

BTU Cottbus-Senftenberg

Dozent: Prof. G. Averkov

Semester: WiSe 20/21 – SoSe 21

---

# Lineare Algebra

---

24. April 2022 (10:04 Uhr)

# Inhaltsverzeichnis

<b>1 Mathematische Grundbegriffe</b>	<b>22</b>
1.1 Aussagen . . . . .	22
1.1.1 Aussage . . . . .	22
1.1.2 Logische Verknüpfungen . . . . .	23
1.2 Mengen . . . . .	25
1.2.1 Menge . . . . .	25
1.2.2 Zahlenmengen . . . . .	26
1.2.3 Definition durch eine Bedingung . . . . .	28
1.2.4 Die leere Menge . . . . .	28
1.2.5 Potenzmenge . . . . .	29
1.2.6 Mengenoperationen . . . . .	29
1.2.7 Disjunkte Mengen . . . . .	29
1.3 Tupel . . . . .	30

1.4 Abbildungen . . . . .	32
1.4.1 Abbildung . . . . .	32
1.4.2 Bild und Urbild . . . . .	33
1.4.3 Injektivitat, Surjektivitat und Bijektivitat . . . . .	35
1.4.4 Umkehrfunktion . . . . .	36
1.4.5 Komposition . . . . .	37
1.4.6 Identische Abbildung . . . . .	37
1.4.7 Vereinigung und Durchschnitt einer indexierten Men- genfamilie . . . . .	38
1.4.8 Summen und Produkte . . . . .	39
1.5 Pradikate . . . . .	41
1.5.1 Pradikat . . . . .	41
1.5.2 Quantoren . . . . .	41
1.6 Relationen . . . . .	43
1.6.1 Relation . . . . .	43

1.6.2 Äquivalenzrelation . . . . .	44
<b>2 Ausgewählte algebraische Strukturen</b>	<b>48</b>
2.1 Gruppen . . . . .	48
2.1.1 Binäroperationen . . . . .	48
2.1.2 Gruppe . . . . .	50
2.1.3 Das linke und das rechte Inverse . . . . .	55
2.1.4 Eindeutigkeit des linken neutralen Elements . . . . .	55
2.1.5 Neutralität von links und rechts . . . . .	56
2.1.6 Eindeutigkeit des inversen Elements . . . . .	57
2.1.7 Zusammenfassung der Gruppeneigenschaften . . . . .	57
2.1.8 Das Inverse des Produkts und der Potenz . . . . .	59
2.1.9 Zyklische Gruppen . . . . .	60
2.1.10 Untergruppen . . . . .	61

2.2 Ringe . . . . .	62
2.2.1 Ring . . . . .	62
2.2.2 Polynomring in einer Unbestimmten . . . . .	64
2.2.3 Restklassenringe . . . . .	71
2.3 Körper . . . . .	73
2.3.1 Körper . . . . .	73
2.3.2 Restklassenkörper . . . . .	74
2.3.3 Körper Komplexer Zahlen . . . . .	75
<b>A Vollständige Induktion</b>	<b>79</b>
<b>B Eliminationsverfahren zur Lösung von LGS</b>	<b>82</b>
B.1 Lineare Gleichungssysteme (LGS) . . . . .	82
B.2 Elementartransformationen eines LGS . . . . .	84
B.2.1 Gauß-Verfahren . . . . .	86
B.2.2 Gauß-Jordan-Verfahren . . . . .	93

<b>3 Vektorräume</b>	<b>95</b>
3.1 Vektorraum (VR) . . . . .	95
3.1.1 $\mathbb{K}^n$ und $\mathbb{K}^X$ . . . . .	95
3.1.2 Vektorraum über $\mathbb{K}$ . . . . .	97
3.2 Untervektorräume (UVR) . . . . .	99
3.2.1 Untervektorraum . . . . .	99
3.2.2 Kriterium für Untervektorräume . . . . .	101
3.2.3 Durchschnitt von Untervektorräumen . . . . .	102
3.2.4 Vereinigung von Untervektorräumen . . . . .	102
3.3 Linearkombinationen . . . . .	103
3.3.1 Linearkombination . . . . .	103
3.3.2 Lineare Unabhängigkeit und Abhängigkeit . . . . .	105
3.3.3 Kriterien für lineare Unabhängigkeit . . . . .	107
3.3.4 Eigenschaften der linearen Abhängigkeit . . . . .	108

<b>3.4 Basis und Dimension</b>	110
3.4.1 Erzeugendensysteme und Basen	110
3.4.2 Charakterisierungen der Basis-Eigenschaft und die Fol-	
gerungen daraus	110
3.4.3 Basisauswahlsatz	113
3.4.4 Austauschlemma	114
3.4.5 Austauschsatz	116
3.4.6 Dimension	119
3.4.7 Dimension von Untervektorräumen	120
3.4.8 Ergänzung zu einer Basis	122
<b>3.5 Rang</b>	123
3.5.1 Matrizen und ihr Rang	123
3.5.2 Elementartransformationen von Vektorsystemen	125
3.5.3 Der Rang der transponierten Matrix	128

3.6	Summen von Vektorräumen . . . . .	134
3.6.1	Summe von Vektorräumen . . . . .	134
3.6.2	Dimensionsformel für Summe von zwei Untervektor-	
	räumen . . . . .	135
3.6.3	Direkte Summe . . . . .	137
3.6.4	Charakterisierung der direkten Summe . . . . .	139
3.6.5	Direkter Summand . . . . .	140
3.6.6	Direkte Summe endlich vieler Vektorräume . . . . .	140
3.7	Projektive Räume * . . . . .	143
3.7.1	Projektive Räume . . . . .	144
3.7.2	Zwei projektive Punkte definieren genau eine Gerade	145
3.7.3	Zwei projektive Geraden einer projektiven Ebene schnei-	
	den sich in genau einem Punkt . . . . .	146

<b>4 Lineare Abbildungen</b>	<b>151</b>
4.1 Beispiele von linearen Abbildungen . . . . .	151
4.1.1 $90^\circ$ -Drehung . . . . .	152
4.1.2 Projektion in $\mathbb{R}^3$ . . . . .	152
4.1.3 Projektion von $\mathbb{R}^3$ nach $\mathbb{R}^2$ . . . . .	152
4.1.4 Punktspiegelung . . . . .	152
4.1.5 Spiegelung an einer Geraden . . . . .	153
4.1.6 Scherung . . . . .	153
4.1.7 Streckung . . . . .	153
4.1.8 Zyklische Verschiebung der Komponenten . . . . .	153
4.2 Lineare Abbildungen für allgemeine Vektorräume . . . . .	154
4.2.1 Lineare Abbildung . . . . .	154
4.2.2 Lineare Abbildungen und die Grundbegriffe für Vek- torräume . . . . .	155
4.2.3 Komposition von linearen Abbildungen . . . . .	158

4.2.4	Vektorräume der linearen Abbildungen . . . . .	159
4.2.5	Lineare Abbildungen eines Vektorraums . . . . .	160
4.3	Matrizenmultiplikation und lineare Abbildungen für Räume	
$\mathbb{K}^n$	. . . . .	162
4.3.1	Multiplikation von Matrizen . . . . .	162
4.3.2	Matrix einer linearen Abbildung von $\mathbb{K}^n$ nach $\mathbb{K}^m$ . .	166
4.3.3	Matrix der Komposition von linearen Abbildungen .	168
4.3.4	Rechenregeln für Matrizen . . . . .	170
4.3.5	Die Einheitsmatrix . . . . .	172
4.3.6	Invertierbarkeit von Matrizen . . . . .	173
4.3.7	Eigenschaften der inversen Matrizen . . . . .	175
4.3.8	Allgemeine lineare Gruppe . . . . .	175
4.3.9	Elementartransformationen linearer Gleichungssysteme und Matrizenmultiplikation . . . . .	176

4.4 Bild, Kern und verwandte Begriffe . . . . .	179
4.4.1 Bild, Kern und Faser . . . . .	179
4.4.2 Beschreibung der Injektivität und Surjektivität durch das Bild und den Kern . . . . .	180
4.4.3 Rang einer linearen Abbildung . . . . .	182
4.4.4 Nichtleere Fasern sind Verschiebungen vom Kern . .	183
4.4.5 Affine Unterräume . . . . .	184
4.4.6 Der Rangsatz . . . . .	186
4.4.7 Die Dimension der Faser . . . . .	188
4.4.8 Klassifikation endlichdimensionaler Vektorräume . .	188
4.4.9 Injektivität und Surjektivität linearer Abbildungen eines endlichdimensionalen Vektorraums . . . . .	189
4.4.10 Verfahren zur Invertierung von Matrizen . . . . .	191
4.4.11 Rang der Komposition von linearen Abbildungen bzw. des Matrixprodukts . . . . .	198

4.4.12 Rang und Lösbarkeit von linearen Gleichungssystemen	201
4.4.13 Faktorisierungssatz	203
4.4.14 Quotientenräume	205
4.5 Koordinatensysteme	211
4.5.1 Basisdarstellung von Vektoren	211
4.5.2 Basiswechsel	212
4.5.3 Wiederholter Basiswechsel	215
4.5.4 Basendarstellung von linearen Abbildungen	216
4.5.5 Basendarstellung einer Komposition von linearen Ab-	
bildungen	218
4.5.6 Basiswechsel für lineare Abbildungen	218
<b>5 Determinanten</b>	<b>220</b>
5.1 Grundlagen	220
5.1.1 Geometrische Motivation	220

5.1.2	Definierende Eigenschaften . . . . .	221
5.1.3	Weitere Eigenschaften . . . . .	223
5.2	Leibniz-Formel . . . . .	236
5.2.1	Permutationen und Determinanten . . . . .	236
5.2.2	Zerlegung von Permutation in Produkte von Trans-	
positionen . . . . .	238	
5.2.3	Vorzeichen von Permutationen . . . . .	239
5.2.4	Vorzeichen und das Produkt von Permutationen . . .	242
5.2.5	Leibniz-Formel . . . . .	243
5.2.6	Die Determinante der Transponierten Matrix . . . .	249
5.3	Determinante, Rang, Minoren und Invertierung von Matrizen	251
5.3.1	Cramer'sche Regel . . . . .	251
5.3.2	Entwicklung nach Zeilen und Spalten . . . . .	264
5.3.3	Die komplementäre Matrix . . . . .	276
5.3.4	Rang und Minoren . . . . .	283

5.3.5	Der Satz von Binet-Cauchy	294
5.3.6	Die Formel von Sylvester	306
<b>6</b>	<b>Eigenwerte und Eigenvektoren</b>	<b>307</b>
6.1	Grundlagen	307
6.1.1	Beispiele und Motivation	308
6.1.2	Definition von Eigenwerten und Eigenvektoren	313
6.1.3	Diagonalisierbarkeit von linearen Abbildungen	317
6.1.4	Diagonalisierbarkeit: eine hinreichende Bedingung	324
6.1.5	Eigenraum	328
6.2	Das charakteristische Polynom	330
6.2.1	Das charakteristische Polynom einer Matrix	332
6.2.2	Rechtfertigung der Definition vom charakteristischen Polynom	335

6.2.3 Nullstellen, Grad und Koeffizienten des charakteristischen Polynoms . . . . .	340
6.2.4 Das charakteristische Polynom, die Determinante und die Spur von linearen Abbildungen . . . . .	345
6.2.5 Der Satz von Cayley-Hamilton . . . . .	347
6.3 Diagonalisierbarkeit . . . . .	359
6.3.1 Eine notwendige und eine hinreichende Bedingung . . . . .	359
6.3.2 Vielfachheit von Nullstellen und Zerlegbarkeit in Linearfaktoren . . . . .	361
6.3.3 Ungleichungen für die algebraische und die geometrische Vielfachheit . . . . .	366
6.3.4 Charakterisierung der Diagonalisierbarkeit . . . . .	367
6.4 Die Jordansche Normalform . . . . .	373
6.4.1 Die Voraussetzungen . . . . .	373
6.4.2 Das Ziel und der Ansatz . . . . .	374

6.4.3	Über das Addieren eines Vielfachen der identischen Abbildung . . . . .	377
6.4.4	Das Lemma von Fitting . . . . .	380
6.4.5	Haupträume . . . . .	391
6.4.6	Hauptraumzerlegung . . . . .	394
6.4.7	Hauptraumzerlegung für Matrizen . . . . .	396
6.4.8	Jordansche Normalform für nilpotente Abbildungen .	402
6.4.9	Jordansche Normalform für allgemeine lineare Abbil- dungen . . . . .	424
6.4.10	Jordansche Normalform für Matrizen . . . . .	427
<b>7</b>	<b>Euklidische Räume und quadratische Formen</b>	<b>431</b>
7.1	Euklidische Räume . . . . .	432
7.1.1	Euklidische Räume über $\mathbb{R}$ und $\mathbb{C}$ : Motivation und Definitionen . . . . .	432

7.1.2	Die Ungleichung von Cauchy-Schwarz und der Winkel zwischen Vektoren . . . . .	440
7.1.3	Eigenschaften der Norm und des Abstands . . . . .	445
7.1.4	Orthogonale bzw. Orthonormale Systeme und das Gram-Schmidt-Verfahren . . . . .	447
7.1.5	Orthogonale Projektion . . . . .	451
7.1.6	Orthogonale Untervektorräume, Summen und Orthogonalkomplement . . . . .	454
7.2	Lineare Abbildungen Euklidischer Räume . . . . .	456
7.2.1	Adjungierte Abbildung . . . . .	457
7.2.2	Lineare Isometrien von Euklidischen Räumen . . . . .	463
7.2.3	Diagonalisierung linearer Isometrien . . . . .	466
7.2.4	Selbstdjungierte Abbildungen . . . . .	470
7.2.5	Diagonalisierbarkeit von selbstdjungierten Abbildungen komplexer Euklidischer Räume . . . . .	471

7.2.6	Diagonalisierung von selbstadjungierten Abbildungen reeller Euklidischer Räume . . . . .	473
7.3	Quadratische Formen . . . . .	479
7.3.1	Motivation und Grundbegriffe . . . . .	479
7.3.2	Darstellung quadratischer Formen durch symmetri- sche bilineare Formen . . . . .	484
7.3.3	Basisdarstellungen bilinearer und quadratischer Formen	486
7.3.4	Diagonalisierung von quadratischen Formen . . . . .	487
7.3.5	Definitheit, Semidefinitheit und Indefinitheit . . . . .	493
7.3.6	Charakterisierung der Definitheit, Semidefinitheit und Indefinitheit mit Hilfe der Eigenwerte . . . . .	495
7.3.7	Charakterisierung der Definitheit, Semidefinitheit und Indefinitheit durch die Koeffizienten des charakteris- tischen Polynoms . . . . .	497
7.3.8	Charakterisierung der Definitheit mit Hauptminoren	500

7.3.9	Signatur . . . . .	509
<b>8</b>	<b>Ganzzahlige lineare Algebra</b>	<b>511</b>
8.1	Lösung einer diophantschen linearen Gleichung . . . . .	511
8.1.1	Unimodulare Elementartransformationen . . . . .	512
8.1.2	Der größte gemeinsame Teiler von zwei Zahlen . . . . .	514
8.1.3	Der Chinesische Restsatz für zwei Restklassenringe .	519
8.1.4	Eine Anwendung: die RSA-Verschlüsselung . . . . .	523
8.1.5	Eine diophantsche Gleichung in 2 Variablen . . . . .	529
8.1.6	Eine diophantsche Gleichung in $n$ Variablen . . . . .	530
8.2	Lösung von diophantschen linearen Gleichungssystemen . .	533
8.2.1	Hermit'sche Normalform . . . . .	533
8.2.2	Ein Algorithmus zur Lösung von diophantschen li- nearen Gleichungssystemen . . . . .	538
8.2.3	Anwendungen zu modularen Gleichungen . . . . .	540

8.3 Gitter . . . . .	542
8.3.1 Charakterisierung von unimodularen Matrizen . . . . .	543
8.3.2 Gitter, ihre Dimension und die Determinante . . . . .	545
8.3.3 Erzeugung rationalen Gittern . . . . .	547
8.3.4 Gitter als diskrete Untergruppen von $\mathbb{R}^n$ . . . . .	548
<b>9 Tensoren und duale Vektorräume</b>	<b>549</b>
9.1 Einstieg: Tensoren als multidimensionale Arrays . . . . .	549
9.1.1 Tensoren und ihre Grundoperationen . . . . .	549
9.1.2 Tensorprodukt und Determinanten . . . . .	552
9.1.3 Tensoren und die Euklidische Struktur . . . . .	555
9.2 Tensoren und multilinear Abbildungen . . . . .	555
9.2.1 Duale Vektorräume und duale Basen . . . . .	555
9.2.2 Tensoren und multilinear Formen . . . . .	557

<b>9.2.3 Koordinatentransformationen für Basendarstellungen</b>	
<b>multilinearer Formen</b>	559
<b>9.3 Tensoren und multilineare Abbildungen</b>	560
<b>9.4 Anwendungen der Tensoren</b>	562
<b>9.4.1 Quantenberechnungen</b>	562
<b>9.5 Verallgemeinertes Hookesches Gesetz</b>	562
<b>9.6 Tensorrang und schnelle Multiplikation von Matrizen</b>	562
<b>9.7 Quantitative Biologie</b>	562
 <b>10 Matroide</b>	 562
 <b>11 Verschiedenes</b>	 562
<b>11.1 Singulärwertzerlegung</b>	562

# 1 Mathematische Grundbegriffe

## 1.1 Aussagen

### 1.1.1 Aussage

Eine Aussage ist ein Satz, der entweder wahr oder falsch ist. Etwa:

- $2 < 1$  (falsch)
- $2 = 1$  (falsch)
- $2 > 1$  (wahr)

Keine Aussagen:

- Hallo!
- Was gibt's Neues?

**Bem** (Benennung von Aussagen). Bei der Kategorisierung von beweisbaren mathematischen Aussagen gibt es die folgenden Tendenzen:

- Proposition: relativ einfach zu beweisen
- Theorem, Satz: bemerkenswert oder schwer zu beweisen
- Lemma, Hilfssatz: Hilfsaussage, die technisch ist und beim Beweis eines Theorems eingesetzt wird.

### 1.1.2 Logische Verknüpfungen

Seien  $A$  und  $B$  Aussagen. Dann definiert man anhand von  $A$  und  $B$  die folgenden Aussagen:

- $A \wedge B$  Konjunktion („und“)
- $A \vee B$  Disjunktion („oder“)

- $A \Rightarrow B$  Implikation
- $A \Leftrightarrow B$  Äquivalenz
- $A \dot{\vee} B$  ausschließende Disjunktion
- $\neg A, \bar{A}$  Negation (Verneinung)

Bsp.

- $x, y \in \mathbb{R}, x = y \Rightarrow x^2 = y^2$  (wahr)
- $x, y \in \mathbb{R}, x^2 = y^2 \Rightarrow x = y$  (falsch)

## 1.2 Mengen

### 1.2.1 Menge

Eine Menge ist eine Ansammlung von Objekten. Diese Objekte nennt man Elemente der Menge. (intuitive Beschreibung, keine exakte Definition)

Eine Weise, Mengen zu definieren, ist durch die Auflistung ihrer Elemente. Dabei stehen die geschweiften Klammer für Mengen, die drei Punkte bedeuten „usw“.

- $\{1, 2, 5, 7\}$
- $\{1\}$
- $\{1, \{2, 5\}, \{6\}\}$
- $\{1, 2, 3, \dots\}$

Ist  $x$  Element der Menge  $X$ , so schreibt man  $x \in X$ . Ist  $x$  nicht Element der Menge  $X$ , so schreibt man  $x \notin X$ .

Seien  $A$  und  $B$  Mengen. Dann ist  $A$  genau dann eine Teilmenge von  $B$ , wenn jedes Element von  $A$  auch Element von  $B$  ist ( $A \subseteq B \Leftrightarrow \forall x \in A : x \in B$ ).<sup>1</sup> Zwei Mengen  $A$  und  $B$  heißen genau dann gleich, wenn  $A \subseteq B$  und  $B \subseteq A$  gilt.  $A$  heißt genau dann echte Teilmenge einer Menge  $B$ , wenn  $A \subseteq B$  und  $A \neq B$  erfüllt sind. Bezeichnung:  $A \subsetneq B$ .

### 1.2.2 Zahlenmengen

$\mathbb{N} := \{1, 2, 3, \dots\}$  die Menge der natürlichen Zahlen<sup>2</sup>

---

<sup>1</sup>In einigen mathematischen Quellen bezeichnet man die Inklusion als  $\subset$  und nicht als  $\subseteq$ . Es gibt aber auch Quellen, in denen  $\subset$  die strikte Inklusion bezeichnet. Daher ziehe ich persönlich  $\subseteq$  vor.

<sup>2</sup>Die Definition des Begriffs natürliche Zahl sowie die Bedeutung der jeweiligen Bezeichnung  $\mathbb{N}$  für die Menge der natürlichen Zahlen sind abhängig von einer konkreten

$$\mathbb{N}_0 := \{0, 1, 2, \dots\}$$

$\mathbb{Z} := \{0, 1, -1, 2, -2, \dots\}$  ganze Zahlen

$\mathbb{Q} := \left\{ \frac{p}{q} : p \in \mathbb{Z}, q \in \mathbb{N} \right\}$  rationale Zahlen

$\mathbb{R}$  reelle Zahlen

$\mathbb{C}$  komplexe Zahlen

$$\Rightarrow \mathbb{N} \subseteq \mathbb{N}_0 \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R} \subseteq \mathbb{C}$$

---

Quelle. Manche Quellen definieren die Menge der natürlichen Zahlen als  $\{0, 1, 2, \dots\}$ , es ist mittlerweile sogar die ISO-Norm 80000-2, es gibt aber trotz der ISO-Norm sehr viele Quellen, in denen 0 nicht in die Menge der natürlichen Zahl aufgenommen wird. Tendenziell wählt man  $\mathbb{N} = \{1, 2, 3, \dots\}$  in der Analysis und  $\mathbb{N} = \{0, 1, 2, \dots\}$  in der theoretischen Informatik, Mengenlehre sowie der diskreten Mathematik.

### 1.2.3 Definition durch eine Bedingung

Eine weitere Weise, Mengen zu definieren, ist durch Bedingungen. Format:

$$\{\text{AUSDRUCK} : \text{BEDINGUNG}\}$$

Die Lesart für den Doppelpunkt ist “sodass” bzw. “mit der Bedingung”.

Zum Beispiel beschreibt die Formel

$$\{k^2 : k \in \mathbb{N}, k \text{ ungerade}\}$$

die Menge aller Quadrate von positiven ungeraden Zahlen. Eine andere Beschreibung für dieselbe Menge ist

$$\{(2i - i)^2 : i \in \mathbb{N}\}.$$

### 1.2.4 Die leere Menge

Die leere Menge ist die Menge, die keine Elemente enthält. Bezeichnung:  $\emptyset$ .

### 1.2.5 Potenzmenge

Sei  $X$  eine Menge. Dann ist die Potenzmenge von  $X$  die Menge aller Teilmengen von  $X$ . Bezeichnung:  $2^X$ . Formal:  $2^X := \{A : A \subseteq X\}$ .

Anmerkung: Ist  $|X| = n \in \mathbb{N}_0$ , so gilt  $|2^X| = 2^n$ .

### 1.2.6 Mengenoperationen

Seien  $A, B$  Mengen. Dann heißt

- $A \cap B := \{x : (x \in A) \wedge (x \in B)\}$  Durchschnitt von  $A$  und  $B$
- $A \cup B := \{x : (x \in A) \vee (x \in B)\}$  Vereinigung von  $A$  und  $B$
- $A \setminus B := \{x : (x \in A) \wedge (x \notin B)\}$  Mengendifferenz von  $A$  und  $B$

### 1.2.7 Disjunkte Mengen

Seien  $A, B$  Mengen.  $A$  und  $B$  heißen genau dann disjunkt, wenn  $A \cap B = \emptyset$ .

## 1.3 Tupel

Für Objekte  $a, b$  kann man das *geordnete Paar*  $(a, b)$  definieren. Für Objekte  $a, b, c, d$  definiert man die Gleichheit  $(a, b) = (c, d)$  durch  $a = c$  und  $b = d$ .  $a$  heißt das erste Element des Paares  $(a, b)$  und  $b$  heißt das zweite Element.

Für Mengen  $X, Y$  definiert man das *Kreuzprodukt*  $A \times B$  durch  $A \times B := \{(x, y) : x \in X, y \in Y\}$ . Analog definiert man geordnete Tupel und das Kreuzprodukt  $A \times B \times C$  von Mengen  $X, Y$  und  $Z$ . Noch allgemeiner kann man für jedes  $n \in \mathbb{N}$  geordnete  $n$ -Tupel  $(x_1, \dots, x_n)$  einführen und das Kreuzprodukt  $X_1 \times \dots \times X_n := \{(x_1, \dots, x_n) : x_1 \in X_1, \dots, x_n \in X_n\}$  von Mengen  $X_1, \dots, X_n$ .

Für eine Menge  $X$  führt man die Bezeichnung

$$X^n := \underbrace{X \times \dots \times X}_{n \text{ mal}} = \{(x_1, \dots, x_n) : x_1, \dots, x_n \in X\}.$$

Das Element  $x_i$  mit  $i \in \{1, \dots, n\}$  im  $n$ -Tupel  $(x_1, \dots, x_n)$  heißt die  $i$ -te

*Komponente* des Tupels.

**Bsp.** Geometrische Veranschaulichung einiger Tupel:

- $[0, 1] \times [0, 2]$  ist Rechteck in  $\mathbb{R}^2$   
 $\{0\} \times [0, 2]$  ist eine Kante dieses Rechtecks  
 $\{0, 1\} \times \{0, 2\}$  sind Eckpunkte dieses Rechtecks
- $[0, 1]^3$  ist Würfel in  $\mathbb{R}^3$   
 $[0, 1]^2 \times \{0\}$  ist eine Seitenfläche des Würfels  
 $[0, 1]^2 \times \{1\}$  ist gegenüberliegende Seitenfläche (Facette)  
 $\{0\}^2 \times [0, 1]$  ist eine Kante des Würfels
- $[0, 1]^4$  ist ein 4-dimensionaler Würfel in  $\mathbb{R}^4$

## 1.4 Abbildungen

### 1.4.1 Abbildung

Seien  $X, Y$  Mengen. Eine Abbildung  $f$  von  $X$  nach  $Y$  ist eine Vorschrift, die jedem  $x \in X$  genau ein Element aus  $Y$  zuordnet. Dieses Element aus  $Y$  wird durch  $f(x)$  bezeichnet. Wenn  $f$  eine Abbildung von  $X$  nach  $Y$  ist, dann bezeichnet man das durch:  $f : X \rightarrow Y$ . Die Menge  $X$  heißt der Definitionsbereich von  $f$ ,  $Y$  heißt der Wertebereich von  $f$ .

- $f : \mathbb{R} \rightarrow \mathbb{R}$ ,  
$$f(x) := x^2 - 2x + 7 \quad \forall x \in \mathbb{R}$$
- $f : \mathbb{R} \setminus \{1\} \rightarrow \mathbb{R}$ ,  
$$f(x) := \frac{1}{x-1} \quad \forall x \in \mathbb{R}$$
- $\text{sign} : \mathbb{R} \rightarrow \mathbb{R} = \{1, 0, -1\}$
- $f : 2^{\mathbb{N}} \rightarrow \mathbb{N}$ ,  $f(A) := \min(A) \quad \forall A \subseteq \mathbb{N}$ , z.B.  $f(\{1, 7, 43\}) = 1$

- $f : \mathbb{N} \rightarrow 2^{\mathbb{N}}, f(k) := \{1, \dots, k\} \forall k \in \mathbb{N}$

Zwei Abbildungen  $f, g : X \rightarrow Y$  heißen gleich, falls  $f(x) = g(x)$  für alle  $x \in X$  gilt. Durch  $Y^X$  bezeichnet man die Menge aller Abbildungen von  $X$  nach  $Y$ .

#### 1.4.2 Bild und Urbild

Seien  $X, Y, A, B$  Mengen mit  $A \subseteq X$  und  $B \subseteq Y$ . Sei  $f : X \rightarrow Y$ . Dann heißt  $f(A) := \{f(x) : x \in A\}$  das Bild von  $A$  bzgl.  $f$  und  $f^{-1}(B) := \{x \in X : f(x) \in B\}$  das Urbild von  $B$  bzgl.  $f$ .

**Bsp.** Sei  $f : \mathbb{R} \rightarrow \mathbb{R}$  mit  $f(x) := x^2$  für alle  $x \in \mathbb{R}$ .

- $f([1, 2]) = [1, 4] \nearrow$  Abb. 1
- $f^{-1}([1, 4]) = [1, 2] \cup [-2, -1] \nearrow$  Abb. 2

- $f^{-1}([-7, 8]) = \{x \in \mathbb{R} : f(x) \in [-7, 8]\}$   
 $= \{x \in \mathbb{R} : -7 \leq f(x) \leq 8\}$   
 $= \{x \in \mathbb{R} : -7 \leq x^2 \leq 8\}$   
 $= \{x \in \mathbb{R} : x^2 \leq 8\}$   
 $= \{x \in \mathbb{R} : |x| \leq \sqrt{8}\}$   
 $= [-\sqrt{8}, \sqrt{8}]$

**Bem** (Intervalle). Seien  $a, b \in \mathbb{R}$  mit  $a \leq b$ . Dann können Intervalle wie folgt definiert werden:

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

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

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

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

### 1.4.3 Injektivitat, Surjektivitat und Bijektivitat

Seien  $X, Y$  Mengen und sei  $f : X \rightarrow Y$ . Dann heit  $f$ :

- injektiv, falls fr alle  $x_1, x_2 \in X : x_1 \neq x_2$  die Bedingung  $f(x_1) \neq f(x_2)$  gilt.
- surjektiv, falls fr jedes  $y \in Y$  ein  $x \in X$  mit der Eigenschaft  $f(x) = y$  existiert.
- bijektiv, falls  $f$  injektiv und surjektiv ist.

**Bsp.** Untersuche folgende Funktionen auf Bijektivitat:

- $f : \mathbb{R} \rightarrow \mathbb{R}, f(x) := x^2$  fr alle  $x \in \mathbb{R}$   
surjektiv ? nein,  $-1 \neq f(x)$  fr alle  $x \in \mathbb{R}$   
injektiv ? nein,  $f(x) = f(-x)$  fr alle  $x \in \mathbb{R}$

- $f : \mathbb{R} \rightarrow [0, +\infty), f(x) := x^2$  für alle  $x \in \mathbb{R}$   
 surjektiv ? ja  
 injektiv ? nein (analog)
- $f : \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R}, f(x) = \frac{1}{x}$  für alle  $x \in \mathbb{R}$   
 surjektiv ? nein, 0 wird nicht angenommen  
 injektiv ? ja
- $f : \mathbb{R} \rightarrow \mathbb{R}, f(x) = 2x + 3$  für alle  $x \in \mathbb{R}$   
 bijektiv ? ja

#### 1.4.4 Umkehrfunktion

Seien  $X, Y$  Mengen und sei  $f : X \rightarrow Y$  bijektiv. Die Abbildung, die jedem  $y \in Y$  das eindeutige  $x \in X$  mit  $f(x) = y$  zuordnet, heißt die Umkehrabbildung von  $f$  und wird durch  $f^{-1}$  bezeichnet.

**Bsp.** Die Umkehrung von  $f : \mathbb{R} \rightarrow \mathbb{R} : f(x) := 2x + 3$  ist  $f^{-1}(x) = \frac{x-3}{2}$  ( $x \in \mathbb{R}$ ).

### 1.4.5 Komposition

Seien  $X, Y, Z$  Mengen,  $f : X \rightarrow Y$  und  $g : Y \rightarrow Z$ . Dann heißt  $g \circ f : X \rightarrow Z$  mit  $(g \circ f)(x) := g(f(x))$  für alle  $x \in X$  die Komposition von  $g$  und  $f$ .

**Bsp.** Seien  $f : \mathbb{R} \rightarrow \mathbb{R} : f(x) = 2x+3$  für alle  $x \in \mathbb{R}$  und  $g : \mathbb{R} \rightarrow \mathbb{R} : g(x) = x^2$  für alle  $x \in \mathbb{R}$ . Dann ist  $(f \circ g)(x) = 2x^2 + 3$  und  $(g \circ f)(x) = (2x + 3)^2$ .

### 1.4.6 Identische Abbildung

Sei  $X$  eine Menge. Dann heißt die Abbildung  $\text{id}_X : X \rightarrow X$  mit  $\text{id}_X(x) := x$  für alle  $x \in X$  die identische Abbildung auf  $X$ . Man schreibt auch häufig  $\text{id}$ , wenn  $X$  nicht angegeben werden muss.

**Bem.** Seien  $X, Y$  Mengen und sei  $f : X \rightarrow Y$  bijektiv. Dann gilt

- $f \circ f^{-1} = \text{id}_Y$ ,
- $f^{-1} \circ f = \text{id}_X$ .

#### 1.4.7 Vereinigung und Durchschnitt einer indexierten Mengenfamilie

Eine Familie bzw. Schar  $(A_i)_{i \in I}$  von Teilmengen von  $X$ , die durch die Menge  $I$  indexiert sind, ist eine Abbildung  $i \mapsto A_i$  von  $I$  nach  $2^X$ .

Für die Familie  $(A_i)_{i \in I}$  definiert man

den Durchschnitt:  $\bigcap_{i \in I} A_i := \{x \in X : x \in A_i \text{ für alle } i \in I\}$ ,

die Vereinigung:  $\bigcup_{i \in I} A_i := \{x \in X : x \in A_i \text{ für ein } i \in I\}$ .

**Bsp.**  $A_t := [t - 1, t + 1]$  ist die Menge aller Punkte in  $\mathbb{R}$ , deren Entfernung von  $t \in \mathbb{R}$  höchstens 1 ist. Dann ist  $\bigcap_{t \in [-1, 1]} A_t = \{0\}$ , denn 0 ist der einzige Punkt in  $\mathbb{R}$ , dessen Entfernung zu allen Punkten in  $[-1, 1]$  höchstens 1 ist. Des Weiteren ist  $\bigcup_{t \in [-1, 1]} A_t = [-2, 2]$ , denn alle Punkte im Intervall  $[-2, 2]$ , und nur diese Punkte auf der reellen Achse  $\mathbb{R}$ , haben den Abstand höchstens 1 zu einem Punkt aus  $[-1, 1]$ .

#### 1.4.8 Summen und Produkte

Eine Menge  $X$  heißt endlich, falls  $X = \emptyset$  oder falls eine bijektive Abbildung von  $\{1, \dots, n\}$  nach  $X$  existiert mit  $n \in \mathbb{N}$ . Der Wert  $n$  heißt die Anzahl der Elemente (Kardinalität) von  $X$  und wird durch  $|X|$  bezeichnet.

Man setzt die Kardinalität von  $\emptyset$  gleich 0.  $|X|$  ist wohl definiert, d.h. eine Menge kann nicht zwei unterschiedliche Kardinalitäten haben.

Sei  $X$  eine nichtleere endliche Menge. Dann kann  $X$  als  $X = \{x_1, \dots, x_n\}$  dargestellt werden mit  $x_i \neq x_j \Leftrightarrow i \neq j$  für alle  $i, j \in \{1, \dots, n\}$ .

Für eine Abbildung  $f : X \rightarrow \mathbb{R}$  definiert man

$$\sum_{x \in X} f(x) := f(x_1) + \dots + f(x_n),$$

$$\prod_{x \in X} f(x) := f(x_1) \cdot \dots \cdot f(x_n).$$

Im Fall  $X = \emptyset$  definiert man für  $f : X \rightarrow \mathbb{R}$  und  $\sum_{x \in X} f(x) = 0$  und  $\prod_{x \in X} f(x) = 1$ . Die Summe und das Produkt über eine Menge  $X$  sind wohldefiniert (d.h., die beiden Werte sind von der Nummerierung  $x_1, \dots, x_n$  der Elemente von  $X$  unabhängig).

F'ru  $n \in \mathbb{N}_0$  stehen die Bezeichnungen  $\sum_{i=1}^n$  bzw.  $\prod_{i=1}^n$  für die Summe bzw. das Produkt über alle ganzzahligen  $i$  mit  $1 \leq i \leq n$ .

## 1.5 Prädikate

### 1.5.1 Prädikat

Sei  $X$  Menge. Dann heißt  $P : X \rightarrow \{\text{falsch}, \text{wahr}\}$  *Prädikat* auf  $X$ . Etwa  $P : \mathbb{N} \rightarrow \{\text{falsch}, \text{wahr}\}$ ,  $P(k) := \text{„}k(k+1)\text{ ist durch }3\text{ teilbar“}$  für alle  $k \in \mathbb{N}$ .

Durch ein Prädikat  $P : X \rightarrow \{\text{falsch}, \text{wahr}\}$  kann man die Menge  $\{x \in X : P(x)\}$  definieren.

### 1.5.2 Quantoren

$\forall x \in X : P(x)$  für ein Prädikat  $P$  auf eine Menge  $X$  steht für die Aussage „die Bedingung  $P(x)$  gilt für alle  $x \in X$ .“  $\forall$  heißt das *Allgemeinheitsquantor* (Bedeutung: für  $\forall$ le).

$\exists x \in X : P(x)$  bezeichnet die Aussage „die Bedingung  $P(x)$  gilt für ein

$x \in X$ .“  $\exists$  heißt *Existenzquantor* (Bedeutung: es  $\exists$ istert).

**Bem.** Negierung von Aussagen:

$$(i) \overline{\forall x \in X : P(x)} \Leftrightarrow \exists x \in X : \overline{P(x)}$$

$$(ii) \overline{\exists x \in X : P(x)} \Leftrightarrow \forall x \in X : \overline{P(x)}$$

**Bem.**  $\forall$  und  $\exists$  lassen sich kombinieren. Etwa, wenn man ein Prädikat  $P$  auf  $X \times Y$  hat ( $X, Y$  Mengen), so kann man die Aussagen  $(\forall x \in X \exists y \in Y : P(x, y))$ ,  $(\exists x \in X \forall y \in Y : P(x, y))$  usw. einführen.

**Bsp.** Sei  $(a_n)_{n \in \mathbb{N}}$  Folge reeller Zahlen (mit anderen Worten:  $a : \mathbb{N} \rightarrow \mathbb{R}$ ) und sei  $\alpha \in \mathbb{R}$ . Dann heißt  $\alpha$  Limes von  $(a_n)_{n \in \mathbb{N}}$ , falls das Folgende gilt:

$$\forall \epsilon \in \mathbb{R}^+ \exists N \in \mathbb{N} \forall n \in \mathbb{N} : ((n \geq N) \Rightarrow (|a_n - \alpha| < \epsilon))$$

## 1.6 Relationen

### 1.6.1 Relation

Seien  $X, Y$  Mengen. Dann heißt eine Teilmenge  $R$  von  $X \times Y$  eine (binäre) *Relation* zwischen (den Elementen von)  $X$  und  $Y$ .

Wenn für  $x \in X$  und  $y \in Y$  die Bedingung  $(x, y) \in R$  gilt, so schreibt man  $xRy$ . Wenn  $X = Y$ , dann sagt man, dass  $R$  eine (binäre) Relation auf  $X$  ist.

Bsp.

- $X$  - Menge von Fahrzeugen  
 $Y$  - Menge von Features von Fahrzeugen

	Ersatzrad	Radio	Navi	Automatik
$f_1$	1	1	1	1
$f_2$	1	1	1	0
$f_3$	0	0	1	1
$f_4$	0	1	1	0

- $\leq, <, \geq, >$  auf  $\mathbb{R}$
- $\subseteq$  als Relation auf  $2^X$  für eine Menge  $X$
- Für  $a, b \in \mathbb{N}$  schreibt man  $a|b$ , wenn  $b$  durch  $a$  ohne Rest teilbar ist.

### 1.6.2 Äquivalenzrelation

Sei  $X$  Menge und  $\sim$  eine Relation auf  $X$ . Dann heißt  $\sim$  eine Äquivalenzrelation, falls:

- (i)  $\sim$  ist *reflexiv*, d.h.  $x \sim x$  für alle  $x \in X$ .

- (ii)  $\sim$  ist *symmetrisch*, d.h.  $x \sim y$  ist äquivalent zu  $y \sim x$  für alle  $x \in X$ .
- (iii)  $\sim$  ist *transitiv*, d.h. aus  $x \sim y$  und  $y \sim z$  folgt  $x \sim z$  für alle  $x, y, z \in X$ .

Für eine Äquivalenzrelation  $\sim$  auf einer Menge  $X$  und ein  $x \in X$  heißt

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

die Äquivalenzklasse von  $x$  bzgl.  $\sim$ . Die Menge aller Äquivalenzklassen von  $\sim$  ist

$$X/\sim := \{[x]_{\sim} : x \in X\}.$$

**Bsp.**

- Sei  $V$  endliche Menge und sei  $\binom{V}{2} := \{\{u, v\} : u, v \in V, u \neq v\}$ . Das Paar  $(V, E)$  mit  $E \subseteq \binom{V}{2}$  heißt *Graph* mit Knotenmenge  $V$  und Knotenmenge  $E$ .

$$G = (V, E), G = \{1, \dots, 6\}, E = \{\{1, 2\}, \{2, 3\}, \{3, 4\}, \{4, 1\}, \{1, 3\}, \{5, 6\}\}$$

Für  $a, b \in V$  heißt  $b$  von  $a$  aus *erreichbar* (im Graphen  $G = (V, E)$ ), falls ein  $k \in \mathbb{N}_0$  und Elemente  $u_0, \dots, u_k \in V$  existieren mit  $u_0 = a$ ,  $u_k = b$  und  $\{u_i, u_{i+1}\} \in E$  für alle  $i \in \mathbb{N}_0$  mit  $i < k$ .

Die Erreichbarkeit ist eine Äquivalenzklasse auf  $V$ . Die Äquivalenzklassen (Zusammenhangskomponenten) für dieses Beispiel sind  $\{1, 2, 3, 4\}$  und  $\{5, 6\}$ .

- Sei  $m \in \mathbb{N}$ . Für  $a, b \in \mathbb{Z}$  sagt man, dass  $a$  kongruent zu  $b$  modulo  $m$  ist, falls  $a - b \in m\mathbb{Z}$ , wobei  $m\mathbb{Z} := \{mz : z \in \mathbb{Z}\}$ .

Schreibweise:  $a \equiv b \pmod{m}$ .

Die Kongruenz modulo  $m$  ist eine Äquivalenzrelation auf  $\mathbb{Z}$ .

- Sei  $\sim$  Relation auf  $\mathbb{Z} \times \mathbb{N}$ , definiert durch  $(a, b) \sim (c, d)$  für  $a, c \in \mathbb{Z}, b, d \in \mathbb{N}$ , wenn  $ad = bc$  gilt.

Diese Relation ist eine Äquivalenzrelation (Aufgabe).

D.h. jede rationale Zahl ist eine Äquivalenzklasse von diesem  $\sim$ .

## 2 Ausgewählte algebraische Strukturen

Die algebraischen Strukturen dieses Kapitels haben zwei Formen.  $(X, *)$  – Grundmenge  $X$ , die mit einer Verknüpfung  $*$  ausgestattet ist, und  $(X, \cdot, +)$  – Grundmenge, die mit zwei Verknüpfungen  $\cdot$  und  $+$  ausgestattet ist.

### 2.1 Gruppen

#### 2.1.1 Binäroperationen

Sei  $X$  Menge und sei  $* : X \times X \rightarrow X$ . Dann heißt  $*$  eine *Binäroperation* (bzw. *Binärverknüpfung*) auf  $X$ . Die Struktur  $(X, *)$  mit einer Binäroperation  $X \times X \rightarrow X$  nennt man in Algebra ein *Magma*.

**Bsp.**  $+, -, \cdot$  sind binäre Operationen auf  $\mathbb{Q}$ ,  $/$  ist eine binäre Operation auf  $\mathbb{Q} \setminus \{0\}$ , aber nicht auf  $\mathbb{Q}$ .

Man schreibt  $x * y$  statt  $*(x, y)$  für  $x, y \in X$ .  $*$  heißt *assoziativ*, wenn

$$x * (y * z) = (x * y) * z$$

für alle  $x, y, z \in X$ . Ein Paar  $(X, *)$ , für welches  $* : X \times X \rightarrow X$  assoziativ ist, nennt man in Algebra eine *Halbgruppe*.

**Bsp.**  $+$  und  $\cdot$  sind beide assoziativ auf  $\mathbb{Z}$ , sowie  $\mathbb{Q}$ ,  $\mathbb{R}$  und  $\mathbb{C}$ .  $-$  ist nicht assoziativ auf  $\mathbb{Z}$ .

Für eine gegebene Menge  $X$  ist  $\circ$  eine assoziative Operationen auf der Menge aller Abbildungen von  $X$  nach  $X$ .

**Bem.** Bei einer Assoziativen Binäroperation hat die Wahl der Klammerung von  $x_1 * \dots * x_n$  mit  $n \in \mathbb{N}$  und  $x_1, \dots, x_n \in X$  keinen Einfluss auf den Wert dieses Ausdrucks. Zum Beispiel kann man mit einer mehrfachen Anwendung des Assoziativgesetzes zeigen, dass  $a * (b * (c * d)) = ((a * b) * c) * d$  für alle  $a, b, c, d \in X$  gilt. Wie oft soll man das Assoziativgesetz anwenden, um dieses Gesetz herzuleiten?

## 2.1.2 Gruppe

Sei  $G$  Menge und  $* : G \times G \rightarrow G$ . Dann heißt  $(G, *)$  Gruppe, falls:

(G1)  $*$  ist assoziativ, d.h.  $a * (b * c) = (a * b) * c$  für alle  $a, b, c \in G$ .

(G2) Es existiert ein Element  $e \in G$ , sodass für jedes  $a \in G$ :

(a)  $e * a = a$

(b) ein  $b \in G$  existiert mit  $b * a = e$ .

Ein  $e \in G$  mit  $e * a = a$  für alle  $a \in G$  heißt ein *linkes neutrales Element*.

Für  $a \in G$  heißt jedes  $b \in G$  mit  $b * a = e$  ein *linkes inverses Element* von  $a$  bzgl.  $e$ .

Eine Gruppe  $(G, *)$  heißt *abelsch* (oder *kommutativ*), falls  $a * b = b * a$  für alle  $a, b \in G$  gilt.

**Bem.** Oft wird die Gruppenoperation  $*$  durch  $\cdot$  bezeichnet. In diesem Fall sagt man, dass die Gruppe multiplikativ geschrieben wird. Man benutzt

dann auch die Standardbegriffe und -bezeichnungen für die Multiplikation:

$$ab := a \cdot b \quad (2.1.1)$$

$$a^n := \underbrace{a \cdot \dots \cdot a}_{n \text{ mal}} \quad (\text{für } n \in \mathbb{N}) \quad (2.1.2)$$

$$a^0 := e. \quad (2.1.3)$$

Damit die Definition  $a^0 = e$  korrekt ist, muss noch bewiesen werden, dass das  $e$  aus (G2)(a) eindeutig ist. Die Eindeutigkeit wird in Kürze bewiesen.

Eine multiplikativ geschriebene Gruppe muss nicht kommutativ sein (d.h.  $ab = ba$  muss in einer Gruppe  $(G, \cdot)$  nicht gelten).

Bsp.

- Sei  $n \in \mathbb{N}$ . Wir bezeichnen durch  $S_n$  die Menge aller bijektiven Abbildungen von  $\{1, \dots, n\}$  in  $\{1, \dots, n\}$ . Die Elemente von  $S_n$  nennt man *Permutationen* von  $\{1, \dots, n\}$ .

Man führt die Operation  $\cdot$  auf  $S_n$  ein, durch  $\sigma \cdot \tau := \sigma \circ \tau$  für alle  $\sigma, \tau \in S_n$ .

Wie sieht  $S_n$  aus? Für  $n = 1, 2, 3, \dots$

$$S_1 : |S_1| = 1$$

$i$	$e(i)$
1	1

$$S_2 : |S_2| = 2$$

$i$	$e(i)$	$\sigma(i)$
1	1	2
2	2	1

$$S_3 : |S_3| = 6$$

$i$	$e(i)$	$\sigma(i)$	$\tau(i)$	$\phi_1(i)$	$\phi_2(i)$	$\phi_3(i)$
1	1	3	2	1	3	2
2	2	1	3	3	2	1
3	3	2	1	2	1	3

Was ist  $\tau\phi_1$ ? Was ist  $\phi_1\tau$ ? Sind  $\tau\phi_1$  und  $\tau\phi_1$  gleich?

$S_n$ : Die Anzahl der Elemente in  $S_n$  ist  $|S_n| := n! = \prod_{i=1}^n i$  für jedes  $n \in \mathbb{N}$ .

Anmerkung:  $0! = 1$ .

- Horizontale Spiegelung  $h$  und vertikale Spiegelung  $v$  auf ein Rechteck:

Wir generieren die Menge aller Operationen, die durch Komposition von  $v$  und  $h$  (in einer beliebigen Reihenfolge) erzeugbar sind.

Die Gruppe wird genau so wie  $S_n$  multiplikativ geschrieben.

$$h^2 = e \quad (\text{neutral})$$

$$v^2 = e$$

$$hv = vh \quad (\text{Drehung um } 180^\circ)$$

Damit ist  $\{e, h, v, hv\}$  die Menge aller Operationen. Deren Verknüpfungen lauten:

$\cdot$	e	h	v	hv
e	e	h	v	hv
h	h	e	hv	v
v	v	hv	e	h
hv	hv	v	h	e

- Sei  $G$  die Menge aller Funktionen  $f : \mathbb{R} \rightarrow \mathbb{R}$  der Form  $f(x) := ax + b$  mit  $a \in \mathbb{R} \setminus \{0\}$  und  $b \in \mathbb{R}$ . Dann ist  $(G, \circ)$  eine Gruppe mit unendlich vielen Elementen.

### 2.1.3 Das linke und das rechte Inverse

**Prop.** Sei  $(G, \cdot)$  Gruppe, sei  $e \in G$  ein linkes neutrales Element von  $(G, \cdot)$  und seien  $a, b \in G$  Elemente mit  $ab = e$ . Dann gilt  $ba = e$ .

*Beweis.* Weil  $e$  ein linkes neutrales Element ist, existiert nach (G2)(b) ein  $c \in G$  mit  $ca = e$ . Nach (G2)(a) gilt  $ba = e(ba)$ . Wir setzen für  $e$  den Ausdruck  $ca$  ein und erhalten somit  $ba = (ca)(ba)$ . Nach dem Assoziativgesetz (G1) gilt  $ba = (ca)(ba) = c(a(ba))$ . Eine weitere Anwendung von (G1) ergibt  $ba = c((ab)a)$ . Es folgt  $ba = c(ea)$  und weil  $e$  ein linkes neutrales Element ist, hat man  $ba = ca = e$ .  $\square$

### 2.1.4 Eindeutigkeit des linken neutralen Elements

**Prop.** Sei  $(G, \cdot)$  Gruppe und seien  $e, e'$  linke neutrale Elemente von  $(G, \cdot)$ . Dann gilt  $e = e'$ .

*Beweis.* Weil  $e$  von links neutral ist, gilt  $ee' = e'$ . Nach 2.1.3 gilt  $e'e = e'$ ,

weil  $e'$  von links neutral ist. Andererseits gilt  $e'e = e$ , weil  $e'$  von links neutral ist. Es folgt  $e = e'$ .  $\square$

Nun kann man von *dem* linken neutralen Element sprechen. Dieses wird im Folgenden durch  $e$  bezeichnet.

### 2.1.5 Neutralität von links und rechts

**Prop.** *Sei  $(G, \cdot)$  Gruppe und sei  $e$  das linke neutrale Element von  $(G, \cdot)$ . Dann gilt  $ae = a$  für alle  $a \in G$  (d.h.  $e$  ist auch von rechts neutral).*

*Beweis.* Nach (G2)(b) existiert ein  $b \in G$  mit  $ba = e$ . Dann gilt

$$ae = a(ba) \stackrel{(G1)}{=} (ab)a \stackrel{(2.1.3)}{=} ea = a.$$

Ein Element  $e$  mit  $ae = ea = a$  für alle  $a \in G$  heißt *neutral*. Man kann also von dem neutralen Element  $e$  der Gruppe  $(G, \cdot)$  sprechen.

## 2.1.6 Eindeutigkeit des inversen Elements

**Prop.** Sei  $(G, \cdot)$  Gruppe, sei  $a \in G$  und seien  $a, b \in G$  linke Inverse zu  $a$ , d.h.  $ba = e$  und  $ca = e$ . Dann gilt  $b = c$ .

*Beweis.* Nach 2.1.3 gilt  $ac = e$ . Wegen 2.1.5 hat man  $b = be$ . Das Einsetzen von  $ac$  für  $e$  ergibt  $b = b(ac)$ .  $b = b(ac) \stackrel{(G1)}{=} (ba)c = ec = c$ .  $\square$

Mit der Berücksichtigung von 2.1.3 und 2.1.6 ergibt sich, dass jedes Element  $a$  einer Gruppe  $(G, \cdot)$  ein eindeutiges Element  $a^{-1}$  besitzt mit  $a^{-1}a = aa^{-1} = e$ .  $a^{-1}$  heißt das *Inverse* von  $a$ .

## 2.1.7 Zusammenfassung der Gruppeneigenschaften

In einer Gruppe  $(G, \cdot)$  gilt:

(i)  $(ab)c = a(bc)$

(ii) Es existiert genau ein  $e \in G$  mit  $ae = ea = a \quad \forall a \in G$

(iii) Für jedes  $a \in G$  existiert genau ein  $a^{-1} \in G$  mit  $aa^{-1} = a^{-1}a = e$

Wenn die Gruppe kommutativ ist, dann wird die Gruppenoperation auch oft durch  $+$  bezeichnet. In diesem Fall verwendet man die Bezeichnungen und Begriffe für die Addition:

- Das Inverse zu  $a \in G$  wird durch  $-a$  bezeichnet und das Negative von  $a$  genannt.
- $na := \underbrace{a + \dots + a}_{n\text{-mal}}$  für alle  $n \in \mathbb{N}$  und  $a \in G$
- Das neutrale Element wird durch  $0$  bezeichnet.

In einer kommutativen Gruppe  $(G, \cdot)$  gilt:

- $(a \cdot b) \cdot c = a \cdot (b \cdot c) \quad \forall a, b, c \in G$
- $a \cdot b = b \cdot a \quad \forall a, b \in G$

- Es existiert genau eine  $0 \in G$  mit  $a + 0 = a \quad \forall a \in G$
- für jedes  $a \in G$  existiert genau ein  $-a \in G$  mit  $a + (-a) = 0$

**Bsp.** Kommutative Gruppen:  $(\mathbb{Z}, +)$ ,  $(\mathbb{Q}, +)$ ,  $(\mathbb{R}, +)$ ,  $(\mathbb{Q} \setminus \{0\}, \cdot)$ ,  $(\mathbb{R} \setminus \{0\}, \cdot)$ .

### 2.1.8 Das Inverse des Produkts und der Potenz

**Prop.** Sei  $(G, \cdot)$  Gruppe, Seien  $a, b \in G$  und  $n \in \mathbb{N}$ , dann gilt

$$(ab)^{-1} = b^{-1}a^{-1}, \tag{2.1.4}$$

$$(a^n)^{-1} = (a^{-1})^n. \tag{2.1.5}$$

*Beweis.* Übungsaufgabe □

Mit der Berücksichtigung dieser Proposition setzen wir

$$\begin{aligned} (a)^{-n} &:= (a^n)^{-1} = (a^{-1})^n \quad \text{für } n \in \mathbb{N}, \\ a^0 &:= e. \end{aligned}$$

## 2.1.9 Zyklische Gruppen

Eine Gruppe  $(G, \cdot)$  heißt zyklisch, falls  $a \in G$  existiert, sodass jedes  $x \in G$  als  $x = a^n$  mit  $n \in \mathbb{Z}$  darstellbar ist.

Für eine additiv geschriebene Gruppe  $(G, +)$  lässt sich diese Eigenschaft so hinschreiben: es existiert ein  $a \in G$ , sodass jedes  $x \in G$  als  $x = na$  mit  $n \in \mathbb{Z}$  darstellbar ist.)

**Bsp.**  $(\mathbb{Z}, +)$  ist zyklisch.  $(\mathbb{Q}, +)$  ist nicht zyklisch.

**Prop.** Sei  $m \in \mathbb{N}$ . Seien  $A, B \in \mathbb{Z}/m\mathbb{Z}$  (Restklassen). Dann existiert eine eindeutige Restklasse  $C \in \mathbb{Z}/m\mathbb{Z}$  mit  $a + b \in C$  für alle  $a \in A, b \in B$ .

Die Klasse  $C$  aus der Proposition wird durch  $A + B$  bezeichnet; mit  $A, B$  wie in der Proposition. Somit hat man eine Addition auf  $\mathbb{Z}/m\mathbb{Z}$  eingeführt.

**Prop.** Sei  $m \in \mathbb{N}$ . Dann ist  $(\mathbb{Z}/m\mathbb{Z}, +)$  eine zyklische Gruppe.

*Beweis.* Die Gruppeneigenschaften kann man direkt verifizieren. Die Gruppe ist zyklisch, weil sie offensichtlich durch  $[1]$  (die Restklasse von 1) erzeugt wird.  $\square$

### 2.1.10 Untergruppen

Sei  $(G, \cdot)$  Gruppe und  $\emptyset \neq H \subseteq G$ . Dann heißt  $H$  Untergruppe von  $(G, \cdot)$ , falls für alle  $a, b \in H$  gilt:

- (i)  $ab \in H$  (Abgeschlossenheit),
- (ii)  $a^{-1} \in H$  (Existenz der Inversen innerhalb von  $H$ ).

**Bem.**  $H$  ist genau dann eine Untergruppe von  $(G, \cdot)$ , wenn  $\cdot$  eingeschränkt werden kann und  $(H, \cdot)$  eine Gruppe ist.

## 2.2 Ringe

### 2.2.1 Ring

Sei  $R$  nichtleere Menge und  $+$  sowie  $\cdot$  Binäroperationen auf  $R$ . Dann heißt  $(R, +, \cdot)$  Ring, falls:

(R1)  $(R, +)$  ist eine kommutative Gruppe

(R2)  $\cdot$  ist assoziativ

(R3) Es gelten die Distributivgesetze

$$(a) (a + b)c = ac + bc$$

$$(b) c(a + b) = ca + cb$$

Ein Ring  $(R, +, \cdot)$  heißt kommutativ, falls  $\cdot$  kommutativ ist. Ein Ring  $(R, +, \cdot)$  heißt ein Ring mit 1 bzw. unitärer Ring, falls ein Element  $1 \in R$  mit  $1 \cdot a = a \cdot 1 = a \quad \forall a \in R$  existiert.

Bsp.

- Kommutative Ringe mit 1:

$$(\mathbb{Z}, +, \cdot), \quad (\mathbb{Q}, +, \cdot), \quad (\mathbb{R}, +, \cdot).$$

- Sei  $X$  nichtleere Menge. Dann ist  $(\mathbb{R}^X, +, \cdot)$  mit

$$(f + g)(x) := f(x) + g(x) \quad \forall x \in X$$
$$(f \cdot g)(x) := f(x) \cdot g(x) \quad \forall x \in X$$

ein Ring.

- Polynomringe und Matrizenringe (später)

**Bem.** In einem Ring  $(R, +, \cdot)$  gilt

$$a \cdot 0 = 0 \cdot a = 0. \tag{2.2.1}$$

*Beweis.*

$$\begin{aligned} a \cdot 0 &= a \cdot (0 + 0) = a \cdot 0 + a \cdot 0 \\ \Rightarrow \quad \underbrace{a \cdot 0 + (-a \cdot 0)}_0 &= a \cdot 0 + \underbrace{a \cdot 0 + (-a \cdot 0)}_0 \\ \Rightarrow \quad 0 &= a \cdot 0 + 0 = a \cdot 0 \end{aligned} \quad \square$$

### 2.2.2 Polynomring in einer Unbestimmten

Sei  $R$  kommutativer Ring und  $x$  ein formales Symbol (das man auch eine Unbestimmte bzw. eine formale Variable nennt.). Die Menge  $R[x]$  der Polynome in  $x$  mit Koeffizienten in  $x$  besteht aus den formalen Ausdrücken der Form

$$f = \sum_{i=0}^{\infty} c_i x^i \tag{2.2.2}$$

mit  $c_i \in R$  für alle  $i \in \mathbb{N}_0$  und  $c_i \neq 0$  nur für endlich viele  $i \in \mathbb{N}_0$ .

Die Elemente von  $R[x]$  heißen Polynome in der Unbestimmten  $x$  mit Koeffizienten in  $R$ , der Ausdruck  $x^i$  heißt Monom vom Grad  $i$ , und der Ausdrücke  $c_i x^i$  bei  $c_i \neq 0$  heißt der Term von  $f$  vom Grad  $i$ .

Für  $f$  in (2.2.2) definiert man den Grad  $\deg f$  von  $f$  als den maximalen Grad eines Terms in  $f$ . Im Fall, dass  $f$  ein Nullpolynom ist (das Polynom mit  $c_i = 0$  für alle  $i \in \mathbb{N}_0$ ), setzt man  $\deg f = -\infty$ . Ein Polynom vom Grad höchstens  $N \in \mathbb{N}_0$  lässt sich somit als  $\sum_{i=0}^N c_i x^i$  darstellen.

Die Gleichheit der Polynome wird durch den Koeffizientenvergleich definiert.

In  $R[x]$  definiert man  $+$  und  $\cdot$  für  $f = \sum_{i=0}^{\infty} c_i x^i$  und  $g = \sum_{i=0}^{\infty} d_i x^i$  durch:

$$f + g := \sum_{i=0}^{\infty} (c_i + d_i) x^i \tag{2.2.3}$$

$$f \cdot g := \sum_{i,j \in \mathbb{N}_0} c_i d_j x^{i+j} = \sum_{k=0}^{\infty} \left( \sum_{i=0}^k c_i d_{i-k} \right) x^k \tag{2.2.4}$$

**Prop.** Sei  $R$  ein kommutativer Ring und  $x$  eine Unbestimmte, dann ist

$R[x]$  ein kommutativer Ring. (Ohne Beweis).

**Bsp.** Der Fall  $R = \mathbb{Z}$ :

$$10 \in \mathbb{Z}[x] \text{ denn } 10 = \sum_{i=0}^{\infty} c_i x^i \quad \text{mit } c_0 = 10, c_1 = c_2 = \dots = 0.$$

$$5x \in \mathbb{Z}[x] \text{ denn } 5x = \sum_{i=0}^{\infty} c_i x^i \quad \text{mit } c_0 = 0, c_1 = 5, c_2 = c_3 = \dots = 0$$

$$1 - 5x + x^3 \in \mathbb{Z}[x] \text{ denn } c_0 = 1, c_1 = -5, c_2 = 0, c_3 = 1, c_4 = c_5 = \dots = 0.$$

**Bsp.**

$$f = 1 + 2x + 5x^2$$

$$g = 3 - 7x + 6x^2 + x^3$$

$$f + g = (1 + 3) + (2 + (-7))x + (5 + 6)x^2 + (0 + 1)x^5$$

$$\begin{aligned} f \cdot g &= (1 \cdot 3) + (1 \cdot (-7) + 3 \cdot 2)x + (1 \cdot 6 + 2 \cdot (-7) + 5 \cdot 3)x^2 \\ &\quad + (1 \cdot 1 + 2 \cdot 6 + 5 \cdot (-7))x^3 + (2 \cdot 1 + 5 \cdot 6)x^4 + (5 \cdot 1)x^5 \end{aligned}$$

Bem.

$$\begin{array}{ll} f \xrightarrow{D^0} f & f \xrightarrow{D^2} f'' \\ f \xrightarrow{D} f' & f \xrightarrow{D^3} f''' \end{array}$$

$D$  kann man in ein Polynom aus  $R[x]$  einsetzen.

$$\begin{aligned} 6(x-1)(x-1) &= x^2 - 2x + 1 \\ \Leftrightarrow (x^1 - x^0)(x^1 - x^0) &= x^2 - 2x^1 + x^0 \end{aligned}$$

Für  $x$  den Operator  $D$  einsetzen:

$$\begin{aligned} \underbrace{(D^1 - D^0)(D^1 - D^0)}_{\text{Operator}} &= \underbrace{D^2 - 2D^1 + D^0}_{\text{derselbe Operator}} \\ (D^1 - D^0)(D^1 - D^0)f &= (f' - f)' - (f' - f) \\ (D^2 - 2D^1 + D^0)f &= f'' - 2f' + f \\ \Rightarrow (f' - f)' - (f' - f) &= f'' - 2f' + f \end{aligned}$$

(Ende der Nebenbemerkung)

Sei  $f \in R[x]$ ,  $f = \sum_{i=0}^{\infty} c_i x^i$ . Sei  $a \in R$ . Dann heißt  $f(a) := \sum_{i=0}^{\infty} c_i a^i$  der Wert von  $f$  an der Stelle  $a$ .

Somit bestimmt  $f$  die Funktion  $a \in R \mapsto f(a)$  von  $R$  nach  $R$ . Diese Funktion nennt man die durch  $f$  bestimmte Polynomfunktion.

Seien  $f, g \in R[x]$  und sei  $f = \sum_{i=0}^{\infty} c_i x^i$ . Wir definieren  $f(g(x))$  durch

$$f(g(x)) := \sum_{i=0}^{\infty} c_i g(x)^i \in R[x].$$

Sei  $f \in R[x]$  und sei  $a \in R$ . Dann heißt  $a$  eine *Wurzel* (oder *Nullstelle*) von  $f$ , falls  $f(a) = 0$  gilt.

**Prop.** *Sei  $R$  ein kommutativer Ring, sei  $x$  eine Unbekannte und sei  $f \in R[x] \setminus 0$ . Sei  $a \in R$  eine Nullstelle von  $f$ . Dann existiert ein eindeutiges  $g \in R[x]$  mit  $f(x) = (x - a)g(x)$ .*

*Beweis.*  $a$  ist Nullstelle von  $f(x) \Rightarrow x = 0$  ist Nullstelle von  $f(x + a)$ . Sei  $f(x + a) = \sum_{i=0}^{\infty} c_i x^i$  mit  $c_i \in R$  für alle  $i \in \mathbb{N}_0$ .

$$\begin{aligned} 0 &= f(x + a) = c_0 + c_1 \cdot 0^1 + \dots = c_0 \\ \Rightarrow f(x + a) &= c_1 x^1 + \dots + c_n x_n \quad (\text{für } n = \deg(f) \in \mathbb{N}) \\ &= x \underbrace{(c_1 x^0 + \dots + c_n x^{n-1})}_{=: h \in R[x]} \end{aligned}$$

Man hat  $f(x + a) = xh(x)$ . Wir setzen für  $x$  das Polynom  $x - a$  ein und erhalten  $f(x) = f(x - a + a) = (x - a)h(x - a)$ . Das Polynom  $g(x) := h(x - a)$  erfüllt die Behauptung.

Wir zeigen die Eindeutigkeit von  $g$ . Seien  $g, \tilde{g} \in R[x]$  mit  $f(x) = (x - a)g(x) = (x - a)\tilde{g}(x)$ . Wir setzen für  $x$  das Polynom  $x + a$  ein:  $f(x + a) = xg(x + a) = x\tilde{g}(x + a)$ . Nach Konstruktion sieht man, dass  $g(x + a) = \tilde{g}(x + a) = h(x)$  gilt. Wir setzen nun für  $x$  das Polynom  $x - a$  ein:  $g(x) = \tilde{g}(x)$ . Das zeigt die Eindeutigkeit.  $\square$

**Bem.** Das Polynom  $g$  aus der vorigen Proposition lässt sich durch Division von Polynomen ermitteln. Ein konkretes Beispiel:

**Bsp.**  $f := x^3 - 5x^2 + 7x - 2, a = 2$ . Was ist  $f(a)$ ? Wenn  $f(a) = 0$ , was ist  $g$  mit  $f(x) = (x - a)g(x)$ ?

$$\begin{array}{rcl}
 x^3 - 5x^2 + 7x - 2 & = & (x - 2) (x^2 - 3x + 1) \quad \Rightarrow \quad f(2) = 0 \\
 - x^3 + 2x^2 & & \\
 \hline
 - 3x^2 + 7x & & \\
 3x^2 - 6x & & \\
 \hline
 x - 2 & & \\
 - x + 2 & & \\
 \hline
 0 & &
 \end{array}$$

$$\begin{array}{rcl}
 x^3 - 5x^2 + 7x - 2 = (x - 3)(x^2 - 2x + 1) + 1 & \implies & f(3) = 1 \\
 - x^3 + 3x^2 \\
 \hline
 - 2x^2 + 7x \\
 - 2x^2 - 6x \\
 \hline
 x - 2 \\
 - x + 3 \\
 \hline
 1
 \end{array}$$

### 2.2.3 Restklassenringe

**Prop.** Sei  $m \in \mathbb{N}$  und seien  $A, B \in \mathbb{Z}/m\mathbb{Z}$ . Dann existiert eine eindeutige Restklasse  $C \in \mathbb{Z}/m\mathbb{Z}$  mit  $a \cdot b \in C$  für alle  $a \in A$  und  $b \in B$ .

Für  $C, A, B$  wie oben schreibt man  $C := A \cdot B$ . Somit hat man eine Multiplikation auf  $\mathbb{Z}/m\mathbb{Z}$  eingeführt.

*Beweis.* Übungsaufgabe. □

**Prop.** Sei  $m \in \mathbb{N}$ . Dann ist  $(\mathbb{Z}/m\mathbb{Z}, +, \cdot)$  ein kommutativer Ring mit 1.

*Beweis.* Die Ringeigenschaften lassen sich direkt verifizieren.  $\square$

**Bsp**  $(\mathbb{Z}/2\mathbb{Z})$ . Wir schreiben Einfachheit halber  $x$  an der Stelle von  $[x]$ . Die Rechentafeln für  $\mathbb{Z}/2\mathbb{Z}$  sind:

+	0	1	
0	0	1	0
1	1	0	0

·	0	1	
0	0	0	0
1	0	1	0

**Bsp**  $(\mathbb{Z}/3\mathbb{Z})$ . Mit der Vereinbarung wie in den vorigen Tafeln haben die Rechentafeln für  $\mathbb{Z}/3\mathbb{Z}$  die folgende Form:

+	0	1	2	
0	0	1	2	0
1	1	2	0	0
2	2	0	1	0

·	0	1	2	
0	0	0	0	0
1	0	1	2	0
2	0	2	1	0

## 2.3 Körper

### 2.3.1 Körper

Sei  $\mathbb{K}$  Menge. Seien  $+$  und  $\cdot$  Binäroperationen auf  $\mathbb{K}$ .  $\mathbb{K}$  heißt *Körper*, falls

- $(\mathbb{K}, +)$  ist eine kommutative Gruppe (mit dem neutralen Element 0)
- $(\mathbb{K} \setminus \{0\}, \cdot)$  ist eine kommutative Gruppe (mit dem neutralen Element 1)
- Es gilt das Distributivgesetz  $a(b + c) = ab + ac$  (mit  $a, b, c \in \mathbb{K}$ ).

Mit anderen Worten: Ein Körper  $(\mathbb{K}, +, \cdot)$  ist ein kommutativer Ring mit 1, in dem  $0 \neq 1$  gilt und alle Elemente aus  $\mathbb{K} \setminus \{0\}$  invertierbar sind.

Bsp.

- $(\mathbb{Z}, +, \cdot)$  ist Ring, kein Körper

- $(\mathbb{Q}, +, \cdot)$  ist Körper
- $(\mathbb{R}[x], +, \cdot)$  ist Ring, kein Körper
- $(\mathbb{R}, +, \cdot)$  ist Körper, auf dem die Analysis aufgebaut wird
- $(\mathbb{Z}/m\mathbb{Z}, +, \cdot)$  mit  $m \in \mathbb{N}$  ist Ring

### 2.3.2 Restklassenkörper

Seien  $a, b \in \mathbb{Z}$ . Dann heißt im Fall, dass  $a$  und  $b$  nicht beide gleich 0 sind, das größte  $k \in \mathbb{N}$ , das  $a$  sowie  $b$  teilt, der größte gemeinsame Teiler von  $a$  und  $b$ . Im Fall  $a = b = 0$  setzt man den größten gemeinsamen Teiler von  $a$  und  $b$  gleich 0. Die Bezeichnungen dazu sind:  $\gcd(a, b)$  und  $\text{ggT}(a, b)$ .

**Prop.** Seien  $a, b \in \mathbb{Z}$ . Dann existieren  $x, y \in \mathbb{Z}$  mit  $xa + yb = \text{ggT}(a, b)$ .

*Beweis.* Übungsaufgabe. Diese Aussage hat einen konstruktiven (algorithmischen) Beweis. Der Algorithmus heißt der erweiterte Euklidische Algorithmus.  $\square$

**Thm.** Sei  $m \in \mathbb{N}$ .  $\mathbb{Z}/m\mathbb{Z}$  ist genau dann ein Körper, wenn  $m$  eine Primzahl ist.

*Beweis.* Wenn  $m$  keine Primzahl ist, dann ist  $\mathbb{Z}/m\mathbb{Z}$  kein Körper (nicht nullteilerfrei). Wenn  $m$  eine Primzahl ist, ist  $\mathbb{Z}/m\mathbb{Z}$  ein Körper.  $\square$

### 2.3.3 Körper Komplexe Zahlen

Man nennt einen Körper  $\mathbb{K}$  *algebraisch abgeschlossen*, wenn jedes Polynom  $f \in \mathbb{K}[x]$  mit  $\deg f > 0$  eine Nullstelle hat.

**Bem.**  $\mathbb{R}$  ist *nicht* algebraisch abgeschlossen.

**Bem** (Intuition zu  $\mathbb{C}$ ). Man führt ein formales Symbol  $i$  ein (imaginäre Einheit), mit der Eigenschaft  $i^2 = -1$  ( $i$  ist Nullstelle von  $x^2 + 1$ ).

$$C := \{a + ib : a, b \in \mathbb{R}\}$$

**Bem** (Formale Definition von  $\mathbb{C}$ ). Die Menge  $\mathbb{C}$  aller komplexen Zahlen ist als  $\mathbb{R} \times \mathbb{R}$  mit der Addition

$$(a_1, b_1) + (a_2, b_2) := (a_1 + a_2, b_1 + b_2)$$

und der Multiplikation

$$(a_1, b_1) \cdot (a_2, b_2) := (a_1 a_2 - b_1 b_2, a_1 b_2 + a_2 b_1)$$

definiert ( $\forall a_1, a_2, b_1, b_2 \in \mathbb{R}$ ).

Jede reelle Zahl  $a \in \mathbb{R}$  wird als  $(a, 0) \in \mathbb{C}$  definiert. Wir interpretieren also  $\mathbb{R}$  als Teilmenge von  $\mathbb{C}$ .

Die Zahl  $i := (0, 1) \in \mathbb{C}$  heißt die imaginäre Einheit. Somit lässt sich jedes  $z \in \mathbb{C}$  als  $z = a + ib$  mit  $a, b \in \mathbb{R}$  schreiben. Dabei heißt  $a$  der Realteil

von  $z$  ( $\Re(z)$  bzw.  $\operatorname{Re}(z)$ ) und  $b$  der Imaginärteil von  $z$  ( $\Im(z)$  bzw.  $\operatorname{Im}(z)$ ).

Die Zahl  $\bar{z} = a - ib$  heißt komplex konjugiert zu  $z$ .

**Prop.**  $\mathbb{C}$  ist ein Körper.

*Beweis.*  $\mathbb{C}$  ist ein kommutativer Ring mit 1 (lässt sich direkt verifizieren).

Wir zeigen, dass für jedes  $z \in \mathbb{C} \setminus \{0\}$  eine Zahl  $w \in \mathbb{C}$  mit  $z \cdot w = 1$  existiert. Sei  $z = a + ib$  mit  $a, b \in \mathbb{R}$ . Man setze  $w := \frac{a - ib}{a^2 + b^2}$ . Dann ist

$$\begin{aligned} z \cdot w &= (a + ib)(a - ib)(a^2 + b^2)^{-1} \\ &= (a^2 + b^2)(a^2 + b^2)^{-1} \\ &= 1 \end{aligned}$$

□

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

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

der Betrag von  $z$ .

**Thm.** Sei  $f \in \mathbb{C}[x]$  mit  $n := \deg f > 0$ . Dann existieren  $\lambda_1, \lambda_2, \dots, \lambda_n, c \in \mathbb{C}$  mit

$$f(x) = c(x - \lambda_1) \cdot \dots \cdot (x - \lambda_n). \quad (2.3.1)$$

**Bem.** Wenn ein  $\lambda \in \mathbb{C}$  in der Folge  $\lambda_1, \lambda_2, \dots, \lambda_n$   $k$ -mal vorkommt, dann heißt  $\lambda$  Nullstelle von  $f$  der Vielfachheit  $k$ .

**Bsp.**

$$f := x^2 - 2x + 1 = (x - 1)^2 \quad \text{mit Nullstelle 1 der Vielfachheit 2}$$

$$f := x^2 + 3x + 2 = (x + 1)(x + 2) \quad \text{mit Nullstellen -1, -2 der Vielfachheit von je 1}$$

$$f := x^2 + 4 \quad \text{mit Nullstellen } 2i, -2i \text{ der Vielfachheit von je 1}$$

# A Vollständige Induktion

Sei  $P : \mathbb{N} \rightarrow \{\text{falsch}, \text{wahr}\}$ . Dann sind die folgenden Bedingungen äquivalent:

- (a)  $P(n)$  gilt für alle  $n \in \mathbb{N}$ .
- (b)  $P(1)$  gilt, und für alle  $n \in \mathbb{N}$  gilt die Implikation  $P(n) \Rightarrow P(n+1)$ .

Wenn man (a) mit Hilfe von (b) zeigt, so sagt man, dass man die *vollständige Induktion* über  $n$  benutzt. Die Aussage  $P(1)$  nennt man den *Induktionsanfang*, die Annahme,  $P(n)$  sei erfüllt, die *Induktionsvoraussetzung*, und die Herleitung von  $P(n+1)$  aus  $P(n)$  den *Induktionsschritt*.

Wenn man die Äquivalenz von (a) und (b) für das Prädikat  $P(1) \wedge \dots \wedge P(n)$  an der Stelle von  $P(n)$  benutzt, so erhält man, dass (a) auch zur

folgenden Behauptung äquivalent ist:

- (c)  $P(1)$  gilt, und für alle  $n \in \mathbb{N}$  gilt die Implikation  $P(1) \wedge \cdots \wedge P(n) \Rightarrow P(n+1)$ .

Aussage (c) ist eine Variante der vollständigen Induktion, mit der Induktionsannahme,  $P(k)$  sei für alle  $k \in \mathbb{N}$  mit  $k \leq n$  erfüllt.

Des Weiteren benutzt man naheliegende Varianten der vollständigen Induktion für Aussagen der Form “ $P(n)$  gilt für alle  $n \geq n_0$ ” für Prädikate  $P : \{n \in \mathbb{Z} : n \geq n_0\} \rightarrow \{\text{falsch}, \text{wahr}\}$  und  $n_0 \in \mathbb{Z}$ . Bei der Induktion in diesem Fall ist die Aussage  $P(n_0)$  der Induktionsanfang.

**Bsp.** Wir zeigen den Existenz-Teil des Fundamentalsatzes der Arithmetik: Jede natürliche Zahl ist Produkt endlich vieler Primzahlen. Etwas formaler heißt das: jedes  $n \in \mathbb{N}$  besitzt die Faktorisierung  $\prod_{i=1}^k p_i$  mit  $k \in \mathbb{N}_0$ , wobei alle  $p_i$  in dieser Faktorisierung Primzahlen sind. Wir zeigen die Aussage durch Induktion über  $n$ . Die Aussage gilt für  $n = 1$  mit  $k = 0$  ( $1$  ist

Produkt von 0 Primzahlen). Sei  $n \in \mathbb{N}$  und man nehme an, allen Zahlen aus  $\{1, \dots, n\}$  lassen sich in endlich viele Primfaktoren zerlegen. Wir betrachten nun die Zahl  $n+1$ . Ist  $n+1$  Primzahl, so ist sie Produkt  $n+1 = \prod_{i=1}^k p_k$  mit  $k = 1$  und  $p_1 = n+1$ . Wenn  $n+1$  keine Primzahl ist, so existieren  $a, b \in \mathbb{N}$  mit  $a, b \geq 2$  und  $n+1 = ab$ . Aus  $a, b \geq 2$  und  $ab = n+1$  folgt, dass  $a$  und  $b$  in  $\{1, \dots, n\}$  liegen: denn man hat  $a = \frac{n+1}{b} \leq \frac{n+1}{2} \leq n$ , und analog auch  $b \leq n$ . Nach der Induktionsvoraussetzung lassen sich  $a$  und  $b$  in endlich viele Primfaktoren zerlegen:  $a = \prod_{i=1}^s q_i$  mit  $b = \prod_{j=1}^t r_j$ , wobei  $s, t \in \mathbb{N}$  und alle  $q_i$  und  $r_j$  Primzahlen sind. Dann ist  $n+1 = ab = q_1 \cdot \dots \cdot q_s \cdot r_1 \cdot \dots \cdot r_t$  Faktorisierung von  $n+1$  in  $s+t \in \mathbb{N}$  Primfaktoren.

## B Eliminationsverfahren zur Lösung von LGS

### B.1 Lineare Gleichungssysteme (LGS)

Im Folgenden seien  $\mathbb{K}$  ein Körper,  $m, n \in \mathbb{N}$  und  $a_{ij}, b_i \in \mathbb{K}$  für alle  $i \in \{1, \dots, m\}, j \in \{1, \dots, n\}$ . Wir definieren ein *lineares Gleichungssystem* bzgl. des Körpers  $\mathbb{K}$  als ein System der Form

$$\sum_{j=1}^n a_{ij}x_j = b_i \quad (i = 1, \dots, m). \quad (\text{B.1.1})$$

mit unbekannten  $x_1, \dots, x_n$  aus  $\mathbb{K}$ . Die Menge

$$X := \{(x_1, \dots, x_n) \in \mathbb{K}^n : (\text{B.1.1}) \text{ gilt}\} \quad (\text{B.1.2})$$

heißt die Lösungsmenge von (B.1.1).

Das Ziel ist die folgenden Fragen algorithmisch zu beantworten:

(i) Ist  $X = \emptyset$ ?

(ii) Besteht  $X$  aus genau einem Element aus  $\mathbb{K}^n$ ?

Außerdem wollen wir in dem Fall  $X \neq \emptyset$  eine explizite Beschreibung von  $X$  ermitteln. Man schreibt (B.1.1) oft als eine Tabelle:

	$x_1$	$\dots$	$x_n$	
$z_1$	$a_{1,1}$	$\dots$	$a_{1,n}$	$b_1$
$z_2$	$a_{2,1}$	$\dots$	$a_{2,n}$	$b_2$
$\vdots$	$\vdots$		$\vdots$	$\vdots$
$z_m$	$a_{m,1}$	$\dots$	$a_{m,n}$	$b_m$

Hierbei nutzen wir  $z_i$  als eine (optionale) Bezeichnung für die  $i$ -te Gleichung ( $z$  wie Zeile).

**Bsp.** Die Gleichung

$$\begin{cases} x_1 + x_2 = 15 \\ x_1 - 2x_2 = 3 \end{cases}$$

wird tabellarisch als

	$x_1$	$x_2$	
$z_1$	1	1	15
$z_2$	1	-2	3

dargestellt.

## B.2 Elementartransformationen eines LGS

Für Gleichungen  $z_i$  und  $z_k$  mit  $i, k \in \{1, \dots, m\}$  und einem Parameter  $\alpha \in \mathbb{K}$  definieren wir die Gleichung

$$\begin{aligned}
 z_i + \alpha z_k &:= \sum_{j=1}^n a_{ij}x_j + \alpha \left( \sum_{j=1}^n a_{kj}x_j \right) \\
 &= \sum_{j=1}^n (\alpha_{ij} + \alpha a_{kj}) x_j = b_i + \alpha b_k.
 \end{aligned} \tag{B.2.1}$$

Wir führen die folgenden 3 Arten der Elementartransformationen für (B.1.1) ein:

**Typ 1.** Für  $i, k \in \{1, \dots, n\}$  werden die Gleichungen  $z_i$  und  $z_k$  vertauscht.

**Bezeichnung:**  $z_i \leftrightarrow z_k$ .

**Typ 2.** Für  $i \in \{1, \dots, m\}$  und  $\alpha \in \mathbb{K} \setminus \{0\}$  wird  $z_i$  durch  $\alpha z_i$  ersetzt.

**Bezeichnung:**  $z_i := \alpha z_i$

**Typ 3.** Für  $i, k \in \{1, \dots, m\}$  mit  $i \neq k$  und für  $\alpha \in \mathbb{K}$  wird  $z_i$  durch  $z_i + \alpha z_k$  ersetzt.

**Bezeichnung:**  $z_i := z_i + \alpha z_k$ .

**Prop.** *Elementartransformationen eines linearen Gleichungssystems ändern die Lösungsmenge des Systems nicht.*

*Beweis.*

Typ 1 ist klar.

Typ 2: Sei  $i \in \{1, \dots, m\}$  und  $\alpha \in \mathbb{K} \setminus \{0\}$ . Ist  $z_i$  erfüllt, dann ist auch  $\alpha z_i$  erfüllt. Umgekehrt: Ist  $\alpha z_i$  erfüllt, dann ist auch  $\alpha^{-1}(\alpha z_i) = z_i$  erfüllt.

Typ 3: Seien  $i, k \in \{1, \dots, m\}$  und  $\alpha \in \mathbb{K}$ . Wenn  $z_i$  und  $z_k$  erfüllt sind, dann sind auch  $z_i + \alpha z_k$  und  $z_k$  erfüllt. Umgekehrt: Wenn  $z_i + \alpha z_k$  und  $z_k$  erfüllt sind, dann sind auch  $(z_i + \alpha z_k) - \alpha z_k = z_i$  und  $z_k$  erfüllt.  $\square$

Für  $i$  und  $j$  sagen wir, dass die Gleichung  $z_i$  die Unbekannte  $x_j$  *enthält*, falls  $a_{ij} \neq 0$  ist. Ansonsten sagen wir, dass  $z_i$  die Unbekannte  $x_i$  *nicht enthält*.

### B.2.1 Gauß-Verfahren

Wir sagen, dass die Gleichung  $z_i$  die  $j$ -te Variable des LGS *enthält*, wenn  $a_{ij} \neq 0$  gilt. Wenn das LGS eine Variable  $x_j$  in einer Gleichung enthält, so kann die Variable aus allen anderen Gleichungen des Systems durch Trans-

formationen vom Typ 3 entfernt werden. Man erhält somit ein äquivalentes LGS, in dem  $x_j$  in einer einzigen Gleichung enthalten ist.

Der Grundgedanke der Eliminationsverfahren zur Lösung der LGS ist, dass man nun  $x_j$  aus der Gleichung  $z_i$  eindeutig ermitteln kann, sobald die Werte aller anderen Variablen festgelegt sind. Die Eliminationsverfahren basieren auf der iterativen Anwendung der vorigen Beobachtung.

Das *Gauß-Verfahren* geht nach dem folgenden Muster vor.

- In der ersten Iteration wird eine Variable  $x_{j_1}$  gefunden, die in einer der Gleichungen  $z_1, \dots, z_m$  enthalten ist. Durch das Vertauschen der Gleichungen können wir sicher stellen, dass  $x_{j_1}$  in  $z_1$  enthalten ist. Nun wird  $x_{j_1}$  aus  $z_2, \dots, z_m$  mit Hilfe der Transformationen vom Typ 3 entfernt.
- In der zweiten Iteration wird eine Variable  $x_{j_2}$  gefunden, die einer

der Gleichungen  $z_2, \dots, z_m$  enthalten ist. Durch das Vertauschen der Gleichungen  $z_2, \dots, z_m$  können wir sicher stellen, dass  $x_{j_2}$  in  $z_2$  enthalten ist. Die Variable  $x_{j_2}$  wird aus den Gleichungen  $z_3, \dots, z_m$  mit Hilfe der Transformationen vom Typ 3 entfernt.

- Nach der  $k$ -ten Iteration hat man im aktuellen LGS  $k$  Variablen  $x_{j_1}, \dots, x_{j_k}$  fixiert, wobei  $x_{j_s}$  in der Gleichung  $z_s$  aber nicht in den Gleichungen  $z_i$  mit  $i > s$  enthalten ist.
- Hat man nach der  $k$ -ten Iteration in den Gleichungen  $z_i$  mit  $i > k$  keine Variablen mehr, so terminiert das Verfahren. Dies ist spätestens in der  $m$ -ten Iteration der Fall, da nach  $m$  Iterationen alle  $m$  Gleichungen des LGS bereits abgearbeitet sind.

Da die Elementartransformationen die Lösungsmenge des LGS nicht ändern, erhalten wir nach jeder Iteration sowie nach der Terminierung

des Verfahrens ein LGS, das zum Ausgangssystem äquivalent ist.

Wenn man im Gauß-Verfahren die Variablen in der Reihenfolge  $x_1, \dots, x_n$  abarbeitet, d.h., bei der Wahl einer Variable  $x_{j_k}$  wählt in der  $k$ -ten Iteration aus den Gleichungen  $z_{k+1}, \dots, z_n$  wählt man stets eine Variable mit dem kleinstmöglichen Index, dann erzeugt das Gauß-Verfahren nach seiner Terminierung ein LGS in der sogenannten Zeilenstufenform.

Man sagt, dass ein LGS in Unbekannten  $x_1, \dots, x_n$  die *Zeilenstufenform* hat, wenn für gewisse Indizes  $1 \leq j_1 < \dots < j_r \leq n$  folgende Eigenschaften erfüllt sind:

- Für jedes  $i = 1, \dots, r$  enthält die  $i$ -te Gleichung des LGS die Variable  $x_{j_i}$  aber nicht die Variablen  $x_1, \dots, x_{j_i-1}$ .
- Für jedes  $i > r$  enthält die  $i$ -te Gleichung des LGS keine Variablen.

A 4x7 matrix in row echelon form. The matrix has 4 rows and 7 columns. The first three columns contain non-zero entries in the first three rows, while the fourth column is all zeros. The fourth column contains only zeros below the first row. The fifth column contains non-zero entries in the first two rows, while the third and fourth columns are all zeros. The sixth column contains non-zero entries in the first three rows, while the seventh column is all zeros. The seventh column contains only zeros below the second row. The entries are marked with asterisks (\*). The first three columns have circled pivots at their top-left positions. Green arrows point from the first three columns to the right, indicating the progression of the elimination process. Braces on the left and right sides group the columns into four groups of two columns each, with the last group being a single column.

