



Vorlesung

# Lineare Algebra I+II

gehalten im WS 2024/25 und im SS 2025

---

Prof. Dr. Tobias Harks

Universität Passau

Institut für Mathematik

Dr.-Hans-Kapfinger-Straße 30

94032 Passau

Email: [tobias.harks@uni-passau.de](mailto:tobias.harks@uni-passau.de)

22. Juli 2025



# Vorwort

Dieses Skript ist in diesem WS 24/25 die wesentliche Grundlage dieser Veranstaltung. Es beruht in großen Teilen auf dem Buch "Lineare Algebra" von Prof. Dr. Gerd Fischer (Uni Düsseldorf). An manchen Stellen sind Änderungen in der Notation und der Inhalte vorgenommen worden. Des Weiteren sind manche Beispiele und Inhalte aus dem Skript von Prof. Dr. Peter Quast (Uni Augsburg) und Prof. Dr. Tobias Kaiser (Uni Passau) übernommen worden.

Das Skript wird während des Semesters fertiggestellt und daher wird es zu kontinuierlichen Veränderungen kommen!

Passau, Februar 2025  
Prof. Dr. Tobias Harks



# Inhaltsverzeichnis

<b>1</b>	<b>Einführung</b>	<b>7</b>
1.1	Einführendes Beispiel . . . . .	7
1.2	Lineare Gleichungssysteme und Matrizen . . . . .	10
1.3	Das Eliminationsverfahren von Gauss . . . . .	14
<b>2</b>	<b>Grundbegriffe</b>	<b>21</b>
2.1	Quantoren, Implikationen und Äquivalenzen . . . . .	21
2.2	Mengen und Abbildungen . . . . .	23
2.2.1	Äquivalenzrelationen . . . . .	28
2.3	Gruppen, Ringe und Körper . . . . .	30
2.3.1	Gruppen . . . . .	30
2.3.2	Ringe . . . . .	33
2.3.3	Körper . . . . .	34
2.3.4	Komplexe Zahlen . . . . .	35
2.4	Vektorräume . . . . .	38
2.4.1	Untervektorräume . . . . .	41
2.4.2	Abschluss von Untervektorräumen und Lineare Unabhängigkeit . . . . .	41
2.5	Basen und Dimension . . . . .	44
2.5.1	Erzeugendensystem und Basis eines Vektorraumes . . . . .	44
2.5.2	Dimension eines Vektorraumes . . . . .	48
2.5.3	Anwendungen auf Matrizen . . . . .	53
<b>3</b>	<b>Lineare Abbildungen</b>	<b>59</b>
3.1	Bild, Faser und Kern von Homomorphismen . . . . .	64
3.2	Dimensionsformel . . . . .	67
3.3	Lineare Gleichungssysteme . . . . .	68
3.4	Lineare Abbildungen und Matrizen . . . . .	72
3.5	Quotientenvektorräume . . . . .	76
3.6	Multiplikation von Matrizen . . . . .	79
3.7	Koordinatentransformation bei Basiswechsel . . . . .	87
<b>4</b>	<b>Determinanten</b>	<b>95</b>
4.1	Definition und Grundlegende Eigenschaften . . . . .	95

## 6 | Inhaltsverzeichnis

4.2	Existenz und Eindeutigkeit . . . . .	102
4.3	Laplacescher Entwicklungssatz . . . . .	103
4.4	Leibnizsche Formel . . . . .	107
<b>5</b>	<b>Eigenwerte und Normalformen von Endomorphismen</b>	<b>109</b>
5.1	Polynome . . . . .	110
5.2	Eigenwerte . . . . .	114
5.3	Das Charakteristische Polynom . . . . .	117
5.4	Diagonalisierung . . . . .	120
5.5	Simultane Diagonalisierung . . . . .	127
5.6	Trigonalisierung . . . . .	128
5.7	Potenzen von Endomorphismen . . . . .	134
5.8	Die Jordansche Normalform . . . . .	143
<b>6</b>	<b>Euklidische und unitäre Vektorräume</b>	<b>155</b>
6.1	Bilinearformen . . . . .	155
6.2	Reelle Skalarprodukte . . . . .	158
6.3	Komplexe Skalarprodukte . . . . .	162
6.4	Orthogonalität . . . . .	164
6.5	Volumina und Haddamard Ungleichung . . . . .	169
6.6	Adjungierte Abbildungen . . . . .	171
6.7	Selbstadjungierte Endomorphismen . . . . .	175
6.8	Unitäre und orthogonale Abbildungen . . . . .	179
6.9	Symmetrische Bilinearformen und positiv definite Matrizen . . . . .	186
6.10	Singulärwertzerlegung . . . . .	191

## Kapitel 1

# Einführung

Im Wesentlichen befasst sich die Lineare Algebra mit der Theorie der Lösung von linearen Gleichungssystemen (LGS).

### 1.1 Einführendes Beispiel

Wir werden dies nun an einem kleinen Beispiel erläutern:  
Betrachte folgendes lineare Gleichungssystem:

$$\begin{aligned}2x + 3y &= 1 \\ x + 2y &= 1\end{aligned}$$

in den Variablen  $x, y \in \mathbb{R}$ .<sup>1</sup> Aus der Schule kennen wir das Eliminationsverfahren: Wir lösen die erste Gleichung nach  $x$  auf

$$x = \frac{1}{2} - \frac{3}{2}y$$

und setzen den Term für  $x$  in die 2. Gleichung ein:

$$\frac{1}{2} - \frac{3}{2}y + 2y = 1.$$

Diese Gleichung können wir nach  $y$  auflösen und erhalten

$$y = 1.$$

Nach dem Einsetzen von  $y = 1$  in den Term  $x = \frac{1}{2} - \frac{3}{2}y$  erhalten wir  $x = -1$  und somit ergibt sich eine Lösung  $x = -1, y = 1$ . Folgende Fragen treten unmittelbar auf:

- Ist dies die einzige Lösung dieses Gleichungssystems? Bei dieser Frage muss natürlich geklärt ein, in welcher Menge gesucht werden darf (im Beispiel suchen wir Werte für  $x$  und  $y$  im Zahlenbereich  $\mathbb{R}$ ).
- Kann es sein, dass für ein LGS keine oder (unendlich) viele Lösungen existieren?

---

<sup>1</sup>In dieser Vorlesung gehen wir davon aus, dass die Mengen  $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$  (aus der Schule!?) bekannt sind. Eine präzise Definition von  $\mathbb{R}$  wird üblicherweise in der Analysis 1 Vorlesung erfolgen.

## 8 | Kapitel 1. Einführung

- Welche Struktur haben Lösungsmengen ?
- Welcher Zusammenhang besteht zwischen der Struktur von Lösungen und den Eingabeparametern (hier die auftretenden Koeffizienten und die Zahlen der rechten Seite des LGS)?
- Kann man Lösungen von LGS effizient (mit einem Algorithmus) berechnen ?

Wir illustrieren einige der aufgeworfenen Fragen mittels eines weiteren Beispiels. Angenommen wir parametrisieren das obige LGS mittels zweier Parameter  $a, b \in \mathbb{R}$  in der Form:

$$\begin{aligned}2x + 3y &= 1 \\ x + ay &= b\end{aligned}$$

wiederum in den Variablen  $x, y \in \mathbb{R}$ .

Mittels gleichem Vorgehen erhalten wir

$$\frac{1}{2} - \frac{3}{2}y + ay = b.$$

Diese Gleichung können wir nach  $y$  auflösen und erhalten für  $a \neq \frac{3}{2}$ :

$$y = \frac{1}{a - \frac{3}{2}} \left( b - \frac{1}{2} \right).$$

Einsetzen des Wertes für  $y$  in die 2. Gleichung liefert die Gleichung:

$$2x + 3 \left( \frac{1}{a - \frac{3}{2}} \left( b - \frac{1}{2} \right) \right) = 1,$$

welche wir nach  $x$  auflösen können:

$$x = 1/2 - \left( \frac{3/2}{a - \frac{3}{2}} \left( b - \frac{1}{2} \right) \right).$$

Für die konkreten Werte  $a = b = 1$  ergibt sich z.B. die Lösung

$$y = \left( \frac{1}{-1/2} \right) \cdot \frac{1}{2} = -1, x = 1/2 + 3/2 = 2.$$

Bei unserer Rechnung haben wir (außer der Bedingung  $a \neq \frac{3}{2}$ ) "keine Information verloren", so dass es für ein Paar von beliebigen aber festen Parametern  $(a, b)$  mit  $a \neq \frac{3}{2}, b \in \mathbb{R}$  genau eine Lösung der Form

$$\left( 1/2 - \left( \frac{3/2}{a - \frac{3}{2}} \left( b - \frac{1}{2} \right) \right), \frac{1}{a - \frac{3}{2}} \left( b - \frac{1}{2} \right) \right)$$

gibt. Was passiert denn nun für den Fall  $a = 3/2, b \in \mathbb{R}$  ? In diesem Fall, bei Beibehaltung

der Eliminationslogik, ergibt sich:

$$\frac{1}{2} - \frac{3}{2}y + ay = b$$

was

$$\frac{1}{2} = b$$

ergibt. Wenn jedoch eingangs  $b \neq 1/2$  galt, bekommen wir hier einen Widerspruch und somit kann unser LGS für den Fall  $a = 3/2, b \neq 1/2$  keine Lösung besitzen. Bleibt nun der Fall  $a = 3/2, b = 1/2$  übrig. Für diesen Fall erfüllt offensichtlich jedes  $y \in \mathbb{R}$  die 2. Gleichung

$$\frac{1}{2} - \frac{3}{2}y + ay = b$$

und somit ist – bei Erfüllung der 1. Gleichung – diese 2. Gleichung immer erfüllt und enthält daher keine neuen Informationen. Somit können wir die gesamte Lösungsmenge basierend auf der 1. Gleichung darstellen:

$$\{(1/2 - 3/2y, y) | y \in \mathbb{R}\} \subset \mathbb{R}^2.$$

Diese Menge ist eine Gerade im  $\mathbb{R}^2$  und enthält offensichtlich unendlich viele Lösungen des LGS.

Wir haben also gesehen, dass abhängig von der Wahl der Parameter  $a, b$  des LGS, die Lösungsmenge entweder einelementig (im Fall  $a \neq \frac{3}{2}, b \in \mathbb{R}$ ), leer (im Fall  $a = \frac{3}{2}, b \neq 1/2$ ) oder unendlich gross (im Fall  $a = \frac{3}{2}, b = 1/2$ ) ist und in diesem letzten Fall einer Gerade im  $\mathbb{R}^2$  entspricht.

In Verlaufe der Vorlesung werden wir den Einfluss der Eingabeparameter auf die Struktur der Lösungsmengen systematisch untersuchen. An dieser Stelle sei allerdings explizit darauf hingewiesen, dass wir in der Vorlesung die obigem Fragestellungen aus einem abstrakteren Standpunkt aus beleuchten (die Rechenoperationen  $+$  und  $\cdot$  werden allgemeiner aufgefasst und axiomatisiert und die Menge  $\mathbb{R}$  in der wir Lösungen suchen wird allgemeiner aufgefasst was zu den Begriffen [Gruppen](#), [Ringe](#), [Körper](#), [Vektorräume](#) führt). Die Lösungen eines LGS werden abstrakter aufgefasst als Nullstellen von ([affinen](#)) [linearen Abbildungen](#). Diese Abstraktionen vereinfachen manche Beweise aber vor allem führen sie zu einer sehr schönen und mächtigen Theorie, die in anderen Gebieten der Mathematik wie auch in anderen Wissenschaftsgebieten zahlreiche Anwendungen findet.

## 1.2 Lineare Gleichungssysteme und Matrizen

Wie oben schon beschrieben, wollen wir mehrere Gleichungen mit mehreren Variablen lösen, und daher bewegen wir uns im **reellen Standardraum** der Dimension<sup>2</sup>  $n$ :

$$\mathbb{R}^n = \left\{ x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} : x_i \in \mathbb{R}, i = 1, \dots, n \right\}.$$

$\mathbb{R}^n$  ist die Gesamtheit der geordneten Tupel (auch **Vektoren** genannt) von reellen Zahlen. Die Ordnung ist wichtig an dieser Stelle, denn zwei Elemente  $x, y \in \mathbb{R}^n$  sind genau dann gleich, d.h.  $x = y$ , wenn  $x_i = y_i$  für alle  $i = 1, \dots, n$ . In der linearen Algebra werden wir mit den Vektoren rechnen und die grundlegenden Operationen sind die **Addition** von zwei Vektoren und **Multiplikation** mit einer Zahl definiert durch:

$$x + y := \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}, \quad \lambda \cdot x := \begin{pmatrix} \lambda x_1 \\ \vdots \\ \lambda x_n \end{pmatrix} \quad \text{für } \lambda \in \mathbb{R}.$$

**Definition 1.1 (Lineare Gleichungssysteme).** Seien  $n, m \in \mathbb{N} := \{1, 2, \dots\}$ . Eine **lineare Gleichung** über  $\mathbb{R}^n$  in den **Variablen**  $x_1, \dots, x_n$  ist eine Gleichung der Form

$$a_1 x_1 + \dots + a_n x_n = b,$$

wobei  $a_1, \dots, a_n \in \mathbb{R}$  die **Koeffizienten** der Variablen sind. Ein Tupel  $x^* = \begin{pmatrix} x_1^* \\ \vdots \\ x_n^* \end{pmatrix}$  ist eine Lösung der Gleichung, falls gilt:

$$a_1 x_1^* + \dots + a_n x_n^* = b.$$

Ein **lineares Gleichungssystem (LGS)** über  $\mathbb{R}^n$  in den **Variablen**  $x_1, \dots, x_n$  besteht aus  $m$  solchen linearen Gleichungen:

$$\begin{aligned} a_{11}x_1 + \dots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + \dots + a_{2n}x_n &= b_2 \\ &\vdots \\ &\vdots \end{aligned}$$

<sup>2</sup>Was genau unter dem Begriff **Dimension** zu verstehen ist, wird später erklärt.

$$a_{m1}x_1 + \dots + a_{mn}x_n = b_m,$$

wobei wiederum  $a_{ij} \in \mathbb{R}$ ,  $1 \leq i \leq m$ ,  $1 \leq j \leq n$  die Koeffizienten der Variablen sind.

Wir werden nun eine kompakte Schreibweise für LGS mittels **Matrizen** einführen.

**Definition 1.2.** Seien  $n, m \in \mathbb{N}$ . Eine  $m \times n$  Matrix mit Einträgen aus  $\mathbb{R}$  ist gegeben durch

$$A := (a_{ij})_{1 \leq i \leq m, 1 \leq j \leq n}.$$

Dargestellt werden Matrizen mittels eines Schemas, der sogenannten Matrix-Darstellung, welche durch  $m$  **Zeilen** und  $n$  **Spalten** gekennzeichnet ist:

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix}$$

Die Menge der  $m \times n$  Matrizen mit Einträgen aus  $\mathbb{R}$  wird mit  $M(m \times n, \mathbb{R})$  bezeichnet.

Wir werden die Zeilenvektoren der Matrix mit  $a_i$ ,  $i = 1, \dots, m$  und die Spaltenvektoren  $a^j \in \mathbb{R}^n$ ,  $j = 1, \dots, n$  bezeichnen. Es gilt also<sup>3</sup>

$$a_i = (a_{i1}, \dots, a_{in}) \text{ und } a^j = \begin{pmatrix} a_{1j} \\ \vdots \\ a_{mj} \end{pmatrix}.$$

Desweiteren erhalten wir:

$$A = (a_i)_{i=1, \dots, m} = (a^j)_{j=1, \dots, n}.$$

Die rechte Seite eines LGS wird nun mit einem Vektor  $b$  mit  $m$  Einträgen beschrieben:

$$b = \begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix} \in \mathbb{R}^m.$$

<sup>3</sup>Formal ist  $a_i$  also eine  $1 \times n$ -Matrix und  $a^j$  eine  $m \times 1$ -Matrix aber es ist üblich  $a_i$  als Element des  $\mathbb{R}^n$  und  $a^j$  als Element des  $\mathbb{R}^m$  zu identifizieren.

## 12 | Kapitel 1. Einführung

Wenn wir die Variable (Unbekannte)  $x$  auch als Spaltenvektor auffassen:

$$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$$

und ein Produkt zwischen einer Matrix  $A$  und einem Vektor  $x$  erklären:

$$A \cdot x := \begin{pmatrix} a_{11}x_1 + \cdots + a_{1n}x_n \\ a_{21}x_1 + \cdots + a_{2n}x_n \\ \vdots \\ a_{m1}x_1 + \cdots + a_{mn}x_n \end{pmatrix},$$

so erhalten wir die folgende kompakte Schreibweise für ein LGS:

$$Ax = b.$$

In dieser Schreibweise nennt man  $A$  die **Koeffizientenmatrix**. Für das LGS  $Ax = b$  nennt man das Tupel  $(A, b)$  die **erweiterte Koeffizientenmatrix**:

$$(A, b) := \begin{pmatrix} a_{11} & \cdots & a_{1n} & b_1 \\ a_{21} & \cdots & a_{2n} & b_2 \\ \vdots & & \vdots & \\ a_{m1} & \cdots & a_{mn} & b_m \end{pmatrix}.$$

An dieser Stelle sei erwähnt, dass es in der Mathematik üblich ist, die Summation von Termen wie z.B.  $a_{11}x_1 + \cdots + a_{1n}x_n$  mit dem Summensymbol  $\sum$  abzukürzen:

$$c_1 + \cdots + c_k =: \sum_{j=1}^k c_j.$$

Also erhalten wir:

$$a_{11}x_1 + \cdots + a_{1n}x_n = \sum_{j=1}^n a_{1j}x_j.$$

**Definition 1.3.** Für  $A \in M(m \times n, \mathbb{R})$ ,  $b \in \mathbb{R}^m$  bezeichne  $\text{Lös}(A, b)$  die Lösungsmenge des LGS  $Ax = b$ , d.h.,

$$\text{Lös}(A, b) := \{x \in \mathbb{R}^n \mid Ax = b\}.$$

Nun werden wir uns der Frage widmen, wie sich die Lösungsmenge  $\text{Lös}(A, b)$  eines LGS berechnen lässt.

**Beispiel 1.4.** Betrachte ein LGS in den Variablen  $x_1, x_2, x_3$ :

$$x_1 + 2x_2 + 2x_3 = 6$$

$$2x_2 + 2x_3 = 4$$

$$x_1 + 2x_2 + x_3 = 6$$

Dieses LGS hat nicht die schöne Gestalt, dass man direkt eine Variable isolieren kann. Wir können aber die 3. Gleichung verändern, indem wir die 1. von der 3. subtrahieren:

$$x_1 + 2x_2 + 2x_3 = 6$$

$$2x_2 + 2x_3 = 4$$

$$-x_3 = 0$$

Indem wir die 3. Gleichung mit  $-1$  multiplizieren erhalten wir  $x_3^* = 0$ . Einsetzen in die 2. Gleichung liefert  $x_2^* = 2$  und wiederum Einsetzen in die 1. Gleichung liefert:  $x_1^* = 2$ .

Eine wichtige Frage ist, ob durch obige Gleichungsumformungen eventuell Lösungen verloren gehen können.

**Definition 1.5 (Elementare Zeilenumformungen).** Sei  $A \in M(m \times n, \mathbb{R})$ . Folgende Umformungen heißen **elementare Zeilenumformungen**:

1. Vertauschung zweier Zeilen (Typ 1).
2. Multiplikation einer Zeile mit einer Zahl  $\lambda \in \mathbb{R}^* := \mathbb{R} \setminus \{0\}$  (Typ 2).
3. Addition des  $\lambda$ -fachen der  $j$ -ten Zeile zur  $k$ -ten Zeile, wobei  $k \neq j$  und  $\lambda \in \mathbb{R}$  (Typ 3).

Wir erhalten folgenden Satz.

**Satz 1.6.** Sei  $A \in M(m \times n, \mathbb{R})$ ,  $b \in \mathbb{R}^m$  und nehme an, dass  $(A', b')$  durch endlich viele elementare Zeilenumformungen der erweiterten Koeffizientenmatrix  $(A, b)$  entstanden ist. Dann gilt:

$$\text{Lös}(A, b) = \text{Lös}(A', b').$$

*Beweis.* Da wir nur endlich viele elementare Zeilenumformungen betrachten, können wir uns auf zwei aufeinander folgende Zeilenumformungen beschränken, denn dann gilt die gewünschte Eigenschaft ja weiterhin (nach Anwendung eines der Typen I-III).<sup>4</sup>

Typ 1 ist klar.

Typ 2: auch klar, da für alle  $\lambda \in \mathbb{R}^*$ :  $\sum_{j=1}^n a_{1j}x_j = b_j \Leftrightarrow \sum_{j=1}^n \lambda a_{1j}x_j = \lambda b_j$ .

<sup>4</sup>Vgl. das Prinzip der Induktion, siehe Satz 2.6.

Typ 3: Betrachte  $j \neq k$  und  $\lambda \in \mathbb{R}$ . Da alle anderen Gleichungen unverändert bleiben, ist folgendes zu zeigen:

$$\begin{aligned} a_{k1}x_1 + \cdots + a_{kn}x_n &= b_k \\ a_{j1}x_1 + \cdots + a_{jn}x_n &= b_j \end{aligned} \quad (1.1)$$

gilt, genau dann wenn für  $\lambda \in \mathbb{R}$  gilt:

$$\begin{aligned} a_{k1}x_1 + \cdots + a_{kn}x_n &= b_k \\ (a_{j1} + \lambda a_{k1})x_1 + \cdots + (a_{jn} + \lambda a_{kn})x_n &= b_j + \lambda b_k. \end{aligned} \quad (1.2)$$

Sei (1.1) erfüllt für  $x \in \mathbb{R}^n$ . Addiere  $\lambda a_{k1}x_1 + \cdots + \lambda a_{kn}x_n = \lambda b_k$  zur zweiten Gleichung von (1.1) und wir erhalten  $(a_{j1} + \lambda a_{k1})x_1 + \cdots + (a_{jn} + \lambda a_{kn})x_n = b_j + \lambda b_k$ . Umgekehrt, sei (1.2) erfüllt für  $x \in \mathbb{R}^n, \lambda \in \mathbb{R}$ . Subtrahiere das  $\lambda$ -fache der ersten Gleichung von der zweiten Gleichung von (1.2) und wir erhalten (1.1).  $\square$

### 1.3 Das Eliminationsverfahren von Gauss

Nun werden wir ein konkretes Verfahren kennenlernen, welches die Menge  $\text{Lös}(A, b)$  angibt. Die Idee zur Bestimmung einer Lösung ist es, das LGS mittels elementarer Zeilenumformungen so darzustellen, dass iterativ eine Variable explizit bestimmbar ist und dann diese in die übrigen Gleichungen eingesetzt werden kann. Wir veranschaulichen dies am Beispiel 1.4.

$$\begin{aligned} x_1 + 2x_2 + 2x_3 &= 6 \\ 2x_2 + 2x_3 &= 4 \\ x_1 + 2x_2 + x_3 &= 6 \end{aligned}$$

Mittels der erweiterten Koeffizientenmatrix erhalten wir alternativ:

$$(A, b) = \left( \begin{array}{ccc|c} 1 & 2 & 2 & 6 \\ 0 & 2 & 2 & 4 \\ 1 & 2 & 1 & 6 \end{array} \right)$$

Im Folgenden werden wir die Abtrennung der letzten Spalte mit  $|$  weglassen. Nach einer elementaren Zeilenoperation hatten wir das LGS überführt in die Form:

$$\begin{aligned} x_1 + 2x_2 + x_3 &= 6 \\ 2x_2 + 2x_3 &= 4 \\ -x_3 &= 0 \end{aligned}$$

mit der erweiterten Koeffizientenmatrix

$$(A', b') = \begin{pmatrix} 1 & 2 & 1 & 6 \\ 0 & 2 & 2 & 4 \\ 0 & 0 & -1 & 0 \end{pmatrix}$$

In dieser sogenannten **Zeilenstufenform** der erweiterten Koeffizientenmatrix konnten wir nun direkt die Lösung für  $x_3$  von der dritten Zeile ablesen  $x_3^* = 0$  und diesen Wert in die 2. Gleichung einsetzen, um  $x_2^* = 2$  zu bestimmen, und so weiter. Um dieses Vorgehen zu formalisieren, werden wir den Begriff der **Zeilenstufenform einer Matrix** definieren.

**Definition 1.7 (Zeilenstufenform).** Eine Matrix  $A \in M(m \times n, \mathbb{R})$  ist in **Zeilenstufenform**, falls folgendes gilt:

1. Es gibt eine Zahl  $r \in \mathbb{N} \cup \{0\}$  mit  $0 \leq r \leq m$ , so dass  $a_i \neq 0, 1 \leq i \leq r$  und  $a_i = 0, i = r + 1, \dots, m$ .
2. Für  $1 \leq i \leq r$  definiere

$$j_i := \min\{j : a_{ij} \neq 0\}$$

als den kleinsten Index einer Spalte  $1 \leq j \leq n$ , so dass ein Eintrag  $a_{ij} \neq 0$  vorkommt. Für diese Zahlen fordern wir:

$$j_1 < j_2 < \dots < j_r.$$

Die Zahl  $r$  gibt die Anzahl der Stufen von  $A$  an und die besonderen Einträge  $a_{ij_i}$  für  $1 \leq i \leq r$  heißen **Pivots**. Wichtig ist, dass der Fall  $r = 0$  zulässig ist. In diesem Fall sind alle Einträge von  $A$  gleich 0!

Wir geben zunächst ein konkretes Beispiel.

$$A = \begin{pmatrix} 0 & 1 & 2 & 3 & 4 & 0 & 4 \\ 0 & 0 & 2 & 3 & 4 & 6 & 0 \\ 0 & 0 & 0 & 0 & 0 & 6 & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Hier ist  $m = 4, n = 7, r = 3, j_1 = 2, j_2 = 3, j_3 = 6$ . In Abbildung 1.3 ist eine allgemeine Matrix in Zeilenstufenform angegeben. Wir zeigen nun, dass jede Matrix  $A$  durch elementare Zeilenumformungen in eine Zeilenstufenmatrix überführt werden kann.

**Lemma 1.8.** Sei  $A \in M(m \times n, \mathbb{R})$  und  $s \in \{0, \dots, m\}$  und  $j_1, j_2, \dots, j_s \in \{1, \dots, n\}$ , so dass Folgendes gilt:

$$A = \begin{pmatrix} a_{11} & (*) & (*) & (*) & (*) & (*) & (*) & \dots & \dots & \dots & \dots & \dots & (*) \\ 0 & a_{22} & (*) & (*) & (*) & (*) & (*) & \dots & \dots & \dots & \dots & \dots & (*) \\ 0 & 0 & 0 & a_{34} & (*) & (*) & (*) & \dots & \dots & \dots & \dots & \dots & (*) \\ 0 & 0 & 0 & 0 & 0 & 0 & a_{47} & (*) & \dots & \dots & \dots & \dots & (*) \\ \vdots & & & & \vdots & & & & \vdots & & & & \vdots \\ 0 & 0 & 0 & \dots & \dots & \dots & \dots & \dots & 0 & a_{rl} & (*) & \dots & (*) \\ 0 & 0 & 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & 0 & 0 & 0 \\ 0 & 0 & 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & 0 & 0 & 0 \\ \vdots & & & & \vdots & & & & \vdots & & & & \vdots \\ 0 & 0 & 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & 0 & 0 & 0 \end{pmatrix}$$

Abbildung 1.1: Allgemeine Matrix in Zeilenstufenform.

1.  $j_1 < j_2 < \dots < j_s$ ,
2.  $a_{ij} = 0$  für alle  $i \in \{1, \dots, s\}, j = 1, \dots, j_i - 1$ ,  $a_{i,j_i} \neq 0$
3.  $a_{ij} = 0$  für alle  $i \in \{s+1, \dots, m\}, j \in \{1, 2, \dots, j_s\}$ .

Dann ist  $A$  entweder schon in Zeilenstufenform oder es gibt eine Matrix  $A'$ , die durch elementare Zeilenumformungen aus  $A$  hervorgeht und für die gilt: Es gibt  $j_{s+1} \in \{0, \dots, n\}$ , so dass gilt:

- $j_1 < j_2 < \dots < j_s < j_{s+1}$ ,
- $a'_{ij} = 0$  für alle  $i \in \{1, \dots, s+1\}, j = 1, \dots, j_i - 1$ , und  $a'_{i,j_i} \neq 0$
- $a'_{ij} = 0$  für alle  $i \in \{s+2, \dots, m\}, j \in \{1, 2, \dots, j_{s+1}\}$ .

*Beweis.* Falls  $a_{ij} = 0$  für alle  $i \in \{s+1, \dots, m\}, j \in \{j_s+1, \dots, n\}$ , so hat  $A$  Zeilenstufenform und das Lemma ist bewiesen. Anderenfalls sei

$$j_{s+1} := \min\{j \in \{1, \dots, n\} \mid \exists i \in \{s+1, \dots, m\} \text{ mit } a_{ij} \neq 0\}.$$

Wegen 3. gilt  $j_{s+1} > j_s$ . Sei

$$k := \min\{i \in \{s+1, \dots, m\} : a_{i,j_{s+1}} \neq 0\}.$$

Wir können nun o.B.d.A. annehmen, dass  $a_{s+1,j_{s+1}} \neq 0$  (sonst vertausche  $(s+1)$ -te Zeile

mit der  $k$ -ten Zeile). Für  $i \in \{s+2, \dots, m\}$  setze

$$\lambda_i := -\frac{a_{i,j_{s+1}}}{a_{s+1,j_{s+1}}}.$$

Indem nun für  $i \in \{s+2, \dots, m\}$  das  $\lambda_i$ -fache der  $(s+1)$ -ten Zeile zur  $i$ -ten Zeile addiert, erhält man eine Matrix  $A'$  mit den gewünschten Eigenschaften.  $\square$

Wir erhalten nun den zentralen Satz der oben versprochene Ergebnis liefert.

**Satz 1.9.** Sei  $A \in M(m \times n, \mathbb{R})$ . Dann ist  $A$  durch elementare Zeilenumformungen in eine Zeilenstufenmatrix  $A'$  überführbar.

*Beweis.*  $A$  erfüllt die Voraussetzungen von Lemma 1.8 mit  $s = 0$ . Durch höchstens  $m$ -malige Anwendung der elementaren Zeilenumformung wie im Beweis von Lemma 1.8 erhält man eine Zeilenstufenmatrix  $A'$ . Beachte, dass die Anzahl der Stufen durch eine Anwendung der elementaren Zeilenumformung (wie im Beweis von Lemma 1.8) um 1 erhöht wird.  $\square$

**Bemerkung 1.10.** Die iterative Anwendung der elementaren Zeilenumformung wie im Beweis von Lemma 1.8 ergibt den sogenannten **Gauss-Algorithmus**.

Wir geben ein Beispiel an.

**Beispiel 1.11.**

$$A = \begin{pmatrix} 0 & 0 & 1 & 2 & 9 \\ 0 & 3 & 4 & 5 & 9 \\ 0 & 6 & 7 & 8 & 9 \\ 0 & 9 & 9 & 9 & 9 \end{pmatrix}$$

Für  $A$  ist im Sinne von Lemma 1.8  $s = 0$ . Wir erhalten  $j_{s+1} = j_1 = 2$  und  $k = 2$ , weil  $a_{12} = 0$  aber  $a_{22} \neq 0$ . Durch Vertauschen der 1. mit der 2. Zeile erhalten wir (der o.B.d.A. Schritt!):

$$\begin{pmatrix} 0 & 3 & 4 & 5 & 9 \\ 0 & 0 & 1 & 2 & 9 \\ 0 & 6 & 7 & 8 & 9 \\ 0 & 9 & 9 & 9 & 9 \end{pmatrix}$$

Addiere nun zur 3. Zeile das  $-2$ -fache der 1. Zeile und zur 4. Zeile das  $-3$ -fache der 1.

Zeile und wir erhalten:

$$\begin{pmatrix} 0 & 3 & 4 & 5 & 9 \\ 0 & 0 & 1 & 2 & 9 \\ 0 & 0 & -1 & -2 & -9 \\ 0 & 0 & -3 & -6 & -18 \end{pmatrix}$$

An dieser Stelle habe wir die Anzahl der Stufen von 0 auf 1 erhöht mit dem Pivotelement  $a'_{12} = 3$ . In der nächsten Iteration starten wir nun mit  $s = 1, j_s = 2$ . Wir bestimmen  $j_{s+1} = 3, k = 2$ . Addiere nun zur 3. Zeile die 2. Zeile und zur 4. Zeile das 3-fache der 2. Zeile und wir erhalten:

$$\begin{pmatrix} 0 & 3 & 4 & 5 & 9 \\ 0 & 0 & 1 & 2 & 9 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 9 \end{pmatrix}$$

Vertausche 3. und 4. Zeile und erhalte:

$$\begin{pmatrix} 0 & 3 & 4 & 5 & 9 \\ 0 & 0 & 1 & 2 & 9 \\ 0 & 0 & 0 & 0 & 9 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

Diese Matrix ist in Zeilenstufenform.

Wir erhalten nun folgenden Zusammenhang zur Lösbarkeit des LGS  $Ax = b$ :

**Satz 1.12.** Sei  $A \in M(m \times n, \mathbb{R})$  in Zeilenstufenform,  $b \in \mathbb{R}^m$ ,  $r \in \{0, \dots, m\}$  die Anzahl der Stufen und  $j_1 < \dots < j_r$  die Indizes der Pivotspalten. Dann gilt:

1.  $\text{Lös}(A, b) \neq \emptyset \Leftrightarrow b_{r+1} = \dots = b_m = 0$ .
2. Falls  $\text{Lös}(A, b) \neq \emptyset$ , so gilt

$$\text{Lös}(A, b) = \left\{ x \in \mathbb{R}^n \mid x_{j_i} = \frac{1}{a_{i,j_i}} \left( b_i - \sum_{j=j_i+1}^n a_{ij} x_j \right) \text{ für } i \in \{1, \dots, r\} \right\}.$$

Insbesondere sind alle  $x_j$  für  $j \notin \{j_1, \dots, j_r\}$  frei wählbar, während die Werte für  $x_{j_1}, \dots, x_{j_r}$  durch die obigen Gleichungen festgelegt sind.

*Beweis.* Falls  $\text{Lös}(A, b) \neq \emptyset$  so muss natürlich  $b_{r+1} = \dots = b_m = 0$ . Alle anderen Implikationen ergeben sich durch sukzessive Einsetzen der Werte von  $x_{j_i}$  (in der Reihenfolge

$l = r, \dots, 1$ ) in die jeweils obigen Gleichungen.  $\square$

Insgesamt erhalten wir nun einen Algorithmus, der die Lösungsmenge  $\text{Lös}(A, b)$  bestimmt:

1. Die erweiterte Koeffizientenmatrix  $(A, b)$  wird mit Hilfe elementarer Zeilenumformungen auf Zeilenstufenform  $(A', b')$  gebracht. Mit Satz 1.9 wissen wir, dass  $\text{Lös}(A, b) = \text{Lös}(A', b')$  gilt.
2. Mithilfe von Satz 1.12 wird dann  $\text{Lös}(A', b')$  gelöst.



## Kapitel 2

# Grundbegriffe

### 2.1 Quantoren, Implikationen und Äquivalenzen

Mathematische Symbole dienen häufig als Abkürzungen bzw. als kompakte Schreibweise für bestimmte Sachverhalte oder Beziehungen von Aussagen zueinander. Die beiden wichtigsten Beispiele sind die Quantoren:

$\forall$  bedeutet “für alle“

$\exists$  bedeutet “es existiert (mindestens ein)“

$\exists!$  bedeutet “es existiert genau ein“.

#### Beispiel 2.1.

$$\forall z \in \mathbb{Q} \exists (p \in \mathbb{Z}, q \in \mathbb{N}) : z = p/q$$

In Worten: für jede rationale Zahl  $z$  existiert (mindestens) ein Paar  $p, q$  mit  $p$  ganzzahlig und  $q$  natürliche Zahl, so dass  $z$  dargestellt werden kann in der Form  $z = p/q$ .

Eine Aussage wird hier nicht exakt definiert aber wir geben einige Beispiele:

- Der Himmel ist blau (je nach Wetterlage wahr oder falsch)
- $2 > 1$  (wahr)
- 7 ist durch 3 teilbar (falsch)

Wir nehmen an, dass eine Aussage stets entweder wahr oder falsch ist, aber nicht beides zugleich. Es gibt natürlich manche Aussagen, von denen nicht bekannt ist ob sie wahr oder falsch sind (siehe die offenen Probleme der Mathematikforschung). Wenn aus einer Aussage  $A$  die Aussage  $B$  folgt, so schreibt man

$$A \Rightarrow B.$$

Es wird auch gesagt:  $A$  impliziert  $B$ . Etwas formaler können wir  $A \Rightarrow B$  als Aussage auffassen und dann mittels einer Wahrheitstafel definieren, ob sie abhängig von der Wahr-

heitsbelegung von A und B wahr oder falsch ist:

A	B	A $\Rightarrow$ B
w	w	w
w	f	f
f	w	w
f	f	w

$A \Rightarrow B$  ist wahr, falls die Voraussetzung A falsch oder die Folgerung B wahr ist.  $A \Rightarrow B$  ist falsch, falls die Voraussetzung A wahr und die Folgerung B falsch ist.

Die Beziehung

$$A \Leftarrow B$$

bedeutet dass B die Aussage A impliziert und ist gleichbedeutend mit

$$B \Rightarrow A.$$

Aussagen können logisch mittels der “und” und “oder” Beziehung verknüpft werden um neue Aussagen zu generieren:

$A \wedge B$  : Aussage A und Aussage B

$A \vee B$  : Aussage A oder B (oder beide)

$\neg A$  : das Gegenteil von Aussage A (Negation)

Je nach Wahrheitsbelegung von A und B können wir die Wahrheitsbelegung von obigen Aussagen definieren:

A	B	A $\wedge$ B
w	w	w
w	f	f
f	w	f
f	f	f

Wenn sowohl aus A die Aussage B folgt als auch B die Aussage A impliziert, dann heissen A und B **äquivalent** bzw. **A gilt genau dann, wenn B gilt** und man schreibt:

$$A \Leftrightarrow B.$$

Wir können  $A \Leftrightarrow B$  auch als die Aussage  $(A \Rightarrow B) \wedge (B \Rightarrow A)$  auffassen.

- Übung 2.2.** 1. Stelle die Aussagen  $A \vee B$ ,  $\neg A$ ,  $A \Leftrightarrow B$  abhängig von der Wahrheitsbelegung von  $A$  und  $B$  mittels Wahrheitstafeln dar.
2. Zeige:  $(A \Rightarrow B) \Leftrightarrow (\neg B \Rightarrow \neg A)$ .

Einige wichtige Beweistechniken finden sich hier:

**Bemerkung 2.3.** Es gibt verschiedene Beweistechniken, um eine Aussage der Form  $A \Rightarrow B$  zu beweisen.

1. Direkter Beweis: Unter der Annahme, dass  $A$  wahr ist, schliesst man direkt, dass auch  $B$  wahr sein muss.

Beispiel:

$$n \in \mathbb{N} \Rightarrow n + (n + 1) + (n + 2) \text{ durch } 3 \text{ teilbar.}$$

Beweis: der Nachfolger der Zahl  $n \in \mathbb{N}$  ist die natürliche Zahl  $n + 1$ . Wir rechnen:  $n + (n + 1) + (n + 2) = 3n + 3 = 3 \cdot (n + 1)$  und somit ist diese Implikation wahr.

2. Beweis per Kontraposition:  $\neg B \Rightarrow \neg A$ .

Beispiel:  $a^2$  ungerade für  $a \in \mathbb{N} \Rightarrow a$  ungerade. Angenommen es gilt  $a$  gerade. Somit  $\exists k \in \mathbb{N}$  mit  $a = 2 \cdot k$ . Wir rechnen  $a^2 = (2 \cdot k)^2 = 2^2 \cdot k^2$  und somit ist  $a^2$  durch 2 teilbar also eine gerade Zahl.

3. Beweis per Widerspruch: Wir nehmen an, dass die Aussage  $A \Rightarrow B$  falsch ist und führen dies zum Widerspruch. Dies bedeutet (siehe Wahrheitstafeln aus den Übungen), dass  $A$  wahr ist,  $B$  jedoch falsch ist, und dies zusammen zu einem Widerspruch gebracht wird.

Beispiel:  $a^2$  ungerade für  $a \in \mathbb{N} \Rightarrow a$  ungerade. Sei nun  $a^2$  ungerade und  $a$  gerade. Wie vorhin erhalten wir  $a^2 = (2 \cdot k)^2 = 2^2 \cdot k^2$  und somit  $a^2$  gerade, Widerspruch!

## 2.2 Mengen und Abbildungen

Eine intuitive Definition einer Menge wurde von Georg Cantor gegeben: **Eine Menge ist eine Zusammenfassung von bestimmten unterscheidbaren Objekten unserer Anschauung oder unseres Denkens zu einem Ganzen.**

Sei eine Menge mit  $M$  bezeichnet. Ein Objekt  $x$  dieser Menge wird als **Element von  $M$**  bezeichnet und wir schreiben mathematisch  $x \in M$ . Falls  $x$  kein Element von  $M$  ist, schreiben wir  $x \notin M$ . Endliche Mengen können durch Angabe ihrer Elemente angegeben werden in der Form

$$M := \{x_1, x_2, \dots, x_n\}$$

wobei  $n \in \mathbb{N}$  dann die Anzahl der Elemente ist, falls die Elemente **paarweise verschieden sind**, d.h.,  $x_i \neq x_j \forall i \neq j$ . Die einfachste unendliche Menge sind die **natürlichen Zahlen**:

$$\mathbb{N} := \{1, 2, 3, \dots\}.$$

Unendliche abzählbare Mengen können in der Form

$$M := \{x_1, x_2, \dots\}$$

beschrieben werden, das heißt sie werden durch die natürlichen Zahlen fortlaufend indiziert.

**Beispiel 2.4.** Wir geben einige Beispiele von solchen Mengen:

- Leere Menge:  $\emptyset$
- Menge  $\mathbb{N} = \{1, 2, 3, \dots\}$  der natürlichen Zahlen.
- Menge  $\mathbb{Z} = \{\dots, -2, -1, 0, 1, 2, \dots\}$  der ganzen Zahlen.
- Menge  $\mathbb{Q} = \{\frac{z}{n} \mid z \in \mathbb{Z}, n \in \mathbb{N}\}$  der rationalen Zahlen.
- Menge aller Primzahlen zwischen 2 und 10 (also  $M = \{3, 5, 7\}$ ).

Die natürlichen Zahlen können miteinander addiert und multipliziert werden (wie aus der Schule bekannt) und weiterhin gilt, dass  $\mathbb{N}$  **total geordnet** ist, d.h. wir führen eine Relation<sup>1</sup>  $<$  zwischen zwei Elementen ein, welche folgendermaßen definiert ist

$$m < n \Leftrightarrow \exists k \in \mathbb{N} : m + k = n.$$

Die Relation  $<$  definiert eine (starke/strikte) totale Ordnung auf  $\mathbb{N}$ , also eine transitive Relation, sodass für alle  $m, n \in \mathbb{N}$  genau eines von  $m < n$ ,  $m = n$  und  $m > n$  gilt. Wir wollen nun folgendes Prinzip als wahr annehmen:

**| Annahme 2.5.** Jede nichtleere Teilmenge von  $\mathbb{N}$  hat ein kleinstes Element.

Mit diesem Prinzip können wir das Beweisprinzip der **vollständigen Induktion** einführen.

**Satz 2.6 (Vollständige Induktion).** Für  $n \in \mathbb{N}$  sei  $A(n)$  eine Aussage. Folgende Bedingungen seien erfüllt:

- Induktionsanfang:  $A(1)$  ist wahr.
- Induktionsschluss: Sei  $n \in \mathbb{N}$ : Wenn  $A(n)$  wahr ist, dann ist auch  $A(n + 1)$  wahr.

Dann ist  $A(n)$  wahr für alle  $n \in \mathbb{N}$ .

*Beweis.* Wir zeigen dass das Prinzip der vollständigen Induktion aus dem Prinzip des kleinsten Elements folgt. Definiere dazu die Menge

$$M := \{n \in \mathbb{N} \mid A(n) \text{ ist falsch.}\}.$$

<sup>1</sup>Vgl. Kapitel 2.2.1, wo Relationen und Abbildungen eingeführt werden.

Falls  $M = \emptyset$ , ist nichts mehr zu zeigen. Sei also  $M \neq \emptyset$ . Dann ist natürlich  $M$  eine Teilmenge der natürlichen Zahlen und hat mit Annahme 2.5 ein kleinstes Element, sagen wir  $k \in \mathbb{N}$ . Da  $A(1)$  nach Voraussetzung wahr ist, muss  $k > 1$  gelten. Setze  $\ell := k - 1 \in \mathbb{N}$ . Nach Definition von  $M$  ist  $A(\ell)$  wahr und mit dem Induktionsschluss auch  $A(\ell+1) = A(k)$ , Widerspruch. Also muss  $M = \emptyset$  gelten.  $\square$

**Definition 2.7.** Seien  $X, Y$  zwei Mengen.  $X$  ist eine Teilmenge von  $Y$ , d.h.  $X \subset Y$ , genau dann, wenn gilt

$$X \subset Y : \Leftrightarrow x \in X \Rightarrow x \in Y.$$

Der Doppelpunkt vor  $\Leftrightarrow$  ist so zu verstehen, dass das Objekt (oder in unserem Fall eine wahre Aussage) links vom Doppelpunkt definiert wird. Die Negation von  $X \subset Y$  wird geschrieben als  $X \not\subset Y$ . Es gilt

$$X = Y \Leftrightarrow X \subset Y \text{ und } Y \subset X.$$

Ist  $X \subset Y$  und  $X \neq Y$  so schreiben wir  $X \subsetneq Y$ .

Eine besondere Menge stellt die **leere Menge**  $\emptyset = \{ \}$  dar. Die leere Menge ist Teilmenge jeder Menge  $Y$ , weil die Aussage  $\forall x \in \emptyset : x \in Y$  wahr ist.

Teilungen werden häufig über bestimmte Eigenschaften der Elemente definiert, z.B. in der Form

$$\{x \in X : \text{Eigenschaft } E(x) \text{ ist erfüllt}\}$$

Die **Menge aller Teilmengen** einer Menge  $Y$  heisst **Potenzmenge**:

$$\mathcal{P}(Y) := \{X : X \subset Y\}.$$

**Definition 2.8 (Mengenoperationen).** Für zwei Mengen  $X, Y$  definieren wir:

1. Vereinigung  $X \cup Y$  von  $X$  und  $Y$  als

$$X \cup Y := \{x : (x \in X) \vee (x \in Y)\}$$

Für Indexmengen  $I$  können wir die Vereinigung von Mengen  $X_i, i \in I$  definieren als:

$$\cup_{i \in I} X_i := \{x : (\exists i \in I : x \in X_i)\}.$$

2. Durchschnitt  $X \cap Y$  von  $X$  und  $Y$  als

$$X \cap Y := \{x : (x \in X) \wedge (x \in Y)\}$$

Für nicht-leere Indexmengen  $I$  können wir den Durchschnitt von Mengen  $X_i, i \in I$  definieren als:

$$\cap_{i \in I} X_i := \{x : (\forall i \in I : x \in X_i)\}.$$

3. Differenzmenge  $X \setminus Y$  von  $X$  und  $Y$  als

$$X \setminus Y := \{x : (x \in X) \wedge (x \notin Y)\}$$

4. Kartesische Produkt  $X \times Y$  von  $X$  und  $Y$  als

$$X \times Y := \{(x, y) : (x \in X) \wedge (y \in Y)\}$$

In natürlicher Weise lässt sich das kartesische Produkt auf  $n$  Mengen  $X_1, \dots, X_n$  erweitern:

$$X_1 \times \dots \times X_n := \{(x_1, \dots, x_n) : (x_i \in X_i) \forall i \in \{1, \dots, n\}\}$$

Ist  $X_1 = \dots = X_n$ , so schreiben wir  $X^n := X \times \dots \times X$ .

**Definition 2.9.** Eine **Abbildung**  $f$  zwischen zwei Mengen  $X$  und  $Y$  ist eine **Zuordnung** von Elementen aus  $X$  auf Elemente von  $Y$ , die jedem Element  $x \in X$  ein **eindeutiges** Element  $f(x) \in Y$  zuordnet. Man schreibt

$$f : X \rightarrow Y, x \mapsto f(x).$$

Hier heisst  $X$  der Definitionsbereich (oder Urbildbereich) und  $Y$  der Wertebereich (oder Zielmenge) von  $f$ .

