

**Diskrete Strukturen**  
und  
**Lineare Algebra für Informatiker**

Skript zur Vorlesung

Dr. Timo Hanke  
Prof. Dr. Gerhard Hiß  
Dr. Frank Lübeck (Korrekturen und Ergänzungen)  
Lehrstuhl D für Mathematik  
Lehrstuhl für Algebra und Zahlentheorie  
RWTH Aachen

Letzte Aktualisierung:  
20. August 2024

Unter der freundlichen Mithilfe von:  
Wolf-Daniel Andres, Grischa Studzinski und Florian Weingarten.

# Inhaltsverzeichnis

<b>Erster Teil: Grundlagen</b>	<b>2</b>
<b>1 Mathematische Grundbegriffe</b>	<b>5</b>
1.1 Aussagen . . . . .	5
1.2 Mengen . . . . .	11
1.3 Beweisprinzipien . . . . .	17
1.4 Abbildungen . . . . .	20
1.5 Relationen . . . . .	31
<b>2 Algebraische Strukturen</b>	<b>39</b>
2.1 Gruppen . . . . .	39
2.2 Ringe . . . . .	46
2.3 Polynome . . . . .	50
2.4 Teilbarkeitslehre in kommutativen Ringen . . . . .	53
2.5 Der Euklidische Algorithmus . . . . .	61
2.6 Restklassenringe . . . . .	65
2.7 Permutationen . . . . .	75
<b>3 Lineare Gleichungssysteme und Matrizen</b>	<b>83</b>
3.1 Matrizen . . . . .	83
3.2 Matrix-Arithmetik . . . . .	85
3.3 Lineare Gleichungssysteme . . . . .	90
3.4 Der Gauß-Algorithmus . . . . .	96
<b>Zweiter Teil: Diskrete Mathematik</b>	<b>107</b>
<b>Einleitung</b>	<b>111</b>
<b>4 Kombinatorik</b>	<b>113</b>
4.1 Permutationen und Kombinationen . . . . .	113
4.2 Binomialkoeffizienten . . . . .	117
4.3 Kombinatorische Beweisprinzipien . . . . .	121

4.4	Stirling'sche Zahlen . . . . .	125
<b>5</b>	<b>Graphentheorie</b>	<b>129</b>
5.1	Grundbegriffe . . . . .	129
5.2	Distanz und gewichtete Graphen . . . . .	135
5.3	Hamiltonkreise und Eulertouren . . . . .	141
5.4	Bäume . . . . .	145
	 <b>Dritter Teil: Lineare Algebra</b>	 <b>149</b>
	<b>Einleitung</b>	<b>153</b>
<b>6</b>	<b>Vektorräume und lineare Abbildungen</b>	<b>155</b>
6.1	Vektorräume . . . . .	155
6.2	Basis und Dimension . . . . .	158
6.3	Lineare Abbildungen . . . . .	176
6.4	Lineare Abbildungen und Matrizen . . . . .	187
6.5	Lineare Gleichungssysteme und Matrizen II . . . . .	208
6.6	Anwendung: Lineare Codes . . . . .	218
<b>7</b>	<b>Determinanten und Eigenvektoren</b>	<b>223</b>
7.1	Determinanten . . . . .	223
7.2	Eigenwerte und Eigenvektoren . . . . .	229
7.3	Der PageRank-Algorithmus . . . . .	238
7.4	Diagonalisierbarkeit und Trigonalisierbarkeit . . . . .	243
7.5	Der Satz von Cayley-Hamilton . . . . .	251
7.6	Ausblick: Normalformen . . . . .	259
<b>8</b>	<b>Euklidische und Unitäre Vektorräume</b>	<b>263</b>
8.1	Euklidische und unitäre Vektorräume . . . . .	263
8.2	Orthogonalität . . . . .	268
8.3	Positiv definite Matrizen . . . . .	275
8.4	Unitäre und orthogonale Abbildungen . . . . .	279
8.5	Der Spektralsatz . . . . .	286
8.6	Approximation . . . . .	293

# Grundlagen



# Kapitel 1

## Mathematische Grundbegriffe

### 1.1 Aussagen

#### 1.1.1 Definition und Beispiele

**Definition.** *Mathematische Aussagen* oder kurz *Aussagen* sind sprachliche Ausdrücke, die auch Formeln und Symbole enthalten können, und die einen eindeutigen *Wahrheitswert* besitzen, der entweder *wahr* oder *falsch* lautet.

**Beispiel.** Mathematische Aussagen sind:

- (i) ‘ $2 + 3 = 5$ ’ (wahr)
- (ii) ‘Alle Punkte auf einem Kreis haben den gleichen Abstand zum Mittelpunkt’ (wahr)
- (iii) ‘Jede gerade ganze Zahl größer als 2 ist Summe zweier Primzahlen’ (unbekannt)
- (iv) ‘Jede reelle Zahl ist ein Quadrat einer reellen Zahl’ (falsch)
- (v) ‘Es gibt eine ganze Zahl, deren Quadrat gleich ihrem Doppelten ist’ (wahr)

Die Aussage (iii) ist eine mathematische Aussage, denn sie besitzt einen Wahrheitswert, auch wenn uns dieser nicht bekannt ist. Die *Goldbach’sche Vermutung* besagt, dass der Wahrheitswert von (iii) wahr lautet. Keine mathematischen Aussagen sind dagegen ‘Aachen ist schön’ und ‘Die Mensapreise sind zu hoch’.

### 1.1.2 Zusammensetzung und Verneinung

**Definition a.** Für beliebige Aussagen  $A$  und  $B$  definieren wir die Wahrheitswerte für folgende *zusammengesetzte Aussagen*:

- (i) Die *Negation* (*Verneinung*)  $\neg A$  ist genau dann wahr, wenn  $A$  falsch ist.
- (ii) Die *Konjunktion* (*und-Verknüpfung*)  $A \wedge B$  ist genau dann wahr, wenn sowohl  $A$  als auch  $B$  wahr ist.
- (iii) Die *Disjunktion* (*oder-Verknüpfung*)  $A \vee B$  ist genau dann wahr, wenn  $A$  oder  $B$  wahr ist oder beide wahr sind.
- (iv) Das *exklusive oder*  $A \text{ xor } B$  ist genau dann wahr, wenn  $A$  oder  $B$  wahr ist, aber nicht beide wahr sind.
- (v) Die *Subjunktion* (*wenn-dann-Verknüpfung*)  $A \rightarrow B$  ist genau dann falsch, wenn  $A$  wahr ist und  $B$  falsch ist.
- (vi) Die *Bijunktion* (*genau-dann-Verknüpfung*)  $A \leftrightarrow B$  ist genau dann wahr, wenn  $A$  und  $B$  den gleichen Wahrheitswert besitzen.

**Sprechweise.** Zu  $\neg A$  sagt man „nicht  $A$ “, zu  $A \wedge B$  „ $A$  und  $B$ “, zu  $A \vee B$  „ $A$  oder  $B$ “, zu  $A \text{ xor } B$  „ $A$  x-or  $B$ “ oder „entweder  $A$  oder  $B$ “, zu  $A \rightarrow B$  „wenn  $A$  dann  $B$ “, zu  $A \leftrightarrow B$  „ $A$  gilt genau dann, wenn  $B$  gilt“.

**Wahrheitstafel.** Die Wahrheitswerte der eingeführten zusammengesetzten Aussagen können in folgender Tabelle zusammengefasst werden. Dies ist ein Beispiel für eine *Wahrheitstafel*. Wir schreiben 1 bzw. 0 für die Wahrheitswerte *wahr* bzw. *falsch*.

$A$	$B$	$\neg A$	$A \wedge B$	$A \vee B$	$A \text{ xor } B$	$A \rightarrow B$	$A \leftrightarrow B$
1	1	0	1	1	0	1	1
1	0	0	0	1	1	0	0
0	1	1	0	1	1	1	0
0	0	1	0	0	0	1	1

**Beispiel.**

- (i) Die Verneinung von ' $2+3=5$ ' lässt sich als 'Es gilt nicht, dass  $2+3=5$  ist' oder kürzer als ' $2+3$  ist ungleich 5' formulieren.



- (ii) Die Verneinung von ‘Das Glas ist voll’ lässt sich als ‘Das Glas ist nicht voll’ formulieren, nicht aber als ‘Das Glas ist leer’.
- (iii) Die Verneinung von ‘Alle Gläser sind voll’ lässt sich als ‘Nicht alle Gläser sind voll’, oder als ‘Es gibt ein Glas, das nicht voll ist’ formulieren.
- (iv) ‘Wenn  $2 + 3 = 6$ , dann ist  $2 + 3 = 7$ ’ ist wahr.

**Definition b.** Seien  $A$  und  $B$  Aussagen.

- (i) Ist  $A \rightarrow B$  wahr, dann schreiben wir  $A \Rightarrow B$  und sagen: “*Aus  $A$  folgt  $B$* ” oder “ *$A$  impliziert  $B$* ” oder “*Wenn  $A$ , dann  $B$* ” oder “ *$A$  ist hinreichend für  $B$* ” oder “ *$B$  ist notwendig für  $A$* ”.
- (ii) Ist  $A \leftrightarrow B$  wahr, dann schreiben wir  $A \Leftrightarrow B$  und sagen: “ *$A$  genau dann, wenn  $B$* ” oder “ *$A$  dann und nur dann, wenn  $B$* ” oder “ *$A$  ist notwendig und hinreichend für  $B$* ”.

### 1.1.3 Tautologien

**Definition.** (i) Ein *logischer Term* ist ein Ausdruck bestehend aus Variablen  $A, B, \dots$  und den Konstanten 1 und 0, die verknüpft sind mit den Symbolen  $\neg, \wedge, \vee, \text{ xor }, \rightarrow, \leftrightarrow$  (und Klammern). Durch Belegung der Variablen mit Wahrheitswerten bekommt der Term selbst einen Wahrheitswert.

- (ii) Zwei logische Terme  $S$  und  $T$ , definiert auf derselben Variablenmenge, heißen *logisch äquivalent* oder *wertverlaufsgleich*, geschrieben  $S \equiv T$ , wenn  $S$  und  $T$  denselben Wahrheitswert haben für jede Belegung der Variablen.

- (iii) Ein logischer Term  $T$  heißt *Tautologie*, wenn  $T \equiv w$ .

**Beispiel a.** Im Folgenden stellen wir einige einfache logischen Äquivalenzen zusammen.

- (i)
  - $A \wedge (B \wedge C) \equiv (A \wedge B) \wedge C$
  - $A \vee (B \vee C) \equiv (A \vee B) \vee C$
- (ii)
  - $A \wedge 1 \equiv A$
  - $A \vee 0 \equiv A$
- (iii)
  - $A \wedge B \equiv B \wedge A$

- $A \vee B \equiv B \vee A$
- (iv) •  $A \wedge A \equiv A$
- $A \vee A \equiv A$
- (v) •  $A \wedge \neg A \equiv 0$
- $A \vee \neg A \equiv 1$
- (vi) •  $A \wedge (B \vee C) \equiv (A \wedge B) \vee (A \wedge C)$
- $A \vee (B \wedge C) \equiv (A \vee B) \wedge (A \vee C)$
- (vii) •  $A \wedge (A \vee B) \equiv A$
- $A \vee (A \wedge B) \equiv A$

**Beispiel b.** Durch logische Äquivalenzen lassen sich logische Symbole durch andere ersetzen.

$$(i) \quad A \text{ xor } B \equiv (A \wedge \neg B) \vee (\neg A \wedge B).$$

Wir sagen daher, dass xor durch  $\neg, \wedge, \vee$  ausgedrückt werden kann.

$$(ii) \quad A \rightarrow B \equiv \neg(A \wedge \neg B).$$

$$(iii) \quad A \leftrightarrow B \equiv \neg(A \text{ xor } B).$$

*Übung a.* Man zeige, dass xor durch  $\neg, \vee$  ausgedrückt werden kann.

**Beispiel c.**  $A \wedge \neg B$  und  $A \rightarrow ((B \rightarrow \neg C) \vee D)$  sind logische Terme, aber keine Tautologien.  $(A \rightarrow B) \leftrightarrow \neg(A \wedge \neg B)$  ist eine Tautologie. Bedeutsame Tautologien sind:

(i) Modus Ponens:

$$(A \wedge (A \rightarrow B)) \rightarrow B$$

(ii) Tertium non datur (Gesetz des ausgeschlossenen Dritten):

$$A \vee \neg A$$

(iii) de Morgan-Gesetze:

$$\neg(A \wedge B) \leftrightarrow (\neg A \vee \neg B),$$

$$\neg(A \vee B) \leftrightarrow (\neg A \wedge \neg B)$$

(iv) Kontrapositionsgesetz:

$$(A \rightarrow B) \leftrightarrow (\neg B \rightarrow \neg A)$$

**Bemerkung a.** Es seien  $S, T$  logische Terme. Dann gilt  $S \equiv T$  genau dann,  $S \leftrightarrow T$  eine Tautologie ist. Insbesondere ist  $S \Leftrightarrow T$  für jede Belegung der Variablen in  $S$  und  $T$ .

Übung b.

- (i) Man schreibe die Tautologien auf, die von Beispiel **b** geliefert werden.
- (ii) Ist  $(A \leftrightarrow B) \leftrightarrow ((A \rightarrow B) \wedge (B \rightarrow A))$  eine Tautologie?
- (iii) Man folgere aus den Tautologien des Beispiels durch Einsetzen, dass auch  $\neg(A \wedge \neg A)$  eine Tautologie ist.
- (iv) Gelten die „Distributivgesetze“  $A \vee (B \wedge C) \equiv (A \vee B) \wedge (A \vee C)$  und  $A \wedge (B \vee C) \equiv (A \wedge B) \vee (A \wedge C)$ ?

**Bemerkung b.** Tautologien helfen bei Beweisen: Aus Modus Ponens folgt  $(A \wedge (A \rightarrow B)) \Rightarrow B$ ; zeigt man also, dass  $A$  wahr ist und  $A \Rightarrow B$  gilt (d.h. dass  $A \rightarrow B$  wahr ist), so folgt, dass auch  $B$  wahr ist.

Möchte man  $A \Rightarrow B$  zeigen, so kann man nach dem Kontrapositionsgesetz anstelle dessen auch  $\neg B \Rightarrow \neg A$  zeigen (z.B. statt ‘Wenn  $x$  kein Quadrat ist, dann  $x < 0$ ’ zeigt man ‘Wenn  $x \geq 0$ , dann  $x$  ein Quadrat’).

### 1.1.4 Aussageformen

**Definition.** Eine *Aussageform* ist ein sprachlicher Ausdruck, der Variablen enthält, und der für jede Belegung aller vorkommenden Variablen mit konkreten Objekten zu einer Aussage wird. (Diese letzte Bedingung führt dazu, dass die Auswahl der Objekte, mit denen die Variablen belegt werden können, i.A. eingeschränkt ist; siehe Beispiel (i) unten.)

**Bemerkung.** Eine Aussageform ist selbst keine Aussage. Die Zusammensetzung von Aussageformen mittels  $\neg, \wedge, \vee, \rightarrow, \leftrightarrow$ , etc. ist wieder eine Aussageform.

**Beispiel.**

- (i) ‘Wenn  $x > 0$ , dann ist  $x$  ein Quadrat.’ ist eine Aussageform. Wird die Variable  $x$  mit einer beliebigen reellen Zahl belegt, so erhalten wir eine Aussage (einen eindeutigen Wahrheitswert).

Hier setzen wir implizit voraus, dass die Variable  $x$  nur mit Objekten belegt wird, für die die Aussage  $x > 0$  Sinn ergibt (definiert ist).

- (ii) ‘Person  $x$  hat mindestens 50% der Klausur-Punkte erzielt’ ist eine Aussageform. Wird die Variable  $x$  mit einer beliebigen Person belegt, so erhalten wir eine Aussage (einen eindeutigen Wahrheitswert).
- (iii) Es sei  $A(x)$  die Aussageform ‘Person  $x$  hat in der Klausur volle Punktzahl erzielt’ und  $B(x)$  die Aussageform ‘Person  $x$  hat das Modul bestanden’. Dann ist auch  $A(x) \rightarrow B(x)$  eine Aussageform.

Für jede Belegung der Variable  $x$  mit einer Person ist  $A(x) \rightarrow B(x)$  eine wahre Aussage. Es gilt also  $A(x) \Rightarrow B(x)$ . Das liegt daran, dass der Fall  $A(x)$  wahr und  $B(x)$  falsch (der einzige Fall in dem  $A(x) \rightarrow B(x)$  falsch ist) nicht vorkommt.

- (iv) Es sei  $A(t)$  die Aussageform ‘Der Projektor im Hörsaal ist zum Zeitpunkt  $t$  aus’ und  $B(t)$  die Aussageform ‘Der Hörsaal ist zum Zeitpunkt  $t$  leer’.

Für jede Belegung der Variable  $t$  mit einem Zeitpunkt ist  $A(t) \rightarrow B(t)$  eine Aussage. Deren Wahrheitswert hängt allerdings von  $t$  ab. Wann ist sie falsch?

**Bemerkung.** Wenn  $A(x) \rightarrow B(x)$  unabhängig von  $x$  stets wahr ist (wie in Beispiel (iii)), gilt also  $A(x) \Rightarrow B(x)$ , dann drückt dies offensichtlich einen kausalen Zusammenhang aus.

### 1.1.5 Sprachliche Konventionen

Wir einigen uns auf folgende Konventionen

- (i) Wir sagen: „Die Aussage  $A$  *gilt*“, falls  $A$  den Wahrheitswert 1 hat (also wahr ist).
- (ii)  $A := B$  bedeutet: Das Symbol  $A$  wird durch das Symbol  $B$  definiert.
- (iii)  $A :\Leftrightarrow B$  bedeutet: Die Aussage  $A$  wird durch die Aussage  $B$  definiert ( $A$  hat per Definition den gleichen Wahrheitswert wie  $B$ ).
- (iv) *Ein* bedeutet stets „mindestens ein“ und ist von „genau ein“ zu unterscheiden.
- (v) In einer Aufzählung von Objekten  $x_1, \dots, x_n$  heißen  $x_1, \dots, x_n$  *paarweise verschieden*, wenn keine zwei Objekte der Aufzählung gleich sind (d.h. wenn in der Aufzählung keine Wiederholungen vorkommen). Davon zu unterscheiden ist „verschieden“ im Sinne von „nicht alle gleich“. Wenn wir von „ $n$  verschiedenen Objekten  $x_1, \dots, x_n$ “ sprechen, impliziert das, dass  $x_1, \dots, x_n$  paarweise verschieden sind.

## 1.2 Mengen

### 1.2.1 Definition und Beispiele

“Unter einer *Menge* verstehen wir jede Zusammenfassung  $M$  von bestimmten wohlunterscheidbaren Objekten unserer Anschauung oder unseres Denkens [welche die *Elemente* von  $M$  genannt werden] zu einem Ganzen.”

Georg Cantor, 1895

Bei der Auslegung von Cantor’s Begriff einer „Zusammenfassung“ ist allerdings Vorsicht geboten. Das wusste schon Cantor selbst und zeigte, dass die Betrachtung der „Menge aller Mengen“ zu einem Widerspruch führt: nach der *Zweiten Cantor’schen Antinomie* wäre sie „größer“ als sie selbst. Man kann auch ohne Betrachtung der „Größe“ einer Menge einen rein logischen Widerspruch aus der „Menge aller Mengen“ ableiten, die *Russel’sche Antinomie* (siehe Übung **b** unten). Wir einigen uns auf die folgende Definition des Mengenbegriffs.

**Definition a.** Eine *Menge*  $M$  ist etwas, zu dem jedes beliebige Objekt  $x$  entweder *Element* der Menge ist, geschrieben  $x \in M$ , oder nicht, geschrieben  $x \notin M$ .

Mengen sind also gerade dadurch gekennzeichnet, dass ‘ $x \in M$ ’ für jedes Objekt  $x$  eine Aussage ist (einen eindeutigen Wahrheitswert hat), also gerade dadurch, dass ‘ $x \in M$ ’ eine Aussageform ist. Umgekehrt ist für jede Aussageform  $A(x)$  die Zusammenfassung aller  $x$ , für die  $A(x)$  wahr ist, eine Menge (vgl. Schreibweise (iii) unten).

**Bemerkung a.** Mengen, die sich selbst enthalten führen nicht per se zu einem Widerspruch. In der weitverbreitetsten Mengenlehre (der *Zermelo-Fraenkel-Mengenlehre*), der wir uns anschließen wollen, sind Mengen, die sich selbst als Elemente enthalten, allerdings nicht erlaubt.

**Definition b.** Sind  $M, N$  zwei Mengen, so heißt  $N$  eine *Teilmenge* von  $M$  und  $M$  eine *Obermenge* von  $N$ , geschrieben  $N \subseteq M$ , wenn für alle  $x \in N$  gilt:  $x \in M$ . Das Zeichen  $\subseteq$  bzw. die Aussage  $N \subseteq M$  heißt *Inklusion*.

Zwei Mengen  $M$  und  $N$  heißen *gleich*, geschrieben  $M = N$ , wenn  $M \subseteq N$  und  $N \subseteq M$ .

Eine Menge  $M$  heißt *endlich*, wenn  $M$  nur endlich viele Elemente besitzt. Man schreibt in diesem Fall  $|M|$  für die Anzahl der Elemente von  $M$ . Anderenfalls heißt  $M$  *unendlich* und man schreibt  $|M| = \infty$ .

**Schreibweise.** Es folgen die gebräuchlichsten Methoden, Mengen zu beschreiben.

- (i) *Aufzählen.* Die Elemente werden aufgelistet und mit Mengenklammern eingeschlossen. Reihenfolge und Wiederholungen spielen bei der Mengenaufzählung keine Rolle, z.B.

$$\{1, 3, 17\} = \{3, 1, 17\} = \{1, 3, 17, 1, 3\}.$$

- (ii) *Beschreiben.* Mengen können durch Worte beschrieben werden, etwa:

$$\text{Menge der natürlichen Zahlen} = \{1, 2, 3, 4, 5, \dots\}$$

$$\text{Menge der ganzen Zahlen} = \{\dots, -2, -1, 0, 1, 2, \dots\}$$

- (iii) *Aussondern.* Es sei  $M$  eine Menge. Ist  $A(x)$  eine Aussageform, so bezeichnet

$$\{x \in M \mid A(x)\}$$

diejenige Teilmenge von  $M$ , die aus allen Elementen besteht, für die  $A(x)$  wahr ist (gesprochen „Menge aller  $x$  aus  $M$  mit  $A(x)$ “). Benennen wir beispielsweise die Menge der natürlichen Zahlen mit  $\mathbb{N}$ , so ist  $\{n \in \mathbb{N} \mid n \text{ ist ungerade}\}$  die Menge der ungeraden natürlichen Zahlen, also  $\{1, 3, 5, 7, \dots\}$ .

- (iv) *Abbilden.* Seien  $M$  und  $N$  Mengen und  $f(x)$  für jedes  $x \in M$  ein Element aus  $N$ . (Wir greifen hier dem Begriff der *Abbildung* vor.) Dann ist

$$\{f(x) \mid x \in M\}$$

eine Teilmenge von  $N$  (insbesondere eine Menge), die Menge aller Elemente der Form  $f(x)$  von  $N$ , wobei  $x$  alle Elemente aus  $M$  durchläuft. Ist z.B.  $\mathbb{N}$  die Menge der natürlichen Zahlen, dann ist  $\{n^2 \mid n \in \mathbb{N}\}$  die Menge der Quadratzahlen (hier ist  $M = N = \mathbb{N}$ ). Ist  $\mathbb{R}$  die Menge der reellen Zahlen, dann ist  $\{|x| \mid x \in \mathbb{R}\}$  die Menge der nicht-negativen reellen Zahlen. Abbilden und Aussondern können kombiniert werden, sodass z.B.  $\{n^2 \mid n \in \mathbb{N}, n \text{ ungerade}\}$  die Menge aller Quadrate von ungeraden natürlichen Zahlen bezeichnet, also  $\{1, 9, 25, 49, \dots\}$ .

**Bemerkung b.** In der Regel schreibt man die Menge

$$\{n^2 \mid n \in \{m \in \mathbb{N} \mid m \text{ ungerade}\}\}$$

auch kurz und intuitiv als  $\{n^2 \mid n \in \mathbb{N} \text{ ungerade}\}$ , ohne sich Gedanken über Abbilden und Aussondern zu machen. Man muss beide Schreibweisen aber penibel trennen, wenn man die Menge beispielsweise in ein Computeralgebra-System eingeben möchte.

**Beispiel.** Häufig auftretende Mengen sind:

Symbol	Beschreibung	Definition
$\emptyset$	leere Menge	$\{\}$
$\mathbb{N}$	natürliche Zahlen	$\{1, 2, 3, \dots\}$
$\mathbb{N}_0$	natürliche Zahlen einschl. 0	$\{0, 1, 2, 3, \dots\}$
$\underline{n}$	$n$ -elementige Menge, $n \in \mathbb{N}_0$	$\{1, 2, \dots, n\}$ , $\underline{0} := \emptyset$
$\mathbb{P}$	Primzahlen	$\{2, 3, 5, 7, 11, 13, \dots\}$
$\mathbb{Z}$	ganze Zahlen	$\{\dots, -2, -1, 0, 1, 2, \dots\}$
$\mathbb{Q}$	rationale Zahlen	$\{\frac{a}{b} : a \in \mathbb{Z}, b \in \mathbb{N}\}$
$\mathbb{R}$	reelle Zahlen	$\{a_1 a_2 \dots a_r, b_1 b_2 \dots : a_i, b_i \in \{0, 1, \dots, 9\}\}$
$\mathbb{R}_{>0}$	positive reelle Zahlen	$\{x \in \mathbb{R} \mid x > 0\}$
$\mathbb{R}_{\geq 0}$	nicht-negative reelle Zahlen	$\{x \in \mathbb{R} \mid x \geq 0\}$
$\mathbb{C}$	komplexe Zahlen	$\{a + bi : a, b \in \mathbb{R}\}$

Nur die erste und vierte der Mengen der Tabelle sind endlich, nämlich  $|\emptyset| = 0$  und  $|\underline{n}| = n$  für alle  $n \in \mathbb{N}_0$ . Es gilt:

$$\emptyset = \underline{0} \subseteq \underline{1} \subseteq \underline{2} \subseteq \dots \subseteq \mathbb{N} \subseteq \mathbb{N}_0 \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}.$$

*Übung a.* Was gilt für eine Menge  $M$ :

- (i)  $x \in M$  xor  $x \notin M$  für alle  $x$ ?
- (ii)  $x \in M \Leftrightarrow \neg(x \notin M)$ ?
- (iii)  $\neg(x \in M) \Leftrightarrow x \notin M$ ?

*Übung b* (Russel's Antinomie). Die „Menge aller Mengen“ würde als Teilmenge enthalten die „Menge“  $\mathcal{M}$  aller Mengen, die sich nicht selbst als Element enthalten. Ist dann  $\mathcal{M} \in \mathcal{M}$  oder  $\mathcal{M} \notin \mathcal{M}$ ?

### 1.2.2 Quantifizierte Aussagen

Es sei  $A(x)$  eine Aussageform. Nach Definition 1.1.4 ist  $A(x)$  für jedes  $x$  eine Aussage. Setzt man in  $A(x)$  für  $x$  in ein konkretes Objekt ein, so sagt man,  $x$  wird *spezifiziert*. Zwei weitere Möglichkeiten, aus  $A(x)$  eine Aussage zu machen, bestehen darin,  $x$  zu *quantifizieren*:

‘Für alle  $x \in M$  gilt  $A(x)$ ’ und ‘Es gibt ein  $x \in M$ , für das  $A(x)$  gilt’.

Hierbei ist  $M$  eine Menge. Diese sprachlichen Ausdrücke sind Aussagen, denn  $x$  ist keine (freie) Variable mehr!

**Beispiel.**

- (i) Sei  $A(x)$  die Aussageform ' $x > 5$ '. Dann ist 'Es existiert ein  $x \in \mathbb{N}$  mit  $A(x)$ ' wahr, weil z.B.  $A(7)$  wahr ist. Dagegen ist 'Für alle  $x \in \mathbb{N}$  gilt  $A(x)$ ' falsch, weil z.B.  $A(2)$  falsch ist.
- (ii) Sei  $A(t)$  die Aussageform 'Zum Zeitpunkt  $t$  gilt: Projektor ist aus  $\rightarrow$  Hörsaal ist leer'. Ist  $t$  ein konkreter Zeitpunkt, an dem der Projektor an ist oder der Hörsaal leer, so ist die Aussage  $A(t)$  wahr. Da es solche Zeitpunkte gibt, ist 'Es gibt eine Zeit  $t$  mit  $A(t)$ ' wahr. Ist  $t$  dagegen ein konkreter Zeitpunkt, an dem der Projektor aus ist und der Hörsaal nicht leer, so ist die Aussage  $A(t)$  falsch. Da es auch solche Zeitpunkte gibt, ist auch 'Es gibt eine Zeit  $t$  mit  $\neg A(t)$ ' wahr und 'Für alle Zeiten  $t$  gilt  $A(t)$ ' falsch.
- (iii) Die Verneinung von 'Für alle  $x \in M$  gilt  $A(x)$ ' lässt sich als 'Es existiert  $x \in M$  mit  $\neg A(x)$ ' bzw. 'Es existiert  $x \in M$  für das  $A(x)$  nicht gilt' formulieren. Die Verneinung von 'Für alle  $x \in \mathbb{R}$  gilt  $x^2 > 0$ ' lässt sich als 'Es existiert ein  $x \in \mathbb{R}$  mit  $x^2 \leq 0$ ' formulieren.
- (iv) Die Verneinung von 'Es existiert ein  $x \in M$  mit  $A(x)$ ' lässt sich als 'Für alle  $x \in M$  gilt  $\neg A(x)$ ' formulieren. Die Verneinung von 'Es gibt eine Person im Hörsaal, die ihr Handy aus hat' lässt sich als 'Alle Personen im Hörsaal haben ihr Handy an' formulieren.

**Bemerkung.** Gelegentlich schreibt man (missbräuchlich) nur eine Aussageform  $A(x)$  auf, meint damit aber die Aussage 'Für alle  $x \in M$  gilt  $A(x)$ '. Das geht nur, wenn die Menge  $M$  aus dem Zusammenhang klar ist.

*Übung.* Wie lautet der Wahrheitswert der Aussagen 'Für alle  $x \in \emptyset$  gilt  $A(x)$ ' und 'Es gibt  $x \in \emptyset$  mit  $A(x)$ '?

**1.2.3 Konstruktion von Mengen**

**Definition** (Mengenoperationen). Es seien  $M, N$  beliebige Mengen.

- (i)  $M \cap N := \{x \in M \mid x \in N\}$  heißt *Durchschnitt* von  $M$  und  $N$ .
- (ii)  $M \cup N := \{x \mid x \in M \text{ oder } x \in N\}$  heißt *Vereinigung* von  $M$  und  $N$ .
- (iii)  $M \setminus N := \{x \in M \mid x \notin N\}$  heißt die *Differenzmenge*, gesprochen „ $M$  ohne  $N$ “.



- (iv)  $M \times N := \{(x, y) \mid x \in M \text{ und } y \in N\}$  heißt *kartesisches Produkt* von  $M$  und  $N$ .

Hierbei ist  $(x, y)$  ein *geordnetes Paar*. Zwei geordnete Paare  $(x, y)$  und  $(x', y')$  sind genau dann gleich, wenn  $x = x'$  und  $y = y'$ .

- (v)  $\text{Pot}(M) := \{S \mid S \subseteq M\}$  heißt die *Potenzmenge* von  $M$ .

**Beispiel.**

- (i) Die leere Menge ist Teilmenge jeder beliebigen Menge (auch von sich selbst).
- (ii) Es gilt:

$$\begin{aligned}\text{Pot}(\emptyset) &= \{\emptyset\}, \\ \text{Pot}(\{1\}) &= \{\emptyset, \{1\}\}, \\ \text{Pot}(\{1, 2\}) &= \{\emptyset, \{1\}, \{2\}, \{1, 2\}\}, \\ &\vdots\end{aligned}$$

- (iii) Für Mengen  $M$  und  $N$  gilt:

- $M \cap N = N \Leftrightarrow N \subseteq M$ .
- $M \cup N = N \Leftrightarrow M \subseteq N$ .

**Bemerkung.** Für Mengen  $L, M, N$  gelten folgende Rechenregeln.

- (i)    •  $L \cap (M \cap N) = (L \cap M) \cap N$   
       •  $L \cup (M \cup N) = (L \cup M) \cup N$
- (ii)    •  $L \cap M = M \cap L$   
       •  $L \cup M = M \cup L$
- (iii)    •  $L \cap L = L$   
       •  $L \cup L = L$
- (iv)    •  $L \cap (M \cup N) = (L \cap M) \cup (L \cap N)$   
       •  $L \cup (M \cap N) = (L \cup M) \cap (L \cup N)$
- (v)    •  $L \cap (L \cup M) = L$   
       •  $L \cup (L \cap M) = L$

*Übung.*

- (i) Wie viele Elemente hat  $\text{Pot}(\underline{n})$  für  $n \in \mathbb{N}_0$ ?

### 1.2.4 Indexmengen

**Definition a.** Es sei  $n \in \mathbb{N}$ . Für Zahlen  $a_1, \dots, a_n$ , Mengen  $M_1, \dots, M_n$  und Aussagen  $A_1, \dots, A_n$  definieren wir:

- (i)  $\sum_{i=1}^n a_i := a_1 + \dots + a_n$
- (ii)  $\prod_{i=1}^n a_i := a_1 \cdot \dots \cdot a_n$
- (iii)  $\bigcup_{i=1}^n M_i := M_1 \cup \dots \cup M_n$
- (iv)  $\bigcap_{i=1}^n M_i := M_1 \cap \dots \cap M_n$
- (v)  $\bigvee_{i=1}^n A_i := A_1 \vee \dots \vee A_n$
- (vi)  $\bigwedge_{i=1}^n A_i := A_1 \wedge \dots \wedge A_n$

Diese *Aufzähl Schreibweisen* können teilweise auf beliebige *Indexmengen*  $I$  verallgemeinert werden, die auch unendlich sein dürfen:

**Definition b.** Für jedes  $i \in I$  sei  $M_i$  eine Menge.

- (i) Wir definieren  $\bigcup_{i \in I} M_i$  durch

$$x \in \bigcup_{i \in I} M_i :\Leftrightarrow \text{es gibt } i \in I \text{ mit } x \in M_i$$

- (ii) Wir definieren  $\bigcap_{i \in I} M_i$  durch

$$x \in \bigcap_{i \in I} M_i :\Leftrightarrow \text{für alle } i \in I \text{ gilt } x \in M_i$$

Es ist auch sinnvoll, den Begriff „paarweise verschieden“ für beliebig indizierte Objekte auszudehnen:

**Definition c.** Für jedes  $i \in I$  sei  $x_i$  ein Objekt. Die Objekte  $x_i, i \in I$ , heißen *paarweise verschieden*, wenn für alle  $i, j \in I$  gilt:  $x_i = x_j \Rightarrow i = j$ .

**Beispiel.** (i) Die Zahlen  $n^2, n \in \mathbb{N}$ , sind paarweise verschieden.

- (ii) Die Zahlen  $n^2, n \in \mathbb{Z}$ , sind nicht paarweise verschieden.

### 1.2.5 Mengenpartitionen

**Definition.**

- (i) Zwei Mengen  $A, B$  heißen *disjunkt*, wenn  $A \cap B = \emptyset$ .
- (ii) Mengen  $M_i, i \in I$ , heißen *paarweise disjunkt*, wenn für alle  $i, j \in I$  mit  $i \neq j$  gilt:  $M_i \cap M_j = \emptyset$ .
- (iii) Es sei  $\mathcal{M}$  eine Menge von Mengen ( $\mathcal{M}$  darf hier unendlich sein). Die Elemente von  $\mathcal{M}$  heißen *paarweise disjunkt*, wenn je zwei davon disjunkt sind, d.h. wenn für alle  $M, M' \in \mathcal{M}$  mit  $M \neq M'$  gilt:  $M \cap M' = \emptyset$ .
- (iv) Es sei  $M$  eine Menge. Eine *Partition* von  $M$  ist eine Menge  $\mathcal{P}$  nicht-leerer, paarweise disjunkter Teilmengen von  $M$  mit  $M = \bigcup_{C \in \mathcal{P}} C$ . Die Elemente  $C \in \mathcal{P}$  heißen *Teile* der Partition.

**Bemerkung.** Für jede Partition  $\mathcal{P}$  von  $M$  ist  $\mathcal{P} \subseteq \text{Pot}(M) \setminus \{\emptyset\}$ .

**Beispiel.**

- (i)  $\mathcal{P} = \{\{n \in \mathbb{N} \mid n \text{ gerade}\}, \{n \in \mathbb{N} \mid n \text{ ungerade}\}\}$  stellt eine Partition von  $\mathbb{N}$  mit zwei Teilen dar.
- (ii)  $\mathcal{P} = \{\{n \in \mathbb{N} \mid n \text{ hat } k \text{ Dezimalstellen}\} \mid k \in \mathbb{N}\}$  stellt eine Partition von  $\mathbb{N}$  mit unendlich vielen Teilen dar.
- (iii) Die einzige Partition von  $\emptyset$  ist  $\mathcal{P} = \emptyset$ .

*Übung.* Man mache sich klar:

- (i) Sind  $M, N$  endliche, disjunkte Mengen, so gilt  $|M \cup N| = |M| + |N|$ .
- (ii) Sind  $M_1, \dots, M_n$  endliche, paarweise disjunkte Mengen, so gilt

$$|\bigcup_{i=1}^n M_i| = \sum_{i=1}^n |M_i|.$$

## 1.3 Beweisprinzipien

### 1.3.1 Direkter Beweis

**Prinzip.** Ziel:  $A \Rightarrow B$  (d.h.  $A \rightarrow B$  ist wahr).

Um das Ziel zu zeigen, nehmen wir an, dass  $A$  wahr ist und folgern daraus mittels logischer Schlüsse, dass  $B$  wahr ist. Wenn das gelungen ist, ist  $A \Rightarrow B$  bewiesen.

**Beispiel.** Für alle  $n \in \mathbb{N}$  gilt:  $n$  ungerade  $\Rightarrow n^2$  ungerade.

*Beweis.* Sei  $n \in \mathbb{N}$  beliebig, sei  $A$  die Aussage ‘ $n$  ist ungerade’ und  $B$  die Aussage ‘ $n^2$  ist ungerade’. Wir nehmen an,  $A$  ist wahr, d.h.  $n$  ist ungerade. Wir folgern, dass  $B$  wahr ist: Da  $n$  ungerade ist, existiert ein  $k \in \mathbb{N}$  mit  $n = 2k - 1$ . Dann ist  $n^2 = (2k - 1)^2 = 4k^2 - 4k + 1 = 2(2k^2 - 2k) + 1$ , eine ungerade Zahl. Damit ist gefolgert, dass  $B$  wahr ist. Nach dem Beweisprinzip des direkten Beweises ist also  $A \Rightarrow B$  wahr. Da  $n \in \mathbb{N}$  beliebig gewählt war, gilt dies für alle  $n \in \mathbb{N}$ .  $\square$

*Übung.* Was passiert, wenn sich aus  $A$  ein Widerspruch folgern lässt,  $A$  also falsch ist?

### 1.3.2 Beweis durch Kontraposition

**Prinzip.** Ziel:  $A \Rightarrow B$ .

Stattdessen zeigen wir,  $\neg B \Rightarrow \neg A$ . Wenn das gelungen ist, ist  $A \Rightarrow B$  bewiesen.

*Beweis des Prinzips.* Dieses Prinzip beruht auf der bekannten Tautologie  $(A \rightarrow B) \Leftrightarrow (\neg B \rightarrow \neg A)$  aus Beispiel 1.1.3.  $\square$

**Beispiel.** Für alle  $n \in \mathbb{N}$  gilt:  $n^2$  gerade  $\Rightarrow n$  gerade.

*Beweis.* Sei  $n \in \mathbb{N}$  beliebig, sei  $A$  die Aussage ‘ $n^2$  ist gerade’ und  $B$  die Aussage ‘ $n$  ist gerade’. Wir zeigen  $\neg B \Rightarrow \neg A$ : Dies ist gleichbedeutend mit ‘ $n$  ist ungerade  $\Rightarrow n^2$  ist ungerade’ und wurde schon in Beispiel 1.3.1 gezeigt. Damit gilt nach dem Beweisprinzip der Kontraposition auch  $A \Rightarrow B$ . Da  $n \in \mathbb{N}$  beliebig gewählt war, gilt dies für alle  $n \in \mathbb{N}$ .  $\square$

### 1.3.3 Beweis durch Widerspruch

**Prinzip.** Ziel:  $A$  ist wahr.

Wir zeigen, dass  $\neg A \Rightarrow (B \wedge \neg B)$  gilt. Wenn das gelungen ist, ist auch  $A$  wahr. ( $B \wedge \neg B$  ist hier der Widerspruch und die Aussage  $B$  kann frei gewählt werden.)

*Beweis des Prinzips.*  $B \wedge \neg B$  ist stets falsch (vgl. Übung 1.1.3 b). Wenn  $\neg A \Rightarrow (B \wedge \neg B)$  gilt, ist also  $\neg A \rightarrow (B \wedge \neg B)$  wahr. Das kann nur der Fall sein, wenn  $\neg A$  falsch ist (vgl. Definition von  $\rightarrow$ ), d.h.  $A$  wahr ist.  $\square$

**Beispiel.** Es sei  $A$  die Aussage  $\sqrt{2} \notin \mathbb{Q}$ .

*Beweis.* Wir nehmen an,  $\neg A$  ist wahr, d.h.  $\sqrt{2} \in \mathbb{Q}$ . Dann gibt es  $n, m \in \mathbb{N}$ , die nicht beide gerade sind und  $\sqrt{2} = m/n$  erfüllen ( $\sqrt{2}$  wird als Bruch geschrieben und dieser gekürzt). Seien solche  $n, m$  gewählt. Durch Quadrieren folgt  $2n^2 = m^2$ , d.h.  $m^2$  ist gerade. Also ist  $m$  gerade nach Beispiel 1.3.2. Sei  $k \in \mathbb{N}$  mit  $m = 2k$ . Dann gilt  $2n^2 = m^2 = 4k^2$ , also  $n^2 = 2k^2$ , d.h.  $n^2$  ist gerade. Also ist  $n$  gerade nach Beispiel 1.3.2. Insgesamt wurde gezeigt, dass sowohl  $n$  als auch  $m$  gerade sind. Das ist ein Widerspruch (die Aussage  $B$  kann hier ‘ $n$  und  $m$  sind nicht beide gerade’ gewählt werden). Also ist die Annahme  $\sqrt{2} \in \mathbb{Q}$  falsch, und damit ist die Behauptung  $\sqrt{2} \notin \mathbb{Q}$  wahr.  $\square$

### 1.3.4 Vollständige Induktion

**Prinzip.** Ziel: Für alle  $n \in \mathbb{N}$  gilt  $A(n)$ .

Wir zeigen als *Induktionsanfang*, dass  $A(1)$  wahr ist, und als *Induktionsschritt* die Implikation  $A(n) \Rightarrow A(n+1)$  für alle  $n \in \mathbb{N}$ . Dann ist  $A(n)$  für alle  $n \in \mathbb{N}$  wahr. Man spricht präziser von einer vollständigen Induktion *über  $n$* . Im Induktionsschritt nennt man die Aussage  $A(n)$  die *Induktionsvoraussetzung*.

*Beweis des Prinzips.* Das Prinzip beruht auf der folgenden Eigenschaft von  $\mathbb{N}$ , die wir als gegeben annehmen:

Für jede Teilmenge  $A \subseteq \mathbb{N}$  gilt: Ist  $1 \in A$  und ist für jedes  $n \in A$  auch  $n+1 \in A$ , dann ist  $A = \mathbb{N}$ .

Bei der vollständigen Induktion zeigen wir gerade, dass die Menge  $A := \{n \in \mathbb{N} \mid A(n) \text{ ist wahr}\}$  diese Bedingung erfüllt, also gleich  $\mathbb{N}$  ist.  $\square$

**Bemerkung.** Eine alternative Möglichkeit, die Aussage ‘Für alle  $n \in \mathbb{N}$  gilt  $A(n)$ ’ zu zeigen, wäre, ein *beliebiges*  $n \in \mathbb{N}$  zu wählen und dann  $A(n)$  mit einem der Prinzipien 1.3.1–1.3.3 zu beweisen. (Genau so wurde in Beispiel 1.3.1 und 1.3.2 vorgegangen.) Da vollständige Induktion nur für  $\mathbb{N}$  möglich ist, ist diese Alternative sogar der einzige Weg, um Aussagen ‘Für alle  $x \in M$  gilt  $A(x)$ ’ zu zeigen, bei denen die Menge  $M$  „größer“ als  $\mathbb{N}$  ist.

**Beispiel.** Für alle  $n \in \mathbb{N}$  gilt  $\sum_{i=1}^n i = \frac{n(n+1)}{2}$ .

*Beweis.* Wir führen eine vollständige Induktion über  $n$ . Sei also  $A(n)$  die Aussageform  $\sum_{i=1}^n i = \frac{n(n+1)}{2}$ .

Induktionsanfang: Es ist  $\sum_{i=1}^1 i = 1 = \frac{1 \cdot 2}{2}$ , d.h.  $A(1)$  ist wahr.

Induktionsschritt: Sei jetzt  $n \in \mathbb{N}$  beliebig. Wir zeigen  $A(n) \Rightarrow A(n+1)$  mittels eines direkten Beweises. Wir nehmen an, dass  $A(n)$  wahr ist, d.h.

dass  $\sum_{i=1}^n i = \frac{n(n+1)}{2}$  gilt. Dieses ist die Induktionsvoraussetzung (kurz IV). Dann ist

$$\begin{aligned} \sum_{i=1}^{n+1} i &= \left( \sum_{i=1}^n i \right) + (n+1) \stackrel{IV}{=} \frac{n(n+1)}{2} + (n+1) = \frac{n(n+1) + 2(n+1)}{2} \\ &= \frac{(n+1)(n+2)}{2}. \end{aligned}$$

Der Induktionsschritt ist damit erledigt, weil dies genau die Aussage  $A(n+1)$  ist.  $\square$

**Bemerkung.** Es gibt verschiedene Varianten der Induktion, z.B.

- (i) Der Induktionsanfang kann bei  $n_0 \in \mathbb{N}$  statt bei 1 gemacht werden. Damit wird die Aussage  $A(n)$  für alle  $n \geq n_0$  gezeigt.
- (ii) Als Induktionsvoraussetzung kann  $A(1) \wedge \dots \wedge A(n)$  anstelle von  $A(n)$  verwendet werden, was unter Umständen stärker ist.
- (iii) Es gibt die vollständige Induktion nicht nur für  $\mathbb{N}$  sondern auch eine sog. *strukturelle Induktion*, die z.B. über einen „Termaufbau“ geführt werden kann. Dies spielt in der Logik und bei formalen Sprachen eine Rolle.
- (iv) In der Informatik beweist man die Korrektheit von Algorithmen häufig mit sog. *Schleifeninvarianten*. Im Prinzip beweist man damit die Korrektheit des Algorithmus durch Induktion über die Anzahl der Schleifendurchläufe, und die Schleifeninvariante hat die Rolle der Induktionsvoraussetzung.

*Übung.* Man zeige mittels vollständiger Induktion, dass sich eine Tafel Schokolade mit  $n$  Stücken stets durch  $(n-1)$ -maliges Durchbrechen in Einzelstücke zerlegen lässt. Hier wird vorausgesetzt, dass einmaliges Durchbrechen ein einzelnes Stück in genau zwei Teile zerlegt. Hinweis: Verwenden Sie als Induktionsvoraussetzung  $A(1) \wedge \dots \wedge A(n)$ .

## 1.4 Abbildungen

### 1.4.1 Definition und Beispiele

**Definition a.** Seien  $M, N$  Mengen. Eine *Abbildung*  $f$  von  $M$  nach  $N$  ist eine „Vorschrift“ (z.B. eine Formel), die jedem  $x \in M$  genau ein Element  $f(x) \in N$  zuordnet, geschrieben

$$f : M \rightarrow N, \quad x \mapsto f(x).$$

Es heißen:  $M$  der *Definitionsbereich* von  $f$ ,  $N$  der *Zielbereich* oder *Wertebereich* von  $f$ ,  $f(x)$  das *Bild* von  $x$  unter  $f$ ,  $x$  ein *Urbild* von  $f(x)$  unter  $f$ .

Zur Angabe einer Abbildung gehört die Angabe von Definitions- und Zielbereich dazu, d.h. zwei Abbildungen  $f : M \rightarrow N$  und  $g : M' \rightarrow N'$  sind nur dann gleich, wenn  $M = M'$ ,  $N = N'$  und  $f(x) = g(x)$  für alle  $x \in M$ .

Die Menge aller Abbildungen von  $M$  nach  $N$  wird mit  $\text{Abb}(M, N)$  oder mit  $N^M$  bezeichnet.

### Beispiel a.

- (i)  $f : \mathbb{N} \rightarrow \mathbb{R}, i \mapsto i^2$ .
- (ii) Es sei  $M$  eine Menge von Glasperlen, und sei  $F$  die Menge aller Farben. Dann gibt es die Abbildung  $f : M \rightarrow F, x \mapsto \text{Farbe von } x$ .
- (iii) Für jede Menge  $A$  von Personen gibt es die Abbildung  $J : A \rightarrow \mathbb{Z}, p \mapsto \text{Geburtsjahr von } p$ .
- (iv) Die Addition in  $\mathbb{Z}$  kann als die Abbildung

$$\mathbb{Z} \times \mathbb{Z} \rightarrow \mathbb{Z}, \quad (x, y) \mapsto x + y$$

aufgefasst werden.

- (v) Für jede Menge  $M$  gibt es die *Identitätsabbildung*

$$\text{id}_M : M \rightarrow M, x \mapsto x.$$

- (vi) Betrachten wir die Abbildungen

$$\begin{aligned} f & : \mathbb{R} \rightarrow \mathbb{R}, & x & \mapsto \sqrt{x^2}, \\ g & : \mathbb{R} \rightarrow \mathbb{R}, & x & \mapsto |x|, \\ h & : \mathbb{R} \rightarrow \mathbb{R}_{\geq 0}, & x & \mapsto |x|, \end{aligned}$$

so ist  $f = g \neq h$ .

- (vii)  $\text{Abb}(\mathbb{R}, \mathbb{R}) = \{\mathbb{R} \rightarrow \mathbb{R}\} =$  Menge aller reellen Funktionen.
- (viii) Für jede Menge  $N$  existiert genau eine Abbildung  $\emptyset \rightarrow N$ .
- (ix) Für jede nicht-leere Menge  $M$  existiert keine Abbildung  $M \rightarrow \emptyset$ .

**Bemerkung.** Eine Abbildung  $f : \mathbb{N} \rightarrow N$  wird auch *Folge in  $N$*  genannt. Oft benutzt man für Folgen die Schreibweise  $a_1, a_2, a_3, \dots$  oder  $(a_i)_{i \in \mathbb{N}}$ , wobei  $a_i$  für das Bild  $f(i) \in N$  steht. Die Folge aus Beispiel a.(i) würde also auch geschrieben als  $1, 4, 9, 16, \dots$  oder als  $(i^2)_{i \in \mathbb{N}}$ .

Die Menge aller Folgen in  $N$  wird daher auch als  $\text{Abb}(\mathbb{N}, N)$  oder  $N^{\mathbb{N}}$  geschrieben. Beispielsweise ist  $\{0, 1\}^{\mathbb{N}}$  die Menge der Binärfolgen (manchmal auch geschrieben als  $2^{\mathbb{N}}$ ),  $\mathbb{R}^{\mathbb{N}}$  die Menge der reellen Folgen, usw.

**Definition b.** Es sei  $M$  eine Menge und  $n \in \mathbb{N}$ . Ein  $n$ -Tupel über  $M$  ist eine Abbildung  $t : \underline{n} \rightarrow M$ . Wie bei Folgen schreiben wir  $(x_1, \dots, x_n)$  oder  $(x_i)_{i \in \underline{n}}$  für  $t$ , wobei  $x_i := t(i)$  ist für  $i \in \underline{n}$ . Wir setzen  $M^n := M^{\underline{n}} = \text{Abb}(\underline{n}, M)$ .

**Beispiel b.** (i) Das 5-Tupel  $(1, -3, 0, 0, 27)$  über  $\mathbb{Z}$  ist z.B. die Abbildung  $t : \underline{5} \rightarrow \mathbb{Z}$  mit  $t(1) = 1, t(2) = -3, t(3) = t(4) = 0, t(5) = 27$ .

(ii) Für jede Menge  $N$  kann  $N^2$  mit  $N \times N$  identifiziert werden. (Hier wird das 2-Tupel  $(x, y) \in N^2$ , d.i. die Abbildung  $\{1, 2\} \rightarrow N, 1 \mapsto x, 2 \mapsto y$ , mit dem **geordneten Paar**  $(x, y) \in N \times N$  identifiziert.)

Schließlich können wir mit dem Abbildungsbegriff auch kartesische Produkte von mehr als zwei Mengen definieren.

**Definition c.** Es sei  $n \in \mathbb{N}$  und  $M_i$  eine Menge für alle  $i \in \underline{n}$ . Wir setzen

$$M := \bigcup_{i \in \underline{n}} M_i$$

und definieren

$$M_1 \times \dots \times M_n := \{f : \underline{n} \rightarrow M \mid f(i) \in M_i \text{ für alle } i \in \underline{n}\},$$

und nennen  $M_1 \times \dots \times M_n$  das *kartesische Produkt* der Mengen  $M_1, \dots, M_n$ .

Wie bei Folgen schreiben wir  $(x_1, \dots, x_n)$  oder  $(x_i)_{i \in \underline{n}}$  für  $f \in M_1 \times \dots \times M_n$ , wobei  $x_i := f(i)$  ist für  $1 \leq i \leq n$ .

Es ist also  $M_1 \times \dots \times M_n$  die Menge aller  $n$ -Tupel  $(x_1, \dots, x_n) = (x_i)_{i \in \underline{n}} \in M^n$  mit  $x_i \in M_i$  für  $i \in \underline{n}$ .

**Beispiel c.** Für jede Menge  $M$  und jede natürliche Zahl  $n \geq 2$  kann  $M^n$  mit dem  $n$ -fachen kartesischen Produkt  $M \times \dots \times M$  (mit  $n$  Faktoren) identifiziert werden.

Ersetzt man in Definition c die Menge  $\underline{n}$  durch eine beliebige Indexmenge  $I$ , erhält man das kartesische Produkt über  $I$ .



- Definition d.** (i) Es seien  $I$  und  $M$  Mengen. Eine Abbildung  $f : I \rightarrow M$  wird gelegentlich auch mit  $(x_i)_{i \in I}$  notiert, wobei  $x_i := f(i)$  ist für  $i \in I$ . In diesem Fall nennen wir  $(x_i)_{i \in I}$  eine durch  $I$  indizierte *Familie* in  $M$ .
- (ii) Es sei  $I$  eine Menge und  $M_i$  eine Menge für alle  $i \in I$ . Wir setzen

$$M := \bigcup_{i \in I} M_i$$

und definieren

$$\prod_{i \in I} M_i := \{f : I \rightarrow M \mid f(i) \in M_i \text{ für alle } i \in I\},$$

und nennen  $\prod_{i \in I} M_i$  das *kartesische Produkt* der Mengen  $M_i, i \in I$ .

In der oben eingeführten Schreibweise gilt also

$$\prod_{i \in I} M_i := \{(x_i)_{i \in I} \mid x_i \in M_i \text{ für alle } i \in I\}.$$

*Übung.*

- (i) Bestimmen Sie  $|\text{Abb}(N, M)|$  für endliche Mengen  $N$  und  $M$ .
- (ii) Wie viele Elemente hat  $M_1 \times \cdots \times M_n$  für  $n \in \mathbb{N}$  und endliche Mengen  $M_1, \dots, M_n$ ?
- (iii) Wie viele Elemente hat  $M^n$  für  $n \in \mathbb{N}$  und eine endlichen Menge  $M$ ?

### 1.4.2 Definition durch Rekursion

Folgen auf einer Menge können *rekursiv* definiert werden.

**Beispiel.** (i) Auf  $\mathbb{R}_{>0}$  existiert genau eine Folge  $(a_n)_{n \in \mathbb{N}}$  mit

$$a_1 := 1 \text{ und } a_{n+1} := 1 + \frac{1}{a_n} \text{ für } n \geq 1.$$

(ii) Es sei  $a \in \mathbb{R}$ . Es gibt genau eine Folge  $x = (x_n)_{n \in \mathbb{N}}$  in  $\mathbb{R}$  mit

$$x_1 = a \text{ und } x_{n+1} = a \cdot x_n \text{ für } n \geq 1.$$

Wir schreiben:  $a^n := x_n$  für das  $n$ -te Glied dieser Folge.

Alternativ verwenden wir für dieses Vorgehen oft auch die Sprechweise:

Für  $a \in \mathbb{R}$  definieren wir die *Potenzen*  $a^n$  für  $n \in \mathbb{N}$  *rekursiv* durch:

$$a^1 := a \text{ und } a^{n+1} := a \cdot a^n \text{ für } n \geq 1.$$

Die Definition durch Rekursion beruht auf dem folgenden Satz.

**Satz.** Es sei  $N$  eine Menge,  $f: N \rightarrow N$  Abbildung und  $a \in N$ .  
Dann gibt es genau eine Folge  $(a_n)_{n \in \mathbb{N}}$  in  $N$  mit:

- $a_1 = a$
- $a_{n+1} = f(a_n)$  für  $n \in \mathbb{N}$ .

Dieser *Rekursionssatz von Dedekind* kann durch vollständige Induktion bewiesen werden. Wir verzichten hier auf einen Beweis.

**Bemerkung.** Wir erhalten die Folgen aus Beispiel 1.4.2 mittels der folgenden Abbildungen.

- (i)  $f: \mathbb{R}_{>0} \rightarrow \mathbb{R}_{>0}$ ,  $x \mapsto 1 + 1/x$ .
- (ii)  $f: \mathbb{R} \rightarrow \mathbb{R}$ ,  $x \mapsto ax$ .

### 1.4.3 Bild und Urbild

**Definition.** Es sei  $f: M \rightarrow N$  eine Abbildung.

- (i) Für jede Teilmenge  $X \subseteq M$  heißt  $f(X) := \{f(x) \mid x \in X\}$  das *Bild von  $X$  unter  $f$* .
- (ii) Das Bild  $f(M)$  von  $M$  unter  $f$  wird schlicht das *Bild von  $f$*  genannt.
- (iii) Für jede Teilmenge  $Y \subseteq N$  heißt  $f^{-1}(Y) := \{x \in M \mid f(x) \in Y\}$  das *Urbild von  $Y$  unter  $f$* .
- (iv) Die Mengen  $f^{-1}(\{y\})$  mit  $y \in N$  heißen die *Fasern von  $f$* .

Die Schreibweise  $f^{-1}$  für das Urbild hat im Allgemeinen nichts mit Umkehrabbildungen zu tun.

**Beispiel.** Die Faser der Abbildung  $J$  aus Beispiel 1.4.1a.(iii) zu 2000 ist die Menge aller Personen, die im Jahr 2000 geboren sind.

**Bemerkung a.** Die nicht-leeren Fasern einer Abbildung bilden eine Partition des Definitionsbereichs.

### 1.4.4 Injektive und surjektive Abbildungen

**Definition.** Es sei  $f : M \rightarrow N$  eine Abbildung.

- (i)  $f$  heißt *surjektiv*, falls  $f(M) = N$ .
- (ii)  $f$  heißt *injektiv*, falls für alle  $x, x' \in M$  gilt:  $f(x) = f(x') \Rightarrow x = x'$ .
- (iii)  $f$  heißt *bijektiv*, falls  $f$  injektiv und surjektiv ist.

**Bemerkung a.** Eine Abbildung  $f : M \rightarrow N$  ist per Definition injektiv, surjektiv bzw. bijektiv, wenn jedes Element  $y \in N$  höchstens ein, mindestens ein bzw. genau ein Urbild hat. Das ist genau dann der Fall, wenn alle Fasern von  $f$  höchstens ein, mindestens ein bzw. genau ein Element besitzen, also genau dann, wenn für jedes  $y \in N$  die Gleichung  $f(x) = y$  höchstens eine, mindestens eine bzw. genau eine Lösung  $x \in M$  hat.

**Beispiel.**

- (i)  $f : \mathbb{Z} \rightarrow \mathbb{Z}, z \mapsto 2z$  ist injektiv, aber nicht surjektiv.
- (ii)  $f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto 2x$  ist bijektiv.
- (iii)  $f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto x^2$  ist weder injektiv noch surjektiv. In der Tat ist  $f(\mathbb{R}) = \mathbb{R}_{\geq 0}$ , also  $f$  nicht surjektiv. Weiter ist z.B.  $f(2) = 4 = f(-2)$  aber  $2 \neq -2$ , folglich ist  $f$  nicht injektiv.
- (iv) Es sei  $f : M \rightarrow F$  die Abbildung aus Beispiel (1.4.1)a.(ii). Die Faser  $f^{-1}(\{\text{rot}\})$  ist die Menge der roten Perlen in  $M$ . Es ist  $f$  genau dann injektiv, wenn von jeder Farbe höchstens eine Perle in  $M$  vorkommt, wenn also keine zwei Perlen aus  $M$  die gleiche Farbe haben. Weiter ist  $f$  genau dann surjektiv, wenn von jeder Farbe (mindestens) eine Perle in  $M$  vorkommt.
- (v) Die Abbildung  $\emptyset \rightarrow N$  ist injektiv. Sie ist genau dann surjektiv, wenn  $N = \emptyset$ .
- (vi) Hashfunktionen (bzw. „Checksummen“ oder „Fingerprints“), z.B. die bekannte  $\text{md5} : \{\text{Texte}\} \rightarrow \{0, 1\}^{128}$ , die einen 128-bit Hashwert produziert, sind nicht injektiv (da verschiedene Texte gleichen Hashwert haben können), sind surjektiv (um alle Hashwerte auszunutzen), und haben „gleich große“ Fasern (das macht gerade eine gute Hashfunktion aus!).
- (vii) Verschlüsselungsfunktionen, etwa  $\text{crypt} : \{0, 1\}^k \rightarrow \{0, 1\}^k$ , sind injektiv, damit eine eindeutige Entschlüsselung möglich ist.

*Übung.* Man mache sich klar, dass eine Abbildung  $f : M \rightarrow N$  genau dann injektiv ist, wenn für alle  $x_1, \dots, x_r \in M$  gilt:

$$x_1, \dots, x_r \text{ paarweise verschieden} \Leftrightarrow f(x_1), \dots, f(x_r) \text{ paarweise verschieden.}$$

### 1.4.5 Einschränkung

**Definition.** Es sei  $f : M \rightarrow N$  eine Abbildung und  $M' \subseteq M$ . Dann heißt die Abbildung

$$f|_{M'} : M' \rightarrow N, \quad x \mapsto f(x)$$

die *Einschränkung* von  $f$  auf  $M'$ .

**Bemerkung.** Jede Abbildung kann durch Einschränkung auf eine geeignete Teilmenge des Definitionsbereichs injektiv gemacht werden. Z.B. ist für  $f$  aus Beispiel 1.4.4(iii) die Einschränkung  $f|_{\mathbb{R}_{\geq 0}}$  injektiv, ebenso wie die Einschränkung  $f|_{\mathbb{R}_{\leq 0}}$ .

### 1.4.6 Komposition

**Definition.** Es seien  $M, N, L$  Mengen. Weiter seien  $f : M \rightarrow N$  und  $g : N \rightarrow L$  zwei Abbildungen. Dann heißt die Abbildung

$$g \circ f : M \rightarrow L, \quad x \mapsto (g \circ f)(x) := g(f(x))$$

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

$$\begin{array}{ccccc} & & g \circ f & & \\ & \searrow & \text{---} & \nearrow & \\ M & \xrightarrow{f} & N & \xrightarrow{g} & L \end{array}$$

$$x \xrightarrow{f} f(x) \xrightarrow{g} g(f(x))$$

**Beispiel.** Für die Abbildungen

$$\begin{aligned} f : \mathbb{R} &\rightarrow \mathbb{R}_{\geq 0}, & x &\mapsto (x-3)^2, \\ g : \mathbb{R}_{\geq 0} &\rightarrow \mathbb{R}, & x &\mapsto \sqrt{x} \end{aligned}$$

ergeben sich die Kompositionen

$$\begin{aligned} g \circ f : \mathbb{R} &\rightarrow \mathbb{R}, & x &\mapsto \sqrt{(x-3)^2} = |x-3|, \\ f \circ g : \mathbb{R}_{\geq 0} &\rightarrow \mathbb{R}_{\geq 0}, & x &\mapsto (\sqrt{x}-3)^2 \end{aligned}$$

**Bemerkung.** Es seien  $f, g, h$  Abbildungen.

- (i) Es gilt  $(h \circ g) \circ f = h \circ (g \circ f)$ , sofern beide Seiten der Gleichung definiert sind. Daher kann die Komposition auch ohne Klammern kurz als  $h \circ g \circ f$  geschrieben werden.

### 1.4.7 Umkehrabbildungen

**Definition.** Es seien  $f : M \rightarrow N$  und  $g : N \rightarrow M$  Abbildungen. Dann heißt  $g$  eine *linksseitige (rechtsseitige) Umkehrabbildung von  $f$* , wenn  $g \circ f = \text{id}_M$  (wenn  $f \circ g = \text{id}_N$ ). Wir sprechen schlicht von einer *Umkehrabbildung von  $f$* , wenn  $g$  sowohl links- als auch rechtsseitige Umkehrabbildung von  $f$  ist.

**Satz a.** Sei  $f : M \rightarrow N$  eine Abbildung und sei  $M$  nicht leer.

- (i)  $f$  besitzt genau dann eine linksseitige Umkehrabbildung, wenn  $f$  injektiv ist.
- (ii)  $f$  besitzt genau dann eine rechtsseitige Umkehrabbildung, wenn  $f$  surjektiv ist.
- (iii)  $f$  besitzt genau dann eine Umkehrabbildung, wenn  $f$  bijektiv ist.

**Bemerkung.** Existiert eine Umkehrabbildung, so ist sie eindeutig bestimmt (Übung). Links- und rechtsseitige Umkehrabbildungen sind im Allgemeinen nicht eindeutig (Beispiel unten).

**Schreibweise.** Falls  $f$  bijektiv ist, so wird die eindeutige Umkehrabbildung mit  $f^{-1}$  bezeichnet. Die ist nicht zu verwechseln mit dem Urbild, das ebenfalls mit  $f^{-1}$  bezeichnet wird. Was gemeint ist, ergibt sich aus dem Zusammenhang.

*Beweis.* (i) Es sind zwei Richtungen zu zeigen, wir zeigen zuerst den „wenn“-Teil. Dazu nehmen wir an,  $f$  sei injektiv und konstruieren eine linksseitige Umkehrabbildung  $g$ . Wähle  $x_0 \in M$  beliebig ( $M \neq \emptyset$ ) und definiere  $g : N \rightarrow M$  durch

$$g(y) := \begin{cases} x & \text{falls } y = f(x) \text{ für ein } x \in M, \\ x_0 & \text{falls } y \notin f(M), \end{cases}$$

Das  $x$  in der ersten Zeile ist eindeutig, da  $f$  injektiv ist, also ist  $g$  wohldefiniert. Damit gilt  $(g \circ f)(x) = g(f(x)) = x$  für alle  $x \in M$ , d.h.  $g \circ f = \text{id}_M$  wie gewünscht.

Wir zeigen jetzt die andere Richtung, den „genau dann“-Teil. Dazu nehmen wir an,  $g : N \rightarrow M$  sei eine linksseitige Umkehrabbildung und folgern, dass  $f$  injektiv ist. Aus  $g \circ f = \text{id}_M$  folgt, dass für alle  $x, x' \in M$  gilt:

$$f(x) = f(x') \Rightarrow g(f(x)) = g(f(x')) \Rightarrow \underbrace{(g \circ f)(x)}_{=\text{id}_M} = \underbrace{(g \circ f)(x')}_{=\text{id}_M} \Rightarrow x = x'.$$

Also ist  $f$  tatsächlich injektiv und der Beweis beendet.

(ii), (iii) siehe Vorlesung. □

### Beispiel.

(i)  $f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto 2x$  ist bijektiv mit der Umkehrabbildung

$$f^{-1} : \mathbb{R} \rightarrow \mathbb{R}, \quad x \mapsto \frac{1}{2}x$$

(ii)  $f : \mathbb{R}_{\geq 0} \rightarrow \mathbb{R}_{\geq 1}, x \mapsto x^2 + 1$  ist bijektiv mit der Umkehrabbildung

$$f^{-1} : \mathbb{R}_{\geq 1} \rightarrow \mathbb{R}_{\geq 0}, \quad x \mapsto \sqrt{x - 1}$$

(iii)  $f : \mathbb{Z} \rightarrow \mathbb{Q}, a \mapsto a$  ist injektiv, aber nicht surjektiv.

$$g \circ f = \text{id}_{\mathbb{Z}} \quad : \quad \text{z.B. } g(x) := \lfloor x \rfloor \text{ oder } g(x) := \lceil x \rceil$$

$$f \circ g = \text{id}_{\mathbb{Q}} \quad : \quad \text{nicht möglich}$$

(Hier bezeichnet  $\lfloor x \rfloor$  die größte ganze Zahl  $\leq x$ , und  $\lceil x \rceil$  die kleinste ganze Zahl  $\geq x$ .)

(iv)  $f : \mathbb{R} \rightarrow \mathbb{R}_{\geq 0}, x \mapsto |x|$  ist surjektiv, aber nicht injektiv.

$$g \circ f = \text{id}_{\mathbb{R}} \quad : \quad \text{nicht möglich}$$

$$f \circ g = \text{id}_{\mathbb{R}_{\geq 0}} \quad : \quad \text{z.B. } g(x) := x \text{ oder } g(x) := -x$$

**Satz b.** *Es seien  $f : M \rightarrow N$  und  $g : N \rightarrow L$  zwei bijektive Abbildungen. Wenn  $g \circ f$  definiert ist, so ist  $g \circ f$  ebenfalls bijektiv und es gilt:*

$$(g \circ f)^{-1} = f^{-1} \circ g^{-1}.$$

*Beweis.* als Übung. □

*Übung.*

(i) Zeigen Sie die restlichen Teile der Bemerkung.

(ii) Zeigen Sie den Satz.

(iii) Gilt der Satz auch, wenn man bijektiv durch injektiv ersetzt?

(iv) Gilt der Satz auch, wenn man bijektiv durch surjektiv ersetzt?

### 1.4.8 Abbildungen einer Menge in sich

Sind  $f, g : M \rightarrow M$  zwei Abbildungen einer Menge  $M$  in sich, so kann man stets die Kompositionen  $f \circ g$  und  $g \circ f$  bilden.

**Definition.** Es sei  $f : M \rightarrow M$  eine Abbildung und es sei  $n \in \mathbb{N}$ . Dann setzen wir

$$f^n := \underbrace{f \circ \dots \circ f}_{n\text{-mal}}, \quad f^0 := \text{id}_M.$$

Falls  $f$  bijektiv ist, so definieren wir auch  $f^{-n} := (f^{-1})^n$ .

**Bemerkung.**

- (i) Es gilt  $f^n(x) = f(f(\dots f(x)))$ .
- (ii) Für bijektive Abbildungen einer Menge in sich selbst haben wir die üblichen Potenzrechenregeln:

$$f^{a+b} = f^a \circ f^b \quad \text{und} \quad f^{ab} = (f^a)^b \quad \text{für alle } a, b \in \mathbb{Z}.$$

### 1.4.9 Die Mächtigkeit von Mengen

**Definition a.** Zwei Mengen  $M$  und  $N$  heißen *gleichmächtig*, wenn eine bijektive Abbildung  $M \rightarrow N$  existiert.

*Übung a.*  $\mathbb{N}, \mathbb{Z}$  und  $\mathbb{Q}$  sind gleichmächtig.

**Satz a (Cantor).** Für jede Menge  $M$  sind  $M$  und  $\text{Pot}(M)$  nicht gleichmächtig.

*Beweis.* Sei  $f$  eine beliebige Abbildung  $f : M \rightarrow \text{Pot}(M)$ . Definiere  $A_f := \{x \in M \mid x \notin f(x)\} \in \text{Pot}(M)$ . Angenommen, es gibt  $m \in M$  mit  $f(m) = A_f$ . Falls  $m \in A_f$ , so folgt  $m \notin f(m) = A_f$  (Widerspruch). Falls  $m \notin A_f = f(m)$ , so folgt  $m \in A_f$  (Widerspruch). Also ist  $f$  nicht surjektiv.  $\square$

*Übung b.* Man folgere aus dem Satz:

- (i)  $\mathbb{N}$  und  $\mathbb{R}$  sind nicht gleichmächtig.
- (ii) Die Zusammenfassung aller Mengen ist keine Menge.

Nun können wir eine exakte Definition der Endlichkeit einer Menge geben.

**Definition b.** Es sei  $M$  eine Menge.

- (i)  $M$  heißt *endlich*, wenn  $M$  gleichmächtig zu  $\underline{n}$  für ein  $n \in \mathbb{N}_0$  ist (Erinnerung:  $\underline{0} = \emptyset$ ).

In diesem Fall definieren wir  $|M| := n$ , und nennen  $|M|$  die *Mächtigkeit* von  $M$  (oder die *Anzahl der Elemente* von  $M$ ).

- (ii)  $M$  heißt *unendlich*, wenn  $M$  nicht endlich ist.

Für Abbildungen zwischen endlichen Mengen gibt es Beziehungen zwischen deren Mächtigkeit.

**Bemerkung a.** Es seien  $M, N$  endliche Mengen und  $f : M \rightarrow N$  eine Abbildung. Dann gelten  $|f(M)| \leq |M|$  und  $|f(M)| \leq |N|$ .

*Übung c.* Man folgere aus Bemerkung a, dass für eine injektive Abbildung  $f : M \rightarrow N$  stets  $|M| \leq |N|$  ist, und für eine surjektive Abbildung  $f : M \rightarrow N$  stets  $|M| \geq |N|$  ist.

Genauer kann man bei Abbildungen zwischen endlichen Mengen Injektivität, Surjektivität und Bijektivität wie folgt charakterisieren.

**Satz b.** Es sei  $f : M \rightarrow N$  eine Abbildung und  $M, N$  endlich.

- (i)  $f$  injektiv  $\Leftrightarrow |f(M)| = |M|$ .  
 (ii)  $f$  surjektiv  $\Leftrightarrow |f(M)| = |N|$ .  
 (iii) Ist  $|M| = |N|$ , dann sind äquivalent:

- $f$  injektiv
- $f$  surjektiv
- $f$  bijektiv

Darauf beruht das berühmte Dedekind'sche Schubfachprinzip:

**Bemerkung b.** Werden  $m$  Objekte auf  $n$  Schubfächer verteilt, und ist  $m > n$ , dann gibt es ein Schubfach, welches mindestens zwei Objekte enthält.

Dies ist genau die Aussage: Sind  $M, N$  endliche Mengen mit  $|M| > |N|$ , und  $f : M \rightarrow N$  eine Abbildung, dann ist  $f$  nicht injektiv.

*Übung.* Bestimmen Sie die Anzahl injektiver Abbildungen von  $\underline{m}$  nach  $\underline{n}$ .

*Übung.* Es sei  $f : M \rightarrow N$  eine Abbildung zwischen endlichen Mengen. Dann gilt

$$|M| < |N| \Rightarrow f \text{ nicht surjektiv.}$$



### 1.4.10 Kombinatorische Strukturen als Abbildungen

Tupel, Permutationen, Kombinationen und Multimengen (Definition erst in späterem Kapitel) können mit Abbildungen bestimmter Art identifiziert werden.

**Beispiel.** Es sei  $A$  eine Menge und  $k \in \mathbb{N}$ .

- (i) Eine  $k$ -Permutation aus  $A$  ist eine injektive Abbildung  $\underline{k} \rightarrow A$ . Die Permutation  $(a_1, \dots, a_k)$  entspricht der Abbildung  $f : \underline{k} \rightarrow A, i \mapsto a_i$ .
- (ii) Ist  $|A| = n \in \mathbb{N}$ , so ist eine Permutation aus  $A$  eine bijektive Abbildung  $\underline{n} \rightarrow A$ . Die Permutation  $(a_1, \dots, a_n)$  entspricht der Abbildung  $f : \underline{n} \rightarrow A, i \mapsto a_i$ .
- (iii) Eine  $k$ -Kombination aus  $A$  ist eine Abbildung  $A \rightarrow \{0, 1\}$  mit  $|f^{-1}(\{1\})| = k$  (die Faser zu 1 hat  $k$  Elemente). Die Kombination  $M \subseteq A$  entspricht der Abbildung  $f : A \rightarrow \{0, 1\}$  mit  $f(a) = 0$  falls  $a \notin M$  und  $f(a) = 1$  falls  $a \in M$ . Die Abbildung  $f$  bezeichnet man auch als *charakteristische Funktion* von  $M$ .
- (iv) Eine  $k$ -Multimenge ist eine Abbildung  $A \rightarrow \mathbb{N}_0$  mit  $\sum_{a \in A} f(a) = k$ . Die Multimenge  $M \subseteq A$  entspricht der Abbildung  $f : A \rightarrow \mathbb{N}_0$ , wobei  $f(a)$  angibt, wie oft  $a$  in  $M$  vorkommt. Die Abbildung  $f$  wird als *Häufigkeitsfunktion* von  $M$  bezeichnet.

*Übung.* Eine  $k$ -elementige Teilmenge  $M$  von  $A$  kann als  $k$ -Kombination oder als  $k$ -Multimenge aufgefasst werden. Vergleichen Sie die charakteristische Funktion von  $M$  mit der Häufigkeitsfunktion von  $M$ .

## 1.5 Relationen

### 1.5.1 Definition und Beispiele

Relationen drücken Beziehungen zwischen Elementen von zwei Mengen aus, z.B. wäre „liegt in“ eine Relation zwischen  $\{\text{Städte}\}$  und  $\{\text{Länder}\}$ . In der Informatik werden Relationen z.B. in relationalen Datenbanken verwendet.

**Definition.** Es seien  $M$  und  $N$  zwei Mengen.

- (i) Eine Teilmenge  $R \subseteq M \times N$  heißt *Relation zwischen  $M$  und  $N$* , oder kürzer *Relation auf  $M$*  falls  $M = N$ . Für  $(x, y) \in R$  schreiben wir auch  $xRy$  und sagen „ $x$  steht in Relation zu  $y$  bzgl.  $R$ “.

- (ii) Eine Relation  $R \subseteq M \times M$  auf  $M$  heißt
- (R) *reflexiv*, falls  $xRx$  für alle  $x \in M$ ,
  - (R') *antireflexiv*, falls nicht  $xRx$  für alle  $x \in M$ ,
  - (S) *symmetrisch*, falls  $xRy \Rightarrow yRx$  für alle  $x, y \in M$ ,
  - (A) *antisymmetrisch*, falls  $(xRy \wedge yRx) \Rightarrow x = y$  für alle  $x, y \in M$ ,
  - (T) *transitiv*, falls  $(xRy \wedge yRz) \Rightarrow xRz$  für alle  $x, y, z \in M$ .
- (iii) Eine Relation, die (R), (S) und (T) erfüllt, heißt *Äquivalenzrelation*.
- (iv) Eine Relation, die (R), (A) und (T) erfüllt, heißt *(partielle) Ordnung*.
- (v) Eine Relation, die (R) und (T) erfüllt, heißt *Präordnung*.
- (vi) Eine Ordnung heißt *Totalordnung*, wenn  $xRy \vee yRx$  für alle  $x, y \in M$ .

### Beispiel.

- (i)  $M = \mathbb{R}$  und  $R = „\leq“$ , d.h.  $(x, y) \in R$  genau dann, wenn  $x \leq y$ .  
 $„\leq“$  ist reflexiv, antisymmetrisch und transitiv, also eine Ordnung.  
 $„\leq“$  ist sogar eine Totalordnung.
- (ii)  $M = \mathbb{R}$  und  $R = „<“$ , d.h.  $(x, y) \in R \Leftrightarrow x < y$ .  
 $„<“$  ist antisymmetrisch(!) und transitiv, aber weder reflexiv noch symmetrisch.
- (iii)  $M = \text{Pot}(N)$  und  $R = „\subseteq“$ .  
 $„\subseteq“$  ist eine Ordnung. Falls  $|N| \geq 2$ , so ist  $„\subseteq“$  jedoch keine Totalordnung, da z.B. für  $\{1\}, \{2\} \in \text{Pot}\{1, 2\}$  weder  $\{1\} \subseteq \{2\}$  noch  $\{2\} \subseteq \{1\}$  gilt.
- (iv)  $M = \mathbb{Z}$ . Die *Teilbarkeitsrelation*  $„|“$  ist erklärt durch  $x \mid y$  genau dann, wenn ein  $z \in \mathbb{Z}$  existiert mit  $xz = y$ . Sie ist reflexiv und transitiv, also eine Präordnung. Sie ist nicht antisymmetrisch, denn  $1 \mid -1$  und  $-1 \mid 1$  obwohl  $1 \neq -1$ . Also ist  $„|“$  keine Ordnung auf  $\mathbb{Z}$ .
- (v) Die *Teilbarkeitsrelation*  $„|“$  ist eine Ordnung auf  $\mathbb{N}$ , aber keine Totalordnung.
- (vi) Auf jeder Menge  $M$  stellt die *Gleichheit*  $„=“$  eine Äquivalenzrelation dar mit  $R = \{(x, x) \mid x \in M\}$ .

- (vii) Auf einer Menge  $M$  von Personen können zwei Relationen  $V$  und  $G$  erklärt werden durch:

$$xVy :\Leftrightarrow x \text{ ist verwandt mit } y,$$

$$xGy :\Leftrightarrow x \text{ hat das gleiche Geburtsdatum (Tag und Monat) wie } y.$$

Beide sind Äquivalenzrelationen. Ersetzt man „verwandt“ durch „erstgradig verwandt“, so ist  $V$  nicht mehr transitiv.

- (viii) Jede Abbildung  $f : M \rightarrow N$  kann als Relation zwischen  $M$  und  $N$  aufgefasst werden:

$$f = \{(x, f(x)) \mid x \in M\}.$$

Abbildungen sind also eine spezielle Art von Relationen.

- (ix) Für jede Abbildung  $f : M \rightarrow N$  kann man eine Relation  $R_f$  auf  $M$  erklären durch

$$xR_fy :\Leftrightarrow f(x) = f(y) \text{ (d.h. } x \text{ und } y \text{ liegen in derselben Faser von } f).$$

$R_f$  ist eine Äquivalenzrelation.

- (x)  $M = \mathbb{Z}$ . Die *Paritätsrelation* „ $\equiv_2$ “, definiert durch

$$x \equiv_2 y :\Leftrightarrow x - y \text{ gerade}$$

ist eine Äquivalenzrelation auf  $\mathbb{Z}$ .

- (xi) Sei  $M$  eine Menge und  $\leq$  eine Präordnung auf  $M$ . Definiere Relation  $\diamond$  auf  $M$  durch

$$x \diamond y :\Leftrightarrow x \leq y \text{ und } y \leq x.$$

Dann ist  $\diamond$  eine Äquivalenzrelation auf  $M$ .

*Übung.* Durch welche Datenstruktur würden Sie eine Relation auf einer endlichen Menge in einem Computerprogramm repräsentieren? Wie prüfen Sie anhand dieser Datenstruktur, ob die Relation reflexiv, symmetrisch bzw. antisymmetrisch ist?

*Übung.* Es seien  $R$  eine Relation auf  $A$  und  $A' \subseteq A$ . Dann ist  $R' := R \cap (A' \times A')$  eine Relation auf  $A'$ . Man mache sich klar, dass jede der Eigenschaften aus Teil (ii) der Definition beim Übergang von  $R$  zu  $R'$  erhalten bleibt.

*Übung.* Welche Bedingung muss eine Relation  $R \subseteq N \times M$  erfüllen, damit sie im Sinne von Beispiel (ix) als eine Abbildung von  $N$  nach  $M$  aufgefasst werden kann? Unter welcher Bedingung ist diese Abbildung injektiv, surjektiv bzw. bijektiv? Welche Relation gehört im bijektiven Fall zur Umkehrabbildung?

### 1.5.2 Äquivalenzrelationen

**Definition.** Es sei  $\sim$  eine Äquivalenzrelation auf  $M$ . Für  $x \in M$  heißt

$$[x] := [x]_{\sim} := \{y \in M \mid x \sim y\}$$

die *Äquivalenzklasse von  $\sim$  zu  $x$* . Die Menge aller Äquivalenzklassen von  $\sim$  wird mit  $M/\sim$  bezeichnet.

**Bemerkung.** Es sei  $\sim$  eine Äquivalenzrelation auf  $M$ . Dann gilt für alle  $x, y \in M$ :

- (i)  $x \in [x]_{\sim}$ ,
- (ii)  $y \in [x]_{\sim} \Leftrightarrow x \in [y]_{\sim}$ ,
- (iii)  $y \in [x]_{\sim} \Rightarrow [y]_{\sim} = [x]_{\sim}$ .

Wegen (iii) bezeichnet man jedes Element einer Äquivalenzklasse als ein *Repräsentant* derselben.

*Beweis.* als Übung. □

**Beispiel.**

- (i) Für die Gleichheitsrelation auf einer Menge  $M$  ist  $[x]_{=} = \{x\}$  und  $M/_{=} = \{\{x\} \mid x \in M\}$ .
- (ii) Für die Äquivalenzrelationen  $V$  und  $G$  aus Beispiel (1.5.1)(vii) gilt für jede Person  $P$  der Menge:

$$\begin{aligned} [P]_V &= \{\text{Verwandte von } P\}, \\ [P]_G &= \{\text{Personen, die am gleichen Tag Geburtstag feiern wie } P\}. \end{aligned}$$

- (iii) Es sei  $f : N \rightarrow M$  eine Abbildung und  $R_f$  die Äquivalenzrelation aus Beispiel (1.5.1)(ix). Dann ist

$$[x]_{R_f} = \{x' \in N \mid f(x) = f(x')\} = f^{-1}(\{x\}),$$

für jedes  $x \in N$ , und  $M/R_f$  ist die Menge der nicht-leeren Fasern von  $f$ .

- (iv) Für die Paritätsrelation aus Beispiel (1.5.1)(x) ist

$$\begin{aligned} [0]_{\equiv_2} &= \{a \in \mathbb{Z} \mid a \text{ gerade}\}, \\ [1]_{\equiv_2} &= \{a \in \mathbb{Z} \mid a \text{ ungerade}\}, \end{aligned}$$

und  $M/\equiv_2 = \{[0]_{\equiv_2}, [1]_{\equiv_2}\}$ .

- (v) Betrachte die Teilbarkeitsrelation „ $|$ “ auf  $\mathbb{Z}$ . Dies ist eine Präordnung. Sei  $\diamond$  die daraus gemäß Beispiel (1.5.1)(xi) gebildete Äquivalenzrelation (d.h.  $z \diamond z' \Leftrightarrow z \mid z'$  und  $z' \mid z$ ). Dann ist  $[z]_\diamond = \{z, -z\}$ .

Offensichtlich „partitioniert“ eine Äquivalenzrelation die Menge.

**Satz.** *Es sei  $M$  eine Menge.*

- (i) *Ist  $\sim$  eine Äquivalenzrelation auf  $M$ , so ist  $M/\sim$  eine Partition von  $M$ .*
- (ii) *Ist  $\mathcal{P}$  eine Partition von  $M$ , so existiert eine Äquivalenzrelation  $\sim$  auf  $M$  mit  $M/\sim = \mathcal{P}$ .*

*Die Äquivalenzrelationen auf  $M$  entsprechen also den Partitionen von  $M$ .*

*Beweis.*

- (i) Sei  $\sim$  eine Äquivalenzrelation auf  $M$  und setze  $\mathcal{P} := M/\sim$ . Wegen  $x \in [x]_\sim$  sind alle Äquivalenzklassen nicht leer und ihre Vereinigung ist ganz  $M$ . Es bleibt zu zeigen, dass die Äquivalenzklassen paarweise disjunkt sind (vgl. Definition (1.2.5)(iv)). Betrachte also zwei beliebige Klassen  $[x]_\sim, [y]_\sim$  mit  $x, y \in M$ . Zu zeigen ist:

$$[x]_\sim \neq [y]_\sim \Rightarrow [x]_\sim \cap [y]_\sim = \emptyset,$$

bzw. die Kontraposition

$$[x]_\sim \cap [y]_\sim \neq \emptyset \Rightarrow [x]_\sim = [y]_\sim.$$

Ist aber  $z \in [x]_\sim \cap [y]_\sim$ , so folgt daraus nach Teil (iii) der Bemerkung  $[x]_\sim = [z]_\sim = [y]_\sim$ .

- (ii) Durch die Vorschrift

$$x \sim y :\Leftrightarrow x \text{ und } y \text{ liegen in demselben Teil der Partition}$$

wird eine Äquivalenzrelation definiert. (Man überprüfe das!)  
Die Äquivalenzklassen sind offensichtlich genau die Teile von  $\mathcal{P}$ .

□

### 1.5.3 Partielle Ordnungen

Es sei  $\trianglelefteq$  eine partielle Ordnung auf  $M$ .

**Definition.** Ein Element  $m \in M$  heißt *minimales Element*, falls kein  $m' \in M$  existiert mit  $m' \neq m$  und  $m' \trianglelefteq m$ .

Ein Element  $m \in M$  wird heißt *kleinstes Element* oder *Minimum*, falls für alle  $m' \in M$  gilt:  $m \trianglelefteq m'$ .

Analog definiert man *maximales Element* und *größtes Element* (Übung). Ein *größtes Element* von  $M$  heißt auch *Maximum*.

**Bemerkung a.** Nach Definition bedeutet

$$\begin{array}{ll} m \text{ Minimum von } M : & \text{für alle } x \in M \text{ gilt } m \trianglelefteq x. \\ m \text{ minimal in } M : & \text{für alle } x \in M \text{ gilt } x \trianglelefteq m \Rightarrow x = m. \end{array}$$

Minimal zu sein ist also zu verstehen als „kein anderes ist kleiner“. Minimum zu sein ist also zu verstehen als „alle anderen sind größer“.

**Beispiel.** Wir betrachten die Teilbarkeitsrelation „|“ auf  $\mathbb{N}$ . Minimal zu sein bzgl. „|“ bedeutet „kein anderes ist Teiler“. Minimum zu sein bzgl. „|“ bedeutet „alle anderen sind Vielfache“.

- (i) Die Menge  $\{2, 3, 4, 6\}$  besitzt kein Minimum, hat aber die minimalen Elemente 2 und 3.
- (ii) Die Menge  $\{2, 3, 5\}$  besitzt kein Minimum, und jedes Element ist minimal.
- (iii) Die Menge  $\{2, 4, 6\}$  besitzt das Minimum 2, und 2 ist das einzige minimale Element.

**Satz.** Es sei  $\trianglelefteq$  eine partielle Ordnung auf  $M$ .

- (i) Jedes Minimum von  $M$  ist minimal in  $M$ .
- (ii) Existiert ein Minimum von  $M$ , so ist es das einzige minimale Element in  $M$ . Insbesondere ist das Minimum eindeutig.
- (iii) Bei einer Totalordnung ist jedes minimale Element in  $M$  auch Minimum von  $M$  (die Begriffe minimal und Minimum sind bei Totalordnungen also identisch).

*Beweis.* (i) Ist  $m$  ein Minimum und  $x \trianglelefteq m$ , so folgt  $x = m$  wegen der Antisymmetrie ( $m \trianglelefteq x \wedge x \trianglelefteq m \Rightarrow x = m$ ).

(ii) Sei  $m$  ein Minimum und sei  $m'$  minimal. Da  $m$  Minimum ist, gilt  $m \trianglelefteq m'$ . Da  $m'$  minimal ist, folgt daraus  $m = m'$ .

(iii) Sei  $\trianglelefteq$  eine Totalordnung auf  $M$  und sei  $m \in M$  minimal. Zu zeigen ist  $m \trianglelefteq x$  für alle  $x \in M$ . Sei also  $x \in M$  beliebig. Bei einer Totalordnung ist  $m \trianglelefteq x$  oder  $x \trianglelefteq m$ . Im ersten Fall sind wir fertig. Im zweiten Fall folgt  $x = m$ , da  $m$  minimal ist, also  $x \trianglelefteq m$  wegen der Reflexivität.  $\square$

**Bemerkung b.** Jede nicht-leere Teilmenge von  $\mathbb{N}$  hat bzgl. der Ordnung  $\leq$  ein Minimum (auf den Beweis verzichten wir hier). Man sagt die Ordnung  $\leq$  auf  $\mathbb{N}$  ist eine *Wohlordnung*.

*Übung.* Jede endliche Menge mit partieller Ordnung hat ein minimales Element.

*Übung.* Formuliere die Definition, Bemerkung **a** und den Satz für *maximale Elemente* und *größte Elemente* (auch *Maxima* genannt) aus.

*Übung.* Wir können die Begriffe minimal und Minimum auch definieren, wenn die Relation keine Ordnung ist. Zeigen Sie am Beispiel der Teilbarkeitsrelation auf  $\mathbb{Z}$  (die keine Ordnung ist), dass dann der Satz nicht mehr gilt.





# Kapitel 2

## Algebraische Strukturen

### 2.1 Gruppen

#### 2.1.1 Strukturen und Verknüpfungen

**Definition.** Eine *Verknüpfung* auf einer Menge  $M$  ist eine Abbildung

$$M \times M \rightarrow M.$$

Eine *algebraische Struktur* ist eine Menge mit ein oder mehreren Verknüpfungen.

**Beispiel a.**

- (i)  $- : \mathbb{Z} \times \mathbb{Z} \rightarrow \mathbb{Z}, (x, y) \mapsto x - y$  ist eine Verknüpfung auf  $\mathbb{Z}$ .
- (ii) Für jede Menge  $N$  ist  $\circ$  eine Verknüpfung auf  $\text{Abb}(N, N)$ .
- (iii)  $\wedge$  ist eine Verknüpfung auf  $B = \{0, 1\}$  (wenn wir 0 und 1 als Wahrheitswerte definieren, und  $\wedge$  durch die zugehörige Wahrheitstafel definiert ist).
- (iv) Es sei  $A$  eine beliebige Menge und  $n \in \mathbb{N}_0$ . Sind  $a_1, \dots, a_n \in A$ , so nennen wir  $a_1 \cdots a_n$  ein *Wort* der Länge  $n$  über dem Alphabet  $A$  (formal ist  $a_1 \cdots a_n$  das  $n$ -Tupel  $(a_1, \dots, a_n)$ , wobei wir in diesem Kontext die Klammern und Kommas in der Notation der Tupel weglassen).

Es bezeichne  $\epsilon$  das Wort der Länge 0 (das leere Wort oder leere Tupel), und  $A^*$  die Menge aller Wörter über  $A$  (beliebiger Länge) einschließlich  $\epsilon$ . Dann wird durch

$$a_1 \cdots a_n * b_1 \cdots b_m := a_1 \cdots a_n b_1 \cdots b_m$$

eine Verknüpfung auf  $A^*$  definiert, die *Verkettung* oder *Konkatenation*.

**Schreibweise.** Es seien  $M$  eine Menge,  $\bullet$  eine Verknüpfung auf  $M$ ,  $m \in M$ , und  $A, B \subseteq M$ .

- (i)  $m \bullet A := \{m \bullet a \mid a \in A\} \subseteq M$
- (ii)  $A \bullet m := \{a \bullet m \mid a \in A\} \subseteq M$
- (iii)  $A \bullet B := \{a \bullet b \mid a \in A, b \in B\} \subseteq M$

**Beispiel b.**

$$7\mathbb{Z} = \{7a \mid a \in \mathbb{Z}\} = \{\dots, -14, -7, 0, 7, 14, \dots\},$$

$$2 + 7\mathbb{Z} = \{2 + 7a \mid a \in \mathbb{Z}\} = \{\dots - 12, -5, 2, 9, 16, \dots\}.$$

### 2.1.2 Monoide

**Definition a.** Es sei  $M$  eine Menge mit einer Verknüpfung

$$\bullet : M \times M \rightarrow M, (x, y) \mapsto x \bullet y.$$

Wir nennen  $(M, \bullet)$  ein *Monoid*, wenn folgende Axiome gelten:

- (G1)  $(x \bullet y) \bullet z = x \bullet (y \bullet z)$  für alle  $x, y, z \in M$ .
- (G2) Es existiert  $e \in M$  mit  $e \bullet x = x = x \bullet e$  für alle  $x \in M$ .

Das Monoid heißt *abelsch* oder *kommutativ*, wenn zusätzlich gilt:

- (G4)  $x \bullet y = y \bullet x$  für alle  $x, y \in M$ .

Man nennt (G1) das Assoziativgesetz und (G4) das Kommutativgesetz.

**Bemerkung.** Das Element  $e$  in (G2) ist eindeutig und wird das *neutrale Element* von  $M$  genannt.

*Beweis.* Sind  $e, e' \in M$  zwei Elemente wie in (G2), so gilt einerseits  $e \bullet e' = e$  und andererseits  $e \bullet e' = e'$ , also  $e = e'$ .  $\square$

**Schreibweise.**

- (i) In einem Monoid  $(M, \bullet)$  gilt  $a_1 \bullet a_2 \bullet \dots \bullet a_n := (\dots ((a_1 \bullet a_2) \bullet a_3) \bullet \dots a_n)$  (oder jede andere Klammerung).
- (ii) In einem abelschen Monoid benutzt man häufig  $+$  als Verknüpfungszeichen, schreibt 0 statt  $e$  und  $na$  ( $n \in \mathbb{N}$ ) als Abkürzung für  $\underbrace{a + a + \dots + a}_{n\text{-mal}}$ .

- (iii) Falls  $\cdot$  als Verknüpfungszeichen benutzt wird, schreibt man häufig 1 statt  $e$  und  $a^n$  ( $n \in \mathbb{N}$ ) als Abkürzung für  $\underbrace{a \cdot a \cdot \dots \cdot a}_{n\text{-mal}}$ .

**Beispiel a.** Es sei  $A$  eine beliebige Menge,  $B := \{0, 1\}$ .

- (i)  $(\mathbb{N}, +)$  ist kein Monoid, da (G2) nicht gilt.
- (ii)  $(\mathbb{Z}, -)$  ist kein Monoid, da (G1) nicht gilt.
- (iii)  $(\mathbb{N}_0, +)$  ist ein abelsches Monoid mit neutralem Element 0.
- (iv)  $(\mathbb{R}, \cdot)$  ist ein abelsches Monoid mit neutralem Element 1.
- (v) Für jede nicht-leere Menge  $A$  ist  $(\text{Abb}(A, A), \circ)$  ein Monoid mit neutralem Element  $\text{id}_A$ .
- (vi)  $(B, \wedge)$  ist ein abelsches Monoid mit neutralem Element 1.
- (vii)  $(B, \vee), (B, \text{xor})$  sind abelsche Monoide mit neutralem Element 0.
- (viii)  $(B, \Rightarrow)$  ist kein Monoid, da (G1) nicht gilt. (man prüfe nach, dass z.B.  $(0 \Rightarrow 0) \Rightarrow 0$  ungleich  $0 \Rightarrow (0 \Rightarrow 0)$  ist).
- (ix)  $(A^*, *)$  ist Monoid mit neutralem Element  $\epsilon$ .

*Übung.* Es sei  $A$  eine nicht-leere Menge. Man zeige, dass  $(\text{Abb}(A, A), \circ)$  genau dann abelsch ist wenn  $|A| = 1$  ist.

### 2.1.3 Inverse und Einheiten

**Definition.** Es seien  $(M, \bullet)$  ein Monoid mit neutralem Element  $e$  und  $a \in M$ .

- (i) Gibt es  $b \in M$  mit  $a \bullet b = e$ , so heißt  $a$  *rechtsinvertierbar* und  $b$  *rechtsinvers* zu  $a$  bzw.  $b$  ein *Rechtsinverses* von  $a$ .
- (ii) Gibt es  $b \in M$  mit  $b \bullet a = e$ , so heißt  $a$  *linksinvertierbar* und  $b$  *linksinvers* zu  $a$  bzw.  $b$  ein *Linksinverses* von  $a$ .
- (iii) Ist  $a$  sowohl links- als auch rechtsinvertierbar, so heißt  $a$  eine *Einheit*.
- (iv) Gibt es  $b \in M$  mit  $b \bullet a = e = a \bullet b$ , so heißt  $a$  *invertierbar* und  $b$  *invers* zu  $a$  bzw.  $b$  ein *Inverses* von  $a$ .

**Bemerkung.** Es seien  $(M, \bullet)$  ein Monoid und  $a \in M$ . Dann ist  $a$  genau dann eine Einheit, wenn  $a$  invertierbar ist. In diesem Fall ist jedes Linksinverse von  $a$  auch Rechtsinverses, und umgekehrt. Weiter ist das Inverse von  $a$  eindeutig durch  $a$  bestimmt und wird mit  $a^{-1}$  bezeichnet. Wir bezeichnen die Menge der Einheiten von  $M$  mit  $M^\times$ .