Hierbei nennt man  $x_{j_1}, \dots, x_{j_r}$  in manchen Quellen die *Leitvariablen* und alle anderen Variablen die *freien Variablen*. In der linearen Optimierung nennt man  $x_{j_1}, \dots, x_{j_r}$  die Basisvariablen und alle anderen Variablen Nichtbasisvariablen. Die Koeffizienten  $a_{1,j_1}, \dots, a_{r,j_r} \in \mathbb{K} \setminus \{0\}$  nennt man die *Pivotelemente*. Es ist klar, dass man durch Anwendung der Transformationen vom Typ 2 alle Pivotelemente gleich 1 setzen kann.

**Prop.** *Ist LGS ein System in einer Zeilenstufenform mit Leitvariablen  $x_{j_1}, \dots, x_{j_r}$  und  $j_1 < \dots < j_r$ , so gilt:*

- (a) *Das LGS hat genau dann keine Lösungen, wenn es eine Gleichung der Form  $0 = b_i$  mit der rechten Seite  $b_i \in \mathbb{K} \setminus \{0\}$  enthält.*
- (b) *Das LGS hat genau eine eindeutige Lösung, wenn es keine Gleichung der Form  $0 = b_i$  mit  $b_i \in \mathbb{K} \setminus \{0\}$  und keine freien Variablen enthält.*
- (c) *Wenn das LGS mehr als eine Lösung hat, dann gibt es zu jeder Wahl*

*der Werte für freie Variablen eine eindeutige Wahl der Werte der Leitvariablen, für welche das System erfüllt ist.*

*Beweis.* (a) Wenn das System eine Gleichung der Form  $0 = b_i$  mit der rechten Seite  $b_i \neq 0$  enthält, ist das System nicht lösbar. Wenn das System keine solche Gleichungen enthält, so können wir alle freien Variablen (falls man welche hat) gleich 0 setzen. Aus der Gleichung  $z_r$  lässt sich nun der Wert von  $x_{i_r}$  ermitteln, der Gleichung  $z_r$  erfüllt, aus der Gleichung  $z_{r-1}$  der Wert von  $x_{i_{r-1}}$  usw., bis die Werte aller Leitvariablen so gewählt sind, dass die Gleichungen  $z_1, \dots, z_r$  erfüllt sind. Die Gleichungen  $z_i > 0$  mit  $i > r$  haben die Form  $0 = 0$  und sind somit automatisch erfüllt.

(b) Systeme mit Gleichungen der Form  $0 = b_i$  mit einer Rechten Seite  $b_i \neq 0$  haben keine Lösungen. Systeme ohne solche Gleichungen und mit freien Variable, können die freien Variablen auf mindestens zwei verschiedene weisen belegt werden. Zu jeder dieser Belegung finden sich wie im Beweis von (a) die Belegungen der Leitvariablen, die as System erfüllen. So

konstruiert man zwei verschiedene Lösungen des Systems. Alle sonstigen Systeme enthalten weder Gleichungen der Form  $0 = b_i$  mit  $b_i \neq 0$  noch freie Variablen. Bei diesen Systemen ergibt sich ein eindeutiger Wert von  $x_{i_r}$  aus  $z_r$ , dann ein eindeutiger Wert von  $x_{i_{r-1}}$  aus  $z_{r-1}$  usw., bis die Werte von allen Variablen eindeutig festgelegt sind.

(c) folgt direkt aus den vorigen Überlegungen zur Wahl der Werte von Leitvariablen in Abhängigkeit von den freien Variablen.  $\square$

**Bsp.** Lösen Sie das LGS bzgl.  $\mathbb{Q}$ ,  $\mathbb{R}$  und  $\mathbb{Z}/7\mathbb{Z}$ .

$$\begin{cases} x_1 + 2x_2 + 3x_3 = 4 \\ 2x_1 + x_2 - x_3 = 5 \\ x_1 + x_2 + x_3 = 2 \end{cases}$$

$$\left| \begin{array}{ccc|c} x_1 & x_2 & x_3 & \\ \hline 1 & 2 & 3 & 4 \\ 2 & 1 & -1 & 5 \\ 1 & 1 & 1 & 2 \end{array} \right|$$

$z_2 := z_2 - 2z_1$   
 $z_3 := z_3 - z_1$

	$x_1$	$x_2$	$x_3$	
$\bar{z}_1$	1	2	3	4
$\bar{z}_2$	0	-3	-7	-3
$\bar{z}_3$	0	-1	-2	-2

Typ 2, 3.

	$x_1$	$x_2$	$x_3$	
$\bar{z}_1$	1	2	3	4
$\bar{z}_2$	0	1	2	2
$\bar{z}_3$	0	-3	-7	-3

$$\bar{z}_3 := \bar{z}_3 + 3\bar{z}_2$$

	$x_1$	$x_2$	$x_3$	
$\bar{z}_1$	1	2	3	4
$\bar{z}_2$	0	1	2	2
$\bar{z}_3$	0	0	-1	3

$$x_3 = -3$$

$$x_2 = 8$$

$$x_1 = -3$$

$$\begin{aligned}
 x_2 + 2x_3 &= 2 \\
 x_2 &= 2 - 2x_3 \\
 &= 2 - 2 \cdot (-3) \\
 &= 8
 \end{aligned}$$

$$x_1 + 2x_2 + 3x_3 = 4$$

$$x_1 = 4 - 2x_2 - 3x_3$$

$$\begin{aligned}
 &= 4 - 2 \cdot 8 - 3 \cdot (-3) \\
 &= 4 - 16 + 9
 \end{aligned}$$

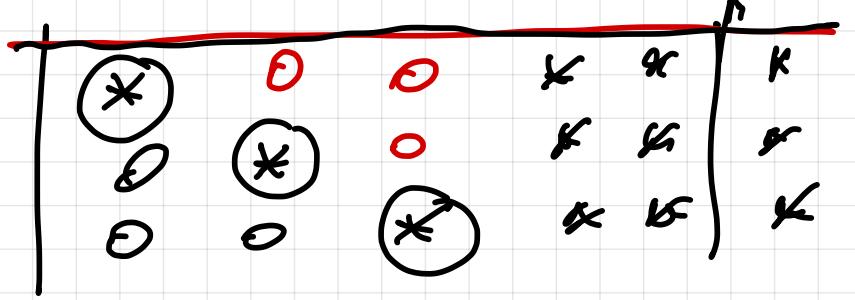
## B.2.2 Gauß-Jordan-Verfahren

Das Gauß-Jordan-Verfahren ist eine geringfügige Modifikation des Gauß-Verfahrens bei der man in der  $k$ -ten Iteration die  $k$ -te Leitvariable  $x_{j_k}$  nicht nur aus den nachfolgenden Gleichungen  $z_i$  mit  $i > k$  sondern auch aus den vorigen Gleichungen  $z_i$  mit  $i < k$  mit Transformationen vom Typ 3 entfernt. Wenn die Variablen in der Reihenfolge  $x_1, \dots, x_n$  abgearbeitet werden, so überführt das Gauß-Jordan-Verfahren das System in die sogenannte reduzierte Zeilenstufenform. Wir sagen, dass ein System in einer reduzierten Stufenform ist, wenn es in einer Stufenform ist und jede Leitvariable in genau einer Gleichung enthalten ist<sup>3</sup>

**Bsp.** Geben Sie eine parametrische Beschreibung der Geraden in  $\mathbb{R}^3$ , die

---

<sup>3</sup>in manchen Quellen fordert man zusätzlich, dass alle Pivotelemente gleich 1 sind. Das lässt sich mit der Verwendung der Transformationen vom Typ 3 sicherstellen.



als Lösungsmenge des LGS

$$\begin{cases} x_1 + 2x_2 + 3x_3 = 4 \\ 2x_1 + x_2 - x_3 = 5 \end{cases}$$

gegeben ist.

	$x_1$	$x_2$	$x_3$	
$z_1$	1	2	3	4
$z_2$	2	1	-1	5

$$z_2 := z_2 - 2z_1$$

	$\cancel{1}$	2	3	
$z_2$	0	$\cancel{-3}$	-7	-3

$$z_2 := -\frac{1}{3} z_2$$

	$\cancel{1}$	2	3	
$z_2$	0	$\cancel{\frac{7}{3}}$	$\frac{7}{3}$	1

$$z_1 := z_1 - 2z_2$$

	$\cancel{1}$	0	$-\frac{5}{3}$	2
$z_2$	0	$\cancel{1}$	$\frac{7}{3}$	1

$$\begin{aligned} 3 - 2 \cdot \frac{7}{3} &= \\ = \frac{9}{3} - \frac{14}{3} &= -\frac{5}{3} \end{aligned}$$

Stufenform

Stufenform

Reduzierte  
Stufenform

$$\left\{ \begin{array}{l} x_1 - \frac{5}{3}x_3 = 2 \\ x_2 + \frac{7}{3}x_3 = 1 \end{array} \right.$$

$$\left\{ \begin{array}{l} x_1 = 2 + \frac{5}{3}x_3 \\ x_2 = 1 - \frac{7}{3}x_3 \end{array} \right.$$

Die Lösungsmenge ist

$$\left\{ \left( 2 + \frac{5}{3}t, 1 - \frac{7}{3}t, t \right) : t \in \mathbb{R} \right\}$$

### 3 Vektorräume

In diesem Kapitel sei  $\mathbb{K}$  ein beliebiger Körper.

#### 3.1 Vektorraum (VR)

##### 3.1.1 $\mathbb{K}^n$ und $\mathbb{K}^X$

(Konkrete Vektorräume)

Sei  $n \in \mathbb{N}$ . Die Menge  $\mathbb{K}^n$  wird

mit der Addition

$$+ : \mathbb{K}^n \times \mathbb{K}^n \rightarrow \mathbb{K}^n$$

und der Skalarmultiplikation

$$\cdot : \mathbb{K} \times \mathbb{K}^n \rightarrow \mathbb{K}^n$$

ausgestattet, die durch

$$(x_1, \dots, x_n) + (y_1, \dots, y_n) := (x_1 + y_1, \dots, x_n + y_n)$$

$$\alpha(x_1, \dots, x_n) := (\alpha x_1, \dots, \alpha x_n)$$

für alle  $(x_1, \dots, x_n), (y_1, \dots, y_n) \in \mathbb{K}^n$  und alle  $\alpha \in \mathbb{K}$  definiert sind.

Sei  $X$  nichtleere Menge ( $X$  kann unendlich sein). Wir definieren in  $\mathbb{K}^X$   
 $+ : \mathbb{K}^X \times \mathbb{K}^X \rightarrow \mathbb{K}^X$  und  $\cdot : \mathbb{K} \times \mathbb{K}^X \rightarrow \mathbb{K}^X$  durch:

$$(f + g)(x) := f(x) + g(x)$$
$$(\alpha \cdot f)(x) := \alpha \cdot f(x)$$

TODO: Bild mit Vektoren  $u = (3, 1)$ ,  $v = (1, 2)$ ,  $u + v$ ,  $2u$  und  $-v$ .

**Bsp.**  $\mathbb{K}^X$  benutzt man unter anderem in Kontexten, bei denen man die Variablen nicht mit Nummern sondern mit anderen Labels versehen will, in  $\mathbb{R}^X$  mit  $X = \{CHF, CNY, EUR, GBP\}$  kann man etwa den Vektor  $v \in \mathbb{R}^X$  der aktuellen Währungskurse Betrachten mit  $v(CHF) = 1.10$ ,  $v(CNY) = 0.15$ ,  $v(EUR) = 1.19$ ,  $v(GBP) = 1.34$ . Solche Kontexte sind zum Beispiel in der kombinatorischen Optimierung reichlich vorhanden. Natürlich kann man in diesem Beispiel die Variablen nummerieren und mit  $\mathbb{R}^4$  an der Stelle von  $\mathbb{R}^X$  arbeiten.

### 3.1.2 Vektorraum über $\mathbb{K}$

(Abstrakter Vektorraum).

Abstrakt definierte Vektorräume geben einem einen konzeptuellen Zugang zur Theorie der Vektorräume, ohne dass man in konkreten Situationen an die Koordinaten (wie bei  $\mathbb{K}^n$ ) denken muss.

Eine Menge  $V$  mit Addition  $+ : V \times V \rightarrow V$  und Skalarmultiplikation  $\cdot : \mathbb{K} \times V \rightarrow V$  heißt Vektorraum über  $\mathbb{K}$ , falls:

(V1)  $(V, +)$  ist abelsche Gruppe

(V2)  $(\lambda + \mu)v = \lambda v + \mu v$  für alle  $\lambda, \mu \in \mathbb{K}$  und  $v, w \in V$

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

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

$$1 \cdot v = v$$

Kurze Erinnerung an abelsche Gruppen: da hat man eine  $0$ , zu jedem  $u$  findet man das  $-u$ , und  $+$  ist kommutativ und assoziativ.

$v + 0 = v$   
 $v + (-v) = 0$   
 $(u + v) + w = u + (v + w)$   
 $u + v = v + u$   
 $\forall u, v, w \in V$

$$\mathbb{R}^2 = \{(x, y) : x, y \in \mathbb{R}\}$$

$$\{(x, y, z) \in \mathbb{R}^3 : x + y + z = 0\}$$

**Bem** (Begriffe / Bezeichnungen).

in Fall wo  $V = \mathbb{K}^n$

- $0 \in V$  heißt Nullvektor  $(0, \dots, 0)$ .
- Elemente von  $V$  heißen Vektoren
- Elemente von  $\mathbb{K}$  heißen Skalare
- $\cdot$  wird oft in Formeln weggelassen ( $\cdot$  höhere Priorität als  $+$ )

**Prop.** Sei  $V$  Vektorraum über  $\mathbb{K}$ . Sei  $\lambda \in \mathbb{K}$  und  $v \in V$ , dann:

(i)  $\lambda \cdot v = 0 \Leftrightarrow \lambda = 0$  oder  $v = 0$

(ii)  $(-1) \cdot v = -v$

*Beweis.* Übungsaufgabe. □

## 3.2 Untervektorräume (UVR)

### 3.2.1 Untervektorraum

Sei  $V$  Vektorraum über  $\mathbb{K}$  und sei  $W \subseteq V$ . Dann heißt  $W$  Untervektorraum von  $V$ , wenn:

$$(\text{UV1}) \quad W \neq 0$$

$$(\text{UV2}) \quad v, w \in W \Rightarrow v + w \in W$$

$$(\text{UV3}) \quad \alpha \in \mathbb{K}, v \in W \Rightarrow \alpha \cdot v \in W$$

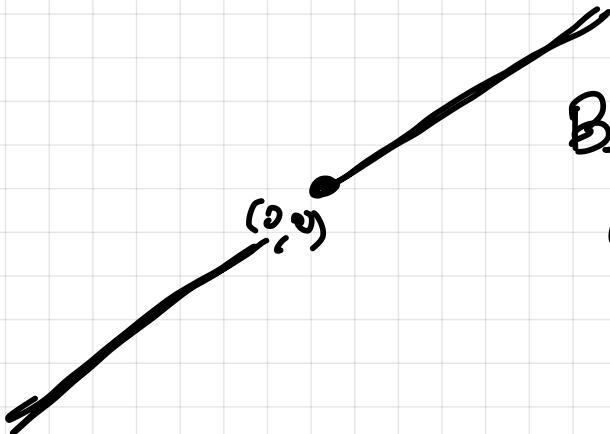
**Bem.** Ein LGS, dessen alle rechten Seiten gleich 0 sind, nennt man homogen. Die Lösungsmenge eines LGS in  $n$  Variablen ist ein UVR von  $\mathbb{K}^n$ .

**Bsp.** Sei  $\mathbb{K} = \mathbb{Z}/2\mathbb{Z} = \{0, 1\}$ . In der Codierungstheorie heißen die UVR

Bsp.

$\mathbb{R}^2$

$\{(0,0)\}$  ist Untervektorraum von  $\mathbb{R}^2$



Beliebige Gerade, die den Nullpunkt  
enthält.

$\mathbb{R}^2$  ist auch Untervektorraum von  $\mathbb{R}^2$ .

von  $\mathbb{K}^n$  binäre lineare Codes der Länge  $n$ . Die Lösungsmenge

$$\begin{aligned} W &= \{(x_1, x_2, x_3) \in \mathbb{K}^3 : x_1 + x_2 + x_3 = 0\} \\ &= \{(0, 0, 0), (1, 0, 1), (0, 1, 1), (1, 1, 0)\}. \end{aligned}$$

der homogenen Gleichung  $x_1 + x_2 + x_3 = 0$  ist etwa ein binärer linearer Code der Länge 3. Bei einer digitalen Kommunikation mit der Verwendung von  $W$ , kann man zwei Bits  $(a, b) \in \mathbb{K}^2$  als den Code  $x = (a, b, a+b) \in W$  verschicken. Empfängt man infolge einer Störung nicht  $x$  sondern einen anderen Vektor  $y \in \mathbb{K}^3$ , der nicht zu  $W$  gehört, so weiß man, dass während der Übertragung ein Fehler aufgetreten ist. Durch unser Beispiel-Code  $W$  kann  $y$  nicht zu  $x$  korrigiert werden, es gibt aber andere Codes, mit denen man eine gewisse Anzahl an Fehlern in  $y$  korrigieren kann.

**Bsp.** In Machine Learning und Statistik können Abhängigkeiten in einer Datenmenge  $u_1, \dots, u_N \in \mathbb{R}^n$  mit Hilfe von Untervektorräumen von  $U$  von  $\mathbb{R}^n$  beschrieben. Mehr dazu später? TODO: Ein Bild für  $n = 2$ . Stich-

punkt: lineare Regression. Hierzu gibt es zahlreiche Anwendungsmöglichkeiten (Abhängigkeiten in Kundenbewertungen bei Netflix, Amazon usw.)

### 3.2.2 Kriterium für Untervektorräume

**Thm.** *Sei  $V$  Vektorraum über  $\mathbb{K}$  und sei  $W \subseteq V$ . Dann sind die folgenden Bedingungen äquivalent:*

- (i)  *$W$  ist Untervektorraum von  $V$*
- (ii)  *$+$  und  $\cdot$  lassen sich von  $V$  auf  $W$  einschränken und darüber hinaus ist  $W$  Vektorraum über  $\mathbb{K}$  bezüglich dieser Operationen.*

*Beweis.* Folgt direkt aus der Definition. □

### 3.2.3 Durchschnitt von Untervektorräumen

**Prop.** Sei  $V$  Vektorraum über  $\mathbb{K}$ . Sei  $I$  nichtleere Menge und sei  $W_i$  Untervektorraum von  $V$  für jedes  $i \in I$ . Dann ist auch

$$W := \bigcap_{i \in I} W_i$$

ein Untervektorraum von  $V$ .

*Beweis.* Erst  $0 \in W$  nachweisen, der Rest folgt direkt.  $\square$

### 3.2.4 Vereinigung von Untervektorräumen

**Prop.** Sei  $V$  Vektorraum über  $\mathbb{K}$ . Seien  $W$  und  $W'$  Untervektorräume von  $V$  derart, dass  $W \cup W'$  auch ein Untervektorraum von  $V$  ist. Dann gilt

$$W \subseteq W' \text{ oder } W' \subseteq W.$$

*Beweis.* Wir nehmen  $W \not\subseteq W'$  an und zeigen  $W' \subseteq W$ . Fixiere  $w \in W \setminus W'$ .

Sei  $w' \in W'$  beliebig. Zu zeigen:  $w' \in W$ .

Da  $W \cup W'$  Untervektorraum ist, folgt aus  $w, w' \in W \cup W'$ , dass  $w + w' \in W \cup W'$ . Es gilt  $w + w' \in W \setminus W'$ , denn sonst  $w + w' \in W' \Rightarrow w = (w + w') - \underbrace{w'}_{\in W'} \in W'$  ( $\Rightarrow \not\subseteq$  zur Wahl von  $w \in W \setminus W'$ ).

Es folgt  $w + w' \in W \Rightarrow w, w + w' \in W$  und weil  $W$  Untervektorraum ist, gilt  $w' = (w + w') - \underbrace{w}_{\in W} \in W$ .  $\square$

### 3.3 Linearkombinationen

#### 3.3.1 Linearkombination

Sei  $r \in \mathbb{N}_0$  und seien  $v_1, \dots, v_r$  Vektoren in einem Vektorraum  $V$  über  $\mathbb{K}$ .

Seien  $\lambda_1, \dots, \lambda_r \in \mathbb{K}$ . Dann heißen  $\lambda_1 v_1 + \dots + \lambda_r v_r$  die Linearkombination von  $v_1, \dots, v_r$  mit Koeffizienten  $\lambda_1, \dots, \lambda_r$ .

Im Fall  $r = 0$  setzt man die Linearkombination gleich  $0 \in V$ . Sei  $I$  Menge und  $v_i \in V$  für jedes  $i \in I$ . Dann heißt

$$\text{lin}_{\mathbb{K}}(v_i)_{i \in I} := \{\lambda_1 v_{i_1} + \dots + \lambda_r v_{i_r} : \lambda_1, \dots, \lambda_r \in \mathbb{K}, i_1, \dots, i_r \in I, r \in \mathbb{N}_0\} \quad (3.3.1)$$

die *lineare Hülle* von  $(v_i)_{i \in I}$ . Insbesondere gilt

$$\text{lin}_{\mathbb{K}}(v_1, \dots, v_r) := \{\lambda_1 v_1 + \dots + \lambda_r v_r : \lambda_1, \dots, \lambda_r \in \mathbb{K}\}. \quad (3.3.2)$$

Hier und im folgenden: wenn die Wahl des Index  $\mathbb{K}$  klar ist, wird der Index weg gelassen.

Man kann auch von der linearen Hülle einer Menge  $M \subseteq V$  reden:

$$\text{lin}(M) := \{\lambda_1 v_1 + \dots + \lambda_r v_r : \lambda_1, \dots, \lambda_r \in \mathbb{K}, v_1, \dots, v_r \in M, r \in \mathbb{N}_0\} \quad (3.3.3)$$

**Prop.** *Sei  $V$  Vektorraum über  $\mathbb{K}$ . Sei  $I$  Menge und sei  $v_i \in V$  für jedes  $i \in I$ . Dann gilt:*

- (a)  $\text{lin}(v_i)_{i \in I}$  ist Untervektorraum von  $V$ .
- (b) Ist  $W$  Untervektorraum von  $V$  mit  $v_i \in W$  für alle  $i \in I$ , dann ist  $\text{lin}(v_i)_{i \in I} \subseteq W$ .

*Beweis.* Direkt. □

**Bsp.** Beispiele für  $\mathbb{R}^3$

- $\text{lin}(v)$  mit  $v \neq 0$  ist eine Gerade durch 0.
- $\text{lin}(v, w)$  mit  $v \neq 0$  und  $w \notin \text{lin}(v)$  ist eine Ebene durch 0.

### 3.3.2 Lineare Unabhängigkeit und Abhängigkeit

Sei  $r \in \mathbb{N}_0$  und seien  $v_1, \dots, v_r$  Vektoren in einem Vektorraum über  $\mathbb{K}$ . Die Vektoren  $v_1, \dots, v_r$  heißen *linear unabhängig*, falls für alle  $\lambda_1, \dots, \lambda_r \in \mathbb{K}$  aus  $\lambda_1 v_1 + \dots + \lambda_r v_r = 0$  die Gleichungen  $\lambda_1 = \dots = \lambda_r = 0$  folgen.

Bsp.

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \in \mathbb{R}^2$$

Sind diese drei Vektoren linear unabhängig?

Nlin.

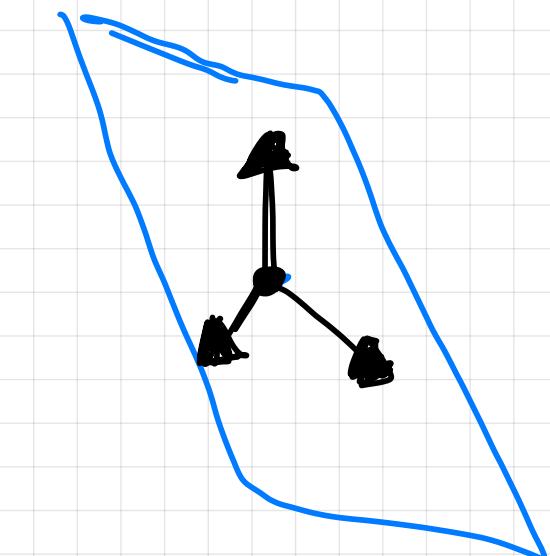
$$1 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 1 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} + (-1) \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

$1, 1, -1$  sind nicht alle gleich 0.

Bsp.

$$\begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \in \mathbb{R}^3$$

Sind diese Vektoren linear unabhängig?



Wir lösen die Vektorgleichung

$$\lambda_1 \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix} + \lambda_2 \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} + \lambda_3 \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\left\{ \begin{array}{l} \lambda_1 + \lambda_2 = 0 \\ \lambda_1 - \lambda_2 + \lambda_3 = 0 \\ -2\lambda_1 - \lambda_3 = 0 \end{array} \right.$$

	$\lambda_1$	$\lambda_2$	$\lambda_3$	
$z_1$	(1)	1	0	0
$z_2$	1	-1	1	0
$z_3$	-2	0	-1	0
<hr/>				
$z_1$	(1)	1	0	0
$z_2$	0	-2	(1)	0
$z_3$	0	2	-1	0
<hr/>				
$z_1$	(1)	1	0	0
$z_2$	0	-2	(1)	0
$z_3$	0	0	0	0

$z_2 := z_2 - z_1$ ,  
 $z_3 := z_3 + 2z_1$

$z_2 := z_2 - z_1$

$$\left\{ \begin{array}{l} \textcircled{\lambda}_1 + \lambda_2 = 0 \\ -2\lambda_2 + \textcircled{\lambda}_3 = 0 \\ 0 = 0 \end{array} \right.$$

$$\boxed{\begin{array}{l} \lambda_1 = -\lambda_2 \\ \lambda_3 = 2\lambda_2 \end{array}}$$

Ich kann L.P.  $\lambda_2 = 1$ ,  $\lambda_1 = -1$ ,  $\lambda_3 = 2$  fassen

$\Rightarrow$  Die drei Vektoren sind linear  
abhängig.

Sei  $I$  Menge und  $v_i \in V$  für jedes  $i \in I$ . Die Vektoren  $v_i$  mit  $i \in I$  heißen linear unabhängig, falls für jede endliche Teilmenge  $J$  von  $I$  die Vektoren  $v_j$  mit  $j \in J$  linear unabhängig sind. Die Vektoren  $v_i$  mit  $i \in I$  heißen *linear abhängig*, falls sie nicht linear unabhängig sind.

Vektoren  $v_1, \dots, v_r \in V$  mit  $r \in \mathbb{N}_0$  sind genau dann linear abhängig, wenn  $\lambda_1, \dots, \lambda_r$  mit  $\lambda_1 v_1 + \dots + \lambda_r v_r = 0$  und  $(\lambda_1, \dots, \lambda_r) \neq (0, \dots, 0)$  existieren.

Ein System aus Nullvektoren ist linear unabhängig.

**Bsp.** Vektoren  $v_1 = (2, -1, 3)$ ,  $v_2 = (1, 2, 0)$  und  $v_3 = (-4, 7, -9)$  aus  $\mathbb{R}^3$  sind linear abhängig, denn  $3v_1 - 2v_2 + v_3 = (0, 0, 0)$ . Je, zwei der drei Vektoren  $v_1, v_2, v_3$  sind aber linear unabhängig.

**Bem.** Für gegebene  $v_1, \dots, v_r \in \mathbb{K}^n$  ist die Vektorgleichung  $\lambda_1 v_1 + \dots + \lambda_r v_r = 0$ , wenn man sie komponentenweise ausschreibt, ein LGS aus  $n$  Gleichungen in  $r$  Unbekannten  $\lambda_1, \dots, \lambda_r \in \mathbb{K}$ . Ob  $v_1, \dots, v_r$  linear (un)abhängig sind kann also mit Hilfe des Gauß-Verfahrens getestet werden.

Was ist ein Vektor ?

Was ist eine Zahl ?

Varianten des Gauß-Verfahrens....

Bsp.

$$\left. \begin{array}{l} \sin^2 x \\ \cos^2 x \\ 1 \end{array} \right\} \text{das sind Vektoren im Raum } \mathbb{R}^3$$

Sind  $\sin^2 x, \cos^2 x, 1$  linear unabhängig?

Diese drei Vektoren sind linear abhängig, denn

$$1 \cdot \sin^2 x + 1 \cdot \cos^2 x + (-1) \cdot 1 = 0$$

Bsp.

find

$$\left( \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \end{array} \right), \left( \begin{array}{c} 1 \\ 1 \\ 1 \\ 1 \end{array} \right), \left( \begin{array}{c} 1 \\ -1 \\ 1 \\ -1 \end{array} \right)$$



linear unabhängig?

$$\lambda_1 \left( \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \end{array} \right) + \lambda_2 \left( \begin{array}{c} 1 \\ 1 \\ 1 \\ 1 \end{array} \right) + \lambda_3 \left( \begin{array}{c} 1 \\ -1 \\ 1 \\ -1 \end{array} \right) = \left( \begin{array}{c} 0 \\ 0 \\ 0 \\ 0 \end{array} \right)$$

Was sind mögliche  $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$

	$\lambda_1$	$\lambda_2$	$\lambda_3$	
$z_1$	1	1	1	0
$z_2$	2	1	-1	0
$z_3$	3	1	1	0
$z_4$	4	1	-1	0
$z_1$	1	1	1	0
$z_2$	1	0	-2	0
$z_3$	2	0	0	0
$z_4$	3	0	-2	0

$$z_2 := z_2 - z_1$$

$$z_3 := z_3 - z_1$$

$$z_4 := z_4 - z_1$$

$$z_1 := z_1 - z_2$$

$$z_3 := z_3 - 2z_2$$

$$z_4 := z_4 - 3z_2$$

$z_1$	0	1	3	0
$z_2$	1	0	-2	0
$z_3$	0	0	4	0
$z_4$	0	0	4	0
<hr/>	<hr/>	<hr/>	<hr/>	<hr/>
$z_1$	0	1	3	0
$z_2$	1	0	-2	0
$z_3$	0	0	1	0
$z_4$	0	0	1	0
<hr/>	<hr/>	<hr/>	<hr/>	<hr/>
$z_1$	0	1	0	0
$z_2$	1	0	0	0
$z_3$	0	0	1	0
$z_4$	0	0	0	0

$$z_3 := \frac{1}{4} z_3$$

$$z_4 := \frac{1}{4} z_2$$

$$z_1 := z_1 - 3 z_3$$

$$z_2 := z_2 + 2 z_3$$

$$z_4 := z_4 - z_3$$

$$\left. \begin{array}{l} z_2 = 0 \\ z_1 = 0 \\ z_3 = 0 \\ 0 = 0 \end{array} \right\} \Rightarrow$$

$\begin{pmatrix} 1 \\ 2 \\ 3 \\ 4 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ -1 \\ 1 \\ -1 \end{pmatrix}$  sind  
 lineare unabhängige.

### 3.3.3 Kriterien für lineare Unabhängigkeit

**Thm.** Sei  $V$  Vektorraum über  $\mathbb{K}$ . Sei  $I$  Menge und sei  $v_i \in V$  für jedes  $i \in I$ . Dann sind die folgenden Bedingungen äquivalent:

- (i) Vektoren  $v_i$  mit  $i \in I$  sind linear unabhängig.
- (ii) Jedes  $v \in \text{lin}(v_i)_{i \in I}$  lässt sich in eindeutiger Weise als Linearkombination von  $(v_i)_{i \in I}$  darstellen.

*Beweis.*

(ii)  $\Rightarrow$  (i): Angenommen, (ii) gilt. Dann lässt sich der Nullvektor in eindeutiger Weise als Linearkombination der Vektoren  $v_i$  mit  $i \in I$ . Diese eindeutige Weise ist: alle Koeffizienten = 0. Also sind  $v_i$  mit  $i \in I$  linear unabhängig.

(i)  $\Rightarrow$  (ii): Angenommen, (i) gilt. Wir zeigen (ii) durch einen Widerspruchsbeweis. Wir nehmen an, (ii) gilt nicht, d.h. für ein  $v \in V$  gilt  $v =$

Bsp.

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \end{pmatrix} \in \mathbb{R}^2$$

$$\begin{pmatrix} 4 \\ 3 \end{pmatrix} = 1 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 3 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} + 0 \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

$$\begin{pmatrix} 4 \\ 3 \end{pmatrix} = 2 \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + 1 \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} + 2 \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

Keine eindeutige Lösung für  $\begin{pmatrix} 4 \\ 3 \end{pmatrix}$ , da  
das System aus den 3 Vektoren  
linear abhängig ist.

$\sum_{i \in I} \lambda_i v_i = \sum_{i \in I} \mu_i v_i$ , sodass  $\lambda_i, \mu_i \in \mathbb{K}$  und nur endlich viele dieser Skalare von Null verschieden sind, und für mindestens ein  $k \in I$  die Bedingung  $\lambda_k \neq \mu_k$  erfüllt ist.

Dann  $0 = v - v = \sum_{i \in I} \lambda_i v_i - \sum_{i \in I} \mu_i v_i = \sum_{i \in I} (\lambda_i - \mu_i) v_i$  mit  $\lambda_k - \mu_k \neq 0$  für  $k$  wie oben.  $\Rightarrow (v_i)_{i \in I}$  sind linear abhängig  $\Rightarrow \not\subseteq$  zu (i).  $\square$

### 3.3.4 Eigenschaften der linearen Abhängigkeit

**Prop.** Sei  $V$  Vektorraum über  $\mathbb{K}$ . Dann gilt:

- (i) Ein einziger Vektor  $v$  ist genau dann linear abhängig, wenn  $v = 0$ .
- (ii) Gehört  $0 \in V$  zu einer Familie von Vektoren, so sind diese linear abhängig.
- (iii) Kommt der gleiche Vektor in einer Familie von Vektoren mehrmals

vor, so sind die Vektoren der Familie linear abhangig.

(iv) Fur  $r \in \mathbb{N}, r \geq 2$  sind Vektoren  $v_1, \dots, v_r \in V$  genau dann linear abhangig, wenn ein  $j \in \{1, \dots, r\}$  existiert mit  $v_j \in \text{lin}(v_i)_{i \in \{1, \dots, r\} \setminus \{j\}}$ .

Beweis. (i), (ii) und (iii) sind klar.

(iv) Sind  $v_1, \dots, v_r$  linear abhangig, so gilt  $0 = \lambda_1 v_1 + \dots + \lambda_r v_r$  mit gewissen  $\lambda_1, \dots, \lambda_r \in \mathbb{K}$  und mit  $\lambda_k \neq 0$  fur ein  $k \in \{1, \dots, r\} \Rightarrow -\lambda_k v_k = \sum_{i \in \{1, \dots, r\} \setminus \{k\}} \lambda_i v_i \Rightarrow v_k = \sum_{i \in \{1, \dots, r\} \setminus \{k\}} -\frac{\lambda_i}{\lambda_k} v_i$ .

Umgekehrt: wenn  $v_k \in \text{lin}(v_i)_{i \in \{1, \dots, r\} \setminus \{k\}}$ , d.h.  $v_k = \sum_{i \in \{1, \dots, r\} \setminus \{k\}} \lambda_i v_i$  mit  $\lambda_i \in \mathbb{K}$  fur alle  $i \in \{1, \dots, r\} \setminus \{k\}$ , so kann man  $\lambda_k := -1$  setzen und erhalt  $0 = \sum_{i=1}^r \lambda_i v_i$  mit  $(\lambda_1, \dots, \lambda_r) \neq (0, \dots, 0)$ . Also sind  $v_1, \dots, v_r$  linear abhangig.  $\square$

*cepshape  
\lin*

zu Prop. 3.3.4 (iv)

Bsp.  $u = \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}, v = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, w = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} \in \mathbb{R}^3$

wir wissen bereits, dass diese drei Vektoren linear abhängig sind.

Wie können wir einen dieser Vektoren als eine Linearkombination der beiden anderen darstellen.

	$\lambda_1$	$\lambda_2$	$\lambda_3$	
$z_1$	1	1	0	0
$z_2$	1	-1	1	0
$z_3$	-2	0	-1	0
$z_1$	1	1	0	0
$z_2$	2	0	1	0
$z_3$	-2	0	-1	0
$z_1$	1	1	0	0
$z_2$	2	0	1	0
$z_3$	0	0	0	0

$$z_2 := z_2 + 2z_1$$

$$z_3 := z_3 + z_2$$

$$\begin{cases} \lambda_1 + \lambda_2 = 0 \\ 2\lambda_1 + \lambda_3 = 0 \end{cases}$$

$$\lambda_1 u + \lambda_2 v + \lambda_3 w = 0 \Leftrightarrow \begin{aligned} \lambda_2 &= -\lambda_1 \\ \lambda_3 &= -2\lambda_1 \end{aligned}$$

wir setzen  $\lambda_1 = -1$  und erhalten

$$(-1) \cdot u + 1 \cdot v + 2 \cdot w = 0$$

$$\Rightarrow u = 1 \cdot v + 2 \cdot w$$

Wir haben  $u$  als Linearkombination  
 $v$  und  $w$  dargestellt.

Kontrollk:  $\begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix} = 1 \cdot \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} + 2 \cdot \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$

Bsp.

$v_1, v_2, v_3, v_4, v_5$

$$\lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3 + \lambda_4 v_4 + \lambda_5 v_5 = 0$$

↓ Gauß-Jordan

$$\lambda_2 = \lambda_1 + \lambda_3 - 2\lambda_5$$

$$\lambda_4 = 2\lambda_1 - \lambda_3 + \lambda_5$$

Welche Vektoren sind bei der Erzeugung  
der Linearkombination redundant?

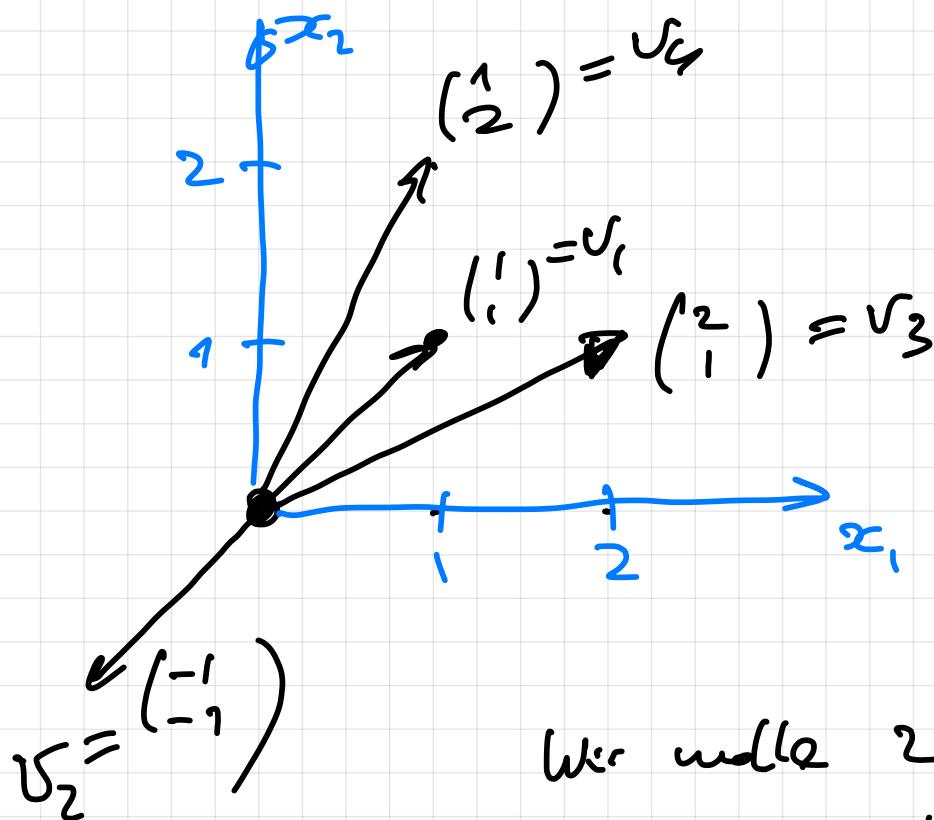
Antwort:  $v_1, v_3, v_5$  sind redundant

$$\left. \begin{array}{l} \lambda_1 = -1 \\ \lambda_3 = 0 \\ \lambda_5 = 0 \end{array} \right] \Rightarrow \begin{array}{l} \lambda_2 = -1 \\ \lambda_4 = -2 \end{array} \Rightarrow \begin{array}{l} -v_1 - v_2 - 2v_5 = 0 \\ v_1 = -v_2 - 2v_4 \end{array}$$

$$\left. \begin{array}{l} \lambda_1 = 0 \\ \lambda_3 = -1 \\ \lambda_5 = 0 \end{array} \right\} \Rightarrow \begin{array}{l} \lambda_2 = -1 \\ \lambda_4 = 1 \end{array} \Rightarrow \begin{array}{l} -v_3 - v_2 + v_4 = 0 \\ \Rightarrow v_3 = -v_2 + v_4 \end{array}$$

$$\left. \begin{array}{l} \lambda_1 = 0 \\ \lambda_3 = 0 \\ \lambda_5 = -1 \end{array} \right\} \Rightarrow \begin{array}{l} \lambda_2 = 2 \\ \lambda_4 = -1 \end{array} \Rightarrow \begin{array}{l} -v_5 + 2v_2 - v_4 = 0 \\ \Rightarrow v_5 = 2v_2 - v_4 \end{array}$$

Bsp.



Wir wollen 2 der Vektoren wählen und mit Hilfe der gewählten Vektoren die beiden anderen darstellen.

$$\lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3 + \lambda_4 v_4 = 0$$

	$\lambda_1$	$\lambda_2$	$\lambda_3$	$\lambda_4$	
$x_1$	1	-1	2	1	0
$x_2$	1	-1	1	2	0

Gauß-Jordan

$$z_2 := z_2 - z_1$$

	$\lambda_1$	$\lambda_2$	$\lambda_3$	$\lambda_4$	
$\lambda_1$	(1)	-1	2	1	0
$\lambda_2$	0	0	-1	(1)	0
$\lambda_3$	(1)	-9	3	0	0
$\lambda_4$	0	0	-1	(1)	0

$$\lambda_1 := \lambda_1 - \lambda_2$$

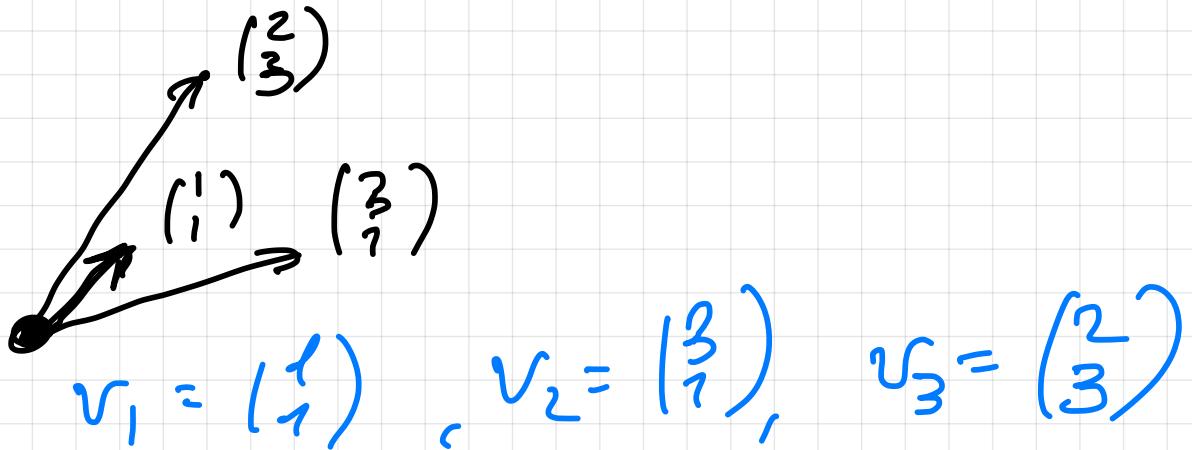
$$\boxed{\begin{aligned}\lambda_1 &= \lambda_2 - 3\lambda_3 \\ \lambda_4 &= \lambda_3\end{aligned}}$$

Das Ergebnis  
von  
Gauß-Jordan.

$$\left. \begin{aligned}\lambda_2 &= -1 \\ \lambda_3 &= 0\end{aligned} \right\} \Rightarrow \left. \begin{aligned}\lambda_1 &= -1 \\ \lambda_4 &= 0\end{aligned} \right\} \Rightarrow \begin{aligned}-v_2 - v_1 &= 0 \\ v_2 &= -v_1\end{aligned}$$

$$\left. \begin{aligned}\lambda_2 &= 0 \\ \lambda_3 &= -1\end{aligned} \right\} \Rightarrow \left. \begin{aligned}\lambda_1 &= 3 \\ \lambda_4 &= -1\end{aligned} \right\} \Rightarrow \begin{aligned}-v_3 + 3v_1 - v_4 &= 0 \\ v_3 &= 3v_1 - v_4\end{aligned}$$

Bsp.



$$\lambda_1 v_1 + \lambda_2 v_2 + \lambda_3 v_3 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

$$\left( \begin{array}{cccc|c} & \lambda_1 & \lambda_2 & \lambda_3 & \\ \bar{z}_1 & 1 & 3 & 2 & 0 \\ \bar{z}_2 & 1 & 1 & 3 & 0 \\ \hline \bar{z}_1 & -2 & 0 & -7 & 0 \\ \bar{z}_2 & 1 & 1 & 3 & 0 \\ \hline \bar{z}_1 & -2 & 0 & -7 & 0 \\ \bar{z}_2 & 0 & 1 & -\frac{1}{2} & 0 \end{array} \right)$$

$\left. \begin{array}{l} -2\lambda_1 - 7\lambda_3 = 0 \\ \lambda_2 - \frac{1}{2}\lambda_3 = 0 \end{array} \right\}$

$$z_1 := z_1 - 3z_2$$

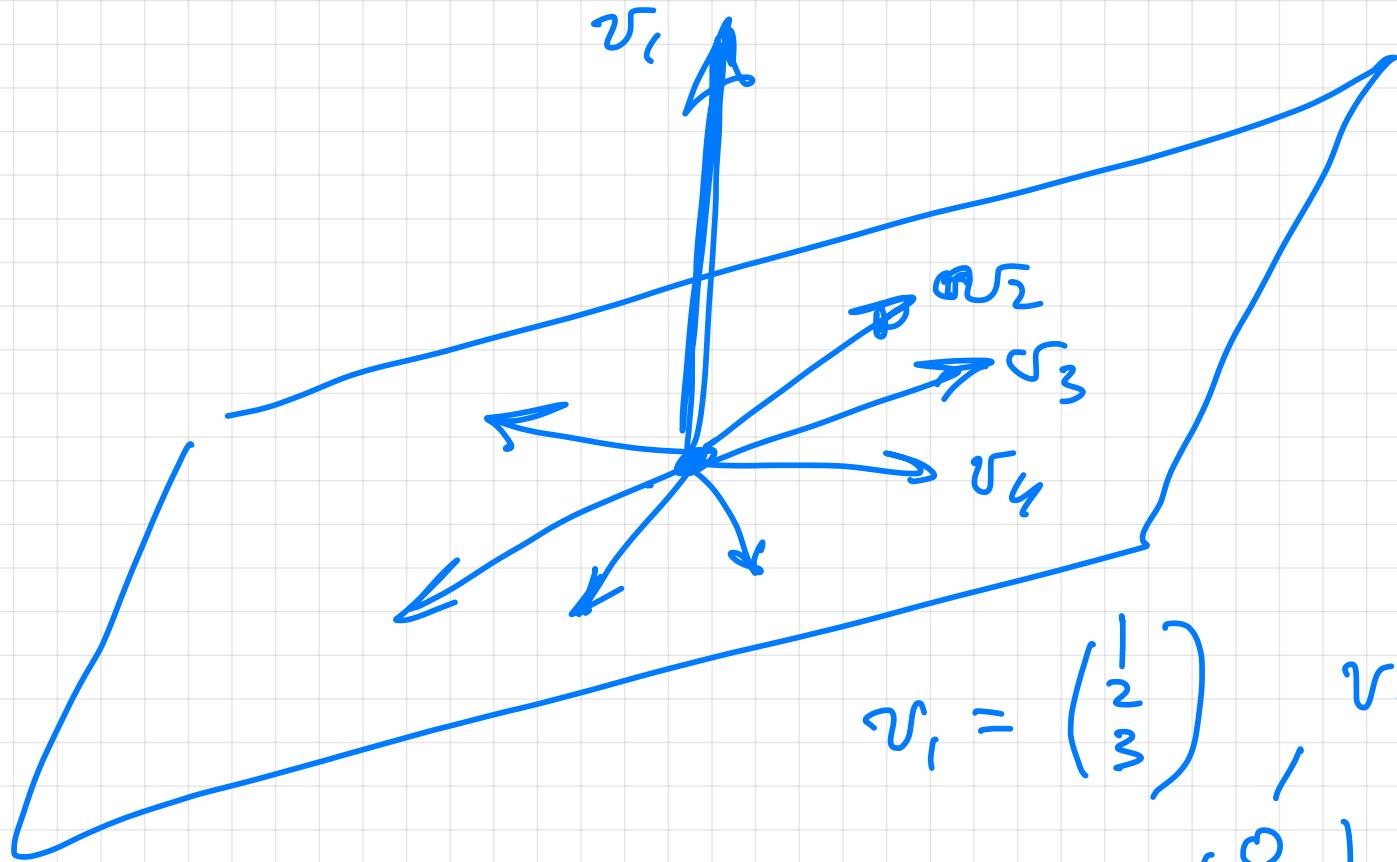
$$z_2 := z_2 + \frac{1}{2}z_1$$

$$\begin{aligned} \lambda_1 &= -\frac{7}{2}\lambda_3 \\ \lambda_2 &= \frac{1}{2}\lambda_3 \end{aligned}$$

$$\lambda_3 = -1 \implies \begin{cases} \lambda_1 = \frac{1}{2} \\ \lambda_2 = -\frac{1}{2} \end{cases} \Rightarrow \frac{1}{2}v_1 - \frac{1}{2}v_2 - v_3 = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$



$$v_3 = \frac{1}{2}v_1 - \frac{1}{2}v_2$$



$$v_1 = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

$$v_2 = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \quad v_3 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix},$$

$$v_4 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \quad v_5 = \begin{pmatrix} 2 \\ -1 \\ -1 \end{pmatrix}$$

## 3.4 Basis und Dimension

### 3.4.1 Erzeugendensysteme und Basen

Sei  $I$  Menge und seien  $v_i$  mit  $i \in I$  Vektoren aus einem VR  $V$  über  $\mathbb{K}$ . Dann heißt die Familie  $(v_i)_{i \in I}$  *Erzeugendensystem* von  $V$ , falls  $\text{lin}(v_i)_{i \in I} = V$ .

Die Familie  $(v_i)_{i \in I}$  heißt *Basis* von  $V$ , falls  $(v_i)_{i \in I}$  ein linear unabhängiges Erzeugendensystem von  $V$  ist. Ist  $(v_i)_{i \in I}$  Basis und  $I$  eine endliche Menge, so heißt  $|I|$  die Länge der Basis  $(v_i)_{i \in I}$ .

Ein VR  $V$  heißt *endlich erzeugt*, falls  $V$  eine endliche Basis besitzt.

### 3.4.2 Charakterisierungen der Basis-Eigenschaft und die Folgerungen daraus

**Thm.** Sei  $V$  Vektorraum über  $\mathbb{K}$ , sei  $n \in \mathbb{N}$  und  $v_1, \dots, v_n \in V$ . Dann sind die folgenden Bedingungen äquivalent:

- (i)  $v_1, \dots, v_n$  bilden eine Basis.

Bsp.

$$\begin{aligned} V &= \{ f \in \mathbb{R}[x] : \deg f \leq 2 \} \\ &= \{ ax^2 + bx + c : a, b, c \in \mathbb{R} \} \end{aligned}$$

Das ist ein VR über  $\mathbb{R}$ .

$x^2, x, 1$  ist eine Basis dieses Raums.

Erzählt: jedes  $f \in V$  ist in  $\text{lin}(1, x, x^2)$ .

Linear unabhängig:

$$a \cdot x^2 + b \cdot x + c \cdot 1 = 0 \quad (a, b, c \in \mathbb{R})$$

$$\implies a = b = c = 0.$$

$x^2 - 1, \quad x + 1, \quad x - 1$  . Sind diese  
Vektoren eine Basis von  $V$ ?

Linear unabhängigkeit:

$$\lambda_1 (x^2 - 1) + \lambda_2 (x + 1) + \lambda_3 (x - 1) = 0$$

$(\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}).$

Wir lösen diese Gleichung in  $\lambda_1, \lambda_2, \lambda_3$ .

$$\lambda_1 \cdot x^2 + (\lambda_2 + \lambda_3) \cdot x + (-\lambda_1 + \lambda_2 - \lambda_3) \cdot 1 = 0$$

$$\Downarrow \left\{ \begin{array}{l} \lambda_1 = 0 \\ \lambda_2 + \lambda_3 = 0 \\ -\lambda_1 + \lambda_2 - \lambda_3 = 0 \end{array} \right. \Leftrightarrow \left\{ \begin{array}{l} \lambda_1 = 0 \\ \lambda_2 + \lambda_3 = 0 \\ \lambda_2 - \lambda_3 = 0 \end{array} \right. \Rightarrow \lambda_1 = \lambda_2 = \lambda_3 = 0$$

$x^2 - 1, x + 1, x - 1$  ist sogar eine Basis!

Warum?

(ii)  $\text{lin}(v_1, \dots, v_n) = V$  und  $V \neq \text{lin}(v_i)_{i \in \{1, \dots, n\} \setminus \{r\}}$  für alle  $r \in \{1, \dots, n\}$ .

(iii) Zu jedem  $v \in V$  gibt es eindeutige  $\lambda_1, \dots, \lambda_n \in \mathbb{K}$  mit  $v = \lambda_1 v_1 + \dots + \lambda_n v_n$ .

(iv)  $v_1, \dots, v_n$  sind linear unabhängig und für jedes  $v \in V$  sind  $v_1, \dots, v_n, v$  linear abhängig.

*Beweis.* (i)  $\Rightarrow$  (ii): Sei (i) erfüllt. Angenommen, (ii) wäre nicht erfüllt. Dann gilt  $v_r := \sum_{i \in \{1, \dots, n\} \setminus \{r\}} \lambda_i v_i$  für ein  $r \in \{1, \dots, n\}$  und gewisse  $\lambda_i \in \mathbb{K}$  mit  $i \in \{1, \dots, n\} \setminus \{r\}$ . Dann gilt  $0 = \sum_{i=1}^n \lambda_i v_i$ , wobei  $\lambda_r := -1$  gesetzt wird. Es gilt  $(\lambda_1, \dots, \lambda_n) \neq (0, \dots, 0)$ . D.h.  $v_1, \dots, v_n$  sind linear abhängig, Widerspruch zu (i).

(ii)  $\Rightarrow$  (iii): Sei (ii) erfüllt. Angenommen, (iii) wäre nicht erfüllt. Dann existiert  $v \in V$  derart, dass  $v = \sum_{i=1}^n \lambda_i v_i = \sum_{i=1}^n \mu_i v_i$  mit  $\lambda_1, \dots, \lambda_n, \mu_1, \dots, \mu_n \in \mathbb{K}$  und mit  $\lambda_r \neq \mu_r$  für mindestens ein  $r \in \{1, \dots, n\}$ . Dann  $0 =$

$$\sum_{i=1}^n (\lambda_i - \mu_i) v_i \Rightarrow 0 = \sum_{i=1}^n \frac{\lambda_i - \mu_i}{\lambda_r - \mu_r} v_i \Rightarrow v_r = \sum_{i \in \{1, \dots, n\} \setminus \{r\}} \frac{\lambda_i - \mu_i}{\lambda_r - \mu_r} v_i \Rightarrow \\ v_r \in \text{lin}(v_i)_{i \in \{1, \dots, n\} \setminus \{r\}} \Rightarrow \text{Widerspruch zu (ii).}$$

(iii)  $\Rightarrow$  (iv): Angenommen, (iii) gilt. Zu zeigen: (iv). Anwendung von (iii) für den Fall  $v = 0$  ergibt die lineare Unabhängigkeit von  $v_1, \dots, v_n$ . Sei  $v \in V$  beliebig. Nach (iii) existieren  $\lambda_1, \dots, \lambda_r \in \mathbb{K}$  mit  $v = \sum_{i=1}^r \lambda_i v_i \Rightarrow \lambda_1 v_1 + \dots + \lambda_r v_r + (-1)v = 0$ , wobei  $\lambda_i \neq 0$ . D.h.  $v_1, \dots, v_r, v$  sind linear abhängig.

(iv)  $\Rightarrow$  (i): Angenommen, (iv) gilt. Wir zeigen (i). Es muss gezeigt werden, dass  $V = \text{lin}(v_1, \dots, v_n)$  gilt. Sei  $v \in V$  beliebig. Wegen (iv) sind  $v_1, \dots, v_n, v$  linear abhängig. D.h.  $\lambda_1 v_1 + \lambda_n v_n + \lambda v = 0$  für gewisse  $\lambda_1, \dots, \lambda_n, \lambda \in \mathbb{K}$  mit  $\lambda_i \neq 0$ . Es gilt  $\lambda \neq 0$ , denn sonst wären  $v_1, \dots, v_n$  linear abhängig, was der Bedingung (iv) widerspricht. Es folgt  $v = \sum_{i=1}^n \frac{\lambda_i}{\lambda} v_i \in \text{lin}(v_1, \dots, v_n)$ .  $\square$

**Kor.** Sei  $V$  ein Vektorraum über  $\mathbb{K}$ , der nicht endlich erzeugt ist. Dann

besitzt  $V$  ein unendliches System von linear unabhängigen Vektoren.

*Beweis.* Ein System aus Vektoren  $v_1, \dots, v_n \in V$  mit  $n = 0$  ist linear unabhängig. Für ein beliebiges linear unabhängiges System  $v_1, \dots, v_n \in V$  mit  $n \in \mathbb{N}_0$  kann man ein  $v \in V$  finden, sodass die  $n + 1$  Vektoren  $v_1, \dots, v_n, v$  linear unabhängig sind. Denn sonst wäre  $v_1, \dots, v_n$  nach dem vorigen Theorem eine Basis und dadurch der Raum  $V$  endlich erzeugt.

Das iterative Anwenden der vorigen Prozeduren ergibt ein unendliches System von linear unabhängigen Vektoren.  $\square$

### 3.4.3 Basisauswahlsatz

**Thm.** Sei  $V$  ein endlich erzeugter Vektorraum über  $\mathbb{K}$ . Dann besitzt  $V$  eine Basis.

*Beweis.* Sei  $v_1, \dots, v_m$  mit  $m \in \mathbb{N}_0$  ein System von Vektoren aus  $V$ , die  $V$  erzeugen. Wenn man in  $\{v_1, \dots, v_m\}$  ein  $v_i$  findet ( $i \in \{1, \dots, m\}$ ), sodass

$v_i \in \text{lin}(v_j)_{j \in \{1, \dots, m\} \setminus \{i\}}$ , dann ersetzt man  $v_1, \dots, v_m$  durch  $\text{lin}(v_j)_{j \in \{1, \dots, m\} \setminus \{i\}}$ .

Nach dieser Änderung bleibt das System erzeugend.

Man wiederholt den vorigen Schritt solange, bis  $v_i \notin \text{lin}(v_j)_{j \in \{1, \dots, m\} \setminus \{i\}}$  für alle  $i \in \{1, \dots, m\}$  gilt. Nach Theorem 3.4.2 ist  $v_1, \dots, v_m$  mit den vorigen Eigenschaften eine Basis.  $\square$

### 3.4.4 Austauschlemma

**Lem.** Sei  $V$  Vektorraum über  $\mathbb{K}$  und sei  $v_1, \dots, v_r$  mit  $r \in \mathbb{N}$  Basis von  $V$ .

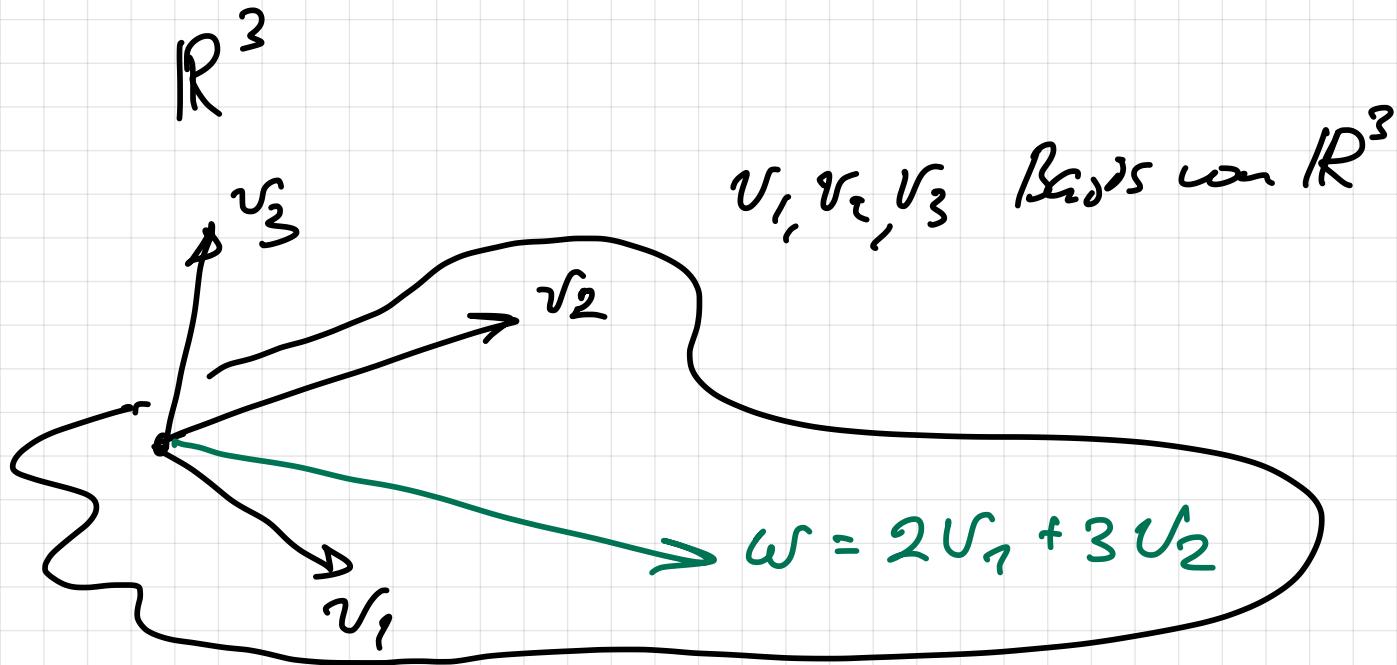
Sei  $w = \lambda_1 v_1 + \dots + \lambda_r v_r$  und  $k \in \{1, \dots, r\}$  ein Index mit  $\lambda_k \neq 0$ . Dann ist  $v_1, \dots, v_{k-1}, w, v_{k+1}, \dots, v_r$  eine Basis von  $V$ .

*Beweis.* Man zeige zunächst, dass  $v_1, \dots, v_{k-1}, w, v_{k+1}, \dots, v_r$  erzeugend sind:

Sei  $v \in V$  beliebig. Da  $v_1, \dots, v_r$  eine Basis ist, gilt  $v = \mu_1 v_1 + \dots + \mu_r v_r$  für gewisse  $\mu_1, \dots, \mu_r \in \mathbb{K}$ . Es gilt:

$$v_k = \frac{1}{\lambda_k}w - \sum_{i \in \{1, \dots, r\} \setminus \{k\}} \frac{\lambda_i}{\lambda_k} v_i$$

Bsp.



$w, v_1, v_3$  Basis von  $\mathbb{R}^3$

$w, v_2, v_3$  Basis von  $\mathbb{R}^3$

$w, v_1, v_2$  keine Basis von  $\mathbb{R}^3$

Bsp.  $v_1, v_2, v_3, v_4$  Basis von  $\mathbb{R}^4$

$$w = \frac{1}{2}v_1 + \frac{2}{3}v_3$$

$w, v_1, v_2, v_3, v_4$  Basis von  $\mathbb{R}^5$

$w, v_1, v_2, v_4$  keine Basis von  $\mathbb{R}^4$

$w, v_1, v_2, v_4$  eine Basis von  $\mathbb{R}^4$

$w, v_1, v_2, v_3$  keine Basis von  $\mathbb{R}^4$

Bsp.  $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}$  Basis von  $\mathbb{R}^3$ .

$\begin{matrix} \parallel \\ a_1 \end{matrix}, \quad \begin{matrix} \parallel \\ a_2 \end{matrix}, \quad \begin{matrix} \parallel \\ a_3 \end{matrix}$

$$v = \begin{pmatrix} 1 \\ 2 \\ -3 \end{pmatrix}$$

Wie kann man  $v$  in  $a_1, a_2, a_3$  einsetzen, sodass  
daraus eine Basis entsteht?

Wir bestimmen eine Darstellung von  $v$  in der Basis  $a_1, a_2, a_3$

$$v = \lambda_1 a_1 + \lambda_2 a_2 + \lambda_3 a_3$$

$$\left| \begin{array}{ccc|c} 1 & \lambda_1 & \lambda_2 & \lambda_3 \\ 1 & 1 & 0 & 1 \\ 1 & -1 & 1 & 2 \\ 1 & 0 & -1 & -3 \end{array} \right|$$

Durch das Einsetzen der Darstellung für  $v_k$  in die Gleichung für  $v$  erhält man:

$$\begin{aligned} v &= \left( \mu_1 - \mu_k \frac{\lambda_1}{\lambda_k} \right) v_1 + \dots + \left( \mu_{k-1} - \mu_k \frac{\lambda_{k-1}}{\lambda_k} \right) v_{k-1} + \frac{\mu_k}{\lambda_k} w \\ &\quad + \left( \mu_{k+1} - \mu_k \frac{\lambda_{k+1}}{\lambda_k} \right) v_{k+1} + \dots + \left( \mu_r - \mu_k \frac{\lambda_r}{\lambda_k} \right) v_r \end{aligned}$$

Das zeigt, dass  $v_1, \dots, v_{k-1}, w, v_{k+1}, \dots, v_r$  ein Erzeugendensystem ist.

Es bleibt zu zeigen, dass  $v_1, \dots, v_{k-1}, w, v_{k+1}, \dots, v_r$  linear unabhängig sind: Seien  $\mu_1, \dots, \mu_{k-1}, \mu, \mu_{k+1}, \dots, \mu_r \in \mathbb{K}$  Skalare mit

$$\mu w + \sum_{i \in \{1, \dots, r\} \setminus \{k\}} \mu_i v_i = 0.$$

Das ergibt

$$\mu \lambda_k v_k + \sum_{i \in \{1, \dots, r\} \setminus \{k\}} (\mu_i + \mu \lambda_i) v_i = 0$$

Weil  $v_1, \dots, v_r$  eine Basis bilden, gilt:

$$\begin{aligned}\mu\lambda_k &= 0 \\ \mu_i + \mu\lambda_i &= 0 \quad \forall i \in \{1, \dots, r\} \setminus \{k\}\end{aligned}$$

Aus  $\mu\lambda_k = 0$  und  $\lambda_k \neq 0$  folgt  $\mu = 0$ . Dann gilt  $\mu_i = \mu\lambda_i = 0$  für alle  $i \in \{1, \dots, r\} \setminus \{k\}$ . Das heißt:  $\mu$  und alle  $\mu_i$  sind 0. Das zeigt die lineare Unabhängigkeit von  $v_1, \dots, v_{k-1}, w, v_{k+1}, \dots, v_r$ .  $\square$

### 3.4.5 Austauschsatz

**Thm.** Sei  $V$  Vektorraum über  $\mathbb{K}$ . Sei  $v_1, \dots, v_r$  eine Basis von  $V$  mit  $r \in \mathbb{N}$ . Seien  $w_1, \dots, w_n \in V$  mit  $n \in \mathbb{N}_0$  linear unabhängig. Dann gilt:

- (a)  $n \leq r$
- (b) Es existieren  $i_1, \dots, i_n \in \{1, \dots, r\}$ ,  $i_1 < \dots < i_n$ , sodass man nach dem Austausch  $v_{i_j}$  gegen  $w_j$  für alle  $j = 1, \dots, n$  aus  $v_1, \dots, v_r$  wieder

eine Basis von  $V$  erhält.

*Beweis.* Induktion über  $n$ :

*Induktionsanfang:*

$n = 0$  klar.

*Induktionsvoraussetzung:*

Sei  $n \in \mathbb{N}$  und seien (a) und (b) mit  $n - 1$  an der Stelle von  $n$  erfüllt.

*Induktionsschritt:* Wir betrachten linear unabhängige Vektoren  $w_1, \dots, w_n \in V$ . Dann sind  $w_1, \dots, w_{n-1}$  linear unabhängig und die Induktionsannahme (Teil (a)) ergibt, dass  $n - 1 \leq r$  gilt. Die Induktionsannahme (Teil (b)) ergibt, dass nach einer geeigneten Umnummerierung der Vektoren  $v_1, \dots, v_r$  das System  $w_1, \dots, w_{n-1}, v_n, \dots, v_r$  eine Basis ist.

Die Gleichheit  $n - 1 = r$  gilt nicht. Wenn  $n - 1 = r$  gilt, dann ist  $w_1, \dots, w_{n-1}, v_n, \dots, v_r$  gleich  $w_1, \dots, w_{n-1}$ . D.h.  $w_1, \dots, w_{n-1}$  ist eine Basis.  $w_1, \dots, w_n$  sind linear unabhängig. Das widerspricht Theorem 3.4.2.(i) $\Rightarrow$ (iv).

Aus  $n - 1 \leq r$  und  $n - 1 \neq r$  folgt  $n \leq r$ .

Da  $w_1, \dots, w_{n-1}, v_n, \dots, v_r$  eine Basis ist, gilt  $w_n = \lambda_1 w_1 + \dots + \lambda_{n-1} w_{n-1} + \lambda_n v_n + \dots + \lambda_r v_r$  für gewisse  $\lambda_1, \dots, \lambda_r \in \mathbb{K}$ .

Wenn  $\lambda_n = \dots = \lambda_r = 0$  gelten würde, dann wäre  $w_n \in \text{lin}(w_1, \dots, w_{n-1})$ , was der linearen Unabhängigkeit von  $w_1, \dots, w_n$  widerspricht. D.h.  $\lambda_k \neq 0$  für ein  $k \in \{n, \dots, r\}$ .

Nach einer wiederholten Umnummerierung kann man annehmen, dass  $k = n$  ist, d.h.  $\lambda_n \neq 0$ . Nach dem Basisaustauschlemma kann man in der Basis  $w_1, \dots, w_{n-1}, v_n, \dots, v_r$  den Vektor  $v_n$  gegen  $w_n$  austauschen. Es folgt:  $w_1, \dots, w_n, v_{n+1}, \dots, v_r$  bilden eine Basis.  $\square$

**Bem.** Der Beweis des vorigen Theorems ist konstruktiv. Das heißt, der Beweis führt zu einem Rechenverfahren.

### 3.4.6 Dimension

**Kor.** Sei  $V$  ein endlich erzeugter Vektorraum über  $\mathbb{K}$ . Dann existiert ein  $n \in \mathbb{N}_0$ , sodass jede Basis von  $V$  aus genau  $n$  Vektoren besteht.

*Beweis.* Nach dem Basisauswahlsatz besitzt  $V$  eine Basis  $v_1, \dots, v_n \in V$  mit  $n \in \mathbb{N}_0$ . Der Vektorraum  $V$  enthält keine unendlichen linear unabhängigen Systeme. Denn hätte  $V$  ein unendliches linear unabhängiges System, dann könnte man in diesem System  $r$  linear unabhängige Vektoren  $w_1, \dots, w_r$  wählen mit  $r \in \mathbb{N}, r > n$ . Das widerspricht aber dem Basisaustauschsatz, aus dem die Ungleichung  $r \leq n$  folgt.

Es folgt, dass jedes linear unabhängige System (unter anderem jede Basis) aus endlich vielen Vektoren besteht. Sei  $w_1, \dots, w_l$  eine beliebige Basis von  $V$  mit  $l \in \mathbb{N}_0$ . Das doppelte Anwenden des Basisaustauschsatzes (Teil (a)) zu den Systemen  $v_1, \dots, v_n$  und  $w_1, \dots, w_l$  ergibt  $n \leq l$  und  $l \leq n$ . Also gilt  $l = n$ .  $\square$

Sei  $V$  Vektorraum über  $\mathbb{K}$ . Dann heißt

$$\dim_{\mathbb{K}}(V) = \begin{cases} \infty & \text{falls } V \text{ nicht endlich erzeugt ist} \\ n \in \mathbb{N}_0 & \text{falls } V \text{ eine Basis aus } n \text{ Vektoren besitzt} \end{cases} \quad (3.4.1)$$

**Bem.** Sei  $n \in \mathbb{N}$ . Dann gilt:

$$\dim(\mathbb{K}^n) = n \quad \text{z.B. Stelle} \quad (3.4.2)$$

Die Vektoren  $e_1, \dots, e_n \in \mathbb{K}^n$  mit  $e_i := (0, \dots, 0, 1, 0, \dots, 0)$  (1 an Stelle  $i$ ) mit  $i \in \{1, \dots, n\}$  bilden eine Basis von  $\mathbb{K}^n$ . Die Basiseigenschaften sind direkt verifizierbar.  $e_1, \dots, e_n$  heißt die *Standardbasis* von  $\mathbb{K}^n$ .

### 3.4.7 Dimension von Untervektorräumen

**Kor.** Sei  $V$  Vektorraum über  $\mathbb{K}$ , der endlich erzeugt ist. Sei  $W$  Untervektorraum von  $V$ . Dann gilt:

- (i)  $\dim(W) \leq \dim(V)$

Bsp.

Diagram illustrating a basis in  $\mathbb{R}^2$ . A coordinate system shows two vectors originating from the origin:  $e_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$  (red arrow) and  $e_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$  (blue arrow). A third vector  $\begin{pmatrix} 3 \\ 2 \end{pmatrix}$  (black arrow) is shown, which can be expressed as a linear combination of  $e_1$  and  $e_2$ :

$$\begin{pmatrix} 3 \\ 2 \end{pmatrix} = x \cdot \begin{pmatrix} 1 \\ 0 \end{pmatrix} + y \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix}$$

$$= x \cdot e_1 + y \cdot e_2$$

$$\begin{pmatrix} 3 \\ 2 \end{pmatrix} = 3 \cdot e_1 + 2 \cdot e_2$$

Diagram illustrating a basis in a 2D plane. Two vectors  $a_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$  (green arrow) and  $a_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$  (green arrow) are shown as a basis. A vector  $\begin{pmatrix} x \\ y \end{pmatrix}$  (black arrow) is shown, which can be expressed as a linear combination of  $a_1$  and  $a_2$ :

$a_1, a_2$  auch eine Basis von  $\mathbb{R}^2$

$$\begin{pmatrix} x \\ y \end{pmatrix} = x' \cdot a_1 + y' a_2$$

Was ist  $x', y'$  in Abhängigkeit von  $x, y$ ?

$$\Downarrow \quad \left\{ \begin{array}{l} x = x' + y' \quad (\underline{\text{I}}) \\ y = x' - y' \quad (\underline{\text{II}}) \end{array} \right.$$

$$\Updownarrow \quad \left\{ \begin{array}{l} x' = \frac{x+y}{2} \\ y' = \frac{x-y}{2} \end{array} \right. \quad \begin{array}{l} (\underline{\text{I}}) + (\underline{\text{II}}) \\ \hline 2 \\ (\underline{\text{I}}) - (\underline{\text{II}}) \\ \hline 2 \end{array}$$

$$\begin{pmatrix} 3 \\ 2 \end{pmatrix} = \frac{5}{2} a_1 + \frac{1}{2} a_2$$

$$\begin{pmatrix} 30 \\ 28 \\ 31 \\ 31 \\ 33 \end{pmatrix}$$

$$\begin{matrix} 30 \\ \text{dann } -2 \\ \text{dann } +3 \\ \text{dann } +0 \\ \text{dann } +2 \end{matrix} \quad \left. \right\}$$

Zu welcher Basis gehören diese Werte?

$$30 \cdot e_1 + 28 \cdot e_2 + 31 e_3 + 31 \cdot e_4 + 33 e_5$$

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{pmatrix} =$$

$$\begin{aligned}
 & x_1 \cdot a_1 + \\
 & (x_2 - x_1) \cdot a_2 + \\
 & (x_3 - x_2) \cdot a_3 + \\
 & (x_4 - x_3) \cdot a_4 + \\
 & (x_5 - x_4) \cdot a_5
 \end{aligned}$$

$$\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = a_1$$

$$\begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = a_2$$

$$\begin{pmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 1 \end{pmatrix} = a_3$$

$$\begin{pmatrix} 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{pmatrix} = a_4$$

$$\begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{pmatrix} = a_5$$

$$(ii) \dim(W) = \dim(V) \Leftrightarrow W = V$$

*Beweis.* (i)  $W$  ist endlich erzeugt. Denn sonst hätte  $W$  ein unendliches linear unabhängiges System. Das widerspricht aber dem Basisaustauschsatz (Teil (a)).

Also hat  $W$  eine endliche Basis. Die Anzahl der Vektoren in dieser Basis ist nach dem Basisaustauschsatz (Teil (a)) nicht größer als die Anzahl der Vektoren in den Basen von  $V$ . Das zeigt  $\dim(W) \leq \dim(V)$ .

(ii) Für  $W = V$  ist trivialerweise  $\dim(W) = \dim(V)$ . Umgekehrt, sei  $\dim(W) = \dim(V)$ . Sei  $v_1, \dots, v_n$  ( $n \in \mathbb{N}_0$ ) eine Basis von  $W$ . Wenn  $W \neq V$  wäre, dann wäre  $v_1, \dots, v_n, v$  mit einem beliebigen  $v \in V \setminus W$  linear unabhängig. Das widerspricht dem Basisaustauschsatz (Teil (a)).  $\square$

**Bsp.**

- $V = \mathbb{R}^3, \mathbb{K} = \mathbb{R}, \dim(\mathbb{R}^3) = 3$ . Untervektorräume von  $\mathbb{R}^3$  haben Dimension 0, 1, 2 oder 3.  $\{0\}$  der einzige UVR der Dimension 0.  $\mathbb{R}^3$  der einzige UVR der Dimension 3. (triviale Fälle)

$\text{lin}(a)$  mit  $a \in \mathbb{R}^3 \setminus \{0\}$  sind Geraden durch den Nullpunkt: die UVRs der Dimension 1.  $\text{lin}(a, b)$  mit linear unabhängigen  $a$  und  $b$  aus  $\mathbb{R}^3$  sind die Ebenen durch den Nullpunkt die UVRs der Dimension 2.

- Sei  $x$  eine Unbestimmte. Dann ist  $\mathbb{K}[x]$  Vektorraum über  $\mathbb{K}$ .  $\dim \mathbb{K}[x] = \infty$ , denn  $1, x^2, x^3, \dots$  sind unendlich viele linear unabhängige Vektoren aus  $\mathbb{K}[x]$ .
- $\dim_{\mathbb{Q}}(\mathbb{R}) = \infty$

### 3.4.8 Ergänzung zu einer Basis

**Thm** (Basisergänzungssatz). *Sei  $V$  ein endlichdimensionaler Vektorraum über  $\mathbb{K}$ . Seien  $v_1, \dots, v_n \in V$  ( $n \in \mathbb{N}_0$ ) linear unabhängig. Dann existieren*

$v_{n+1}, \dots, v_r \in V$  mit  $r = \dim(V)$  derart, dass  $v_1, \dots, v_r$  eine Basis ist.

*Beweis.* Sei  $w_1, \dots, w_r$  eine beliebige Basis von  $V$ . Durch die Verwendung des Basisaustauschsatzes für  $v_1, \dots, v_n$  und  $w_1, \dots, w_r$  erhält man eine Basis  $v_1, \dots, v_r$  mit Vektoren  $v_{n+1}, \dots, v_r$  aus  $\{w_1, \dots, w_r\}$ .  $\square$

## 3.5 Rang

### 3.5.1 Matrizen und ihr Rang

Seien  $m, n \in \mathbb{N}$ . Wir setzen  $\mathbb{K}^{m \times n} := \mathbb{K}^{\{1, \dots, m\} \times \{1, \dots, n\}}$ . Die Elemente  $A$  von  $\mathbb{K}^{m \times n}$  heißen  $m \times n$  Matrizen bzw. *Matrizen* der Größe  $m \times n$ . [Wie Vektoren aus  $\mathbb{K}^l$ , die Komponenten haben aber zwei Indizes.]

Schreibweise:  $A = (a_{ij})_{i=1, j=1}^{m, n}$  bedeutet  $A$  ist Matrix mit der Komponente  $a_{ij}$  in der Position  $(i, j)$  für alle  $i \in \{1, \dots, m\}, j \in \{1, \dots, n\}$ . Oder

alternativ:

$$A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ a_{21} & \cdots & a_{2n} \\ \vdots & & \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \quad (3.5.1)$$

$i$  heißt der *Zeilenindex* und  $j$  heißt der *Spaltenindex* der Komponente  $a_{ij}$ .

$A$  kann durch durch Folge ihrer Spalten bzw. Zeilen definiert werden:

$$A = \begin{pmatrix} | & & | \\ s_1 & \cdots & s_n \\ | & & | \end{pmatrix} = \begin{pmatrix} - & z_1 & - \\ & \vdots & \\ - & z_m & - \end{pmatrix} \quad (3.5.2)$$

mit  $s_1, \dots, s_n \in \mathbb{K}^{m \times 1}$  und  $z_1, \dots, z_m \in \mathbb{K}^{1 \times n}$ .  $z_i$  heißt die  $i$ -te Zeile von  $A$ ,  $s_j$  heißt die  $j$ -te Spalte von  $A$ .

$\mathbb{K}^r$  mit  $r \in \mathbb{N}$  wird im Folgenden oft als  $\mathbb{K}^{r \times 1}$  oder  $\mathbb{K}^{1 \times r}$  interpretiert. [In der Literatur wird die Interpretierung als  $\mathbb{K}^{r \times 1}$  meistens vorgezogen.]

Der Wert  $\text{rang}(A) = \dim(\text{lin}(s_1, \dots, s_n))$  heißt der *Rang* von  $A$ .

### 3.5.2 Elementartransformationen von Vektorsystemen

Analog zu den Elementartransformationen von linearen Gleichungssystemen führen wir Elementartransformationen von endlichen Vektorsystemen  $v_1, \dots, v_n$  ( $n \in \mathbb{N}_0$ ) eines Vektorraums  $V$  über  $\mathbb{K}$  ein:

**Typ 1.** Für  $i, j \in \{1, \dots, n\}$  werden  $v_i$  und  $v_j$  vertauscht.

**Typ 2.** Für  $i \in \{1, \dots, n\}$  und  $\alpha \in \mathbb{K} \setminus \{0\}$  wird  $v_i$  durch  $\alpha v_i$  ersetzt.

**Typ 3.** Für  $i, j \in \{1, \dots, n\}, i \neq j$  und  $\alpha \in \mathbb{K}$  wird  $v_i$  durch  $v_i + \alpha v_j$  ersetzt.

**Lem.** *Elementartransformationen eines (endlichen) Vektorsystems aus einem Vektorraum  $V$  über  $\mathbb{K}$  ändern die lineare Hülle dieses Systems nicht.*

*Beweis.* Für Typ 1 ist das klar.

Typ 2: O.B.d.A. nehmen wir an, dass  $v_1$  und  $\alpha v_1$  (mit  $\alpha \in \mathbb{K} \setminus \{0\}$ ) ersetzt

wird. Zu zeigen:  $\text{lin}(v_1, \dots, v_n) = \text{lin}(\alpha v_1, v_2, \dots, v_n)$ .

$$\begin{aligned}\text{lin}(\alpha v_1, v_2, \dots, v_n) &= \{\lambda_1(\alpha v_1) + \lambda_2 v_2 + \dots + \lambda_n v_n : \lambda_1, \dots, \lambda_n \in \mathbb{K}\} \\ &= \{(\lambda_1 \alpha) v_1 + \lambda_2 v_2 + \dots + \lambda_n v_n : \lambda_1, \dots, \lambda_n \in \mathbb{K}\} \\ &\subseteq \text{lin}(v_1, \dots, v_n).\end{aligned}$$

$$\begin{aligned}\text{Umgekehrt: } \text{lin}(v_1, \dots, v_n) &= \{\lambda_1 v_1 + \dots + \lambda_n v_n : \lambda_1, \dots, \lambda_n \in \mathbb{K}\} \\ &= \left\{ \frac{\lambda_1}{\alpha} (\alpha v_1) + \lambda_2 v_2 + \dots + \lambda_n v_n : \lambda_1, \dots, \lambda_n \in \mathbb{K} \right\} \\ &\subseteq \text{lin}(\alpha v_1, v_2, \dots, v_n).\end{aligned}$$

Typ 3: Der Einfachheit halber betrachten wir ein System aus 2 Vektoren, d.h.  $v_1, v_2$ .  $v_2$  wird durch  $v_2 + \alpha v_1$  ersetzt mit  $\alpha \in \mathbb{K}$ . Zu zeigen:  $\text{lin}(v_1, v_2) =$

$$\text{lin}(v_1, v_2 + \alpha v_1).$$

$$\begin{aligned}\text{lin}(v_1, v_2 + \alpha v_1) &= \{\lambda_1 v_1 + \lambda_2(v_2 + \alpha v_1) : \lambda_1, \lambda_2 \in \mathbb{K}\} \\ &= \{(\lambda_1 + \lambda_2\alpha)v_1 + \lambda_2 v_2 : \lambda_1, \lambda_2 \in \mathbb{K}\} \\ &\subseteq \text{lin}(v_1, v_2) \\ \text{lin}(v_1, v_2) &= \{\lambda_1 v_1 + \lambda_2 v_2 : \lambda_1, \lambda_2 \in \mathbb{K}\} \\ &= \{\lambda_1 v_1 + \lambda_2(v_2 + \alpha v_1 - \alpha v_1) : \lambda_1, \lambda_2 \in \mathbb{K}\} \\ &= \{(\lambda_1 - \lambda_2\alpha)v_1 + \lambda_2(v_2 + \alpha v_1) : \lambda_1, \lambda_2 \in \mathbb{K}\} \\ &\subseteq \text{lin}(v_1, v_2 + \alpha v_1).\end{aligned}$$

□

### 3.5.3 Der Rang der transponierten Matrix

Seien  $m, n \in \mathbb{N}$  und sei  $A = (a_{ij})_{i=1,j=1}^{m,n} \in \mathbb{K}^{m \times n}$ . Dann heißt  $A^\top = (a_{ij}^\top)_{i=1,j=1}^{m,n} \in \mathbb{K}^{n \times m}$  mit  $a_{ij}^\top = a_{ji}$  für alle  $i, j$ . D.h., wenn

$$A = \begin{pmatrix} | & & | \\ s_1 & \cdots & s_n \\ | & & | \end{pmatrix} = \begin{pmatrix} - & z_1 & - \\ \vdots & & \vdots \\ - & z_m & - \end{pmatrix},$$

dann ist

$$A^\top = \begin{pmatrix} - & s_1 & - \\ \vdots & & \vdots \\ - & s_n & - \end{pmatrix} = \begin{pmatrix} | & & | \\ z_1 & \cdots & z_m \\ | & & | \end{pmatrix}.$$

**Thm.** Seien  $m, n \in \mathbb{N}$  und sei  $A \in \mathbb{K}^{m \times n}$ . Dann gilt  $\text{rang}(A) = \text{rang}(A^\top)$ .

*Beweis.* Sei  $A = (s_1, \dots, s_n) = (z_1, \dots, z_m)^\top$ . Man betrachte die Gleichung  $\lambda_1 s_1 + \dots + \lambda_n s_n = 0$  mit Unbekannten  $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ . Diese Vektorglei-

chung ist nichts anderes als das lineare Gleichungssystem

$$\sum_{j=1}^n a_{ij} \lambda_j = 0 \quad \forall i \in \{1, \dots, m\}.$$

Wir überführen dieses System in eine reduzierte Stufenform. O.B.d.A. seien  $(1, 1), \dots, (r, r)$  die Positionen der Pivotelemente und seien alle Pivotelemente gleich 1:

$\lambda_1 \dots \lambda_n$							$\lambda_1 \dots \lambda_r \lambda_{r+1} \dots \lambda_n$		
$\lambda_1 \dots \lambda_n$							1 * ... *	0	
$a_{11} \dots a_{1n}$	0	Gauß-Jordan						... : : :	: : : :
$\vdots \ddots \vdots$	$\vdots$							1 * ... *	0
$a_{m1} \dots a_{mn}$	0							0 0 ... 0	0
							... : : :	: : : :	
							0 0 ... 0	0	

Sei  $A'$  die Matrix, die der reduzierten Stufenform entspricht und sei

$$A' = \begin{pmatrix} | & & | \\ s'_1 & \cdots & s'_n \\ | & & | \end{pmatrix} = \begin{pmatrix} - & z'_1 & - \\ & \vdots & \\ - & z'_m & - \end{pmatrix}.$$

Dann ist  $\dim(\text{lin}(z'_1, \dots, z'_m)) = \dim(\text{lin}(z_1, \dots, z_m))$ , weil Elementartransformationen der Vektoren eines Systems die lineare Hülle des Systems nicht ändern und  $z'_1, \dots, z'_m$  aus  $z_1, \dots, z_m$  durch die Anwendung von Elementartransformationen entstanden sind.

Die  $r$  Vektoren  $z'_1, \dots, z'_r$  bilden eine Basis von  $\text{lin}(z'_1, \dots, z'_m)$ :

- (i) Erzeugendensystem: Offensichtlich gilt  $\text{lin}(z'_1, \dots, z'_r) = \text{lin}(z'_1, \dots, z'_m)$ , da  $z'_i$  mit  $i > r$  Nullvektoren sind.
- (ii) Lineare Unabhängigkeit: Wenn  $\mu_1, \dots, \mu_r \in \mathbb{K}$  und  $\mu_1 z'_1 + \dots + \mu_r z'_r = 0$ , dann folgt  $\mu_1 = \dots = \mu_r = 0$ , denn  $\mu_i$  ist die  $i$ -te Komponente von  $\mu_1 z'_1 + \dots + \mu_r z'_r \quad \forall i \in \{1, \dots, r\}$ .

