden. Der *Wertebereich* ist die Menge aller angenommenen Werte, so wie der *Definitionsbereich* die Menge aller definierten Stellen ist (die Sprechweisen sind in der Literatur nicht einheitlich, aber zwischen  $B$  und  $\text{Bild}(f)$  wird immer unterschieden). In  $f: A \rightarrow B$  wird nur verlangt, dass  $B$  eine Obermenge des Wertebereichs ist; die Bezeichnung als „Wertevorrat“ oder „Zielmenge“ deutet dies an. Die unsymmetrische Behandlung von  $A$  und  $B$  liegt daran, dass oft viele verschiedene Abbildungen zwischen fest gewählten Mengen  $A$  und  $B$  eingeführt und untersucht werden. Sie besitzen unterschiedliche Bilder und oft ist das Bild einer Abbildung zunächst auch gar nicht bekannt.



Eine Funktion  $f: A \rightarrow B$  bildet jedes Element  $a$  des Definitionsbereichs  $A$  auf ein Element  $f(a)$  des Wertevorrats  $B$  ab. Der Wertebereich  $\text{Bild}(f)$  ist die Menge aller Funktionswerte (im Diagramm rechts grau dargestellt). Er kann mit  $B$  zusammenfallen oder eine echte Teilmenge von  $B$  sein.

### Beispiele

Wir können  $\sin: \mathbb{R} \rightarrow \mathbb{R}$ ,  $\sin: \mathbb{R} \rightarrow [-1, 1]$  schreiben, nicht aber  $\sin: \mathbb{R} \rightarrow [0, 1]$ .

Ebenso gilt  $\text{const}_0^{\mathbb{N}}: \mathbb{N} \rightarrow \mathbb{N}$ ,  $\text{const}_0^{\mathbb{N}}: \mathbb{N} \rightarrow \{0, 1\}$ ,  $\text{id}_A: A \rightarrow B$  für alle  $B \supseteq A$  usw.



## Angabe von Abbildungen

Um sich selbst und die mathematische Mitwelt nicht unglücklich zu machen, müssen Abbildungen immer genau angegeben werden:

### Beispiel

Die Sprechweise „die Funktion  $1/x$ “ ist ungenau. Es wird nicht klar, welchen Definitionsbereich die Funktion hat. Weiter benötigt man oft auch ein Funktionszeichen  $f, g, h, F, G, H, \dots$ . Eindeutig sind, mit  $\mathbb{R}^* = \mathbb{R} - \{0\}$ :

(1) Sei  $f: \mathbb{R}^* \rightarrow \mathbb{R}$  mit  $f(x) = 1/x$  (für alle  $x \in \mathbb{R}^*$ ). (*ordentlich*)

(2) Sei  $f: \mathbb{R}^* \ni x \mapsto 1/x \in \mathbb{R}$ . (*kompakt*)

(3) Sei  $f$  die  $1/x$ -Funktion auf  $\mathbb{R}^*$ . (*Term mit Definitionsbereich*)

Im Kontext von  $\mathbb{R}$  nicht unbedingt üblich, aber kurz und exakt sind zudem:

(4) Sei  $f = \{(x, 1/x) \mid x \in \mathbb{R}^*\}$ . (*extensional*)

(5) Sei  $f = (1/x)_{x \in \mathbb{R}^*}$ . (*familiär*)

Das Objekt, über das man redet, wird andernfalls nicht klar. Und Genauigkeit vermeidet Fehlvorstellungen: Die Frage, ob  $f$  unstetig an der Stelle 0 ist, welche viele Anfänger falsch mit „ja“ beantworten, wird hinfällig. Denn  $f$  ist im Nullpunkt nicht definiert, die Stetigkeit oder Unstetigkeit einer Funktion wird aber nur für Elemente des Definitionsbereichs erklärt. Die Funktion  $f$  ist (überall) stetig.

## Familien

Familien und Funktionen sind ein und dasselbe: Ist  $f$  eine Funktion auf einer Menge  $A$ , so gilt  $f = (f(a))_{a \in A} = (f(a) \mid a \in A)$ . Umgekehrt ist eine Familie  $(b_a \mid a \in A)$  die Funktion  $f$  auf  $A$  mit  $f(a) = b_a$  für alle  $a \in A$ . Die Familien-Notation bringt aber eine eigene Dynamik mit sich. Die Analysis käme ohne Folgen  $(x_n)_{n \in \mathbb{N}}$  (Familien mit der Indexmenge  $\mathbb{N}$ ) nicht aus; es wäre viel zu umständlich, sie jedes Mal in der Form  $f: \mathbb{N} \rightarrow \mathbb{R}$  anzugeben.

### Beispiel

„Die Folge  $(2n)_{n \in \mathbb{N}}$  ist monoton steigend.“ ist viel prägnanter als die gleichwertige Aussage „Die Funktion  $f: \mathbb{N} \rightarrow \mathbb{R}$  mit  $f(n) = 2n$  für alle  $n$  ist monoton steigend.“

Familien sind nützlich zur Parametrisierung. Eine Familie  $(f_t \mid t \in \mathbb{R})$  kann zum Beispiel eine suggestive Notation für Funktionen  $f_t: \mathbb{R}^3 \rightarrow \mathbb{R}$  sein, die von einem zeitlichen Parameter  $t$  abhängen. Oft verwendete „neutrale“ Indexmengen sind  $I$  und  $J$ .

Viele Notationen für Familien sind fast selbsterklärend, etwa

$$\bigcup_{i \in I} A_i = \bigcup \{A_i \mid i \in I\} = \{a \mid \text{es gibt ein } i \in I \text{ mit } a \in A_i\},$$

$$(A_i \cap B_i \mid i \in I) = (C_i \mid C_i = A_i \cap B_i \text{ für alle } i \in I)$$

für gegebene Familien  $(A_i \mid i \in I)$  und  $(B_i \mid i \in I)$  von Mengen.

## 1.7 Umgang mit Funktionen

**Definition** (*Komposition, Bild, Urbild, Einschränkung, abgeschlossen unter, Produkte*)

### Komposition

Sind  $f, g$  Funktionen mit  $\text{Bild}(f) \subseteq \text{Def}(g)$ , so heißt

$$g \circ f = \{ (a, g(f(a))) \mid a \in \text{Def}(f) \} \quad (\text{gelesen: „} g \text{ kringel } f \text{“, „} g \text{ nach } f \text{“})$$

die *Komposition* oder *Verknüpfung* von  $f$  und  $g$ .

### Bild und Urbild

Ist  $f$  eine Funktion, so definieren wir für alle  $X \subseteq \text{Def}(f)$  und alle Mengen  $Y$ :

$$f[X] = f(X) = \{ f(x) \mid x \in X \}, \quad (\text{Bild von } X)$$

$$f^{-1}[Y] = f^{-1}(Y) = \{ a \in \text{Def}(f) \mid f(a) \in Y \}. \quad (\text{Urbild von } Y)$$

Für alle  $y$  heißt das Urbild  $f^{-1}[\{y\}]$  auch die *Faser* von  $f$  über  $y$ .

### Einschränkung

Ist  $f$  eine Funktion und  $C \subseteq \text{Def}(f)$ , so heißt

$$f|_C = \{ (a, f(a)) \mid a \in C \} \quad (\text{gelesen: „} f \text{ eingeschränkt auf } C \text{“})$$

die *Einschränkung* von  $f$  auf  $C$ .

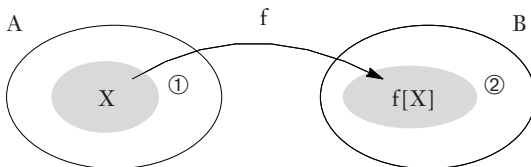
### Funktionenmengen und allgemeine Produkte

Für Mengen  $A$  und  $B$  definieren wir

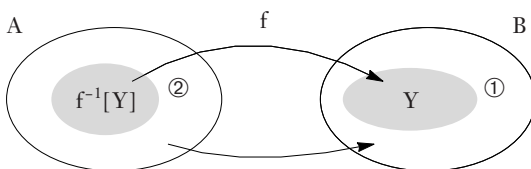
$$B^A = {}^A B = \{ f \mid f : A \rightarrow B \}.$$

Ist  $(B_i)_{i \in I}$  eine Familie, so definieren wir

$$\prod_{i \in I} B_i = \{ (b_i)_{i \in I} \mid b_i \in B_i \text{ für alle } i \in I \}. \quad (\text{kartesisches Produkt von } (B_i)_{i \in I})$$



Bei der Formung des Bildes einer Menge unter einer Funktion  $f : A \rightarrow B$  startet man mit einer Teilmenge  $X$  des Definitionsbereichs  $A$  von  $f$  (①) und sammelt alle Funktionswerte  $f(x)$  mit  $x \in X$  (②).



Ist ein Urbild zu bestimmen, so startet man mit einer Menge  $Y$  (①) und sammelt alle Stellen  $x$  des Definitionsbereichs von  $f$ , die durch Anwendung von  $f$  in der Menge  $Y$  landen (②). Alle anderen Elemente von  $A$  landen außerhalb von  $Y$ .

Für  $f: A \rightarrow B$  und  $g: B \rightarrow C$  gilt  $g \circ f: A \rightarrow C$ ,  $(g \circ f)(a) = g(f(a))$  für alle  $a \in A$ . Die Verknüpfung  $g \circ f$  beschreibt die Hintereinanderausführung von  $f$  und  $g$ : zuerst  $f$ , dann  $g$ , also „ $g$  nach  $f$ “, obwohl  $g$  vor  $f$  steht. Sie ist assoziativ,

$$h \circ (g \circ f) = (h \circ g) \circ f \quad \text{für alle } f: A \rightarrow B, g: B \rightarrow C, h: C \rightarrow D,$$

sodass wir  $h \circ g \circ f$  schreiben können. Es gilt  $(h \circ g \circ f)(a) = h(g(f(a)))$  für alle  $a \in A$ .

Für die Einschränkung  $f|C$  gilt:  $f|C$  ist wie  $f$ , aber nur noch für  $a \in C$  definiert.

## Beispiele

- (1) Seien  $f, g: \mathbb{N} \rightarrow \mathbb{N}$  mit  $f(n) = n + 1$ ,  $g(n) = n^2$  für alle  $n$ . Dann gilt  $f \circ g \neq g \circ f$ , da

$$(g \circ f)(n) = g(f(n)) = (n + 1)^2, \quad (f \circ g)(n) = f(g(n)) = n^2 + 1 \quad \text{für alle } n.$$

- (2) Für  $f: \mathbb{N} \rightarrow \mathbb{N}$  mit  $f(n) = n^2$  für alle  $n$  gilt

$$f[\{0, 4, 7\}] = \{0, 16, 49\}, \quad f^{-1}[\{1, \dots, 10\}] = \{1, 2, 3\}.$$

- (3) Für  $\cos: \mathbb{R} \rightarrow \mathbb{R}$  ist  $\cos_0 = \cos|_{[0, \pi]}$  ein „Ausschnitt“ des Kosinus, der zur Definition des Arkuskosinus verwendet werden kann (vgl. 1.9).

- (4)  $\mathbb{R}^{\mathbb{R}}$  ist die Menge aller reellen Funktionen  $f: \mathbb{R} \rightarrow \mathbb{R}$ . Ebenso ist  $\mathbb{R}^{\mathbb{N}}$  die Menge aller reellen Folgen  $(x_n)_{n \in \mathbb{N}}$  und  $\{0, 1\}^{\mathbb{N}}$  die Menge aller 0-1-Folgen.

- (5) Ist  $B_n = \{0, \dots, n\}$  für alle  $n$ , so ist  $\prod_{n \in \mathbb{N}} B_n$  die Menge aller  $f: \mathbb{N} \rightarrow \mathbb{N}$  mit  $f(n) \leq n$  für alle  $n$ .

- (6) Ist  $B_{2n} = A$  und  $B_{2n+1} = B$  für alle  $n$ , so ist  $\prod_{n \in \mathbb{N}} B_n$  die Menge aller Folgen, die abwechselnd Werte in den Mengen  $A$  und  $B$  annehmen.

- (7) Ist  $B_a = B$  für alle  $a \in A$ , so ist  $\prod_{a \in A} B_a = \prod_{a \in A} B = B^A$ .

Häufig verwendete Regeln für Bild und Urbild einer Funktion  $f$  sind:

$f[\text{Def}(f)] = \text{Bild}(f)$	$f^{-1}[\text{Bild}(f)] = \text{Def}(f)$
$f[f^{-1}[Y]] \subseteq Y$	$f^{-1}[f[X]] \supseteq X$
$f[X_1 \cap X_2] \subseteq f[X_1] \cap f[X_2]$	$f^{-1}[Y_1 \cap Y_2] = f^{-1}[Y_1] \cap f^{-1}[Y_2]$
$f[X_1 \cup X_2] = f[X_1] \cup f[X_2]$	$f^{-1}[Y_1 \cup Y_2] = f^{-1}[Y_1] \cup f^{-1}[Y_2]$
$f[X_1 - X_2] \supseteq f[X_1] - f[X_2]$	$f^{-1}[Y_1 - Y_2] = f^{-1}[Y_1] - f^{-1}[Y_2]$

## Beispiel

Für die Funktion  $f = \{(0, 2), (1, 2)\}$  gilt

$$f[\{0\} \cap \{1\}] = \emptyset \neq \{2\} = \{2\} \cap \{2\} = f[\{0\}] \cap f[\{1\}].$$

## 1.8 Operationen und Abgeschlossenheit

### Definition (Operation, abgeschlossen unter)

#### Operationen oder Verknüpfungen

Gilt  $f : A^n \rightarrow A$  für ein  $n \in \mathbb{N}$ , so heißt  $f$  eine ( $n$ -stellige) *Operation* oder *Verknüpfung* auf  $A$ . Wir schreiben

$$f(a_1, \dots, a_n) \text{ statt } f((a_1, \dots, a_n)) \text{ für alle } a_1, \dots, a_n \in A.$$

Ist  $n = 2$  und  $f$  ein Zeichen wie  $+$ ,  $\cdot$ ,  $\circ$ , ..., so schreiben wir auch

$$a + b, a \cdot b, a \circ b, \dots \text{ statt } +(a, b), \cdot(a, b), \circ(a, b), \dots \text{ für alle } a, b \in A.$$

#### Abgeschlossenheit einer Menge unter einer Operation

Ist  $f : A^n \rightarrow A$  eine Operation und  $B \subseteq A$ , so heißt  $B$  *abgeschlossen* unter  $f$ , falls

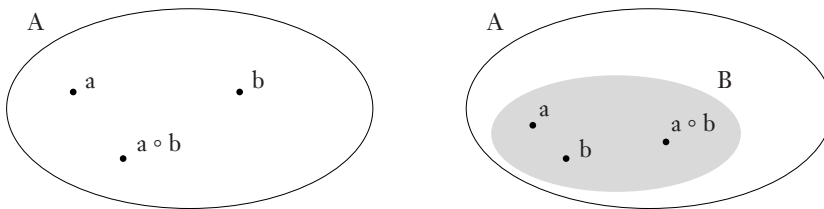
$$f(a_1, \dots, a_n) \in B \text{ für alle } a_1, \dots, a_n \in B. \quad (\text{Abgeschlossenheitsbedingung})$$

#### Abschluss einer Menge unter einer Operation

Ist  $f : A^n \rightarrow A$  und  $B \subseteq A$ , so setzen wir:

$$\langle B \rangle = \langle B \rangle_f = \text{„die } \subseteq\text{-kleinste unter } f \text{ abgeschlossene Obermenge von } B\text{“}.$$

Die Menge  $\langle B \rangle$  heißt der *Abschluss* von  $B$  unter  $f$  oder die von  $f$  und  $B$  *erzeugte* Teilmenge von  $A$ . Ist  $B = \{b_1, \dots, b_m\}$ , so schreiben wir  $\langle b_1, \dots, b_m \rangle$  für  $\langle B \rangle$ .



Bei einer zweistelligen Operation  $\circ : A^2 \rightarrow A$  werden je zwei Elemente  $a$  und  $b$  von  $A$  auf ein Element  $a \circ b$  von  $A$  abgebildet;  $a$  und  $b$  müssen dabei nicht verschieden sein und es kann  $a \circ b \neq b \circ a$  gelten.

Eine Teilmenge  $B$  von  $A$  ist abgeschlossen unter der Operation, wenn die Anwendung der Operation nicht aus  $B$  herausführt, d. h., für alle  $a, b \in B$  ist  $a \circ b$  wieder ein Element von  $B$ .

Eine Funktion der Form  $f : A^n \rightarrow B$  heißt eine  $n$ -stellige Funktion auf  $A$  und man schreibt  $f(a_1, \dots, a_n)$  statt  $f((a_1, \dots, a_n))$ , um die Lesbarkeit zu vereinfachen. Operationen sind  $n$ -stellige Funktionen auf  $A$ , die Werte in  $A$  annehmen. Jedem  $n$ -Tupel  $(a_1, \dots, a_n)$  mit Einträgen in  $A$  wird ein  $a = f(a_1, \dots, a_n) \in A$  zugeordnet.

In der Algebra spielen vor allem zweistellige Operationen eine wichtige Rolle. Der Leser denke an die Addition und Multiplikation auf den Zahlenmengen  $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$ . Es gilt zum Beispiel  $+: \mathbb{N}^2 \rightarrow \mathbb{N}$  und  $\cdot: \mathbb{R}^2 \rightarrow \mathbb{R}$ . Die „innere Notation“  $n + m$  bzw.  $x \cdot y$  liefert die vertrauten Ausdrücke:  $n + (m + k)$  ist viel besser lesbar als  $+(n, +(m, k))$ .

Auch einstellige Operationen sind von Interesse. Hier gilt einfach  $f : A \rightarrow A$ . Ein Beispiel ist die Nachfolgerfunktion  $S : \mathbb{N} \rightarrow \mathbb{N}$  mit  $S(n) = n + 1$  für alle  $n$ .

### Abgeschlossenheit einer Menge $B$ unter $f$

Seien  $f : A^n \rightarrow A$  und  $B \subseteq A$ . Die Abgeschlossenheit von  $B$  bedeutet, dass die Anwendung von  $f$  auf je  $n$  Elemente in  $B$  stets Werte in  $B$  liefert: „ $f$  führt nicht aus  $B$  heraus.“ Gleichwertig ist, dass die Einschränkung von  $f$  auf  $B^n$  eine Operation auf  $B$  ist:

$$f|B^n : B^n \rightarrow B. \quad (\text{Abgeschlossenheitsbedingung, Umformulierung})$$

#### Beispiele

- (1) Die Menge  $G$  der geraden Zahlen ist abgeschlossen unter der Addition  $+$  auf  $\mathbb{N}$ , da  $n + m \in G$  für alle  $n, m \in G$ . Dagegen ist  $U = \mathbb{N} - G$  nicht abgeschlossen unter  $+$ , da  $1, 3 \in U$ , aber  $1 + 3 \notin U$ .
- (2)  $\mathbb{R}^+ = \{x \in \mathbb{R} \mid x > 0\}$  und  $\mathbb{R}_0^+ = \mathbb{R}^+ \cup \{0\}$  sind abgeschlossen unter der reellen Multiplikation, da das Produkt zweier positiver (nichtnegativer) Zahlen positiv (nichtnegativ) ist. Ebenso ist das Intervall  $[0, 1]$  abgeschlossen unter  $\cdot$ . Dagegen sind  $\{x \in \mathbb{R} \mid x \leq 0\}$  und  $[0, 2]$  nicht abgeschlossen unter  $\cdot$ .
- (3) Für jede Operation  $f : A^n \rightarrow A$  sind  $A$  und  $\emptyset$  abgeschlossen unter  $f$ .

### Abschluss einer Menge $B$ unter $f$

Für jedes  $B \subseteq A$  existiert eine bzgl. der Inklusion  $\subseteq$  kleinste unter  $f$  abgeschlossene Obermenge  $\langle B \rangle = \langle B \rangle_f$ . Sie lässt sich auf zwei äquivalente Arten konstruieren:

#### Konstruktion „von oben“: Schnittbildung

Es gilt  $\langle B \rangle = \bigcap \{C \mid C \supseteq B \text{ und } C \text{ ist abgeschlossen unter } f\}$ .

#### Konstruktion „von unten“: wiederholte Anwendung von $f$

Wir setzen  $B_0 = B$  und  $B_{k+1} = B_k \cup f[B_k^n]$  für alle  $k \in \mathbb{N}$ . Dann gilt

$$\langle B \rangle = \bigcup_{k \in \mathbb{N}} B_k.$$

#### Beispiele

- (1) Sei  $S : \mathbb{N} \rightarrow \mathbb{N}$  die Nachfolgerbildung, und sei  $B = \{4, 9\}$ . Dann gilt

$$B_0 = \{4, 9\}, \quad B_1 = B_0 \cup \{5, 10\}, \quad B_2 = B_1 \cup \{6, 11\}, \quad \dots$$

$$\text{Damit ist } \langle 4, 9 \rangle = \{n \in \mathbb{N} \mid n \geq 4\}.$$

- (2) Sei  $+: \mathbb{N}^2 \rightarrow \mathbb{N}$  die Addition, und sei  $B = \{4, 9\}$ . Dann gilt

$$(\#) \quad \langle 4, 9 \rangle = \{n4 + m9 \mid n, m \geq 1\}.$$

Denn die Menge  $M$  rechts ist abgeschlossen unter  $+$ , sodass  $\langle 4, 9 \rangle \subseteq M$ .

Ist  $C$  eine unter  $+$  abgeschlossene Obermenge von  $\{4, 9\}$ , so sind  $n4, m9$  und  $n4 + m9$  Elemente von  $C$  für alle  $n, m \geq 1$ . Also gilt  $M \subseteq C$  für alle unter  $+$  abgeschlossenen  $C \supseteq \{4, 9\}$ , sodass  $M \subseteq \langle 4, 9 \rangle$ .

## 1.9 Abbildungseigenschaften

**Definition** (*injektiv, surjektiv, bijektiv, Umkehrfunktion*)

*Abbildungseigenschaften*

Eine Funktion  $f$  heißt

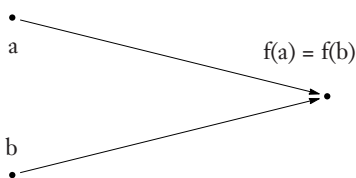
*injektiv*, falls  $f(a) \neq f(b)$  für alle  $a \neq b$  in  $\text{Def}(f)$ , (*Linkseindeutigkeit*)

*surjektiv nach B*, falls  $\text{Bild}(f) = B$ ,

*bijektiv nach B*, falls  $f$  ist injektiv und surjektiv nach B.

*Umkehrfunktion*

Ist  $f$  injektiv, so heißt  $f^{-1} = \{ (b, a) \mid (a, b) \in f \}$  die *Umkehrfunktion* von  $f$ .



Die Injektivität ist genau dann verletzt, wenn zwei verschiedene Stellen auf denselben Wert abgebildet werden.

$f$	
$a_1$	$b_1$
$a_2$	$b_2$
$a_3$	$b_3$
$a_4$	$b_4$
$a_5$	$b_5$

$f^{-1}$	
$b_1$	$a_1$
$b_2$	$a_2$
$b_3$	$a_3$
$b_4$	$a_4$
$b_5$	$a_5$

Ist  $f$  injektiv, so entsteht die Umkehrfunktion  $f^{-1}$  anschaulich durch Vertauschen der Spalten.

Die Injektivität von  $f$  bedeutet in der Tabelleninterpretation, dass auch in der rechten Spalte kein  $y$  mehrfach vorkommt. Sie ist eine Eigenschaft der Funktion  $f$  und benötigt im Gegensatz zur Surjektivität und Bijektivität keine Nennung einer Menge  $B$ .

Für eine Funktion $f : A \rightarrow B$ bedeutet ...	
<i>injektiv</i>	kein Wert (in $B$ ) wird mehrfach angenommen
<i>surjektiv</i>	jeder Wert in $B$ wird mindestens einmal angenommen
<i>bijektiv</i>	jeder Wert in $B$ wird genau einmal angenommen

Ist  $f : A \rightarrow B$  bijektiv, so stellt  $f$  eine vollständige Paarbildung zwischen den Elementen der Mengen  $A$  und  $B$  her. Jedem Element von  $A$  entspricht durch den Vermittler oder Paarbilder  $f$  genau ein Element von  $B$  und umgekehrt. Damit sind die Mengen  $A$  und  $B$  anschaulich gleich groß. Diese Anschauung werden wir im nächsten Abschnitt präzisieren.

Ist  $f : A \rightarrow B$ , so ist  $f : A \rightarrow \text{Bild}(f)$  surjektiv.

Ist  $f : A \rightarrow B$  injektiv, so ist  $f : A \rightarrow \text{Bild}(f)$  bijektiv.

Sind  $f : A \rightarrow B$  und  $g : B \rightarrow C$  injektiv (surjektiv, bijektiv), so auch  $g \circ f : A \rightarrow C$ .

### Beispiele

- (1) Für jede Menge  $A$  ist  $\text{id}_A : A \rightarrow A$  bijektiv.
- (2) Seien  $G = \{2n \mid n \in \mathbb{N}\}$  und  $U = \mathbb{N} - G$ . Dann sind bijektiv:  
 $f : \mathbb{N} \rightarrow G$  mit  $f(n) = 2n$ ,  $g : \mathbb{N} \rightarrow U$  mit  $g(n) = 2n + 1$ ,  
 $h : G \rightarrow U$  mit  $h(n) = n + 1$ ,  $k : U \rightarrow G$  mit  $k(n) = n - 1$ .
- (3)  $\cos : \mathbb{R} \rightarrow [-1, 1]$  ist surjektiv, aber nicht injektiv,  
 $\cos|_{[0, \pi]} : [0, \pi] \rightarrow \mathbb{R}$  ist injektiv, aber nicht surjektiv,  
 $\cos|_{[0, \pi]} : [0, \pi] \rightarrow [-1, 1]$  ist bijektiv.

### Umkehrfunktionen

Die Bildung der Umkehrfunktion bedeutet in der

*Tabelleninterpretation:* Vertauschung der Spalten

*Zuordnungsinterpretation:* Ändern der Abbildungsrichtung

*kartesische Interpretation:* Spiegelung an der Winkelhalbierenden

*Pfeilinterpretation:* Umkehrung aller Pfeile des Feldes

*Dienstboten-Interpretation:* Bote soll  $a$ , gegeben  $f(a)$ , wieder zurückbringen.

Damit durch den Spaltentausch eine Funktion entsteht, ist es notwendig (und hinreichend), dass in der ursprünglichen rechten Spalte kein  $y$  mehrfach vorkommt.

Wichtige Eigenschaften sind:

Ist  $f : A \rightarrow B$  injektiv, so ist  $f^{-1} : \text{Bild}(f) \rightarrow A$  bijektiv.

Ist  $f : A \rightarrow B$  bijektiv, so ist auch  $f^{-1} : B \rightarrow A$  bijektiv.

$f^{-1} \circ f = \text{id}_A$ ,  $f \circ f^{-1} = \text{id}_B$ , wobei  $A = \text{Def}(f)$ ,  $B = \text{Bild}(f)$ .

### Beispiele

- (1) Sei  $f : \mathbb{R}^* \rightarrow \mathbb{R}$  mit  $f(x) = 1/x$  für alle  $x \in \mathbb{R}^*$ . Dann ist  $f$  injektiv. Für alle  $x, y \in \mathbb{R}^*$  gilt  $y = 1/x$  genau dann, wenn  $x = 1/y$ . Damit ist  $f^{-1} = f$ , was man kartesisch durch Spiegelung an der Winkelhalbierenden schön einsehen kann.
- (2) Die Kosinusfunktion muss man vor einer Umkehrung geeignet einschränken. Mit der injektiven Funktion  $\cos_0 = \cos|_{[0, \pi]}$  kann man den Arkuskosinus  $\arccos : [-1, 1] \rightarrow \mathbb{R}$  durch  $\arccos = \cos_0^{-1}$  definieren. Es gilt  $\cos(\arccos(x)) = x$  für alle  $x \in [-1, 1]$ , aber  $\arccos(\cos(x)) \neq x$  für  $x \notin [0, \pi]$ .

## 1.10 Mächtigkeitsvergleiche

### Definition (Mächtigkeitsbegriffe)

#### Mächtigkeitsvergleich

Für Mengen  $A, B$  definieren wir:

$|A| \leq |B|$ , falls es gibt eine Injektion  $f: A \rightarrow B$ ,

$|A| = |B|$ , falls es gibt eine Bijektion  $f: A \rightarrow B$ ,

$|A| < |B|$ , falls  $|A| \leq |B|$  und  $|A| \neq |B|$ .

Gilt  $|A| \leq |B|$  ( $|A| < |B|$ ), so sagen wir, dass die *Mächtigkeit* von  $A$  *kleinergleich* (*kleiner*) der *Mächtigkeit* von  $B$  ist. Gilt  $|A| = |B|$ , so sagen wir, dass  $A$  und  $B$  *gleichmächtig* sind oder die *gleiche Mächtigkeit* besitzen.

#### Endlichkeit, Unendlichkeit, Abzählbarkeit, Überabzählbarkeit

Eine Menge  $A$  heißt

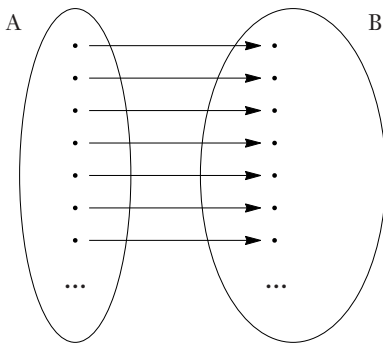
*endlich*, falls  $|A| < |\mathbb{N}|$ ,

*abzählbar*, falls  $|A| \leq |\mathbb{N}|$ ,

*unendlich*, falls  $|\mathbb{N}| \leq |A|$ ,

*abzählbar unendlich*, falls  $|A| = |\mathbb{N}|$ ,

*überabzählbar*, falls  $|\mathbb{N}| < |A|$ .



Die Mächtigkeit von  $A$  ist kleinergleich der Mächtigkeit von  $B$ , falls es eine Injektion von  $A$  nach  $B$  gibt. Anschaulich passt dann die Menge  $A$  in die Menge  $B$  hinein. Aus dem Mächtigkeitsvergleich  $|A| \leq |B|$  folgt alles Weitere, denn die Gleichmächtigkeit  $|A| = |B|$  ist äquivalent zu  $|A| \leq |B|$  und  $|B| \leq |A|$  (Satz von Cantor-Bernstein). Insgesamt ergibt sich aus den Abbildungseigenschaften „injektiv, surjektiv, bijektiv“ eine Größenlehre für Mengen ohne Verwendung von Zahlen.

Die Abbildungseigenschaften „injektiv, surjektiv, bijektiv“ sind geeignet, zwei Mengen ihrer Größe nach zu vergleichen, ohne ihre Elemente zählen zu müssen. Anschaulich bedeutet  $|A| \leq |B|$ , dass jedes Element von  $A$  einen Partner in  $B$  findet, wenn ein geschickter Vermittler  $f$  die Zuordnung übernimmt. Die Gleichmächtigkeit  $|A| = |B|$  bedeutet, dass die Elemente von  $A$  mit den Elementen von  $B$  vollständig gepaart werden. Und  $|A| < |B|$  bedeutet, dass eine Partnerfindung wie in  $|A| \leq |B|$  möglich ist, aber jede Partnerfindung einsame  $b$  in  $B$  zurücklässt.



Die Endlichkeit einer Menge  $A$  ist äquivalent dazu, dass  $A$  in der Form  $A = \{a_1, \dots, a_n\}$  mit paarweise verschiedenen  $a_k$  geschrieben werden kann (wobei  $n = 0$  für  $A = \emptyset$  zugelassen ist). Man schreibt dann auch  $|A| = n$ . Für endliche Mengen gilt:

### Dirichletsches Schubfach- oder Taubenschlagprinzip

Sind  $A, B$  endlich mit  $|A| = |B|$  und ist  $f : A \rightarrow B$ , so sind äquivalent:

(a)  $f : A \rightarrow B$  ist injektiv, (b)  $f : A \rightarrow B$  ist surjektiv, (c)  $f : A \rightarrow B$  ist bijektiv.

Die Bezeichnung „Schubfachprinzip“ lässt sich wie folgt illustrieren:

### Beispiel

Verteilt man  $n$  Kugeln  $a_1, \dots, a_n$  auf  $m < n$  Fächer, so gibt es ein Fach, das mindestens zwei Kugeln enthält. Denn sonst wäre  $f : \{a_1, \dots, a_n\} \rightarrow \{1, \dots, m\}$  mit  $f(a_k) = \text{„die Fachnummer von } a_k\text{“}$  injektiv, aber wegen  $m < n$  nicht surjektiv.

Diese Ergebnisse sind anschaulich klar. Große Überraschungen bergen dagegen die unendlichen Mengen. Man kann zeigen:

$$|\mathbb{N}| = |\mathbb{Z}| = |\mathbb{Q}| = |\mathbb{N}^2| = |\mathbb{N}^3| = \left| \bigcup_{n \in \mathbb{N}} \mathbb{N}^n \right|,$$

wobei die letzte Menge die Menge aller endlichen Tupel natürlicher Zahlen ist und damit über eine geeignete Zahlenkodierung von Buchstaben und Satzzeichen jedes Buch (aufgefasst als Zeichenfolge) als Element enthält. Eine Universalbibliothek, die alle Bücher enthält, ist abzählbar. Dagegen zeigt man in der Analysis, dass die reellen Zahlen überabzählbar sind:  $|\mathbb{N}| < |\mathbb{R}|$ . Dieses Ergebnis kann man auch so formulieren:

*Jede Folge  $x_0, x_1, x_2, \dots, x_n, \dots$  reeller Zahlen lässt eine reelle Zahl aus.*

Ebenso gilt  $|\mathbb{R}| < |\mathbb{R}^{\mathbb{R}}|$ , sodass gilt:

*Jede Familie  $(f_x)_{x \in \mathbb{R}}$  reeller Funktionen  $f_x : \mathbb{R} \rightarrow \mathbb{R}$  lässt eine reelle Funktion aus.*

Es gibt also auch im Unendlichen verschiedene Mächtigkeitsstufen. Wichtige allgemeine Ergebnisse der Mächtigkeitsstufen, die für alle Mengen  $A, B$  gelten, sind:

$$|A| < |\mathcal{P}(A)| \quad (\text{Satz von Cantor})$$

$$|A| \leq |B| \text{ und } |B| \leq |A| \text{ impliziert } |A| = |B| \quad (\text{Satz von Cantor-Bernstein})$$

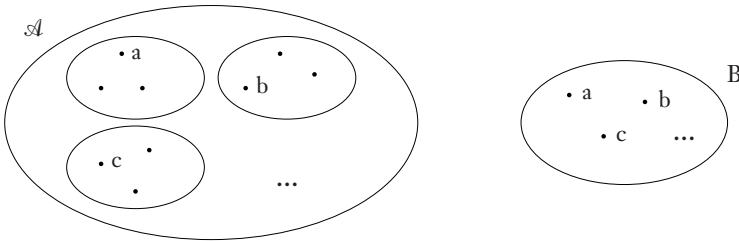
$$|A| \leq |B| \text{ oder } |B| \leq |A| \quad (\text{Vergleichbarkeitssatz von Cantor-Zermelo})$$

Während die beiden ersten Sätze trickreich, aber elementar bewiesen werden können, müssen zum Beweis des Vergleichbarkeitssatzes schwere Geschütze aufgeföhren werden. Der Satz ist äquivalent zum Auswahlaxiom, das wir nun besprechen werden.

## 1.11 Das Auswahlaxiom

**Axiom** (*Auswahlaxiom von Ernst Zermelo*)

Ist  $\mathcal{A}$  eine Menge nichtleerer und paarweise disjunkter Mengen, so gibt es eine Menge  $B$ , die mit jedem  $A \in \mathcal{A}$  genau ein Element gemeinsam hat.



Eine Auswahlmenge  $B$  „pickt“ aus jeder Menge des Systems  $\mathcal{A}$  genau ein Element heraus. Ist  $\mathcal{A}$  endlich, so ist die Existenz von  $B$  induktiv beweisbar. Dagegen muss man selbst für Systeme  $\mathcal{A}$ , deren Elemente alle zweielementig sind (Systeme von „Sockenpaaren“), im Allgemeinen das Auswahlaxiom heranziehen, um eine Auswahlmenge zu erhalten.

Zwei Zeilen. Unzählige Diskussionen. Und dies seit der Einführung 1908. Das Auswahlaxiom ist das umstrittenste Prinzip der Mathematik. Die Meinungen umfassen:

- (a) Offensichtlich wahr/richtig/korrekt.
- (b) Nicht akzeptabel.
- (c) Offensichtlich richtig, aber in seinen Konsequenzen nicht akzeptabel.
- (d) Einfach notwendig für viele mathematische Unternehmungen.
- (e) Richtig, aber eine Schwäche der Präsentation der mengentheoretischen Axiome.

Eine Menge  $B$  wie im Auswahlaxiom nennt man auch eine *Auswahlmenge* für das Mengensystem  $\mathcal{A}$ . Das System wird als eine Zerlegung von  $A = \bigcup \mathcal{A}$  vorausgesetzt (vgl. 0.4). Aus jedem „Land“  $A \in \mathcal{A}$  wird in einer Auswahlmenge genau ein „Bewohner“ ausgewählt und alle ausgewählten Bewohner bilden die Auswahlmenge  $B$ . Damit gilt:

### Existenz von Repräsentantensystemen

Jede Äquivalenzrelation besitzt ein vollständiges Repräsentantensystem.

Diese Folgerung ist sogar äquivalent zum Auswahlaxiom. Gleiches gilt für:

### Existenz von Auswahlfunktionen

Ist  $(A_i \mid i \in I)$  eine Familie nichtleerer Mengen, so ist  $\prod_{i \in I} A_i \neq \emptyset$ , d. h., es gibt eine Funktion  $f$  auf  $I$  mit  $f(i) \in A_i$  für alle  $i \in I$ .

Diese Aussage lässt sich als unendliche Verallgemeinerung des Satzes auffassen, dass ein kartesisches Produkt  $A_1 \times \dots \times A_n$  nichtleerer Mengen nichtleer ist.

Das Auswahlaxiom eröffnet eine neue Möglichkeit zur Konstruktion von Funktionen (vgl. Abschnitt 1.5):

**Möglichkeit 4: „ein (beliebiges)  $b$  mit ...“**

Sei  $\mathcal{E}(a, b)$  eine Eigenschaft und  $A$  eine Menge mit

(++) Für alle  $a \in A$  gibt es ein  $b$  mit  $\mathcal{E}(a, b)$ .

Dann gibt es eine Funktion  $f$  auf  $A$ , sodass  $\mathcal{E}(a, f(a))$  für alle  $a \in A$  erfüllt ist.

Eine solche Funktion  $f$  können wir in der Praxis einfach definieren durch

$f(a) =$  „ein  $b$  mit  $\mathcal{E}(a, b)$ “ für alle  $a \in A$ .

Durch das „ein“ wird ein abstrakter Auswahlakt angedeutet. Dass zu jedem  $a$  mindestens ein  $b$  mit der Eigenschaft  $\mathcal{E}(a, b)$  existiert, muss vorab bewiesen werden.

**Beispiele**

(1) Ist  $\sim$  eine Äquivalenz auf  $A$ , so können wir  $f$  auf  $A/\sim$  definieren durch:

$f(B) =$  „ein  $b \in B$ “ für alle  $B \in A/\sim$ .

Welche  $b$  aus den Äquivalenzklassen gewählt werden, bleibt unbestimmt. Wir wissen aber, dass  $\text{Bild}(f)$  ein Repräsentantensystem von  $\sim$  ist.

(2) Sei  $f : A \rightarrow B$  surjektiv. Dann definieren wir  $g : B \rightarrow A$  durch

$g(b) =$  „ein  $a \in A$  mit  $f(a) = b$ “.

Welche Urbilder gewählt werden, bleibt unbestimmt. Wir wissen aber, dass  $g \circ f = \text{id}_A$ . (Die Aussage: „Ist  $f : A \rightarrow B$  surjektiv, so existiert ein  $g : B \rightarrow A$  mit  $g \circ f = \text{id}_A$ .“ ist erneut äquivalent zum Auswahlaxiom.)

(3) Ist  $(A_n \mid n \in \mathbb{N})$  eine Familie abzählbar unendlicher Mengen, so sei

$(a_{n,k} \mid k \in \mathbb{N}) =$  „eine Bijektion  $f : \mathbb{N} \rightarrow A_n$ “ für alle  $n \in \mathbb{N}$ .

Damit kann die Diagonalaufzählung

$a_{0,0}, a_{0,1}, a_{1,0}, a_{0,2}, a_{1,1}, a_{2,0}, a_{0,3}, a_{1,2}, a_{2,1}, a_{3,0}, \dots$

von  $\bigcup_{n \in \mathbb{N}} A_n$  definiert werden. Dies zeigt (mit Hilfe des Auswahlaxioms), dass eine abzählbare Vereinigung abzählbar unendlicher Mengen abzählbar ist.

Im nächsten Abschnitt werden wir das zum Auswahlaxiom äquivalente Zornsche Lemma vorstellen. Im dritten Kapitel können wir mit Hilfe des Zornschen Lemmas zeigen, dass jeder Vektorraum eine Basis besitzt. Diese fundamentale Aussage der Linearen Algebra ist ohne Auswahlaxiom nicht beweisbar (sie ist überraschenderweise ebenfalls äquivalent zum Auswahlaxiom). Manche Mathematiker bemerken scherzhaft:

„Ob jeder Vektorraum eine Basis besitzt, hängt vom Dozenten ab.“

## 1.12 Das Zornsche Lemma

### Satz (Zornsches Lemma)

Sei  $\leq$  eine partielle Ordnung auf  $A$ . Es gelte:

- (#) Ist  $B \subseteq A$  linear geordnet durch  $\leq$ , so existiert eine obere Schranke von  $B$ .  
(Kettenbedingung)

Dann existiert ein maximales Element der Ordnung. Genauer existiert für alle  $a_0 \in A$  ein maximales Element  $b$  in  $A$  mit  $a_0 \leq b$ .

Eine Teilmenge  $B$  von  $A$  wird durch  $\leq$  linear geordnet, falls je zwei Elemente von  $B$  vergleichbar sind:

Für alle  $a, b \in B$  gilt  $a \leq b$  oder  $b \leq a$ .

Man nennt eine solche Menge  $B$  eine *Kette* in  $A$ . Das Zornsche Lemma besagt:

*Hat jede Kette in  $A$  eine obere Schranke, so existiert ein maximales Element in  $A$ .*

Es wird nicht behauptet, dass  $\max(A)$  existiert. Behauptet wird lediglich: Es gibt mindestens ein  $a \in A$ , sodass kein  $b \in A$  größer als  $a$  ist.

Die leere Menge gilt als Kette und hat nach (#) eine obere Schranke  $a$  (jedes  $a \in A$  ist eine solche). Damit ist eine Ordnung, die die Kettenbedingung erfüllt, nichtleer.

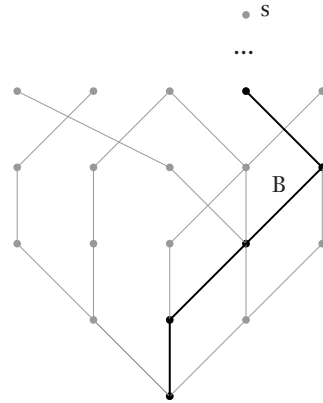


Illustration der Kettenbedingung (#): Die Kette  $B$  der partiellen Ordnung hat die obere Schranke  $s$ .

### Beispiele

- (1) Sei  $A = \mathcal{P}(\{1, \dots, 100\})$  geordnet durch  $\subseteq$ . Dann ist

$$B = \{\emptyset, \{1\}, \{1, 4\}, \{1, 4, 10, 12\}, \{1, 4, 10, 12, 44, 45, 50\}\}$$

eine Kette. Dies gilt nicht für  $C = \{\{1\}, \{1, 4\}, \{1, 2, 10\}\}$ .

- (2) Die durch  $\subseteq$  geordnete Menge  $A = \mathcal{P}(\mathbb{N}) - \{\mathbb{N}\}$  hat keine maximalen Elemente. Die Ordnung erfüllt die Kettenbedingung nicht. Denn

$$B = \{\{0\}, \{0, 1\}, \{0, 1, 2\}, \dots\} = \{\{0, \dots, n\} \mid n \in \mathbb{N}\}$$

ist eine Kette, besitzt aber keine obere Schranke in  $A$ .

- (3) Auf  $\mathbb{N}$  sei  $\leq = \{(n, n + 2k) \mid n, k \in \mathbb{N}, n \neq 0\} \cup \{(2k, 0) \mid k \in \mathbb{N}\}$ , sodass

$$1 < 3 < 5 < \dots, \quad 2 < 4 < 6 < \dots < 0.$$

Dann ist  $0$  ein maximales Element von  $\mathbb{N}$ . Die Kettenbedingung ist verletzt, da die Kette  $\{1, 3, 5, \dots\}$  keine obere Schranke besitzt.

### Anschaulicher Beweis des Zornschen Lemmas

Wir setzen uns auf ein beliebiges  $a_0 \in A$  und blicken von dort aus nach oben, d. h., wir betrachten  $X_0 = \{a \in A \mid a > a_0\}$ . Sehen wir nichts ( $X_0 = \emptyset$ ), so ist  $a_0$  maximal in  $A$  und wir sind fertig. Andernfalls wählen wir ein beliebiges  $a_1 \in X_0$ . Es gilt dann

$$a_0 < a_1.$$

Nun klettern wir nach  $a_1$  hoch und blicken nach oben. Ist  $X_1 = \{a \in A \mid a > a_1\}$  leer, so ist  $a_1$  maximal und wir sind fertig. Andernfalls wählen wir ein  $a_2 \in X_1$ . Dann gilt

$$a_0 < a_1 < a_2.$$

So in der Ordnung hochklettern und finden wir entweder ein maximales Element  $a_n$  wie gewünscht, oder aber wir erhalten eine unendliche Kette

$$a_0 < a_1 < \dots < a_n < \dots$$

Nun hilft uns die Kettenbedingung weiter. Denn  $B = \{a_n \mid n \in \mathbb{N}\}$  ist linear geordnet. Nach (#) existiert also eine obere Schranke von  $B$ . Wir wählen eine derartige Schranke, die wir  $a_\omega$  nennen (wobei  $\omega$  an  $\infty$  erinnert). Wir klettern nun nach  $a_\omega$  und blicken von dort erneut nach oben. Ist  $X_\omega = \{a \in A \mid a > a_\omega\}$  leer, so ist  $a_\omega$  maximal. Andernfalls wählen wir ein beliebiges  $a_{\omega+1} \in X_\omega$  und wiederholen das Verfahren des Hochkletterns, wobei wir an „Limesstellen“ des Hochkletterns die Kettenbedingung (#) zu Hilfe rufen. Irgendwann (das kann sehr, sehr, sehr lange dauern) finden wir schließlich ein maximales Element  $a_\alpha$ , denn sonst könnten wir wieder weiterklettern und  $a_{\alpha+1}$  bilden. Das Element  $a_\alpha > a_0$  ist wie gewünscht.

Zunächst sieht alles wie eine übliche Rekursion aus. Aber wir sind nach unendlich vielen Schritten noch nicht unbedingt fertig. In der  $\subseteq$ -Ordnung auf  $\mathcal{P}(\mathbb{N})$  können wir über

$$a_0 = \{0\}, \quad a_1 = \{0, 2\}, \quad a_2 = \{0, 2, 4\}, \quad \dots$$

hochklettern, haben dann aber das maximale Element  $\mathbb{N}$  noch nicht gefunden. Erst

$$a_\omega = \{0, 2, 4, \dots\}, \quad a_{\omega+1} = \{0, 2, 4, \dots, 1\}, \quad a_{\omega+2} = \{0, 2, 4, \dots, 1, 3\}, \quad \dots$$

liefert  $a_{\omega+\omega} = \mathbb{N}$ .

Der anschauliche Beweis kann unter Wahrung der Idee streng geführt werden, wenn man statt  $\mathbb{N}$  die transfiniten Zahlen

$$0, 1, 2, \dots, n, \dots, \omega, \omega + 1, \omega + 2, \dots, \omega + \omega, \omega + \omega + 1, \dots, \dots, \dots, \dots, \dots, \dots, \dots$$

verwendet, mit denen jede noch so große partielle Ordnung durchwandert werden kann. Da diese Zahlen schwierig sind, wird das Zornsche Lemma oft entweder gar nicht bewiesen, oder es wird ein unanschaulicher Beweis geführt, der die transfiniten Zahlen vermeidet. Ist das Zornsche Lemma als eine Art Axiom aber einmal da, so kann es als Werkzeug verwendet werden, um in ähnlichen Situationen ein transfinite Hochklettern zu vermeiden. Zu diesem Zweck ist es von den Algebraikern ins Leben gerufen worden: Genuss des Transfiniten ohne transfiniten Zahlen.

Obiger Beweis verwendet „wir wählen ...“ und damit das Auswahlaxiom. Man kann zeigen, dass das Zornsche Lemma äquivalent zum Auswahlaxiom ist.



## 2. Kapitel

---

# Algebraische Strukturen

---

## 2.1 Halbgruppen

### Definition (Halbgruppe, Assoziativgesetz)

Seien  $H$  eine Menge und  $\circ : H^2 \rightarrow H$  eine (zweistellige) Operation auf  $H$ . Dann heißt das Paar  $(H, \circ)$  eine *Halbgruppe*, falls gilt:

$$(a \circ b) \circ c = a \circ (b \circ c) \quad \text{für alle } a, b, c \in H. \quad (\text{Assoziativgesetz})$$

Eine Halbgruppe ist also eine mit einer assoziativen Operation ausgestattete Menge. Wir schreiben  $a \circ b$  statt  $\circ(a, b)$ . Im Begriff „Operation auf  $H$ “ ist enthalten, dass der Wertebereich von  $\circ$  eine Teilmenge von  $H$  ist (vgl. 1.8). Es gilt also

$$a \circ b \in H \quad \text{für alle } a, b \in H.$$

### Andere Schreibweisen

Das Zeichen  $\circ$  steht für eine beliebige Operation und hat oft nichts mit der Komposition von Funktionen zu tun. Ist  $H$  eine Menge von Funktionen, so ist jedoch  $\circ$  die Komposition von Funktionen, wenn nichts anderes gesagt wird. Statt  $\circ$  können wir ein beliebiges anderes Zeichen verwenden. Typische Operationszeichen sind  $*$ ,  $\cdot$ ,  $+$ . Das Gleiche gilt für die zugrunde gelegte Menge. Man kann schreiben:

„Sei  $(M, +)$  eine Halbgruppe.“

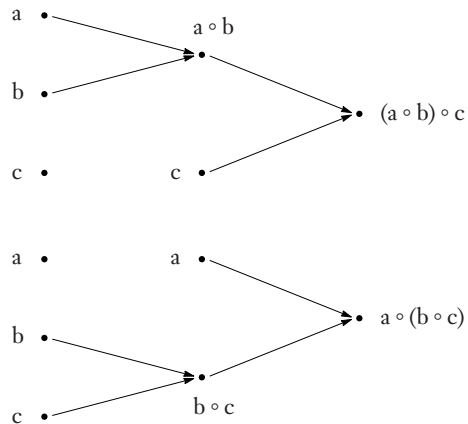
Dies bedeutet, dass  $+: M^2 \rightarrow M$  und dass  $a + (b + c) = (a + b) + c$  für alle  $a, b, c \in M$ . Auch hier hat die Operation  $+$  in vielen Fällen nichts mit der Addition auf einer Zahlenmenge wie  $\mathbb{N}$  oder  $\mathbb{R}$  zu tun. Für das Zeichen  $+$  sind bestimmte Notationen reserviert, die wir im Folgenden kennenlernen werden.

### Vereinfachung der Notation

Anstelle von  $(H, \circ)$  schreibt und sagt man oft auch nur  $H$ . Eine Operation ist dann stillschweigend mit dabei. So sagt man zum Beispiel: „Ist  $H$  eine Halbgruppe, so gilt  $a \circ (b \circ c) = (a \circ b) \circ c$  für alle  $a$  und  $b$  in  $H$ .“ Diese bewusste Verwechslung einer Struktur  $(H, \circ)$  mit ihrer Trägermenge  $H$  wird in vielen Fällen durchgeführt. Sie ist in der Regel ungefährlich und erleichtert die Sprechweise.

### Weglassen des Operationszeichens

Ist das Operationszeichen von  $+$  verschieden, so lässt man es oft weg. So schreibt man zum Beispiel  $ab$  statt  $a \cdot b$  und  $a(bc)$  statt  $a * (b * c)$  usw.



Zweistufige „Verarbeitung“ von drei Objekten  $a, b, c$ .  
Bei einer assoziativen Operation ist  $(a \circ b) \circ c = a \circ (b \circ c)$ .



**Beispiele**

- (1)  $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$  bilden mit der üblichen Addition Halbgruppen. Das Gleiche gilt für die Multiplikation.
- (2) Ist  $G = \{2n \mid n \in \mathbb{N}\}$  die Menge der geraden und  $U = \mathbb{N} - G$  die Menge der ungeraden Zahlen, so sind  $(G, +)$ ,  $(G, \cdot)$  und  $(U, \cdot)$  mit der üblichen Addition und Multiplikation Halbgruppen. Dagegen ist  $(U, +)$  keine Halbgruppe, da + wegen  $1 + 1 \notin U$  keine Operation auf  $U$  ist.
- (3) Ist  $A$  eine Menge und  $H = \{f \mid f: A \rightarrow A\}$ , so ist  $(H, \circ)$  eine Halbgruppe. Gleiches gilt für  $H' = \{f \mid f: A \rightarrow A \text{ ist injektiv}\}$ .
- (4) Setzen wir  $a \circ b = |b - a|$  für alle  $a, b \in \mathbb{Z}$ , so ist  $(\mathbb{Z}, \circ)$  keine Halbgruppe, da zum Beispiel  $(1 \circ 2) \circ 3 = 1 \circ 3 = 2$ , aber  $1 \circ (2 \circ 3) = 1 \circ 1 = 0$ .
- (5) Sind  $H_1$  und  $H_2$  Halbgruppen und ist  $H = H_1 \times H_2$ , so setzen wir
 
$$(a, b) \circ (c, d) = (ac, bd) \quad \text{für alle } (a, b), (c, d) \in H.$$
 Dann ist  $(H, \circ)$  eine Halbgruppe. Sie heißt das *Produkt* von  $H_1$  und  $H_2$ .

Das Assoziativgesetz ist ein unverzichtbarer Begleiter bei den allermeisten algebraischen Unternehmungen. Seine Wirkung können wir so zusammenfassen:

*Wir dürfen Klammern weglassen.*

Da nämlich  $a \circ (b \circ c) = (a \circ b) \circ c$  für alle Elemente einer Halbgruppe gilt, können wir kurz  $a \circ b \circ c$  schreiben. Allgemein gilt  $(s \circ t) \circ u = s \circ (t \circ u)$  für alle Terme  $s, t, u$ , sodass wir einfach  $s \circ t \circ u$  oder  $stu$  schreiben können. Ein Term ist dabei ein aus Variablen, dem Operationszeichen und Klammern aufgebauter Ausdruck wie  $(a \circ a) \circ (b \circ (c \circ a))$ .

**Beispiel**

In Halbgruppen ist  $a \circ b \circ c \circ d$  unzweideutig, da

$$((a \circ b) \circ c) \circ d = (a \circ b) \circ (c \circ d) = a \circ (b \circ (c \circ d)) = a \circ ((b \circ c) \circ d) = \dots$$

Wir führen ein:

**Potenzen und Produkte**

Ist  $H$  eine Halbgruppe, so definieren wir für alle  $a, a_1, \dots, a_n \in H$  und  $n \geq 1$  rekursiv:

$$a^1 = a, \quad a^{n+1} = a^n \circ a, \quad (\text{Potenzen})$$

$$\prod_{1 \leq k \leq 1} a_k = a_1, \quad \prod_{1 \leq k \leq n+1} a_k = \left(\prod_{1 \leq k \leq n} a_k\right) \circ a_{n+1}. \quad (\text{Produkt})$$

Induktiv zeigt man die folgenden Potenzregeln:

$$(a^n)^m = a^{mn}, \quad a^n a^m = a^{n+m} \quad \text{für alle } a \in H \text{ und } n, m \geq 1.$$

**Beispiele**

$$a^2 \circ b^2 = a \circ a \circ b \circ b, \quad (a \circ b)^2 = a \circ b \circ a \circ b.$$

## 2.2 Monoide

### Definition (Monoid, neutrales Element)

Eine Halbgruppe  $(M, \circ)$  heißt *Monoid*, falls gilt:

Es gibt ein  $e \in M$ , sodass für alle  $a \in M$  gilt:  $a \circ e = e \circ a = a$ .

(Existenz eines neutralen Elements)

Ein derartiges  $e$  heißt ein *neutrales Element* des Monoids.

Monoide sind also Halbgruppen, die ein zusätzliches Axiom erfüllen: Die Tafel von  $\circ$  enthält eine triviale Zeile und Spalte. So unscheinbar die Eigenschaft

$$a \circ e = e \circ a = a \quad \text{für alle } a \in A$$

sein mag, so wichtig ist die Existenz eines „nichts verändernden“ oder „neutralen“ Elements für alles Weitere. Eine wichtige Beobachtung ist:

$\circ$	e	a	b	c	...
e	e	a	b	c	...
a	a	$a \circ a$	$a \circ b$	$a \circ c$	...
b	b	$b \circ a$	$b \circ b$	$b \circ c$	...
c	c	$c \circ a$	$c \circ b$	$c \circ c$	...
...	...	...	...	...	...

Ein neutrales Element  $e$  in der Verknüpfungstafel der Operation

### Eindeutigkeit des neutralen Elements

Sind  $e$  und  $e'$  neutrale Elemente eines Monoids  $(M, \circ)$ , so gilt  $e = e'$ .

Sind nämlich  $e$  und  $e'$  neutral, so gilt  $e = e \circ e' = e'$ , wobei wir beim ersten Gleichheitszeichen die Neutralität von  $e'$  verwenden und beim zweiten die Neutralität von  $e$ . Wir können also fortan schreiben:

„Sei  $e$  das neutrale Element des Monoids  $(M, \circ)$ .“

### Zeichenwahl für das neutrale Element

Das Zeichen für das neutrale Element eines Monoids ist prinzipiell beliebig. Für die Operationszeichen  $\circ, *, \cdot, \dots$  wird neben  $e$  oft 1 und für das Operationszeichen  $+$  zumeist 0 verwendet.

In Monoiden können wir die für Halbgruppen erklärte Potenzierung erweitern:

### Der Exponent Null und das leere Produkt

Ist  $M$  ein Monoid mit neutralem Element  $e$ , so setzen wir für alle  $a \in M$ :

$$a^0 = e, \quad \prod_{1 \leq k \leq 0} a_k = e.$$

Die Regeln  $(a^n)^m = a^{nm}$  und  $a^n a^m = a^{n+m}$  gelten nun für alle  $a \in M$  und  $n, m \in \mathbb{N}$ . Es gilt  $\prod_{1 \leq k \leq n} a = a^n$ . Da das leere Produkt  $e$  ist, ist hier auch  $n = 0$  zulässig.

Für Monoide wie  $(\mathbb{N}, \cdot)$  und  $(\mathbb{R}, \cdot)$  mit neutralem Element 1 gilt nach Definition wie gewohnt  $a^0 = 1$  für alle  $a$ , einschließlich  $0^0 = 1$ .

## Beispiele

- (1)  $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$  mit der Addition  $+$  sind Monoide mit neutralem Element 0. Gleiches gilt für die Multiplikation, wobei dann 1 neutral ist.
- (2)  $\mathbb{N}^* = \mathbb{N} - \{0\}$  ist mit der Multiplikation ein Monoid mit neutralem Element 1.
- (3)  $\mathbb{N}^*$  ist mit der Addition eine Halbgruppe, aber kein Monoid.
- (4) Ist  $e$  beliebig, so ist  $\{e\}$  ein Monoid, wenn wir  $e \circ e = e$  definieren.
- (5) Ist  $H$  eine Halbgruppe und  $e \notin H$ , so können wir die Operation auf  $H$  zu einer Operation auf  $M = H \cup \{e\}$  fortsetzen, indem wir definieren:  

$$a \circ e = e \circ a = a \quad \text{für alle } a \in M.$$
 Dann ist  $M$  ein Monoid mit neutralem Element  $e$ .
- (6) Ist  $A$  eine Menge und  $M = \{f \mid f: A \rightarrow A\}$ , so ist  $(M, \circ)$  ein Monoid. Das neutrale Element ist die Identität  $\text{id}_A: A \rightarrow A$ .
- (7) Seien  $M_1, M_2$  Monoide mit den neutralen Elementen  $e_1$  bzw.  $e_2$ . Dann ist das Produkt  $M = M_1 \times M_2$  der Halbgruppen  $M_1$  und  $M_2$  ein Monoid mit neutralem Element  $(e_1, e_2)$  (vgl. 2.1).
- (8) Sei  $(M, \cdot)$  ein Monoid mit neutralem Element  $e$ . Dann definieren wir eine Operation  $\circ$  auf der Potenzmenge  $\mathcal{P}(M) = \{A \mid A \subseteq M\}$  von  $M$  durch  

$$A \circ B = \{a \cdot b \mid a \in A \text{ und } b \in B\} \quad \text{für alle } A, B \subseteq M.$$
 Dann ist  $(\mathcal{P}(M), \circ)$  ein Monoid mit neutralem Element  $\{e\}$ .
- (9) Für jede Menge  $M$  ist  $(\mathcal{P}(M), \cup)$  ein Monoid mit neutralem Element  $\emptyset$  und  $(\mathcal{P}(M), \cap)$  ein Monoid mit neutralem Element  $M$ .

Das folgende Beispiel zeigt, dass es nicht genügt, lediglich die Existenz eines einseitig neutralen Elements  $e$  mit

„ $a \circ e = a$  für alle  $a \in M$ “ oder „ $e \circ a = a$  für alle  $a \in M$ “

in der Definition eines Monoids zu fordern.

## Beispiel

Auf der Menge  $H = \{0, 1\}$  definieren wir:

$$0 \circ 0 = 1 \circ 0 = 0, \quad 0 \circ 1 = 1 \circ 1 = 1$$

nach der Devise „der zweite Faktor setzt sich durch“. Dann ist  $(H, \circ)$  eine Halbgruppe, aber kein Monoid, denn weder 0 noch 1 sind neutral. Für alle  $a \in H$  gilt

$0 \circ a = a$  und  $1 \circ a = a$ , sodass 0 und 1 sog. linksneutrale Elemente sind. Analoges gilt für die Operation „der erste Faktor setzt sich durch“.

$\circ$	0	1
0	0	1
1	0	1

## 2.3 Gruppen

### Definition (Gruppe, inverse Elemente)

Ein Monoid  $(G, \circ)$  heißt eine *Gruppe*, falls für das neutrale Element  $e$  von  $G$  gilt:

Für alle  $a \in G$  existiert ein  $b \in G$  mit  $a \circ b = b \circ a = e$ . (Existenz inverser Elemente)

Gilt  $a \circ b = b \circ a = e$ , so heißt  $b$  *invers* zu  $a$ .

Gruppen sind also Monoide, die ein weiteres Axiom erfüllen: Jede Zeile und Spalte der Operationstafel enthält spiegelsymmetrisch einen Eintrag  $e$ . Wie für das neutrale Element eines Monoids gilt:

### Eindeutigkeit des Inversen

Sind  $b$  und  $b'$  invers zu  $a$  in der Gruppe  $G$ , so gilt  $b = b'$ .

Denn sind  $b$  und  $b'$  invers zu  $a$ , so ist

$$b = e \circ b = b' \circ a \circ b = b' \circ e = b'.$$

Damit können wir definieren:

### Die Inversennotation $a^{-1}$

In einer Gruppe bezeichnen wir das eindeutig bestimmte Inverse von  $a$  mit  $a^{-1}$ .

Um  $b = a^{-1}$  für Elemente  $a, b$  einer Gruppe zu zeigen, genügt der Nachweis von  $a \circ b = e$ . Denn dann ist  $b = e \circ b = a^{-1} \circ a \circ b = a^{-1} \circ e = a^{-1}$ . Ebenso folgt aus  $b \circ a = e$ , dass  $b = a^{-1}$ .

In Gruppen können wir die Potenzierung erneut erweitern:

### Negative Exponenten

Ist  $G$  eine Gruppe, so setzen wir  $a^{-n} = (a^{-1})^n$  für alle  $a \in G$  und  $n \in \mathbb{N}$ .

Die Potenzregeln aus 2.1 und 2.2 gelten nun für alle ganzen Zahlen  $n, m$ . Allgemeinere Exponentiationen  $a^q$  mit  $q \in \mathbb{Q}$  oder  $a^x$  mit  $x \in \mathbb{R}$  sind nur unter zusätzlichen Voraussetzungen möglich und fallen in das Aufgabengebiet der Analysis.

### Beispiele

- (1) Jede Menge  $G = \{a\}$  mit  $a \circ a = a$  ist eine Gruppe. Es gilt  $a = a^{-1} = e$ .
- (2)  $\mathbb{Z}$ ,  $\mathbb{Q}$ ,  $\mathbb{R}$ ,  $\mathbb{C}$  sind mit der Addition Gruppen. Ebenso sind  $\mathbb{Q}^*$ ,  $\mathbb{R}^*$ ,  $\mathbb{C}^*$  Gruppen unter der Multiplikation, wobei der Stern die Entfernung der Null bedeutet.
- (3) Ist  $(G, \circ)$  eine Gruppe, so auch  $(G, *)$ , wobei  $a * b = b \circ a$  für alle  $a, b \in G$ .

$\circ$	e	a	b	c	...
e	e	a	b	c	...
a	a		e		...
b	b	e			...
c	c			e	...
...	...	...	...	...	...

Nach Definition einer Gruppe taucht in jeder Zeile und jeder Spalte mindestens einmal das neutrale Element  $e$  des Monoids in spiegelsymmetrischer Weise auf. Stärker gilt, dass  $e$  jeweils genau einmal auftaucht und dass auf die Forderung des spiegelsymmetrischen Auftretens verzichtet werden kann (vgl. den vereinfachten Nachweis der Gruppenaxiome unten).

- (4) Für alle  $n \geq 1$  bildet  $\mathbb{R}^n$  mit der komponenten- oder punktweisen Addition  $(x_1, \dots, x_n) + (y_1, \dots, y_n) = (x_1 + y_1, \dots, x_n + y_n)$  eine Gruppe. Das Element  $(0, \dots, 0)$  ist neutral und  $(-x_1, \dots, -x_n)$  ist invers zu  $(x_1, \dots, x_n)$ .
- (5) Für alle  $m \geq 1$  bildet  $\mathbb{Z}_m = \mathbb{Z}/\equiv_m$  unter der Addition  $[a] + [b] = [a + b]$  von Restklassen eine Gruppe.  $\mathbb{Z}_m^* = \mathbb{Z}_m - \{0\}$  bildet unter der Multiplikation  $[a] \cdot [b] = [ab]$  genau dann eine Gruppe, wenn  $m$  eine Primzahl ist.
- (6) Sind  $G_1$  und  $G_2$  Gruppen, so ist auch das Produkt  $G = G_1 \times G_2$  eine Gruppe. Für alle  $(a, b) \in G$  gilt  $(a, b)^{-1} = (a^{-1}, b^{-1})$ .

Eine eigene Definition verdient die Erweiterung des Permutationsbegriffs aus 1.5:

**Definition** (*symmetrische Gruppe, Permutationen,  $S_A, S_n$* )

Seien  $A$  eine Menge und  $S_A = \{f \mid f: A \rightarrow A \text{ ist bijektiv}\}$ . Dann heißt  $(S_A, \circ)$  die *symmetrische Gruppe* oder *Permutationsgruppe* von  $A$ . Jedes Element von  $S_A$  heißt eine *Permutation* auf  $A$ . Weiter schreiben wir  $S_n$  statt  $S_{\{1, \dots, n\}}$ .

Das neutrale Element von  $S_A$  ist  $\text{id}_A$ . Für alle  $f \in S_A$  ist die Umkehrfunktion  $f^{-1}$  invers zu  $f$ , sodass die Lesarten von  $f^{-1}$  als Umkehrfunktion oder Inverses übereinstimmen.

Jedes Monoid  $(M, \circ)$  gibt Anlass zur Definition einer Gruppe: Wir nennen ein  $a \in M$  *invertierbar*, falls es ein  $b \in M$  gibt mit  $a \circ b = b \circ a = e$ . Dann ist

$$M^\times = \{a \in M \mid \text{es gibt ein } b \in M \text{ mit } a \circ b = b \circ a = e\} \quad (\text{Gruppe der invertierbaren Elemente})$$

mit der von  $G$  ererbten Operation eine Gruppe.

**Beispiel**

Sei  $A$  eine Menge. Dann gilt für das Monoid  $M = \{f \mid f: A \rightarrow A\}$  unter Komposition:

$$\{f \in M \mid \text{es gibt ein } g \in M \text{ mit } g \circ f = \text{id}\} = \{f \in M \mid f \text{ ist injektiv}\},$$

$$\{f \in M \mid \text{es gibt ein } g \in M \text{ mit } f \circ g = \text{id}\} = \{f \in M \mid f \text{ ist surjektiv}\},$$

$$M^\times = \{f \in M \mid f \text{ ist invertierbar}\} = \{f \in M \mid f \text{ ist bijektiv}\} = S_A.$$

In diesem Zusammenhang ist überraschend:

**Vereinfachter Nachweis der Gruppenaxiome**

Eine Halbgruppe  $G$  ist eine Gruppe, falls gilt:

(G1) Es gibt ein  $e \in G$ , sodass  $a \circ e = a$  für alle  $a \in G$ .

(G2) Ist  $e$  wie in (G1), so gilt: Für alle  $a \in G$  gibt es ein  $b \in G$  mit  $a \circ b = e$ .

Sei nämlich  $e$  wie in (G1). Ist nun  $a \in G$  beliebig, so gibt es nach (G2) ein  $b$  mit  $a \circ b = e$  und ein  $c$  mit  $b \circ c = e$ . Dann gilt  $b \circ a = b \circ a \circ e = b \circ a \circ b \circ c = b \circ e \circ c = b \circ c = e$ , und damit  $e \circ a = a \circ b \circ a = a \circ e = a$ . Dies zeigt, dass  $G$  eine Gruppe ist.

## 2.4 Rechenregeln in Gruppen

### Satz (Kürzungs- und Inversenregeln)

Sei  $(G, \circ)$  eine Gruppe. Dann gelten für alle  $a, b, c \in G$ :

#### Kürzungsregeln

$$a \circ b = a \circ c \text{ impliziert } b = c,$$

$$b \circ a = c \circ a \text{ impliziert } b = c.$$

$$(x \circ a)^{-1} \circ b \circ c = a^{-1} \circ d \circ c$$

$$(x \circ a)^{-1} \circ b = a^{-1} \circ d$$

Kürzen von  $c$

#### Lösbarkeit von Gleichungen

Die Gleichungen

$$a \circ x = b \text{ bzw. } x \circ a = b$$

sind eindeutig lösbar durch

$$x = a^{-1} \circ b \text{ bzw. } x = b \circ a^{-1}.$$

$$a^{-1} \circ x^{-1} \circ b = a^{-1} \circ d$$

Inversenregel

$$x^{-1} \circ b = d$$

Kürzen von  $a^{-1}$

$$b = x \circ d$$

Mult. mit  $x$  von links

$$b \circ d^{-1} = x$$

Mult. mit  $d^{-1}$  von rechts

#### Inversenregeln

$$(a^{-1})^{-1} = a, (a \circ b)^{-1} = b^{-1} \circ a^{-1}.$$

Schrittweises Auflösen einer gegebenen Gleichung (erste Zeile) nach  $x$  in einer Gruppe. Rechts steht eine Begründung für den gerade durchgeführten Schritt.

Diese Regeln sind ständig im Einsatz. Ihre Beweise sind kurz und instruktiv.

### Beweis der Kürzungsregeln

Gilt  $a \circ b = a \circ c$ , so gilt  $a^{-1} \circ a \circ b = a^{-1} \circ a \circ c$  und damit  $e \circ b = e \circ c$  und damit  $b = c$ . Das Argument können wir kurz so zusammenfassen:

*Multiplizieren der Gleichung „ $a \circ b = a \circ c$ “ mit  $a^{-1}$  von links entfernt das  $a$ .*

Dies erklärt auch den Namen „Kürzungsregel“. Analoges gilt für die zweite Regel.

### Beweis der eindeutigen Lösbarkeit von Gleichungen

Einsetzen von  $a^{-1} \circ b$  für  $x$  zeigt, dass eine Lösung von  $a \circ x = b$  vorliegt. Gilt umgekehrt  $a \circ y = b$  für ein  $y$ , so zeigt die Multiplikation mit  $a^{-1}$  von links, dass  $y = a^{-1} \circ b$ . Analoges gilt für die zweite Gleichung.

### Beweis der Inversenregeln

Für die erste Regel beobachten wir, dass für alle  $a \in G$  gilt:

$$a^{-1} \circ a = a \circ a^{-1} = e,$$

sodass  $a$  das eindeutige Inverse von  $a^{-1}$  ist. Damit ist  $a = (a^{-1})^{-1}$ . Eine doppelte Invertierung darf man also streichen. Das Inverse des Inversen von  $a$  ist  $a$ .

Die zweite Regel folgt aus

$$(a \circ b) \circ (b^{-1} \circ a^{-1}) = a \circ b \circ b^{-1} \circ a^{-1} = a \circ e \circ a^{-1} = a \circ a^{-1} = e.$$

Der Leser beachte, dass sich die Reihenfolge beim Invertieren umkehrt. Das folgende Beispiel (2) zeigt, dass dies beachtet werden muss.

### Beispiele

(1) Sei  $(G, \circ)$  eine Gruppe und  $a, b, c \in G$ . Dann gilt:

$$(a \circ b^{-1})^{-1} = (b^{-1})^{-1} \circ a^{-1} = b \circ a^{-1},$$

$$(a \circ b \circ c)^{-1} = (a \circ (b \circ c))^{-1} = (b \circ c)^{-1} \circ a^{-1} = c^{-1} \circ b^{-1} \circ a^{-1}.$$

(2) Sei  $S_3$  die symmetrische Gruppe auf  $\{1, 2, 3\}$ , und seien  $a = (2, 3, 1)$ ,  $b = (1, 3, 2)$  und  $c = a \circ b = (2, 1, 3)$  (vgl. 1.5 zur Notation und 2.3 zu  $S_n$ ). Dann gilt

$$a^{-1} = (3, 1, 2), \quad b^{-1} = b, \quad c^{-1} = c = b^{-1} \circ a^{-1}, \quad a^{-1} \circ b^{-1} = (3, 2, 1).$$

$$\text{Also gilt } (a \circ b)^{-1} = c^{-1} = c \neq a^{-1} \circ b^{-1}.$$

Die eindeutige Lösbarkeit von Gleichungen führt zu einer bemerkenswerten kombinatorischen Eigenschaft der Operationstafel einer Gruppe (Gruppentafel):

### Bijektivität der Translationen

Ist  $G$  eine Gruppe und  $a \in G$ , so ist die *Linkstranslation*  $\ell_a : G \rightarrow G$  bijektiv, wobei

$$\ell_a(b) = a \circ b \quad \text{für alle } b \in G.$$

Gleiches gilt für die *Rechtstranslation*  $r_a : G \rightarrow G$  mit  $r_a(b) = b \circ a$  für alle  $b \in G$ .

Anschaulich interpretiert bedeutet dies:

*In den Zeilen und Spalten einer Gruppentafel stehen Permutationen von  $G$ .*

Umgekehrt gilt:

### Charakterisierung von Gruppen

Eine Halbgruppe  $H \neq \emptyset$  ist genau dann eine Gruppe, wenn alle  $\ell_a : H \rightarrow H$  und  $r_a : H \rightarrow H$  bijektiv sind. Zudem kann „bijektiv“ durch „surjektiv“ ersetzt werden.

### Beispiele

(1) Sei  $V = \{e, a, b, c\}$  mit paarweise verschiedenen  $e, a, b, c$ . Wir definieren  $\circ$  auf  $V$  durch die Tafel rechts. Man überprüft, dass  $\circ$  assoziativ ist. Da in den Zeilen und Spalten Permutationen stehen, ist  $V$  eine Gruppe. Sie heißt die *Kleinsche Vierergruppe*.

$\circ$	e	a	b	c
e	e	a	b	c
a	a	e	c	b
b	b	c	e	a
c	c	b	a	e

(2) In der Tafel rechts stehen in allen Zeilen und Spalten Permutationen, aber die Operation ist nicht assoziativ, da

$$(1 \circ 1) \circ 2 = 3 \circ 2 = 1, \quad 1 \circ (1 \circ 2) = 1 \circ 2 = 2.$$

Es liegt also keine Halbgruppe und damit auch keine Gruppe vor.

$\circ$	1	2	3
1	3	2	1
2	1	3	2
3	2	1	3

## 2.5 Kommutative Operationen

### Definition (kommutative Operation, abelsche Struktur)

Eine Operation  $\circ : M^2 \rightarrow M$  auf einer Menge  $M$  heißt *kommutativ*, falls gilt:

$$a \circ b = b \circ a \quad \text{für alle } a, b \in M. \quad (\text{Kommutativgesetz})$$

Wir nennen dann  $(H, \circ)$  *kommutativ* oder *abelsch*.

Dem Gesetz entspricht erneut eine anschauliche Eigenschaft der Operationstafel: Es gilt genau dann, wenn die Tafel symmetrisch ist, d.h. die Spiegelung an der Diagonalen die Tafel nicht ändert.

$\circ$	a	b	c	...
a	$a^2$	$a \circ b$	$a \circ c$	...
b	$a \circ b$	$b^2$	$b \circ c$	...
c	$a \circ c$	$b \circ c$	$c^2$	...
...	...	...	...	...

### Beispiele

(1) Die Monoide

$$(\mathbb{N}, +), (\mathbb{Z}, +), (\mathbb{Q}, +), (\mathbb{R}, \cdot), (\mathbb{R}^*, \cdot)$$

sind kommutativ. Die Restklassengruppen  $(\mathbb{Z}_m, +)$  und  $(\mathbb{Z}_m^*, \cdot)$  sind abelsch für alle  $m \geq 1$ . Ebenso ist die Kleinsche Vierergruppe  $V = \{e, a, b, c\}$  abelsch.

(2) Sind  $H_1$  und  $H_2$  kommutative Halbgruppen, so auch  $H_1 \times H_2$ .

(3) Eine Permutationsgruppe  $S_A$  ist genau dann abelsch, wenn  $A$  höchstens zwei Elemente hat. Für die Gruppe  $S_3$  gilt zum Beispiel

$$(2, 3, 1) \circ (1, 3, 2) = (2, 1, 3), \quad (1, 3, 2) \circ (2, 3, 1) = (3, 2, 1).$$

Eine kommutative Operation bringt viele Vereinfachungen mit sich. In Analogie zum Assoziativgesetz können wir die Wirkung der Kommutativität so zusammenfassen:

*Wir dürfen beliebig umordnen.*

In einer kommutativen Halbgruppe gilt beispielsweise

$$(a \circ b)^2 = a \circ b \circ a \circ b = a \circ a \circ b \circ b = a^2 \circ b^2.$$

Beim zweiten „=“ wird die Kommutativität  $b \circ a = a \circ b$  benutzt. Gilt sie nicht, so ist das „Reinziehen“ des Exponenten in der Regel nicht erlaubt. Dies ist auch die einzige Schwierigkeit, die das Gesetz bereitet: Man darf es nicht anwenden, wenn es nicht gilt.

Allgemein gilt:

### Potenzierung in kommutativen Strukturen

Ist  $H$  eine kommutative Halbgruppe, so gilt:

$$(a \circ b)^n = a^n \circ b^n \quad \text{für alle } a, b \text{ und } n \geq 1.$$

Ist  $H$  ein Monoid oder eine Gruppe, so gilt dies für alle  $n \in \mathbb{N}$  bzw. alle  $n \in \mathbb{Z}$ .



**Beispiel**

In einer abelschen Gruppe  $G$  gilt

$$(a \circ b)^{-1} = a^{-1} \circ b^{-1} \quad \text{für alle } a, b \in G.$$

Der Leser vergleiche dies mit dem Gegenbeispiel für die  $S_3$  in Abschnitt 2.4.

In kommutativen Strukturen sind spezielle Notationen üblich. Die drei folgenden Bemerkungen stellen das Wichtigste hierzu zusammen.

**Verwendung des Additionszeichens**

Das Additionszeichen  $+$  wird ausschließlich für kommutative Operationen verwendet. Andere Operationszeichen wie  $\circ, *, \cdot$  können sowohl für kommutative als auch für nichtkommutative Operationen verwendet werden.

**Notationen für das Pluszeichen**

In additiv notierten (und also kommutativen) Strukturen schreiben wir

$$na \text{ statt } a^n, \quad (\text{Vervielfachung})$$

$$-a \text{ statt } a^{-1}, \quad (\text{additive Inverse})$$

$$a - b \text{ statt } a + (-b), \quad (\text{Subtraktion})$$

$$\sum_{1 \leq k \leq n} a_k \text{ statt } \prod_{1 \leq k \leq n} a_k, \quad \sum_{1 \leq k \leq 0} a_k = 0. \quad (\text{Summe})$$

**Notationen für ein kommutatives Multiplikationszeichen**

In abelschen Gruppen  $(G, \cdot)$  mit neutralem Element 1 schreiben wir auch

$$1/a \text{ statt } a^{-1}, \quad (\text{Bruchnotation})$$

$$a/b \text{ statt } a \cdot 1/b. \quad (\text{Division})$$

Unsere Rechengesetze lassen sich mit den neuen Notationen umschreiben. Wir geben exemplarisch einige Übersetzungen an.

**Beispiele**

- (1) In einer abelschen Gruppe  $(G, +)$  gilt für alle  $a, b \in G$  und  $n, m \in \mathbb{Z}$ :

$$m(na) = (mn)a, \quad na + ma = (n+m)a, \quad n(a+b) = na + nb,$$

$$-(-a) = a, \quad -(a+b) = -b - a = -a - b.$$

- (2) In einer abelschen Gruppe  $(G, \cdot)$  mit neutralem Element 1 gilt für alle  $a, b \in G$ :

$$\frac{a}{a} = 1, \quad \frac{1}{1/a} = a, \quad \frac{1}{ab} = \frac{1}{b} \cdot \frac{1}{a} = \frac{1}{a} \cdot \frac{1}{b},$$

$$\frac{1}{a/b} = \frac{b}{a} \quad (\text{da } (a b^{-1})^{-1} = b a^{-1}, \text{ vgl. 2.4}).$$

## 2.6 Untergruppen

### Definition (Untergruppe)

Sei  $(G, \circ)$  eine Gruppe, und sei  $H \subseteq G$ . Dann heißt  $H$  eine *Untergruppe* von  $G$ , falls  $H$  zusammen mit der Operation von  $G$  eine Gruppe bildet, d. h. falls  $(H, \circ|_{H^2})$  eine Gruppe ist.

Zu jeder algebraischen Struktur gibt es Unterstrukturen, und wir könnten auch Unterhalbgruppen und Untermonoide betrachten. Wir beschränken uns hier auf Gruppen.

Ist  $(H, \circ|_{H^2})$  eine Gruppe, so gilt  $\circ|_{H^2} : H^2 \rightarrow H$  und damit

$a \circ b \in H$  für alle  $a, b \in H$ .

Eine Untergruppe  $H$  ist also abgeschlossen unter  $\circ$  (vgl. 1. 8).

Je nach Kontext fassen wir eine Untergruppe  $H$  von  $G$  als Teilmenge von  $G$  oder als vollwertige Gruppe auf.

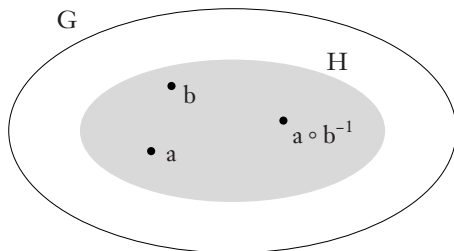


Illustration des Untergruppenkriteriums: Eine nichtleere Teilmenge  $H$  einer Gruppe ist genau dann eine Untergruppe, wenn für je zwei Elemente  $a$  und  $b$  in  $H$  auch  $a \circ b^{-1}$  ein Element von  $H$  ist.

### Beispiele

Wir betrachten die abelsche Gruppe  $(\mathbb{Z}, +)$ .

- (1)  $H = \{ 2a + 1 \mid a \in \mathbb{Z} \}$  ist nicht abgeschlossen unter  $+$ , da  $1 + 1 \notin H$ . Also ist  $+|_{H^2}$  keine Operation auf  $H$  und damit  $H$  keine Untergruppe von  $\mathbb{Z}$ .
- (2)  $\mathbb{N} \subseteq \mathbb{Z}$  ist abgeschlossen unter  $+$ , da  $n + m \in \mathbb{N}$  für alle  $n, m \in \mathbb{N}$ . Aber  $\mathbb{N}$  ist keine Untergruppe von  $\mathbb{Z}$ , da  $(\mathbb{N}, +)$  keine Gruppe ist.
- (3)  $H = \{ 2a \mid a \in \mathbb{Z} \}$  ist abgeschlossen unter  $+$ . Die Operation  $+$  ist nach wie vor assoziativ,  $0 \in H$  ist neutral und  $-2a \in H$  ist invers zu  $2a \in H$ . Also ist  $H$  eine Untergruppe von  $\mathbb{Z}$ .

Das folgende Kriterium erleichtert den Nachweis, ob eine Menge  $H \subseteq G$  eine Untergruppe bildet oder nicht.

Untergruppenkriterium
$H \subseteq G$ ist genau dann eine Untergruppe von $G$ , wenn gilt: (UG1) $H \neq \emptyset$ . (UG2) Für alle $a, b \in H$ ist $a \circ b^{-1} \in H$ .

## Beispiele

- (1)  $\mathbb{Z}$  ist eine Untergruppe von  $(\mathbb{Q}, +)$ ,  $\mathbb{Q}$  ist eine Untergruppe von  $(\mathbb{R}, +)$  und  $\mathbb{R}$  ist eine Untergruppe von  $(\mathbb{C}, +)$ .
- (2) Für jede Gruppe  $G$  sind  $\{e\}$  und  $G$  die sog. *trivialen Untergruppen* von  $G$ .
- (3)  $\{(x_1, x_2, 0) \mid x_1, x_2 \in \mathbb{R}\}$  ist eine Untergruppe von  $(\mathbb{R}^3, +)$ .
- (4) Wir betrachten noch einmal die Gruppe  $(\mathbb{Z}, +)$ . Sei  $m \in \mathbb{N}$  und

$$m\mathbb{Z} = \{ma \mid a \in \mathbb{Z}\} = \{am \mid a \in \mathbb{Z}\} = \mathbb{Z}m$$

die Menge der ganzzahligen Vielfachen von  $m$ . Dann ist  $m\mathbb{Z} \neq \emptyset$  und es gilt

$$am - bm = (a - b)m \in m\mathbb{Z} \quad \text{für alle } am, bm \in m\mathbb{Z}.$$

Nach dem Untergruppenkriterium ist also  $m\mathbb{Z}$  eine Untergruppe von  $(\mathbb{Z}, +)$ . Man kann zeigen, dass alle Untergruppen von  $(\mathbb{Z}, +)$  von der Form  $m\mathbb{Z}$  sind. Die Beweisidee ist: Ist  $H \neq \{0\}$  eine Untergruppe von  $(\mathbb{Z}, +)$ , so setzen wir

$$m = \min_{a \in H, a \neq 0} |a|.$$

Aus den Abgeschlossenheitseigenschaften von  $H$  folgt  $am \in H$  für alle  $a \in \mathbb{Z}$ . Also ist  $m\mathbb{Z} \subseteq H$ . Eine Division mit Rest zeigt, dass es kein  $b \in H - m\mathbb{Z}$  gibt: Ansonsten wäre  $b = am + c$  für ein  $0 < c < m$  und damit  $c = b - am \in H$ . Also ist  $H = m\mathbb{Z}$ .

- (5) Seien  $G$  eine Gruppe und  $a \in G$ . Dann ist der Abschluss

$$\langle a \rangle = \{a^n \mid n \in \mathbb{Z}\}$$

der Menge  $\{a\}$  unter der Gruppenoperation eine Untergruppe von  $G$  (vgl. 1.8). Denn es gilt  $a^0 = e \in \langle a \rangle$  und für alle  $a^n, b^m \in \langle a \rangle$  ist

$$a^n \circ (a^m)^{-1} = a^n \circ a^{-m} = a^{n-m} \in \langle a \rangle.$$

Allgemein definieren wir (mit den Begriffsbildungen aus 1.8):

### Definition (erzeugte Untergruppe, zyklisch)

Sei  $G$  eine Gruppe und  $A \subseteq G$ . Dann heißt der Abschluss  $\langle A \rangle$  von  $A$  unter  $\circ$  die von  $A$  erzeugte Untergruppe. Gilt  $G = \langle A \rangle$ , so wird  $G$  von  $A$  erzeugt.  $G$  heißt *zyklisch*, falls  $G$  von einem Element  $a$  erzeugt wird, d. h., es gibt ein  $a$  mit  $\langle a \rangle = G$ .

Jede zyklische Gruppe  $G = \langle a \rangle$  ist abelsch, da  $a^n a^m = a^{m+n} = a^m a^n$  für alle  $n, m \in \mathbb{Z}$ .

## Beispiele

- (1) Für  $(\mathbb{Z}, +)$  und  $m \in \mathbb{N}$  gilt  $\langle m \rangle = \langle -m \rangle = \{am \mid a \in \mathbb{Z}\} = m\mathbb{Z}$ .
- (2) Die Kleinsche Vierergruppe  $V = \{e, a, b, c\}$  ist abelsch, aber nicht zyklisch, da
 
$$\langle e \rangle = \{e\}, \quad \langle a \rangle = \{e, a\}, \quad \langle b \rangle = \{e, b\}, \quad \langle c \rangle = \{e, c\}.$$

## 2.7 Normalteiler und Faktorgruppen

### Definition (Nebenklassen, Normalteiler, Faktorgruppe)

Seien  $G$  eine Gruppe und  $H$  eine Untergruppe von  $G$ .

*Äquivalenzrelationen und Nebenklassen bzgl.  $H$*

Wir definieren zwei Äquivalenzrelationen  $\sim$  und  $\approx$  auf  $G$  durch

$$\begin{aligned} a \sim b, & \quad \text{falls } ba^{-1} \in H, \\ a \approx b, & \quad \text{falls } a^{-1}b \in H \end{aligned} \quad \text{für alle } a, b \in G.$$

Für alle  $a \in G$  heißen

$a/\sim = Ha = \{ha \mid h \in H\}$  die *H-Links- oder a-Rechtsnebenklasse* und

$a/\approx = aH = \{ah \mid h \in H\}$  die *H-Rechts- oder a-Linksnebenklasse*

von  $a$  in  $G$  bzgl.  $H$ .

### Normalteiler

Die Untergruppe  $H$  heißt ein *Normalteiler* von  $G$ , falls  $\sim = \approx$ , d. h., falls

$$aH = Ha \quad \text{für alle } a \in G. \quad (\text{Normalteiler-Bedingung})$$

Wir nennen dann  $aH = Ha$  die *Nebenklasse* von  $a$  in  $G$  bzgl.  $H$  und definieren  $G/H = \{aH \mid a \in G\}$  und eine Operation  $\cdot : (G/H)^2 \rightarrow G/H$  durch

$$aH \cdot bH = (ab)H \quad \text{für alle } aH, bH \in G/H.$$

Schließlich heißt  $(G/H, \cdot)$  die *Faktorgruppe* von  $G$  bzgl.  $H$ .

Jede Untergruppe  $H$  von  $G$  induziert zwei Äquivalenzrelationen  $\sim$  und  $\approx$  auf  $G$ . Für alle  $a, b \in G$  sind äquivalent:

- (1)  $ba^{-1} \in H$ .
- (2) Es gibt ein  $h \in H$  mit  $ba^{-1} = h$ .
- (3) Es gibt ein  $h \in H$  mit  $b = ha$ .
- (4)  $b \in Ha = \{ha \mid h \in H\}$ .

Dies zeigt, dass  $a/\sim = Ha$ . Analog gilt  $a/\approx = aH$ .

Für Normalteiler ist  $\sim = \approx$ . Es gilt dann  $aH = Ha$  für alle  $a \in G$  oder gleichwertig

$$ba^{-1} \in H \quad \text{für alle } a \in G \text{ und } b \in H. \quad (\text{Normalteiler-Bedingung, Umformulierung})$$

Nicht jede Untergruppe ist ein Normalteiler:

aH			H				G
• a			• e				

Die Relation  $\approx$  zerlegt  $G$  in Äquivalenzklassen  $aH$ . Alle Äquivalenzklassen sind gleichmächtig (Satz von Lagrange).

**Beispiele**

(1) Für jede Gruppe  $G$  sind die trivialen Untergruppen  $\{e\}$  und  $G$  Normalteiler.

(2) Ist die Gruppe  $G$  abelsch, so gilt

$$aH = \{ah \mid h \in H\} = \{ha \mid h \in H\} = Ha.$$

Damit ist jede Untergruppe einer abelschen Gruppe ein Normalteiler.

(3) Die Untergruppe  $H = \{(1, 2, 3), (1, 3, 2)\}$  von  $S_3$  ist kein Normalteiler, da

$$(3, 2, 1)H = \{(3, 2, 1) \circ (1, 2, 3), (3, 2, 1) \circ (1, 3, 2)\} = \{(3, 2, 1), (3, 1, 2)\},$$

$$H(3, 2, 1) = \{(1, 2, 3) \circ (3, 2, 1), (1, 3, 2) \circ (3, 2, 1)\} = \{(3, 2, 1), (2, 3, 1)\}.$$

Dagegen ist  $H' = \{(1, 2, 3), (2, 3, 1), (3, 1, 2)\}$  ein Normalteiler der  $S_3$ .

Für einen Normalteiler  $H$  kann auf der Menge der Nebenklassen

$$G/H = \{aH \mid a \in G\} = \{Ha \mid a \in G\}$$

eine Operation erklärt werden:  $aH \cdot bH = (ab)H$ . Dass  $H$  ein Normalteiler ist, ist wichtig:

**Beispiel**

Für  $H \subseteq S_3$ , wie oben würde  $aH \cdot bH = (ab)H$  implizieren, dass

$$(3, 2, 1)H \cdot (3, 2, 1)H = (1, 2, 3)H = H \neq (2, 3, 1)H = (3, 1, 2)H \cdot (3, 1, 2)H,$$

obwohl  $(3, 2, 1)H = \{(3, 2, 1), (3, 1, 2)\} = (3, 1, 2)H$ . Mit anderen Worten:

$(ab)H$  hängt von der Wahl von  $a, b$  und nicht nur von  $aH$  und  $bH$  ab, sodass – wie man in solchen Situationen sagt – die Multiplikation nicht wohldefiniert ist.

Ist  $H$  ein Normalteiler von  $G$ , so ist  $G/H$  eine Gruppe. Die Nebenklasse  $eH = H$  ist neutral und das Inverse von  $aH$  ist  $a^{-1}H$ . Im Allgemeinen ist  $(ab)H \neq (ba)H$ , sodass die Faktorgruppe  $G/H$  nicht notwendig abelsch ist.

Für additiv notierte Gruppen  $(G, +)$  haben Nebenklassen die Gestalt

$$a + H = \{a + h \mid h \in H\} = \{h + a \mid h \in H\} = H + a.$$

In dieser Form werden sie uns auch in der Vektorraumtheorie begegnen (vgl. 3.11, 4.5).

**Beispiel**

Für alle  $m \geq 1$  ist  $m\mathbb{Z}$  ein Normalteiler von  $(\mathbb{Z}, +)$ . Es gilt:

$$\mathbb{Z}/m\mathbb{Z} = \{a + m\mathbb{Z} \mid a \in \mathbb{Z}\} = \{[a]_m \mid a \in \mathbb{Z}\} = \mathbb{Z}_m.$$

Allgemeine Normalteiler und ihre Faktorgruppen  $G/H$  werden wir im Homomorphiesatz noch einmal betrachten (vgl. 4.4). In der Algebra spielen sie eine Schlüsselrolle bei der Klassifikation von endlichen Gruppen und der Untersuchung der Frage, ob polynomielle Gleichungen durch Wurzelziehen lösbar sind.

## 2.8 Ringe

### Definition (Ring, kommutativer Ring, Nullteilerfreiheit)

Sei  $R$  eine Menge, und seien  $+: R^2 \rightarrow R$  und  $\cdot: R^2 \rightarrow R$  Operationen auf  $R$ . Dann heißt das Tripel  $(R, +, \cdot)$  ein *Ring (mit 1)*, falls gilt:

- (a)  $(R, +)$  ist eine abelsche Gruppe.
- (b)  $(R, \cdot)$  ist ein Monoid.
- (c) Für alle  $a, b, c \in R$  gilt:

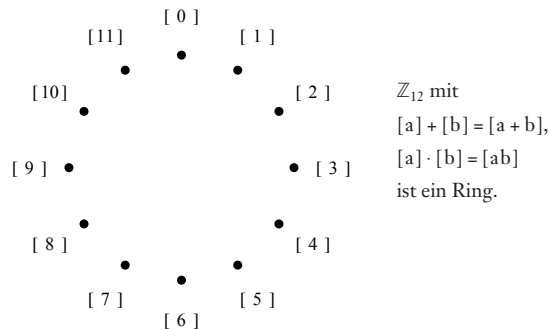
$$a(b + c) = ab + ac, \quad (a + b)c = ac + bc. \quad (\text{Distributivgesetze})$$

Ein Ring  $R$  heißt *kommutativ*, wenn  $\cdot$  kommutativ ist. Er heißt *nullteilerfrei*, falls

$$a \cdot b = 0 \text{ impliziert } a = 0 \text{ oder } b = 0 \text{ für alle } a, b \in R. \quad (\text{Nullteilerfreiheit})$$

Ein  $a \in R$  heißt *invertierbar* oder eine *Einheit*, wenn es ein  $b$  gibt mit  $ab = ba = 1$ . Die multiplikative Gruppe aller Einheiten in  $R$  wird mit  $R^\times$  bezeichnet.

Ringe sind unsere erste algebraische Struktur, bei der mehr als eine Operation vorhanden ist. Ein Ring ist ausgestattet mit einer (kommutativen) Addition und einer (nicht notwendig kommutativen) Multiplikation. Die Distributivgesetze verbinden die beiden Operationen.



### Konventionen

„Mal“ bindet stärker als „Plus“, der Malpunkt kann weggelassen werden, 0 ist das neutrale Element von  $(R, +)$ , 1 das neutrale Element von  $(R, \cdot)$ .

### Rechenregeln in Ringen

$$a \cdot 0 = 0 = 0 \cdot a, \quad (-a) \cdot b = -(ab) = a \cdot (-b), \quad (-a)(-b) = ab.$$

Diese Regeln ergeben sich aus den Beobachtungen:

$$a \cdot 0 = a(0 + 0) = a \cdot 0 + a \cdot 0, \text{ sodass } 0 = a \cdot 0, \quad (\text{analog für } 0a)$$

$$0 = (a - a)b = ab + (-a)b, \text{ sodass } -(ab) = (-a)b, \quad (\text{analog für } a(-b)),$$

$$(-a)(-b) = -(a(-b)) = --(ab) = ab.$$

Die Distributivgesetze schreiben also den Wert für die Multiplikation mit 0 vor und erzwingen „Minus mal Minus gleich Plus“.

**Warnung**

Die Regel  $(-a)b = -(ab)$  kann nicht mit dem Assoziativgesetz begründet werden: Das Argument „man darf beliebig Klammern setzen“ ist hier nicht korrekt, da das Minuszeichen kein Element von  $R$  ist, sondern additiv Inverse bezeichnet.

**Beispiele**

- (1)  $R = \{0\}$  mit  $0 + 0 = 0$  und  $0 \cdot 0 = 0$  ist der sog. *Nullring* oder *triviale Ring*. Er ist der einzige Ring, der  $1 = 0$  erfüllt.
- (2)  $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$  (mit  $+$  und  $\cdot$  wie üblich) sind kommutative nullteilerfreie Ringe. In  $\mathbb{Z}$  sind genau die Elemente  $1$  und  $-1$  Einheiten, sodass  $\mathbb{Z}^\times = \{-1, 1\}$ .
- (3) Für alle  $m \geq 1$  bildet  $\mathbb{Z}_m$  mit  $[a] + [b] = [a + b]$  und  $[a] \cdot [b] = [ab]$  einen Ring mit  $0 = [0]$  und  $1 = [1]$ .  $\mathbb{Z}_1$  ist der Nullring. Ist  $m = ab$ , so gilt  $[a] \cdot [b] = [0] = 0$ . Also ist  $\mathbb{Z}_m$  genau dann nullteilerfrei, wenn  $m = 1$  oder  $m$  prim ist.
- (4) Ist  $M$  eine Menge, so ist  $(\mathcal{P}(M), \Delta, \cap)$  mit der symmetrischen Differenz
 
$$A \Delta B = (A - B) \cup (B - A) \quad \text{für alle } A, B \subseteq M$$
 als Addition ein kommutativer Ring mit  $0 = \emptyset$ ,  $1 = M$ ,  $-A = A^c$  für alle  $A \subseteq M$ .
- (5) Sind  $R_1$  und  $R_2$  Ringe, so ist  $R = R_1 \times R_2$  mit den Produktoperationen ein Ring.

Wichtige nichtkommutative Ringe werden wir in Kapitel 5 kennenlernen.

Weitere Rechenregeln in Ringen
--------------------------------

Für alle kommutierenden  $a, b \in R$  (d. h.  $ab = ba$ ) und alle  $n \in \mathbb{N}$  gilt:

$$a^n - b^n = (a - b)(a^{n-1} + a^{n-2}b^1 + \dots + a^1b^{n-2} + b^{n-1}) = (a^{n-1} + a^{n-2}b^1 + \dots + a^1b^{n-2} + b^{n-1})(a - b),$$

$$a^n - 1 = (a - 1)(a^0 + a^1 + \dots + a^{n-2} + a^{n-1}) = (a^0 + a^1 + \dots + a^{n-2} + a^{n-1})(a - 1),$$

$$(a + b)^n = \sum_{0 \leq k \leq n} \binom{n}{k} a^{n-k} b^k. \quad (\text{Binomischer Lehrsatz})$$

Die erste Regel kann durch Ausmultiplizieren bewiesen werden, die zweite folgt durch Setzen von  $b = 1$  aus der ersten. Der binomische Lehrsatz lässt sich durch Induktion nach  $n$  zeigen. Der Leser wird vielleicht erkennen, dass die zweite Regel die Formel für die endliche geometrische Reihe in  $\mathbb{R}$  oder  $\mathbb{C}$  liefert:

$$\sum_{k \leq n} a^k = \frac{1 - a^{n+1}}{1 - a} \quad \text{für alle } a \in \mathbb{R} \text{ bzw. alle } a \in \mathbb{C} \text{ mit } a \neq 1.$$

## 2.9 Körper

### Definition (Divisionsbereich, Schiefkörper, Körper)

Ein Ring  $(K, +, \cdot)$  heißt ein *Divisionsbereich* oder *Schiefkörper*, falls für  $K^* = K - \{0\}$  gilt:

$(K^*, \cdot)$  ist eine Gruppe (wobei  $(K^*, \cdot)$  wieder kurz für  $(K^*, \cdot|K^{*2})$  steht).