*Beweis.* Per Definition ist jedes invertierbare Element eine Einheit. Sei umgekehrt  $a$  eine Einheit, etwa  $b, b' \in M$  mit  $b \bullet a = e$  und  $a \bullet b' = e$ . Dann folgt  $b = b \bullet e = b \bullet (a \bullet b') = (b \bullet a) \bullet b' = e \bullet b' = b'$ . Also ist  $b = b'$  und somit  $a$  invertierbar. Mit  $b = b'$  sind auch alle weiteren Aussagen der Bemerkung gezeigt.  $\square$

**Beispiel.** Es sei  $A$  eine nicht-leere Menge. Wir betrachten ein Element  $f : A \rightarrow A$  des Monoids  $(\text{Abb}(A, A), \circ)$ .

- (i)  $f$  ist genau dann rechtsinvertierbar, wenn  $f$  surjektiv ist.
- (ii)  $f$  ist genau dann linksinvertierbar, wenn  $f$  injektiv ist.
- (iii)  $f$  ist genau dann invertierbar, wenn  $f$  bijektiv ist.

*Übung a.* Es seien  $(M, \bullet)$  ein Monoid,  $a \in M$ , und  $m_a$  die Abbildung

$$m_a : M \rightarrow M, x \mapsto a \bullet x.$$

Man zeige:

- (i)  $m_a$  ist genau dann surjektiv, wenn  $a$  rechtsinvertierbar ist.
- (ii) Ist  $a$  linksinvertierbar, so ist  $m_a$  injektiv.

Man gebe ein Beispiel dafür an, dass die Umkehrung von (ii) nicht gilt.

*Übung b.* Es seien  $(M, \bullet)$  ein Monoid,  $a, a' \in M^\times$  und  $c \in M$ . Man zeige:

- (i) Es gilt  $a^{-1} \in M^\times$  und  $(a^{-1})^{-1} = a$ .
- (ii) Es gilt  $a \bullet a' \in M^\times$  und  $(a \bullet a')^{-1} = a'^{-1} \bullet a^{-1}$ .
- (iii) Die Gleichung  $a \bullet x = c$  hat eine eindeutige Lösung  $x \in M$ .
- (iv) Die Gleichung  $x \bullet a = c$  hat eine eindeutige Lösung  $x \in M$ .
- (v) Aus  $a \bullet c = e$  folgt  $c = a^{-1}$ .
- (vi) Aus  $c \bullet a = e$  folgt  $c = a^{-1}$ .

### 2.1.4 Gruppen

**Definition a.** Ein Monoid  $(G, \bullet)$ , in dem alle Elemente invertierbar sind, heißt *Gruppe*. D.h. in einer Gruppe gilt:

(G3) Für alle  $x \in G$  existiert  $x' \in G$  mit  $x \bullet x' = e = x' \bullet x$ .

**Beispiel a.**

- (i)  $(\mathbb{Z}, +)$  ist eine abelsche Gruppe.
- (ii)  $(\mathbb{N}_0, +)$  ist keine Gruppe, da (G3) nicht gilt.
- (iii)  $(\mathbb{R}, \cdot)$  ist keine Gruppe, da (G3) nicht gilt.
- (iv)  $(\mathbb{R} \setminus \{0\}, \cdot)$  und  $(\mathbb{R}_{>0}, \cdot)$  sind abelsche Gruppen.
- (v)  $(\mathbb{Z} \setminus \{0\}, \cdot)$  und  $(\mathbb{N}, \cdot)$  sind keine Gruppen.
- (vi) Sei  $A$  eine nicht-leere Menge und

$$S_A := \{f \in \text{Abb}(A, A) \mid f \text{ ist invertierbar}\}.$$

Dann ist  $(S_A, \circ)$  eine Gruppe, die *symmetrische Gruppe auf  $A$* . Ist  $A = \underline{n}$  für ein  $n \in \mathbb{N}$ , dann schreiben wir  $S_n := S_{\underline{n}}$  und nennen  $S_n$  die *symmetrische Gruppe auf  $n$  Ziffern*.

- (vii)  $(B, \wedge), (B, \vee)$  sind keine Gruppen.
- (viii)  $(B, \text{xor})$  ist eine Gruppe.
- (ix)  $(A^*, *)$  ist keine Gruppe, da (G3) nicht gilt.

**Schreibweise.**

- (i) In einer abelschen Gruppe benutzt man häufig  $+$  als Verknüpfungszeichen, schreibt  $-a$  für das Inverse von  $a$ , und benutzt die Abkürzungen:  $a - b := a + (-b)$ ,  $(-n)a := n(-a)$  für  $n \in \mathbb{N}$ ,  $0a := 0$ .
- (ii) Falls  $\cdot$  als Verknüpfungszeichen benutzt wird, schreibt man  $a^{-1}$  für das Inverse von  $a$ ,  $1$  statt  $e$ , lässt  $\cdot$  einfach weg, und benutzt die Abkürzungen:  $a^{-n} := (a^{-1})^n$  für  $n \in \mathbb{N}$ ,  $a^0 := 1$ . Falls die Gruppe abelsch ist, kann man auch  $a/b$  für  $ab^{-1}$  schreiben.

**Bemerkung.** Ist  $(M, \bullet)$  ein Monoid, so ist  $(M^\times, \bullet)$  eine Gruppe. Die Gruppe  $(M^\times, \bullet)$  wird *Einheitengruppe* von  $M$  genannt.

*Beweis.* Zu zeigen ist, dass  $\bullet$  eine Verknüpfung auf  $M^\times$  ist, und dass für  $a \in M^\times$  auch das Inverse von  $a$  in  $M^\times$  liegt. Beides wurde in Übung 2.1.3b gezeigt.  $\square$

**Satz.** Es sei  $(G, \cdot)$  eine Gruppe und  $a, b \in G$ .

- (i) Für alle  $c \in G$  gilt:  $a = b \Leftrightarrow a \cdot c = b \cdot c$  und  $a = b \Leftrightarrow c \cdot a = c \cdot b$ . („Multiplikation“ von links oder rechts in einer Gruppe ist eine Äquivalenzumformung.)
- (ii) Die Gleichung  $a \cdot x = b$  hat eine eindeutige Lösung  $x \in G$  (ebenso die Gleichung  $x \cdot a = b$ ).

*Beweis.*

- (i) Die Implikation  $a = b \Rightarrow a \cdot c = b \cdot c$  ist trivial. Damit folgt aber auch  $a \cdot c = b \cdot c \Rightarrow (a \cdot c) \cdot c^{-1} = (b \cdot c) \cdot c^{-1}$ , und die rechte Seite lautet  $a = b$ . Die Äquivalenz  $a = b \Leftrightarrow c \cdot a = c \cdot b$  verläuft entsprechend mit Multiplikation von  $c^{-1}$  auf der linken Seite.
- (ii) Nach (i) gilt  $a \cdot x = b$  genau dann, wenn  $x = a^{-1} \cdot (a \cdot x) = a^{-1} \cdot b$  ist. Entsprechend gilt  $x \cdot a = b$  genau dann, wenn  $x = (x \cdot a) \cdot a^{-1} = b \cdot a^{-1}$  ist.

$\square$

*Übung a.* Bestimmen Sie zu allen Beispielen von Monoiden und Gruppen die neutralen bzw. inversen Elemente.

*Übung b.* Es sei  $A$  eine nicht-leere endliche Menge. Man zeige, dass  $S_A$  genau dann abelsch ist, wenn  $|A| \leq 2$ .

*Übung c.* Es seien  $(G, \cdot)$  eine Gruppe und  $a \in G$ . Ist die Abbildung  $\lambda_a : G \rightarrow G, x \mapsto a \cdot x$  injektiv, surjektiv, bijektiv?

### 2.1.5 Untergruppen

**Definition.** Es sei  $(G, \cdot)$  eine Gruppe. Eine Teilmenge  $H \subseteq G$  heißt *Untergruppe von  $G$* , geschrieben  $H \leq G$ , wenn gilt:

- (U1)  $e \in H$ .
- (U2) Für alle  $x, y \in H$  ist auch  $x \cdot y^{-1} \in H$ . (Wir sagen:  $H$  ist *abgeschlossen* bzgl.  $\cdot$  und Invertieren.)

In diesem Fall ist  $H$  selbst eine Gruppe bzgl. der Verknüpfung  $\cdot$  aus  $G$ .

**Beispiel.**

- (i) Für jedes  $n \in \mathbb{N}_0$  ist  $n\mathbb{Z} := \{nz \mid z \in \mathbb{Z}\}$  eine Untergruppe von  $(\mathbb{Z}, +)$  (z.B.  $3\mathbb{Z} = \{\dots, -9, -6, -3, 0, 3, 6, 9, \dots\}$ ).
- (ii)  $\mathbb{N}$  ist keine Untergruppe von  $(\mathbb{Z}, +)$ .
- (iii)  $H := \{\pi \in S_n \mid \pi(n) = n\}$  ist eine Untergruppe von  $(S_n, \circ)$ .
- (iv)  $\mathbb{Q}_{>0}$  ist eine Untergruppe von  $(\mathbb{R}_{>0}, \cdot)$ .
- (v)  $\mathbb{N}$  ist keine Untergruppe von  $(\mathbb{R}_{>0}, \cdot)$ .

*Beweis.*

- (i) (U1):  $e = 0 = n \cdot 0 \in n\mathbb{Z}$ .  
(U2):  $nx - ny = n(x - y) \in n\mathbb{Z}$ .
- (ii) (U1) gilt nicht, denn  $e = 0 \notin \mathbb{N}$ .
- (iii) (U1):  $e = \text{id}_n$  lässt  $n$  fest, also  $\text{id}_n \in H$ .  
(U2): Seien  $\sigma, \pi \in H$ , d.h.  $\sigma(n) = n$  und  $\pi(n) = n$ . Aus  $\pi(n) = n$  folgt  $\pi^{-1}(n) = n$ . Weiter ergibt sich  $\sigma \circ \pi^{-1}(n) = \sigma(\pi^{-1}(n)) = \sigma(n) = n$ , d.h.  $\sigma \circ \pi^{-1} \in H$ .
- (iv) (U1):  $e = 1 \in \mathbb{Q}_{>0}$ .  
(U2): Sind  $x, y \in \mathbb{Q}_{>0}$ , so ist auch  $xy^{-1} \in \mathbb{Q}_{>0}$ .
- (v) (U2) gilt nicht, da z.B.  $2^{-1} \notin \mathbb{N}$ .

□

**2.1.6 Kartesische Produkte**

**Satz.** Es seien  $(G, \cdot)$  eine Gruppe und  $M$  eine Menge. Die Menge  $\text{Abb}(M, G) = \{f : M \rightarrow G\}$  wird zu einer Gruppe  $(\text{Abb}(M, G), \bullet)$ , wenn man die Verknüpfung

$$\bullet : \text{Abb}(M, G) \times \text{Abb}(M, G) \rightarrow \text{Abb}(M, G), (f, g) \mapsto f \bullet g$$

*durch*

$$(f \bullet g)(x) := f(x) \cdot g(x) \text{ für alle } x \in M$$

definiert. Da  $\bullet$  durch  $\cdot$  definiert ist schreibt man in der Regel  $(\text{Abb}(M, G), \cdot)$ . Ist  $(G, \cdot)$  abelsch, so ist auch  $(\text{Abb}(M, G), \cdot)$  abelsch.

**Beispiel.** Es sei  $(G, \cdot)$  eine Gruppe. Die Gruppe  $(G^n, \cdot)$  ist dann die Menge

$$G^n = \{n\text{-Tupel über } G\} = \{(a_1, \dots, a_n) \mid a_i \in G\}$$

mit *komponentenweiser* Verknüpfung, d.h.

$$(a_1, \dots, a_n) \cdot (b_1, \dots, b_n) := (a_1 \cdot b_1, \dots, a_n \cdot b_n).$$

$G^n$  wird das *n-fache kartesische Produkt* von  $G$  genannt.

*Übung.* Es seien  $(G, \bullet)$  und  $(G', \circ)$  zwei Gruppen. Man zeige, dass die Menge  $G \times G'$  mit der Verknüpfung

$$(g_1, g'_1) \cdot (g_2, g'_2) := (g_1 \bullet g_2, g'_1 \circ g'_2)$$

wieder eine Gruppe ist.

## 2.2 Ringe

### 2.2.1 Definition und Beispiele

In  $\mathbb{Z}, \mathbb{Q}, \mathbb{R}$  und  $\mathbb{C}$  gibt es zwei Verknüpfungen  $+$  und  $\cdot$ , die mittels der Distributivgesetze miteinander verbunden sind. Die entsprechende Abstraktion der Rechenregeln führt zu den Begriffen Ring und Körper.

**Definition.** Eine Menge  $R$  mit zwei Verknüpfungen

$$+ : R \times R \rightarrow R, \quad \text{und} \quad \cdot : R \times R \rightarrow R$$

heißt *Ring*, wenn folgende Bedingungen erfüllt sind:

(R1)  $(R, +)$  ist eine abelsche Gruppe.

(R2)  $(R, \cdot)$  ist ein Monoid.

(R3)  $x \cdot (y + z) = x \cdot y + x \cdot z$  und  $(x + y) \cdot z = x \cdot z + y \cdot z$  für alle  $x, y, z \in R$ .

Der Ring heißt *kommutativ*, wenn zusätzlich gilt:

(R4)  $x \cdot y = y \cdot x$  für alle  $x, y \in R$ .

**Beispiel.**

(i)  $(\mathbb{Z}, +, \cdot)$  ist ein kommutativer Ring.

(ii)  $R = \{0\}$  mit  $0 + 0 := 0$  und  $0 \cdot 0 := 0$  bildet den trivialen Ring.



- (iii) Nicht-kommutative Ringe begegnen uns in der Linearen Algebra, z.B. der „Matrizenring“ und der „Endomorphismenring“.

**Bemerkung.** Die Gleichungen aus (R3) heißen *Distributivgesetze*. Man vereinbart in einem Ring, dass  $\cdot$  stärker bindet als  $+$ , d.h.  $a \cdot b + c$  steht für  $(a \cdot b) + c$ , und  $a + b \cdot c$  für  $a + (b \cdot c)$ . (Punktrechnung geht vor Strichrechnung.) Dies spart Klammern und wurde in obiger Formulierung von (R3) bereits benutzt! Ferner wird vereinbart, dass  $\cdot$  weggelassen werden kann, d.h.  $ab$  steht für  $a \cdot b$ . Das neutrale Element der Gruppe  $(R, +)$  wird mit 0 bezeichnet und *Nullelement* bzw. kurz *Null* von  $R$  genannt. Das neutrale Element des Monoids  $(R, \cdot)$  wird mit 1 bezeichnet und *Einselement* bzw. kurz *Eins* von  $R$  genannt. Wir nennen  $-a$  das *additive Inverse* oder *negative Element* von  $a$ .

*Übung a.* Es sei  $R$  ein Ring. Man zeige:

- (i)  $0 \cdot a = a \cdot 0 = 0$  für alle  $a \in R$ .
- (ii)  $-a = (-1) \cdot a$  und  $a = (-1) \cdot (-a)$  für alle  $a \in R$ .
- (iii)  $a \cdot (-b) = (-a) \cdot b = -(a \cdot b)$  für alle  $a, b \in R$ .

*Beweis.* Aus  $0 = 0 + 0$  und dem Distributivgesetz folgt  $0 \cdot a = (0 + 0) \cdot a = 0 \cdot a + 0 \cdot a$ . Addition von  $-(0 \cdot a)$  auf beiden Seiten liefert  $0 = 0 \cdot a$ .

$$a + (-1) \cdot a = 1 \cdot a + (-1) \cdot a = (1 + (-1)) \cdot a = 0 \cdot a = 0. \quad \square$$

*Übung b.* Es sei  $R$  ein Ring. Für jedes  $n \in \mathbb{N}$  und  $a \in R$  definieren wir

$$na := \underbrace{a + \dots + a}_{n\text{-mal}}.$$

Man zeige:  $-(na) = n(-a)$ . Wie definiert man sinnvoll  $na$  für alle  $n \in \mathbb{Z}$ ?

*Übung c.* Man zeige: Ist  $R$  ein Ring mit  $1 = 0$ , so ist  $R = \{0\}$ .

*Übung d.* Es seien  $R, S$  zwei Ringe,  $n \in \mathbb{N}$  und  $M$  eine Menge. Wie sind die Verknüpfungen zu definieren, mit denen auch  $R \times S, R^n$  und  $\text{Abb}(M, R)$  zu einem Ring werden?

## 2.2.2 Einheitengruppe

**Definition.** Es sei  $R$  ein Ring. Die Begriffe *invertierbar*, *Einheit*, *Einheitengruppe* und die Notation  $R^\times$  beziehen sich auf das Monoid  $(R, \cdot)$ .

**Beispiel.**

- (i)  $\mathbb{Z}^\times = \{1, -1\}$ .
- (ii) In jedem Ring  $R$  ist  $1, -1 \in R^\times$ . (1 und  $-1$  können aber gleich sein, wie wir an den Beispielen  $\mathbb{F}_2$  und  $\mathbb{F}_4$  unten sehen werden.)

*Übung a.* Es seien  $R$  kommutativ,  $a \in R$ , und  $m_a$  bezeichne die Abbildung  $m_a : R \rightarrow R, x \mapsto ax$ . Man zeige die Äquivalenz folgender Aussagen:

- (i)  $a$  ist Einheit.
- (ii)  $m_a$  ist bijektiv.
- (iii) Die Gleichung  $ax = b$  ist für alle  $b \in R$  eindeutig lösbar.

Insbesondere gilt für jedes  $a \in R^\times$ :  $ax = 0 \Rightarrow x = 0$ .

*Übung b.* Es seien  $R$  kommutativ,  $a, b \in R$ . Man zeige:  $ab \in R^\times \Leftrightarrow a \in R^\times \wedge b \in R^\times$ . Hieraus folgt, dass auch  $R \setminus R^\times$  unter der Multiplikation abgeschlossen ist.

### 2.2.3 Nullteiler

Es sei  $R$  ein kommutativer Ring.

**Definition.** Ein Element  $a \in R$  heißt *Nullteiler* von  $R$ , wenn  $b \in R \setminus \{0\}$  existiert mit  $ab = 0$ . Der Ring  $R$  heißt *nullteilerfrei*, wenn er keine Nullteiler außer 0 enthält. Der Ring  $R$  heißt *Integritätsbereich*, wenn  $1 \neq 0$  und  $R$  nullteilerfrei ist.

**Bemerkung.** Sei  $a \in R$  und bezeichne  $m_a$  die Abbildung  $m_a : R \rightarrow R, x \mapsto ax$ . Dann sind äquivalent:

- (i)  $a$  ist kein Nullteiler von  $R$ .
- (ii) Für alle  $x \in R$  gilt:  $ax = 0 \Rightarrow x = 0$ .
- (iii) Für alle  $x, x' \in R$  gilt:  $ax = ax' \Rightarrow x = x'$ . (Kürzungsregel)
- (iv) Für alle  $b \in R$  hat die Gleichung  $ax = b$  höchstens eine Lösung.
- (v)  $m_a$  ist injektiv.

Insbesondere sind Einheiten keine Nullteiler (nach Übung 2.2.2 ist  $m_a$  für Einheiten bijektiv). Weiter ist  $R$  genau dann nullteilerfrei, wenn für alle  $a, b \in R$  gilt:

$$ab = 0 \Rightarrow (a = 0 \vee b = 0).$$

*Beweis.* Übung. □

**Beispiel.**  $\mathbb{Z}$ , alle Körper sowie der triviale Ring sind nullteilerfrei.

*Beweis.* als Übung. □

*Übung a.* Man zeige: 0 ist genau dann kein Nullteiler, wenn  $R$  der triviale Ring ist.

*Übung b.* Man zeige für alle  $a, b \in R$ : Ist  $a$  ein Nullteiler, so auch  $ab$ . Gilt auch die Umkehrung?

## 2.2.4 Körper

**Definition.** Ein kommutativer Ring  $R$  heißt *Körper*, wenn  $1 \neq 0$  und  $R^\times = R \setminus \{0\}$  gilt.

Ein Körper ist also ein nicht-trivialer Ring, in dem jedes von 0 verschiedene Element invertierbar ist.

**Beispiel a.**  $(\mathbb{Q}, +, \cdot)$ ,  $(\mathbb{R}, +, \cdot)$  und  $(\mathbb{C}, +, \cdot)$  sind Körper. Im Unterschied zu  $\mathbb{Q}$  erfüllt  $\mathbb{R}$  noch die „Vollständigkeitsaxiome“ und die „Anordnungsaxiome“, die man in der Analysis-Vorlesung lernt.  $(\mathbb{Z}, +, \cdot)$  ist kein Körper.

Es gibt aber auch endliche Körper.

**Beispiel b.** Definiert man auf der Menge  $\{0, 1\}$  zwei Abbildungen  $+, \cdot$  durch die *Verknüpfungstafeln*

$+$	$0$	$1$
$0$	$0$	$1$
$1$	$1$	$0$

$\cdot$	$0$	$1$
$0$	$0$	$0$
$1$	$0$	$1$

so entsteht ein Körper (man prüfe alle Axiome nach!). Wir bezeichnen diesen Körper mit  $\mathbb{F}_2$ .

Identifiziert man 0 mit „falsch“ und 1 mit „wahr“, dann stellt man außerdem fest, dass  $+$  gerade der Verknüpfung xor entspricht, und  $\cdot$  der Verknüpfung  $\wedge$ .

**Beispiel c.** Die Menge  $\mathbb{F}_4 := \{0, 1, a, b\}$  mit den Verknüpfungstafeln

$+$	$0$	$1$	$a$	$b$
$0$	$0$	$1$	$a$	$b$
$1$	$1$	$0$	$b$	$a$
$a$	$a$	$b$	$0$	$1$
$b$	$b$	$a$	$1$	$0$

$\cdot$	$0$	$1$	$a$	$b$
$0$	$0$	$0$	$0$	$0$
$1$	$0$	$1$	$a$	$b$
$a$	$0$	$a$	$b$	$1$
$b$	$0$	$b$	$1$	$a$

bildet einen Körper.

*Beweis.* Übung. □

**Bemerkung.**

- (i) Die Tafeln in Beispiel c sind bis auf Benennung der Elemente  $a, b$  eindeutig, d.h. es gibt genau einen Körper mit 4 Elementen (siehe Vorlesung oder vgl. [3], §2.2, 37-38, leicht lesbar).
- (ii) Es gibt für jede Primzahlpotenz  $p^n$  genau einen Körper mit  $p^n$  Elementen (ohne Beweis). Dieser wird mit  $\mathbb{F}_{p^n}$  bezeichnet (das  $\mathbb{F}$  steht hier für „field“, engl. für Körper). Für  $n = 1$  werden diese Körper weiter unten konstruiert:  $\mathbb{F}_p$  ist identisch mit dem dort eingeführten „Restklassenring“  $\mathbb{Z}_p$ .  
*Achtung:*  $\mathbb{F}_{p^n}$  für  $n > 1$  wird in dieser Vorlesung nicht behandelt und ist insbesondere nicht identisch mit  $\mathbb{Z}_{p^n}$ , denn  $\mathbb{Z}_{p^n}$  ist für  $n > 1$  kein Körper.
- (iii) Endliche Körper sind für die Informatik von besonderer Bedeutung, etwa in der Kodierungstheorie. Es sei daran erinnert, dass man ein Bit als Element des Körper  $\mathbb{F}_2$  auffassen kann, ein Byte als Element des Körpers  $\mathbb{F}_{256}$ , usw.

*Übung.* Sind  $K, L$  zwei Körper, so ist der Ring  $K \times L$  (mit komponentenweisen Operationen) *kein* Körper.

*Beweis.* Es gilt  $(1, 0) \cdot (0, 1) = (0, 0)$ . Nach Übung 2.2.2a ist  $(1, 0)$  keine Einheit in  $K \times L$ . □

## 2.3 Polynome

In diesem Abschnitt sei  $K$  ein Körper. Der hier eingeführte Polynomring über  $K$  ist ein besonders wichtiges Beispiel für einen Integritätsbereich.

### 2.3.1 Definition und Beispiele

**Definition.**

- (i) Ein *Polynom* über  $K$  in der *Unbestimmten*  $X$  ist ein Ausdruck der Form

$$f = \sum_{i=0}^n a_i X^i$$

mit  $a_i \in K$  für alle  $i = 0, \dots, n$ . Die  $a_i$  heißen die *Koeffizienten* des Polynoms, insbesondere heißt  $a_0$  der *konstante* oder *absolute Koeffizient*.

(Koeffizienten, die gleich 0 sind können beliebig hinzugefügt oder weggelassen werden, ohne den Ausdruck zu verändern.)

- (ii) Zwei Polynome  $f = \sum_{i=0}^n a_i X^i$  und  $g = \sum_{i=0}^n b_i X^i$  sind genau dann *gleich*, wenn  $a_i = b_i$  für alle  $i = 0, \dots, n$ .
- (iii) Die Menge aller Polynome über  $K$  wird mit  $K[X]$  bezeichnet.
- (iv) Sind alle Koeffizienten von  $f \in K[X]$  gleich 0, so heißt  $f$  das *Nullpolynom*, geschrieben  $f = 0$ .
- (v) Ist  $f \in K[X]$  nicht das Nullpolynom, dann wird das größte  $i \in \mathbb{N}_0$ , für das  $a_i \neq 0$  ist, der *Grad* von  $f$  genannt und mit  $\deg f$  bezeichnet. Für das Nullpolynom setzen wir  $\deg 0 := -\infty$ .
- (vi) Ist  $\deg f = n \geq 0$ , so heißt  $a_n$  der *Leitkoeffizient* oder *Hauptkoeffizient* von  $f$ .
- (vii) Ein Polynom heißt *normiert*, wenn der Hauptkoeffizient gleich 1 ist.
- (viii) Ein Polynom  $f$  heißt *linear*, wenn  $\deg f = 1$ .
- (ix) Ein Polynom  $f$  heißt *konstant*, wenn  $\deg f \leq 0$  ist.

**Schreibweise.** Der Kürze halber schreibt man  $X^i$  statt  $1X^i$ ,  $X$  statt  $X^1$ ,  $a_0$  statt  $a_0X^0$ , und  $0X^i$  lässt man weg.

**Beispiel.**

$$(i) \quad f = 1X^4 + 0X^3 - \frac{1}{3}X^2 + 1X^1 - 2X^0 = X^4 - \frac{1}{3}X^2 + X - 2 \in \mathbb{R}[X].$$

$$(ii) \quad g = 1X^2 + 1X^1 + 0X^0 = X^2 + X \in \mathbb{F}_2[X].$$

**Bemerkung a.** Jedes Polynom  $f \in K[X]$  definiert eine Abbildung  $K \rightarrow K$  dadurch, dass man das „Einsetzen“ in die Unbestimmte als Zuordnungsvorschrift wählt. Diese Abbildung bezeichnen wir ebenfalls mit  $f$ , sprechen aber zur Unterscheidung von der *Polynomfunktion* zu  $f$ . Für jedes  $a \in K$  nennen wir  $f(a)$  den *Wert von  $f$  an der Stelle  $a$* .

Die Polynome  $f$  und  $g$  aus dem Beispiel haben die Polynomfunktionen

$$f : \mathbb{R} \rightarrow \mathbb{R}, \quad a \mapsto f(a) = a^4 - \frac{1}{3}a^2 + a^1 - 2a^0 = a^4 - \frac{1}{3}a^2 + a - 2.$$

$$g : \mathbb{F}_2 \rightarrow \mathbb{F}_2, \quad a \mapsto g(a) = a^2 - a = 0.$$

(Man beachte, dass  $a^2 - a = 0$  für alle  $a \in \mathbb{F}_2$ .)

Verschiedene Polynome können dieselbe Polynomfunktion haben (z.B.  $g$  und das Nullpolynom). Aus diesem Grund sind Polynomfunktionen und Polynome zu unterscheiden.

**Bemerkung b.** Jedes  $a \in K$  kann als konstantes Polynom  $aX^0$  aufgefasst werden. Auf diese Weise wird  $K$  zu einer Teilmenge von  $K[X]$ .

**Bemerkung c.** Für eine mathematisch präzise Definition des Polynombegriffs betrachtet man

$$K^{(\mathbb{N}_0)} := \{(a_i)_{i \in \mathbb{N}_0} \in K^{\mathbb{N}_0} \mid a_i = 0 \text{ für fast alle } i \in \mathbb{N}_0\}.$$

Hier bedeutet *fast alle*, wie in der Analysis, *alle, bis auf endlich viele*. Eine Folge  $(a_i)_{i \in \mathbb{N}_0}$  liegt also genau dann in  $K^{(\mathbb{N}_0)}$ , wenn ein  $N \in \mathbb{N}_0$  existiert mit  $a_i = 0$  für alle  $i \geq N$ .

Das Polynom  $f = \sum_{i=0}^n a_i X^i \in K[X]$  kann durch die Folge seiner Koeffizienten

$$(a_0, a_1, a_2, \dots, a_n, 0, 0, 0, \dots) \in K^{(\mathbb{N}_0)}$$

definiert werden. Dies führt zu der Definition  $K[X] := K^{(\mathbb{N}_0)}$ .

In dieser Formulierung gilt dann für die Unbestimmte:

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

Konstante Polynome entsprechen den Folgen

$$a_0 X^0 = (a_0, 0, 0, 0, \dots), \quad a_0 \in K.$$

### 2.3.2 Der Polynomring

Für Polynome gibt es eine natürliche Addition und Multiplikation, die aus der Menge  $K[X]$  einen Ring macht.

**Definition.** Für beliebige Polynome  $f = \sum_{i=0}^n a_i X^i$  und  $g = \sum_{i=0}^m b_i X^i$  aus  $K[X]$  wird deren Summe und Produkt definiert als:

$$f + g := \sum_{i=0}^{\max\{n,m\}} (a_i + b_i) X^i,$$

$$f \cdot g := \sum_{i=0}^{n+m} c_i X^i \text{ mit } c_i := \sum_{k=0}^i a_k b_{i-k}.$$

**Bemerkung.**

- (i) Mit dieser Addition und Multiplikation wird  $K[X]$  ein kommutativer Ring. (Man prüfe die Ringaxiome nach!) Das neutrale Element der Addition ist das Nullpolynom, und das neutrale Element der Multiplikation ist das konstante Polynom  $1 = 1X^0$ .

- (ii) Man kann Polynome auch als endliche Folgen auffassen, etwa das Polynom  $-3X^2 + X + 2$  als die Folge  $(2, 1, -3, 0, 0, \dots)$ . Somit haben wir die Inklusion  $K[X] \subseteq \text{Abb}(\mathbb{N}, K)$ . Dabei stimmt die Addition in  $K[X]$  mit der punktweisen Addition in  $\text{Abb}(\mathbb{N}, K)$  überein, nicht aber die Multiplikation.
- (iii)  $K$  ist ein „Teilring“ von  $K[X]$ .
- (iv) Es gelten die Gradformeln

$$\begin{aligned}\deg(f + g) &\leq \max\{\deg f, \deg g\}, \\ \deg(f \cdot g) &= \deg f + \deg g.\end{aligned}$$

- (v)  $K[X]$  ist nullteilerfrei.
- (vi) In  $K[X]$  gilt die Kürzungsregel, d.h. für alle  $f, g, h \in K[X]$  mit  $f \neq 0$  gilt:

$$fg = fh \Rightarrow g = h.$$

- (vii) Die Einheitengruppe des Ringes  $K[X]$  lautet

$$\begin{aligned}K[X]^\times &= \{f \in K[X] \mid \deg f = 0\} = \{\text{konstante Polynome} \neq 0\} \\ &= K^\times = K \setminus \{0\}.\end{aligned}$$

*Beweis.* Übung. □

## 2.4 Teilbarkeitslehre in kommutativen Ringen

Hier definieren wir die Teilbarkeitsrelation in kommutativen Ringen und leiten einige Eigenschaften davon her. Wir werden die Ergebnisse hauptsächlich auf den Ring  $\mathbb{Z}$  der ganzen Zahlen und den Polynomring  $K[X]$  über dem Körper  $K$  anwenden.

### 2.4.1 Teilbarkeitsrelation

Es sei  $R$  ein kommutativer Ring.

**Definition a.** Es seien  $a, b \in R$ . Wir sagen  $a$  *teilt*  $b$  bzw.  $b$  *ist Vielfaches von*  $a$ , geschrieben  $a \mid b$ , wenn ein  $x \in R$  existiert mit  $ax = b$ .

**Bemerkung a.** Die Relation  $|$  auf  $R$  ist reflexiv und transitiv. Für alle  $a, b, c \in R$  und alle  $u, v \in R^\times$  gelten:

- (i)  $a | b \Rightarrow a | bc$ ,
- (ii)  $(a | b \wedge a | c) \Rightarrow a | b + c$ ,
- (iii)  $a | 0$ ,
- (iv)  $0 | a \Leftrightarrow a = 0$ ,
- (v)  $a | b \Leftrightarrow ua | vb$ .

*Beweis.* Siehe Vorlesung. □

*Übung a.* Es seien  $a, b, c \in R$ . Man zeige, dass aus zwei der folgenden Aussagen die dritte folgt:

- (i)  $a | b$
- (ii)  $a | c$
- (iii)  $a | b + c$

**Definition b.** Wir nennen  $a, b \in R$  *assoziiert*, geschrieben  $a \sim b$ , wenn ein  $u \in R^\times$  existiert mit  $au = b$ . Aus  $a \sim b$  folgt offensichtlich  $a | b$  und  $b | a$ .

Erinnerung (vgl. Abschnitt 2.2.3):  $R$  heißt Integritätsbereich, wenn  $1 \neq 0$  ist und aus  $ab = 0$  für  $a, b \in R$  folgt:  $a = 0$  oder  $b = 0$ .

**Bemerkung b.** Es sei  $R$  ein Integritätsbereich und  $a, b \in R$ . Dann sind äquivalent:

- (i)  $a | b$  und  $b | a$ ,
- (ii)  $a \sim b$ .

*Beweis.* (i)  $\Rightarrow$  (ii): Es seien  $x, y \in R$  mit  $b = ax$  und  $a = by$ . Dann ist  $b = ax = byx$ . Ausklammern von  $b$  liefert  $b(1 - yx) = 0$ . Ist  $b = 0$ , dann auch  $a = by = 0$  und es gilt  $a \sim b$ . Sei nun  $b \neq 0$ . Da  $R$  ein Integritätsbereich ist folgt  $1 - yx = 0$ , also  $yx = 1$ . Damit ist  $x \in R^\times$  und somit  $a \sim b$ .

(ii)  $\Rightarrow$  (i): Das ist Bemerkung b. □

**Beispiel.** (i) In  $\mathbb{Z}$  gilt:  $a \sim b \Leftrightarrow |a| = |b|$ . Die Relation  $|$  auf  $\mathbb{Z}$  ist also nicht antisymmetrisch.

(ii) Die Relation  $|$  auf  $\mathbb{N}$  ist eine partielle Ordnung.



- (iii) Es sei  $K$  ein Körper und  $R = K[X]$  der Polynomring über  $K$  in der Unbestimmten  $X$ . Auf der Menge der normierten Polynome aus  $K[X]$  bildet  $|$  eine partielle Ordnung. Aus  $f | g$  folgt offensichtlich  $\deg f \leq \deg g$ .

*Übung b.* (i) Man zeige, dass  $\sim$  eine Äquivalenzrelation auf  $R$  ist. Die Äquivalenzklassen  $[a]_{\sim}$  bzgl.  $\sim$  heißen die *Assoziiertenklassen* von  $R$ .

(ii) Wie sehen die Assoziiertenklassen von  $\mathbb{Z}$  aus?

(iii) Wie sehen die Assoziiertenklassen von  $K[X]$  aus?

(iv) Die Assoziiertenklasse von 1 ist  $R^\times$ .

(v) Nach Teil (v) von Bemerkung a hängt die Relation  $a | b$  nur von den Assoziiertenklassen von  $a$  und  $b$  ab. Die Relation  $|$  lässt sich also als eine Relation  $|\sim$  auf der Menge  $R/\sim$  der Assoziiertenklassen von  $R$  auffassen. Man zeige, dass  $|\sim$  reflexiv und transitiv ist.

*Übung c.* Man zeige: Ist  $b$  eine Einheit in  $R$  und  $a | b$ , so ist auch  $a$  eine Einheit.

## 2.4.2 Ideale

Es sei  $R$  ein kommutativer Ring.

**Definition.** Eine Teilmenge  $I \subseteq R$  von  $R$  heißt *Ideal* von  $R$ , falls gilt:

- (i)  $I$  ist Untergruppe der additiven Gruppe  $(R, +)$ .
- (ii)  $RI \subseteq I$ , das heißt  $ra \in I$  für alle  $r \in R$  und  $a \in I$ .

**Bemerkung.** Für Elemente  $a_1, \dots, a_k \in R$  definieren wir

$$(a_1, \dots, a_k) = \{r_1 a_1 + \dots + r_k a_k \mid r_1, \dots, r_k \in R\}.$$

Dann ist  $(a_1, \dots, a_k)$  das kleinste Ideal, das  $a_1, \dots, a_k$  enthält, und wird *das von  $a_1, \dots, a_k$  erzeugte Ideal* genannt. Ideale, die von einem Element erzeugt werden, d.h. Ideale von der Form  $(a)$ , heißen *Hauptideale*. Für alle  $a, b \in R$  gelten:

- (i)  $a | b \Leftrightarrow (a) \supseteq (b)$
- (ii)  $a \sim b \Rightarrow (a) = (b)$

*Beweis.* Als Übung. □

*Übung a.* Man zeige, dass mit zwei Idealen  $I, J \subseteq R$  auch  $I \cap J$  ein Ideal von  $R$  ist.

*Übung b.* Es sei  $R$  ein Integritätsbereich. Man zeige, dass für alle  $a, b \in R$  gilt:

$$(a) = (b) \Leftrightarrow a \sim b.$$

Es gibt also eine Bijektion zwischen  $R/\sim$  (die Menge der Assoziiertenklassen) und der Menge der Hauptideale  $\mathcal{P}$  von  $R$ , die die Relation  $|\sim$  in die partielle Ordnung  $\supseteq$  auf  $\mathcal{P}$  überführt. Insbesondere ist  $|\sim$  eine partielle Ordnung.

Frage: Gibt es einen nicht nullteilerfreien Ring, der dies erfüllt?

### 2.4.3 Division mit Rest in $\mathbb{Z}$

**Satz.** Für alle  $a, b \in \mathbb{Z}$  mit  $b \neq 0$  existieren eindeutig bestimmte  $q, r \in \mathbb{Z}$  mit  $a = qb + r$  und  $0 \leq r < |b|$ .

*Beweis.* Wegen  $a = qb + r \Leftrightarrow a = (-q)(-b) + r$  können wir oBdA  $b \geq 0$  annehmen. Eindeutigkeit: Angenommen, wir haben  $q, q', r, r' \in \mathbb{Z}$  mit  $qb + r = q'b + r'$  und  $0 \leq r, r' < b$ . Nach Annahme ist  $(q - q')b = r' - r$ , also  $b \mid r' - r$ . Ebenfalls nach Annahme ist  $0 \leq r' - r < b$ . Es folgt  $r' - r = 0$  bzw.  $r' = r$ . Da  $\mathbb{Z}$  nullteilerfrei ist und  $b \neq 0$ , folgt aus  $(q - q')b = 0$  auch  $q = q'$ .

Existenz: Wähle  $q$  maximal mit  $qb \leq a$  und setze  $r := a - qb$ . (Die Wahl von  $q$  bedeutet  $q := \lfloor a/b \rfloor$ .) Damit ist  $r \geq 0$  klar. Wir zeigen  $r < b$  mit einem Widerspruchsbeweis. Angenommen  $r \geq b$ . Dann ist  $a = r + qb \geq (q + 1)b$ . Das steht im Widerspruch zur Maximalität von  $q$ , also ist die Annahme  $r \geq b$  falsch und die Behauptung  $r < b$  bewiesen.  $\square$

**Beispiel.**  $-237 = (-12) \cdot 21 + 15$ ,  $0 \leq 15 < 21$ .

Man beachte  $(-11) \cdot 21 = -231 > -237$  und  $(-12) \cdot 21 = -252 \leq -237$ .

### 2.4.4 Division mit Rest in $K[X]$

In diesem Abschnitt sei  $K$  ein Körper. wir haben in  $K[X]$  ein analoges Ergebnis zur Division mit Rest in  $\mathbb{Z}$ , die *Polynomdivision*.

**Satz.** Es seien  $f, g \in K[X]$  mit  $g \neq 0$ . Dann existieren eindeutige  $q, r \in K[X]$  mit  $f = qg + r$  und  $\deg r < \deg g$ .

*Beweis.* Eindeutigkeit: Angenommen, wir haben  $q, q', r, r' \in K[X]$  mit  $qg + r = q'g + r'$  und  $\deg r, r' < \deg g$ . Dann ist  $(q - q')g = r' - r$ , also

$$\deg(q - q') + \deg g = \deg(r' - r) \leq \max\{\deg r', \deg r\} < \deg g.$$

Es folgt  $\deg(q - q') < 0$ , d.h.  $q - q' = 0$ . Somit ist  $q = q'$  und  $r = r'$ .

Existenz: Wir können  $\deg f \geq \deg g$  annehmen, denn sonst ist  $f = 0 \cdot g + f$  und  $\deg f < \deg g$ . Zusammen mit der Voraussetzung  $g \neq 0$  haben wir  $\deg f \geq \deg g \geq 0$  und können somit eine vollständige Induktion nach  $\deg f$  führen.

Induktionsanfang ( $\deg f = 0$ ): Dann ist auch  $\deg g = 0$ , d.h.  $f$  und  $g$  sind beide konstant und ungleich 0. Für  $f = a_0$  und  $g = b_0$  mit  $a_0, b_0 \in K \setminus \{0\}$  gilt aber  $f = \frac{a_0}{b_0} \cdot g + 0$  und  $\deg 0 = -\infty < 0 = \deg g$ . Damit ist der Induktionsanfang erledigt.

Induktionsschritt: Sei jetzt  $\deg f = n > 0$  und sei die Existenz von  $q$  und  $r$  für alle  $f$  mit  $\deg f < n$  bereits bewiesen (Ind.Vor.). Es seien  $f = \sum_{i=0}^n a_i X^i$  und  $g = \sum_{i=0}^m b_i X^i$  mit  $a_n, b_m \neq 0$  und  $m \leq n$ . Setzt man

$$f' := f - \frac{a_n}{b_m} X^{n-m} g,$$

so ist  $\deg f' < n$ . Nach Induktionsvoraussetzung gibt es  $q', r \in K[X]$  mit  $f' = q'g + r$  und  $\deg r < \deg g$ . Es folgt  $f = (\frac{a_n}{b_m} X^{n-m} + q')g + r$ , d.h.  $q := \frac{a_n}{b_m} X^{n-m} + q'$  und  $r$  sind wie gewünscht.  $\square$

**Bemerkung.** In der Formulierung und im Beweis des Satzes benutzen wir die Konvention  $-\infty = \deg 0 < \deg h$  für alle  $0 \neq h \in K[X]$ . Wem diese Konvention missfällt, darf stattdessen die Aussage wie folgt lesen: *Dann existieren eindeutig bestimmte  $q, r \in K[X]$  mit  $f = qg + r$  und  $r = 0$  oder  $r \neq 0$  und  $\deg r < \deg g$ .*

**Beispiel.**  $f = 2X^3 - 9X^2 + 4X, g = X^2 - 3X - 4 \in \mathbb{Q}[X]$ . Wir dividieren  $f$  durch  $g$  mit Rest:

$$\begin{array}{r} (2X^3 \quad -9X^2 + 4X) : (X^2 - 3X - 4) = 2X - 3 \\ -(2X^3 \quad -6X^2 - 8X) \\ \hline \quad -3X^2 + 12X \\ \quad -(-3X^2 + 9X + 12) \\ \hline \qquad \qquad 3X - 12 \end{array}$$

Also

$$f = \underbrace{(2X - 3)}_q \cdot g + \underbrace{3X - 12}_r, \quad \deg r = 1 < 2 = \deg g.$$

### 2.4.5 Nullstellen

Wie im vorigen Abschnitt sei  $K$  ein Körper. Bevor wir die Theorie weiterentwickeln, führen wir den wichtigen Begriff des Ringhomomorphismus ein.

**Definition a.** Es seien  $R$  und  $S$  zwei kommutative Ringe. Eine Abbildung  $\varphi : R \rightarrow S$  heißt *Ringhomomorphismus*, wenn gilt:

- (i)  $\varphi(r + r') = \varphi(r) + \varphi(r')$  für alle  $r, r' \in R$ ;
- (ii)  $\varphi(rr') = \varphi(r)\varphi(r')$  für alle  $r, r' \in R$ ;
- (iii)  $\varphi(1) = 1$ .

Ein wichtiger Ringhomomorphismus ist das Einsetzen eines Körperelements in Polynome.

**Bemerkung a.** Es sei  $x \in K$  fest. Wir betrachten die Abbildung

$$\tau_x : K[X] \rightarrow K, \quad f \mapsto f(x).$$

Jedem Polynom  $f \in K[X]$  wird also der Wert der durch  $f$  definierten Polynomfunktion an der Stelle  $x$  zugeordnet. Dann ist  $\tau_x$  ein Ringhomomorphismus, der *Einsetzungshomomorphismus* zu  $x$ .

*Beweis.* Wir weisen die Bedingungen aus Definition a für  $\tau_x$  nach. Dazu seien  $f = \sum_{i=0}^n a_i X^i$  und  $g = \sum_{i=0}^n b_i X^i$  in  $K[X]$ .

- (i) Es ist zu zeigen:  $\tau_x(f + g) = \tau_x(f) + \tau_x(g)$ . Dies ist äquivalent zu  $(f + g)(x) = f(x) + g(x)$ . Es ist  $f + g = \sum_{i=0}^n (a_i + b_i) X^i$ , also

$$\begin{aligned} (f + g)(x) &= \sum_{i=0}^n (a_i + b_i) x^i \\ &= \sum_{i=0}^n a_i x^i + \sum_{i=0}^n b_i x^i \\ &= f(x) + g(x). \end{aligned}$$

- (ii) Es ist zu zeigen:  $\tau_x(fg) = \tau_x(f)\tau_x(g)$ . Dies ist äquivalent zu  $(fg)(x) = f(x)g(x)$ . Es ist  $fg = \sum_{i=0}^{2n} c_i X^i$ , mit  $c_k = \sum_{i=0}^k a_i b_{k-i}$ . Es gilt:

$$\begin{aligned} f(x)g(x) &= \left( \sum_{i=0}^n a_i x^i \right) \cdot \left( \sum_{i=0}^n b_i x^i \right) \\ &= \sum_{i=0}^{2n} c_i x^i \\ &= fg(x), \end{aligned}$$

wobei die zweite Gleichung aus dem Distributivgesetz in  $K$  folgt.

- (iii) Das Einselement in  $K[X]$  ist das konstante Polynom  $1 = 1X^0$ , also ist  $\tau_x(1) = 1$ .

□

**Definition b.** Wir sagen  $a \in K$  ist *Nullstelle* eines Polynoms  $f \in K[X]$ , wenn  $f(a) = 0$  gilt, d.h. wenn die durch  $f$  definierte Polynomfunktion an der Stelle  $a$  den Wert 0 hat.

**Satz.** Es seien  $f \in K[X]$  und  $a \in K$ . Dann gilt:

$$f(a) = 0 \Leftrightarrow X - a \text{ teilt } f.$$

*Beweis.*  $\Leftarrow$ : Es sei  $f = (X - a) \cdot g$  mit  $g \in K[X]$ . Da  $\tau_a$  (das Einsetzen von  $a$ ) ein Homomorphismus ist, folgt  $f(a) = (a - a) \cdot g(a) = 0$ .

$\Rightarrow$ : Es sei  $f(a) = 0$ . Nach Polynomdivision gibt es eindeutig bestimmte  $q, r \in K[X]$  mit  $f = q \cdot (X - a) + r$  und  $\deg r < \deg(X - a) = 1$ . Das bedeutet, dass  $r$  konstant ist, also  $r = r_0 \in K$ . Da  $\tau_a$  ein Homomorphismus ist, folgt  $0 = f(a) = q(a)(a - a) + r(a) = q(a) \cdot 0 + r(a) = r_0$ . Somit ist  $r$  das Nullpolynom und  $f = (X - a) \cdot q$ . □

**Definition c.** Es seien  $0 \neq f \in K[X]$  und  $a \in K$ . Die Teiler von  $f$  der Form  $X - a$  werden *Linearfaktoren* von  $f$  genannt. Weiter heißt

$$\max\{n \in \mathbb{N}_0 \mid (X - a)^n \text{ teilt } f\}$$

die *Vielfachheit* von  $a$  als Nullstelle von  $f$ .

**Bemerkung b.** Wegen der Gradformel aus Bemerkung (2.3.2) ist die Vielfachheit stets  $\leq \deg f$ , also insbesondere endlich. Der Satz besagt, dass  $a$  genau dann Nullstelle von  $f \neq 0$  ist, wenn  $a$  Vielfachheit  $\geq 1$  hat.

### 2.4.6 Zerlegung in Linearfaktoren

Weiterhin sei  $K$  ein Körper.

**Satz.** Es sei  $0 \neq f \in K[X]$ . Sind  $a_1, \dots, a_l$  paarweise verschiedene Nullstellen von  $f$  mit den Vielfachheiten  $n_1, \dots, n_l$ , so gilt

$$f = (X - a_1)^{n_1} \cdots (X - a_l)^{n_l} \cdot g \tag{2.1}$$

für ein  $0 \neq g \in K[X]$  mit  $g(a_1), \dots, g(a_l) \neq 0$ .

*Beweis.* Wenn  $f$  die Zerlegung (2.1) hat, dann folgt  $g(a_i) \neq 0$  aus der Maximalität der  $n_i$ . In der Tat, falls  $g(a_i) = 0$  dann würde  $g$  nach Satz 2.4.5 von  $X - a_i$  geteilt werden, woraus  $(X - a_i)^{n_i+1} \mid f$  folgt.

Wir zeigen nun per Induktion nach  $l$ , dass die Zerlegung (2.1) existiert. Für  $l = 1$  folgt das aus der Definition der Vielfachheit. Sei also  $l > 1$  und die Behauptung für  $l - 1$  bereits bewiesen. Dann gibt es  $0 \neq g \in K[X]$  mit  $f = (X - a_1)^{n_1} \cdots (X - a_{l-1})^{n_{l-1}} \cdot g$ . Setzen wir  $h := (X - a_1)^{n_1} \cdots (X - a_{l-1})^{n_{l-1}}$ , so erhalten wir  $f = hg$ . Da die  $a_1, \dots, a_l$  paarweise verschieden sind, ist  $h(a_l) = (a_l - a_1)^{n_1} \cdots (a_l - a_{l-1})^{n_{l-1}} \neq 0$ . Nach Voraussetzung gilt  $(X - a_l)^{n_l} \mid f = hg$ . Das folgende Lemma zeigt, dass dann  $g$  von  $(X - a_l)^{n_l}$  geteilt wird. Damit ist die Behauptung bewiesen.  $\square$

**Lemma.** *Es seien  $g, h \in K[X]$ ,  $a \in K$  und  $n \in \mathbb{N}$ . Aus  $(X - a)^n \mid hg$  und  $h(a) \neq 0$  folgt  $(X - a)^n \mid g$ .*

*Beweis.* Induktion nach  $n$ . Sei  $n = 1$ : Wegen  $X - a \mid hg$  gilt  $h(a)g(a) = (hg)(a) = 0$ , also  $g(a) = 0$  denn  $h(a) \neq 0$ . Nach Satz 2.4.5 bedeutet das  $X - a \mid g$ .

Sei nun  $n > 1$  und die Behauptung für  $n - 1$  bereits bewiesen. Sei  $(X - a)^n \mid hg$ . Da insbesondere  $X - a \mid hg$ , so folgt nach der Überlegung für  $n = 1$ , dass  $X - a \mid g$ . Sei  $g = (X - a) \cdot g'$ , also  $(X - a)^n \mid hg = (X - a)hg'$ . Mit der Kürzungsregel in  $K[X]$  folgt  $(X - a)^{n-1} \mid hg'$ , und daraus nach Induktionsvoraussetzung  $(X - a)^{n-1} \mid g'$ . Insgesamt also  $(X - a)^n \mid g$ .  $\square$

**Folgerung.** *Es sei  $0 \neq f \in K[X]$ . Sind  $a_1, \dots, a_l$  paarweise verschiedene Nullstellen von  $f$  mit den Vielfachheiten  $n_1, \dots, n_l$ , so gilt  $\sum_{i=1}^l n_i \leq \deg f$ .*

*Das heißt, jedes Polynom  $f$  hat höchstens  $\deg f$  viele Nullstellen, wenn jede Nullstelle mit ihrer Vielfachheit gezählt wird.*

*Beweis.* Folgt sofort aus dem Satz.  $\square$

**Definition.** Es sei  $0 \neq f \in K[X]$ . Wir sagen  $f$  zerfällt vollständig in Linearfaktoren (über  $K$ ), wenn es paarweise verschiedenen Nullstellen  $a_1, \dots, a_l$  gibt, deren Vielfachheiten  $\sum_{i=1}^l n_i = \deg f$  erfüllen. Das ist genau dann der Fall, wenn es eine Zerlegung

$$f = c(X - a_1)^{n_1} \cdots (X - a_l)^{n_l}$$

gibt mit  $c \in K$  konstant.

### 2.4.7 Fundamentalsatz der Algebra

**Satz.** *Jedes Polynom  $f \in \mathbb{C}[X]$  zerfällt vollständig in Linearfaktoren.*

**Beispiel.**

$$\begin{aligned} f(X) &= X^4 - 1 = (X^2 - 1)(X^2 + 1) \\ &= (X + 1)(X - 1)(X^2 + 1) \\ &= (X + 1)(X - 1)(X - i)(X + i) \end{aligned}$$

**Folgerung.** Jedes Polynom  $f \in \mathbb{R}[X]$  besitzt eine Zerlegung  $f = f_1 \dots f_l$  mit allen  $f_i \in \mathbb{R}[X]$  und  $\deg f_i \leq 2$ .

*Beweis.* Für  $z = a + bi \in \mathbb{C}$  mit  $a, b \in \mathbb{R}$  heißt  $\bar{z} = a - bi$  das *konjugierte Element* zu  $z$ . Offensichtlich gilt  $\bar{\bar{z}} = z$  genau dann, wenn  $z \in \mathbb{R}$ . Da die Abbildung  $\mathbb{C} \rightarrow \mathbb{C}, z \mapsto \bar{z}$  ein Ringisomorphismus ist (man prüfe das nach!), gilt  $f(\bar{z}) = \overline{f(z)}$ . Folglich ist  $f(z) = 0 \Leftrightarrow f(\bar{z}) = 0$ . Die komplexen (nicht-reellen) Nullstellen treten also in Paaren, bestehend aus  $z$  und  $\bar{z}$ , auf. Somit hat  $f$  nach dem Fundamentalsatz eine Zerlegung der Form

$$f = c(X - a_1) \cdots (X - a_r)(X - z_1)(X - \bar{z}_1) \cdots (X - z_s)(X - \bar{z}_s)$$

mit  $a_1, \dots, a_r \in \mathbb{R}, z_1, \dots, z_s \in \mathbb{C} \setminus \mathbb{R}, c \in \mathbb{C}$  und  $r + 2s = \deg f$ . Man sagt, das Polynom  $f$  hat  $r$  reelle Nullstellen und  $s$  Paare komplex-konjugierter Nullstellen. Die Behauptung folgt nun, weil

$$(X - z)(X - \bar{z}) = X^2 - (z + \bar{z})X + z\bar{z} \in \mathbb{R}[X].$$

(Man prüfe nach, dass  $z + \bar{z}$  und  $z\bar{z}$  tatsächlich reell sind!) □

## 2.5 Der Euklidische Algorithmus

In diesem Abschnitt sei  $K$  ein Körper und  $R$  einer der beiden folgenden Ringe:  $R = \mathbb{Z}$ , der Ring der ganzen Zahlen, oder  $R = K[X]$ , der Polynomring in der Unbestimmten  $X$  über dem Körper  $K$ . Der Ring  $R$  ist kommutativ und nullteilerfrei, also ein Integritätsbereich.

### 2.5.1 Der ggT

Wir ziehen eine erste Folgerung aus der Division mit Rest. Dazu führen wir die folgende Notation ein.

**Notation.** Sei  $R = \mathbb{Z}$  oder  $R = K[X]$ . Für  $0 \neq a \in R$  setzen wir

$$\nu(a) := \begin{cases} |a|, & \text{falls } R = \mathbb{Z} \\ \deg a, & \text{falls } R = K[X] \end{cases}$$

**Bemerkung a.** Sei  $I$  ein Ideal in  $R$ . Dann existiert  $g \in R$  mit  $I = (g) = gR$ .

*Beweis.* Ist  $I = \{0\}$ , nehme  $g = 0$ . Sei also  $I \neq \{0\}$  und  $g \in I \setminus \{0\}$  mit  $\nu(g)$  minimal unter allen Elementen aus  $I \setminus \{0\}$ . Sei  $f \in I$ . Wir müssen zeigen:  $f \in (g)$ , d.h.  $g \mid f$ . Dazu dividieren wir  $f$  durch  $g$  mit Rest. Aus den Sätzen 2.4.3 und 2.4.4 erhalten wir  $q, r \in R$  mit  $f = qg + r$  und  $r = 0$  oder  $\nu(r) < \nu(g)$ . Angenommen,  $r \neq 0$ . Dann ist  $r = f - qg \in I \setminus \{0\}$  und  $\nu(r) < \nu(g)$ , im Widerspruch zur Wahl von  $g$ .  $\square$

**Bemerkung b.** Integritätsbereiche, in denen jedes Ideal ein Hauptideal ist, werden *Hauptidealringe* genannt. Die Ringe  $\mathbb{Z}$  und  $K[X]$  sind also Hauptidealringe.

**Bemerkung c.** Es seien  $f, g \in R$  mit  $g \neq 0$ . Betrachte die Menge  $D$  der positiven bzw. normierten gemeinsamen Teiler von  $f$  und  $g$ , d.h.

$$D := \{d \in \mathbb{N} \mid d \text{ teilt } f \text{ und } d \text{ teilt } g\}$$

falls  $R = \mathbb{Z}$  bzw.

$$D := \{d \in K[X] \mid d \text{ teilt } f, d \text{ teilt } g \text{ und } d \text{ normiert}\}$$

falls  $R = K[X]$ . Dann hat  $D$  bzgl. der Ordnung  $\mid$  ein Maximum.

*Beweis.* Betrachte das Ideal  $(f, g) = \{\lambda f + \mu g \mid \lambda, \mu \in R\}$  (siehe Bemerkung 2.4.2). Nach Bemerkung a existiert ein  $d \in R$  mit  $(f, g) = (d)$ . Insbesondere ist  $d = \lambda f + \mu g$  mit geeigneten  $\lambda, \mu \in R$ . Wir können oBdA  $d$  als positiv bzw. normiert annehmen. Dann ist  $d \in D$  wegen  $(f) \subseteq (f, g) = (d)$  und  $(g) \subseteq (f, g) = (d)$  (siehe Bemerkung 2.4.2). Sei nun  $d' \in D$ . Aus  $d' \mid f$  und  $d' \mid g$  folgt  $d' \mid \lambda f + \mu g = d$  nach Bemerkung 2.4.1 a. Damit ist  $d$  das Maximum von  $D$ .  $\square$

**Notation.** Sei  $0 \neq f \in K[X]$ . Wir schreiben  $|f|$  für das eindeutig bestimmte normierte Polynom in der Assoziiertenklasse von  $f$ , d.h.  $|f| = a_n^{-1}f$ , falls  $a_n$  den Leitkoeffizienten von  $f$  bezeichnet. Für  $f = 0$  sei  $|f| = 0$ .

**Definition.** Seien  $f, g \in R$ . Der *größte gemeinsame Teiler*, geschrieben  $\text{ggT}(f, g)$  von  $f$  und  $g$  ist definiert durch

$$\text{ggT}(f, g) := \max D$$

mit  $D$  wie in Bemerkung c, falls  $g \neq 0$ , und

$$\text{ggT}(f, 0) := |f|,$$

falls  $g = 0$ . Wir nennen  $f$  und  $g$  *teilerfremd*, wenn  $\text{ggT}(f, g) = 1$  ist.



**Bemerkung d.** Für alle  $a, b \in R$  gilt:

- (i)  $\text{ggT}(a, b) = \text{ggT}(b, a)$ ,
- (ii)  $\text{ggT}(a, b) = \text{ggT}(|a|, |b|)$ ,
- (iii)  $\text{ggT}(a, 0) = |a|$ ,
- (iv)  $a = qb + r \Rightarrow \text{ggT}(a, b) = \text{ggT}(b, r)$ .

*Beweis.* Sei  $a = qb + r$  bzw.  $r = a - qb$ . Nach Bemerkung 2.4.1 a gilt sowohl  $d \mid a, b \Rightarrow d \mid r$  und  $d \mid b, r \Rightarrow d \mid a$ . Die gemeinsamen Teiler von  $a, b$  sind also identisch mit den gemeinsamen Teilern von  $b, r$ .  $\square$

## 2.5.2 Das kgV

**Bemerkung a.** Es seien  $f, g \in R$ ,  $f, g \neq 0$ . Betrachte die Menge  $V$  der positiven bzw. normierten gemeinsamen Vielfachen von  $f$  und  $g$ , d.h.

$$V := \{v \in \mathbb{N} \mid f \text{ teilt } v \text{ und } g \text{ teilt } v\}$$

falls  $R = \mathbb{Z}$  bzw.

$$V := \{v \in K[X] \setminus \{0\} \mid f \text{ teilt } v, g \text{ teilt } v \text{ und } v \text{ normiert}\}$$

falls  $R = K[X]$ . Dann hat  $V$  bzgl. der Ordnung  $\mid$  ein Minimum.

*Beweis.* Betrachte das Ideal  $(f) \cap (g)$  (siehe Übung 2.4.2 a). Nach Bemerkung 2.5.1 a existiert ein  $v \in R$  mit  $(f) \cap (g) = (v)$ . Wir können oBdA  $v$  als positiv bzw. normiert annehmen. Dann ist  $v \in V$  wegen  $(v) \subseteq (f)$  und  $(v) \subseteq (g)$  (siehe Bemerkung 2.4.2). Sei nun  $v' \in V$ . Aus  $f \mid v'$  und  $g \mid v'$  folgt  $v' \in (f) \cap (g) = (v)$  nach Bemerkung 2.4.1 a. Damit gilt  $v \mid v'$  und  $v$  ist das Minimum von  $V$ .  $\square$

**Definition.** Seien  $f, g \in R$ . Das *kleinste gemeinsame Vielfache*, geschrieben  $\text{kgV}(f, g)$  von  $f$  und  $g$  ist definiert durch

$$\text{kgV}(f, g) := \min V$$

mit  $V$  wie Bemerkung a, falls  $f, g \neq 0$  sind, und

$$\text{kgV}(f, g) = 0,$$

falls  $f = 0$  oder  $g = 0$  ist.

**Bemerkung b.** Mit der Notation aus Abschnitt 2.5.1 gilt für alle  $a, b \in R$ :

- (i)  $\text{kgV}(a, b) = \text{kgV}(b, a)$ ,
- (ii)  $\text{kgV}(a, b) = \text{kgV}(|a|, |b|)$ ,
- (iii)  $\text{kgV}(a, 0) = 0$ .

*Übung.* Es seien  $a, b \in R$  nicht beide gleich 0. Man zeige

$$\text{kgV}(a, b) = \frac{|ab|}{\text{ggT}(a, b)}.$$

### 2.5.3 Der Euklidische Algorithmus

Hier stellen wir den Euklidische Algorithmus vor. Dieser berechnet nicht nur  $\text{ggT}(a, b)$  für  $a, b \in R$ , sondern auch eine Darstellung  $\text{ggT}(a, b) = \lambda a + \mu b$  mit  $\lambda, \mu \in R$ .

**Beispiel.** Wie lautet  $\text{ggT}(91, 168)$ ?

Rechnung:

$$168 = 1 \cdot 91 + 77$$

$$91 = 1 \cdot 77 + 14$$

$$77 = 5 \cdot 14 + 7$$

$$14 = 2 \cdot 7 + 0.$$

Nach Bemerkung (2.5.1)(d) gilt somit

$$\text{ggT}(168, 91) = \text{ggT}(91, 77) = \text{ggT}(77, 14) = \text{ggT}(14, 7) = \text{ggT}(7, 0) = 7.$$

Rückwärts Einsetzen:

$$7 = 77 - 5 \cdot 14$$

$$= 77 - 5 \cdot (91 - 1 \cdot 77) = -5 \cdot 91 + 6 \cdot 77$$

$$= -5 \cdot 91 + 6 \cdot (168 - 1 \cdot 91) = 6 \cdot 168 - 11 \cdot 91.$$

Somit gilt  $\text{ggT}(91, 168) = (-11) \cdot 91 + 6 \cdot 168$ .

**Beispiel** (Fortsetzung von Beispiel (2.4.4)). Wir dividieren  $g$  durch  $r$  mit Rest:

$$\begin{array}{r} (X^2 - 3X - 4) : (3X - 12) = \frac{1}{3}X + \frac{1}{3} = \frac{1}{3}(X + 1) \\ -(X^2 - 4X) \\ \hline X - 4 \\ -(X - 4) \\ \hline 0 \end{array}$$

D.h.  $g = \frac{1}{3}(X+1) \cdot r + 0$ . Damit ist  $r$  bis auf Normierung der ggT von  $f$  und  $g$ , also  $\text{ggT}(f, g) = X - 4$ . Rückwärtseinsetzen liefert weiterhin die Darstellung:

$$\begin{aligned}\text{ggT}(f, g) &= X - 4 = \frac{1}{3}(3X - 12) = \frac{1}{3}(f - (2X - 3)g) \\ &= \frac{1}{3} \cdot f - \frac{1}{3}(2X - 3) \cdot g.\end{aligned}$$

Im folgenden Algorithmus benutzen wir die Notation  $\nu$  aus (2.5.1).

**Algorithmus.** Es seien  $a, b \in R$  mit  $b \neq 0$ . Die folgende Prozedur liefert  $d, \lambda, \mu \in R$  mit  $d = \text{ggT}(a, b) = \lambda a + \mu b$ .

EUKLID( $a, b$ )

- 1 Bestimme  $q, r$  mit  $a = qb + r$  und  $\nu(r) < \nu(b)$ .
- 2 **if**  $r = 0$
- 3     **then return**  $(|b|, 0, |b|/b)$
- 4     **else**  $(d, \lambda, \mu) \leftarrow \text{EUKLID}(b, r)$
- 5     **return**  $(d, \mu, \lambda - q\mu)$

*Beweis.* 1. Es sei  $a = qb + r$ .

3. Falls  $r = 0$ , dann  $b \mid a$ , also  $\text{ggT}(a, b) = |b| = 0 \cdot a + |b|/b \cdot b$ .

4. Sei  $r > 0$  und  $d = \text{ggT}(b, r) = \lambda b + \mu r$ .

5. Nach Bemerkung (2.5.1)(d) ist  $d = \text{ggT}(a, b)$ . Außerdem gilt  $d = \lambda b + \mu(a - qb) = \mu a + (\lambda - q\mu)b$ .  $\square$

**Bemerkung.** Der größte gemeinsame Teiler wurde ohne Verwendung des Begriffs „Primzahl“ definiert und kann mit dem Euklidischen Algorithmus ohne Kenntnis der Primfaktorzerlegung berechnet werden.

*Übung a.* Es seien  $a, b \in \mathbb{N}$ . Die Koeffizienten  $\lambda, \mu$  in der Darstellung

$$\text{ggT}(a, b) = \lambda a + \mu b$$

sind nicht eindeutig. Geben Sie ein Beispiel an. Zeigen Sie weiter, dass  $\lambda, \mu$  unter der Zusatzbedingung  $-b/d < \lambda \leq 0$  und  $0 < \mu \leq a/d$  eindeutig werden.

## 2.6 Restklassenringe

In diesem Abschnitt führen wir die wichtigsten Konstruktionen von kommutativen Ringen und Körpern ein. Dies sind Restklassenringe von ganzen Zahlen bzw. Polynomringen.

### 2.6.1 Kongruenz modulo $n$

**Definition.** Für jedes  $n \in \mathbb{N}$  definieren wir auf  $\mathbb{Z}$  eine Relation  $\equiv_n$  durch

$$a \equiv_n b :\Leftrightarrow n \mid a - b.$$

Statt  $a \equiv_n b$  schreibt man auch  $a \equiv b \pmod{n}$  und sagt „ $a$  kongruent  $b$  modulo  $n$ “.

**Bemerkung.** Es gilt  $a \equiv_n b$  genau dann, wenn  $a$  und  $b$  bei Division durch  $n$  denselben Rest lassen.

*Beweis.* Seien  $a = qn + r$  und  $b = q'n + r'$  mit  $0 \leq r, r' < n$ . Dann ist  $a - b = (q - q')n + (r - r')$  und  $|r - r'| < n$ . Nach Bemerkung (2.4.1) gilt  $n \mid a - b$  genau dann, wenn  $n \mid r - r'$ . Wegen  $|r - r'| < n$  ist das genau dann der Fall, wenn  $r - r' = 0$ .  $\square$

**Beispiel.** Ist 14 kongruent 23 modulo 3 ( $14 \equiv_3 23$ )? Ja, weil  $14 - 23 = -9$  Vielfaches von 3 ist. Alternativ kann man die Reste bei Division durch 3 vergleichen:  $14 = 4 \cdot 3 + 2$  und  $23 = 7 \cdot 3 + 2$ . Sie stimmen überein (beide = 2).

**Satz.** Für jedes  $n \in \mathbb{N}$  ist die Relation  $\equiv_n$  eine Äquivalenzrelation.

*Beweis.* Klar aus der Bemerkung.  $\square$

### 2.6.2 Restklassen modulo $n$

Es sei  $n \in \mathbb{N}$  in diesem Abschnitt fest gewählt.

**Definition a.** Es sei  $a \in \mathbb{Z}$ . Wir setzen

$$a \bmod n := r,$$

für den eindeutig bestimmten Rest  $r \in \mathbb{Z}$  bei der Division mit Rest von  $a$  durch  $n$ . Es ist also  $a \bmod n = r$  genau dann, wenn  $a = qn + r$  mit  $0 \leq r < n$  ist.

**Definition b.** Die Äquivalenzklasse von  $a \in \mathbb{Z}$  bzgl.  $\equiv_n$  wird mit  $\bar{a}$  bezeichnet und wird die *Restklasse von  $a$  modulo  $n$*  genannt.

**Bemerkung.** Die Restklasse  $\bar{a}$  besteht aus allen ganzen Zahlen, die bei Division durch  $n$  denselben Rest lassen wie  $a$ . Es gilt

$$\bar{a} = a + n\mathbb{Z} = \{\dots, a - 2n, a - n, a, a + n, a + 2n, \dots\}.$$

Dividiert man  $a$  durch  $n$  mit Rest, etwa  $a = qn + r$  mit  $0 \leq r < n$ , so ist  $\bar{a} = \bar{r}$ . Mit Definition [a](#) gilt also

$$\bar{a} = \overline{a \bmod n}.$$

Der Rest  $a \bmod n$  ist weiterhin der kleinste nicht-negative Repräsentant von  $\bar{a}$ . Folglich hat jede Restklasse modulo  $n$  genau einen Repräsentanten zwischen 0 und  $n - 1$  (nämlich  $a \bmod n$  für die Restklasse, die  $a$  enthält). Es gibt also genau  $n$  verschiedene Restklassen modulo  $n$ :  $\bar{0}, \bar{1}, \dots, \overline{n-1}$ .

**Beispiel a.** Für  $n = 3$  ist  $\overline{14} = \{\dots, 5, 8, 11, 14, 17, 20, 23, \dots\} = \overline{23}$ . Wegen  $14 = 4 \cdot 3 + 2$  ist  $\overline{14} = \bar{2}$ , und 2 ist der kleinste nicht-negative Repräsentant von  $\overline{14}$ .

**Definition c.** Die Menge der Restklassen modulo  $n$  wird mit  $\mathbb{Z}_n$  (oder  $\mathbb{Z}/(n)$ ) bezeichnet, also  $\mathbb{Z}_n := \{\bar{0}, \bar{1}, \dots, \overline{n-1}\}$ . Es gilt  $|\mathbb{Z}_n| = n$ .

**Beispiel b.**  $\mathbb{Z}_3 = \{\bar{0}, \bar{1}, \bar{2}\}$ .

### 2.6.3 Rechnen mit Restklassen

Wir möchten auf der Menge der Restklassen modulo  $n$  zwei Verknüpfungen  $+$  und  $\cdot$  einführen mittels der Definition

$$\bar{a} + \bar{b} := \overline{a + b} \quad \text{und} \quad \bar{a} \cdot \bar{b} := \overline{a \cdot b}. \quad (*)$$

Das Problem in dieser Definition ist, dass sie – auf den ersten Blick – von der Wahl der Repräsentanten  $a$  und  $b$  abzuhängen scheint. Der folgende Satz zeigt, dass dem nicht so ist. Nur aufgrund des Satzes handelt es sich bei  $(*)$  überhaupt um eine gültige Definition.

**Satz.** Sei  $n \in \mathbb{N}$  fest. Sind  $a, a', b, b' \in \mathbb{Z}$  mit  $\bar{a} = \bar{a'}$  und  $\bar{b} = \bar{b'}$ , so gilt:

$$(i) \quad \overline{a + b} = \overline{a' + b'},$$

$$(ii) \quad \overline{a \cdot b} = \overline{a' \cdot b'}.$$

*Beweis.* Nach Voraussetzung ist  $n \mid a - a'$  und  $n \mid b - b'$ . Nach Bemerkung (2.4.1) (ii) teilt  $n$  auch  $(a - a') + (b - b') = (a + b) - (a' + b')$ , also gilt (i). Nach Bemerkung (2.4.1) (i) und (ii) teilt  $n$  auch  $(a - a')b' + (b - b')a = (a \cdot b - a' \cdot b')$ , also gilt (ii).  $\square$

**Folgerung.** Die Menge  $\mathbb{Z}_n$  bildet bzgl. der Verknüpfungen aus  $(*)$  einen kommutativen Ring.

*Beweis.* Addition und Multiplikation in  $\mathbb{Z}_n$  sind über die entsprechenden Operationen aus  $\mathbb{Z}$  definiert. Daher werde Assoziativ-, Kommutativ- und Distributivgesetze von  $\mathbb{Z}$  „geerbt“. Weiter ist die 0 in  $\mathbb{Z}_n$  die Restklasse  $\bar{0}$ , das negative Element zu  $\bar{a}$  ist  $\overline{-a}$ , und die 1 in  $\mathbb{Z}_n$  ist die Restklasse  $\bar{1}$ . Damit prüft man alle Axiome leicht nach.  $\square$

**Definition.** Der Ring  $(\mathbb{Z}_n, +, \cdot)$  mit den Verknüpfungen aus  $(*)$  wird *Restklassenring modulo  $n$*  genannt.

**Bemerkung.** Das praktische Rechnen mit Restklassen geschieht am besten auf die folgende Weise. Es seien  $0 \leq i, j < n$  Elemente aus  $\mathbb{Z}$ , die die Restklassen  $\bar{i}$  und  $\bar{j}$  repräsentieren.

- (i) Zur Addition von  $\bar{i}$  und  $\bar{j}$ , addiere  $i$  und  $j$  in  $\mathbb{Z}$  und dividiere das Ergebnis mit Rest durch  $n$ . Der Rest ist der Repräsentant der Restklasse  $\bar{i} + \bar{j}$ . In Formeln:

$$\bar{i} + \bar{j} = \overline{(i + j) \bmod n}.$$

Diese Rechnung wird noch durch folgende Überlegung vereinfacht. Ist  $i + j < n$ , dann ist  $(i + j) \bmod n = i + j$ . Andernfalls ist  $(i + j) \bmod n = n - (i + j)$ . Ist  $i > 0$ , dann ist  $\overline{n - i}$  das Additive Inverse von  $\bar{i}$ .

- (ii) Zur Multiplikation von  $\bar{i}$  und  $\bar{j}$ , multipliziere  $i$  und  $j$  in  $\mathbb{Z}$  und dividiere das Ergebnis mit Rest durch  $n$ . Der Rest ist der Repräsentant der Restklasse  $\bar{i} \cdot \bar{j}$ . In Formeln:

$$\bar{i} \cdot \bar{j} = \overline{(i \cdot j) \bmod n}.$$

Bei dieser Rechnung müssen nur nicht-negative ganze Zahlen kleiner als  $n(n - 1)$  betrachtet werden.

**Beispiel.**

- (i)  $\mathbb{Z}_2 = \{\bar{0}, \bar{1}\}$ , wobei

$$\begin{aligned}\bar{0} &= 2\mathbb{Z} = \{\text{gerade ganze Zahlen}\}, \\ \bar{1} &= 1 + 2\mathbb{Z} = \{\text{ungerade ganze Zahlen}\}.\end{aligned}$$

Die Verknüpfungstabellen von  $\mathbb{Z}_2$  lauten (beachte  $\bar{1} + \bar{1} = \overline{1 + 1} = \bar{2} = \bar{0}$ ):

$$\begin{array}{c|cc} + & \bar{0} & \bar{1} \\ \hline \bar{0} & \bar{0} & \bar{1} \\ \bar{1} & \bar{1} & \bar{0} \end{array} \quad \text{und} \quad \begin{array}{c|cc} \cdot & \bar{0} & \bar{1} \\ \hline \bar{0} & \bar{0} & \bar{0} \\ \bar{1} & \bar{0} & \bar{1} \end{array}$$

Aus der Tabelle für  $+$  liest man z.B. ab, dass „gerade+ungerade immer ungerade ergibt“ und dass „ungerade+ungerade immer gerade ergibt“. Diese Aussagen sind hiermit auch bewiesen (genauer durch obigen Satz)!

Identifiziert man  $\bar{0}$  mit falsch und  $\bar{1}$  mit wahr, so entspricht  $+$  gerade xor und  $\cdot$  entspricht  $\wedge$ . Damit ist gezeigt, dass auch  $(B, \text{xor}, \wedge)$  einen kommutativen Ring bildet, und dass dieser als identisch mit dem Ring  $(\mathbb{Z}_2, +, \cdot)$  angesehen werden kann.

(ii) Die Verknüpfungstabellen von  $\mathbb{Z}_4$  lauten

$+$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$		$\cdot$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$
$\bar{0}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	und	$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$	$\bar{0}$
$\bar{1}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{0}$		$\bar{1}$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$
$\bar{2}$	$\bar{2}$	$\bar{3}$	$\bar{0}$	$\bar{1}$		$\bar{2}$	$\bar{0}$	$\bar{2}$	$\bar{0}$	$\bar{2}$
$\bar{3}$	$\bar{3}$	$\bar{0}$	$\bar{1}$	$\bar{2}$		$\bar{3}$	$\bar{0}$	$\bar{3}$	$\bar{2}$	$\bar{1}$

(iii) In  $\mathbb{Z}_7$  gilt:

$$\begin{aligned}
 \bar{3} + \bar{5} &= \bar{8} = \bar{1}, \\
 \bar{3} - \bar{5} &= \bar{3} + (-\bar{5}) = \bar{3} + \overline{-5} = \overline{3-5} = \overline{-2} = \bar{5}, \\
 \bar{6} \cdot \bar{5} &= \overline{30} = \bar{2}, \\
 \bar{6} \cdot \bar{5} &= \overline{-1} \cdot \bar{5} = \overline{-5} = \bar{2}, \\
 \bar{6}^{100000} &= \overline{-1}^{100000} = \overline{(-1)^{100000}} = \bar{1}.
 \end{aligned}$$

(iv) In  $\mathbb{Z}_6$  gilt  $\bar{3} \cdot \bar{2} = \bar{6} = \bar{0}$ , aber  $\bar{3} \neq \bar{0}$  und  $\bar{2} \neq \bar{0}$ . Die Restklasse  $\bar{0}$  ist aber die 0 in  $\mathbb{Z}_6$ , d.h.  $\mathbb{Z}_6$  ist nicht nullteilerfrei!  $\mathbb{Z}_6$  ist auch ein Gegenbeispiel zur Kürzungsregel (vgl. Bemerkung 2.2.3):  $\bar{2} \cdot \bar{3} = \bar{4} \cdot \bar{3}$ , aber  $\bar{2} \neq \bar{4}$ .

(v) In  $\mathbb{Z}_6$  ist  $\bar{5}$  eine Einheit, denn  $\bar{5} \cdot \bar{5} = \bar{1}$  und  $\bar{1}$  ist die 1. Neben  $\bar{1}$  ist  $\bar{5}$  sogar die einzige Einheit (man prüfe das nach!), also  $\mathbb{Z}_6^\times = \{\bar{1}, \bar{5}\}$ . Man beachte  $\bar{5} = -\bar{1}$ .

*Übung.* Was für ein Ring ist  $\mathbb{Z}_1$ ?

## 2.6.4 Gleichungen in $\mathbb{Z}_n$

**Beispiel a.** Für welche  $b \in \mathbb{Z}$  ist  $\bar{9} \cdot x = \bar{b}$  in  $\mathbb{Z}_{15}$  lösbar?

$x$	$\bar{0}$	$\bar{1}$	$\bar{2}$	$\bar{3}$	$\bar{4}$	$\bar{5}$	$\bar{6}$	$\bar{7}$	$\bar{8}$	$\bar{9}$	$\bar{10}$	$\bar{11}$	$\bar{12}$	$\bar{13}$	$\bar{14}$
$\bar{9} \cdot x$	$\bar{0}$	$\bar{9}$	$\bar{3}$	$\bar{12}$	$\bar{6}$	$\bar{0}$	$\bar{9}$	$\bar{3}$	$\bar{12}$	$\bar{6}$	$\bar{0}$	$\bar{9}$	$\bar{3}$	$\bar{12}$	$\bar{6}$

Antwort: Es gibt genau dann eine Lösung, wenn  $\bar{b} = \bar{0}, \bar{3}, \bar{6}, \bar{9}, \bar{12}$ . Für  $b = 3$  gibt es z.B. die Lösungen  $x = \bar{2}, \bar{7}, \bar{12}$ .

**Satz.** Es seien  $n \in \mathbb{N}$  und  $a, b \in \mathbb{Z}$  gegeben. Die Gleichung  $\bar{a} \cdot x = \bar{b}$  in  $\mathbb{Z}_n$  ist genau dann lösbar, wenn  $\text{ggT}(a, n) \mid b$ .

*Beweis.* Sei  $\bar{a} \cdot x = \bar{b}$  lösbar, etwa  $\lambda \in \mathbb{Z}$  mit  $\bar{a} \cdot \bar{\lambda} = \bar{b}$ . D.h.  $n \mid \lambda a - b$ . Aus  $\text{ggT}(a, n) \mid \lambda a$  und  $\text{ggT}(a, n) \mid \lambda a - b$  folgt  $\text{ggT}(a, n) \mid b$  (vgl. Übung 2.4.1a).

Sei umgekehrt  $\text{ggT}(a, n) \mid b$ , etwa  $c \in \mathbb{Z}$  mit  $\text{ggT}(a, n) \cdot c = b$ . Nach Algorithmus 2.5.3 gibt es  $\lambda, \mu \in \mathbb{Z}$  mit  $\text{ggT}(a, n) = \lambda a + \mu n$ . Multiplikation mit  $c$  liefert  $b = (c\lambda)a + (c\mu)n$ . In  $\mathbb{Z}_n$  bedeutet das  $\bar{b} = \overline{c\lambda} \cdot \bar{a}$ , d.h.  $x = c\lambda$  ist eine Lösung.  $\square$

**Beispiel b.** Löse  $\bar{6} \cdot x = \bar{9}$  in  $\mathbb{Z}_{15}$ . Rechnung: Mit dem euklidischen Algorithmus berechnet man  $\text{ggT}(6, 15) = \bar{3} = 1 \cdot 15 - 2 \cdot 6$ . Multiplikation mit 3 liefert  $9 = 3 \cdot 15 - 6 \cdot 6$ . Modulo 15 ergibt sich  $\bar{9} = \bar{0} - \bar{6} \cdot \bar{6}$ . Folglich ist  $x = -\bar{6} = \overline{-6} = \bar{9}$  eine Lösung. Die Lösung ist nicht eindeutig, z.B. ist auch  $\bar{6} \cdot \bar{4} = \bar{24} = \bar{9}$  oder  $\bar{6} \cdot \bar{14} = \bar{6} \cdot \overline{-1} = \overline{-6} = \bar{9}$ .

**Definition.** Ein Element  $p \in \mathbb{N}$  heißt *Primzahl*, wenn  $p > 1$  ist und 1 und  $p$  die einzigen Teiler von  $p$  in  $\mathbb{N}$  sind.

**Folgerung.** Es seien  $n \in \mathbb{N}$  und  $a \in \mathbb{Z}$ .

$$(i) \quad \bar{a} \in \mathbb{Z}_n^\times \Leftrightarrow \text{ggT}(a, n) = 1.$$

$$(ii) \quad \mathbb{Z}_n \text{ ist genau dann ein Körper, wenn } n \text{ eine Primzahl ist.}$$

*Beweis.* Übung unter Verwendung des Satzes.  $\square$

**Beispiel c.**  $\mathbb{Z}_9^\times = \{\bar{1}, \bar{2}, \bar{4}, \bar{5}, \bar{7}, \bar{8}\}$ . Es gilt:

$$\begin{array}{ll} \bar{1}^{-1} = \bar{1}, & \bar{8}^{-1} = \bar{8}, \\ \bar{2}^{-1} = \bar{5}, & \bar{7}^{-1} = \bar{4}, \\ \bar{4}^{-1} = \bar{7}, & \bar{5}^{-1} = \bar{2}. \end{array}$$

*Übung a.* Wie lauten alle Einheiten von  $\mathbb{Z}_{11}$  und ihre Inversen? Ist  $\mathbb{Z}_{11}$  ein Körper?

*Übung b.* Man zeige, dass für teilerfremde  $a, b \in \mathbb{Z}$  stets gilt:  $a \mid bc \Rightarrow a \mid c$ .  
*Hinweis:* Man rechne in  $\mathbb{Z}_a$ .

*Übung c.* Wie viele Einheiten hat  $\mathbb{Z}_{p^n}$ , wenn  $p$  eine Primzahl ist?

*Übung d.* Es sei  $n > 1$ . Man zeige, dass  $\bar{a} \in \mathbb{Z}_n$  genau dann Nullteiler ist, wenn  $\text{ggT}(a, n) \neq 1$ .



*Übung e.* Man zeige, dass in  $\mathbb{Z}_n$  jedes Element entweder Einheit oder Nullteiler ist.

*Übung f.* Man prüfe folgende Aussage aus Übung 2.2.3b in verschiedenen  $\mathbb{Z}_n$  nach:  $(a) = (b) \Leftrightarrow a \sim b$ ?

## 2.6.5 Die Euler'sche Funktion

**Definition.** Für  $n \in \mathbb{N}$  definiere

$$\varphi(n) := |\mathbb{Z}_n^\times| = |\{a \in \mathbb{Z} \mid 0 \leq a < n, \text{ggT}(a, n) = 1\}|.$$

Die Abbildung  $\varphi : \mathbb{N} \rightarrow \mathbb{N}$  heißt die *Euler'sche  $\varphi$ -Funktion*.

**Bemerkung a.** (i) Für alle  $m, n \in \mathbb{N}$  mit  $\text{ggT}(m, n) = 1$  gilt  $\varphi(mn) = \varphi(m)\varphi(n)$ .

(ii) Für alle Primzahlen  $p$  gilt  $\varphi(p^k) = p^{k-1}(p-1)$ .

*Beweis.* (i) Ohne Beweis. (ii) Als Übung. (Kombinatorik!) □

**Beispiel.**  $\varphi(9) = \varphi(3^2) = 3^1(3-1) = 3 \cdot 2 = 6$ .

$\varphi(20) = \varphi(4) \cdot \varphi(5) = 2^1(2-1) \cdot 5^0(5-1) = 2 \cdot 4 = 8$ .

**Bemerkung b.** Es sei  $G$  eine endliche abelsche Gruppe und  $x \in G$ . Dann ist  $x^{|G|} = 1$ .

*Beweis.* Es sei  $|G| = m$  und  $G = \{g_1, g_2, \dots, g_m\}$ . Dann ist auch  $G = \{xg_1, xg_2, \dots, xg_m\}$ . Wir setzen  $a := \prod_{i=1}^m g_i \in G$  und erhalten

$$a = \prod_{i=1}^m g_i = \prod_{i=1}^m (xg_i) = x^{|G|} \prod_{i=1}^m g_i = x^{|G|} a,$$

wobei beide mittlere Gleichungen benutzen, dass  $G$  abelsch ist. Multiplikation mit  $a^{-1}$  liefert die Behauptung. □

Daraus ergeben sich zwei wichtige Resultate der elementaren Zahlentheorie.

**Satz a.** (Satz von Euler) *Es seien  $n \in \mathbb{N}$  und  $a \in \mathbb{Z}$  mit  $\text{ggT}(a, n) = 1$ . Dann ist*

$$a^{\varphi(n)} \equiv_n 1.$$

*Beweis.* Wegen  $\text{ggT}(a, n) = 1$  ist  $\bar{a} \in (\mathbb{Z}_n)^\times$ . Aus Bemerkung b ergibt sich mit  $\varphi(n) = |(\mathbb{Z}_n)^\times|$ , dass  $\bar{a}^{\varphi(n)} = \bar{1}$  ist in  $(\mathbb{Z}_n)^\times$ . Daraus folgt die Behauptung. □

Wir spezialisieren noch auf den Fall, dass  $n$  eine Primzahl ist.

**Satz b.** (Kleiner Satz von Fermat) *Es seien  $p \in \mathbb{P}$  und  $a \in \mathbb{Z}$  mit  $p \nmid a$ . Dann ist*

$$a^{p-1} \equiv_p 1.$$

### 2.6.6 Restklassenringe von $K[X]$

In diesem Abschnitt sei  $K$  ein Körper. Die Konstruktionen und Aussagen aus den Abschnitten 2.6.1 – 2.6.3 lassen sich auf den Fall des Ringes  $K[X]$  übertragen. Wir geben die wichtigsten Resultate ohne Beweise an. Diese lassen sich wie beim Ring der ganzen Zahlen führen.

**Definition.** Für jedes  $g \in K[X] \setminus \{0\}$  definieren wir auf  $K[X]$  eine Relation  $\equiv_g$  durch

$$f \equiv_g h :\Leftrightarrow g \mid f - h.$$

Statt  $f \equiv_g h$  schreibt man auch  $f \equiv h \pmod{g}$  und sagt „ $f$  kongruent  $h$  modulo  $g$ “.

**Bemerkung.** Es gilt  $f \equiv_g h$  genau dann, wenn  $f$  und  $h$  bei Division durch  $g$  denselben Rest lassen.

**Beispiel.** (i)  $X^2 - 1 \equiv_{X^2-1} 0$

(ii)  $X^2 \equiv_{X^2-1} 1$

(iii)  $X^4 - X^2 + 1 \equiv_{X^2-1} 1$

**Satz.** Für jedes  $g \in K[X] \setminus \{0\}$  ist die Relation  $\equiv_g$  eine Äquivalenzrelation.

**Definition a.** Es sei  $g \in K[X] \setminus \{0\}$ . Die Äquivalenzklasse von  $f \in K[X]$  bzgl.  $\equiv_g$  wird die *Restklasse von  $f$  modulo  $g$*  genannt und mit  $\bar{f}$  bezeichnet. Wir schreiben

$$K[X]/(g) := \{\bar{f} \mid f \in K[X]\}$$

für die Menge der Restklassen modulo  $g$ .

**Definition b.** Es sei  $d \in \mathbb{N}_0$ . Wir setzen

$$K[X]_{<d} := \{f \in K[X] \mid \deg f < d\}.$$

Insbesondere ist  $K[X]_{<0} = \{0\}$  und  $K[X]_{<1}$  die Menge der konstanten Polynome.

**Definition c.** Es seien  $g \in K[X] \setminus \{0\}$  und  $f \in K[X]$ . Wir setzen

$$f \bmod g := r,$$

für den eindeutig bestimmten Rest  $r \in K[X]$  bei der Division mit Rest von  $f$  durch  $g$ . Es ist also  $f \bmod g = r$  genau dann, wenn  $f = qg + r$  mit  $\deg r < \deg g$  ist.

**Bemerkung.** Es seien  $g \in K[X] \setminus \{0\}$  und  $n := \deg g$ .

Die Restklasse  $\bar{f}$  von  $f \in K[X]$  modulo  $g$  besteht aus den Polynomen, die bei Division durch  $g$  denselben Rest lassen wie  $f$ . Dies ist die Menge

$$\bar{f} = f + gK[X] := \{f + gh \mid h \in K[X]\}.$$

Dividiert man  $f$  durch  $g$  mit Rest, etwa  $f = qg + r$  mit  $\deg r < \deg g$ , so ist  $f \equiv_g r$ . Mit Definition [c](#) gilt also

$$\bar{f} = \overline{f \bmod g}.$$

Der Rest  $f \bmod g$  ist weiterhin der Repräsentant der Restklasse  $\bar{f}$  von kleinstem Grad. Folglich hat jede Restklasse modulo  $g$  genau einen Repräsentanten aus  $K[X]_{<n}$  (nämlich  $f \bmod g$  für die Restklasse, die  $f$  enthält). Es gibt also eine Bijektion zwischen der Menge der Restklassen modulo  $g$  und  $K[X]_{<n}$ .

**Satz.** Es sei  $g \in K[X] \setminus \{0\}$ . Sind  $f, f', h, h' \in K[X]$  mit  $f \equiv_g f'$  und  $h \equiv_g h'$ , so gilt:

$$(i) \quad f + h \equiv_g f' + h',$$

$$(ii) \quad f \cdot h \equiv_g f' \cdot h'.$$

**Folgerung.** Es sei  $g \in K[X] \setminus \{0\}$ . Die Menge  $K[X]/(g)$  bildet bzgl. der folgenden Verknüpfungen einen kommutativen Ring.

$$(i) \quad \bar{f} + \bar{h} := \overline{f + h}, \quad f, h \in K[X];$$

$$(ii) \quad \bar{f} \cdot \bar{h} := \overline{f \cdot h}, \quad f, h \in K[X];$$

**Definition.** Es sei  $g \in K[X] \setminus \{0\}$ . Der Ring  $(K[X]/(g), +, \cdot)$  mit den obigen Verknüpfungen wird *Restklassenring von  $K[X]$  modulo  $g$*  genannt.

**Bemerkung.** Es sei  $g \in K[X] \setminus \{0\}$  und  $n := \deg g$ . Das praktische Rechnen in  $K[X]/(g)$  geschieht am besten auf die folgende Weise. Seien  $f, h$  Elemente aus  $K[X]_{<n}$ , die die Restklassen  $\bar{f}$  und  $\bar{h}$  repräsentieren.

- (i) Zur Addition von  $\bar{f}$  und  $\bar{h}$ , addiere  $f$  und  $h$  in  $K[X]$ . Die Summe ist Repräsentant der Restklasse  $\bar{f} + \bar{h}$ . In Formeln:

$$\bar{f} + \bar{h} = \overline{f + h}.$$

Das Additive Inverse von  $\bar{f}$  ist  $\overline{-f}$ .

- (ii) Zur Multiplikation von  $\bar{f}$  und  $\bar{h}$ , multipliziere  $f$  und  $h$  in  $K[X]$  und dividiere das Ergebnis mit Rest durch  $g$ . Der Rest ist der Repräsentant der Restklasse  $\bar{f} \cdot \bar{h}$ . In Formeln:

$$\bar{f} \cdot \bar{h} = \overline{(f \cdot h) \bmod g}.$$

Bei dieser Rechnung müssen nur Polynome vom Grad höchstens  $2(n-1)$  betrachtet werden.

Was sind die Einheiten in  $K[X]/(g)$ ?

**Satz.** Es sei  $g \in K[X] \setminus \{0\}$  und  $f \in K[X]$ . Dann gilt:

$$\bar{f} \text{ ist Einheit in } K[X]/(g) \Leftrightarrow \text{ggT}(f, g) = 1.$$

**Definition.** Ein Element  $g \in K[X]$  heißt *irreduzibel*, wenn  $g \neq 0$  ist,  $\deg g \geq 1$  ist und es gilt: die einzigen Teiler von  $g$  sind Einheiten oder assoziiert zu  $g$ . Mit anderen Worten: Ist  $g = fh$  mit  $f, h \in K[X]$ , dann ist  $f \in K^\times$  oder  $h \in K^\times$ .

**Folgerung.** Es sei  $g \in K[X] \setminus \{0\}$ . Dann gilt:

$K[X]/(g)$  ist genau dann ein Körper, wenn  $g$  irreduzibel ist.

Mithilfe dieser Folgerung können wir weitere Körper definieren.

**Beispiel.** (i)  $\mathbb{F}_4 := \mathbb{F}_2[X]/(X^2 + X + 1)$  ist ein Körper mit vier Elementen. Wir setzen  $\alpha := \bar{X} \in \mathbb{F}_4$ . Die Elemente von  $\mathbb{F}_4$  sind  $0, 1, \alpha, 1 + \alpha$ . Es bestehen folgende Verknüpfungstafeln.

+	0	1	$\alpha$	$1 + \alpha$
0	0	1	$\alpha$	$1 + \alpha$
1	1	0	$1 + \alpha$	$\alpha$
$\alpha$	$\alpha$	$1 + \alpha$	0	1
$1 + \alpha$	$1 + \alpha$	$\alpha$	1	0

·	0	1	$\alpha$	$1 + \alpha$
0	0	0	0	0
1	0	1	$\alpha$	$1 + \alpha$
$\alpha$	0	$\alpha$	$1 + \alpha$	1
$1 + \alpha$	0	$1 + \alpha$	1	$\alpha$

Vergleiche diese Tafeln mit Beispiel 2.2.4 c.

- (ii)  $\mathbb{C} := \mathbb{R}[X]/(X^2 + 1)$  ist ein Körper, der Körper der komplexen Zahlen. Wir setzen  $i := \overline{X} \in \mathbb{C}$  und identifizieren  $\bar{r}$  mit  $r$  für  $r \in \mathbb{R}$ . Dann ist  $i^2 = -1$ , und jedes Element  $z \in \mathbb{C}$  hat eine eindeutige Darstellung als

$$z := a + bi$$

mit  $a, b \in \mathbb{R}$ . Wir nennen  $a$  den *Realteil* von  $z$  und  $b$  den *Imaginärteil* von  $z$ . Die Abbildung

$$\mathbb{C} \rightarrow \mathbb{C}, \quad a + bi \mapsto a - bi$$

heißt *komplexe Konjugation*.

## 2.7 Permutationen

### 2.7.1 Definition und Beispiele

Es sei  $A$  eine endliche Menge und  $|A| = n$ . Wir nummerieren die Elemente von  $A$  und schreiben  $A = \{a_1, a_2, \dots, a_n\}$ .

**Definition.** Eine bijektive Abbildung  $\pi : A \rightarrow A$  heißt *Permutation von  $A$* . Wir verwenden für Permutationen die Schreibweise

$$\pi = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \\ \pi(a_1) & \pi(a_2) & \cdots & \pi(a_n) \end{pmatrix}.$$

Die Menge aller Permutationen von  $A$  wird mit  $S_A$  bezeichnet, also

$$S_A := \{\pi : A \rightarrow A \mid \pi \text{ bijektiv}\}.$$

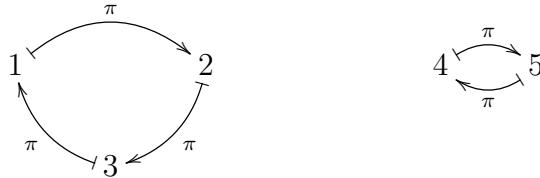
In dem wichtigen Spezialfall  $A = \underline{n}$  schreiben wir kurz  $S_n$  statt  $S_{\underline{n}}$ .

**Bemerkung.**

- (i) Wenn  $|A| = n$ , dann  $|S_A| = n!$ . Das gilt auch für  $n = 0$ , denn  $S_{\emptyset}$  hat genau ein Element (nach Beispiel 1.4.1viii existiert genau eine Abbildung  $\emptyset \rightarrow \emptyset$  und die ist bijektiv).
- (ii) Die Komposition von Permutationen von  $A$  ist wieder eine Permutation von  $A$ . Bei Permutationen sagt man statt Komposition auch *Produkt* und lässt das Zeichen  $\circ$  einfach weg.

**Beispiel.**

- (i) Die Permutation  $\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 1 & 5 & 4 \end{pmatrix} \in S_5$  lässt sich so veranschaulichen:



- (ii) Ist  $\psi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 4 & 3 & 1 & 2 \end{pmatrix}$  und  $\pi$  wie oben dann ergeben sich als Kompositionen

$$\pi \circ \psi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 5 & 1 & 2 & 3 \end{pmatrix} \quad \text{und} \quad \psi \circ \pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 3 & 5 & 2 & 1 \end{pmatrix}.$$

**2.7.2 Der Träger einer Permutation**

**Definition.** Für  $\pi \in S_A$  heißt

$$T_\pi := \{a \in A \mid \pi(a) \neq a\} \subseteq A$$

der *Träger* von  $\pi$ .