Also gilt  $\dim(\text{lin}(z_1, \dots, z_m)) = \dim(\text{lin}(z'_1, \dots, z'_m)) = r$ .

Wir zeigen, dass  $s_1, \dots, s_r$  eine Basis von  $\text{lin}(s_1, \dots, s_n)$  bilden:

- (i) Lineare Unabhängigkeit: Betrachte  $\lambda_1, \dots, \lambda_r \in \mathbb{K}$  mit  $\lambda_1 s_1 + \dots + \lambda_r s_r = 0$ , was dasselbe ist wie  $\lambda_1 s_1 + \dots + \lambda_r s_r + \lambda_{r+1} s_{r+1} + \dots + \lambda_n s_n = 0$  mit  $\lambda_{r+1} = \dots = \lambda_n = 0$ .

Das System  $\lambda_1 s_1 + \dots + \lambda_n s_n = 0$  ist äquivalent zum System in der reduzierten Stufenform. Durch die Verwendung letzterer folgen aus  $\lambda_{r+1} = \dots = \lambda_n = 0$  die Gleichungen  $\lambda_1 = \dots = \lambda_r = 0$ .

- (ii) Erzeugendensystem: Um dies zu zeigen, verifizieren wir, dass  $s_{r+1}, \dots, s_n \in \text{lin}(s_1, \dots, s_r)$  gilt. Dafür fixieren wir  $i \in \{r+1, \dots, n\}$  und möchten  $\lambda_1, \dots, \lambda_r \in \mathbb{K}$  mit  $s_i = \lambda_1 s_1 + \dots + \lambda_r s_r$  bestimmen.

Die Gleichung  $s_i = \lambda_1 s_1 + \dots + \lambda_r s_r$  ist äquivalent zu

$$\lambda_1 s_1 + \dots + \lambda_r s_r + \lambda_{r+1} s_{r+1} + \dots + \lambda_n s_n = 0 \quad (3.5.3)$$

mit  $\lambda_i = -1$  und  $\lambda_j = 0 \quad \forall j \in \{r+1, \dots, n\} \setminus \{i\}$ .

Aus der reduzierten Stufenform ergibt sich eine Wahl von  $\lambda_1, \dots, \lambda_r$ , für die (3.5.3) erfüllt ist.

Also gilt:

$$\begin{aligned} r &= \dim(\text{lin}(s_1, \dots, s_n)) = \text{rang}(A) \\ r &= \dim(\text{lin}(z_1, \dots, z_n)) = \text{rang}(A^\top) \\ \Rightarrow \quad \text{rang}(A) &= \text{rang}(A^\top). \end{aligned}$$

□

**Kor.** Elementartransformationen der Zeilen sowie Spalten einer Matrix mit Komponenten in  $\mathbb{K}$  ändern den Rang nicht.

*Beweis.* Für Spalten folgt die Aussage aus 3.5.2. Elementartransformationen ändern  $\text{rang}(A^\top)$  nicht. Da aber  $\text{rang}(A) = \text{rang}(A^\top)$  gilt, folgt daraus die Behauptung für Zeilen. □

**Bem.** Der Beweis des vorigen Theorems enthält ein Verfahren zur Bestimmung einer Basis  $v_{i_1}, \dots, v_{i_r}$  von  $\text{lin}(v_1, \dots, v_n)$  für Vektorsysteme  $v_1, \dots, v_r$  mit  $v_i \in \mathbb{K}^m$  und  $n, m \in \mathbb{N}$ .

## 3.6 Summen von Vektorräumen

### 3.6.1 Summe von Vektorräumen

Sei  $V$  Vektorraum über  $\mathbb{K}$  und seien  $W_1, \dots, W_k$  Untervektorräume von  $V$  ( $k \in \mathbb{N}$ ). Wir definieren:

$$W_1 + \dots + W_k := \{w_1 + \dots + w_k : w_1 \in W_1, \dots, w_k \in W_k\} \quad (3.6.1)$$

**Bem.** Die Summe  $W_1 + \dots + W_k$  ist Untervektorraum von  $V$  und es gilt:

- (i)  $W_1 + \dots + W_k = \text{lin}(W_1 \cup \dots \cup W_k)$
- (ii)  $\dim(W_1 + \dots + W_k) \leq \dim(W_1) + \dots + \dim(W_k)$

### 3.6.2 Dimensionsformel für Summe von zwei Untervektorräumen

**Thm.** Sei  $V$  Vektorraum über  $\mathbb{K}$  und seien  $W_1, W_2$  endlichdimensionale Untervektorräume von  $V$ . Dann gilt:

$$\dim(W_1 + W_2) = \dim(W_1) + \dim(W_2) - \dim(W_1 \cap W_2) \quad (3.6.2)$$

*Beweis.* Sei  $v_1, \dots, v_m$  eine beliebige Basis von  $W_1 \cap W_2$  (d.h.  $m = \dim(W_1 \cap W_2)$ ). Wir ergänzen  $v_1, \dots, v_m$  zu einer Basis  $v_1, \dots, v_m, w_1, \dots, w_k$  von  $W_1$  (d.h.  $\dim(W_1) = m + k$ ) und zu einer Basis  $v_1, \dots, v_m, w'_1, \dots, w'_l$  von  $W_2$  (d.h.  $\dim(W_2) = m + l$ ).

Es reicht zu zeigen, dass  $\mathcal{B} = (v_1, \dots, v_m, w_1, \dots, w_k, w'_1, \dots, w'_l)$  eine Basis von  $W_1 + W_2$  bilden:

- (i)  $\mathcal{B}$  ist ein Erzeugendensystem für  $W_1 + W_2$ , denn jedes  $a \in W_1 + W_2$  ist  $a = a_1 + a_2$  mit  $a_1 \in W_1 = \text{lin}(v_1, \dots, v_m, w_1, \dots, w_k)$ ,  $a_2 \in W_2 = \text{lin}(v_1, \dots, v_m, w'_1, \dots, w'_l)$ .

(ii) Lineare Unabhängigkeit von  $\mathcal{B}$ :

- (a)  $v_1, \dots, v_m$  ist Basis von  $W_1 \cap W_2$
- (b)  $v_1, \dots, v_m, w_1, \dots, w_k$  ist Basis von  $W_1$
- (c)  $v_1, \dots, v_m, w'_1, \dots, w'_l$  ist Basis von  $W_2$

Seien  $\lambda_1, \dots, \lambda_m, \mu_1, \dots, \mu_k, \mu'_1, \dots, \mu'_l \in \mathbb{K}$ , sodass  $\lambda_1 v_1 + \dots + \lambda_m v_m + \mu_1 w_1 + \dots + \mu_k w_k + \mu'_1 w'_1 + \dots + \mu'_l w'_l = 0$ .

Betrachte  $v := \lambda_1 v_1 + \dots + \lambda_m v_m + \mu_1 w_1 + \dots + \mu_k w_k$ . Einerseits ist  $v \in W_1$ . Andererseits gilt  $v = (-\mu'_1)w'_1 + \dots + (-\mu'_l)w'_l \in W_2$ , also  $v \in W_1 \cap W_2$ .

Dann ist aber auch  $v = \lambda'_1 v_1 + \dots + \lambda'_m v_m$  für gewisse  $\lambda'_1, \dots, \lambda'_m \in \mathbb{K}$ .

Aus der Charakterisierung der Basiseigenschaften folgt  $\lambda_1 = \lambda'_1, \dots, \lambda_m = \lambda'_m$  und  $\mu_1 = \dots = \mu_k = 0$ .

Dies ergibt  $\lambda_1 v_1 + \dots + \lambda_m v_m + \mu'_1 w'_1 + \dots + \mu'_l w'_l = 0$ . Da aber

$v_1, \dots, v_m, w'_1, \dots, w'_l$  eine Basis von  $W_2$  ist, folgt  $\lambda_1 = \dots = \lambda_m = \mu'_1 = \dots = \mu'_l = 0$ .  $\square$

### 3.6.3 Direkte Summe

**Lem.** Sei  $V$  Vektorraum über  $\mathbb{K}$ , seien  $W_1, W_2$  Untervektorräume von  $V$  und sei  $V = W_1 + W_2$ . Dann sind die folgenden Bedingungen äquivalent:

- (i)  $W_1 \cap W_2 = \{0\}$
- (ii) Jedes  $v \in V$  ist eindeutig als  $v = w_1 + w_2$  mit  $w_1 \in W_1$  und  $w_2 \in W_2$  darstellbar.
- (iii)  $w_1$  und  $w_2$  sind linear unabhängig für alle  $w_1 \in W_1 \setminus \{0\}$  und  $w_2 \in W_2 \setminus \{0\}$ .

*Beweis.* (i)  $\Rightarrow$  (ii): Sei (i) erfüllt und seien  $w_1, w'_1 \in W_1, w_2, w'_2 \in W_2$  mit  $v = w_1 + w_2 = w'_1 + w'_2$ . Aus dieser Gleichung folgt  $w_1 - w'_1 = w'_2 - w_2$ .

Es folgt  $w_1 - w'_1 = w'_2 - w_2 \in W_1 \cap W_2 = \{0\}$ . D.h.  $w_1 - w'_1 = w'_2 - w_2 = 0$ . Es folgt  $w_1 = w'_1, w_2 = w'_2$ .

(ii) $\Rightarrow$ (iii): Sei (ii) erfüllt. Seien  $w_1 \in W_1 \setminus \{0\}$  und  $w_2 \in W_2 \setminus \{0\}$ . Wenn  $w_1$  und  $w_2$  linear abhängig wären, dann hätte man  $0 = \lambda_1 w_1 + \lambda_2 w_2$  mit  $\lambda_1, \lambda_2 \in \mathbb{K}$  und  $(\lambda_1, \lambda_2) \neq (0, 0)$ . Also  $0 = 0 + 0 = \lambda_1 w_1 + \lambda_2 w_2$  mit  $\lambda_1 w_1 \neq 0$  oder  $\lambda_2 w_2 \neq 0$ . Das ist  $\nsubseteq$  zu (ii) für den Fall  $v = 0$ .

(iii) $\Rightarrow$ (i): Wir zeigen  $\neg(i) \Rightarrow \neg(iii)$ . Sei  $W_1 \cap W_2 \neq \{0\}$ . Wir fixieren ein  $v \in (W_1 \cap W_2) \setminus \{0\}$ . Für  $w_1 = v \in W_1 \setminus \{0\}$  und  $w_2 = v \in W_2 \setminus \{0\}$  gilt:  $w_1$  und  $w_2$  sind linear abhängig.

□

Ein Vektorraum  $V$  über  $\mathbb{K}$  heißt *direkte Summe* von Untervektorräumen  $W_1$  und  $W_2$  von  $V$ , wenn  $V = W_1 + W_2$  und  $W_1 \cap W_2 = \{0\}$  gilt. Bezeichnung:  $V = W_1 \oplus W_2$ .

### 3.6.4 Charakterisierung der direkten Summe

**Thm.** Sei  $V$  endlichdimensionaler Vektorraum über  $\mathbb{K}$  mit Untervektorräumen  $W_1$  und  $W_2$ . Dann sind die folgenden Bedingungen äquivalent:

- (i)  $V = W_1 \oplus W_2$
- (ii) Es gibt Basen  $w_1, \dots, w_k$  von  $W_1$  und  $w'_1, \dots, w'_l$  von  $W_2$  ( $k, l \in \mathbb{N}_0$ ), sodass  $w_1, \dots, w_k, w'_1, \dots, w'_l$  eine Basis von  $V$  ist.
- (iii)  $V = W_1 + W_2$  und  $\dim(V) = \dim(W_1) + \dim(W_2)$ .

*Beweis.* (i) $\Rightarrow$ (ii):  $V = W_1 \oplus W_2$  bedeutet, dass  $V = W_1 + W_2$  und  $W_1 \cap W_2 = \{0\}$ . Der Beweis der Dimensionsformel für die Summe von zwei Vektorräumen für den Fall  $\dim(W_1 \cap W_2) = 0$  ergibt, dass  $w_1, \dots, w_k, w'_1, \dots, w'_l$  aus (ii) eine Basis von  $V$  ist.

(ii) $\Rightarrow$ (iii): klar.

(iii) $\Rightarrow$ (i): Sei (iii) erfüllt. Aus der Dimensionsformel folgt  $\dim(W_1 \cap W_2) = 0$ . Das heißt  $W_1 \cap W_2 = \{0\}$ .

□

### 3.6.5 Direkter Summand

**Kor.** Sei  $V$  Vektorraum über  $\mathbb{K}$  mit  $\dim(V) < \infty$ . Sei  $W$  Untervektorraum von  $V$ . Dann existiert ein Untervektorraum  $W'$  von  $V$  mit  $V = W \oplus W'$ .

*Beweis.* Sei  $v_1, \dots, v_m$  Basis von  $W$  ( $m \in \mathbb{N}_0$ ). Wir erweitern diese Basis zu einer Basis  $v_1, \dots, v_n$  von  $V$  ( $n \geq m, n \in \mathbb{N}_0$ ). Die Behauptung gilt für  $W' := \text{lin}(v_{m+1}, \dots, v_n)$ . □

### 3.6.6 Direkte Summe endlich vieler Vektorräume

Sei  $V$  Vektorraum über  $\mathbb{K}$ , sei  $W_1, \dots, W_k$  ( $k \in \mathbb{N}$ ) Untervektorräume von  $V$ . Dann heißt  $V$  direkte Summe von  $W_1, \dots, W_k$  (Bezeichnung:  $V = W_1 \oplus$

$\dots \oplus W_k$ ), wenn die folgenden Bedingungen erfüllt sind:

$$(\text{DS1}) \quad V = W_1 + \dots + W_k$$

$$(\text{DS2}) \quad \begin{aligned} &\text{Wenn } w_1 \in W_1, \dots, w_k \in W_k \text{ und } w_1 + \dots + w_k = 0, \\ &\text{dann folgt } w_1 = \dots = w_k = 0. \end{aligned}$$

**Bsp.**  $\mathbb{R}^2$  ist keine direkte Summe von  $\text{lin}(e_1), \text{lin}(e_2), \text{lin}(e_1 + e_2)$ .

**Bsp.** Wenn  $v_1, \dots, v_n$  ( $n \in \mathbb{N}$ ) eine Basis von  $V$  ist, dann gilt:  $V = \text{lin}(v_1) \oplus \dots \oplus \text{lin}(v_n)$ .

**Thm.** Für Untervektorräume  $W_1, \dots, W_k$  ( $k \in \mathbb{N}$ ) eines endlichdimensionalen Vektorraums  $V$  über  $\mathbb{K}$  sind folgende Bedingungen äquivalent:

$$(i) \quad V = W_1 \oplus \dots \oplus W_k$$

$$(ii) \quad \text{Ist für jedes } i \in \{1, \dots, k\} \text{ eine Basis } (v_{i,1}, \dots, v_{i,r_i}) \text{ mit } r_i \in \mathbb{N}_0 \text{ von}$$

$W_i$  gegeben, so ist  $\mathcal{B} = (\underbrace{v_{1,1}, \dots, v_{1,r_1}}_{\text{Basis von } W_1}, \dots, \underbrace{v_{k,1}, \dots, v_{k,r_k}}_{\text{Basis von } W_k})$  eine Basis von  $V$ .

(iii)  $V = W_1 + \dots + W_k$  und  $\dim(V) = \dim(W_1) + \dots + \dim(W_k)$ .

*Beweis.* (i) $\Rightarrow$ (ii): Sei (i) erfüllt.  $\mathcal{B}$  ist ein Erzeugendensystem von  $V$  (wegen  $V = W_1 + \dots + W_k$ ). Es bleibt zu zeigen, dass  $\mathcal{B}$  ein linear unabhängiges System ist.

$\sum_{i=1}^k \sum_{j=1}^{r_i} \mu_{ij} v_{ij} = 0 \quad \forall \mu_{ij} \in \mathbb{K}$ . Sei  $w_i := \sum_{j=1}^{r_i} \mu_{ij} v_{ij}$ , dann gilt  $\sum_{i=1}^k w_i = 0$ .

Da die Summe  $v = w_1 + \dots + w_k$  direkt ist, folgt  $w_1 = \dots = w_k = 0$ . Das heißt  $w_i = \sum_{j=1}^{r_j} \mu_{ij} v_{ij} = 0 \quad \forall i \in \{1, \dots, k\}$ . Da  $v_{i1}, \dots, v_{ir_j}$  eine Basis bilden, folgt  $\mu_{ij} = 0 \quad \forall i \forall j$ .

(ii) $\Rightarrow$ (i): Sei (ii) erfüllt. Sei  $w_i \in W_i$  für  $i \in \{1, \dots, k\}$  und sei  $\sum_{i=1}^k w_i = 0$ . Da  $v_{i1}, \dots, v_{ir_j}$  eine Basis von  $w_i$  ist, gilt  $w_i = \sum_{j=1}^{r_j} \mu_{ij} v_{ij}$  für gewisse

$$\mu_{ij} \in \mathbb{K}.$$

Durch einsetzen erhält man  $\sum_{i=1}^n \sum_{j=1}^{r_j} \mu_{ij} v_{ij} = 0$ . Da  $\mathcal{B}$  eine Basis ist, folgt  $\mu_{ij} = 0 \quad \forall i \forall j$ . Das bedeutet  $w_1 = \dots = w_k = 0$ .

(ii) $\Rightarrow$ (iii): klar, denn die Dimension ist die Anzahl der Vektoren in einer Basis des Raumes.

(iii) $\Rightarrow$ (ii): Sei (iii) erfüllt.  $\mathcal{B}$  ist ein Erzeugendensystem.  $\mathcal{B}$  kann nicht linear abhängig sein, sonst könnte man in  $\mathcal{B}$  eine Basis von  $V$  wählen, die weniger als  $\dim(W_1) + \dots + \dim(W_k)$  Vektoren hat, was der Gleichung  $\dim(V) = \dim(W_1) + \dots + \dim(W_k)$  widerspricht.  $\square$

### 3.7 Projektive Räume \*

\* = optionales Thema

### 3.7.1 Projektive Räume

Für einen Vektorraum  $V$  ist der *projektive Raum*  $\mathbf{P}(V)$  von  $V$  die Menge der Äquivalenzklassen auf  $V \setminus \{0\}$  mit der Äquivalenz  $x \sim y$  von  $x$  und  $y$ , die durch die Bedingung definiert ist, dass ein  $\lambda \in \mathbb{K} \setminus \{0\}$  existiert, welches die Gleichung  $x = \lambda y$  erfüllt. Der projektive Raum  $\mathbf{P}(\mathbb{K}^{n+1})$  wird als  $\mathbb{K}\mathbf{P}^n$  bezeichnet; man nennt ihn der *n-dimensionale projektive Raum* über dem Körper  $\mathbb{K}$ . Man nennt die Elemente des projektiven Raums *Punkte*. Die Äquivalenzklasse von  $(x_0, \dots, x_n) \in \mathbb{K}^{n+1}$  in  $\mathbf{P}(\mathbb{K}^{n+1})$  wird als  $(x_0 : \dots : x_n)$  bezeichnet. Für  $p = (x_0 : \dots : x_n)$  nennt man  $x_0, \dots, x_n$  die *projektiven Koordinaten* des Punktes  $p$ . Die *projektiven Koordinaten* sind nur bis auf eine multiplikative Konstante bestimmt, denn  $(x_0 : \dots : x_n) = (\lambda x_0 : \dots : \lambda x_n)$  für alle  $\lambda \in \mathbb{K} \setminus \{0\}$ .

Ist  $U$  ein UVR von  $V$ , dann nennt man  $\mathbf{P}(U) \subseteq \mathbf{P}(V)$  einen *projektiven Untervektorraum* von  $U$  der Dimension  $\dim(U) - 1$ . Ein 0-dimensionaler projektiver Raum besteht aus einem einzigen Punkt. Ein 1-dimensionaler

projektive Unterräume nennt man (*projektive*) *Gerade* und 2-dimensionale projektive Unterräume nennt man (*projektive*) *Ebene*.

### 3.7.2 Zwei projektive Punkte definieren genau eine Gerade

Die Abbildung  $(x_1, \dots, x_n) \mapsto (1 : x_1 : \dots : x_n)$  ist die Standardeinbettung von  $\mathbb{K}^n$  in  $\mathbb{KP}^n$ . Die Einbettung deckt alle Punkte  $(x_0 : x_1 : \dots : x_n)$  ab, für welche  $x_0$  ungleich 0 sind. Für Körper wie  $\mathbb{K} = \mathbb{R}$  und  $\mathbb{K} = \mathbb{C}$  werden alle anderen Punkte  $(0 : x_1 : \dots : x_n)$  als Punkte im Unendlichen interpretiert.

**Prop.** *Zwei verschiedene Punkte eines projektiven Raums liegen in genau einer projektiven Geraden.*

*Beweis.* Zwei Punkte in  $\mathbf{P}(V)$  können als 0-dimensionale projektive Unterräume  $\mathbf{P}(A)$  und  $\mathbf{P}(B)$  definiert werden, wobei  $A$  und  $B$  zwei verschiedene 1-dimensionale Untervektorräume von  $V$  sind.

*Existenz:* Der Vektorraum  $A + B$  hat Dimension  $\dim(A) + \dim(B) = 2$ .

Also ist  $\mathbf{P}(A + B)$  die projektive Gerade, welche die Punkte  $\mathbf{P}(A)$  und  $\mathbf{P}(B)$  enthält.

*Eindeutigkeit:* Wir betrachten nun eine beliebige projektive Gerade, d.h., den projektiven Raum  $\mathbf{P}(U)$ , der durch einen 2-dimensionalen Untervektorraum  $U$  von  $V$  definiert ist. Wenn  $\mathbf{P}(A)$  und  $\mathbf{P}(B)$  in  $\mathbf{P}(U)$  liegen, dann gilt  $A \subseteq U$  und  $B \subseteq U$ . Dann folgt aber auch  $A + B \subseteq U$ . Da  $A + B$  und  $U$  zwei-dimensional sind, erhalten wir  $A + B = U$ . Das bedeutet  $\mathbf{P}(U) = \mathbf{P}(A + B)$ .  $\square$

### 3.7.3 Zwei projektive Geraden einer projektiven Ebene schneiden sich in genau einem Punkt

**Prop.** *Zwei verschiedene projektive geraden einer projektiven Ebene schneiden sich in genau einem Punkt.*

*Beweis.* Wir betrachten eine projektive Ebene  $\mathbf{P}(V)$ , das heißt,  $V$  ist 3-dimensionaler Vektorraum. Zwei verschiedene projektive Geraden  $\mathbf{P}(U)$

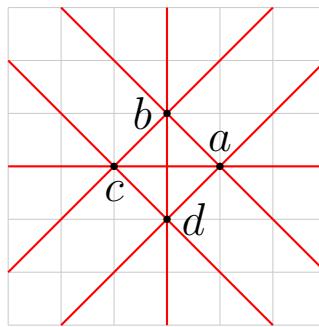
und  $\mathbf{P}(W)$  sind durch zwei verschiedene zwei-dimensionale Untervektorräume  $U$  und  $W$  von  $V$  gegeben. Die Summe  $U + W$  ist der gesamte  $V$ , denn sonst wäre die Dimension von  $U + W$  geringer als  $U + W$ . Dann wäre die Dimension von  $U + W$  gleich 2, denn  $U + W$  die zwei-dimensionale Räume  $U$  und  $W$  als Untervektorräume enthält. Dann würde man aber  $U + W = U = W$  haben, was der Annahme  $U \neq W$  widerspricht.

*Existenz:* Aus  $3 = \dim(V) = \dim(U + W) = \dim(U) + \dim(W) - \dim(U \cap W) = 2 + 2 - \dim(U \cap W)$  erhalten wir, dass  $\dim(U \cap W) = 1$  ist. Somit ist  $\mathbf{P}(U \cap W)$  der Punkt, der in den beiden projektiven Geraden  $\mathbf{P}(U)$  und  $\mathbf{P}(W)$  enthalten ist.

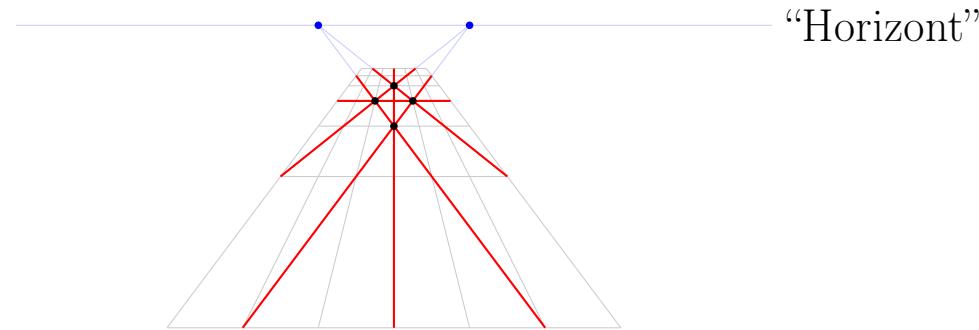
*Eindeutigkeit:* Sei  $\mathbf{P}(A)$  ein projektiver Punkt, in dem sich  $\mathbf{P}(U)$  und  $\mathbf{P}(W)$  scheiden. Das heißt,  $A$  ist ein-dimensionaler Untervektorraum von  $V$  mit  $A \subseteq U$  und  $A \subseteq W$ . Oben haben wir gezeigt, dass  $U \cap W$  ein-dimensionale ist. Aus  $A \subseteq U \cap W$  und  $\dim(A) = \dim(U \cap W) = 1$  folgt nun die Gleichheit  $A = U \cap W$ . Das heißt,  $\mathbf{P}(A) = \mathbf{P}(U \cap W)$ .  $\square$

**Bsp.** Vier Punkte  $(1, 0), (-1, 0), (0, 1), (0, -1)$  werden in die reelle projektive Ebene  $\mathbb{RP}^2$  eingebettet. Wir betrachten also die Punkte  $a = (1 : 1 : 0), b = (1 : 0 : 1), c = (1 : -1 : 0), d = (1 : 0 : -1)$ . Je zwei von diesen vier Punkten definieren eine Gerade und je zwei gerade schneiden sich in genau einem Punkt. Wir bezeichnen die Geraden als  $ab, ac, ad, bc, bd, cd$ . Nun wollen wir die Schnittpunkte der paaren dieser Geraden berechnen. Bei manchen Paaren ist das klar;  $ab$  und  $ac$  schneiden sich in  $a$ ,  $bc$  und  $bd$  schneiden sich in  $b$  usw. Die interessanten Schnittpunkte sind die, die nicht in  $\{a, b, c, d\}$  liegen. Wir wollen also bestimmen, wo sich  $ab$  und  $cd$ ,  $ac$  und  $bd$  sowie  $ad$  und  $bc$  schneiden.

Hier das Bild in  $\mathbb{R}^2$ : in  $\mathbb{R}^2$  schneiden sich zum Beispiel die Geraden  $ab$  und  $cd$  nicht,  $bc$  und  $ad$  schneiden sich ebenfalls nicht.



Das projektive Bild dazu, in dem sich jede Wahl von zwei der sechs Geraden einen Schnittpunkt hat. Die Schnittpunkte  $ab \cap cd$  und  $bc \cap ad$  haben die Form  $(0 : x_1 : x_2)$  (sie befinden sich im Unendlichen).



Der Schnittpunkt von projektiven Geraden  $ab$  und  $cd$  ist  $(0 : -1 : 1)$ .

In der Sprache der Vektorräume heißt das, dass der Vektor  $(0, -1, 1)$  im Schnitt von  $\text{lin}((1, 1, 0), (1, 0, 1))$  und  $\text{lin}((1, -1, 0), (1, 0, -1))$  liegt.

Der Schnittpunkt von  $bc$  und  $ad$  ist  $(0 : 1 : 1)$ .

**Bsp.** Im Fall vom Körper  $\mathbb{K} = \{0, 1\}$  lässt sich der projektive Raum  $\mathbf{P}(V)$  mit  $V \setminus \{0\}$  identifizieren. Die projektive Ebene  $\mathbb{K}\mathbf{P}^2$  für  $\mathbb{K} = \{0, 1\}$  heißt **Fano-Ebene**. Die Punkte und Geraden der Fano-Ebene können durch das folgende Diagramm: IM AUFBAU.

# **4 Lineare Abbildungen**

Lineare Abbildungen sind Abbildungen von einem Vektorraum in einen Vektorraum, die die lineare Struktur erhalten. Lineare Abbildungen nennt man in verschiedenen Teilgebieten der Mathematik auch lineare Transformationen, lineare Operatoren oder auch Homomorphismen von Vektorräumen.

## **4.1 Beispiele von linearen Abbildungen**

In diesem Abschnitt sei der zugrundeliegende Körper  $\mathbb{K} = \mathbb{R}$ .

IM AUFBAU: Bilder - Quadrat mit einem Piktogramm drin.

#### **4.1.1 90°-Drehung**

90°-Drehung mit Drehzentrum in 0 im Gegenuhrzeigersinn.

$$F : \mathbb{R}^2 \rightarrow \mathbb{R}^2, (x_1, x_2) \mapsto (-x_2, x_1)$$

#### **4.1.2 Projektion in $\mathbb{R}^3$**

$$F : \mathbb{R}^3 \rightarrow \mathbb{R}^3, (x_1, x_2, x_3) \mapsto (x_1, x_2, 0)$$

#### **4.1.3 Projektion von $\mathbb{R}^3$ nach $\mathbb{R}^2$**

$$F : \mathbb{R}^3 \rightarrow \mathbb{R}^2, (x_1, x_2, x_3) \mapsto (x_1, x_2)$$

#### **4.1.4 Punktspiegelung**

Spiegelung an der 0.

$$F : x = (x_1, x_2) \mapsto -x = (-x_1, -x_2)$$

#### **4.1.5 Spiegelung an einer Geraden**

Spiegelung an  $\mathbb{R} \times \{0\}$  ( $x_1$ -Achse).  $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2, (x_1, x_2) \mapsto (x_1, -x_2)$

#### **4.1.6 Scherung**

$F : \mathbb{R}^2 \rightarrow \mathbb{R}^2, (x_1, x_2) \mapsto (x_1 + x_2, x_2)$

#### **4.1.7 Streckung**

$F : \mathbb{R}^2 \rightarrow \mathbb{R}^2, F : (x_1, x_2) \mapsto (x_1, 2 \cdot x_2)$

#### **4.1.8 Zyklische Verschiebung der Komponenten**

Sei  $n \in \mathbb{N}$ .  $F : \mathbb{R}^n \rightarrow \mathbb{R}^n, (x_1, \dots, x_n) \mapsto (x_n, x_1, \dots, x_{n-1})$ .

Die vorigen Abbildungen haben das Folgende gemeinsam: jede Komponente

von  $F(x)$  ist Linearkombination der Komponenten von  $x$  mit konstanten Koeffizienten.

## 4.2 Lineare Abbildungen für allgemeine Vektorräume

### 4.2.1 Lineare Abbildung

Seien  $V$  und  $W$  Vektorräume über  $\mathbb{K}$ . Sei  $F : V \rightarrow W$ . Die Abbildung  $F$  heißt linear, falls:

$$(L1) \quad F(v + w) = F(v) + F(w) \quad \forall v, w \in V$$

$$(L2) \quad F(\lambda v) = \lambda F(v) \quad \forall v \in V, \forall \lambda \in \mathbb{K}$$

**Bem.** Offensichtlich gelten (L1) und (L2) genau dann, wenn folgendes gilt:

$$(L) \quad F(\lambda v + \mu w) = \lambda F(v) + \mu F(w) \quad \forall v, w \in V, \forall \lambda, \mu \in \mathbb{K}$$

#### 4.2.2 Lineare Abbildungen und die Grundbegriffe für Vektorräume

**Prop.** Seien  $V$  und  $W$  Vektorräume über  $\mathbb{K}$ . Sei  $F : V \rightarrow W$  linear. Dann gilt:

- (i)  $F(0) = 0$  und  $F(v - w) = F(v) - F(w)$   $\forall v, w \in V$
- (ii)  $F(\lambda_1 v_1 + \dots + \lambda_n v_n) = \lambda_1 F(v_1) + \dots + \lambda_n F(v_n)$   
 $\forall n \in \mathbb{N}_0, \lambda_1, \dots, \lambda_n \in \mathbb{K}, v_1, \dots, v_n \in \mathbb{K}$
- (iii) Ist  $(v_i)_{i \in I}$  eine linear abhängige Familie von Vektoren aus  $V$ , dann ist  $(F(v_i))_{i \in I}$  eine linear abhängige Familie von Vektoren aus  $W$ .
- (iv) Ist  $V'$  Untervektorraum von  $V$ , dann ist  $F(V')$  Untervektorraum von  $W$ .
- (v) Ist  $W'$  Untervektorraum von  $W$ , dann ist  $F^{-1}(W')$  Untervektorraum von  $V$ .

(vi)  $\dim(F(V)) \leq \dim(V)$

(vii) Ist  $F$  bijektiv, dann ist auch  $F' : W \rightarrow V$  linear.

*Beweis.* (i)  $F(x) = F(0 + x) = F(0) + F(x) \quad \forall x \in V \Rightarrow F(0) = F(x) - F(x) = 0$

(ii) folgt durch das iterative Anwenden von (L1) und (L2):

$$\begin{aligned} F(\lambda_1 v_1 + \cdots + \lambda_n v_n) &= F(\lambda_1 v_1 + \cdots + \lambda_{n-1} v_{n-1}) + F(\lambda_n v_n) \\ &= F(\lambda_1 v_1 + \cdots + \lambda_{n-1} v_{n-1}) + \lambda_n F(v_n) \end{aligned}$$

und so weiter für alle  $n \in \mathbb{N}$ .

(iii) Wenn  $(v_i)_{i \in I}$  linear abhängig ist, dann existieren  $\lambda_{i_1}, \dots, \lambda_{i_n} \in \mathbb{K}$  (nicht alle Null) und  $v_{i_1}, \dots, v_{i_n}$  mit  $i_1, \dots, i_n \in I$  ( $n \in \mathbb{N}$ ), sodass  $\lambda_{i_1} v_{i_1} + \cdots + \lambda_{i_n} v_{i_n} = 0$ . Das Anwenden von  $F$  mit der Berücksichtigung von (i) und (ii) ergibt

$$0 = F(0) = F(\lambda_{i_1} v_{i_1} + \cdots + \lambda_{i_n} v_{i_n}) = \lambda_{i_1} F(v_{i_1}) + \cdots + \lambda_{i_n} F(v_{i_n}).$$

Das zeigt die lineare Abhangigkeit von  $(F(v_i))_{i \in I}$ .

- (iv)  $F(V') \neq \emptyset$ , denn  $0 \in V'$  und somit  $0 = F(0) \in F(V')$ . Wenn  $w_1, w_2 \in F(V')$ , dann gilt  $w_1 = F(v_1), w_2 = F(v_2)$  fur gewisse  $v_1, v_2 \in V'$ .  $V'$  ist Untervektorraum von  $V$  und somit  $v_1 + v_2 \in V' \Rightarrow w_1 + w_2 = F(v_1) + F(v_2) = F(v_1 + v_2) \in F(V')$ .

Analog zeigt man auch, dass fur jedes  $\lambda \in \mathbb{K}$  und jedes  $w \in F(V')$  die Bedingung  $\lambda w \in F(V')$  gilt.

- (v)  $F^{-1}(W') \neq \emptyset$ , denn  $0 \in F^{-1}(W')$  (d.h.  $0 = F(0) \in W'$ ).

Seien  $v_1, v_2 \in F^{-1}(W')$ , d.h.  $F(v_1), F(v_2) \in W'$ . Es folgt  $F(v_1 + v_2) = F(v_1) + F(v_2) \in W'$ . D.h.  $v_1 + v_2 \in F^{-1}(W')$ .

Analog zeigt man, dass  $\lambda v \in F^{-1}(W')$  fur alle  $\lambda \in \mathbb{K}$  und alle  $v \in F^{-1}(W')$  gilt.

- (vi) Ubung.

(vii) Übung. □

#### 4.2.3 Komposition von linearen Abbildungen

**Prop.** Seien  $U, V$  und  $W$  Vektorräume über  $\mathbb{K}$ . Seien  $G : U \rightarrow V$  und  $F : V \rightarrow W$  linear. Dann ist auch  $F \circ G$  linear.

*Beweis.* Seien  $u_1, u_2 \in U$ . Dann ist wegen der Linearität von  $G$  und  $F$

$$\begin{aligned}(F \circ G)(u_1 + u_2) &= F(G(u_1 + u_2)) \\&= F(G(u_1) + G(u_2)) \\&= F(G(u_1)) + F(G(u_2)) \\&= (F \circ G)(u_1) + (F \circ G)(u_2).\end{aligned}$$

Analog verifiziert man, dass  $(F \circ G)(\lambda u) = \lambda(F \circ G)(u)$  für alle  $\lambda \in \mathbb{K}, u \in U$  gilt. □

#### 4.2.4 Vektorräume der linearen Abbildungen

Seien  $U$  und  $V$  Vektorräume über  $\mathbb{K}$ . Wir bezeichnen durch  $\text{Lin}_{\mathbb{K}}(V, W)$  die Menge aller linearen Abbildungen von  $V$  nach  $W$ .

**Prop.** *Seien  $V$  und  $W$  Vektorräume über  $\mathbb{K}$ . Die Menge  $\text{Lin}(V, W)$  mit den Operationen  $(F + G)(v) := F(v) + G(v) \quad \forall v \in V$  (Addition) und  $(\lambda \cdot F)(v) := \lambda \cdot F(v) \quad \forall \lambda \in \mathbb{K}, v \in V$  (Skalarmultiplikation) ist ein Vektorraum.*

*Beweis.* Wir bemerken, dass die Menge  $W^V$  mit der Addition und Skalarmultiplikation wie oben ein Vektorraum ist (der Beweis ist direkt). Weil  $\text{Lin}(V, W) \subseteq W^V$  gilt, reicht es zu zeigen, dass  $\text{Lin}(V, W)$  ein Untervektorraum von  $W^V$  ist.

Die Nullabbildung von  $V$  nach  $W$  (d.h.  $F$  mit  $F(v) = 0$  für jedes  $v \in V$ ) gehört zu  $\text{Lin}(V, W)$ . Daher ist  $\text{Lin}(V, W) \neq \emptyset$ .

Seien  $F_1, F_2 \in \text{Lin}(V, W)$ . Zu zeigen:  $F_1 + F_2 \in \text{Lin}(V, W)$ . Seien  $v_1, v_2 \in$

$V$ . Dann ist

$$\begin{aligned}(F_1 + F_2)(v_1 + v_2) &= F_1(v_1 + v_2) + F_2(v_1 + v_2) \\&= F_1(v_1) + F_1(v_2) + F_2(v_1) + F_2(v_2) \\&= (F_1(v_1) + F_2(v_1)) + (F_1(v_2) + F_2(v_2)) \\&= (F_1 + F_2)(v_1) + (F_1 + F_2)(v_2).\end{aligned}$$

Analog zeigt man, dass  $(F_1 + F_2)(\lambda v) = \lambda(F_1 + F_2)(v)$   $\forall \lambda \in \mathbb{K}, v \in V$  gilt, und dass für  $\mu \in \mathbb{K}$  und  $F \in \text{Lin}(V, W)$  die Bedingung  $\mu F \in \text{Lin}(V, W)$  erfüllt ist.  $\square$

#### 4.2.5 Lineare Abbildungen eines Vektorraums

Sei  $V$  Vektorraum über  $\mathbb{K}$ . Wir führen die Bezeichnung  $\text{Lin}_{\mathbb{K}}(V) := \text{Lin}_{\mathbb{K}}(V, V)$  ein. Wir nennen die Elemente von  $\text{Lin}_{\mathbb{K}}(V)$  die linearen Abbildungen des Vektorraums  $V$ .

**Prop.** Sei  $V$  Vektorraum über  $\mathbb{K}$ . Dann ist die Menge  $\text{Lin}(V)$  mit den Operationen  $(F + G)(v) := F(v) + G(v) \quad \forall v \in V, \forall F, G \in \text{Lin}(V)$  (Addition) und  $(F \cdot G)(v) := (F \circ G)(v) \quad \forall v \in V, \forall F, G \in \text{Lin}(V)$  (Multiplikation) ein Ring mit Eins.

*Beweis.* Da  $(V, +)$  eine kommutative Gruppe ist, ist auch  $(\text{Lin}(V), +)$  eine kommutative Gruppe. Die Multiplikation ist assoziativ, weil  $\circ$  (Komposition) assoziativ ist. Die beiden Distributivgesetze gelten (der Beweis ist direkt). Das neutrale Element ist  $\text{id}_V$  (identische Abbildung auf  $V$ ).  $\square$

## 4.3 Matrizenmultiplikation und lineare Abbildungen für Räume $\mathbb{K}^n$

Eine lineare Abbildung von  $\mathbb{K}^n$  nach  $\mathbb{K}^m$  kann durch eine  $m \times n$  Matrix beschrieben werden. Die Komposition von zwei linearen Abbildungen  $F$  und  $G$  mit  $\mathbb{K}^n \xrightarrow{G} \mathbb{K}^p \xrightarrow{F} \mathbb{K}^m$  entspricht dann dem Produkt von zwei Matrizen.

### 4.3.1 Multiplikation von Matrizen

Seien  $A = (a_{ij})_{i=1,j=1}^{m,n} \in \mathbb{K}^{m \times n}$  und  $B = (b_{jk})_{j=1,k=1}^{n,p} \in \mathbb{K}^{n \times p}$  ( $m, n, p \in \mathbb{N}$ ). Dann heißt  $C = (c_{ik})_{i=1,k=1}^{m,p} := A \cdot B \in \mathbb{K}^{m \times p}$  mit

$$c_{ik} := \sum_{j=1}^n a_{ij} b_{jk}$$

das Produkt von  $A$  und  $B$ . Das heißt,  $c_{ik}$  hängt von der  $i$ -ten Zeile von  $A$  und der  $k$ -ten Spalte von  $B$  ab:

$$\begin{pmatrix} & & \\ a_{i1} & \dots & a_{in} \end{pmatrix} \cdot \begin{pmatrix} b_{1k} \\ \vdots \\ b_{nk} \end{pmatrix} = \begin{pmatrix} & \\ & c_{ik} \end{pmatrix}$$

**Bsp** ( $2 \times 2$  Matrizen).

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

$$c_{11} = a_{11}b_{11} + a_{12}b_{21}$$

$$c_{12} = a_{11}b_{12} + a_{12}b_{22}$$

$$c_{21} = a_{21}b_{11} + a_{22}b_{21}$$

$$c_{22} = a_{21}b_{12} + a_{22}b_{22}$$

**Bem.** Die Matrix-Vektor-Multiplikation ist ein Spezialfall der Matrix-Matrix-Multiplikation. Für  $A = (a_{ij}) \in \mathbb{K}^{m \times n}$  und  $x = (x_1, \dots, x_n) \in \mathbb{K}^n$  Dann ist

$Ax$  ein Element von  $\mathbb{K}^m$  mit

$$Ax = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ & \vdots & \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11}x_1 + \dots + a_{1n}x_n \\ \vdots \\ a_{m1}x_1 + \dots + a_{mn}x_n \end{pmatrix}$$

wobei man hier die Elemente aus  $\mathbb{K}^n$  und  $\mathbb{K}^m$  als Spalten interpretiert.

Nun kann auch ein LGS mit  $m$  Gleichungen und  $n$  Unbekannten kompakt als  $Ax = b$  beschrieben werden.

**Bem.** Wenn wir im Produkt  $AB$  die Matrix  $B$  spaltenweise als

$$B = \begin{pmatrix} | & & | \\ b_1 & \cdots & b_p \\ | & & | \end{pmatrix}$$

so kann auch das Produkt  $AB$  spaltenweise als

$$AB = \begin{pmatrix} | & | \\ Ab_1 & \cdots & Ab_p \\ | & | \end{pmatrix}$$

dargestellt werden. Das heißt: die Matrix  $A$  (er erste Faktor) wirkt auf jede Spalte der Matrix  $B$  (der zweite Faktor)

**Bem.** Durch ein Matrix-Vektor-Produkt  $Ax$  mit  $A \in \mathbb{K}^{m \times n}$  und  $x \in \mathbb{K}^n$  können Linearkombinationen von Vektoren aus  $\mathbb{K}^m$  beschrieben werden, indem man die Matrix  $A$  spaltenweise interpretiert. Ist

$$A = \begin{pmatrix} | & | \\ a_1 & a_n \\ | & | \end{pmatrix}$$

und  $x = (x_1, \dots, x_n) \in \mathbb{K}^n$  so gilt

$$Ax = x_1a_1 + \dots + x_na_n.$$

Das heißt,  $Ax$  ist die Linearkombination der Spalten von  $A$ , deren Koeffizienten die Komponenten des Vektors  $x$  sind.

### 4.3.2 Matrix einer linearen Abbildung von $\mathbb{K}^n$ nach $\mathbb{K}^m$

**Prop.** Sei  $F : \mathbb{K}^m \rightarrow \mathbb{K}^n$  linear ( $m, n \in \mathbb{N}$ ). Dann existiert genau eine Matrix  $A \in \mathbb{K}^{m \times n}$  mit  $F(x) = Ax$  für alle  $x \in \mathbb{K}^n$ . Es gilt

$$A = \begin{pmatrix} | & & | \\ F(e_1) & \cdots & F(e_n) \\ | & & | \end{pmatrix}$$

*Beweis.* Sei  $x = (x_1, \dots, x_n) \in \mathbb{K}^n$ . Dann ist  $x = x_1e_1 + \dots + x_ne_n$ . Somit gilt

$$F(x) = F(x_1e_1 + \dots + x_ne_n) = x_1F(e_1) + \dots + x_nF(e_n) = \begin{pmatrix} | & & | \\ F(e_1) & \cdots & F(e_n) \\ | & & | \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}.$$

Es bleibt die Eindeutigkeit von  $A$  zu zeigen. Sei  $F(x) = Ax$  für alle  $x \in \mathbb{K}^n$ . Für  $x = e_i$  ( $i \in \{1, \dots, n\}$ ) gilt  $F(e_i) = Ae_i$ . Dabei ist  $Ae_i$  die  $i$ -te Spalte von  $A$ . Das zeigt die Eindeutigkeit.  $\square$

Für  $F$  und  $A$  wie in der vorigen Proposition nennen wir  $A$  die Matrix von  $F$ .  $F$  heißt die durch  $A$  definierte lineare Abbildung.

Bsp.

- Identische Abbildung:  $F : \mathbb{K}^3 \rightarrow \mathbb{K}^3$  mit  $F(x) = x$  für alle  $x \in \mathbb{K}^3$ .

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

- Drehung um 90 Grad.  $\mathbb{K} = \mathbb{R}$ ,  $(x_1, x_2) \mapsto (-x_2, x_1)$

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

- Projektion auf die Ebene  $\mathbb{R}^2 \times \{0\}$  innerhalb von  $\mathbb{R}^3$ .  $(x_1, x_2, x_3) \mapsto (x_1, x_2, 0)$

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

- Projektion auf die ersten beiden Koordinaten als Abbildung von  $\mathbb{R}^3$  nach  $\mathbb{R}^2$ .  $(x_1, x_2, x_3) \mapsto (x_1, x_2)$

$$\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} \mapsto \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

### 4.3.3 Matrix der Komposition von linearen Abbildungen

**Thm.** Seien  $l, m, n \in \mathbb{N}$ , seien  $G : \mathbb{K}^n \rightarrow \mathbb{K}^m$  und  $F : \mathbb{K}^m \rightarrow \mathbb{K}^l$  linear. Sei  $B \in \mathbb{K}^{m \times n}$  die Matrix von  $G$  und sei  $A \in \mathbb{K}^{l \times m}$  die Matrix von  $F$ . Dann

ist  $AB$  die Matrix von  $F \circ G$ .

*Beweis.* Es gilt  $F(y) = Ay$  und  $G(x) = Bx$  für alle  $x \in \mathbb{K}^n$  und alle  $y \in \mathbb{K}^m$ . Daher gilt  $(F \circ G)(x) = F(G(x)) = F(Bx) = A \cdot (B \cdot x)$ .

Zu zeigen:  $A \cdot (B \cdot x) = (A \cdot B) \cdot x$ . Sei  $x = (x_1, \dots, x_n)$ , sei  $A = (a_{ij})_{i=1, j=1}^{l, m}$  und sei  $B = (b_{jk})_{j=1, k=1}^{m, n}$ . Man hat:

$$\text{Die } i\text{-te Komponente von } A \cdot (B \cdot x) \text{ ist } \sum_{j=1}^m a_{ij} \underbrace{\left( \sum_{k=1}^n b_{jk} x_k \right)}_{j\text{-te Komponente von } Bx},$$

$$\text{Die } i\text{-te Komponente von } (A \cdot B) \cdot x \text{ ist } \sum_{k=1}^n \underbrace{\left( \sum_{j=1}^m a_{ij} b_{jk} \right)}_{\text{die Komponente in der Position } (i, k) \text{ von } AB} x_k$$

Durch das Auflösen der Klammern sieht man, dass diese beiden geschach-

telten Summen mit der Summe

$$\sum_{\substack{j \in \{1, \dots, m\} \\ k \in \{1, \dots, n\}}} a_{ij} b_{jk} x_k \quad (4.3.1)$$

übereinstimmen. Das ergibt die Behauptung.  $\square$

#### 4.3.4 Rechenregeln für Matrizen

Für  $m, n \in \mathbb{N}$  ist die Menge  $\mathbb{K}^{m \times n}$  aller Matrizen der Größe  $m \times n$  einen Vektorraum der Form  $\mathbb{K}^X$  mit  $X = \{1, \dots, m\} \times \{1, \dots, n\}$ . Diese Vektorräume wurden in 3.1.1 eingeführt. Somit hat man für  $\mathbb{K}^{m \times n}$  Addition und Skalarmultiplikation. Etwa:

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} + \begin{pmatrix} a'_{11} & a'_{12} \\ a'_{21} & a'_{22} \end{pmatrix} = \begin{pmatrix} a_{11} + a'_{11} & a_{12} + a'_{12} \\ a_{21} + a'_{21} & a_{22} + a'_{22} \end{pmatrix} \quad (4.3.2)$$

$$\lambda \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} \lambda a_{11} & \lambda a_{12} \\ \lambda a_{21} & \lambda a_{22} \end{pmatrix} \quad (4.3.3)$$

**Thm.** Seien  $m, n, r, s \in \mathbb{N}$ . Seien  $A, A' \in \mathbb{K}^{m \times n}$ ,  $B, B' \in \mathbb{K}^{r \times s}$  und  $\lambda \in \mathbb{K}$ . Dann gilt:

- (i)  $A \cdot (B + B') = A \cdot B + A \cdot B'$
- (ii)  $(A + A') \cdot B = A \cdot B + A' \cdot B$
- (iii)  $A \cdot (\lambda B) = (\lambda A) \cdot B = \lambda(A \cdot B)$
- (iv)  $(A \cdot B) \cdot C = A \cdot (B \cdot C)$
- (v)  $(A \cdot B)^\top = B^\top \cdot A^\top$

*Beweis.* Die ersten drei Formeln sind einfach. Die letzten beiden Formeln sind Übungsaufgaben.  $\square$

### 4.3.5 Die Einheitsmatrix

Seien  $i, j \in \mathbb{N}$ . Dann heißt

$$\delta_{ij} = \begin{cases} 1 & \text{falls } i = j \\ 0 & \text{falls } i \neq j \end{cases} \quad (4.3.4)$$

das *Kronecker Delta* von  $(i, j)$ .

Für  $n \in \mathbb{N}$  heißt die Matrix  $I_n := (\delta_{ij})_{i,j=1}^n$  die Einheitsmatrix der Größe  $n \times n$ . Wenn die Wahl von  $n$  klar ist, schreibt man auch  $I$ . Oder auch:  $\mathbb{I}$ .  
Etwa:

$$I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad (4.3.5)$$

Für jede Matrix  $A \in \mathbb{K}^{m \times n}$  ( $m, n \in \mathbb{N}$ ) gilt offensichtlich  $I_m A = A I_n = A$ .  
Die durch  $I_n$  definierte Abbildung ist die identische Abbildung auf  $\mathbb{K}^n$ .

**Prop.** Sei  $n \in \mathbb{N}$ . Dann ist  $\mathbb{K}^{n \times n}$  mit der Addition und Multiplikation von Matrizen ein Ring mit Eins.

Das neutrale Element bzgl. der Multiplikation ist die Einheitsmatrix  $I_n$ . Die Matrix in  $\mathbb{K}^{n \times n}$  ( $n \in \mathbb{N}$ ), deren Einträge alle gleich Null sind, wird die Nullmatrix genannt und durch  $O$  bezeichnet.

*Beweis.* Alle Ringeigenschaften lassen sich direkt verifizieren. Es sei bemerkt, dass man im Wesentlichen keinen Unterschied zwischen  $\mathbb{K}^{n \times n}$  und  $\text{Lin}(\mathbb{K}^n)$  hat. Man sagt: die Ringe  $\mathbb{K}^{n \times n}$  und  $\text{Lin}(\mathbb{K}^n)$  sind isomorph (die genaue Definition wird erstmal nicht angegeben;  $\rightsquigarrow$  gleich, aber nicht das-selbe).  $\square$

#### 4.3.6 Invertierbarkeit von Matrizen

Sei  $n \in \mathbb{N}$  und  $A \in \mathbb{K}^{n \times n}$ . Dann heißt  $A$  invertierbar, falls eine Matrix  $B \in \mathbb{K}^{n \times n}$  mit  $BA = AB = I$  existiert. Für eine invertierbare Matrix  $A$  ist

die Matrix  $B$  wie oben eindeutig bestimmt (Übungsaufgabe). Diese Matrix wird die *inverse Matrix* von  $A$  genannt und durch  $A^{-1}$  bezeichnet.

**Prop.** Sei  $n \in \mathbb{N}$ , sei  $F : \mathbb{K}^n \rightarrow \mathbb{K}^n$  linear uns sei  $A$  die Matrix von  $F$ . Dann sind die folgenden Bedingungen äquivalent:

(i)  $F$  ist invertierbar.

(ii)  $A$  ist invertierbar.

Wenn (i) und (ii) gelten, dann ist  $A^{-1}$  die Matrix der Abbildung  $F^{-1}$ .

*Beweis.* Direkte Folgerung von 4.3.2 und 4.3.3. □

### 4.3.7 Eigenschaften der inversen Matrizen

**Prop.** Sei  $n \in \mathbb{N}$ , seien  $A, B \in \mathbb{K}^{n \times n}$  invertierbare Matrizen. Dann sind  $A^{-1}, A^\top$  und  $AB$  invertierbar und es gilt:

$$(A^{-1})^{-1} = A \tag{4.3.6}$$

$$(A^\top)^{-1} = (A^{-1})^\top \tag{4.3.7}$$

$$(AB)^{-1} = B^{-1}A^{-1} \tag{4.3.8}$$

*Beweis.* Übungsaufgabe. □

### 4.3.8 Allgemeine lineare Gruppe

Für  $n \in \mathbb{N}$  sei („General Linear“)

$$\mathrm{GL}_n(\mathbb{K}) := \{A \in \mathbb{K}^{n \times n} : A \text{ ist invertierbar}\}. \tag{4.3.9}$$

**Prop.** Sei  $n \in \mathbb{N}$ . Dann gilt  $A \cdot B \in \mathrm{GL}_n(\mathbb{K})$  für alle  $A, B \in \mathrm{GL}_n(\mathbb{K})$  und darüber hinaus ist  $(\mathrm{GL}_n(\mathbb{K}), \cdot)$  eine Gruppe.

### 4.3.9 Elementartransformationen linearer Gleichungssysteme und Matrizenmultiplikation

Wir betrachten ein beliebiges LGS  $\sum_{j=1}^n a_{ij}x_j = b_i \quad \forall i \in \{1, \dots, m\}$  ( $m, n \in \mathbb{N}$ ) in Unbekannten  $x_1, \dots, x_n \in \mathbb{K}$ , mit  $a_{ij}, b_i \in \mathbb{K} \quad \forall i, j$ . Das System kann als  $Ax = b$  geschrieben werden, wobei

$$A = (a_{ij})_{i=1, j=1}^{m, n}, \quad b = (b_i)_{i=1}^m, \quad x = (x_j)_{j=1}^n.$$

Wir beschreiben die 3 Typen der Elementartransformationen von LGS durch Matrizenmultiplikation.

**Typ 1.** (Vertauschen der Gleichungen  $i$  und  $j$  mit  $i, j \in \{1, \dots, m\}, i \neq j$ .)

Das neue System hat die Form  $(T_1 A)x = T_1 b$ , wobei  $T_1$  die Einheitsmatrix mit vertauschten Zeilen  $i$  und  $j$  ist.  $T_1$  ist invertierbar und  $(T_1)^{-1} = T_1$ .

**Typ 2.** (Für  $i \in \{1, \dots, m\}$  und  $\alpha \in \mathbb{K} \setminus \{0\}$  wird die  $i$ -te Gleichung mit  $\alpha$  multipliziert.)

$$T_2(\alpha) = \begin{pmatrix} 1 & & & \\ & \ddots & & \\ & & 1 & \\ & & & \alpha \\ & & & & 1 \\ & & & & & \ddots \\ & & & & & & 1 \end{pmatrix} = I_n + (\alpha - 1)e_i e_i^\top$$

$T_2(\alpha)$  ist umkehrbar und  $T_2(\alpha)^{-1} = T_2(\frac{1}{\alpha})$ .

**Typ 3.** (Für  $i, j \in \{1, \dots, m\}$  und  $\alpha \in \mathbb{K}$  wird zu der  $i$ -ten Gleichung das  $\alpha$ -fache der  $j$ -ten Gleichung addiert.)

$$T_3(\alpha) = I_n + \alpha e_i e_j^\top$$

$Ax = b$  wird zu  $(T_3(\alpha)A)x = T_3(\alpha)b$  überführt.  $T_3$  ist invertierbar und  $T_3(\alpha)^{-1} = T_3(-\alpha)$ .

## 4.4 Bild, Kern und verwandte Begriffe

### 4.4.1 Bild, Kern und Faser

Seien  $V$  und  $W$  Vektorräume über  $\mathbb{K}$  und sei  $F : V \rightarrow W$  linear. Sei  $w \in W$ . Dann definieren wir das *Bild* (4.4.1) von  $F$ , den *Kern* (4.4.2) von  $F$ , sowie die *Faser* (4.4.3) von  $F$  über  $w$ :

$$\text{im}(F) := F(V) \tag{4.4.1}$$

$$\ker(F) := F^{-1}(\{0\}) = \{x \in V : F(x) = 0\} \tag{4.4.2}$$

$$F^{-1}(w) := F^{-1}(\{w\}) = \{x \in V : F(x) = w\} \tag{4.4.3}$$

Analog führt man  $\text{im}(A)$  und  $\ker(A)$  für Matrizen  $A \in \mathbb{K}^{m \times n}$  ( $m, n \in \mathbb{N}$ ) ein:

$$\text{im}(A) := \{Ax : x \in \mathbb{K}^n\} \subseteq \mathbb{K}^m \tag{4.4.4}$$

$$\ker(A) := \{x \in \mathbb{K}^n : Ax = 0\} \subseteq \mathbb{K}^n \tag{4.4.5}$$

**Bsp.**  $F : \mathbb{K}^2 \rightarrow K^2$  mit  $F(x_1, x_2) = (x_2, x_2)$ .

$$\text{im}(F) = \{(t, t) : t \in \mathbb{K}\}$$

$$\ker(F) = \{(t, 0) : t \in \mathbb{K}\} = \mathbb{K} \times \{0\}$$

$$F^{-1}(w) = \begin{cases} \emptyset & \text{falls } w \notin \text{im}(F) \\ \mathbb{K} \times \{t\} & \text{falls } w \in \text{im}(F), \text{ d.h. } w = (t, t) \text{ mit } t \in \mathbb{K} \end{cases}$$

#### 4.4.2 Beschreibung der Injektivität und Surjektivität durch das Bild und den Kern

**Prop.** Seien  $V$  und  $W$  Vektorräume über  $\mathbb{K}$  und sei  $F : V \rightarrow W$  linear.  
Dann gilt:

(a)  $F$  ist surjektiv  $\Leftrightarrow \text{im}(F) = W$ .

(b)  $F$  ist injektiv  $\Leftrightarrow \ker(F) = \{0\}$ .

- (c) Ist  $F$  injektiv und sind  $v_1, \dots, v_n \in V$  ( $n \in \mathbb{N}_0$ ) linear unabhängig, so sind auch  $F(v_1), \dots, F(v_n)$  linear unabhängig.

*Beweis.* (a) ist klar.

- (b) Wenn  $F$  injektiv ist, dann folgt offensichtlich  $\ker(F) = \{0\}$ . Umgekehrt: sei  $\ker(F) = \{0\}$  und seien  $u, v \in V$  mit  $F(u) = F(v)$ . Dann gilt  $F(u) - F(v) = F(u - v) = 0$ , d.h.  $u - v \in \ker(F)$  und somit  $u - v = 0$  (also  $u = v$ ).
- (c) Seien  $\alpha_1, \dots, \alpha_n \in \mathbb{K}$  mit  $\alpha_1 F(v_1) + \dots + \alpha_n F(v_n) = 0$ . Dann folgt  $F(\alpha_1 v_1 + \dots + \alpha_n v_n) = 0$ . Da  $F$  injektiv ist, folgt  $\alpha_1 v_1 + \dots + \alpha_n v_n = 0$ . Aus der linearen Unabhängigkeit von  $v_1, \dots, v_n$  folgt  $\alpha_1 = \dots = \alpha_n = 0$ .  $\square$

### 4.4.3 Rang einer linearen Abbildung

Seien  $V$  und  $W$  Vektorräume über  $\mathbb{K}$  und sei  $F : V \rightarrow W$  linear. Dann heißt

$$\text{rang}(F) := \dim(\text{im}(F)) \quad (4.4.6)$$

der *Rang* von  $F$ .

**Bem.** Wenn  $A \in \mathbb{K}^{m \times n}$  ( $m, n \in \mathbb{N}$ ) und  $F : \mathbb{K}^n \rightarrow \mathbb{K}^m$  mit  $F(x) = Ax \forall x \in \mathbb{K}^n$ , dann gilt  $\text{rang}(A) = \text{rang}(F)$ .

*Beweis.*  $F(e_1), \dots, F(e_n)$  sind die  $n$  Spalten von  $A$ . Somit

$$\text{rang}(A) = \dim(\text{lin}(F(e_1), \dots, F(e_n))), \quad \text{und}$$

$$\begin{aligned} \text{lin}(F(e_1), \dots, F(e_n)) &= \{\alpha_1 F(e_1) + \dots + \alpha_n F(e_n) : \alpha_1, \dots, \alpha_n \in \mathbb{K}\} \\ &= \{F(\alpha_1 e_1 + \dots + \alpha_n e_n) : \alpha_1, \dots, \alpha_n \in \mathbb{K}\} \\ &= \{F(x) : x \in \mathbb{K}^n\} \\ &= \text{im}(F). \end{aligned}$$

Somit erhält man die Gleichheit von  $\text{rang}(A)$  und  $\text{rang}(F)$ .  $\square$

#### 4.4.4 Nichtleere Fasern sind Verschiebungen vom Kern

Sei  $V$  Vektorraum über  $\mathbb{K}$ , sei  $X \subseteq V$  und sei  $a \in V$ . Dann heißt

$$a + X := X + a := \{a + x : x \in X\} \quad (4.4.7)$$

die *Verschiebung* von  $X$  um den Vektor  $a \in V$ .

**Prop.** Seien  $V$  und  $W$  Vektorräume über  $\mathbb{K}$ , sei  $F : V \rightarrow W$  linear, seien  $u \in V$  und  $w = F(u) \in W$ . Dann gilt  $F^{-1}(w) = u + \ker(F)$ .

*Beweis.* Sei  $v \in F^{-1}(w)$ , d.h.  $F(v) = w$ . Dann gilt  $v = u + (v - u)$  mit  $F(v - u) = F(v) - F(u) = w - w = 0$ , d.h.  $v - u \in \ker(F)$ . Das zeigt  $F^{-1}(w) \subseteq u + \ker(F)$ .

Umgekehrt: sei  $v \in \ker(F)$ , d.h.  $F(v) = 0$ . Dann gilt für  $u+v$ :  $F(u+v) = F(u) + F(v) = w + 0 = w$ , d.h.  $u+v \in F^{-1}(w)$ . Das ergibt  $u + \ker(F) \subseteq F^{-1}(w)$ .  $\square$

**Kor.** Seien  $A \in \mathbb{K}^{m \times n}$  und  $b \in \mathbb{K}^m$  ( $m, n \in \mathbb{N}$ ). Sei  $X = \{x \in \mathbb{K}^n : Ax = b\}$  nicht leer. Sei  $x^* \in X$ . Dann gilt  $X = x^* + X_0$  mit  $X_0 = \{x \in \mathbb{K}^n : Ax = 0\} = \ker(A)$ .

*Beweis.* Umformulierung der vorigen Proposition mit Matrizen.  $\square$

#### 4.4.5 Affine Unterräume

Eine Teilmenge  $X$  eines Vektorraums  $V$  über  $\mathbb{K}$  nennt man *affiner Unterraum* von  $V$ , falls  $X = \emptyset$  oder  $X$  eine Verschiebung eines Untervektorräums von  $V$  ist (d.h.  $X = a + U$  für ein  $a \in V$  und einen Untervektorraum  $U$  von  $V$ ).

Wir nennen die Menge  $X - X := x - x' : x, x' \in X$  die *Differenzenmenge* von  $X$  (die Menge der Differenzen).

**Prop.** Sei  $V$  Vektorraum über  $\mathbb{K}$  und sei  $X$  nichtleerer affiner Unterraum von  $V$ . Dann ist  $X - X$  Untervektorraum von  $V$  und es gilt  $X - p = X - X$

für alle  $p \in X$ .

*Beweis.*  $X = a + U$  für ein  $a \in V$  und einen Untervektorraum  $U$  von  $V$ . Das ergibt  $X - a = U$  und  $a \in X$ . Somit hat man  $X - X = (a + U) - (a + U) = U - U = U$ .

Sei  $p \in X$  beliebig. Dann gilt

$$X - p = U + \underbrace{a - p}_{\in X - X = U} = U. \quad \square$$

Aus der Proposition folgt, dass ein nichtleerer affiner Unterraum  $X$  von  $V$  durch einen eindeutigen Untervektorraum  $U$  definiert ist, und dass *jede* Verschiebung von  $X$ , die 0 enthält, mit  $U$  übereinstimmt.

Wir definieren die Dimension eines affinen Unterraums  $X$  durch

$$\dim_{\mathbb{K}}(X) := \begin{cases} -1 & \text{falls } X = \emptyset \\ \dim_{\mathbb{K}}(X - X) & \text{falls } X \neq \emptyset \end{cases} \quad (4.4.8)$$

#### 4.4.6 Der Rangsatz

**Thm.** Seien  $V$  und  $W$  Vektorräume über  $\mathbb{K}$  mit  $\dim(V) < \infty$ . Sei  $F : V \rightarrow W$  linear. Dann gilt

$$\dim(V) = \dim(\text{im}(F)) + \dim(\ker(F)). \quad (4.4.9)$$

*Beweis.* Sei  $v_1, \dots, v_k$  Basis von  $\ker(F)$  mit  $k \in \mathbb{N}_0$ . Sei  $w_1, \dots, w_r$  Basis von  $\text{im}(F)$  mit  $r \in \mathbb{N}_0$ . Sei  $u_1 \in F^{-1}(w_1), \dots, u_r \in F^{-1}(w_r)$ , d.h.  $F(u_i) = w_i$  für  $i \in \{r, \dots, k\}$ . Wir zeigen, dass  $u_1, \dots, u_r, v_1, \dots, v_k$  eine Basis von  $V$  ist.

Sei  $v \in V$  beliebig. Dann gilt  $F(v) = \mu_1 w_1 + \dots + \mu_r w_r$  für gewisse  $\mu_1, \dots, \mu_r \in \mathbb{K}$ . Wir setzen  $v' := \mu_1 u_1 + \dots + \mu_r u_r$ . Es gilt  $F(v') = F(v)$  und somit  $F(v - v') = 0$ . D.h.  $v - v' \in \ker(F)$ .

Es existieren  $\lambda_1, \dots, \lambda_k \in \mathbb{K}$  mit  $v - v' = \lambda_1 v_1 + \dots + \lambda_k v_k$ . Es folgt:

$$\begin{aligned} v &= v' + \lambda_1 v_1 + \dots + \lambda_k v_k \\ &= \mu_1 u_1 + \dots + \mu_r u_r + \lambda_1 v_1 + \dots + \lambda_k v_k. \end{aligned}$$

Es bleibt die lineare Unabhängigkeit von  $u_1, \dots, u_r, v_1, \dots, v_k$  zu zeigen.

Seien  $\mu_1, \dots, \mu_r, \lambda_1, \dots, \lambda_k \in \mathbb{K}$  und

$$\mu_1 u_1 + \dots + \mu_r u_r + \lambda_1 v_1 + \dots + \lambda_k v_k = 0 \quad (4.4.10)$$

Wir wenden  $F$  zur linken und rechten Seite an und erhalten

$$\underbrace{\mu_1 F(u_1)}_{w_1} + \dots + \underbrace{\mu_r F(u_r)}_{w_r} + \underbrace{\lambda_1 F(v_1)}_0 + \dots + \underbrace{\lambda_k F(v_k)}_0 = 0,$$

d.h.  $\mu_1 w_1 + \mu_r w_r = 0$ . Es folgt  $\mu_1 = \dots = \mu_r = 0$ . Wir setzen  $\mu_i = 0$  ( $i \in \{r, \dots\}$ ) in (4.4.10) ein und erhalten  $\lambda_1 v_1 + \dots + \lambda_k v_k = 0$ . Die lineare Unabhängigkeit von  $v_1, \dots, v_k$  ergibt  $\lambda_1 = \dots = \lambda_k = 0$ .  $\square$

#### 4.4.7 Die Dimension der Faser

**Kor.** Seien  $V$  und  $W$  Vektorräume über  $\mathbb{K}$  mit  $\dim(V) < \infty$ . Sei  $F : V \rightarrow W$  linear. Sei  $w \in W$  und sei  $F^{-1}(w) \neq \emptyset$ . Dann gilt

$$\dim(F^{-1}(w)) = \dim(V) - \dim(\text{im}(F)). \quad (4.4.11)$$

*Beweis.* Aufgabe. □

#### 4.4.8 Klassifikation endlichdimensionaler Vektorräume

**Kor.** Seien  $V$  und  $W$  endlichdimensionale Vektorräume über  $\mathbb{K}$ . Eine bijektive Abbildung von  $V$  nach  $W$  existiert genau dann, wenn  $\dim(V) = \dim(W)$  gilt.

*Beweis.* Aufgabe. □

#### 4.4.9 Injektivität und Surjektivität linearer Abbildungen eines endlichdimensionalen Vektorraums

**Thm.** Seien  $V$  und  $W$  endlichdimensionale Vektorräume über  $\mathbb{K}$  mit  $n := \dim(V) = \dim(W)$  und  $F : V \rightarrow W$  linear. Dann sind die folgenden Bedingungen äquivalent:

(i)  $F$  ist injektiv.

(ii)  $F$  ist surjektiv.

(iii)  $F$  ist bijektiv.

(iv)  $\text{rang}(F) = n$ .

*Beweis.* (i)  $\Rightarrow$  (iv): Angenommen  $F$  ist injektiv. Dann ist  $\ker(F) = \{0\}$

und daher  $\dim(\ker(F)) = 0$ . Nach dem Rangsatz ist

$$\begin{aligned}\operatorname{rang}(F) &= \dim(\operatorname{im}(F)) \\ &= \dim(V) - \dim(\ker(F)) \\ &= n - n \\ &= 0.\end{aligned}$$

(iv)  $\Rightarrow$  (ii):  $\operatorname{rang}(F) = n \Rightarrow \dim(F(V)) = \dim(W) \Rightarrow F(V) = W$ . Dann ist  $F$  surjektiv.

(ii)  $\Rightarrow$  (iii): Angenommen  $F$  ist surjektiv, d.h.  $\operatorname{im}(F) = W$ .  $\Rightarrow \dim(\operatorname{im}(F)) = \dim(W) = n \Rightarrow \dim(\ker(F)) = \dim(V) - \dim(\operatorname{im}(F)) = n - n = 0 \Rightarrow \ker(F) = \{0\} \Rightarrow F$  ist injektiv  $\Rightarrow F$  bijektiv.

(iii)  $\Rightarrow$  (i) ist trivial.  $\square$

**Kor.** Sei  $A \in \mathbb{K}^{n \times n}$  ( $n \in \mathbb{N}$ ).  $A$  ist genau dann invertierbar, wenn  $\operatorname{rang}(A) = n$  gilt.

*Beweis.* Sei  $F : \mathbb{K}^n \rightarrow \mathbb{K}^n$  mit  $F(x) = Ax$  für alle  $x \in \mathbb{K}^n$ .

$$\begin{aligned} & \text{rang}(A) = n \\ \Leftrightarrow & F \text{ ist surjektiv} \\ \Leftrightarrow & F \text{ ist bijektiv} \\ \Leftrightarrow & A \text{ ist invertierbar.} \end{aligned} \quad \square$$

#### 4.4.10 Verfahren zur Invertierung von Matrizen

Wir wollen entscheiden, ob eine gegebene Matrix  $A \in \mathbb{K}^{n \times n}$  ( $n \in \mathbb{N}$ ) invertierbar ist, und gegebenenfalls  $A^{-1}$  berechnen.

Ist  $A$  invertierbar, so hat das System  $Ax = y$  mit den Vektoren der Unbekannten  $x$  für jede Wahl von  $y \in \mathbb{K}^n$  eine eindeutige Lösung. Diese Lösung ist  $x = A^{-1}Ax = A^{-1}y$ .

Bei  $Ax = y$  “sagt uns” die Matrix  $A$  für das gegebene  $x \in \mathbb{K}^n$ , was  $y \in \mathbb{K}^n$  ist. Die Matrix  $A^{-1}$  (falls vorhanden) “sagt uns” für das gegebene  $y$ ,

was das eindeutige  $x$  dazu ist. Die Komponenten von  $x$  und  $y$  können wir als zwei Gruppen mit je  $n$  Variablen auffassen, die miteinander durch drei Gleichungen verlinkt sind (denn  $Ax = y$  ist ein LGS aus  $n$  Gleichungen). Mit dem Gauß-Jordan-Verfahren können wir versuchen, das System  $Ax = y$  nach  $x$  auflösen.

Wenn das Gauß-Jordan-Verfahren es schafft, dieses System nach den  $x$ -Variablen aufzulösen, ist unsere Matrix invertierbar. Gelingt das Auflösen nicht, so findet man heraus, für welche Wahl der Werte für  $y$ -Variablen man keine dazu passenden Werte für  $x$ -Variablen findet. Im System  $Ax = y$  sind die rechten Seiten der Gleichungen von einem Vektor  $y$  aus  $n$  Variablen abhängig. Im Gegensatz zur Lösung von  $Ax = b$  für ein festes  $b$  sind die rechten Seiten bei der Bearbeitung von  $Ax = y$  mit einem variablen Vektor  $y$  im Rahmen des Gauß-Jordan-Verfahrens keine Konstanten sondern lineare Funktionen in  $y$ . Der Verlauf des Gauß-Verfahrens ist aber völlig analog.

**Bsp.** Ist  $A = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 1 & 0 & 2 \end{pmatrix} \in \mathbb{R}^{3 \times 3}$  invertierbar? Ggf., was ist  $A^{-1}$ ?  $Ax = y$

ist das LGS

$$\left\{ \begin{array}{lcl} 2x_1 + x_2 & = & y_1 \\ 2x_2 & = & y_2 \\ x_1 + 2x_3 & = & y_3 \end{array} \right.$$

das wir mit Gauß-Jordan nach  $x_1, x_2, x_3$  auflösen wollen:

	$x_1$	$x_2$	$x_3$	$y_1$	$y_2$	$y_3$	
$g_1$	2	1	0	1	0	0	$g_1 := \frac{1}{2} g_1$
$g_2$	0	2	0	0	1	0	$g_2 := \frac{1}{2} g_2$
$g_3$	1	0	2	0	0	1	
$g_1$	1	$\frac{1}{2}$	0	$\frac{1}{2}$	0	0	$g_1 := g_1 - \frac{1}{2} g_2$
$g_2$	0	1	0	0	$\frac{1}{2}$	0	
$g_3$	1	0	2	0	0	1	
$g_1$	1	0	0	$\frac{1}{2}$	$-\frac{1}{2}$	0	$g_3 := g_3 - g_1$
$g_2$	0	1	0	0	$\frac{1}{2}$	0	
$g_3$	1	0	2	$-\frac{1}{2}$	$\frac{1}{4}$	1	
$g_1$	1	0	0	$\frac{1}{2}$	$-\frac{1}{2}$	0	$g_3 := \frac{1}{2} g_3$
$g_2$	0	1	0	0	$\frac{1}{2}$	0	
$g_3$	0	0	2	$-\frac{1}{2}$	$\frac{1}{4}$	1	
$g_1$	1	0	0	$\frac{1}{2}$	$-\frac{1}{2}$	0	
$g_2$	0	1	0	0	$\frac{1}{2}$	0	
$g_3$	0	0	1	$-\frac{1}{4}$	$\frac{1}{8}$	$\frac{1}{2}$	

Das Gauß-Jordan-Verfahren sagt uns also:

$$\begin{cases} x_1 = & \frac{1}{2}y_1 - \frac{1}{2}y_2 \\ x_2 = & \frac{1}{2}y_2 \\ x_3 = & -\frac{1}{4}y_1 + \frac{1}{8}y_2 + \frac{1}{2}y_3. \end{cases}$$

Oder in der Matrix-Form:

$$\underbrace{\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}}_x = \underbrace{\begin{pmatrix} \frac{1}{2} & -\frac{1}{4} & 0 \\ 0 & \frac{1}{2} & 0 \\ -\frac{1}{4} & \frac{1}{8} & \frac{1}{2} \end{pmatrix}}_{A^{-1}} \underbrace{\begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}}_y$$

**Bsp.** Ist die Matrix

$$A = \begin{pmatrix} -2 & 1 & 1 \\ 1 & -2 & 1 \\ 1 & 1 & -2 \end{pmatrix} \in \mathbb{R}^{3 \times 3}$$

invertierbar? Wir verwenden Gauß-Jordan-Verfahren um  $Ax = y$  für ein Variables  $y$  nach  $x$  auflösen. Die Gleichung  $Ax = y$  ist

$$\left\{ \begin{array}{l} -2x_1 + x_2 + x_3 = y_1 \\ x_1 + -2x_2 + x_3 = y_2 \\ x_1 + x_2 + -2x_3 = y_3 \end{array} \right.$$

Wenn wir es schaffen, durch das Gauß-Jordan-Verfahren dieses System nach den  $x$ -Variablen aufzulösen, ist unsere Matrix invertierbar. Gelingt das Auflösen nicht, so finden wir heraus, für welche Wahl der Werte für  $y$ -Variablen man keine dazu passenden Werte für  $x$ -Variablen findet. Im System  $Ax = y$  ist sind die rechten Seiten der Gleichungen von  $y$  abhängig. Beim Transformieren bleiben die rechten Seiten lineare Funktionen in  $y$ :

	$x_1$	$x_2$	$x_3$	$y_1$	$y_2$	$y_3$	
$g_1$	-2	1	1	1	0	0	$g_1 := g_1 + 2g_3$
$g_2$	1	-2	1	0	1	0	$g_2 := g_2 - g_3$
$g_3$	1	1	-2	0	0	1	
$g_1$	0	3	-3	1	0	2	$g_1 := g_1 + g_2$
$g_2$	0	-3	3	0	1	-1	
$g_3$	1	1	-2	0	0	1	
$g_1$	0	0	0	1	1	1	
$g_2$	0	-3	3	0	1	-1	
$g_3$	1	1	-2	0	0	1	

An dieser Stelle können wir nun das Gauß-Jordan-Verfahren abbrechen, da man bereits sieht, dass  $A$  nicht invertierbar ist. Die erste Gleichung im letzten erhaltenen System ist  $0 = y_1 + y_2 + y_3$ . Man findet also nicht zu jedem  $y \in \mathbb{R}^3$  ein  $x$  mit  $Ax = y$ . Wenn  $y_1 + y_2 + 3 \neq 0$ , zum Beispiel, für  $y_1 = 1, y_2 = 0, y_3 = 0$ , dann gibt es kein  $x$  mit  $Ax = y$ .

#### 4.4.11 Rang der Komposition von linearen Abbildungen bzw. des Matrixprodukts

**Thm.** Seien  $U, V$  und  $W$  endlichdimensionale Vektorräume über  $\mathbb{K}$ . Seien  $F : V \rightarrow W$  und  $G : U \rightarrow V$  linear. Dann gilt:

$$\text{rang}(F) + \text{rang}(G) - \dim(V) \leq \text{rang}(F \circ G) \leq \min\{\text{rang}(F), \text{rang}(G)\} \quad (4.4.12)$$

*Beweis.* Es gilt  $G(U) \subseteq V$ , und somit auch  $F(G(U)) \subseteq F(V)$ . Also ist

$$\text{rang}(F \circ G) \leq \text{rang}(F) \Leftrightarrow \dim(F(G(U))) \leq \dim(F(V)).$$

Da durch die Anwendung von  $F$  zu  $G(U)$  ein Untervektorraum entsteht, dessen Dimension höchstens die Dimension von  $G(U)$  sein kann, gilt

$$\text{rang}(F \circ G) \leq \text{rang}(G) \Leftrightarrow \dim(F(G(U))) \leq \dim(G(U)).$$

Außerdem ist

$$\begin{aligned}
 & \text{rang}(F) + \text{rang}(G) - \dim(V) \leq \text{rang}(F \circ G) \\
 \Leftrightarrow & \dim(F(V)) + \dim(G(U)) - \dim(V) \leq \dim(F(G(U))) \\
 \Leftrightarrow & \underbrace{\dim(G(U)) - \dim(F(G(U)))}_{\dim(\ker(F|_{G(U)}))} \leq \underbrace{\dim(V) - \dim(F(V))}_{\dim(\ker(F))} \\
 \Leftrightarrow & \dim(\{x \in G(U) : F(x) = 0\}) \leq \dim(\{x \in V : F(x) = 0\}),
 \end{aligned}$$

weil aus  $G(U) \subseteq V$  die Inklusion  $\ker(F|_{G(U)}) \subseteq \ker(F)$  und somit  $\dim(\ker(F|_{G(U)})) \leq \dim(\ker(F))$  folgt.  $\square$

**Bem** (Ergänzung zu mathematischen Grundlagen). Seien  $X, Y$  Mengen. Sei  $A \subseteq X$  und  $f : X \rightarrow Y$ . Die Abbildung  $f|_A : A \rightarrow Y$  mit  $(f|_A)(x) := f(x) \quad \forall x \in A$  heißt die *Einschränkung* von  $f$  auf  $A$ . (Gesprochen „ $f$  eingeschränkt auf  $A$ “).

**Kor.** Seien  $m, n, k \in \mathbb{N}$  und seien  $A \in \mathbb{K}^{m \times n}, B \in \mathbb{K}^{n \times k}$ . Dann gilt:

- (a)  $\text{rang}(A) + \text{rang}(B) - n \leq \text{rang}(AB) \leq \min\{\text{rang}(A), \text{rang}(B)\}$
- (b) Wenn  $m = n$  gilt und  $A$  invertierbar ist, dann gilt  $\text{rang}(AB) = \text{rang}(B)$ .
- (c) Wenn  $n = k$  und  $B$  invertierbar ist, dann gilt  $\text{rang}(AB) = \text{rang}(A)$ .

*Beweis.* (a) folgt direkt aus dem vorigen Theorem.

- (b) Ist  $m = n$  und  $A$  invertierbar, so gilt  $\text{rang}(A) = n$  und somit  $\text{rang}(A) + \text{rang}(B) - n = \text{rang}(B)$ . Es gilt  $\text{rang}(B) \leq k$  (aus der Definition) und  $\text{rang}(B) \leq n$  (wegen  $\text{rang}(B) = \text{rang}(B^\top) \leq n$ ). Somit gilt  $\min\{\text{rang}(A), \text{rang}(B)\} = \text{rang}(B)$ .
- (c) Der Beweis von (c) ist analog zum Beweis von (b). □

#### 4.4.12 Rang und Lösbarkeit von linearen Gleichungssystemen

Für ein LGS  $Ax = b$  mit  $n$  Unbekannten und  $m$  Gleichungen. Nennt man  $A$  die *Matrix* des Systems. Seien  $a_1, \dots, a_n$  die Spalten von  $A$ . Man nennt die Matrix

$$(A | b) := \begin{pmatrix} | & | & | \\ a_1 & \cdots & a_n & b \\ | & | & | \end{pmatrix}$$

die *erweiterte Matrix* des Systems  $Ax = b$ . Das Gauß-Verfahren operiert auf der erweiterten Matrix.

**Thm.** Seien  $m, n \in \mathbb{N}$ , sei  $A \in \mathbb{K}^{m \times n}$  und  $b \in \mathbb{K}^m$ . Dann gilt:

- (a)  $\text{rang}(A) \leq \text{rang}(A | b) \leq \text{rang}(A) + 1$
- (b) Das System  $Ax = b$  besitzt genau dann eine Lösung  $x$ , wenn  $\text{rang}(A | b) = \text{rang}(A)$  gilt.

- (c) Das System  $Ax = b$  besitzt genau dann eine eindeutige Lösung, wenn  $n = \text{rang}(A) = \text{rang}(A | b)$  gilt.

*Beweis.* Seien  $a_1, \dots, a_n$  Spalten von  $A$ .

- (a)  $\text{rang}(A) \leq \text{rang}(A | b)$  gilt, weil alle Spalten von  $A$  auch Spalten von  $(A | b)$  sind.  $\text{rang}(A | b) \leq \text{rang}(A) + 1$  gilt, durch das Hinzufügen einer Spalte zur Matrix die Dimension des Spaltenraums um höchstens 1 wächst.
- (b) Die Lösbarkeit von  $Ax = b$  bedeutet  $b \in \text{lin}(a_1, \dots, a_n)$ . Es ist nicht schwer zu sehen, dass  $b \in \text{lin}(a_1, \dots, a_n)$  als  $\text{lin}(a_1, \dots, a_n) = \text{lin}(a_1, \dots, a_n, b)$  umformuliert werden kann.
- (c) Die Lösbarkeit von  $Ax = b$  ist nach b) äquivalent zu  $\text{rang}(A) = \text{rang}(A | b)$  ist. Für eindeutige Lösbarkeit muss  $\dim(\ker(A)) = 0$  sein, denn sonst hat man für eine Lösung  $x$  und  $h \in \ker(A) \setminus \{0\}$  eine wei-

tere Lösung  $x + h$ . Umgekehrt hat man zwei verschiedene Lösungen  $x$  und  $x'$ , so ist  $x - x' \in \ker(A) \setminus \{0\}$ .

Nach dem Rangsatz hat man  $\text{rang}(A) + \dim(\ker(A)) = n$  gilt. Die Bedingung  $\dim(\ker(A)) = 0$  ist also äquivalent zu  $\text{rang}(A) = n$ .

□

#### 4.4.13 Faktorisierungssatz

**Thm.** Seien  $V, W$  Vektorräume über einem Körper  $\mathbb{K}$  und sei  $\dim(V) < \infty$ . Ferner sei  $F : V \rightarrow W$  eine lineare Abbildung. Sei  $\mathcal{A} = (u_1, \dots, u_r, v_1, \dots, v_k)$  Basis von  $V$  mit  $\ker(F) = \text{lin}(v_1, \dots, v_k)$ , wobei  $r, k \in \mathbb{N}_0$ . Sei  $U := \text{lin}(u_1, \dots, u_r)$ . Dann gilt:

- (a)  $V = U \oplus \ker(F)$
- (b) Die Abbildung  $\tilde{F} : U \rightarrow \text{im}(F)$  mit  $\tilde{F}(u) = F(u) \forall u \in U$  ist eine lineare Bijektion.

- (c) Es existiert eine eindeutige lineare Abbildung  $P : V \rightarrow U$  mit  $P(u + v') = u \forall u \in U, v' \in \ker(F)$ .
- (d) Es gilt  $F(v) = \tilde{F}(P(v)) \forall v \in V$ .

*Beweis.* (a) Folgt direkt aus der Charakterisierung der direkten Summe.

- (b) Linearität von  $\tilde{F}$  klar, wegen Linearität von  $F$ .  $\tilde{F}$  ist surjektiv, denn  $\text{im}(F) = F(V) \stackrel{(a)}{=} F(U \oplus \ker(F)) = F(U) + F(\ker(F)) = F(U) = \tilde{F}(U)$ .  $\tilde{F}$  ist auch injektiv, denn aus  $F(u) = \tilde{F}(u)$  mit  $u, u' \in U$  folgt  $F(u - u') = F(u) - F(u') = \tilde{F}(u) - \tilde{F}(u') = 0$ . Also ist  $u - u' \in \ker(F)$  und da  $U \cap \ker(F) = \{0\}$ , folgt  $u = u'$ .
- (c) Existenz und Eindeutigkeit folgen aus der Darstellung  $V = U \oplus \ker(F)$ . Zur Linearität: seien  $u_1, u_2 \in U, v'_1, v'_2 \in \ker(F)$ . Dann ist  $P(u_1 + v'_1 + u_2 + v'_2) = P(u_1 + u_2 + v'_1 + v'_2) = u_1 + u_2$  sowie  $P(u_1 + u_2) + P(v'_1 + v'_2) = u_1 + u_2$ . Analog zeigt man  $P(\lambda(u + v')) = \lambda P(u + v')$ .

(d) Nach (a) ist  $v$  eindeutig als  $v = u + v'$  mit  $u \in U, v' \in \ker(F)$  darstellbar.  $F(u + v') = F(u) + F(v') = F(u)$ . Andererseits ist  $\tilde{F}(P(u + v')) = \tilde{F}(u) = F(u)$ .  $\square$

#### 4.4.14 Quotientenräume

Sei  $\mathbb{K}$  ein Körper,  $V$  ein Vektorraum und sei  $U$  ein Untervektorraum von  $V$ . Für  $v, v' \in V$  schreiben wir  $v \equiv v' \pmod{U}$ , falls  $v - v' \in U$ . Die Äquivalenz modulo  $U$  ist tatsächlich eine Äquivalenzrelation:

- $v \equiv v$ , da  $v - v = 0 \in U$  ( $\forall v \in V$ )
- ist  $v \equiv v'$  also  $v - v' \in U$ , dann ist  $-(v - v') = v' - v \in U$ , also  $v' \equiv v$
- ist  $v \equiv v'$  und  $v' \equiv v''$  also  $v - v', v' - v'' \in U$ , dann ist  $(v - v') + (v' - v'') = v - v'' \in U$ , also  $v \equiv v''$

Für die Äquivalenzklasse  $[v] := [v]_U$  von  $v \in V$  gilt  $[v] = v + U$ . Die Menge aller Äquivalenzklassen wird mit  $V/U$  (angelehnt an  $\mathbb{Z}/m\mathbb{Z}$ ) bezeichnet.