Wir können mit einer Abbildung  $f : X \rightarrow Y$  auch Teilmengen  $M \subset X$  abbilden bzw. für Mengen  $N \subset Y$  ihre Urbilder unter  $f$  angeben:

$$f(M) := \{y \in Y : \exists x \in M \text{ mit } f(x) = y\} \subset Y, f^{-1}(N) := \{x \in X : f(x) \in N\} \subset X.$$

Hier heisst  $f(M) \subset Y$  das **Bild** von  $M \subset X$  unter  $f$  und  $f^{-1}(N) \subset X$  das **Urbild** von  $N \subset Y$  unter  $f$ . Für  $M \subset X$  ist die Abbildung

$$f|_M : M \rightarrow Y, x \mapsto f(x),$$

die **Beschränkung** von  $f$  auf  $M$ .

Für zwei Abbildungen  $f : X \rightarrow Y$  und  $g : X \rightarrow Y$  definieren wir

$$f = g :\Leftrightarrow \forall x \in X : f(x) = g(x).$$

Mit  $\text{Abb}(X, Y)$  bezeichnen wir die **Menge der Abbildungen** von  $X$  nach  $Y$ .

Für einelementige Mengen  $N = \{y\} \subset Y$  schreibt man oft auch kurz  $f^{-1}(y) := f^{-1}(\{y\})$ . Aber Vorsicht:  $f^{-1}$  ist trotzdem keine Abbildung von  $Y$  nach  $X$ , da auch  $f^{-1}(y)$  eine Teilmenge von  $X$  (und kein Element von  $X$ ) ist. Insbesondere kann  $f^{-1}(y)$  mehrere Elemente enthalten oder leer sein.

Wir werden nun wichtige Eigenschaften von Abbildungen festhalten.

**Definition 2.10** (Eigenschaften von Abbildungen). Eine Abbildung  $f : X \rightarrow Y$  heisst

- **injektiv**, falls für alle  $x, x' \in X$  gilt:  $f(x) = f(x') \Rightarrow x = x'$ .
- **surjektiv**, falls es für alle  $y \in Y$  ein  $x \in X$  gibt, mit  $f(x) = y$ , d.h.,  $f(X) = Y$ .
- **bijektiv**, falls  $f$  injektiv und surjektiv ist.

Falls  $f$  bijektiv ist, so gibt es zu jedem  $y \in Y$  genau ein  $x \in X$  mit  $f(x) = y$  und in diesem Fall können wir eine **Umkehrabbildung**

$$f^{-1} : Y \rightarrow X, y \mapsto x = f^{-1}(y) \text{ mit } f(x) = y$$

definieren.<sup>2</sup> Wir halten folgendes Lemma fest.

**Lemma 2.11.** Wenn  $f : X \rightarrow Y$  bijektiv ist, dann ist auch  $f^{-1} : Y \rightarrow X$  bijektiv.

*Beweis.* Wir zeigen zuerst, dass  $f^{-1}$  injektiv ist. Dazu sei  $f^{-1}(y) = f^{-1}(y')$ . Wir erhalten

$$y = f(f^{-1}(y)) = f(f^{-1}(y')) = y'.$$

Um Surjektivität von  $f^{-1}$  nachzuweisen betrachten wir  $x \in X$  beliebig. Wir erhalten  $x = f^{-1}(f(x))$  und somit ist  $f^{-1}$  surjektiv.  $\square$

Im obigen Beweis hatte wir zwei Abbildungen hintereinander ausgeführt. Diese Operation werden wir nun formal einführen und ein eigenes Symbol dafür spendieren.

**Definition 2.12.** Für zwei Abbildungen  $g : X \rightarrow Y$  und  $f : Y \rightarrow Z$  definieren wir die **Hintereinanderschaltung**  $f \circ g$  von  $f$  und  $g$  durch

$$f \circ g : X \rightarrow Z, x \mapsto f(g(x)).$$

Eine wichtige Abbildung ist die sogenannte **Identität**, welche für eine Menge  $X$  als  $\text{id}_X$  definiert ist:

$$\text{id}_X : X \rightarrow X, x \mapsto x.$$

Wir erhalten eine alternative Charakterisierung von Injektivität, Surjektivität und Bijektivität.

**Lemma 2.13.** Sei  $f : X \rightarrow Y$  eine Abbildung mit  $X \neq \emptyset$ . Dann gilt:

1.  $f$  injektiv  $\Leftrightarrow$  es gibt eine Abbildung  $g : Y \rightarrow X$  mit  $g \circ f = \text{id}_X$ .
2.  $f$  surjektiv  $\Leftrightarrow$  es gibt eine Abbildung  $g : Y \rightarrow X$  mit  $f \circ g = \text{id}_Y$ .

<sup>2</sup>Ausnahmsweise verwenden wir hier eine Doppelnotation,  $f^{-1}$  sowohl für Urbilder als auch für Umkehrabbildungen im Falle bijektiver Abbildungen.

3.  $f$  bijektiv  $\Leftrightarrow$  es gibt eine Abbildung  $g : Y \rightarrow X$  mit  $g \circ f = \text{id}_X$  und  $f \circ g = \text{id}_Y$ . In diesem Fall ist  $g = f^{-1}$ .

*Beweis.* Zu 1:  $\Rightarrow$ : Sei  $f$  injektiv, d.h., zu jedem  $y \in f(X)$  gibt es genau ein  $x$  mit  $f(x) = y$  und wir definieren  $g(y) := x$ . Für  $y \in Y \setminus f(X)$  definieren wir  $g(y) := x_0$  für  $x_0 \in X$  beliebig. Dies ergibt eine Abbildung  $g : Y \rightarrow X$  mit  $g \circ f = \text{id}_X$ .

$\Leftarrow$ : Sei umgekehrt  $g : Y \rightarrow X$  mit  $g \circ f = \text{id}_X$  gegeben. Für  $x, x' \in X$  mit  $f(x) = f(x')$  rechnen wir nach

$$x = \text{id}_X(x) = g \circ f(x) = g \circ f(x') = \text{id}_X(x') = x'.$$

□

**Übung 2.14.** Beweise die verbliebenen Punkte 2. und 3.

Eine wichtige Klasse von Abbildungen sind die **Projektionen**.

**Beispiel 2.15.** Eine Abbildung der Form

$$\pi_i : X_1 \times \cdots \times X_n \rightarrow X_i, (x_1, \dots, x_n) \mapsto x_i$$

heißt **Projektion** und bildet Tupel der Form

$$x = (x_1, \dots, x_n) \text{ mit } x_i \in X_i, i = 1, \dots, n,$$

auf ihre  $i$ -te Komponente ab.

### 2.2.1 Äquivalenzrelationen

Für manche Anwendungen möchte man je zwei Elemente einer Menge bezüglich bestimmter Charakteristika "sortieren". Zum Beispiel sind im Supermarkt alle Äpfel, Bananen, Birnen, etc. jeweils in einem eigenen Bereich (z.B. Apfelkorb, Bananenkorb etc.) zusammengefasst. Je zwei Elemente können nun verglichen werden und entweder sind sie von der gleichen Charakteristik (z.B. zwei Äpfel) oder eben nicht. Eine solche Zerlegung von Mengen kann mittels **Äquivalenzrelationen** modelliert werden. Man schreibt allgemein  $x \sim y$  für eine **Relation** zwischen zwei Elementen  $x, y \in X$  einer Menge  $X \neq \emptyset$  und für unser Beispiel oben würde  $x \sim y$  bedeuten, dass  $x$  und  $y$  die gleiche Charakteristik besitzen (wie z.B. zwei Äpfel). Eine Relation ist durch ihren **Graphen**  $R \subset X \times X$  beschrieben, wobei

$$(x, y) \in R \Leftrightarrow x \sim y. \quad (2.1)$$

Eine Relation auf  $X$  kann somit durch die Angabe einer solchen Teilmenge  $R$  beschrieben werden, wobei das Zeichen  $\sim$  dann durch (2.1) erklärt ist.

**Definition 2.16.** Eine Relation  $\sim$  heisst **Äquivalenzrelation** auf einer Menge  $X \neq \emptyset$ , wenn für beliebige  $x, y, z \in X$  gilt:

A1:  $x \sim x$  (Reflexivität)

A2:  $x \sim y \Rightarrow y \sim x$  (Symmetrie)

A3:  $(x \sim y \text{ und } y \sim z) \Rightarrow x \sim z$  (Transitivität)

Für  $x \sim y$  sagt man auch:  $x$  und  $y$  sind **äquivalent**.

Wir kommen nun auf unser "Obst" Beispiel zurück und werden für die speziellen Teilmengen der Äpfel, Bananen, etc. einen eigenen Begriff einführen: **Äquivalenzklassen**.

**Definition 2.17.** Für eine Äquivalenzrelation  $\sim$  auf einer Menge  $X \neq \emptyset$  heisst eine Teilmenge  $A \subset X$  eine **Äquivalenzklasse**, falls gilt:

1.  $A \neq \emptyset$

2.  $x, y \in A \Rightarrow x \sim y$

3.  $x \in A, y \in X, x \sim y \Rightarrow y \in A$

**Lemma 2.18.** Sei  $\sim$  eine Äquivalenzrelation auf einer Menge  $X \neq \emptyset$ . Dann gehört jedes Element  $x \in X$  zu genau einer Äquivalenzklasse  $A$  bezüglich  $\sim$ . Insbesondere trägt das Mengensystem der Äquivalenzklassen eine spezielle **laminare Struktur**, das heisst,

$$A, A' \text{ Äquivalenzklassen} \Rightarrow A = A' \text{ oder } A \cap A' = \emptyset.$$

*Beweis.* Für ein festes  $a \in X$  definiere

$$A := \{x \in X : x \sim a\}.$$

Wir werden zeigen, dass  $A$  eine Äquivalenzklassen ist, die  $a$  enthält. Weil  $\sim$  reflexiv ist, gilt natürlich  $a \in A$  und somit  $A \neq \emptyset$ . Für  $x, y \in A$  gilt  $x \sim a$  und  $y \sim a$ , also folgt mit A2 und A3 dass  $x \sim y$  gilt. Sei nun  $x \in A, y \in X$  und  $x \sim y$ . Es folgt, dass  $x \sim a$  und mit  $x \sim y$  sowie A2 und A3 folgt, dass auch  $y \sim a$  gilt, was  $y \in A$  zeigt. Also ist  $A$  in der Tat eine Äquivalenzklasse die  $a$  enthält.

Wir müssen noch zeigen, dass zwei Äquivalenzklassen  $A, A'$  entweder gleich oder disjunkt sind. Angenommen  $a \in A' \cap A$ . Für  $x \in A$  gilt  $a \sim x$  und mit  $a \in A'$  gilt auch  $x \in A'$ , also  $A \subset A'$ . Dieses Argument können wir auch mit  $x \in A'$  beginnen lassen, was dann zu  $A' \subset A$  führt, also insgesamt  $A = A'$ .  $\square$

Der obige Beweis nutzt die Konstruktion einer Äquivalenzklasse, in der für ein festes  $a \in X$  die Menge  $\{x \in X : x \sim a\}$  definiert wird. Hierfür werden wir nun ein eigene

Notation benutzen:

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

Eine Äquivalenzrelation  $R$  liefert also eine Zerlegung der Grundmenge  $X$  in disjunkte Äquivalenzklassen. Die Menge der Äquivalenzklassen wird mit  $X/R$  oder alternativ auch mit  $X/\sim$  bezeichnet und man nennt sie die **Quotientenmenge**. Ordnet man jedem Element  $a \in X$  die zugehörige Äquivalenzklasse  $[a]$  zu, in der das Element  $a$  enthalten ist, so erhält man eine kanonische Abbildung

$$X \rightarrow X/\sim, a \mapsto [a].$$

Das Urbild eines Elements  $[a] \in X/\sim$  ist wieder  $[a] \subseteq X$  und jedes  $x \in [a]$  heißt **Repräsentant** der Äquivalenzklasse  $[a]$ .

**Beispiel 2.19 (Restklassen).** Das Rechnen mit Restklassen ist im Alltag verankert. Wenn es 9 Uhr morgens ist, dann sagen wir häufig, dass es in 4 Stunden 1 Uhr sein wird. Die Uhrzeit wird also häufig modulo 12 dargestellt. Sei nun  $n \in \mathbb{N}$ . Wir definieren die Äquivalenzrelation  $\sim_n$  auf  $\mathbb{Z}$  wie folgt:

$$x \sim_n x' :\Leftrightarrow x = x' \pmod n \Leftrightarrow n \text{ teilt } x - x' \Leftrightarrow \exists a \in \mathbb{Z} : x - x' = an.$$

Wir sehen, dass es genau  $n$  Äquivalenzklassen der Form  $[0], [1], \dots, [n-1]$  gibt. Der entsprechende Quotientenraum ist gegeben durch  $\mathbb{Z}_n := \mathbb{Z}/\sim_n = \{[0], [1], \dots, [n-1]\}$ .

## 2.3 Gruppen, Ringe und Körper

Nun wenden wir uns der Frage zu, wie man mit Elementen einer Menge "rechnen" kann. Aus der Schule ist bekannt, dass man Zahlen addieren, subtrahieren oder multiplizieren kann. Diese Operationen werden wir nun etwas allgemeiner auf abstrakten Mengen axiomatisch einführen.

### 2.3.1 Gruppen

Dazu sei  $G$  eine Menge. Eine Verknüpfung von zwei Elementen  $a, b \in G$  wird als eine Abbildung  $*$  :  $G \times G \rightarrow G$  aufgefasst, die  $(a, b)$  auf ein Element  $*(a, b) \in G$  abbildet:

$$* : G \times G \rightarrow G, (a, b) \mapsto *(a, b).$$

Zur Vereinfachung schreiben wir im Folgenden  $a * b := *(a, b)$ .

Wir geben ein einfaches Beispiel.

**Beispiel 2.20.** • Sei  $G = \mathbb{N}$  und  $a, b \in \mathbb{N}$ . Definiere  $a * b := a + b$ . Es ist leicht einzusehen, dass tatsächlich  $a * b \in G$  gilt für alle  $(a, b) \in G \times G$ .

- Sei  $G = \mathbb{R}$  und  $a, b \in \mathbb{R}$ . Definiere  $a * b := a \cdot b$ .

Mengen mit Verknüpfungen  $*$ , die besondere Eigenschaften aufweisen, werden im Folgenden mit eigenen Definitionen ausgezeichnet.

**Definition 2.21.** Ein Tupel  $(G, *)$  heisst **Gruppe**, wenn die folgenden Axiome erfüllt sind:

**G.1**  $(a * b) * c = a * (b * c)$  für alle  $a, b, c \in G$  (Assoziativgesetz)

**G.2** Es gibt ein **neutrales Element**  $e \in G$  mit den folgenden Eigenschaften

(i)  $e * a = a$  für alle  $a \in G$ .

(ii) Zu jedem  $a \in G$  gibt es ein  $a^{-1} \in G$  mit  $a^{-1} * a = e$ . Ein Element  $a^{-1}$  wird zu **a inverses Element** genannt.

Eine Gruppe heisst **abelsch** oder **kommutativ**, falls außerdem noch  $a * b = b * a$  für alle  $a, b \in G$  gilt.

Mit **G.1** kann man nun bei Mehrfachprodukten die Klammern einfach weglassen:

$$a * (b * c) = (a * b) * c = a * b * c.$$

Bisher wurde a priori nicht verlangt, dass  $e$  oder  $a^{-1}$  für ein gegebenes  $a \in G$  eindeutig sein müssen. Dies folgt aber unmittelbar, denn es gelten folgende Eigenschaften für eine Gruppe.

**Lemma 2.22.** Sei  $(G, *)$  eine Gruppe. Dann gilt:

1. Für alle  $a \in G$  gilt:  $a * e = a$  für ein neutrales Element  $e \in G$ . Das neutrale Element  $e \in G$  ist eindeutig bestimmt.
2. Für jedes  $a \in G$  ist das Inverse  $a^{-1}$  eindeutig bestimmt. Für alle  $a, b \in G$  gilt:
  - $a * a^{-1} = e$
  - $(a^{-1})^{-1} = a$
  - $(a * b)^{-1} = b^{-1} * a^{-1}$
3. Es gelten folgende Kürzungsregeln:
  - $a * x = a * y \Rightarrow x = y$
  - $x * a = y * a \Rightarrow x = y$

*Beweis.* Nach **G.2** gibt es ein neutrales Element  $e \in G$ , so dass für es für beliebiges  $a \in G$  ein Inverses  $a^{-1}$  und zu  $a^{-1}$  wiederum ein Inverses  $(a^{-1})^{-1}$  mit  $(a^{-1})^{-1} * a^{-1} = e$  gibt. Daraus ergibt sich:

$$a * a^{-1} = e * (a * a^{-1}) \quad (\text{wegen (i)})$$

$$\begin{aligned}
&= ((a^{-1})^{-1} * a^{-1}) * (a * a^{-1}) && \text{(siehe oben)} \\
&= (a^{-1})^{-1} * (a^{-1} * (a * a^{-1})) && \text{(wegen G.1)} \\
&= (a^{-1})^{-1} * ((a^{-1} * a) * a^{-1}) && \text{(wegen G.1)} \\
&= (a^{-1})^{-1} * (e * a^{-1}) && \text{(wegen (ii))} \\
&= (a^{-1})^{-1} * a^{-1} && \text{(wegen (i))} \\
&= e. && \text{(wegen (ii))}
\end{aligned}$$

Somit folgt

$$a * e = a * (a^{-1} * a) = (a * a^{-1}) * a = e * a = a.$$

Ist  $e'$  ein weiteres neutrales Element, so gilt

$$e' = e' * e \stackrel{(i)}{=} e.$$

Mit diesen Überlegungen haben wir 1. und den ersten Listenpunkt von 2. bewiesen.

Nun zeigen wir die restlichen Punkte von 2. Ist  $a'$  ein anderes Inverses von  $a$ , so gilt:

$$a' = a' * e \quad \text{(wegen 1.)}$$

$$= a' * (a * a^{-1}) \quad (2.2)$$

$$= (a' * a) * a^{-1} \quad (2.3)$$

$$= e * a^{-1} \quad (2.4)$$

$$= a^{-1}. \quad (2.5)$$

Somit ist Eindeutigkeit des Inversen gezeigt. Wir sehen auch, dass mit  $a * a^{-1} = e$  folgt, dass  $a$  inverses Element zu  $a^{-1}$  ist, d.h.,  $(a^{-1})^{-1} = a$ . Des Weiteren gilt:

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

Der Beweis der Kürzungsregeln beruht nun auf der Verknüpfung von links bzw. rechts der ersten Gleichung mit dem inversen Element  $a^{-1}$ .  $\square$

Wir geben nun einige (Nicht-)Beispiele von Gruppen an.

**Beispiel 2.23.** 1.  $G = \{-1, 1\}$  mit  $a * b := a \cdot b$ .

2.  $G = \mathbb{Z}$  mit  $a * b := a + b$ .

3.  $G = \mathbb{R}^n, n \geq 1$  mit  $a * b := a + b$ .

4.  $G = \mathbb{R}^* := \mathbb{R} \setminus \{0\}$  mit  $a * b := a \cdot b$ .

Keine Gruppen sind:

1.  $G = \mathbb{N}$  mit  $a * b := a + b$ , da mit neutralem Element 0 die Inversen nicht existieren.

2.  $G = \mathbb{R}$  mit  $a * b := a \cdot b$ , da das neutrale Element 1 ist, aber 0 kein Inverses besitzt.

**Beispiel 2.24.** Sei  $X \neq \emptyset$  und  $G := \{f : X \rightarrow X, f \text{ ist bijektiv}\}$ . Dann ist  $(G, *)$  mit  $f * f' := f \circ f'$  eine Gruppe. Der Ausdruck  $f \circ f'$  ist in Def. 2.12 erklärt!

**Übung 2.25.** Verifiziere in Beispiel 2.24, dass  $(G, *)$  eine Gruppe ist.

**Beispiel 2.26 (Restklassen).** Für  $[m], [m'] \in \mathbb{Z}_n, n \in \mathbb{N}$  definiere die Addition und Multiplikation gemäß:

$$[m] + [m'] := [m + m'] \text{ und } [m] \cdot [m'] := [mm'].$$

**Übung 2.27.** • Zeige, dass Addition und Multiplikation auf  $\mathbb{Z}_n$  wohldefiniert (das heißt unabhängig von der Wahl der Repräsentanten) ist.

- Zeige, dass  $(\mathbb{Z}_n, +)$  eine abelsche Gruppe ist.
- Zeige, dass  $(\mathbb{Z}_n, \cdot)$  i.A. keine Gruppe ist.
- Zeige, dass  $(\mathbb{Z}_p^*, \cdot)$  für  $p$  Primzahl eine abelsche Gruppe ist. Hier ist  $\mathbb{Z}_p^* := \mathbb{Z}_p \setminus \{0\}$ .

### 2.3.2 Ringe

Bei Gruppen gibt es nur eine einzige Verknüpfung. Beim Rechnen mit Zahlen kennen wir jedoch (mindestens) zwei elementare Verknüpfungen, nämlich Addition und Multiplikation. Mit einer Verallgemeinerung von Gruppen auf zwei Verknüpfungen werden wir uns nun beschäftigen.

**Definition 2.28.** Ein Tupel  $(R, +, \cdot)$  mit zwei Verknüpfungen  $+$  und  $\cdot$  der Form

$$+ : R \times R \rightarrow R, (a, b) \mapsto a + b \quad (\text{Addition})$$

$$\cdot : R \times R \rightarrow R, (a, b) \mapsto a \cdot b \quad (\text{Multiplikation})$$

heißt **Ring**, wenn folgendes gilt:

R1:  $(R, +)$  ist eine Abelsche Gruppe.

R2:  $(R, \cdot)$  erfüllt **G.1**, d.h., die Multiplikation  $\cdot$  ist assoziativ.

R3: Es gelten die **Distributivgesetze**, d.h., für alle  $a, b, c \in R$  gilt

$$a \cdot (b + c) = a \cdot b + a \cdot c \text{ und } (a + b) \cdot c = a \cdot c + b \cdot c. \quad (2.6)$$

Ein Ring  $R$  heißt kommutativ, wenn stets  $a \cdot b = b \cdot a$  gilt.

Ein Element  $1 \in R$  heißt **Einselement** (oder neutrales Element bzgl. Multiplikation), wenn  $1 \cdot a = a \cdot 1 = a$  für alle  $a \in R$  gilt. Um unnötige Klammern zu vermeiden, legen wir fest, dass bei der Anwendung der Verknüpfungen die Multiplikation Vorrang vor der Addition hat.

**Beispiel 2.29.** •  $(\mathbb{Z}, +, \cdot)$  ist ein kommutativer Ring mit Einselement.

- Für  $n \in \mathbb{N}$  ist  $(\mathbb{Z}_n, +, \cdot)$  ein kommutativer Ring mit Einselement.

**Übung 2.30.** Verifiziere die obigen Aussagen.

Man beachte, dass es in Ringen bezüglich der Multiplikation kein Inverses geben muss, siehe zum Beispiel den Ring der ganzen Zahlen  $(\mathbb{Z}, +, \cdot)$ . Wenn wir hingegen den Ring der reellen Zahlen betrachten  $(\mathbb{R}, +, \cdot)$  ist dies jedoch erfüllt. Dies führt uns zu dem Konzept der **Körper** (fields auf Englisch).

### 2.3.3 Körper

**Definition 2.31.** Ein Tupel  $(K, +, \cdot)$  mit zwei Verknüpfungen  $+$  und  $\cdot$  der Form

$$+ : K \times K \rightarrow K, (a, b) \mapsto a + b \quad (\text{Addition})$$

$$\cdot : K \times K \rightarrow K, (a, b) \mapsto a \cdot b \quad (\text{Multiplikation})$$

heißt **Körper**, wenn folgendes gilt:

K1:  $(K, +)$  ist eine Abelsche Gruppe. Das neutrale Element der Addition wird mit  $0 \in K$  bezeichnet und das zu  $a \in K$  Inverse mit  $-a$ .

K2:  $(K^*, \cdot)$  ist eine Abelsche Gruppe, wobei  $K^* := K \setminus \{0\}$ . Das neutrale Element der Multiplikation ist hier  $1 \in K^*$ . Das zu  $a \in K^*$  inverse Element wird mit  $a^{-1}$  oder auch  $1/a$  bezeichnet und man schreibt  $b/a = a^{-1}b = ba^{-1}$ . Beachte, dass wir soeben und auch im Folgenden das explizite Symbol  $\cdot$  der Multiplikation weglassen und einfach  $a \cdot b = ab$  schreiben.

K3: Es gelten die **Distributivgesetze**, d.h., für alle  $a, b, c \in K$  gilt

$$a \cdot (b + c) = a \cdot b + a \cdot c \quad \text{und} \quad (a + b) \cdot c = a \cdot c + b \cdot c. \quad (2.7)$$

Wir erhalten folgende wichtige Eigenschaften von Körpern.

**Lemma 2.32.** Sei  $(K, +, \cdot)$  ein Körper und  $a, b, x, y \in K$  beliebig. Dann gilt:

1.  $1 \neq 0$ , d.h. auch, dass ein Körper immer mindestens zwei Elemente hat.
2.  $0 \cdot a = a \cdot 0 = 0$
3.  $a \cdot b = 0 \Leftrightarrow a = 0$  oder  $b = 0$
4.  $a \cdot (-b) = -(a \cdot b)$  und  $(-a) \cdot (-b) = a \cdot b$
5.  $x \cdot a = y \cdot a$  und  $a \neq 0 \Rightarrow x = y$

*Beweis.* 1. ist trivial weil  $0 \notin K^*$ .

Zu 2.:  $0a = (0 + 0)a \stackrel{(2.7)}{=} 0a + 0a \Rightarrow 0a = 0$  (Gleicher Trick für  $a0!$ ).

Zu 3.:  $\Rightarrow$ : Angenommen  $a \neq 0$  und  $b \neq 0$  und  $a \cdot b = 0$ . Somit gilt  $a, b \in K^*$  und weil  $(K^*, \cdot)$  eine Gruppe ist, gilt  $a \cdot b \in K^*$ , Widerspruch.

$\Leftarrow$ : Folgt direkt aus Punkt 2.

Zu 4.:

$$ab + a(-b) \stackrel{(2.7)}{=} a(b + (-b)) = a0 \stackrel{\text{Punkt 2.}}{=} 0$$

Somit gilt  $a(-b) = -ab$ . Wir erhalten weiterhin:

$$(-a)(-b) + a(-b) \stackrel{(2.7)}{=} (-a + a)(-b) = 0 \Rightarrow (-a)(-b) = -(a(-b)) = -(-ab) = ab,$$

wobei der vorletzte Schritt die vorher gezeigte Beziehung  $a(-b) = -ab$  nutzt.

Zu 5.: Angenommen  $x = 0$  oder  $y = 0$ . Dann folgt mit  $xa = ya = 0$ ,  $a \neq 0$  und Punkt 3., dass  $x = y = 0$  gelten muss. Also seien  $x, y \in K^*$ .

$$xa = ya \Rightarrow xa + (-ya) = 0.$$

Wegen  $-ya \stackrel{(K^*, \cdot) \text{ Abelsch}}{=} -(ay) \stackrel{\text{Punkt 4.}}{=} (-y)a$  erhalten wir

$$xa + (-y)a = 0 \stackrel{(2.7)}{\Rightarrow} (x + (-y))a = 0 \stackrel{a \neq 0}{\Rightarrow} x + (-y) = 0 \Rightarrow x = y.$$

□

**Beispiel 2.33.** Die rationalen und die reellen Zahlen bilden mit der Addition und Multiplikation (wie wir sie aus der Schule kennen) jeweils einen Körper, d.h.,  $(\mathbb{Q}, +, \cdot)$  und  $(\mathbb{R}, +, \cdot)$  sind Körper.

### 2.3.4 Komplexe Zahlen

Eine weiteres wichtiges Beispiel eines Körpers sind die **komplexen Zahlen**. Zur Motivation der komplexen Zahlen betrachten wir die Gleichung  $x^2 = -1$ , welche für  $x \in \mathbb{R}$  keine Lösung hat. Die Grundidee ist nun eine Erweiterung des Körpers  $(\mathbb{R}, +, \cdot)$  einzuführen, so dass die obige Gleichung in dem erweiterten Körper lösbar wird.

**Definition 2.34 (Komplexe Zahlen).** Setze  $\mathbb{C} := \mathbb{R} \times \mathbb{R}$  und definiere auf  $\mathbb{C}$  folgendermaßen eine Addition und Multiplikation gemäß:

$$+ : \mathbb{C} \times \mathbb{C} \rightarrow \mathbb{C}, (a, b) + (a', b') := (a + a', b + b') \quad (\text{Addition})$$

$$\cdot : \mathbb{C} \times \mathbb{C} \rightarrow \mathbb{C}, (a, b) \cdot (a', b') := (aa' - bb', ab' + a'b), \quad (\text{Multiplikation})$$

wobei auf der rechten Seite jeweils  $+$  und  $\cdot$  wie im Körper  $(\mathbb{R}, +, \cdot)$  verwendet wird.

| **Satz 2.35.**  $(\mathbb{C}, +, \cdot)$  ist ein Körper.

*Beweis.* Assoziativgesetz und Kommutativgesetz für die jeweiligen Gruppen  $(\mathbb{C}, +)$  und  $(\mathbb{C}^*, \cdot)$  ergeben sich jeweils mittels derer aus  $(\mathbb{R}, +)$  bzw.  $(\mathbb{R}^*, \cdot)$ . Das neutrale Element der Addition ist  $(0, 0)$ , das Inverse der Addition bzgl.  $z = (a, b)$  ist  $-z := (-a, -b)$ . Das Einselement der Multiplikation ist  $(1, 0)$  und das Inverse einer Zahl  $z = (a, b) \in \mathbb{C}^*$  ist gegeben durch

$$z^{-1} := \left( \frac{a}{a^2 + b^2}, \frac{-b}{a^2 + b^2} \right).$$

□

Da  $(a, 0) + (b, 0) = (a + b, 0)$  und  $(a, 0) \cdot (b, 0) = (ab, 0)$  gilt, kann man jede reelle Zahl  $a \in \mathbb{R}$  mit der komplexen Zahl  $(a, 0) \in \mathbb{C}$  identifizieren. In diesem Sinne kann  $\mathbb{R}$  dann als Teilmenge von  $\mathbb{C}$  aufgefasst werden. Per Konvention setzten wir  $i := (0, 1)$  und so erhält man für eine beliebige komplexe Zahl  $z = (a, b)$  die übliche Darstellung

$$z = (a, b) = (a, 0) + (0, 1)(b, 0) =: a + ib, \quad (a, b \in \mathbb{R}).$$

Für die komplexe Zahl  $i \in \mathbb{C}$  gilt:

$$i^2 := i \cdot i = (0, 1) \cdot (0, 1) = (-1, 0) = -1,$$

wobei die letzte Gleichung die Identifikation von Zahlen der Form  $(a, 0) \in \mathbb{C}$  als  $a \in \mathbb{R}$  benutzt. Also löst  $i \in \mathbb{C}$  wie vorab versprochen die Gleichung  $x^2 = -1$ . Die Zahl  $i \in \mathbb{C}$  wird auch als **imaginäre Zahl** bezeichnet.

| **Lemma 2.36.** Zu jeder Zahl  $z = (a, b) \in \mathbb{C}$  sind  $a, b \in \mathbb{R}$  die eindeutigen Zahlen, so dass gilt  $z = a + ib$ .

*Beweis.* Oben hatten wir schon gesehen, dass die Darstellung  $z = a + ib$  korrekt ist. Bleibt die Eindeutigkeit zu zeigen. Dazu seien  $a, b, a', b' \in \mathbb{R}$  mit  $z = a + ib = a' + ib'$ . Dann folgt:

$$(a, b) = (a, 0) + (0, b) = a + ib = a' + ib' = (a', 0) + (0, b') = (a', b').$$

Nach Definition des Kartesischen Produkts  $\mathbb{R} \times \mathbb{R}$  folgt, dass  $a = a'$  und  $b = b'$ . □

| **Definition 2.37.** Für eine komplexe Zahl  $z = a + ib \in \mathbb{C}$  mit  $a, b \in \mathbb{R}$  definieren wir:

1.  $\operatorname{Re}(z) := a \in \mathbb{R}$  (Realteil)
2.  $\operatorname{Im}(z) := b \in \mathbb{R}$  (Imaginärteil)
3.  $\bar{z} := a - ib$  (Komplex Konjugierte)

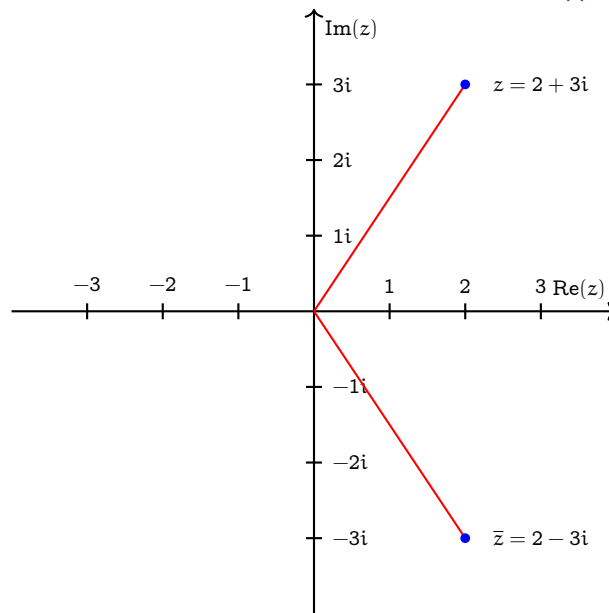


Abbildung 2.1: Illustration der komplexen Zahl  $z = 2 + 3i$  und ihrer Konjugierten  $\bar{z}$ .

**Übung 2.38.** Seien  $z, w \in \mathbb{C}$ . Beweise folgende Rechenregeln:

1.  $\overline{z + w} = \bar{z} + \bar{w}$ .
2.  $\overline{zw} = \bar{z}\bar{w}$ .
3.  $\overline{-z} = -\bar{z}$ .
4.  $\overline{\frac{1}{z}} = \frac{1}{\bar{z}}$ .
5.  $z = \bar{z} \Leftrightarrow z \in \mathbb{R}$ .
6.  $\overline{\bar{z}} = z$ .
7.  $z + \bar{z} = 2\operatorname{Re}(z), z - \bar{z} = 2i\operatorname{Im}(z)$ .
8.  $z\bar{z} = (\operatorname{Re}(z))^2 + (\operatorname{Im}(z))^2$ .

**Definition 2.39.** Für  $z = a + ib \in \mathbb{C}$  definiere den Betrag von  $z$  durch  $|z| := \sqrt{a^2 + b^2}$ .

Geometrisch entspricht  $|z|$  dem Euklidischen Abstand von  $(0, 0)$  zu  $z$  in der Gaußschen Zahlenebene.

**Lemma 2.40.** Seien  $z, w \in \mathbb{C}$ . Dann gilt:

1.  $|z| = |\bar{z}| = \sqrt{z\bar{z}}$
2.  $|\operatorname{Re}(z)| \leq |z|, |\operatorname{Im}(z)| \leq |z|$ .

3. Für  $z \in \mathbb{R}$  sind der komplexe und der reelle Betrag gleich.
4.  $|z| \geq 0$  und  $|z| = 0 \Leftrightarrow z = 0$ .
5.  $|-z| = |z|$ .
6.  $|zw| = |z||w|$ .

*Beweis.* Übung. □

**Satz 2.41 (Dreiecksungleichung).** Seien  $z, w \in \mathbb{C}$ . Dann gilt die Dreiecksungleichung:

$$|z + w| \leq |z| + |w|.$$

*Beweis.*

$$\begin{aligned}
 |z + w|^2 &\stackrel{\text{Lemma 2.40,1.}}{=} (z + w)\overline{(z + w)} \stackrel{\text{Übung 2.38,2.}}{=} (z + w)(\bar{z} + \bar{w}) \\
 &= z\bar{z} + z\bar{w} + \bar{z}w + w\bar{w} \stackrel{\text{Übung 2.38,7.}}{=} |z|^2 + 2\operatorname{Re}(z\bar{w}) + |w|^2 \\
 &\leq \stackrel{\text{Lemma 2.40,2.}}{|z|^2 + 2|z\bar{w}| + |w|^2} \\
 &= \stackrel{\text{Lemma 2.40,6.}}{|z|^2 + 2|z||w| + |w|^2} = (|z| + |w|)^2.
 \end{aligned}$$

□

## 2.4 Vektorräume

Nun wollen wir die Theorie des Rechnens in Körpern erweitern und auch **Vektoren** zulassen.

Dazu fangen wir zunächst mit einfachen Beispielen an.

**Beispiel 2.42.** Das Standardbeispiel eines Vektorraumes (Definition folgt später!) ist der **Standardraum**:

$$\mathbb{K}^n := \left\{ x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} : x_i \in \mathbb{K}, i = 1, \dots, n \right\}.$$

wobei  $\mathbb{K}$  ein Körper ist. Zum Beispiel erhalten wir den Spezialfall  $\mathbb{K}^n = \mathbb{R}^n$  für  $\mathbb{K} = \mathbb{R}$ . Auf  $\mathbb{K}^n$  können wir mit Hilfe der Verknüpfungen  $+$ ,  $\cdot$  auf  $\mathbb{K}$  zwei neue Verknüpfungen auf

$K^n \times K^n$  definieren um auch in  $K^n$  rechnen zu können:

$$\begin{aligned} + : K^n \times K^n &\rightarrow K^n, (x, y) \mapsto x + y && \text{(Vektor-Addition)} \\ \cdot : K \times K^n &\rightarrow K^n, (\lambda, x) \mapsto \lambda \cdot x, && \text{(Skalar-Multiplikation)} \end{aligned}$$

wobei

$$x + y := \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}, \lambda \cdot x := \begin{pmatrix} \lambda x_1 \\ \vdots \\ \lambda x_n \end{pmatrix} \text{ für } \lambda \in K.$$

Beachte, dass hier die Addition auf der rechten Seite durch die Addition im Körper  $K$  komponentenweise definiert ist. Die Multiplikation auf der rechten Seite ist durch die Multiplikation im Körper  $(K, \cdot)$  und hier wird das Symbol  $\cdot$  der Einfachheit weggelassen. Man nennt die Verknüpfung  $+$  bzgl.  $K^n \times K^n$  **innere Verknüpfung** und die Verknüpfung  $\cdot$  bzgl.  $K \times K^n$  **äußere Verknüpfung**.

Allgemeiner werden wir nun den Begriff des **Vektorraumes** definieren.

**Definition 2.43.** Sei  $K$  ein Körper. Eine Menge  $V$  zusammen mit einer inneren Verknüpfung

$$+ : V \times V \rightarrow V, (v, w) \mapsto v + w \quad \text{(Vektor-Addition)}$$

und einer äußere Verknüpfung

$$\cdot : K \times V \rightarrow V, (\lambda, x) \mapsto \lambda \cdot x, \quad \text{(Skalar-Multiplikation)}$$

heißt  **$K$ -Vektorraum**, falls folgendes gilt:

V1:  $(V, +)$  ist eine Abelsche Gruppe. Das neutrale Element der Addition wird mit  $0 \in V$  bezeichnet und das zu  $v \in V$  Inverse mit  $-v$ .

V2: Die Skalarmultiplikation muss für alle  $\lambda, \mu \in K$  und  $v, w \in V$  folgende Bedingungen erfüllen:

$$\begin{aligned} (\lambda + \mu) \cdot v &= \lambda \cdot v + \mu \cdot v, \quad \lambda \cdot (v + w) = \lambda \cdot v + \lambda \cdot w \\ \lambda \cdot (\mu \cdot v) &= (\lambda\mu) \cdot v, \quad 1 \cdot v = v. \end{aligned}$$

Überprüfe, dass  $K^n$  tatsächlich ein  $K$ -Vektorraum ist. Hier ist z.B. das neutrale Element der

Addition gegeben durch  $0 = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$ . Beachte, dass in der Definition eines  $K$ -Vektorraumes

die die beiden Verknüpfungen  $+$ ,  $\cdot$  genauso bezeichnet wurden wie diejenigen des Körpers  $(K, +, \cdot)$ .

**Lemma 2.44.** In einem  $K$ -Vektorraum gelten folgende Rechenregeln.

1.  $0 \cdot v = 0$ .
2.  $\lambda \cdot 0 = 0$ .
3.  $\lambda \cdot v = 0 \Rightarrow \lambda = 0$  oder  $v = 0$ .
4.  $(-1) \cdot v = -v$ .

*Beweis.* Zu 1.:  $0 \cdot v = (0 + 0) \cdot v = 0 \cdot v + 0 \cdot v \Rightarrow 0 \cdot v = 0$ .

Zu 2.:  $\lambda \cdot 0 = \lambda \cdot (0 + 0) = \lambda \cdot 0 + \lambda \cdot 0 \Rightarrow \lambda \cdot 0 = 0$

Zu 3.: Sei  $\lambda \cdot v = 0$ . Wenn  $\lambda \neq 0$ , so gibt es ein multiplikatives Inverses  $\lambda^{-1}$  und wir erhalten:

$$v = 1 \cdot v = (\lambda^{-1}\lambda) \cdot v = \lambda^{-1} \cdot (\lambda \cdot v) = \lambda^{-1} \cdot 0 = 0.$$

Zu 4.:

$$v + (-1) \cdot v = 1 \cdot v + (-1) \cdot v = (1 - 1) \cdot v = 0 \cdot v = 0 \Rightarrow (-1) \cdot v = -v.$$

□

Aus diesen Rechenregeln sollte nun klar sein, dass man durchaus die Notation schlank halten kann und es immer klar sein sollte, in welchen Mengen man welche Verknüpfung ausführt. Wir werden daher auch das neutrale Element der Addition  $0$  nicht mehr in fetter Schrift schreiben, sondern  $0 = \mathbf{0}$  benutzen. Wir geben nun einige Beispiele für  $K$ -Vektorräume.

**Beispiel 2.45.** •  $\mathbb{Q}^n, \mathbb{R}^n, \mathbb{C}^n$  bilden mit den Verknüpfungen  $+$ ,  $\cdot$  einen Standard-Vektorraum.

- Sei  $(K, +, \cdot)$  ein Körper. Dann ist bildet die Menge  $M(m \times n, K)$  bezüglich der Addition:

$$A + B := \begin{pmatrix} a_{11} + b_{11} & \dots & a_{1n} + b_{1n} \\ \vdots & & \vdots \\ a_{m1} + b_{m1} & \dots & a_{mn} + b_{mn} \end{pmatrix}$$

und Skalar-Multiplikation

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

für alle  $\lambda \in K$  und alle  $A, B \in V$  einen  $K$ -Vektorraum.

### 2.4.1 Untervektorräume

Wir werden nun Teilmengen eines  $K$ -Vektorraumes  $W \subset V$  identifizieren, welche die Vektorraumstruktur bzgl.  $V$  **erben**, das heisst, dass  $W$  mit den Verknüpfungen von  $V$  selbst einen  $K$ -Vektorraum bildet.

**Definition 2.46.** Sei  $V$  ein  $K$ -Vektorraum und  $W \subset V$ .  $W$  heisst **Untervektorraum** von  $V$ , falls gilt:

$$\text{UV1: } W \neq \emptyset.$$

$$\text{UV2: } v, w \in W \Rightarrow v + w \in W.$$

$$\text{UV3: } v \in W, \lambda \in K \Rightarrow \lambda \cdot v \in W.$$

**Übung 2.47.** Sei  $V$  ein  $K$ -Vektorraum und  $W \subset V$  ein Untervektorraum. Dann ist  $W$  ebenfalls ein  $K$ -Vektorraum und besitzt das gleiche neutrale Element  $0$  wie  $V$ .

**Beispiel 2.48.** Sei  $A \in M(m \times n, K)$  für einen Körper  $(K, +, \cdot)$ . Dann ist die Lösungsmenge des **homogenen LGS**

$$W := \{x \in K^n \mid Ax = 0\}$$

ein Untervektorraum des  $K^n$ .

**Lemma 2.49.** Sei  $V$  ein  $K$ -Vektorraum und  $I \neq \emptyset$  eine Indexmenge und für jedes  $i \in I$  sei  $W_i$  ein Untervektorraum von  $V$ . Dann ist

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

wieder ein Untervektorraum von  $V$ .

*Beweis.* Da  $0 \in W_i$  für alle  $i \in I$  gilt auch  $0 \in W$  und insbesondere  $W \neq \emptyset$ . Für  $v, w \in W$  sind auch  $v, w \in W_i$  für alle  $i \in I$  und somit auch  $v + w$  was wiederum  $v + w \in W$  zeigt. Analoges zeigt man für  $\lambda w$  mit  $\lambda \in K$ .  $\square$

Beachte dass Vereinigungen von Untervektorräumen im Allgemeinen keine Untervektorräume mehr sind.

### 2.4.2 Abschluss von Untervektorräumen und Lineare Unabhängigkeit

Teilmengen eines  $K$ -Vektorraumes  $V$ , die keinen Untervektorraum bilden, können zu einem solchen **abgeschlossen** werden.

Für eine Menge  $I$  nennen wir eine Abbildung

$$\phi : I \rightarrow X, i \mapsto x_i := \phi(i)$$

eine **Familie** von Elementen in  $X$ . Jedes  $i \in I$  ist eindeutig einem Element  $x_i \in X$  zugeordnet aber  $x_i = x_j, i \neq j$  ist erlaubt. Wir schreiben zur Abkürzung für eine Familie  $(x_i)_{i \in I}$ .

Sei nun  $V$  ein  $K$ -Vektorraum und  $(v_i)_{i \in I}$  eine **Familie** von Vektoren  $v_i \in V, i \in I$  für eine Indexmenge  $I$ . Ist z.B.  $I = \{1, \dots, r\}$ , so erhalten wir die endliche Menge von Vektoren  $v_1, \dots, v_r$ .

**Definition 2.50 (Linearkombination, Span).** 1. Ein Vektor  $v$  heisst **Linearkombination** der Vektoren  $v_1, \dots, v_r$ , falls es Skalare  $\lambda_1, \dots, \lambda_r \in K$  gibt mit

$$v = \lambda_1 v_1 + \dots + \lambda_r v_r.$$

2. Für eine allgemeine Familie  $(v_i)_{i \in I}$  definieren wir

$$\text{span}_K((v_i)_{i \in I})$$

als die Menge der  $v \in V$ , die sich aus einer endlichen Teilfamilie von  $(v_i)_{i \in I}$  linear kombinieren lassen, d.h. zu  $v \in V$  muss es eine Zahl  $r \in \mathbb{N}$  geben sowie Indizes  $i_1, \dots, i_r \in I$  und Skalare  $\lambda_1, \dots, \lambda_r \in K$  mit

$$v = \lambda_1 v_{i_1} + \dots + \lambda_r v_{i_r}.$$

Beachte, dass die Auswahl der Teilfamilie  $(i_j)_{j=1, \dots, r}$  von  $I$  durchaus von  $v$  abhängen darf!  $\text{span}_K((v_i)_{i \in I})$  wird auch der von  $(v_i)_{i \in I}$  **aufgespannte Raum** genannt.

3. Falls  $I = \emptyset$ , so setzt man  $\text{span}_K((v_i)_{i \in \emptyset}) := \{0\}$ .

4. Für eine endliche Familie, also,  $I = \{1, \dots, r\}$  schreibt man

$$\text{span}_K((v_i)_{i \in I}) = \text{span}_K(v_1, \dots, v_r) = K v_1 + \dots + K v_r.$$

Falls der Körper  $K$  aus dem Kontext klar ist, lässt man die Bezeichnung weg und schreibt  $\text{span}(v_1, \dots, v_r)$ .

5. Für eine Teilmenge  $M \subset V$  eines  $K$ -Vektorraumes  $V$  ist  $\text{span}(M)$  definiert als die Menge aller **endlichen** Linearkombinationen von Vektoren aus  $M$ .

**Bemerkung 2.51.** Sei  $V$  ein  $K$ -Vektorraum und  $(v_i)_{i \in I}$  eine Familie von Vektoren aus  $V$ . Dann gilt:

1.  $\text{span}((v_i)_{i \in I}) \subset V$  ist ein Untervektorraum.
2. Sei  $W \subset V$  ein Untervektorraum und  $v_i \in W \forall i \in I$ . Dann gilt  $\text{span}((v_i)_{i \in I}) \subset W$ .

Aus dem zweiten Punkt folgt, dass  $\text{span}((v_i)_{i \in I})$  der **kleinste** Untervektorraum von  $V$  ist, der alle  $(v_i)_{i \in I}$  enthält.

**Beispiel 2.52.** Betrachte den Standardraum  $K^n$  für einen Körper  $(K, +, \cdot)$ . Für  $I =$

$\{1, \dots, n\}$  setze  $e_i = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$ , wobei die 1 genau an der  $i$ -ten Stelle steht. Dann gilt

$$\text{span}((e_i)_{i \in I}) = K^n.$$

**Definition 2.53 (Lineare Unabhängigkeit).** Sei  $V$  ein  $K$ -Vektorraum. Eine endliche Familie  $(v_1, \dots, v_r)$  von Vektoren aus  $V$  heisst **linear unabhängig**, falls für  $\lambda_i \in K, i = 1, \dots, r$  gilt

$$\lambda_1 v_1 + \dots + \lambda_r v_r = 0 \Rightarrow \lambda_1 = \dots = \lambda_r = 0.$$

Eine Familie  $(v_i)_{i \in I}$  heisst **linear unabhängig**, falls jede endliche Teilfamilie linear unabhängig ist. Die Familie  $(v_i)_{i \in I}$  heisst **linear abhängig**, falls sie nicht linear unabhängig ist. Per Konvention spannt die leere Familie den Nullvektorraum auf und wird als linear unabhängig bezeichnet.