**Beispiel.**

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\ 5 & 8 & 3 & 6 & 2 & 7 & 4 & 1 & 9 & 11 & 10 \end{pmatrix}, \quad T_\pi = \{1, 2, 4, 5, 6, 7, 8, 10, 11\}.$$

**Bemerkung.** Es seien  $\pi, \psi \in S_A$ .

- (i)  $\pi(T_\pi) = T_\pi$ .
- (ii) Gilt  $T_\pi \subseteq B$ , so kann  $\pi$  auch als Element von  $S_B$  aufgefasst werden.
- (iii) Haben  $\pi$  und  $\psi$  disjunkte Träger, so gilt  $\pi \circ \psi = \psi \circ \pi$ .

*Beweis.*

- (i) Es reicht, die Inklusion  $\pi(T_\pi) \subseteq T_\pi$  zu zeigen. Daraus folgt schon die Gleichheit, da es sich um endliche Mengen handelt und da  $|\pi(T_\pi)| = |T_\pi|$  wegen der Injektivität von  $\pi$  gilt (vgl. Bem. 1.4.4a). Sei also  $a$  ein beliebiges Element aus  $T_\pi$ . Da  $\pi(a) \neq a$  und  $\pi$  injektiv, folgt  $\pi(\pi(a)) \neq \pi(a)$ . Das bedeutet gerade  $\pi(a) \in T_\pi$ . Da  $a \in T_\pi$  beliebig war, ist  $\pi(T_\pi) \subseteq T_\pi$  gezeigt.

(ii) klar.

(iii) als Übung.

□

### 2.7.3 Zykel und Transpositionen

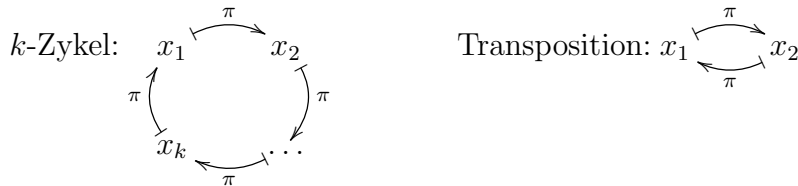
**Definition.** Es seien  $x_1, x_2, \dots, x_k \in A$  paarweise verschieden. Die Permutation  $\sigma \in S_A$  mit

$$\sigma(x) = \begin{cases} x_{i+1} & \text{falls } x = x_i \text{ und } i < k, \\ x_1 & \text{falls } x = x_k, \\ x & \text{falls } x \neq x_1, x_2, \dots, x_k, \end{cases}$$

heißt *Zykel der Länge  $k$*  oder kurz  *$k$ -Zykel* von  $S_A$ . Wir verwenden für  $\sigma$  die Schreibweise

$$\sigma = (x_1, x_2, \dots, x_k).$$

Die 2-Zykel heißen auch *Transpositionen* von  $S_A$ .



**Bemerkung.**

- (i) Es gilt stets  $(x_1, x_2, \dots, x_k)^k = \text{id}$ .
- (ii) Es gilt stets  $(x_1, x_2, \dots, x_k)^{-1} = (x_k, x_{k-1}, \dots, x_1)$ .
- (iii) Für Transpositionen  $\tau$  gilt  $\tau^{-1} = \tau$ .
- (iv) Jeder 1-Zykel ist die Identität.
- (v) Jeder  $k$ -Zykel lässt sich als Produkt von  $k-1$  Transpositionen schreiben:

$$(x_1, x_2, \dots, x_k) = (x_1, x_2)(x_2, x_3) \cdots (x_{k-1}, x_k).$$

Eine solche Zerlegung ist im Allgemeinen nicht eindeutig (vgl. Beispiel [a](#) unten).

**Beispiel a.** Der 4-Zykel  $\sigma := (1, 5, 2, 4) \in S_5$  ist die Permutation

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 4 & 3 & 1 & 2 \end{pmatrix}.$$

Es gilt

$$\begin{aligned} \sigma^{-1} &= \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 5 & 3 & 2 & 1 \end{pmatrix} = (4, 2, 5, 1), \\ \sigma^2 &= \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 1 & 3 & 5 & 4 \end{pmatrix} = (1, 2)(5, 4), \\ \sigma^3 &= \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 5 & 3 & 2 & 1 \end{pmatrix} = (1, 4, 2, 5), \\ \sigma^4 &= \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 1 & 2 & 3 & 4 & 5 \end{pmatrix} = \text{id}. \end{aligned}$$

Es gilt

$$\sigma = (1, 5)(5, 2)(2, 4) = (1, 4)(1, 2)(1, 5).$$

**Beispiel b.** Möchte man eine Liste von  $n$  Elementen ordnen (z.B. eine Liste von Wörtern nach alphabetischer Reihenfolge), so ist eine Permutation  $\pi \in S_n$  zu finden, die die (ungeordnete) Liste in ihre geordnete Reihenfolge überführt. Das  $i$ -te Wort der ungeordneten Liste steht in der geordneten Liste dann an  $\pi(i)$ -ter Stelle. Ein Sortieralgorithmus findet  $\pi$  im Allgemeinen nicht in einem Schritt, sondern nimmt nacheinander eine Reihe von Vertauschungen vor; er konstruiert somit  $\pi$  als ein Produkt  $\pi_1 \circ \dots \circ \pi_r$  einzelner (einfacherer) Umordnungen  $\pi_i$ . Der *Bubblesort*-Algorithmus kommt dabei z.B. mit Transpositionen  $\pi_i$  aus. Damit das immer funktioniert muss sich jede Permutation als Produkt von Transpositionen schreiben lassen. Davon überzeugen wir uns mit Hilfe von Satz 2.7.4 unten.

#### 2.7.4 Zerlegung in Zykel

**Satz.** Jede Permutation  $\pi \in S_A$  lässt sich als Produkt von Zykeln schreiben, deren Träger paarweise disjunkt sind. Bis auf Reihenfolge und bis auf Erwähnung von 1-Zykeln ist diese Zerlegung eindeutig.

*Beweis.* Siehe Beispiel. □

Man spricht kurz von einer Zerlegung von  $\pi$  in paarweise disjunkte Zykeln.



**Beispiel a.** Für  $\pi$  aus Beispiel (2.7.2) haben wir die Zerlegung

$$\begin{aligned}\pi &= (1, 5, 2, 8)(3)(4, 6, 7)(9)(10, 11) \\ &= (1, 5, 2, 8)(4, 6, 7)(10, 11).\end{aligned}$$

Die Träger der drei Zykeln lauten  $\{1, 5, 2, 8\}$ ,  $\{4, 6, 7\}$ ,  $\{10, 11\}$  und sind paarweise disjunkt. Die einzelnen Zykeln zerlegen sich weiter in Produkte von Transpositionen, z.B.  $(1, 5, 2, 8) = (1, 5)(5, 2)(2, 8)$  und  $(4, 6, 7) = (4, 6)(6, 7)$ , also

$$\pi = (1, 5)(5, 2)(2, 8)(4, 6)(6, 7)(10, 11).$$

Die Zykelschreibweise lässt sich besonders leicht „potenzieren“:

$$\begin{aligned}\pi &= (1, 5, 2, 8)(4, 6, 7)(10, 11), \\ \pi^2 &= (1, 2)(4, 7, 6)(5, 8), \\ \pi^3 &= (1, 8, 2, 5)(10, 11), \\ \pi^4 &= (4, 6, 7), \\ &\vdots \\ \pi^{11} &= (1, 8, 2, 5)(4, 7, 6)(10, 11) = \pi^{-1}, \\ \pi^{12} &= \text{id}.\end{aligned}$$

**Definition.** Es sei  $\pi \in S_A$ . Die *Zykelzahl* von  $\pi \in S_A$  ist die Anzahl der Zykeln inklusive aller 1-Zykeln, die bei einer Zerlegung von  $\pi$  in paarweise disjunkte Zykeln auftreten.

Die Zykelzahl ist gemäß obigem Satz eindeutig bestimmt. Sie hängt allerdings nicht nur von  $\pi$  sondern auch von  $A$  ab!

**Beispiel b.** Die Zykelzahl von  $\pi$  aus Beispiel a ist 5. Da sich die Identität  $\text{id} \in S_n$  in lauter 1-Zykeln zerlegt, hat sie die Zykelzahl  $n$ . Die Zykelzahl der (einzigen) Permutation  $\emptyset \rightarrow \emptyset$  wird als 0 definiert.

### 2.7.5 Das Signum

Wir bezeichnen hier mit  $I_A$  die Menge der 2-elementigen Teilmengen einer Menge  $A$ , d.h.  $I_A = \{\{i, j\} \subseteq A \mid i \neq j\}$ . Wir schreiben  $I_n$  für  $I_{\underline{n}}$ . (In der Kombinatorik lernen wir, dass  $|I_n| = \frac{n(n-1)}{2}$ .)

**Definition.** Sei  $\pi \in S_n$ . Das *Signum* von  $\pi$  ist definiert als

$$\text{sgn } \pi := \prod_{\{i, j\} \in I_n} \frac{\pi(i) - \pi(j)}{i - j}.$$

Man beachte, dass  $\text{sgn } \pi$  wohldefiniert ist, weil sich jeder einzelne Quotient nicht ändert, wenn man  $i$  und  $j$  vertauscht.

**Beispiel a.** Für  $\pi = \text{id} \in S_n$  sind alle Faktoren des Produktes gleich 1, also  $\text{sgn id} = 1$ . Für  $n = 2$  und  $\pi = (1, 2)$  ist  $I_n = \{\{1, 2\}\}$ , also  $\text{sgn}(1, 2) = \frac{2-1}{1-2} = -1$ .

**Bemerkung a.**

- (i) Es gilt stets  $\text{sgn } \pi = \pm 1$ .
- (ii) Wir nennen  $\pi$  *gerade*, falls  $\text{sgn } \pi = 1$  und *ungerade* falls  $\text{sgn } \pi = -1$ .
- (iii) Es gilt  $\text{sgn } \pi = \text{sgn } \pi'$  wobei  $\pi' := \pi|_{T_\pi} \in S_{T_\pi}$ .

*Beweis.* (i) Da  $\pi$  bijektiv ist, gilt  $\{\{\pi(i), \pi(j)\} \subseteq \underline{n} \mid i \neq j\} = I_n$ . D.h. wenn  $\{i, j\}$  die Menge  $I_n$  durchläuft, so durchläuft auch  $\{\pi(i), \pi(j)\}$  genau die Menge  $I_n$ . Folglich sind  $\prod_{\{i,j\} \in I_n} (\pi(i) - \pi(j))$  und  $\prod_{\{i,j\} \in I_n} (i - j)$  (Zähler und Nenner) bis auf Vorzeichen gleich, und somit  $|\text{sgn } \pi| = 1$ .

(iii) Es sei  $T = T_\pi$  (der Träger von  $\pi$ ) und  $F = \underline{n} \setminus T$  (die Fixpunkte von  $\pi$ ). Die Menge  $I_n$  partitioniert sich in  $I_n = I_T \cup I_F \cup \{\{i, j\} \mid i \in T, j \in F\}$ . Folglich zerlegt sich das Produkt aus der Definition von  $\text{sgn } \pi$  in die drei Teilprodukte

$$\begin{aligned} \prod_{\{i,j\} \in I_T} \frac{\pi(i) - \pi(j)}{i - j} &= \text{sgn } \pi', \\ \prod_{\{i,j\} \in I_F} \frac{\pi(i) - \pi(j)}{i - j} &= \prod_{\{i,j\} \in I_F} \frac{i - j}{i - j} = 1, \\ \prod_{i \in T, j \in F} \frac{\pi(i) - \pi(j)}{i - j} &= \underbrace{\prod_{j \in F} \prod_{i \in T} \frac{\pi(i) - j}{i - j}}_{=1} = 1. \end{aligned}$$

Man beachte in der letzten Gleichung, dass wenn  $i$  die Menge  $T$  durchläuft, dann auch  $\pi(i)$  genau die Menge  $T$  durchläuft. Damit ist  $\text{sgn } \pi = \text{sgn } \pi'$  gezeigt.  $\square$

**Beispiel b.** Für jede Transposition  $\pi = (a, b) \in S_n$  ist  $\text{sgn } \pi = -1$ .

*Beweis.* Es ist  $T_\pi = \{a, b\}$  und  $\pi' = \pi|_{\{a,b\}} \in S_{\{a,b\}}$ . Somit gilt

$$\text{sgn } \pi = \text{sgn } \pi' = \frac{\pi(a) - \pi(b)}{a - b} = \frac{b - a}{a - b} = -1.$$

$\square$

**Satz.** Für alle  $\pi, \psi \in S_n$  gilt  $\operatorname{sgn}(\pi \circ \psi) = \operatorname{sgn} \pi \cdot \operatorname{sgn} \psi$ .

*Beweis.* Es gilt

$$\begin{aligned} \operatorname{sgn}(\pi \circ \psi) &= \prod_{\{i,j\} \in I_n} \frac{(\pi \circ \psi)(i) - (\pi \circ \psi)(j)}{i - j} \\ &= \prod_{\{i,j\} \in I_n} \left( \frac{\pi(\psi(i)) - \pi(\psi(j))}{\psi(i) - \psi(j)} \cdot \frac{\psi(i) - \psi(j)}{i - j} \right) \end{aligned}$$

Da mit  $\{i, j\}$  auch  $\{\psi(i), \psi(j)\}$  genau die Menge  $I_n$  durchläuft ist dieses Produkt gleich  $\operatorname{sgn} \pi \cdot \operatorname{sgn} \psi$ .  $\square$

**Folgerung a.** Für alle  $\pi, \psi \in S_n$  gelten:

- (i)  $\operatorname{sgn} \pi^{-1} = \operatorname{sgn} \pi$ .
- (ii)  $\operatorname{sgn}(\psi^{-1} \pi \psi) = \operatorname{sgn} \pi$ .

*Beweis.* Der Satz und die Tatsache  $\operatorname{sgn} \operatorname{id} = 1$ .  $\square$

**Bemerkung b.** Aus Folgerung a(ii) geht hervor, dass eine Umbenennung der Elemente von  $\underline{n}$  (hier vorgenommen durch  $\psi$ ) das Signum von  $\pi$  nicht ändert. Die Definition des Signum stellt sich deshalb (nachträglich) als unabhängig von der Nummerierung innerhalb der Menge  $\underline{n}$  heraus.

Diese Tatsache kann man ausnutzen, um die Definition des Signum auf Permutationen  $\pi \in S_A$  (anstatt nur  $\pi \in S_n$ ) auszudehnen: wähle eine beliebige Bijektion  $\varphi : A \rightarrow \underline{n}$  und setze  $\operatorname{sgn} \pi := \operatorname{sgn}(\varphi \circ \pi \circ \varphi^{-1})$  (beachte  $\varphi \circ \pi \circ \varphi^{-1} \in S_n$ ).

**Folgerung b.** Es sei  $\pi \in S_A$ .

- (i) Ist  $\pi = \tau_1 \circ \dots \circ \tau_r$  mit Transpositionen  $\tau_i$ , so gilt  $\operatorname{sgn} \pi = (-1)^r$ .
- (ii) Ist  $\pi$  ein  $k$ -Zykel so gilt  $\operatorname{sgn} \pi = (-1)^{k-1}$ .
- (iii)  $\pi$  ist genau dann gerade, wenn in jeder Darstellung von  $\pi$  als Produkt von Transpositionen die Anzahl der Transpositionen gerade ist.

*Beweis.* (i) folgt aus dem Satz und Beispiel b. (ii) folgt aus (i) und Bemerkung 2.7.3(v). (iii) folgt aus (i).  $\square$

**Beispiel c.** Um das Signum der Permutation  $\pi$  aus Beispiel (2.7.2) zu berechnen, benutzen wir die Zerlegung  $\pi = (1, 5, 2, 8)(4, 6, 7)(10, 11)$  aus Beispiel (2.7.3). Dann ergibt sich aus dem Satz und Folgerung b(ii):

$$\operatorname{sgn} \pi = \operatorname{sgn}(1, 5, 2, 8) \cdot \operatorname{sgn}(4, 6, 7) \cdot \operatorname{sgn}(10, 11) = (-1)^3 \cdot (-1)^2 \cdot (-1)^1 = (-1)^6 = 1.$$

- Übung.* (i) Es seien  $\pi \in S_A$  und  $\varphi : A \rightarrow \underline{n}$  eine Bijektion. Man zeige, dass  $\operatorname{sgn}(\varphi \circ \pi \circ \varphi^{-1})$  unabhängig von der Wahl der Bijektion  $\varphi$  ist.
- (ii) Es seien  $\pi \in S_n$  und  $n \leq m$ . Fasse  $\pi$  als Element von  $S_m$  auf. Hängt  $\operatorname{sgn} \pi$  von  $m$  ab?
- (iii) Man zeige: Hat  $\pi \in S_n$  die Zykelzahl  $z$ , so gilt  $\operatorname{sgn} \pi = (-1)^{n-z}$ .

# Kapitel 3

## Lineare Gleichungssysteme und Matrizen

### 3.1 Matrizen

Es sei  $R$  ein kommutativer Ring mit  $1 \neq 0$ .

**Definition.**

- (i) Eine  $(m \times n)$ -Matrix  $A$  über  $R$  ist ein rechteckiges „Schema“ von  $m \cdot n$  Elementen  $a_{ij} \in R$  der Form

$$A = (a_{ij})_{\substack{1 \leq i \leq m, \\ 1 \leq j \leq n}} := \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix}$$

Die  $a_{ij} \in R$ ,  $1 \leq i \leq m$ ,  $1 \leq j \leq n$ , heißen die *Koeffizienten* oder *Einträge* von  $A$ .

- (ii) Zwei  $(m \times n)$ -Matrizen  $A = (a_{ij})$  und  $B = (b_{ij})$  über  $R$  heißen *gleich*, geschrieben  $A = B$ , wenn  $a_{ij} = b_{ij}$  für alle  $1 \leq i \leq m$  und alle  $1 \leq j \leq n$ . Die Menge aller  $(m \times n)$ -Matrizen über  $R$  wird mit  $R^{m \times n}$  bezeichnet.

- (iii) Es sei  $A = (a_{ij}) \in R^{m \times n}$ .

Die  $(1 \times n)$ -Matrix  $z_i := (a_{i1} \ a_{i2} \ \dots \ a_{in})$  heißt *i-te Zeile* von  $A$ .

Die  $(m \times 1)$ -Matrix  $s_j := \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix}$  heißt *j-te Spalte* von  $A$ .

- (iv) Eine  $(1 \times n)$ -Matrix wird auch (Zeilen-)  $n$ -Tupel und eine  $(m \times 1)$ -Matrix wird (Spalten-)  $m$ -Tupel genannt. Wir setzen (vgl. Definition 1.4.1b):

$$R^n := R^{n \times 1} = \text{Menge aller Spalten-}n\text{-Tupel über } R.$$

- (v) Eine  $(m \times n)$ -Matrix  $A = (a_{ij})$  mit allen  $a_{ij} = 0$  wird *Nullmatrix* genannt, geschrieben  $A = 0$ .

**Bemerkung.**

- (i) Im Index gilt „Zeile vor Spalte“, d.h.  $a_{ij}$  steht in der  $i$ -ten Zeile und  $j$ -ten Spalte.
- (ii) Eine  $(m \times n)$ -Matrix  $A = (a_{ij})$  über  $R$  kann als Abbildung

$$a : \underline{m} \times \underline{n} \rightarrow R, (i, j) \mapsto a(i, j) := a_{ij}$$

aufgefasst werden. Das steht in Analogie zu den  $n$ -Tupeln, die man ebenfalls als Abbildung auffassen kann (vgl. Definition 1.4.1b).

**Schreibweise.** Sind  $z_1, \dots, z_m$  die Zeilen und  $s_1, \dots, s_n$  die Spalten von  $A$ , so schreiben wir auch:

$$A = \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_m \end{pmatrix} = (s_1 \quad s_2 \quad \dots \quad s_n) = (s_1, s_2, \dots, s_n).$$

Diese Vereinbarung ist Teil einer flexiblen Schreibweise, nach der eine Matrix aus Blöcken, die selbst Matrizen sind, zusammengebaut werden kann. Man kann z.B.

$$M = \begin{pmatrix} A & B \\ C & D \end{pmatrix}$$

bilden, wenn  $A$  und  $B$  ebenso wie  $C$  und  $D$  jeweils gleich viele Zeilen haben, und  $A$  und  $C$  ebenso wie  $B$  und  $D$  jeweils gleich viele Spalten.

**Beispiel.**

- (i)  $\begin{pmatrix} 2 & -1 \\ 4 & 0 \\ 5 & 3 \end{pmatrix}$  ist eine  $(3 \times 2)$ -Matrix.
- (ii)  $\begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$  ist die  $(2 \times 3)$ -Nullmatrix.
- (iii)  $\underbrace{\begin{pmatrix} 2 \\ 4 \\ 5 \end{pmatrix}}_{(3 \times 1)} \neq \underbrace{(2 \quad 4 \quad 5)}_{(1 \times 3)}.$

## 3.2 Matrix-Arithmetik

In dem ganzen Abschnitt ist  $R$  ein kommutativer Ring mit  $1 \neq 0$  und  $R^{m \times n}$  die Menge der  $m \times n$ -Matrizen über  $R$ .

### 3.2.1 Die Grundrechenarten

**Schreibweise.** Es sei  $A \in R^{m \times n}$ . Für  $1 \leq i \leq m$  und  $1 \leq j \leq n$  bezeichnen wir – wie üblich – mit  $a_{ij}$  den  $(i, j)$ -Eintrag von  $A$ , d.h. den Eintrag in der  $i$ -ten Zeile und  $j$ -ten Spalte. (Merke: „Zeile vor Spalte“).

Sei umgekehrt  $a$  eine Abbildung  $a : \underline{m} \times \underline{n} \rightarrow R, (i, j) \mapsto a(i, j)$ . Dann bezeichnen wir mit

$$(a(i, j)) := (a(i, j))_{ij} := (a(i, j))_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}}$$

diejenige Matrix  $A \in R^{m \times n}$  mit  $a_{ij} = a(i, j)$  für alle  $1 \leq i \leq m, 1 \leq j \leq n$ .

**Definition.** Es seien  $A \in R^{m \times n}$  und  $r \in R$ .

- (i)  $A^t := (a_{ji})_{\substack{1 \leq i \leq n \\ 1 \leq j \leq m}} \in R^{n \times m}$  heißt die *Transponierte* von  $A$ .
- (ii)  $r \cdot A := (r \cdot a_{ij})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} \in R^{m \times n}$  heißt (*skalares*) *Vielfaches* von  $A$ .
- (iii) Für jedes  $B = (b_{ij}) \in R^{m \times n}$  definieren wir die *Summe*  $A + B := (a_{ij} + b_{ij})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq n}} \in R^{m \times n}$ .
- (iv) Für jedes  $B = (b_{ij}) \in R^{n \times l}, l \in \mathbb{N}$ , definieren wir das *Produkt*  $A \cdot B := (c_{ij})_{\substack{1 \leq i \leq m \\ 1 \leq j \leq l}} \in R^{m \times l}$  durch

$$c_{ij} := \sum_{k=1}^n a_{ik} b_{kj} \text{ für alle } i, j.$$

**Beispiel a.**

$$\begin{pmatrix} 2 & 3 & 0 \\ -1 & 0 & 2 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 & -1 \\ 1 & 0 & 1 \\ 2 & -1 & 0 \end{pmatrix} = \begin{pmatrix} 3 & 2 & 1 \\ 4 & -3 & 1 \end{pmatrix}$$

$$\begin{pmatrix} 2 & 3 & 0 \\ -1 & 0 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 1 & 0 \\ 2 & -1 & 0 & 3 \\ 0 & 2 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 8 & -3 & 2 & 9 \\ -1 & 4 & -1 & 2 \end{pmatrix}$$

$$\begin{pmatrix} 2 & 3 & 0 \\ -1 & 0 & 2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & 1 \\ 2 & -1 & 0 \end{pmatrix} \text{ nicht definiert}$$

**Bemerkung a.**

- (i) Die Zeilen von  $A^t$  erhält man aus den Spalten von  $A$  (in gleicher Reihenfolge), und umgekehrt. Spalten addiert haben in der Schreibweise  $\mathbb{L}(A, b) = s + \mathbb{L}(A, 0)$ .
- (ii) Wir identifizieren  $R^{1 \times 1}$  mit  $R$ , also die  $1 \times 1$ -Matrix  $(a)$  über  $R$  mit dem Ringelement  $a \in R$ .
- (iii) Das Produkt  $A \cdot B$  ist nur definiert, wenn Spaltenzahl von  $A$  gleich Zeilenzahl von  $B$  ist.

$$\cdot : R^{m \times n} \times R^{n \times l} \rightarrow R^{m \times l}$$

Spezialfälle:

$$\begin{array}{ll} \cdot : R^{m \times n} \times R^n \rightarrow R^m & l = 1 \text{ (Matrix} \cdot \text{Spalte=Spalte)} \\ \cdot : R^{1 \times n} \times R^{n \times l} \rightarrow R^{1 \times l} & m = 1 \text{ (Zeile} \cdot \text{Matrix=Zeile)} \\ \cdot : R^{1 \times n} \times R^n \rightarrow R = R^{1 \times 1} & l = m = 1 \text{ (Skalarprodukt)} \\ \cdot : R^m \times R^{1 \times l} \rightarrow R^{m \times l} & n = 1 \text{ (Spalte} \cdot \text{Zeile=Matrix)} \end{array}$$

Der Fall  $l = m = 1$  ist das Skalarprodukt aus der Schule, nur dass hier einer der Vektoren als Zeile geschrieben wird.

- (iv) Es seien  $A \in R^{m \times n}$  und  $B \in R^{n \times l}$ . Bezeichnet  $z_i$  die  $i$ -te Zeile von  $A$  und  $s_j$  die  $j$ -te Spalte von  $B$ , so gilt

$$A \cdot B = (z_i \cdot s_j)_{\substack{1 \leq i \leq m, \\ 1 \leq j \leq l}} \in R^{m \times l}.$$

Hier bezeichnet  $\cdot$  in  $z_i \cdot s_j$  die Matrixmultiplikation (also das Skalarprodukt), und die  $(1 \times 1)$ -Matrix  $z_i \cdot s_j$  wird ihrem Eintrag identifiziert.

**Beispiel b.**

$$(i) \quad l = 1: \begin{pmatrix} 1 & 0 & -2 \\ 3 & 2 & 0 \end{pmatrix} \cdot \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 3 \\ 11 \end{pmatrix}$$

$$(ii) \quad m = 1: (1 \quad 0 \quad -2) \cdot \begin{pmatrix} 0 & 1 \\ -1 & 0 \\ 1 & 1 \end{pmatrix} = (-2 \quad -1)$$

$$(iii) \quad l = m = 1 \text{ (Skalarprodukt): } (1 \quad 0 \quad -2) \cdot \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix} = 3 + 0 + 0 = 3$$



$$(iv) \ n = 1: \begin{pmatrix} 3 \\ 1 \\ 0 \end{pmatrix} \cdot (1 \ 0 \ -2) = \begin{pmatrix} 3 & 0 & -6 \\ 1 & 0 & -2 \\ 0 & 0 & 0 \end{pmatrix}$$

### 3.2.2 Quadratische Matrizen

**Definition.** Es sei  $n \in \mathbb{N}$ .

(i) Eine  $n \times n$ -Matrix heißt *quadratisch*.

(ii) Die  $n$ -reihige *Einheitsmatrix* ist definiert als  $E_n := (\delta_{ij})_{1 \leq i, j \leq n}$  mit

$$\delta_{ij} := \begin{cases} 1 & \text{falls } i = j, \\ 0 & \text{falls } i \neq j. \end{cases}$$

$$\text{Es gilt } E_n = \begin{pmatrix} 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{pmatrix} \in R^{n \times n}, \text{ z.B. } E_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

(iii) Quadratische Matrizen der Formen

$$\begin{pmatrix} \star & & 0 \\ & \ddots & \\ 0 & & \star \end{pmatrix}, \quad \begin{pmatrix} \star & \cdots & \star \\ & \ddots & \vdots \\ 0 & & \star \end{pmatrix}, \quad \text{bzw.} \quad \begin{pmatrix} \star & & 0 \\ \vdots & \ddots & \\ \star & \cdots & \star \end{pmatrix}$$

mit beliebigen Einträgen  $\star \in R$  heißen *Diagonalmatrix*, *obere Dreiecksmatrix*, bzw. *untere Dreiecksmatrix*.

### 3.2.3 Der Matrizenring

**Satz.** Es seien  $n, m, l, p \in \mathbb{N}$ . Es bezeichne  $0$  die  $m \times n$ -Nullmatrix. Für alle  $A, A' \in R^{m \times n}, B, B' \in R^{n \times l}, C \in R^{l \times p}$  und  $r \in R$  gilt:

(i)  $(R^{m \times n}, +)$  ist abelsche Gruppe mit neutralem Element  $0$ .

(ii)  $(A \cdot B) \cdot C = A \cdot (B \cdot C)$

(iii)  $E_m \cdot A = A = A \cdot E_n$

(iv)  $(A + A') \cdot B = A \cdot B + A' \cdot B$

(v)  $A \cdot (B + B') = A \cdot B + A \cdot B'$

(vi)  $r \cdot (A \cdot B) = (r \cdot A) \cdot B = A \cdot (r \cdot B)$

$$(vii) \quad (A^t)^t = A$$

$$(viii) \quad (A + A')^t = A^t + (A')^t$$

$$(ix) \quad (A \cdot B)^t = B^t \cdot A^t$$

*Beweis.* (i) ist klar, weil  $+$  einträgenweise definiert ist (vgl. §2.1.6).

(ii) Auf beiden Seiten ergibt sich der  $(i, j)$ -Eintrag  $\sum_{\alpha=1}^n \sum_{\beta=1}^l a_{i\alpha} b_{\alpha\beta} c_{\beta j}$  (Rechnung als Übung).

(iii) Nach Bemerkung 3.2.1iv ist  $E_m \cdot A = (z_i \cdot s_j)_{ij}$ , wobei  $z_i$  die  $i$ -te Zeile von  $E_m$  ist und  $s_j$  die  $j$ -te Spalte von  $A$ . Es gilt

$$z_i = (0 \cdots 0 \underbrace{1}_{\text{Pos. } i} 0 \cdots 0) \quad \text{und} \quad s_j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{mj} \end{pmatrix},$$

also

$$z_i \cdot s_j = 0 \cdot a_{1j} + \cdots + 1 \cdot a_{ij} + 0 + \cdots + 0 = a_{ij}.$$

Damit ist  $E_m \cdot A = (a_{ij}) = A$  gezeigt. Genauso verfährt man mit  $A \cdot E_n = A$ .  
(iv)

$$\begin{aligned} (A + B) \cdot C &= \left( \sum_{k=1}^n (a_{ik} + b_{ik}) c_{kj} \right)_{ij} \\ &= \left( \sum_{k=1}^n a_{ik} c_{kj} + \sum_{k=1}^n b_{ik} c_{kj} \right)_{ij} \\ &= \left( \sum_{k=1}^n a_{ik} c_{kj} \right)_{ij} + \left( \sum_{k=1}^n b_{ik} c_{kj} \right)_{ij} = AC + BC. \end{aligned}$$

(v) genauso wie (iv).

(vi) Übung (Ansatz wie in (iv)).

(vii) und (viii) sind klar.

(ix)

$$\begin{aligned} (A \cdot B)^t &= \left( \sum_{k=1}^n a_{ik} b_{kj} \right)_{ij}^t = \left( \sum_{k=1}^n a_{ik} b_{kj} \right)_{ji} \\ &\quad \parallel \\ B^t \cdot A^t &= (b_{ji})_{ij} \cdot (a_{ji})_{ij} = \left( \sum_{k=1}^n b_{ki} a_{jk} \right)_{ij} \end{aligned}$$

□

*Übung.* Für welche Teile des Satzes braucht man, dass  $R$  kommutativ ist?

**Folgerung.** Es sei  $n \in \mathbb{N}$ . Dann wird  $R^{n \times n}$  mit der Matrix-Addition und Matrix-Multiplikation aus Definition (3.2.1) zu einem Ring, dem Matrizenring. Die neutralen Elemente sind  $0 \in R^{n \times n}$  bzgl. der Addition und  $E_n \in R^{n \times n}$  bzgl. der Multiplikation.

*Beweis.* Die Eigenschaften (i)–(v) aus Satz 3.2.3. □

**Bemerkung.**

(i)  $R^{1 \times 1}$  kann mit  $R$  identifiziert werden.

(ii)  $R^{n \times n}$  ist für  $n \geq 2$  nicht kommutativ. Für  $n = 2$  sieht man das an

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \neq \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix},$$

und ein solches Beispiel lässt sich für jedes  $n \geq 2$  finden.

(iii)  $R^{n \times n}$  ist für  $n \geq 2$  nicht nullteilerfrei (sogar wenn  $R$  ein Körper ist). Es gibt sogar  $A \in R^{n \times n}$ ,  $A \neq 0$ , mit  $A^2 = 0$ , wie man an dem Beispiel

$$A = \begin{pmatrix} \cdots & 0 & 1 \\ & 0 & 0 \\ & & \vdots \end{pmatrix} \text{ sieht. Insbesondere ist } R^{n \times n} \text{ für } n \geq 2 \text{ kein Körper.}$$

(iv)  $R^{n \times n}$  ist auch mit komponentenweiser Multiplikation ein Ring (sogar ein kommutativer Ring). Dieser Ring ist aber nicht besonders interessant. Mit komponentenweiser Multiplikation ist man nicht auf quadratische Matrizen beschränkt, auch  $R^{m \times n}$  wird damit zu einem Ring.

### 3.2.4 Die lineare Gruppe

**Definition.** Die Einheitsgruppe des Matrizenringes  $R^{n \times n}$  (vgl. §2.2.2) wird die *allgemeine lineare Gruppe* über  $R$  vom Grad  $n$  genannt, geschrieben

$$\mathrm{GL}_n(R) := (R^{n \times n})^\times = \{A \in R^{n \times n} \mid A \text{ invertierbar}\}.$$

Die invertierbaren Matrizen heißen auch *regulär*. Das inverse Element zu  $A \in \mathrm{GL}_n(R)$  wird die *inverse Matrix* zu  $A$  genannt, oder die *Inverse* von  $A$ .

**Beispiel.**  $A = \begin{pmatrix} 1 & 2 \\ -1 & -1 \end{pmatrix} \in \mathbb{Q}^{2 \times 2}$  ist regulär:

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

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

Also ist  $A^{-1} = \begin{pmatrix} -1 & -2 \\ 1 & 1 \end{pmatrix}$ .

**Bemerkung a.** Mit  $A \in \text{GL}_n(R)$  ist auch  $A^t \in \text{GL}_n(R)$  und  $(A^t)^{-1} = (A^{-1})^t$ .

*Beweis.* Nach Satz 3.2.3ix gilt

$$A^t \cdot (A^{-1})^t = (A^{-1} \cdot A)^t = E_n^t = E_n,$$

und

$$(A^{-1})^t \cdot A^t = (A \cdot A^{-1})^t = E_n^t = E_n.$$

□

*Übung.* Es seien  $A, B \in R^{n \times n}$ .

- (i) Kann man aus  $A \cdot B = E_n$  schließen, dass  $A$  regulär und  $B$  die Inverse von  $A$  ist?
- (ii) Wenn  $A$  als regulär vorausgesetzt wird, ist dann  $B$  notwendigerweise die Inverse von  $A$ ? Was hat das mit Übung 2.1.3 zu tun?

## 3.3 Lineare Gleichungssysteme

### 3.3.1 Lineare Gleichungssysteme

In diesem Abschnitt sei  $K$  ein Körper.

**Definition.** Ein *lineares Gleichungssystem* über  $K$ , kurz LGS, hat die Form

$$\begin{array}{ccccccc} a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\ a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\ & & & & & & \vdots & & \\ a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m \end{array}$$

mit  $a_{ij}, b_j \in K$  (die *Koeffizienten* des LGS). Das sind  $m$  Gleichungen in den  $n$  *Unbekannten*  $x_1, \dots, x_n$ .

Eine *Lösung* des LGS ist ein Spalten- $n$ -Tupel  $\begin{pmatrix} s_1 \\ \vdots \\ s_n \end{pmatrix} \in K^n (= K^{n \times 1})$  derart,

dass alle  $m$  Gleichungen erfüllt sind, wenn  $s_i$  für  $x_i$  eingesetzt wird ( $i = 1, \dots, n$ ). Die Menge aller Lösungen wird mit  $\mathbb{L}$  bezeichnet. Das LGS heißt *homogen*, wenn  $b_1 = b_2 = \dots = b_m = 0$ , sonst *inhomogen*.

**Aufgabe:** Gegeben  $a_{ij}$  und  $b_i$ , bestimme alle Lösungen!

**Beispiel a.** Es sei  $K = \mathbb{R}$  und  $n = 2$ ; statt  $x_1, x_2$  nimm  $x, y$ .

$$x^2 + y^2 = 1 \quad \text{und} \quad xy = 1 \quad \text{sind nicht linear.}$$

**Beispiel b.**  $n = 2, m = 2$ .

$$\begin{array}{lll} \text{(i)} & \begin{array}{l} x + y = 2 \\ x - y = 0 \end{array} & \text{(ii)} \begin{array}{l} x + y = 2 \\ x + y = 0 \end{array} & \text{(iii)} \begin{array}{l} x + y = 2 \\ 3x + 3y = 6 \end{array} \end{array}$$

Lösung:

(i) Aus  $x - y = 0$  folgt  $x = y$ . Einsetzen in  $x + y = 2$  liefert  $2x = 2$ , also  $x = 1$ . Ergebnis:  $\mathbb{L} = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}$  (genau eine Lösung).

(ii) Es folgt der Widerspruch  $0 = 2$ . Ergebnis:  $\mathbb{L} = \emptyset$  (keine Lösung).

(iii) Aus  $x + y = 2$  folgt  $y = 2 - x$ . Einsetzen in  $3x + 3y = 6$  liefert  $3x + 6 - 3x = 6$ , also  $6 = 6$ . Das ist redundant und  $x$  bleibt „frei“. Ergebnis:  $\mathbb{L} = \left\{ \begin{pmatrix} x \\ 2 - x \end{pmatrix} \mid x \in K \right\}$  (mehr als eine Lösung).

Die gezeigten Lösungswege mittels Auflösen und Einsetzen nennt man *algebraische Lösungswege*. Es gibt auch *geometrische Lösungswege*, die aber in der Vorlesung nicht thematisiert werden.

### 3.3.2 Äquivalenzumformungen

In diesem Abschnitt sei wieder  $K$  ein Körper. Unsere Untersuchungen zielen darauf ab, die folgenden Frage zu beantworten.

- (i) Wie löst man Gleichungen mit beliebig vielen Unbekannten? (Eine algebraische Lösung ist bevorzugt.)
- (ii) Gibt es einen systematischen Weg?

(iii) Wie viele Lösungen kann es dabei geben?

**Satz.** Die Lösungsmenge eines LGS ändert sich nicht, wenn

- (i) zwei Gleichungen vertauscht werden, oder
- (ii) das  $c$ -fache ( $c \in K$ ) einer Gleichung zu einer anderen addiert wird, oder
- (iii) eine Gleichung mit einem  $c \in K$  ( $c \neq 0$ ) multipliziert wird.

Diese Umformungen heißen Äquivalenzumformungen.

*Beweis.* Die Aussagen (i) und (iii) sind klar. Um (ii) zu beweisen, können wir wegen (i) annehmen, dass die betreffenden Gleichungen die ersten beiden sind, also

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \end{aligned}$$

Nach der Umformung in (ii) werden daraus die beiden Gleichungen

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ (a_{21} + ca_{11})x_1 + (a_{22} + ca_{12})x_2 + \cdots + (a_{2n} + ca_{1n})x_n &= b_2 + cb_1 \end{aligned}$$

Ist nun  $\begin{pmatrix} s_1 \\ \vdots \\ s_n \end{pmatrix} \in K^n$  mit

$$\sum_{j=1}^n a_{1j}s_j = b_1$$

und

$$\sum_{j=1}^n a_{2j}s_j = b_2,$$

dann gilt auch

$$\sum_{j=1}^n (a_{1j} + ca_{2j})s_j = \sum_{j=1}^n a_{1j}s_j + c \sum_{j=1}^n a_{2j}s_j = b_1 + cb_2.$$

Damit ist jede Lösung des ursprünglichen LGS auch eine Lösung des umgeformten LGS. Das ursprüngliche LGS erhält man aus dem Umgeformten LGS durch Addition des  $(-c)$ -fachen der ersten Gleichung auf die zweite. Damit ist jede Lösung des umgeformten LGS auch eine Lösung des ursprünglichen LGS. Daraus folgt die Behauptung.  $\square$

**Beispiel.** Äquivalenzumformungen am Beispiel 3.3.1b:

$$\begin{array}{ccc}
 \begin{array}{l} x + y = 2 \\ x - y = 0 \end{array} \left| \begin{array}{l} \cdot (-1) \\ \longleftarrow \end{array} \right]_+ & \Longleftrightarrow & \begin{array}{l} x + y = 2 \\ -2y = -2 \end{array} \left| \begin{array}{l} \\ \cdot (-\frac{1}{2}) \end{array} \right. \\
 \Longleftrightarrow & & \Longleftrightarrow \\
 \begin{array}{l} x + y = 2 \\ y = 1 \end{array} \left| \begin{array}{l} \longleftarrow \\ \cdot (-1) \end{array} \right]_+ & & \begin{array}{l} x = 1 \\ y = 1 \end{array}
 \end{array}$$

Die Lösungsmenge lautet also  $\mathbb{L} = \left\{ \begin{pmatrix} 1 \\ 1 \end{pmatrix} \right\}$ .

**Bemerkung.**

- (i) Äquivalenzumformungen sind eine (bessere) Alternative zum „Auflösen und Einsetzen“.
- (ii) Wir haben in dem Beispiel nur mit den Koeffizienten des LGS gerechnet. Wir können uns sparen, die Unbekannten mit aufzuschreiben, wenn wir die Koeffizienten am „richtigen Platz“ belassen ( $\rightarrow$  Matrix eines LGS).

### 3.3.3 Die Koeffizientenmatrix

Es sei  $K$  ein beliebiger Körper.

**Definition.** Gegeben sei das LGS über  $K$ :

$$\begin{array}{ccccccc}
 a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1n}x_n & = & b_1 \\
 a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2n}x_n & = & b_2 \\
 & & & & & & \vdots & & \\
 a_{m1}x_1 & + & a_{m2}x_2 & + & \cdots & + & a_{mn}x_n & = & b_m
 \end{array}$$

mit  $a_{ij}, b_i \in K$  für alle  $1 \leq i \leq m, 1 \leq j \leq n$ . Die Matrix  $A := (a_{ij}) \in K^{m \times n}$  heißt die *Koeffizientenmatrix*, und das Spalten- $m$ -Tupel  $b := (b_i) \in K^m$  heißt die *rechte Seite* des LGS. Als *erweiterte Koeffizientenmatrix* bezeichnen wir die Matrix

$$(A, b) = \begin{pmatrix} a_{11} & \cdots & a_{1n} & b_1 \\ \vdots & & \vdots & \vdots \\ a_{m1} & \cdots & a_{mn} & b_m \end{pmatrix} \in K^{m \times (n+1)}.$$

Für die Lösungsmenge des LGS schreiben wir  $\mathbb{L}(A, b)$ .

**Bemerkung.**

(i) Eine *Lösung* des LGS ist ein Spalten- $n$ -Tupel  $s := \begin{pmatrix} s_1 \\ \vdots \\ s_n \end{pmatrix} \in K^n$  mit

$$\sum_{j=1}^n a_{ij}s_j = b_i \text{ für jedes } i = 1, \dots, m.$$

(ii)  $\mathbb{L}(A, b) \subseteq K^n$ .

(iii) Die „Namen“ der Unbekannten spielen jetzt keine Rolle mehr.

**Beispiel.**  $K = \mathbb{Q}$  und  $n = m = 4$ . Das LGS

$$\begin{array}{cccccccl} x_1 & + & 2x_2 & & & + & x_4 & = & 1 \\ x_1 & + & 2x_2 & + & 2x_3 & + & 3x_4 & = & 5 \\ 2x_1 & + & 4x_2 & & & & & + & 3x_4 & = & 5 \\ & & & & 3x_3 & + & 2x_4 & = & 3 \end{array}$$

hat die erweiterte Koeffizientenmatrix

$$\begin{pmatrix} 1 & 2 & 0 & 1 & 1 \\ 1 & 2 & 2 & 3 & 5 \\ 2 & 4 & 0 & 3 & 5 \\ 0 & 0 & 3 & 2 & 3 \end{pmatrix} \in \mathbb{Q}^{4 \times 5}.$$

Man zeigt mit Äquivalenzumformungen (Rechnung siehe Vorlesung):

$$\mathbb{L} = \left\{ \left( \begin{pmatrix} -2 - 2t \\ t \\ -1 \\ 3 \end{pmatrix} \right) \middle| t \in \mathbb{Q} \right\} = \left\{ \left( \begin{pmatrix} -2 \\ 0 \\ -1 \\ 3 \end{pmatrix} + t \cdot \begin{pmatrix} -2 \\ 1 \\ 0 \\ 0 \end{pmatrix} \right) \middle| t \in \mathbb{Q} \right\} \subseteq \mathbb{Q}^4.$$

### 3.3.4 Matrixmultiplikation und LGS

Wir setzen weiter voraus, dass  $K$  ein Körper ist. Wir wollen hier zeigen, wie man lineare Gleichungssysteme mithilfe von Matrizen formulieren kann.

**Bemerkung a.** Es sei  $A = (a_{ij}) \in K^{m \times n}$  und  $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in K^n$ . Nach



Definition der Matrixmultiplikation (Spezialfall  $l = 1$ ) ist

$$A \cdot x = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} \in K^m \quad \text{mit } b_i = \sum_{j=1}^n a_{ij}x_j \text{ für } i = 1, \dots, m.$$

Aus diesem Grund schreiben wir das LGS über  $K$  mit erweiterter Koeffizientenmatrix  $(A, b) \in K^{m \times (n+1)}$  formal als Matrixgleichung

$$A \cdot x = b,$$

wobei  $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$  ein Spalten- $n$ -Tupel ist, das aus Unbekannten besteht. Eine

Lösung von  $A \cdot x = b$  ist ein Element  $s \in K^n$  mit  $As = b$ . Die Lösungsmenge von  $A \cdot x = b$  ist also gegeben durch

$$\mathbb{L}(A, b) = \{s \in K^n \mid As = b\}.$$

**Beispiel.** Das LGS

$$\begin{array}{rrrrrcl} 2x_1 & + & x_2 & - & x_3 & = & 5 \\ x_1 & - & x_2 & & & = & 1 \end{array}$$

wird als Matrixgleichung geschrieben:

$$\underbrace{\begin{pmatrix} 2 & 1 & -1 \\ 1 & -1 & 0 \end{pmatrix}}_A \cdot \underbrace{\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}}_x = \underbrace{\begin{pmatrix} 5 \\ -1 \end{pmatrix}}_b.$$

**Schreibweise.** Es sei  $A \in K^{m \times n}$ . Wir schreiben

- (i)  $\varphi_A$  für die Abbildung  $\varphi_A : K^n \rightarrow K^m, x \mapsto A \cdot x$ .
- (ii)  $Ax = b$  für das lineare Gleichungssystem mit erweiterter Koeffizientenmatrix  $(A, b)$ .

**Bemerkung b.**(i) Für jedes  $s \in \mathbb{L}(A, b)$  gilt

$$\mathbb{L}(A, b) = s + \mathbb{L}(A, 0) := \{s + u \mid u \in \mathbb{L}(A, 0)\}.$$

(ii) Das Bild von  $\varphi_A$  lautet  $\varphi_A(K^n) = \{b \in K^m \mid Ax = b \text{ lösbar}\}.$ (iii) Die Faser von  $\varphi_A$  zu  $b \in K^m$  lautet

$$\varphi_A^{-1}(\{b\}) = \{s \in K^n \mid As = b\} = \mathbb{L}(A, b).$$

*Beweis.* (i) Es sei  $s \in \mathbb{L}(A, b)$ , d.h.  $s \in K^n$  mit  $As = b$ . Für ein beliebiges  $t \in K^n$  folgt unter Benutzung von Satz 3.2.3(v):

$$\begin{aligned} t \in \mathbb{L}(A, b) &\Leftrightarrow At = b \Leftrightarrow At = As \\ &\Leftrightarrow A(t - s) = 0 \Leftrightarrow t - s \in \mathbb{L}(A, 0) \Leftrightarrow t \in s + \mathbb{L}(A, 0). \end{aligned}$$

□

## 3.4 Der Gauß-Algorithmus

Es sei  $K$  ein beliebiger Körper.

### 3.4.1 Zeilentransformationen

Wir führen die Äquivalenzumformungen eines LGS jetzt nur noch für seine erweiterte Koeffizientenmatrix durch.

**Definition.** Es seien  $m, n \in \mathbb{N}$ . Eine *elementare Zeilentransformation* ist eine Abbildung

$$t : K^{m \times n} \rightarrow K^{m \times n}, \quad A \mapsto t(A),$$

von einem der drei Typen  $\tau, \alpha, \mu$ , wobei  $1 \leq i, j \leq m$  und  $c \in K$ :

- (i)  $\tau_{ij}$ : vertauscht die  $i$ -te und  $j$ -te Zeile von  $A$ .
- (ii)  $\alpha_{ij}(c), i \neq j$ : addiert das  $c$ -fache der  $j$ -ten Zeile zur  $i$ -ten Zeile von  $A$ .
- (iii)  $\mu_i(c)$  mit  $c \neq 0$ : multipliziert die  $i$ -te Zeile von  $A$  mit  $c$ .

Wir schreiben  $A \rightsquigarrow B$ , wenn die Matrix  $B$  aus  $A$  durch eine endliche Folge von elementaren Zeilentransformationen hervorgeht.

**Beispiel.**  $K = \mathbb{Q}, m = 3, n = 4$ .

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 1 & 1 \\ -1 & -1 & 5 & 6 \end{pmatrix} \xrightarrow{\tau_{23}} \begin{pmatrix} 1 & 2 & 3 & 4 \\ -1 & -1 & 5 & 6 \\ 0 & 0 & 1 & 1 \end{pmatrix} \xrightarrow{\alpha_{12}(2)} \begin{pmatrix} -1 & 0 & 13 & 16 \\ -1 & -1 & 5 & 6 \\ 0 & 0 & 1 & 1 \end{pmatrix} \xrightarrow{\mu_2(-1)} \begin{pmatrix} -1 & 0 & 13 & 16 \\ 1 & 1 & -5 & -6 \\ 0 & 0 & 1 & 1 \end{pmatrix}$$

**Bemerkung.**

- (i) Jede elementare Zeilentransformation  $t$  ist *umkehrbar*, d.h. es gibt eine elementare Zeilentransformation  $t'$  so dass gilt:  $t \circ t' = t' \circ t = \text{id}_{K^{m \times n}}$ .
- (ii) Die Relation  $\rightsquigarrow$  ist eine Äquivalenzrelation auf  $K^{m \times n}$ . Gilt  $A \rightsquigarrow B$ , so nennen wir  $A$  und  $B$  *Gauß-äquivalent*.

*Beweis.* Übung (vgl. Satz 3.3.2). **Umkehrung von  $\alpha_{ij}(c)$  ist  $\alpha_{ij}(-c)$ .**  $\square$

**Satz.** Es seien  $(A, b), (A', b') \in K^{m \times (n+1)}$  die erweiterten Koeffizientenmatrizen zweier linearer Gleichungssysteme. Es gilt:

$$(A, b) \rightsquigarrow (A', b') \implies \mathbb{L}(A, b) = \mathbb{L}(A', b').$$

*Beweis.* Elementare Zeilentransformationen der erweiterten Koeffizientenmatrix stellen Äquivalenzumformungen des LGS im Sinne von Satz 3.3.2 dar. Damit gilt

$$\mathbb{L}(A, b) = \mathbb{L}(\tau_{ij}(A, b)) = \mathbb{L}(\alpha_{ij}(c)(A, b)) = \mathbb{L}(\mu_i(c)(A, b)).$$

Die Behauptung ergibt sich durch Induktion nach der Anzahl der angewendeten elementaren Zeilentransformationen.  $\square$

*Übung.* Gilt auch die Umkehrung des Satzes, d.h. folgt aus  $\mathbb{L}(A, b) = \mathbb{L}(A', b')$ , dass  $(A, b) \rightsquigarrow (A', b')$ ?

*Beweis.* Nein, z.B.  $(A, b) = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$  und  $(A', b') = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ .

$\square$

**Frage.** Es seien  $A, A' \in K^{m \times n}$ . Folgt aus  $\mathbb{L}(A, 0) = \mathbb{L}(A', 0)$ , dass  $A \rightsquigarrow A'$ ?

### 3.4.2 Zeilenstufenform

Weiterhin sei  $K$  ein beliebiger Körper.

Ziel: Bringe eine gegebene Matrix durch eine Folge elementarer Zeilentransformationen auf eine „einfache“ bzw. „praktische“ Gestalt (hängt vom Problem ab). Für LGS ist folgende Gestalt „praktisch“.

**Definition.** Es sei  $A \in K^{m \times n}$ . Für  $i = 1, \dots, m$  bezeichne  $z_i$  die  $i$ -te Zeile von  $A$ . Definiere  $k_i \in \{1, \dots, n+1\}$  als die Anzahl der führenden Nullen von  $z_i$  plus 1. Dann sagen wir  $A$  hat *Zeilenstufenform*, wenn

$$k_1 < k_2 < \dots < k_r < k_{r+1} = \dots = k_m = n+1$$

für ein  $0 \leq r \leq m$  ist. Wir nennen  $r$  die *Stufenzahl* von  $A$  und  $k_1, \dots, k_r$  die *Stufenindizes*.

**Bemerkung.** Die Definition von  $k_i$  bedeutet, dass  $z_i$  die Form

$$z_i = (0 \quad \dots \quad 0 \quad \blacksquare \quad \star \quad \dots \quad \star)$$

hat, wobei  $\blacksquare$  und  $\star$  beliebige Einträge aus  $K$  sind, aber  $\blacksquare \neq 0$  ist, und  $\blacksquare$  genau an der  $k_i$ -ten Stelle steht. Enthält  $z_i$  nur Nullen, so ist  $k_i = n+1$ .

Eine Matrix hat demnach Zeilenstufenform, wenn sie so aussieht:

$$\left( \begin{array}{ccc|cccccccccccc} 0 & \dots & 0 & \blacksquare & \star & \dots & \star & \star & \star & \dots & \star & \star & \dots & \star \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 & \blacksquare & \star & \dots & \star & \star & \dots & \star \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & & & & & & \\ \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \ddots & & \star & \vdots & & \vdots \\ 0 & \dots & 0 & 0 & 0 & \dots & 0 & 0 & \dots & 0 & \blacksquare & \star & \dots & \star \\ \hline 0 & \dots & 0 & 0 & \dots & & 0 & 0 & \dots & & 0 & & \dots & 0 \\ \vdots & & \vdots & \vdots & & & \vdots & \vdots & & & \vdots & & \vdots & \\ 0 & \dots & 0 & 0 & \dots & & 0 & 0 & \dots & & 0 & & \dots & 0 \end{array} \right)$$

Die  $\blacksquare$  bilden die „Stufen“ und  $k_i$  ist der Spaltenindex der  $i$ -ten Stufe. Es gilt

$$r = \text{Anzahl Stufen} = \text{Anzahl nicht-Null-Zeilen}.$$

Null-Zeilen dürfen in der Zeilenstufenform nur am unteren Ende der Matrix vorkommen, und es gibt genau  $m - r$  davon. Insbesondere hat die Nullmatrix aus  $K^{m \times n}$  Zeilenstufenform mit Stufenzahl  $r = 0$ .

**Frage.** Es seien  $A, A' \in K^{m \times n}$  in Zeilenstufenform. Folgt aus  $A \rightsquigarrow A'$ , dass  $A$  und  $A'$  gleiche Stufenzahl (Stufenindizes) haben?

### 3.4.3 Gauß-Algorithmus I

Es sei  $K$  ein Körper.

**Satz.** Jede Matrix  $A \in K^{m \times n}$  kann durch eine Folge elementarer Zeilentransformationen (vom Typ  $\tau$  und  $\alpha$ ) auf Zeilenstufenform gebracht werden.

**Bemerkung a.** Der Satz besagt, dass jede Matrix  $A$  Gauß-äquivalent zu einer Matrix in Zeilenstufenform ist. Die Zeilenstufenform ist allerdings nicht eindeutig. Jede Matrix, die Gauß-äquivalent zu  $A$  und in Zeilenstufenform ist, nennen wir *eine Zeilenstufenform von  $A$* .

**Algorithmus** (Gauß). Es sei  $A = (a_{ij}) \in K^{m \times n}$ . Für  $j = 1, \dots, n$  bezeichne  $s_j$  die  $j$ -te Spalte von  $A$ . Die folgenden Schritte überführen  $A$  in Zeilenstufenform.

1. Ist  $A$  die Nullmatrix oder eine  $(1 \times n)$ -Matrix, dann Stopp.
2. Setze  $k := \min\{j \mid 1 \leq j \leq n, s_j \neq 0\}$ .
3. Wähle ein  $i$  mit  $a_{ik} \neq 0$  und wende  $\tau_{1i}$  an. ( $\tau_{11}$  ist erlaubt.)
4. Für jedes  $i = 2, \dots, m$  wende  $\alpha_{i1}(-\frac{a_{ik}}{a_{1k}})$  an.
5. Führe die Schritte 1. – 5. rekursiv mit der Matrix  $(a_{ij})_{\substack{2 \leq i \leq m \\ k < j \leq n}} \in K^{(m-1) \times (n-k)}$  aus.

**Bemerkung b.**

- (i) Der Gauß-Algorithmus ist ein Algorithmus, der Matrizen auf Zeilenstufenform bringt. Das Lösen von linearen Gleichungssystemen ist eine wichtige Anwendung, die wir in den Abschnitten (3.4.4) und (3.4.5) herausarbeiten werden, aber bei weitem nicht die einzige Anwendung.
- (ii) Der Gauß-Algorithmus verändert nicht die Größe einer Matrix. Insbesondere dürfen Null-Zeilen (streng genommen) nicht einfach weggelassen werden. Beim Lösen von (homogenen und inhomogenen) linearen Gleichungssystemen ist das aber trotzdem sinnvoll, da Null-Zeilen redundante Gleichungen repräsentieren.
- (iii) Es folgt eine Erläuterung der einzelnen Schritte:
  1. Jede Nullmatrix und jede  $1 \times n$ -Matrix ist in Zeilenstufenform.
  2. Die  $k$ -te Spalte ist die erste Spalte von links, die nicht komplett aus Nullen besteht.

3. Falls in der  $k$ -ten Spalte ganz oben eine Null steht, dann tausche die oberste Zeile gegen eine andere, so dass das nicht mehr der Fall ist.
  4. Addiere geeignete Vielfache der obersten Zeile zu allen anderen Zeilen, so dass alle anderen Zeilen Null-Einträge in der  $k$ -ten Spalte bekommen.
  5. Mache rekursiv weiter mit der Teilmatrix, die in der zweiten Zeile und der  $k + 1$ -ten Spalte beginnt.
- (iv) Die  $k$ 's aus allen rekursiven Durchläufen sind genau die Stufenindizes  $k_1, \dots, k_r$  der Zeilenstufenform, die am Ende herauskommt. Insbesondere durchläuft der Algorithmus genau  $r$  Rekursionsschritte.
- (v) Nach den Schritten 3. und 4. wird die transformierte Matrix wieder mit  $(a_{ij})$  bezeichnet.

**Beispiel.**

$$\begin{pmatrix} 1 & -2 & 3 & 4 & 2 \\ 2 & -4 & 6 & 9 & 1 \\ -1 & 2 & -1 & -3 & -6 \\ 1 & -2 & 5 & 4 & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & -2 & 3 & 4 & 2 \\ 0 & 0 & 2 & 1 & -4 \\ 0 & 0 & 0 & -1 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

(Rechnung siehe Vorlesung)

### 3.4.4 Homogene LGS

Als Anwendung des Gauß-Algorithmus stellen wir ein Lösungsverfahren für homogene Lineare Gleichungssysteme vor.

**Anwendung** (Lösungsverfahren für homogene LGS).

Gegeben sei ein homogenes LGS mit Koeffizientenmatrix  $A \in K^{m \times n}$ .

1. Bringe  $A$  mittels elementarer Zeilentransformationen auf Zeilenstufenform (z.B. mit Algorithmus 3.4.3).
2. Die  $r$  Unbekannten, die zu den Spalten mit den Stufenindizes  $k_1, \dots, k_r$  gehören, werden *abhängig* genannt, die anderen  $n - r$  Unbekannten werden *frei* genannt.
3. Ersetze die freien Unbekannten durch Parameter  $t_1, \dots, t_{n-r} \in K$ .
4. Löse von unten nach oben nach den abhängigen Unbekannten auf (*Rückwärtssubstitution*).

**Beispiel.** Für die Matrix  $A \in \mathbb{Q}^{4 \times 5}$  aus Beispiel 3.4.3 ergibt sich:

$$A \rightsquigarrow \begin{pmatrix} 1 & -2 & 3 & 4 & 2 \\ 0 & 0 & 2 & 1 & -4 \\ 0 & 0 & 0 & -1 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad \mathbb{L}(A, 0) = \left\{ \begin{pmatrix} 2t_1 - \frac{31}{2}t_2 \\ t_1 \\ \frac{1}{2}t_2 \\ 3t_2 \\ t_2 \end{pmatrix} \mid t_1, t_2 \in \mathbb{Q} \right\}.$$

(Rechnung siehe Vorlesung)

**Bemerkung.**

- (i) Ein homogenes LGS hat immer eine Lösung, nämlich die *triviale Lösung*  $0 \in K^n$ .
- (ii) Hat ein homogenes LGS weniger Gleichungen als Unbekannte ( $m < n$ ), so gibt es nicht-triviale Lösungen.

**Die Umkehrung dieser Aussage gilt nicht!**

Erklärung: In Zeilenstufenform ist immer  $r \leq m$ . Aus  $m < n$  folgt also  $r < n$  bzw.  $n - r > 0$ . Da  $n - r$  die Anzahl der freien Unbekannten ist, gibt es mehr als eine Lösung.

- (iii) Für ein homogenes LGS sind folgende Aussagen äquivalent:

- Das LGS ist nicht-trivial lösbar.
- $\mathbb{L} \neq \{0\}$ .
- Das LGS ist nicht eindeutig lösbar.
- Es gibt freie Unbekannte ( $n - r > 0$ ).

Vorsicht bei der Aussage „das LGS hat unendlich viele Lösungen“: der Körper kann endlich sein!

*Übung.* Es seien  $A, A' \in K^{m \times n}$  in Zeilenstufenform mit  $A \rightsquigarrow A'$ . Man zeige als teilweise Antwort auf Frage 3.4.2: Hat  $A$  die Stufenzahl  $n$ , so hat auch  $A'$  die Stufenzahl  $n$ .

### 3.4.5 Inhomogene LGS

**Bemerkung.** Nicht jedes inhomogene LGS hat eine Lösung. Über jedem Körper ist z.B.  $0 \cdot x = 1$  unlösbar. Allgemein ist die lineare Gleichung  $a \cdot x = b$  genau dann lösbar, wenn  $a \neq 0$  oder  $b = 0$  ist.

**Anwendung** (Lösungsverfahren für inhomogene LGS).

Gegeben sei ein homogenes LGS mit erweiterter Koeffizientenmatrix  $(A, b) \in K^{m \times (n+1)}$ . Man bringe  $(A, b)$  mittels elementarer Zeilentransformationen auf Zeilenstufenform (z.B. mit Algorithmus 3.4.3).

Lösungsentscheidung. Es seien  $k_1, \dots, k_r$  die Stufenindizes der Zeilenstufenform. Die Lösbarkeit kann am Index  $k_r$  abgelesen werden: Ist  $r > 0$  und  $k_r = n + 1$ , so ist das LGS unlösbar. In der Tat hat dann die  $r$ -te Zeile, welche die unterste Nicht-Null-Zeile ist, die Form  $(0 \ \cdots \ 0 \ \blacksquare)$ . Sie entspricht einer nach der Bemerkung unlösbaren Gleichung  $0 \cdot x_1 + \cdots + 0 \cdot x_n = b \neq 0$ . Ist dagegen  $r = 0$  oder  $k_r \leq n$ , so ist das LGS lösbar.

Lösungsmenge. Man betrachtet zunächst nur das homogene System (d.h. man ignoriert die Spalte  $b$  bzw. setzt sie gleich 0). Gemäß Anwendung 3.4.4 definiert man freie und abhängige Unbekannte und bestimmt die Lösungsmenge  $\mathbb{L}(A, 0)$ . Weiter bestimmt man eine beliebige Lösung  $s \in \mathbb{L}(A, b)$ , z.B. indem alle freien Unbekannten gleich 0 gesetzt werden. Die Lösungsmenge ergibt sich dann als

$$\mathbb{L}(A, b) = \{s + u \mid u \in \mathbb{L}(A, 0)\} = s + \mathbb{L}(A, 0). \quad (3.1)$$

*Beweis.* Die Lösungsentscheidung ist klar. Gleichung (3.1) wird in Bemerkung 3.3.4 b bewiesen.  $\square$

**Beispiel.**  $n = m = 4$ .

$$A = \begin{pmatrix} 1 & -2 & 3 & 4 \\ 2 & -4 & 6 & 9 \\ -1 & 2 & -1 & -3 \\ 1 & -2 & 5 & 4 \end{pmatrix} \in \mathbb{Q}^{4 \times 4}, \quad b = \begin{pmatrix} 2 \\ 1 \\ -6 \\ 1 \end{pmatrix} \in \mathbb{Q}^4.$$

Wie in Beispiel 3.4.3 haben wir

$$(A, b) \rightsquigarrow \begin{pmatrix} 1 & -2 & 3 & 4 & 2 \\ 0 & 0 & 2 & 1 & -4 \\ 0 & 0 & 0 & -1 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

Damit ergibt sich

$$\mathbb{L}(A, b) = \left\{ \begin{pmatrix} 2t + \frac{31}{2} \\ t \\ -\frac{1}{2} \\ -3 \end{pmatrix} \mid t \in \mathbb{Q} \right\} = \left\{ \begin{pmatrix} \frac{31}{2} \\ 0 \\ -\frac{1}{2} \\ -3 \end{pmatrix} + t \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \end{pmatrix} \mid t \in \mathbb{Q} \right\}.$$



(Rechnung siehe Vorlesung.) Wie in (3.1) schreiben wir auch:

$$\mathbb{L}(A, b) = \underbrace{\begin{pmatrix} \frac{31}{2} \\ 0 \\ -\frac{1}{2} \\ -3 \end{pmatrix}}_{\text{spezielle Lsg.}} + \underbrace{\mathbb{Q} \cdot \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \end{pmatrix}}_{\mathbb{L}(A, 0)}$$

Wir notieren noch ein Kriterium für die eindeutige Lösbarkeit von linearen Gleichungssystemen.

**Bemerkung.** Es sei  $A'$  eine Zeilenstufenform von  $A$ . Folgende Aussagen sind äquivalent:

- (i)  $Ax = b$  hat für jedes  $b \in K^m$  höchstens eine Lösung.
- (ii)  $Ax = 0$  ist eindeutig lösbar (nur trivial).
- (iii)  $A'$  hat Stufenzahl  $n$ .
- (iv)  $\varphi_A$  ist injektiv. (Zur Definition von  $\varphi_A$  siehe Schreibweise 3.3.4.)

Insbesondere ist dann  $m \geq n$ .

*Beweis.* (i) $\Rightarrow$ (ii): Setze  $b := 0$ .

(ii) $\Rightarrow$ (iii): Da es keine freien Unbekannten geben kann, muss  $A'$  Stufenzahl  $n$  haben.

(iii) $\Rightarrow$ (iv): Da  $A'$  Stufenzahl  $n$  hat, gibt es keine freien Unbekannten, also höchstens eine Lösung.

(iv) $\Rightarrow$ (i): Klar aus der Definition von  $\varphi_A$  (vgl. auch Bemerkung 3.3.4 b).  $\square$

*Übung.* Wie sieht die reduzierte Zeilenstufenform von  $A$  aus, wenn die Aussagen der Bemerkung gelten?

### 3.4.6 Reduzierte Zeilenstufenform

Beim Lösen von (homogenen oder inhomogenen) LGS mit den vorgestellten Verfahren kann man auch die Rückwärtssubstitution durch elementare Zeilentransformationen darstellen.

**Beispiel.** Wir formen die Zeilenstufenform aus Beispiel 3.4.5 weiter um:

$$(A, b) \rightsquigarrow \begin{pmatrix} 1 & -2 & 3 & 4 & 2 \\ 0 & 0 & 2 & 1 & -4 \\ 0 & 0 & 0 & -1 & 3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & -2 & 0 & 0 & \frac{31}{2} \\ 0 & 0 & 1 & 0 & -\frac{1}{2} \\ 0 & 0 & 0 & 1 & -3 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

(Rechnung siehe Vorlesung.)

Hieraus kann man die Lösungsmenge ohne weitere Rechnung direkt ablesen:

$$\mathbb{L}(A, b) = \begin{pmatrix} \frac{31}{2} \\ 0 \\ \frac{1}{2} \\ -3 \end{pmatrix} + \mathbb{Q} \cdot \begin{pmatrix} 2 \\ 1 \\ 0 \\ 0 \end{pmatrix}$$

Um das zu systematisieren machen wir die folgende

**Definition.** Es sei  $A \in K^{m \times n}$ .

- (i)  $A$  hat *reduzierte Zeilenstufenform*, wenn  $A$  Zeilenstufenform hat (vgl. 3.4.2) und zusätzlich gilt:

Für alle  $1 \leq j \leq r$  ist  $a_{1k_j} = a_{2k_j} = \dots = a_{j-1,k_j} = 0, a_{jk_j} = 1$

- (ii)  $A$  hat *Normalform*, wenn  $A$  reduzierte Zeilenstufenform hat und zusätzlich gilt:

Für alle  $1 \leq i \leq r$  ist  $k_i = i$ .

**Bemerkung.**

- (i) Eine Matrix hat reduzierte Zeilenstufenform, wenn sie so aussieht:

$$\left( \begin{array}{cccc|cccccccccccc} 0 & \cdots & 0 & 1 & \star & \cdots & \star & 0 & \star & \cdots & 0 & \star & \cdots & \star \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 1 & \star & \cdots & 0 & \star & \cdots & \star \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & & & & \vdots & & \\ \vdots & & \vdots & \vdots & \vdots & & \vdots & \vdots & \ddots & & 0 & \vdots & & \vdots \\ 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 & \cdots & 0 & 1 & \star & \cdots & \star \\ \hline 0 & \cdots & 0 & 0 & \cdots & & 0 & 0 & \cdots & & 0 & \cdots & 0 & \\ \vdots & & \vdots & \vdots & & & \vdots & \vdots & & & \vdots & \vdots & & \\ 0 & \cdots & 0 & 0 & \cdots & & 0 & 0 & \cdots & & 0 & \cdots & 0 & \end{array} \right)$$

wobei  $\star$  beliebige Einträge aus  $K$  sind.

- (ii) Eine Matrix hat Normalform, wenn sie so aussieht:

$$\left( \begin{array}{ccccc|c} 1 & 0 & 0 & \cdots & 0 & \\ 0 & 1 & 0 & \cdots & 0 & \\ 0 & 0 & \ddots & & \vdots & \star \\ \vdots & \vdots & & 1 & 0 & \\ 0 & 0 & \cdots & 0 & 1 & \\ \hline & & 0 & & & 0 \end{array} \right)$$

wobei  $\star$  ein beliebiger „Block“ ist.

### 3.4.7 Gauß-Algorithmus II

**Satz.** Jede Matrix  $A \in K^{m \times n}$  kann durch eine Folge elementarer Zeilentransformationen (vom Typ  $\tau, \alpha$  und  $\mu$ ) auf reduzierte Zeilenstufenform gebracht werden. Mit Spaltenvertauschungen kann  $A$  weiter auf Normalform gebracht werden.

*Übung.* Man schreibe die einzelnen Schritte eines Algorithmus auf, der eine gegebene Matrix in Zeilenstufenform auf reduzierte Zeilenstufenform bringt (mittels elementarer Zeilentransformationen).

**Bemerkung a.** Beim Lösen von (homogenen und inhomogenen) linearen Gleichungssystemen darf man auch Spalten vertauschen, wenn man über die Zuordnung zwischen Spalten und Unbekannten in geeigneter Weise Buch führt und die „ $b$ -Spalte“ an ihrer Stelle belässt. Spaltenvertauschungen gehören üblicherweise nicht zum Gauß-Algorithmus.

**Beispiel a.** Spaltenvertauschungen können die Rechnung abkürzen. Z.B. kann man

$$(A, b) := \begin{pmatrix} x_1 & x_2 & x_3 & b \\ 2 & 1 & -1 & 2 \\ -2 & 0 & 1 & -6 \\ 1 & 0 & 0 & 3 \end{pmatrix}$$

allein durch Spaltenvertauschungen auf die Zeilenstufenform

$$\begin{pmatrix} x_2 & x_3 & x_1 & b \\ 1 & -1 & 2 & 2 \\ 0 & 1 & -2 & -6 \\ 0 & 0 & 1 & 3 \end{pmatrix}$$

bringen. Weiter kommt man in zwei Schritten zur reduzierten Zeilenstufenform:

$$\begin{pmatrix} x_2 & x_3 & x_1 & b \\ 1 & -1 & 2 & 2 \\ 0 & 1 & -2 & -6 \\ 0 & 0 & 1 & 3 \end{pmatrix} \rightsquigarrow \begin{pmatrix} x_2 & x_3 & x_1 & b \\ 1 & 0 & 0 & -4 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 3 \end{pmatrix}$$

Diese ist gleichzeitig Normalform, und man liest als Lösungsmenge ab:

$$\mathbb{L}(A, b) = \left\{ \begin{pmatrix} 3 \\ -4 \\ 0 \end{pmatrix} \right\}.$$

(Man achte auf die Reihenfolge der Einträge!)

**Beispiel b.** Über  $K = \mathbb{Q}$  sei die folgende erweiterte Koeffizientenmatrix in Normalform gegeben:

$$(A, b) = \begin{pmatrix} 1 & 0 & 0 & 1 & 0 & 2 \\ 0 & 1 & 0 & 2 & 1 & 4 \\ 0 & 0 & 1 & 0 & -1 & 6 \end{pmatrix}.$$

Die Lösungsmenge kann man direkt ohne jede Rechnung ablesen:

$$\mathbb{L}(A, b) = \begin{pmatrix} 2 \\ 4 \\ 6 \\ 0 \\ 0 \end{pmatrix} + \mathbb{Q} \cdot \begin{pmatrix} 1 \\ 2 \\ 0 \\ -1 \\ 0 \end{pmatrix} + \mathbb{Q} \cdot \begin{pmatrix} 0 \\ 1 \\ -1 \\ 0 \\ -1 \end{pmatrix}.$$

(Erläuterung in der Vorlesung.)

**Bemerkung b.** Das Beispiel lässt sich wie folgt verallgemeinern. Es sei

$$A = \left( \begin{array}{c|c} E_r & C \\ \hline 0 & 0 \end{array} \right) \in K^{m \times n}$$

(also  $C \in K^{r \times (n-r)}$ ) in Normalform. Weiter sei

$$b = \begin{pmatrix} b' \\ b'' \end{pmatrix} \in K^m$$

mit  $b' \in K^r$  und  $b'' \in K^{m-r}$ .

(Wir stellen uns vor, dass die Matrix  $(A, b)$  aus der erweiterten Koeffizientenmatrix eines LGS durch elementare Umformungen und Spaltenvertauschungen der ersten  $n$  Spalten entstanden ist. Dann kann aus der Lösungsmenge  $\mathbb{L}(A, b)$  die Lösungsmenge des ursprünglichen LGS gemäß Bemerkung a und Beispiel a bestimmt werden.)

Mit obigen Notationen gilt:

$$(i) \quad \mathbb{L}(A, 0) = \left\{ \begin{pmatrix} C \\ -E_{n-r} \end{pmatrix} t \mid t \in K^{n-r} \right\}.$$

$$(ii) \quad \mathbb{L}(A, b) = \emptyset \Leftrightarrow b'' \neq 0.$$

$$(iii) \quad \text{Ist } b'' = 0, \text{ dann ist } \begin{pmatrix} b' \\ 0 \end{pmatrix} \in \mathbb{L}(A, b).$$

*Beweis.* Aus den Formeln für die Matrixmultiplikation folgt

$$\left( \begin{array}{c|c} E_r & C \\ \hline 0 & 0 \end{array} \right) \cdot \left( \begin{array}{c} C \\ -E_{n-r} \end{array} \right) = 0 \in K^{m \times (n-r)}.$$

Daraus ergibt sich (i). (Alternativ kann Anwendung 3.4.4 zum Beweis von (i) verwendet werden.) Aussage (ii) ist die Lösbarkeitsentscheidung für das LGS mit erweiterter Koeffizientenmatrix  $(A, b)$  in Zeilenstufenform (siehe Anwendung 3.4.5). Die letzte Aussage folgt aus der Gestalt von  $(A, b)$ .  $\square$



# Diskrete Mathematik





# Einleitung

In der Mathematik ist der Begriff „**diskret**“ als gegensätzlich zu „kontinuierlich“ zu verstehen. *Diskret* werden solche Strukturen genannt, die endlich sind oder – falls unendlich – zumindest schrittweise abzählbar; als *kontinuierlich* dagegen solche, die nicht schrittweise abzählbar sind. In diesem Sinne ist z.B. die Zahlenmenge der natürlichen Zahlen  $\{1, 2, 3, \dots\}$  diskret, während die Zahlenmenge der reellen Zahlen (Dezimalbrüche) kontinuierlich ist. Letzteres wird veranschaulicht, indem man sich die reellen Zahlen als eine kontinuierliche Zahlengerade (von  $-\infty$  bis  $+\infty$  mit 0 in der „Mitte“) vorstellt. Auf dieser reellen Zahlengerade sind dann die natürlichen Zahlen als eine abzählbare Folge von Punkten zu finden.

In dieses Schema passen insbesondere die in der Elektro- bzw. Informationstechnik verwendeten Begriffe „digital“ und „analog“. Ein „digitaler Wert“ ist auf einer diskreten Menge definiert (mit den Elementen 0 und 1, also sogar auf einer endlichen Menge), während ein analoger Wert auf einem Kontinuum (z.B. auf einem bestimmten Abschnitt der reellen Zahlengerade) definiert ist.

Unter den mathematischen Disziplinen beschäftigt sich die **Analysis** mit kontinuierlichen Strukturen (insbesondere mit den reellen Zahlen) und die **Diskrete Mathematik** mit diskreten Strukturen. Die diskrete Mathematik, obwohl in der Form des Studiums der natürlichen Zahlen schon im Altertum präsent, wird aber erst seit dem 20. Jahrhundert als eigenständiges Gebiet betrachtet. So wie eine besondere Motivation für die Entwicklung der Analysis auf Anwendungen in der Physik zurückgeht, gilt das gleiche für die diskrete Mathematik und die Informatik. Offensichtlich sind die in der Informatik beschriebenen und untersuchten Objekte wie Digitalcomputer, Programme (Algorithmen), formale Sprachen, etc. diskreter Natur, während die in der klassischen Physik untersuchten Prozesse kontinuierlicher Natur sind (bzw. sich als kontinuierlich vorgestellt werden).

Wichtige diskrete Strukturen bzw. Objekte, die in dieser Vorlesung behandelt werden, sind endliche Mengen und Summen (Kap. Kombinatorik), endliche Graphen (Kap. Graphentheorie), das Zahlssystem der ganzen Zahlen und Polynome (beides Kap. Algebraische Strukturen).



# Kapitel 4

## Kombinatorik

### 4.1 Permutationen und Kombinationen

Es sei  $A$  in diesem Abschnitt eine endliche Menge mit  $|A| = n$ .

#### 4.1.1 Permutationen

**Definition a.** Es sei  $k \in \mathbb{N}, k \leq n$ . Eine  $k$ -Permutation aus  $A$  ist eine geordnete Auswahl von  $k$  verschiedenen Elementen aus  $A$ . Eine  $n$ -Permutation aus  $A$  wird auch kurz *Permutation von  $A$*  genannt.

Mit „geordneter Auswahl“ ist gemeint, dass es auf die Reihenfolge der Auswahl ankommt. Mathematisch ist eine  $k$ -Permutation aus  $A$  ein  $k$ -Tupel über  $A$  (vgl. Definition 1.4.1 b), dessen Einträge paarweise verschieden sind. Dementsprechend werden  $k$ -Permutationen in derselben Schreibweise wie Tupel notiert.

Eine Permutation von  $A$  kann auch als eine „Anordnung“ von  $A$  aufgefasst werden.

**Beispiel a.**

- (i)  $(4, 3, 2)$ ,  $(4, 2, 3)$  und  $(3, 5, 1)$  sind verschiedene 3-Permutationen aus  $\underline{5}$ .
- (ii)  $(1, 2, 1)$  ist keine Permutation.
- (iii)  $(1, 3, 5, 2, 4)$  und  $(5, 4, 3, 2, 1)$  sind Permutationen von  $\underline{5}$ .
- (iv) Die Medaillenverteilung nach einem 100m-Lauf mit 8 Läufern ist eine 3-Permutation aus  $\underline{8}$ .
- (v) Die aktuelle Bundesligatabelle ist eine Permutation von  $\underline{18}$ .

**Definition b.** Für  $n \in \mathbb{N}$  heißt

$$n! := 1 \cdot 2 \cdot \dots \cdot n$$

die *Fakultät von  $n$* . Wir setzen  $0! := 1$ .

**Satz.** Es sei  $k \in \mathbb{N}, k \leq n$ . Die Anzahl der  $k$ -Permutationen aus  $A$  beträgt  $\frac{n!}{(n-k)!}$ . Die Anzahl der Permutationen von  $A$  beträgt  $n!$ .

*Beweis.* Wir bilden alle  $k$ -Tupel  $(a_1, \dots, a_k)$  über  $A$  mit paarweise verschiedenen Einträgen. Dabei gibt es

$$\begin{array}{ll} n & \text{Möglichkeiten für } a_1, \\ n-1 & \text{Möglichkeiten für } a_2, \\ \vdots & \\ n-(k-1) & \text{Möglichkeiten für } a_k, \end{array}$$

also  $n(n-1) \cdots (n-k+1) = \frac{n!}{(n-k)!}$  Möglichkeiten insgesamt.  $\square$

**Beispiel b.**

- (i) Die Anzahl der 2-Permutationen aus  $\underline{3}$  ist  $\frac{3!}{(3-2)!} = 6$ .
- (ii) Es gibt genau  $\frac{8!}{(8-3)!} = 6 \cdot 7 \cdot 8 = 336$  mögliche Medaillenverteilungen (Gold, Silber, Bronze) auf 8 Läufer.
- (iii) Es gibt  $18! \approx 6,4 \cdot 10^{15}$  mögliche Bundesligatabellen aus 18 Mannschaften.

### 4.1.2 Kombinationen

**Definition.** Es sei  $k \in \mathbb{N}, k \leq n$ . Eine  $k$ -Kombination aus  $A$  ist eine ungeordnete Auswahl von  $k$  verschiedenen Elementen aus  $A$ .

Mit „ungeordneter Auswahl“ ist gemeint, dass es auf die Reihenfolge der Auswahl nicht ankommt. Mathematisch ist eine  $k$ -Kombination aus  $A$  eine  $k$ -elementige Teilmenge von  $A$ . Dementsprechend werden  $k$ -Kombinationen in derselben Schreibweise wie Mengen notiert.

**Beispiel a.**

- (i) Es sei  $A = \underline{5} = \{1, 2, 3, 4, 5\}$ . Dann sind  $\{4, 3, 2\} = \{4, 2, 3\}$  und  $\{3, 5, 1\}$  verschiedene 3-Kombinationen aus  $A$ .
- (ii) Ein ausgefüllter Lottoschein ist eine 6-Kombination aus  $\underline{49}$ .

(iii) Die Bundesliga-Absteiger bilden eine 3-Kombination aus 18.

(iv) Eine Skathand ist eine 10-Kombination aus 32.

**Satz.** Es sei  $k \in \mathbb{N}$  mit  $k \leq n$ . Die Anzahl der  $k$ -Kombinationen aus  $A$  beträgt  $\frac{n!}{k!(n-k)!}$ .

*Beweis.* Aus einer  $k$ -Kombination wird durch Anordnung eine  $k$ -Permutation. Jede  $k$ -Kombination kann gemäß Satz 4.1.1 auf  $k!$  Arten angeordnet werden. Z.B.

$$\{2, 3, 4\} \subseteq \underline{5} \xrightarrow{\text{Anordnung}} (2, 3, 4), (2, 4, 3), (3, 2, 4), (3, 4, 2), (4, 2, 3), (4, 3, 2)$$

$$\{1, 3\} \subseteq \underline{5} \xrightarrow{\text{Anordnung}} (1, 3), (3, 1)$$

Also gilt

$$k! \cdot \#k\text{-Kombinationen} = \#k\text{-Permutationen}.$$

Da die rechte Seite nach Satz 4.1.1 gleich  $\frac{n!}{(n-k)!}$  ist, folgt die Behauptung durch Division durch  $k!$ .  $\square$

**Beispiel b.**

- (i) Die Anzahl der 2-Kombinationen aus 4 ist  $\frac{4!}{2!(4-2)!} = 6$ .
- (ii) Es gibt  $\frac{49!}{6!43!} = 13983816$  Möglichkeiten, einen Lottoschein auszufüllen.
- (iii) Es gibt  $\frac{18!}{3!15!} = 816$  Möglichkeiten, drei von 18 Mannschaften absteigen zu lassen.
- (iv) Es gibt  $\frac{32!}{10!22!} \approx 64512240$  mögliche Skathände.

### 4.1.3 Tupel

**Bemerkung.** Es sei  $A$  eine Menge und  $k \in \mathbb{N}$ . Ein  $k$ -Tupel über  $A$  ist eine geordnete Auswahl von  $k$  beliebigen (nicht notwendigerweise verschiedenen) Elementen aus  $A$ .

**Beispiel.**

- (i) Eine natürliche Zahl mit maximal  $k$  Dezimalstellen ist ein  $k$ -Tupel über  $\{0, 1, \dots, 9\}$ .
- (ii) Das Resultat einer Klausur mit  $k$  Teilnehmern und 11 möglichen Noten (von 1.0 bis 5.0) ist ein  $k$ -Tupel über 11. Nummeriert man die Teilnehmer von 1 bis  $k$  und ist  $a_i$  die Note von Teilnehmer  $i$ , dann ist das Resultat das Tupel  $(a_1, \dots, a_k)$ .

(iii) Teilmengen von  $\underline{n}$  können durch  $n$ -Tupel über  $\{0, 1\}$  „kodiert“ werden.

Erklärung: Der Teilmenge  $M \subseteq \underline{n}$  wird das  $n$ -Tupel  $\iota_M := (x_1, \dots, x_n) \in \{0, 1\}^n$  zugeordnet, das folgendermaßen definiert ist:

$$x_i := \begin{cases} 1, & \text{falls } i \in M, \\ 0, & \text{falls } i \notin M. \end{cases}$$

„Kodiert“ bedeutet hier, dass die Abbildung

$$\text{Pot}(\underline{n}) \rightarrow \{0, 1\}^n, \quad M \mapsto \iota_M$$

eine Bijektion ist.

Z.B. werden  $\{2, 4\} \subseteq \underline{5}$  das 5-Tupel  $(0, 1, 0, 1, 0)$ , und  $\{2, 3\} \subseteq \underline{3}$  das 3-Tupel  $(0, 1, 1)$  zugeordnet.

**Satz.** Es sei  $A$  eine endliche Menge mit  $|A| = n$  und  $k \in \mathbb{N}$ . Die Anzahl der  $k$ -Tupel über  $A$  beträgt  $n^k$ .

*Beweis.* Klar. □

**Folgerung.**  $|\text{Pot}(A)| = 2^n$ .

*Beweis.* Der Satz und Beispiel (iii). □

#### 4.1.4 Multimengen

**Definition.** Es sei  $k \in \mathbb{N}$ . Eine  $k$ -Multimenge über  $A$  ist eine ungeordnete Auswahl von  $k$  beliebigen (nicht notwendigerweise verschiedenen) Elementen aus  $A$ .

**Schreibweise.** Eine Multimenge ist eine „Menge mit Wiederholungen“ und wird mit den modifizierten Mengenklammern  $\{^*$  und  $^*\}$  notiert.

**Bemerkung.** Eine  $k$ -Multimenge über  $A$  kann kodiert werden durch ein  $n$ -Tupel über  $\mathbb{N}_0$ , dessen Einträge sich zu  $k$  aufsummieren. Dazu nummeriert man die Elemente von  $A$ , etwa  $A = \{a_1, \dots, a_n\}$ , und gibt im  $i$ -ten Eintrag des Tupels an, wie oft  $a_i$  in der Multimenge vorkommt. Wir nennen dieses Tupel das *Häufigkeitstupel* der Multimenge.

**Beispiel.**

- (i) Ein Lostopf ist eine Multimenge, aber in der Regel keine Menge, da gewisse Lose mehrfach vorkommen können (z.B. Nieten).

- (ii) Das Resultat eines Kniffel-Wurfs (Wurf mit 5 Würfeln gleichzeitig) ist eine 5-Multimenge über 6. Der Wurf  $\begin{smallmatrix} \square & \square \\ \bullet & \bullet \\ \square & \square \end{smallmatrix} \quad \square \quad \begin{smallmatrix} \square \\ \bullet \\ \square \end{smallmatrix}$  bedeutet z.B. die Multimenge  $\{^*2, 1, 4, 1, 2^*\} = \{^*1, 1, 2, 2, 4^*\}$ . Als 6-Tupel über  $\mathbb{N}_0$  geschrieben bedeutet dieser Wurf  $(2, 2, 0, 1, 0, 0)$ .
- (iii) Der Notenspiegel einer Klausur mit  $k$  Teilnehmern und 11 möglichen Noten (von 1.0 bis 5.0) ist eine  $k$ -Multimenge über 11. Der Notenspiegel ist das anonymisierte Resultat der Klausur. Nummeriert man die Teilnehmer von 1 bis  $k$  und ist  $a_i$  die Note von Teilnehmer  $i$ , dann ist der Notenspiegel die  $k$ -Multimenge  $\{^*a_1, \dots, a_k^*\}$ . Üblicherweise wird ein Notenspiegel als Tabelle der Häufigkeiten der einzelnen Noten angegeben. Diese Tabelle ist gerade das oben erwähnte Häufigkeitstupel von  $A$ , ein 11-Tupel über  $\mathbb{N}_0$ .

**Satz.** Es sei  $A$  eine endliche Menge mit  $|A| = n$  und  $k \in \mathbb{N}$ . Die Anzahl der  $k$ -Multimengen über  $A$  beträgt  $\frac{(n+k-1)!}{k!(n-1)!}$ .

*Beweis.* Es sei  $k \in \mathbb{N}$ . Wir zählen die  $n$ -Tupel  $(l_1, \dots, l_n)$  über  $\mathbb{N}_0$  mit  $\sum_{i=1}^n l_i = k$ . Dazu kodieren wir Tupel dieser Art als  $(n+k-1)$ -Tupel über  $\{0, 1\}$ , indem wir für jedes Komma eine Null und für jedes  $l_i > 0$  genau  $l_i$  viele Einsen schreiben. Aus dem Tupel  $(2, 2, 0, 1, 0, 0)$  wird z.B. das Wort 1101100100. Offensichtlich gehört zu jedem  $n$ -Tupel über  $\mathbb{N}_0$ , dessen Einträge sich zu  $k$  aufsummieren, ein  $(n+k-1)$ -Tupel mit  $k$  Einsen und  $n-1$  Nullen. Umgekehrt entsteht jedes  $(n+k-1)$ -Tupel mit  $k$  Einsen und  $n-1$  Nullen aus einem  $n$ -Tupel über  $\mathbb{N}_0$ , dessen Einträge sich zu  $k$  aufsummieren.

Ein  $(n+k-1)$ -Tupel aus  $k$  Einsen und  $n-1$  Nullen ist eindeutig durch die Positionen der  $k$  vielen Einsen gegeben, entspricht also einer  $k$ -Kombination aus  $\underline{n+k-1}$ . Gemäß Satz 4.1.2 lautet die gesuchte Anzahl somit  $\frac{(n+k-1)!}{k!(n-1)!}$ .  $\square$

## 4.2 Binomialkoeffizienten

Es seien in diesem Abschnitt  $n, k \in \mathbb{N}_0$ .

### 4.2.1 Definition und Binomischer Lehrsatz

**Definition.** Für  $k \leq n$  heißt

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

der *Binomialkoeffizient* „ $n$  über  $k$ “.

Nach Satz 4.1.2 ist  $\binom{n}{k}$  gleich der Anzahl der  $k$ -Kombinationen aus einer  $n$ -elementigen Menge. Insbesondere ist  $\binom{n}{k}$  stets eine ganze Zahl. Es gilt  $\binom{n}{0} = \binom{n}{n} = 1$  für alle  $n \in \mathbb{N}_0$  und  $\binom{n}{1} = \binom{n}{n-1} = n$  für alle  $n \in \mathbb{N}$ .

**Schreibweise.** Es sei  $R$  ein kommutativer Ring. Für  $a \in R$  und  $z \in \mathbb{Z}$  schreiben wir

$$z.a := \begin{cases} \underbrace{a + a + \cdots + a}_{z \text{ Summanden}}, & \text{falls } z \in \mathbb{N} \\ 0, & \text{falls } z = 0 \\ -(-z.a), & \text{falls } z < 0 \end{cases}$$

Meist lassen wir den Punkt weg, d.h. wir schreiben  $za$  statt  $z.a$ .

Ist  $z = xy$  für  $x, y \in \mathbb{Z}$ , dann gilt  $z.a = x.(y.a)$  für alle  $a \in R$ .

**Satz** (Binomischer Lehrsatz). *Es sei  $R$  ein kommutativer Ring. Für  $a, b \in R$  und  $n \in \mathbb{N}_0$  gilt*

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

*Beweis.* Wir betrachten den Ausdruck  $(a + b)^n = (a + b) \cdots (a + b)$  und nummerieren die Klammern mit  $1, \dots, n$ . Für jeden der Summanden, die beim Ausmultiplizieren entstehen, wird aus jeder der  $n$  Klammern entweder das  $a$  oder das  $b$  ausgewählt. Bezeichnen wir mit  $I$  die Menge der Nummern der Klammern, aus denen  $a$  ausgewählt wird, so gilt

$$(a + b)^n = \sum_{I \subseteq \underline{n}} a^{|I|} b^{n-|I|}.$$

Hier durchläuft  $I$  alle Teilmengen von  $\underline{n}$ , d.h.  $I$  durchläuft  $\text{Pot}(\underline{n})$ . Also hat die Summe  $|\text{Pot}(\underline{n})| = 2^n$  Summanden. Wir fassen nun jeweils alle Summanden  $a^k b^{n-k}$  für gleiches  $k$  zusammen. Da es genau  $\binom{n}{k}$  Teilmengen  $I \subseteq \underline{n}$  mit  $|I| = k$  gibt, erhalten wir

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

□



**Korollar.** Es sei  $R$  ein Ring und  $p$  eine Primzahl mit  $p \cdot a = 0$  für alle  $a \in R$  (z.B.  $R = \mathbb{F}_p$  der Körper mit  $p$  Elementen). Dann ist

$$(a + b)^p = a^p + b^p$$

für alle  $a, b \in R$ .

*Beweis.* Nach dem binomische Lehrsatz gilt

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

Für  $0 < k < p$  ist

$$\binom{p}{k} = \frac{p \cdot (p-1) \cdots (p-k+1)}{k!}$$

von der Form  $xp$  für ein  $x \in \mathbb{N}$ , also  $\binom{p}{k} \cdot a^k b^{p-k} = 0$ . □

*Übung.* Man zeige mit Hilfe des Binomischen Lehrsatzes die Identität

$$\sum_{k=1}^n (-1)^k \binom{n}{k} = -1 \quad (n \geq 1).$$

### 4.2.2 Das Pascal'sche Dreieck

**Satz.** Für alle  $n, m \in \mathbb{N}_0$  gelten:

- (i)  $\binom{n}{k} = \binom{n}{n-k}$  für alle  $0 \leq k \leq n$ ,
- (ii)  $\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$  für alle  $1 \leq k \leq n-1$ ,
- (iii)  $\sum_{i=0}^k \binom{m}{i} \binom{n}{k-i} = \binom{m+n}{k}$  für alle  $0 \leq k \leq n, m$ . (Vandermonde-Identität)

*Beweis.* (i) Ist klar nach Definition.

(ii) Es sei  $1 \leq k \leq n$ . Wir teilen alle  $k$ -elementigen Teilmengen  $I \subseteq \underline{n}$  auf in solche  $I$ , die  $n$  enthalten, und solche  $I$ , die  $n$  nicht enthalten. Die  $I$ 's der zweiten Art sind Teilmengen von  $\underline{n-1}$ , also gibt es davon  $\binom{n-1}{k}$  viele. Die  $I$ 's der ersten Art sind die Vereinigung von  $\{n\}$  mit einer  $(k-1)$ -elementigen Teilmengen von  $\underline{n-1}$ , also gibt es davon  $\binom{n-1}{k-1}$  viele. Da die Anzahl aller  $k$ -elementigen Teilmengen von  $\underline{n}$  genau  $\binom{n}{k}$  beträgt, folgt die Behauptung.

(iii) Als Übung. □

Die Binomialkoeffizienten lassen sich im sog. *Pascal'schen Dreieck* anordnen:

$n = 0:$				1				
$n = 1:$			1		1			
$n = 2:$			1		2		1	
$n = 3:$		1		3		3	1	
$n = 4:$		1	4		6		4	1
$n = 5:$	1		5	10		10	5	1
$n = 6:$	1	6	15	20	15	6	1	

*Übung.*

- (i) Man zeige Teil (ii) des Satzes durch direktes Einsetzen der Definition und Umformung.
- (ii) Man zeige mittels vollständiger Induktion, dass  $\binom{n}{k}$  eine ganze Zahl ist.  
*Hinweis:* Verwende den Satz.
- (iii) Man zeige den Binomischen Lehrsatz mittels vollständiger Induktion.  
*Hinweis:* Verwende den Satz.
- (iv) Man zeige mittels vollständiger Induktion, dass  $\sum_{i=1}^n i = \binom{n+1}{2}$ .  
*Hinweis:* Verwende den Satz.
- (v) Es seien  $n_1, \dots, n_r \in \mathbb{N}$  und  $n = \sum_{i=1}^r n_i$ . Man zeige:  $\sum_{i=1}^r \binom{n_i}{2} \leq \binom{n-r+1}{2}$ . Ist die Ungleichung scharf?
- (vi) Man zeige mittels vollständiger Induktion:

$$\sum_{k=1}^n (-1)^k \binom{n}{k} = -1 \quad (n \geq 1).$$

*Hinweis:* Verwende den Satz.

- (vii) Man zeige die Vandermonde-Identität mit einem kombinatorischen Beweis. *Hinweis:* Verallgemeinere den Beweis von Teil (ii) des Satzes.
- (viii) Man zeige die Vandermonde-Identität mittels vollständiger Induktion.

## 4.3 Kombinatorische Beweisprinzipien

Wir formulieren nun systematisch einige kombinatorische Beweisprinzipien. Zum Teil wurden diese Prinzipien in den Beweisen der §§1–2 und in den Übungen schon angewendet.

### 4.3.1 Summenregel

**Prinzip.** Für disjunkte, endliche Mengen  $A$  und  $B$  gilt stets

$$|A \cup B| = |A| + |B|.$$

Das Prinzip lässt sich sofort auf endlich viele Mengen verallgemeinern: Für paarweise disjunkte, endliche Mengen  $A_1, \dots, A_r$  gilt stets

$$\left| \bigcup_{i=1}^r A_i \right| = \sum_{i=1}^r |A_i|.$$

**Beispiel.**

- (i) Der Beweis von Satz 4.2.2(ii).
- (ii) Ist  $A \subseteq M$ , so hat die *Komplementärmenge*  $M \setminus A$  die Mächtigkeit  $|M| - |A|$ .

*Übung.*

- (i) Wie viele Teilmengen von  $\underline{6}$  gibt es, die höchstens 4 Elemente enthalten?
- (ii) Man zeige mit Hilfe der Summenregel, dass  $\sum_{i=0}^n \binom{n}{i} = 2^n$ .

### 4.3.2 Produktregel

**Prinzip.** Für zwei beliebige endliche Mengen  $A$  und  $B$  gilt stets

$$|A \times B| = |A| \cdot |B|.$$

Das Prinzip lässt sich sofort auf endlich viele Mengen verallgemeinern: Für endliche Mengen  $A_1, \dots, A_r$  gilt stets

$$|A_1 \times \cdots \times A_r| = \prod_{i=1}^r |A_i|.$$

Insbesondere gilt für jede endliche Menge und jedes  $n \in \mathbb{N}$ :

$$|A^n| = |A|^n.$$

**Beispiel.**

- (i) Der Beweis von Satz 4.1.2.
- (ii) Der Beweis von Satz 4.1.4.

Übung a. Wie viele Tippreihen mit genau 4 Richtigen gibt es für eine feste Lotto-Ziehung?