Ist der Ring zudem kommutativ, so heißt  $(K, +, \cdot)$  ein *Körper*.

Ein Körper ist also ein kommutativer nichttrivialer Ring, in dem alle von Null verschiedenen Elemente Einheiten sind, d. h., es gilt  $K^\times = K^* = K - \{0\}$ .

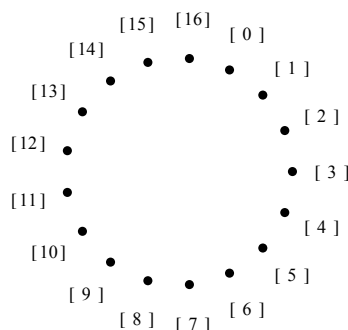
Die Sonderbehandlung der 0 ist unvermeidlich: Würde ein multiplikatives Inverses  $0^{-1}$  der 0 existieren, so würde gelten:

$$0 = 0 \cdot 0^{-1} = 1.$$

Beim ersten „=“ verwenden wir die in allen Ringen gültige Regel  $0a = 0$  und beim zweiten „=“ die Definition des multipli-

kativen Inversen. Damit kann die Null nur im Nullring  $\{0\}$  invertierbar sein!

Wir fassen den Körperbegriff noch einmal zusammen.  $(K, +)$  erfüllt vier Axiome: Assoziativität, Existenz eines neutralen Elements, Existenz von Inversen, Kommutativität. Gleiches gilt für  $(K^*, \cdot)$ . Zudem gelten zwei Distributivgesetze. Damit ergeben sich insgesamt zehn *Körperaxiome*. Automatisch gilt:



$\mathbb{Z}_{17}$  mit

$$[a] + [b] = [a + b],$$

$$[a] \cdot [b] = [ab]$$

ist ein Körper. Es gilt zum Beispiel

$$[5] \cdot [7] + [16] =$$

$$[35] + [16] =$$

$$[1] + [16] =$$

$$[17] = [0] = 0.$$

### Nullteilerfreiheit in Körpern

Für alle  $a, b \in K$  gilt:

$$ab = 0 \text{ impliziert } a = 0 \text{ oder } b = 0.$$

Ist nämlich  $ab = 0$  und  $a \neq 0$ , so existiert  $a^{-1}$ , sodass

$$b = a^{-1}0 = 0.$$

In einem Körper  $K$  stehen alle vier Grundrechenarten zur Verfügung:  $+, -, \cdot : K^2 \rightarrow K$  wie in jedem Ring, und zusätzlich auch eine Division  $/ : K \times K^* \rightarrow K$  vermöge

$$a/b = a b^{-1} \quad \text{für alle } a, b \in K \text{ mit } b \neq 0.$$

Die Bruchnotation  $a/b$  ist aufgrund der Kommutativität möglich, da  $a \cdot 1/b = 1/b \cdot a$  (vgl. auch 2.5). Es gelten die vertrauten Rechengesetze:



### Rechenregeln in Körpern (Bruchrechnen)

Für alle  $a, c \in K$  und  $b, d \in K^*$  gilt:

$$\frac{a}{b} + \frac{c}{d} = \frac{ad + bc}{bd}; \quad \frac{a}{b} \cdot \frac{c}{d} = \frac{ac}{bd}; \quad \frac{a/b}{c/d} = \frac{ad}{bc}, \text{ falls } c \neq 0.$$

#### Beispiele

- (1)  $\mathbb{Q}, \mathbb{R}, \mathbb{C}$  sind mit den üblichen Operationen Körper. Für alle  $p \geq 1$  ist der Restklassenring  $\mathbb{Z}_p$  genau dann ein Körper, wenn  $p$  eine Primzahl ist; der Körper  $\mathbb{Z}_p$  heißt dann der *Restklassenkörper* modulo  $p$ .
- (2) Seien  $K = \{0, 1\}$ ,  $+$  wie in  $\mathbb{Z}_2$  und  $\cdot$  definiert durch  $a \cdot b = a$  für alle  $a, b \in K$ . Dann sind  $(K, +)$ ,  $(K^*, \cdot)$  mit  $K^* = K - \{0\} = \{1\}$  abelsche Gruppen, aber es gilt nur ein Distributivgesetz. Damit ist  $(K, +, \cdot)$  kein Ring und insbesondere kein Körper.
- (3) Auf dem  $\mathbb{R}^4$  kann eine nichtkommutative Multiplikation  $\cdot$  erklärt werden, sodass die sog. *hamiltonschen Quaternionen*  $\mathbb{H} = (\mathbb{R}^4, +, \cdot)$  einen Schiefkörper bilden.

Der *Satz von Wedderburn* besagt, dass jeder endliche Schiefkörper bereits ein Körper ist. Damit fallen Schiefkörper und Körper im Endlichen zusammen.

### Die Charakteristik eines Körpers

Grob gesprochen ist eine Struktur  $(K, +, \cdot)$  ein Körper, wenn die „üblichen Rechenregeln“ gelten. Dabei ist aber Vorsicht geboten. Denn in  $\mathbb{Z}_p$  mit einer Primzahl  $p$  gilt

$$1 + \dots + 1 \text{ (p-oft)} = [1] + \dots + [1] \text{ (p-oft)} = [p] = 0.$$

Damit schließen die Körperaxiome nicht aus, dass wir durch Aufsummieren der Eins die Null erhalten! Dies motiviert:

#### Definition (Charakteristik eines Körpers)

Sei  $K$  ein Körper. Gibt es ein  $m \geq 1$  mit  $m1 = \sum_{1 \leq k \leq m} 1 = 0$ , so setzen wir

$\text{char}(K) =$  „das kleinste  $m \geq 1$  mit  $m1 = 0$ “. (Charakteristik von  $K$ )

Andernfalls setzen wir  $\text{char}(K) = 0$ .

Ist  $\text{char}(K) \neq 0$ , so ist  $\text{char}(K) \geq 2$ , da  $0 \neq 1$  gilt. Ist nun  $\text{char}(K) = nm$  mit  $n, m > 1$ , so ist  $(nm)1 = (n1)(m1) = 0$ , also  $n1 = 0$  oder  $m1 = 0$ , da  $K$  nullteilerfrei ist. Nach Minimalität ist dann also  $n = 1$  oder  $m = 1$ . Damit ist  $\text{char}(K)$  eine Primzahl. Die Restklassenkörper  $\mathbb{Z}_p$  zeigen, dass jede Primzahl als Charakteristik vorkommt. Die Anzahl der Elemente eines endlichen Körpers muss dagegen keine Primzahl sein. Es gilt:

#### Klassifikation endlicher Körper

Die Mächtigkeiten endlicher Körper sind genau die Zahlen  $p^n$  mit  $p$  prim und  $n \geq 1$ .

## 2.10 Angeordnete Körper

**Definition** (*angeordnete Körper, Betrag, positiv, negativ, anordenbar*)

*Angeordneter Körper*

Sei  $(K, +, \cdot)$  ein Körper und  $\leq$  eine lineare Ordnung auf  $K$ . Dann heißt  $(K, +, \cdot, <)$  oder kurz  $K$  ein *angeordneter Körper*, falls für alle  $a, b, c \in K$  gilt:

$$(a) \quad a < b \text{ impliziert } a + c < b + c, \quad (\text{Translationsinvarianz})$$

$$(b) \quad 0 < a, b \text{ impliziert } 0 < ab. \quad (\text{Positivitätsregel})$$

Für alle  $a \in K$  heißt dann

$$|a| = \begin{cases} a & \text{falls } a \geq 0, \\ -a & \text{sonst.} \end{cases}$$

der *Betrag* von  $a$ . Gilt  $a > 0$ ,  $a < 0$ ,  $a \geq 0$ , so heißt  $a$  *positiv*, *negativ* bzw. *nichtnegativ*.

Das *Vorzeichen*  $\text{sgn}(a) \in \{1, 0, -1\}$  von  $a$  ist definiert als  $\text{sgn}(a) = 1$ , falls  $a > 0$ ,  $\text{sgn}(0) = 0$  und  $\text{sgn}(a) = -1$ , falls  $a < 0$ . Weiter seien

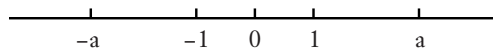
$$K^+ = \{a \in K \mid a > 0\}, \quad K_0^+ = \{a \in K \mid a \geq 0\}.$$

*Anordenbarer Körper*

Ein Körper  $(K, +, \cdot)$  heißt *anordenbar*, falls es eine lineare Ordnung auf  $K$  gibt, sodass  $(K, +, \cdot, <)$  ein angeordneter Körper ist.

Zu den algebraischen Operationen eines Körpers tritt nun also noch eine lineare Ordnung. So wie die Distributivgesetze die Addition und die Multiplikation miteinander verbinden, so verbinden die beiden *Anordnungsaxiome* (a) und (b) die Ordnung mit der Arithmetik.

Für alle  $a, b \in K$  gelten:



$$1 < a \text{ impliziert } -a < -1$$

In einem angeordneten Körper gelten die von den rationalen und reellen Zahlen vertrauten Regeln für Ungleichungen, etwa das Umdrehen des Vorzeichens bei Multiplikation mit einer negativen Zahl.

### Eigenschaften des Betrags

$$|a| \geq 0, \quad |a| = 0 \text{ genau dann, wenn } a = 0$$

$$|ab| = |a| |b| \quad (\text{Produktregel})$$

$$|a + b| \leq |a| + |b| \quad (\text{Dreiecksungleichung})$$

$$||a| - |b|| \leq |a \pm b| \quad (\text{umgekehrte Dreiecksungleichung})$$

Die beiden Anordnungsaxiome genügen, um alle vertrauten Eigenschaften für Ungleichungen herleiten zu können:

### Rechenregeln in angeordneten Körpern

Für alle  $a, b, c \in K$  gilt:

$$0 < 1, \quad -1 < 0,$$

$$0 < a, b \text{ impliziert } 0 < a + b,$$

$$a, b < 0 \text{ impliziert } a + b < 0 \text{ und } 0 < ab,$$

$$a \leq 0 \text{ und } b \geq 0 \text{ impliziert } ab \leq 0,$$

$$a < b \text{ impliziert } ca < cb, \quad \text{falls } c > 0,$$

$$a < b \text{ impliziert } cb < ca, \quad \text{falls } c < 0,$$

$$0 < a \text{ und } b > 1 \text{ impliziert } a < ab,$$

$$0 < a \text{ und } b < 1 \text{ impliziert } ab < a.$$

Exemplarisch beweisen wir hier:

#### Multiplikation einer Ungleichung $a < b$ mit $c < 0$

Aus  $c < 0$  folgt  $0 = c - c < -c$  und aus  $a < b$  folgt  $0 = a - a < b - a$  aus der Translationsinvarianz. Damit gilt nach der Positivitätsregel, dass

$$0 < (-c)(b - a) = ca - cb.$$

Wieder nach Translationsinvarianz gilt also  $cb < ca - cb + cb = ca$ .

#### Beispiele

- (1)  $\mathbb{Q}$  und  $\mathbb{R}$  sind unter den üblichen Ordnungen angeordnete Körper.
- (2) Die Restklassenkörper  $\mathbb{Z}_p$  sind nicht anordenbar. Denn aus  $0 < 1$  und der Translationsinvarianz folgt induktiv, dass  $0 < n \cdot 1$ . In  $\mathbb{Z}_p$  gilt aber  $p \cdot 1 = 0$  für das  $p$ -Fache der 1. Allgemeiner zeigt das Argument, dass  $\text{char}(K) = 0$  gilt, wenn  $K$  anordenbar ist.
- (3)  $\mathbb{C}$  ist nicht anordenbar. Denn in jedem angeordneten Körper gilt  $-1 < 0$  und  $a^2 > 0$  für alle  $a$ . Da  $i^2 = -1$  in  $\mathbb{C}$  gilt, kann  $\mathbb{C}$  nicht anordenbar sein.

Das Zahlssystem  $\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R}$  ist durch Verbesserungen gekennzeichnet: In  $\mathbb{Z}$  können wir subtrahieren, in  $\mathbb{Q}$  dividieren, in  $\mathbb{R}$  Suprema und Infima bilden. Beim Übergang von  $\mathbb{R}$  nach  $\mathbb{C}$  gewinnen wir die Lösbarkeit von Gleichungen (vgl. 2.12), aber es geht zum ersten Mal auch etwas verloren: die Ordnung der Zahlen, ein Größer und Kleiner.

#### Bemerkung

Es gibt durchaus lineare Ordnungen auf  $\mathbb{C} = \mathbb{R}^2$ . Die lexikographische Ordnung  $<_{\text{lex}}$  ist ein Beispiel (vgl. 1.3). Sie erfüllt die Translationsinvarianz. Es gilt  $i >_{\text{lex}} 0$ , aber  $i^2 = -1 <_{\text{lex}} 0$ , sodass die Positivitätsregel verletzt ist.

## 2.11 Polynomringe und Polynomfunktionen

**Definition** (*Polynom, Polynomring, Koeffizient, Unbestimmte, Grad, normiert*)

Sei  $R$  ein kommutativer Ring. Wir setzen:

$$R^{(\mathbb{N})} = \text{„die Menge aller Folgen } p = (p_n)_{n \in \mathbb{N}} \text{ in } R, \text{ die schließlich gleich } 0 \text{ sind“} = \\ \{ (p_n)_{n \in \mathbb{N}} \in R^{\mathbb{N}} \mid \text{es gibt ein } n_0, \text{ sodass } p_n = 0 \text{ für alle } n \geq n_0 \}.$$

Für alle  $p = (p_n)_{n \in \mathbb{N}}, q = (q_n)_{n \in \mathbb{N}}$   
in  $R^{(\mathbb{N})}$  definieren wir

$$p + q = (p_n + q_n)_{n \in \mathbb{N}},$$

$$p \cdot q = (\sum_{m \leq n} p_m q_{n-m})_{n \in \mathbb{N}}.$$

Der so entstehende Ring  $R^{(\mathbb{N})}$   
heißt der *Polynomring* über  $R$ .  
Die Elemente von  $R^{(\mathbb{N})}$  heißen  
*Polynome*. Für  $p \in R^{(\mathbb{N})}$  und  $n \geq 0$   
heißt  $p_n$  der  $n$ -te *Koeffizient* von  $p$ .

*Einführung einer Unbestimmten*

Wir setzen

$$X = (0, 1, 0, 0, 0, \dots)$$

und schreiben oft  $R[X]$  anstelle  
von  $R^{(\mathbb{N})}$ .

*Grad eines Polynoms*

Wir definieren die *Gradfunktion*

$$\deg : R^{(\mathbb{N})} \rightarrow \mathbb{N} \cup \{-\infty\} \text{ durch}$$

$$\deg(p) = \begin{cases} k & \text{falls } p_k \neq 0 \text{ und } p_n = 0 \text{ für alle } n \geq k, \\ -\infty & \text{falls } p_n = 0 \text{ für alle } n. \end{cases}$$

Wir nennen  $\deg(p)$  den *Grad* von  $p$ . Hat  $p$  den Grad  $k \geq 0$ , so heißt  $p_k$  der  
*Leitkoeffizient* von  $p$ . Ist der Leitkoeffizient gleich 1, so heißt  $p$  *normiert*.

Im Polynomring  $R[X] = R^{(\mathbb{N})}$  gilt  $0 = (0, 0, 0, \dots)$  (Nullpolynom) und  $1 = (1, 0, 0, 0, \dots)$ .  
Die Unbestimmte  $X$  ist ein spezielles Polynom. Für alle  $n \geq 0$  und  $p, q \in R[X]$  gilt

$$X^n = (0, \dots, 0, 1, 0, 0, 0, \dots), \quad \text{mit } n \text{ Nullen vor der } 1,$$

$$p = \sum_{n \leq \deg(p)} p_n X^n = p_0 + p_1 X + \dots + p_n X^n, \quad \text{wobei } n = \deg(p),$$

$$p \cdot q = \sum_{n \leq \deg(p) + \deg(q)} c_n X^n, \quad \text{wobei } c_n = \sum_{m \leq n} p_m q_{n-m}.$$

Diese „termartigen“ Darstellungen beherrschen den Umgang mit Polynomen.

0	0					
0	0	0				
0	0	0	0			
2	6	-2	4	0		
2	6	-2	4	0	0	
1	3	-1	2	0	0	0
	3	-1	2	0	0	0

Zur diagonalen  
Bildung der  
Koeffizienten bei  
Multiplikation

$$\begin{aligned} (3, -1, 2, 0, 0, 0, \dots) \cdot (1, 2, 2, 0, 0, 0, \dots) &= \\ (2X^2 - X + 3) \cdot (2X^2 + 2X + 1) &= \\ 4X^4 + (-2 + 4)X^3 + (6 - 2 + 2)X^2 + (6 - 1)X + 3 &= \\ 4X^4 + 2X^3 + 6X^2 + 5X + 3 &= \\ (3, 5, 6, 2, 4, 0, 0, \dots) \end{aligned}$$

**Beispiele**

Sei  $R = \mathbb{Z}$ . Wir schreiben kurz  $(p_1, \dots, p_n)$  statt  $(p_1, \dots, p_n, 0, 0, 0, \dots)$ . Dann gilt

$$(1, 0, 1) = 1 + X^2,$$

$$(1, 1)(1, 1) = (1 + X)(1 + X) = (1 + X)^2 = 1 + 2X + X^2 = (1, 2, 1),$$

$$(1, 1)(1, -1) = (1 + X)(1 - X) = 1 - X^2 = (1, 0, -1),$$

$$(3, -1, 2) \cdot (1, 2, 2) = (3, 5, 6, 2, 4) \text{ (vgl. obiges Diagramm).}$$

Der Grad eines Polynoms ist ein Maß für seine Komplexität. Dabei ist  $\deg(0) = -\infty$  eine nützliche Konvention. Die Polynome der ersten Grade haben die Formen:

$$(0, 0, 0, \dots) = 0 \quad (\text{Grad } -\infty)$$

$$(p_0, 0, 0, 0, \dots) = p_0 \quad p_0 \neq 0, \quad (\text{Grad } 0)$$

$$(p_0, p_1, 0, 0, 0, \dots) = p_0 + p_1 X \quad p_1 \neq 0. \quad (\text{Grad } 1)$$

Rechenregeln für den Grad
$\deg(p + q) \leq \max(\deg(p), \deg(q)), \quad \deg(p \cdot q) \leq \deg(p) + \deg(q).$ Ist der Ring $R$ nullteilerfrei, so gilt Gleichheit für das Produkt.

Durch Einsetzen von Ringelementen für  $X$  in  $\sum_{k \leq n} p_k X^k$  erhalten wir Funktionen:

**Definition** (*Polynomfunktionen, Nullstelle*)

Für  $p = \sum_{k \leq \deg(p)} p_k X^k \in R[X]$  ist die *Polynomfunktion*  $f_p : R \rightarrow R$  definiert durch

$$f_p(x) = \sum_{k \leq \deg(p)} p_k x^k \quad \text{für alle } x \in R.$$

Gilt  $f_p(w) = 0$  für ein  $w \in R$ , so heißt  $w$  eine *Nullstelle* von  $p$ .

In der Regel sind  $p$  und  $f_p$  zu unterscheiden:

**Beispiel**

Für  $p = (X - [0])(X - [1])(X - [2]) \in \mathbb{Z}_3[X]$  gilt  $f_p(x) = 0$  für alle  $x \in \mathbb{Z}_3$ . Damit gilt  $f_p = f_0$  mit dem Nullpolynom  $0$ , aber es gilt  $p \neq 0$ . Gleiches gilt für  $p^n$ ,  $n \geq 1$ .

Für viele Ringe  $R$  und insbesondere für jeden unendlichen Körper kann man aber zeigen, dass  $f_p \neq f_q$  für  $p \neq q$ . Eine Identifizierung von  $p$  und  $f_p$  ist dann möglich. Sie wird speziell in analytischen Überlegungen (für  $R = \mathbb{R}$  oder  $R = \mathbb{C}$ ) oft durchgeführt.

**Allgemeinere Einsetzungen**

Ist  $A$  ein zweiter Ring und  $\cdot : R \times A \rightarrow A$  erklärt, so ist  $\sum_{k \leq n} p_k x^k \in A$  für alle  $p_k \in R$  und  $x \in A$  definiert. Wir erhalten so für alle  $p \in R[X]$  eine Polynomfunktion der Form  $f_p : A \rightarrow A$ . Ein Beispiel werden wir in Kapitel 8 kennenlernen, wenn wir Matrizen in ein Polynom einsetzen.

## 2.12 Division und Nullstellen von Polynomen

**Satz** (*Polynomdivision, Abspalten einer Nullstelle, Anzahl der Nullstellen*)

Sei  $R$  ein kommutativer Ring, und sei  $R[X] = R^{(\mathbb{N})}$  der Polynomring über  $R$ .

*Polynomdivision*

Seien  $a, b \in R[X]$ . Der Leitkoeffizient von  $b$  sei eine Einheit in  $R$ . Dann gibt es eindeutig bestimmte  $q, r \in R[X]$  mit

$$a = qb + r, \quad \deg(r) < \deg(b). \quad (\text{Polynomdivision mit Rest})$$

*Abspalten von Nullstellen*

Ist  $p \in R[X]$  und  $w \in R$  eine Nullstelle von  $p$ , so gibt es genau ein  $q \in R[X]$  mit

$$p = (X - w)q. \quad (\text{Abspaltung einer Nullstelle})$$

Allgemeiner gibt es ein  $k$ -Tupel  $(w_1, \dots, w_k) \in R^k$ ,  $k \leq \deg(p)$ , mit

$$p = (X - w_1)(X - w_2) \dots (X - w_k) q, \quad (\text{vollständige Nullstellenabspaltung})$$

sodass  $q$  ein nullstellenfreies Polynom vom Grad  $n - k$  ist. Ist  $k = n$ , so gilt

$$p = p_n(X - w_1)(X - w_2) \dots (X - w_n). \quad (\text{Zerlegung in Linearfaktoren})$$

Zur Division von  $a = a_0$  durch  $b$  finden wir Polynome  $q_i \in R[X]$  mit

$$a_0 = q_1 b + a_1, \quad \deg(a_0) > \deg(a_1),$$

$$a_1 = q_2 b + a_2, \quad \deg(a_1) > \deg(a_2),$$

...

$$a_{m-1} = q_m b + a_m,$$

$$\deg(a_{m-1}) \geq \deg(b) > \deg(a_m).$$

Dann gilt wie gewünscht

$$a_0 = q_1 b + a_1 = (q_1 + q_2)b + a_2 =$$

$$\dots = (q_1 + \dots + q_m)b + a_m.$$

Die Polynome  $q_i, a_i$  sind definiert durch

$$q_i = \frac{\text{Leitkoeffizient}(a_{i-1})}{\text{Leitkoeffizient}(b)} X^{\deg(a_{i-1}) - \deg(b)},$$

$$a_i = a_{i-1} - q_i b \quad \text{für alle } 1 \leq i \leq n.$$

*Zur Eindeutigkeit:* Ist  $a = \bar{q}b + \bar{r}$  mit  $\deg(\bar{r}) < \deg(b)$ , so gilt  $(q - \bar{q})b = \bar{r} - r$  und damit  $\text{grad}(q - \bar{q}) + \text{grad}(b) = \text{grad}(\bar{r} - r) < \text{grad}(b)$ . Also ist  $\text{grad}(q - \bar{q}) = -\infty$ , sodass  $q = \bar{q}$ ,  $r = \bar{r}$ .

Polynomdivision in  $\mathbb{Z}[X]$  für

$$a = 3X^4 - 4X^3 + 7X^2 - 11X + 5,$$

$$b = X^3 + 2X - 1.$$

$X^4$	$X^3$	$X^2$	$X^1$	1	
3	-4	7	-11	5	$a = a_0$
3	0	6	-3	0	$3Xb$
	-4	1	-8	5	$a_1$
	-4	0	-8	4	$-4b$
		1	0	1	$a_2$

$$a = (3X - 4)b + (X^2 + 1)$$

*Zur Abspaltung von Nullstellen:* Ist  $w$  eine Nullstelle von  $p$ , so gilt  $p = q(X - w) + r$  mit  $\deg(r) < \deg(X - w) = 1$ , sodass  $r = a$  für ein  $a \in R$ . Wegen

$$0 = f_p(w) = f_q(w)(w - w) + f_r(w) = 0 + a = a$$

gilt  $r = 0$ . Dies zeigt, dass sich eine Nullstelle abspalten lässt. Wiederholtes Abspalten von Nullstellen liefert die restlichen Aussagen.

### Beispiele

- (1) Das Polynom  $X^2 - 2 \in \mathbb{Q}[X]$  hat aufgrund der Irrationalität von  $\sqrt{2}$  keine Nullstelle. Als Polynom in  $\mathbb{R}[X]$  hat  $X^2 - 2$  die Nullstellen  $\sqrt{2}$  und  $-\sqrt{2}$ , sodass  $X^2 - 2 = (X - \sqrt{2})(X + \sqrt{2})$ .
- (2) Das Polynom  $X^2 + 1 \in \mathbb{R}[X]$  hat keine Nullstelle. Als Polynom in  $\mathbb{C}[X]$  hat  $X^2 + 1$  die Nullstellen  $i$  und  $-i$ , sodass  $X^2 + 1 = (X - i)(X + i)$ .
- (3) Ist  $p \in \mathbb{R}[X]$  ein Polynom ungeraden Grades, so hat  $p$  eine Nullstelle. Denn aufgrund der ungeraden höchsten Potenz gibt es  $a, b \in \mathbb{R}$ , sodass  $f_p(a) > 0$  und  $f_p(b) < 0$ . Nach dem Zwischenwertsatz der Analysis hat also  $f_p$  eine Nullstelle.

In der vollständigen Nullstellenabspaltung müssen die  $w_1, \dots, w_k$  nicht paarweise verschieden sein. Für eine Nullstelle  $w$  von  $p$  gibt die *algebraische Vielfachheit*

$$\mu_p(w) = \max(\{m \geq 1 \mid \text{es gibt ein } q \in R[X] \text{ mit } p = (X - w)^m q\})$$

an, wie oft der Faktor  $(X - w)$  in der vollständigen Nullstellenabspaltung erscheint.

Es gilt der für Algebra und Analysis gleichermaßen unentbehrliche

### Satz (Fundamentalsatz der Algebra)

■ Jedes Polynom  $p \in \mathbb{C}[X]$  zerfällt in Linearfaktoren.

Zerfällt für einen Körper  $K$  jedes Polynom  $p \in K[X]$  in Linearfaktoren, so heißt  $K$  *algebraisch abgeschlossen*. Im Gegensatz zu  $\mathbb{C}$  sind  $\mathbb{Q}$  und  $\mathbb{R}$  nicht algebraisch abgeschlossen. Der Körper  $\mathbb{A}$  der algebraischen Zahlen ist algebraisch abgeschlossen. Denn man kann zeigen, dass jedes Polynom  $p \in \mathbb{A}[X]$  nur algebraische Nullstellen besitzt.

Der Fundamentalsatz liefert auch eine wertvolle Erkenntnis für reelle Polynome. Ist nämlich  $p \in \mathbb{C}[X]$  ein Polynom mit Koeffizienten in  $\mathbb{R} \subseteq \mathbb{C}$ , so ist mit  $w$  auch die Konjugierte  $\bar{w} = \operatorname{Re}(w) - i \operatorname{Im}(w)$  von  $w$  eine Nullstelle von  $p$ . Nun hat

$$(X - w)(X - \bar{w}) = X^2 - (w + \bar{w})X + w\bar{w} = X^2 - 2\operatorname{Re}(w)X + |w|^2$$

reelle Koeffizienten. Durch eine derartige Zusammenfassung von Paaren erhält man:

### Satz (Zerlegung eines reellen Polynoms)

Jedes Polynom  $p \in \mathbb{R}[X]$  vom Grad  $n \geq 0$  lässt sich in der Form

$$p = a_n (X - w_1) \dots (X - w_k) q_1 \dots q_{(n-k)/2}$$

schreiben, mit nullstellenfreien Polynomen  $q_j$  zweiten Grades der Form

$$q_j = X^2 - 2b_jX + c_j, \quad c_j > 0.$$





# 3. Kapitel

---

## Vektorräume

---

### 3.1 Vektorräume

**Definition** (*K-Vektorraum, Vektor, Skalar*)

Seien  $(V, +)$  eine abelsche Gruppe,  $K$  ein Körper und  $\cdot : K \times V \rightarrow V$ . Dann heißt  $(V, +, \cdot)$  oder kurz  $V$  ein *K-Vektorraum*, falls für alle  $\alpha, \beta \in K$  und  $v, w \in V$  gilt:

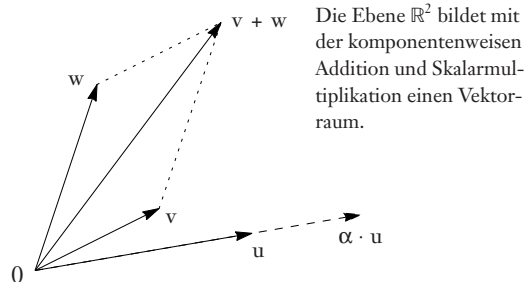
- (a)  $1 \cdot v = v$ ,
- (b)  $\alpha \cdot (\beta \cdot v) = (\alpha \beta) \cdot v$ ,
- (c)  $\alpha \cdot (v + w) = (\alpha \cdot v) + (\alpha \cdot w)$ ,
- (d)  $(\alpha + \beta) \cdot v = (\alpha \cdot v) + (\beta \cdot v)$ . (Axiome für die Skalarmultiplikation)

Die Elemente von  $V$  nennen wir *Vektoren* und die Elemente von  $K$  *Skalare*. Der Körper  $K$  heißt der *Skalarenkörper* von  $V$ . Die Abbildung  $+$  :  $V^2 \rightarrow V$  heißt die *Vektoraddition* und  $\cdot : K \times V \rightarrow V$  die *Skalarmultiplikation* von  $V$ .

In einem Vektorraum sind also vier Abbildungen vorhanden: eine Addition in  $V$ , eine Addition und Multiplikation in  $K$  sowie eine Skalarmultiplikation  $\cdot$ , die es erlaubt, einen Vektor  $v \in V$  mit einem Skalar  $\alpha \in K$  zu „skalieren“, sodass ein Vektor  $w = \alpha \cdot v$  entsteht.

Die geforderten Eigenschaften (die sog. *Vektorraumaxiome*) umfassen zehn Körper-, vier Gruppenaxiome und die Axiome (a) bis (d) für die Skalarmultiplikation. Diese 18 Axiome lassen sich kurz so zusammenfassen: Auf einer abelschen Gruppe  $(V, +)$  ist eine Skalarmultiplikation mit guten Recheneigenschaften erklärt.

Wir können den skalaren Malpunkt weglassen und nach (b) zum Beispiel  $\alpha \beta v$  schreiben. Es ist ungefährlich, den Skalar 0 und den *Nullvektor* 0 in  $V$  gleich zu bezeichnen. Gleiches gilt für die Additionen bzw. Subtraktionen in  $K$  und  $V$ . Möglich sind diese Vereinfachungen, weil für alle Skalare  $\alpha$  und Vektoren  $v$  gilt:



$\alpha v = 0$ genau dann, wenn $\alpha = 0$ oder $v = 0$
$(-\alpha) v = \alpha(-v) = -\alpha v$

Für Skalare stehen die vier Grundrechenarten zur Verfügung, von einem Produkt von Vektoren  $v$  und  $w$  ist in den Vektorraumaxiomen dagegen nicht die Rede. Vermutlich aus der Schule bekannt sind das Skalarprodukt  $\circ : V \times V \rightarrow \mathbb{R}$  für  $V = \mathbb{R}^2$  oder  $V = \mathbb{R}^3$  sowie das Kreuzprodukt  $\times : \mathbb{R}^3 \times \mathbb{R}^3 \rightarrow \mathbb{R}^3$ . (Statt  $v \circ w$  sind auch  $v \bullet w$  oder  $\langle v, w \rangle$  üblich.) Vor allem das Skalarprodukt wird später eine wichtige Rolle spielen (vgl. Kapitel 6).

Für Skalare werden griechische Buchstaben wie  $\alpha, \beta, \lambda, \mu, \dots$  verwendet. Dadurch können Dekorationen wie Pfeile oder Striche über den Vektoren entfallen.

### Beispiele

- (1) Jeder Körper  $K$  ist ein  $K$ -Vektorraum. Vektoren und Skalare sind in diesem Fall identisch. Speziell ist  $\mathbb{R}$  ein  $\mathbb{R}$ -Vektorraum und  $\mathbb{C}$  ein  $\mathbb{C}$ -Vektorraum.
- (2) Ist  $K$  ein Körper und  $n \in \mathbb{N}$ , so ist  $K^n$  mit
 
$$(a_1, \dots, a_n) + (b_1, \dots, b_n) = (a_1 + b_1, \dots, a_n + b_n),$$

$$\alpha (a_1, \dots, a_n) = (\alpha a_1, \dots, \alpha a_n) \text{ für alle } \alpha \in K, (a_1, \dots, a_n), (b_1, \dots, b_n) \in K^n$$
 ein  $K$ -Vektorraum (wobei  $K^0 = \{0\}$ ). Speziell gilt dies für  $\mathbb{R}^n$  und  $\mathbb{C}^n$ .
- (3)  $\mathbb{C}$  ist ein  $\mathbb{R}$ -Vektorraum: Die Vektoren sind komplexe Zahlen, die reellen Zahlen dienen als Skalare. Analog ist  $\mathbb{R}$  ein  $\mathbb{Q}$ -Vektorraum: Die Vektoren sind reelle Zahlen, als Skalare sind nur rationale Zahlen zugelassen. Allgemein gilt: Sind  $L, K$  Körper mit  $L \subseteq K$ , so ist  $K$  ein  $L$ -Vektorraum.
- (4) Sei  $(V, +, \cdot)$  ein  $\mathbb{C}$ -Vektorraum. Wir definieren  $\alpha * v = \overline{\alpha} v$  für  $\alpha \in \mathbb{C}$  und  $v \in V$ . Dann ist  $(V, +, *)$  ein  $\mathbb{C}$ -Vektorraum.
- (5) Sei  $M$  eine nichtleere Menge. Wir definieren  $1 \cdot A = A$ ,  $0 \cdot A = \emptyset$  für alle  $A \subseteq M$ . Dann ist  $(\mathcal{P}(M), \Delta, \cdot)$  ein  $K$ -Vektorraum über dem Körper  $K = \{0, 1\}$ . Die Vektoren sind Teilmengen von  $M$  und die Skalare stets 0 oder 1.

Auch viele Ringe führen zu Vektorräumen. Jeder Ring  $R$  eignet sich als Menge von Vektoren, mit der Ringaddition als Vektoraddition. Ist nun  $K \subseteq R$  ein Körper, so liefert die Ringmultiplikation  $\cdot : R \times R \rightarrow R$  durch Einschränkung auf  $K \times R$  eine Skalarmultiplikation  $\cdot : K \times R \rightarrow R$ , die (a) – (d) erfüllt. Der Ring  $R$  wird so zu einem  $K$ -Vektorraum. Allerdings enthält nicht jeder Ring einen Körper (so etwa der Restklassenring  $\mathbb{Z}_4$ ). Ein wichtiges Beispiel ist jedoch:

### Der Polynomring $K[X]$ als $K$ -Vektorraum

Sei  $K[X]$  der Polynomring über einem Körper  $K$ . Dann bilden die konstanten Polynome (also die Polynome vom Grad kleinergleich 0) einen Körper, den wir mit  $K$  identifizieren können.  $K[X]$  ist damit ein  $K$ -Vektorraum.

Schließlich betrachten wir noch einige Gegenbeispiele.

### Beispiele zu den Axiomen für die Skalarmultiplikation

- (1) Definieren wir auf  $\mathbb{R}$  die Skalarmultiplikation durch  $\alpha \cdot x = 0$ , so gelten (b), (c), (d), aber (a) ist verletzt.
- (2) Definieren wir auf  $\mathbb{C}$  die Skalarmultiplikation durch  $\alpha \cdot z = \operatorname{Re}(\alpha)z$ , so gelten (a), (c), (d), aber (b) ist verletzt.
- (3) Definieren wir auf  $\mathbb{C}^2$  die Skalarmultiplikation durch  $\alpha \cdot (z_1, z_2) = (\alpha z_1, \alpha z_2)$  für  $z_2 \neq 0$  und  $\alpha \cdot (z_1, 0) = (\overline{\alpha} z_1, 0)$ , so gelten (a), (b), (d), aber (c) ist verletzt.
- (4) Definieren wir auf  $\mathbb{R}$  die Skalarmultiplikation durch  $\alpha \cdot x = x$ , so gelten (a), (b), (c), aber (d) ist verletzt.

### 3.2 Unterräume

#### Definition (Unterraum)

Sei  $V$  ein  $K$ -Vektorraum, und sei  $U \subseteq V$ . Dann heißt  $U$  ein *Unterraum* oder *Untervektorraum* von  $V$ , falls gilt:

- (a)  $U$  ist eine Untergruppe der abelschen Gruppe  $V$ .
- (b)  $U$  ist abgeschlossen unter der Skalarmultiplikation, d. h., für alle  $\alpha \in K$  und  $u \in U$  gilt  $\alpha u \in U$ .

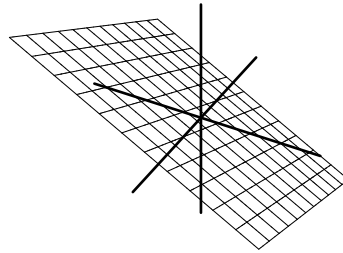
Unterräume verhalten sich zu Vektorräumen so wie Untergruppen zu Gruppen. Die Definition besagt:

*$U$  ist mit der von  $V$  ererbten Vektoraddition und Skalarmultiplikation ein  $K$ -Vektorraum.*

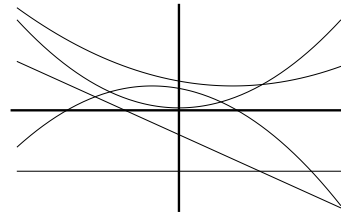
Die Gültigkeit der Axiome für die Skalarmultiplikation müssen wir nicht fordern. Sie überträgt sich von  $V$  auf jede Teilmenge  $U$  von  $V$ .

Wie für Untergruppen sehen wir je nach Kontext einen Unterraum als Teilmenge eines Vektorraumes oder als vollständigen Vektorraum an.

Das Analogon zum Untergruppenkriterium ist:



Eine Ebene durch 0 ist ein Unterraum des  $\mathbb{R}^3$ .



Die Polynome vom Grad kleinergleich 2 bilden einen Unterraum des  $\mathbb{R}[X]$ . Im Diagramm identifizieren wir sie mit Polynomfunktionen.

#### Unterraumkriterium

$U \subseteq V$  ist genau dann ein Unterraum von  $V$ , falls gilt:

- (U1)  $U \neq \emptyset$ .
- (U2) Für alle  $u, w \in U$  gilt  $u + w \in U$ .
- (U3) Für alle  $\alpha \in K$  und  $u \in U$  gilt  $\alpha u \in U$ .

Die Aussagen (U1) und (U2) gelten, falls  $U$  eine Untergruppe von  $V$  ist, und (U3) ist genau die Aussage (b) der Definition. Sind (U1) – (U3) erfüllt, so gilt für alle  $u, w \in U$ , dass

$$u - w = u + (-w) = u + (-1)w \in U,$$

sodass das Untergruppenkriterium anwendbar ist und Teil (a) der Definition liefert.

## Beispiele

- (1) Für jeden Vektorraum  $V$  sind  $\{0\}$  und  $V$  Unterräume von  $V$ .

- (2) Sei  $V = \mathbb{R}^3$ . Dann sind

$$U = \{(x_1, 0, 0) \in V \mid x_1 \in \mathbb{R}\} \text{ und } W = \{(x_1, x_2, 0) \in V \mid x_1, x_2 \in \mathbb{R}\}$$

Unterräume von  $V$ . Allgemeiner bildet jede Gerade und jede Ebene in  $V$  durch den Nullpunkt einen Unterraum von  $V$ . Geraden und Ebenen, die nicht durch den Nullpunkt verlaufen, bilden dagegen keine Unterräume (vgl. auch 3.12).

- (3) Allgemeiner als (2): Seien  $n \geq 1$  und  $I \subseteq \{1, \dots, n\}$ . Dann ist

$$U = \{(x_1, \dots, x_n) \in \mathbb{R}^n \mid x_i = 0 \text{ für alle } i \in I\}$$

ein Unterraum von  $\mathbb{R}^n$ .

- (4)  $\mathbb{Q}$  ist eine Untergruppe von  $(\mathbb{R}, +)$ , aber kein Unterraum des  $\mathbb{R}$ -Vektorraumes  $\mathbb{R}$ . Die Abgeschlossenheit unter Skalarmultiplikation ist verletzt: Ist  $\alpha$  irrational, so ist  $\alpha \cdot 1 = \alpha \notin \mathbb{Q}$ . Dagegen ist  $\mathbb{Q}$  ein Unterraum des  $\mathbb{Q}$ -Vektorraumes  $\mathbb{R}$ , bei dem nur rationale Zahlen als Skalare für Vektoren (reelle Zahlen) auftauchen.

- (5) Für alle  $n$  ist  $U_n = \{v \in K[X] \mid \deg(v) \leq n\}$  ein Unterraum des  $K$ -Vektorraums  $K[X]$  aller Polynome über  $K$ .

- (6) Sind  $U$  und  $W$  Unterräume von  $V$ , so ist auch der Durchschnitt  $U \cap W$  ein Unterraum von  $V$ . Allgemeiner gilt: Ist  $(U_i)_{i \in I}$  eine Familie von Unterräumen von  $V$ , so ist auch

$$U = \bigcap_{i \in I} U_i = \{v \in V \mid v \in U_i \text{ für alle } i \in I\}$$

ein Unterraum von  $V$ .

- (7) Sind  $U$  und  $W$  Unterräume von  $V$ , so ist  $U \cup W$  im Allgemeinen nicht abgeschlossen unter der Vektoraddition und damit kein Unterraum von  $V$ . Sind zum Beispiel  $U$  und  $W$  zwei verschiedene Geraden der Ebene durch 0, so ist  $U \cup W$  keine Untergruppe von  $(\mathbb{R}^2, +)$ . Denn sind  $u \in U$  und  $w \in W$  beide ungleich dem Nullvektor, so ist  $u + w$  kein Element von  $U \cup W$ .

- (8) Ist  $(U_i)_{i \in I}$  eine Familie von Unterräumen von  $V$  und gilt die Vergleichbarkeit

$$U_i \subseteq U_j \text{ oder } U_j \subseteq U_i \text{ für alle } i, j \in I,$$

so ist auch

$$\bigcup_{i \in I} U_i = \{v \in V \mid \text{es gibt ein } i \in I \text{ mit } v \in U_i\}$$

ein Unterraum von  $V$ . Ist  $I$  endlich, so ist die Vereinigung einfach gleich dem größten Element der durch die  $U_i$  gebildeten  $\subseteq$ -Kette. Es gibt aber auch Beispiele für unendliche Ketten, so etwa  $(U_n)_{n \in \mathbb{N}}$  wie in Beispiel (5). Hier ist

$$\bigcup_{n \in \mathbb{N}} U_n = K[X].$$

### 3.3 Produkte von Vektorräumen

#### Definition (Produkte und Potenzen)

##### Endliche Produkte

Seien  $V_1, \dots, V_n$ ,  $n \geq 1$ ,  $K$ -Vektorräume, und sei

$$W = V_1 \times \dots \times V_n = \{(v_1, \dots, v_n) \mid v_i \in V_i \text{ für alle } 1 \leq i \leq n\}.$$

Für alle  $v = (v_1, \dots, v_n)$ ,  $w = (w_1, \dots, w_n) \in W$  und  $\alpha \in K$  setzen wir:

$$v + w = (v_1 + w_1, \dots, v_n + w_n),$$

$$\alpha \cdot v = (\alpha v_1, \dots, \alpha v_n).$$

$(W, +, \cdot)$  heißt das *Produkt* der Vektorräume  $V_1, \dots, V_n$ .

##### Familien-Produkte

Sei  $(V_i)_{i \in I}$  eine Familie von  $K$ -Vektorräumen, und sei

$$W = \prod_{i \in I} V_i = \{(v_i)_{i \in I} \mid v_i \in V_i \text{ für alle } i \in I\}.$$

Für alle  $v = (v_i)_{i \in I}$ ,  $w = (w_i)_{i \in I} \in W$  und  $\alpha \in K$  setzen wir

$$v + w = (v_i + w_i)_{i \in I}, \quad \alpha \cdot v = (\alpha v_i)_{i \in I}.$$

$(W, +, \cdot)$  heißt das *Produkt* der Vektorräume  $(V_i)_{i \in I}$ .

##### Potenzen

Sind in einem Produkt alle  $V_i$  gleich einem Vektorraum  $V$ , so schreiben wir auch

$$V^n \text{ statt } V_1 \times \dots \times V_n,$$

$$V^I \text{ statt } \prod_{i \in I} V_i.$$

$V^n$  bzw.  $V^I$  heißt die  $n$ - bzw.  $I$ -fache *Potenz* von  $V$ . Weiter setzen wir

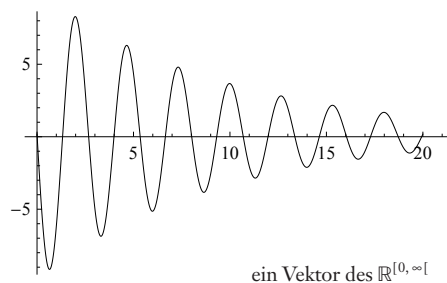
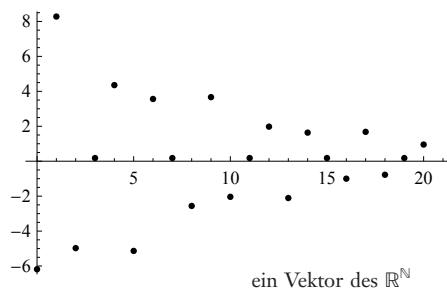
$$V^0 = \{0\}.$$

Ist  $K$  ein Körper, so schreiben wir  $K^n$  und  $K^I$  für die Potenzen des  $K$ -Vektorraumes  $K$ .

Kurz: Auf den Produkten

$$V_1 \times \dots \times V_n, \quad \prod_{i \in I} V_i$$

wird eine *komponenten-* oder *punktweise* Vektoraddition und Skalarmultiplikation erklärt. Die Produkte werden dadurch zu  $K$ -Vektorräumen.



**Bemerkung 1**

In der Produktbildung ist es wichtig, dass alle beteiligten Vektorräume denselben Skalarenkörper  $K$  besitzen.  $K$  hängt nicht vom Index  $i$  ab.

**Bemerkung 2**

Ist  $I = \{1, \dots, n\}$ , so können wir das Produkt  $V_1 \times \dots \times V_n$  mit  $\prod_{i \in I} V_i$  und die Potenz  $V^n$  mit der Potenz  $V^{\{1, \dots, n\}}$  identifizieren. Dadurch werden die endlichen Produkte zu Spezialfällen der allgemeinen Produkte. Dies gilt auch für  $n = 0$  und  $I = \emptyset$ , wenn wir  $0$  mit der leeren Menge identifizieren.

Explizit wollen wir noch einmal die Körperpotenzen

$$K^n = \{(x_1, \dots, x_n) \mid x_i \in K \text{ für alle } 1 \leq i \leq n\},$$

$$K^I = \{(x_i)_{i \in I} \mid x_i \in K \text{ für alle } i \in I\} = \{f \mid f: I \rightarrow K\}$$

notieren. Sie spielen eine fundamentale Rolle in der Linearen Algebra, und wir werden ihnen noch oft begegnen. Am wichtigsten sind hier die Körper  $K = \mathbb{R}$  und  $K = \mathbb{C}$ .

**Beispiele**

- (1)  $\mathbb{R} \times \mathbb{R} \times \mathbb{R} = \mathbb{R}^3$ ,  $V \times V \times V \times V = V^4$  usw.
- (2) Die Vektoren von  $V = \mathbb{R}^2 \times \mathbb{R}^3$  haben die Form  $((x_1, x_2), (y_1, y_2, y_3))$ . Identifizieren wir diese Vektoren mit  $(x_1, x_2, y_1, y_2, y_3)$ , so wird  $V$  zum  $\mathbb{R}^5$ .
- (3) Beispiele für Potenzen  $V^I$  sind  $\mathbb{R}^{\{0, 2, 4\}}$ ,  $\mathbb{R}^{\mathbb{N}}$ ,  $\mathbb{C}^{\mathbb{N}}$ ,  $\mathbb{R}^{\mathbb{R}}$  und  $\mathbb{C}^{\mathbb{C}}$ .
- (4) Die Vektoren des  $\mathbb{R}^{\mathbb{N}}$  sind die reellen Folgen  $(x_n)_{n \in \mathbb{N}}$ . Die Addition und Skalarmultiplikation auf dem  $\mathbb{R}^{\mathbb{N}}$  sind wie in der Analysis erklärt durch
 
$$(x_n)_{n \in \mathbb{N}} + (y_n)_{n \in \mathbb{N}} = (x_n + y_n)_{n \in \mathbb{N}} \quad \text{und} \quad \alpha(x_n)_{n \in \mathbb{N}} = (\alpha x_n)_{n \in \mathbb{N}}.$$
 Analoges gilt für  $\mathbb{C}^{\mathbb{N}}$ .
- (5) Für  $I \subseteq \mathbb{R}$  sind die Vektoren des  $\mathbb{R}^I$  reellwertige Funktionen der Form  $f: I \rightarrow \mathbb{R}$ . Die Addition und Skalarmultiplikation fällt erneut mit den analytischen Operationen  $f + g$  und  $\alpha f$  zusammen. Typische Fälle sind  $I = [0, 1]$  und  $I = \mathbb{R}$ . Analoges gilt für  $\mathbb{C}^I$  mit  $I \subseteq \mathbb{C}$ , etwa  $I = [0, 1]^2$ ,  $I = \{z \in \mathbb{C} \mid |z| = 1\}$  oder  $I = \mathbb{C}$ .

Der Polynomring  $K[X] = K^{(\mathbb{N})}$  ist ein Unterraum der Potenz  $K^{\mathbb{N}}$ . Er besteht aus allen Folgen  $(x_n)_{n \in \mathbb{N}}$  in  $K$ , deren Träger  $\{n \mid x_n \neq 0\}$  endlich ist. Allgemein definieren wir:

**Definition** (die Vektorräume  $V^{(I)}$ )

Für jeden  $K$ -Vektorraum  $V$  und jede Menge  $I$  sei

$$V^{(I)} = \{(v_i)_{i \in I} \in V^I \mid \{i \in I \mid v_i \neq 0\} \text{ ist endlich}\}.$$

Der Vektorraum  $V^{(I)}$  ist ein Unterraum des  $V^I$ . Ist  $I$  endlich, so gilt  $V^{(I)} = V^I$ . Andernfalls ist  $V^{(I)}$  eine echte Teilmenge des  $V^I$ .

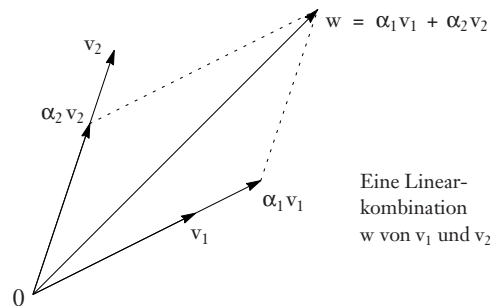
### 3.4 Linearkombinationen und Erzeugendensysteme

**Definition** (*Linearkombination, Spann, Erzeugendensystem*)

Sei  $V$  ein  $K$ -Vektorraum.

- (a) Ein  $w \in V$  heißt eine *Linearkombination* von  $v_1, \dots, v_n \in V$ , falls  $\alpha_1, \dots, \alpha_n \in K$  existieren mit  $w = \alpha_1 v_1 + \dots + \alpha_n v_n$ .
- (b) Für  $A \subseteq V$  definieren wir den *Spann* von  $A$  durch
 
$$\text{span}(A) = \langle A \rangle = \{w \in V \mid w \text{ ist eine Linearkombination von Vektoren in } A\}.$$
 Für eine Familie  $(v_i)_{i \in I}$  in  $V$  sei  $\text{span}((v_i)_{i \in I}) = \text{span}(\{v_i \mid i \in I\})$ .
- (c) Ein  $A \subseteq V$  heißt *erzeugend* oder ein *Erzeugendensystem* von  $V$ , falls  $\text{span}(A) = V$ . Eine Familie  $(v_i)_{i \in I}$  in  $V$  heißt *erzeugend*, falls  $\{v_i \mid i \in I\}$  erzeugend ist.

Linearkombinationen verallgemeinern die Summen  $v_1 + \dots + v_n$ , die in allen Gruppen erklärt sind. In einem Vektorraum können die Summanden „skaliert“ oder „gewichtet“ werden.



Eine Linearkombination  $w$  von  $v_1$  und  $v_2$

**Beispiele**

- (1) Im  $\mathbb{R}^3$  ist
 
$$(2, 3, 3) = 2(1, 1, 1) + 1(0, 1, 1)$$
 eine Linearkombination der Vektoren  $(1, 1, 1), (0, 1, 1)$ .
- (2) Da die leere Summe gleich 0 ist, gilt  $\text{span}(\emptyset) = \{0\}$ . Weiter ist  $\text{span}(0) = \{0\}$ ,  $\text{span}(v) = \{\alpha v \mid \alpha \in K\}$ ,  $\text{span}(v, w) = \{\alpha v + \beta w \mid \alpha, \beta \in K\}$ .
- (3) Ist  $v \in \mathbb{R}^3$ ,  $v \neq 0$ , so ist  $\text{span}(v)$  eine Gerade durch den Nullpunkt. Ist dann  $w \in \mathbb{R}^3$  ein Vektor mit  $w \notin \text{span}(v)$ , so ist  $\text{span}(v, w)$  eine Ebene durch den Nullpunkt.
- (4) Es gilt  $\text{span}(v_1, \dots, v_n) = \text{span}(-v_1, \dots, -v_n)$ , da für alle  $\alpha_1, \dots, \alpha_n$  in  $K$  gilt, dass
 
$$\alpha_1 v_1 + \dots + \alpha_n v_n = (-\alpha_1)(-v_1) + \dots + (-\alpha_n)(-v_n).$$
- (5) Es gilt  $\text{span}(v_1, v_2, v_1 + v_2) = \text{span}(v_1, v_2)$ , da für alle  $\alpha_1, \alpha_2, \alpha_3$  gilt, dass
 
$$\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 (v_1 + v_2) = (\alpha_1 + \alpha_3)v_1 + (\alpha_2 + \alpha_3)v_2.$$

Eine nützliche Notation für Familien ist:

**Die Summen  $\sum_{i \in I} \alpha_i v_i$** 

Für alle Familien  $(v_i)_{i \in I}$  in  $V$  und alle  $(\alpha_i)_{i \in I} \in K^{(I)}$  sei  $\sum_{i \in I} \alpha_i v_i = \sum_{i \in I, \alpha_i \neq 0} \alpha_i v_i$ .

In  $\sum_{i \in I} \alpha_i v_i$  ist die Menge aller  $i$  mit  $\alpha_i \neq 0$  stets endlich, sodass sich die Summe auf eine endliche Summe reduziert. Es gilt  $\text{span}((v_i)_{i \in I}) = \{\sum_{i \in I} \alpha_i v_i \mid (\alpha_i)_{i \in I} \in K^{(I)}\}$ .



### Eigenschaften des Spans

Für alle  $A \subseteq V$  ist  $\text{span}(A)$  der kleinste Unterraum  $U$  von  $V$  mit  $U \supseteq A$ .

Ist  $A \subseteq B$ , so gilt  $\text{span}(A) \subseteq \text{span}(B)$ .

Ist  $B \subseteq \text{span}(A)$  und  $A \subseteq \text{span}(B)$ , so ist  $\text{span}(A) = \text{span}(B)$ .

Es gilt  $\text{span}(V) = V$ , sodass jeder Vektorraum ein Erzeugendensystem besitzt. Für weitere Beispiele definieren wir:

**Definition** (die Standardvektoren  $e_i$ )

(a) Sei  $V = K^n$  für ein  $n \geq 1$ . Dann definieren wir

$$e_1 = (1, 0, \dots, 0), e_2 = (0, 1, 0, \dots, 0), \dots, e_n = (0, \dots, 0, 1).$$

(b) Sei  $V = K[X] = K^{(\mathbb{N})}$ . Dann definieren wir  $e_n \in V$  für alle  $n \in \mathbb{N}$  durch

$$e_n = X^n = (0, \dots, 0, 1, 0, 0, 0, \dots) \text{ mit } n \text{ Nullen vor der } 1.$$

(c) Sei  $V = K^{(I)}$ . Dann definieren wir  $e_i \in V$  für alle  $i \in I$  durch  $e_i(i) = 1$  und  $e_i(j) = 0$  für alle  $i \neq j$ .

### Beispiele

(1) Sei  $V = K^n$  mit  $n \geq 1$ . Dann ist  $\{e_1, \dots, e_n\}$  erzeugend. Gleiches gilt für

$$v_1 = (1, 0, \dots, 0), v_2 = (1, 1, 0, \dots, 0), v_3 = (1, 1, 1, 0, \dots, 0), \dots, v_n = (1, \dots, 1).$$

(2) Sei  $V = K[X] = K^{(\mathbb{N})}$ . Dann ist  $\{e_n \mid n \in \mathbb{N}\}$  erzeugend. Weiter gilt

$$\text{span}(e_0, \dots, e_n) = \{w \in K[X] \mid \text{grad}(w) \leq n\}.$$

(3) In  $V = K^{(I)}$  ist  $\{e_i \mid i \in I\}$  erzeugend. Ist  $I$  unendlich, so ist  $K^{(I)} \neq K^I$  und  $\{e_i \mid i \in I\}$  nicht erzeugend in  $K^I$ . Zum Beispiel liegt  $v = (1)_{i \in I}$  nicht im Spann der  $e_i$ , falls  $I$  unendlich ist (da Linearkombinationen stets endliche Summen sind).

Die Begriffsbildungen sind auch außerhalb der Linearen Algebra von Bedeutung:

### Exkurs: Trigonometrische Polynome

Sei  $V = \{f : \mathbb{R} \rightarrow \mathbb{R} \mid f(x) = f(x + 2\pi) \text{ für alle } x \in \mathbb{R}\}$  der  $\mathbb{R}$ -Vektorraum der  $2\pi$ -periodischen Funktionen.  $V$  ist ein Unterraum des  $\mathbb{R}^{\mathbb{R}}$ . Wir betrachten

$$v_k = \text{„die Funktion } g \text{ auf } \mathbb{R} \text{ mit } g(x) = \cos(kx) \text{ für alle } x\text{“ für } k \in \mathbb{N},$$

$$w_k = \text{„die Funktion } g \text{ auf } \mathbb{R} \text{ mit } g(x) = \sin(kx) \text{ für alle } x\text{“ für } k \in \mathbb{N} - \{0\}.$$

Für jedes  $n$  heißen die Linearkombinationen von  $v_0, v_1, w_1, \dots, v_n, w_n$  die *trigonometrischen Polynome* vom Grad kleinergleich  $n$ . Sie lassen sich schreiben als

$$f(x) = a_0 + \sum_{1 \leq k \leq n} (a_k \cos(kx) + b_k \sin(kx)), \quad x \in \mathbb{R}, \text{ mit } a_k, b_k \in \mathbb{R},$$

und werden in der Analysis zur Approximation von Funktionen in  $V$  verwendet.

### 3.5 Lineare Unabhängigkeit

**Definition** (*linear unabhängig, linear abhängig*)

Sei  $V$  ein  $K$ -Vektorraum.

- (a) Ein Tupel  $(v_1, \dots, v_n)$  von Vektoren in  $V$  heißt *linear unabhängig*, falls für alle  $\alpha_1, \dots, \alpha_n \in K$  gilt:

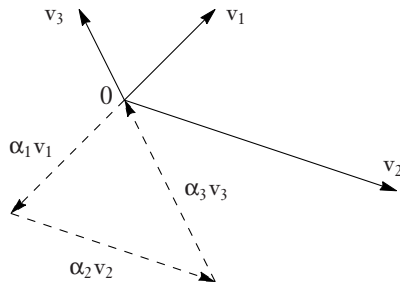
$$\alpha_1 v_1 + \dots + \alpha_n v_n = 0 \text{ impliziert } \alpha_1 = \dots = \alpha_n = 0. \\ \text{(eindeutige Nulldarstellung, Nullbedingung)}$$

- (b) Eine Menge  $A \subseteq V$  heißt *linear unabhängig*, falls jedes Tupel von paarweise verschiedenen Vektoren in  $A$  linear unabhängig ist. Eine Familie  $(v_i)_{i \in I}$  in  $V$  heißt *linear unabhängig*, falls  $(v_{i_1}, \dots, v_{i_n})$  für alle paarweise verschiedenen  $i_1, \dots, i_n \in I$ ,  $n \geq 1$ , linear unabhängig ist.

Andernfalls heißt  $(v_1, \dots, v_n)$  bzw.  $A$  bzw.  $(v_i)_{i \in I}$  *linear abhängig*.

Ein Tupel  $(v_1, \dots, v_n)$  ist also linear unabhängig, wenn der Nullvektor nur trivial als Linearkombination dargestellt werden kann. Somit ist  $(v_1, \dots, v_n)$  linear abhängig, wenn es  $\alpha_1, \dots, \alpha_n \in K$  gibt mit

- (a)  $0 = \alpha_1 v_1 + \dots + \alpha_n v_n$ ,  
(b)  $\alpha_i \neq 0$  für mindestens ein  $i$ .



$(v_1, v_2, v_3)$  sind linear abhängig, da eine nicht-triviale Darstellung der 0 existiert.

#### Formulierungen der linearen Unabhängigkeit

Für alle  $A \subseteq V$  sind äquivalent:

- (a)  $A$  ist linear unabhängig.  
(b) Für alle  $v \in A$  ist  $v \notin \text{span}(A - \{v\})$ . (Spannbedingung)

Für jede Familie  $(v_i)_{i \in I}$  in  $V$  sind äquivalent:

- (a)  $(v_i)_{i \in I}$  ist linear unabhängig.  
(b) Für alle  $(\alpha_i)_{i \in I} \in K^{(I)}$  gilt:  
 $\sum_{i \in I} \alpha_i v_i = 0 \text{ impliziert } \alpha_i = 0 \text{ für alle } i \in I$ .  
(c) Für alle  $(\alpha_i)_{i \in I}, (\beta_i)_{i \in I} \in K^{(I)}$  gilt:  
 $\sum_{i \in I} \alpha_i v_i = \sum_{i \in I} \beta_i v_i \text{ impliziert } \alpha_i = \beta_i \text{ für alle } i \in I$ .

(Eindeutigkeit der Darstellung als Linearkombination)

Die Spannbedingung lässt sich besonders griffig formulieren:

„Kein Vektor von  $A$  liegt im Spann der anderen.“

„Verkleinern wir  $A$ , so verkleinern wir den Spann.“

In Familien-Schreibweise lautet die Spannbedingung:  $v_j \notin \text{span}((v_i)_{i \in I, i \neq j})$  für alle  $j \in I$ . Zur Überprüfung der linearen Unabhängigkeit ist der Nachweis der Nullbedingung aber oft einfacher als der Nachweis der Spannbedingung.

Wir betrachten nun „versteckte Details“ der Begriffsbildung und erste Beispiele.

### Formale Feinheiten

- (1) Die leere Menge ist linear unabhängig. Die Nullbedingung ist leer.
- (2)  $(0)$  ist linear abhängig, da  $0 = 1 \cdot 0$  eine nichttriviale Darstellung der 0 ist. Auch die Spannbedingung zeigt dies, da  $0 \in \{0\} = \text{span}(\emptyset) = \text{span}(\{0\} - \{0\})$ . Allgemein ist jedes  $A \subseteq V$  mit  $0 \in A$  linear abhängig.
- (3) Ist  $v \neq 0$ , so ist  $(v)$  linear unabhängig, da aus  $0 = \alpha v$  folgt, dass  $\alpha = 0$ .
- (4) Ist  $v_1 = v_n$  und  $n \geq 1$ , so ist  $(v_1, \dots, v_n)$  linear abhängig, da  $0 = 1 v_1 - 1 v_n$ . Ist  $v = w \neq 0$ , so ist  $(v, w)$  linear abhängig, aber  $\{v, w\} = \{w\}$  ist linear unabhängig.

### Beispiele

- (1) Für Vektoren  $v_1 = (x_1, y_1, z_1)$ ,  $v_2 = (x_2, y_2, z_2)$ ,  $v_3 = (x_3, y_3, z_3) \in \mathbb{R}^3$  ist  $(v_1, v_2, v_3)$  genau dann linear unabhängig, wenn für alle  $\alpha_1, \alpha_2, \alpha_3 \in \mathbb{R}$  gilt:

$$\alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3 = 0 \text{ impliziert } \alpha_1 = \alpha_2 = \alpha_3 = 0.$$

Dies ist gleichwertig dazu, dass das lineare Gleichungssystem

$$\alpha_1 x_1 + \alpha_2 x_2 + \alpha_3 x_3 = 0$$

$$\alpha_1 y_1 + \alpha_2 y_2 + \alpha_3 y_3 = 0$$

$$\alpha_1 z_1 + \alpha_2 z_2 + \alpha_3 z_3 = 0$$

in den reellen Unbestimmten  $\alpha_1, \alpha_2, \alpha_3$  nur die Lösung  $\alpha_1 = \alpha_2 = \alpha_3 = 0$  besitzt. Analoges gilt für  $n$  Vektoren  $v_1, \dots, v_n$  des  $\mathbb{R}^n$ ,  $n \geq 1$ .

- (2) Seien  $V = \mathbb{R}^{\mathbb{R}}$ ,  $v = \sin$  und  $w = \cos$ . Dann ist  $(v, w)$  linear unabhängig. Denn sei  $\alpha \sin + \beta \cos = 0$ . Dann gilt  $\alpha \sin(x) + \beta \cos(x) = 0$  für alle  $x \in \mathbb{R}$ , speziell also

$$\alpha \sin(0) + \beta \cos(0) = 0, \quad \alpha \sin(\pi/2) + \beta \cos(\pi/2) = 0.$$

Aus  $\sin(0) = \cos(\pi/2) = 0$  und  $\sin(\pi/2) = \cos(0) = 1$  folgt nun  $\alpha = \beta = 0$ .

Allgemeiner ist die Menge der Sinus- und Kosinusfunktionen aus dem Exkurs im letzten Abschnitt linear unabhängig.

- (3) Sei  $V = K^{(\mathbb{I})}$ . Dann ist  $\{e_i \mid i \in \mathbb{I}\}$  linear unabhängig. Speziell ist die Menge  $\{1, X, X^2, X^3, \dots\} = \{e_0, e_1, e_2, \dots\}$  der Monome linear unabhängig im  $K$ -Vektorraum  $K[X] = K^{(\mathbb{N})}$ .

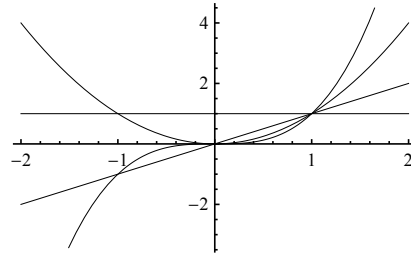
### 3.6 Basen und Koordinatenvektoren

#### Definition (Basis)

Sei  $V$  ein  $K$ -Vektorraum.

- (a) Ein Tupel  $(v_1, \dots, v_n)$  von Vektoren in  $V$  heißt eine *Basis* von  $V$ , wenn  $(v_1, \dots, v_n)$  linear unabhängig und erzeugend ist.
- (b) Ebenso heißt eine Menge  $B \subseteq V$  bzw. eine Familie  $(v_i)_{i \in I}$  in  $V$  eine *Basis* von  $V$ , wenn sie linear unabhängig und erzeugend ist.

Die Vektoren einer Basis nennen wir auch *Basisvektoren*. Ist eine Basis  $B$  endlich, so heißt die Anzahl  $|B|$  ihrer Elemente die *Länge* von  $B$ .



„Basis“ bedeutet:

*Jeder Vektor liegt im Spann von  $B$  und jeder Vektor in  $B$  wird für diese Eigenschaft wirklich gebraucht.*

$(1, X, X^2, X^3)$  ist eine Basis des Vektorraums  $V \subseteq \mathbb{R}[X]$  aller Polynome über  $\mathbb{R}$  vom Grad kleiner gleich 3. Im Diagramm sind die Polynomfunktionen dargestellt.

#### Äquivalenzen für endliche Basen

Für jedes Tupel  $(v_1, \dots, v_n)$  in  $V$  sind äquivalent:

- (a)  $(v_1, \dots, v_n)$  ist eine Basis von  $V$ .
- (b)  $(v_1, \dots, v_n)$  ist linear unabhängig und für alle  $v \in V$  ist  $(v_1, \dots, v_n, v)$  linear abhängig. *(maximal linear unabhängig)*
- (c)  $(v_1, \dots, v_n)$  ist erzeugend und für alle  $i$  ist  $(v_1, \dots, v_{i-1}, v_{i+1}, \dots, v_n)$  nicht erzeugend. *(minimal erzeugend)*
- (d) Jeder Vektor  $v$  in  $V$  besitzt eine eindeutige Darstellung der Form  $v = \alpha_1 v_1 + \dots + \alpha_n v_n$  mit  $\alpha_1, \dots, \alpha_n \in K$ .  
*(Existenz und Eindeutigkeit der Darstellung als Linearkombination)*

Analoge Äquivalenzen lassen sich auch für beliebige Mengen und Familien angeben. Für Mengen  $B \subseteq V$  lautet (b):

$B$  ist ein  $\subseteq$ -maximales Element von  $\{A \subseteq V \mid A \text{ ist linear unabhängig}\}$ .

Für Familien  $(v_i)_{i \in I}$  in  $V$  lautet (d):

Für jeden Vektor  $v$  in  $V$  existiert eine eindeutige Darstellung der Form  $v = \sum_{i \in I} \alpha_i v_i$ .

Die Eindeutigkeit erlaubt folgende fundamentale Definition:

**Definition** (Koordinatenabbildung  $\Phi_B$ , Koordinatenvektor)

Sei  $V$  ein  $K$ -Vektorraum.

- (a) Sei  $B = (v_1, \dots, v_n)$  eine Basis. Dann definieren wir  $\Phi_B : V \rightarrow K^n$  durch

$$\Phi_B(v) = v_B = \text{„der Vektor } (\alpha_1, \dots, \alpha_n) \in K^n \text{ mit } v = \alpha_1 v_1 + \dots + \alpha_n v_n\text{“}.$$

- (b) Sei  $B = (v_i)_{i \in I}$  eine Basis. Dann definieren wir  $\Phi_B : V \rightarrow K^{(I)}$  durch

$$\Phi_B(v) = v_B = \text{„der Vektor } (\alpha_i)_{i \in I} \in K^{(I)} \text{ mit } v = \sum_{i \in I} \alpha_i v_i\text{“}.$$

Wir nennen  $\Phi_B(v) = v_B$  den *Koordinatenvektor* von  $v$  bzgl.  $B$ . Für alle  $i$  heißt der Skalar  $\alpha_i = v_B(i)$  der  $v_i$ -*Anteil* von  $v$  bzgl.  $B$ .

Ein Koordinatenvektor ist also ein Element eines Vektorraumes  $K^n$  oder allgemeiner des  $K^{(I)}$ . Ist  $V = K^n$  bzw.  $V = K^{(I)}$ , so gehören  $v$  und  $v_B$  demselben Vektorraum an. Im Allgemeinen leben sie in verschiedenen Räumen. Die Reihenfolge oder Indizierung der Basisvektoren spielt für Koordinatenvektoren eine Rolle.

**Beispiele**

- (1) Die Standardvektoren  $e_1 = (1, 0, 0)$ ,  $e_2 = (0, 1, 0)$ ,  $e_3 = (0, 0, 1)$  bilden eine Basis  $B = (e_1, e_2, e_3)$  des  $\mathbb{R}^3$ . Für alle  $(x, y, z) \in \mathbb{R}^3$  gilt

$$(x, y, z) = (x, 0, 0) + (0, y, 0) + (0, 0, z) = x e_1 + y e_2 + z e_3.$$

Damit ist  $v = v_B$  für alle  $v \in \mathbb{R}^3$ . Man nennt  $B$  die *kanonische Basis* oder die *Standardbasis* des  $\mathbb{R}^3$ . Analoges gilt für  $\mathbb{R}^n$  und  $\mathbb{C}^n$  für alle  $n$ .

- (2) Die Vektoren  $v_1 = (0, 0, 1)$ ,  $v_2 = (0, 1, 1)$ ,  $v_3 = (1, 1, 1)$  bilden ebenfalls eine Basis  $C = (v_1, v_2, v_3)$  des  $\mathbb{R}^3$ . Es gilt zum Beispiel

$$(1, 2, 3) = 1 \cdot v_1 + 1 \cdot v_2 + 1 \cdot v_3,$$

sodass  $(1, 2, 3)_C = (1, 1, 1)$ . Ebenso ist, mit der kanonischen Basis  $(e_1, e_2, e_3)$ ,

$$(0, 0, 1)_C = e_1, \quad (0, 1, 1)_C = e_2, \quad (1, 1, 1)_C = e_3.$$

- (3) Sei  $V = K[X] = K^{(\mathbb{N})}$ . Dann ist  $B = (e_n)_{n \in \mathbb{N}}$  eine Basis von  $V$ . Allgemein ist  $(e_i)_{i \in I}$  eine Basis des  $K^{(I)}$ . Für alle  $v$  gilt  $v = \sum_{i \in I} v(i) e_i$ , sodass  $v_B = v$ . Wir nennen  $B$  wieder die *kanonische Basis* oder *Standardbasis* des Vektorraums  $K^{(I)}$ .

- (4)  $e_1 = (1, 0) = 1$  und  $e_2 = (0, 1) = i$  bilden eine Basis  $B = (e_1, e_2)$  des  $\mathbb{R}$ -Vektorraums  $\mathbb{C} = \mathbb{R}^2$ . Dagegen ist  $(e_1)$  eine Basis des  $\mathbb{C}$ -Vektorraums  $\mathbb{C}$ , da sich jedes  $v \in \mathbb{C}$  eindeutig als  $v = v e_1$  schreiben lässt.

- (5) Eine Basis von  $\mathbb{R}^{\mathbb{N}}$  oder des  $\mathbb{Q}$ -Vektorraums  $\mathbb{R}$  ist nicht zu sehen (vgl. 3.9).

**Warnung: Bestimmten Artikel vermeiden**

Von Anfängern hört man oft: „ $(e_1, e_2, e_3)$  ist die Basis des  $\mathbb{R}^3$ .“ Der Wunsch nach Eindeutigkeit ist verständlich, aber die Aussage ist analog zu: „Die Katze ist das Tier.“ Also bitte „ $(e_1, e_2, e_3)$  ist eine Basis des  $\mathbb{R}^3$ “, so wie „Die Katze ist ein Tier.“

### 3.7 Austauschlemma und Austauschsatz

**Satz** (*Austauschlemma und Austauschsatz von Ernst Steinitz*)

Sei  $V$  ein  $K$ -Vektorraum, und sei  $(v_1, \dots, v_n)$  eine Basis von  $V$ . Dann gilt:

*Austauschlemma*

Ist  $v \in V$  und  $1 \leq i \leq n$  derart, dass in der Linearkombination

$$v = \alpha_1 v_1 + \dots + \alpha_n v_n$$

der Skalar  $\alpha_i$  von 0 verschieden ist, so ist auch

$$v_1, \dots, v_{i-1}, v, v_{i+1}, \dots, v_n$$

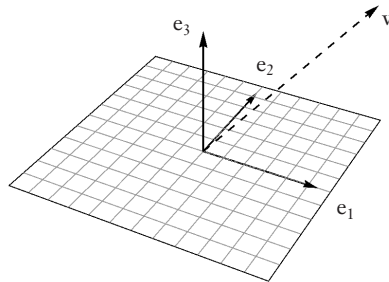
eine Basis von  $V$ .

*Austauschsatz von Ernst Steinitz*

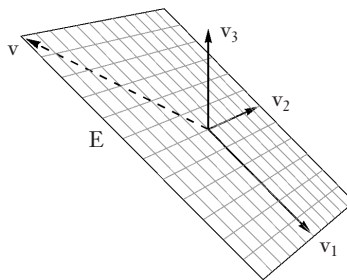
Ist  $(w_1, \dots, w_k)$  linear unabhängig in  $V$ , so ist  $k \leq n$  und es gibt  $n - k$  Vektoren unter den Basisvektoren  $v_1, \dots, v_n$ , sodass diese Vektoren zusammen mit  $(w_1, \dots, w_k)$  eine Basis von  $V$  bilden.

Das Austauschlemma besagt, dass man einen Vektor  $v_i$  einer Basis  $B$  gegen einen Vektor  $v$  austauschen darf, wenn  $v$  einen nichttrivialen  $v_i$ -Anteil bzgl.  $B$  aufweist, wenn also die  $i$ -te Komponente des Koordinatenvektors  $v_B$  von  $v$  von 0 verschieden ist.

Das Austauschlemma dient als Grundlage für einen Beweis des Austauschsatzes (vgl. Beispiel 3). Dieser besagt, dass man ein linear unabhängiges  $k$ -Tupel in eine Basis der Länge  $n$  integrieren kann, indem man gewisse Basisvektoren durch die Vektoren des Tupels ersetzt (anders formuliert: das Tupel mit Basisvektoren zu einer Basis erweitert). Ein wichtiger Bestandteil der Aussage des Austauschsatzes ist, dass  $k$  kleinergleich  $n$  ist. Es kann also nicht mehr linear unabhängige Vektoren als Elemente in einer Basis geben. Das ist zwar glaubhaft, aber keineswegs klar.



In der Basis  $(e_1, e_2, e_3)$  des  $\mathbb{R}^3$  kann jeder Basisvektor gegen  $v = e_1 + e_2 + e_3$  ausgetauscht werden:  $(v, e_2, e_3), (e_1, v, e_3), (e_1, e_2, v)$  sind Basen des  $\mathbb{R}^3$ .



In der Basis  $(v_1, v_2, v_3)$  des  $\mathbb{R}^3$  können  $v_1$  und  $v_2$  gegen  $v$  ausgetauscht werden, nicht aber  $v_3$ . Der Vektor  $v$  liegt in der von  $v_1$  und  $v_2$  aufgespannten Ebene  $E$  und hat damit den  $v_3$ -Anteil 0.

**Beispiele**

- (1) Sei  $(e_1, e_2, e_3)$  die Standardbasis des  $\mathbb{R}^3$ . Dann ist für alle  $v = (x, y, z) \in \mathbb{R}^3$  mit  $z \neq 0$  auch  $(e_1, e_2, v)$  eine Basis des  $\mathbb{R}^3$ . Der Vektor  $e_3$  von  $B$  lässt sich also durch jeden Vektor mit einem Höhenanteil ungleich 0 ersetzen.
- (2) Die „ $\alpha_i \neq 0$ “-Bedingung im Austauschlemma ist auch notwendig dafür, dass der Austausch „ $v_i$  gegen  $v$ “ eine Basis hinterlässt. Zum Beispiel können wir in der Basis  $((1, 0, 0), (1, 1, 0), (1, 1, 1))$  des  $\mathbb{R}^3$  den dritten Vektor nicht gegen einen Vektor  $(x_1, x_2, 0)$  austauschen, ohne die Basiseigenschaft zu zerstören.
- (3) Wir betrachten die Standardbasis  $(e_1, e_2, e_3, e_4)$  des  $\mathbb{R}^4$  und das linear unabhängige Paar  $(w_1, w_2)$  mit  $w_1 = (1, 2, 0, 0)$  und  $w_2 = (3, 6, 2, -1)$ . Es gilt
- $$w_1 = (1, 2, 0, 0) = 1 e_1 + 2 e_2 + 0 e_3 + 0 e_4.$$
- Nach dem Austauschlemma ist  $(w_1, e_2, e_3, e_4)$  eine Basis. Nun gilt
- $$w_2 = (3, 6, 2, -1) = 3 w_1 + 0 e_2 + 2 e_3 - e_4,$$
- sodass nach dem Austauschlemma  $(w_1, e_2, w_2, e_4)$  eine Basis ist (nicht aber  $(w_1, w_2, e_3, e_4)$ ). Damit haben wir  $(w_1, w_2)$  in die Basis  $(e_1, e_2, e_3, e_4)$  integriert.
- (4) Sei  $V$  ein Vektorraum derart, dass für jedes  $k \geq 1$  ein linear unabhängiges Tupel  $(w_1, \dots, w_k)$  existiert. Dann besitzt  $V$  keine endliche Basis. Denn wäre  $(v_1, \dots, v_n)$  eine Basis, so wären nach dem Austauschsatz  $k \leq n$  für alle  $k$ , was nicht sein kann.

Zwei einfache, aber wichtige Folgerungen aus dem Austauschsatz sind:

Für jeden Vektorraum $V$ , der eine endliche Basis besitzt, gilt:
<p style="text-align: center;"><b>Längensatz</b></p> <p style="text-align: center;">Je zwei Basen <math>B_1</math> und <math>B_2</math> haben die gleiche Länge.</p>
<p style="text-align: center;"><b>Basisergänzungssatz</b></p> <p style="text-align: center;">Ist <math>A</math> linear unabhängig, so existiert eine Basis <math>B \supseteq A</math>.</p>

Sind nämlich  $(v_1, \dots, v_n)$  und  $(w_1, \dots, w_k)$  zwei Basen, so gilt  $k \leq n$  nach dem Austauschsatz, da  $(w_1, \dots, w_k)$  linear unabhängig ist. Analog gilt  $n \leq k$  und damit  $k = n$ . Dies zeigt den Längensatz. Der Basisergänzungssatz ist eine Abschwächung des Austauschsatzes.

Die Ergebnisse sind auch für Vektorräume, die keine endliche Basis besitzen, richtig, wobei im Längensatz „gleiche Mächtigkeit“ an die Stelle von „gleiche Länge“ tritt. Zum Beweis müssen dann allerdings andere Methoden verwendet werden (vgl. hierzu auch Abschnitt 3.9).

### 3.8 Die Dimension

#### Definition (Dimension eines Vektorraumes)

Ein Vektorraum  $V$  heißt *endlich-dimensional*, in Zeichen  $\dim(V) < \infty$ , falls eine endliche Basis von  $V$  existiert. Andernfalls heißt  $V$  *unendlich-dimensional*, in Zeichen  $\dim(V) = \infty$ . Ist  $V$  endlich-dimensional und  $(v_1, \dots, v_n)$  eine Basis von  $V$ , so heißt  $V$  *n-dimensional*, in Zeichen  $\dim(V) = n$ .

Die Unterscheidung zwischen „ $\dim(V) < \infty$ “ und „ $\dim(V) = \infty$ “ ist einfach möglich, die Setzung von „ $\dim(V) = n$ “ beruht dagegen auf dem Längensatz in 3.7.

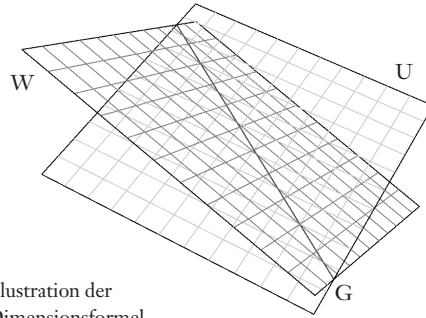


Illustration der  
Dimensionsformel

$\dim(U) + \dim(W) = \dim(U \cap W) + \dim(\text{span}(U \cup W))$   
anhand von Ebenen  $U$  und  $W$  im  $\mathbb{R}^3$  durch  $0$ , deren Durchschnitt eine Gerade  $G$  ist:  $2 + 2 = 1 + 3$ .

#### Bemerkung

Ist  $V$  *endlich erzeugt*, d. h. gibt es  $v_1, \dots, v_n \in V$  mit

$$\text{span}(v_1, \dots, v_n) = V,$$

so ist  $V$  endlich-dimensional.

Denn wir können einen

Vektor  $v_i$ , der im Spann der Vektoren  $v_j$ ,  $j \neq i$ , liegt, streichen, ohne den Spann  $V$  zu verkleinern. So lässt sich  $(v_1, \dots, v_n)$  schrittweise zu einer Basis reduzieren.

#### Beispiele

(1) Ist  $V = \{0\}$ , so gilt  $\dim(V) = 0$ . Denn die leere Menge ist eine Basis von  $V$ .

(2) Ist  $V = \mathbb{R}^n$ , so gilt  $\dim(V) = n$ , denn  $(e_1, \dots, e_n)$  ist eine Basis von  $V$ . Analog ist  $\dim(V) = n$  für  $\mathbb{C}^n$ .

(3) Ist  $V$  der  $\mathbb{R}$ -Vektorraum  $\mathbb{C}^n$ , so gilt  $\dim(V) = 2n$ . Eine Basis ist

$$((1, 0, \dots, 0), (i, 0, \dots, 0), \dots, (0, \dots, 0, 1), (0, \dots, 0, i)).$$

(4) Sei  $M = \{a_1, \dots, a_n\}$  eine nichtleere Menge mit genau  $n$  Elementen, und sei  $V$  der  $\{0, 1\}$ -Vektorraum  $\mathcal{P}(M)$  mit der symmetrischen Differenz als Vektoraddition und der Skalarmultiplikation  $0 \cdot A = \emptyset$  und  $1 \cdot A = A$  für alle  $A \subseteq M$ . Dann gilt  $\dim(V) = n$ , denn  $(\{a_1\}, \dots, \{a_n\})$  ist eine Basis von  $V$ .

(5) Ein Produktraum  $K^I$  ist genau dann endlich-dimensional, wenn  $I$  endlich ist.

(6) Der Vektorraum  $K[X] = K^{(\mathbb{N})}$  ist unendlich-dimensional.

(7) Der  $\mathbb{Q}$ -Vektorraum  $\mathbb{R}$  ist unendlich-dimensional. Denn für alle  $v_1, \dots, v_n \in \mathbb{R}$  ist

$$\text{span}(v_1, \dots, v_n) = \{\alpha_1 v_1 + \dots + \alpha_n v_n \mid \alpha_1, \dots, \alpha_n \in \mathbb{Q}\}$$

abzählbar und damit ungleich  $\mathbb{R}$ . Allgemein gilt, dass ein überabzählbarer Vektorraum  $V$  über einem abzählbaren Körper unendlich-dimensional ist.



Hat man die Dimension eines endlich-dimensionalen Vektorraums  $V$  als  $n$  bestimmt, so ist der Nachweis, dass  $n$  Vektoren eine Basis bilden, nur noch halb so aufwendig. Es genügt zu zeigen, dass die Vektoren linear unabhängig oder erzeugend sind. Das „oder“ wird automatisch zum „und“:

Satz von der Halbierung der Arbeit
<p>Ist <math>\dim(V) = n</math>, so sind für alle <math>v_1, \dots, v_n \in V</math> äquivalent:</p> <ul style="list-style-type: none"> <li>(a) <math>(v_1, \dots, v_n)</math> ist eine Basis von <math>V</math>.</li> <li>(b) <math>(v_1, \dots, v_n)</math> ist linear unabhängig.</li> <li>(c) <math>(v_1, \dots, v_n)</math> ist erzeugend.</li> </ul>

Wir betrachten schließlich noch Unterräume von endlich-dimensionalen Vektorräumen. Ist  $(v_1, \dots, v_n)$  eine Basis, so existieren Unterräume  $U_0, \dots, U_n$  der Dimensionen  $0, \dots, n$ :

$$U_0 = \{0\} = \text{span}(\emptyset), \quad U_1 = \text{span}(v_1), \quad U_2 = \text{span}(v_1, v_2), \quad \dots, \quad U_n = V = \text{span}(v_1, \dots, v_n).$$

Wichtige Ergebnisse über die Dimension von Unterräumen sind:

Dimension von Unterräumen
<p>Ist <math>V</math> endlich-dimensional und <math>U</math> ein Unterraum von <math>V</math>, so ist <math>U</math> endlich-dimensional und <math>\dim(U) \leq \dim(V)</math>. Ist <math>\dim(U) = \dim(V)</math>, so ist <math>U = V</math>.</p>
<p>Sind <math>U, W</math> Unterräume von <math>V</math>, so gilt die <i>Dimensionsformel</i>:</p> $\dim(U) + \dim(W) = \dim(U \cap W) + \dim(\text{span}(U \cup W)).$

Der Leser vergleiche die Dimensionsformel für Unterräume mit der Anzahlformel für endliche Mengen  $A, B$ :  $|A| + |B| = |A \cap B| + |A \cup B|$ .

### Beispiele

- (1) Sei  $V = \mathbb{R}^3$ , und seien  $U$  und  $W$  zwei verschiedene Ebenen durch den Nullpunkt. Dann ist  $\text{span}(U \cup W) = V$ , denn ist  $(u_1, u_2)$  eine Basis von  $U$ , so ist  $(u_1, u_2, v)$  für alle  $v \in W - U$  eine Basis von  $V$ . Nach der Dimensionsformel ist
 
$$\dim(U \cap W) = \dim(U) + \dim(W) - \dim(\text{span}(U \cup W)) = 2 + 2 - 3 = 1.$$
 Damit ist  $U \cap W$  eine Gerade durch den Nullpunkt.
- (2) Sind  $U, W$  Unterräume eines Vektorraums  $V$  mit  $U \cap W = \{0\}$ , so ist
 
$$\dim(\text{span}(U \cup W)) = \dim(U) + \dim(W) - \dim(\{0\}) = \dim(U) + \dim(W).$$

### 3.9 Die Existenz von Basen

**Satz** (*allgemeiner Basisexistenz- und Basisergänzungssatz*)

Jeder Vektorraum  $V$  besitzt eine Basis. Genauer gilt: Ist  $A_0 \subseteq V$  linear unabhängig, so existiert eine Basis  $B$  von  $V$  mit  $B \supseteq A_0$ .

Unter den unendlich-dimensionalen Vektorräumen konnten wir bislang nur für die Vektorräume der Form  $K^{(\mathbb{I})}$  eine Basis angeben. Der Satz besagt nun, dass jeder Vektorraum eine Basis besitzt, und stärker, dass der Basisergänzungssatz für jeden Vektorraum gilt. Speziell gibt es Basen des  $\mathbb{R}^{\mathbb{N}}$ , des  $\mathbb{R}^{\mathbb{R}}$  und des  $\mathbb{Q}$ -Vektorraums  $\mathbb{R}$ .

Der Basisexistenzsatz nimmt eine besondere Stellung in der Linearen Algebra ein: Er lässt sich nur mit Hilfe des Auswahlaxioms beweisen (und ist sogar äquivalent zu diesem Axiom, vgl. 1.11). Die natürliche Frage „Hat jeder Vektorraum eine Basis?“ ist überraschenderweise mit den Grundlagen der Mathematik verknüpft.

Der Beweis des Satzes wird üblicherweise mit Hilfe des Zornschen Lemmas geführt (vgl. 1.12). Abgesehen von diesem abstrakten Hilfsmittel ist der Beweis überraschend kurz. Er ist zudem typisch für andere Anwendungen des Zornschen Lemmas, wie sie in der Algebra und Funktionalanalysis auftauchen. Sei

$$\mathcal{B} = \{A \subseteq V \mid A \text{ ist linear unabhängig}\}.$$

Ist  $\mathcal{A}$  eine bzgl. der Inklusion linear geordnete Teilmenge von  $\mathcal{B}$ , d.h., gilt  $A_1 \subseteq A_2$  oder  $A_2 \subseteq A_1$  für alle  $A_1, A_2 \in \mathcal{A}$ , so ist die Menge

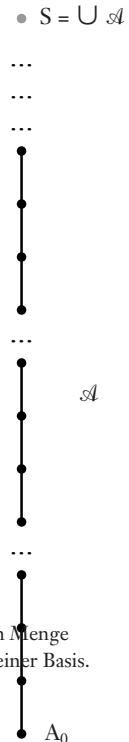
$$\bigcup \mathcal{A} = \{v \in V \mid \text{es gibt ein } A \in \mathcal{A} \text{ mit } v \in A\}$$

linear unabhängig, also ein Element von  $\mathcal{B}$ . Das Zornsche Lemma liefert nun die Existenz eines  $\subseteq$ -maximalen Elements  $B \in \mathcal{B}$ . Nach Konstruktion ist  $B$  eine linear unabhängige Menge von Vektoren, die sich nicht mehr vergrößern lässt, ohne die lineare Unabhängigkeit zu zerstören. Damit ist  $B$  eine Basis von  $V$ . Der allgemeine Basisergänzungssatz wird genauso bewiesen, wobei man nun mit dem folgenden Mengensystem arbeitet:

$$\mathcal{B}_{A_0} = \{A \subseteq V \mid A \text{ ist linear unabhängig und } A_0 \subseteq A\}.$$

Die Vektorraumtheorie bleibt unvollständig, wenn die Frage der Existenz von Basen nicht angesprochen wird. Andererseits muss ein Anfänger hier auch nicht zu tief einsteigen. Wir beenden diese Sektion mit zwei Exkursen, die sich an interessierte Leser wenden, die mehr wissen wollen.

Das durch die Inklusion  $\subseteq$  partiell geordnete System aller linear unabhängigen Teilmengen von  $V$  erfüllt die Kettenbedingung, denn die Vereinigung einer Kette linear unabhängiger Mengen ist linear unabhängig. Damit ist das Zornsche Lemma anwendbar. Im Diagramm ist  $\mathcal{A}$  eine Kette und  $S = \bigcup \mathcal{A}$ .  $S$  ist eine obere Schranke von  $\mathcal{A}$ . Ketten können sehr lang sein, eine mit Hilfe der natürlichen Zahlen gebildete schrittweise Erweiterung einer linear unabhängigen Menge  $A_0$  führt in der Regel nicht zu einer Basis.



### Exkurs I: Hamel-Basen

Wir betrachten den  $\mathbb{Q}$ -Vektorraum  $\mathbb{R}$ . Der Skalarenkörper ist hier „künstlich“ auf die rationalen Zahlen beschränkt, die Vektoren sind dagegen beliebige reelle Zahlen. Eine Basis  $B$  dieses Vektorraums nennt man auch eine *Hamel-Basis*. Ist  $B \subseteq \mathbb{R}$  eine Hamel-Basis, so lässt sich jede reelle Zahl  $x$  eindeutig schreiben als

$$x = q_1 b_1 + \dots + q_n b_n$$

mit  $n \geq 0$  und  $q_i \in \mathbb{Q}^*$ ,  $b_i \in B$  für alle  $1 \leq i \leq n$ . Die explizite Angabe einer Hamel-Basis ist unmöglich, die Basis  $B$  bleibt abstrakt, nur das Auswahlaxiom garantiert die Existenz. Hamel-Basen erlauben jedoch bemerkenswerte Konstruktionen. Für ein Beispiel betrachten wir *additive Funktionen*  $f: \mathbb{R} \rightarrow \mathbb{R}$ , also Funktionen mit der Eigenschaft

$$f(x+y) = f(x) + f(y) \quad \text{für alle } x, y \in \mathbb{R}.$$

Jede Gerade  $g: \mathbb{R} \rightarrow \mathbb{R}$  durch den Nullpunkt ist additiv, und man kann zeigen, dass eine stetige additive Funktion eine Gerade durch den Nullpunkt ist. Mit Hilfe von Hamel-Basen lassen sich nun aber auch unstetige additive Funktionen konstruieren. Wir definieren hierzu  $f: \mathbb{R} \rightarrow \mathbb{R}$  durch

$$f(x) = q_1 + \dots + q_n \in \mathbb{Q}, \quad \text{mit } x = q_1 b_1 + \dots + q_n b_n \text{ wie oben.}$$

Die rationale Zahl  $f(x)$  ist also die Summe der Einträge des Koordinatenvektors  $v_B$  von  $x$  bzgl.  $B$ . Die Funktion  $f$  ist additiv. Aber es gilt  $\text{Bild}(f) = \mathbb{Q}$ , denn ist  $b \in B$  beliebig, so gilt  $f(qb) = q$  für alle  $q \in \mathbb{Q}$ . Nach dem Zwischenwertsatz ist  $f$  unstetig, denn eine stetige Funktion nimmt mit  $q < r$  auch alle Werte in  $[q, r]$  an.

### Exkurs II: Moduln

Ein Vektorraum ist mit einem Skalarenkörper  $K$  ausgestattet. Allgemeiner kann man statt eines Körpers einen Ring zugrunde legen. Die Axiome bleiben gleich. Statt von Vektorräumen spricht man dann von *Moduln*. („Modul“ wird auf dem „o“ betont, nicht auf dem „u“.) Ein Modul fühlt sich an wie ein Vektorraum, wir dürfen aber im Allgemeinen nicht mehr durch Skalare  $\alpha \neq 0$  dividieren. Bemerkenswerterweise ist der Basisexistenzsatz für Moduln nicht mehr gültig. Es gibt sogar endliche Moduln, die keine Basis besitzen. Ein Beispiel liefert das Rechnen in  $\mathbb{Z}$  modulo 4, also der Modul  $\mathbb{Z}_4 = \{ [0], [1], [2], [3] \}$  über dem Skalarenring  $\mathbb{Z}$ . Die Skalarmultiplikation wird wie üblich durch  $a[k] = [ak]$  für alle  $a \in \mathbb{Z}$  und  $[k] \in \mathbb{Z}_4$  erklärt. Dieser Modul hat keine Basis, denn für alle  $[k] \in \mathbb{Z}_4$  ist  $([k])$  linear abhängig, da  $0 = [0] = [4k] = 4[k]$  eine nichttriviale Darstellung der Null ist.

Interessant ist auch der Modul  $\mathbb{Z}$  über dem Skalarenring  $\mathbb{Z}$ . Hier gilt:

- (a)  $(1)$  ist eine Basis,
- (b)  $(2, 3)$  ist erzeugend (da  $a = a \cdot 3 - a \cdot 2$  für alle  $a \in \mathbb{Z}$ ),
- (c)  $(2, 3)$  ist linear abhängig (da  $3 \cdot 2 - 2 \cdot 3 = 0$ ),
- (d) es gibt keine Basis  $B \subseteq (2, 3)$  (da weder  $(2)$  noch  $(3)$  erzeugend ist).

### 3.10 Summen von Vektorräumen

#### Definition (äußere und innere Summen)

##### Äußere Summe

Sei  $(V_i)_{i \in I}$  eine Familie von Vektorräumen. Dann definieren wir die *äußere Summe*  $W$  der Vektorräume  $V_i$  durch

$$W = \{ f \in \prod_{i \in I} V_i \mid \text{supp}(f) \text{ ist endlich} \}, \text{ wobei}$$

$$\text{supp}(f) = \{ i \in I \mid f(i) \neq 0 \}$$

der *Träger* von  $f$  ist. In Zeichen schreiben wir

$$W = \bigoplus_{i \in I} V_i.$$

##### Innere Summe

Sei  $V$  ein Vektorraum, und seien  $W_1, \dots, W_n$  Unterräume von  $V$ . Dann setzen wir

$$W_1 + \dots + W_n = \{ w_1 + \dots + w_n \mid w_i \in W_i \text{ für alle } i \in I \}.$$

Allgemeiner definieren wir für eine Familie  $(W_i)_{i \in I}$  von Unterräumen von  $V$ :

$$\sum_{i \in I} W_i = \{ \sum_{j \in J} w_j \mid J \subseteq I \text{ ist endlich, } w_j \in W_j \text{ für alle } j \in J \}.$$

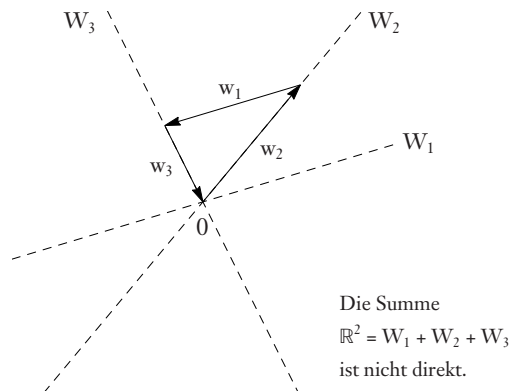
Die Unterräume  $W_1 + \dots + W_n$  bzw.  $\sum_{i \in I} W_i$  von  $V$  nennen wir die *innere Summe* der Unterräume  $W_i$ .

Eine innere Summe heißt *direkt*, falls jeder Vektor  $w_1 + \dots + w_n$  bzw.  $\sum_{j \in J} w_j$  der Definition der Summe nur dann gleich 0 ist, wenn alle Summanden  $w_j$  null sind. Wir schreiben dann

$$W = W_1 \oplus \dots \oplus W_n \text{ bzw.}$$

$$W = \bigoplus_{i \in I} W_i.$$

Die Summen lassen sich mit bekannten Konstruktionen erläutern:



##### Äußere Summen

Sei  $W = \bigoplus_{i \in I} V_i$  eine äußere Summe. Ist  $I$  endlich, so ist  $W = \prod_{i \in I} V_i$ , d. h., die äußere Summe ist dann einfach das endliche Produkt der  $V_i$ . Ist

$I$  unendlich, so ist  $W$  ein Unterraum von  $\prod_{i \in I} V_i$ . Der Unterraum  $W$  besteht aus allen Vektoren des Produkts, die an höchstens endlich vielen Stellen von 0 verschieden sind. Damit sind die Vektorräume  $V^{(I)}$  (vgl. 3.3) spezielle äußere Summen:

$$V^{(I)} = \bigoplus_{i \in I} V = \{ f \in V^I \mid \text{supp}(f) \text{ ist endlich} \}.$$

$$\text{Insbesondere ist } K[X] = K^{(\mathbb{N})} = \bigoplus_{n \in \mathbb{N}} K.$$

### Innere Summen

Die innere Summe kann man auch über den Spann erklären, denn

$$W_1 + \dots + W_n = \text{span}(W_1 \cup \dots \cup W_n), \quad \sum_{i \in I} W_i = \text{span}(\bigcup_{i \in I} W_i).$$

Die innere Summe der Unterräume  $W_i$  ist also der kleinste Unterraum von  $V$ , der alle Unterräume  $W_i$  umfasst.

### Direkte innere Summen

Die Direktheit einer inneren Summe  $W = W_1 + \dots + W_n$  lässt sich mit Hilfe des Begriffs der linearen Unabhängigkeit so formulieren:

Picken wir aus den Summanden  $W_i$  je einen von 0 verschiedenen Vektor  $w_i$  heraus, so ist  $(w_1, \dots, w_n)$  stets linear unabhängig.

Analog bedeutet die Direktheit für eine allgemeine Summe  $W = \bigoplus_{i \in I} W_i$ :

Picken wir aus endlich vielen Summanden  $W_j$ ,  $j \in J$ , je einen von 0 verschiedenen Vektor  $w_j$  heraus, so ist  $(w_j)_{j \in J}$  stets linear unabhängig in  $W$ .