Wir werden die folgende Sprechweise (ohne den Begriff "Familie") verwenden: die Vektoren  $v_1, \dots, v_n$  sind linear unabhängig (oder linear abhängig).

**Lemma 2.54.** Sei  $V$  ein  $K$ -Vektorraum. Für eine Familie  $(v_i)_{i \in I}$  in  $V$  sind folgende Aussagen äquivalent:

1.  $(v_i)_{i \in I}$  sind linear unabhängig.
2. Jeder Vektor  $v \in \text{span}((v_i)_{i \in I})$  besitzt eine eindeutige Linearkombination von Vektoren aus  $(v_i)_{i \in I}$ .

*Beweis.* (2)  $\Rightarrow$  (1): Folgt sofort mit Beweis per Widerspruch, da lineare Abhängigkeit impliziert, dass es verschiedene Linearkombinationen des Nullvektors geben muss.

(2)  $\Leftarrow$  (1): Sei  $v \in \text{span}((v_i)_{i \in I})$  durch zwei Arten linear kombinierbar ist, d.h., es gibt  $\lambda_i, \mu_i, i \in J$  mit endlicher Teilmenge  $J \subset I$  und  $\lambda_i = \mu_i = 0 \forall i \in I \setminus J$  so dass gilt:

$$v = \sum_{i \in J} \lambda_i v_i = \sum_{i \in J} \mu_i v_i \Leftrightarrow \sum_{i \in J} (\lambda_i - \mu_i) v_i = 0.$$

Mit der linearen Unabhängigkeit von  $(v_i)_{i \in I}$  folgt  $\lambda_i - \mu_i = 0$  für alle  $i \in J$  und somit ist die Linearkombinationen von  $v$  eindeutig.  $\square$

Wir erhalten folgendes Korollar als Kontraposition des obigen Lemmas.

**Korollar 2.55.** Eine Familie  $(v_i)_{i \in I}$  in  $V$  ist linear abhängig genau dann, wenn wenigstens ein Element der Familie als (endliche) Linearkombination der anderen Elemente der Familie dargestellt werden kann.

**Übung 2.56.** Beweise Korollar 2.55.

**Bemerkung 2.57.**

1. Ein einzelner Vektor  $v \in V$  ist genau dann linear unabhängig, wenn  $v \neq 0$ .
2. Ist  $0 \in V$  Teil der Familie, so ist sie linear abhängig, insbesondere ist  $0 \in V$  linear abhängig.
3. Kommt ein Vektor mehrfach in einer Familie vor, so ist sie linear abhängig.

## 2.5 Basen und Dimension

In diesem Abschnitt werden wir uns mit der “Grösse” eines Vektorraumes beschäftigen und diese durch die Angabe der **Dimension** messen. Wir starten zunächst mit den grundlegenden Begriffen eines **Erzeugendensystems** und der **Basis** eines Vektorraumes.

### 2.5.1 Erzeugendensystem und Basis eines Vektorraumes

**Definition 2.58.** Eine Familie  $\mathcal{B} = (v_i)_{i \in I}$  in einem  $K$ -Vektorraum  $V$  heisst **Erzeugendensystem** von  $V$ , falls gilt  $V = \text{span}(\mathcal{B})$ .  $\mathcal{B}$  heisst **Basis** von  $V$ , wenn sie ein linear unabhängiges Erzeugendensystem von  $V$  ist.  $V$  heisst **endlich erzeugt**, falls es ein endliches Erzeugendensystem, d.h.,  $\mathcal{B} = (v_1, \dots, v_n)$  mit  $V = \text{span}(\mathcal{B})$  gibt. Für eine endliche Basis  $\mathcal{B} = (v_1, \dots, v_n)$  nennt man  $n$  die **Länge** der Basis.

Wir werden nun die Länge einer Basis eines Vektorraumes näher untersuchen. Hier stellt sich die Frage, ob es Basen verschiedener Länge überhaupt geben kann. Bevor wir diese grundlegende Frage beantworten, werden zunächst einige Strukturresultate von Basen hergeleitet.

**Satz 2.59.** Für eine Familie von Vektoren  $\mathcal{B} = (v_1, \dots, v_n)$  in einem  $K$ -Vektorraum  $V \neq \emptyset$  sind folgende Aussagen äquivalent:

1.  $\mathcal{B}$  ist eine Basis, d.h. ein linear unabhängiges Erzeugendensystem.
2.  $\mathcal{B}$  ist eine unverkürzbares Erzeugendensystem, d.h., für jedes  $r \in \{1, \dots, n\}$  ist  $\mathcal{B} = (v_1, \dots, v_{r-1}, v_{r+1}, \dots, v_n)$  kein Erzeugendensystem.
3. Zu jedem  $v \in V$  gibt es eindeutig bestimmte  $\lambda_1, \dots, \lambda_n \in K$  mit

$$v = \lambda_1 v_1 + \dots + \lambda_n v_n.$$

4.  $\mathcal{B}$  ist nicht verlängerbar linear unabhängig, d.h., für jeden Vektor  $v \in V$  gilt, dass die Familie  $(v_1, \dots, v_n, v)$  linear abhängig ist.

*Beweis.* (1) $\Rightarrow$ (2): Sei  $\mathcal{B}$  ein linear unabhängiges Erzeugendensystem und sei  $\mathcal{B}$  verkürzbar für  $r = 1$  (o.E.). Somit gibt es  $\lambda_2, \dots, \lambda_n$  mit

$$v_1 = \lambda_2 v_2 + \dots + \lambda_n v_n \Rightarrow 0 = -v_1 + \lambda_2 v_2 + \dots + \lambda_n v_n,$$

im Widerspruch zur Linearen Unabhängigkeit von  $\mathcal{B}$ .

(2) $\Rightarrow$ (3): Sei  $\mathcal{B}$  ein Erzeugendensystem. Angenommen es gibt zwei verschiedene Linearkombinationen für ein  $v \in V$ :

$$v = \lambda_1 v_1 + \dots + \lambda_n v_n = \mu_1 v_1 + \dots + \mu_n v_n,$$

wobei o.E.  $\mu_1 \neq \lambda_1$ . Subtraktion der beiden Terme und Division durch  $\lambda_1 - \mu_1$  ergibt

$$v_1 = \frac{\mu_2 - \lambda_2}{\lambda_1 - \mu_1} v_2 + \dots + \frac{\mu_n - \lambda_n}{\lambda_1 - \mu_1} v_n,$$

so dass  $\mathcal{B}$  verkürzbar ist.

(3) $\Rightarrow$ (4): Aus (3) folgt, dass  $\mathcal{B}$  linear unabhängig ist. Sei nun  $v \in V$  beliebig. Wir erhalten

$$v = \lambda_1 v_2 + \dots + \lambda_n v_n \Rightarrow \lambda_1 v_2 + \dots + \lambda_n v_n + (-1)v = 0$$

also sind  $(v_1, \dots, v_n, v)$  linear abhängig.

(4) $\Rightarrow$ (1): Sei  $\mathcal{B}$  unverlängerbar und linear unabhängig. Für jedes  $v \in V$  ist also  $(v_1, \dots, v_n, v)$  linear abhängig. Das bedeutet, dass es  $(\lambda_1, \dots, \lambda_n, \lambda) \neq 0$  gibt mit

$$\lambda_1 v_1 + \dots + \lambda_n v_n + \lambda v = 0.$$

Den Fall  $\lambda = 0$  können wir ausschließen, da ja sonst auch  $(v_1, \dots, v_n)$  linear abhängig wäre. Also folgt  $\lambda \neq 0$  und wir erhalten

$$v = -\frac{\lambda_1}{\lambda} v_1 - \dots - \frac{\lambda_n}{\lambda} v_n.$$

Somit ist  $\mathcal{B}$  ein Erzeugendensystem. □

Wir erhalten eine direkte Folgerung aus dem letzten Beweis.

**Korollar 2.60.** Ist  $V$  nicht endlich erzeugt, so gibt es eine unendliche lineare unabhängige Familie.

*Beweis.* Wir zeigen zunächst, dass es für beliebiges  $n \in \mathbb{N}$  mit linear unabhängigen Vektoren  $v_1, \dots, v_n$  einen weiteren Vektor  $v \in V$  gibt, so dass  $(v_1, \dots, v_n, v)$  auch linear unabhängig ist. Falls es keinen solchen Vektor  $v \in V$  gibt, ist für jedes  $v \in V$  somit  $(v_1, \dots, v_n, v)$

linear abhängig und nach dem Beweis von (4) $\Rightarrow$ (1) wäre  $(v_1, \dots, v_n)$  ein Erzeugendensystem, Widerspruch.

Mit Hilfe dieser Aussage konstruieren wir nun rekursiv für jedes  $n \in \mathbb{N} \cup \{0\}$  eine Familie  $(v_k^n)_{k \in \{1, \dots, n\}}$  von linear unabhängigen Vektoren aus  $V$  mit der zusätzlichen Eigenschaft, dass  $v_k^n = v_k^m$  für alle  $n \leq m$  und  $k \in \{1, \dots, n\}$  gilt: Für  $n = 0$  ist  $\{1, \dots, n\} = \emptyset$  und daher nichts zu konstruieren. Existiert bereits die Familie  $(v_k^n)_{k \in \{1, \dots, n\}}$ , so erhalten wir durch Anwendung obiger Aussage einen Vektor  $v \in V$ , sodass  $(v_1^n, \dots, v_n^n, v)$  linear unabhängig ist, und konstruieren  $(v_k^{n+1})_{k \in \{1, \dots, n+1\}}$  durch Setzen von  $v_k^{n+1} := v_k^n$  für  $k \leq n$  und  $v_{n+1}^{n+1} := v$ . Die geforderten Eigenschaften gelten dann per Konstruktion (und Induktion). Mit Hilfe dieser Familien definieren wir nun eine weitere Familie  $(v_n)_{n \in \mathbb{N}}$  mittels  $v_n := v_n^n$  für alle  $n \in \mathbb{N}$ . Diese Familie ist offenbar unendlich groß. Ferner ist sie linear unabhängig, denn für jede endliche Teilmenge  $J \subset \mathbb{N}$  existiert  $m := \max\{j \in J\}$  und es gilt  $v_j = v_j^m$  für alle  $j \in J$ . Damit ist  $(v_j)_{j \in J}$  linear unabhängig als Teilfamilie der linear unabhängigen Familie  $(v_n^m)_{n \in \mathbb{N}}$ .  $\square$

Allgemein gilt folgender Satz über die Existenz von Basen.

**I Satz 2.61.** Jeder  $K$ -Vektorraum besitzt eine Basis.

Ein Beweis des obigen Satzes muss auch mit dem Fall von unendlichen Erzeugendensystemen umgehen und daher wird häufig das Lemma von Zorn aus der Mengenlehre benutzt. Wir werden hier nur einen Beweis für das etwas schwächere Existenzresultat für **endlich erzeugte**  $K$ -Vektorräume zeigen.

**Satz 2.62.** Sei  $V$  ein  $K$ -Vektorraum. Dann enthält jedes endliche Erzeugendensystem eine endliche Basis. Insbesondere besitzt jeder endlich erzeugte  $K$ -Vektorraum eine endliche Basis.

*Beweis.* Sei  $E = \{v_1, \dots, v_n\}$  ein Erzeugendensystem von  $V$ . Wenn  $E$  linear unabhängig ist, dann ist  $E$  schon eine Basis. Ist  $E$  linear abhängig, dann gibt es  $v_j \in E$  für ein  $j \in \{1, \dots, n\}$ , welches als Linearkombination der Elemente von  $E_1 = E \setminus \{v_j\}$  darstellbar ist (vgl. Korollar 2.55) und somit ist  $E_1$  wiederum ein Erzeugendensystem von  $V$ . Indem die gleiche Argumentation wieder auf  $E_1$  angewendet wird erhalten wir ein  $E_2$ . Dies wird nun iteriert und wir erhalten eine Inklusionskette:

$$E \supseteq E_1 \supseteq E_2 \supseteq \dots$$

Da  $E$  endlich viele Elemente enthält muss die Kette nach endlich vielen Schritten abbrechen und wir erhalten ein  $E_s \subset E$ , welches eine Basis mit endlich vielen Elementen ist.  $\square$

Im Folgenden werden wir nun den Steinitzischen Basisaustauschsatz beweisen, der aufzeigt, wie man zwei Basen jeweils ineinander überführen kann. Wir starten zunächst mit dem sogenannten Austauschlemma.

**Lemma 2.63.** Sei  $V$  ein  $K$ -Vektorraum mit der Basis  $\mathcal{B} = (v_1, \dots, v_r)$  und sei  $w = \lambda_1 v_1 + \dots + \lambda_r v_r$ . Falls es  $k \in \{1, \dots, r\}$  mit  $\lambda_k \neq 0$  gibt, so ist

$$\mathcal{B}' := (v_1, \dots, v_{k-1}, w, v_{k+1}, \dots, v_r)$$

wieder eine Basis von  $V$ .

*Beweis.* O.E. nehmen wir an dass  $k = 1$  gilt. Wir erhalten eine Darstellung von  $v_1$  durch

$$v_1 = \frac{1}{\lambda_1} w - \frac{\lambda_2}{\lambda_1} \dots - \frac{\lambda_r}{\lambda_1} v_r.$$

Sei nun  $v \in V$  beliebig. Wir erhalten (da  $\mathcal{B}$  Erzeugendensystem)

$$\begin{aligned} v &= \mu_1 v_1 + \dots + \mu_r v_r \\ &= \frac{\mu_1}{\lambda_1} w + \left( \mu_2 - \frac{\mu_1 \lambda_2}{\lambda_1} \right) v_2 + \dots + \left( \mu_r - \frac{\mu_1 \lambda_r}{\lambda_1} \right) v_r. \end{aligned}$$

Somit ist  $\mathcal{B}'$  ein Erzeugendensystem. Nun zeigen wir, dass  $\mathcal{B}'$  linear unabhängig ist. Sei also

$$\mu w + \mu_2 v_2 + \dots + \mu_r v_r = 0.$$

Wir setzen  $w = \lambda_1 v_1 + \dots + \lambda_r v_r$  ein und erhalten:

$$\mu \lambda_1 v_1 + (\mu \lambda_2 + \mu_2) v_2 + \dots + (\mu \lambda_r + \mu_r) v_r = 0.$$

Es folgt

$$\mu \lambda_1 = (\mu \lambda_2 + \mu_2) = \dots = (\mu \lambda_r + \mu_r) = 0,$$

da  $\mathcal{B}$  linear unabhängig war. Wegen  $\lambda_1 \neq 0$  folgt  $\mu = 0$  und damit  $\mu_2 = \dots = \mu_r = 0$ .  $\square$

**Satz 2.64 (Steinitz'scher Austauschatz).** Sei  $V$  ein  $K$ -Vektorraum mit der Basis  $\mathcal{B} = (v_1, \dots, v_r)$  und sei  $(w_1, \dots, w_n)$  eine linear unabhängige Familie in  $V$ . Dann gilt  $n \leq r$  und es gibt Indizes  $i_1, \dots, i_n \in \{1, \dots, r\}$ , so dass man nach Austausch von  $v_{i_j}$  gegen  $w_j, j = 1, \dots, n$  wieder eine Basis von  $V$  erhält. O.E. sei  $i_j = j, j = 1, \dots, n$  und wir erhalten, dass

$$\mathcal{B}^* = (w_1, \dots, w_n, v_{n+1}, \dots, v_r)$$

eine Basis von  $V$  ist.

*Beweis.* Wir führen eine Induktion bzgl.  $n \in \mathbb{N} \cup \{0\}$  durch. Für  $n = 0$  ist nichts zu beweisen. Sei  $n \geq 1$  und wir nehmen an, dass die Aussage des Satzes für  $n - 1$  wahr ist. Da  $w_1, \dots, w_{n-1}$  linear unabhängig sind, ergibt die Induktionsannahme, dass

$$(w_1, \dots, w_{n-1}, v_n, \dots, v_r)$$

eine Basis von  $V$  ist. Also gibt es  $\lambda_1, \dots, \lambda_r \in K$  mit

$$w_n = \lambda_1 w_1 + \dots + \lambda_{n-1} w_{n-1} + \lambda_n v_n + \dots + \lambda_r v_r.$$

Falls  $\lambda_n = \dots = \lambda_r = 0$  wäre, so wären  $(w_1, \dots, w_n)$  linear abhängig, Widerspruch. Also gibt es wenigstens ein  $i \in \{n, \dots, r\}$  mit  $\lambda_i \neq 0$ . Insbesondere gilt also  $n \leq r$ . Darüber hinaus können wir o.E. annehmen, dass  $\lambda_n \neq 0$  gilt. Laut Lemma 2.63 können wir also  $v_n$  gegen  $w_n$  tauschen. Daher ist  $\mathcal{B}^*$  eine Basis von  $V$ .  $\square$

Wir erhalten nun wichtige Folgerungen.

- Korollar 2.65.**
1. Hat ein  $K$ -Vektorraum  $V$  eine endliche Basis, so ist jede Basis von  $V$  endlich.
  2. Je zwei endliche Basen eines  $K$ -Vektorraums  $V$  haben die gleiche Länge.

*Beweis.* Zu (1): Sei  $\mathcal{B} = (v_1, \dots, v_r)$  endliche Basis und  $(w_i)_{i \in I}$  beliebige Basis. Ist  $I$  unendlich, so gibt es Indizes  $i_j, j = 1, \dots, r+1$  mit  $w_{i_1}, \dots, w_{i_{r+1}}$  linear unabhängig, im Widerspruch zum Austauschsatz (siehe  $n \leq r$ ).

Zu (2): Sind  $\mathcal{B} = (v_1, \dots, v_r)$  und  $\mathcal{B}' = (w_1, \dots, w_k)$  zwei Basen, so wende jeweils den Austauschsatz an, was  $k \leq r$  und auch  $r \leq k$  impliziert.  $\square$

Nun wollen wir uns mit dem Problem befassen, wie man eine Basis aus einer Menge linear unabhängiger Vektoren “aufbaut”.

**Satz 2.66 (Basisergänzungssatz).** In einem endlich erzeugten Vektorraum  $V$  seien linear unabhängige Vektoren  $w_1, \dots, w_n$  gegeben. Dann gibt es ein  $r \geq n$  und Vektoren  $w_{n+1}, \dots, w_r \in V$ , so dass

$$\mathcal{B} = (w_1, \dots, w_n, w_{n+1}, \dots, w_r)$$

eine Basis von  $V$  ist.

*Beweis.* Sei  $(v_1, \dots, v_r)$  eine Basis von  $V$  (existiert nach Satz 2.62). Nun wenden wir den Satz 2.64 an und erhalten (nach geeigneter Umnummerierung der Indizes) eine Basis mittels

$$w_{n+1} := v_{n+1}, \dots, w_r := v_r. \quad \square$$

## 2.5.2 Dimension eines Vektorraumes

**Definition 2.67.** Ist  $V$  ein  $K$ -Vektorraum, so heißt

$$\dim_K(V) := \begin{cases} \infty, & \text{falls } V \text{ keine endliche Basis besitzt,} \\ r, & \text{falls } V \text{ eine Basis der Länge } r \text{ besitzt.} \end{cases} \quad (2.8)$$

$\dim_K(V)$  heißt **Dimension** von  $V$  über  $K$ . Falls  $K$  aus dem Kontext klar ist, schreibt man auch  $\dim(V)$ .

**Korollar 2.68.** Ist  $W \subset V$  ein Untervektorraum eines endlich erzeugten Vektorraumes  $V$ , so ist auch  $W$  endlich erzeugt und es gilt  $\dim(W) \leq \dim(V)$ .

Gilt sogar  $\dim(W) = \dim(V)$ , so ist  $W = V$ .

*Beweis.* Wir zeigen zunächst, dass  $W$  endlich erzeugt ist: Wäre dies nämlich nicht der Fall, so gäbe es nach Korollar 2.60 eine unendliche linear unabhängige Familie in  $W$ . Diese wäre dann aber auch eine unendliche linear unabhängige Familie in  $V$  – ein Widerspruch zum Austauschsatz, da  $V$  endlich erzeugt ist (und daher eine endliche Basis hat).

Sei nun  $(w_1, \dots, w_n)$  eine Basis von  $W$  und  $(v_1, \dots, v_r)$  eine Basis von  $V$ . Nach Satz 2.64 gilt dann  $\dim(W) = n \leq r = \dim(V)$ .

Gelte nun  $n = \dim(W) = \dim(V) = r$ . Wäre nun  $W \neq V$ , so gäbe es einen Vektor  $v \in V \setminus W$ , so dass  $(w_1, \dots, w_n, v)$  linear unabhängig sind. Das steht allerdings erneut im Widerspruch zum Austauschsatz.  $\square$

**Beispiel 2.69.** Betrachte den Standardraum  $K^n$  für einen Körper  $(K, +, \cdot)$ .  $K^n$  besitzt die kanonische Basis  $(e_1, \dots, e_n)$

**Beispiel 2.70 (Unendlich Dimensionaler Vektorraum).** Betrachte den unendlich-dimensionalen Raum der Polynomfunktionen mit endlichem Grad über  $\mathbb{R}$ . Formal definiere

$$\mathcal{P}_d := \{f : \mathbb{R} \rightarrow \mathbb{R} \mid \exists \text{ Koeffizienten } a_0, a_1, \dots, a_d \in \mathbb{R}, \text{ so dass gilt } f(x) = \sum_{i=0}^d a_i x^i \text{ für alle } x \in \mathbb{R}\}$$

als die Menge der Polynomfunktionen mit maximalem Grad  $d$ . Es gilt, dass  $\mathcal{P}_d$  mit den Verknüpfungen

$$+ : \mathcal{P}_d \times \mathcal{P}_d \rightarrow \mathcal{P}_d, f + g := f(x) + g(x) \text{ für alle } x \in \mathbb{R} \quad (\text{Addition})$$

$$\cdot : \mathbb{R} \times \mathcal{P}_d \rightarrow \mathcal{P}_d, \lambda f := \lambda f(x) \text{ für alle } x \in \mathbb{R}, \quad (\text{Skalar-Multiplikation})$$

einen  $\mathbb{R}$ -Vektorraum bildet. Definiere für  $i \in \mathbb{N} \cup \{0\}$  ein Monom vom Grad  $i$  durch

$$m_i : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto x^i.$$

Die Monome  $\mathcal{B}_d = (m_0, m_1, m_2, \dots, m_d)$ , bilden dann eine Basis von  $\mathcal{P}_d$  (benutze hier, dass ein Polynom vom Maximalgrad  $d$  höchstens  $d$ -viele reelle Nullstellen besitzt oder identisch 0 ist). Definieren wir

$$\mathcal{P} := \cup_{d \in \mathbb{N}} \mathcal{P}_d$$

so erhalten wir den unendlich-dimensionalen Vektorraum aller Polynomfunktionen mit endlichem Grad. Eine unendlich grosse Basis ist hier gegeben durch  $\mathcal{B} = (m_0, m_1, m_2, \dots)$ .

Nun wenden wir uns der Dimensionsformel für Untervektorräume zu.

**Definition 2.71.** Sei  $V$  ein  $K$ -Vektorraum mit Untervektorräumen  $W_1, \dots, W_r \subset V$ . Dann heißt

$$W_1 + \dots + W_r := \{v \in V \mid v = v_1 + \dots + v_r, v_i \in W_i, i = 1, \dots, r\}$$

die **Summe** von  $W_1, \dots, W_r$ .

**Übung 2.72.** Folgende Aussagen gelten:

1.  $W_1 + \dots + W_r \subset V$  ist ein Untervektorraum von  $V$ .
2.  $W_1 + \dots + W_r = \text{span}(W_1 \cup \dots \cup W_r)$ .
3.  $\dim(W_1 + \dots + W_r) \leq \sum_{i=1}^r \dim(W_i)$ .

Für den Fall  $r = 2$  werden wir für die letzte obige Ungleichung zeigen, wann diese straff ist.

**Satz 2.73 (Dimensionsformel).** Für endlichdimensionale Untervektorräume  $W_1, W_2 \subset V$  gilt

$$\dim(W_1 + W_2) = \dim(W_1) + \dim(W_2) - \dim(W_1 \cap W_2).$$

*Beweis.* Sei  $(v_1, \dots, v_m)$  eine Basis von  $W_1 \cap W_2$ . Wir benutzen Satz 2.66 um sie jeweils zu Basen

$$(v_1, \dots, v_m, w_1, \dots, w_k) \text{ und } (v_1, \dots, v_m, w'_1, \dots, w'_\ell)$$

von  $W_1$  und  $W_2$  zu ergänzen. Wir müssen nun zeigen, dass

$$\mathcal{B} = (v_1, \dots, v_m, w_1, \dots, w_k, w'_1, \dots, w'_\ell)$$

eine Basis von  $W_1 + W_2$  ist. Es folgt unmittelbar, dass  $W_1 + W_2$  von  $\mathcal{B}$  erzeugt wird. Wir zeigen nun lineare Unabhängigkeit der Vektoren in  $\mathcal{B}$ . Sei

$$\lambda_1 v_1 + \dots + \lambda_m v_m + \mu_1 w_1 + \dots + \mu_k w_k + \mu'_1 w'_1 + \dots + \mu'_\ell w'_\ell = 0. \quad (2.9)$$

Setzen wir

$$v := \lambda_1 v_1 + \dots + \lambda_m v_m + \mu_1 w_1 + \dots + \mu_k w_k,$$

so ist  $v \in W_1$  und  $-v = \mu'_1 w'_1 + \dots + \mu'_\ell w'_\ell \in W_2$  und somit  $v \in W_1 \cap W_2$ . Also gilt

$$v = \lambda'_1 v_1 + \dots + \lambda'_m v_m$$

mit  $\lambda'_1, \dots, \lambda'_m \in K$ . Da  $(v_1, \dots, v_m, w_1, \dots, w_k)$  eine Basis von  $W_1$  ist, folgt dass die obige Linearkombination eindeutig ist und somit gilt:

$$\mu_1 = \dots = \mu_k = 0.$$

Wenn wir das in (2.9) einsetzen, so erhalten wir (unter Benutzung, dass  $(v_1, \dots, v_m, w'_1, \dots, w'_k)$  Basis von  $W_2$  ist und somit linear unabhängig)

$$\lambda_1 = \dots = \lambda_m = \mu'_1 = \dots = \mu'_l = 0.$$

□

Wir zeigen nun ein weiteres Strukturresultat bezüglich zweier Untervektorräume.

**Lemma 2.74.** Ist  $V = W_1 + W_2$  für zwei Untervektorräume  $W_1, W_2$ , so sind äquivalent:

1.  $W_1 \cap W_2 = \{0\}$ .
2. Jedes  $v \in V$  ist eindeutig darstellbar als  $v = w_1 + w_2$  mit  $w_1 \in W_1, w_2 \in W_2$ .
3. Jedes Paar von Vektoren  $w_1 \in W_1 \setminus \{0\}, w_2 \in W_2 \setminus \{0\}$  ist linear unabhängig.

*Beweis.* Zu (1). $\Rightarrow$ (2).: Sei  $v = w_1 + w_2 = w'_1 + w'_2$  mit  $w_1 \neq w'_1$  (o.E.). Dann folgt

$$0 \neq w_1 - w'_1 = w'_2 - w_2 \in W_1 \cap W_2.$$

Also ist  $W_1 \cap W_2 \neq \{0\}$ , Widerspruch.

Zu (2). $\Rightarrow$ (3).: Sind  $w_1, w_2$  linear abhängig, so erhält man verschiedene Darstellungen des Nullvektors.

Zu (3). $\Rightarrow$ (1).: Ist  $0 \neq v \in W_1 \cap W_2$ , so folgt:  $1v + (-1)v = 0$ , Widerspruch. □

**Definition 2.75.** Ein Vektorraum  $V$  heisst **direkte Summe** von zwei Untervektorräumen  $W_1$  und  $W_2$  (geschrieben  $V = W_1 \oplus W_2$ ), wenn gilt

$$V = W_1 + W_2 \text{ und } W_1 \cap W_2 = \{0\}.$$

**Satz 2.76.** Sei  $V$  ein endlichdimensionaler  $K$ -Vektorraum mit Untervektorräumen  $W_1$  und  $W_2$ , so sind folgende Bedingungen äquivalent:

1.  $V = W_1 \oplus W_2$ .
2. Es gibt Basen  $(w_1, \dots, w_k)$  von  $W_1$  und  $(w'_1, \dots, w'_l)$  von  $W_2$ , so dass  $(w_1, \dots, w_k, w'_1, \dots, w'_l)$  eine Basis von  $V$  ist.
3.  $V = W_1 + W_2$  und  $\dim V = \dim W_1 + \dim W_2$ .

*Beweis.* Übung mittels Dimensionsformel. □

Wir erhalten folgendes Korollar.

**Korollar 2.77.** Sei  $V$  ein endlichdimensionaler  $K$ -Vektorraum und  $W$  ein Untervektorraum. Dann gibt es einen (i.A. nicht eindeutigen) Untervektorraum  $W' \subset V$  mit

$$V = W \oplus W'.$$

*Beweis.* Wir wählen eine Basis  $(v_1, \dots, v_m)$  von  $W$  und ergänzen sie dann zu einer Basis  $(v_1, \dots, v_m, v_{m+1}, \dots, v_n)$  von  $V$  und definieren

$$W' := \text{span}(v_{m+1}, \dots, v_n).$$

□

Wir werden nun die direkte Summe von mehr als zwei Vektorräumen betrachten.

**Definition 2.78.** Ein Vektorraum  $V$  heisst **direkte Summe** von Untervektorräumen  $W_1, \dots, W_k$  (geschrieben  $V = W_1 \oplus \dots \oplus W_k$ ), wenn gilt:

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

Nichtnullvektoren in  $(w_1, \dots, w_k)$  mit  $w_i \in W_i, i = 1 \dots, k$  sind linear unabhängig. (DS2)

Die Bedingung DS2 besagt, dass für alle  $(w_1, \dots, w_k) \in W_1 \times \dots \times W_k$  jede Teilfamilie  $(w_j)_{j \in J}$  mit  $J \subset \{1, \dots, k\}$  und  $w_j \neq 0$  für jedes  $j \in J$  linear unabhängig ist. Insbesondere gilt:

$$w_1 + \dots + w_k = 0 \Rightarrow w_1 = \dots = w_k = 0.$$

**Beispiel 2.79.** Falls  $(v_1, \dots, v_n)$  eine Basis von  $V$  ist, so gilt

$$V = K v_1 \oplus \dots \oplus K v_k.$$

**Satz 2.80.** Für Untervektorräume  $W_1, \dots, W_k$  eines endlichdimensionalen  $K$ -Vektorraum  $V$  sind äquivalent:

1.  $V = W_1 \oplus \dots \oplus W_k$ .
2. Ist für jedes  $i \in \{1, \dots, k\}$  eine Basis  $(v_1^i, \dots, v_{r_i}^i)$  von  $W_i$  gegeben, so ist

$$\mathcal{B} := (v_1^1, \dots, v_{r_1}^1, v_1^2, \dots, v_{r_2}^2, \dots, v_1^k, \dots, v_{r_k}^k)$$

eine Basis von  $V$ .

3.  $V = W_1 + \dots + W_k$  und  $\dim V = \dim W_1 + \dots + \dim W_k$ .

*Beweis.* Zu (1) $\Rightarrow$ (2):  $\mathcal{B}$  ist ein Erzeugendensystem wegen  $V = W_1 + \dots + W_k$ . Zum Nachweis der linearen Unabhängigkeit sei

$$\mu_1^1 v_1^1 + \dots + \mu_{r_1}^1 v_{r_1}^1 + \dots + \mu_{r_k}^k v_{r_k}^k = 0.$$

Setze

$$w_i := \mu_1^i v_1^i + \dots + \mu_{r_i}^i v_{r_i}^i$$

so gilt

$$w_1 + \dots + w_k = 0.$$

Mittels (DS2) erhalten wir  $w_1 = \dots = w_k = 0$ . Also gilt  $\mu_1^i v_1^i + \dots + \mu_{r_i}^i v_{r_i}^i = 0$  und daraus folgt  $\mu_1^i = \dots = \mu_{r_i}^i = 0$  für alle  $i \in \{1, \dots, k\}$ .

Zu (2) $\Leftrightarrow$ (3): klar!

Zu (2) $\Rightarrow$ (1): Wir müssen nur (DS2) zeigen. Betrachte dazu eine Teilmenge  $J \subset \{1, \dots, k\}$  und für jedes  $i \in J$  einen Vektor  $w_i \in W_i$  mit:

$$0 \neq w_i = \mu_1^i v_1^i + \dots + \mu_{r_i}^i v_{r_i}^i \in W_i. \quad (2.10)$$

Sei o.E.  $\mu_1^i \neq 0$  für alle  $i \in J$ . Für

$$\sum_{i \in J} \lambda_i w_i = 0,$$

erhalten wir durch Einsetzen von (2.10):

$$\sum_{i \in J} \sum_{\ell=1}^{r_i} \lambda_i \mu_\ell^i v_\ell^i = 0.$$

Weil  $\mathcal{B}$  eine Basis ist, folgt  $\lambda_i \mu_1^i = 0$  für alle  $i \in J$ , und somit  $\lambda_i = 0$  für alle  $i \in J$ .  $\square$

### 2.5.3 Anwendungen auf Matrizen

Wir betrachten den Standardraum  $K^n$  für einen Körper  $(K, +, \cdot)$ . Seien  $a_1, \dots, a_m \in K^n$  gegeben und sei

$$W = \text{span}(a_1, \dots, a_m).$$

Fassen wir die  $a_i, i = 1, \dots, m$  als Zeilenvektoren auf, so ergeben sie untereinander geschrieben die Matrix

$$A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ a_{21} & \dots & a_{2n} \\ \vdots & & \vdots \\ a_{m1} & \dots & a_{mn} \end{pmatrix} \in M(m \times n, K).$$

**Definition 2.81.** Für  $A \in M(m \times n, K)$  mit zugehörigen Zeilen  $a_1, \dots, a_m \in K^n$  heisst

$$\text{ZR}(A) := \text{span}(a_1, \dots, a_m) \subset K^n$$

der **Zeilenraum** von  $A$ .

Wir kommen nun auf die in Definition 1.5 eingeführten elementaren Zeilenoperationen zurück.

**Lemma 2.82.** Ist  $B$  durch  $A$  durch endlich viele elementare Zeilenumformungen entstanden, so gilt

$$\text{ZR}(A) = \text{ZR}(B).$$

*Beweis.* Wir müssen wieder nur überprüfen, dass sich der erzeugte Untervektorraum nach der Anwendung einer elementaren Zeilenumformung nicht verändert hat, d.h.,  $a_i = b_i$  für alle  $i \notin \{j, k\}$  und  $b_j$  und  $b_k$  entstehen durch Anwendung des Typs 1, 2, oder 3 für  $j \neq k$  und o.E.  $j < k$ . Offensichtlich gilt

$$\text{span}(a_1, \dots, a_m) = \text{span}(b_1, \dots, b_m)$$

für Typ 1 Umformungen durch Umsortieren der Linearkombinationen. Weiterhin gilt für Typ 2 Umformungen mit  $\lambda \in K \setminus \{0\}$

$$\text{span}(b_1, \dots, b_m) = \text{span}(\lambda a_1, \dots, a_m) = \text{span}(a_1, \dots, a_m),$$

wobei o.E. die erste Zeile mit  $\lambda \in K \setminus \{0\}$  skaliert wurde. Betrachte nun den Typ 3, d.h.,  $B$  entsteht durch das addieren des  $\lambda$ -fachen der  $j$  Zeile zur  $k$ -ten Zeile von  $A$ . Sei  $v \in \text{span}(a_1, \dots, a_m)$ . Wir erhalten:

$$\begin{aligned} v &= \sum_{i=1}^m \mu_i a_i = \sum_{i \in \{1, \dots, m\} \setminus \{j, k\}} \mu_i a_i + \mu_j a_j + \mu_k a_k \\ &= \sum_{i \in \{1, \dots, m\} \setminus \{j, k\}} \mu_i a_i + (\mu_j - \mu_k \lambda) a_j + \mu_k (a_k + \lambda a_j) \\ &= \sum_{i \in \{1, \dots, m\} \setminus \{j, k\}} \mu_i b_i + (\mu_j - \mu_k \lambda) b_j + \mu_k b_k \end{aligned}$$

Also gilt  $v \in \text{span}(b_1, \dots, b_m)$ . Die andere Richtung  $\text{span}(b_1, \dots, b_m) \subset \text{span}(a_1, \dots, a_m)$  folgt analog.  $\square$

Wir erhalten als Verallgemeinerung von Satz 1.9 folgendes Resultat bzgl. der Umwandlung einer Matrix in Zeilenstufenform.

**Satz 2.83.** Sei  $A \in M(m \times n, K)$ . Dann ist  $A$  durch elementare Zeilenumformungen in eine Zeilenstufenmatrix  $A'$  überführbar.

Wenn wir nun ausgehend von einer Matrix  $A \in M(m \times n, K)$  diese in Zeilenstufenform  $B$  gebracht haben, so sind die von 0 verschiedenen Zeilen  $b_1, \dots, b_r$  eine Basis des Zeilenraums  $ZR(A) = ZR(B)$  (überprüfe, dass die Vektoren  $b_i, i = 1, \dots, r$  linear unabhängig sind!).

Analog zu elementaren Zeilenumformungen können wir auch elementare Spaltenumformungen definieren. Beachte, dass der Spaltenraum einer Matrix analog zum Zeilenraum definiert ist, d.h.,  $SR(A) := \text{span}(a^1, \dots, a^n)$ , allerdings gilt

$$SR(A) \subset K^m, \text{ wo hingegen } ZR(A) \subset K^n.$$

Wir erhalten folgendes Korollar bzgl. des Spaltenraumes bei elementaren Spaltenumformungen.

**Korollar 2.84.** Der Spaltenraum und damit auch der Spaltenrang bleibt unter elementaren Spaltenumformungen unverändert.

*Beweis.* Dies kann genauso wie in Lemma 2.82 bewiesen werden. □

Allerdings bleibt der Zeilenraum unter elementaren Spaltenumformungen i.A. nicht unverändert. Analog bleibt der Spaltenraum unter elementaren Zeilenumformungen i.A. nicht unverändert. Wir geben ein Beispiel.

**Beispiel 2.85.** Betrachte folgende Matrix

$$A = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} \in M(2 \times 2, \mathbb{R}).$$

Vertauschen wir nun die beiden Spalten erhalten wir

$$B = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \in M(2 \times 2, \mathbb{R}).$$

Der Spann der Zeilen von  $A$  ist die  $x$ -Achse im  $\mathbb{R}^2$ . Der Spann der Zeilen von  $B$  ist die  $y$ -Achse im  $\mathbb{R}^2$ .

**Definition 2.86.** Für  $A \in M(m \times n, K)$  sei

$$\begin{aligned} \text{Zeilenrang } A &:= \dim(ZR(A)) \\ \text{Spaltenrang } A &:= \dim(SR(A)). \end{aligned}$$

Wir erhalten folgenden Satz bzgl. des Zeilen- oder Spaltenrangs bei Anwendung von elementaren Zeilen- oder Spaltenumformungen.

**Satz 2.87.** Der Zeilenrang bleibt unter endlich vielen elementaren Spaltenumformungen unverändert. Der Spaltenrang bleibt unter endlich vielen elementaren Zeilenumformungen unverändert.

*Beweis.* Wir zeigen nur die erste Behauptung, die zweite folgt analog. Seien  $A = (a_{ij})$ ,  $B = (b_{ij}) \in M(m \times n, K)$ , wobei  $B$  aus  $A$  durch Anwendung einer elementaren Spaltenumformung hervorgeht.

1. Fall:  $B$  entsteht durch Vertauschung der  $\ell$ -ten und  $k$ -ten Spalte von  $A$ , o.E.,  $k < \ell$ . Dann gilt

$$b_{ij} = \begin{cases} a_{ij}, & j \notin \{\ell, k\} \\ a_{i\ell}, & j = k \\ a_{ik}, & j = \ell \end{cases}$$

Sei nun  $I \subset \{1, \dots, m\}$  und  $(\mu_i)_{i \in I}$  eine Familie in  $K$ . Dann gilt

$$\begin{aligned} \sum_{i \in I} \mu_i a_i = 0 &\Leftrightarrow \sum_{i \in I} \mu_i a_{ij} = 0 \text{ für alle } j \in \{1, \dots, n\} \\ &\Leftrightarrow \sum_{i \in I} \mu_i b_{ij} = 0 \text{ für alle } j \in \{1, \dots, n\} \\ &\Leftrightarrow \sum_{i \in I} \mu_i b_i = 0. \end{aligned}$$

Also folgt, dass  $\{a_i | i \in I\}$  eine maximal linear unabhängige Teilmenge von  $\{a_1, \dots, a_m\}$  ist, genau dann, wenn  $\{b_i | i \in I\}$  eine maximal linear unabhängige Teilmenge von  $\{b_1, \dots, b_m\}$  ist. Somit folgt  $\dim(\text{ZR}(A)) = \dim(\text{ZR}(B))$ .

2. Fall:  $B$  entsteht aus  $A$  durch Multiplikation der  $k$ -ten Spalte für  $k \in \{1, \dots, n\}$  mit  $\lambda \in K^*$ . Dann gilt:

$$b_{ij} = \begin{cases} a_{ij}, & \text{falls } j \neq k \\ \lambda a_{ik}, & \text{falls } j = k. \end{cases}$$

Sei  $I \subset \{1, \dots, m\}$  und  $(\mu_i)_{i \in I}$  eine Familie in  $K$ . Dann gilt

$$\begin{aligned} \sum_{i \in I} \mu_i a_i = 0 &\Leftrightarrow \sum_{i \in I} \mu_i a_{ij} = 0 \text{ für alle } j \in \{1, \dots, n\} \\ &\Leftrightarrow \sum_{i \in I} \mu_i b_{ij} = 0 \text{ für alle } j \in \{1, \dots, n\} \setminus \{k\} \text{ und } \sum_{i \in I} \frac{\mu_i}{\lambda} b_{ik} = 0 \\ &\Leftrightarrow \sum_{i \in I} \mu_i b_i = 0, \end{aligned}$$

wobei wir nutzen, dass gilt  $\sum_{i \in I} \frac{\mu_i}{\lambda} b_{ik} = 0 \Leftrightarrow \sum_{i \in I} \mu_i b_{ik} = 0$ . Der Rest des Beweises geht wie im Fall 1.

3. Fall:  $B$  entsteht aus  $A$  durch Addition des  $\lambda$ -fachen der  $\ell$ -ten Spalte zur  $k$ -ten Spalte

für  $\ell, k \in \{1, \dots, n\}, \lambda \in K$ . Wir erhalten:

$$b^j = \begin{cases} a^j, & \text{falls } j \neq k \\ a^k + \lambda a^\ell, & \text{falls } j = k. \end{cases}$$

Sei  $I \subset \{1, \dots, m\}$  und  $(\mu_i)_{i \in I}$  eine Familie in  $K$ . Wir zeigen wiederum:  $\sum_{i \in I} \mu_i a_i = 0 \Leftrightarrow \sum_{i \in I} \mu_i b_i = 0$ .

$\Rightarrow$ : Es gilt  $\sum_{i \in I} \mu_i a_i = 0 \Leftrightarrow \sum_{i \in I} \mu_i a_{ij} = 0$  für alle  $j \in \{1, \dots, n\}$ . Daraus folgt  $\sum_{i \in I} \mu_i (a_{ik} + \lambda a_{i\ell}) = 0$  und somit  $\sum_{i \in I} \mu_i b_i = 0$ .

$\Leftarrow$ : Aus  $\sum_{i \in I} \mu_i b_i = 0$  folgt  $\sum_{i \in I} \mu_i b_{ij} = 0$  für alle  $j \in \{1, \dots, n\}$ . Somit gilt  $\sum_{i \in I} \mu_i a_{ij} = 0$  für alle  $j \in \{1, \dots, n\} \setminus \{k\}$  und

$$0 = \sum_{i \in I} \mu_i (a_{ik} + \lambda a_{i\ell}) = \sum_{i \in I} \mu_i a_{ik} + \lambda \sum_{i \in I} \mu_i a_{i\ell}.$$

Mit  $\ell \neq k$  gilt  $\sum_{i \in I} \mu_i a_{ik} = 0$  und somit auch  $\sum_{i \in I} \mu_i a_{i\ell} = 0$ . □

**Definition 2.88.** Eine Matrix der Form

$$H = \begin{pmatrix} a_{11} & & 0 & * \\ & \ddots & & * \\ 0 & & a_{rr} & * \\ 0 & 0 & 0 & 0 \end{pmatrix} \in M(m \times n, K).$$

mit  $a_{ii} \neq 0, i = 1, \dots, r$  ist in **Hermitescher Normalform**.

**Satz 2.89.** Jede Matrix  $A \in M(m \times n, K)$  lässt sich durch elementare Zeilenumformungen und Vertauschung von Spalten in Hermitesche Normalform bringen.

**Übung 2.90.** Beweise Satz 2.89 in dem der Algorithmus von Gauss erweitert wird. (Dies führt zum **Gauss-Jordan** Algorithmus.)

**Satz 2.91.** Sei Matrix  $A \in M(m \times n, K)$ . Dann gilt:

$$\text{Zeilenrang } A = \text{Spaltenrang } A.$$

*Beweis.* Mit Satz 2.87 sowie Lemma 2.82 und Korollar 2.84 folgt, dass elementare Zeilenumformungen und Spaltenvertauschungen weder den Zeilenrang noch den Spaltenrang ändern. Wird also  $A$  durch elementare Zeilenumformungen und Spaltenvertauschungen in eine Matrix  $B$  in Hermitescher Normalform überführt, so haben  $A$  und  $B$  den gleichen Zeilen- wie Spaltenrang. Da  $B$  Zeilenstufenform hat, gilt, dass der Zeilenrang von  $A$  gleich der Anzahl  $r$  der nicht-null Zeilen ist. Weiterhin sieht man, dass die Spaltenvektoren

$b^{r+1}, \dots, b^n$  sich als Linearkombination der Spalten  $b^1, \dots, b^r$  darstellen lassen und somit ist auch der Spaltenrang von  $B$  gleich  $r$ . □

**Definition 2.92.** Für eine Matrix  $A \in M(m \times n, K)$  ist der **Rang** von  $A$  definiert als

$$\text{rang } A := \text{Zeilenrang } A = \text{Spaltenrang } A.$$

## Kapitel 3

# Lineare Abbildungen

Nun werden wir uns mit abstrakten **linearen Abbildungen zwischen  $K$ -Vektorräumen** beschäftigen. Etwas später wird dann auch ausführlich der Zusammenhang von linearen Abbildungen und Matrizen (bzw. von Abbildungen der Form  $x \mapsto Ax$ ) diskutiert.

**Definition 3.1.** Eine Abbildung  $F : V \rightarrow W$  zwischen  $K$ -Vektorräumen  $V$  und  $W$  heißt **linear** (oder  **$K$ -linear** bzw. **Homomorphismus von  $K$ -Vektorräumen**), wenn für alle  $v, w \in V, \lambda \in K$  gilt:

$$\text{L1: } F(v + w) = F(v) + F(w)$$

$$\text{L2: } F(\lambda v) = \lambda F(v).$$

**Übung 3.2.** Zeige dass L1 und L2 äquivalent sind zu

$$F(\lambda v + \mu w) = \lambda F(v) + \mu F(w) \text{ für alle } v, w \in V, \lambda, \mu \in K. \quad (\text{L})$$

Wir werden nun einige wichtige Spezialfälle von linearen Abbildungen diskutieren.

**Definition 3.3.** Wir nennen eine linear Abbildung  $F : V \rightarrow W$  zwischen  $K$ -Vektorräumen  $V$  und  $W$  einen

1. **Isomorphismus**, falls  $F$  bijektiv ist,
2. **Endomorphismus**, falls  $V = W$ ,
3. **Automorphismus**, falls  $V = W$  und  $F$  bijektiv ist.

Wir geben zwei Beispiele.

**Beispiel 3.4 (Transposition von Matrizen).** Die Darstellung von Vektoren als Zeilen- oder Spaltenvektoren ist eine Konvention (wir hatten  $\mathbb{R}^n$  als den Raum der  $n$ -dimensionalen Spaltenvektoren eingeführt). Das Vertauschen von Zeilen mit Spalten in einer Matrix nennt man **Transposition** und ist für viele Rechnungen extrem hilfreich.