**Thm.** *Sei  $\mathbb{K}$  ein Körper,  $V$  ein  $\mathbb{K}$ -Vektorraum und sei  $U$  ein Untervektorraum von  $V$ . Dann existiert eine eindeutige Addition und Skalarmultiplikation auf  $V/U$ , sodass  $V/U$  mit diesen Operationen ein Vektorraum und die kanonische Abbildung  $\rho : V \rightarrow V/U, v \mapsto [v] \forall v \in V$  eine lineare Abbildung ist. Mit diesen Operationen auf  $V/U$  gilt:*

- (a)  $\rho$  ist surjektiv
- (b)  $\ker(\rho) = U$
- (c)  $\dim(V/U) = \dim(V) - \dim(U)$

*Beweis.* Eindeutigkeit der Addition und Skalarmultiplikation: Damit  $\rho$  linear ist, muss  $[v + w] = [v] + [w] \quad \forall v, w \in V$  gelten. D.h. die Summe von  $[v]$  und  $[w]$  muss  $[v + w]$  sein. Außerdem muss für die Linearität von  $\rho$  auch

$[\lambda v] = \lambda[v] \quad \forall \lambda \in \mathbb{K}, v \in V$  gelten. D.h. das Skalarprodukt  $\lambda[v]$  muss als  $[\lambda v]$  definiert werden.

Existenz der Addition und Skalarmultiplikation: zu zeigen ist: gilt für  $v, v', w, w' \in V$ , dass  $[v] = [v']$  und  $[w] = [w']$ , so gilt  $[v + w] = [v' + w']$ . Das gilt, da  $v - v' \in U$  und  $w - w' \in U$  auch  $(v' + w') - (v + w) \in U$  impliziert. Ebenso ist dann auch  $\lambda v' - \lambda v \in U$  und somit  $[\lambda v] = [\lambda v']$ . Das zeigt die Existenz der Skalarmultiplikation.

- (a) folgt direkt aus der Definition der kanonischen Basis.
- (b) offensichtlich ist  $[0]$  der Nullvektor von  $V/U$ . Hat man nun  $v \in \ker(\rho)$ , so ist  $[v] = [0]$ , also  $v - 0 \in U$  und damit  $v \in U$ .
- (c) Wegen der Linearität von  $\rho$  gilt  $\dim(\text{im}(\rho)) + \dim(\ker(\rho)) = \dim(V)$ . Hier ist wegen (a)  $\dim(\text{im}(\rho)) = \dim(V/U)$ . Wegen (b) ist  $\dim(\ker(\rho)) = \dim(U)$ . Zusammengenommen folgt  $\dim(V/U) + \dim(U) = \dim(V)$ .

□

**Bsp.** Stellen wir uns vor, die Positionen von drei Rennautos auf einer Strecke werden durch einen Vektor  $(p_1, p_2, p_3) \in \mathbb{R}^3$  beschrieben:  $p_i$  ist die Position vom  $i$ -ten Rennauto (die Strecke ist als die reelle Achse  $\mathbb{R}$  modelliert). Wir betrachten nun im Vektorraum  $V = \mathbb{R}^3$  den Untervektorraum  $U = \{(t, t, t) : t \in \mathbb{R}^3\}$ . Welche Informationen beschreibt uns der Vektorraum  $V/U$ ? Mit anderen Worten: was heißt es, dass man den Vektor  $(p_1, p_2, p_3)$  modulo  $U$  kennt. Wenn etwa  $(p_1, p_2, p_3)$  kongruent zu  $(5, -3, 2)$  modulo  $U$  ist, dann heißt es dass man  $p_1 = t + 5, p_2 = t - 3, p_3 = t + 2$  für ein  $t \in \mathbb{R}$  gilt. Somit kennen wir zwar nicht die genaue Position der drei Rennautos auf der Strecke, wir kennen aber die relative Position der Rennautos zueinander. Das erste Rennauto ist ganz vorne, dann kommt das dritte Auto, dass  $5 - 2 = 3$  drei Einheiten zurückliegt, und das zweite Auto liegt  $2 - (-3) = 5$  Einheiten hinter dem ersten Rennauto. Wir sehen also durch die Angabe von  $(p_1, p_2, p_3)$  modulo  $U$  die drei Rennautos relativ zueinander, in welchem Teil der Strecke sie sich befinden.

$$p_2 \quad p_3 = p_2 + 5 \quad p_1 = p_3 + 3$$

Wo ist die 0???

Trotz der Tatsache, dass die Positionen  $(p_1, p_2, p_3)$  modulo  $U$  keine exakten Angaben der Werte von  $p_1, p_2, p_3$  sind, kann man modulo  $U$  rechnen. Hat zum Beispiel nach einer Weile, das erste Auto 1 Einheit, das zweite Auto 2 Einheiten und das dritte Auto 4 Einheiten zurückgelegt, so ergibt sich der Vektor der neuen Positionen  $(5 + 1, -3 + 2, 2 + 4) = (6, -1, 6)$  modulo  $U$ : das dritte Auto hat das erste Auto eingeholt, das zweite liegt 7 Einheiten zurück. Nach wie vor kennt man aber die durch die Angabe der Positionen modulo  $U$  keine Information über die genauen Positionen auf der Strecke.

**Bsp** (Ein weiterführendes Beispiel für Quotientenräume). Analog zu Polynomringen mit einer Unbestimmten kann man auch Polynomringe mit

mehr als einer Unbestimmten einführen. Wir betrachten etwa  $\mathbb{R}[x, y]$  – den Polynomring der Polynome in Unbestimmten  $x$  und  $y$  mit Koeffizienten in  $\mathbb{R}$ :

$$1, \quad 1 - 2xy + y^2, \quad xy + 3x^2 + x^4y^5 \quad \in \mathbb{R}[x, y]$$

$\mathbb{R}[x, y]$  ist ein Vektorraum über  $\mathbb{R}$ . (Rechnen auf dem Kreis:)  $U := \{(x^2 + y^2 - 1) \cdot f(x, y) : f(x, y) \in \mathbb{R}[x, y]\}$  ist ein Untervektorraum von  $\mathbb{R}[x, y]$ .  $\mathbb{R}[x, y]/U$  entspricht einer Struktur in der mit Berücksichtigung der Bedingung  $x^2 + y^2 = 1$  gerechnet wird.

Etwa werden  $x+y$  und  $x+y+(x^2+y^2-1)\cdot 2$  gleich, wenn man  $x^2+y^2 = 1$  fordert. Genauer gilt:

$$x + y \equiv x + y + (x^2 + y^2 - 1) \cdot 2 \mod U$$

## 4.5 Koordinatensysteme

### 4.5.1 Basisdarstellung von Vektoren

Sei  $V$  endlichdimensionaler Vektorraum über  $\mathbb{K}$  und sei  $n := \dim(V)$ . Sei  $\mathcal{B} = (b_1, \dots, b_n)$  eine Basis von  $V$  und  $x \in V$ . Der Vektor  $(\beta_1, \dots, \beta_n) \in \mathbb{K}^n$  mit  $x = \beta_1 b_1 + \dots + \beta_n b_n$  heißt der Vektor der Koordinaten von  $x$  in der Basis  $\mathcal{B}$ . Bezeichnung:  $x_{\mathcal{B}}$ . Die Abbildung  $x \mapsto x_{\mathcal{B}}$  von  $V$  nach  $\mathbb{K}^n$  ist eine lineare Bijektion.

**Bsp** (RoboCup Soccer). Inhalt...

### 4.5.2 Basiswechsel

Sei  $V$  endlichdimensionaler Vektorraum über  $\mathbb{K}$  mit  $n := \dim(V)$ . Seien  $\mathcal{A} = (a_1, \dots, a_n)$  und  $\mathcal{B} = (b_1, \dots, b_n)$  Basen von  $V$ . Wir führen die Matrix

$$T_{\mathcal{B} \leftarrow \mathcal{A}} := \begin{pmatrix} | & & | \\ (a_1)_{\mathcal{B}} & \cdots & (a_n)_{\mathcal{B}} \\ | & & | \end{pmatrix} \quad (4.5.1)$$

ein, deren  $j$ -te Spalte mit  $(a_j)_{\mathcal{B}}$  übereinstimmt ( $\forall j \in \{1, \dots, n\}$ ). Das heißt, bei der Angabe durch die Komponenten

$$T_{\mathcal{B} \leftarrow \mathcal{A}} = \begin{pmatrix} \tau_{11} & \cdots & \tau_{1n} \\ \vdots & & \vdots \\ \tau_{n1} & \cdots & \tau_{nn} \end{pmatrix} \quad (4.5.2)$$

erfüllen die Komponenten die Gleichungen

$$a_j = \tau_{1j}b_1 + \dots + \tau_{nj}b_n \quad \forall i \in \{1, \dots, n\}. \quad (4.5.3)$$

für  $j \in \{1, \dots, n\}$ .

Die Matrix  $T_{\mathcal{B} \leftarrow \mathcal{A}}$  heißt *Basiswechselmatrix* oder *Matrix des Wechsels* von der Basis  $\mathcal{A}$  zur Basis  $\mathcal{B}$ . Hier ist  $\mathcal{A}$  die “alte Basis” und  $\mathcal{B}$  die “neue Basis”. Durch die Matrix  $T_{\mathcal{B} \leftarrow \mathcal{A}}$  werden also die Vektoren der alten Basis in der neuen Basis geschrieben. Diese Beschreibungen helfen beim Übergang von der alten zur neuen Basis:

**Thm.** *Sei  $V$  endlichdimensionaler Vektorraum über  $\mathbb{K}$ , seien  $\mathcal{A}$  und  $\mathcal{B}$  Basen von  $V$  und sei  $x \in V$ . Dann gilt:*

$$x_{\mathcal{B}} = T_{\mathcal{B} \leftarrow \mathcal{A}} x_{\mathcal{A}}. \quad (4.5.4)$$

*Beweis.* Seien  $x_{\mathcal{A}} = (\alpha_1, \dots, \alpha_n)$ ,  $x_{\mathcal{B}} = (\beta_1, \dots, \beta_n)$ ,  $T_{\mathcal{B} \leftarrow \mathcal{A}} = (\tau_{ij})_{i,j=1}^n$ . D.h.

$$\begin{aligned} x &= \alpha_1 a_1 + \dots + \alpha_n a_n \\ &= \beta_1 b_1 + \dots + \beta_n b_n \\ a_j &= \tau_{1j} b_1 + \dots + \tau_{nj} b_n \quad \forall j. \end{aligned}$$

Es folgt:

$$\begin{aligned}
x &= \sum_{j=1}^n \alpha_j a_j \\
&= \sum_{j=1}^n \alpha_j \left( \sum_{i=1}^n \tau_{ij} b_i \right) \\
&= \sum_{i,j \in \{1, \dots, n\}} \tau_{ij} \alpha_j b_i \\
&= \sum_{i=1}^n \left( \sum_{j=1}^n \tau_{ij} \alpha_j \right) b_i.
\end{aligned}$$

Andererseits gilt  $x = \sum_{i=1}^n \beta_i b_i$ . Es folgt  $\beta_i = \sum_{j=1}^n \tau_{ij} \alpha_j$  für alle  $i \in \{1, \dots, n\}$ . Mit anderen Worten gilt  $x_B = T_{B \leftarrow A} x_A$ .  $\square$

### 4.5.3 Wiederholter Basiswechsel

**Thm.** Sei  $V$  endlichdimensionaler Vektorraum über  $\mathbb{K}$  und seien  $\mathcal{A}, \mathcal{B}$  und  $\mathcal{C}$  Basen von  $V$ . Dann gilt:

$$T_{C \leftarrow A} = T_{C \leftarrow B} T_{B \leftarrow A}. \quad (4.5.5)$$

*Beweis.* Übung. □

**Kor.** Sei  $V$  endlichdimensionaler Vektorraum über  $\mathbb{K}$  und seien  $\mathcal{A}$  und  $\mathcal{B}$  Basen von  $V$ . Dann ist  $T_{B \leftarrow A}$  invertierbar und es gilt:

$$(T_{B \leftarrow A})^{-1} = T_{A \leftarrow B}. \quad (4.5.6)$$

*Beweis.*  $T_{A \leftarrow A}$  ist die Einheitsmatrix. Die Behauptung folgt aus dem vorigen Theorem mit  $\mathcal{C} = \mathcal{A}$ . □

#### 4.5.4 Basendarstellung von linearen Abbildungen

Seien  $V$  und  $W$  endlichdimensionale Vektorräume über  $\mathbb{K}$ . Sei  $m := \dim(W)$  und  $n := \dim(V)$ . Sei  $\mathcal{A} = (a_1, \dots, a_m)$  Basis von  $W$  und  $\mathcal{B} = (b_1, \dots, b_n)$  Basis von  $V$ . Sei  $F : V \rightarrow W$  linear.

Wir definieren die Matrix  $F_{\mathcal{A}, \mathcal{B}} = (\phi_{ij})_{i=1, j=1}^{m, n} \in \mathbb{K}^{m \times n}$ , deren  $j$ -te Spalte mit  $F(b_j)_{\mathcal{A}}$  übereinstimmt (für alle  $j \in \{1, \dots, n\}$ ). D.h.

$$F_{\mathcal{A}, \mathcal{B}} = \begin{pmatrix} \phi_{11} & \cdots & \phi_{1n} \\ & \vdots & \\ \phi_{m1} & \cdots & \phi_{mn} \end{pmatrix} \quad (4.5.7)$$

und

$$F(b_j) = \phi_{1j}a_1 + \dots + \phi_{mj}a_m \quad \forall j \in \{1, \dots, n\}. \quad (4.5.8)$$

**Thm.** *Seien  $V$  und  $W$  endlichdimensionale Vektorräume über  $\mathbb{K}$ , sei  $\mathcal{A}$  Basis von  $W$  und  $\mathcal{B}$  Basis von  $V$ . Sei  $F : V \rightarrow W$  linear. Sei  $x \in V$  und*

$y = F(x)$ . Dann gilt:

$$y_{\mathcal{A}} = F_{\mathcal{A}, \mathcal{B}} x_{\mathcal{B}}. \quad (4.5.9)$$

*Beweis.* Sei  $\mathcal{B} = (b_1, \dots, b_n)$ . Sei  $\mathcal{A} = (a_1, \dots, a_m)$ . Sei  $F_{\mathcal{A}, \mathcal{B}} = (\phi_{ij})_{i=1, j=1}^{m, n}$ . Sei  $x_{\mathcal{B}} = (\beta_1, \dots, \beta_n)$ , d.h.  $x = \beta_1 b_1 + \dots + \beta_n b_n$ . Sei  $y_{\mathcal{A}} = (\alpha_1, \dots, \alpha_n)$ , d.h.  $y = \alpha_1 a_1 + \dots + \alpha_n a_n$ .

$$\begin{aligned} y &= F(x) = F(\beta_1 b_1 + \dots + \beta_n b_n) = \beta_1 F(b_1) + \dots + \beta_n F(b_n) \\ &= \sum_{j=1}^n \beta_j F(b_j) = \sum_{j=1}^n \beta_j \left( \sum_{i=1}^m \phi_{ij} a_i \right) = \sum_{i=1}^m \left( \sum_{j=1}^n \phi_{ij} \beta_j \right) a_i. \end{aligned}$$

Andererseits gilt  $y = \sum_{i=1}^m \alpha_i a_i$ . Es folgt  $\alpha_i = \sum_{j=1}^n \phi_{ij} \beta_j$  für alle  $i \in \{1, \dots, m\}$ . D.h.  $y_{\mathcal{A}} = F_{\mathcal{A}, \mathcal{B}} x_{\mathcal{B}}$ .  $\square$

$F_{\mathcal{A}, \mathcal{B}}$  heißt die Darstellung von  $F$  in  $\mathcal{A}$  und  $\mathcal{B}$ . Die Abbildung  $F \mapsto F_{\mathcal{A}, \mathcal{B}}$  ist eine lineare Bijektion von  $\text{Lin}(V, W)$  nach  $\mathbb{K}^{m \times n}$ .

Für lineare Abbildungen  $F : V \rightarrow V$  und Basen  $\mathcal{B}$  von  $V$  führen wir die Bezeichnung  $F_{\mathcal{B}} := F_{\mathcal{B}, \mathcal{B}}$  ein.

#### 4.5.5 Basendarstellung einer Komposition von linearen Abbildungen

**Thm.** Seien  $U, V$  und  $W$  endlichdimensionale Vektorräume über  $\mathbb{K}$ . Seien  $F : V \rightarrow W$  und  $G : U \rightarrow V$  linear. Sei  $\mathcal{A}$  Basis von  $W$ ,  $\mathcal{B}$  Basis von  $V$  und  $\mathcal{C}$  Basis von  $U$ . Dann gilt:

$$(F \circ G)_{\mathcal{A}, \mathcal{C}} = F_{\mathcal{A}, \mathcal{B}} \cdot F_{\mathcal{B}, \mathcal{C}}. \quad (4.5.10)$$

*Beweis.* Übung (?), folgt aus dem vorigen Theorem.  $\square$

#### 4.5.6 Basiswechsel für lineare Abbildungen

**Thm.** Seien  $V$  und  $W$  endlichdimensionale Vektorräume über  $\mathbb{K}$ . Seien  $\mathcal{A}$  und  $\mathcal{A}'$  Basen von  $W$ . Seien  $\mathcal{B}$  und  $\mathcal{B}'$  Basen von  $V$ . Sei  $F : V \rightarrow W$  linear. Dann gilt:

$$F_{\mathcal{A}', \mathcal{B}'} = T_{\mathcal{A}' \leftarrow \mathcal{A}} \cdot F_{\mathcal{A}, \mathcal{B}} \cdot T_{\mathcal{B} \leftarrow \mathcal{B}'} \quad (4.5.11)$$

*Beweis.* Sei  $x \in V$  und  $y = F(x)$ . Es gilt  $y_{\mathcal{A}} = F_{\mathcal{A}, \mathcal{B}} \cdot x_{\mathcal{B}}$  und  $y_{\mathcal{A}'} = F_{\mathcal{A}', \mathcal{B}'} \cdot x_{\mathcal{B}'}$ . Mit der Verwendung von  $x_{\mathcal{B}} = T_{\mathcal{B} \leftarrow \mathcal{B}'} \cdot x_{\mathcal{B}'}$  folgt  $y_{\mathcal{A}} = F_{\mathcal{A}, \mathcal{B}} \cdot T_{\mathcal{B} \leftarrow \mathcal{B}'} \cdot x_{\mathcal{B}'}$ . Multiplikation mit  $T_{\mathcal{A}' \leftarrow \mathcal{A}}$  von links ergibt  $y_{\mathcal{A}'} = T_{\mathcal{A}' \leftarrow \mathcal{A}} \cdot y_{\mathcal{A}} = T_{\mathcal{A}' \leftarrow \mathcal{A}} \cdot F_{\mathcal{A}, \mathcal{B}} \cdot T_{\mathcal{B} \leftarrow \mathcal{B}'} \cdot x_{\mathcal{B}'}$ . Es folgt  $F_{\mathcal{A}', \mathcal{B}'} \cdot x_{\mathcal{B}'} = T_{\mathcal{A}' \leftarrow \mathcal{A}} \cdot F_{\mathcal{A}, \mathcal{B}} \cdot T_{\mathcal{B} \leftarrow \mathcal{B}'} \cdot x_{\mathcal{B}'}$ . Sei  $\mathcal{B}' = (b'_1, \dots, b'_n)$ . Im Fall  $x = b'_i$  gilt  $x_{\mathcal{B}'} = e_i$ . Die vorige Gleichung ergibt, dass die  $i$ -ten Spalten von  $F_{\mathcal{A}', \mathcal{B}'}$  und  $T_{\mathcal{A}' \leftarrow \mathcal{A}} \cdot F_{\mathcal{A}, \mathcal{B}} \cdot T_{\mathcal{B} \leftarrow \mathcal{B}'}$  übereinstimmen. Weil  $i$  beliebig ist, folgt die Behauptung.  $\square$

# 5 Determinanten

## 5.1 Grundlagen

### 5.1.1 Geometrische Motivation

Sei  $n \in \mathbb{N}$  und  $a_1, \dots, a_n \in \mathbb{R}^n$ . Sei  $P(a_1, \dots, a_n) := \{\sum_{i=1}^n \alpha_i a_i : \alpha_1, \dots, \alpha_n \in [0, 1]\}$ . Wenn  $a_1, \dots, a_n$  linear unabhängig sind, dann heißt  $P(a_1, \dots, a_n)$  *Parallelotop* (*Parallelepiped* im Fall  $n = 3$ , *Parallelogramm* im Fall  $n = 2$ ). Sei  $V(a_1, \dots, a_n)$  das *Volumen* von  $P(a_1, \dots, a_n)$  (Flächeninhalt im Fall  $n = 2$ ). Volumen wird in der Analysis eingeführt.

- $V(a_1, \dots, a_n)$  ist „zum Teil linear“ in den Vektorargumenten  $a_1, \dots, a_n$ .  
 $V(a_1, a'_2 + a''_2) = V(a_1, a'_2) + V(a_1, a''_2)$ , wenn  $a'_2$  und  $a''_2$  auf derselben Seite von  $\text{lin}(a_1)$  liegen (siehe Bild).
- $V(\lambda a_1, a_2) = \lambda V(a_1, a_2)$ , falls  $\lambda \geq 0$  (siehe Bild).

- Betrachten wir den Fall  $a_1 = (l, 0), a_2 = (s, h)$  mit  $l, s, h \in \mathbb{R}$ .  
 $V(a_1, a_2) = |lh|$ .

Es stellt sich heraus, dass  $\pm V(a_1, \dots, a_n)$  mit einem Vorzeichen, das von der Orientierung des Systems  $a_1, \dots, a_n$  abhängig ist, linear in jedem der  $n$  Vektoren  $a_1, \dots, a_n$  ist.

### 5.1.2 Definierende Eigenschaften

**Thm.** Sei  $n \in \mathbb{N}$ . Es existiert genau eine Funktion  $\det : \overbrace{\mathbb{K}^n \times \dots \times \mathbb{K}^n}^{n \text{ mal}} \rightarrow \mathbb{K}$  mit den folgenden Eigenschaften:

(D1)  *$\det$  ist linear in jedem der  $n$  Vektorargumente, d.h.:*

$$\begin{aligned} \det(a_1, \dots, a_{i-1}, \alpha u + \beta v, a_{i+1}, \dots, a_n) &= \alpha \det(a_1, \dots, a_{i-1}, u, a_{i+1}, \dots, a_n) \\ &\quad + \beta \det(a_1, \dots, a_{i-1}, v, a_{i+1}, \dots, a_n) \end{aligned}$$

für alle  $a_1, \dots, a_n, u, v \in \mathbb{K}^n$ ,  $\alpha, \beta \in \mathbb{K}$  und  $i \in \{1, \dots, n\}$ .

(D2)  $\det$  ist alternierend, d.h. wenn  $a_i = a_j$  für zwei Indizes  $i, j \in \{1, \dots, n\}$  mit  $i \neq j$ , dann  $\det(a_1, \dots, a_n) = 0$  für alle  $a_1, \dots, a_n \in \mathbb{K}^n$ .

(D3)  $\det(e_1, \dots, e_n) = 1$ .

Der Beweis wird später gegeben. Zunächst wird  $\det$  als eine beliebige Funktion mit (D1), (D2), (D3) behandelt (ohne zu verifizieren, ob eine solche Funktion existiert oder eindeutig ist).

Sei  $n \in \mathbb{N}$  und sei  $A = (a_{ij})_{i,j=1}^n = (a_1 | \dots | a_n) \in \mathbb{K}^{n \times n}$ . Wir nennen  $\det(A) := \det(a_1, \dots, a_n)$  die *Determinante* von  $A$ . Eine weitere Bezeichnung:

$$\begin{vmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{vmatrix} \quad (5.1.1)$$

### 5.1.3 Weitere Eigenschaften

**Thm.** Sei  $n \in \mathbb{N}$ . Sei  $A = (a_{ij}) \in \mathbb{K}^{n \times n}$ . Dann gilt:

- (D4)  $\det(\lambda A) = \lambda^n \det(A)$  für alle  $\lambda \in \mathbb{K}$ .
- (D5) Enthält  $A$  eine Nullspalte, dann gilt  $\det(A) = 0$ .
- (D6) Entsteht  $B$  durch Vertauschung von zwei Spalten von  $A$ , dann gilt  $\det(B) = -\det(A)$ .

Mit anderen Worten ändern die Elementartransformationen von Spalten vom Typ 1 das Vorzeichen der Determinante, da sich die Orientierung der Vektoren ändert.

- (D7) Entsteht  $B$  aus  $A$  durch Addition der  $\lambda$ -fachen  $j$ -ten Spalte zur  $i$ -ten Spalte (mit  $\lambda \in \mathbb{K}$ ,  $i, j \in \{1, \dots, n\}$  und  $i \neq j$ ), so gilt  $\det(B) = \det(A)$ .

*Mit anderen Worten ändern die Elementartransformationen von Spalten vom Typ 3 die Determinante nicht.*

(D8) *Ist  $A$  eine obere Dreiecksmatrix, so gilt  $\det(A) = a_{11} \cdot \dots \cdot a_{nn}$ .*

(D9) *Ist  $n \geq 2$  und hat  $A$  die Form  $A = \begin{pmatrix} B & C \\ O & D \end{pmatrix}$  mit  $B \in \mathbb{K}^{k \times k}$ ,  $C \in \mathbb{K}^{k \times (n-k)}$  und  $D \in \mathbb{K}^{(n-k) \times (n-k)}$  und  $k \in \{1, \dots, n-1\}$ , dann gilt  $\det(A) = \det(B) \det(D)$ .*

(D10)  $\det(A) = 0 \Leftrightarrow \text{rang}(A) < n$ .

(D11) *Für jede  $B \in \mathbb{K}^{n \times b}$  gilt  $\det(AB) = \det(A) \det(B)$ . Insbesondere, wenn  $A$  invertierbar ist, dann gilt  $\det(A) \neq 0$  und  $\det(A^{-1}) = \det(A)^{-1}$ .*

*Beweis.*

(D4) folgt aus (D1):

$$\det(\lambda A) = \det(\lambda a_1, \dots, \lambda a_n) = \lambda^n \det(a_1, \dots, a_n) = \lambda^n \det(A).$$

(D5) Der Einfachheit halber sei  $a_1 = 0$ . Dann gilt:

$$\det(A) = \det(0, a_2, \dots, a_n) = \det(0 \cdot 0, a_2, \dots, a_n) = 0 \cdot \det(0, a_2, \dots, a_n) = 0.$$

(D6) Der Einfachheit halber wird angenommen, dass  $B$  durch Vertauschen der ersten und zweiten Spalte von  $A$  entsteht.

$$\begin{aligned} 0 &= \det(a_1 + a_2, a_1 + a_2, a_3, \dots, a_n) \\ &= \det(a_1, a_1 + a_2, a_3, \dots, a_n) + \det(a_2, a_1 + a_2, a_3, \dots, a_n) \\ &= \det(a_1, a_1, a_3, \dots, a_n) + \det(a_1, a_2, a_3, \dots, a_n) \\ &\quad + \det(a_2, a_1, a_3, \dots, a_n) + \det(a_2, a_2, a_3, \dots, a_n) \\ &= 0 + \det(A) + \det(B) + 0 \\ \Rightarrow \det(B) &= -\det(A). \end{aligned}$$

(D7) O.B.d.A. sei  $i = 1$  und  $j = 2$ .

21.01.2015

$$\begin{aligned}
\det(B) &= \det(a_1 + \lambda a_2, a_2, \dots, a_n) \\
&= \det(a_1, a_2, \dots, a_n) + \lambda \det(a_2, a_2, \dots, a_n) \\
&= \det(a_1, a_2, \dots, a_n) + \lambda \cdot 0 \\
&= \det(A).
\end{aligned}$$

(D8) Sei  $A = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ & \ddots & \vdots \\ & & a_{nn} \end{pmatrix}$  eine obere Dreiecksmatrix.

Wir betrachten den Fall, dass  $a_{ii} = 0$  ist, für ein  $i \in \{1, \dots, n\}$ . Wir fixieren das kleinstmögliche  $i$  wie oben, d.h.  $a_{jj} \neq 0 \forall j \in \{1, \dots, n\}$  mit  $j < i$ .

Wir verwenden eine Spaltentransformation, um mit Hilfe der  $(i - 1)$ -ten Spalte die  $(i - 1)$ -te Komponente der  $i$ -ten Spalte durch 0 zu ersetzen. Wir verwenden anschließend eine Spaltentransformation vom Typ 3, um mit Hilfe der  $(i - 2)$ -ten Spalte die  $(i - 2)$ -te Komponente

der  $i$ -ten Spalte durch 0 zu ersetzen. Die iterative Fortsetzung dieser Prozedur erzeugt eine Matrix, deren  $i$ -te Spalte eine 0-Spalte ist.

Eine Matrix mit einer 0-Spalte hat die Determinante 0. Da die Spaltentransformationen vom Typ 3 die Determinante nicht ändern, folgt  $\det(A) = 0$ .

Wir betrachten nun den Fall, dass  $a_{ii} \neq 0 \forall i \in \{1, \dots, n\}$ . Mit Hilfe von Spaltentransformationen vom Typ 3 ersetzt man die zweite Komponente der Spalten 3 bis  $n$  durch 0, usw.

Mit dieser Prozedur wird  $A$  zur Diagonalmatrix konvertiert. D.h. es gilt:

$$\begin{aligned}
\det(A) &= \det \begin{pmatrix} a_{11} & & \\ & \ddots & \\ & & a_{nn} \end{pmatrix} \\
&= \det(a_{11}e_1, \dots, a_{nn}e_n) \\
&= a_{11} \cdots a_{nn} \cdot \det(e_1, \dots, e_n) \\
&= a_{11} \cdots a_{nn}.
\end{aligned}$$

(D9) Sei  $A = \begin{pmatrix} B & C \\ O & D \end{pmatrix}$  mit  $B \in \mathbb{K}^{k \times k}$ ,  $D \in \mathbb{K}^{(n-k) \times (n-k)}$  und  $k \in \{1, \dots, n\}$ .

Die Matrizen  $B$  und  $D$  lassen sich durch Spaltentransformationen vom Typ 1 (Vertauschung) und Typ 3 zu oberen Dreiecksmatrizen  $B'$  und  $D'$  konvertieren. (Im Wesentlichen verwendet man dabei das Gaußverfahren.)

Den Spaltentransformationen, die man bei der Konvertierung von  $B$  zu  $B'$  und  $D$  zu  $D'$  benutzte, kann man Spaltentransformationen der

Matrix  $A$  zuordnen, welche die Matrix  $A$  zu einer Matrix der Form  
 $A' = \begin{pmatrix} B' & C' \\ O & D' \end{pmatrix}$  überführen.

Die Matrizen  $A'$ ,  $B'$  und  $D'$  sind diagonal:

$$\det(A') = \det(B') \det(D')$$

Sei  $s$  die Anzahl der Transformationen vom Typ 1, die bei der Konvertierung von  $B$  benutzt wurden. Sei  $t$  die Anzahl der Transformationen vom Typ 1, die bei der Konvertierung von  $D$  benutzt wurden.

Dann ist  $s + t$  die Anzahl der Transformationen vom Typ 1, die bei der Konvertierung von  $A$  benutzt wurden. Es gilt:

$$\det(A') = (-1)^{s+t} \det(A)$$

$$\det(B') = (-1)^s \det(B)$$

$$\det(D') = (-1)^t \det(D)$$

Die vorigen Gleichungen und  $(*)$  ergeben  $\det(A) = \det(B) \det(D)$ .

(D10) Im Fall  $n = 1$  ist die Behauptung trivial. Sei  $n \geq 2$ . Im Fall  $\text{rang}(A) < n$  ist eine der Spalten von  $A$  Linearkombination der restlichen Spalten. O.B.d.A. sei  $a_1 = \sum_{i=2}^n \lambda_i a_i$  mit  $\lambda_2, \dots, \lambda_n \in \mathbb{K}$ .

$$\begin{aligned}\det(A) &= \det(a_1, \dots, a_n) \\ &= \det\left(\sum_{i=2}^n \lambda_i a_i, a_2, \dots, a_n\right) \\ &= \sum_{i=2}^n \lambda_i \det(a_i, a_2, \dots, a_n) \\ &= \sum_{i=2}^n \lambda_i \cdot 0 = 0\end{aligned}$$

Im Fall  $\text{rang}(A) = n$  zeigen wir, dass  $\det(A) \neq 0$  ist. Wir konvertieren  $A$  zu einer oberen Dreiecksmatrix mit der Verwendung von Spaltentransformationen vom Typ 1 und 3. Dafür wird im Wesentlichen das Gaußverfahren benutzt.

$$A = \begin{pmatrix} & & | \\ a_1 & \cdots & a_n \\ & & | \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}$$

Die  $n$ -te Komponente einer der Spalten  $a_1, \dots, a_n$  ist ungleich Null, denn sonst wäre  $\text{rang}(A) < n$ .

Durch die Verwendung von Spaltentransformationen vom Typ 1 kann man annehmen, dass  $a_{m,n} \neq 0$  ist. Mit Transformationen vom Typ 3 kann man anschließend die Komponenten  $a_{n,1}, \dots, a_{n,n-1}$  durch 0 ersetzen.

$$A' = \begin{pmatrix} & & * \\ B & & \vdots \\ & & * \\ 0 & \cdots & 0 & a_{nn} \end{pmatrix}, \text{ wobei } B \in \mathbb{K}^{(n-1) \times (n-1)} \text{ und } a_{nn} \neq 0.$$

Nach dieser Modifizierung von  $A$  ist der Rang der Matrix  $(a_{ij})_{i,j=1}^{n-1} \in \mathbb{K}^{(n-1) \times (n-1)}$  gleich  $n - 1$ . Denn sonst wären die  $n - 1$  Spalten dieser

Matrix und somit auch die ersten  $n-1$  Spalten von  $A$  linear abhangig, was der Bedingung  $\text{rang}(A) = n$  widersprache.

Durch die iterative Fortsetzung der vorigen Schritte wird  $A$  zu einer oberen Dreiecksmatrix konvertiert, deren Diagonalelemente alle ungleich Null sind. Es folgt  $\det(A) \neq 0$ .

(D11) Im Fall  $\text{rang}(B) < n$  gilt  $\text{rang}(AB) \leq \text{rang}(B) < n$ . Somit gilt  $\det(AB) = 0 = \det(A) \cdot 0 = \det(A) \det(B)$ .

Im Fall  $\text{rang}(B) = n$  ist  $B$  invertierbar. Die Matrix  $B$  lsst sich durch Spaltentransformationen vom Typ 1,2,3 zur Einheitsmatrix  $I_n$  konvertieren.

Jede Elementartransformation der Spalten kann als Matrixmultiplikation dargestellt werden. Sei  $C = (c_1, \dots, c_n) \in \mathbb{K}^{n \times n}$  beliebig.

22.01.2015

Typ 1: Das Vertauschen der  $i$ -ten und  $j$ -ten Spalte von  $C$  mit  $i, j \in \{1, \dots, n\}, i \neq j$ . Diese Transformation kann als  $C \mapsto C \cdot S$  beschreiben werden, wobei  $S$  aus der Einheitsmatrix entsteht, indem man die Elemente in den Positionen  $(i, i)$  und  $(j, j)$  durch Nullen und die Elemente in den Positionen  $(i, j)$  und  $(j, i)$  durch Einsen ersetzt.

Für die Determinanten gilt:

$$\det(CS) = -\det(C)$$

$$\det(S) = -1$$

Typ 2: Die  $i$ -te Spalte von  $C = (c_1, \dots, c_n)$  wird durch  $\lambda c_i$  ersetzt, mit  $\lambda \in \mathbb{K} \setminus \{0\}$ . Diese Transformation wird als  $C \mapsto C \cdot S$  beschrieben, wobei  $S$  aus der Einheitsmatrix entsteht, indem man das Element in der Position  $(i, i)$  durch  $\lambda$  ersetzt.

Für die Determinanten gilt:

$$\det(CS) = \lambda \det(C)$$

$$\det(S) = \lambda$$

Typ 3:  $i, j \in \{1, \dots, n\}, i \neq j$ .  $C = (c_1, \dots, c_n)$ . Die  $i$ -te Spalte wird durch  $c_i + \lambda c_j$  ersetzt, mit  $\lambda \in \mathbb{K}$ . Diese Transformation wird als  $C \mapsto CS$  beschrieben, wobei  $S$  aus der Einheitsmatrix entsteht, indem das Element in der Position  $(j, i)$  durch  $\lambda$  ersetzt.

Für die Determinanten gilt:

$$\det(CS) = \det(C)$$

$$\begin{aligned}\det(S) &= \det(e_1, \dots, e_{i-1}, e_i + \lambda e_j, e_{i+1}, \dots, e_n) \\ &= \det(e_1, \dots, e_n) + \lambda \det(e_1, \dots, e_{i-1}, e_j, e_{i+1}, \dots, e_n) \\ &= 1 + \lambda \cdot 0 = 1\end{aligned}$$

Da  $\text{rang}(B) = n$ , existieren Matrizen  $S_1, \dots, S_t \in \mathbb{K}^{n \times n}$  ( $t \in \mathbb{N}_0$ ), die den Spaltentransformationen vom Typ 1, 2, 3 entsprechen und für die  $B = S_1 \cdot \dots \cdot S_t$  gilt. Es folgt:

$$\begin{aligned}
\det(AB) &= \det(AS_1 \cdots S_t) \\
&= \det(AS_1 \cdots S_{t-1}) \det(S_t) \\
&\quad \vdots \\
&= \det(A) \det(S_1) \cdots \det(S_t) \\
&= \det(A) \det(S_1 S_2) \cdots \det(S_t) \\
&\quad \vdots \\
&= \det(A) \det(S_1 \cdots S_t) \\
&= \det(A) \det(B).
\end{aligned}$$

Für invertierbare Matrizen  $A$  gilt  $\det(A^{-1}) = \det(A)^{-1}$ , denn  $1 = \det(I) = \det(A^{-1}A) = \det(A^{-1})\det(A)$ .  $\square$

## 5.2 Leibniz-Formel

### 5.2.1 Permutationen und Determinanten

Für  $n \in \mathbb{N}$  bezeichnet  $S_n$  die Menge aller bijektiven Abbildungen von  $\{1, \dots, n\}$  nach  $\{1, \dots, n\}$ . Die Elemente von  $S_n$  heißen *Permutationen*.  $S_n$  bildet eine Gruppe bzgl. der Komposition. Die Gruppe  $S_n$  schreiben wir multiplikativ.

Das neutrale Element von  $S_n$  wird durch  $e$  bezeichnet. Die Permutationen  $\sigma \in S_n$  werden folgendermaßen tabellarisch dargestellt:

$$\begin{bmatrix} 1 & \cdots & n \\ \sigma(1) & \cdots & \sigma(n) \end{bmatrix} \quad (5.2.1)$$

Bsp.  $\begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{bmatrix}$

Sei  $A = (a_1, \dots, a_n) = (a_{ij})_{i,j=1}^n \in \mathbb{K}^{n \times n}$  ( $n \in \mathbb{N}$ ). Wegen  $a_j = a_{1j}e_1 + \dots +$

$a_{nj}e_n \forall j \in \{1, \dots, n\}$  gilt:

$$\begin{aligned}
 \det(A) &= \det(a_1, \dots, a_n) \\
 &= \det\left(\sum_{i_1=1}^n a_{i_1,1}e_{i_1}, \dots, \sum_{i_n=1}^n a_{i_n,n}e_{i_n}\right) \\
 &= \sum_{i_1, \dots, i_n \in \{1, \dots, n\}} a_{i_1,1} \cdots a_{i_n,n} \cdot \det(e_{i_1}, \dots, e_{i_n})
 \end{aligned}$$

Wenn zwei Indizes unter  $i_1, \dots, i_n$  den gleichen Wert haben, gilt  $\det(e_{i_1}, \dots, e_{i_n}) = 0$ . Ansonsten beschreiben  $i_1, \dots, i_n$  eine Permutation  $\sigma$  mit  $\sigma(1) = i_1, \dots, \sigma(n) = i_n$ . Es folgt:

$$\det(A) = \sum_{\sigma \in S_n} a_{\sigma(1),1} \cdots a_{\sigma(n),n} \cdot \underbrace{\det(e_{\sigma(1)}, \dots, e_{\sigma(n)})}_{\text{muss noch berechnet werden}}$$

### 5.2.2 Zerlegung von Permutation in Produkte von Transpositionen

Eine Permutation  $\tau \in S_n$  ( $n \in \mathbb{N}$ ) heißt Transposition der Elemente  $i \in \{1, \dots, n\}$  und  $j \in \{1, \dots, n\}$  mit  $i \neq j$ , falls  $\tau$  alle Elemente von  $\{1, \dots, n\}$  außer  $j$  unverändert lässt und die Elemente  $i$  und  $j$  vertauscht.  
D.h.:

$$\begin{aligned}\tau(k) &= k \quad \forall k \in \{1, \dots, n\} \setminus \{i, j\} \\ \tau(i) &= j \\ \tau(j) &= i\end{aligned}$$

**Prop.** *Jede Permutation ist Produkt endlich vieler Transpositionen.*

*Beweis.* Für jede Transposition  $\tau \in S_n$  gilt  $\tau^2 = e$ . Wir führen Induktion über  $n \in \mathbb{N}$ . Im Fall  $n = 1$  ist  $e$  das einzige Element von  $S_n$ ;  $e$  ist das Produkt von 0 Permutationen. Sei  $n \in \mathbb{N}$  mit  $n \geq 2$ . Sei  $\sigma \in S_n$  beliebig.

Im Fall  $\sigma(n) = n$ , dann  $\sigma$  als Permutation von  $\{1, \dots, n-1\}$  interpretiert werden und die Behauptung folgt nach der Induktionsvoraussetzung.

Im Fall  $\sigma(n) \neq n$  existiert ein  $i \in \{1, \dots, n-1\}$  mit  $\sigma(i) = n$ . Sei  $\tau$  die Transposition der Elemente  $i$  und  $n$ . Dann gilt  $(\sigma\tau)(n) = n$ , sodass  $\sigma\tau$  als Permutation von  $\{1, \dots, n-1\}$  interpretiert werden kann. Nach der Induktionsvoraussetzung gilt  $\sigma\tau = \tau_1 \cdots \tau_k$  für gewisse Transpositionen  $\tau_1, \dots, \tau_k \in S_n$ . Die Multiplikation mit  $\tau$  ergibt  $\sigma = \sigma\tau\tau = \tau_1 \cdots \tau_k \cdot \tau$ .  $\square$

### 5.2.3 Vorzeichen von Permutationen

Für eine Menge  $X$  wird durch  $\binom{X}{2}$  die Menge aller zweielementigen Teilmengen von  $X$  bezeichnet. D.h.  $\binom{X}{2} = \{\{x, x'\} : x, x' \in X, x \neq x'\}$ .

Sei  $n \in \mathbb{N}$ , sei  $N := \{1, \dots, n\}$  und sei  $\sigma \in S_n$ . Wir nennen  $I \in \binom{N}{2}$  den *Fehlstand* von  $\sigma$ , falls  $I = \{i, j\}$  mit  $i < j$  und  $\sigma(i) > \sigma(j)$  gilt.

Bsp.  $\begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{bmatrix}$  hat die Fehlstände  $\{2, 3\}$  und  $\{1, 3\}$ .

Ist  $k$  die Anzahl der Fehlstände von  $\sigma$ , so nennt man  $\text{sign}(\sigma) = (-1)^k$  das *Vorzeichen* von  $\sigma$ . Man sagt  $\sigma$  ist *gerade*, wenn  $\text{sign}(\sigma) = 1$  ist, und  $\sigma$  ist *ungerade*, wenn  $\text{sign}(\sigma) = -1$  ist.

**Prop.**

28.01.2015

Sei  $n \in \mathbb{N}$ ,  $N = \{1, \dots, n\}$ ,  $\sigma \in S_n$ . Dann gilt:

$$\text{sign}(\sigma) = \prod_{\{i,j\} \in \binom{N}{2}} \frac{\sigma(j) - \sigma(i)}{j - i} \quad (5.2.2)$$

Die rechte Seite ist wohldefiniert. (Beim Tausch von  $i$  und  $j$  ändert sich der Wert nicht.)

*Beweis.* Seien  $i, j \in \mathbb{N}$  mit  $i < j$ . Ist  $\{i, j\}$  Fehlstand in  $\sigma$ , dann gilt:

$$\frac{\sigma(j) - \sigma(i)}{j - i} = (-1) \cdot \frac{|\sigma(j) - \sigma(i)|}{|j - i|} < 0$$

Ist  $\{i, j\}$  kein Fehlstand von  $\sigma$ , dann gilt:

$$\frac{\sigma(j) - \sigma(i)}{j - i} = \frac{|\sigma(j) - \sigma(i)|}{|j - i|}$$

D.h. die rechte Seite von (5.2.2) ist

$$\begin{aligned} & (-1)^k \prod_{\{i,j\} \in \binom{N}{2}} \frac{|\sigma(j) - \sigma(i)|}{|j - i|} \\ &= \text{sign}(\sigma) \prod_{\{i,j\} \in \binom{N}{2}} \frac{|\sigma(j) - \sigma(i)|}{|j - i|} \\ &= \text{sign}(\sigma) \left( \frac{\prod_{\{i,j\} \in \binom{N}{2}} |\sigma(j) - \sigma(i)|}{\prod_{\{i,j\} \in \binom{N}{2}} |j - i|} \right) \end{aligned}$$

Die Produkte im Zähler und Nenner sind gleich: Die Abbildung  $\{i, j\} \mapsto \{\sigma(i), \sigma(j)\}$  ist eine Bijektion auf  $\binom{N}{2}$ . Durch Variablensubstitution  $i' = \sigma(i)$  und  $j' = \sigma(j)$  erhält man, dass das Produkt im Zähler  $\prod_{\{i',j'\} \in \binom{N}{2}} |j' - i'|$  ist.  $\square$

## 5.2.4 Vorzeichen und das Produkt von Permutationen

**Thm.** Sei  $n \in \mathbb{N}$ ,  $\sigma, \tau \in S_n$ . Dann gilt:

$$\text{sign}(\sigma\tau) = \text{sign}(\sigma) \cdot \text{sign}(\tau) \quad (5.2.3)$$

*Beweis.* Sei  $N = \{1, \dots, n\}$ . Dann ist

$$\begin{aligned} \text{sign}(\sigma\tau) &= \prod_{\{i,j\} \in \binom{N}{2}} \frac{\sigma(\tau(j)) - \sigma(\tau(i))}{j - i} \\ &= \prod_{\{i,j\} \in \binom{N}{2}} \left( \frac{\sigma(\tau(j)) - \sigma(\tau(i))}{\tau(j) - \tau(i)} \cdot \frac{\tau(j) - \tau(i)}{j - i} \right) \\ &= \underbrace{\prod_{\{i,j\} \in \binom{N}{2}} \frac{\sigma(\tau(j)) - \sigma(\tau(i))}{\tau(j) - \tau(i)}}_{\text{sign}(\sigma) \text{ [Substituiere } \tau(j)=j', \tau(i)=i']} \cdot \underbrace{\prod_{\{i,j\} \in \binom{N}{2}} \frac{\tau(j) - \tau(i)}{j - i}}_{\text{sign}(\tau)} \\ &= \text{sign}(\sigma) \cdot \text{sign}(\tau) \end{aligned}$$

□

**Bem.** Für jede Transposition  $\tau$  gilt  $\text{sign}(\tau) = -1$  (Übungsaufgabe). Jede Permutation  $\sigma \in S_n$  kann man als Produkt  $\sigma = \tau_1 \cdots \tau_k$  von  $k \in \mathbb{N}_0$  Transpositionen darstellen.

Nach der vorigen Proposition gilt  $\text{sign}(\sigma) = 1 \Leftrightarrow k$  gerade.

### 5.2.5 Leibniz-Formel

Seien  $b_1, \dots, b_n \in \mathbb{K}^n$ . Ist  $\tau \in S_n$  eine Transposition, so gilt nach (D6):

$$\det(b_{\tau(1)}, \dots, b_{\tau(n)}) = -\det(b_1, \dots, b_n)$$

Mit Berücksichtigung der vorigen Bemerkung folgt

$$\det(b_{\sigma(1)}, \dots, b_{\sigma(n)}) = \text{sign}(\sigma) \det(b_1, \dots, b_n) \quad \forall \sigma \in S_n \quad (5.2.4)$$

Unter anderem gilt für die Standardbasis:

$$\det(e_{\sigma(1)}, \dots, e_{\sigma(n)}) = \text{sign}(\sigma) \cdot \det(e_1, \dots, e_n) = \text{sign}(\sigma)$$

D.h. für  $A = (a_1, \dots, a_n) = (a_{ij}) \in \mathbb{K}^{n \times n}$  gilt wegen 5.2.1 die sogenannte Leibniz-Formel:

$$\det(A) = \sum_{\sigma \in S_n} \text{sign}(\sigma) a_{\sigma(1),1} \cdots a_{\sigma(n),n} \quad (5.2.5)$$

Bsp.

$n = 2$ :

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \in \mathbb{K}^{2 \times 2}$$

$$\begin{aligned} \det(A) &= a_{11} \cdot a_{22} & \text{sign} \begin{bmatrix} 1 & 2 \\ 1 & 2 \end{bmatrix} &= 1 \\ &\quad - a_{21} \cdot a_{12} & \text{sign} \begin{bmatrix} 1 & 2 \\ 2 & 1 \end{bmatrix} &= -1 \end{aligned}$$

$n = 3$ :

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$$

$$\begin{aligned} \det(A) &= a_{11}a_{22}a_{33} & \text{sign } \begin{bmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{bmatrix} &= 1 \\ &\quad - a_{11}a_{32}a_{23} & \text{sign } \begin{bmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{bmatrix} &= -1 \\ &\quad - a_{21}a_{12}a_{33} & \text{sign } \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{bmatrix} &= -1 \\ &\quad - a_{31}a_{22}a_{13} & \text{sign } \begin{bmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{bmatrix} &= -1 \\ &\quad + a_{21}a_{32}a_{13} & \text{sign } \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{bmatrix} &= 1 \\ &\quad + a_{31}a_{12}a_{23} & \text{sign } \begin{bmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{bmatrix} &= 1 \end{aligned}$$

Das Muster:  $+ \begin{bmatrix} \diagup & \diagdown \\ \diagdown & \diagup \end{bmatrix}$  und  $- \begin{bmatrix} \diagup & \diagdown \\ \diagup & \diagdown \end{bmatrix}$ .

D.h. (D1), (D2), (D3)  $\Rightarrow$  Leibniz-Formel (5.2.5). Es bleibt zu zeigen, dass die durch die Leibniz-Formel definierte Funktion (D1), (D2), (D3) erfüllt ( $\Leftarrow$ ):

(D1) Schreibe o.B.d.A.  $a_1$  als Linearkombination  $a_1 = \alpha u + \beta v$  mit  $\alpha, \beta \in$

$\mathbb{K}, u, v \in \mathbb{K}^n$ .

$$\begin{aligned}
\det(A) &= \sum_{\sigma \in S_n} \text{sign}(\sigma) a_{\sigma(1),1} \cdots a_{\sigma(n),n} \\
&= \sum_{\sigma \in S_n} (\alpha u_{\sigma(1),1} + \beta v_{\sigma(1),1}) a_{\sigma(2),2} \cdots a_{\sigma(n),n} \\
&= \alpha \sum_{\sigma \in S_n} \text{sign}(\sigma) u_{\sigma(1),1} a_{\sigma(2),2} \cdots a_{\sigma(n),n} \\
&\quad + \beta \sum_{\sigma \in S_n} \text{sign}(\sigma) v_{\sigma(1),1} a_{\sigma(2),2} \cdots a_{\sigma(n),n}
\end{aligned}$$

(D2) O.B.d.A. sei  $a_1 = a_2$ . Sei  $\tau \in S_n$  Transposition von 1 und 2. Die Abbildung  $\sigma \mapsto \sigma\tau$  auf  $S_n$  ist eine Bijektion, auch eine Bijektion von

$\{\sigma \in S_n : \text{sign}(\sigma) = -1\}$  nach  $\{\sigma \in S_n : \text{sign}(\sigma) = 1\}$ . Es folgt:

$$\begin{aligned}\det(A) &= \sum_{\sigma \in S_n} \text{sign}(\sigma) a_{\sigma(1),1} \cdots a_{\sigma(n),n} \\ &= \sum_{\substack{\sigma \in S_n \\ \text{sign}(\sigma)=1}} a_{\sigma(1),1} \cdots a_{\sigma(n),n} - \sum_{\substack{\sigma \in S_n \\ \text{sign}(\sigma)=-1}} a_{\sigma(1),1} \cdots a_{\sigma(n),n}\end{aligned}$$

Aus  $a_1 = a_2$  folgt  $a_{\sigma(2),1} = a_{\sigma(2),2}$  und  $a_{\sigma(1),1} = a_{\sigma(1),2}$ . Somit sind die beiden Summen des vorigen Ausdrucks gleich, d.h. der Ausdruck ist 0.

(D3) Für  $A$  wird  $I = (\delta_{ij})_{i,j=1}^n$  eingesetzt:

10.04.2015

$$\det(I) = \sum_{\sigma \in S_n} \text{sign}(\sigma) \underbrace{\delta_{\sigma(1),1} \cdots \delta_{\sigma(n),n}}_{\substack{\neq 0 \text{ nur dann,} \\ \text{wenn } \sigma(1)=1}} \underbrace{\neq 0 \text{ nur dann,} \\ \text{wenn } \sigma(n)=n}_{\substack{\neq 0, \text{ nur wenn } \sigma(1)=1, \dots, \sigma(n)=n, \\ \text{d.h. wenn } \sigma = \text{id} = e \text{ ist.}}}$$

Daher ist die vorige Summe = 1.

### 5.2.6 Die Determinante der Transponierten Matrix

Sei  $A = (a_1, \dots, a_n) = (a_{ij})_{i,j=1}^n \in \mathbb{K}^{n \times n}$  ( $n \in \mathbb{N}$ ). Die Leibniz-Formel besagt:

$$\begin{aligned}\det(A) &= \sum_{\sigma \in S_n} \text{sign}(\sigma) a_{\sigma(1),1} \cdots a_{\sigma(n),n} \\ &= \sum_{\sigma \in S_n} \text{sign}(\sigma) \prod_{k=1}^n a_{\sigma(k),k}\end{aligned}$$

Im Produkt wird  $i$  für  $\sigma(k)$  eingesetzt. Mit  $\text{sign}(\sigma^{-1}) = \text{sign}(\sigma)$  folgt:

$$\begin{aligned}\det(A) &= \sum_{\sigma \in S_n} \text{sign}(\sigma) \prod_{i=1}^n a_{i,\sigma^{-1}(i)} \\ &= \sum_{\sigma \in S_n} \text{sign}(\sigma^{-1}) \prod_{i=1}^n a_{i,\sigma^{-1}(i)}\end{aligned}$$

Im Summanden kann  $\sigma^{-1}$  durch  $\sigma$  ersetzt werden, da die Abbildung  $\sigma \mapsto \sigma^{-1}$  eine Bijektion von  $S_n$  nach  $S_n$  ist. D.h. man hat eine weitere Form der

Leibniz-Formel:

$$\det(A) = \sum_{\sigma \in S_n} \text{sign}(\sigma) \prod_{i=1}^n a_{i,\sigma(i)}$$

Sei  $A^\top = (a_{i,j}^\top)_{i,j=1}^n$  mit  $a_{ij}^\top = a_{ji}$   $\forall i, j = 1, \dots, n$ . Aus der vorigen Formel ergibt sich die Gleichung:

$$\det(A) = \sum_{\sigma \in S_n} \text{sign}(\sigma) \prod_{i=1}^n a_{\sigma(i),i}^\top = \det(A^\top)$$

wegen der Leibniz-Formel aus 5.2.5.

**Thm.** Sei  $A \in \mathbb{K}^{n \times n}$ . Dann gilt:

$$\det(A^\top) = \det(A). \quad (5.2.6)$$

## 5.3 Determinante, Rang, Minoren und Invertierung von Matrizen

### 5.3.1 Cramer'sche Regel

Die Formel für die Determinante (Leibniz-Formel), die wir im letzten Abschnitt hergeleitet haben, führt nun zu Formeln für die Lösungen der Grundaufgaben der linearen Algebra. Wie wir bereits wissen, ist die Aufgabe  $Ax = b$  (Lösung von linearen Gleichungssystemen) eine zentrale Aufgabe, die man innerhalb der linearen Algebra (und darüber hinaus) verstehen und behandeln will. Bei Systemen mit  $n$  Gleichungen und  $n$  unbekannten hat man im Fall einer regulären Matrix  $A \in \mathbb{K}^{n \times n}$  genau eine Lösung. Wir sind auf der Suche nach einer Formel für diese eindeutige Lösung in Abhängigkeit von  $A$  und  $b$ .

Wir denken an das LGS  $Ax = b$  mit der rechten Seite  $b \in \mathbb{K}^n$  und der

regulären Matrix  $A \in \mathbb{K}^{n \times n}$  spaltenweise und schreiben daher  $Ax = b$  als

$$a_1x_1 + \cdots + a_nx_n = b$$

mit Hilfe der  $n$  Spalten  $a_1, \dots, a_n \in \mathbb{K}^n$  von  $A$  und der  $n$  unbekannten Komponenten  $x_1, \dots, x_n$  des Vektors  $x$  um. Für jede einzelne Variable  $x_i$  wollen wir nun eine Formel für deren Wert herleiten. Um den Wert für  $x_i$  herzuleiten, betrachten wir

$$\Delta_i := \det(a_1, \dots, a_{i-1}, b, a_{i+1}, \dots, a_n).$$

Das ist die Determinante der Matrix, welche aus  $A$  durch das Einfügen der Spalte  $b$  an der Stelle der  $i$ -ten Spalte entsteht. Da  $b$  Linearkombination der Spalten  $a_1, \dots, a_n$  mit den Koeffizienten  $x_1, \dots, x_n$  ist, gilt

$$\Delta_i = \det(a_1, \dots, a_{i-1}, a_1x_1 + \cdots + a_nx_n, a_{i+1}, \dots, a_n).$$

Weil die Determinante in jedem ihrer  $n$  Vektorargumente linear ist, gilt

$$\Delta_i = x_1 \det(a_1, \dots, a_{i-1}, a_1, a_{i+1}, \dots, a_n) + \cdots + x_n \det(a_1, \dots, a_{i-1}, a_n, a_{i+1}, \dots, a_n)$$

Die Determinante ist ein alternierendes Funktional: sind zwei der  $n$  Vektorargumente gleich, dann ist die Determinante gleich 0. Also überlebt in der vorigen Summe auf der rechten Seite genau ein Term, und zwar der  $i$ -te. Wir erhalten somit

$$\Delta_i = x_i \det(a_1, \dots, a_n) = x_i \det(A).$$

Wie wir bereits wissen, ist die Matrix  $A$  genau dann regulär, wenn  $\det(A) \neq 0$  ist. Unter der Voraussetzung, dass  $A$  regulär ist, können wir also die Gleichung  $\Delta_i = x_i \det(A)$  durch  $\det(A)$  teilen, und erhalten somit die Formel

$$x_i = \frac{\det(a_1, \dots, a_{i-1}, b, a_{i+1}, \dots, a_n)}{\det(A)} \quad (5.3.1)$$

für den Wert der  $i$ -ten Unbekannten der Lösung von einem LGS  $Ax = b$  mit einer regulären Matrix  $A$ . Gleichung (5.3.1), mit  $i \in \{1, \dots, n\}$ , heißt die *Cramer'sche Regel*.

Die Cramer'sche Regel ist im allgemeinen Fall kein Konkurrent zum Gauß-Verfahren zur Lösung von linearen Gleichungssystemen. Wenn man

etwa ein konkretes  $100 \times 100$  System  $Ax = b$  lösen möchte, so müsste man nach der Cramer'schen Regel 101 Determinanten ausrechnen, und das Letztere würde man wahrscheinlich auch mit dem Gauß-Verfahren machen. Anstatt 101-mal das Gauß-Verfahren für die Berechnung der Determinante der  $100 \times 100$  Matrizen auszuführen, wäre es viel sinnvoller ein mal das Gauß-Verfahren zur Lösung von  $Ax = b$  auszuführen. Die Cramer'sche Regel ist also im Allgemeinen keine Alternative zum Gauß-Verfahren, sie ist mehr ein weiterer Beitrag zur Theorie der linearen Gleichungssysteme, der anderen Zwecken dient. Zum Beispiel kann man die Cramer'sche Regel für qualitative theoretische Analyse benutzen. Hier ein Beispiel. Die Leibniz-Formel zeigt, dass  $\det(A)$  eine stetige Funktion in den Komponenten von  $A$  ist. Die Cramer'sche Regel beschreibt die Lösung  $x$  von  $Ax = b$  als Quotienten der Determinanten und zeigt somit, dass  $x$  als Funktion von  $A$  und  $b$  ebenfalls eine stetige Funktion in den Komponenten von  $A$  und  $b$  ist. Des Weiteren sehen wir, dass  $x$  linear von  $b$  abhängig ist, und

kriegen durch die Formel (5.3.1) eine genaue Darstellung der Abhangigkeit. Des Weiteren bietet die Cramer'sche Regel Moglichkeiten,  $Ax = b$  in den parametrischen/symbolischen Fallen zu behandeln. Das bedeutet, dass manche/alle der Komponenten von  $A$  und  $b$  keine feste Werte sind, sondern Parameter/Symbole. In diesen Fallen schafft, man dann eine Formel fur  $x$  in Abhangigkeit von den Parametern/Symbolen herzuleiten.

**Bsp** (Eine Gleichung und eine Unbekannte). Ein lineares Gleichungssystem in einer Unbekannten und mit einer Gleichung hat die Form  $ax = b$  mit  $a, b \in \mathbb{K}$  und einer Unbekannten  $x \in \mathbb{K}$ . Diese System kann man naturlich ohne jegliches Wissen aus der linearen Algebra problemlos behandeln. Im Fall  $a \neq 0$ , ist  $x = \frac{b}{a}$  die eindeutige Losung. Diese Formel ist Spezialfall  $n = 1$  der Cramer'schen Regel. Der Wert  $a$  ist etwa die Determinante der  $1 \times 1$  Matrix  $A = (a)$ .

Schon der Fall von zwei unbekannten und zwei Gleichungen

$$\begin{cases} a_{11}x_1 + a_{12}x_2 = b_1 \\ a_{21}x_1 + a_{22}x_2 = b_2 \end{cases} \quad (5.3.2)$$

ist nicht so angenehm, wenn man für diesen Fall, für allgemeine Koeffizienten die Formel für die Lösung  $x_1, x_2$  in Abhängigkeit von den Koeffizienten der linken und rechten Seite ohne Benutzung der Theorie herleiten möchte. Im Prinzip ist so eine Herleitung möglich. Man kann unter Annahme  $a_{11} \neq 0$ ,  $x_1$  aus der ersten Gleichung in Abhängigkeit von  $x_2$  darstellen. Dann kann man dann die Darstellung für  $x_1$  in die zweite Gleichung einsetzen und anschließend aus der Gleichung, die auf diese Weise entsteht, die Formel für  $x_2$  und dann für  $x_1$  bestimmen (probieren Sie das mal aus!). So eine Herleitung ist ziemlich umständlich und gar nicht systematisch. Die Benutzung der Cramer'schen Regel führt uns dagegen direkt zu den gewünschten Formeln:

**Bsp** (Zwei Gleichungen und zwei Unbekannten). Das System (5.3.2) hat

genau dann eine eindeutige Lösung, wenn die Determinante

$$\Delta := \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{21}a_{12}$$

der Matrix der linken Seite ungleich 0 ist. In diesem Fall ist die eindeutige Lösung

$$x_1 = \frac{\Delta_1}{\Delta} \quad x_2 = \frac{\Delta_2}{\Delta}$$

mit

$$\Delta_1 := \begin{vmatrix} b_1 & a_{12} \\ b_2 & a_{22} \end{vmatrix} = b_1a_{22} - b_2a_{12} \quad \Delta_2 := \begin{vmatrix} a_{11} & b_1 \\ a_{21} & b_2 \end{vmatrix} = a_{11}b_2 - a_{21}b_1.$$

**Bsp** (Interpolation mit univariaten affinen Funktionen). Interpolation ist die Berechnung einer Funktion (aus einer vorab festgelegten Menge von Funktionen), die an gegebenen Stellen (den sogenannten Stützstellen) vorgegebene Werte haben soll. Mehrere Grundformen der Interpolation führen

direkt zu linearen Gleichungssystemen. (Interpolation wird im Mathematik-Studium meistens in dem Kurs Numerik behandelt.)

Wir betrachten in diesem Beispiel Interpolation mit univariaten affinen Funktionen. Ein Beispiel aus dem Uni-Leben: der Vorlesungsleiter legt fest, dass man beim Erreichen von 95% aller Punkte für die Klausur eine 1 bekommt, und beim Erreichen von 50% aller Punkte eine 4 bekommt. Das kann man als die Vorgabe  $1 \mapsto 0,95$  und  $4 \mapsto 0,5$  beschreiben. Wie viel Prozent aller Punkte soll man für andere Noten wie etwa 2 und 3 erreichen? Eine logische Regel wäre, das durch eine affine Funktion  $f(x) = ax + b$  zu beschrieben, welche die Interpolationsbedingungen  $f(1) = 0,9$  und  $f(4) = 0,5$  erfüllt. Die Stützstellen sind also hier 1 und 4 und die Stützwerte jeweils 0,9 und 0,5. Lösen wir diese Aufgabe ganz allgemeine für zwei beliebige Stützstellen und -Werte.

Wir bestimmen die Funktion  $f : \mathbb{K} \rightarrow \mathbb{K}$  der Form  $f(x) = ax + b$  mit

den Koeffizienten  $a, b \in \mathbb{K}$  derart, dass die Bedingungen

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

für gegebene  $x_1, x_2, y_1, y_2 \in \mathbb{K}$  mit  $x_1 \neq x_2$  erfüllt sind. Die beiden Interpolationsbedingungen sind lineare Gleichungen

$$\begin{cases}x_1 \cdot a + 1 \cdot b = y_1 \\x_2 \cdot b + 1 \cdot b = y_2\end{cases}$$

in den unbekannten Koeffizienten  $a$  und  $b$ . Lassen Sie sich an dieser Stelle nicht durch die Bezeichnungen verwirren. Wir haben für die Interpolation andere Bezeichnungen eingeführt, die nicht auf die Bezeichnungen bei der Diskussion von allgemeinen linearen Gleichungssystemen abgestimmt sind. Nun sind bei uns  $a$  und  $b$  die Unbekannten und  $x_1, x_2, y_1, y_2$  gegebene Werte,

durch welche die Koeffizienten unseres LGS festgelegt werden. Mit Matrizen und Vektoren kann das System als

$$\begin{pmatrix} x_1 & 1 \\ x_2 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$$

geschrieben werden. Somit ist die Determinante der Matrix der linken Seite

$$\Delta := \begin{vmatrix} x_1 & 1 \\ x_2 & 1 \end{vmatrix} = x_1 - x_2.$$

Man hat  $\Delta \neq 0$ , da die beiden Stützstellen zwei verschiedene Werte sind. Die Anwendung der Cramer'schen Regel ergibt die Formeln

$$a = \frac{1}{\Delta} \begin{vmatrix} y_1 & 1 \\ y_2 & 1 \end{vmatrix} = \frac{y_1 - y_2}{x_1 - x_2} \quad b = \frac{1}{\Delta} \begin{vmatrix} x_1 & y_1 \\ x_2 & y_2 \end{vmatrix} = \frac{x_1 y_2 - x_2 y_1}{x_1 - x_2}$$

für die beiden Koeffizienten  $a$  und  $b$  der interpolierenden Funktionen  $f(x) = ax + b$ .

**Bsp** (Interpolation mit quadratischen Funktionen). Geben Sie analog zum vorigen Beispiel Formeln für die Koeffizienten  $a, b, c$  bei der Interpolation mit quadratischen Funktionen  $f(x) = ax^2 + bx + c$  und den Interpolationsbedingungen  $f(x_1) = y_1$ ,  $f(x_2) = y_2$  und  $f(x_3) = y_3$  für drei paarweise verschiedene Stützstellen  $x_1, x_2, x_3 \in \mathbb{K}$ .

**Bsp** (Interpolation mit affinen Funktionen von zwei Variablen). Wir bestimmen jene Funktion  $f : \mathbb{K}^2 \rightarrow \mathbb{K}$ , die durch  $f(x, y) = ax + by + c \quad \forall (x, y) \in \mathbb{K}^2$  ( $a, b, c \in \mathbb{K}$  sind die Koeffizienten) und die Interpolationsbedingungen

$$f(x_i, y_i) = z_i \quad \forall i \in \{1, 2, 3\}$$

gegeben ist, für  $x_i, y_i, z_i \in \mathbb{K}$ . Der Graph der Funktion  $f$  ist eine Ebene. Geometrisch beschrieben, geht es also um die Suche nach einer Ebenen, die drei gegebene Punkte  $(x_1, y_1, z_1), (x_2, y_2, z_2), (x_3, y_3, z_3)$ . Eine solche Ebene muss nicht immer existieren oder eindeutig sein: ein Problem hat man, wenn die drei Punkte nicht kollinear sind. Wir können also nicht erwarten, dass

eindeutige  $a, b, c$  ohne Zusatzbedingungen existieren. Die Interpolationsbedingungen schreiben wir wie in den vorigen Beispielen als ein LGS in der Matrixform:

$$\begin{pmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = \begin{pmatrix} z_1 \\ z_2 \\ z_3 \end{pmatrix} \quad (5.3.3)$$

Für die Existenz einer eindeutigen Lösung  $a, b, c$  dieses Systems ist also notwendig und hinreichend, dass die Determinante der Matrix der linken Seite

$$\Delta := \begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix}$$

ungleich 0 ist. Geometrisch bedeutet die Bedingung  $\Delta \neq 0$ , dass die Stützstellen  $(x_1, y_1), (x_2, y_2)$  und  $(x_3, y_3)$  keine kollinearen Punkte sind. In diesem

Fall ergibt die Cramer'sche Regel die Formeln

$$a = \frac{1}{\Delta} \cdot \begin{vmatrix} z_1 & y_1 & 1 \\ z_2 & y_2 & 1 \\ z_3 & y_3 & 1 \end{vmatrix} \quad b = \frac{1}{\Delta} \cdot \begin{vmatrix} x_1 & z_1 & 1 \\ x_2 & z_2 & 1 \\ x_3 & z_3 & 1 \end{vmatrix} \quad c = \frac{1}{\Delta} \cdot \begin{vmatrix} x_1 & y_1 & z_1 \\ x_2 & y_2 & z_2 \\ x_3 & y_3 & z_3 \end{vmatrix}$$

für die gesuchten Koeffizienten  $a, b, c$ , der Funktion  $f$ .

**Bem.** In der Numerik findet die Interpolation natürliche Anwendungen, wenn der zugrundeliegende Körper der Körper der reellen oder komplexen Zahlen ist. Dieser Kontext ist unter den Mathe-Studierenden am meisten bekannt, weil die Interpolation ein Standard-Thema aus der Numerik ist. Wie wir aber oben gesehen haben, kann man die Interpolation genau so gut bzgl. eines beliebigen zugrundeliegenden Körpers diskutieren, ohne das man was ändern muss (die Interpolationsformeln bleiben unverändert). Interpolation bzgl. endlicher Körper benutzt man in der Kodierungstheorie

und Kryptographie, Interpolation in exakter Arithmetik, bzgl. des Körpers  $\mathbb{Q}$ , wird in Kombinatorik benutzt.

### 5.3.2 Entwicklung nach Zeilen und Spalten

Nun möchten wir wieder die Berechnung der Determinanten diskutieren. Bis jetzt hatten wir dafür zwei Werkzeuge zur Verfügung: das Gauß-Verfahren und die Leibniz-Formel. Wie bereits oben besprochen ist die Leibniz-Formel eigentlich kein Konkurrent für das Gauß-Verfahren, sondern ein theoretisches Hilfsmittel, das in den symbolischen/parametrischen Kontexten eingesetzt werden kann.

Die Entwicklung nach einer Spalte ist eine weitere Formel, mit der man eine  $n \times n$  Determinante durch gewisse Determinanten der Größe  $(n - 1) \times (n - 1)$  darstellen kann.

Wir betrachten eine  $n \times n$  Matrix  $A = (a_{ij})_{i,j=1}^n \in \mathbb{K}^{n \times n}$  und bezeichnen

deren Spalten als  $a_1, \dots, a_n$ , sodass man

$$\det(A) = \det(a_1, \dots, a_n)$$

hat. Wir fixieren wir nun eine beliebige Spalte  $a_j$ , nach der wir die Determinante entwickeln wollen. Die Spalte  $a_j$  ist Linearkombination der Vektoren der Standardbasis:

$$a_j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{nj} \end{pmatrix} = \sum_{i=1}^n a_{ij} e_i.$$

Aus der Linearität in  $a_j$  folgt also

$$\begin{aligned} \det(A) &= \det(a_1, \dots, a_n) \\ &= \det(a_1, \dots, a_{j-1}, \sum_{i=1}^n a_{ij} e_i, a_{j+1}, \dots, a_n) \\ &= \sum_{i=1}^n a_{ij} \underbrace{\det(a_1, \dots, a_{j-1}, e_i, a_{j+1}, \dots, a_n)}_{=: \Delta_{i,j}}. \end{aligned}$$

Wir haben also  $\det(A)$  als Linearkombination der Determinanten  $\Delta_{i,j}$  beschrieben, mit

$$\Delta_{i,j} = \det \begin{pmatrix} & & 0 & & \\ NW & & \vdots & & NO \\ & & 0 & & \\ a_{i,1} & \cdots & a_{i,j-1} & 1 & a_{i,j+1} & \cdots & a_{i,n} \\ & & & 0 & & & \\ SW & & \vdots & & SO \\ & & 0 & & \end{pmatrix}$$

Die Matrix die wir oben sehen hat in der Spalte  $j$  überall Nullen bis auf die Position  $(i, j)$ , in der eine 1 steht. Ansonsten hat die Matrix die gleichen Komponenten wie  $A$ . Bei dieser Matrix kann man bzgl. der Position  $(i, j)$  vier Untermatrizen  $NW, NO, SW, SO$  einführen, bzgl. der vier Himmelsrichtungen Nordwest, Nordost, Südost und Südwest. Bemerkung zu Bezeichnungen: Die Matrizen  $NW, NO, SO, SW$  hängen von der Wahl

von  $(i, j)$  ab, was man aber an unseren Bezeichnungen nicht direkt sieht (wir wollen komplizierte Bezeichnungen vermeiden).

Es bleibt nun, die Determinante  $\Delta_{i,j}$  zu einer  $(n - 1) \times (n - 1)$  zu konvertieren. Die Eins in der Position  $(i, j)$  kann man benutzt werden, um alle anderen Komponenten in der  $i$ -ten Zeile mit Hilfe der Transformationen vom Typ 3 durch 0 zu ersetzen. Da in der  $j$ -ten Spalte nur die Komponente in der Position  $(i, j)$  ungleich 0 ist, ändert sich hierbei nur die  $i$ -te Spalte der Matrix. Man erhält:

$$\Delta_{i,j} = \det \begin{pmatrix} & & 0 \\ NW & \vdots & NO \\ & 0 \\ 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\ & 0 \\ SW & \vdots & SO \\ & 0 \end{pmatrix}$$

Als Nächstes wollen wir Zeilen und Spalten so vertauschen, dass nach dem Vertauschen die Eins in der Position  $(i, j)$  in der Ecke oben links landet. Schrittweise geht man folgendermaßen vor: die  $j$ -te Spalte ‘hüpft’  $j - 1$  mal über die davorstehenden Spalten und wird nach  $j - 1$  mal Hüpfen die erste Spalte. Anschließend hüpfst die  $i$ -te Zeile  $i - 1$  mal über die darüber liegenden

Zeilen und wird nach  $i - 1$  mal Hüpfen die erste Zeile. Schematisch:

$$\begin{array}{c}
 \left( \begin{array}{ccccccccc} & & 0 & & & & & & \\ NW & : & NO & & & & & & \\ & & 0 & & & & & & \\ 0 & \dots & 0 & 1 & 0 & \dots & 0 & & \\ & & & 0 & & & & & \\ SW & : & SO & & & & & & \\ & & 0 & & & & & & \end{array} \right) \\
 \rightsquigarrow \left( \begin{array}{ccccc} 0 & & & & \\ : & NW & NO & & \\ 0 & & & & \\ 1 & 0 & \dots & & 0 \\ 0 & & & & \\ : & SW & SO & & \\ 0 & & & & \\ 1 & 0 & \dots & & 0 \\ 0 & & & & \\ : & NW & NO & & \\ & & & & \\ SW & & SO & & \end{array} \right) \text{ durch } j - 1 \text{ Spaltenvertauschungen} \\
 \rightsquigarrow \left( \begin{array}{ccccc} 1 & 0 & \dots & & 0 \\ 0 & & & & \\ : & NW & NO & & \\ & & & & \\ SW & & SO & & \end{array} \right) \text{ 269} \\
 \rightsquigarrow \text{durch } i - 1 \text{ Zeilenvertauschungen}
 \end{array}$$

Wir haben  $j - 1$  mal Spalten und  $i - 1$  mal Zeilen vertauscht. Jedes Vertauschen ändert das Vorzeichen der Determinante. Wir erhalten also die Gleichung

$$\det(A) = (-1)^{i+j} \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & NW & NO & \\ & SW & SO & \\ 0 & & & \end{pmatrix}$$

Die  $(n-1) \times (n-1)$  Matrix aus den Blöcken  $NW, NO, SW, SO$  entsteht aus  $A$  durch Streichen der  $i$ -ten Zeile und  $j$ -ten Spalte. Wir bezeichnen diese

Matrix durch  $A'_{i,j}$ . Man hat

$$\begin{aligned}\Delta_{i,j} &= (-1)^{i+j} \det \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & A'_{i,j} & \\ 0 & & & \end{pmatrix} \\ &= (-1)^{i+j} \det(1) \det(A'_{i,j}) \quad \text{wegen (D9)} \\ &= (-1)^{i+j} \det(A'_{i,j})\end{aligned}$$

Das ergibt die Entwicklung der Determinante von  $A$  nach der  $j$ -ten Spalte:

$$\det(A) = \sum_{i=1}^n a_{i,j} \cdot (-1)^{i+j} \det(A'_{i,j}) \quad (5.3.4)$$

Bei der Berechnung von Determinanten spielen die Zeilen die gleiche Rolle wie die Spalten (vgl. 5.2.6). Daher: eine analoge Formel bzgl. Zeilen. Wir halten also den folgenden Satz:

**Thm** (Der Entwicklungssatz von Laplace). Sei  $A = (a_{ij})_{i,j=1}^n \in \mathbb{K}^{n \times n}$  mit  $n \in \mathbb{N}$ . Dann gilt:

$$\det(A) = \sum_{i=1}^n a_{ij} (-1)^{i+j} \det(A'_{ij}) \quad \forall j \in \{1, \dots, n\} \quad (5.3.5)$$

$$= \sum_{j=1}^n a_{ij} (-1)^{i+j} \det(A'_{ij}) \quad \forall i \in \{1, \dots, n\}, \quad (5.3.6)$$

wobei  $A'_{ij} \in \mathbb{K}^{(n-1) \times (n-1)}$  aus  $A$  durch das Streichen der  $i$ -ten Zeile und  $j$ -ten Spalte entsteht.

**Bsp** ( $n = 2$ ). Die Anwendung der Laplace-Entwicklung, egal für welche Zeile oder Spalte, für eine allgemeine  $2 \times 2$  ergibt logischerweise die Leibniz-Formel  $\det(A) = a_{11}a_{22} - a_{21}a_{12}$  im Fall  $n = 2$ .

**Bsp** ( $n = 3$ ). Wenn wir die Determinante einer allgemeinen  $3 \times 3$  Matrix

etwa nach der ersten Spalte entwickeln, erhalten wir

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

Theoretisch reicht die Laplace-Entwicklung allein aus, um Determinante einer Matrix beliebiger Größe rekursiv zu berechnen: denn die Determinante einer  $1 \times 1$  Matrix ist einfach der einzige Wert in der Matrix und die Determinante einer  $n \times n$  wird in der Laplace-Entwicklung durch die Determinanten von  $(n-1) \times (n-1)$  Matrizen ausgedrückt. Es ist nicht schwer zu verstehen, dass eine solche Methode für genügend große  $n$  viel zu aufwändig ist und keine alternative zum Gauß-Verfahren bietet. Dennoch ist die Laplace-Entwicklung als ein weiterer theoretischer Baustein nützlich. Es gelten hier die gleichen Kommentare, die wir bereits bei der Diskussion der Leibniz-Formel und der Cramer'schen Regel gegeben haben. Des Weiteren kann die Leibniz-Formel bei konkreten Berechnungen eingesetzt werden,

wenn eine Spalte oder Zeile der zugrundeliegenden Matrix dünn besetzt ist (= wenig Nichtnull-Elemente hat).

**Bsp.** Sind zwei Vektoren  $v_1, v_2 \in \mathbb{K}^3$  linear unabhängig dann ist die lineare Hülle von  $v_1$  und  $v_2$  eine Ebene, die den Nullpunkt,  $v_1$  und  $v_2$  enthält. Wie sieht die Gleichung dieser Ebene für eine allgemeine Wahl von  $v_1$  und  $v_2$  aus?

Wir definieren  $v_1$  und  $v_2$  durch deren Koordinaten:  $v_1 = (x_1, y_1, z_1)$ ,  $v_2 = (x_2, y_2, z_2)$ . Sei  $v = (x, y, z) \in \mathbb{K}^3$  beliebiger Vektor. Der Vektor  $v$  liegt genau dann in  $\text{lin}(v_1, v_2)$ , wenn  $v, v_1, v_2$  linear abhängig sind. Die Vektoren  $v, v_1, v_2$  sind genau dann linear unabhängig, wenn die  $3 \times 3$  Matrix mit Spalten  $v, v_1, v_2$  singulär ist. Das Letztere ist äquivalent zu  $\det(v, v_1, v_2) = 0$ , oder mit Koordinaten:

$$\begin{vmatrix} x & x_1 & x_2 \\ y & y_1 & y_2 \\ z & z_1 & z_2 \end{vmatrix} = 0.$$

Die Gleichung der Ebene können wir in der Form  $ax + by + cz = 0$ , mit Koeffizienten  $a, b, c \in \mathbb{K}$ , mühelos hinschreiben. Wenn man die vorige Determinante nach der ersten Spalte entwickelt erhält man

$$\begin{vmatrix} y_1 & y_2 \\ z_1 & z_2 \end{vmatrix} x - \begin{vmatrix} x_1 & x_2 \\ z_1 & z_2 \end{vmatrix} y + \begin{vmatrix} x_1 & x_2 \\ y_1 & y_2 \end{vmatrix} z = 0.$$

D.h., die Gleichung der Ebene  $\text{lin}(v_1, v_2)$  ist  $ax + by + cz = 0$ , mit den Koeffizienten

$$a = y_1z_2 - z_1y_2, \quad b = -(x_1z_2 - z_1x_2), \quad c = x_1y_2 - y_1x_2.$$

Wie man auch an den Beispielen aus diesem Abschnitt sieht, sind die wenigen Grundformeln für die Determinanten eine weiterreichende Zusammenfassung verschiedenster Formeln, die man sonst in der Literatur sieht. Wenn man lediglich diese wenigen Grundformeln kennt und ihren Hintergrund versteht, erspart man sich die Mühe, vieler anderen Formeln zu merken.

### 5.3.3 Die komplementäre Matrix

Unser nächstes Ziel ist Formeln für die Komponenten der inversen Matrix herzuleiten. Es ist klar, dass man solche Formeln aus der Cramer'schen Regel herleiten kann. Die Spalten der inversen Matrix sind die Lösungen der  $n$  Gleichungssysteme  $Ax = e_j$  mit  $j \in \{1, \dots, n\}$ . Man kann also  $n$  mal die Cramer'sche Regel anwenden und dann die Determinanten, die bei der Cramer'schen Regel im Zähler des Quotienten stehen, nachbearbeiten. Sie können gerne diesen Weg der Herleitung ausprobieren. Alternativ kann man einfach die Beweisidee der Cramer'schen Regel auf unsere neue Situation anpassen, ohne auf die Regel direkt zuzugreifen.

Wie in den vorigen Paragraphen betrachten wir eine  $n \times n$  Matrix  $A$ , deren Komponenten wir als  $a_{ij}$  und die Spalten als  $a_1, \dots, a_n$  bezeichnen. Wir betrachten eine beliebige Spalte  $a_j$  von  $A$  und tauschen diese gegen eine Spalte  $a_k$  der Matrix  $A$  aus. So entsteht die Matrix mit Spalten  $a_1, \dots, a_{j-1}, a_k, a_{j+1}, \dots, a_n$ . Ist  $k = j$ , dann ist das direkt die Matrix  $A$ ,

deren Determinante gleich  $\det(A)$  ist. Bei  $k \neq j$  haben wir eine Matrix, in der die Spalte  $a_k$  der Matrix  $A$  zweimal vorkommt. Somit ist die Determinante gleich 0. Mit dem Kronecker-Delta, kann man unsere Beobachtung als die Gleichung

$$\det(a_1, \dots, a_{j-1}, a_k, a_{j+1}, \dots, a_n) = \det(A) \cdot \delta_{jk}$$

hinschreiben. Wie bei der Herleitung der Leibniz-Formel wird nun die Spalte  $a_k$ , die wir als das  $j$ -te Vektorargument der Determinante eingesetzt haben in der Standardbasis dargestellt:

$$a_k = \begin{pmatrix} a_{1k} \\ \vdots \\ a_{nk} \end{pmatrix} = \sum_{i=1}^n a_{ik} e_i.$$

Aus der Linearität der Determinante in ihren Vektorargumenten erhält man

$$\begin{aligned}\det(A) \cdot \delta_{jk} &= \det(a_1, \dots, a_{j-1}, a_k, a_{j+1}, \dots, a_n) \\ &= \sum_{i=1}^n a_{ik} \det(a_1, \dots, a_{j-1}, e_i, a_{j+1}, \dots, a_n)\end{aligned}$$

Die Determinante  $\det(a_1, \dots, a_{j-1}, e_i, a_{j+1}, \dots, a_n)$  haben wir bereits bei der Herleitung der Laplace-Entwicklung in [5.3.2] behandelt. Diese Determinante ist gleich  $(-1)^{i+j} \det(A'_{ij})$ , wobei  $A'_{ij}$  die Matrix bezeichnet, die aus  $A$  durch das Streichen der  $i$ -ten Zeile und der  $j$ -ten Spalte entsteht. Das ergibt die Gleichung

$$\det(A) \cdot \delta_{jk} = \sum_{i=1}^n a_{ik} (-1)^{i+j} \det(A'_{ij}). \quad (5.3.7)$$

Wenn man genau hinschaut, erkennt man, dass (5.3.7) im Wesentlichen die Formel zur Berechnung der Komponenten der inversen Matrix ist. Die inverse Matrix  $B = (b_{ji})_{j,i=1}^n$  von  $A$  ist die eindeutige Matrix mit  $BA = I$ ,

oder komponentenweise geschrieben:  $\sum_{i=1}^n b_{ji}a_{ik} = \delta_{jk}$ . Wir können bei einer invertierbaren Matrix  $A$  Formel (5.3.7) durch  $\det(A)$  teilen, da die Determinante einer invertierbaren Matrix ungleich 0 ist. Aus (5.3.7) folgt somit die Darstellung  $b_{ji} = \frac{1}{\det(A)}(-1)^{i+j} \det(A'_{ij})$  der Komponenten der Matrix  $B = A^{-1}$ .

Wie man sieht, braucht man nur am Ende unserer Herleitung die Invertierbarkeit als eine Voraussetzung, unter der man durch  $\det(A)$  teilen kann. Formel (5.3.7) gilt für eine allgemeine  $n \times n$  Matrix  $A$ . Diese Formel motiviert uns die sogenannte *komplementäre Matrix*  $A^\# = (a_{ji}^\#)_{j,i=1}^n \in \mathbb{K}^{n \times n}$  einzuführen, um (5.3.7) mit Hilfe der komplementären Matrix in der Matrixform darstellen zu können. Wir setzen  $a_{ji}^\# := (-1)^{i+j} \det(A'_{ij})$  für  $i, j \in \{1, \dots, n\}$ . Formel (5.3.7) hat also die Matrixform

$$\det(A)I = A^\# \cdot A. \quad (5.3.8)$$

Man kann auch die Version  $A \cdot A^\# = \det(A)I$  dieser Formel analog herleiten. Probieren Sie das aus!

Das folgende Theorem fasst nun unsere Überlegungen zusammen:

**Thm.** Sei  $A \in \mathbb{K}^{n \times n}$ . Für die komplementäre Matrix  $A^\#$  von  $A$  gilt:

$$A \cdot A^\# = A^\# \cdot A = \det(A)I. \quad (5.3.9)$$

Insbesondere, wenn  $A$  invertierbar ist, gilt

$$A^{-1} = \frac{1}{\det(A)} \cdot A^\#. \quad (5.3.10)$$

Beschreiben wir Formel (5.3.10) noch kurz mit Worten: Wenn wir die Komponente in der Position  $(i, j)$  ( $i$ -te Zeile und  $j$ -te Spalte) der inversen Matrix von  $A$  mit ermitteln wollen, dann streichen wir die  $i$ -te Spalte und  $j$ -te Zeile (hier aufpassen!) von  $A$ , berechnen die Determinante der so entstandenen  $(n - 1) \times (n - 1)$  Matrix, teilen diese Determinante durch die Determinante von  $A$ , und passen das Vorzeichen des Ergebnis durch die Multiplikation mit  $(-1)^{i+j}$  an. Multiplikation mit  $(-1)^{i+j}$  ist Änderung des Vorzeichens

nach dem Schachbrettmuster, etwa für  $n = 4$ :

+	-	+	-
-	+	-	+
+	-	+	-
-	+	-	+

**Bsp** ( $n = 2$ ). Für  $A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \in \mathbb{K}^{2 \times 2}$  hat die komplementäre Matrix die Form  $A^\# = \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}$ . Wenn  $A$  invertierbar ist, dann ist  $\det(A) = a_{11}a_{22} - a_{21}a_{12} \neq 0$ . Die Gleichung  $A^{-1} = \frac{1}{\det(A)} A^\#$  kann dann komponentenweise als

$$A^{-1} = \frac{1}{a_{11}a_{22} - a_{21}a_{12}} \begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}.$$

ausgeschrieben werden.