### Verhältnis von äußeren und direkten inneren Summen

Ist  $W = W_1 \oplus \dots \oplus W_n$  eine direkte innere Summe und  $W^* = \bigoplus_{1 \leq i \leq n} W_i$  die äußere Summe der Vektorräume  $W_i$ , so haben wir die natürliche Korrespondenz

$$(w_1, \dots, w_n) \in W^* \quad \simeq \quad w_1 + \dots + w_n \in W.$$

Aufgrund der Direktheit von  $W$  liefert diese Entsprechung eine Bijektion

$$\varphi: W^* \rightarrow W, \quad \varphi(w_1, \dots, w_n) = w_1 + \dots + w_n \quad \text{für alle } (w_1, \dots, w_n) \in W^*.$$

(Genauer ist  $\varphi$  ein Vektorraum-Isomorphismus zwischen  $W$  und  $W^*$  im Sinne von 4.5.) Analoges gilt für allgemeine äußere und direkte innere Summen. Damit ist die doppelte Verwendung des Zeichens  $\oplus$  in der Regel harmlos, wenn man die Unterschiede der beiden Konstruktionen vor Augen hat.

### Beispiele

- (1) Sind  $W_1$  eine Gerade und  $W_2$  eine Ebene im  $\mathbb{R}^3$  durch den Nullpunkt mit  $W_1 \cap W_2 = \{0\}$ , so gilt  $\mathbb{R}^3 = W_1 \oplus W_2$ . Ebenso ist  $\mathbb{R}^3 = \{0\} \oplus \mathbb{R}^3$ .
- (2) Sind  $W_1, W_2, W_3 \subseteq \mathbb{R}^2$  paarweise verschiedene Geraden durch den Nullpunkt, so gilt  $W_i \cap W_j = \{0\}$  für alle  $i \neq j$ . Aber die Summe  $W_1 + W_2 + W_3 = \mathbb{R}^2$  ist nicht direkt.
- (3) Aus der Dimensionsformel in 3.8 folgt, dass eine endliche innere Summe  $W = W_1 + \dots + W_n$  in einem Vektorraum der endlichen Dimension  $m$  genau dann direkt ist, wenn  $\dim(W_1) + \dots + \dim(W_n) = m$ .

### 3.11 Quotientenräume

#### Definition (Quotientenraum)

Sei  $V$  ein  $K$ -Vektorraum, und sei  $U$  ein Unterraum von  $V$ . Dann definieren wir eine Äquivalenzrelation  $\sim$  auf  $V$  durch

$$v \sim w, \text{ falls } v - w \in U \quad \text{für alle } v, w \in V.$$

Auf der Faktorisierung  $V/U = \{ [v] \mid v \in V \} = \{ v/\sim \mid v \in V \}$  definieren wir

$$[v] + [w] = [v + w] \quad \text{für alle } v, w \in V,$$

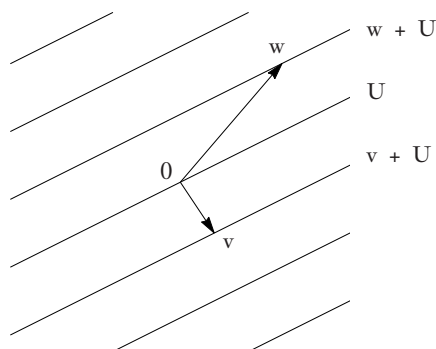
$$\alpha \cdot [v] = [\alpha v] \quad \text{für alle } \alpha \in K \text{ und } v \in V.$$

Der so entstehende Vektorraum  $V/U$  heißt der *Quotientenraum* von  $V$  modulo  $U$ .

Eine Äquivalenzklasse  $[v]$  nennen wir auch eine *Nebenklasse* von  $V$  bzgl.  $U$ .

Die Idee ist, die Vektoren in  $U$  als „unwesentlich“ zu betrachten und Vektoren  $v$  und  $w$  in  $V$  miteinander zu identifizieren, deren „Unterschied“  $v - w$  unwesentlich ist.

Die Relation  $\sim$  ist eine Äquivalenz auf  $V$  und die Abbildungen  $+$  und  $\cdot$  sind wohldefiniert. Durch sie wird  $V/U$  zu einem  $K$ -Vektorraum. Die Klassen  $[v]$  sind die Vektoren dieses Raums, die Skalare sind einfach die Skalare von  $V$ . Der Nullvektor des Quotientenraumes ist  $U$ . Für alle  $v \in V$  gilt



Für eine Gerade  $U$  durch  $0$  in  $\mathbb{R}^2$  besteht  $\mathbb{R}^2/U$  aus allen zu  $U$  parallelen Geraden  $[v] = v + U$ .

$$[v] = v + U, \text{ wobei } v + U = \{ v + u \mid u \in U \}.$$

Mit Blick auf die Faktorgruppen in 2.7 ist die Konstruktion nicht neu:  $(V/U, +)$  ist die Faktorgruppe der Gruppe  $(V, +)$  bzgl. der Untergruppe  $U$ . Da  $(V, +)$  abelsch ist, ist  $U$  ein Normalteiler. Im Unterschied zur reinen Gruppentheorie kann auf der Faktorgruppe  $V/U$  zudem eine Skalarmultiplikation erklärt werden, sodass  $V/U$  zu einem Vektorraum wird.

Eigenschaften der Nebenklassen	
$[0] = 0 + U = U = 0$	
$[u] = U$ für alle $u \in U$	
$\sum_{1 \leq i \leq n} \alpha_i [v_i] = [\sum_{1 \leq i \leq n} \alpha_i v_i]$ für alle $v_1, \dots, v_n \in V, \alpha_1, \dots, \alpha_n \in K$	

**Beispiele**

- (1) Für  $U = \{0\}$  ist  $[v] = \{v\}$  für alle  $v \in V$  und damit

$$V/U = \{[v] \mid v \in V\} = \{\{v\} \mid v \in V\}.$$

- (2) Für  $U = V$  ist  $[v] = V$  für alle  $v \in V$  und damit  $V/U = \{V\} = \{[0]\} = \{0\}$ .

- (3) Ist  $U$  eine Gerade durch den Nullpunkt in der Ebene  $V = \mathbb{R}^2$ , so ist eine Nebenklasse  $[v] = v + U$  eine zu  $U$  parallele Gerade. Der Quotientenraum  $V/U$  besteht aus allen zu  $U$  parallelen Geraden. Analoges gilt für Geraden oder Ebenen durch den Nullpunkt in  $\mathbb{R}^3$ .

- (4) Sei  $V$  der  $\mathbb{R}$ -Vektorraum aller (Riemann-)integrierbaren  $2\pi$ -periodischen Funktionen  $f: \mathbb{R} \rightarrow \mathbb{R}$  ( $V$  ist ein Unterraum des  $\mathbb{R}^{\mathbb{R}}$ ). Dann ist

$$U = \{f \in V \mid \int_0^{2\pi} |f(x)| dx = 0\}$$

ein Unterraum von  $V$ . Zwei Funktionen  $f, g \in V$  sind äquivalent modulo  $U$ , falls

$$\int_0^{2\pi} |f(x) - g(x)| dx = 0.$$

Der Quotientenraum  $V/U$  spielt in der Analysis in der Theorie der Fourier-Reihen eine Rolle. Allgemein werden Quotientenräume dieser Art in der Funktionalanalysis studiert.

Wir betrachten noch, wie sich Basen unter einer Faktorisierung  $V/U$  verhalten. Sei hierzu  $V$  ein endlich-dimensionaler  $K$ -Vektorraum, und sei  $U$  ein Unterraum von  $V$  mit  $\dim(U) = k$ . Weiter sei  $(u_1, \dots, u_k)$  eine Basis von  $U$  und  $B = (v_1, \dots, v_n, u_1, \dots, u_k)$  eine Basis von  $V$ . Dann gilt für alle Skalare  $\alpha_i$  und  $\beta_j$

$$\begin{aligned} [\sum_{1 \leq i \leq n} \alpha_i v_i + \sum_{1 \leq j \leq k} \beta_j u_j] &= [\sum_{1 \leq i \leq n} \alpha_i v_i] + [\sum_{1 \leq j \leq k} \beta_j u_j] = \\ \sum_{1 \leq i \leq n} \alpha_i [v_i] + \sum_{1 \leq j \leq k} \beta_j 0 &= \sum_{1 \leq i \leq n} \alpha_i [v_i], \end{aligned}$$

sodass man den  $U$ -Anteil eines Vektors bezüglich der Basis  $B$  vernachlässigen kann. Die Nebenklassen

$$[v_1], \dots, [v_n] \in V/U$$

bilden eine Basis  $B_U = ([v_1], \dots, [v_n])$  des Quotientenraums  $V/U$ . Also gilt

$$\dim(V/U) = \dim(V) - \dim(U).$$

Ist  $v_B = (\alpha_1, \dots, \alpha_n, \beta_1, \dots, \beta_k) \in K^{n+k}$  der Koordinatenvektor eines Vektors  $v \in V$  bezüglich der Basis  $B$ , so ist

$$v_{B_U} = (\alpha_1, \dots, \alpha_n) \in K^n$$

der Koordinatenvektor von  $[v] \in V/U$  bezüglich  $B_U$ .

### 3.12 Affine Unterräume und Koordinaten

**Definition** (*affiner Unterraum, affine Kombination*)

Sei  $V$  ein  $K$ -Vektorraum.

- (a) Ein  $A \subseteq V$  heißt ein *affiner Unterraum* von  $V$ , falls  $A$  leer ist oder ein  $v \in V$  und ein Unterraum  $U$  von  $V$  existieren mit

$$A = v + U = \{v + u \mid u \in U\}.$$

- (b) Ein  $w \in V$  heißt eine *affine Kombination* der Vektoren  $v_1, \dots, v_n$  in  $V$ , falls Skalare  $\alpha_1, \dots, \alpha_n$  existieren mit:

$$w = \alpha_1 v_1 + \dots + \alpha_n v_n,$$

$$\alpha_1 + \dots + \alpha_n = 1.$$

Die nichtleeren affinen Unterräume von  $V$  sind also die um einen Vektor  $v$  „verschobenen“ Unterräume von  $V$  (also alle Nebenklassen, vgl. 3.11). Der Vektor  $v$  ist im Gegensatz zu  $U$  nicht eindeutig bestimmt. Es gilt

$v + U = v' + U'$  genau dann, wenn  $U = U'$  und  $v - v' \in U$ .

Dass die leere Menge als affiner Unterraum gilt, ist eine nützliche Konvention (vgl. Abschnitt 4.8). Im Kontrast dazu ist die leere Menge kein Unterraum von  $V$ .

**Beispiele**

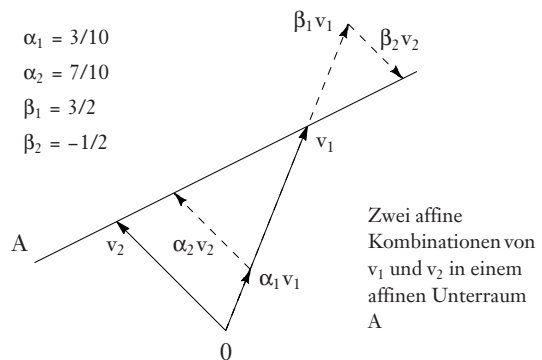
- (1) Die affinen Unterräume von  $\mathbb{R}$  sind  $\emptyset$  und alle einpunktigen Mengen  $\{x\}$  (denn es gilt  $\{x\} = x + U$  für den Unterraum  $U = \{0\}$  von  $\mathbb{R}$ ).
- (2) Die affinen Unterräume von  $\mathbb{R}^2$  sind  $\emptyset$ , alle einpunktigen Mengen  $\{v\}$  und alle Geraden  $\{v_0 + \alpha v_1 \mid \alpha \in \mathbb{R}\}$  in der Ebene.

Affine Kombinationen sind zunächst lediglich spezielle Linearkombinationen. Den Zusammenhang mit affinen Unterräumen zeigt:

#### Charakterisierung der affinen Unterräume

Sei  $V$  ein  $K$ -Vektorraum, und sei  $A \subseteq V$ . Dann sind äquivalent:

- (a)  $A$  ist ein affiner Unterraum von  $V$ .
- (b)  $A$  ist abgeschlossen unter affinen Kombinationen: Für alle  $v_1, \dots, v_n \in A$  und  $\alpha_1, \dots, \alpha_n \in K$  mit  $\alpha_1 + \dots + \alpha_n = 1$  ist  $\alpha_1 v_1 + \dots + \alpha_n v_n \in A$ .





Die Äquivalenz ist klar für  $A = \emptyset$ . Ist  $A = v + U$  ein affiner Unterraum von  $V$ , so haben affine Kombinationen mit Vektoren in  $A$  die Form

$$\alpha_1(v + u_1) + \dots + \alpha_n(v + u_n) = 1v + \alpha_1 u_1 + \dots + \alpha_n u_n \in v + U = A,$$

sodass  $A$  abgeschlossen unter affinen Kombinationen ist. Gilt umgekehrt (b) und ist  $v \in A$  beliebig, so ist  $U = \{w - v \mid w \in A\}$  wegen  $w - w = 0 \in U$  und

$$\alpha(w_1 - v) + \beta(w_2 - v) = (1 - \alpha - \beta)v + \alpha w_1 + \beta w_2 - v = w' - v$$

ein Unterraum von  $V$ . Zudem gilt  $A = \{v + w - v \mid w \in A\} = v + U$ .

Aus der Charakterisierung erhalten wir:

### Erzeugung von affinen Räumen

Für alle  $v_0, v_1, \dots, v_n \in V$  ist  $A = \{w \mid w \text{ ist eine affine Kombination von } v_0, \dots, v_n\}$  der kleinste affine Unterraum von  $V$ , der  $v_0, \dots, v_n$  als Elemente enthält. Es gilt

$$A = v_0 + \text{span}(v_0 - v_1, \dots, v_0 - v_n).$$

Wir erweitern nun noch den Basisbegriff auf affine Räume. Dabei beschränken wir uns auf den endlich-dimensionalen Fall.

### Definition (affine Basis, $\dim(A)$ , affine und baryzentrische Koordinatenvektoren)

Sei  $A = v_0 + U$  ein affiner Unterraum von  $V$ , und seien  $v_1, \dots, v_n \in A$ . Dann heißt  $(v_0, v_1, \dots, v_n)$  eine *affine Basis* und  $n$  die *Dimension* von  $A$ , falls  $(v_1 - v_0, \dots, v_n - v_0)$  eine Basis von  $U$  ist. Für alle  $w \in A$  heißt das eindeutige  $n$ -Tupel  $(\alpha_1, \dots, \alpha_n)$  mit

$$w = v_0 + \alpha_1(v_1 - v_0) + \dots + \alpha_n(v_n - v_0)$$

der *affine Koordinatenvektor* und das eindeutige  $(n+1)$ -Tupel  $(\lambda_0, \dots, \lambda_n)$  mit

$$w = \lambda_0 v_0 + \lambda_1 v_1 + \dots + \lambda_n v_n, \quad \lambda_0 + \dots + \lambda_n = 1.$$

der *baryzentrische Koordinatenvektor* von  $w$  bzgl.  $(v_0, \dots, v_n)$ .

Es gilt  $\lambda_0 = 1 - (\alpha_1 + \dots + \alpha_n)$  und  $\lambda_k = \alpha_k$  für alle  $1 \leq k \leq n$ . In affinen Koordinaten ist der Vektor  $v_0$  als „Ursprung“ des affinen Raums  $A$  ausgezeichnet, in baryzentrischen Koordinaten sind die Vektoren  $v_0, \dots, v_n$  gleichberechtigt. Das Wort „Baryzentrum“ bedeutet „Schwerpunkt“. Die Namensgebung illustriert:

### Beispiel

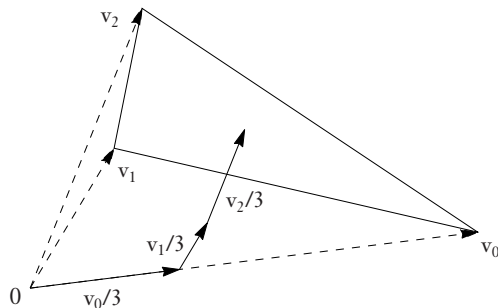
Für alle  $v_0, v_1, v_2 \in \mathbb{R}^2$  ist

$$w = v_0/3 + v_1/3 + v_2/3$$

der Schwerpunkt des durch die Vektoren  $v_0, v_1, v_2$  definierten Dreiecks  $D$ . Es gilt

$$D = \{ \lambda_0 v_0 + \lambda_1 v_1 + \lambda_2 v_2 \mid$$

$$\lambda_0 + \lambda_1 + \lambda_2 = 1, \lambda_0, \lambda_1, \lambda_2 \geq 0 \}.$$





## 4. Kapitel

---

# Lineare Abbildungen

---

## 4.1 Gruppenhomomorphismen

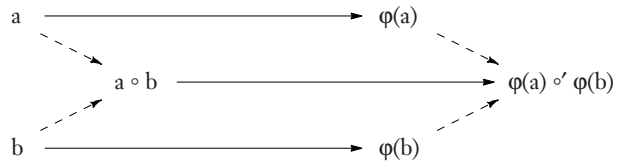
### Definition (Gruppenhomomorphismus)

Seien  $(G, \circ)$ ,  $(G', \circ')$  Gruppen. Eine Abbildung  $\varphi : G \rightarrow G'$  heißt ein (Gruppen-) Homomorphismus, falls

$$\varphi(a \circ b) = \varphi(a) \circ' \varphi(b) \quad \text{für alle } a, b \in G. \quad (\text{Homomorphiebedingung})$$

Strukturerhaltende Abbildungen gehören wie die Unterstrukturen zu den Grundmotiven der Mathematik. Wir beschränken uns hier auf die Gruppen, allgemeiner könnten wir auch Homomorphismen zwischen Halbgruppen betrachten.

Die Grundidee ist:



$$\text{Die Homomorphiebedingung } \varphi(a \circ b) = \varphi(a) \circ' \varphi(b)$$

*Die Anwendung der Abbildung und die Ausführung der Operation sind vertauschbar.*

Ist  $\varphi : G \rightarrow G'$  eine Abbildung und sind  $a, b \in G$ , so können wir zuerst  $c = a \circ b$  bilden und dann  $\varphi$  anwenden. Wir erhalten so  $\varphi(c) \in G'$ . Wir können aber auch zuerst  $a$  und  $b$  mit Hilfe von  $\varphi$  nach  $G'$  schicken und dort  $\varphi(a) \circ' \varphi(b)$  bilden. Die Homomorphiebedingung besagt, dass beide Wege zu dem selben Element von  $G'$  führen:

$$\varphi(a \circ b) = \varphi(c) = \varphi(a) \circ' \varphi(b).$$

Häufig gebraucht werden:

$$\varphi(e) = e', \quad \varphi(a^{-1}) = \varphi(a)^{-1} \quad \text{für alle } a \in G.$$

Diese Eigenschaften lassen sich wie folgt einsehen. Es gilt

$$\varphi(e) = \varphi(e \circ e) = \varphi(e) \circ' \varphi(e), \quad \text{sodass } e' = \varphi(e),$$

$$e' = \varphi(e) = \varphi(a \circ a^{-1}) = \varphi(a) \circ' \varphi(a^{-1}), \quad \text{sodass } \varphi(a^{-1}) = \varphi(a)^{-1}.$$

### Notationen

- (1) Abbildungen zwischen Gruppen notieren wir auch in der Form  $\varphi : (G, \circ) \rightarrow (G', \circ')$ . Dabei ist  $\text{Def}(\varphi) = G$  und  $\text{Bild}(\varphi) \subseteq G'$ .
- (2) Umgekehrt erleichtert es oft die Notation, die Operationen gar nicht zu erwähnen und etwa multiplikativ  $\varphi(ab) = \varphi(a)\varphi(b)$  zu schreiben, obwohl die Operationen in  $G$  und  $G'$  verschieden sein können.

**Beispiele**

- (1) Wir definieren
- $\varphi : (\mathbb{R}^2, +) \rightarrow (\mathbb{R}, +)$
- durch

$$\varphi(a, b) = a \quad \text{für alle } (a, b) \in \mathbb{R}^2.$$

Die Abbildung  $\varphi$  beschreibt die Projektion auf die erste Koordinate. Für alle  $(a, b), (c, d) \in \mathbb{R}^2$  gilt:

$$\varphi((a, b) + (c, d)) = \varphi(a + c, b + d) = a + c = \varphi(a, b) + \varphi(c, d).$$

Also ist  $\varphi$  ein Homomorphismus.

- (2) Die reelle Exponentialfunktion
- $\exp : (\mathbb{R}, +) \rightarrow (\mathbb{R}^*, \cdot)$
- zur Basis
- $e$
- ist ein Homomorphismus. Denn nach dem Additionstheorem der Analysis gilt

$$\exp(x + y) = e^{x+y} = e^x \cdot e^y = \exp(x) \cdot \exp(y) \quad \text{für alle } x, y \in \mathbb{R}.$$

Gleiches gilt für die komplexe Exponentialfunktion  $\exp : (\mathbb{C}, +) \rightarrow (\mathbb{C}^*, \cdot)$ .

- (3) Die komplexe Konjugation
- $\varphi : (\mathbb{C}, +) \rightarrow (\mathbb{C}, +)$
- mit

$$\varphi(x + iy) = x - iy \quad \text{für alle } x + iy \in \mathbb{C}$$

ist ein Homomorphismus. Denn für alle  $x_1 + iy_1, x_2 + iy_2 \in \mathbb{C}$  gilt

$$\begin{aligned} \varphi((x_1 + iy_1) + (x_2 + iy_2)) &= \varphi((x_1 + x_2) + i(y_1 + y_2)) = \\ &= x_1 + x_2 - i(y_1 + y_2) = x_1 - iy_1 + x_2 - iy_2 = \varphi(x_1 + iy_1) + \varphi(x_2 + iy_2). \end{aligned}$$

Das Gleiche gilt, wenn wir  $(\mathbb{C}^*, \cdot)$  statt  $(\mathbb{C}, +)$  zugrunde legen.

- (4) Wir betrachten
- $(\mathbb{Z}, +)$
- , eine beliebige Gruppe
- $(G, \circ)$
- und ein beliebiges
- $a \in G$
- . Nun definieren wir
- $\varphi : \mathbb{Z} \rightarrow G$
- durch

$$\varphi(n) = a^n \quad \text{für alle } n \in \mathbb{Z}.$$

Dann gilt

$$\varphi(n + m) = a^{n+m} = a^n \circ a^m = \varphi(n) \circ \varphi(m) \quad \text{für alle } n, m \in \mathbb{Z}.$$

Also ist  $\varphi$  ein Homomorphismus. Spezialfälle sind

$$\varphi : (\mathbb{Z}, +) \rightarrow (\mathbb{Z}_p^*, \cdot), \quad \varphi(n) = [a]^n, \quad \text{wobei } [a] = [a]_p \in \mathbb{Z}_p^*, \quad p \text{ prim,}$$

$$\varphi : (\mathbb{Z}, +) \rightarrow (\mathbb{Z}_m, +), \quad \varphi(n) = n[a], \quad \text{wobei } [a] = [a]_m \in \mathbb{Z}_m, \quad m \geq 1.$$

- (5) Für alle Gruppen
- $G, G'$
- ist
- $\varphi : G \rightarrow G'$
- mit
- $\varphi(a) = e'$
- für alle
- $a \in G$
- ein Homomorphismus, der sog.
- triviale Homomorphismus*
- von
- $G$
- nach
- $G'$
- .

- (6) Für jede Gruppe
- $G$
- ist die Identität
- $\text{id} : G \rightarrow G$
- ein Homomorphismus.

- (7) Sind
- $\varphi : G \rightarrow G'$
- und
- $\psi : G' \rightarrow G''$
- Homomorphismen, so ist auch die Komposition
- $\psi \circ \varphi : G \rightarrow G''$
- ein Homomorphismus.

## 4.2 Mono-, Epi-, Iso-, Endo- und Automorphismen

### Definition (Typen von Homomorphismen, isomorphe Gruppen)

Seien  $G, G'$  Gruppen.

- (a) Ein Homomorphismus  $\varphi : G \rightarrow G'$  heißt

*Monomorphismus*, falls  $\varphi$  injektiv ist,

*Epimorphismus*, falls  $\varphi$  surjektiv ist,

*Isomorphismus*, falls  $\varphi$  bijektiv ist,

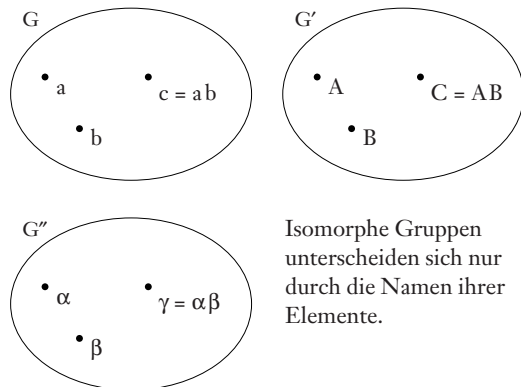
*Endomorphismus*, falls  $G = G'$ ,

*Automorphismus*, falls  $G = G'$  und  $\varphi$  bijektiv ist.

- (b)  $G$  und  $G'$  heißen *isomorph*, in Zeichen  $G \cong G'$ , falls ein Isomorphismus  $\varphi : G \rightarrow G'$  existiert.

Die aus dem Griechischen stammenden Vorsilben bestimmen die Bedeutung: „mono“ steht für „allein, einzig, nur“ (Monolog, Monokultur), „epi“ für „auf“ (Epidemie), „iso“ für „gleich, entsprechend“ (Isobaren, Isomere), „endo“ für „innerhalb“ (endogen, Endogamie), „auto“ für „selbst“ (autonom, Autodidakt). Homomorphismen werden oft mit griechischen Buchstaben wie  $\varphi, \psi, \pi, \Phi, \Psi, \dots$  bezeichnet.

Isomorphismen sind von besonderer Bedeutung. Ein Isomorphismus ist eine strukturerhaltende Bijektion. Er bringt, wie jede Bijektion, die Elemente zweier Mengen in eine 1-1-Korrespondenz, sodass jedem Element  $a$  der einen Menge genau ein Element  $\varphi(a)$  der anderen Menge entspricht. Zusätzlich erhält diese Korrespondenz die Struktur der Menge gemäß der Homomorphiebedingung. Eine anschauliche Interpretation ist:



*Ein Isomorphismus  $\varphi : G \rightarrow G'$  ändert die Namen der Elemente:  
 $a \in G$  erhält den neuen Namen  $\varphi(a) \in G'$ .*

Die Isomorphie von  $G$  und  $G'$  bedeutet dann:

*$G$  und  $G'$  unterscheiden sich, bei einer geeigneten Umbenennung,  
 lediglich durch die Namen ihrer Elemente.*

Der Zusatz „geeignet“ ist hier wichtig: Die Isomorphie zweier Gruppen involviert einen Existenzquantor, denn  $G \cong G'$  bedeutet: Es gibt einen Isomorphismus  $\varphi : G \rightarrow G'$ .

Wir betrachten einige Beispiele. Die ersten sieben Beispiele entsprechen dabei den sieben Beispielen des vorherigen Abschnitts.

### Beispiele

- (1) Die Projektion  $\varphi : (\mathbb{R}^2, +) \rightarrow (\mathbb{R}, +)$  auf die erste Koordinate ist ein Epimorphismus, aber kein Monomorphismus (da zum Beispiel  $\varphi(0, 1) = \varphi(0, 0) = 0$ ).
- (2) Die reelle Exponentialfunktion  $\exp : (\mathbb{R}, +) \rightarrow (\mathbb{R}^*, \cdot)$  ist ein Monomorphismus, aber kein Epimorphismus (da zum Beispiel  $-1$  nicht angenommen wird). Dagegen ist  $\exp : (\mathbb{R}, +) \rightarrow (\mathbb{R}^+, \cdot)$  ein Isomorphismus.
- (3) Die komplexe Konjugation  $\varphi : (\mathbb{C}, +) \rightarrow (\mathbb{C}, +)$  ist ein Automorphismus.
- (4) Die Potenzierung  $\varphi : (\mathbb{Z}, +) \rightarrow (G, \circ)$  zur Basis  $a \in G$  mit  $\varphi(n) = a^n$  für alle  $n$  ist im Allgemeinen weder ein Mono- noch ein Epimorphismus. Wird  $G$  durch  $a$  erzeugt, so ist  $\varphi$  ein Epimorphismus. Ist  $a^n \neq e$  für alle  $n \geq 1$ , so ist  $\varphi$  ein Monomorphismus.
- (5) Der triviale Homomorphismus  $\varphi : G \rightarrow G'$  ist für  $G \neq \{e\}$  kein Monomorphismus und für  $G' \neq \{e'\}$  kein Epimorphismus.
- (6) Für alle Gruppen  $G$  ist die Identität  $\text{id} : G \rightarrow G$  ein Automorphismus. Damit ist  $G$  isomorph zu sich selbst.
- (7) Sind  $\varphi : G \rightarrow G'$  und  $\psi : G' \rightarrow G''$  Homomorphismen eines bestimmten Typs, so hat auch  $\psi \circ \varphi : G \rightarrow G''$  diesen Typ. Insbesondere ist die Isomorphie-Relation transitiv: Ist  $G$  isomorph zu  $G'$  und weiter  $G'$  isomorph zu  $G''$ , so ist  $G$  isomorph zu  $G''$ .
- (8) Ist  $\varphi : G \rightarrow G'$  ein Isomorphismus, so auch  $\varphi^{-1} : G' \rightarrow G$ . Damit ist die Isomorphie symmetrisch: Ist  $G$  isomorph zu  $G'$ , so ist  $G'$  isomorph zu  $G$ .
- (9) Ist  $N$  ein Normalteiler von  $G$ , so ist die Abbildung  $\pi : G \rightarrow G/N$  mit  $\pi(a) = aN$  für alle  $a \in G$  wohldefiniert und ein Epimorphismus. Sie heißt die *natürliche Projektion* von  $G$  auf  $G/N$ .

Ist also  $\mathcal{G}$  eine Menge von Gruppen, ist die Relation  $\cong$  eine Äquivalenzrelation auf  $\mathcal{G}$ . Die Beispiele zeigen weiter, dass Gruppenhomomorphismen Anlass zur Definition neuer Gruppen geben. Für jede Gruppe  $G$  ist

$$\text{Aut}(G) = \{ \varphi : G \rightarrow G \mid \varphi \text{ ist ein Automorphismus} \}$$

eine Gruppe unter der Komposition. Sie heißt die *Automorphismengruppe* von  $G$  und ist eine Untergruppe der Permutationsgruppe  $S_G$  aller Bijektionen von  $G$  nach  $G$  (vgl. 2.3). Für  $G \neq \{e\}$  ist  $\text{Aut}(G)$  eine echte Untergruppe von  $S_G$ . Denn sind  $a, e \in G$  verschieden, so ist die Permutation  $f : G \rightarrow G$  mit  $f(a) = e$ ,  $f(e) = a$ ,  $f(b) = b$  für alle anderen  $b$  kein Automorphismus (da  $f(e) \neq e$ ). Weiter ist  $\text{Aut}(G)$  die Gruppe der invertierbaren Elemente des Monoids  $\text{End}(G) = \{ \varphi : G \rightarrow G \mid \varphi \text{ ist ein Endomorphismus} \}$ .

### 4.3 Kern und Bild

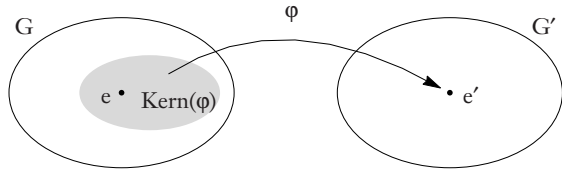
**Definition** (*Kern und Bild eines Homomorphismus*)

Sei  $\varphi : G \rightarrow G'$  ein Gruppenhomomorphismus. Dann setzen wir

$$\text{Kern}(\varphi) = \{a \in G \mid \varphi(a) = e'\},$$

$$\text{Bild}(\varphi) = \{\varphi(a) \mid a \in G\}.$$

Die Mengen  $\text{Kern}(\varphi)$  und  $\text{Bild}(\varphi)$  heißen der *Kern* bzw. das *Bild* von  $\varphi$ .

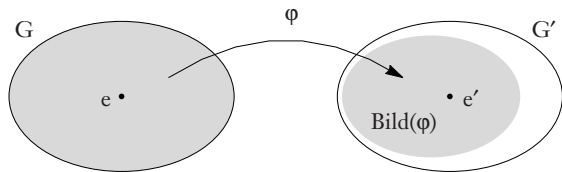


Neben  $e$  können weitere Elemente von  $G$  auf das neutrale Element  $e'$  von  $G'$  abgebildet werden. Die Menge dieser Elemente ist  $\text{Kern}(\varphi)$ .

Nach Definition ist der Kern von  $\varphi : G \rightarrow G'$  eine Teilmenge des Definitionsbereichs  $G$  von  $\varphi$ , während das Bild von  $\varphi$  eine Teilmenge des Wertevorrats  $G'$  von  $\varphi$  ist. Als Faustregel kann

*Kern links und Bild rechts*

helfen. Es ist wichtig, die beiden Welten  $G$  und  $G'$  zu trennen, wenn  $G \neq G'$ .



In  $\text{Bild}(\varphi)$  werden alle Werte von  $\varphi$  gesammelt.

Kern und Bild lassen sich mit den allgemeinen Abbildungsbegriffen beschreiben:

$$\text{Kern}(\varphi) = \varphi^{-1}[\{e'\}] = \text{„das Urbild von } \{e'\} \text{ unter } \varphi\text{“} = \text{„die Faser von } \varphi \text{ über } e'\text{“},$$

$$\text{Bild}(\varphi) = \varphi[G] = \text{„der Wertebereich von } \varphi\text{“}.$$

Damit sind Kern und Bild streng genommen nichts Neues. Sie spielen aber für die Beschreibung von Homomorphismen eine so bedeutende Rolle, dass sich eine eigene Begriffsbildung lohnt. Wichtige Eigenschaften sind:

<b>Kern und Bild eines Gruppenhomomorphismus <math>\varphi : G \rightarrow G'</math></b>	
<p><math>\text{Kern}(\varphi)</math> ist ein Normalteiler von <math>G</math>.</p> <p><math>\text{Bild}(\varphi)</math> ist eine Untergruppe von <math>G'</math>.</p>	
<p><math>\varphi</math> ist genau injektiv, wenn <math>\text{Kern}(\varphi) = \{e\}</math>.</p> <p><math>\varphi</math> ist genau dann surjektiv, wenn <math>\text{Bild}(\varphi) = G'</math>.</p>	

Wir weisen zur Illustration die Eigenschaften des Kerns nach. Dabei notieren wir die Gruppenoperationen von  $G$  und  $G'$  multiplikativ.



### Kern( $\varphi$ ) ist ein Normalteiler von $G$

Wegen  $\varphi(e) = e'$  ist  $e \in \text{Kern}(\varphi)$  und damit  $\text{Kern}(\varphi) \neq \emptyset$ . Für  $a, b \in \text{Kern}(\varphi)$  ist

$$\varphi(ab^{-1}) = \varphi(a) \varphi(b^{-1}) = \varphi(a) \varphi(b)^{-1} = e' e'^{-1} = e',$$

sodass  $ab^{-1} \in \text{Kern}(\varphi)$ . Nach dem Untergruppenkriterium ist also  $\text{Kern}(\varphi)$  eine Untergruppe von  $G$ .

Für „Normalteiler“ seien  $a \in G$  und  $b \in \text{Kern}(\varphi)$ . Dann gilt

$$\varphi(aba^{-1}) = \varphi(a) \varphi(b) \varphi(a^{-1}) = \varphi(a) e \varphi(a)^{-1} = \varphi(a) \varphi(a)^{-1} = e'.$$

Also gilt  $aba^{-1} \in \text{Kern}(\varphi)$ . Dies zeigt, dass  $\text{Kern}(\varphi)$  ein Normalteiler ist (vgl. 2.7).

### $\varphi$ ist genau dann injektiv, wenn $\text{Kern}(\varphi) = \{e\}$

Sei  $\varphi$  injektiv. Wegen  $\varphi(e) = e'$  gilt  $\{e\} \subseteq \text{Kern}(\varphi)$ . Da  $\varphi$  injektiv ist, hat  $e'$  höchstens ein Urbild unter  $\varphi$ , sodass  $\text{Kern}(\varphi) \subseteq \{e\}$ . Damit ist  $\text{Kern}(\varphi) = \{e\}$ .

Sei umgekehrt  $\text{Kern}(\varphi) = \{e\}$  und seien  $a, b \in G$  mit  $\varphi(a) = \varphi(b)$ . Dann gilt

$$\varphi(ab^{-1}) = \varphi(a) \varphi(b)^{-1} = \varphi(a) \varphi(a)^{-1} = e'.$$

Also ist  $ab^{-1} \in \text{Kern}(\varphi)$  und damit  $ab^{-1} = e$ , also  $a = b$ . Dies zeigt, dass  $\varphi$  injektiv ist.

Die Untergruppe  $\text{Bild}(\varphi)$  ist im Allgemeinen kein Normalteiler. Ist  $G$  eine Gruppe und  $H$  eine Untergruppe von  $G$ , die kein Normalteiler ist, so ist die Identität  $\text{id} : H \rightarrow G$  ein Homomorphismus, dessen Bild kein Normalteiler ist.

### Beispiele

- (1) Für die Projektion  $\varphi : (\mathbb{R}^2, +) \rightarrow (\mathbb{R}, +)$  auf die erste Koordinate gilt

$$\text{Kern}(\varphi) = \{0\} \times \mathbb{R} = \{(0, y) \mid y \in \mathbb{R}\}, \quad \text{Bild}(\varphi) = \mathbb{R}.$$

- (2) Sei  $\varphi : (\mathbb{Z}, \cdot) \rightarrow (\mathbb{Z}_5^*, \cdot)$  definiert durch

$$\varphi(a) = [2]^a \quad \text{für alle } a \in \mathbb{Z} \quad (\text{wobei } [2] = [2]_5),$$

sodass  $\varphi(0) = [1]$ ,  $\varphi(1) = [2]$ ,  $\varphi(2) = [4]$ ,  $\varphi(3) = [8] = [3]$ ,  $\varphi(4) = [16] = [1]$ .  
Dann gilt (vgl. 2.6):

$$\text{Kern}(\varphi) = 4\mathbb{Z} = \{0, 4, -4, 8, -8, \dots\}, \quad \text{Bild}(\varphi) = \mathbb{Z}_5^*.$$

- (3) Sei  $\varphi : (\mathbb{Z}, +) \rightarrow (\mathbb{Z}_{12}, +)$  definiert durch

$$\varphi(a) = a[2] \quad \text{für alle } a \in \mathbb{Z} \quad (\text{wobei nun } [2] = [2]_{12}),$$

sodass  $\varphi(3) = [6]$ ,  $\varphi(6) = [12] = [0]$ ,  $\varphi(-1) = -[2] = [10]$  usw. Dann gilt

$$\text{Kern}(\varphi) = 6\mathbb{Z} = \{0, 6, -6, \dots\},$$

$$\text{Bild}(\varphi) = \{[0], [2], [4], [6], [8], [10]\}.$$

## 4.4 Der Homomorphiesatz

### Satz (Homomorphiesatz)

Sei  $\varphi : G \rightarrow G'$  ein Gruppenhomomorphismus.

*Homomorphiesatz für Epimorphismen oder Isomorphiesatz*

Ist  $\varphi$  ein Epimorphismus, so sind  $G/\text{Kern}(\varphi)$  und  $G'$  isomorph. Genauer ist

$$\psi : G/\text{Kern}(\varphi) \rightarrow G', \quad \psi(a \text{ Kern}(\varphi)) = \varphi(a) \quad \text{für alle } a \in G$$

ein Isomorphismus.

*Allgemeiner Homomorphiesatz*

Sei  $N \subseteq \text{Kern}(\varphi)$  ein Normalteiler von  $G$  und  $\pi : G \rightarrow G/N$  die natürliche Projektion, d.h.

$$\pi(a) = aN \quad \text{für alle } a \in G.$$

Dann existiert genau ein Homomorphismus  $\psi : G/N \rightarrow G'$  mit

$$\varphi = \psi \circ \pi.$$

Zur Illustration der Sätze sind kommutative Diagramme nützlich. Zwei Abbildungswege führen von  $G$  zu  $G'$ , und die Kommutativität des Diagramms bedeutet, dass diese Wege gleich sind.

Um uns den Sätzen zu nähern, betrachten wir eine Surjektion

$$f : A \rightarrow A'$$

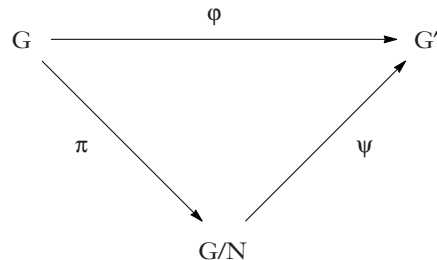
zwischen beliebigen Mengen  $A$  und  $A'$ . Wir stellen uns  $f$  als eine Färbung vor, die jedes Element  $a$  von  $A$  mit einer Farbe  $f(a)$  in  $A'$  einfärbt. Die Surjektivität bedeutet, dass jede

Farbe in  $A'$  tatsächlich als Farbe eines Elements in  $A$  vorkommt. Der Farbkasten  $A'$  wird also voll ausgenutzt. Wir definieren nun eine Äquivalenzrelation auf  $A$  durch

$$a \sim b, \quad \text{falls} \quad f(a) = f(b) \quad \text{für alle } a, b \in A.$$

Sie entspricht der Identifizierung von Elementen mit der gleichen Farbe. Für jedes  $a \in A$  ist  $a/\sim$  die Menge aller  $b \in A$ , die die Farbe  $f(a)$  haben. Da  $f$  surjektiv ist, ist die Anzahl der Farbklassen  $a/\sim$  gleich der Anzahl der Farben: Es gilt  $|A/\sim| = |A'|$ . Genauer ist die Abbildung  $g : A/\sim \rightarrow A'$  mit

$$g(a/\sim) = f(a) \quad \text{für alle } a \in A$$



Es gilt  $\varphi = \psi \circ \pi$ . Sind in einem Diagramm alle Wege, die von einer Menge in eine andere durch Anwendung von Funktionen und ihrer Komposition führen, gleich, so heißt das Diagramm kommutativ. Für den Homomorphiesatz liegt ein kommutatives Diagramm vor.

die sich aufdrängende oder, wie Mathematiker gerne sagen, „kanonische“ Bijektion zwischen den beiden Mengen. Färben wir also die Bücher einer Bibliothek mit fünf Farben, so haben wir genau fünf Farbklassen (Mengen von Büchern gleicher Farbe) vorliegen. Der Vorgang, die Menge der blauen Bücher auf die Farbe „blau“ abzubilden, ist so natürlich, dass er eigentlich kaum der Rede wert ist. Damit sind wir der abstrakten, aber letztendlich auch einfachen Aussage des Epimorphiesatzes bereits sehr nahe. Im Unterschied zu reinen Surjektionen bleibt nun zusätzlich die algebraische Struktur erhalten, wir betrachten also strukturerhaltende Färbungen. Die Operation  $\circ'$  auf  $G'$  können wir als algebraische „Farbmischung“ ansehen: Erhält  $a$  die Farbe  $\varphi(a)$  und  $b$  die Farbe  $\varphi(b)$ , so erhält  $a \circ b$  die „Mischfarbe“  $\varphi(a) \circ' \varphi(b)$ .

### Beispiele

- (1) Sei  $\varphi : (\mathbb{R}^2, +) \rightarrow (\mathbb{R}, +)$  die Projektion auf die erste Koordinate. Für alle  $x \in \mathbb{R}$  sei  $S_x = \{x\} \times \mathbb{R} = \{(x, y) \mid y \in \mathbb{R}\}$  die Senkrechte durch  $(x, 0)$ . Dann ist  $\varphi$  ein Epimorphismus mit  $\text{Kern}(\varphi) = S_0$  ( $y$ -Achse). Es gilt

$$G/\text{Kern}(\varphi) = \{(x, y) + S_0 \mid (x, y) \in \mathbb{R}^2\} = \{S_x \mid x \in \mathbb{R}\},$$

$$\psi(S_x) = x \text{ für alle } x \in \mathbb{R}.$$

Die Faktorgruppe  $\mathbb{R}^2/\text{Kern}(\varphi)$  besteht also aus allen zur  $x$ -Achse senkrechten Geraden. Der Isomorphismus  $\psi : \mathbb{R}^2/\text{Kern}(\varphi) \rightarrow \mathbb{R}$  gibt den Schnittpunkt dieser Geraden mit der  $x$ -Achse an.

- (2) Sei  $\varphi : (\mathbb{Z}, +) \rightarrow (\mathbb{Z}_{10}, +)$  der Vervielfachungs-Homomorphismus mit

$$\varphi(a) = a[3]_{10} \text{ für alle } a \in \mathbb{Z}.$$

Dann gilt  $\text{Kern}(\varphi) = 10\mathbb{Z}$  und  $\text{Bild}(\varphi) = \mathbb{Z}_{10}$ . Damit ist

$$\mathbb{Z}/\text{Kern}(\varphi) = \mathbb{Z}/10\mathbb{Z} = \mathbb{Z}_{10}.$$

Es gilt also  $G/\text{Kern}(\varphi) = G'$ . Für  $\psi : \mathbb{Z}_{10} \rightarrow \mathbb{Z}_{10}$  gilt

$$\psi([a]_{10}) = \varphi(a) = a[3]_{10} \text{ für alle } a \in \mathbb{Z}.$$

Der konstruierte Isomorphismus  $\psi : \mathbb{Z}_{10} \rightarrow \mathbb{Z}_{10}$  ist nicht die Identität.

Im allgemeinen Homomorphiesatz identifizieren wir für  $N \neq \text{Kern}(\varphi)$  weniger Objekte miteinander als möglich. In Analogie zu den Farben: Wir teilen die gleichfarbigen Bücher zusätzlich in „Taschenbücher“ und „fester Einband“ oder in verschiedene Sprachen ein. Wir halten damit einige Merkmale fest, die  $\varphi$  ignoriert.

### Beispiel

Sei  $\varphi : (\mathbb{Z}, +) \rightarrow (\mathbb{Z}_8, +)$  definiert durch  $\varphi(a) = a[2]_8$  für alle  $a$ , sodass  $\text{Kern}(\varphi) = 4\mathbb{Z}$ . Für den Normalteiler  $N = 16\mathbb{Z} \subseteq \text{Kern}(\varphi)$  von  $\mathbb{Z}$  gilt  $\mathbb{Z}/N = \mathbb{Z}_{16}$ . Für die Abbildungen  $\pi : \mathbb{Z} \rightarrow \mathbb{Z}_{16}$ ,  $\psi : \mathbb{Z}_{16} \rightarrow \mathbb{Z}_8$  wie im Homomorphiesatz ist

$$\pi(a) = [a]_{16}, \quad \psi([a]_{16}) = \varphi(a) = a[2]_8 \text{ für alle } a \in \mathbb{Z}.$$

## 4.5 Lineare Abbildungen

### Definition (lineare Abbildung)

Seien  $V, W$   $K$ -Vektorräume. Dann heißt eine Abbildung  $f : V \rightarrow W$  *linear*, falls gilt:

(a)  $f : (V, +) \rightarrow (W, +)$  ist ein Gruppenhomomorphismus, d. h.,

$$f(v + w) = f(v) + f(w) \quad \text{für alle } v, w \in V,$$

(b)  $f(\alpha v) = \alpha f(v)$  für alle  $\alpha \in K$  und  $v \in V$ . (Skalierungseigenschaft)

Eine lineare Abbildung  $f : V \rightarrow W$  ist also ein Homomorphismus zwischen den additiven Vektorgruppen, der zusätzlich die Skalarmultiplikation respektiert. Für lineare Abbildungen sind  $f, g, F, G, \dots$  üblicher als  $\varphi, \psi, \pi, \dots$

Die Bedingungen (a) und (b) lassen sich zusammenfassen:

$$\begin{array}{ccc}
 \alpha v & \xrightarrow{\quad} & \alpha f(v) \\
 & \searrow \text{dashed} & \searrow \text{dashed} \\
 & \alpha v + \beta w & \xrightarrow{\quad} \alpha f(v) + \beta f(w) \\
 & \nearrow \text{dashed} & \nearrow \text{dashed} \\
 \beta w & \xrightarrow{\quad} & \beta f(w)
 \end{array}$$

(+)  $f(\alpha v + \beta w) = \alpha f(v) + \beta f(w)$  für alle  $\alpha, \beta \in K$  und  $v, w \in V$ .

Setzt man  $\alpha = \beta = 1$ , so erhält man (a);  $w = 0, \beta = 0$  liefert (b).

Da eine lineare Abbildung  $f : V \rightarrow W$  ein Homomorphismus ist, stehen die Begriffe und Ergebnisse der vorangehenden Abschnitte zur Verfügung:

$f$ ist ein Mono-, Epi-, Iso-, Endo- bzw. Automorphismus, wenn $f$ injektiv, $f$ surjektiv, $f$ bijektiv, $V = W$ bzw. $V = W$ und $f$ bijektiv ist.
$\text{Kern}(f) = \{ v \in V \mid f(v) = 0 \}$ ist ein Unterraum von $V$ , $\text{Bild}(f) = \{ f(v) \mid v \in V \}$ ist ein Unterraum von $W$ .
<p style="text-align: center;"><i>Isomorphiesatz für Vektorräume</i></p> <p>Ist <math>f : V \rightarrow W</math> ein Epimorphismus, so ist <math>g : V/\text{Kern}(f) \rightarrow W</math> mit</p> $g(v + \text{Kern}(f)) = f(v) \quad \text{für alle } v \in V$ <p>ein Isomorphismus zwischen dem Quotientenraum <math>V/\text{Kern}(f)</math> und <math>W</math>.</p>

Neu kommt hinzu:

$f$  ist genau dann ein Monomorphismus, wenn  $f$  lineare Unabhängigkeit erhält.

$f$  ist genau dann ein Epimorphismus, wenn  $f$  Erzeugendensysteme erhält.

$f$  ist genau dann ein Isomorphismus, wenn  $f$  Basen erhält.

Die Erhaltungseigenschaften bedeuten: Ist  $A \subseteq V$  linear unabhängig (erzeugend, eine Basis), so ist auch  $f[A] = \{f(v) \mid v \in A\}$  linear unabhängig (erzeugend, eine Basis). Wir sagen auch: „Ein Isomorphismus übersetzt Basen in Basen“ usw.

### Beispiele

- (1) Die Projektion  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$  auf die erste Koordinate ist linear.
- (2) Die Vertauschung  $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$  mit  $f(x, y) = (y, x)$  für alle  $(x, y) \in \mathbb{R}^2$  ist linear.
- (3) Die Drehung  $f_\varphi: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ , die  $(x, y) \in \mathbb{R}^2$  auf den um den Winkel  $\varphi$  gegen den Uhrzeigersinn gedrehten Vektor abbildet, ist linear.
- (4) Sei  $K$  ein Körper und seien  $n, m \geq 1$ . Weiter seien  $\alpha_{i,j} \in K$  für alle  $1 \leq i \leq m$  und alle  $1 \leq j \leq n$ . Wir definieren  $f: K^n \rightarrow K^m$  durch

$$f(x_1, \dots, x_n) = (y_1, \dots, y_m), \quad \text{wobei}$$

$$y_1 = \alpha_{1,1} x_1 + \alpha_{1,2} x_2 + \dots + \alpha_{1,n} x_n,$$

$$y_2 = \alpha_{2,1} x_1 + \alpha_{2,2} x_2 + \dots + \alpha_{2,n} x_n,$$

...

$$y_m = \alpha_{m,1} x_1 + \alpha_{m,2} x_2 + \dots + \alpha_{m,n} x_n.$$

Dann ist  $f$  linear. Wir werden in 4.7 sehen, dass jede lineare Abbildung zwischen den Vektorräumen  $K^n$  und  $K^m$  so definiert werden kann.

- (5) Seien  $[a, b] \subseteq \mathbb{R}$  und  $p \in [a, b]$ . Wir betrachten den  $\mathbb{R}$ -Vektorraum

$$V = \{f: [a, b] \rightarrow \mathbb{R} \mid f \text{ ist differenzierbar in } p\} \quad \text{und}$$

$$D: V \rightarrow \mathbb{R} \quad \text{mit} \quad D(f) = f'(p) \quad \text{für alle } f \in V.$$

Dann ist  $D$  linear. Ebenso ist für  $W = \{f: [a, b] \rightarrow \mathbb{R} \mid f \text{ ist integrierbar}\}$  die Abbildung  $I: W \rightarrow \mathbb{R}$  linear, wobei

$$I(f) = \int_a^b f(x) dx \quad \text{für alle } f \in W.$$

### Exkurs: Die Skalierungseigenschaft muss gefordert werden

Ist  $f: (V, +) \rightarrow (W, +)$  ein Homomorphismus und  $K \supseteq \mathbb{Q}$ , so gilt

$$f(qv) = q f(v) \quad \text{für alle } q \in \mathbb{Q} \text{ und } v \in V,$$

wie man durch Verallgemeinerung von

$$f(v+v) = f(v) + f(v) = 2f(v) \quad \text{und} \quad f(v) = f(v/2 + v/2) = 2f(v/2)$$

zeigt. Mit Hilfe einer Hamel-Basis des  $\mathbb{Q}$ -Vektorraumes  $\mathbb{R}$  lässt sich ein Homomorphismus konstruieren, der die Skalierungseigenschaft verletzt. Sei hierzu  $f: \mathbb{R} \rightarrow \mathbb{Q}$  wie in 3.9 additiv und unstetig mit  $\text{Bild}(f) = \mathbb{Q}$ . Dann gilt (a), aber (b) ist verletzt.

Denn für  $v \in \mathbb{R}$  mit  $f(v) = 1$  gilt  $f(\sqrt{2}v) \neq \sqrt{2}f(v) = \sqrt{2}$ , da  $f(\sqrt{2}v) \in \mathbb{Q}$ .

## 4.6 Konstruktion linearer Abbildungen

### Satz (Konstruktionssatz)

Seien  $V, W$   $K$ -Vektorräume,  
 $(v_i)_{i \in I}$  eine Basis von  $V$  und  
 $(w_i)_{i \in I}$  eine Familie in  $W$ .  
 Dann gibt es genau eine  
 lineare Abbildung  $f : V \rightarrow W$   
 mit der Eigenschaft

$$f(v_i) = w_i \quad \text{für alle } i \in I.$$

Anders formuliert:

*Die Werte einer linearen Abbildung  
 $f$  lassen sich auf einer Basis  
 beliebig vorschreiben, und  $f$  ist durch  
 diese Werte eindeutig bestimmt.*

Zum Nachweis der Existenz  
 setzen wir

$$f(v) = \sum_{i \in I} \alpha_i w_i \quad \text{für alle } v = \sum_{i \in I} \alpha_i v_i \in V.$$

Durch die Eindeutigkeit der Darstellung von Vektoren in  $V$  bzgl.  $(v_i)_{i \in I}$  entsteht so eine wohldefinierte Abbildung  $f : V \rightarrow W$  mit  $f(v_i) = w_i$  für alle  $i \in I$ . Man überprüft leicht, dass  $f$  linear ist. Sind umgekehrt  $f, g : V \rightarrow W$  linear mit

$$f(v_i) = w_i = g(v_i) \quad \text{für alle } i \in I,$$

so gilt für alle  $v = \sum_{i \in I} \alpha_i v_i \in V$ , dass

$$f(v) = \sum_{i \in I} \alpha_i f(v_i) = \sum_{i \in I} \alpha_i w_i = \sum_{i \in I} \alpha_i g(v_i) = g(v).$$

### Beispiele

- (1) Sei  $(e_1, e_2, e_3)$  die kanonische Basis des  $\mathbb{R}^3$ . Dann gibt es genau eine lineare Abbildung  $f : \mathbb{R}^3 \rightarrow \mathbb{R}^4$  mit

$$f(e_1) = (1, 0, 1, 1), \quad f(e_2) = (1, 0, 1, 1), \quad f(e_3) = (0, 1, 0, 1).$$

- (2) Seien  $f, g : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  lineare Abbildungen mit

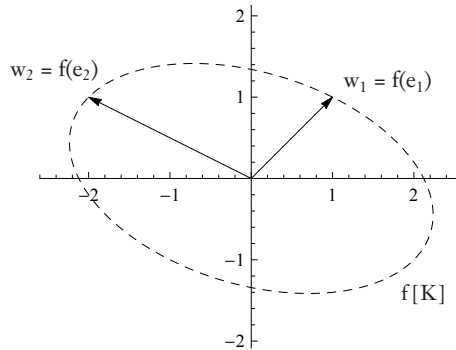
$$f(1, 1) = g(1, 1), \quad f(1, 2) = g(1, 2).$$

Dann gilt  $f = g$ , da  $(1, 1)$  und  $(1, 2)$  eine Basis des  $\mathbb{R}^2$  bilden.

- (3) Sei  $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  eine lineare Abbildung mit

$$f(1, 0) = (0, 1), \quad f(0, 1) = (-1, 0).$$

Dann ist  $f$  die Drehung um  $\pi/2$  gegen den Uhrzeigersinn.



Es gibt genau eine lineare Abbildung  $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  mit  $f(e_1) = w_1 = (1, 1)$  und  $f(e_2) = w_2 = (-2, 1)$ . Damit ist durch die beiden Werte insbesondere das Bild  $f[K]$  des Einheitskreises  $K = \{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1\}$  festgelegt. Wir werden in Kapitel 8 zeigen, dass  $f[K]$  für jede Wahl von  $w_1$  und  $w_2$  eine Ellipse ist.

Eine wichtige Folgerung des Konstruktionssatzes ist:

### Fortsetzungssatz für lineare Abbildungen

Seien  $V, W$   $K$ -Vektorräume und sei  $U$  ein Unterraum von  $V$ . Weiter sei  $f: U \rightarrow W$  linear. Dann gibt es eine lineare Abbildung  $g: V \rightarrow W$  mit  $g|_U = f$ .

Ergänzen wir nämlich eine Basis  $(v_i)_{i \in J}$  von  $U$  zu einer Basis  $(v_i)_{i \in I}$ ,  $I \supseteq J$ , von  $V$  nach dem Basisergänzungssatz (vgl. 3.7 und 3.9), so ist, mit einer beliebigen Familie  $(w_i)_{i \in I-J}$  in  $W$ , die eindeutige lineare Abbildung  $g: V \rightarrow W$  mit

$$g(v_i) = \begin{cases} f(v_i) & \text{falls } i \in J, \\ w_i & \text{falls } i \in I - J \end{cases}$$

wie gewünscht. Speziell gilt dies für  $(w_i)_{i \in I-J}$  mit  $w_i = 0$  für alle  $i \in I - J$ .

### Beispiele

- (1) Seien  $V = W = \mathbb{R}^3$ ,  $U = \mathbb{R}^2 \times \{0\}$  und sei  $f: U \rightarrow \mathbb{R}^3$  die Drehung in der  $x$ - $y$ -Ebene um  $\pi/2$  gegen den Uhrzeigersinn. Wir betrachten nun die aus  $e_1 = (1, 0, 0)$  und  $e_2 = (0, 1, 0)$  gebildete Basis von  $U$  und ergänzen diese um  $e_3 = (0, 0, 1)$  zu einer Basis von  $\mathbb{R}^3$ . Der Vektor  $w_3 = e_3$  liefert als Fortsetzung  $g: \mathbb{R}^3 \rightarrow \mathbb{R}^3$  von  $f$  die Drehung um  $\pi/2$  um die  $z$ -Achse im  $\mathbb{R}^3$  gegen den Uhrzeigersinn. Das Bild von  $g$  ist  $\mathbb{R}^3$ . Der Vektor  $w_3 = 0$  liefert dagegen als Fortsetzung  $g$  die Projektion im  $\mathbb{R}^3$  „ $(x, y, z)$  nach  $(x, y, 0)$ “ auf die  $x$ - $y$ -Ebene, gefolgt von der Drehung  $f$  um  $\pi/2$  gegen den Uhrzeigersinn. Das Bild von  $g$  ist hier  $U$ .

- (2) Seien  $V = W = \mathbb{R}^2$ ,  $U = \mathbb{R} \times \{0\}$  und  $f: U \rightarrow \mathbb{R}^2$  die Identität auf  $U$ , sodass

$$f(x, 0) = (x, 0) \quad \text{für alle } x \in \mathbb{R}.$$

Wir ergänzen die Basis  $(e_1)$  von  $U$  zur kanonischen Basis  $(e_1, e_2)$  von  $V$ . Dann liefert der Vektor  $w_2 = 0$  die Fortsetzung  $g: \mathbb{R}^2 \rightarrow \mathbb{R}^2$  von  $f$  mit

$$g(x, y) = g((x, 0) + (0, y)) = (x, 0) + y g(0, 1) = (x, 0) \quad \text{für alle } (x, y) \in \mathbb{R}^2.$$

Ergänzen wir dagegen  $(e_1)$  zur Basis  $(e_1, (1, 1))$  des  $\mathbb{R}^2$ , so liefert der Vektor  $w_2 = 0$  die Fortsetzung  $g$  von  $f$  mit

$$g(x, y) = g((x - y, 0) + y(1, 1)) = (x - y, 0) + y g(1, 1) = (x - y, 0)$$

$$\text{für alle } (x, y) \in \mathbb{R}^2.$$

Wir halten also fest:

### Warnung

Für die Vorgabe „ $w_i = 0$  für alle  $i \in I - J$ “ gilt im Allgemeinen nicht, dass  $g(v) = 0$  für alle  $v \in V - U$ . Weiter hängt die Fortsetzung  $g$  von  $f$  auch für diese Vorgabe in der Regel von der Basis  $(v_i)_{i \in I}$  ab. Man kann also nicht von *der* Nullfortsetzung von  $f$  sprechen. Eindeutig ist  $g =$  „die Nullfortsetzung von  $f$  bzgl. der Basis  $(v_i)_{i \in I}$ “.

## 4.7 Darstellung linearer Abbildungen

### Satz (Darstellungssatz)

Seien  $K$  ein Körper und  $n, m \geq 1$ . Weiter sei  $f: K^n \rightarrow K^m$  linear. Dann gibt es eindeutig bestimmte  $\alpha_{i,j} \in K$ ,  $1 \leq i \leq m$ ,  $1 \leq j \leq n$ , sodass

$$f(x_1, \dots, x_n) = (y_1, \dots, y_m) \text{ mit}$$

$$y_1 = \alpha_{1,1} x_1 + \alpha_{1,2} x_2 + \dots + \alpha_{1,n} x_n,$$

$$y_2 = \alpha_{2,1} x_1 + \alpha_{2,2} x_2 + \dots + \alpha_{2,n} x_n,$$

...

$$y_m = \alpha_{m,1} x_1 + \alpha_{m,2} x_2 + \dots + \alpha_{m,n} x_n.$$

Ist  $e_j$  der  $j$ -te kanonische Einheitsvektor des  $K^n$ , so gilt

$$(+)\ f(e_j) = (\alpha_{1,j}, \dots, \alpha_{m,j}) = \text{„}j\text{-te Spalte des rechteckigen } \alpha_{i,j}\text{-Schemas“}.$$

Zum Beweis verwenden wir (+) zur Definition der  $\alpha_{i,j}$ . Dann gilt für alle  $(x_1, \dots, x_n) \in K^n$ :

$$f(x_1, \dots, x_n) = f(x_1 e_1 + \dots + x_n e_n) = x_1 f(e_1) + \dots + x_n f(e_n) =$$

$$x_1 (\alpha_{1,1}, \dots, \alpha_{m,1}) + \dots + x_n (\alpha_{1,n}, \dots, \alpha_{m,n}) =$$

$$(x_1 \alpha_{1,1} + \dots + x_n \alpha_{1,n}, \dots, x_1 \alpha_{m,1} + \dots + x_n \alpha_{m,n}).$$

Damit gelten die  $y$ -Gleichungen des Satzes (wobei wir dort der Konvention folgen, die  $\alpha_{i,j}$  vor den  $x_j$  zu notieren). Einsetzen der Basisvektoren  $e_j$  für  $(x_1, \dots, x_n)$  in die Gleichungen zeigt, dass die  $\alpha_{i,j}$  eindeutig bestimmt sind.

### Beispiel

Sei  $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$  die Drehung um  $\pi/4$  gegen den Uhrzeigersinn. Mit  $\beta = 1/\sqrt{2}$  gilt

$$f(1, 0) = \beta (1, 1),$$

$$f(0, 1) = \beta (-1, 1).$$

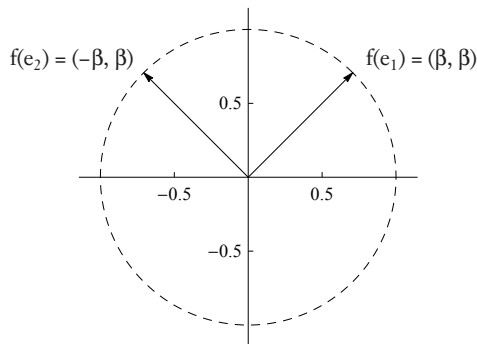
Folglich ist

$$f(x, y) = (\bar{x}, \bar{y})$$

mit

$$\bar{x} = \beta x - \beta y,$$

$$\bar{y} = \beta x + \beta y.$$



Die Drehung um  $\pi/4$  gegen den Uhrzeigersinn ist bestimmt durch die Bilder der Basisvektoren  $e_1$  und  $e_2$ . Mit Hilfe der Koordinaten dieser beiden Werte können alle Werte leicht berechnet werden.



Der Darstellungssatz gilt allgemeiner in der folgenden Form:

### Allgemeiner Darstellungssatz

Sei  $f: V \rightarrow W$  eine lineare Abbildung, und seien  $(v_j)_{j \in J}$  und  $(w_i)_{i \in I}$  Basen von  $V$  bzw.  $W$ . Dann gibt es eindeutige Skalare  $\alpha_{i,j}$ ,  $(i, j) \in I \times J$ , sodass

$$f\left(\sum_{j \in J} \lambda_j v_j\right) = \sum_{i \in I} \mu_i w_i, \quad \text{mit}$$

$$\mu_i = \sum_{j \in J} \alpha_{i,j} \lambda_j \quad \text{für alle } i \in I.$$

Die  $\alpha_{i,j}$  sind definiert durch

$$(+)\quad f(v_j) = \sum_{i \in I} \alpha_{i,j} w_i$$

für alle  $j \in J$ , d. h., die Spalten des  $\alpha_{i,j}$ -Schemas sind die Koordinatenvektoren bzgl.  $(w_i)_{i \in I}$  der Bilder der Basisvektoren  $(v_j)_{j \in J}$ .

f	$v_1$	$v_2$	...	$v_j$	...
$w_1$	$\alpha_{1,1}$	$\alpha_{1,2}$	...	$\alpha_{1,j}$	...
$w_2$	$\alpha_{2,1}$	$\alpha_{2,2}$	...	$\alpha_{2,j}$	...
...	...	...	...	...	...
$w_i$	$\alpha_{i,1}$	$\alpha_{i,2}$	...	$\alpha_{i,j}$	...
...	...	...	...	...	...

$\alpha_{i,j}$  ist der  $w_i$ -Anteil von  $f(v_j)$  bzgl.  $(w_i)_{i \in I}$ .

Merkregel: Bei der  $\alpha_{i,j}$ -Darstellung von  $f: V \rightarrow W$  verweist der Index  $j$  immer auf  $V$  und der Index  $i$  immer auf  $W$ :  $v_j \in V$ ,  $w_i \in W$ .

Obiger Darstellungssatz entspricht dem Spezialfall

$$V = K^n, \quad (v_j)_{j \in J} = (e_1, \dots, e_j, \dots, e_n) = \text{„die kanonische Basis des } K^n\text{“},$$

$$W = K^m, \quad (w_i)_{i \in I} = (e_1, \dots, e_i, \dots, e_m) = \text{„die kanonische Basis des } K^m\text{“}.$$

Im allgemeinen Satz können  $V$  und  $W$  unendlich-dimensional sein. Da auch die Basen beliebig sind, liefert dieser Satz aber auch im Endlich-Dimensionalen etwas Neues:

### Beispiel

Sei wieder  $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$  die Drehung um  $\pi/4$  gegen den Uhrzeigersinn. Seien

$$v_1 = (1, 0), \quad v_2 = \beta(1, 1), \quad w_1 = \beta(1, 1), \quad w_2 = \beta(-1, 1),$$

wobei  $\beta = 1/\sqrt{2}$ . Für die Basen  $(v_1, v_2)$  von  $V = \mathbb{R}^2$  und  $(w_1, w_2)$  von  $W = \mathbb{R}^2$  gilt

$$f(v_1) = w_1 = 1 w_1 + 0 w_2,$$

$$f(v_2) = (0, 1) = \frac{1}{2\beta} (0, 2\beta) = \frac{1}{2\beta} w_1 + \frac{1}{2\beta} w_2.$$

Damit lauten die  $\alpha_{i,j}$  gemäß „Koordinatenvektoren der Bilder liefern die Spalten“:

$$\alpha_{1,1} = 1, \quad \alpha_{1,2} = \frac{1}{2\beta},$$

$$\alpha_{2,1} = 0, \quad \alpha_{2,2} = \frac{1}{2\beta}.$$

## 4.8 Fasern und lineare Gleichungssysteme

**Definition** (Urbildmenge  $L_f(w)$  eines Vektors unter einer linearen Abbildung)

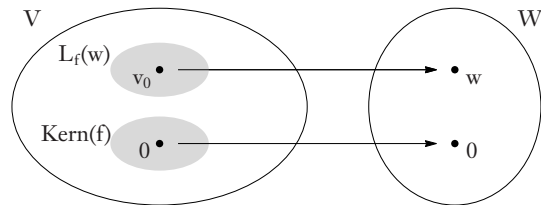
Seien  $V, W$   $K$ -Vektorräume, und sei  $f: V \rightarrow W$  linear. Weiter sei  $w \in W$ . Dann setzen wir

$$L_f(w) = f^{-1}[\{w\}] = \{v \in V \mid f(v) = w\}. \quad (\text{Faser von } f \text{ über } w)$$

Der Kern

$$\text{Kern}(f) = \{v \in V \mid f(v) = 0\}$$

einer linearen Abbildung besteht aus allen Vektoren von  $V$ , die auf den Nullvektor abgebildet werden. Nun lassen wir anstelle des Nullvektors einen beliebigen Vektor  $w$  aus  $W$  zu und sammeln in



$$L_f(w) = v_0 + \text{Kern}(f)$$

$$L_f(w) = \{v \in V \mid f(v) = w\}$$

alle Vektoren von  $V$ , die auf  $w$  abgebildet werden. Es gilt  $\text{Kern}(f) = L_f(0)$ . Die Menge  $L_f(w)$  ist nichts anderes als die in 1.7 eingeführte Faser von  $f$  über  $w$ . Wir werden gleich sehen, dass Fasern eng mit der Lösung von linearen Gleichungssystemen verknüpft sind, was die Wahl des Buchstabens „ $L$ “ (für Lösungsmenge) motiviert.

Wir hatten gezeigt, dass der Kern einer linearen Abbildung ein Unterraum von  $V$  ist. Speziell ist  $0 \in L_f(0)$ . Allgemeine Fasern  $L_f(w)$  können dagegen leer sein. Es gilt:

$L_f(w)$ ist ein affiner Unterraum von $V$ .
<p>Ist <math>L_f(w) \neq \emptyset</math> und <math>v_0</math> ein beliebiges Element von <math>L_f(w)</math>, so gilt</p> $L_f(w) = v_0 + L_f(0) = v_0 + \text{Kern}(f).$ <p>Ist also <math>\text{Kern}(f)</math> endlich-dimensional, so ist <math>\dim(L_f(w)) = \dim(\text{Kern}(f))</math>.</p>

Die Aussagen ergeben sich aus dem Homomorphiesatz, lassen sich aber auch direkt einsehen: Für  $L_f(w) = \emptyset$  ist nichts zu zeigen. Sei also  $v_0 \in L_f(w)$ . Ist nun  $v \in \text{Kern}(f)$ , so gilt

$$f(v_0 + v) = f(v_0) + f(v) = w + 0 = w,$$

sodass  $v_0 + v \in L_f(w)$ . Damit ist  $v_0 + \text{Kern}(f) \subseteq L_f(w)$ . Ist umgekehrt  $v \in L_f(w)$ , so gilt

$$f(v - v_0) = f(v) - f(v_0) = w - w = 0,$$

sodass  $v = v_0 + (v - v_0) \in v_0 + \text{Kern}(f)$ . Damit ist auch  $L_f(w) \subseteq v_0 + \text{Kern}(f)$ . Dies zeigt, dass  $L_f(w)$  der durch „beliebiges Element + Kern“ gegebene affine Unterraum von  $V$  ist.

**Beispiele**

- (1) Sei  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$  die Projektion auf die erste Komponente. Dann gilt  
 $L_f(1) = (1, 0) + \text{Kern}(f) = (1, 4) + \text{Kern}(f)$  mit  $\text{Kern}(f) = \{ (0, y) \mid y \in \mathbb{R} \}$ .
- (2) Sei  $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$  die Drehung um  $\pi/2$ ,  $f(x, y) = (-y, x)$ . Dann gilt  
 $L_f(1, 1) = (1, -1) + \text{Kern}(f) = (1, -1) + \{0\} = \{(1, -1)\}$ .
- (3) Sei  $f: \mathbb{R}^3 \rightarrow \mathbb{R}^2$  definiert durch  $f(x, y, z) = (x, x)$ . Dann gilt  $L_f(0, 1) = \emptyset$ .

**Zusammenhang mit linearen Gleichungssystemen**

Ein lineares Gleichungssystem wird oft in der Form

$$\alpha_{1,1} x_1 + \alpha_{1,2} x_2 + \dots + \alpha_{1,n} x_n = b_1$$

...

$$\alpha_{m,1} x_1 + \alpha_{m,2} x_2 + \dots + \alpha_{m,n} x_n = b_m$$

präsentiert, mit gegebenen Elementen  $\alpha_{i,j}$ ,  $b_i$  eines Körpers  $K$ . Die  $\alpha_{i,j}$  heißen dann die *Koeffizienten* und  $b = (b_1, \dots, b_m) \in K^m$  die *rechte Seite* oder der *Zielvektor* des Systems. Ist  $b = 0$ , so heißt das System *homogen*, andernfalls heißt es *inhomogen*. Die *Lösungsmenge*  $L$  des Systems besteht aus allen Vektoren  $x = (x_1, \dots, x_n) \in K^n$ , für die alle Gleichungen erfüllt sind. Definieren wir nun  $f: K^n \rightarrow K^m$  durch die linke Seite des Systems (sodass die  $\alpha_{i,j}$  die darstellenden Koeffizienten von  $f$  sind, vgl. 4.7), so gilt

$$L = L_f(b), \text{ wobei } b = (b_1, \dots, b_m).$$

Damit kann man ein Gleichungssystem auch kompakt in der Form

$$f(x) = b \quad (\text{Abbildungsnotation für Gleichungssysteme})$$

notieren. Unsere Ergebnisse zeigen:  $L$  ist ein affiner Unterraum des  $K^n$ . Ist  $L \neq \emptyset$  und  $x_0$  irgendeine Lösung des Systems (eine „spezielle Lösung“), so gilt

$$L = x_0 + L_0, \quad (\text{Lösungsmenge} = \text{spezielle Lösung} + \text{homogene Lösungsmenge})$$

wobei  $L_0 = L_f(0)$  die Lösungsmenge des zugeordneten homogenen Systems ist, bei dem die  $\alpha_{i,j}$  gleich bleiben, aber alle  $b_i = 0$  sind.

**Beispiele**

- (1) Beispiel (1) oben entspricht dem System  $1x + 0y = 1$ .
- (2) Beispiel (2) oben entspricht dem System:
- $$\begin{array}{rcl} 0x - 1y & = & 1 \\ 1x + 0y & = & 1 \end{array}$$
- (3) Beispiel (3) oben entspricht dem unlösbaren System:
- $$\begin{array}{rcl} 1x + 0y + 0z & = & 0 \\ 1x + 0y + 0z & = & 1 \end{array}$$

In den Spalten der Systeme stehen die Bilder der Basisvektoren  $e_1, \dots, e_n$  unter  $f$ .

## 4.9 Isomorphie von Vektorräumen

### Satz (Isomorphiesätze)

#### Isomorphiesatz für endlich-dimensionale Vektorräume

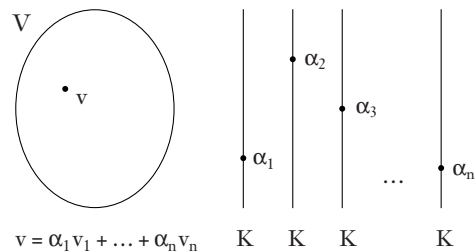
Ist  $V$  ein endlich-dimensionaler  $K$ -Vektorraum und  $n = \dim(V)$ , so ist  $V$  isomorph zum  $K$ -Vektorraum  $K^n$ .

#### Allgemeiner Isomorphiesatz

Ist  $V$  ein  $K$ -Vektorraum und  $(v_i)_{i \in I}$  eine Basis von  $V$ , so ist  $V$  isomorph zum  $K$ -Vektorraum  $K^{(I)}$  und weiter zu jedem Vektorraum  $K^{(J)}$  mit  $|I| = |J|$ . Insbesondere gilt: Zwei  $K$ -Vektorräume  $V$  und  $W$  sind genau dann isomorph, wenn sie gleichmächtige Basen besitzen, d. h., wenn es eine Basis  $(v_i)_{i \in I}$  von  $V$ , eine Basis  $(w_j)_{j \in J}$  von  $W$  und eine Bijektion  $b : I \rightarrow J$  gibt.

Die Vektorräume  $K^n$  und allgemeiner  $K^{(I)}$  sind Könige im Reich aller  $K$ -Vektorräume. Bis auf die „Namen der Vektoren“ ist jeder endlich-dimensionale Vektorraum ein  $K^n$  und jeder unendlich-dimensionale Vektorraum ein  $K^{(I)}$  mit  $I = \mathbb{N}, \mathbb{R}$  usw. Man sagt auch:

*In den Isomorphieklassen der  $K$ -Vektorräume gibt es kanonische Repräsentanten.*



Ist  $V$   $n$ -dimensional, so liefert der Übergang von einem Vektor  $v \in V$  zu seinem Koordinatenvektor  $(\alpha_1, \dots, \alpha_n) \in K^n$  bzgl. einer Basis  $(v_1, \dots, v_n)$  von  $V$  einen Isomorphismus zwischen  $V$  und  $K^n$ . Die Senkrechten des Diagramms kann man sich als Regler eines Mischpults vorstellen, mit denen man alle Vektoren in  $V$  einstellen kann.

### Beispiele

- (1) Ist  $V$   $n$ -dimensional, so gilt  $V \cong K^n$ . Ist also  $K$  endlich, so hat  $V$  genau  $|K|^n$ -viele Vektoren.

- (2) Für den  $\mathbb{R}$ -Vektorraum  $\mathbb{C}^n$  ist eine Basis gegeben durch

$$e_1 = (1, 0, \dots, 0), \dots, e_n = (0, \dots, 0, 1), e_{n+1} = (i, 0, \dots, 0), e_{2n} = (0, \dots, 0, i).$$

Damit ist der  $\mathbb{R}$ -Vektorraum  $\mathbb{C}^n$  isomorph zum  $\mathbb{R}$ -Vektorraum  $\mathbb{R}^{2n}$ .

- (3) Ist  $V$  ein  $K$ -Vektorraum mit einer abzählbar unendlichen Basis, so ist  $V$  isomorph zum  $K$ -Vektorraum  $K[X] = K^{(\mathbb{N})}$  aller Polynome über  $K$ .

Hinsichtlich des endlich-dimensionalen Satzes betrachten wir eine Basis  $B = (v_1, \dots, v_n)$  von  $V$  und die Koordinatenabbildung  $\Phi_B : V \rightarrow K^n$  mit  $\Phi(v_i) = e_i$  für alle  $i$ , d. h.

$$\Phi_B(\alpha_1 v_1 + \dots + \alpha_n v_n) = \alpha_1 e_1 + \dots + \alpha_n e_n = (\alpha_1, \dots, \alpha_n).$$

Diese Zuordnung ist bijektiv, da jedem Vektor genau ein Koordinatenvektor entspricht und umgekehrt (vgl. 3.6).

Ist allgemeiner  $(v_i)_{i \in I}$  eine Basis von  $V$  und  $B = (e_i)_{i \in I}$  die kanonische Basis des  $K^{(I)}$ , so ist  $\Phi_B : V \rightarrow K^{(I)}$  mit  $\Phi_B(v_i) = e_i$  für alle  $i \in I$  bijektiv, sodass  $V$  und  $K^{(I)}$  isomorph sind. Ist  $b : I \rightarrow J$  bijektiv, so ist auch die lineare Abbildung  $g : K^{(I)} \rightarrow K^{(J)}$  mit  $g(e_i) = e_{b(i)}$  für alle  $i \in I$  bijektiv, sodass  $K^{(I)}$  und  $K^{(J)}$  isomorph sind.

Das Ergebnis ist so stark, dass man fast ein wenig enttäuscht sein könnte. Konzentriert man sich auf endlich-dimensionale  $K$ -Vektorräume mit den Skalarenkörpern  $K = \mathbb{R}$  oder  $K = \mathbb{C}$ , so gibt es bis auf Isomorphie nur die Beispiele

$$\mathbb{R}^0, \mathbb{R}^1, \mathbb{R}^2, \dots, \mathbb{R}^n, \dots \quad \text{und} \quad \mathbb{C}^0, \mathbb{C}^1, \mathbb{C}^2, \dots, \mathbb{C}^n, \dots$$

So viel Aufwand für so wenig? Die Skepsis ist nicht berechtigt:

- (a) Dass die Welt einfacher ist, als sie sein könnte, bleibt erfreulich.
- (b) Ohne den allgemeinen Vektorraumbegriff kann man gar nicht sehen, dass viele Strukturen bis auf Isomorphie der  $\mathbb{R}^n$ ,  $\mathbb{C}^n$  oder allgemeiner der  $K^n$  sind (man denke etwa an die Polynome über  $K$  vom Grad kleiner als  $n$ ).
- (c) Der  $K^n$  stellt zwar Codes für Vektoren in  $V$  zur Verfügung, kann aber oft  $V$  nicht vollständig ersetzen, da dadurch eine auf  $V$  vorhandene zusätzliche Struktur verloren gehen würde.
- (d) Für Vektorräume wie den  $\mathbb{R}^{\mathbb{N}}$  oder  $\mathbb{R}^{\mathbb{R}}$ , die eine überabzählbare Basis besitzen, bleibt der Isomorphiesatz abstrakt (vgl. den folgenden Exkurs).

### Exkurs: Basen des $K^I$ für unendliche Indexmengen $I$

Ist  $B$  eine Basis eines unendlich-dimensionalen  $K$ -Vektorraums, so sind die Mengen  $V$  und  $B \times K$  gleichmächtig. (Beweisidee: Eine Basis  $B$  kodiert alle Vektoren in  $V$  durch Tupel der Form  $(b_1, \dots, b_n, \alpha_1, \dots, \alpha_n) \in B^n \times K^n$ ,  $n \in \mathbb{N}$ , und davon gibt es genau  $B \times K$  viele, wenn  $B$  oder  $K$  unendlich ist.) Ist nun  $K^I$  ein  $K$ -Vektorraum mit einer unendlichen Indexmenge  $I$ , so existiert eine linear unabhängige Menge der Mächtigkeit von  $K$ , etwa  $\{g_\alpha : I \rightarrow K \mid \alpha \in K\}$  mit

$$g_\alpha(n) = \alpha^n \quad \text{für alle } n \in \mathbb{N}, \quad g_\alpha(i) = 0 \quad \text{für alle } i \in I - \mathbb{N},$$

wobei wir ohne Einschränkung  $\mathbb{N} \subseteq I$  annehmen. Also ist die Mächtigkeit einer Basis  $B$  von  $K^I$  größergleich der Mächtigkeit von  $K$  und damit gilt

$$|K^I| = |B \times K| = |B|. \quad (\text{Satz von Erdős-Kaplansky})$$

Der  $\mathbb{R}$ -Vektorraum  $\mathbb{R}^{\mathbb{N}}$  aller unendlichen reellen Folgen hat also Basen der Mächtigkeit von  $\mathbb{R}^{\mathbb{N}}$ . Da  $\mathbb{R}^{\mathbb{N}}$ ,  $\mathbb{R}$  und  $\mathcal{P}(\mathbb{N}) = \{A \mid A \subseteq \mathbb{N}\}$  gleichmächtig sind, gilt also

$$\mathbb{R}^{\mathbb{N}} \cong \mathbb{R}^{(\mathbb{R}^{\mathbb{N}})} \cong \mathbb{R}^{(\mathbb{R})} \cong \mathbb{R}^{(\mathcal{P}(\mathbb{N}))}.$$

Analog hat der  $\mathbb{R}$ -Vektorraum  $\mathbb{R}^{\mathbb{R}}$  aller reellen Funktionen Basen der Mächtigkeit von  $\mathbb{R}^{\mathbb{R}}$ . Die Mengen  $\mathbb{R}^{\mathbb{R}}$  und  $\mathcal{P}(\mathbb{R}) = \{A \mid A \subseteq \mathbb{R}\}$  sind gleichmächtig, sodass

$$\mathbb{R}^{\mathbb{R}} \cong \mathbb{R}^{(\mathbb{R}^{\mathbb{R}})} = \mathbb{R}^{(\mathcal{P}(\mathbb{R}))}.$$

## 4.10 Die Dimensionsformel

### Satz (Dimensionsformel für lineare Abbildungen)

Seien  $V, W$  endlich-dimensionale  $K$ -Vektorräume, und sei  $f : V \rightarrow W$  linear. Dann gilt

$$\dim(V) = \dim(\text{Kern}(f)) + \dim(\text{Bild}(f)). \quad (\text{Dimensionsformel})$$

Ist die Dimension  $n$  eines  $K$ -Vektorraumes  $V$  einmal bestimmt, so erleichtert die Dimensionsformel die Untersuchung linearer Abbildungen von  $V$  in einen beliebigen anderen  $K$ -Vektorraum  $W$ . Kennt man nämlich  $m = \dim(\text{Kern}(f))$ , so kennt man  $\dim(\text{Bild}(f)) = n - m$ .

Analog errechnet sich die Dimension des Kerns aus der des Bildes.

Die Addition auf der rechten Seite der Formel soll nicht darüber hinwegtäuschen, dass die Dimensionen in zwei verschiedenen Vektorräumen berechnet werden, wenn  $f$  kein Endomorphismus ist.

Um die Dimensionsformel einzusehen, betrachten wir eine Basis  $(u_1, \dots, u_m)$  des Unterraums  $U = \text{Kern}(f)$  von  $V$ . Ist  $m = n = \dim(V)$ , so ist  $U = V$  und damit  $\text{Bild}(f) = \{0\}$  und die Aussage „ $n = n + 0$ “ der Dimensionsformel klar. Andernfalls ergänzen wir die Basis von  $U$  zu einer Basis  $(u_1, \dots, u_m, v_1, \dots, v_k)$  von  $V$ , sodass  $n = m + k$ . Für alle

$$u = \alpha_1 u_1 + \dots + \alpha_m u_m, \quad v = \beta_1 v_1 + \dots + \beta_k v_k$$

in  $V$  gilt dann

$$(+)\quad f(u + v) = f(u) + f(v) = 0 + f(v) = f(v) = \beta_1 f(v_1) + \dots + \beta_k f(v_k).$$

Wir setzen nun  $w_1 = f(v_1), \dots, w_k = f(v_k)$ . Dann folgt aus (+):

(a)  $\text{Bild}(f) = \text{span}(w_1, \dots, w_k)$ . Denn jeder Vektor  $f(u + v)$  des Bildes hat die Form

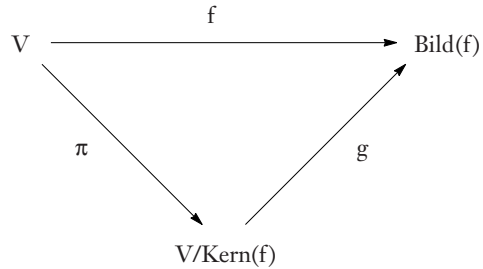
$$f(u + v) = f(v) = \beta_1 f(v_1) + \dots + \beta_k f(v_k),$$

(b)  $(w_1, \dots, w_k)$  ist linear unabhängig in  $W$ . Denn sind  $\beta_1, \dots, \beta_k \in K$  mit

$$f(v) = \beta_1 f(v_1) + \dots + \beta_k f(v_k) = 0,$$

so ist  $v \in U = \text{Kern}(f)$  und damit  $\beta_1 = \dots = \beta_k = 0$ .

Damit ist  $(w_1, \dots, w_k)$  eine Basis von  $\text{Bild}(f)$ , sodass  $k = \dim(\text{Bild}(f))$ .



Anwendung des Homomorphiesatzes: Ist  $(u_1, \dots, u_m)$  eine Basis von  $U = \text{Kern}(f)$  und  $(u_1, \dots, u_m, v_1, \dots, v_k)$  eine Basis von  $V$ , so ist  $(v_1 + U, \dots, v_k + U)$  eine Basis von  $V/U$  (vgl. 3.11). Da  $g : V/U \rightarrow \text{Bild}(f)$  ein Isomorphismus ist, gilt  $\dim(\text{Bild}(f)) = \dim(V/U) = k = \dim(V) - \dim(U)$ .

Die Dimensionsformel lässt sich auch durch Anwendung des Isomorphiesatzes für Vektorräume beweisen (vgl. das obige Diagramm).

### Beispiele

- (1) Sei  $f: \mathbb{R}^{12} \rightarrow \mathbb{R}^7$  ein Epimorphismus. Dann gilt  $\dim(\text{Kern}(f)) = 5$ .
- (2) Sei  $f: V \rightarrow W$  ein Epimorphismus zwischen endlich-dimensionalen Vektorräumen. Dann gilt
 
$$\dim(W) = \dim(V) - \dim(\text{Kern}(f)) \leq \dim(V).$$
- (3) Sind  $f: V \rightarrow W$  und  $g: W \rightarrow U$  Epimorphismen, so gilt
 
$$\begin{aligned} \dim(V) &= \dim(\text{Kern}(f)) + \dim(\text{Bild}(f)) = \\ &\dim(\text{Kern}(f)) + \dim(W) = \\ &\dim(\text{Kern}(f)) + \dim(\text{Kern}(g)) + \dim(\text{Bild}(g)) = \\ &\dim(\text{Kern}(f)) + \dim(\text{Kern}(g)) + \dim(U). \end{aligned}$$
- (4) Sind  $f_i: V_i \rightarrow V_{i+1}$  Epimorphismen für  $1 \leq i \leq n$  mit  $V_{n+1} = \{0\}$ , so gilt
 
$$\dim(V) = \dim(\text{Kern}(f_1)) + \dim(\text{Kern}(f_2)) + \dots + \dim(\text{Kern}(f_n)).$$

Ein wichtige Anwendung der Dimensionsformel werden wir im nächsten Kapitel kennenlernen („Zeilenrang = Spaltenrang“).

Für endliche Mengen  $A, B$  mit  $|A| = |B|$  und eine Funktion  $f: A \rightarrow B$  sind die Eigenschaften „injektiv“, „surjektiv“, „bijektiv“ nach dem Schubfachprinzip äquivalent (vgl. 1.10). Aus der Dimensionsformel erhalten wir folgendes Analogon für endlich-dimensionale Vektorräume:

Ist  $\dim(V) = \dim(W) < \infty$  und  $f: V \rightarrow W$  linear, so sind äquivalent:

- (a)  $f$  ist ein Monomorphismus.
- (b)  $f$  ist ein Epimorphismus.
- (c)  $f$  ist ein Isomorphismus.

Denn mit  $m = \dim(\text{Kern}(f))$ ,  $k = \dim(\text{Bild}(f))$  ist

$$\dim(W) = \dim(V) = m + k.$$

Folglich gilt

$m = 0$  (d.h.,  $f$  ist ein Monomorphismus) *genau dann, wenn*

$k = \dim(W)$  (d.h.,  $f$  ist ein Epimorphismus).

## 4.11 Lineare Abbildungen als Vektoren

### Definition (lineare Operatoren)

Seien  $V, W$   $K$ -Vektorräume. Dann definieren wir die  $K$ -Vektorräume

$$\operatorname{Hom}(V, W) = \operatorname{Hom}_K(V, W) = \{f : V \rightarrow W \mid f \text{ ist linear}\},$$

$$\operatorname{End}(V) = \operatorname{End}_K(V) = \operatorname{Hom}(V, V).$$

Die Elemente von  $\operatorname{Hom}(V, W)$  heißen auch *lineare Operatoren* von  $V$  nach  $W$ .

Die Menge  $\operatorname{Hom}(V, W)$  ist ein Unterraum des Vektorraums  $W^V$  aller Funktionen von  $V$  nach  $W$ . Für alle  $f, g$  in  $\operatorname{Hom}(V, W)$  und alle  $\alpha \in K$  sind  $f + g, \alpha f : V \rightarrow W$  definiert durch

$$(f + g)(v) = f(v) + g(v),$$

$$(\alpha f)(v) = \alpha f(v) \quad \text{für alle } v \in V.$$

$E_{3,2}$	$v_1$	$v_2$	$v_3$	$v_4$
$w_1$	0	0	0	0
$w_2$	0	0	0	0
$w_3$	0	1	0	0

Sind  $(v_1, v_2, v_3, v_4)$  und  $(w_1, w_2, w_3)$  Basen von  $V$  bzw.  $W$ , so können wir  $E_{2,3} \in \operatorname{Hom}(V, W)$  durch Abbilden von  $v_2$  auf  $w_3$  und Nullfortsetzung definieren. Alle  $E_{i,j}$  bilden eine Basis von  $\operatorname{Hom}(V, W)$  der Länge  $4 \cdot 3 = 12$ .

### Beispiele

- (1) Der Vektorraum  $\operatorname{End}_{\mathbb{R}}(\mathbb{R}^3)$  besteht aus allen linearen Abbildungen des dreidimensionalen Raums in sich selbst. Darunter fallen zum Beispiel Drehungen um eine Achse durch 0, Streckungen, die Spiegelung am Nullpunkt oder an einer Geraden oder Ebene durch den Nullpunkt und Projektionen auf derartige Geraden und Ebenen.
- (2) Die Menge aller Automorphismen  $f : V \rightarrow V$  ist eine Teilmenge von  $\operatorname{End}(V)$ . Sie bildet aber keinen Unterraum von  $\operatorname{End}(V)$ , da die Addition zweier Bijektionen im Allgemeinen keine Bijektion mehr ist. Für die Bijektionen  $f, g : V \rightarrow V$  mit  $f(v) = v, g(v) = -v$  für alle  $v \in V$  gilt zum Beispiel  $f + g = 0$ .

Die Idee, aus linearen Abbildungen, die Vektoren eines Vektorraums  $V$  auf Vektoren eines Vektorraumes  $W$  abbilden, einen Vektorraum zu konstruieren, dessen Vektoren also lineare Abbildungen sind, ist sicher gewöhnungsbedürftig. Derartige Konstruktionen tauchen in der Mathematik aber häufiger auf. Nach der axiomatischen Untersuchung von algebraischen Strukturen studiert man Abbildungen zwischen Strukturen und stellt dann oft fest, dass diese Abbildungen selbst wieder eine algebraische Struktur besitzen. Mit der Automorphismengruppe  $\operatorname{Aut}(G)$  haben wir bereits ein Beispiel kennengelernt (vgl. 4.2).

Sind  $f : V \rightarrow W$  und  $g : W \rightarrow U$  lineare Abbildungen, so ist auch die Komposition

$$g \circ f : V \rightarrow U$$

linear. Insbesondere ist für alle  $f, g \in \operatorname{End}(V)$  auch  $g \circ f \in \operatorname{End}(V)$ . Der Vektorraum  $\operatorname{End}(V)$  kann also mit einer Multiplikation  $\circ$  versehen werden. Wir können Vektoren in  $\operatorname{End}(V)$  nicht nur addieren und skalieren, sondern auch multiplizieren (im Gegensatz zu den Vek-



toren des, als Beispiel,  $\mathbb{R}^5$ ). Für alle  $f, g, h \in \text{End}(V)$  und alle Skalare  $\alpha$  gilt, wenn wir die Komposition  $\circ$  multiplikativ schreiben:

$$(a) (fg)h = f(gh),$$

$$(b) f(g + h) = fg + fh, \quad (f + g)h = fh + gh,$$

$$(c) \alpha(fg) = (\alpha f)g = f(\alpha g).$$

Wird ein  $K$ -Vektorraum  $A$  mit einer Multiplikation  $\cdot : A^2 \rightarrow A$  versehen, sodass (a) – (c) gilt, so heißt  $A$  eine (*assoziative*) *Algebra* auf  $K$ . Der  $K$ -Vektorraum  $\text{End}(V)$  ist also eine  $K$ -Algebra unter der Komposition von Abbildungen.

### Beispiel

Ist  $f \in \text{End}(V)$ , so auch  $f^2 = f \circ f$ ,  $f^3 = f^2 \circ f$  usw. Damit ist für alle  $n$  und alle Skalare  $\alpha_0, \dots, \alpha_n$  die Abbildung

$$g = \alpha_n f^n + \alpha_{n-1} f^{n-1} + \dots + \alpha_1 f + \alpha_0 f^0$$

ein Element von  $\text{End}(V)$ , wobei  $f^0 = \text{id}_V$ .

Wir bestimmen nun noch die Dimension von  $\text{Hom}(V, W)$ . Hier gilt:

$$\text{Ist } \dim(V) = n \text{ und } \dim(W) = m, \text{ so ist } \dim(\text{Hom}(V, W)) = nm.$$

Sind nämlich  $(v_1, \dots, v_n)$  und  $(w_1, \dots, w_m)$  Basen von  $V$  bzw.  $W$ , so sei  $E_{i,j} : V \rightarrow W$  für alle  $1 \leq i \leq m$  und  $1 \leq j \leq n$  die eindeutige lineare Abbildung mit

$$E_{i,j}(v_j) = w_i, \quad E_{i,j}(v_k) = 0 \text{ für alle } k \neq j.$$

Dann ist  $B = (E_{i,j})_{1 \leq i \leq m, 1 \leq j \leq n}$  eine Basis von  $\text{Hom}(V, W)$  der Länge  $nm$ . Die eine Abbildung  $E_{i,j}$  darstellenden Koeffizienten weisen genau eine Eins und ansonsten nur Nullen auf (vgl. obiges Diagramm und 4. 7). Ist  $f \in \text{Hom}(V, W)$ , so gilt

$$f = \sum_{1 \leq i \leq m, 1 \leq j \leq n} \alpha_{i,j} E_{i,j},$$

mit den darstellenden Koeffizienten  $\alpha_{i,j}$  von  $f$ . Diese Koeffizienten sind also die Koordinaten von  $f$  bzgl. der Basis  $B$  von  $\text{Hom}(V, W)$ .

### Beispiel

Für  $V = W = \mathbb{R}^3$  und die kanonischen Basen gilt  $E_{2,3}(e_3) = e_2$  und allgemein  $E_{2,3}(x, y, z) = (0, z, 0) = (\bar{x}, \bar{y}, \bar{z})$  mit

$$\bar{x} = 0x + 0y + 0z$$

$$\bar{y} = 0x + 0y + 1z$$

$$\bar{z} = 0x + 0y + 0z.$$

## 4.12 Dualräume und duale Abbildungen

**Definition** (*Dualraum, lineares Funktional, duale Basis*)

Der Dualraum  $V^*$

Sei  $V$  ein  $K$ -Vektorraum. Dann definieren wir den Dualraum  $V^*$  von  $V$  durch

$$V^* = \text{Hom}(V, K) = \{ f \mid f : V \rightarrow K \text{ ist linear} \}.$$

Die Elemente von  $V^*$  heißen auch *lineare Funktionale*.

Die Dualbasis  $v_1^*, \dots, v_n^*$

Ist  $\dim(V) < \infty$  und  $(v_1, \dots, v_n)$  eine Basis von  $V$ , so definieren wir für alle  $1 \leq j \leq n$ :

$v_j^* =$  „das eindeutige  $f \in V^*$  mit  $f(v_j) = 1$  und  $f(v_k) = 0$  für alle  $k \neq j$ “.

Das Tupel  $(v_1^*, \dots, v_n^*)$  heißt die zu  $(v_1, \dots, v_n)$  *duale Basis*.

Wir betrachten hier einen Spezialfall von  $\text{Hom}(V, W)$ : Der Zielraum  $W$  ist nun der Skalarenkörper  $K$  von  $V$ .

### Beispiel

Für  $V = \mathbb{R}^2$  besteht  $V^*$  aus allen linearen  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ . Für jedes  $f \in V^*$  gilt

$$f(x, y) = x f(1, 0) + y f(0, 1) = ax + by \quad \text{für alle } x, y \in \mathbb{R}, \text{ wobei}$$

$$a = f(e_1) = f(1, 0), \quad b = f(e_2) = f(0, 1).$$

Damit ist  $f$  die Ebene durch den Ursprung mit der Steigung  $a$  entlang der  $x$ -Achse und der Steigung  $b$  entlang der  $y$ -Achse. Der Dualraum  $V^*$  von  $\mathbb{R}^2$  besteht aus allen diesen Ebenen. Analog besteht  $V^*$  für  $V = \mathbb{R}$  aus allen Geraden  $f : \mathbb{R} \rightarrow \mathbb{R}$  mit  $f(0) = 0$ .

Ist  $n = \dim(V) < \infty$ , so ist  $\dim(V^*) = \dim(\text{Hom}(V, K)) = n$  1.11, sodass  $V \cong V^*$ .

Die  $*$ -Operation ordnet einem Basisvektor  $v_j$  ein Element  $v_j^*$  des Dualraums zu. Nützlich ist hier das *Kronecker-Symbol*  $\delta_{jk}$ , das definiert ist durch  $\delta_{jj} = 1$  und  $\delta_{jk} = 0$  falls  $j \neq k$ . Für alle  $1 \leq j, k \leq n$  und  $\alpha_1, \dots, \alpha_n \in K$  gilt also

$$v_j^*(v_k) = \delta_{jk},$$

$$v_j^*(\alpha_1 v_1 + \dots + \alpha_n v_n) = \alpha_j.$$

Die lineare Abbildung  $v_j^* : V \rightarrow \mathbb{R}$  pickt für jedes  $v \in V$  die  $j$ -te Koordinate von  $v$  bzgl.  $B = (v_1, \dots, v_n)$  heraus (vgl. 3.6). Somit ist

$$(v_1^*(v), \dots, v_n^*(v)) = \Phi_B(v) = (\alpha_1, \dots, \alpha_n)$$

für alle  $v = \alpha_1 v_1 + \dots + \alpha_n v_n \in V$ .

Für unendlich-dimensionale Vektorräume  $V$  ist die Isomorphie  $V \cong V^*$  nicht mehr gültig. Der Dualraum  $V^*$  ist dann substantiell größer als  $V$ .