**Definition 3.5.** Sei  $A \in M(m \times n, K)$ . Die Matrix

$$A^T := (a'_{ij}) \in M(n \times m, K) \text{ mit } a'_{ij} := a_{ji}, i = 1, \dots, m, j = 1, \dots, n$$

heißt **Transponierte Matrix** von  $A$ .

$$A = \begin{pmatrix} 1 & 0 & 3 \\ 0 & 2 & 1 \end{pmatrix} \text{ mit Transponierter } A^T = \begin{pmatrix} 1 & 0 \\ 0 & 2 \\ 3 & 1 \end{pmatrix}.$$

Die Transposition kann als Isomorphismus

$$^T : M(m \times n, K) \rightarrow M(n \times m, K), A \mapsto A^T$$

aufgefasst werden, für welchen die folgenden Rechenregeln gelten:

1.  $(A + B)^T = A^T + B^T$
2.  $(\lambda A)^T = \lambda A^T$
3.  $(A^T)^T = A$ .

**Beispiel 3.6 (Matrix-Vektor-Produkt).** Jede Matrix  $A \in M(m \times n, K)$  definiert eine Abbildung  $F_A : K^n \rightarrow K^m, x \mapsto A \cdot x$ . Durch Nachrechnen zeigt man leicht, dass jede solche Abbildung linear ist, d.h. für beliebige Spaltenvektoren  $x, y \in K^n$  und Skalare  $\lambda \in K$  gilt  $A \cdot (x + y) = A \cdot x + A \cdot y$  und  $A \cdot (\lambda x) = \lambda(A \cdot x)$ :

Seien  $x, y \in K^n$  zwei (Spalten-)Vektoren. Dann gilt:

$$\begin{aligned} A(x + y) &\stackrel{(\Delta)}{=} A(x_j + y_j)_{j=1, \dots, n} \stackrel{(*)}{=} \left( \sum_{j=1}^n a_{ij}(x_j + y_j) \right)_{i=1, \dots, m} \stackrel{(\#)}{=} \left( \sum_{j=1}^n (a_{ij}x_j + a_{ij}y_j) \right)_{i=1, \dots, m} \\ &\stackrel{(\#)}{=} \left( \left( \sum_{j=1}^n a_{ij}x_j \right) + \left( \sum_{j=1}^n a_{ij}y_j \right) \right)_{i=1, \dots, m} \\ &\stackrel{(\Delta)}{=} \left( \left( \sum_{j=1}^n a_{ij}x_j \right) \right)_{i=1, \dots, m} + \left( \left( \sum_{j=1}^n a_{ij}y_j \right) \right)_{i=1, \dots, m} \stackrel{(*)}{=} Ax + Ay. \end{aligned}$$

Dabei wenden wir in  $(*)$  die Definition des Matrix-Vektorprodukts an, in  $(\Delta)$  die der Vektoraddition und in  $(\#)$  die Rechenregeln in Körpern.

Analog gilt für einen (Spalten-)Vektor  $x \in K^n$  und ein Skalar  $\lambda \in K$ :

$$\begin{aligned} A(\lambda x) &\stackrel{(\circ)}{=} A(\lambda x_j)_{j=1, \dots, n} \stackrel{(*)}{=} \left( \sum_{j=1}^n a_{ij}(\lambda x_j) \right)_{i=1, \dots, m} \stackrel{(\#)}{=} \left( \lambda \sum_{j=1}^n a_{ij}x_j \right)_{i=1, \dots, m} \\ &\stackrel{(\circ)}{=} \lambda \left( \sum_{j=1}^n a_{ij}x_j \right)_{i=1, \dots, m} \stackrel{(*)}{=} \lambda(Ax). \end{aligned}$$

Hierbei wenden wir in (\*) die Definition des Matrix-Vektorprodukts an, in (○) die der Skalarmultiplikation von Vektoren und in (#) die Rechenregeln in Körpern.

Wir zeigen nun einige fundamentale Eigenschaften von linearen Abbildungen.

**Lemma 3.7.** Sei  $F : V \rightarrow W$  eine lineare Abbildung zwischen  $K$ -Vektorräumen  $V$  und  $W$ . Dann gilt:

1.  $F(0) = 0$  und  $F(v - w) = F(v) - F(w)$ .
2.  $F(\lambda_1 v_1 + \dots + \lambda_n v_n) = \lambda_1 F(v_1) + \dots + \lambda_n F(v_n)$ .
3. Ist die Familie  $(v_i)_{i \in I}$  linear abhängig in  $V$ , so ist  $(F(v_i))_{i \in I}$  linear abhängig in  $W$ .
4. Sind  $V' \subset V$  und  $W' \subset W$  Untervektorräume, so sind auch  $F(V') \subset W$  und  $F^{-1}(W') \subset V$  Untervektorräume.
5.  $\dim(F(V)) \leq \dim V$ .
6. Ist  $F$  ein Isomorphismus, so ist  $F^{-1} : W \rightarrow V$  linear.

*Beweis.* Zu 1.:  $F(0) = F(0 \cdot 0) = 0 \cdot F(0) = 0$ .

$$F(v - w) = F(v + (-w)) = F(v) + (-1)F(w) = F(v) - F(w).$$

Zu 2.: Wende wiederholt (L) an.

Zu 3.: Ist für  $i_j \in I, \lambda_j \in K, j = 1, \dots, n$

$$\lambda_1 v_{i_1} + \dots + \lambda_n v_{i_n} = 0,$$

so folgt mit 2. und  $F(0) = 0$ :

$$\lambda_1 F(v_{i_1}) + \dots + \lambda_n F(v_{i_n}) = 0.$$

Zu 4.: Mit  $0 \in V'$  gilt  $0 = F(0) \in F(V')$ . Sind  $w, w' \in F(V')$ , so gibt es  $v, v' \in V'$  mit  $F(v) = w$  und  $F(v') = w'$ . Wir erhalten mit  $v + v' \in V'$ :

$$w + w' = F(v) + F(v') = F(v + v') \in F(V').$$

Für  $\lambda \in K$  gilt wegen  $\lambda v \in V'$ :

$$\lambda w = \lambda F(v) = F(\lambda v) \in F(V').$$

Mit  $0 \in W'$  und  $F(0) = 0$  gilt  $0 \in F^{-1}(W')$ . Wenn  $v, v' \in F^{-1}(W')$ , so gilt  $F(v), F(v') \in W'$ . Also gilt  $F(v + v') = F(v) + F(v') \in W'$  und somit  $v + v' \in F^{-1}(W')$ . Analog folgt  $\lambda v \in F^{-1}(W')$ .

Zu 5.: Gilt  $w_1 = F(v_1), \dots, w_n = F(v_n) \in F(V)$  linear unabhängig, so sind nach 3. auch  $v_1, \dots, v_n \in V$  linear unabhängig.

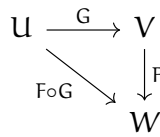


Abbildung 3.1: Illustration der Komposition von linearen Abbildungen.

Zu 6.: Seien  $w, w' \in W$  und  $\lambda, \mu \in K$ . Falls  $w = F(v)$  und  $w' = F(v')$ , so ist  $v = F^{-1}(w)$  und  $v' = F^{-1}(w')$  und mit Anwendung von  $F^{-1}$  auf beide Seiten der Gleichung

$$F(\lambda v + \mu v') = \lambda w + \mu w'$$

folgt

$$\lambda F^{-1}(w) + \mu F^{-1}(w') = F^{-1}(\lambda w + \mu w').$$

□

Die Komposition zweier linearen Abbildungen ergibt wieder eine lineare Abbildung:

**Lemma 3.8.** Seien  $U, V, W$  jeweils  $K$ -Vektorräume und  $G : U \rightarrow V$ ,  $F : V \rightarrow W$  lineare Abbildungen. Dann ist auch die folgende Abbildung linear:

$$F \circ G : U \rightarrow W.$$

*Beweis.* Für  $u, u' \in U$  gilt:

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

Analog wird  $(F \circ G)(\lambda u) = \lambda(F \circ G)(u)$  gezeigt. □

Die Menge der linearen Abbildungen zwischen zwei  $K$ -Vektorräumen  $V$  und  $W$  wird folgendermaßen bezeichnet:

$$\text{Hom}_K(V, W) := \{F : V \rightarrow W \mid F \text{ ist ein } K\text{-Homomorphismus}\}.$$

**Übung 3.9.** Die Menge  $\text{Hom}_K(V, W)$  mit den Verknüpfungen

$$(F + G)(v) := F(v) + G(v), \quad (\lambda F)(v) := \lambda F(v) \text{ für } F, G \in \text{Hom}_K(V, W), v \in V, \lambda \in K,$$

ist selbst wieder ein  $K$ -Vektorraum.

Wir befassen uns nun mit der Existenz von linearen Abbildungen, so dass eine vorgegebene Menge im Bild der linearen Abbildung auftaucht.

**Satz 3.10.** Gegeben seien  $K$ -Vektorräume  $V$  und  $W$  sowie  $v_1, \dots, v_r \in V$  und  $w_1, \dots, w_r \in W$ . Dann gilt:

1. Sind  $v_1, \dots, v_r$  linear unabhängig, so gibt es mindestens eine lineare Abbildung

$$F : V \rightarrow W \text{ mit } F(v_i) = w_i \text{ für } i = 1, \dots, r.$$

2. Ist  $\mathcal{B} = (v_1, \dots, v_r)$  eine Basis, so gibt es genau eine lineare Abbildung

$$F : V \rightarrow W \text{ mit } F(v_i) = w_i \text{ für } i = 1, \dots, r.$$

*Beweis.* Wir starten mit 2. Wir definieren zunächst  $F(v_i) := w_i, i = 1, \dots, r$ . Für  $v \in V$  beliebig, gibt es eine eindeutige Darstellung mittels der Basisvektoren (siehe Lemma 2.54, 2.):

$$v = \lambda_1 v_1 + \dots + \lambda_r v_r,$$

und somit erhalten wir eine Abbildung, die auf ganz  $V$  definiert ist:

$$F(v) = \lambda_1 F(v_1) + \dots + \lambda_r F(v_r).$$

Beachte, dass  $F(v)$  wohldefiniert ist, da die Darstellung von  $v$  mithilfe der Vektoren aus  $\mathcal{B}$  eindeutig ist. Wir zeigen nun, dass  $F$  eine lineare Abbildung ist. Dazu seien  $v, v' \in V$ . Wir erhalten

$$\begin{aligned} F(v + v') &= F\left(\sum_{i=1}^r \lambda_i v_i + \lambda'_i v_i\right) \\ &= F\left(\sum_{i=1}^r (\lambda_i + \lambda'_i) v_i\right) \\ &= \sum_{i=1}^r (\lambda_i + \lambda'_i) w_i \\ &= \sum_{i=1}^r \lambda_i w_i + \sum_{i=1}^r \lambda'_i w_i \\ &= F(v) + F(v'). \end{aligned}$$

Weiterhin gilt

$$F(\lambda v) = \sum_{i=1}^r \lambda \lambda_i w_i = \lambda \left(\sum_{i=1}^r \lambda_i w_i\right) = \lambda F(v).$$

Zu 1: Wir ergänzen  $v_1, \dots, v_r$  zu einer Basis

$$(v_1, \dots, v_r, v_{r+1}, \dots, v_n)$$

und geben beliebige Vektoren  $w_{r+1}, \dots, w_n \in W$  vor. Bei der Anwendung von 2. erhalten

wir ein  $F$  mit

$$F(v_i) = w_i, i = 1, \dots, n.$$

□

### 3.1 Bild, Faser und Kern von Homomorphismen

**Definition 3.11.** Sei  $F : V \rightarrow W$  eine lineare Abbildung zwischen  $K$ -Vektorräumen  $V$  und  $W$ . Dann nennen wir:

1.  $\text{Im } F := F(V)$  das **Bild** von  $F$ .
2.  $F^{-1}(w) := \{v \in V \mid F(v) = w\}$  die **Faser** von  $w \in W$ .
3.  $\text{Ker } F := \{v \in V \mid F(v) = 0\} = F^{-1}(0)$  den **Kern** von  $F$ .

**Satz 3.12.** Sei  $F : V \rightarrow W$  linear. Dann gilt:

1.  $\text{Im } F \subset W$  und  $\text{Ker } F \subset V$  sind Untervektorräume.
2.  $F$  surjektiv  $\Leftrightarrow \text{Im } F = W$ .
3.  $F$  injektiv  $\Leftrightarrow \text{Ker } F = \{0\}$ .
4. Ist  $F$  injektiv und sind  $v_1, \dots, v_n \in V$  linear unabhängig, so sind auch die Bilder  $F(v_1), \dots, F(v_n)$  linear unabhängig.

*Beweis.* Zu (1): Folgt direkt aus Lemma 3.7.4 durch die Wahl  $V' = V$  und  $W' = \{0\}$ .

Zu (2): Das ist ja die Definition von Surjektivität.

Zu (3):  $\Rightarrow$  ist klar. Für  $\Leftarrow$ : Seien  $v, v' \in V$  mit  $F(v) = F(v')$ . Dann gilt für  $v - v' \in V$ :

$$0 = F(v) - F(v') = F(v - v') \Rightarrow v - v' \in \text{Ker } F = \{0\},$$

also  $v = v'$ .

Zu (4):

$$\mu_1 F(v_1) + \dots + \mu_n F(v_n) = 0 \Leftrightarrow F(\mu_1 v_1 + \dots + \mu_n v_n) = 0.$$

Mit der Injektivität von  $F$  (nämlich  $\text{Ker } F = \{0\}$ ) bekommen wir  $\mu_1 v_1 + \dots + \mu_n v_n = 0$  und somit folgt aus der linearen Unabhängigkeit von  $v_1, \dots, v_n$ , dass  $\mu_1 = \dots = \mu_n = 0$ . □

**Definition 3.13.** Der **Rang** einer linearen Abbildung  $F$  ist definiert als

$$\text{rang } F := \dim(\text{Im } F) \tag{3.1}$$

Wenn wir uns die lineare Abbildung  $F_A$  anschauen, die durch eine Matrix  $A \in M(m \times n, K)$

mittels

$$F_A : K^n \rightarrow K^m, x \mapsto Ax,$$

definiert wird, so ist der Rang der Abbildung  $F_A$  gleich dem Rang der Matrix  $A$ , d.h. es gilt  $\text{rang } F_A = \text{rang } A$ . Für die kanonische Basis  $(e_1, \dots, e_n)$  von  $K^n$  sind nämlich

$$Ae_1, \dots, Ae_n$$

die Spalten von  $A$ . Somit folgt

$$\text{Im } F_A = F_A(K^n) = \text{span}(Ae_1, \dots, Ae_n) = \text{SR}(A),$$

wobei  $\text{SR}(A)$  den Spaltenraum von  $A$  bezeichnet.

Wir werden nun die Begriffe, Bild, Kern und Faser einer linearen Abbildung veranschaulichen.

**Beispiel 3.14.** Betrachte folgende lineare Abbildung:

$$F : \mathbb{R}^2 \rightarrow \mathbb{R}^2, \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \mapsto \begin{pmatrix} -2x_1 + 2x_2 \\ -x_1 + x_2 \end{pmatrix}.$$

Wir erhalten offensichtlich

$$\text{Im } F = \mathbb{R} \begin{pmatrix} 2 \\ 1 \end{pmatrix} \text{ und } \text{Ker } F = \mathbb{R} \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

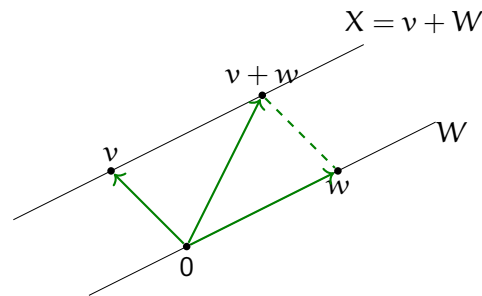
Für ein festes  $b \in \mathbb{R}$  gilt natürlich  $\begin{pmatrix} 2b \\ b \end{pmatrix} \in \text{Im } F$  und wir erhalten

$$F^{-1} \left( \begin{pmatrix} 2b \\ b \end{pmatrix} \right) = \begin{pmatrix} 0 \\ b \end{pmatrix} + \mathbb{R} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ b \end{pmatrix} + \text{Ker } F.$$

Die Faser von  $\begin{pmatrix} 2b \\ b \end{pmatrix}$  verläuft also **parallel** zum Kern von  $F$ .

**Lemma 3.15.** Sei  $F : V \rightarrow W$  linear und  $w \in \text{Im } F$  und  $u \in F^{-1}(w)$  beliebig. Dann gilt

$$F^{-1}(w) = u + \text{Ker } F := \{u + v \mid v \in \text{Ker } F\}.$$



*Beweis.* Für  $v' \in F^{-1}(w)$  gilt

$$\begin{aligned} F(v') &= F(u) \Rightarrow F(v' - u) = 0 \Rightarrow v := v' - u \in \text{Ker } F \\ \Rightarrow v' &= u + v \in u + \text{Ker } F. \end{aligned}$$

Umgekehrt, sei  $v' = u + v \in u + \text{Ker } F$ . Somit folgt  $F(v') = F(u) + 0 = F(u) = w$  und es gilt  $v' \in F^{-1}(w)$ .  $\square$

**Definition 3.16.** Eine Teilmenge  $X$  eines  $K$ -Vektorraums  $V$  heißt **affiner Unterraum**, falls es ein  $v \in V$  und einen Untervektorraum  $W$  gibt mit

$$X = v + W := \{u \in V \mid \exists w \in W \text{ mit } u = v + w\}.$$

Per Konvention ist die leere Menge auch ein affiner Unterraum.

**Lemma 3.17.** Sei  $X = v + W \subset V$  ein affiner Unterraum. Dann gilt:

1. Für ein beliebiges  $v' \in X$  ist  $X = v' + W$ .
2. Ist  $v' \in V$  und  $W' \subset V$  ein Untervektorraum mit  $v + W = v' + W'$ , so gilt  $W = W'$  und  $v' - v \in W$ .

Das Lemma besagt, dass zu einem affinen Unterraum  $v + W$  der Untervektorraum  $W$  eindeutig bestimmt ist, nur der „Aufhängepunkt“  $v$  kann in  $X$  frei gewählt werden.

*Beweis.* Zu (1): Sei  $v' = v + w'$ . Wir zeigen beide Inklusionen.  
 $X \subset v' + W$ :

$$\begin{aligned} u \in X &\Rightarrow u = v + w \text{ mit } w \in W \\ &\Rightarrow u = v' + (w - w') \\ &\Rightarrow u \in v' + W. \end{aligned}$$

$X \supset v' + W : u = v' + w \in v' + W \Rightarrow u = v + (w + w') \in v + W.$

Zu (2): Definiere

$$X - X := \{u - u' \mid u \in X, u' \in X\}$$

als die Menge der Differenzen. Wir erhalten

$$X - X = W \text{ und } X - X = W' \text{ und somit } W = W'.$$

Wegen  $v + W = v' + W$  gibt es ein  $w \in W$  mit  $v' = v + w$  also  $v' - v = w \in W$ .  $\square$

Lemma 3.17 zeigt, dass für einen nicht-leeren affinen Unterraum, der zugehörige Untervektorraum eindeutig bestimmt ist. Daher ist die Dimension eines nicht-leeren affinen Unterraums  $X = v + W$  mittels

$$\dim(X) := \dim(W)$$

wohldefiniert. Für den leeren affinen Unterraum  $X = \emptyset$  definieren wir zusätzlich  $\dim(\emptyset) := -1$ .

### 3.2 Dimensionsformel

**Satz 3.18.** Sei  $F : V \rightarrow W$  linear und  $\dim(V) < \infty$ . Sind Basen  $(v_1, \dots, v_k)$  von  $\text{Ker } F$  und  $(w_1, \dots, w_r)$  von  $\text{Im } F$ , sowie beliebige Vektoren  $u_1 \in F^{-1}(w_1), \dots, u_r \in F^{-1}(w_r)$  gegeben, so ist

$$\mathcal{B} := (u_1, \dots, u_r, v_1, \dots, v_k) \text{ eine Basis von } V.$$

Insbesondere gilt die [Dimensionsformel](#)

$$\dim(V) = \dim(\text{Im } F) + \dim(\text{Ker } F).$$

*Beweis.* Für  $v \in V$  sei

$$F(v) = \mu_1 w_1 + \dots + \mu_r w_r \text{ und } v' := \mu_1 u_1 + \dots + \mu_r u_r.$$

Mit  $F(v) = F(v')$  folgt  $v - v' \in \text{Ker } F$  und somit

$$v - v' = \lambda_1 v_1 + \dots + \lambda_k v_k.$$

Somit gilt

$$v = v' + \lambda_1 v_1 + \dots + \lambda_k v_k = \mu_1 u_1 + \dots + \mu_r u_r + \lambda_1 v_1 + \dots + \lambda_k v_k.$$

Also wird  $V$  von  $\mathcal{B}$  erzeugt. Zur linearen Unabhängigkeit, sei

$$\mu_1 u_1 + \dots + \mu_r u_r + \lambda_1 v_1 + \dots + \lambda_k v_k = 0.$$

Wende  $F$  an und wir erhalten:

$$\mu_1 w_1 + \cdots + \mu_r w_r = 0.$$

Mit der linearen Unabhängigkeit von  $(w_1, \dots, w_r)$  gilt  $\mu_1 = \cdots = \mu_r = 0$  und somit erhalten wir

$$\lambda_1 v_1 + \cdots + \lambda_k v_k = 0.$$

Dies impliziert mit der linearen Unabhängigkeit von  $(v_1, \dots, v_k)$ , dass  $\lambda_1 = \cdots = \lambda_k = 0$  gilt.  $\square$

Wir erhalten die Folgerung:

**Korollar 3.19.** Seien  $V, W$  endlichdimensionale  $K$ -Vektorräume. Dann gelten folgende Aussagen:

1. Ist  $F : V \rightarrow W$  linear, so gilt  $\dim(F^{-1}(w)) = \dim(V) - \dim(\text{Im } F)$  für alle  $w \in W$  mit  $F^{-1}(w) \neq \emptyset$ .
2. Es existiert ein Isomorphismus  $F : V \rightarrow W \Leftrightarrow \dim(V) = \dim(W)$ .
3. Sei  $F : V \rightarrow W$  linear und  $\dim(V) = \dim(W)$ . Dann sind äquivalent:
  - $F$  injektiv.
  - $F$  surjektiv.
  - $F$  bijektiv.

*Beweis.* Zu (1): Wir hatten in Lemma 3.15 gesehen, dass für  $w \in \text{Im } F$  und  $u \in F^{-1}(w)$  gilt  $F^{-1}(w) = u + \text{Ker } F$ . Die Aussage folgt dann unmittelbar mit Satz 3.18.

Zu (2):  $\Rightarrow$ : Sei  $F$  bijektiv, d.h., insbesondere surjektiv und injektiv. Mit Satz 3.12 folgt, dass  $\text{Im } F = W$  und  $\text{Ker } F = \{0\}$ . Satz 3.18 impliziert also  $\dim(V) = \dim(W)$ .

$\Leftarrow$ : Es gelte  $\dim(V) = \dim(W)$ . Wähle eine Basis  $\mathcal{B} = (v_1, \dots, v_n)$  von  $V$  und eine Basis  $\mathcal{B}' = (w_1, \dots, w_n)$  von  $W$ . Mit Satz 3.10 gibt es eine eindeutige lineare Abbildung  $F : V \rightarrow W$ , die  $F(v_i) = w_i, i = 1, \dots, n$  erfüllt. Somit gilt  $\text{Im } F = W$  und  $\text{Ker } F = \{0\}$  und  $F$  ist ein Isomorphismus.

Zu (3):  $F$  injektiv  $\Leftrightarrow \text{Ker } F = \{0\} \stackrel{\Leftrightarrow}{\dim(V)=\dim(W)} \text{Im } F = W \Leftrightarrow F$  surjektiv.  $F$  bijektiv  $\Leftrightarrow F$  injektiv  $\wedge$  surjektiv.  $\square$

### 3.3 Lineare Gleichungssysteme

Wir kommen nun zurück auf lineare Gleichungssysteme (LGS) über einen Körper  $K$ :

$$A \cdot x = b \text{ für eine Matrix } A \in M(m \times n, K) \text{ und } b \in K^m$$

in der Variable  $x \in K^n$ . Für  $b = 0$  nennen wir das zugehörige LGS  $A \cdot x = 0$  **homogen**. Wir erinnern uns an die Lösungsraum:

$$\text{Lös}(A, b) := \{x \in K^n \mid Ax = b\} \subset K^n.$$

Wir werden nun die Theorie der linearen Abbildungen auf LGS anwenden indem wir die folgende **von A induzierte** lineare Abbildung studieren:

$$F_A : K^n \rightarrow K^m, x \mapsto A \cdot x.$$

Wir sehen sofort, dass  $\text{Lös}(A, b) = F_A^{-1}(b)$  und  $\text{Lös}(A, 0) = \text{Ker } F_A$  gilt. Mit der Bezeichnung  $\text{rang } A = \text{rang } F_A = \text{Spaltenrang } A = \text{Zeilenrang } A$  folgt aus der bisher entwickelten Theorie:

**Korollar 3.20.** Gegeben sei ein LGS  $Ax = b$  mit  $A \in M(m \times n, K)$  und  $b \in K^m$  sowie  $\text{rang } A = r$ . Dann gilt:

1.  $\text{Lös}(A, 0)$  ist ein Untervektorraum von  $K^n$  der Dimension  $n - r$ .
2.  $\text{Lös}(A, b)$  ist entweder leer oder ein affiner Unterraum von  $K^n$  der Dimension  $n - r$ .  
Ist  $v \in \text{Lös}(A, b)$  beliebig, so gilt

$$\text{Lös}(A, b) = v + \text{Lös}(A, 0).$$

Um zu entscheiden, ob  $\text{Lös}(A, b)$  nicht-leer ist, schauen wir uns den Rang der erweiterten Koeffizientenmatrix an:

$$\text{rang}(A) =: r \leq \text{rang}(A, b) \leq r + 1.$$

**Satz 3.21.**  $\text{Lös}(A, b)$  ist nicht-leer  $\Leftrightarrow r = \text{rang}(A, b)$ .

*Beweis.* Wir betrachten die beiden durch  $A$  und  $A'$  induzierten linearen Abbildungen

$$F_A : K^n \rightarrow K^m, x \mapsto A \cdot x \text{ und } F_{A'} : K^{n+1} \rightarrow K^m, x' \mapsto A' \cdot x',$$

wobei  $A' = (A, b)$  gesetzt wird. Sind nun  $(e_1, \dots, e_n)$  und  $(e'_1, \dots, e'_{n+1})$  die kanonischen Basen, so erhalten wir

$$F_A(e_1) = F_{A'}(e'_1), \dots, F_A(e_n) = F_{A'}(e'_n) \text{ und } F_{A'}(e'_{n+1}) = b.$$

Es gilt also  $b \in \text{Im } F_{A'}$  und mit  $\text{Im } F_A \subset \text{Im } F_{A'}$  gilt natürlich  $\text{rang}(A) \leq \text{rang}(A')$ . Somit folgt

$$r = \text{rang}(A) = \text{rang}(A') \Leftrightarrow \text{Im } F_A = \text{Im } F_{A'} \Leftrightarrow b \in \text{Im } F_A \Leftrightarrow \text{Lös}(A, b) \neq \emptyset.$$

□

$$A = \left( \begin{array}{cccccccccccc|cccc} a_{11} & (*) & (*) & (*) & (*) & (*) & (*) & \dots & \dots & \dots & \dots & \dots & b_1 \\ 0 & a_{22} & (*) & (*) & (*) & (*) & (*) & \dots & \dots & \dots & \dots & \dots & b_2 \\ 0 & 0 & a_{33} & (*) & (*) & (*) & (*) & \dots & \dots & \dots & \dots & \dots & b_3 \\ 0 & 0 & 0 & a_{44} & (*) & (*) & (*) & (*) & \dots & \dots & \dots & \dots & \vdots \\ \vdots & & & & \vdots & & & & & & & & \vdots \\ 0 & 0 & 0 & \dots & \dots & \dots & \dots & \dots & 0 & a_{rr} & (*) & \dots & b_r \\ 0 & 0 & 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & 0 & 0 & b_{r+1} \\ 0 & 0 & 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & 0 & 0 & \vdots \\ \vdots & & & & \vdots & & & & & & & & \vdots \\ 0 & 0 & 0 & \dots & \dots & \dots & \dots & \dots & \dots & \dots & 0 & 0 & b_m \end{array} \right)$$

Abbildung 3.2: Allgemeine Matrix in Zeilenstufenform nach eventueller Vertauschung von Spalten.

Wir betrachten nun wieder die Situation, dass wir die erweiterte Matrix  $(A, b)$  in Zeilenstufenform gegeben haben. Wie wir schon gesehen hatten, kann jede erweiterte Matrix  $(A, b)$  in Zeilenstufenform gebracht werden, ohne dass sich der Lösungsraum  $\text{Lös}(A, b)$  ändert. In Abbildung 3.3 ist eine allgemeine Matrix in Zeilenstufenform angegeben, bei der wir annehmen, dass die ersten  $r$  Spalten die Pivot-Spalten sind. Dies kann o.E. durch eventuelle Vertauschung von Spalten erreicht werden. In der Darstellung 3.3 gilt

$$\text{rang}(A) = \text{rang}(A, b) \Leftrightarrow b_{r+1} = \dots = b_m = 0.$$

Im Folgenden werden wir uns die Parametrisierung des Lösungsraums für eine Matrix in Zeilenstufenform, für die obige Bedingung erfüllt ist, genauer anschauen. Zur Erinnerung:

$$\text{Lös}(A, b) = \left\{ x \in K^n \mid x_i = \frac{1}{a_{ii}} \left( b_i - \sum_{j=i+1}^n a_{ij} x_j \right) \text{ für } i \in \{1, \dots, r\} \right\}.$$

Wenn wir nun  $\text{Lös}(A, b)$  bzgl. der freien Variablen  $x_{r+1}, \dots, x_n$  parametrisieren, so setzen wir  $\lambda_1 := x_{r+1}, \dots, \lambda_k := x_n$  mit  $n - r =: k$ . Hier sind  $x_1, \dots, x_r$  die gebundene Variablen, welche eindeutig dadurch festgelegt sind, welche Werte die freien Variablen  $\lambda_i, i = 1, \dots, k$  angenommen haben. Zur Berechnung von  $x_1, \dots, x_r$  beginnen wir mit der  $r$ -ten Gleichung:

$$a_{rr}x_r + a_{r,r+1}\lambda_1 + \dots + a_{rn}\lambda_n = b_r$$

woraus folgt

$$x_r(\lambda) := \frac{1}{a_{rr}}(b_r - a_{r,r+1}\lambda_1 - \dots - a_{rn}\lambda_k),$$

wobei  $\lambda = (\lambda_1, \dots, \lambda_k)$ . Eingesetzt in die  $(r-1)$ -te Gleichung führt zu

$$x_{r-1}(\lambda) := d_{r-1,r-1}b_{r-1} + d_{r-1,r}b_r + c_{r-1,1}\lambda_1 + \dots + c_{r-1,k}\lambda_k,$$

wobei die auftretenden Zahlen  $c$  und  $d$  von den Einträgen in den Zeilen  $r-1$  und  $r$  abhängen. Iterieren wir dieses Vorgehen erhalten wir schließlich

$$x_1(\lambda) := d_{11}b_1 + \dots + d_{1r}b_r + c_{11}\lambda_1 + \dots + c_{1k}\lambda_k.$$

Wenn wir nun die Einträge  $d, c$  in Matrizen eintragen, erhalten wir

$$D' := (d_{ij}) \in M(r \times r, K), \quad C' := (c_{ij}) \in M(r \times k, K).$$

Wir erweitern diese Matrizen zu

$$C := \begin{pmatrix} C' \\ E_k \end{pmatrix} \in M(n \times k, K) \text{ und } D := \begin{pmatrix} D' \\ 0 \end{pmatrix} \in M(n \times r, K),$$

wobei  $E_k$  die  $k$ -dimensionale Einheitsmatrix darstellt:

$$E_k = \begin{pmatrix} 1 & & 0 \\ & \ddots & \\ 0 & & 1 \end{pmatrix} \in M(k \times k, K).$$

Das ergibt lineare Abbildungen:

$$\phi : K^r \rightarrow K^n, \quad b \mapsto D \cdot b \text{ und } \Phi_0 : K^k \rightarrow K^n, \quad \lambda \mapsto C \cdot \lambda,$$

wobei  $b = \begin{pmatrix} b_1 \\ \vdots \\ b_r \end{pmatrix} \in K^r$ . Insgesamt erhalten wir gemäß obiger Rechnung eine Abbildung

$$\begin{aligned} \Phi_b : K^k &\rightarrow K^n \\ \lambda = (\lambda_1, \dots, \lambda_k) &\mapsto \phi(b) + \Phi_0(\lambda). \end{aligned}$$

Nach der Konstruktion von  $\Phi_b$  gilt für alle  $b \in K^r$ :

$$\Phi_b(K^k) \subset \text{Lös}(A, b) = \phi(b) + \text{Lös}(A, 0).$$

Für  $b = 0$  gilt, dass  $\Phi_0$  injektiv ist, weil  $\text{rang } C = k$  und somit  $\text{Ker } C = \{0\}$ . Weiterhin ist  $\dim(\text{Lös}(A, 0)) = k$  (siehe Korollar 3.20) und somit folgt:

$$\Phi_0(K^k) = \text{Lös}(A, 0) \text{ und } \Phi_b(K^k) = \text{Lös}(A, b).$$

Wir halten den folgenden Satz über die Parametrisierung von Lösungsräumen von LGS mit Matrizen in Zeilenstufenform fest:

**Satz 3.22.** Sei  $(A, b)$  in Zeilenstufenform mit  $\text{rang } A = r$  und  $b \in K^r$ . Dann hat die oben konstruierte Abbildung  $\Phi_b : K^{n-r} \rightarrow \text{Lös}(A, b) \subset K^n$  folgende Eigenschaften:

1.  $\Phi_b$  ist für jedes  $b \in K^r$  bijektiv.
2. Es gibt einen Homomorphismus  $\phi : K^r \rightarrow K^n$ , sodass für alle  $b \in K^r$  gilt:

$$\Phi_b = \phi(b) + \Phi_0 \text{ und } \text{Lös}(A, b) = \phi(b) + \text{Lös}(A, 0).$$

Ein Spezialfall bei der Lösung von LGS tritt auf, wenn die Lösung **eindeutig** ist.

**Satz 3.23.** Gegeben sei ein LGS  $Ax = b$  mit  $A \in M(m \times n, K)$  und  $b \in K^m$ . Dann sind folgende Bedingungen äquivalent:

1. Das LGS  $Ax = b$  ist eindeutig lösbar.
2.  $\text{rang } A = \text{rang}(A, b) = n$ .

*Beweis.* Wir hatten schon gesehen, dass gilt  $\text{Lös}(A, b) \neq \emptyset \Leftrightarrow \text{rang } A = \text{rang}(A, b)$ . Bzgl. der Eindeutigkeit nutzen wir:  $\text{Lös}(A, b) = \phi(b) + \text{Lös}(A, 0)$  mit  $\dim(\text{Lös}(A, 0)) = n - \text{rang } A$ . □

### 3.4 Lineare Abbildungen und Matrizen

In diesem Abschnitt wird es darum gehen, lineare Abbildungen mit Hilfe von Matrizen darzustellen. Wir starten mit einem Korollar des Darstellungssatz bzgl. vorgegebener Basen, siehe Satz 3.10.

**Korollar 3.24.** Gegeben sind  $K$ -Vektorräume  $V$  und  $W$ . Sei  $(v_1, \dots, v_r)$  eine Basis von  $V$  und  $w_1, \dots, w_r \in W$ . Für die eindeutige lineare Abbildung

$$F : V \rightarrow W \text{ mit } F(v_i) = w_i \text{ für } i = 1, \dots, r.$$

gelten folgende Eigenschaften:

- (a)  $\text{Im } F = \text{span}(w_1, \dots, w_r)$
- (b)  $F$  injektiv  $\Leftrightarrow w_1, \dots, w_r$  linear unabhängig.

*Beweis.* Zu (a):  $\text{Im } F \subset \text{span}(w_1, \dots, w_r)$  ist trivial. Ist umgekehrt  $w = \mu_1 w_1 + \dots + \mu_r w_r$  so folgt  $w = F(\mu_1 v_1 + \dots + \mu_r v_r)$ .

Zu (b):  $\Rightarrow$ : Angenommen,  $w_1, \dots, w_r$  sind linear abhängig. Dann gibt es  $(\mu_1, \dots, \mu_r) \neq 0$  mit

$$\mu_1 w_1 + \dots + \mu_r w_r = 0,$$

und somit  $F(\mu_1 v_1 + \dots + \mu_r v_r) = 0$ , also ist  $F$  nicht injektiv.

$\Leftarrow$ : Sei  $F(v) = 0$ . Mit  $v = \lambda_1 v_1 + \dots + \lambda_r v_r$  folgt

$$F(\lambda_1 v_1 + \dots + \lambda_r v_r) = \lambda_1 w_1 + \dots + \lambda_r w_r = 0.$$

Da  $w_i, i = 1, \dots, r$  linear unabhängig, folgt  $\lambda_i = 0, i = 1, \dots, r$  und somit gilt  $v = 0$ .  $\square$

Wir erhalten weitere Folgerungen aus Satz 3.10.

**Korollar 3.25.** Zu jeder linearen Abbildung  $F : K^n \rightarrow K^m$  gibt es genau eine Matrix  $A \in M(m \times n, K)$  mit  $F(x) = A \cdot x$  für alle  $x \in K^n$ .

*Beweis.*  $e_1, \dots, e_n$  bildet eine Basis des  $K^n$  und wir können  $w_i := F(e_i), i = 1, \dots, n$  setzen. Wir schreiben nun  $F(e_1), \dots, F(e_n)$  als Spaltenvektoren in eine Matrix  $A$ . Die resultierende Abbildung  $F_A : K^n \rightarrow K^m$  ist linear und erfüllt  $F_A(e_i) = w_i, i = 1, \dots, n$ . Satz 3.10 besagt, dass es genau eine lineare Abbildung  $F : K^n \rightarrow K^m$  mit  $F(e_i) = w_i$  für alle  $i = 1, \dots, n$  gibt und somit folgt  $F = F_A$ .  $\square$

Jede lineare Abbildungen zwischen Standardvektorräumen ist also als Matrix-Abbildung darstellbar.

**Korollar 3.26.** Sei  $V$  ein  $K$ -Vektorraum mit Basis  $\mathcal{B} = (v_1, \dots, v_n)$ . Dann gibt es genau einen Isomorphismus

$$\Phi_{\mathcal{B}} : K^n \rightarrow V \text{ mit } \Phi_{\mathcal{B}}(e_j) = v_j, j = 1, \dots, n.$$

Die Abbildung  $\Phi_{\mathcal{B}}$  heißt **Koordinatensystem** und wir werden uns in den folgenden Abschnitten weiter damit beschäftigen.

Nun werden wir den Darstellungssatz von linearen Abbildungen mittels Matrizen für allgemeine  $K$ -Vektorräume  $V$  erweitern.

**Satz 3.27.** Seien  $V$  und  $W$  jeweils  $K$ -Vektorräume mit Basen  $\mathcal{A} = (v_1, \dots, v_n)$  und  $\mathcal{B} = (w_1, \dots, w_m)$ . Dann gilt:

1. Zu jeder linearen Abbildung  $F : V \rightarrow W$  gibt es genau eine Matrix  $A \in M(m \times n, K)$  mit

$$F(v_j) = \sum_{i=1}^m a_{ij} w_i \text{ für } j = 1, \dots, n. \quad (3.2)$$

2. Die resultierende Abbildung

$$M_{\mathcal{B}}^{\mathcal{A}} : \text{Hom}(V, W) \rightarrow M(m \times n, K), F \mapsto M_{\mathcal{B}}^{\mathcal{A}}(F) := A \quad (3.3)$$

ist ein Isomorphismus zwischen  $K$ -Vektorräumen.

*Beweis.* Zu 1: Da  $\mathcal{B}$  Basis von  $W$  sind die Linearkombinationen in (3.2) eindeutig und somit auch die Einträge der Matrix  $A$  eindeutig bestimmt.

Zu 2: Seien  $F, G : V \rightarrow W$  zwei lineare Abbildungen und  $A, B \in M(m \times n, K)$  die darstellende Matrizen von  $F$  bzw.  $G$  bzgl. der Basen  $\mathcal{A}$  und  $\mathcal{B}$  (siehe 3.2). Dann gilt:

$$(F + G)(v_j) = F(v_j) + G(v_j) = \sum_{i=1}^m a_{ij}w_i + \sum_{i=1}^m b_{ij}w_i = \sum_{i=1}^m (a_{ij} + b_{ij})w_i$$

und somit

$$M_{\mathcal{B}}^{\mathcal{A}}(F + G) = A + B = M_{\mathcal{B}}^{\mathcal{A}}(F) + M_{\mathcal{B}}^{\mathcal{A}}(G).$$

Analog kann  $M_{\mathcal{B}}^{\mathcal{A}}(\lambda F) = \lambda M_{\mathcal{B}}^{\mathcal{A}}(F)$  für  $\lambda \in K$  gezeigt werden.  $M_{\mathcal{B}}^{\mathcal{A}}$  ist also eine lineare Abbildung. Bijektivität von  $M_{\mathcal{B}}^{\mathcal{A}}$  folgt aus Satz 3.10: Für  $A \in M(m \times n, K)$  beliebig, erhalten wir  $n$  Vektoren in  $W$ :

$$w'_j := \sum_{i=1}^m a_{ij}w_i, j = 1, \dots, n.$$

Da  $\mathcal{A}$  eine Basis ist, folgt mit Satz 3.10, dass es eine eindeutige lineare Abbildung  $F$  gibt mit  $F(v_j) = w'_j, j = 1, \dots, n$ . Wir erhalten  $M_{\mathcal{B}}^{\mathcal{A}}(F) = A$  (Surjektivität) mit einzigem  $F$  (Injektivität).  $\square$

**Beispiel 3.28.** Betrachten wir zwei Standardvektorräume  $K^n$  und  $K^m$  zusammen mit ihren kanonischen Basen  $\mathcal{A} = (e_1, \dots, e_n)$  bzw.  $\mathcal{B} = (e'_1, \dots, e'_m)$ . Dann gilt für jede Matrix  $A \in M(m \times n, K)$  und die dadurch definierte lineare Abbildung  $F_A : K^n \rightarrow K^m, x \mapsto Ax$ , dass  $M_{\mathcal{B}}^{\mathcal{A}}(F_A) = A$ .

Obiger Satz zeigt, wie wir lineare Abbildungen als Vektoren in einem Vektorraum auffassen können. Z.B. können wir eine Basis von  $\text{Hom}(V, W)$  durch folgende Setzung bekommen:

$$F_i^j : V \rightarrow W, F_i^j(v_k) := \begin{cases} w_i, & \text{für } k = j, \\ 0, & \text{sonst.} \end{cases}$$

Mit dieser Definition gilt  $M_{\mathcal{B}}^{\mathcal{A}}(F_i^j) = E_i^j$ , wobei

$$E_i^j = \begin{pmatrix} 0 \dots 0 & \dots & 0 \\ \vdots & \vdots & \vdots \\ 0 \dots 0 & \dots & 0 \\ 0 \dots 1 & \dots & 0 \\ 0 \dots 0 & \dots & 0 \\ \vdots & \vdots & \vdots \\ 0 \dots 0 & \dots & 0 \end{pmatrix} \in M(m \times n, K),$$

und die 1 steht in der  $i$ -ten Zeile und der  $j$ -ten Spalte. Die so definierten  $m \cdot n$  linearen Abbildungen  $F_i^j$  bilden eine Basis des  $\text{Hom}(V, W)$ . Die zur Linearkombination benötigten Skalare in  $K$  zur Darstellung einer beliebigen linearen Abbildung  $F$  stehen genau in der Matrix  $M_{\mathcal{B}}^{\mathcal{A}}(F)$ .

**Korollar 3.29.** Sei  $F : V \rightarrow W$  linear,  $n = \dim V$ ,  $m = \dim W$  und  $r = \dim(\text{Im } F)$ . Dann gibt es Basen  $\mathcal{A}$  von  $V$  und  $\mathcal{B}$  von  $W$ , so dass gilt:

$$M_{\mathcal{B}}^{\mathcal{A}}(F) = \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix} \in M(m \times n, K).$$

*Beweis.* Seien  $w_1, \dots, w_r$  eine Basis von  $\text{Im } F$ . Wir wählen  $\mathcal{A}$  wie in Satz 3.18:

$$\mathcal{A} = (u_1, \dots, u_r, v_1, \dots, v_k),$$

wobei  $u_i \in F^{-1}(w_i)$ ,  $i = 1, \dots, r$  und  $v_j \in \text{Ker } F$ ,  $j = 1, \dots, k$ . Des Weiteren ergänzen wir  $(w_1, \dots, w_r)$  zu einer Basis von  $W$ :

$$\mathcal{B} := (w_1, \dots, w_r, w_{r+1}, \dots, w_m).$$

Somit erhalten wir, dass  $M_{\mathcal{B}}^{\mathcal{A}}(F)$  die geforderte Struktur besitzt. □

Für Endomorphismen ( $F : V \rightarrow V$ ) können wir  $\mathcal{A} = \mathcal{B}$  setzen und erhalten  $M_{\mathcal{B}}^{\mathcal{A}}(F) = M_{\mathcal{B}}^{\mathcal{B}}(F) =: M_{\mathcal{B}}(F)$ . Der Isomorphismus  $M_{\mathcal{B}} : \text{End}(V) \rightarrow M(n \times n, K)$  ist charakterisiert durch

$$F(v_j) = \sum_{i=1}^n a_{ij} v_i \text{ für } j = 1, \dots, n,$$

mit  $\mathcal{B} = (v_1, \dots, v_n)$  und  $A = M_{\mathcal{B}}(F)$ .

**3.5 Quotientenvektorräume**

Wir zeigen nun, dass sich eine lineare Abbildung  $F : V \rightarrow W$  für eine direkte Summe  $V = U \oplus \text{Ker } F$  **faktorisieren** lässt indem eine Hintereinanderschaltung einer eingeschränkten Abbildung  $F|_U$  und einer Projektion  $P$  von  $V$  auf  $U$  konstruiert wird.

**Satz 3.30 (Faktorisierungssatz).** Sei  $F : V \rightarrow W$  linear und

$$\mathcal{B} = (u_1, \dots, u_r, v_1, \dots, v_k) \text{ eine Basis von } V \text{ mit } \text{Ker } F = \text{span}(v_1, \dots, v_k)$$

und  $U = \text{span}(u_1, \dots, u_r)$ . Dann gilt:

1.  $V = U \oplus \text{Ker } F$ .
2.  $F|_U : U \rightarrow \text{Im } F$  ist ein Isomorphismus.
3. Sei  $P : U \oplus \text{Ker } F \rightarrow U, v = u + v' \mapsto u$  die Projektion auf den ersten Summanden. Dann gilt  $F = (F|_U) \circ P$ .

*Beweis.* Zu 1.: Diese Aussage folgt direkt aus Satz 2.76.

Zu 2.: Mit  $\text{Ker } F|_U = \text{Ker } F \cap U = \{0\}$  gilt, dass  $F|_U$  injektiv ist und somit ein Isomorphismus ist ( $F|_U : U \rightarrow \text{Im } F$  ist per Konstruktion surjektiv).

Zu 3.: Sei  $v = u + v' \in V$  mit  $u \in U$  und  $v' \in \text{Ker } F$ . Dann gilt  $P(v) = u$  und wir erhalten

$$F(v) = F(u) + F(v') = F(u) = F|_U(u) = F|_U(P(v)).$$

□

In Lemma 3.15 wurde gezeigt, dass die Fasern von  $F$  die zum Untervektorraum  $\text{Ker } F$  parallel verschobenen affinen Untervektorräume sind. Nun werden zu einem gegebenen Untervektorraum  $U \subset V$  eine lineare Abbildung  $F$  mit  $\text{Ker } F = U$  konstruieren, wobei nun die Aufgabe darin besteht ein geeignetes Bild  $W$  für ein solches  $F$  zu finden. Es wird sich herausstellen, dass  $W = V/U$  ein passender Kandidat ist, wobei  $V/U$  den **Quotientenraum** von  $V$  nach  $U$  bezeichnet.