**Satz.** Es sei  $\mathcal{A}$  eine Multimenge mit  $r$  verschiedenen Elementen  $a_1, \dots, a_r$ , wobei  $a_i$  mit Häufigkeit  $k_i$  auftritt. Sei  $k = k_1 + \dots + k_r$ , die „Mächtigkeit“ von  $\mathcal{A}$ . Die Anzahl der Anordnungen von  $\mathcal{A}$  beträgt dann

$$\frac{k!}{k_1! \cdots k_r!}.$$

1. *Beweis.* Wir betrachten statt  $\mathcal{A}$  zunächst die Menge

$$A = \{a_{11}, \dots, a_{1k_1}, a_{21}, \dots, a_{2k_2}, \dots, a_{r1}, \dots, a_{rk_r}\},$$

in der die  $a_{ij}$  als verschieden angenommen werden. Offensichtlich ist  $|A| = k$ . Nach Satz 4.1.1 gibt es  $k!$  verschiedene Anordnungen von  $A$ . Jede Anordnung von  $\mathcal{A}$  entsteht aus einer Anordnung von  $A$ , indem man, für jedes  $i$ , alle  $a_{ij}$  durch  $a_i$  ersetzt. Diese Ersetzung, durchgeführt für ein festes  $i$ , macht genau  $k_i!$  verschiedene Anordnungen von  $A$  gleich. Nach der Produktregel macht diese Ersetzung, durchgeführt für alle  $i$ , also genau  $k_1! \cdots k_r!$  verschiedene Anordnungen von  $A$  gleich. Daraus ergibt sich die Formel  $\frac{k!}{k_1! \cdots k_r!}$  für die Zahl der Anordnungen von  $\mathcal{A}$ .  $\square$

2. *Beweis.* Jede Anordnung von  $\mathcal{A}$  entsteht auf eindeutige Weise aus folgendem Prozess: Wir wählen eine  $k_1$ -Kombination von  $\underline{k}$ ; diese gibt die Positionen in der Anordnung an, an denen wir  $a_1$  eintragen. (Es muss genau  $k_1$  Positionen in der Anordnung geben, an denen  $a_1$  steht.) Wir wählen dann eine  $k_2$ -Kombination aus den verbleibenden  $k - k_1$  Positionen, um dort  $a_2$  einzutragen, usw. Nach Produktregel gibt es für diesen Prozess genau  $\binom{k}{k_1} \binom{k-k_1}{k_2} \binom{k-k_1-k_2}{k_3} \cdots \binom{k-k_1-\dots-k_{r-1}}{k_r}$ . (Der letzte Faktor ist identisch  $\binom{k_r}{k_r} = 1$ .) Durch Einsetzen und Kürzen ergibt sich die Formel.  $\square$

Übung b. (i) Wie viele verschiedene Wörter kann man durch Anordnung der Buchstaben P, I, Z, Z, A gewinnen?

- (ii) Wie viele Möglichkeiten gibt es, aus 25 Fußballspielern zwei Mannschaftsaufstellungen (erste und zweite Mannschaft) mit je 11 Spielern zu machen?
- (iii) Auf einem Kongress gibt es einen Hauptredner, der dreimal vortragen soll, und drei Nebenredner, die je zweimal vortragen sollen. Wie viele Vortragsprogramme sind möglich?

### 4.3.3 Inklusions-Exklusions-Prinzip

**Prinzip.** Für zwei beliebige endliche Mengen  $A$  und  $B$  gilt stets

$$|A \cup B| = |A| + |B| - |A \cap B|.$$

Das Prinzip lässt sich auf endlich viele Mengen verallgemeinern:

**Satz.** Für endliche Mengen  $A_1, \dots, A_r$  gilt die Formel

$$\begin{aligned} \left| \bigcup_{i=1}^r A_i \right| &= \sum_{k=1}^r (-1)^{k-1} \sum_{1 \leq i_1 < \dots < i_k \leq r} |A_{i_1} \cap A_{i_2} \cap \dots \cap A_{i_k}| \\ &= \sum_{k=1}^r (-1)^{k-1} \sum_{I \subseteq \underline{r}, |I|=k} \left| \bigcap_{i \in I} A_i \right|. \end{aligned}$$

*Beweis.* Setze  $A := \bigcup_{i=1}^r A_i$ . Wir rechnen nach, dass jedes Element  $a \in A$  auf der rechten Seite der Formel tatsächlich genau einmal gezählt wird. Sei also  $a$  ein beliebiges fest gewähltes Element aus  $A$ . Definiere  $I_a$  als die Menge der Indizes  $i$  aller Mengen  $A_i$ , die  $a$  enthalten, d.h.

$$I_a := \{i \in \underline{r} \mid a \in A_i\}.$$

In der Formel werden Ausdrücke der Form  $|\bigcap_{i \in I} A_i|$  für bestimmte Indexmengen  $I \subseteq \underline{r}$  aufsummiert. Sei  $I \subseteq \underline{r}$  eine beliebige solche Indexmenge. Dann wird das Element  $a$  in  $|\bigcap_{i \in I} A_i|$  genau 1-mal gezählt, wenn  $a \in \bigcap_{i \in I} A_i$ , sonst 0-mal. Weiter gilt  $a \in \bigcap_{i \in I} A_i$  genau dann wenn  $i \in I_a$  für alle  $i \in I$ , also genau dann wenn  $I \subseteq I_a$ . Der Anteil von  $a$  an dem Ausdruck

$$\sum_{I \subseteq \underline{r}, |I|=k} \left| \bigcap_{i \in I} A_i \right|$$

für festes  $k$  beträgt somit

$$\sum_{I \subseteq I_a, |I|=k} 1 + \sum_{I \not\subseteq I_a, |I|=k} 0 = \sum_{I \subseteq I_a, |I|=k} 1,$$

also genau die Anzahl der  $k$ -elementigen Teilmengen von  $I_a$ . Diese Zahl hängt nur von  $|I_a|$  ab, und beträgt  $\binom{|I_a|}{k}$  falls  $k \leq |I_a|$  und 0 falls  $k > |I_a|$ . Der Anteil von  $a$  an der gesamten rechten Seite beträgt somit

$$\sum_{k=1}^{|I_a|} (-1)^{k-1} \binom{|I_a|}{k}.$$

Nach Übung 4.2 gilt für alle  $m \in \mathbb{N}$ :

$$\sum_{k=1}^m (-1)^k \binom{m}{k} = -1.$$

Damit ist gezeigt, dass  $a$  auf der gesamten rechten Seite genau einmal gezählt wurde.  $\square$

Für  $r = 3$  ergibt sich z.B.

$$\begin{aligned} |A_1 \cup A_2 \cup A_3| &= +|A_1| + |A_2| + |A_3| \\ &\quad - |A_1 \cap A_2| - |A_1 \cap A_3| - |A_2 \cap A_3| \\ &\quad + |A_1 \cap A_2 \cap A_3|. \end{aligned}$$

**Beispiel.** Wie viele Zahlen zwischen 1 und 100 sind durch 2, 3 oder 5 teilbar? Wir haben die Menge

$$A = \{i \in \mathbb{N} \mid i \leq 100, 2|i \vee 3|i \vee 5|i\}$$

zu zählen. Leicht zählbar sind die Mengen

$$A_n := \{i \in \mathbb{N} \mid i \leq 100, n|i\},$$

für alle  $n \in \mathbb{N}$  ist nämlich  $|A_n| = \lfloor \frac{100}{n} \rfloor$ . Da  $A$  die Vereinigung  $A = A_2 \cup A_3 \cup A_5$  ist, ergibt sich nach dem Inklusions-Exklusions-Prinzip

$$|A| = |A_2| + |A_3| + |A_5| - |A_2 \cap A_3| - |A_2 \cap A_5| - |A_3 \cap A_5| + |A_2 \cap A_3 \cap A_5|.$$

Es bleibt, die verschiedenen Durchschnitte zu zählen. Nun ist jede natürliche Zahl  $i$  genau dann durch 2 und 3 teilbar, wenn sie durch 6 teilbar ist. D.h.  $A_2 \cap A_3 = A_6$ . Analog ergibt sich  $A_2 \cap A_5 = A_{10}$ ,  $A_3 \cap A_5 = A_{15}$ ,  $A_2 \cap A_3 \cap A_5 = A_{30}$ . (Man beachte, dass 2, 3 und 5 Primzahlen sind; allgemein gilt  $A_n \cap A_m = A_{\text{kgV}(n,m)}$  für beliebige  $n, m \in \mathbb{N}$ .) Also

$$\begin{aligned} |A| &= |A_2| + |A_3| + |A_5| - |A_6| - |A_{10}| - |A_{15}| + |A_{30}| \\ &= 50 + 33 + 20 - 16 - 10 - 6 + 3 = 74. \end{aligned}$$

*Übung a.* Die Bevölkerung von Aachen, die arbeitet oder studiert, betrage 150000. Wenn davon 20% studieren und 90% arbeiten, Wie viele Aachener Studenten arbeiten dann neben ihrem Studium?

*Übung b.* Es seien  $n_1, \dots, n_r \in \mathbb{N}$  und  $n = \sum_{i=1}^r n_i$ . Man gebe einen kombinatorischen Beweis für die Ungleichung  $\sum_{i=1}^r \binom{n_i}{2} \leq \binom{n-r+1}{2}$  aus Übung 4.2.2v. *Hinweis:* Betrachte Mengen  $A_1, \dots, A_r$  Mengen mit  $|A_i| = n_i$ , für die es ein Element  $a$  gibt, so dass für alle  $i, j \in \underline{r}$  mit  $i \neq j$  gilt:  $A_i \cap A_j = \{a\}$ .

*Beweis.* Nach dem Inklusions-Exklusions-Prinzip ist  $|\bigcup_{i=1}^r A_i| = n - r + 1$ . Bezeichne  $\text{Pot}_2(A)$  die Menge der 2-elementigen Teilmengen von  $A$ . Zu zeigen ist:  $\sum_{i=1}^r |\text{Pot}_2(A_i)| \leq |\text{Pot}_2(\bigcup_{i=1}^r A_i)|$ . Trivialerweise ist  $\bigcup_{i=1}^r \text{Pot}_2(A_i) \subseteq \text{Pot}_2(\bigcup_{i=1}^r A_i)$ . Wegen  $|A_i \cap A_j| = 1$  für  $i \neq j$  sind die Mengen  $\text{Pot}_2(A_i)$  für  $i = 1, \dots, r$  paarweise disjunkt. Die Behauptung folgt also mit der Summenregel.  $\square$

### 4.3.4 Schubfachprinzip

**Prinzip.** Verteilt man  $n$  Elemente auf  $m$  Schubladen und ist  $n > m$ , so enthält eine Schublade mindestens zwei Elemente.

**Beispiel.** In jeder Menge von 13 Personen gibt es zwei, die im gleichen Monat Geburtstag haben.

## 4.4 Stirling'sche Zahlen

Die Binomialkoeffizienten wurden eingeführt, da sie beim Zählen von Teilmengen bzw. Multimengen fester Mächtigkeit auftreten. Die Stirling'schen Zahlen stellen zwei weitere Arten von Zählkoeffizienten dar. Sie treten auf beim Zählen von Partitionen mit fester Anzahl von Teilen bzw. beim Zählen von Permutationen mit fester Zykelzahl.

### 4.4.1 Stirling-Zahlen zweiter Art

**Definition.** Es seien  $n, k \in \mathbb{N}_0$ . Wir definieren

$$S_{n,k} := \text{Anzahl der Partitionen von } \underline{n} \text{ mit genau } k \text{ Teilen.}$$

Die Zahlen  $S_{n,k}$  heißen *Stirling-Zahlen zweiter Art*. Partitionen mit  $k$  Teilen nennen wir auch kurz *k-Partitionen*.

**Beispiel.** Wie viele Möglichkeiten gibt es,  $n$  Studenten auf  $k$  Tutoriengruppen aufzuteilen, wobei keine Gruppe leer bleiben soll? Eine solche Aufteilung ist eine  $k$ -Partition von  $\underline{n}$ , somit gibt es  $S_{n,k}$  Möglichkeiten.

**Bemerkung.** Für alle  $n, k \in \mathbb{N}_0$  gelten:

- (i)  $S_{n,n} = 1$ ,
- (ii)  $S_{n,0} = 0$  falls  $n > 0$ ,
- (iii)  $S_{n,k} = 0$  falls  $k > n$ .

- (i) Es gibt genau eine  $n$ -Partition von  $\underline{n}$ . Das gilt auch für  $n = 0$ , da es genau eine Partition der leeren Menge gibt, und die hat 0 Teile.
- (ii) Eine Partition einer nicht-leeren Menge muss mindestens 1 Teil haben.
- (iii) Eine Partition von  $\underline{n}$  kann höchstens  $n$  Teile haben.

☐

*Beweis.* Es sei  $T_1 \cup \dots \cup T_k = \underline{n}$  eine  $k$ -Partition von  $\underline{n}$ . Wir nehmen o.B.d.A. an, dass  $n$  in  $T_k$  liegt (die Nummerierung der Teile spielt keine Rolle). Entfernt man  $n$  aus  $T_k$  und  $\underline{n}$ , so bekommt man  $T_1 \cup \dots \cup T_{k-1} \cup (T_k \setminus \{n\}) = \underline{n-1}$ . Je nachdem, ob  $T_k \setminus \{n\}$  leer ist oder nicht, ist dies eine  $(k-1)$ -Partition oder eine  $k$ -Partition von  $\underline{n-1}$ . Umgekehrt entsteht jede  $k$ -Partition von  $\underline{n}$  auf eine der folgenden Arten:

- Hinzufügen des Teiles  $\{n\}$  zu einer  $(k-1)$ -Partition von  $\underline{n-1}$ ,
- Hinzufügen des Elementes  $n$  zu einem der Teile einer  $k$ -Partition von  $n-1$ .

Keine Partition kann auf beide Arten entstehen, denn bei a) liegt  $n$  stets in einem Teil der Mächtigkeit 1, bei b) stets in einem Teil der Mächtigkeit  $> 1$ . Folglich ist die Anwendung der Summenregel erlaubt. Bei a) gibt es  $S_{n-1,k-1}$  viele  $(k-1)$ -Partitionen von  $n-1$ , die jeweils auf eindeutige Art um den Teil  $\{n\}$  ergänzt werden. Bei b) gibt es  $S_{n-1,k}$  viele  $k$ -Partitionen von  $n-1$ , bei denen auf jeweils  $k$  verschiedene Arten das Element  $n$  zu einem der Teile hinzugefügt wird. Nach Produktregel entstehen durch b) also  $kS_{n-1,k}$  viele  $k$ -Partitionen von  $n$ . Mit der Summenregel ergibt sich schließlich die Formel  $S_{n,k} = S_{n-1,k-1} + kS_{n-1,k}$ .  $\square$

Die Zahlen  $S_{n,k}$  lassen sich im sog. *Stirling-Dreieck zweiter Art* anordnen:

$n = 0:$				1				
$n = 1:$				0		1		
$n = 2:$			0		1		1	
$n = 3:$			0		1		3	
$n = 4:$			0		1		7	
$n = 5:$			0		1		15	
$n = 6:$			0		1		31	



*Übung.* Man zeige:

- (i) Die Anzahl der surjektiven Abbildungen  $\underline{m} \rightarrow \underline{k}$  beträgt  $k! \cdot S_{m,k}$ .
- (ii) Es gilt  $\sum_{k=0}^n S_{m,k} \cdot \frac{n!}{(n-k)!} = n^m$ .  
*Tipp:*  $n^m$  ist die Anzahl aller Abbildungen  $\underline{m} \rightarrow \underline{n}$ .

#### 4.4.2 Stirling-Zahlen erster Art

**Definition.** Es seien  $n, k \in \mathbb{N}_0$ . Wir definieren

$$s_{n,k} := \text{Anzahl der Permutationen aus } S_n \text{ mit Zykelzahl } k.$$

Die Zahlen  $s_{n,k}$  heißen *Stirling-Zahlen erster Art*.

**Beispiel.** Bei einem Treffen von  $n$  Philosophen teilen sich diese in  $k$  Diskussionsgruppen auf (Gruppen mit nur einer Person sind erlaubt). Die Teilnehmer jeder Gruppe setzen sich im Kreis hin und philosophieren über ein Thema. Wie viele mögliche Sitzordnungen gibt es? Antwort:  $s_{n,k}$ .

**Bemerkung.** Für alle  $n, k \in \mathbb{N}_0$  gelten:

- (i)  $s_{n,n} = 1$ ,
- (ii)  $s_{n,0} = 0$  falls  $n > 0$ ,
- (iii)  $s_{n,k} = 0$  falls  $k > n$ .

*Beweis.*

- (i) Hat  $\pi \in S_n$  die Zykelzahl  $n$ , so müssen alle Zykeln die Länge 1 haben, also ist  $\pi = \text{id}$ . Das gilt auch für  $n = 0$ , denn das einzige Element aus  $S_0$  hat Zykelzahl 0.
- (ii) Die Zykelzahl eines Elementes von  $S_n$  mit  $n > 0$  ist stets  $> 0$ .
- (iii) Die Zykelzahl eines Elementes von  $S_n$  kann höchstens  $n$  betragen.

□

**Satz.** Für alle  $n, k \in \mathbb{N}$  gilt  $s_{n,k} = s_{n-1,k-1} + (n-1)s_{n-1,k}$ .

*Beweis.* Eine Modifikation des Beweises von Satz 4.4.1 (Übung). □

Die Zahlen  $s_{n,k}$  lassen sich im sog. *Stirling-Dreieck erster Art* anordnen:

$n = 0:$								1													
$n = 1:$								0					1								
$n = 2:$								0				1			1						
$n = 3:$								0			2			3		1					
$n = 4:$								0		6			11		6		1				
$n = 5:$								0		24			50		35		10		1		
$n = 6:$								0		120			274		225		85		15		1

Übung.

- (i) Man führe den Beweis des Satzes aus.
- (ii) Man zeige  $\sum_{k=0}^n s_{n,k} = n!$ .

# Kapitel 5

## Graphentheorie

### 5.1 Grundbegriffe

#### 5.1.1 Ungerichtete Graphen

**Definition a.** Ein (ungerichteter) *Graph* ist ein Paar  $G = (V, E)$ , bestehend aus einer endlichen Menge  $V$  und einer Menge  $E$  von zweielementigen Teilmengen von  $V$ . Die Elemente von  $V$  werden *Knoten* (engl. *vertex*) genannt, die Elemente von  $E$  *Kanten* (engl. *edge*). Es heißt  $n_G := |V|$  die *Knotenzahl* und  $m_G := |E|$  die *Kantenzahl* von  $G$ .

**Bemerkung a.** Das mathematische Modell für eine Kante zwischen den Knoten  $u, v \in V$  ist hier die zweielementige Teilmenge  $\{u, v\} = \{v, u\} \subseteq V$ . Das bedeutet, dass unsere Definition keine sog. „Schlingen“ zulässt, d.h. Kanten von einem Knoten zu sich selbst. Für die Kante  $\{u, v\}$  verwenden wir alternativ auch die Schreibweise  $uv$  bzw.  $vu$ .

Ein weiteres mögliches mathematisches Modell für die Kanten ist, die Kantenmenge als eine symmetrische, antireflexive Relation auf der Knotenmenge aufzufassen.

Erlaubt ist der Graph  $G = (\emptyset, \emptyset)$ .

**Bemerkung b.** Andere verbreitete Definitionen von Graphen erlauben gerichtete Kanten, Schlingen, Mehrfachkanten, gewichtete Kanten, gefärbte Kanten, usw. Entsprechend muss das mathematische Modell für die Kantenmenge variiert werden.

*Übung a.* Jede Relation auf einer Menge  $V$  kann als ein gerichteter Graph (mit erlaubten Schlingen) veranschaulicht werden. Man mache sich klar, was jede einzelne der folgenden Eigenschaften der Relation für das Aussehen dieses Graphen bedeuten: symmetrisch, antisymmetrisch, reflexiv, antireflexiv, transitiv, Äquivalenzrelation, Totalordnung.

*Übung b.* Was wäre ein mathematisches Modell für einen ungerichteten Graphen mit Mehrfachkanten bzw. mit gewichteten Kanten?

In diesem und den folgenden Abschnitten sei  $G = (V, E)$  stets ein Graph.

**Definition b.**

- (i) Ist  $uv \in E$  eine Kante, so werden  $u$  und  $v$  die *Endknoten* von  $uv$  genannt. In diesem Fall heißen  $u$  und  $v$  *adjazent* oder *benachbart*, sowie  $u$  *Nachbar* von  $v$  und umgekehrt.
- (ii) Die Menge aller Nachbarn von  $v \in V$  wird mit  $\Gamma(v) := \Gamma_G(v)$  bezeichnet.
- (iii)  $G$  heißt *vollständiger* Graph, wenn je zwei beliebige Knoten adjazent sind, also genau dann, wenn  $m_G = \binom{n_G}{2}$ .
- (iv) Eine Kante  $e \in E$  heißt *inzident* zu einem Knoten  $v \in V$ , wenn  $v$  ein Endknoten von  $e$  ist.
- (v) Zwei verschiedene Kanten heißen *inzident*, wenn sie einen gemeinsamen Endknoten haben.

*Übung c.* In jedem Graph  $G$  gilt  $m_G \leq \binom{n_G}{2}$ .

### 5.1.2 Datenstruktur für Graphen

Es sei  $G = (V, E)$  ein Graph mit  $V = \{1, \dots, n\}$  und  $E = \{e_1, \dots, e_m\}$ .

**Definition.** Die *Adjazenzmatrix* von  $G$  ist die Matrix

$$A := \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \in \{0, 1\}^{n \times n} \text{ mit } a_{ij} := \begin{cases} 1 & \text{falls } ij \in E, \\ 0 & \text{falls } ij \notin E. \end{cases}$$

Die *Adjazenzliste* von  $G$  ist die Liste  $\Gamma := (\Gamma(1), \Gamma(2), \dots, \Gamma(n))$ .

Die *Inzidenzmatrix* von  $G$  ist die Matrix

$$B := \begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1m} \\ \vdots & & & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nm} \end{pmatrix} \in \{0, 1\}^{n \times m} \text{ mit } b_{ij} := \begin{cases} 1 & \text{falls } i \in e_j, \\ 0 & \text{falls } i \notin e_j. \end{cases}$$

**Bemerkung.** Die Adjazenzmatrix enthält 0 entlang der Diagonalen von  $a_{11}$  bis  $a_{nn}$  und ist spiegelsymmetrisch zu dieser Diagonalen. Die  $j$ -te Spalte der Inzidenzmatrix enthält genau zwei Einsen, nämlich zu den beiden Endknoten der Kante  $e_j$ .

**Beispiel.** Siehe Vorlesung.

### 5.1.3 Teilgraphen

Es sei  $G = (V, E)$  ein Graph.

**Definition.** Ein Graph  $G' = (V', E')$  wird *Teilgraph* von  $G$  genannt, geschrieben  $G' \leq G$ , wenn  $V' \subseteq V$  und  $E' \subseteq E$ .

**Beispiel.** Ist  $V' \subseteq V$ , so wird durch  $E' := E \cap \{uv \mid u, v \in V'\}$  ein Teilgraph  $(V', E')$  von  $G$  definiert. Dieser wird der *auf  $V'$  induzierte Teilgraph von  $G$*  genannt, geschrieben  $G|_{V'}$ .

### 5.1.4 Der Grad

Es sei  $G = (V, E)$  ein Graph.

**Definition.** Wir definieren den *Grad* von  $v \in V$  als  $\deg(v) := |\Gamma(v)|$ , also die Anzahl der Nachbarn von  $v$  bzw. die Anzahl der zu  $v$  inzidenten Kanten. Knoten mit Grad 0 heißen *isoliert*.

**Bemerkung.**  $\sum_{v \in V} \deg(v) = 2m_G$ .

**Folgerung.** In jedem Graphen ist die Anzahl der Knoten mit ungeradem Grad gerade.

**Beispiel.** Die Anzahl der Personen auf einer Party, die einer ungeraden Zahl von Gästen die Hand geben, ist gerade. (Aufgrund dieses Beispiel wird die Folgerung auch „Handschlagslemma“ genannt.)

### 5.1.5 Kantenzüge, Pfade, Kreise, Touren

Es sei  $G = (V, E)$  ein Graph.

**Definition.** Es sei  $l \in \mathbb{N}_0$ .

- (i) Ein *Kantenzug der Länge  $l$  in  $G$*  ist ein Tupel  $(v_0, v_1, \dots, v_l)$  von Knoten mit  $v_i v_{i+1} \in E$  für alle  $i = 0, \dots, l-1$ . Zu einem Kantenzug  $(v_0, \dots, v_l)$  sagen wir auch genauer *Kantenzug von  $v_0$  nach  $v_l$*  oder  *$v_0$ - $v_l$ -Kantenzug*, und die Knoten  $v_0, v_l$  werden sein *Anfangs-* bzw. *Endknoten* genannt. Der Kantenzug heißt *geschlossen* falls  $v_0 = v_l$ .
- (ii) Ein Kantenzug  $(v_0, \dots, v_l)$  heißt *Pfad der Länge  $l$  in  $G$* , falls die Knoten  $v_0, \dots, v_l$  paarweise verschieden sind. Zu einem Pfad  $(v_0, \dots, v_l)$  sagen wir auch genauer *Pfad von  $v_0$  nach  $v_l$*  oder  *$v_0$ - $v_l$ -Pfad*, und die Knoten  $v_0, v_l$  werden sein *Anfangs-* bzw. *Endknoten* genannt,

- (iii) Ein *Kreis der Länge  $l$  in  $G$*  ist ein geschlossener Kantenzug  $(v_0, \dots, v_l)$ , für den  $l \geq 3$  und  $(v_0, \dots, v_{l-1})$  ein Pfad ist.
- (iv) Eine *Tour der Länge  $l$  in  $G$*  ist ein geschlossener Kantenzug  $(v_0, \dots, v_l)$ , für den die Kanten  $v_0v_1, v_1v_2, \dots, v_{l-1}v_l$  paarweise verschieden sind.

**Bemerkung.**

- (i) Für jeden Knoten  $v \in V$  ist  $(v)$  ein  $v$ - $v$ -Pfad der Länge 0.
- (ii) Jeder Kreis ist eine Tour, aber nicht umgekehrt.
- (iii) Ist  $(v_0, \dots, v_l)$  ein Kreis, so ist auch  $(v_1, \dots, v_{l-1}, v_0, v_1)$  ein Kreis. Diese beiden Kreise sind formal verschieden! Liest man das Tupel  $(v_0, \dots, v_l)$  aber als Zykel  $(v_0 \ v_1 \ \dots \ v_{l-1})$ , also als eine Permutation von  $V$ , so liefern beide Kreise denselben Zykel.
- (iv) Ist  $(v_0, \dots, v_l)$  ein Kreis, so ist auch  $(v_l, \dots, v_0)$  ein Kreis. Diese beiden Kreise sind formal verschieden!

**Beispiel.** Siehe Vorlesung.

*Übung.* Ist eine Kante  $e \in E$  Teil von zwei *verschiedenen* Kreisen von  $G = (V, E)$ , so besitzt auch  $(V, E \setminus \{e\})$  einen Kreis. Hier ist zunächst geeignet zu definieren, was es heißt, dass  $e$  Teil eines Kreises ist, und wann zwei Kreise als gleich anzusehen sind.

Alternativ: Definiere eine *Kreiszahl*  $k$  von  $e$  und von  $G$  und zeige  $k_{G'} = k_G - k_e$  für  $G' = (V, E \setminus e)$ .

### 5.1.6 Zusammenhang und Komponenten

Es sei  $G = (V, E)$  ein Graph.

**Definition.** Die *Zusammenhangsrelation*  $\sim$  auf  $V$  wird definiert durch

$$u \sim v :\Leftrightarrow \text{ es gibt einen } u\text{-}v\text{-Kantenzug in } G.$$

$G$  heißt *zusammenhängend*, falls  $u \sim v$  für alle  $u, v \in V$ , anderenfalls *unzusammenhängend*.

**Bemerkung a.** Offensichtlich ist  $\sim$  eine Äquivalenzrelation (Übung). Wir lesen  $u \sim v$  auch als „ $u$  ist verbunden mit  $v$ “ oder „ $u$  und  $v$  hängen zusammen“. Für alle  $u, v \in V$  gilt:

$$u \sim v \Leftrightarrow \text{ es gibt einen } u\text{-}v\text{-Pfad in } G.$$

*Beweis.*  $\Rightarrow$ : Sei  $u \sim v$  und sei  $(v_0, v_1, \dots, v_l)$  mit  $v_0 = u$  und  $v_l = v$  ein  $u$ - $v$ -Kantenzug in  $G$  von minimaler Länge  $l$ . Angenommen  $(v_0, v_1, \dots, v_l)$  ist kein Pfad, d.h.  $v_i = v_j$  für geeignete  $0 \leq i < j \leq l$ . Dann ist  $(v_0, \dots, v_i, v_{j+1}, \dots, v_l)$  ein  $u$ - $v$ -Kantenzug der Länge  $l - (j - i) < l$  im Widerspruch zur Minimalität von  $l$ . Also ist die Annahme falsch und  $(v_0, v_1, \dots, v_l)$  ein Pfad.

$\Leftarrow$ : trivial.  $\square$

*Übung.* Besitzt  $G$  einen Knoten vom Grad  $n_G - 1$ , so ist  $G$  zusammenhängend.

**Definition.** Die *Zusammenhangskomponenten* oder kurz *Komponenten* von  $G$  sind die induzierten Teilgraphen  $G|_U$ , wobei  $U$  die Äquivalenzklassen von  $V$  bzgl.  $\sim$  durchläuft. Die Anzahl der Äquivalenzklassen von  $\sim$  bezeichnen wir als *Komponentenzahl*  $r_G$  von  $G$ . Es heißt  $G_v := G|_{[v]_\sim}$  die *Zusammenhangskomponente* von  $v \in V$ . Komponenten, die aus einem einzelnen Knoten bestehen, nennen wir *trivial*.

**Beispiel.** Siehe Vorlesung.

**Bemerkung b.**  $G$  ist genau dann zusammenhängend, wenn  $r_G \leq 1$ . Eine Komponente ist genau dann trivial, wenn sie keine Kanten enthält. Ein Knoten ist genau dann isoliert, wenn seine Zusammenhangskomponente trivial ist.

### 5.1.7 Die Zahlen $n_G, m_G, r_G$

Es sei  $G = (V, E)$  ein Graph.

**Lemma.** Für alle  $u, v \in V$  gilt:

$$(i) \quad r_{(V, E)} - 1 \leq r_{(V, E \cup \{uv\})} \leq r_{(V, E)}.$$

$$(ii) \quad r_{(V, E \setminus \{uv\})} - 1 \leq r_{(V, E)} \leq r_{(V, E \setminus \{uv\})}.$$

*Beweis.* i) Die neue Kante  $uv$  kann höchstens zwei Komponenten verbinden. ii) folgt aus i).  $\square$

**Satz a** (Untere Schranke für  $m_G$ ).  $m_G \geq n_G - r_G$ .

*Beweis.* Wir führen eine Induktion nach  $m_G$ . In einem Graph ohne Kanten ( $m_G = 0$ ) sind alle Komponenten trivial, also  $r_G = n_G$ . Sei nun  $m_G > 0$  und die Behauptung für kleineres  $m_G$  bereits bewiesen. Wähle ein  $e \in E$  und setze  $G' := (V, E \setminus \{e\})$ . Nach Teil (ii) des Lemmas gilt  $r_{G'} - 1 \leq r_G$ . Mit der Induktionsvoraussetzung, angewendet auf  $G'$ , folgt  $n_G = n_{G'} \leq m_{G'} + r_{G'} = m_G - 1 + r_{G'} \leq m_G + r_G$ .  $\square$

**Folgerung a.** Ist  $G$  zusammenhängend, dann ist  $m_G \geq n_G - 1$ .

**Satz b** (Obere Schranke für  $m_G$ ).  $m_G \leq \binom{n_G+1-r_G}{2}$ .

*Beweis.* Für  $r_G = 1$  ist die Aussage  $m_G \leq \binom{n_G}{2}$  klar. Für allgemeines  $r_G$  folgt sie daraus durch Summation über die Komponenten mittels Übung 4.2.2v.  $\square$

**Folgerung b.** Ist  $G$  unzusammenhängend, so gilt  $m_G \leq \binom{n_G-1}{2}$ .

*Übung.* Man zeige Folgerung b mittels vollständiger Induktion nach  $n_G$ .

*Beweis.* Sei  $n = n_G$ . Für  $n = 1$  ist die Aussage trivial, für  $n = 2$  lautet sie  $0 \leq 0$  (Induktionsanfang). Sei nun  $n \geq 3$ . Sei oBdA  $V = \underline{n}$ .

Falls  $n$  isoliert ist, so folgt  $m_G \leq \binom{n-1}{2}$  aus der Betrachtung von  $G|_{\underline{n-1}}$ . Sei also  $n$  nicht isoliert und  $G$  unzusammenhängend. Dann ist a)  $\deg n \leq n - 2$  (Lemma), und b)  $G|_{\underline{n-1}}$  unzusammenhängend (sonst  $G$  zusammenhängend). Mit Induktionsvoraussetzung, angewendet auf  $G' := G|_{\underline{n-1}}$ , folgt  $m_G = m_{G'} \leq \binom{n-2}{2} + (n-2) = \frac{(n-2)(n-3)}{2} + (n-2) = \binom{n-1}{2}$ .  $\square$

### 5.1.8 Brücken

**Bemerkung a.** Es seien  $e = uv \in E$ ,  $G' = (V, E \setminus \{e\})$ . Folgende Aussagen sind äquivalent:

- (i)  $u \not\sim v$  in  $G'$ ,
- (ii)  $r_{G'} > r_G$ .

**Definition.** Eine Kante  $e = uv \in E$  heißt *Brücke* von  $G$ , wenn die Bedingungen aus Bemerkung a erfüllt sind, sonst *Nicht-Brücke* von  $G$ .

**Beispiel.** Ist  $\deg u = 1$ , so ist die einzige zu  $u$  inzidente Kante eine Brücke. Weitere Beispiele inkl. Bilder siehe Vorlesung.

**Bemerkung b.** Es seien  $e = uv \in E$ ,  $G' = (V, E \setminus \{e\})$ . Folgende Aussagen sind äquivalent:

- (i)  $e$  ist keine Brücke von  $G$ ,
- (ii)  $u \sim v$  in  $G'$ ,
- (iii)  $r_{G'} = r_G$ ,
- (iv) es gibt einen  $u$ - $v$ -Kantenzug in  $G$ , der nicht über  $e$  führt,



(v) es gibt einen  $u$ - $v$ -Pfad in  $G$ , der nicht über  $e$  führt,

(vi)  $e$  ist Teil eines Kreises in  $G$ .

*Beweis.* Die Äquivalenz (iv)  $\Leftrightarrow$  (v) benutzt Bemerkung (5.1.6). Der Rest ist trivial.  $\square$

**Satz.** Ist  $u \in V$  zu  $l$  Brücken inzident ( $l \in \mathbb{N}$ ), so besitzt  $G$  mindestens  $l$  von  $u$  verschiedene Knoten von ungeradem Grad.

**Folgerung.** Haben in einem Graphen alle Knoten geraden Grad, so besitzt er keine Brücken.

*Beweis des Satzes.* Seien  $e_1, \dots, e_l \in E$  zu  $u$  inzidente Brücken in  $G$ ,  $e_i = uv_i$ . Setze  $G' = (V, E \setminus \{e_1, \dots, e_l\})$ . In  $G'$  liegen die Knoten  $v_1, \dots, v_l$  in verschiedenen Zusammenhangskomponenten  $G'_{v_i}$ . Behauptung: jede der Komponenten  $G'_{v_i}$  enthält einen Knoten von ungeradem Grad in  $G$ . In der Tat, falls  $\deg_G(v_i)$  gerade ist, so ist  $\deg_{G'}(v_i)$  ungerade. Nach dem Handschlagslemma, angewendet auf  $G'_{v_i}$ , enthält dann  $G'_{v_i}$  einen weiteren Knoten  $v'_i \neq v_i$  mit  $\deg_{G'}(v'_i)$  ungerade. Wegen  $v'_i \neq v_i$  ist aber  $\deg_G(v'_i) = \deg_{G'}(v'_i)$ , also ungerade.  $\square$

## 5.2 Distanz und gewichtete Graphen

### 5.2.1 Distanz

Es sei  $G = (V, E)$  ein Graph.

**Definition.** Für alle  $v, w \in V$  mit  $v \sim w$  definieren wir die *Distanz* zwischen  $v$  und  $w$  als

$$d(v, w) := \min\{l \in \mathbb{N}_0 \mid \text{in } G \text{ existiert ein } v\text{-}w\text{-Pfad der Länge } l\} \in \mathbb{N}_0.$$

Für alle  $v, w \in V$  mit  $v \not\sim w$  wird  $d(v, w) := \infty$  gesetzt.

**Bemerkung.** Für alle  $v, w \in V$  gelten:

$$(i) \quad d(v, w) = 0 \Leftrightarrow v = w,$$

$$(ii) \quad d(v, w) < \infty \Leftrightarrow v \sim w.$$

$G$  ist genau dann zusammenhängend, wenn  $d(v, w) < \infty$  für alle  $v, w \in V$ .

```

BREITENSUCHE( $\Gamma, w$ )
1  initialisiere array  $d[1, \dots, n]$  mit allen Einträgen gleich  $\infty$ 
2  initialisiere array  $p[1, \dots, n]$  mit allen Einträgen gleich NIL
3  initialisiere leere queue  $Q$  (FIFO)
4   $d[w] \leftarrow 0$ 
5  INSERT( $Q, w$ )
6  while  $Q$  ist nicht leer
7  do  $v \leftarrow$  EXTRACT( $Q$ )
8      for  $u \in \Gamma(v)$ 
9      do if  $d[u] = \infty$ 
10         then INSERT( $Q, u$ )
11              $d[u] \leftarrow d[v] + 1$ 
12              $p[u] \leftarrow v$ 
13 return  $d, p$ 

```

Abbildung 5.1: Prozedur Breitensuche

### 5.2.2 Breitensuche

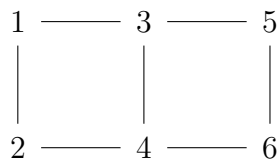
Die *Breitensuche* ist ein Algorithmus, der, beginnend bei einer Wurzel  $w \in V$ , alle Knoten der Zusammenhangskomponente von  $w$  mit aufsteigender Distanz durchläuft. Er eignet sich also zur Berechnung der Zusammenhangskomponenten von  $G$ , insbesondere zur Bestimmung der Brücken und zur Prüfung des Graphen auf Zusammenhang. Außerdem können mit der Breitensuche die Distanzen  $d(v, w)$  sowie kürzeste Pfade von  $v$  nach  $w$  für jeden Knoten  $v$  bestimmt werden. Die kürzesten Pfade von jedem  $v$  zu  $w$  können dadurch angegeben werden, dass man jedem  $v \in V$  einen *Vorgänger* mit kleinerer Distanz zu  $w$  zuordnet. Aus den Vorgängern erhält man dann umgekehrt einen kürzesten Kantenzug von  $v$  nach  $w$ , indem man, ausgehend von  $v$ , sukzessive zum jeweiligen Vorgänger übergeht.

**Algorithmus.** Es sei  $G$  ein Graph mit Knotenmenge  $V = \{1, \dots, n\}$ , gegeben als Adjazenzliste  $\Gamma = (\Gamma(1), \dots, \Gamma(n))$ , und es sei  $w \in V$ . Die in Abbildung 5.2.2 dargestellte Prozedur BREITENSUCHE berechnet zu jedem  $v \in V$  die Distanz  $d(v) := d(v, w)$  sowie einen Vorgänger  $p(v)$  in einem kürzesten  $v$ - $w$ -Pfad.

Die verwendete Datenstruktur *queue* ist eine Warteschlange im „First-in-first-out“-Modus. Der Aufruf INSERT( $Q, x$ ) hängt das Element  $x$  am Ende der Warteschlange ein, der Aufruf EXTRACT( $Q$ ) entnimmt das Element, das am Anfang der Warteschlange steht.

**Bemerkung a.** Da der Verlauf der Breitensuche davon abhängt, in welcher Reihenfolge Knoten in die Warteschlange aufgenommen werden, spielt die Anordnung der Knoten in den Mengen  $\Gamma(v)$ ,  $v \in V$  der Adjazenzliste eine Rolle; diese Anordnung bestimmt, in welcher Reihenfolge die Knoten in der **for**-Schleife bearbeitet werden. An folgendem Beispiel wird deutlich, wie die Anordnung der Adjazenzlisten den Verlauf und das Ergebnis für  $p$ , nicht aber das Ergebnis für  $d$  beeinflusst.

**Beispiel.** Betrachte folgenden Graph mit  $V = \underline{6}$  und Wurzel  $w = 1$ :



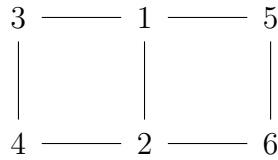
Die erste Tabelle zeigt den Ablauf der Breitensuche, wenn die Mengen  $\Gamma(v)$  mit aufsteigender Nummerierung angeordnet sind. Jede Zeile entspricht dabei einem Durchlauf der **while**-Schleife und gibt folgendes an: die Zustände der Datenstrukturen  $d, p, Q$  zu Beginn der **while**-Schleife, das von **EXTRACT** gelieferte  $v$ , die Liste  $\Gamma(v)$  der Nachbarn von  $v$ , und die Teilliste der  $u \in \Gamma(v)$  mit  $d[u] = \infty$ .

$d$	$p$	$Q$	$v$	$\Gamma(v)$	$d[u] = \infty$
$(0, \infty, \infty, \infty, \infty, \infty)$	$(-, -, -, -, -, -)$	$(1)$	1	$(2, 3)$	$(2, 3)$
$(0, 1, 1, \infty, \infty, \infty)$	$(-, 1, 1, -, -, -)$	$(2, 3)$	2	$(1, 4)$	$(4)$
$(0, 1, 1, 2, \infty, \infty)$	$(-, 1, 1, 2, -, -)$	$(3, 4)$	3	$(1, 4, 5)$	$(5)$
$(0, 1, 1, 2, 2, \infty)$	$(-, 1, 1, 2, 3, -)$	$(4, 5)$	4	$(2, 3, 6)$	$(6)$
$(0, 1, 1, 2, 2, 3)$	$(-, 1, 1, 2, 3, 4)$	$(5, 6)$	5	$(3, 6)$	$()$
$(0, 1, 1, 2, 2, 3)$	$(-, 1, 1, 2, 3, 4)$	$(6)$	6	$(4, 5)$	$()$
$(0, 1, 1, 2, 2, 3)$	$(-, 1, 1, 2, 3, 4)$	$()$			

Die nächste Tabelle zeigt den Ablauf, wenn die Mengen  $\Gamma(v)$  mit absteigender Nummerierung angeordnet sind.

$d$	$p$	$Q$	$v$	$\Gamma(v)$	$d[u] = \infty$
$(0, \infty, \infty, \infty, \infty, \infty)$	$(-, -, -, -, -, -)$	$(1)$	1	$(3, 2)$	$(3, 2)$
$(0, 1, 1, \infty, \infty, \infty)$	$(-, 1, 1, -, -, -)$	$(3, 2)$	3	$(5, 4, 1)$	$(5, 4)$
$(0, 1, 1, 2, 2, \infty)$	$(-, 1, 1, 3, 3, -)$	$(2, 5, 4)$	2	$(4, 1)$	$()$
$(0, 1, 1, 2, 2, \infty)$	$(-, 1, 1, 3, 3, -)$	$(5, 4)$	5	$(6, 3)$	$(6)$
$(0, 1, 1, 2, 2, 3)$	$(-, 1, 1, 3, 3, 5)$	$(4, 6)$	4	$(6, 3, 2)$	$()$
$(0, 1, 1, 2, 2, 3)$	$(-, 1, 1, 3, 3, 5)$	$(6)$	6	$(5, 4)$	$()$
$(0, 1, 1, 2, 2, 3)$	$(-, 1, 1, 3, 3, 5)$	$()$			

Übung a. Wir betrachten den folgenden Graph mit  $V = \underline{6}$  und Wurzel  $w = 1$ :



Man beschreibe den Verlauf der Breitensuche mit einer Tabelle wie im Beispiel, wobei die Adjazenzlisten mit aufsteigender Nummerierung angeordnet sind.

Übung b. Geben Sie eine Schleifeninvariante für die while-Schleife in der Breitensuche an.

**Bemerkung b.** Die *Tiefensuche* wird realisiert, wenn man die queue (FIFO) durch einen stack (LIFO=„Last-in-first-out“) ersetzt. Geht es nur um die Bestimmung der Zusammenhangskomponente von  $w$  bzw. um die Prüfung des gesamten Graphen auf Zusammenhang, dann spielt es keine Rolle, ob Breiten- oder Tiefensuche verwendet wird.

### 5.2.3 Dijkstras Algorithmus

**Definition.** Ein (ungerichteter) *gewichteter Graph* ist ein Tripel  $G = (V, E, f)$ , wobei  $(V, E)$  ein Graph ist und  $w$  eine *Gewichtsfunktion*  $f : E \rightarrow \mathbb{R}_{\geq 0}$ . Für jede Teilmenge  $T \subseteq E$  und jeden Kantenzug  $z = (v_0, \dots, v_l)$  in  $G$  definieren wir deren *Gewichte* als  $f(T) := \sum_{e \in T} f(e)$  bzw.  $f(z) := \sum_{i=1}^l f(v_{i-1}v_i)$ .

Für alle  $v, w \in V$  mit  $v \sim w$  definieren wir die *Distanz* zwischen  $v$  und  $w$  als

$$d(v, w) := \min\{f(z) \mid z \text{ ist } v\text{-}w\text{-Pfad in } G\} \in \mathbb{R}_{\geq 0}.$$

Für alle  $v, w \in V$  mit  $v \not\sim w$  wird  $d(v, w) := \infty$  gesetzt.

Der Algorithmus von *Dijkstra* (1959) ist eine modifizierte Form der Breitensuche, die, beginnend bei einer Wurzel  $w \in V$ , für jeden Knoten der Zusammenhangskomponente von  $w$  die Distanz  $d(v, w)$  sowie einen  $v$ - $w$ -Pfad  $z$  mit minimalem Gewicht, d.h. mit  $f(z) = d(v, w)$ , berechnet.

**Algorithmus.** Es sei  $G = (V, E, f)$  ein gewichteter Graph mit Knotenmenge  $V = \{1, \dots, n\}$ , gegeben als Adjazenzliste  $\Gamma$ , und es sei  $w \in V$ . Die in der Abbildung unten dargestellte Prozedur DIJKSTRA berechnet zu jedem  $v \in V$  die Distanz  $d(v) := d(v, w)$  sowie einen Vorgänger  $p(v)$  in einem  $v$ - $w$ -Pfad von minimalem Gewicht.

Die verwendete Datenstruktur **priority queue** ist eine *Vorrangwarteschlange*, bei der jedem ihrer Element ein *Prioritätswert* zugeordnet ist. Der Aufruf  $\text{INSERT}(Q, x, k)$  fügt das Element  $x$  in die Warteschlange ein und ordnet  $x$  die Priorität  $k \geq 0$  zu. Falls  $x$  bereits in der Warteschlange enthalten ist, wird nur die Priorität neu auf  $k$  gesetzt. Der Aufruf  $\text{EXTRACTMIN}(Q)$  entnimmt das Element mit der niedrigsten Priorität.

*Beweis.* Zum Beweis der Korrektheit führen wir die folgende Terminologie ein, die sich auf den Verlauf des Algorithmus bezieht. Ein Knoten  $v \in V$  heißt *besucht*, falls  $v \notin Q$  ist. Ist  $u \in V$ , dann heißt ein  $w$ - $u$ -Pfad  $z$  in  $G$  *bekannt*, falls alle Vorgänger von  $u$  auf  $z$  besuchte Knoten sind. Nach dem Ende einer while-Schleife existiert ein bekannter  $w$ - $u$ -Pfad genau dann, wenn  $d[u] < \infty$  ist. Die Korrektheit ergibt sich aus den folgenden beiden Aussagen.

(a) Nach dem Ende einer jeden while-Schleife gilt für alle  $u \in V$  mit  $d[u] < \infty$ :  $d[u] = \min\{f(z) \mid z \text{ ist bekannter } w\text{-}u\text{-Pfad in } G\}$ .

(b) Wird  $v$  aus  $Q$  in einer while-Schleife extrahiert, dann ist  $d[v]$  bereits die Distanz zu  $w$ . Anschließend wird  $d[v]$  nicht mehr verändert.

Wir beweisen beide Aussagen durch Induktion über die Anzahl  $n$  der durchlaufenen while-Schleifen. Die Aussagen sind offensichtlich richtig für  $n = 0$  (d.h. vor dem Durchlaufen der ersten while-Schleife). Wir nehmen an, dass die Aussagen richtig sind nach dem Durchlaufen der  $n$ -ten while-Schleife. Es sei  $v$  der in der  $(n+1)$ -ten while-Schleife extrahierte Knoten und  $u \in \Gamma(v)$ . Es seien  $d$  und  $D$  die Werte von  $d[u]$  nach der  $n$ -ten bzw.  $(n+1)$ -ten Schleife. Ist  $d[v] + f(uv) \geq d$ , dann ist  $D = d$  und gibt keinen neuen bekannten  $w$ - $u$ -Pfad. In diesem Fall folgt die Aussage (a) aus der Induktionsvoraussetzung. Ist  $d[v] + f(uv) < d$ , dann ist  $D = d[v] + f(uv)$ , und es gibt einen neuen  $w$ - $u$ -Pfad der Distanz  $D$ , in dem  $v$  der Vorgänger von  $u$  ist. Es sei  $z$  ein bekannter  $w$ - $u$ -Pfad und  $v'$  der Vorgänger von  $u$  auf  $z$ . Mit  $z_1$  bezeichnen wir das  $w$ - $v'$ -Anfangsstück von  $z$ . Ist  $v' = v$ , dann ist  $f(z) = f(z_1) + f(uv) \geq d[v] + f(uv) = D$  nach Induktionsvoraussetzung, da  $z_1$  ein bekannter  $w$ - $v$ -Pfad ist. Es sei nun  $v' \neq v$ . Da  $v'$  ein besuchter Knoten ist, wurde  $v'$  in einer früheren while-Schleife extrahiert. Nach Induktionsvoraussetzung gilt also  $d[v'] = d(w, v')$ . Außerdem ist  $d[v'] + f(uv') \geq d$ , da  $u \in \Gamma(v')$  und  $v'$  ein besuchter Knoten ist. Wir erhalten

$$f(z) = f(z_1) + f(uv') \geq d(w, v') + f(uv') = d[v'] + f(uv') \geq d > D.$$

Also gilt die Behauptung aus (a) auch in diesem Fall.

Wir beweisen nun noch (b). Vor der while-Schleife ist  $v$  ein nicht-besuchter Knoten und es gilt:

$$d[v] = \min\{d[v'] \mid v' \in V \text{ nicht besucht}\}.$$

```

DIJKSTRA( $\Gamma, w, f$ )
1  initialisiere array  $d[1, \dots, n]$  mit allen Einträgen gleich  $\infty$ 
2  initialisiere array  $p[1, \dots, n]$  mit allen Einträgen gleich NIL
3  initialisiere priority queue  $Q$  mit Elementen  $1, \dots, n$  und allen Prioritäten  $= \infty$ 
4   $d[w] \leftarrow 0$ 
5  INSERT( $Q, w, d[w]$ )
6  while  $Q$  nicht leer
7  do  $v \leftarrow \text{EXTRACTMIN}(Q)$ 
8      for  $u \in \Gamma[v]$ 
9      do if  $d[v] + f(uv) < d[u]$ 
10         then  $d[u] \leftarrow d[v] + f(uv)$ 
11              $p[u] \leftarrow v$ 
12             INSERT( $Q, u, d[u]$ )
13 return  $d, p$ 

```

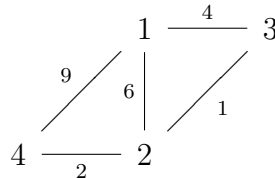
Abbildung 5.2: Prozedur Dijkstra

Angenommen, es existiert ein  $w$ - $v$ -Pfad  $z$  mit  $f(z) < d[v]$ . Aus der Induktionsvoraussetzung für Aussage (a) folgt, dass  $z$  nicht bekannt ist. Sei  $u$  der erste (von  $w$  aus gesehen) nicht besuchte Knoten auf  $z$ . Mit  $z_1$  bezeichnen wir das  $w$ - $u$ -Anfangsstück von  $z$  und mit  $z_2$  das  $u$ - $v$ -Endstück von  $z$ . Aus der Induktionsvoraussetzung für Aussage (a) folgt  $d[u] \leq f(z_1)$ , da  $z_1$  ein bekannter  $w$ - $u$ -Pfad ist. Wir erhalten

$$d[v] > f(z) = f(z_1) + f(z_2) \geq d[u] + f(z_2) \geq d[u],$$

im Widerspruch zur Wahl von  $v$ . Damit ist  $d[v] = d(w, v)$ . Es ist klar, dass dann  $d[v]$  im weiteren Verlauf des Algorithmus nicht mehr geändert wird.  $\square$

**Beispiel.** Betrachte folgenden gewichteten Graph mit  $V = \underline{4}$  und Wurzel  $w = 1$ :

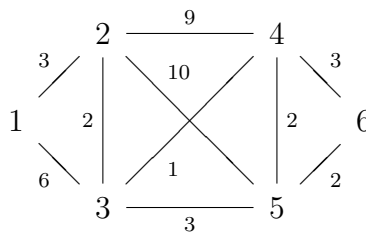


Die erste Tabelle zeigt den Ablauf des Dijkstra-Algorithmus, wenn die Adjazenzlisten mit aufsteigender Nummerierung angeordnet sind. Jede Zeile entspricht dabei einem Durchlauf der **while**-Schleife und gibt folgendes an: die Zustände der Datenstrukturen  $d, p, Q$  zu Beginn der **while**-Schleife, das von EXTRACTMIN gelieferte  $v$ , dessen Adjazenzliste  $\Gamma(v)$ , und die Teilliste

der  $u \in \Gamma(v)$  mit  $d[v] + f(uv) < d[u]$ . Die Vorrangwarteschlange  $Q$  wird jetzt als Menge geschrieben. Die Prioritäten in  $Q$  brauchen nicht extra aufgelistet werden, da sie mit den Werten  $d[v]$  übereinstimmen.

$d$	$p$	$Q$	$v$	$\Gamma(v)$	$d[v] + f(uv) < d[u]$
$(0, \infty, \infty, \infty)$	$(-, -, -, -)$	$\{1, 2, 3, 4\}$	1	$(2, 3, 4)$	$(2, 3, 4)$
$(0, 6, 4, 9)$	$(-, 1, 1, 1)$	$\{2, 3, 4\}$	3	$(1, 2)$	$(2)$
$(0, 5, 4, 9)$	$(-, 3, 1, 1)$	$\{2, 4\}$	2	$(1, 3, 4)$	$(4)$
$(0, 5, 4, 7)$	$(-, 3, 1, 2)$	$\{4\}$	4	$(1, 2)$	$()$
$(0, 5, 4, 7)$	$(-, 3, 1, 2)$	$\{\}$			

*Übung a.* Betrachte folgenden gewichteten Graph mit  $V = \underline{6}$  und Wurzel  $w = 1$ :



Man beschreibe den Verlauf des Dijkstra-Algorithmus mit einer Tabelle wie im Beispiel, wobei die Adjazenzlisten mit aufsteigender Nummerierung angeordnet sind.

*Übung b.* Geben Sie eine Schleifeninvariante für die while-Schleife im Dijkstra-Algorithmus an und versuchen Sie damit, die Korrektheit zu beweisen.

## 5.3 Hamiltonkreise und Eulertouren

Es sei  $G = (V, E)$  ein Graph.

### 5.3.1 Definition und Beispiele

**Definition.**

- (i) Ein Kreis der Länge  $n_G$  in  $G$  heißt *Hamiltonkreis*.
- (ii) Eine Tour der Länge  $m_G$  in  $G$  heißt *Eulertour*.

**Bemerkung.**

- (i) Ein geschlossener Kantenzug  $(v_0, \dots, v_l)$  ist genau dann ein Hamiltonkreis, wenn in der Auflistung  $v_0, \dots, v_{l-1}$  jeder Knoten aus  $V$  genau einmal vorkommt. Existiert ein Hamiltonkreis, so ist  $G$  zusammenhängend und jeder Knoten hat  $\text{Grad} \geq 2$ .

- (ii) Ein geschlossener Kantenzug  $(v_0, \dots, v_l)$  ist genau dann eine Eulertour, wenn in der Auflistung  $v_0v_1, v_1v_2, \dots, v_{l-1}v_l$  jede Kante aus  $E$  genau einmal vorkommt. Existiert eine Eulertour, so hat  $G$  höchstens eine nicht-triviale Komponente.
- (iii) Ein (nicht notwendig geschlossener) Kantenzug  $(v_0, \dots, v_l)$  heißt *Eulerzug*, wenn in der Auflistung  $v_0v_1, v_1v_2, \dots, v_{l-1}v_l$  jede Kante aus  $E$  genau einmal vorkommt.

**Beispiel a.** Jeder Graph ohne Kanten hat eine Eulertour der Länge 0. Der Graph „Haus vom Nikolaus“ besitzt einen Eulerzug, aber keine Eulertour. Weitere Beispiele inkl. Bilder siehe Vorlesung.

**Beispiel b.** Das Straßennetz einer Stadt sei durch einen Graphen modelliert, in dem die Knoten den Kreuzungen entsprechen und die Kanten den Straßenabschnitten. Der Fahrer eines Schneeräumfahrzeuges sucht dann eine Eulertour.

*Übung.* Ist jeder Hamiltonkreis eine Eulertour?

### 5.3.2 Eulertouren

**Bemerkung.** Existiert in  $G$  eine Eulertour, so gelten:

- (i) alle Knoten haben geraden Grad,
- (ii) höchstens eine Komponente ist nicht-trivial.

*Beweis.* Hat man  $v$  über eine Kante erreicht, dann muss man  $v$  über eine andere Kante wieder verlassen. Hat man  $v$  nicht erreicht, so ist  $v$  isoliert, also  $\deg v = 0$  gerade.  $\square$

Unser Ziel ist es nun, die Umkehrung dieser Bemerkung zu zeigen unter der notwendigen Voraussetzung, dass  $G$  höchstens eine nicht-triviale Komponente hat.

**Satz.** Jeder Graph, der höchstens eine nicht-triviale Komponente besitzt (z.B. ein zusammenhängender Graph) und genau zwei Knoten  $u, v$  mit ungeradem Grad, besitzt auch einen  $u$ - $v$ -Eulerzug.

*Beweis.* Es sei  $G$  ein solcher Graph. Da  $\deg u$  und  $\deg v$  ungerade (also  $> 0$ ) sind, liegen  $u$  und  $v$  in der einzigen nicht-trivialen Komponente von  $G$ . Wir führen Induktion nach  $m_G$ . Für  $m_G = 1$  ist  $E = \{uv\}$  und die Aussage trivial (Induktionsanfang). Sei nun  $m_G > 1$  und die Behauptung für kleineres  $m_G$  bereits bewiesen. O.B.d.A. nehmen wir  $\deg(u) \geq \deg(v) > 1$ ,



dann ist sogar  $\deg(u) \geq 3$ . Nach Satz 5.1.8 ist  $u$  zu höchstens einer Brücke inzident. Es existiert also eine Kante  $e = uw$  mit  $w \neq v$ , so dass  $e$  keine Brücke ist. Ist  $\deg(u) = 1$ , dann auch  $\deg(v) = 1$ . In diesem Fall ist  $u$  zu einer einzigen Kante  $e = uw$  inzident, und  $w \neq v$  wegen  $m_G > 1$ . Betrachte  $G' := (V, E \setminus \{e\})$ . Auch  $G'$  hat höchstens eine nicht-triviale Komponente, da  $e$  entweder keine Brücke oder  $u$  isoliert ist. In  $G'$  sind außerdem  $w$  und  $v$  die einzigen Knoten mit ungeradem Grad. Nach Induktionsvoraussetzung besitzt  $G'$  also einen Eulerzug von  $w$  nach  $v$ . Begonnen mit  $e = uw$  liefert dies einen Eulerzug von  $u$  nach  $v$  in  $G$ .  $\square$

**Folgerung.** *Ein Graph besitzt genau dann eine Eulertour, wenn er höchstens eine nicht-triviale Komponente besitzt (z.B. wenn er zusammenhängend ist) und alle Knoten geraden Grad haben.*

**Algorithmus** (Fleury, „Schneeräumen, 1883“). *Es sei  $G = (V, E)$  ein zusammenhängender Graph, dessen Knoten geraden Grad haben. Die in der Abbildung unten dargestellte Prozedur FLEURY berechnet eine Eulertour.*

Der Aufruf  $\text{Append}(T, x)$  hängt das Element  $x$  am Ende der Liste  $T$  an. Die Korrektheit des Algorithmus ergibt sich aus einer Modifikation des Beweises von Satz 5.3.2.

FLEURY( $V, E$ )

```

1  initialisiere leere Liste  $T$ 
2   $v \leftarrow$  beliebiger Knoten aus  $V$ 
3  APPEND( $T, v$ )
4  while  $E$  ist nicht leer
5  do if  $\deg v = 1$ 
6      then  $w \leftarrow$  einziger Nachbar von  $v$ 
7      else  $w \leftarrow$  ein Nachbar von  $v$  mit  $vw$  keine Brücke
8      APPEND( $T, w$ )
9       $E \leftarrow E \setminus \{vw\}$ 
10      $v \leftarrow w$ 
11 return  $T$ 
```

Abbildung 5.3: Prozedur Fleury

### 5.3.3 Hamiltonkreise

**Bemerkung.** Existiert in  $G$  ein Hamiltonkreis, so ist  $G$  zusammenhängend und  $n_G \geq 3$ .

**Satz.** Es sei  $n_G \geq 3$  und  $G$  zusammenhängend. Falls für alle  $u, v \in V$  mit  $u \neq v$  und  $uv \notin E$  gilt

$$\deg u + \deg v \geq n_G,$$

so besitzt  $G$  einen Hamiltonkreis.

*Beweis.* Setze  $n := n_G$ . Es sei  $(v_1, \dots, v_n)$  eine beliebige Permutation der Knotenmenge  $V$ . Wegen  $n \geq 3$  können wir  $(v_1, \dots, v_n, v_1)$  als Kreis der Länge  $n$  in dem vollständigen Graphen mit Knotenmenge  $V$  auffassen. Von den  $n$  Kanten dieses Kreises seien  $r$  in  $E$ . Ist  $r = n$ , dann ist dieser Kreis ein Hamiltonkreis in  $G$ . Sei also  $r < n$ , o.B.d.A. etwa  $v_1v_2 \notin E$ .

Behauptung: Es existiert ein  $i \in \{3, \dots, n\}$  so, dass  $v_1v_{i-1}, v_2v_i \in E$ .

Dann ist  $(v_1, v_{i-1}, v_{i-2}, \dots, v_2, v_i, v_{i+1}, \dots, v_n, v_1)$  ein Kreis im vollständigen Graphen, von dessen Kanten mindestens  $r+1$  in  $E$  liegen. In der Tat wurden  $v_1v_2, v_{i-1}v_i$  ersetzt durch  $v_1v_{i-1}, v_2v_i$ . Nach endlich vielen Schritten kommen wir zu einem Kreis der Länge  $n$  im vollständigen Graphen, dessen sämtliche Kanten in  $E$  liegen, d.h. zu einem Hamiltonkreis in  $G$ .

Wir zeigen nun die Behauptung. Die Bedingung an  $i \in \{3, \dots, n\}$  lautet  $v_i \in \Gamma(v_2)$  und  $v_{i-1} \in \Gamma(v_1)$ . Setze  $S := \{1 \leq j \leq n \mid v_j \in \Gamma(v_2)\}$  und  $T := \{2 \leq j \leq n+1 \mid v_{j-1} \in \Gamma(v_1)\}$ . Es gilt  $|S| = \deg v_2$  und  $|T| = \deg v_1$ . Wegen  $v_1v_2 \notin E$  gilt nach Voraussetzung  $|S| + |T| \geq n$ . Wegen  $1, 2 \notin S \cup T$  ist  $|S \cup T| < n$ . Aus  $n > |S \cup T| = |S| + |T| - |S \cap T|$  (Inklusions-Exklusions-Prinzip) folgt demnach  $|S \cap T| \geq 1$  und daraus die Behauptung.  $\square$

**Bemerkung.** Erfüllt ein Graph die Voraussetzungen des Satzes und ist  $n$  gerade, so gilt  $m_G \geq \frac{n^2}{4}$ . Wegen  $\frac{n^2}{4} > \frac{n(n-1)}{4} = \frac{1}{2} \binom{n}{2}$  bedeutet das, dass der Graph mindestens halb so viele Kanten enthält, wie der vollständige Graph mit gleicher Knotenzahl.

*Beweis.* Es bezeichne  $S$  die Gradsumme des Graphen  $G$ , also  $S = 2m_G$ . Die Behauptung ist dann  $S \geq \frac{n^2}{2}$ . Es sei  $k$  der minimale Grad der unter allen Knoten auftritt. Ist  $k \geq \frac{n}{2}$ , dann folgt  $S \geq n \cdot \frac{n}{2} = \frac{n^2}{2}$  wie gefordert. Sei also  $k = \frac{n}{2} - l$  mit  $l \in \mathbb{N}$ . Wähle einen Knoten  $v$  mit Grad  $k$ . Partitioniere die Knotenmenge in  $V = V_0 \cup V_1$ , wobei  $V_0 := \Gamma(v) \cup \{v\}$  und  $V_1 := V \setminus V_0$ . Für alle  $w \in V_1$  gilt nach Voraussetzung  $n \leq \deg v + \deg w = k + \deg w$ , also  $\deg w \geq n - k = \frac{n}{2} + l$ . Nach Wahl von  $k$  gilt  $\deg w \geq k = \frac{n}{2} - l$  für alle  $w \in V_0$ . Ausserdem ist  $|V_0| = k + 1 = \frac{n}{2} - l + 1$  und  $|V_1| = n - (k + 1) = \frac{n}{2} + l - 1$ .

Es folgt

$$\begin{aligned}
 S &\geq |V_0|(\frac{n}{2} - l) + |V_1|(\frac{n}{2} + l) = \\
 &= ((\frac{n}{2} - l) + 1)(\frac{n}{2} - l) + ((\frac{n}{2} + l) - 1)(\frac{n}{2} + l) \\
 &= (\frac{n}{2} - l)^2 + (\frac{n}{2} + l)^2 + (\frac{n}{2} - l) - (\frac{n}{2} + l) \\
 &= 2\frac{n^2}{4} + 2l^2 - 2l = \frac{n^2}{2} + 2(l^2 - l) \geq \frac{n^2}{2}.
 \end{aligned}$$

□

*Übung.* Man zeige: Erfüllt ein Graph die Voraussetzungen des Satzes, so gilt  $d_w(v) \leq 2$  für alle  $v, w \in V$ .

*Übung.* Man zeige: Erfüllt ein Graph die Voraussetzungen des Satzes und ist  $n$  ungerade, so gilt  $m_G \geq \frac{n^2-1}{4}$ . *Tip: Modifiziere den Beweis der Bemerkung.*

*Übung.* Geben Sie einen Graphen mit 6 Knoten und 9 Kanten an, der die Voraussetzungen des Satzes erfüllt.

*Übung.* Versuchen Sie, für kleines  $n$ , einen Graphen mit ungerader Knotenzahl  $n$  und Kantenzahl  $\frac{n^2-1}{4}$  anzugeben, der die Voraussetzungen des Satzes erfüllt.

## 5.4 Bäume

Es sei  $G = (V, E)$  ein Graph mit  $n_G > 0$ .

### 5.4.1 Definition und Beispiele

**Definition.**  $G$  heißt *kreisfrei* bzw. *Wald*, falls  $G$  keine Kreise enthält. Ein zusammenhängender Wald heißt *Baum*. Die Knoten eines Waldes mit Grad  $\leq 1$  heißen *Blätter*.

**Beispiel.** Siehe Vorlesung.

**Bemerkung.**

- (i) Ein Graph ist genau dann kreisfrei, wenn jede Kante eine Brücke ist.
- (ii) Jeder Baum mit mehr als einem Knoten hat mindestens zwei Blätter.
- (iii) Jeder Baum mit mehr als zwei Knoten hat höchstens  $n_G - 1$  Blätter.

*Beweis.*

- (i) Bemerkung 5.1.8b.
- (ii) Es seien  $G$  ein Baum und  $(v_0, v_1, \dots, v_l)$  ein beliebiger maximaler Pfad in  $G$ . (Ein maximaler Pfad ist einer, der sich nicht „verlängern“ lässt.) Wenn  $\deg v_0 > 1$ , dann hat  $v_0$  einen Nachbarn  $w \neq v_1$ . Wäre  $w = v_i$  für ein  $2 \leq i \leq l$ , so gäbe es einen Kreis in  $G$ , im Widerspruch dazu, dass  $G$  ein Baum ist. Somit ist auch  $(w, v_0, \dots, v_l)$  ein Pfad, im Widerspruch zur Maximalität von  $(v_0, \dots, v_l)$ . Folglich ist  $\deg v_0 \leq 1$ , d.h.  $v_0$  ist ein Blatt, und dasselbe gilt aus Symmetriegründen für  $v_l$ .
- (iii) Es sei  $G$  ein Baum mit  $n$  Blättern, d.h. jeder Knoten ist Blatt. Dann ist  $n \geq \sum_{v \in V} \deg v = 2m_G \geq 2(n-1) = 2n-2$ , wobei die zweite Ungleichung nach Folgerung 5.1.7a gilt. Daraus folgt  $n \leq 2$ .

□

## 5.4.2 Kantenzahl

Ist  $r$  die Komponentenzahl von  $G$ , so gilt nach Satz 5.1.7a  $r \geq n_G - m_G$ . Als Zusatz erhalten wir:

**Satz.** *Es gilt  $r = n_G - m_G$  genau dann, wenn  $G$  kreisfrei ist.*

**Lemma.** *Es sei  $e \in E$ . Für Komponentenzahl  $l$  von  $(V, E \setminus \{e\})$  gilt dann  $l \leq r + 1$ . Weiter ist  $l = r + 1$  genau dann, wenn  $e$  eine Brücke ist.*

*Beweis.* Nach Lemma (5.4.2), angewendet auf  $(V, E \setminus \{e\})$ , ist  $r \geq l - 1$ , d.h.  $l \leq r + 1$ . Per Definition ist  $e$  genau dann eine Brücke, wenn  $l > r$ , also genau dann, wenn  $l = r + 1$ . □

*Beweis des Satzes.* Zunächst sei  $G$  kreisfrei. Wir zeigen  $r = n_G - m_G$  per Induktion nach  $m_G$  (genauso wie im Beweis von Satz 5.1.7a). Ist  $m_G = 0$ , so besteht jede Komponente aus einem einzelnen Knoten, also gilt  $r = n_G$ . Sei nun  $m_G > 0$  und die Behauptung für kleineres  $m_G$  bereits bewiesen. Wähle ein  $e \in E$  und setze  $G' := (V, E \setminus \{e\})$ . Sei  $l$  die Komponentenzahl von  $G'$ . Da  $G$  kreisfrei ist, ist  $e$  eine Brücke und nach dem Lemma gilt  $l = r + 1$ . Da  $G$  kreisfrei ist, ist auch  $G'$  kreisfrei und nach Induktionsvoraussetzung (angewendet auf  $G'$ ) gilt  $l = n_G - (m_G - 1) = n_G - m_G + 1$ . Zusammen also  $r = n_G - m_G$ .

Nun sei  $G$  nicht kreisfrei. Dann existiert eine nicht-Brücke  $e$ . Der Graph  $G' := (V, E \setminus \{e\})$  hat dann dieselbe Komponentenzahl wie  $G$ . Nach Satz 5.1.7a (angewendet auf  $G'$ ) gilt also  $r \geq n_G - (m_G - 1) > n_G - m_G$ . □

**Folgerung.** Ein Graph ist genau dann ein Baum, wenn mindestens zwei der folgenden Bedingungen erfüllt sind.

- (i)  $G$  ist kreisfrei,
- (ii)  $G$  ist zusammenhängend,
- (iii)  $m_G = n_G - 1$ .

*Beweis.* Zu zeigen ist, dass aus je zwei der Bedingungen die dritte folgt. Das wird offensichtlich, wenn man Bedingung (i) dem Satz gemäß durch die Bedingung  $r = n_G - m_G$  ersetzt, sowie (ii) durch die Bedingung  $r = 1$ .  $\square$

**Bemerkung.** Jeder zusammenhängende Graph erfüllt nach Folgerung 5.1.7a  $m_G \geq n_G - 1$ . Jeder kreisfreie Graph erfüllt nach Satz 5.4.2  $m_G = n_G - r \leq n_G - 1$ . Ein Baum ist also ein zusammenhängender Graph mit minimal möglicher Kantenzahl und ein kreisfreier Graph mit maximal möglicher Kantenzahl.

### 5.4.3 Spannbäume

**Definition.** Ein Teilgraph  $G' = (V', E')$  von  $G$  heißt *Spannbaum* von  $G$  (engl. *spanning tree*), wenn  $G'$  ein Baum ist und  $V' = V$ .

**Beispiel.** Siehe Vorlesung.

**Satz.** Jeder zusammenhängende Graph hat einen Spannbaum.

*Beweis.* Die Breitensuche mit beliebiger Wurzel  $w$  liefert für jedes  $v \in V \setminus \{w\}$  einen „Vorgänger“  $p(v)$ , der kleinere Distanz zu  $w$  hat. Die Kantenmenge  $E' := \{vp(v) \mid v \in V, v \neq w\}$  liefert dann einen Spannbaum  $(V, E')$  von  $G$ : Einerseits ist  $(V, E')$  zusammenhängend weil jeder Knoten über seine Vorgänger mit  $w$  verbunden ist. Andererseits sind die Kanten  $vp(v)$  mit  $v \in V \setminus \{w\}$  paarweise verschieden, also  $|E'| = n_G - 1$ . Nach Folgerung (5.4.2) ist  $(V, E')$  ein Baum.  $\square$

**Bemerkung.** Die Blätterzahl der Spannbäume eines Graphen ist durch den Graphen nicht eindeutig festgelegt.

Zwei weitere Algorithmen zur Generierung eines Spannbaumes bieten sich an.

**Algorithmus a** (Sukzessives Entfernen von Kanten). *Es sei  $(V, E)$  ein zusammenhängender Graph. Beginnend mit der Kantenmenge  $B := E$  werden sukzessive solche Kanten aus  $B$  entfernt, die keine Brücken in  $(V, B)$  sind. Wenn das nicht mehr möglich ist, dann ist  $(V, B)$  ein Spannbaum.*

*Beweis.* Zu Beginn des Algorithmus ist  $(V, B) = (V, E)$ , also  $(V, B)$  zusammenhängend. Für jedes  $e \in B$  sind dann äquivalent:

- (i)  $e$  ist keine Brücke,
- (ii)  $(V, B \setminus \{e\})$  ist zusammenhängend.

(Das ist die Definition von Brücke.) Entfernt man aus  $B$  stets Kanten  $e$  dieser Art, so bleibt  $(V, B)$  während des gesamten Verlaufs des Algorithmus zusammenhängend. Wenn das Abbruchkriterium erfüllt ist, d.h. wenn es keine Kanten dieser Art mehr gibt, dann ist  $(V, B)$  nach Bemerkung (5.4.1) kreisfrei. Somit ist  $(V, B)$  dann ein Baum mit Knotenmenge  $V$ , also Spannbaum von  $(V, E)$ .  $\square$

**Algorithmus b** (Sukzessives Hinzufügen von Kanten). *Beginnend mit der leeren Kantenmenge  $B := \emptyset$  werden sukzessive solche Kanten aus  $E$  zu  $B$  hinzugefügt, deren Endknoten in verschiedenen Komponenten von  $(V, B)$  liegen. Wenn das nicht mehr möglich ist, dann ist  $(V, B)$  ein Spannbaum.*

*Beweis.* Zu Beginn des Algorithmus ist  $(V, B)$  kreisfrei ( $B = \emptyset$ ). Für jedes  $e \in E$  sind dann äquivalent:

- (i)  $e$  hat Endknoten in verschiedenen Komponenten von  $(V, B)$ ,
- (ii)  $(V, B \cup \{e\})$  ist kreisfrei.

(Dies ist Bemerkung 5.1.8b.) Fügt man zu  $B$  stets Kanten  $e$  dieser Art hinzu, so bleibt  $(V, B)$  während des gesamten Verlaufs des Algorithmus kreisfrei. Wenn das Abbruchkriterium erfüllt ist, d.h. wenn es keine Kanten dieser Art mehr gibt, so ist  $(V, B)$  zusammenhängend. Somit ist  $(V, B)$  ein Baum mit Knotenmenge  $V$ , also Spannbaum von  $(V, E)$ .  $\square$

**Beispiel.** Siehe Vorlesung.

#### 5.4.4 Minimale Spannbäume

Es sei nun  $G = (V, E)$  ein zusammenhängender Graph mit einer *Gewichtsfunktion*

$$f : E \rightarrow \mathbb{R}_{\geq 0}.$$

Für jede Teilmenge  $T \subseteq E$  definieren wir

$$f(T) := \sum_{e \in T} f(e).$$

**Definition.** Ein *minimaler Spannbaum* von  $G$  ist ein Spannbaum  $(V, B)$  von  $G$  mit minimalem Gewicht  $w(B)$  unter allen Spannbäumen von  $G$ .

**Beispiel.** Siehe Vorlesung.

**Satz** (Kruskal). Die „Greedy-Version“ von Algorithmus 5.4.3b, die in jedem einzelnen Schritt unter allen möglichen hinzufügbaren Kanten eine mit geringstem Gewicht hinzufügt, liefert einen minimalen Spannbaum von  $G$ .

Die „Greedy-Version“ von Algorithmus 5.4.3b wird auch *Algorithmus von Kruskal* genannt.

**Lemma** (Austauschlemma). Es seien  $(V, A)$  und  $(V, B)$  zwei Bäume mit derselben Knotenmenge  $V$ . Für jedes  $a \in A \setminus B$  gibt es ein  $b \in B \setminus A$  so, dass  $(V, B \cup \{a\} \setminus \{b\})$  auch ein Baum ist.

*Beweis.* (Es reicht sogar, dass  $(V, A)$  kreisfrei ist.) Nach Folgerung (5.4.2) ist  $|B| = n_G - 1$ . Sei  $a \in A \setminus B$ . Dann ist  $(V, B \cup \{a\})$  zusammenhängend, aber wegen  $|B \cup \{a\}| = |B| + 1 > n_G - 1$  kein Baum, enthält also einen Kreis. Wähle einen Kreis in  $(V, B \cup \{a\})$  und darin eine Kante  $b$ , die nicht in  $A$  liegt. (Da  $(V, A)$  kreisfrei ist, können nicht alle Kanten des Kreises in  $A$  liegen.) Da  $b$  Teil eines Kreises in  $(V, B \cup \{a\})$  ist, ist auch  $(V, B \cup \{a\} \setminus \{b\})$  zusammenhängend (vgl. Bemerkung 5.1.8b). Wegen  $|B \cup \{a\} \setminus \{b\}| = |B| = n_G - 1$  ist  $(V, B \cup \{a\} \setminus \{b\})$  nach Folgerung (5.4.2) ein Baum.  $\square$

*Beweis des Satzes.* Es sei  $(V, A)$  der vom Greedy-Algorithmus produzierte Spannbaum, und es sei  $(V, B)$  ein minimaler Spannbaum von  $G$ . Weiter sei  $a_1, \dots, a_{n-1}$  die Aufzählung der Kanten aus  $A$  in der Reihenfolge, wie sie vom Greedy-Algorithmus ausgewählt wurden. Falls  $A \neq B$ , so existiert  $1 \leq i \leq n - 1$  mit  $a_1, \dots, a_{i-1} \in B$  und  $a_i \notin B$ . Wir nehmen weiter an, dass  $(V, B)$  unter allen minimalen Spannbäumen ein solcher ist, für den  $i$  maximal ist. Nach dem Austauschlemma gibt es ein  $b \in B \setminus A$  so, dass  $(V, B \cup \{a_i\} \setminus \{b\})$  auch ein Baum ist. Aus der Maximalität von  $i$  folgt, dass  $(V, B \cup \{a_i\} \setminus \{b\})$  kein minimaler Spannbaum ist. Also ist  $f(a_i) > f(b)$ . Da  $(V, \{a_1, \dots, a_{i-1}, b\})$  als Teilgraph von  $(V, B)$  kreisfrei ist, hätte der Greedy-Algorithmus also im  $i$ -ten Schritt  $b$  anstatt  $a_i$  anwählen muss. Da dies ein Widerspruch ist, muss  $A = B$  sein.  $\square$

**Beispiel.** Siehe Vorlesung.





# Lineare Algebra



# Einleitung

In der **Algebra** geht es um Gleichungen mit „Unbekannten“ und darum, wie man sie umformt bzw. sogar löst. Mit „Lösen“ meinen wir hier immer exakte Lösungen, im Gegensatz zu angenäherten Lösungen, wie man sie etwa in der *Numerik* betrachtet. In der Regel werden Gleichungen immer über einem bestimmten Zahlbereich betrachtet, etwa über  $\mathbb{Z}, \mathbb{Q}, \mathbb{R}, \mathbb{C}$  oder einem endlichen Zahlbereich (z.B.  $\mathbb{Z}_p$ ). Typisch für die Algebra ist das *Abstrahieren* von konkreten Zahlbereichen, das es ermöglicht, mit den gleichen Methoden unabhängig vom Zahlbereich arbeiten zu können. Soweit möglich werden wir auch algorithmische Aspekte des Lösens von Gleichungen hervorheben.

Man nennt einen „Ausdruck“ oder eine „Gleichung“ mit Unbekannten **linear**, wenn die Unbekannten (in ihrer Gesamtheit) linear (insbesondere mit dem Exponenten 1) auftreten, sonst *nicht-linear*.

**Beispiel.** Es seien  $x, y, z$  Unbekannte und  $a$  eine Konstante. Die folgenden Gleichungen bezeichnet man als *linear*:

$$3x = 2y, \quad a^2x = 2y, \quad ax - y = 7z.$$

Dagegen sind *nicht-linear*:

$$3x^2 = 2y^2, \quad 3^x = 2y, \quad \sin x = 2y.$$

Die *Lineare Algebra* hat ihren Ursprung in der Beschäftigung mit linearen Gleichungen, in einem gewissen Sinne also mit der „einfachsten“ Art von Gleichungen. Im Vergleich zu anderen mathematischen Disziplinen bietet sie dafür auch die mit Abstand erfolgreichsten Lösungsansätze. Probleme aus der Praxis, die nicht-linear sind, werden in der Regel zuerst „linearisiert“. Mit Linearen Gleichungssystemen, deren Lösungsmengen und den zugehörigen Lösungsverfahren haben wir uns bereits ausführlich in Kapitel 3 dieser Vorlesung beschäftigt. In dieser Vorlesung sollen die strukturellen Aspekte der Linearen Algebra vorgestellt und untersucht werden. Was damit gemeint ist, soll in den folgenden Beispielen angedeutet werden.

**Beispiel.**

- (i) Gegeben sei die Gleichung  $3x = 2y$  mit  $x, y \in \mathbb{R}$ . Es gilt  $3x = 2y$  genau dann, wenn  $y = 3x/2$  ist. Die Lösungsmenge

$$\mathbb{L} := \left\{ \begin{pmatrix} x \\ y \end{pmatrix} \mid 3x = 2y, x, y \in \mathbb{R} \right\}$$

ergibt sich also zu

$$\mathbb{L} = \left\{ \begin{pmatrix} x \\ \frac{3}{2}x \end{pmatrix} \mid x \in \mathbb{R} \right\}.$$

Diese Menge kann als eine Gerade im 2-dimensionalen Raum aufgefasst werden.

- (ii) Gegeben sei die Gleichung  $x - 4 = 0$  mit  $x \in \mathbb{R}$ . Die Lösungsmenge  $\mathbb{L} := \{x \mid x - 4 = 0, x \in \mathbb{R}\}$  ergibt sich zu  $\mathbb{L} = \{4\}$ . Diese Menge kann als ein Punkt im 1-dimensionalen Raum aufgefasst werden.

Ein Zusammenhang zwischen Algebra und Geometrie besteht also darin, dass Lösungsmengen von Gleichungen mit Unbekannten als geometrische Objekte im „Raum“ aufgefasst werden können. Dabei entspricht die Anzahl der Unbekannten gerade der **Dimension** des Raumes. Dieser Zusammenhang ist für die Einführung des zentralen Begriffes der Linearen Algebra verantwortlich: dem **Vektorraum**.

**Beispiel.** Es sei  $f : N \rightarrow M$  eine Abbildung. Weiter sei  $x \in N$  eine Unbekannte und  $c \in M$  eine Konstante. Für die (nicht notwendigerweise lineare) Gleichung  $f(x) = c$  und ihre Lösungsmenge  $\mathbb{L} := \{x \in N \mid f(x) = c\}$  gilt:

- (i)  $f(x) = c$  ist genau dann lösbar, wenn  $c$  im Bild von  $f$  liegt.
- (ii)  $\mathbb{L} = f^{-1}(\{c\}) = \{\text{Urbilder von } c \text{ unter } f\}$ , d.h.  $\mathbb{L}$  ist die Faser von  $f$  zu  $c$ .
- (iii)  $f$  ist genau dann injektiv, wenn  $f(x) = c$  für jedes  $c \in M$  höchstens eine Lösung hat.
- (iv)  $f$  ist genau dann surjektiv, wenn  $f(x) = c$  für jedes  $c \in M$  mindestens eine Lösung hat.
- (v)  $f$  ist genau dann bijektiv, wenn  $f(x) = c$  für jedes  $c \in M$  genau eine Lösung hat.

Offensichtlich kann man Gleichungen also auch mit Abbildungen in Verbindung bringen. Im Fall von linearen Gleichungen sind das die **linearen Abbildungen**.

# Kapitel 6

## Vektorräume und lineare Abbildungen

### 6.1 Vektorräume

#### 6.1.1 Definition und Beispiele

Es sei  $K$  ein Körper.

**Definition.** Es sei  $(V, +)$  eine abelsche Gruppe. Dann heißt  $V$  ein  $K$ -Vektorraum oder Vektorraum über  $K$ , wenn eine skalare Multiplikation

$$\cdot : K \times V \rightarrow V, (a, v) \mapsto a \cdot v = av$$

definiert ist, sodaß für alle  $a, b \in K$  und  $v, w \in V$  gelten:

$$(V1) \quad (a + b)v = av + bv;$$

$$(V2) \quad a(v + w) = av + aw;$$

$$(V3) \quad a(bv) = (ab)v;$$

$$(V4) \quad 1v = v.$$

Die Elemente von  $V$  heißen *Vektoren*, die Elemente von  $K$  heißen *Skalare*.

Der Deutlichkeit halber unterscheiden wir in der untenstehenden Folgerung in der Notation zwischen dem Nullelement von  $V$ , dem *Nullvektor*, und dem Nullelement von  $K$ . Ersteres bezeichnen wir mit  $\mathbf{o}$ , letzteres wie üblich mit  $0$ .

**Folgerung.** Es sei  $V$  ein  $K$ -Vektorraum. Für alle  $a \in K, v \in V$  gelten:

$$(W1) \quad 0 \cdot v = \mathbf{o};$$

$$(W2) \quad a \cdot \mathbf{o} = \mathbf{o};$$

$$(W3) \quad -v = (-1)v;$$

$$(W4) \quad (-a)v = -(av);$$

$$(W5) \quad av = \mathbf{o} \Leftrightarrow a = 0 \text{ oder } v = \mathbf{o}.$$

$$\text{Beweis. (W1)} \quad 0 \cdot v = (0 + 0)v \stackrel{(V1)}{=} 0 \cdot v + 0 \cdot v \stackrel{\text{Satz 2.1.4}}{=} 0 \cdot v = \mathbf{o}$$

$$(W2) \quad a \cdot \mathbf{o} \stackrel{(W1)}{=} a \cdot (0 \cdot \mathbf{o}) \stackrel{(V3)}{=} (a \cdot 0) \cdot \mathbf{o} = 0 \cdot \mathbf{o} \stackrel{(W1)}{=} \mathbf{o}$$

$$(W3) \quad v + (-1)v \stackrel{(V4)}{=} 1v + (-1)v \stackrel{(V1)}{=} (1 + (-1))v = 0 \cdot v \stackrel{(W1)}{=} \mathbf{o}, \text{ also } -v = (-1)v \\ \text{nach Satz 2.1.4.}$$

$$(W4) \quad av + (-a)v \stackrel{(V1)}{=} (a + (-a))v = 0 \cdot v \stackrel{(W1)}{=} \mathbf{o}$$

$$(W5) \quad \text{„}\Leftarrow\text{“: (W1) und (W2).}$$

$$\text{„}\Rightarrow\text{“: Sei } av = \mathbf{o} \text{ und } a \neq 0. \text{ Zu zeigen: } v = \mathbf{o}.$$

$$v \stackrel{(V4)}{=} 1v \stackrel{a \neq 0}{=} (a^{-1}a)v \stackrel{(V3)}{=} a^{-1}(av) \stackrel{V.or.}{=} a^{-1}\mathbf{o} \stackrel{(W2)}{=} \mathbf{o}.$$

□

*Übung.* In welcher der Folgerungen wird benutzt, dass  $K$  ein Körper ist statt nur ein Ring?

Jetzt heben wir die Unterscheidung in der Notation auf, und bezeichnen die Nullelemente von  $V$  und von  $K$  mit dem gleichen Symbol  $0$ .

### Beispiel.

- (i)  $V = \{0\}$  ist ein  $K$ -Vektorraum (für jeden Körper  $K$ ) mit der skalaren Multiplikation  $a \cdot 0 = 0$  für alle  $a \in K$ . Er wird der *triviale*  $K$ -Vektorraum genannt.
- (ii) Sind  $K \subseteq L$  zwei beliebige Körper (insbesondere auch für  $K = L$ ), dann ist  $L$  ein  $K$ -Vektorraum mit

$$\cdot : K \times L \rightarrow L, (a, b) \mapsto ab$$

(Die skalare Multiplikation ist hier gerade die Multiplikation in  $L$ .)  
z.B.:  $K$  ist  $K$ -Vektorraum,  $\mathbb{C}$  ist sowohl  $\mathbb{R}$ -Vektorraum als auch  $\mathbb{Q}$ -Vektorraum,  $\mathbb{R}$  ist  $\mathbb{Q}$ -Vektorraum, ...

(iii)  $(K^{m \times n}, +)$  ist  $K$ -Vektorraum mit

$$\cdot : K \times K^{m \times n} \rightarrow K^{m \times n}, (a, A) \mapsto aA$$

(Die skalare Multiplikation ist das skalare Vielfache von Matrizen aus Definition 3.2.1.)

Speziell: die Elemente von  $K^n = K^{n \times 1}$  und  $K^{1 \times n}$  heißen *Spaltenvektoren* bzw. *Zeilenvektoren*.

- (iv)  $\mathbb{R}^3$  der 3-dimensionale „euklidische Raum“, in dem man sich Vektoren als Pfeile ausgehend von einem „Ursprung“ vorstellt. Die Addition von Vektoren entspricht dem „Hintereinanderhängen“ von Pfeilen, die skalare Multiplikation dem „Verlängern“. Nicht jeder Vektorraum erlaubt jedoch eine geometrische Vorstellung.
- (v) Sei  $M$  beliebige Menge. Nach Satz 2.1.6 ist  $(\text{Abb}(M, K), +)$  eine abelsche Gruppe.  $\text{Abb}(M, K)$  wird zu einem  $K$ -Vektorraum mit der skalaren Multiplikation:

$$\cdot : K \times \text{Abb}(M, K) \rightarrow \text{Abb}(M, K), (a, f) \mapsto af, (af)(x) := af(x)$$

Wichtige Beispiele hiervon sind die  $\mathbb{R}$ -Vektorräume:

$$\text{Abb}(\mathbb{R}, \mathbb{R}) := \{f : \mathbb{R} \rightarrow \mathbb{R}\} = \text{reelle Funktionen}$$

$$\text{Abb}(\mathbb{N}, \mathbb{R}) := \{f : \mathbb{N} \rightarrow \mathbb{R}\} = \{(a_1, a_2, a_3, \dots) \mid a_i \in \mathbb{R}, i \in \mathbb{N}\} = \text{reelle Folgen.}$$

*Übung.* Wie ist die skalare Multiplikation in Beispiel (v) zu definieren? Man vergleiche Beispiel (iii) mit dem Spezialfall  $M = \underline{m} \times \underline{n}$  von Beispiel (v).

### 6.1.2 Untervektorräume

**Definition.** Es sei  $V$  ein  $K$ -Vektorraum,  $U \subseteq V$ . Dann heißt  $U$  *Untervektorraum* bzw. *Unterraum* von  $V$ , geschrieben  $U \leq V$ , wenn gelten:

(UV1)  $U \neq \emptyset$ ;

(UV2)  $u + u' \in U$  für alle  $u, u' \in U$  ;

(UV3)  $au \in U$  für alle  $a \in K, u \in U$ .

**Bemerkung.** Ein Unterraum ist also abgeschlossen unter Addition und unter skalarer Multiplikation. Jeder Unterraum von  $V$  enthält 0 und ist selbst ein  $K$ -Vektorraum bzgl. der Addition und der skalaren Multiplikation von  $V$ . (Man zeige dies zur Übung!)

**Beispiel.** Es sei  $V$  ein  $K$ -Vektorraum,

- (i)  $\{0\} \leq V$  und  $V \leq V$ .
- (ii) Für jedes  $v \in V$  ist  $K \cdot v := \{av \mid a \in K\} \leq V$ .
- (iii) Für  $U := \{(a_1, \dots, a_n) \in K^{1 \times n} \mid \sum_{i=1}^n a_i = 0\}$  ist  $U \leq K^{1 \times n}$ .
- (iv) Definiere

$$C(\mathbb{R}) := \{f : \mathbb{R} \rightarrow \mathbb{R} \mid f \text{ stetig} \}$$

$$C^\infty(\mathbb{R}) := \{f : \mathbb{R} \rightarrow \mathbb{R} \mid f \text{ beliebig oft stetig differenzierbar} \}$$

$$\text{Pol}(\mathbb{R}) := \{f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto a_n x^n + \dots + a_1 x^1 + a_0 \mid a_i \in \mathbb{R}, n \in \mathbb{N}_0\}$$

Mit  $\text{Pol}(\mathbb{R})$  wird also die Menge der Polynomfunktionen auf  $\mathbb{R}$  bezeichnet (siehe Bemerkung 2.3.1a für die Definition von Polynomfunktion.)

Dann ist  $\text{Pol}(\mathbb{R}) \leq C^\infty(\mathbb{R}) \leq C(\mathbb{R}) \leq \text{Abb}(\mathbb{R}, \mathbb{R})$ .

- (v)  $V = \mathbb{R}^2$  (euklidische Ebene)  
Geraden durch 0 sind Untervektorräume von  $V$ . Geraden, die nicht durch 0 gehen, sind keine Untervektorräume von  $V$ .
- (vi) Sei  $V$  ein  $K$ -Vektorraum und  $U_1, U_2 \leq V$ . Dann ist  $U_1 \cap U_2 \leq V$ . Setzen wir  $U_1 + U_2 := \{u_1 + u_2 \mid u_i \in U_i\} \subseteq V$ , dann ist  $U_1 + U_2 \leq V$ .
- (vii) Für jede Matrix  $A \in K^{m \times n}$  ist  $\mathbb{L}(A, 0)$  ein Unterraum von  $K^n$ .

**Bemerkung.** Es sei  $A \in K^{m \times n}$ . Der Unterraum  $\mathbb{L}(A, 0)$  von  $K^n$  heißt *Nullraum* von  $A$ .

## 6.2 Basis und Dimension

### 6.2.1 Linearkombinationen und Erzeugnis

Es sei  $K$  ein Körper,  $V$  ein  $K$ -Vektorraum.



**Definition.**

- (i) Es seien  $n \in \mathbb{N}$  und  $v_1, \dots, v_n \in V$ . Eine *Linearkombination des Tupels*  $(v_1, \dots, v_n)$  ist ein Vektor aus  $V$  der Form  $a_1v_1 + \dots + a_nv_n$  mit  $a_1, \dots, a_n \in K$ . Die Elemente  $a_1, \dots, a_n \in K$  heißen die *Koeffizienten* der Linearkombination. Ist  $v \in V$  mit  $v = a_1v_1 + \dots + a_nv_n$ , so sagen wir  $v$  wird durch die Linearkombination *dargestellt*. Die Linearkombination heißt *trivial*, wenn  $a_1 = \dots = a_n = 0$ , sonst *nicht-trivial*. Um Fallunterscheidungen zu vermeiden, definieren wir die Linearkombinationen des leeren Tupels von Vektoren als den Nullvektor.
- (ii) Sei  $M \subseteq V$ ,  $n \in \mathbb{N}_0$ . Eine *Linearkombination aus  $M$*  ist eine Linearkombination von  $(v_1, \dots, v_n)$  mit Vektoren  $v_1, \dots, v_n \in M$ . Wir benutzen die Konvention, dass  $(v_1, \dots, v_n)$  für  $n = 0$  das leere Tupel ist. Für  $M = \emptyset$  kann nur das leere Tupel von Elementen aus  $M$  gebildet werden.
- Die Menge  $\langle M \rangle$  aller Linearkombinationen aus  $M$  heißt die *lineare Hülle* von  $M$  oder das *Erzeugnis* von  $M$  oder der von  $M$  *erzeugte* oder *aufgespannte* Unterraum.
- (iii) Gibt es zu gegebenem  $v \in V$  eine Linearkombination von  $(v_1, \dots, v_n)$  (bzw. aus  $M$ ), die  $v$  darstellt, so sagen wir,  $v$  *lässt sich aus*  $(v_1, \dots, v_n)$  (bzw. aus  $M$ ) *linear kombinieren*.

**Bemerkung.** Linearkombinationen eines festen Tupels von Vektoren mit verschiedenen Koeffiziententupeln können denselben Vektor darstellen. Beispielsweise sind  $1v + (-1)v$  und  $0v + 0v$  Linearkombinationen des Tupels  $(v, v)$  mit verschiedenen Koeffizientenpaaren  $(1, -1)$  bzw.  $(0, 0)$ , die aber beide den Nullvektor darstellen.

Gleiches gilt für Linearkombinationen aus Mengen:  $6v$  und  $3(2v)$  sind zwei Linearkombinationen aus  $\{v, 2v\}$ , die als Linearkombination verschieden sind (weil sie Linearkombinationen verschiedener 1-Tupel sind), aber denselben Vektor darstellen. Ein weiteres Beispiel bilden die Linearkombinationen  $1v + 1(-v)$  und  $0v$  aus  $\{v, -v\}$ .

**Schreibweise.**  $\langle v_1, \dots, v_n \rangle := \langle \{v_1, \dots, v_n\} \rangle$ .

**Beispiel.**

- (i)  $\langle \emptyset \rangle = \{0\}$ .
- (ii) Für  $v \in V$  ist  $\langle v \rangle = Kv$  (siehe Beispiel 6.1.2(ii)).

- (iii) Es sei  $V$  der „euklidische Raum“, also der  $\mathbb{R}$ -Vektorraum  $\mathbb{R}^3$ . Für jedes  $v \in V \setminus \{0\}$  ist das Erzeugnis  $\langle v \rangle$  eine Gerade durch den Ursprung. Ist weiter  $w \in V$  und  $w \notin \langle v \rangle$ , so ist  $\langle v, w \rangle$  eine Ebene.

$$(iv) \quad V = \mathbb{R}^3, v_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, v_2 = \begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix}.$$

$$v_1 + v_2 = \begin{pmatrix} 0 \\ -2 \\ 2 \end{pmatrix}, v_1 - v_2 = \begin{pmatrix} 2 \\ 0 \\ -2 \end{pmatrix} \text{ sind Linearkombination von } (v_1, v_2).$$

$$\langle v_1, v_2 \rangle = \left\{ a \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} + b \begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix} \mid a, b \in \mathbb{R} \right\} = \left\{ \begin{pmatrix} a - b \\ -a - b \\ 2b \end{pmatrix} \mid a, b \in \mathbb{R} \right\}.$$

$$(v) \quad K = \mathbb{R}, V = C^\infty(\mathbb{R}), v_1 = \text{id}_{\mathbb{R}}, v_2 = \sin.$$

$$\begin{aligned} \langle v_1, v_2 \rangle &= \{a \cdot \text{id}_{\mathbb{R}} + b \cdot \sin \mid a, b \in \mathbb{R}\} \\ &= \{f : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto ax + b \sin(x) \mid a, b \in \mathbb{R}\} \end{aligned}$$

*Übung a.* Es seien  $v_1, v_2$  wie in Beispiel (iv). Wie prüft man, ob ein gegebenes  $v \in V$  in  $\langle v_1, v_2 \rangle$  liegt? Man zeige weiter:

$$\langle v_1, v_2 \rangle = \left\{ \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \mid a_1 + a_2 + a_3 = 0 \right\}.$$

**Satz.** Es sei  $M \subseteq V$ .

- (i)  $M \subseteq \langle M \rangle$ .
- (ii)  $\langle M \rangle \leq V$ .
- (iii)  $M \subseteq U \leq V \Rightarrow \langle M \rangle \subseteq U$ .  
*D.h.  $\langle M \rangle$  ist der kleinste Untervektorraum von  $V$ , der  $M$  enthält.*  
*D.h.  $\langle M \rangle$  ist das Minimum (bzgl. der Relation  $\subseteq$ , vgl. Definition 1.5.3) aller Untervektorräume von  $V$ , die  $M$  enthalten.*
- (iv)  $M \leq V \Leftrightarrow M = \langle M \rangle$ .
- (v)  $\langle \langle M \rangle \rangle = \langle M \rangle$ .