$v_j^*$	$v_1$	$\dots$	$v_{j-1}$	$v_j$	$v_{j+1}$	$\dots$	$v_n$
	0	$\dots$	0	1	0	$\dots$	0

Die linearen Abbildungen  $v_j^*$  sind „Koordinatenpicker“:

$$v_1^*(\alpha_1 v_1 + \dots + \alpha_n v_n) = \alpha_1$$

$\dots$

$$v_j^*(\alpha_1 v_1 + \dots + \alpha_n v_n) = \alpha_j$$

$\dots$

$$v_n^*(\alpha_1 v_1 + \dots + \alpha_n v_n) = \alpha_n$$

**Beispiel**

Sei  $V = \mathbb{R}^{(\mathbb{N})}$ , und sei  $(e_n)_{n \in \mathbb{N}}$  die kanonische Basis von  $V$ . Die linearen Funktionale  $e_n^* \in V^*$  können genau wie oben definiert werden. Die Familie  $(e_n^*)_{n \in \mathbb{N}}$  ist linear unabhängig in  $V^*$ , aber nicht mehr erzeugend: Ist  $f: V \rightarrow \mathbb{R}$  linear mit  $f(e_n) \neq 0$  für unendlich viele  $n$ , so ist  $f \notin \text{span}((e_n^*)_{n \in \mathbb{N}})$ . Analoges gilt für  $V = K^{(\mathbb{I})}$ ,  $\mathbb{I}$  unendlich.

Mit Hilfe der Dualräume führen wir ein:

**Definition** (*duale Abbildung*)

Seien  $V, W$  beliebige  $K$ -Vektorräume, und sei  $f: V \rightarrow W$  linear. Dann ist die *duale Abbildung*  $f^*: W^* \rightarrow V^*$  von  $f$  für alle  $g \in W^*$  definiert durch

$$f^*(g) = g \circ f. \quad (\text{Pullback von } g \text{ durch } f)$$

Ein  $g \in W^*$  wird durch Vorschalten eines festen  $f \in \text{Hom}(V, W)$  zu einem linearen Funktional  $f^*(g) \in V^*$  zurückgezogen. Sind  $V$  und  $W$  endlich-dimensional mit  $\dim(V) = n$  und  $\dim(W) = m$ , so liefert die Dimensionsformel, dass

$$m - \dim(\text{Bild}(f^*)) = \dim(\text{Kern}(f^*)) =$$

$$\dim(\{ g \in W^* \mid g \circ f = 0 \}) =$$

$$\dim(\{ g \in W^* \mid g(w) = 0 \text{ für alle } w \in \text{Bild}(f) \}) = m - \dim(\text{Bild}(f)).$$

Wir erhalten:

**Dimensionen des dualen Bildes und Kernes**

$$\dim(\text{Bild}(f^*)) = \dim(\text{Bild}(f)), \quad \dim(\text{Kern}(f^*)) = m - n + \dim(\text{Kern}(f)).$$

Aus den Formeln folgt, dass sich die Eigenschaften „Epimorphismus“ und „Monomorphismus“ beim Wechsel zwischen  $f$  und  $f^*$  austauschen.

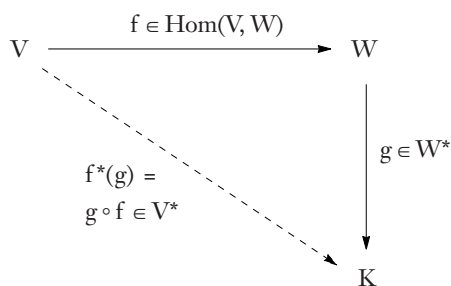
**Exkurs: Bidualräume**

Zu jedem Vektorraum  $V$  kann man den Dualraum  $V^*$  bilden, und damit lässt sich auch der Dualraum  $(V^*)^* = V^{**}$  von  $V^*$  bilden, der sog. *Bidualraum* von  $V$ . Er besteht aus allen linearen  $F: V^* \rightarrow K$ . Ein  $F \in V^{**}$  weist jedem linearen  $f: V \rightarrow K$  einen Skalar  $F(f) \in K$  zu. Das ist gar nicht so wild, wie es zunächst aussieht: Ist  $v \in V$  beliebig, so definieren wir das Element  $F_v: V^* \rightarrow K$  des Bidualraums  $V^{**}$  durch

$$F_v(f) = f(v) \quad \text{für alle } f \in V^*.$$

Die Funktion  $F_v$  pickt aus jeder linearen Abbildung  $f: V \rightarrow K$  den Wert  $f(v)$  heraus. Ist  $\dim(V) < \infty$  (und also  $V \cong V^* \cong V^{**}$ ), so ist jedes Element von  $V^{**}$  von der Form  $F_v$ . Genauer ist dann die Abbildung  $\Psi: V \rightarrow V^{**}$  ein Isomorphismus, wobei

$$\Psi(v) = F_v \quad \text{für alle } v \in V.$$





# 5. Kapitel

---

## Matrizen

---

## 5.1 Matrizen

**Definition** (*Matrix, Einträge, Spalten, Zeilen,  $K^{m \times n}$* )

Seien  $K$  ein Körper und  $m, n \geq 1$ . Eine Familie

$$A = (a_{i,j})_{1 \leq i \leq m, 1 \leq j \leq n}$$

in  $K$  nennen wir die  $m \times n$ -Matrix über  $K$  mit den *Einträgen*  $A(i, j) = a_{i,j} \in K$  an den *Stellen*  $(i, j)$ . Wir notieren  $A$  in Form einer Tabelle mit  $m$  Zeilen und  $n$  Spalten:

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \dots & \dots & \dots & \dots \\ a_{m,1} & a_{m,2} & \dots & a_{m,n} \end{pmatrix}.$$

Die Vektoren  $(a_{1,j}, \dots, a_{m,j}) \in K^m$  und  $(a_{i,1}, \dots, a_{i,n}) \in K^n$  heißen die *Spalten* bzw. *Zeilen* von  $A$ . Gilt  $n = m$ , so heißt  $A$  *quadratisch*. Wir schreiben kurz

$$K^{m \times n} \text{ anstelle von } K^{\{1, \dots, m\} \times \{1, \dots, n\}} = \{A \mid A \text{ ist eine } m \times n\text{-Matrix über } K\}$$

für den  $K$ -Vektorraum aller  $m \times n$ -Matrizen mit Einträgen in  $K$ .

Eine Matrix ist formal eine Tabelle von Körper-Elementen. So wie man einen Vektor  $x \in \mathbb{R}^{12}$  als Liste von reellen Zahlen mit zwölf Einträgen auffassen kann, so kann man eine Matrix  $A \in \mathbb{R}^{3 \times 4}$  als Tabelle mit drei Zeilen und vier Spalten auffassen, deren Einträge aus reellen Zahlen bestehen. Die fundamentale Bedeutung dieser Tabellen für die Lineare Algebra ergibt sich durch ihren engen Zusammenhang mit linearen Abbildungen. Bei der Untersuchung linearer Abbildungen sind uns Matrizen schon mehrfach begegnet (vgl. 4.7, 4.11, 4.12). In diesem Kapitel werden wir die Darstellung einer linearen Abbildung durch eine Matrix genauer untersuchen. Weit über die Lineare Algebra hinaus haben Matrizen vielfältige Anwendungen, insbesondere spielen sie in der Analysis, der Graphentheorie und der Wahrscheinlichkeitstheorie eine wichtige Rolle. Immer dann, wenn doppelt indizierte Objekte auftreten, kommen Matrizen ins Spiel. Matrizen gehören zu den Grundbegriffen der Mathematik.

### Notationen und Konventionen

- (1) Sind  $m, n$  aus dem Kontext heraus klar, so schreiben wir kurz  $A = (a_{i,j}) = (a_{ij})$ . Statt  $a_{ij}$  schreiben wir alternativ auch  $A(i, j)$ . Als Familie ist eine Matrix  $A$  eine Funktion von  $\{1, \dots, m\} \times \{1, \dots, n\}$  nach  $K$ , sodass  $A(i, j)$  wohldefiniert ist.
- (2) Matrizen werden oft mit großen Buchstaben  $A, B, C, \dots$  bezeichnet und ihre Einträge automatisch entsprechend mit  $a_{ij}, b_{ij}, c_{ij}, \dots$ . Im Folgenden läuft der Zeilenindex  $i$  von 1 bis  $m$  und der Spaltenindex  $j$  von 1 bis  $n$ . Die Entsprechungen sind wie im Alphabet:  $m$  kommt vor  $n$  und  $i$  vor  $j$ .
- (3) Matrizen werden oft auch mit eckigen statt runden Klammern notiert.

Der Vektorraum  $K^{m \times n}$  ist der Produktraum  $K^I$  für  $I = \{1, \dots, m\} \times \{1, \dots, n\}$  (vgl. 3.3). Für alle  $A = (a_{ij})$ ,  $B = (b_{ij}) \in K^{m \times n}$  und  $\lambda \in K$  gilt

$$A + B = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} + \begin{pmatrix} b_{11} & \dots & b_{1n} \\ \dots & \dots & \dots \\ b_{m1} & \dots & b_{mn} \end{pmatrix} = \begin{pmatrix} a_{11} + b_{11} & \dots & a_{1n} + b_{1n} \\ \dots & \dots & \dots \\ a_{m1} + b_{m1} & \dots & a_{mn} + b_{mn} \end{pmatrix},$$

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

### Beispiele

(1) Die *Nullmatrix*  $0 \in K^{m \times n}$  ist definiert durch  $0(i, j) = 0$  für alle  $i, j$ .

(2) Die *Einheitsmatrix*  $E_n \in K^{n \times n}$  ist definiert durch

$$E_n(i, i) = 1 \text{ für alle } i, \quad E_n(i, j) = 0 \text{ für alle } i \neq j.$$

Mit Hilfe des Kronecker-Deltas  $\delta_{ij}$  gilt  $E_n(i, j) = \delta_{ij}$  für alle  $i, j$ . Die Spalten und Zeilen von  $E_n$  sind die Standardvektoren  $e_1, \dots, e_n$ .

$$\begin{pmatrix} a_1 & & & \\ & a_2 & & \\ & & \dots & \\ & & & a_n \end{pmatrix}$$

Die Diagonalmatrix  $\text{diag}(a_1, \dots, a_n) \in K^{n \times n}$ . Nichtspezifizierte Einträge sind gleich null.

(3) Ein  $A \in K^{n \times n}$  heißt *Diagonalmatrix*, falls  $A(i, j) = 0$  für alle  $i \neq j$ . Wir schreiben  $\text{diag}(a_1, \dots, a_n)$  für die Diagonalmatrix  $A$  mit  $A(i, i) = a_i$  für alle  $i$ . Speziell gilt  $E_n = \text{diag}(1, \dots, 1)$ .

(4) Ein  $A \in K^{n \times n}$  heißt eine *obere Dreiecksmatrix*, falls  $A(i, j) = 0$  für alle  $i > j$ . Analog ist eine *untere Dreiecksmatrix* durch „ $A(i, j) = 0$  für alle  $i < j$ “ definiert.

(5) Wir definieren  $E_{ij} \in K^{m \times n}$  für alle  $i = 1, \dots, m$  und  $j = 1, \dots, n$  als die Matrix, die genau an der Stelle  $(i, j)$  den Eintrag 1 besitzt und sonst nur Nulleinträge aufweist. Es gilt also  $E_{ij}(i', j') = \delta_{(i, j), (i', j')}$  für alle  $i', j'$ . Die Matrizen  $E_{ij}$  bilden die Standardbasis des  $mn$ -dimensionalen Vektorraums  $K^{m \times n}$ . Für alle  $A \in K^{m \times n}$  gilt  $A = \sum_{1 \leq i \leq m, 1 \leq j \leq n} a_{ij} E_{ij}$ .

Wichtig für das Folgende ist:

### Einbettung des $K^m$ in den $K^{m \times 1}$

Sei  $m \geq 1$ . Wir vereinbaren:

$$(x_1, \dots, x_m) \in K^m \text{ wird identifiziert mit } \begin{pmatrix} x_1 \\ \dots \\ x_m \end{pmatrix} \in K^{m \times 1}.$$

Damit gilt  $K^m = K^{m \times 1}$ . In den folgenden Abschnitten wird klar werden, warum wir den Vektorraum  $K^{m \times 1}$  (einspaltige Matrizen) gegenüber dem auf den ersten Blick vielleicht naheliegenderen Vektorraum  $K^{1 \times m}$  (einzeilige Matrizen) bevorzugen.

## 5.2 Matrizen und lineare Abbildungen

**Definition** (*Matrix-Vektor-Produkt, zugeordnete Abbildung, darstellende Matrix*)

Seien  $K$  ein Körper und  $m, n \geq 1$ .

*Matrix-Vektor-Produkt*

Für  $A \in K^{m \times n}$  und  $x = (x_1, \dots, x_n) \in K^n$  definieren wir das *Matrix-Vektor-Produkt*  $Ax \in K^m$  von  $A$  mit  $x$  durch

$$(+)\quad Ax = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \dots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + \dots + a_{1n}x_n \\ \dots \\ a_{m1}x_1 + \dots + a_{mn}x_n \end{pmatrix}.$$

*Zugeordnete lineare Abbildung*

Ist  $A \in K^{m \times n}$ , so heißt  $f_A : K^n \rightarrow K^m$ ,

$f_A(x) = Ax$  für alle  $x \in K^n$ ,

die  $A$  *zugeordnete lineare Abbildung*.

Weiter setzen wir  $\text{Kern}(A) = \text{Kern}(f_A)$ ,

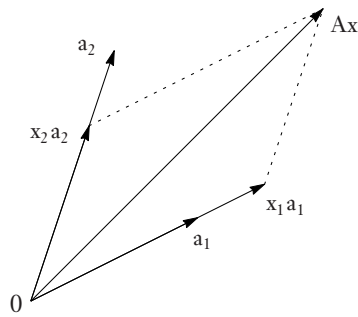
$\text{Bild}(A) = \text{Bild}(f_A)$ .

*Darstellende Matrix*

Ist  $f : K^n \rightarrow K^m$  linear, so heißt

$A_f = (f(e_1) \ \dots \ f(e_n)) \in K^{m \times n}$

die  $f$  *darstellende Matrix*.



Das Produkt  $Ax$  für  $A = (a_1 \ a_2) \in \mathbb{R}^{2 \times 2}$  mit den Spalten  $a_1 = (a_{11}, a_{21})$  und  $a_2 = (a_{21}, a_{22})$

In (+) fassen wir wie vereinbart  $x \in K^n$  als  $n \times 1$ -Matrix und die  $m \times 1$ -Matrix rechts als Element des  $K^m$  auf. Es entsteht eine Abbildung  $f_A$  von  $K^n$  nach  $K^m$ . Die Berechnung von  $f_A(x) = Ax$  lässt sich durch „Zeile mal Spalte“ ( $m$ -mal durchgeführt) beschreiben. Die wichtige andere Lesart

$$(++)\quad Ax = x_1 \begin{pmatrix} a_{11} \\ \dots \\ a_{m1} \end{pmatrix} + x_2 \begin{pmatrix} a_{12} \\ \dots \\ a_{m2} \end{pmatrix} + \dots + x_n \begin{pmatrix} a_{1n} \\ \dots \\ a_{mn} \end{pmatrix} = x_1 a_1 + \dots + x_n a_n.$$

zeigt, dass  $Ax$  eine Linearkombination der Spalten  $a_1, \dots, a_n$  von  $A$  ist. Aus beiden Darstellungen lässt sich ablesen, dass

$$A(x + y) = Ax + Ay, \quad A(\lambda x) = \lambda Ax \quad \text{für alle } x, y \in K^n \text{ und } \lambda \in K.$$

Damit ist  $f_A : K^n \rightarrow K^m$  eine lineare Abbildung. Dass wir umgekehrt einer linearen Abbildung  $f : K^n \rightarrow K^m$  eine Matrix  $A_f \in K^{m \times n}$  zuordnen können, haben wir im Darstellungssatz in 4.7 schon gesehen:

*Die Spalten von  $A_f$  sind die Bilder der kanonischen Basisvektoren  $e_1, \dots, e_n$  unter  $f$ .*



Die  $f_A$  darstellende Matrix ist  $A$  selbst. Denn nach Definition des Matrix-Vektor-Produkts  $Ax$  sind  $Ae_1, \dots, Ae_n$  die Spalten von  $A$ , sodass

$$A_{f_A} = \begin{pmatrix} f_A(e_1) & \dots & f_A(e_n) \end{pmatrix} = \begin{pmatrix} Ae_1 & \dots & Ae_n \end{pmatrix} = A.$$

Genauer gilt:

### Isomorphie von Matrizen und linearen Abbildungen

Die Abbildung  $\Psi: \text{Hom}(K^n, K^m) \rightarrow K^{m \times n}$  mit

$$\Psi(f) = A_f \quad \text{für alle } A \in K^{m \times n}$$

ist ein Isomorphismus mit  $\Psi^{-1}(A) = f_A$  für alle  $A \in K^{m \times n}$ .

Damit haben wir unser Tabellen-Verständnis von Matrizen substantiell erweitert:

*Matrizen sind (im Sinne eines Isomorphismus) lineare Abbildungen.*

### Beispiele

- (1) Sei  $n \geq 1$ . Dann gilt  $E_n x = x$  für alle  $x \in K^n$ , sodass  $f_{E_n} = \text{id}_{K^n}$ . Für eine Diagonalmatrix  $A = \text{diag}(a_1, \dots, a_n)$  gilt

$$Ax = (a_1 x_1, \dots, a_n x_n) \quad \text{für alle } x = (x_1, \dots, x_n) \in K^n.$$

- (2) Die Matrix-Vektor-Produkte mit den Matrizen

$$A = \begin{pmatrix} 1 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \quad C = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$$

über  $\mathbb{R}$  beschreiben: die Projektion  $f_A: \mathbb{R}^2 \rightarrow \mathbb{R}$ ,  $f_A(x, y) = x$ , auf die erste Komponente; die Einbettung  $f_B: \mathbb{R} \rightarrow \mathbb{R}^2$ ,  $f_B(x) = (x, x)$ ; die Vertauschung  $f_C: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ ,  $f_C(x, y) = (y, x)$ , der Komponenten.

- (3) Das Matrix-Vektor-Produkt mit der reellen Matrix

$$A = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}$$

beschreibt die Drehung im  $\mathbb{R}^2$  um den Winkel  $\varphi$  gegen den Uhrzeigersinn.

- (4) Für die Matrizen  $E_{ij}$  der Standardbasis des  $K^{m \times n}$  gilt

$$E_{ij} x = (0, \dots, 0, x_j, 0, \dots, 0) = x_j e_i \in K^m \quad \text{für alle } x \in K^n,$$

wobei  $x_j$  an der  $i$ -ten Stelle steht. Das Matrix-Vektor-Produkt mit  $E_{ij}$  pickt also die Komponente  $x_j$  aus  $x \in K^n$  heraus und platziert sie an der  $i$ -ten Stelle.

- (5) Mit den Bezeichnungen aus 4.11 gilt

$$\Psi(E_{ij}) = E_{i,j} \in \text{Hom}(K^n, K^m) \quad \text{für alle } i, j,$$

wobei  $E_{i,j}$  bezüglich der Standardbasen des  $K^n$  und  $K^m$  definiert ist.

### 5.3 Die Matrizenmultiplikation

#### Definition (Matrizenprodukt)

Seien  $K$  ein Körper und  $k, m, n \geq 1$ . Wir definieren für alle  $A = (a_{ir}) \in K^{m \times k}$  und  $B = (b_{rj}) \in K^{k \times n}$  das *Matrizenprodukt*  $AB = A \cdot B = (c_{ij}) \in K^{m \times n}$  durch

$$c_{ij} = \sum_{1 \leq r \leq k} a_{ir} b_{rj} = a_{i1} b_{1j} + \dots + a_{ik} b_{kj} \quad \text{für alle } 1 \leq i \leq m, 1 \leq j \leq n.$$

$$\begin{aligned} AB &= \begin{pmatrix} a_{11} & \dots & a_{1k} \\ \dots & \dots & \dots \\ a_{m1} & \dots & a_{mk} \end{pmatrix} \cdot \begin{pmatrix} b_{11} & \dots & b_{1n} \\ \dots & \dots & \dots \\ b_{k1} & \dots & b_{kn} \end{pmatrix} = \\ &= \begin{pmatrix} a_{11}b_{11} + \dots + a_{1k}b_{k1} & \dots & a_{11}b_{1n} + \dots + a_{1k}b_{kn} \\ \dots & \dots & \dots \\ a_{m1}b_{11} + \dots + a_{mk}b_{k1} & \dots & a_{m1}b_{1n} + \dots + a_{mk}b_{kn} \end{pmatrix} = \\ &= \begin{pmatrix} Ab_1 & \dots & Ab_n \end{pmatrix}, \quad \text{mit den Spalten } b_1, \dots, b_n \text{ der Matrix } B. \end{aligned}$$

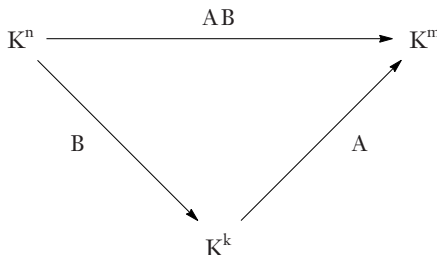
Die Produktbildung erfolgt gemäß „Zeile mal Spalte“,  $mn$ -mal durchgeführt. In den Spalten von  $AB$  stehen die Matrix-Vektor-Produkte von  $A$  mit den Spalten von  $B$ . Insbesondere ist  $Ax$  für  $A \in K^{m \times n}$  und  $x \in K^n = K^{n \times 1}$  der Spezialfall der Matrizenmultiplikation mit einem einspaltigen zweiten Faktor. Das Produkt  $AB$  ist nur erklärt, wenn die Zeilenzahl von  $A$  mit der Spaltenzahl von  $B$  übereinstimmt. Unentbehrlich ist:

#### Motivation der Matrizenmultiplikation

Sind  $f: K^n \rightarrow K^k$  und  $g: K^k \rightarrow K^m$  lineare Abbildungen, so gilt

$$A_{g \circ f} = A_g \cdot A_f. \quad (\text{Kompositionssatz für darstellende Matrizen})$$

Die darstellende Matrix der Komposition  $g \circ f$  ist also das Produkt der darstellenden Matrizen von  $g$  und  $f$ . Sind umgekehrt  $A \in K^{m \times k}$ ,  $B \in K^{k \times n}$ , so gilt  $f_{AB} = f_A \circ f_B$ .



Die Multiplikation ist so gemacht, dass

$$f_{AB} = f_A \circ f_B.$$

Konvention: In Diagrammen schreiben wir oft einfach  $C$  statt  $f_C$ . Dies ist suggestiv und besser lesbar. Manche Autoren identifizieren generell  $C$  und  $f_C$ .

Ohne explizites Nachrechnen ergibt sich aus der Assoziativität der Komposition von Funktionen, dass die Matrizenmultiplikation assoziativ ist.

**Beispiele**

- (1) Für alle  $A \in K^{m \times n}$  gilt  $AE_n = (Ae_1 \dots Ae_n) = A$  und analog  $E_m A = A$ . Speziell ist  $AE_n = E_n A = A$  für alle  $A \in K^{n \times n}$ .

- (2) Sei  $K$  ein Körper. Dann gilt:

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \neq 0,$$

$$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = 0.$$

- (3) Für  $A = \text{diag}(a_1, \dots, a_n)$ ,  $B = \text{diag}(b_1, \dots, b_n) \in K^{n \times n}$  gilt

$$AB = \text{diag}(a_1 b_1, \dots, a_n b_n) = BA.$$

Die Diagonalmatrizen des  $K^{n \times n}$  sind also abgeschlossen unter der Matrizenmultiplikation. Ebenso ist das Produkt zweier unterer (oberer) Dreiecksmatrizen des  $K^{n \times n}$  wieder eine untere (obere) Dreiecksmatrix des  $K^{n \times n}$ .

- (4) Beschreiben  $A_\varphi, A_\psi \in \mathbb{R}^{2 \times 2}$  die Drehungen um  $\varphi$  bzw.  $\psi$ , so beschreibt  $A_\varphi A_\psi$  die Drehung um  $\varphi + \psi$  (vgl. 5.2). Es gilt  $A_\varphi A_\psi = A_{\varphi + \psi} = A_{\psi + \varphi} = A_\psi A_\varphi$ . Aus

$$A_\varphi A_\psi = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix} \begin{pmatrix} \cos \psi & -\sin \psi \\ \sin \psi & \cos \psi \end{pmatrix} = \begin{pmatrix} \cos(\varphi + \psi) & -\sin(\varphi + \psi) \\ \sin(\varphi + \psi) & \cos(\varphi + \psi) \end{pmatrix} = A_{\varphi + \psi}$$

erhalten wir die Additionstheoreme für den Sinus und Kosinus:

$$\cos(\varphi + \psi) = A_{\varphi + \psi}(1, 1) = (A_\varphi A_\psi)(1, 1) = \cos \varphi \cos \psi - \sin \varphi \sin \psi,$$

$$\sin(\varphi + \psi) = A_{\varphi + \psi}(2, 1) = (A_\varphi A_\psi)(2, 1) = \sin \varphi \cos \psi + \cos \varphi \sin \psi.$$

Gilt  $m = n = k$ , so ist die Matrizenmultiplikation eine Operation auf der Menge  $K^{n \times n}$  aller quadratischen Matrizen mit je  $n$  Zeilen und Spalten. Algebraische Eigenschaften dieser Operation sind:

- (1) Die Menge  $K^{n \times n}$  bildet mit der Addition und Multiplikation von Matrizen einen Ring. Die Nullmatrix  $0$  ist additiv neutral und die Einheitsmatrix  $E_n = \text{diag}(1, \dots, 1)$  multiplikativ neutral.
- (2) Der  $K$ -Vektorraum  $K^{n \times n}$  bildet mit der Multiplikation von Matrizen eine  $K$ -Algebra.

Beispiel (2) zeigt, dass der Matrizenring im Allgemeinen weder kommutativ noch nullteilerfrei ist. In  $K^{n \times n}$  sind wie in jedem Ring die Potenzen  $A^k$  definiert:

$$A^0 = E_n, \quad A^{k+1} = A^k A \quad \text{für alle } k \in \mathbb{N}.$$

## 5.4 Darstellende Matrizen für beliebige Basen

### Definition (darstellende Matrix bzgl. zweier Basen)

Seien  $V, W$  endlich-dimensionale  $K$ -Vektorräume,  $\mathcal{A} = (v_1, \dots, v_n)$ ,  $\mathcal{B} = (w_1, \dots, w_m)$  Basen von  $V$  bzw.  $W$  und  $f : V \rightarrow W$  linear. Dann ist die  $f$  bzgl. der Basen  $\mathcal{A}$  und  $\mathcal{B}$  darstellende Matrix

$$A = A_f^{\mathcal{A}, \mathcal{B}} = \text{„}A_f \text{ bzgl. } \mathcal{A}, \mathcal{B}\text{“}$$

definiert als

$$(\Phi_{\mathcal{B}}(f(v_1)) \dots \Phi_{\mathcal{B}}(f(v_n))) \in K^{m \times n},$$

mit der Koordinatenabbildung

$$\Phi_{\mathcal{B}} : W \rightarrow K^m.$$

Die Matrix  $A$  berechnet, gegeben die  $\mathcal{A}$ -Koordinaten  $x \in K^n$  von  $v \in V$ , die  $\mathcal{B}$ -Koordinaten  $Ax \in K^m$  von  $f(v) \in W$ . Ihre Definition lautet in Kurzform:

*Die Spalten von  $A$  sind die  $\mathcal{B}$ -Koordinaten der Bilder der Basisvektoren in  $\mathcal{A}$ .*

Die Matrix  $A$  lässt sich aufstellen, wenn wir die Vektoren  $f(v_i)$  als Linearkombinationen bzgl.  $\mathcal{B}$  schreiben:

$$f(v_1) = a_{11} w_1 + \dots + a_{m1} w_m,$$

...

$$f(v_n) = a_{1n} w_1 + \dots + a_{mn} w_m.$$

Die Darstellung von  $f(v_1)$  liefert die erste Spalte von  $A$ , die Darstellung  $f(v_2)$  die zweite Spalte von  $A$  usw. Der Leser vergleiche den allgemeinen Darstellungssatz in 4.7.

Die Definition von  $A_f^{\mathcal{A}, \mathcal{B}}$  verallgemeinert die Definition von  $A_f$  aus 5.2. Dort hatten wir  $V = K^n$ ,  $W = K^m$  und die Standardbasen betrachtet. Die Koordinatenabbildungen sind in diesem Fall die Identitäten.

Wir erhalten:

### Isomorphie von Matrizen und linearen Abbildungen, allgemeine Form

Für  $V, W, \mathcal{A}, \mathcal{B}$  wie oben ist die Abbildung  $\Psi : \text{Hom}(V, W) \rightarrow K^{m \times n}$  mit

$$\Psi(f) = \text{„}A_f \text{ bzgl. } \mathcal{A}, \mathcal{B}\text{“} \quad \text{für alle linearen } f : V \rightarrow W$$

ein Isomorphismus mit  $\Psi^{-1}(A) = \Phi_{\mathcal{B}}^{-1} \circ f_A \circ \Phi_{\mathcal{A}}$  für alle  $A \in K^{m \times n}$ .

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ \Phi_{\mathcal{A}} \downarrow & & \downarrow \Phi_{\mathcal{B}} \\ K^n & \xrightarrow{A} & K^m \end{array}$$

Für  $A = A_f$  bzgl.  $\mathcal{A}, \mathcal{B}$  und die Koordinatenabbildungen  $\Phi_{\mathcal{A}}$  und  $\Phi_{\mathcal{B}}$  (vgl. 3.6, 4.9) gilt

$$f = \Phi_{\mathcal{B}}^{-1} \circ f_A \circ \Phi_{\mathcal{A}}.$$

Die Matrix  $A$  rechnet die Koordinaten um.

$$\begin{array}{ccccc} V & \xrightarrow{f} & U & \xrightarrow{g} & W \\ \Phi_{\mathcal{A}} \downarrow & & \downarrow \Phi_{\mathcal{C}} & & \downarrow \Phi_{\mathcal{B}} \\ K^n & \xrightarrow{A} & K^k & \xrightarrow{B} & K^m \\ & \searrow & & \nearrow & \\ & & BA & & \end{array}$$

Die Multiplikation entspricht der Komposition: Für  $A = A_f$  bzgl.  $\mathcal{A}, \mathcal{C}$  und  $B = A_g$  bzgl.  $\mathcal{C}, \mathcal{B}$  gilt  $BA = A_{g \circ f}$  bzgl.  $\mathcal{A}, \mathcal{B}$ .

**Beispiel**

Im  $\mathbb{R}^2$  seien  $v_1 = (1, 1)$  und  $v_2 = (1, 2)$ . Wir betrachten die durch  $f(v_1) = e_1$ ,  $f(v_2) = e_2$  eindeutig definierte lineare Abbildung  $f: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ . Es gilt

$$A_f = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \quad \text{bzgl. } (v_1, v_2), (e_1, e_2), \quad \text{da } f(v_1) = e_1 + 0e_2, f(v_2) = 0e_1 + e_2,$$

$$A_f = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix} \quad \text{bzgl. } (v_1, v_2), (v_1, v_2), \quad \text{da } f(v_1) = 2v_1 - v_2, f(v_2) = -v_1 + v_2.$$

Die Definition „ $A_f$  bzgl.  $\mathcal{A}, \mathcal{B}$ “ trägt der Gleichberechtigung aller Basen Rechnung. Folgende Überlegung zeigt jedoch, dass wir eine beliebig vorgegebene Abbildung  $f$  sehr einfach darstellen können, wenn wir die Basen  $\mathcal{A}$  und  $\mathcal{B}$  geschickt wählen:

**Die Normalformdarstellung**

Sei  $f: V \rightarrow W$  linear, und seien  $v_1, \dots, v_r \in V$  derart, dass  $w_1 = f(v_1), \dots, w_r = f(v_r)$  eine Basis des Unterraums  $\text{Bild}(f)$  von  $W$  bilden. Wir ergänzen nun die  $v_j$  zu einer Basis  $\mathcal{A} = (v_1, \dots, v_n)$  von  $V$ , indem wir eine Basis  $(v_{r+1}, \dots, v_n)$  von  $\text{Kern}(f)$  anfügen (ist  $r = n$ , so entfällt dieser Schritt). Weiter ergänzen wir die  $w_i$  beliebig zu einer Basis  $\mathcal{B} = (w_1, \dots, w_m)$  von  $W$ . Dann gilt nach Konstruktion

$$\Phi_{\mathcal{B}} f(v_1) = e_1, \dots, \Phi_{\mathcal{B}} f(v_r) = e_r, \quad \Phi_{\mathcal{B}} f(v_{r+1}) = \dots = \Phi_{\mathcal{B}} f(v_n) = 0.$$

Damit gilt bzgl.  $\mathcal{A}, \mathcal{B}$

$$A_f = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} \in K^{m \times n}, \text{ wobei } r = \dim(\text{Bild}(f)). \quad (\text{Normalformdarstellung})$$

Ist  $f: V \rightarrow W$  ein Isomorphismus, so ist die darstellende Matrix gleich  $E_n$ .

Dies motiviert:

**Definition** (*äquivalente Matrizen*)

Zwei Matrizen  $A, A' \in K^{m \times n}$  heißen *äquivalent*, falls sie bzgl. geeigneter Basen dieselbe Abbildung darstellen, d. h., falls es  $K$ -Vektorräume  $V, W$  mit  $n = \dim(V)$ ,  $m = \dim(W)$ , ein lineares  $f: V \rightarrow W$  und Basen  $\mathcal{A}, \mathcal{A}'$  von  $V$  und  $\mathcal{B}, \mathcal{B}'$  von  $W$  gibt mit

$$A = A_f \text{ bzgl. } \mathcal{A}, \mathcal{B}, \quad A' = A_f \text{ bzgl. } \mathcal{A}', \mathcal{B}'.$$

Für alle  $m, n$  liegt (wie der Name suggeriert) eine Äquivalenzrelation auf  $K^{m \times n}$  vor. Ein vollständiges Repräsentantensystem wird gegeben durch

$$0 = \begin{pmatrix} E_0 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} E_1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} E_2 & 0 \\ 0 & 0 \end{pmatrix}, \dots, \begin{pmatrix} E_k & 0 \\ 0 & 0 \end{pmatrix} \in K^{m \times n}, \quad \text{mit } k = \min(m, n).$$

## 5.5 Invertierbare Matrizen

### Definition (Invertierbarkeit, Inverse, allgemeine lineare Gruppe)

Seien  $K$  ein Körper und  $n \geq 1$ . Ein  $A \in K^{n \times n}$  heißt *invertierbar*, falls es ein  $B \in K^{n \times n}$  gibt mit  $AB = BA = E_n$ . Die Matrix  $B$  heißt dann die zu  $A$  *inverse Matrix* und wird mit  $A^{-1}$  bezeichnet. Eine nicht invertierbare Matrix nennt man *singulär*. Weiter heißt

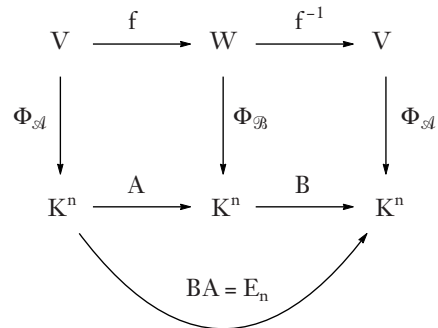
$$GL(n, K) = \{A \in K^{n \times n} \mid A \text{ ist invertierbar}\}$$

die *allgemeine lineare Gruppe* vom Grad  $n$  über  $K$ .

Die Gruppe  $GL(n, K)$  besteht aus den Einheiten des Matrizenrings  $K^{n \times n}$  („GL“ steht für „general linear“). Nach den Rechenregeln in Gruppen gilt für alle  $A, B \in GL(n, K)$ :

$$(A^{-1})^{-1} = A, (AB)^{-1} = B^{-1}A^{-1}.$$

Jeweils äquivalent zur Invertierbarkeit von  $A \in K^{n \times n}$  sind die Bedingungen:



Die Invertierung entspricht der Umkehrabbildung:  $(A_f)^{-1} = A_{f^{-1}}$ .

$f_A : K^n \rightarrow K^n$ ist bijektiv (gleichwertig: injektiv, surjektiv).
Die Spalten von $A$ bilden eine Basis des $K^n$ .
Es gibt ein $B \in K^{n \times n}$ mit $AB = E_n$ oder $BA = E_n$ .

Die beiden ersten Kriterien folgen aus der Definition. Das nicht selbstverständliche dritte Kriterium ergibt sich daraus, dass  $f_A \circ f_B = \text{id}$  impliziert, dass  $f_A$  surjektiv und  $f_B$  injektiv ist (vgl. hierzu die Diskussion von  $M^\times$  in 2.3).

### Beispiele

- (1) Eine Diagonalmatrix  $A = \text{diag}(a_1, \dots, a_n)$  ist genau dann invertierbar, wenn alle  $a_i$  von null verschieden sind. In diesem Fall gilt  $A^{-1} = \text{diag}(a_1^{-1}, \dots, a_n^{-1})$ .
- (2) Die Drehmatrizen  $A_\varphi \in \mathbb{R}^{2 \times 2}$  sind invertierbar mit  $A_\varphi^{-1} = A_{-\varphi}$ . Sie bilden eine Untergruppe von  $GL(2, \mathbb{R})$ .

$$(3) \text{ Für } A \in GL(2, \mathbb{R}) \text{ gilt } A^{-1} = \frac{1}{a_{11}a_{22} - a_{21}a_{12}} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix},$$

vorausgesetzt, der Nenner ist ungleich 0. Diese Formel wird durch die Einführung von Determinanten verständlich (vgl. 7.1).

- (4) Die Summe  $A + B$  von  $A, B \in GL(n, K)$  ist im Allgemeinen nicht invertierbar, wie  $A = E_n$  und  $B = -E_n$  zeigen. Mit  $A$  ist aber stets auch  $-A$  invertierbar.

Wir betrachten zwei Anwendungen invertierbarer Matrizen.

### Eindeutig lösbar lineare Gleichungssysteme

Eine Matrix  $A \in K^{n \times n}$  ist genau dann invertierbar, wenn das Gleichungssystem

$$Ax = b$$

für alle  $b \in K^n$  eine eindeutige Lösung besitzt. Denn genau dann ist  $f_A : K^n \rightarrow K^n$  bijektiv. Ist umgekehrt  $A^{-1}$  bekannt, so gilt

$$Ax = b \text{ genau dann, wenn } x = A^{-1}b, \quad (\text{Lösen durch Invertierung})$$

wie die Multiplikation von links mit  $A^{-1}$  zeigt. Kennt man  $A^{-1}$ , so kann man  $Ax = b$  für jede rechte Seite  $b$  durch Berechnung von  $A^{-1}b$  lösen.

### Berechnung von Koordinatenvektoren

Seien  $\mathcal{A} = (a_1, \dots, a_n)$  eine Basis des  $K^n$  und  $\Phi_{\mathcal{A}} : K^n \rightarrow K^n$  die zugehörige Koordinatenabbildung. Wir bilden die  $n \times n$ -Matrix  $A$  mit den Basisvektoren als Spalten:

$$A = \begin{pmatrix} a_1 & \dots & a_n \end{pmatrix}.$$

Für alle  $x, y \in K^n$  gilt  $\Phi_{\mathcal{A}}(x) = y$  genau dann, wenn

$$x = y_1 a_1 + \dots + y_n a_n = Ay.$$

Die Matrix  $A$  ist also die darstellende Matrix von  $\Phi_{\mathcal{A}}^{-1}$  (bzgl. der Standardbasen), da  $\Phi_{\mathcal{A}}^{-1}(y) = Ay$  für alle  $y$ . Damit ist  $A^{-1}$  die darstellende Matrix von  $\Phi_{\mathcal{A}}$ , sodass

$$(+)\ \Phi_{\mathcal{A}}(x) = A^{-1}x \text{ für alle } x \in K^n. \quad (\text{Koordinatenberechnung durch Invertierung})$$

Alternativ können wir so argumentieren:  $A$  ist die darstellende Matrix von  $\text{id} : K^n \rightarrow K^n$  bzgl. der Basen  $\mathcal{A}$  und  $\mathcal{B} = (e_1, \dots, e_n)$ , denn die Spalten von  $A$  sind die Koordinatenvektoren von  $\text{id}(a_i)$  bzgl.  $\mathcal{B}$ . Das kommutative Diagramm rechts liefert (+).

$$\begin{array}{ccc} V = K^n & \xrightarrow{\text{id}} & W = K^n \\ \Phi_{\mathcal{A}} \downarrow & & \downarrow \Phi_{\mathcal{B}} = \text{id} \\ K^n & \xrightarrow{A} & K^n \end{array}$$

Es bleiben die Fragen:

*Wie berechnet man  $A^{-1}$  für  $A \in GL(n, K)$ ?*

*Wie überprüft man, ob  $A \in K^{n \times n}$  invertierbar ist?*

Der Ansatz „ $AB = E_n$ “ mit einer unbekannten Matrix  $B \in K^{n \times n}$  führt zu  $n$  linearen Gleichungssystemen

$$Ax = e_1, \dots, Ax = e_n,$$

deren Lösungen die Spalten von  $B = A^{-1}$  bilden. Eine effektive Möglichkeit zur Bestimmung von  $A^{-1}$  werden wir im folgenden Abschnitt kennenlernen.

5.6
Die Elementarmatrizen

Definition
(Elementarmatrizen und ihre Typen)

Seien  $K$  ein Körper und  $n \geq 1$ . Für alle  $1 \leq i, j \leq n$  und  $\lambda \in K$  sei  $W_{ij}(\lambda) \in K^{n \times n}$  die Matrix, die aus  $E_n$  durch Überschreiben des  $(i, j)$ -Eintrags mit  $\lambda$  hervorgeht. Wir nennen ein  $W \in K^{n \times n}$  eine *Elementarmatrix*, falls  $W$  von einem der folgenden Typen ist:

- Additionstyp

$W = W_{ij}(\lambda)$

mit  $\lambda \in K, i \neq j$ ,
- Multiplikationstyp

$W = W_{ii}(\lambda)$

mit  $\lambda \in K^*$ .

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & \lambda & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$W_{24}(\lambda) \in K^{5 \times 5}$$

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & \lambda & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$W_{44}(\lambda) \in K^{5 \times 5}$$

Die Bezeichnung  $W_{ij}(\lambda)$  steht für „write  $\lambda$  at  $(i, j)$  in  $E_n$ “. Der Additionstyp enthält einen Eintrag  $\lambda$  außerhalb der mit Einsen gefüllten Diagonalen. Diese Matrizen sind obere oder untere Dreiecksmatrizen. Der Multiplikationstyp entsteht aus der Einheitsmatrix  $E_n$  durch Ersetzung einer Eins durch einen von Null verschiedenen Skalar. Die Namensgebung wird durch die Wirkung der Multiplikation einer Matrix  $A$  mit einer Elementarmatrix erklärt:

Matrizenprodukte mit Elementarmatrizen $W \in K^{m \times m}$ von links	
Typ	$WA$ entsteht aus $A \in K^{m \times n}$ durch ...
$W = W_{ij}(\lambda), i \neq j$	Addition des $\lambda$ -Fachen der $j$ -ten Zeile zur $i$ -ten Zeile
$W = W_{ii}(\lambda)$	Multiplikation der $i$ -ten Zeile mit $\lambda$

Analoge Aussagen mit „Spalte“ statt „Zeile“ gelten für Produkte mit Elementarmatrizen  $W \in K^{n \times n}$  von rechts. Zu beachten ist lediglich, dass in  $AW$  für  $W = W_{ij}(\lambda)$  das  $\lambda$ -Fache der  $i$ -ten Spalte zur  $j$ -ten Spalte von  $A \in K^{m \times n}$  addiert wird.

Die Elementarmatrizen sind invertierbar und ihre Inversen sind Elementarmatrizen. Es gilt:

Typ	inverse Matrix
$W_{ij}(\lambda), i \neq j$	$W_{ij}(-\lambda)$
$W_{ii}(\lambda), \lambda \neq 0$	$W_{ii}(1/\lambda)$



Die Elementarmatrizen eignen sich zur Manipulation und Vereinfachung von allgemeinen Matrizen. Ein Paradebeispiel ist die Invertierung einer Matrix  $A \in GL(n, K)$ . Wir können Elementarmatrizen  $L_1, \dots, L_k \in GL(n, K)$  finden, die durch Linksmultiplikation  $A$  schrittweise ausräumen, sodass

$$L_k \dots L_1 A = E_n.$$

Dann ist  $L_k \dots L_1 = A^{-1}$ . Wegen  $L_k \dots L_1 = L_k \dots L_1 E_n$  können wir also  $A^{-1}$  bestimmen, indem wir simultan zur Umformung von  $A$  die Matrix  $E_n$  in analoger Weise behandeln: Aus  $A$  wird  $E_n$  und aus  $E_n$  wird  $A^{-1}$ . Wir führen das Verfahren an einem Beispiel vor (genauer und allgemeiner wird das „Ausräumen“ in 5.12 behandelt).

### Beispiel: Invertierung einer Matrix

$$\begin{aligned} A_0 &= \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = E_3 & A_4 &= \begin{pmatrix} 1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 1 \\ 1 & -2 & 1 \end{pmatrix} \\ A_1 &= \begin{pmatrix} 1 & 1 & 0 \\ 0 & -1 & 1 \\ 1 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} & A_5 &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & -1 & 1 \\ 0 & -1 & 1 \\ 1 & -2 & 1 \end{pmatrix} \\ A_2 &= \begin{pmatrix} 1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & -2 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix} & A_6 &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & -1 \\ 1 & -2 & 1 \end{pmatrix} \\ A_3 &= \begin{pmatrix} 1 & 1 & 0 \\ 0 & -1 & 1 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & -2 & 1 \end{pmatrix} & A_7 &= \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & -1 & 1 \\ 0 & 1 & -1 \\ -1 & 2 & -1 \end{pmatrix} = A^{-1} \end{aligned}$$

Es gilt  $A_0 = A, \quad A_1 = L_1 A_0, \quad A_2 = L_2 A_1, \quad \dots,$   
 $A_7 = E_3 = L_7 \dots L_1 A_0 = A^{-1} A$

mit Additionstypen  $L_1, \dots, L_5$  und Multiplikationstypen  $L_6, L_7$ . Das Verfahren lässt sich auf jedes  $A \in K^{n \times n}$  anwenden, um zu testen, ob  $A$  invertierbar ist: Wird eine Nullzeile oder Nullspalte erreicht, so ist  $A$  singular (denn Matrizen mit Nullzeilen oder Nullspalten sind singular, und ist  $A$  invertierbar, so auch alle  $A_i$ ).

Analog kann man  $A^{-1} = R_1 \dots R_k$  durch Spaltenoperationen von rechts gewinnen.

Wir fassen unsere Überlegungen noch einmal in dem folgenden überraschenden Ergebnis zusammen:

**Satz** (Erzeugung von  $GL(n, K)$  durch Elementarmatrizen)

■ Jede invertierbare Matrix ist ein Produkt von Elementarmatrizen.

## 5.7 Die Permutationsmatrizen

### Definition (Permutationsmatrix, Transpositionsmatrix)

Seien  $K$  ein Körper,  $n \geq 1$  und  $\pi \in S_n = \{\sigma \mid \sigma: \{1, \dots, n\} \rightarrow \{1, \dots, n\} \text{ ist bijektiv}\}$ . Dann heißt die Matrix

$$P_\pi = \begin{pmatrix} e_{\pi(1)} & \dots & e_{\pi(n)} \end{pmatrix} \in K^{n \times n}$$

die zu  $\pi$  gehörige *Permutationsmatrix*. Ist  $\pi$  eine Transposition, so heißt  $P_\pi$  eine *Transpositionsmatrix*. Vertauscht eine Transposition  $\pi$  die Zahlen  $i \neq j$ , so schreiben wir auch kurz  $P_{ij}$  für die zugehörige Transpositionsmatrix.

$\begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}$	<p>Die Matrix links ist die Permutationsmatrix <math>P_\pi</math> für <math>\pi = (3, 4, 1, 5, 6, 2)</math>.</p> <p>Die Matrix rechts ist die Transpositionsmatrix <math>P_{26}</math>.</p>	$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix}$
--	---	--

Die Spalten der Permutationsmatrix  $P_\pi$  sind die gemäß  $\pi = (\pi(1), \dots, \pi(n))$  umgeordneten kanonischen Einheitsvektoren  $e_1, \dots, e_n$ . Jede Zeile und jede Spalte von  $P_\pi$  hat genau einen Eins-Eintrag und sonst nur Nullen. Jede Permutationsmatrix geht aus  $E_n$  durch Vertauschung von Spalten hervor. Bei den spezielleren Transpositionsmatrizen werden genau zwei verschiedene Spalten ausgetauscht. Für alle  $i \neq j$  gilt  $P_{ij} = P_{ji}$ .

### Beispiele

- (1) Für  $n = 2$  gibt es neben  $E_2$  nur noch die Permutationsmatrix

$$P_{12} = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \text{ die zudem eine Transpositionsmatrix ist. Es gilt } P_{12}^2 = E_2.$$

- (2) Für  $n = 3$  gibt es neben  $E_3$  noch die fünf Permutationsmatrizen

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

Für alle  $n$  gibt es genau  $n! = |S_n|$  viele Permutationsmatrizen und genau  $n(n-1)/2$  Transpositionsmatrizen („2 aus  $n$ “). Die Transpositionsmatrizen sind durch genau zwei Null-Einträge auf der Diagonalen charakterisiert.

Es gilt  $P_\pi(i, j) = 1$  genau dann, wenn  $\pi(j) = i$ . Dies ist äquivalent zu  $\pi^{-1}(i) = j$ . Damit sind die Zeilen von  $P_\pi$  die gemäß  $\pi^{-1}$  angeordneten Einheitsvektoren  $e_1, \dots, e_n$ :

$$P_\pi = \begin{pmatrix} e_{\pi(1)} & \dots & e_{\pi(n)} \end{pmatrix} = \begin{pmatrix} e_{\pi^{-1}(1)} \\ \dots \\ e_{\pi^{-1}(n)} \end{pmatrix}.$$

Man rechnet nach, dass für alle  $\pi, \sigma \in S_n$  gilt:

$$P_\pi P_\sigma = P_{\pi \circ \sigma}, \quad P_\pi^{-1} = P_{\pi^{-1}}. \quad (\text{Kompositions- und Invertierungsregel})$$

Permutationsmatrizen wirken auf Vektoren und andere Matrizen wie folgt.

### Matrix-Vektor-Produkt

Für alle  $\pi \in S_n$  und alle  $x \in K^n$  gilt

$$P_\pi x = \begin{pmatrix} e_{\pi(1)} & \dots & e_{\pi(n)} \end{pmatrix} x = x_1 e_{\pi(1)} + \dots + x_n e_{\pi(n)} = (x_{\pi^{-1}(1)}, \dots, x_{\pi^{-1}(n)}).$$

Die  $i$ -te Komponente von  $x$  sitzt in  $y = P_\pi x$  an der Stelle  $j = \pi(i)$ . Damit sitzt an der  $j$ -ten Stelle von  $y$  die  $i = \pi^{-1}(j)$ -te Komponente von  $x$ .

### Matrizenprodukt von links und rechts

Für alle  $A \in K^{m \times n}$ ,  $\pi \in S_n$  und  $\sigma \in S_m$  gilt

$$AP_\pi = \begin{pmatrix} A e_{\pi(1)} & \dots & A e_{\pi(n)} \end{pmatrix} = \begin{pmatrix} a_{\pi(1)} & \dots & a_{\pi(n)} \end{pmatrix} \quad \text{mit den Spalten } a_1, \dots, a_n \text{ von } A,$$

$$P_\sigma A = \begin{pmatrix} a_{\sigma^{-1}(1)} \\ \dots \\ a_{\sigma^{-1}(m)} \end{pmatrix} \quad \text{mit den Zeilen } a_1, \dots, a_m \text{ von } A.$$

Die Multiplikation mit  $P_\pi$  von rechts vertauscht also die Spalten von  $A$ , während die Multiplikation mit  $P_\sigma$  von links die Zeilen von  $A$  vertauscht. Speziell sind in  $AP_{ij}$  die Spalten  $i$  und  $j$  vertauscht und in  $P_{ij}A$  die Zeilen  $i$  und  $j$ .

Wie jede invertierbare Matrix lässt sich eine Permutationsmatrix als Produkt von Elementarmatrizen darstellen. Dies lässt sich aber auch leicht direkt einsehen. Für die Transpositionen gilt

$$P_{ij} = W_{ij}(-1) W_{ij}(1) W_{ji}(-1) W_{ji}(1) \quad \text{für alle } i \neq j.$$

Stellt man nun ein  $\pi \in S_n$  als Komposition von Transpositionen dar, so ergibt sich eine Darstellung von  $P_\pi$  als Produkt von Transpositionsmatrizen.

Da das Vertauschen von Zeilen und Spalten speziell beim Umgang mit Gleichungssystemen als elementare Operation angesehen wird, werden die Transpositionsmatrizen oft als weiterer Typ von Elementarmatrizen zugelassen.

## 5.8 Basiswechsel und Transformationsformel

### Definition (Transformationsmatrix eines Basiswechsels)

Sei  $V$  ein  $n$ -dimensionaler  $K$ -Vektorraum, und seien  $\mathcal{A} = (v_1, \dots, v_n)$  und  $\mathcal{B} = (w_1, \dots, w_n)$  Basen von  $V$ . Dann heißt die darstellende Matrix  $T \in K^{n \times n}$  der Identität  $\text{id} : V \rightarrow V$  bzgl.  $\mathcal{A}$  und  $\mathcal{B}$  die *Transformationsmatrix* oder *Übergangsmatrix* des Basiswechsels von  $\mathcal{A}$  nach  $\mathcal{B}$ .

In den Spalten einer  $f : V \rightarrow W$  darstellenden Matrix stehen die Koordinatenvektoren bzgl.  $\mathcal{B}$  der Bilder der Basisvektoren in  $\mathcal{A}$  unter  $f$ . Bei einem Basiswechsel ist  $f$  die Identität. Damit gilt für die Transformationsmatrix  $T$ :

$$\begin{array}{ccc} V & \xrightarrow{\text{id}} & V \\ \Phi_{\mathcal{A}} \downarrow & & \downarrow \Phi_{\mathcal{B}} \\ K^n & \xrightarrow{T} & K^n \end{array}$$

*In den Spalten von  $T$  stehen die neuen Koordinaten von  $v_1, \dots, v_n$ .*

Für die  $T$  zugeordnete lineare Abbildung  $f_T : K^n \rightarrow K^n$  gilt

$$f_T = \Phi_{\mathcal{B}} \circ \text{id} \circ \Phi_{\mathcal{A}}^{-1} = \Phi_{\mathcal{B}} \circ \Phi_{\mathcal{A}}^{-1}.$$

Damit ist  $Tx = \Phi_{\mathcal{B}}(\Phi_{\mathcal{A}}^{-1}(x))$  für alle  $x \in K^n$ . Weiter lesen wir aus dem Diagramm ab:

*$T^{-1}$  ist die Transformationsmatrix des Basiswechsels von  $\mathcal{B}$  nach  $\mathcal{A}$ .*

Ein wichtiger Spezialfall ist:

### Basiswechsel für $V = K^n$

Schreiben wir die Basisvektoren als Spalten in zwei Matrizen  $A, B \in GL(n, K)$ ,

$$A = \begin{pmatrix} v_1 & \dots & v_n \end{pmatrix}, \quad B = \begin{pmatrix} w_1 & \dots & w_n \end{pmatrix},$$

so sind  $A^{-1}$  und  $B^{-1}$  die darstellenden Matrizen der Koordinatenabbildungen, d. h., für alle  $x \in K^n$  gilt

$$\Phi_{\mathcal{A}}(x) = A^{-1}x, \quad \Phi_{\mathcal{B}}(x) = B^{-1}x.$$

Folglich ist

$$T = B^{-1}A.$$

Ist eine der beteiligten Basen die Standardbasis des  $K^n$ , so gilt:

$$\begin{array}{ccc} K^n & \xrightarrow{E_n} & K^n \\ A^{-1} \downarrow & & \downarrow B^{-1} \\ K^n & \xrightarrow{T} & K^n \end{array}$$

Für ...	gilt ...	In den Spalten von $T$ stehen ...
$\mathcal{A} = (e_1, \dots, e_n)$	$T = B^{-1}$	die neuen Koordinaten der Standardvektoren
$\mathcal{B} = (e_1, \dots, e_n)$	$T = A$	die alten Basisvektoren

### Beispiele

- (1) Die Drehmatrix  $A_\varphi \in \mathbb{R}^{2 \times 2}$  ist die Transformationsmatrix für jeden Basiswechsel des  $\mathbb{R}^2$ , für den die neue Basis  $\mathcal{B}$  aus den um den Winkel  $\varphi$  gedrehten Vektoren der alten Basis  $\mathcal{A}$  besteht.
- (2) Ist  $\mathcal{A} = (e_1, \dots, e_n)$  und  $\mathcal{B} = (e_{\pi(1)}, \dots, e_{\pi(n)})$  für eine Permutation  $\pi \in S_n$ , so ist  $P_\pi$  die Transformationsmatrix des Basiswechsels von  $\mathcal{A}$  nach  $\mathcal{B}$ . Diesen Wechsel kann man sich als Umnummerierung der Koordinatenachsen vorstellen.

Als Nächstes untersuchen wir, wie sich die darstellende Matrix einer linearen Abbildung beim Wechsel der Basen verändert.

### Die Transformationsformel

Seien  $V, W$   $K$ -Vektorräume,  $n = \dim(V)$ ,  $m = \dim(W)$ ,  $\mathcal{A}, \mathcal{A}'$  Basen von  $V$  und  $\mathcal{B}, \mathcal{B}'$  Basen von  $W$ . Weiter sei  $f: V \rightarrow W$  eine lineare Abbildung. Wir setzen

$A = A_f$  bzgl.  $\mathcal{A}, \mathcal{B}$ ,

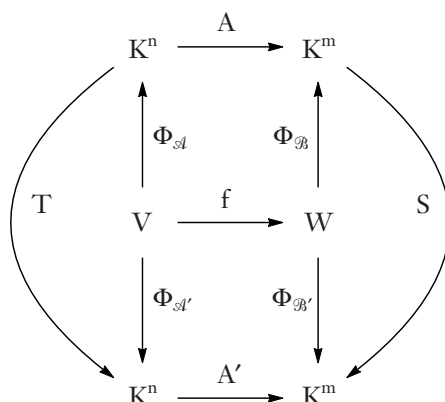
$A' = A_f$  bzgl.  $\mathcal{A}', \mathcal{B}'$ ,

$T =$  „die Transformationsmatrix von  $\mathcal{A}$  nach  $\mathcal{A}'$ “,

$S =$  „die Transformationsmatrix von  $\mathcal{B}$  nach  $\mathcal{B}'$ “.

Das Diagramm zeigt:

$A' = S A T^{-1}$ . (Transformationsformel)



Eine wichtige Anwendung der Formel ist:

### Charakterisierung der Äquivalenz

Zwei Matrizen  $A, A' \in K^{m \times n}$  sind genau dann äquivalent, wenn es  $S \in GL(m, K)$  und  $T \in GL(n, K)$  gibt mit  $A' = S A T^{-1}$ . Denn genau in diesem Fall stellen  $A$  und  $A'$  die gleiche lineare Abbildung für geeignete Basen dar.

Der Spezialfall  $V = W$ ,  $\mathcal{A} = \mathcal{B}$  und  $\mathcal{A}' = \mathcal{B}'$  motiviert:

### Definition (ähnliche Matrizen)

Zwei Matrizen  $A, A' \in K^{n \times n}$  heißen *ähnlich*, falls es ein  $S \in GL(n, K)$  gibt mit  $A' = S A S^{-1}$ .

Nach der Transformationsformel sind  $A, A'$  genau dann ähnlich, wenn es eine lineare Abbildung  $f: V \rightarrow V$  und Basen  $\mathcal{A}, \mathcal{A}'$  von  $V$  gibt mit  $A = A_f$  bzgl.  $\mathcal{A}, \mathcal{A}$  und  $A' = A_f$  bzgl.  $\mathcal{A}', \mathcal{A}'$ . Wir werden in 5.12 sehen, wie man  $S$  und  $T^{-1}$  für ein gegebenes  $A$  so berechnen kann, dass  $B = S A T^{-1}$  in Normalform ist. Mit dem Problem, ein  $S$  zu finden, für welches  $S A S^{-1}$  möglichst einfach ist, befassen wir uns im achten Kapitel.

## 5.9 Die Transposition

### Definition (transponierte Matrix, Transposition, symmetrische Matrix)

Seien  $K$  ein Körper,  $m, n \geq 1$  und  $A \in K^{m \times n}$ . Dann ist die zu  $A$  *transponierte Matrix*

$A^t \in K^{n \times m}$  definiert durch  $A^t(i, j) = A(j, i)$  für alle  $1 \leq i \leq m$ ,  $1 \leq j \leq n$ . Ist  $m = n$  und  $A^t = A$ , so heißt  $A$  *symmetrisch*.

Die Transposition vertauscht Zeilen und Spalten. Ist  $A$  quadratisch, so geht  $A^t$  durch Spiegelung an der Diagonale aus  $A$  hervor. Die Transposition, die  $A \in K^{m \times n}$  auf  $A^t \in K^{n \times m}$  abbildet, ist ein Vektorraum-Isomorphismus. Es gilt

$$(A^t)^t = A, \quad (\lambda A)^t = \lambda A^t,$$

$$(A + B)^t = A^t + B^t, \quad (AB)^t = B^t A^t.$$

Für  $A \in GL(n, K)$  gilt zudem

$$(A^t)^{-1} = (A^{-1})^t.$$

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

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

### Beispiele

(1) Jede Diagonalmatrix ist symmetrisch. Obere (untere) Dreiecksmatrizen werden durch Transposition zu unteren (oberen) Dreiecksmatrizen.

(2) Die symmetrischen  $n \times n$ -Matrizen bilden einen Unterraum des  $K^{n \times n}$ . Das Produkt zweier symmetrischer Matrizen kann unsymmetrisch sein.

$$\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} 2 & 4 \\ 3 & 7 \end{pmatrix}$$

Wir betrachten zwei nützliche Produktbildungen, die sich mit Hilfe der Transposition elegant einführen und handhaben lassen.

### Das Produkt $x^t y$

Seien  $K$  ein Körper und  $n \geq 1$ . Dann gilt für alle  $x, y \in K^n$

$$x^t y = \begin{pmatrix} x_1 & \dots & x_n \end{pmatrix} \begin{pmatrix} y_1 \\ \dots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 y_1 + \dots + x_n y_n \end{pmatrix} = x_1 y_1 + \dots + x_n y_n \in K.$$

Dabei verwenden wir unsere Konvention, ein  $x = (x_1, \dots, x_n) \in K^n$  als einspaltige Matrix zu lesen. Diese Matrix wird durch die Transposition zu einer einzeiligen Matrix. Als Merkregel gilt, dass das „t“ bei  $(x_1, \dots, x_n)^t$  einfach die Kommata löscht. Insgesamt definiert  $x^t y$  eine Abbildung von  $K^n \times K^n$  nach  $K$ .

**Das Produkt  $x y^t$** 

Seien  $K$  ein Körper und  $m, n \geq 1$ . Dann gilt für alle  $x \in K^m$  und  $y \in K^n$

$$x y^t = \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} (y_1 \ \dots \ y_n) = \begin{pmatrix} x_1 y_1 & \dots & x_1 y_n \\ \dots & \dots & \dots \\ x_m y_1 & \dots & x_m y_n \end{pmatrix} = (x_i y_j)_{ij} \in K^{m \times n}.$$

Die Spalten von  $x y^t$  sind skalare Vielfache von  $x$ , die Zeilen skalare Vielfache von  $y$ . Das Produkt  $x y^t$  definiert eine Abbildung von  $K^m \times K^n$  nach  $K^{m \times n}$ .

**Beispiele**

- (1) Für die kanonischen Basisvektoren  $e_1, \dots, e_n$  des  $K^n$  gilt

$$e_i^t e_j = \delta_{ij} \in \{0, 1\}, \quad e_i e_j^t = \delta_{ij} E_n \in K^{n \times n}.$$

- (2) Das Produkt  $C = AB$  für  $A \in K^{m \times k}$ ,  $B \in K^{k \times n}$  können wir definieren durch

$$c_{ij} = a_i^t b_j \text{ mit der } i\text{-ten Zeile } a_i \text{ von } A \text{ und der } j\text{-ten Spalte } b_j \text{ von } B.$$

- (3) Wir wissen schon, dass  $A e_j$  die  $j$ -Spalte von  $A \in K^{m \times n}$  ist. Nun ergänzen wir:

$$e_i^t A = \text{„die } i\text{-te Zeile von } A\text{“} \quad \text{für alle } 1 \leq i \leq m.$$

Weiter gilt  $e_i^t A e_j = a_{ij}$  für alle  $i, j$ .

Zwischen  $f_A : K^n \rightarrow K^m$  und  $f_{A^t} : K^m \rightarrow K^n$  besteht keine offensichtliche Beziehung. Ist  $m = n$  und  $A$  symmetrisch, so ist  $f_A = f_{A^t}$ . Für Permutationsmatrizen gilt  $P_\pi^t = P_\pi^{-1}$ , im Allgemeinen hat  $A^t$  aber nichts mit einer Umkehrabbildung zu tun. Den Schlüssel zum Verständnis liefern erst die Dualräume (vgl. 4.12):

**Satz (Dualitätssatz für  $A^t$ )**

Seien  $V, W$  endlich-dimensionale  $K$ -Vektorräume und  $\mathcal{A} = (v_1, \dots, v_n)$ ,  $\mathcal{B} = (w_1, \dots, w_m)$  Basen von  $V, W$ . Weiter seien  $f : V \rightarrow W$  linear und  $A$  die  $f$  bzgl.  $\mathcal{A}, \mathcal{B}$  darstellende Matrix. Dann gilt:

$A^t$  ist die darstellende Matrix der dualen Abbildung  $f^* : W^* \rightarrow V^*$  bzgl.  $\mathcal{B}^*, \mathcal{A}^*$ .

Denn für alle  $1 \leq i \leq m$  gilt

$$f^*(w_i^*) = w_i^* \circ f = a_{i1} v_1^* + \dots + a_{in} v_n^*.$$

Damit sind die Zeilen von  $A$  die Koordinatenvektoren der Bilder von  $w_1^*, \dots, w_m^*$  unter  $f^*$ . Diese Vektoren sind die Spalten der  $f^*$  bzgl.  $\mathcal{B}^*, \mathcal{A}^*$  darstellenden Matrix  $B$ , sodass  $B = A^t$ .

$$\begin{array}{ccc} V & \xrightarrow{f} & W \\ \downarrow \Phi_{\mathcal{A}} & & \downarrow \Phi_{\mathcal{B}} \\ K^n & \xrightarrow{A} & K^m \\ \\ V^* & \xleftarrow{f^*} & W^* \\ \downarrow \Phi_{\mathcal{A}^*} & & \downarrow \Phi_{\mathcal{B}^*} \\ K^n & \xleftarrow{A^t} & K^m \end{array}$$

## 5.10 Der Rang

### Definition (Rang einer Matrix)

Seien  $K$  ein Körper,  $m, n \geq 1$  und  $A \in K^{m \times n}$ . Dann heißt

$$\text{rang}(A) = \dim(\text{Bild}(A))$$

der *Rang* oder *Spaltenrang* der Matrix  $A$ .

$$\begin{pmatrix} a_{11} & \dots & \dots & a_{1j} & \dots & \dots & a_{1n} \\ \dots & \dots & \dots & a_{2j} & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{m1} & \dots & \dots & a_{mj} & \dots & \dots & a_{mn} \end{pmatrix}$$

Der Rang einer Matrix ist definiert als die Dimension des Bildes der zugeordneten linearen Abbildung  $f_A: K^n \rightarrow K^m$ . Die Bezeichnung „Spaltenrang“ wird durch

Sind die markierten Spalten linear unabhängig und zerstört die Hinzunahme einer weiteren Spalte die lineare Unabhängigkeit, so ist ihre Anzahl der Rang von  $A$ .

$$f_A(x) = Ax = x_1 \begin{pmatrix} a_{11} \\ \dots \\ a_{m1} \end{pmatrix} + \dots + x_n \begin{pmatrix} a_{1n} \\ \dots \\ a_{mn} \end{pmatrix} = x_1 a_1 + \dots + x_n a_n$$

klar:  $\text{Bild}(A)$  ist der von den Spalten von  $A$  aufgespannte Unterraum  $U$  des  $K^m$  und  $\text{rang}(A) = \dim(U)$ . Da  $U$  von  $n$  Vektoren erzeugt wird und ein Unterraum des  $K^m$  ist, gilt

$$0 \leq \text{rang}(A) \leq \min(n, m). \quad (\text{Rangabschätzung})$$

Gilt  $\text{rang}(A) = \min(m, n)$ , so hat  $A$  *vollen Rang*. Dies ist zum Beispiel der Fall, wenn die Spalten von  $A$  linear unabhängig sind. Allgemein ist  $\text{rang}(A)$  die Mächtigkeit einer bezüglich der Inklusion maximalen linear unabhängigen Menge von Spalten von  $A$ . Etwas salopp sagt man auch, dass  $\text{rang}(A)$  die Anzahl der linear unabhängigen Spalten von  $A$  ist.

### Beispiele

- (1) Die Nullmatrix ist die einzige Matrix des  $K^{m \times n}$  mit Rang 0.
- (2) Die Drehmatrizen  $A_\varphi \in \mathbb{R}^{2 \times 2}$  haben für alle Winkel  $\varphi$  den Rang 2.
- (3) Der Rang einer Diagonalmatrix ist die Anzahl ihrer von null verschiedenen Diagonaleinträge. Allgemein gilt für eine obere oder untere Dreiecksmatrix  $A$ :  

$$\text{rang}(A) = |\{i \mid A(i, i) \neq 0\}|. \quad (\text{Rangformel für Dreiecksmatrizen})$$
- (4) Für alle von 0 verschiedenen  $x \in K^m$  und  $y \in K^n$  gilt  $\text{rang}(xy^t) = 1$ . Umgekehrt ist jede Matrix des  $K^{m \times n}$  vom Rang 1 von der Form  $xy^t$  mit  $x, y \neq 0$ .
- (5)  $\text{rang}(E_n + (-E_n)) = 0$ ,  $\text{rang}(E_{11}) + \dots + \text{rang}(E_{nn}) = n = \text{rang}(E_{11} + \dots + E_{nn})$ .

Für alle  $A, B \in K^{m \times n}$  gilt  $0 \leq \text{rang}(A + B) \leq \text{rang}(A) + \text{rang}(B)$  (Subadditivität). Das letzte Beispiel zeigt, dass keine besseren allgemeinen Abschätzungen möglich sind.



### Die Dimensionsformel

Für alle  $A \in K^{m \times n}$  gilt nach der Dimensionsformel für  $f_A : K^n \rightarrow K^m$ :

$$n = \dim(\text{Kern}(A)) + \text{rang}(A),$$

$$\text{rang}(A) = n \quad \text{genau dann, wenn} \quad f_A \text{ ist injektiv,}$$

$$\text{rang}(A) = m \quad \text{genau dann, wenn} \quad f_A \text{ ist surjektiv,}$$

$$\text{rang}(A) = m = n \quad \text{genau dann, wenn} \quad f_A \text{ ist bijektiv.}$$

### Rang eines Produkts

Für alle  $A \in K^{m \times n}$ ,  $S \in GL(m, K)$ ,  $T \in GL(n, K)$  gilt

$$\text{rang}(A) = \text{rang}(SA) = \text{rang}(AT) = \text{rang}(SAT),$$

da sich die Dimension des Bildes einer linearen Abbildung durch Vor- und Nachschalten von Isomorphismen nicht ändert. Speziell haben äquivalente Matrizen den gleichen Rang. Allgemein gilt für  $A \in K^{m \times k}$ ,  $B \in K^{k \times n}$  nur die Abschätzung

$$\text{rang}(A) + \text{rang}(B) - k \leq \text{rang}(AB) \leq \min(\text{rang}(A), \text{rang}(B)).$$

Der Begriff des Spaltenrangs legt es nahe, auch den Zeilenrang einer Matrix zu betrachten. Er ist definiert durch

$$\text{Zeilenrang}(A) = \text{rang}(A^t).$$

Der Zeilenrang der Matrix  $A$  ist die Dimension des von den Zeilen von  $A$  erzeugten Unterraums des  $K^m$ . Bemerkenswerterweise gilt:

$$\begin{pmatrix} a_{11} & \dots & \dots & \dots & \dots & \dots & a_{1n} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{i1} & a_{i2} & \dots & \dots & \dots & \dots & a_{in} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ a_{m1} & \dots & \dots & \dots & \dots & \dots & a_{mn} \end{pmatrix}$$

Der Zeilenrang wird analog zum Spaltenrang definiert.

**Satz** (Zeilenrang gleich Spaltenrang)

■ Für alle  $A \in K^{m \times n}$  gilt  $\text{Zeilenrang}(A) = \text{rang}(A)$ .

### Erster Beweis mit Hilfe der dualen Abbildung

$$\text{rang}(A^t) = \dim(\text{Bild}(f_{A^t})) = \dim(\text{Bild}(f_A^*)) = \dim(\text{Bild}(f_A)) \quad (\text{nach 5.4 und 4.12}).$$

### Zweiter Beweis mit Hilfe der Normalform

Ist  $A \in K^{m \times n}$  und  $B$  die zu  $A$  äquivalente Matrix in Normalform mit  $r$  Einsen, so gibt es invertierbare  $S, T$  mit  $B = S A T^{-1}$ . Offenbar gilt  $\text{rang}(B) = r = \text{Zeilenrang}(B)$ . Aus der Äquivalenz von  $A$  und  $B$  und den Regeln für die Transposition ergibt sich

$$\text{rang}(A) = \text{rang}(S A T^{-1}) = \text{rang}(B) = r,$$

$$\text{Zeilenrang}(A) = \text{rang}(A^t) = \text{rang}((T^t)^{-1} A^t S^t) = \text{rang}(B^t) = \text{Zeilenrang}(B) = r.$$

Einen dritten Beweis werden wir in 5.12 kennenlernen.

## 5.11 Die Zeilenstufenform

### Definition (Zeilenstufenform, Pivots, ausgeräumt, diagonal)

Seien  $K$  ein Körper,  $m, n \geq 1$  und  $A \in K^{m \times n}$ . Weiter seien  $a_1, \dots, a_m$  die Zeilen von  $A$ . Wir sagen, dass  $A$  *Zeilenstufenform* ist, falls gilt:

- (a) Es gibt ein  $0 \leq r \leq \min(m, n)$  mit  $a_1, \dots, a_r \neq 0$  und  $a_{r+1} = \dots = a_m = 0$ .
- (b) Es gilt  $p(1) < \dots < p(r)$  für  $p(i) = \min(\{j \mid A(i, j) \neq 0\})$ .

Die Einträge  $A(1, p(1)), \dots, A(r, p(r))$  heißen dann die *Pivots* von  $A$ . Sind alle Pivots gleich 1 und alle Einträge über den Pivots gleich 0, so heißt  $A$  *reduziert*. Gilt  $p(i) = i$  für alle  $i$ , so hat  $A$  *diagonale Pivots*.

$$\begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & 4 & 0 & 1 & 0 & 2 & 0 & 0 & 2 \\ 0 & 0 & 0 & 0 & 0 & 6 & 0 & 0 & 0 & 3 & 0 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Eine  $6 \times 12$ -Matrix in Zeilenstufenform mit  $r = 4$  und Pivots 1, 4, 6, -1 an den Stellen (1, 3), (2, 5), (3, 6), (4, 9). Die Waagrechten des Linienzugs können beliebig lang sein, die Senkrechten haben dagegen stets die Länge eins. Die Matrix kann mit 0-Spalten beginnen und mit 0-Zeilen enden.

Die Pivots einer Matrix in Zeilenstufenform sitzen an den Stufen einer fallenden Treppe, unter der sich nur Nullen befinden. Aus der Treppenanordnung ergibt sich:

### Rang einer Matrix in Zeilenstufenform

Hat  $A \in K^{n \times m}$  Zeilenstufenform mit  $r$  Pivots, so sind die  $r$  Zeilen sowie die  $r$  Spalten, die Pivots enthalten, linear unabhängig. Es gilt  $\text{rang}(A) = r$ .

### Beispiele

$$\begin{pmatrix} 1 & 0 & -2 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

reduziert

$$\begin{pmatrix} 1 & 0 & -1 & 1 & 0 & 0 \\ 0 & 2 & 0 & 0 & 1 & 2 \\ 0 & 0 & 2 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

diagonale Pivots

$$\begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 2 \\ 0 & 0 & 1 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

reduziert mit diagonalen Pivots

Die Zeilenstufenform ist im Hinblick auf lineare Gleichungssysteme von Interesse. Liegt ein lineares Gleichungssystem  $Ax = b$  mit einer Koeffizientenmatrix  $A \in K^{m \times n}$  in Zeilenstufenform und beliebiger rechter Seite  $b \in K^m$  vor, so können wir die Lösbarkeit des Systems direkt ablesen und die Lösungen vergleichsweise einfach bestimmen:

**Lösbarkeitskriterium**

Sei  $A \in K^{m \times n}$  in Zeilenstufenform mit  $r$  Pivots, d.h.  $m - r$  Nullzeilen. Weiter sei  $b \in K^m$ . Dann ist  $Ax = b$  genau dann lösbar, wenn

$$b = \begin{pmatrix} \bar{b} \\ 0 \end{pmatrix}, \quad \bar{b} \in K^r$$

lösbare rechte Seiten

$$(+)\quad b_{r+1} = \dots = b_m = 0.$$

Gilt (+), so ist der Lösungsraum  $L_A(b) = \{x \in K^n \mid Ax = b\}$  ein  $(n - r)$ -dimensionaler affiner Unterraum des  $K^n$  (vgl. 4.8). Die Lösungen lassen sich dann in Abhängigkeit der Qualität der Zeilenstufenform auf verschiedene Weisen bestimmen:

**Form I: A ist reduziert mit diagonalen Pivots**

Es gilt  $A = \begin{pmatrix} E_r & \bar{A} \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} E_r & \bar{a}_{r+1} & \dots & \bar{a}_n \\ 0 & & 0 & \end{pmatrix}$ ,  $\bar{a}_{r+1}, \dots, \bar{a}_n \in K^r$ , sodass

$$L_A(b) = \left\{ \begin{pmatrix} \bar{b} \\ 0 \end{pmatrix} + \lambda_1 \begin{pmatrix} -\bar{a}_{r+1} \\ e_1 \end{pmatrix} + \dots + \lambda_{n-r} \begin{pmatrix} -\bar{a}_n \\ e_{n-r} \end{pmatrix} \mid \lambda_1, \dots, \lambda_{n-r} \in K \right\}$$

Dabei sind  $e_1, \dots, e_{n-r}$  die kanonischen Basisvektoren des  $K^{n-r}$ .

**Form II: A hat diagonale Pivots**

Wir finden Lösungen durch „Rückwärts-Substitution“: Beliebige vorgegebene  $x_{r+1}, \dots, x_n \in K$  ergänzen wir durch

$$x_r = \frac{b_r - a_{r,r+1}x_{r+1} - \dots - a_{r,n}x_n}{a_{r,r}},$$

...

$$x_1 = \frac{b_1 - a_{1,2}x_2 - \dots - a_{1,n}x_n}{a_{1,1}}$$

zu einer Lösung  $x = (x_1, \dots, x_n)$  von  $Ax = b$ .

**Form III: Allgemeiner Fall**

Liegen die Pivots nicht auf der Diagonalen, so können wir dies durch Vertauschung der Spalten von  $A$  erreichen. Es gibt eine Permutationsmatrix  $P = P_\pi \in K^{n \times n}$  derart, dass  $AP$  diagonale Pivots besitzt. Dann gilt

$$L_A(b) = \{x \mid Ax = b\} = \{Px' \mid APx' = b\} = \{Px' \mid x' \in L_{AP}(b)\} = \{(x'_{\pi^{-1}(1)}, \dots, x'_{\pi^{-1}(n)}) \mid x' \in L_{AP}(b)\},$$

sodass wir die Lösungen von  $L_A(b)$  durch die Form I oder II erhalten können. Die rechte Seite  $b$  bleibt bei dieser „Umbenennung der Variablen“ unverändert.

Die Überführung eines beliebigen Systems  $Ax = b$  in ein äquivalentes System  $A'x' = b'$  mit  $A'$  in Zeilenstufenform besprechen wir im folgenden Abschnitt.

## 5.12 Eliminationsverfahren

### Satz (Überführung in Zeilenstufenform)

Seien  $K$  ein Körper,  $n, m \geq 1$  und  $A \in K^{m \times n}$ . Dann gilt:

- (a) Es gibt Elementarmatrizen  $L_1, \dots, L_k \in K^{m \times m}$  derart, dass

$$L_k \dots L_1 A \in K^{m \times n}$$

in Zeilenstufenform mit Eins-Pivots ist. Weiter kann eine reduzierte Zeilenstufenform erreicht werden.

- (b) Es gibt Elementarmatrizen  $L_1, \dots, L_k \in K^{m \times m}$  und eine Permutationsmatrix  $P \in K^{n \times n}$  derart, dass

$$L_k \dots L_1 A P \in K^{m \times n}$$

in reduzierter Zeilenstufenform mit diagonalen Pivots ist.

Derartige Elementarmatrizen können durch *Eliminationsverfahren* gefunden werden. „Eliminieren eines Eintrags  $a_{ij}$ “ bedeutet, dass  $a_{ij} = 0$  durch Multiplikation mit einer Elementarmatrix erreicht wird. Die Verfahren ähneln dem Verfahren zum Invertieren einer Matrix (vgl. 5.6). Da nun rechteckige Matrizen vorliegen und Nullzeilen und -spalten auftreten können, sind einige Modifikationen nötig. Folgende Algorithmen leisten das Gewünschte. Viele Varianten sind denkbar (vgl. die LR-Zerlegung in Überblick 5).

### Die Gauß-Elimination

Sei  $A \in K^{m \times n}$ . Wir betrachten die erste Spalte  $j^*$  von  $A$ , die der Zeilenstufenform entgegensteht. Weiter sei  $(i^*, j^*)$  die Stelle eines erwünschten Pivots.

- (1) Durch Linksmultiplikation mit einem Additionstyp  $W_{i^*,i}(\lambda)$  mit  $i > i^*$  gefolgt von einem Multiplikationstyp  $W_{i^*,j^*}(\lambda)$  erreichen wir  $a_{i^*,j^*} \neq 0$  und  $a_{i^*,j^*} = 1$ . (Durch  $i > i^*$  wird die Zeilenstufenform links der Spalte  $j^*$  bewahrt.)

- (2) Durch Linksmultiplikation mit den Additionstypen

$$W_{ij^*}(-a_{i1}) \quad \text{für } i = i^* + 1, \dots, m$$

eliminieren wir alle Einträge unterhalb des aktuellen Pivots.

Wir wiederholen das Verfahren, bis eine Zeilenstufenform erreicht ist.

### Die Gauß-Jordan-Elimination

Bei dieser Variante eliminieren wir in (2) zusätzlich alle Einträge oberhalb des aktuellen Pivots durch Linksmultiplikation mit  $W_{ij^*}(-a_{ij^*})$  für  $i = 1, \dots, i^* - 1$ .

### Permutation der Spalten

Durch Rechtsmultiplikation mit einer Permutationsmatrix können wir die Spalten der erzeugten Zeilenstufenform so umordnen, dass diagonale Pivots entstehen.

Wir diskutieren einige Anwendungen der Ergebnisse.

### Lösen eines Gleichungssystems $Ax = b$

Für  $A \in K^{m \times n}$  und  $b \in K^m$  sei  $(A \mid b) = (A \ b) \in K^{m \times (n+1)}$  die um die Spalte  $b$  erweiterte Koeffizientenmatrix. Wir bringen nun  $A$  in diagonale Zeilenstufenform, wobei wir alle Additionen und Multiplikationen auch an  $b$  durchführen. Ein Beispiel für die Gauß-Jordan-Elimination mit  $A \in \mathbb{R}^{3 \times 4}$ ,  $b = (-1, 1, -2) \in \mathbb{R}^3$  ist:

$$\begin{aligned}
 (0) \quad & \left( \begin{array}{cccc|c} 1 & -1 & 1 & 0 & -1 \\ 1 & -1 & 1 & -1 & 1 \\ 0 & 0 & 1 & 1 & -2 \end{array} \right) & (1) \quad & \left( \begin{array}{cccc|c} 1 & -1 & 1 & 0 & -1 \\ 0 & 0 & 0 & -1 & 2 \\ 0 & 0 & 1 & 1 & -2 \end{array} \right) & (2) \quad & \left( \begin{array}{cccc|c} 1 & -1 & 1 & 0 & -1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 & -2 \end{array} \right) \\
 (3) \quad & \left( \begin{array}{cccc|c} 1 & -1 & 1 & 0 & -1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & -2 \end{array} \right) & (4) \quad & \left( \begin{array}{cccc|c} 1 & -1 & 0 & 0 & -1 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & -2 \end{array} \right) & (5) \quad & \left( \begin{array}{cccc|c} 1 & 0 & 0 & -1 & -1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & -2 \end{array} \right)
 \end{aligned}$$

Die Lösungen des Systems  $A_5 x' = b_5$  in (5) sind gegeben durch

$$x' = \begin{pmatrix} -1 \\ 0 \\ -2 \\ 0 \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ 0 \\ 0 \\ 1 \end{pmatrix}, \quad \lambda \in \mathbb{R}.$$

Rückgängigmachen der Permutation  $A_5 = A_4 P_\pi$  mit  $\pi = (1, 3, 4, 2)$  in (4)  $\rightarrow$  (5) liefert die Lösungen  $x = P_\pi x'$  des ursprünglichen Systems  $Ax = b$  in (0) (vgl. 5.11):

$$x = \begin{pmatrix} -1 \\ 0 \\ 0 \\ -2 \end{pmatrix} + \lambda \begin{pmatrix} 1 \\ 1 \\ 0 \\ 0 \end{pmatrix}, \quad \lambda \in \mathbb{R}.$$

### Zeilenrang gleich Spaltenrang, dritter Beweis

Sei  $A \in K^{m \times n}$ . Man zeigt (1) mit Hilfe des Austauschlemmas, dass die Linksmultiplikation mit einer Elementarmatrix weder den Zeilen- noch den Spaltenrang von  $A$  ändert, (2) dass eine Matrix in Zeilenstufenform mit  $r$  Pivots den Zeilen- und Spaltenrang  $r$  besitzt. Überführt man also  $A$  in eine Zeilenstufenform  $B = L_k \dots L_1 A$  mit  $r$  Pivots, so gilt

$$\text{rang}(A) = \text{rang}(B) = r = \text{Zeilenrang}(B) = \text{Zeilenrang}(A).$$

### Gewinnung der Normalform

Ist  $A \in K^{m \times n}$  und  $L_k \dots L_1 A P$  mit  $r$  Pivots wie in (b), so können wir die Spalten  $r+1, \dots, n$  durch Rechtsmultiplikation mit Elementarmatrizen  $R_1 \dots R_s$  eliminieren. Dann ist  $L_k \dots L_1 A P R_1 \dots R_s$  die zu  $A$  äquivalente Matrix in Normalform.



## 6. Kapitel

---

# Euklidische und unitäre Vektorräume

---

## 6.1 Das kanonische Skalarprodukt im $\mathbb{R}^n$

**Definition** (*Skalarprodukt, orthogonal, euklidische Norm, Länge, normiert*)

Sei  $n \geq 1$ . Dann heißt die Abbildung  $\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$  mit

$$\langle x, y \rangle = x^t y = x_1 y_1 + \dots + x_n y_n$$

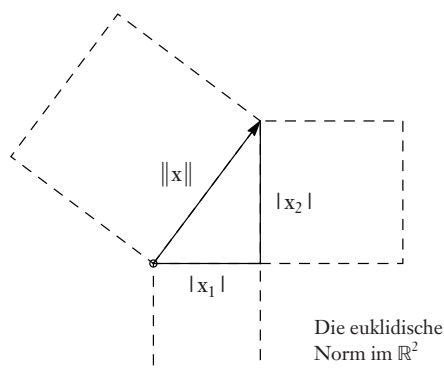
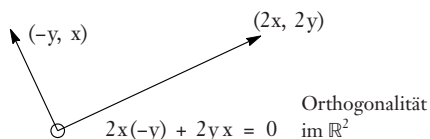
das *kanonische Skalarprodukt* oder *kanonische innere Produkt* des  $\mathbb{R}^n$ .

Sind  $x, y \in \mathbb{R}^n$  mit  $\langle x, y \rangle = 0$ , so sagen wir, dass  $x$  und  $y$  *orthogonal* sind oder *aufeinander senkrecht stehen*. Weiter definieren wir die *euklidische Norm*

$\|\cdot\| : \mathbb{R}^n \rightarrow [0, \infty[$  des  $\mathbb{R}^n$  durch

$$\|x\| = \sqrt{\langle x, x \rangle} = \sqrt{x_1^2 + \dots + x_n^2}.$$

Die Zahl  $\|x\|$  heißt die *euklidische Norm* oder *Länge* von  $x$ . Gilt  $\|x\| = 1$ , so heißt  $x$  *normiert*.



In diesem Kapitel untersuchen wir geometrische Begriffe wie Länge, Orthogonalität, Winkel in Vektorräumen über den Skalarkörpern  $K = \mathbb{R}$  oder  $K = \mathbb{C}$ . Sie werden mit Hilfe eines Skalarprodukts eingeführt.

In den beiden ersten Abschnitten betrachten wir besonders wichtige Skalarprodukte auf dem  $\mathbb{R}^n$  und  $\mathbb{C}^n$ . Ihre Eigenschaften motivieren die allgemeine Definition in Abschnitt 6.3. Ein erstes Beispiel für die geometrische Kraft des Skalarprodukts ist die Einführung der euklidischen Norm oder Länge. Für alle  $x \in \mathbb{R}^n$  ist  $\|x\| = \sqrt{\langle x, x \rangle}$  definiert, da

$$\langle x, x \rangle = x_1^2 + \dots + x_n^2 \geq 0.$$

Der Satz des Pythagoras motiviert die Bezeichnung als *Länge*.

### Zur Punkt-Notation für Abbildungen

Die Notation  $\langle \cdot, \cdot \rangle : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$  bedeutet die Abbildung  $F : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}$  mit  $F(x, y) = \langle x, y \rangle$  für alle  $x, y$ . Analoges gilt für die Norm  $\|\cdot\|$ . Die Punkt-Notation erlaubt es, viele Abbildungen unkompliziert anzugeben. Für alle  $y \in \mathbb{R}^n$  ist zum Beispiel  $\langle \cdot, y \rangle : \mathbb{R}^n \rightarrow \mathbb{R}$  die Abbildung  $G : \mathbb{R}^n \rightarrow \mathbb{R}$  mit  $G(x) = \langle x, y \rangle$  für alle  $x$ .

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

Die Abbildungen $\langle x, \cdot \rangle, \langle \cdot, y \rangle : \mathbb{R}^n \rightarrow \mathbb{R}$ sind linear.	<i>Bilinearität</i>
$\langle x, y \rangle = \langle y, x \rangle$	<i>Symmetrie</i>
$\langle x, x \rangle > 0$ für alle $x \neq 0$	<i>positive Definitheit</i>



Für die Norm gilt eine der wichtigsten Ungleichungen der Mathematik:

### Cauchy-Schwarz-Ungleichung

Für alle  $x, y \in \mathbb{R}^n$  und  $\lambda \in \mathbb{R}$  gilt aufgrund der Bilinearität

$$0 \leq \|x - \lambda y\|^2 = \langle x - \lambda y, x - \lambda y \rangle = \|x\|^2 - 2\lambda \langle x, y \rangle + \lambda^2 \|y\|^2.$$

Ist nun  $y \neq 0$  und  $\lambda = \langle x, y \rangle / \|y\|^2$ , so erhalten wir die Cauchy-Schwarz-Ungleichung

$$|\langle x, y \rangle| \leq \|x\| \|y\|,$$

die auch für  $y = 0$  gilt. Gleichheit gilt genau dann, wenn  $x$  und  $y$  linear abhängig sind.

Damit können wir einführen:

### Winkel

Für alle  $x, y \in \mathbb{R}^n - \{0\}$  heißt

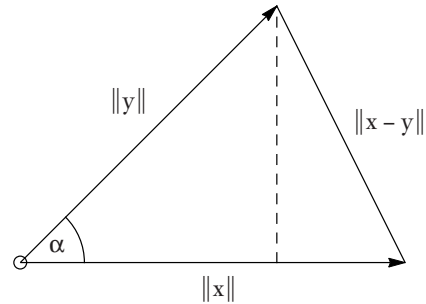
$$\alpha(x, y) = \arccos\left(\frac{\langle x, y \rangle}{\|x\| \|y\|}\right) \in [0, \pi]$$

der von  $x, y$  *eingeschlossene Winkel*.

Nach Cauchy-Schwarz ist der Bruch ein Element von  $[-1, 1]$ , sodass der Arkuskosinus anwendbar ist. Zur

Motivation der Formel betrachten wir das von zwei Vektoren  $x, y \in \mathbb{R}^2$  gebildete Dreieck mit den Seitenlängen  $\|x\|, \|y\|, \|x - y\|$ . Mit Hilfe des Kosinussatzes lässt sich  $\cos(\alpha)$  durch die Seitenlängen ausdrücken. Die Formel folgt nun aus

$$\|x - y\|^2 = \|x\|^2 + \|y\|^2 - 2 \langle x, y \rangle.$$



Der Kosinussatz:

$$\|x - y\|^2 = \|x\|^2 + \|y\|^2 - 2\|x\|\|y\|\cos(\alpha) + \|y\|^2$$

Am Vorzeichen des Skalarprodukts lässt sich ablesen, ob der Winkel  $\alpha$  stumpf oder spitz ist, und  $\alpha$  ist genau dann gleich  $\pi/2 = \arccos(0)$ , wenn  $\langle x, y \rangle = 0$ . Dass die Orthogonalität von  $x, y \in \mathbb{R}^n$  durch die einfache (kosinusfreie) Bedingung  $x_1 y_1 + \dots + x_n y_n = 0$  eingefangen wird, gehört zu den Wundern der Linearen Algebra. Eine überraschende Folgerung ist, dass die Lösungen eines Gleichungssystems  $Ax = 0$  für  $A \in \mathbb{R}^{m \times n}$  aus genau den Vektoren des  $\mathbb{R}^n$  besteht, die auf allen Zeilen  $a_1, \dots, a_m \in \mathbb{R}^n$  von  $A$  senkrecht stehen. Denn die Komponenten von  $Ax \in \mathbb{R}^m$  sind die Skalarprodukte  $\langle a_1, x \rangle, \dots, \langle a_m, x \rangle$ .

### Beispiele

- (1) Das Skalarprodukt mit  $e_1, \dots, e_n \in \mathbb{R}^n$  liefert für jedes  $x \in \mathbb{R}^n$  die Komponenten  $\langle e_1, x \rangle = x_1, \dots, \langle e_n, x \rangle = x_n$ . Insbesondere gilt  $\langle e_i, e_j \rangle = \delta_{ij}$  für alle  $i, j$ . Die Vektoren  $e_1, \dots, e_n$  sind also normiert und paarweise orthogonal zueinander.

- (2) Für  $x = (1, 0)$ ,  $y = (1, 1)$ ,  $z = (-1, 1)$  gilt

$$\langle x, y \rangle = 1, \quad \langle x, z \rangle = -1, \quad \|x\| = 1, \quad \|y\| = \|z\| = \sqrt{2},$$

$$\alpha(x, y) = \arccos(1/\sqrt{2}) = \pi/4, \quad \alpha(x, z) = \arccos(-1/\sqrt{2}) = 3\pi/4.$$

## 6.2 Das kanonische Skalarprodukt im $\mathbb{C}^n$

**Definition** (*Skalarprodukt, orthogonal, euklidische Norm, normiert*)

Sei  $n \geq 1$ . Dann heißt die Abbildung  $\langle \cdot, \cdot \rangle : \mathbb{C}^n \times \mathbb{C}^n \rightarrow \mathbb{C}$  mit

$$\langle z, w \rangle = \overline{z_1} w_1 + \dots + \overline{z_n} w_n \quad \text{für alle } z, w \in \mathbb{C}^n$$

das *kanonische Skalarprodukt* oder *kanonische innere Produkt* des  $\mathbb{C}^n$ . Zwei Vektoren  $z, w \in \mathbb{C}^n$  heißen *orthogonal* oder *stehen senkrecht aufeinander*, falls  $\langle z, w \rangle = 0$ .

Weiter ist die *euklidische Norm*  $\|\cdot\| : \mathbb{C}^n \rightarrow [0, \infty[$  definiert durch

$$\|z\| = \sqrt{\langle z, z \rangle} \quad \text{für alle } z \in \mathbb{C}^n.$$

Gilt  $\|z\| = 1$ , so heißt  $z$  *normiert*.

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

Die Abbildung $\langle z, \cdot \rangle : \mathbb{C}^n \rightarrow \mathbb{C}$ ist linear.	
Die Abbildung $\langle \cdot, w \rangle : \mathbb{C}^n \rightarrow \mathbb{C}$ ist antilinear, d. h.	<i>Sesquilinearität</i>
$\langle z + \lambda z', w \rangle = \langle z, w \rangle + \overline{\lambda} \langle z', w \rangle$ für alle $z' \in \mathbb{C}^n$ , $\lambda \in \mathbb{C}$ .	
$\langle z, w \rangle = \overline{\langle w, z \rangle}$	<i>Hermitizität</i>
$\langle z, z \rangle > 0$ für alle $z \neq 0$ .	<i>positive Definitheit</i>

„Sesqui“ bedeutet „anderthalb“ und deutet an, dass die doppelte Linearität der reellen Version modifiziert werden muss. Könnte man sich das Leben nicht einfacher machen und auf die Konjugation in der Definition verzichten? Die Antwort ist „nein“. Es gilt

$$i \cdot i = -1 \quad \text{für } n = 1, \quad 1 \cdot 1 + i \cdot i = 1 - 1 = 0 \quad \text{für } n = 2,$$

sodass die positive Definitheit ohne Konjugation verletzt ist. Weiter wird die Konjugation für die Definition der Norm benötigt: Für alle  $z \in \mathbb{C}^n$  gilt

$$\langle z, z \rangle = \overline{z_1} z_1 + \dots + \overline{z_n} z_n = |z_1|^2 + \dots + |z_n|^2 \geq 0,$$

sodass das komplexe Skalarprodukt von  $z$  mit sich selbst eine nichtnegative reelle Zahl ist, deren reelle Wurzel wir ziehen können. Im Allgemeinen ist  $\text{Im}(\langle z, w \rangle) \neq 0$ .

Sind  $x, y \in \mathbb{R}^n \subseteq \mathbb{C}^n$ , so stimmen das reelle und komplexe Skalarprodukt der beiden Vektoren überein. Die komplexe Version setzt also die reelle fort.

### Bemerkung

Oft wird die Konjugation auch in der zweiten Komponente durchgeführt. Beide Definitionen sind gleich gut und erzeugen denselben Orthogonalitätsbegriff, da

$$\overline{z_1} w_1 + \dots + \overline{z_n} w_n = 0 \quad \text{genau dann, wenn} \quad z_1 \overline{w_1} + \dots + z_n \overline{w_n} = 0.$$

Wir konjugieren im Folgenden immer in der ersten Komponente.

**Beispiele**

- (1) Für die Standardbasisvektoren  $e_1, \dots, e_n$  des  $\mathbb{C}^n$  gilt  $\langle e_i, e_j \rangle = \delta_{ij}$ , sodass diese Vektoren normiert und paarweise orthogonal sind.
- (2) Für  $n = 1$  gilt  $\langle i, z \rangle = -i z$ . Für  $n = 2$  gilt  $\langle (i, i), (z, w) \rangle = -i z - i w = -i(z + w)$ .
- (3) Sind  $m, n \geq 1$  und  $A \in \mathbb{C}^{m \times n}$ , so gilt  $Az = 0$  für  $z \in \mathbb{C}^n$  genau dann, wenn  $(z_1, \dots, z_n)$  senkrecht auf allen Zeilen der *konjugierten Matrix*  $\bar{A} = (\bar{a}_{ij})_{ij}$  steht.

Wie im reellen Fall ist unverzichtbar:

**Cauchy-Schwarz-Ungleichung**

Für alle  $z, w \in \mathbb{C}^n$  gilt

$$|\langle z, w \rangle| \leq \|z\| \|w\|. \quad (\text{Cauchy-Schwarz-Ungleichung})$$

Gleichheit gilt genau dann, wenn  $z$  und  $w$  linear abhängig sind.

Der Beweis kann analog geführt werden, wobei man nun verwendet, dass

$$\begin{aligned} 0 \leq \|z - \lambda w\|^2 &= \|z\|^2 - \lambda \langle z, w \rangle - \overline{\lambda \langle z, w \rangle} + |\lambda|^2 \|w\|^2 = \\ &= \|z\|^2 - 2 \operatorname{Re}(\lambda \langle z, w \rangle) + |\lambda|^2 \|w\|^2. \end{aligned}$$

Aus der Cauchy-Schwarz-Ungleichung gewinnen wir:

**Dreiecksungleichung**

Für alle  $z, w \in \mathbb{C}^n$  gilt

$$\begin{aligned} \|z + w\|^2 &= \|z\|^2 + 2 \operatorname{Re}(\langle z, w \rangle) + \|w\|^2 \leq \|z\|^2 + 2 |\langle z, w \rangle| + \|w\|^2 \leq \\ &= \|z\|^2 + 2 \|z\| \|w\| + \|w\|^2 = (\|z\| + \|w\|)^2, \end{aligned}$$

sodass aufgrund der Monotonie der reellen Quadratfunktion gilt, dass

$$\|z + w\| \leq \|z\| + \|w\|.$$

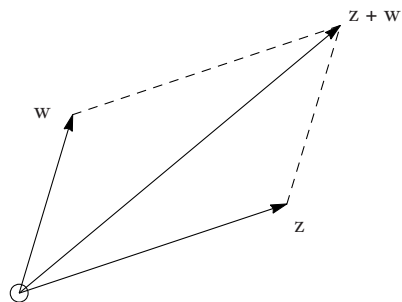
Nützlich sind auch die Abschätzungen

$$\|z\| - \|w\| \leq \|z \pm w\| \leq \|z\| + \|w\|.$$

Sie folgen aus der Dreiecksungleichung, da

$$\|z \pm w\| \leq \|z\| + \|\pm w\| = \|z\| + \|w\|,$$

$$\|z\| = \|z \pm w \mp w\| \leq \|z \pm w\| + \|w\|.$$



Die Dreiecksungleichung

$$\|z + w\| \leq \|z\| + \|w\|.$$

Der direkte Weg ist der kürzeste.

Da die euklidische Norm des  $\mathbb{C}^n$  die des  $\mathbb{R}^n$  fortsetzt, gelten alle Ungleichungen auch für die euklidische Norm des  $\mathbb{R}^n$ . Dies kann man natürlich auch direkt aus der reellen Ungleichung von Cauchy-Schwarz herleiten.