**Bsp** ( $n = 3$ ). Für eine  $3 \times 3$  Matrix  $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \in \mathbb{K}^{3 \times 3}$  hat die komplementäre Matrix die Form

$$A^\# = \begin{pmatrix} \left| \begin{matrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{matrix} \right| - \left| \begin{matrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{matrix} \right| & \left| \begin{matrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{matrix} \right| \\ - \left| \begin{matrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{matrix} \right| & \left| \begin{matrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{matrix} \right| - \left| \begin{matrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{matrix} \right| \\ \left| \begin{matrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{matrix} \right| - \left| \begin{matrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{matrix} \right| & \left| \begin{matrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{matrix} \right| \end{pmatrix}.$$

Teilen durch  $\det(A)$  ergibt eine Formel für  $A^{-1}$ .

### 5.3.4 Rang und Minoren

Wir haben in diesem Kapitel schon Einiges, was wie im Teil 1 des Kurses behandelt hatten (Lösung von Linearen Gleichungssystemen, Invertierung von Matrizen), mit den Determinanten in Verbindung gesetzt, und auf diesem Wege Formeln für entsprechende Grundaufgaben der linearen Algebra erhalten. In diesem Paragraphen ist der Rang dran. Der Rang ist die Dimension der linearen Hülle der Spalten (bzw. Zeilen) der Matrix. Die Berechnung des Ranges (= der Dimension) gehört ebenfalls zu den Grundaufgaben der linearen Algebra. Ist es möglich, den Rang durch die Berechnung von Determinanten zu bestimmen? Ja! Und zwar sollen wir die Determinanten der Untermatrizen verschiedener Größe ausrechnen.

Eine *Untermatrix* einer Matrix  $A$  ist eine Matrix die durch das Streichen einer Auswahl an Spalten und Zeilen von  $A$  entsteht. Hat  $A$  eine Größe  $m \times n$  und sind  $I, J$  Indexmengen  $I \subseteq \{1, \dots, m\}$  und  $J \subseteq \{1, \dots, n\}$ , so ordnet man dem Paar  $I, J$  die Untermatrix zu, die aus  $A$  durch Streichen

der Zeilen mit Nummern außerhalb  $I$  und der Spalten mit Nummern außerhalb von  $J$  entsteht. Ist  $I = \{i_1, \dots, i_k\}$  mit  $1 \leq i_1 < \dots < i_k \leq m$ ,  $J = \{j_1, \dots, j_l\}$  mit  $1 \leq j_1 < \dots < j_l \leq n$ , und  $A = (a_{ij}) \in \mathbb{K}^{m \times n}$ , dann ist die Untermatrix von  $A$  zu den Indexmengen  $I$  und  $J$  die  $k \times l$  Matrix der Form

$$\begin{pmatrix} a_{i_1 j_1} & \cdots & a_{i_1 j_l} \\ \vdots & & \vdots \\ a_{i_k j_1} & \cdots & a_{i_k j_l} \end{pmatrix} \in \mathbb{K}^{k \times l}$$

Zum Beispiel, die Untermatrix von

$$\begin{pmatrix} 2 & 3 & 0 & 4 \\ 1 & \mathbf{6} & -1 & 5 \\ 3 & 2 & 0 & 0 \\ 8 & \mathbf{7} & \mathbf{5} & 6 \end{pmatrix}$$

zu den Indexmengen  $I = \{2, 4\}$  und  $J = \{2, 3\}$  ist  $\begin{pmatrix} 6 & -1 \\ 7 & 5 \end{pmatrix}$ .

Die Determinante einer  $k \times k$  Untermatrix von  $A$  heißt ein  $k \times k$  Minor von  $A$ . Bei der oben angegebenen Matrix ist also

$$\begin{vmatrix} 6 & -1 \\ 7 & 5 \end{vmatrix} = 6 \cdot 5 - 7 \cdot (-1) = 37$$

ein  $2 \times 2$  Minor. Der Wert

$$\begin{vmatrix} -1 & 5 \\ 0 & 0 \end{vmatrix} = 0$$

ist ein weiterer  $2 \times 2$  Minor. Einzelne Komponenten sind  $1 \times 1$  Minoren.

Das folgende Theorem erstellt die Verbindung des Rangs mit den Minoren. Der Rang einer  $m \times n$  ist ein Wert zwischen 0 und  $\min\{m, n\}$ . Aus dem folgenden Theorem folgt, dass der Rang durch die Minoren der Matrix eindeutig bestimmt ist.

**Thm.** Sei  $A \in \mathbb{K}^{m \times n}$  und  $r$  natürliche Zahl mit  $r \leq \min\{m, n\}$ . Dann sind die folgenden Bedingungen äquivalent:

(i)  $\text{rang}(A) \geq r$

(ii) *mindestens ein  $r \times r$  Minor von  $A$  ist ungleich 0.*

*Beweis.* Wir bemerken zuerst, dass die beiden Bedingungen nach dem Vertauschen der Zeilen und Spalten der Matrix  $A$  unverändert bleiben. Wenn eine Eigenschaft bzw. Wert bzgl. einer Operation sich nicht ändert, so sagt man in Mathematik, dass diese Eigenschaft bzw. Wert bzgl. dieser Operation *invariant* ist. Bei der Bedingung (i) folgt die Invarianz im Fall des Vertauschens der Spalten direkt aus der Definition. Beim Vertauschen der Zeilen benutzt man zusätzlich die Eigenschaft, dass der Zeilen- und Spaltenrang einer Matrix gleich sind. Die Invarianz der Bedingung (ii) kann man ebenfalls ziemlich direkt verifizieren. Entsteht Matrix  $B$  aus  $A$  durch das Vertauschen von Zeilen oder Spalten und ist  $M$  ein  $r \times r$  Minor von  $A$ , dann ist  $M$  oder  $-M$  ein  $r \times r$  Minor von  $B$ . Die Umkehrung gilt ebenfalls: ist  $M$  ein  $r \times r$  Minor von  $B$ , dann ist  $M$  oder  $-M$  ein  $r \times r$  Minor von  $A$ .

Das zeigt, dass die Eigenschaft (ii) unverändert bleibt, wenn wir  $A$  durch  $B$  austauschen.

(ii)  $\Rightarrow$  (i): Sei (ii) erfüllt. Wegen der oben besprochenen Invarianz, können wir nach einem geeigneten Vertauschen der Zeilen und Spalten von  $A$  annehmen, dass der Minor der Größe  $r \times r$  oben links ungleich Null ist, das ist der Minor bzgl. der ersten  $r$  Zeilen und Spalten. Das heißt, für  $(a_{ij})_{i=1,\dots,m, j=1,\dots,n} \in \mathbb{R}^{m \times n}$  ist die Determinante der Untermatrix

$$(a_{ij})_{i,j=1,\dots,r}$$

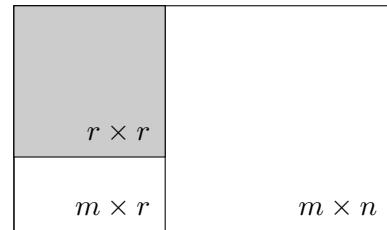
der Größe  $r \times r$  ungleich Null. Das heißt, dass die Untermatrix den Rang  $r$  hat, die Zeilen der Untermatrix bilden also eine Basis von  $\mathbb{K}^r$ . Es folgt, dass die  $m$  Zeilen der  $m \times r$  Untermatrix

$$(a_{ij})_{i=1,\dots,m, j=1,\dots,r}$$

den Vektorraum  $\mathbb{K}^r$  aufspannen. Es folgt, dass diese  $m \times r$  Untermatrix den Rang  $r$  hat. Daraus folgt, dass die  $r$  Spalten dieser  $m \times r$  Untermatrix

einen Vektorraum in  $\mathbb{K}^m$  der Dimension  $r$  aufspannen. Das sind aber die ersten  $r$  der insgesamt  $n$  Spalten von  $A$ . Das bedeutet, dass die  $n$  Spalten von  $A$  einen Vektorraum der Dimension mindestens  $r$  aufspannen. Also ist  $\text{rang}(A) \geq r$ , d.h., (i) ist erfüllt.

Hier eine schematische Darstellung der Untermatrizen, die wir im vorigen Argument benutztten:



(i)  $\Rightarrow$  (ii): sei (i) erfüllt. Um (ii) herzuleiten, können wir im Wesentlichen unseren Beweis von (ii)  $\Rightarrow$  (i) umkehren. Die Spalten von  $A$  spannen einen Raum der Dimension mindestens  $r$  auf. Wir können aus den Spalten von  $A$  eine Basis des Spaltenraums auswählen. Diese Basis hat mindestens  $r$  Vektoren, beliebige  $r$  Vektoren aus der Basis, spannen einen  $r$ -dimensional Vek-

torraum auf. Da wir die Spalten beliebig vertauschen können, können wir ohne Beschränkung der Allgemeinheit annehmen, dass die ersten  $r$  Spalten von  $A$  solche  $r$  linear unabhängige Vektoren sind. Das bedeutet, die  $m \times r$  Untermatrix

$$(a_{ij})_{i=1,\dots,m, j=1,\dots,r}$$

hat den Rang  $r$ . Die Zeilen dieser Matrix spannen somit den  $r$ -dimensionalen Vektorraum  $\mathbb{K}^r$  auf. Man findet also unter den Zeilen der vorigen  $m \times r$  Untermatrix eine Basis von  $\mathbb{K}^r$ . Da wir Zeilen von  $A$  beliebig vertauschen können, können wir annehmen, dass die ersten  $r$  Zeilen der vorigen  $m \times r$  Untermatrix eine Basis von  $\mathbb{K}^r$  bilden. Das bedeutet, dass die  $r \times r$  Untermatrix

$$(a_{ij})_{i=1,\dots,r, j=1,\dots,r}$$

den Rang  $r$  hat. Daraus folgt, dass die Determinante dieser  $r \times r$  Untermatrix ungleich 0 ist. Somit ist (ii) erfüllt.  $\square$

**Bsp.** Die Intuition hinter dem vorigen Theorem wird aus der folgenden Beschreibung klar. Betrachten wir eine Nadel und eine Scheibe im dreidimensionalen Raum. Die Nadel ist eindimensional und die Scheibe ist zweidimensional. Nun wird die Nadel und die Scheibe von vorne, von oben und von der Seite fotografiert. Für die beiden Objekte kriegen wir je drei Fotos und müssen nun die Nadel von der Scheibe auseinanderhalten. Ein oder zwei Fotos reichen im Allgemeinen nicht aus. Wenn z.B. die Scheibe horizontal platziert ist, so sieht man an den Fotos von der Seite und von vorne die zweite Dimension nicht. Mit anderen Worten ist eine entsprechende Projektion der Scheibe eindimensional, obwohl die Scheibe zweidimensional ist. Da wir aber drei Fotos aus drei zueinander senkrechten Richtungen machen, sehen wir an mindestens einem der drei Fotos, dass die Scheibe mindestens zwei Dimensionen hat. Auf diese Weise lässt sich eine Scheibe von einer Nadel unterscheiden.

Was hier informell beschrieben ist ist eine Anwendung des vorigen Theo-

rems im Fall  $r = 2$ . Wir können die Anwendung aber auch rein formal illustrieren. Betrachten wir eine beliebige  $3 \times 2$  Matrix

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{pmatrix}.$$

Der Rang von  $A$  ist 0, 1 oder 2. Wie können wir die Matrizen mit dem Rang höchstens 1 und die Matrizen mit dem Rang 2 auseinander halten? Die geometrische Interpretation ist die Folgende. Der Nullvektor 0 und die beiden Spalten  $a_1, a_2$  von  $A$  bilden im nichtentarteten Fall die Ecken eines Dreiecks. Im entarteten Fall kann sich das Dreieck zu einer Strecke (Rang 1) oder sogar zu einem Punkt (Rang 0) entarten. Laut dem vorigen Theorem hat  $A$  genau dann Rang 2, wenn mindestens einer der drei  $2 \times 2$  Minoren

$$M_1 = \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}, \quad M_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{31} & a_{32} \end{vmatrix}, \quad M_3 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$$

ungleich 0 ist. Mit anderen Worten hat  $A$  genau dann den Rang höchstens 1, wenn  $M_1 = M_2 = M_3 = 0$  erfüllt ist (alle drei  $2 \times 2$  Minoren gleich 0). Wenn die Standardbasisvektoren  $e_1$  nach vorne,  $e_2$  zur Seite und  $e_3$  nach oben gerichtet sind, dann entspricht der Minor  $M_1$  dem Foto von vorne,  $M_2$  dem Foto von der Seite und  $M_3$  dem Foto von oben. Man kann nämlich sehen, dass  $\frac{1}{2}M_i$  bis auf das Vorzeichen die Fläche des Schattens unseres Dreiecks bei der Projektion in Richtung  $e_i$ . Haben alle drei Projektion Fläche 0, dann ist unser Dreieck entartet.

**Bsp.** Bei der Diskussion der Laplace-Entwicklung haben wir bereit einen durch zwei Vektoren aufgespannten zweidimensionalen Vektorraum in  $\mathbb{K}^3$  (eine Ebene) eine Beschreibung durch eine Gleichung hergeleitet. Nun wollen wir eine durch ihre Richtung gegebene Gerade mit Gleichungen beschreiben.

Sei  $v \in \mathbb{K}^3 \setminus \{0\}$ . Wir beschreiben die Gerade  $\text{lin}(v)$  durch lineare Gleichungen. Sei  $v = (a, b, c)$ ,  $a, b, c \in \mathbb{K}$ .

Der Vektor  $p = (x, y, z) \in \mathbb{K}^3$  liegt genau dann auf der Geraden  $\text{lin}(v)$ , wenn die Vektoren  $p$  und  $v$  linear abhängig sind. Das ist genau dann der Fall, wenn der Rang der Matrix

$$\begin{pmatrix} x & a \\ y & b \\ z & c \end{pmatrix}$$

mit den Spalten  $p$  und  $v$  kleiner als zwei ist. Nach dem vorigen Theorem gilt das letztere genau dann, wenn alle 3 Minoren unserer  $3 \times 2$  Matrix gleich 0 sind. Das ergibt die Beschreibung von  $\text{lin}(v)$  durch das System

$$\begin{cases} xb - ya = 0 \\ xc - za = 0 \\ yc - zb = 0 \end{cases}$$

von 3 linearen Gleichungen. Intuitiv ist es klar, dass man eine Gerade im dreidimensionalen Raum durch nur zwei Gleichungen beschreiben kann.

Eine der drei Gleichungen ist tatsächlich redundant, nur weiß man nicht genau welche, wenn man die Koordinaten  $a, b, c$  des Richtungsvektors  $v$  nicht genau spezifiziert, sonder allgemein hält.

Versuchen Sie z.B. im Fall von  $\mathbb{K} = \mathbb{R}$  und  $a = 1, b = 2, c = 3$  eine Beschreibung der Geraden  $\text{lin}(v)$  mit nur zwei Gleichungen herzuleiten. Dabei können Ihnen ihre Kenntnisse über den Rang bestimmt helfen.

### 5.3.5 Der Satz von Binet-Cauchy

Im Abschluss dieses Kapitels diskutieren wir einen weiteren allgemeinen Satz über die Determinanten, den sogenannten Satz von Binet-Cauchy. Für quadratische Matrizen  $A$  und  $B$  der gleichen Größe haben wir am Anfang des Kapitels die Formel  $\det(AB) = \det(A)\det(B)$  hergeleitet. Wenn wir  $B$  durch  $B^\top$  austauschen und dann  $\det(B^\top) = \det(B)$  benutzen, können wir die Formel auch als  $\det(AB^\top) = \det(A)\det(B)$  formulieren. Es stellt sich heraus, dass für die vorige Formel eine Verallgemeinerung für nicht-

quadratische Matrizen (gleicher Größe) existiert. Diese Verallgemeinerung ist der Gegenstand des Satzes von Cauchy-Binet.

Die geometrische Interpretation des Satzes von Binet-Cauchy in einem Spezialfall

**Thm** (Satz von Binet-Cauchy). *Man betrachte  $m \times n$  Matrizen  $A = (a_1, \dots, a_n)$ ,  $B = (b_1, \dots, b_n) \in \mathbb{K}^{m \times n}$  mit  $m \leq n$ . Dann erfüllt die Determinante der  $m \times m$  Matrix  $AB^\top$  die Gleichung.*

$$\det(AB^\top) = \sum_{1 \leq i_1 < \dots < i_m \leq n} \det(a_{i_1}, \dots, a_{i_m}) \det(b_{i_1}, \dots, b_{i_m}). \quad (5.3.11)$$

(Die Summe geht über alle  $i_1, \dots, i_m \in \{1, \dots, n\}$  mit  $1 \leq i_1 < \dots < i_m \leq n$  und hat somit  $\binom{n}{m}$  Summanden).

Im Satz von Binet-Cauchy wird die Determinante von  $AB^\top$  mit der Verwendung aller  $m \times m$  Minoren von  $A$  und  $B$  dargestellt. Die jeweiligen Minoren von  $A$  und  $B$  werden miteinander multipliziert, und alle solche Produkte von Minoren werden zusammengerechnet.

*Beweis des Theorems.* Wir leiten die Formel mit der Verwendung der Grund-eigenschaften der Determinante (die Determinante ist linear in jeder Spalte und alternierend bzgl. der Spalten). Die Spalten von  $AB^\top$  sind Linearkombinationen der Spalten von  $A$ . Dh., wenn wir die Komponenten von  $B$  als  $b_{ij}$  bezeichnen, so erhalten wir

$$\begin{aligned} AB^\top &= \begin{pmatrix} | & & | \\ a_1 & \cdots & a_n \\ | & & | \end{pmatrix} \begin{pmatrix} b_{11} & \cdots & b_{m1} \\ \vdots & & \vdots \\ b_{1n} & \cdots & b_{mn} \end{pmatrix} \\ &= \begin{pmatrix} | & & | \\ \sum_{j=1}^n a_j b_{1j} & \cdots & \sum_{j=1}^n a_j b_{mj} \\ | & & | \end{pmatrix} \end{aligned}$$

D.h. für  $k \in \{1, \dots, m\}$  hat die  $k$ -te Spalte von  $AB^\top$  die Form  $a_1 b_{k1} + \dots + a_n b_{kn} = \sum_{j=1}^n a_j b_{kj}$ . Jeder der  $m$  Spalten von  $AB^\top$  also also Linearkombination von jeweils  $n$  Vektoren. Eine Linearkombination ist eine Summe, und

wir wollen nun all diese  $m$  Summen als eine ‘große Summe’ zusammenfassen. Wir erhalten mit der Verwendung der Linearität in den Spalten die Gleichung:

$$\begin{aligned}\det(AB^\top) &= \det\left(\sum_{j_1=1}^n a_{j_1} b_{1,j_1}, \dots, \sum_{j_m=1}^n a_{j_m} b_{m,j_m}\right) \\ &= \sum_{j_1, \dots, j_m \in \{1, \dots, n\}} \det(a_{j_1}, \dots, a_{j_m}) \cdot b_{1,j_1} \cdots b_{m,j_m}.\end{aligned}$$

Die vorige Summe der  $m^n$  Summanden hat viele Nullen. Da die Determinante alternierendes Funktional ist gilt: Wenn mindestens 2 der Werte  $j_1, \dots, j_m$  gleich sind, ist  $\det(a_{j_1}, \dots, a_{j_m}) = 0$ . Daher kann man sich in der vorigen Summe auf  $j_1, \dots, j_m \in \{1, \dots, m\}$  beschränken, deren Werte paarweise verschieden sind. In diesem Fall kann die  $m$ -elementige Menge  $\{j_1, \dots, j_m\}$  „aufsteigend nummeriert“ werden, d.h. es existieren  $i_1, \dots, i_m \in \{1, \dots, n\}$  mit  $i_1 < \dots < i_m$  und  $\{j_1, \dots, j_m\} = \{i_1, \dots, i_m\}$ .

Sei  $\sigma \in S_m$  eine Permutation mit  $j_1 = i_{\sigma(1)}, \dots, j_m = i_{\sigma(m)}$ . [Bspw. für

$j_1 = 3, j_2 = 1, j_3 = 7$  sind  $i_1 = 1, i_2 = 3, i_3 = 7$  und  $\sigma(1) = 2, \sigma(2) = 1, \sigma(3) = 3.$ ] Die Permutation  $\sigma$  und die Indizes  $i_1, \dots, i_m$  sind durch die Wahl von  $j_1, \dots, j_m$  eindeutig bestimmt. Wir setzen für  $j_1 = i_{\sigma(1)}, \dots, j_m = i_{\sigma(m)}$  ein und erhalten

$$\begin{aligned}
\det(AB^\top) &= \sum_{\substack{\sigma \in S_m \\ 1 \leq i_1 < \dots < i_m \leq n}} \det(a_{i_{\sigma(1)}}, \dots, a_{i_{\sigma(m)}}) \cdot b_{1,i_{\sigma(1)}} \cdots b_{m,i_{\sigma(m)}} \\
&= \sum_{\substack{\sigma \in S_m \\ 1 \leq i_1 < \dots < i_m \leq n}} \det(a_{i_1}, \dots, a_{i_m}) \operatorname{sign}(\sigma) \cdot b_{1,i_{\sigma(1)}} \cdots b_{m,i_{\sigma(m)}} \\
&= \sum_{1 \leq i_1 < \dots < i_m \leq n} \det(a_{i_1}, \dots, a_{i_m}) \underbrace{\sum_{\sigma \in S_m} \operatorname{sign}(\sigma) \cdot b_{1,i_{\sigma(1)}} \cdots b_{m,i_{\sigma(m)}}}_{\text{Leibnizformel für } (b_{i_1}, \dots, b_{i_m})} \\
&= \sum_{1 \leq i_1 < \dots < i_m \leq n} \det(a_{i_1}, \dots, a_{i_m}) \det(b_{i_1}, \dots, b_{i_m}) \quad \square
\end{aligned}$$

*Ein anderer Beweis des Satzes von Binet-Cauchy.* Wir bezeichnen als  $A_{I,J}$  die Untermatrix zu den Zeilen bzw. Splaten, die durch  $I$  bzw.  $J$  indexiert

sind. des Weiteren seien  $[m] := \{1, \dots, m\}$  und  $[n] := \{1, \dots, n\}$  und sei  $\binom{[n]}{m}$  die Menge aller  $m$ -elementigen Teilmengen von  $[n]$ . Dann kann die Binet-Cauchy-Formel als

$$\det(AB^\top) = \sum_{I \in \binom{[n]}{m}} \det(A_{[m], I}) \det(B_{[m], I})$$

formuliert werden. Wir können  $AB^\top$  mit Hilfe der Zeilen  $B_{1,[n]}, \dots, B_{m,[n]}$  von  $B$  wie folgt darstellen:

$$(AB_{1,[n]}^\top, \dots, AB_{m,[n]}^\top).$$

Da die Determinante linear in den Spalten ist zeigt das, dass

$$\det(AB^\top) = \det(AB_{1,[n]}^\top, \dots, AB_{m,[n]}^\top)$$

als Funktion der Spalten von  $B$  in jeder der  $m$  Zeilen von  $B$  linear ist. Des Weiteren ist die Funktion alternierend in der Zeilen von  $B$ , weil die Determinante alternierend ist.

Das Gleiche gilt für die rechte Seite der Formel, die wir beweisen, die rechte Seite ist linear in den Zeilen von  $B$  und alternierend, denn

$$\sum_{I \in \binom{[n]}{m}} \det(A_{[m],I}) \det(B_{[m],I}) = \sum_{I \in \binom{[n]}{m}} \det(A_{[m],I}) \det(B_{1,I}, \dots, B_{m,I}).$$

Aus der multilinearität in den Zeilen von  $B$  folgt, dass die Formel im allgemeinen Fall gilt, sobald wir sie im Fall verifizieren, bei dem  $B$  Vektoren aus der Standardbasis von  $\mathbb{K}^n$  sind: also für  $B_{1,[n]}, \dots, B_{m,[n]} \in \{e_1^\top, \dots, e_n^\top\}$ . Sei also  $B_{i,[n]} = e_{k_i}^\top$  mit  $k_1, \dots, k_m \in [n]$ . Da die beiden Seiten alternierend sind, kriegen wir auf der linken und rechten Seite eine 0, wenn  $B$  zwei identische Zeilen hat. Wir können also annehmen, dass  $k_1, \dots, k_m$  paarweise verschieden sind. Da die linken und die rechten Seiten alternierend sind, verursacht das vertauschen von zwei Zeilen von  $B$  eine Vorzeichenänderung auf beiden Seiten. Wir können also durch das Vertauschen der Zeilen, die Formel auf den Fall  $k_1 < \dots < k_m$  zurückführen. Nach diesen Veränderungen ist die

linke Seite

$$\begin{aligned}
\det(AB^\top) &= \det(AB_{1,[n]}^\top \cdots AB_{m,[n]}^\top) \\
&= \det(A(e_{k_1} \cdots e_{k_m})) \\
&= \det(Ae_{k_1}, \dots, Ae_{k_m}) \\
&= \det(A_{[m],\{k_1, \dots, k_m\}}).
\end{aligned}$$

Die rechte Seite ist

$$\sum_{I \in \binom{[n]}{m}} \det(A_{[m],I}) \det(B_{1,I}, \dots, B_{m,I}) = \sum_{I \in \binom{[n]}{m}} \det(A_{[m],I}) \det((e_{k_1} \cdots e_{k_m})_{I,[m]}),$$

wobei die Matrix  $(e_{k_1} \cdots e_{k_m})_I$  im Fall  $I \neq \{k_1, \dots, k_m\}$  die Determinante 0 hat, weil dann ein Element  $i \in I \setminus \{k_1, \dots, k_m\}$  existiert, woraus folgt, dass die genannte Matrix eine Nullzeile besitzt. Auf der rechten Seite bleibt also nur der Term für  $I = \{k_1, \dots, k_m\}$ , un dieser Term ist gleich  $\det(A_{[m],\{k_1, \dots, k_m\}})$ .  $\square$

**Bem.** Besonders interessant ist der Fall  $A = B$  des Satzes Binet-Cauchy. In diesem Fall ist die rechte Seite der Formel die Summe der Quadrate aller  $m \times m$  Minoren von  $A$ . Wenn unser Körper der Körper der reellen Zahlen ist, wissen wir also, dass  $\det(AA^\top)$  ein nicht-negativer Wert ist. Wir können also die Wurzel ziehen und kommen zur Gleichung

$$\sqrt{\det(AA^\top)} = \sqrt{\sum_{1 \leq i_1 < \dots < i_m \leq n} \det(a_{i_1}, \dots, a_{i_m})^2}.$$

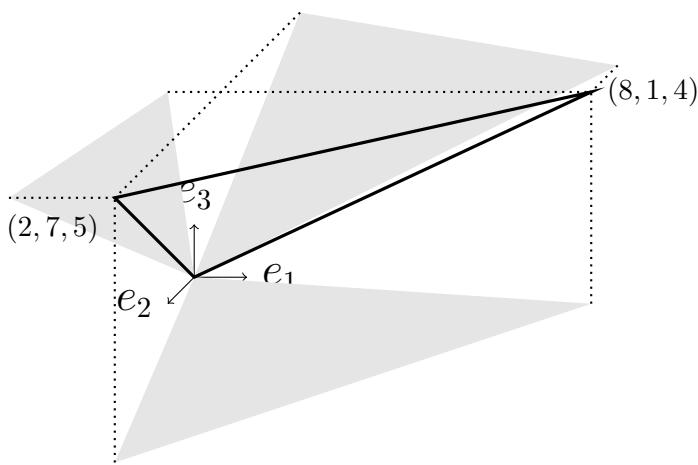
Die vorige Gleichung hat eine Geometrische Bedeutung. Sie ist im gewissen Sinne eine weitreichende Verallgemeinerung des Satzes von Pythagoras, über den Sie in der Schule gehört haben könnten.

Sind  $z_1, \dots, z_m$  die Zeilen von  $A$ , dann ist der Wert  $\sqrt{\det(AA^\top)}$  das  $m$ -dimensionale Volumen des sogenannten Hyperparallelepipeds  $[0, 1]z_1 + \dots + [0, 1]z_m = \{\sum_{i=1}^m \alpha_i z_i : \alpha_i, \dots, \alpha_m \in [0, 1]\}$ , das durch die Vektoren  $z_1, \dots, z_m$  aufgespannt ist. Die Formel von Binet-Cauchy beschreibt das Volumen des  $m$ -dimensionalen Hyperparallelepipeds in der Dimension  $n$

durch die Volumina seiner ‘Schatten’ beim Projizieren auf  $m$ -dimensionale Räume, welche durch die Standardbasisvektoren aufgespannt sind.

Am einfachsten lässt sich der Fall von Parallelogrammen in der Dimension drei veranschaulichen. Wir sehen die Photos einer Scheibe von vorne, von der Seite und von oben, messen die Flächen der Bilder der Scheibe auf allen drei Fotos und wollen anhand dieser Flächen die Fläche der Scheibe ausrechnen. Das geht tatsächlich: die Fläche der Scheibe ist die Wurzel aus der Summe der Quadranten der Flächen der Bilder.

**Bsp.** Berechnen wir die Fläche des Dreiecks mit Ecken  $(0, 0, 0)$ ,  $(2, 7, 5)$  und  $(8, 1, 4)$  ist mit der Verwendung der linken Seite der Binet-Cauchy-Formel.



Die Projektionen auf die drei Koordinatenebenen sind die drei Dreiecke:

- In Richtung  $e_1$ : Dreieck  $T_1$  mit den Ecken  $(0, 0)$ ,  $(7, 5)$  und  $(1, 4)$ ,
- In Richtung  $e_2$ : Dreieck  $T_2$  mit den Ecken  $(0, 0)$ ,  $(2, 5)$  und  $(8, 4)$ ,
- In Richtung  $e_3$ : Dreieck  $T_3$  mit den Ecken  $(0, 0)$ ,  $(2, 7)$  und  $(8, 1)$ .

Die Flächen der drei Dreiecke ergeben sich aus den drei  $2 \times 2$  Minoren der

Matrix

$$A = \begin{pmatrix} 2 & 7 & 5 \\ 1 & 8 & 4 \end{pmatrix}$$

Die drei Minoren sind

$$\begin{vmatrix} 7 & 5 \\ 8 & 4 \end{vmatrix} = -12, \quad \begin{vmatrix} 2 & 5 \\ 1 & 4 \end{vmatrix} = 3, \quad \begin{vmatrix} 2 & 7 \\ 1 & 8 \end{vmatrix} = 9.$$

Die Fläche von  $T_1$ ,  $T_2$  und  $T_3$  ist somit jeweils  $6$ ,  $\frac{3}{2}$  und  $\frac{9}{2}$ . Die Fläche von unserem Dreieck in der Dimension drei ist somit

$$\sqrt{6^2 + \left(\frac{3}{2}\right)^2 + \left(\frac{9}{2}\right)^2} = \sqrt{58,5}.$$

Man kommt auf den selben Wert, wenn man die Fläche unseres Dreiecks mit Hilfe der linken Seite der Binet-Cauchy-Formel als  $\frac{1}{2}\sqrt{\det(AA^\top)}$  berechnet. (Die Berechnung durch die Fläche der drei Projektion haben wir hier zur Veranschaulichung der Binet-Cauchy-Formel gegeben. Es ist keine Empfehl-

lung: man muss die Fläche nicht unbedingt auf diese Weise berechnen, die Verwendung von  $\sqrt{\det(AA^\top)}$  führt einen schneller zum Ziel.)

### 5.3.6 Die Formel von Sylvester

**Thm.** Seien  $A \in \mathbb{K}^{m \times n}$  und  $B \in \mathbb{K}^{n \times m}$ . Dann gilt  $\det(I_m + AB) = \det(I_n + BA)$ .

*Beweis.* Es gilt

$$\begin{aligned} \begin{pmatrix} I_m & 0 \\ B & I_n \end{pmatrix} \begin{pmatrix} I_m & 0 \\ 0 & I_n - BA \end{pmatrix} \begin{pmatrix} I_m & A \\ 0 & I_n \end{pmatrix} &= \begin{pmatrix} I_m & A \\ B & I_n \end{pmatrix} \\ &= \begin{pmatrix} I_m & 0 \\ B & I_n \end{pmatrix} \begin{pmatrix} I_m - AB & 0 \\ 0 & I_n \end{pmatrix} \begin{pmatrix} I_m & A \\ 0 & I_n \end{pmatrix} \end{aligned}$$

Durch die Anwendung von  $\det$  zu dieser Gleichung erhalten wir die gewünschte Gleichung.  $\square$

# 6 Eigenwerte und Eigenvektoren

Die Rolle der Eigenwerte und Eigenvektoren in der linearen Algebra und deren Anwendungen innerhalb und außerhalb der Mathematik ist absolut zentral. Egal ob man ein lineares Gleichungssystem auf Stabilität analysiert, ein System von linearen Differentialgleichungen löst, Eigenschaften einer Markov-Kette untersucht, oder ein iteratives Verfahren benutzt, um ein nichtlineares Optimierungsproblem zu lösen, braucht man Eigenwerte und Eigenvektoren. Eigenwerte und Eigenvektoren geben einem in der Regel die wichtigste Information über die zugrundeliegende lineare Abbildung eines Vektorraums bzw. über die zugrundeliegende Matrix.

## 6.1 Grundlagen

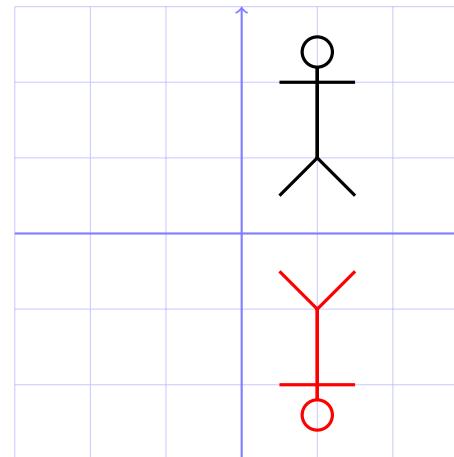
In diesem Abschnitt geben wir die Grunddefinitionen Eigenwert und Eigenvektor und beweisen die ersten, nützlichen Resultate darüber.

### 6.1.1 Beispiele und Motivation

Das Ziel, dass man bei der Berechnung von Eigenwerten und Eigenvektoren verfolgt ist, eine gegebene lineare Abbildung eines endlichdimensionalen Vektorraums als „Zusammensetzung“ von linearen Abbildungen auf Vektorräumen einer kleineren Dimension darstellen. Das geht nicht immer, aber es gibt Situationen, in denen man sogar die gegebene lineare Abbildung als Zusammensetzung der linearen Abbildungen eindimensionaler Vektorräume darstellen kann (das ist der günstigste Fall). Etwas formaler beschrieben, ist man auf der Suche nach einer Basis  $\mathcal{B}$  für eine gegebene lineare Abbildung  $F : V \rightarrow V$  eines endlichdimensionalen Vektorraums  $V$ , in der die Matrix  $F_{\mathcal{B}}$  der Abbildung eine möglichst einfache Struktur hat. Am glücklichsten ist man, wenn man eine Basis findet, in der die Matrix der Abbildung diagonal ist.

Beispiele:

- (i) Die Spiegelung  $x = (x_1, x_2) \mapsto (x_1, -x_2)$  auf  $\mathbb{K}^2$  bzgl. der  $x_1$ -Achse zerfällt in die folgenden zwei lineare Abbildungen eindimensionaler Unterräume von  $\mathbb{K}^2$ :

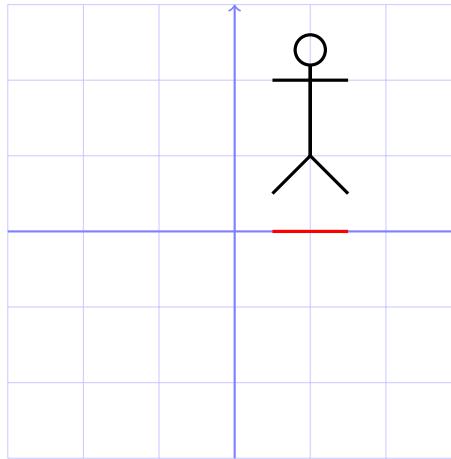


- (a) die Identische Abbildung  $x \mapsto x$  auf der Geraden  $\text{lin}(e_1) = \mathbb{K} \times \{0\}$  und  
(b) die Spiegelung  $x \mapsto -x$  auf der Geraden  $\text{lin}(e_2) = \{0\} \times \mathbb{K}$ .

Diese Abbildung des Raums  $\mathbb{K}^2$  ist also eine „Zusammensetzung“ von

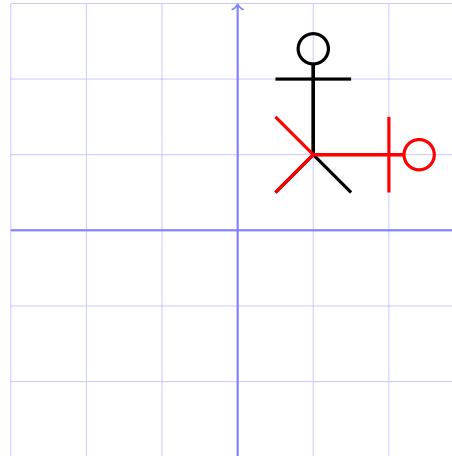
zwei linearen Abbildungen von eindimensionalen Räumen. Mit anderen Worten: die Abbildung schickt  $e_1$  auf  $e_1$  und  $e_2$  auf  $-e_2$ . Die Abbildung ist durch diese Vorgabe eindeutig beschrieben, da eine lineare Abbildung durch die Angabe ihrer Wirkung auf einer Basis eindeutig festgelegt ist. Da  $e_1$  auf  $e_1$  geschickt wird, fällt man bei der Anwendung der Abbildung auf den Vektoren aus der Geraden  $\text{lin}(e_1)$  nicht aus dieser Geraden aus. Ebenso fällt man nicht aus der Geraden  $\text{lin}(e_2)$  aus.

- (ii) Ein weiteres ähnliches Beispiel ist die Projektion  $x = (x_1, x_2) \mapsto (x_1, 0)$  auf die  $x_1$ -Achse in  $\mathbb{K}^2$ . Diese zerfällt in
  - (a) Die identische Abbildung  $x \mapsto x$  auf der  $x_1$ -Achse  $\text{lin}(e_1) = \mathbb{K} \times \{0\}$  und
  - (b) Die Nullabbildung  $x \mapsto 0$  auf der  $x_2$ -Achse  $\text{lin}(e_2) = \{0\} \times \mathbb{K}$ .



- (iii) Bei den vorigen beiden Abbildungen war die Zerlegung direkt erkennbar, weil die zugrundeliegenden Untervektorräume von  $\mathbb{K}^2$  Koordinatenachsen waren. Hier noch ein Beispiel, in dem man ebenfalls eine Zerlegung findet, aber nicht bzgl. der Koordinaten Achsen. Wir betrachten die Abbildung  $x = (x_1, x_2) \mapsto (x_2, x_1)$  auf  $\mathbb{K}^2$ . Diese Abbildung ist wie auch das erste Beispiel ebenfalls eine Spiegelung.
- (a)  $x \mapsto x$  auf der Geraden  $\{(x_1, x_2) \in \mathbb{K}^2 : x_1 = x_2\}$  und

(b)  $x \mapsto -x$  auf der Geraden  $\{(x_1, x_2) \in \mathbb{K}^2 : x_1 = -x_2\}$ .



(iv) Nun betrachten wir ein Links-Shift der Koordinaten  $(x_1, x_2) \mapsto (x_2, 0)$  auf  $\mathbb{K}^2$ . Die Vektoren in der  $x_1$ -Achse bleiben nach der Anwendung der Abbildung in der  $x_1$ -Achse, denn sie werden auf 0 abgebildet:  $x \mapsto 0$  auf  $\text{lin}(e_1) = \mathbb{K} \times \{0\}$ . Man findet aber keinen einen anderen eindimensionalen Untervektorraum von  $\mathbb{K}^2$  mit der Eigenschaft, dass die Vektoren dieses Untervektorraums nach der Anwendung der Abbil-

dung im fixierten Untervektorraum bleiben. Diese Abbildung lässt sich also nicht als Zusammensetzung von Abbildungen eindimensionaler Untervektorräume darstellen.

### 6.1.2 Definition von Eigenwerten und Eigenvektoren

Nun haben wir genug Motivation für die Begriffe Eigenwert und Eigenvektor. Sei  $V$  Vektorraum über  $\mathbb{K}$  und sei  $F : V \rightarrow V$  lineare Abbildung. Ein Paar  $(\lambda, v)$  aus einem Skalar  $\lambda \in \mathbb{K}$  und einem Nichtnullvektor  $v \in V \setminus \{0\}$  mit der Eigenschaft

$$F(v) = \lambda v \tag{6.1.1}$$

heißt *Eigenpaar* von  $F$ . Dabei heißt  $\lambda$  *Eigenwert* von  $F$  und  $v$  *Eigenvektor* zum Eigenwert  $\lambda$  von  $F$ .

Sie können nun Beispiele aus 6.1.1 nochmals betrachten und alle Eigenpaare der in 6.1.1 angegebenen linearen Abbildungen finden. Wenn wir in einem Eigenpaar  $(\lambda, v)$  den Eigenvektor  $v$  mit einem Nichtnullwert  $\alpha \in$

$\mathbb{K} \setminus \{0\}$  multiplizieren, so erhalten wir ein Eigenpaar  $(\lambda, \alpha v)$ , das sich im Wesentlichen von dem ursprünglichen Eigenpaar  $(\lambda, v)$  nicht unterscheidet. Mit anderen Worten ist nur die Richtung eines Eigenvektors wichtig. Die genaue Skalierung des Eigenvektors  $v$  spielt keine Rolle.

Analog werden diese Begriffe Eigenpaar, Eigenwert und Eigenvektor für eine quadratische Matrix eingeführt. Die Eigenwerte, Eigenvektoren und Eigenpaare von  $A \in \mathbb{K}^{n \times n}$  sind jene der zugehörigen Abbildung  $x \mapsto Ax$ , oder ausformuliert:  $(\lambda, v)$  mit  $\lambda \in \mathbb{K}$  und  $v \in \mathbb{K}^n \setminus \{0\}$  ist *Eigenpaar* von  $A$ , wenn die Bedingung  $Av = \lambda v$  erfüllt ist.

**Bem.** Wie bereits in Teil 1 des Kurses erwähnt ermöglichen Matrizen eine konkrete Beschreibung der linearen Abbildung im “konkreten Raum”  $\mathbb{K}^n$ . Da man im Raum  $\mathbb{K}^n$  die Standardbasis  $e_1, \dots, e_n$  hat, ist es naheliegend die Wirkung einer Abbildung des Raums  $\mathbb{K}^n$  genau in dieser Basis festzuliegen. Genau dass macht die Matrix  $A$  der Abbildung  $x \mapsto Ax$ . Die Standardbasis ist aber nicht notwendigerweise eine Basis, in der man die zugrundeliegen-

de Abbildung am besten versteht. Das ist einer der Gründe, warum der abstrakte Zugang zur linearen Algebra durch allgemeine Vektorräume  $V$  günstig ist: man legt keine feste Basis fest. Des Weiteren muss man noch bedenken, dass nicht jeder Vektorraum eine Standardbasis hat. Tatsächlich hat zum Beispiel die Ebene  $E := \{(x, y, z) : x + 2y - 3z = 0\} \subseteq \mathbb{R}^3$  keine Standardbasis, man kann aber stets eine Basis  $b_1, b_2$  wählen und danach diese Ebene als  $\mathbb{R}^2$  behandeln. Man kann z.B. eine 90-Grad Drehung  $F : E \rightarrow E$  in der Ebene  $E$  betrachten und analysieren. Wenn wir die lineare Algebra ausschließlich konkret auf der Basis von Vektoren aus  $\mathbb{K}^n$  und Matrizen entwickelt hätten, hätten wir keine Sprache zur Verfügung zur Einführung und der Behandlung der Drehung  $F$  wie oben.

**Bem.** Es ist klar, dass der Begriff vom Kern  $\ker(F)$  einer linearen Abbildung  $F : V \rightarrow V$  mit Eigenvektoren zusammenhängt. Nichtnullvektoren aus dem Kern sind genau die Eigenwerte von  $F$  zum Eigenwert 0.

**Bem.** Ein Eigenvektor  $v$  zum Eigenwert 1 hat die Eigenschaft  $F(v) = v$ , das

heißt,  $v$  ist ein sogenannter Fixpunkt der Abbildung  $v$ . Solche Fixpunkte sind oft von einem besonderen Interesse. Der Vektor der Google-Pageranks der Internetseiten ist ein Fixpunkt einer linearen Abbildung, die durch die Vernetzung der Internetseiten bestimmt ist.

**Bem.** Die Aufgaben  $Ax = b$  (Lösung von linearen Gleichungssystemen) und  $Av = \lambda v$  (Eigenwertaufgabe) sind die zwei Grund-Rechenaufgaben der linearen Algebra. In der Numerik werden Sie verschiedene Methoden zur Lösung dieser zwei Aufgaben lernen. Die beiden Aufgaben sind nämlich der Hauptgegenstand der Numerischen Linearen Algebra.

Hierbei wählt man  $\mathbb{C}$  als den zugrundeliegenden Körper und führt die Berechnungen in der Praxis annähernd mit Hilfe der sogenannten Gleitkommazahlen. Die Aufgabe  $Av = \lambda v$  kann etwas konkreter in verschiedenen Varianten gestellt werden. In der numerischen Mathematik interessiert man sich oft für den Eigenwert  $\lambda$  mit dem höchsten Betrag und einen zugehörigen Eigenvektor. Alle Rechenaufgaben aus der linearen Algebra, die wir

bis jetzt betrachtet haben, konnten wir durch das Gauß-Verfahren (bzw. eine Variation davon) lösen. Die Eigenwertaufgabe ist aber die Aufgabe, bei der das Gauß-Verfahren **nicht** direkt hilft (man braucht weitere Ideen). Die Aufgabe ist eigentlich nicht linear, denn  $\lambda$  ist unbekannt,  $v$  ist ebenfalls unbekannt und die Aufgabe  $Av = \lambda v$  enthält das **Produkt** dieser unbekannten Objekte. Sobald man das Unbekannte  $v$  oder  $\lambda$  bestimmt hat, ist man im Wesentlichen durch. Kennt man das  $v$ , so ermittelt man  $\lambda$  als Streckungsfaktor der parallelen Vektoren  $Av$  und  $v$ . Hat man  $\lambda$  ermittelt, so reduziert sich die Bestimmung eines zugehörigen Eigenvektors zur Lösung des LGS  $Av = \lambda v$  für einen unbekannten Vektor  $v$ .

### 6.1.3 Diagonalisierbarkeit von linearen Abbildungen

Hier und im Folgenden sei  $n \in \mathbb{N}$ . Sei  $V$  ein  $n$ -dimensionaler Vektorraum über  $\mathbb{K}$ . Eine Abbildung  $F \in \text{Lin}(V)$  heißt diagonalisierbar, wenn eine Basis  $\mathcal{B}$  von  $V$  existiert, für welche die Matrix  $F_{\mathcal{B}}$  diagonal ist. Die Diagonalisier-

baren Abbildungen sind die einfachsten Abbildungen in der Theorie der Eigenwerte und Eigenvektoren.

Die nachfolgende Proposition formuliert Diagonalisierbarkeit in der Sprache der Eigenwerte und Eigenvektoren.

**Prop.** *Sei  $V$  ein  $n$ -dimensionaler Vektorraum und sei  $F \in \text{Lin}(V)$ . Dann sind die folgenden Bedingungen äquivalent:*

- (i)  *$F$  ist diagonalisierbar.*
- (ii)  *$V$  besitzt eine Basis aus Eigenvektoren von  $F$ .*

*Beweis.* Um diese Proposition herzuleiten, reicht es aus, die Definition von  $F_{\mathcal{B}}$  auszuschreiben. Die Matrix  $F_{\mathcal{B}}$  der Abbildung  $F$  bzgl. der Basis  $\mathcal{B} = (b_1, \dots, b_n)$  entsteht folgendermaßen: man wenden  $F$  zu den Basisvektoren  $b_1, \dots, b_n$  und schreibt das Ergebnis jedes mal wieder in der Basis  $b_1, \dots, b_n$  auf. Die  $i$ -te Spalte enthält die Darstellung von  $F(b_i)$  in der Basis  $b_1, \dots, b_n$ . Die Eigenschaft, dass  $F_{\mathcal{B}}$  diagonal mit Diagonalelementen

$\lambda_1, \dots, \lambda_n$  ist, bedeutet als, dass  $f(b_i) = \lambda_i b_i$  erfüllt. Das heißt,  $\mathcal{B}$  ist Basis aus Eigenvektoren von  $F$ , und die jeweiligen Diagonalelemente sind die entsprechenden Eigenwerte.  $\square$

**Bsp** (aus 6.1.1). Sei  $F(x_1, x_2) = (x_2, x_1)$ . Dann ist  $b_1 = (1, 1)$  Eigenvektor zum Eigenwert 1,  $b_2 = (-1, 1)$  Eigenvektor zum Eigenwert -1 und  $\mathcal{B} = (b_1, b_2)$  eine Basis von  $\mathbb{Q}^2$ . D.h.

$$F_{\mathcal{B}} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

**Bsp** (aus 6.1.1). Sei  $F(x_1, x_2) = (x_2, 0)$ . Dies ist ein Beispiel einer “problematischen” Abbildung, welche man nicht diagonalisieren kann. Zeigen wir das durch ein Widerspruchsargument. Angenommen,  $F$  hätte eine Basis  $b_1, b_2$  aus Eigenvektoren mit den jeweiligen Eigenwerten  $\lambda_1$  und  $\lambda_2$ . Die zweifache Anwendung von  $F$  zu einem beliebigen Vektor aus  $\mathbb{K}^2$  ergibt null ( $(x_1, x_2) \mapsto (x_2, 0) \mapsto (0, 0)$ ). Das heißt  $F^2(x) = 0$  für alle  $x$ . Wir haben

aber  $F(b_i) = \lambda_i b_i$  und somit  $F^2(b_i) = F(\lambda_i b_i) = \lambda_i F(b_i) = \lambda_i^2 b_i$ . Also ist  $\lambda_i^2 b_i = 0$ . Da  $b_i$  kein Nullvektor ist, folgt  $\lambda_i^2 = 0$ . Daraus folgt  $\lambda_i = 0$ . Wir haben  $\lambda_1 = \lambda_2 = 0$ . Unser Abbildung schickt somit den Basisvektor  $b_i$  auf den Vektor  $0b_i = 0$ . Somit ist  $F$  eine Nullabbildung, das widerspricht aber zu  $F(e_2) = e_1 \neq 0$ .

Durch die vorigen Beispiele sehen wir, dass man unproblematische (diagonalisierbare) sowie problematische (nichtdiagonalisierbare) Abbildungen hat. Wir sind an der Theorie interessiert, die uns helfen würde, die beiden Fälle auseinanderzuhalten. Im Rest des Kapitels werden Werkzeuge dazu entwickelt.

Im Rest dieses Abschnitts besprechen wir noch kurz die Diagonalisierbarkeit von quadratischen Matrizen. Eine Matrix  $A \in \mathbb{K}^{n \times n}$  heißt diagonalisierbar, wenn die lineare Abbildung  $x \mapsto Ax$  auf  $\mathbb{K}^n$  diagonalisierbar ist, mit anderen Worten:  $\mathbb{K}^n$  besitzt eine Basis  $b_1, \dots, b_n$  mit  $Ab_i = \lambda_i b_i \forall i \in \{1, \dots, n\}$ , wobei  $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ . Aus dieser Definition wird

deutlich, dass die Eigenschaft von der Wahl des Körpers abhängig ist. Eine Matrix  $A \in \mathbb{R}^{n \times n}$  ist auch eine Matrix aus  $\mathbb{C}^{n \times n}$ . Es gibt Fälle, in den  $A$  bzgl.  $\mathbb{R}$  nicht diagonalisierbar ist, aber bzgl. des größeren Körpers  $\mathbb{C}$  diagonalisierbar. Mehr dazu später.

**Prop.** *Sei  $A \in \mathbb{K}^{n \times n}$ . Die Matrix  $A$  ist genau dann diagonalisierbar (bzgl.  $\mathbb{K}$ ), wenn eine invertierbare Matrix  $B \in \mathbb{K}^{n \times n}$  existiert, sodass  $B^{-1}AB$  diagonal ist.*

*Beweis.* Wie der vorige Beweis brauchen wir auch hier, lediglich die Diagonalisierbarkeit in eine andere Sprache zu übersetzen, und zwar in die Sprache der Matrizen. Sei  $A$  diagonalisierbar, d.h. es existiert eine Basis  $b_1, \dots, b_n$  von  $\mathbb{K}^n$  mit  $Ab_i = \lambda_i b_i \quad \forall i \in \{1, \dots, n\}$ , wobei  $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ .

Sei  $B = (b_1, \dots, b_n)$ . Dann ist

$$A \begin{pmatrix} | & & | \\ b_1 & \cdots & b_n \\ | & & | \end{pmatrix} = \begin{pmatrix} | & & | \\ \lambda_1 b_1 & \cdots & \lambda_n b_n \\ | & & | \end{pmatrix} = \begin{pmatrix} | & & | \\ b_1 & \cdots & b_n \\ | & & | \end{pmatrix} \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

$B$  ist invertierbar, weil die Spalten von  $B$  eine Basis bilden (vgl. LA I). Wir multiplizieren die vorige Gleichung mit  $B^{-1}$  von links und erhalten

$$B^{-1}AB = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix} \quad (6.1.2)$$

Umgekehrt: sei  $B \in \mathbb{K}^{n \times n}$  eine invertierbare Matrix, für welche  $B^{-1}AB$  diagonal ist, d.h. (6.1.2) gilt mit  $\lambda_1, \dots, \lambda_n \in \mathbb{K}$ . Es folgt:  $b_1, \dots, b_n$  ist eine Basis von  $\mathbb{K}^n$  mit  $Ab_i = \lambda_i b_i \ \forall i \in \{1, \dots, n\}$ .  $\square$

Motiviert durch die vorige Proposition führen wir den Begriff der Ähnlichkeit von Matrizen ein. Matrix  $A \in \mathbb{K}^{n \times n}$  und  $\tilde{A} \in \mathbb{K}^{n \times n}$  heißen *ähnlich*,

wenn eine invertierbare Matrix  $B \in \mathbb{K}^{n \times n}$  existiert mit  $B^{-1}AB = \tilde{A}$ . Die Ähnlichkeit von Matrizen ist eine Äquivalenzrelation (Zeigen Sie das).

**Bsp.**  $\tilde{A} = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, A = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \in \mathbb{Q}^2$ . Seien  $b_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, b_2 = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ .

Dann ist  $Ab_1 = b_1, Ab_2 = -b_2$ . D.h. für  $B = (b_1, b_2)$  ist  $AB = B\tilde{A} \Leftrightarrow B^{-1}AB = \tilde{A}$ . Die Matrix  $\tilde{A}$  beschreibt eine Spiegelung an der  $x_2$ -Achse, die Matrix  $A$  eine Spiegelung an der Achse  $x_1 = x_2$ . Die Matrizen machen also tatsächlich was Ähnliches.

In Bezug auf die Theorie der Eigenwerte und Eigenvektoren unterscheiden sich zwei verschiedene ähnliche Matrizen voneinander nicht (d.h., sie haben komplett die gleichen Eigenschaften in dieser Theorie).

### 6.1.4 Diagonalisierbarkeit: eine hinreichende Bedingung

Wir diskutieren hier eine Bedingung, die in vielen Situationen hilft, Diagonalisierbarkeit zu erkennen. Bei der Diagonalisierung geht es um die Suche nach einer Basis, in der die Abbildung diagonal ist. Die Vektoren der Basis sollen Eigenvektoren sein. Wir sind als auf der Suche nach linear unabhängigen Eigenvektoren. Das folgende Lemma zeigt, dass wir im Fall von verschiedenen Eigenwerten, die lineare Unabhängigkeit geschenkt bekommen.

**Lem.** *Sei  $F : V \rightarrow V$  lineare Abbildung über einem Vektorraum  $V$  und seien  $(\lambda_1, v_1), \dots, (\lambda_m, v_m)$  Eigenpaare von  $F$  mit der Eigenschaft, dass  $\lambda_1, \dots, \lambda_m$  paarweise verschieden sind. Dann sind  $v_1, \dots, v_m$  linear unabhängig.*

*Beweis.* Induktion über  $m \in \mathbb{N}$ . Für  $m = 1$  ist die Behauptung trivial. Angenommen, die Behauptung gilt für  $m - 1$  Eigenpaare mit  $m \geq 2$ .

Wir betrachten nun  $m$  Eigenpaare mit  $m$  verschiedenen Eigenwerten. Seien  $\alpha_1, \dots, \alpha_m \in \mathbb{K}$  beliebige Skalare mit

$$\alpha_1 v_1 + \dots + \alpha_m v_m = 0 \quad (6.1.3)$$

Wir multiplizieren  $(6.1.3)$  mit  $\lambda_m$  und erhalten:

$$\lambda_m \alpha_1 v_1 + \dots + \lambda_m \alpha_{m-1} v_{m-1} + \lambda_m \alpha_m v_m = 0$$

Die Anwendung von  $F$  zu  $(6.1.3)$  ergibt:

$$\lambda_1 \alpha_1 v_1 + \dots + \lambda_m \alpha_m v_m = 0$$

In der Differenz der vorigen zwei Gleichungen verschwindet der Term mit  $v_m$ :

$$(\lambda_1 - \lambda_m) \alpha_1 v_1 + \dots + (\lambda_{m-1} - \lambda_m) \alpha_{m-1} v_{m-1} = 0$$

Nach Induktionsvoraussetzung sind  $v_1, \dots, v_{m-1}$  linear unabhängig. Also folgt aus der vorigen Gleichung, dass alle Koeffizienten der linearen Kom-

bination auf der linken Seite gleich 0 sind:

$$(\lambda_1 - \lambda_m)\alpha_1 = \dots = (\lambda_{m-1} - \lambda_m)\alpha_{m-1} = 0$$

Da die Eigenwerte paarweise verschieden sind, ist  $\lambda_i - \lambda_m \neq 0$  für alle  $i \in \{1, \dots, m-1\}$ , erhält man

$$\alpha_1 = \dots = \alpha_{m-1} = 0$$

Das Einsetzen in (6.1.3) ergibt  $\alpha_m v_m = 0$ . Mit  $v_m \neq 0$  folgt  $\alpha_m = 0$ . Wir haben aus (6.1.3)  $\alpha_1 = \dots = \alpha_m = 0$  erhalten. D.h.,  $v_1, \dots, v_m$  sind linear unabhängig.  $\square$

**Thm.** Sei  $F : V \rightarrow V$  lineare Abbildung eines  $n$ -dimensionalen Vektorraums  $V$ , welche  $n$  paarweise verschiedene Eigenwerte  $\lambda_1, \dots, \lambda_n \in \mathbb{K}$  besitzt. Dann ist  $F$  diagonalisierbar.

*Beweis.* Das vorige Lemma ergibt, dass die Eigenvektoren  $b_1, \dots, b_n$  zu den Eigenwerten  $\lambda_1, \dots, \lambda_n$  eine Basis von  $V$  bilden. Die Matrix von  $F$  bzgl. dieser Basis ist diagonal.  $\square$

Wann ist eine hinreichende Bedingung für eine gewünschte Eigenschaft nützlich? In der Regel wünscht man sich eine möglichst allgemeine hinreichende Bedingung, also eine Bedingung, die nicht zu restriktiv ist. Aus dieser Perspektive ist die oben präsentierte hinreichende Bedingung der Diagonalisierbarkeit zumindest bzgl. des Körpers der komplexen Zahlen **sehr** nützlich. Denn diese Bedingung ist bzgl.  $\mathbb{C}$  generisch erfüllt: wenn Sie etwa eine Matrix  $A$  aus  $\mathbb{C}^{n \times n}$  generieren, in dem Sie die  $n^2$  Komponenten von  $A$  aus einem Bereich, etwa aus dem Segment  $[-1, 1]$  gleichmäßig und unabhängig ziehen, dann kriegen Sie mit Wahrscheinlichkeit 1 eine Matrix mit  $n$  paarweise verschiedenen Eigenwerten in  $\mathbb{C}$ . “Generisch erfüllt” ist aber nicht nicht das selbe wie (bedingungslos) “erfüllt”. Wenn sie zwei unabhängige gleichmäßig verteilte Werte  $x, y$  aus  $[0, 1]$  ziehen, dann ist die Eigenschaft  $x \neq y$  generisch erfüllt (es ist also unwahrscheinlich, dass  $x = y$  gilt). Es gibt zwar Werte  $x, y \in [0, 1]$  mit  $x = y$ , diese Werte bilden eine eindimensionale Ausnahme im Quadrat aller Paare  $(x, y) \in [0, 1]^2$ .

### 6.1.5 Eigenraum

Alle Eigenwerte zu einem festen Eigenwert plus der Nullvektor bilden einen Untervektorraum, den sogenannten Eigenraum. Sei  $F : V \rightarrow V$  lineare Abbildung eines Vektorraums  $V$  und sei  $\lambda \in \mathbb{K}$ . Dann heißt

$$\text{Eig}(F, \lambda) := \{v \in V : F(v) = \lambda v\} \quad (6.1.4)$$

der *Eigenraum* von  $F$  bzgl.  $\lambda$ . Wir setzen in dieser Definition nicht voraus, dass  $\lambda$  Eigenwert von  $F$  ist. Wenn  $\lambda$  kein Eigenwert ist, dann enthält  $\text{Eig}(F, \lambda)$  nichts außer dem Nullvektor. Fassen wir die Grundeigenschaften der Eigenräume zusammen:

**Prop.** Für eine lineare Abbildung  $F : V \rightarrow V$  eines Vektorraums  $V$  und ein  $\lambda \in V$  gilt:

- (i)  $\text{Eig}(F, \lambda)$  ist Untervektorraum von  $V$ .
- (ii)  $\lambda$  ist genau dann Eigenwert von  $F$ , wenn  $\text{Eig}(F, \lambda) \neq \{0\}$  gilt.

(iii) Wenn  $\lambda$  Eigenwert von  $F$  ist, dann gilt für  $v \in V$ :

$$v \text{ ist Eigenvektor von } F \text{ zu } \lambda \iff v \in \text{Eig}(F, \lambda) \setminus \{0\}$$

(iv)  $\text{Eig}(F, \lambda) = \ker(\lambda \text{id}_V - F)$

(v) Wenn  $\lambda_1, \lambda_2 \in \mathbb{K}$  verschieden sind, dann gilt  $\text{Eig}(F, \lambda_1) \cap \text{Eig}(F, \lambda_2) = \{0\}$ .

*Beweis.* (i) - (iv) klar, (v) folgt aus dem Lemma in 6.1.4 im Fall  $m = 2$ .  $\square$

Durch die Einführung von  $\text{Eig}(F, \lambda)$  und die Bemerkung  $\text{Eig}(F, \lambda) = \ker(\lambda \text{id} - F)$  schaffen wir mehr Ordnung in unserer Diskussion der Eigenwerte und Eigenvektoren. Jedem Eigenwert  $\lambda$  wird nun ein Vektorraum zugeordnet  $\text{Eig}(F, \lambda)$ , in dem alle Eigenvektoren zu  $\lambda$  enthalten sind und der als Kern der Abbildung  $\lambda \text{id} - F$  beschrieben werden kann. Wegen (v) gibt es paarweise keine “Überlappungen” der Räume  $\text{Eig}(F, \lambda)$  (bis auf den Nullvektor, der in jedem Vektorraum enthalten ist).

## 6.2 Das charakteristische Polynom

Betrachten wir die Version der Eigenwertaufgabe, in der man alle Eigenwerte bestimmen möchte. Es stellt sich heraus, dass die Eigenwerte die Nullstellen eines besonderen Polynoms sind, das man einer Matrix bzw. einer linearen Abbildung zuordnet. Dieses Polynom wird das charakteristische Polynom genannt. Wenn man das charakteristische Polynome ausrechnen kann, so muss man dann “nur noch” die Nullstellen davon berechnen können.

Am Anfang vom Teil 1 des Kurses haben wir uns kurz mit Polynomen beschäftigt. Hier eine kurze Zusammenfassung vom Wissen über Polynome, das wir in diesem Abschnitt benötigen:

## Wiederholung

Sei  $t$  eine Unbestimmte.

- Der Polynomring  $\mathbb{K}[t]$  enthält die Monome  $t^0, t^1, t^2, t^3, \dots$ , wobei  $t^0$  mit 1 aus  $\mathbb{K}$  identifiziert wird, und endliche Linearkombinationen dieser Monome mit Koeffizienten aus  $\mathbb{K}$ . Das Polynom ist ein formaler Ausdruck (wird also durch die Angabe der Koeffizienten für seine Monome festgelegt).
- Addition und Multiplikation von Polynomen erfolgt komponentenweise.
- Die Gleichheit zweier Polynome aus  $\mathbb{K}[t]$  wird durch den Koeffizientenvergleich definiert.
- Ein Polynom  $f \in \mathbb{K}[t]$  kann an einem  $\lambda \in \mathbb{K}$  ausgewertet werden. Somit definiert jedes  $f \in \mathbb{K}[t]$  die Polynomfunktion  $\lambda \mapsto f(\lambda)$  auf  $\mathbb{K}$ .
- Polynom ist nicht das gleiche wie Polynomfunktion. Nehmen wir als Beispiel den Körper  $\mathbb{K} = \{0, 1\}$ <sup>331</sup>. Die Polynome  $f = t$  und  $g = t^2$  in  $\mathbb{K}[t]$  bestimmen die selbe Polynomfunktion (die identische Funktion), das sind aber trotzdem verschiedene Polynome.

### 6.2.1 Das charakteristische Polynom einer Matrix

In Kapitel 5 über die Determinanten haben aus den Formeln für die Determinanten Formeln für Grundaufgaben der linearen Algebra hergeleitet. Nun sind die Eigenwerte dran, wir erstellen mit der Verwendung der Determinanten eine Gleichung für die Eigenwerte.

**Prop.** *Sei  $A \in \mathbb{K}^{n \times n}$ ,  $\lambda \in \mathbb{K}$ . Dann sind folgende Aussagen äquivalent:*

- (i)  *$\lambda$  ist Eigenwert von  $A$ .*
- (ii)  *$\det(\lambda I - A) = 0$ .*

*Beweis.*  $\lambda$  ist genau dann Eigenwert von  $v$ , wenn die Gleichung  $Av = \lambda v$  eine Nichtrnull-Lösung  $v$  besitzt. Das bedeutet, dass der Kern der quadratischen Matrix  $\lambda I - A$  ungleich Null ist. Das Letztere ist äquivalent zu  $\det(\lambda I - A) = 0$ , wie wir aus dem Kapitel 5 über die Determinanten wissen.  $\square$

Hier und im Folgenden sei  $t$  eine Unbestimmte. Motiviert durch die vorige Proposition möchten wir das charakteristische Polynom  $p_A \in \mathbb{K}[t]$  einer Matrix  $A \in \mathbb{K}^{n \times n}$  durch die Gleichung

$$p_A := \det(tI - A) \quad (6.2.1)$$

einführen. Da wir mathematisch sauber arbeiten wollen, können wir das  $p_A$  auf diese Weise noch nicht bedenkenlos definieren. Schauen wir uns den Ausdruck  $\det(tI - A)$  genauer an.  $tI - A$  ist die Matrix, deren Komponenten zum Ring  $\mathbb{K}[t]$  der Polynome gehören. Etwa im Fall  $n = 2$ :

$$\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} = \begin{pmatrix} t - a_{11} & -a_{12} \\ -a_{21} & t - a_{22} \end{pmatrix}$$

Wir betrachten also die Determinante einer Matrix, deren Komponenten zum kommutativen Ring  $\mathbb{K}[t]$  gehören. Die Theorie der Determinanten haben wir aber für die Matrizen entwickelt, deren Elemente zu einem Körper

gehören. Streng genommen können wir also von dieser Theorie ohne extra Begründung keinen Gebrauch machen. Es ist klar, dass man in die Leibniz-Formel auch Matrizen über einem Ring einsetzen kann, denn in der Leibniz-Formel wird nirgendwo dividiert (es reicht als aus, einen kommutativen Ring zu haben, in dem im Gegensatz zu einem Körper nicht jedes Element invertierbar sein muss). Also könnte man im Prinzip  $p_A$  durch die Leibniz-Formel einführen. Aber, das wäre noch kein Ausweg. Denn wir wollen die Theorie der Determinanten benutzen, deren Gültigkeit wir nur im Fall von Matrizen über einem Körper verifiziert. Wir können die Situation ziemlich einfach retten: und zwar werden wir den Ring  $\mathbb{K}[t]$  in einen Körper einbetten.

### 6.2.2 Rechtfertigung der Definition vom charakteristischen Polynom

Es ist lehrreich einen Vergleich mit den ganzen Zahlen zu machen. Von einer Matrix mit ganzzahligen Komponenten  $A \in \mathbb{Z}^{n \times n}$  können wir die Determinante berechnen, da der Ring  $\mathbb{Z}$  der ganzen Zahlen zum Körper  $\mathbb{Q}$  rationaler Zahlen erweitert werden kann. Also gehört  $A \in \mathbb{Z}^{n \times n}$  zu  $\mathbb{Q}^{n \times n}$  und somit hat  $A$  die Determinante bzgl. des Körpers der rationalen Zahlen. Diese Determinante kann unter anderem durch die Leibniz-Formel berechnet werden, in der nicht dividiert wird. Das zeigt also, dass die Determinante einer ganzzahligen Matrix eine ganze Zahl ist. Obwohl wir die rationalen Zahlen als Hilfstruktur benutzt haben, ist unsere Eingabe  $A \in \mathbb{Z}^{n \times n}$  sowie Rückgabe  $\det(A)$  ganzzahlig.

In der Theorie der Ringe und Körper gibt es eine Verallgemeinerung der Konstruktion der rationalen Zahlen aus den ganzen Zahlen. Auf Basis welcher kommutative Ringe  $R$  kann man Quotienten  $\frac{a}{b}$  mit  $a, b \in R$  und

$b \neq 0$  einführen? Wenn der Ring  $R$  Nullteiler hat, hat man ein Problem. Nehmen wir etwa an,  $a$  und  $b$  sind zwei Elemente von einem kommutativen Ring  $R$  mit eins, die nicht gleich null sind und deren Produkt  $ab$  gleich null ist (in  $\mathbb{Z}/6\mathbb{Z}$  sind es zum Beispiel die Restklassen von 2 und 3). Die Quotienten können wir in einem solchen Ring nicht sinnvoll einführen. Im Prinzip könnten wir  $\frac{1}{a}$  und  $\frac{1}{b}$  betrachten, weil die Nenner ungleich null sind, wenn wir aber die beiden Quotienten multiplizieren kriegen wir ein Problem:  $\frac{1}{a} \cdot \frac{1}{b} = \frac{1}{ab} = \frac{1}{0}$ . Im Ring der ganzen Zahlen hat man keine Nullteiler, daher lassen sich die Quotienten (also die rationalen Zahlen) korrekt einführen.

Genau so hat man auch im Polynomring  $\mathbb{K}[t]$  keine Nullteiler, und somit hat man auch eine wohldefinierte Weise die Quotienten  $\frac{f}{g}$  von Polynomen  $f, g \in \mathbb{K}[t]$  mit  $g \neq 0$  einzuführen. Die Menge solcher Quotienten, ausgestattet mit den natürlichen Operationen  $+$  und  $\cdot$ , bildet den sogenannten Quotientenkörper  $\mathbb{K}(t)$ . Es ist keine Überraschung, dass die Addition und

Multiplikation folgendermaßen eingeführt werden:

$$\frac{f_1}{g_1} + \frac{f_2}{g_2} := \frac{f_1g_2 + f_2g_1}{g_1g_2} \quad (6.2.2)$$

$$\frac{f_1}{g_1} \cdot \frac{f_2}{g_2} := \frac{f_1f_2}{g_1g_2} \quad (6.2.3)$$

Des Weiteren lässt sich ein und der selbe Quotient auf mehrere Weisen hinschreiben. Wir müssen also klären, wann zwei Quotienten gleich sind. Hier auch keine Überraschungen:

$$\frac{f_1}{g_1} = \frac{f_2}{g_2} \Leftrightarrow f_1g_2 = f_2g_1. \quad (6.2.4)$$

Die Konstruktion des Quotientenkörper kann natürlich auch noch formaler beschrieben werden: ein Quotient ist eine Äquivalenzklasse aus Paaren (Zähler,Nenner) (mit Zähler aus  $\mathbb{K}[t]$  und Nenner aus  $\mathbb{K}[t] \setminus \{0\}$ ), wobei zwei solche Paare  $(f_1, g_1)$  und  $(f_2, g_2)$  äquivalent sind, wenn  $f_1g_2 = f_2g_1$  gilt. Sie können sich gerne überlegen, dass die oben angeführten Konstruktionen mathematisch korrekt sind: die eingeführte Relation ist tatsächlich

eine Äquivalenzrelation,  $+$  und  $\cdot$  sind unabhängig davon, durch wie man den Quotienten konkret dargestellt und  $\mathbb{K}(t)$  ist tatsächlich ein Körper.

Der Ring  $\mathbb{K}[t]$  wird als Unterstruktur von  $\mathbb{K}(t)$  aufgefasst, denn  $f \in \mathbb{K}[t]$  wird natürlicherweise mit dem Quotienten  $\frac{f}{1}$  identifiziert.

Nun hat unser Definition des charakteristischen Polynoms  $\mathbb{K}[t]$  eine eindeutige Interpretation. Man hat

$$p_A := \det(tI - A).$$

Die Elemente der Matrix  $tI - A$  gehören zum Ring  $\mathbb{K}[t]$  und somit auch zum Körper  $\mathbb{K}(t)$  der rationalen Funktionen. Somit wissen wir dass  $\det(tI - A)$  bzgl. des Körpers  $\mathbb{K}(t)$  definiert. Aber aus der Leibniz-Formel (in der nicht dividiert wird) wissen wir, dass  $\det(tI - A)$  zum Ring  $\mathbb{K}[t]$  gehört.

**Bem.**

08.05.2015

Zur Berechnung von charakteristischen Polynomen können jede Methode

aus dem Kapitel 5 benutzen: Leibniz-Formel, Laplace-Entwicklung, Gauß-Verfahren bzgl. des Körpers  $\mathbb{K}(t)$ .

**Bsp.** Hier einige Beispiele der charakteristischen Polynome:

Matrix	Das charakteristische Polynom
$\begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$	$\det \begin{pmatrix} t & 0 \\ 0 & t \end{pmatrix} = t^2$
$\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$	$\det \begin{pmatrix} t & 1 \\ 0 & t \end{pmatrix} = t^2$
$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$	$\det \begin{pmatrix} t-1 & 0 \\ 0 & t \end{pmatrix} = (t-1)t = t^2 - t$
$\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$	$\det \begin{pmatrix} t-1 & -1 \\ 0 & t-1 \end{pmatrix} = (t-1)^2$
$\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$	$\det \begin{pmatrix} t-1 & -1 \\ -1 & t-1 \end{pmatrix} = (t-1)^2 - 1 = t(t-2)$

**Bem.** Ist  $R$  ein kommutativer Ring mit 1 und ohne Nullteiler, so bildet die Menge der Quotienten  $\frac{a}{b}$  mit  $a \in R$  und  $b \in \mathbb{R} \setminus \{0\}$ , zusammen mit  $+$  und  $\cdot$  (die genauso wie oben eingeführt werden), den sogenannten Quotientenkörper  $\text{Quot}(R)$  des Rings  $R$ . Anhand dieses Körpers kann man die Determinanten der Matrizen aus  $R^{n \times n}$  einführen.

### 6.2.3 Nullstellen, Grad und Koeffizienten des charakteristischen Polynoms

**Prop.** Sei  $A \in \mathbb{K}^{n \times n}$  mit  $n \in \mathbb{N}$  und  $\lambda \in \mathbb{K}$ . Dann sind die folgenden Aussagen äquivalent:

- (i)  $\lambda$  ist Eigenwert von  $A$ .
- (ii)  $\lambda$  ist Nullstelle von  $p_A$ , d.h.  $p_A(\lambda) = 0$ .

*Beweis.* Die Aussage ist nichts anders als eine Umformulierung von Proposition aus 6.2.1 ( $\lambda$  Eigenwert von  $A \Leftrightarrow \det(\lambda I - A) = 0$ ).  $\square$

**Bsp.** Im charakteristischen Polynom sind alle Eigenwerte als seine Nullstellen gespeichert. Das Polynom enthält aber noch mehr Information über die Matrix.

Betrachten wir den binären Körper  $\mathbb{K} = \{0, 1\}$ . Hier zwei Matrizen:

$$A = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \in \mathbb{K}^{3 \times 3}, \quad O = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \in \mathbb{K}^{3 \times 3}.$$

Diese Matrizen sind schon sehr grundlegend unterschiedlich, denn keine Matrix ist ähnlich zur Nullmatrix (im buchstäblichen Sinne der Definition der Ähnlichkeit).

Berechnen wir die charakteristischen Polynome der beiden Matrizen (man

beachte, dass im binären Körper  $-1 = 1$  gilt):

$$p_A = \det \begin{pmatrix} t & 0 & 0 \\ 1 & t & 1 \\ 0 & 1 & t+1 \end{pmatrix} = t^2(t+1) + t = t^3 + t^2 + t = t(t^2 + t + 1)$$

$$p_O = \det \begin{pmatrix} t & 0 & 0 \\ 0 & t & 0 \\ 0 & 0 & t \end{pmatrix} = t^3$$

Die beiden Polynome sind nur einer Stelle Null: an der Null. Das heißt, die beiden Matrizen haben einen einzigen Eigenwert: den Null. Anhang der Eigenwerte sind also die Matrizen nicht unterscheidbar. Auch anhand der Auswertung des charakteristischen Polynoms an den Werten aus  $\mathbb{K}$  nicht:  $p_A(0) = p_O(0) = 0$  und  $p_A(1) = p_O(1) = 1$ . Aber  $p_A$  und  $p_O$  sind zwei verschiedene Polynome. Wir erkennen also den Unterschied von  $A$  und  $O$  an den charakteristischen Polynomen in diesem Fall.

Welche Informationen sind im charakteristischen Polynom als Koeffizienten enthalten? Zwei interessanten Koeffizienten sind die Koeffizienten der Monome  $t^0$  und  $t^{n-1}$ :

**Prop.** Sei  $A = (a_{ij})_{i,j=1,\dots,n} \in \mathbb{K}^{n \times n}$ . Dann hat das charakteristische Polynom die Form

$$p_A = t^n - (a_{11} + \cdots + a_{nn})t^{n-1} + \cdots + (-1)^n \det(A),$$

d.h., der Koeffizient vom Monom  $t^n$  ist 1, der Koeffizient vom Monom  $t^{n-1}$  ist  $-(a_{11} + \cdots + a_{nn})$ , und der Koeffizient vom Monom  $t^0$  ist  $(-1)^n \det(A)$ .

*Beweis.* Der Koeffizient vor  $t^0$  ist  $p_A(0) = \det(0 \cdot I - A) = \det(-A) = (-1)^n \det(A)$ .

Nach der Leibniz-Formel gilt:

$$p_A = \det(tI - A) = \sum_{\sigma \in S_n} \text{sign}(\sigma) \underbrace{\prod_{i=1}^n (t\delta_{i,\sigma(i)} - a_{i,\sigma(i)})}_{=: f_\sigma}$$

Beim Polynom  $f_\sigma$  trägt jedes  $i$  mit  $\sigma(i) = i$  ein Faktor  $t - a_{ii}$  bei. Die Faktoren zu  $i$  mit  $\sigma(i) \neq i$  sind Konstanten. Somit ist der Grad von  $f_\sigma$  gleich der Anzahl der Fixpunkte von  $\sigma$ , das sind die Indizes  $i$  mit  $\sigma(i) = i$ . Die identische Permutation hat  $n$  Fixpunkte, alle anderen weniger als  $n - 1$ . Daraus folgt, dass die Koeffizienten von  $p_A$  bei den Monomen  $t^n$  und  $t^{n-1}$  die gleichen sind wie beim Polynom  $f_\sigma = (t - a_{11}) \cdots (t - a_{nn})$  zur identischen Permutation  $\sigma$ . Das ergibt die gewünschten Ausdrücke für die Koeffizienten der Monome  $t^n$  und  $t^{n-1}$ .  $\square$

Den Wert  $\text{tr}(A) := a_{11} + \dots + a_{nn}$  nennt man die *Spur* der Matrix  $A = (a_{ij}) \in \mathbb{K}^{n \times n}$  (die Bezeichnung stammt vom englischen Wort „trace“). Wir habe also

$$p_A = t^n - \text{tr}(A)t^{n-1} + \cdots + (-1)^n \det(A).$$

Durch das charakteristische Polynom erkennt man also die Spur sowie die Determinante von  $A$ .

## 6.2.4 Das charakteristische Polynom, die Determinante und die Spur von linearen Abbildungen

Ähnliche Matrizen haben das gleiche charakteristische Polynom:

**Lem.** Seien  $A, \tilde{A} \in \mathbb{K}^{n \times n}$  ähnliche Matrizen. Dann haben  $A$  und  $\tilde{A}$  das gleiche charakteristische Polynome. Des Weiteren sind die Determinante und die Spur von  $A$  und  $\tilde{A}$  gleich.

*Beweis.* Laut der Definition der Ähnlichkeit gilt  $\tilde{A} = B^{-1}AB$  für eine invertierbare Matrix  $B \in \mathbb{K}^{n \times n}$ . Es reicht, die Behauptung über das charakteristische Polynom zu zeigen, denn die Informationen über die Determinante und die Spur ist in den Koeffizienten des charakteristischen Polynoms enthalten. Die Gleichheit der charakteristischen Polynome wird direkt ve-

rifiziert:

$$\begin{aligned}
p_{\tilde{A}} &= \det(tI - \tilde{A}) \\
&= \det(tI - B^{-1}AB) \\
&= \det(tB^{-1}IB - B^{-1}AB) \\
&= \det(B^{-1}(tI - A)B) \\
&= \det(B^{-1}) \det(tI - A) \det(B) \\
&= \det(tI - A) = p_A
\end{aligned}$$
□

Sei  $V$  ein  $n$ -dimensionaler Vektorraum über  $\mathbb{K}$  mit  $n \in \mathbb{N}$ . Sei  $F \in \text{Lin}(V)$ . Aus dem vorigen Lemma folgt, dass für zwei beliebige Basen  $\mathcal{A}$  und  $\mathcal{B}$  von  $V$  die Matrizen  $F_{\mathcal{A}}$  und  $F_{\mathcal{B}}$  das gleiche charakteristische Polynom, die gleiche Determinante und die gleich Spur haben (vgl. den Abschnitt über den Basiswechsel aus LA I:  $F_{\mathcal{A}} = T_{\mathcal{A} \leftarrow \mathcal{B}} F_{\mathcal{B}} T_{\mathcal{B} \leftarrow \mathcal{A}}$  mit  $T_{\mathcal{A} \leftarrow \mathcal{B}} T_{\mathcal{B} \leftarrow \mathcal{A}} = I$ , denn  $F_{\mathcal{A}}$  und  $F_{\mathcal{B}}$  sind ähnlich).

D.h. wir können  $\det(F)$ ,  $\text{tr}(F)$  und  $p_F$  (das charakteristische Polynom von

$F$ ) als

$$\begin{aligned}\det(F) &:= \det(F_{\mathcal{B}}) \\ \operatorname{tr}(F) &:= \operatorname{tr}(F_{\mathcal{B}}) \\ p_F &:= p_{F_{\mathcal{B}}}\end{aligned}\tag{6.2.5}$$

definieren, wobei die rechten Seiten von der Basis  $\mathcal{B}$  von  $V$  unabhängig sind.

### 6.2.5 Der Satz von Cayley-Hamilton

In ein Polynom  $f = \sum_{i=0}^d c_i t^i \in \mathbb{K}[t]$  kann man nicht nur Werte, sondern auch kompliziertere Objekte wie z.B. quadratische Matrizen einsetzen. Wir definieren die Auswertung  $f(A)$  von  $f$  an einer Matrix  $A \in \mathbb{K}^{n \times n}$  als

$$f(A) := \sum_{i=0}^d c_i A^i.$$

Für  $f = t^2 - t + 1$  erhält man z.B.  $f(A) = A^2 - A^1 + A^0 = A^2 - A + I$ .

**Bem.** Die Menge  $\mathbb{K}[A] = \{p(A) : p \in \mathbb{K}[t]\}$  ist ein kommutativer Ring mit 1, wenn wir  $\mathbb{K}[A]$  mit der Addition und Multiplikation von Matrizen ausstatten. Der Ring  $\mathbb{K}[A]$  ‘lebt’ innerhalb des Rings  $\mathbb{K}^{n \times n}$ . Im Gegenteil zu  $\mathbb{K}[A]$  ist  $\mathbb{K}^{n \times n}$  aber für alle  $n \geq 2$  nicht kommutativ. Dass  $\mathbb{K}[A]$  kommutativ ist, lässt sich direkt prüfen: das liegt daran, dass  $A^i A^j = A^j A^i$  für alle  $i, j \in \mathbb{N}_0$  erfüllt ist, denn die linke und die rechte Seite sind gleich  $A^{i+j}$ .

**Bsp.** Hier einige Beispiele von Formeln, die in  $\mathbb{K}[A]$  gelten:

$$(A + I)^2 = A^2 + 2A + I$$

$$(A + I)(A - I) = A^2 - I$$

$$(A + I)^3 = A^3 + 3A^2 + 3A + I$$

$$(A - I)(A^2 + A + I) = A^3 - I$$

Wenn man die Matrix  $A$  durch eine Zahl ersetzt und  $I$  durch die Zahl 1, so erhält man die bekannte Formeln aus der Schule, die man durch das Ausmultiplizieren herleiten kann. Auch die hier angegebenen Formeln können

durch Ausmultiplizieren hergeleitet werden. Die oben präsentierten 4 Formeln gelten für eine beliebige Matrix  $A$ . Es gibt natürlich auch Formeln, die für konkrete Matrizen spezifisch sind. Zum Beispiel gilt  $A^2 = I$  wenn  $x \mapsto Ax$  eine Spiegelung ist (an einem Punkt oder an einer Ebene oder noch allgemeiner) und es gilt  $A^2 = A$  wenn  $A$  eine Projektion ist (auf eine Ebene auf eine Gerade...)

**Bem.** Interessanterweise ist  $\mathbb{K}[A]$  nicht nur einfach ein kommutativer Ring mit 1, die Struktur von  $\mathbb{K}[A]$  ist sogar noch etwas reichhaltiger. Es ist nämlich klar, dass  $\mathbb{K}[A]$  auch ein Vektorraum ist:  $\mathbb{K}[A]$  ist Untervektorraum des Vektorraums  $\mathbb{K}^{n \times n}$ , was man direkt verifizieren kann. Einen Ring mit einer kompatiblen Vektorraum nennt man eine Algebra.  $\mathbb{K}[A]$  ist eine Algebra über dem Körper  $\mathbb{K}$ .

**Bem.** Wenn man einen neuen (endlich-dimensionalen) Vektorraum entdeckt bzw. einführt, dann ist eine der ersten quantitativen Fragen, die man stellen kann ist: was ist die Dimension? Wir können also auch im Fall von

$\mathbb{K}[A]$  fragen: was ist die Dimension von  $\mathbb{K}[A]$ ? Das wissen wir zwar noch nicht, aber wir können zumindest eine erste Abschätzung machen:  $\mathbb{K}[A]$  ist Untervektorraum von  $\mathbb{K}^{n \times n}$ , daher kann die Dimension von  $\mathbb{K}[A]$  nicht höher als  $\dim(\mathbb{K}^{n \times n}) = n^2$  sein. Wir werden in Kürze feststellen, dass der genau wert deutlich kleiner ist.

Wenn eine Formel oder eine Aussage einfach formulierbar oder verblüffend ist, so merkt man sie am leichtesten. Der Satz von Cayley-Hamilton ist sehr einfach formuliert *und* sehr verblüffend!

**Thm** (Der Satz von Cayley-Hamilton für Matrizen). *Sei  $A \in \mathbb{K}^{n \times n}$ . Dann gilt:*

$$p_A(A) = 0. \quad (6.2.6)$$

Nochmals mit Worten: wenn wir in das charakteristische Polynom einer Matrix die Matrix selbst einsetzen, kriegen wir die Null-Matrix.

Ein pragmatischer Mensch würde vielleicht fragen: *wieso* tun wir das denn? Es gibt verschiedene Anwendungen. Bevor wir aber sie diskutieren, wollen wir aber zuerst den Satz beweisen.

*Beweis.* Wir werden mit den Komponenten von  $A$  arbeiten. Sei also  $A = (a_{ij}) \in \mathbb{K}^{n \times n}$ . Des Weiteren sei  $B(t) = (b_{ij}(t)) \in \mathbb{K}[t]^{n \times n}$  die komplementäre Matrix von  $(tI - A)^\top$ . Dass wir  $tI - A$  ins Spiel bringen, ist logisch, denn  $p_A$  ist die Determinante von  $tI - A$ , und unsere Behauptung involviert  $p_A$ . Man beachte aber, dass wir  $tI - A$  transponiert haben.