*Beweis.* (i)  $v$  ist Linearkombination von  $(v)$ .

- (ii) Wegen  $0 \in \langle M \rangle$  ist  $\langle M \rangle \neq \emptyset$ . Mit  $u, u' \in \langle M \rangle$  und  $a \in K$  sind offenbar auch  $u + u'$  und  $au$  Linearkombination aus  $M$ .
- (iii) Es sei  $M \subseteq U \leq V$ . Jede Linearkombination  $a_1 v_1 + \dots + a_r v_r$  mit  $v_1, \dots, v_r \in M$  liegt dann in  $U$ , d.h.  $\langle M \rangle \subseteq U$ .
- (iv) Ist  $M \leq V$  so kann in (iii)  $U = M$  gewählt werden, also  $M \subseteq \langle M \rangle \subseteq M$ , also  $M = \langle M \rangle$ . Ist  $M = \langle M \rangle$ , so gilt  $M \leq V$  nach (ii).
- (v) folgt aus (iii) mit Wahl  $U = \langle M \rangle$ .

□

*Übung* b. Man zeige, dass für alle Teilmengen  $M \subseteq V$  und alle  $v \in V$  gilt:  
 $v \in \langle M \rangle \iff \langle M \cup \{v\} \rangle = \langle M \rangle$ .

## 6.2.2 Zeilenraum und Spaltenraum

Es seien  $K$  ein Körper und  $A \in K^{m \times n}$  mit Zeilen  $z_1, \dots, z_m \in K^{1 \times n}$  und Spalten  $s_1, \dots, s_n \in K^m$ .

### Bemerkung.

- (i) Sei  $V = K^m$ , also  $s_1, \dots, s_n \in V$ . Wir haben

$$Ax = A \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \sum_{i=1}^n x_i s_i \text{ für jedes } x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in K^n,$$

d.h.  $Ax$  ist die Linearkombination von  $(s_1, \dots, s_n)$  mit Koeffizienten  $x_1, \dots, x_n$ .

- (ii) Sei  $W = K^{1 \times n}$ , also  $z_1, \dots, z_m \in W$ . Wir haben

$$yA = (y_1, \dots, y_m)A = \sum_{i=1}^m y_i z_i \text{ für jedes } y = (y_1, \dots, y_m) \in K^{1 \times m},$$

d.h.  $yA$  ist die Linearkombination von  $(z_1, \dots, z_m)$  mit Koeffizienten  $y_1, \dots, y_m$ .

**Beispiel.**  $A = \begin{pmatrix} 1 & 0 & -2 \\ 3 & 2 & 0 \end{pmatrix}$ .

$$A \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix} = 1 \cdot \begin{pmatrix} 1 \\ 3 \end{pmatrix} + 0 \cdot \begin{pmatrix} 0 \\ 2 \end{pmatrix} + (-1) \cdot \begin{pmatrix} -2 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 \\ 3 \end{pmatrix} - \begin{pmatrix} -2 \\ 0 \end{pmatrix} = \begin{pmatrix} 3 \\ 3 \end{pmatrix}$$

$$\begin{aligned}
\begin{pmatrix} 2 & -1 \end{pmatrix} A &= 2 \cdot \begin{pmatrix} 1 & 0 & -2 \end{pmatrix} + (-1) \cdot \begin{pmatrix} 3 & 2 & 0 \end{pmatrix} \\
&= \begin{pmatrix} 2 & 0 & -4 \end{pmatrix} - \begin{pmatrix} 3 & 2 & 0 \end{pmatrix} \\
&= \begin{pmatrix} -1 & -2 & -4 \end{pmatrix}
\end{aligned}$$

**Definition.**

- (i)  $\text{ZR}(A) := \langle \{z_1, \dots, z_m\} \rangle \leq K^{1 \times n}$  heißt *Zeilenraum von A*.
- (ii)  $\text{SR}(A) := \langle \{s_1, \dots, s_n\} \rangle \leq K^m$  heißt *Spaltenraum von A*.

*Übung.*

- (i)  $\text{SR}(A) = \{Ax \mid x \in K^n\}$ .
- (ii)  $b \in \text{SR}(A) \iff Ax = b$  lösbar.

### 6.2.3 Lineare Abhängigkeit

Es sei  $K$  ein Körper,  $V$  ein  $K$ -Vektorraum.

**Definition.** Es seien  $n \in \mathbb{N}$ ,  $\mathcal{T} = (v_1, \dots, v_n)$  ein  $n$ -Tupel über  $V$ . Eine *lineare Abhängigkeit von  $\mathcal{T}$*  ist eine nicht-triviale Linearkombination von  $\mathcal{T}$ , die den Nullvektor darstellt. Wir nennen  $\mathcal{T}$  *linear abhängig*, falls eine lineare Abhängigkeit von  $\mathcal{T}$  existiert, sonst *linear unabhängig*. Per Konvention ist das leere Tupel über  $V$  linear unabhängig.

Es sei nun  $M \subseteq V$ . Wir nennen  $M$  *linear abhängig*, wenn ein linear abhängiges Tupel  $(v_1, \dots, v_n)$  mit **paarweise verschiedenen**  $v_1, \dots, v_n \in M$  existiert. Andernfalls heißt  $M$  linear unabhängig. Per Konvention ist die leere Menge linear unabhängig.

**Bemerkung.**

- (i)  $(v_1, \dots, v_n)$  ist genau dann linear abhängig, wenn  $a_1, \dots, a_n \in K$  existieren, nicht alle  $a_i = 0$ , mit  $\sum_{i=1}^n a_i v_i = 0$ .
- (ii)  $(v_1, \dots, v_n)$  ist genau dann linear unabhängig, wenn jede Linearkombination von  $(v_1, \dots, v_n)$ , die 0 darstellt, trivial ist. D.h. wenn gilt:

$$\sum_{i=1}^n a_i v_i = 0 \Rightarrow a_1 = \dots = a_n = 0.$$

- (iii)  $M$  ist genau dann linear abhängig, wenn paarweise verschiedene  $v_1, \dots, v_n \in M$  existieren ( $n \in \mathbb{N}$ ) sowie  $(a_1, \dots, a_n) \in K^{1 \times n} \setminus \{0\}$  mit  $\sum_{i=1}^n a_i v_i = 0$ .
- (iv)  $M$  linear abhängig  $\Rightarrow$  jedes  $M' \supseteq M$  linear abhängig
- (v)  $M$  linear unabhängig  $\Rightarrow$  jedes  $M' \subseteq M$  linear unabhängig

**Beispiel.**

- (i)  $0 \in M \Rightarrow M$  linear abhängig
- (ii)  $(\dots, v, \dots, v, \dots)$  linear abhängig
- (iii)  $v \neq 0 \Rightarrow \{v\}$  linear unabhängig
- (iv)  $\emptyset$  ist linear unabhängig
- (v) Es sei  $K = \mathbb{Q}$  und  $V = \mathbb{Q}^2$ . Dann ist  $\left\{ \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 3 \\ 4 \end{pmatrix} \right\}$  linear unabhängig:

Seien  $a_1, a_2 \in \mathbb{Q}$  mit  $a_1 \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix} + a_2 \cdot \begin{pmatrix} 3 \\ 4 \end{pmatrix} = 0$ , d.h.  $\begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} \cdot \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = 0$ .

$$\begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} \xrightarrow{\text{Gau\ss}} \begin{pmatrix} 1 & 3 \\ 0 & -2 \end{pmatrix}$$

$\Rightarrow \begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} \cdot x = 0$  ist eindeutig lösbar (d.h. nur trivial lösbar).

Also folgt  $a_1 = a_2 = 0$ .

Dagegen ist  $\left\{ \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 3 \\ 4 \end{pmatrix}, \begin{pmatrix} 5 \\ 6 \end{pmatrix} \right\}$  linear abhängig:

$$-\begin{pmatrix} 1 \\ 2 \end{pmatrix} + 2\begin{pmatrix} 3 \\ 4 \end{pmatrix} - \begin{pmatrix} 5 \\ 6 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

- (vi) Die Spalten einer Matrix  $A \in K^{m \times n}$  sind genau dann linear unabhängig, wenn  $Ax = 0$  nur trivial lösbar ist.

*Übung* a. Es sei  $A \in K^{m \times n}$ . Man zeige:

- (i) Ist  $A$  in Zeilenstufenform, und seien  $z_1, \dots, z_r \in K^{1 \times m}$ , die nicht-Null-Zeilen von  $A$ . Dann ist  $(z_1, \dots, z_r)$  linear unabhängig.

- (ii) Ist  $A$  in Zeilenstufenform, und seien  $s_1, \dots, s_r \in K^n$  die Spalten von  $A$ , die zu den Stufenindizes gehören. Dann ist  $(s_1, \dots, s_r)$  linear unabhängig.
- (iii) Die Zeilen von  $E_n$  sind linear unabhängig.
- (iv) Die Spalten von  $E_n$  sind linear unabhängig.

*Übung b.* Man definiere eine „Spaltenstufenform“ von  $A$  und zeige, dass die nicht-Null-Spalten einer Matrix in Spaltenstufenform linear unabhängig sind.

*Übung c.* Man zeige, dass es für jede linear abhängige Menge  $M \subseteq V$  ein  $v \in M$  mit  $\langle M \setminus \{v\} \rangle = \langle M \rangle$  gibt.

*Übung d.* Es seien  $K = \mathbb{R}, V = \mathbb{R}^2, u_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, u_2 = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, u_3 = \begin{pmatrix} -3 \\ -3 \end{pmatrix}$ . Man zeige, dass die Menge  $M = \{u_1, u_2, u_3\}$  linear abhängig ist. Für welche  $w \in M$  gilt  $\langle M \rangle = \langle M \setminus \{w\} \rangle$ ?

*Übung e.* Es seien  $u, v \in V$ . Wann ist  $(u, v)$  linear abhängig? Wann ist  $\{u, v\}$  linear abhängig?

### 6.2.4 Basen

Es sei  $K$  ein Körper,  $V$  ein  $K$ -Vektorraum.

**Definition.** Eine Teilmenge  $M \subseteq V$  heißt *Erzeugendensystem* von  $V$ , wenn  $V = \langle M \rangle$  ist. Eine *Basis* von  $V$  ist ein linear unabhängiges Erzeugendensystem von  $V$ .

Die *Länge* eines Erzeugendensystems  $M$  ist definiert als die Zahl  $|M|$  falls  $|M| < \infty$ , und sonst als  $\infty$  (unendlich). Wenn es ein endliches Erzeugendensystem gibt, so nennen wir den Vektorraum *endlich erzeugt*.

**Beispiel.**

- (i) Die leere Menge  $\emptyset$  ist Basis des trivialen  $K$ -Vektorraums  $\{0\}$ .
- (ii)  $V = K^n$ . Für  $1 \leq i \leq n$  sei

$$e_i := \begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \quad \text{der } i\text{-te Einheitsvektor (1 an der } i\text{-ten Stelle).}$$

Dann ist  $\{e_1, \dots, e_n\}$  Basis von  $K^n$ , genannt die *Standardbasis*.

- (iii)  $V = K^{m \times n}$ . Für  $1 \leq i \leq m$ ,  $1 \leq j \leq n$  sei  $E_{ij} \in K^{m \times n}$  die Matrix mit einer 1 an Position  $(i, j)$  und Nullen sonst. Dann ist

$$\{E_{11}, \dots, E_{1n}, E_{21}, \dots, E_{2n}, E_{31}, \dots, E_{mn}\}$$

eine Basis von  $V$ , genannt die *Standardbasis*.

- (iv)  $\{1, i\}$  ist eine Basis des  $\mathbb{R}$ -Vektorraums  $\mathbb{C}$  (hier ist  $i \in \mathbb{C}$  mit  $i^2 = -1$ ).
- (v)  $\{1\}$  ist eine Basis des  $\mathbb{C}$ -Vektorraums  $\mathbb{C}$ .
- (vi) Der  $\mathbb{Q}$ -Vektorraum  $\mathbb{R}$  ist nicht endlich-erzeugt.
- (vii) Der  $\mathbb{R}$ -Vektorraum  $\text{Pol}(\mathbb{R})$  hat die unendliche Basis  $\{1, x, x^2, x^3, \dots\}$ .

Beweis zu Basen von  $\mathbb{C}$ : Elemente haben eindeutige Form  $a + bi$ .

**Satz.** Für  $B \subseteq V$  sind äquivalent:

- (i)  $B$  ist eine Basis von  $V$ .
- (ii)  $B$  ist ein minimales Erzeugendensystem von  $V$ .  
(D.h. keine echte Teilmenge von  $B$  ist Erzeugendensystem von  $V$ .)
- (iii)  $B$  ist eine maximale linear unabhängige Teilmenge von  $V$ .  
(D.h. keine echte Obermenge von  $B$  in  $V$  ist linear unabhängig.)

Wir benötigen für den Beweis des Satzes das folgende Lemma, welches einen Zusammenhang zwischen den Begriffen Erzeugnis und lineare Abhängigkeit herstellt.

**Lemma.** Es sei  $M \subseteq V$ ,  $v \notin M$ .

- (i) Ist  $v \in \langle M \rangle$ , dann ist  $M \cup \{v\}$  linear abhängig.
- (ii) Wenn  $M$  linear unabhängig und  $M \cup \{v\}$  linear abhängig ist, dann ist  $v \in \langle M \rangle$ .

*Beweis.* (i) Sei  $v = \sum_{i=1}^n a_i v_i$  mit  $v_i \in M$ . Durch Zusammenfassung von Summanden können wir die  $v_i$  als paarweise verschieden annehmen. Nach Voraussetzung  $v \notin M$  ist auch  $v$  von allen  $v_i$  verschieden. Es folgt, dass  $1v - \sum_{i=1}^n a_i v_i = 0$  eine lineare Abhängigkeit von  $M \cup \{v\}$  ist.

(ii) Sei  $\sum_{i=1}^n a_i v_i$  eine lineare Abhängigkeit von  $v_i \in M \cup \{v\}$ . Da  $M$  linear unabhängig ist, kommt  $v$  in dieser Linearkombination mit Koeffizient  $\neq 0$  vor (sonst wäre schon  $M$  linear abhängig). Wir nehmen oBdA  $v = v_1$  an. Dann folgt  $v = -\sum_{i=2}^n \frac{a_i}{a_1} v_i$ . Wegen  $v_2, \dots, v_n \neq v$  sind  $v_2, \dots, v_n \in M$ , also  $v \in \langle M \rangle$ .  $\square$

*Übung a.* Es sei  $M \subseteq V$  linear unabhängig. Gilt für jedes  $v \in V$ :  $M \cup \{v\}$  linear abhängig  $\Leftrightarrow v \in \langle M \rangle$ ?

*Beweis des Satzes.* (i) $\Rightarrow$ (ii): Sei  $B$  eine Basis von  $V$ . Dann ist  $B$  Erzeugendensystem von  $V$ . Um zu zeigen, dass  $B$  minimal mit dieser Eigenschaft ist, wählen wir ein beliebiges  $v \in B$ , setzen  $M := B \setminus \{v\}$  und zeigen, dass  $M$  kein Erzeugendensystem von  $V$  ist. In der Tat, wegen  $v \notin M$  and  $M \cup \{v\} = B$  linear unabhängig, gilt nach Lemma (i) (Kontraposition):  $v \notin \langle M \rangle$ .

(ii) $\Rightarrow$ (i): Sei  $B$  ein minimales Erzeugendensystem. Annahme:  $B$  ist linear abhängig. Dann gibt es nach Übung 6.2.3c ein  $v \in B$  mit  $\langle B \setminus \{v\} \rangle = \langle B \rangle$ . Dies steht im Widerspruch zur Minimalität von  $B$ , also ist die Annahme falsch, d.h.  $B$  ist linear unabhängig.

(i) $\Rightarrow$ (iii) Sei  $B$  eine Basis von  $V$ . Dann ist  $B$  linear unabhängige Teilmenge von  $V$ . Um zu zeigen, dass  $B$  maximal mit dieser Eigenschaft ist, wählen wir ein beliebiges  $v \in V \setminus B$  und zeigen, dass  $B \cup \{v\}$  linear abhängig ist. Das ist gerade die Aussage von Lemma (i).

(iii) $\Rightarrow$ (ii): Sei  $B$  maximale linear unabhängige Teilmenge von  $V$ . Um zu zeigen, dass  $B$  auch Erzeugendensystem von  $V$  ist, wählen wir ein beliebiges  $v \in V$  und zeigen  $v \in \langle B \rangle$ . Sei oBdA  $v \notin B$  (sonst ist  $v \in B \subseteq \langle B \rangle$  klar). Nach Voraussetzung ist dann  $B \cup \{v\}$  linear abhängig, also  $v \in \langle B \rangle$  nach Lemma (ii).  $\square$

**Folgerung.** (*Basisauswahl*) Jedes endliche Erzeugendensystem von  $V$  enthält eine Basis von  $V$ . Insbesondere hat jeder endlich-erzeugte Vektorraum eine endliche Basis.

*Beweis.* Es sei  $M$  ein endliches Erzeugendensystem von  $V$ , also  $M \subseteq V$  mit  $\langle M \rangle = V$ . Wenn  $M$  keine Basis ist, dann ist  $M$  nach dem Satz kein minimales Erzeugendensystem. Also gibt es eine echte Teilmenge  $M' \subsetneq M$  mit  $\langle M' \rangle = V$ . Da  $M$  endlich ist, kommt man nach endlich vielen Wiederholungen dieses Schlusses zu einem minimalen Erzeugendensystem, also zu einer Basis.  $\square$

**Bemerkung.** Die Folgerung gilt auch ohne die Annahme, dass  $M$  endlich ist. Insbesondere hat jeder Vektorraum eine Basis. Für den Beweis benötigt man allerdings das *Lemma von Zorn* (siehe Vorlesung *Mathematische Logik I*).

**Algorithmus.** Die Prozedur *Basisauswahl* in untenstehender Abbildung liefert zu jedem endlichen Erzeugendensystem  $M$  von  $V$  eine Teilmenge  $B \subseteq M$ , die Basis von  $V$  ist.

*Übung.* Man zeige die Korrektheit des Algorithmus *Basisauswahl*, d.h. dass er abbricht und eine Basis liefert. Man erkläre weiterhin, mit welchem Ansatz man die Schleifenbedingung prüft und wie man eine nicht-triviale Linearkombination in Zeile 3 findet. (*Hinweis:* Man löse ein geeignetes LGS.)



BASISAUSWAHL( $M$ )

```

1   $B \leftarrow M$ 
2  while  $B$  linear abhängig
3  do Wähle lineare Abhängigkeit  $\sum_{i=1}^n a_i v_i$  von  $B$ 
4       $v \leftarrow$  ein beliebiges  $v_i$  mit  $a_i \neq 0$ 
5       $B \leftarrow B \setminus \{v\}$ 
6  return  $B$ 

```

Abbildung 6.1: Prozedur Basisauswahl

### 6.2.5 Dimension

Es sei  $K$  ein Körper,  $V$  ein **endlich-erzeugter**  $K$ -Vektorraum,  $B$  eine endliche Basis von  $V$  (existiert nach Folgerung 6.2.4), und  $n = |B|$ .

**Lemma.** Für jede Teilmenge  $M \subseteq V$  gilt:

- (i)  $|M| > n \Rightarrow M$  linear abhängig
- (ii)  $M$  linear unabhängig  $\Rightarrow |M| \leq n$
- (iii)  $M$  erzeugt  $V \Rightarrow |M| \geq n$

Insbesondere ist jede Basis von  $V$  endlich.

*Beweis.* (i) Sei  $B = \{v_1, \dots, v_n\}$ . Wegen  $|M| > n$  gibt es paarweise verschiedene  $w_1, \dots, w_{n+1} \in M$ . Schreibe jedes  $w_j$  als Linearkombination der Basisvektoren  $v_1, \dots, v_n$ :

$$w_j = \sum_{i=1}^n a_{ij} v_i, \quad a_{ij} \in K, j = 1, \dots, n+1.$$

Betrachte das homogene LGS über  $K$

$$\sum_{j=1}^{n+1} a_{ij} x_j = 0 \quad \text{für } i = 1, \dots, n,$$

bestehend aus  $n$  Gleichungen in  $n+1$  Unbekannten  $x_1, \dots, x_{n+1}$ . Da es mehr Unbekannte als Gleichungen gibt, besitzt es eine nicht-triviale Lösung (Bemerkung 3.4.4(ii)), d.h. es gibt  $c_1, \dots, c_{n+1} \in K$ , nicht alle  $c_j = 0$ , mit

$$\sum_{j=1}^{n+1} a_{ij} c_j = 0 \quad \text{für } i = 1, \dots, n.$$

Es folgt

$$\begin{aligned} \sum_{j=1}^{n+1} c_j w_j &= \sum_{j=1}^{n+1} c_j \sum_{i=1}^n a_{ij} v_i \stackrel{(V2)}{=} \sum_{j=1}^{n+1} \sum_{i=1}^n c_j (a_{ij} v_i) \\ &\stackrel{(V3)}{=} \sum_{i=1}^n \sum_{j=1}^{n+1} (c_j a_{ij}) v_i \stackrel{(V1)}{=} \sum_{i=1}^n \left( \sum_{j=1}^{n+1} c_j a_{ij} \right) v_i = \sum_{i=1}^n 0 v_i = 0. \end{aligned}$$

Da nicht alle  $c_j = 0$  und die  $w_j$  paarweise verschieden sind, ist  $M$  linear abhängig.

(ii) ist die Kontraposition von (i). Dies impliziert, dass jede Basis von  $V$  endlich ist.

(iii) Wir nehmen oBdA an, dass  $M$  endlich ist (sonst ist  $|M| \geq n$  klar). Als endliches Erzeugendensystem von  $V$  enthält  $M$  eine (endliche) Basis  $B'$  von  $V$  (Folgerung 6.2.4). Teil (ii) mit  $B$  als  $M$  und  $B'$  als  $B$  besagt:  $B$  linear unabhängig  $\Rightarrow |B| \leq |B'|$ . Also  $n = |B| \leq |B'| \leq |M|$ .  $\square$

**Satz a.** *Es gilt*

$$\begin{aligned} n &= \max\{|M| \mid M \subseteq V \text{ linear unabhängig}\} \\ &= \min\{|M| \mid M \subseteq V, \langle M \rangle = V\} \end{aligned}$$

Insbesondere haben alle Basen die gleiche Länge.

*Beweis.* Bezeichne das angegebene Maximum mit  $l$  und das Minimum mit  $k$ . Da  $B$  linear unabhängiges Erzeugendensystem ist, gilt  $k \leq n \leq l$ . Nach dem Lemma gilt  $l \leq n \leq k$ . Zusammen folgt die Gleichheit.  $\square$

**Definition.** Die Zahl  $n$  wird *Dimension* von  $V$  genannt, geschrieben  $\dim V$  bzw. genauer  $\dim_K V$ . Für nicht endlich-erzeugte Vektorräume setzen wir  $\dim_K V := \infty$ .

**Folgerung a.** *Für jede Teilmenge  $M \subseteq V$  sind äquivalent:*

- (i)  $M$  ist Basis von  $V$ .
- (ii)  $M$  ist linear unabhängig und  $|M| = n$ .
- (iii)  $M$  erzeugt  $V$  und  $|M| = n$ .

Frage:  $V \subsetneq W$  mit gleicher endlicher Dimension.  $V = W$ ?

*Beweis.* Das folgt aus dem Satz und der Charakterisierung in Satz 6.2.4. (Details als Übung)  $\square$

**Beispiel.**

- (i)  $\dim_K \{0\} = 0$ . (Basis:  $\emptyset$ .)
- (ii)  $\dim_K K^n = n$ . (Standardbasis  $e_1, \dots, e_n$ .)
- (iii)  $\dim_K K^{m \times n} = nm$ . (Basis:  $\{E_{ij} \mid 1 \leq i \leq m, 1 \leq j \leq n\}$ .)
- (iv)  $\dim_{\mathbb{Q}} \mathbb{R} = \dim_{\mathbb{Q}} \mathbb{C} = \infty$ . (Basis: unbekannt.)
- (v) Die  $\mathbb{R}$ -Vektorräume  $\text{Abb}(\mathbb{R}, \mathbb{R})$ ,  $C(\mathbb{R})$ ,  $C^\infty(\mathbb{R})$ ,  $\text{Pol}(\mathbb{R})$ ,  $\text{Abb}(\mathbb{N}, \mathbb{R})$  haben  $\dim_{\mathbb{R}} = \infty$ .
- (vi)  $\dim_{\mathbb{R}} \mathbb{C} = 2$ ,  $\dim_{\mathbb{C}} \mathbb{C} = 1$ .  
 $\{1, i\}$  ist Basis von  $\mathbb{C}$  als  $\mathbb{R}$ -Vektorraum, aber  $\{1, i\}$  ist linear abhängig über  $\mathbb{C}$ :  
 $a \cdot 1 + b \cdot i = 0 \xrightarrow{a, b \in \mathbb{R}} a = b = 0$ .  
 $a \cdot 1 + b \cdot i = 0$  für  $a = i, b = -1 \in \mathbb{C}$ .
- (vii) Es seien  $v_1, \dots, v_n \in V$  paarweise verschieden.  
 Sind  $\{v_1, \dots, v_n\}$  linear unabhängig, dann ist  $\dim_K \langle v_1, \dots, v_n \rangle = n$   
 (Basis von  $\langle v_1, \dots, v_n \rangle$ :  $\{v_1, \dots, v_n\}$ ).

*Übung a.* Man überlege sich ein Beispiel, in dem  $V$  zugleich Vektorraum über zwei verschiedenen Körpern ist und dabei unterschiedliche Dimensionen hat.

*Übung b.* Es sei  $M$  ein endliches Erzeugendensystem von  $V$ . Man zeige:  $n = \max\{|M'| \mid M' \subseteq M, M' \text{ linear unabhängig}\}$ . Wir sagen dazu: „ $n$  ist die Maximalzahl linear unabhängiger Elemente von  $M$ “.

**Satz b.** *Es sei  $V$  ein endlich-dimensionaler  $K$ -Vektorraum.*

*Eine Teilmenge  $M \subseteq V$  ist genau dann Basis von  $V$ , wenn jedes  $v \in V$  eine eindeutige Darstellung als Linearkombination von Elementen aus  $M$  besitzt.*

*Beweis.* 1.  $M$  ist genau dann linear unabhängig, wenn jedes  $v \in V$  höchstens eine Darstellung als Linearkombination von Elementen aus  $M$  besitzt. (Details siehe Vorlesung).

2.  $M$  erzeugt genau dann  $V$ , wenn jedes  $v \in V$  mindestens eine Darstellung als Linearkombination von Elementen aus  $M$  besitzt. (klar)  $\square$

**Folgerung b.** *Ist  $K$  ein endlicher Körper mit  $q$  Elementen und  $\dim_K V = n$ , so gilt  $|V| = q^n$ . Im Allgemeinen misst die Dimension die „Größe“ eines Vektorraums (nicht nur für endliche Körper).*

### 6.2.6 Basisergänzung

Es seien  $K$  ein Körper und  $V$  ein **endlich-dimensionaler**  $K$ -Vektorraum.

**Satz.** *Jede linear unabhängige Teilmenge von  $V$  lässt sich zu einer Basis ergänzen.*

*Beweis.* Es sei  $M \subseteq V$  linear unabhängig. Wenn  $M$  keine Basis ist, dann ist  $M$  nach Satz 6.2.4 nicht maximal linear unabhängig, besitzt also eine echte Obermenge  $M' \supsetneq M$  in  $V$ , die linear unabhängig ist. Da  $|M'| \leq \dim V < \infty$  (Teil (ii) von Satz 6.2.5a), gelangt man nach endlich vielen Wiederholungen dieses Schlusses zu einer maximal linear unabhängigen Teilmenge von  $V$ , also zu einer Basis.  $\square$

**Algorithmus.** *Die Prozedur Basisergänzung der untenstehenden Abbildung liefert zu jeder linear unabhängigen Teilmenge  $M \subseteq V$  (insbesondere auch zu  $M = \emptyset$ ) eine Obermenge  $B \supseteq M$ , die Basis von  $V$  ist.*

*Beweis.* Nach Voraussetzung ist  $B$  in Schritt 1 linear unabhängig. Wir zeigen, dass ‘ $B$  linear unabhängig’ eine Schleifeninvariante ist: Aus  $B$  linear unabhängig und  $v \notin \langle B \rangle$  folgt nach Teil (ii) von Lemma 6.2.4, dass auch  $B \cup \{v\}$  linear unabhängig ist. Nach Teil (ii) von Satz 6.2.5a ist also stets  $|B| \leq \dim V < \infty$ . Da  $B$  mit jedem Durchlauf größer wird, bricht die Schleife ab. Bei Abbruch ist  $\langle B \rangle = V$ , also  $B$  eine Basis. (Statt  $\langle B \rangle = V$  ist als Abbruchkriterium auch  $|B| = \dim V$  erlaubt, sofern  $\dim V$  bekannt ist.)  $\square$

BASISERGÄNZUNG( $M$ )

```

1   $B \leftarrow M$ 
2  while  $\langle B \rangle \neq V$ 
3    do  $v \leftarrow$  beliebiges Element aus  $V \setminus \langle B \rangle$ 
4       $B \leftarrow B \cup \{v\}$ 
5  return  $B$ 
```

Abbildung 6.2: Prozedur Basisergänzung

**Beispiel a.**  $V := \mathbb{R}^3$ ,  $M := \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\}$ .

Wähle  $v \notin \langle M \rangle$ , z.B.  $v = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ .

$$M' := M \cup \{v\} = \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} \right\}$$

Wähle  $w \notin \langle M' \rangle$ , z.B.  $w = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$ .

$$M'' := M' \cup \{w\} = \left\{ \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

$M''$  ist Basis weil  $|M''| = 3 = \dim V$  (Folgerung 6.2.5a).

**Folgerung.** Für jeden Unterraum  $U \leq V$  gelten:

(i)  $\dim_K U \leq \dim_K V$ ,

(ii)  $\dim_K U = \dim_K V \Rightarrow U = V$ .

*Beweis.* Sei  $B$  eine Basis von  $U$ . Dann ist  $|B| = \dim_K U$ , und  $B$  ist auch linear unabhängig als Teilmenge von  $V$ . Nach Satz 6.2.5a folgt  $\dim_K U = |B| \leq \dim_K V$ . Falls  $|B| = \dim_K U = \dim_K V$ , so ist  $B$  nach Folgerung 6.2.5a auch Basis von  $V$ . Also  $U = \langle B \rangle = V$ .  $\square$

**Beispiel b.** Mit dem Dimensionsbegriff und der Folgerung kann man sehr leicht alle  $U$  Unterräume von  $\mathbb{R}^3$  bestimmen:

$\dim U = 0 :$	$U = \langle \emptyset \rangle = \{0\}$	(Ursprung)
$\dim U = 1 :$	$U = \langle v \rangle, v \neq 0$	(alle Geraden durch 0)
$\dim U = 2 :$	$U = \langle v, w \rangle, v, w \neq 0, w \notin \langle v \rangle$	(alle Ebenen durch 0)
$\dim U = 3 :$	$U = V$	(ganz $\mathbb{R}^3$ )

### 6.2.7 Innere direkte Summen

Es sei  $V$  ein **endlich-dimensionaler**  $K$ -Vektorraum. Wir erinnern an die Summe von Untervektorräumen: Sind  $U_1, U_2 \leq V$ , dann ist

$$U_1 + U_2 := \{u_1 + u_2 \mid u_i \in U_i, i = 1, 2\}$$

ein Untervektorraum von  $V$ , die *Summe* von  $U_1$  und  $U_2$ .

**Bemerkung a.** Es seien  $U_1, U_2 \leq V$ . Dann ist

$$U_1 + U_2 = \langle U_1 \cup U_2 \rangle,$$

d.h.  $U_1 + U_2$  ist der kleinste Untervektorraum von  $V$ , der  $U_1$  und  $U_2$  enthält.

**Satz a.** Es seien  $U_1, U_2 \leq V$  Untervektorräume von  $V$ . Wir setzen  $W := U_1 + U_2$ . Dann sind die folgenden Aussagen äquivalent:

(i)  $U_1 \cap U_2 = \{0\}$ .

(ii) Jedes  $w \in W$  besitzt eine **eindeutige** Darstellung als

$$w = u_1 + u_2$$

mit  $u_i \in U_i$ ,  $i = 1, 2$ .

(iii) Sind  $B_i$  Basen von  $U_i$  für  $i = 1, 2$ , dann ist  $B_1 \cap B_2 = \emptyset$  und

$$B_1 \cup B_2$$

eine Basis von  $W$ .

*Beweis.* Siehe Vorlesung. □

**Definition a.** Es seien  $U_1, U_2 \leq V$  mit  $V = U_1 + U_2$  und  $U_1 \cap U_2 = \{0\}$ . Dann schreiben wir

$$V = U_1 \oplus U_2$$

und nennen  $V$  die (*innere*) direkte Summe von  $U_1$  und  $U_2$ .

**Beispiel a.** Es sei  $K = \mathbb{R}$  und  $V = \mathbb{R}^3$ . Wir betrachten die folgenden Untervektorräume von  $V$ .

$$U_1 = \left\{ \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} \mid x, y \in \mathbb{R} \right\},$$

$$U_2 = \mathbb{R} \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \left\langle \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right\rangle,$$

$$U_3 = \mathbb{R} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \left\langle \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\rangle,$$

$$U_4 = \left\{ \begin{pmatrix} 0 \\ y \\ z \end{pmatrix} \mid y, z \in \mathbb{R} \right\}.$$

Dann ist  $V = U_1 \oplus U_2$  und  $V = U_1 \oplus U_3$ , aber  $V$  ist keine innere direkte Summe von  $U_1$  und  $U_4$ , da  $U_1 \cap U_4 \neq \{0\}$  ist. Allerdings gilt  $V = U_1 + U_4$ .

**Folgerung a.** Ist  $V$  endlich erzeugt und

$$V = U_1 \oplus U_2$$

mit  $U_1, U_2 \leq V$ , dann ist

$$\dim_K V = \dim_K U_1 + \dim_K U_2.$$

*Beweis.* Wir benutzen die Notation aus Satz a. Die Basen  $B_i$  sind endlich für  $i = 1, 2$ , da die Untervektorräume  $U_1$  und  $U_2$  des endlich-dimensionalen Vektorraums  $V$  endlich-dimensional sind. Wegen  $B_i \subseteq U_i$  für  $i = 1, 2$  ist  $B_1 \cap B_2 = \emptyset$ , denn ein Element aus  $B_1 \cap B_2$  liegt in  $U_1 \cap U_2 = \{0\}$  und ist als Element einer Basis nicht der Nullvektor. Die Behauptung folgt nun aus Satz a.  $\square$

**Satz b.** Es sei  $V$  endlich-dimensional. Es seien  $U, U'$  Untervektorräume von  $V$  und  $\{v_1, \dots, v_m\}$  eine Basis von  $U \cap U'$ , die wir zu einer Basis

$$\{v_1, \dots, v_m, w_1, \dots, w_n\} \text{ von } U$$

und zu einer Basis

$$\{v_1, \dots, v_m, w'_1, \dots, w'_{n'}\} \text{ von } U'$$

ergänzen. Dann ist

$$\{v_1, \dots, v_m, w_1, \dots, w_n, w'_1, \dots, w'_{n'}\}$$

eine Basis von  $U + U'$ .

*Beweis.* Setze  $B := \{v_1, \dots, v_m, w_1, \dots, w_n, w'_1, \dots, w'_{n'}\}$ . Weil  $B$  Erzeugendensysteme von  $U$  und  $U'$  enthält, ist  $\langle B \rangle = \langle U \cup U' \rangle = U + U'$ , d.h.  $B$  ist eine Erzeugendensystem von  $U + U'$ . Wir zeigen nun, dass  $B$  linear unabhängig ist. Seien dazu  $a_1, \dots, a_m, b_1, \dots, b_n, b'_1, \dots, b'_{n'} \in K$  mit

$$\sum_{i=1}^m a_i v_i + \sum_{i=1}^n b_i w_i + \sum_{i=1}^{n'} b'_i w'_i = 0.$$

Dann ist

$$\sum_{i=1}^m a_i v_i + \sum_{i=1}^n b_i w_i = \sum_{i=1}^{n'} (-b'_i) w'_i. \quad (6.1)$$

Der Vektor auf der linken Seite von (6.1) liegt in  $U$ , der Vektor auf der rechten Seite von (6.1) in  $U'$ . Also ist  $\sum_{i=1}^{n'} (-b'_i) w'_i \in U \cap U'$ . Da  $\{v_1, \dots, v_m\}$  eine Basis von  $U \cap U'$  ist, existieren  $c_1, \dots, c_m \in K$  mit

$$\sum_{i=1}^{n'} (-b'_i) w'_i = \sum_{i=1}^m c_i v_i.$$

Hieraus folgt

$$\sum_{i=1}^m c_i v_i + \sum_{i=1}^{n'} b'_i w'_i = 0.$$

Da  $\{v_1, \dots, v_m, w'_1, \dots, w'_{n'}\}$  eine Basis von  $U'$  und damit linear unabhängig ist, folgt aus der letzten Gleichung  $c_i = 0$  für alle  $1 \leq i \leq m$  und  $b'_i = 0$  für alle  $1 \leq i \leq n'$ . Setzt man dies in Gleichung (6.1) ein, erhält man

$$\sum_{i=1}^m a_i v_i + \sum_{i=1}^n b_i w_i = 0.$$

Aus der linearen Unabhängigkeit von  $\{v_1, \dots, v_m, w_1, \dots, w_n\}$  ergibt sich daraus  $a_i = 0$  für alle  $1 \leq i \leq m$  und  $b_i = 0$  für alle  $1 \leq i \leq n$ . Damit ist gezeigt, dass  $B$  linear unabhängig ist.  $\square$

**Folgerung b** (Dimensionssatz für Summen von Untervektorräumen). *Unter den Voraussetzungen von Satz b gilt*

$$\dim_K(U + U') = \dim_K U + \dim_K U' - \dim_K(U \cap U').$$

*Beweis.* Mit den Bezeichnungen von Satz b ist  $\dim_K(U \cap U') = m$ ,  $\dim_K U = m + n$ ,  $\dim_K U' = m + n'$  und  $\dim_K(U + U') = m + n + n'$ . Also ist

$$\begin{aligned} \dim_K U + \dim_K U' - \dim_K(U \cap U') &= m + n + m + n' - m \\ &= m + n + n' \\ &= \dim_K(U + U'), \end{aligned}$$

was zu beweisen war.  $\square$

*Übung a.* Es sei  $V$  endlich-dimensional und  $U, U' \leq V$  mit

$$\dim U + \dim U' > \dim V.$$

Zeigen Sie, dass  $U \cap U' \neq \{0\}$  ist.

Mithilfe von inneren direkten Summen können wir Projektionen definieren.

**Definition b.** Es seien  $U_1, U_2 \leq V$  mit  $V = U_1 \oplus U_2$ . Für  $i = 1, 2$  ist die *Projektion* von  $V$  auf  $U_i$  definiert durch

$$\text{pr}_i: V \rightarrow U_i, \quad u_1 + u_2 \mapsto u_i.$$

(Nach Satz a besitzt jedes Element  $v \in V$  eine eindeutige Darstellung als  $v = u_1 + u_2$  mit  $u_1 \in U_1$  und  $u_2 \in U_2$ .)



**Beispiel b.** Es seien  $V$  und  $U_i$ ,  $i = 1, 2, 3$  wie in Beispiel **a**. Betrachten wir die Zerlegung  $V = U_1 \oplus U_2$ , dann ist

$$\text{pr}_1\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = \begin{pmatrix} x \\ y \\ 0 \end{pmatrix},$$

denn  $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ z \end{pmatrix}$  ist die eindeutige Zerlegung von  $v = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$  als

$v = u_1 + u_2$  mit  $u_1 = \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} \in U_1$  und  $u_2 = \begin{pmatrix} 0 \\ 0 \\ z \end{pmatrix} \in U_2$ .

Betrachten wir dagegen die Zerlegung  $V = U_1 \oplus U_3$ , dann ist

$$\text{pr}_1\left(\begin{pmatrix} x \\ y \\ z \end{pmatrix}\right) = \begin{pmatrix} x - z \\ y - z \\ 0 \end{pmatrix},$$

denn  $\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x - z \\ y - z \\ 0 \end{pmatrix} + \begin{pmatrix} z \\ z \\ z \end{pmatrix}$  ist die eindeutige Zerlegung von  $v = \begin{pmatrix} x \\ y \\ z \end{pmatrix}$

als  $v = u_1 + u_3$  mit  $u_1 = \begin{pmatrix} x - z \\ y - z \\ 0 \end{pmatrix} \in U_1$  und  $u_3 = \begin{pmatrix} z \\ z \\ z \end{pmatrix} \in U_3$ .

**Definition c.** Es sei  $U \leq V$ . Ein Untervektorraum  $U' \leq V$  heißt *Komplement* zu  $U$  in  $V$ , falls  $V = U \oplus U'$  ist.

**Beispiel c.** Es seien  $V$  und  $U_i$ ,  $i = 1, 2, 3, 4$  wie in Beispiel **a**. Dann sind  $U_2$  und  $U_3$  Komplemente zu  $U_1$  in  $V$ . Dagegen ist  $U_4$  kein Komplement zu  $U_1$  in  $V$ .

**Bemerkung b.** Es sei  $V$  endlich-dimensional und  $U \leq V$ . Dann existiert ein Komplement zu  $U$  in  $V$ .

*Beweis.* Wähle eine Basis  $\{v_1, \dots, v_m\}$  von  $U$ , und ergänze diese zu einer Basis  $\{v_1, \dots, v_m, v_{m+1}, \dots, v_n\}$  von  $V$ . Dies ist nach dem Basisergänzungssatz immer möglich. Setzen wir  $U' := \langle v_{m+1}, \dots, v_n \rangle$ , dann gilt  $V = U \oplus U'$  nach Satz **a**.  $\square$

## 6.3 Lineare Abbildungen

### 6.3.1 Homomorphismen

Homomorphismen sind „strukturhaltende Abbildungen“.

**Definition a.** Es seien  $(G, \circ)$  und  $(H, \bullet)$  zwei Gruppen. Eine Abbildung  $\varphi : G \rightarrow H$  heißt *Gruppen-Homomorphismus*, wenn für alle  $x, y \in G$  gilt:

$$\varphi(x \circ y) = \varphi(x) \bullet \varphi(y).$$

**Definition b.** Seien  $R$  und  $S$  zwei Ringe. Eine Abbildung  $\varphi : R \rightarrow S$  heißt *Ring-Homomorphismus*, wenn gelten:

- (i)  $\varphi$  ist Gruppenhomomorphismus  $(R, +) \rightarrow (S, +)$
- (ii)  $\varphi(xy) = \varphi(x)\varphi(y)$  für alle  $x, y \in R$
- (iii)  $\varphi(1) = 1$

Ein „Körper-Homomorphismus“ ist schlicht ein Ringhomomorphismus zwischen zwei Körpern (jeder Körper ist ein kommutativer Ring).

**Definition c.** Ein Homomorphismus  $\varphi$  zwischen zwei Strukturen heißt *Monomorphismus* wenn er injektiv ist, *Epimorphismus* wenn er surjektiv ist, und *Isomorphismus* wenn er bijektiv ist. Existiert ein Isomorphismus  $\varphi : A \rightarrow B$  dann heißen  $A$  und  $B$  *isomorph*, geschrieben  $A \cong B$ .

**Bemerkung.** Oft untersucht man Strukturen (Gruppen, Ringe, Vektorräume, etc.) nur „bis auf Isomorphie“, d.h. man unterscheidet nicht zwischen isomorphen Strukturen. Dies ist dadurch begründet, dass isomorphe Strukturen im Prinzip durch „Umbenennung der Elemente“ (vermittelt durch einen Isomorphismus) auseinander hervorgehen.

**Beispiel a.**

- (i) Für jede Untergruppe  $U$  von  $(G, \circ)$  ist die Abbildung

$$\varphi : (U, \circ) \rightarrow (G, \circ), \quad x \mapsto x$$

ein Gruppen-Monomorphismus, z.B.  $(\mathbb{Z}, +) \rightarrow (\mathbb{R}, +), x \mapsto x$ .

- (ii)  $\exp : (\mathbb{R}, +) \rightarrow (\mathbb{R}_{>0}, \cdot), x \mapsto \exp(x) = e^x$  ist Gruppen-Isomorphismus ( $e^{x+y} = e^x \cdot e^y$  für alle  $x, y \in \mathbb{R}$ ). Daher gilt  $(\mathbb{R}, +) \cong (\mathbb{R}_{>0}, \cdot)$ .

- (iii) Es sei  $(G, \cdot)$  Gruppe,  $n \in \mathbb{N}$ . Für jedes  $1 \leq i \leq n$  ist

$$\epsilon_i : G \rightarrow G^n, \quad x \mapsto (e, \dots, e, \underbrace{x}_{i\text{-te Position}}, e, \dots, e)$$

ein Gruppen-Monomorphismus und

$$\rho_i : G^n \rightarrow G, \quad (x_1, \dots, x_n) \mapsto x_i$$

ein Gruppen-Epimorphismus. Es gilt stets  $\rho_i \circ \epsilon_i = \text{id}_G$ .

- (iv) Gegeben seien ein Körper  $K$ , eine abelsche Gruppe  $(V, +)$  und eine Abbildung

$$\cdot : K \times V \rightarrow V, (\lambda, v) \mapsto \lambda \cdot v = \lambda v.$$

Das Vektorraumaxiom (V2) besagt, dass für jedes  $\lambda \in K$  die Abbildung

$$f_\lambda : (V, +) \rightarrow (V, +), v \mapsto \lambda v$$

ein Gruppen-Homomorphismus ist.

Das Vektorraumaxiom (V1) besagt, dass für jedes  $v \in V$  die Abbildung

$$g_v : (K, +) \rightarrow (V, +), \lambda \mapsto \lambda v$$

ein Gruppen-Homomorphismus ist.

- (v) Es gibt  $n^n$  Abbildungen  $\mathbb{Z}_n \rightarrow \mathbb{Z}_n$ , aber nur  $n$  Gruppen-Homomorphismen  $(\mathbb{Z}_n, +) \rightarrow (\mathbb{Z}_n, +)$ .

*Übung a.* Für jeden Gruppenhomomorphismus  $\varphi : G \rightarrow H$  gelten:

- (i)  $\varphi(e_G) = e_H$ ,
- (ii)  $\varphi(g^{-1}) = \varphi(g)^{-1}$  für alle  $g \in G$ ,
- (iii)  $\varphi$  ist genau dann injektiv, wenn für alle  $g \in G$  gilt:  $\varphi(g) = e_H \Rightarrow g = e_G$ .

**Beispiel b.**

- (i) Für jedes  $n \in \mathbb{N}$  ist die Abbildung  $\varphi : \mathbb{Z} \rightarrow \mathbb{Z}_n, a \mapsto \bar{a}$  ein Ring-Epimorphismus.
- (ii) Für jedes  $a \in \mathbb{Z}$  mit  $a \neq 1$  ist die Abbildung

$$m_a : \mathbb{Z} \rightarrow \mathbb{Z}, x \mapsto a \cdot x$$

kein Ring-Homomorphismus, da  $m_a(1) = a \neq 1$ .

- (iii)  $\varphi : \mathbb{Z} \rightarrow \mathbb{Z}, a \mapsto a^3$  ist kein Ring-Homomorphismus, weil z.B.  $\varphi(1+1) = 8 \neq 2 = \varphi(1) + \varphi(1)$ .
- (iv)  $\varphi : \mathbb{Z}_3 \rightarrow \mathbb{Z}_3, a \mapsto a^3$  ist ein Ring-Isomorphismus.
- (v)  $\mathbb{F}_p \rightarrow \mathbb{F}_p, x \mapsto x^p$  ist ein Ring-Isomorphismus.
- (vi)  $\mathbb{Q} \rightarrow \mathbb{R}, x \mapsto x$  ist ein Ring-Monomorphismus.
- (vii)  $\mathbb{C} \rightarrow \mathbb{C}, z \mapsto \bar{z}$  ist ein Ring-Isomorphismus (komplexe Konjugation:  $\overline{a+bi} = a-bi$ ).

*Übung* b. Jeder Körperhomomorphismus ist injektiv.

*Beweis.* Es seien  $K, L$  zwei Körper und  $\varphi : K \rightarrow L$  ein Ringhomomorphismus. Angenommen,  $\varphi$  ist nicht injektiv, d.h. es gibt  $x, y \in K$  mit  $\varphi(x) = \varphi(y)$ , obwohl  $x \neq y$ . Dann ist  $z := x - y \neq 0$  und  $\varphi(z) = \varphi(x) - \varphi(y) = 0$  und

$$1 = \varphi(1) = \varphi(z \cdot z^{-1}) = \varphi(z) \cdot \varphi(z^{-1}) = 0 \cdot \varphi(z^{-1}) = 0$$

Widerspruch. □

### 6.3.2 Lineare Abbildungen

In diesem Abschnitt sei  $K$  ein Körper.

**Definition.** Es seien  $V, W$  zwei  $K$ -Vektorräume.

- (i) Eine Abbildung  $\varphi : V \rightarrow W$  heißt *lineare Abbildung* oder *Vektorraum-Homomorphismus*, falls für alle  $v, v' \in V$  und alle  $\lambda \in K$  gelten:

$$\varphi(v + v') = \varphi(v) + \varphi(v'), \quad \varphi(\lambda v) = \lambda \varphi(v).$$

Die Menge aller Homomorphismen  $V \rightarrow W$  wird mit  $\text{Hom}(V, W)$  bezeichnet.

- (ii) Ein Vektorraum-Homomorphismus  $\varphi : V \rightarrow V$  heißt *Endomorphismus* von  $V$ . Die Menge  $\text{Hom}(V, V)$  aller Endomorphismen von  $V$  wird mit  $\text{End}(V)$  bezeichnet.

**Bemerkung.**

- (i) Für jeden Vektorraum-Homomorphismus  $\varphi : V \rightarrow W$  gilt  $\varphi(0) = 0$ .
- (ii) Im folgenden meinen wir mit *Homomorphismus* stets einen Vektorraum-Homomorphismus.

**Beispiel.**

- (i) Lineare Abbildungen  $\mathbb{R}^2 \rightarrow \mathbb{R}^2$  sind z.B.: Drehungen um 0, Spiegelungen an Geraden durch 0, Projektionen auf Koordinatenachsen. Nicht linear sind dagegen Translationen (Verschiebungen).
- (ii) Betrachte die Abbildungen  $\varphi_1, \dots, \varphi_4 : \mathbb{R}^3 \rightarrow \mathbb{R}^2$  mit

$$\begin{aligned}\varphi_1 : \begin{pmatrix} a \\ b \\ c \end{pmatrix} &\mapsto \begin{pmatrix} a \\ b \end{pmatrix}, & \varphi_2 : \begin{pmatrix} a \\ b \\ c \end{pmatrix} &\mapsto \begin{pmatrix} 1+a \\ b \end{pmatrix}, \\ \varphi_3 : \begin{pmatrix} a \\ b \\ c \end{pmatrix} &\mapsto \begin{pmatrix} a+c \\ b \end{pmatrix}, & \varphi_4 : \begin{pmatrix} a \\ b \\ c \end{pmatrix} &\mapsto \begin{pmatrix} a \\ b^2 \end{pmatrix}.\end{aligned}$$

Davon sind  $\varphi_1, \varphi_3$  linear,  $\varphi_2, \varphi_4$  dagegen nicht. Die Abbildung  $\varphi_1$  ist gerade die Projektion des  $\mathbb{R}^3$  auf die  $e_1$ - $e_2$ -Ebene.

- (iii) Die Transpositionsabbildung

$$(\cdot)^t : K^{m \times n} \rightarrow K^{n \times m}, A \mapsto A^t$$

ist linear. Spezialfälle sind:

$$\begin{aligned}(\cdot)^t &: K^{1 \times n} \rightarrow K^n, z \mapsto z^t \\ (\cdot)^t &: K^m \rightarrow K^{1 \times m}, s \mapsto s^t\end{aligned}$$

- (iv) Für jede Matrix  $A \in K^{m \times n}$  ist die Abbildung  $\varphi_A : K^n \rightarrow K^m, x \mapsto Ax$  linear.
- (v) Betrachte die  $\mathbb{R}$ -Vektorräume  $\text{Abb}(\mathbb{R}, \mathbb{R})$  (reelle Funktionen) und  $\mathbb{R}$ . Für jedes fest gewählte  $a \in \mathbb{R}$  ist die Abbildung

$$\epsilon_a : \text{Abb}(\mathbb{R}, \mathbb{R}) \rightarrow \mathbb{R}, \quad f \mapsto f(a)$$

linear und wird *Einsetzungshomomorphismus* genannt.

- (vi) Betrachte die  $\mathbb{R}$ -Vektorräume  $C^\infty(\mathbb{R})$  (beliebig oft stetig differenzierbare reelle Funktionen) und  $\mathbb{R}$ . Die *Ableitungsabbildung*

$$\text{diff} : C^\infty(\mathbb{R}) \rightarrow C^\infty(\mathbb{R}), \quad f \mapsto f'$$

ist linear. (Das ist eine bekannte Ableitungsregel aus der Analysis.)

*Übung.* Es sei  $f : V \rightarrow W$ .

- (i)  $f$  ist genau dann linear, wenn für alle  $n \in \mathbb{N}, \lambda_i \in K, v_i \in V$  gilt:  
 $f(\sum_{i=1}^n \lambda_i v_i) = \sum_{i=1}^n \lambda_i f(v_i)$ .

Sei nun  $f$  linear.

- (ii) Ist  $g : W \rightarrow U$  linear, so auch  $g \circ f : V \rightarrow U$ .  
 (iii) Ist  $f$  bijektiv (Isomorphismus), so ist auch  $f^{-1}$  Isomorphismus.

### 6.3.3 Kern und Bild

Hier seien  $K$  ein Körper und  $V, W$  Vektorräume über  $K$ .

**Definition.** Es sei  $\varphi \in \text{Hom}(V, W)$ .

- (i)  $\text{Kern } \varphi := \{v \in V \mid \varphi(v) = 0\}$  heißt *Kern von  $\varphi$* .  
 (ii)  $\text{Bild } \varphi := \varphi(V) = \{\varphi(v) \mid v \in V\}$  heißt *Bild von  $\varphi$* .

**Bemerkung.** Für jedes  $\varphi \in \text{Hom}(V, W)$  gilt:

- (i)  $\text{Kern } \varphi \leq V$ .  
 (ii)  $\text{Bild } \varphi \leq W$ .  
 (iii)  $\varphi$  injektiv  $\Leftrightarrow \text{Kern } \varphi = \{0\}$ .  
 (iv)  $\varphi$  surjektiv  $\Leftrightarrow \text{Bild } \varphi = W$ .  
 (v) Für jedes  $v \in V$  und  $w = \varphi(v)$  gilt:

$$\varphi^{-1}(\{w\}) = v + \text{Kern } \varphi.$$

*Beweis.* (siehe Vorlesung)

□

**Beispiel.**

- (i) Es sei  $A \in K^{m \times n}$ . Dann ist  $\text{Kern } \varphi_A = \mathbb{L}(A, 0)$ .  
 (ii) Der Kern der Projektion  $\varphi_1$  aus Beispiel 6.3.2(ii) ist die von  $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$  erzeugte Gerade. Das Bild ist ganz  $\mathbb{R}^2$ .

- (iii) Der Kern des Einsetzungshomomorphismus  $\epsilon_a$  aus Beispiel 6.3.2(v) besteht genau aus denjenigen reellen Funktionen, die bei  $a$  eine Nullstelle haben. Das Bild ist ganz  $\mathbb{R}$ , weil es zu jedem  $x \in \mathbb{R}$  eine reelle Funktion gibt, die an der Stelle  $a$  den Wert  $x$  annimmt (z.B. die konstante Funktion mit dem Wert  $x$ ).
- (iv) Der Kern der Ableitungsabbildung aus Beispiel (6.3.2)(vi) besteht genau aus den konstanten reellen Funktionen. (Der Kern ist 1-dimensional!) Das Bild ist ganz  $C^\infty(\mathbb{R})$ , weil jede stetige reelle Funktion eine Stammfunktion hat.

Zur Berechnung von Kern und Bild mit Hilfe von Koordinaten siehe Abschnitt (6.4.7).

*Übung.* Für jedes  $\varphi \in \text{Hom}(V, W)$  gilt:

- (i)  $U \leq V \Rightarrow \varphi(U) \leq W$ .
- (ii)  $U \leq W \Rightarrow \varphi^{-1}(U) \leq V$ .
- (iii)  $M \subseteq V \Rightarrow \varphi(\langle M \rangle) = \langle \varphi(M) \rangle$ .
- (iv)  $M \subseteq V$  linear unabhängig,  $\varphi$  injektiv  $\Rightarrow \varphi(M)$  linear unabhängig.
- (v)  $U \leq V \Rightarrow \dim \varphi(U) \leq \dim U$ .
- (vi)  $U \leq V$  und  $\varphi$  injektiv  $\Rightarrow \dim \varphi(U) = \dim U$ .
- (vii)  $U \leq \text{Bild } \varphi \Rightarrow \dim \varphi^{-1}(U) \geq \dim U$ .

*Beweis.* Es gilt  $\varphi(\sum_{i=1}^n \lambda_i v_i) = \sum_{i=1}^n \lambda_i \varphi(v_i)$ . Daraus lassen sich (i)–(iv) folgern.

Z.B. (iv): Es seien  $M$  linear unabhängig und  $\varphi$  injektiv. Sei  $\sum_{i=1}^n \lambda_i \varphi(v_i) = 0_W$  eine lineare Abhängigkeit in  $\varphi(M)$ , d.h.  $\lambda_i \in K$ ,  $v_i \in M$  und  $\varphi(v_i)$  paarweise verschieden. Dann sind auch die  $v_i$  paarweise verschieden (das ist klar für jede Abbildung  $\varphi$ ) und  $\varphi(\sum_{i=1}^n \lambda_i v_i) = 0_W$ . Da  $\varphi$  injektiv ist, folgt  $\sum_{i=1}^n \lambda_i v_i = 0$ . Da  $M$  linear unabhängig ist, sind alle  $\lambda_i = 0$ . Damit ist gezeigt, dass  $\varphi(M)$  linear unabhängig ist.

Z.B. (v): Wähle Basis  $B$  von  $U$ . Nach (iii) ist  $\varphi(U) = \langle \varphi(B) \rangle$ , also  $\dim \varphi(U) \leq |\varphi(B)| \leq |B| = \dim U$ .  $\square$

### 6.3.4 Existenz linearer Abbildungen

Es sei  $K$  ein Körper und  $V, W$  zwei  $K$ -Vektorräume.

**Frage.** Es seien Vektoren  $v_1, \dots, v_n \in V$  und  $w_1, \dots, w_n \in W$  gegeben. Gibt es eine lineare Abbildung  $V \rightarrow W$  mit  $v_i \mapsto w_i$ ?

**Satz a.** *Es sei  $B$  eine Basis von  $V$ . Zu jedem  $v \in B$  sei ein  $w_v \in W$  gegeben. Dann existiert eine eindeutige lineare Abbildung  $\varphi : V \rightarrow W$  mit  $\varphi(v) = w_v$  für alle  $v \in B$ .*

*Beweis.* Eindeutigkeit: Sei  $\varphi : V \rightarrow W$  linear mit  $\varphi(v) = w_v$  für alle  $v \in B$ . Wir zeigen, dass damit  $\varphi$  schon auf ganz  $V$  festgelegt ist. Sei dazu  $v \in V$  beliebig, etwa  $v = \sum_{i=1}^n \lambda_i v_i$  mit  $v_1, \dots, v_n \in B$  paarweise verschieden und  $\lambda_i \in K \setminus \{0\}$ . Nach Satz 6.4.1 sind die  $\lambda_i$  eindeutig bestimmt. Aus der Linearität von  $\varphi$  folgt:

$$\varphi(v) = \sum_{i=1}^n \lambda_i \varphi(v_i) = \sum_{i=1}^n \lambda_i w_{v_i}. \quad (6.2)$$

Wegen der Eindeutigkeit der  $\lambda_i$  ist damit auch  $\varphi(v)$  eindeutig festgelegt.

Existenz: Benutzen wir die Gleichung (6.2) als Definition für ein  $\varphi : V \rightarrow W$ , so bleibt nur noch zu prüfen, dass das so definierte  $\varphi$  auch linear ist (Übung).  $\square$

**Bemerkung.**

- (i) Man liest Satz a auch so: Jede Abbildung  $f : B \rightarrow W$  lässt sich eindeutig zu einer linearen Abbildung  $\varphi : V \rightarrow W$  fortsetzen.
- (ii) Die lineare Unabhängigkeit von  $B$  wird in Satz a nur bei der Eindeutigkeit gebraucht; die Tatsache, dass  $B$  Erzeugendensystem dagegen nur bei der Existenz. Somit gilt: Ist  $B$  linear unabhängig (statt Basis), so gibt es in Satz a mindestens ein solches  $\varphi$  (statt genau ein); ist  $B$  Erzeugendensystem, so gibt es in Satz a höchstens ein solches  $\varphi$ .

**Beispiel.**  $V = \mathbb{R}^3, B = (e_1, e_2, e_3)$ . Wähle  $w_1 = e_2, w_2 = -e_1, w_3 = 0$ . Nach Satz a gibt es genau einen Endomorphismus von  $\mathbb{R}^3$  mit  $e_i \mapsto w_i$  für  $i = 1, 2, 3$ . Frage: Was für eine Abbildung ist das?

Antwort: Sei  $\varphi$  die Projektion auf die  $e_1$ - $e_2$ -Ebene gefolgt von einer  $90^\circ$ -Drehung um die  $e_3$ -Achse. Wir wissen aus vorherigen Beispielen, dass  $\varphi$  linear ist. Da  $\varphi$  auch  $\varphi(e_i) = w_i$  für  $i = 1, 2, 3$  erfüllt, muss es der gesuchte Endomorphismus sein (weil er eindeutig ist). Die Abbildungsvorschrift lautet:

$$\varphi : \begin{pmatrix} a \\ b \\ c \end{pmatrix} \mapsto \begin{pmatrix} -b \\ a \\ 0 \end{pmatrix}.$$



**Satz b.** Sei  $W$  ein beliebiger nicht-trivialer  $K$ -Vektorraum. Eine Teilmenge  $B \subseteq V$  ist genau dann eine Basis von  $V$ , wenn sich jede Abbildung  $f : B \rightarrow W$  eindeutig zu einer linearen Abbildung  $\varphi : V \rightarrow W$  fortsetzen lässt.

*Beweis.* Eine Richtung wurde bereits in Satz a gezeigt. Wir setzen nun voraus, jede Abbildung  $f : B \rightarrow W$  lasse sich eindeutig zu einer linearen Abbildung  $\varphi : V \rightarrow W$  fortsetzen, und folgern, dass  $B$  Basis ist.

Annahme:  $B$  ist linear abhängig, etwa  $\sum_{i=1}^n \lambda_i v_i = 0_V$  mit  $n \in \mathbb{N}$ ,  $v_1, \dots, v_n \in B$  paarweise verschieden und  $\lambda_i \in K \setminus \{0\}$ . Nach Voraussetzung gibt es  $\varphi \in \text{Hom}(V, W)$  mit  $\varphi(v_1) \neq 0_W$  und  $\varphi(v_i) = 0_W$  für  $i = 2, \dots, n$ . Dann folgt  $0_W = \varphi(0_V) = \varphi(\sum \lambda_i v_i) = \sum \lambda_i \varphi(v_i) = \lambda_1 \varphi(v_1) \neq 0_W$ . Da dies ein Widerspruch ist, ist die Annahme falsch, also  $B$  linear unabhängig.

Annahme:  $B$  ist keine Basis. Ergänze die linear unabhängige Menge  $B$  zu einer Basis  $B'$  von  $V$  und wähle ein  $v \in B' \setminus B$ . Dann existieren nach Satz a mindestens zwei verschiedene Fortsetzungen  $\varphi$  von  $f$ , nämlich ein  $\varphi$  mit  $\varphi(v) = 0_W$  und eins mit  $\varphi(v) \neq 0_W$ . Da dies ein Widerspruch zur Voraussetzung ist, ist die Annahme falsch, also  $B$  eine Basis von  $V$ .  $\square$

*Übung a.* Es sei  $\varphi : V \rightarrow W$  linear und surjektiv (Epimorphismus). Man zeige, dass eine lineare Abbildung  $\psi : W \rightarrow V$  existiert mit  $\varphi \circ \psi = \text{id}_W$ .

Hinweis: Übung 6.3.2 und Satz a.

*Übung b.* Es sei  $\varphi : V \rightarrow W$  linear und injektiv (Monomorphismus). Man zeige, dass eine lineare Abbildung  $\psi : W \rightarrow V$  existiert mit  $\psi \circ \varphi = \text{id}_V$ .

Hinweis: Übung 6.3.2 und Satz a.

### 6.3.5 Monomorphismen und Epimorphismen

Es seien  $V, W$  zwei beliebige  $K$ -Vektorräume und  $\varphi \in \text{Hom}(V, W)$ . Wir erinnern daran, dass eine injektive lineare Abbildung auch Monomorphismus, und eine surjektive lineare Abbildung auch Epimorphismus genannt wird.

**Satz.** Es sei  $B$  eine beliebige Basis von  $V$ .

(i) Folgende Aussagen sind äquivalent:

1.  $\varphi$  ist Monomorphismus.
2. Für jede Teilmenge  $M \subseteq V$  gilt:  
 $M$  linear unabhängig  $\Rightarrow \varphi(M)$  linear unabhängig.
3.  $\varphi(B)$  ist linear unabhängig und  $\varphi|_B$  ist injektiv.

In diesem Fall gilt  $\dim V \leq \dim W$ .

(ii) Folgende Aussagen sind äquivalent:

1.  $\varphi$  ist Epimorphismus.
2. Für jede Teilmenge  $M \subseteq V$  gilt:  
 $M$  erzeugt  $V \Rightarrow \varphi(M)$  erzeugt  $W$ .
3.  $\varphi(B)$  ist Erzeugendensystem von  $W$ .

In diesem Fall gilt  $\dim V \geq \dim W$ .

(iii) Folgende Aussagen sind äquivalent:

1.  $\varphi$  ist Isomorphismus.
2. Für jede Teilmenge  $M \subseteq V$  gilt:  
 $M$  Basis von  $V \Rightarrow \varphi(M)$  Basis von  $W$ .
3.  $\varphi(B)$  ist Basis von  $W$  und  $\varphi|_B$  ist injektiv.

In diesem Fall gilt  $\dim V = \dim W$ .

*Beweis.*

- (i) 1.  $\Rightarrow$  2. wurde bereits in Übung 6.3.3iv gezeigt.  
 2.  $\Rightarrow$  1. Sei  $\varphi(v) = 0_W, v \in V$ . Wir zeigen  $v = 0_V$ . Dann ist nachgewiesen:  $\varphi$  ist injektiv, d.h. Monomorphismus. Wende 2. mit  $M = \{v\}$  an. Da  $\varphi(M) = \{0_W\}$  linear abhängig ist, ist auch  $M = \{v\}$  linear abhängig, d.h.  $v = 0_V$ .  
 3.  $\Rightarrow$  1. Sei  $\varphi(v) = 0_W, v \in V$ . Wir zeigen  $v = 0_V$ ; dann ist  $\varphi$  Monomorphismus. Schreibe  $v = \sum_{i=1}^n \lambda_i v_i$  mit  $n \in \mathbb{N}, v_1, \dots, v_n \in B$  paarweise verschieden und  $\lambda_i \in K$ . Da  $\varphi|_B$  injektiv ist, sind auch  $\varphi(v_1), \dots, \varphi(v_n)$  paarweise verschieden. Da  $\varphi(B)$  linear unabhängig ist und  $0_W = \varphi(v) = \sum_{i=1}^n \lambda_i \varphi(v_i)$ , sind alle  $\lambda_i = 0$ , also  $v = 0_V$ .  
 1.  $\wedge$  2.  $\Rightarrow$  3. ist trivial.  
 Nach Übung 6.3.3vi ist für injektives  $\varphi$ :  $\dim V = \dim \varphi(V) \leq \dim W$ .

(ii) als Übung.

- (iii) 1.  $\Leftrightarrow$  3. und 1.  $\Rightarrow$  2. folgen aus i) und ii).  
 2.  $\Rightarrow$  1. folgt aus i) und ii) zusammen mit Basisergänzung und Basisauswahl (Details als Übung).  
 Die Dimensionsgleichung folgt auch aus i) und ii).

□

*Übung.* Sei  $V$  endlich-dimensional.

- (i)  $\varphi|_B$  injektiv  $\Leftrightarrow |\varphi(B)| = |B|$ .
- (ii)  $\varphi$  injektiv  $\Leftrightarrow$  für alle  $U \leq V$  gilt  $\dim \varphi(U) = \dim U \Leftrightarrow \dim \varphi(V) = \dim V$ .

**Beispiel.**

- (i) Im  $\mathbb{R}^2$  sind Drehungen um 0 und Spiegelungen an Ursprungsgeraden stets Isomorphismen.
- (ii) Die Projektion  $\varphi_1$  aus Beispiel (6.3.2)(ii) ist Epimorphismus, aber kein Monomorphismus.
- (iii) Die Codierungsabbildung  $\varphi_G : \mathbb{F}_2^k \rightarrow \mathbb{F}_2^n$  (vgl. 6.6) ist ein Monomorphismus, aber kein Epimorphismus.
- (iv) Mit  $\varphi$  Isomorphismus, ist auch  $\varphi^{-1}$  linear (Übung) und somit Isomorphismus.
- (v) Die Transpositionsabbildung

$$^t : K^{m \times n} \rightarrow K^{n \times m}, A \mapsto A^t$$

ist ein Isomorphismus.

- (vi) Für  $\varphi \in \text{Hom}(V, W)$  gilt:  $\varphi$  Monomorphismus  $\Leftrightarrow \varphi : V \rightarrow \varphi(V)$  Isomorphismus.

**6.3.6 Endlich-dimensionale Vektorräume**

Es seien  $V, W$  zwei endlich-dimensionale  $K$ -Vektorräume und  $\varphi \in \text{Hom}(V, W)$ .

**Definition.**

$$\begin{array}{ll} \text{Rg } \varphi := \dim(\text{Bild } \varphi) & \text{Rang von } \varphi. \\ \text{Def } \varphi := \dim(\text{Kern } \varphi) & \text{Defekt von } \varphi. \end{array}$$

**Beispiel.** Die Projektion  $\varphi_1$  aus Beispiel (6.3.2)(ii) hat den Rang 2, da das Bild gleich  $\mathbb{R}^2$  ist, und den Defekt 1, da der Kern die Gerade  $\langle e_3 \rangle$  ist.

**Bemerkung.**

- (i) Es gilt  $\text{Rg } \varphi \leq \dim W$ , und man hat genau dann  $\text{Rg } \varphi = \dim W$ , wenn  $\varphi$  ein Epimorphismus ist.
- (ii) Es gilt  $\text{Def } \varphi \leq \dim V$ , und man hat genau dann  $\text{Def } \varphi = 0$ , wenn  $\varphi$  ein Monomorphismus ist.
- (iii) Es gilt  $\text{Rg } \varphi \leq \dim V$ , und man hat genau dann  $\text{Rg } \varphi = \dim V$ , wenn  $\varphi$  ein Monomorphismus ist.

Im Spezialfall  $\dim V = \dim W = n$  ergibt sich:

$$\operatorname{Rg} \varphi = n \Leftrightarrow \operatorname{Def} \varphi = 0 \Leftrightarrow \varphi \text{ Isomorphismus,}$$

d.h. die Begriffe Monomorphismus, Epimorphismus und Isomorphismus sind in diesem Fall äquivalent.

*Beweis.* (i) und (ii) sind unmittelbar klar aus den Definitionen der Begriffe. (iii): Es sei  $B$  eine beliebige Basis von  $V$ . Da  $\varphi(V)$  von  $\varphi(B)$  erzeugt wird, gilt  $\dim \varphi(V) \leq |\varphi(B)|$  mit Gleichheit genau dann, wenn  $\varphi(B)$  Basis von  $\varphi(V)$  ist. Weiter ist  $|\varphi(B)| \leq |B| = \dim V$  mit Gleichheit genau dann, wenn  $\varphi|_B$  injektiv ist. Also  $\dim \varphi(V) \leq |\varphi(B)| \leq |B| = \dim V$  mit

$$\begin{aligned} \dim \varphi(V) = \dim V &\Leftrightarrow \dim \varphi(V) = |\varphi(B)| \wedge |\varphi(B)| = |B| \\ &\Leftrightarrow \varphi(B) \text{ Basis von } \varphi(V) \wedge \varphi|_B \text{ injektiv.} \end{aligned}$$

Nach Satz 6.3.5 ist letzteres äquivalent zu  $\varphi : V \rightarrow \varphi(V)$  Isomorphismus, bzw. zu  $\varphi : V \rightarrow W$  Monomorphismus.

Im Fall  $\dim V = \dim W = n$  ergibt sich:

$$\operatorname{Def} \varphi = 0 \stackrel{(ii)}{\Leftrightarrow} \varphi \text{ Monomorphismus} \stackrel{(iii)}{\Leftrightarrow} \operatorname{Rg} \varphi = n \stackrel{(i)}{\Leftrightarrow} \varphi \text{ Epimorphismus.}$$

□

**Satz.**  $V \cong W \Leftrightarrow \dim V = \dim W$ .

*Beweis.*  $\Rightarrow$  wurde schon in Satz 6.3.5 (iii) gezeigt. Sei nun  $\dim V = \dim W = n$ . Wähle beliebige Basen  $\{v_1, \dots, v_n\}$  von  $V$  und  $\{w_1, \dots, w_n\}$  von  $W$ . Nach Satz 6.3.4 gibt es eine lineare Abbildung  $\varphi : V \rightarrow W$  mit  $v_i \mapsto w_i$  für  $i = 1, \dots, n$ . Nach Satz 6.3.5 (iii) ist dieses  $\varphi$  ein Isomorphismus. □

**Folgerung.** Ist  $n = \dim V$ , dann ist  $V \cong K^n$ .

**Beispiel.**

(i)  $\mathbb{C} \cong \mathbb{R}^2$  (als  $\mathbb{R}$ -Vektorraum).

(ii)  $\mathbb{R}^{2 \times 2} \cong \mathbb{R}^4$ .

(iii)  $\mathbb{R}^{n \times m} \cong \mathbb{R}^{nm} \cong \mathbb{R}^{m \times n}$ .

(iv)  $K^{1 \times n} \cong K^n$ .

### 6.3.7 $\text{Hom}(V, W)$ als Vektorraum

Es seien  $V$  und  $W$  zwei  $K$ -Vektorräume.

**Satz.** Die Menge  $\text{Hom}(V, W)$  wird selbst zu einem  $K$ -Vektorraum, wenn man Addition und skalare Multiplikation punktweise definiert, d.h.

$$\begin{aligned}(\varphi + \psi)(x) &:= \varphi(x) + \psi(x), \\ (\lambda\varphi)(x) &:= \lambda\varphi(x).\end{aligned}$$

*Beweis.* Zunächst ist  $\text{Abb}(V, W)$  mit punktweiser Addition und punktweiser skalarer Multiplikation ein  $K$ -Vektorraum. Es bleibt zu zeigen, dass die Teilmenge  $\text{Hom}(V, W) \subseteq \text{Abb}(V, W)$  die Unterraumbedingungen erfüllt. Man rechnet leicht nach, dass mit  $\varphi, \psi$  linear auch  $\varphi + \psi$  und  $\lambda\varphi$  wieder linear sind (Übung).  $\square$

**Frage.** Welche Dimension hat  $\text{Hom}(V, W)$ ? Wie sieht eine Basis aus?

### 6.3.8 Der Endomorphismenring

**Bemerkung.** Die Komposition von linearen Abbildungen ist wieder linear. Genauer, sind  $U, V, W$  drei  $K$ -Vektorräume,  $\varphi \in \text{Hom}(V, W)$  und  $\psi \in \text{Hom}(U, V)$ , so ist  $\varphi \circ \psi \in \text{Hom}(U, W)$ .

*Beweis.* Übung.  $\square$

Wir betrachten nun  $\text{End}(V) = \text{Hom}(V, V)$  für einen  $K$ -Vektorraum  $V$ . Gemäß Satz 6.3.7 ist  $\text{End}(V)$  selbst ein  $K$ -Vektorraum mit der punktweisen Addition und skalaren Multiplikation. Im Falle von  $\text{End}(V)$  haben wir zusätzlich noch die Verknüpfung  $\circ$  (Komposition).

**Satz.**  $(\text{End}(V), +, \circ)$  ist ein Ring. Die neutralen Elemente sind die Nullabbildung  $(0 : V \rightarrow V, v \mapsto 0)$  und die Identität  $(1 = \text{id}_V)$ .

**Definition.**  $(\text{End}(V), +, \circ)$  wird der *Endomorphismenring von  $V$*  genannt. Die Einheitengruppe  $\text{Aut}(V) := (\text{End}(V)^\times, \circ)$  heißt *Automorphismengruppe*, deren Elemente *Automorphismen*.

## 6.4 Lineare Abbildungen und Matrizen

In diesem Abschnitt sei  $K$  ein Körper und  $V$  ein  $K$ -Vektorraum.

### 6.4.1 Koordinaten

**Definition a.** Es sei  $B = \{v_1, \dots, v_n\}$  eine Basis von  $V$  mit  $|B| = n$ . Ist  $v \in V$  und  $v = \sum_{i=1}^n \lambda_i v_i$  die eindeutige Darstellung von  $v$  als Linearkombination der Basisvektoren (siehe Satz 6.2.5b), dann werden die (eindeutigen) Koeffizienten  $\lambda_1, \dots, \lambda_n$  die *Koordinaten* von  $v$  bzgl.  $B$  genannt.

**Definition b.** Es seien  $v_1, \dots, v_n \in V$ . Das  $n$ -Tupel  $\mathcal{B} = (v_1, \dots, v_n)$  heißt *geordnete Basis* von  $V$ , wenn  $\{v_1, \dots, v_n\}$  eine Basis von  $V$  ist und  $v_1, \dots, v_n$  paarweise verschieden sind. In diesem Fall definieren wir die *Koordinatenabbildung* bzgl.  $\mathcal{B}$  als

$$\kappa_{\mathcal{B}} : V \rightarrow K^n, v \mapsto \begin{pmatrix} \lambda_1 \\ \vdots \\ \lambda_n \end{pmatrix},$$

wobei  $v = \sum_{i=1}^n \lambda_i v_i$ . Das Bild  $\kappa_{\mathcal{B}}(v) \in K^n$  heißt der *Koordinatenvektor* von  $v$  bzgl.  $\mathcal{B}$ .

**Bemerkung a.**

- (i) Jeder endlich-dimensionale Vektorraum besitzt eine geordnete Basis, also auch eine Koordinatenabbildung.
- (ii) Koordinatenabbildungen sind stets bijektiv.
- (iii) Es sei  $n = \dim V$ . Jeder Isomorphismus  $V \rightarrow K^n$  ist die Koordinatenabbildung  $\kappa_{\mathcal{B}}$  bzgl. einer geeigneten geordneten Basis  $\mathcal{B}$  von  $V$ .

*Beweis.* (ii) Wegen  $\dim V = \dim K^n$  ist  $V \cong K^n$  (siehe Folgerung 6.3.6). Es sei  $\varphi : V \rightarrow K^n$  ein beliebiger Isomorphismus. Dann ist auch  $\varphi^{-1} : K^n \rightarrow V$  ein Isomorphismus (Beispiel 6.3.5(iv)). Da  $\{e_1, \dots, e_n\}$  eine Basis von  $K^n$  ist und  $\varphi^{-1}$  ein Isomorphismus, ist nach Satz 6.3.5(iii)  $\{\varphi^{-1}(e_1), \dots, \varphi^{-1}(e_n)\}$  eine Basis von  $V$ . Für die geordnete Basis  $\mathcal{B} := (\varphi^{-1}(e_1), \dots, \varphi^{-1}(e_n))$  und die Koordinatenabbildung  $\kappa_{\mathcal{B}}$  gilt dann  $\kappa_{\mathcal{B}} : \varphi^{-1}(e_i) \mapsto e_i$  für  $i = 1, \dots, n$ . Die linearen Abbildungen  $\kappa_{\mathcal{B}}$  und  $\varphi$  stimmen also auf den Basiselementen von  $\mathcal{B}$  überein. Laut Satz 6.3.4 a handelt es sich daher um dieselben Abbildungen, d.h.  $\kappa_{\mathcal{B}} = \varphi$ .  $\square$

Auf Grund der in Teil (iii) der Bemerkung erwähnten Tatsache lassen sich sämtliche Rechnungen in endlich-dimensionalen Vektorräumen via Koordinatenabbildungen auf Rechnungen in  $K^n$  zurückführen.

**Beispiel.**

- (i) Betrachte  $\mathbb{C}$  als  $\mathbb{R}$ -Vektorraum mit der geordneten Basis  $\mathcal{B} = (1, i)$ . Die Koordinatenabbildung zu  $\mathcal{B}$  lautet:

$$\kappa_B : \mathbb{C} \rightarrow \mathbb{R}^2, a + bi \mapsto \begin{pmatrix} a \\ b \end{pmatrix}.$$

(Daher kommt die Interpretation von  $\mathbb{C}$  als „komplexe Ebene“.)

- (ii) Betrachte den  $\mathbb{R}$ -Vektorraum  $\mathbb{R}^{2 \times 2}$  mit der geordneten Basis

$$B = (E_{11}, E_{12}, E_{21}, E_{22}) = \left( \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right).$$

Die Koordinatenabbildung zu  $\mathcal{B}$  lautet:

$$\kappa_B : \mathbb{R}^{2 \times 2} \rightarrow \mathbb{R}^4, \begin{pmatrix} a & b \\ c & d \end{pmatrix} \mapsto \begin{pmatrix} a \\ b \\ c \\ d \end{pmatrix}.$$

- (iii) Betrachte den  $\mathbb{R}$ -Vektorraum  $\mathbb{R}^2$  mit der geordneten Basis

$$B = \left( \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \right).$$

Die Koordinatenabbildung zu  $\mathcal{B}$  lautet:

$$\kappa_B : \mathbb{R}^2 \rightarrow \mathbb{R}^2, \begin{pmatrix} a \\ b \end{pmatrix} \mapsto \begin{pmatrix} \frac{a+b}{2} \\ \frac{a-b}{2} \end{pmatrix}.$$

*Übung: Man prüfe das nach!*

## 6.4.2 Die Abbildungsmatrix

In diesem Abschnitt werden alle Vektorräume  $V$  als endlich-dimensional und nicht-trivial vorausgesetzt, also  $0 < \dim < \infty$ . Mit Basen sind immer geordnete Basen gemeint und wir bezeichnen diese mit  $\mathcal{A}, \mathcal{B}, \mathcal{C}, \dots$

Ist  $\mathcal{C} = (v_1, \dots, v_n)$  eine geordnete Basis von  $V$  und  $\varphi \in \text{GL}(V)$ , so bezeichne  $\varphi(\mathcal{C})$  das Tupel  $(\varphi(v_1), \dots, \varphi(v_n))$ . Man beachte, dass die Zuordnung

$$\text{GL}(V) \rightarrow \{\text{geordnete Basen von } V\}, \quad \varphi \mapsto \varphi(\mathcal{C})$$

eine Bijektion ist (das folgt aus den Sätzen 6.3.4 a und 6.3.5).

**Definition.** Es seien  $V$  und  $W$  zwei  $K$ -Vektorräume mit  $\dim V = n$  und  $\dim W = m$  und mit den geordneten Basen  $\mathcal{B} = (v_1, \dots, v_n)$  von  $V$  und  $\mathcal{C}$  von  $W$ .

Die *Abbildungsmatrix* von  $\varphi$  bzgl.  $\mathcal{B}$  und  $\mathcal{C}$  ist definiert als

$$M(\varphi) := M_{\mathcal{C}}^{\mathcal{B}}(\varphi) := (s_1, \dots, s_n), \quad s_i := \kappa_{\mathcal{C}}(\varphi(v_i)).$$

Ist  $V = W$  und  $\mathcal{B} = \mathcal{C}$ , so sagen wir kurz Abbildungsmatrix bzgl.  $\mathcal{B}$  und schreiben auch  $M_{\mathcal{B}}(\varphi)$  statt  $M_{\mathcal{B}}^{\mathcal{B}}(\varphi)$ . Wir verwenden die Schreibweise  $M(\varphi)$ , wenn  $\mathcal{B}$  und  $\mathcal{C}$  fest gewählt sind.

**Bemerkung a.** Die Abbildung

$$M_{\mathcal{C}}^{\mathcal{B}} : \text{Hom}(V, W) \rightarrow K^{m \times n}, \quad \varphi \mapsto M_{\mathcal{C}}^{\mathcal{B}}(\varphi)$$

ist ein  $K$ -Vektorraum-Isomorphismus. Insbesondere hat  $\text{Hom}(V, W)$  die Dimension  $mn$ .

*Beweis.* Nach Satz 6.3.4 a wird  $\varphi$  ein-eindeutig durch das Tupel  $\varphi(\mathcal{B}) := (\varphi(v_1), \dots, \varphi(v_n))$  beschrieben. Da  $\kappa_{\mathcal{C}}$  bijektiv ist, wird  $\varphi(\mathcal{B})$  wiederum ein-eindeutig durch  $M_{\mathcal{C}}^{\mathcal{B}}(\varphi)$  beschrieben. Die Abbildung  $M_{\mathcal{C}}^{\mathcal{B}}$  ist somit eine Bijektion und, wie man leicht nachrechnet, auch linear:  $M_{\mathcal{C}}^{\mathcal{B}}(\varphi + \psi) = M_{\mathcal{C}}^{\mathcal{B}}(\varphi) + M_{\mathcal{C}}^{\mathcal{B}}(\psi)$  und  $M_{\mathcal{C}}^{\mathcal{B}}(a\varphi) = aM_{\mathcal{C}}^{\mathcal{B}}(\varphi)$  für alle  $\varphi, \psi \in \text{Hom}(V, W)$  und alle  $a \in K$ .  $\square$

**Satz.** Die Abbildungsmatrix  $M_{\mathcal{C}}^{\mathcal{B}}(\varphi)$  ist die eindeutige Matrix  $M \in K^{m \times n}$  mit der Eigenschaft

$$\kappa_{\mathcal{C}}(\varphi(v)) = M \cdot \kappa_{\mathcal{B}}(v) \text{ für alle } v \in V, \quad (6.3)$$

**Bemerkung b.** Eine äquivalente Formulierung der Bedingung (6.3) ist:

$$\kappa_{\mathcal{C}} \circ \varphi = \varphi_M \circ \kappa_{\mathcal{B}} \quad \text{bzw.} \quad \varphi = \kappa_{\mathcal{C}}^{-1} \circ \varphi_M \circ \kappa_{\mathcal{B}}.$$

(Diagramm siehe Vorlesung.)

*Beweis des Satzes.* Laut Satz 6.3.4 a ist die Bedingung (6.3) äquivalent dazu, dass  $\kappa_{\mathcal{C}}(\varphi(v_i)) = M \cdot \kappa_{\mathcal{B}}(v_i)$  für alle Basisvektoren  $v_i$  aus  $\mathcal{B}$  gilt. Für  $1 \leq i \leq n$  ist aber  $M \cdot \kappa_{\mathcal{B}}(v_i) = Me_i = i\text{-te Spalte von } M$ , und, per Definition,  $\kappa_{\mathcal{C}}(\varphi(v_i)) = s_i$ . Somit ist (6.3) äquivalent zu  $M = M_{\mathcal{C}}^{\mathcal{B}}(\varphi)$ .  $\square$

Im folgenden Beispiel wird dargestellt, wie die Abbildungsmatrix das effektive Rechnen mit der linearen Abbildung mittels Koordinaten erlaubt.

**Beispiel.** Es bezeichne  $\mathcal{E}$  die Standardbasis von  $K^n$ , d.h.  $\mathcal{E} = (e_1, \dots, e_n)$ . Dann ist die zugehörige Koordinatenabbildung  $\kappa_{\mathcal{E}} : K^n \rightarrow K^n$  die Identität.



- (i) Wie lautet die Abbildungsmatrix  $S_0$  der Spiegelung von  $\mathbb{R}^2$  an der  $e_1$ -Achse bzgl.  $\mathcal{E}$ ? Wir haben  $\mathcal{E} = (e_1, e_2) = \left( \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix} \right)$ . Die Spiegelung an der  $e_1$ -Achse ist eine lineare Abbildung  $\sigma_0 : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  mit den Bildern der Basisvektoren  $\sigma_0(e_1) = e_1$  und  $\sigma_0(e_2) = -e_2$ . Die Koordinatenvektoren dieser Bilder lauten

$$\kappa_{\mathcal{E}}(\sigma_0(e_1)) = \sigma_0(e_1) = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \kappa_{\mathcal{E}}(\sigma_0(e_2)) = \sigma_0(e_2) = \begin{pmatrix} 0 \\ -1 \end{pmatrix},$$

$$\text{also ist } S_0 = M_{\mathcal{E}}^{\mathcal{E}}(\sigma_0) = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

$$\text{Probe: } \sigma_0\left(\begin{pmatrix} a \\ b \end{pmatrix}\right) = \begin{pmatrix} a \\ -b \end{pmatrix} = S_0 \cdot \begin{pmatrix} a \\ b \end{pmatrix}.$$

- (ii) Wie lautet die Abbildungsmatrix  $R_{\alpha}$  der Drehung von  $\mathbb{R}^2$  um den Winkel  $\alpha$  gegen den Uhrzeigersinn? Diese Drehung ist eine lineare Abbildung  $\rho_{\alpha} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  mit den Bildern der Basisvektoren  $\rho_{\alpha}(e_1) = \cos \alpha \cdot e_1 + \sin \alpha \cdot e_2$  und  $\rho_{\alpha}(e_2) = -\sin \alpha \cdot e_1 + \cos \alpha \cdot e_2$ . Die Koordinatenvektoren dieser Bilder lauten

$$\kappa_{\mathcal{E}}(\rho_{\alpha}(e_1)) = \rho_{\alpha}(e_1) = \begin{pmatrix} \cos \alpha \\ \sin \alpha \end{pmatrix}, \quad \kappa_{\mathcal{E}}(\rho_{\alpha}(e_2)) = \rho_{\alpha}(e_2) = \begin{pmatrix} -\sin \alpha \\ \cos \alpha \end{pmatrix},$$

$$\text{also ist } R_{\alpha} = M_{\mathcal{E}}^{\mathcal{E}}(\rho_{\alpha}) = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}.$$

$$\text{Probe: } \rho_{\alpha}\left(\begin{pmatrix} a \\ b \end{pmatrix}\right) = \begin{pmatrix} a \cos \alpha - b \sin \alpha \\ a \sin \alpha + b \cos \alpha \end{pmatrix} = R_{\alpha} \cdot \begin{pmatrix} a \\ b \end{pmatrix}.$$

- (iii) Es sei  $A \in K^{m \times n}$ . Für  $\varphi_A : K^n \rightarrow K^m, x \mapsto A \cdot x$  ist

$$M_{\mathcal{E}}^{\mathcal{E}}(\varphi_A) = A.$$

- (iv) Betrachte  $\varphi = \text{id}_V$ . Für jede Basis  $\mathcal{B}$  von  $V$  ist

$$M_{\mathcal{B}}^{\mathcal{B}}(\text{id}_V) = E_n,$$

wobei  $n = \dim V$  ist.

- (v) Es sei

$$V = \text{Pol}_n(\mathbb{R}) \leq \text{Abb}(\mathbb{R}, \mathbb{R})$$

der Vektorraum der Polynomfunktionen

$$\{f : \mathbb{R} \rightarrow \mathbb{R}, f(x) = a_n x^n + \cdots + a_1 x^1 + a_0 \mid a_0, \dots, a_n \in \mathbb{R}\}.$$

$V$  ist  $\mathbb{R}$ -Vektorraum der Dimension  $n+1$  mit Basis  $\mathcal{B} = (p_0, p_1, \dots, p_n)$ , mit  $p_i : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto x^i$  für  $0 \leq i \leq n$ . Die *Ableitungsabbildung*

$$\text{diff} : V \rightarrow V, f \mapsto f'$$

ist linear (vgl. Beispiel 6.3.2(vi)). Wir bestimmen die Abbildungsmatrix von  $\text{diff}$  bezüglich  $\mathcal{B}$ . Die Bilder der Basisvektoren sind

$$\text{diff}(p_i) = \begin{cases} 0 & \text{falls } i = 0, \\ ip_{i-1} & \text{falls } i > 0. \end{cases}$$

Übersetzt in Koordinaten bzgl.  $\mathcal{B}$  bedeutet das

$$M_{\text{diff}} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ & 0 & 2 & & \\ \vdots & \vdots & 0 & & \\ & & \vdots & \ddots & n \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix} \in \mathbb{R}^{(n+1) \times (n+1)}.$$

Probe: Für  $f = a_n p_n + \cdots + a_1 p_1 + a_0 p_0 \in \text{Pol}_n(\mathbb{R})$  ist  $f' = na_n p_{n-1} + \cdots + 2a_2 p_1 + a_1 p_0$ . Also wie gewünscht

$$\kappa_{\mathcal{B}}(f') = \begin{pmatrix} a_1 \\ 2a_2 \\ \vdots \\ na_n \\ 0 \end{pmatrix} = M_{\text{diff}} \cdot \begin{pmatrix} a_0 \\ \vdots \\ a_n \end{pmatrix} = M_{\text{diff}} \cdot \kappa_{\mathcal{B}}(f).$$

### 6.4.3 Der Produktsatz

Es seien  $\varphi : V \rightarrow W$  und  $\psi : W \rightarrow U$  zwei lineare Abbildungen zwischen  $K$ -Vektorräumen. Von  $U, V, W$  seien die geordneten Basen  $\mathcal{A}, \mathcal{B}, \mathcal{C}$  fest gewählt.

**Satz.**  $M_{\mathcal{A}}^{\mathcal{B}}(\psi \circ \varphi) = M_{\mathcal{A}}^{\mathcal{C}}(\psi) \cdot M_{\mathcal{C}}^{\mathcal{B}}(\varphi)$ .

*Beweis.* Man beachte, dass stets  $\varphi_{\mathcal{A}} \circ \varphi_{\mathcal{B}} = \varphi_{\mathcal{AB}}$  ist, sofern das Matrixprodukt  $\mathcal{AB}$  existiert. Damit ergibt sich

$$\begin{aligned} \psi \circ \varphi &= \kappa_{\mathcal{A}}^{-1} \circ \psi_{M(\psi)} \circ \kappa_{\mathcal{C}} \circ \kappa_{\mathcal{C}}^{-1} \circ \varphi_{M(\varphi)} \circ \kappa_{\mathcal{B}} \\ &= \kappa_{\mathcal{A}}^{-1} \circ (\varphi_{M(\psi)} \circ \varphi_{M(\varphi)}) \circ \kappa_{\mathcal{B}} \\ &= \kappa_{\mathcal{A}}^{-1} \circ \varphi_{M(\psi) \cdot M(\varphi)} \circ \kappa_{\mathcal{B}}. \end{aligned}$$

Aus Satz 6.4.2 folgt  $M(\psi \circ \varphi) = M(\psi) \cdot M(\varphi)$ . □

**Folgerung a.** Falls  $\dim V = \dim W$  (d.h.  $M_{\mathcal{C}}^{\mathcal{B}}(\varphi)$  quadratisch) ist, so ist  $\varphi$  genau dann ein Isomorphismus, wenn  $M_{\mathcal{C}}^{\mathcal{B}}(\varphi)$  invertierbar ist. In diesem Fall gilt  $M_{\mathcal{B}}^{\mathcal{C}}(\varphi^{-1}) = (M_{\mathcal{C}}^{\mathcal{B}}(\varphi))^{-1}$ .

*Beweis.* Es sei  $\dim V = \dim W = n$ , d.h.  $M := M_{\mathcal{C}}^{\mathcal{B}}(\varphi)$  eine  $n \times n$ -Matrix. Wegen  $\varphi = \kappa_{\mathcal{C}}^{-1} \circ \varphi_M \circ \kappa_{\mathcal{B}}$  (siehe Bemerkung 6.4.2 b), ist  $\varphi$  genau dann ein Isomorphismus, wenn  $\varphi_M$  einer ist, denn  $\kappa_{\mathcal{B}}$  und  $\kappa_{\mathcal{C}}$  sind Isomorphismen.

Es sei zuerst  $M$  invertierbar. Dann ist  $\varphi_M$  bijektiv mit der Umkehrabbildung  $\varphi_M^{-1}$ . Also ist  $\varphi_M$  und damit auch  $\varphi$  ein Isomorphismus. Sei nun  $\varphi$  ein Isomorphismus. Aus dem Produktsatz und Beispiel (6.4.2)(iv) folgt  $M_{\mathcal{C}}^{\mathcal{B}}(\varphi) \cdot M_{\mathcal{B}}^{\mathcal{C}}(\varphi^{-1}) = M_{\mathcal{C}}^{\mathcal{C}}(\text{id}_W) = E_n$ , und  $M_{\mathcal{B}}^{\mathcal{C}}(\varphi^{-1}) \cdot M_{\mathcal{C}}^{\mathcal{B}}(\varphi) = M_{\mathcal{B}}^{\mathcal{B}}(\text{id}_V) = E_n$ . Also ist  $M = M_{\mathcal{C}}^{\mathcal{B}}(\varphi)$  invertierbar und es gilt  $M_{\mathcal{B}}^{\mathcal{C}}(\varphi^{-1}) = (M_{\mathcal{C}}^{\mathcal{B}}(\varphi))^{-1}$ .  $\square$

**Folgerung b.** Der Vektorraum-Isomorphismus

$$M_{\mathcal{B}}^{\mathcal{B}} : \text{End}(V) \rightarrow K^{n \times n}, \quad \varphi \mapsto M_{\mathcal{B}}^{\mathcal{B}}(\varphi)$$

ist auch Ringisomorphismus. Die Einschränkung

$$M_{\mathcal{B}}^{\mathcal{B}} : \text{GL}(V) \rightarrow \text{GL}_n(K), \quad \varphi \mapsto M_{\mathcal{B}}^{\mathcal{B}}(\varphi)$$

ist Gruppenisomorphismus.

*Beweis.* Um zu sehen, dass  $M$  ein Ringhomomorphismus ist, sind  $M_{\text{id}} = E_n$ ,  $M(\varphi + \psi) = M(\varphi) + M(\psi)$ , und  $M(\varphi \circ \psi) = M(\varphi) \cdot M(\psi)$  zu prüfen. Alle drei Bedingungen wurden bereits gezeigt (die letzte durch den Produktsatz).

Die Einschränkung eines Ringisomorphismus auf die Einheitsgruppen ist stets Gruppenisomorphismus.  $\square$

**Beispiel a.** Es sei  $\sigma_{\alpha} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  die Spiegelung an der um den Winkel  $\alpha$  gegen den Uhrzeigersinn gedrehten  $e_1$ -Achse (Zeichnung siehe Vorlesung). Wie lautet die Abbildungsmatrix  $S_{\alpha}$  von  $\sigma_{\alpha}$  bzgl.  $\mathcal{E}$ ? Idee: Wir schreiben  $\sigma_{\alpha}$  als die Komposition  $\rho_{\alpha} \circ \sigma_0 \circ \rho_{-\alpha}$  und verwenden den Produktsatz. Die Abbildungsmatrizen von  $\rho_{\alpha}$  und  $\sigma_0$  sind bereits aus Beispiel (6.4.2) bekannt. Es ergibt sich:

$$\begin{aligned} S_{\alpha} &= M(\sigma_{\alpha}) = M(\rho_{\alpha}) \cdot M(\sigma_0) \cdot M(\rho_{-\alpha}) = R_{\alpha} \cdot S_0 \cdot R_{-\alpha} \\ &= \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix} \\ &= \begin{pmatrix} \cos^2 \alpha - \sin^2 \alpha & 2 \sin \alpha \cos \alpha \\ 2 \sin \alpha \cos \alpha & \sin^2 \alpha - \cos^2 \alpha \end{pmatrix} \end{aligned}$$

Für  $\alpha = 30$  erhalten wir wegen  $\sin 30 = \sin \frac{\pi}{6} = \frac{1}{2}$  und  $\cos 30 = \cos \frac{\pi}{6} = \frac{\sqrt{3}}{2}$  die Matrix  $S_{30} = \frac{1}{2} \begin{pmatrix} 1 & \sqrt{3} \\ \sqrt{3} & -1 \end{pmatrix}$ .

**Beispiel b.** Da  $\rho_\alpha$  bijektiv ist, ist  $R_\alpha = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}$  invertierbar. Die Umkehrabbildung von  $\rho_\alpha$  ist  $\rho_{-\alpha}$ , folglich

$$R_\alpha^{-1} = M(\rho_\alpha)^{-1} = M(\rho_\alpha^{-1}) = M(\rho_{-\alpha}) = R_{-\alpha} = \begin{pmatrix} \cos \alpha & \sin \alpha \\ -\sin \alpha & \cos \alpha \end{pmatrix}.$$

Man beachte, dass die Gleichung  $R_\alpha \cdot R_{-\alpha} = E_2$  äquivalent ist zu:  $\sin^2 \alpha + \cos^2 \alpha = 1$ .