Sei  $V$  ein  $K$ -Vektorraum und  $U \subset V$  ein Untervektorraum. Für  $v, v' \in V$  definieren wir eine **Äquivalenzrelation modulo  $U$**  wie folgt:

$$v \sim_U v' :\Leftrightarrow v' - v \in U.$$

**| Übung 3.31.** Zeige dass  $\sim_U$  eine Äquivalenzrelation auf  $V$  definiert.

Es folgt unmittelbar, dass die Äquivalenzklasse eines  $v \in V$  dem affinen Untervektorraum

$$\{v' \in V \mid v' \sim_U v\} = v + U$$

entspricht (benutze hier, dass  $v' - v \in U \Leftrightarrow \exists u \in U : v' = v + u$ .) Die Menge der Äquivalenzklassen wird mit  $V/U$  bezeichnet und heißt der **Quotientenraum** von  $V$  nach  $U$ . Diese Bezeichnung ergibt insofern Sinn als dass alle Elemente aus  $U$  in der Äquivalenzklasse von 0 liegen und in diesem Sinne herausdividiert werden. Wir führen nun die **kanonische Abbildung**  $\rho$  ein, welche jedem Vektor  $v \in V$  die zugehörige Äquivalenzklasse zuordnet:

$$\rho : V \rightarrow V/U = \{v + U \mid v \in V\}, v \mapsto \rho(v) = v + U.$$

**Satz 3.32.** Sei  $V$  ein  $K$ -Vektorraum und  $U \subset V$  ein Untervektorraum. Dann gibt es eindeutig definierte Verknüpfungen  $\tilde{+} : V/U \times V/U \rightarrow V/U$  und  $\tilde{\cdot} : K \times V/U \rightarrow V/U$ , so dass

- $V/U$  ein  $K$ -Vektorraum und
- $\rho : V \rightarrow V/U, v \mapsto \rho(v) = v + U$  linear wird.

Weiterhin gilt:

1.  $\rho$  ist surjektiv.
2.  $\text{Ker } \rho = U$ .
3. Für  $\dim V < \infty$  gilt  $\dim(V/U) = \dim V - \dim U$
4. Ist  $F : V \rightarrow W$  eine lineare Abbildung mit  $U \subset \text{Ker } F$ , so existiert genau eine Abbildung  $G : V/U \rightarrow W$  mit  $F = G \circ \rho$  und  $\text{Ker } G = (\text{Ker } F)/U$ .

*Beweis.* Wir konstruieren zunächst die beiden Verknüpfungen  $\tilde{+}$  und  $\tilde{\cdot}$ : Da  $\rho$  linear sein soll, muss gelten:

$$\begin{aligned} (v + U) \tilde{+} (w + U) &= \rho(v) \tilde{+} \rho(w) = \rho(v + w) = (v + w) + U \\ \lambda \tilde{\cdot} (v + U) &= \lambda \tilde{\cdot} \rho(v) = \rho(\lambda v) = \lambda v + U \end{aligned}$$

Somit gibt es nur eine Möglichkeit die beiden Verknüpfungen zu definieren:

$$(v + U) \tilde{+} (w + U) := (v + w) + U, \lambda \tilde{\cdot} (v + U) := \lambda v + U.$$

Wir müssen nun zeigen, dass diese Definition wohldefiniert ist, d.h. von der Wahl des Repräsentanten  $v$  und  $w$  unabhängig ist. Dazu seien  $v'$  und  $w'$  weitere Repräsentanten aus der jeweiligen Äquivalenzklasse von  $v$  und  $w$ :

$$v + U = v' + U \text{ und } w + U = w' + U.$$

Wir erhalten  $v' - v \in U, w' - w \in U$  und  $(v' + w') - (v + w) \in U$  und somit

$$(v + w) + U = (v' + w') + U.$$

Analog folgt auch  $\lambda v + U = \lambda v' + U$ . Die Vektorraumaxiome für  $V/U$  können ohne Probleme nachgeprüft werden (Übung). Der Nullvektor in  $V/U$  ist  $U$ , denn es gilt

$$(v + U) \tilde{+} U = (v + U) \tilde{+} (0 + U) = (v + 0) + U = v + U.$$

Das zu  $v + U$  additive Inverse ist  $-v + U$ . Diese Rechnungen verdeutlichen, dass wir mit Untervektorräumen genau so wie mit Vektoren rechnen können und somit lassen wir nun die Symbole  $\tilde{+}$  und  $\tilde{\cdot}$  von nun an weg und schreiben wieder wie üblich  $+$  und  $\cdot$ .

Zu (1): Dies folgt direkt aus der Definition von  $\rho$  und  $V/U$ .

Zu (2): Für  $v \in V$  gilt:  $v \in \text{Ker } \rho \iff v + U = \rho(v) = U \iff v \in U$ .

Zu (3): Wir wenden die Dimensionsformel aus Satz 3.18 an und erhalten:

$$\dim V = \dim(\text{Ker } \rho) + \dim(\text{Im } \rho) = \dim(U) + \dim(V/U).$$

Zu (4): Wir definieren  $G : V/U \rightarrow W$  durch

$$G(v + U) := F(v).$$

$G$  ist wohldefiniert, denn für  $v + U = v' + U$  folgt

$$v - v' \in U \subset \text{Ker } F \text{ also } F(v) = F(v'),$$

und somit ist der Wert  $G(v + U)$  unabhängig von der Auswahl des Repräsentanten der Äquivalenzklasse. Die Linearität von  $G$  folgt ebenso aus der Linearität von  $F$ . Die Definition von  $G$  impliziert dann direkt  $F = G \circ \rho$ . Weiterhin bekommen wir

$$v + U \in \text{Ker } G \iff v \in \text{Ker } F \iff v + U \in (\text{Ker } F)/U,$$

woraus  $\text{Ker } G = (\text{Ker } F)/U$  folgt. □

Wir betrachten im Folgenden die direkte Summe von zwei Untervektorräumen  $V = V_1 \oplus V_2$ . Zur Darstellung des Quotientenraumes  $V/V_2$  können wir nun die kanonische Abbildung eingeschränkt auf den direkten Summand  $V_1$  benutzen.

**Satz 3.33.** Sei  $V = V_1 \oplus V_2$  und  $\rho : V \rightarrow V/V_2$  die kanonische Abbildung. Dann ist

$$\rho' := \rho|_{V_1} : V_1 \rightarrow V/V_2 \text{ ein Isomorphismus.}$$

*Beweis.* Jedes  $v \in V$  hat die eindeutige Darstellung  $v = v_1 + v_2$  mit  $v_1 \in V_1, v_2 \in V_2$ . Weiterhin erhalten wir:

$$\rho(v) = \rho(v_1 + v_2) = v_1 + v_2 + V_2 = v_1 + V_2 = \rho'(v_1).$$

Somit ist  $\rho'$  bijektiv. □

### 3.6 Multiplikation von Matrizen

Wir hatten in Lemma 3.8 gesehen, dass die Komposition von linearen Abbildungen wieder linear ist. Wir werden nun eine solche Komposition betrachten, bei der die involvierten Abbildungen durch Matrizen dargestellt sind. Dazu betrachten wir lineare Matrix-Abbildungen der Form

$$F_B : K^r \rightarrow K^n, x \mapsto B \cdot x, \text{ und } F_A : K^n \rightarrow K^m, x \mapsto A \cdot x,$$

für Matrizen  $A \in M(m \times n, K)$  und  $B \in M(n \times r, K)$ . Für  $y = B \cdot x$  und  $z = A \cdot y$  rechnen wir nach

$$y_j = b_{j1}x_1 + \cdots + b_{jr}x_r, \quad j = 1, \dots, n \quad (3.4)$$

$$z_i = a_{i1}y_1 + \cdots + a_{in}y_n, \quad i = 1, \dots, m \quad (3.5)$$

Mit dieser Rechnung können wir die Komposition  $F_A(F_B(x))$  darstellen. Setzen wir (3.4) in (3.5) ein, so erhalten wir

$$\begin{aligned} z_i &= a_{i1}(b_{11}x_1 + \cdots + b_{1r}x_r) + \cdots + a_{in}(b_{n1}x_1 + \cdots + b_{nr}x_r) \\ &= (a_{i1}b_{11} + \cdots + a_{in}b_{n1})x_1 + \cdots + (a_{i1}b_{1r} + \cdots + a_{in}b_{nr})x_r. \end{aligned}$$

Setzen wir nun

$$c_{ik} := a_{i1}b_{1k} + \cdots + a_{in}b_{nk}, \quad k = 1, \dots, n, i = 1, \dots, m,$$

so erhalten wir

$$\begin{aligned} z_i &= \sum_{j=1}^n a_{ij}y_j = \sum_{j=1}^n \left( \sum_{k=1}^r b_{jk}x_k \right) = \sum_{j=1}^n \left( \sum_{k=1}^r a_{ij}b_{jk}x_k \right) \\ &= \sum_{k=1}^r \left( \sum_{j=1}^n a_{ij}b_{jk} \right) x_k = \sum_{k=1}^r c_{ik}x_k. \end{aligned}$$

**Definition 3.34.** Seien  $A \in M(m \times n, K)$  und  $B \in M(n \times r, K)$ . Dann ist das Produkt von  $A$  und  $B$  definiert durch

$$A \cdot B := C \text{ wobei } C = (c_{ik})_{i=1, \dots, m, k=1, \dots, r} \text{ und } c_{ik} = \sum_{j=1}^n a_{ij}b_{jk} \quad (3.6)$$

Wir leiten nun weitere Rechenregeln für die Matrizenmultiplikation her.

$$\begin{array}{c} \uparrow m \\ \left( \begin{array}{ccc} \leftarrow n \rightarrow \\ \bullet \dots \bullet \\ \leftarrow n \rightarrow \end{array} \right) \\ \downarrow \\ \text{Zeile } i \end{array} \times \begin{array}{c} \uparrow n \\ \left( \begin{array}{ccc} \leftarrow r \rightarrow \\ \bullet \\ \vdots \\ \bullet \\ \leftarrow r \rightarrow \end{array} \right) \\ \downarrow \\ \text{Spalte } k \end{array} = \begin{array}{c} \uparrow m \\ \left( \begin{array}{ccc} \leftarrow r \rightarrow \\ \bullet \\ \vdots \\ \bullet \\ \leftarrow r \rightarrow \end{array} \right) \\ \downarrow \\ c_{ik} \end{array}$$

Abbildung 3.3: Illustration der Multiplikation  $A \cdot B = C$ .

**Lemma 3.35.** Sind Matrizen  $A, A' \in M(m \times n, K)$  und  $B, B' \in M(n \times r, K), C \in M(r \times s, K)$  und  $\lambda \in K$  gegeben, so gilt:

1.  $A \cdot (B + B') = A \cdot B + A \cdot B'$  und  $(A + A') \cdot B = A \cdot B + A' \cdot B$
2.  $A \cdot (\lambda B) = (\lambda A) \cdot B = \lambda(A \cdot B)$
3.  $(A \cdot B) \cdot C = A \cdot (B \cdot C)$ .
4.  $(A \cdot B)^T = B^T \cdot A^T$ .
5.  $E_m \cdot A = A \cdot E_n = A$ .

**Übung 3.36.** • Beweise Lemma 3.35.

- Zeige, dass mit den Rechenregeln aus Lemma 3.35 folgt, dass die Menge  $M(n \times n, K)$  mit der Addition und der Multiplikation von Matrizen einen Ring  $(M(n \times n, K), +, \cdot)$  bildet.

**Bemerkung 3.37.** Es ist bei obigen Rechenregeln, insbesondere bei Punkt 4. Vorsicht geboten! Es gilt i.A.  $(A \cdot B)^T \neq A^T \cdot B^T$ ! Der Ausdruck rechts ist für  $m \neq n$  nicht einmal definiert, aber selbst für  $m = n$  gilt i.A. nicht die Gleichheit.

Wir werden nun untersuchen, wie der Rang der Matrix  $C = A \cdot B$  von den Rängen der Matrizen  $A$  und  $B$  abhängt.

**Lemma 3.38.** Ist  $A \in M(m \times n, K)$  und  $B \in M(n \times r, K)$ , so gilt

$$\text{rang } A + \text{rang } B - n \leq \text{rang } (A \cdot B) \leq \min\{\text{rang } A, \text{rang } B\}.$$

*Beweis.* Wenn wir die Matrizen als lineare Abbildungen auffassen, erhalten wir folgendes Diagramm. Wir definieren nun  $F' := F_A|_{\text{Im } B}$ . Dann gilt

$$\text{Im } F' = \text{Im } (F_{A \cdot B}) \text{ und } \text{Ker } F' = \text{Ker } F_A \cap \text{Im } F_B.$$

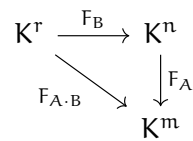


Abbildung 3.4: Illustration der Komposition von A und B.

Wenden wir die Dimensionsformel aus Satz 3.18 an, so erhalten wir

$$\text{rang}(A \cdot B) = \text{rang } F' = \dim(\text{Im } F_B) - \dim(\text{Ker } F') = \text{rang } B - \dim(\text{Ker } F'). \quad (3.7)$$

Somit folgt sofort  $\text{rang}(A \cdot B) \leq \text{rang } B$ . Weiterhin gilt  $\text{Im } F_{A \cdot B} \subset \text{Im } F_A$ . Dazu sei  $z \in \text{Im } F_{A \cdot B}$ , d.h. es existiert  $x \in K^r$  mit  $(A \cdot B) \cdot x = (\sum_{j=1}^n a_{ij} y_j(x))_{i \in \{1, \dots, m\}} = z$ , wobei  $y_j(x) = b_j \cdot x$ . Wähle also  $\tilde{x}_j := y_j(x)$ ,  $j = 1, \dots, n$  und wir erhalten mit  $A \cdot \tilde{x} = z$ , dass  $z \in \text{Im } F_A$ . Es folgt  $\text{rang}(A \cdot B) \leq \text{rang } A$  und insgesamt  $\text{rang}(A \cdot B) \leq \min\{\text{rang } A, \text{rang } B\}$ .

Mit  $\text{Ker } F' \subset \text{Ker } F_A$  folgt aus (3.7)

$$\text{rang}(A \cdot B) \geq \text{rang } B - \dim(\text{Ker } F_A) = \text{rang } B + \text{rang } A - n.$$

Hier folgt die letzte Gleichung aus der Dimensionsformel für die lineare Abbildung  $F_A$ . Dies ergibt  $\text{rang } A + \text{rang } B - n \leq \text{rang}(A \cdot B)$ . Diese Abschätzung ist straff, falls  $\text{Ker } F' = \text{Ker } A$ .  $\square$

Wir wenden uns nun den kanonischen Isomorphismen  $\text{Hom}(K^n, K^n) \rightarrow M(n \times n, K)$  für den Standardraum  $K^n$  zu. Zur Erinnerung: sei  $\mathcal{B} = (v_1, \dots, v_n)$  eine Basis von  $K^n$ . Dann gibt es genau einen Isomorphismus  $\Phi_{\mathcal{B}} : K^n \rightarrow K^n$  mit  $\Phi_{\mathcal{B}}(e_i) = v_i$ ,  $i = 1, \dots, n$  und  $\Phi_{\mathcal{B}}$  kann durch die quadratische Matrix  $(\Phi_{\mathcal{B}}(e_i)_{i=1, \dots, n})$  beschrieben werden. Ein Isomorphismus  $F : K^n \rightarrow K^n$  besitzt eine lineare Umkehrabbildung  $F^{-1}$  mit

$$F \circ F^{-1} = F^{-1} \circ F = \text{id}_{K^n},$$

wobei für eine allgemeine Menge die Abbildung  $\text{id}_V$  die **Identitätsabbildung** auf  $V$  definiert, d.h.,

$$\text{id}_V : V \rightarrow V, v \mapsto v.$$

Wenn wir diesen Zusammenhang auf die Multiplikation von Matrizen anwenden, erhalten wir:

**Definition 3.39.** Eine Matrix  $A \in M(n \times n, K)$  heißt **invertierbar** oder auch **regulär**, wenn es ein  $A^{-1} \in M(n \times n, K)$  gibt mit der Eigenschaft

$$A \cdot A^{-1} = A^{-1} \cdot A = E_n.$$

Wir sehen sofort, dass diese Definition genau das Richtige tut: Sei  $F_A$  bzw.  $F_{A^{-1}}$  die durch

$A$  und  $A^{-1}$  definierte lineare Matrix-Abbildung. Dann erhalten wir

$$F_A(F_{A^{-1}}(x)) = A \cdot (A^{-1} \cdot x) = (A \cdot A^{-1}) \cdot x = E_n \cdot x = x$$

also gilt

$$F_A \circ F_{A^{-1}} = \text{id}_{K^n}.$$

Wir werden auch nachweisen, dass die suggestive Notation  $A^{-1}$  auch **erlaubt** ist, denn die Menge der invertierbaren Matrizen  $A \in M(n \times n, K)$  bilden eine Gruppe bzgl. der Matrizenmultiplikation und somit existiert für jede Matrix  $A$  ein eindeutiges Inverses, welches wir mit  $A^{-1}$  bezeichnen haben.

**Satz 3.40.** Die Menge

$$GL(n, K) := \{A \in M(n \times n, K) \mid A \text{ invertierbar}\}.$$

bildet mit der Verknüpfung der Matrizenmultiplikation eine Gruppe mit neutralem Element  $E_n$ . Diese Gruppe heißt auch **allgemeine lineare Gruppe**.

*Beweis.* Als Erstes zeigen wir, dass die Matrizenmultiplikation auf  $GL(n, K) \times GL(n, K)$  abgeschlossen ist, d.h., für alle  $A, B \in GL(n, K)$  muss auch  $A \cdot B \in GL(n, K)$  gelten. Dazu seien  $A^{-1}$  und  $B^{-1}$  mit der Eigenschaft

$$A \cdot A^{-1} = A^{-1} \cdot A = B \cdot B^{-1} = B^{-1} \cdot B = E_n.$$

Dann folgt aus Lemma 3.35.3.:

$$(B^{-1} \cdot A^{-1}) \cdot (A \cdot B) = E_n = (A \cdot B) \cdot (B^{-1} \cdot A^{-1}).$$

Also ist  $A \cdot B$  invertierbar und somit in  $GL(n, K)$ . Wir müssen nun noch die Gruppenaxiome G1 und G2 nachweisen. G1 folgt direkt aus Lemma 3.35.3.  $E_n$  ist das neutrale Element und zu jedem  $A \in GL(n, K)$  gibt es per Definition ein Inverses. Wie in Kapitel 2.3.1 gezeigt, ist das Inverse eindeutig und somit war es zulässig, die Schreibweise  $A^{-1}$  zu verwenden. Insbesondere gilt  $(A^{-1})^{-1} = A$  für alle  $A \in GL(n, K)$ .  $\square$

**Satz 3.41.** Für eine Matrix  $A \in M(n \times n, K)$  sind folgende Aussagen äquivalent:

1.  $A$  ist invertierbar.
2.  $A^\top$  ist invertierbar.
3. Spaltenrang von  $A$  ist gleich  $n$ .
4. Zeilenrang von  $A$  ist gleich  $n$ .

Des Weiteren gilt  $(A^\top)^{-1} = (A^{-1})^\top$ .

*Beweis.* 1.  $\Rightarrow$  2. : Wir rechnen nach:

$$(A^{-1})^T \cdot A^T = (A \cdot A^{-1})^T = E_n^T = E_n.$$

2.  $\Rightarrow$  1. :

$$E_n = (A^T)^{-1} \cdot A^T = (A \cdot ((A^T)^{-1})^T)^T \Rightarrow A \cdot ((A^T)^{-1})^T = E_n^T = E_n.$$

Somit gibt es ein zu  $A$  Inverses  $((A^T)^{-1})^T$  und mit der Eindeutigkeit des Inversen folgt ausserdem

$$((A^T)^{-1})^T = A^{-1} \Leftrightarrow (A^T)^{-1} = (A^{-1})^T.$$

1.  $\Rightarrow$  3. : Die Beziehung  $A \cdot A^{-1} = E_n$  impliziert, dass  $e^j \in \text{Im}(F_A)$ ,  $j = 1, \dots, n$  und somit  $\text{rang}(A) = n$  gelten muss.

1.  $\Leftarrow$  3. : Aus  $\text{Im}(F_A) = K^n$  folgt, dass es  $A^{-1}$  geben muss mit  $A \cdot A^{-1} = E_n$ .

2.  $\Leftrightarrow$  4. folgt aus obigem Argument nach Transposition.  $\square$

Wir leiten nun einen Algorithmus her, um für  $A \in \text{GL}(n, K)$  die Inverse  $A^{-1}$  zu berechnen. Mittels der Gleichung  $A \cdot A^{-1} = E_n$  können wir die  $n$  Spalten von  $A^{-1}$  als Lösungen der folgenden LGS auffassen:

$$A \cdot x^j = e^j, j = 1, \dots, n, \quad (3.8)$$

wobei  $x^j := \begin{pmatrix} x_{1j} \\ \vdots \\ x_{nj} \end{pmatrix}.$

**Bemerkung 3.42.** Seien  $z^j, j = 1, \dots, n$  Lösungen des LGS (3.8). Dann gilt  $A^{-1} = (z^j)_{j=1, \dots, n}$ .

Diese Einsicht nutzen wir im Folgenden, um den Gauss-Algorithmus zur Lösung von LGS auf unseren Fall anzupassen. Ein wichtiges Hilfsmittel hierzu sind die sogenannten **Elementarmatrizen**.

**Definition 3.43 (Elementarmatrizen).** Eine Matrix in  $M(n \times n, K)$  heisst **Elementarmatrix**, wenn sie von einem der drei folgenden Typen ist:

1. Typ 1: Sei  $i, j \in \{1, \dots, n\}$ . Definiere

$$E_{i \leftrightarrow j} := E_n - E_i^i - E_j^j + E_i^j + E_j^i,$$

wobei  $E_i^j \in M(n \times n, K)$  mit  $(e_i^j)_{lk} = 0$  für  $(l, k) \neq (i, j)$  und  $(e_i^j)_{lk} = 1$  für  $(l, k) = (i, j)$ . In diesem Fall entsteht  $E_{i \leftrightarrow j}$  aus  $E_n$  durch vertauschen der  $i$ -ten und  $j$ -ten Zeile.

2. Typ 2: Sei  $i \in \{1, \dots, n\}$  und  $\lambda \in K^*$ . Definiere

$$E_i(\lambda) := E_n + (\lambda - 1)E_i^i.$$

In diesem Fall entsteht  $E_i(\lambda)$  aus  $E_n$  durch Multiplikation der  $i$ -ten Zeile mit  $\lambda$ .

3. Typ 3: Sei  $i, j \in \{1, \dots, n\}$  mit  $i \neq j$  und sei  $\lambda \in K$ . Definiere

$$E_i^j(\lambda) := E_n + \lambda E_i^j.$$

$E_i^j(\lambda)$  entsteht aus  $E_n$  durch Addition des  $\lambda$ -fachen der  $j$ -ten Zeile zur  $i$ -ten Zeile.

Wir halten folgende Bemerkung fest:

**Bemerkung 3.44.** 1. Eine Elementarmatrix ist regulär. Es gilt weiterhin:

- (a)  $(E_{i \leftrightarrow j})^{-1} = (E_{j \leftrightarrow i})^{-1} = E_{i \leftrightarrow j} = E_{j \leftrightarrow i}$ .
  - (b)  $(E_i(\lambda))^{-1} = E_i(\frac{1}{\lambda})$ .
  - (c)  $(E_i^j(\lambda))^{-1} = E_i^j(-\lambda)$
2. Sei  $A \in M(n \times n, K)$  eine Elementarmatrix,  $B \in M(n \times r, K)$  und  $B' := A \cdot B$ . Dann gilt:
- (a) Ist  $A = E_{i \leftrightarrow j}$ , so entsteht  $B'$  aus  $B$  durch Vertauschen der  $i$ -ten und  $j$ -ten Zeile.
  - (b) Ist  $A = E_i(\lambda)$ , so entsteht  $B'$  aus  $B$  durch Multiplikation der  $i$ -ten Zeile mit  $\lambda$ .
  - (c) Ist  $A = E_i^j(\lambda)$ , so entsteht  $B'$  aus  $B$  durch Addition des  $\lambda$ -fachen der  $j$ -ten Zeile zur  $i$ -ten Zeile.
3. Sei  $A \in M(n \times n, K)$  eine Elementarmatrix,  $B \in M(r \times n, K)$  und  $B' := B \cdot A$ . Dann gilt:
- (a) Ist  $A = E_{i \leftrightarrow j}$ , so entsteht  $B'$  aus  $B$  durch Vertauschen der  $i$ -ten und  $j$ -ten Spalte.
  - (b) Ist  $A = E_i(\lambda)$ , so entsteht  $B'$  aus  $B$  durch Multiplikation der  $i$ -ten Spalte mit  $\lambda$ .
  - (c) Ist  $A = E_i^j(\lambda)$ , so entsteht  $B'$  aus  $B$  durch Addition des  $\lambda$ -fachen der  $i$ -ten Spalte zur  $j$ -ten Spalte.

**Satz 3.45.** Eine Matrix  $A \in M(n \times n, K)$  ist genau dann regulär, wenn sich  $A$  durch (endlich viele) elementare Zeilenumformungen in die Einheitsmatrix überführen lässt. Ist  $A$  regulär, so erhalten wir  $A^{-1}$ , indem wir dieselben elementaren Zeilenumformungen auf  $E_n$  anwenden.

*Beweis.*  $\Leftarrow$ : Wenn  $B$  aus  $A$  durch elementare Zeilenumformungen hervorgeht, so ändert sich der Rang nicht und es gilt  $\text{rang}(A) = \text{rang}(B)$ . Wenn sich  $B$  nun durch elementare Zeilenumformungen in  $E_n$  überführen lässt, so gilt  $\text{rang}(A) = \text{rang}(E_n) = n$ , also folgt mit Satz 3.41, dass  $A$  regulär ist.

$\Rightarrow$ : Ist  $A \in M(n \times n, K)$  regulär, so gilt mit Satz 3.41, dass  $\text{rang}(A) = n$ . Anwendung des Gauss-Algorithmus führt auf eine Zeilenstufenform mit  $n$  Pivotzeilen so dass gilt  $a'_{ii} \neq 0, i = 1, \dots, n$ :

$$A' = \begin{pmatrix} a'_{11} & * & \dots & * \\ 0 & a'_{22} & * & * \\ 0 & & \ddots & \vdots \\ 0 & & \dots & 0 & a'_{nn} \end{pmatrix} \in M(n \times n, K).$$

Jetzt können wir noch ein geeignetes Vielfaches der 2. bis zur  $n$ -ten Zeile zur 1. Zeile addieren, um in der 1. Zeile den Vektor  $(a'_{11}, 0, \dots, 0)$  zu erhalten. Gleichmaßen können wir mit den Zeilen 2 bis  $n$  umgehen. Insgesamt erhalten wir so die Diagonalmatrix

$$A'' = \begin{pmatrix} a'_{11} & 0 & \dots & 0 \\ 0 & a'_{22} & 0 & \dots & 0 \\ 0 & & \ddots & & \vdots \\ 0 & & \dots & 0 & a'_{nn} \end{pmatrix} \in M(n \times n, K).$$

Nun multiplizieren wir die jeweilige Zeile mit  $1/a'_{ii}$  und erhalten  $E_n$ .

Der obige Beweis zeigt nun, dass es Elementarmatrizen  $B_1, \dots, B_N \in M(n \times n, K)$  gibt mit

$$B_N \cdots B_1 \cdot A = E_n.$$

Somit gilt

$$B_N \cdots B_1 = B_N \cdots B_1 \cdot E_n = A^{-1}.$$

□

Wir geben nun ein Beispiel an.

**Beispiel 3.46.** Wir möchten die Matrix  $A$  invertieren, oder entscheiden, dass  $A$  nicht vollen Rang hat, also nicht regulär ist.

$$A = \begin{pmatrix} 1 & -1 & 0 \\ 3 & -3 & 1 \\ 1 & -2 & 2 \end{pmatrix}$$

Wir schreiben  $A$  und die Einheitsmatrix  $E_n$  nebeneinander.

$$\left[ \begin{array}{ccc|ccc} 1 & -1 & 0 & 1 & 0 & 0 \\ 3 & -3 & 1 & 0 & 1 & 0 \\ 1 & -2 & 2 & 0 & 0 & 1 \end{array} \right] \quad (3.9)$$

Dann führen wir mittels elementarer ZUF  $A$  auf die Einheitsmatrix zurück (bzw. erkennen, dass eine Nullzeile entsteht). Dies entspricht der Operation  $B_N \cdots B_1 \cdot A$ . Wenn wir diese elementaren ZUFs simultan auf  $E_n$  anwenden, so erhalten wir die gewünschte Inverse  $B_N \cdots B_1 \cdot E_n = A^{-1}$ .

$$\begin{array}{ccc} \left[ \begin{array}{ccc|ccc} 1 & -1 & 0 & 1 & 0 & 0 \\ 3 & -3 & 1 & 0 & 1 & 0 \\ 1 & -2 & 2 & 0 & 0 & 1 \end{array} \right] & \xrightarrow{\text{II.}-3\cdot\text{I, III.}-\text{I.}} & \left[ \begin{array}{ccc|ccc} 1 & -1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & -3 & 1 & 0 \\ 0 & -1 & 2 & -1 & 0 & 1 \end{array} \right] \\ \xrightarrow{\text{II.} \leftrightarrow \text{III.}} & & \left[ \begin{array}{ccc|ccc} 1 & -1 & 0 & 1 & 0 & 0 \\ 0 & -1 & 2 & -1 & 0 & 1 \\ 0 & 0 & 1 & -3 & 1 & 0 \end{array} \right] & \xrightarrow{\text{I.}-\text{II.}} & \left[ \begin{array}{ccc|ccc} 1 & 0 & -2 & 2 & 0 & -1 \\ 0 & -1 & 2 & -1 & 0 & 1 \\ 0 & 0 & 1 & -3 & 1 & 0 \end{array} \right] \\ \xrightarrow{\text{I.}+2\cdot\text{III, II.}-2\cdot\text{III.}} & & \left[ \begin{array}{ccc|ccc} 1 & 0 & 0 & -4 & 2 & -1 \\ 0 & -1 & 0 & 5 & -2 & 1 \\ 0 & 0 & 1 & -3 & 1 & 0 \end{array} \right] & \xrightarrow{(-1)\cdot\text{II.}} & \left[ \begin{array}{ccc|ccc} 1 & 0 & 0 & -4 & 2 & -1 \\ 0 & 1 & 0 & -5 & 2 & -1 \\ 0 & 0 & 1 & -3 & 1 & 0 \end{array} \right] \end{array}$$

Somit ist  $A$  invertierbar und es gilt

$$A^{-1} = \begin{pmatrix} -4 & 2 & -1 \\ -5 & 2 & -1 \\ -3 & 1 & 0 \end{pmatrix}.$$

**Lemma 3.47.** Sei  $A \in M(m \times n, K)$ . Dann gilt für beliebige  $B \in GL(m, K), C \in GL(n, K)$ :

$$\text{rang}(A) = \text{rang}(B \cdot A \cdot C).$$

*Beweis.* Übung. □

### 3.7 Koordinatentransformation bei Basiswechsel

Wir kommen nun auf den Begriff des **Koordinatensystems** zurück. Sei  $V$  ein  $K$ -Vektorraum mit Basis  $\mathcal{B} = (v_1, \dots, v_n)$ . Wir erhalten gemäß Korollar 3.26 einen eindeutigen Isomorphismus

$$\Phi_{\mathcal{B}} : K^n \rightarrow V \text{ mit } \Phi_{\mathcal{B}}(e_j) = v_j, j = 1, \dots, n,$$

wobei  $e_j, j = 1, \dots, n$  wieder die kanonischen Basisvektoren des  $K^n$  bezeichnen. Wir erhalten

$$\Phi_{\mathcal{B}}((x_1, \dots, x_n)^T) = x_1 v_1 + \dots + x_n v_n.$$

$\Phi_{\mathcal{B}}$  heißt das durch  $\mathcal{B}$  bestimmte **Koordinatensystem** von  $V$ . Für  $v = \Phi_{\mathcal{B}}((x_1, \dots, x_n)^T) = x_1 v_1 + \dots + x_n v_n \in V$  heißt der Vektor  $(x_1, \dots, x_n)^T$  die **Koordinaten von  $v$  (bzgl.  $\mathcal{B}$ )** und es gilt

$$x = (x_1, \dots, x_n)^T = \Phi_{\mathcal{B}}^{-1}(v).$$

Manchmal ist es wichtig verschiedene Koordinaten (also Koordinaten von  $v$  bezüglich verschiedener Basen) ineinander umzuwandeln.

**Definition 3.48 (Transformationsmatrix).** Seien zwei Basen  $\mathcal{A} = (v_1, \dots, v_n)$  und  $\mathcal{B} = (w_1, \dots, w_n)$  eines  $K$ -Vektorraums  $V$  gegeben. Die **Transformationsmatrix** des Basiswechsels von  $\mathcal{A}$  nach  $\mathcal{B}$  ist gegeben durch die Matrix  $T_{\mathcal{B}}^{\mathcal{A}} \in M(n \times n, K)$ , so dass für alle

Spaltenindizes  $1 \leq j \leq n$  die Spalte  $\begin{pmatrix} t_{1j} \\ \vdots \\ t_{nj} \end{pmatrix}$  die Koordinaten der Darstellung von  $v_j$  in

der Basis  $\mathcal{B}$  ist:

$$v_j = \sum_{i=1}^n t_{ij} w_i = t_{1j} w_1 + \dots + t_{nj} w_n = \Phi_{\mathcal{B}}((t_{1j}, \dots, t_{nj})^T).$$

**Satz 3.49.** Seien zwei Basen  $\mathcal{A} = (v_1, \dots, v_n)$  und  $\mathcal{B} = (w_1, \dots, w_n)$  eines  $K$ -Vektorraums  $V$  sowie die zugehörige Transformationsmatrix  $T_{\mathcal{B}}^{\mathcal{A}}$  des Basiswechsels von  $\mathcal{A}$  nach  $\mathcal{B}$  gegeben. Für  $v \in V$  sei  $v = x_1 v_1 + \dots + x_n v_n = y_1 w_1 + \dots + y_n w_n$ . Dann gilt:

1.  $y = T_{\mathcal{B}}^{\mathcal{A}} \cdot x$ .
2.  $T_{\mathcal{B}}^{\mathcal{A}} \in \text{GL}(n, K)$ .
3.  $x = (T_{\mathcal{B}}^{\mathcal{A}})^{-1} \cdot y$ .

*Beweis.* Zu 1.:

$$\sum_{i=1}^n y_i w_i = \sum_{i=1}^n x_i v_i = x_1 \left( \sum_{i=1}^n t_{i1} w_i \right) + \dots + x_n \left( \sum_{i=1}^n t_{in} w_i \right)$$

$$\begin{aligned}
&= \sum_{j=1}^n (x_1 t_{j1} + \cdots + x_n t_{jn}) w_j \\
&= \sum_{j=1}^n (t_{j1} x_1 + \cdots + t_{jn} x_n) w_j.
\end{aligned}$$

Da  $\mathcal{B}$  eine Basis ist, ist die Linearkombination von  $v$  mit Vektoren aus  $\mathcal{B}$  eindeutig und wir erhalten für alle  $i = 1, \dots, n$ :

$$y_i = t_{i1} x_1 + \cdots + t_{in} x_n.$$

Zu 2.: Für  $0 \in V$  erhalten wir die eindeutige Darstellung

$$0 = 0 \cdot v_1 + \cdots + 0 \cdot v_n = 0 \cdot w_1 + \cdots + 0 \cdot w_n.$$

Somit gilt  $y = 0 \Leftrightarrow x = 0$ . Fassen wir  $y = T_{\mathcal{B}}^A \cdot x$  als LGS in der Variablen  $x$  auf, so folgt mit  $y = 0$ , dass es nur die triviale Lösung  $x = 0$  gibt, und somit gilt  $\text{rang}(T_{\mathcal{B}}^A) = n$  und mit Satz 3.41 folgt  $T_{\mathcal{B}}^A \in \text{GL}(n, K)$ .

Zu 3.: trivial. □

Wir können die Aussage von Satz 3.49 das kommutative Diagramm von Isomorphismen in Abb. 3.5 veranschaulichen:

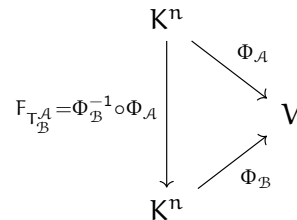


Abbildung 3.5: Kommutatives Diagramm von Isomorphismen.

**Bemerkung 3.50.** Seien  $\mathcal{A}, \mathcal{B}, \mathcal{C}$  Basen von  $V$  der Länge  $n$ . Dann gilt:

1.  $T_{\mathcal{A}}^{\mathcal{A}} = E_n$ .
2.  $T_{\mathcal{C}}^{\mathcal{A}} = T_{\mathcal{C}}^{\mathcal{B}} \cdot T_{\mathcal{B}}^{\mathcal{A}}$ .
3.  $T_{\mathcal{A}}^{\mathcal{B}} = (T_{\mathcal{B}}^{\mathcal{A}})^{-1}$ .

*Beweis.* Übungen. □

**Beispiel 3.51.** Wir wenden den obigen Satz nun auf die Situation  $V = K^n$  an. In diesem Fall sind  $\Phi_{\mathcal{A}}$  und  $\Phi_{\mathcal{B}}$  durch Matrixabbildungen mit den Matrizen  $A$  und  $B$  gegeben, wobei die Spaltenvektoren durch die jeweiligen Basisvektoren in  $\mathcal{A}$  und  $\mathcal{B}$  gegeben sind. Falls  $\mathcal{A}$  die kanonische Basis ist (also  $v_i = e_i, i = 1, \dots, n$ ), so folgt sogar  $T = B^{-1} \cdot E = B^{-1}$ . Wir erhalten folgendes Diagramm in Abb. 3.6:

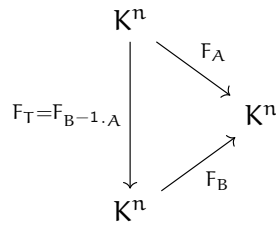


Abbildung 3.6: Kommutatives Diagramm von Isomorphismen zwischen Standardräumen  $\mathbb{K}^n$ .

**Beispiel 3.52.** Wir werden noch konkreter und betrachten im  $\mathbb{R}^2$  die kanonische Basis  $\mathcal{A} = (e_1, e_2)$  und die Basis  $\mathcal{B} = (w_1, w_2) = ((2, 1)^\top, (1, 3)^\top)$ . Unser Ziel ist es, den Vektor  $v = -e_1 - e_2 = (-1) \cdot e_1 + (-1) \cdot e_2$  bezüglich der Basis  $\mathcal{B}$  darzustellen. Wir haben  $\Phi_{\mathcal{A}}((-1, -1)^\top) = v$ . Gesucht ist  $y \in \mathbb{R}^2$  mit  $\Phi_{\mathcal{B}}((y_1, y_2)^\top) = v$ . Wir erhalten

$$\Phi_{\mathcal{A}}(x) = E_2 \cdot x, \quad \Phi_{\mathcal{B}}(y) = B \cdot y, \quad \text{mit } B = \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}.$$

Um die Transformationsmatrix  $T_{\mathcal{B}}^{\mathcal{A}} = B^{-1}$  zu erhalten, müssen wir also die Inverse  $B^{-1}$  berechnen:

$$B^{-1} = \begin{pmatrix} 3/5 & -1/5 \\ -1/5 & 2/5 \end{pmatrix}.$$

Somit folgt

$$y = B^{-1}x \Leftrightarrow \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 3/5 & -1/5 \\ -1/5 & 2/5 \end{pmatrix} \cdot \begin{pmatrix} -1 \\ -1 \end{pmatrix} = \begin{pmatrix} -2/5 \\ -1/5 \end{pmatrix}$$

und wir erhalten

$$v = -\frac{2}{5}w_1 - \frac{1}{5}w_2 = -\frac{2}{5} \begin{pmatrix} 2 \\ 1 \end{pmatrix} - \frac{1}{5} \begin{pmatrix} 1 \\ 3 \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \end{pmatrix}.$$

Nun wenden wir uns dem Zusammenhang von Koordinatensystemen und darstellender Matrizen für lineare Abbildungen zu (vgl. Satz 3.27).

**Satz 3.53.** Sei  $F: V \rightarrow W$  eine lineare Abbildung und seien  $\mathcal{A}$  und  $\mathcal{B}$  Basen von  $V$  und  $W$ . Dann gilt:

$$\Phi_{\mathcal{B}} \circ F_{M_{\mathcal{B}}^{\mathcal{A}}(F)} = F \circ \Phi_{\mathcal{A}}, \quad \text{und somit } F_{M_{\mathcal{B}}^{\mathcal{A}}(F)} = \Phi_{\mathcal{B}}^{-1} \circ F \circ \Phi_{\mathcal{A}}. \quad (3.10)$$

Diese Aussage wird illustriert mit folgendem kommutativen Diagramm in Abb. 3.7:

$$\begin{array}{ccc}
 \mathbb{K}^n & \xrightarrow{\Phi_{\mathcal{A}}} & V \\
 F_{M_{\mathcal{B}}^{\mathcal{A}}(F)} \downarrow & & \downarrow F \\
 \mathbb{K}^m & \xrightarrow{\Phi_{\mathcal{B}}} & W
 \end{array}$$

Abbildung 3.7: Kommutatives Diagramm von Abbildungen für darstellende Matrixabbildung  $F_{M_{\mathcal{B}}^{\mathcal{A}}(F)}$ .

*Beweis.* Mit Satz 3.10, reicht es zu zeigen, dass die beiden Abbildungen auf der kanonischen Basis  $(e_1, \dots, e_n)$  übereinstimmen. Sei  $M_{\mathcal{B}}^{\mathcal{A}}(F) = A = (a_{ij})$ . Wir erhalten

$$(\Phi_{\mathcal{B}} \circ F_{M_{\mathcal{B}}^{\mathcal{A}}(F)})(e_j) = \Phi_{\mathcal{B}}(M_{\mathcal{B}}^{\mathcal{A}}(F) \cdot e_j) = \Phi_{\mathcal{B}}((a_{1j}, \dots, a_{mj})^T) = \sum_{i=1}^m a_{ij} w_i,$$

$$\text{sowie } F(\Phi_{\mathcal{A}}(e_j)) = F(v_j) = \sum_{i=1}^m a_{ij} w_i.$$

Die Gleichung  $F_{M_{\mathcal{B}}^{\mathcal{A}}(F)} = \Phi_{\mathcal{B}}^{-1} \circ F \circ \Phi_{\mathcal{A}}$  erhalten wir durch Anwendung von  $\Phi_{\mathcal{B}}^{-1}$  von links auf die obige Gleichung.  $\square$

**Bemerkung 3.54.** Wir sehen nun auch, dass die darstellenden Matrizen  $M_{\mathcal{B}}^{\mathcal{A}}(F)$  eine Verallgemeinerung der Transformationsmatrizen sind, denn für  $V = W$  und  $F = \text{id}_V$  gilt

$$M_{\mathcal{B}}^{\mathcal{A}}(\text{id}_V) = T_{\mathcal{B}}^{\mathcal{A}}.$$

Nun werden wir den obigen Zusammenhang von Koordinatentransformation und darstellender Matrix auf die Komposition von linearen Abbildungen erweitern. Wie sich herausstellen wird, kann die darstellende Matrix der Komposition von linearen Abbildungen als Produkt der jeweiligen darstellenden Matrizen berechnet werden.

**Satz 3.55.** Gegeben seien  $\mathbb{K}$ -Vektorräume  $U, V$  und  $W$  mit  $\dim(U) = r, \dim(V) = n, \dim(W) = m$ , Basen  $\mathcal{A}, \mathcal{B}$  und  $\mathcal{C}$  sowie lineare Abbildungen  $G : U \rightarrow V$  und  $F : V \rightarrow W$ . Dann gilt:

$$M_{\mathcal{C}}^{\mathcal{A}}(F \circ G) = M_{\mathcal{C}}^{\mathcal{B}}(F) \cdot M_{\mathcal{B}}^{\mathcal{A}}(G).$$

*Beweis.* Wir führen den Beweis dadurch, dass wir zeigen, dass das untenstehende Diagramm kommutativ ist, wobei  $A := M_{\mathcal{C}}^{\mathcal{B}}(F), B := M_{\mathcal{B}}^{\mathcal{A}}(G)$ . Diese Behauptung folgt aus der Kommutativität des oberen wie unteren Vierecks mittels Abb.3.7 (siehe Satz 3.53). Das rechte bzw. linke Dreieck kommutiert nach Definition der Komposition von linearen Abbildungen.  $\square$

Nun kommen wir zur **Transformationsformel**, welche einen Zusammenhang zwischen der "neuen" Darstellungsmatrix einer linearen Abbildung  $F$  bei einem Basiswechsel (von  $\mathcal{A}$

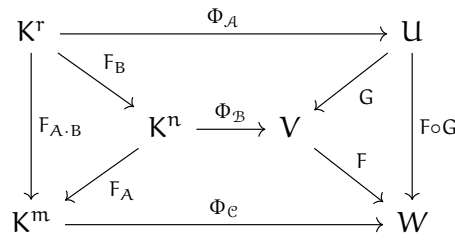


Abbildung 3.8: Kommutatives Diagramm bzgl. Komposition von Abbildungen und ihren darstellenden Matrizen.

nach  $\mathcal{A}'$  von  $V$  und von  $\mathcal{B}$  nach  $\mathcal{B}'$  von  $W$ ) und der "alten" Darstellungsmatrix  $M_{\mathcal{B}}^{\mathcal{A}}(F)$  sowie der Transformationsmatrizen  $T_{\mathcal{A}'}^{\mathcal{A}}, T_{\mathcal{B}'}^{\mathcal{B}}$  herstellt.

**Satz 3.56.** Sei  $F : V \rightarrow W$  eine lineare Abbildung und  $\mathcal{A} = (v_1, \dots, v_n), \mathcal{A}' = (v'_1, \dots, v'_n)$  Basen von  $V$  und  $\mathcal{B} = (w_1, \dots, w_m), \mathcal{B}' = (w'_1, \dots, w'_m)$  Basen von  $W$ . Dann ist das untenstehende Diagramm kommutativ und es gilt

$$M_{\mathcal{B}'}^{\mathcal{A}'}(F) = T_{\mathcal{B}'}^{\mathcal{B}} \cdot M_{\mathcal{B}}^{\mathcal{A}}(F) \cdot (T_{\mathcal{A}'}^{\mathcal{A}})^{-1}.$$

*Beweis.* Wir argumentieren wiederum, dass das folgende Diagramm kommutativ ist. Das

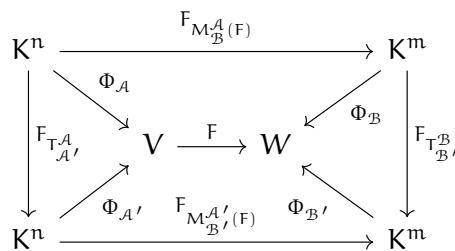


Abbildung 3.9: Kommutatives Diagramm bzgl. Komposition von Abbildungen und ihren darstellenden Matrizen.

obere sowie das untere Viereck kommutiert mit Satz 3.53. Das linke sowie das rechte Dreieck kommutieren mit Satz 3.49.

Wir geben nun neben diesem "Beweis per Diagramm" auch einen weiteren "rechnerischen" Beweis an. Dazu sei  $A := M_{\mathcal{B}}^{\mathcal{A}}(F), A' := M_{\mathcal{B}'}^{\mathcal{A}'}(F), T := T_{\mathcal{B}'}^{\mathcal{B}}, S^{-1} := T_{\mathcal{A}'}^{\mathcal{A}}$ . Wir wollen also die Beziehung

$$A' = T \cdot A \cdot S^{-1}$$

beweisen, wobei  $A = (a_{ij}), A' = (a'_{ij}) \in M(m \times n, K), T = (t_{ij}) \in GL(m, K), S^{-1} = (s_{ij}) \in GL(n, K)$ . Wir erhalten per Definition von  $A$  und  $A'$  für  $i, j \in \{1, \dots, n\}$ :

$$F(v_i) = \sum_{k=1}^m a_{ki} w_k, \quad F(v'_j) = \sum_{\ell=1}^m a'_{\ell j} w'_\ell.$$

Wir rechnen nun die Terme aus:

$$F(v'_j) = F\left(\sum_{i=1}^n s_{ij}v_i\right) = \sum_{i=1}^n s_{ij}F(v_i) = \sum_{i=1}^n s_{ij}\left(\sum_{k=1}^m a_{ki}w_k\right) = \sum_{k=1}^m \left(\sum_{i=1}^n a_{ki}s_{ij}\right)w_k.$$

Des Weiteren gilt  $w_k = \sum_{\ell=1}^m t_{\ell k}w'_\ell$  und somit

$$\begin{aligned} F(v'_j) &= \sum_{k=1}^m \left(\sum_{i=1}^n a_{ki}s_{ij}\right) \left(\sum_{\ell=1}^m t_{\ell k}w'_\ell\right) \\ &= \sum_{\ell=1}^m \left(\sum_{k=1}^m t_{\ell k} \left(\sum_{i=1}^n a_{ki}s_{ij}\right)\right) w'_\ell. \end{aligned}$$

Daraus folgt für  $1 \leq \ell \leq m, 1 \leq j \leq n$ :

$$a'_{\ell j} = \sum_{k=1}^m t_{\ell k} \left(\sum_{i=1}^n a_{ki}s_{ij}\right),$$

also  $A' = T \cdot A \cdot S^{-1}$ . □

**Bemerkung 3.57.** Für den Spezialfall  $V = W, \mathcal{A} = \mathcal{B}, \mathcal{A}' = \mathcal{B}'$  folgt  $A' = S \cdot A \cdot S^{-1}$  mit  $S = T_{\mathcal{A}'}^{\mathcal{A}}$ .