Aus den Grundeigenschaften der komplementären Matrizen erhalten wir die Gleichungen

$$\sum_{j=1}^n b_{ij}(t)(t\delta_{jk} - a_{kj}) = \det(tI - A) \cdot \delta_{ik}$$

für alle  $i, k \in \{1, \dots, n\}$ . Die Gleichungen enthalten das charakteristische Polynom  $p_A = \det(tI - A)$  auf der rechten Seiten. Der Plan ist nun, in

die aufgestellten polynomiellen Gleichungen die Matrix  $A$  an der Stelle der Unbestimmten  $t$  einzusetzen. Durch das Einsetzen erhalten wir

$$\sum_{j=1}^n b_{ij}(A) \cdot (\delta_{jk}A - a_{kj}I) = p_A(A)\delta_{ik}$$

Multipliziere nun von rechts mit  $e_k$ :

$$\sum_{j=1}^n b_{ij}(A) \cdot \underbrace{(\delta_{jk}A - a_{kj}I)e_k}_{\delta_{jk}Ae_k - a_{kj}e_k} = p_A(A)\delta_{ik}e_k$$

und summiere über  $k \in \{1, \dots, n\}$ :

$$\sum_{k=1}^n \sum_{j=1}^n b_{ij}(A) \cdot (\delta_{jk}Ae_k - a_{kj}e_k) = p_A(A) \sum_{k=1}^n \delta_{ik}e_k \quad (6.2.7)$$

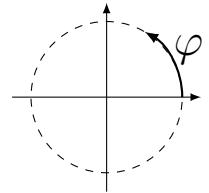
Uns bleibt es, die linke und die rechte Seite der letzten Gleichung zu vereinfachen.  $\delta_{ik}$  ist nur für  $i = k$  gleich 1 und ansonsten gleich 0. Daher ist die rechte Seite gleich  $p_A(A)e_i$ . Die linke Seite von (6.2.7) sieht etwas

umständlich aus kann aber deutlich vereinfacht werden. Wir verändern die Reihenfolge der beiden Summe und erhalten dann

$$\sum_{j=1}^n b_{ij}(A) \cdot \left( \underbrace{\sum_{k=1}^n \delta_{jk} A e_k}_{Ae_j} - \underbrace{\sum_{k=1}^n a_{kj} e_k}_{Ae_j} \right) = 0.$$

Also ist  $p_A(A)e_i = 0$  für alle  $i$ . Folglich ist  $p_A(A)$  die  $n \times n$  Null-Matrix.  $\square$

**Bsp.** Der Satz von Cayley-Hamilton ist sehr allgemein. Allgemeine Aussagen können manchmal überwältigend sein. In solchen Situationen kann es helfen, konkrete Beispiele zur allgemeinen Aussage zu betrachten. Betrachten wir die Matrix der Drehung um einen Winkel  $\varphi$  am Ursprung. Hier die Zusammenfassung, wie die Matrix aussieht und wie Sie auf Standardbasisvektoren wirkt:



$$e_1 \mapsto \begin{pmatrix} \cos \varphi \\ \sin \varphi \end{pmatrix}$$

$$e_2 \mapsto \begin{pmatrix} -\sin \varphi \\ \cos \varphi \end{pmatrix}$$

$$A = \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix}$$

$A^2$  dreht einen Vektor um  $2\varphi$ . Demnach ist  $A^2 = \begin{pmatrix} \cos 2\varphi & -\sin 2\varphi \\ \sin 2\varphi & \cos 2\varphi \end{pmatrix}$ . Au-

ßerdem gilt:

$$\begin{aligned} p_A &= \det \begin{pmatrix} t - \cos \varphi & \sin \varphi \\ -\sin \varphi & t - \cos \varphi \end{pmatrix} \\ &= t^2 - 2t \cos \varphi + \cos^2 \varphi + \sin^2 \varphi \\ &= t^2 - 2t \cos \varphi + 1 \end{aligned}$$

Nach Cayley-Hamilton ist  $A^2 - 2 \cos \varphi \cdot A + I = 0$ , d.h.

$$\begin{pmatrix} \cos 2\varphi & -\sin 2\varphi \\ \sin 2\varphi & \cos 2\varphi \end{pmatrix} = 2 \cos \varphi \begin{pmatrix} \cos \varphi & -\sin \varphi \\ \sin \varphi & \cos \varphi \end{pmatrix} - \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

Demzufolge ist

$$\cos(2\varphi) = 2 \cos^2 \varphi - 1$$

und

$$\sin(2\varphi) = 2 \sin \varphi \cos \varphi.$$

Das sind die bekannten Formeln für den Kosinus und Sinus des doppelten Winkels. Es kann sein, dass Sie diese Formeln bereits in der Schule hatten (in Mathematik oder Physik). Diese Formeln haben wir aus dem Satz von Cayley-Hamilton hergeleitet. Es gibt natürlich auch andere, direktere Möglichkeiten, die beiden Formeln herzuleiten. Uns geht in diesem Beispiel darum, das neue Wissen, das wir gerade erwerben mit dem alten Wissen zu verbinden (je mehr Verbindungen sehen wir, desto besser verstehen wir den Stoff).

Können wir auch Formeln für den Kosinus und Sinus des dreifachen Winkel herleiten? Der dreifache Winkel hängt mit  $A^3$  zusammen. Wir müssen also herausfinden, wie  $A^3$  von  $A$  abhängt. Wir wissen bereits, dass  $A^2 = 2 \cos \varphi A - I$  gilt. Daher ist  $A^3 = A \cdot A^2 = A(2 \cos \varphi A - I) = 2 \cos \varphi A^2 - A$ . Und dann können wir für  $A^2$  den Ausdruck  $2 \cos \varphi A - I$  noch einmal einsetzen. Das ergibt:

$$A^3 = 2 \cos \varphi (2 \cos \varphi A - I) = (4 \cos^2 \varphi - 1)A - 2 \cos \varphi I.$$

Wenn man nun die letzte Formel Komponentenweise ausschreibt sieht man, wie man  $\cos 3\varphi$  und  $\sin 3\varphi$  mit Hilfe von  $\cos \varphi$  und  $\sin \varphi$  darstellen kann.

**Bem.** Durch Cayley-Hamilton wird  $A^n$  ( $A \in \mathbb{K}^{n \times n}$ ) als Linearkombination von  $A^0, \dots, A^{n-1}$  dargestellt. Man kann hierdurch die Potenzen  $A^k$  mit Hilfe der Multiplikation von Polynomen ausrechnen (vgl. Beispiel oben). Cayley-Hamilton hilft uns also im Ring  $\mathbb{K}[A]$  zu rechnen.

**Bem.** Sei  $A \in \mathbb{K}^{n \times n}$ . Als Vektorraum über  $\mathbb{K}$  hat  $\mathbb{K}[A]$  die Dimension höchstens  $n$ . (Übungsaufgabe)

**Bem.** Wir können verschiedene interessante algebraische Strukturen als  $\mathbb{K}[A]$  umsetzen. Die Algebra  $\mathbb{K}[A]$  ist also oft eine konkrete Matrix-basierte ‘Umsetzung’ einer abstrakten algebraischen Struktur. Natürlich hat man auch die Verbindung in die andere Richtung: denn jede quadratische Matrix  $A$  wird mit der Algebra  $\mathbb{K}[A]$  in Verbindung gesetzt. Somit wird  $A$  ein Element der Algebraischen Struktur  $\mathbb{K}[A]$ .

Quadratische Matrizen sind konkrete Analoga der linearen Abbildungen eines Vektorraums  $V$ . Wir können natürlich auch eine lineare Abbildung  $F : V \rightarrow V$  in ein Polynom  $p = \sum_{i=1}^d c_i t^i \in \mathbb{K}[t]$  einsetzen: wir definieren  $p(F)$  als  $p(F) = \sum_{i=0}^d c_i F^i$ . Auch in diesem Fall wird der Ring  $\mathbb{K}[F] := \{q(F) : q \in \mathbb{K}[t]\}$  eingeführt. Der Ring  $\mathbb{K}[F]$  ist auch ein Untervektorraum von  $\text{Lin}(V)$ .

In der Sprache der linearen Abbildungen wird der Satz von Cayley-Hamilton folgendermaßen formuliert:

**Thm** (Satz von Cayley-Hamilton für lineare Abbildungen). *Sei  $V$  endlich-dimensionaler Vektorraum und  $F : V \rightarrow V$  lineare Abbildung. Dann gilt:*

$$p_F(F) = 0 \tag{6.2.8}$$

*Beweis.* Übungsaufgabe. □

## 6.3 Diagonalisierbarkeit

### 6.3.1 Eine notwendige und eine hinreichende Bedingung

Sei  $f \in \mathbb{K}[t] \setminus \{0\}$ .  $f$  zerfällt in Linearfaktoren, wenn  $f$  als  $f = c(t - \mu_1) \cdots (t - \mu_n)$  mit  $c \in \mathbb{K} \setminus \{0\}$  und  $\mu_1, \dots, \mu_n \in \mathbb{K}$  darstellbar ist. Diese Eigenschaft ist von der Wahl von  $\mathbb{K}$  abhängig.

Bsp.

- (i)  $f = 3t^2 - 6t + 3 \in \mathbb{K}[t]$  zerfällt für  $\mathbb{K} = \mathbb{Q}$  in Linearfaktoren, denn es ist

$$f = 3(t - 1)^2 \quad (c = 3 \text{ und } \mu_1 = \mu_2 = 1).$$

- (ii)  $f = t^2 - 2 \in \mathbb{K}[t]$  zerfällt für  $\mathbb{K} = \mathbb{Q}$  nicht in Linearfaktoren, aber bzgl.  $\mathbb{K} = \mathbb{R}$ :

$$f = (t - \sqrt{2})(t + \sqrt{2}).$$

**Thm.** Sei  $V$  Vektorraum über  $\mathbb{K}$  mit  $\dim(V) = n \in \mathbb{N}$ . Sei  $F \in \text{Lin}(V)$ . Dann gilt:

- (i) Ist  $F$  diagonalisierbar, so zerfällt  $p_F$  in Linearfaktoren.
- (ii) Ist  $p_F = (t - \mu_1) \cdots (t - \mu_n)$  mit paarweise verschiedenen  $\mu_1, \dots, \mu_k \in \mathbb{K}$ , d.h.  $p_F$  zerfällt in Linearfaktoren und hat  $n$  paarweise verschiedene Nullstellen, dann ist  $F$  diagonalisierbar.

*Beweis.*

- (i) Sei  $F$  diagonalisierbar, d.h. es existiert eine Basis  $\mathcal{B}$  von  $V$ , für welche die Matrix  $F_{\mathcal{B}}$  diagonal ist.

D.h.  $F_{\mathcal{B}} = \text{diag}(\mu_1, \dots, \mu_n)$  mit  $\mu_1, \dots, \mu_n \in \mathbb{K}$  und es gilt  $p_F = p_{F_{\mathcal{B}}} = \det(tI - F_{\mathcal{B}}) = (t - \mu_1) \cdots (t - \mu_n)$ . Also zerfällt  $p_F$  in Linearfaktoren.

- (ii)  $\mu_1, \dots, \mu_n$  sind paarweise verschiedene Eigenwerte von  $F$ . Aus Theorem 6.1.4 ergibt sich die Behauptung.  $\square$

### 6.3.2 Vielfachheit von Nullstellen und Zerlegbarkeit in Linearfaktoren

**Prop.** Sei  $f \in \mathbb{K}[t] \setminus \{0\}$  und sei  $\lambda \in \mathbb{K}$  Nullstelle von  $f$ . Dann existiert eine durch  $f$  und  $\lambda$  eindeutig bestimmte Zahl  $r \in \mathbb{N}$  und ein Polynom  $g \in \mathbb{K}[t] \setminus \{0\}$  mit  $f = (t - \lambda)^r g$  und  $g(\lambda) \neq 0$ .

*Beweis.* Aufgabe. □

Die Zahl  $r$  aus der vorigen Proposition nennt man die *algebraische Vielfachheit* der Nullstelle  $\lambda$  von  $f$ .

**Bem.**  $r$  und  $g$  zu  $f$  lassen sich durch die iterative Verwendung der Polynomdivision bestimmen (vgl. Schule).

**Prop.** Sei  $f \in \mathbb{K}[t] \setminus \{0\}$ . Seien  $\lambda_1, \dots, \lambda_k$  ( $k \in \mathbb{N}_0$ ) die paarweise verschiedenen Nullstellen von  $f$ . Dann kann  $f$  als das Produkt

$$f = (t - \lambda_1)^{r_1} \cdots (t - \lambda_k)^{r_k} g \quad (6.3.1)$$

dargestellt werden, wobei  $r_1, \dots, r_k \in \mathbb{N}$  mit  $g \in \mathbb{K}[t] \setminus \{0\}$  keine Nullstellen in  $\mathbb{K}$  hat. Die vorige Darstellung ist bis auf die Nummerierung der Terme  $(t - \lambda_i)^{r_i}$  ( $i \in \{1, \dots, k\}$ ) eindeutig.

*Beweis.* Aufgabe. □

**Bem.** Die Darstellung (6.3.1) lässt sich mit Hilfe der Polynomdivision bestimmen.

**Bem.**  $f$  zerfällt genau dann in Linearfaktoren, wenn in der Darstellung (6.3.1)  $g$  eine Konstante ist, d.h.  $g \in \mathbb{K} \setminus \{0\}$ .

**Bem.** Zur Berechnung von (6.3.1) muss die Menge  $\{\lambda_1, \dots, \lambda_k\}$  von  $f$  berechnet werden. Möglichkeiten dafür sind:

- (i)  $\mathbb{K}$  ist endlich  $\rightsquigarrow$  alle Elemente von  $\mathbb{K}$  durchprobieren (die direkte Methode).

(ii)  $\mathbb{K} = \mathbb{Q}$ . Man hat die folgende notwendige Bedingung. Sei  $f \in \mathbb{Z}[t] \setminus \{0\}$  mit  $f = c_0 t^0 + \dots + c_d t^d$  ( $c_0, \dots, c_d \in \mathbb{Z}, c_d \neq 0, d \in \mathbb{N}$ ). Sei  $\lambda = \frac{a}{b}$  mit teilerfremden  $a, b \in \mathbb{Z} \setminus \{0\}$  Nullstelle von  $f$ , d.h.  $f(\lambda) = 0$ . Dann gilt:

- $a$  teilt  $c_0$
- $b$  teilt  $c_d$

Dies sind wiederum endlich viele Möglichkeiten.

Bsp.

$$(i) A = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \in \mathbb{K}^{3 \times 3} \text{ mit } \mathbb{K} = \mathbb{Z}/2\mathbb{Z}. \text{ Dann ist}$$

$$p_A = \det \begin{pmatrix} t & 0 & 0 \\ 1 & t & 1 \\ 0 & 1 & t+1 \end{pmatrix} = t^2(t+1) + t = t^3 + t^2 + t = t \underbrace{(t^2 + t + 1)}_{\text{zerfällt nicht in Linearfaktoren}}.$$

Es folgt:  $A$  ist nicht diagonalisierbar (bzgl.  $\mathbb{K} = \mathbb{Z}/2\mathbb{Z}$ ).

(ii)  $A = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \in \mathbb{K}^{3 \times 3}$  mit  $\mathbb{K} = \mathbb{Q}$ . Dann ist

$$p_A = \det \begin{pmatrix} t & 0 & 0 \\ -1 & t & -1 \\ 0 & -1 & t-1 \end{pmatrix} = t^2(t-1)-t = t^3-t^2-t = t \underbrace{(t^2-t-1)}_{\text{hat keine Nullstellen in } \mathbb{Q}}.$$

Es folgt:  $A$  ist bzgl.  $\mathbb{K} = \mathbb{Q}$  nicht diagonalisierbar.

(iii) Sei  $A$  wie oben und  $\mathbb{K} = \mathbb{R}$ . Man hat drei unterschiedliche Nullstellen:  $0, \frac{1}{2} \pm \frac{\sqrt{5}}{2}$ . D.h.  $p_A = t(t - \frac{1}{2} - \frac{\sqrt{5}}{2})(t - \frac{1}{2} + \frac{\sqrt{5}}{2})$ . Es folgt:  $A$  ist diagonalisierbar.

(iv) Sei  $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in \mathbb{K}^{2 \times 2}$ ,  $\mathbb{K} = \mathbb{R}$ .

Es ist  $p_A = t^2 + 1$ .  $\lambda^2 + 1 \neq 0 \quad \forall \lambda \in \mathbb{R} \Rightarrow$  bzgl.  $\mathbb{K} = \mathbb{R}$  zerfällt  $A$  nicht in Linearfaktoren.

(v) Sei  $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in \mathbb{K}^{2 \times 2}$ ,  $\mathbb{K} = \mathbb{C}$ . Sei  $i$  die imaginäre Einheit, d.h.  $i^2 + 1 = 0$ .

Es ist  $p_A = t^2 + 1 = (t - i)(t + i)$ .  $\Rightarrow A$  ist diagonalisierbar.

(vi) Sei  $A = O \in \mathbb{K}^{2 \times 2}$  und  $\mathbb{K}$  beliebig. Dann ist  $p_A = t^2$ .  $A$  ist diagonalisierbar (das Theorem aus 6.3.1 kann das aber nicht entscheiden).

(vii) Sei  $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \in \mathbb{K}^{2 \times 2}$  und  $\mathbb{K}$  beliebig. Dann ist  $p_A = t^2$ .  $A$  ist nicht diagonalisierbar (vgl. Kapitel 6...).

### 6.3.3 Ungleichungen für die algebraische und die geometrische Vielfachheit

Sei  $V$  ein endlichdimensionaler Vektorraum über  $\mathbb{K}$  und sei  $F \in \text{Lin}(V)$ . Für die Diagonalsierbarkeit von  $F$  ist es notwendig, aber im allgemeinen Fall nicht hinreichend, dass  $p_F$  in Linearfaktoren zerfällt. Wir sind auf der Suche nach einer Bedingung an  $F$ , die in Kombination mit der vorigen Bedingung an das charakteristische Polynom, die Diagonalsierbarkeit von  $F$  charakterisiert.

Sei  $\lambda$  Eigenwert von  $F$ , d.h.  $p_F(\lambda) = 0$ . Die Vielfachheit von  $\lambda$  als Nullstelle von  $p_F$  nennt man die *algebraische Vielfachheit* des Eigenwerts  $\lambda$  von  $F$ .

Die Dimension von  $\text{Eig}(F, \lambda) = \ker(F - \lambda \text{id})$  nennt man die *geometrische Vielfachheit* des Eigenwerts  $\lambda$  von  $F$ .

**Thm.** Sei  $V$   $n$ -dimensionaler Vektorraum über  $\mathbb{K}$  mit  $n \in \mathbb{N}$  und sei  $F \in$

$\text{Lin}(V)$ . Sei  $\lambda$  Eigenwert von  $F$ . Sei  $k$  die geometrische Vielfachheit des Eigenwerts  $\lambda$  von  $F$  und  $l$  die algebraische Vielfachheit. Dann gilt:  $k \leq l$ .

*Beweis.* Sei  $b_1, \dots, b_k$  eine Basis von  $\text{Eig}(F, \lambda) = \{v \in V : F(v) = \lambda v\}$ . Wir erweitern  $b_1, \dots, b_k$  zu einer Basis  $\mathcal{B} = (b_1, \dots, b_n)$  von  $V$ . Die Matrix  $F_{\mathcal{B}}$  hat die folgende Struktur:

$$F_{\mathcal{B}} = \begin{pmatrix} \lambda I_k & C \\ O & D \end{pmatrix}$$

mit  $C \in \mathbb{K}^{k \times (n-k)}$  und  $D \in \mathbb{K}^{(n-k) \times (n-k)}$ . Es folgt:  $p_F = p_{F_{\mathcal{B}}} = \det(tI - F_{\mathcal{B}}) = \det((t - \lambda)I)_{k \times k} \det(tI - D)_{(n-k) \times (n-k)} = (t - \lambda)^k \det(tI - D)$ . Es folgt, dass die algebraische Vielfachheit von  $l$  mindestens  $k$  ist.  $\square$

### 6.3.4 Charakterisierung der Diagonalisierbarkeit

**Thm.** Sei  $V$  ein  $n$ -dimensionaler Vektorraum über  $\mathbb{K}$  mit  $n \in \mathbb{N}$ . Sei  $F \in \text{Lin}(V)$ . Dann sind die folgenden Bedingungen äquivalent:

- (i)  $F$  ist diagonalisierbar.
- (ii) Das charakteristische Polynom von  $F$  zerfällt in Linearfaktoren und für jeden Eigenwert  $\lambda$  von  $F$  sind die geometrische und die algebraische Vielfachheit von  $\lambda$  gleich.
- (iii)  $V$  ist direkte Summe der Eigenräume zu den Eigenwerten von  $F$ .

*Beweis.*

(i) $\Rightarrow$ (ii): Sei  $F$  diagonalisierbar. Dann existiert eine Basis  $\mathcal{B} = (b_1, \dots, b_n)$  und Werte  $\mu_1, \dots, \mu_n \in \mathbb{K}$  mit  $F(b_i) = \mu_i b_i$  für alle  $i \in \{1, \dots, n\}$ .  $\Rightarrow$

$$F_{\mathcal{B}} = \begin{pmatrix} \mu_1 & & \\ & \ddots & \\ & & \mu_n \end{pmatrix} \quad \Rightarrow \quad p_F = p_{F_{\mathcal{B}}} = \det(tI - F_{\mathcal{B}}) = (t - \mu_1) \cdots (t - \mu_n),$$

also zerfällt  $p_F$  in Linearfaktoren. Sei  $\lambda$  beliebiger Eigenwert von  $F$  und sei  $r \in \mathbb{K}$  die algebraische Vielfachheit von  $\lambda$ . Dann gilt  $\lambda =$

$\mu_i$  für genau  $r$  unterschiedliche Indizes  $i \in \{1, \dots, n\}$ . O.B.d.A. sei  $\lambda = \mu_1 = \dots = \mu_r$  und  $\lambda \neq \mu_i$  für  $i > r$ . Somit sind  $b_1, \dots, b_r$  linear unabhängige Eigenvektoren zu  $\lambda$ , sodass die Ungleichung  $r = \dim(\text{lin}(b_1, \dots, b_r)) \leq \dim(\text{Eig}(F, \lambda)) =: l$  erfüllt ist. Es folgt, dass die Gleichung  $r = l$  erfüllt ist, da die Ungleichung  $l \leq r$  in 6.3.3 gezeigt wurde.

(ii) $\Rightarrow$ (iii): Sei (ii) erfüllt. Seien  $\lambda_1, \dots, \lambda_k$  ( $k \in \mathbb{N}$ ) die paarweise verschiedenen Nullstellen von  $p_F$ . Wegen (ii) gilt  $p_F = (t - \lambda_1)^{r_1} \cdots (t - \lambda_k)^{r_k}$  mit  $r_1, \dots, r_k \in \mathbb{N}$ . Wegen (ii) gilt auch  $r_i = \dim(\text{Eig}(F, \lambda_i))$  für alle  $i \in \{1, \dots, n\}$ .

Wir zeigen, dass die Summe der Räume  $\text{Eig}(F, \lambda_i)$  mit  $i \in \{1, \dots, k\}$  direkt ist. Seien  $v_i \in \text{Eig}(F, \lambda_i)$   $\forall i \in \{1, \dots, k\}$  Vektoren mit  $v_1 + \dots + v_k = 0$ . Da Eigenvektoren zu paarweise verschiedenen Eigenwerten linear unabhängig sind (Lemma 6.1.4), folgt  $v_1 = \dots = v_k = 0$ . Also

ist die Summe der Räume  $\text{Eig}(F, \lambda_i)$  mit  $i \in \{1, \dots, k\}$  direkt. Für die Dimension dieser Summe gilt:

$$\begin{aligned}
 & \dim(\text{Eig}(F, \lambda_1) \oplus \dots \oplus \text{Eig}(F, \lambda_k)) \\
 = & \dim(\text{Eig}(F, \lambda_1)) + \dots + \dim(\text{Eig}(F, \lambda_k)) \quad \text{nach Theorem 3.6.4} \\
 = & r_1 + \dots + r_k \quad \text{wegen (ii)} \\
 = & n = \dim(V) \\
 \Rightarrow & \text{Eig}(F, \lambda_1) \oplus \dots \oplus \text{Eig}(F, \lambda_k) = V
 \end{aligned}$$

(iii) $\Rightarrow$ (i): Sei (iii) erfüllt und seien  $\lambda_1, \dots, \lambda_k$  die Eigenwerte wie oben. Sei  $\mathcal{B}_i$  eine Basis von  $\text{Eig}(F, \lambda_i)$  für  $i \in \{1, \dots, k\}$ . Wegen (iii) erhält man durch das Zusammenfügen der Systeme  $\mathcal{B}_1, \dots, \mathcal{B}_k$  eine Basis  $\mathcal{B}$  von  $V$ . Für diese Basis gilt

$$F_{\mathcal{B}} = \begin{pmatrix} \lambda_1 I_{r_1} & & \\ & \ddots & \\ & & \lambda_k I_{r_k} \end{pmatrix}$$

mit  $r_i = \dim(\text{Eig}(F, \lambda_i)) \in \mathbb{N} \quad \forall i \in \{1, \dots, k\}$ . Also ist  $F$  diagonalisierbar.  $\square$

**Bem.** Das vorige Theorem zusammen mit Methoden zur Bestimmung von Nullstellen und Faktorisierung von Polynomen führt zu Rechenverfahren, welche die Diagonalisierbarkeit entscheiden können und gegebenenfalls eine Basis  $\mathcal{B}$  finden, für die  $F_{\mathcal{B}}$  diagonal ist:

1. Bestimme alle Eigenwerte der gegebenen Matrix und die jeweiligen algebraischen Vielfachheiten.
2. Berechne zu jedem der Eigenwerte eine Basis des Eigenraums (und somit auch die geometrische Vielfachheit).
3. Ist für einen der Eigenwerte die geometrische Vielfachheit ungleich der algebraischen Vielfachheit, dann ist die gegebene Matrix nicht diagonalisierbar.

4. Sonst ist die Matrix Diagonal in der Basis, die durch das Zusammenfügen der Basen der Eigenräume entsteht.

## 6.4 Die Jordansche Normalform

### 6.4.1 Die Voraussetzungen

Das Hauptresultat dieses Abschnitts wird mit den folgenden Voraussetzungen formuliert.

Sei  $V$  ein  $n$ -dimensionaler Vektorraum über  $\mathbb{K}$  mit  $n \in \mathbb{N}$ . Sei  $F \in \text{Lin}(V)$  eine Abbildung, deren charakteristisches Polynom in Linearfaktoren zerfällt, d.h.

$$p_F = (t - \lambda_1)^{r_1} \cdots (t - \lambda_k)^{r_k}$$

mit paarweise verschiedenen  $\lambda_1, \dots, \lambda_k \in \mathbb{K}$  und  $r_1, \dots, r_k \in \mathbb{N}$ .

Es sei daran erinnert, dass die vorige Voraussetzung im Fall  $\mathbb{K} = \mathbb{C}$  für jede lineare Abbildung  $F \in \text{Lin}(V)$  erfüllt ist (im Gegenteil zum Fall  $\mathbb{K} = \mathbb{R}$ ).

### 6.4.2 Das Ziel und der Ansatz

Unter den Voraussetzungen aus 6.4.1 wollen wir eine Basis  $\mathcal{B}$  von  $V$  finden, für welche die Matrix  $F_{\mathcal{B}}$  möglichst wenig Nichtnullkomponenten außerhalb der Diagonalen hat (die genaue Struktur von  $F_{\mathcal{B}}$  wird später beschrieben).

Im allgemeinen Fall wird die gesuchte Basis  $\mathcal{B}$  nicht nur aus Vektoren der Eigenräume  $\text{Eig}(F, \lambda_i) = \ker(F - \lambda_i \text{id})$  bestehen, sondern auch aus Vektoren der Räume  $\ker(F - \lambda_i \text{id})^2, \ker(F - \lambda_i \text{id})^3$ , usw. Beachte hier: wenn eine einfache Anwendung einer Abbildung einen Vektor nicht auf 0 bringt, dann könnte es immer noch sein, dass eine zweifache oder eine dreifache Anwendung den Vektor auf 0 bringt. Mit anderen Worten wird der Kern der Potenz einer linearen Abbildung im Allgemeinen größer, wenn der Exponent wächst.

Wenn ein Vektor  $v \in \ker(F - \lambda_i \text{id}) \setminus \{0\}$  in  $\mathcal{B}$  aufgenommen wird, so hat

$F_{\mathcal{B}}$  bzgl. des Vektors  $v$  die Struktur

$$F_{\mathcal{B}} = v \begin{pmatrix} v \\ 0 \\ \vdots \\ 0 \\ \lambda_i \\ 0 \\ \vdots \\ 0 \end{pmatrix},$$

denn  $v$  ist Eigenvektor von  $F$  zum Eigenwert  $\lambda_i$ , sodass man  $F(v) = \lambda_i v$  hat.

Was passiert, wenn man nicht genug Eigenvektoren hat um eine Basis daraus zu bilden? Wenn etwa ein Vektor  $w \in \ker(F - \lambda_i \text{id})^2 \setminus \ker(F - \lambda_i \text{id})$  in  $\mathcal{B}$  aufgenommen wird, dann wird die Konstruktion zu sein, dass man neben  $w$  zusätzlich auch den Vektor  $v := (F - \lambda_i \text{id})(w)$  mit aufnimmt. So gilt

$v \neq 0$  und  $(F - \lambda_i \text{id})(v) = 0$ , d.h.  $F(v) = \lambda_i v$  für  $v$  und  $(F - \lambda_i \text{id})(w) = v$ , d.h.  $F(w) = \lambda_i w + v$ . Somit hat  $F_{\mathcal{B}}$  bzgl.  $v$  und  $w$  die folgende Struktur

$$F_{\mathcal{B}} = \begin{pmatrix} v & w \\ v & w \end{pmatrix} \left( \begin{array}{cc|cc} & & v & w \\ & & 0 & 0 \\ & & \vdots & \vdots \\ & & 0 & 0 \\ & & \lambda_i & 1 \\ & & 0 & \lambda_i \\ & & \vdots & 0 \\ & & \vdots & \vdots \\ & & 0 & 0 \end{array} \right).$$

In den Spalten für  $v$  und  $w$ : zweimal  $\lambda_i$  auf der Diagonalen und eine einzige Nichtnullkomponente (die 1) außerhalb der Diagonale. Zwar sehen wir in den beiden Spalten auch außerhalb der Diagonale Nichtnull-Elemente (in dieser Beispielsituation, nur ein Nichtnull-Element), aber nur wenige.

### 6.4.3 Über das Addieren eines Vielfachen der identischen Abbildung

Die Änderungen, die dadurch entstehen, dass man zur einer linearen Abbildung ein vielfaches der identischen Abbildung dazu addieren, können sehr einfach verfolgt werden. Bzgl. der Diagonalisierbarkeit ändert sich zum Beispiel gar nichts, das charakteristische Polynom ändert sich durch eine Verschiebung, Eigenwerte verschieben sich ebenfalls. In der folgenden Proposition werden die Auswirkungen der Änderung von  $F$  zu  $F + \alpha \text{id}$  genau dargestellt.

**Prop** (Verschiebungstrick). *Sei  $F : V \rightarrow V$  lineare Abbildung eines  $n$ -dimensionalen Vektorraums, mit  $n \in \mathbb{N}$ , und man betrachte die Abbildung  $G := F + \alpha \text{id}$  mit  $\alpha \in \mathbb{K}$ . Dann gilt:*

- (i) *Für jedes  $\lambda \in \mathbb{K}$  gilt:  $\lambda$  ist Eigenwert von  $F \Leftrightarrow \lambda + \alpha$  ist Eigenwert von  $G$ .*

(ii) Für jedes  $\lambda \in \mathbb{K}$  gilt:  $\text{Eig}(F, \lambda) = \text{Eig}(G, \lambda + \alpha)$ .

Insbesondere ist die geometrische Vielfachheit von jedem Eigenwert  $\lambda$  von  $F$  gleich der geometrischen Vielfachheit des entsprechenden Eigenwertes  $\lambda + \alpha$  von  $G$ .

(iii) Die algebraische Vielfachheit von jedem Eigenwert  $\lambda$  von  $F$  ist gleich der algebraischen Vielfachheit des entsprechenden Eigenwertes  $\lambda + \alpha$  von  $G$ .

(iv) Für die charakteristischen Polynome von  $F$  und  $G$  gilt:  $p_G(t) = p_F(t - \alpha)$ .

(v) Für jede Basis  $\mathcal{B}$  von  $V$  gilt:  $G_{\mathcal{B}} = F_{\mathcal{B}} + \alpha I$ .

Beweis.

(i,ii) Sei  $v \in V \setminus \{0\}$ . Dann gilt:

$$\begin{aligned} & \lambda \in \mathbb{K} \text{ ist Eigenwert von } F \text{ zu } v \\ \Leftrightarrow & F(v) = \lambda v \\ \Leftrightarrow & F(v) + \alpha v = (\lambda + \alpha)v \\ \Leftrightarrow & G(v) = (\lambda + \alpha)v \\ \Leftrightarrow & \lambda + \alpha \text{ ist Eigenwert von } G \text{ zu } v \end{aligned}$$

(v) Sei  $\mathcal{B}$  Basis von  $V$ . Dann gilt  $G_{\mathcal{B}} = (F + \alpha \text{id})_{\mathcal{B}} = F_{\mathcal{B}} + \alpha \text{id}_{\mathcal{B}} = F_{\mathcal{B}} + \alpha I$ .

(iv) Es gilt  $p_G(t) = p_{G_{\mathcal{B}}}(t) = \det(tI - G_{\mathcal{B}}) = \det(tI - F_{\mathcal{B}} - \alpha I) = \det((t - \alpha)I - F_{\mathcal{B}}) = p_{F_{\mathcal{B}}}(t - \alpha) = p_F(t - \alpha)$ .

(iii) folgt direkt aus (iv). □

Unter den Voraussetzungen aus 6.4.1 können wir nun mit Hilfe der vorigen Proposition Behauptungen über  $G = F - \lambda_i \text{id}$  (mit Eigenwert 0, d.h.  $G$  ist nicht invertierbar) zu Behauptungen über  $F$  konvertieren.

#### 6.4.4 Das Lemma von Fitting

Das Lemma von Fitting ist das Herzstück der Theorie der Jordanschen Normalformen (JNF). Der Kontext zu diesem Lemma ist so. Wir beschäftigen uns mit einem Eigenwert  $\lambda_i$  von  $F \in \text{Lin}(V)$  und betrachten dafür die Abbildung  $G := F - \lambda_i \text{id}$ . Der Kern von  $G$  enthält die Eigenvektoren von  $F$  zu  $\lambda_i$ , aber es kann sein, dass wir noch weitere Vektoren brauchen, die mit  $\lambda_i$  zusammenhängen, aber keine Eigenvektoren sind, um am Ende die JNF von  $F$  aufzubauen. In dem Lemma geht es darum, den zugrundeliegenden Vektorraum  $V$  in Summe von zwei Vektorräumen zu zerlegen. Der eine Vektorraum, der im Lemma als  $U_d$  bezeichnet wird, enthält die Eigenvektoren von  $F$  zu  $\lambda_i$  und die weiteren Vektoren, die mit  $\lambda_i$  zusammenhängen. Der andere Vektorraum, der im Lemma als  $W_d$  bezeichnet wird, ist der “Rest”. Was ist die konzeptuelle Beschreibung von  $U_d$ ? Der Raum  $U_d$  besteht aus den Vektoren, die auf 0 abgebildet werden, wenn man zu diesen Vektoren  $G$  oft genug anwendet (bei Eigenvektoren von  $F$  zum Eigenwert  $\lambda_i$  ist einmal

schon genug). Im Lemma taucht nur  $G$  auf, nicht das  $F$  (das  $F$  hat man, wenn wir das Lemma später anwenden). Nun zum anderen Vektorraum  $W_d$ . Wenn wir  $G$  zu  $V$  anwenden, erhalten wir das Bild von  $V$  bzgl.  $G$ :  $V$  schrumpft also zu einem potenziell kleineren Vektorraum  $\text{im}(G)$ . Wir bleiben aber nicht bei  $\text{im}(G)$  sondern wenden zu  $\text{im}(G)$  die Abbildung  $G$  nochmal an, und so weiter. So entsteht durch die iterative Anwendung von  $G$  eine Folge von Vektorräumen, die ineinander geschachtelt sind. Es ist also eine Art Matrjoschka (<https://de.wikipedia.org/wiki/Matrjoschka>) aus Vektorräumen. Unser Räume sind aber endlich-dimensional, beim Schrumpfen verringert sich die Dimension. Irgendwann erreichen wir also den Raum  $W_d$ , der nicht mehr weiter geschrumpft wird. Der Raum  $W_d$  ist die kleinste Puppe in unserer Matrjoschka.

**Lem.** *Sei  $V$  ein  $n$ -dimensionaler Vektorraum über  $\mathbb{K}$  mit  $n \in \mathbb{N}$ . Sei  $G \in \text{Lin}(V)$  eine nicht-invertierbare Abbildung (d.h. 0 ist Eigenwert von  $G$ ). Sei  $r \in \mathbb{N}$  die algebraische Vielfachheit des Eigenwerts 0 von  $G$ . Für  $i \in \mathbb{N}_0$*

seien

$$U_i := \ker(G^i) \quad \text{und} \quad W_i := \operatorname{im}(G^i).$$

Dann existiert ein Wert  $d \in \{1, \dots, r\}$  mit den folgenden Eigenschaften:

(a1)  $\{0\} = U_0 \subsetneq U_1 \subsetneq \dots \subsetneq U_d = U_{d+1} = \dots$

(a2)  $V = W_0 \supsetneq W_1 \supsetneq \dots \supsetneq W_d = W_{d+1} = \dots$

(b)  $V = U_d \oplus W_d$

(c1)  $G(U_d) \subseteq U_d$  und  $S := G|_{U_d} \in \operatorname{Lin}(U_d)$  erfüllt die Bedingung  $S^d = 0$ .

(c2)  $G(W_d) = W_d$  und  $T := G|_{W_d} \in \operatorname{Lin}(W_d)$  ist eine Bijektion.

(d) Für die charakteristischen Polynome von  $G$ ,  $S$  und  $T$  gilt:

$$p_G = p_S p_T$$

$$p_S = t^r$$

$$p_T(0) \neq 0$$

(e)  $\dim(U_d) = r$  und  $\dim(W_d) = n - r$ .

**Bem** (zu den Bezeichnungen).

$X \subsetneq Y$  bedeutet  $X \subseteq Y, X \neq Y$  und  $X \supsetneq Y$  bedeutet  $X \supseteq Y, X \neq Y$ .

$G|_{U_d} \in \text{Lin}(U_d)$  bedeutet, dass wir die Einschränkung von  $G$  auf  $U_d$  nicht (wie sonst üblich) als eine Abbildung von  $U_d$  nach  $V$  interpretieren wollen, sondern als eine Abbildung von  $U_d$  nach  $U_d$ .

*Beweis.*

(a1) Zunächst wird (a1) für ein  $d \in \mathbb{N}$  gezeigt. Am Ende des Beweises von diesem Lemma wird die Ungleichung  $d \leq r$  nachgewiesen.

Für jedes  $i \in \mathbb{N}_0$  gilt  $U_i \subseteq U_{i+1}$ , denn für jedes  $x \in V$  folgt aus  $G^i(x) = 0$  die Gleichung  $G^{i+1}(x) = G(G^i(x)) = G(0) = 0$ . Somit hat man die Gleichheit  $U_i = U_{i+1}$  genau dann, wenn  $\dim(U_i) = \dim(U_{i+1})$  gilt, und eine strikte Inklusion  $U_i \subsetneq U_{i+1}$ , wenn  $\dim(U_i) < \dim(U_{i+1})$ .

gilt. Weil  $V$  endlich-dimensional und  $(\dim(U_i))_{i \in \mathbb{N}}$  monoton ist, gilt  $\dim(U_i) = \dim(U_{i+1})$  für ein  $i \in \mathbb{N}_0$ .

Wir zeigen nun, dass aus der Gleichheit  $U_i = U_{i+1}$  die Gleichheit  $U_{i+1} = U_{i+2}$  folgt. Sei  $x \in U_{i+2}$ , d.h.  $G^{i+2}(x) = 0$ . Somit gilt  $G^{i+1}(G(x)) = 0$ , d.h.  $G(x) \in U_{i+1}$ . Damit ist  $G(x) \in U_i \Leftrightarrow G^i(G(x)) = 0 \Leftrightarrow G^{i+1}(x) = 0 \Leftrightarrow x \in U_{i+1}$ . Also gilt  $U_{i+2} \subseteq U_{i+1}$  und somit auch  $U_{i+1} = U_{i+2}$ . Wir haben (a1) für ein  $d \in \mathbb{N}$  nachgewiesen.

- (a2) Nach dem Rangsatz gilt  $\dim(U_i) + \dim(W_i) = \dim(V) = n$  für jedes  $i \in \mathbb{N}_0$ . Somit hat man  $\dim(U_i) = \dim(U_{i+1})$  genau dann, wenn  $\dim(W_i) = \dim(W_{i+1})$  gilt, und  $\dim(U_i) < \dim(U_{i+1})$  genau dann, wenn  $\dim(W_i) > \dim(W_{i+1})$  gilt. Somit folgt (a2) aus (a1).
- (b) Wir zeigen, dass die Summe von  $U_d$  und  $W_d$  direkt ist, d.h.  $U_d \cap W_d = \{0\}$ . Sei  $x \in U_d \cap W_d$ , d.h.  $G^d(x) = 0$  und  $x = G^d(v)$  für ein  $v \in V$ . Durch Einsetzen erhält man  $G^d(G^d(v)) = 0$ , d.h.  $G^{2d}(v) = 0$ . Also

$v \in U_{2d}$  und wegen  $U_{2d} = U_d$  erhält man  $G^d(v) = 0$ . D.h.  $x = 0$ .

Wir haben  $U_d \cap W_d = \{0\}$  gezeigt, also ist die Summe von  $U_d$  und  $W_d$  direkt. Nach dem Rangsatz gilt  $\dim(U_d) + \dim(W_d) = n$ . Es folgt  $\dim(U_d \oplus W_d) = \dim(U_d) + \dim(W_d) = n = \dim(V)$ , d.h.  $U_d \oplus W_d = V$ .

(c1) Aus der Definition der  $U_i$  folgt: 29.05.2015

$$\begin{aligned}
 G(U_i) &= \{G(x) : x \in V, G^i(x) = 0\} \\
 &= \{G(x) : x \in V, G^{i-1}(G(x)) = 0\} \\
 &\subseteq \{y \in V : G^{i-1}(y) = 0\} \\
 &= U_{i-1}
 \end{aligned} \tag{6.4.1}$$

Somit gilt  $G(U_d) \subseteq U_{d-1} \subseteq U_d$ . Aus der Definitionen von  $S$  und  $U_d$  folgt  $S^d(x) = G^d(x) = 0$  für alle  $x \in U_d$ .

(c2) Wir zeigen  $G(W_i) = W_{i+1}$  für alle  $i \in \mathbb{N}_0$ . Es gilt:

$$\begin{aligned}
G(W_i) &= \{G(y) : y \in W_i\} \\
&= \{G(y) : y = G^i(x), x \in V\} \\
&= \{G(G^i(x)) : x \in V\} \\
&= \{G^{i+1}(x) : x \in V\} \\
&= W_{i+1}
\end{aligned} \tag{6.4.2}$$

Demnach ist  $G(W_d) = W_{d+1} = W_d$  nach (a).  $\Rightarrow T := G|_{W_d} \in \text{Lin}(W_d)$  ist surjektiv und somit auch bijektiv (vgl. Kapitel 4).

- (d) Sei  $\mathcal{A}$  eine Basis von  $U_d$  und sei  $\mathcal{B}$  eine Basis von  $W_d$ . Sei  $(\mathcal{A}, \mathcal{B})$  die Basis, die durch das Zusammenfügen der Systeme  $\mathcal{A}$  und  $\mathcal{B}$  entsteht.

Die Matrix von  $G$  in der Basis  $(\mathcal{A}, \mathcal{B})$  hat die Form

$$\begin{aligned}
 F_{(\mathcal{A}, \mathcal{B})} &= \begin{pmatrix} S_{\mathcal{A}} & O \\ O & T_{\mathcal{B}} \end{pmatrix} \\
 \Rightarrow p_G &= \det \left( tI - \begin{pmatrix} S_{\mathcal{A}} & O \\ O & T_{\mathcal{B}} \end{pmatrix} \right) \\
 &= \det \begin{pmatrix} tI - S_{\mathcal{A}} & O \\ O & tI - T_{\mathcal{B}} \end{pmatrix} \\
 &= \det(tI - S_{\mathcal{A}}) \det(tI - T_{\mathcal{B}}) \\
 &= p_S p_T
 \end{aligned}$$

Es ist  $p_T(0) \neq 0$ , da  $T$  bijektiv ist und somit 0 nicht als Eigenwert hat. Wir zeigen nun, dass  $p_S = t^r$  gilt. Dafür werden wir die Basis  $\mathcal{A}$  von  $U_d$  auf die folgende iterative Weise wählen:

Iteration 1: Wähle eine Basis  $\mathcal{A}_1$  von  $U_1$ .

Iteration 2: Erweitere die Basis  $\mathcal{A}_1$  mit einem System  $\mathcal{A}_2$  zu einer Basis  $(\mathcal{A}_1, \mathcal{A}_2)$  von  $U_2$ , und so weiter ...

Iteration  $d$ : Erweitere die Basis  $(\mathcal{A}_1, \dots, \mathcal{A}_{d-1})$  mit einem System  $\mathcal{A}_d$  zur Basis  $(\mathcal{A}_1, \dots, \mathcal{A}_d)$  von  $U_d$ .

Wir setzen  $\mathcal{A} = (\mathcal{A}_1, \dots, \mathcal{A}_d)$ . Die Matrix  $S_{\mathcal{A}}$  hat nun die folgende Struktur:

$$S_{\mathcal{A}} = \begin{pmatrix} & \mathcal{A}_1 & \mathcal{A}_2 & & \mathcal{A}_{d-1} & \mathcal{A}_d \\ \mathcal{A}_1 & O & * & \dots & * & * \\ \mathcal{A}_2 & O & O & \dots & * & * \\ \vdots & \vdots & \ddots & & \vdots & \vdots \\ \mathcal{A}_{d-1} & O & O & \dots & O & * \\ \mathcal{A}_d & O & O & \dots & O & O \end{pmatrix},$$

wobei  $O$  einen Null-Block bezeichnet und  $*$  einen beliebigen Block. Somit ist  $S_{\mathcal{A}}$  eine obere Dreiecksmatrix mit Nullen auf der Diagonale

und man hat  $p_S = t^{\dim(U_d)}$ . Da  $p_G = p_S p_T$  mit  $p_T(0) \neq 0$  gilt, ist  $\dim(U_d)$  die algebraische Vielfachheit der Nullstelle 0 von  $p_G$ , d.h.  $\dim(U_d) = r$ .

(e) folgt aus dem Beweis von (d) (vgl.  $\dim(U_d) = r$ ) und aus (b).

Es bleibt die Ungleichung  $d \leq r$  zu zeigen. Da  $(\mathcal{A}_1, \dots, \mathcal{A}_d)$  eine Basis von  $U_d$  ist mit  $\dim(U_d) = r$ , und jedes dieser  $d$  Systeme mindestens einen Vektor enthält (vgl.  $U_1 \subsetneq \dots \subsetneq U_d$  in (a1)), gilt  $d \leq r$ .  $\square$

**Bem** (zu den entarteten Fällen). Wenn  $W_d = \{0\}$ , setzen wir  $p_T = 1$  (das macht man generell bei Abbildungen auf einem 0-dimensionalen Raum).

**Bsp.** Sei  $G \in \mathbb{K}^{3 \times 3}$  mit  $\mathbb{K} = \mathbb{R}$  die Matrix

$$\begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 1 \\ 1 & 0 & -1 \end{pmatrix}.$$

Wir interpretieren  $G$  als lineare Abbildung  $x \mapsto Gx$  des Raumes  $\mathbb{R}^3$ . Es ist  $p_G = \det(tI - G) = (t + 1)^3 - 1 = t^3 + 3t^2 + 3t = t(t^2 + 3t + 1)$ . Somit hat die Nullstelle 0 von  $p_G$  die algebraische Vielfachheit 1 und es gilt

$$\begin{aligned}\{0\} &= \ker(G^0) \subsetneq \ker(G^1) = \ker(G^2) = \dots \\ \mathbb{R}^3 &= \text{im}(G^0) \supsetneq \text{im}(G^1) = \text{im}(G^2) = \dots\end{aligned}$$

da  $d \leq r = 1$  im Lemma von Fitting. Wir berechnen eine Basis von  $\ker(G^1)$ . Es ist  $\ker(G^1) = \text{lin}((1, 1, 1)^\top)$ . Da  $\mathbb{R}^3 = \ker(G^1) \oplus \text{im}(G^1)$  gilt, ist  $\dim(\text{im}(G^1)) = 2$ . Dabei ist

$$\text{im}(G^1) = \text{lin}(\underbrace{(-1, 0, 1)^\top, (1, -1, 0)^\top, (0, 1, -1)^\top}_{\text{die Spalten von } G}) = \text{lin}(\underbrace{(-1, 0, 1)^\top, (1, -1, 0)^\top}_{\text{linear unabhängig}}),$$

d.h.  $(-1, 0, 1)^\top, (1, -1, 0)^\top$  ist eine Basis von  $\text{im}(G^1)$ . Nun können wir  $G$  in

der Basis  $\mathcal{B}$  aus  $b_1 = (1, 1, 1)^\top$ ,  $b_2 = (-1, 0, 1)^\top$ ,  $b_3 = (1, -1, 0)^\top$  darstellen:

$$G_{\mathcal{B}} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & * & * \\ 0 & * & * \end{pmatrix}$$

**Bsp.** Sei  $G(x_1, x_2, x_3, x_4) = (x_2, x_3, x_3 + x_4, x_4)$  und  $G \in \text{Lin}(\mathbb{R}^4)$ .

#### 6.4.5 Haupträume

Unter den Voraussetzungen aus [6.4.1](#) kann man nun das Lemma von Fitting auf die Abbildungen  $G := F - \lambda_i \text{id}$  für  $i = 1, \dots, k$  anwenden. Die Abbildung kann dann aus den  $k$  Abbildungen  $F|_{U_d} \in \text{Lin}(U_d)$ , mit  $U_d$  wie im Lemma 6.4.4, zusammengesetzt werden. Wir analysieren eine solche Abbildung  $F|_{U_d}$ .

**Thm.** *Sei  $V$  ein  $\mathbb{K}$ -Vektorraum mit  $\dim(V) = n \in \mathbb{N}$ ,  $F \in \text{Lin}(V)$  und  $\lambda \in \mathbb{K}$  Eigenwert von  $F$  mit algebraischer Vielfachheit  $r \in \mathbb{N}$ . Sei  $H :=$*

$\ker((F - \lambda \text{id})^r)$ . Dann gilt:

- (i)  $F(H) \subseteq H$
- (ii) Die Abbildung  $F|_H \in \text{Lin}(H)$  hat das charakteristische Polynom  $(t - \lambda)^r$ .

Beweis.

- (i) Betrachte die Abbildung  $G := F - \lambda \text{id} \in \text{Lin}(V)$ . Nach dem Verschiebungstrick 6.4.3 hat  $G$  den Eigenwert 0 mit algebraischer Vielfachheit  $r$ . Somit erfüllt  $G$  die Voraussetzungen des Lemmas von Fitting:

Betrachte die Räume  $U_i = \ker(G^i)$  wie im Lemma. Es gilt:  $H = U_r$  und weil  $d \leq r$  auch  $H = U_d$ . Nach (c1) im Lemma gilt  $G(H) \subseteq H$ , und somit auch  $F(H) = (G + \lambda \text{id})(H) \subseteq H$ .

- (ii) Nach der Behauptung (d) des Lemmas von Fitting ist das charakteristische Polynom von  $G|_H \in \text{Lin}(H)$  gleich  $t^r$ . Nach dem Verschie-

bungstrick ist das charakteristische Polynom von  $F|_H = (G + \lambda \text{id})|_H \in \text{Lin}(H)$  gleich  $(t - \lambda)^r$ .  $\square$

Der Raum  $H$  aus dem vorigen Theorem wird *Hauptraum* von  $F$  zu  $\lambda$  genannt. Es gilt:

$$\text{Eig}(F, \lambda) = \ker(F - \lambda \text{id}) \subseteq H = \ker((F - \lambda \text{id})^r) \quad (6.4.3)$$

Vektoren aus  $H$  werden auch manchmal Hauptvektoren von  $F$  zum Eigenwert  $\lambda$  genannt. Die Hauptvektoren können in Stufen zerlegt werden: hierbei wird gezählt, wie oft man  $F - \lambda \text{id}$  zum Vektor iterativ anwenden soll, um aus dem Vektor den Nullvektor zu erhalten. Die Eigenvektoren sind die Hauptvektoren der Stufe eins. Im Allgemeinen hat man aber auch Hauptvektoren höherer Stufen.

### 6.4.6 Hauptraumzerlegung

Das folgende Theorem ist der erste Schritt auf dem Weg zur JNF. Wenn das charakteristische Polynom von  $F$  in lineare Faktoren zerfällt und  $F$   $k$  verschiedene Eigenwerte hat, so kann man  $F$  in  $k$  lineare Abbildungen zerfallen lassen: um JNF aufzustellen, muss man dann noch für jede der  $k$  Abbildungen eine einfache Basisdarstellung bestimmen.

**Thm.** *Unter den Voraussetzungen 6.4.1 ist  $V$  direkte Summe der Haupträume von  $F$  zu den Eigenwerten  $\lambda_1, \dots, \lambda_k \in \mathbb{K}$ , d.h.  $V = H_1 \oplus \dots \oplus H_k$  mit  $H_i = \ker((F - \lambda_i \text{id})^{r_i})$  für  $i = 1, \dots, k$ .*

*Beweis.* Um die Formeln etwas zu vereinfachen, führe wir den Beweis exemplarisch für  $k = 3$ . Dann gilt  $p_F = (t - \lambda_1)^{r_1}(t - \lambda_2)^{r_2}(t - \lambda_3)^{r_3}$ . Man zeige zuerst, dass die Summe der Räume  $H_1, H_2, H_3$  direkt ist und anschließend, dass diese Summe mit  $V$  übereinstimmt.

Betrachte beliebige Vektoren  $v_1 \in H_1, v_2 \in H_2, v_3 \in H_3$  mit  $v_1 + v_2 + v_3 =$

0 und zeige, dass  $v_1 = v_2 = v_3 = 0$  gilt. Durch Anwendung von  $(F - \lambda_1 \text{id})^{r_1}$  zu  $v_1 + v_2 + v_3 = 0$  erhält man  $(F - \lambda_1 \text{id})^{r_1}(v_2 + v_3) = 0$ , da  $v_1 \in H_1$ . Nun wende  $(F - \lambda_2 \text{id})^{r_2}$  darauf an:  $(F - \lambda_2 \text{id})^{r_2} \circ (F - \lambda_1 \text{id})^{r_1}(v_2 + v_3) = 0$ . Da die Menge  $\mathbb{K}[F]$  ein kommutativer Ring bzgl.  $+$  und  $\circ$  auf  $\text{Lin}(V)$  ist, kann die Reihenfolge der beiden Abbildungen vertauscht werden:  $(F - \lambda_1 \text{id})^{r_1} \circ (F - \lambda_2 \text{id})^{r_2}(v_2 + v_3) = 0$ . Somit ist  $(F - \lambda_1 \text{id})^{r_1} \circ (F - \lambda_2 \text{id})^{r_2}(v_3) = 0$ , da  $v_2 \in H_2$ .

Aus dem Theorem 6.4.5 folgt, dass  $\lambda_3$  der einzige Eigenwert von  $F|_{H_3} \in \text{Lin}(H_3)$  ist. Somit ist  $\lambda_3 - \lambda_1 \neq 0$  der einzige Eigenwert von  $(F - \lambda_1 \text{id})|_{H_3} \in \text{Lin}(H_3)$  und  $\lambda_3 - \lambda_2 \neq 0$  der einzige Eigenwert von  $(F - \lambda_2 \text{id})|_{H_3} \in \text{Lin}(H_3)$ .

D.h. weder  $(F - \lambda_1 \text{id})|_{H_3} \in \text{Lin}(H_3)$  noch  $(F - \lambda_2 \text{id})|_{H_3} \in \text{Lin}(H_3)$  haben 0 als Eigenwert, d.h. die Abbildungen sind injektiv. Somit ist  $v_3 = 0$ . Analog zeigt man  $v_1 = v_2 = 0$ . Damit ist die Summe von  $H_1, H_2, H_3$  direkt und es ist  $H_1 \oplus H_2 \oplus H_3 \subseteq V$ . Zu zeigen bleibt die Gleichheit. Es gilt nach 6.4.5

oder Fitting

$$\begin{aligned}\dim(H_1 \oplus H_2 \oplus H_3) &= \dim(H_1) + \dim(H_2) + \dim(H_3) \\ &= r_1 + r_2 + r_3 \\ &= n = \dim(V),\end{aligned}$$

und damit  $H_1 \oplus H_2 \oplus H_3 = V$ . □

#### 6.4.7 Hauptraumzerlegung für Matrizen

Das vorige Theorem können wir natürlich in der Sprache der Matrizen formulieren:

**Thm.** *Sei  $A \in \mathbb{K}^{n \times n}$  eine Matrix, deren charakteristisches Polynom in Linearfaktoren zerfällt, d.h.  $p_A = \prod_{k=1}^n (t - \lambda_k)^{r_k}$  mit paarweise verschiedenen*

$\lambda_1, \dots, \lambda_k \in \mathbb{K}$ . Dann existiert eine invertierbare Matrix  $B \in \mathbb{K}^{n \times n}$  mit

$$B^{-1}AB = \begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_k \end{pmatrix}$$

wobei  $A_i \in \mathbb{K}^{r_i \times r_i}$  eine Matrix ist mit charakteristischem Polynom  $p_{A_i} = (t - \lambda_i)^{r_i}$  für jedes  $i \in \{1, \dots, k\}$ .

*Beweis.* Folgt direkt aus dem vorigen Theorem. □

Die Matrix  $\begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_k \end{pmatrix}$  wie im vorigen heißt *blockdiagonal* mit *Diagonallöcken*  $A_1, \dots, A_k$ .

Es bietet sich an, ein Beispiel zu betrachten.

**Bsp.** Die Matrix

$$A = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 3 \\ 0 & 0 & 2 \end{pmatrix} \in \mathbb{Q}^{3 \times 3}$$

hat ein charakteristisches Polynom, das bzgl. des Körpers  $\mathbb{K} = \mathbb{Q}$ , in Linearfaktoren zerfällt: man hat

$$p_A = (t - 1)^2 \cdot (t - 2).$$

Die Theorie sagt uns also, dass die Matrix  $A$  in einer Basis in zwei Blöcke zerfällt, mit den Größen  $2 \times 2$  und  $1 \times 1$ . Mit dem Block der Größe  $1 \times 1$  hat man nicht viel Arbeit: dieser Block ist nichts anderes als der Eigenwert 2, in die Basis wird also ein Eigenvektor zu 2 aufgenommen. Die Bestimmung der größeren Blöcke kann unter Umständen etwas aufwändiger sein.

Wir betrachten die Matrizen  $A - 1 \cdot I$  und bestimmen die Basis des  $2 \times 2$  Blocks. Dafür berechnen wir. Es gilt

$$\dim(\ker(A - 1 \cdot I)^1) = \dim \left( \ker \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 3 \\ 0 & 0 & 1 \end{pmatrix} \right) = 1$$

Wegen

$$(A - 1 \cdot I)^2 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 3 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 3 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 3 \\ 0 & 0 & 3 \\ 0 & 0 & 1 \end{pmatrix}$$

gilt

$$\dim(\ker(A - 1 \cdot I)^2) = 2.$$

Die Dimensionen von  $\ker(A - 1 \cdot I)^i$  für  $i \geq 2$  müssen wir nicht ausrechnen, denn Theorie sagt uns, dass sich diese Dimensionen ab  $i$  gleich der algebraischen Vielfachheit nicht mehr ändern. Die algebraische Vielfachheit von 1 ist 2.

Damit gilt

$$\{0\} = \ker(A - 1 \cdot I)^0 \subsetneq \ker(A - 1 \cdot I)^1 \subsetneq \ker(A - 1 \cdot I)^2 = \ker(A - 1 \cdot I)^3 = \dots$$

Wir werden eine Basis von  $\ker(A - 1 \cdot I)^2$  benutzen. Da wir  $(A - 1 \cdot I)^2$  explizit ausgerechnet haben, sieht man direkt, dass  $e_1, e_2$  eine Basis von  $\ker(A - 1 \cdot I)^2$  ist.

Für die Hauptraumzerlegung von  $A$  brauchen wir noch einen Vektor für den  $1 \times 1$  Block zum Eigenwert 2 (den Eigenvektor zu 2). Da 2 die algebraische Vielfachheit 1 hat, gilt:

$$\{0\} = \ker(A - 2I)^0 \subsetneq \underbrace{\ker(A - 2I)^1}_{\text{Das ist der Hauptraum}} = \ker(A - 2I)^2 = \dots$$

zu 2; seine Dimension ist  
daher gleich 1.

Wir bestimmen nun eine Basis von  $\ker(A - 2I)^1 = \ker \begin{pmatrix} -1 & 1 & 0 \\ 0 & -1 & 3 \\ 0 & 0 & 0 \end{pmatrix}$ .

Man sieht dass der Vektor  $(3, 3, 1)^\top$  ist eine Basis dieses (eindimensionalen) Kerns von  $A - 2I$  bildet.

Die Hauptraumzerlegung erreicht man also mit der Basis  $\mathcal{B} = (e_1, e_2, (3, 3, 1)^\top)$ . D.h. für die invertierbare Matrix  $B = (e_1, e_2, (3, 3, 1)^\top)$  hat man

$$B^{-1}AB = \left( \begin{array}{cc|c} 1 & 1 & 0 \\ 0 & 1 & 0 \\ \hline 0 & 0 & 2 \end{array} \right)$$

Fragen: wie bestimmt man den  $2 \times 2$  Block? Wie würde der  $2 \times 2$  Block aussehen, wenn wir die Basis  $(1, 1, 0)^\top, (1, 0, 0)^\top, (3, 3, 1)^\top$  benutzen?

### 6.4.8 Jordansche Normalform für nilpotente Abbildungen

Eine lineare Abbildung  $G$  eines Vektorraums  $V$  heißt *nilpotent*, falls  $G^i$  für ein  $i \in \mathbb{N}$  eine Nullabbildung ist.

Die Matrix der Form

$$\begin{pmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix}$$

heißt *Jordan-Matrix* der Größe  $n \in \mathbb{N}$  zum Wert  $\lambda \in \mathbb{K}$ . Wir leiten in diesem Abschnitt die JNF für nilpotente lineare Abbildungen her. Im nächsten Paragraphen erhalten wir als Folgerung daraus die JNF für allgemeine lineare Abbildungen. Um die Methode zum Aufbauen einer JNF zu verstehen, lohnt es sich eine konkrete nilpotente Abbildung anzuschauen, die durch ihre JNF gegeben ist.

**Bsp.** Dieses Beispiel soll vorberieten, den Beweis des nachfolgenden Theo-

rems zu verstehen. Wir betrachten eine nilpotente Abbildung eines 6-dimensionalen Vektorraums, die in einer Basis  $\mathcal{B} = (b_1, \dots, b_6)$  durch die folgende Blockdiagonalmatrix gegeben ist:

$$G_{\mathcal{B}} = \begin{matrix} & b_1 & b_2 & b_3 & b_4 & b_5 & b_6 \\ b_1 & 0 & 1 & 0 & & & \\ b_2 & 0 & 0 & 1 & & & \\ b_3 & 0 & 0 & 0 & & & \\ b_4 & & & & 0 & 1 & \\ b_5 & & & & 0 & 0 & \\ b_6 & & & & & & 0 \end{matrix}$$

Die Matrix hat drei Jordan-Matrizen zum Wert 0 als Diagonalblöcke. Die Matrix sagt uns, dass die Abbildung  $G$  auf den Basisvektoren auf die fol-

gende Weise wirkt:

$$\begin{aligned} b_3 &\mapsto b_2 \mapsto b_1 \mapsto 0 \\ b_5 &\mapsto b_4 \mapsto 0 \\ b_6 &\mapsto 0 \end{aligned}$$

Wie hängt diese Wirkung mit den Vektorräumen  $U_1, \dots, U_d$  aus dem Lemma vom Fitting zusammen? Man beachte, dass der Vektorraum  $U_d$  aus dem Lemma von Fitting mit dem gesamten Raum  $\mathbb{K}^6$  übereinstimmt, weil unser Abbildung nilpotent ist (jeder Vektor liegt in  $U_d$ , weil jeder Vektor auf 0 abgebildet wird, wenn man die Abbildung  $G$  zu dem Vektor oft genug anwendet). Wir sehen, dass in unserem Beispiel  $d$  drei ist, denn nach einer dreifachen Anwendung von  $G$  wird jeder Basisvektor gleich 0 (und somit auch jeder andere Vektor unseres Vektorraums), und bei  $b_3$  muss man die Abbildung dreimal anwenden, um den Nullvektor zu erhalten.

Also ist  $U_3 = \mathbb{K}^6$  und  $U_2$  eine echte Teilmenge von  $U_3$ . Der Vektorraum  $U_2$  entsteht aus all den Basisvektoren, bei denen die zweifache Anwendung

der Abbildung ausreicht, um den Nullvektor zu erhalten. Man hat also  $U_2 = \text{lin}(b_2, b_1, b_5, b_4, b_6)$ . Und  $U_1$  entsteht aus den Basisvektoren, bei denn eine Anwendung der Abbildung ausreicht, um den Nullvektor zu erhalten. Man hat also  $U_1 = \text{lin}(b_1, b_4, b_6)$ .

An diesem Beispiel haben wir den Weg zur Konstruktion der JNF rückwärts verfolgt. Bei der Eigentlichen Aufgaben ist eine nilpotente Abbildung  $G$  gegeben. Dann berechnet man  $U_1, \dots, U_d$  und will anschließend eine JNF  $G_{\mathcal{B}}$  und die entsprechende Basis  $\mathcal{B}$  berechnen.

Wie kommt man an die Basis  $\mathcal{B}$  anhand der Vektorräume  $U_1, \dots, U_d$ ? Wir brauchen geeignete Räume, welche die ‘Lücken’ zwischen aufeinanderfolgenden Vektorräumen der Folge  $U_1, \dots, U_d$  ‘stopfen’. In unserem Beispiel stopft  $\text{lin}(b_3)$  die Lücke zwischen  $U_3$  und  $U_2$ ,  $\text{lin}(b_2, b_5)$  stopft die Lücke zwischen  $U_2$  und  $U_1$ , und  $U_1$  ist als  $\text{lin}(b_1, b_4, b_6)$  beschrieben. Die Basisvektoren entstehen also durch das Stopfen der Lücken. Die Lücke zwischen  $U_2$  und  $U_3 = \mathbb{K}^6$  war eindimensional. Wir benötigen also einen Vektor um, eine be-

liebige Basis von  $U_2$ , zu einer Basis vom gesamten Vektorraum zu erweitern. Diesen Vektor haben wir als  $b_3$  in unserem Beispiel bezeichnet. Wir benutzen bei dem Aufbau der Basis durchgängig das folgende Prinzip: Wenn wir einen Vektor zur Basis  $\mathcal{B}$  hinzugefügt haben, kommen auch die Bilder des Vektors in die Basis, welche durch die iterative Anwendung der Abbildung  $G$  entstehen. Wir fügen also auch  $G(b_3) = b_2$  und  $b_1 = G^2(b_3) = G(b_2)$  zur Basis hinzu. Den Vektor  $G^3(b_3) = 0$  dürfen wir natürlich nicht hinzufügen: in Basen hat man keine Nullvektoren. Was gilt nun für die neu hinzugefügten Vektoren:  $b_2$  stopft zum Teil die Lücke zwischen  $U_2$  und  $U_1$ , aber nicht komplett.  $b_1$  liegt in  $U_1$  erzeugt aber diesen Raum nicht. Wir beschäftigen uns also mit der nächsten Lücke: in der Lücke zwischen  $U_2$  und  $U_1$  liegt bereits  $b_2$  und wir fügen  $b_5$  noch hinzu, um diese Lücke auszufüllen. Nach dem Hinzufügen von  $b_5$  fügen wir auch sein Bild  $b_4 = G(b_5)$  hinzu. Der Vektor  $b_4$  liegt in  $U_1$ . Nun hat man zwei Vektoren, und zwar  $b_1$  und  $b_4$ , in  $U_1$  gewählt. Diese beiden Vektoren erzeugt aber nicht  $U_1$ . Es wird als ein

weiterer Vektoren hinzugefügt (in unserem Beispiel ist das nur ein Vektor  $b_6$ ), sodass man nach dieser Ergänzung Vektoren  $b_1, b_4, b_6$  erhält, die  $U_1$  erzeugen.

Während der Konstruktion lässt man also nach und nach Ketten von Vektoren wachsen, wobei jeder der Vektoren einer der Lücken zugeordnet werden kann.

In unserem Beispiel war hatten die Räume  $U_1, \dots, U_d$  die folgenden Dimensionen:  $\dim(U_3) = 6, \dim(U_2) = 5, \dim(U_1) = 3$ . Die Lücken haben also die Dimensionen  $\dim(U_3) - \dim(U_2) = 6 - 5 = 1, \dim(U_2) - \dim(U_1) = 5 - 3 = 2$  und  $\dim(U_1) = 3$ . Wenn man uns  $G$  wie im Beispiel geben würde, so würden wir nach der Bestimmung von  $U_1, U_2, U_3$  die folgende erste Kette konstruieren:

$$b_3 \mapsto b_2 \mapsto b_1 \mapsto 0$$

(Die Wahl des Vektors  $b_3$  der Kette ist übrigens nicht eindeutig.)

Durch diese Kette haben wir in jeder der drei Lücken eine Dimension

abgedeckt.

Dann kämme noch eine Kette hinzu, und wir hätten:

$$\begin{aligned} b_3 &\mapsto b_2 \mapsto b_1 \mapsto 0 \\ b_5 &\mapsto b_4 \mapsto 0 \end{aligned}$$

Lücke Zwischen  $U_3$  und  $U_1$  ist in diesem Punkt komplett ausgefüllt, sowie die Lücke zwischen  $U_2$  und  $U_1$ . Im letzten Schritt würde wir noch eine Kette hinzufügen, um  $U_1$  auszufüllen:

$$\begin{aligned} b_3 &\mapsto b_2 \mapsto b_1 \mapsto 0 \\ b_5 &\mapsto b_4 \mapsto 0 \\ b_6 &\mapsto 0 \end{aligned}$$

Bei diesem Prozess entstand pro Schritt genau eine neue Kette. Im Allgemeinen können Verschiedene Situationen auftreten, etwa, dass man in einem der Schritt gar keine Kette oder mehrere Ketten hinzufügt. Wie viel Ketten in im  $i$ -ten Schritt hinzugefügt werden (für  $i \in \{1, \dots, d\}$ ), hängt von den Dimensionen der Vektorräume  $U_1, \dots, U_d$  ab.

**Thm.** Sei  $G$  eine nilpotente lineare Abbildung eines  $n$ -dimensionalen Vektorraums  $V$  über  $\mathbb{K}$  ( $n \in \mathbb{N}$ ). Dann existiert eine Basis  $\mathcal{B}$  von  $V$  derart, dass  $G_{\mathcal{B}}$  blockdiagonal und jeder Block von  $G_{\mathcal{B}}$  eine Jordan-Matrix zum Wert 0 ist. Die Matrix  $G_{\mathcal{B}}$  ist durch  $G$  bis auf die Reihenfolge der Blöcke eindeutig bestimmt.

*Beweis.* Die Existenz von  $\mathcal{B}$  un die Eindeutigkeit der Matrix  $G_{\mathcal{B}}$  muss gezeigt werden.

Existenz: Bei der Herleitung der Existenz von  $\mathcal{B}$  lehnen wir uns an das Lemma von Fitting an und übernehmen die Bezeichnungen daraus. Das heißt, wir benutzen die Räume  $U_i = \ker(G^i)$  und  $W_i = \text{im}(G^i)$  für und die Konstante  $d \in \mathbb{N}$  mit der Eigenschaft

$$\{0\} = U_0 \subsetneq \dots \subsetneq U_d = U_{d+1} = \dots$$

$$V = W_0 \supsetneq \dots \supsetneq W_d = W_{d+1} = \dots$$

Aus der Nilpotenz von  $G$  folgt  $U_d = V$ . Wegen  $V = U_d \oplus W_d$  hat man somit

$$W_d = \{0\}.$$

Im entarteten Fall  $d = 1$ , ist  $U_1 = V$ , das heißt, der Kern von  $G$  ist der gesamte Vektorraum  $V$ . Das bedeutet,  $G$  ist Nullabbildung, dass  $G_B$  für jede Basis eine Nullmatrix ist. Der Wert 0 ist Jordan-Matrix der Größe 1 zum Wert 1. Die  $n \times n$  Nullmatrix ist also blockdigonal mit  $n$  Jordan-Blöcken der Größe 1 zum Wert 0.

Als Nächstes betrachten wir den nicht-entarteten Fall  $d \geq 2$ . Wir befassen uns mit den ‘Lücken’ in der Kette  $U_1 \subsetneq U_2 \subsetneq \cdots \subsetneq U_{d-1} \subsetneq U_d$  und als erste nehmen wir die Lücke zwischen  $U_{d-1}$  und  $U_d$  unter die Lupe. Diese wird mit einem Vektorraum ‘gestopft’, den wir  $V_d$  nennen. Das heißt, wir wählen einen Untervektorraum  $V_d$  von  $U_d$ , mit der Eigenschaft  $U_d = U_{d-1} \oplus V_d$ .

Wegen  $V_d \subseteq U_d$ , werden die Vektoren aus  $V_d$  durch  $d$ -fache Anwendung von  $G$  auf 0 abgebildet. Daraus folgt  $G(V_d) \subseteq U_{d-1}$ , oder mit Worten: die Vektoren aus  $G(V_d)$  werden durch  $(d - 1)$ -fache Anwendung von  $G$  auf 0 abgebildet. Der Vektorraum  $U_{d-1}$  enthält somit die Untervektorräume  $U_{d-2}$

und  $G(V_d)$  als Untervektorräume. Es stellt sich heraus, dass die Summe der Vektorräume  $U_{d-2}$  und  $G(V_d)$  direkt ist. Um das zu sehen, reicht es  $U_{d-2} \cap G(V_d) = \{0\}$  zu verifizieren. Wir betrachten einen beliebigen Vektor  $x \in G(V_d) \cap U_{d-2}$ . Nach der Wahl von  $x$ , gilt  $x = G(v)$  für ein  $v \in V_d$  und  $G^{d-2}(x) = 0$ . Wenn wir den Ausdruck für  $x$  in  $G^{d-2}(x) = 0$  einsetzen, erhalten wir  $0 = G^{d-2}(G(v)) = G^{d-1}(v)$ . Das ergibt  $v \in U_{d-1}$ . Somit liegt  $v$  in  $V_d$  und  $U_{d-1}$ . Da die Summe von  $V_d$  und  $U_{d-1}$  nach der Wahl von  $V_d$  direkt ist, folgt nun  $v = 0$ . Somit ist auch  $x = G(v)$  gleich 0. Das zeigt  $U_{d-2} \cap G(V_d) = \{0\}$ .

Im Vektorraum  $U_{d-1}$  ist also die direkte Summe  $U_{d-2} \oplus G(V_d)$  als Untervektorraum enthalten. Nun erweitern wir  $G(V_d)$  zu einem Untervektorraum  $V_{d-1}$ , um  $U_{d-1}$  als direkte Summe von  $U_{d-2}$  und  $V_{d-1}$  darzustellen, wir wählen also einen Vektorraum  $V_{d-1}$  mit  $V_{d-1} \supseteq G(V_d)$  und  $U_{d-1} = U_{d-2} \oplus V_{d-1}$ .  $V_{d-1}$  ‘stopft’ somit die Lücke zwischen  $U_{d-2}$  und  $U_{d-1}$  und  $G$  bildet den Vektorraum  $V_d$ , in der Lücke zwischen  $U_d$  und  $U_{d-1}$ , in den Raum  $V_{d-1}$ ,

in der Lücke zwischen  $U_{d-1}$  und  $U_{d-2}$  ab. Wir zeigen noch zusätzlich, dass die Einschränkung von  $G$  auf  $V_d$  injektiv wirkt. Dafür betrachten wir ein beliebiges  $x \in V_d$  mit  $G(x) = 0$ . Wenn wir zur Gleichung  $G(x) = 0$  die Abbildung  $G^{d-2}$  anwenden, erhalten wir  $G^{d-1}(x) = 0$ . Das ergibt  $x \in U_{d-1}$ . Der Vektor  $x$  liegt also in  $V_d \cap U_{d-1}$ . Da die Summe von  $V_d$  und  $U_{d-1}$  direkt ist, hat man  $V_d \cap U_{d-1} = \{0\}$ . Das zeigt  $x = 0$ . Die Abbildung  $G$  ist also in der Tat injektiv auf  $V_d$ .

Nun setzen wir die oben beschriebene Konstruktion der Räume  $V_d, V_{d-1}, \dots$  fort und erhalten die Räume  $V_d, \dots, V_1$  mit den folgenden Eigenschaften:

- (i)  $U_i = U_{i-1} \oplus V_i$  für alle  $i \in \{1, \dots, d\}$
- (ii)  $G|_{V_i}$  ist injektiv und  $G(V_i) \subseteq V_{i-1}$  für alle  $i \in \{2, \dots, d\}$ .
- (iii)  $G|_{V_1}$  ist eine Nullabbildung.

Informell: die Räume  $V_d, \dots, V_1$  stopfen die Lücken, jeder Raum  $V_i$  wird

injektiv in in den nächsten Raum  $V_{i-1}$ , es sei denn das war der Raum  $V_1$  ( $V_1$  wird auf 0 abgebildet).

Insbesondere folgt aus (i) die Gleichheit  $U_i = V_1 \oplus \dots \oplus V_i$  (wegen (i) und  $U_0 = \{0\}$ ) und, für den Fall  $i = d$ , die Gleichheit  $V = U_d = V_1 \oplus \dots \oplus V_d$ . Anhand von  $V_d, \dots, V_1$  können wir nun eine gewünschte Basis  $\mathcal{B}$  fixieren. Das machen wir iterativ folgendermaßen.

Wir fixieren eine Basis  $\mathcal{B}_d$  von  $V_d$ . Da  $G$  auf  $V_d$  injektiv ist, ist  $G(\mathcal{B}_d)$  ein linear unabhängiges System in  $V_{d-1}$ . Da  $G$  auf  $V_{d-1}$  injektiv ist, ist  $G(G(\mathcal{B}_d)) = G^2(\mathcal{B}_d)$  ein linear unabhängiges System in  $V_{d-2}$  usw. Wir erhalten also die folgenden linear unabhängigen Systeme:

$\mathcal{B}_d$  in  $V_d$

$G(\mathcal{B}_d)$  in  $V_{d-1}$

$\vdots$

$G^{d-1}(\mathcal{B}_d)$  in  $V_1$ .

Das System  $\mathcal{B}_d$  ist bereits eine Basis von  $V_d$ , das nächste System  $G(\mathcal{B}_d)$  ist aber im Allgemeinen keine Basis von  $V_{d-1}$ . Wir ergänzen also  $G(\mathcal{B}_d)$  zu einer Basis  $(G(\mathcal{B}_d), \mathcal{B}_{d-1})$  von  $V_{d-1}$  (wenn  $G(\mathcal{B}_d)$  bereits eine Basis von  $V_{d-1}$  ist, ist  $\mathcal{B}_{d-1}$  leer). Nun konstruieren wir anhand  $\mathcal{B}_{d-1}$  linear unabhängige System  $G(\mathcal{B}_{d-1}), \dots, G^{d-2}(\mathcal{B}_{d-1})$  mit den folgenden Eigenschaften:

$\mathcal{B}_d$  ist Basis von  $V_d$

$(G(\mathcal{B}_d), \mathcal{B}_{d-1})$  ist Basis von  $V_{d-1}$

$(G^2(\mathcal{B}_d), G(\mathcal{B}_{d-1}))$  ist linear unabhängiges System in  $V_{d-1}$

$\vdots$

$(G^{d-1}(\mathcal{B}_d), G^{d-2}(\mathcal{B}_{d-2}))$  ist linear unabhängiges System in  $V_1$ .

Dieser Prozess lässt sich iterativ fortsetzen. Wir erhielten eine Basis für  $V_1$ , dann eine Basis für  $V_{d-1}$ , als nächstes erhalten wir eine Basis von  $V_{d-2}$

usw. Am Ende des Prozesses haben wir die Vektorsysteme  $\mathcal{B}_d, \dots, \mathcal{B}_1$ , für welche die folgenden Eigenschaften erfüllt sind:

$\mathcal{B}_d$  ist Basis von  $V_d$

$(G(\mathcal{B}_d), \mathcal{B}_{d-1})$  ist Basis von  $V_{d-1}$

$(G^2(\mathcal{B}_d), G(\mathcal{B}_{d-1}), \mathcal{B}_{d-2})$  ist Basis von  $V_{d-1}$

$\vdots$

$(G^{d-1}(\mathcal{B}_d), G^{d-2}(\mathcal{B}_{d-2}), \dots, \mathcal{B}_1)$  ist Basis von  $V_1$ .

Durch das Zusammenfügen der oben gewählten Basen der Vektorräume  $V_d, \dots, V_1$  entsteht eine Basis  $\mathcal{B}$  von  $V = V_1 \oplus \dots \oplus V_d$ , für welche die Behauptung des Theorems erfüllt ist. Schauen wir uns mal an, wie die Abbildung auf den Vektoren unserer Basis  $\mathcal{B}$  wirkt. Nach der Konstruktion

bilden die Basisvektoren bzgl. der Wirkung von  $G$  die folgenden Ketten:

$$\begin{array}{ccccccc}
 b & \mapsto & G(b) & \mapsto & \dots & \mapsto & G^{i-1}(b) \mapsto G^i(b) = 0 \\
 \cap & & \cap & & & & \cap \\
 \mathcal{B}_i & & G(\mathcal{B}_i) & & & & G^{i-1}(\mathcal{B}_i) \\
 \cap & & \cap & & & & \cap \\
 V_i & & V_{i-1} & & & & V_1
 \end{array}$$

Bei einer passenden Reihung der Vektoren hat also die Matrix  $G_{\mathcal{B}}$  die Blockdiagonalstruktur, und genau  $|\mathcal{B}_i|$  Diagonalblöcke von  $G_{\mathcal{B}}$  sind Jordan-Matrizen der Größe  $i$  zum Wert 0.

Eindeutigkeit: Die Eindeutigkeit ist eine Nebenbemerkung zum Beweis der Existenz. Die Räume  $U_1, \dots, U_d$  und ihre Dimensionen sind eindeutig durch  $G$  bestimmt. Die Räume  $V_1, \dots, V_d$  sind zwar nicht eindeutig durch  $G$  bestimmt, ihre Dimensionen sind aber eindeutig; denn aus  $U_i = U_{i-1} \oplus V_i$  folgt  $\dim(V_i) = \dim(U_i) - \dim(U_{i-1})$ . Ist  $\mathcal{B}$  eine Basis, in der die Matrix  $G_{\mathcal{B}}$  die in der Behauptung beschriebene Struktur hat, so kann man die

Basis  $\mathcal{B}$  genau so wie oben im Existenzbeweis mit Hilfe von Vektorsystemen  $\mathcal{B}_1, \dots, \mathcal{B}_d$  strukturieren und anhand der Vektoren aus  $\mathcal{B}$  entsprechende Vektorräume  $V_d, \dots, V_1$  aufspannen, für welche  $U_i = U_{i-1} \oplus V_i$  erfüllt ist. Die Anzahl der Vektoren in  $\mathcal{B}_d$  ist gleich der Dimension von  $V_d$  und somit eindeutig durch  $G$  bestimmt. Die Anzahl der Vektoren in der Basis  $(G(\mathcal{B}_d), \mathcal{B}_{d-1})$  ist die Dimension von  $V_{d-1}$  und somit eindeutig durch  $G$  bestimmt. Die Anzahl der Vektoren in  $\mathcal{B}_{d-1}$  ist die Anzahl der Vektoren in  $(G(\mathcal{B}_d), \mathcal{B}_{d-1})$  (gleich  $\dim(V_{d-1})$ ) minus die Anzahl der Vektoren in  $G(\mathcal{B}_d)$  (gleich  $\dim(V_d)$ ). Somit ist die Anzahl der Vektoren in  $\mathcal{B}_{d-1}$  gleich  $\dim(V_d) - \dim(V_{d-1})$ , dieser Wert ist eindeutig durch  $G$  bestimmt. Eine iterative Fortsetzung dieses Arguments zeigt, dass die Anzahl der Vektoren in jedem der  $d$  Systeme  $\mathcal{B}_d, \dots, \mathcal{B}_1$  eindeutig durch  $G$  bestimmt ist. Die Anzahl der Vektoren in  $|\mathcal{B}_i|$  ist aber die Anzahl der Diagonalblöcke der Größe  $i$ . Wir erhalten also die Eindeutigkeit.  $\square$

Die Darstellung  $G_{\mathcal{B}}$  von  $G$  aus dem vorigen Theorem heißt die JNF von

G. JNF für allgemeine lineare Abbildungen werden wir im folgenden Paragraphen einführen.

Durch die Identifikation von Matrizen mit linearen Abbildungen können wir natürlich von nilpotenten Matrizen sprechen und von JNF solcher Matrizen.

Bsp.

- (i) Wir berechnen die JNF der (nilpotenten) Matrix

$$G = \begin{pmatrix} 0 & 1 & 3 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix} \in \mathbb{R}^{3 \times 3}.$$

Die Matrix ist tatsächlich nilpotent. Das kann man direkt überprüfen, oder durch Cayley-Hamilton. Das charakteristische Polynom ist  $t^3$ , also gilt nach Cayley-Hamilton  $G^3 = 0$ . Im Lemma von Fitting werden die

Räume  $U_i$  spätestens ab  $i = 3$  gleich  $\mathbb{R}^3$ . Eine genauere Analyse ergibt

$$\{0\} = U_0 \subsetneq U_1 \subsetneq U_2 \subsetneq U_3 = \mathbb{R}^3,$$

da  $\dim(U_1) = 1$  und  $\dim(U_2) = 2$  gilt. Die ‘Lücke’ zwischen  $U_3$  und  $U_2$  sowie zwischen  $U_2$  und  $U_1$  ist also ein-dimensional. Wenn wir also einen Vektor aus  $U_3 \setminus U_2$  und dann  $G$  zu diesem Vektor iterativ anwenden, stopfen wir alle Lücken. Somit hat die JNF von  $G$  einen einzigen Block: die Jordan-Matrix der Größe 3 zum Wert 0:

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

Wie die JNF aussieht wissen wir also, eher wir eine konkrete Basis dazu fixieren. Um eine Basis zu konstruieren, wählen wir einen Vektor aus  $U_3 \setminus U_2$ . Das ist ein Vektor, der durch eine zweifache Anwendung

von  $G$  nicht gleich 0 wird. Es gibt sehr viele solche Vektoren, zum Beispiel

$e_3$  ( $G^2e_3 = 2e_1 \neq 0$ ). Durch die iterative Anwendung von  $G$  zu  $e_3$  entstehen die Vektoren.

$$Ge_3 = \begin{pmatrix} 3 \\ 2 \\ 0 \end{pmatrix} \quad G^2e_3 = \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}$$

Die Matrix hat also die JNF

$$G_B = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$$

in der Basis

$$\mathcal{B} = \left( \left( \begin{pmatrix} 2 \\ 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 2 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \right) \right).$$

(ii) Wir bestimmen die JNF der (nilpotenten) Matrix

$$G = \begin{pmatrix} 4 & 1 & -1 \\ -8 & -2 & 2 \\ 8 & 2 & -2 \end{pmatrix} \in \mathbb{R}^{3 \times 3}.$$

Die Nilpotenz kann mit Hilfe von Cayley-Hamilton verifiziert werden, denn das charakteristische Polynom dieser Matrix ist  $t^3$ , sodass man  $G^3 = 0$  hat. Die Anwendung von Cayley-Hamilton sagt uns aber nicht, welche kleinste Potenz der Matrix  $G$  gleich 0. Es stellt sich heraus, dass in diesem Beispiel bereits die zweite Potenz gleich 0 ist: man kann direkt verifizieren, dass  $G^2 = 0$  ist. Wir brauchen also nur die Räume  $U_1$  und  $U_2$ . Deren Dimensionen sind:

$$\dim(U_1) = \dim(\ker(G)) = 2$$

$$\dim(U_2) = \dim(\ker(G^2)) = 3$$

Wir wählen einen Vektor in  $U_2 \setminus U_1$ , z.B.  $e_1$  ist ein solcher Vektor. Eine iterative Anwendung von  $G$  zu diesem Vektor ergibt noch den Vektor

$$Ge_1 = \begin{pmatrix} 4 \\ -8 \\ 8 \end{pmatrix}$$

Es ist klar, dass  $G^2e_1 = 0$  ist. Zwei Vektoren haben wir bereits in der Basis unserer JNF fixiert. Uns fehlt ein dritter Vektor:

$$\mathcal{B} = \left( \underbrace{\begin{pmatrix} 4 \\ -8 \\ 8 \end{pmatrix}}_{b_1}, \underbrace{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}}_{b_2}, \underbrace{\begin{pmatrix} ? \\ ? \\ ? \end{pmatrix}}_{b_3} \right)$$

Als  $b_3$  wählt man einen Vektor aus  $U_1$ , der zum Vektor  $b_1 \in U_1$  linear

unabhängig ist, z.B.  $b_3 = e_1 - 4e_2$ . Also hat  $G$  in der Basis

$$\mathcal{B} = \left( \underbrace{\begin{pmatrix} 4 \\ -8 \\ 8 \end{pmatrix}}_{b_1}, \underbrace{\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}}_{b_2}, \underbrace{\begin{pmatrix} 1 \\ -4 \\ 0 \end{pmatrix}}_{b_3} \right)$$

die JNF

$$G_{\mathcal{B}} = \begin{pmatrix} 0 & 1 & \\ 0 & 0 & \\ & & 0 \end{pmatrix}$$

Die JNF hat zwei Diagonalblöcke der Größen 2 und 1.

### 6.4.9 Jordansche Normalform für allgemeine lineare Abbildungen

Die Matrix  $F_{\mathcal{B}}$  im folgenden Theorem heißt die JNF von  $F$ .