*Übung.* Warum gilt die Gleichung  $\sigma_\alpha = \sigma_0 \circ \rho_{2\alpha}$ ? Man berechne daraus  $S_\alpha$  mit Hilfe des Produktsatzes. Wie ist das Ergebnis im Vergleich zu Beispiel **a** zu interpretieren?

#### 6.4.4 Die Basiswechselmatrix

Es seien  $\mathcal{B}, \mathcal{B}'$  zwei geordnete Basen von  $V$ ,  $0 < \dim V = n < \infty$ .

**Definition.** Die Matrix  $M_{\mathcal{B}}^{\mathcal{B}'}(\text{id}_V)$  wird *Basiswechselmatrix* oder *Basistransformationsmatrix* genannt, geschrieben  $T_{\mathcal{B}}^{\mathcal{B}'}$ .

**Bemerkung.**

- (i) In den Spalten von  $T_{\mathcal{B}}^{\mathcal{B}'}$  stehen die Basisvektoren aus  $\mathcal{B}'$ , geschrieben in Koordinaten bzgl.  $\mathcal{B}$ .
- (ii)  $T_{\mathcal{B}}^{\mathcal{B}'}$  ist die eindeutige Matrix  $T \in K^{n \times n}$  mit der Eigenschaft  $\kappa_{\mathcal{B}} = \varphi_T \circ \kappa_{\mathcal{B}'}$ , d.h.

$$\kappa_{\mathcal{B}}(v) = T \cdot \kappa_{\mathcal{B}'}(v) \quad \text{für alle } v \in V.$$

- (iii) Es gilt  $T_{\mathcal{B}'}^{\mathcal{B}} = (T_{\mathcal{B}}^{\mathcal{B}'})^{-1}$ .

- (iv) Für jedes  $\varphi \in \text{GL}(V)$  gilt  $M_{\mathcal{B}}^{\mathcal{B}'}(\varphi) = T_{\varphi(\mathcal{B})}^{\mathcal{B}'}$ .

*Beweis.* Die Aussagen ergeben sich aus der Definition und den Bemerkungen 6.4.2, sowie aus Folgerung 6.4.3 a:

$$T_{\mathcal{B}}^{\mathcal{B}'} = M_{\mathcal{B}}^{\mathcal{B}'}(\text{id}_V) = (M_{\mathcal{B}'}^{\mathcal{B}}((\text{id}_V)^{-1}))^{-1} = (M_{\mathcal{B}'}^{\mathcal{B}}(\text{id}_V))^{-1} = (T_{\mathcal{B}'}^{\mathcal{B}})^{-1}.$$

□

**Folgerung.** Für festes  $\mathcal{B}$  ist folgende Abbildung eine Bijektion:

$$\{\text{geordnete Basen von } V\} \rightarrow \text{GL}_n(K), \quad \mathcal{B}' \mapsto T_{\mathcal{B}'}^{\mathcal{B}}.$$

*Beweis.* Nach Teil (iv) der Bemerkung sind die Zuordnungen  $\mathrm{GL}(V) \rightarrow \mathrm{GL}_n(K), \varphi \mapsto M_{\mathcal{B}}^{\mathcal{B}}(\varphi)$  und  $\varphi \mapsto T_{\varphi(\mathcal{B})}^{\mathcal{B}}$  identisch. Die Aussage folgt aus der Tatsache, dass sowohl  $\mathrm{GL}(V) \rightarrow \mathrm{GL}_n(K), \varphi \mapsto M_{\mathcal{B}}^{\mathcal{B}}(\varphi)$  als auch  $\mathrm{GL}(V) \rightarrow \{\text{geordnete Basen von } V\}, \varphi \mapsto \varphi(\mathcal{B})$  bijektiv sind. (Diagramm siehe Vorlesung.)  $\square$

**Beispiel.** Es sei  $V = \mathbb{R}^2$ . Wir betrachten die Basen  $\mathcal{B} = \mathcal{E} = (e_1, e_2)$  und  $\mathcal{B}' = (v_1, v_2) = \left(\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \end{pmatrix}\right)$ . Dann gilt:

$$T := T_{\mathcal{B}}^{\mathcal{B}'} = \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix}, \quad T_{\mathcal{B}'}^{\mathcal{B}} = T^{-1} = \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix}^{-1} = \dots = \frac{1}{5} \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}.$$

Wir bestimmen die Koordinaten von  $v := \begin{pmatrix} -1 \\ 3 \end{pmatrix}$  bzgl.  $\mathcal{B}'$  mittels Basiswechselmatrix. Es gilt  $\kappa_{\mathcal{B}}(v) = \begin{pmatrix} -1 \\ 3 \end{pmatrix}$ , also Teil (ii) der Bemerkung

$$\kappa_{\mathcal{B}'}(v) = T_{\mathcal{B}'}^{\mathcal{B}} \cdot \kappa_{\mathcal{B}}(v) = \frac{1}{5} \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix} \begin{pmatrix} -1 \\ 3 \end{pmatrix} = \frac{1}{5} \begin{pmatrix} 5 \\ 5 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

Per Definition ist  $\kappa_{\mathcal{B}'}(v) = \begin{pmatrix} a \\ b \end{pmatrix}$  gerade der Lösungsvektor der Gleichung  $av_1 + bv_2 = v$ . Das liefert uns die Probe:  $1 \cdot v_1 + 1 \cdot v_2 = v$ .  $\checkmark$

*Anmerkung:* Natürlich kann  $\kappa_{\mathcal{B}'}(v)$  als Lösung des linearen Gleichungssystems  $av_1 + bv_2 = v$  auch direkt (ohne Verwendung von Basiswechselmatrizen) bestimmt werden. In Matrixschreibweise lautet dieses Gleichungssystem  $T \cdot \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} -1 \\ 3 \end{pmatrix}$ , d.h. die Basiswechselmatrix taucht auch hier wieder auf.

### 6.4.5 Der Basiswechselsatz

Es seien  $\mathcal{B}, \mathcal{B}'$  geordnete Basen von  $V$  und  $\mathcal{C}, \mathcal{C}'$  geordnete Basen von  $W$ .

**Satz.** Für jede lineare Abbildung  $\varphi : V \rightarrow W$  gilt:

$$M_{\mathcal{C}'}^{\mathcal{B}'}(\varphi) = T_{\mathcal{C}'}^{\mathcal{C}} \cdot M_{\mathcal{C}}^{\mathcal{B}}(\varphi) \cdot T_{\mathcal{B}}^{\mathcal{B}'}.$$

*Beweis.* Nach Definition (6.4.4) und Satz 6.4.3 gilt

$$T_{\mathcal{C}'}^{\mathcal{C}} \cdot M_{\mathcal{C}}^{\mathcal{B}}(\varphi) \cdot T_{\mathcal{B}}^{\mathcal{B}'} = M_{\mathcal{C}'}^{\mathcal{C}}(\mathrm{id}_W) \cdot M_{\mathcal{C}}^{\mathcal{B}}(\varphi) \cdot M_{\mathcal{B}}^{\mathcal{B}'}(\mathrm{id}_V) = M_{\mathcal{C}'}^{\mathcal{B}'}(\mathrm{id}_W \circ \varphi \circ \mathrm{id}_V).$$

$\square$

**Beispiel.** Wir betrachten die lineare Abbildung  $\varphi_A : \mathbb{R}^2 \rightarrow \mathbb{R}^2$  mit  $A = \begin{pmatrix} -3/5 & 4/5 \\ 4/5 & 3/5 \end{pmatrix}$ . Wir haben  $M_{\mathcal{E}}^{\mathcal{E}}(\varphi_A) = A$  (vgl. Beispiel (6.4.2)(iii)). *Frage:* Gibt es eine Basis  $\mathcal{B}$  von  $V$  so, dass  $M_{\mathcal{B}}^{\mathcal{B}}(\varphi_A)$  besonders „einfach“ wird? Dies ist in der Tat der Fall für die Basis  $\mathcal{B} = (v_1, v_2) = \left( \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} -2 \\ 1 \end{pmatrix} \right)$ , wie folgende Rechnung zeigt. Aus Beispiel (6.4.4) wissen wir

$$T_{\mathcal{B}}^{\mathcal{E}} = \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix}, \quad T_{\mathcal{E}}^{\mathcal{B}} = \frac{1}{5} \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix}.$$

Nach obigem Satz gilt folglich

$$\begin{aligned} M_{\mathcal{B}}^{\mathcal{B}}(\varphi_A) &= T_{\mathcal{E}}^{\mathcal{B}} \cdot M_{\mathcal{E}}^{\mathcal{E}}(\varphi_A) \cdot T_{\mathcal{B}}^{\mathcal{E}} = \frac{1}{5} \begin{pmatrix} 1 & 2 \\ -2 & 1 \end{pmatrix} \frac{1}{5} \begin{pmatrix} -3 & 4 \\ 4 & 3 \end{pmatrix} \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix} \\ &= \frac{1}{25} \begin{pmatrix} 5 & 10 \\ 10 & -5 \end{pmatrix} \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix} = \frac{1}{5} \begin{pmatrix} 1 & 2 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} 1 & -2 \\ 2 & 1 \end{pmatrix} = \frac{1}{5} \begin{pmatrix} 5 & 0 \\ 0 & -5 \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}. \end{aligned}$$

Die Abbildung  $\varphi_A$  ist also eine Spiegelung an der  $v_1$ -Achse! (Bild siehe Vorlesung.) Wie man die „richtige“ Basis  $\mathcal{B}$  systematisch auffindet, werden wir erst später sehen, im Kapitel über Eigenvektoren.

### 6.4.6 Zeilen- und Spaltenraum und der Rang von Matrizen

In diesem Abschnitt sei  $K$  ein Körper und  $A \in K^{m \times n}$ .

**Definition a.** von  $A$ . Die folgenden vier Unterräume heißen die *Fundamentallräume* von  $A$ :

$\text{SR}(A) := \langle s_1, \dots, s_n \rangle$	Spaltenraum
$\text{ZR}(A) := \langle z_1, \dots, z_m \rangle$	Zeilenraum
$\mathbb{L}(A, 0) := \{x \in K^n \mid Ax = 0\}$	Rechts-Nullraum
$\mathbb{L}^0(A) := \{y \in K^{1 \times m} \mid yA = 0\}$	Links-Nullraum

**Bemerkung a.** Es gelten

$$\begin{aligned} \text{SR}(A) &\leq K^m, & \mathbb{L}(A, 0) &\leq K^n, \\ \text{ZR}(A) &\leq K^{1 \times n}, & \mathbb{L}^0(A) &\leq K^{1 \times m}. \end{aligned}$$

Weiter lässt sich jeder Unterraum  $U \leq K^m$  als Spaltenraum einer  $m \times l$ -Matrix über  $K$  schreiben, wobei  $l = \dim_K U$ . Ebenso lässt sich jeder Unterraum  $U \leq K^{1 \times n}$  als Zeilenraum einer  $l \times n$ -Matrix über  $K$  schreiben, wobei  $l = \dim_K U$ .

*Beweis.* Man wählt eine Basis von  $U$  aus und trägt die Basisvektoren in die Spalten (bzw. Zeilen) von  $A$  ein.  $\square$

**Bemerkung b.** Wie üblich sei  $\varphi_A : K^n \rightarrow K^m$ ,  $x \mapsto Ax$ . Dann ist  $\text{SR}(A) = \text{Bild } \varphi_A$ .

*Beweis.* Dies ist eine unmittelbare Folgerung aus der Definition der Matrixmultiplikation: Wie oben seien  $s_1, \dots, s_n$  die Spalten von  $A$  und es sei

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in K^n.$$

Dann ist

$$Ax = \sum_{j=1}^n x_j s_j \in \text{SR}(A),$$

und jedes Element aus  $\text{SR}(A)$  ist von dieser Form.  $\square$

**Definition b.** Es sei  $\varphi_A$  wie in obiger Bemerkung definiert. Mit *Rang von  $A$* , geschrieben  $\text{Rg } A$ , bezeichnen wir den Rang von  $\varphi_A$ , d.h.

$$\text{Rg } A := \text{Rg } \varphi_A = \dim \text{Bild } \varphi_A,$$

wobei die letzte Gleichheit die Definition von  $\text{Rg } \varphi_A$  ist. Mit obiger Bemerkung ist der Rang von  $A$  gleich der Dimension des Spaltenraums von  $A$ , also  $\text{Rg } A = \dim \text{SR}(A)$ .

Gelegentlich nennen wir  $\dim \text{ZR}(A)$  den *Zeilenrang* von  $A$  und  $\text{Rg } A = \dim \text{SR}(A)$  auch den *Spaltenrang* von  $A$ .

**Bemerkung.** Nach Übung 6.2.5b ist der Zeilenrang (Spaltenrang) gleich der Maximalzahl linear unabhängiger Zeilen (Spalten) von  $A$ .

**Beispiel a.**

- (i) Es sei  $A \in K^{m \times n}$ . Dann ist  $\text{Kern } \varphi_A = \mathbb{L}(A, 0)$  und  $\text{Bild } \varphi_A = \text{SR}(A)$ . Die Bemerkung liefert bereits bekannte Aussagen:

- $\mathbb{L}(A, 0) \leq K^n$ ,  $\text{SR}(A) \leq K^m$ .

- $\varphi_A$  injektiv  $\Leftrightarrow Ax = 0$  nur trivial lösbar.
- $\varphi_A$  surjektiv  $\Leftrightarrow \text{SR}(A) = K^m$ .
- $\mathbb{L}(A, b) = s + \mathbb{L}(A, 0)$  für jedes  $s \in \mathbb{L}(A, b)$ .

(ii) Es sei  $A \in K^{m \times n}$ . Für das Bild unter der Transpositionsabbildung gilt:

$$\text{SR}(A)^t = \text{ZR}(A^t), \quad \text{ZR}(A)^t = \text{SR}(A^t).$$

$$\mathbb{L}(A, 0)^t = \mathbb{L}^0(A^t), \quad \mathbb{L}^0(A)^t = \mathbb{L}(A^t, 0).$$

(In der Tat:  $y \in \mathbb{L}^0(A) \Leftrightarrow yA = 0 \Leftrightarrow (yA)^t = 0 \Leftrightarrow A^t y^t = 0 \Leftrightarrow y^t \in \mathbb{L}(A^t, 0)$ .)

(iii) Der Kern der Projektion  $\varphi_1$  aus Beispiel (6.3.2)(ii) ist die von  $\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$  erzeugte Gerade. Das Bild ist ganz  $\mathbb{R}^2$ .

**Frage.** Zu den vier Fundamentalräumen stellen sich folgende Fragen: Welche Zusammenhänge bestehen zwischen ihnen? Wie bestimmt man Basis und Dimension? Wie hängen die Dimensionen mit dem Rang zusammen? Wie testet man für ein gegebenes Element aus  $K^m, K^n, K^{1 \times n}$  bzw.  $K^{1 \times m}$  ob es in dem entsprechenden Fundamentalraum liegt? Lässt sich jeder Unterraum von  $K^n$  als Nullraum einer Matrix schreiben?

Nach Definition ist  $\mathbb{L}(A, 0)$  die Lösungsmenge des homogenen LGS mit Koeffizientenmatrix  $A$ . Eine Basis und die Dimension von  $\mathbb{L}(A, 0)$  werden wir weiter unten in Abschnitt 6.5.1 berechnen. Der Links-Nullraum von  $A$  kann aus  $\mathbb{L}(A^t, 0)$  bestimmt werden (siehe Beispiel a(ii)). Die Tests lauten:  $b \in \text{SR}(A) \Leftrightarrow Ax = b$  lösbar,  $c \in \text{ZR}(A) \Leftrightarrow xA = c$  lösbar, wobei die letzte Gleichung durch Transponieren auf ein LGS mit Koeffizientenmatrix  $A^t$  und rechter Seite  $c^t$  zurückgeführt werden kann.

**Beispiel.** Es seien  $A$  in Zeilenstufenform,  $r = \text{Rg } A$  und  $1 \leq k_1 < \dots < k_r \leq n$  die Stufenindizes von  $A$ . Da die Zeilen ungleich 0 von  $A$  linear unabhängig sind (Teil (i) von Übung 6.2.3a), und  $\text{ZR}(A)$  erzeugen, gilt  $\dim \text{ZR}(A) = r$ . Alle Spalten von  $A$  haben die ihre Einträge ungleich 0 innerhalb der ersten  $r$  Komponenten; also gilt  $\text{SR}(A) \leq \langle e_1, \dots, e_r \rangle$ , und damit  $\dim \text{SR}(A) \leq \dim \langle e_1, \dots, e_r \rangle = r$  (Folgerung 6.2.6). Seien nun wieder  $s_1, \dots, s_n$  die Spalten von  $A$ . Da das Tupel  $(s_{k_1}, \dots, s_{k_r})$  linear unabhängig ist (Teil (ii) von Übung 6.2.3a), gilt  $\dim \text{SR}(A) \geq r$ . Zusammen also  $\dim \text{ZR}(A) = \dim \text{SR}(A) = \text{Rg } A$ .



In untenstehendem Lemma wird gezeigt, dass elementare Zeilentransformationen an einer Matrix folgende Dinge nicht ändern: den Nullraum, den Zeilenraum und den Rang.

**Lemma.** Falls  $A' \in K^{m \times n}$  aus  $A$  durch eine Folge elementarer Zeilentransformationen hervorgeht, wir schreiben  $A \rightsquigarrow A'$ , dann gelten:

$$(i) \mathbb{L}(A, 0) = \mathbb{L}(A', 0),$$

(ii) Sind  $s_1, \dots, s_n$  die Spalten von  $A$ ,  $s'_1, \dots, s'_n$  die Spalten von  $A'$ , und  $1 \leq i_1, \dots, i_l \leq n$ , dann gilt:

$$\{s_{i_1}, \dots, s_{i_l}\} \text{ linear unabhängig} \Leftrightarrow \{s'_{i_1}, \dots, s'_{i_l}\} \text{ linear unabhängig},$$

$$(iii) \dim \text{SR}(A) = \dim \text{SR}(A'),$$

$$(iv) \text{ZR}(A) = \text{ZR}(A'),$$

$$(v) \dim \text{ZR}(A) = \dim \text{ZR}(A').$$

*Beweis.* (i) Ist bekannt aus Satz 3.4.1.

(ii) Durch Herausstreichen von Spalten aus  $A$  und  $A'$  (und aus dem ganzen Prozess  $A \rightsquigarrow A'$ ) können wir oBdA annehmen, dass  $l = n$  und  $i_j = j$  ist für  $j = 1, \dots, n$ . Mit dieser Vereinfachung ist die Behauptung äquivalent zu:  $Ax = 0$  nur trivial lösbar  $\Leftrightarrow A'x = 0$  nur trivial lösbar. Das folgt wiederum aus (i).

(iii) Folgt aus (ii) und der Bemerkung.

(iv) und (v) als Übung. □

*Übung a.* Man mache sich klar, dass elementare Zeilentransformationen an einer Matrix folgende Dinge ändern können: den Spaltenraum, den Raum  $\mathbb{L}^0$ , die lineare (Un)abhängigkeit von Zeilen.

**Folgerung.** Entsteht  $A'$  aus  $A$  durch eine Folge elementarer Zeilen- und Spaltentransformationen, dann ist  $\text{Rg } A = \text{Rg } A'$ .

**Satz.** Es gilt stets:

$$(i) \text{Rg}(A) = \dim \text{SR}(A) = \dim \text{ZR}(A).$$

$$(ii) \text{Rg } A + \dim \mathbb{L}(A, 0) = n.$$

*Beweis.* (i) ist das Beispiel und das Lemma. (ii) wurde in Folgerung 6.5.1 gezeigt. □

*Übung b.* Man zeige, dass  $\text{Rg } A$  genau die Maximalzahl linear unabhängiger Spalten (Zeilen) von  $A$  ist. Weiter folgere man:  $\text{Rg } A = \text{Rg } A^t$ .

*Übung c.* Man zeige die Korrektheit folgender Methode zur Basisauswahl: Sei  $M = \{v_1, \dots, v_l\} \subseteq K^m$ ,  $U = \langle M \rangle$ . Trage  $v_1, \dots, v_l$  in die Spalten einer  $m \times l$ -Matrix ein und bringe diese auf Zeilenstufenform. Seien  $k_1, \dots, k_r$  die Stufenindizes. Dann ist  $\{v_{k_1}, \dots, v_{k_r}\} \subseteq M$  eine Basis von  $U$ .

*Übung d.* Es seien  $A, A'$  beide in Zeilenstufenform und  $A \rightsquigarrow A'$ . Haben  $A$  und  $A'$  dann stets identische Stufenindizes?

*Übung e.* Eine quadratische Matrix  $A \in K^{n \times n}$  ist genau dann invertierbar, wenn  $\text{Rg } A = n$  ist.

### 6.4.7 Berechnung von Kern und Bild

**Bemerkung.** Es seien Basen  $\mathcal{B}$  von  $V$  und  $\mathcal{C}$  von  $W$  fest gewählt mit zugehörigen Koordinatenabbildungen  $\kappa_{\mathcal{B}}$  und  $\kappa_{\mathcal{C}}$ .

- (i)  $\kappa_{\mathcal{B}}(\text{Kern } \varphi) = \mathbb{L}(M(\varphi), 0)$  bzw.  $\text{Kern } \varphi = \kappa_{\mathcal{B}}^{-1}(\mathbb{L}(M(\varphi), 0))$ .
- (ii)  $\kappa_{\mathcal{C}}(\text{Bild } \varphi) = \text{SR}(M(\varphi))$  bzw.  $\text{Bild } \varphi = \kappa_{\mathcal{C}}^{-1}(\text{SR}(M(\varphi)))$ .
- (iii)  $\text{Rg } \varphi = \text{Rg } M(\varphi)$  und  $\text{Def } \varphi = \text{Def } M(\varphi)$ .

*Beweis.* Wir zeigen exemplarisch die Aussagen über Kern  $\varphi$  und Def  $\varphi$ , die Aussagen über Bild  $\varphi$  und Rg  $\varphi$  gehen ähnlich. (i): Die Aussage ist  $v \in \text{Kern } \varphi \Leftrightarrow \kappa_{\mathcal{B}}(v) \in \mathbb{L}(M(\varphi), 0)$ . Das folgt aus der Injektivität von  $\kappa_{\mathcal{C}}$  und der Bedingung (6.3):

$$\varphi(v) = 0 \Leftrightarrow \kappa_{\mathcal{C}}(\varphi(v)) = 0 \Leftrightarrow M(\varphi)\kappa_{\mathcal{B}}(v) = 0.$$

(iii): Als Monomorphismus erhält  $\kappa_{\mathcal{B}}$  die Dimension von Unterräumen (Übung 6.3.5 oder Bemerkung 6.3.6iii). Daher folgt aus (i):

$$\text{Def } \varphi = \dim(\text{Kern } \varphi) = \dim(\mathbb{L}(M(\varphi), 0)) = \text{Def } M(\varphi).$$

□

**Folgerung.** Für  $\varphi \in \text{Hom}(V, W)$  gilt stets:  $\text{Rg } \varphi + \text{Def } \varphi = \dim V$ .

*Beweis.* Wegen Teil (iii) der Bemerkung folgt dies aus der entsprechenden Gleichung für Matrizen aus Satz 6.4.6. □

**Beispiel.** Wir berechnen mittels Koordinaten Kern und Bild der Ableitungsabbildung

$$\text{diff} : \text{Pol}_n(\mathbb{R}) \rightarrow \text{Pol}_n(\mathbb{R}), f \mapsto f'$$

(vergleiche Beispiel 6.4.2.) Die Abbildungsmatrix von  $\text{diff}$  bzgl.  $\mathcal{B}$  wurde in Beispiel (6.4.2) bestimmt und lautet

$$M_{\text{diff}} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ & 0 & 2 & & \\ \vdots & \vdots & 0 & & \\ & & \vdots & \ddots & n \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix} \in \mathbb{R}^{(n+1) \times (n+1)}.$$

Man sieht sofort

$$\mathbb{L}(M(\varphi), 0) = \left\langle \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \right\rangle, \quad \text{und} \quad \text{SR}(M(\varphi)) = \left\langle \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 2 \\ \vdots \\ 0 \end{pmatrix}, \dots, \begin{pmatrix} 0 \\ \vdots \\ n \\ 0 \end{pmatrix} \right\rangle,$$

also  $\mathbb{L}(M(\varphi), 0) = \langle e_1 \rangle$  und  $\text{SR}(M(\varphi)) = \langle e_1, \dots, e_n \rangle$ . Rückübersetzt in Vektoren aus  $\text{Pol}_n(\mathbb{R})$  bedeutet das

$$\begin{aligned} \text{Kern}(\text{diff}) &= \langle \kappa_{\mathcal{B}}^{-1}(e_1) \rangle = \langle 1 \rangle = \text{Pol}_0(\mathbb{R}) \text{ (konstante Funktionen)} \\ \text{Bild}(\text{diff}) &= \langle \kappa_{\mathcal{B}}^{-1}(e_1), \dots, \kappa_{\mathcal{B}}^{-1}(e_n) \rangle = \text{Pol}_{n-1}(\mathbb{R}). \end{aligned}$$

Man sieht  $\text{Rg}(\text{diff}) = n$  und  $\text{Def}(\text{diff}) = 1$ , die Gleichung  $\text{Rg}(\text{diff}) + \text{Def}(\text{diff}) = n + 1 = \dim \text{Pol}_n(\mathbb{R})$  ist also erfüllt.

### 6.4.8 Äquivalenz und Ähnlichkeit von Matrizen

Wir erinnern an den Basiswechselsatz. Es seien  $V$  und  $W$  zwei endlich-dimensionale  $K$ -Vektorräume und  $\varphi : V \rightarrow W$  eine lineare Abbildung. Weiter seien  $\mathcal{B}, \mathcal{B}'$  Basen von  $V$  und  $\mathcal{C}, \mathcal{C}'$  Basen von  $W$ . Mit  $T$  und  $S$  bezeichnen wir die Basiswechselmatrizen  $T := M_{\mathcal{B}'}^{\mathcal{B}}(\text{id}_V)$  bzw.  $S := M_{\mathcal{C}'}^{\mathcal{C}}(\text{id}_W)$ . Dann liefert der Basiswechselsatz 6.4.5

$$M_{\mathcal{C}'}^{\mathcal{B}'}(\varphi) = S \cdot M_{\mathcal{C}}^{\mathcal{B}}(\varphi) \cdot T.$$

Es sei nun  $W = V$  und  $\mathcal{C} = \mathcal{B}$ ,  $\mathcal{C}' = \mathcal{B}'$ . Wir schreiben für  $M_{\mathcal{B}}^{\mathcal{B}}(\varphi)$  kurz  $M_{\mathcal{B}}(\varphi)$ . Dann liefert der Basiswechselsatz

$$M_{\mathcal{B}'}(\varphi) = T^{-1} \cdot M_{\mathcal{B}}(\varphi) \cdot T.$$

Dies Betrachtungen legen folgende Definitionen nahe.

**Definition.**

- (i) Es seien  $A, B \in K^{m \times n}$ . Wir nennen  $A$  und  $B$  *äquivalent*, wenn ein  $S \in \mathrm{GL}_m(K)$  und ein  $T \in \mathrm{GL}_n(K)$  existiert mit  $B = SAT$ .
- (ii) Es seien  $A, B \in K^{n \times n}$ . Wir nennen  $A$  und  $B$  *ähnlich*, wenn ein  $T \in \mathrm{GL}_n(K)$  existiert mit  $B = T^{-1}AT$ .

*Übung a.* Man zeige, dass sowohl Äquivalenz als auch Ähnlichkeit von Matrizen Äquivalenzrelationen sind.

*Übung b.* Man zeige, dass äquivalente bzw. ähnliche Matrizen denselben Rang haben.

Der Basiswechselsatz 6.4.5 und Folgerung 6.4.4 implizieren: Zwei Matrizen aus  $K^{m \times n}$  sind genau dann äquivalent, wenn sie die gleiche lineare Abbildung von  $V$  nach  $W$  beschreiben, nur bezüglich verschiedener Basen. Zwei Matrizen aus  $K^{n \times n}$  sind genau dann ähnlich, wenn sie den gleichen Endomorphismus von  $V$  beschreiben, nur bezüglich verschiedener Basen.

**Folgerung.** Es seien  $m, n \in \mathbb{N}$ . Weiter sei für jedes  $0 \leq r \leq \min(m, n)$  mit  $Q_r \in K^{m \times n}$  die folgende Matrix bezeichnet:

$$Q_r := \left( \begin{array}{c|c} E_r & 0 \\ \hline 0 & 0 \end{array} \right).$$

Dann gelten:

- (i) Jede Matrix  $A \in K^{m \times n}$  ist zu der Matrix  $Q_r$  mit  $r = \mathrm{Rg} A$  äquivalent.
- (ii) Zwei Matrizen  $A, B \in K^{m \times n}$  sind genau dann äquivalent, wenn  $\mathrm{Rg} A = \mathrm{Rg} B$  ist.
- (iii) Die Anzahl der Äquivalenzklassen in  $K^{m \times n}$  ist  $\min(m, n) + 1$ .

*Beweis.* (i) Betrachte  $\varphi_A : K^n \rightarrow K^m$ . Wähle eine Basis

$$\mathcal{B} = (v_1, \dots, v_r, v_{r+1}, \dots, v_n)$$

von  $K^n$ , so dass  $(v_{r+1}, \dots, v_n)$  eine Basis von  $\mathrm{Kern} \varphi_A$  ist. Setze  $w_j := \varphi_A(v_j)$  für  $1 \leq j \leq r$ . Dann ist  $(w_1, \dots, w_r)$  eine Basis von  $\mathrm{Bild} \varphi_A$  (siehe Vorlesung) und kann zu einer Basis

$$\mathcal{C} = (w_1, \dots, w_r, w_{r+1}, \dots, w_m)$$

von  $W$  ergänzt werden. Mit dieser Notation ist  $M_{\mathcal{C}}^{\mathcal{B}}(\varphi_A) = Q_r$ , und wegen  $r = \dim(\text{Bild } \varphi_A) = \text{Rg } A$  folgt die Behauptung.

(ii) Sind  $A$  und  $B$  äquivalent, dann ist  $\text{Rg } A = \text{Rg } B$  nach Übung b. Ist umgekehrt  $\text{Rg } A = r = \text{Rg } B$ , dann ist nach (i) sowohl  $A$  als auch  $B$  zu  $Q_r$  äquivalent.

(iii) Jede Äquivalenzklasse enthält genau eine Matrix  $Q_r$  für  $0 \leq r \leq \min(m, n)$ .  $\square$

Die Frage nach der Ähnlichkeit zweier quadratischer Matrizen vom selben Format ist nicht so einfach zu behandeln, und wird im Rahmen dieser Vorlesung auch nicht vollständig zu beantworten sein.

### 6.4.9 Elementare Transformationen, Matrixmultiplikation und Matrixinversion

Hier stellen wir alternative Zugänge zu einigen Ergebnissen aus sowie Ergänzungen zu Abschnitt 6.4.6 vor. Wir erinnern an die in Abschnitt 3.4.1 eingeführten elementaren Zeilentransformationen einer Matrix. Es sei  $A \in K^{m \times n}$ ,  $c \in K$  und  $1 \leq i \neq j \leq m$ . Die Transformation  $\tau_{ij}$  vertauscht die  $i$ -te und  $j$ -te Zeile von  $A$ , die Transformation  $\alpha_{ij}(c)$  addiert das  $c$ -fache der  $j$ -ten Zeile zur  $i$ -ten Zeile von  $A$  und die Transformation  $\mu_i(c)$  mit  $c \neq 0$  multipliziert die  $i$ -te Zeile von  $A$  mit  $c$ . Diese Transformationen können durch Praemultiplikation mit geeigneten invertierbaren Matrizen aus  $K^{m \times m}$  realisiert werden.

**Definition.** Es sei  $m \in \mathbb{N}$  und  $c \in K$ . Wir definieren Matrizen aus  $K^{m \times m}$  wie folgt:

- (i)  $T_{ij}$  entstehe aus  $E_m$  durch Anwenden von  $\tau_{ij}$ .
- (ii)  $A_{ij}(c)$  entstehe aus  $E_m$  durch Anwenden von  $\alpha_{ij}(c)$ .
- (iii)  $M_i(c)$  für  $c \neq 0$  entstehe aus  $E_m$  durch Anwenden von  $\mu_i(c)$ .

Die so definierten Matrizen heißen *Elementarmatrizen*.

**Bemerkung a.** Mit der Notation aus der Definition gilt:  $T_{ij}, A_{ij}(c), M_i(c) \in \text{GL}_m(K)$  und

- (i)  $T_{ij}^{-1} = T_{ij}$ ,
- (ii)  $A_{ij}(c)^{-1} = A_{ij}(-c)$ ,
- (iii)  $M_i(c)^{-1} = M_i(c^{-1})$ .

*Beweis.* Übung. □

**Bemerkung b.** Die Bezeichnungen seien wie in der Bemerkung **a**, und es sei  $A \in K^{m \times n}$ . Dann gelten:

- (i) Entsteht  $A'$  aus  $A$  durch die Transformation  $\tau_{ij}$ , dann ist  $A' = T_{ij}A$ .
- (ii) Entsteht  $A'$  aus  $A$  durch die Transformation  $\alpha_{ij}(c)$ , dann ist  $A' = A_{ij}(c)A$ .
- (iii) Entsteht  $A'$  aus  $A$  durch die Transformation  $\mu_i(c)$ , dann ist  $A' = M_i(c)A$ .

*Beweis.* Dies ist eine unmittelbare Folgerung aus der Definition der Matrixmultiplikation. □

**Folgerung a.** Es seien  $A, A' \in K^{m \times n}$ . Entsteht  $A'$  aus  $A$  durch eine Folge elementarer Zeilentransformationen, dann existiert  $S \in \text{GL}_m(K)$  mit  $A' = SA$ .

*Beweis.* Nach Voraussetzung und Bemerkung **b** existiert eine Folge von Matrizen  $S_1, S_2, \dots, S_k$ , wobei jede Matrix  $S_j$  eine der Matrizen aus der Definition ist, mit

$$A' = S_k S_{k-1} \cdots S_1 A.$$

Da alle  $S_j$  nach Bemerkung **a** invertierbar sind, ist  $S := S_k S_{k-1} \cdots S_1 \in \text{GL}_m(K)$ , woraus die Behauptung folgt. □

**Beispiel a.**

$$A = \begin{pmatrix} 2 & -1 & 1 \\ -4 & 2 & -1 \end{pmatrix} \xrightarrow{\alpha_{21}(2)} \begin{pmatrix} 2 & -1 & 1 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{\alpha_{12}(-1)} \begin{pmatrix} 2 & -1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = A'$$

Für  $S := A_{12}(-1) \cdot A_{21}(2) = \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} -1 & -1 \\ 2 & 1 \end{pmatrix}$  muss nach der Folgerung gelten:  $S \cdot A = A'$ . (Man prüfe das nach!)

**Bemerkung c.** Es seien  $A \in K^{m \times n}$  und  $S \in \text{GL}_m(K)$ . Dann ist  $\text{SR}(SA) \cong \text{SR}(A)$ .

*Beweis.* Die Abbildung  $\text{SR}(A) \rightarrow \text{SR}(SA)$ ,  $y \mapsto Sy$  ist ein Isomorphismus:

- Die Abbildung ist linear aufgrund der Regeln für Matrix-Arithmetik.

- Surjektivität: Es sei  $z \in \text{SR}(SA)$ , etwa  $z = (SA)x$  für ein  $x \in K^n$ . Dann ist  $z = S(Ax)$  im Bild der Abbildung.
- Injektivität: Ist  $Sy = Sy'$  für  $y, y' \in \text{SR}(A)$ , dann ist  $y = (S^{-1}S)y = S^{-1}(Sy) = S^{-1}(Sy') = (S^{-1}S)y' = y'$ .

□

**Folgerung b.** *Elementare Zeilentransformationen und elementare Spaltentransformationen ändern den Rang einer Matrix nicht.*

*Beweis.* Die erste Aussage ergibt sich aus Folgerung a und Bemerkung c. Die zweite Aussage ergibt sich aus der Tatsache, dass elementare Spaltentransformationen einer Matrix ihren Spaltenraum gleich lassen. □

**Folgerung c.** *Es sei  $A \in K^{m \times n}$ . Durch eine Folge elementarer Zeilen- und Spaltentransformationen kann  $A$  auf die Gestalt  $\left( \begin{array}{c|c} E_r & 0 \\ \hline 0 & 0 \end{array} \right)$  transformiert werden. Dann ist  $r = \text{Rg } A$ .*

Natürlich lassen sich auch elementare **Spaltentransformationen** durch Matrixmultiplikationen realisieren.

**Bemerkung d.** Die Bezeichnungen seien wie in der Bemerkung a, und es sei  $A \in K^{l \times m}$ . Dann gelten:

- Entsteht  $A'$  aus  $A$  durch Vertauschen der  $i$ -ten mit der  $j$ -ten Spalte, dann ist  $A' = AT_{ij}$ .
- Entsteht  $A'$  aus  $A$  durch die Addition des  $c$ -fachen der  $i$ -ten Spalte zur  $j$ -ten, dann ist  $A' = AA_{ij}(c)$ .
- Entsteht  $A'$  aus  $A$  durch Multiplikation der  $i$ -ten Spalte mit  $0 \neq c \in K$ , dann ist  $A' = AM_i(c)$ .

**Folgerung d.** *Es seien  $A, A' \in K^{m \times n}$ . Entsteht  $A'$  aus  $A$  durch eine Folge elementarer Zeilen- und Spaltentransformationen, dann existieren  $S \in \text{GL}_m(K)$  und  $T \in \text{GL}_n(K)$  mit  $A' = SAT$ .*

*Beweis.* Analog zum Beweis von Folgerung a. □

Zum Schluss dieses Abschnitts stellen wir noch einen Algorithmus zum Invertieren von Matrizen vor.

**Bemerkung e.** Es sei  $A \in K^{n \times n}$ . Dann sind äquivalent:

- (i)  $A \in \text{GL}_n(K)$ .
- (ii)  $A$  kann durch elementare Zeilentransformationen in  $E_n$  überführt werden.

*Beweis.* (i)  $\Rightarrow$  (ii): Überführe  $A$  durch eine Folge elementarer Zeilentransformationen in Zeilenstufenform  $A'$ . Wir haben  $\text{Rg } A' = \text{Rg } A = n$  nach Folgerung [b](#) und Übung [6.4.6e](#). Also hat  $A'$  die Stufenindizes  $1, 2, \dots, n$  und lässt sich durch elementare Zeilentransformationen in  $E_n$  überführen.

(ii)  $\Rightarrow$  (i): Wir haben  $n = \text{Rg } E_n = \text{Rg } A$  nach Folgerung [b](#), und damit  $A \in \text{GL}_n(K)$  nach Übung [6.4.6e](#).  $\square$

### Algorithmus.

**Eingabe:**  $A \in K^{n \times n}$ .

**Ausgabe:**  $A^{-1}$ , falls  $A \in \text{GL}_n(K)$ , und **fail**, sonst.

1. Überführe  $C := (A \mid E_n) \in K^{n \times 2n}$  durch elementare Zeilentransformationen in die Matrix  $C'$  in Zeilenstufenform.

Falls  $C'$  eine Zeile mit  $n$  führenden Nullen besitzt, Return **fail**.

2. Überführe  $C'$  durch elementare Zeilentransformationen in die Matrix

$$(E_n \mid B) \in K^{n \times 2n}.$$

Return  $B$ .

*Beweis.* Wir beweisen die Korrektheit dieses Algorithmus. Die Matrix  $C'$  ist von der Form  $C' = (A' \mid B')$ , wobei  $A' \in K^{n \times n}$  in Zeilenstufenform ist und aus  $A$  durch eine Folge elementarer Zeilentransformationen entstanden ist. Ist  $A$  nicht invertierbar, dann ist  $\text{Rg } A' = \text{Rg } A < n$  nach Übung [6.4.6e](#). Dann ist eine Zeile von  $A'$  die Nullzeile und wir geben **fail** zurück.

Ist dagegen  $A$  invertierbar, dann ist  $\text{Rg } A' = \text{Rg } A = n$ , und die Zeilenstufenform  $A'$  hat die Stufenindizes  $1, 2, \dots, n$ . Die Matrix  $C'$  lässt sich also durch elementare Zeilentransformationen auf die Form  $(E_n \mid B)$  bringen. Nach Folgerung [a](#) existiert  $S \in \text{GL}_n(K)$  mit  $S(A \mid E_n) = (E_n \mid B)$ . Nach den Regeln der Matrixmultiplikation ist  $S(A \mid E_n) = (SA \mid S)$ . Es folgt  $SA = E_n$  und  $S = B$ . Aus  $SA = E_n$  ergibt sich  $A^{-1} = E_n A^{-1} = (SA)A^{-1} = S(AA^{-1}) = SE_n = S$ , d.h.  $B = S = A^{-1}$ .  $\square$



**Beispiel.**  $A = \begin{pmatrix} 1 & 2 \\ -1 & -1 \end{pmatrix}$ , wie in Beispiel (3.2.4).

$$\begin{array}{cc|cc}
 & A & & E_2 \\
 1 & 2 & 1 & 0 \\
 -1 & -1 & 0 & 1 \\
 \hline
 1 & 2 & 1 & 0 \\
 1 & 0 & 1 & 1 \\
 \hline
 1 & 0 & -1 & -2 \\
 0 & 1 & 1 & 1 \\
 E_2 & & A^{-1} & 
 \end{array}$$

*Übung a.* Man mache sich folgendes klar: Das Schema in der Anwendung Matrix-Inversion ist identisch mit dem Schema für das Lösen von inhomogenen linearen Gleichungssystemen, nur dass es hier mehrere rechte Seiten gibt (eine für jede Spalte auf der rechten Seite). Jede Spalte von  $B$  lässt sich daher als Lösung eines inhomogenen LGS auffassen. Aus dieser Interpretation lässt sich die Gleichung  $AB = E_n$  ablesen.

*Übung b.* Man prüfe  $A$  auf Regularität und berechne die Inverse:

$$A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & -1 & 2 \end{pmatrix}.$$

*Beweis.*

$$A^{-1} = \frac{1}{6} \begin{pmatrix} 5 & -3 & -1 \\ -1 & 3 & -1 \\ -3 & 3 & 3 \end{pmatrix}.$$

□

### 6.4.10 Nullraum vs. Spaltenraum

Es sei  $U \leq K^n$ . In 6.4.6 wurde gefragt, ob sich  $U$  als Spaltenraum einer Matrix  $B$  und als Nullraum einer Matrix  $A$  schreiben lässt. Ersteres kommt der Angabe eines Erzeugendensystems gleich (wenn  $B$  dabei minimale Spaltenzahl hat, sogar der Angabe einer Basis) und ist offensichtlich möglich. Letzteres stellt eine Beschreibung von  $U$  durch *definierende Gleichungen* dar. Da jede Zeile von  $A$  eine Gleichung darstellt, ist eine minimale Zeilenzahl gewünscht. Beide Schreibweisen haben ihre Vor- und Nachteile, etwa beim Test  $c \in U$  für gegebenes  $c \in K^n$ .

**Beispiel a.** Es sei  $E$  eine Ebene im euklidischen Raum  $\mathbb{R}^3$ . Die Schreibweise  $E = \text{SR}(B)$  mit  $B = \begin{pmatrix} u_1 & v_1 \\ u_2 & v_2 \\ u_3 & v_3 \end{pmatrix}$  ist die *Parameterform* von  $E$ , wobei  $u, v$  die Richtungsvektoren sind. Die Schreibweise  $E = \mathbb{L}(A, 0)$  mit  $A = \begin{pmatrix} a_1 & a_2 & a_3 \end{pmatrix}$  ist die *Hesse'sche Normalenform* von  $E$ , wobei  $\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix}$  der Normalenvektor ist.

Dieser Abschnitt zeigt, wie sich ein gegebener Spaltenraum als ein Nullraum schreiben lässt, womit dann beide Ausgangsfragen positiv beantwortet sind. Es seien  $A \in K^{m \times n}$  und  $B \in K^{n \times l}$ .

**Lemma.**

$$(i) \text{ SR}(B) \subseteq \mathbb{L}(A, 0) \Leftrightarrow AB = 0.$$

$$(ii) \text{ SR}(B) = \mathbb{L}(A, 0) \Leftrightarrow AB = 0 \text{ und } \text{Rg } A + \text{Rg } B = n.$$

*Beweis.* (i) Offensichtlich gilt  $AB = 0$  genau dann, wenn jede Spalte von  $B$  in  $\mathbb{L}(A, 0)$  liegt. Da  $\mathbb{L}(A, 0)$  abgeschlossen unter Linearkombinationen ist (es ist ein Unterraum), bedeutet letzteres gerade  $\text{SR}(B) \subseteq \mathbb{L}(A, 0)$ .

(ii) Es gilt  $\dim \text{SR}(B) = \text{Rg } B$ , und nach Satz 6.4.6 gilt  $\dim \mathbb{L}(A, 0) = n - \text{Rg } A$ . Also haben  $\text{SR}(B)$  und  $\mathbb{L}(A, 0)$  genau dann gleiche Dimension, wenn  $\text{Rg } A + \text{Rg } B = n$  ist.  $\square$

**Satz.**  $\text{SR}(B) = \mathbb{L}(A, 0) \Leftrightarrow \mathbb{L}(B^t, 0) = \text{SR}(A^t)$ .

*Beweis.* Durch zweimalige Anwendung des Lemmas folgt:

$$\begin{aligned} \text{SR}(B) = \mathbb{L}(A, 0) &\Leftrightarrow AB = 0 \text{ und } \text{Rg } A + \text{Rg } B = n \\ &\Leftrightarrow B^t A^t = 0 \text{ und } \text{Rg } A^t + \text{Rg } B^t = n \\ &\Leftrightarrow \text{SR}(A^t) = \mathbb{L}(B^t, 0). \end{aligned}$$

$\square$

**Beispiel b.** Für  $B = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$  ist  $A$  gesucht mit  $\mathbb{L}(A, 0) = \text{SR}(B)$ . Nach dem

Satz ist das äquivalent zu  $\mathbb{L}(B^t, 0) = \text{SR}(A^t)$ . Nun kann  $A^t$  mit dem Verfahren aus 6.5.1 bestimmt werden. Als Ergebnis erhält man:

$$A = \begin{pmatrix} 1 & -1 & 0 \\ 1 & 0 & -1 \end{pmatrix}.$$

*Übung.* Es sei  $U \leq K^n$  und  $\dim U = d$ . Man zeige, dass  $U$  Nullraum einer Matrix  $A \in K^{(n-d) \times n}$  ist.

## 6.5 Lineare Gleichungssysteme und Matrizen II

Hier werden wir die Theorie der Linearen Gleichungssysteme noch einmal systematisch im Lichte von Vektorräumen, linearen Abbildungen und Matrizen betrachten.

### 6.5.1 Der Lösungsraum eines homogenen LGS

Es seien  $K$  ein Körper und  $A \in K^{m \times n}$ . Bekanntlich ist  $\mathbb{L}(A, 0)$  ein Unterraum von  $K^n$ . Wir wollen Basis und Dimension dieses *Lösungsraumes* bestimmen.

**Satz.** *Ist die Matrix  $A$  in Normalform, also*

$$A = \left( \begin{array}{cccc|ccc} 1 & 0 & \dots & 0 & c_{1(r+1)} & \dots & c_{1n} \\ 0 & 1 & & 0 & c_{2(r+1)} & \dots & c_{2n} \\ \vdots & & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & c_{r(r+1)} & \dots & c_{rn} \\ 0 & \dots & \dots & 0 & 0 & \dots & 0 \\ 0 & & \ddots & 0 & 0 & \ddots & 0 \\ 0 & \dots & \dots & 0 & 0 & \dots & 0 \end{array} \right),$$

$\underbrace{\hspace{10em}}_r$ 
 $\underbrace{\hspace{10em}}_{n-r}$

so bilden die Spalten der Matrix

$$L := \left( \begin{array}{ccc} c_{1(r+1)} & \dots & c_{1n} \\ c_{2(r+1)} & \dots & c_{2n} \\ \vdots & \ddots & \vdots \\ c_{r(r+1)} & \dots & c_{rn} \\ -1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & -1 \end{array} \right),$$

eine Basis von  $\mathbb{L}(A, 0)$ . Insbesondere gilt  $\mathbb{L}(A, 0) = \text{SR}(L)$  und  $\dim \mathbb{L}(A, 0) = n - r$ .

*Beweis.* Es seien  $v_{r+1}, \dots, v_n$  die Spalten von  $L$ . Wegen  $A \cdot L = 0$  (nachprüfen!) gilt  $v_{r+1}, \dots, v_n \in \mathbb{L}(A, 0)$ . An der Form von  $L$  erkennt man, dass  $B := \{v_{r+1}, \dots, v_n\}$  linear unabhängig ist (Teil (ii) von Übung 6.2.3a) und  $|B| = n - r$ . Wir zeigen nun, dass für ein beliebiges  $w \in \mathbb{L}(A, 0) \setminus B$  die Menge  $B \cup \{w\}$  linear abhängig ist; dann ist gezeigt, dass  $B$  eine maximal

linear unabhängige Teilmenge von  $\mathbb{L}(A, 0)$ , also Basis von  $\mathbb{L}(A, 0)$  ist (Satz 6.2.4). Sei also

$$w = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix} \in \mathbb{L}(A, 0) \setminus B, \quad w' := w + w_{r+1}v_{r+1} + \dots + w_nv_n \in \mathbb{L}(A, 0).$$

Wir zeigen  $w' = 0$ ; dann folgt, dass  $B \cup \{w\}$  linear abhängig ist. Man rechnet nach, dass für geeignete  $w'_1, \dots, w'_r \in K$ :

$$w' = \begin{pmatrix} w'_1 \\ \vdots \\ w'_r \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in K^n, \quad Aw' = \begin{pmatrix} w'_1 \\ \vdots \\ w'_r \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in K^m.$$

Aus  $Aw' = 0$  folgt also wie gewünscht  $w' = 0$ . □

**Folgerung.** *Alle Zeilenstufenformen von  $A$  haben dieselbe Stufenzahl  $r$  und es gilt*

$$\dim \mathbb{L}(A, 0) = n - r.$$

*Wir sprechen im folgenden auch einfach von der Stufenzahl von  $A$  (anstatt von der Stufenzahl der Zeilenstufenform von  $A$ ). Sie ist genau dann gleich  $n$ , wenn  $Ax = 0$  nur trivial lösbar ist.*

**Beispiel.** Es sei  $A = \begin{pmatrix} 1 & 1 & 1 \end{pmatrix}$ . Dann bilden die Spalten von  $L = \begin{pmatrix} 1 & 1 \\ -1 & 0 \\ 0 & -1 \end{pmatrix}$  einen Basis von  $\mathbb{L}(A, 0)$ . Insbesondere gilt  $\mathbb{L}(A, 0) = \text{SR}(L)$  und  $\dim \mathbb{L}(A, 0) = 3 - 1 = 2$ .

*Achtung: Hat man, um  $A$  auf Normalform zu bringen, Spaltenvertauschungen gemacht, so muss man diese in Form von Zeilenvertauschungen in der Basis von  $\mathbb{L}(A, 0)$  wieder rückgängig machen.*

**Beispiel a.** Es sei

$$A = \begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}.$$

Dann bilden die Spalten von

$$B = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ -1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}$$

eine Basis von  $\mathbb{L}(A, 0)$ .

### 6.5.2 Lösbarkeitskriterien

Hier wollen wir verschiedene Lösbarkeitskriterien für lineare Gleichungssysteme vorstellen. Es sei  $A \in K^{m \times n}$ .

**Satz a.** *Es sei  $b \in K^m$  und  $(A \mid b)$  die erweiterte Koeffizientenmatrix. Dann sind folgende Aussagen äquivalent:*

- 1.)  $Ax = b$  ist lösbar.
- 2.)  $b \in \text{SR}(A)$ .
- 3.)  $\text{SR}(A \mid b) = \text{SR}(A)$ .
- 4.)  $\text{Rg}(A \mid b) = \text{Rg}(A)$ .

*Beweis.* Siehe Vorlesung. □

**Satz b.** *Es sei  $A'$  eine Zeilenstufenform von  $A$ . Folgende Aussagen sind äquivalent:*

- 1.)  $Ax = b$  hat für jedes  $b \in K^m$  mindestens eine Lösung.
- 2.)  $A'$  hat Stufenzahl  $m$ .
- 3.)  $\varphi_A$  ist surjektiv.
- 4.)  $\text{Rg } A = m$ .

*Beweis.* Nach Folgerung 6.4.9a gibt es  $S \in \text{GL}_m(K)$  mit  $A' = SA$ .

1.) $\Rightarrow$ 2.) Setze  $b := \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix} \in K^m$ . Sei  $x$  eine Lösung von  $Ax = S^{-1}b$ . Dann

ist  $A'x = S Ax = S S^{-1}b = b$ . Das geht nur, wenn  $A'$  genau  $m$  Stufen hat.

2.) $\Rightarrow$ 1.) Sei  $b \in K^m$  beliebig. Da  $A'$  genau  $m$  Stufen hat, gibt es  $x \in K^m$  mit  $A'x = Sb$ . Dann ist  $Ax = S^{-1}A'x = S^{-1}Sb = b$ .

1.) $\Leftrightarrow$ 3.) ist klar nach Bemerkung 3.3.4.

3.) $\Leftrightarrow$ 4.) ist die Definition von Rang.  $\square$

*Übung a.* Wie sieht die Normalform von  $A$  aus, wenn die Aussagen von Satz b gelten?

*Übung b.* Man zeige, dass die Aussagen von Satz b äquivalent dazu sind, dass  $e_1, \dots, e_m$  im Bild von  $\varphi_A$  liegen.

*Übung c.* Es sei  $n = m$ . Man zeige, dass die Aussagen von Satz b sowie die Aussagen aus Folgerung 3.3.4 äquivalent sind denen aus Satz 6.5.3, und mache sich die Bedeutung klar.

*Beweis.* Zeilenstufenform und reduzierte Zeilenstufenform haben offensichtlich gleiche Stufenzahl. Eine quadratische  $n \times n$ -Matrix in reduzierter Zeilenstufenform hat genau dann die Stufenzahl  $n$ , wenn sie die Einheitsmatrix ist.  $\square$

Analog zu Satz b kann man den Fall charakterisieren, in dem es das Lineare Gleichungssystem  $Ax = b$  für jede rechte Seite höchstens eine Lösung hat.

**Satz c.** *Es sei  $A'$  eine Zeilenstufenform von  $A$ . Folgende Aussagen sind äquivalent:*

1.)  $Ax = b$  hat für jedes  $b \in K^m$  höchstens eine Lösung.

2.)  $Ax = 0$  hat genau eine Lösung.

3.)  $A'$  hat Stufenzahl  $n$ .

4.)  $\varphi_A$  ist injektiv.

5.)  $\text{Rg } A = n$ .

*Beweis.* Siehe Vorlesung.  $\square$

### 6.5.3 Quadratische Koeffizientenmatrizen

Es sei  $K$  ein Körper und  $A$  die Koeffizientenmatrix eines linearen Gleichungssystems über  $K$ . Wir nehmen an, dass  $A$  quadratisch ist, d.h. das System hat genauso viele Unbekannte wie Gleichungen. Sei  $A \in K^{n \times n}$ .

**Bemerkung.** Wenn  $A$  invertierbar ist, dann ist  $A \cdot x = b$  für jedes  $b \in K^n$  eindeutig lösbar, und die Lösung lautet  $x = A^{-1}b$ .

*Beweis.* Eindeutigkeit: Aus  $Ax = b = Ax'$  folgt  $x = E_n x = (A^{-1}A)x = A^{-1}(Ax) = A^{-1}(Ax') = (A^{-1}A)x' = E_n x' = x'$ .

Existenz:  $A(A^{-1})x = (AA^{-1})x = E_n x = x$ . □

**Beispiel.** Löse  $\begin{pmatrix} 1 & 2 \\ -1 & -1 \end{pmatrix} \cdot x = b$  für verschiedene  $b \in K^2$ . Da  $A$  invertierbar ist (vgl. Beispiel 3.2.4), ist  $Ax = b$  für jedes  $b \in K^2$  eindeutig lösbar. Die Lösung erhält man einfach durch Multiplikation mit  $A^{-1} = \begin{pmatrix} -1 & -2 \\ 1 & 1 \end{pmatrix}$ :

$$\begin{aligned} Ax = \begin{pmatrix} 1 \\ 0 \end{pmatrix} &\Rightarrow x = A^{-1} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}. \\ Ax = \begin{pmatrix} 0 \\ 1 \end{pmatrix} &\Rightarrow x = A^{-1} \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} -2 \\ 1 \end{pmatrix}. \\ Ax = \begin{pmatrix} -1 \\ 1 \end{pmatrix} &\Rightarrow x = A^{-1} \begin{pmatrix} -1 \\ 1 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \end{pmatrix}. \end{aligned}$$

*Beweis.* Übung. □

Wir werden jetzt die Umkehrung der Bemerkung beweisen, also zeigen: Wenn  $Ax = 0$  nur trivial lösbar ist, dann ist  $A$  invertierbar. Ein Verfahren zur Bestimmung der Inversen von  $A$  wurde in Algorithmus 6.4.9 hergeleitet. Die Äquivalenz von (i) und (iv) des folgenden Satzes wurde bereits in Bemerkung 6.4.9e bewiesen.

**Satz.** Es seien  $A \in K^{n \times n}$  und  $A'$  eine reduzierte Zeilenstufenform von  $A$ . Folgende Aussagen sind äquivalent:

- (i)  $A$  ist invertierbar.
- (ii)  $A \cdot x = 0$  ist eindeutig lösbar (nur trivial lösbar).
- (iii)  $A' = E_n$ .
- (iv)  $A \rightsquigarrow E_n$ .
- (v)  $A$  ist das Produkt von Elementarmatrizen.

*Beweis.* Wir machen einen Ringschluß.

(i)  $\Rightarrow$  (ii) Da  $A$  invertierbar ist gilt:  $Ax = 0 \Rightarrow x = E_n x = A^{-1}Ax = A^{-1}0 = 0$ .

(ii)  $\Leftrightarrow$  (iii)  $Ax = 0$  ist genau dann eindeutig lösbar, wenn es keine freien Unbekannten gibt, also genau dann wenn  $A'$  genau  $n$  Stufen hat, also genau dann

wenn  $A' = E_n$  ist.

(iii)  $\Rightarrow$  (iv)  $A \rightsquigarrow A' = E_n$ .

(iv)  $\Rightarrow$  (v) Wegen  $A \rightsquigarrow E_n$  gibt es nach Definition 6.4.9 und Bemerkung 6.4.9b Elementarmatrizen  $S_1, \dots, S_r$  so, dass  $E_n = S_r \cdots S_1 A$  ist. Es folgt  $A = S_1^{-1} \cdots S_r^{-1}$  und die  $S_i^{-1}$  sind Elementarmatrizen nach Bemerkung 6.4.9a.

(v)  $\Rightarrow$  (i) Da  $\text{GL}_n(K)$  eine Gruppe ist, sind mit Elementarmatrizen auch deren Produkte wieder invertierbar.  $\square$

### 6.5.4 Matrixgleichungen

Es seien in diesem ganzen Abschnitt  $K$  ein Körper und  $m \in \mathbb{N}$ . Für  $i \leq m$

bezeichne  $e_i$  das Element  $\begin{pmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{pmatrix} \in K^n$  mit dem 1-Eintrag an der  $i$ -ten Stelle.

**Definition.** Es sei  $A \in K^{m \times n}$ . Gleichungen der Form

$$A \cdot X = B \quad \text{und} \quad X \cdot A = B$$

mit gegebener Matrix  $B$  und unbekannter Matrix  $X$  bezeichnen wir als *Matrixgleichungen*. Dabei ist im ersten Fall  $B \in K^{m \times l}$  und  $X \in K^{n \times l}$  für ein  $l \in \mathbb{N}$ , und im zweiten Fall  $B \in K^{l \times n}$  und  $X \in K^{l \times m}$ .

**Bemerkung.** Es seien  $b_1, \dots, b_l \in K^m$  die Spalten von  $B \in K^{m \times l}$ .

- (i) Ein wichtiger Spezialfall sind die Gleichungen  $A \cdot X = E_m$  und  $X \cdot A = E_n$ .
- (ii) Die Lösungen von  $A \cdot X = B$  sind genau diejenigen Matrizen  $X \in K^{n \times l}$ , deren Spalten  $x_1, \dots, x_l \in K^n$

$$A \cdot x_i = b_i$$

lösen für jedes  $1 \leq i \leq l$ .

Die Gleichung  $A \cdot X = B$  kann daher als Zusammenfassung mehrerer linearer Gleichungssysteme interpretiert werden, und zwar eines für jede Spalte von  $B$ . Dabei haben alle linearen Gleichungssysteme dieselbe Koeffizientenmatrix  $A$ , aber verschiedene rechte Seiten. Entsprechend wird  $A \cdot X = B$  genauso gelöst wie einzelne lineare Gleichungssysteme: mit dem Gauß-Algorithmus.



- (iii) Ist  $A \in K^{n \times n}$ , so kommt das Schema zur Matrix-Inversion dem Lösen der Matrixgleichung  $A \cdot X = E_n$  gleich. Nach (ii) entspricht diese Gleichung wiederum den einzelnen Gleichungssystemen  $Ax = e_i$  für  $i = 1, \dots, n$ .
- (iv) Die Gleichung  $A \cdot X = B$  ist genau dann eindeutig lösbar, wenn  $Ax_i = b_i$  für jedes  $1 \leq i \leq l$  eindeutig lösbar ist. Gemäß Satz 3.3.4 ist das genau dann der Fall, wenn  $A \cdot X = B$  lösbar ist und  $A \cdot x = 0$  nur trivial lösbar.
- (v) Die Gleichung  $X \cdot A = B$  kann durch Transposition in eine Gleichung der Form  $A' \cdot X = B'$  überführt werden, denn sie ist äquivalent zu  $A^t \cdot X^t = B^t$ . Hat man  $X^t$  gefunden (wie in (ii) beschrieben), so erhält man  $X$  durch Transponieren, denn  $X = (X^t)^t$ .

**Beispiel a.** Es sei  $K = \mathbb{R}$ . Löse

$$\begin{pmatrix} 2 & -1 & 1 \\ -4 & 2 & -1 \end{pmatrix} \cdot X = \begin{pmatrix} 2 & 0 \\ -3 & -2 \end{pmatrix}.$$

Eine Lösung ist  $X = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 1 & -2 \end{pmatrix}$ . Alle Lösungen erhält man, indem zu jeder

Spalte von  $X$  beliebige Vielfache von  $\begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}$  addiert werden.

(Rechnung als Übung.)

**Beispiel b.** Es sei  $K = \mathbb{R}$ . Löse

$$X \cdot \begin{pmatrix} 2 & -4 \\ -1 & 2 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 2 & -3 \\ 0 & -2 \end{pmatrix}.$$

Wir lösen stattdessen  $A^t \cdot Y = B^t$ . Gemäß Beispiel a ist  $Y = \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 1 & -2 \end{pmatrix}$

eine Lösung. Somit ist  $X = Y^t = \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & -2 \end{pmatrix}$  eine Lösung der Ausgangsgleichung. Alle Lösungen erhält man, indem zu jeder Zeile von  $X$  beliebige Vielfache von  $(1, 2, 0)$  addiert werden.

*Übung.* Man zeige, dass  $AX = B$  genau dann eindeutig lösbar ist, wenn  $m = n$  und  $A$  invertierbar ist. In diesem Fall ist  $X = A^{-1}B$  die eindeutige Lösung.

### 6.5.5 Links- und Rechtsinverse

**Definition.** Es sei  $A \in K^{m \times n}$ .

- (i) Gibt es eine Matrix  $B \in K^{n \times m}$  mit  $A \cdot B = E_m$ , so heißt  $A$  *rechtsinvertierbar* und  $B$  eine *Rechtsinverse* zu  $A$ .
- (ii) Gibt es eine Matrix  $B \in K^{n \times m}$  mit  $B \cdot A = E_n$ , so heißt  $A$  *linksinvertierbar* und  $B$  eine *Linksinverse* zu  $A$ .

**Bemerkung.** Links- und Rechtsinverse müssen nicht existieren, und wenn sie existieren müssen sie nicht eindeutig sein. Die Bedeutung ihrer Existenz wird im folgenden Satz klar.

**Satz.** Es sei  $A \in K^{m \times n}$  und  $b \in K^m$ .

- (i)  $A$  besitzt genau dann eine Rechtsinverse  $R$ , wenn die Aussagen aus Satz 6.5.2b gelten. In diesem Fall ist  $n \geq m$ ,  $\varphi_R$  eine rechtsseitige Umkehrabbildung von  $\varphi_A$  und  $R \cdot b$  eine Lösung von  $A \cdot x = b$ .
- (ii)  $A$  besitzt genau dann eine Linksinverse  $L$ , wenn die Aussagen aus Satz 6.5.2c gelten. In diesem Fall ist  $m \geq n$ ,  $\varphi_L$  eine linksseitige Umkehrabbildung von  $\varphi_A$  und  $L \cdot b$  die einzig mögliche Lösung von  $A \cdot x = b$ .

*Beweis.* (i) Ist  $R$  eine Rechtsinverse von  $A$ , so ist  $A \cdot (R \cdot b) = (A \cdot R) \cdot b = E_m \cdot b = b$ , d.h.  $R \cdot b$  ist eine Lösung von  $A \cdot x = b$ . Damit gelten auch die Aussagen aus Satz 6.5.2b. Umgekehrt findet man die  $i$ -te Spalte von  $R$  als Lösung von  $Ax = e_i$  (vgl. Bemerkung 6.5.4).

(ii) Ist  $L$  eine Linksinverse von  $A$  und  $A \cdot x = b$  lösbar, so folgt  $x = E_n \cdot x = (L \cdot A) \cdot x = L \cdot (A \cdot x) = L \cdot b$ . Damit gelten die Aussagen aus Satz 6.5.2c. Wir zeigen nun die Umkehrung, also sei  $Ax = 0$  eindeutig lösbar. Nach Abschnitt 3.4.4 hat  $A$  die reduzierte Zeilenstufenform  $\begin{pmatrix} E_n \\ 0 \end{pmatrix}$ . Nach Folgerung 6.4.9a gibt es  $S \in \text{GL}_m(K)$  mit  $SA = \begin{pmatrix} E_n \\ 0 \end{pmatrix}$ . Die oberen  $n$  Zeilen von  $S$  bilden somit eine Linkinverse von  $A$ .  $\square$

**Beispiel a** (Rechtsinverse). Es sei  $A = \begin{pmatrix} 2 & -1 & 1 \\ -4 & 2 & -1 \end{pmatrix} \in \mathbb{Q}^{2 \times 3}$ . Wir bestimmen zunächst eine Rechtsinverse  $R \in \mathbb{Q}^{3 \times 2}$  von  $A$ . Die  $i$ -te Spalte von  $R$

ist eine Lösung von  $Ax = e_i$ . Wir berechnen:

$$\begin{aligned}\mathbb{L}(A, \begin{pmatrix} 1 \\ 0 \end{pmatrix}) &= \begin{pmatrix} -1/2 \\ 0 \\ 2 \end{pmatrix} + \mathbb{Q} \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} . \\ \mathbb{L}(A, \begin{pmatrix} 0 \\ 1 \end{pmatrix}) &= \begin{pmatrix} -1/2 \\ 0 \\ 1 \end{pmatrix} + \mathbb{Q} \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} .\end{aligned}$$

Somit existiert  $R$ , ist aber nicht eindeutig. Als eine ganzzahlige Lösung lesen wir z.B.

$$R = \begin{pmatrix} 0 & 0 \\ 1 & 1 \\ 2 & 1 \end{pmatrix}$$

ab. Wir berechnen nun Lösungen von  $A \cdot x = b$  für  $b = \begin{pmatrix} 2 \\ -3 \end{pmatrix}$  und  $b = \begin{pmatrix} 0 \\ -2 \end{pmatrix}$  mit Hilfe der Rechtsinversen:

$$x = R \begin{pmatrix} 2 \\ -3 \end{pmatrix} = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}, \quad y = R \begin{pmatrix} 0 \\ -2 \end{pmatrix} = \begin{pmatrix} 0 \\ -2 \\ -2 \end{pmatrix}.$$

Als  $2 \times 3$ -Matrix kann  $A$  keine Linksinverse besitzen.

**Beispiel b** (Linksinverse). Es sei  $A = \begin{pmatrix} 2 & -4 \\ -1 & 2 \\ 1 & -1 \end{pmatrix} \in \mathbb{Q}^{2 \times 3}$ . Wir bestimmen eine Linksinverse  $L \in \mathbb{Q}^{2 \times 3}$  von  $A$ , indem wir  $A^t X = E_2$  lösen und  $L := X^t$  setzen. Ein solches  $X$  wurde bereits in Beispiel [a](#) berechnet. Daraus bekommen wir  $L = \begin{pmatrix} 0 & 1 & 2 \\ 0 & 1 & 1 \end{pmatrix}$ , dieses  $L$  ist aber nicht eindeutig. Wir berechnen

nun Lösungen von  $A \cdot x = b$  für  $b = e_1, e_2, e_3$ ,  $\begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$  mit Hilfe der Linksinversen.  $Lb$  ist der einzige Kandidat für eine Lösung und wird zur Probe

eingesetzt:

$$\begin{aligned}
 Le_1 &= \begin{pmatrix} 0 \\ 0 \end{pmatrix}, & \text{Probe: } A \begin{pmatrix} 0 \\ 0 \end{pmatrix} &= \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \neq \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \\
 Le_2 &= \begin{pmatrix} 1 \\ 1 \end{pmatrix}, & \text{Probe: } A \begin{pmatrix} 1 \\ 1 \end{pmatrix} &= \begin{pmatrix} -2 \\ \star \\ \star \end{pmatrix} \neq \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \\
 Le_3 &= \begin{pmatrix} 2 \\ 1 \end{pmatrix}, & \text{Probe: } A \begin{pmatrix} 2 \\ 1 \end{pmatrix} &= \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \\
 L \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} &= \begin{pmatrix} 8 \\ 5 \end{pmatrix}, & \text{Probe: } A \begin{pmatrix} 8 \\ 5 \end{pmatrix} &= \begin{pmatrix} -4 \\ \star \\ \star \end{pmatrix} \neq \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}
 \end{aligned}$$

Also gilt

$$\mathbb{L}(A, e_1) = \mathbb{L}(A, e_2), \mathbb{L}(A, \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}) = \emptyset, \mathbb{L}(A, e_3) = \left\{ \begin{pmatrix} 2 \\ 1 \end{pmatrix} \right\}.$$

Als  $3 \times 2$ -Matrix kann  $A$  keine Rechtsinverse besitzen.

**Folgerung.** Für  $A \in K^{n \times n}$  sind äquivalent:

- (i)  $A$  ist invertierbar.
- (ii)  $A$  besitzt eine Linksinverse.
- (iii)  $A$  besitzt eine Rechtsinverse.

*Beweis.* Dies folgt aus obigem Satz zusammen mit den Sätzen 6.5.2b und 6.5.2c.  $\square$

*Übung.* Man zeige, dass jede Matrix  $A$ , die sowohl eine Links- als auch eine Rechtsinverse besitzt, quadratisch und invertierbar ist, und dass dann die Links- und Rechtsinversen eindeutig sind und mit  $A^{-1}$  übereinstimmen.

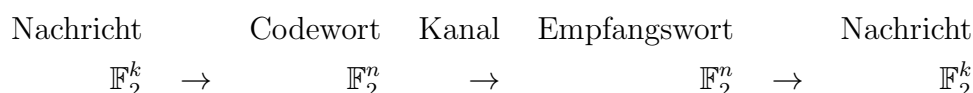
## 6.6 Anwendung: Lineare Codes

Ein Sender schickt Bitfolgen über einen Kanal, der evtl. fehlerbehaftet ist, zu einem Empfänger. Über den Kanal werden folgende Annahmen gemacht:

- (i) Es gehen keine Bits verloren.

- (ii) Die Fehlerwahrscheinlichkeit für ein einzelnes Bit, bei der Übertragung zu kippen, ist  $< 1/2$ .

Um Fehler erkennen bzw. sogar korrigieren zu können, wird nach folgendem Schema vorgegangen ( $n > k$ ):



Die Bitfolge wird also in Nachrichten fester Länge ( $k$  bit) zerlegt, und jede Nachricht als ein Codewort ( $n$  bit) codiert. Die Menge  $C$  aller möglichen Codewörter ist eine echte Teilmenge von  $\mathbb{F}_2^n$ . Ist das Empfangswort kein Codewort, so kann der Empfänger mit Sicherheit davon ausgehen, dass ein Fehler bei der Übertragung stattgefunden hat (und evtl. eine erneute Sendung anfordern). Übertragungsfehler, bei denen ein Codewort in ein anderes übergeht, können allerdings nicht erkannt werden. Die Idee ist nun,  $C$  so zu wählen, dass sich verschiedene Codewörter an hinreichend vielen Stellen unterscheiden. Dadurch wird es unwahrscheinlich, dass ein Codewort durch einen Übertragungsfehler in ein anderes Codewort übergeht.

**Definition a.** Ein Unterraum  $C \leq \mathbb{F}_2^n$  heißt (*binärer*) *linearer Code* der Länge  $n$ . Die Elemente von  $C$  heißen *Codewörter*.

**Bemerkung a.** In einem Code  $C$  der Dimension  $k$  gibt es  $2^k$  Codewörter.

**Codierung.** Der Sender schreibt den Code  $C$  als Spaltenraum  $\text{SR}(G)$  mit  $G \in \mathbb{F}_2^{n \times k}$  und minimaler Spaltenzahl. Dann ist  $\dim C = k$  und die Spalten von  $G$  sind linear unabhängig. Insbesondere ist  $Gx = 0$  nur trivial lösbar (vgl. Bemerkung (6.2.2)). Der Kern der Abbildung

$$\varphi_G : \mathbb{F}_2^k \rightarrow \mathbb{F}_2^n, \quad v \mapsto Gv$$

ist also  $\{0\}$ , d.h.  $\varphi_G$  ist injektiv. Das Bild von  $\varphi_G$  ist per Definition  $\text{SR}(G) = C$ . Daher kann  $\varphi_G$  als Codierungsabbildung verwendet werden. Die Matrix  $G$  wird *Generatormatrix* von  $C$  genannt.

**Dekodierung.** Der Empfänger schreibt den Code  $C$  als Nullraum  $\mathbb{L}_0(H)$  mit  $H \in \mathbb{F}_2^{l \times n}$  und minimaler Zeilenzahl. Dann ist  $l = n - \dim C = n - k$ . Zur Prüfung des Empfangswortes  $w \in \mathbb{F}_2^n$  berechnet der Empfänger  $Hw \in \mathbb{F}_2^l$ , das Prüfergebnis. Ist  $Hw \neq 0$ , so liegt mit Sicherheit ein Fehler vor, da  $w$  kein Codewort ist. Ist  $Hw = 0$ , so ist  $w$  ein Codewort und der Empfänger geht davon aus, dass kein Fehler vorliegt (was mit einer gewissen Wahrscheinlichkeit richtig ist). Die Matrix  $H$  wird *Kontrollmatrix* von  $C$  genannt.

**Bemerkung b.** Weder Generator- noch Kontrollmatrix sind eindeutig durch den Code  $C$  definiert. Die Größen beider Matrizen sind aber durch Länge und Dimension von  $C$  bestimmt.

Einen Übertragungsfehler auf einem Wort  $c \in \mathbb{F}_2^n$  stellen wir uns als Addition im Vektorraum  $\mathbb{F}_2^n$  eines *Fehlervektors*  $\epsilon \in \mathbb{F}_2^n$  vor. Da die Addition über  $\mathbb{F}_2$  der xor-Verknüpfung entspricht, verändert die Addition von  $\epsilon$  zu  $c$  genau die Einträge von  $c$ , die an einer Position stehen, an der  $\epsilon$  eine 1 enthält.

**Bemerkung c.** Es bleiben genau die Übertragungsfehler  $\epsilon \in \mathbb{F}_2^n$  unerkannt, für die  $H\epsilon = 0$  ist, d.h. die selbst Codewörter sind.

*Beweis.* Sei  $c \in C$  das gesendete Codewort und  $w = c + \epsilon$  das Empfangswort. Der Fehler  $\epsilon$  bleibt unentdeckt, wenn das Prüfergebnis  $Hw = H(c + \epsilon) = Hc + H\epsilon = 0$  ist. Wegen  $c \in C$  ist  $Hc = 0$ , also  $Hw = 0$  genau dann, wenn  $H\epsilon = 0$ .  $\square$

**Definition b.** Wir nennen einen Fehlervektor  $\epsilon \in \mathbb{F}_2^n$  *einfach*, wenn  $\epsilon$  genau einen 1-Eintrag enthält und sonst nur Nullen.

**Satz.**

- (i) Sind alle Spalten von  $H$  ungleich 0, so werden alle 1-fachen Übertragungsfehler erkannt.
- (ii) Sind alle Spalten von  $H$  ungleich 0 und paarweise verschieden, so können alle 1-fachen Übertragungsfehler vom Empfänger korrigiert werden.

*Beweis.* Es sei  $\epsilon$  ein 1-facher Fehlervektor, d.h.  $\epsilon = e_i$  für einen Einheitsvektor  $e_i$ . Das Prüfergebnis lautet dann  $H(c + \epsilon) = Hc + H\epsilon = He_i = i$ -te Spalte von  $H$ , und zwar unabhängig vom Codewort  $c \in C$ . Sind alle Spalten von  $H$  ungleich 0, so ist also  $He_i \neq 0$  für alle  $i$ , d.h. alle 1-fachen Fehler werden erkannt.

Sind die Spalten von  $H$  ausserdem paarweise verschieden, so erlaubt das Prüfergebnis  $He_i$  einen eindeutigen Rückschluß auf die Stelle  $i$ , durch Vergleich des Prüfergebnisses mit allen Spalten von  $H$ .  $\square$

**Beispiel a.** (3-facher Wiederholungscode)  $C = \left\{ \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \right\} \leq \mathbb{F}_2^3$  ist ein

Code der Länge 3 mit Dimension 1. Eine Generator- und Kontrollmatrix sind z.B.

$$G = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \quad H = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}.$$

An  $H$  erkennt man, dass alle 1-fachen Fehler korrigiert werden können. Z.B. für  $\epsilon = e_3$ :

$$v = (1) \mapsto c = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \mapsto w = c + \epsilon = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \mapsto Hw = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \text{3-te Spalte von } H.$$

**Beispiel b.** (Konstruktion von Codes) Wichtige Eigenschaften des Codes sind nach dem Satz an der Kontrollmatrix zu erkennen. Aus diesem Grund konstruieren wir nun einen Code indirekt über seine Kontrollmatrix. Angenommen, die Zahl  $n - k$  (also die Zeilenzahl von  $H$ ) sei fest. Für einen effizienten Code ist die Zahl der Codewörter,  $2^k$ , zu maximieren. Wegen  $k = n - (n - k)$  ist also  $n$ , die Spaltenzahl von  $H$ , zu maximieren.

Wollen wir etwa einen möglichst effizienten Code mit  $n - k = 3$  konstruieren, der alle 1-fachen Fehler korrigieren kann, so müssen wir in die Spalten von  $H$  genau die verschiedenen Spaltenvektoren aus  $\mathbb{F}_2^3$  ungleich 0 eintragen. Das sind die Binärzahlen 1 bis 7, also

$$H = \begin{pmatrix} 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix}.$$

Der Code  $C = \mathbb{L}_0(H)$  heißt *Hamming-Code*. Eine Generatormatrix ist z.B.

$$G = \begin{pmatrix} 1 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Die Rechnung zur Bestimmung von  $G$  wurde bereits in Beispiel (6.5.1) gemacht.





# Kapitel 7

## Determinanten und Eigenvektoren

### 7.1 Determinanten

In diesem Paragraphen seien  $R$  ein kommutativer Ring und  $n \in \mathbb{N}$ . Wir betrachten quadratische  $n \times n$ -Matrizen über  $R$  und bezeichnen mit  $s_1, \dots, s_n \in R^n$  jeweils die Spalten der Matrix. Die Einträge einer  $n \times n$ -Matrix bezeichnen wir mit  $a_{ij}$  oder  $a_{i,j}$ .

#### 7.1.1 Definition und Eigenschaften

Wir beginnen mit der Definition der Determinante einer Matrix.

**Definition.** Die *Determinante* ist die Abbildung

$$\det : R^{n \times n} \rightarrow R, \quad A \mapsto \det(A),$$

definiert durch

$$\det(A) = \sum_{\pi \in S_n} \operatorname{sgn}(\pi) a_{\pi(1),1} a_{\pi(2),2} \cdots a_{\pi(n),n} = \sum_{\pi \in S_n} \operatorname{sgn}(\pi) \prod_{i=1}^n a_{\pi(i),i}.$$

**Bemerkung.** (i) Die in obiger Definition verwendete Formel für die Determinante einer Matrix heißt die *Leibniz-Formel*.

(ii) Für  $\det(A)$  schreiben wir auch  $|A|$ .

**Beispiel.** Für  $n = 2$  ergibt sich

$$\det A = a_{11}a_{22} - a_{21}a_{12}.$$

Für  $n = 3$  ergibt sich die als *Regel von Sarrus* (franz. Mathematiker, 1795-1861) bekannte Formel

$$\det A = a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33}.$$

Man zeichne sich zu diesen Formeln ein Bild (siehe Vorlesung)! Achtung: die Bilder lassen sich nicht auf  $n \geq 4$  verallgemeinern, da die Leibniz-Formel dann zu viele Summanden hat.

Ein Beispiel für  $n = 3$ :

$$\begin{vmatrix} 1 & 2 & 3 \\ 0 & 1 & 2 \\ -1 & 0 & 5 \end{vmatrix} = (5 - 4 + 0) - (-3 + 0 + 0) = 1 - (-3) = 4.$$

**Folgerung.** *Es gilt stets  $\det A = \det A^t$ .*

*Beweis.* Wenn  $\pi$  die Menge  $S_n$  durchläuft, so durchläuft auch  $\pi^{-1}$  die Menge  $S_n$ . Da ausserdem  $\operatorname{sgn} \pi = \operatorname{sgn} \pi^{-1}$  gilt, haben wir nach der Leibniz-Formel  $\det A = \sum_{\pi \in S_n} \operatorname{sgn}(\pi) \prod_{i=1}^n a_{\pi^{-1}(i), i}$ . Wenn  $i$  die Zahlen  $1, \dots, n$  durchläuft, so durchläuft für festes  $\pi \in S_n$  auch  $\pi(i)$  genau  $1, \dots, n$ , weil  $\pi$  eine Bijektion ist. Also gilt  $\prod_{i=1}^n a_{\pi^{-1}(i), i} = \prod_{i=1}^n a_{\pi^{-1} \circ \pi(i), \pi(i)} = \prod_{i=1}^n a_{i, \pi(i)}$ . Dies zeigt  $\det A = \sum_{\pi \in S_n} \operatorname{sgn}(\pi) \prod_{i=1}^n a_{\pi(i), i} = \det A^t$ .  $\square$

### 7.1.2 Formale Eigenschaften der Determinantenabbildung

In diesem Abschnitt fassen identifizieren wir eine Matrix  $A \in R^{n \times n}$  mit dem  $n$ -Tupel ihrer Spalten  $(s_1, \dots, s_n)$ , also  $R^{n \times n}$  mit  $R^n \times R^n \times \dots \times R^n$  ( $n$  Faktoren).

**Satz.** *Die Determinante aus 7.1.1 erfüllt folgende Eigenschaften.*

*Für alle  $1 \leq i, j \leq n$  und  $\lambda \in R$  gelten:*

- a)  $\det(\dots, s_j + s'_j, \dots) = \det(\dots, s_j, \dots) + \det(\dots, s'_j, \dots)$ .
- b)  $\det(\dots, \lambda s_j, \dots) = \lambda \det(\dots, s_j, \dots)$ .
- c)  $\det(\dots, s, \dots, s, \dots) = 0$ .
- d)  $\det(E_n) = 1$ .

*Man sagt, die Determinante ist multilinear (Eigenschaft a) und b)), alternierend (Eigenschaft c)) und normiert (Eigenschaft d)).*

*Beweis.* Für c) siehe Vorlesung; die anderen Aussagen als Übung.  $\square$

**Folgerung.** Es gelten die folgenden Aussagen.

- (i)  $\det(\dots, s_i, \dots, s_j, \dots) = -\det(\dots, s_j, \dots, s_i, \dots)$ .
- (ii)  $\det(s_{\pi(1)}, \dots, s_{\pi(n)}) = \operatorname{sgn}(\pi) \cdot \det(s_1, \dots, s_n)$  für alle  $\pi \in S_n$ .
- (iii)  $\det(e_{\pi(1)}, \dots, e_{\pi(n)}) = \operatorname{sgn}(\pi)$  für alle  $\pi \in S_n$ .
- (iv)  $\det(\dots, 0, \dots) = 0$ .
- (v)  $\det(\lambda A) = \lambda^n \det(A)$ .
- (vi)  $\det(\dots, s_i + \lambda s_j, \dots, s_j, \dots) = \det(\dots, s_i, \dots, s_j, \dots)$ .

*Beweis.* (i) Nach Satz c) und a) haben wir

$$\begin{aligned} 0 &= \det(\dots, s_i + s_j, \dots, s_i + s_j, \dots) \\ &= \det(\dots, s_i, \dots, s_i, \dots) + \det(\dots, s_i, \dots, s_j, \dots) \\ &\quad + \det(\dots, s_j, \dots, s_i, \dots) + \det(\dots, s_j, \dots, s_j, \dots) \\ &= \det(\dots, s_i, \dots, s_j, \dots) + \det(\dots, s_j, \dots, s_i, \dots). \end{aligned}$$

Daraus folgt die Behauptung.

- (ii) Schreibe  $\pi = \tau_1 \circ \tau_2 \circ \dots \circ \tau_l$  mit Transpositionen  $\tau_i$ ,  $1 \leq i \leq l$ . Aus  $\operatorname{sgn}(\pi) = (-1)^l$  folgt zusammen mit (i) die Behauptung.
- (iii) Ist ein Spezialfall von (ii).
- (iv) und (v) ergeben sich aus Satz b).
- (vi) folgt aus Satz a) und b). □

**Bemerkung.** Es sei  $D : R^{n \times n} \rightarrow R$  eine Abbildung, die die Bedingungen a) – c) aus dem Satz erfüllt. Dann ist  $D = D(E_n) \cdot \det$ .

*Beweis.* Als Übung. □

### 7.1.3 Laplace-Entwicklung

**Definition.** Es seien  $A \in R^{n \times n}$ ,  $1 \leq i, j \leq n$ . Wir betrachten die  $(n-1) \times (n-1)$ -Untermatrix  $A_{ij}$  von  $A$ , die durch Streichen der  $i$ -ten Zeile und der  $j$ -ten Spalte entsteht, also

$$A_{ij} := \begin{pmatrix} a_{1,1} & \cdots & a_{1,j-1} & a_{1,j+1} & \cdots & a_{1,n} \\ \vdots & & & & & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j-1} & a_{i-1,j+1} & \cdots & a_{i-1,n} \\ a_{i+1,1} & \cdots & a_{i+1,j-1} & a_{i+1,j+1} & \cdots & a_{i+1,n} \\ \vdots & & & & & \vdots \\ a_{n,1} & \cdots & a_{n,j-1} & a_{n,j+1} & \cdots & a_{n,n} \end{pmatrix}.$$

Die Determinante  $|A_{ij}|$  dieser Untermatrix heißt *Minor von A zu ij*, geschriebenen  $\text{Minor}_{ij}(A)$ .

**Satz.** Für alle  $1 \leq i, j \leq n$  gilt:

$$\det A = \sum_{k=1}^n (-1)^{i+k} a_{ik} \text{Minor}_{ik}(A) = \sum_{k=1}^n (-1)^{k+j} a_{kj} \text{Minor}_{kj}(A).$$

Die erste Summe nennt man die Laplace-Entwicklung nach der  $i$ -ten Zeile, die zweite Summe nennt man die Laplace-Entwicklung nach der  $j$ -ten Spalte.

*Beweis.* Lasse ich weg. □

**Beispiel.**

$$\begin{aligned} \begin{vmatrix} 1 & 2 & 3 & 4 \\ -1 & -1 & 0 & 1 \\ 4 & 0 & 3 & -1 \\ 2 & 0 & -1 & 1 \end{vmatrix} &= -2 \begin{vmatrix} -1 & 0 & 1 \\ 4 & 3 & -1 \\ 2 & -1 & 1 \end{vmatrix} + (-1) \begin{vmatrix} 1 & 3 & 4 \\ 4 & 3 & -1 \\ 2 & -1 & 1 \end{vmatrix} \\ &= -2 \left( - \begin{vmatrix} 3 & -1 \\ -1 & 1 \end{vmatrix} + \begin{vmatrix} 4 & 3 \\ 2 & -1 \end{vmatrix} \right) - \left( \begin{vmatrix} 3 & -1 \\ -1 & 1 \end{vmatrix} - 3 \begin{vmatrix} 4 & -1 \\ 2 & 1 \end{vmatrix} + 4 \begin{vmatrix} 4 & 3 \\ 2 & -1 \end{vmatrix} \right) \end{aligned}$$

Dabei wurde die  $4 \times 4$ -Matrix nach der 2. Spalte entwickelt, und die  $3 \times 3$ -Matrizen jeweils nach der 1. Zeile. Nun ist es sinnvoll, nach gleichen  $2 \times 2$ -Determinanten zu sortieren, bevor diese nach der Formel aus Beispiel 7.1.1 berechnet werden:

$$= \begin{vmatrix} 3 & -1 \\ -1 & 1 \end{vmatrix} - 6 \begin{vmatrix} 4 & 3 \\ 2 & -1 \end{vmatrix} + 3 \begin{vmatrix} 4 & -1 \\ 2 & 1 \end{vmatrix} = 2 - 6 \cdot (-10) + 3 \cdot 6 = 2 + 60 + 18 = 80.$$

**Folgerung a** (Kästchensatz). Ist  $A$  von der Form  $A = \left( \begin{array}{c|c} B & C \\ \hline 0 & D \end{array} \right)$ , so gilt

$$\det A = \det(B) \cdot \det(D).$$

*Beweis.* Wir führen eine Induktion nach  $k$ , wobei  $B \in R^{k \times k}$ . Für  $k = 1$  ist die Aussage klar, wenn man nach der ersten Spalte entwickelt. Sei nun  $k > 1$  und die Behauptung für  $k - 1$  bereits bewiesen. Da  $a_{i1} = 0$  für alle  $i > k$  und  $a_{i1} = b_{i1}$  für alle  $i \leq k$ , ergibt die Entwicklung nach der ersten Spalte:

$$\det A = \sum_{i=1}^k (-1)^{i+1} b_{i1} \text{Minor}_{i1}(A).$$

Sei  $i \leq k$ . Wegen  $A_{i1} = \left( \begin{array}{c|c} B_{i1} & C \\ \hline 0 & D \end{array} \right)$  gilt nach Induktionsvoraussetzung  $\text{Minor}_{i1}(A) = |A_{i1}| = |B_{i1}| \cdot |D|$ , folglich  $\text{Minor}_{i1}(A) = \text{Minor}_{i1}(B) \cdot |D|$ . Das zeigt

$$\det A = \sum_{i=1}^k (-1)^{i+k} b_{i1} \text{Minor}_{i1}(B) |D| = \left( \sum_{i=1}^k b_{i1} \text{Minor}_{i1}(B) \right) |D| = |B| \cdot |D|.$$

□

**Folgerung b.** Obere und untere Dreiecksmatrizen haben als Determinante das Produkt der Diagonaleinträge, d.h.  $\det A = a_{11}a_{22} \cdots a_{nn}$ .

*Beweis.* Induktion nach  $n$  mit dem Kästchensatz, der im Induktionsschritt mit einer  $1 \times 1$ -Matrix  $B$  angewendet wird. □

### 7.1.4 Determinante und Gauß-Algorithmus

Die Regel **b)** aus Satz 7.1.2 und die Regeln **(i)** und **(vi)** aus Folgerung 7.1.2 erlauben es, die Matrix wie beim Gauß-Algorithmus auf eine Dreiecksform zu bringen, woraus sich dann mit der Folgerung aus dem Kästchensatz die Determinante leicht ablesen lässt. Dabei sind sowohl Spalten- als auch Zeilentransformationen erlaubt (auch gemischt). Man beachte jedoch, dass der Gauß-Algorithmus im Allgemeinen nur über Körpern funktioniert.

**Beispiel.**

$$\begin{aligned} A = \begin{vmatrix} 3 & 0 & -2 \\ 6 & 0 & 1 \\ -9 & -2 & 5 \end{vmatrix} &= 3 \begin{vmatrix} 1 & 0 & -2 \\ 2 & 0 & 1 \\ -3 & -2 & 5 \end{vmatrix} = -3 \begin{vmatrix} 1 & -2 & 0 \\ 2 & 1 & 0 \\ -3 & 5 & -2 \end{vmatrix} = -3 \begin{vmatrix} 1 & 0 & 0 \\ 2 & 5 & 0 \\ -3 & -1 & -2 \end{vmatrix} \\ &= (-3) \cdot 1 \cdot 5 \cdot (-2) = 30. \end{aligned}$$

**Folgerung.** Ist  $R = K$  ein Körper, so gilt:  $\det A \neq 0 \Leftrightarrow \text{Rg } A = n$ . Also:

$$\text{GL}_n(K) = \{A \in K^{n \times n} \mid \det A \neq 0\}.$$

*Beweis.* Geht  $A'$  aus  $A$  durch eine Folge elementarer Zeilen- und Spaltentransformationen hervor, so ist  $\det A' = \lambda \det A$  für ein  $\lambda \in K \setminus \{0\}$  (siehe Regeln **b)**, **(i)**, **(vi)** aus 7.1.2). In diesem Fall gilt also  $\det A = 0 \Leftrightarrow \det A' = 0$ . Wir können daher o.B.d.A. annehmen, dass  $A$  in Zeilenstufenform, also eine obere Dreiecksmatrix ist. Eine obere Dreiecksmatrix der Größe  $n \times n$  hat offensichtlich genau dann Rang  $n$  (volle Stufenzahl), wenn alle Einträge entlang der Hauptdiagonalen ungleich 0 sind. Nach der Folgerung aus dem Kästchensatz ist das genau dann der Fall, wenn  $\det A \neq 0$ . (Man beachte, dass Körper nullteilerfrei sind.) □

### 7.1.5 Produktsatz

**Satz.** Es gilt  $\det(AB) = \det A \cdot \det B$  für alle  $A, B \in R^{n \times n}$ .

*Beweis.* Wir sehen  $A$  als fest an und definieren die Abbildung

$$D : R^{n \times n} \rightarrow R, \quad B \mapsto \det(AB).$$

Man prüft leicht nach, dass  $D$  multilinear und alternierend ist. Das liegt daran, dass für die Spalten  $s_1, \dots, s_n$  von  $A$  gilt  $D(s_1, \dots, s_n) = \det(As_1, \dots, As_n)$ , dass  $\det$  multilinear und alternierend ist, und dass die Matrixmultiplikation mit  $A$  ebenfalls linear ist (Rechnung als Übung). Nach Bemerkung 7.1.2 gilt also  $D(B) = \det B \cdot D(E_n)$ . Wegen  $D(B) = \det(AB)$  und  $D(E_n) = \det(AE_n) = \det A$  ist das genau die Behauptung.  $\square$

**Folgerung.**

- (i) Ist  $A \in R^{n \times n}$  invertierbar, so ist auch  $\det A \in R$  invertierbar, und es gilt  $\det(A^{-1}) = (\det A)^{-1}$ .
- (ii) Ähnliche Matrizen haben dieselbe Determinante.
- (iii) Die Einschränkung der Abbildung  $\det$  auf  $\mathrm{GL}_n(R)$  ist ein Gruppenhomomorphismus

$$\det : (\mathrm{GL}_n(R), \cdot) \rightarrow (R^\times, \cdot).$$

Da alle Abbildungsmatrizen eines Endomorphismus  $\varphi$  ähnlich zueinander sind (vgl. 6.4.8), somit dieselbe Determinante haben, können wir diese als die *Determinante von  $\varphi$*  auffassen.

**Definition.** Für  $\varphi \in \mathrm{End}(V)$  setzen wir  $\det \varphi := \det M_{\mathcal{B}}^{\mathcal{B}}(\varphi)$ , wobei  $\mathcal{B}$  eine beliebige Basis von  $V$  ist. (Die Definition ist unabhängig von der Wahl von  $\mathcal{B}$ .)

*Frage:* Gilt auch die Umkehrung von Teil (i) der Folgerung, d.h. ist  $A$  invertierbar, wenn  $\det A$  invertierbar ist?

### 7.1.6 Cramer'sche Regel und Adjunktenformel

**Definition.** Es sei  $\tilde{A} := (\tilde{a}_{ij})$  definiert durch

$$\tilde{a}_{ij} := (-1)^{i+j} \mathrm{Minor}_{ji}(A),$$

$1 \leq i, j \leq n$ . Die Matrix  $\tilde{A}$  wird *komplementäre Matrix* oder *Adjunkte* von  $A$  genannt, auch geschrieben als  $\mathrm{adj}(A)$ .

**Satz a.** Für jedes  $A \in R^{n \times n}$  gilt  $A\tilde{A} = (\det A)E_n = \tilde{A}A$ .

*Beweis.* Siehe Vorlesung. □

**Folgerung.** Für  $A \in R^{n \times n}$  gilt:  $A$  invertierbar  $\Leftrightarrow \det A$  invertierbar. In diesem Fall ist  $\det(A)^{-1} = \det(A^{-1})$  und  $A^{-1} = \det(A)^{-1}\tilde{A}$ .

*Beweis.* Es sei  $A$  invertierbar. Dann ist  $1 = \det(E_n) = \det(A \cdot A^{-1}) = \det(A) \cdot \det(A^{-1})$ . Also ist  $\det(A)$  invertierbar und  $\det(A)^{-1} = \det(A^{-1})$ .

Es sei nun  $\det(A)$  invertierbar. Nach dem Satz a gilt  $A \cdot (\det(A)^{-1}\tilde{A}) = E_n = (\det(A)^{-1}\tilde{A}) \cdot A$ . Also ist  $A$  invertierbar und es gilt  $A^{-1} = \det(A)^{-1}\tilde{A}$ . □

**Satz b.** Cramer'sche Regel:

Es sei  $A \in \text{GL}_n(R)$  und  $b \in R^n$ . Es sei  $b \in R^n$  und  $s_1, \dots, s_n \in R^n$  bezeichne die Spalten von  $A$ . Die eindeutige Lösung  $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in R^n$  von  $Ax = b$  lautet

$$x_j := \frac{1}{\det A} \det(s_1, \dots, s_{j-1}, b, s_{j+1}, \dots, s_n).$$

*Beweis.* Da  $A$  invertierbar ist, existiert eine Lösung  $x$  von  $Ax = b$ , und es ist  $b = \sum_{i=1}^n x_i s_i$ . Aus Satz 7.1.2 a) – c) ergibt sich

$$\begin{aligned} \det(s_1, \dots, s_{j-1}, b, s_{j+1}, \dots, s_n) &= \sum_{i=1}^n x_i \det(s_1, \dots, s_{j-1}, s_i, s_{j+1}, \dots, s_n) \\ &= x_j \det A. \end{aligned}$$

□

**Bemerkung a.** Falls die Matrix  $A$  ganzzahlige Einträge hat, dann liegt ein Vorteil der Cramer'schen Regel und der Adjunktenformel darin, dass die Nenner aller im Lösungsvektor bzw. in der Inversen auftretenden Brüche bereits in dem Term  $\frac{1}{\det A}$  stecken. Die restliche Rechnung kommt ohne Brüche aus. Insbesondere sind alle Kofaktoren wieder ganzzahlig.

## 7.2 Eigenwerte und Eigenvektoren

In dem gesamten Abschnitt seien  $K$  ein Körper,  $n \in \mathbb{N}$ ,  $A \in K^{n \times n}$ ,  $V$  ein  $K$ -Vektorraum mit  $0 < \dim V = n < \infty$  und  $\varphi \in \text{End}(V)$ .

### 7.2.1 Das charakteristische Polynom

**Definition a.** Es sei  $A = (a_{ij}) \in K^{n \times n}$ . Man nennt

$$\det(X \cdot E_n - A) = \begin{vmatrix} X - a_{11} & -a_{12} & \cdots & -a_{1n} \\ -a_{21} & X - a_{22} & \cdots & -a_{2n} \\ \vdots & & \ddots & \\ -a_{n1} & \cdots & & X - a_{nn} \end{vmatrix} \in K[X]$$

das *charakteristische Polynom* von  $A$ , geschrieben  $\chi_A$ .

**Beispiel.**

$$(i) \quad A = \begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 1 & -1 & 2 \end{pmatrix} \in \mathbb{Q}^{3 \times 3}. \quad \chi_A = X^3 - 6X^2 + 11X - 6 \in \mathbb{Q}[X].$$

(Rechnung siehe Vorlesung.)

$$(ii) \quad \chi_{E_n} = (X - 1)^n = X^n - X^{n-1} + \dots + (-1)^{n-1}X + (-1)^n.$$

(Rechnung mit Kästchensatz).