### 6.3 Allgemeine Skalarprodukte

**Definition** (*Skalarprodukt, inneres Produkt, euklidisch, unitär, orthogonal*)

*Skalarprodukt für reelle Vektorräume*

Sei  $V$  ein  $\mathbb{R}$ -Vektorraum. Eine Abbildung  $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{R}$  heißt ein *Skalarprodukt* oder *inneres Produkt* auf  $V$ , falls für alle  $v, w, v', w' \in V$  und  $\lambda \in \mathbb{R}$  gilt:

- (a)  $\langle v + \lambda v', w \rangle = \langle v, w \rangle + \lambda \langle v', w \rangle,$   
 $\langle v, w + \lambda w' \rangle = \langle v, w \rangle + \lambda \langle v, w' \rangle,$  (Bilinearität)
- (b)  $\langle v, w \rangle = \langle w, v \rangle,$  (Symmetrie)
- (c)  $\langle v, v \rangle > 0$  für alle  $v \neq 0.$  (positive Definitheit)

Ein mit einem Skalarprodukt ausgestatteter  $\mathbb{R}$ -Vektorraum heißt *euklidisch*.

*Skalarprodukt für komplexe Vektorräume*

Sei  $V$  ein  $\mathbb{C}$ -Vektorraum. Eine Abbildung  $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{C}$  heißt ein *Skalarprodukt* oder *inneres Produkt* auf  $V$ , falls für alle  $v, w, v', w' \in V$  und  $\lambda \in \mathbb{C}$  gilt:

- (a)  $\langle v + \lambda v', w \rangle = \langle v, w \rangle + \lambda \langle v', w \rangle,$   
 $\langle v, w + \lambda w' \rangle = \langle v, w \rangle + \lambda \langle v, w' \rangle,$  (Sesquilinearität)
- (b)  $\langle v, w \rangle = \overline{\langle w, v \rangle},$  (Hermitizität)
- (c)  $\langle v, v \rangle > 0$  für alle  $v \neq 0.$  (positive Definitheit)

Ein mit einem Skalarprodukt ausgestatteter  $\mathbb{C}$ -Vektorraum heißt *unitär*.

*Orthogonalität*

Zwei Vektoren  $v, w$  eines euklidischen oder unitären Vektorraums  $V$  heißen *orthogonal* oder *stehen senkrecht aufeinander*, falls  $\langle v, w \rangle = 0$ .

Die essentiellen Eigenschaften der kanonischen Skalarprodukte des  $\mathbb{R}^n$  und  $\mathbb{C}^n$  motivieren die Definition abstrakter Skalarprodukte. Die Skalarenkörper sind  $\mathbb{R}$  oder  $\mathbb{C}$  (um Wurzeln aus  $\langle v, v \rangle$  ziehen zu können), der Vektorraum  $V$  ist ansonsten beliebig.

Wegen  $\langle 0, 0 \rangle = \langle 0 - 0, 0 \rangle = \langle 0, 0 \rangle - \langle 0, 0 \rangle = 0$  und der positiven Definitheit gilt für alle Vektoren  $v$ :  $\langle v, v \rangle = 0$  genau dann, wenn  $v = 0$ .

Die Orthogonalität steht im Zentrum der Theorie. Als Erstes halten wir fest:

#### Orthogonalität impliziert lineare Unabhängigkeit

Sind  $v_1, \dots, v_n$  von null verschiedene und paarweise orthogonale Elemente eines euklidischen oder unitären Vektorraums  $V$ , so ist  $(v_1, \dots, v_n)$  linear unabhängig.

Denn für alle  $1 \leq i \leq n$  und alle Skalare  $\alpha_1, \dots, \alpha_n$  gilt

$$\langle v_i, \alpha_1 v_1 + \dots + \alpha_n v_n \rangle = \alpha_1 \langle v_i, v_1 \rangle + \dots + \alpha_n \langle v_i, v_n \rangle = \alpha_i \langle v_i, v_i \rangle,$$

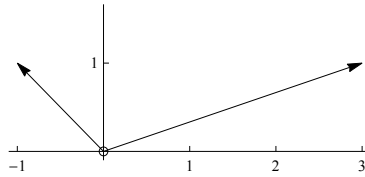
so dass  $\alpha_1 v_1 + \dots + \alpha_n v_n = 0$  wegen  $\langle v_i, v_i \rangle > 0$  nur möglich ist, wenn alle  $\alpha_i$  null sind.

## Beispiele

- (1) Ist  $D = \text{diag}(d_1, \dots, d_n) \in \mathbb{R}^{n \times n}$  mit Diagonaleinträgen  $d_i > 0$ , so definiert

$$\langle x, y \rangle_D = x^t D y = d_1 x_1 y_1 + \dots + d_n x_n y_n \quad \text{für alle } x, y \in \mathbb{R}^n.$$

ein Skalarprodukt auf dem  $\mathbb{R}^n$ . Die Koordinatenprodukte werden mit den Gewichten  $d_i$  versehen. Für  $D = E_n$  ergibt sich das kanonische Skalarprodukt.



Die Vektoren  $(3, 1)$  und  $(-1, 1)$  sind orthogonal für  $D = (1, 3)$ .

- (2) Auf dem Vektorraum  $\mathbb{R}^{m \times n}$  aller reellen  $m \times n$ -Matrizen definiert

$$\langle A, B \rangle = (A^t B)_{11} + \dots + (A^t B)_{nn} = \sum_{i,j} a_{ij} b_{ij} \quad \text{für alle } A, B \in \mathbb{R}^{m \times n}$$

ein Skalarprodukt. Es entsteht, wenn wir die Matrizen des  $\mathbb{R}^{m \times n}$  durch Aneinanderfügen der Zeilen (oder Spalten) in Vektoren des  $\mathbb{R}^{mn}$  verwandeln und dann das kanonische Skalarprodukt des  $\mathbb{R}^{mn}$  verwenden.

- (3) Der euklidische Vektorraum  $\ell_{\mathbb{R}}^2$  der *quadratsummierbaren Folgen* in  $\mathbb{R}$  ist definiert durch

$$\ell_{\mathbb{R}}^2 = \{ (x_n)_{n \in \mathbb{N}} \in \mathbb{R}^{\mathbb{N}} \mid \sum_n |x_n|^2 < \infty \},$$

$$\langle (x_n)_{n \in \mathbb{N}}, (y_n)_{n \in \mathbb{N}} \rangle = \sum_n x_n y_n \quad \text{für alle } x, y \in \ell_{\mathbb{R}}^2.$$

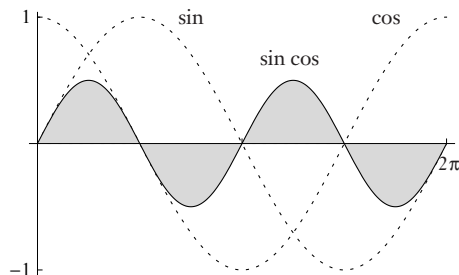
Analog ist der unitäre Vektorraum  $\ell_{\mathbb{C}}^2$  aller quadratsummierbaren Folgen in  $\mathbb{C}$  definiert, wobei nun  $\langle (z_n)_{n \in \mathbb{N}}, (w_n)_{n \in \mathbb{N}} \rangle = \sum_n \overline{z_n} w_n$  für alle  $(z_n)_{n \in \mathbb{N}}, (w_n)_{n \in \mathbb{N}} \in \ell_{\mathbb{C}}^2$ .

Die Vektorräume  $\mathbb{R}^{(\mathbb{N})}$  bzw.  $\mathbb{C}^{(\mathbb{N})}$  aller Folgen mit endlichem Träger sind Teilräume des  $\ell_{\mathbb{R}}^2$  bzw.  $\ell_{\mathbb{C}}^2$  und damit ebenfalls euklidisch bzw. unitär.

- (4) Sei  $I = [a, b]$  mit  $a < b$  ein reelles Intervall, und sei  $V = \mathcal{C}(I, \mathbb{R})$  der  $\mathbb{R}$ -Vektorraum aller stetigen Funktionen  $f: I \rightarrow \mathbb{R}$ . Dann definiert

$$\langle f, g \rangle = \int_a^b f(x) g(x) dx$$

ein Skalarprodukt auf  $V$ . Ist  $I = [0, 2\pi]$ , so sind die auf  $I$  eingeschränkten Sinus- und Kosinusfunktionen orthogonal und insbesondere linear unabhängig.



Orthogonalität in  $V$  bedeutet, dass der signierte Flächeninhalt des Produkts gleich null ist.

- (5) Für  $I = [a, b] \subseteq \mathbb{R}$  und den  $\mathbb{C}$ -Vektorraum  $V = \mathcal{C}(I, \mathbb{C})$  aller stetigen  $f: I \rightarrow \mathbb{C}$  erhält man ein Skalarprodukt wie in (4), wenn man im Integral  $\overline{f(x)}g(x)$  statt  $f(x)g(x)$  verwendet.

## 6.4 Normierte Vektorräume

### Definition (Norm auf $V$ , induzierte Norm eines Skalarprodukts)

Sei  $V$  ein  $K$ -Vektorraum mit  $K = \mathbb{R}$  oder  $K = \mathbb{C}$ . Eine Abbildung  $\|\cdot\| : V \rightarrow [0, \infty[$  heißt eine *Norm* auf  $V$ , falls für alle  $v, w \in V$  und  $\lambda \in K$  gilt:

$$(a) \quad \|\lambda v\| = |\lambda| \|v\|, \quad (\text{Homogenität})$$

$$(b) \quad \|v + w\| \leq \|v\| + \|w\|, \quad (\text{Dreiecksungleichung})$$

$$(c) \quad \|v\| > 0, \text{ falls } v \neq 0. \quad (\text{Definitheit})$$

Ein  $v \in V$  heißt *normiert*, falls  $\|v\| = 1$ . Für alle  $v \in V - \{0\}$  heißt  $N(v) = v/\|v\|$  die *Normierung* von  $v$ .

Ist  $V$  euklidisch oder unitär, so heißt die Abbildung  $\|\cdot\| : V \rightarrow [0, \infty[$  mit

$$(+)\quad \|v\| = \sqrt{\langle v, v \rangle} \quad \text{für alle } v \in V$$

die durch das Skalarprodukt *induzierte Norm* oder die *euklidische Norm* auf  $V$ .

Der Begriff der Norm ist durch die essentiellen Eigenschaften der euklidischen Länge eines Vektors im  $\mathbb{R}^n$  oder  $\mathbb{C}^n$  motiviert. Wie für das allgemeine Skalarprodukt ist  $V$  ein beliebiger  $K$ -Vektorraum mit  $K = \mathbb{R}$  oder  $K = \mathbb{C}$ .

Dass die durch (+) definierte Abbildung tatsächlich eine Norm ist, folgt aus der Linearität und Definitheit des Skalarprodukts. Für  $K = \mathbb{C}$  gilt zum Beispiel

$$\|\lambda v\|^2 = \langle \lambda v, \lambda v \rangle = \lambda \bar{\lambda} \langle v, v \rangle = |\lambda|^2 \|v\|^2 \quad \text{für alle } \lambda \in \mathbb{C}, v \in V,$$

woraus sich die Homogenität ergibt. Die Dreiecksungleichung ist schwieriger zu zeigen. Man beweist hierzu genau wie für die kanonischen Skalarprodukte:

### Cauchy-Schwarz-Ungleichung

Ist  $V$  euklidisch oder unitär, so gilt

$$|\langle v, w \rangle| \leq \|v\| \|w\| \quad \text{für alle } v, w \in V. \quad (\text{Cauchy-Schwarz-Ungleichung})$$

Gleichheit gilt genau dann, wenn  $v$  und  $w$  linear abhängig sind.

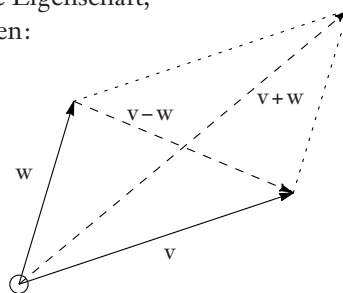
Genau wie in 6.2 ergibt sich nun die Dreiecksungleichung, und erneut gilt die allgemeinere Form  $\|v\| - \|w\| \leq \|v \pm w\| \leq \|v\| + \|w\|$ . Jedes Skalarprodukt induziert also eine Norm. Umgekehrt brauchen wir eine zusätzliche Eigenschaft, um aus einer Norm ein Skalarprodukt zu erzeugen:

### Definition (Parallelogramm-Gleichung)

Eine Norm auf  $V$  erfüllt die *Parallelogramm-Gleichung*, falls für alle  $v, w \in V$  gilt:

$$\|v + w\|^2 + \|v - w\|^2 = 2\|v\|^2 + 2\|w\|^2.$$

Damit können wir durchführen:



### Polarisation von Pascual Jordan und John von Neumann

Ist  $V$  ein normierter  $K$ -Vektorraum mit Parallelogramm-Gleichung, so wird ein Skalarprodukt auf  $V$  definiert durch die *Polarisations-Gleichungen*

$$4 \langle v, w \rangle = \begin{cases} \|v + w\|^2 - \|v - w\|^2 & \text{falls } K = \mathbb{R}, \\ \|v + w\|^2 - \|v - w\|^2 + i(\|iv + w\|^2 - \|iv - w\|^2) & \text{falls } K = \mathbb{C}. \end{cases}$$

Weiter gilt: Die Norm eines Skalarprodukt erfüllt die Parallelogramm-Gleichung, und die Polarisation rekonstruiert das Skalarprodukt aus der Norm. Kurz:

$$\text{Skalarprodukt} = \text{Norm} + \text{Parallelogramm-Gleichung}.$$

### Beispiele

- (1) Die *Summennorm* oder *Manhattan-Norm* auf dem  $K^n$  ist definiert durch

$$\|x\| = |x_1| + \dots + |x_n| \quad \text{für alle } x \in K^n.$$

- (2) Die *Maximumsnorm* auf dem  $K^n$  ist definiert durch

$$\|x\|_\infty = \max_{1 \leq i \leq n} |x_i| \quad \text{für alle } x \in K^n.$$

- (3) Für alle  $p \in [1, \infty]$  ist die *p-Norm* auf dem  $K^n$  definiert durch

$$\|x\|_p = (|x_1|^p + \dots + |x_n|^p)^{1/p} \quad \text{für alle } x \in K^n.$$

Die Summen- und die euklidische Norm sind die  $p$ -Normen für  $p = 1$  bzw.  $p = 2$ . Lediglich die 2-Norm erfüllt die Parallelogramm-Gleichung, sodass die  $p$ -Norm für  $p \neq 2$  von keinem Skalarprodukt abstammt.

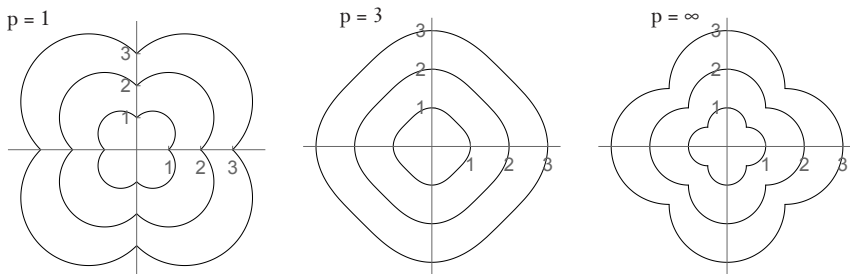


Illustration der Normen  $\|\cdot\|_p$  auf dem  $\mathbb{R}^2$  für  $p = 1, 3, \infty$ . Gezeigt sind alle  $\|v\|_p v \in \mathbb{R}^2$ , wobei  $v$  die Kreise der Radien  $r = 1, 2, 3$  der 2-Norm durchläuft (übliche Kreise). Für andere  $p$  ergeben sich ähnliche Bilder. Die Werte  $p = 1$  und  $p = \infty$  bilden die Extremfälle.

- (4) Sei  $V = \mathcal{C}([0, 1], K)$  der  $K$ -Vektorraum aller stetigen  $f: [0, 1] \rightarrow K$ . Dann sind die *Maximumsnorm* und für alle  $p \geq 1$  die *p-Norm* auf  $V$  definiert durch

$$\|f\|_\infty = \max_{x \in [0, 1]} |f(x)|, \quad \|f\|_p = \left( \int_0^1 |f(x)|^p dx \right)^{1/p}.$$

Diese Normen sind die kontinuierlichen Analoga der Normen auf dem  $K^n$ .

## 6.5 Normen im Endlich-Dimensionalen

### Satz (Äquivalenzsatz für Normen)

Sei  $V$  ein endlich-dimensionaler  $K$ -Vektorraum mit  $K = \mathbb{R}$  oder  $K = \mathbb{C}$ , und seien  $\|\cdot\|, \|\cdot\|' : V \rightarrow [0, \infty[$  Normen auf  $V$ . Dann sind die beiden Normen äquivalent, d. h., es gibt reelle Zahlen  $c, d > 0$  mit

$$(+)\quad c\|v\| \leq \|v\|', \quad d\|v\|' \leq \|v\| \quad \text{für alle } v \in V.$$

Zur Illustration betrachten wir die zu einer Norm auf  $V$  gehörige *Einheitskugel*

$$B = \{v \in V \mid \|v\| \leq 1\}.$$

(Die Namensgebung stammt von der euklidischen Norm auf dem  $\mathbb{R}^3$ . Im Allgemeinen ist  $B$  nur in einem abstrakten Sinn kugelförmig.) Der Äquivalenzsatz besagt, dass die Einheitskugeln  $B$  und  $B'$  zweier Normen  $\|\cdot\|$  bzw.  $\|\cdot\|'$  auf  $V$  nach einer geeigneten Skalierung ineinander Platz haben. Definieren wir für  $c > 0$  und  $A \subseteq V$  die Skalierung  $cA$  durch

$$cA = \{cv \mid v \in A\},$$

so sind äquivalent:

$$(a) \quad c\|v\| \leq \|v\|' \quad \text{für alle } v \in V.$$

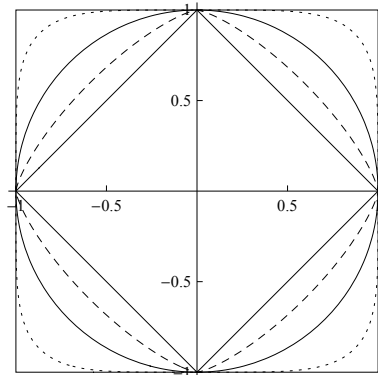
$$(b) \quad cB' \subseteq B.$$

Damit ist (+) äquivalent zu

$$cB' \subseteq B \quad \text{und} \quad dB \subseteq B'.$$

Äquivalent zu (+) ist auch, dass  $c, C > 0$  existieren mit

$$c\|v\| \leq \|v\|' \leq C\|v\| \quad \text{für alle } v \in V.$$



Die Einheitskugeln der  $p$ -Normen für  $p = 1, 4/3, 2, 5, \infty$  (von innen nach außen)

### Beispiele

- (1) Für die Normen  $\|\cdot\|_p$  für  $p = 1, 2, \infty$  auf dem  $K^n$  gelten die Abschätzungen:

$$\|x\|_\infty \leq \|x\|_2 \leq \|x\|_1 \leq \sqrt{n} \|x\|_2 \leq n \|x\|_\infty.$$

- (2) Für den unendlich-dimensionalen  $\mathbb{R}$ -Vektorraum  $V = \mathcal{C}([0, 1], \mathbb{R})$  ist der Satz nicht mehr richtig. Für alle  $n \geq 1$  sei  $f_n : [0, 1] \rightarrow \mathbb{R}$  mit

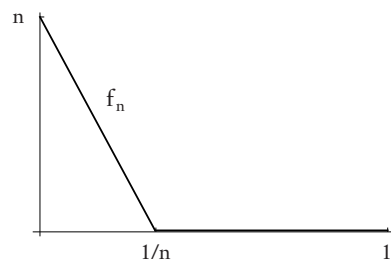
$$f_n(x) = \max(n - n^2 x, 0) \quad \text{für alle } x.$$

$$\text{Dann gilt } \|f_n\|_\infty = n, \|f_n\|_1 = 1/2$$

(Integral von  $|f|$ ) für alle  $n \geq 1$ .

Folglich gibt es kein  $d > 0$  mit

$$d\|f\|_\infty < \|f\|_1 \quad \text{für alle } f \in V.$$





Wir diskutieren zwei analytische Anwendungen des Satzes.

### Komponentenweise Konvergenz

Ist  $V$  ein normierter Vektorraum,  $(x_k)_{k \in \mathbb{N}}$  eine Folge in  $V$  und  $x \in V$ , so schreiben wir

$$\lim_{k \rightarrow \infty} x_k = x \quad (\text{bzgl. } \|\cdot\|),$$

falls  $\lim_{k \rightarrow \infty} \|x_k - x\| = 0$  in  $\mathbb{R}$  gilt. Wir sagen dann, dass die Folge  $(x_k)_{k \in \mathbb{N}}$  *unter der Norm von  $V$  gegen  $x$  konvergiert*. Mit Hilfe des Äquivalenzsatzes können wir den Konvergenzbegriff für endlich-dimensionale Vektorräume identifizieren und zeigen, dass er nicht von der Norm abhängt:

Ist  $V = K^n$ , so sind äquivalent:

$$(a) \quad x = \lim_{k \rightarrow \infty} x_k,$$

$$(b) \quad \lim_{k \rightarrow \infty} (x_k)_j = x_j \quad \text{für alle } 1 \leq j \leq n.$$

(komponentenweise Konvergenz in  $K = \mathbb{R}$  bzw.  $K = \mathbb{C}$ )

Zum Beweis verwenden wir, dass  $\|x - x_k\| \leq c \|x - x_k\|_\infty$  für ein geeignetes  $c$  gilt, und dass die Konvergenz bzgl. der Maximumsmetrik die komponentenweise Konvergenz ist. Allgemeiner gilt für Folgen in einem endlich-dimensionalen normierten Vektorraum  $V$  mit Basis  $\mathcal{A} = (v_1, \dots, v_n)$  und zugehöriger Koordinatenabbildung  $\Phi_{\mathcal{A}} : V \rightarrow K^n$ :

$$\lim_{k \rightarrow \infty} x_k = x \quad \text{genau dann, wenn} \quad \lim_{k \rightarrow \infty} \Phi_{\mathcal{A}}(x_k)_j = \Phi_{\mathcal{A}}(x)_j \quad \text{für alle } 1 \leq j \leq n.$$

Statt komponentenweiser Konvergenz spricht man deswegen auch von koordinatenweiser Konvergenz. Ist zum Beispiel  $V$  ein  $\mathbb{R}$ -Vektorraum mit einer Basis  $(v_1, v_2, v_3)$ , so konvergiert eine Folge in  $V$  unter jeder Norm genau dann gegen einen Vektor  $v = \alpha_1 v_1 + \alpha_2 v_2 + \alpha_3 v_3$ , wenn die drei reellen Koordinatenfolgen der Folge in  $\mathbb{R}$  gegen  $\alpha_1$ ,  $\alpha_2$  und  $\alpha_3$  konvergieren.

### Homomorphismen sind Lipschitz-stetig

Seien  $V = K^n$ ,  $W = K^m$  normiert durch  $\|\cdot\|_V$  bzw.  $\|\cdot\|_W$ , und sei  $f : V \rightarrow W$  linear. Dann gilt für alle  $x, y \in V$  aufgrund der Linearität von  $f$  und der Dreiecksungleichung:

$$\|f(x) - f(y)\|_W = \|(x_1 - y_1)f(e_1) + \dots + (x_n - y_n)f(e_n)\|_W \leq$$

$$|x_1 - y_1| \|f(e_1)\|_W + \dots + |x_n - y_n| \|f(e_n)\|_W \leq s \|x - y\|_\infty \leq s c \|x - y\|_V,$$

wobei wir

$$s = \|f(e_1)\|_W + \dots + \|f(e_n)\|_W$$

setzen und für die Konstante  $c$  den Äquivalenzsatz bemühen. Damit ist  $f$  Lipschitz-stetig mit der Lipschitz-Konstanten  $L = s c$ . Allgemeiner zeigt man in dieser Weise, dass jeder Homomorphismus  $f : V \rightarrow W$  zwischen normierten endlich-dimensionalen Vektorräumen  $V$  und  $W$  Lipschitz-stetig ist. Die Lipschitz-Konstante hängt dabei von den gewählten Normen ab.

## 6.6 Orthonormalbasen

### Definition (Orthogonalbasis, Orthonormalbasis)

Sei  $V$  ein euklidischer oder unitärer Vektorraum. Eine Basis  $(v_i)_{i \in I}$  von  $V$  heißt eine *Orthogonalbasis*, falls  $\langle v_i, v_j \rangle = 0$  für alle  $i, j \in I$  mit  $i \neq j$ . Gilt zusätzlich  $\|v_i\| = 1$  für alle  $i$ , so heißt  $(v_i)_{i \in I}$  eine *Orthonormalbasis*.

Eine Orthonormalbasis ist also eine Basis aus normierten Vektoren, die paarweise aufeinander senkrecht stehen. Kompakt kann man dies durch

$$\langle v_i, v_j \rangle = \delta_{ij} \quad \text{für alle } i, j \in I \quad (\text{Orthonormalitätsbedingung})$$

zum Ausdruck bringen. Da die Orthogonalität die lineare Unabhängigkeit nach sich zieht, ist eine orthogonale Familie  $(v_i)_{i \in I}$  in  $V - \{0\}$  bereits dann eine Orthogonalbasis, wenn sie erzeugend ist. Weiter ist dann  $(N(v_i))_{i \in I}$  eine Orthonormalbasis. Jede Orthogonalbasis lässt sich also durch Normierung in eine Orthonormalbasis überführen. Mit der Konstruktion von Orthogonalbasen werden wir uns im nächsten Abschnitt beschäftigen. Zunächst wollen wir wichtige Eigenschaften festhalten und Beispiele kennenlernen.

Ist  $(v_1, \dots, v_n)$  eine Orthonormalbasis von  $V$ , so gilt für alle  $v \in V$ :

$$v = \langle v_1, v \rangle v_1 + \dots + \langle v_n, v \rangle v_n, \quad (\text{Koordinatenbestimmung})$$

$$\|v\|^2 = |\langle v_1, v \rangle|^2 + \dots + |\langle v_n, v \rangle|^2. \quad (\text{Parseval-Gleichung})$$

Ist  $\Phi: V \rightarrow K^n$  die Koordinatenabbildung bzgl.  $(v_1, \dots, v_n)$ , so gilt

$$\langle v, w \rangle = \langle \Phi(v), \Phi(w) \rangle_{\text{kanonisch}} \quad \text{für alle } v, w \in V.$$

In diesem Sinn ist  $V$  isomorph zum  $K^n$  mit dem kanonischen Skalarprodukt.

Ist  $(v_i)_{i \in I}$  eine Orthonormalbasis von  $V$ , so ist für alle  $v \in V$  die Menge aller Indizes  $i$  mit  $\langle v_i, v \rangle \neq 0$  endlich, und es gilt:

$$v = \sum_{i \in I} \langle v_i, v \rangle v_i, \quad (\text{Koordinatenbestimmung})$$

$$\|v\|^2 = \sum_{i \in I} |\langle v_i, v \rangle|^2. \quad (\text{Parseval-Gleichung})$$

Die Aussagen ergeben sich für  $K = \mathbb{C}$  und  $v = \sum_i \alpha_i v_i$ ,  $w = \sum_i \beta_i v_i$  aus

$$\langle v_i, v \rangle = \langle v_i, \sum_j \alpha_j v_j \rangle = \sum_j \alpha_j \langle v_i, v_j \rangle = \sum_j \alpha_j \delta_{ij} = \alpha_i \quad \text{für alle } i \in I,$$

$$\langle v, v \rangle = \langle \sum_i \alpha_i v_i, v \rangle = \sum_i \overline{\alpha_i} \langle v_i, v \rangle = \sum_i |\alpha_i|^2 = \sum_i |\langle v_i, v \rangle|^2,$$

$$\langle v, w \rangle = \sum_i \overline{\alpha_i} \langle v_i, w \rangle = \sum_i \overline{\alpha_i} \beta_i = \langle \Phi(v), \Phi(w) \rangle_{\text{kanonisch}} \quad \text{für endliche } I.$$

Für jedes  $i$  pickt  $\langle v_i, \cdot \rangle: V \rightarrow K$  die  $i$ -Koordinate von  $v$  bzgl.  $(v_i)_{i \in I}$  heraus. Es gilt also  $\langle v_i, \cdot \rangle = v_i^*$  mit den linear unabhängigen dualen Vektoren  $v_i^* \in V^*$  (vgl. 3.12).

## Beispiele

- (1) Die Standardbasis  $(e_1, \dots, e_n)$  des  $\mathbb{C}^n$  ist eine Orthonormalbasis bzgl. des kanonischen Skalarprodukts. Die Parseval-Gleichung schreibt sich als
- $$\|z\|^2 = |\langle e_1, z \rangle|^2 + \dots + |\langle e_n, z \rangle|^2 = |z_1|^2 + \dots + |z_n|^2 \quad \text{für alle } z \in \mathbb{C}^n.$$
- (2) Ist  $(v_1, v_2)$  eine Orthonormalbasis des  $\mathbb{R}^2$  bzgl. des kanonischen Skalarprodukts, so gibt es ein  $\alpha \in [0, 2\pi[$  mit  $v_1 = (\cos \alpha, \sin \alpha)$  (Polarkoordinaten). Dann gilt
- $$v_2 = (-\sin \alpha, \cos \alpha) \quad \text{oder} \quad v_2 = (\sin \alpha, -\cos \alpha).$$
- (3) Die Orthonormalbasen des  $\mathbb{R}^3$  bzgl. des kanonischen Skalarprodukts lassen sich als normierte rechtwinklige Dreibeine mit Spitze am Nullpunkt beschreiben.
- (4) Sei  $n \in \mathbb{N}$ . Wir betrachten den  $(2n + 1)$ -dimensionalen unitären Vektorraum der *trigonometrischen Polynome* von Grad kleinergleich  $n$ :

$$V = \{ f: \mathbb{R} \rightarrow \mathbb{C} \mid \text{es gibt } a_{-n}, \dots, a_n \in \mathbb{C} \text{ mit } f(x) = \sum_{-n \leq k \leq n} a_k e^{ikx} \text{ für alle } x \},$$

$$\langle f, g \rangle = \frac{1}{2\pi} \int_0^{2\pi} \overline{f(x)} g(x) dx \quad \text{für alle } f, g \in V.$$

Die (als Terme notierten) Funktionen  $e^{ikx}$ ,  $-n \leq k \leq n$  bilden eine Orthonormalbasis von  $V$ . Für alle  $f \in V$  gilt

$$a_k = \frac{1}{2\pi} \int_0^{2\pi} e^{-ikx} f(x) dx \quad \text{für alle } -n \leq k \leq n, \quad (\text{Koeffizientenberechnung})$$

$$\frac{1}{2\pi} \int_0^{2\pi} |f(x)|^2 dx = \sum_{-n \leq k \leq n} |a_k|^2. \quad (\text{Parseval-Gleichung})$$

- (5) Im  $\mathbb{R}$ -Vektorraum aller reellen Polynomfunktionen

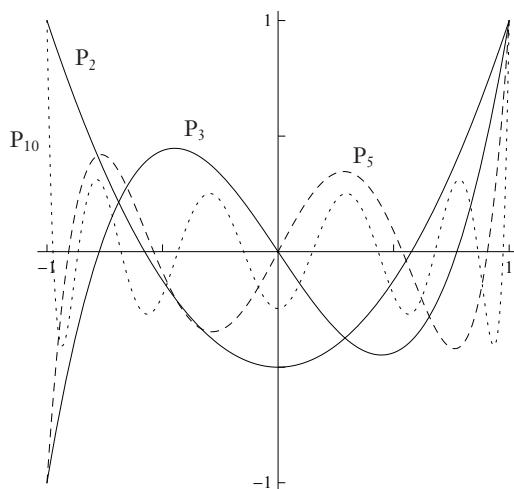
$$V = \{ f: \mathbb{R} \rightarrow \mathbb{R} \mid \text{es gibt } a_0, \dots, a_n \in \mathbb{R} \text{ mit } f(x) = \sum_{k \leq n} a_k x^k \text{ für alle } x \} \text{ mit}$$

$$\langle f, g \rangle = \int_{-1}^1 f(x) g(x) dx$$

definieren wir die *Legendre-Polynome*  $P_n$  rekursiv durch

$$\begin{aligned} P_0(x) &= 1, \quad P_1(x) = x, \\ (n+1) P_{n+1}(x) &= \\ (2n+1)x P_n(x) - n P_{n-1}(x). \end{aligned}$$

Man kann zeigen, dass  $\langle P_n, P_m \rangle = 2/(2n+1) \delta_{nm}$ , sodass die  $P_n$  eine Orthogonalbasis von  $V$  bilden. Sie sind in der Physik bedeutsam. Eine mathematische Motivation werden wir in 6.7 kennenlernen.



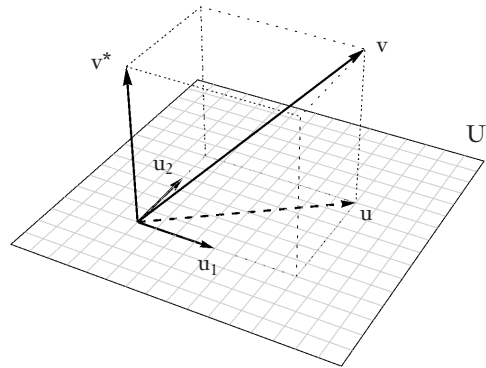
## 6.7 Das Orthonormalisierungsverfahren

### Satz (Existenz von Orthonormalbasen)

Sei  $V$  ein euklidischer oder unitärer Vektorraum, der eine abzählbare Basis besitzt. Dann besitzt  $V$  eine Orthonormalbasis.

Der Satz ist das „orthogonale Analogon“ zum Basisexistenzsatz (3. 9). Wir werden unten sehen, dass wir diesmal auf eine Dimensionsvoraussetzung nicht verzichten können.

Zum Beweis betrachten wir einen endlich-dimensionalen Unterraum  $U$  von  $V$ . Wir nehmen an, dass  $U$  eine Orthonormalbasis  $(u_1, \dots, u_k)$  besitzt. Nun sei  $v \in V - U$  beliebig. Dann steht der Vektor



$$u = \langle u_1, v \rangle u_1 + \langle u_2, v \rangle u_2$$

$$v^* = v - u \quad \text{mit} \quad u = \sum_{1 \leq i \leq k} \langle u_i, v \rangle u_i \in U$$

senkrecht auf allen  $u_j$  (vgl. das Diagramm), da

$$\langle u_j, v^* \rangle = \langle u_j, v \rangle - \sum_{1 \leq i \leq k} \langle u_i, v \rangle \langle u_j, u_i \rangle = \langle u_j, v \rangle - \langle u_j, v \rangle = 0 \quad \text{für alle } 1 \leq j \leq k.$$

Wegen  $v \notin U$  ist  $v^* \neq 0$ , und damit ist

$$(u_1, \dots, u_k, N(v^*)) \quad (\text{orthonormale Erweiterung})$$

eine Orthonormalbasis des Unterraums  $\text{span}(u_1, \dots, u_k, v^*) = \text{span}(U \cup \{v\})$ . Die Argumentation liefert folgendes Verfahren zur Konstruktion einer Orthonormalbasis:

### Das Orthonormalisierungsverfahren von Gram-Schmidt

Sei  $(v_1, \dots, v_n)$  eine Basis von  $V$ . Dann definieren wir rekursiv:

$$u_1 = N(v_1),$$

$$u_{k+1} = N(v_{k+1} - \sum_{1 \leq i \leq k} \langle u_i, v_{k+1} \rangle u_i) \quad \text{für alle } 1 \leq k \leq n-1.$$

Dann ist  $(u_1, \dots, u_n)$  eine Orthonormalbasis von  $V$ . Zudem gilt

$$\text{span}(u_1, \dots, u_k) = \text{span}(v_1, \dots, v_k) \quad \text{für alle } k \leq n.$$

Die Orthonormalbasis  $(u_1, \dots, u_n)$  heißt die *Gram-Schmidt-Orthonormalisierung* von  $(v_1, \dots, v_n)$ . Das Verfahren kann analog für eine abzählbar unendliche Basis  $(v_n)_{n \in \mathbb{N}}$  von  $V$  durchgeführt werden und liefert dann eine Orthonormalbasis  $(u_n)_{n \in \mathbb{N}}$  von  $V$ .

Als Korollar erhält man die sog. QR-Zerlegung einer invertierbaren Matrix, die wir in Überblick 7 diskutieren. Die Summen  $\sum_{1 \leq i \leq k} \langle u_i, v_{k+1} \rangle u_i$  werden wir im nächsten Abschnitt genauer betrachten.

**Beispiele**

- (1) Wir betrachten den  $\mathbb{R}^3$  mit dem kanonischen Skalarprodukt. Das Verfahren von Gram-Schmidt liefert für die Basis  $(v_1, v_2, v_3) = ((1, 0, 0), (1, 1, 0), (1, 1, 1))$ :

$$u_1 = v_1 = (1, 0, 0),$$

$$u_2 = N(v_2 - \langle u_1, v_2 \rangle u_1) = N(v_2 - u_1) = (0, 1, 0),$$

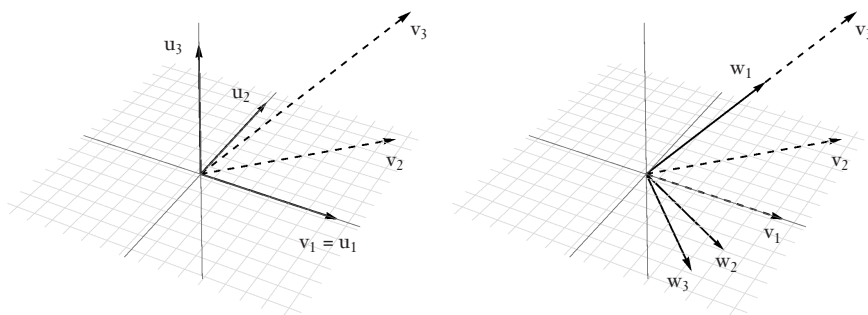
$$u_3 = N(v_3 - \langle u_1, v_3 \rangle u_1 - \langle u_2, v_3 \rangle u_2) = N(v_3 - u_1 - u_2) = (0, 0, 1).$$

Die Orthonormalisierung ergibt also die kanonische Basis des  $\mathbb{R}^3$ . Wenden wir dagegen das Verfahren auf die umgeordnete Basis  $(v_3, v_2, v_1)$  an, so erhalten wir

$$w_1 = N(v_3) = \alpha(1, 1, 1) \quad \text{mit } \alpha = 1/\sqrt{3},$$

$$w_2 = N(v_2 - \langle w_1, v_2 \rangle w_1) = \beta(1, 1, -2) \quad \text{mit } \beta = 1/\sqrt{6},$$

$$w_3 = N(v_1 - \langle w_1, v_1 \rangle w_1 - \langle w_2, v_1 \rangle w_2) = \gamma(1, -1, 0) \quad \text{mit } \gamma = 1/\sqrt{2}.$$



- (2) Wir betrachten den  $\mathbb{R}$ -Vektorraum  $V$  aller reellen Polynomfunktionen mit

$$\langle f, g \rangle = \int_{-1}^1 f(x) g(x) dx \quad \text{für alle } f, g \in V$$

und die abzählbar unendliche Basis  $(1, x, x^2, x^3, \dots)$ . Das Orthonormalisierungsverfahren liefert die normierten Legendre-Polynome  $N(P_0), N(P_1), N(P_2), \dots$  (vgl. 8.7).

**Exkurs: Ein euklidischer Vektorraum ohne Orthonormalbasis**

Sei  $\ell^2$  der euklidische Vektorraum aller quadratsummierbaren Folgen in  $\mathbb{R}$  (vgl. 6.3).

*Annahme*,  $\ell^2$  besitzt eine Orthonormalbasis  $(v_i)_{i \in I}$ . Für alle  $n$  ist dann  $e_n \in \ell^2$  eine Linearkombination von Vektoren der Basis. Insgesamt werden zur Darstellung aller  $e_n$  nur abzählbar viele  $v_i$  verwendet. Da  $I$  überabzählbar ist (vgl. 4.9), gibt es ein nicht verwendetes  $v_{i^*}$ . Dann gilt  $\langle v_{i^*}, e_n \rangle = 0$  für alle  $n$ , sodass  $v_{i^*} = 0$ , *Widerspruch*.

Woran scheitert ein allgemeiner Existenzbeweis? Das Zornsche Lemma liefert eine maximale orthonormale Familie  $(u_i)_{i \in I}$ , aber nun ist  $U = \text{span}(\{u_i \mid i \in I\}) \neq V$  möglich. Denn die Bildung  $v^* = v - \sum_{i \in I} \langle u_i, v \rangle u_i$  ist für ein  $v \in V - U$  im Allgemeinen nicht mehr möglich, da es unendlich viele  $i \in I$  mit  $\langle u_i, v \rangle \neq 0$  geben kann. Ein Beispiel in  $\ell^2$  ist  $u_n = e_n$  für  $n$  und  $v$  mit  $v(k) = 1$  für alle  $k$ .

## 6.8 Orthogonale Komplemente und Projektionen

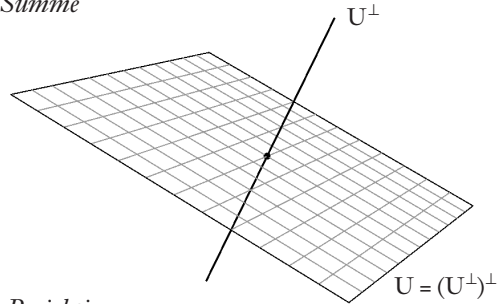
### Definition (orthogonales Komplement, orthogonale Projektion)

Sei  $V$  ein euklidischer oder unitärer Vektorraum.

#### Orthogonale Unterräume und orthogonale Summe

Zwei Unterräume  $U$  und  $W$  von  $V$  heißen *orthogonal*, falls  $\langle u, w \rangle = 0$  für alle  $u \in U$  und  $w \in W$ .

$V$  heißt *orthogonale Summe* einer Familie von Unterräumen  $(U_i)_{i \in I}$ , falls  $V = \sum_{i \in I} U_i$  und die  $U_i$  paarweise orthogonal sind.



#### Orthogonales Komplement und orthogonale Projektion

Ist  $U$  ein Unterraum von  $V$ , so heißt

$$U^\perp = \{ v \in V \mid \langle v, u \rangle = 0 \text{ für alle } u \in U \}$$

das *orthogonale Komplement* von  $U$  in  $V$ . Die Abbildung  $P_U : U + U^\perp \rightarrow U$  mit

$$P_U(v) = \text{„das eindeutige } u \in U \text{ mit } v - u \in U^\perp\text{“ für alle } v \in U + U^\perp$$

heißt die *orthogonale Projektion* von  $U + U^\perp$  auf  $U$ .

Orthogonale Komplemente in  $\mathbb{R}^3$

Mit Hilfe orthogonaler Unterräume lässt sich ein euklidischer oder unitärer Vektorraum übersichtlich organisieren. Sind  $U$  und  $W$  orthogonal, so ist  $U \cap W = \{0\}$ . Denn für alle Vektoren  $u \in U \cap W$  gilt  $0 = \langle u, u \rangle$  und damit  $u = 0$ . Weiter gilt:

#### Orthogonale Summen sind direkt

Ist  $V = \sum_{i \in I} U_i$  eine orthogonale Summe und sind  $u_i \in U_i$  mit  $\sum_{i \in I} u_i = 0$ , so gilt

$$0 = \langle \sum_{i \in I} u_i, \sum_{i \in I} u_i \rangle = \sum_{i, j \in I} \langle u_i, u_j \rangle = \sum_{i \in I} \langle u_i, u_i \rangle = \sum_{i \in I} \|u_i\|^2,$$

sodass  $u_i = 0$  für alle  $i \in I$ . Damit ist  $V = \oplus_{i \in I} U_i$  (vgl. 3. 10).

Insbesondere ist die orthogonale Summe  $U + U^\perp$  direkt, sodass die orthogonale Projektion  $P_U : U + U^\perp \rightarrow U$  wohldefiniert ist. Wichtig ist:

Ist  $U$  endlich-dimensional, so ist  $U + U^\perp = V$ .

Zum Beweis seien  $(u_1, \dots, u_k)$  eine Orthonormalbasis von  $U$ ,  $v \in V$  beliebig und

$$v^* = v - u \text{ mit } u = \sum_{1 \leq i \leq k} \langle u_i, v \rangle u_i \in U.$$

Der Vektor  $v^*$  steht senkrecht auf allen  $u_i$ , sodass  $v = u + v^* \in U + U^\perp$ .

Ist  $V$  endlich-dimensional, so gilt also  $\dim(U) + \dim(U^\perp) = \dim(V)$ . Weiter ist dann  $(U^\perp)^\perp = U$ . Allgemein gilt nur  $U \subseteq (U^\perp)^\perp$ , vgl. das folgende Beispiel (2).

Die orthogonale Projektion  $P_U : U + U^\perp \rightarrow U$  ist linear und surjektiv. Weiter gilt  $P_U|_U = \text{id}_U$  und  $P_U \circ P_U = P_U$  (Idempotenz). Wichtig sind darüber hinaus:

Ist $(u_i)_{i \in I}$ eine Orthonormalbasis von $U$ , so gilt für alle $v \in U + U^\perp$ :	
$P_U(v) = \sum_{i \in I} \langle u_i, v \rangle u_i.$	(Berechnungsformel)
Für alle $v \in U + U^\perp$ gilt $\ v - P_U(v)\  = \min_{u \in U} \ v - u\ .$	
(Bestapproximation)	

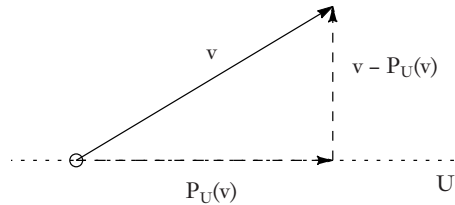
Die Rekursionsformel des Gram-Schmidt-Verfahrens können wir nun schreiben als

$$u_{k+1} = N(v_{k+1} - P_{U_k}(v_{k+1})), \text{ mit}$$

$$U_k = \text{span}(v_1, \dots, v_k) = \text{span}(u_1, \dots, u_k).$$

In Kurzform lautet das Verfahren also:

*Projiziere und normalisiere die Differenz.*



## Beispiele

- (1) Im  $\mathbb{C}$ -Vektorraum  $V$  aller stetigen Funktionen von  $[0, 2\pi]$  nach  $\mathbb{C}$  mit

$$\langle f, g \rangle = \frac{1}{2\pi} \int_0^{2\pi} \overline{f(x)} g(x) dx \quad \text{für alle } f, g \in V$$

erzeugen die orthonormalen Vektoren  $e^{ikx}$ ,  $-n \leq k \leq n$  für alle  $n \geq 1$  einen Unterraum  $U_n$ . Für alle  $f \in V$  ist  $P_{U_n}(f)$  die  $n$ -te *Fourier-Approximation* an  $f$ :

$$P_U(f) = \sum_{-n \leq k \leq n} \langle e^{ikx}, f \rangle e^{ikx} = \sum_{-n \leq k \leq n} c_k e^{ikx}, \text{ mit}$$

$$c_k = \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-ikx} dx \quad \text{für alle } -n \leq k \leq n.$$

- (2) Im euklidischen Vektorraum  $V = \mathbb{R}^{(\mathbb{N})}$  aller reellen Folgen mit endlichem Träger sei

$$U = \{ a_1 e_1 + \dots + a_n e_n \mid n \geq 1, a_1 + \dots + a_n = 0 \}$$

der Unterraum aller Folgen, deren Folgenglieder sich zu 0 aufsummieren. Ist nun  $v \in U^\perp$ , so gilt wegen  $e_i - e_j \in U$  für  $i \neq j$ , dass

$$v(i) - v(j) = \langle v, e_i - e_j \rangle = 0 \quad \text{für alle } i \neq j.$$

Also ist  $v$  konstant damit gleich 0. Folglich ist  $U^\perp = \{0\}$  und  $(U^\perp)^\perp = V \neq U$ .

Da  $U$  und  $V$  eine abzählbar unendliche Dimension besitzen, existieren Orthonormalbasen der beiden Räume. Eine Orthonormalbasis von  $U$  lässt sich aber wegen  $U^\perp = \{0\}$  nicht zu einer Orthonormalbasis von  $V$  fortsetzen. Das orthogonale Analogon des Basisergänzungssatzes ist also nicht mehr gültig.

## 6.9 Orthogonale Homomorphismen und Matrizen

**Definition** (*orthogonaler Homomorphismus, orthogonale Matrix, unitäre Matrix*)

*Orthogonale Homomorphismen*

Seien  $V, W$  euklidische bzw. unitäre Vektorräume. Eine lineare Abbildung  $f : V \rightarrow W$  heißt ein *orthogonaler Homomorphismus*, falls

$$(+)\quad \langle f(v), f(w) \rangle_W = \langle v, w \rangle_V \quad \text{für alle } v, w \in V.$$

*Orthogonale und unitäre Matrizen*

Eine Matrix  $Q \in \mathbb{R}^{n \times n}$  heißt *orthogonal*, falls  $\langle Qx, Qy \rangle = \langle x, y \rangle$  für alle  $x, y \in \mathbb{R}^n$ .

Analog heißt eine Matrix  $U \in \mathbb{C}^{n \times n}$  *unitär*, falls  $\langle Ux, Uy \rangle = \langle x, y \rangle$  für alle  $x, y \in \mathbb{C}^n$ .

Dabei werden die kanonischen Skalarprodukte des  $\mathbb{R}^n$  bzw.  $\mathbb{C}^n$  verwendet.

Wie für alle algebraischen Strukturen sind strukturerhaltende Abbildungen von Interesse. Eine Abbildung  $f : V \rightarrow W$  zwischen Vektorräumen mit Skalarprodukt erhält die Struktur, wenn  $f$  linear ist und das Skalarprodukt von  $V$  im Sinne von (+) respektiert.

Orthogonale Homomorphismen sind injektiv, da  $f(v) = 0$  impliziert, dass  $\langle f(v), f(v) \rangle = 0$  und damit  $\langle v, v \rangle = 0$ . Nach positiver Definitheit ist also  $\text{Kern}(f) = \{0\}$  und somit  $f$  injektiv.

### Orthogonalität und Längentreue sind äquivalent

Ist  $f : V \rightarrow W$  orthogonal, so gilt  $\|f(v)\|^2 = \langle f(v), f(v) \rangle = \langle v, v \rangle = \|v\|^2$  und damit

$$\|f(v)\| = \|v\| \quad \text{für alle } v \in V. \quad (\text{Längentreue})$$

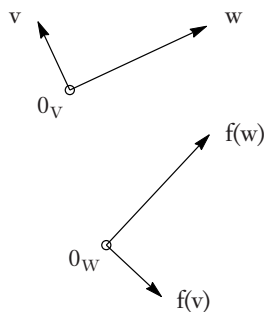
Ist umgekehrt  $f : V \rightarrow W$  linear und längentreu, so gilt im Fall  $K = \mathbb{R}$  nach der Polarisationsformel

$$4\langle f(x), f(y) \rangle = \|f(x) + f(y)\|^2 - \|f(x) - f(y)\|^2 =$$

$$\|f(x + y)\|^2 - \|f(x - y)\|^2 =$$

$$\|x + y\|^2 - \|x - y\|^2 = 4\langle x, y \rangle.$$

Also ist  $f$  orthogonal. Analoges gilt für  $K = \mathbb{C}$ .



Die Brücke zu den Matrizen ist gegeben durch:

- (1) Ist  $V$  ein endlich-dimensionaler euklidischer bzw. unitärer Vektorraum, so ist  $V$  orthogonal isomorph zum  $K^n$  mit dem kanonischen Skalarprodukt, d. h., es existiert ein Isomorphismus  $f : V \rightarrow K^n$  mit (+). Ist  $(v_1, \dots, v_n)$  eine Orthonormalbasis von  $V$ , so ist das lineare  $f : V \rightarrow K^n$  mit  $f(v_i) = e_i$  für alle  $i$  ein solcher Isomorphismus.
- (2) Eine Matrix  $Q \in \mathbb{R}^{n \times n}$  ist genau dann orthogonal, wenn  $f_Q : \mathbb{R}^n \rightarrow \mathbb{R}^n$  orthogonal ist. Analoges gilt für eine unitäre Matrix  $U \in \mathbb{C}^{n \times n}$ .

Dass das Matrix-Vektor-Produkt das kanonische Skalarprodukt nicht verändert, lässt sich durch eine Reihe von äquivalenten Bedingungen zum Ausdruck bringen:



<i>Charakterisierungen der Orthogonalität von <math>Q \in \mathbb{R}^{n \times n}</math></i>
$\ Qx\  = \ x\ $ für alle $x \in \mathbb{R}^n$ ( <i>Längentreue</i> )
Die Spalten von $Q$ bilden eine Orthonormalbasis.
$Q^{-1} = Q^t$ ( <i>Invertierung durch Transposition</i> )
Die Zeilen von $Q$ bilden eine Orthonormalbasis.

$$\begin{pmatrix} a_{11} & \dots & \dots & a_{1n} \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ a_{n1} & \dots & \dots & a_{nn} \end{pmatrix} \cdot \begin{pmatrix} a_{11} & \dots & \dots & a_{n1} \\ \dots & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ a_{1n} & \dots & \dots & a_{nn} \end{pmatrix} = E_n$$

zur Eigenschaft  $Q^{-1} = Q^t$

Orthogonale Matrizen sind auch winkeltreu, da sie das Skalarprodukt erhalten. Aus der Winkeltreue folgt im Allgemeinen aber nicht die Orthogonalität, wie etwa eine Streckung  $\lambda E_n$  zeigt. Typische Argumente, die die Verwendung und das Wechselspiel der Bedingungen illustrieren, sind:

(1) Ist  $Q$  orthogonal, so gilt  $\langle Qe_i, Qe_j \rangle = \langle e_i, e_j \rangle = \delta_{ij}$  für die Spalten  $Qe_1, \dots, Qe_n$  von  $Q$ , sodass die Spalten von  $Q$  eine Orthonormalbasis des  $\mathbb{R}^n$  bilden.

(2) Bilden die Zeilen  $q_1, \dots, q_n$  von  $Q$  eine Orthonormalbasis, so ist  $Q$  orthogonal, da

$$\begin{aligned} \langle Qx, Qy \rangle &= \langle (\langle q_1, x \rangle, \dots, \langle q_n, x \rangle), (\langle q_1, y \rangle, \dots, \langle q_n, y \rangle) \rangle = \\ &\langle q_1, x \rangle \langle q_1, y \rangle + \dots + \langle q_n, x \rangle \langle q_n, y \rangle = \langle \langle q_1, x \rangle q_1 + \dots + \langle q_n, x \rangle q_n, y \rangle = \langle x, y \rangle. \end{aligned}$$

Für unitäre Matrizen gelten analoge Charakterisierungen, wobei wir hinsichtlich der Inversenbildung alle Einträge der Matrix bei der Transponierung zusätzlich zu konjugieren haben. Definieren wir also für eine beliebige Matrix  $A \in \mathbb{C}^{n \times n}$  die *adjungierte Matrix*  $A^* \in \mathbb{C}^{n \times n}$  durch  $A^*(i, j) = \overline{a_{ji}}$ , so gilt für unitäre Matrizen also

$$U^{-1} = U^*. \quad (\text{Invertierungsregel für unitäre Matrizen})$$

Die orthogonalen bzw. unitären Matrizen bilden die Untergruppen  $O(n)$  von  $GL(n, \mathbb{R})$  bzw.  $U(n)$  von  $GL(n, \mathbb{C})$ . Wir werden sie später noch genauer untersuchen.

## Beispiele

- (1) Wir betrachten die Orthonormalbasis  $(w_1, w_2, w_3)$  des  $\mathbb{R}^3$  mit  $w_1 = \alpha(1, 1, 1)$ ,  $w_2 = \beta(1, 1, -2)$ ,  $w_3 = \gamma(1, -1, 0)$ , wobei  $\alpha = 1/\sqrt{3}$ ,  $\beta = 1/\sqrt{6}$ ,  $\gamma = 1/\sqrt{2}$  (vgl. 6.7). Ist  $Q$  die Matrix mit den Spalten  $w_1, w_2, w_3$ , so ist  $Q$  orthogonal und

$$Q Q^t = \begin{pmatrix} \alpha & \beta & \gamma \\ \alpha & \beta & -\gamma \\ \alpha & -2\beta & 0 \end{pmatrix} \begin{pmatrix} \alpha & \alpha & \alpha \\ \beta & \beta & -2\beta \\ \gamma & -\gamma & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = E_3.$$

- (2) Die Abbildung  $f: \ell_{\mathbb{R}}^2 \rightarrow \ell_{\mathbb{R}}^2$  mit  $f(x_0, x_1, x_2, \dots) = (0, x_0, x_1, \dots)$  ist orthogonal, aber nicht surjektiv.

## 6.10 Der Rieszsche Darstellungssatz

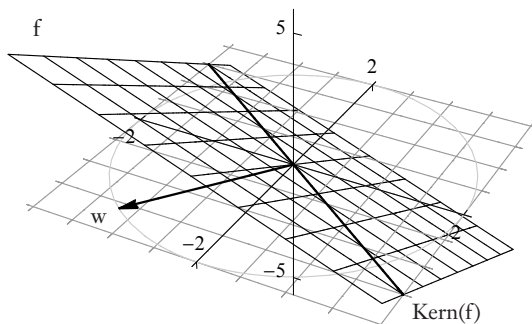
### Satz (Rieszscher Darstellungssatz)

Sei  $V$  ein endlich-dimensionaler euklidischer oder unitärer Vektorraum, und sei  $f \in V^*$ .

Dann gibt es ein eindeutiges  $w \in V$  mit  $f = \langle w, \cdot \rangle$ , d.h.

$$f(v) = \langle w, v \rangle \text{ für alle } v \in V.$$

Für alle  $w \in V$  ist  $\langle w, \cdot \rangle \in V^*$ . Der Satz besagt, dass umgekehrt jedes  $f \in V^*$  von der eindeutigen Form  $\langle w, \cdot \rangle$  ist, falls  $V$  endlich-dimensional ist. Wir nennen  $w$  den *darstellenden* oder *Riesz-Vektor* von  $f$ .



$$f(x, y) = \frac{-6x - 8y}{5}$$

$$w = (-6/5, -8/5), \quad \|w\| = 2$$

### Konstruktion des darstellenden Vektors

Wir nehmen  $K = \mathbb{C}$  an und betrachten eine Orthonormalbasis  $(v_1, \dots, v_n)$  von  $V$  und die Dualbasis  $(v_1^*, \dots, v_n^*)$  von  $V^*$ . Dann gibt es eindeutige  $\alpha_1, \dots, \alpha_n \in \mathbb{C}$  mit

$$f = \alpha_1 v_1^* + \dots + \alpha_n v_n^*.$$

Es gilt  $(\alpha_1, \dots, \alpha_n) = \Phi_{(v_1^*, \dots, v_n^*)}(f)$  mit der Koordinatenabbildung  $\Phi_{(v_1^*, \dots, v_n^*)} : V^* \rightarrow \mathbb{C}^n$ . Weiter ist  $\alpha_1 = f(v_1), \dots, \alpha_n = f(v_n)$ , sodass die  $\alpha_i$  durch Auswerten von  $f$  auf den Basisvektoren berechnet werden können. Für alle  $v = \lambda_1 v_1 + \dots + \lambda_n v_n \in V$  gilt

$$f(v) = f(\lambda_1 v_1 + \dots + \lambda_n v_n) = (\alpha_1 v_1^* + \dots + \alpha_n v_n^*)(\lambda_1 v_1 + \dots + \lambda_n v_n) =$$

$$\alpha_1 \lambda_1 + \dots + \alpha_n \lambda_n = \langle \overline{\alpha_1} v_1 + \dots + \overline{\alpha_n} v_n, v \rangle,$$

wobei wir im letzten Schritt die Orthonormalität der Basis verwenden. Damit ist

$$w = \overline{\alpha_1} v_1 + \dots + \overline{\alpha_n} v_n.$$

(Identifikation des darstellenden Vektors)

Zur Eindeutigkeit beobachten wir, dass für alle  $w \neq u$  die Abbildungen  $\langle w, \cdot \rangle$  und  $\langle u, \cdot \rangle$  verschieden sind. Denn ist  $\langle w, \cdot \rangle = \langle u, \cdot \rangle$ , so ist  $\langle w - u, \cdot \rangle$  die Nullabbildung, sodass insbesondere  $\langle w - u, w - u \rangle = 0$  und damit  $w = u$  nach positiver Definitheit. Für  $K = \mathbb{R}$  bleibt die Argumentation gleich, wobei die Konjugationen wegfallen.

### Beispiele

- (1) Sei  $V = \mathbb{R}^n$  mit dem kanonischen Skalarprodukt und der Standardbasis  $(e_1, \dots, e_n)$ , und sei  $f : \mathbb{R}^n \rightarrow \mathbb{R}$  linear. Dann gilt für alle  $v = (\lambda_1, \dots, \lambda_n) \in \mathbb{R}^n$ :

$$f(v) = f(\lambda_1 e_1 + \dots + \lambda_n e_n) = \lambda_1 f(e_1) + \dots + \lambda_n f(e_n) =$$

$$f(e_1) \lambda_1 + \dots + f(e_n) \lambda_n = \langle (f(e_1), \dots, f(e_n)), v \rangle,$$

sodass  $w = (f(e_1), \dots, f(e_n)) \in \mathbb{R}^n$  der darstellende Vektor von  $f$  ist.

- (2) Für den  $\mathbb{R}^2$  mit dem kanonischen Skalarprodukt ist ein lineares  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$  eine Ebene durch den Nullpunkt. Der Riesz-Vektor  $w \in \mathbb{R}^2$  ist

$$w = (f(1, 0), f(0, 1)).$$

Dieser Vektor steht senkrecht auf  $\text{Kern}(f)$  (dem Schnitt von  $f$  mit der  $x$ - $y$ -Ebene) und zeigt in die Richtung des stärksten Anstiegs der Ebene. In der Sprache der Analysis ist  $w$  der Gradient von  $f$  im Punkt 0. Obiges Diagramm visualisiert die Situation für ein konkretes  $f$ .

Der Rieszsche Darstellungssatz ist für unendlich-dimensionale Vektorräume nicht mehr ohne zusätzliche Voraussetzungen gültig:

### Beispiel

Sei  $V$  der  $\mathbb{R}$ -Vektorraum der reellen Polynomfunktionen auf  $\mathbb{R}$  mit

$$\langle f, g \rangle = \int_{-1}^1 f(x) g(x) dx \quad \text{für alle } f, g \in V.$$

Wir betrachten das lineare Funktional  $F: V \rightarrow \mathbb{R}$  mit

$$F(f) = f(0) \quad \text{für alle } f \in V. \quad (\text{Auswertung am Nullpunkt})$$

*Annahme*, es gibt ein  $g \in V$  mit  $F(f) = \langle g, f \rangle$  für alle  $f \in V$ . Dann gilt  $\langle g, x^2 g \rangle = 0$ , da das Polynom  $x^2 g$  im Nullpunkt gleich 0 ist. Damit ist aber

$$\int_{-1}^1 x^2 g(x)^2 dx = \langle g, x^2 g \rangle = 0.$$

Dies ist nur möglich, wenn  $g = 0$ . Dann ist aber  $F = \langle g, \cdot \rangle = 0$ , *Widerspruch*.

### Exkurs I: Der Darstellungssatz für stetige Funktionale auf Hilbert-Räumen

Ist ein euklidischer oder unitärer Vektorraum  $V$  bzgl. der durch das Skalarprodukt induzierten Norm vollständig (im Sinne der Konvergenz von Cauchy-Folgen in  $V$ ), so nennt man  $V$  einen *Hilbert-Raum*. So ist beispielsweise der  $\ell^2(\mathbb{C})$  ein Hilbert-Raum. Der Rieszsche Darstellungssatz gilt nun für Hilbert-Räume, wenn man sich auf stetige Funktionale  $f: V \rightarrow K$  beschränkt. Jedes stetige Funktional hat also die eindeutige Form  $\langle w, \cdot \rangle$  und umgekehrt sind alle  $\langle w, \cdot \rangle$  stetige Funktionale.

### Exkurs II: Bra-Vektoren und Ket-Vektoren (Dirac-Notation)

In der mathematischen Physik schreibt man die lineare Abbildung  $\langle w, \cdot \rangle: V \rightarrow \mathbb{C}$  oft als *Bra-Vektor* in der Form  $\langle w |$ . Weiter schreibt man  $v \in V$  als *Ket-Vektor* in der Form  $| v \rangle$ . Die Sprechweisen sind durch Bra-Ket  $\sim$  bracket motiviert: Ein Bra-Vektor lässt sich auf einen Ket-Vektor anwenden:  $\langle w | | v \rangle = \langle w | v \rangle = \langle w, v \rangle$ . Ist nun  $(v_i)_{i \in I}$  eine Orthonormalbasis von  $V$ , so gilt für alle  $v = \sum_i \alpha_i v_i$ ,  $w = \sum_i \beta_i v_i \in V$

$$\sum_{i \in I} \langle w | v_i \rangle \langle v_i | v \rangle = \sum_{i \in I} \overline{\beta_i} \alpha_i = \langle w | v \rangle.$$

Damit lässt sich  $\sum_{i \in I} | v_i \rangle \langle v_i |$  als Identität interpretieren. Insgesamt entsteht ein suggestiver Kalkül, der insbesondere in der Quantenmechanik verwendet wird.

## 6.11 Der adjungierte Endomorphismus

### Definition (adjungierte Abbildung, selbstadjungiert)

Seien  $V, W$  euklidische oder unitäre Vektorräume, und sei  $f : V \rightarrow W$  ein Homomorphismus. Dann heißt der im Fall der Existenz eindeutig bestimmte Homomorphismus  $f^* : W \rightarrow V$  mit

$$(+)\quad \langle f^*(w), \cdot \rangle_V = \langle w, f(\cdot) \rangle_W \quad \text{für alle } w \in W, \text{ d. h.}$$

$$\langle f^*(w), v \rangle_V = \langle w, f(v) \rangle_W \quad \text{für alle } v \in V \text{ und alle } w \in W$$

der zu  $f$  *adjungierte Homomorphismus*. Gilt  $V = W$  und  $f = f^*$ , so heißt der Endomorphismus  $f : V \rightarrow V$  *selbstadjungiert*.

Zur Motivation der Bedingung (+) betrachten wir die zu  $f : V \rightarrow W$  duale Abbildung (vgl. 4.12)

$$f^\circ : W^* \rightarrow V^*,$$

$$f^\circ(g) = g \circ f \quad \text{für alle } g \in W^*$$

(wir schreiben  $f^\circ$ , da wir  $f^*$  für die Adjungierte reservieren). Jedes lineare Funktional der rieszschen Form

$$\langle w, \cdot \rangle : W \rightarrow K$$

können wir mit Hilfe von  $f^\circ$  zu einem linearen Funktional

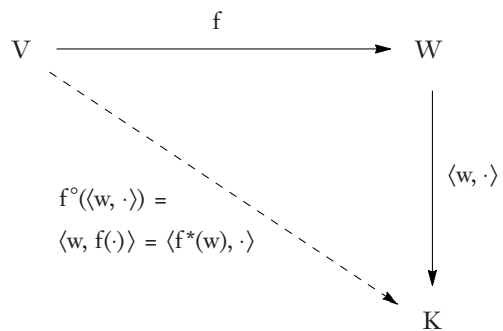
$$f^\circ(\langle w, \cdot \rangle) = \langle w, \cdot \rangle \circ f = \langle w, f(\cdot) \rangle : V \rightarrow K$$

zurückziehen. Nun hoffen wir, dass dieses Funktional erneut von der Form  $\langle u, \cdot \rangle : V \rightarrow K$  ist, für ein im Fall der Existenz eindeutig bestimmtes  $u \in V$ . Existiert  $u$ , so gilt

$$\langle u, \cdot \rangle = f^\circ(\langle w, \cdot \rangle) = \langle w, f(\cdot) \rangle.$$

Dies ist genau die Bedingung (+) für den Vektor  $w$  und  $f^*(w) = u$ . Wir fassen zusammen:

*Der adjungierte Homomorphismus liefert die Umrechnung von rieszschen linearen Funktionalen auf  $W$  zu rieszschen linearen Funktionalen auf  $V$  gemäß des durch  $f$  gegebenen Pullbacks von Funktionalen.*



Die Adjunktion als Pullback

Da für endlich-dimensionale Vektorräume jedes lineare Funktional die rieszsche Form hat, zeigt unsere Argumentation:

Sind  $V, W$  endlich-dimensional, so existiert  $f^*$ .

In Matrizenform lässt sich die Adjungierte einfach handhaben:

### Die darstellende Matrix von $f^*$ ist $A^t$ bzw. $A^*$

Sei  $K = \mathbb{R}$ , und seien  $(v_1, \dots, v_n)$  und  $(w_1, \dots, w_m)$  Orthonormalbasen von  $V$  bzw.  $W$ . Weiter sei  $A \in K^{m \times n}$  die darstellende Matrix von  $f : V \rightarrow W$  bzgl. dieser Basen und  $B \in K^{n \times m}$  die darstellende Matrix von  $f^* : W \rightarrow V$  bzgl.  $(w_1, \dots, w_m), (v_1, \dots, v_n)$ . Dann gilt („Die Spalten sind die Koordinatenvektoren der Bilder der Basisvektoren.“):

$$B(j, i) = \langle v_j, f^*(w_i) \rangle = \langle f^*(w_i), v_j \rangle = \langle w_i, f(v_j) \rangle = A(i, j) \quad \text{für alle } i, j.$$

Damit ist  $B = A^t$  (vgl. den Dualitätssatz in 5.9). Für  $K = \mathbb{C}$  ist  $B = A^*$ , da beim zweiten Gleichheitszeichen eine komplexe Konjugation auftritt. Ist  $f : V \rightarrow V$  selbstadjungiert und  $K = \mathbb{R}$ , so gilt  $A = A^t$ , d.h.,  $A$  ist symmetrisch. Für  $K = \mathbb{C}$  erhalten wir  $A = A^*$ . Matrizen mit dieser Eigenschaft nennt man *hermitesch*.

Für die Adjungierte gelten  $(\lambda f + g)^* = \lambda f^* + g^*$  und  $(g \circ f)^* = f^* \circ g^*$ . Weiter ist

$$\begin{aligned} \text{Kern}(f) &= \{ v \in V \mid \langle w, f(v) \rangle = 0 \text{ für alle } w \in W \} = \\ &= \{ v \in V \mid \langle f^*(w), v \rangle = 0 \text{ für alle } w \in W \} = \text{Bild}(f^*)^\perp. \end{aligned}$$

Für endlich-dimensionale Vektorräume  $V$  und  $W$  ergibt sich nach der Dimensionsformel und  $\dim(\text{Bild}(f^*)^\perp) = \dim(V) - \dim(\text{Bild}(f^*))$  also  $\dim(\text{Bild}(f)) = \dim(\text{Bild}(f^*))$  (vgl. 5.9).

Dass  $f^*$  im Fall der Existenz eindeutig bestimmt ist, folgt unabhängig vom Rieszschen Darstellungssatz aus  $\langle v, \cdot \rangle \neq \langle u, \cdot \rangle$  für alle  $v \neq u$  in  $V$ . Für unendlich-dimensionale Vektorräume kann ein adjungierter Homomorphismus existieren oder nicht:

### Beispiele

- (1) Ist  $V = \ell_{\mathbb{R}}^2$  und  $f : V \rightarrow V$  mit  $f(x_0, x_1, \dots) = (x_1, x_2, \dots)$  (Linksshift), so ist  $f^* : V \rightarrow V$  mit  $f^*(x_0, x_1, \dots) = (0, x_0, x_1, \dots)$  (Rechtsshift) die Adjungierte von  $f$ . Es gilt  $f \circ f^* = \text{id}_V$ , aber  $f^* \circ f \neq \text{id}_V$ .

- (2) Sei  $V$  der  $\mathbb{R}$ -Vektorraum der reellen Polynomfunktionen auf  $\mathbb{R}$  mit dem Integral von  $-1$  bis  $1$  über  $f(x)g(x)$  als Skalarprodukt. Sei  $D : V \rightarrow V$  mit  $D(f) = f'$  für alle  $f \in V$  (Ableitungsoperator). *Annahme*, die Adjungierte  $D^* : V \rightarrow V$  von  $D$  existiert. Dann gilt nach partieller Integration

$$\langle D^*f, g \rangle = \langle f, Dg \rangle = f(1)g(1) - f(-1)g(-1) - \langle Df, g \rangle, \quad \text{also}$$

$$\langle (D^* + D)f, g \rangle = f(1)g(1) - f(-1)g(-1) \quad \text{für alle } f, g \in V.$$

Damit ist  $\langle (D + D^*)f, (x-1)^2(x+1)^2(D + D^*)f \rangle = 0$  für alle  $f$ , also  $D + D^* = 0$  (denn für  $g \neq 0$  ist das Integral über  $g^2(x-1)^2(x+1)^2$  von  $-1$  bis  $1$  größer als  $0$ ). Also gilt  $D^* = -D$ , sodass  $D^*1 = 0$ . Da  $D^*$  die Adjungierte von  $D$  ist, gilt  $0 = \langle D^*1, x \rangle = \langle 1, Dx \rangle = \langle 1, 1 \rangle = 2$ , *Widerspruch*.

Man kann mit dem allgemeinen Rieszschen Darstellungssatz zeigen, dass  $f^*$  für ein stetiges lineares  $f : V \rightarrow W$  zwischen Hilbert-Räumen  $V, W$  immer existiert (vgl. 6.10).

## 6.12 Sesquilinearformen

### Definition (Sesquilinearform, symmetrische und hermitesche Form, Definitheit)

Sei  $V$  ein  $K$ -Vektorraum mit  $K = \mathbb{R}$  oder  $K = \mathbb{C}$ . Eine Abbildung  $\varphi : V \times V \rightarrow K$  heißt eine *Sesquilinearform*, falls für alle  $v, w \in V$  und  $\lambda \in K$  gilt:

$$\varphi(v + \lambda v', w) = \varphi(v, w) + \bar{\lambda} \varphi(v', w), \quad \varphi(v, w + \lambda w') = \varphi(v, w) + \lambda \varphi(v, w').$$

Gilt zusätzlich  $\varphi(v, w) = \overline{\varphi(w, v)}$  für alle  $v, w \in V$ , so heißt  $\varphi$  eine *symmetrische Bilinearform*, falls  $K = \mathbb{R}$ , bzw. eine *hermitesche Form*, falls  $K = \mathbb{C}$ . Eine solche Form heißt

<i>positiv definit</i> ,	falls	$\varphi(v, v) > 0$ für alle $v \in V$ mit $v \neq 0$ ,
<i>positiv semidefinit</i> ,	falls	$\varphi(v, v) \geq 0$ für alle $v \in V$ ,
<i>negativ (semi-) definit</i> ,	falls	$-\varphi$ positiv (semi-) definit ist,
<i>indefinit</i> ,	falls	$v, w \in V$ existieren mit $\varphi(v, v) > 0$ und $\varphi(w, w) < 0$ .

Die Definition verallgemeinert den Begriff eines Skalarprodukts. Ein Skalarprodukt ist eine symmetrische bzw. hermitesche Form, die positiv definit ist.

### Beispiel

Seien  $n \geq 1$  und  $A \in K^{n \times n}$ . Dann wird eine Sesquilinearform  $\varphi$  auf  $K^n$  definiert durch

$$\varphi(v, w) = \langle v, Aw \rangle_{\text{kanonisch}} \quad \text{für alle } v, w \in K^n.$$

Die Form  $\varphi$  ist genau dann symmetrisch bzw. hermitesch, wenn die Matrix  $A$  dies ist.

Ist  $V$  endlich-dimensional, so sind die Formen dieses Beispiels im folgenden Sinn bereits alle Formen:

### Die gramsche Matrix einer Sesquilinearform

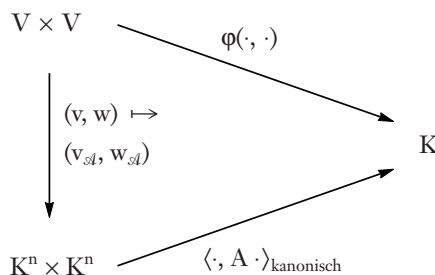
Sei  $\varphi : V \times V \rightarrow K$  eine Sesquilinearform und sei  $\mathcal{A} = (v_1, \dots, v_n)$  eine Basis von  $V$ . Dann ist die *gramsche Matrix*  $A_\varphi = A_{\varphi, \mathcal{A}} \in K^{n \times n}$  von  $\varphi$  bzgl.  $\mathcal{A}$  definiert durch

$$A_\varphi(i, j) = \varphi(v_i, v_j) \quad \text{für alle } i, j.$$

Die Form  $\varphi$  ist genau dann symmetrisch bzw. hermitesch, wenn  $A_\varphi$  dies ist. Mit der Koordinatenabbildung  $\Phi_{\mathcal{A}} : V \rightarrow K^n$  und  $v_{\mathcal{A}} = \Phi_{\mathcal{A}}(v)$  gilt

$$(+)\quad \varphi(v, w) = \langle v_{\mathcal{A}}, A_\varphi w_{\mathcal{A}} \rangle_{\text{kanonisch}}$$

für alle  $v, w \in V$ . Definieren wir umgekehrt eine Form  $\varphi$  durch (+) mit einer beliebigen Matrix  $A$  des  $K^{n \times n}$ , so ist  $A_\varphi = A$ .



Wir betrachten zwei Stufen, die zwischen beliebigen Sesquilinearformen und vollwertigen Skalarprodukten liegen, genauer.

## I. Symmetrische und hermitesche Formen

Wir notieren diese Formen wie Skalarprodukte oft als  $\langle \cdot, \cdot \rangle : V \times V \rightarrow K$ . Die Begriffe „orthogonal“ und „Orthogonalbasis“ sind wie früher definiert, und erneut stellt sich die Frage nach der Existenz einer Orthogonalbasis. Das Verfahren von Gram-Schmidt kann an einer Division durch  $\langle v, v \rangle = 0$  für ein  $v \neq 0$  scheitern. Dennoch ist es richtig, dass endlich-dimensionale Vektorräume, die mit einer symmetrischen oder hermiteschen Form versehen sind, eine Orthogonalbasis bzgl. dieser Form besitzen. Ist  $\mathcal{A} = (v_1, \dots, v_n)$  eine solche Orthogonalbasis, so ist die gramsche Matrix  $A \in K^{n \times n}$  der Form bzgl.  $\mathcal{A}$  eine Diagonalmatrix. Wir werden in Kapitel 8 bei der Diskussion der Hauptachsentransformation darauf zurückkommen.

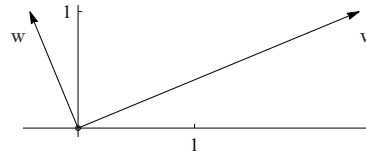
### Beispiel

Für die Bilinearform  $\langle \cdot, \cdot \rangle = \langle \cdot, A \cdot \rangle_{\text{kanonisch}}$  auf  $\mathbb{R}^2$  mit der Matrix  $A \in \mathbb{R}^{2 \times 2}$  rechts gilt

$$A = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

$$\langle (x_1, y_1), (x_2, y_2) \rangle = x_1(x_2 + y_2) + y_1(x_2 - y_2).$$

Die Vektoren  $v = (1 + \sqrt{2}, 1)$ ,  $w = (1 - \sqrt{2}, 1)$  bilden eine Orthogonalbasis von  $\mathbb{R}^2$  bzgl. der Form  $\langle \cdot, \cdot \rangle$  (und bzgl.  $\langle \cdot, \cdot \rangle_{\text{kanonisch}}$ ).



## II. Positiv semidefinite symmetrische und hermitesche Formen

Für diese Formen gilt die Cauchy-Schwarz-Ungleichung

$$|\langle v, w \rangle| \leq \|v\| \|w\| \quad \text{für alle } v, w \in V, \text{ wobei } \|u\| = \sqrt{\langle u, u \rangle}.$$

Gleichheit kann nun auch für linear unabhängige  $v, w$  eintreten (man betrachte die Nullform). Die Halb- oder Seminorm  $\|\cdot\| : V \rightarrow K$  erfüllt die Homogenität und die Dreiecksungleichung, aber  $\|v\| = 0$  ist für  $v \neq 0$  möglich. In diesem Fall lässt sich  $v$  nicht normieren.

### Beispiel

Für reelle  $a < b$  und  $V = \{ f : [a, b] \rightarrow \mathbb{C} \mid f \text{ ist Riemann-integrierbar} \}$  definiert

$$\langle f, g \rangle = \int_a^b \overline{f(x)} g(x) dx \quad \text{für alle } f, g \in V$$

eine positiv semidefinite hermitesche Form. Die Form ist nicht positiv definit, da  $\langle f, f \rangle = 0$  gilt, wenn  $f$  an höchstens abzählbar vielen Stellen ungleich null ist. Diese Form spielt insbesondere in der Theorie der Fourier-Reihen eine wichtige Rolle.

Durch Faktorisierung kann man die positive Semidefinitheit zur positiven Definitheit verstärken: Ist  $U = \{ u \in V \mid \langle u, u \rangle = 0 \}$ , so wird auf dem Faktorraum  $V/U$  ein Skalarprodukt durch  $\langle v + U, w + U \rangle = \langle v, w \rangle$  für alle  $v, w \in V$  definiert.





# 7. Kapitel

---

## Determinanten

---

## 7.1 $2 \times 2$ -Determinanten

### Definition (Determinantenfunktion, Determinante einer $2 \times 2$ -Matrix)

Seien  $K$  ein Körper und  $K^{2 \times 2}$  die Menge der  $2 \times 2$ -Matrizen über  $K$ . Dann heißt eine Abbildung  $\det : K^{2 \times 2} \rightarrow K$  eine *Determinantenfunktion* auf  $K^{2 \times 2}$ , falls gilt:

#### Multilinearität in den Spalten

Für alle  $a, b, c \in K^2$  und alle  $\lambda \in K$  gilt

$$\det(a; \lambda b + c) = \lambda \det(a; b) + \det(a; c),$$

$$\det(\lambda a + b; c) = \lambda \det(a; c) + \det(b; c),$$

d. h., für alle  $a, b \in K^2$  sind die folgenden Abbildungen linear:

$$\det(a; \cdot) : K^2 \rightarrow K, \quad \det(\cdot; b) : K^2 \rightarrow K.$$

#### Alternation

Für alle  $a \in K^2$  gilt  $\det(a; a) = 0$ .

#### Normiertheit

Es gilt  $\det E_2 = 1$ .

Für alle  $A \in K^{2 \times 2}$  heißt dann  $\det A$  die *Determinante* der Matrix  $A$ .

Das Thema dieses Kapitels sind „gute“ Funktionen  $\det : K^{n \times n} \rightarrow K$  mit

$\det A = 0$  genau dann, wenn  $A$  singulär.

Dann ist  $\det A \neq 0$  äquivalent dazu, dass das lineare Gleichungssystem  $Ax = b$  für alle  $b \in K^n$  eindeutig lösbar ist. Die Lösbarkeit ist also durch die Determinante festgelegt (determinare = bestimmen).

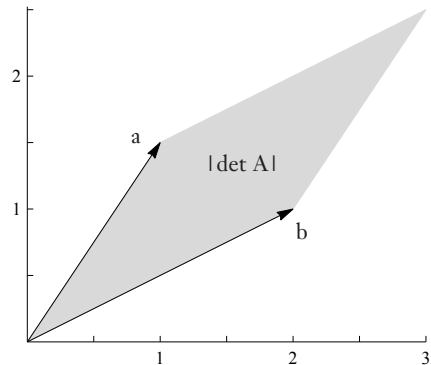
Anstelle einer direkten Definition verfolgen wir, wie seit Karl Weierstraß 1886 üblich, einen axiomatischen Zugang. Dabei hat sich das Trio

„multilinear, alternierend, normiert“

als besonders geeignet erwiesen. Wir untersuchen es in diesem Abschnitt für den Spezialfall  $n = 2$ , im nächsten Abschnitt werden wir eine beliebige Dimension  $n \geq 1$  zulassen.

### Notation

Wir trennen die Spalten einer Matrix oft durch Strichpunkte voneinander ab, wenn dies der Lesbarkeit dient.



$$\det A = a_1 b_2 - b_1 a_2 = 1 \cdot 1 - 2 \cdot 3/2 = -2$$

Die Determinante der Matrix  $A$  mit den Spalten  $a$  und  $b$  ist dem Betrag nach der Flächeninhalt des von  $a$  und  $b$  aufgespannten Parallelogramms. In 5.9 untersuchen wir die geometrische Bedeutung der Determinante genauer.

### Determinanten als bilineare Abbildungen

Fassen wir Matrizen des  $K^{2 \times 2}$  als Elemente von  $K^2 \times K^2$  auf, so ist eine Determinantenfunktion  $\det : K^2 \times K^2 \rightarrow K$  eine bilineare Funktion (vgl. Kapitel 6). Die Alternation bedeutet, dass für alle  $a \in K^2$  das Paar  $(a, a)$  auf null abgebildet wird. Dies steht im starken Kontrast zur positiven Definitheit  $\langle v, v \rangle > 0$  für  $v \neq 0$  eines Skalarprodukts.

Zur Illustration der Konsequenzen und des Zusammenspiels der drei grundlegenden Eigenschaften zeigen wir:

### Existenz und Eindeutigkeit der $2 \times 2$ -Determinantenfunktion

Ist  $\det : K^{2 \times 2} \rightarrow K$  eine Determinantenfunktion, so gilt für alle  $a, b \in K^2$  und  $\lambda \in K$ :

- (1)  $\det(a; b + \lambda a) = \det(a; b) + \lambda \det(a; a) = \det(a; b)$ ,
- (2)  $0 = \det(a + b; a + b) = \det(a; a) + \det(a; b) + \det(b; a) + \det(b; b) = \det(a; b) + \det(b; a)$ , sodass  $\det(a; b) = -\det(b; a)$ .

Die Determinante bleibt also bei der Addition des  $\lambda$ -Fachen einer Spalte zu einer anderen unverändert und bei einer Spaltenvertauschung ändert sich das Vorzeichen. Ist nun  $A \in K^{2 \times 2}$  die Matrix mit den Spalten  $a = (a_1, a_2)$  und  $b = (b_1, b_2)$ , so können wir im Fall  $a_1 \neq 0$  die Matrix durch Addition des  $\lambda = -b_1/a_1$ -Fachen der ersten Spalte auf die zweite Spalte auf die Form

$$\begin{pmatrix} a_1 & 0 \\ a_2 & b_2 - \lambda a_2 \end{pmatrix} \quad \text{und im Fall } b_2 \neq \lambda a_2 \text{ weiter auf } \begin{pmatrix} a_1 & 0 \\ 0 & b_2 - \lambda a_2 \end{pmatrix}$$

bringen, ohne die Determinante zu verändern. Nach Multilinearität und Normierung ist die Determinante einer Diagonalmatrix das Produkt ihrer Diagonaleinträge, sodass

$$\det A = a_1 \cdot (b_2 - \lambda a_2) = a_1 b_2 - b_1 a_2.$$

Dieselbe Formel ergibt sich für alle anderen Fälle bei analoger Argumentation. Umgekehrt ist die durch diese Formel definierte Funktion auf  $K^{2 \times 2}$  multilinear, alternierend und normiert. Damit existiert auf dem  $K^{2 \times 2}$  genau eine Determinantenfunktion.

### Beispiele

$$(1) \det \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} = -2, \quad \det \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} = -1, \quad \det \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} = 1 \quad \text{für alle } \alpha \in \mathbb{R}.$$

- (2) Die Determinantenfunktion ist multilinear, aber nicht linear. Es gilt
 
$$\det E_2 + \det E_2 = 2 \neq 4 = \det(2E_2).$$

## 7.2 $n \times n$ -Determinanten

### Definition (Determinantenfunktion, Determinante einer $n \times n$ -Matrix)

Seien  $K$  ein Körper und  $n \geq 1$ . Dann heißt eine Abbildung  $\det : K^{n \times n} \rightarrow K$  eine *Determinantenfunktion* auf  $K^{n \times n}$ , falls gilt:

#### Multilinearität in den Spalten

Für alle  $1 \leq k \leq n$  und alle  $a_1, \dots, a_k, a_{k+1}, \dots, a_n \in K^n$  ist die Abbildung

$$\det \begin{pmatrix} a_1 & \dots & a_{k-1} & \cdot & a_{k+1} & \dots & a_n \end{pmatrix} : K^n \rightarrow K$$

linear.

#### Alternation

Für  $A \in K^{n \times n}$  mit zwei gleichen Spalten gilt  $\det A = 0$ .

$$\det \begin{pmatrix} \dots & \lambda a_1 + b_1 & \dots \\ \dots & \dots & \dots \\ \dots & \lambda a_n + b_n & \dots \end{pmatrix} =$$

#### Normiertheit

Es gilt  $\det E_n = 1$ .

$$\lambda \det \begin{pmatrix} \dots & a_1 & \dots \\ \dots & \dots & \dots \\ \dots & a_n & \dots \end{pmatrix} + \det \begin{pmatrix} \dots & b_1 & \dots \\ \dots & \dots & \dots \\ \dots & b_n & \dots \end{pmatrix}$$

Aus den Determinantenaxiomen

„*multilinear, alternierend, normiert*“

lassen sich herleiten:

$\det(\text{diag}(d_1, \dots, d_n)) = d_1 \cdot \dots \cdot d_n.$	(Diagonalprodukt I)
Ist $W_{ij}(\lambda)$ , $i \neq j$ , ein Additionstyp (vgl. 5.6), so gilt $\det(AW_{ij}(\lambda)) = \det(A).$	(Spaltenaddition)
Ist $P_{ij}$ eine Transpositionsmatrix (vgl. 5.7), so gilt $\det(AP_{ij}) = -\det(A).$	(Spaltentausch)
$\det(A) \neq 0$ genau dann, wenn $A \in GL(n, K).$	(Invertierbarkeit)
Die Determinante einer oberen bzw. unteren Dreiecksmatrix ist das Produkt der Diagonaleinträge.	(Diagonalprodukt II)
$\det(\lambda A) = \lambda^n \det A$ , $\det(-A) = (-1)^n \det(A).$	(Skalierung)

Wie im Fall  $n = 2$  bleibt eine Determinante also bei der Addition des  $\lambda$ -fachen einer Spalte zu einer anderen gleich, während sie beim Tausch zweier Spalten ihr Vorzeichen ändert. Wir werden später sehen, dass diese Eigenschaften auch für die Zeilen gelten.

Die Regeln für Diagonalmatrizen, für Spaltenadditionen und für Spaltenvertauschungen lassen sich wie im Fall  $n = 2$  einsehen. Damit können wir zeigen:

### Invertierbarkeit = Nichtverschwinden der Determinante

Ist  $A$  nicht invertierbar, so ist eine Spalte  $a_k$  von  $A$  eine Linearkombination der anderen. Aufgrund der Multilinearität und der Alternation der Determinante gilt dann

$$(+)\det A = \det \left( a_1; \dots a_{k-1}; \sum_{i \neq k} \alpha_i a_i; a_{k+1}; \dots a_n \right) = 0.$$

Ist  $A$  invertierbar, so lässt sich  $A$  mit Hilfe von Spaltenadditionen  $W_{ij}(\lambda)$ ,  $i \neq j$ , in eine Diagonalmatrix  $B$  mit Diagonaleinträgen  $b_{ii} \neq 0$  überführen. Die Determinante bleibt dabei gleich. Damit gilt

$$(++)\det A = \det B = b_{11} \cdot \dots \cdot b_{nn} \neq 0.$$

Die Argumentation liefert mehr:

<i>Existenz- und Eindeigkeitssatz</i>
Für alle $n \geq 1$ existiert genau eine Determinantenfunktion auf dem $K^{n \times n}$ .

Denn auf den singulären Matrizen sind Determinantenfunktionen  $\det$  und  $\det'$  gleich 0. Und für eine invertierbare Matrix  $A$  gilt  $\det A = b_{11} \cdot \dots \cdot b_{nn} = \det' A$ , da wir zur Herleitung von  $(++)$  nur die Determinantenaxiome eingesetzt haben. Umgekehrt können wir  $(+)$  und  $(++)$  zur Definition von  $\det(A)$  verwenden und dann die Determinantenaxiome beweisen (wobei die Multilinearität etwas Arbeit erfordert). Andere Beweise der Existenz und Eindeigkeit werden wir später kennenlernen.

### Berechnung von Determinanten durch Überführung in Dreiecksmatrizen

Eine Matrix  $A$  können wir durch Spaltenadditionen in eine Dreiecksmatrix  $B$  überführen, ohne die Determinante zu verändern. Dann ist  $\det(A) = b_{11} \cdot \dots \cdot b_{nn}$ .

Zum Beweis dieser zweiten Diagonalprodukt-Regel beobachten wir:

Ist  $A$  singulär, so hat  $B$  eine Null auf der Diagonalen, sodass  $\det A = 0 = b_{11} \cdot \dots \cdot b_{nn}$ .

Andernfalls können wir  $B$  durch weitere Spaltenadditionen in eine Diagonalmatrix  $C$  verwandeln, ohne die Diagonaleinträge  $b_{ii}$  oder die Determinante zu verändern.

Dann gilt  $\det A = \det B = \det C = c_{11} \cdot \dots \cdot c_{nn} = b_{11} \cdot \dots \cdot b_{nn}$ .

### Beispiel

Ausräumen oberhalb der Diagonalen mit Hilfe von Spaltenadditionen zeigt:

$$\det \begin{pmatrix} 1 & 1 & -1 & -1 \\ 1 & 4 & 2 & 5 \\ 1 & 3 & 0 & 3 \\ 3 & 6 & 0 & 1 \end{pmatrix} = \det \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 3 & 3 & 6 \\ 1 & 2 & 1 & 4 \\ 3 & 3 & 3 & 4 \end{pmatrix} = \det \begin{pmatrix} 1 & 0 & 0 & 0 \\ 1 & 3 & 0 & 0 \\ 1 & 2 & -1 & 0 \\ 3 & 3 & 0 & -2 \end{pmatrix} = 6.$$

### 7.3 Das Vorzeichen einer Permutation

**Definition** (*Vorzeichen, gerade, ungerade, alternierende Gruppe*)

Seien  $n \geq 1$  und  $S_n$  die Gruppe der Permutationen auf  $\{1, \dots, n\}$ . Dann ist die *Vorzeichenfunktion*  $\text{sgn} : S_n \rightarrow \{-1, 1\}$  definiert durch

$$\text{sgn}(\sigma) = \prod_{1 \leq i < j \leq n} \frac{\sigma(j) - \sigma(i)}{j - i} \quad \text{für alle } \sigma \in S_n.$$

Wir nennen  $\text{sgn}(\sigma)$  das *Vorzeichen* oder *Signum* der Permutation  $\sigma$ . Eine Permutation  $\sigma$  heißt *gerade*, falls  $\text{sgn}(\sigma) = 1$ , und *ungerade*, falls  $\text{sgn}(\sigma) = -1$ . Wir setzen:

$$A_n = \{\sigma \in S_n \mid \text{sgn}(\sigma) = 1\}. \quad (\text{alternierende Gruppe})$$

Permutationen und ihre Vorzeichen spielen in der Theorie der Determinanten eine wichtige Rolle. Wir werden im nächsten Abschnitt Determinantenfunktionen mit Hilfe von Permutationen explizit definieren. In diesem Abschnitt treffen wir die nötigen algebraischen Vorbereitungen.

Aufgrund der Bijektivität einer Permutation  $\sigma : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$  gilt

$$\{\{i, j\} \mid 1 \leq i < j \leq n\} = \{\{\sigma(i), \sigma(j)\} \mid 1 \leq i < j \leq n\}.$$

Hieraus liest man ab, dass der Zähler und der Nenner des Produkts

$$\prod_{1 \leq i < j \leq n} \frac{\sigma(j) - \sigma(i)}{j - i}$$

abgesehen von den Vorzeichen dieselben Faktoren enthalten. Damit ist  $\text{sgn}(\sigma) \in \{-1, 1\}$ . Nennen wir ein Paar  $(i, j)$  mit  $i < j$  einen *Fehlstand* von  $\sigma$ , falls  $\sigma(i) > \sigma(j)$ , so gilt also:

$$\text{Ist } k \text{ die Anzahl der Fehlstände von } \sigma, \text{ so ist } \text{sgn}(\sigma) = (-1)^k.$$

#### Beispiele

- (1) Die Permutation  $(1, \dots, n)$  hat keine Fehlstände und damit das Vorzeichen  $(-1)^0 = 1$ .
- (2) Die Permutation  $\sigma = (2, 3, \dots, n, 1)$  hat die Fehlstände  $(1, n), \dots, (n-1, n)$ . Damit ist  $\text{sgn}(\sigma) = (-1)^{n-1}$ .
- (3) Die Permutation  $\sigma = (n, \dots, 1)$  hat  $n(n-1)/2$  Fehlstände. Gilt  $n \equiv 0 \pmod{4}$  oder  $n \equiv 1 \pmod{4}$ , so ist  $\text{sgn}(\sigma) = 1$ . Andernfalls ist  $\text{sgn}(\sigma) = -1$ .
- (4) Ist  $\tau \in S_n$  die Transposition, die  $i < j$  vertauscht, so enthält die Produktformel der Definition von  $\text{sgn}(\tau)$  genau einen Faktor  $-1$  und sonst nur Einsen. Damit ist  $\text{sgn}(\tau) = -1$ .

### Homomorphie der Vorzeichenfunktion

Für alle  $\pi, \sigma \in S_n$  gilt  $\operatorname{sgn}(\pi \circ \sigma) = \operatorname{sgn}(\pi) \operatorname{sgn}(\sigma)$ . Die Abbildung  $\operatorname{sgn} : S_n \rightarrow \{-1, 1\}$  ist also ein Gruppenhomomorphismus. Speziell gilt  $\operatorname{sgn}(\sigma^{-1}) = \operatorname{sgn}(\sigma)^{-1}$  für alle  $\sigma \in S_n$ . Weiter ist  $A_n = \operatorname{Kern}(\operatorname{sgn})$ , sodass  $A_n$  ein Normalteiler von  $S_n$  ist.

Hieraus ergeben sich neue Möglichkeiten zur Berechnung des Vorzeichens. Ist  $\sigma \in S_n$  beliebig, so können wir ausgehend von  $(1, 2, \dots, n)$  durch Anwendung von Transpositionen oder der Identität Permutationen der Form

$$(\sigma(1), \dots, (\sigma(1), \sigma(2), \dots), \dots, (\sigma(1), \dots, \sigma(n))) \quad (\text{schrittweises Einstellen der Werte})$$

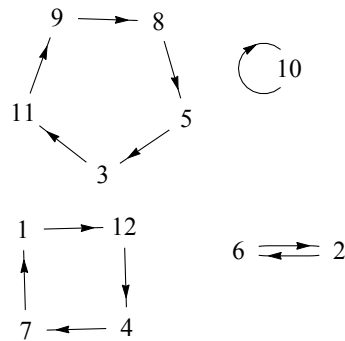
erzeugen. Damit ist jede Permutation das Produkt von höchstens  $n - 1$  Transpositionen ( $\sigma(n)$  ist automatisch richtig, wenn alle anderen Werte richtig sind). Da jede Transposition das Vorzeichen  $-1$  besitzt, erhalten wir:

$$\text{Ist } \sigma = \tau_k \circ \dots \circ \tau_1 \text{ mit Transpositionen } \tau_i, \text{ so ist } \operatorname{sgn}(\sigma) = (-1)^k.$$

Eine anschauliche Analyse liefert die Zerlegung einer Permutation in Zyklen. Ist  $\sigma \in S_n$  und  $i \in \{1, \dots, n\}$ , so können wir aufgrund der Injektivität von  $\sigma$  die Bahn

$$B(i) = \{i, \sigma(i), \sigma^2(i), \dots, \sigma^k(i) = i\}$$

bilden. Die Permutation  $\pi$  mit  $\pi(j) = \sigma(j)$  für  $j \in B(i)$  und  $\sigma(j) = j$  für  $j \notin B(i)$  heißt der von  $i$  erzeugte *Zyklus von  $\sigma$* . Jede Permutation ist das Produkt ihrer (untereinander kommutierenden) Zyklen. Hat eine Bahn  $B$  genau  $k$  Elemente, so hat der zugehörige Zyklus das Vorzeichen  $(-1)^{k-1}$  (vgl. Beispiel (2)). Da sich die Bahnlängen zu  $n$  aufsummieren, gilt:



Ein  $\sigma \in S_{12}$  mit vier Bahnen. Es gilt  
 $\operatorname{sgn}(\sigma) = (-1)^{12-4} = 1.$

$$\text{Hat } \sigma \in S_n \text{ genau } m \text{ Bahnen, so ist } \operatorname{sgn}(\sigma) = (-1)^{n-m}.$$

### Beispiele

- (1) Die Permutation  $(1, \dots, n)$  hat die Bahnen  $\{1\}, \dots, \{n\}$  und damit das Vorzeichen  $(-1)^{n-n} = 1$ . Die Zyklen der Bahnen sind jeweils die Identität.
- (2)  $\sigma = (2, 3, \dots, n, 1)$  hat nur die eine Bahn  $\{1, 2, \dots, n\}$ , sodass  $\operatorname{sgn}(\sigma) = (-1)^{n-1}$ .
- (3)  $\sigma = (7, \dots, 1)$  hat die Bahnen  $\{1, 7\}, \{2, 6\}, \{3, 5\}, \{4\}$ , sodass  $\operatorname{sgn}(\sigma) = -1$ .
- (4) Die Transposition, die  $i$  und  $j$  vertauscht, hat die Bahnen  $\{i, j\}$  und  $\{k\}$  mit  $k \in \{1, \dots, n\} - \{i, j\}$ . Das Vorzeichen ist also  $(-1)^{n-(n-1)} = -1$ .

## 7.4 Die Leibniz-Formel

### Satz (Formel von Leibniz)

Seien  $K$  ein Körper und  $n \geq 1$ . Dann definiert

$$\det A = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{\sigma(1), 1} \dots a_{\sigma(n), n} \quad \text{für alle } A \in K^{n \times n} \quad (\text{Leibniz-Formel})$$

die eindeutige Determinantenfunktion auf dem  $K^{n \times n}$ .

Die Leibniz-Formel erfordert eine Summation über alle Elemente der symmetrischen Gruppe  $S_n$ . Da die Mächtigkeit von  $S_n$  gleich  $n!$  ist und mit  $n$  sehr schnell wächst, ist die Leibniz-Formel keine praktikable Rechenformel. Dagegen ist sie ein wertvolles Element der Theorie.

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & a_{15} \\ a_{21} & a_{22} & a_{23} & a_{24} & a_{25} \\ a_{31} & a_{32} & a_{33} & a_{34} & a_{35} \\ a_{41} & a_{42} & a_{43} & a_{44} & a_{45} \\ a_{51} & a_{52} & a_{53} & a_{54} & a_{55} \end{pmatrix}$$

Die Leibniz-Determinante einer  $(5 \times 5)$ -Matrix hat  $5! = 120$  Summanden. Der im Diagramm dargestellte Summand  $a_{21} a_{52} a_{43} a_{14} a_{35}$  gehört zur Permutation  $\sigma = (2, 5, 4, 1, 3)$  mit  $\operatorname{sgn}(\sigma) = 1$ .

### Motivation der Formel

Wir nehmen an, dass  $\det' : K^{n \times n} \rightarrow K$  eine Determinantenfunktion ist, und zeigen, dass die Leibniz-Formel für  $\det'$  gelten muss. Eine Verifikation der Determinantenaxiome zeigt, dass durch die Formel tatsächlich eine Determinantenfunktion definiert wird. Dies liefert einen zweiten Beweis der Existenz und Eindeutigkeit.