**Thm.** *Seien die Voraussetzungen aus 6.4.1 erfüllt. Dann existiert eine Basis  $\mathcal{B}$  von  $V$ , sodass  $F_{\mathcal{B}}$  blockdiagonal ist, wobei die Diagonalblöcke Jordan-Matrizen zu den Werten  $\lambda_1, \dots, \lambda_k$  sind und für jedes  $i \in \{1, \dots, k\}$  die Gesamtgröße der Jordan-Matrizen zum Wert  $\lambda_i$  gleich  $r_i$  ist. Die Matrix  $F_{\mathcal{B}}$  wie oben ist bis auf die Reihenfolge der Blöcke durch  $F$  eindeutig bestimmt.*

*Beweis.*

Existenz von  $\mathcal{B}$ : Wir betrachten die Hauträume  $H_i := \ker(F - \lambda_i \text{id})^{r_i}$  mit  $i \in \{1, \dots, k\}$ . Die Abbildung  $F_i := F|_{H_i} \in \text{Lin}(H_i)$  hat das charakteristische Polynom  $(t - \lambda_i)^{r_i}$ . Daher hat die Abbildung  $G_i := F_i - \lambda_i \text{id}$  das charakteristische Polynom  $t^{r_i}$ . Nach dem Satz von Cayley-Hamilton ist  $G_i^{r_i} = 0 \Rightarrow G_i$  ist nilpotent. Nach Theorem 6.4.8 besitzt

$H_i$  eine Basis  $\mathcal{B}_i$ , für welche die Matrix  $(G_i)_{\mathcal{B}_i}$  Jordansche Normalform hat (mit Jordan-Matrizen zum Wert 0). Demnach hat  $(F_i)_{\mathcal{B}_i} = (G_i)_{\mathcal{B}_i} + \lambda_i I$  Jordansche Normalform (mit Jordan-Matrizen zum Wert  $\lambda_i$ ).

Da  $V = H_1 \oplus \dots \oplus H_k$  gilt, ist  $\mathcal{B} = (\mathcal{B}_1, \dots, \mathcal{B}_k)$  die gesuchte Basis von  $V$ .

Eindeutigkeit von  $F_{\mathcal{B}}$  (Beweisskizze): Man kehrt die obige Konstruktion um.

Erklärung an einem Beispiel:

$$F_{\mathcal{B}} = \begin{pmatrix} b_1 & b_2 & b_3 & b_4 & b_5 & b_6 \\ b_1 & \lambda_1 & 1 & & & \\ b_2 & 0 & \lambda_1 & & & \\ b_3 & & & \lambda_1 & & \\ b_4 & & & & \lambda_2 & 1 & 0 \\ b_5 & & & & 0 & \lambda_2 & 1 \\ b_6 & & & & 0 & 0 & \lambda_2 \end{pmatrix}$$

Man sieht:  $b_1, b_2, b_3$  ist Basis von  $H_1$  und  $b_4, b_5, b_6$  ist Basis von  $H_2$ .

Darüber hinaus sieht man, dass  $(F - \lambda_1 \text{id})|_{H_1} \in \text{Lin}(H_1)$  und  $(F - \lambda_2 \text{id})|_{H_2}$  nilpotent sind.

Somit folgt die Eindeutigkeit von  $F_{\mathcal{B}}$  aus der Eindeutigkeit von den Hauträumen und aus der Eindeutigkeit im nilpotenten Fall.  $\square$

### 6.4.10 Jordansche Normalform für Matrizen

Natürlich kann man die Existenz der JNF komplett in der Sprache der Matrizen formulieren. Das machen wir im folgenden Theorem.

**Thm.** Sei  $A \in \mathbb{K}^{n \times n}$  ( $n \in \mathbb{N}$ ) eine Matrix, deren charakteristisches Polynom in Linearfaktoren zerfällt, d.h.  $p_A = \prod_{i=1}^k (t - \lambda_i)^{r_i}$  mit  $k \in \mathbb{N}$ ,  $\lambda_1, \dots, \lambda_k \in \mathbb{K}$  und  $r_1, \dots, r_k \in \mathbb{N}$ . Dann existiert eine reguläre Matrix  $B \in \mathbb{K}^{n \times n}$  derart, dass  $B^{-1}AB$  blockdiagonal ist, wobei die Diagonalblöcke Jordan-Matrizen zu den Werten  $\lambda_1, \dots, \lambda_k$  sind und für jedes  $i \in \{1, \dots, k\}$  die Gesamtgröße der Jordan-Matrizen zum Wert  $\lambda_i$  gleich  $r_i$  ist. Die Matrix  $B^{-1}AB$  wie oben ist bis auf die Reihenfolge der Blöcke durch  $A$  eindeutig bestimmt.

*Beweis.* Folgt direkt aus Theorem 6.4.9. □

**Bem** (zur Berechnung von Jordanschen Normalformen für Matrizen, deren charakteristisches Polynom zerfällt). Das folgende ‘Rezept’ zur Konstrukti-

on von JNF einer Matrix  $A$  ist eine Zusammenfassung der Überlegungen in diesem Abschnitt. Wir nehmen an, das charakteristische Polynom zerfällt in lineare Faktoren.

- (i) Man bestimmt die Hauptaumzerlegung, d.h. eine reguläre Matrix  $C$  mit

$$C^{-1}AC = \begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_k \end{pmatrix}$$

sodass  $A_i$  das charakteristische Polynom  $(t - \lambda_i)_i^r$  hat.

- (ii) Für jedes  $i \in \{1, \dots, k\}$  eine invertierbare Matrix  $B_i$  bestimmen, sodass für die nilpotente Matrix  $A_i - \lambda_i I$  die Matrix  $B_i^{-1}(A_i - \lambda_i I)B_i$  Jordansche Normalform hat. Dann hat auch  $J_i = B_i^{-1}A_iB_i$  Jordansche Normalform (mit Jordan-Matrizen zum Wert  $\lambda_i$ ).

(iii) Betrachte nun:

$$\begin{aligned}
 C^{-1}AC &= \begin{pmatrix} A_1 & & \\ & \ddots & \\ & & A_k \end{pmatrix} \\
 &= \begin{pmatrix} B_1 J_1 B_1^{-1} & & \\ & \ddots & \\ & & B_k J_k B_k^{-1} \end{pmatrix} \\
 &= \begin{pmatrix} B_1 & & \\ & \ddots & \\ & & B_k \end{pmatrix} \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_k \end{pmatrix} \begin{pmatrix} B_1 & & \\ & \ddots & \\ & & B_k \end{pmatrix}^{-1}
 \end{aligned}$$

Das heißt, die JNF von  $A$  ist die Matrix

$$B^{-1}AB = \begin{pmatrix} J_1 & & \\ & \ddots & \\ & & J_k \end{pmatrix},$$

wobei die Basis durch die Matrix

$$B = C \begin{pmatrix} B_1 & & \\ & \ddots & \\ & & B_k \end{pmatrix}$$

dargestellt ist.

**Bem.** Unser Weg zur Herleitung der Existenz von Jordanischen Normalformen war im Verhältnis zu anderen Überlegungen, die wir in diesem Kurs gemacht hatten, relativ lang: wir begannen mit dem Lemma von Fitting, dann kam die Hauptraumzerlegung, dann die Jordanische Normalform für nilpotente Abbildungen, und am Ende Jordanische Normalform für allgemeine Abbildungen. Die JNF ist ein Beispiel einer Aussage, an dem man sieht, dass die Behandlung mancher Fakten gewisse Zeit und Geduld erfordert.

## 7 Euklidische Räume und quadratische Formen

Wir nennen Abbildungen mit Werten aus  $\mathbb{K}$  Funktionen. In diesem Kapitel behandeln wir die folgenden besondere Klassen von Funktionen: bilineare Funktionen, Skalarprodukte, quadratische Formen. Die drei Klassen hängen untereinander zusammen und werden daher gemeinsam behandelt. Einen Vektorraum mit einem Skalarprodukt nennt man Euklidisch. Die Diskussion der Skalarprodukte nennt man somit die Theorie Euklidischer Räume.

## 7.1 Euklidische Räume

### 7.1.1 Euklidische Räume über $\mathbb{R}$ und $\mathbb{C}$ : Motivation und Definitionen

Vektorräume über Körpern  $\mathbb{R}$  und  $\mathbb{C}$  können mit einer Zusatzstruktur ausgestattet werden, welche uns erlaubt, verschiedene Begriffe Euklidischer Geometrie wie Euklidischer Abstand, Winkel, Drehung, Senkrechte Vektoren, Orthogonale Spiegelung und Orthogonale Projektion einzuführen. Vektorräume mit einer solchen Zusatzstruktur nennt man Euklidisch. Euklidischer Räume über  $\mathbb{R}$  sind anschaulicher als die Euklidische Räume über  $\mathbb{C}$ , da wir oft den physikalischen Raum der uns umgibt als einen drei-dimensionalen Euklidischen Raum über  $\mathbb{R}$  modellieren. Euklidische Räume über  $\mathbb{C}$  sind zwar etwas weniger anschaulich aber in vielen Hinsichten ähnlich zu den Euklidischen Räumen über  $\mathbb{R}$ . Im Folgenden werden wir die Euklidischen Räume über  $\mathbb{R}$  und  $\mathbb{C}$  gleichzeitig behandeln. Wir wollen den etwas tech-

nischeren Fall vom zugrundeliegenden Körper  $\mathbb{C}$  aus zwei Gründen nicht weglassen: einerseits braucht man diesen Fall in verschiedenen Anwendungen (diskrete Fourier-Transformation, Quantenphysik), anderseits braucht man manche Resultate für Euklidische Vektorräume über  $\mathbb{C}$  an einigen Stellen als Hilfsmittel zur Behandlung der Euklidischen Räume über  $\mathbb{R}$ .

Wodurch soll die Euklidische Struktur definiert werden? Wir haben Begriffe, die für uns untechnisch erscheinen und mit denen wir gut vertraut sind, wie zum Beispiel Abstand und Winkel. Es hat sich während der Entwicklung der linearen Algebra herausgestellt, dass es nicht besonders bequem ist, sich beim *Definieren* von Euklidischen Räumen an diese Begriffe anzulehnen. Es gibt aber einen auf den ersten Blick weniger intuitiven Begriff, der als Grundlage der Euklidischen Struktur sehr gut geeignet ist. Dieser Begriff ist das Skalarprodukt. Skalarprodukte haben eine natürliche Definition aus der Perspektive der linearen Algebra: das ist ihr technischer Vorteil. Des Weiteren kann man auf Skalarprodukten basierend die

restlichen relevanten Begriffe aus der Euklidischen Geometrie sehr bequem einführen.

Um Euklidische Räume über  $\mathbb{R}$  einzuführen, führen wir zuerst bilineare Formen ein. Für einen Vektorraum  $V$ , heißt eine Funktion  $f : V \times V \rightarrow \mathbb{K}$  *bilineare Form* auf  $V$ , falls  $f$  in beiden Vektorargumenten linear ist, d.h.  $f(\alpha x + \beta y, u) = \alpha f(x, u) + \beta f(y, u)$  sowie  $f(v, \alpha x + \beta y) = \alpha f(v, x) + \beta f(v, y) \quad \forall \alpha, \beta \in \mathbb{K}, \forall x, y, u, v \in V$ . Eine bilineare Form  $f$  heißt *symmetrisch*, falls  $f(u, v) = f(v, u)$  für alle  $u, v \in V$  gilt. Setzen wir in eine bilineare Form  $f(u, v)$  für  $u$  oder  $v$  den Nullvektor ein, so erhalten wir natürlich  $f(u, v) = 0$ .

Sei nun  $V$  Vektorraum über  $\mathbb{R}$ . Dann heißt eine symmetrische bilineare Form  $\langle \cdot, \cdot \rangle$  auf  $V$  *Skalarprodukt*, falls  $\langle u, u \rangle \geq 0$  und die Gleichung  $\langle u, u \rangle = 0$  nur für den Nullvektor erfüllt ist. Ein Vektorraum  $V$  über  $\mathbb{R}$ , der mit einem Skalarprodukt  $\langle \cdot, \cdot \rangle$  ausgestattet ist, heißt (*reeller*) *Euklidischer Raum*. Die Funktion  $\|u\| := \sqrt{\langle u, u \rangle}$  von  $\mathbb{R}^n$  nach  $\mathbb{R}$  heißt *Norm*. Geometrisch ist die

Norm von  $u$  die Länge des Vektors  $u$ . Der Wert  $\|u - v\|$  heißt *Abstand* zwischen den Punkten  $u$  und  $v$ .

Der Raum  $\mathbb{R}^n$ ,  $n \in \mathbb{N}$ , wird mit folgendem Standard-Skalarprodukt ausgestattet: Für  $x = (x_j)_{j=1}^n$  und  $y = (y_j)_{j=1}^n$  setzt man

$$\langle x, y \rangle := x_1 y_1 + \dots + x_n y_n.$$

Die Norm von  $x$  ist also

$$\|x\| = \sqrt{x_1^2 + \dots + x_n^2}.$$

In Vektorform ist  $\langle x, y \rangle = x^\top y$  und  $\|x\| = \sqrt{x^\top x}$  mit Spaltenvektoren  $x$  und  $y$ . Im Raum  $\mathbb{R}^n$  kann man viele verschiedene Skalarprodukte einführen. Für jede reguläre Matrix  $A \in \mathbb{R}^{n \times n}$  ist die bilineare Form  $\langle Ax, Ay \rangle$  auch ein Skalarprodukt, das eine (andere) Euklidische Struktur in  $\mathbb{R}^n$  definiert. Wenn wir aber über  $\mathbb{R}^n$  als einen Euklidischen Raum sprechen, meinen wir stillschweigend, dass  $\mathbb{R}^n$  die Euklidische Struktur durch das Standard-Skalarprodukt festgelegt ist.

Als nächste Schritt wollen wir möglichst analog Euklidische Räume über  $\mathbb{C}$  einführen. Hier eine kurze Zusammenfassung der Operationen, die wir für komplexe Zahlen benutzen werden:

**Wiederholung** (Komplexe Zahlen). Sei  $i$  die imaginäre Einheit in  $\mathbb{C}$ .

Sei  $z = a + ib \in \mathbb{C}$  mit  $a, b \in \mathbb{R}$ . Es gilt

- $z_1 = a_1 + b_1i, z_2 = a_2 + b_2i \in \mathbb{C} \Rightarrow z_1 \cdot z_2 = (a_1 + b_1i)(a_2 + b_2i) = a_1a_2 - b_1b_2 + (a_1b_1 + a_2b_1)i$
- $|z| = \sqrt{a^2 + b^2}$
- $\bar{z} = a - bi$  heißt die zu  $z = a + bi$  komplex konjugierte Zahl.

Für  $z_1, z_2 \in \mathbb{C}$  gilt  $\overline{z_1 + z_2} = \overline{z_1} + \overline{z_2}$  und  $\overline{z_1 \cdot z_2} = \overline{z_1} \cdot \overline{z_2}$ .

Skalarprodukte auf Vektorräumen über  $\mathbb{C}$  kann man *nicht* über bilineare Formen einführen. Die Funktion  $x^\top y$  ist zum Beispiel bilinear auf  $\mathbb{C}^n$ .

Darauf basierend kann man aber keine Norm einführen, denn für  $x \in \mathbb{C}^n$  ist  $x^\top x$  im Allgemeinen keine nicht-negative reelle Zahl. Wir können uns aber eine andere Funktion anschauen, und zwar Für  $x = (x_1, \dots, x_n)$  und  $y = (y_1, \dots, y_n)$  aus  $\mathbb{C}^n$ , setzen wir  $\langle x, y \rangle := x_1\overline{y_1} + \dots + x_n\overline{y_n}$ . In Vektorschreibweise ist  $\langle x, y \rangle = x^\top \bar{y}$  mit Spaltenvektoren  $x$  und  $y$ , wobei  $\bar{y} = (\overline{y_1}, \dots, \overline{y_n})$ . Dies ist das *Standard-Skalarprodukt* im Raum  $\mathbb{C}^n$ . Es ist klar, dass man für diese Funktion  $\langle x, x \rangle \geq 0$  für alle  $x \in \mathbb{C}^n$  hat. Wir werden sehen, dass das das richtige Konzept eines Skalarprodukts ist. Man beachte auch, dass die Funktion  $\langle x, y \rangle := x^\top \bar{y}$  das Standard-Skalarprodukt auf  $\mathbb{R}^n$  erweitern. Denn  $\mathbb{R}^n$  ist Teilmenge von  $\mathbb{C}^n$ : setzen wir also reelle Vektoren  $x, y \in \mathbb{R}^n$  in  $x^\top \bar{y}$  ein, so erhalten wir  $x^\top y$  (das Standard-Skalarprodukt von  $\mathbb{R}^n$ ). Diese Beobachtung zeigt auch Folgendes: hat man eine Aussage über Vektorräume über  $\mathbb{C}$  bewiesen, so ist eine entsprechende Aussage über Vektorräume über  $\mathbb{R}$  somit auch abgedeckt.

Die obigen Überlegungen für  $\mathbb{C}^n$  führen zu den folgenden Begriffen

für allgemeine Vektorräume über  $\mathbb{C}$ .

Sei  $V$  Vektorraum über  $\mathbb{C}$ . Eine Funktion  $f : V \times V \rightarrow \mathbb{C}$  heißt  $1\frac{1}{2}$ -lineare Form auf  $V$ , falls  $f$  folgende Bedingungen erfüllt:

- (i)  $f$  ist linear im ersten Vektorargument, d.h.

$$f(\alpha x + \beta y, u) = \alpha f(x, u) + \beta f(y, u) \quad \forall \alpha, \beta \in \mathbb{K}, \forall x, y, u, v \in V.$$

- (ii)  $f$  ist  $\frac{1}{2}$ -linear im zweiten Vektorargument, d.h.

$$f(v, \alpha x + \beta y) = \overline{\alpha} f(v, x) + \overline{\beta} f(v, y) \quad \forall \alpha, \beta \in \mathbb{K}, \forall x, y, u, v \in V.$$

Eine  $1\frac{1}{2}$ -lineare Form  $f$  auf  $V$  nennt man *hermitesch*, wenn  $f(u, v) = \overline{f(v, u)}$  für alle  $u, v \in V$  erfüllt ist. Insbesondere hat man für hermitische Formen  $f(u, u) \in \mathbb{R}$  für alle  $u \in V$ .

Eine hermitesch  $1\frac{1}{2}$ -lineare Form  $\langle \cdot, \cdot \rangle$  auf Vektorräumen über  $\mathbb{C}$  heißt Skalarprodukt, falls  $\langle u, u \rangle \geq 0$  für alle  $u \in V$  gilt und die Gleichung  $\langle u, u \rangle = 0$  nur für den Nullvektor erfüllt ist. Ein Vektorraum  $V$  über  $\mathbb{C}$ , der mit einem

Skalarprodukt ausgestattet ist, heißt (*komplexer*) *Euklidischer Raum*. Oft nennt man solche Räume auch *unitäre Räume*. Die Norm und der Abstand werden auf dieselbe Weise eingeführt, wie im reellen Euklidischen Raum.

Bei den folgenden Betrachtungen bedeutet die Annahme, dass  $V$  ein Euklidischer Raum über  $\mathbb{K}$  ist, dass  $\mathbb{K}$  entweder  $\mathbb{R}$  oder  $\mathbb{C}$  sein soll.

**Bem.** Man behandelt in der Schule die sogenannten binomischen Formeln:<sup>4</sup> die erste  $(a + b)^2 = a^2 + b^2$ , die zweite  $(a - b)^2 = a^2 - 2ab + b^2$  und die dritte  $(a - b)(a + b) = a^2 - b^2$ . Diese Formeln (für Werte  $a, b \in \mathbb{R}$ ) können

---

<sup>4</sup>Diese Namen sind Schulbegriffe, die man außerhalb der Schule in der Fachwelt der Mathematik eigentlich nicht benutzt.

auf Vektorräume über  $\mathbb{R}$  folgendermaßen übertragen werden. Es gilt

$$\begin{aligned}\|a + b\|^2 &= \|a\|^2 + 2 \langle a, b \rangle + \|b\|^2, \\ \|a - b\|^2 &= \|a\|^2 - 2 \langle a, b \rangle + \|b\|^2, \\ \langle a - b, a + b \rangle &= \|a\|^2 - \|b\|^2\end{aligned}$$

für beliebige Vektoren  $a, b \in V$  eines Euklidischen Raums  $V$  über  $\mathbb{R}$ . Entsprechende Formeln für Euklidische Räume über  $\mathbb{C}$  kann man auch herleiten (die linken Seiten bleiben gleich, die rechten Seiten sehen aber etwas technischer aus).

### 7.1.2 Die Ungleichung von Cauchy-Schwarz und der Winkel zwischen Vektoren

Die Ungleichung von Cauchy-Schwarz gehört zu den bekanntesten und nützlichsten Ungleichungen in Mathematik.

**Thm.** Sei  $V$  Euklidischer Raum über  $\mathbb{K}$ . Dann gilt:

$$|\langle u, v \rangle| \leq \|u\| \cdot \|v\| \quad \forall u, v \in V, \quad (7.1.1)$$

wobei Gleichheit genau dann erfüllt ist, wenn  $u$  und  $v$  linear abhängig sind.

*Beweis.* Hier nur für  $\mathbb{K} = \mathbb{R}$ . ( $\mathbb{K} = \mathbb{C}$  Aufgabe)

Seien  $u, v \in V \setminus \{0\}$ , da sonst die Behauptung trivial ist. Für jedes  $\lambda \in \mathbb{R}$  gilt

$$0 \leq \|u - \lambda v\|^2 = \|u\|^2 - 2\lambda \langle u, v \rangle + \lambda^2 \|v\|^2$$

Wir haben also die Ungleichung

$$\|u\|^2 - 2\lambda \langle u, v \rangle + \lambda^2 \|v\|^2 \geq 0$$

erhalten, die von einem Parameter  $\lambda \in \mathbb{R}$  abhängig ist. Wir wollen nun die Wahl von  $\lambda$  fixieren, die uns die stärkste Ungleichung gibt. Dafür soll

die linke Seite minimiert werden. Um das beste  $\lambda$  zu fixieren, kann man also z.B. Methoden aus der Analysis (Ableitungen usw.) benutzen. Da die linke Seite lediglich eine quadratische Funktion in  $\lambda$  ist, kommt auch ohne Analysis aus. Wir teilen die Gleichung durch  $\|v\|^2$  und erhalten

$$\frac{\|u\|^2}{\|v\|^2} - 2\lambda \frac{\langle u, v \rangle}{\|v\|} + \lambda^2 \geq 0$$

Dann machen wir die quadratische Ergänzung und erhalten.

$$\frac{\|u\|^2}{\|v\|^2} - \frac{\langle u, v \rangle^2}{\|v\|^4} + \left( \lambda - \frac{\langle u, v \rangle}{\|v\|^2} \right)^2 \geq 0$$

Diese Ungleichung wird am schärfsten, wenn der Klammerausdruck 0 ist. Durch Einsetzen von  $\lambda = \frac{\langle u, v \rangle}{\|v\|^2}$  erhalten wir also

$$\|u\|^2 \|v\|^2 \geq \langle u, v \rangle^2,$$

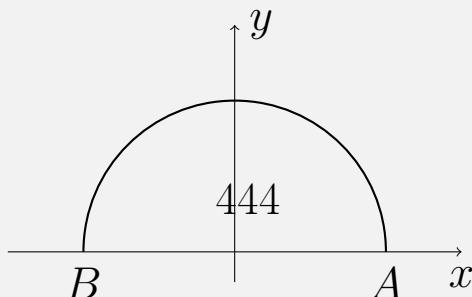
was der Ungleichung in der Behauptung äquivalent ist.

Wir müssen noch den Gleichheitsfall charakterisieren. Lineare Abhängigkeit von  $u$  und  $v$  ist für die Gleichheit offensichtlich hinreichend. Umgekehrt: gilt die Gleichheit  $\|u\|\|v\| = |\langle u \rangle v|$  für  $u, v \in V$  so hat man nach der vorigen Herleitung  $\|u - \lambda v\| = 0$  für  $\lambda = \frac{\langle u, v \rangle}{\|v\|^2}$ . Aus  $\|u - \lambda v\| = 0$  folgt  $u - \lambda v = 0$ , das heißt:  $u$  und  $v$  sind linear abhängig.  $\square$

Sei  $V$  ein reeller Euklidischer Raum und  $u, v \in V \setminus \{0\}$ . Nach der Ungleichung von Cauchy-Schwarz ist  $\left\langle \frac{u}{\|u\|}, \frac{v}{\|v\|} \right\rangle = \frac{\frac{u}{\|u\|} \cdot \frac{v}{\|v\|}}{\|u\|\|v\|}$  ein Wert in  $[-1, 1]$ . Ein Wert aus  $[-1, 1]$  kann eindeutig als  $\cos \phi$  mit  $\phi \in [0, \pi]$  dargestellt werden. Der Wert  $\phi$  heißt der Winkel zwischen  $u$  und  $v$ .

Das heißt, wir definieren den *Winkel*  $\phi$  zwischen  $u$  und  $v$  als einen eindeutigen Wert  $\phi \in [0, \pi]$  mit  $\cos \phi = \frac{\langle u, v \rangle}{\|u\|\|v\|}$ .

**Wiederholung** (Kosinus). Falls man aus der Schule die geometrische Bedeutung von Kosinus nicht weiß, hier eine kurze Erklärung. Wir betrachten die obere Hälfte des Kreises mit Radius eins und Zentrum in  $(0, 0)$ . Dieser Bogen hat die Länge  $\pi = 3,1415926\ldots$ . Wir laufen vom rechten Endpunkt  $A$  aus entlang des Bogens. Nach  $\pi$  Einheiten kommen wir zum linken Endpunkt  $B$ . Nach  $\pi/2$  Einheiten stehen wir ganz oben auf dem Bogen. Wenn wir  $\phi \in [0, \pi]$  fixieren und genau  $\phi$  Einheiten zurücklegen so landen wir irgendwo auf dem Bogen. Dann schauen wir uns die  $x$ -Koordinate an. Ist  $x = 1$  so haben wir uns gar nicht bewegt und unser  $\phi$  ist  $0$ . Ist  $x = -1$  so haben wir den gesamten Bogen zurückgelegt und unser  $\phi$  ist  $\pi$ . Im Allgemeinen ist unser  $x$  ein Wert zwischen  $-1$  und  $1$ . Den Wert  $x$  in Abhängigkeit von  $\phi$  nennt man Kosinus  $\phi$ . Wie wir oben besprochen haben gilt  $\cos(0) = 1$ ,  $\cos(\pi/2) = 0$  und  $\cos(\pi) = -1$ .



Bei Berechnungen mit Kosinus (und Sinus) mit Taschenrechner und

### 7.1.3 Eigenschaften der Norm und des Abstands

**Prop.** Sei  $V$  Euklidischer Raum über  $\mathbb{K}$ . Dann gilt:

$$(i) \ \|u\| > 0 \text{ und } \|u\| = 0 \Leftrightarrow u = 0 \quad \forall u \in V$$

$$(ii) \ \|\alpha u\| = |\alpha| \|u\| \quad \forall \alpha \in \mathbb{K}, \forall u \in V$$

$$(iii) \ \|u + v\| \leq \|u\| + \|v\| \quad \forall u, v \in V$$

*Beweis.* Übung. □

Eine Funktion, welche die drei Bedingungen der vorigen Proposition erfüllt, nennt man eine *Norm* auf  $V$ .

Die Ungleichung (iii) wird Dreiecksungleichung genannt. Wenn man über die Abstände zwischen Punkten  $a, b, c \in V$  redet, kann man diese umschreiben:  $\|a - c\| \leq \|a - b\| + \|b - c\|$ . Auch der Gleichheitsfall lässt sich analog beschreiben. (Übung)

Auf einem Vektorraum  $V$  kann man verschiedene Normen definieren, nicht nur die Euklidischen. Eine Euklidische Norm kommt aber stets mit einem Skalarprodukt, und dieses kann aus der Norm abgelesen werden, wie die folgende Proposition zeigt.

**Prop.** *Sei  $V$  Euklidischer Raum über  $\mathbb{R}$ . Dann gilt:*

$$\langle u, v \rangle = \frac{1}{4} (\|u + v\|^2 - \|u - v\|^2) \quad \forall u, v \in V \quad (7.1.2)$$

*Beweis.* Aufgabe. □

Das Skalarprodukt wird somit eindeutig durch die Euklidische Norm bestimmt, wie man in der vorigen Proposition im Fall des Körpers der reellen Zahlen zeigt. Eine analoge Aussage für die komplexen Zahlen ist wie folgt:

**Prop.** *Sei  $V$  Euklidischer Raum über  $\mathbb{C}$  und  $i$  die imaginäre Einheit. Dann*

gilt:

$$\langle u, v \rangle = \frac{1}{4} (\|u + v\|^2 - \|u - v\|^2 + i\|u + iv\|^2 - i\|u - iv\|^2) \quad \forall u, v \in V \quad (7.1.3)$$

*Beweis.* Aufgabe. □

#### 7.1.4 Orthogonale bzw. Orthonormale Systeme und das Gram-Schmidt-Verfahren

Vektoren  $u$  und  $v$  eines Euklidischen Raumes  $V$  heißen *orthogonal* (senkrecht), falls  $\langle u, v \rangle = 0$  gilt. Ein Vektorsystem in  $V$  heißt orthogonal, falls die Vektoren dieses Systems ungleich 0 und paarweise orthogonal sind. Jedes orthogonale System ist linear unabhängig (Aufgabe). Ein orthogonales System Vektorsystem heißt *orthonormal*, falls die Norm der Vektoren dieses Systems gleich 1 ist. Die Standard-Basis  $e_1, \dots, e_n$  von  $\mathbb{R}^n$  ( $n \in \mathbb{N}$ ) ist eine Orthonormalbasis (d.h. Basis und orthonormales System).

Hat man in einem  $n$ -dimensionalen Euklidischen Raum  $V$  ( $n \in \mathbb{N}$ ) eine Orthonormalbasis  $\mathcal{B} = (b_1, \dots, b_n)$ , so lässt sich für jeden Vektor  $u \in V$  die Darstellung von  $u$  in  $\mathcal{B}$  durch die Berechnung von  $n$  Skalarprodukten bestimmen:

$$u_{\mathcal{B}} = \begin{pmatrix} \langle u, b_1 \rangle \\ \vdots \\ \langle u, b_n \rangle \end{pmatrix}.$$

Das ist sehr praktisch. Wenn man etwa im Raum  $\mathbb{K}^n$  einen Vektor in einer beliebigen Basis darstellen möchte, muss man ein lineares Gleichungssystem lösen (der Aufwand ist kubisch in  $n$ , Berechnung von  $n$  Skalarprodukten hat dagegen einen quadratischen Aufwand). Die Berechnungen von Skalarprodukten und Normen können in einer gegebenen Basis durchgeführt werden. Für  $u, v \in V$  gilt  $\langle u, v \rangle = \langle u_{\mathcal{B}}, v_{\mathcal{B}} \rangle = \sum_{k=1}^n \langle u, b_k \rangle \overline{\langle v, b_k \rangle}$ . Daraus folgt auch  $\|x\| = \|x_{\mathcal{B}}\|$ . Somit können die Berechnungen im abstrakten Euklidischen Raum  $V$  durch Berechnungen im "konkreten" Euklidischen Raum  $\mathbb{K}^n$  ersetzt

werden.

Die Orthonormalbasen sind sehr nützlich. Wenn ein Euklidischer Vektorraum durch eine beliebige Basis gegeben ist, so kann man darauf basierend stets eine Orthonormalbasis ausrechnen.

**Thm** (Existenz einer Orthonormalbasis). *Sei  $V$   $n$ -dimensionaler Euklidischer Raum über  $\mathbb{K}$  ( $n \in \mathbb{N}$ ). Dann besitzt  $V$  eine Orthonormalbasis.*

*Beweis.* Es genügt, die Existenz einer Orthogonalbasis zu zeigen (jede Orthogonalbasis kann durch Normierung in eine Orthonormalbasis überführt werden). Wir fixieren eine beliebige Basis  $a_1, \dots, a_n$  von  $V$  und konstruieren anhand dieser Basis eine Orthogonalbasis  $b_1, \dots, b_n$ .

Wir gehen iterativ vor. In der  $i$ -ten Iteration hat man bereits eine Orthogonalbasis  $b_1, \dots, b_{i-1}$  von  $\text{lin}(a_1, \dots, a_{i-1})$  zur Verfügung, und erweitert diese Basis zu einer Orthogonalbasis  $b_1, \dots, b_i$  von  $\text{lin}(a_1, \dots, a_i)$ .

In der ersten Iteration wird  $b_1 = a_1$  gesetzt. In der  $i$ -ten Iteration ( $i \geq 2$ ) wählen fixieren wir  $b_i$  mit Hilfe der Formel  $b_i = a_i - \sum_{k=1}^{i-1} \gamma_k b_k$ , bei der

die Koeffizienten  $\gamma_1, \dots, \gamma_{i-1} \in \mathbb{K}$  festgelegt werden müssen. Die Intuition hinter der Formel ist wie folgt: Wir lassen von  $a_i$  ausgehen, den Vektor  $b_i$  parallel zu dem Raum  $\text{lin}(a_1, \dots, a_{i-1})$  “gleiten”, bis er orthogonal zu  $\text{lin}(a_1, \dots, a_{i-1})$  wird.

Die Werte  $\gamma_1, \dots, \gamma_{i-1}$  werden aus den Bedingungen  $\langle b_i, b_j \rangle = 0$  mit  $j \in \{1, \dots, i-1\}$  bestimmt. Da  $b_1, \dots, b_{i-1}$  ein Orthogonalsystem ist, erhält man aus den vorigen Bedingungen die Darstellung  $\gamma_j = \langle a_i, b_j \rangle / \langle b_j, b_j \rangle \quad \forall j \in \{1, \dots, i-1\}$ . Der so gewählte Vektor  $b_i$  ist ungleich 0. Denn sonst wäre  $a_i \in \text{lin}(b_1, \dots, b_{i-1}) = \text{lin}(a_1, \dots, a_{i-1})$ , was der linearen Unabhängigkeit von  $a_1, \dots, a_n$  widerspricht. Nach  $n$  Iterationen hat man die gewünschte Basis konstruiert.  $\square$

**Bem.** Das Verfahren aus dem vorigen Beweis heißt das Gram-Schmidt-Verfahren. Das Verfahren kann auch in der Sprache der Matrizen formuliert werden. Dann nennt man es die  $QR$ -Zerlegung.

**Kor** (Ergänzung von orthogonalen bzw. orthonormalen Systemen). *Sei  $b_1, \dots, b_m$*

$(m \in \mathbb{N}_0)$  ein orthogonales bzw. orthonormales System in einem  $n$ -dimensionalen Euklidischen Raum  $V$  ( $n \in \mathbb{N}$ ). Dann ist  $m \leq n$  und  $b_1, \dots, b_m$  kann zu einer Orthogonal- bzw. Orthonormalbasis von  $V$  erweitert werden.

*Beweis.* Erweitere  $b_1, \dots, b_m$  zu einer beliebigen Basis von  $V$  und orthogonalisiere diese dann diese anschließend mit dem Gram-Schmidt.  $\square$

### 7.1.5 Orthogonale Projektion

Orthogonale Projektion auf einen Untervektorraum ist eine Operation, die man oft in der Praxis benötigt. Lineare Regression in der Statistik sowie die Methode der kleinsten Quadrate sind z.B. nichts anderes als die Umsetzung der orthogonalen Projektion im Kontext der Statistik bzw. Numerik.

Durch das folgende Theorem wird die orthogonale Projektion eingeführt.

**Thm.** Sei  $V$  ein  $n$ -dimensionaler Euklidischer Raum über  $\mathbb{K}$  und  $U$  ein Untervektorraum von  $V$ . Dann gilt:

- (i) Zu jedem  $x \in V$  existiert ein eindeutiges  $y \in U$  derart, unter allen Vektoren in  $U$  der Vektor  $y$  bzgl. der Euklidischen Norm zu  $x$  am nächsten liegt. (D.h.  $\|x - y\| \leq \|x - u\| \quad \forall u \in U$ ) gilt.
- (ii) Für jedes  $x \in V$  ist  $y$  aus (i) der eindeutige Vektor  $y$  mit der Eigenschaft, dass  $x - y$  zu jedem Vektor aus  $U$  orthogonal ist.
- (iii) Die Abbildung  $x \mapsto y$  ist linear.

*Beweis.* Sei  $u_1, \dots, u_m$  eine Orthonormal-Basis von  $U$ . Wir zeigen, dass für den Vektor

$$y = \langle x, u_1 \rangle u_1 + \dots + \langle x, u_m \rangle u_m \quad (7.1.4)$$

die Behauptungen des Theorems erfüllt sind. Wir beginnen mit (ii). Wenn  $y - x$  zu jedem Basisvektor  $u_i$  von  $U$  orthogonal ist, dann ist  $y - x$  zu allen Vektoren aus  $U$  orthogonal. Für  $y$ , das durch die vorige Formel gegeben ist, hat man tatsächlich  $\langle y - x, u_i \rangle$  für alle  $i = 1, \dots, m$ . Umgekehrt, wenn  $y$  ein Vektor aus  $U$  ist, für welchen  $\langle y - x \rangle u$  für alle  $u \in U$  gilt, dann gilt

$\langle y - x \rangle u_i$  für alle  $i = 1, \dots, m$ . Wenn wir nun  $y$  als Linearkombination  $y = \langle y \rangle u_1 u_1 + \dots + \langle y, u_m \rangle u_m$  der Basisvektoren darstellen, so erhalten wir aus den Gleichungen  $\langle y - x \rangle u_i = 0$ , dass  $\langle y \rangle u_i = \langle x \rangle u_i$ .

Um nun (i) herzuleiten, müssen wir überprüfen, dass der von uns gewählte Vektor  $y$  ein eindeutiger Vektor ist, für welchen die Eigenschaft in (i) erfüllt ist. In (i) arbeiten wir mit Abständen. Wir sind auf der Such nach einem Vektor in  $U$ , der zu  $x$  am nächsten Ist. Wir minimieren also  $\|x - u\|$  für  $u \in U$ . Es ist besser das Quadrat der Norm zu minimieren (denn die Norm hatte eine Wurzel in der Definition, die wir gerne vermeiden wollen). Das heißt, wir minimieren  $\|x - u\|^2$  für  $u \in U$ . Nun benutzen wir das  $y$  aus (7.1.4) und formen den Ausdruck  $\|x - u\|^2$  folgendermaßen um:

$$\|x - u\|^2 = \|(x - y) + (y - u)\|^2 = \|x - y\|^2 + \langle x - y \rangle y - u + \langle y - u \rangle x - y + \|y - u\|^2.$$

Der Vektor  $y - u$  liegt in  $U$  und, wie wir in (ii) gesehen haben, ist  $x - y$

orthogonal zu jedem Vektor aus  $U$ . Also hat man

$$\|x - u\|^2 = \|x - y\|^2 + \|y - u\|^2.$$

Der quadrierte Abstand von  $x$  und  $u$  ergibt sich somit als der quadrierte Abstand von  $x$  und  $y$  plus der Zusatz  $\|y - u\|^2$ , der von  $u$  abhängig ist. Der Zusatz ist genau dann am kleinsten wenn  $y = u$  gleich ist. Dies zeigt (i).

(iii) ist das Nebenprodukt des Beweises von (i) und (ii): die rechte Seite von (7.1.4) ist linear in  $x$ . □

### 7.1.6 Orthogonale Untervektorräume, Summen und Orthogonalkomplement

Im Teil 1 des Kurses haben wir Zerlegungen von Vektorräumen in direkte Summanden eingeführt. Nun führen wir Zerlegungen von Euklidischen Räumen in orthogonale direkte Summanden ein.

Sei  $V$  Euklidischer Raum über  $\mathbb{K}$  und  $X$  Teilmenge von  $V$  ( $X$  muss kein

Untervektorraum sein.) Dann heißt  $X^\perp := \{v \in V : \langle v, x \rangle = 0 \ \forall x \in X\}$  das *Orthogonalkomplement* von  $X$  in  $V$ . Mit anderen Worten ist  $X^\perp$  die Menge der Vektoren aus  $V$  die zu allen Vektoren in  $X$  orthogonal sind.  $X^\perp$  ist ein Untervektorraum von  $V$ . (Aufgabe)

Zwei Untervektorräume  $U, W$  von  $V$  heißen *orthogonal*, wenn  $\langle u, w \rangle = 0 \ \forall u \in U, w \in W$ . Schreibweise:  $U \perp W$ . Ein Vektor  $v \in V$  und ein Untervektorraum  $U$  von  $V$  heißen orthogonal, falls  $\langle v, u \rangle = 0 \ \forall u \in U$ . Schreibweise:  $v \perp U$ .

Die Summe von Untervektorräumen  $U_1, \dots, U_k$  von  $V$  heißt orthogonal, falls  $U_1, \dots, U_k$  paarweise orthogonal sind. Schreibweise:  $U_1 \oplus \dots \oplus U_k$ . Jede orthogonale Summe ist direkt. (Aufgabe)

Wenn wir nun einen Untervektorraum  $U$  von  $V$  fixieren, können wir  $V$  in die orthogonale direkte Summe von  $U$  und  $U^\perp$  zerlegen:

**Thm.** *Sei  $V$  ein  $n$ -dimensionaler Euklidischer Raum über  $\mathbb{K}$  und  $U$  ein Untervektorraum von  $V$ . Dann gilt  $V = U \oplus U^\perp$  und  $(U^\perp)^\perp = U$ .*

*Beweis.* Es ist  $U \perp U^\perp$  laut Definition von  $(\cdot)^\perp$ . Betrachte eine Orthonormalbasis  $u_1, \dots, u_m$  von  $U$  und ergänze diese zu einer Orthonormalbasis  $u_1, \dots, u_n$  von  $V$ . Man kann zeigen, dass  $U^\perp = \text{lin}(u_{m+1}, \dots, u_n)$  gilt. (Aufgabe, einfach). D.h.  $(\text{lin}(u_1, \dots, u_m))^\perp = \text{lin}(u_{m+1}, \dots, u_n)$ . Daraus folgt:

$$\begin{aligned} ((\text{lin}(u_1, \dots, u_m))^\perp)^\perp &= (\text{lin}(u_{m+1}, \dots, u_n))^\perp \\ &= \text{lin}(u_1, \dots, u_m), \end{aligned}$$

d.h.  $(U^\perp)^\perp = U$ .

□

## 7.2 Lineare Abbildungen Euklidischer Räume

Manche Abbildungen eines Euklidischen Raums haben ganz besondere Eigenschaften bzgl. der Euklidischen Struktur. Die Bewegungen (unter anderem Drehungen und orthogonale Spiegelungen) sind zum Beispiel sehr interessant: diese Abbildungen ändern die Längen der Vektoren nicht. Außerdem

gibt es Abbildungen, die einer *Orthogonalbasis* durch eine Diagonalmatrix darstellbar sind.

### 7.2.1 Adjungierte Abbildung

Die adjungierte Abbildung ist ein nützliches technisches Werkzeug, mit dem man die oben erwähnten Klassen spezieller linearer Abbildungen schön beschreiben kann. In dem folgenden Theorem wird die adjungierte Abbildung durch die Gleichung  $\langle F(u), v \rangle = \langle u, F^*(v) \rangle$  charakterisiert:

**Thm** (Existenz und Eindeutigkeit der adjungierten Abbildung). *Sei  $V$  ein  $n$ -dimensionaler Euklidischer Raum über  $\mathbb{K}$ . Zu jedem  $F \in \text{Lin}(V)$  existiert eine eindeutige Abbildung  $F^* \in \text{Lin}(V)$  mit*

$$\langle F(u), v \rangle = \langle u, F^*(v) \rangle \quad \forall u, v \in V \tag{7.2.1}$$

*Für die Operation  $F \mapsto F^*$  auf  $\text{Lin}(V)$  gilt für alle  $F, G \in \text{Lin}(G)$  und  $\alpha, \beta \in \mathbb{K}$  Folgendes:*

$$(i) \quad (F^*)^* = F$$

$$(ii) \quad (\alpha F + \beta G)^* = \overline{\alpha} F^* + \overline{\beta} G^*$$

$$(iii) \quad (F \circ G)^* = G^* \circ F^*$$

$$(iv) \quad \text{id}^* = \text{id}$$

(v) Ist  $F$  invertierbar, so ist auch  $F^*$  invertierbar, und es gilt  $(F^*)^{-1} = (F^{-1})^*$ .

Die Abbildung  $F^*$  heißt die adjungierte Abbildung von  $F$ .

*Beweis.* Sei  $b_1, \dots, b_n$  eine Orthonormalbasis von  $V$ . Wenn  $F^* \in \text{Lin}(V)$  eine Abbildung ist, die (7.2.1) erfüllt, dann ist  $\langle F^*(x), b_i \rangle = \overline{\langle b_i, F^*(x) \rangle} = \overline{\langle F(b_i), x \rangle} = \langle x, F(b_i) \rangle$  und es gilt:

$$F^*(x) = \sum_{i=1}^n \langle F^*(x), b_i \rangle b_i = \sum_{i=1}^n \langle x, F(b_i) \rangle b_i,$$

Dies zeigt die Eindeutigkeit: wenn ein  $F^*$  mit (7.2.1) existiert, dann ist das  $F^*$  in unserer festen Orthonormalbasis eindeutig wie oben beschrieben.

Es bleibt zu zeigen, dass das so definierte  $F^*$  die Bedingung (7.2.1) erfüllt (Existenz). Wir setzen unsere Darstellung von  $F^*$  in die rechte Seite von

(7.2.1) und transformieren dann die rechte Seite zur linken Seite wie folgt:

$$\begin{aligned}
 \langle u, F^*(v) \rangle &= \left\langle u, \sum_{i=1}^n \langle v, F(b_i) \rangle b_i \right\rangle && (\text{Einsetzen des Ausdrucks f\"ur } F^*(v)) \\
 &= \sum_{i=1}^n \overline{\langle v, F(b_i) \rangle} \langle u, b_i \rangle && (\text{nach } \frac{1}{2}\text{-Linearit\"at im zweiten Vektorargument}) \\
 &= \sum_{i=1}^n \langle F(b_i), v \rangle \langle u, b_i \rangle \\
 &= \left\langle \sum_{i=1}^n \langle u, b_i \rangle F(b_i), v \right\rangle && (\text{nach Linearit\"at im ersten Vektorargument}) \\
 &= \left\langle F \left( \sum_{i=1}^n \langle u, b_i \rangle b_i \right), v \right\rangle && (\text{nach der Linearit\"at von } F) \\
 &= \langle F(u), v \rangle.
 \end{aligned}$$

(i), (ii) und (iv) sind Aufgaben. F\"ur (iii) betrachte  $\langle (F \circ G)(u), v \rangle = \langle F(G(u)), v \rangle = \langle G(u), F^*(v) \rangle = \langle u, G^*(F^*(v)) \rangle = \langle u, (G^* \circ F^*)(v) \rangle$  und

für (v) ist  $(F^{-1})^* \circ F^* = (F \circ F^{-1})^* = \text{id} \Rightarrow F^*$  ist invertierbar und  $(F^*)^{-1} = (F^{-1})^*$ .  $\square$

Die folgenden Proposition präsentiert Dualitätsrelationen zwischen dem Kern und dem Bild der Abbildungen  $F$  und  $F^*$ .

**Prop.** *Sei  $V$  endlich-dimensionaler Euklidischer Raum über  $\mathbb{K}$ . Sei  $F \in \text{Lin}(V)$ . Dann gilt:*

$$\begin{aligned}\ker(F^*) &= \text{im}(F)^\perp, \\ \text{im}(F^*) &= \ker(F)^\perp.\end{aligned}$$

*Beweis.* Es ist

$$\begin{aligned}\text{im}(F)^\perp &= \{u \in V : \langle u, y \rangle = 0 \quad \forall y \in \text{im}(F)\} \\ &= \{u \in V : \langle u, F(v) \rangle \quad \forall v \in V\} \\ &= \{u \in V : \langle F^*(u), v \rangle \quad \forall v \in V\} \\ &= \{u \in V : F^*(u) = 0\} = \ker(F^*).\end{aligned}$$

Die zweite Gleichung kann man analog beweisen, oder alternativ aus der ersten folgern (mittels (i) aus dem vorigen Theorem und Theorem 7.1.6).

□

**Bem.**

- Sei  $A \in \mathbb{C}^{n \times n}$  ( $n \in \mathbb{N}$ ). Die adjungierte Abbildung zu  $x \mapsto Ax$  ist die Abbildung  $x \mapsto A^*x$  mit  $A^* = \overline{A}^\top$ . (hier bezeichnet  $\overline{A} \in \mathbb{C}^{n \times n}$  die Matrix, welche durch die Konjugation der Komponenten von  $A$  entsteht).
- Sei  $A \in \mathbb{R}^{n \times n}$  ( $n \in \mathbb{N}$ ). Die adjungierte Abbildung zu  $x \mapsto Ax$  (auf  $\mathbb{R}^n$ ) ist die Abbildung  $x \mapsto A^\top x$ . Bzgl. des Körpers  $\mathbb{R}$  benutzen wir als  $A^*$  als eine andere Bezeichnung für  $A^\top$ .

### 7.2.2 Lineare Isometrien von Euklidischen Räumen

In diesem Paragraphen geht es um die linearen Abbildungen, welche die Längen der Vektoren nicht verändert: solche Abbildungen nennt man die linearen Bewegungen oder die linearen Isometrien (Isometrie heißt Abstand erhalten).

Man kann die linearen Isometrien mit Hilfe der Norm definieren und auf zwei Weisen äquivalent beschreiben (mit Skalarprodukten und mit Hilfe der adjungierten Abbildung).

**Thm** (Charakterisierung von linearen Isometrien). *Sei  $V$  ein endlich-dimensionaler Euklidischer Raum über  $\mathbb{K}$ . Sei  $F \in \text{Lin}(V)$ . Dann sind die folgenden Bedingungen äquivalent:*

- (i)  $\|F(u)\| = \|u\| \quad \forall u \in V$
- (ii)  $\langle F(u), F(v) \rangle = \langle u, v \rangle \quad \forall u, v \in V$

(iii)  $F$  ist invertierbar mit  $F^{-1} = F^*$ .

*Beweis.*

(i) $\Rightarrow$ (ii): folgt aus Proposition 7.1.3 von Seite 445 (das Skalarprodukt kann mit Hilfe der Normen gewisser Linearkombination von  $u$  und  $v$  beschrieben werden, das ergibt dass die linke und die rechte Seite der Gleichung in (ii) identisch sind).

(ii) $\Rightarrow$ (iii):  $\langle F(u), F(v) \rangle = \langle u, v \rangle \Rightarrow \langle (F^* \circ F)(u), v \rangle = \langle u, v \rangle = \langle u, \text{id}(v) \rangle$ ,  
d.h.  $F^* \circ F = \text{id}^* = \text{id}$  und  $F^{-1} = F^*$ .

(iii) $\Rightarrow$ (i):  $\|F(u)\|^2 = \langle F(u), F(u) \rangle = \langle u, (F^* \circ F)(u) \rangle = \langle u, \text{id}(u) \rangle = \langle u, u \rangle = \|u\|^2$ .  $\square$

Lineare Abbildungen, die die Bedingungen (i) – (iii) wie oben erfüllen, heißen *lineare Isometrien*. Lineare Isometrien reeller Euklidischer Räume heißen *orthogonale Abbildungen*. Lineare Isometrien komplexer Euklidischer Räume heißen *unitäre Abbildungen*.

Wenn für  $A \in \mathbb{C}^{n \times n}$  ( $n \in \mathbb{N}$ ) die Abbildung  $x \mapsto Ax$  auf  $\mathbb{C}^n$  unitär ist, dann nennt man die Matrix  $A$  *unitär*. Wenn für  $A \in \mathbb{R}^{n \times n}$  ( $n \in \mathbb{N}$ ) die Abbildung  $x \mapsto Ax$  auf  $\mathbb{R}^n$  orthogonal ist, dann nennt man die Matrix  $A$  *orthogonal*.

Die folgende Proposition charakterisiert die unitären Matrizen auf mehrere Weisen:

**Prop.** *Sei  $A \in \mathbb{C}^{n \times n}$ . Dann sind die folgenden Bedingungen äquivalent:*

- (i)  *$A$  ist unitär.*
- (ii)  *$AA^* = I$ .*
- (iii) *Die Zeilen von  $A$  bilden eine Orthonormalbasis von  $\mathbb{C}^n$ .*
- (iv) *Die Spalten von  $A$  bilden eine Orthonormalbasis von  $\mathbb{C}^n$ .*

*Beweis.* Aufgabe. □

Eine entsprechende Proposition kann auch für orthogonale Matrizen formuliert werden:

**Prop.** *Sei  $A \in \mathbb{R}^{n \times n}$ . Dann sind die folgenden Bedingungen äquivalent:*

- (i)  *$A$  ist orthogonal.*
- (ii)  *$AA^\top = I$ .*
- (iii) *Die Zeilen von  $A$  bilden eine Orthonormalbasis von  $\mathbb{R}^n$ .*
- (iv) *Die Spalten von  $A$  bilden eine Orthonormalbasis von  $\mathbb{R}^n$ .*

*Beweis.* Aufgabe. □

### 7.2.3 Diagonalisierung linearer Isometrien

Es stellt sich heraus, dass alle linearen Isometrien über  $\mathbb{C}$  diagonalisierbar sind. Es gilt sogar eine stärkere Eigenschaft: jede lineare Isometrie ist

in einer passend gewählten Orthonormalbasis durch eine Diagonalmatrix dargestellt.

**Bem.** Sei  $F \in \text{Lin}(V)$  eine lineare Isometrie eines endlich-dimensionalen Euklidischen Raumes  $V$  über  $\mathbb{K}$ . Sei  $\lambda \in \mathbb{K}$  Eigenwert von  $F$ . Dann gilt  $|\lambda| = 1$  (Aufgabe).

Ist  $u$  Eigenvektor einer linearen Isometrie  $F$ , dann kann man zeigen, dass  $F$  als eine lineare Isometrie innerhalb des Orthogonalkomplements von  $u$  aufgefasst werden kann:

**Prop.** *Sei  $F \in \text{Lin}(V)$  eine lineare Isometrie eines endlich-dimensionalen Euklidischen Raumes  $V$  über  $\mathbb{K}$ . Sei  $u \in V \setminus \{0\}$  Eigenvektor von  $F$  und man betrachte das Orthogonalkomplement zum Eigenvektor  $u$*

$$u^\perp := \{v \in V : \langle v, u \rangle = 0\}.$$

Dann gilt  $F(u^\perp) \subseteq u^\perp$ .

*Beweis.* Sei  $y \in F(u^\perp)$ , d.h.  $y = F(x)$  mit einem  $x \in V$  derart, dass  $\langle x, u \rangle = 0$  gilt. Für  $y$  wie oben gilt  $\langle y, u \rangle = \langle F(x), u \rangle = \langle x, F^*(u) \rangle = \langle x, F^{-1}(u) \rangle$ . Es ist  $F(u) = \lambda u$  mit  $\lambda \in \mathbb{K}$ ,  $|\lambda| = 1$ , d.h.  $F^{-1}(u) = \frac{1}{\lambda}u$ . Somit gilt  $\langle y, u \rangle = \langle x, \frac{1}{\lambda}u \rangle = \frac{1}{\lambda} \langle x, u \rangle = 0$ , d.h.  $y \in u^\perp$ .  $\square$

**Thm** (Spektralsatz für lineare Isometrie eines komplexen Euklidischen Raums).  
*Sei  $V$  ein  $n$ -dimensionaler Euklidischer Raum über  $\mathbb{C}$ . Sei  $F \in \text{Lin}(V)$  eine Isometrie. Dann existiert eine Orthonormalbasis  $\mathcal{B}$  von  $V$ , für welche die Matrix  $F_{\mathcal{B}}$  diagonal ist und die Diagonaleinträge von  $F_{\mathcal{B}}$  den Betrag 1 haben.*

*Beweis.* Das charakteristische Polynom von  $F$  ist ein Polynom vom Grad  $n \geq 1$  aus  $\mathbb{C}[t]$ . Ein solches Polynom hat mindestens eine Nullstelle (vgl. Kapitel 2). D.h.  $F$  besitzt einen Eigenwert  $\lambda \in \mathbb{C}$ . Sei  $u \in V \setminus \{0\}$  Eigenvektor zu  $\lambda$  mit  $\|u\| = 1$ . Nach der vorigen Proposition gilt  $F(u^\perp) \subseteq u^\perp$ . Somit ist  $F|_{u^\perp} \in \text{Lin}(u^\perp)$  wohldefiniert. Da  $F$  eine lineare Isometrie ist, ist auch  $F|_{u^\perp}$  eine lineare Isometrie.

Der Raum  $V$  kann als  $V = \text{lin}(u) \oplus u^\perp$  dargestellt werden (vgl. 7.1.6), wobei  $\dim(\text{lin}(u)) = 1$  und  $\dim(u^\perp) = n - 1$  gilt. Nun kann die Basis  $\mathcal{B} = (b_1, \dots, b_n)$  rekursiv konstruiert werden: Man wählt  $b_1 = u$  und als  $(b_2, \dots, b_n)$  wählt man eine Orthonormalbasis von  $u^\perp$  aus Eigenvektoren von  $F|_{u^\perp}$ . Da jeder Eigenwert von  $F$  den Betrag 1 hat, haben die Diagonaleinträge von  $F_{\mathcal{B}}$  ebenfalls den Betrag 1.  $\square$

**Bsp.** Aufgabe: Diagonalisieren Sie die Matrix

$$\begin{pmatrix} \cos \phi & -\sin \phi \\ \sin \phi & \cos \phi \end{pmatrix}$$

mit  $\phi \in \mathbb{R}$  bzgl.  $\mathbb{C}$  mit Hilfe einer Orthonormalbasis. Diese Matrix ist im Fall  $\phi \in \mathbb{R} \setminus \pi\mathbb{Z}$  bzgl.  $\mathbb{R}$  nicht diagonalisierbar!

### 7.2.4 Selbstadjungierte Abbildungen

Nun diskutieren wir eine weitere wichtige Klasse von linearen Abbildungen eines Euklidischen Raums: die Klasse der selbstadjungierten Abbildungen.

Sei  $V$  endlich-dimensionaler Euklidischer Raum über  $\mathbb{K}$  und sei  $F \in \text{Lin}(V)$ . Die Abbildung  $F$  heißt *selbstadjungiert*, falls  $F = F^*$  gilt, d.h.  $\langle F(u), v \rangle = \langle u, F(v) \rangle \quad \forall u, v \in V$ . Eine Matrix  $A \in \mathbb{C}^{n \times n}$  heißt *selbstadjungiert*, falls  $A = A^*$  gilt. (Insbesondere für  $A \in \mathbb{R}^{n \times n}$  ist  $A =$  selbstadjungiert  $\Leftrightarrow A$  symmetrisch).

Die Eigenwerte selbstadjungierter Abbildungen sind immer reell:

**Prop.** *Sei  $V$  endlich-dimensionaler Euklidischer Raum über  $\mathbb{C}$ . Sei  $F \in \text{Lin}(V)$  selbstadjungiert und sei  $\lambda \in \mathbb{C}$  Eigenwert von  $F$ . Dann folgt:  $\lambda \in \mathbb{R}$ .*

*Beweis.* Sei  $u$  Eigenvektor zu  $\lambda$ . Dann gilt:  $\lambda \langle u, u \rangle = \langle \lambda u, u \rangle = \langle F(u), u \rangle = \langle u, F(u) \rangle = \langle u, \lambda u \rangle = \bar{\lambda} \langle u, u \rangle$  Es folgt  $\lambda = \bar{\lambda}$  und somit  $\lambda \in \mathbb{R}$ .  $\square$

Ist  $u$  ein Eigenvektor einer selbstadjungierten Abbildung, so kann man

$F$  als eine Abbildung des Orthogonalkomplements  $u^\perp$  auffassen:

**Prop.** *Sei  $V$  endlich-dimensionaler Euklidischer Raum über  $\mathbb{C}$ . Sei  $F \in \text{Lin}(V)$  selbstadjungiert und sei  $u \in V \setminus \{0\}$  Eigenvektor von  $F$ . Dann gilt  $F(u^\perp) \subseteq u^\perp$ .*

*Beweis.* Sei  $\lambda$  Eigenwert zu  $u$ . Sei  $y \in F(u^\perp)$ , d.h.  $y = F(x)$  mit  $x \in V$  und  $\langle x, u \rangle = 0$ . Es gilt  $\langle y, u \rangle = \langle F(x), u \rangle = \langle x, F(u) \rangle = \langle x, \lambda u \rangle = \bar{\lambda} \langle x, u \rangle = 0$ . D.h.  $y \in u^\perp$ .  $\square$

### 7.2.5 Diagonalisierbarkeit von selbstadjungierten Abbildungen komplexer Euklidischer Räume

Selbstadjugierte Abbildungen sind in einer passend gewählten Orthonormalbasis durch eine Diagonalmatrix dargestellt:

**Thm.** *Sei  $V$  ein  $n$ -dimensionaler Euklidischer Raum über  $\mathbb{C}$ . Sei  $F \in \text{Lin}(V)$  selbstadjungiert. Dann existiert eine Orthonormalbasis  $\mathcal{B}$  von  $V$ ,*

für welche die Matrix  $F_{\mathcal{B}}$  diagonal ist und die Diagonalelemente von  $F_{\mathcal{B}}$  zu  $\mathbb{R}$  gehören.

*Beweis.* Man konstruiert eine Orthonormalbasis  $\mathcal{B}$  aus Eigenvektoren von  $F$  nach einem ähnlichen Muster wie im Beweis von Theorem 7.2.3. Da alle Eigenwerte von  $F$  reell sind, sind auch die Diagonalelemente von  $F_{\mathcal{B}}$  reell.  $\square$

**Kor.** Sei  $A \in \mathbb{C}^{n \times n}$  mit  $n \in \mathbb{N}$  selbstadjungiert. Dann existiert eine unitäre Matrix  $U \in \mathbb{C}^{n \times n}$ , sodass  $U^*AU$  eine Diagonalmatrix ist, deren Diagonalelemente in  $\mathbb{R}$  liegen.

*Beweis.* Folgt direkt aus dem vorigen Theorem.  $\square$

### 7.2.6 Diagonalisierung von selbstadjungierten Abbildungen reeller Euklidischer Räume

Nun diskutieren wir noch die selbstadjungierte Abbildungen reeller Euklidischer Räume.

**Thm.** *Sei  $V$   $n$ -dimensionaler Euklidischer Raum über  $\mathbb{R}$ . Sei  $F \in \text{Lin}(V)$  selbstadjungiert. Dann existiert eine Orthonormalbasis  $\mathcal{B}$  von  $V$ , für welche die Matrix  $F_{\mathcal{B}}$  diagonal ist.*

*Beweis.* Sei  $\mathcal{A}$  eine beliebige Orthonormalbasis von  $V$ . Da  $F$  selbstadjungiert ist, ist die Matrix  $F_{\mathcal{A}}$  symmetrisch: Sei  $\mathcal{A} = (a_1, \dots, a_n)$ , dann ist  $F(a_j) = \sum_{i=1}^n \langle F(a_j), a_i \rangle a_i$  und  $(F_{\mathcal{A}})_{ij} = \langle F(a_j), a_i \rangle = \langle a_j, F(a_i) \rangle = \langle F(a_i), a_j \rangle = (F_{\mathcal{A}})_{ji}$ .  $F_{\mathcal{A}}$  als eine Matrix aus  $\mathbb{C}^{n \times n}$  ist selbstadjungiert. Das charakteristische Polynom von  $F_{\mathcal{A}}$  hat mindestens eine Nullstelle  $\lambda \in \mathbb{C}$ . Nach 7.2.4 gilt  $\lambda \in \mathbb{R}$ . Dann ist  $\lambda \in \mathbb{R}$  ein Eigenwert von  $F$ . Wir betrachten einen Eigenvektor  $u \in V$  mit  $\|u\| = 1$ . Ab dieser Stelle verläuft der Beweis

analog zum Beweis des Theorems 7.2.3. □

Das vorige Theorem wird auch das Theorem über die Hauptachsentransformation genannt. Dazu ein Beispiel:

**Bsp.** Wir betrachten die Gleichung  $34x^2 + 24xy + 41y^2 = 25$  in Unbekannten  $x, y \in \mathbb{R}$ . Diese Gleichung beschreibt eine Ellipse. Wir möchten herausfinden, wie diese Ellipse ausgerichtet ist. Die Gleichung kann folgendermaßen mit Matrizen und Vektoren geschrieben werden:

$$\begin{pmatrix} x & y \end{pmatrix} \cdot \underbrace{\begin{pmatrix} 34/25 & 12/25 \\ 12/25 & 41/25 \end{pmatrix}}_A \cdot \begin{pmatrix} x \\ y \end{pmatrix} = 1. \quad (7.2.2)$$

Die Abbildung  $F(u) = Au$  aus  $\text{Lin}(\mathbb{R}^2)$  ist selbstadjungiert, da  $A = A^\top$  gilt.  $\mathbb{R}^2$  besitzt eine Orthonormalbasis aus Eigenvektoren von  $F$ . Nach etwas Rechnen stellt sich heraus, dass  $a = (4/5, 3/5)^\top$  Eigenvektor zum Eigenwert 2 und  $b = (-3/5, 4/5)$  Eigenvektor zum Eigenwert 1 ist. Das die beiden

Vektoren  $a$  und  $b$  Länge 1 haben, ist also  $a, b$  die gesuchte Orthonormalbasis (die Orthonormalbasis ist bis das Ändern von  $a$  zu  $-a$  und/oder  $b$  zu  $-b$  eindeutig). Ein Vektor  $u \in \mathbb{R}^2$  hat die Koordinaten  $\langle u, a \rangle$  und  $\langle u, b \rangle$  in unserer Basis  $a, b$ . Das bedeutet

$$u = \langle u, a \rangle a + \langle u, b \rangle b$$

Somit hat man

$$\begin{aligned}\langle F(u), u \rangle &= \langle \langle u, a \rangle 2a + \langle u, b \rangle b, \langle u, a \rangle a + \langle u, b \rangle b \rangle \\ &= 2\underbrace{\langle u, a \rangle}_x^2 + \underbrace{\langle u, b \rangle}_y^2\end{aligned}$$

Die Gleichung im neuen Koordinatensystem lautet:

$$2\tilde{x}^2 + \tilde{y}^2 = 1$$

Offenbar sind die Vektoren  $a$  und  $b$  die Richtungen der sogenannten Hauptachsen der Ellipse, welche durch die Gleichung (7.2.2) beschrieben ist.

**Bsp** (Hauptkomponentenanalyse). Bei der Analyse, wie eine Konfiguration von Punkten  $p_1, \dots, p_k \in \mathbb{R}^n$  ( $k \geq 2$ ) im Raum “ausgerichtet” ist, kann die sogenannte Hauptkomponentenanalyse benutzt werden. Dafür zentriert man zuerst die Originalkonfiguration zur Konfiguration  $u_1, \dots, u_k$  mit  $u_i := p_i - \bar{p}$  und  $\bar{p} := \frac{1}{k} \sum_{i=1}^k p_i$ . Jedem  $u_i$  wird die symmetrische Matrix  $u_i u_i^\top$  vom Rang eins zugeordnet. Der Kern dieser symmetrischen Matrix ist das Orthogonalkomplement von  $u_i$ ; im Fall  $u_i \neq 0$  ist  $u_i$  der Eigenvektor zum einzigen Nichnulleigenwert, der gleich  $u_i^\top u_i = \|u_i\|^2$  ist. Die Matrix  $u_i u_i^\top$  behält also die Information die Richtung von  $u_i$  und die Länge von  $u_i$ . Das heißt, für eine einpunktige Konfiguration  $u$  wird durch die Angabe einer Orthonormalbasis  $b_1, \dots, b_n$  aus Eigenvektoren zu den Eigenwerten  $\lambda, 0, \dots, 0$  festgestellt, dass die einpunktige Konfiguration  $u$  entlang  $b_1$  ausgerichtet ist der Punkt  $u$  in einer der beiden Positionen  $\pm\sqrt{\lambda}b_1$  liegt.

Durch die symmetrische Matrix

$$C = \frac{1}{k-1} \sum_{i=1}^k u_i u_i^\top$$

werden all die Rang-Eins-Matrizen zu den Vektoren  $u_1, \dots, u_k$  zu einer Matrix zusammengefasst. Die Eigenpaare von  $C$  vermitteln die Information darüber, wie die Konfiguration der Vektoren  $u_1, \dots, u_k$  ausgerichtet ist. Die Angabe einer Orthonormalbasis  $b_1, \dots, b_n$  zu Eigenwerten  $\lambda_1 \geq \dots \geq \lambda_n$  geben eine Zusammenfassung, welche “Tendenzen” man bei der Ausrichtung und Positionierung der Vektoren aus  $u_1, \dots, u_k$  hat.

Hierzu noch einige Kommentare: Aus den Ergebnissen im nächsten Kapitel wird klar, dass alle Eigenwerte von  $C$  nichtnegativ sind. Der Konstante Faktor  $\frac{1}{k-1}$  ist lediglich für eine Skalierung zuständig. Wieso man den Faktor  $\frac{1}{k-1}$  und nicht etwa  $\frac{1}{k}$  fixiert, wird in den Kursen zur Statistik und Wahrscheinlichkeitstheorie erklärt (der Unterschied bei der Nutzung des Faktors  $\frac{1}{k}$  ist sehr gering, weil das Verhältnis  $\frac{k-1}{k}$  für große  $k$  nahezu 1 ist). Die Ma-

trix  $C \in \mathbb{R}^{n \times n}$  nennt man die Stichproben-Kovarianzmatrix zur Stichprobe  $p_1, \dots, p_k$  mit Umfang  $k$ .

**Thm.** *Sei  $A \in \mathbb{R}^{n \times n}$  symmetrisch mit  $n \in \mathbb{N}$ . Dann existiert eine orthogonale Matrix  $U \in \mathbb{R}^{n \times n}$ , d.h.  $U^\top U = I$ , für welche  $U^\top A U$  diagonal ist.*

*Beweis.* Folgt direkt aus dem vorigen Theorem.  $\square$

**Bem.** Das vorige Theorem wird ebenfalls Theorem über Hauptachsentransformationen genannt.

**Kor.** *Zwei Eigenvektoren zu unterschiedlichen Eigenwerten einer selbstadjungierten Abbildung sind zueinander orthogonal.*

*Beweis.* Aufgabe.  $\square$

## 7.3 Quadratische Formen

### 7.3.1 Motivation und Grundbegriffe

In der linearen Algebra haben wir uns bis jetzt oft mit linearen Abbildungen und linearen Funktionen beschäftigt. Lineare Funktionen sind Polynomalfunktionen zu Polynomen vom Grad 1. Man kann natürlich die weiteren Schritte machen und die komplizierteren Fälle vom Grad 2, 3 usw. untersuchen. Hierbei ist der Fall vom Grad 2 aus verschiedenen Gründen besonders wichtig. Einen der Gründe haben wir bereits gesehen: die Norm zum Quadrat in Euklidischen Vektorräumen ist eine quadratische Funktion.

Wir wollen uns also die quadratischen Funktionen genauer anschauen. Etwas präziser beschrieben geht es um die homogenen quadratischen Funktionen. Untersuchungen quadratischer Funktionen könnte man quadratische Algebra nennen, wir werden aber sehen, dass man die quadratische Algebra recht zügig zur linearen Algebra reduzieren kann, sodass man den

Begriff quadratische Algebra gar nicht benutzt. Das Zusammenspiel des nichtlinearen und linearen Falls ist oft so: Lineare Algebra führt immer zu nichtlinearen Funktionen und Problemen, während nichtlineare Probleme oft mit Hilfe der linearen Algebra gelöst werden. Beispiele dafür sind wie folgt. Das charakteristische Polynom ist nicht linear. Die Determinante als Funktion der  $n^2$  Komponenten ist ebenfalls nicht linear. Die Eigenwertaufgabe  $Ax = \lambda x$  ist nicht linear, da man  $\lambda$  und  $x$  multipliziert und  $\lambda$  sowie  $x$  unbekannt sind. Beispiele in die andere Richtung: um nichtlineare Probleme zu lösen, werden diese oft linearisiert. Die allgemeine Algebra (die man ganz grob als Theorie der nichtlinearen Algebraischen Mengen und Polynomen beliebiger Polynome beschreiben kann) benutzt die lineare Algebra als eine “Schlüsseltechnologie”.

Homogene quadratische Funktionen finden eine wichtige Anwendung bei der Lösung multivariaten Optimierungsaufgaben. Um die Anwendung zu illustrieren beginnen wir zunächst mit einer univariaten Funktion  $f : \mathbb{R} \rightarrow \mathbb{R}$ ,

die genügend oft differenzierbar ist. Wir wollen bestimmen, ob  $f$  an einer Stelle, etwa an der 0, ein lokales Minimum erreicht. Durch die Taylorreihenentwicklung vom Grad 2 können wir die Funktion wie folgt in Terme zerlegen:

$$f(x) = \underbrace{f(0)}_{\text{Konstante (egal)}} + \underbrace{f'(0)x}_{\text{der lineare Term}} + \underbrace{\frac{f''(0)}{2}x^2}_{\text{der quadratische Term}} + \underbrace{R(x)}_{\text{Restterm}}$$

Die Bedingung  $f'(0)$  ist notwendig für ein lokales Minimum, es ist also notwendig, dass der lineare Term identisch null ist. Wenn die notwendige Bedingung  $f'(0) = 0$  erfüllt ist, weiß man, dass die Bedingung  $f''(0) > 0$  für ein lokales Minimum in 0 hinreichend ist (das alles wissen Sie wahrscheinlich bereits aus dem Analysis-Kurs). Nun schauen wir uns die multivariate Situation an, etwa den Fall von zwei Variablen. Betrachten wir eine zweivariate Funktion  $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ , die genügend oft differenzierbar ist. Die Taylorentwicklung vom Grad 2 im Nullpunkt sieht für diese Funktion sieht

so aus:

$$f(x_1, x_2) = \underbrace{f(0)}_{\text{Konstante}} + \underbrace{(\partial_1 f)(0)x_1 + (\partial_2 f)(0)x_2}_{\text{der lineare Term}} + \underbrace{\frac{(\partial_1^2 f)(0)}{2}x_1^2 + (\partial_1 \partial_2 f)(0)x_1 x_2 + \frac{(\partial_2^2 f)(0)}{2}x_2^2}_{\text{der quadratische Term}} + \underbrace{R(x)}_{\text{Restterm}}$$

Auch im Multivariatenfall ist es notwendig für ein lokales Minimum in einem Punkt, dass der lineare Term in diesem Punkt identisch null ist. Die notwendige Bedingung für ein lokales Minimum in 0 ist somit  $(\partial_1 f)(0) = (\partial_2 f)(0) = 0$ . Im multivariaten Fall hat man eine hinreichende Bedingung, die analog zur präsentierten Bedingung im univariaten Fall ist. Wenn der lineare Term identisch null, ist dann reicht es aus, dass der quadratische Term für alle  $x$ , die Ungleich 0 ist, strikt positiv ist. Wir wollen also bestimmen können, wann eine homogene quadratische Funktion für alle  $x \neq 0$  strikt positiv ist. Im Folgenden werden wir verschiedene Mittel dazu entwickeln.

Nun sind wir soweit, die Grunddefinitionen zu geben. Für den “konkreten” Vektorraum  $\mathbb{K}^n$  können quadratische Formen mit Hilfe von Matrizen eingeführt werden. Eine Funktion  $q : \mathbb{K}^n \rightarrow \mathbb{K}$  heißt *quadratische Form*, falls  $q$  als  $q(x) = x^\top Ax$  für eine Matrix  $A \in \mathbb{K}^{n \times n}$  beschrieben werden kann. Mit Komponenten  $x = (x_i)_{i=1}^n$  beschrieben, hat man den Ausdruck  $q(x) = \sum_{i,j=1}^n a_{i,j}x_i x_j$ . Das heißt,  $q(x)$  ist Linearkombination der Produkte  $x_i x_j$  mit  $i, j \in \{1, \dots, n\}$ . Ein Beispiel:

$$q(x_1, x_2) = x_1^2 - 5x_1 x_2 + 6x_2^2$$

ist eine quadratische Form auf  $\mathbb{R}^2$ . Was ist eine Matrix  $A$  dazu?

(Das Wort Form steht übrigens in Algebra für homogene Funktionen/Polynome. Die Determinante einer  $3 \times 3$  Matrix als Funktion ihrer 9 Komponenten ist zum Beispiel eine kubische Form.)

Für abstrakte Vektorräume werden die quadratischen Formen mit Hilfe der bilinearen Formen eingeführt. Es sei daran erinnert, dass wir die bilinearen Formen bereits zur Einführung von Euklidischen Räumen über  $\mathbb{R}$

benutzt haben. Sei  $V$  Vektorraum über  $\mathbb{K}$ . Eine Funktion  $q : V \rightarrow \mathbb{K}$  heißt *quadratische Form*, falls eine bilineare Form  $b$  auf  $V$  existiert, für die  $q(v) = b(v, v) \quad \forall v \in V$  gilt. Wir setzen also in eine bilineare Form als ersten und zweiten Vektorargument den selben Vektor ein, um eine quadratische Form zu erhalten.

### 7.3.2 Darstellung quadratischer Formen durch symmetrische bilineare Formen

Bei der Diskussion quadratischer Formen entsteht ein technisches Problem im Bezug auf den gewählten Körper  $\mathbb{K}$ . Wenn man sich zum Beispiel voll auf die Körper  $\mathbb{R}$  und  $\mathbb{C}$  konzentriert, lässt sich dieses Problem komplett vermeiden, wenn man aber auch die endlichen Körper mitberücksichtigen will, dann muss man sich mit dem Problem auseinandersetzen. Quadratische Formen über endlichen Körpern sind für Anwendungen in der Kodierungstheorie und Kryptographie wichtig, das ist einer der Gründe, warum

man die Theorie der quadratischen Formen für allgemeine Körper entwickeln will. Das Problem ist wie folgt: In manchen endlichen Körpern gilt die Gleichung  $1 + 1 = 0$ . Das einfachste und wichtigste Beispiel dafür ist der binäre Körper  $\mathbb{K} = \{0, 1\}$ . Der Fall des Körpers mit  $1 + 1 = 0$  in der Theorie von quadratischen Formen ist ein Ausnahmefall. Wir werden die Theorie für den “Regelfall”  $1 + 1 \neq 0$  entwickeln.

Die folgende Proposition zeigt Folgendes. Während jede symmetrische bilineare Form eine quadratische Form bestimmt, lässt sich die bilineare Form aus der quadratischen Form in dem “Regelfall” rekonstruieren:

**Prop** (Polarisationsformel). *Sei im Körper  $\mathbb{K}$  die Bedingung  $1 + 1 \neq 0$  erfüllt. Sei  $V$  Vektorraum über  $\mathbb{K}$ . Zu jeder quadratischen Form  $q : V \rightarrow \mathbb{K}$  existiert eine eindeutige symmetrische bilineare Form  $f$  mit  $q(v) = f(v, v) \quad \forall v \in V$ . Diese bilineare Form  $f$  ist durch die Gleichung  $f(u, v) = \frac{1}{2}(q(u + v) - q(u) - q(v))$  bestimmt.*

*Beweis.* Aufgabe. □

**Bem.**  $\frac{1}{2}$  ist das inverse Element. Für die Körper  $\mathbb{R}$  und  $\mathbb{C}$  ist  $\frac{1}{2}$  im ganz normalen Sinne (wie in der Schule). Für allgemeine  $\mathbb{K}$  sollte  $\frac{1}{2}$  im Sinne der Arithmetik von  $\mathbb{K}$  verstanden werden, das das inverse Element zu  $1 + 1$ . Das ist zum Beispiel die Restklasse von 3 in  $\mathbb{K} = \mathbb{Z}/5\mathbb{Z}$ .

### 7.3.3 Basisdarstellungen bilinearer und quadratischer Formen

Genauso wie bei den anderen Themen der linearen Algebra ist auch bei quadratischen Formen eine Möglichkeit, passende Koordinaten (= eine passende Basis) zu wählen, sehr wichtig. Wir definieren also die Basisdarstellung bilinearer und quadratischer Formen.

Sei  $V$  ein  $n$ -dimensionaler Vektorraum über  $\mathbb{K}$  mit  $n \in \mathbb{N}$  und sei  $f : V \times V \rightarrow \mathbb{K}$  bilinear. Sei  $\mathcal{B} = (b_1, \dots, b_n)$  eine Basis von  $V$ . Die Matrix  $f_{\mathcal{B}} = (f(b_i, b_j))_{i,j=1}^n$  heißt die Matrix von  $f$  in der Basis  $\mathcal{B}$ .

Für alle  $u, v \in V$  gilt:  $f(u, v) = u_{\mathcal{B}}^\top f_{\mathcal{B}} v_{\mathcal{B}}$  (direkt verifizierbar). Für Basen  $\mathcal{A}$  und  $\mathcal{B}$  von  $V$  hat man die folgende Form (für den Basiswechsel):

$f_{\mathcal{B}} = T_{\mathcal{A} \leftarrow \mathcal{B}}^\top f_{\mathcal{A}} T_{\mathcal{A} \leftarrow \mathcal{B}}$  (Beweis: Aufgabe). Beachten Sie hier das folgende: bei Basiswechsel für Matrizen und Abbildungen, hatten wir die inverse der Basiswechsel-Matrix benutzt. Hier benutzen wir dagegen die transponierte Matrix.

Im Fall eines Körpers  $\mathbb{K}$  mit  $1 + 1 \neq 0$ , wird die Darstellung einer quadratischen Form  $q$  auf  $V$  durch  $q_{\mathcal{B}} = f_{\mathcal{B}}$  definiert, wo  $f$  die symmetrische bilineare Form zu  $q$  bezeichnet. D.h.  $q_{\mathcal{B}} = (f(b_i, b_j))_{i,j=1,\dots,n} = (\frac{1}{2}(q(b_i + b_j) - q(b_i) - q(b_j)))_{i,j=1}^n$ .

### 7.3.4 Diagonalisierung von quadratischen Formen

Die einfachsten quadratischen Formen sind die diagonalen Formen, wie etwa  $q(x_1, x_2, x_3) = x_1^2 + 2x_2^3 + 5x_3^2$  oder  $r(x_1, x_2) = x_1^2 - 2x_2^2$ . Im Fall des Körpers  $\mathbb{R}$  können wir für solche Formen für uns wichtige Eigenschaften sehr einfach nachweisen:  $q$  oben ist zum Beispiel für alle  $x \in \mathbb{R}^3 \setminus \{0\}$  strikt positiv,  $r$  dagegen kann positive sowie negative Werte annehmen. Wenn also  $q$  ein

quadratischer Term ist, in einer Taylorentwicklung, bei der lineare Term gleich 0 ist, dann wissen wir, dass die Funktion, die wir entwickelt haben, im gegebenen Punkt ihr lokales Minimum erreicht.

Das folgende Theorem präsentiert eine wichtige Botschaft: jede quadratische Form ist in passenden Koordinaten diagonal. Betrachten wir kurz als Beispiel die ganz einfache quadratische Form  $h(x_1, x_2) = x_1 x_2$  auf  $\mathbb{R}^2$ . Die kann als  $h(x_1, x_2) = \frac{1}{2}(x_1 + x_2)^2 - \frac{1}{2}(x_1 - x_2)^2$  dargestellt werden. Die quadratische Form  $h$  ist somit diagonal in den Koordinaten  $y_1 = x_1 + x_2$  und  $y_2 = x_1 - x_2$ .

**Thm.** *Sei in  $\mathbb{K}$   $1 + 1 \neq 0$  erfüllt. Sei  $V$   $n$ -dimensionaler Vektorraum und  $q : V \rightarrow \mathbb{K}$  eine quadratische Form. Dann existiert eine Basis  $\mathcal{B}$  von  $V$ , für welche die Matrix  $q_{\mathcal{B}}$  diagonal ist.*

*Beweis.* Induktion über  $n$ . Der Fall  $n = 1$  ist trivial. Sei die Behauptung für festes  $n - 1$  erfüllt. Sei  $f$  symmetrische bilineare Form zu  $q$ . Wenn  $q$  (und somit auch  $f$ ) identisch 0 ist, dann gilt die Behauptung für eine

beliebige Basis von  $V$ . Ansonsten wähle ein  $u \in V \setminus \{0\}$  mit  $q(u) \neq 0$ . Betrachte die lineare Funktion  $L(x) = f(u, x)$   $\forall x \in V$  von  $V$  nach  $\mathbb{K}$ . Da  $L(u) = f(u, u) \neq 0$ , ist  $\text{im}(L) = \mathbb{K}$ . Somit folgt aus dem Rangsatz  $n = \dim(\text{im}(L)) + \dim(\ker(L)) \Rightarrow \dim(\ker(L)) = n - 1$ . Nach Induktionsvoraussetzung existiert eine Basis  $(b_2, \dots, b_n)$  von  $\ker(L)$ , für welche die Matrix der quadratischen Form  $q|_{\ker(L)}$  diagonal ist. Die Behauptung gilt für  $\mathcal{B} = (b_1, \dots, b_n)$  mit  $b_1 = u$ .  $\square$

**Bsp** (und Bemerkung). Sei  $n \in \mathbb{N}$ ,  $A \in \mathbb{K}^{n \times n}$  symmetrisch,  $1 + 1 \neq 0$  in  $\mathbb{K}$ , und  $q(x) = x^\top Ax$  eine quadratische Form auf  $\mathbb{K}^n$ . Wie diagonalisiert man  $q$ ?

Etwa für  $A = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix} \in \mathbb{R}^{3 \times 3}$ , d.h.  $q(x_1, x_2, x_3) = 2(x_1x_2 + x_2x_3 + x_3x_1)$ .

Das Vorgehen ist ähnlich zum Gauß-Verfahren. Allerdings werden in jedem

Schritt eine Spalten- und eine entsprechende Zeilenoperation ausgeführt:  
 (Bezeichnungen für Spaltenoperationen:  $s_i \leftrightarrow s_j$ , usw. Für Zeilen:  $z_i \leftrightarrow z_j$ , usw.)

$$\begin{array}{lll}
 \rightarrow s_1 := s_1 + s_2 & \rightarrow \begin{pmatrix} 1 & 1 & 1 \\ 1 & 0 & 1 \\ 2 & 1 & 0 \end{pmatrix} & \rightarrow z_1 := z_1 + z_2 \quad \rightarrow \begin{pmatrix} 2 & 1 & 2 \\ 1 & 0 & 1 \\ 2 & 1 & 0 \end{pmatrix} \\
 \rightarrow s_2 := s_2 - \frac{1}{2}s_1 & \rightarrow \begin{pmatrix} 2 & 0 & 2 \\ 1 & -\frac{1}{2} & 1 \\ 2 & 0 & 0 \end{pmatrix} & \rightarrow z_2 := z_2 - \frac{1}{2}z_1 \quad \rightarrow \begin{pmatrix} 2 & 0 & 2 \\ 0 & -\frac{1}{2} & 0 \\ 2 & 0 & 0 \end{pmatrix} \\
 \rightarrow s_3 := s_3 - s_1 & \rightarrow \begin{pmatrix} 2 & 0 & 0 \\ 0 & -\frac{1}{2} & 1 \\ 2 & 0 & -2 \end{pmatrix} & \rightarrow z_3 := z_3 - z_1 \quad \rightarrow \begin{pmatrix} 2 & 0 & 0 \\ 0 & -\frac{1}{2} & 0 \\ 0 & 0 & -2 \end{pmatrix}
 \end{array}$$

Für die Diagonalmatrix  $D = \text{diag}(2, -1/2, -2)$  gilt  $D = M_3^\top M_2^\top M_1^\top A M_1 M_2 M_3$  bzw.  $D = M^\top A M$ , wobei  $M = M_1 M_2 M_3$  und

$$M_1 = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad M_3 = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

$$M_2 = \begin{pmatrix} 1 & -\frac{1}{2} & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad M = \begin{pmatrix} 1 & -\frac{1}{2} & -1 \\ 1 & \frac{1}{2} & -1 \\ 0 & 0 & 1 \end{pmatrix}$$

D.h. für  $u = (u_1, u_2, u_3) \in \mathbb{R}^3$  gilt:

$$q(Mu) = (Mu)^\top = AMu = u^\top M^\top A M u = u^\top Du = 2u_1^2 - \frac{1}{2}u_2^2 - 2u_3^2$$

Um  $M$  auszurechnen genügt es, dieselben Spaltentransformationen mit der

Einheitsmatrix beginnend auszuführen:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{I \cdot M_1} \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{I \cdot M_1 M_2} \begin{pmatrix} 1 & -\frac{1}{2} & 0 \\ 1 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix} \xrightarrow{I \cdot M_1 M_2 M_3} \begin{pmatrix} 1 & -\frac{1}{2} & -1 \\ 1 & \frac{1}{2} & -1 \\ 0 & 0 & 1 \end{pmatrix} = M$$

Wählt man nun  $x = Mu$ , so gilt  $2(x_1x_2 + x_2x_3 + x_3x_1) = q(x) = q(Mu) = 2u_1^2 - \frac{1}{2}u_2^2 - 2u_3^2$ .

Allgemeiner: Wenn man nur Nullen auf der Diagonale hat, muss man dort Nichtnullelemente erzeugen. Wenn man ein Nichtnullelement dort hat, kann man es verwenden, um die restlichen Elemente der Zeile / Spalte durch 0 zu ersetzen.

**Bem.** Im Fall  $\mathbb{K} = \mathbb{R}$  kann eine gegebene quadratische Form  $q : \mathbb{R}^n \rightarrow \mathbb{R}$  sogar mit Hilfe einer Orthogonal-Matrix diagonalisiert werden. In diesem Fall müssen aber die Eigenwerte der symmetrischen Matrix  $A$  berechnet werden.

### 7.3.5 Definitheit, Semidefinitheit und Indefinitheit

Sei  $V$  Vektorraum über  $\mathbb{R}$  und  $q : V \rightarrow \mathbb{R}$  eine quadratische Form.  $q$  heißt

- *positiv definit*, wenn  $q(v) > 0 \quad \forall v \in V \setminus \{0\}$
- *positiv semidefinit*, wenn  $q(v) \geq 0 \quad \forall v \in V$
- *negativ definit*, wenn  $q(v) < 0 \quad \forall v \in V \setminus \{0\}$
- *negativ semidefinit*, wenn  $q(v) \leq 0 \quad \forall v \in V$
- *indefinit*, wenn  $x, y \in V$  existieren mit  $q(x) > 0$  und  $q(y) < 0$ .