Die Beziehung  $A' = T \cdot A \cdot S^{-1}$  setzt die beiden Matrizen  $A$  und  $A'$  in ein spezielles Verhältnis, welches wir uns nun genauer anschauen werden.

**Definition 3.58.** Seien  $A, A' \in M(m \times n, K)$ .  $A$  und  $A'$  heißen **äquivalent** (in Zeichen  $A \sim A'$ ), falls es Matrizen  $S \in GL(m, K)$  und  $T \in GL(n, K)$  gibt mit

$$A' = S \cdot A \cdot T^{-1}.$$

Im Spezialfall  $n = m$  nennen wir  $A$  und  $A'$  **ähnlich** (in Zeichen  $A \approx A'$ ), falls es  $S \in GL(n, K)$  gibt mit

$$A' = S \cdot A \cdot S^{-1}.$$

Wir halten folgende Eigenschaften fest.

- Lemma 3.59.**
1.  $\sim$  ist eine Äquivalenzrelation auf  $M(m \times n, K)$ .
  2.  $\approx$  ist eine Äquivalenzrelation auf  $M(n \times n, K)$ .
  3. Sind  $A, A' \in M(n \times n, K)$  dann gilt:  $A \approx A' \Rightarrow A \sim A'$ .

*Beweis.* Übung. □

**Satz 3.60.** 1. Seien  $A, A' \in M(m \times n, K)$  und seien  $V$  und  $W$   $K$ -Vektorräume der Dimension  $n$  und  $m$ . Dann sind äquivalent:

- (a)  $A \sim A'$ .  
 (b) Es existieren  $F \in \text{Hom}(V, W)$ , Basen  $\mathcal{A} = (v_1, \dots, v_n)$ ,  $\mathcal{A}' = (v'_1, \dots, v'_n)$  von  $V$  sowie Basen  $\mathcal{B} = (w_1, \dots, w_m)$ ,  $\mathcal{B}' = (w'_1, \dots, w'_m)$  von  $W$ , so dass  $A = M_{\mathcal{B}}^{\mathcal{A}}(F)$ ,  $A' = M_{\mathcal{B}'}^{\mathcal{A}'}(F)$ .

2. Sei nun  $n = m$ . Dann sind äquivalent:

- (a)  $A \approx A'$ .  
 (b) Es existieren  $F \in \text{End}(V)$ , Basen  $\mathcal{A} = (v_1, \dots, v_n)$ ,  $\mathcal{A}' = (v'_1, \dots, v'_n)$  von  $V$ , so dass  $A = M_{\mathcal{A}}^{\mathcal{A}}(F)$ ,  $A' = M_{\mathcal{A}'}^{\mathcal{A}'}(F)$ .

*Beweis.* 1b.  $\Rightarrow$  1a. sowie 2b.  $\Rightarrow$  2a. folgen direkt aus Satz 3.56. Für die umgekehrte Richtung wählen wir zunächst eine Basis  $\mathcal{A}$  von  $V$  und eine Basis  $\mathcal{B}$  von  $W$ . Wir definieren dann die gesuchte lineare Abbildung wie im Beweis von Satz 3.56:

$$F(v_j) := \sum_{i=1}^m a_{ij} w_i \text{ für alle } j = 1, \dots, n.$$

Für beliebige  $v \in V$  mit Koordinatendarstellung  $\Phi_{\mathcal{A}}^{-1}(v) = (b_1, \dots, b_n)$  setzen wir

$$F(v) = F\left(\sum_{j=1}^n b_j v_j\right) := \sum_{j=1}^n b_j F(v_j) = \sum_{i=1}^m \left(\sum_{j=1}^n a_{ij} b_j\right) w_i.$$

Dann nutzen wir die Transformationsmatrizen  $S$  und  $T = (t_{ij})$ , um  $\mathcal{A}'$  und  $\mathcal{B}'$  mittels

$$v'_j := \sum_{i=1}^n s_{ij} v_i, j = 1, \dots, n \text{ sowie } w'_i := \sum_{j=1}^m t_{ij} w_j, i = 1, \dots, m$$

zu berechnen. Somit erhalten wir die gewünschte Eigenschaft  $A = M_{\mathcal{B}}^{\mathcal{A}}(F)$ ,  $A' = M_{\mathcal{B}'}^{\mathcal{A}'}(F)$ .  $\square$

**Satz 3.61.** 1. Sei  $A \in M(m \times n, K)$ . Dann ist  $A$  äquivalent zu der Matrix der Gestalt (wir sagen auch **Normalform**)

$$N := \begin{pmatrix} E_r & 0 \\ 0 & 0 \end{pmatrix}, \text{ wobei } r = \text{rang}(A).$$

2. Seien  $A, A' \in M(m \times n, K)$ . Dann gilt  $A \sim A' \Leftrightarrow \text{rang}(A) = \text{rang}(A')$ .

*Beweis.* Zu 1: Seien  $\mathcal{A} = (v_1, \dots, v_n)$  und  $\mathcal{B} = (w_1, \dots, w_m)$  Basen von  $V$  bzw.  $W$ . Wie im Beweis von Satz 3.60 konstruiert, gibt es eine lineare Abbildung  $F \in \text{Hom}(V, W)$ , so

94 | Kapitel 3. Lineare Abbildungen

dass  $A = M_{\mathcal{B}}^{\mathcal{A}}(F)$  gilt. Mit Korollar 3.29 erhalten wir Basen  $\mathcal{A}'$  und  $\mathcal{B}'$ , so dass  $N = M_{\mathcal{B}'}^{\mathcal{A}'}(F)$  die gewünschte Normalform hat. Nach Satz 3.60 gilt  $A \sim A'$ .

Zu 2:  $\Rightarrow$ :  $A \sim A'$  bedeutet per Definition  $A' = S \cdot A \cdot T^{-1}$  für  $S \in GL(m, K)$  und  $T \in GL(n, K)$ . Mit Lemma 3.47 folgt  $\text{rang}(A) = \text{rang}(A')$ .

$\Leftarrow$ : Ist  $\text{rang}(A) = \text{rang}(A')$  so folgt mit dem Beweis von 1., dass  $A \sim N$  und  $A' \sim N$ . Da  $\sim$  eine Äquivalenzrelation ist, folgt  $A \sim A'$ .  $\square$

## Kapitel 4

# Determinanten

### 4.1 Definition und Grundlegende Eigenschaften

Wir betrachten im folgenden quadratische Matrizen  $A = (a_{ij}) \in M(n \times n, K)$ , wobei  $a_i, i = 1, \dots, n$  die Zeilen von  $A$  bezeichnen und  $a^j, j = 1, \dots, n$  die Spalten von  $A$ .

**Definition 4.1.** Sei  $K$  ein Körper. Eine Abbildung

$$\det : M(n \times n, K) \rightarrow K, A \mapsto \det(A),$$

heißt **Determinante**, falls  $\det$  die folgenden Eigenschaften erfüllt:

D1:  $\det$  ist **linear in jeder Zeile**. Für die  $i$ -te Zeile bedeutet das

$$\det \begin{pmatrix} \vdots \\ \lambda a_i + \mu a'_i \\ \vdots \end{pmatrix} = \lambda \det \begin{pmatrix} \vdots \\ a_i \\ \vdots \end{pmatrix} + \mu \det \begin{pmatrix} \vdots \\ a'_i \\ \vdots \end{pmatrix},$$

wobei  $a_i$  und  $a'_i$  Zeilenvektoren einer Matrix aus  $M(n \times n, K)$  sind und  $\lambda, \mu \in K$

D2:  $\det$  ist **alternierend**, d.h.  $\det(A) = 0$ , falls  $i, k \in \{1, \dots, n\}$  existieren mit  $i \neq k$  und  $a_i = a_k$ .

D3:  $\det(E_n) = 1$  (Normierung).

Der Wert  $\det(A) \in K$  heißt **Determinante von  $A$** .

Im Folgenden werden wir zunächst einige fundamentale Eigenschaften der Determinante beweisen. Im Anschluss daran werden wir dann die Existenz und Eindeutigkeit der Determinante bzgl. der Axiome D1, D2 und D3 zeigen.

**Bemerkung 4.2.** 1. Für  $A = (\lambda) \in K$  gilt  $\det(\lambda) = \lambda$ .

2. Für  $A \in M(n \times n, K)$  und  $\lambda \in K$  gilt  $\det(\lambda A) = \lambda^n \det(A)$ .

*Beweis.* Für die erste Aussage rechnen wir nach:

$$\det(\lambda) = \det(1 \cdot \lambda) \stackrel{D1}{=} \lambda \det(1) \stackrel{D3}{=} \lambda \cdot 1 = \lambda.$$

Für die zweite Aussage wenden wir n-mal D1 an:

$$\det(\lambda A) = \det \begin{pmatrix} \lambda a_1 \\ \vdots \\ \lambda a_n \end{pmatrix} \stackrel{D1}{=} \lambda \det \begin{pmatrix} a_1 \\ \lambda a_2 \\ \vdots \\ \lambda a_n \end{pmatrix} \stackrel{D1}{=} \dots \stackrel{D1}{=} \lambda^n \det(A).$$

□

- Lemma 4.3.**
1. Typ1: Entsteht  $A'$  aus  $A$  durch Vertauschung der  $i$ -ten und  $k$ -ten Zeile für  $i, k \in \{1, \dots, n\}$ . Dann gilt  $\det(A') = -\det(A)$ .
  2. Typ2: Sei  $i \in \{1, \dots, n\}$  und  $\lambda \in K^*$ .  $A'$  entsteht durch Multiplikation von Zeile  $i$  mit  $\lambda$ . Dann gilt  $\det(A') = \lambda \det(A)$ .
  3. Typ3: Sei  $i, k \in \{1, \dots, n\}$  mit  $i \neq j$  und sei  $\lambda \in K$ .  $A'$  entsteht aus  $A$  durch Addition des  $\lambda$ -fachen der  $k$ -ten Zeile zur  $i$ -ten Zeile. Dann gilt  $\det(A') = \det(A)$ .

*Beweis.* Zu 1.: Sei

$$A = \begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_k \\ \vdots \end{pmatrix} \quad \text{und} \quad A' = \begin{pmatrix} \vdots \\ a_k \\ \vdots \\ a_i \\ \vdots \end{pmatrix}.$$

Aus der Beziehung

$$\det \begin{pmatrix} \vdots \\ a_i \\ \vdots \\ a_i \\ \vdots \end{pmatrix} = \det \begin{pmatrix} \vdots \\ a_k \\ \vdots \\ a_k \\ \vdots \end{pmatrix} \stackrel{D2}{=} 0$$

folgt

$$\begin{aligned} \det(A) + \det(A') &= \det \begin{pmatrix} \vdots \\ \mathbf{a}_i \\ \vdots \\ \mathbf{a}_k \\ \vdots \end{pmatrix} + \det \begin{pmatrix} \vdots \\ \mathbf{a}_k \\ \vdots \\ \mathbf{a}_i \\ \vdots \end{pmatrix} + \det \begin{pmatrix} \vdots \\ \mathbf{a}_i \\ \vdots \\ \mathbf{a}_i \\ \vdots \end{pmatrix} + \det \begin{pmatrix} \vdots \\ \mathbf{a}_k \\ \vdots \\ \mathbf{a}_k \\ \vdots \end{pmatrix} \\ &= \det \begin{pmatrix} \vdots \\ \mathbf{a}_i \\ \vdots \\ \mathbf{a}_i + \mathbf{a}_k \\ \vdots \end{pmatrix} + \det \begin{pmatrix} \vdots \\ \mathbf{a}_k \\ \vdots \\ \mathbf{a}_i + \mathbf{a}_k \\ \vdots \end{pmatrix} = \det \begin{pmatrix} \vdots \\ \mathbf{a}_i + \mathbf{a}_k \\ \vdots \\ \mathbf{a}_i + \mathbf{a}_k \\ \vdots \end{pmatrix} = 0. \end{aligned}$$

2. folgt direkt aus D1.

Zu 3.:

$$A' = \begin{pmatrix} \vdots \\ \mathbf{a}_i + \lambda \mathbf{a}_k \\ \vdots \\ \mathbf{a}_k \\ \vdots \end{pmatrix}$$

erhalten wir

$$\det(A') = \det \begin{pmatrix} \vdots \\ \mathbf{a}_i + \lambda \mathbf{a}_k \\ \vdots \\ \mathbf{a}_k \\ \vdots \end{pmatrix} = \det \begin{pmatrix} \vdots \\ \mathbf{a}_i \\ \vdots \\ \mathbf{a}_k \\ \vdots \end{pmatrix} + \lambda \det \begin{pmatrix} \vdots \\ \mathbf{a}_k \\ \vdots \\ \mathbf{a}_k \\ \vdots \end{pmatrix} = \det \begin{pmatrix} \vdots \\ \mathbf{a}_i \\ \vdots \\ \mathbf{a}_k \\ \vdots \end{pmatrix} = \det(A).$$

□

**Korollar 4.4.** Falls  $A'$  aus  $A$  durch endlich viele Anwendungen von Typ 1,2,3 elementare Zeilenumformungen entsteht, so gilt

$$\det(A) = 0 \Leftrightarrow \det(A') = 0.$$

**Lemma 4.5.** Gilt  $a_i = 0$  für ein  $i \in \{1, \dots, n\}$ , so folgt  $\det(A) = 0$ .

*Beweis.* Für  $n = 1$  gilt die Aussage wegen Bemerkung 4.2. Für  $n \geq 2$  sei  $a_i = 0$  für  $i \in \{1, \dots, n\}$ . Sei  $k \in \{1, \dots, n\}$ ,  $k \neq i$ . Sei  $A'$  die Matrix die durch die Addition der  $k$ -ten Zeile zur  $i$ -ten Zeile von  $A$  entsteht. Wir erhalten

$$\det(A) \stackrel{\text{Lemma 4.3,(3)}}{=} \det(A') \stackrel{D2}{=} 0.$$

□

**Definition 4.6.** Eine Matrix  $A \in M(m \times n, K)$  heißt **obere Dreiecksmatrix**, falls sie folgende Gestalt hat:

$$A = \begin{pmatrix} a_{11} & * & * \\ & \ddots & * \\ 0 & & a_{rr} \end{pmatrix}.$$

**Lemma 4.7.** Sei  $A \in M(m \times n, K)$  eine obere Dreiecksmatrix. Dann gilt

$$\det(A) = a_{11} \cdots a_{nn} = \prod_{j=1}^n a_{jj}.$$

*Beweis.* Wir betrachten zunächst den Fall, dass gilt  $\prod_{j=1}^n a_{jj} \neq 0$ . Elementare Zeilenumformungen vom Typ 3 lassen die Determinante unverändert (siehe Lemma 4.3) und somit gilt

$$\det(A) = \det \begin{pmatrix} a_{11} & & 0 \\ & \ddots & \\ 0 & & a_{rr} \end{pmatrix}.$$

Anwendung von Lemma 4.3.2 ergibt

$$\det \begin{pmatrix} a_{11} & & 0 \\ & \ddots & \\ 0 & & a_{rr} \end{pmatrix} = \prod_{j=1}^n a_{jj} \det(E_n) = \prod_{j=1}^n a_{jj}.$$

Falls nun  $a_{jj} = 0$  für ein  $j \in \{1, \dots, n\}$ , so gilt natürlich  $\prod_{j=1}^n a_{jj} = 0$ . Wir müssen also zeigen, dass  $\det(A) = 0$  gilt. Wir setzen

$$j = \arg \max\{\ell \in \{1, \dots, n\} | a_{\ell\ell} = 0\}.$$

Wir erhalten somit

$$A = \begin{pmatrix} a_{11} & * & * & * & * & * & * & & \\ & \ddots & * & * & * & * & * & * & * \\ & & & a_{j-1,j-1} & * & * & * & * & * \\ \hline & & & & 0 & a_{j,j+1} & * & * & a_{jn} \\ & & & & & a_{j+1,j+1} & * & * & * \\ & & & & & & \ddots & * & * \\ 0 & & & & & & & & a_{rr} \end{pmatrix}.$$

Weil  $a_{kk} \neq 0$  für  $k > j$  können wir  $A$  mittels elementarer Zeilenumformungen vom Typ 3 in folgende Gestalt bringen:

$$A = \begin{pmatrix} a_{11} & * & * & * & * & * & * & & \\ & \ddots & * & * & * & * & * & * & * \\ & & & a_{j-1,j-1} & * & * & * & * & * \\ \hline 0 & \dots & & & 0 & 0 & 0 & \dots & 0 \\ & & & & & a_{j+1,j+1} & * & * & * \\ & & & & & & \ddots & * & * \\ 0 & & & & & & & & a_{rr} \end{pmatrix}.$$

Wir erhalten  $\det(A) = \det(A') = 0$ . □

**Lemma 4.8.** Sei  $A \in M(n \times n, K)$ . Dann sind die folgenden Aussagen äquivalent:

1.  $\det(A) \neq 0$ .
2.  $\text{rang}(A) = n$ .
3.  $A \in \text{GL}(n, K)$ .

*Beweis.* Durch Anwenden von elementarem ZUF können wir  $A$  in eine obere Dreiecksmatrix  $A'$  überführen, wobei sich ggf. nur das Vorzeichen der Determinante ändert. Wir erhalten:

$$\begin{aligned} A \in \text{GL}(n, K) &\Leftrightarrow \text{rang}(A) = n \\ &\Leftrightarrow \text{rang}(A') = n \\ &\Leftrightarrow a'_{jj} \neq 0, j \in \{1, \dots, n\} \end{aligned}$$

$$\Leftrightarrow \det(A') = \prod_{j=1}^n a'_{jj} \neq 0$$

$$\Leftrightarrow \det(A) \neq 0,$$

wobei die letzte Äquivalenz aus Kor.4.4 folgt. □

**Lemma 4.9.** Es gelten folgende Eigenschaften der Elementarmatrizen:

1.  $\det(E_{i \leftrightarrow j}) = -1$ .
2.  $\det(E_i(\lambda)) = \lambda$ .
3.  $\det(E_i^j(\lambda)) = 1$ .

Für beliebige Matrizen  $A \in M(n \times n, K)$  gilt:

1.  $\det(E_{i \leftrightarrow j}A) = -\det(A) = \det(E_{i \leftrightarrow j}) \det(A)$ .
2.  $\det(E_i(\lambda)A) = \lambda \det(A) = \det(E_i(\lambda)) \det(A)$ .
3.  $\det(E_i^j(\lambda)A) = \det(A) = \det(E_i^j(\lambda)) \det(A)$ .

Für jede Elementarmatrix  $E$  gilt  $\det(E) = \det(E^T)$ .

*Beweis.* Übungen. □

**Satz 4.10.** Seien  $A, B \in M(n \times n, K)$ . Dann gilt:

$$\det(AB) = \det(A) \det(B).$$

*Beweis.* Wir führen eine Fallunterscheidung durch. Angenommen  $\text{rang}(A) < n$ . Mit  $\text{Im}(AB) \subset \text{Im}(A)$  folgt  $\text{rang}(AB) \leq \text{rang}(A) < n$  also mit Lemma 4.8  $\det(AB) = \det(A) \det(B) = 0$ . Somit nehmen wir nun  $\text{rang}(A) = n$  an. Mit Satz 3.45 folgt, dass

$$A = C_1 \cdots C_k$$

für Elementarmatrizen  $C_i, i = 1, \dots, k$ . Mit Lemma 4.9 folgt

$$\det(A) = \det(C_1 \cdots C_k) = \det(C_1) \cdots \det(C_k)$$

und somit

$$\det(AB) = \det(C_1 \cdots C_k \cdot B) = \det(C_1) \cdots \det(C_k) \det(B) = \det(A) \det(B).$$

□

**Bemerkung 4.11.** Weil  $K$  ein Körper ist, gilt:  $\det(AB) = \det(BA)$ .

**Satz 4.12.** Seien  $A \in M(n \times n, K)$ . Dann gilt:

$$\det(A^T) = \det(A).$$

*Beweis.* Falls  $\text{rang}(A) = \text{rang}(A^T) < n$ , so erhalten wir  $\det(A^T) = \det(A) = 0$ . Somit nehmen wir nun  $\text{rang}(A) = \text{rang}(A^T) = n$  an. Mit der Darstellung  $A = C_1 \cdots C_k$  für Elementarmatrizen  $C_i, i = 1, \dots, k$  und Lemma 4.9 gilt

$$\begin{aligned} \det(A^T) &= \det(C_k^T \cdots C_1^T) \\ &= \det(C_k^T) \cdots \det(C_1^T) \\ &= \det(C_k) \cdots \det(C_1) \\ &= \det(C_1) \cdots \det(C_k) \\ &= \det(C_1 \cdots C_k) \\ &= \det(A). \end{aligned}$$

□

**Bemerkung 4.13.** Aus der Eigenschaft  $\det(A^T) = \det(A)$  folgt, dass die Axiome D1 und D2 auch für Spalten der Matrix  $A$  gelten müssen.

**Satz 4.14.** Sei  $n = k + \ell$  mit  $k, \ell \in \mathbb{N}$  und  $A \in M(k \times k, K), B \in M(\ell \times \ell, K), C \in M(k \times \ell, K)$ . Dann gilt:

$$\det \begin{pmatrix} A & C \\ 0_{\ell, k} & B \end{pmatrix} = \det(A) \det(B).$$

*Beweis.* Wir wenden das Gauss-Verfahren auf die beiden Matrizen  $(A \ C)$  und  $(0_{\ell, k} \ B)$  an und erhalten  $(A' \ C')$  und  $(0_{\ell, k} \ B')$ , wobei jeweils  $A'$  und  $B'$  obere Dreiecksmatrizen sind. Seien jeweils  $r$  und  $s$  die Anzahl der benutzten Zeilenvertauschungen. Dann gilt mit Lemma 4.3:

$$\det(A) = (-1)^r \det(A') \quad \text{und} \quad \det(B) = (-1)^s \det(B').$$

Für die obere Dreiecksmatrix  $\begin{pmatrix} A' & C' \\ 0_{\ell, k} & B' \end{pmatrix}$  erhalten wir mit Lemma 4.7:

$$\det \begin{pmatrix} A' & C' \\ 0_{\ell, k} & B' \end{pmatrix} = \prod_{i=1}^k a'_{ii} \cdot \prod_{i=1}^{\ell} b'_{ii} = \det(A') \det(B').$$

Insgesamt folgt:

$$\det \begin{pmatrix} A & C \\ 0_{\ell,k} & B \end{pmatrix} = (-1)^{r+s} \det \begin{pmatrix} A' & C' \\ 0_{\ell,k} & B' \end{pmatrix} = (-1)^r \det(A') (-1)^s \det(B') = \det(A) \det(B).$$

□

## 4.2 Existenz und Eindeutigkeit

Wir beginnen mit der Eindeutigkeit der Determinantenfunktion.

**Satz 4.15.** Sei  $n \in \mathbb{N}$ . Dann gibt es **höchstens** eine Funktion  $\det : M(n \times n, K) \rightarrow K$ , welche die Axiome D1, D2 und D3 erfüllt.

*Beweis.* Seien  $\det_1$  und  $\det_2$  zwei Determinantenfunktionen und sei  $A \in M(n \times n, K)$  beliebig. Mit dem Gauss-Verfahren können wir  $A$  in eine obere Dreiecksmatrix  $A'$  verwandeln. Seien dazu  $k$  Zeilenvertauschungen nötig. Mit Lemma 4.7 erhalten wir

$$\det_1(A) = (-1)^k \det_1(A') = (-1)^k \prod_{i=1}^n a'_{ii} = (-1)^k \det_2(A') = \det_2(A).$$

□

Nun kommen wir zur Existenz der Determinantenfunktion.

**Satz 4.16.** Sei  $n \in \mathbb{N}$ . Dann gibt es **mindestens** eine Funktion  $\det : M(n \times n, K) \rightarrow K$ , welche die Axiome D1, D2 und D3 erfüllt.

*Beweis.* Wir geben einen Beweis per Induktion bzgl.  $n \in \mathbb{N}$  unter Ausnutzung der **Laplaceschen Entwicklungsformel**

1. Induktionsanfang  $n = 1$ : Für diesen Fall erfüllt  $\det(\lambda) := \lambda, \lambda \in K$  die Axiome D1, D2 und D3.
2. Induktionsvoraussetzung: Wir nehmen an, dass es für beliebiges aber festes  $n \in \mathbb{N}$  eine Determinantenfunktion  $\det_n : M(n \times n, K) \rightarrow K$  gibt, welche die Axiome D1, D2 und D3 erfüllt.
3. Induktionsschluss  $n \rightarrow n + 1$ : Für eine beliebige aber feste Zahl  $j \in \{1, \dots, n + 1\}$  definieren wir

$$\det_{n+1} : M((n + 1) \times (n + 1), K) \rightarrow K, A \mapsto \det_{n+1}(A) := \sum_{i=1}^{n+1} (-1)^{i+j} a_{ij} \det_n(A_{ij}),$$

wobei  $A_{ij}$  die Matrix in  $M(n \times n, K)$ , die aus  $A$  durch Weglassen der  $i$ -ten Zeile und  $j$ -ten Spalte entsteht.

Wir überprüfen nun die drei Axiome D1, D2 und D3. Zu D3: Wir schreiben  $E_{n+1} = (\delta_{ij})$ , wobei  $\delta_{ij} = \begin{cases} 1, & \text{falls } i = j, \\ 0, & \text{sonst.} \end{cases}$  Wir erhalten:

$$\det_{n+1}(E_{n+1}) = \sum_{i=1}^{n+1} (-1)^{i+j} \delta_{ij} \det_n(E_{ij}) = (-1)^{2j} \det_n(E_{jj}) = 1,$$

wobei  $E_{ij}$  aus  $E_{n+1}$  durch Weglassen der  $i$ -ten Zeile und  $j$ -ten Spalte entsteht.

Zu D2: Seien die  $r$ -te und  $s$ -te Zeile in  $A$  gleich für  $r < s$ . Dann hat für  $i \neq \{r, s\}$ , die Matrix  $A_{ij}$  zwei gleiche Zeilen und nach IV gilt  $\det_n(A_{ij}) = 0$ . Mit  $a_{rj} = a_{sj}$  gilt:

$$\det_n(A_{rj}) = (-1)^{s-r-1} \det_n(A_{sj}),$$

da bei der Transformation von  $A_{rj}$  zu  $A_{sj}$  die Zeilen mit Index  $r+1$  bis  $s$  jeweils einmal getauscht werden müssen. Insgesamt folgt:

$$\begin{aligned} \det_{n+1}(A) &= (-1)^{r+j} (-1)^{s-r-1} a_{rj} \det_n(A_{sj}) + (-1)^{s+j} (-1)^{s+j} a_{rj} \det_n(A_{sj}) \\ &= ((-1)^{r+j} (-1)^{s-r-1} + (-1)^{s+j}) a_{rj} \det_n(A_{sj}) \\ &= ((-1)^{r+j+s-r-1} + (-1)^{s+j}) a_{rj} \det_n(A_{sj}) \\ &= \underbrace{((-1)^{j+s-1} + (-1)^{s+j})}_{=0} a_{rj} \det_n(A_{sj}) \\ &= 0. \end{aligned}$$

Zu D1: Die Abbildung

$$A \mapsto \sum_{i=1}^{n+1} (-1)^{i+j} a_{ij} \det_n(A_{ij})$$

ist linear in den Zeilen  $a_r$ ,  $r = 1, \dots, n+1$ , falls für alle  $i, j \in \{1, \dots, n+1\}$  die Abbildung

$$A \mapsto a_{ij} \det_n(A_{ij})$$

linear in den Zeilen ist. Die Linearität in der  $r$ -ten Zeile für  $r \neq i$  folgt nach Induktionsvoraussetzung aus der Linearität von  $\det_n(A_{ij})$  in den Zeilen. Für  $r = i$  ist die Abbildung

$$M((n+1) \times (n+1), K) \rightarrow K, A \mapsto a_{ij}$$

linear in der  $i$ -ten Zeile, während  $A_{ij}$  nicht von  $a_{ij}$  abhängt. □

### 4.3 Laplacescher Entwicklungssatz

Wir halten nun die im Beweis von Satz 4.16 eingeführte Formel für die Determinante fest:

**Bemerkung 4.17.** 1. Entwicklung der Determinante nach der  $j$ -ten Spalte:

$$\det(A) := \sum_{i=1}^n (-1)^{i+j} a_{ij} \det(A_{ij}),$$

wobei  $A_{ij}$  die Matrix in  $M(n \times n, K)$ , die aus  $A$  durch Weglassen der  $i$ -ten Zeile und  $j$ -ten Spalte entsteht.

2. Mit Satz 4.12 können wir die Determinante auch nach der  $i$ -ten Zeile entwickeln:

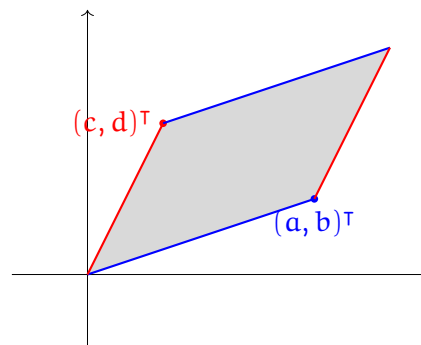
$$\det(A) := \sum_{j=1}^n (-1)^{i+j} a_{ij} \det(A_{ij}).$$

Diese Formeln ergeben eine rekursive Vorschrift, um die Determinante zu berechnen. Um  $\det(A)$  zu berechnen, müssen wir  $\det(A_{ij})$  berechnen, wobei die Matrizen  $A_{ij} \in M((n-1) \times (n-1), K)$  nun eine Dimension "verloren" haben. Die jeweiligen Werte  $\det(A_{ij})$  können wir wiederum mit Hilfe der Determinanten von Matrizen der Form  $(A_{ij})_{\ell k} \in M((n-2) \times (n-2), K)$  berechnet werden, und so weiter. In diesem Sinne ist es nützlich, die Determinanten der "kleinen" Matrizen zu kennen.

**Übung 4.18.** Für  $A \in M(2 \times 2, K)$  gilt

$$\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = ad - bc.$$

Wir können für  $(a, b)^T, (c, d)^T \in \mathbb{R}^2$  im  $\mathbb{R}^2$  den Wert  $|\det(A)|$  als die Fläche des von  $(0, 0)^T, (a, b)^T, (a, b)^T + (c, d)^T, (c, d)^T$  aufgespannten Parallelogramms interpretieren:



Für die rekursive Berechnung der Determinante ist es auch nützlich sich das durch  $(-1)^{i+j}$

induzierte Vorzeichenschema vor Augen zu führen:

$$\begin{pmatrix} + & - & + & \dots \\ - & + & - & \dots \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

**Beispiel 4.19.** Berechne

$$\det \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & 0 & 0 \\ 1 & 2 & 1 & 0 \end{pmatrix} = 0.$$

Wir werden nun eine Beziehung von Determinanten zu Lösungen von LGS der Form  $Ax = b$  herstellen. Dazu definieren wir den Begriff der **komplementären Matrix** von  $A$ .

**Definition 4.20.** Sei  $A \in M(n \times n, K)$ . Sei  $A_{ij}^+ \in M(n \times n, K)$  die Matrix, die aus  $A$  durch Ersetzen der  $i$ -ten Zeile und  $j$ -ten Spalte durch einen Einheitsvektor mit der 1 an der Stelle  $(i, j)$  entsteht (siehe Abbildung unten):

$$A_{ij}^+ = \begin{pmatrix} a_{11} & \dots & a_{1,j-1} & 0 & a_{1,j+1} & \dots & a_{1n} \\ \vdots & & & & & & \vdots \\ a_{i-1,1} & \dots & a_{i-1,j-1} & 0 & a_{i-1,j+1} & \dots & a_{i-1,n} \\ 0 & \dots & 0 & 1 & 0 & \dots & 0 \\ a_{i+1,1} & \dots & a_{i+1,j-1} & 0 & a_{i+1,j+1} & \dots & a_{i+1,n} \\ \vdots & & & & & & \vdots \\ a_{n1} & \dots & a_{n,j-1} & 0 & a_{n,j+1} & \dots & a_{nn} \end{pmatrix}.$$

Die Matrix  $A^\# = (a_{ij}^\#)$  mit  $a_{ij}^\# := \det(A_{ij}^+)$  heißt **komplementäre Matrix** von  $A$ .

**Lemma 4.21.** Die folgenden Aussagen gelten:

1.  $\det(A_{ij}^+) = (-1)^{i+j} \det(A_{ij})$ .
2.  $\det(A_{ij}^+) = \det(a^1, \dots, a^{j-1}, e^i, a^{j+1}, \dots, a^n)$ .

*Beweis.* Zu 1: Die Entwicklung der Determinante  $\det(A_{ij}^+)$  nach der  $i$ -ten Zeile oder der  $j$ -ten Spalte liefert direkt das gewünschte Ergebnis.

Zu 2: Durch Addition von Vielfachen der  $j$ -ten Spalte zu den anderen Spalten können wir  $(a^1, \dots, a^{j-1}, e^i, a^{j+1}, \dots, a^n)$  in  $A_{ij}^+$  überführen.  $\square$

**Satz 4.22.** Ist  $A \in M(n \times n, K)$  und  $A^\#$  die komplementäre Matrix, so ist

$$A^\# \cdot A = A \cdot A^\# = (\det(A)) \cdot E_n.$$

*Beweis.* Wir berechnen die Komponenten von  $A^\# \cdot A = (b_{ik})$ :

$$\begin{aligned} b_{ik} &= \sum_{j=1}^n a_{ij}^\# a_{jk} = \sum_{j=1}^n a_{jk} \det(A_{ji}^+) \\ &= \sum_{j=1}^n a_{jk} \cdot \det(a^1, \dots, a^{i-1}, e^j, a^{i+1}, \dots, a^n) \end{aligned} \quad (\text{Lem. 4.21})$$

$$= \det(a^1, \dots, a^{i-1}, \sum_{j=1}^n a_{jk} e^j, a^{i+1}, \dots, a^n) \quad (\text{D1})$$

$$\begin{aligned} &= \det(a^1, \dots, a^{i-1}, a^k, a^{i+1}, \dots, a^n) \\ &= \delta_{ik} \cdot \det(A). \end{aligned} \quad (\text{D2})$$

Es folgt somit  $A^\# \cdot A = (\det(A)) \cdot E_n$ . Der Wert  $A \cdot A^\#$  kann analog berechnen.  $\square$

Mit Satz 4.22 folgt unmittelbar, dass sich die Inverse der Matrix  $A$  berechnen lässt.

**Korollar 4.23.** Die Inverse einer Matrix  $A \in GL(n, K)$  kann dargestellt werden durch:

$$A^{-1} = \frac{1}{\det(A)} A^\# = \frac{1}{\det(A)} C^T,$$

wobei  $C_{ij} := (-1)^{i+j} \det(A_{ij})$ . Die letzte Gleichung folgt aus Lemma 4.21.

**Satz 4.24 (Cramersche Regel).** Sei  $A \in GL(n, K)$ ,  $b \in K^n$  und  $x \in K^n$  die eindeutige Lösung des LGS  $Ax = b$ . Dann gilt für alle  $i = 1, \dots, n$ :

$$x_i = \frac{\det(a^1, \dots, a^{i-1}, b, a^{i+1}, \dots, a^n)}{\det(A)}.$$

*Beweis.* Wir erhalten

$$Ax = b \Leftrightarrow x = A^{-1}b.$$

Mit  $A^{-1} = \frac{1}{\det(A)} \cdot A^\#$  folgt

$$a_{ij}^{-1} = \frac{\det(A_{ji}^+)}{\det(A)} = \frac{\det(a^1, \dots, a^{i-1}, e^j, a^{i+1}, \dots, a^n)}{\det(A)}.$$

Somit folgt für  $x_i$

$$x_i = \sum_{j=1}^n b_j \frac{\det(A_{ji}^+)}{\det(A)} \stackrel{D1}{=} \frac{\det(a^1, \dots, a^{i-1}, b, a^{i+1}, \dots, a^n)}{\det(A)}.$$

□

## 4.4 Leibnizsche Formel

Zum Abschluss werden wir nun die **Leibnizsche Formel** für die Determinante kennenlernen. Sie bietet eine geschlossene Formel und hat den Vorteil, dass wir direkt sehen, dass sie für  $K = \mathbb{R}$  eine stetige Funktion ist.

Sei  $S_n$  die Menge aller Permutationen auf der Menge  $\{1, \dots, n\}$ . Zur Erinnerung: eine Permutation ist eine bijektive Abbildung

$$\sigma : \{1, \dots, n\} \rightarrow \{1, \dots, n\}.$$

Für  $\sigma \in S_n$  definieren wir

$$\alpha(\sigma) := |\{(j, k) \in \{1, \dots, n\} \times \{1, \dots, n\} : j < k \text{ und } \sigma(j) > \sigma(k)\}|.$$

Beachte dass für eine Menge  $M = \{x_1, \dots, x_k\}$  der Ausdruck  $|M|$  die Anzahl ihrer Elemente definiert, also  $|M| = k$ . Die Abbildung  $\alpha$  zählt also die Anzahl der Zahlenpaare  $(j, k)$  bei der  $\sigma$  die Größenrelation umgedreht hat. Wir definieren das **Signum** von  $\sigma$  als

$$\text{Sign}(\sigma) := (-1)^{\alpha(\sigma)}.$$

**Definition 4.25.** Wir führen folgende Begriffe ein:

1. Eine **Nachbarvertauschung**  $\tau$  ist eine Permutation, welche nur zwei benachbarte Zahlen miteinander vertauscht und alle anderen Zahlen unverändert lässt. Dann gilt  $\alpha(\tau) = 1$ ,  $\text{Sign}(\tau) = -1$ , sowie

$$\alpha(\tau \circ \sigma) = \alpha(\sigma) \pm 1,$$

da  $\tau$  entweder einen weiteren Fehlstand hervorruft oder einen existierenden unter  $\sigma$  behebt. Für eine Nachbarvertauschung gilt:

$$\text{Sign}(\tau \circ \sigma) = (-1)^{\alpha(\sigma) \pm 1} = (-1)\text{Sign}(\sigma) = \text{Sign}(\tau)\text{Sign}(\sigma) \quad (4.1)$$

2. Eine **Transposition**  $\tau$  ist eine Permutation, welche zwei (nicht-notwendig benachbarte) Zahlen miteinander vertauscht und alle anderen Zahlen unverändert lässt. Liegen für eine Transposition  $\tau$  zwischen den beiden vertauschten Zahlen genau  $r$  weitere Zahlen, so können wir  $\tau$  mittels  $2r + 1$  Nachbarvertauschungen realisieren.

Es folgt:

$$\text{Sign}(\tau) = (-1)^{2r+1} = -1. \quad (4.2)$$

Jede Permutation lässt sich mittels endlicher vieler Transpositionen darstellen (ohne Beweis<sup>1</sup>)  $\sigma = \tau_1 \circ \dots \circ \tau_k$  und es gilt mit (4.1):

$$\text{Sign}(\sigma = \tau_1 \circ \dots \circ \tau_k) = (-1)^k.$$

Mit der Einsicht, dass  $\{-1, 1\}$  bzgl. der gewöhnlichen Multiplikation eine Gruppe ist, können wir schließen, dass

$$\text{Sign} : S_n \rightarrow \{-1, 1\},$$

ein Gruppenhomomorphismus ist.

**Satz 4.26 (Leibnizsche Formel).** Sei  $A = (a_{ij}) \in M(n \times n, K)$ . Dann gilt:

$$\det(A) = \sum_{\sigma \in S_n} \text{Sign}(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)} = \sum_{\sigma \in S_n} \text{Sign}(\sigma) \prod_{i=1}^n a_{i\sigma(i)}. \quad (4.3)$$

*Beweis.* Eigenschaft D1 folgt direkt aus der Linearität der einzelnen Summanden. Für Eigenschaft D2, seien  $j$ -te und  $k$ -te Zeile gleich und  $\tau$  gerade die Vertauschung von  $j$  und  $k$ . Setzen wir  $A_n := \{\sigma \in S_n : \text{Sign}(\sigma) = 1\}$  und  $\bar{A}_n := \{\sigma \circ \tau : \sigma \in A_n\}$ , so erhalten wir  $S_n = A_n \cup \bar{A}_n$ . Es folgt mit  $\text{Sign}(\sigma) = 1$  für  $\sigma \in A_n$  sowie  $\text{Sign}(\rho) = \text{Sign}(\sigma) \cdot \text{Sign}(\tau) = -1$  für  $\rho = \sigma \circ \tau \in \bar{A}_n$ :

$$\begin{aligned} \det(A) &= \sum_{\sigma \in A_n} \text{Sign}(\sigma) a_{1\sigma(1)} \cdots a_{n\sigma(n)} + \sum_{\rho \in \bar{A}_n} \text{Sign}(\rho) a_{1\rho(1)} \cdots a_{n\rho(n)} \\ &= \sum_{\sigma \in A_n} a_{1\sigma(1)} \cdots a_{n\sigma(n)} - \sum_{\sigma \in A_n} a_{1\sigma(\tau(1))} \cdots a_{n\sigma(\tau(n))} \\ &= 0. \end{aligned}$$

Die letzte Gleichung folgt aus folgender Überlegung: für  $\ell \notin \{j, k\}$  gilt die Gleichheit  $a_{\ell\sigma(\tau(\ell))} = a_{\ell\sigma(\ell)}$  und für  $\ell \in \{j, k\}$  gilt mit  $a_j = a_k$  die Gleichheit  $a_{j\sigma(\tau(j))} = a_{j\sigma(k)} = a_{k\sigma(k)}$ , sowie  $a_{k\sigma(\tau(k))} = a_{k\sigma(j)} = a_{j\sigma(j)}$ .

Eigenschaft D3 gilt, da mit  $A = E_n = (\delta_{ij})$  nur die Permutation  $\text{id} \in S_n$  einen nicht-null Beitrag von  $a_{11} \cdots a_{nn} = 1$  für die Summe in (4.3) leistet.  $\square$

<sup>1</sup>Ein ausführlicher Beweis findet sich zum Beispiel in Fischer und Springborn, Lineare Algebra, 19 Auflage (ab Seite 205).

## Kapitel 5

# Eigenwerte und Normalformen von Endomorphismen

Unser Ziel ist es, die Theorie der **Normalformen** für Darstellungsmatrizen von Endomorphismen fortzuführen.

Dabei wird sich herausstellen, dass die sogenannten **Eigenwerte** und **Eigenvektoren** einer Matrix  $A \in M(n \times n, K)$  von zentraler Bedeutung sind. Ein Eigenvektor  $v \in K^n, v \neq 0$  zum Eigenwert  $\lambda \in K$  ist durch folgende Eigenschaft definiert:

$$Av = \lambda v.$$

Diese Eigenschaft können wir umformulieren, so dass  $v$  folgendes homogene LGS löst:

$$(A - \lambda E_n)v = 0.$$

Wir wissen, dass die Existenz einer Lösung  $v \in K^n, v \neq 0$  für das LGS gleichbedeutend ist mit

$$\det(A - \lambda E_n) = 0.$$

Somit spielen die Nullstellen der Determinantenfunktion von speziellen in  $\lambda \in K$  **parametrisierten** Matrizen eine große Rolle. Wie in der Leibnizformel zu sehen, kann die Determinante als Summe von Produkten von Körperelementen aufgefasst werden und für obere Dreiecksmatrizen ergibt sich die Determinante sogar als Produkt der Diagonalelemente (siehe Lemma 4.7). Dies führt dann auf die Untersuchung von Nullstellen eines **Polynoms** in der **Unbekannten**  $\lambda \in K$ .

Bevor wir nun Polynome und schließlich die Theorie der Eigenwerte und Eigenvektoren entwickeln, werden wir zunächst Determinanten von Endomorphismen einführen. Sei dazu ein Endomorphismus  $F : V \rightarrow V$  mit  $\dim V = n < \infty$  gegeben. Für zwei Basen  $\mathcal{A}$  und  $\mathcal{B}$  von  $V$  gilt mit Satz 3.56, dass es  $S \in GL(n, K)$  gibt mit

$$\mathcal{M}_{\mathcal{A}}(F) = S\mathcal{M}_{\mathcal{B}}(F)S^{-1}.$$

Unter Ausnutzung von Satz 4.10 gilt somit:

$$\det \mathcal{M}_{\mathcal{A}}(F) = \det \mathcal{M}_{\mathcal{B}}(F),$$

und daher können wir einem Endomorphismus eine Determinante zuordnen:

**Definition 5.1.** Sei  $F \in \text{End}(V)$  und  $\mathcal{B}$  beliebige Basis von  $V$ . Dann ist die Determinante von  $F$  definiert als

$$\det F := \det \mathcal{M}_{\mathcal{B}}(F)$$

und wir erhalten eine wohldefinierte Abbildung

$$\begin{aligned} \det : \text{End}(V) &\rightarrow K \\ F &\mapsto \det F. \end{aligned}$$

**Bemerkung 5.2.** Für  $F \in \text{End}(V)$  sind äquivalent:

- $F$  ist surjektiv.
- $\det F \neq 0$

| **Übung 5.3.** Beweise Bemerkung 5.2.

## 5.1 Polynome

**Definition 5.4.** Ein **Polynom** mit Koeffizienten in  $K$  in der Unbekannten  $t$  ist ein Ausdruck der Gestalt

$$f = a_0 + a_1 t + \cdots + a_n t^n,$$

wobei  $a_0, a_1, \dots, a_n \in K$ . Mit  $K[t]$  bezeichnen wir die Menge solcher Polynome. Sind alle Koeffizienten  $a_i = 0, i = 0, \dots, n$ , so sprechen wir vom **Nullpolynom** und schreiben  $f = 0$ . Der **Grad** eines Polynoms  $f$  ist definiert durch

$$\deg f := \begin{cases} -\infty, & \text{falls } f = 0 \\ \max\{i \in \mathbb{N} : a_i \neq 0\}, & \text{sonst.} \end{cases}$$

Zwei Polynome  $f = a_0 + a_1 t + \cdots + a_n t^n \in K[t]$  und  $g = b_0 + b_1 t + \cdots + b_m t^m \in K[t]$  sind gleich, wir schreiben  $f = g$ , falls gilt

$$\deg f = \deg g \text{ und } a_i = b_i \text{ für alle } i = 1, \dots, \deg f.$$

Die Rolle der Unbekannten  $t$  und die Potenzen  $t^k$  sind allerdings nicht hinreichend geklärt. Wir werden im Folgenden  $t = \lambda \in K$  wählen und somit ist  $\lambda^k \in K$  wohldefiniert und wir erhalten

$$f(\lambda) := a_0 + a_1 \lambda + \cdots + a_n \lambda^n \in K.$$

Somit induziert ein Polynom  $f$  eine Abbildung

$$\tilde{f}: K \rightarrow K, \lambda \mapsto f(\lambda),$$

und insgesamt erhalten wir

$$\sim: K[t] \rightarrow \text{Abb}(K, K), f \mapsto \tilde{f}.$$

Wir geben ein Beispiel, welches verdeutlicht warum wir die Unterscheidung zwischen  $f$  und  $\tilde{f}$  benötigen.

**Beispiel 5.5.** Sei  $(K, \cdot, +)$  ein Körper mit  $K = \{0, 1\}$  und sei  $f = t + t^2$ . Wir erhalten für  $f \in K[t]$ :

$$f(0) = 0 + 0 = 0 \text{ und } f(1) = 1 + 1 = 0.$$

Also ist  $\tilde{f}$  die Nullabbildung, allerdings ist  $f \neq 0$ , weil  $a_1 = a_2 \neq 0$ . Wir sehen hiermit, dass die Abbildung  $\sim: K[t] \rightarrow \text{Abb}(K, K)$  nicht injektiv ist.

**Satz 5.6.** Das Tupel  $(K[t], +, \cdot)$  ist für  $f = a_0 + a_1t + \dots + a_nt^n$  und  $g = b_0 + b_1t + \dots + b_mt^m$  mittels

$$\begin{aligned} f + g &:= (a_0 + b_0) + \dots + (a_n + b_n)t^n \\ f \cdot g &:= c_0 + c_1t + \dots + c_{n+m}t^{n+m} \text{ mit } c_k = \sum_{i+j=k} a_i b_j \end{aligned}$$

ein kommutativer Ring, wobei für die Definition der Addition o.E.  $m = n$  angenommen werden kann. Weiterhin gilt

$$\deg(f \cdot g) = \deg f + \deg g,$$

wobei wir per Konvention  $n - \infty = -\infty + m = -\infty + (-\infty) = -\infty$  setzen.  $K[t]$  heißt **Polynomring**.

*Beweis.* Übung. □

Bei Ringen ist die Division nicht erklärt (das Inverse bzgl. der multiplikativen Verknüpfung muss nicht existieren), siehe z.B. den Ring der ganzen Zahlen. Allerdings können wir versuchen, eine **Division mit Rest** zu definieren.