Sei also  $A \in K^{n \times n}$ . Mit den kanonischen Basisvektoren  $e_1, \dots, e_n$  gilt

$$\det' A = \det' \left( \sum_i a_{i1} e_i; \quad \dots; \quad \sum_i a_{in} e_i \right) =_{(a)}$$

$$\sum_{1 \leq i_1, \dots, i_n \leq n} a_{i_1, 1} \dots a_{i_n, n} \det' (e_{i_1}; \dots; e_{i_n}) =_{(b)}$$

$$\sum_{\sigma \in S_n} a_{\sigma(1), 1} \dots a_{\sigma(n), n} \det' (e_{\sigma(1)}; \dots; e_{\sigma(n)}) =_{(c)}$$

$$\sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{\sigma(1), 1} \dots a_{\sigma(n), n} \det' E_n =_{(d)}$$

$$\sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{\sigma(1), 1} \dots a_{\sigma(n), n}.$$

Dabei verwenden wir:

- (a)  $n$ -mal die Multilinearität zur Darstellung als Summe der Länge  $n^n$ ,
- (b) die Alternation zur Reduktion der Summe auf  $n! = |S_n|$  Permutationen,
- (c) für jedes  $\sigma \in S_n$   $k(\sigma)$  Spaltenvertauschungen, die die vorliegende Matrix in  $E_n$  überführen und durch den Faktor  $\operatorname{sgn}(\sigma) = (-1)^{k(\sigma)}$  korrigiert werden,
- (d) die Normierung.



Wir bestimmen einige uns schon bekannte und einige neue Determinanten mit Hilfe der Leibniz-Formel.

### Beispiele

- (1) Ist  $A = \text{diag}(d_1, \dots, d_n)$  eine Diagonalmatrix und  $\sigma \in S_n$  nicht die Identität, so gibt es ein  $i$  mit  $\sigma(i) \neq i$  und daher  $a_{\sigma(i), i} = 0$ . Damit trägt lediglich die Identität etwas zur Leibniz-Summe bei, sodass  $\det A = \text{sgn}(\text{id}) a_{11} \dots a_{nn} = d_1 \dots d_n$ .
- (2) Allgemeiner als Beispiel (1) zeigt die Leibniz-Formel, dass die Determinante einer oberen oder unteren Dreiecksmatrix das Produkt ihrer Diagonaleinträge ist.
- (3) Ist  $\sigma \in S_n$  und ist

$$A = \text{diag}(d_1, \dots, d_n) P_\sigma = (d_1 e_{\sigma(1)}; \dots; d_n e_{\sigma(n)})$$

eine umgeordnete Diagonalmatrix, so trägt lediglich die Permutation  $\sigma$  etwas zur Leibniz-Summe bei. Damit ist

$$\det A = \text{sgn}(\sigma) d_1 \dots d_n.$$

Speziell ist  $\det(P_\sigma) = \text{sgn}(\sigma)$  (was wir im Übergang von (c) zu (d) oben schon verwendet haben).

- (4) Für  $n = 2$  gibt es genau die Permutationen  $(1, 2)$  und  $(2, 1)$ . Damit gilt für alle  $A \in K^{n \times n}$

$$\det A = \text{sgn}(1, 2) a_{11} a_{22} + \text{sgn}(2, 1) a_{21} a_{12} = a_{11} a_{22} - a_{21} a_{12}.$$

Damit haben wir die in 7.1 gefundene Formel für  $2 \times 2$ -Matrizen reproduziert.

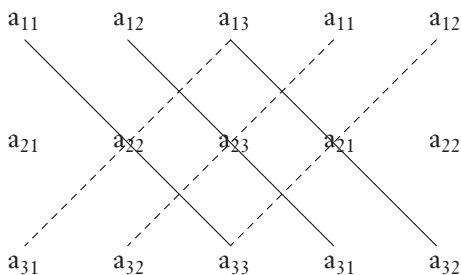
- (5) Für  $n = 3$  gibt es genau sechs Permutationen:

$(1, 2, 3), (2, 3, 1), (3, 1, 2)$  haben das Vorzeichen 1,

$(3, 2, 1), (2, 1, 3), (1, 3, 2)$  haben das Vorzeichen  $-1$ .

Damit gilt für alle  $A \in K^{3 \times 3}$

$$\begin{aligned} \det A = & + a_{11} a_{22} a_{33} + a_{21} a_{32} a_{13} + a_{31} a_{12} a_{23} \\ & - a_{31} a_{22} a_{13} - a_{21} a_{12} a_{33} - a_{11} a_{32} a_{23}. \end{aligned} \quad (\text{Regel von Sarrus})$$



Merkhilfe zur Regel von Sarrus: Die Produkte entlang der drei durchgezogenen (gestrichelten) Diagonalen haben ein positives (negatives) Vorzeichen.

## 7.5 Multiplikation und Transposition

### Satz (Multiplikationssatz und Transpositionssatz für Determinanten)

Seien  $K$  ein Körper und  $n \geq 1$ .

#### Multiplikationssatz

Für alle  $A, B \in K^{n \times n}$  gilt

$$\det(AB) = \det A \det B.$$

$$GL(n, K) \xrightarrow{\det} K^*$$

$$\det(AB) = \det(A) \det(B), \quad \det(A^{-1}) = \det(A)^{-1}$$

#### Transpositionssatz

Für alle  $A \in K^{n \times n}$  gilt

$$\det A^t = \det A.$$

Die Kernaussage des Multiplikationssatzes ist, dass die Determinantenfunktion ein Gruppenhomomorphismus von  $GL(n, K)$  in die multiplikative Gruppe  $K^*$  ist.

Die beiden Aussagen gehören zu den überraschenden Folgerungen der Determinantenaxiome. Sie lassen sich wie folgt beweisen.

### Beweis des Multiplikationssatzes

Ist  $B \in K^{n \times n}$  mit  $\det B = 0$  und  $A \in K^{n \times n}$  beliebig, so ist  $AB$  singulär. Folglich ist

$$\det(AB) = 0 = \det A \cdot 0 = \det A \det B.$$

Sei also  $B \in K^{n \times n}$  mit  $\det B \neq 0$ . Wir definieren  $\det' : K^{n \times n} \rightarrow K$  durch

$$\det' A = \frac{\det(AB)}{\det B} \quad \text{für alle } A \in K^{n \times n}.$$

Dann gelten alle Determinantenaxiome für  $\det'$ . Aufgrund der Eindeutigkeit einer Determinantenfunktion auf  $K^{n \times n}$  ist  $\det' = \det$  und damit

$$\det(AB) = \det' A \det B = \det A \det B \quad \text{für alle } A \in K^{n \times n}.$$

Das Argument ist ein Paradebeispiel der Nützlichkeit einer möglichst einfachen axiomatischen Charakterisierung.

### Beweis des Transpositionssatzes

Sei  $A \in K^{n \times n}$ . Dann gilt

$$\det A^t = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{1, \sigma(1)} \cdots a_{n, \sigma(n)} \stackrel{(1)}{=} \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{\sigma^{-1}(1), 1} \cdots a_{\sigma^{-1}(n), n} \stackrel{(2)}{=}$$

$$\sum_{\sigma \in S_n} \operatorname{sgn}(\sigma^{-1}) a_{\sigma^{-1}(1), 1} \cdots a_{\sigma^{-1}(n), n} \stackrel{(3)}{=} \sum_{\pi \in S_n} \operatorname{sgn}(\pi) a_{\pi(1), 1} \cdots a_{\pi(n), n} = \det A.$$

Dabei haben wir verwendet: (1)  $a_{1, \sigma(1)} \cdots a_{n, \sigma(n)}$  und  $a_{\sigma^{-1}(1), 1} \cdots a_{\sigma^{-1}(n), n}$  besitzen dieselben Faktoren, (2)  $\operatorname{sgn}(\sigma^{-1}) = \operatorname{sgn}(\sigma)^{-1}$  für alle  $\sigma \in S_n$ , (3)  $\sigma^{-1}$  durchläuft die Gruppe  $S_n$  bijektiv, wenn dies für  $\sigma$  der Fall ist.

Mit Hilfe von Elementarmatrizen können wir das Ergebnis auch anders gewinnen:

**Alternativer Beweis des Transpositionssatzes**

Für alle  $\lambda \in K$  und  $i, j$  mit  $i \neq j$  gilt:

$$\det(W_{ij}(\lambda)) = \det(W_{ij}(\lambda)^t) = 1.$$

Ist  $A \in K^{n \times n}$ , so gibt es Additionstypen  $L_1, \dots, L_k$  und eine Dreiecksmatrix  $B$  mit

$$A L_1 \dots L_k = B.$$

Dann ist

$$L_k^t \dots L_1^t A^t = B^t.$$

Da die Diagonalen der Dreiecksmatrizen  $B$  und  $B^t$  übereinstimmen, gilt  $\det B = \det B^t$ . Da alle  $L_i$  und  $L_i^t$  die Determinante 1 haben, liefert der Multiplikationssatz

$$\det A = \det B = \det B^t = \det A^t.$$

$$W_{ij}(\lambda) = \begin{pmatrix} 1 & & & \\ & 1 & & \lambda \\ & & \dots & \\ & & & 1 \\ & & & & 1 \end{pmatrix}$$

$$W_{ji}(\lambda) = \begin{pmatrix} 1 & & & \\ & 1 & & \\ & & \dots & \\ & \lambda & & 1 \\ & & & & 1 \end{pmatrix}$$

$$\det(W_{ij}(\lambda)) = 1 = \det(W_{ji}(\lambda))$$

$$W_{ij}(\lambda) = W_{ji}(\lambda)^t$$

Da sich beim Transponieren Spalten und Zeilen austauschen, ergibt sich:

*Die für Spalten formulierten Determinantenaxiome und die daraus abgeleiteten Spaltenregeln gelten analog auch für Zeilen.*

Die Determinantenfunktion ist also auch in den Zeilen multilinear und alternierend. Damit bleibt die Determinante bei Addition des  $\lambda$ -Fachen einer Zeile zu einer anderen unverändert, wechselt beim Tausch von zwei Zeilen das Vorzeichen und skaliert mit  $\lambda$ , wenn eine Zeile mit  $\lambda$  multipliziert wird.

**Beispiel**

Für eine Matrix des  $K^{3 \times 3}$  mit den Zeilen  $a, b, c$  gilt

$$\det \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \det \begin{pmatrix} a^t & b^t & c^t \end{pmatrix} = -\det \begin{pmatrix} c^t & b^t & a^t \end{pmatrix} = -\det \begin{pmatrix} c \\ b \\ a \end{pmatrix}.$$

**Spaltenaxiome oder Zeilenaxiome?**

Oft werden die Determinantenaxiome auch als Zeileneigenschaften formuliert und die Spalteneigenschaften gefolgert. Bei der axiomatischen Bevorzugung der Zeilen steht der Zusammenhang mit linearen Gleichungssystemen im Vordergrund, bei der Bevorzugung der Spalten die natürliche Übersetzung der Multilinearität einer Abbildung  $f: V^n \rightarrow W$  in die Sprache der Matrizen (mit  $V = K^n = K^{n \times 1}$ ,  $V^n = K^{n \times n}$  ist  $f: K^{n \times n} \rightarrow K$ ). Letztendlich gilt: Beide Zugänge liefern dieselbe Determinantenfunktion und sind damit äquivalent.

## 7.6 Der Entwicklungssatz von Laplace

### Satz (Spalten- und Zeilenentwicklung)

Seien  $K$  ein Körper und  $n \geq 2$ .  
 Für alle  $A \in K^{n \times n}$  und  $1 \leq i, j \leq n$   
 sei  $A'_{ij} \in K^{(n-1) \times (n-1)}$  die Matrix,  
 die aus  $A$  durch Streichen der  $i$ -  
 ten Zeile und  
 $j$ -ten Spalte entsteht. Dann gilt  
 für alle Matrizen  $A \in K^{n \times n}$  und  
 alle Spaltenindizes  $1 \leq j \leq n$

$$A'_{ij} = \begin{pmatrix} a_{11} & \dots & a_{1j} & \dots & a_{1n} \\ \dots & \dots & \dots & \dots & \dots \\ \hline a_{i1} & \dots & a_{ij} & \dots & a_{in} \\ \hline \dots & \dots & \dots & \dots & \dots \\ a_{n1} & \dots & a_{nj} & \dots & a_{nn} \end{pmatrix}$$

$$\det A = \sum_{1 \leq i \leq n} (-1)^{i+j} a_{ij} \det A'_{ij}. \quad (\text{Entwicklung nach der } j\text{-ten Spalte})$$

Analog gilt für alle Zeilenindizes  $1 \leq i \leq n$

$$\det A = \sum_{1 \leq j \leq n} (-1)^{i+j} a_{ij} \det A'_{ij}. \quad (\text{Entwicklung nach der } i\text{-ten Zeile})$$

Der Entwicklungssatz stellt eine weitere Möglichkeit der Berechnung von Determinanten dar. Besonders geeignet ist er für Matrizen, die eine Zeile oder Spalte mit vielen Nulleinträgen besitzen.

### Beweis des Entwicklungssatzes

Wesentliches Hilfsmittel sind die  $n \times n$ -Matrizen

$$A_{ij} = \begin{pmatrix} a_{11} & \dots & 0 & \dots & a_{1n} \\ \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ a_{n1} & \dots & 0 & \dots & a_{nn} \end{pmatrix} \in K^{n \times n},$$

bei denen die  $i$ -te Zeile von  $A$  mit  $e_j$  und die  $j$ -te Spalte von  $A$  mit  $e_i$  überschrieben ist. Die Determinanten der Matrizen  $A_{ij}$  und  $A'_{ij}$  stimmen bis auf ein von der Stelle  $(i, j)$  abhängiges Vorzeichen überein: Es gilt

$$\det A_{ij} = \det (a_1 \dots e_i \dots a_n) = (-1)^{i-1+j-1} \det \begin{pmatrix} 1 & 0 \\ 0 & A'_{ij} \end{pmatrix} = (-1)^{i+j} \det A'_{ij},$$

wobei wir im zweiten Schritt eine  $(i-1)$ -malige Zeilen- und eine  $(j-1)$ -malige Spaltenvertauschung durchführen. Ist nun  $j$  festgewählt, so gilt

$$\det A = (a_1; \dots; \sum_i a_{ij} e_i; \dots; a_n) = \sum_i a_{ij} \det A_{ij} = \sum_i (-1)^{i+j} a_{ij} \det A'_{ij}.$$

Die Zeilenentwicklung zeigt man analog.

Die im Entwicklungssatz von Laplace auftauchenden Vorzeichen  $(-1)^{i+j}$  haben eine schachbrettartige Verteilung (vgl. das Diagramm rechts).

Die Spalten- oder Zeilenentwicklung kann mehrfach hintereinander durchgeführt werden. Die Beispiele (3) und (4) illustrieren dieses Vorgehen.

+	-	+	-	...
-	+	-	+	...
+	-	+	-	...
-	+	-	+	...
...	...	...	...	...

### Beispiele

- (1) Entwickeln wir  $A \in K^{2 \times 2}$  nach der ersten Spalte, so erhalten wir

$$\det A = a_{11} \det A_{11}' - a_{21} \det A_{21}' = a_{11} a_{22} - a_{21} a_{12}.$$

- (2) Entwickeln wir  $A \in K^{3 \times 3}$  nach der ersten Zeile, so erhalten wir

$$\begin{aligned} \det A &= a_{11} \det A_{11}' - a_{12} \det A_{12}' + a_{13} \det A_{13}' = \\ &= a_{11} \det \begin{pmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{pmatrix} - a_{12} \det \begin{pmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{pmatrix} + a_{13} \det \begin{pmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix} = \\ &= a_{11} a_{22} a_{33} - a_{11} a_{23} a_{32} - a_{12} a_{21} a_{33} + a_{12} a_{23} a_{31} + a_{13} a_{21} a_{32} - a_{13} a_{22} a_{31}, \end{aligned}$$

also erneut die Regel von Sarrus (vgl. 7.4).

- (3) Zweimaliges Entwickeln nach der zweiten Zeile liefert

$$\det \begin{pmatrix} 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & -1 \end{pmatrix} = \det \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & -1 \end{pmatrix} = \det \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} = -2.$$

- (4) Entwickeln nach der dritten und dann nach der zweiten Spalte ergibt

$$\begin{aligned} \det \begin{pmatrix} 1 & 2 & 0 & 3 \\ 4 & 5 & 1 & 7 \\ 1 & -2 & 0 & 1 \\ 2 & 0 & 0 & 4 \end{pmatrix} &= -\det \begin{pmatrix} 1 & 2 & 3 \\ 1 & -2 & 1 \\ 2 & 0 & 4 \end{pmatrix} = \\ &= 2 \det \begin{pmatrix} 1 & 1 \\ 2 & 4 \end{pmatrix} + 2 \det \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} = 2 \cdot 2 + 2 \cdot (-2) = 0. \end{aligned}$$

## 7.7 Komplementärmatrizen und die Regel von Cramer

### Definition (Komplementärmatrix)

Seien  $n \geq 1$ ,  $K$  ein Körper und  $A \in K^{n \times n}$ . Dann definieren wir die zu  $A$  *komplementäre Matrix* oder die *Adjunkte* von  $A$  als die Matrix  $A^\# \in K^{n \times n}$  mit

$$a_{ij}^\# = \det A_{ji} \quad \text{für alle } i, j.$$

Die Matrix  $A^\#$  entsteht aus  $A$  durch Ersetzen aller Einträge  $a_{ij}$  durch die Determinanten der Matrizen  $A_{ji}$  und anschließendes Transponieren. Nach den Ergebnissen aus 7.6 gilt für alle  $n \geq 2$ :

$$a_{ij}^\# = (-1)^{i+j} \det A'_{ji}. \quad A^\# = \begin{pmatrix} \det A_{11}' & -\det A_{21}' & \det A_{31}' & \dots \\ -\det A_{12}' & \det A_{22}' & -\det A_{32}' & \dots \\ \det A_{13}' & -\det A_{23}' & \det A_{33}' & \dots \\ \dots & \dots & \dots & \dots \end{pmatrix}$$

Den Entwicklungssatz von Laplace können wir nun so schreiben:

$$\det A = \sum_{1 \leq i \leq n} a_{ij} a_{ji}^\# = (A^\# A)(j, j) \quad \text{für alle } j,$$

$$\det A = \sum_{1 \leq j \leq n} a_{ij} a_{ji}^\# = (A A^\#)(i, i) \quad \text{für alle } i.$$

Die Diagonalen von  $A^\# A$  und  $A A^\#$  sind also konstant gleich  $\det(A)$ . Allgemein gilt

$$(A^\# A)(i, j) = \sum_k a_{ik}^\# a_{kj} = \sum_k a_{kj} \det(A_{ki}) =$$

$$\sum_k a_{kj} \det(a_1 \dots e_k \dots a_n) = \det(a_1 \dots a_j \dots a_n) = \delta_{ij} \det A \quad \text{für alle } i, j,$$

wobei  $a_1, \dots, a_n$  die Spalten von  $A$  sind und  $e_k$  und  $a_j$  in der  $i$ -ten Spalte stehen. Analoges gilt für  $A A^\#$ . Damit haben wir:

$$A^\# A = \det(A) E_n = A A^\# \quad \text{für alle } A \in K^{n \times n}.$$

### Beispiele

- (1) Für alle  $A \in K^{2 \times 2}$  berechnet sich die komplementäre Matrix zu

$$A^\# = \begin{pmatrix} \det A_{11} & \det A_{21} \\ \det A_{12} & \det A_{22} \end{pmatrix} = \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}.$$

- (2) Für  $A = E_n$  gilt  $\det A_{ij} = \delta_{ij}$  für alle  $i, j$ . Also ist  $E_n^\# = E_n$ .

- (3) Für  $A = \text{diag}(d_1, \dots, d_n)$  ist  $A^\# = \text{diag}(a_{11}^\#, \dots, a_{nn}^\#)$  mit  $a_{ii}^\# = \prod_{j \neq i} a_{jj}$ .

- (4) Für alle  $A \in K^{n \times n}$  gilt  $\det(A^\#) \det(A) = \det(A^\# A) = \det(\det(A) E_n) = \det(A)^n$ .  
Für invertierbare  $A$  ist also  $\det(A^\#) = \det(A)^{n-1}$ .

Die Diagonale von  $A^\# A$  und  $AA^\#$  liefert den Entwicklungssatz von Laplace. Die Kenntnis des gesamten Produkts erlaubt die Berechnung von  $A^{-1}$  mit Hilfe der Komplementärmatrix. Multiplizieren wir nämlich  $\det(A) E_n = A^\# A$  von rechts mit  $A^{-1}$ , so erhalten wir:

$$A^{-1} = \frac{A^\#}{\det A} \quad \text{für alle } A \in GL(n, K).$$

Eine klassische Anwendung ist:

### Die Regel von Cramer

Seien  $A \in GL(n, K)$  und  $b \in K^n$ . Für alle  $1 \leq j \leq n$  sei  $A_j \in K^{n \times n}$  die Matrix, die aus  $A$  entsteht, wenn die  $j$ -te Spalte von  $A$  durch  $b$  ersetzt wird. Dann ist  $(x_1, \dots, x_n) \in K^n$  mit

$$x_j = \frac{\det A_j}{\det A} \quad \text{für alle } j$$

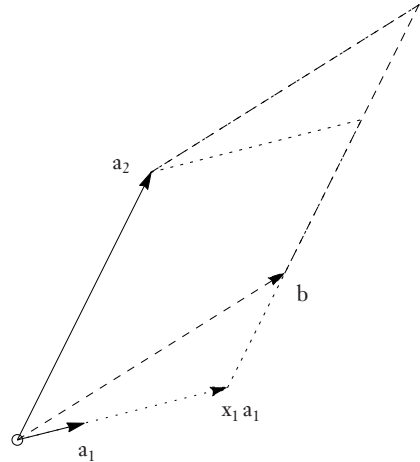
die eindeutige Lösung des linearen Gleichungssystems  $Ax = b$ .

Schreiben wir nämlich die Lösung des Systems in der Form  $A^{-1}b$ , so gilt für alle  $j$  (vgl. die Berechnung von  $A^\# A(i, j)$ )

$$\det(A) (A^{-1}b)_j = (A^\# b)_j =$$

$$\sum_i b_i \det(A_{ij}) = \det \begin{pmatrix} a_1 & \dots & b & \dots & a_n \end{pmatrix}$$

mit  $b$  in der  $j$ -ten Spalte. Dies zeigt die Regel von Cramer.



Ist  $A \in GL(2, \mathbb{R})$  und  $Ax = b$ , so gilt nach der Regel von Cramer

$$\det \begin{pmatrix} x_1 a_1 & a_2 \end{pmatrix} = x_1 \det A = \det \begin{pmatrix} b & a_2 \end{pmatrix}.$$

Dies bedeutet, dass die beiden von  $x_1 a_1$ ,  $a_2$  und  $b$ ,  $a_2$  aufgespannten Parallelogramme denselben Flächeninhalt haben. Analoges gilt für  $x_2$ .

### Beispiel

Für  $A \in GL(2, K)$  benötigt die Regel von Cramer die Determinanten der Matrizen

$$A_1 = \begin{pmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{pmatrix}, \quad A_2 = \begin{pmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{pmatrix}, \quad A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}.$$

Für das lineare Gleichungssystem

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 5 \\ 6 \end{pmatrix}$$

gilt  $\det A_1 = 8$ ,  $\det A_2 = -9$ ,  $\det A = -2$ . Damit ist  $(-4, 9/2)$  die Lösung des Systems.

## 7.8 Die speziellen linearen Gruppen

### Definition (die Gruppen $SL(n, K)$ )

Für einen Körper  $K$  und  $n \geq 1$  sei  $SL(n, K) = \{ A \in GL(n, K) \mid \det A = 1 \}$ .  
(spezielle lineare Gruppe)

Die Menge  $SL(n, K)$  ist als Kern von  
 $\det : GL(n, K) \rightarrow K^*$

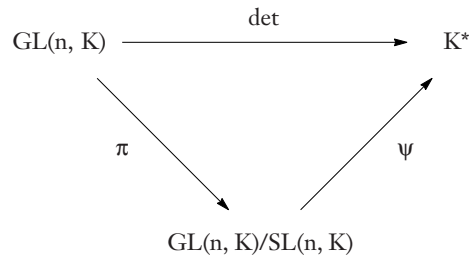
ein Normalteiler von  $GL(n, K)$ . Wegen

$\det(\text{diag}(a, 1, \dots, 1)) = a$  für alle  $a \in K$

ist  $\det : GL(n, K) \rightarrow K^*$  ein Epimorphismus. Nach dem Homomorphiesatz sind  $GL(n, K)/SL(n, K)$  und  $K^*$  isomorph.

Die Matrizen  $\text{diag}(a, 1, \dots, 1)$ ,  $a \in K^*$ , bilden ein vollständiges Repräsentantensystem.

Im Fall  $n = 1$  gilt  $SL(1, K) = \{ 1 \}$ . Für alle  $n \geq 2$  gilt die folgende Charakterisierung:



### Die elementare lineare Gruppe

Ist  $K$  ein Körper und  $n \geq 2$ , so ist die *elementare lineare Gruppe* definiert durch

$E(n, K) = \{ A \in GL(n, K) \mid A \text{ ist ein Produkt von Additionstypen } W_{ij}(\lambda), i \neq j, \lambda \in K \}$ .

Wegen  $W_{ij}(\lambda)^{-1} = W_{ij}(-\lambda)$  für  $i \neq j$  ist  $E(n, K)$  in der Tat eine Gruppe. Es gilt:

$E(n, K) = SL(n, K)$ .

Die Inklusion „ $\subseteq$ “ folgt aus dem Multiplikationssatz und  $\det(W_{ij}(\lambda)) = 1$  für  $i \neq j$ .

Für „ $\supseteq$ “ verwenden wir, dass sich jedes  $A \in GL(n, K)$  durch Multiplikation mit Additionstypen in eine Diagonalmatrix überführen lässt. Die Determinante bleibt dabei unverändert. Damit bleibt zu zeigen, dass jede Diagonalmatrix der Form

$\text{diag}(d_1, \dots, d_n), d_1 \dots d_n = 1$

ein Element von  $E(n, K)$  ist. Dies verifiziert man für  $n = 2$  direkt. Die allgemeine Behauptung kann man nun induktiv mit Hilfe folgender Faktorisierung zeigen:

$\text{diag}(d_1, \dots, d_n) = \text{diag}(d_1, d_2 \dots d_n, 1, \dots, 1) \text{diag}(1, (d_3 \dots d_n)^{-1}, d_3, \dots, d_n)$ .

Wir betrachten nun noch den Spezialfall  $K = \mathbb{R}$  und die Determinanten orthogonaler Matrizen. Wir setzen für alle  $n \geq 1$ :

$O(n) = \{ Q \in \mathbb{R}^{n \times n} \mid Q \text{ ist orthogonal} \}$ , (orthogonale Gruppe)

$SO(n) = O(n) \cap SL(n, \mathbb{R})$ . (spezielle orthogonale Gruppe)

Da die Orthogonalität durch  $Q^{-1} = Q^t$  charakterisiert ist, haben alle orthogonalen Matrizen die Determinante  $\pm 1$ . Für alle Matrizen  $W$  mit  $\det(W) = -1$  gilt

$O(n) = SO(n) \cup \{ QW \mid Q \in SO(n) \}$ .



Die Gruppen  $O(n)$  und  $SO(n)$  lassen sich für  $n \leq 3$  anschaulich beschreiben:

$n$	$Q \in O(n)$	$f_Q : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ist ...
1	$\det Q = 1$	die Identität
1	$\det Q = -1$	die Spiegelung am Nullpunkt
2	$\det Q = 1$	die Drehung um einen Winkel $\alpha$
2	$\det Q = -1$	die Spiegelung an einer Geraden durch 0
3	$\det Q = 1$	die Rotation um eine Achse durch 0
3	$\det Q = -1$	eine Rotationsspiegelung, d. h., die Rotation um eine Achse durch 0 plus eine (vor oder nach der Rotation durchgeführte) Spiegelung an der zur Rotationsachse senkrechten Ebene

	$\det Q = 1$	$\det Q = -1$
$n = 1$	$\begin{pmatrix} 1 \end{pmatrix}$	$\begin{pmatrix} -1 \end{pmatrix}$
$n = 2$	$\begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}$	$\begin{pmatrix} \cos \alpha & \sin \alpha \\ \sin \alpha & -\cos \alpha \end{pmatrix}$
$n = 3$	$\begin{pmatrix} \cos \alpha & -\sin \alpha & 0 \\ \sin \alpha & \cos \alpha & 0 \\ 0 & 0 & 1 \end{pmatrix}$	$\begin{pmatrix} \cos \alpha & -\sin \alpha & 0 \\ \sin \alpha & \cos \alpha & 0 \\ 0 & 0 & -1 \end{pmatrix}$

Die Tabelle zeigt die Form aller Elemente von  $O(1)$  und  $O(2)$  sowie wichtige Elemente von  $O(3)$ . Die Matrix für  $n = 2$  und  $\det Q = -1$  ist eine Drehung um  $-\alpha$  gefolgt von einer Spiegelung an der  $x$ -Achse, also eine Spiegelung an der Geraden durch 0 mit dem Winkel  $\alpha/2$ . Die Matrix für  $n = 3$  und  $\det Q = -1$  ist eine Rotation um die  $z$ -Achse plus eine Spiegelung an der  $x$ - $y$ -Ebene.

Viele andere Klassifikationen sind möglich. Für alle  $n \geq 1$  und alle  $Q \in O(n)$  gilt zum Beispiel, dass  $f_Q : \mathbb{R}^n \rightarrow \mathbb{R}^n$  eine Komposition von höchstens  $n$  Spiegelungen an Unterräumen des  $\mathbb{R}^n$  der Dimension  $n - 1$  ist (Hyperebenen im  $\mathbb{R}^n$ ).

Für den Körper  $K = \mathbb{C}$  definieren wir analog:

$$U(n) = \{ U \in \mathbb{C}^{n \times n} \mid U \text{ ist unitär} \}, \quad (\text{unitäre Gruppe})$$

$$SU(n) = U(n) \cap SL(n, \mathbb{C}). \quad (\text{spezielle unitäre Gruppe})$$

Unitäre Matrizen sind durch  $U^{-1} = U^*$  charakterisiert. Wegen  $\det(U^*) = \overline{\det(U)}$  gilt also  $|\det(U)| = 1$  für alle  $U \in U(n)$ . Die Determinante einer unitären Matrix hat damit die Form  $e^{i\varphi}$ . Durch Multiplikation mit einer Matrix  $W$  mit  $\det(W) = e^{-i\varphi}$  erhält man eine Matrix  $UW$  in  $SU(n)$ . Ist  $W_\varphi = W_{11}(e^{i\varphi})$ , so gilt

$$U(n) = \{ UW_\varphi \mid U \in SU(n), \varphi \in [0, 2\pi[ \}.$$

## 7.9 Volumina von Parallelotopen

### Satz (geometrische Bedeutung der Determinante)

Seien  $1 \leq r \leq n$ ,  $a_1, \dots, a_r \in \mathbb{R}^n$  und  $A = \begin{pmatrix} a_1 & \dots & a_r \end{pmatrix} \in \mathbb{R}^{n \times r}$ . Weiter sei

$$P = P(a_1, \dots, a_r) = \{x \in \mathbb{R}^n \mid \text{es gibt } 0 \leq \lambda_1, \dots, \lambda_r \leq 1 \text{ mit } x = \sum_i \lambda_i a_i\}$$

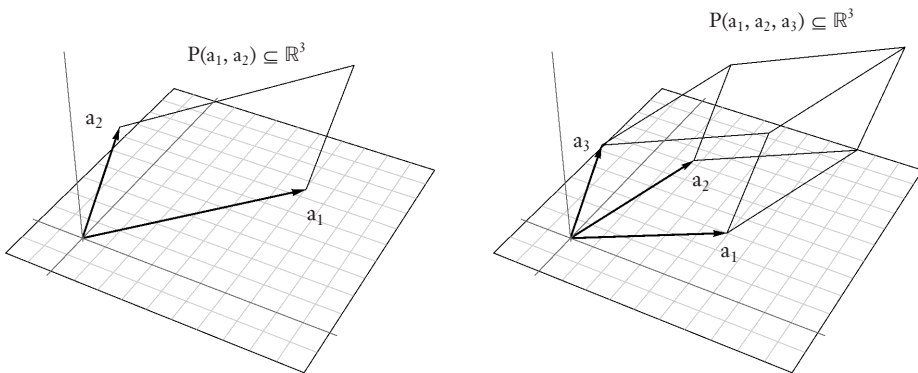
das von  $a_1, \dots, a_r$  aufgespannte Parallelotop der Dimension  $\dim(\text{span}(a_1, \dots, a_r)) \leq r$ .

Dann gilt für die gramsche Matrix  $A^t A \in \mathbb{R}^{r \times r}$

$$(+)\ \text{vol}_r(P)^2 = \det(A^t A) = \det \begin{pmatrix} \langle a_1, a_1 \rangle & \langle a_1, a_2 \rangle & \dots & \langle a_1, a_r \rangle \\ \langle a_2, a_1 \rangle & \langle a_2, a_2 \rangle & \dots & \langle a_2, a_r \rangle \\ \dots & \dots & \dots & \dots \\ \langle a_r, a_1 \rangle & \langle a_r, a_2 \rangle & \dots & \langle a_r, a_r \rangle \end{pmatrix} \geq 0,$$

wobei  $\text{vol}_r(\cdot)$  das  $r$ -dimensionale Volumen im  $\mathbb{R}^n$  bezeichnet und in der Determinante das kanonische Skalarprodukt des  $\mathbb{R}^n$  verwendet wird. Insbesondere gilt für  $r = n$

$$(++)\ \text{vol}_n(P) = |\det A|.$$



Das Ergebnis setzt voraus, dass  $\text{vol}_r(P(a_1, \dots, a_r))$  erklärt ist. Ohne Anleihe bei der Maß- und Integrationstheorie kann dies auf folgende Art geschehen:

### Rekursive Definition des Volumens von Parallelotopen

Sei  $n \geq 1$ . Für  $r = 1$  sei  $\text{vol}_1(P(a_1)) = |a_1|$ . Rekursiv definieren wir nun in Verallgemeinerung von „Grundseite mal Höhe“ und „Grundfläche mal Höhe“:

$$\text{vol}_{r+1}(P(a_1, \dots, a_{r+1})) = \text{vol}_r(P(a_1, \dots, a_r)) \cdot h, \text{ wobei}$$

$$h = \|a_{r+1} - \text{pr}_U(a_{r+1})\| \text{ mit } U = \text{span}(a_1, \dots, a_r).$$

Es gilt  $\text{vol}_r(P(a_1, \dots, a_r)) = 0$  genau dann, wenn  $(a_1, \dots, a_r)$  linear abhängig ist. Genau in diesem Fall ist auch die gramsche Determinante  $\det(A^t A)$  gleich 0. Allgemein zeigt man (+) durch Induktion nach  $r \leq n$ . Aus (+) folgt nun, dass  $\text{vol}_r(P(a_1, \dots, a_r))$  nur von der Menge  $P$  und nicht von der Reihenfolge der  $a_i$  abhängt. Im Fall  $r = n$  gilt  $A^t, A \in \mathbb{R}^{n \times n}$  und  $\det(A^t A) = \det(A^t) \det(A) = \det(A)^2$ , woraus sich (++) ergibt.

**Beispiele**

- (1) Für das von den Vektoren  $a_1 = (1, 1, 1)$  und  $a_2 = (2, 1, -1)$  des  $\mathbb{R}^3$  aufgespannte Parallelogramm  $P \subseteq \mathbb{R}^3$  gilt

$$\text{vol}_2(P)^2 = \det \begin{pmatrix} \langle a_1, a_1 \rangle & \langle a_1, a_2 \rangle \\ \langle a_2, a_1 \rangle & \langle a_2, a_2 \rangle \end{pmatrix} = \det \begin{pmatrix} 3 & 2 \\ 2 & 6 \end{pmatrix} = 14.$$

Damit hat  $P$  den Flächeninhalt  $\sqrt{14}$ .

- (2) Sei  $P \subseteq \mathbb{R}^3$  das von  $a_1 = (1, 1, 1)$ ,  $a_2 = (2, 1, -1)$  und  $a_3 = (1, 0, -1)$  aufgespannte Parallelepip. Wegen

$$\det \begin{pmatrix} 1 & 2 & 1 \\ 1 & 1 & 0 \\ 1 & -1 & -1 \end{pmatrix} = \det \begin{pmatrix} 1 & 2 & 1 \\ 0 & -1 & -1 \\ 0 & -3 & -2 \end{pmatrix} = \det \begin{pmatrix} 1 & 2 & 1 \\ 0 & -1 & -1 \\ 0 & 0 & 1 \end{pmatrix} = -1$$

gilt  $\text{vol}_3(P) = 1$ .

Die Volumenformel (++) lässt sich auch direkt mit Hilfe der Determinantenaxiome sehr anschaulich erklären (und umgekehrt lassen sich die Axiome dadurch motivieren). Ist  $n = 2$ , so gilt für alle  $a_1, a_2, a'_1, a'_2 \in \mathbb{R}^2$  und  $\lambda \in \mathbb{R}$  mit  $\text{vol} = \text{vol}_2$ :

$$(1) \text{vol}(P(\lambda a_1, a_2)) = |\lambda| \text{vol}(P(a_1, a_2)) = \text{vol}(P(a_1, \lambda a_2)), \quad (\text{Streckung})$$

$$\text{vol}(P(a_1 + a'_1, a_2)) = \text{vol}(P(a_1, a_2)) + \text{vol}(P(a'_1, a_2)),$$

$$\text{vol}(P(a_1, a_2 + a'_2)) = \text{vol}(P(a_1, a_2)) + \text{vol}(P(a_1, a'_2)), \quad (\text{Additivität})$$

$$(2) \text{vol}(P(a_1, a_1)) = 0, \quad (\text{degenerierter Fall})$$

$$(3) \text{vol}(P(e_1, e_2)) = 1. \quad (\text{Normierung})$$

Bis auf den Betrag bei  $\lambda$  entsprechen diese Eigenschaften genau den Determinantenaxiomen. Analoge Überlegungen gelten für andere Dimensionen.

Im Unterschied zum Volumen ist die Determinante vorzeichenbehaftet. Anders als das Volumen ändert sie ihr Vorzeichen, wenn zwei aufspannende Vektoren vertauscht werden. Die Determinante  $\det(A)$  enthält damit auch eine Information über die Orientierung von  $P$ .

**Die Volumenveränderung einer linearen Abbildung**

Ist  $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$  linear, so ist das Bild des Einheitswürfels  $P = P(e_1, \dots, e_n)$  unter  $f$  das Parallelotop  $P_f = P(f(e_1), \dots, f(e_n))$ . Das Volumen von  $P_f$  ist der Betrag der Determinante der  $f$  darstellenden Matrix  $A$  (bzgl. der Standardbasis). Wegen

$$\text{vol}_n(P) = 1, \quad |\det(A)| = \text{vol}_n(P_f)$$

können wir also  $|\det(A)|$  als Maß für die durch die lineare Abbildung bewirkte Volumenveränderung ansehen.

## 7.10 Das Kreuzprodukt

### Definition (Kreuzprodukt)

Seien  $a, b \in \mathbb{R}^3$ . Dann gibt es nach dem Rieszschen Darstellungssatz genau einen Vektor  $w \in \mathbb{R}^3$ , der das lineare Funktional

$$\det \begin{pmatrix} a & b & \cdot \end{pmatrix} : \mathbb{R}^3 \rightarrow \mathbb{R}$$

darstellt. Wir schreiben  $w = a \times b$  und nennen  $w$  das *Kreuzprodukt* von  $a$  und  $b$ .

Nach Definition gilt also (mit dem kanonischen Skalarprodukt)

$$\langle a \times b, c \rangle = \det \begin{pmatrix} a & b & c \end{pmatrix} \quad \text{für alle } a, b, c \in \mathbb{R}^3. \text{ Für } c = e_1, e_2, e_3 \text{ erhalten wir}$$

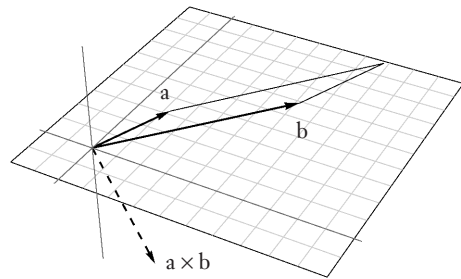
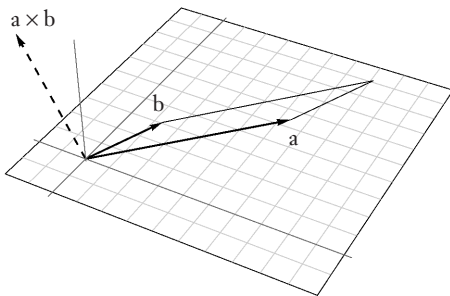
$$(a \times b)_1 = \langle a \times b, e_1 \rangle = \det \begin{pmatrix} a & b & e_1 \end{pmatrix} = \det \begin{pmatrix} a_2 & b_2 \\ a_3 & b_3 \end{pmatrix},$$

$$(a \times b)_2 = \langle a \times b, e_2 \rangle = \det \begin{pmatrix} a & b & e_2 \end{pmatrix} = -\det \begin{pmatrix} a_1 & b_1 \\ a_3 & b_3 \end{pmatrix},$$

$$(a \times b)_3 = \langle a \times b, e_3 \rangle = \det \begin{pmatrix} a & b & e_3 \end{pmatrix} = \det \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \end{pmatrix},$$

sodass

$$a \times b = \begin{pmatrix} a_2 b_3 - a_3 b_2 \\ a_3 b_1 - a_1 b_3 \\ a_1 b_2 - a_2 b_1 \end{pmatrix}.$$



Der Vektor  $a \times b$  steht senkrecht auf  $a$  und  $b$  und hat die Länge des Flächeninhalts des von  $a$  und  $b$  aufgespannten Parallelogramms. Die Richtung von  $a \times b$  kann mit der Rechten-Hand-Regel (Drei-Finger-Regel) ermittelt werden:  $a$  entspricht dem Daumen,  $b$  dem Zeigefinger und  $a \times b$  dem Mittelfinger der rechten Hand.

### Beispiel

Für alle  $b \in \mathbb{R}^3$  gilt

$$e_1 \times b = \begin{pmatrix} 0 \\ -b_3 \\ b_2 \end{pmatrix}, \quad e_2 \times b = \begin{pmatrix} b_3 \\ 0 \\ -b_1 \end{pmatrix}, \quad e_3 \times b = \begin{pmatrix} -b_2 \\ b_1 \\ 0 \end{pmatrix}.$$

Inbesondere ist  $e_1 \times e_2 = e_3$ ,  $e_1 \times e_3 = -e_2$ ,  $e_2 \times e_3 = e_1$ . Das Kreuzprodukt ist nicht assoziativ, da zum Beispiel  $e_1 \times (e_1 \times e_2) = -e_2 \neq 0 = (e_1 \times e_1) \times e_2$ .

<i>Eigenschaften des Kreuzprodukts</i>	
$\langle a \times b, a \rangle = 0, \quad \langle a \times b, b \rangle = 0$	<i>Orthogonalität</i>
$\text{vol}_2(P(a, b)) = \ a \times b\ , \quad \text{vol}_3(P(a, b, c)) =  \langle a \times b, c \rangle $	<i>Volumenformeln</i>
$\cos \alpha = \frac{\langle a, b \rangle}{\ a\  \ b\ }, \quad \sin \alpha = \frac{\ a \times b\ }{\ a\  \ b\ }$	<i>Winkel</i>
$Aa \times Ab = \det(A) (A^t)^{-1} (a \times b)$	<i>Transformation</i>
$Qa \times Qb = Q(a \times b)$	<i>Rotation</i>
$a \times b = -b \times a$	<i>Antikommutativität</i>
$a \times (b \times c) = b \langle a, c \rangle - c \langle a, b \rangle$	<i>bac-minus-cab-Regel</i>
$(\lambda a + b) \times c = \lambda (a \times c) + b \times c$ $a \times (\lambda b + c) = \lambda (a \times b) + a \times c$	<i>Bilinearität</i>
$u \times u = 0$	<i>Alternation</i>
$a \times (b \times c) + b \times (c \times a) + c \times (a \times b) = 0$	<i>Jacobi-Identität</i>

In dieser Tabelle sind  $a, b, c \in \mathbb{R}^3$ ,  $A \in GL(3, \mathbb{R})$ ,  $Q \in SO(3)$  beliebig, wobei für den von  $a$  und  $b$  eingeschlossenen Winkel  $\alpha$  vorausgesetzt wird, dass  $a, b \neq 0$ . Weiter ist  $P(a, b)$  das von  $a, b$  aufgespannte Parallelogramm und  $P(a, b, c)$  das von  $a, b, c$  aufgespannte Parallelepipet.

Die Transformation lässt sich elegant so zeigen: Für alle  $a, b, c$  gilt

$$\begin{aligned} \langle Aa \times Ab, c \rangle &= \det \begin{pmatrix} Aa & Ab & c \end{pmatrix} = \det \left( A \cdot \begin{pmatrix} a & b & A^{-1}c \end{pmatrix} \right) = \\ &= \det A \cdot \det \begin{pmatrix} a & b & A^{-1}c \end{pmatrix} = \det A \cdot \langle a \times b, A^{-1}c \rangle = \det A \cdot \langle (A^t)^{-1}(a \times b), c \rangle. \end{aligned}$$

Die Rotation ergibt sich nun aus  $\det(Q) = 1$  und  $(Q^t)^{-1} = Q$  für  $Q \in SO(3)$ .

### Das verallgemeinerte $(n - 1)$ -stellige Kreuzprodukt im $\mathbb{R}^n$

Mit Hilfe des Rieszschen Darstellungssatzes kann für jede Dimension  $n \geq 2$  ein Kreuzprodukt  $a_1 \times \dots \times a_{n-1} \in \mathbb{R}^n$  erklärt werden durch

$$\langle a_1 \times \dots \times a_{n-1}, a \rangle = \det \begin{pmatrix} a_1 & \dots & a_{n-1} & a \end{pmatrix} \quad \text{für alle } a_1, \dots, a_{n-1}, a \in \mathbb{R}^n.$$

Es gilt zum Beispiel die Orthogonalität  $a_1 \times \dots \times a_{n-1} \in \text{span}(a_1, \dots, a_{n-1})^\perp$  und die Volumenformel  $\text{vol}_n(P(a_1, \dots, a_n)) = |\langle a_1 \times \dots \times a_{n-1}, a_n \rangle|$ .

## 7.11 Positive Definitheit

### Satz (Charakterisierungen der positiven Definitheit)

Seien  $n \geq 1$  und  $A \in \mathbb{C}^{n \times n}$  eine hermitesche Matrix. Dann sind äquivalent:

- (a)  $A$  ist positiv definit, d. h., für das kanonische Skalarprodukt auf dem  $\mathbb{C}^n$  gilt
 
$$\langle z, Az \rangle > 0 \quad \text{für alle } z \in \mathbb{C}^n - \{0\}.$$
- (b) Für die Matrizen  $A_k = (a_{ij})_{1 \leq i, j \leq k} \in \mathbb{C}^{k \times k}$  gilt
 
$$\det(A_k) > 0 \quad \text{für alle } 1 \leq k \leq n. \quad (\text{Hauptminorenkriterium})$$
- (c)  $A$  lässt sich durch Multiplikation mit Additionstypen  $W_{ij}(\lambda)$ ,  $i > j$ , in eine Dreiecksmatrix  $B$  mit positiven reellen Diagonaleinträgen verwandeln.
- (d) Es gibt eine Dreiecksmatrix  $L \in GL(n, \mathbb{C})$  mit  $A = L^* L$ . (Cholesky-Zerlegung)
- (e) Es gibt ein  $B \in GL(n, \mathbb{C})$  mit  $A = B^* B$ .

Analoges gilt für symmetrische Matrizen  $A \in \mathbb{R}^{n \times n}$ .

Die Zahlen  $\det(A_k)$  heißen die *Hauptminoren* von  $A$ . Nach (b) sind alle  $A_k$  und damit  $A = A_n$  invertierbar, wenn  $A$  positiv definit ist.

Der Satz erlaubt für eine gegebene Hermitesche Matrix  $A \in \mathbb{C}^{n \times n}$  (oder symmetrische Matrix  $A \in \mathbb{R}^{n \times n}$ ) die Beantwortung von:

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} & \dots \\ a_{21} & a_{22} & a_{23} & a_{24} & \dots \\ a_{31} & a_{32} & a_{33} & a_{34} & \dots \\ a_{41} & a_{42} & a_{43} & a_{44} & \dots \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix}$$

### Ist $A$ positiv definit?

Für kleine  $n$  ist das Hauptminorenkriterium geeignet, um die positive Definitheit von  $A$  zu entscheiden. Für größere  $n$  überführen wir  $A$  durch Spaltenadditionen in eine Dreiecksmatrix  $B = A L_1 \dots L_k$ . Dann ist  $A$  genau dann positiv definit, wenn alle Diagonaleinträge  $\lambda_i$  von  $B$  reell und zudem positiv sind. Die Cholesky-Zerlegung  $A = L^* L$  ist im positiv definiten Fall gegeben durch

$$L = (L_1 \dots L_k W_{nn}(\mu_n) \dots W_{11}(\mu_1))^{-1} \quad \text{mit } \mu_i = \sqrt{\lambda_i}.$$

Dass die Abschwächung (e) von (d) die positive Definitheit impliziert, folgt aus

$$\langle x, B^* B x \rangle = \langle B^{**} x, B x \rangle = \langle B x, B x \rangle > 0 \quad \text{für } B \in GL(n, \mathbb{C}) \text{ und } x \neq 0.$$

Für die in 6.12 untersuchten Sesquilinearformen gilt:

### Positiv definite Formen

Seien  $V$  ein  $\mathbb{C}$ -Vektorraum,  $\varphi: V \times V \rightarrow \mathbb{C}$  eine hermitesche Form,  $\mathcal{A} = (v_1, \dots, v_n)$  eine Basis von  $V$  und  $A = (\varphi(v_i, v_j))_{ij} \in \mathbb{C}^{n \times n}$  die gramsche Matrix von  $\varphi$  bzgl.  $\mathcal{A}$ .

Dann ist  $\varphi$  genau dann positiv definit, wenn eine (alle) der Aussagen (a) – (e) gelten.

Analoges gilt für eine symmetrische Form  $\varphi: V \times V \rightarrow \mathbb{R}$  auf einem  $\mathbb{R}$ -Vektorraum.

**Beispiele**

- (1) Ist  $A \in \mathbb{R}^{n \times n}$  symmetrisch und positiv definit, so sind alle Diagonaleinträge von  $A$  positiv, da  $a_{ii} = \langle e_i, A e_i \rangle > 0$  für alle  $1 \leq i \leq n$ . Dass diese Eigenschaft nicht hinreichend ist, zeigt die Matrix  $A \in \mathbb{R}^{2 \times 2}$  mit den Spalten  $(1, 2)$ ,  $(2, 1)$ .

- (2) Da  $A^t A$  und  $AA^t$  für alle  $A \in GL(3, \mathbb{R})$  positiv definit sind, gilt dies mit

$$A = \begin{pmatrix} 1 & -1 & -1 \\ 1 & 1 & 2 \\ 1 & 1 & 0 \end{pmatrix} \quad \text{für} \quad A^t A = \begin{pmatrix} 3 & 1 & 1 \\ 1 & 3 & 3 \\ 1 & 3 & 5 \end{pmatrix} \quad \text{und} \quad A A^t = \begin{pmatrix} 3 & -2 & 0 \\ -2 & 6 & 2 \\ 0 & 2 & 2 \end{pmatrix}.$$

- (3) Auf  $V = \mathbb{R}^2$  sei die symmetrische Bilinearform  $\varphi$  definiert durch

$$\varphi(v, w) = v_1 w_1 - v_2 w_2 \quad \text{für alle } v, w \in \mathbb{R}^2.$$

Für die Basen  $\mathcal{A} = (e_1, e_2)$  und  $\mathcal{B} = (e_1, (2, 1))$  sind

$$A_{\varphi, \mathcal{A}} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad A_{\varphi, \mathcal{B}} = \begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix}$$

die zugehörigen gramschen Matrizen. Sie sind nicht positiv definit.

- (4) Die Definitheit einer Matrix spielt in der mehrdimensionalen Analysis bei der Suche nach lokalen Extrema eine Rolle. Für ein zweimal stetig differenzierbares  $f: \mathbb{R}^2 \rightarrow \mathbb{R}$  und  $(x, y) \in \mathbb{R}^2$  sind der Gradient  $\text{grad}(f)(x, y) = \nabla f(x, y) \in \mathbb{R}^2$  und die Hesse-Matrix  $H_f(x, y) \in \mathbb{R}^{2 \times 2}$  von  $f$  an der Stelle  $(x, y)$  definiert durch

$$\text{grad}(f)(x, y) = (\partial_1 f(x, y), \partial_2 f(x, y)),$$

$$H_f(x, y) = \begin{pmatrix} \partial_1 \partial_1 f(x, y) & \partial_1 \partial_2 f(x, y) \\ \partial_1 \partial_2 f(x, y) & \partial_2 \partial_2 f(x, y) \end{pmatrix},$$

wobei  $\partial_1$  und  $\partial_2$  die partiellen Ableitungen nach der ersten bzw. zweiten Koordinate bezeichnen. Ist  $(x, y) \in \mathbb{R}^2$  ein kritischer Punkt von  $f$ , d. h.  $\text{grad}(f)(x, y) = 0$ , so hat  $f$  in  $(x, y)$  eine lokale Minimalstelle (bzw. Maximalstelle), wenn  $H_f(x, y)$  (bzw.  $-H_f(x, y)$ ) positiv definit ist. Für  $f$  mit  $f(x, y) = x^2 + xy + y^2$  gilt

$$\partial_1 f(x, y) = 2x + y,$$

$$\partial_2 f(x, y) = 2y + x,$$

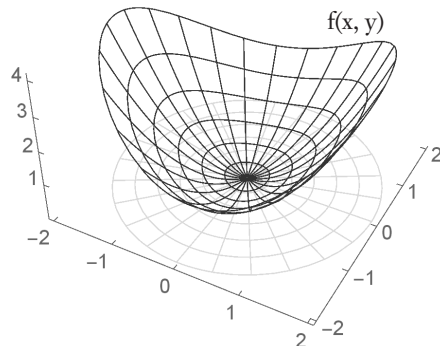
$$\partial_1 \partial_1 f(x, y) = \partial_2 \partial_2 f(x, y) = 2,$$

$$\partial_1 \partial_2 f(x, y) = \partial_2 \partial_1 f(x, y) = 1.$$

Im kritischen Punkt  $0 = (0, 0)$  ist

$$H_f(0) = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \text{ positiv definit.}$$

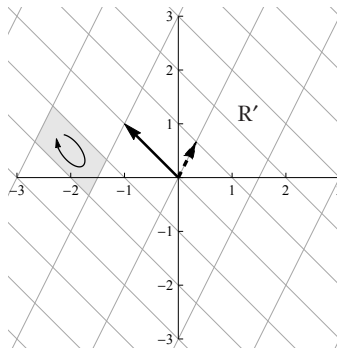
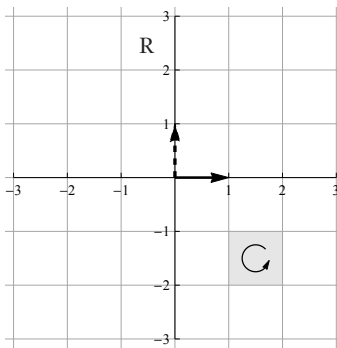
Also ist  $0$  eine lokale Minimalstelle.



## 7.12 Die Determinante eines Endomorphismus

### Definition (Determinante eines Endomorphismus)

Seien  $V$  ein Vektorraum der Dimension  $n \geq 1$ ,  $f: V \rightarrow V$  linear und  $A \in K^{n \times n}$  die darstellende Matrix von  $f$  bzgl. einer beliebigen Basis  $\mathcal{A}$  von  $V$ . Dann heißt  $\det(f) = \det(A)$  die *Determinante* von  $f$ .



$$A = \begin{pmatrix} -1 & 1/3 \\ 1 & 2/3 \end{pmatrix}$$

$$\mathcal{A} = (e_1, e_2)$$

$$f(x) = Ax$$

$$R' = f[R]$$

$$\det(f) = -1$$

$$|\det(f)| = \text{vol}_2(R')$$

$\det(f)$  misst die durch  $f$  bewirkte Veränderung eines orientierten Volumens.

Die Determinante  $\det(f)$  eines Endomorphismus hängt nicht von der Wahl der Basis ab, da ähnliche Matrizen dieselbe Determinante besitzen: Sind  $\mathcal{A}, \mathcal{A}'$  Basen von  $V$  und  $A$  und  $A'$  die darstellenden Matrizen von  $f$  bzgl. dieser Basen, so gibt es nach der Transformationsformel eine Matrix  $S \in GL(n, K)$  mit

$$A' = S A S^{-1}.$$

Nach dem Multiplikationssatz ist  $\det A' = \det S \det A \det S^{-1} = \det A$ .

Allgemeine Eigenschaften sind:

$\det(f \circ g) = \det(f) \det(g)$	<i>Verknüpfung</i>
$\det(f) \neq 0$ genau dann, wenn $f$ ist ein Automorphismus In diesem Fall ist $\det(f^{-1}) = \det(f)^{-1}$ .	<i>Umkehrung</i>

Ist  $V$  euklidisch oder unitär, so können wir  $f^*: V \rightarrow V$  bilden (vgl. 6.11). Ist  $\mathcal{A}$  eine Orthonormalbasis  $\mathcal{A}$  von  $V$ , so ist  $A^t$  (für  $K = \mathbb{R}$ ) bzw.  $A^*$  (für  $K = \mathbb{C}$ ) die darstellende Matrix von  $f^*$  bzgl.  $\mathcal{A}$ . Damit gilt:

$\det(f^*) = \det(f)$	<i>für <math>V</math> euklidisch</i>
$\det(f^*) = \overline{\det(f)}$ Ist $f$ selbstadjungiert, so ist $\det(f) = \det(f^*) \in \mathbb{R}$ .	<i>für <math>V</math> unitär</i>



**Beispiele**

- (1) Die Identität  $\text{id}_V : V \rightarrow V$  hat bzgl. jeder Basis die darstellende Matrix  $E_n$ . Es gilt  $\det(\text{id}_V) = 1$ .

- (2) Seien  $K$  ein Körper,  $V = K^n$  und  $f : V \rightarrow V$ ,

$$f(x_1, \dots, x_n) = f(x_2, x_1, \dots) \quad \text{für alle } (x_1, \dots, x_n) \in V.$$

die Vertauschung der beiden ersten Komponenten. Die  $f$  bzgl.  $(e_1, \dots, e_n)$  darstellende Matrix  $A$  hat die Spalten  $e_2, e_1, e_3, \dots, e_n$ . Damit ist

$$\det(f) = \det(A) = -1.$$

- (3) Sei  $V$  der  $\mathbb{R}$ -Vektorraum der reellen Polynomfunktionen vom Grad kleiner gleich  $n-1$  und sei  $D : V \rightarrow V$  der Ableitungsendomorphismus,

$$D(f) = f' \quad \text{für alle } f \in V.$$

Bzgl. der Basis  $(1, x, \dots, x^{n-1})$  ist die obere Dreiecksmatrix

$$A = \begin{pmatrix} 0 & 1 & & & \\ & 0 & 2 & & \\ & & \dots & \dots & \\ & & & 0 & n-1 \\ & & & & 0 \end{pmatrix}$$

$$D(1) = 0$$

$$D(x) = 1$$

$$D(x^2) = 2x$$

$$\dots$$

$$D(x^{n-1}) = (n-1)x^{n-2}$$

die darstellende Matrix von  $f$  (die Spalten sind die Koordinatenvektoren der Bilder der Basisvektoren). Damit ist  $\det(f) = 0$ .

- (4) Seien  $K$  ein Körper,  $V = K^{2 \times 2}$  und  $f : K^{2 \times 2} \rightarrow K^{2 \times 2}$  die Transposition,

$$f(A) = A^t \quad \text{für alle } A \in K^{2 \times 2}.$$

Dann bilden die Matrizen

$$E_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad E_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad E_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad E_4 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

eine Basis von  $V$ . Die darstellende Matrix von  $f$  bzgl. dieser Basis ist

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Damit ist  $\det(f) = \det(A) = -\det(E_4) = -1$ .



# 8. Kapitel

---

## Eigenwerte

---

## 8.1 Eigenwerte und Eigenvektoren

### Definition (Eigenwert, Eigenvektor, Eigenraum, Spektrum)

#### Eigenwerte und Eigenvektoren für Endomorphismen

Seien  $V$  ein  $K$ -Vektorraum und  $f : V \rightarrow V$  ein Endomorphismus. Weiter seien  $\lambda \in K$  und  $v \in V - \{0\}$ . Dann heißt  $\lambda$  ein *Eigenwert* und  $v$  ein *Eigenvektor* von  $f$  (zum Eigenwert  $\lambda$ ), falls  $f(v) = \lambda v$ . Wir setzen

$$\sigma(f) = \{ \lambda \in K \mid \lambda \text{ ist ein Eigenwert von } f \}, \quad (\text{Spektrum von } f)$$

$$\begin{aligned} \text{Eig}(f, \lambda) &= \{ v \in V \mid v \text{ ist ein Eigenvektor von } f \text{ zum Eigenwert } \lambda \} \cup \{0\} = \\ &= \{ v \in V \mid f(v) = \lambda v \} \quad \text{für alle } \lambda \in \sigma(f). \end{aligned} \quad (\text{Eigenraum von } f \text{ bzgl. } \lambda)$$

Die Dimension des Unterraums  $\text{Eig}(f, \lambda)$  heißt die *geometrische Vielfachheit* des Eigenwerts  $\lambda$  von  $f$ .

#### Eigenwerte und Eigenvektoren für Matrizen

Seien  $K$  ein Körper,  $n \geq 1$  und  $A \in K^{n \times n}$ . Dann heißt ein  $\lambda \in K$  ein *Eigenwert* und  $x \in K^n - \{0\}$  ein *Eigenvektor* von  $A$ , falls  $Ax = \lambda x$ , d. h., falls  $\lambda$  ein Eigenwert und  $x$  ein Eigenvektor des Endomorphismus  $f_A : K^n \rightarrow K^n$  ist. Ebenso sind das *Spektrum* und die *Eigenräume* von  $A$  definiert durch

$$\sigma(A) = \sigma(f_A), \quad \text{Eig}(A, \lambda) = \text{Eig}(f_A, \lambda) \quad \text{für alle } \lambda \in \sigma(A).$$

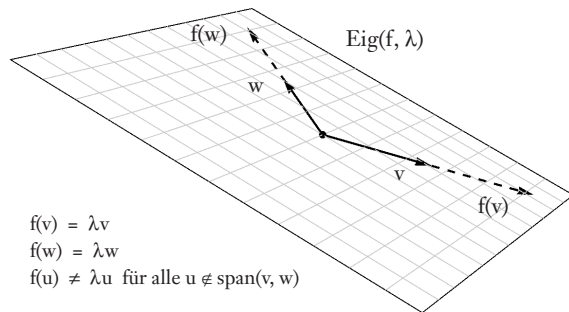
$\dim(\text{Eig}(A, \lambda))$  heißt die *geometrische Vielfachheit* des Eigenwerts  $\lambda$  von  $A$ .

Eigenwerte und Eigenvektoren sind nützlich, um einen Endomorphismus möglichst einfach darzustellen: Auf einem Eigenraum  $\text{Eig}(f, \lambda)$  ist  $f$  die schlichte Skalierung um den Faktor  $\lambda$ . Sind  $v_1, \dots, v_n$  Eigenvektoren von  $f$  zu den Eigenwerten  $\lambda_1, \dots, \lambda_n$ , so gilt

$$(+)\quad f(\alpha_1 v_1 + \dots + \alpha_n v_n) = \lambda_1 \alpha_1 v_1 + \dots + \lambda_n \alpha_n v_n \quad \text{für alle } \alpha_1, \dots, \alpha_n \in K.$$

Ist  $(v_1, \dots, v_n)$  eine Basis von  $V$ , so können wir  $f(w)$  für jeden Vektor  $w$  durch (+) angeben.

Die „Eigen-Begriffe“ übertragen sich in natürlicher Weise von Endomorphismen zwischen endlich-dimensionalen Vektorräumen auf Matrizen. Allgemein spielen Eigenvektoren, Eigenwerte und Spektren aber auch für unendlich-dimensionale Vektorräume eine wichtige Rolle, etwa in der Funktionalanalysis und der Quantenmechanik.



## Die Rolle des Nullvektors und der Null des Skalarenkörpers

Es gilt  $f(0) = 0 = \lambda 0$  für alle  $\lambda \in K$ . Da man nicht möchte, dass jeder Skalar  $\lambda$  ein Eigenwert von  $f$  ist, schließt man den Nullvektor  $0 \in V$  als Eigenvektor aus. In die Eigenräume  $\text{Eig}(\lambda, f)$  nimmt man ihn dagegen mit auf, damit diese Unterräume von  $V$  sind. Der Skalar  $0 \in K$  ist als Eigenwert zugelassen:  $f(v) = 0v = 0$  ist für  $v \neq 0$  eine wichtige Information über  $f$ . Der zugehörige Eigenraum  $\text{Eig}(0, f)$  ist der Kern von  $f$ .

## Grundlegende Eigenschaften

Eigenvektoren $v_1, \dots, v_k$ zu paarweise verschiedenen Eigenwerten $\lambda_1, \dots, \lambda_k$ sind linear unabhängig.	<i>lineare Unabhängigkeit</i>
Die Summe aller Eigenräume ist direkt.	<i>Summe</i>
$\text{Eig}(f, \lambda) = \text{Kern}(f - \lambda \text{Id}_V)$ $\text{Eig}(A, \lambda) = \text{Kern}(A - \lambda E_n)$	<i>Kerndarstellung</i>

Die lineare Unabhängigkeit zeigt man induktiv. Der Induktionsschritt von  $k - 1$  nach  $k$  wird für eine gegebene Nulldarstellung  $\alpha_1 v_1 + \dots + \alpha_k v_k = 0$  durch Subtraktion von

$$\lambda_1 \alpha_1 v_1 + \dots + \lambda_k \alpha_k v_k = 0 \quad (\text{Anwendung von } f \text{ auf die Nulldarstellung})$$

$$\lambda_k \alpha_1 v_1 + \dots + \lambda_k \alpha_k v_k = 0 \quad (\text{Multiplikation der Nulldarstellung mit } \lambda_k)$$

getragen. Die Direktheit von  $\bigoplus_{\lambda \in \sigma(f)} \text{Eig}(f, \lambda)$  folgt nun aus der linearen Unabhängigkeit. Schließlich ist  $f(v) = \lambda v$  äquivalent zu  $f(v) - \lambda v = 0$  und damit zu  $(f - \lambda \text{Id}_V)(v) = 0$ . Letzteres besagt, dass  $v$  im Kern des Endomorphismus  $f - \lambda \text{Id}_V$  liegt. Analoges gilt für Matrizen.

Die folgenden Fragen sind also äquivalent:

*Welche Eigenwerte besitzt  $A$ ?      Für welche  $\lambda$  ist  $A - \lambda E_n$  singulär?*

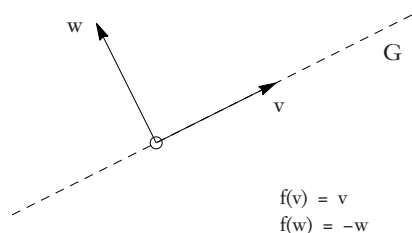
## Beispiele

- Sei  $f_\varphi : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  die Drehung um den Winkel  $\varphi \in [0, 2\pi[$ . Dann gilt:  
Ist  $\varphi = 0$ , so ist  $f_0(x) = x$  für alle  $x$ ; damit ist  $\sigma(f_0) = \{1\}$  und  $\text{Eig}(f_0, 1) = \mathbb{R}^2$ .  
Ist  $\varphi = \pi$ , so ist  $f_\pi(x) = -x$  für alle  $x$ ; damit ist  $\sigma(f_\pi) = \{-1\}$  und  $\text{Eig}(f_\pi, -1) = \mathbb{R}^2$ .  
Für alle anderen  $\varphi$  ist  $\sigma(f_\varphi) = \emptyset$ .

- Sei  $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  die Spiegelung an einer Geraden  $G$  durch 0.  
Dann ist  $f(x) = x$  für alle  $x \in G$  und  $f(x) = -x$  für jedes  $x$ , das senkrecht auf  $G$  steht. Es gibt keine weiteren Eigenvektoren, sodass

$$\sigma(f) = \{1, -1\},$$

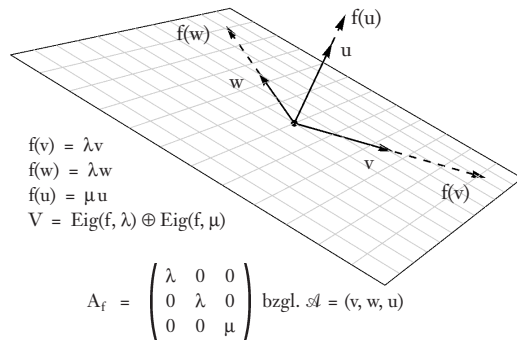
$$\text{Eig}(f, 1) = G, \text{Eig}(f, -1) = G^\perp.$$



## 8.2 Die Diagonalisierbarkeit

### Definition (diagonalisierbare Endomorphismen und Matrizen)

Seien  $V$  ein  $n$ -dimensionaler Vektorraum und  $f : V \rightarrow V$  ein Endomorphismus. Dann heißt  $f$  *diagonalisierbar*, falls eine Basis  $(v_1, \dots, v_n)$  aus Eigenvektoren existiert. Analog heißt eine Matrix  $A \in K^{n \times n}$  *diagonalisierbar*, falls  $f_A : K^n \rightarrow K^n$  dies ist.



Die Diagonalisierbarkeit ist die optimale Eigenschaft im Sinne der einfachen Darstellung. Dies (und die Namensgebung) wird illustriert durch:

<i>Charakterisierungen der Diagonalisierbarkeit von <math>f : V \rightarrow V</math></i>
Es gibt eine Basis $\mathcal{A} = (v_1, \dots, v_n)$ aus Eigenvektoren von $f$ .
$V = \bigoplus_{\lambda \in \sigma(f)} \text{Eig}(f, \lambda)$
$\sum_{\lambda \in \sigma(f)} \dim(\text{Eig}(f, \lambda)) = n$
Es gibt eine Basis $\mathcal{A}$ von $V$ , sodass die darstellende Matrix $D = A_f^{\mathcal{A}, \mathcal{A}}$ von $f$ bzgl. $\mathcal{A}$ , $\mathcal{A}$ eine Diagonalmatrix ist.

<i>Charakterisierungen der Diagonalisierbarkeit von <math>A \in K^{n \times n}</math></i>
Es gibt eine Basis $\mathcal{A} = (v_1, \dots, v_n)$ aus Eigenvektoren von $A$ .
$K^n = \bigoplus_{\lambda \in \sigma(f)} \text{Eig}(A, \lambda)$
$\sum_{\lambda \in \sigma(f)} \dim(\text{Eig}(A, \lambda)) = n$
$A$ ist ähnlich zu einer Diagonalmatrix $D$ , d. h., es gibt ein $S \in \text{GL}(n, K)$ , sodass $D = S A S^{-1}$ eine Diagonalmatrix ist.

Für die Diagonalmatrizen der vierten Formulierung gilt zusätzlich:

*In der Diagonalen von  $D$  stehen die Eigenwerte von  $f$ . Ist die geometrische Vielfachheit von  $\lambda$  gleich  $k$ , so kommt  $\lambda$  genau  $k$ -oft in der Diagonalen vor.*

Einen Endomorphismus zu diagonalisieren bedeutet, eine Basis von  $V$  zu finden, sodass  $A_f^{\mathcal{A}, \mathcal{A}}$  diagonal ist. Wir erinnern hierzu an den in 5.8 betrachteten Spezialfall der Transformationsformel

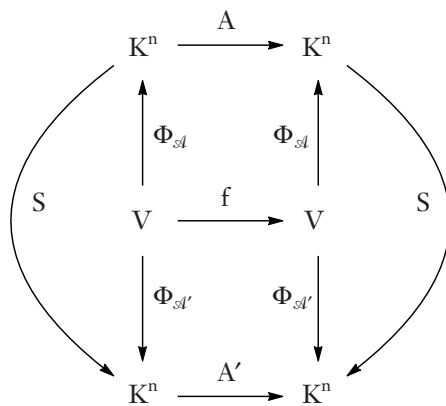
$$A' = S A S^{-1}.$$

Im Unterschied zu „SAT<sup>-1</sup>“ halten wir die Basen  $\mathcal{A}$  und  $\mathcal{A}'$  beim Übergang von links nach rechts fest, sodass nur zwei statt vier Basen im Spiel sind. Gute  $A_f^{\mathcal{A}, \mathcal{A}}$  sind schwieriger zu konstruieren als gute  $A_f^{\mathcal{A}, \mathcal{B}}$  (vgl. 5.4). Die Suche nach guten Darstellungen  $A_f^{\mathcal{A}} := A_f^{\mathcal{A}, \mathcal{A}}$  eines Endomorphismus ist als *Normalformproblem* bekannt. Die Frage nach der Diagonalisierbarkeit ist die wichtigste Instanz des Problems und der Ausgangspunkt für alle weiteren Fragen, die sich stellen, wenn die Diagonalisierung nicht möglich ist.

Für Matrizen halten wir fest:

Ist $\mathcal{A}$ eine Basis von $V$ , so besitzen $A = A_f^{\mathcal{A}}$ und $f$ dieselben Eigenwerte.
Ähnliche Matrizen besitzen dieselben Eigenwerte.

Genauer stimmen für jeden Eigenwert  $\lambda$  auch die geometrischen Vielfachheiten überein. Der Beweis dieser Aussagen folgt aus dem kommutativen Diagramm oben unter Verwendung der Eigenschaft „ $\Phi(\lambda v) = \lambda \Phi(v)$  für alle  $\lambda \in K, v \in V$ “, die jeder Koordinatenisomorphismus  $\Phi : V \rightarrow K^n$  erfüllt.



### Beispiel

Wir betrachten die Matrix

$$A = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix} \in \mathbb{R}^{2 \times 2}.$$

Mit  $w = \sqrt{2}$  sind  $x_1 = (w, 1)$ ,  $x_2 = (-w, 1)$  Eigenvektoren zu den Eigenwerten  $\lambda_1 = 1 + w$  und  $\lambda_2 = 1 - w$  (eine Möglichkeit, Eigenwerte und Eigenvektoren zu finden, diskutieren wir im nächsten Abschnitt). Für die Eigenbasis  $\mathcal{A} = (x_1, x_2)$  ist die darstellende Matrix von  $f_A$  bzgl.  $\mathcal{A}$  die Diagonalmatrix  $D = \text{diag}(1 + w, 1 - w)$  („die Spalten sind die Koordinaten der Bilder der Basisvektoren“). Ist  $T$  die Matrix mit den Spalten  $x_1$  und  $x_2$  (vgl. 5.8), so gilt für  $S = T^{-1}$ :

$$S A S^{-1} = \frac{1}{2} \begin{pmatrix} 1/w & 1 \\ -1/w & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} w & -w \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1+w & 0 \\ 0 & 1-w \end{pmatrix} = D.$$

### 8.3 Das charakteristische Polynom

#### Definition (*charakteristisches Polynom*)

*Charakteristisches Polynom einer Matrix*

Seien  $K$  ein Körper und  $A \in K^{n \times n}$ ,  
 $n \geq 1$ . Dann heißt

$$p_A = \det(A - X E_n) \in K[X]$$

das *charakteristische Polynom* von  $A$ .

$$p_A = \det \begin{pmatrix} a_{11} - X & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - X & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & \dots & \dots & a_{nn} - X \end{pmatrix}$$

*Charakteristisches Polynom eines Endomorphismus*

Sind  $V$  ein  $n$ -dimensionaler  $K$ -Vektorraum und  $f : V \rightarrow V$  ein Endomorphismus, so heißt

$$p_f = \det(A - X E_n) \in K[X]$$

das *charakteristische Polynom* von  $f$ , wobei  $A \in K^{n \times n}$  die darstellende Matrix von  $f$  bzgl. einer beliebigen Basis von  $V$  ist.

Die Definition ist durch die Beobachtung motiviert, dass ein Skalar  $\lambda$  genau dann ein Eigenwert von  $A$  ist, wenn  $A - \lambda E_n$  singular ist. Dies ist genau dann der Fall, wenn

$$p_A(\lambda) = \det(A - \lambda E_n) = 0.$$

Mit anderen Worten:

Die Nullstellen von  $p_A$  sind die Eigenwerte von  $A$ .

#### Bemerkung

Wir haben Determinanten nur für Matrizen mit Einträgen aus einem Körper  $K$  eingeführt. Hier benötigen wir sie für Matrizen mit Einträgen im Polynomring  $K[X]$ . Folgende Lösungen sind möglich: (1) Man entwickelt die Determinantentheorie allgemeiner für Matrizen über Ringen. (2) Man erweitert den Polynomring  $K[X]$  zum Körper  $K(X)$  der rationalen Funktionen, dessen Elemente in der Form  $P(X)/Q(X)$  mit  $P(X), Q(X) \in K[X]$  dargestellt werden können. Wegen  $K[X] \subseteq K(X)$  ist dann die benötigte Determinante erklärt.

Die Leibniz-Formel zeigt, dass  $p_A$  tatsächlich ein Polynom (vom Grad  $n$ ) ist:

$$p_A = \det(A - X E_n) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) (a_{\sigma(1), 1} - X \delta_{\sigma(1), 1}) \dots (a_{\sigma(n), n} - X \delta_{\sigma(n), n}) = b_0 + b_1 X^1 + \dots + b_n X^n,$$

mit gewissen Koeffizienten  $b_0, \dots, b_n \in K$ , von denen wir drei einfach angeben können:

$$b_0 = \det(A), \quad b_{n-1} = (-1)^{n-1} (a_{11} + \dots + a_{nn}), \quad b_n = (-1)^n.$$



Um zu zeigen, dass das charakteristische Polynom  $p_f$  nicht von der Wahl der Basis abhängt, betrachten wir ein  $S \in GL(n, K)$ . Dann gilt für  $A' = SAS^{-1}$ :

$$\begin{aligned}\det(A - XE_n) &= \det(S) \det(A - XE_n) \det(S^{-1}) = \det(S (A - XE_n) S^{-1}) = \\ \det(SAS^{-1} - S XE_n S^{-1}) &= \det(SAS^{-1} - XE_n) = \det(A' - XE_n).\end{aligned}$$

Mit anderen Worten:

Ähnliche Matrizen besitzen dasselbe charakteristische Polynom.

Die Darstellung von  $b_{n-1}$  motiviert einen neuen Begriff: Die Summe der Diagonaleinträge einer Matrix  $A$  heißt die *Spur* von  $A$ ,

$$\text{spur}(A) = a_{11} + \dots + a_{nn}.$$

Da die charakteristischen Polynome ähnlicher Matrizen gleich sind, folgt: Ähnliche Matrizen besitzen die gleiche Spur. Ist  $A$  diagonalisierbar, so ist die Spur von  $A$  also die Summe  $\lambda_1 + \dots + \lambda_n$  der (in ihrer Vielfachheit gezählten) Eigenwerte von  $A$ .

### Beispiel: Die Dimension $n = 2$

Für  $n = 2$  gilt

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

$$X^2 - (a_{11} + a_{22})X + a_{11} a_{22} - a_{12} a_{21} = X^2 - \text{spur}(A)X + \det(A).$$

Ist  $K = \mathbb{R}$ , so entscheidet die Diskriminante  $d = \text{spur}(A)^2 - 4 \det(A)$  über die Existenz von Eigenwerten:

Ist  $d < 0$ , so hat  $A$  keine Eigenwerte.

Ist  $d = 0$ , so ist  $\text{spur}(A)/2$  der einzige Eigenwert von  $A$ .

Ist  $d > 0$ , so hat  $A$  die zwei Eigenwerte  $\lambda_{1,2} = (\text{spur}(A) \pm \sqrt{d})/2$ .

Für die in 8.2 betrachtete Matrix  $A = \begin{pmatrix} 1 & 2 \\ 1 & 1 \end{pmatrix} \in \mathbb{R}^{2 \times 2}$  ist  $d = 8$  und wir erhalten

$$p_A = X^2 - 2X - 1 = (X - (1 + w))(X - (1 - w)) \text{ mit } w = \sqrt{2}.$$

Zugehörige Eigenvektoren kann man nun durch Lösen der Gleichungssysteme

$$(A - (1 + w)E_2)x = 0, \quad (A - (1 - w)E_2)x = 0$$

finden. Die Eigenräume sind Geraden durch 0.

### 8.4 Das Diagonalisierbarkeitskriterium

**Satz** (*Diagonalisierbarkeitskriterium, Übereinstimmung der Vielfachheiten*)

Seien  $V$  ein  $n$ -dimensionaler  $K$ -Vektorraum und  $f : V \rightarrow V$  ein Endomorphismus. Dann sind äquivalent:

- (a)  $f$  ist diagonalisierbar.
- (b) Das charakteristische Polynom  $p_f \in K[X]$  von  $f$  zerfällt in Linearfaktoren,

$$p_f = (-1)^n (X - \lambda_1)^{\mu_1} (X - \lambda_2)^{\mu_2} \dots (X - \lambda_k)^{\mu_k}, \text{ mit } \lambda_i \neq \lambda_j \text{ für } i \neq j,$$

und für alle Eigenwerte  $\lambda$  von  $f$  ist die geometrische Vielfachheit von  $\lambda$  gleich der algebraischen Vielfachheit von  $\lambda$  als Nullstelle von  $p_f$ :

$$\dim(\text{Eig}(f, \lambda_i)) = \mu_i \text{ für alle } 1 \leq i \leq k.$$

Wir wissen, dass die Nullstellen von  $p_f$  die Eigenwerte von  $f$  sind:  $\sigma(f) = \{\lambda \mid p_f(\lambda) = 0\}$ . Weiter wissen wir, dass  $f$  genau dann diagonalisierbar ist, wenn  $V$  die direkte Summe aller Eigenräume ist. Die Diagonalisierbarkeit von  $f$  ist also gleichwertig zu

$$\sum_{\lambda \in \sigma(f)} \dim(\text{Eig}(f, \lambda)) = n.$$

Gilt (b) wie im Satz, so ist dies erfüllt, da dann

$$\sum_{\lambda \in \sigma(f)} \dim(\text{Eig}(f, \lambda)) = \sum_{1 \leq i \leq k} \mu_i = n.$$

Um auch „(a) impliziert (b)“ zu zeigen, beobachten wir, dass für alle Endomorphismen  $f$  und alle Eigenwerte  $\lambda$  von  $f$  unabhängig vom Zerfallen von  $p_f$  in Linearfaktoren gilt:

Die geometrische Vielfachheit von  $\lambda$  ist kleinergleich der algebraischen Vielfachheit der Nullstelle  $\lambda$  von  $p_f$ .

Um die Ungleichung einzusehen, ergänzen wir eine Basis  $(v_1, \dots, v_k)$  von  $\text{Eig}(f, \lambda)$  zu einer Basis  $\mathcal{A} = (v_1, \dots, v_n)$  von  $V$ . Dann hat die darstellende Matrix von  $f$  bzgl.  $\mathcal{A}$  die Blockform

$$A = \begin{pmatrix} \lambda E_k & B \\ 0 & C \end{pmatrix}, \quad k = \dim(\text{Eig}(f, \lambda)).$$

Da wir zur Berechnung von  $p_f$  eine beliebige darstellende Matrix verwenden können, ist

$$p_f = \det(A - X E_n) = \det(\lambda E_k - X E_k) \det(C - X E_{n-k}) = (\lambda - X)^k p_C, \text{ sodass } k \leq \mu_{p_f}(\lambda).$$

Ist  $f$  diagonalisierbar, so gilt also

$$n = \sum_{\lambda \in \sigma(f)} \dim(\text{Eig}(f, \lambda)) \leq \sum_{p_f(\lambda) = 0} \mu_{p_f}(\lambda) \leq n.$$

Dies ist nur möglich, wenn  $\mu_{p_f}(\lambda) = \dim(\text{Eig}(f, \lambda))$  für alle  $\lambda \in \sigma(f)$ , sodass (b) gilt.

In  $\mathbb{C}$  zerfällt jedes Polynom in Linearfaktoren (Fundamentalsatz der Algebra, vgl. 2.12).  
Damit erhalten wir:

<i>Diagonalisierbarkeitskriterium für Endomorphismen <math>f: \mathbb{C} \rightarrow \mathbb{C}</math></i>
$f$ ist genau dann diagonalisierbar, wenn $\dim(\text{Eig}(f, \lambda)) = \mu_{\text{pf}}(\lambda)$ für alle $\lambda \in \sigma(f)$ .

Die folgenden Beispiele zeigen, dass die geometrische Vielfachheit echt kleiner sein kann als die algebraische Vielfachheit.

### Beispiele

(1) Für  $\varepsilon \geq 0$  sei

$$A(\varepsilon) = \begin{pmatrix} 0 & \varepsilon + 1 \\ \varepsilon - 1 & 2 \end{pmatrix}.$$

Es gilt

$$p_{A(\varepsilon)} = \det \begin{pmatrix} -X & \varepsilon + 1 \\ \varepsilon - 1 & 2 - X \end{pmatrix} =$$

$$(X - 1)^2 - \varepsilon^2,$$

sodass  $\sigma(A(\varepsilon)) = \{1 + \varepsilon, 1 - \varepsilon\}$  mit Eigenvektoren

$$v(\varepsilon) = (1, 1), \quad w(\varepsilon) = ((1 + \varepsilon)/(1 - \varepsilon), 1).$$

Im Grenzfall  $\varepsilon = 0$  erhalten wir eine doppelte Nullstelle des charakteristischen Polynoms, deren geometrische Vielfachheit gleich 1 ist.

(2) Seien  $a, b, c \in \mathbb{R}$ . Für die obere Dreiecksmatrix

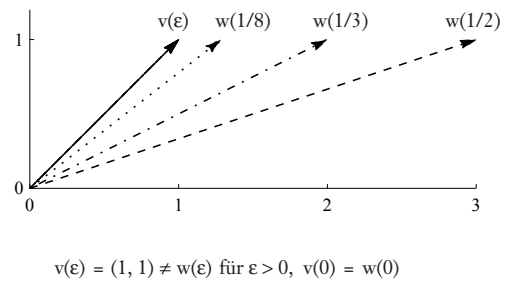
$$A = \begin{pmatrix} a & b \\ 0 & c \end{pmatrix} \in \mathbb{R}^{2 \times 2}$$

ist  $p_f = (a - X)(c - X)$ . Damit gilt  $\sigma(f) = \{a, c\}$ . Ist  $a \neq c$ , so ist  $A$  diagonalisierbar; die Vektoren  $v_1 = e_1 = (1, 0)$  und  $v_2 = (b/(c - a), 1)$  bilden eine Eigenbasis.

Ist  $a = c$ , so ist  $a$  eine doppelte Nullstelle von  $p_f$  und

$$\dim(\text{Eig}(A, a)) = \dim \text{Kern}(A - aE_2) = 2 - \text{rang} \begin{pmatrix} 0 & b \\ 0 & 0 \end{pmatrix} \in \{1, 2\}.$$

Im Fall  $a = c$  ist also  $A$  genau dann diagonalisierbar, wenn  $b = 0$ . Für  $b = 0$  sind  $e_1$  und  $e_2$  Eigenvektoren, für  $b \neq 0$  ist  $\text{Eig}(A, a) = \text{span}(e_1)$ .



## 8.5 Die Trigonalisierung

### Satz (Trigonalisierungssatz, Schur-Zerlegung)

Seien  $V$  ein  $n$ -dimensionaler  $K$ -Vektorraum,  $n \geq 1$ , und  $f : V \rightarrow V$  ein Endomorphismus. Dann sind äquivalent:

- (a)  $V$  besitzt eine Basis  $\mathcal{A}$  derart, dass die darstellende Matrix  $A$  von  $f$  bzgl.  $\mathcal{A}$  eine obere Dreiecksmatrix ist. (Schur-Zerlegung)
- (b) Das charakteristische Polynom  $p_f$  zerfällt in Linearfaktoren.

Ist  $A$  wie in (a), so stehen auf der Diagonale von  $A$  die Eigenwerte von  $f$ . Genauer gilt:  $a_{11}, \dots, a_{nn}$  ist eine Aufzählung der Eigenwerte von  $f$ , in der jeder Eigenwert  $\lambda$  genau  $\mu_f(\lambda)$  mal erscheint. Insbesondere ist

$$\det(f) = \det(A) = a_{11} \dots a_{nn}.$$

$$A = \begin{pmatrix} \lambda_1 & a_{12} & a_{13} & a_{14} & a_{15} & a_{16} \\ 0 & \lambda_1 & a_{23} & a_{24} & a_{25} & a_{26} \\ 0 & 0 & \lambda_1 & a_{34} & a_{35} & a_{36} \\ 0 & 0 & 0 & \lambda_2 & a_{45} & a_{46} \\ 0 & 0 & 0 & 0 & \lambda_2 & a_{56} \\ 0 & 0 & 0 & 0 & 0 & \lambda_3 \end{pmatrix}$$

$$p_A = (-1)^6 (X - \lambda_1)^3 (X - \lambda_2)^2 (X - \lambda_3)$$

$\lambda_1, \lambda_2, \lambda_3$  paarweise verschieden

Wir wissen, dass  $f$  genau dann diagonalisierbar ist, wenn  $p_f$  in Linearfaktoren zerfällt und die algebraischen und geometrischen Vielfachheiten übereinstimmen (vgl. 8.4). Lassen wir die Vielfachheitsforderung fallen, so erhalten wir Trigonalisierbarkeit (Darstellbarkeit durch eine Dreiecksmatrix).

Im Fall  $K = \mathbb{C}$  ist (b) immer erfüllt. Damit kann also jeder Endomorphismus  $V$  eines endlich-dimensionalen  $\mathbb{C}$ -Vektorraums durch eine obere Dreiecksmatrix dargestellt werden. Allgemeiner gilt dies für jeden algebraisch abgeschlossenen Körper, etwa den Körper  $K = \mathbb{A}$  der algebraischen Zahlen.

Für Matrizen lautet das Ergebnis:

Ist  $K$  ein Körper,  $n \geq 1$  und  $A \in K^{n \times n}$  beliebig, so sind äquivalent:

- (a) Es gibt ein  $S \in GL(n, K)$ , sodass  $SAS^{-1}$  eine obere Dreiecksmatrix ist.
- (b)  $p_A$  zerfällt in Linearfaktoren.

Ist  $A$  wie (a), so können wir  $A$  zur Berechnung von  $p_f$  verwenden. Da die Determinante einer Dreiecksmatrix das Produkt ihrer Diagonaleinträge ist, gilt

$$p_f = \det(A - XE_n) = (a_{11} - X) \dots (a_{nn} - X),$$

sodass  $p_f$  in Linearfaktoren zerfällt. Diese Überlegung zeigt auch die Behauptung über die Diagonaleinträge von  $A$ . Die Implikation von (b) nach (a) lässt sich durch Induktion über die Dimension von  $V$  konstruktiv beweisen:

**Konstruktion der Basis  $\mathcal{A}$  und der Dreiecksmatrix  $A$** 

Im Fall  $n = 1$  ist die  $1 \times 1$ -Matrix  $A$  mit  $a_{11} = \lambda_1$  wie gewünscht, wobei  $p_f = \lambda_1 - X$ . Im Induktionsschritt von  $n - 1$  nach  $n$  sei  $\lambda_1$  ein Eigenwert und  $v_1$  ein zugehöriger Eigenvektor von  $f$ . Weiter sei  $\mathcal{B} = (v_1, u_2, \dots, u_n)$  eine Basis von  $V$ . Die  $f$  bzgl.  $\mathcal{B}$  darstellende Matrix hat in der ersten Spalte die gewünschte Form

$$B = \begin{pmatrix} \lambda_1 & b_{12} & \dots & b_{1n} \\ 0 & & B' & \end{pmatrix} \quad \text{mit } B' \in K^{(n-1) \times (n-1)}.$$

Wir setzen  $U = \text{span}(u_2, \dots, u_n) \subseteq V$  und definieren  $g : U \rightarrow U$  durch

$$g(u_j) = f(u_j) - b_{1j}v_1 = b_{2j}u_2 + \dots + b_{nj}u_n \quad \text{für alle } 2 \leq j \leq n.$$

Es gilt  $p_f = (\lambda_1 - X)p_g$ , sodass  $p_g$  in Linearfaktoren zerfällt. Nach Induktionsvoraussetzung gibt es eine Basis  $\mathcal{A}' = (v_2, \dots, v_n)$  von  $U$  derart, dass die darstellende Matrix  $A' \in K^{(n-1) \times (n-1)}$  von  $g$  bzgl.  $\mathcal{A}'$  eine obere Dreiecksmatrix ist. Nun ist  $\mathcal{A} = (v_1, \dots, v_n)$  wie gewünscht, denn die darstellende Matrix von  $f$  bzgl.  $\mathcal{A}$  hat die Form

$$A = \begin{pmatrix} \lambda_1 & a_{12} & \dots & a_{1n} \\ 0 & & A' & \end{pmatrix}.$$

**Beispiel**

Wir betrachten den nicht diagonalisierbaren Endomorphismus  $f_C : \mathbb{R}^3 \rightarrow \mathbb{R}^3$  mit

$$C = \begin{pmatrix} 0 & -1 & 0 \\ 1 & 2 & 0 \\ 0 & 1 & 1 \end{pmatrix}, \quad p_C = -(X-1)^3, \quad \sigma(f) = \{1\}, \quad \mu_{p_C}(1) = 3, \quad \dim(\text{Eig}(C, 1)) = 1.$$

Den Eigenvektor  $v_1 = e_3$  zum Eigenwert  $\lambda_1 = 1$  ergänzen wir durch  $u_2 = e_1, u_3 = e_2$  zur Basis  $\mathcal{B} = (v_1, u_2, u_3) = (e_3, e_1, e_2)$  des  $\mathbb{R}^3$ . Dann ist

$$B = \begin{pmatrix} \lambda_1 & b_{12} & b_{13} \\ 0 & & B' \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & -1 \\ 0 & 1 & 2 \end{pmatrix}, \quad B' = \begin{pmatrix} 0 & -1 \\ 1 & 2 \end{pmatrix}$$

die darstellende Matrix von  $f_C$  bzgl.  $\mathcal{B}$ . Seien  $U = \text{span}(e_1, e_2)$  und  $g : U \rightarrow U$  mit  $g(e_1) = e_2, g(e_2) = -e_1 + 2e_2$ . Dann hat  $g$  den Eigenvektor  $w_1 = (1, -1, 0)$  zum Eigenwert 1 und wird bzgl. der Basis  $\mathcal{A}' = (w_1, e_1)$  von  $U \subseteq \mathbb{R}^3$  durch die Dreiecksmatrix

$$A' = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$$

dargestellt. Die darstellende Matrix von  $f_C$  bzgl.  $\mathcal{A} = (v_1, w_1, e_1)$  ist

$$A = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -1 \\ 0 & 0 & 1 \end{pmatrix} = S C S^{-1} \quad \text{mit} \quad S^{-1} = (v_1 \ w_1 \ e_1) = \begin{pmatrix} 0 & 1 & 1 \\ 0 & -1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

## 8.6 Der Spektralsatz

### Satz (Spektralsatz)

#### Spektralsatz für selbstadjungierte Endomorphismen

Seien  $V$  ein euklidischer oder unitärer  $n$ -dimensionaler Vektorraum und  $f : V \rightarrow V$  ein Endomorphismus. Dann sind äquivalent:

- (a)  $f = f^*$ , d. h., es gilt  $\langle f(v), w \rangle = \langle v, f(w) \rangle$  für alle  $v, w \in V$ .
- (b)  $\sigma(f) \subseteq \mathbb{R}$  und  $V$  besitzt eine Orthonormalbasis aus Eigenvektoren von  $f$ .

#### Spektralsatz für symmetrische bzw. hermitesche Matrizen

Seien  $n \geq 1$  und  $A \in \mathbb{R}^{n \times n}$  bzw.  $A \in \mathbb{C}^{n \times n}$ . Dann sind äquivalent:

- (a)  $A = A^*$  (im Fall  $K = \mathbb{R}$  also  $A = A^t$ ).
- (b)  $\sigma(f) \subseteq \mathbb{R}$  und es gibt eine orthogonale bzw. unitäre Matrix  $S$  derart, dass  $S A S^{-1}$  diagonal ist.

Selbstadjungierte Endomorphismen sind also nicht nur diagonalisierbar, sondern sogar orthogonal diagonalisierbar: Es gibt eine Eigenbasis, die eine Orthonormalbasis von  $V$  ist. In der Sprache der Matrizen bedeutet dies: Eine hermitesche Matrix  $A$  ist nicht nur ähnlich zu einer Diagonalmatrix  $D = \text{diag}(\lambda_1, \dots, \lambda_n)$ , sondern der Übergang  $D = S A S^{-1}$  kann sogar mit einer orthogonalen bzw. unitären Matrix  $S$  erreicht werden, sodass  $S^{-1} = S^*$ .

### Beweis des Spektralsatzes für $K = \mathbb{C}$

Ist  $f$  selbstadjungiert,  $\lambda \in \sigma(f)$  und  $v$  ein Eigenvektor von  $f$  zum Eigenwert  $\lambda$ , so gilt

$$\lambda \langle v, v \rangle = \langle v, \lambda v \rangle = \langle v, f(v) \rangle = \langle f(v), v \rangle = \langle \lambda v, v \rangle = \bar{\lambda} \langle v, v \rangle \quad \text{mit } \langle v, v \rangle \neq 0,$$

sodass  $\lambda = \bar{\lambda}$  und damit  $\lambda \in \mathbb{R}$ . Das Polynom  $p_f$  hat in  $\mathbb{C}$  eine Nullstelle, und diese ist nach dem Gezeigten reell. Mit diesen Beobachtungen kann durch Induktion nach  $n$  bewiesen werden, dass  $V$  eine Orthonormalbasis aus Eigenvektoren besitzt. Im Induktionsschritt von  $n - 1$  nach  $n$  betrachten wir  $\lambda$  und  $v \neq 0$  mit  $f(v) = \lambda v$  und setzen

$$U = \text{span}(v)^\perp = \{ u \in V \mid \langle u, v \rangle = 0 \}.$$

Für alle  $u \in U$  gilt

$$\langle f(u), v \rangle = \langle u, f(v) \rangle = \langle u, \lambda v \rangle = \lambda \langle u, v \rangle = 0,$$

sodass  $f[U] \subseteq U$ . Damit ist  $f|_U : U \rightarrow U$  ein selbstadjungierter Endomorphismus, der nach Induktionsvoraussetzung eine Orthonormalbasis aus Eigenvektoren besitzt. Ergänzen wir eine solche Basis um  $v$ , so erhalten wir wegen  $V = U \oplus \text{span}(v)$  eine Orthonormalbasis aus Eigenvektoren von  $f$  für ganz  $V$ .

Ist umgekehrt  $(v_1, \dots, v_n)$  eine Orthonormalbasis von  $V$  aus Eigenvektoren von  $f$ , so ist

$$\langle f(v_i), v_j \rangle = \langle \lambda_i v_i, v_j \rangle = \bar{\lambda}_i \langle v_i, v_j \rangle = \lambda_i \delta_{ij} = \langle v_i, \lambda_j v_j \rangle = \langle v_i, f(v_j) \rangle \quad \text{für alle } i, j.$$

Hieraus ergibt sich, dass  $f$  selbstadjungiert ist.

**Beispiel**

Seien  $a, b \in \mathbb{R}$  mit  $a^2 + b^2 = 1$ . Weiter sei  $\alpha$  der von  $(a, b)$  und  $(1, 0)$  eingeschlossene Winkel. Dann beschreibt die symmetrische Matrix

$$A = \begin{pmatrix} a & b \\ b & -a \end{pmatrix} = \begin{pmatrix} \cos \alpha & \sin \alpha \\ \sin \alpha & -\cos \alpha \end{pmatrix} \in O(2), \quad \det(A) = -1,$$

die Spiegelung an der Geraden  $G$  durch 0 mit dem Winkel  $\beta = \alpha/2$  (vgl. 7.8). Damit hat  $A$  die Eigenwerte  $\lambda_{1,2} = \pm 1$  und zugehörige normierte Eigenvektoren

$$v_1 = N(a+1, b) = (\cos \beta, \sin \beta),$$

$$v_2 = N(-b, a+1) = (-\sin \beta, \cos \beta),$$

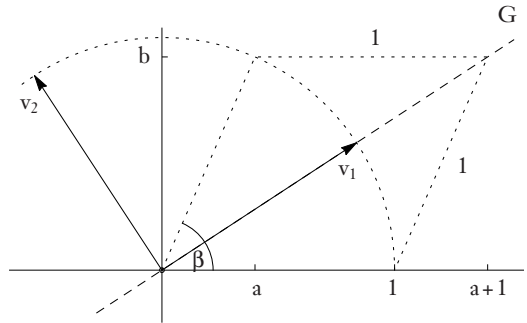
mit  $N(v) = v/\|v\|$ . Ist  $T$  die Matrix mit den Spalten  $v_1$  und  $v_2$ , so ist  $T$  die Drehmatrix in  $SO(2)$  um den Winkel  $\beta$ . Für  $S = T^{-1} = T^t$  gilt also

$$S A S^{-1} = \begin{pmatrix} \cos \beta & \sin \beta \\ -\sin \beta & \cos \beta \end{pmatrix} \begin{pmatrix} a & b \\ b & -a \end{pmatrix} \begin{pmatrix} \cos \beta & -\sin \beta \\ \sin \beta & \cos \beta \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Schreibt man eine beliebige symmetrische Matrix  $B \in \mathbb{R}^{2 \times 2}$  als

$$B = \text{diag}(d, d) + rA, \quad \text{mit } d = \text{spur}(B)/2, \quad r = \|(b_{11} - b_{22})/2, b_{21}\|,$$

so hat  $A$  die gerade untersuchte Form. Man kann nun ablesen, dass  $B$  die Eigenwerte  $d \pm r$  und die Eigenvektoren  $v_1, v_2$  wie oben besitzt.

**Der Spektralsatz für normale Endomorphismen und Matrizen**

Für  $K = \mathbb{R}$  ist  $\sigma(f) \subseteq \mathbb{R}$  immer richtig, sodass die Existenz einer orthonormalen Eigenbasis äquivalent zur Selbstadjungiertheit von  $f$  ist. Für  $K = \mathbb{C}$  liefert das Streichen von „ $\sigma(f) \subseteq \mathbb{R}$ “ in (b) eine echte Abschwächung, die sich ebenfalls durch eine Adjungiertheits-Bedingung einfangen lässt. Zur Motivation beobachten wir: Für alle  $f \in \text{End}(V)$  sind  $f \circ f^*$  und  $f^* \circ f$  selbstadjungiert. Im Allgemeinen ist aber  $f \circ f^* \neq f^* \circ f$ . Man nennt  $f$  *normal*, falls  $f \circ f^* = f^* \circ f$ . Gleichwertig dazu ist, dass  $\langle f(v), f(w) \rangle = \langle f^*(v), f^*(w) \rangle$  für alle  $v, w \in V$ . Wichtige Beispiele neben den selbstadjungierten Endomorphismen sind unitäre  $f$ , da dann  $f \circ f^* = f \circ f^{-1} = f^{-1} \circ f = f^* \circ f$ . Äquivalent sind nun:

- (a)  $f$  ist normal. (b)  $V$  besitzt eine Orthonormalbasis aus Eigenvektoren von  $f$ .