Durch Diagonalisierung einer quadratischen Form kann entschieden werden, welchen der obigen Typen sie hat. Außerdem gilt:

$$q \text{ positiv definit} \Leftrightarrow -q \text{ negativ definit}$$

$$q \text{ positiv semidefinit} \Leftrightarrow -q \text{ negativ semidefinit}$$

Außerdem benutzt man auch für symmetrische Matrix  $A \in \mathbb{R}^{n \times n}$  Begriffe wie oben (positiv semidefinite usw.) im Bezug auf die zugrundeliegende quadratische Form  $q(x) = x^\top Ax$ .

**Bem** (Der Zusammenhang der quadratischen Formen und selbstadjungierten Abbildungen). Die vorigen Begriffe können auch für selbst-adjungierten Abbildungen definiert werden. Ist  $F : V \rightarrow V$  eine selbstadjungierte linearen Abbildung auf einem Euklidischen Raum über  $\mathbb{K}$ , so ist unabhängig davon ob  $\mathbb{K} = \mathbb{C}$  oder  $\mathbb{K} = \mathbb{R}$  ist die Funktion  $q_F(u) := \langle F(u), u \rangle$  reellwertig, denn  $\langle F(u), u \rangle = \langle u, F(u) \rangle = \overline{\langle u, F(u) \rangle}$ . Im Fall von  $\mathbb{K} = \mathbb{R}$  ist  $q_F(u)$  eine quadratische Form (bei  $\mathbb{K} = \mathbb{C}$  aber nicht). Man nennt  $F$  positiv-semidefinit wenn  $q_F(u) \geq 0$  für alle  $u \in V$  erfüllt ist. Die anderen Begriffe, die wir oben für quadratische Formen eingefürt haben, wie positiv definit, negativ definit usw. werden analog definiert.

### 7.3.6 Charakterisierung der Definitheit, Semidefinitheit und Indefinitheit mit Hilfe der Eigenwerte

Theorie der Eigenwerte spielt auch bei der Untersuchungen quadratischer Formen eine sehr wichtige Rolle, wie das folgende Theorem zeigt.

**Thm.** Sei  $q : \mathbb{R}^n \rightarrow \mathbb{R}$  quadratische Form mit  $q(x) = x^\top A x \quad \forall x \in \mathbb{R}^n$ , wobei  $A \in \mathbb{R}^{n \times n}$  eine symmetrische Matrix ist. Dann gilt:

- (i)  $q$  positiv definit  $\Leftrightarrow$  alle Eigenwerte von  $A$  sind positiv,
- (ii)  $q$  positiv semidefinit  $\Leftrightarrow$  alle Eigenwerte von  $A$  sind nichtnegativ,
- (iii)  $q$  negativ definit  $\Leftrightarrow$  alle Eigenwerte von  $A$  sind negativ,
- (iv)  $q$  negativ semidefinit  $\Leftrightarrow$  alle Eigenwerte von  $A$  sind nichtpositiv,
- (v)  $q$  indefinit  $\Leftrightarrow A$  hat positive und negative Eigenwerte.

*Beweis.* Da  $A$  symmetrisch ist, existiert eine Orthogonalmatrix  $U$ , für welche die Matrix  $U^\top AU$  diagonal ist, d.h.  $D := U^\top AU = \text{diag}(\lambda_1, \dots, \lambda_n)$  mit  $\lambda_1, \dots, \lambda_n \in \mathbb{R}$ . Es ist  $p_A(t) = p_D(t) = (t - \lambda_1) \cdots (t - \lambda_n)$ . Daher ist  $\{\lambda_1, \dots, \lambda_n\}$  die Menge aller Eigenwerte von  $A$ . Betrachte zunächst (i):

$$\begin{aligned}
 & q \text{ ist positiv definit} \\
 \Leftrightarrow & q(v) > 0 \quad \forall v \in \mathbb{R}^n \setminus \{0\} \\
 \Leftrightarrow & q(Ux) > 0 \quad \forall x \in \mathbb{R}^n \setminus \{0\} \\
 \Leftrightarrow & (Ux)^\top AUx > 0 \quad \forall x \in \mathbb{R}^n \setminus \{0\} \\
 \Leftrightarrow & x^\top U^\top AUx > 0 \quad \forall x \in \mathbb{R}^n \setminus \{0\} \\
 \Leftrightarrow & x^\top Dx > 0 \quad \forall x \in \mathbb{R}^n \setminus \{0\} \\
 \Leftrightarrow & \lambda_1 x_1^2 + \dots + \lambda_n x_n^2 > 0 \quad \forall (x_1, \dots, x_n) \in \mathbb{R}^n \setminus \{0\} \\
 \Leftrightarrow & \lambda_1 > 0, \dots, \lambda_n > 0
 \end{aligned}$$

Die restlichen Behauptungen werden analog bewiesen.  $\square$

Das vorige Theorem wird oft bei Rechenaufgaben für zwei-variate quadratische Formen benutzt, weil in diesem Fall die Gleichung  $p_A(\lambda) = 0$  eine quadratische Gleichung ist und man die Formel für die Lösungen der quadratischen Gleichung bereits in der Schule kennengelernt hat.

### **7.3.7 Charakterisierung der Definitheit, Semidefinitheit und Indefinitheit durch die Koeffizienten des charakteristischen Polynoms**

Wenn man mit Hilfe von 7.3.7 entscheiden möchte, ob die gegebene Matrix positiv (semi)definit oder negative (semi)definit ist, muss man die Eigenwerte von  $A$  ausrechnen: das ist keine leichte Aufgabe. In dem folgenden Theorem wird gezeigt, wenn man die Aufgabe bewältigt, ohne dass man die Eigenwerte ausrechnen muss. Wenn man das charakteristische Polynom  $p_A$  der symmetrischen Matrix  $A$  kennt, reicht nämlich ein Blick auf die Koeffizienten von  $p_A$ .

**Thm.** Sei  $q : \mathbb{R}^n \rightarrow \mathbb{R}$  quadratische Form mit  $q(x) = x^\top Ax \quad \forall x \in \mathbb{R}^n$  für eine symmetrische Matrix  $A \in \mathbb{R}^{n \times n}$  mit  $n \in \mathbb{N}$ . Dann gilt:

- (i)  $q$  ist negativ definit  $\Leftrightarrow$  die Koeffizienten von  $p_A$  sind alle positiv.
- (ii)  $q$  ist negativ semidefinit  $\Leftrightarrow$  die Koeffizienten von  $p_A$  sind alle nicht-negativ.
- (iii)  $q$  ist positiv definit  $\Leftrightarrow$  die Koeffizienten von  $p_{-A}(t) = (-1)^n p_A(-t)$  sind alle positiv.
- (iv)  $q$  ist positiv semidefinit  $\Leftrightarrow$  die Koeffizienten von  $p_{-A}(t) = (-1)^n p_A(-t)$  sind alle nicht-negativ.

*Beweis.* Wir betrachten  $\lambda_1, \dots, \lambda_n$  wie im vorigen Beweis, dann ist  $p_A(t) = (t - \lambda_1) \cdots (t - \lambda_n)$ . Sei  $\mu_i = -\lambda_i$  für  $i \in \{1, \dots, n\}$ . Dann gilt

$$p_A(t) = (t + \mu_1) \cdots (t + \mu_n) = t^n + (\mu_1 + \dots + \mu_n)t^{n-1} + \dots + \mu_1 \cdots \mu_n t^0$$

und demnach:

(i)  $q$  ist negativ definit

$$\Leftrightarrow \lambda_i < 0 \quad \forall i \in \{1, \dots, n\}$$

$$\Leftrightarrow \mu_i > 0 \quad \forall i \in \{1, \dots, n\}$$

$\Rightarrow$  die Koeffizienten von  $p_A$  sind alle positiv

Umgekehrt: seien alle Koeffizienten von  $p_A$  positiv. Für  $\lambda \in \mathbb{R}$  mit  $\lambda \geq 0$  gilt  $p_A(\lambda) > 0$ .  $\Rightarrow p_A$  hat keine nicht-negativen Nullstellen.  $\Rightarrow$  alle Nullstellen  $\lambda_1, \dots, \lambda_n$  sind negativ.  $\Rightarrow q$  ist negativ definit.

(ii) wird analog gezeigt.

(iii) folgt aus (i).

(iv) folgt aus (ii). □

### 7.3.8 Charakterisierung der Definitheit mit Hauptminoren

Wenn man eine nicht allzu große symmetrische Matrix hat und man herausfindet möchte, ob die quadratische Form dazu positiv oder negativ definit ist, kann man noch das folgende Kriterium benutzen, das auf der Berechnung von sogenannten Hauptminoren basiert.

**Thm.** Sei  $q : \mathbb{R}^n \rightarrow \mathbb{R}$  quadratische Form mit  $q(x) = x^\top Ax \quad \forall x \in \mathbb{R}^n$  für eine symmetrische Matrix  $A = (a_{ij})_{i,j=1}^n$  mit  $n \in \mathbb{N}$ . Für jedes  $k \in \{1, \dots, n\}$  sei  $A_k := (a_{ij})_{i,j=1}^k \in \mathbb{R}^{k \times k}$ . Dann gilt:

- (i)  $q$  ist positiv definit  $\Leftrightarrow \det(A_k) > 0 \quad \forall k \in \{1, \dots, n\}$ .
- (ii)  $q$  ist negativ definit  $\Leftrightarrow (-1)^k \det(A_k) > 0 \quad \forall k \in \{1, \dots, n\}$ .

*Beweis.* (ii) folgt durch Anwendung von (i) zur quadratischen Form  $-q$ ; zeige also (i):

,,  $\Rightarrow$ “ Sei  $q$  positiv definit. Sei  $k \in \{1, \dots, n\}$ . Dann ist die quadratische Form  $q(x_1, \dots, x_k, 0, \dots, 0)$  positiv definit und durch die Matrix  $A_k$  gegeben.  $\det(A_k) > 0$ , denn  $\det(A_k)$  ist ein Koeffizient von  $p_{-A_k}(t)$ ; vgl. Theorem 7.1.3 (iii).

,,  $\Leftarrow$ “ Durch Induktion über  $n \in \mathbb{N}$ .

Sei  $n \geq 2$  und sei die Aussage (Implikation „ $\Leftarrow$ “) mit  $n - 1$  an der Stelle von  $n$  erfüllt. Sei  $q : \mathbb{R}^n \rightarrow \mathbb{R}$  eine quadratische Form mit  $\det(A_k) > 0 \ \forall k \in \{1, \dots, n\}$ . Aus der Induktionsvoraussetzung folgt, dass die quadratische Form  $q(x_1, \dots, x_{n-1}, 0)$  auf  $\mathbb{R}^{n-1}$  positiv definit ist, d.h.  $\det(A_1), \dots, \det(A_{n-1}) > 0$ . Wir betrachten eine Orthogonalmatrix  $\tilde{U} \in \mathbb{R}^{(n-1) \times (n-1)}$ , für welche die Matrix  $\tilde{U}^\top A_{n-1} \tilde{U}$  diagonal ist, d.h.  $\tilde{D} := \tilde{U}^\top A_{n-1} \tilde{U} = \text{diag}(d_1, \dots, d_{n-1})$  mit  $d_1, \dots, d_{n-1} \in \mathbb{R}$ . Da  $q(x_1, \dots, x_{n-1}, 0)$  positiv definit ist und durch die Matrix  $A_{n-1}$  darstellbar ist, folgt  $d_1, \dots, d_{n-1} > 0$ .

Wir führen die Matrix  $U := \begin{pmatrix} \tilde{U} & 0 \\ 0 & 1 \end{pmatrix} \in \mathbb{R}^{n \times n}$  ein. Die Matrix  $U$  benutzen wir für einen Koordinatenwechsel. Konkret bedeutet es, dass wir die Matrix  $R := U^\top AU$  betrachten. Die Matrix  $R$  hat einen Diagonalblock der Größe  $(n - 1) \times (n - 1)$ . Genauer sieht die Struktur von  $R$

so aus:

$$\begin{aligned}
R &= \begin{pmatrix} \tilde{U}^\top & 0 \\ \vdots & \\ 0 & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix} \begin{pmatrix} & & & a_{1,n} \\ & A_{n-1} & & \vdots \\ & & & a_{n-1,n} \\ a_{n,1} & \cdots & a_{n,n-1} & a_{n,n} \end{pmatrix} \begin{pmatrix} \tilde{U} & 0 \\ \vdots & \\ 0 & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix} \\
&= \begin{pmatrix} \tilde{U}^\top A_{n-1} \tilde{U} & * \\ \vdots & * \\ * & \cdots & * & a_{nn} \end{pmatrix} \\
&= \begin{pmatrix} \tilde{D} & * \\ \vdots & * \\ * & \cdots & * & a_{nn} \end{pmatrix} \\
&= \begin{pmatrix} d_1 & & c_1 \\ & \ddots & \vdots \\ & & d_{n-1} & c_{n-1} \\ c_1 & \cdots & c_{n-1} & a_{nn} \end{pmatrix}
\end{aligned}$$

mit gewissen  $c_1, \dots, c_{n-1} \in \mathbb{R}$ . Um zu zeigen, dass  $q$  positiv ist, reicht es zu zeigen, dass die quadratische Form  $r(x) = x^\top Rx$  positiv definit ist. Um das zu sehen, diagonalisieren wir  $r$ . Aus der Darstellung von  $R$  folgt

$$\begin{aligned} r(x_1, \dots, x_n) &= (d_1 x_1^2 + 2c_1 x_1 x_n) + \dots + (d_{n-1} x_{n-1}^2 + 2c_{n-1} x_{n-1} x_n) + a_{nn} x_n^2 \\ &= \sum_{i=1}^{n-1} (d_i x_i^2 + 2c_i x_i x_n) + a_{nn} x_n^2 \end{aligned}$$

Wir haben  $n-1$  Summanden, und der  $i$ -te Summende ist eine quadratische Funktion in  $x_i$ . Wir machen für diese Funktion die quadratische Ergänzung und erhalten die Darstellung

$$r(x_1, \dots, x_n) = \sum_{i=1}^{n-1} d_i \left( (x_i + \frac{c_i}{d_i} x_n)^2 - \left( \frac{c_i}{d_i} x_n \right)^2 \right) + a_{nn} x_n^2$$

Im vorigen Ausdruck hat man viele Terme der Form Konstante mal  $x_n^2$ . Wir fassen all diese Terme zusammen und erhalten die übersicht-

liche Darstellung

$$r(x_1, \dots, x_n) = \sum_{i=1}^{n-1} d_i \left( x_i + \frac{c_i}{d_i} x_n \right)^2 + \left( a_{nn} - \sum_{i=1}^{n-1} \frac{c_i^2}{d_i} \right) x_n^2. \quad (7.3.1)$$

Das vorige ist die Gewünschte Diagonalisierung. Die Koeffizienten  $d_1, \dots, d_{n-1}$  sind positiv. Wir sollen also nur noch herausfinden, was der Koeffizient von  $x_n^2$  ist. Dafür werden wir die Determinante von  $R$  in Abhängigkeit vom Koeffizienten von  $x_n^2$  darstellen. Es sei bemerkt, dass  $\det(R) = \det(U^\top AU) = \det(A)$  gilt. Laut unserer Voraussetzung gilt  $\det(A) = \det(A_n) > 0$ . Das heißt  $\det(R) > 0$ . Andererseits können wir nun eine Formel für  $\det(R)$  mit der Verwendung der elementaren Zeilentransformationen ausrechnen. Als Erstes ziehen wir aus der

$i$ -ten Zeile den Faktor  $d_i$  heraus, für  $i \in \{1, \dots, n-1\}$ , und erhalten:

$$\begin{aligned} \det(R) &= \det \begin{pmatrix} d_1 & & c_1 \\ & \ddots & \vdots \\ & & d_{n-1} & c_{n-1} \\ c_1 & \cdots & c_{n-1} & a_{nn} \end{pmatrix} \\ &= d_1 \cdots d_{n-1} \det \begin{pmatrix} 1 & & c_1/d_1 \\ & \ddots & \vdots \\ & & 1 & c_{n-1}/d_{n-1} \\ c_1 & \cdots & c_{n-1} & a_{nn} \end{pmatrix} \end{aligned}$$

Um zur oberen Dreiecksstruktur zu kommen, müssen wir die Koeffizienten  $c_1, \dots, c_{n-1}$  in der letzten Zeile entfernen. Dafür wird von der letzten Zeile eine passende Linearkombination der restlichen Zeilen

abgezogen. Wir erhalten

$$\begin{aligned}\det(R) &= d_1 \cdots d_{n-1} \det \begin{pmatrix} 1 & & c_1/d_1 \\ \ddots & & \vdots \\ & 1 & c_{n-1}/d_{n-1} \\ 0 & \cdots & 0 & a_{nn} - \sum_{i=1}^{n-1} c_i^2/d_i \end{pmatrix} \\ &= \underbrace{d_1 \cdots d_{n-1}}_{>0} \left( a_{nn} - \sum_{i=1}^{n-1} \frac{c_i^2}{d_i} \right).\end{aligned}$$

Wir sehen also das der positive Wert  $\det(R)$  das Produkt der positiven Werte  $d_1, \dots, d_{n-1}$  und des Koeffizienten  $a_{nn} - \sum_{i=1}^{n-1} \frac{c_i^2}{d_i}$ . Es folgt

$$a_{nn} - \sum_{i=1}^{n-1} \frac{c_i^2}{d_i} > 0.$$

Aus (7.3.1) folgt nun, dass die quadratische Form  $r$  positive semidefinit ist. Es bleibt zu zeigen, dass  $r$  sogar positiv definit ist. Ist

$r(x_1, \dots, x_n) = 0$  so hat man  $x_i + \frac{c_i}{d_i}x_n = 0$  für alle  $i \in \{1, \dots, n-1\}$  und  $x_n = 0$ . Daraus folgt offensichtlich  $x_1 = \dots = x_n = 0$ . Also ist  $r(x) > 0$  für alle  $x \in \mathbb{R}^n \setminus \{0\}$ , was den Beweis abschließt.  $\square$

Man beachte an dieser Stelle, dass es um das Hauptminoren-Kriterium ausschließlich um die Definitheit geht. Man kann nämlich die Semidefiniteit anhand von Hauptminoren nicht entscheiden. Betrachten wir zum Beispiel die folgenden Matrizen

$$\begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Die quadratische Form der ersten Matrix ist positiv semidefinit, die quadratische Form der zweiten Matrix ist negativ semidefinit, und die quadratische Form der dritten Matrix ist indefinit. Diese Unterschiede lassen sich an den Hauptminoren nicht erkennen, denn bei jeder dieser drei Matrizen sind alle drei Hauptminoren gleich 0.

Semidefinitheit kann aber trotzdem mit Minoren charakterisiert werden, man braucht aber mehr Minoren (nicht nur die Hauptminoren).

### 7.3.9 Signatur

**Thm.** Sei  $A \in \mathbb{R}^{n \times n}$  symmetrische Matrix mit  $s$  positiven und  $t$  negativen Eigenwerten (mit Berücksichtigung der Vielfachheit). Sei  $U \in \mathbb{R}^{n \times n}$  reguläre Matrix, für welche  $D = U^\top AU$  Diagonalmatrix ist. Dann hat  $D$  genau  $s$  positive und genau  $t$  negative Diagonalelemente.

*Beweis.* Seien  $u_1, \dots, u_n$  die Spalten von  $U$ , seien  $d_1, \dots, d_n$  die Diagonalelemente von  $D$ . Seien oBdA  $d_1 > 0, \dots, d_i > 0$  und  $d_{i+1} \leq 0, \dots, d_n \leq 0$ . Wir betrachten die Vektorräume  $V := \text{lin}(u_1, \dots, u_i)$  und  $W := \text{lin}(u_{i+1}, \dots, u_n)$  und die quadratische Form  $q(x) = x^\top Ax$  zu  $A$ .

Es ist klar, dass die Einschränkung  $q|_V$  von  $q$  auf  $V$  positiv definit ist. Analog ist die Einschränkung  $q|_W$  von  $q$  auf  $W$  negativ semidefinit.

Seien nun  $v_1, \dots, v_n$  eine Basis aus Eigenvektoren von  $A$ , wobei  $v_1, \dots, v_s$  die Eigenvektoren zu positiven Eigenwerten von  $A$  sind und  $v_{s+1}, \dots, v_n$  die Eigenvektoren zu nichtnegativen Eigenwerten. Wir setzen  $V' := \text{lin}(v_1, \dots, v_s)$  und  $W' := \text{lin}(v_{s+1}, \dots, v_{s+n})$ . Man hat:  $q|_{V'}$  ist positiv definit,  $q|_{W'}$  ist negativ semidefinit.

Die Vektorräume  $V$  und  $W'$  schneiden sich nur in 0. Es folgt  $n \geq \dim(V \oplus W') = \dim(V) + \dim(W') = i + n - s$ .

Die Vektorräume  $V'$  und  $W$  scheiden sich auch nur in 0. Es folgt  $n \geq \dim(V' \oplus W) = \dim(V') + \dim(W) = s + n - i$ .

Wir erhalten also  $i = s$ , das heißt,  $D$  hat genau  $s$  positive Diagonalelemente. Analog kann auch gezeigt werden, dass  $D$  genau  $t$  negative Diagonalelemente hat.  $\square$

Die Information darüber, wie viel positive und wie viel negative Eigenwerte (mit Berücksichtigung der Vielfachheit) eine symmetrische Matrix  $A \in \mathbb{R}^{n \times n}$  hat, nennt man die *Signatur* der Matrix  $A$ . Man spricht auch

von der Signatur der quadratischen Form  $q(x) = x^\top Ax$ .

## 8 Ganzzahlige lineare Algebra

Wir haben uns bis jetzt mit der linearen Algebra über einem Körper  $\mathbb{K}$  beschäftigt. Der Stoff, der in diesem Kapitel präsentiert wird, kann man als Lineare Algebra über dem Ring  $\mathbb{Z}$  der ganzen Zahlen auffassen. Lineare Algebra über  $\mathbb{Z}$  hat gewisse Ähnlichkeiten zur linearen Algebra über  $\mathbb{K}$ , da man aber bei der ganzzahligen linearen Algebra einen Ring  $\mathbb{Z}$  zugrunde legt, und keinen Körper, hat man auch Unterschiede.

### 8.1 Lösung einer diophantischen linearen Gleichung

Ein System bzw. Gleichung heißt diopantsch, wenn die Unbekannten dieser Gleichung ganzzahlig sind. Wir beschäftigen uns zuerst mit einer diophant-

schen linearen Gleichung. Bereits für diesen Fall gibt es sehr interessante Anwendungen.

### 8.1.1 Unimodulare Elementartransformationen

Für lineare Gleichungssysteme, Vektorsysteme, Spalten sowie Zeilen einer Matrix haben wir Elementartransformationen vom Typ 1,2 und 3 eingeführt. Nun führen wir unimodulare Elementartransformationen vom Typ 1,2 und 3. Wir beschränken uns auf Systeme von Vektoren. Seien  $a_1, \dots, a_k \in \mathbb{Z}^n$ . Die unimodularen Elementartransformationen sind wie folgt definiert:

Typ 1: Vertauschen  $a_i \leftrightarrow a_j$  von zwei Vektoren  $a_i$  und  $a_j$  mit  $1 \leq i, j \leq k$ .

Typ 2: Ersetzen von  $a_i$  durch  $-a_j$ , das heißt  $a_j := -a_i$  mit  $1 \leq j \leq k$ .

Typ 3: Addieren eines ganzzahligen Vielfachen eines Vektors zu einem anderen Vektor:  $a_j := a_j + \beta a_i$  mit  $\beta \in \mathbb{Z}$ ,  $1 \leq i, j \leq k$  und  $i \neq j$ .

Jeder dieser drei Transformationen kann mit Hilfe der Matrix  $A = (a_1, \dots, a_k) \in \mathbb{Z}^{n \times k}$  als die Operationen  $A \mapsto AE$  beschrieben werden, wobei  $E$  abhängig vom Typ die folgenden Matrizen ist:

Für Typ 1:  $E = I - (e_i + e_j)(e_i + e_j)^\top$  für  $i \neq j$  und  $E = I$  für  $i = j$ .

Für Typ 2:  $E = I - 2e_i e_i^\top$

Für Typ 3:  $E = I + \beta e_i e_j^\top$ .

Die jeweiligen Matrizen nennen wir unimodulare elementare Matrizen vom Type 1, 2 und 3. Hierbei merken wir, dass die inversen der unimodularen elementaren Transformationen ebenfalls unimodulare elementare Transformationen sind. Das kann man zum Beispiel an den jeweiligen Matrizen verfolgen. Die Matrix  $E$  beim Typ 1 und 2 ist selbst-inverse. Die Inverse zur Matrix  $E = I + \beta e_i e_j^\top$  im Typ 3 ist  $E^{-1} = I - \beta e_i e_j^\top$ . Des Weiteren merken wir, dass unabhängig vom Typ der Transformation die jeweilige Matrix

$E$  eine ganzzahlige Matrix ist, deren Determinanten betragsmäßig gleich eins ist. Solche Matrizen werden wir im folgenden genauer untersuchen.

### 8.1.2 Der größte gemeinsame Teiler von zwei Zahlen

Wir betrachten  $v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} \in \mathbb{Z}^2 \setminus \{0\}$ . Der größte gemeinsame Teiler  $\text{ggT}(v) = \text{ggT}(v_1, v_2)$  der eindeutige Wert  $g \in \mathbb{N}$  derart, dass jeder gemeinsame Teiler von  $v_1$  und  $v_2$  auch den Wert  $g$  teilt.

**Thm.** Für jedes  $v \in \mathbb{Z}^2 \setminus \{0\}$  gibt es ein  $u \in \mathbb{Z}^2$  mit  $u^\top v = \text{ggT}(v)$ . Darüber hinaus existiert ein Algorithmus, der für ein gegebenes  $v$  den Vektor  $u$  und  $\text{ggT}(v)$  berechnet.

*Beweis.* Wir zeigen zunächst, dass ein Algorithmus zur Berechnung von  $\text{ggT}(v)$  existiert. Anschließend ergänzen wir den Algorithmus, zu einem Algorithmus, der noch zusätzlich den Vektor  $u$  berechnet.

Anwendung unimodularer Elementartransformationen zur Komponenten von  $v = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix}$  ändern den größten gemeinsame Teiler nicht. Das ist klar für Typ 1, weil die Transformation vom Typ 1 lediglich  $v_1$  und  $v_2$  vertauscht. Für Typ 2 ist es ebenfalls klar, weil der größte gemeinsame Teiler nicht von den Vorzeichen von  $v_1$  und  $v_2$  abhängig ist. Beim Typ 3 müssen wir  $\text{ggT}(v_1, v_2 + \alpha v_1) = \text{ggT}(v_1, v_2)$  für  $\alpha \in \mathbb{Z}$  nachweisen. Jede natürliche Zahl, die  $v_1$  und  $v_2$  teilt, teilt auch  $v_1$  und  $v_2 + \alpha v_1$ . Umgekehrt: jede natürliche Zahl, die  $v_1$  und  $v_2 + \alpha v_1$  teilt, teilt auch  $v_1$  und  $v_2 = (v_2 + \alpha v_1) - \alpha v_1$ . Also ändern die unimodularen Elementartransformationen den größten gemeinsamen Teiler nicht.

Unser Algorithmus zur Berechnung von  $\text{ggT}(v)$  transformiert die Komponenten von  $v$  mit unimodularen Elementartransfomrationen solange, bis eine der Komponenten gleich 0 wird.

Nach einer geeigneten Anwendung der Transformationen vom Typ 1 und

$2$  können wir  $0 \leq v_1 \leq v_2$  erzwingen. Ist  $v_1 = 0$ , so wissen wir dass  $\text{ggT}(v) = \text{ggT}(0, v_2) = v_2$  ist. Ansonsten teilen wir  $v_2$  durch  $v_1$  mit Rest:  $v_2 = qv_1 + r$  mit  $q, r \in \mathbb{Z}$ ,  $q \geq 0$  und  $0 \leq r < v_1$ . Durch die Transformation  $v_2 := v_2 - qv_1$  und anschließendes Vertauschen von  $v_1$  und  $v_2$  wird der Wert von  $v_1$  geringer. Nach endlich vielen Runden werden wir also  $v_1 = 0$  erzwingen können. Da man für  $v_2 \in \mathbb{N}$  offensichtlich  $\text{ggT}(0, v_2) = v_2$ , haben wir also den größten gemeinsamen Teiler ausgerechnet.

Um einen geeigneten Vektor zu bestimmen, können wir den obigen Prozess folgendermaßen ergänzen. Wir können den Prozess als Veränderungen eines linearen Gleichungssystems interpretieren. Man startet mit einem trivialen System  $Ix = v$ , deren linke Seite durch die Einheitsmatrix  $I \in \mathbb{Z}^{2 \times 2}$  bestimmt ist und transformiert dieses System mit unimodularen Elementartransformationen solange, bis auf der rechten Seite einer der Koeffizienten gleich  $0$  ist. Das entspricht der iterativen Anwendung der unimodularen Elementaren Matrizen  $E_1, \dots, E_t$  welche das Originalsystem  $Ix = v$  zum

System  $E_t \cdots E_1 x = E_t \cdots E_1 v$  überführen, bei dem  $E_t \cdots E_1 v = \begin{pmatrix} 0 \\ g \end{pmatrix}$  für ein  $g \in \mathbb{N}$  erfüllt ist. Sei  $U = E_t \cdots E_1$ . Dann hat das transformierte System die Form

$$Ux = \begin{pmatrix} 0 \\ g \end{pmatrix}$$

Da wir mit dem System  $Ix = v$  gestartet haben, wissen wir, dass  $x = v$  die Lösung des transformierten Systems ist. Wir können also als  $u^\top$  die zweite Zeile der Matrix  $U$  festlegen.  $u^\top v = \text{ggT}(v)$ .  $\square$

**Bem.** Der Algorithmus zur Bestimmung von  $\text{ggT}(v)$  aus dem vorigen Beweis heißt der Euklidische Algorithmus. Der Algorithmus zur Bestimmung von  $u$  und  $\text{ggT}(v)$  aus dem vorigen Theorem heißt der erweiterte Euklidische Algorithmus.

**Bsp.** Ein Beispiel zum erweiterten Euklidischen Algorithmus.  $\text{ggT}(119, 170)$ .

	$x_1$	$x_2$	
(1)	1	0	119
(2)	0	1	170

Teilen mit Rest:  $170 = 1 \cdot 119 + 51$ . Daher  $(2) := (2) - (1)$ .

	$x_1$	$x_2$	
(1)	1	0	119
(2)	-1	1	51

Teilen mit Rest  $119 = 2 \cdot 51 + 17$ . Daher  $(1) := (1) - (2)$ .

	$x_1$	$x_2$	
(1)	3	-2	17
(2)	-1	1	51

Teilen mit Rest  $51 = 3 \cdot 17 + 0$ . Daher  $(2) := (2) - 3(1)$ .

	$x_1$	$x_2$	
(1)	3	-2	17
(2)	-10	7	0

Das bedeutet  $\text{ggT}(119, 170) = 17$  und  $3 \cdot 119 - 2 \cdot 170 = 17$ .

### 8.1.3 Der Chinesische Restsatz für zwei Restklassenringe

In diesem Abschnitt arbeiten wir mit verschiedenen Restklassenringen Gleichzeitig. Wir bezeichnen als  $[x]_m$  die Restklasse von  $x \in \mathbb{Z}$  im Ring  $\mathbb{Z}/m\mathbb{Z}$  mit  $m \in \mathbb{N}$ .

**Bem** (Kongruenzen). Seien  $q_1, q_2 \in \mathbb{N}$  Zahlen mit  $\text{ggT}(q_1, q_2) = 1$  und  $b_1, b_2 \in \mathbb{Z}$ . Wir bestimmen die Lösungsmenge des folgenden Systems der Kongruenzen

$$\begin{aligned} x &\equiv b_1 \pmod{q_1}, \\ x &\equiv b_2 \pmod{q_2}. \end{aligned}$$

mit einem unbekannten  $x \in \mathbb{Z}$ . Ein solches Kongruenzensystem behandelt man zum Beispiel im Rahmen des RSA-Verfahrens zur Verschlüsselung von Daten.

Mit dem erweiterten Euklidischen Algorithmus können  $a_1, a_2 \in \mathbb{Z}$  mit  $1 = a_1q_1 + a_2q_2$  bestimmt werden. Da  $a_1q_1$  kongruent zu 1 modulo  $q_2$  ist und  $a_2q_2$  kongruent zu 1 modulo  $q_1$ , sehen wir dass  $b_1a_2q_2 + b_2a_1q_1$  eine Lösung unseres Kongruenzensystems ist. Man erhält alle Lösungen, wenn man zu dieser Lösungen alle möglichen Lösungen des homogenen Kongruenzensystems

$$\begin{aligned}x &\equiv 0 \pmod{q_1}, \\x &\equiv 0 \pmod{q_2}.\end{aligned}$$

Beim homogenen Kongruenzensystem handelt es sich um die Zahlen  $x$ , die durch  $q_1$  und durch  $q_2$  teilbar ist. Da  $q_1$  und  $q_2$  teilerfremd sind, sind genau die Zahlen die durch das Produkt  $q_1q_2$  teilbar sind. Die Lösungsmenge

unseres System der Kongruenzen ist somit

$$\{b_1a_2q_2 + b_2a_1q_1 + zq_1q_2 : z \in \mathbb{Z}\}$$

**Bem.** Das Kongruenzensystem aus der vorigen Bemerkung ist vollständig gelöst, aber wir können aber noch über die Struktur der vorigen Beschreibung nachdenken. Wir sehen ist die Lösungsmenge eine Äquivalenzklasse aus dem Restklassenring  $\mathbb{Z}/q_1q_2\mathbb{Z}$  ist. Anderseits spielt die Angabe von  $b_i$  nur modulo  $q_i$  eine Rolle, weil das addieren zu  $b_i$  eines Vielfachen von  $q_i$  die entsprechende Kongruenz nicht ändert. Unser Beschreibung der Lösungsmenge des Kongruenzensystems zeigt also, dass die Abbildung

$$[x]_{q_1q_2} \mapsto ([b_1]_{q_1}, [b_2]_{q_2}) = ([x]_{q_1}, [x]_{q_2}).$$

von  $\mathbb{Z}/q_1q_2\mathbb{Z}$  nach  $\mathbb{Z}/q_1 \times \mathbb{Z}/q_2$  eine Bijektion ist, wobei hier  $[x]_m$  für die Restklasse von  $x$  in  $\mathbb{Z}/m\mathbb{Z}$  steht.

Diese Bijektion ist ein sogenannter Ringsomorphismus. Denn  $\mathbb{Z}/q_1q_2\mathbb{Z}$  ist ein Ring, und  $\mathbb{Z}/q_1\mathbb{Z} \times \mathbb{Z}/q_2\mathbb{Z}$  als kartesisches Produkt von zwei Ringen

kann ebenfalls mit einer natürlichen Ringstruktur ausgestattet werden, indem man die Elemente daraus komponentenweise addiert und multipliziert. Die Abbildung  $[x]_{q_1 q_2} \mapsto ([x]_{q_1}, [x]_{q_2})$ , ist ein sogenannter Ringhomomorphismus, denn sie erhält die Ringstruktur (die Null, die Eins, die Addition und die Multiplikation). Die bijektiven Ringhomomorphismen nennt man Ringsomorphismen.

Für konkrete Berechnungen hat der oben aufgestellte Ringsomorphismus folgende Auswirkungen. Angenommen, wir möchte ein System  $Ax = b$  mit  $A \in \mathbb{Z}^{m \times n}$  und  $b \in \mathbb{Z}^m$  modulo  $q_1 q_2$  lösen, und seien  $q_1$  und  $q_2$  Primzahlen. Dann können wir dieses System einmal bzgl. des Körpers  $\mathbb{Z}/q_1\mathbb{Z}$  lösen, und einmal bzgl. des Körpers  $\mathbb{Z}/q_2\mathbb{Z}$  und dann aus den so ermittelten Lösungen in den beiden Fällen, die Lösung in  $\mathbb{Z}/q_1 q_2\mathbb{Z}$  berechnen.

**Thm** (Der chinesische Restsatz für zwei Restklassenringe). *Seien  $q_1, q_2 \in \mathbb{N}$  teilerfremd. Dann ist die Abbildung*

$$[x]_{q_1 q_2} \mapsto ([x]_{q_1}, [x]_{q_2})$$

ein Ringisomorphismus von  $\mathbb{Z}/q_1q_2\mathbb{Z}$  nach  $\mathbb{Z}/q_1\mathbb{Z} \times \mathbb{Z}/q_2\mathbb{Z}$ .

#### 8.1.4 Eine Anwendung: die RSA-Verschlüsselung

**Thm** (Der kleine Satz von Fermat). *Sei  $p$  Primzahl. Im Körper  $\mathbb{K} = \mathbb{Z}/p\mathbb{Z}$  gilt  $a^{p-1} = 1$  für alle  $a \in \mathbb{K} \setminus \{0\}$ .*

*Beweis.* Wir benutzen die Bijektion  $x \in \mathbb{K} \setminus \{0\} \leftrightarrow y \in \mathbb{K} \setminus \{0\}$  mit  $y = ax$  und  $x = a^{-1}x$ .

Das Produkt aller Elemente aus  $\mathbb{K} \setminus \{0\}$  ist somit einerseits  $u := \prod_{x \in \mathbb{K} \setminus \{0\}} x$  und andererseits  $\prod_{x \in \mathbb{K} \setminus \{0\}} (ax) = a^{p-1} \prod_{x \in \mathbb{K} \setminus \{0\}} x = a^{p-1}u$ .

Die Gleichung  $a^{p-1}u = u$  kann durch  $u$  geteilt werden, weil  $u$  als Produkt von Nichtnullelementen von  $\mathbb{K}$  ungleich 0 ist. Das ergibt  $a^{p-1} = 1$ .  $\square$

Das RSA-Kryptosystem (die etwa im SSH-Protokoll benutzt wird) basiert auf zwei verschiedenen Primzahlen  $p$  und  $q$ , die pseudozufällig generiert werden, und der Öffentlichkeit unbekannt sind, und einem Wert  $e \in \mathbb{N}$

mit  $\text{ggT}(e, (p - 1)(q - 1)) = 1$ . Wie man  $p, q$  und  $e$  genau generiert, ist wichtig, wird aber hier nicht im Detail diskutiert. Was wir lediglich dazu bemerken ist Folgendes: man kann effizient testen, ob eine gegebenen natürliche Zahl eine Primzahl ist, und die Primzahlen sind dicht genug innerhalb der natürlichen Zahlen verteilt, sodass man beim sukzessiven Durchsuchen ungerader natürlicher Zahlen schnell genug auf eine Primzahl stößt.

Der Wert  $N = pq$  ist der sogenannte Modul des Kryptosystems. Das Paar  $(e, N)$  nennt man den öffentlichen Schlüssel. Der private Schlüssel  $(d, N)$  wird generiert, indem man anhand  $e$  einen Wert  $d \in \mathbb{N}$  mit  $ed \equiv 1 \pmod{(p - 1)(q - 1)}$ . Da  $e$  und  $(p - 1)(q - 1)$  teilerfremd sind, gibt es einen solchen Wert, und dieser Wert kann man mit Hilfe des erweiterten Euklidischen Algorithmus generieren.

Nachdem  $e$  und  $d$  generiert sind, kann die verschlüsselte Kommunikation zwischen zwei Teilnehmenden, die man standardmäßig Alice und Bob nennt, wie folgt aufgebaut werden.

Alice behält  $(d, N)$  als privaten Schlüssel und veröffentlicht  $(e, N)$  als den sogenannten öffentlichen Schlüssel. Wenn Bob eine verschlüsselte Nachricht  $m \in \mathbb{Z}/N\mathbb{Z}$  an Alice schickt, so schickt er  $c = m^e$ . Beim Empfang von  $c$  berechnet Alice  $c^d \in \mathbb{Z}/N\mathbb{Z}$  mit Hilfe des privaten Schlüssels  $(d, N)$  und erhält auf diese Weise den Wert  $c^d = (m^e)^d = m^{ed}$ . Es stellt sich heraus, dass  $m^{ed} = m$  gilt. Das heißt, Alice hat die Nachricht  $m$  von Bob erhalten.

Dass  $m^{ed} = m$  gilt, ist die Grundlage des RSA-Kryptosystems.

**Prop.** Seien  $p, q$  Primzahlen mit  $p \neq q$  und  $e, d \in \mathbb{N}$  Werte mit  $ed \equiv 1 \pmod{N}$  und  $N := pq$ . Dann gilt  $m^{ed} = m$  für alle  $m \in \mathbb{Z}/N\mathbb{Z}$ .

*Beweis.* Nach dem Chinesischen Restsatz ist der Ring  $\mathbb{Z}/N\mathbb{Z}$  isomorph zum Ring  $\mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/q\mathbb{Z}$ . Wir können also durch den Isomorphismus aus dem Chinesischen Restsatz die Nachricht  $m \in \mathbb{Z}/N\mathbb{Z}$  als ein Paar  $(a, b) \in \mathbb{Z}/p\mathbb{Z} \times \mathbb{Z}/q\mathbb{Z}$  darstellen. So reduziert sich die Gleichung  $m^{ed} = m$  zu zwei Gleichungen  $a^{ed} = a$  mit  $a \in \mathbb{Z}/p\mathbb{Z}$  und  $b^{ed} = b$  mit  $b \in \mathbb{Z}/q\mathbb{Z}$ . Etwas weniger abstrakt lässt sich der Übergang von  $m$  zu  $a$  und  $b$  so erklären:

$m$  ist die Restklasse einer ganzen Zahl  $z \in \mathbb{Z}$  modulo  $N$ . Die Bedingung  $m^{ed} = m$  kann dann als die Teilbarkeit von  $z^{ed} - z$  durch  $N = pq$  formuliert werden. Das Letztere ist Äquivalent zur Teilbarkeit von  $z^{ed} - z$  durch  $p$  und durch  $q$ . Wir können also als  $a$  die Restklasse von  $z$  modulo  $p$  und als  $b$  die Restklasse von  $z$  modulo  $q$  fixieren.

Wegen  $ed \equiv 1 \pmod{(p-1)(q-1)}$  gilt  $ed = 1 + k(p-1)(q-1)$  für ein  $k \in \mathbb{Z}_{\geq 0}$ . Die Gleichungen für  $a$  und  $b$  können also als

$$\begin{aligned} a^{1+k(p-1)(q-1)} &= a, \\ b^{1+k(p-1)(q-1)} &= b. \end{aligned}$$

umformuliert werden. Die beiden Gleichungen sind symmetrisch bis auf das Vertauschen der Rollen von  $p$  und  $q$ . Es reicht also eine davon zu zeigen. Ist  $a = 0$ , so ist die Gleichung für  $a$  trivial. Im Fall  $a \in \mathbb{K} \setminus \{0\}$  erhalten wir aus dem kleinen Satz von Fermat

$$a^{1+k(p-1)(q-1)} = a \cdot (a^{p-1})^{k(q-1)} = a \cdot 1^{k(q-1)} = a.$$

□

**Bem** (Elektronische Signaturen mit dem RSA-Kryptosystem). Um eine Nachricht zu signieren, benutzt man eine öffentlich verfügbare Funktion  $h$  (eine sogenannte Hashfunktion), welche eine Nachricht  $M$  einer beliebigen Bitlänge, die signiert wird, auf einen Wert  $m = h(M) \in \mathbb{Z}/N\mathbb{Z}$  abbildet (der Hashwert hat keine beliebige Bitlänge, weil  $N$  festgelegt ist).

Alice schickt eine signierte Nachricht an Bob wie folgt. Sie signiert ihre Nachricht  $M$  mit  $s = m^d$  mit Hilfe ihres Privatschlüssels. Bob erhält die Nachricht  $M$  mit der Signatur  $s$ . Um die Echtheit der Nachricht zu überprüfen, benutzt Bob den öffentlichen Schlüssel  $(e, N)$  von Alice. Bob berechnet zuerst  $s^e$ . Wie oben gezeigt ist  $s^e = (m^d)^e = m^{ed} = m$ . Bob erhält also den Hashwert der Nachricht  $M$ . Zur Prüfung der Nachricht kann nun der Wert  $h(M)$  mit  $m$  zu verglichen werden.

**Bsp.** Ein reales Beispiel eines öffentlichen Schlüssels. Der Modul im Dezimalsystem ist eine 617-stellige Zahl:

$$N = 316944941931319194507627067708094688403544346666724240087815 \\ 735154316257460965852803682506212526317946314538155651204243 \\ 340980505298750717927449356077755915943710216498246601390424 \\ 37373082830123871741534477795846680213349536805990124654019 \\ 474582260883704618655528914785850965434061131819506418338731 \\ 293581466779781458757319881728466979661277553474903218087349 \\ 304497929069589881438700538136624963724757910769077994156998 \\ 684944501323144019033703886630577589883833146072520896676339 \\ 508364526110987991992811450967808280730177742672819293812614 \\ 151991755996613548301019856966248512826059816155818536279614 \\ 88141725060004451$$

Exponent im Dezimalsystem:  $e = 65\,537$ .

Um diese Verschlüsselung zu brechen, würde man das  $N$  als Produkt von zwei Primzahlen darstellen müssen:  $N = pq$ . Aus  $p$  und  $q$  kann man dann

$(p - 1)(q - 1)$  berechnen und dann den privaten Schlüssel  $d$  mit  $ed \equiv 1 \pmod{(p - 1)(q - 1)}$  mit Hilfe des erweiterten Euklidischen Algorithmus ausrechnen.

### 8.1.5 Eine diophantsche Gleichung in 2 Variablen

**Thm.** Sei  $a \in \mathbb{Z}^2 \setminus \{0\}$  und  $b \in \mathbb{Z}$ . Die Gleichung  $a^\top x = b$  hat genau dann eine ganzzahlige Lösung  $x \in \mathbb{Z}^2$ , wenn  $b$  durch  $\text{ggT}(a)$  teilbar ist. Darüber hinaus existiert ein Algorithmus, der für die gegebene Gleichung  $a^\top x = b$  entscheidet, ob diese Gleichung eine ganzzahlige Lösung besitzt und ggf. eine solche Lösung bestimmt.

*Beweis.* Für jedes  $x \in \mathbb{Z}^2$  ist  $a^\top x$  durch  $\text{ggT}(a)$  teilbar. Die Teilbarkeit von  $b$  durch  $\text{ggT}(a)$  ist somit notwendig für die Existenz einer ganzzahligen Lösung  $x \in \mathbb{Z}^2$ . Nach Theorem 8.1.2 kann ein  $u \in \mathbb{Z}^2$  mit  $a^\top u = \text{ggT}(a)$  mit Hilfe des erweiterten Euklidischen Algorithmus berechnet werden. Ist

$b$  durch  $\text{ggT}(a)$  teilbar, dann ist  $x = \frac{b}{\text{ggT}(a)}u$  eine ganzzahlige Lösung des Systems  $a^\top x = b$ .  $\square$

**Bem.** Überlegen Sie sich, wie man alle ganzzahlige Lösungen einer solchen Gleichung  $a^\top x = b$  bestimmen kann.

**Bsp.** IM AUFBAU: Beispiel aus [8.1.2] fortsetzen.

### 8.1.6 Eine diophantsche Gleichung in $n$ Variablen

Für  $v \in \mathbb{Z}^n \setminus \{0\}$  ist der  $\text{ggT}(v)$  der Wert  $g \in \mathbb{N}$  mit der Eigenschaft, dass jeder Teiler der Komponenten von  $v$  auch den Wert  $g$  teilt.

**Thm.** Für jedes  $v \in \mathbb{Z}^n \setminus \{0\}$  gibt es ein  $u \in \mathbb{Z}^n$  mit  $u^\top v = \text{ggT}(v)$ . Darüber hinaus existiert ein Algorithmus, der für ein gegebenes  $v$  einen Vektor  $u$  und  $\text{ggT}(v)$  berechnet.

*Beweis.* Der Fall  $n = 1$  ist trivial. Wir müssen unseren Beweis des Theorems [8.1.2] vom Fall  $n = 2$  auf das allgemeine  $n \geq 2$  erweitern. Genau so wie

im Fall  $n = 2$  sehen, wir dass die unimodularen Elementartransformationen der Komponenten von  $v$  den Wert  $\text{ggT}(v)$  nicht ändern. Wir wollen mit solchen Transformationen den Vektor  $v$  zum Vektor der Form  $(g, 0, \dots, 0)$  mit  $g \in \mathbb{N}$  iterativ überführen. Dafür kann man zum Beispiel folgendermaßen vorgehen: Hat der Vektor  $v$  genau eine Nichtnull-Komponente so kann man erzwingen, dass diese Komponente positiv ist und dann ist man fertig. Hat  $v$  zwei Nichtnullkomponenten  $v_i$  und  $v_j$ , so kann man auf diesen Komponenten den Euklidischen Algorithmus für zwei Werte laufen lassen, bis eine der beiden Komponenten gleich 0 wird. Auf diese Weise löscht man iterativ die Nichtnullkomponenten aus, bis am Ende nur eine Nichtnullkomponente übrig bleibt. Um aus diesem Verfahren noch zusätzlich den Vektor  $u$  auszurechnen, kann man genauso wie im Fall  $n = 2$  mit dem trivialen System  $Ix = v$  beginnen un dieses System iterativ mit unimodularen Elementar-

transformationen verändern, bis man am Ende zum System der Form

$$Ux = \begin{pmatrix} g \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

mit  $U \in \mathbb{Z}^{n \times n}$  gelangt. Es ist klar, dass wir  $u^\top$  als erste Zeile der Matrix  $U$  fixieren können.  $\square$

**Kor.** Sei  $a \in \mathbb{Z}^n \setminus \{0\}$  und  $b \in \mathbb{Z}$ . Die Gleichung  $a^\top x = b$  besitzt genau dann eine ganzzahlige Lösung  $x \in \mathbb{Z}^n$ , wenn  $b$  durch  $\text{ggT}(a)$  teilbar ist. Darüberhinaus existiert ein Algorithmus, der entscheidet ob  $a^\top x = b$  eine ganzzahlige Lösung besitzt und ggf. eine solche Lösung bestimmt.

*Beweis.* Der Beweis ist Analog zum Beweis im Fall  $n = 2$ , der bereits diskutiert worden ist. Die Teilbarkeit von  $b$  durch  $\text{ggT}(a)$  ist notwendig für

die Existenz einer ganzzahligen Lösung: ist  $x \in \mathbb{Z}^n$  und  $a^\top x = b$ , so ist  $a^\top x$  durch  $\text{ggT}(a)$  teilbar. Somit ist auch  $b$  durch  $\text{ggT}(a)$  teilbar.

Ist  $b$  durch  $\text{ggT}(a)$  teilbar, so benutzen wir den Algorithmus aus dem Beweis des vorigen Theorems, um ein  $u \in \mathbb{Z}^n$  mit  $a^\top u = \text{ggT}(a)$  zu bestimmen. Dann ist  $x = \frac{b}{\text{ggT}(a)}u$  eine ganzzahlige Lösung von  $a^\top x = b$ .  $\square$

## 8.2 Lösung von diophantschen linearen Gleichungssystemen

### 8.2.1 Hermit'sche Normalform

Man sagt, dass  $H = (h_{ij}) \in \mathbb{Z}^{m \times n}$  in einer HNF ist, wenn  $m \leq n$ ,  $h_{ii} > 0$  und  $h_{ij} \geq 0$  und für alle  $i, j$  gilt und  $h_{ij} < h_{jj}$  für alle  $i, j$  mit  $1 \leq i < j \leq m$  erfüllt ist.

**Thm.** *Es existiert ein Algorithmus, der eine gegebene ganzzahlige Matrix  $A \in \mathbb{Z}^{m \times n}$  mit dem vollen Zeilenrang mit Hilfe von unimodularen Ele-*

mentartransformationen der Splaten in eine HNF überführt. Des Weiteren existiert ein Algorithmus der Neben einer HNF  $H$  von  $A$  auch eine Matrix  $U \in \mathbb{Z}^{n \times n}$  mit  $|\det(U)| = 1$  berechnet, für welche  $H = AU$  gilt.

*Beweis.* *Phase 1:* Wir können den erweiterten Euklidischen Algorithmus benutzen, um mit Hilfe von uETS (= unimodularen Elementartransformationen) die erste Zeile von  $A$  ins Format  $(a_{11}, 0, \dots, 0)$  zu überführen. Dann befindet sich unter den Komponenten  $a_{2,2}, \dots, a_{n,2}$  mindestens eine Nicht-nullkomponente, weil sonst die zweite Zeile linear abhängig von der ersten Zeile wäre, was der Bedingung  $\text{rang}(A) = m$  widerspricht. Wir können die Spalten Nummer 2 bis  $n$  mit uET so verändern, dass nach dieser Änderung die zweite Zeile die Form  $(a_{1,2}, a_{2,2}, 0, \dots, 0)$  mit  $a_{2,2} > 0$  hat. In der  $k$ -ten

Iteration hat die Untermatrix von  $A$  aus den ersten  $k$  Zeilen die Form

$$\begin{pmatrix} a_{1,1} & 0 & \cdots & \cdots & 0 \\ a_{2,1} & a_{2,2} & 0 & \cdots & \cdots & 0 \\ \vdots & & \ddots & \ddots & & \vdots \\ a_{k,1} & \cdots & \cdots & a_{k,k} & 0 & \cdots & 0 \end{pmatrix}$$

Ist  $k = m$ , so ist die erste Phase des Verfahrens abgeschlossen. Ist  $k < m$ , dann betrachtet man die Zeile Nummer  $k + 1$  von  $A$  und verändert durch die uET der letzten  $n - k$  Spalten die Matrix so, dass  $a_{k+1,k+1} > 0$  und  $a_{k+1,j} = 0$  für  $j > k+1$  gilt. Nach endlich vielen Schritten überführt man die Matrix  $A$  in die Form:

$$A = \begin{pmatrix} a_{1,1} & 0 & \cdots & \cdots & 0 \\ a_{2,1} & a_{2,2} & 0 & \cdots & \cdots & 0 \\ \vdots & & \ddots & \ddots & & \vdots \\ a_{m,1} & \cdots & \cdots & a_{m,m} & 0 & \cdots & 0 \end{pmatrix}$$

*Phase 2:* In der zweiten Phase werden Elemente unterhalb von  $a_{1,1}, \dots, a_{m,m}$  modifiziert. Dabei arbeitet man die Spalten Nummer  $1, \dots, m-1$  in dieser Reihenfolge ab. Die Elemente  $a_{j+1,j}, \dots, a_{m,j}$  der  $j$ -ten Spalten mit  $j = 1, \dots, m-1$  werden in dieser Reihenfolge folgendermaßen modifiziert: Man teilt  $a_{i,j}$  mit  $i = j+1, \dots, m$  durch  $a_{i,i}$  mit Rest, erhält die Darstellung  $a_{i,j} = qa_{i,i} + r$  mit  $q, r \in \mathbb{Z}$  und  $0 \leq r < a_{i,i}$  und zieht von der  $j$ -ten Spalte  $q$  mal die  $i$ -te Spalte ab. Auf diese Weise wird die Bedingung  $0 \leq a_{i,j} < a_{i,i}$  sichergestellt, ohne dass die entsprechende Bedingung für die Elemente der selben Spalte oberhalb von  $a_{i,j}$  beeinträchtigt wird. Nach der Terminierung der zweiten Phase befindet sich die Matrix  $A$  in der HNF.

*Berechnung der Matrix  $U$ :* Um die Matrix  $U$  aus der Behauptung zu berechnen, kann man jede uET der Spalten von  $A$  als die Transformation  $A \mapsto AU_i$  für eine elementare unimodulare Matrix  $U_i \in \mathbb{Z}^{n \times n}$  auffassen. Während der Ausführung wird also für eine gegebene Matrix  $A$  eine Folge  $U_1, \dots, U_t$  von elementaren unimodularen Matrizen bestimmt, für die

$H = AU_1 \cdots U_t$  in einer HNF ist. Da jede Matrix  $U_i$  ganzzahlig ist, und die Determinante  $\pm 1$  hat hat die Matrix  $U = U_1 \cdots U_t$  die gewünschten Eigenschaften.  $\square$

**Bem.** Für die praktische Berechnung der Matrix  $U$  aus dem vorigen Beweis startet man mit der Matrix  $I_n \in \mathbb{Z}^{n \times n}$  und führt für diese dieselben Spaltentransformationen durch wie mit der Matrix  $A$ . Das bedeutet, dass man mit der Matrix

$$\begin{pmatrix} A \\ I_n \end{pmatrix}$$

startet, die uET der Spalten durchführt und nach der Terminierung die Matrix

$$\begin{pmatrix} H \\ U \end{pmatrix}$$

erhält, wobei  $H$  die HNF von  $A$  ist und  $U \in \mathbb{Z}^{n \times n}$  die Matrix mit  $|\det(U)| = 1$  ist, die  $A$  in die HNF  $H$  überführt.

Beispiel in SageMath:

```
A=matrix(ZZ, [(18,24,102,4)])
H,U=A.transpose().hermite_form(transformation=True)
print(H)
print(U)
print(U*A.transpose())
```

### 8.2.2 Ein Algorithmus zur Lösung von diophantschen linearen Gleichungssystemen

**Thm.** Sei  $A \in \mathbb{Z}^{m \times n}$  und sei  $b \in \mathbb{Z}^m$ . Es existieren ein Algorithmus, der verifiziert, ob das System  $Ax = b$  eine ganzzahlige Lösung  $x \in \mathbb{Z}^n$  besitzt und ggf. eine solche Lösung berechnet.

*Beweis.* Wir können zuerst verifizieren, ob  $Ax = b$  eine Lösung in  $\mathbb{R}^n$  besitzt. Ist das nicht der Fall, so hat man auch keine Lösung in  $\mathbb{Z}^n$ . Hat  $Ax = b$

eine Lösung in  $\mathbb{R}^n$  so können wir in  $Ax = b$  nur die nichtredundanten Gleichungen behalten. Dann kommt man zum Fall  $Ax = b$  mit  $\text{rang}(A) = m$ .  
 MEHR DETAILS. Man benutzt überführt die Matrix  $A$  in die HNF  $H$  und berechnet dabei noch eine Matrix  $U \in \mathbb{Z}^{n \times n}$  mit  $|\det(U)| = 1$  und  $H = AU$ . Wir führen die neuen Variablen  $y$  mit  $x = Uy$  ein und überführen dadurch  $Ax = b$  zur Gleichung  $AUy = b$ , das ist die Gleichung  $Hy = b$ . Es sei bemerkt, dass  $x \in \mathbb{Z}^n$  zu  $y \in \mathbb{Z}^n$  äquivalent ist.

Sei  $H = (T \ 0)$ , wobei  $T$  eine untere Dreiecksmatrix. Die letzte Gleichung hat eine ganzzahlige Lösung genau dann wenn  $T^{-1}b$  ganzzahlig ist.

In diesem Fall ist  $y = \begin{pmatrix} T^{-1}b \\ 0 \end{pmatrix}$  eine ganzzahlige Lösung von  $Hy = b$ . Also ist

$$U \begin{pmatrix} T^{-1}b \\ 0 \end{pmatrix}$$

eine ganzzahlige Lösung von  $Ax = b$ . □

## Bsp. IM AUFBAU

### 8.2.3 Anwendungen zu modularen Gleichungen

**Bem.** Es gibt einen effizienten Algorithmus, der für gegebene  $b_1, \dots, b_n \in \mathbb{Z}$  und  $q_1, \dots, q_n \in \mathbb{Z}$  testet, ob das System der Kongruenzgleichungen

$$x \equiv b_1 \pmod{q_1}, \dots, x \equiv b_n \pmod{q_n}$$

eine Lösung  $x \in \mathbb{Z}$  besitzt und ggf. eine solche Lösung berechnet.

Die Kongruenzgleichung  $z \equiv b_i \pmod{q_i}$  kann man mit der Verwendung einer Zusatzvariablen  $y_i \in \mathbb{Z}$  als  $z + q_i y_i = b_i$  geschrieben werden. Das überführt das System der Kongruenzen zu einem diophantschen linearen

System

$$\begin{pmatrix} 1 & q_1 & & \\ 1 & & q_2 & \\ \vdots & & & \ddots \\ 1 & & & q_n \end{pmatrix} \begin{pmatrix} x \\ y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix},$$

das man mit Hilfe von HNF lösen kann.

**Bem.** Man betrachte den Restklassenring  $\mathbb{Z}_q := \mathbb{Z}/q\mathbb{Z}$  mit  $q \in \mathbb{N}$ . Sei  $A \in \mathbb{Z}^{m \times n}$  und  $b \in \mathbb{Z}^m$ . Es gibt einen effizienten Algorithmus, der entscheidet, ob das System  $Ax = b$  eine Lösung  $x \in \mathbb{Z}_q^n$  besitzt und ggf. eine solche Lösung berechnet. Ist  $q$  eine Primzahl, so ist  $\mathbb{Z}_q$  ein Körper, sodass man einfach das Gauß-Verfahren benutzen kann. Ansonsten ist  $\mathbb{Z}_q$  kein Körper, sodass man eine andere Methode benötigt. Wir können die Gleichung  $Ax = b$  mit den Unbekannten  $x \in \mathbb{Z}_q^n$  als die Gleichung  $Ax + qy = b$  mit den Unbekannten

$x \in \mathbb{Z}^n$  und  $y \in \mathbb{Z}^m$  interpretieren. Zu dieser Gleichung, die man auch als

$$\begin{pmatrix} A & qI_m \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = b$$

formulieren kann, kann man die oben entwickelten HNF-basierten Methoden benutzen.

### 8.3 Gitter

Ein Standardbeispiel eines Gitters ist  $\mathbb{Z}^n$ , es gibt aber auch andere Gitter. Ein Basiswechsel in einem Gitter erfolgt durch unimodulare Matrizen. Wir beginnen mit einer Charakterisierung von unimodularen Matrizen, führen danach die Gitter ein und diskutieren einige ihrer Eigenschaften.

### 8.3.1 Charakterisierung von unimodularen Matrizen

Wir nennen eine Matrix  $U \in \mathbb{Z}^{n \times n}$  unimodular, wenn  $\det(U) \in \{-1, 1\}$  gilt.

**Thm.** Für  $U \in \mathbb{Z}^{n \times n}$  sind die folgenden Bedingungen äquivalent:

- (i)  $U$  ist unimodular
- (ii)  $U^\top$  ist unimodular
- (iii)  $U^{-1}$  ist unimodular.
- (iv)  $U$  ist regulär mit  $U^{-1} \in \mathbb{Z}^{n \times n}$
- (v)  $\{Uz : z \in \mathbb{Z}^n\} = \mathbb{Z}^n$ .
- (vi) Die HNF von  $U$  ist  $I$ .

*Beweis.* Die Äquivalenz von (i), (ii) und (iii) folgt aus den Formeln für die Determinante.

(iv)  $\Rightarrow$  (v) folgt aus der Bijektion  $x \in \mathbb{Z}^n \leftrightarrow y \in \mathbb{Z}^n$  mit  $y = Ux$  und  $x = U^{-1}y$ .

(v)  $\Rightarrow$  (iv): Aus (v) folgt die Existenz von  $b_1, \dots, b_n \in \mathbb{Z}^n$  mit  $Ub_i = e_i$ . Die Matrix mit den Spalten  $b_1, \dots, b_n$  ist die ganzzahlige inverse Matrix von  $U$ .

(i)  $\Leftrightarrow$  (iv): Die Richtung (i)  $\Rightarrow$  (iv) folgt aus  $U^{-1} = U^\# / \det(U) = U^\#$ . Umgekehrt: ist (iv) erfüllt, so gilt  $1 = \det(UU^{-1}) = \det(U)\det(U^{-1})$ , mit  $\det(U), \det(U^{-1}) \in \mathbb{Z} \setminus \{0\}$ . Es folgt  $\det(U) \in \{-1, 1\}$ .

(i)  $\Leftrightarrow$  (vi): Bei einer regulären quadratischen ganzzahligen Matrix ist die HNF  $H$  gleich  $I$  gilt genau dann, wenn alle Diagonalelemente von  $H$  gleich 1 sind, das letztere ist äquivalent zu  $\det(H) = 1$ . Die HNF  $H$  von  $U$  lässt sich als  $H = UV$  für eine unimodulare Matrix  $V$  darstellen. Das bedeutet, dass bei der HNF von  $\det(H) = \det(U)$  gilt. Das zeigt die Äquivalenz von

(i) und (vi). □

### 8.3.2 Gitter, ihre Dimension und die Determinante

Sind  $b_1, \dots, b_t \in V$  linear unabhängige Vektoren in einem endlich-dimensionalen Vektorraums  $V$  über einem Körper  $\mathbb{K}$  mit  $\mathbb{Q} \subseteq \mathbb{K} \subseteq \mathbb{R}$ , so nennt

$$\Lambda := \{z_1 b_1 + \dots + z_t b_t : z_1, \dots, z_t \in \mathbb{Z}\}$$

ein Gitter in  $V$ . Die lineare Hülle  $\text{lin}(\Lambda) = \text{lin}(b_1, \dots, b_t)$  ist  $t$ -dimensional und ihre Dimension heißt der Rang oder die Dimension des Gitters  $\Lambda$ . Jedes Gitter ist eine Untegruppe von  $(V, +)$ , aber nicht umgekehrt:  $\{z_1 + \sqrt{2}z_2 : z_1, z_2 \in \mathbb{Z}\}$  ist Untegruppe von  $(\mathbb{R}, +)$  aber kein Gitter.

Für eine Matrix  $B \in \mathbb{R}^{n \times t}$  vom Rang  $t$  ist

$$\Lambda(B) := \{Bz : z \in \mathbb{Z}^t\}$$

die Gruppe, die durch die Spalten von  $B$  erzeugt ist. Sind die Spalten von  $B$  linear unabhängig, so ist  $\Lambda(B)$  ein  $t$ -dimensionales Gitter.

**Thm.** Seien  $A, B \in \mathbb{R}^{n \times n}$  reguläre Matrizen mit  $\Lambda(A) = \Lambda(B)$ . Dann ist  $A^{-1}B$  eine unimodulare Matrix und es gilt insbesondere  $|\det(A)| = |\det(B)|$ .

*Beweis.* Da die Spalten von  $A$  in  $\Lambda(B)$  liegen gilt  $A = BU$  für eine Matrix  $U \in \mathbb{Z}^{n \times n}$ . Analog: da die Spalten von  $B$  in  $\Lambda(A)$  liegen, gilt  $B = AV$  für ein  $V \in \mathbb{Z}^{n \times n}$ . Es folgt  $\det(A) = \det(B)\det(U)$  und  $\det(B) = \det(A)\det(V)$ , woraus man  $\det(A)\det(B) = \det(A)\det(B)\det(U)\det(V)$  folgert. Es gilt also  $\det(U), \det(V) \in \{-1, 1\}$ . Das ergibt die gewünschten Behauptungen.

□

Wie im vorigen Theorem gezeigt wurde ist bei der Darstellstellung eines  $n$ -dimensionalen Gitters  $\Lambda \subseteq \mathbb{Z}^{n \times n}$  als  $\Lambda = \Lambda(A)$  durch eine reguläre Matrix  $A \in \mathbb{Z}^{n \times n}$  der Wert  $|\det(A)|$  nur von  $\Lambda$  abhängig. Diesen Wert, den man als  $\det(\Lambda)$  bezeichnet, nennt man die Determinante von  $\Lambda$ .

**Bem.** Für nicht-volldimensionale Gitter kann man ebenfalls die Determi-

nante einführen...

### 8.3.3 Erzeugung rationalen Gittern

**Thm.** Sei  $V$  endlich-dimensionaler Vektorraum über  $\mathbb{Q}$  und seien  $a_1, \dots, a_m \in V$  ( $m \in \mathbb{N}$ ). Dann ist die Menge

$$\Lambda := \{z_1 a_1 + \dots + z_m a_m : z_1, \dots, z_m \in \mathbb{Z}\}$$

ein Gitter.

*Beweis.* Wenn die lineare Hülle von  $a_1, \dots, a_m$  nicht mit  $V$  übereinstimmt, können wir uns auf den Untervektorraum  $\text{lin}(a_1, \dots, a_m)$  einschränken. Daher können wir oBdA voraussetzen, dass  $V = \text{lin}(a_1, \dots, a_m)$  gilt. Ist  $n = \dim(V)$ , so können wir  $V$  mit  $\mathbb{Q}^n$  identifizieren, d.h., wir nehmen an,  $V = \mathbb{Q}^n$ . Da wir die Vektoren aus  $a_1, \dots, a_m \in \mathbb{Q}^n$  mit dem gemeinsamen Nenner aller Komponenten dieser Vektoren multiplizieren können, können wir

$a_1, \dots, a_m \in \mathbb{Z}^n$  annehmen. Wir können mit einer unimodularen Matrix  $U \in \mathbb{Z}^{m \times m}$  die Matrix  $A = (a_1, \dots, a_m) \in \mathbb{Z}^{n \times m}$  in die HNF  $H$  überführen. Das bedeutet  $H = (T \ 0) = AU$ . Dann ist

$$\Lambda = \{Az : z \in \mathbb{Z}^m\} = \{AUz : z \in \mathbb{Z}^n\} = \{Hz : z \in \mathbb{Z}^n\} = \{Ty : y \in \mathbb{Z}^m\}.$$

Das zeigt, dass  $\Lambda$  ein Gitter ist, dass durch die lineare Unabhängigen Spalten der unteren Dreiecksmatrix  $T$  erzeugbar ist.  $\square$

### 8.3.4 Gitter als diskrete Untergruppen von $\mathbb{R}^n$

Ein Teilmenge  $X$  von  $\mathbb{R}^n$  nennen wir beschränkt, wenn für eine positive reelle Zahl  $\rho > 0$  die Bedingung  $\|x\| \leq \rho$  für alle  $x \in X$  erfüllt ist. Ein Teilmenge  $X$  von  $\mathbb{R}^n$  nennen wir diskret, wenn jede beschränkte Teilmenge von  $X$  endlich ist. Das ist äquivalent dazu, dass die Menge  $X$  keine Häufungspunkte besitzt.

**Thm.** Eine Untergruppe  $G$  von  $(\mathbb{R}^n, +)$  ist genau dann ein Gitter, wenn  $G$  diskret ist.

*Beweis.* Beweis nach Gruber-Lekkerkerker? □

## 9 Tensoren und duale Vektorräume

### 9.1 Einstieg: Tensoren als multidimensionale Arrays

#### 9.1.1 Tensoren und ihre Grundoperationen

Genauso wie man Vektorräume “konkret” als  $\mathbb{K}^n$  oder abstrakt als ein  $n$ -dimensionaler Vektorraum  $V$  über  $\mathbb{K}$  behandeln kann, kann man auch Tensoren konkret und abstrakt behandeln. Wir beginnen die Diskussion der Tensoren in der konkreten Umsetzung. Seien  $I_1, \dots, I_k$  endliche Indexmengen. Wir definieren als  $\mathbb{K}^{I_1 \times \dots \times I_k}$  die Menge aller Abbildungen von  $I_1 \times \dots \times I_k$  nach  $\mathbb{K}$ . Im Fall,  $I_j = \{1, \dots, n_j\}$ , bezeichnen wir diese Menge als  $\mathbb{K}^{n_1 \times \dots \times n_k}$ .

Für  $k = 1$ , erhalten wir die Menge  $\mathbb{K}^{n_1}$  der Vektoren in der Dimension  $n_1$  und für  $k = 2$  die Menge der Matrizen der Größe  $n_1 \times n_2$ . Man nennt die Elemente von  $\mathbb{K}^{n_1 \times \cdots \times n_k}$  Tensoren der Ordnung  $k$ , man sagt auch, dass Elemente aus dieser Menge Tensoren der Größe  $n_1 \times \cdots \times n_k$  über  $\mathbb{K}$  sind. Die Menge  $\mathbb{K}^{n_1 \times \cdots \times n_k}$  aller Tensoren besitzt eine natürliche Struktur eines  $\mathbb{K}$ -Vektorraums, sodass man Linearkombinationen von Tensoren mit Koeffizienten in  $\mathbb{K}$  betrachten kann.

Darüberhinaus kann man Tensoren auf zwei verschiedene Weisen multiplizieren. Eine davon ist eine Verallgemeinerung der Matrix-Multiplikation. Für Matrizen  $A = (a_{ij}) \in \mathbb{K}^{I \times J}$  und  $B = (b_{jk}) \in \mathbb{K}^{J \times K}$ , welche durch beliebige endliche Mengen  $I, J$  und  $K$  indexiert sind, ist das Produkt  $AB = C = (c_{ik}) \in K^{I \times K}$  durch

$$c_{ik} = \sum_{j \in J} a_{i,j} b_{j,k}$$

gegeben. Im Fall von Tensoren  $A \in \mathbb{K}^{I_1 \times \cdots \times I_m \times j_1 \times \cdots \times J_n}$  und  $B \in \mathbb{K}^{J_1 \times \cdots \times J_n \times K_1 \times \cdots \times K_s}$

können wir das Produkt  $C$  genau so einführen, indem wir  $A$  als  $I \times J$ -Matrix und  $B$  als  $J \times K$ -Matrix mit  $I = I_1 \times \cdots \times I_m$ ,  $J = J_1 \times \cdots \times j_n$  und  $K = K_1 \times \cdots \times K_s$  auffassen.

Eine weitere Multiplikation, die wir für Tensoren einführen ist die Tensormultiplikation:

$$\otimes : \mathbb{K}^{I_1 \times \cdots \times I_m} \times \mathbb{K}^{J_1 \times \cdots \times J_n} \rightarrow \mathbb{K}^{I_1 \times \cdots \times I_m \times J_1 \times \cdots \times J_n}.$$

Für  $A = (a_{i_1 \dots i_m}) \in \mathbb{K}^{I_1 \times \cdots \times I_m}$  und  $B = (b_{j_1 \dots j_n}) \in \mathbb{K}^{J_1 \times \cdots \times J_n}$  ist das Tensorprodukt  $A \otimes B = C = (c_{i_1 \dots i_m j_1 \dots j_n})$  durch

$$c_{i_1 \dots i_m j_1 \dots j_n} = a_{i_1 \dots i_m} b_{j_1 \dots j_n}$$

definiert. Das Tensorprodukt ist somit eine bilineare Abbildung, das heißt, die Operation  $A \otimes B$  ist in  $A$  und in  $B$  linear.

MULTIPLIKATION VON ZWEI TENSOREN ENTLANG DER DIMENSIONEN.

Seien  $n_1 \times \cdots \times n_s$  und  $m_1 \times \cdots \times n_t$  Formate der Tensoren,  $S \subseteq \{1, \dots, s\}$  und  $T \subseteq \{1, \dots, t\}$  Teilmengen mit der gleichen Kardinalität, und  $\phi : S \rightarrow T$  bijektive Abbildung. Dann können wir eine Multiplikation entlang  $\phi : S \rightarrow T$  definieren.....

### 9.1.2 Tensorprodukt und Determinanten

Wir bemerken zuerst, dass man die Determinante einer durch eine beliebige Meng  $I$  indexierten quadratischen Matrix  $A = (a_{ij}) \in \mathbb{K}^{I \times I}$  genauso wie im Fall  $I = \{1, \dots, n\}$ , den wir diskutiert haben, einführen kann. Wenn wir die Menge  $I$  durch eine bijektive Abbildung  $\phi : \{1, \dots, n\} \rightarrow I$  mit  $n = |I|$  nummerieren, so können wir  $\det(A)$  als die Determinante der Matrix  $M = (a_{\phi(i)\phi(j)})_{i,j=1,\dots,n}$  einführen. Es stellt sich heraus, dass  $\det(A)$  auf diese Weise wohldefiniert ist, weil die Wahl der bijektiven Abbildung keine Auswirkung auf den Wert von  $\det(A)$  hat. Für das fixierte  $\phi$  kann jede andere bijektive Abbildung von  $I$  nach  $\{1, \dots, n\}$  als die Komposition  $\phi \circ \sigma$

mit einer Permutation  $\sigma \in S_n$  realisiert werden. Es stellt sich heraus, dass die Matrix  $N = (a_{\phi(\sigma(i))\phi(\sigma(j))})_{i,j=1,\dots,n}$  die gleiche Determinante wie  $M$  hat, denn wegen der Formel  $\det(a_{\sigma(1)}, \dots, a_{\sigma(n)}) = \text{sign}(\sigma) \det(a_1, \dots, a_n)$ , die man Zeilen- sowie Spaltenweise benutzen kann, gilt

$$\det(N) = \text{sign}(\sigma) \det(a_{\phi(i)\phi(\sigma(j))}) = \text{sign}(\sigma)^2 \det(a_{\phi(i)\phi(j)}) = \text{sign}(\sigma)^2 \det(M) = \det(M).$$

Betrachten wir nun zwei quadratische Matrizen  $A \in \mathbb{K}^{I \times I}$  und  $B \in \mathbb{K}^{J \times J}$ . Ihr Tensorprodukt  $A \otimes B$  ist Tensor im Format  $I \times I \times J \times J$ . Für die Zwecke dieses Abschnitts lohnt es sich aber die 4 Indizes bzw. Indexbereiche anders zu reihen: wir führen also  $\otimes$  als

$$\otimes : \mathbb{K}^{I \times I} \times \mathbb{K}^{J \times J} \rightarrow \mathbb{K}^{I \times J \times I \times J}$$

ein, mit  $(a_{i_1 i_2}) \otimes (b_{j_1 j_2}) = (a_{i_1 j_1 i_2 j_2})$ . Bemerkung, solche Anpassungen der Reihenfolge der Indexmengen sind in der Tensorrechnung üblich: man behält also eine gewisse Flexibilität bei der Einführung der Operation  $\otimes$ . Mit dieser Definition kann also  $A \times B$  als eine durch  $I \times J$ -indexierte quadratische

Matrix interpretiert werden. Dieses Matrix nennt man auch manchmal das Kronecker-Produkt von  $A$  und  $B$ .

**Lem.**  $(A \otimes B)(C \otimes D) = (AC \otimes BD)$ .

**Thm.** Für  $A \in \mathbb{K}^{m \times m}$  und  $B \in \mathbb{K}^{n \times n}$  gilt  $\det(A \otimes B) = \det(A)^n \det(B)^m$ .

*Beweis.* Wegen  $A \otimes B = (A \otimes I_n)(I_m \otimes B)$  gilt

$$\det(A \otimes B) = \det(A \otimes I_n) \det(I_m \otimes B).$$

Hierbei entspricht  $A \times I_n$  einer Blockdiagonalenmatrix mit  $n$  Diagonallöcken, die gleich  $A$  sind, und  $I_m \times B$  entspricht einer Blockdiagonalblöcken Matrix mit  $m$  Diagonallöcken, die gleich  $B$  sind. Wir erhalten also  $\det(A \times I_n) = \det(A)^n$  und  $\det(I_m \otimes B) = \det(B)^m$ .  $\square$

## IM AUFBAU

### 9.1.3 Tensoren und die Euklidische Struktur

#### IM AUFBAU

## 9.2 Tensoren und multilinear Abbildungen

### 9.2.1 Duale Vektorräume und duale Basen

Für einen Vektorraum  $V$  über  $\mathbb{K}$  wird der Raum  $\text{Lin}(V, \mathbb{K})$  der duale Raum von  $V$  genannt und als  $V^*$  bezeichnet.

**Prop.** Sei  $V$   $n$ -dimensionaler Vektorraum über  $\mathbb{K}$  mit  $n \in \mathbb{N}$ . Dann gilt:

- (i) Zu jeder Basis  $\mathcal{B} = (b_1, \dots, b_n)$  von  $V$  gibt es eine eindeutige Basis  $\mathcal{B}^* = (b_1^*, \dots, b_n^*)$  von  $V^*$  mit  $b_i^*(b_j) = \delta_{i,j}$  für alle  $1 \leq i, j \leq n$ .
- (ii) Zu jeder Basis  $\mathcal{F} = (f_1, \dots, f_n)$  von  $V^*$  gibt es eine eindeutige Basis  $\mathcal{F}^* = (f_1^*, \dots, f_n^*)$  von  $V$  mit  $f_i(f_j^*) = \delta_{i,j}$  für alle  $i, j = 1, \dots, n$ .

*Beweis.* (i) Wir setzen  $b_i^*(x) := e_i^\top x_{\mathcal{B}}$ . Nach der Definition von  $x_{\mathcal{B}}$  gilt  $x = \sum_{i=1}^n (e_i^\top x_{\mathcal{B}}) b_i = \sum_{i=1}^n b_i^*(x) b_i$ . Für  $x = b_j$  gilt  $x_{\mathcal{B}} = e_j$ , was die Existenz von  $\mathcal{B}^*$  verifiziert. Zur Eindeutigkeit: Ist  $b_i^*(b_j) = \delta_{i,j}$  für alle  $i, j$ , so gilt  $b_i^*(x) = b_i^*(\sum_{j=1}^n (e_j^\top x_{\mathcal{B}}) b_j) = e_i^\top x_{\mathcal{B}}$ .

Der Beweis von (ii) ist absolut analog, da man einfach die Rollen von  $V$  und  $*$  vertausche kann.  $\square$

Wir führen für die bilineare Form  $V^* \times V$  durch  $(f, v) := f(v)$  ein. Durch die symmetrische Schreibweise  $(f, v)$  wird hervorgehoben, dass  $V$  bzgl.  $V^*$  die gleiche Rolle spielt wie  $V^*$  bzgl.  $V$ .

**Bsp.** Duale Basis zur Basis  $f_1(x) = x_1 + x_2, f_2(x) = x_1 + x_3, f_3(x) = x_2 + x_3$  von  $(\mathbb{R}^3)^*$ . Man hat

$$\begin{pmatrix} f_1(x) \\ f_2(x) \\ f_3(x) \end{pmatrix} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} x$$

Gesucht werden Vektoren  $b_1, b_2, b_3 \in \mathbb{R}^3$  mit  $f_i(b_j) = \delta_{i,j}$ . Das bedeutet, dass wir die Vektoren  $b_1, b_2, b_3$  aus der Gleichung

$$\begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} b_1 & b_2 \\ b^3 & \end{pmatrix} = I_3$$

ermitteln sollen. Es reicht also, die Matrix auf der linken Seite zu invertieren.  
IM AUFBAU.

### 9.2.2 Tensoren und multilinear Formen

Seien  $V_1, \dots, V_r$  endlichdimensionale Vektorräume über  $\mathbb{K}$  und sei  $F : V_1 \times \dots \times V_k \mapsto \mathbb{K}$  eine Funktion, die in jedem der  $r$  Vektorargumente linear ist. Wir nennen eine solche Funktion eine multilinear Form. Bilineare Formen haben wir bereits betrachtet. Sei  $n_i = \dim(V_i)$  und wir fixieren in  $V_i$  eine Basis  $\mathcal{B}_i = (b_{i,1}, \dots, b_{i,n_i})$ . Wir können  $F$  im System  $\mathcal{B} = (\mathcal{B}_1, \dots, \mathcal{B}_r)$  der  $r$

Basen folgendermaßen darstellen. Ist  $v_i = \beta_{i,1}b_1 + \cdots + b_{i,n_i}b_{i,n_i}$ , so hat man

$$\begin{aligned} F(v_1, \dots, v_r) &= F\left(\sum_{j_1=1}^{n_1} \beta_{1,j_1} b_{1,j_1}, \dots, \sum_{j_r=1}^{n_r} b_{1,j_r}\right) \\ &= \sum_{\substack{j_1=1, \dots, n_1 \\ \vdots \\ j_r=1, \dots, n_r}} \beta_{1,j_1} \cdots \beta_{r,j_r} F(b_{j_1}, \dots, b_{j_r}). \end{aligned}$$

Die Multilineare Abbildung ist somit eindeutig durch die Angabe vom Tensor

$$F_{\mathcal{B}} = (F(b_{j_1}, \dots, b_{j_r}))_{j_1, \dots, j_r} \in \mathbb{K}^{n_1 \times \cdots \times n_r}$$

definiert. Diesen Tensor nennt man die Darstellung von  $F$  in den Basen  $\mathcal{B} = (\mathcal{B}_1, \dots, \mathcal{B}_r)$ . Darüber hinaus kann man das Auswerten von  $F$  auf

$(v_1, \dots, v_r)$  als das Skalarprodukt

$$F(v_1, \dots, v_r) = \langle F_{\mathcal{B}}, (v_1)_{\mathcal{B}_1} \otimes \cdots \otimes (v_r)_{\mathcal{B}_r} \rangle$$

darstellen.

### 9.2.3 Koordinatentransformationen für Basendarstellungen multilinearer Formen

IM AUFBAU:

$$F_{\mathcal{B}} = F_{\mathcal{A}} \cdot (T_{\mathcal{A}_1 \leftarrow \mathcal{B}_1} \otimes \cdots \otimes T_{\mathcal{A}_r \leftarrow \mathcal{B}_r}).$$

(Die Reihung der Dimensionen soll in dieser Formel geklärt werden)

Sei  $\text{Lin}(V_1, \dots, V_r; \mathbb{K})$  die Menge aller multilinearen Abbildungen von  $V_1 \times \cdots \times V_r$  nach  $\mathbb{K}$ . Diese Menge hat eine natürliche Struktur eines Vektorraums über  $\mathbb{K}$  mit dieser Struktur ist

IM AUFBAU: Isomorphismus

## 9.3 Tensoren und multilineare Abbildungen

IM AUFBAU



## **9.4 Anwendungen der Tensoren**

### **9.4.1 Quantenberechnungen**

## **9.5 Verallgemeinertes Hookesches Gesetz**

## **9.6 Tensorrang und schnelle Multiplikation von Matrizen**

## **9.7 Quantitative Biologie**

# **10 Matroide**

# **11 Verschiedenes**

## **11.1 Singulärwertzerlegung**

Die Singulärwertzerlegung formuliert man in der Regel mit Matrizen, es gibt aber eine entsprechende Formulierung mit linearen Abbildungen auf

Euklidischen Räumen:

**Thm.** Seien  $V, W$  endlich-dimensionale Euklidische Räume über  $\mathbb{K}$  und  $F : V \rightarrow W$  lineare Abbildung. Dann ist die Darstellung  $F_{\mathcal{A}, \mathcal{B}}$  von  $F$  bzgl. gewisser Orthonormalbasen  $\mathcal{A} = (a_1, \dots, a_m)$  und  $\mathcal{B} = (b_1, \dots, b_n)$  eine  $m \times n$  Matrix, deren Diagonalelemente nichnegative reelle Zahlen sind und die Elemente außerhalb der Diagonale gleich 0 sind. Mit anderen Worten:  $\langle F(a_i), b_j \rangle$  liegt in  $\mathbb{R}_{\geq 0}$  für  $i = j$  und ist 0 für  $i \neq j$ .

*Beweis.* Es gilt  $V = \ker(F)^\perp \oplus \ker(F)$  und  $W = \text{im}(F) \oplus \text{im}(F)^\perp$ . Daher kann die Basis von  $V$  sowie die Basis für  $W$  aus jeweils zwei orthonormalen Basen zusammengesetzt werden:  $\mathcal{A} = (\mathcal{A}', \mathcal{A}'')$  aus der Basis  $\mathcal{A}'$  von  $\text{im}(F)$  und einer Basis  $\mathcal{A}''$  von  $\text{im}(F)^\perp$  und  $\mathcal{B} = (\mathcal{B}', \mathcal{B}'')$  aus einer Basis  $\mathcal{B}'$  von  $\ker(F)^\perp$  und einer Basis  $\mathcal{B}''$  von  $\ker(F)$ .

Dabei können die orthonormale Basen von  $\ker(F)$  und  $\text{im}(F)^\perp$  beliebig gewählt werden, da die Behauptung für jede Wahl erfüllt ist. Somit kann  $F$  durch die Einschränkung ersetzt werden, bei der  $V$  auf  $\ker(F)^\perp$  und  $W$  auf

$\text{im}(F)$  einingeschränkt sind. Eine solche Einschränkung ist eine Bijektion, weil sie injektiv ist un die Dimension von  $\ker(F)^\top$  und  $\text{im}(F)$  (nach dem Rangsatz) gleich sind.

Auf diese Weise reduzieren wir den Fall von einem allgemeinen  $F$  auf den Fall eines bijektiven  $F$  mit  $m = \dim(W) = \dim(V) = n$ . In diesem Fall betrachten wir die Abbildung  $F^* \circ F : V \rightarrow V$ . Diese Abbildung ist als Komposition von zwei bijektiven Abbildungen bijektiv und sie ist selbstadjungiert. Nach dem Spektralsatz für selbstadjungierte Abbildungen besitzt  $F^* \circ F$  eine Orthonormalbasis  $\mathcal{B} = (b_1, \dots, b_n)$  aus den Eigenvektoren. Sei  $\lambda_i$  der Eigenwert von  $F^* \circ F$  zu  $b_i$ . Es gilt  $\lambda_i > 0$ , denn einerseits ist  $\langle (F^* \circ F)(b_i), b_i \rangle = \langle F(b_i), F(b_i) \rangle > 0$  wegen  $F(b_i) \neq 0$  und anderseits ist  $\langle (F^* \circ F)(b_i), b_i \rangle = \langle \lambda_i b_i, b_i \rangle = \lambda_i \langle b_i, b_i \rangle$  mit  $\langle b_i, b_i \rangle > 0$ . Es stellt sich heraus, dass die Behauptung mit  $\mathcal{A} = (a_1, \dots, a_n)$  und  $a_i = \frac{1}{\sqrt{\lambda_i}} F(b_i)$  erfüllt ist. Das System  $\mathcal{A}$  ist eine Orthonormalbasis von  $W$ , denn die Anzahl der

Vektoren in diesem System ist  $n = \dim(W)$  und man hat

$$\begin{aligned}
\langle a_i, a_j \rangle &= \left\langle \frac{1}{\sqrt{\lambda_i}} F(b_i), \frac{1}{\sqrt{\lambda_j}} F(b_j) \right\rangle \\
&= \left\langle \frac{1}{\sqrt{\lambda_i}} b_i, \frac{1}{\sqrt{\lambda_j}} (F^* \circ F)(b_j) \right\rangle \\
&= \left\langle \frac{1}{\sqrt{\lambda_i}} b_i, \sqrt{\lambda_j} b_j \right\rangle \\
&= \sqrt{\frac{\lambda_j}{\lambda_i}} \langle b_i, b_j \rangle \\
&= \delta_{ij}.
\end{aligned}$$

Des Weiteren gilt  $\langle F(b_i), a_j \rangle = \langle \sqrt{\lambda_i} a_i, a_j \rangle = \sqrt{\lambda_i} \langle a_i, a_j \rangle$ : dieser Wert ist  $\sqrt{\lambda_i}$  für  $i = j$  und 0 für  $i \neq j$ .  $\square$

Die Singulärwertzerlegung hat zahlreiche Anwendungen: z.B. in Statistik und Bildverarbeitung, aber auch in der numerischen linearen Algebra und

deren Anwendungen.