**Bemerkung.**

- (i) Das charakteristische Polynom ist normiert vom Grad  $n$ , d.h. hat die Form

$$\chi_A = X^n + c_{n-1}X^{n-1} + \dots + c_1X + c_0.$$

- (ii) Es gilt  $c_{n-1} = -\text{Spur}(A)$ , wobei  $\text{Spur}(A) := a_{11} + \dots + a_{nn}$ .

- (iii) Für die Polynomfunktion zu  $\chi_A$  gilt:

$$\chi_A(\lambda) = \det(\lambda E - A) \quad \text{für alle } \lambda \in K.$$

- (iv) Es gilt  $c_0 = (-1)^n \det A$ .

*Beweis.* (i) und (ii) ergeben sich aus der Leibniz-Formel, denn  $X^n$  und  $X^{n-1}$  können in der Leibniz-Formel nur für  $\pi = \text{id}$  entstehen.

(iii) Folgt aus der Leibniz-Formel und der Tatsache, dass der Einsetzungshomomorphismus  $\tau_\lambda$  ein Ring-Homomorphismus ist. (Zur Erinnerung: Es seien  $f, g \in K[X]$ . Dann ist  $\tau_\lambda(f) := f(\lambda)$  und es gilt  $(f + g)(\lambda) = f(\lambda) + g(\lambda)$  und  $(fg)(\lambda) = f(\lambda)g(\lambda)$ .) Wir schreiben  $X \cdot E_n - A = (f_{ij})_{1 \leq i, j \leq n}$  mit  $f_{ij} \in K[X]$  für alle  $1 \leq i, j \leq n$  (tatsächlich ist  $f_{ij} = -a_{ij}$  für  $1 \leq i \neq j \leq n$  und  $f_{ii} = X - a_{ii}$  für  $1 \leq i \leq n$ ). Dann gilt:



$$\begin{aligned}
\chi_A(\lambda) &= \left( \sum_{\pi \in S_n} \operatorname{sgn}(\pi) \prod_{i=1}^n f_{\pi(i),i}(\lambda) \right) = \sum_{\pi \in S_n} \operatorname{sgn}(\pi) \left( \prod_{i=1}^n f_{\pi(i),i}(\lambda) \right) \\
&= \sum_{\pi \in S_n} \operatorname{sgn}(\pi) \prod_{i=1}^n f_{\pi(i),i}(\lambda) = |\lambda E - A|.
\end{aligned}$$

(iv) Folgt aus (iii), wenn man  $\lambda = 0$  setzt.  $\square$

**Folgerung.** *Es gelten*

(i)  $\chi_A = \chi_{A^t},$

(ii)  $A, B$  *ähnlich*  $\Rightarrow \chi_A = \chi_B.$

*Beweis.* (i) Nach Folgerung 7.1.1 gilt  $\det A = \det A^t$ . Wegen  $(XE - A)^t = XE - A^t$  gilt also  $\chi_A = \det(XE - A) = \det(XE - A^t) = \chi_{A^t}$ .

(ii) Angenommen  $A, B \in K^{n \times n}$  sind ähnlich, d.h. es gibt  $T \in \operatorname{GL}_n(K)$  mit  $B = T^{-1}AT$ . Wir zeigen zunächst, dass dann auch  $XE - B, XE - A \in K[X]^{n \times n}$  ähnlich sind:

$$\begin{aligned}
XE - B &= X(T^{-1}ET) - T^{-1}AT = T^{-1}XET - T^{-1}AT \\
&= T^{-1}(XET - AT) = T^{-1}(XE - A)T.
\end{aligned}$$

Nach Folgerung 7.1.5(ii) ist somit  $\chi_B = \det(XE - B) = \det(XE - A) = \chi_A$ .  $\square$

*Übung.* Gilt auch die Umkehrung von Teil (ii) der Folgerung?

Wegen Teil (ii) der Folgerung ist folgende Definition unabhängig von der Wahl von  $\mathcal{B}$ .

**Definition b.** Dann heißt  $\chi_\varphi := \chi_{M_{\mathcal{B}}^{\mathcal{B}}(\varphi)} \in K[X]$  das *charakteristische Polynom* von  $\varphi$ , wobei  $\mathcal{B}$  eine beliebige Basis von  $V$  ist.

**Beispiel.** Die Drehung  $\rho_\alpha$  von  $\mathbb{R}^2$  um  $\alpha$  im Uhrzeigersinn hat

$$\det \rho_\alpha = 1, \quad \chi_{\rho_\alpha} = X^2 - 2(\cos \alpha)X + 1.$$

Begründung: Nach Beispiel 6.4.2(ii) besitzt  $\rho_\alpha$  eine Abbildungsmatrix der Form  $R_\alpha = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}$ . Aus  $\det R_\alpha = \sin^2 \alpha + \cos^2 \alpha = 1$  und  $\chi_{R_\alpha} = X^2 - 2(\cos \alpha)X + 1$  folgt die Behauptung.

### 7.2.2 Eigenwerte von Endomorphismen

**Definition.** Wir definieren für jedes  $a \in K$  den Unterraum

$$V_a(\varphi) := \{v \in V \mid \varphi(v) = a \cdot v\} \leq V.$$

Wir nennen  $a$  einen *Eigenwert* von  $\varphi$ , wenn  $V_a(\varphi) \neq \{0\}$ . Ist  $a$  ein Eigenwert, so heißt  $V_a(\varphi)$  der *Eigenraum* von  $\varphi$ . Die Vektoren  $0 \neq v \in V_a(\varphi)$  heißen *Eigenvektoren* von  $\varphi$  zum *Eigenwert*  $a$ .

**Bemerkung.**

- (i) Es gilt  $V_a(\varphi) = \text{Kern}(\varphi - a \cdot \text{id}) \leq V$ .
- (ii) Ein Eigenvektor von  $\varphi$  ist ein Vektor, dessen „Richtung“ sich unter der Abbildung nicht ändert.
- (iii) Die Eigenvektoren von  $\varphi$  zum Eigenwert 1 sind genau die von 0 verschiedenen Fixpunkte von  $\varphi$ . Demnach ist 1 genau dann ein Eigenwert von  $\varphi$ , wenn  $\varphi$  Fixpunkte  $\neq 0$  hat.
- (iv) Die Eigenvektoren von  $\varphi$  zum Eigenwert 0 sind genau die von 0 verschiedenen Elemente von  $\text{Kern } \varphi$ . Demnach ist 0 ist genau dann ein Eigenwert von  $\varphi$ , wenn  $\varphi$  nicht-trivialen Kern hat.

**Beispiel.**

- (i) Die Spiegelung des  $\mathbb{R}^2$  an einer Ursprungsgeraden hat die Eigenwerte 1 und  $-1$ . Der Eigenraum zu 1 ist die Spiegelgerade, der Eigenraum zu  $-1$  ist die Ursprungsgerade senkrecht zur Spiegelgeraden.
- (ii) Die Drehung des  $\mathbb{R}^2$  um einen Winkel, der kein Vielfaches von  $180^\circ$  ist, hat keine Eigenwerte.

*Übung.* Welche Eigenwerte haben Projektion und Scherung?

### 7.2.3 Eigenwerte von Matrizen

**Definition.** Wir definieren für jedes  $a \in K$  den Unterraum

$$V_a(A) := V_a(\varphi_A) = \{x \in K^n \mid Ax = ax\} \leq K^n.$$

Wir nennen  $a$  einen *Eigenwert* von  $A$ , wenn  $V_a(A) \neq \{0\}$ . Ist  $a$  ein Eigenwert von  $A$ , so heißt  $V_a(A)$  der *Eigenraum* von  $A$  zum *Eigenwert*  $a$ . Die Vektoren  $0 \neq v \in V_a(A)$  heißen *Eigenvektoren* von  $A$  zum *Eigenwert*  $a$ .

**Bemerkung.**

- (i)  $V_a(A) = \mathbb{L}(A - aE, 0) \leq K^n$ .
- (ii) Die Eigenvektoren zu 1 sind also gerade die nicht-trivialen Lösungen der Gleichung  $Ax = x$ , und die Eigenvektoren zu 0 sind die nicht-trivialen Elemente aus  $\mathbb{L}(A, 0)$ .
- (iii)  $a$  ist genau dann ein Eigenwert von  $A$ , wenn  $\text{Rg}(A - aE) < n$  ist, bzw. wenn  $\text{Def}(A - aE) > 0$  ist, bzw. wenn  $\det(A - aE) = 0$  ist.

**Beispiel.**

- (i)  $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in \mathbb{R}^2$  hat keine Eigenwerte, weil  $A - aE = \begin{pmatrix} -a & 1 \\ -1 & -a \end{pmatrix}$  die Zeilenstufenform  $\begin{pmatrix} 1 & a \\ 0 & 1 + a^2 \end{pmatrix}$  hat, und somit für alle  $a \in \mathbb{R}$  den Rang 2 besitzt. Geometrisch interpretiert beschreibt diese Matrix eine Drehung um 90 im Uhrzeigersinn.

Für die komplexen Zahlen  $a = i$  und  $a = -i$  ist  $1 + a^2 = 0$ , also hat die Matrix  $A - aE$  den Rang 1. Über  $\mathbb{C}$  besitzt  $A$  somit die Eigenwerte  $i$  und  $-i$ .

- (ii) Eine obere Dreiecksmatrix  $A = \begin{pmatrix} a_{11} & & * \\ & \ddots & \\ 0 & & a_{nn} \end{pmatrix} \in K^{n \times n}$  hat die Eigenwerte  $a_{11}, \dots, a_{nn}$ , denn  $A - aE = \begin{pmatrix} a_{11} - a & & * \\ & \ddots & \\ 0 & & a_{nn} - a \end{pmatrix}$  hat genau dann Determinante 0, wenn  $a = a_{ii}$  ist für ein  $1 \leq i \leq n$ .

Die Methode, mittels Gauß-Algorithmus den Rang von  $A - aE$  in Abhängigkeit vom Parameter  $a$  zu berechnen, kann bei größeren Matrizen etwas umständlich werden. Eine Alternative bietet das charakteristische Polynom (siehe 7.2.5 unten).

**7.2.4 Berechnung der Eigenräume**

**Bemerkung.** Eigenräume von Endomorphismen und Matrizen hängen offensichtlich zusammen, und zwar über Koordinatenabbildungen. Ist  $\mathcal{B}$  eine geordnete Basis von  $V$ , so gilt für jedes  $\varphi \in \text{End}_K(V)$ :

$$\kappa_{\mathcal{B}}(V_a(\varphi)) = V_a(M_{\mathcal{B}}^{\mathcal{B}}(\varphi)).$$

Somit lässt sich die Berechnung von  $V_a(\varphi)$  stets auf die Berechnung von  $V_a(M_B^B(\varphi))$ , also auf den Eigenraum einer Matrix, zurückführen. Insbesondere folgt, dass  $\varphi$  und  $M_B^B(\varphi)$  dieselben Eigenwerte haben.

*Beweis.* Da  $\kappa_B$  ein Isomorphismus ist, und unter Benutzung von  $\kappa_B(\varphi(v)) = M_B^B(\varphi) \cdot \kappa_B(v)$ , gilt:

$$\begin{aligned} v \in V_a(\varphi) &\Leftrightarrow \varphi(v) = av \Leftrightarrow \kappa_B(\varphi(v)) = \kappa_B(av) \\ &\Leftrightarrow M_B^B(\varphi) \cdot \kappa_B(v) = a\kappa_B(v) \Leftrightarrow \kappa_B(v) \in V_a(M_B^B(\varphi)). \end{aligned}$$

□

Wir berechnen die Eigenräume in Beispielen, deren Eigenwerte wir schon kennen.

### Beispiel.

- (i) Es sei  $\sigma_0$  die Spiegelung von  $\mathbb{R}^2$  an der  $e_1$ -Achse, von der wir wissen, dass sie die Eigenwerte 1 und  $-1$  besitzt. Die Abbildungsmatrix von  $\sigma_0$  bzgl. der Standardbasis lautet  $A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$ . Wir berechnen  $V_{-1}(\sigma_0) = V_{-1}(A)$ . Aus

$$A - (-1)E_2 = A + E_2 = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$$

liest man  $V_{-1}(A) = \mathbb{L}(A + E_2, 0) = \langle \begin{pmatrix} 0 \\ 1 \end{pmatrix} \rangle = \langle e_2 \rangle$  ab. Wie vermutet, ergibt sich als Eigenraum zum Eigenwert 1 also die Ursprungsgerade senkrecht zur Spiegelgeraden.

- (ii)  $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in \mathbb{C}^{2 \times 2}$  hat die Eigenwerte  $i$  und  $-i$ .

$V_i(A)$ :

$$\begin{pmatrix} i & -1 \\ 1 & i \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & i \\ i & -1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & i \\ 0 & 0 \end{pmatrix}, \text{ also } V_i(A) = \langle \begin{pmatrix} i \\ -1 \end{pmatrix} \rangle.$$

$V_{-i}(A)$ :

$$\begin{pmatrix} -i & -1 \\ 1 & -i \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & -i \\ i & 1 \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & -i \\ 0 & 0 \end{pmatrix}, \text{ also } V_{-i}(A) = \langle \begin{pmatrix} i \\ 1 \end{pmatrix} \rangle.$$

- (iii) Es sei  $\sigma_\alpha$  die Spiegelung von  $\mathbb{R}^2$  an der um  $\alpha$  gedrehten  $e_1$ -Achse. Nach Beispiel 6.4.3a hat  $\sigma_\alpha$  bzgl. der Standardbasis die Matrix

$$S_\alpha = \begin{pmatrix} \cos^2 \alpha - \sin^2 \alpha & 2 \sin \alpha \cos \alpha \\ 2 \sin \alpha \cos \alpha & \sin^2 \alpha - \cos^2 \alpha \end{pmatrix}.$$

Als Übung berechne man den Eigenraum zum Eigenwert 1. (Wir wissen bereits, dass die Spiegelgerade, also die um  $\alpha$  gedrehte  $e_1$ -Achse, herauskommen muss.)

### 7.2.5 Eigenwerte als Nullstellen von $\chi$

**Satz.** Die Eigenwerte von  $\varphi$  bzw.  $A$  sind genau die Nullstellen des charakteristischen Polynoms  $\chi_\varphi$  bzw.  $\chi_A$ .

*Beweis.* Für  $A$ : Genau dann ist  $a$  Eigenwert von  $A$ , wenn  $\mathbb{L}(A - aE, 0)$  nicht-trivial ist, d.h. genau dann, wenn  $\det(A - aE) = 0$  ist, bzw.  $\det(aE - A) = 0$  ist. Nach Bemerkung 7.2.1 ist  $\det(aE - A) = \chi_A(a)$ . Somit ist (i) gezeigt.

Für  $\varphi$ : Setze  $A = M_{\mathcal{B}}^{\mathcal{B}}(\varphi)$  für eine beliebige Basis  $\mathcal{B}$ . Per Definition ist  $\chi_\varphi = \chi_A$ . Nach Bemerkung 7.2.4 haben  $\varphi$  und  $A$  dieselben Eigenwerte. Damit ist alles gezeigt.  $\square$

**Folgerung.**

- (i) Es gibt höchstens  $n$  verschiedene Eigenwerte von  $A$  und von  $\varphi$ .
- (ii)  $A$  und  $A^t$  haben dieselben Eigenwerte.
- (iii) Ähnliche Matrizen haben gleiche Eigenwerte.

*Beweis.* Das charakteristische Polynom hat Grad  $n$  und damit höchstens  $n$  verschiedene Nullstellen. Die Matrizen  $A$  und  $A^t$  haben dasselbe charakteristische Polynom. Das gleiche trifft auf ähnliche Matrizen zu.  $\square$

*Übung.* Haben  $A, A^t$  und zu  $A$  ähnliche Matrizen auch dieselben Eigenvektoren wie  $A$ ? Wenn nicht, finde man ein Gegenbeispiel.

**Beispiel.** Wir überprüfen den Satz für die bisherigen Beispiele aus 7.2.3 und 7.2.4:

- (i)  $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ ,  $\chi_A = X^2 + 1$ . Dieses Polynom hat keine Nullstellen in  $\mathbb{R}$ , zerfällt aber über  $\mathbb{C}$  in  $\chi_A = (X + i)(X - i)$ .

(ii) Die Spiegelungsmatrix  $A = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$  hat  $\chi_A = (X-1)(X+1)$ . Dieses Polynom hat die Nullstellen  $\pm 1$ .

(iii)  $A = \begin{pmatrix} a_{11} & & * \\ & \ddots & \\ 0 & & a_{nn} \end{pmatrix}$ ,  $\chi_A = (X - a_{11}) \cdots (X - a_{nn})$ . Dieses Polynom hat die Nullstellen  $a_{11}, \dots, a_{nn}$ .

(iv)  $\chi_{E_n} = (X - 1)^n$ .

(v) Als Übung berechne man das charakteristische Polynom der Matrix

$$S_\alpha = \begin{pmatrix} \cos^2 \alpha - \sin^2 \alpha & 2 \sin \alpha \cos \alpha \\ 2 \sin \alpha \cos \alpha & \sin^2 \alpha - \cos^2 \alpha \end{pmatrix}$$

aus Beispiel 7.2.4(iii). (Es muss  $(X-1)(X+1)$  herauskommen, weil  $S_\alpha$  eine Spiegelung beschreibt, und somit  $\pm 1$  die einzigen Eigenwerte sind.)

## 7.2.6 Vielfachheit von Eigenwerten

**Definition.** Es sei  $a$  ein Eigenwert von  $A$  bzw. von  $\varphi$ . Die Vielfachheit von  $a$  als Nullstelle von  $\chi_A$  bzw.  $\chi_\varphi$  wird (*algebraische*) *Vielfachheit* von  $a$  genannt, geschrieben  $m_a(A)$  bzw.  $m_a(\varphi)$ .

Die Dimension von  $V_a(A)$  bzw.  $V_a(\varphi)$  wird *geometrische Vielfachheit* von  $a$  genannt, geschrieben  $g_a(A)$  bzw.  $g_a(\varphi)$ .

**Beispiel a.** Wir betrachten die Spiegelung des  $\mathbb{R}^3$  an der  $e_1$ - $e_2$ -Ebene. Der Eigenraum zum Eigenwert 1 ist die  $e_1$ - $e_2$ -Ebene, also 2-dimensional. Somit hat der Eigenwert 1 die geometrische Vielfachheit 2.

Die Abbildungsmatrix bzgl.  $(e_1, e_2, e_3)$  lautet  $\begin{pmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ . Also ist  $\chi = (x+1)(x-1)^2$ , und die algebraische Vielfachheit von 1 ist ebenfalls 2.

Die Aussage aus Folgerung 7.2.5 gilt auch, wenn man jeden Eigenwert mit seiner Vielfachheit zählt:

**Folgerung.** Es gilt stets  $\sum_a m_a(\varphi) \leq n$  bzw.  $\sum_a m_a(A) \leq n$ , wobei die Summe über alle Eigenwerte  $a$  gebildet wird.

*Beweis.* Zählt man die Nullstellen mit ihrer Vielfachheit, so hat ein Polynom vom Grad  $n$  höchstens  $n$  Nullstellen.  $\square$

*Übung a.* Berechne alle Eigenwerte, Eigenräume und Vielfachheiten der Eigenwerte von  $A = \begin{pmatrix} -3 & 0 & 0 \\ 2 & -3 & 1 \\ 10 & 0 & 2 \end{pmatrix}$  bzw.  $B = \begin{pmatrix} -3 & 0 & 0 \\ 1 & -3 & 1 \\ 10 & 0 & 2 \end{pmatrix}$ .

**Satz.** Es gilt stets  $g_a(\varphi) \leq m_a(\varphi)$  bzw.  $g_a(A) \leq m_a(A)$ .

*Beweis.* Wir zeigen die Aussage für  $\varphi$  (die Aussage für  $A$  folgt daraus, indem man  $\varphi_A$  betrachtet). Es sei  $a$  ein Eigenwert von  $\varphi$ , und  $g = g_a(\varphi)$  sei seine geometrische Vielfachheit. Wir wählen eine geordnete Basis  $(v_1, \dots, v_g)$  von  $V_a(\varphi)$  und ergänzen diese zu einer Basis  $\mathcal{B} := (v_1, \dots, v_n)$  von  $V$ . Dann hat  $A := M_{\mathcal{B}}^{\mathcal{B}}(\varphi)$  die Form

$$A = \left( \begin{array}{ccc|c} a & & 0 & \\ & \ddots & & * \\ 0 & & a & \\ \hline & 0 & & B \end{array} \right),$$

und

$$XE - A = \left( \begin{array}{ccc|c} X - a & & 0 & \\ & \ddots & & * \\ 0 & & X - a & \\ \hline & 0 & & XE - B \end{array} \right),$$

wobei oben links jeweils ein  $g \times g$ -Block steht. Nach dem Kästchensatz folgt  $\chi_{\varphi} = \det(XE - A) = (X - a)^g \det(XE - B) = (X - a)^g \chi_B$ . Nach Definition der algebraischen Vielfachheit ist somit  $m_a(\varphi) \geq g$ .  $\square$

*Übung b.* Haben  $A, A^t$  und zu  $A$  ähnliche Matrizen dieselben geometrischen Vielfachheiten?

### 7.2.7 Spiegelungen

**Definition.** Es sei  $1 \neq -1$  in  $K$ . Ein Endomorphismus  $\varphi \in \text{End}(V)$  heißt eine *Spiegelung*, falls gelten:

- (i) 1 und  $-1$  sind Eigenwerte von  $\varphi$ , und
- (ii)  $g_1(\varphi) = n - 1$ .

( $\varphi$  ist Spiegelung an  $V_1(\varphi)$ , der sogenannten *Spiegelungshyperebene*.)

**Bemerkung.** Es sei  $\varphi$  eine Spiegelung. Wähle  $0 \neq v_1 \in V_{-1}(\varphi)$  und wähle eine geordnete Basis  $(v_2, \dots, v_n)$  von  $V_1(\varphi)$ . Wegen  $\varphi(v_1) = -v_1 \neq v_1$  ist  $v_1 \notin V(1, \varphi)$ . Daraus folgt, dass  $\mathcal{B} := (v_1, \dots, v_n)$  linear unabhängig, wegen  $\dim V = n$  also sogar Basis von  $V$  ist. Da alle Basisvektoren Eigenvektoren sind, ist  $M_{\mathcal{B}}^{\mathcal{B}}(\varphi)$  eine Diagonalmatrix, nämlich

$$M_{\mathcal{B}}^{\mathcal{B}}(\varphi) = \begin{pmatrix} -1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{pmatrix}.$$

## 7.3 Der PageRank-Algorithmus

### 7.3.1 Einleitung und Idee

Gegeben seien  $n$  Webseiten  $S_1, \dots, S_n$ , die sich untereinander verlinken, wobei wir keine Links einer Seite auf sich selbst zulassen. Wir veranschaulichen die Situation durch einen gerichteten Graphen ohne Schleifen mit Knotenmenge  $\{S_1, \dots, S_n\}$  und Kantenmenge  $\{S_j \rightarrow S_i \mid S_j \text{ verlinkt auf } S_i\}$ . Es sei  $n_j$  die Anzahl der Kanten, die von  $S_j$  ausgehen. Wir nehmen an, dass  $n_j \geq 1$  ist für jedes  $j$  (Bemerkung 7.3.4 unten erklärt, wie man auf diese Voraussetzung verzichten kann). Definiere die *Link-Matrix*  $L = (l_{ij})_{ij} \in \mathbb{R}^{n \times n}$  durch

$$l_{ij} := \begin{cases} \frac{1}{n_j} & \text{falls } S_j \text{ auf } S_i \text{ verlinkt und } i \neq j, \\ 0 & \text{sonst.} \end{cases}$$

Die  $j$ -te Spalte von  $L$  enthält die Links, die von der Seite  $S_j$  ausgehen, und die Spaltensumme ist  $\sum_{i=1}^n l_{ij} = n_j \cdot \frac{1}{n_j} = 1$  für alle  $j$ . In der Praxis ist  $L$  *dünn besetzt*, d.h. enthält viele Nullen.

**Idee.** Die Seiten stimmen selbst über ihre Wichtigkeit ab.

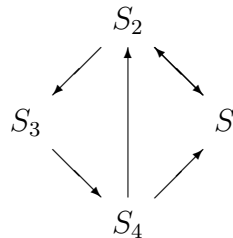
1. Ansatz: Jede Seite hat eine Stimme, die sie gleichmäßig auf diejenigen Seiten verteilt, die von ihr verlinkt werden. Das *Gewicht* (= Wichtigkeit) der Seite  $S_i$  ergibt sich dann als die Zeilensumme  $x_i := \sum_{j=1}^n l_{ij}$ . Problem: “Unwichtige Seiten”, die sich gegenseitig verlinken, werden wichtig.

2. Ansatz: Wie 1., aber jede Seite hat genau so viele Stimmen, wie ihrem Gewicht entspricht. Das Gewicht von  $S_i$  lautet dann  $x_i := \sum_{j=1}^n l_{ij} x_j$ . Der

*Gewichtsvektor*  $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$  erfüllt also die Bedingung  $x = Lx$ , ist also Eigenvektor von  $L$  zum Eigenwert 1.



**Beispiel.**



$$L = \begin{pmatrix} 0 & \frac{1}{2} & 0 & \frac{1}{2} \\ 1 & 0 & 0 & \frac{1}{2} \\ 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix} \quad x = \begin{pmatrix} 3 \\ 4 \\ 2 \\ 2 \end{pmatrix} = Lx$$

**Frage.** Gibt es immer einen Eigenvektor zum Eigenwert 1? Gibt es einen Eigenvektor mit Einträgen  $\geq 0$ ? Ist er eindeutig?

### 7.3.2 Stochastische Matrizen

**Definition.** Eine reelle Matrix  $M \in \mathbb{R}^{m \times n}$  heißt *positiv* (bzw. *negativ*, *nicht-negativ*, *nicht-positiv*), geschrieben  $M > 0$  (bzw.  $M < 0$ ,  $M \geq 0$ ,  $M \leq 0$ ), wenn alle Einträge von  $M > 0$  (bzw.  $< 0$ ,  $\geq 0$ ,  $\leq 0$ ) sind. Wir definieren  $l(M)$  als die Summe aller Einträge von  $M$ .

Eine quadratische nicht-negative reelle Matrix  $M$  heißt *stochastische Matrix*, wenn  $l(s) = 1$  ist für jede Spalte  $s$  von  $M$ .

**Satz.** Jede stochastische Matrix  $M \in \mathbb{R}^{n \times n}$  hat einen nicht-negativen Eigenvektor zum Eigenwert 1.

*Beweis.* Die Matrix  $M^t$  hat Zeilensummen gleich 1, also ist  $\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$  ein Eigenvektor zum Eigenwert 1. Nach Folgerung 7.2.5 hat damit auch  $M$  den Eigenwert 1.

Sei nun  $M = (a_{ij})_{ij}$  und  $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$  ein Eigenvektor von  $A$  zum

Eigenwert 1. Wegen  $x \neq 0$  ist mindestens ein  $x_i \neq 0$ . Wir nehmen an, dass ein  $x_i > 0$  ist (sonst ersetze  $x$  durch  $-x$ ).

Durch Permutieren der Basisvektoren der Standardbasis können wir erreichen:  $x_1, \dots, x_r \leq 0$  und  $x_{r+1}, \dots, x_n > 0$  für ein  $0 \leq r \leq n-1$ . (Eine Permutation der Standardbasis ist ein Basiswechsel, der auf  $M$  die Anwendung derselben Permutation auf die Spalten und ihrer Inversen auf die Zeilen bewirkt. So entsteht wieder eine stochastische Matrix.) Wir zerlegen nun  $M$  und  $x$  in Blöcke  $M = \left( \begin{array}{c|c} A & B \\ \hline C & D \end{array} \right)$ ,  $x = \begin{pmatrix} y \\ z \end{pmatrix}$  derart, dass  $A$  quadratisch ist

und  $A$  und  $y$  genau  $r$  Zeilen haben. Dann sind  $A, B, C, D \geq 0, y \leq 0$  und  $z > 0$ . Aus  $Mx = x$  folgt  $Ay + Bz = y$ , also  $l(y) = l(Ay + Bz) = l(Ay) + l(Bz)$ . Offensichtlich gilt

$$l\left(\left(\frac{A}{C}\right)y\right) = l\left(\left(\frac{Ay}{Cy}\right)\right) = l(Ay) + l(Cy).$$

Wegen  $\sum_{i=1}^n a_{ij} = 1$  für alle  $j = 1, \dots, n$  ist andererseits

$$l\left(\left(\frac{A}{C}\right)y\right) = \sum_{i=1}^n \left(\sum_{j=1}^r a_{ij} x_j\right) = \sum_{j=1}^r \left(\sum_{i=1}^n a_{ij}\right) x_j = \sum_{j=1}^r x_j = l(y).$$

Durch Gleichsetzen von  $l(y)$  folgt  $l(Cy) = l(Bz)$ . Wegen  $Cy \leq 0$  und  $Bz \geq 0$  ist das nur möglich, wenn  $Cy = 0$  und  $Bz = 0$  ist. Man rechnet leicht nach, dass dann mit  $\begin{pmatrix} y \\ z \end{pmatrix}$  auch  $\begin{pmatrix} -y \\ z \end{pmatrix} \geq 0$  ein Eigenvektor zum Eigenwert 1 ist.  $\square$

**Beispiel.** Die Matrix  $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$  zeigt, dass der nicht-negative Eigenvektor  $x$  zum Eigenwert 1 nicht eindeutig sein muss, denn  $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$  und  $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$  haben beide diese Eigenschaft.

Das Phänomen tritt z.B. auf, wenn der “Link-Graph” nicht zusammenhängend ist. Dann hat die Link-Matrix (bei geeigneter Nummerierung der Seiten) eine Blockstruktur der Form

$$\begin{pmatrix} L_1 & & 0 \\ & \ddots & \\ 0 & & L_k \end{pmatrix}.$$

### 7.3.3 Markov-Prozesse

Es sei  $M \in \mathbb{R}^{n \times n}$  und  $v \in \mathbb{R}^n$ .

**Bemerkung.** Es sei  $M$  stochastisch. Dann gilt für alle  $i \geq 0$ :

$$(i) \quad v \geq 0 \Rightarrow M^i v \geq 0.$$

$$(ii) \quad l(M^i v) = l(v).$$

*Beweis.* Als Übung (die Argumente sind dieselben wie im Beweis von Satz 7.3.2).  $\square$

**Satz.** Konvergiert die Folge  $v, Mv, M^2v, \dots$  in  $\mathbb{R}^n$  gegen  $x$ , so ist  $x$  ein Eigenvektor von  $M$  zum Eigenwert 1.

*Beweisskizze.* Mit Konvergenz in  $\mathbb{R}^n$  ist “komponentenweise Konvergenz” gemeint (das wird in der mehrdimensionalen Analysis genau definiert). Konvergiert  $M^i v \rightarrow x$ , so auch die Teilfolge  $M^{i+1} v \rightarrow x$ . Andererseits konvergiert  $M^{i+1} v = M(M^i v) \rightarrow Mx$  (hier wird benötigt, dass  $v \mapsto Mv$  eine stetige Abbildung ist). Folglich  $Mx = x$ , d.h.  $x$  ist Eigenvektor von  $M$  zum Eigenwert 1.  $\square$

**Definition.** Sind  $M \in \mathbb{R}^{n \times n}$  und  $v \in \mathbb{R}^n$  stochastisch, so wird die Folge  $v, Mv, M^2v, \dots$  in  $\mathbb{R}^n$  Markov-Prozess mit Anfangswert  $v$  genannt.

**Folgerung.** Konvergiert der Markov-Prozess  $v, Mv, M^2v, \dots \rightarrow x$ , so ist  $x$  ein nicht-negativer Eigenvektor von  $M$  zum Eigenwert 1 mit  $l(x) = 1$ .

*Beweisskizze.* Aufgrund des Satzes ist  $x$  Eigenvektor von  $M$  zum Eigenwert 1. Die Aussagen  $x \geq 0$  und  $l(x) = 1$  folgen aus der Tatsache, dass  $M^i v \geq 0$  und  $l(M^i v) = 1$  für alle  $i \geq 0$  ist (siehe Bemerkung).  $\square$

**Beispiel a.** Es sei  $L$  die Link-Matrix aus Beispiel 7.3.1. Dann ist

$$L^i v \xrightarrow{i \rightarrow \infty} \frac{1}{11} \begin{pmatrix} 3 \\ 4 \\ 2 \\ 2 \end{pmatrix} \text{ z.B. für } v = \begin{pmatrix} \frac{1}{4} \\ \frac{1}{4} \\ \frac{1}{4} \\ \frac{1}{4} \end{pmatrix}.$$

Die Interpretation ist die eines “Zufallssurfer”, der sich zu Beginn mit Wahrscheinlichkeit  $v_i$  auf der Seite  $S_i$  aufhält, und dann in jedem Schritt zufällig mit gleicher Wahrscheinlichkeit einen beliebigen Link auf  $S_i$  verfolgt. Nach gewisser Zeit hält er sich mit Wahrscheinlichkeit  $x_i$  auf der Seite  $S_i$  auf.

Es gilt sogar

$$L^i \xrightarrow{i \rightarrow \infty} \frac{1}{11} \begin{pmatrix} 3 & 3 & 3 & 3 \\ 4 & 4 & 4 & 4 \\ 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 2 \end{pmatrix}, \text{ also } L^i v \xrightarrow{i \rightarrow \infty} \frac{1}{11} \begin{pmatrix} 3 \\ 4 \\ 2 \\ 2 \end{pmatrix}$$

für alle Anfangswerte  $v$ .

*Übung.* Was passiert in dem Modell des Zufallssurfers, wenn man zulässt, dass eine Webseite keine herausführenden Links besitzt ( $n_j = 0$ )?

**Beispiel b.** Angenommen zu  $n_1$  vorhandenen Webseiten erzeugen wir  $n_2$  neue Webseiten, die sich gegenseitig verlinken, auf die aber sonst kein Link führt. Insgesamt gibt es also  $n = n_1 + n_2$  Webseiten. In der Praxis ist die Anzahl der neuen Webseiten immer klein im Verhältnis zur Gesamtzahl, d.h.  $\frac{n_1}{n} \approx 1$  und  $\frac{n_2}{n} \approx 0$ . Der “Link-Graph” zerfällt dann in zwei Zusammenhangskomponenten mit  $n_1$  bzw.  $n_2$  vielen Knoten. Entsprechend hat die Link-Matrix  $L$  die Blockform  $\begin{pmatrix} L_1 & 0 \\ 0 & L_2 \end{pmatrix}$ , wobei  $L_1$  ein  $n_1 \times n_1$ -Block ist und  $L_2$  der  $n_2 \times n_2$ -Block

$$L_2 = \frac{1}{n_2 - 1} \begin{pmatrix} 0 & 1 & \cdots & 1 \\ 1 & 0 & & 1 \\ \vdots & & & \vdots \\ 1 & \cdots & 1 & 0 \end{pmatrix}.$$

Sowohl  $L_1$  als auch  $L_2$  ist wieder eine stochastische Matrix. Aufgrund der Blockstruktur ist der Eigenvektor zu 1 nicht eindeutig. Z.B. gibt es einen Eigenvektor zu 1 der Form  $x = \begin{pmatrix} 0 \\ x_2 \end{pmatrix}$  mit  $x_2$  Eigenvektor von  $L_2$ . Als Gewichtsvektor interpretiert würde dieser nur den neuen Webseiten Bedeutung zumessen.

Das kann nicht passieren, wenn man den Markov-Prozess mit “gleichverteiltem” Anfangsvektor  $v = \frac{1}{n} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$  bildet. Angenommen  $v, Lv, L^2v, \dots$

konvergiert gegen  $x \in \mathbb{R}^n$ . Zerlegt man  $v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$  und  $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$  mit

$v_1, x_1 \in \mathbb{R}^{n_1}, v_2, x_2 \in \mathbb{R}^{n_2}$ , so ist  $v_2 = \frac{1}{n} \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$  und  $v_2, L_2v_2, L_2^2v_2, \dots$  konver-

giert gegen  $x_2$ . Da  $L_2$  eine stochastische Matrix ist, folgt nach der Bemerkung  $l(x_2) = l(v_2) = n_2/n \approx 0$ . D.h. die neuen Webseiten sind selbst in ihrer Gesamtheit gemäß des Gewichtsvektors  $x$  unbedeutend.

### 7.3.4 Positive stochastische Matrizen

**Satz.** Ist  $M$  eine positive stochastische Matrix  $M$ , so ist

- (i) jeder Eigenvektor zu 1 entweder positiv oder negativ,
- (ii) der Eigenraum zu 1 ein-dimensional,

(iii) der Eigenvektor  $x$  zu 1 mit  $l(x) = 1$  eindeutig bestimmt und positiv.

*Beweis.* Wir führen den Beweis zunächst genau wie im Satz 7.3.2 bis zu der Stelle  $Bz = 0$ . Wegen  $z > 0$  folgt daraus  $B = 0$ . Eine positive Matrix enthält keine Nullmatrix als echte Teilmatrix, daher ist  $r = 0$ . Das bedeutet  $x_1, \dots, x_n > 0$  bzw.  $x > 0$ . Wir haben damit (i) gezeigt (denn  $x$  wurde als beliebiger Eigenvektor zu 1 angenommen und dann entweder  $x$  oder  $-x$  betrachtet). Die untenstehende Übung zeigt, dass daraus (ii) folgt. Somit gibt es genau einen Eigenvektor  $x$  zu 1 mit  $l(x) = 1$ . Nach (i) ist  $x > 0$ .  $\square$

**Beispiel.** Die Matrix  $\begin{pmatrix} 1 & 0 & \frac{1}{3} \\ 0 & 1 & \frac{1}{3} \\ 0 & 0 & \frac{1}{3} \end{pmatrix}$  ist eine stochastische Matrix, die keinen positiven Eigenvektor zum Eigenwert 1 besitzt.

*Übung.* Jeder 2-dimensionale Unterraum von  $\mathbb{R}^n$  enthält Vektoren die weder positiv noch negativ sind.

**Bemerkung.** Ohne Beweis sei folgende Tatsache erwähnt: Ist  $M$  eine positive stochastische Matrix, so konvergiert der Markov-Prozess  $M^i v$  für jeden stochastischen Anfangswert  $v$ . Aus Folgerung 7.3.3 ergibt sich, dass der Grenzwert  $x$  dann der eindeutige Vektor aus Teil (iii) des obigen Satzes ist. Da dies auch für die Anfangswerte  $e_i$  gilt, konvergiert  $M^i$  folglich gegen die Matrix, deren Spalten alle identisch  $x$  sind (vgl. Beispiel 7.3.3a).

In der Anwendung ist es sinnvoll, ein kleines  $\alpha > 0$  zu wählen und in der Link-Matrix zu allen Einträgen  $\alpha/n$  zu addieren (und danach die Spalten wieder auf Spaltensumme 1 zu "normieren"). Die Interpretation von  $\alpha$  ist die Wahrscheinlichkeit, dass der Zufallssurfer keinem vorhandenen Link folgt, sondern auf eine beliebige andere Seite wechselt. Die Link-Matrix ist dann stets positiv (und der Link-Graph zusammenhängend). Dadurch konvergiert der Markov-Prozess zu einem eindeutigen Grenzwert unabhängig vom Anfangswert. Außerdem kann man auf die Voraussetzung  $n_j > 0$  für alle  $j$  verzichten.

## 7.4 Diagonalisierbarkeit und Trigonalisierbarkeit

In dem gesamten Abschnitt seien  $K$  ein Körper,  $n \in \mathbb{N}$ ,  $A \in K^{n \times n}$ ,  $V$  ein  $K$ -Vektorraum mit  $0 < \dim V = n < \infty$  und  $\varphi \in \text{End}(V)$ .

Zusammenhang:

- (i)  $1 \times 1$ -Blöcke  $\leftrightarrow$  Eigenvektoren

- (ii) Diagonalmatrizen  $\leftrightarrow$  Basis aus Eigenvektoren

*Frage:* Wann besitzt  $V$  eine Basis aus Eigenvektoren von  $\varphi$ ?

### 7.4.1 Diagonalisierbare Endomorphismen und Matrizen

**Bemerkung a.** Eine Basis  $\mathcal{B}$  von  $V$  besteht genau dann aus Eigenvektoren von  $\varphi$ , wenn  $M_{\mathcal{B}}(\varphi)$  eine Diagonalmatrix ist.

**Definition.**

- (i)  $\varphi$  heißt *diagonalisierbar*, wenn eine Basis von  $V$  existiert, die aus Eigenvektoren von  $\varphi$  besteht.
- (ii)  $A$  heißt *diagonalisierbar*, wenn  $A$  ähnlich zu einer Diagonalmatrix ist. Ist  $T \in \mathrm{GL}_n(K)$  und  $T^{-1}AT$  eine Diagonalmatrix, so sagt man  $A$  wird durch  $T$  diagonalisiert.

**Satz.** Für jede Basis  $\mathcal{B}$  von  $V$  gilt:

$$\varphi \text{ diagonalisierbar} \Leftrightarrow M_{\mathcal{B}}(\varphi) \text{ diagonalisierbar.}$$

*Beweis.*  $\Rightarrow$ : Sei  $\varphi$  diagonalisierbar. Wähle eine Basis  $\mathcal{C}$  von  $V$  aus Eigenvektoren von  $\varphi$  und sei  $T := M_{\mathcal{B}}^{\mathcal{C}}(\mathrm{id}_V)$  die Basiswechselmatrix. Dann ist  $T^{-1}M_{\mathcal{B}}(\varphi)T = M_{\mathcal{C}}(\varphi)$  eine Diagonalmatrix, also ist  $M_{\mathcal{B}}(\varphi)$  diagonalisierbar.

$\Leftarrow$ : Sei  $M_{\mathcal{B}}(\varphi)$  durch  $T \in \mathrm{GL}_n(K)$  diagonalisierbar, d.h.  $T^{-1}M_{\mathcal{B}}(\varphi)T$  sei eine Diagonalmatrix. Nach Folgerung 6.4.4 gibt es eine Basis  $\mathcal{C}$  von  $V$  mit  $M_{\mathcal{B}}^{\mathcal{C}}(\mathrm{id}_V) = T$ . Dann ist  $M_{\mathcal{C}}(\varphi) = T^{-1}M_{\mathcal{B}}(\varphi)T$  eine Diagonalmatrix, also ist  $\varphi$  diagonalisierbar.  $\square$

**Bemerkung b.**  $A$  ist genau dann diagonalisierbar, wenn  $\varphi_A$  diagonalisierbar ist, also genau dann, wenn eine Basis von  $K^n$  aus Eigenvektoren von  $A$  existiert.

Ist  $\mathcal{C} = (v_1, \dots, v_n)$  eine solche Basis aus Eigenvektoren von  $A$ , so wird  $A$  durch  $T := M_{\mathcal{E}}^{\mathcal{C}}(\mathrm{id}_{K^n})$  diagonalisiert, wobei  $\mathcal{E}$  die Standardbasis von  $K^n$  bezeichnet (die Spalten von  $T$  lauten  $v_1, \dots, v_n$ ). Genauer:

$$T^{-1}AT = \begin{pmatrix} a_1 & & \\ & \ddots & \\ & & a_n \end{pmatrix},$$

wobei  $v_i$  Eigenvektor von  $A$  zum Eigenwert  $a_i$  ist,  $1 \leq i \leq n$ .

*Beweis.* Dies folgt aus dem Satz, indem man  $\varphi = \varphi_A$  und  $\mathcal{B} = \mathcal{E}$  wählt. Da  $v_i$  Eigenvektor von  $A$  zum Eigenwert  $a_i$  ist, hat  $M_{\mathcal{C}}(\varphi_A) = T^{-1}AT$  die besagte Diagonalform.  $\square$

**Beispiel a.** Ist  $A = \begin{pmatrix} -3 & 0 & 0 \\ 2 & -3 & 1 \\ 10 & 0 & 2 \end{pmatrix} \in \mathbb{Q}^{3 \times 3}$  diagonalisierbar?

Aus Beispiel 7.2.6 sind die Eigenvektoren  $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$  und  $\begin{pmatrix} -1 \\ 0 \\ 2 \end{pmatrix}$  zum Eigenwert  $-3$  und  $\begin{pmatrix} 0 \\ 1 \\ 5 \end{pmatrix}$  zum Eigenwert  $2$  bekannt. Da diese drei Vektoren linear unabhängig sind, also eine Basis bilden, wird  $A$  durch die Matrix  $T = \begin{pmatrix} 0 & 0 & -1 \\ 1 & 1 & 0 \\ 5 & 0 & 2 \end{pmatrix}$  diagonalisiert:  $T^{-1}AT = \begin{pmatrix} 2 & & \\ & -3 & \\ & & -3 \end{pmatrix}$ .

**Beispiel b.**

(i)  $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in K^{2 \times 2}$  ist nicht diagonalisierbar.

(ii)  $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in K^{2 \times 2}$  ist nicht diagonalisierbar für  $K = \mathbb{R}$ , aber diagonalisierbar für  $K = \mathbb{C}$ .

Anschaulich beschreibt (i) eine *Scherung* und (ii) eine Drehung um  $90^\circ$ . Für beides gibt es über  $\mathbb{R}$  keine Basis aus Eigenvektoren.

*Beweis.* (i) Wegen  $\chi_A = (X - 1)^2$  lautet der einzig mögliche Eigenwert  $1$ . Wäre  $A$  diagonalisierbar, so gäbe es ein  $T \in \text{GL}_2(K)$ , mit  $T^{-1}AT = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = E_2$ . Daraus folgt aber der Widerspruch  $A = E_2$ .

(ii) Nach Beispiel 7.2.4 hat  $A$  überhaupt keine Eigenvektoren über  $\mathbb{R}$ , also kann auch keine Basis aus Eigenvektoren existieren. Über  $\mathbb{C}$  existieren aber die beiden linear unabhängigen Eigenvektoren  $\begin{pmatrix} i \\ -1 \end{pmatrix}$  zu  $i$  und  $\begin{pmatrix} i \\ 1 \end{pmatrix}$  zu  $-i$ . Somit gilt  $T^{-1}AT = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}$  für  $T = \begin{pmatrix} i & i \\ -1 & 1 \end{pmatrix}$ .  $\square$

### 7.4.2 Kriterien

Wir beschränken uns auf Matrizen; für Endomorphismen geht alles analog.

*Erinnerung an die Vielfachheiten von Eigenwerten.*

Für jeden Eigenwert  $a$  von  $A$  gilt:

$$1 \leq g_a(A) \leq m_a(A) \leq n.$$

Ist  $l$  die Anzahl der verschiedenen Eigenwerte von  $A$ , so folgt:

$$l \leq \sum_a g_a(A) \leq \sum_a m_a(A) \leq n, \quad (7.1)$$

wobei die Summation über die verschiedenen Eigenwerte von  $A$  läuft.

**Satz.** *Folgende Aussagen sind äquivalent:*

- (i)  $A$  ist diagonalisierbar.
- (ii)  $\sum_a g_a(A) = n$ .
- (iii)  $\chi_A$  zerfällt vollständig in Linearfaktoren und für jeden Eigenwert  $a$  von  $A$  ist  $g_a(A) = m_a(A)$ .

*Beweis.* (ii)  $\Leftrightarrow$  (iii): Wegen (7.1) ist (ii) äquivalent zu  $\sum_a g_a(A) = \sum_a m_a(A)$  und  $\sum_a m_a(A) = n$ . Ersteres ist äquivalent dazu, dass  $g_a(A) = m_a(A)$  für jeden Eigenwert  $a$  von  $A$  gilt. Letzteres ist äquivalent dazu, dass  $\chi_A$  vollständig in Linearfaktoren zerfällt.

(i)  $\Rightarrow$  (ii): Sei  $A$  diagonalisierbar. Die Basisvektoren einer Basis aus Eigenvektoren stammen aus den Eigenräumen. Der Eigenraum zum Eigenwert  $a$  liefert höchstens  $g_a(A)$  linear unabhängige Vektoren. Damit folgt  $n \leq \sum_a g_a(A)$  und wegen (7.1) auch die Gleichheit.

(ii)  $\Rightarrow$  (i): Sei  $\sum_a g_a(A) = n$ . Wähle zu jedem Eigenwert  $a$  eine Basis  $B_a$  des Eigenraums  $V_a(A)$ . Dann ist  $|B_a| = g_a(A)$ . Nach folgendem Lemma ist  $B := \cup_a B_a$  linear unabhängig und  $|B| = \sum_a g_a(A) = n$ , also  $B$  eine Basis aus Eigenvektoren von  $A$ .  $\square$

*Übung.*

- (i) Eigenräume zu paarweise verschiedenen Eigenwerten sind *disjunkt*, d.h. der Schnitt ist gleich  $\{0\}$ .
- (ii) Eigenvektoren zu paarweise verschiedenen Eigenwerten sind linear unabhängig.



**Lemma.** Es seien  $a_1, \dots, a_l$  paarweise verschiedene Eigenwerte von  $A$ . Seien  $B_i \subseteq V_{a_i}(A)$  linear unabhängig,  $i = 1, \dots, l$ . Dann sind die  $B_i$  paarweise disjunkt und  $B_1 \cup \dots \cup B_l$  ist linear unabhängig.

*Beweis.* Nach Teil (i) der Übung sind die  $B_i$  paarweise disjunkt. Angenommen  $\sum_{j=1}^m \lambda_j v_j = 0$  ist eine lineare Abhängigkeit in  $B$ , d.h.  $v_1, \dots, v_m \in B$  paarweise verschieden,  $m \geq 1$ , und  $\lambda_1, \dots, \lambda_m \in K \setminus \{0\}$ . Durch Zusammenfassen der Summanden aus jeweils demselben Eigenraum  $V_{a_i}(A)$  bekommen wir eine Summe  $w_1 + \dots + w_l = 0$  mit  $w_i \in V_{a_i}(A)$ . Es sind nicht alle  $w_i = 0$ , denn sonst wäre für dasjenige  $i_0$  mit  $v_1 \in B_{i_0}$  die Gleichungen  $w_{i_0} = 0$  eine lineare Abhängigkeit in  $B_{i_0}$ . Somit ist  $\{w_1, \dots, w_l\}$  linear abhängig, im Widerspruch zu Teil (ii) der Übung. Also ist die Annahme falsch, d.h.  $B$  ist linear unabhängig.  $\square$

**Beispiel.** Wir betrachten die Matrizen aus den Beispielen 7.4.1 erneut.

- (i)  $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in K^{2 \times 2}$  ist nicht diagonalisierbar, obwohl  $\chi_A = (X - 1)^2$  vollständig zerfällt. Man rechnet leicht nach, dass  $g_1(A) = 1 < 2 = m_1(A)$  ist.

Die Tatsache, dass  $\chi_A$  vollständig in Linearfaktoren zerfällt, ist also allein nicht hinreichend für die Diagonalisierbarkeit von  $A$ .

- (ii)  $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \in \mathbb{R}^{2 \times 2}$  ist nicht diagonalisierbar, weil  $\chi_A = X^2 + 1$  nicht vollständig zerfällt.

- (iii) Für  $A = \begin{pmatrix} -3 & 0 & 0 \\ 2 & -3 & 1 \\ 10 & 0 & 2 \end{pmatrix}$  und  $B = \begin{pmatrix} -3 & 0 & 0 \\ 1 & -3 & 1 \\ 10 & 0 & 2 \end{pmatrix}$  gilt  $\chi_A = \chi_B = (X - 2)(X + 3)^2$ . Nach Übung 7.2.6a ist  $g_2(A) + g_{-3}(A) = 1 + 2 = 3$  und  $g_2(B) + g_{-3}(B) = 1 + 1 = 2$ . Somit ist  $A$  diagonalisierbar und  $B$  nicht.

Insbesondere sind  $A$  und  $B$  nicht ähnlich, haben aber gleiches charakteristisches Polynom.

### 7.4.3 Ein hinreichendes Kriterium

**Folgerung.** Wenn  $\chi_A$  vollständig in paarweise verschiedene Linearfaktoren zerfällt, dann ist  $A$  diagonalisierbar.

*Beweis.* In diesem Fall ist  $m_a(A) = 1$  für alle Eigenwerte  $a$  von  $A$ , also offensichtlich  $g_a(A) = m_a(A)$ . Die Aussage ergibt sich also aus Satz 7.4.2.  $\square$

**Beispiel a.**  $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 2 & 2 & -1 \end{pmatrix}$  hat  $\chi_A = (X^2 - 2)(X + 1)$ . Über  $\mathbb{Q}$  zerfällt

dieses Polynom nicht vollständig in Linearfaktoren ( $\sqrt{2} \notin \mathbb{Q}$ ), somit ist  $A$  nicht diagonalisierbar über  $\mathbb{Q}$ .

Über  $\mathbb{R}$  dagegen zerfällt das Polynom vollständig in paarweise verschiedene Linearfaktoren ( $\chi_A = (X - \sqrt{2})(X + \sqrt{2})(X + 1)$ ), somit ist  $A$  diagonalisierbar über  $\mathbb{R}$ .

**Beispiel b.**  $A = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix} \in K^{2 \times 2}$  hat  $\chi_A = (X - 1)(X + 1)$ .

Ist  $1 \neq -1$  (z.B. für  $K = \mathbb{Q}, \mathbb{R}, \mathbb{C}$  oder  $\mathbb{F}_p$  mit  $p \neq 2$ ), so zerfällt also  $\chi_A$  in paarweise verschiedene Linearfaktoren, also ist  $A$  diagonalisierbar.

Ist dagegen  $1 = -1$  (z.B. in  $K = \mathbb{F}_2$ ), so ist das nicht der Fall. Die Matrix ist auch nicht diagonalisierbar, wie bereits in Beispiel 7.4.1b gezeigt wurde.

#### 7.4.4 Trigonalisierbarkeit

**Definition.**

(i)  $\varphi$  heißt *trigonalisierbar*, wenn eine Basis von  $V$  existiert, bzgl. der die Abbildungsmatrix von  $\varphi$  eine obere Dreiecksmatrix ist.

(ii)  $A$  heißt *trigonalisierbar*, wenn  $A$  ähnlich zu einer oberen Dreiecksmatrix ist.

Ist  $T \in \text{GL}_n(K)$  so dass  $T^{-1}AT$  eine obere Dreiecksmatrix ist, so sagt man  $A$  wird durch  $T$  *trigonalisiert*.

**Bemerkung.** Für jede beliebige Basis  $\mathcal{B}$  von  $V$  gilt:

$$\varphi \text{ trigonalisierbar} \Leftrightarrow M_{\mathcal{B}}(\varphi) \text{ trigonalisierbar.}$$

**Satz.**  $A$  ist genau dann trigonalisierbar, wenn  $\chi_A$  vollständig in Linearfaktoren zerfällt.

*Beweis.* Sei  $A$  trigonalisierbar, also ähnlich zu einer oberen Dreiecksmatrix  $D$  mit den Diagonaleinträgen  $a_1, \dots, a_n$ . Laut Kästchensatz ist  $\chi_A = \chi_D = (X - a_1) \cdots (X - a_n)$ .

Wir zeigen die Umkehrung mittels Induktion nach  $n$ . Der Induktionsanfang  $n = 1$  ist trivial. Sei nun  $n > 1$  und die Aussage für  $n - 1$  bereits bewiesen. Es zerfalle  $\chi_A$  vollständig, etwa  $\chi_A = (X - a_1) \cdots (X - a_n)$ . Wegen  $\chi_A(a_1) = 0$  ist  $a_1$  ein Eigenwert von  $A$ . Wähle einen Eigenvektor

$v_1 \in K^n$  von  $A$  zum Eigenwert  $a_1$  und ergänze diesen zu einer geordneten Basis  $\mathcal{B} = (v_1, \dots, v_n)$  von  $K^n$ . Dann hat  $M_{\mathcal{B}}(\varphi_A)$  die Form  $\left(\begin{array}{c|c} a_1 & * \\ \hline 0 & D \end{array}\right) \in K^{n \times n}$  mit  $D \in K^{(n-1) \times (n-1)}$ . Also gilt mit dem Kästchensatz:

$$(X - a_1)(X - a_2) \cdots (X - a_n) = \chi_A = \chi_{M_{\mathcal{B}}(\varphi_A)} = (X - a_1)\chi_D.$$

Mit der Kürzungsregel (im nullteilerfreien Polynomring  $K[X]$ ) folgt  $\chi_D = (X - a_2) \cdots (X - a_n)$ , d.h.  $\chi_D$  zerfällt vollständig in Linearfaktoren. Nach Induktionsvoraussetzung gibt es  $S \in \mathrm{GL}_{n-1}(K)$  so, dass  $S^{-1}DS$  eine obere Dreiecksmatrix ist. Setze  $T := \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & S \end{array}\right) \in K^{n \times n}$ . Dann ist  $\det T = \det S \neq 0$  (Kästchensatz), d.h.  $T \in \mathrm{GL}_n(K)$ . Weiter ist  $T^{-1} = \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & S^{-1} \end{array}\right)$  und

$$\begin{aligned} T^{-1}M_{\mathcal{B}}(\varphi_A)T &= \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & S^{-1} \end{array}\right) \left(\begin{array}{c|c} a_1 & * \\ \hline 0 & D \end{array}\right) \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & S \end{array}\right) \\ &= \left(\begin{array}{c|c} a_1 & * \\ \hline 0 & SD \end{array}\right) \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & S \end{array}\right) = \left(\begin{array}{c|c} a_1 & * \\ \hline 0 & S^{-1}DS \end{array}\right) \end{aligned}$$

Also ist  $M_{\mathcal{B}}(\varphi_A)$ , und somit  $A$ , trigonalisierbar.  $\square$

**Folgerung.** Über  $\mathbb{C}$  ist jede quadratische Matrix trigonalisierbar.

*Beweis.* Der Fundamentalsatz der Algebra.  $\square$

### 7.4.5 Begleitmatrix

Sei  $f = X^n + a_{n-1}X^{n-1} + \dots + a_1X + a_0 \in K[X]$  ein normiertes Polynom.

**Definition.** Die *Begleitmatrix* von  $f$  ist definiert als

$$C(f) := \begin{pmatrix} 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & \cdots & 0 & -a_1 \\ \vdots & & \ddots & & \vdots \\ 0 & 0 & \cdots & 1 & -a_{n-1} \end{pmatrix}.$$

**Satz.**  $\chi_{C(f)} = f$ .

*Beweis.* Übung.  $\square$

### 7.4.6 Anwendung: lineare rekursive Folgen

**Definition.** Eine Folge  $(a_k)_{k \in \mathbb{N}_0}$  in  $K$ , die definiert ist durch eine lineare Rekursionsgleichung

$$a_{k+n} = c_0 a_k + c_1 a_{k+1} + \dots + c_{n-1} a_{k+n-1}$$

für  $k \in \mathbb{N}_0$  sowie durch die Anfangswerte  $a_0, \dots, a_{n-1}$ , heißt *lineare rekursive Folge*. Das normierte Polynom

$$f := X^n - c_{n-1}X^{n-1} - \dots - c_1X - c_0$$

heißt *charakteristisches Polynom* der Anfangsdaten  $(a_0, \dots, a_{n-1})$ .

**Bemerkung.** Sei  $(a_k)_{k \in \mathbb{N}_0}$  eine lineare rekursive Folge mit charakteristischem Polynom  $f = X^n - c_{n-1}X^{n-1} - \dots - c_1X - c_0$ .

(i) Es gilt für alle  $k \geq 0$ :

$$\begin{pmatrix} a_{k+1} \\ \vdots \\ a_{k+n} \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & & 0 \\ \vdots & & \ddots & \ddots & \vdots \\ 0 & 0 & \dots & 0 & 1 \\ c_0 & c_1 & \dots & c_{n-2} & c_{n-1} \end{pmatrix}}_{=:C} \begin{pmatrix} a_k \\ \vdots \\ a_{k+n-1} \end{pmatrix},$$

bzw.

$$\begin{pmatrix} a_{k+1} \\ \vdots \\ a_{k+n} \end{pmatrix} = C^{k+1} \begin{pmatrix} a_0 \\ \vdots \\ a_{n-1} \end{pmatrix}.$$

(ii)  $C = C(f)^t$  und  $\chi_C = f$ .

(iii) Ist  $C$  diagonalisierbar, etwa  $T^{-1}CT = \begin{pmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{pmatrix} =: D$  mit  $T \in \text{GL}_n(K)$ , so folgt

$$C^k = (TDT^{-1})^k = TD^kT^{-1} = T \begin{pmatrix} d_1^k & & \\ & \ddots & \\ & & d_n^k \end{pmatrix} T^{-1}.$$

Daraus ergibt sich eine geschlossene Formel für  $a_k$ .

**Beispiel** (Die Fibonacci-Folge). Wir betrachten die Rekursionsgleichung  $f_k = f_{k-1} + f_{k-2}$  mit den Anfangsgliedern  $f_0 = 0, f_1 = 1$ . Das charakteristische Polynom lautet  $f = X^2 - X - 1$ . Für alle  $k \in \mathbb{N}_0$  gilt

$$\begin{pmatrix} f_k \\ f_{k+1} \end{pmatrix} = C^k \begin{pmatrix} 0 \\ 1 \end{pmatrix} \quad \text{mit } C = \begin{pmatrix} 0 & 1 \\ 1 & 1 \end{pmatrix}.$$

Die Nullstellen von  $f$  (Eigenwerte von  $C$ ) lauten  $d_1 = \frac{1+\sqrt{5}}{2}$  und  $d_2 = \frac{1-\sqrt{5}}{2}$ . Eigenvektoren zu  $d_i$  berechnet man so:

$$C - d_i E = \begin{pmatrix} -d_i & 1 \\ 1 & 1 - d_i \end{pmatrix} \rightsquigarrow \begin{pmatrix} 1 & 1 - d_i \\ 0 & d_i - d_i^2 + 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 - d_i \\ 0 & 0 \end{pmatrix},$$

also  $v_i = \begin{pmatrix} d_i - 1 \\ 1 \end{pmatrix}$ . Somit ist

$$D := T^{-1}CT = \begin{pmatrix} d_1 & 0 \\ 0 & d_2 \end{pmatrix} \quad \text{für } T = \begin{pmatrix} d_1 - 1 & d_2 - 1 \\ 1 & 1 \end{pmatrix}.$$

Man berechnet  $f_k$  als den  $(1, 2)$ -Eintrag von  $C^k = TD^kT^{-1}$ :

$$\begin{aligned} TD^k &= \begin{pmatrix} d_1^k(d_1 - 1) & d_2^k(d_2 - 1) \\ d_1^k & d_2^k \end{pmatrix} \\ \det T &= (d_1 - 1) - (d_2 - 1) = d_1 - d_2 = \sqrt{5} \\ T^{-1} &= \frac{1}{\sqrt{5}} \begin{pmatrix} 1 & 1 - d_2 \\ -1 & d_1 - 1 \end{pmatrix} \\ f_k &= \frac{1}{\sqrt{5}} (d_1^k(d_1 - 1)(1 - d_2) + d_2^k(d_2 - 1)(d_1 - 1)) \end{aligned}$$

Wegen  $-1 = f(1) = (1 - d_1)(1 - d_2)$  folgt die Formel

$$f_k = \frac{1}{\sqrt{5}}(d_1^k - d_2^k) = \frac{1}{\sqrt{5}} \left( \left( \frac{1 + \sqrt{5}}{2} \right)^k - \left( \frac{1 - \sqrt{5}}{2} \right)^k \right).$$

## 7.5 Der Satz von Cayley-Hamilton

Es sei  $V$  in diesem Paragraphen ein endlich-dimensionaler  $K$ -Vektorraum mit  $0 < \dim_K V = n < \infty$ . Weiter seien  $\varphi \in \text{End}(V)$  und  $A \in K^{n \times n}$ .

### 7.5.1 Einsetzungshomomorphismus

**Definition.** Die Abbildungen

$$\begin{aligned}\tau_A : K[X] &\rightarrow K^{n \times n}, & A &\mapsto f(A), \\ \tau_\varphi : K[X] &\rightarrow \text{End}(V), & \varphi &\mapsto f(\varphi)\end{aligned}$$

werden jeweils *Einsetzungshomomorphismus* genannt. Ist

$$f = \sum_{i=0}^m a_i X^i \in K[X],$$

dann ist

$$\tau_A(f) = f(A) = \sum_{i=0}^m a_i A^i \in K^{n \times n}$$

und

$$\tau_\varphi(f) = f(\varphi) = \sum_{i=0}^m a_i \varphi^i \in K^{n \times n}.$$

Es gelten dabei die Konventionen  $A^0 = E_n$  und  $\varphi^0 = \text{id}_V$ .

**Bemerkung.** Die Einsetzungshomomorphismen sind sowohl Ring- als auch Vektorraum-Homomorphismen.

**Frage.** Was ist  $\chi_A(A)$  und  $\chi_\varphi(\varphi)$ ?

**Beispiel.**

(i)  $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \in \mathbb{Q}^{2 \times 2}$ ,  $f = X^2 - 5X - 2 \in \mathbb{Q}[X]$ . Es gilt  $A^0 = E_2$ , also

$$f(A) = A^2 - 5A - 2E_2 = \begin{pmatrix} 7 & 10 \\ 15 & 22 \end{pmatrix} - 5 \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} - \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}.$$

(ii)  $\varphi : \mathbb{Q}^2 \rightarrow \mathbb{Q}^2$ ,  $e_1 \mapsto e_1 + e_2$ ,  $e_2 \mapsto 2e_1 + e_2$ ,  $f = X^2 - 2X - 1 \in \mathbb{Q}[X]$ . Es gilt  $\varphi^0 = \text{id}_{\mathbb{Q}^2} =: \text{id}$ , also  $f(\varphi) = \varphi^2 - 2\varphi - \text{id}$ . Wir berechnen

$$\begin{aligned}f(\varphi)(e_1) &= \dots = 0, \\ f(\varphi)(e_2) &= \dots = 0,\end{aligned}$$

also ist  $f(\varphi) = 0$  (Nullabbildung).

## 7.5.2 Invariante Unterräume

**Definition.** Ein Unterraum  $U \leq V$  heißt *invariant unter  $\varphi$*  bzw.  *$\varphi$ -invariant*, wenn  $\varphi(U) \subseteq U$  ist.

Ist  $U \leq V$  invariant unter  $\varphi$ , dann bezeichnen wir mit  $\varphi_U$  den Endomorphismus von  $U$  der durch Einschränkung von  $\varphi$  auf  $U$  (im Definitions- und Wertebereich) definiert ist:

$$\varphi_U : U \rightarrow U, \quad u \mapsto \varphi(u).$$

**Beispiel.**

- (i) Die Drehung des  $\mathbb{R}^3$  an der  $e_3$ -Achse hat die invarianten Unterräume  $\langle e_3 \rangle$  (die Drehachse) und  $\langle e_1, e_2 \rangle$  (die Drehebene). Die Abbildungsmatrix bzgl.  $(e_1, e_2, e_3)$  hat die Form

$$\left( \begin{array}{c|c} 1 & 0 \\ \hline 0 & R_\alpha \end{array} \right),$$

wobei  $R_\alpha$  die übliche  $2 \times 2$ -Drehmatrix ist (vgl. Beispiel 6.4.2).

- (ii) Es sei  $\mathcal{B} = (v_1, \dots, v_n)$  eine geordnete Basis von  $V$ . Zerlege  $M = M_{\mathcal{B}}(\varphi)$  in Blöcke  $M = \left( \begin{array}{c|c} A & B \\ \hline C & D \end{array} \right)$ , so dass  $A$  ein  $r \times r$ -Block ist. Genau dann ist  $U := \langle v_1, \dots, v_r \rangle$  invariant unter  $\varphi$ , wenn  $C = 0$  ist. In diesem Fall ist  $A = M_{(v_1, \dots, v_r)}(\varphi_U)$ . Entsprechend ist  $\langle v_{r+1}, \dots, v_n \rangle$  genau dann  $\varphi$ -invariant, wenn  $B = 0$  ist.

- (iii)  $\{0\}$  und  $V$  sind stets  $\varphi$ -invariant.

- (iv) Sei  $0 \neq v \in V$  beliebig und sei  $r \in \mathbb{N}$  maximal mit

$$(v, \varphi(v), \varphi^2(v), \dots, \varphi^{r-1}(v))$$

linear unabhängig. Dann besitzt  $\varphi^r(v)$  eine Darstellung

$$\varphi^r(v) = \sum_{i=0}^{r-1} a_i \varphi^i(v)$$

mit  $a_0, \dots, a_{r-1} \in K$  (vgl. Lemma 6.2.4). Folglich ist

$$U_v := \langle v, \varphi(v), \dots, \varphi^{r-1}(v) \rangle$$

ein  $\varphi$ -invarianter Unterraum. Es hat  $U_v$  die Dimension  $r$  und

$$\mathcal{B} = (v, \varphi(v), \varphi^2(v), \dots, \varphi^{r-1}(v))$$

ist eine geordnete Basis von  $U_v$ .

Es ist  $U_v$  der kleinste  $\varphi$ -invariante Unterraum, der  $v$  enthält. Im Allgemeinen kann  $U_v = V$  sein. Falls  $v$  ein Eigenvektor von  $\varphi$  ist, so ist  $U_v = \langle v \rangle$  (der Fall  $r = 1$ ).

- (v) Es sei  $f \in K[X]$ . Dann ist  $\text{Kern } f(\varphi)$  ein  $\varphi$ -invarianter Unterraum (denn  $f(\varphi)(v) = 0 \Rightarrow f(\varphi)(\varphi(v)) = \varphi(f(\varphi)(v)) = \varphi(0) = 0$ ), z.B.:  
 für  $f = a \in K \setminus \{0\}$  ist  $\text{Kern } f(\varphi) = \{0\}$ ,  
 für  $f = X - a$  ist  $\text{Kern } f(\varphi) = \text{Kern}(\varphi - a \cdot \text{id}_V) = V_a(\varphi)$ .

**Frage.** Gibt es nicht-triviale  $\varphi$ -invariante Unterräume und wie findet man sie (von möglichst kleiner Dimension)?

**Lemma.** a) Für jeden  $\varphi$ -invarianten Unterraum  $U \leq V$  gilt  $\chi_{\varphi_U} \mid \chi_\varphi$ .  
 b) Für  $f, g \in K[X]$  gilt:  $f \mid g \Rightarrow \text{Kern } f(\varphi) \leq \text{Kern } g(\varphi)$ .

*Beweis.* a) Wähle eine geordnete Basis von  $U$  und ergänze diese zu einer Basis von  $V$ . Dann folgt die Aussage aus Beispiel (ii) und dem Kästchensatz für charakteristische Polynome.

b) Sei  $h \in K[X]$  mit  $g = h \cdot f$  und  $v \in \text{Kern } f(\varphi)$ . Dann ist  $g(\varphi) = (h \cdot f)(\varphi) = h(\varphi) \circ f(\varphi)$ . Also  $g(\varphi)(v) = h(\varphi)(f(\varphi)(v)) = h(\varphi)(0) = 0$ .  $\square$

### 7.5.3 Satz von Cayley-Hamilton

Wir bestimmen nun  $\text{Kern } \chi_\varphi(\varphi)$ .

**Lemma.** Für jedes  $0 \neq v \in V$  ist  $\chi_{\varphi_{U_v}}(\varphi)(v) = 0$ .

Insbesondere folgt  $v \in U_v \leq \text{Kern } \chi_{\varphi_{U_v}}(\varphi)$ .

*Beweis.* Es seien alle Notationen wie in Beispiel 7.5.2iv. Die Abbildungsmatrix von  $\varphi_{U_v}$  bzgl.  $\mathcal{B}$  hat die Form

$$\begin{pmatrix} 0 & \cdots & 0 & a_0 \\ 1 & 0 & \cdots & 0 & a_1 \\ 0 & \ddots & \ddots & & \vdots \\ 0 & & 1 & 0 & a_{r-2} \\ 0 & \cdots & 0 & 1 & a_{r-1} \end{pmatrix} \in K^{r \times r}.$$

Da dies eine Begleitmatrix ist gilt nach Satz 7.4.5:

$$\chi_{\varphi_{U_v}} = X^r - a_{r-1}X^{r-1} - \cdots - a_1X - a_0.$$



Einsetzen von  $\varphi$  liefert

$$(\chi_{\varphi_{U_v}})(\varphi) = \varphi^r - a_{r-1}\varphi^{r-1} - \dots - a_1\varphi - a_0\text{id}_V.$$

Damit ist klar, dass  $v$  im Kern von  $\chi_{\varphi_{U_v}}(\varphi)$  liegt:

$$\begin{aligned} (\chi_{\varphi_{U_v}})(\varphi)(v) &= (\varphi^r - a_{r-1}\varphi^{r-1} - \dots - a_1\varphi - a_0\text{id}_V)(v) \\ &= \varphi^r(v) - a_{r-1}\varphi^{r-1}(v) - \dots - a_1\varphi(v) - a_0v = 0. \end{aligned}$$

□

**Satz.** Es gilt stets  $\chi_\varphi(\varphi) = 0$  bzw.  $\chi_A(A) = 0$ .

In anderen Worten lautet die Aussage: Kern  $\chi_\varphi(\varphi) = V$ .

*Beweis.* Wir beweisen nur die Aussage für  $\varphi$ ; für  $A$  folgt sie durch Übergang zu  $\varphi_A$ . Sei  $0 \neq v \in V$  beliebig. Nach obigem Lemma sowie Lemma 7.5.2 gilt:  $v \in \text{Kern } \chi_{\varphi_{U_v}}(\varphi) \leq \text{Kern } \chi_\varphi(\varphi)$ . □

**Bemerkung a.** Es sei  $f \in K[X]$ . Dann ist  $\dim U_v \leq \deg f$  für alle  $v \in \text{Kern } f(\varphi)$ .

*Beweis.* Ist  $f = a_r X^r + \dots + a_1 X + a_0$  mit  $a_r \neq 0$  und  $f(\varphi)(v) = a_r \varphi^r(v) + \dots + a_1 \varphi(v) + a_0 v = 0$ , so ist  $(v, \varphi(v), \dots, \varphi^r(v))$  linear abhängig, also  $\dim U_v \leq r = \deg f$ . □

**Folgerung.** Sei  $\chi_\varphi = f_1 \cdots f_r$  mit  $f_i \in K[X]$ ,  $\deg f_i \geq 1$ . Sei

$$m = \max\{\deg f_i \mid i = 1, \dots, r\}.$$

Dann existiert ein  $\varphi$ -invarianter Unterraum  $U$  mit  $0 < \dim U \leq m$ .

*Beweis.* Für  $r = 1$  ist  $m = \deg f_1 = \deg \chi_\varphi = n$  und  $V$  selbst ein  $\varphi$ -invarianter Unterraum der Dimension  $n$ . Wir zeigen die Behauptung für  $r = 2$  (allgemein führe man Induktion nach  $r$ ). Sei  $\chi_\varphi = f \cdot g$ . Nach dem Satz von Cayley-Hamilton ist  $\chi_\varphi(\varphi) = f(\varphi) \circ g(\varphi) = 0$ . Wähle ein beliebiges  $0 \neq v \in V$ . Dann gilt  $f(\varphi)(g(\varphi)(v)) = 0$ . Falls  $g(\varphi)(v) = 0$  ist, so ist  $0 < \dim U_v \leq \deg g \leq m$  nach Bemerkung a. Falls  $v' := g(\varphi)(v) \neq 0$  ist, so ist  $f(\varphi)(v') = 0$ , also  $0 < \dim U_{v'} \leq \deg f \leq m$  nach Bemerkung a. □

*Übung* (Berechnung von  $A^{-1}$ ). Es sei  $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \in \mathbb{Q}^{2 \times 2}$ . Man berechne  $A^{-1}$  mit Hilfe des Satzes von Cayley-Hamilton. Dann ist  $\chi_A = X^2 - 5X - 2 \in \mathbb{Q}[X]$ , nach dem Satz von Cayley-Hamilton also  $A^2 - 5A - 2E_2 = 0$ . Es folgt  $A(A - 5E_2) = A^2 - 5A = 2E_2$ , also

$$A^{-1} = \frac{1}{2}(A - 5E_2) = \frac{1}{2} \begin{pmatrix} -4 & 2 \\ 3 & -1 \end{pmatrix}.$$

Wir schließen mit einer allgemeinen Bemerkung zu invarianten Unterräumen.

**Bemerkung b.**

- (i) Gibt es einen  $\varphi$ -invarianten Unterraum  $U$ , so besitzt  $\chi_\varphi$  einen Teiler vom Grad  $\dim U$  (nämlich  $\chi_{\varphi_U}$ ).
- (ii) Ist  $f$  ein Teiler von  $\chi_\varphi$ , so ist  $\text{Kern } f(\varphi)$  ein invarianter Unterraum von  $\varphi$  der Dimension  $\geq \deg f$ . Es stellt sich heraus (Satz 7.6b unten), dass die Dimensionen von  $\text{Kern } f(\varphi)$ , wobei  $f$  die Teiler von  $\chi_\varphi$  durchläuft, bereits  $\varphi$  wesentlich charakterisieren.
- (iii) Nicht jeder  $\varphi$ -invariante Unterraum  $U$  hat die Form  $\text{Kern } f(\varphi)$  für ein Polynom  $f$ .
- (iv) Ist  $\chi_\varphi = f \cdot g$  mit  $f, g$  teilerfremd, dann ist  $V = \text{Kern } f(\varphi) \oplus \text{Kern } g(\varphi)$ .

*Beweis.* (i) wurde in Lemma 7.5.2 gezeigt.

Für lineare Polynome  $f = X - a$  kennen wir die Aussage (ii) bereits, denn  $\text{Kern } f(\varphi)$  ist dann der Eigenraum zu  $a$  und seine Dimension ist die geometrische Vielfachheit von  $a$ . Bekanntlich ist die geometrische Vielfachheit  $\geq 1 = \deg f$  (kann aber auch  $> 1$  sein). Für beliebiges  $f$  ergibt sich der Beweis erst im Rahmen der Normalformtheorie in der Linearen Algebra II.

(iii) sieht man schon am Beispiel der Identität  $\varphi = \text{id}_V$ , denn dafür ist  $\text{Kern } f(\varphi)$  stets  $\{0\}$  oder ganz  $V$ , während es  $\varphi$ -invariante Unterräume jeder Dimension  $\leq \dim V$  gibt.  $\square$

### 7.5.4 Das Minimalpolynom

Hier beantworten wir die Frage nach der Menge der Polynome  $f \in K[X]$  mit  $f(\varphi) = 0$  bzw.  $f(A) = 0$ .

**Satz a.** *Es existiert ein  $0 \neq \mu_\varphi \in K[X]$  mit*

- (i)  $\mu_\varphi$  ist normiert;
- (ii)  $\mu_\varphi(\varphi) = 0$ ;
- (iii)  $\mu_\varphi \mid f$  für alle  $f \in K[X]$  mit  $f(\varphi) = 0$ .

*Durch diese Bedingungen ist  $\mu_\varphi$  eindeutig bestimmt.*

*Beweis.* Betrachte  $I := \{f \in K[X] \mid f(\varphi) = 0\}$ . Dann ist  $I$  ein Ideal in  $K[X]$ . Da  $K[X]$  ein Hauptidealring ist, existiert ein  $\mu_\varphi \in I$  mit  $I = \mu_\varphi K[X] = I$  (siehe Bemerkungen 2.5.1 a,b). Da wir  $\mu_\varphi$  als normiert annehmen können, und  $\mu_\varphi$  dadurch eindeutig bestimmt ist, ist der Satz bewiesen.  $\square$

**Definition a.**  $\mu_\varphi$  heißt das *Minimalpolynom* von  $\varphi$ . Das *Minimalpolynom* von  $A$  ist definiert als  $\mu_{\varphi_A}$ .

**Bemerkung a.** Es gelten

$$(i) \quad \mu_\varphi \mid \chi_\varphi$$

$$(ii) \quad \mu_A \mid \chi_A$$

Insbesondere ist  $\deg \mu_\varphi, \deg \mu_A \leq n$ .

*Beweis.* Dies folgt aus dem Satz von Cayley-Hamilton.  $\square$

**Bemerkung b.** Es sei  $f \in K[X] \setminus K$  normiert. Dann ist  $\mu_{C(f)} = f$ .

*Beweis.* Sei  $A := C(f)$  und  $\varphi := \varphi_A \in \text{End}_K(K^n)$ . Für  $v = e_1 \in K^n$  ist  $(v, \varphi(v), \dots, \varphi^{n-1}(v)) = (e_1, e_2, \dots, e_n)$  linear unabhängig. Damit folgt  $\deg \mu_\varphi \geq n$  aus Bemerkung 7.5.3 a. Andererseits gilt  $\mu_\varphi \mid \chi_\varphi$  nach Bemerkung a und  $\chi_\varphi = \chi_A = \chi_{C(f)} = f$  nach Satz 7.4.5. Also ist  $\mu_\varphi = f$ .  $\square$

**Beispiel a.** (i) Ist  $A \in K^{n \times n}$  mit  $A^m = 0$  für ein  $m \in \mathbb{N}$ , dann ist  $\mu_A = X^k$  für ein  $k \leq m$ .

$$(ii) \quad \text{Ist } A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \text{ dann ist } \mu_A = X^2.$$

$$(iii) \quad \text{Ist } A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \text{ dann ist } \mu_A = X^2 + 1 = \chi_A.$$

(iv) Sei  $n \in \mathbb{N}$ ,  $A \in K^{n \times n}$  und  $a \in K$ . Dann:

$$\mu_A = X - a \Leftrightarrow A = aE_n.$$

In diesem Fall ist  $\chi_A = (X - a)^n$ .

**Bemerkung c.** Ist  $a \in K$  Eigenwert von  $\varphi$ , dann ist  $\mu_\varphi(a) = 0$ .

*Beweis.* Sei  $\mu_\varphi = \sum_{i=0}^m b_i X^i$ . Sei  $0 \neq v \in V$  Eigenvektor zum Eigenwert  $a$  von  $\varphi$ . Dann gilt

$$0 = \sum_{i=0}^m b_i \varphi^i(v) = \sum_{i=0}^m b_i a^i v = \left( \sum_{i=0}^m b_i a^i \right) v.$$

Wegen  $v \neq 0$  folgt  $\mu_\varphi(a) = \sum_{i=0}^m b_i a^i = 0$ .  $\square$

**Satz b.** Die folgenden Aussagen sind äquivalent:

(i)  $\varphi$  diagonalisierbar.

(ii)  $\mu_\varphi = \prod_{i=1}^m (X - a_i)$  mit  $a_1, \dots, a_m \in K$  paarw. versch.

Analoge Aussagen gelten für  $\mu_A$ .

*Beweis.* “(i)  $\Rightarrow$  (ii)” Es sei  $\varphi$  diagonalisierbar und  $a_1, \dots, a_m$  die verschiedenen Eigenwerte von  $\varphi$ . Nach Bemerkung c ist  $f_i := X - a_i$  ein Teiler von  $\mu_\varphi$  für alle  $1 \leq i \leq m$ . Damit ist  $f := \prod_{i=1}^m f_i = \prod_{i=1}^m (X - a_i)$  ein Teiler von  $\mu_\varphi$ .

Da  $V$  eine Basis aus Eigenvektoren von  $\varphi$  besitzt, kann jedes  $v \in V$  geschrieben werden als  $v = w_1 + w_2 + \dots + w_m$  mit  $w_i \in V_{a_i}(\varphi)$  für  $1 \leq i \leq m$ . Wegen  $f_i(\varphi)(w_i) = (\varphi - a_i \text{id}_V)(w_i) = \varphi(w_i) - a_i w_i = 0$  folgt  $f(\varphi)(w_i) = 0$  für alle  $1 \leq i \leq m$  aus Lemma 7.5.2 b). Also gilt auch  $f(\varphi)(v) = 0$  für alle  $v \in V$ , d.h.  $f(\varphi) = 0$ . Damit ist  $\mu_\varphi$  ein Teiler von  $f$  nach Satz a).

“(ii)  $\Leftarrow$  (i)” Wir beweisen die Aussage für den Spezialfall  $m = 2$ ; der allgemeine Fall folgt daraus mit Induktion über  $m$ . Es seien  $a := a_1$  und  $b := a_2$  die Nullstellen von  $\mu_\varphi$ . Dann ist  $a \neq b$  und  $\mu_\varphi = X^2 - (a+b)X + ab$ . Daraus folgt  $\varphi^2 = (a+b)\varphi - ab \text{id}_V$ .

Wir zeigen:  $V = V_a(\varphi) \oplus V_b(\varphi)$ . Dann besitzt  $V$  eine Basis aus Eigenvektoren von  $\varphi$ , so dass  $\varphi$  diagonalisierbar ist. Wegen  $a \neq b$  ist  $V_a(\varphi) \cap V_b(\varphi) = \{0\}$ . Es genügt also,  $V = V_a(\varphi) + V_b(\varphi)$  zu zeigen. Sei dazu  $v \in V$ . Dann ist

$$v = \frac{1}{b-a} [(\varphi(v) - av) - (\varphi(v) - bv)].$$

Wegen  $\varphi^2 = (a+b)\varphi - ab \text{id}_V$  folgt

$$\begin{aligned} \varphi(\varphi(v) - av) &= \varphi^2(v) - a\varphi(v) \\ &= (a+b)\varphi(v) - abv - a\varphi(v) \\ &= b(\varphi(v) - av), \end{aligned}$$

d.h.  $\varphi(v) - av \in V_b(\varphi)$ . Analog ist  $\varphi(v) - bv \in V_a(\varphi)$ , woraus die Behauptung folgt.  $\square$

**Beispiel b.** Sei  $A \in \mathbb{C}^{n \times n}$  mit  $A^4 = E_n$ . Dann ist  $A$  diagonalisierbar.

*Beweis.* Es gilt  $\mu_A \mid X^4 - 1$ , da  $A^4 - E_n = 0$  ist. Nun ist  $X^4 - 1 = (X - 1)(X + 1)(X - \sqrt{-1})(X + \sqrt{-1})$ , also zerfällt  $X^4 - 1$  und damit auch  $\mu_A$  in paarweise verschiedene Linearfaktoren.  $\square$

## 7.6 Ausblick: Normalformen

In der Linearen Algebra II werden *Normalformen* von quadratischen Matrizen diskutiert, die eine solche Matrix bis auf Ähnlichkeit charakterisieren. Wir geben hier eine Zusammenfassung der Resultate ohne Beweise.

**Bemerkung.** Für ähnliche Matrizen  $A, B \in K^{n \times n}$  stimmen überein:

- (i) Determinante und Spur,
- (ii) die charakteristischen Polynome,
- (iii) die Eigenwerte mit algebraischer Vielfachheit,
- (iv) die geometrischen Vielfachheiten, also die Defekte von  $A - aE$  und  $B - aE$  wobei  $a$  ein beliebiger Eigenwert  $a$  von  $A$  ist,
- (v) die Defekte von  $f(A)$  und  $f(B)$  wobei  $f$  ein Teiler von  $\chi_A$  ist,
- (vi) die Ränge von  $f(A)$  und  $f(B)$  wobei  $f$  ein Teiler von  $\chi_A$  ist.

*Übung.* Man zeige die letzten beiden Teile der Bemerkung.

**Definition.** Die Matrix  $A \in K^{n \times n}$  heißt *zerlegbar*, wenn  $A$  ähnlich zu einer Matrix in Blockform  $\begin{pmatrix} B & 0 \\ 0 & C \end{pmatrix}$  mit einer  $r \times r$ -Matrix  $B$  und  $0 < r < n$  ist. Anderenfalls heißt  $A$  *unzerlegbar*.

**Satz a** (Allgemeine Normalform). *Es seien  $A, B \in K^{n \times n}$ .*

(i)  *$A$  ist ähnlich zu einer Matrix der Form  $\begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_r \end{pmatrix}$  mit unzerlegbaren quadratischen Blöcken  $A_1, \dots, A_r$ .*

(ii) *Jede unzerlegbare Matrix  $A$  ist ähnlich zu einer Begleitmatrix (zum Polynom  $\chi_A$ ).*

Eine zu  $A$  ähnliche Matrix der Form  $\begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_r \end{pmatrix}$  mit Begleitmatrizen  $A_1, \dots, A_r$  wird allgemeine Normalform von  $A$  genannt.

(iii) Die allgemeine Normalform von  $A$  ist bis auf Reihenfolge der  $A_i$  eindeutig durch  $A$  bestimmt.

**Satz b.** Für  $A, B \in K^{n \times n}$  sind äquivalent:

- (i)  $A$  und  $B$  sind ähnlich.
- (ii)  $A$  und  $B$  haben gleiche allgemeine Normalformen.
- (iii)  $\chi_A = \chi_B$  und  $\text{Rg } f(A) = \text{Rg } f(B)$  für jeden Teiler  $f$  von  $\chi_A$ .

**Beispiel.** Die Matrizen  $\begin{pmatrix} 1 & 1 & & \\ & 1 & & \\ & & 1 & 1 \\ & & & 1 \end{pmatrix}$  und  $\begin{pmatrix} 1 & 1 & & \\ & 1 & 1 & \\ & & 1 & \\ & & & 1 \end{pmatrix}$  haben gleiche geometrische Vielfachheiten, sind aber nicht ähnlich, weil sie das Kriterium (iii) aus Satz b für  $f = (X - 1)^2$  nicht erfüllen.

*Übung.* Man folgere aus Satz b, dass  $A$  und  $A^t$  ähnlich sind.

**Satz c** (Jordan'sche Normalform). Es sei  $A \in K^{n \times n}$  und  $\chi_A$  zerfalle vollständig in Linearfaktoren (z.B.  $K = \mathbb{C}$ ). Ist  $A$  unzerlegbar, so sind alle Eigenwerte von  $A$  identisch, etwa gleich  $a \in K$ , und  $A$  ist ähnlich zu dem

Jordan-Block  $J_n(a) = \begin{pmatrix} a & 1 & & \\ & a & & \\ & & \ddots & 1 \\ & & & a \end{pmatrix}$ . Eine zu  $A$  ähnliche Matrix der Form

$\begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_r \end{pmatrix}$  mit Jordan-Blöcken  $A_1, \dots, A_r$  wird Jordan'sche Normalform von  $A$  genannt. Die Jordan'sche Normalform von  $A$  ist bis auf Reihenfolge der Jordan-Blöcke eindeutig durch  $A$  bestimmt.

Als Anwendung der Jordan'schen Normalform kann man zeigen:

**Folgerung a.** Jede homogene lineare Differentialgleichung über  $\mathbb{C}$  mit gegebenen Anfangswerten hat eine geschlossene Lösung.

**Folgerung b.** Jede linear rekursive Folge  $(x_k)_{k \in \mathbb{N}_0}$  über  $\mathbb{C}$  besitzt eine geschlossene Formel für  $x_k$ .

*Beweisskizze.* Es sei  $f$  das charakteristische Polynom der Anfangsdaten zur Folge  $(x_k)$  und  $C$  die Transponierte der Begleitmatrix von  $f$  (vgl. § 7.4.6). Dann existiert eine Jordan'sche Normalform  $D$  von  $C$ . Es reicht zu zeigen, dass es geschlossene Formeln für die Einträge von  $D^k$  gibt. Wir können o.B.d.A. annehmen, dass  $D$  ein einzelner Jordan-Block  $J_m(a)$  ist. Dann kann man zeigen:

$$D^k = \begin{pmatrix} \binom{k}{0} a^k & \binom{k}{1} a^{k-1} & \cdots & \binom{k}{k} a^{k-m} \\ 0 & \binom{k}{0} a^k & \binom{k}{1} a^{k-1} & \vdots \\ 0 & 0 & \ddots & \binom{k}{1} a^{k-1} \\ 0 & 0 & 0 & \binom{k}{0} a^k \end{pmatrix}$$

Für  $D = J_4(a)$  ist z.B.

$$D = \begin{pmatrix} a & 1 & 0 & 0 \\ 0 & a & 1 & 0 \\ 0 & 0 & a & 1 \\ 0 & 0 & 0 & a \end{pmatrix}, \quad D^2 = \begin{pmatrix} a^2 & 2a & 1 & 0 \\ 0 & a^2 & 2a & 1 \\ 0 & 0 & a^2 & 2a \\ 0 & 0 & 0 & a^2 \end{pmatrix}$$

$$D^3 = \begin{pmatrix} a^3 & 3a^2 & 3a & 1 \\ 0 & a^3 & 3a^2 & 3a \\ 0 & 0 & a^3 & 3a^2 \\ 0 & 0 & 0 & a^3 \end{pmatrix}, \quad D^4 = \begin{pmatrix} a^4 & 4a^3 & 6a^2 & 4a \\ 0 & a^4 & 4a^3 & 6a^2 \\ 0 & 0 & a^4 & 4a^3 \\ 0 & 0 & 0 & a^4 \end{pmatrix}$$

□





# Kapitel 8

## Euklidische und Unitäre Vektorräume

In diesem Kapitel sei stets  $K = \mathbb{R}$  oder  $K = \mathbb{C}$  und  $V$  sei ein  $K$ -Vektorraum. Wir bezeichnen mit  $\bar{\phantom{x}}$  die komplexe Konjugation, d.h.  $\overline{x + yi} = x - yi$  für  $x, y \in \mathbb{R}$  und  $i = \sqrt{-1}$ . Für  $z \in \mathbb{C}$  gilt  $\bar{\bar{z}} = z \Leftrightarrow z \in \mathbb{R}$ . Diese Notation wird auf Matrizen über  $K$  ausgedehnt: Ist  $A = (a_{ij}) \in K^{n \times n}$ , dann sei  $\bar{A} = (\bar{a}_{ij}) \in K^{n \times n}$ .

### 8.1 Euklidische und unitäre Vektorräume

#### 8.1.1 Skalarprodukte

**Definition.** Eine Abbildung  $\langle -, - \rangle : V \times V \rightarrow K$  heißt *Skalarprodukt* auf  $V$ , wenn für alle  $a, b \in K$  und  $v, w, w_1, w_2 \in V$  gelten:

$$(S1) \quad \langle v, aw_1 + bw_2 \rangle = \bar{a}\langle v, w_1 \rangle + \bar{b}\langle v, w_2 \rangle.$$

$$(S2) \quad \langle v, w \rangle = \overline{\langle w, v \rangle},$$

$$(S3) \quad \langle v, v \rangle > 0 \text{ für alle } v \neq 0.$$

Ist auf  $V$  ein Skalarprodukt  $\langle -, - \rangle$  definiert und ist  $V$  endlich-dimensional, so heißt  $V$ , genauer  $(V, \langle -, - \rangle)$ , ein *euklidischer Vektorraum*, falls  $K = \mathbb{R}$  ist und ein *unitärer Vektorraum*, falls  $K = \mathbb{C}$  ist.

**Bemerkung.** Aus der Definition eines Skalarproduktes folgt sofort für alle  $v, w, v_1, v_2 \in V$  und  $a, b \in K$ :

$$(i) \quad \langle av_1 + bv_2, w \rangle = a\langle v_1, w \rangle + b\langle v_2, w \rangle.$$

$$(ii) \quad \langle v, 0 \rangle = \langle 0, v \rangle = 0.$$

- (iii)
- $\langle v, v \rangle \geq 0$
- , und
- $\langle v, v \rangle = 0 \Leftrightarrow v = 0$
- .

Man sagt, ein Skalarprodukt ist eine *positiv definite, symmetrische Bilinearform*, falls  $K = \mathbb{R}$  ist und eine *positiv definite, Sesquilinearform*, falls  $K = \mathbb{C}$  ist.

**Beispiel.**

- (i) Standard-Skalarprodukt auf
- $K^n$
- :

$$\left\langle \begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}, \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \right\rangle = \sum_{i=1}^n a_i \bar{b}_i \in K.$$

Man nennt  $K^n$ , ausgestattet mit dem Standardskalarprodukt, den *n-dimensionalen euklidischen Raum*, falls  $K = \mathbb{R}$  ist und den *n-dimensionalen unitären Raum*, falls  $K = \mathbb{C}$  ist. Das Standardskalarprodukt lässt sich als Matrixprodukt einer Zeile mit einer Spalte schreiben:

$$\langle x, y \rangle = x^t \cdot \bar{y}$$

für  $x, y \in K^n$ . Ist umgekehrt  $A \in K^{m \times n}$  mit Zeilen  $z_1, \dots, z_m$  und  $x \in K^n$ , so gilt

$$A \cdot \bar{x} = \begin{pmatrix} \langle z_1^t, x \rangle \\ \vdots \\ \langle z_m^t, x \rangle \end{pmatrix}.$$

- (ii) Auf dem
- $\mathbb{R}$
- Vektorraum
- $C^0([0, 1])$
- ist ein Skalarprodukt definiert durch

$$\langle f, g \rangle = \int_0^1 f(t)g(t)dt.$$

- (iii) Ist
- $\langle -, - \rangle$
- ein Skalarprodukt auf
- $V$
- und
- $\varphi \in \text{End}(V)$
- ein Isomorphismus, d.h.
- $\varphi \in \text{Aut}(V)$
- , so wird durch

$$\langle v, w \rangle^\varphi := \langle \varphi(v), \varphi(w) \rangle$$

ein neues Skalarprodukt  $\langle -, - \rangle^\varphi$  definiert.

- (iv) Unterräume euklidischer (unitärer) Vektorräume sind bzgl. der Einschränkung des Skalarproduktes wieder euklidische (unitäre) Vektorräume.

*Beweis.* Die Nachweise, dass es sich um Skalarprodukte handelt, sind eine leichte Übungsaufgabe.  $\square$

### 8.1.2 Die Norm (Länge)

Es sei  $\langle -, - \rangle$  ein Skalarprodukt auf  $V$ .

**Definition.** Die *Norm* oder *Länge* von  $v \in V$  ist definiert als

$$\|v\| := \sqrt{\langle v, v \rangle}.$$

Wir sagen  $v$  ist *normiert*, wenn  $\|v\| = 1$  ist.

**Bemerkung.**

Es sei  $v \in V$  und  $a \in K$ .

- (i)  $\langle v, v \rangle = \|v\|^2$ .
- (ii)  $\|v\| \geq 0$ , und  $\|v\| = 0 \Leftrightarrow v = 0$ .
- (iii)  $\|av\| = |a| \cdot \|v\|$ . Insbesondere ist  $\frac{v}{\|v\|}$  normiert, falls  $v \neq 0$  ist.

**Beispiel.** (i) Die Länge eines Vektor  $v = \begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{R}^2$  bzgl. des Standard-Skalarproduktes ist  $\|v\| = \sqrt{a^2 + b^2}$ .

- (ii) Die Länge von  $f \in C^0([0, 1])$  bzgl. des Skalarproduktes aus Beispiel 8.1.1 ist  $\|f\| = \int_0^1 f^2(t) dt$ .

### 8.1.3 Cauchy-Schwarz'sche Ungleichung

Es sei  $\langle -, - \rangle$  ein Skalarprodukt auf  $V$ .

**Satz** (Cauchy-Schwarz'sche Ungleichung). Für alle  $v, w \in V$  gilt

$$|\langle v, w \rangle| \leq \|v\| \cdot \|w\|.$$

Weiter ist  $|\langle v, w \rangle| = \|v\| \cdot \|w\|$  genau dann, wenn  $(v, w)$  linear abhängig ist.

*Beweis.* Beide Aussagen sind trivial, wenn  $w = 0$  ist. Sei also  $w \neq 0$  und  $a = \langle v, w \rangle / \langle w, w \rangle$ . Die Ungleichung ist äquivalent zu  $|\langle v, w \rangle|^2 \leq \langle v, v \rangle \langle w, w \rangle$ . Wir zeigen  $\langle v, v \rangle \langle w, w \rangle - |\langle v, w \rangle|^2 \geq 0$ . Dies ergibt sich unter Beachtung von  $\bar{a} = \langle w, v \rangle / \langle w, w \rangle$  und  $|\langle v, w \rangle|^2 = \langle v, w \rangle \langle w, v \rangle$  wie folgt:

$$\begin{aligned} 0 &\leq \langle w, w \rangle \langle v - aw, v - aw \rangle \\ &= \langle w, w \rangle (\langle v, v \rangle - \bar{a} \langle v, w \rangle - a \langle w, v \rangle + a\bar{a} \langle w, w \rangle) \\ &= \langle v, v \rangle \langle w, w \rangle - \langle w, v \rangle \langle v, w \rangle - \langle v, w \rangle \langle w, v \rangle + \langle v, w \rangle \langle w, v \rangle \\ &= \langle v, v \rangle \langle w, w \rangle - |\langle v, w \rangle|^2. \end{aligned}$$

Offensichtlich gilt Gleichheit genau dann, wenn  $\langle v - aw, v - aw \rangle = 0$  ist, also genau dann, wenn  $v = aw$  ist.

□

**Folgerung.** Für alle  $v, w \in V$  gilt:

- (i) (Dreiecksungleichung)  $\|v + w\| \leq \|v\| + \|w\|$ .
- (ii) (umgekehrte Dreiecksungleichung)  $|\|v\| - \|w\|| \leq \|v - w\|$ .
- (iii) (Polarisationsformeln)

- Für  $K = \mathbb{R}$  ist

$$\langle v, w \rangle = \frac{1}{2}(\|v + w\|^2 - \|v\|^2 - \|w\|^2).$$

- Für  $K = \mathbb{C}$  ist

$$\langle v, w \rangle = \frac{1}{4}(\|v + w\|^2 - \|v - w\|^2) + i\frac{1}{4}(\|v + iw\|^2 - \|v - iw\|^2).$$

- (iv) (Parallelogramm-Identität)  $\|v + w\|^2 + \|v - w\|^2 = 2\|v\|^2 + 2\|w\|^2$ .