Analog nennt man eine Matrix  $A \in \mathbb{C}^{n \times n}$  *normal*, falls  $A A^* = A^* A$ . Die Normalität von  $A \in \mathbb{C}^{n \times n}$  ist äquivalent zur Existenz einer unitären Matrix  $S$ , für die  $S A S^{-1}$  diagonal ist. Normalität für reelle Matrizen diskutieren wir im Überblick 10.

## 8.7 Hauptachsentransformation und Trägheitssatz

### Satz (Hauptachsentransformation, Trägheitssatz von Sylvester)

Sei  $K = \mathbb{R}$  oder  $K = \mathbb{C}$ , und sei  $A \in K^{n \times n}$  symmetrisch bzw. hermitesch mit Eigenwerten  $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ . Dann gilt (mit dem kanonischen Skalarprodukt):

#### Hauptachsentransformation, Version I

Es gibt eine Orthonormalbasis  $(x_1, \dots, x_n)$  des  $K^n$  mit  $\langle x_i, Ax_j \rangle = \lambda_i \delta_{ij}$  für alle  $i, j$ .

#### Hauptachsentransformation, Version II

Es gibt eine Orthogonalbasis  $(y_1, \dots, y_n)$  des  $K^n$  mit  $\langle y_i, Ay_j \rangle = \alpha_i \delta_{ij}$  für alle  $i, j$ , wobei  $\alpha_i = \text{sgn}(\lambda_i) \in \{-1, 0, 1\}$ .

#### Trägheitssatz von Sylvester

Ist  $(v_1, \dots, v_n)$  eine Orthogonalbasis des  $K^n$  bzgl.  $\langle \cdot, A \cdot \rangle$  (d.h.  $(v_1, \dots, v_n)$  ist eine Basis des  $K^n$  mit  $\langle v_i, Av_j \rangle = 0$  für  $i \neq j$ ), so gilt

$$(+) \quad |\{i \mid \langle v_i, Av_i \rangle > 0\}| = |\{i \mid \lambda_i > 0\}|, \quad \text{wobei } \diamond \in \{>, <, =\}.$$

Weiter gilt: Für alle  $S \in GL(n, K)$  haben  $A$  und  $S^*AS$  dieselben (in ihrer Vielfachheit gezählten) Anzahlen an positiven und negativen Eigenwerten.

Die Hauptachsentransformation ergibt sich aus dem Spektralsatz: Ist  $(x_1, \dots, x_n)$  eine Orthonormalbasis des  $K^n$  mit  $Ax_i = \lambda_i x_i$  für alle  $i$ , so gilt Version I. Version II erhält man durch Setzen von

$$y_i = \frac{1}{\sqrt{|\lambda_i|}} x_i \quad \text{für } \lambda_i \neq 0, \quad y_i = x_i \quad \text{sonst.}$$

Die Normierungseigenschaft  $\langle y_i, Ay_i \rangle \in \{1, 0, -1\}$  gewinnt man also auf Kosten der Normiertheit der  $y_i$ .

Für den Trägheitssatz sind zusätzliche Argumente nötig. Wir begnügen uns hier mit einem Beispiel zur Illustration der Voraussetzung des Satzes:

### Beispiel

Die Matrix  $A \in \mathbb{R}^{2 \times 2}$  rechts hat die Eigenwerte  $\lambda_{1,2} = \pm \sqrt{2}$ . Für die Basis  $(v_1, v_2)$  des  $\mathbb{R}^2$  mit  $v_1 = (1, 0)$  und  $v_2 = (1, 1)$  gilt  $\langle v_1, Av_1 \rangle = 1$  und  $\langle v_2, Av_2 \rangle = 2$ . Dies zeigt, dass (+) ohne die Orthogonalitätsvoraussetzung verletzt sein kann.

$$A = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}$$

$$p_A = X^2 - 2$$

Bezeichnen  $s^+$ ,  $s^-$ ,  $s^0$  die Anzahlen der positiven, negativen bzw. Null-Eigenwerte von  $A$ , so heißt  $(s^+, s^-)$  die *Signatur* oder der *Typ* von  $A$  ( $s^0$  berechnet sich durch  $n - s^+ - s^-$ ). Der Trägheitssatz besagt, dass  $A$  und  $B = S^*AS$  für alle  $S \in GL(n, K)$  dieselbe Signatur haben (nicht nur für orthogonale bzw. unitäre  $S$ , für die  $S^* = S^{-1}$  gilt, sodass  $A$  und  $B$  ähnlich sind und folglich dieselben Eigenwerte besitzen).

Wir betrachten einige Anwendungen.



### 1. Eigenwertkriterium für positive Definitheit

Eine symmetrische bzw. hermitesche Matrix  $A \in K^{n \times n}$  ist genau dann positiv definit, wenn alle Eigenwerte  $\lambda_1, \dots, \lambda_n$  positiv sind. Denn ist  $(x_1, \dots, x_n)$  wie in der Hauptachsentransformation I und  $x = \sum_i \alpha_i x_i \in K^n$ , so ist

$$\langle x, Ax \rangle = \sum_i \lambda_i |\alpha_i|^2.$$

Die Summe rechts ist genau dann für alle  $x \neq 0$  positiv, wenn alle  $\lambda_i$  positiv sind.

### 2. Kongruente Matrizen

Zwei Matrizen  $A, B$  des  $K^{n \times n}$  mit  $K = \mathbb{R}$  oder  $K = \mathbb{C}$  heißen *kongruent*, wenn es ein  $S \in GL(n, K)$  gibt mit  $B = S^*AS$ . Die Hauptachsentransformation II zeigt: Jede symmetrische bzw. hermitesche Matrix  $A$  ist kongruent zu einer Diagonalmatrix der Form  $D = \text{diag}(1, \dots, 1, -1, \dots, -1, 0, \dots, 0)$  mit  $s^+$ ,  $s^-$  bzw.  $s^0$  Einträgen 1, -1 bzw. 0. Ein  $S$  mit  $D = S^*AS$  wird aus geeignet angeordneten Basisvektoren  $y_i$  gebildet.

### 3. Quadriken

Eine Funktion  $q: \mathbb{R}^n \rightarrow \mathbb{R}$  heißt *quadratisch*, falls es eine symmetrische Matrix  $A \in \mathbb{R}^{n \times n}$ ,  $A \neq 0$ , einen Vektor  $b \in \mathbb{R}^n$  und einen Skalar  $c \in \mathbb{R}$  gibt mit

$$q(x) = \langle x, Ax \rangle + \langle b, x \rangle + c \quad \text{für alle } x \in \mathbb{R}^n \quad (\text{mit dem kanonischen Skalarprodukt}).$$

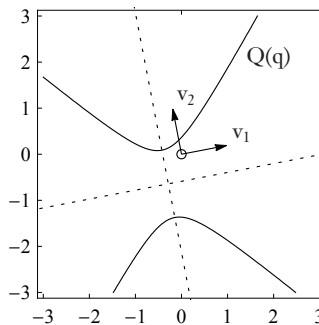
Die Menge  $Q(q) = \{x \in \mathbb{R}^n \mid q(x) = 0\}$  heißt die durch  $q$  definierte *Quadrik* des  $\mathbb{R}^n$ . Die Hauptachsentransformation liefert ein  $S \in SO(n)$  (mit Eigenvektoren von  $A$  als Spalten), sodass  $Q(q)$  für einen gewissen Translationsvektor  $v \in \mathbb{R}^n$  in die Quadrik  $Q(p) = \{S^t(x+v) \mid x \in Q(q)\}$  übergeht, mit

$$p(x) = \langle x, Dx \rangle + \langle (0, \dots, 0, b_{r+1}', \dots, b_n'), x \rangle \quad \text{oder} \quad p(x) = \langle x, Dx \rangle + c',$$

wobei  $r = \text{rang}(A) = s^+ + s^-$ ,  
 $D = \text{diag}(\lambda_1, \dots, \lambda_r, 0, \dots, 0)$ .

Mit Hilfe dieser Normalformdarstellung lassen sich Quadriken klassifizieren.

Für  $n = 2$  ergeben sich Kegelschnitte (Überblick 9).



$q(x) = \langle x, Ax \rangle + \langle b, x \rangle + c$   
 mit  
 $A = \begin{pmatrix} 3 & 1 \\ 1 & -2 \end{pmatrix},$   
 $b = (3, -2), c = 1$   
 und Eigenvektoren  $v_1, v_2$  von  $A$

### 4. Übertragung auf Sesquilinearformen

Für eine symmetrische bzw. hermitesche Form

$\varphi: V \times V \rightarrow K$  und eine Basis  $\mathcal{A} = (v_1, \dots, v_n)$  von  $V$  gilt nach 6.12

$$\varphi(v, w) = \langle v_{\mathcal{A}}, A_{\varphi} w_{\mathcal{A}} \rangle_{\text{kanonisch}} \quad \text{für alle } v, w \in V, \quad A_{\varphi}(i, j) = \varphi(v_i, v_j), \quad A_{\varphi} = A_{\varphi}^*.$$

Die Hauptachsentransformation liefert: Es gibt eine Basis  $\mathcal{A}$  von  $V$ , die eine Orthogonalbasis bzgl.  $\varphi$  ist, d.h.,  $A_{\varphi}$  ist diagonal. Ist  $V$  mit einem Skalarprodukt versehen, so kann  $\mathcal{A}$  als Orthonormalbasis von  $V$  gewählt werden. Der Trägheitssatz liefert: Die Anzahlen der  $i$  mit  $\varphi(v_i, v_i) \diamond 0$  sind für jede Orthogonalbasis  $\mathcal{A}$  von  $V$  bzgl.  $\varphi$  gleich.

## 8.8 Die Singulärwertzerlegung

### Satz (Singulärwertzerlegung)

Seien  $V, W$  endlich-dimensionale euklidische oder unitäre Vektorräume,  $n = \dim(V)$ ,  $m = \dim(W)$ . Weiter sei  $f: V \rightarrow W$  ein Homomorphismus. Dann gibt es Orthonormalbasen  $\mathcal{A}$  und  $\mathcal{B}$  von  $V$  bzw.  $W$  derart, dass die darstellende Matrix  $A \in K^{m \times n}$  von  $f$  bzgl.  $\mathcal{A}, \mathcal{B}$  von der rechteckig diagonalen Form

$$A = \begin{pmatrix} \text{diag}(\sigma_1, \dots, \sigma_r) & 0 \\ 0 & 0 \end{pmatrix}$$

ist, mit  $r = \dim(\text{Bild}(f))$  und reellen bis auf die Reihenfolge eindeutig bestimmten Diagonaleinträgen  $\sigma_1, \dots, \sigma_r > 0$ .

Die Diagonaleinträge

$$\sigma_1 > 0, \dots, \sigma_r > 0, \sigma_{r+1} = a_{r+1, r+1} = 0, \dots, \sigma_{n'} = a_{n', n'} = 0$$

mit  $n' = \min(m, n)$  heißen die *Singulärwerte* von  $f$ .

Im Unterschied zur in 5.4 erreichten Normalformdarstellung (mit der Matrix  $E_r$  oben links) verlangen wir hier Orthonormalbasen, was die Konstruktion erschwert. Im Gegensatz zum Normalformproblem für Endomorphismen sind in der Singulärwertzerlegung jedoch unterschiedliche Basen links und rechts zugelassen (auch im Fall  $V = W$ ), was die Aufgabe erleichtert. Der folgende Beweis zeigt, wie sich die Singulärwertzerlegung aus einer durch den Spektralsatz gelieferten Orthonormalbasis von  $f^* \circ f$  ergibt.

### Konstruktion der Singulärwertzerlegung

Der Endomorphismus  $f^* \circ f: V \rightarrow V$  ist selbstadjungiert, sodass nach dem Spektralsatz eine Orthonormalbasis  $\mathcal{A} = (v_1, \dots, v_n)$  aus Eigenvektoren von  $f^* \circ f$  existiert. Für die zugehörigen Eigenwerte  $\lambda_1, \dots, \lambda_n$  gilt

$$\lambda_j = \langle v_j, f^*(f(v_j)) \rangle_V = \langle f(v_j), f(v_j) \rangle_W \geq 0 \quad \text{für alle } 1 \leq j \leq n.$$

Durch Umordnung erreichen wir, dass  $\lambda_1, \dots, \lambda_r > 0, \lambda_{r+1} = \dots = \lambda_n = 0$  für

$$r = \dim(\text{Bild}(f^* \circ f)) = \dim(\text{Bild}(f)) \leq \min(m, n).$$

Wir setzen nun

$$\sigma_j = \sqrt{\lambda_j}, \quad w_j = \frac{f(v_j)}{\sigma_j} \quad \text{für } 1 \leq j \leq r.$$

Für die Vektoren  $w_1, \dots, w_r$  gilt

$$\langle w_j, w_k \rangle_W = \frac{\langle f(v_j), f(v_k) \rangle_W}{\sigma_j \sigma_k} = \frac{\langle v_j, f^*(f(v_k)) \rangle_V}{\sigma_j \sigma_k} = \frac{\lambda_k}{\sigma_j \sigma_k} \langle v_j, v_k \rangle_V = \delta_{jk}.$$

Ergänzen wir sie zu einer Orthonormalbasis  $\mathcal{B}$  von  $W$ , so gilt  $\langle w_i, f(v_j) \rangle_W = \sigma_j \delta_{ij}$  für alle  $1 \leq i \leq m, 1 \leq j \leq n$ , sodass  $A = A_f$  bzgl.  $\mathcal{A}, \mathcal{B}$  die gewünschte Form hat.

Wir formulieren das Ergebnis noch explizit für Matrizen. Dabei notieren wir Matrizen  $A \in K^{m \times n}$  der Form des Satzes kurz als  $\text{diag}(\sigma_1, \dots, \sigma_r, 0, \dots, 0)$ .

<i>Singulärwertzerlegung für Matrizen</i>
Für alle $A \in \mathbb{R}^{m \times n}$ gibt es $S \in O(m)$ , $T \in O(n)$ mit $SAT^t = SAT^{-1} = \text{diag}(\sigma_1, \dots, \sigma_r, 0, \dots, 0) \in \mathbb{R}^{m \times n}$ mit positiven $\sigma_j$ .
Für alle $A \in \mathbb{C}^{m \times n}$ gibt es $S \in O(m)$ , $T \in U(n)$ mit $SAT^* = SAT^{-1} = \text{diag}(\sigma_1, \dots, \sigma_r, 0, \dots, 0) \in \mathbb{C}^{m \times n}$ mit positiven $\sigma_j$ .

### Die Determinante von $f$

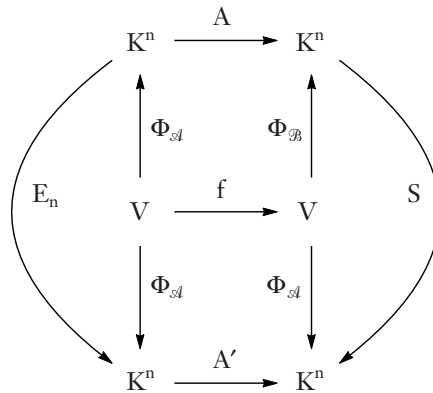
Ist  $V = W$ , so ist  $\det(f)$  definiert.

Ist  $S$  die orthogonale bzw. unitäre Transformationsmatrix des Basiswechsels von  $\mathcal{B}$  nach  $\mathcal{A}$ , so ist  $A' = SA$  die darstellende Matrix von  $f$  bzgl.  $\mathcal{A}$ ,  $\mathcal{A}$ .

Folglich ist

$$\det(f) = \det(SA) = \det(S) \det(A) =$$

$$\det(S) \sigma_1 \dots \sigma_n = \pm \sigma_1 \dots \sigma_n.$$



### Beispiel

Wir betrachten den  $\mathbb{R}^2$  mit dem kanonischen Skalarprodukt. Sei  $f = f_A : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  mit

$$A = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}, \text{ sodass } A^t A = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = 2 E_2.$$

Dann hat die Matrix  $A$  keine Eigenwerte. Die Abbildung  $f^* \circ f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  erfüllt  $(f^* \circ f)(x) = A^t A x = 2x$  für alle  $x$ , hat also die Eigenwerte  $\lambda_{1,2} = 2$  und die orthonormale Eigenbasis  $\mathcal{A} = (e_1, e_2)$ . Wir setzen

$$\sigma_{1,2} = \sqrt{2}, \quad \alpha = 1/\sqrt{2}, \quad w_1 = \alpha f(e_1) = \alpha(1, -1), \quad w_2 = \alpha f(e_2) = \alpha(1, 1).$$

Die darstellende Matrix von  $f$  bzgl. der Orthonormalbasen  $\mathcal{A}$  und  $\mathcal{B} = (w_1, w_2)$  ist  $\text{diag}(\sigma_1, \sigma_2)$ . Die zugehörige Matrizenversion lautet

$$S A T^t = \begin{pmatrix} \alpha & -\alpha \\ \alpha & \alpha \end{pmatrix} \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} \sqrt{2} & 0 \\ 0 & \sqrt{2} \end{pmatrix} = \text{diag}(\sigma_1, \sigma_2),$$

wobei die Spalten von  $S^t$ ,  $T^t \in O(2)$  aus den Vektoren in  $\mathcal{B}$  bzw.  $\mathcal{A}$  gebildet sind.

## 8.9 Lineare Abbildungen und Ellipsen

### Satz (Bild eines Kreises unter einer linearen Abbildung)

Seien  $S^1 = \{x \in \mathbb{R}^2 \mid x_1^2 + x_2^2 = 1\}$  der Einheitskreis im  $\mathbb{R}^2$  und  $A \in \mathbb{R}^{2 \times 2}$ . Dann ist

$$E_A = \{Ax \mid x \in S^1\}$$

eine Ellipse. Die Singulärwerte  $a, b$  von  $A$  sind die Längen der Halbachsen von  $E_A$ . Ist  $A = PDQ$  mit  $D = \text{diag}(a, b)$  und  $Q, P \in O(2)$ , so zeigen die Spalten  $p_1, p_2$  von  $P$  in die Richtungen der Halbachsen.

Aus der Definition von  $f_A(x) = Ax$  ergibt sich, dass  $f_A$  das von  $e_1$  und  $e_2$  aufgespannte Quadrat in das von  $f(e_1)$  und  $f(e_2)$  aufgespannte Parallelogramm verwandelt. Dass das Bild des Einheitskreises unter  $f_A$  eine Ellipse ist, ist elementar nur für spezielle Matrizen leicht einzusehen. Für orthogonale Matrizen (Drehungen und Spiegelungen) ist nichts zu zeigen, da das Bild des Einheitskreises hier wieder der Einheitskreis ist. Für eine Diagonalmatrix  $D = \text{diag}(a, b)$  ist die Aussage ebenfalls klar. Denn es gilt

$$E_D = \{Dx \mid x \in S^1\} = \{(a x_1, b x_2) \mid x \in S^1\} = \{(a \cos t, b \sin t) \mid t \in [0, 2\pi[ \},$$

und die rechte Seite ist die parametrisierte Darstellung einer achsenparallelen Ellipse mit den Halbachsen  $a$  und  $b$  (die auch 0 sein können, sodass die Ellipse degeneriert ist). Im Fall  $a, b > 0$  erhalten wir die äquivalente Darstellung

$$E_D = \left\{ (x, y) \in \mathbb{R}^2 \mid \left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 = 1 \right\}.$$

Mit Hilfe der Singulärwertzerlegung können wir nun allgemein zeigen, dass jede beliebige lineare Abbildung Kreise in Ellipsen verwandelt (wegen  $f_A(\alpha x) = \alpha f_A(x)$  genügt es, dies für  $S^1$  zu beweisen). Zudem erkennen wir die geometrische Bedeutung der Singulärwerte.

### Beweis des Satzes

Sei  $A \in \mathbb{R}^{2 \times 2}$ , und seien  $a, b \geq 0$  die Singulärwerte von  $f_A$ . Weiter sei  $D = \text{diag}(a, b)$ . Die Singulärwertzerlegung liefert  $P, Q \in O(2)$  mit

$$(+)\quad A = PDQ.$$

Da  $Q$  orthogonal ist, gilt  $f_Q[S^1] = S^1$ . Nach (+) gilt also

$$E_A = \{Ax \mid x \in S^1\} = \{PDQx \mid x \in S^1\} = \{PDx \mid x \in S^1\} = f_P[E_D].$$

Wegen  $P \in O(2)$  ist  $P$  entweder eine Drehung oder eine Spiegelung an einer Geraden durch den Nullpunkt. In beiden Fällen ist  $f_P[E]$  eine Ellipse mit Mittelpunkt 0, deren Halbachsen durch die Vektoren

$$a(\cos \alpha, \sin \alpha), \quad b(-\sin \alpha, \cos \alpha)$$

gegeben sind. Dabei ist  $\alpha$  der Drehwinkel bzw. das Doppelte des Winkels, den die Spiegelungsgerade mit der  $x$ -Achse einschließt.

**Beispiel**

Wir betrachten die Drehmatrizen

$$R(\alpha) = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}$$

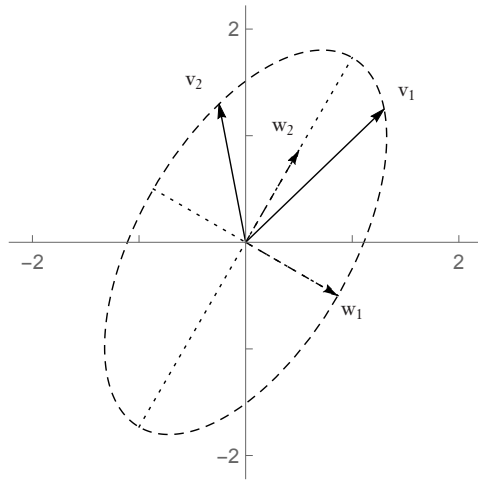
in  $SO(2)$  und definieren

$$A = R(-\pi/6) \operatorname{diag}(1, 2) R(\pi/3) =$$

$$\frac{1}{4} \begin{pmatrix} 3\sqrt{3} & -1 \\ 5 & 3\sqrt{3} \end{pmatrix}.$$

Die Matrix  $A$  beschreibt eine Drehung um 60 Grad, gefolgt von einer Streckung um den Faktor 2 in  $y$ -Richtung und einer Drehung um  $-30$  Grad.

Die Singulärwerte von  $A$  sind 1 und 2. Das Bild von  $S^1$  unter  $f_A$  ist eine Ellipse mit Halbachsenlängen 1 und 2, die in die Richtung der Spalten von  $R(-\pi/6)$  zeigen.



$v_1, v_2$  sind die Spalten von  $A$ ;  $w_1, w_2$  die von  $R(-\pi/6)$

Das Ergebnis lässt sich wie folgt verallgemeinern:

**Satz** (Bild der Sphäre unter einer linearen Abbildung)

Sei  $S^{n-1} = \{x \in \mathbb{R}^n \mid \|x\| = 1\}$  die (euklidische) Einheitssphäre im  $\mathbb{R}^n$ , wobei  $n \geq 1$ . Weiter sei  $A \in \mathbb{R}^{n \times n}$ . Dann ist  $E_A = \{Ax \mid x \in S^{n-1}\}$  ein  $n$ -dimensionales Ellipsoid. Sind  $\sigma_1, \dots, \sigma_n \geq 0$  die Singulärwerte von  $A$ , und ist  $A = PDQ$  mit  $D = \operatorname{diag}(\sigma_1, \dots, \sigma_n)$  und  $Q, P \in O(n)$ , so haben die Halbachsen von  $E_A$  die Längen  $\sigma_1, \dots, \sigma_n$ . Die Spalten  $p_1, \dots, p_n$  von  $P$  zeigen in die Richtungen der Halbachsen.

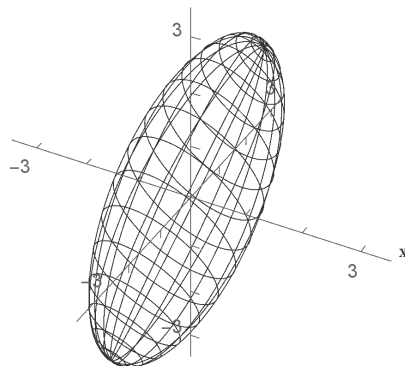
**Beispiel**

Sei  $R_x(\alpha) \in SO(3)$  die Drehung um  $\alpha$  um die  $x$ -Achse. Wir setzen

$$A = R(-\pi/3) \operatorname{diag}(1, 2, 3) =$$

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 3\sqrt{3}/2 \\ 0 & -\sqrt{3} & 3/2 \end{pmatrix}.$$

Die Matrix  $A$  beschreibt eine Streckung um die Faktoren 1, 2, 3 in  $x, y, z$ , gefolgt von einer Drehung um  $-60$  Grad um die  $x$ -Achse. Das Bild von  $S^2$  unter  $f_A$  ist ein Ellipsoid mit Halbachsenlängen 1, 2, 3.



## 8.10 Minimalpolynome und der Satz von Cayley-Hamilton

**Satz** (Existenz des Minimalpolynoms, Satz von Cayley-Hamilton)

Seien  $K$  ein Körper,  $n \geq 1$ ,  $A \in K^{n \times n}$  und

$$I_A = \{ p \in K[X] \mid p(A) = 0 \}.$$

Dann gibt es genau ein Polynom  $m_A \in I_A$  mit

$$(+)\quad m_A \text{ ist normiert und } I_A = \{ q m_A \mid q \in K[X] \}. \quad (\text{Existenz des Minimalpolynoms})$$

$$\text{Weiter gilt } p_A \in I_A, \text{ d.h. } p_A(A) = 0. \quad (\text{Satz von Cayley-Hamilton})$$

$$p_A = \pm (X - \lambda_1)^{\mu_1} \dots (X - \lambda_k)^{\mu_k}$$

$$m_A = (X - \lambda_1)^{v_1} \dots (X - \lambda_k)^{v_k}$$

Zerfällt  $p_A$  in Linearfaktoren, so hat  $m_A$  Exponenten  $v_i$  mit  $1 \leq v_i \leq \mu_i$ . Ist  $A$  diagonalisierbar, so gilt  $v_i = 1$  für alle  $i$ . In 8.11 werden wir die  $v_i$  allgemein charakterisieren.

Während wir bislang Körperelemente in Polynome des Rings  $K[X]$  eingesetzt haben (jedes  $\alpha \in K$  und  $p \in K[X]$  liefert ein  $p(\alpha) \in K$ ), so setzen wir nun quadratische Matrizen einer bestimmten Dimension  $n$  in die Polynome von  $K[X]$  ein: Ist

$$p = \alpha_0 X^0 + \alpha_1 X + \dots + \alpha_k X^k \in K[X],$$

so ist für alle  $A \in K^{n \times n}$  die Auswertung  $p(A)$  definiert durch

$$p(A) = \alpha_0 E_n + \alpha_1 A + \dots + \alpha_k A^k \in K^{n \times n}.$$

Nun halten wir ein Matrix  $A \in K^{n \times n}$  fest und werten alle Polynome  $p \in K[X]$  an der Stelle  $A$  aus. Die Menge  $I_A$  aller  $p$  mit  $p(A) = 0$  ist ein Ideal in  $K[X]$ , d.h.,  $I_A$  ist eine Untergruppe von  $(K[X], +)$  und für alle  $p \in I_A$  und  $q \in K[X]$  ist  $qp \in I_A$ . Zudem gilt  $I_A \neq \{0\}$ . Denn der Vektorraum  $K^{n \times n}$  hat die Dimension  $d = n^2$ , sodass  $(E_n, A, A^2, \dots, A^d)$  linear abhängig ist. Es gibt also eine nichttriviale Nulldarstellung

$$0 = \alpha_0 E_n + \alpha_1 A + \dots + \alpha_d A^d.$$

Also ist  $A$  eine Nullstelle von  $p = \alpha_0 + \alpha_1 X + \dots + \alpha_d X^d$ ,  $1 \leq \deg(p) \leq n^2$ . Der Satz von Cayley-Hamilton besagt stärker, dass ein Polynom  $p \neq 0$  vom Grad  $n$  in  $I_A$  als Element enthalten ist:  $A$  ist Nullstelle des charakteristischen Polynoms  $p_A$ .

### Beispiele

(1) Sei  $A \in K^{2 \times 2}$  mit den Zeilen  $(a, b)$ ,  $(c, d)$ . Wegen  $p_A = X^2 - \text{spur}(A)X + \det(A)$  gilt

$$p_A(A) = \begin{pmatrix} a^2 + bc & ab + bd \\ ac + cd & bc + d^2 \end{pmatrix} - (a + d) \begin{pmatrix} a & b \\ c & d \end{pmatrix} + (ad - bc) E_2 = 0.$$

(2) Ist  $D = \text{diag}(d_1, \dots, d_n)$ , so ist  $p_D = (d_1 - X) \cdot \dots \cdot (d_n - X)$ . Damit ist

$$p_D(D) = (d_1 E_n - D) \dots (d_n E_n - D) = \text{diag}(0, \dots, 0) = 0.$$

- (3) Sind  $A, B \in K^{n \times n}$  ähnlich,  $B = SAS^{-1}$ , so gilt  $B^k = S A^k S^{-1}$  für alle  $k \geq 0$ .  
Ist  $p = \sum_{i \leq k} \alpha_i X^i \in K[X]$ , so gilt

$$p(B) = \sum_{i \leq k} \alpha_i (SAS^{-1})^k = \sum_{i \leq k} \alpha_i S A^k S^{-1} = S p(A) S^{-1}.$$

Damit sind  $p(A)$  und  $p(B)$  wieder ähnlich.

Das eindeutig bestimmte Polynom  $m_A$  mit der Eigenschaft (+) des Satzes heißt das *Minimalpolynom* von  $A$ . Wichtige Eigenschaften dieses Polynoms sind:

$p_A$  und  $m_A$  haben dieselben Nullstellen:

$$\sigma(A) = \{ \lambda \in K \mid m_A(\lambda) = 0 \}.$$

$A$  ist genau dann diagonalisierbar, wenn  $m_A = \prod_{\lambda \in \sigma(A)} (X - \lambda E_n)$ .

### Beispiele

- (1) Für die Matrix

$$A = \begin{pmatrix} 1 & 0 & 3 & 5 \\ 0 & 1 & 1 & 4 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 2 \end{pmatrix} \in \mathbb{R}^{4 \times 4} \text{ gilt } p_A = (X-1)^3(X-2).$$

Also ist  $m_A = (X-1)^k(X-2)$  mit  $k \in \{1, 2, 3\}$ . Einsetzen zeigt, dass  $(A - E_4)(A - 2E_4) \neq 0$ ,  $(A - E_4)^2(A - 2E_4) = 0$ . Ist  $m_A = (X-1)^2(X-2)$ .

- (2) Sei  $P \in K^{n \times n}$  eine *Projektion*, d. h., es gilt  $P^2 = P$ . Dann gilt  $P^2 - P = 0$ , sodass  $X^2 - X = X(X-1) \in I_P$ . Damit ist das Minimalpolynom eines der drei Polynome  $X$ ,  $X-1$ ,  $X^2 - X$ .
- (3) Ein  $A \in K^{n \times n}$  heißt *nilpotent*, falls es ein  $k \geq 1$  gibt mit  $A^k = 0$ . Beispiele liefern alle Dreiecksmatrizen, deren Diagonaleinträge alle null sind, etwa

$$A = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad A^2 = \begin{pmatrix} 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad A^3 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \quad A^4 = 0.$$

Ist  $r$  minimal mit  $A^r = 0$ , so ist  $m_A = X^r$ ,  $\sigma(A) = \{0\}$ ,  $p_A = (-1)^n X^n$ . Ist  $A \neq 0$ , so ist  $r > 1$  und damit  $A$  nicht diagonalisierbar.

### Minimalpolynom eines Endomorphismus

Wie für das charakteristische Polynom kann man auch das Minimalpolynom  $m_f$  eines Endomorphismus  $f: V \rightarrow V$ ,  $V$  endlich-dimensional, definieren: Man setzt  $m_f = m_A$ , wobei  $A$  die darstellende Matrix von  $f$  bzgl. einer beliebigen Basis von  $V$  ist.

## 8.11 Haupträume und Hauptraumzerlegung

**Definition** (*Index, Hauptraum, Hauptvektor*)

*Haupträume von Endomorphismen*

Seien  $V$  ein  $n$ -dimensionaler  $K$ -Vektorraum,  $n \geq 1$ , und  $f: V \rightarrow V$  ein Endomorphismus. Weiter sei  $\lambda \in \sigma(f)$ . Dann setzen wir

$$H_k = H_k(f, \lambda) = \text{Kern}((f - \lambda \text{Id}_V)^k) \quad \text{für alle } k \geq 1,$$

$$i(\lambda) = i(f, \lambda) = \min(\{k \geq 1 \mid H_{k+1} = H_k\}), \quad (\text{Index von } f \text{ bzgl. } \lambda)$$

$$H(f, \lambda) = H_{i(\lambda)}. \quad (\text{Hauptraum von } f \text{ bzgl. } \lambda)$$

Die Elemente von  $H(f, \lambda)$  heißen die *Hauptvektoren* von  $f$  zum Eigenwert  $\lambda$ .

*Haupträume für Matrizen*

Für eine Matrix  $A \in K^{n \times n}$  definieren wir

$$H_k(A, \lambda) = H_k(f_A, \lambda), \quad H(A, \lambda) = H(f_A, \lambda).$$

Das Ziel dieses und des folgenden Abschnitts ist es, die Trigonalisierung eines Endomorphismus, dessen charakteristisches Polynom in Linearfaktoren zerfällt, noch zu verbessern. Wir streben eine Darstellung durch eine obere Dreiecksmatrix an, deren Einträge außerhalb der Diagonalen und der Nebendiagonalen verschwinden. Das entscheidende Hilfsmittel hierzu ist die Verallgemeinerung von Eigenräumen zu Haupträumen. Es gilt

$$\text{Eig}(f, \lambda) = H_1 \subset H_2 \subset \dots \subset H_{i(\lambda)} = H_{i(\lambda)+1} = H(f, \lambda),$$

$$\text{Bild}(f - \lambda \text{Id}_V) \supset \text{Bild}((f - \lambda \text{Id}_V)^2) \supset \dots \supset \text{Bild}((f - \lambda \text{Id}_V)^{i(\lambda)}).$$

Bis zum Index  $i(\lambda)$  liegen strikte Inklusionen vor. Aufgrund der endlichen Dimension von  $V$  muss irgendwann Gleichheit eintreten. Wichtige Eigenschaften sind:

$v \in H_{k+1}$ genau dann, wenn $f(v) - \lambda v \in H_k$	
$f[H_k] \subseteq H_k, \quad f[H(f, \lambda)] \subseteq H(f, \lambda)$	(Invarianz)
$\dim(H(f, \lambda)) = \mu_{\text{pf}}(\lambda)$	
Der Index $i(\lambda)$ ist der Exponent des Linearfaktors $(X - \lambda)$ des Minimalpolynoms $m_f$ .	
Sind $U_1, \dots, U_{i(\lambda)}$ mit $H_k = U_1 \oplus \dots \oplus U_k$ für alle $k$ , so ist $\dim(U_1) \geq \dots \geq \dim(U_{i(\lambda)})$ .	
$V = H(f, \lambda) \oplus \text{Bild}((f - \lambda \text{Id}_V)^{i(\lambda)})$	

Die letzte Eigenschaft ist die Keimzelle eines Beweises von:



**Satz** (Hauptraumzerlegung)

Sei  $f: V \rightarrow V$  wie oben. Zerfällt  $p_f$  in Linearfaktoren, so existiert eine Basis von  $V$  bestehend aus Hauptvektoren von  $f$  und es gilt

$$V = \bigoplus_{\lambda \in \sigma(f)} H(f, \lambda).$$

$$\begin{pmatrix} A(\lambda_1) & & & \\ & A(\lambda_2) & & \\ & & \dots & \\ & & & A(\lambda_m) \end{pmatrix}$$

Fügen wir Basen der verschiedenen Haupträume  $H(f, \lambda_1), \dots, H(f, \lambda_m)$  aneinander, so erhalten wir eine Basis  $\mathcal{A}$  von  $V$ . Aufgrund der Invarianz der Haupträume hat die darstellende Matrix von  $f$  bzgl.  $\mathcal{A}$  eine diagonale Blockform. Es gilt  $A(\lambda_j) \in K^{\mu_j \times \mu_j}$  wobei  $\mu_j = \mu_{p_f}(\lambda_j)$ .

**Beispiele**

- (1) Für die nicht diagonalisierbare  $(2 \times 2)$ -Matrix  $A$  aus 8.4 gilt

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 2 \end{pmatrix}, \quad p_A = (X-1)^2,$$

$$\sigma(A) = \{1\}, \quad \text{Eig}(A, 1) = \text{span}(e_1 - e_2), \quad \text{Kern}((A - E_2)^2) = \text{Kern}(0) = \mathbb{R}^2.$$

Also ist  $i(1) = 2$  und  $H(A, 1) = \mathbb{R}^2$  (wie es nach dem Satz auch sein muss, da nur ein Eigenwert existiert). Jede Basis von  $\mathbb{R}^2$  ist eine Basis aus Hauptvektoren.

- (2) Für  $A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{pmatrix}$  gilt  $p_A = -(X-2)^2(X-3)$ ,  $\sigma(f) = \{2, 3\}$ ,

$$(A - 2E_3)^2 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix} = (A - 2E_3)^3,$$

$$\text{Eig}(A, 2) = \text{span}(e_1), \quad H(A, 2) = \text{span}(e_1, e_2),$$

$$\text{Bild}(A - 2E_3) = \text{span}(e_1, e_2 + e_3), \quad \text{Bild}((A - 2E_3)^2) = \text{span}(e_1 + e_2 + e_3),$$

$$\text{Eig}(A, 3) = H(A, 3) = \text{span}(e_1 + e_2 + e_3), \quad \text{Bild}(A - 3E_3) = \text{span}(e_1, e_2),$$

$$\mathbb{R}^3 = H(A, 2) \oplus H(A, 3) = \text{span}(e_1, e_2) \oplus \text{span}(e_1 + e_2 + e_3).$$

Das Beispiel zeigt, dass

$$H_k(A, \lambda) \cap \text{Bild}((A - \lambda E_n)^k) \neq \{0\}$$

für  $k < i(\lambda)$  gelten kann.

$$\begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} \quad \mathcal{A} = (e_1, e_2, e_1 + e_2 + e_3) \text{ ist eine Basis aus Hauptvektoren. Die Matrix links stellt } f_A \text{ bzgl. } \mathcal{A} \text{ dar.}$$

- (3) Sei  $A \in K^{n \times n}$  diagonalisierbar. Dann gibt es ein  $S \in \text{GL}(n, K)$  mit

$$A = S \text{diag}(\lambda_1, \dots, \lambda_n) S^{-1}, \quad \sigma(A) = \{\lambda_1, \dots, \lambda_n\}.$$

Für alle  $\lambda$  und  $m \geq 1$  gilt  $(A - \lambda E_n)^m = S \text{diag}((\lambda_1 - \lambda)^m, \dots, (\lambda_n - \lambda)^m) S^{-1}$ , sodass  $\text{Kern}((A - \lambda E_n)^m) = \text{Kern}(A - \lambda E_n)$ . Damit ist  $i(\lambda) = 1$  und  $H(A, \lambda) = \text{Eig}(A, \lambda)$  für alle  $\lambda \in \sigma(A)$ .

## 8.12 Die Jordan-Normalform

### Satz (Konstruktion von Jordan-Ketten)

Seien  $V$  ein  $n$ -dimensionaler  $K$ -Vektorraum,  $f \in \text{End}(V)$ ,  $\lambda \in \sigma(f)$  und  $U_1, \dots, U_{i(\lambda)}$  Unterräume von  $V$  mit  $H_k = U_1 \oplus \dots \oplus U_k$  für alle  $1 \leq k \leq i(\lambda)$ . Weiter seien  $k \in \{1, \dots, i(\lambda)\}$  und  $v_k \in U_k$  beliebig. Sind dann  $v_{k-1} \in U_{k-1}, \dots, v_1 \in U_1$  rekursiv definiert durch

$$v_{j-1} = f(v_j) - \lambda v_j$$

so ist  $\mathcal{A} = (v_1, \dots, v_k)$  linear unabhängig und es gilt  $f[W] \subseteq W$  für  $W = \text{span}(\mathcal{A}) \subseteq H_k$ . Die darstellende Matrix  $J_k(\lambda)$  von  $f|_W$  bzgl. der Basis  $\mathcal{A}$  von  $W$  hat die bidiagonale Form rechts.

$$J_k(\lambda) = \begin{pmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & & & & \lambda \end{pmatrix}$$

Man nennt  $(v_1, \dots, v_k)$  die *Jordan-Kette* von  $f$  zum Startwert  $v_k$  und  $J_k(\lambda) \in K^{k \times k}$  einen *Jordan-Block*. Durch Konstruktion von  $r = \dim(\text{Eig}(f, \lambda))$  Jordan-Ketten

$$(v_{1,1}, \dots, v_{1,k(1)}), \dots, (v_{r,1}, \dots, v_{r,k(r)}), \quad i(\lambda) = k(1) \geq \dots \geq k(r) \geq 1,$$

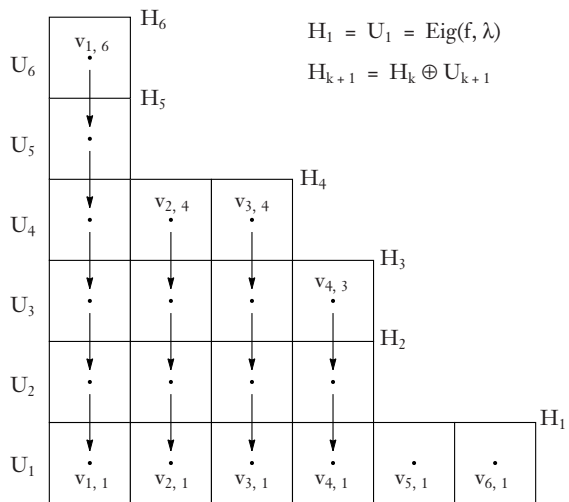
$$k(1) + \dots + k(r) = \mu_{\text{pf}}(\lambda),$$

erhält man eine Basis  $\mathcal{J}(\lambda) =$

$$(v_{1,1}, \dots, v_{1,k(1)}, \dots, v_{r,1}, \dots, v_{r,k(r)})$$

des Hauptraums  $H(f, \lambda)$ . Die Startwerte  $v_{1,k(1)}, \dots, v_{r,k(r)}$  wählt man in  $U_k$  mit  $k \leq i(\lambda)$  so groß wie möglich (sodass jede Jordan-Kette eine Spalte im Diagramm rechts ausfüllt). Die darstellende Matrix  $J(\lambda)$  von  $f|_{H(f, \lambda)}$  bzgl.  $\mathcal{J}(\lambda)$  ist aus Jordan-Blöcken  $J_{k(1)}(\lambda), \dots, J_{k(r)}(\lambda)$  gebildet. Ihre geringe Anzahl an von 0 verschiedenen Einträgen außerhalb der Diagonalen ist optimal in ihrer Ähnlichkeitsklasse. Die in der Rekursion zuletzt konstruierten Glieder  $v_{1,1}, \dots, v_{r,1}$  der Ketten bilden eine Basis des Eigenraums  $\text{Eig}(f, \lambda)$ . Die entsprechenden Spalten der Matrix  $J(\lambda)$  haben genau einen Eintrag  $\lambda$ .

Für alle Eigenwerte  $\lambda$  durchgeführt ergibt sich:



Die Jordan-Ketten werden durch wiederholte Anwendung von  $f - \lambda \text{Id}_V$  auf frei gewählte Startwerte konstruiert. Dabei fällt man in jedem Schritt von  $U_k$  nach  $U_{k-1}$ . Die Länge der Ketten ist durch die Dimensionen der  $U_k$  festgelegt. Im Beispiel des Diagramms ist die algebraische Vielfachheit von  $\lambda$  gleich 19 und  $J(\lambda)$  hat eine 6-4-4-3-1-1-Form, mit  $5 + 3 + 3 + 2 + 0 + 0 = 13 = 19 - \dim(\text{Eig}(f, \lambda))$  Einsen in der Nebendiagonalen.

**Satz** (*Jordan-Normalform*)

Sei  $f: V \rightarrow V$  wie oben. Zerfällt  $p_f$  in Linearfaktoren, so existiert eine Basis  $\mathcal{J}$  von  $V$  derart, dass die darstellende Matrix  $J$  von  $f$  bzgl.  $\mathcal{J}$  die folgende Form hat:

$$J = \begin{pmatrix} J(\lambda_1) & & \\ & \dots & \\ & & J(\lambda_m) \end{pmatrix}, \quad \sigma(f) = \{\lambda_1, \dots, \lambda_m\}, \quad \lambda_i \neq \lambda_j \text{ für } i \neq j.$$

**Beispiel**

Wir untersuchen  $f_A: \mathbb{R}^6 \rightarrow \mathbb{R}^6$  für die Matrix  $A$ .

*Berechnung des charakteristischen Polynoms*

$$p_A = (X - 1)^6, \quad \sigma(A) = \{1\}, \quad \mu_{p_A}(1) = 6$$

*Schrittweise Berechnung des Hauptraums*

$$H_k = \text{Kern}((A - E_6)^k)$$

$$H_1 = U_1 \quad U_1 = \text{span}(e_1 + e_4, e_1 + e_3 - e_5)$$

$$H_2 = H_1 \oplus U_2 \quad U_2 = \text{span}(e_1 + e_5, e_1 - e_6)$$

$$H_3 = H_2 \oplus U_3 \quad U_3 = \text{span}(2e_1 - e_2)$$

$$H_4 = H_3 \oplus U_4 \quad U_4 = \text{span}(e_4)$$

Aus den Dimensionen der  $U_k$  ergibt sich, dass  $J(1)$  eine 4-2-Blockform hat.

*Bildung von Jordan-Ketten, Start in  $U_4$  und  $U_2$*

$$v_4 = e_4 = (0, 0, 0, 1, 0, 0)$$

$$v_3 = Av_4 - v_4 = (0, 1, 1, 0, 1, 1)$$

$$v_2 = Av_3 - v_3 = (1, 0, 0, 0, 1, 0)$$

$$v_1 = Av_2 - v_2 = (1, 0, 0, 1, 0, 0)$$

$$w_2 = e_1 - e_6 = (1, 0, 0, 0, 0, -1)$$

$$w_1 = Aw_2 - w_2 = (0, 0, -1, 1, 1, 0)$$

*Basis und Transformationsmatrix*

$$\mathcal{J} = (v_1, v_2, v_3, v_4, w_1, w_2)$$

$$S^{-1} = \begin{pmatrix} v_1 & v_2 & v_3 & v_4 & w_1 & w_2 \end{pmatrix}$$

$$J = J(1) = S A S^{-1}$$

$$A = \begin{pmatrix} 1 & -1 & 1 & 0 & 1 & 0 \\ -1 & -1 & 2 & 1 & 1 & -1 \\ -1 & -3 & 3 & 1 & 1 & 0 \\ 0 & -1 & 1 & 1 & 1 & -1 \\ -1 & 0 & 2 & 1 & 2 & -2 \\ -1 & -2 & 2 & 1 & 1 & 0 \end{pmatrix}$$

$$A - E_6 = \begin{pmatrix} 0 & -1 & 1 & 0 & 1 & 0 \\ -1 & -2 & 2 & 1 & 1 & -1 \\ -1 & -3 & 2 & 1 & 1 & 0 \\ 0 & -1 & 1 & 0 & 1 & -1 \\ -1 & 0 & 2 & 1 & 1 & -2 \\ -1 & -2 & 2 & 1 & 1 & -1 \end{pmatrix}$$

$$(A - E_6)^2 = \begin{pmatrix} -1 & -1 & 2 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ -1 & -2 & 2 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$(A - E_6)^3 = \begin{pmatrix} -1 & -2 & 2 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & -2 & 2 & 1 & 1 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$(A - E_6)^4 = 0$$

$$J(1) = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$



---

# Überblick und Zusammenfassung

---

## 1. Algebraische Grundstrukturen

Für eine Operation  $\circ : G^2 \rightarrow G$  auf einer Menge  $G$  betrachten wir die Eigenschaften:

(1)	$\forall a, b, c \quad a \circ (b \circ c) = (a \circ b) \circ c$	Assoziativgesetz
(2)	$\exists e \forall x \quad x \circ e = e \circ x = x$	Existenz eines neutralen Elements
(3)	$\forall a \exists b \quad a \circ b = b \circ a = e$	Existenz inverser Elemente
(4)	$\forall a, b \quad a \circ b = b \circ a$	Kommutativgesetz

Die Quantoren beziehen sich dabei auf Elemente in  $G$ . In (3) ist  $e$  ein neutrales Element von  $G$  wie in (2).

$(G, \circ)$ oder kurz $G$ heißt ...	falls gilt:
Halbgruppe	(1)
Monoid	(1), (2)
Gruppe	(1), (2), (3)
kommutativ oder abelsch	(4)

Für zwei Operationen  $+, \cdot : R^2 \rightarrow R$  auf einer Menge  $R$  betrachten wir:

(D)	$\forall a, b, c \quad a(b + c) = ab + ac$	erstes Distributivgesetz
	$\forall a, b, c \quad (b + c)a = ba + ca$	zweites Distributivgesetz

Ist  $(R, +)$  ein Monoid mit neutralem Element  $0$ , so sei  $R^* = R - \{0\}$ .

$(R, +, \cdot)$ oder kurz $R$ heißt ...	falls gilt:
Ring (mit Eins)	$(R, +)$ ist abelsche Gruppe, $(R^*, \cdot)$ ist Monoid, (D)
Schiefkörper	$(R, +)$ ist abelsche Gruppe, $(R^*, \cdot)$ ist Gruppe, (D)
Körper	$(R, +)$ ist abelsche Gruppe, $(R^*, \cdot)$ ist abelsche Gruppe, (D)

## 2. Die Kongruenz modulo m

### Teilbarkeit in den ganzen Zahlen

Eine ganze Zahl  $a$  heißt *teilbar* durch eine ganze Zahl  $m$ , in Zeichen  $m \mid a$  (gelesen:  $m$  ist ein Teiler von  $a$ ), falls es eine ganze Zahl  $d$  gibt mit  $dm = a$ .

### Kongruenz modulo m

Sei  $m \geq 1$  eine natürliche Zahl. Zwei ganze Zahlen  $a$  und  $b$  heißen *kongruent modulo m*, falls  $m \mid (a - b)$ . Dies ist gleichbedeutend damit, dass  $a$  und  $b$  denselben Rest bei Division durch  $m$  haben. Sind  $a$  und  $b$  kongruent modulo  $m$ , so schreiben wir

$$a \equiv_m b \quad \text{oder} \quad a \equiv b \pmod{m}.$$

Wir setzen:

$$[a] = [a]_m = a / \equiv_m = \{ b \in \mathbb{Z} \mid b \equiv_m a \} = \{ \dots, a - 2m, a - m, a, a + m, a + 2m, \dots \},$$

*(Restklasse von a modulo m)*

$$\mathbb{Z}/m\mathbb{Z} = \mathbb{Z}_m = \{ [a] \mid a \in \mathbb{Z} \} = \{ [0], \dots, [m-1] \},$$

$$[a] + [b] = [a + b], \quad [a] \cdot [b] = [a \cdot b] \quad \text{für alle } a, b \in \mathbb{Z}.$$

### Rechenregeln

Für alle  $a, b, c, x, y \in \mathbb{Z}$  gelten:

- (a)  $a \equiv_m a, \quad a \equiv_m b$  *genau dann, wenn*  $b \equiv_m a$ ,  
 $a \equiv_m b$  und  $b \equiv_m c$  *impliziert*  $a \equiv_m c$ ,
- (b)  $a \equiv_m b$  *genau dann, wenn*  $a - b \equiv_m 0$ ,
- (c)  $a \equiv_m b$  und  $c \equiv_m d$  *impliziert*  $xa + yc \equiv_m xb + yd$  und  $ac \equiv_m bd$ ,
- (d)  $a \equiv_m b$  *impliziert*  $p(a) \equiv_m p(b)$  für jedes Polynom  $p \in \mathbb{Z}[X]$ .

### Algebraische Eigenschaften

Die Menge  $\mathbb{Z}_m$  bildet mit den Restklassenoperationen  $+$  und  $\cdot$ :

- (a) eine abelsche Gruppe  $(\mathbb{Z}_m, +)$ ,
- (b) einen Ring  $(\mathbb{Z}_m, +, \cdot)$  mit  $0 = [0]$  und  $1 = [1]$ ,
- (c) genau dann eine abelsche Gruppe  $(\mathbb{Z}_m - \{ [0] \}, \cdot)$ , wenn  $m$  prim ist,
- (d) genau dann einen Körper  $(\mathbb{Z}_m, +, \cdot)$ , wenn  $m$  prim ist.

### 3. Matrizen

#### Schreibweisen für Matrizen

$$A = (a_{i,j})_{1 \leq i \leq m, 1 \leq j \leq n} = (a_{ij})_{ij} = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \dots & \dots & \dots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \in K^{m \times n}$$

$A(i, j) = a_{ij}$  = „Eintrag in der i-ten Zeile und j-ten Spalte“

Nichtspezifizierte Einträge sind 0.

$(v_1 \dots v_n) = (v_1; \dots; v_n) \in K^{m \times n}$  hat die Spalten  $v_1, \dots, v_n \in K^m$ .

$(v_1, \dots, v_m) \in K^{m \times n}$  hat die Zeilen  $v_1, \dots, v_m \in K^n$ .

#### Die Matrizenmultiplikation

Für  $A \in K^{m \times r}$ ,  $B \in K^{r \times n}$  ist das Produkt  $C = AB \in K^{m \times n}$  definiert durch

$$c_{ij} = \sum_{1 \leq k \leq r} a_{ik} b_{kj} = a_{i1} b_{1j} + \dots + a_{ir} b_{rj} \quad (i\text{-te Zeile mal } j\text{-te Spalte}).$$

#### n-Tupel als einspaltige Matrizen

$$(x_1, \dots, x_n) \in K^n \text{ wird identifiziert mit } \begin{pmatrix} x_1 \\ \dots \\ x_n \end{pmatrix} \in K^{n \times 1}$$

Das Matrix-Vektor-Produkt  $Ax = A(x_1, \dots, x_n)$  ist damit ein Matrizenprodukt.

#### Die Transposition

$$A^t = (a_{ji})_{ij} = \begin{pmatrix} a_{11} & \dots & a_{m1} \\ \dots & \dots & \dots \\ a_{1n} & \dots & a_{mn} \end{pmatrix} \in K^{n \times m}$$

$$(x_1, \dots, x_n)^t = \begin{pmatrix} x_1 \\ \dots \\ x_n \end{pmatrix}^t = (x_1 \dots x_n) \in K^{1 \times n}$$



### Konjugierte und adjungierte Matrizen

$$\overline{A} = (\overline{a_{ij}})_{ij}, \quad A^* = \overline{A}^t = (\overline{a_{ji}})_{ij} = \begin{pmatrix} \overline{a_{11}} & \dots & \overline{a_{m1}} \\ \dots & \dots & \dots \\ \overline{a_{1n}} & \dots & \overline{a_{mn}} \end{pmatrix} \text{ für } A \in \mathbb{C}^{m \times n}$$

Man erklärt diese Operationen auch für reelle Matrizen:

$$\overline{A} = A, \quad A^* = A^t \text{ für } A \in \mathbb{R}^{m \times n}$$

### Spezielle quadratische Matrizen für einen fest gewählten Körper K

$$\text{diag}(d_1, \dots, d_n) = (d_i \delta_{ij})_{ij} = \begin{pmatrix} d_1 & & & \\ & d_2 & & \\ & & \dots & \\ & & & d_n \end{pmatrix} \in K^{n \times n}$$

$$E_n = (\delta_{ij})_{ij} = \text{diag}(1, \dots, 1) \in K^{n \times n}$$

$$\text{obere Dreiecksmatrix: } \begin{pmatrix} a_{11} & \dots & \dots & a_{1n} \\ & a_{22} & \dots & a_{2n} \\ & & \dots & \dots \\ & & & a_{nn} \end{pmatrix} \in K^{n \times n}$$

$$\text{untere Dreiecksmatrix: } \begin{pmatrix} a_{11} & & & \\ a_{21} & a_{22} & & \\ \dots & \dots & \dots & \\ a_{n1} & \dots & \dots & a_{nn} \end{pmatrix} \in K^{n \times n}$$

### Elementarmatrizen für K und $n \geq 1$ fest gewählt

$$\text{Additionstyp: } W_{ij}(\lambda) = E_n + \lambda E_{i,j}, \quad i \neq j$$

$$\text{Multiplikationstyp: } W_{ii}(\lambda) = E_n + (\lambda - 1) E_{i,i}$$

wobei  $E_{i,j} =$  „die  $n \times n$ -Matrix über K mit Eintrag 1 an der Stelle  $(i, j)$  und 0 sonst“

Bemerkung: Manchmal werden nur die Additionstypen als Elementarmatrizen betrachtet oder auch weitere Typen dazugerechnet, etwa Transpositionsmatrizen.

## 4. Matrizen und lineare Abbildungen

### Die einer Matrix zugeordnete lineare Abbildung

Für  $A \in K^{m \times n}$  ist  $f_A : K^n \rightarrow K^m$  definiert durch  $f_A(x) = Ax$  für alle  $x \in K^n$ .

### Die darstellende Matrix einer linearen Abbildung

Für  $f : V \rightarrow W$  linear und Basen  $\mathcal{A} = (v_1, \dots, v_n)$ ,  $\mathcal{B} = (w_1, \dots, w_m)$  von  $V$  bzw.  $W$  ist die darstellende Matrix  $A_f^{\mathcal{A}, \mathcal{B}} = „A_f \text{ bzgl. } \mathcal{A}, \mathcal{B}“$  definiert durch

$$A_f^{\mathcal{A}, \mathcal{B}} = \begin{pmatrix} \Phi_{\mathcal{B}}(f(v_1)) & \dots & \Phi_{\mathcal{B}}(f(v_n)) \end{pmatrix} \in K^{m \times n},$$

wobei  $\Phi_{\mathcal{B}}(w) = (\alpha_1, \dots, \alpha_m) \in K^m$  der Koordinatenvektor von  $w \in W$  bzgl.  $\mathcal{B}$  ist, d. h.

$$w = \alpha_1 w_1 + \dots + \alpha_m w_m.$$

Kurz: Die Spalten von  $A_f^{\mathcal{A}, \mathcal{B}}$  sind die Koordinaten (bzgl.  $\mathcal{B}$ ) der Bilder der Vektoren von  $\mathcal{A}$ .

Wir schreiben oft „ $A_f$  bzgl.  $\mathcal{A}$ “ statt „ $A_f$  bzgl.  $\mathcal{A}, \mathcal{A}$ “ sowie  $A_f^{\mathcal{A}}$  statt  $A_f^{\mathcal{A}, \mathcal{A}}$ .

Im Fall  $W = K^m$  kann man  $\Phi_{\mathcal{B}}(y)$  berechnen, indem man die Vektoren  $w_1, \dots, w_m$  der Basis  $\mathcal{B}$  als Spalten in eine Matrix  $S^{-1}$  schreibt und die zu  $S^{-1}$  inverse Matrix  $S$  bestimmt (vgl. 5.5). Dann gilt  $\Phi_{\mathcal{B}}(y) = Sy$  für alle  $y \in K^m$ , sodass

$$A_f^{\mathcal{A}, \mathcal{B}} = \begin{pmatrix} S f(v_1) & \dots & S f(v_n) \end{pmatrix} = S \begin{pmatrix} f(v_1) & \dots & f(v_n) \end{pmatrix}.$$

Darstellende Matrizen im Fall $V = K^n$ und $W = K^m$
Für $f : K^n \rightarrow K^m$ linear und die Standardbasen $\mathcal{A}, \mathcal{B}$ gilt $A_f^{\mathcal{A}, \mathcal{B}} = \begin{pmatrix} f(e_1) & \dots & f(e_n) \end{pmatrix}$ .
<p>Für <math>f = f_A : K^n \rightarrow K^m</math>, <math>T^{-1} = \begin{pmatrix} v_1 &amp; \dots &amp; v_n \end{pmatrix}</math>, <math>S^{-1} = \begin{pmatrix} w_1 &amp; \dots &amp; w_m \end{pmatrix}</math> gilt</p> $A_f^{\mathcal{A}, \mathcal{B}} = S \begin{pmatrix} f(v_1) & \dots & f(v_n) \end{pmatrix} = S \begin{pmatrix} A v_1 & \dots & A v_n \end{pmatrix} = S A T^{-1}.$ <p>Die Matrizen <math>A</math> und <math>A_f^{\mathcal{A}, \mathcal{B}}</math> sind also äquivalent.</p>
<p>Für <math>f = f_A : K^n \rightarrow K^n</math>, <math>S^{-1} = \begin{pmatrix} v_1 &amp; \dots &amp; v_n \end{pmatrix}</math> gilt</p> $A_f^{\mathcal{A}} = S A S^{-1}.$ <p>Die Matrizen <math>A</math> und <math>A_f^{\mathcal{A}}</math> sind also ähnlich.</p>

## 5. Umformungen mit Elementarmatrizen

Seien  $K$  ein Körper,  $n \geq 1$  und  $A \in GL(n, K)$ .

### Umformung mit Additionstypen: Diagonalisierung

Es gibt Additionstypen  $L_1, \dots, L_k$  derart, dass  $L_k \dots L_1 A$  eine Diagonalmatrix ist. Analog gibt es Additionstypen  $R_1, \dots, R_k$  derart, dass  $A R_1 \dots R_k$  diagonal ist.

### Umformung mit Additions- und Multiplikationstypen: Invertierung

Es gibt Additionstypen  $L_1, \dots, L_k$  und Multiplikationstypen  $M_\ell, \dots, M_1$  derart, dass

$$M_\ell \dots M_1 L_k \dots L_1 A = E_n.$$

Analoges gilt für die Multiplikation von rechts.

### Umformung mit unteren Additionstypen und Permutationen: LR-Zerlegung

Es gibt Additionstypen  $L_1, \dots, L_k$  und eine Permutationsmatrix  $P$  mit

- (a)  $L_i$  ist eine untere Dreiecksmatrix für alle  $i$ .
- (b)  $L_k \dots L_1 A P = R$  mit einer oberen Dreiecksmatrix  $R$ .

Für die untere Dreiecksmatrix  $L = (L_k \dots L_1)^{-1}$  gilt dann

$$A P = L R \quad (LR\text{-Zerlegung})$$

Jede invertierbare Matrix kann also nach einer geeigneten Spaltenvertauschung als Produkt einer unteren und einer oberen Dreiecksmatrix geschrieben werden.

Die Ergebnisse lassen sich durch verschiedene Strategien des Ausräumens der Matrix  $A$  beschreiben. Hierbei spielen die Einträge auf der Diagonalen eine wichtige Rolle:

Sei  $B$  die durch das Ausräumen der Spalten  $1, \dots, i-1$  unterhalb der Diagonalen produzierte Matrix. Ist  $b_{ii} \neq 0$ , so räumen wir unterhalb von  $(i, i)$  mit Hilfe von unteren Dreiecksmatrizen  $W_{ki}(\lambda)$ ,  $k > i$ , aus. Ist  $b_{ii} = 0$ , so haben wir zwei Möglichkeiten:

- (1) Wir bringen einen Eintrag  $b_{ki} \neq 0$ ,  $k > i$ , unterhalb von  $(i, i)$  an die Stelle  $(i, i)$  (Multiplikation mit einer oberen Dreiecksmatrix  $W_{ik}(\lambda)$  von links).
- (2) Wir bringen einen Eintrag  $b_{ij} \neq 0$ ,  $i < j$ , rechts von  $(i, i)$  an die Stelle  $(i, i)$  (Multiplikation mit einer Transpositionsmatrix  $P_{ij}$  von rechts).

Die erste Version führt zu einer oberen Dreiecksmatrix  $L_m \dots L_1 A$ . Analoges Ausräumen oberhalb der Diagonalen liefert die Diagonalisierung. Höchstens  $n$  Multiplikationen mit Multiplikationstypen  $M_i$  verwandeln die Diagonalmatrix schließlich in  $E_n$ .

Die zweite Version ergibt die LR-Zerlegung. Sie benötigt im Allgemeinen eine Permutation der Spalten (Umbenennung der Variablen), kommt dafür aber mit unteren Dreiecksmatrizen  $L_i$  aus. Die LR-Zerlegung ist beim numerischen Lösen von linearen Gleichungssystemen von Interesse.

## 6. Matrizengruppen

Die allgemeinen linearen Gruppen lassen sich in vielerlei Weise charakterisieren. Für einen Körper  $K$  und  $n \geq 1$  gilt beispielsweise

$$\begin{aligned} \mathrm{GL}(n, K) &= \{ A \in K^{n \times n} \mid A \text{ ist invertierbar} \} = \{ A \in K^{n \times n} \mid \mathrm{rang}(A) = n \} = \\ &= \{ A \in K^{n \times n} \mid \text{die Zeilen (Spalten) von } A \text{ bilden eine Basis von } K^n \} = \\ &= \{ A \in K^{n \times n} \mid f_A : K^n \rightarrow K^n \text{ ist bijektiv} \} = \{ A \in K^{n \times n} \mid \det(A) \neq 0 \} = \\ &= \{ A \in K^{n \times n} \mid \text{für alle } b \in K^n \text{ ist } Ax = b \text{ eindeutig lösbar} \}. \end{aligned}$$

In der linearen Algebra spielen Untergruppen von  $\mathrm{GL}(n, K)$  eine wichtige Rolle. Wir betrachten einige von ihnen für  $K = \mathbb{R}$  und ein beliebiges  $n \geq 1$ .

### Diagonalmatrizen

Die invertierbaren Diagonalmatrizen

$$\mathrm{Diag}(n) = \{ \mathrm{diag}(a_1, \dots, a_n) \in \mathbb{R}^{n \times n} \mid a_1, \dots, a_n \in \mathbb{R}^* \}$$

bilden eine abelsche Untergruppe von  $\mathrm{GL}(n, \mathbb{R})$ .

### Dreiecksmatrizen

Die invertierbaren unteren Dreiecksmatrizen

$$\Delta_u(n, \mathbb{R}) = \{ A \in \mathbb{R}^{n \times n} \mid a_{ij} = 0 \text{ für alle } i < j, a_{ii} \neq 0 \text{ für alle } i \}$$

sind eine Untergruppe von  $\mathrm{GL}(n, \mathbb{R})$ . Eine Untergruppe von  $\Delta_u(n, \mathbb{R})$  ist

$$\{ A \in \Delta_u(n, \mathbb{R}) \mid a_{ii} = 1 \text{ für alle } i \} = \{ A \in \Delta_u(n, \mathbb{R}) \mid A \text{ ist unipotent} \},$$

wobei ein  $A \in \mathbb{R}^{n \times n}$  *unipotent* heißt, falls  $(A - E_n)^n = 0$ . Die Additionstypen  $W_{ij}(\lambda)$ , die bei der Gauß-Elimination verwendet werden, sind beispielsweise unipotent. Analoges gilt für die invertierbaren oberen Dreiecksmatrizen  $\Delta_o(n, \mathbb{R})$ .

### Die orthogonale Gruppe

Die orthogonalen Matrizen

$$\mathrm{O}(n) = \{ A \in \mathbb{R}^{n \times n} \mid A A^t = E_n = A^t A \} = \{ A \in \mathrm{GL}(n, \mathbb{R}) \mid A^{-1} = A^t \}$$

sind eine Untergruppe von  $\mathrm{GL}(n, \mathbb{R})$ . Für alle  $A \in \mathrm{O}(n)$  gilt  $|\det(A)| = 1$ .

### Die spezielle lineare Gruppe

Die spezielle lineare Gruppe

$$\mathrm{SL}(n, \mathbb{R}) = \{ A \in \mathbb{R}^{n \times n} \mid \det(A) = 1 \}$$

ist der Kern von  $\det : \mathrm{GL}(n, \mathbb{R}) \rightarrow \mathbb{R}^*$  und somit ein Normalteiler von  $\mathrm{GL}(n, \mathbb{R})$ . Sie wird von den Additionstypen  $W_{ij}(\lambda)$  erzeugt (vgl. 7.8). Schließlich ist die spezielle orthogonale Gruppe definiert durch

$$\mathrm{SO}(n) = \mathrm{O}(n) \cap \mathrm{SL}(n, \mathbb{R}).$$

## 7. Matrixzerlegungen

Wir diskutieren einige wichtige Zerlegungen  $A = BC$  einer Matrix  $A \in GL(n, \mathbb{C})$ . Mit den üblichen Änderungen (transponiert/adjungiert, symmetrisch/hermitesch, orthogonal/unitär) ergeben sich analoge Zerlegungen für reelle invertierbare Matrizen.

### LR-Zerlegung und Cholesky-Zerlegung

Das Gauß-Eliminationsverfahren liefert eine Zerlegung

$$AP = LR$$

mit einer Permutationsmatrix  $P$ , unteren Dreiecksmatrix  $L$  und oberen Dreiecksmatrix  $R$  (vgl. 5.12 und Überblick 5). Ist  $A$  hermitesch und positiv definit, so kann  $P = E_n$  gewählt werden und die LR-Zerlegung vereinfacht sich dann zur Cholesky-Zerlegung  $A = LL^*$  (vgl. 7.11).

### Die QR-Zerlegung

Das Verfahren von Gram-Schmidt (vgl. 6.7) liefert für eine Basis  $(a_1, \dots, a_n)$  des  $\mathbb{C}^n$  eine Orthonormalbasis  $(q_1, \dots, q_n)$  bzgl. des kanonischen Skalarprodukts mit

$$(+)\ \text{span}(a_1, \dots, a_k) = \text{span}(q_1, \dots, q_k) \text{ für alle } k \leq n.$$

Sind  $a_1, \dots, a_n$  die Spalten von  $A$ , so erhalten wir die Zerlegung

$$A = QR$$

mit einer unitären Matrix  $Q$  und einer oberen Dreiecksmatrix  $R$ . Dabei hat  $Q$  die Spalten  $q_1, \dots, q_n$ . Dass  $R = Q^*A$  eine obere Dreiecksmatrix ist, folgt aus (+).

### Die Polarzerlegungen

Ist  $SAT^{-1} = SAT^* = \text{diag}(\sigma_1, \dots, \sigma_n) = D$  die Singulärwertzerlegung von  $A$  (vgl. 8.8) mit  $S, T \in U(n)$  und positiven Singulärwerten  $\sigma_k$ , so gilt

$$(a)\ A = S^*DT,$$

$$(b)\ Q = S^*T \text{ ist unitär,}$$

$$(c)\ P_1 = S^*DS \text{ und } P_2 = TDT^* \text{ sind hermitesch und positiv definit.}$$

Damit ergeben sich die *linke* bzw. *rechte Polarzerlegung* von  $A$ :

$$A = (S^*DS)(S^*T) = P_1Q, \quad A = (S^*T)(T^*DT) = QP_2.$$

Die Determinante von  $A$  berechnet sich zu  $\det(A) = (\sigma_1 \dots \sigma_n) \det(Q)$ , was wir wegen  $\sigma_1, \dots, \sigma_n > 0$  und  $|\det(Q)| = 1$  als Verallgemeinerung der Polardarstellung der komplexen Zahl  $\det(A)$  lesen können.

## 8. Die Sesquilinearformen $\langle \cdot, A \cdot \rangle$ und positive Definitheit

Mit den kanonischen Skalarprodukten gilt:

$$\langle x, Ay \rangle = \sum_{1 \leq i, j \leq n} x_i a_{ij} y_j \quad \text{für alle } x, y \in \mathbb{R}^n \text{ und } A \in \mathbb{R}^{n \times n},$$

$$\langle z, Aw \rangle = \sum_{1 \leq i, j \leq n} \overline{z_i} a_{ij} w_j \quad \text{für alle } z, w \in \mathbb{C}^n \text{ und } A \in \mathbb{C}^{n \times n}.$$

Oft gebraucht wird:

Seitenwechsel
$\langle x, Ay \rangle = x^t (A y) = (x^t A) y = (A^t x)^t y = \langle A^t x, y \rangle \quad \text{für } A \in \mathbb{R}^{n \times n}$ $\langle z, Aw \rangle = z^* (A w) = (z^* A) w = (A^* z)^* w = \langle A^* z, w \rangle \quad \text{für } A \in \mathbb{C}^{n \times n}$
$\langle x, Ay \rangle = \langle Ax, y \rangle \quad \text{für } A \in \mathbb{R}^{n \times n} \text{ symmetrisch } (A = A^t)$ $\langle z, Aw \rangle = \langle Az, w \rangle \quad \text{für } A \in \mathbb{C}^{n \times n} \text{ hermitesch } (A = A^*)$

Eine symmetrische bzw. hermitesche Matrix  $A$  ist positiv definit, wenn die Sesquilinearform  $\langle \cdot, A \cdot \rangle : K^n \times K^n \rightarrow K$  dies ist, d.h. wenn  $\langle x, Ax \rangle > 0$  für alle  $x \neq 0$  (vgl. 6.12, 7.11). Der Seitenwechsel liefert:

### Positive Definitheit kongruenter Matrizen

Ist  $A \in \mathbb{R}^{n \times n}$  symmetrisch und  $S \in GL(n, K)$ , so gilt

$$\langle x, S^t A S x \rangle = \langle Sx, A(Sx) \rangle \quad \text{für alle } x \in \mathbb{R}^n.$$

Mit  $A$  ist also auch  $S^t A S$  positiv definit (da  $y = Sx$  mit  $x$  alle Vektoren des  $\mathbb{R}^n - \{0\}$  durchläuft). Analog ist für eine positiv definite Matrix  $A \in \mathbb{C}^{n \times n}$  und  $S \in GL(n, \mathbb{C})$  auch  $S^* A S$  positiv definit. Kurz:

Positive Definitheit vererbt sich auf kongruente Matrizen.

### Eigenwertkriterium der Definitheit

Eine symmetrische oder hermitesche Matrix  $A$  ist genau dann positiv definit, wenn alle Eigenwerte  $\lambda_1, \dots, \lambda_n$  von  $A$  positiv sind. Denn nach dem Spektralsatz ist  $A$  kongruent zur Diagonalmatrix  $\text{diag}(\lambda_1, \dots, \lambda_n)$ . (Dies kann man auch direkt mit Hilfe einer Orthonormalbasis aus Eigenvektoren von  $A$  einsehen, vgl. 8.7.)

### Positive Definitheit von $A^t A$ bzw. $A^* A$

Für alle  $A \in GL(n, \mathbb{R})$  ist  $A^t A$  positiv definit, da  $\langle x, A^t A x \rangle = \langle Ax, Ax \rangle > 0$  für alle  $x \neq 0$ . Analog ist  $A^* A$  positiv definit für alle  $A \in GL(n, \mathbb{C})$ .

## 9. Quadriken in Normalform für $n = 2$

Die Quadriken  $Q(q) = \{ (x, y) \mid q(x, y) = 0 \} \neq \emptyset$  mit  $q : \mathbb{R}^2 \rightarrow \mathbb{R}$  in Normalform sind:

Typ 1  $q(x, y) = \lambda x^2 + \mu y^2 + c, \quad \lambda, \mu \neq 0$

Typ 2  $q(x, y) = \lambda x^2 + by, \quad \lambda \neq 0, b \neq 0$

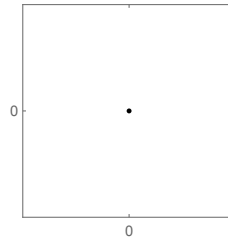
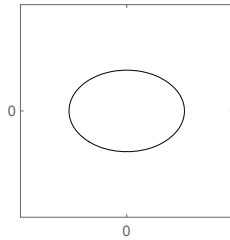
Typ 3  $q(x, y) = \lambda x^2 + c, \quad \lambda \neq 0$

Ellipse

Typ 1

$$\operatorname{sgn}(\lambda) = \operatorname{sgn}(\mu)$$

$$\operatorname{sgn}(c) = -\operatorname{sgn}(\lambda)$$



Punkt

Typ 1

$$\operatorname{sgn}(\lambda) = \operatorname{sgn}(\mu)$$

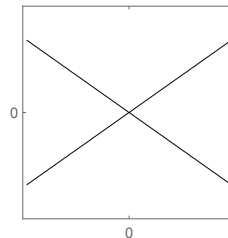
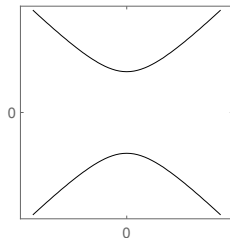
$$c = 0$$

Hyperbel

Typ 1

$$\operatorname{sgn}(\lambda) \neq \operatorname{sgn}(\mu)$$

$$c \neq 0$$



Kreuzende Geraden

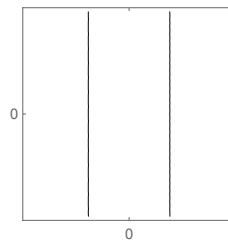
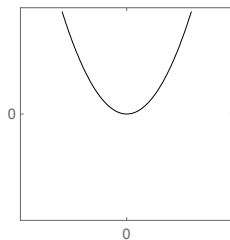
Typ 1

$$\operatorname{sgn}(\lambda) \neq \operatorname{sgn}(\mu)$$

$$c = 0$$

Parabel

Typ 2

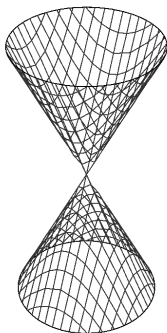


Parallele Geraden

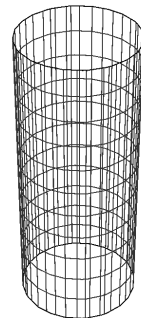
Typ 3

$$\operatorname{sgn}(\lambda) \neq \operatorname{sgn}(c)$$

Doppelgerade im  
Fall  $c = 0$



In den ersten fünf Fällen und im sechsten Fall mit  $c = 0$  ist die Quadrik  $Q(q)$  ein Kegelschnitt, d. h. der Schnitt eines Doppelkegels mit einer Ebene. Doppelgeraden treten beim Schnitt eines Zylinders („Kegelspitze im Unendlichen“) mit einer Ebene auf.



## 10. Normalformen

Wir stellen einige Normalformen von Matrizen zusammen, d.h., wir geben möglichst einfache Repräsentanten für wichtige Äquivalenzrelationen an. Dabei beschränken wir uns zunächst auf Matrizen in  $\mathbb{C}^{n \times n}$ ,  $n \geq 1$ .

### Äquivalente Matrizen

#### Definition

**|**  $A, B \in \mathbb{C}^{n \times n}$  sind *äquivalent*, falls  $S, T \in GL(n, \mathbb{C})$  existieren mit  $B = S A T^{-1}$ .

#### Bedeutung

$A$  und  $B$  stellen dieselbe lineare Abbildung bzgl. verschiedener Basen dar, d.h., es gibt ein lineares  $f: \mathbb{C}^n \rightarrow \mathbb{C}^n$  und Basen  $\mathcal{A}, \mathcal{B}, \mathcal{A}', \mathcal{B}'$  des  $\mathbb{C}^n$  mit  $A = A_f^{\mathcal{A}, \mathcal{B}}$ ,  $B = A_f^{\mathcal{A}', \mathcal{B}'}$ .

#### Normalformen

Sei  $A \in \mathbb{C}^{n \times n}$ , und sei  $r = \text{rang}(A)$ . Dann gibt es  $S, T \in GL(n, \mathbb{C})$  mit

$$S A T^{-1} = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}. \quad (\text{Normalformdarstellung})$$

Weiter gibt es  $S, T \in U(n)$  und  $\sigma_1, \dots, \sigma_r > 0$  mit

$$S A T^{-1} = \begin{pmatrix} \text{diag}(\sigma_1, \dots, \sigma_r) & 0 \\ 0 & 0 \end{pmatrix}. \quad (\text{Singularwertzerlegung})$$

### Kongruente Matrizen

#### Definition

**|**  $A, B \in \mathbb{C}^{n \times n}$  sind *kongruent*, falls ein  $S \in GL(n, \mathbb{C})$  existiert mit  $B = S^* A S$ .

#### Bedeutung

$A$  und  $B$  stellen dieselbe Sesquilinearform bzgl. einer Basis des  $\mathbb{C}^n$  dar, d.h., es gibt eine Sesquilinearform  $\varphi: \mathbb{C}^n \times \mathbb{C}^n \rightarrow \mathbb{C}$  und Basen  $\mathcal{A}, \mathcal{B}$  mit

$$\varphi(x, y) = \langle \Phi_{\mathcal{A}}(x), A \Phi_{\mathcal{A}}(y) \rangle_{\text{kanonisch}} = \langle \Phi_{\mathcal{B}}(x), B \Phi_{\mathcal{B}}(y) \rangle_{\text{kanonisch}} \quad \text{für alle } x, y \in \mathbb{C}^n,$$

#### Normalform (für hermitesche Matrizen)

Ist  $A$  hermitesch, so existiert ein  $S \in GL(n, \mathbb{C})$  mit

$$S^* A S = \begin{pmatrix} E_{s^+} & & \\ & -E_{s^-} & \\ & & 0 \end{pmatrix}, \quad (\text{Hauptachsentransformation, Trägheitssatz})$$

wobei  $s^+$  und  $s^-$  die Anzahlen der positiven bzw. negativen Eigenwerte von  $A$  sind.



## Ähnliche Matrizen

### Definition

**I**  $A, B \in \mathbb{C}^{n \times n}$  sind *ähnlich*, falls ein  $S \in \text{GL}(n, \mathbb{C})$  existiert mit  $B = S A S^{-1}$ .

### Bedeutung

$A$  und  $B$  stellen dieselbe lineare Abbildung bezüglich einer Basis dar, d. h., es gibt ein lineares  $f: \mathbb{C}^n \rightarrow \mathbb{C}^n$  und Basen  $\mathcal{A}, \mathcal{B}$  des  $\mathbb{C}^n$  mit

$$A = A_f^{\mathcal{A}} = A_f^{\mathcal{A}, \mathcal{A}}, \quad B = A_f^{\mathcal{B}} = A_f^{\mathcal{B}, \mathcal{B}}.$$

### Normalformen

Sei  $A \in \mathbb{C}^n$  und seien  $\lambda_1, \dots, \lambda_n$  die in ihrer algebraischen Vielfachheit gezählten Eigenwerte  $A$ . Dann gilt:

- (1) Es gibt ein  $S \in \text{GL}(n, \mathbb{C})$  und  $b_{ij} \in \mathbb{C}$  mit  $i < j$  mit

$$S A S^{-1} = \begin{pmatrix} \lambda_1 & b_{12} & \dots & \dots & b_{1n} \\ & \lambda_2 & b_{23} & \dots & b_{2n} \\ & & \dots & \dots & \dots \\ & & & \lambda_{n-1} & b_{n-1, n} \\ & & & & \lambda_n \end{pmatrix}. \quad (\text{Trigonalisierung, Schur-Zerlegung})$$

- (2) Ist für jeden Eigenwert  $\lambda$  die geometrische Vielfachheit  $\dim(\text{Eig}(f, \lambda))$  gleich der algebraischen Vielfachheit  $\mu_f(\lambda)$ , so existiert ein  $S \in \text{GL}(n, \mathbb{C})$  mit

$$S A S^{-1} = \text{diag}(\lambda_1, \dots, \lambda_n). \quad (\text{Diagonalisierung})$$

- (3) Ist  $A$  normal (d. h.  $AA^* = A^*A$ ), so gibt es ein  $S \in \text{U}(n)$  mit

$$S A S^{-1} = \text{diag}(\lambda_1, \dots, \lambda_n). \quad (\text{unitäre Diagonalisierung, spektrale Zerlegung})$$

Insbesondere gilt dies für hermitesche  $A$  ( $A = A^*$ ) und für unitäre  $A$  ( $A^* = A^{-1}$ ). Genau für die hermiteschen Matrizen sind alle  $\lambda_i$  reell.

- (4) Es gibt ein  $S \in \text{GL}(n, \mathbb{C})$  mit

$$S A S^{-1} = \begin{pmatrix} J(\lambda_1) & & \\ & \dots & \\ & & J(\lambda_m) \end{pmatrix}, \quad (\text{Jordan-Normalform})$$

wobei nun  $\lambda_1, \dots, \lambda_m$ ,  $m \leq n$ , die paarweise verschiedenen Eigenwerte von  $A$  bezeichnen und jedes  $J(\lambda_i)$  eine aus  $\dim(\text{Eig}(A, \lambda_i))$  Jordan-Blöcken zusammengesetzte Bidiagonalmatrix ist. Die algebraische Vielfachheit des Eigenwerts  $\lambda_i$  entspricht der Zeilen- und Spaltenzahl von  $J(\lambda_i)$ . Die Matrix  $SAS^{-1}$  ist eine obere Dreiecksmatrix, sodass die Jordan-Normalform die Trigonalisierung verfeinert.

## Normalformen für reelle Matrizen

Wir betrachten nun reelle Matrizen  $A \in \mathbb{R}^{n \times n}$ . Die Ergebnisse für die Äquivalenz und die Kongruenz bleiben gleich (mit „symmetrisch“ statt „hermitesch“). Für die Ähnlichkeit ergeben sich Unterschiede, da das charakteristische Polynom  $p_A$  über  $\mathbb{R}$  im Allgemeinen nicht in Linearfaktoren zerfällt:

- (1) Eine Trigonalisierung ist im Allgemeinen nicht möglich.
- (2) Die Diagonalisierung gilt unter den Vielfachheitsvoraussetzungen, wenn  $p_A$  in Linearfaktoren zerfällt.
- (3) Eine orthogonale Diagonalisierung ist genau dann möglich, wenn  $A$  symmetrisch ist. Wie für  $\mathbb{C}$  ist dies ein zentrales Ergebnis der Linearen Algebra, vgl. 8.6.
- (4) Die Jordan-Normalform ist als Verstärkung der Trigonalisierung im Allgemeinen nicht mehr erreichbar. Sie gilt, falls  $p_A$  in reelle Linearfaktoren zerfällt.

Eine Matrix  $A \in \mathbb{R}^{n \times n}$  heißt *normal*, falls  $AA^t = A^t A$ . Im Gegensatz zum komplexen Fall ist die Normalität nicht mehr hinreichend für die Diagonalisierbarkeit. Das Beste, was man erreichen kann, ist eine Diagonalform mit  $(2 \times 2)$ -Kästchen: Es gibt ein  $S \in O(n)$  mit

$$S A S^{-1} = \begin{pmatrix} \lambda_1 & & & & \\ & \dots & & & \\ & & \lambda_k & & \\ & & & B_1 & \\ & & & & \dots \\ & & & & & B_s \end{pmatrix}, \quad (\text{Normalform für normale reelle Matrizen})$$

wobei die  $\lambda_i$  die Eigenwerte von  $A$  sind und die  $B_i \in \mathbb{R}^{2 \times 2}$  die schiefsymmetrische Form

$$B_i = \begin{pmatrix} a_i & -b_i \\ b_i & a_i \end{pmatrix}, \quad b_i \neq 0$$

besitzen. Dies zeigt man so: Wegen  $A \in \mathbb{R}^{n \times n} \subseteq \mathbb{C}^{n \times n}$  gibt es eine Orthonormalbasis  $\mathcal{A}$  des  $\mathbb{C}^n$  aus Eigenvektoren von  $A$ . Da  $A$  reell ist, kann  $\mathcal{A} = (x_1, \dots, x_k, z_1, \overline{z_1}, \dots, z_s, \overline{z_s})$  mit  $x_i \in \mathbb{R}^n$  erreicht werden. Ersetzt man die komplexen  $z_i$ -Paare durch  $\sqrt{2} \operatorname{Re}(z_i), \sqrt{2} \operatorname{Im}(z_i) \in \mathbb{R}^n$ , so erhält man eine Orthonormalbasis  $\mathcal{A}'$  des  $\mathbb{R}^n$ , die  $A$  in Normalform bringt. Dabei ist  $a_i = \operatorname{Re}(\lambda)$  und  $b_i = \operatorname{Im}(\lambda)$  für einen komplexen Eigenwert  $\lambda$  von  $A \in \mathbb{C}^{n \times n}$ .

Für den wichtigen Spezialfall einer orthogonalen Matrix  $A \in O(n)$  kann man

$$B_i = \begin{pmatrix} \cos \alpha_i & -\sin \alpha_i \\ \sin \alpha_i & \cos \alpha_i \end{pmatrix}, \quad \alpha_i \text{ kein Vielfaches von } \pi,$$

schreiben (da dann  $a_i^2 + b_i^2 = 1$ ). Diese Matrizen stellen Drehungen dar. Da die Eigenwerte einer orthogonalen Matrix den Betrag 1 haben, lässt sich die Normalform von  $A \in O(n)$  als „ $\pm 1$ -Kette + Drehkästchen“ beschreiben (vgl. hierzu die Tabelle in 7.8).

## 11. Blockstrukturen

Seien  $A \in \mathbb{K}^{n \times n}$  und  $1 \leq k \leq n$ . Dann definiert

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

mit  $A_{11} \in \mathbb{K}^{k \times k}$ ,  $A_{12} \in \mathbb{K}^{k \times (n-k)}$ ,  $A_{21} \in \mathbb{K}^{(n-k) \times k}$ ,  $A_{22} \in \mathbb{K}^{(n-k) \times (n-k)}$  eine  $2 \times 2$ -Blockstruktur der Matrix  $A$ . Die Matrix  $A$  wird in vier Matrizen aufgeteilt, die linke obere Matrix  $A_{11}$  legt dabei die Struktur fest. Blockstrukturen können den Umgang mit großen Matrizen wesentlich vereinfachen. Wir diskutieren einige Beispiele.

### Die Blockmultiplikation

Das Produkt zweier Blockmatrizen lässt sich über die Produkte der einzelnen Blöcke gemäß „Zeile mal Spalte“ berechnen:

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} \begin{pmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{pmatrix} = \begin{pmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{pmatrix}.$$

### Blockdreiecksmatrizen

Eine Blockmatrix der Form

$$\begin{pmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{pmatrix} \text{ oder } \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix}$$

nennt man eine (*untere bzw. obere*) *Blockdreiecksmatrix*. Für diese Matrizen gilt

$$\det(A) = \det(A_{11}) \det(A_{22}), \quad \sigma(A) = \sigma(A_{11}) \cup \sigma(A_{22}).$$

In 8.4 haben wir die Blockdeterminantenformel im Beweis verwendet, dass die geometrische Vielfachheit eines Eigenwerts kleinergleich der algebraischen ist.

### Das Schur-Komplement

Ist der linke obere Block  $A_{11}$  von  $A$  invertierbar, so gilt

$$\begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = \begin{pmatrix} E_k & 0 \\ A_{21} A_{11}^{-1} & E_{n-k} \end{pmatrix} \begin{pmatrix} A_{11} & A_{12} \\ 0 & S \end{pmatrix} \quad (\text{Block-LR-Zerlegung})$$

wobei  $S = A_{22} - A_{21} A_{11}^{-1} A_{12}$  das *Schur-Komplement* von  $A_{11}$  in  $A$  ist. Es gilt

$$\det(A) = \det(A_{11}) \det(S), \quad \sigma(A) = \sigma(A_{11}) \cup \sigma(S),$$

sodass insbesondere die Invertierbarkeit von  $A$  äquivalent zur Invertierbarkeit von  $S$  ist. Ist der rechte untere Block  $A_{22}$  invertierbar, so führen analoge Überlegungen zum Schur-Komplement  $A_{11} - A_{12} A_{22}^{-1} A_{21}$  von  $A_{22}$  in  $A$ .

## 12. Berechnung und Bestimmung

### Matrizen

Berechnung von Koordinaten 5.5, 5.6

Invertierung einer Matrix 5.6

Basiswechsel, Umrechnung auf neue Koordinaten 5.7

Lösen eines Gleichungssystems 5.11, 5.12

Bestimmung der Normalform (bzgl. der Äquivalenz von Matrizen) 5.12

LR-Zerlegung Überblick 5

### Euklidische und unitäre Vektorräume

Orthonormalisierung (Gram-Schmidt) 6.7

QR-Zerlegung Überblick 7

Bestimmung des darstellenden (Riesz-) Vektors 6.10

### Determinanten

Berechnung einer Determinante 7.2, 7.4, 7.6

Volumen eines Parallelotops 7.9

Bestimmung der Definitheit 7.11, 8.7, Überblick 8

### Eigenwerte

Berechnung der Eigenwerte und Eigenvektoren 8.3

Trigonalisierung (Schur-Zerlegung) 8.5

Spektralzerlegung (orthogonale bzw. unitäre Diagonalisierung) 8.6

Hauptachsentransformation 8.7

Singulärwertzerlegung 8.8

Hauptraumzerlegung, Exponenten des Minimalpolynoms 8.11

Jordan-Normalform 8.12

---

# Ausblicke zu Eigenwerten

---

## 1. Eigenwerte ohne Determinanten

In der Linearen Algebra werden traditionell Eigenwerte, Eigenvektoren und Normalformen nach der Diskussion der Determinanten behandelt: Die Eigenwerte einer Matrix  $A$  werden als die Nullstellen des charakteristischen Polynoms  $p_A$  identifiziert, welches mit Hilfe der Determinantenfunktion definiert wird. Ein alternativer Aufbau, der die Verwendung von Determinanten vermeidet, ist möglich. An der Spitze steht dabei das folgende Argument:

### Alternativer Beweis der Existenz eines Eigenwerts einer komplexen Matrix

Sei  $A \in \mathbb{C}^{n \times n}$ . Wir zeigen, dass  $A$  einen Eigenwert  $\lambda \in \mathbb{C}$  besitzt. Hierzu sei  $v \in \mathbb{C}^n - \{0\}$  beliebig. Wegen  $\dim(\mathbb{C}^n) = n$  gibt es ein kleinstes  $m \leq n$ , sodass

$$(v, Av, A^2v, \dots, A^m v)$$

linear abhängig ist. Dann existieren  $\alpha_0, \dots, \alpha_m \in \mathbb{C}$ ,  $\alpha_m \neq 0$ , mit

$$(\text{+}) \quad 0 = \alpha_0 v + \alpha_1 Av + \dots + \alpha_m A^m v = (\alpha_0 E_n + \alpha_1 A + \dots + \alpha_m A^m) v.$$

Nach dem Fundamentalsatz der Algebra gibt es  $\lambda_1, \dots, \lambda_m \in \mathbb{C}$  mit

$$\alpha_0 + \alpha_1 X + \dots + \alpha_m X^m = \alpha_m (X - \lambda_m) \dots (X - \lambda_1).$$

Nach (+) gilt also

$$(\text{++}) \quad \alpha_m (A - \lambda_m E_n) \dots (A - \lambda_1 E_n) v = 0.$$

Ist nun  $1 \leq j \leq m$  minimal mit

$$(A - \lambda_j E_n) \dots (A - \lambda_1 E_n) v = 0,$$

so ist  $\lambda_j$  ein Eigenwert von  $A$  zum Eigenvektor

$$v_j = (A - \lambda_{j-1} E_n) \dots (A - \lambda_1 E_n) v \neq 0, \quad \text{mit } v_j = v \text{ im Fall } j = 1.$$

(Allgemein ist jedes  $\lambda_i$  ein Eigenwert von  $A$ , da man sonst (++) mit  $(A - \lambda_i E_n)^{-1}$  multiplizieren könnte und durch Ausmultiplizieren eine nichttriviale Nulldarstellung mit den Vektoren  $v, Av, A^2v, \dots, A^{m-1}v$  erhalten würde.)

Aufbauend auf diesem Argument haben C. G. Broyden 1975 und S. Axler 1995 eine determinantenfreie Eigenwert- und Normalformentheorie entwickelt, siehe

C. G. Broyden: Basic Matrices, The Maxmillan Press, London 1975,

S. Axler: Down with Determinants!, American Mathematical Monthly 102 (1995), S.139–154.

Wir verweisen auch auf

Garry J. Tee: Up with Determinants!, IMAGE (The Bulletin of the International Linear Algebra Society) 30 (2003), S. 7–11.

## 2. Eigenwerte ohne Fundamentalsatz

Unsere Beweise für die Existenz von Eigenwerten beruhen auf dem Fundamentalsatz der Algebra. Für den wichtigen Spezialfall einer symmetrischen oder hermiteschen Matrix ist es möglich, den Einsatz des Fundamentalsatzes (und von Determinanten) durch ein analytisches Argument zu ersetzen. Für die folgende Diskussion setzen wir Grundkenntnisse der mehrdimensionalen Differentialrechnung voraus.

Seien  $n \geq 1$  und  $A \in \mathbb{R}^{n \times n}$ . Dann heißt die Funktion  $R_A : \mathbb{R}^n - \{0\} \rightarrow \mathbb{R}$  mit

$$R_A(x) = \frac{\langle x, Ax \rangle}{\langle x, x \rangle} \quad \text{für alle } x \in \mathbb{R}^n - \{0\}$$

der *Rayleigh-Quotient* von  $A$  (dabei wird das kanonische Skalarprodukt verwendet). Wichtige Eigenschaften sind:

- (1) Die Funktion  $R_A : \mathbb{R}^n - \{0\} \rightarrow \mathbb{R}$  ist stetig.
- (2)  $R_A(\alpha x) = R_A(x)$  für alle  $\alpha \in \mathbb{R}^*$  und  $x \in \mathbb{R}^n - \{0\}$ . (Homogenität)
- (3) Ist  $\lambda$  ein Eigenwert von  $A$  und  $x$  ein zugehöriger Eigenvektor, so gilt

$$R_A(x) = \frac{\langle x, Ax \rangle}{\langle x, x \rangle} = \frac{\lambda \langle x, x \rangle}{\langle x, x \rangle} = \lambda.$$

### Analytischer Beweis der Existenz eines Eigenwerts für symmetrische Matrizen

Sei  $A \in \mathbb{R}^{n \times n}$  symmetrisch, und sei  $S^{n-1} = \{x \in \mathbb{R}^n \mid \|x\| = 1\}$  die Einheitskugel im  $\mathbb{R}^n$  (mit der euklidischen Norm). Da  $R_A$  stetig und  $S^{n-1}$  kompakt ist, nimmt  $R_A$  auf  $S^{n-1}$  in einem Punkt  $x \in S^{n-1}$  ihr Maximum an. Aufgrund der Homogenität ist dieses Maximum global, sodass der Gradient von  $R_A$  im Punkt  $x$  gleich dem Nullvektor ist:

$$0 = \text{grad}(R_A)(x) = 2 \frac{\langle x, x \rangle Ax - \langle x, Ax \rangle x}{\langle x, x \rangle^2}.$$

(Bei der Berechnung des Gradienten verwenden wir, dass  $A$  symmetrisch ist.)

Damit ist  $\langle x, x \rangle Ax - \langle x, Ax \rangle x = 0$ , sodass

$$Ax = \frac{\langle x, Ax \rangle}{\langle x, x \rangle} x = R_A(x) x.$$

Dies zeigt, dass  $R_A(x)$  ein Eigenwert von  $A$  und  $x$  ein zugehöriger Eigenvektor ist.

Genauer ist  $R_A(x)$  der größte Eigenwert von  $A$ .

Durch Bildung von orthogonalen Unterräumen ergibt sich induktiv, dass jede symmetrische Matrix eine Orthonormalbasis aus Eigenvektoren besitzt. Analoge Überlegungen gelten für hermitesche Matrizen  $A \in \mathbb{C}^{n \times n}$  (dann ist  $R_A$  auf  $\mathbb{C}^n - \{0\}$  definiert, aber nach wie vor reellwertig). Insgesamt ergibt sich ein Beweis des Spektralsatzes, der den Fundamentalsatz der Algebra nicht verwendet.

### 3. Gershgorin-Kreise und die Lage der Eigenwerte

Wir betrachten eine Matrix  $A \in \mathbb{C}^{n \times n}$ , einen Eigenwert  $\lambda \in \sigma(A)$  und einen zugehörigen Eigenvektor  $z \in \mathbb{C}^n - \{0\}$ . Wegen  $Az = \lambda z$  gilt  $\sum_j a_{ij} z_j = \lambda z_i$  für alle  $i$ , sodass

$$\sum_{j \text{ mit } j \neq i} a_{ij} z_j = \lambda z_i - a_{ii} z_i = (\lambda - a_{ii}) z_i \text{ für alle } i.$$

Wir betrachten nun eine im Betrag maximale Komponente des Eigenvektors  $z$ . Sei also  $i^*$  derart, dass

$$|z_{i^*}| = \max_{1 \leq j \leq n} |z_j|.$$

Wegen  $z \neq 0$  ist  $z_{i^*} \neq 0$ . Damit gilt

$$(+)|\lambda - a_{i^*i^*}| \leq \sum_{j \neq i^*} |a_{i^*j}| \frac{|z_j|}{|z_{i^*}|} \leq \sum_{j \neq i^*} |a_{i^*j}|.$$

Damit haben wir den Abstand von  $\lambda$  zum Diagonaleintrag  $a_{i^*i^*}$  abgeschätzt. Definieren wir also für alle  $1 \leq i \leq n$  den  $i$ -ten *Gershgorin-Kreis* von  $A$  durch

$$G(i) = \{w \in \mathbb{C} \mid |w - a_{ii}| \leq \sum_{j \neq i} |a_{ij}|\},$$

so liegt jeder Eigenwert von  $A$  nach (+) in mindestens einem Gershgorin-Kreis von  $A$ . Diese Kreise überdecken also das Spektrum von  $A$ :

$$\sigma(A) \subseteq \bigcup_{1 \leq i \leq n} G(i).$$

Hat  $A$  im Betrag kleine Einträge außerhalb der Diagonalen, so haben die Gershgorin-Kreise einen kleinen Radius, sodass wir in einfacher Weise eine recht genaue Auskunft über die Lage der Eigenwerte von  $A$  erhalten können.