**Satz 5.7.** Für  $f, g \in K[t]$  mit  $g \neq 0$  gibt es eindeutig bestimmte Polynome  $q, r \in K[t]$  mit

$$f = q \cdot g + r \text{ und } \deg r < \deg g.$$

Die obige Beziehung können wir als Division mit Rest in folgender Schreibweise auffassen:

$$\frac{f}{g} = q + \frac{r}{g}.$$

*Beweis.* Wir beginnen mit der Eindeutigkeit. Seien  $q, r, q', r' \in K[t]$  mit

$$f = q \cdot g + r = q' \cdot g + r' \text{ und } \deg r < \deg g, \deg r' < \deg g.$$

Subtraktion der beiden Gleichungen ergibt:

$$0 = (q - q') \cdot g + (r - r'), \text{ also } (q - q') \cdot g = r' - r. \quad (5.1)$$

Nach Definition 5.4 gilt, dass zwei gleiche Polynome den gleichen Grad haben und somit gilt:

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

Also ist  $\deg(q - q') < 0$  und damit  $\deg(q - q') = -\infty$ . Mit der Definition des Grads ergibt sich daraus direkt  $q = q'$ . Mit (5.1) folgt dann auch  $r = r'$ .

Um die Existenz von  $q$  und  $r$  nachzuweisen, werden wir nun das Polynomdivisionsverfahren angeben und seine Korrektheit zeigen. Dazu schreiben wir die Polynome nach fallenden Potenzen auf

$$f = a_n t^n + \dots + a_1 t + a_0, \quad g = b_m t^m + \dots + b_1 t + b_0,$$

wobei  $a_n \neq 0, b_m \neq 0$ . Falls  $n < m$ , so wähle  $q = 0$  und  $f = r$ , da  $f = 0 \cdot g + r$  und  $\deg f < \deg g$ . Für den interessanteren Fall  $n \geq m$  teilen wir zunächst die jeweiligen höchsten Terme um den Term höchster Potenz von  $q$  zu erhalten:

$$q_1 := \frac{a_n}{b_m} t^{n-m}.$$

Im nächsten Schritt betrachten wir

$$f_1 := f - q_1 g.$$

Nach Konstruktion ist  $\deg f_1 < \deg f$ . Falls  $\deg f_1 < m$ , so können wir  $q = q_1$  und  $r = f_1$  setzen. Anderenfalls wiederholen wir den Schritt mit  $f_1$ :

$$f_2 := f_1 - q_2 \cdot g \text{ mit } \deg f_2 < \deg f_1.$$

Die Grade nehmen bei jedem dieser Schritte um mindestens 1 ab und somit erhalten wir ein  $k \leq n - m$  mit

$$f_k := f_{k-1} \cdot g - q_k \cdot g \text{ und } \deg f_k < m.$$

Setzen wir die rekursiv definierten Terme  $f_k$  ineinander ein, erhalten wir

$$f = f_1 + q_1 g = (f_2 + q_2 \cdot g) + q_1 \cdot g = f_2 + (q_1 + q_2) \cdot g = \dots = f_k + (q_1 + \dots + q_k) \cdot g.$$

Somit erhalten wir schlussendlich

$$q := q_1 + \dots + q_k \text{ und } r := f_k. \quad \square$$

Wir interessieren uns nun für **Nullstellen** eines Polynoms  $f \in K[t]$  über dem Körper  $K$ . Das heißt, dass wir folgende Gleichung in der Unbestimmten  $\lambda \in K$  lösen wollen:

$$f(\lambda) = 0, \lambda \in K.$$

**Lemma 5.8.** Ist  $\lambda \in K$  eine Nullstelle von  $f \in K[t]$ , so gibt es ein eindeutig bestimmtes Polynom  $g \in K[t]$  mit den Eigenschaften:

1.  $f = (t - \lambda) \cdot g$ .
2.  $\deg g = (\deg f) - 1$ .

*Beweis.* Wir dividieren  $f$  durch  $(t - \lambda)$  mit Rest und erhalten eindeutig bestimmte  $g, r \in K[t]$  mit

$$f = (t - \lambda) \cdot g + r, \text{ und } \deg r < \deg(t - \lambda) = 1.$$

Wir erhalten also  $r = \alpha_0$  mit  $\alpha_0 \in K$ . Mittels  $f(\lambda) = 0$  erhalten wir

$$0 = (\lambda - \lambda) \cdot g(\lambda) + r = 0 + \alpha_0,$$

und somit ist  $r = \alpha_0 = 0$  und 1. folgt. Mit der Beziehung

$$\deg f = \deg(t - \lambda) + \deg g = 1 + \deg g$$

folgt auch 2. □

Wir diskutieren nun einige Folgerungen.

**Korollar 5.9.** Sei  $K$  ein Körper und  $f \in K[t]$  und sei  $k$  die Anzahl der Nullstellen von  $f$ . Falls  $f \neq 0$ , so gilt  $k \leq \deg f$ .

*Beweis.* Wir führen eine Induktion bzgl.  $\deg f \in \mathbb{N}$  durch. Sei  $\deg f = 0$ . Dann ist  $f = \alpha_0$  ein konstantes Polynom und hat keine Nullstelle. Sei nun  $\deg f = n$  und für alle Polynome  $g \in K[t]$  mit  $\deg g \leq n - 1$  gelte die Induktionsvoraussetzung. Wenn  $f$  keine Nullstelle hat, so gilt die zu beweisende Aussage. Ist nun  $\lambda \in K$  eine Nullstelle von  $f$ , so erhalten wir nach Division durch  $(t - \lambda)$  ein  $g \in K[t]$  mit

$$f = (t - \lambda) \cdot g, \text{ und } \deg g = n - 1.$$

Nun ist ersichtlich, dass alle verbleibenden Nullstellen von  $f$  auch Nullstellen von  $g$  sein müssen. Ist  $\ell$  die Anzahl der Nullstellen von  $g$ , so gilt nach Induktionsvoraussetzung  $\ell \leq \deg g = n - 1$  und insgesamt erhalten wir

$$k \leq 1 + \ell \leq n. \quad \square$$

Ist  $\lambda$  Nullstelle von  $f$ , so kann  $\lambda$  auch Nullstelle von  $g$  sein. In diesem Fall sprechen wir

von einer **mehrfachen Nullstelle** und die **Vielfachheit** bezeichnet die Häufigkeit mit der eine Nullstelle auftritt.

**Definition 5.10.** Sei  $f \in K[t]$  ungleich dem Nullpolynom und  $\lambda \in K$ . Die Zahl

$$\mu(f; \lambda) := \max\{\ell \in \mathbb{N} : f = (t - \lambda)^\ell \cdot g \text{ mit } g \in K[t]\}$$

heißt die **Vielfachheit** der Nullstelle  $\lambda \in K$  von  $f$ .

Sind  $\lambda_1, \dots, \lambda_k \in K$  die verschiedenen Nullstellen von  $f$  und ist  $\mu(f; \lambda_i) = \ell_i$ , so gilt

$$f = (t - \lambda_1)^{\ell_1} \cdots (t - \lambda_k)^{\ell_k} \cdot g,$$

wobei  $g$  ein Polynom mit Grad  $n - \sum_{i=1}^k \ell_i$  und ohne Nullstellen ist. Wir halten hier eine wichtige Aussage von Nullstellen von Polynomen in  $\mathbb{C}[t]$  fest:

**Satz 5.11 (Fundamentalsatz der Algebra).** Jedes Polynom  $f \in \mathbb{C}[t]$  mit  $\deg f > 1$  besitzt mindestens eine Nullstelle in  $\mathbb{C}$ .

Ein Beweis findet sich in den Standardlehrbüchern der Algebra, der Funktionentheorie oder der Analysis. Wir können von einem komplexen Polynom mit Grad  $> 1$  eine komplexe Nullstelle herausdividieren und dann wiederum den Existenzsatz für den Faktor  $g$  anwenden.

**Korollar 5.12.** Jedes Polynom  $f \in \mathbb{C}[t]$  mit  $\deg f > 0$  zerfällt in Linearfaktoren, d.h. es gibt  $a \in \mathbb{C}$  und  $\lambda_i \in \mathbb{C}, i = 1, \dots, n$  mit  $n = \deg f$  und

$$f = a(t - \lambda_1) \cdots (t - \lambda_n).$$

Es wird sich als nützlich erweisen, dass es für jeden Körper  $K$  einen sogenannten **Oberkörper**  $L \supset K$  gibt, so dass  $P \in L[t]$  in Linearfaktoren zerfällt.

**Satz 5.13.** Zu jedem Körper  $K$  gibt es einen Körper  $L \supset K$  (**Oberkörper**), so dass  $P \in L[t]$  in Linearfaktoren zerfällt. Der **Zerfallungskörper** oder auch **algebraischer Abschluss** wäre ein solcher Oberkörper. Für  $\mathbb{R}$  ist  $\mathbb{C}$  ein solcher Oberkörper.

Ein Beweis findet sich in den Standardlehrbüchern der Algebra.

## 5.2 Eigenwerte

Nun befassen wir uns wieder mit linearen Abbildungen und ihren Eigenschaften.

**Definition 5.14 (Eigenwert).** Sei  $F$  ein Endomorphismus eines  $K$ -Vektorraums  $V$ . Ein  $\lambda \in K$  heißt **Eigenwert** von  $F$ , wenn es  $v \in V, v \neq 0$  gibt mit

$$F(v) = \lambda v.$$

Ein solcher Nichtnullvektor  $v \in V$  heißt **Eigenvektor** von  $F$  zum Eigenwert  $\lambda$ .

**Lemma 5.15.** Seien  $F \in \text{End}(V)$  mit  $\dim V = n$  und  $v_i, i = 1, \dots, m$  Eigenvektoren zu paarweise verschiedenen Eigenwerten  $\lambda_i, i = 1, \dots, m$ . Dann sind  $v_i, i = 1, \dots, m$  linear unabhängig und somit gilt  $m \leq n$ .

*Beweis.* Wir führen eine Induktion bzgl.  $m \in \mathbb{N}$  durch. Für  $m = 1$  ist mit  $v_1 \neq 0$  der Induktionsanfang klar. Sei  $m \geq 2$  und die Aussage für  $m - 1$  wahr. Sei

$$\alpha_1 v_1 + \dots + \alpha_m v_m = 0. \quad (5.2)$$

Wende  $F$  an und wir erhalten:

$$\alpha_1 \lambda_1 v_1 + \dots + \alpha_m \lambda_m v_m = 0.$$

Wir können allerdings (5.2) auch mit  $\lambda_1$  multiplizieren und erhalten:

$$\alpha_1 \lambda_1 v_1 + \dots + \alpha_m \lambda_1 v_m = 0.$$

Subtraktion beider Gleichungen ergibt:

$$\alpha_2 (\lambda_2 - \lambda_1) v_2 + \dots + \alpha_m (\lambda_m - \lambda_1) v_m = 0.$$

Nach Induktionsannahme sind  $v_1, \dots, v_{m-1}$  linear unabhängig und somit folgt

$$\alpha_2 (\lambda_2 - \lambda_1) = \dots = \alpha_m (\lambda_m - \lambda_1) = 0$$

Da alle Eigenwerte paarweise verschieden sind, folgt  $\alpha_i = 0, i = 2, \dots, m$ . Setzen wir diese Werte in (5.2) ein, folgt  $\alpha_1 v_1 = 0 \Rightarrow_{v_1 \neq 0} \alpha_1 = 0$ .  $\square$

Das obige Lemma zeigt, dass es höchstens  $\dim V = n$  viele Eigenwerte geben kann. Die Menge der Eigenvektoren ist i.A. jedoch größer.

**Definition 5.16.** Ist  $F \in \text{End}(V)$  und  $\lambda \in K$ , so nennen wir

$$\text{Eig}(F; \lambda) := \{v \in V : F(v) = \lambda v\}$$

den **Eigenraum** von  $F$  bzgl.  $\lambda$ .

**Lemma 5.17.** 1.  $\text{Eig}(F; \lambda) \subset V$  ist ein Untervektorraum.

2.  $\lambda$  ist Eigenwert von  $F \Leftrightarrow \text{Eig}(V; \lambda) \neq \{0\}$ .

3.  $\text{Eig}(F; \lambda) \setminus \{0\}$  ist die Menge der zu  $\lambda \in K$  gehörigen Eigenvektoren von  $F$ .

4.  $\text{Eig}(F; \lambda) = \text{Ker}(F - \lambda \text{id}_V)$ .

5. Sind  $\lambda_1, \lambda_2 \in K$  zwei verschiedene Eigenwerte, so gilt

$$\text{Eig}(F; \lambda_1) \cap \text{Eig}(F; \lambda_2) = \{0\}.$$

*Beweis.* Nur der letzte Punkt bedarf einer Erläuterung. Seien  $F(v) = \lambda_1 v$  und  $F(v) = \lambda_2 v$ . Dann gilt  $(\lambda_1 - \lambda_2)v = 0 \Rightarrow v = 0$ .  $\square$

Wir werden nun einen Zusammenhang herstellen zwischen den Eigenwerten und Eigenvektoren eines Endomorphismus  $F \in \text{End}(V)$  mit  $\dim V = n$  und denen der Matrixabbildung  $F_A : K^n \rightarrow K^n$  für eine darstellende Matrix  $A = \mathcal{M}_{\mathcal{B}}(F)$  für eine beliebige Basis  $\mathcal{B} = (v_1, \dots, v_n)$  von  $V$ . Dazu nutzen wir wieder den eindeutigen Isomorphismus (siehe Korollar 3.26)

$$\Phi_{\mathcal{B}} : K^n \rightarrow V \text{ mit } \Phi_{\mathcal{B}}(e_j) = v_j, j = 1, \dots, n,$$

wobei  $e_j, j = 1, \dots, n$  wieder die kanonischen Basisvektoren des  $K^n$  bezeichnen.  $\Phi_{\mathcal{B}}$  war das durch  $\mathcal{B}$  bestimmte **Koordinatensystem** von  $V$  und für  $v = \Phi_{\mathcal{B}}((x_1, \dots, x_n)^T) = x_1 v_1 + \dots + x_n v_n \in V$  heißt der Vektor  $(x_1, \dots, x_n)^T$  **die Koordinaten von  $v$  (bzgl.  $\mathcal{B}$ )** und es gilt

$$x = (x_1, \dots, x_n)^T = \Phi_{\mathcal{B}}^{-1}(v).$$

**Satz 5.18.** Sei  $V$  ein  $K$ -Vektorraum mit  $\dim V = n$ ,  $F \in \text{End}(V)$  und  $A = \mathcal{M}_{\mathcal{B}}(F)$  für eine beliebige Basis  $\mathcal{B} = (v_1, \dots, v_n)$  von  $V$ . Dann sind folgende Aussagen äquivalent:

1.  $\lambda \in K$  ist ein Eigenwert mit Eigenvektor  $v \in V$  von  $F$ .

2.  $\lambda \in K$  ist ein Eigenwert mit Eigenvektor  $x := \Phi_{\mathcal{B}}^{-1}(v) \in K^n$  von  $F_A$ .

*Beweis.* 1. $\Rightarrow$ 2.: Sei  $\lambda \in K$  Eigenwert von  $F$  mit Eigenvektor  $v \in V$ . Für  $x := \Phi_{\mathcal{B}}^{-1}(v) \in K^n$  erhalten wir:

$$\begin{aligned} F_A(x) &= (\Phi_{\mathcal{B}}^{-1} \circ F \circ \Phi_{\mathcal{B}})(x) && \text{(Satz 3.53)} \\ &= \Phi_{\mathcal{B}}^{-1}(F(\Phi_{\mathcal{B}}(x))) && \text{(Definition der Komposition)} \\ &= \Phi_{\mathcal{B}}^{-1}(F(v)) && \text{(Definition von } \Phi_{\mathcal{B}}) \\ &= \Phi_{\mathcal{B}}^{-1}(\lambda v) && (\lambda, v \text{ Eigen(wert,vektor) von } F) \\ &= \lambda \Phi_{\mathcal{B}}^{-1}(v) && \text{(Linearität von } \Phi_{\mathcal{B}}) \\ &= \lambda x. && \text{(Definition von } \Phi_{\mathcal{B}}) \end{aligned}$$

2.  $\Rightarrow$  1.: Sei  $\lambda \in K$  Eigenwert von  $F_A$  mit Eigenvektor  $x \in K^n$ . Wir erhalten für  $v := \Phi_B(x)$ :

$$\begin{aligned}
 F(v) &= (\Phi_B \circ F_A \circ \Phi_B^{-1})(v) && \text{(Satz 3.53)} \\
 &= \Phi_B(F_A(\Phi_B^{-1}(v))) && \text{(Definition der Komposition)} \\
 &= \Phi_B(F_A(x)) && \text{(Definition von } \Phi_B) \\
 &= \Phi_B(\lambda x) && (\lambda, x \text{ Eigen(wert,vektor) von } F_A) \\
 &= \lambda \Phi_B(x) && \text{(Linearität von } \Phi_B) \\
 &= \lambda v. && \text{(Definition von } \Phi_B)
 \end{aligned}$$

□

Mit obigem Satz können wir leicht die Eigenwerte und Eigenvektoren von ähnlichen Matrizen in Beziehung zueinander setzen.

**Korollar 5.19.** Seien  $A, A' \in M(n \times n, K)$  ähnlich. Dann besitzen  $F_A$  und  $F_{A'}$  die exakt gleichen Eigenwerte.

*Beweis.* Mit Satz 3.60 folgt, dass es Basen  $\mathcal{A}, \mathcal{A}'$  von  $V$  sowie  $F \in \text{End}(V)$  gibt, so dass die ähnlichen Matrizen jeweils als Darstellungsmatrix von  $F$  bzgl. dieser Basen auftreten, d.h.

$$A = \mathcal{M}_{\mathcal{A}}(F), \quad A' = \mathcal{M}_{\mathcal{A}'}(F).$$

Aus Satz 5.18 folgt, dass  $F_A$  und  $F_{A'}$  die gleichen Eigenwerte haben (wie auch  $F$ ). □

**Übung 5.20.** Zeige, dass die Umkehrung der Aussage von Korollar 5.19 im Allgemeinen nicht gilt, d.h., es gibt Matrizen  $A, A' \in M(n \times n, K)$  mit gleichen Eigenwerten, die jedoch nicht ähnlich sind.

### 5.3 Das Charakteristische Polynom

Nun wenden wir uns der Frage zu, wie wir für einen Endomorphismus Eigenwerte und zugehörige Eigenräume berechnen können.

**Satz 5.21.** Für  $F \in \text{End}(V)$  mit  $\dim V = n$  und  $\lambda \in K$  sind äquivalent:

1.  $\lambda$  ist Eigenwert von  $F$ .
2.  $\det(F - \lambda \text{id}_V) = 0$ .

*Beweis.* Es gibt  $v \in V, v \neq 0$  mit  $F(v) = \lambda v$  genau dann, wenn folgende Bedingungen erfüllt sind:

$$\begin{aligned}
 &\exists v \in V \setminus \{0\} : F(v) = \lambda v \\
 \Leftrightarrow &\exists v \in V \setminus \{0\} : (F - \lambda \text{id}_V)(v) = 0 && \text{(Linearität von } F - \lambda \text{id}_V)
 \end{aligned}$$

$$\begin{aligned}
&\Leftrightarrow \text{Ker}(F - \lambda \text{id}_V) \neq \{0\} && \text{(Definition des Kerns)} \\
&\Leftrightarrow \text{Im}(F - \lambda \text{id}_V) \neq V && \text{(Dimensionsformel)} \\
&\Leftrightarrow \text{rang}(F - \lambda \text{id}_V) < \dim V && \text{(Definition des Ranges)} \\
&\Leftrightarrow \det(F - \lambda \text{id}_V) = 0 && \text{(Bemerkung 5.2)}
\end{aligned}$$

□

Satz 5.21 verdeutlicht, dass die Suche nach Eigenwerten auf die Suche nach Nullstellen der Abbildung

$$\tilde{P}_F : K \rightarrow K, \lambda \mapsto \det(F - \lambda \text{id}_V)$$

zurückgeführt werden kann. Diese Abbildung wird auch **charakteristische Funktion** von  $F$  genannt. Wir werden nun zeigen, dass diese Abbildung durch ein Polynom beschrieben werden kann.

Sei  $\mathcal{A}$  eine Basis von  $V$  mit  $\dim V = n$  und  $A = \mathcal{M}_{\mathcal{A}}(F)$ . Für eine Unbestimmte  $t$  definieren wir

$$P_A(t) := \det(A - tE_n) = \det \begin{pmatrix} a_{11} - t & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} - t & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} - t \end{pmatrix} \quad (5.3)$$

In dieser Darstellung können wir die Elemente der Matrix als Elemente des Rings  $K[t]$  der Polynome über  $K$  ansehen. Wir erhalten dann mittels der Leibnizformel aus Satz 4.26

$$P_A(t) = (a_{11} - t) \cdots (a_{nn} - t) + Q(t),$$

wobei der erste Summand zur identischen Permutation  $\text{id}$  gehört und  $Q(t)$  die restlichen Terme der Summe über alle anderen Permutationen darstellt. Hier ist zu beachten, dass alle Permutation in  $S_n \setminus \{\text{id}\}$  mindestens zwei Diagonalelemente des oben stehenden Matrixausdrucks nicht enthalten. Somit ist  $Q(t)$  ein Polynom in der Unbestimmten  $t$  vom Grad höchstens  $n - 2$ . Wir können nun auch den ersten Summand ausrechnen:

$$(a_{11} - t) \cdots (a_{nn} - t) = (-1)^n t^n + (-1)^{n-1} (a_{11} + \cdots + a_{nn}) t^{n-1} + Q'(t),$$

wobei  $Q'(t)$  ein Polynom in der Unbestimmten  $t$  vom Grad höchstens  $n - 2$ . Insgesamt gibt es also  $\alpha_0, \dots, \alpha_n \in K$  mit

$$P_A(t) = \alpha_n t^n + \dots + \alpha_1 t + \alpha_0,$$

wobei

$$\begin{aligned}
\alpha_n &= (-1)^n \\
\alpha_{n-1} &= (-1)^{n-1} (a_{11} + \cdots + a_{nn})
\end{aligned}$$

$$\alpha_0 = \det A.$$

Wir definieren die **Spur** von  $A$  durch  $\text{Spur}(A) := \alpha_{11} + \cdots + \alpha_{nn}$  und somit gilt  $\alpha_{n-1} = (-1)^{n-1} \text{Spur}(A)$ . Insgesamt erhalten wir, dass  $P_A(t)$  ein Element des Polynomrings  $K[t]$  ist und wir nennen

$$P_A(t) = \det(A - tE_n) \in K[t]$$

das **charakteristische Polynom** der Matrix  $A \in M(n \times n, K)$ . Setzen wir nun für die Unbestimmte  $t$  ein Körperelement  $\lambda \in K$  ein, so erhalten wir die Abbildung

$$\tilde{P}_A : K \rightarrow K, \lambda \mapsto P_A(\lambda).$$

Wir betrachten nun wiederum den Endomorphismus  $F \in \text{End}(V)$  mit  $\dim V = n$  und  $A = \mathcal{M}_{\mathcal{A}}(F)$ . Für jedes  $\lambda \in K$  gilt, dass die darstellende Matrix der linearen Abbildung  $F - \lambda \text{id}_V$  bezüglich einer Basis  $\mathcal{A}$  von  $V$  die folgende Form hat:

$$\mathcal{M}_{\mathcal{A}}(F - \lambda \text{id}_V) = A - \lambda E_n.$$

Somit gilt

$$\det(F - \lambda \text{id}_V) = \det(A - \lambda E_n) = P_A(\lambda),$$

was wiederum bedeutet, dass die charakteristische Funktion von  $F$  durch das charakteristische Polynom von  $A$  beschrieben wird. Wir zeigen nun, dass das charakteristische Polynom welches die charakteristische Funktion von  $F$  beschreibt, unabhängig von der Wahl der Basis  $\mathcal{A}$  von  $V$  ist (Erinnerung: mit Satz 3.60 gilt, dass zwei Matrizen  $A, B$  ähnlich sind, falls es Basen  $\mathcal{A}, \mathcal{B}$  von  $V$  gibt mit  $A = \mathcal{M}_{\mathcal{A}}(F)$  und  $B = \mathcal{M}_{\mathcal{B}}(F)$ ).

**| Lemma 5.22.** Ähnliche Matrizen haben das gleiche charakteristische Polynom.

*Beweis.* Sei  $A \approx B$ , d.h., es gibt  $S \in GL(n, k)$  mit

$$B = SAS^{-1}.$$

Indem wir nun formal im Ring  $M(n \times n, K[t])$  rechnen, erhalten wir

$$S \cdot t \cdot E_n \cdot S^{-1} = t \cdot E_n.$$

Somit folgt

$$B - t \cdot E_n = SAS^{-1} - S \cdot t \cdot E_n \cdot S^{-1} = S(A - t \cdot E_n)S^{-1}.$$

Wir wenden nun auf beiden Seiten der Gleichung die Determinante an und erhalten:

$$\det(B - t \cdot E_n) = \det(S(A - t \cdot E_n)S^{-1}) = \det S \cdot \det(A - t \cdot E_n) \cdot \det S^{-1} = \det(A - t \cdot E_n).$$

Somit folgt  $P_A(t) = P_B(t)$ . □

Mit diesem Lemma (und obiger Bemerkung) können wir einem Endomorphismus  $F \in \text{End}(V)$  ohne Schwierigkeiten ein charakteristisches Polynom zuordnen.

**Definition 5.23.** Sei  $F \in \text{End}(V)$  und  $\mathcal{A}$  Basis von  $V$  mit  $A = \mathcal{M}_{\mathcal{A}}(F)$ . Dann heißt

$$P_F(t) := P_A(t)$$

das **charakteristische Polynom** von  $F$ .

Wir halten die bisherigen Ergebnisse in folgendem Satz fest:

**Satz 5.24.** Sei  $F \in \text{End}(V)$  mit  $\dim V = n$ . Dann gilt für  $P_F(t) \in K[t]$ :

1.  $\deg P_F = n$ .
2.  $\tilde{P}_F$  beschreibt die Abbildung  $K \rightarrow K, \lambda \mapsto \det(F - \lambda \text{id}_V)$ .
3. Die Nullstellen von  $P_F$  sind die Eigenwerte von  $F$ .
4. Ist  $A = \mathcal{M}_{\mathcal{A}}(F)$ , so gilt  $P_F(t) = P_A(t) = \det(A - t \cdot E_n)$ .

Für einen Eigenwert  $\lambda \in K$  wollen wir uns nun überlegen, wie wir den zugehörigen Eigenraum  $\text{Eig}(F_A; \lambda)$  berechnen können. Wir schauen uns dazu den einfachen Fall  $F_A : K^n \rightarrow K^n$  an.

**Bemerkung 5.25.** Sei  $F_A \in \text{End}(V)$  mit  $A \in M(n \times n, K)$ . Dann ist  $\text{Eig}(F_A; \lambda)$  für jedes  $\lambda \in K$  der Lösungsraum des LGS:

$$(A - \lambda E_n)x = 0.$$

## 5.4 Diagonalisierung

Wir werden nun Verbindungen der Eigenwerte eines Endomorphismus  $F$  und der Darstellbarkeit von  $F$  mit Hilfe von **einfachen Matrizen** aufzeigen.

- Definition 5.26.**
1. Ein Endomorphismus  $F : V \rightarrow V$  heißt **diagonalisierbar**, wenn es eine Basis  $\mathcal{B}$  von  $V$  aus Eigenvektoren von  $F$  gibt.
  2. Eine Matrix  $A \in M(n \times n, K)$  heißt **diagonalisierbar**, wenn der durch  $A$  beschriebene Endomorphismus  $F_A : K^n \rightarrow K^n$  diagonalisierbar ist, d.h.,  $K^n$  besitzt eine Basis  $\mathcal{B}$  aus Eigenvektoren von  $F_A$ .

Die obige Definition ist durch folgenden Satz motiviert.

**Satz 5.27.** Sei  $\dim V = n < \infty$ . Dann sind äquivalent:

1.  $F \in \text{End}(V)$  ist diagonalisierbar.

2. Es gibt eine Basis  $\mathcal{B} = (v_1, \dots, v_n)$  von  $V$ , so dass gilt:

$$D := \mathcal{M}_{\mathcal{B}}(F) = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

3. Für beliebige Basis  $\mathcal{A}$  von  $V$  ist  $A := \mathcal{M}_{\mathcal{A}}(F)$  diagonalisierbar.

*Beweis.* 1. $\Rightarrow$ 2.: Sei  $\mathcal{B} = (v_1, \dots, v_n)$  eine Basis aus Eigenvektoren von  $V$ . Per Definition gilt, dass es  $\lambda_i \in K, i = 1, \dots, n$  gibt mit

$$F(v_i) = \lambda_i v_i \text{ für alle } i = 1, \dots, n. \quad (5.4)$$

Somit gibt es nach Satz 3.27 genau eine darstellende Matrix  $\mathcal{M}_{\mathcal{B}}$  von  $F$  bzgl.  $\mathcal{B}$ . Diese entspricht wegen (5.4) der obigen Form.

2. $\Rightarrow$ 3.: Mit Satz 3.60 gilt, dass  $A$  und  $D$  ähnlich sind und somit existiert  $S \in GL(n, K)$  mit

$$D = S^{-1}AS.$$

Die Spaltenvektoren  $(s^{-1})^i$  von  $S^{-1}$  sind Eigenvektoren von  $F_A$  bzgl. der Eigenwerte  $\lambda_i, i = 1, \dots, n$ , denn es gilt:

$$D = SAS^{-1} \Leftrightarrow S^{-1}D = AS^{-1} \Leftrightarrow \lambda_i (s^{-1})^i = A(s^{-1})^i \quad \forall i = 1, \dots, n. \quad (5.5)$$

Weiterhin sind die Eigenvektoren  $(s^{-1})^i, i = 1, \dots, n$  linear unabhängig in  $K^n$ , da  $S \in GL(n, K)$ . Somit ist  $A$  diagonalisierbar.

3. $\Rightarrow$ 1.: Sei  $\mathcal{B} = (x_1, \dots, x_n)$  eine Basis von Eigenvektoren von  $F_A$  zu Eigenwerten  $\lambda_1, \dots, \lambda_n$ . und sei  $\Phi_{\mathcal{B}} : K^n \rightarrow V$  der kanonische Isomorphismus. Setze  $v_i := \Phi_{\mathcal{B}}(x_i), i = 1, \dots, n$ . Mit Satz 5.18 gilt, dass  $\lambda_1, \dots, \lambda_n$  auch Eigenwerte von  $F$  mit den Eigenvektoren  $v_1, \dots, v_n$  sind. Da  $\Phi_{\mathcal{B}}$  ein Isomorphismus ist, sind die Vektoren  $v_i, i = 1, \dots, n$  eine Basis von  $V$ .  $\square$

**Satz 5.28.** Falls  $F \in \text{End}(V)$  mit  $\dim V = n$  paarweise verschiedene Eigenwerte  $\lambda_i, i = 1, \dots, n$  hat, so ist  $F$  diagonalisierbar.

*Beweis.* Siehe Lemma 5.15.  $\square$

Wir befassen uns nun mit Kriterien, die eine Diagonalisierung eines Endomorphismus erlauben, siehe Definition 5.26.

**Satz 5.29.** Sei  $F \in \text{End}(V)$  mit  $\dim V = n$ . Dann gilt:

1. Ist  $F$  diagonalisierbar, so gilt

$$P_F = \pm(t - \lambda_1) \cdot \dots \cdot (t - \lambda_n),$$

das heißt, das charakteristische Polynom zerfällt in Linearfaktoren.

2. Ist  $P_F = \pm(t - \lambda_1) \cdot \dots \cdot (t - \lambda_n)$  mit paarweise verschiedenen  $\lambda_i, i = 1, \dots, n$ , so ist  $F$  diagonalisierbar.

*Beweis.* 1.: Per Satz 5.27, folgt dass es eine Basis von Eigenvektoren  $\mathcal{A}$  von  $V$  gibt, so dass  $\mathcal{M}_{\mathcal{A}}(F) = A$ , wobei  $A$  eine Diagonalmatrix mit den Diagonalelementen bestehend aus Eigenwerten  $\lambda_i, i = 1, \dots, n$  von  $F$ . Somit können wir  $P_F = P_A$  ausrechnen und es gilt:

$$P_A(t) = \det(A - tE_n) = \pm(t - \lambda_1) \cdot \dots \cdot (t - \lambda_n).$$

Mit Satz 5.24 gilt, dass die Nullstellen von  $P_F$  den Eigenwerten von  $F$  entsprechen. Somit folgt 2. aus Satz 5.28.  $\square$

Satz 5.29 stellt zwar notwendige und hinreichende Bedingungen für die Diagonalisierbarkeit von  $F \in \text{End}(V)$  auf, allerdings besteht noch eine Lücke für den Fall, dass das charakteristische Polynom  $P_F$  mehrfache Nullstellen besitzt, d.h., wenn gilt

$$P_F = \pm(t - \lambda_1)^{r_1} \cdot \dots \cdot (t - \lambda_k)^{r_k},$$

wobei die  $\lambda_i, i = 1, \dots, k$  paarweise verschieden sind und  $1 \leq r_i \leq n$  für  $i = 1, \dots, k$  sowie  $r_1 + \dots + r_k = n$ . In der Notation von Kapitel 5.1 gilt  $r_i = \mu(P_F; \lambda_i)$ , d.h., dass  $r_i$  die Vielfachheit der Nullstelle  $\lambda_i$  von  $P_F$  angibt. Wie sich herausstellen wird, hängt eine vollständige Charakterisierung der Diagonalisierbarkeit von  $F \in \text{End}(V)$  nicht nur von den Eigenwerten und ihrer Vielfachheit ab, sondern auch von der Struktur der zugehörigen Eigenräume.

**Lemma 5.30.** Ist  $\lambda$  ein Eigenwert von  $F \in \text{End}(V)$  mit  $\dim V = n$ , so gilt:

$$1 \leq \dim \text{Eig}(F; \lambda) \leq \mu(P_F; \lambda).$$

*Beweis.* Sei  $v_1, \dots, v_s$  eine Basis von  $\text{Eig}(F; \lambda)$ . Mit der Definition des Eigenwertes folgt sofort  $1 \leq s$ , da es mindestens einen zugehörigen nicht-null Eigenvektor geben muss. Für die zweite Ungleichungen ergänzen wir  $v_1, \dots, v_s$  zu einer Basis

$$\mathcal{A} = (v_1, \dots, v_s, v_{s+1}, \dots, v_n)$$

von  $V$ . Wir erhalten

$$A := \mathcal{M}_{\mathcal{A}}(F) = \left[ \begin{array}{ccc|c} \lambda & & 0 & * \\ & \ddots & & \vdots \\ 0 & & \lambda & * \\ \hline & & 0 & A' \end{array} \right]$$

und mit Satz 4.14 folgt  $P_F = (t - \lambda)^s \cdot P_{A'}$ . Somit folgt  $\dim \text{Eig}(F; \lambda) = s \leq \mu(P_F; \lambda)$ , da  $\lambda$  durchaus auch als Nullstelle von  $P_{A'}$  auftreten kann.  $\square$

Nun erhalten wir ein Kriterium, welches eine vollständige Charakterisierung der Diagonalisierbarkeit von  $F \in \text{End}(V)$  erlaubt und sich auch rechnerisch überprüfen lässt.

**Satz 5.31.** Sei  $F \in \text{End}(V)$  mit  $\dim V = n$ . Dann sind äquivalent:

1.  $F$  ist diagonalisierbar.
2. (a) Das charakteristische Polynom  $P_F$  zerfällt in Linearfaktoren und  
(b)  $\dim \text{Eig}(F; \lambda) = \mu(P_F; \lambda)$  für alle Eigenwerte  $\lambda$  von  $F$ .
3. Für die paarweise verschiedenen Eigenwerte  $\lambda_1, \dots, \lambda_k$  von  $F$  gilt

$$V = \text{Eig}(F; \lambda_1) \oplus \dots \oplus \text{Eig}(F; \lambda_k).$$

*Beweis.*  $1 \Rightarrow 2$ : Ist  $F$  diagonalisierbar, so folgt mit Satz 5.29 direkt, dass Bedingung (2a) erfüllt ist. Für (2b), betrachten wir eine Basis von  $V$  aus Eigenvektoren  $v_1, \dots, v_n$ . Wir ordnen diese Eigenvektoren den jeweiligen Eigenwerten  $\lambda_1, \dots, \lambda_k$  zu, d.h. wir unterteilen die Basis in

$$(v_1^i, \dots, v_{s_i}^i) \text{ Basis von } \text{Eig}(F; \lambda_i).$$

Setzen wir nun  $r_i := \mu(P_F, \lambda_i)$ , so erhalten wir

$$s_1 + \dots + s_k = n, r_1 + \dots + r_k = n \text{ und } s_i \leq r_i,$$

wobei die letzte Ungleichung aus Lemma 5.30 folgt. Somit folgt  $s_i = r_i$  für alle  $i = 1, \dots, k$  und das entspricht Bedingung (2b).

$2 \Rightarrow 3$ : Sei

$$W := \text{Eig}(F; \lambda_1) + \dots + \text{Eig}(F; \lambda_k).$$

Da mit Lemma 5.15 die Eigenvektoren zu verschiedenen Eigenwerten linear unabhängig sind, folgt nach Definition 2.78, dass

$$W = \text{Eig}(F; \lambda_1) \oplus \dots \oplus \text{Eig}(F; \lambda_k).$$

Mit der Voraussetzung, dass  $P_F$  in Linearfaktoren zerfällt und zudem  $\dim \text{Eig}(F; \lambda_i) = \mu(P_F; \lambda_i)$  für alle  $i = 1, \dots, k$  gilt, folgt  $\dim W = \dim V$  und somit  $V = W$  (vgl. Satz 2.80).

$3 \Rightarrow 1$ : Für die paarweise verschiedenen Eigenwerte  $\lambda_1, \dots, \lambda_k$  von  $F$  gilt sei

$$\mathcal{B}_i := (v_1^i, \dots, v_{s_i}^i)$$

eine Basis von  $\text{Eig}(F; \lambda_i)$  für  $i = 1, \dots, k$ . Somit ist

$$\mathcal{B} = (v_1^1, \dots, v_{s_1}^1, \dots, v_1^k, \dots, v_{s_k}^k)$$



Mit der kanonischen Basis  $\mathcal{K} = (e^1, e^2, e^3)$  erhalten wir

$$A := M_{\mathcal{K}}(F) = \begin{pmatrix} 0 & -1 & 1 \\ -3 & -2 & 3 \\ -2 & -2 & 3 \end{pmatrix}$$

Das charakteristische Polynom berechnet sich zu

$$P_A(t) = -t^3 + t^2 + t - 1 = -(t-1)^2(t+1).$$

Somit sind  $\lambda_1 = 1$  und  $\lambda_2 = -1$  die beiden einzigen Eigenwerte von  $F$ . Der Eigenraum  $\text{Eig}(F; 1)$  ist der Lösungsraum von

$$\begin{pmatrix} -1 & -1 & 1 \\ -3 & -2 & 3 \\ -2 & -2 & 3 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}.$$

Elementare Zeilenumformungen führen auf

$$-x_1 - x_2 + x_3 = 0$$

und somit ist  $\mathcal{B}_1 = \left( \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} \right)$  eine Basis von  $\text{Eig}(F; 1)$ . Dies zeigt

$$\mu(P_F; 1) = 2 = \dim \text{Eig}(F; 1).$$

Für  $\text{Eig}(F; -1)$  erhalten wir

$$\begin{pmatrix} 1 & -1 & 1 \\ -3 & -2 & 3 \\ -2 & -2 & 3 \end{pmatrix} \cdot \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix},$$

was auf die Gleichungen

$$\begin{aligned} x_1 - x_2 + x_3 &= 0 \\ -4x_2 + 6x_3 &= 0 \end{aligned}$$

führt. Somit ist  $\mathcal{B}_2 = \left( \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix} \right)$  eine Basis von  $\text{Eig}(F; -1)$  und es gilt

$$\mu(P_F; -1) = 1 = \dim \text{Eig}(F; -1).$$

Insgesamt ist

$$\mathcal{B} = \left( \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}, \begin{pmatrix} 1 \\ 3 \\ 2 \end{pmatrix} \right)$$

eine Basis von  $\mathbb{R}^3$ . Wenn wir die Basisvektoren aus  $\mathcal{B}$  in eine Matrix  $S^{-1} \in \text{GL}(n, K)$  schreiben, erhalten wir nun die Beziehung

$$AS^{-1} = S^{-1}D \text{ (Übung!),}$$

wobei

$$D := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix}.$$

Multiplikation von links mit  $S \in \text{GL}(n, K)$  ergibt

$$SAS^{-1} = D.$$

Dies ergibt

$$S^{-1} = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & 3 \\ 1 & 1 & 2 \end{pmatrix}, \text{ also } S = \frac{1}{2} \begin{pmatrix} 1 & -1 & 1 \\ -3 & -1 & 3 \\ 1 & 1 & -1 \end{pmatrix},$$

wobei die Spalten von  $S^{-1}$  die Basisvektoren aus  $\mathcal{B}$  sind. Wir erhalten insgesamt

$$D := \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{pmatrix} \text{ und } D = SAS^{-1}.$$

## 5.5 Simultane Diagonalisierung

Wir möchten nun die Frage beantworten, unter welchen Umständen wir zwei gegebene Endomorphismen  $F, G \in \text{End}(V)$  **simultan** diagonalisieren können.

**Definition 5.33.** 1. Seien  $F, G \in \text{End}(V)$  mit  $\dim V = n$ .  $F$  und  $G$  heißen **simultan diagonalisierbar**, falls es eine Basis  $\mathcal{B} = (v_1, \dots, v_n)$  von  $V$  und  $\lambda_i, \mu_i \in K$  für  $i = 1, \dots, n$  gibt, so dass gilt

$$F(v_i) = \lambda_i v_i, G(v_i) = \mu_i v_i \text{ für alle } i = 1, \dots, n.$$

2. Seien  $A, A' \in M(n \times n, K)$ .  $A$  und  $B$  heißen **simultan diagonalisierbar**, falls es  $S \in \text{GL}(n, K)$  gibt, sodass

$$D = S^{-1}AS \text{ und } D' = S^{-1}BS,$$

Diagonalgestalt haben.

Wir führen zur Charakterisierung von simultaner Diagonalisierbarkeit, den Begriff von  $F$ -invarianten Räumen ein.

**Definition 5.34.** Sei  $V$  ein  $K$ -Vektorraum und  $F \in \text{End}(V)$ . Eine Teilmenge  $W \subset V$  heißt  **$F$ -invariant**, falls gilt

$$F(W) \subset W.$$

Bevor wir den allgemeinen Fall anschauen, betrachten wir zunächst den Fall der simultanen Diagonalisierbarkeit von zwei Matrizen. Gibt es zu  $A, B$  ein  $S \in \text{GL}(n, K)$  mit  $D = S^{-1}AS$  und  $D' = S^{-1}BS$ , so folgt aus  $D \cdot D' = D' \cdot D$ , dass

$$B \cdot A = S^{-1}D'SS^{-1}DS = S^{-1}DSS^{-1}D'S = A \cdot B.$$

Somit ist die Kommutativität bzgl. der Multiplikation der Matrixabbildungen schon ein notwendiges Kriterium.

**Satz 5.35.** Seien  $F, G \in \text{End}(V)$  mit  $\dim V = n$  diagonalisierbar und es gelte  $F \circ G = G \circ F$ . Dann sind  $F$  und  $G$  simultan diagonalisierbar.

*Beweis.* Da  $F$  und  $G$  nach Voraussetzung jeweils diagonalisierbar sind, gilt mit Satz 5.31, dass es eine Zerlegung von  $V$  in Eigenräume von  $F$  und  $G$  mit jeweils paarweise verschiedenen Eigenwerten gibt:

$$\begin{aligned} V &= \text{Eig}(F; \lambda_1) \oplus \dots \oplus \text{Eig}(F; \lambda_k) \\ &= \text{Eig}(G; \mu_1) \oplus \dots \oplus \text{Eig}(G; \mu_\ell). \end{aligned}$$

Wir definieren nun

$$W_{ij} := \text{Eig}(F; \lambda_i) \cap \text{Eig}(G; \mu_j) \text{ für } i = 1, \dots, k, j = 1, \dots, \ell.$$

Wir werden nun zeigen, dass die Basen von  $W_{ij}$ ,  $i = 1, \dots, k, j = 1, \dots, \ell$  vereinigt eine Basis von  $V$  sind. Per Konstruktion sind sie ja Eigenvektoren von  $F$  und  $G$  gleichermaßen. Dazu reicht es für ein  $i \in \{1, \dots, k\}$  die Beziehung

$$\text{Eig}(F; \lambda_i) = W_{i1} \oplus \dots \oplus W_{i\ell}$$

zu zeigen, weil ja

$$V = \text{Eig}(F; \lambda_1) \oplus \dots \oplus \text{Eig}(F; \lambda_k).$$

Mit Lemma 5.15 reicht es

$$\text{Eig}(F; \lambda_i) = W_{i1} + \dots + W_{i\ell}$$

zu zeigen. Sei also  $w \in \text{Eig}(F; \lambda_i) \subset V$ . Mit der Voraussetzung gibt es  $w_1, \dots, w_\ell$  mit  $w_j \in \text{Eig}(G; \mu_j)$  für  $j = 1, \dots, \ell$ , so dass gilt

$$w = w_1 + \dots + w_\ell.$$

Wir wollen zeigen, dass  $w_j \in W_{ij}$  für  $j = 1, \dots, \ell$  gilt. Anwendung von  $F$  ergibt

$$F(w_1) + \dots + F(w_\ell) = F(w) \underset{w \in \text{Eig}(F; \lambda_i)}{=} \lambda_i w = \lambda_i w_1 + \dots + \lambda_i w_\ell.$$

Angenommen, die Eigenräume  $\text{Eig}(G; \mu_j)$  wären jeweils  $F$ -invariant, das heißt,

$$F(\text{Eig}(G; \mu_j)) \subset \text{Eig}(G; \mu_j),$$

dann würde bei obiger Gleichung mit der Eindeutigkeit von Linearkombinationen bei direkten Summen (siehe Definition 2.78), die Bedingung  $F(w_j) = \lambda_i w_j$  und somit  $w_j \in W_{ij}$  für alle  $j = 1, \dots, \ell$  gelten. Bleibt also die  $F$ -Invarianz der Eigenräume von  $G$  zu zeigen. Wir rechnen für  $v \in \text{Eig}(G; \mu_j)$  unter Benutzung von  $F \circ G = G \circ F$  nach:

$$G(F(v)) = F(G(v)) = F(\mu_j v) = \mu_j F(v),$$

und somit gilt  $F(v) \in \text{Eig}(G; \mu_j)$ . □

**| Übung 5.36.** Führe eine simultane Diagonalisierung durch.

## 5.6 Trigonalisierung

In diesem Abschnitt werden wir zeigen, dass Endomorphismen, deren charakteristisches Polynom  $P_F$  in Linearfaktoren zerfällt, sich **trigonalisieren** lassen, d.h., dass sich eine obere Dreiecksmatrix als darstellende Matrizen finden lässt.

**Definition 5.37.** 1. Sei  $F \in \text{End}(V)$  mit  $\dim V = n$ .  $F$  heißt **trigonalisierbar**, falls es eine Basis  $\mathcal{B} = (v_1, \dots, v_n)$  von  $V$  gibt, so dass  $A := \mathcal{M}_{\mathcal{B}}(F)$  eine obere Dreiecksmatrix ist, das heißt,  $a_{ij} = 0$  für alle  $i > j$ .

2.  $A \in M(n \times n, K)$  heißt **trigonalisierbar**, falls es  $S \in GL(n, K)$  gibt, sodass

$$D = S^{-1}AS \text{ obere Dreiecksmatrix ist.}$$

Wir werden zur Frage der Trigonalisierbarkeit von Endomorphismen die  $F$ -invarianten Unterräume nutzen. Wenn  $F$  z.B. diagonalisierbar ist und  $\mathcal{B} = (v_1, \dots, v_n)$  eine Basis von Eigenvektoren bildet, so gibt es folgende Zerlegung von  $V$  in  $F$ -invariante Unterräume:

$$V = W_1 \oplus \dots \oplus W_n \text{ mit } W_i = K v_i, i = 1, \dots, n.$$

Im Falle mehrfacher Eigenwerte erhalten wir

$$V = \text{Eig}(F; \lambda_1) \oplus \dots \oplus \text{Eig}(F; \lambda_k) \text{ mit } \dim \text{Eig}(F; \lambda_i) = r_i, i = 1, \dots, k.$$

**Lemma 5.38.** Ist  $W \subset V$  ein  $F$ -invarianter Unterraum, so ist  $P_{F|_W}$  ein Teiler von  $P_F$ .

*Beweis.* Sei  $\mathcal{B}'$  eine Basis von  $W$  und sei  $\mathcal{M}_{\mathcal{B}'}(F|_W)$  die darstellende Matrix von

$$F_W : W \rightarrow W$$

bezüglich  $\mathcal{B}'$ . Wir ergänzen  $\mathcal{B}'$  zu einer Basis  $\mathcal{B}$  von  $V$  und erhalten

$$\mathcal{M}_{\mathcal{B}}(F) = \left[ \begin{array}{c|c} \mathcal{M}_{\mathcal{B}'}(F|_W) & * \\ \hline 0 & A \end{array} \right]$$

Mit Satz 4.14 folgt  $P_F = P_{F|_W} \cdot P_A$ . □

Sei  $A \in M(n \times n, K)$  eine obere Dreiecksmatrix. Für den Endomorphismus  $F_A : K^n \rightarrow K^n$  sind die Untervektorräume

$$W_r := \text{span}(e^1, \dots, e^r)$$

für  $1 \leq r \leq n$  invariant und es gilt

$$W_1 \subset W_2 \subset \dots \subset W_n = K^n,$$

wobei  $\dim W_i = i$  für alle  $i = 1, \dots, n$  gilt. Diese Eigenschaft werden wir nun unter dem Begriff **Fahne** für allgemeine Vektorräume definieren.