*Beweis.* Übung. □

*Übung.* Gibt es ein Skalarprodukt auf  $\mathbb{R}^n$  derart, dass für alle  $x \in \mathbb{R}^n$  gilt:  $\|x\| = \sum_{i=1}^n |x_i|$  wobei  $x^t = (x_1, \dots, x_n)$ ?

### 8.1.4 Winkel

In diesem Abschnitt sei  $K = \mathbb{R}$ !

Es sei  $\langle -, - \rangle$  ein Skalarprodukt auf  $V$ . Nach der Cauchy-Schwarz'schen Ungleichung ist für alle  $v, w \neq 0$  stets  $|\langle v, w \rangle| / (\|v\| \|w\|) \leq 1$ , d.h.

$$-1 \leq \frac{\langle v, w \rangle}{\|v\| \|w\|} \leq 1.$$

Da  $\cos : [0, \pi] \rightarrow [-1, 1]$  bijektiv ist, gibt es ein eindeutiges  $\alpha \in [0, \pi]$  mit  $\cos \alpha = \frac{\langle v, w \rangle}{\|v\| \|w\|}$ .

**Definition.** Der *Winkel* zwischen  $v, w \in V \setminus \{0\}$ , geschrieben  $\angle(v, w)$ , ist das eindeutige  $\alpha \in [0, \pi]$  mit  $\cos \alpha = \frac{\langle v, w \rangle}{\|v\| \|w\|}$ .

Zwei beliebige Vektoren  $v, w \in V$  heißen *orthogonal*, geschrieben  $v \perp w$ , wenn  $\langle v, w \rangle = 0$  ist.

**Beispiel a.** Im 2-dimensionalen euklidischen Raum  $\mathbb{R}^2$  (mit Standard-Skalarprodukt) gilt für jeden normierten Vektor  $v = \begin{pmatrix} a \\ b \end{pmatrix}$ :

$$\cos(\angle(v, e_1)) = \frac{\langle v, e_1 \rangle}{\|v\| \|e_1\|} = a.$$

Die Definition des Winkels stimmt also mit der geometrischen Interpretation überein.

**Beispiel b.** Wir betrachten den  $\mathbb{R}$ -Vektorraum  $V = C^0([-\pi, \pi])$  mit Skalarprodukt

$$\langle f, g \rangle := \frac{1}{\pi} \int_{-\pi}^{\pi} f(x)g(x)dx.$$

Es gelten

$$\begin{aligned} \|\sin\| &= \frac{1}{\pi} \int_{-\pi}^{\pi} \sin^2(x)dx = \frac{1}{\pi} \left[ \frac{x}{2} - \frac{\sin(2x)}{4} \right]_{-\pi}^{\pi} = 1, \\ \|\cos\| &= \frac{1}{\pi} \int_{-\pi}^{\pi} \cos^2(x)dx = \frac{1}{\pi} \left[ \frac{x}{2} + \frac{\sin(2x)}{4} \right]_{-\pi}^{\pi} = 1, \\ \langle \sin, \cos \rangle &= \frac{1}{\pi} \int_{-\pi}^{\pi} \sin(x) \cos(x)dx = \frac{1}{\pi} \left[ -\frac{1}{2} \cos^2 x \right]_{-\pi}^{\pi} = 0. \end{aligned}$$

D.h.  $\sin$  und  $\cos$  sind normiert und orthogonal zueinander. (Dass das letzte Integral 0 ist, sieht man ohne Rechnung schon daran, dass  $\sin(x) \cos(x)$  eine ungerade Funktion ist.)

Definiere nun  $s_t \in V$ ,  $t \in \mathbb{R}$ , durch  $s_t(x) := \sin(x - t)$ , d.h.  $s_t$  ist eine Phasenverschiebung des Sinus um den Winkel  $t$ . Z.B. ist  $s_0 = \sin$ ,  $s_{\pi} = -\sin$ ,  $s_{\pi/2} = -\cos$ ,  $s_{-\pi/2} = \cos$ . Wir berechnen  $\angle(s_0, s_t)$ . Mit dem trigonometrischen Additionstheorem  $\sin(x - t) = \sin x \cos t - \cos x \sin t$  ergibt sich

$$\begin{aligned} \int \sin x \sin(x - t)dx &= \int (\cos t \sin^2 x - \sin t \sin x \cos x)dx \\ &= \cos t \left( \frac{x}{2} - \frac{\sin(2x)}{4} \right) + \sin t \frac{\cos^2 x}{2} + C, \end{aligned}$$

also

$$\langle s_0, s_t \rangle = \frac{1}{\pi} \left[ \cos t \left( \frac{x}{2} - \frac{\sin(2x)}{4} \right) + \sin t \frac{\cos^2 x}{2} \right]_{-\pi}^{\pi} = \cos t.$$

Also gilt  $\angle(s_0, s_t) = t$ , d.h. der Winkel stimmt mit der Phasenverschiebung überein.

**Bemerkung.**

- (i)  $0 \perp v$  für alle  $v \in V$ .
- (ii)  $(v, w)$  linear abhängig  $\Leftrightarrow \angle(v, w) = 0$  oder  $\pi$ .
- (iii) (Pythagoras) Ist  $v \perp w$ , so gilt

$$\|v + w\|^2 = \|v\|^2 + \|w\|^2.$$

*Beweis.* (i)  $\langle 0, v \rangle = 0$  für alle  $v \in V$ .

(ii) Der zweite Teil der Aussage von Satz 8.1.3.

(iii) Direktes Nachrechnen. □

## 8.2 Orthogonalität

In diesem Abschnitt sei  $K = \mathbb{R}$  oder  $K = \mathbb{C}$  und  $V$  ein  $K$ -Vektorraum mit Skalarprodukt  $\langle -, - \rangle$ . Ist  $V$  endlich-dimensional, dann ist  $V$  also ein euklidischer bzw. unitärer Vektorraum.

### 8.2.1 Orthogonalräume

**Definition.** Es seien  $w \in V, M \subseteq V$  und  $U \leq V$ .

- (i)  $v \in V$  heißt *orthogonal* zu  $w$ , geschrieben  $v \perp w$ , wenn  $\langle v, w \rangle = 0$  ist.
- (ii)  $v \in V$  heißt *orthogonal* zu  $M$ , geschrieben  $v \perp M$ , wenn  $\langle v, v' \rangle = 0$  ist für alle  $v' \in M$ .
- (iii) Der *Orthogonalraum* zu  $M$  ist definiert als

$$M^\perp := \{v \in V \mid v \perp M\} \subseteq V.$$

- (iv) Eine Zerlegung  $v = u + u'$  mit  $u \in U$  und  $u' \in U^\perp$  heißt *Orthogonalzerlegung* von  $v \in V$  bzgl.  $U$ .

**Bemerkung.** Es seien  $M \subseteq V$  und  $U \leq V$ .

- (i)  $M^\perp$  ist ein Unterraum von  $V$ .
- (ii)  $M^\perp = \langle M \rangle^\perp$ .
- (iii)  $M \cap M^\perp \subseteq \{0\}$ .

- (iv) Existiert eine Orthogonalzerlegung von  $v \in V$  bzgl.  $U$ , so ist diese eindeutig.

*Beweis.* Übung. □

**Beispiel.**

- (i) Im euklidischen Raum  $\mathbb{R}^3$  ist  $\{v\}^\perp$  für jedes  $v \neq 0$  eine Ebene.

- (ii) Im euklidischen Raum  $\mathbb{R}^n$  ist für jedes  $A \in \mathbb{R}^{n \times l}$ :  $\text{SR}(A)^\perp = \mathbb{L}(A^t, 0)$ .

*Beweis.* (ii) Es seien  $s_1, \dots, s_l \in \mathbb{R}^n$  die Spalten von  $A$ . Dann gilt nach Teil (ii) obiger Bemerkung:  $\text{SR}(A)^\perp = \{s_1, \dots, s_l\}^\perp = \{x \in \mathbb{R}^n \mid s_i^t \cdot x = 0 \text{ für alle } 1 \leq i \leq l\} = \{x \in \mathbb{R}^n \mid (A^t)x = 0\} = \mathbb{L}(A^t, 0) \leq \mathbb{R}^n$ . □

### 8.2.2 Orthogonalsysteme

**Definition.** Ein Tupel  $(v_1, \dots, v_n)$  mit  $v_i \in V$  und  $v_i \neq 0$  für alle  $1 \leq i \leq n$  heißt *Orthogonalsystem*, wenn  $v_1, \dots, v_n$  paarweise orthogonal sind, d.h.  $v_i \perp v_j$  für alle  $1 \leq i \neq j \leq n$  gilt. Sind  $v_1, \dots, v_n$  zusätzlich normiert, so nennen wir  $(v_1, \dots, v_n)$  ein *Orthonormalsystem*.

Ist  $(v_1, \dots, v_n)$  eine Basis von  $V$ , so sprechen wir auch von *Orthogonalbasen* bzw. *Orthonormalbasen*.

**Bemerkung.** Ein Tupel  $\mathcal{B} = (v_1, \dots, v_n)$  mit  $v_i \in V$  für alle  $1 \leq i \leq n$  ist genau dann ein Orthonormalsystem, wenn für alle  $1 \leq i, j \leq n$  gilt:

$$\langle v_i, v_j \rangle = \begin{cases} 1 & \text{falls } i = j, \\ 0 & \text{falls } i \neq j. \end{cases}$$

**Lemma.** Ist  $(v_1, \dots, v_n)$  ein Orthogonalsystem und  $v = \sum_{i=1}^n a_i v_i$ , so gilt für jedes  $1 \leq j \leq n$ :

$$\langle v, v_j \rangle = a_j \langle v_j, v_j \rangle.$$

*Beweis.*  $\langle v, v_j \rangle = \langle \sum_{i=1}^n a_i v_i, v_j \rangle = \sum_{i=1}^n a_i \langle v_i, v_j \rangle = a_j \langle v_j, v_j \rangle$ . □

**Satz.** Es sei  $U \leq V$  und  $(v_1, \dots, v_n)$  eine Orthogonalbasis von  $U$ . Dann gibt es für jedes  $v \in V$  die Orthogonalzerlegung  $v = u + u'$  und es gilt:

$$u = \sum_{i=1}^n \frac{\langle v, v_i \rangle}{\langle v_i, v_i \rangle} \cdot v_i. \quad (8.1)$$

*Beweis.* Es sei  $u$  wie in (8.1),  $u' := v - u$ . Offensichtlich ist  $u \in U$ . Für jedes  $1 \leq j \leq n$  ist nach dem Lemma, angewendet auf  $u$ :

$$\langle u, v_j \rangle = \frac{\langle v, v_j \rangle}{\langle v_j, v_j \rangle} \langle v_j, v_j \rangle = \langle v, v_j \rangle,$$

also

$$\langle u', v_j \rangle = \langle v - u, v_j \rangle = \langle v, v_j \rangle - \langle u, v_j \rangle = 0.$$

Das bedeutet  $u' \in \{v_1, \dots, v_n\}^\perp = \langle v_1, \dots, v_n \rangle^\perp = U^\perp$ .  $\square$

**Beispiel.** Es seien  $v_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$ ,  $v_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$ ,  $v = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$  im euklidischen

Raum  $\mathbb{R}^3$ . Dann ist  $(v_1, v_2)$  eine Orthogonalbasis der Ebene  $U = \langle v_1, v_2 \rangle$ . Wir berechnen die Orthogonalzerlegung von  $v$  bzgl.  $U$ :

$$u = \frac{\langle v, v_1 \rangle}{\langle v_1, v_1 \rangle} v_1 + \frac{\langle v, v_2 \rangle}{\langle v_2, v_2 \rangle} v_2 = \frac{1}{2} \cdot v_1 + 0 \cdot v_2 = \begin{pmatrix} 1/2 \\ 0 \\ 1/2 \end{pmatrix}.$$

$$u' = v - u = \begin{pmatrix} -1/2 \\ 0 \\ 1/2 \end{pmatrix}.$$

*Übung.* Man zeige (evtl. mit Hilfe des Lemmas), dass jedes Orthogonalsystem linear unabhängig ist.

*Beweis.* Ist  $(v_1, \dots, v_n)$  ein Orthogonalsystem so folgt aus  $\sum_{i=1}^n a_i v_i = 0$  nach dem Lemma, angewendet mit  $v := 0 = \sum_{i=1}^n a_i v_i$ , dass  $a_j \langle v_j, v_j \rangle = \langle v, v_j \rangle = \langle 0, v_j \rangle = 0$  für alle  $1 \leq j \leq n$  ist. Wegen  $v_j \neq 0$  ist somit  $a_j = 0$  für alle  $1 \leq j \leq n$ .  $\square$

### 8.2.3 Das Gram-Schmidt-Verfahren

**Satz.** Jeder endlich-dimensionale euklidische Vektorraum besitzt eine Orthogonalbasis.

*Beweis.* Induktion nach  $n = \dim V$ . Für  $n = 1$  ist  $(v)$  für jedes  $v \neq 0$  eine Orthogonalbasis von  $V$ . Sei nun  $n > 1$ . Wähle einen  $n - 1$ -dimensionalen Unterraum  $U \leq V$ . Nach Induktionsvoraussetzung hat  $U$  eine Orthogonalbasis  $(v_1, \dots, v_{n-1})$ . Wähle  $v \in V \setminus U$  beliebig. Nach Satz 8.2.2 gibt es die Orthogonalzerlegung  $v = u + u'$  bzgl.  $U$ . Setze  $\mathcal{B} := (v_1, \dots, v_{n-1}, u')$ . Wegen  $u' \notin U$  (sonst wäre  $v \in U$ ) ist  $\mathcal{B}$  Basis von  $V$ . Wegen  $u' \in U^\perp$  ist  $\mathcal{B}$  ein Orthonormalsystem.  $\square$



Der Beweis des Satzes lässt sich sofort in folgenden rekursiven Algorithmus übersetzen:

**Algorithmus** (Orthogonalisierungsverfahren von Gram-Schmidt). *Es sei  $U$  ein endlich-dimensionaler Unterraum von  $V$ . Die in der Abbildung dargestellte Prozedur GRAM-SCHMIDT berechnet zu jeder Basis  $(v_1, \dots, v_n)$  von  $U$  eine Orthogonalbasis  $(w_1, \dots, w_n)$  von  $U$ .*

```

GRAM-SCHMIDT( $v_1, \dots, v_n$ )
1  if  $n = 1$ 
2    then return  $v_1$ 
3   $w_1, \dots, w_{n-1} \leftarrow \text{GRAM-SCHMIDT}(v_1, \dots, v_{n-1})$ 
4   $v_{n0} \leftarrow \sum_{i=1}^{n-1} \frac{\langle v_n, w_i \rangle}{\langle w_i, w_i \rangle} w_i$ 
5   $w_n \leftarrow v_n - v_{n0}$ 
6  return  $w_1, \dots, w_n$ 

```

Abbildung 8.1: Prozedur Gram-Schmidt

**Beispiel.** Es sei  $V = \mathbb{R}^4$  der 4-dimensional euklidische Raum,  $U = \langle v_1, v_2, v_3 \rangle$  mit

$$v_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}, v_2 = \begin{pmatrix} 1 \\ -1 \\ 2 \\ 0 \end{pmatrix}, v_3 = \begin{pmatrix} -1 \\ 0 \\ -2 \\ 1 \end{pmatrix}.$$

Wir berechnen mit Gram-Schmidt eine Orthogonalbasis von  $U$ :

$$1. w_1 = v_1 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ 0 \end{pmatrix}.$$

$$2. v_{20} = \frac{\langle v_2, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 = \frac{-1}{1} w_1 = \begin{pmatrix} 0 \\ -1 \\ 0 \\ 0 \end{pmatrix}, w_2 = v_2 - v_{20} = \begin{pmatrix} 1 \\ 0 \\ 2 \\ 0 \end{pmatrix}.$$

$$3. v_{30} = \frac{\langle v_3, w_1 \rangle}{\langle w_1, w_1 \rangle} w_1 + \frac{\langle v_3, w_2 \rangle}{\langle w_2, w_2 \rangle} w_2 = \frac{0}{1} w_1 + \frac{-5}{5} w_2 = \begin{pmatrix} -1 \\ 0 \\ -2 \\ 0 \end{pmatrix},$$

$$w_3 = v_3 - v_{30} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \end{pmatrix}.$$

Es ist  $(w_1, w_2, w_3)$  eine Orthogonalbasis von  $U$ .

*Übung.* Was passiert, wenn man das Gram-Schmidt-Verfahren nicht mit einer Basis, sondern nur mit einem Erzeugendensystem von  $U$  beginnt?

### 8.2.4 Die QR-Zerlegung einer Matrix

Als Anwendung des Gram-Schmidt-Verfahrens stellen wir die in der numerischen Analysis verwendete QR-Zerlegung einer Matrix vor.

**Satz.** Es sei  $A \in K^{m \times n}$  mit  $\text{Rg } A = n$ . Dann existiert eine Matrix  $Q \in K^{m \times n}$ , deren Spalten bzgl. des Standardskalarprodukts ein Orthonormalsystem in  $K^m$  bilden, und eine obere Dreiecksmatrix  $R \in \text{GL}_n(K)$  mit  $A = QR$ . Diese Faktorisierung von  $A$  heißt die QR-Zerlegung von  $A$ .

*Beweis.* Es seien  $s_1, \dots, s_n \in K^m$  die Spalten von  $A$ . Wegen  $\text{Rg } A = n$  ist  $(s_1, \dots, s_n)$  linear unabhängig. Seien  $t'_1, \dots, t'_n$  die Elemente aus  $K^m$ , die vom Gram-Schmidt-Verfahren mit Eingabe  $(s_1, \dots, s_n)$  ausgegeben werden. Es ist also

$$t'_j = s_j - \sum_{i=1}^{j-1} \frac{\langle s_j, t'_i \rangle}{\langle t'_i, t'_i \rangle} t'_i \quad (8.2)$$

für alle  $1 \leq j \leq n$  (siehe Abbildung 8.1). Wir setzen nun  $t_j := t'_j / \|t'_j\|$  für  $1 \leq j \leq n$ . Dann ist  $(t_1, \dots, t_n)$  ein Orthonormalsystem in  $K^m$ . Wir definieren  $Q$  als diejenige  $(m \times n)$ -Matrix über  $K$ , deren Spalten  $t_1, \dots, t_n$  sind. Aus Gleichung (8.2) erhalten wir

$$s_j = \sum_{i=1}^{j-1} \langle s_j, t_i \rangle t_i + t'_j \quad (8.3)$$

für alle  $1 \leq j \leq n$ . Da  $(t_1, \dots, t_j)$  eine Orthonormalsystem ist, erhalten wir aus Gleichung (8.3) auch  $\langle s_j, t_j \rangle = \langle t'_j, t_j \rangle = \langle t'_j, t'_j / \|t'_j\| \rangle = \|t'_j\|$ , also  $t'_j = \langle s_j, t_j \rangle t_j$  und  $\langle s_j, t_j \rangle \neq 0$  für alle  $1 \leq j \leq n$ . Definieren wir nun die Matrix  $R = (r_{ij})_{1 \leq i, j \leq n} \in K^{n \times n}$  durch

$$r_{ij} = \begin{cases} \langle s_j, t_i \rangle & \text{falls } i \leq j, \\ 0 & \text{falls } i > j, \end{cases}$$

dann ist  $R \in \text{GL}_n(K)$  und die Faktorisierung  $A = QR$  ergibt sich aus Gleichung (8.3).  $\square$

**Beispiel.** Es sei  $A = \begin{pmatrix} 1 & 1 & 0 \\ 2 & 0 & -1 \\ 2 & 1 & 1 \end{pmatrix}$ . Wenden wir das Gram Schmidt-Verfahren auf die Spalten

$$s_1 = \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}, \quad s_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \quad s_3 = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix}$$

von  $A$  an und normieren die resultierenden Vektoren, erhalten wir

$$t_1 = \frac{1}{3} \begin{pmatrix} 1 \\ 2 \\ 2 \end{pmatrix}, \quad t_2 = \frac{1}{3} \begin{pmatrix} 2 \\ -2 \\ 1 \end{pmatrix}, \quad t_3 = \frac{1}{3} \begin{pmatrix} -2 \\ -1 \\ 2 \end{pmatrix}.$$

Die Matrizen  $Q$  und  $R$  ergeben sich also zu

$$Q = \frac{1}{3} \begin{pmatrix} 1 & 2 & -2 \\ 2 & -2 & -1 \\ 2 & 1 & 2 \end{pmatrix}$$

und

$$R = \begin{pmatrix} \langle s_1, t_1 \rangle & \langle s_2, t_1 \rangle & \langle s_3, t_1 \rangle \\ 0 & \langle s_2, t_2 \rangle & \langle s_3, t_2 \rangle \\ 0 & 0 & \langle s_3, t_3 \rangle \end{pmatrix} = \begin{pmatrix} 3 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}.$$

### 8.2.5 Die Orthogonalprojektion

**Definition.** Es sei  $U \leq V$ . Ein Endomorphismus  $\pi \in \text{End}(V)$  heißt *Projektion* auf  $U \leq V$ , wenn  $\text{Bild}(\pi) = U$  und  $\pi \circ \pi = \pi$  ist. Wir sprechen von einer *Orthogonal-Projektion* (OP), wenn zusätzlich  $\text{Kern}(\pi) = U^\perp$  ist.

**Satz.** Es sei  $U$  ein endlich-dimensionaler Unterraum von  $V$ . Dann gibt es genau eine Orthogonal-Projektion  $\text{pr}_U$  auf  $U$ .

*Beweis.* Eindeutigkeit: Es sei  $\text{pr}$  eine Orthogonalprojektion auf  $U$  und  $v \in V$ . Setzt man  $u := \text{pr}(v) \in U$  und  $u' := v - \text{pr}(v)$ , so gilt  $\text{pr}(u') = \text{pr}(u) - \text{pr}(\text{pr}(v)) = \text{pr}(u) - \text{pr}(u) = 0$ , also  $u' \in U^\perp$ . D.h.  $v = u + u'$  ist eine Orthogonalzerlegung von  $v$  bzgl.  $U$ . Nach Bemerkung 8.2.1 (iv) ist  $u$  eindeutig durch  $v$  bestimmt.

Existenz: Nach Satz 8.2.3 hat  $U$  eine Orthogonalbasis. Nach Satz 8.2.2 hat jedes  $v \in V$  eine Orthogonalzerlegung  $v = u + u'$  bzgl.  $U$ . Definiere eine Abbildung  $\text{pr} : V \rightarrow U$  durch  $\text{pr}(v) := u$ . Als Übung zeige man, dass  $\text{pr}$  linear ist.  $\square$

**Bemerkung.** Ist  $(v_1, \dots, v_n)$  eine Orthogonalbasis von  $U$ , so ist  $\text{pr}_U(v)$  durch die Formel (8.1) für  $u$  gegeben, die wir auch *Projektionsformel* nennen. Ist  $(v_1, \dots, v_r)$  sogar eine Orthonormalbasis von  $U$ , so vereinfacht sich die Projektions-Formel zu

$$\text{pr}_U(v) = \sum_{i=1}^n \langle v, v_i \rangle \cdot v_i.$$

**Beispiel.** Es sei  $U \leq \mathbb{R}^4$  wie in Beispiel 8.2.3. Wir verwenden die dort berechnete Orthogonalbasis  $(w_1, w_2, w_3)$  von  $U$ , um  $\text{pr}_U(v)$  für folgendes  $v$  zu berechnen:

$$v = \begin{pmatrix} 2 \\ -1 \\ 1 \\ 1 \end{pmatrix}, \text{pr}_U(v) = \sum_{i=1}^3 \frac{\langle v, w_i \rangle}{\langle w_i, w_i \rangle} w_i = \frac{-1}{1} w_1 + \frac{5}{5} w_2 + \frac{1}{1} w_3 = \begin{pmatrix} 1 \\ -1 \\ 2 \\ 1 \end{pmatrix}.$$

### 8.2.6 Die Dimensionsformel

Es sei  $U \leq V$  und  $\dim V = n < \infty$ .

**Satz.**  $\dim U + \dim U^\perp = \dim V$ .

*Beweis.* Dies folgt aus der Dimensionsformel für die lineare Abbildung  $\text{pr}_U$  (siehe Folgerung 6.4.7), denn  $\text{Bild}(\text{pr}_U) = U$  und  $\text{Kern}(\text{pr}_U) = U^\perp$ .  $\square$

*Übung.* Man beweise den Satz ohne Verwendung der Dimensionsformel für lineare Abbildungen. Hinweis: Führe eine Basisergänzung mit dem Gram-Schmidt-Verfahren durch.

**Folgerung.**  $(U^\perp)^\perp = U$ .

*Beweis.* Offensichtlich ist  $U \subseteq (U^\perp)^\perp$ . Nach dem Satz gilt außerdem

$$\dim(U^\perp)^\perp = n - \dim U^\perp = n - (n - \dim U) = \dim U.$$

Damit folgt die Gleichheit.  $\square$

**Beispiel.** Betrachte  $\mathbb{R}^n$  und  $\mathbb{R}^m$  jeweils mit dem Standard-Skalarprodukt. Für jedes  $A \in \mathbb{R}^{m \times n}$  gelten:

$$(i) \quad \text{Kern } \varphi_{A^t} = \mathbb{L}(A^t, 0) = \text{SR}(A)^\perp = (\text{Bild } \varphi_A)^\perp \leq \mathbb{R}^m.$$

$$(ii) \quad \text{Bild } \varphi_{A^t} = \text{SR}(A^t) = \mathbb{L}(A, 0)^\perp = (\text{Kern } \varphi_A)^\perp \leq \mathbb{R}^n,$$

*Beweis.* Die beiden “äußeren” Gleichheitszeichen sind jeweils klar aus der Definition von  $\varphi_A$ . Die Gleichung  $\text{SR}(A)^\perp = \mathbb{L}(A^t, 0)$  wurde bereits in Beispiel 8.2.1 gezeigt. Dieselbe Aussage für  $A^t$  statt  $A$  lautet  $\text{SR}(A^t)^\perp = \mathbb{L}(A, 0) \leq \mathbb{R}^n$ . Mit der Folgerung ergibt sich  $\mathbb{L}(A, 0)^\perp = (\text{SR}(A^t)^\perp)^\perp = \text{SR}(A^t)$ .  $\square$

### 8.2.7 Die Orthogonalentwicklung

**Satz.** Es sei  $\mathcal{B} = (v_1, \dots, v_n)$  eine Orthogonalbasis von  $V$  und  $v \in V$ . Dann ist

$$v = \sum_{i=1}^n \frac{\langle v, v_i \rangle}{\langle v_i, v_i \rangle} \cdot v_i, \quad \text{d.h. } \kappa_{\mathcal{B}}(v) = \begin{pmatrix} \frac{\langle v, v_1 \rangle}{\langle v_1, v_1 \rangle} \\ \vdots \\ \frac{\langle v, v_n \rangle}{\langle v_n, v_n \rangle} \end{pmatrix}.$$

Im Fall einer Orthonormalbasis gilt:

$$v = \sum_{i=1}^n \langle v, v_i \rangle \cdot v_i, \quad \kappa_{\mathcal{B}}(v) = \begin{pmatrix} \langle v, v_1 \rangle \\ \vdots \\ \langle v, v_n \rangle \end{pmatrix}.$$

*Beweis.* Betrachtet man die Orthogonalprojektion von  $V$  auf sich selbst, so ist  $\text{pr}_V(v) = v$  und die Aussage ergibt sich aus (8.1).  $\square$

**Beispiel.** Betrachte die Vektoren

$$v_1 = \frac{1}{\sqrt{6}} \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}, v_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, v_3 = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ -1 \\ -1 \end{pmatrix}, v = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} \in \mathbb{R}^3.$$

Dann ist  $\mathcal{B} = (v_1, v_2, v_3)$  eine Orthonormalbasis von  $\mathbb{R}^3$  und die Orthogonalentwicklung von  $v$  nach  $\mathcal{B}$  lautet:

$$v = \sqrt{\frac{2}{3}}e_1 + \sqrt{2}e_2 - \sqrt{\frac{1}{3}}e_3.$$

*Beweis.* Man rechnet nach, dass

$$\langle v_i, v_j \rangle = \begin{cases} 1 & \text{falls } i = j, \\ 0 & \text{falls } i \neq j \end{cases}$$

für  $i, j \in \{1, 2, 3\}$  ist. D.h.  $(v_1, v_2, v_3)$  ist ein ONS, also nach Satz 8.2.2 auch linear unabhängig, also eine ONB. Die Koordinaten bzgl. dieser Basis berechnet man als

$$\begin{aligned} \langle v, v_1 \rangle &= \frac{1 - 1 + 2}{\sqrt{6}} = \sqrt{\frac{2}{3}}, \\ \langle v, v_2 \rangle &= \frac{1 + 1 + 0}{\sqrt{2}} = \sqrt{2}, \\ \langle v, v_3 \rangle &= \frac{1 - 1 - 1}{\sqrt{3}} = -\sqrt{\frac{1}{3}}. \end{aligned}$$

$\square$

### 8.3 Positiv definite Matrizen

Es sei  $V$  in diesem Abschnitt ein  $K$ -Vektorraum mit  $0 < n = \dim V < \infty$  und es sei  $\langle -, - \rangle$  eine Skalarprodukt auf  $V$ .

#### 8.3.1 Die Gram-Matrix

**Definition.** Für jede geordnete Basis  $\mathcal{B} = (v_1, \dots, v_n)$  von  $V$  definieren wir die *Gram-Matrix* von  $\langle -, - \rangle$  bzgl.  $\mathcal{B}$  als

$$G_{\mathcal{B}} := (\langle v_i, v_j \rangle)_{ij} \in K^{n \times n}.$$

**Beispiel.** Wir betrachten  $\mathbb{R}^n$  mit dem Standard-Skalarprodukt  $\langle -, - \rangle$ .

- (i) Es sei  $\mathcal{B} = (s_1, \dots, s_r)$  mit  $s_i \in \mathbb{R}^n$ . Da  $\langle s_i, s_j \rangle = s_i^t \cdot s_j$  für alle  $1 \leq i, j \leq r$  ist, ergibt sich

$$G_{\mathcal{B}} = A^t A,$$

wobei  $A$  die Matrix mit den Spalten  $s_1, \dots, s_r$  ist, also  $A = M_{\mathcal{E}}^{\mathcal{B}}(\text{id}_{\mathbb{R}^n})$  mit der Standardbasis  $\mathcal{E}$ .

- (ii) Bezüglich der Standardbasis  $\mathcal{E}$  ist  $G_{\mathcal{E}} = E_n$ .

- (iii) Auf  $\mathbb{R}^2$  bzgl.  $\mathcal{B} = \left( \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 3 \end{pmatrix} \right)$  ist z.B.

$$G_{\mathcal{B}} = \begin{pmatrix} 1 & -1 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ -1 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -1 & 3 \end{pmatrix}^t \begin{pmatrix} 1 & 0 \\ -1 & 3 \end{pmatrix} = \begin{pmatrix} 2 & -3 \\ -3 & 9 \end{pmatrix}.$$

**Bemerkung.** Es sei  $\mathcal{B}$  eine geordnete Basis von  $V$ .

- (i)  $\mathcal{B}$  Orthogonalbasis  $\Leftrightarrow G_{\mathcal{B}}$  Diagonalmatrix.

- (ii)  $\mathcal{B}$  Orthonormalbasis  $\Leftrightarrow G_{\mathcal{B}} = E_n$ .

- (iii)

$$\langle v, w \rangle = \kappa_{\mathcal{B}}(v)^t \cdot G_{\mathcal{B}} \cdot \overline{\kappa_{\mathcal{B}}(w)} \text{ für alle } v, w \in V. \quad (8.4)$$

- (iv)  $\mathcal{B}$  Orthonormalbasis  $\Rightarrow \langle v, w \rangle = \kappa_{\mathcal{B}}(v)^t \cdot \overline{\kappa_{\mathcal{B}}(w)}$  für alle  $v, w \in V$ .

Nach (iii) ist das Skalarprodukt schon eindeutig durch die Gram-Matrix definiert. Nach (iv) verhält sich jedes Skalarprodukt wie das Standardskalarprodukt, wenn man zu den Koordinatenvektoren bzgl. einer Orthonormalbasis übergeht. Insbesondere ist das Skalarprodukt schon eindeutig durch die Angabe einer Orthonormalbasis definiert.

*Beweis.* Beweis als Übung. □

**Frage.**

1. Wie bekommt man alle Skalarprodukte auf  $V$ ?
2. Welche quadratischen Matrizen treten als Gram-Matrizen auf, d.h. für welche quadratischen Matrizen wird durch die Formel (8.4) ein Skalarprodukt definiert?
3. Wie verhält sich die Gram-Matrix unter Basiswechsel?
4. Wie verhalten sich die Gram-Matrizen verschiedener Skalarprodukte zueinander?

### 8.3.2 Basiswechselsatz

**Satz.** Seien  $\mathcal{B}$  und  $\mathcal{B}'$  zwei geordnete Basen von  $V$ , und  $T := M_{\mathcal{B}}^{\mathcal{B}'}(\text{id}_V)$  die Basiswechslematrix. Dann gilt  $G_{\mathcal{B}'} = T^t \cdot G_{\mathcal{B}} \cdot \overline{T}$ .

*Beweis.* Sei  $\mathcal{B} = (v_1, \dots, v_n)$  und  $\mathcal{B}' = (v'_1, \dots, v'_n)$ . Für jedes  $1 \leq j \leq n$  gilt  $\kappa_{\mathcal{B}}(v'_j) = T e_j$ . Aus Bemerkung 8.3.1(iii) folgt

$$\langle v'_i, v'_j \rangle = \kappa_{\mathcal{B}}(v'_i)^t \cdot G_{\mathcal{B}} \cdot \overline{\kappa_{\mathcal{B}}(v'_j)} = e_i^t (T^t \cdot G_{\mathcal{B}} \cdot \overline{T}) e_j$$

für alle  $1 \leq i, j \leq n$ . Auf der linken Seite obiger Gleichung steht der  $(i, j)$ -Eintrag der Matrix  $G_{\mathcal{B}'}$ , auf der rechten Seite der  $(i, j)$ -Eintrag der Matrix  $T^t \cdot G_{\mathcal{B}} \cdot \overline{T}$ , woraus sich die Behauptung ergibt. □

*Übung.* Wir definieren eine Relation  $\sim$  auf  $\mathbb{R}^{n \times n}$  durch  $A \sim B$ , falls  $T \in \text{GL}_n(\mathbb{R})$  existiert mit  $T^t A T = B$ . Zeige, dass  $\sim$  eine Äquivalenzrelation ist.

*Übung.* Man zeige durch eine Abwandlung des Gauß-Algorithmus und mit Hilfe des Basiswechselsatzes für Gram-Matrizen, dass jeder endlich-dimensionale euklidische Vektorraum eine Orthogonalbasis besitzt.

### 8.3.3 Positiv definite Matrizen

**Definition.** Es sei  $A \in K^{n \times n}$ .

- (i) (Erinnerung:) Es sei  $K = \mathbb{R}$ . Dann heißt  $A$  *symmetrisch*, wenn  $A = A^t$  ist.
- (ii) Es sei  $K = \mathbb{C}$ . Dann heißt  $A^\dagger := \overline{A}^t$  die zu  $A$  adjungierte Matrix. Es heißt  $A$  *hermitesch*, wenn  $A = A^\dagger$  ist.

- (iii) Es sei  $K = \mathbb{R}$ . Dann heißt  $A$  *positiv definit*, wenn  $A$  symmetrisch ist und  $v^t A v > 0$  für alle  $0 \neq v \in \mathbb{R}^n$  ist.
- (iv) Es sei  $K = \mathbb{C}$ . Dann heißt  $A$  *positiv definit*, wenn  $A$  hermitesch ist und  $v^t A \bar{v} > 0$  für alle  $0 \neq v \in \mathbb{C}^n$  ist.

**Satz.** Es seien  $\mathcal{B}$  eine geordnete Basis von  $V$ . Dann ist  $G_{\mathcal{B}}$  positiv definit.

*Beweis.* Offensichtlich ist  $G_{\mathcal{B}}$  symmetrisch bzw. hermitesch, weil das Skalarprodukt diese Eigenschaften hat. Sei nun  $0 \neq x \in K^n$ . Dann existiert  $0 \neq v \in V$  mit  $x = \kappa_{\mathcal{B}}(v)$ . Damit gilt mit Bemerkung 8.3.1(iii):

$$0 < \langle v, v \rangle = \kappa_{\mathcal{B}}(v)^t \cdot G_{\mathcal{B}} \cdot \overline{\kappa_{\mathcal{B}}(v)} = x^t G_{\mathcal{B}} \bar{x},$$

woraus die Behauptung folgt. □

Umgekehrt kann man zu einer positiv definiten  $(n \times n)$ -Matrix ein Skalarprodukt auf  $K^n$  definieren.

**Bemerkung.** Es sei  $A \in K^{n \times n}$  positiv definit. Dann ist

$$\langle -, - \rangle_A : K^n \times K^n \rightarrow K, \quad \langle x, y \rangle := x^t \cdot A \cdot \bar{y}$$

eine Skalarprodukt auf  $K^n$ .

*Beweis.* Die Bedingungen (S1) aus Definition 8.1.1 folgen aus Eigenschaften der Matrixmultiplikation. Bedingung (S2) ist erfüllt, weil  $A$  symmetrisch bzw. hermitesch ist, und Bedingung (S3), weil  $A$  positiv definit ist. □

**Beispiel.**

- (i)  $E_n$  ist positiv definit.
- (ii) Eine Diagonalmatrix  $D$  ist genau dann positiv definit, wenn alle Diagonaleinträge positive reelle Zahlen sind. Das liegt daran, dass  $e_i D \bar{e}_i = e_i D e_i$  gerade der  $i$ -te Diagonaleintrag von  $D$  ist.

- (iii)  $A = \begin{pmatrix} 4 & -2 \\ -2 & 3 \end{pmatrix} \in \mathbb{R}^{2 \times 2}$  ist positiv definit, denn

$$(a, b) A \begin{pmatrix} a \\ b \end{pmatrix} = 4a^2 - 4ab + 3b^2 = (2a - b)^2 + 2b^2 > 0$$

falls  $a \neq 0$  oder  $b \neq 0$ .



(iv)  $A = \begin{pmatrix} 4 & -2 \\ -2 & -1 \end{pmatrix} \in \mathbb{R}^{2 \times 2}$  ist nicht positiv definit, denn

$$(0, 1)A \begin{pmatrix} 0 \\ 1 \end{pmatrix} = -1.$$

(v)  $A = \begin{pmatrix} 4 & -2 \\ -2 & 1 \end{pmatrix} \in \mathbb{R}^{2 \times 2}$  ist nicht positiv definit, denn

$$(1, 2)A \begin{pmatrix} 1 \\ 2 \end{pmatrix} = 0.$$

**Folgerung.**

(i) Es sei  $A \in K^{n \times n}$ . Dann sind äquivalent:

- $A$  ist positiv definit.
- Es existiert  $T \in \text{GL}_n(K)$  mit  $A = T^t \bar{T}$ .

(ii) Für  $A \in \mathbb{R}^{n \times n}$  und  $S \in \text{GL}_n(\mathbb{R})$  sind äquivalent:

- $A$  positiv definit.
- $S^t A S$  positiv definit.

*Beweis.* (i) Sei  $A$  positiv definit. Nach der Bemerkung ist  $\langle -, - \rangle_A$  ein Skalarprodukt auf  $K^n$ . Offensichtlich ist  $A$  die Gram-Matrix von  $\langle -, - \rangle_A$  bzgl. der Standardbasis  $\mathcal{B}$ . Sei  $\mathcal{B}'$  eine Orthonormalbasis von  $K^n$  bzgl.  $\langle -, - \rangle_A$  und  $S = M_{\mathcal{B}}^{\mathcal{B}'}(\text{id}_{K^n})$  die Basiswechselmatrix, d.h. die Spalten von  $S$  sind die Elemente von  $\mathcal{B}'$ . Nach dem Basiswechselsatz folgt  $E_n = S^t \cdot A \cdot \bar{S}$ , da die Gram-Matrix von  $\langle -, - \rangle_A$  bzgl.  $\mathcal{B}'$  die Einheitsmatrix ist. Es folgt  $A = T^t \bar{T}$  mit  $T = S^{-1}$ . Die Umkehrung ist eine leichte Übungsaufgabe.

(ii) Sei  $A$  positiv definit. Nach (i) ist dann  $A = T^t T$  mit  $T \in \text{GL}_n(K)$ , also auch  $S^t A S = S^t T^t T S = (TS)^t (TS)$  positiv definit. Die Richtung  $\Rightarrow$  folgt wegen  $(S^{-1})^t (S^t A S) S^{-1} = A$ .  $\square$

**Beispiel.** Die Matrix  $A = \begin{pmatrix} 2 & -1 \\ -1 & 5 \end{pmatrix} \in \mathbb{R}^{2 \times 2}$  ist positiv definit. In der Tat ist

$$\begin{pmatrix} 2 & -1 \\ -1 & 5 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & -2 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 1 & -2 \end{pmatrix}$$

und  $T := \begin{pmatrix} 1 & 1 \\ 1 & -2 \end{pmatrix}$  ist invertierbar.

Insbesondere ist  $\langle -, - \rangle_A$  ein Skalarprodukt auf  $\mathbb{R}^2$ . Es ergibt sich

$$\left\langle \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \right\rangle_A = (x_1, x_2) \cdot \begin{pmatrix} 2 & -1 \\ -1 & 5 \end{pmatrix} \cdot \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = 2x_1y_1 - x_1y_2 - x_2y_1 + 5x_2y_2.$$

**Frage.** Wie stellt man systematisch fest, ob eine symmetrische Matrix positiv definit ist?

## 8.4 Unitäre und orthogonale Abbildungen

Es seien  $V$  und  $W$  zwei endlich-dimensionale  $K$ -Vektorräume mit Skalarprodukten  $\langle -, - \rangle$ . (Wir verwenden das gleiche Symbol für die Skalarprodukte auf  $V$  und auf  $W$ .)

### 8.4.1 Adjungierte Abbildungen

**Definition.** Es sei  $\varphi : V \rightarrow W$  eine lineare Abbildung. Eine lineare Abbildung  $\psi : W \rightarrow V$  heißt *Adjungierte zu  $\varphi$* , wenn gilt:

$$\langle \varphi(v), w \rangle = \langle v, \psi(w) \rangle$$

für alle  $v \in V$  und  $w \in W$ .

**Satz.** Es sei  $\varphi : V \rightarrow W$  linear. Dann existiert genau eine zu  $\varphi$  adjungierte Abbildung. Diese wird mit  $\varphi^\dagger$  bezeichnet. Sind  $\mathcal{B}$  und  $\mathcal{C}$  Orthogonalnormalbasen von  $V$  bzw.  $W$ , dann gilt

$$M_{\mathcal{B}}^{\mathcal{C}}(\varphi^\dagger) = M_{\mathcal{C}}^{\mathcal{B}}(\varphi)^\dagger.$$

*Beweis.* Wähle geordnete Orthogonalnormalbasen  $\mathcal{B} = (v_1, \dots, v_n)$  von  $V$  und  $\mathcal{C} = (w_1, \dots, w_m)$  von  $W$ . Setze  $A := M_{\mathcal{C}}^{\mathcal{B}}(\varphi)$  und sei  $A = (a_{ij})_{1 \leq i, j \leq n}$ . Definiere  $\psi : W \rightarrow V$  durch die Bilder auf den Elementen von  $\mathcal{C}$  wie folgt (vgl. Satz 6.3.4b). Für alle  $1 \leq j \leq m$  sei

$$\psi(w_j) := \sum_{i=1}^n \bar{a}_{ji} v_i.$$

Dann ist  $M_{\mathcal{B}}^{\mathcal{C}}(\psi) = A^\dagger$ , und wir müssen zeigen, dass  $\psi$  adjungiert zu  $\varphi$  ist. Dazu genügt es wegen der Eigenschaften (S1) aus Definition 8.1.1 zu zeigen:

$$\langle \varphi(v_i), w_j \rangle = \langle v_i, \psi(w_j) \rangle$$

für alle  $1 \leq i \leq n$  und alle  $1 \leq j \leq m$ . Die linke Seite ergibt sich zu

$$\langle \varphi(v_i), w_j \rangle = \left\langle \sum_{k=1}^m a_{ki} w_k, w_j \right\rangle = \sum_{k=1}^m a_{ki} \langle w_k, w_j \rangle = a_{ji},$$

und die rechte Seite zu

$$\langle v_i, \psi(w_j) \rangle = \left\langle v_i, \sum_{k=1}^n \bar{a}_{jk} v_k \right\rangle = \sum_{k=1}^n a_{jk} \langle v_i, v_k \rangle = a_{ji},$$

woraus die Behauptung folgt.

Wir zeigen jetzt die Eindeutigkeit. Seien  $\psi$  und  $\psi'$  Adjungierte zu  $\varphi$ . Dann gilt für alle  $v \in V$  und  $w \in W$ :

$$\langle v, \psi(w) \rangle = \langle \varphi(v), w \rangle = \langle v, \psi'(w) \rangle.$$

Wir erhalten  $\langle v, \psi(w) - \psi'(w) \rangle = 0$  für alle  $v \in V$  und  $w \in W$ , woraus  $\psi(w) = \psi'(w)$  für alle  $w \in W$  folgt.

Wir setzen  $\varphi^\dagger := \psi$ . Seien  $\mathcal{B}'$  und  $\mathcal{C}'$  Orthonormalbasen von  $V$  bzw.  $W$ . Wir definieren  $\psi' : W \rightarrow V$  wie oben mittels der Basis  $\mathcal{C}'$  von  $W$  und der Matrix  $A' := M_{\mathcal{C}'}^{\mathcal{B}'}(\varphi)$ . Dann ist  $M_{\mathcal{B}'}^{\mathcal{C}'}(\psi') = (A')^\dagger$  und  $\psi'$  ist adjungiert zu  $\varphi$ . Wegen der Eindeutigkeit der Adjungierten ist  $\psi' = \varphi^\dagger$ . Damit sind alle Teile des Satzes bewiesen.  $\square$

**Bemerkung.** Es sei  $\varphi : V \rightarrow W$  eine lineare Abbildung. Dann ist  $(\varphi^\dagger)^\dagger = \varphi$ .

*Beweis.* Es ist  $\langle w, (\varphi^\dagger)^\dagger(v) \rangle = \langle \varphi^\dagger(w), v \rangle = \overline{\langle v, \varphi^\dagger(w) \rangle} = \overline{\langle \varphi(v), w \rangle} = \langle w, \varphi(v) \rangle$  für alle  $v \in V$  und  $w \in W$ . Es folgt  $(\varphi^\dagger)^\dagger(v) = \varphi(v)$  für alle  $v \in V$ .  $\square$

*Übung.* Es seien  $U$  ein endlich-dimensionaler  $K$ -Vektorraum mit Skalarprodukt und  $\varphi : V \rightarrow W$  und  $\psi : W \rightarrow U$  lineare Abbildungen. Dann ist  $(\psi \circ \varphi)^\dagger = \varphi^\dagger \circ \psi^\dagger$ .

### 8.4.2 Unitäre und orthogonale Homomorphismen

Es sei  $\varphi \in \text{Hom}_K(V, W)$ .

**Bemerkung.** Es sei  $\mathcal{B}$  eine Basis von  $V$ . Folgende Aussagen sind äquivalent:

- (i)  $\langle \varphi(v), \varphi(w) \rangle = \langle v, w \rangle$  für alle  $v, w \in V$ .
- (ii)  $\langle \varphi(v), \varphi(w) \rangle = \langle v, w \rangle$  für alle  $v, w \in \mathcal{B}$ .
- (iii)  $\|\varphi(v)\| = \|v\|$  für alle  $v \in V$ .

(iv)  $v \in V$  normiert  $\Rightarrow \varphi(v)$  normiert.

In diesem Fall gelten insbesondere:

(v)  $\varphi$  injektiv,

(vi)  $v \perp w \Rightarrow \varphi(v) \perp \varphi(w)$  für alle  $v, w \in V$ ,

(vii)  $\varphi(U^\perp) \subseteq \varphi(U)^\perp$  für alle  $U \leq V$ ,

(viii)  $\mathcal{B}$  Orthonormalbasis  $\Rightarrow \varphi(\mathcal{B})$  Orthonormalsystem,

(ix) Ist  $K = \mathbb{R}$ , dann ist  $\angle(\varphi(v), \varphi(w)) = \angle(v, w)$  für alle  $v, w \in V$ ,

(x) Ist  $\dim V = \dim W$ , dann ist  $\varphi^{-1} = \varphi^\dagger$ ,

(xi) Ist  $V = W$  und  $c$  Eigenwert von  $\varphi$ , dann ist  $c\bar{c} = 1$ .

*Beweis.* Die Äquivalenz von (i) und (ii) bzw. (iii) und (iv) folgt aus Eigenschaft (S1) aus Definition 8.1.1. Die Äquivalenz von (ii) und (iii) folgt aus den Polarisationsformeln. Für den Rest des Beweises nehmen wir an, dass  $\varphi$  eine der Bedingungen (i) – (iv) erfüllt. Sei  $v \in \text{Kern}(\varphi)$ . Dann ist  $0 = \|0\| = \|\varphi(v)\| = \|v\|$ , also  $v = 0$ . Das beweist (v). Die Aussagen (vi) – (ix) sind leichte Übungsaufgaben. Ist  $\dim V = \dim W$ , dann ist  $\varphi$  invertierbar, und es gilt für alle  $v \in V, w \in W$ :

$$\langle \varphi(v), w \rangle = \langle \varphi(v), \varphi(\varphi^{-1}(w)) \rangle = \langle v, \varphi^{-1}(w) \rangle.$$

Aus Satz 8.4.1 folgt  $\varphi^\dagger = \varphi^{-1}$ , und damit die Aussage (x). Sei nun  $V = W$  und  $v \in V$  ein Eigenvektor von  $\varphi$  zum Eigenwert  $c$ . Dann ist  $\langle v, v \rangle = \langle \varphi(v), \varphi(v) \rangle = \langle cv, cv \rangle = c\bar{c}\langle v, v \rangle$ . Aus  $\langle v, v \rangle \neq 0$  folgt die Aussage (xi).  $\square$

**Definition.** Es erfülle  $\varphi \in \text{End}_K(V)$  die Bedingungen aus der Bemerkung.

(i) Ist  $K = \mathbb{C}$ , dann heißt  $\varphi$  *unitär*. Die Menge der unitären Endomorphismen von  $V$  wird mit  $U(V)$  bezeichnet.

(ii) Ist  $K = \mathbb{R}$ , dann heißt  $\varphi$  *orthogonal*. Die Menge der orthogonalen Endomorphismen von  $V$  wird mit  $O(V)$  bezeichnet.

**Beispiel.**

(i) Der Endomorphismus  $-\text{id}_V$  ist unitär (orthogonal), denn

$$\langle -v, -w \rangle = (-1)^2 \langle v, w \rangle = \langle v, w \rangle.$$

- (ii) Ist  $\varphi$  unitär (orthogonal), so ist auch  $\varphi^{-1}$  unitär (orthogonal) (leichte Übung).
- (iii) Es sei  $\dim V = n$  und  $\mathcal{B}$  eine Orthonormalbasis von  $V$ . Betrachtet man  $K^n$  mit dem Standard-Skalarprodukt, so erfüllt  $\kappa_{\mathcal{B}} : V \rightarrow K^n$  die Bedingungen der Bemerkung.
- (iv) Orthogonalprojektionen sind im Allgemeinen nicht injektiv, also nicht orthogonal.
- (v) Eine Spiegelung  $\varphi$  im Sinne von Definition 7.2.7 ist im Allgemeinen nicht orthogonal, sondern nur dann, wenn  $V_{-1}(\varphi) = V_1(\varphi)^{\perp}$  ist. In diesem Abschnitt meinen wir mit *Spiegelung* stets eine orthogonale Spiegelung.
- (vi) Jede Drehung des  $\mathbb{R}^2$  um den Ursprung ist orthogonal, weil sie längen-erhaltend ist.

*Übung a.* Man zeige:  $U(V)$  bzw.  $O(V)$  ist eine Untergruppe von  $\text{Aut}(V)$ .

*Übung b.* Es sei  $K = \mathbb{R}$  und  $\varphi \in \text{Hom}_{\mathbb{R}}(V, W)$ . Man zeige:  $\text{Bild}(\varphi^{\dagger}) = \text{Kern}(\varphi)^{\perp}$  und  $\text{Kern}(\varphi^{\dagger}) = \text{Bild}(\varphi)^{\perp}$ .

*Hinweis:* Beispiel 8.2.6.

### 8.4.3 Unitäre und orthogonale Matrizen

**Bemerkung.** Für  $A \in K^{n \times n}$  sind folgende Aussagen äquivalent:

- (i)  $A^{\dagger}A = E_n$ .
- (ii)  $A \in \text{GL}_n(K)$  und  $A^{-1} = A^{\dagger}$ .
- (iii) Die Spalten von  $A$  bilden eine Orthonormalbasis von  $K^n$  bzgl. des Standard-Skalarprodukts.
- (iv) Die Zeilen von  $A$  bilden eine Orthonormalbasis von  $K^{1 \times n}$  bzgl. des Standard-Skalarprodukts.

In diesem Fall gelten insbesondere:

- (v)  $|\det A| = 1$ .

*Beweis.* Übung. □

**Definition.** (i)  $A \in \mathbb{C}^{n \times n}$  heißt *unitär*, wenn die Bedingungen aus Bemerkung 8.4.3 gelten. Die Menge aller unitären  $n \times n$ -Matrizen

$$U_n(\mathbb{C}) := \{A \in \mathbb{C}^{n \times n} \mid A^\dagger A = E_n\}$$

wird *unitäre Gruppe* genannt.

(ii)  $A \in \mathbb{R}^{n \times n}$  heißt *orthogonal*, wenn die Bedingungen aus Bemerkung 8.4.3 gelten. Die Menge aller orthogonalen  $n \times n$ -Matrizen

$$O_n(\mathbb{R}) := \{A \in \mathbb{R}^{n \times n} \mid A^t A = E_n\}$$

wird *orthogonale Gruppe* genannt. Auch die Bezeichnung  $O(n)$  statt  $O_n(\mathbb{R})$  ist gebräuchlich.

*Übung.* Man zeige  $O_n(\mathbb{R}) \leq \mathrm{GL}_n(\mathbb{R})$  und  $U_n(\mathbb{C}) \leq \mathrm{GL}_n(\mathbb{C})$ . Ist  $A \in O_n(\mathbb{R})$ , dann auch  $A^t$ . Ist  $A \in U_n(\mathbb{C})$ , dann auch  $A^\dagger$ .

**Satz.** Es sei  $K = \mathbb{R}$  und  $\dim V = n > 0$ . Weiter sei  $\mathcal{B}$  eine Orthonormalbasis von  $V$ . Für  $\varphi \in \mathrm{End}(V)$  sind äquivalent:

- (i)  $\varphi \in O(V)$ .
- (ii)  $\varphi(\mathcal{B})$  ist eine Orthonormalbasis von  $V$ .
- (iii)  $M_{\mathcal{B}}(\varphi) \in O_n(\mathbb{R})$ .

Analoge Aussagen gelten für den Fall  $K = \mathbb{C}$  und die Gruppen  $U(V)$  bzw.  $U_n(\mathbb{C})$ .

*Beweis.* Aus Bemerkung 8.4.2 (viii) ergibt sich die Implikation von (i) nach (ii). Es sei  $\mathcal{B} = (v_1, \dots, v_n)$ . Die Spalten von  $M_{\mathcal{B}}(\varphi)$  werden gebildet von den Vektoren  $\kappa_{\mathcal{B}}(\varphi(v_1)), \dots, \kappa_{\mathcal{B}}(\varphi(v_n))$ . Ist  $\varphi(\mathcal{B})$  ist eine Orthonormalbasis von  $V$ , dann ist  $(\kappa_{\mathcal{B}}(\varphi(v_1)), \dots, \kappa_{\mathcal{B}}(\varphi(v_n)))$  eine Orthonormalbasis von  $K^n$  bzgl. des Standard-Skalarprodukts (siehe Beispiel 8.4.2(iii)). Also ist  $M_{\mathcal{B}}(\varphi) \in O_n(\mathbb{R})$  nach Bemerkung (iii). Da  $\mathcal{B}$  eine Orthonormalbasis von  $V$  ist, ist  $G_{\mathcal{B}} = E_n$ . Damit ergibt sich aus Bemerkung 8.3.1(iii):

$$\begin{aligned} \langle \varphi(v), \varphi(w) \rangle &= \kappa_{\mathcal{B}}(\varphi(v))^t \cdot \kappa_{\mathcal{B}}(\varphi(w)) \\ &= \kappa_{\mathcal{B}}(v)^t M_{\mathcal{B}}(\varphi)^t \cdot M_{\mathcal{B}}(\varphi) \kappa_{\mathcal{B}}(w) \\ &= \kappa_{\mathcal{B}}(v)^t \cdot \kappa_{\mathcal{B}}(w) \\ &= \langle v, w \rangle \end{aligned}$$

für alle  $v, w \in V$ . Also ist  $\varphi \in O(V)$ . □

**Beispiel.** In diesem Beispiel sei  $K = \mathbb{R}$ .

- (i)  $A \in O(1) \Leftrightarrow A = (1)$  oder  $A = (-1)$ .
- (ii) Die orthogonalen  $1 \times 1$ -Matrizen sind genau  $(1)$  und  $(-1)$ .
- (iii) Eine Diagonalmatrix ist genau dann orthogonal, wenn alle Diagonaleinträge gleich  $\pm 1$  sind.
- (iv) Die  $2 \times 2$ -Matrizen  $R_\alpha$  (Drehung) und  $S_\alpha$  (Spiegelung) aus Beispiel 6.4.2 sind orthogonal. Z.B.

$$R_{\pi/4} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix} \in O(2).$$

- (v) Die Matrix  $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$  (Scherung) hat  $\det A = 1$ , ist aber nicht orthogonal.

#### 8.4.4 $O(2)$

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

**Satz.** Jede Matrix  $A \in O(2)$  hat die Form  $A = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} = R_\alpha$  oder  $A = \begin{pmatrix} \cos \alpha & \sin \alpha \\ \sin \alpha & -\cos \alpha \end{pmatrix} = S_{\alpha/2}$  mit  $\alpha \in (-\pi, \pi]$ .

**Bemerkung.** Wir betrachten  $\mathbb{R}^2$  bzgl. des Standard-Skalarprodukts.

- (i) Die normierten Vektoren aus  $\mathbb{R}^2$  sind genau die Vektoren der Form  $\begin{pmatrix} \cos \alpha \\ \sin \alpha \end{pmatrix}$  mit  $\alpha \in [-\pi, \pi]$ .
- (ii) Zu jedem normierten  $v = \begin{pmatrix} a \\ b \end{pmatrix} \in \mathbb{R}^2 \setminus \{0\}$  gibt es genau zwei normierte  $w \in \mathbb{R}^2$  mit  $v \perp w$ ; diese lauten  $w = \begin{pmatrix} -b \\ a \end{pmatrix}$  und  $w = \begin{pmatrix} b \\ -a \end{pmatrix}$ .

*Beweis.* (i) Sei  $v = \begin{pmatrix} a \\ b \end{pmatrix}$  normiert, d.h.  $a^2 + b^2 = 1$ . Sei  $\alpha = \angle(v, e_1)$ , d.h.  $\alpha \in [0, \pi]$  mit  $\cos \alpha = \langle v, e_1 \rangle = a$ . Wegen  $\sin^2 = 1 - \cos^2$  (Analysis) folgt  $\sin^2 \alpha = 1 - a^2 = b^2$ , also  $\sin \alpha = \pm b$ . Falls  $b < 0$ , dann ersetze  $\alpha$  durch  $-\alpha$ . So bekommen wir  $\alpha \in (-\pi, \pi]$  mit  $\cos \alpha = a$  und  $\sin \alpha = b$ . Umgekehrt ist jeder Vektor dieser Form normiert wegen  $\sin^2 + \cos^2 = 1$ .

(ii) Wegen  $\dim\langle v \rangle = 1$  ist  $\dim\langle v \rangle^\perp = 2 - 1 = 1$  und jeder 1-dimensionale  $\mathbb{R}$ -Vektorraum enthält genau 2 normierte Vektoren. Die angegebenen  $w$  sind verschieden und erfüllen offenbar  $v \perp w$ .  $\square$

*Beweis des Satzes.* Es seien  $s_1, s_2 \in \mathbb{R}^2$  die Spalten von  $A \in O(2)$ . Dann ist  $\|s_1\| = \|s_2\| = 1$  und  $s_1 \perp s_2$ . Nach Teil (i) der Bemerkung gibt es  $\alpha \in (-\pi, \pi]$  mit  $s_1 = \begin{pmatrix} \cos \alpha \\ \sin \alpha \end{pmatrix}$ . Nach Teil (ii) der Bemerkung folgt  $s_2 = \begin{pmatrix} -\sin \alpha \\ \cos \alpha \end{pmatrix}$  oder  $s_2 = \begin{pmatrix} \sin \alpha \\ -\cos \alpha \end{pmatrix}$ .  $\square$

**Folgerung** (Übersicht über die orthogonalen Endomorphismen von  $\mathbb{R}^2$ ).

Abbildung	Drehung um $\alpha \in (-\pi, \pi) \setminus \{0\}$	Drehung um $\pi$	Spiegelung
Matrix	$R_\alpha = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}$	$R_\pi = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$	$S_\alpha = \begin{pmatrix} \cos 2\alpha & \sin 2\alpha \\ \sin 2\alpha & -\cos 2\alpha \end{pmatrix}$
Determinante	1	1	-1
Spur	$2 \cos \alpha$	-2	0
char. Polynom	$X^2 - 2 \cos(\alpha)X + 1$	$(X + 1)^2$	$(X + 1)(X - 1)$
Eigenwerte	keine	-1, -1	-1, 1
Eigenvektorbasis	keine	$\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}$	$\begin{pmatrix} -\sin \alpha \\ \cos \alpha \end{pmatrix}, \begin{pmatrix} \cos \alpha \\ \sin \alpha \end{pmatrix}$
Diagonalform	nicht diag.bar	$\begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}$	$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}$
Orthogonalität	eigentlich	eigentlich	uneigentlich
selbstadjungiert	nein	ja	ja

**Beispiel.** Welche Art von Abbildung beschreibt  $A = \frac{1}{5} \begin{pmatrix} 4 & 3 \\ 3 & -4 \end{pmatrix}$ ?

1.  $A^t A = E_2 \Rightarrow A \in O(2)$ .
  2.  $\det A = -1 \Rightarrow A$  Spiegelung.
- Wie lautet die Spiegelachse?

Die Spiegelachse ist gerade der Eigenraum  $V_1(A) = \mathbb{L}(A - E, 0) = \langle \begin{pmatrix} 3 \\ 1 \end{pmatrix} \rangle$ .

*Übung.* Bei einer Spiegelung  $A \in O(2)$  ist die Spiegelachse gleich  $\langle v + Av \rangle$  für jedes  $v \in \mathbb{R}^2 \setminus \{0\}$ . Man prüfe die Aussage an dem obigen Beispiel.

## 8.5 Der Spektralsatz

In diesem Abschnitt sei  $V$  ein  $n$ -dimensionaler  $K$ -Vektorraum mit Skalarprodukt  $\langle -, - \rangle$ .



### 8.5.1 Unitäre und orthogonale Diagonalisierbarkeit

**Definition a.** Es seien  $A, B \in K^{n \times n}$ .

- (i) Es sei  $K = \mathbb{R}$ . Dann heißt  $A$  *orthogonal ähnlich* zu  $B$  genau dann, wenn  $T \in O_n(\mathbb{R})$  existiert mit  $B = T^{-1}AT$ .
- (ii) Es sei  $K = \mathbb{C}$ . Dann heißt  $A$  *unitär ähnlich* zu  $B$  genau dann, wenn  $T \in U_n(\mathbb{C})$  existiert mit  $B = T^{-1}AT$ .

**Definition b.** Es sei  $\varphi \in \text{End}_K(V)$  und  $A \in K^{n \times n}$ .

- (i)  $\varphi$  heißt *unitär (bzw. orthogonal) diagonalisierbar*, wenn  $K = \mathbb{C}$  (bzw.  $K = \mathbb{R}$ ) ist und eine Orthonormalbasis von  $V$  aus Eigenvektoren von  $\varphi$  existiert.
- (ii)  $A$  heißt *unitär (bzw. orthogonal) diagonalisierbar*, wenn  $A$  unitär (bzw. orthogonal) ähnlich zu einer Diagonalmatrix ist.

**Bemerkung.** (i) Es seien  $\mathcal{B}$  und  $\mathcal{B}'$  Orthonormalbasen von  $V$  und  $T := M_{\mathcal{B}}^{\mathcal{B}'}(\text{id}_V)$  die Basiswechselmatrix. Dann ist  $T$  unitär bzw. orthogonal.

- (ii) Es sei  $A \in K^{n \times n}$ . Dann ist  $A$  genau dann unitär bzw. orthogonal diagonalisierbar, wenn  $\varphi_A$  es ist.

*Beweis.* (i) Weil  $\text{id}_V$  unitär bzw. orthogonal ist, folgt mit Satz 8.4.1:

$$(M_{\mathcal{B}}^{\mathcal{B}'}(\text{id}_V))^{\dagger} = M_{\mathcal{B}'}^{\mathcal{B}}(\text{id}_V^{\dagger}) = M_{\mathcal{B}'}^{\mathcal{B}}(\text{id}_V^{-1}) = (M_{\mathcal{B}}^{\mathcal{B}'}(\text{id}_V))^{-1}.$$

(ii) Wir beweisen nur den unitären Fall. Der orthogonale geht analog. Eine Matrix  $T \in \text{GL}_n(\mathbb{C})$  ist genau dann unitär, wenn die Spalten von  $T$  eine Orthonormalbasis von  $\mathbb{C}^n$  bzgl. des Standard-Skalarprodukts bilden (siehe Bemerkung 8.4.3). Es sei  $A$  unitär diagonalisierbar, etwa  $T^{-1}AT = D$  mit einer Diagonalmatrix  $D$  und  $T \in U_n(\mathbb{C})$ . Dann bilden die Spalten von  $T$  eine Orthonormalbasis aus Eigenvektoren von  $\varphi_A$ . Ist umgekehrt  $\varphi_A$  orthogonal diagonalisierbar und  $T$  die Matrix, deren Spalten eine Orthonormalbasis aus Eigenvektoren von  $\varphi_A$  bilden, dann ist  $T$  unitär und  $T^{-1}AT$  eine Diagonalmatrix.  $\square$

### 8.5.2 Normale Endomorphismen

**Definition.** Es sei  $\varphi \in \text{End}_K(V)$  und  $A \in K^{n \times n}$ .

- (i)  $\varphi$  heißt *normal*, falls  $\varphi \circ \varphi^{\dagger} = \varphi^{\dagger} \circ \varphi$  ist.

(ii)  $A$  heißt *normal*, falls  $AA^\dagger = A^\dagger A$  ist.

**Beispiel.** (i)  $\begin{pmatrix} 3 & -2 \\ 2 & 3 \end{pmatrix} \in K^{2 \times 2}$  ist normal.

(ii)  $\begin{pmatrix} 1 & i \\ 1 & 1 \end{pmatrix} \in \mathbb{C}^{2 \times 2}$  ist normal.

(iii) Ist  $A \in K^{n \times n}$  unitär bzw. orthogonal, dann ist  $A$  normal.

(iv) Ist  $A \in K^{n \times n}$  hermitesch bzw. symmetrisch, dann ist  $A$  normal.

(v) Ist  $\varphi \in \text{End}_K(V)$  unitär bzw. orthogonal oder hermitesch, dann ist  $\varphi$  normal.

*Beweis.*

$$(i) \begin{pmatrix} 3 & -2 \\ 2 & 3 \end{pmatrix} \cdot \begin{pmatrix} 3 & -2 \\ 2 & 3 \end{pmatrix}^\dagger = \begin{pmatrix} 3 & -2 \\ 2 & 3 \end{pmatrix} \cdot \begin{pmatrix} 3 & 2 \\ -2 & 3 \end{pmatrix} = \begin{pmatrix} 13 & 0 \\ 0 & 13 \end{pmatrix}, \text{ und}$$

$$\begin{pmatrix} 3 & -2 \\ 2 & 3 \end{pmatrix}^\dagger \cdot \begin{pmatrix} 3 & -2 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} 3 & 2 \\ -2 & 3 \end{pmatrix} \cdot \begin{pmatrix} 3 & -2 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} 13 & 0 \\ 0 & 13 \end{pmatrix}.$$

$$(ii) \begin{pmatrix} 1 & i \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & i \\ 1 & 1 \end{pmatrix}^\dagger = \begin{pmatrix} 1 & i \\ 1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ -i & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1+i \\ 1-i & 2 \end{pmatrix}, \text{ und}$$

$$\begin{pmatrix} 1 & i \\ 1 & 1 \end{pmatrix}^\dagger \cdot \begin{pmatrix} 1 & i \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ -i & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & i \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1+i \\ 1-i & 2 \end{pmatrix}.$$

(iii) Dies folgt aus  $AA^{-1} = A^{-1}A$ .

(iv) Ist trivial.

(v) Analog zum Beweis von (iii) und (iv). □

Alle folgenden Ergebnisse für normale Endomorphismen und Matrizen gelten also insbesondere für unitäre, orthogonale und hermitesche Endomorphismen und Matrizen.

**Bemerkung a.** Es sei  $\varphi \in \text{End}_K(V)$ . Dann sind die folgenden Aussagen äquivalent:

(i)  $\varphi$  ist normal.

(ii)  $\langle \varphi^\dagger(v), \varphi^\dagger(w) \rangle = \langle \varphi(v), \varphi(w) \rangle$  für alle  $v, w \in V$ .

(iii)  $\|\varphi^\dagger(v)\| = \|\varphi(v)\|$  für alle  $v \in V$ .

*Beweis.* Wegen der Polarisationsformeln (siehe Folgerung 8.1.3(iii)) sind (ii) und (iii) äquivalent. Wegen  $(\varphi^\dagger)^\dagger = \varphi$  (siehe Bemerkung 8.4.1) gilt  $\langle v, \varphi(w) \rangle = \langle \varphi^\dagger(v), w \rangle$  für alle  $v, w \in V$ . Diese Identität werden wir im Folgenden zweimal benutzen.

(i)  $\Rightarrow$  (ii): Wir haben  $\langle \varphi^\dagger(v), \varphi^\dagger(w) \rangle = \langle \varphi(\varphi^\dagger(v)), w \rangle = \langle \varphi^\dagger(\varphi(v)), w \rangle = \langle \varphi(v), \varphi(w) \rangle$  für alle  $v, w \in V$ .

(ii)  $\Rightarrow$  (i): Wir haben  $\langle \varphi(\varphi^\dagger(v)), w \rangle = \langle \varphi^\dagger(v), \varphi^\dagger(w) \rangle = \langle \varphi(v), \varphi(w) \rangle = \langle \varphi^\dagger(\varphi(v)), w \rangle$  für alle  $v, w \in V$ . Daraus folgt  $\varphi(\varphi^\dagger(v)) = \varphi^\dagger(\varphi(v))$  für alle  $v \in V$ , also  $\varphi^\dagger \circ \varphi = \varphi \circ \varphi^\dagger$ .  $\square$

**Bemerkung b.** Es sei  $\varphi \in \text{End}_K(V)$  normal und  $a \in K$ . Dann ist  $V_a(\varphi) = V_{\bar{a}}(\varphi^\dagger)$ . Insbesondere ist  $a$  genau dann ein Eigenwert von  $\varphi$ , wenn  $\bar{a}$  ein Eigenwert von  $\varphi^\dagger$  ist.

*Beweis.* Wegen  $\varphi^\dagger \circ \varphi = \varphi \circ \varphi^\dagger$  ist  $V_a(\varphi)$  invariant unter  $\varphi^\dagger$ : Sei  $v \in V_a(\varphi)$ , also  $\varphi(v) = av$ . Dann ist  $\varphi(\varphi^\dagger(v)) = \varphi^\dagger(\varphi(v)) = \varphi^\dagger(av) = a\varphi^\dagger(v)$ , d.h.  $\varphi^\dagger(v) \in V_a(\varphi)$ . Damit gilt für alle  $v, w \in V_a(\varphi)$ :

$$\langle w, \varphi^\dagger(v) \rangle = \langle \varphi(w), v \rangle = a\langle w, v \rangle = \langle w, \bar{a}v \rangle.$$

Also ist  $\varphi^\dagger(v) = \bar{a}v$  für alle  $v \in V_a(\varphi)$ , d.h.  $V_a(\varphi) \leq V_{\bar{a}}(\varphi^\dagger)$ . Indem wir die Rolle von  $\varphi$  und  $\varphi^\dagger$  vertauschen, erhalten wir  $V_{\bar{a}}(\varphi^\dagger) \leq V_a((\varphi^\dagger)^\dagger) = V_a(\varphi)$ . Die zweite Aussage folgt aus der ersten.  $\square$

### 8.5.3 Der Spektralsatz

In diesem Unter-abschnitt sei  $\varphi \in \text{End}_K(V)$ .

**Bemerkung a.** Es sei  $U \leq V$  so dass  $U$  und  $U^\perp$  invariant unter  $\varphi$  sind. Dann sind  $U$  und  $U^\perp$  auch  $\varphi^\dagger$ -invariant, und es gilt

$$(\varphi_U)^\dagger = (\varphi^\dagger)_U.$$

Ist außerdem  $\varphi$  normal oder hermitesch, dann ist auch  $\varphi_U$  normal bzw. hermitesch.

*Beweis.* Wir verwenden  $V = U \oplus U^\perp$  und  $(U^\perp)^\perp = U$ . Seien  $u \in U$  und  $u' \in U^\perp$ . Dann ist auch  $\varphi(u) \in U$  und es gilt  $0 = \langle \varphi(u), u' \rangle = \langle u, \varphi^\dagger(u') \rangle$ . Da  $u \in U$  beliebig war, ist  $\varphi^\dagger(u') \in U^\perp$ . Damit ist  $\varphi^\dagger(U^\perp) \leq U^\perp$ . Da auch  $U^\perp$  invariant unter  $\varphi$  ist, ergibt sich  $\varphi^\dagger(U) = \varphi^\dagger((U^\perp)^\perp) \leq (U^\perp)^\perp = U$ .

Die zweite Aussage ergibt sich aus der Eindeutigkeit der adjungierten Abbildung und der Gleichung  $\langle \varphi(u), y \rangle = \langle u, \varphi^\dagger(y) \rangle$  für alle  $u, y \in U$ .

Die dritte Behauptung schließlich folgt aus der zweiten: Ist  $\varphi$  normal, dann folgt  $\varphi_U \circ (\varphi_U)^\dagger = \varphi_U \circ (\varphi^\dagger)_U = (\varphi \circ \varphi^\dagger)_U = (\varphi^\dagger \circ \varphi)_U = (\varphi^\dagger)_U \circ \varphi_U = (\varphi_U)^\dagger \circ \varphi_U$ , also ist auch  $\varphi_U$  normal. Ist  $\varphi$  hermitesch, so erhalten wir  $(\varphi_U)^\dagger = (\varphi^\dagger)_U = \varphi_U$ , also ist auch  $\varphi_U$  hermitesch.  $\square$

**Lemma.** *Es sei  $A \in K^{n \times n}$  hermitesch. Dann ist  $\chi_A = \prod_{i=1}^n (X - a_i)$  mit  $a_i \in \mathbb{R}$  für alle  $1 \leq i \leq n$  (d.h.  $\chi_A$  zerfällt über  $\mathbb{R}$  in Linearfaktoren).*

*Beweis.* Wir fassen  $A$  als Matrix in  $\mathbb{C}^{n \times n}$  auf (auch wenn  $K = \mathbb{R}$  ist) und versehen  $\mathbb{C}^n$  mit dem Standard-Skalarprodukt. Dann ist die Standard-Basis von  $\mathbb{C}^n$  eine Orthonormalbasis und  $\varphi_A : \mathbb{C}^n \rightarrow \mathbb{C}^n$  hermitesch (siehe Satz 8.4.1). Aus dem Fundamentalsatz der Algebra folgt  $\chi_{\varphi_A} = \chi_A = \prod_{i=1}^n (X - a_i)$  mit  $a_i \in \mathbb{C}$  für alle  $1 \leq i \leq n$ . Da  $\varphi_A$  hermitesch ist, ist  $\varphi_A$  auch normal (siehe Beispiel 8.5.2(v)), also ist  $V_{a_i}(\varphi_A) = V_{\bar{a}_i}(\varphi_A^\dagger) = V_{\bar{a}_i}(\varphi_A)$  für alle  $1 \leq i \leq n$  nach Bemerkung 8.5.2b. Sei nun  $1 \leq i \leq n$  und  $0 \neq v \in V_{a_i}(\varphi_A)$ . Dann ist  $a_i v = \varphi_A(v) = \bar{a}_i v$ . Also ist  $a_i = \bar{a}_i$ , d.h.  $a_i \in \mathbb{R}$ .  $\square$

**Folgerung a.** *Ist  $A \in \mathbb{R}^{n \times n}$  symmetrisch, so zerfällt  $\chi_A$  über  $\mathbb{R}$  vollständig in Linearfaktoren.*

**Satz** (Spektralsatz).

- (i) *Ist  $\varphi$  normal, und zerfällt  $\chi_\varphi$  in Linearfaktoren, dann existiert eine Orthonormalbasis von  $V$  aus Eigenvektoren von  $\varphi$ .*
- (ii) *Ist  $\varphi$  hermitesch, dann existiert eine Orthonormalbasis von  $V$  aus Eigenvektoren von  $\varphi$ .*

*Beweis.* Aufgrund des obigen Lemmas genügt es, Aussage (i) zu beweisen. Wir führen den Beweis durch Induktion über  $n$ , wobei der Fall  $n = 1$  trivial ist. Es sei als  $n > 1$  und  $a_1 \in K$  ein Eigenwert von  $\varphi$ , der auch Voraussetzung existiert. Sei  $v_1$  ein normierter Eigenvektor zu  $a_1$ . Setze  $U := \langle v_1 \rangle$ . Dann ist  $V = U \oplus U^\perp$  und  $\dim U^\perp = n - 1$ . Offensichtlich ist  $U$  invariant unter  $\varphi$ . Wir zeigen, dass auch  $U^\perp$  invariant unter  $\varphi$  ist. Sei dazu  $u' \in U^\perp$ . Dann ist  $\langle \varphi(u'), v_1 \rangle = \langle u', \varphi^\dagger(v_1) \rangle = \langle u', \bar{a}_1 v_1 \rangle = a_1 \langle u', v_1 \rangle = 0$ , wobei sich das zweite Gleichheitszeichen aus Bemerkung 8.5.2b ergibt. Also ist  $\varphi(u') \in \{v_1\}^\perp = U^\perp$ . Nach Bemerkung a ist  $\varphi_U$  normal. Weil das charakteristische Polynom von  $\varphi_U$  ein Teiler von  $\chi_\varphi$  ist, zerfällt ersteres in Linearfaktoren. Nach Induktionsvoraussetzung besitzt  $U^\perp$  eine Orthonormalbasis  $(v_2, \dots, v_n)$  aus Eigenvektoren von  $\varphi_U$ . Also ist  $(v_1, \dots, v_n)$  eine Orthonormalbasis von  $V$  aus Eigenvektoren von  $\varphi$ .  $\square$

**Folgerung b.** *Zu jeder reellen symmetrischen Matrix gibt es eine Eigenvektorbasis von  $\mathbb{R}^n$ , die gleichzeitig Orthonormalbasis bzgl. des Standard-Skalarproduktes von  $\mathbb{R}^n$  ist.*

**Bemerkung b.** Man kann sogar die Umkehrung dieser Folgerung des Spektralsatzes zeigen: Gibt es zu  $A \in \mathbb{R}^{n \times n}$  eine orthonormale Eigenvektorbasis, so ist  $A$  symmetrisch.

**Folgerung c.** Es sei  $A \in \mathbb{R}^{n \times n}$  symmetrisch. Dann ist  $A$  genau dann positiv definit, wenn alle Eigenwerte von  $A$  positiv sind.

*Beweis.* Es sei  $\mathcal{B}$  eine orthonormale Eigenvektorbasis zu  $A$  (existiert nach Folgerung b). Die Basiswechselmatrix zur Standard-Basis bezeichnen wir mit  $T$ , d.h. die Spalten von  $T$  sind die Elemente von  $\mathcal{B}$ . Dann ist  $T$  orthogonal nach Bemerkung 8.4.3(iii), also  $T^{-1} = T^t$ . Da  $\mathcal{B}$  Eigenvektorbasis ist, ist  $D := T^{-1}AT$  eine Diagonalmatrix, deren Diagonaleinträge genau die Eigenwerten von  $A$  sind. Nach Teil (ii) von Folgerung 8.3.3 ist  $A$  genau dann positiv definit, wenn  $D$  positiv definit ist. Nach Beispiel 8.3.3(ii) also genau dann, wenn alle Eigenwerte von  $A$  positiv sind.  $\square$

*Übung.* Man überprüfe dieses Kriterium für die positive Definitheit anhand von Beispiel 8.3.3.

**Bemerkung c.** Jede normale Matrix  $A \in \mathbb{C}^{n \times n}$  ist unitär diagonalisierbar. Jede symmetrische Matrix  $A \in \mathbb{R}^{n \times n}$  ist orthogonal diagonalisierbar,

*Beweis.* Dies folgt sofort aus dem Spektralsatz und Bemerkung 8.5.1.  $\square$

*Übung.* Zeige, dass jede orthogonal diagonalisierbare reelle Matrix symmetrisch ist.

### 8.5.4 Die Singulärwertzerlegung

In diesem Unter-abschnitt sei  $W$  eine  $m$ -dimensionaler  $K$ -Vektorraum mit Skalarprodukt. die Bedeutung von  $V$  bleibt wie im gesamten Abschnitt 8.5 beibehalten. Wir erinnern daran, dass nach Lemma 8.5.3 das charakteristische Polynom einer hermiteschen  $(n \times n)$ -Matrix über  $K$  vollständig in Linearfaktoren zerfällt, und dass alle seine Nullstellen reell sind. Die analoge Aussage gilt dann natürlich auch für hermitesche Endomorphismen von  $V$ .

**Bemerkung a.** Es sei  $\varphi \in \text{Hom}_K(V, W)$  und  $A \in K^{m \times n}$ . Dann gelten:

- (i)  $\varphi^\dagger \circ \varphi \in \text{End}_K(V)$  ist hermitesch und die Eigenwerte von  $\varphi^\dagger \circ \varphi$  sind nicht-negative reelle Zahlen.
- (ii)  $A^\dagger A \in K^{n \times n}$  ist hermitesch und die Eigenwerte von  $A^\dagger A$  sind nicht-negative reelle Zahlen.

*Beweis.* Aussage (ii) folgt aus Aussage (i) durch Übergang zu einer Abbildungsmatrix von  $\varphi$  bzgl. Orthonormalbasen von  $V$  und  $W$ . Unter Verwendung von Übung und Bemerkung 8.4.1 erhalten wir  $(\varphi^\dagger \circ \varphi)^\dagger = \varphi^\dagger \circ (\varphi^\dagger)^\dagger = \varphi^\dagger \circ \varphi$ . Also ist  $\varphi^\dagger \circ \varphi$  hermitesch. Sei  $a \in K$  ein Eigenwert von  $\varphi^\dagger \circ \varphi$  und  $v$

ein zugehöriger normierter Eigenvektor. Dann gilt  $\|\varphi(v)\|^2 = \langle \varphi(v), \varphi(v) \rangle = \langle \varphi^\dagger(\varphi(v)), v \rangle = a \langle v, v \rangle = a$ . Damit ist auch die erste Aussage von (i) bewiesen, denn  $\|\varphi(v)\|^2 \geq 0$ .  $\square$

**Definition a.** (i) Es sei  $\varphi \in \text{Hom}_K(V, W)$  und  $A \in K^{m \times n}$ . Die Eigenwerte von  $\varphi^\dagger \circ \varphi$  bzw.  $A^\dagger A$  seien  $a_1, a_2, \dots, a_n$  mit  $a_1 \geq a_2 \geq \dots \geq a_r > 0 = a_{r+1} = \dots = a_n$ . Dann heißt  $(\sqrt{a_1}, \dots, \sqrt{a_r})$  das *Singulärwerttupel* von  $\varphi$  bzw.  $A$  und wird mit  $\sigma(\varphi)$  bzw.  $\sigma(A)$  bezeichnet.

(ii) Es seien  $m, n \in \mathbb{N}$  und  $r \in \mathbb{N}_0$  mit  $r \leq \min\{m, n\}$ , sowie  $\sigma = (\sigma_1, \dots, \sigma_r) \in \mathbb{R}^{1 \times r}$  mit  $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$ . Dann setzen wir

$$\Sigma_\sigma := \Sigma_{\sigma_1, \dots, \sigma_r} := \left( \begin{array}{ccc|c} \sigma_1 & & 0 & \\ & \ddots & & 0 \\ 0 & & \sigma_r & \\ \hline & & 0 & 0 \end{array} \right) \in \mathbb{R}^{m \times n}.$$

**Bemerkung b.** Es seien  $m, n, r$  und  $\sigma$  wie in Teil (ii) von Definition a. Dann ist  $\sigma(\Sigma_\sigma) = \sigma$ .

*Beweis.* Wir haben

$$\Sigma_\sigma^\dagger \Sigma_\sigma = \Sigma_\sigma^t \Sigma_\sigma = \left( \begin{array}{ccc|c} \sigma_1^2 & & 0 & \\ & \ddots & & 0 \\ 0 & & \sigma_r^2 & \\ \hline & & 0 & 0 \end{array} \right) \in \mathbb{R}^{n \times n}.$$

Daraus folgt die Behauptung.  $\square$

**Definition b.** Es seien  $A, B \in K^{m \times n}$ .

- (i) Es sei  $K = \mathbb{C}$ . Dann heißen  $A$  und  $B$  *unitär äquivalent*, wenn unitäre Matrizen  $S \in U_m(\mathbb{C})$  und  $T \in U_n(\mathbb{C})$  existieren mit  $B = SAT$ .
- (ii) Es sei  $K = \mathbb{R}$ . Dann heißen  $A$  und  $B$  *orthogonal äquivalent*, wenn orthogonale Matrizen  $S \in O_m(\mathbb{R})$  und  $T \in O_n(\mathbb{R})$  existieren mit  $B = SAT$ .

**Satz** (Singulärwertzerlegung). *Es sei  $A \in K^{m \times n}$ . Dann ist  $A$  unitär bzw. orthogonal äquivalent zu  $\Sigma_{\sigma(A)}$ .*

*Beweis.* Wir versehen  $K^n$  und  $K^m$  mit dem Standard-Skalarprodukt und betrachten  $\varphi_A : K^n \rightarrow K^m$ . Es sei  $\mathcal{B} = (v_1, \dots, v_n)$  eine Orthonormalbasis von  $K^n$  aus Eigenvektoren von  $A^\dagger A$  (die Existenz dieser Basis folgt aus dem Spektralsatz und Bemerkung a). Es sei  $a_i$  der Eigenwert zu  $v_i$  für  $1 \leq i \leq n$ . Wir können die Nummerierung außerdem so wählen, dass gilt:  $a_1 \geq a_2 \geq \dots \geq a_r > 0 = a_{r+1} = \dots = a_n$ . Dann ist  $\sigma(A) = (\sqrt{a_1}, \dots, \sqrt{a_r})$ . Für  $1 \leq j \leq r$  setzen wir

$$w_j := \frac{Av_j}{\sqrt{a_j}}.$$

Dann gilt für alle  $1 \leq i, j \leq r$ :

$$\begin{aligned} \langle w_i, w_j \rangle &= \frac{1}{\sqrt{a_i a_j}} \langle Av_i, Av_j \rangle \\ &= \frac{1}{\sqrt{a_i a_j}} \langle v_i, A^\dagger A v_j \rangle \\ &= \frac{a_j}{\sqrt{a_i a_j}} \langle v_i, v_j \rangle = \delta_{ij}. \end{aligned}$$

Also bildet  $(w_1, \dots, w_r)$  ein Orthonormalsystem in  $K^m$ , das wir zu einer Orthonormalbasis  $\mathcal{C}$  von  $K^m$  ergänzen. Mit diesen Wahlen ist  $M_{\mathcal{C}}^{\mathcal{B}}(\varphi_A) = \Sigma_{\sigma(A)}$ . Bzgl. der Standard-Basen von  $K^n$  bzw.  $K^m$  hat  $\varphi_A$  die Abbildungsmatrix  $A$ . Da  $\mathcal{B}$  und  $\mathcal{C}$  Orthonormalbasen sind, sind die Basiswechselmatrizen  $T$  in  $K^n$  und  $S$  in  $K^m$  unitär bzw. orthogonal (siehe Teil (i) von Bemerkung 8.5.1). Aus dem Basiswechselsatz (siehe Satz 6.4.4) folgt  $\Sigma_{\sigma(A)} = M_{\mathcal{C}}^{\mathcal{B}}(\varphi_A) = SAT$ , was zu zeigen war.  $\square$

**Folgerung a.** *Es seien  $A, B \in K^{m \times n}$ . Dann sind  $A$  und  $B$  genau dann unitär bzw. orthogonal äquivalent, wenn  $\sigma(A) = \sigma(B)$  ist.*

*Beweis.* Ist  $\sigma(A) = \sigma(B)$ , dann sind  $A$  und  $B$  nach obigem Satz zu  $\Sigma_{\sigma(A)} = \Sigma_{\sigma(B)}$  unitär bzw. orthogonal äquivalent, also auch zueinander.

Es seien  $A$  und  $B$  unitär äquivalent. Dann existieren  $S \in U_m(K)$  und  $T \in U_n(K)$  mit  $B = SAT$ . Dann ist  $B^\dagger B = (SAT)^\dagger SAT = T^\dagger A^\dagger S^\dagger SAT = T^\dagger A^\dagger AT$ , weil  $S^\dagger = S^{-1}$  ist. Damit sind  $A^\dagger A$  und  $B^\dagger B$  unitär ähnlich, und es folgt  $\sigma(A) = \sigma(B)$ .  $\square$

**Folgerung b.** *Es sei  $A \in K^{m \times n}$ . Dann ist  $\text{Rg}(A) = \text{Rg}(A^\dagger A)$ .*

**Folgerung c.** *Die Menge der unitären bzw. orthogonalen Äquivalenzklassen der  $(m \times n)$ -Matrizen über  $K$  steht in Bijektion zu  $\{(\sigma_1, \dots, \sigma_r) \in \mathbb{R}^{1 \times r} \mid \sigma_1 \geq \dots \geq \sigma_r > 0\}$ .*

## 8.6 Approximation

In diesem Unter-Abschnitt sei stets  $K = \mathbb{R}$ . In einem euklidischen Vektorraum  $V$  seien  $M \subset V$  und  $v \in V$  gegeben. Es wird ein Element  $x \in M$  gesucht, dass eine „beste Näherung“ an  $v$  darstellt.

### 8.6.1 Winkelapproximation

Vorausgesetzt, dass  $v$  und alle  $x \in M$  normiert sind, kann man nach dem von  $v$  und  $x$  eingeschlossenen Winkel approximieren. Nach der Cauchy-Schwarz-Ungleichung gilt für normierte Vektoren  $-1 \leq \langle v, x \rangle \leq 1$ , wobei  $\langle v, x \rangle = 1$  genau dann, wenn  $v = w$ ,  $\langle v, x \rangle = 0$  genau dann, wenn  $v \perp w$ , und  $\langle v, x \rangle = -1$  genau dann, wenn  $v = -w$  ist. Für eine beste Approximation ist daher  $\langle v, x \rangle$  zu maximieren.

**Beispiel a.** Gegeben seien  $n$  Dokumente  $D_1, \dots, D_n$  und  $m$  Terme  $T_1, \dots, T_m$ . Wir definieren zu Dokument  $j$  den Vektor

$$d'_j = \begin{pmatrix} d'_{1j} \\ \vdots \\ d'_{mj} \end{pmatrix} \in \mathbb{R}^m, \text{ wobei } d'_{ij} = \begin{cases} 1 & \text{falls } T_i \text{ in } D_j \text{ vorkommt,} \\ 0 & \text{sonst} \end{cases}$$

und normieren zu  $d_j := \frac{d'_j}{\|d'_j\|}$ .

Aus einer Suchanfrage nach den Termen  $T_{i_1}, \dots, T_{i_l}$  wird entsprechend ein Suchvektor

$$q' = \begin{pmatrix} q'_1 \\ \vdots \\ q'_m \end{pmatrix} \in \mathbb{R}^m, \text{ wobei } q'_i = \begin{cases} 1 & \text{falls } i \in \{i_1, \dots, i_l\}, \\ 0 & \text{sonst} \end{cases}$$

gebildet und zu  $q := \frac{q'}{\|q'\|}$  normiert.

Das am besten zur Suchanfrage passende Dokument ist dann  $D_j$  für dasjenige  $j$ , für das  $\langle q, d_j \rangle$  maximal wird.

Man beachte, dass man für das Standard-Skalarprodukt alle  $\langle q, d_j \rangle$  durch eine einzelne Matrixmultiplikation errechnen kann. Schreibt man  $d_1, \dots, d_n$  in die Spalten einer Matrix  $D$ , so ist  $D$  eine Markov-Matrix und es gilt

$$q^t \cdot D = (\langle q, d_1 \rangle, \dots, \langle q, d_n \rangle) \in \mathbb{R}^{1 \times n}.$$

**Beispiel b.** Hat man analoge Audiosignale statt Textdokumenten, so kann man diese als Vektoren  $d'_j \in C^0([0, 1])$  auffassen (z.B. bei Abspiel-Länge 1s). Unter Verwendung des Skalarproduktes  $\langle f, g \rangle = \int_0^1 f(x)g(x)dx$  könnte man analog zum vorherigen Beispiel verfahren.



### 8.6.2 Abstandsapproximation

Im Allgemeinen, d.h. wenn die Vektoren nicht normiert sind, versucht man den Abstand  $\|v - x\|$  für  $x \in M$  zu minimieren.

**Definition.** Man nennt  $d(v, M) := \inf\{\|v - x\| \mid x \in M\}$  den *Abstand* von  $v$  zu  $M$ .

Falls  $M$  ein Unterraum ist, so wird die beste Approximation gerade von der Projektion geliefert:

**Satz.** Es seien  $U \leq V$  und  $v \in V$ . Dann gilt für alle  $u \in U$ :

$$\|v - u\| \geq \|v - \text{pr}_U(v)\|.$$

Insbesondere ist  $d(v, U) = \|v - \text{pr}_U(v)\|$ .

*Beweis.* Setze  $u_0 := \text{pr}_U(v)$ . Dann ist  $v - u_0 \in U^\perp$ . Für jedes  $u \in U$  gilt somit  $(v - u_0) \perp (u_0 - u)$ , also nach Pythagoras:

$$\|v - u\|^2 = \|v - u_0\|^2 + \|u_0 - u\|^2 \geq \|v - u_0\|^2.$$

Daraus folgt die Behauptung.  $\square$

**Bemerkung.** Die Wahl des Skalarproduktes bestimmt die Definition des Abstandes und legt damit das Kriterium fest, nach welchem approximiert wird. Verschiedene Skalarprodukte liefern im Allgemeinen verschiedene beste Approximationen.

**Beispiel.** Ist  $\mathcal{B} = (v_1, \dots, v_n)$  eine Orthogonalbasis von  $V$  und  $U = \langle v_1, \dots, v_r \rangle$  so lässt sich die  $\text{pr}_U$  in Koordinaten bzgl.  $\mathcal{B}$  ohne Rechnung sofort angeben:

$$\kappa_{\mathcal{B}}(v) = (x_1, \dots, x_n)^t \Rightarrow \kappa_{\mathcal{B}}(\text{pr}_U(v)) = (x_1, \dots, x_r, 0, \dots, 0)^t.$$

*Übung.* Wie lautet die Abbildungsmatrix von  $\text{pr}_U$  bzgl. der angegebenen Basis in obigem Beispiel?

### 8.6.3 Datenkompression

Die Abstandsapproximationen durch Elemente eines möglichst kleinen Unterraums  $U$  kann zur verlustbehafteten Datenkompression verwendet werden. Sei dazu eine Orthogonalbasis  $\mathcal{B} = (v_1, \dots, v_n)$  von  $V$  gewählt. Um einen gegebenen "Datensatz"  $v \in V$  mit "Genauigkeit"  $1 \leq r \leq n$  zu komprimieren, wird  $\kappa_{\mathcal{B}}(v) = (x_1, \dots, x_n)^t$  berechnet und nur das Tupel  $(x_1, \dots, x_r)$  gespeichert bzw. übertragen. Gemäß Beispiel 8.6.2 definiert dieses Tupel die beste

Approximation aus dem Unterraum  $U_r := \langle v_1, \dots, v_r \rangle$  an  $v$ . Bei gewählter Genauigkeit  $r$  ergibt sich die Kompressionsrate  $r/n$ . Die Schwierigkeit liegt darin, die Orthogonalbasis  $\mathcal{B}$  so zu wählen, dass sich die für die gegebene Anwendung signifikante Information in den ersten Koordinaten sammelt und die Information aus den letzten Koordinaten verzichtbar ist.

**Beispiel a.** Wir fassen Binärzahlen der Länge  $n$  bit als Vektoren aus  $\mathbb{R}^n$  auf, etwa die Zahl  $1011 \dots 0$  als  $(1, 0, 1, 1, \dots, 0)^t$ . Die “least significant bits” stehen in der Binärzahl rechts, die “most significant bits” links. Betrachtet man das Standard-Skalarprodukt und wählt als Orthogonalbasis die Standardbasis, so läuft das beschriebene Verfahren darauf hinaus, nur die  $r$  “most significant bits” zu speichern.

**Beispiel b.** Bei der *diskreten Kosinustransformation* wird die Orthogonalbasis  $\mathcal{B} = (v_0, \dots, v_{n-1})$  von  $\mathbb{R}^n$  bzgl. Standard-Skalarprodukt gewählt, die definiert ist durch:

$$v_i := \begin{pmatrix} \cos\left(\frac{1}{2n} \cdot i\pi\right) \\ \cos\left(\frac{3}{2n} \cdot i\pi\right) \\ \vdots \\ \cos\left(\frac{2n-1}{2n} \cdot i\pi\right) \end{pmatrix}$$

Bezeichnet  $T_n := {}^{\mathcal{E}}T^{\mathcal{B}}$  die zugehörige Basiswechselmatrix, so ist z.B.

$$T_2 = \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}, \quad T_3 = \begin{pmatrix} 1 & \sqrt{3}/2 & 1/2 \\ 1 & 0 & -1 \\ 1 & -\sqrt{3}/2 & 1/2 \end{pmatrix}.$$

Unter Verwendung geeigneter trigonometrischer Identitäten kann man nachrechnen, dass  $\mathcal{B}$  tatsächlich eine Orthogonalbasis ist ( $\langle v_i, v_j \rangle = 0$  für alle  $i \neq j$ ) und ausserdem

$$\|v_i\| = \begin{cases} \sqrt{n} & \text{falls } i = 0, \\ \sqrt{\frac{n}{2}} & \text{falls } i > 0. \end{cases}$$

Die Idee, die hinter der diskreten Kosinustransformation steckt, wird ersichtlich, wenn man folgende kontinuierliche Variante betrachtet.

**Beispiel c.** Es sei  $V = C^0([0, 1])$  mit dem Skalarprodukt

$$\langle f, g \rangle = \int_0^\pi f(x)g(x)dx.$$

Für jedes  $n \in \mathbb{N}$  ist ein Orthogonalsystem  $\mathcal{B}_n = (f_0, \dots, f_{n-1})$  definiert durch  $f_i(x) := \cos(ix)$ . Unter Verwendung geeigneter trigonometrischer Identitäten kann man nachrechnen, dass  $\mathcal{B}$  tatsächlich ein Orthogonalsystem ist ( $\langle f_i, f_j \rangle = 0$  für alle  $i \neq j$ ) und ausserdem

$$\|f_i\| = \begin{cases} \sqrt{\pi} & \text{falls } i = 0, \\ \sqrt{\frac{\pi}{2}} & \text{falls } i > 0. \end{cases}$$

Die Projektion auf die Unterräume  $U_n = \langle f_0, \dots, f_{n-1} \rangle$  bedeutet, eine gegebene Funktion  $f$  durch eine Überlagerung (Linearkombination) von Kosinusfunktion verschiedener Frequenzen zu approximieren.

**Bemerkung a.** Man sagt, bei der Kosinustransformation wird vom “Zeitraum” in den “Frequenzraum” transformiert. (besser wäre: von der “Zeitbasis” in die “Frequenzbasis”). Verschiedene Modifikationen der diskreten Kosinustransformation werden bei Audio-Codecs verwendet. Eine zweidimensionale Variante ist die Grundlage des JPEG-Verfahrens.



# Literaturverzeichnis

- [1] M. Aigner. *Diskrete Mathematik*. Vieweg, 2004.
- [2] H. Anton. *Lineare Algebra*. Spektrum, 1995.
- [3] A. Beutelspacher. *Lineare Algebra*. Vieweg, 2003.
- [4] G. Fischer. *Lineare Algebra*. Vieweg, 2005.
- [5] S. Teschl G. Teschl. *Mathematik für Informatiker, Band 1*. Springer, 2007.
- [6] K. Jänich. *Lineare Algebra*. Springer, 2003.
- [7] A. Steger. *Diskrete Strukturen*. Springer, 2001.
- [8] K. Meyberg und P. Vachenauer. *Höhere Mathematik*. Springer, 2001.