Genauer lässt sich zeigen:

*Ist ein Kreis  $G(i)$  disjunkt von allen anderen, so enthält er genau einen Eigenwert.*

Für eine noch genauere Beschreibung setzen wir

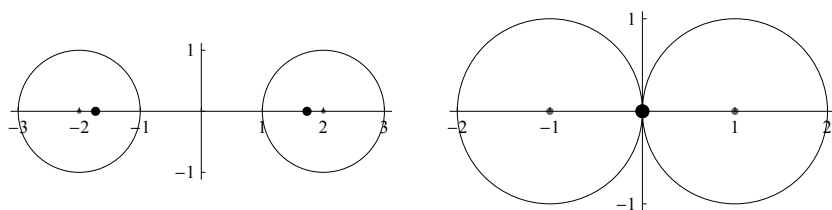
$$G(I) = \bigcup_{i \in I} G(i) \quad \text{für } I \subseteq \{1, \dots, n\}.$$

Dann gilt, mit in ihrer Vielfachheit gezählten Eigenwerten:

*Sind  $I_1$  und  $I_2$  disjunkt mit  $I_1 \cup I_2 = \{1, \dots, n\}$  und  $G(I_1) \cap G(I_2) = \emptyset$ , so enthält  $G(I_1)$  genau  $|I_1|$  und  $G(I_2)$  genau  $|I_2|$  Eigenwerte.*

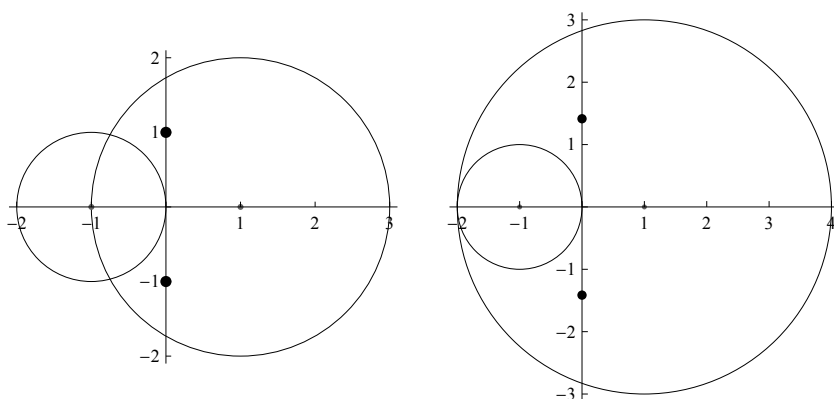
Die folgenden Diagramme zeigen die Gershgorin-Kreise für vier  $2 \times 2$ -Matrizen und eine  $4 \times 4$ -Matrix. Die Mittelpunkte der Kreise sind durch kleine graue Punkte markiert, die Eigenwerte durch größere schwarze Punkte. Die  $2 \times 2$ -Matrizen zeigen, dass die Eigenwerte am Rand der Kreise liegen können und dass ein Kreis im Fall einer Überlappung oder einer Inklusion keinen Eigenwert enthalten muss.





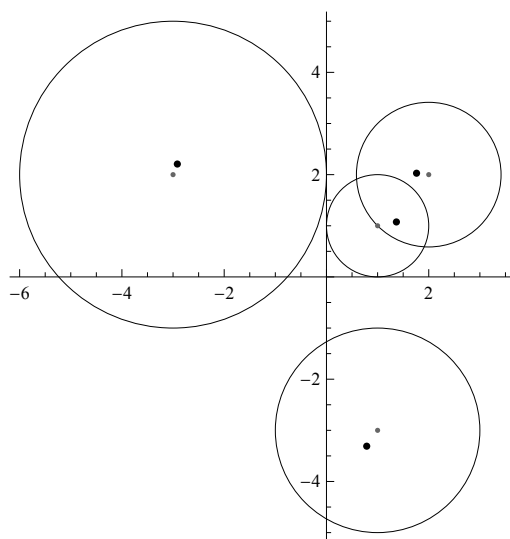
$$A = \begin{pmatrix} 2 & -1 \\ 1 & -2 \end{pmatrix}$$

$$A = \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}$$



$$A = \begin{pmatrix} 1 & -2 \\ 1 & -1 \end{pmatrix}$$

$$A = \begin{pmatrix} 1 & -3 \\ 1 & -1 \end{pmatrix}$$



$$A = \begin{pmatrix} 1-3i & 1 & 0 & i \\ 0 & 1+i & 0 & 1 \\ 0 & 1+i & 2+2i & 0 \\ 2i & 0 & -1 & -3+2i \end{pmatrix}$$

## 4. Matrixnormen

Ist  $n \geq 1$  und  $\|\cdot\|: \mathbb{C}^n \rightarrow [0, \infty[$  eine Norm auf dem  $\mathbb{C}^n$ , so definiert

$$\|A\| = \max \{ \|Az\| \mid z \in S^{n-1} \},$$

eine Norm  $\|\cdot\|: \mathbb{C}^{n \times n} \rightarrow [0, \infty[$ , wobei  $S^{n-1} = \{ z \in \mathbb{C}^n \mid \|z\| = 1 \}$ . Sie heißt die von der Norm  $\|\cdot\|$  *induzierte Matrixnorm* auf  $\mathbb{C}^{n \times n}$ . Da die Einheitskugel  $S^{n-1}$  eine kompakte Teilmenge des  $\mathbb{C}^n$  ist (unter der von der Norm  $\|\cdot\|$  induzierten Metrik), nimmt die stetige Funktion  $F: S^{n-1} \rightarrow \mathbb{R}$ ,  $F(z) = \|Az\|$ , ihr Maximum an, sodass  $\|A\|$  wohldefiniert ist. Die Homogenität, Definitheit und Dreiecksungleichung folgen aus den entsprechenden Eigenschaften der Ausgangsnorm.

### Die Spektral-Normen

Wir nehmen nun an, dass die Norm auf dem  $\mathbb{C}^n$  die euklidische Norm ist und bestimmen schrittweise die Werte der induzierten Matrixnorm. Wir schreiben  $\|\cdot\|$  statt  $\|\cdot\|_2$ .

#### Diagonalmatrizen

Ist  $A = \text{diag}(d_1, \dots, d_n)$  und  $d = \max_k |d_k|$ , so gilt

$$\|Az\|^2 = |d_1 z_1|^2 + \dots + |d_n z_n|^2 \leq d^2 (|z_1|^2 + \dots + |z_n|^2) = d^2 \quad \text{für alle } z \in S^{n-1}.$$

Wegen  $\|Ae_k\| = \|d_k e_k\| = |d_k|$  für alle  $k$  ist also  $\|A\| = d$ .

#### Unitäre Matrizen

Ist  $U \in U(n)$ , so bildet  $f_U$  die Kugel  $S^{n-1}$  bijektiv auf sich selbst ab. Hieraus ergibt sich, dass  $\|U\| = 1$  und allgemeiner

$$\|AU\| = \|A\| = \|UA\| \quad \text{für alle } A \in \mathbb{C}^{n \times n}.$$

#### Normale Matrizen

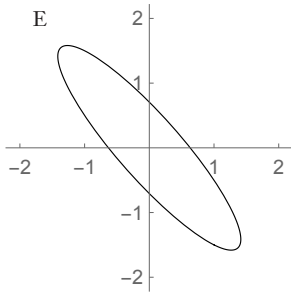
Nach dem Spektralsatz (vgl. 8.6) lässt sich eine hermitesche oder allgemeiner normale Matrix  $A$  unitär diagonalisieren: Es gibt ein  $S \in U(n)$  mit  $A = S^* \text{diag}(\lambda_1, \dots, \lambda_n) S$ , wobei  $\lambda_1, \dots, \lambda_n$  die in ihrer Vielfachheit gezählten Eigenwerte von  $A$  sind. Nach den vorangehenden Überlegungen ist also

$$\|A\| = \max_k |\lambda_k|$$

der betragsmäßig größte Eigenwert von  $A$ . Die euklidisch induzierte Matrixnorm heißt deswegen auch die *Spektral-Norm* auf dem  $\mathbb{C}^{n \times n}$ .

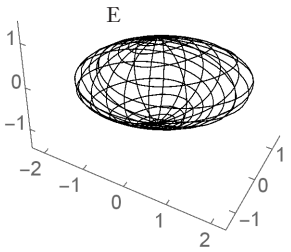
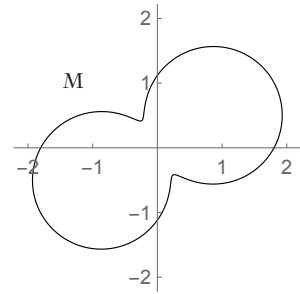
#### Allgemeiner Fall

Ist  $A \in \mathbb{C}^{n \times n}$  beliebig, so liefert eine Singulärwertzerlegung Matrizen  $S, T \in U(n)$  und  $\sigma_1, \dots, \sigma_n \geq 0$  mit  $A = S^* \text{diag}(\sigma_1, \dots, \sigma_n) T$  (vgl. 8.8). Damit gilt  $\|A\| = \max_k \sigma_k$ . Die Spektral-Norm der Matrix  $A$  ist also stets der größte Singulärwert von  $A$ . Die Bezeichnung „Singulärwertnorm“ wäre demnach passender, ist aber nicht üblich.



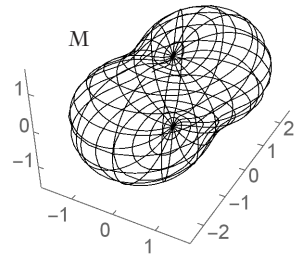
$$A = \begin{pmatrix} 1 & 1 \\ -3/2 & -1/2 \end{pmatrix}$$

$$\|A\| = \frac{\sqrt{9 + \sqrt{65}}}{2} = 2,065\dots$$



$$A = \begin{pmatrix} 1 & 2 & 0 \\ 1 & -2 & 0 \\ 0 & 1 & 1 \end{pmatrix}$$

$$\|A\| = 2,5320\dots$$



$$E = \{ Av \mid \|v\| = 1 \}$$

$$M = \{ \|Av\| v \mid \|v\| = 1 \}$$

Die Diagramme visualisieren die Spektralnorm für reelle Matrizen. Links sind die Bilder der Einheitsvektoren  $v$  unter  $A$  dargestellt, rechts die mit  $\|Av\|$  skalierten Einheitsvektoren  $v$ . Die Spektralnorm ist nach Definition der Radius der kleinsten 2- bzw. 3-dimensionalen Sphäre, die die Menge  $E$  oder gleichwertig  $M$  umfasst.

### Die Spaltensummen- und Zeilensummennormen

Schließlich betrachten wir noch die Summennorm  $\|\cdot\|_1$  und die Maximumsnorm  $\|\cdot\|_\infty$  auf dem  $\mathbb{C}^n$  (vgl. 6.4). Für die zugehörigen induzierten Matrixnormen gilt

$$\|A\|_1 = \max_{1 \leq j \leq n} \sum_{1 \leq i \leq n} |a_{ij}| = \max_{1 \leq j \leq n} \|a_j\|_1, \quad (\text{Spaltensummennorm})$$

$$\|A\|_\infty = \max_{1 \leq i \leq n} \sum_{1 \leq j \leq n} |a_{ij}| = \max_{1 \leq i \leq n} \|b_i\|_1 \quad (\text{Zeilensummennorm})$$

für alle  $A \in \mathbb{C}^{n \times n}$ , wobei die  $a_j$  die Spalten und die  $b_i$  die Zeilen von  $A$  sind. Zur Berechnung von  $\|A\|_1$  summiert man für jede Spalte die Beträge der Einträge und wählt unter den  $n$  Summen den maximalen Wert. Analoges gilt für  $\|A\|_\infty$  mit „Zeilen“ statt „Spalten“.

Die beiden Normen eignen sich zur Abschätzung der Spektralnorm  $\|\cdot\|_2$ , denn für alle  $A \in \mathbb{C}^{n \times n}$  gilt

$$\|A\|_2 \leq \sqrt{\|A\|_1 \|A\|_\infty}. \quad (\text{Schur-Abschätzung oder Schur-Test})$$

Für Diagonalmatrizen ist die Ungleichung eine Gleichung.

## 5. Matrixexponentiale

Für eine Matrix  $A \in \mathbb{C}^{n \times n}$  ist das *Exponential*  $\exp(A) = e^A \in \mathbb{C}^{n \times n}$  definiert durch

$$\exp(A) = \sum_{k \geq 0} \frac{A^k}{k!},$$

wobei die Reihe als Limes der Partialsummen bezüglich einer beliebigen Matrixnorm  $\|\cdot\|$  zu verstehen ist. Die Reihe konvergiert für alle  $A$ , und die Konvergenz ist beschrieben durch die Konvergenz der Einträge in  $\mathbb{C}$ :

$$\exp(A)(i, j) = \lim_{n \rightarrow \infty} \sum_{k \leq n} \frac{A^k(i, j)}{k!} \quad \text{für alle } 1 \leq i, j \leq n.$$

Für alle  $A, B \in \mathbb{C}^{n \times n}$  gilt:

- (a)  $\exp(0) = E_n$ ,
- (b)  $\exp(A)$  ist invertierbar und es gilt  $\exp(A)^{-1} = \exp(-A)$ ,
- (c)  $\exp(A^*) = \exp(A)^*$ ,
- (d)  $\exp(A + B) = \exp(A) \exp(B)$ , falls  $A$  und  $B$  kommutieren (d.h.  $AB = BA$ ),
- (e) ist  $A^m = 0$  für ein  $m$ , so ist  $\exp(A) = \sum_{k < m} A^k/k!$  ein Matrixpolynom vom Grad  $\leq m$ .

Wir bestimmen die Exponentiale  $\exp(A)$  wieder schrittweise.

### Diagonalmatrizen

Ist  $A = \text{diag}(d_1, \dots, d_n)$ , so gilt  $A^k = \text{diag}(d_1^k, \dots, d_n^k)$  für alle  $k \geq 0$  und damit

$$\exp(A) = \text{diag}(\exp(d_1), \dots, \exp(d_n)).$$

Speziell ist  $\exp(\lambda E_n) = \text{diag}(\exp(\lambda), \dots, \exp(\lambda)) = \exp(\lambda) E_n$ .

### Ähnliche Matrizen

Sind  $A, B$  ähnlich und ist  $S \in GL(n, \mathbb{C})$  mit  $A = S^{-1} B S$ , so gilt

$$A^k = S^{-1} B^k S \quad \text{für alle } k \geq 0.$$

Hieraus ergibt sich

$$\exp(A) = S^{-1} \exp(B) S.$$

Damit sind also die Matrizen  $\exp(A)$  und  $\exp(B)$  ebenfalls ähnlich.

### Diagonalisierbare Matrizen

Gilt  $A = S^{-1} \text{diag}(\lambda_1, \dots, \lambda_n) S$  mit  $S \in GL(n, \mathbb{C})$ , so gilt

$$\exp(A) = S^{-1} \text{diag}(\exp(\lambda_1), \dots, \exp(\lambda_n)) S$$

nach den vorangehenden Überlegungen. Mit  $A$  ist also auch  $\exp(A)$  diagonalisierbar.

**Allgemeiner Fall: Bestimmung über eine Jordan-Normalform**

Ist  $A$  nicht diagonalisierbar, so liefert eine Jordan-Normalform, dass

$$\exp(A) = S^{-1} \begin{pmatrix} \exp(J(\lambda_1)) & & \\ & \dots & \\ & & \exp(J(\lambda_m)) \end{pmatrix} S$$

mit  $S \in GL(n, \mathbb{C})$  und Jordan-Blöcken  $J(\lambda_1), \dots, J(\lambda_m)$  (vgl. 8.12). Damit ist die Berechnung von  $\exp(A)$  auf die Berechnung des Exponentials von Jordan-Blöcken reduziert. Zur Berechnung eines  $k \times k$  Jordan-Blocks  $J(\lambda)$  schreiben wir

$$J(\lambda) = \lambda E_k + N \quad \text{mit} \quad N = \begin{pmatrix} 0 & 1 & & \\ & 0 & 1 & \\ & & \dots & \dots \\ & & & 0 & 1 \\ & & & & 0 \end{pmatrix}.$$

Es gilt  $(\lambda E_k) N = N (\lambda E_k)$ , sodass

$$\exp(J(\lambda)) = \exp(\lambda E_k) \exp(N) = \exp(\lambda) \exp(N).$$

Wegen  $N^k = 0$  ist  $\exp(N)$  ein Polynom vom Grad kleiner als  $k$ . So ergibt sich zum Beispiel

$$\exp(J_2(\lambda)) = \exp \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix} = \exp(\lambda) \exp \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \exp(\lambda) \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix},$$

$$\exp(J_3(\lambda)) = \exp \begin{pmatrix} \lambda & 1 & 0 \\ 0 & \lambda & 1 \\ 0 & 0 & \lambda \end{pmatrix} = \exp(\lambda) \exp \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} =$$

$$\exp(\lambda) \left( N^0 + N^1 + \frac{N^2}{2} \right) = \exp(\lambda) \exp \begin{pmatrix} 1 & 1 & 1/2 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$$

Mit Hilfe der Jordan-Normalform lässt sich zudem zeigen, dass

$$\det(\exp(A)) = \exp(\text{spur}(A)) \quad \text{für alle } A \in \mathbb{C}^{n \times n}.$$

Da  $\exp(x) > 0$  für alle  $x \in \mathbb{R}$  gilt, gibt es im Fall  $\det(B) < 0$  kein  $A \in \mathbb{R}^{n \times n}$  mit  $\exp(A) = B$ . Man kann jedoch zeigen, dass  $\exp: \mathbb{C}^{n \times n} \rightarrow GL(n, \mathbb{C})$  surjektiv ist, sodass für alle  $B \in GL(n, \mathbb{C})$  ein (nicht eindeutiger) Matrixlogarithmus  $A \in \mathbb{C}^{n \times n}$  durch  $\exp(A) = B$  erklärt werden kann.

## 6. Lineare Systeme von Differentialgleichungen

Seien  $A \in \mathbb{C}^{n \times n}$  und  $y_0 \in \mathbb{C}^n$ . Wir betrachten das in einer reellen Variablen  $t$  formulierte Anfangswertproblem

$$\dot{y}(t) = Ay(t), \quad y(0) = y_0, \quad \text{d.h.}$$

$$\dot{y}_1(t) = a_{11} y_1(t) + \dots + a_{1n} y_n(t), \quad y_1(0) = (y_0)_1,$$

$$\dot{y}_2(t) = a_{21} y_1(t) + \dots + a_{2n} y_n(t), \quad y_2(0) = (y_0)_2,$$

...

$$\dot{y}_n(t) = a_{n1} y_1(t) + \dots + a_{nn} y_n(t), \quad y_n(0) = (y_0)_n.$$

Die (nach der Theorie der Differentialgleichungen eindeutig bestimmte) Lösung

$$y: \mathbb{R} \rightarrow \mathbb{C}^n, \quad y(t) = (y_1(t), \dots, y_n(t)),$$

des Anfangswertproblems ist gegeben durch

$$(+)\ y(t) = \exp(tA) y_0 \quad \text{für alle } t \in \mathbb{R}.$$

Denn es gilt:

$$(a) \ y(0) = \exp(0A) y_0 = \exp(0) y_0 = y_0,$$

$$(b) \ \frac{d}{dt} \exp(tA) = \frac{d}{dt} \sum_{k \geq 0} \frac{t^k}{k!} A^k = \sum_{k \geq 0} \frac{t^{k-1}}{(k-1)!} A^k = A \exp(tA),$$

sodass

$$\dot{y}(t) = \frac{d}{dt} \exp(tA) y_0 = A \exp(tA) y_0 = A y(t) \quad \text{für alle } t \in \mathbb{R}.$$

Damit kann das Matrixexponential also zur Lösung eines homogenen Systems von linearen Differentialgleichungen eingesetzt werden. Weiter gilt nach den Ergebnissen des letzten Abschnitts: Ist  $A$  diagonalisierbar, so gibt es ein  $S \in GL(n, \mathbb{C})$  mit

$$(++) \ y(t) = S^{-1} \operatorname{diag}(\exp(t\lambda_1), \dots, \exp(t\lambda_n)) S y_0 \quad \text{für alle } t \in \mathbb{R},$$

wobei  $\lambda_1, \dots, \lambda_n$  die Eigenwerte von  $A$  sind. Wir können die in der Form  $(++)$  präsentierte Lösung wie folgt beschreiben:

- (1) Bestimme den Koordinatenvektor  $S y_0 \in \mathbb{C}^n$  des Anfangswerts  $y_0$  bzgl. einer Basis aus Eigenvektoren von  $A$ . (Die Basisvektoren sind die Spalten von  $S^{-1}$ .)
- (2) Multipliziere jede Koordinate von  $S y_0$  mit dem entsprechenden Faktor  $\exp(t\lambda_i) \in \mathbb{C}$ .
- (3) Wechsle durch Anwendung von  $S^{-1}$  in die kanonische Basis zurück.

Schreiben wir  $\lambda_j = \rho_j + i\mu_j$  mit  $\rho_j, \mu_j \in \mathbb{R}$ , so ist das Verhalten der Lösungskomponente  $y_j(t)$  im Wesentlichen von der komplexen Exponentialfunktion

$$\exp(t\lambda_j) = \exp(t\rho_j) \exp(it\mu_j) = \exp(t\rho_j) (\cos(t\mu_j) + i \sin(t\mu_j)), \quad t \in \mathbb{R},$$

bestimmt. Ist  $\rho_j > 0$ , so strebt  $y_j(t)$  für  $t \rightarrow \infty$  exponentiell schnell gegen  $\infty$ . Ist  $\rho_j = 0$ , so oszilliert  $y_j(t)$  wie  $\cos(t\mu_j) + i \sin(t\mu_j)$ . Ist  $\rho_j < 0$ , so konvergiert  $y_j(t)$  für  $t \rightarrow \infty$  exponentiell schnell gegen Null.

Ist  $A$  nicht diagonalisierbar, so kann die Lösung  $y(t) = \exp(tA)y_0$  durch die Exponentiale von Jordan-Blöcken beschrieben werden. Die  $(++)$  entsprechenden Funktionen haben nun die Form  $\exp(t\lambda_j) p_j(t)$ ,  $t \in \mathbb{R}$ , mit einem komplexen Polynom  $p_j: \mathbb{C} \rightarrow \mathbb{C}$ .

### Der harmonische Oszillator

Zu den wichtigsten Beispielen eines Anfangswertproblems der obigen Form zählt

$$\dot{y}(t) = Ay(t), \quad y(0) = y_0 \quad \text{mit} \quad A = \begin{pmatrix} 0 & \omega \\ -\omega & 0 \end{pmatrix}, \quad \omega > 0.$$

Komponentenweise schreibt sich dies als

$$\dot{y}_1 = \omega y_2, \quad \dot{y}_2 = -\omega y_1.$$

Für die zweite Komponente erhalten wir also

$$\ddot{y}_2 = -\omega^2 y_2. \quad (\text{harmonischer Oszillator der Frequenz } \omega)$$

Die Matrix  $A$  ist *schiefsymmetrisch*, d. h.  $A = -A^*$ . Damit ist

$$\exp(tA)^* = \exp(tA^*) = \exp(-tA) = \exp(tA)^{-1}.$$

Als schiefsymmetrische Matrix ist  $A$  normal und damit unitär diagonalisierbar. Man berechnet

$$A = S^* \operatorname{diag}(i\omega, -i\omega) S \quad \text{mit} \quad S = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -i \\ 1 & i \end{pmatrix}.$$

Wenden wir die Lösungsformel (+) auf die beiden Anfangswerte  $y_0 = e_1$  und  $y_0 = e_2$  an, so erhalten wir als Lösungen die trigonometrischen Oszillationen

$$y(t) = \exp(tA) e_1 = \begin{pmatrix} \cos(\omega t) \\ \sin(\omega t) \end{pmatrix} \quad \text{bzw.} \quad y(t) = \exp(tA) e_2 = \begin{pmatrix} -\sin(\omega t) \\ \cos(\omega t) \end{pmatrix}.$$

Für einen beliebigen Anfangswert  $y_0 = \alpha e_0 + \beta e_2$  erhalten wir damit die Lösung

$$y(t) = \exp(tA) y_0 = \alpha \begin{pmatrix} \cos(\omega t) \\ \sin(\omega t) \end{pmatrix} + \beta \begin{pmatrix} -\sin(\omega t) \\ \cos(\omega t) \end{pmatrix}.$$





---

# Anhänge

---

## 1. Junktoren

In der Mathematik werden die Verknüpfungen

*nicht, und, oder, impliziert, äquivalent*

zum Teil anders verwendet als in der Umgangssprache. Wichtig sind:

- (a) „A und B“ ist gleichwertig zu „B und A“,
- (b) „A oder B“ ist gleichwertig zu „B oder A“.
- (c) „A oder B“ ist kein exklusives „entweder A oder B“, sondern bedeutet „eines von beiden oder auch beide“, „mindestens eine der beiden Aussagen ist richtig“.
- (d) „A impliziert B“ bedeutet

„aus A folgt B“, „wenn A gilt, so gilt auch B“,

„A zieht B nach sich“,

„A ist hinreichend für B“, „B ist notwendig für A“.

Die Implikation will keine Kausalität ausdrücken. „A impliziert B“ ist gleichwertig zu „(nicht A) oder B“. Speziell ist die Aussage „A impliziert B“ stets richtig, wenn die Aussage A falsch ist.

- (e) „A ist äquivalent zu B“ bedeutet „A gilt genau dann, wenn B gilt“, „A gilt dann und nur dann, wenn B gilt“, „(A impliziert B) und (B impliziert A)“.

Die folgende Tabelle gibt einen Überblick über die Junktoren.

<i>Zeichen</i>	<i>Bedeutung</i>	<i>Name</i>
$\neg$	nicht/non ...	Negation
$\wedge$	... und ... sowohl ... als auch ...	Konjunktion
$\vee$	oder (nicht exklusiv)	Disjunktion
$\rightarrow$	... impliziert ... aus ... folgt ... wenn ... so auch ...	Implikation
$\leftrightarrow$	... genau dann, wenn ... ... ist äquivalent zu ... ... dann und nur dann, wenn ...	Äquivalenz

Genauer wird die Semantik der mathematischen Junktoren durch die folgenden Wahrheitstafeln festgelegt. Dabei steht „w“ für „wahr, gültig“ und „f“ für „falsch, ungültig“.

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

Ist zum Beispiel A wahr und B falsch, so ist  $\neg A$  falsch,  $A \vee B$  wahr und  $A \rightarrow B$  falsch.

Mit Hilfe von Klammern lassen sich mehrere Aussagen miteinander verbinden:

$(A \wedge \neg B) \wedge C$ ,  $(A \vee B) \rightarrow C$ ,  $A \vee (B \rightarrow C)$ , ...

Um Klammern zu sparen, vereinbart man die Bindungsstärke

$\neg$ ,  $\wedge$ ,  $\vee$ ,  $\rightarrow$ ,  $\leftrightarrow$  (von stark nach schwach bindend),

die man sich durch Magnete vorstellen kann: Der Magnet  $\wedge$  ist stärker als  $\rightarrow$ , sodass zum Beispiel  $A \wedge B \rightarrow C$  die Aussage  $(A \wedge B) \rightarrow C$  ist und nicht etwa  $A \wedge (B \rightarrow C)$ .

Ist eine aus A, B, C, ... zusammengesetzte Aussage für alle Wahrheitswerte „w“ und „f“ von A, B, C, ... wahr, so heißt die Aussage eine *Tautologie*. Ob eine Tautologie vorliegt, kann man mit Hilfe von Wahrheitstafeln überprüfen. Ein Beispiel ist:

A	$\rightarrow$	B	$\leftrightarrow$	$\neg$	A	$\vee$	B
w	w	w	w	f	w	w	w
w	f	f	w	f	w	f	f
f	w	w	w	w	f	w	w
f	w	f	w	w	f	w	f
1		5	2		4	3	

Die Zahlen geben an, in welcher Reihenfolge die Spalten berechnet werden. Die Tafel für  $\vee$  wirkt auf die Spalten 2 und 3, die Tafel für  $\leftrightarrow$  auf die Spalten 1 und 4.

Ist die zuletzt berechnete Spalte – die *Ergebnisspalte* der Tafel – nur mit dem Wert „w“ gefüllt, so heißt die untersuchte Aussage eine *Tautologie* oder *allgemeingültig*. Weitere Beispiele für Tautologien sind:

$\neg \neg A \leftrightarrow A$  (doppelte Verneinung, duplex negatio affirmat)

$A \vee \neg A$  (Prinzip vom ausgeschlossenen Dritten, tertium non datur)

$\neg(A \wedge B) \leftrightarrow \neg A \vee \neg B$ ,  $\neg(A \vee B) \leftrightarrow \neg A \wedge \neg B$  (De-Morgan-Regeln)

$A \rightarrow B \leftrightarrow (\neg B \rightarrow \neg A)$  (Kontrapositionsgesetz)

$(A \rightarrow B) \wedge (\neg A \rightarrow B) \rightarrow B$  (Fallunterscheidung)

## 2. Quantoren

Zur Formulierung mathematischer Aussagen werden neben den Junktoren häufig Quantoren der Form „für alle“ und „es gibt ein“ verwendet. Einige Beispiele sind:

Es gibt ein  $x$  mit  $f(x) = 0$ . *(f besitzt eine Nullstelle)*

Für alle  $n$  gibt es ein  $p \geq n$ , sodass  $p$  und  $p + 2$  prim sind. *(Existenz unendlich vieler Primzahlzwillinge)*

Es gibt ein  $e$  mit  $x \circ e = e \circ x = x$ . *(Existenz eines neutralen Elements)*

Für alle  $x$  gibt es ein  $y$  mit  $x \circ y = y \circ x = e$ . *(Existenz inverser Elemente)*

Wie für die Junktoren können wir eine Tabelle angeben:

Zeichen	Bedeutung	Name
$\forall$	für alle ...	Allquantor
$\exists$	es gibt (mindestens) ein ...	Existenzquantor

Die drei wichtigsten Quantorenregeln sind:

$\neg \forall x A(x)$	$\leftrightarrow$	$\exists x \neg A(x)$
$\neg \exists x A(x)$	$\leftrightarrow$	$\forall x \neg A(x)$
$\exists x \forall y A(x, y)$	$\rightarrow$	$\forall y \exists x A(x, y)$

### Beispiele

(1) Sei  $A(x)$  = „Der Zwerg  $x$  hat rote Haare.“ Dann bedeuten

$\neg \forall x A(x)$ : „Nicht jeder Zwerg hat rote Haare.“

$\exists x \neg A(x)$ : „Es gibt einen Zwerg, der keine roten Haare hat.“

Diese Aussagen sind äquivalent.

(2) Sei  $A(x, y)$  = „Der Lehrer  $x$  unterrichtet das Fach  $y$ .“ Dann bedeuten:

$\exists x \forall y A(x, y)$ : „Es gibt einen Lehrer, der jedes Fach unterrichtet.“

$\forall y \exists x A(x, y)$ : „Jedes Fach wird von mindestens einem Lehrer unterrichtet.“

Die erste Aussage ist für viele Schulen falsch, die zweite für die meisten Schulen richtig.

### 3. Zum Funktionsbegriff

#### Abbildungen und ihre Graphen

Eine Funktion (gleichwertig: Abbildung, Zuordnung, Operator, Familie, Transformation) ist eine rechtseindeutige Relation (vgl. 1.4). Dadurch wird eine Funktion  $f$  mit ihrem Graphen identifiziert:

$$f = \text{graph}(f) = \{(a, b) \mid f(a) = b\} = \{(a, b) \mid (a, b) \in f\}.$$

#### Bemerkung

Diese Definition ist in der mathematischen Grundlagenforschung vorherrschend. In der Algebra wird eine Funktion oft auch als Tripel  $f = (\text{graph}(f), A, B)$  in der Bedeutung  $f: A \rightarrow B$  erklärt. Dadurch ergeben sich kleinere begriffliche und notationelle Unterschiede.

#### Definitionsbereich, Wertebereich, Wertevorrat

Jede Funktion  $f$  hat einen eindeutig bestimmten Definitionsbereich (engl. domain)

$$\text{Def}(f) = \{a \mid f(a) \text{ ist definiert}\} = \{a \mid \text{es gibt ein } b \text{ mit } f(a) = b\}$$

und einen eindeutig bestimmten Wertebereich (engl. range)

$$\text{Bild}(f) = \{f(a) \mid a \in \text{Def}(f)\}.$$

Wir schreiben  $f: A \rightarrow B$ , falls  $A = \text{Def}(f)$  und  $\text{Bild}(f) \subseteq B$ . Die Menge  $B$  heißt dann ein *Wertevorrat* oder eine *Zielmenge* für  $f$ . Ein Wertevorrat ist nicht eindeutig bestimmt. Für jede Obermenge  $B$  von  $\text{Bild}(f)$  gilt  $f: A \rightarrow B$ .

#### Bemerkung

- (a) Die Sprechweisen über Wertebereich und Wertevorrat sind nicht einheitlich.
- (b) In der Tripeldefinition ist  $(\text{graph}(f), A, B) \neq (\text{graph}(f), A, C)$  für  $B \neq C$ .  
Ein Wertevorrat gehört hier zu einer Funktion fest mit dazu.

#### Terme und Variablen

Eine Funktion kann durch einen Term definiert sein, muss es aber nicht. In vielen Fällen ist eine Termdefinition nicht möglich (auch in der Analysis nicht, vgl. 1.5). Um die Sprechweise und Notation zu vereinfachen, wird oft vereinbart, dass eine Funktion mit einem sie definierenden Term gleichgesetzt wird. Der Definitionsbereich muss aber aus dem Kontext heraus klar werden (Beispiel: die Funktion  $x^2$  auf  $[0, \infty[ \subseteq \mathbb{R}$ ). Eine Variable muss bei der Angabe einer Funktion nicht angegeben werden (also einfach  $f$  statt  $f(x)$ ). Auch hier gilt, dass die Angabe einer Variablen suggestiv und notationell vorteilhaft sein kann, man denke etwa an  $x(t)$  für eine zeitabhängige Ortsfunktion  $x: \mathbb{R} \rightarrow \mathbb{R}$  oder die Bedeutung der Variablen bei der Berechnung von Integralen.

## 4. Zahlen

Wir stellen die wesentlichen Strukturmerkmale des Zahlensystems

$$\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}$$

im Überblick vor.

### Die natürlichen Zahlen

Die natürlichen Zahlen sind durch ein Anfangselement 0 und die Nachfolgebildung, die jeder natürlichen Zahl  $n$  ihren Nachfolger  $S(n)$  zuordnet, bestimmt:

$$0, S(0), S(S(0)), S(S(S(0))), \dots, n, S(n), \dots$$

Das Anfangselement ist kein Nachfolger und je zwei verschiedene Zahlen haben verschiedene Nachfolger. Weiter wird jede Zahl  $n$  von der 0 aus erreicht, wenn wir  $S$  oft genug anwenden. Genauer kann man zeigen, dass es bis auf Isomorphie (bis auf die Namen der Zahlen) genau eine Struktur  $(\mathbb{N}, S, 0)$  mit  $S : \mathbb{N} \rightarrow \mathbb{N}$ ,  $0 \in \mathbb{N}$  gibt mit:

$$(a) \quad \forall n \quad S(n) \neq 0, \quad (\text{Nachfolgeraxiom 1})$$

$$(b) \quad \forall n, m \quad (n \neq m \rightarrow S(n) \neq S(m)), \quad (\text{Nachfolgeraxiom 2})$$

$$(c) \quad \forall X \subseteq \mathbb{N} \quad (0 \in X \wedge \forall n \quad (n \in X \rightarrow S(n) \in X) \rightarrow X = \mathbb{N}). \quad (\text{Induktionsaxiom})$$

Die Aussagen (a), (b), (c) heißen die *Dedekind-Peano-Axiome*. Die Funktion  $S$  heißt die *Nachfolgerfunktion* und für alle  $n \in \mathbb{N}$  heißt  $S(n)$  der (*unmittelbare*) *Nachfolger* von  $n$  (das „S“ steht für „successor“). Weiter heißt die 0 das *Anfangselement* von  $\mathbb{N}$ .

Mit Hilfe der Axiome der Mengenlehre kann man eine Struktur  $(\mathbb{N}, S, 0)$  konstruieren, die die Dedekind-Peano-Axiome erfüllt. Für die heute bevorzugte Konstruktion gilt:

$$0 = \emptyset, \quad 1 = S(0) = \{0\}, \quad 2 = S(1) = \{0, 1\}, \quad \dots, \quad S(n) = n \cup \{n\} = \{0, \dots, n\}, \quad \dots$$

Damit sind die natürlichen Zahlen als Mengen eingeführt.

Aus der Nachfolgerfunktion  $S$  lässt sich die gesamte Arithmetik auf  $\mathbb{N}$  gewinnen. Für alle  $m \in \mathbb{N}$  definiert man durch Rekursion nach  $n \in \mathbb{N}$ :

$$m + 0 = m, \quad m + S(n) = S(m + n), \quad (\text{Addition auf } \mathbb{N})$$

$$m \cdot 0 = 0, \quad m \cdot S(n) = m \cdot n + m, \quad (\text{Multiplikation auf } \mathbb{N})$$

$$m^0 = 1, \quad m^{S(n)} = m^n \cdot m. \quad (\text{Exponentiation auf } \mathbb{N})$$

Die bekannten Rechenregeln lassen sich nun mit Hilfe des Induktionsaxioms beweisen.

Auf  $\mathbb{N}$  erhält man eine lineare Ordnung  $\leq$ , indem man für alle  $n, m \in \mathbb{N}$  setzt:

$$n \leq m, \quad \text{falls} \quad \exists k \quad n + k = m. \quad (\text{Ordnung auf } \mathbb{N})$$

Diese Ordnung ist eine *Wohlordnung* auf  $\mathbb{N}$ , d.h., jede nichtleere Teilmenge von  $\mathbb{N}$  besitzt ein kleinstes Element (*Prinzip vom kleinsten Element*).

## Die ganzen Zahlen

Die Idee ist, ganze Zahlen als Paare  $(n, m)$  von natürlichen Zahlen  $n, m$  einzuführen. Dabei steht  $(n, m)$  intuitiv für  $n - m$ , speziell also  $(n, 0)$  für  $n$  und  $(0, m)$  für  $-m$ . Da bei dieser Lesart  $(n, m)$  und  $(n + k, m + k)$  für alle  $k$  gleich sind, ist die Einführung einer Äquivalenzrelation nötig. Für alle  $(n, m), (n', m') \in \mathbb{N}$  setzt man:

$$(n, m) \sim (n', m'), \text{ falls } n + m' = n' + m,$$

$$\mathbb{Z} = \mathbb{N}^2 / \sim, \quad (\text{Menge der ganzen Zahlen})$$

$$[n, m] = (n, m) / \sim, \quad (\text{Vereinfachung der Notation})$$

$$[n, m] + [n', m'] = [n + n', m + m'], \quad (\text{Addition auf } \mathbb{Z})$$

$$[n, m] \cdot [n', m'] = [n n' + m m', n m' + m n']. \quad (\text{Multiplikation auf } \mathbb{Z})$$

Die Struktur  $(\mathbb{Z}, +, \cdot)$  ist ein Ring. Durch Identifikation von  $n$  mit  $[n, 0]$  kann man  $\mathbb{N} \subseteq \mathbb{Z}$  annehmen. Weiter gelten

$$-[n, m] = [m, n],$$

$$n - m = [n, 0] - [m, 0] = [n, 0] + [0, m] = [n, m] \text{ für alle } n, m \in \mathbb{N}.$$

Damit haben wir die ursprüngliche Motivation wiedergefunden. Wir können  $[n, m]$  durch  $n - m$  ersetzen.

Eine lineare Ordnung auf  $\mathbb{Z}$  erhält man durch

$$[n, m] \leq [n', m'], \text{ falls } n + m' \leq m + n' \text{ (in } \mathbb{N}). \quad (\text{Ordnung auf } \mathbb{Z})$$

## Die rationalen Zahlen

Ähnlich verläuft die Konstruktion von  $\mathbb{Q}$ . Die Idee ist, dass das Paar  $(a, b) \in \mathbb{Z} \times \mathbb{Z}^*$  den Bruch  $a/b$  repräsentiert (mit  $\mathbb{Z}^* = \mathbb{Z} - \{0\}$ ). Für alle  $(a, b), (c, d) \in \mathbb{Z} \times \mathbb{Z}^*$  setzen wir:

$$(a, b) \sim (c, d), \text{ falls } ad = bc,$$

$$\mathbb{Q} = (\mathbb{Z} \times \mathbb{Z}^*) / \sim, \quad (\text{Menge der rationalen Zahlen})$$

$$a/b = (a, b) / \sim, \quad (\text{Vereinfachung der Notation})$$

$$a/b + c/d = (ad + bc)/(bd), \quad (\text{Addition auf } \mathbb{Q})$$

$$a/b \cdot c/d = (ac)/(bd). \quad (\text{Multiplikation auf } \mathbb{Q})$$

Es entsteht so der Körper  $(\mathbb{Q}, +, \cdot)$ . Durch Identifikation von  $a/1$  mit  $a$  für alle  $a \in \mathbb{Z}$  erhält man  $\mathbb{Z} \subseteq \mathbb{Q}$ . Es gilt

$$b^{-1} = (b/1)^{-1} = 1/b \text{ für alle } b \in \mathbb{Z}^*,$$

$$a \cdot b^{-1} = a \cdot 1/b = a/b \text{ für alle } a, b \in \mathbb{Z} \text{ mit } b \neq 0.$$

Schließlich wird  $\mathbb{Q}$  zu einem angeordneten Körper durch die Definition

$$a/b \leq c/d, \text{ falls } ad \leq bc \text{ (in } \mathbb{Z}) \text{ für alle } a, b, c, d \in \mathbb{Z} \text{ mit } b, d > 0. \quad (\text{Ordnung auf } \mathbb{Q})$$

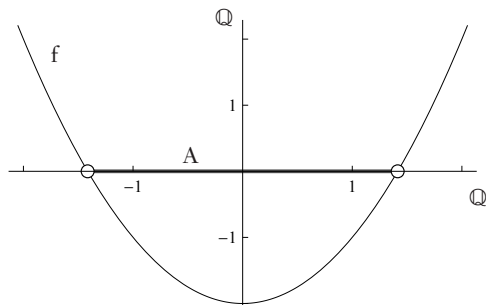
## Die reellen Zahlen

Der Schritt von  $\mathbb{Q}$  nach  $\mathbb{R}$  ist der komplizierteste, aber auch spannendste Erweiterungsschritt. Während  $\mathbb{N}$ ,  $\mathbb{Z}$  und  $\mathbb{Q}$  abzählbar sind, entsteht nun eine überabzählbare Menge von Zahlen.

Ausgangspunkt ist die Existenz irrationaler Größen wie  $\sqrt{2}$ ,  $\sqrt{3}$ ,  $\pi$ ,  $e$ . Die Irrationalität von  $\sqrt{2}$  führt zum Beispiel dazu, dass  $f: \mathbb{Q} \rightarrow \mathbb{Q}$  mit

$$f(q) = q^2 - 2 \text{ für alle } q \in \mathbb{Q}$$

keine Nullstellen besitzt. Der Zahlenstrahl  $\mathbb{Q}$  ist damit kein Kontinuum, wie es in der Analysis benötigt wird. Von einem Kontinuum verlangen wir folgende Eigenschaft:



Jede nichtleere und beschränkte Menge  $X$  von Punkten besitzt ein Supremum (kleinste obere Schranke), d. h., es gibt einen Punkt  $s$  mit:

- (a)  $x \leq s$  für alle  $x \in X$     (b) Ist  $t$  ein Punkt mit  $x \leq t$  für alle  $x \in X$ , so ist  $s \leq t$ .

(lineares Vollständigkeitsaxiom)

Ist  $s$  das Supremum von  $X$ , so schreibt man  $s = \sup(X)$ . Analog wird das Infimum  $s = \inf(X)$  einer beschränkten nichtleeren Menge  $X$  als die größte untere Schranke von  $X$  definiert. Anschaulich ist  $\sup(X)$  der rechte und  $\inf(X)$  der linke Randpunkt der Menge  $X$ . Diese Punkte können der Menge  $X$  als Element angehören oder nicht.

Es ist hier nicht der Ort, die Konstruktion eines Kontinuums im Detail zu diskutieren. Im Wesentlichen besteht sie darin, alle Lücken von  $\mathbb{Q}$  mit neuen „irrationalen“ Zahlen zu füllen. Eine Lücke von  $\mathbb{Q}$  wird dabei durch eine nichtleere und beschränkte Teilmenge  $A$  von  $\mathbb{Q}$  markiert, die kein Supremum oder kein Infimum in  $\mathbb{Q}$  besitzt. Ein Beispiel ist

$$A = \{q \in \mathbb{Q} \mid q^2 < 2\}.$$

Die Menge  $A$  besitzt obere Schranken in  $\mathbb{Q}$  wie zum Beispiel 2,  $3/2$  oder  $1,42$ , aber kein Supremum in  $\mathbb{Q}$ . Jede obere Schranke von  $A$  in  $\mathbb{Q}$  kann noch verkleinert werden. Analog lässt sich jede untere Schranke von  $A$  in  $\mathbb{Q}$  noch vergrößern. Die durch  $A$  markierten Lücken von  $\mathbb{Q}$  entsprechen genau den Nullstellen der oben betrachteten Funktion  $f$ .

Zum Füllen der Lücken von  $\mathbb{Q}$  stehen verschiedene Methoden zur Verfügung. Man kann zum Beispiel Teilmengen von  $\mathbb{Q}$  (Konstruktion von Dedekind) oder Folgen rationaler Zahlen verwenden (Konstruktion von Cantor). Insgesamt entsteht ein angeordneter Körper  $\mathbb{R}$ , der das Vollständigkeitsaxiom erfüllt. Man kann weiter zeigen, dass dieser Körper bis auf Isomorphie eindeutig bestimmt ist, sodass alle Konstruktionen von  $\mathbb{R}$  äquivalent sind. In  $\mathbb{R}$  stehen nun die Zahlgrößen  $\sqrt{2}$ ,  $\sqrt{3}$ ,  $e$ ,  $\pi$ , ... zur Verfügung. Dabei ist „...“ trügerisch, da es überabzählbar viele irrationale Zahlen gibt. Das Vollständigkeitsaxiom ist ohne Überabzählbarkeit nicht zu haben.



## Die komplexen Zahlen

Die Erweiterung der reellen Zahlen  $\mathbb{R}$  zu den komplexen Zahlen  $\mathbb{C}$  ist aus technischer Sicht wieder leicht möglich. Die Idee ist, den reellen Zahlen  $\mathbb{R}$  negative Wurzeln hinzuzufügen, sodass zum Beispiel die Gleichung  $x^2 + 1 = 0$  eine Lösung im erweiterten Zahlbereich besitzt. Diese Erweiterung lässt sich anschaulich durchführen, indem man die Punkte der Ebene als Zahlen auffasst. Man setzt

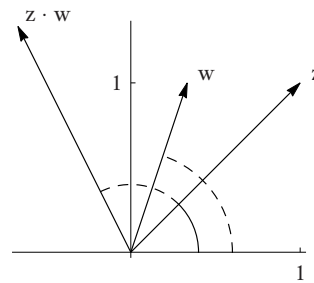
$$\mathbb{C} = \mathbb{R}^2, \quad (\text{Menge der komplexen Zahlen})$$

$$(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2), \quad (\text{komplexe Addition})$$

$$(x_1, y_1) \cdot (x_2, y_2) = (x_1 x_2 - y_1 y_2, x_1 y_2 + x_2 y_1) \quad (\text{komplexe Multiplikation})$$

für alle  $(x_1, y_1), (x_2, y_2) \in \mathbb{C}$ . Die Addition ist die übliche Vektoraddition in  $\mathbb{R}^2$ . Auch die Multiplikation hat eine einfache geometrische Interpretation. Sind  $z, w \in \mathbb{C}$ , so gilt:

- (a) Die Länge von  $z \cdot w$  ist das Produkt der Längen von  $z$  und  $w$ .
- (b) Der (gegen den Uhrzeigersinn gemessene) Winkel, den  $z \cdot w$  mit der positiven  $x$ -Achse einschließt, ist die Summe der entsprechenden Winkel von  $z$  und  $w$ .



Multipliziere die Längen und addiere die Winkel.

Die komplexen Zahlen  $\mathbb{C}$  bilden einen Körper mit  $0 = (0, 0)$  und  $1 = (1, 0)$ . Durch die Identifikation von  $x \in \mathbb{R}$  mit  $(x, 0) \in \mathbb{C}$  erreicht man  $\mathbb{R} \subseteq \mathbb{C}$ . Weiter setzt man

$$\operatorname{Re}(x, y) = x, \quad \operatorname{Im}(x, y) = y, \quad (\text{Realteil bzw. Imaginärteil von } (x, y))$$

$$i = (0, 1). \quad (\text{imaginäre Einheit})$$

Es gilt  $i^2 = (0, 1)^2 = (0 - 1, 0 + 0) = (-1, 0) = -1$ , sodass  $i$  eine Lösung von  $z^2 + 1 = 0$  ist. Allgemein hat jedes Polynom über  $\mathbb{C}$  eine komplexe Nullstelle (Fundamentalsatz der Algebra).

Die imaginäre Einheit dominiert das Rechnen mit den komplexen Zahlen. Es gilt

$$(x, y) = x + iy = \operatorname{Re}(x, y) + i \operatorname{Im}(x, y) \quad \text{für alle } x, y \in \mathbb{R},$$

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

$$x_1 x_2 - y_1 y_2 + i(x_1 y_2 + x_2 y_1) = (x_1 x_2 - y_1 y_2, x_1 y_2 + x_2 y_1) \quad \text{für alle } x_1, y_1, x_2, y_2 \in \mathbb{R}.$$

Die komplexe Multiplikation lässt sich also aus  $i^2 = -1$  durch Ausmultiplizieren motivieren (und rekonstruieren).

Eine wichtige Operation ist die *Konjugation*, die für alle  $z = (x, y) \in \mathbb{C}$  definiert ist durch

$$\bar{z} = \operatorname{Re}(z) - i \operatorname{Im}(z) = x - iy. \quad (\text{komplex Konjugierte von } z)$$

Sie entspricht der Spiegelung an der  $x$ -Achse. Für alle  $z \in \mathbb{C}$  gilt  $z \bar{z} = \operatorname{Re}(z)^2 + \operatorname{Im}(z)^2 \in \mathbb{R}$ , sodass  $z \bar{z}$  das Quadrat der euklidischen Länge  $|z| = \sqrt{x^2 + y^2}$  von  $z = (x, y)$  ist.

## 5. Geometrische Grundlagen

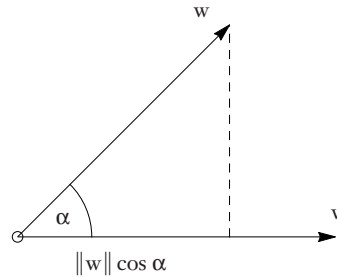
Die euklidische Ebene ist definiert durch

$$\mathbb{R}^2 = \mathbb{R} \times \mathbb{R} = \{ (x, y) \mid x, y \in \mathbb{R} \} = \{ v \mid v = (v_1, v_2) \text{ mit } v_1, v_2 \in \mathbb{R} \}.$$

Für alle Vektoren  $v = (v_1, v_2)$ ,  $w = (w_1, w_2)$  der Ebene definieren wir das euklidische Skalarprodukt  $\langle v, w \rangle$  von  $v$  und  $w$  und die Norm oder Länge  $\|v\|$  von  $v$  durch

$$\langle v, w \rangle = v \bullet w = v_1 w_1 + v_2 w_2,$$

$$\|v\| = \sqrt{\langle v, v \rangle} = \sqrt{v_1^2 + v_2^2}.$$



$$\langle v, w \rangle = \langle w, v \rangle = \|v\| \|w\| \cos \alpha$$

### Kreise und Ellipsen

Für alle  $r > 0$  ist die Kreislinie mit Radius  $r$  und Mittelpunkt  $0$  definiert durch

$$K_r = \{ v \in \mathbb{R}^2 \mid \|v\| = r \} = \{ (x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = r^2 \} = \{ (\cos \alpha, \sin \alpha) \mid \alpha \in [0, 2\pi[ \}.$$

Für alle  $a, b \in \mathbb{R}$  ist

$$E_{a,b} = \{ (ax, by) \mid (x, y) \in K_1 \} = \{ (a \cos \alpha, b \sin \alpha) \mid \alpha \in [0, 2\pi[ \}.$$

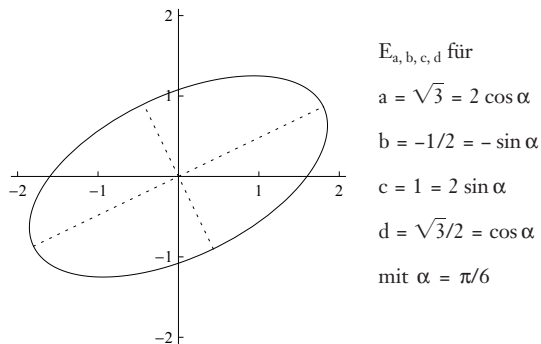
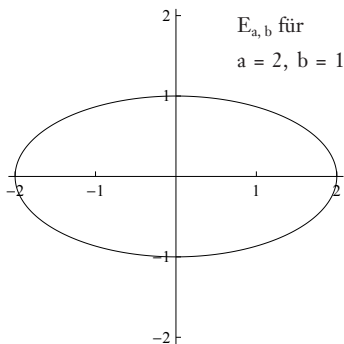
eine achsenparallele Ellipse mit den Halbachsen  $|a|$  und  $|b|$  (ist  $a = 0$  oder  $b = 0$ , so ist die Ellipse degeneriert). Für  $a, b \neq 0$  erhält man die Darstellung

$$E_{a,b} = \left\{ (x, y) \in \mathbb{R}^2 \mid \left( \frac{x}{a} \right)^2 + \left( \frac{y}{b} \right)^2 = 1 \right\}.$$

Allgemeine Ellipsen mit Mittelpunkt  $0$  entstehen aus den achsenparallelen Ellipsen durch Drehung. Sie haben (was keineswegs trivial ist) die Form

$$E_{a,b,c,d} = \{ (ax + by, cx + dy) \mid (x, y) \in K_1 \}, \text{ mit beliebigen } a, b, c, d \in \mathbb{R}.$$

Eine Ellipse erscheint so als das Bild des Einheitskreises unter einer linearen Abbildung (vgl. 4.6 und 8.9).

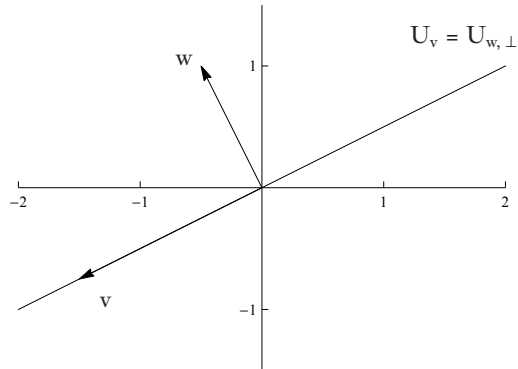


## Geraden

Für alle  $v \in \mathbb{R}^2 - \{0\}$  ist

$$U_v = \{ \alpha v \mid \alpha \in \mathbb{R} \}$$

die durch den Richtungsvektor  $v$  definierte Gerade durch den Nullpunkt. Eine alternative Möglichkeit, eine Gerade zu definieren, ist, alle auf einem bestimmten Vektor  $w \neq 0$  senkrecht stehenden Vektoren zu betrachten:



$$U_{w, \perp} = \{ v \in \mathbb{R}^2 \mid \langle v, w \rangle = 0 \} = \{ (x, y) \in \mathbb{R}^2 \mid xw_1 + yw_2 = 0 \}.$$

## Die Dimension $n = 3$

Im dreidimensionalen Raum  $\mathbb{R}^3 = \mathbb{R}^2 \times \mathbb{R} = \{ (v_1, v_2, v_3) \mid v_1, v_2, v_3 \in \mathbb{R} \}$  sind das euklidische Skalarprodukt und die euklidische Länge definiert durch

$$\langle v, w \rangle = v \bullet w = v_1 w_1 + v_2 w_2 + v_3 w_3,$$

$$\|v\| = \sqrt{\langle v, v \rangle} = \sqrt{v_1^2 + v_2^2 + v_3^2}.$$

Für alle  $r > 0$  ist

$$K_r = \{ v \in \mathbb{R}^3 \mid \|v\| = r \} = \{ (x, y, z) \in \mathbb{R}^3 \mid x^2 + y^2 + z^2 = r^2 \}$$

die Oberfläche einer Kugel mit Radius  $r$  und Mittelpunkt  $0$ . Weiter ist

$$E_{a, b, c} = \{ (ax, by, cz) \mid (x, y, z) \in K_1 \}$$

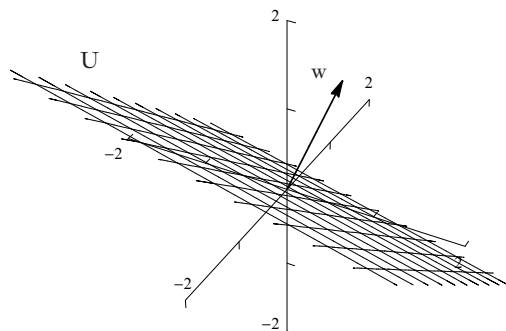
ein achsenparalleles Ellipsoid mit den Halbachsen  $|a|$ ,  $|b|$ ,  $|c|$ . Allgemeine Ellipsoide mit Mittelpunkt  $0$  haben die Form (vgl. 8.9)

$$\{ (a_{11}x + a_{12}y + a_{13}z, a_{21}x + a_{22}y + a_{23}z, a_{31}x + a_{32}y + a_{33}z) \mid (x, y, z) \in K_1 \}.$$

Sind  $u$  und  $v$  Vektoren des  $\mathbb{R}^3$ , die nicht auf einer gemeinsamen Geraden liegen, so ist

$$U = \{ \alpha u + \beta v \mid \alpha, \beta \in \mathbb{R} \}$$

eine Ebene des  $\mathbb{R}^3$ . Alternativ kann man eine Ebene als Menge  $\{ v \in \mathbb{R}^3 \mid \langle v, w \rangle = 0 \}$  aller Vektoren definieren, die auf einem Vektor  $w \neq 0$  senkrecht stehen.



## 6. Die Axiome der Mengenlehre

Die gesamte Mathematik lässt sich aus dem Mengenbegriff entwickeln. Zahlen, Relationen, Funktionen, algebraische Strukturen usw. lassen sich als Mengen einführen. Neben der Gleichheit  $=$  wird dabei nur die Elementbeziehung  $\in$  verwendet. Die Eigenschaften von  $\in$  werden durch Axiome beschrieben, auf die sich ein Mathematiker bewusst oder unbewusst stützt. Wir stellen die weit verbreitete Zermelo-Fraenkel-Axiomatik ZFC kurz vor (Z = Ernst Zermelo, F = Abraham Fraenkel, C = „axiom of choice“ = Auswahlaxiom). Dieses System besteht aus den folgenden Axiomen:

### I. Extensionalitätsaxiom

Zwei Mengen sind genau dann gleich, wenn sie dieselben Elemente besitzen.

Eine Menge ist also durch ihre Elemente vollständig bestimmt.

### II. Existenz der leeren Menge

Es gibt eine Menge, die kein Element enthält.

Die leere Menge wird mit  $\emptyset$  oder  $\{ \}$  bezeichnet.

### III. Paarmengenaxiom

Zu je zwei Mengen  $a, b$  existiert eine Menge  $c$ , die genau  $a$  und  $b$  als Elemente hat.

Wir schreiben  $c = \{ a, b \}$ . Mit Hilfe des Axioms können wir  $(a, b) = \{ \{ a \}, \{ a, b \} \}$  setzen und damit Relationen und Funktionen einführen (dabei ist  $\{ a \} = \{ a, a \}$ ).

Hier und im Folgenden verwenden wir kleine Buchstaben  $a, b, c, \dots$  für Mengen. Da jedes Objekt der Theorie eine Menge ist, ist jedes Objekt auch eine Menge von Mengen, sodass die Unterscheidung zwischen Punkt/Zahl, Menge, Mengensystem in der axiomatischen Mengenlehre streng genommen bedeutungslos ist. Natürlich werden Mengen in konkreten Kontexten suggestiv mit  $M, \mathcal{A}$  usw. bezeichnet.

### IV. Aussonderungsschema

Zu jeder Eigenschaft  $\mathcal{E}$  und jeder Menge  $a$  gibt es eine Menge  $b$ , die genau die Elemente  $c$  von  $a$  enthält, auf die  $\mathcal{E}$  zutrifft.

Wir schreiben  $b = \{ c \in a \mid \mathcal{E}(c) \}$ . Für alle  $d$  gilt  $d \in b$  genau dann, wenn  $d \in a$  und  $\mathcal{E}(d)$ . Da jeder Eigenschaft ein Axiom entspricht, spricht man von einem *Axiomschema*. Das System ZFC umfasst damit unendlich viele Axiome. Das Aussonderungsschema ist ein Ersatz für das inkonsistente Komprehensionsschema, das die Bildung von

$$\{ c \mid \mathcal{E}(c) \} \quad (\text{unbeschränkte Komprehension})$$

und damit die Russell-Zermelo-Komprehension  $\{ c \mid c \notin c \}$  erlaubt. Aussonderung genügt in vielen Fällen, da zumeist ein „großer Bereich“ wie  $\mathbb{N}, \mathbb{R}, \mathbb{R}^2$  usw. untersucht wird, für dessen Teilmengen man sich interessiert. Das Aussonderungsschema wird nun durch weitere Axiome ergänzt, die die Bildung dieser „großen Bereiche“ ermöglichen.

## V. Vereinigungsmengenaxiom

Zu jeder Menge  $a$  existiert eine Menge  $b$ , deren Elemente genau die Elemente der Elemente von  $a$  sind.

Wir schreiben  $b = \bigcup a$  und setzen  $a \cup b = \bigcup \{a, b\}$ .

## VI. Unendlichkeitsaxiom

Es existiert eine Menge  $a$ , die die leere Menge als Element enthält, und die mit jedem ihrer Elemente  $b$  auch  $b \cup \{b\}$  als Element enthält.

Man setzt  $0 = \emptyset$ ,  $1 = \{0\}$ ,  $2 = 1 \cup \{1\}$ , ...,  $n + 1 = n \cup \{n\}$ , ... Aus dem Axiom folgt, dass  $\mathbb{N} = \{0, 1, 2, \dots\}$  existiert. Genauer nennt man eine Menge  $a$  wie im Unendlichkeitsaxiom induktiv und definiert  $\mathbb{N}$  als den Durchschnitt aller induktiven Mengen. Man kann zeigen, dass  $(\mathbb{N}, S, 0)$  mit  $S(n) = n \cup \{n\}$  für alle  $n \in \mathbb{N}$  die Dedekind-Peano-Axiome erfüllt.

## VII. Potenzmengenaxiom

Zu jeder Menge  $a$  existiert die Menge  $b$  aller Teilmengen von  $a$ .

Wir schreiben  $b = \mathcal{P}(a)$ . Das Axiom führt zu überabzählbaren Mengen. Aus  $\mathcal{P}(\mathbb{N})$  gewinnt man die reellen Zahlen  $\mathbb{R}$  und aus  $\mathcal{P}(\mathbb{R})$  die Menge aller Funktionen von  $\mathbb{R}$  nach  $\mathbb{R}$ .

## VIII. Ersetzungsschema

Sei  $\mathcal{E}(a, b)$  eine Eigenschaft derart, dass für jede Menge  $a$  genau eine Menge  $b$  existiert mit  $\mathcal{E}(a, b)$ . Dann existiert für jede Menge  $c$  die Menge, die entsteht, wenn jedes Element  $a$  von  $c$  durch das eindeutige  $b$  mit  $\mathcal{E}(a, b)$  ersetzt wird.

Das Schema erlaubt die Definition von Funktionen  $f$  auf einer Menge  $c$  der Form

$f(a) =$  „das eindeutige  $b$  mit  $\mathcal{E}(a, b)$ “ für alle  $a \in c$

(vgl. 1. 5). Oft genügt das Aussonderungsschema: Weiß man, dass alle  $(a, b)$  mit  $\mathcal{E}(a, b)$  einer Menge  $d$  angehören, so ist  $f = \{(a, b) \in d \mid \mathcal{E}(a, b)\}$ . Das Ersetzungsschema erweist sich aber als eine echte Verstärkung des Aussonderungsschemas.

## IX. Fundierungsaxiom oder Regularitätsaxiom

Jede nichtleere Menge  $a$  besitzt ein Element  $b$  mit  $a \cap b = \emptyset$ .

Das Fundierungsaxiom wird außerhalb der Mengenlehre kaum benutzt. In der Mengenlehre ermöglicht es einen stufenweisen Aufbau des Mengenuniversums, bei dem jeder Menge ein Maß für ihre Komplexität zugewiesen wird.

## X. Auswahlaxiom

Ist  $a$  eine Menge, deren Elemente nichtleer und paarweise disjunkt sind, so existiert eine Menge  $b$ , die mit jedem Element von  $a$  genau ein Element gemeinsam hat.

Wir verweisen den Leser auf 1. 11 für eine Diskussion des Auswahlaxioms.



---

# Literatur

---

**Birkhoff, Garrett / Mac Lane, Saunders** 1965 *A Survey of Modern Algebra*. Macmillan, New York, 3. Auflage 1965.

**Bosch, Siegfried** 2009 *Lineare Algebra*. Springer, Berlin, 4. Auflage 2009.

– 2013 *Algebra*. Springer Spektrum, Berlin, 8. Auflage 2013.

**Deiser, Oliver** 2008 *Reelle Zahlen*. Springer, Berlin, 2. Auflage 2008.

– 2009 *Einführung in die Mengenlehre*. Springer, Berlin, 3. Auflage 2009.

– 2010 *Grundbegriffe der wissenschaftlichen Mathematik*. Springer, Berlin, 2010.

– 2012 *Erste Hilfe in Analysis*. Springer, Berlin, 2012.

– 2013 *Analysis 1*. Springer, Berlin, 2. Auflage 2013.

– 2015 *Analysis 2*. Springer, Berlin, 2. Auflage 2015.

**Deiser, Oliver / Lasser, Caroline** 2015 *Erste Hilfe in linearer Algebra*. Springer, Berlin, 2015.  
Erste Auflage dieses Buches.

**Deiser, Oliver / Lasser, Caroline / Vogt, Elmar / Werner, Dirk** 2016 *12 x 12  
Schlüsselkonzepte zur Mathematik*. Springer, Berlin, 2. Auflage 2016.

**Ebbinghaus, Heinz-Dieter et al.** 1992 *Zahlen*. Springer, Berlin, 3. Auflage 1992.

**Fischer, Gerd** 2012 *Lernbuch Lineare Algebra und Analytische Geometrie*. Springer Spektrum, Berlin, 2. Auflage 2012.

**Halmos, Paul** 1994 *Naïve Mengenlehre*. Vandenhoeck & Ruprecht, Göttingen 5. Auflage 1994.

**Huppert, Bertram / Willems, Wolfgang** 2010 *Lineare Algebra*. Vieweg+Teubner, Wiesbaden, 2. Auflage 2010.

**Jänich, Klaus** 2013 *Lineare Algebra*. Springer, Berlin, korrigierter Nachdruck der 11. Auflage 2013.

**Koecher, Max** 2003 *Lineare Algebra und analytische Geometrie*. Springer, Berlin, korrigierter Nachdruck der 4. Auflage 2003.

**Kowalsky, Hans-Joachim / Michler, Gerhard** 2003 *Lineare Algebra*. De Gruyter, Berlin, 12. Auflage 2003.

- Lipschutz, Seymour** 1988 *Lineare Algebra*. McGraw-Hill, Hamburg, Nachdruck 1988.
- Reiss, Kristina / Schmieder, Gerald** 2007 *Basiswissen Zahlentheorie*. Springer, Berlin, 2. Auflage 2007.
- Lax, Peter** 1997 *Linear Algebra and its Applications*. Wiley, New Jersey, 2. Auflage 1997.
- Scheja, Günter / Storch, Uwe** 1994 *Lehrbuch der Algebra 1*. B. G. Teubner, Stuttgart, 2. Auflage 1994.
- 1988 *Lehrbuch der Algebra 2*. B. G. Teubner, Stuttgart, 1988.
- Stroth, Gernot** 2008 *Lineare Algebra*. Heldermann, Berlin, 2008.
- Trefethen, Lloyd / Bau, David** 1997 *Numerical Linear Algebra*. SIAM, Philadelphia 1997.



# Notationen

<b>Griechisches Alphabet</b>			$\emptyset$	10	$\mathcal{A}^c$	16
Alpha	A	$\alpha$	$N \subseteq M$	11	$a R b$	20
Beta	B	$\beta$	$N \subset M$	11	$\text{Def}(R)$	20
Gamma	$\Gamma$	$\gamma$	$M \supseteq N$	11	$\text{dom}(R)$	20
Delta	$\Delta$	$\delta$	$M \supset N$	11	$\text{Bild}(R)$	20
Epsilon	E	$\epsilon$	$\{x_1, \dots, x_n\}$	12	$\text{rng}(R)$	20
Zeta	Z	$\zeta$	$\{\}$	12	$a R b R c$	20
Eta	H	$\eta$	$(x, y)$	13	$d \mid a$	21
Theta	$\Theta$	$\theta, \vartheta$	$(x, y, z)$	13	$a \equiv_m b$	21
Jota	I	$\iota$	$(x_1, \dots, x_n)$	13	$a \equiv b \text{ mod}(m)$	21
Kappa	K	$\kappa$	$(x, y, z)$	13	$a/\sim$	22
Lambda	$\Lambda$	$\lambda$	$M = \{x \mid \mathcal{E}(x)\}$	14	$A/\sim$	22
My	M	$\mu$	$\emptyset$	15	$[a]_{\sim}$	22
Ny	N	$\nu$	$\{\}$	15	$[a]$	22
Xi	$\Xi$	$\xi$	$\{a\}$	15	$[a]_m$	23
Omikron	O	$\omicron$	$\{a, b\}$	15	$\mathbb{Z}_m$	23
Pi	$\Pi$	$\pi$	$\{a_1, \dots, a_n\}$	15	$X \leq a$	24
Rho	P	$\rho$	$(a, b)$	15	$a \geq X$	24
Sigma	$\Sigma$	$\sigma, \varsigma$	$(a, b, c)$	15	$a \leq X$	24
Tau	T	$\tau$	$A \times B$	15	$X \geq a$	24
Ypsilon	Y	$\upsilon$	$A \times B \times C$	15	$a = \max(X)$	24
Phi	$\Phi$	$\varphi$	$A^2, A^3, A^4$	15	$a = \min(X)$	24
Chi	X	$\chi$	$\{x \in A \mid \mathcal{E}(x)\}$	15	$a = \sup(X)$	24
Psi	$\Psi$	$\psi$	$\mathcal{P}(M)$	15	$a = \inf(X)$	24
Omega	$\Omega$	$\omega$	$A \cap B$	16	$\leq_{\text{lex}}$	25
$x \in M$		10	$A \cup B$	16	$f(a) = b$	26
$x \notin M$		10	$A - B$	16	$fa = b$	26
$\mathbb{N}$		10	$A \setminus B$	16	$f : a \mapsto b$	26
$\mathbb{Z}$		10	$A^c$	16	$a \xrightarrow{f} b$	26
$\mathbb{Q}$		10	$A \Delta B$	16	$\text{id}_A$	26
$\mathbb{R}$		10	$\bigcap \mathcal{A}$	16	$\text{const}_c^A$	26
			$\bigcup \mathcal{A}$	16	$\chi_B^A$	26

$1_B^A$	26	$\sum_{1 \leq k \leq 0} a_k$	55	$W_1 \oplus \dots \oplus W_n$	90
$\begin{pmatrix} 1 & \dots & n \\ b_1 & \dots & b_n \end{pmatrix}$	29	$1/a$	55	$\bigoplus_{i \in I} W_i$	90
$(b_1, \dots, b_n)$	29	$a/b$	55	$v + U$	92
$f : A \rightarrow B$	30	$m\mathbb{Z}$	57	$G \cong G'$	100
$f : A \ni a \mapsto f(a) \in B$	30	$\mathbb{Z}m$	57	$\text{Aut}(G)$	101
$(b_a)_{a \in A}$	30	$Ha$	58	$\text{Kern}(\varphi)$	102
$(b_a \mid a \in A)$	30	$aH$	58	$\text{Bild}(\varphi)$	102
$g \circ f$	32	$G/H$	58	$L_f(w)$	112
$f[X]$	32	$G/H$	58	$\Phi_B$	114
$f^{-1}[Y]$	32	$a + H$	59	$\text{Hom}_K(V, W)$	118
$f C$	32	$\mathbb{Z}/m\mathbb{Z}$	59	$\text{End}_K(V)$	118
$B^A$	32	$R^\times$	60	$V^*$	120
${}^AB$	32	$\mathbb{Z}_m$	61	$V^*$	120
$\prod_{i \in I} B_i$	32	$K^*$	62	$v_j^*$	120
$f(a_1, \dots, a_n)$	34	$a/b$	62	$\delta_{jk}$	120
$a + b, a \cdot b, a \circ b, \dots$	34	$\text{char}(K)$	63	$f^* : W^* \rightarrow V^*$	121
$\langle B \rangle$	34	$\text{sgn}(a)$	64	$A(i, j) = a_{ij}$	124
$\langle b_1, \dots, b_m \rangle$	34	$K^+$	64	$K^{m \times n}$	124
$f^{-1}$	36	$K_0^+$	64	$E_n$	125
$ A  \leq  B $	38	$R^{(\mathbb{N})}$	66	$\text{diag}(a_1, \dots, a_n)$	125
$ A  =  B $	38	$X = (0, 1, 0, 0, 0, \dots)$	66	$E_{ij}$	125
$ A  <  B $	38	$R[X]$	66	$f_A$	126
$a^n$	47	$\deg$	66	$\text{Kern}(A)$	126
$\prod_{1 \leq k \leq l} a_k$	47	$\mu_p(w)$	69	$\text{Bild}(A)$	126
$e$	48	$\prod_{i \in I} V_i$	76	$A_f^{\text{st}}$	130
$1$	48	$V^n$	76	$A^{-1}$	132
$0$	48	$V^I$	76	$\text{GL}(n, K)$	132
$a^0$	48	$K^n$	77	$W_{ij}(\lambda)$	134
$a^{-1}$	50	$K^I$	77	$P_\pi$	136
$a^{-n}$	50	$V^{(I)}$	77	$A^t$	140
$\mathbb{Z}_m$	51	$\text{span}(A)$	78	$\text{rang}(A)$	142
$\mathbb{Z}_m^*$	51	$\langle A \rangle$	78	$\langle x, y \rangle$	150
$S_A$	51	$\sum_{i \in I} \alpha_i v_i$	78	$\ x\ $	150
$S_n$	51	$e_i$	79	$\alpha(x, y)$	151
$M^\times$	51	$\sum_{i \in I} \alpha_i v_i$	82	$\langle z, w \rangle$	152
$na$	55	$\Phi_B(v)$	83	$\overline{A}$	153
$-a$	55	$\dim(V)$	86	$\langle v, w \rangle$	154
$a - b$	55	$\text{supp}(f)$	90	$\ell_{\mathbb{R}}^2$	155
$\sum_{1 \leq k \leq n} a_k$	55	$\bigoplus_{i \in I} V_i$	90	$\ell_{\mathbb{R}}^2$	155

$\ell_{\mathbb{C}}^2$	155	$E_A$	219
$N(v)$	156	$I_A$	220
$\ x\ _{\infty}$	157	$m_f$	222
$\ x\ _p$	157	$H_k(f, \lambda)$	223
$\ f\ _{\infty}$	157	$i(f, \lambda)$	223
$U^{\perp}$	164	$a \equiv_m b$	229
$P_U$	164	$\neg$	256
$A^*$	167	$\wedge$	256
$O(n)$	167	$\vee$	256
$U(n)$	167	$\rightarrow$	256
$f^* : W \rightarrow V$	170	$\leftrightarrow$	256
$f^{\circ} : W^* \rightarrow V^*$	170	$\forall$	259
$A_{\Phi, \mathcal{A}}$	172	$\exists$	259
$\det : K^{n \times n} \rightarrow K$	178	$\mathbb{N}$	261
$\operatorname{sgn} : S_n \rightarrow \{-1, 1\}$	180	$\mathbb{Z}$	261
$A_n$	180	$\mathbb{Q}$	261
$A_{ij}$	186	$\mathbb{R}$	261
$A^{\#}$	188	$\mathbb{C}$	261
$SL(n, K)$	190	$\sup(X)$	263
$E(n, K)$	190	$\inf(X)$	263
$O(n)$	190	$\operatorname{Re}(x, y)$	264
$SO(n)$	190	$\operatorname{Im}(x, y)$	264
$U(n)$	191	$i$	264
$SU(n)$	191	$\bar{z}$	264
$\operatorname{vol}_r(P)$	192	ZFC	267
$\operatorname{vol}_n(P)$	192		
$a \times b$	194		
$\det(f)$	198		
$\sigma(f)$	202		
$\operatorname{Eig}(f, \lambda)$	202		
$\sigma(A)$	202		
$\operatorname{Eig}(A, \lambda)$	202		
$A_f^{\mathcal{A}}$	205		
$p_A$	206		
$p_f$	206		
$K(X)$	206		
$\operatorname{spur}(A)$	207		
$E_A$	218		



---

# Index

---

## A

Abbildung 26, 260  
Abbildungseigenschaften 36  
abelsch 54  
abgebildet 30  
abgeschlossen 34  
Abgeschlossenheitsbedingung 34f  
Abschluss 34  
Abspalten von Nullstellen 68  
abzählbar 38  
abzählbar unendlich 38  
Additionstheoreme 129  
Additionstyp 134  
additive Funktionen 89  
additive Inverse 55  
adjungierte Homomorphismus 170  
adjungierte Matrix 167  
Adjunkte 188  
affine Basis 95  
affine Kombination 94  
affine Koordinatenvektor 95  
affiner Unterraum 94  
ähnlich 139, 204  
Algebra 119, 129  
algebraisch abgeschlossen 69  
Algebraische Grundstrukturen 4, 228  
algebraische Vielfachheit 69  
algebraischen Vielfachheit 208  
allgemeine lineare Gruppe 132  
Allgemeiner Darstellungssatz 111  
Allgemeiner Homomorphiesatz 104  
allgemeingültig 257  
Allquantor 259  
Alternation 176, 178, 195  
alternierende Gruppe 180  
Anfangswertproblem 252  
Angabe von Abbildungen 31  
angeordneter Körper 64, 263  
annimmt 30  
anordenbar, 64

Anordnungsaxiome 64  
Anteil 83  
Antikommutativität 195  
antisymmetrisch 20  
Anwendung 30  
äquivalent 22, 131, 238, 256  
Äquivalenz 22, 256  
Äquivalenzklasse 22  
Äquivalenzrelation 22  
Äquivalenzsatz für Normen 158  
Argument 30  
Assoziativgesetz 46, 228  
Assoziativität 17  
aufeinander senkrecht stehen 150  
Auflistung 15  
Ausräumen 233  
äußere Summe 90  
Aussonderung 15  
Aussonderungsschema 267  
Austauschlemma 84  
Austauschsatz 84  
Auswahlaxiom 40, 268  
Auswahlfunktion 40  
Auswahlmenge 40  
Automorphismengruppe 101  
Automorphismus 100  
Axiome für die Skalarmultiplikation 72  
Axiomschema 267

## B

bac-minus-cab-Regel 195  
Bahn 181  
baryzentrische Koordinatenvektor 95  
Basis 82  
Basisergänzungssatz 85, 88  
Basisexistenz 88  
Basisvektoren 82  
Basiswechsel 138  
Berechnung von Determinanten 179  
Berechnung von Koordinatenvektoren 133

Bestapproximation 165  
 Betrag 64  
 Bidualraum 121  
 bijektiv 36  
 Bild 20, 32, 102  
 Bild des Einheitswürfels 193  
 Bilinearform 172  
 Bilinearität 150, 154, 195  
 Bindungsstärke 257  
 Binomischer Lehrsatz 61  
 Blockstruktur 241  
 Boolesche Mengenoperationen 16  
 Bra-Vektor 169  
 Bruchnotation 62  
 Bruchrechnen 63

## C

Cauchy-Schwarz-Ungleichung 151, 153, 156, 173  
 Charakteristik 63  
 charakteristische Funktion 26  
 charakteristische Polynom 206  
 Cholesky-Zerlegung 196, 235

## D

dann und nur dann 256  
 darstellende Matrix 126, 130  
 Darstellung linearer Abbildungen 2, 110  
 Darstellungssatz 110f, 168  
 Dedekind-Peano-Axiome 261  
 definiert 30  
 Definitheit 156, 236  
 Definitionsbereich 20, 260  
 De-Morgan-Regeln 17, 257  
 Determinante 176, 198  
 Determinantenaxiomen 178  
 Determinantenfunktion 176, 178  
 diagonale Pivots 144  
 diagonalisierbar 204  
 Diagonalisierbarkeitskriterium 208  
 Diagonalisierung 233  
 Diagonalmatrix 125  
 Differenz 16  
 Differenzenregeln 17  
 Dimension 86, 95  
 Dimensionsformel 87  
 Dimensionsformel für lineare Abbildungen 116

Dirac-Notation 169  
 direkt 90  
 disjunkt 16  
 Disjunktion 256  
 Distributivgesetz 228  
 Distributivgesetze 60  
 Distributivität 17  
 Division 62  
 Divisionsbereich 62  
 Divisor 21  
 Doppelkegel 237  
 doppelte Verneinung 257  
 Dorfbarbier 14  
 Drehkästchen 240  
 Drehung 107ff, 127, 191  
 Dreiecksmatrix 125  
 Dreiecksungleichung 64, 153, 156  
 Drei-Finger-Regel 194  
 duale Abbildung 121  
 duale Basis 120  
 Dualitätssatz 141  
 Dualraum 120  
 Durchschnitt 16

## E

echte Obermenge 11  
 echte Teilmenge 11  
 Eigenraum 202  
 Eigenschaften des Kreuzprodukts 195  
 Eigenvektor 202  
 Eigenwert 202, 244f  
 Eigenwertkriterium 215, 236  
 eindeutige Nulldarstellung 80  
 Einermenge 15  
 eingeschlossene Winkel 151  
 Einheit 60  
 Einheitskreis 218  
 Einheitskugel 158  
 Einheitsmatrix 125  
 Einheitssphäre 219  
 Einschränkung 32  
 Einsetzen von Ringelementen 67  
 einstellige Operation 34  
 elementare lineare Gruppe 190  
 Elementarmatrix 134  
 Eliminationsverfahren 146  
 Ellipse 108, 218, 237, 265, 276  
 Ellipsoid 219

endlich 38  
 endlich erzeugt 86  
 endlich-dimensional 86  
 endlicher Körper 63  
 Endomorphismus 100  
 Entwicklungssatz von Laplace 3, 186  
 Epimorphismus 100  
 ergibt 30  
 Ersetzungsschema 268  
 erweiterte Koeffizientenmatrix 147  
 erzeugend 78  
 Erzeugendensystem 78  
 erzeugt 34  
 erzeugte Untergruppe 57  
 es gibt 259  
 euklidisch 154  
 euklidische Ebene 265  
 euklidische Norm 150, 152, 156  
 Existenz eines neutralen Elements 48  
 Existenz inverser Elemente 50  
 Existenz- und Eindeigkeitssatz 179  
 Existenz von Orthonormalbasen 162  
 Existenzquantor 259  
 Exponent Null 48  
 Exponential 250  
 Extensionalitätsaxiom 267  
 Extensionalitätsprinzip 11

## F

Faktorgruppe 58  
 Faktorisierung 22  
 Fallunterscheidung 258  
 Familie 30, 260  
 Faser 32, 112  
 Fehlstand 180  
 Folge 30  
 Folgennotation 30  
 Formel von Leibniz 182  
 Fortsetzungssatz 109  
 Fourier-Approximation 165  
 Frequenz 253  
 Fundamentalsatz der Algebra 69, 264  
 Fundierungsaxiom 268  
 Funktion 26  
 Funktional 120  
 für alle 259

## G

Gauß-Elimination 146  
 Gauß-Jordan-Elimination 146  
 genau dann, wenn 256  
 geometrische Bedeutung der Determinante 192  
 geometrische Reihe 61  
 geometrische Vielfachheit 202  
 geordnetes Paar 13, 15  
 gerade 180  
 Gershgorin-Kreis 246  
 gleiche Mächtigkeit 38  
 gleichmächtig 38  
 Gleichungssystem 113  
 Grad 66  
 Gradfunktion 66  
 Gradient 197  
 gramsche Determinante 192  
 gramsche Matrix 172, 192, 196  
 Gram-Schmidt-Orthonormalisierung 162  
 größte untere Schranke 263  
 Grundrechenarten 62  
 Gruppe 50  
 Gruppe der invertierbaren Elemente 51  
 Gruppenaxiome 51  
 Gruppenhomomorphismus 98

## H

Halbachsen 219  
 Halbgruppe 46  
 Hamel-Basis 89  
 harmonischer Oszillator 253  
 Hasse-Diagramm 24  
 Hauptachsentransformation 173, 214  
 Hauptminoren 196  
 Hauptminorenkriterium 196  
 Hauptraum 223  
 Hauptraumzerlegung 224  
 Hauptvektoren 223  
 hermitesch 171  
 hermitesche Form 172  
 Hermitizität 152, 154  
 Hesse-Matrix 197  
 Hilbert-Raum 169  
 Hintereinanderausführung 33  
 homogen 113  
 Homogenität 156  
 Homomorphiesatz 104  
 Homomorphismus 98

Hyperbel 237  
 Hyperebene 191

## I

Ideal 220  
 Idempotenz 165  
 Identität 26  
 imaginäre Einheit 264  
 Imaginärteil 264  
 Implikation 256  
 impliziert 256  
 indefinit 172  
 Index 223  
 Indexmenge 30  
 Indikatorfunktion 26  
 Induktionsaxiom 261  
 induzierte Matrixnorm 248  
 induzierte Norm 156  
 Infimum 24, 263  
 inhomogen 113  
 injektiv 36  
 Inklusion 11  
 innere Produkt 150, 152  
 innere Summe 90  
 inneres Produkt 154  
 Invarianz 223  
 invers 50  
 Inversennotation 50  
 Inversenregeln 52  
 invertierbar 60, 132  
 invertierbaren Elemente 51  
 Invertierung 233  
 Invertierung einer Matrix 135  
 Invertierungsregel 137  
 irrationale Zahlen 263  
 irreflexiv 20  
 isomorph 100  
 Isomorphiesatz 104, 114  
 Isomorphiesatz für Vektorräume 106  
 Isomorphismus 100

## J

Jacobi-Identität 195  
 Jordan-Block 225  
 Jordan-Kette 225  
 Jordan-Normalform 226, 251  
 Junktoren 5, 256

## K

kanonische Basis 83  
 kanonische innere Produkt 150, 152  
 kanonische Skalarprodukt 150, 152  
 kartesisches Produkt 32, 40  
 Kegelschnitt 237  
 Kern 102  
 Kerndarstellung 203  
 Kette 42  
 Kettenbedingung 42  
 Ket-Vektor 169  
 Klammern 47  
 Klasseneinteilung 16  
 Klassifikation endlicher Körper 63  
 Kleinsche Vierergruppe 53  
 kleinste obere Schranke 263  
 Koeffizient 66  
 Koeffizienten 113  
 Koeffizientenmatrix 147  
 kommutativ 54, 60  
 Kommutativgesetz 54, 228  
 Kommutativität 17  
 Komplement 16, 164  
 komplementär 16  
 komplementäre Matrix 188  
 Komplementierung 17  
 komplexen Zahlen 264  
 komponenten- 76  
 Komponentenweise Konvergenz 159  
 Komposition 32  
 Kompositionssatz für darstellende Matrizen 128  
 kongruent 215  
 kongruent modulo 21, 229  
 Kongruenz 4, 21, 23, 229  
 Konjugation 264  
 konjugierten Matrix 153  
 Konjunktion 256  
 konstante Abbildung 26  
 Konstruktion linearer Abbildungen 2, 108  
 Konstruktion von Cantor 263  
 Konstruktion von Dedekind 263  
 Konstruktion von Jordan-Ketten 225  
 Konstruktionssatz 108  
 Kontinuum 263  
 Kontrapositionsgesetz 257  
 Konvergenz 159  
 Koordinatenabbildung 83, 114



Koordinatenberechnung durch Invertierung 133  
 Koordinatenbestimmung 160  
 Koordinatenpicker 120  
 Koordinatenvektor 83, 95  
 Körper 62, 228  
 Kosinussatz 151  
 Kreislinie 265  
 Kreuzprodukt 15, 194  
 kritischer Punkt 197  
 Kronecker-Delta 125  
 Kronecker-Symbol 120  
 Kuratowski-Paar 13  
 Kürzungsregeln 52  
 K-Vektorraum 72

## L\_\_\_\_\_

Länge 82, 150, 264  
 Längensatz 85  
 Längentreue 166f  
 leere Menge 15  
 leere Produkt 48  
 Legendre-Polynome 161  
 Leibniz-Formel 182  
 Leitkoeffizient 66  
 lexikographische Ordnung 25  
 linear 106  
 linear abhängig 80  
 linear unabhängig 80  
 lineare Abbildung 106  
 lineare Funktionale 120  
 lineare Operatoren 118  
 lineare Ordnung 24  
 Lineare Systeme 4, 252  
 lineares Gleichungssystem 113  
 Linearfaktoren 68  
 Linearkombination 78  
 Linkseindeutigkeit 36  
 Linksnebenklasse 58  
 Linksshift 171  
 Linkstranslation 53  
 Lipschitz-stetig 159  
 Lösbarkeit von Gleichungen 52  
 Lösbarkeitskriterium 145  
 Lösen durch Invertierung 133  
 Lösen eines Gleichungssystems 147  
 Lösungsmenge 112f  
 Lösungsraum 145

LR-Zerlegung 233, 235, 241  
 Lücke 263

## M\_\_\_\_\_

Mächtigkeit 38  
 Mächtigkeitsvergleich 38  
 Manhattan-Norm 157  
 Matrixexponential 4, 250  
 Matrixnorm 248  
 Matrix-Vektor-Produkt 126  
 Matrixzerlegungen 4, 235  
 Matrizenmultiplikation 3, 128  
 Matrizenprodukt 128  
 Matrizenring 129  
 maximal 24  
 maximal linear unabhängig) 82  
 Maximalstelle 197  
 Maximum 24  
 Maximumsnorm 157  
 Mengenkompensation 14  
 Mengensystem 13, 16  
 minimal 24  
 minimal erzeugend 82  
 Minimalpolynom 220f  
 Minimalstelle 197  
 Minimum 24  
 Minus mal Minus 60  
 Modul 89  
 modulo 21f, 92  
 modulo  $m$  4, 229  
 Monoid 48  
 Monomorphismus 100  
 Multilinearität 176, 178  
 Multiplikationssatz 184  
 Multiplikationstyp 134

## N\_\_\_\_\_

nach 30  
 Nachfolgeraxiom 261  
 Nachfolgerfunktion 34, 261  
 Nachweis der Gruppenaxiome 51  
 natürliche Projektion 101, 104  
 $n$ -dimensional 86  
 Nebenklasse 58, 92  
 Negation 256  
 negativ 64  
 negativ (semi-)definit 172  
 neutrales Element 48

nicht 256  
 nichtnegativ 64  
 nilpotent 221  
 non 256  
 Norm 150, 152, 156  
 normal 213, 240  
 Normalform 4, 147, 225, 237  
 Normalformdarstellung 131, 215  
 Normalformproblem 205  
 Normalteiler 58  
 Normalteiler-Bedingung 58  
 normiert 66, 150, 152, 156  
 Normiertheit 176, 178  
 Normierung 156  
 n-stellig 34  
 Nullbedingung 80  
 Nullmatrix 125  
 Nullpolynom 66  
 Nullring 61  
 Nullstelle 67f  
 Nullstellen von  $p_A$  206  
 Nullstellenabspaltung 68  
 nullteilerfrei 60  
 Nullteilerfreiheit 60  
 Nullteilerfreiheit in Körpern 62  
 Nullvektor 72

## O\_\_\_\_\_

obere Dreiecksmatrix 125  
 obere Schranke 24  
 Obermenge 11  
 oder 256  
 Operation 34  
 Operator 118, 260  
 Ordnung 24  
 orthogonal 150, 152, 154, 164, 166  
 Orthogonalbasis 160  
 orthogonale Gruppe 190  
 orthogonale Komplement 164  
 orthogonale Projektion 164  
 orthogonale Summe 164  
 orthogonaler Homomorphismus 166  
 Orthogonalität 195  
 Orthonormalbasis 160  
 Orthonormalisierungsverfahren 162

## P\_\_\_\_\_

Paar 13

Paarmenge 15  
 Paarmengenaxiom 267  
 (paarweise) disjunkt 16  
 Parabel 237  
 Parallelepipet 193  
 Parallelogramm 193, 276  
 Parallelogramm-Gleichung 156  
 Parallelotop 192  
 Parseval-Gleichung 160  
 partielle Ordnung 24  
 Partition 16  
 Peano-Axiome 261  
 Permutation 28, 51  
 Permutation der Spalten 146  
 Permutationsgruppe 51  
 Permutationsmatrix 136  
 Pivots 144  
 p-Norm 157  
 Polarisation 157  
 Polarisations-Gleichungen 157  
 Polarzerlegung 235  
 Polynomdivision 68  
 Polynome 66  
 Polynomfunktion 67  
 Polynomring 66  
 positiv 64  
 positiv definit 172, 196, 215, 236  
 positiv semidefinit 172  
 positive Definitheit 150, 152, 154  
 Positivitätsregel 64  
 Potenz 47, 76  
 Potenzmenge 15  
 Potenzmengenaxiom 268  
 Potenzregeln 47  
 Prinzip vom ausgeschlossenen Dritten 257  
 Prinzip vom kleinsten Element 261  
 Produkt 47, 76  
 Produkte von Vektorräumen 2, 76  
 Produktregel 64  
 Projektion 99, 104, 164, 221  
 Pullback 121  
 Pünktchen-Notation 15  
 Punkt-Notation für Abbildungen 150  
 punktweise 76

## Q\_\_\_\_\_

QR-Zerlegung 162, 235  
 quadratisch 124, 215

quadratsummierbaren Folgen 155  
 Quadrik 215  
 Quadriken 4, 237  
 Quantoren 5, 259  
 Quantorenregeln 259  
 Quaternionen 63  
 Quotientenraum 92

## R

Rang 142  
 Rangformel 142  
 Rayleigh-Quotient 245  
 Realteil 264  
 Rechenregeln für den Grad 67  
 Rechenregeln in angeordneten Körpern 65  
 Rechenregeln in Körpern 63  
 Rechenregeln in Ringen 60f  
 rechte Seite 113  
 Rechte-Hand-Regel 194  
 Rechtseindeutigkeit 26  
 Rechtsnebenklasse 58  
 Rechtsschift 171  
 Rechtstranslation 53  
 reflexiv 20  
 Regel von Cramer 189  
 Regel von Sarrus 183  
 Regeln für  $\Delta$  17  
 Regularitätsaxiom 268  
 Relation 20  
 Repräsentant 22  
 Repräsentantensystem 22, 40  
 Rest 68  
 Restklassenkörper 63  
 Rieszschen Darstellungssatz 194  
 Rieszscher Darstellungssatz 168  
 Riesz-Vektor 168  
 Ring 60, 228  
 Rotation 191, 195  
 Rotationsspiegelung 191  
 Russell-Komprehension 14

## S

Satz des Pythagoras 150  
 Satz von Cantor 39  
 Satz von Cantor-Bernstein 39  
 Satz von Cayley-Hamilton 220  
 Satz von Erdős-Kaplansky 115  
 Satz von Lagrange 58

Satz von Wedderburn 63  
 Schiefkörper 62, 228  
 schiefsymmetrisch 253  
 schließlich gleich 0 66  
 Schranke 24, 263  
 Schubfach- 39  
 Schur-Komplement 241  
 Schur-Test 249  
 Schur-Zerlegung 210  
 Schwerpunkt 95  
 Seitenwechsel 236  
 selbstadjungiert 170  
 Seminorm 173  
 senkrecht 150, 152, 154  
 Sesquilinearform 172, 215  
 Sesquilinearität 152, 154  
 Signatur 214  
 Signum 180  
 Singleton 15  
 singulär 132  
 Singulärwerte 216, 218  
 Singulärwertzerlegung 216  
 Skalare 72  
 Skalarenkörper 72  
 Skalarmultiplikation 72  
 Skalarprodukt 72, 150, 152, 154  
 Skalierungseigenschaft 106  
 Spalten 124  
 Spaltenaxiome 185  
 Spaltenindex 124  
 Spaltenrang 142  
 Spaltensummennorm 249  
 Spann 78  
 Spannbedingung 80  
 Spektral-Norm 248  
 Spektralsatz 212, 245  
 Spektrum 202  
 spezielle lineare Gruppe 190  
 spezielle Lösung 113  
 spezielle orthogonale Gruppe 190  
 spezielle unitäre Gruppe 191  
 Spiegelung 191  
 Spur 207  
 Standardbasis 83, 125  
 Standardvektoren 79  
 stehen senkrecht aufeinander 152, 154  
 Stelle 30  
 Stellen 124

strikte partielle Ordnung 25  
 Struktur 46  
 Subadditivität 142  
 Subtraktion 55  
 Summe 55, 90  
 Summennorm 157  
 Supremum 24, 263  
 surjektiv 36  
 Symmetrie 150, 154  
 symmetrisch 20, 140  
 symmetrische Bilinearform 172  
 symmetrische Differenz 16  
 symmetrische Gruppe 51

## T\_\_\_\_\_

Taubenschlagprinzip 39  
 Tautologie 257  
 teilbar 229  
 Teilbarkeit 21  
 Teiler 21  
 Teilmenge 11  
 TermAuswertung 28  
 Termdefinitionen 28  
 tertium non datur 257  
 total 24  
 Träger 77, 90  
 Trägermenge 46  
 Trägheitssatz 214  
 transfinite Zahlen 43  
 Transformation 260  
 Transformationsformel 139, 205  
 Transformationsmatrix 138  
 transitiv 20  
 Transitivität 11  
 Translation 53  
 Translationsinvarianz 64  
 transponierte Matrix 140  
 Transposition 28, 140  
 Transpositionsmatrix 136  
 Transpositionssatz 184  
 Trigonalisierung 4, 210  
 trigonometrischen Polynome 79, 161  
 Tripel 13, 15  
 triviale Homomorphismus 99  
 Tupel 13  
 Typ 214

## U\_\_\_\_\_

überabzählbar 38  
 Überabzählbarkeit 263  
 Überdeckung 16  
 Überführung in Zeilenstufenform 146  
 Übergangsmatrix 138  
 Umkehrfunktion 36  
 Unbestimmte 66  
 und 256  
 unendlich 38  
 unendlich-dimensional 86  
 Unendlichkeitsaxiom 268  
 ungerade 180  
 unipotent 234  
 unitär 154, 166  
 unitäre Gruppe 191  
 untere Schranke 24  
 Untergruppe 56  
 Untergruppenkriterium 56  
 Unterraum 74  
 Unterraumkriterium 74  
 Untervektorraum 74  
 Urbild 32

## V\_\_\_\_\_

Vektoraddition 72  
 Vektoren 72  
 Vektorraum 72  
 Vektorraumaxiome 72  
 Vereinigung 16  
 Vereinigungsmengenaxiom 268  
 vergleichbar 24  
 Vergleichbarkeitssatz 39  
 Verknüpfung 32, 34  
 Verknüpfungstafel 48  
 Vervielfachung 55  
 Vielfaches 21  
 Vielfachheit 202, 208  
 vollen Rang 142  
 Vollständigkeitsaxiom 263  
 Volumenformeln 195  
 Volumenveränderung 193  
 Vorzeichen 64, 180  
 Vorzeichenfunktion 180

## W\_\_\_\_\_

Wahrheitstafel 257  
 Wahrheitswert 257  
 Wert 26, 30

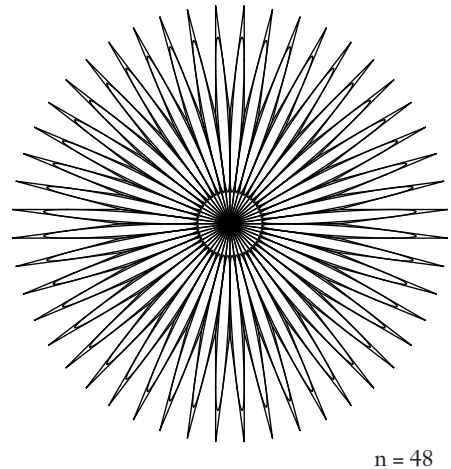
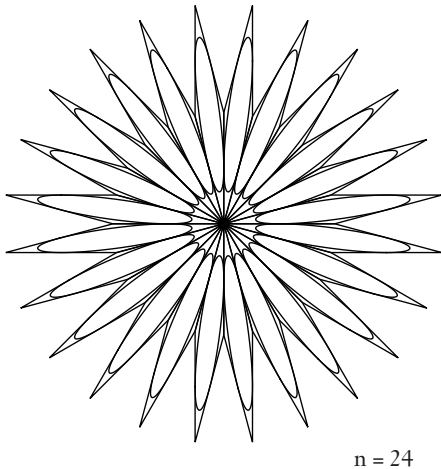
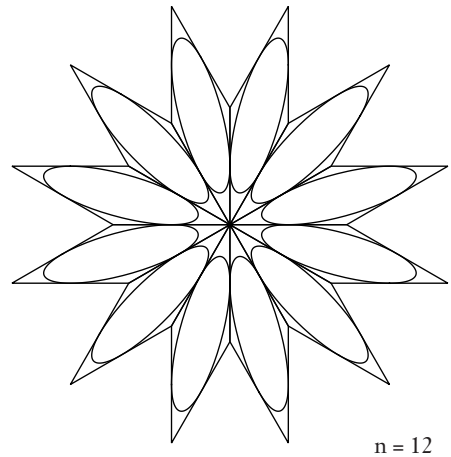
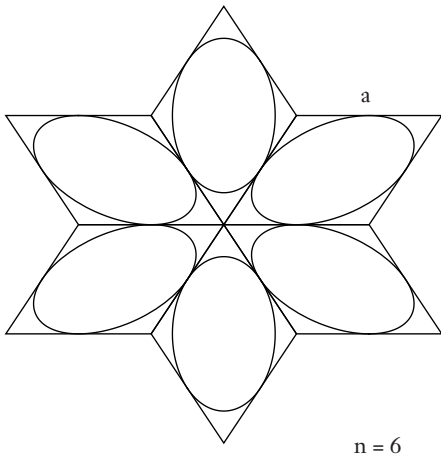
Wertebereich 20, 260  
Wertevorrat 30, 260  
Winkel 151, 195  
Winkeltreue 167  
wohldefiniert 59  
Wohlordnung 261

## **Z**\_\_\_\_\_

Zeilen 124  
Zeilenaxiome 185  
Zeilenindex 124  
Zeilenrang 143  
Zeilenrang gleich Spaltenrang 143, 147  
Zeilenstufenform 144  
Zeilensummennorm 249  
Zerlegung 16  
Zerlegung in Linearfaktoren 68  
Zerlegungen 235  
Zermelo-Fraenkel-Axiomatik 267  
Zielmenge 30, 260  
Zielvektor 113  
Zornsches Lemma 42  
zugeordnete lineare Abbildung 126  
Zuordnung 260  
zwischen 30  
zyklisch 57  
Zyklus 181



# Matrizensterne



## Konstruktionsanleitung

Man wähle ein  $a > 0$  (Zackenlänge),  $n \geq 3$  (Anzahl der Zacken) und setze

$$\alpha = \frac{2\pi}{n}, \quad D_\alpha = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}, \quad A = \begin{pmatrix} a & a \cos \alpha \\ 0 & a \sin \alpha \end{pmatrix},$$

$Q$  = „das Quadrat mit den Ecken  $0, (1, 0), (1, 1), (0, 1)$ “,

$C$  = „der in  $Q$  eingeschriebene Kreis mit Mittelpunkt  $(1/2, 1/2)$  und Radius  $1/2$ “.

Die Sterne entstehen durch Zeichnen der Bilder von  $Q$  (Parallelogramme) und  $C$  (Ellipse  $n$ ) unter den Abbildungen  $f_{A(k)} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ , wobei  $A(k) = D_{k\alpha} A$  für  $0 \leq k \leq n-1$ .