**Definition 5.39.** Eine **Fahne** in einem  $n$ -dimensionalen Vektorraum  $V$  ist eine Kette

$$\{0\} = V_0 \subset V_1 \subset V_2 \subset \dots \subset V_n = V,$$

mit  $\dim V_i = i$  für alle  $i = 1, \dots, n$ . Eine Fahne heißt  $F$ -invariant, falls gilt

$$F(V_i) \subset V_i, i = 1, \dots, n.$$

Wir erhalten nun folgende Charakterisierung der Trigonalisierbarkeit von Endomorphismen.

**Satz 5.40.** Sei  $F \in \text{End}(V)$  mit  $\dim V = n$ . Dann sind äquivalent:

1.  $F$  ist trigonalisierbar.
2. Das charakteristische Polynom  $P_F$  zerfällt in Linearfaktoren.
3. Es gibt eine  $F$ -invariante Fahne von  $V$ .

*Beweis.*  $1. \Rightarrow 2.$ : Da es eine Basis  $\mathcal{B}$  von  $V$  gibt, so dass  $A := \mathcal{M}_{\mathcal{B}}(F)$  eine obere Dreiecksmatrix ist, gilt

$$P_F(t) = \det(\mathcal{M}_{\mathcal{B}}(F) - t \cdot E_n) = \pm(t - a_{11}) \cdot \dots \cdot (t - a_{nn}),$$

für die Diagonalelemente  $a_{ii}, i = 1, \dots, n$  von  $A$ .

$2. \Rightarrow 3.$ : Wir beweisen diese Aussage per Induktion über  $n = \dim V$ . Für  $n = 0, 1$  ist die Aussage klar. Sei also  $n \geq 2$ . Zu  $\lambda_1$  wählen wir einen Eigenvektor  $v_1 \in \text{Eig}(F; \lambda_1)$  und ergänzen ihn zu einer Basis  $\mathcal{B} = (v_1, w_2, \dots, w_n)$  von  $V$ . Wir erhalten für  $V_1 = \text{span}(v_1)$  und  $W = \text{span}(w_2, \dots, w_n)$  dass

$$V = V_1 \oplus W,$$

und insbesondere gilt, dass  $V_1$   $F$ -invariant ist. Die darstellende Matrix  $\mathcal{M}_{\mathcal{B}}(F)$  hat die Form

$$\mathcal{M}_{\mathcal{B}}(F) = \begin{bmatrix} \lambda_1 & a_{12} & \cdots & a_{1n} \\ 0 & \boxed{\phantom{B}} \\ \vdots & & & \\ 0 & & & \end{bmatrix}$$

$W$  ist im allgemeinen nicht  $F$ -invariant aufgrund der Einträge  $a_{12}, \dots, a_{1n}$ . Um nun die Basis  $\mathcal{B}$  zu verändern, definieren wir die beiden linearen Abbildungen  $H : W \rightarrow V_1$  und  $G : W \rightarrow W$  durch

$$H(w_j) := a_{1j}v_1 \text{ und } G(w_j) := a_{2j}w_2 + \cdots + a_{nj}w_n$$

und erhalten

$$F(w) = H(w) + G(w) \text{ für alle } w \in W.$$

Für die charakteristischen Polynome gilt die Beziehung:

$$P_F = (\lambda_1 - t) \cdot P_G, \text{ und somit } P_G = (\lambda_2 - t) \cdot \dots \cdot (\lambda_n - t).$$

Wir wenden nun die Induktionsvoraussetzung auf den Raum  $W$  mit  $\dim W = n - 1$  an und erhalten eine  $G$ -invariante Fahne

$$\{0\} = W_0 \subset W_1 \subset W_2 \cdots \subset W_{n-1} = W.$$

Wir setzen nun  $V_r := V_1 + W_{r-1}$ ,  $r = 1, \dots, n$  und zeigen, dass wir so eine  $F$ -invariante Fahne erhalten. Für  $\mu v_1 \in V_1$ ,  $w \in W_{r-1}$ ,  $\mu \in K$  rechnen wir nach:

$$F(\mu v_1 + w) = \lambda_1 \mu v_1 + F(w) = \lambda_1 \mu v_1 + H(w) + G(w).$$

Wegen  $H(w) \in V_1$  und  $G(w) \in W_{r-1}$  folgt die Behauptung.

3.  $\Rightarrow$  1.: Sei  $\{0\} = V_0 \subset V_1 \subset \dots \subset V_n = V$  eine  $F$ -invariante Fahne von  $V$ . Durch  $(n - 1)$ -maliges Anwenden des Basisergänzungssatzes (Satz 2.66) erhalten wir dann eine Basis  $\mathcal{B} = (v_1, \dots, v_n)$  von  $V_n = V$  mit  $\text{span}(v_1, \dots, v_i) = V_i$  für alle  $i = 1, \dots, n$ . Aufgrund der  $F$ -Invarianz der gegebenen Fahne ist die darstellende Matrix  $M_{\mathcal{B}}(F)$  dann eine obere Dreiecksmatrix, also  $F$  trigonalisierbar.  $\square$

Wir werden nun einen Algorithmus zur Berechnung einer Trigonalisierung angeben.

Input:  $A \in M(n \times n, K)$  und  $P_A(t) = \pm(t - \lambda_1) \cdot \dots \cdot (t - \lambda_n)$  für  $\lambda_i \in K$ ,  $i = 1, \dots, n$ .  
Gesucht:  $S \in GL(n, K)$  mit  $D := S \cdot A \cdot S^{-1}$  obere Dreiecksmatrix.

1. Sei  $\mathcal{B}_1$  die kanonische Basis von  $K^n$  und setze  $W_1 := K^n$ . Berechne zum Eigenwert  $\lambda_1$  einen Eigenvektor  $v_1 \in \text{Eig}(F_A; \lambda_1)$ . Wende das Steinitzsche Austauschlemma (Lemma 2.63) an und bestimme  $j_1 \in \{1, \dots, n\}$ , so dass

$$\mathcal{B}_2 := (v_1, e_1, \dots, \bar{e}_{j_1}, \dots, e_n)$$

eine Basis von  $K^n$  ist, wobei  $\bar{\phantom{e}}$  bedeutet, dass der entsprechende Vektor ausgelassen wird. Setze  $S_1^{-1} := T_{\mathcal{B}_1}^{\mathcal{B}_2}$ . In diesem ersten Schritt besteht  $S_1^{-1}$  aus den Spaltenvektoren der Basis  $\mathcal{B}_2$ . Wir erhalten

$$A_2 := S_1 \cdot A \cdot S_1^{-1} = \begin{bmatrix} \lambda_1 & * & \cdots & * \\ 0 & \boxed{A'_2} & & \\ \vdots & & & \\ 0 & & & \end{bmatrix}$$

2. Im 2. Schritt setzen wir  $\mathcal{B}'_2 := (e_1, \dots, \bar{e}_{j_1}, \dots, e_n)$  und  $W_2 := \text{span}(\mathcal{B}'_2)$  und betrachten den (eindeutigen) Endomorphismus  $F_2 : W_2 \rightarrow W_2$  mit  $\mathcal{M}_{\mathcal{B}'_2}(F_2) = A'_2$ . Für diese Abbildung gilt dann

$$P_{F_2} = P_{A'_2}(t) = \pm(t - \lambda_2) \cdot \dots \cdot (t - \lambda_n).$$

Zu  $\lambda_2$  wähle einen Eigenvektor  $v_2 \in \text{Eig}(F_2; \lambda_2)$ . Wende das Steinitzsche Austauschlemma (Lemma 2.63) an und bestimme  $j_2 \neq j_1$  mit  $j_2 \in \{1, \dots, n\}$ , so dass

$$\mathcal{B}'_3 := (v_2, e_1, \dots, \bar{e}_{j_1}, \dots, \bar{e}_{j_2}, \dots, e_n)$$

eine Basis von  $W_2$  ist und somit

$$\mathcal{B}_3 := (v_1, v_2, e_1, \dots, \bar{e}_{j_1}, \dots, \bar{e}_{j_2}, \dots, e_n)$$

eine Basis von  $K^n$  ist. Setze  $S_2^{-1} := T_{\mathcal{B}_1}^{\mathcal{B}_3}$  und wir erhalten

$$A_3 := S_2 \cdot A \cdot S_2^{-1} = \begin{bmatrix} \lambda_1 & * & * & * & * \\ 0 & \lambda_2 & * & * & * \\ 0 & 0 & \boxed{A'_3} & & \\ 0 & 0 & & & \\ 0 & 0 & & & \end{bmatrix}$$

Bei der Berechnung von  $S_2^{-1}$  können wir benutzen, dass gilt  $T_{\mathcal{B}_1}^{\mathcal{B}_3} = T_{\mathcal{B}_1}^{\mathcal{B}_2} \cdot T_{\mathcal{B}_2}^{\mathcal{B}_3}$  sowie

$$T_{\mathcal{B}_2}^{\mathcal{B}_3} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & \boxed{T_{\mathcal{B}_2}^{\mathcal{B}'_3}} \\ \vdots & & & \\ 0 & & & \end{bmatrix}$$

3. Iteriere dieses Vorgehen und nach höchstens  $n - 1$  Schritten erhalten wir eine obere Dreiecksmatrix

$$D := A_n = S_{n-1} A S_{n-1}^{-1}.$$

Output:  $D, S_{n-1}, S_{n-1}^{-1}$ .

Wir geben ein Beispiel an.

**Beispiel 5.41.** Sei  $K = \mathbb{R}$  und  $n = 3$  und die Matrix  $A$  habe die Gestalt

$$A = \begin{pmatrix} 3 & 4 & 3 \\ -1 & 0 & -1 \\ 1 & 2 & 3 \end{pmatrix}.$$

Wir erhalten  $P_A(t) = -(t-2)^3$  und somit  $\lambda_1 = \lambda_2 = \lambda_3 = 2$ . Wegen  $\text{rang}(A - 2 \cdot E_3) = 2$  gilt

$$\dim \text{Eig}(A; 2) = 1 < 3 = \mu(P_A; 2)$$

und somit ist  $A$  nicht diagonalisierbar. Da  $P_A$  in Linearfaktoren zerfällt, ist  $A$  jedoch trigonalisierbar. Wir wenden obigen Algorithmus an.

1. Schritt: Wir wählen  $v_1 = \begin{pmatrix} 1 \\ -1 \\ 1 \end{pmatrix} \in \text{Eig}(A; 2)$ ,  $j_1 = 1$  und somit  $\mathcal{B}_2 = (v_1, e^2, e^3)$ ,

sowie

$$S_1^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}, S_1 = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}, A_2 = S_1 \cdot A \cdot S_1^{-1} = \begin{pmatrix} 2 & 4 & 3 \\ 0 & 4 & 2 \\ 0 & -2 & 0 \end{pmatrix}.$$

Somit erhalten wir

$$A'_2 = \begin{pmatrix} 4 & 2 \\ -2 & 0 \end{pmatrix}.$$

2. Schritt: Wähle  $v'_2 \in \text{Eig}(F_{A'_2}, 2) \subset \mathbb{R}^2$ , z.B.,  $v'_2 = (1, -1)^T$ . Damit ist  $v_2$  gegeben durch

$$v_2 = 1 \cdot e^2 - 1 \cdot e^3 \in \mathbb{R}^3,$$

und somit  $j_2 = 2$  und  $\mathcal{B}_3 = (v_1, v_2, e^3)$ . Wir erhalten

$$S_2^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 1 & -1 & 1 \end{pmatrix}, A_3 = \begin{pmatrix} 2 & 1 & 3 \\ 0 & 2 & 2 \\ 0 & 0 & 2 \end{pmatrix} = D.$$

Somit ist  $S = S_2$ .

## 5.7 Potenzen von Endomorphismen

Wir können die Eigenwertbedingung  $F(v) = \lambda v$  für ein  $F \in \text{End}(V)$  auch mit einem linearen Polynom

$$L(t) := t - \lambda$$

ausdrücken. Setzen wir für die Unbestimmte  $t$  den Endomorphismus  $F$  ein, so ergibt sich

$$L(F) := F - \lambda \text{id}_V$$

und es gilt  $\text{Eig}(F; \lambda) = \text{Ker } L(F)$ . Wir werden nun das Einsetzen von Endomorphismen  $F \in \text{End}(V)$  in ein Polynom formalisieren. Für ein Polynom  $P \in K[t]$  der Form

$$P(t) = \alpha_r t^r + \cdots + \alpha_1 t + \alpha_0 \in K[t]$$

definieren wir

$$P(F) := \alpha_r F^r + \cdots + \alpha_1 F + \alpha_0 \text{id}_V \in \text{End}(V),$$

wobei

$$F^k := \underbrace{F \circ \cdots \circ F}_{k\text{-mal}}.$$

Für eine Matrix  $A \in M(n \times n, K)$  können wir analog vorgehen:

$$P(A) := \alpha_r A^r + \cdots + \alpha_1 A + \alpha_0 E_n \in M(n \times n, K),$$

wobei  $A^k$  die  $k$ -fache Multiplikation der Matrix darstellt. Insgesamt haben wir nun schon drei verschiedene Arten von **Polynomfunktionen** kennengelernt:

1. Polynomfunktionen über dem Körper  $K$ :

$$K \rightarrow K, \lambda \rightarrow P(\lambda),$$

2. Polynomfunktionen über  $\text{End}(V)$ :

$$\text{End}(V) \rightarrow \text{End}(V), F \rightarrow P(F),$$

3. Polynomfunktionen über  $M(n \times n, K)$ :

$$M(n \times n, K) \rightarrow M(n \times n, K), A \rightarrow P(A).$$

Es ist auffällig, dass  $K$ ,  $\text{End}(V)$  und  $M(n \times n, K)$  jeweils auch  $K$ -Vektorräume sind, auf denen zusätzlich eine Multiplikation definiert ist, die sie zu Ringen macht. Solche Mengen heißen auch  **$K$ -Algebren**.

**| Definition 5.42 ( $K$ -Algebra).** Ein Ring  $(R, +, \cdot)$  mit 1-Element der Multiplikation, der

gleichzeitig ein  $K$ -Vektorraum ist und für den außerdem gilt:

$$\forall \lambda \in K, a, b \in R : \lambda(a \cdot b) = (\lambda a) \cdot b = a \cdot (\lambda b)$$

heißt  $K$ -Algebra (mit 1).

**Definition 5.43.** 1. Seien  $(R, +, \cdot)$  und  $(S, \oplus, \otimes)$  zwei Ringe mit Einselement. Dann heißt eine Abbildung  $\phi : R \rightarrow S$  ein **Homomorphismen von Ringen**, wenn für alle  $a, b \in R$  gilt:

$$\phi(a + b) = \phi(a) \oplus \phi(b)$$

$$\phi(a \cdot b) = \phi(a) \otimes \phi(b)$$

$$\phi(1_R) = 1_S.$$

Bei Homomorphismen von Ringen ist es also egal, ob man erst zwei Elemente verknüpft, und das Ergebnis abbildet, oder erst die zwei Elemente abbildet, und dann die Bilder verknüpft.

2. Seien  $(R, +, \cdot)$  und  $(S, \oplus, \otimes)$  zwei  $K$ -Algebren.  $\phi$  heißt **Homomorphismus von  $K$ -Algebren**, falls zusätzlich zu obigen Bedingungen gilt:  $\phi(\lambda a) = \lambda \phi(a)$  für alle  $\lambda \in K$  und  $a \in R$  ( $\phi$  ist also linear). Falls  $\phi$  zusätzlich bijektiv ist, sprechen wir von einem Isomorphismus von  $K$ -Algebren.

Nun wenden wir uns der Frage zu, wie verschiedene Polynome ein festes Körperlement, einen festgehaltenen Endomorphismus oder eine festgehaltene Matrix abbilden. Dies führt auf die **Einsetzungshomomorphismen**.

**Definition 5.44.** Die **Einsetzungshomomorphismen** für  $x \in K, F \in \text{End}(V)$  und  $A \in M(n \times n, K)$  sind die Abbildungen:

$$\text{ev}_x : K[t] \rightarrow K, \quad \text{ev}_x(P) = P(x)$$

$$\text{ev}_F : K[t] \rightarrow \text{End}(V), \quad \text{ev}_F(P) = P(F)$$

$$\text{ev}_A : K[t] \rightarrow M(n \times n, K), \quad \text{ev}_A(P) = P(A).$$

**Bemerkung 5.45.** 1. Die Einsetzungshomomorphismen sind **Homomorphismen von  $K$ -Algebren**.

2. Sei  $V$  ein  $K$ -Vektorraum mit  $\dim V = n$  und sei  $\mathcal{B}$  Basis von  $V$ . Dann ist

$$\Phi : \text{End}(V) \rightarrow M(n \times n, K), \quad F \mapsto \mathcal{M}_{\mathcal{B}}(F)$$

ein Isomorphismus von  $K$ -Algebren.

Wir werden uns nun weiter mit den Eigenschaften von Polynomen nach Einsetzung von Endomorphismen vertraut machen.

**Bemerkung 5.46.** Sei  $P \in K[t]$  ein Polynom der Form

$$P(t) = \alpha_r t^r + \cdots + \alpha_1 t + \alpha_0 \in K[t]$$

1. Ein Eigenvektor  $v$  von  $F \in \text{End}(V)$  zum Eigenwert  $\lambda \in K$  ist auch ein Eigenvektor von  $P(F) \in \text{End}(V)$  zum Eigenwert  $P(\lambda) \in K$ .
2. Wenn  $F$  diagonalisierbar ist, dann ist es auch  $P(F)$ .
3. Sei  $D \in M(n \times n, K)$  eine Diagonalmatrix mit Diagonalelementen  $\lambda_1, \dots, \lambda_n$ . Dann ist auch  $P(D)$  eine Diagonalmatrix mit Diagonalelementen  $P(\lambda_1), \dots, P(\lambda_n)$ .
4. Für  $A \in M(n \times n, K)$  mit  $B = SAS^{-1}$ ,  $S \in GL(n, K)$  gilt:  $P(B) = SP(A)S^{-1}$ .
5. Sei  $A \in M(n \times n, K)$  diagonalisierbar mit  $A = SDS^{-1}$  für eine Diagonalmatrix  $D$ . Dann gilt  $P(A) = SP(D)S^{-1}$  mit  $P(D)$  wie in 3.
6. Wenn  $A = \mathcal{M}_{\mathcal{B}}(F)$  für eine Basis  $\mathcal{B}$ , dann ist  $P(A) = \mathcal{M}_{\mathcal{B}}(P(F))$ .

*Beweis.* Übungen. □

Wir erhalten folgendes Resultat mit Hilfe des Einsetzungshomomorphismus  $\text{ev}_F$ .

**Lemma 5.47.** Seien  $P$  und  $Q$  aus  $K[t]$ . Dann gilt

$$P(F) \circ Q(F) = Q(F) \circ P(F).$$

*Beweis.* Wir erhalten folgende Kette von Gleichungen:

$$\begin{aligned} P(F) \circ Q(F) &= \text{ev}_F(P) \circ \text{ev}_F(Q) \stackrel{\text{Bem. 5.45}}{=} \text{ev}_F(PQ) \stackrel{K[t] \text{ kommutativ}}{=} \text{ev}_F(QP) \\ &= \text{ev}_F(Q) \circ \text{ev}_F(P) = Q(F) \circ P(F). \end{aligned}$$

□

- Lemma 5.48.**
1. Für  $F, G \in \text{End}(V)$  mit  $F \circ G = G \circ F$  ist  $\text{Ker } G$  ein  $F$ -invarianter Unterraum.
  2. Für  $F \in \text{End}(V)$  und jedes Polynom  $P \in K[t]$  ist  $\text{Ker } P(F)$  ein  $F$ -invarianter Unterraum.

*Beweis.* Zu 1: Sei  $v \in \text{Ker } G$  also  $G(v) = 0$ . Wir müssen zeigen, dass  $F(v) \in \text{Ker } G$  gilt. Es folgt  $G(F(v)) = F(G(v)) = F(0) = 0$ , also  $F(v) \in \text{Ker } G$ .

Zu 2: Mit Lemma 5.47 gilt  $P(F) \circ F = F \circ P(F)$ . Mit Punkt 1 folgt die Behauptung. □

Im Allgemeinen gilt  $\text{Ker } P(F) \neq \{0\}$  nur für spezielle Polynome. Wir werden im Folgenden den Extremfall  $P(F) = 0$  für ein Nichtnullpolynom  $P \in K[t]$  untersuchen.

**Lemma 5.49.** Sei  $V$  ein  $K$ -Vektorraum mit  $\dim V = n$  und  $F \in \text{End}(V)$ . Dann gibt es ein Polynom  $P \in K[t]$  mit  $1 \leq \deg P \leq n^2$  und  $P(F) = 0$ .

*Beweis.* Es gilt  $\dim \text{End}(V) = n^2$  (warum?) und daher sind die  $n^2 + 1$  Endomorphismen

$$\text{id}_V, F, F^2, \dots, F^{n^2}$$

linear abhängig. Es gibt also Elemente  $\alpha_0, \dots, \alpha_{n^2}$  wobei mindestens eines der Elemente nicht null ist mit

$$\alpha_0 \text{id}_V + \alpha_1 F + \dots + \alpha_{n^2} F^{n^2} = 0.$$

Somit gilt für  $P(t) = \alpha_0 + \alpha_1 t + \dots + \alpha_{n^2} t^{n^2}$  die Eigenschaft  $P(F) = 0$ . Die Gradeigenschaft folgt aus der Überlegung, dass der Fall  $\deg P = -\infty$  auf  $\alpha_0 = \dots = \alpha_{n^2} = 0$  führt, Widerspruch.  $\square$

Wir werden nun die Dimensionsschranke  $n^2$  in Lemma 5.49 auf  $n = \dim V$  reduzieren. Dies geschieht mit dem Satz von Cayley-Hamilton. Zunächst beweisen wir den Satz für diagonalisierbare Endomorphismen.

**Satz 5.50 (Satz von Cayley-Hamilton, Diagonalisierbare Endomorphismen).** Sei  $V$  ein  $K$ -Vektorraum mit  $\dim V = n$  und  $F \in \text{End}(V)$  diagonalisierbar mit charakteristischem Polynom  $P_F \in K[t]$  (vom Grad  $n$ ). Dann gilt

$$P_F(F) = 0 \in \text{End}(V).$$

Analog gilt für jede diagonalisierbare Matrix  $A \in M(n \times n, K)$ :

$$P_A(A) = 0 \in M(n \times n, K).$$

*Beweis.* Sei  $(v_1, \dots, v_n)$  eine Basis aus Eigenvektoren zu den Eigenwerten  $\lambda_1, \dots, \lambda_n$  von  $F$ . Für  $v \in V$  beliebig erhalten wir  $v = \alpha_1 v_1 + \dots + \alpha_n v_n$ . Wir erhalten

$$P_F(F)(v) = \sum_{j=1}^n \alpha_j P_F(F)(v_j) \stackrel{\text{Bem. 5.46, 1}}{=} \sum_{j=1}^n \alpha_j P_F(\lambda_j) v_j = 0,$$

wobei wir für die letzte Gleichung  $P_F(\lambda_i) = 0 \in K$  für alle  $i = 1, \dots, n$  ausgenutzt haben. Die Aussage für Matrizen folgt gleichermaßen indem wir den Endomorphismus  $F_A$  betrachten.  $\square$

Nun werden wir den Satz von Cayley-Hamilton in voller Allgemeinheit zeigen.

**Satz 5.51** (Satz von Cayley-Hamilton). Sei  $V$  ein  $K$ -Vektorraum mit  $\dim V = n$  und  $F \in \text{End}(V)$  mit charakteristischem Polynom  $P_F \in K[t]$  (vom Grad  $n$ ). Dann gilt

$$P_F(F) = 0 \in \text{End}(V).$$

Analog gilt für jede Matrix  $A \in M(n \times n, K)$ :

$$P_A(A) = 0 \in M(n \times n, K).$$

*Beweis.* Wir starten mit dem Spezialfall  $K = \mathbb{C}$ . Da jedes Polynom in  $\mathbb{C}$  in Linearfaktoren zerfällt, ist  $F$  trigonalisierbar, d.h. es gibt eine  $F$ -invariante Fahne

$$\{0\} = V_0 \subset V_1 \subset \dots \subset V_n = V.$$

Ist  $\mathcal{B} = (v_1, \dots, v_n)$  eine Basis von  $V$  mit  $V_i = \text{span}(v_1, \dots, v_i)$  für  $i = 1, \dots, n$ , so gilt

$$\mathcal{M}_{\mathcal{B}}(F) = \begin{pmatrix} \lambda_1 & & * \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$$

wobei  $\lambda_1, \dots, \lambda_n \in \mathbb{C}$  die Eigenwerte von  $F$  sind. Mit der Form

$$P_F(t) = (\lambda_1 - t) \cdot \dots \cdot (\lambda_n - t),$$

können wir  $P_F(F)$  schrittweise aufbauen: Für  $i = 1, \dots, n$  definiere

$$G_i := (\lambda_1 \text{id}_V - F) \circ (\lambda_2 \text{id}_V - F) \circ \dots \circ (\lambda_i \text{id}_V - F) \in \text{End}(V).$$

Wir erhalten  $G_n = P_F(F)$ . Wir werden also mittels Induktion über  $i = 1, \dots, n$  zeigen, dass  $G_i(V_i) = \{0\}$  für  $i = 1, \dots, n$ . Wir beginnen mit  $i = 1$ . Es gilt  $F(v_1) = \lambda_1 v_1$  und somit folgt  $G_1(v_1) = (\lambda_1 \text{id}_V - F)(v_1) = 0$ . Sei nun  $i \geq 2$  und  $v \in V_i$ . Dann gibt es  $w \in V_{i-1}$  und  $\mu \in \mathbb{C}$  mit  $v = w + \mu v_i$ . Mit

$$\lambda_i w - F(w) \in V_{i-1} \text{ und } \lambda_i v_i - F(v_i) \in V_{i-1} \text{ (siehe } \mathcal{M}_{\mathcal{B}}(F)\text{!)}$$

gilt nach Induktionsvoraussetzung

$$G_i(w) = (G_{i-1} \circ (\lambda_i \text{id}_V - F))(w) = G_{i-1}(\lambda_i w - F(w)) = 0.$$

sowie

$$G_i(v_i) = (G_{i-1} \circ (\lambda_i \text{id}_V - F))(v_i) = G_{i-1}(\lambda_i v_i - F(v_i)) = 0.$$

Also folgt

$$G_i(v) = G_i(w) + \mu G_i(v_i) = 0.$$

Nun überlegen wir uns, wie wir den Fall  $K = \mathbb{R}$  behandeln können. In diesem Fall zerfällt  $P_F(t)$  über  $\mathbb{R}$  ja nicht unbedingt in Linearfaktoren. Wir wenden nun jedoch eine "Komplexifizierungsmethode" an. Sei

$$A := \mathcal{M}_{\mathcal{B}}(F)$$

für eine beliebige Basis von  $V$  ( $A$  ist also eine Matrix mit Einträgen in  $\mathbb{R}$ ). Indem wir jedoch die Einträge aus  $A$  als komplexe Zahlen auffassen (wie das geht haben wir in Lineare Algebra I kennengelernt) erhalten wir einen Endomorphismus

$$F_A : \mathbb{C}^n \rightarrow \mathbb{C}^n.$$

Für diesen Endomorphismus hatten wir ja bereits  $P_{F_A}(F_A) = 0$  gezeigt. Ferner gilt  $P_{F_A} = P_A = P_F$ , nach Definition des Charakteristischen Polynoms von Endomorphismen und weil  $A$  darstellende Matrix sowohl von  $F_A$  als auch von  $F$  ist.<sup>1</sup> Zusammen zeigt dies, dass  $P_F(F_A) = 0$  und daher  $P_F(A) = 0$  (mit der Definition von  $F_A$ ). Unter Anwendung von Bemerkung 5.46.6 erhalten wir damit

$$\mathcal{M}_{\mathcal{B}}(P_F(F)) = P_F(\mathcal{M}_{\mathcal{B}}(F)) = P_F(A) = 0.$$

Somit folgt dann auch  $P_F(F) = 0 \in \text{End}(V)$  (benutze hier die Definition von  $\mathcal{M}_{\mathcal{B}}(P_F(F))!$ ). Nun betrachten wir einen allgemeinen Körper  $K$ . Wie in Satz 5.13 beschrieben, gibt es einen Oberkörper  $L \supset K$ , sodass  $P_F(t)$  über  $L$  in Linearfaktoren zerfällt. Nun wenden wir die gleiche Argumentation wie oben an (an dieser Stelle gehen wir nicht ins Detail wie wir Elemente aus  $K$  auch als Elemente in  $L$  auffassen).  $\square$

Nun werden wir uns weiter mit den Polynomen befassen, für die gilt  $P(F) = 0$ . Wir hatten gesehen, dass

$$\Phi_F : K[t] \rightarrow \text{End}(V), P(t) \mapsto P(F)$$

ein Homomorphismus von Ringen ist. Es gilt weiterhin folgende Eigenschaft:

**Übung 5.52.** Das Bild  $K[F] = \{P(F) : P(t) \in K[t]\} \subset \text{End}(V)$  ist ein kommutativer Unterring des Ringes  $\text{End}(V)$ .

Wir definieren

$$\mathcal{J}_F := \{P(t) \in K[t] : P(F) = 0\} \subset K[t]$$

als das **Ideal von  $F$** . Mit dem Satz von Cayley-Hamilton gilt offensichtlich, dass  $P_F \in \mathcal{J}_F$ . Wir werden nun die Struktur von  $\mathcal{J}_F$  weiter untersuchen.

**Definition 5.53.** Eine nicht-leere Teilmenge  $\mathcal{J}$  eines kommutativen Rings  $(R, +, \cdot)$  heißt **Ideal**, wenn gilt:

$$I1: P, Q \in \mathcal{J} \Rightarrow P - Q \in \mathcal{J},$$

<sup>1</sup>Wir verwenden hier implizit, dass man  $M(n \times n, \mathbb{R})$  als Teilmenge von  $M(n \times n, \mathbb{C})$  und  $\mathbb{R}[t]$  als Teilmenge von  $\mathbb{C}[t]$  auffassen kann (analog zu  $\mathbb{R} \subset \mathbb{C}$ ).

I2:  $P \in \mathcal{J}, Q \in \mathbb{R} \Rightarrow Q \cdot P \in \mathcal{J}$ .

**Übung 5.54.** Zeige, dass  $\mathcal{J}_F$  ein Ideal ist.

**Satz 5.55.** Zu jedem Ideal  $\mathcal{J} \subset K[t]$  mit  $\mathcal{J} \neq \{0\}$  gibt es ein eindeutiges Polynom  $M \in \mathcal{J}$  mit den folgenden Eigenschaften:

1.  $M$  ist normalisiert, d.h.,  $M = t^d + \dots + a_0$  wobei  $d = \deg M$ .
2. Für jedes  $P \in \mathcal{J}$  gibt es  $Q \in K[t]$  mit  $P = Q \cdot M$ .

$M$  heißt **Minimalpolynom von  $\mathcal{J}$** . Im Fall  $\mathcal{J} = \mathcal{J}_F$  nennen wir  $M$  auch das **Minimalpolynom von  $F$**  und bezeichnen es üblicherweise mit  $M_F$ .

*Beweis.* Sei  $d := \min\{\deg P : P \in \mathcal{J} \setminus \{0\}\} \in \mathbb{N} \cup \{0\}$ . Wähle nun ein beliebiges Element  $P = a_0 + \dots + a_d t^d \in \mathcal{J}$  mit  $\deg P = d$ . Dann gilt also insbesondere  $a_d \neq 0$  und wir können

$$M := \frac{1}{a_d} \cdot P = \frac{a_0}{a_d} + \dots + \frac{a_d}{a_d} t^d$$

setzen. Damit ist  $M$  ein normiertes Polynom und, wegen I2, ein Element von  $\mathcal{J}$ .

Sei nun  $P \in \mathcal{J}$  beliebig. Dann dividieren wir mit Rest durch  $M$  (Satz 5.7) und erhalten  $Q, R \in K[t]$  mit

$$P = Q \cdot M + R,$$

und  $\deg R < \deg M = d$ . Mit I1 und I2 gilt dann

$$R = P - Q \cdot M \in \mathcal{J}.$$

Da aber  $d$  der kleinste Grad eines Nicht-Nullpolynoms in  $\mathcal{J}$  ist, muss dann  $R$  selbst das Nullpolynom sein. Also gilt  $P = Q \cdot M$ . Damit ist die Existenz eines Minimalpolynoms gezeigt.

Es verbleibt die Eindeutigkeit: Sei dazu  $M'$  ein weiteres Minimalpolynom, also insbesondere

$$M = b_0 + \dots + b_{d-1} t^{d-1} + t^d, M' = c_0 + \dots + c_{d-1} t^{d-1} + t^d \in \mathcal{J}.$$

Wegen I1 ist dann  $M - M' \in \mathcal{J}$  mit  $\deg(M - M') \leq d - 1 < d$ . Nach Definition von  $d$  ist also  $M - M' = 0$  und damit  $M' = M$ .  $\square$

Wir schauen uns nun ein Beispiel an, welches verdeutlicht, wie sich das charakteristische Polynom eines Endomorphismus und das zugehörige Minimalpolynom unterscheiden können.



wobei die  $r_i, i = 1, \dots, k$  paarweise verschieden sind. Aus 1. folgt

$$M_A = \pm(t - \lambda_1)^{s_1} \cdot \dots \cdot (t - \lambda_k)^{s_k} \text{ mit } s_i \leq r_i, i = 1, \dots, k,$$

d.h. dass  $M_A$  keine weiteren Linearfaktoren besitzt. Wir zeigen nun, dass keiner fehlt, d.h.,  $s_i \geq 1, i = 1, \dots, k$  (und somit gilt dann  $P_F$  teilt  $M_F^n$ , weil  $r_i \leq n$  für alle  $i=1, \dots, k$ ). Angenommen es gibt ein  $i$  mit  $s_i = 0$ . Sei  $v_i \in \mathbb{C}^n$  Eigenvektor zu  $\lambda_i$ . Dann gilt für  $i \neq j$ :

$$\mathbb{C}^n \ni 0 \neq (F_A - \lambda_j \text{id}_V)(v_i) = (\lambda_i - \lambda_j)v_i \in \text{Eig}(F_A; \lambda_i)$$

und somit

$$M_A(A)(v_i) \neq 0 \text{ im Widerspruch zu } M_A(A) = 0 \in M(n \times n, K).$$

Definiere

$$Q := (t - \lambda_1)^{ns_1 - r_1} \cdot \dots \cdot (t - \lambda_k)^{ns_k - r_k} \in \mathbb{C}[t].$$

Wir erhalten  $M_A^n = \pm Q \cdot P_A$  und da  $M_A, P_A \in K[t]$  folgt  $Q \in K[t]$  (siehe Übung unten).  $\square$

**Übung 5.58.** Beweise folgende Aussage: Sei  $K'$  ein Körper und  $K$  ein Unterkörper von  $K'$ . Sind  $f, g \in K[t], f \neq 0$  und  $q \in K'[t]$  mit  $f = q \cdot g$ , so folgt  $q \in K[t]$ .

**Definition 5.59.** Ein Endomorphismus  $F \in \text{End}(V)$  heisst **nilpotent**, wenn  $F^k = 0$  für ein  $k \in \mathbb{N}$  gilt.

**Satz 5.60.** Sei  $V$  ein  $K$ -Vektorraum mit  $\dim V = n$  und  $F \in \text{End}(V)$ . Dann sind äquivalent:

1.  $F$  ist nilpotent.
2.  $F^d = 0$  für ein  $d$  mit  $1 \leq d \leq n$ .
3.  $P_F = \pm t^n$ .
4. Es gibt eine Basis  $\mathcal{B}$  von  $V$  mit

$$\mathcal{M}_{\mathcal{B}}(F) = \begin{pmatrix} 0 & & * \\ & \ddots & \\ 0 & & 0 \end{pmatrix}.$$

*Beweis.*  $1. \Rightarrow 2.$ : Die Bedingung 1. impliziert  $t^k \in \mathcal{J}_F$  und somit gilt mit Satz 5.55, 2., dass  $M_F = t^d$  für  $d \leq k$ . Da  $P_F \in \mathcal{J}_F$  und  $\deg P_F = n$  folgt ebenso  $d \leq n$ .

$2. \Rightarrow 3.$ : Mit  $M_F = t^d$  für  $1 \leq d \leq n$  und der Aussage von Satz 5.57, 2. folgt, dass  $P_F = \pm t^n$ .

3.  $\Rightarrow$  4.: Folgt mit Satz 5.40, wobei die Eigenwerte 0 von  $F$  in der Diagonale der Darstellungsmatrix auftreten.

4.  $\Rightarrow$  1.: Übung. □

## 5.8 Die Jordansche Normalform

Im Folgenden werden wir die Trigonalisierungstheorie etwas verfeinern. Für trigonalisierbare Endomorphismen  $F$  hatten wir in Satz 5.40 gesehen, dass wir eine Basis  $\mathcal{B}$  finden, sodass die resultierende Darstellungsmatrix  $\mathcal{M}_{\mathcal{B}}(F)$  eine obere Dreiecksform hat. Allerdings gibt dieser Satz keine weitere Auskunft darüber, wie die Einträge oberhalb der Diagonalen aussehen. Dies werden wir nun mit der Jordanschen Normalform weiter spezifizieren. Sei  $V$  ein  $K$ -Vektorraum mit  $\dim V = n \geq 1$  und sei  $F \in \text{End}(V)$  mit

$$P_F = \pm(t - \lambda_1)^{r_1} \cdot \dots \cdot (t - \lambda_k)^{r_k},$$

mit  $r_i, i = 1, \dots, k$  paarweise verschieden. Mit Satz 5.31 gilt, dass

$$V = \text{Eig}(F; \lambda_1) \oplus \dots \oplus \text{Eig}(F; \lambda_k).$$

genau, dann wenn

$$\dim \text{Eig}(F; \lambda_i) = \mu(P_F; \lambda_i) \text{ für alle } i = 1, \dots, k.$$

Für trigonalisierbare Endomorphismen, die nicht diagonalisierbar sind, ist diese Dimensionsgleichung für die Eigenräume nicht erfüllt, d.h., die Dimension von einigen Eigenräumen ist strikt kleiner als die algebraische Vielfachheit des Eigenwertes (Vielfachheit des Eigenwertes als Nullstelle des charakteristischen Polynoms). Wir werden nun den sogenannten **Hauptraum** einführen, der stets den Eigenraum enthält allerdings eine grössere Dimension haben kann.

**Definition 5.61.** Sei  $\lambda$  ein Eigenwert von  $F$  mit algebraischer Vielfachheit  $r \in \mathbb{N}$ . Dann ist der **Hauptraum von  $F$  zu  $\lambda$**  definiert als:

$$\text{Hau}(F; \lambda) := \text{Ker} (F - \lambda \text{id}_V)^r.$$

Um den Hauptraum besser zu verstehen, werden wir nun für einen Eigenwert  $\lambda$  von  $F$  die Potenzen der linearen Abbildung:

$$G := F - \lambda \text{id}_V$$

untersuchen. Da viele der Aussagen dieser Potenzen auch für allgemeine Endomorphismen gelten, lassen wir nun  $G \in \text{End}(V)$  beliebig.

**Lemma 5.62.** Sei  $G \in \text{End}(V)$ . Dann gilt für alle  $k \in \mathbb{N}$ :

$$\text{Ker } G^k \subset \text{Ker } G^{k+1} \text{ und } \text{Im } G^k \supset \text{Im } G^{k+1}.$$

Insbesondere gilt unter den Voraussetzungen von Definition 5.61:

$$\text{Eig}(F; \lambda) \subset \text{Hau}(F; \lambda).$$

*Beweis.* Sei  $v \in \text{Ker } G^k, w \in \text{Im } G^{k+1}$ , d.h.  $\exists u \in V$  mit  $G^{k+1}(u) = w \in V$  und wir setzen  $u' := G(u) \in V$ . Wir erhalten:

$$G^{k+1}(v) = G \circ G^k(v) = G(0) = 0, \quad G^k(u') = G^k(G(u)) = G^{k+1}(u) = w,$$

also  $v \in \text{Ker } G^{k+1}$  und  $w \in \text{Im } G^k$ . □

Mit Lemma 5.62 erhalten wir die folgenden Ketten:

$$\begin{aligned} \{0\} &\subset \text{Ker } G \subset \text{Ker } G^2 \subset \dots \subset \text{Ker } G^\ell \\ V &\supset \text{Im } G \supset \text{Im } G^2 \supset \dots \supset \text{Im } G^\ell. \end{aligned}$$

Für jedes  $\ell$  gilt mittels der Dimensionsformel:

$$\dim \text{Ker } G^\ell + \dim \text{Im } G^\ell = \dim V.$$

Mit  $\dim V = n$  können die beiden Ketten nicht endlos auf- bzw. absteigen. Dies führt auf das Lemma von Fitting.

**Lemma 5.63.** Zu  $G \in \text{End}(V)$  mit  $\dim V = n$  definiere

$$d := \min\{\ell \in \mathbb{N} \cup \{0\} : \text{Ker } G^\ell = \text{Ker } G^{\ell+1}\} \text{ und } r := \mu(P_G; 0).$$

Dann gilt:

1.  $d = \min\{\ell \in \mathbb{N} \cup \{0\} : \text{Im } G^{\ell+1} = \text{Im } G^\ell\}$
2.  $\text{Ker } G^{d+i} = \text{Ker } G^d, \text{Im } G^{d+i} = \text{Im } G^d$  für alle  $i \in \mathbb{N}$ .
3. Die Räume  $U := \text{Ker } G^d$  und  $W := \text{Im } G^d$  sind  $G$ -invariant.
4.  $(G|_U)^d = 0$  und  $G|_W : W \rightarrow W$  ist ein Isomorphismus.
5. Für das Minimalpolynom von  $G|_U$  gilt  $M_{G|_U} = t^d$ .
6.  $V = U \oplus W, \dim U = r \geq d, \dim W = n - r$ .

Insbesondere gibt es eine Basis  $\mathcal{B}$  von  $V$  mit

$$\mathcal{M}_{\mathcal{B}}(G) = \begin{pmatrix} N & 0 \\ 0 & C \end{pmatrix} \text{ mit } N^d = 0 \text{ und } C \in \text{GL}(n - r, K).$$

*Beweis.* Wir betrachten das Diagramm in Abb. 5.1. Mit Hilfe der Dimensionsformel er-

$$\begin{array}{ccc} \text{Ker } G^\ell \subset V & \xrightarrow{G^\ell} & \text{Im } G^\ell \\ & \cap & \cup \\ & & \text{Ker } G^{\ell+1} \subset V \xrightarrow{G^{\ell+1}} \text{Im } G^{\ell+1} \end{array}$$

Abbildung 5.1: Diagramm von Abbildungen und Inklusionen.

halten wir

$$\begin{aligned} \text{Im } G^{\ell+1} = \text{Im } G^\ell &\Leftrightarrow \dim \text{Im } G^{\ell+1} = \dim \text{Im } G^\ell \\ &\Leftrightarrow \dim \text{Ker } G^{\ell+1} = \dim \text{Ker } G^\ell \\ &\Leftrightarrow \text{Ker } G^{\ell+1} = \text{Ker } G^\ell \\ &\Leftrightarrow G|_{\text{Im } G^\ell} : \text{Im } G^\ell \rightarrow \text{Im } G^{\ell+1} \text{ ist ein Isomorphismus.} \end{aligned}$$

Mit

$$\text{Im } G^{\ell+1} = \text{Im } G^\ell \Leftrightarrow \text{Ker } G^{\ell+1} = \text{Ker } G^\ell$$

folgt direkt 1.  $G|_{\text{Im } G^\ell}$  Isomorphismus impliziert 2. Die Aussagen 3. und 4. sind klar.

Zu 5.: Mit 4. gilt, dass  $M_{G|_U} = t^k$  für ein  $k \leq d$ . Angenommen  $M_{G|_U} = t^k$  für ein  $k < d$ .

Wir erhalten mit  $(G|_U)^k = 0$ :

$$U \subset \text{Ker } (G|_U)^k \subset \text{Ker } G^k \subset \dots \subset \text{Ker } G^{d+1} \subset \text{Ker } G^d = U,$$

und somit  $\text{Ker } G^k = \text{Ker } G^{k+1}$  im Widerspruch zur Definition von  $d$ .

Zu 6.: Wir zeigen zunächst  $V = U \oplus W$ . Da nach Dimensionsformel gilt  $\dim V = \dim \text{Ker } G^d + \dim \text{Im } G^d = \dim U + \dim W$ , genügt es dafür zu zeigen, dass  $U \cap W = \{0\}$ . Sei dazu  $v \in U \cap W$ , also  $v \in \text{Ker } G^d$  und  $v \in \text{Im } G^d$  also  $v = G^d(w)$  für ein  $w \in V$ . Mit  $v \in \text{Ker } G^d$  gilt  $0 = G^d(v) = G^d(G^d(w)) = G^{2d}(w)$ . Somit folgt  $w \in \text{Ker } G^{2d} = \text{Ker } G^d$  (siehe 2.), also  $v = G^d(w) = 0$  und die erste Behauptung ist bewiesen.

Per Definition von  $d$  gilt  $\text{Ker } G^k \subsetneq \text{Ker } G^{k+1}$  für alle  $k < d$  und da  $\text{Ker } G^k$  jeweils ein Untervektorraum ist, muss für die Dimensionen gelten:

$$\dim \text{Ker } G^1 < \dim \text{Ker } G^2 < \dots < \dim \text{Ker } G^d,$$

also  $\dim U = \dim \text{Ker } G^d \geq d$ . Für  $r = \mu(P_G; 0)$  erhalten wir  $P_G(t) = (t-0)^r \cdot Q(t)$  mit  $Q(0) \neq 0$  und mit  $V = U \oplus W$  gilt:

$$(t-0)^r \cdot Q(t) = P_G(t) = P_{G|_U}(t) \cdot P_{G|_W}(t).$$

Da  $G|_U$  nilpotent ist (siehe 4.), gilt mit Satz 5.60, dass  $P_{G|_U}(t) = t^m$  für  $m \geq d$  sowie

$m = \dim U$ . Für den Isomorphismus  $G|_W$  (siehe 4.) gilt  $P_{G|_W}(0) \neq 0$  (benutze dass eine darstellende Matrix von  $G|_W$  vollen Rang hat) und mit der Übung 5.64 folgt  $m = r$ .

**Übung 5.64.** Zeige folgende Aussage: Sei  $f, Q, Q' \in K[t]$  mit

$$f(t) = t^r \cdot Q(t) = t^m \cdot Q'(t) \text{ mit } Q(0) \neq 0, \text{ und } Q'(0) \neq 0.$$

Dann gilt  $m = r$ .

Die letzte Aussage folgt mittels der Zerlegung  $V = U \oplus W$  und den Eigenschaften von  $G$  aus 4. □

Nun sind wir so weit, den Satz über die Hauptraumzerlegung zu beweisen.

**Satz 5.65.** Sei  $F \in \text{End}(V)$  mit  $\dim V = n$  und sei

$$P_F = \pm(t - \lambda_1)^{r_1} \cdot \dots \cdot (t - \lambda_k)^{r_k}$$

mit paarweise verschiedenen  $\lambda_1, \dots, \lambda_k \in K$ . Weiterhin sei  $V_i := \text{Hau}(F; \lambda_i) \subset V$  für jedes  $\lambda_i$  der zugehörige Hauptraum von  $F$ . Dann gilt:

1.  $F(V_i) \subset V_i$  und  $\dim V_i = r_i$  für alle  $i = 1, \dots, k$ .
2.  $V = V_1 \oplus \dots \oplus V_k$ .
3.  $F$  hat eine Zerlegung  $F = F_D + F_N$  mit
  - (a)  $F_D$  ist diagonalisierbar.
  - (b)  $F_N$  ist nilpotent.
  - (c)  $F_D \circ F_N = F_N \circ F_D$ .

*Beweis.* Wir werden den Satz durch iteratives Anwenden des Lemmas von Fitting (Lemma 5.63) beweisen. Wir wählen  $G_1 := F - \lambda_1 \cdot \text{id}_V$  und erhalten

$$P_{G_1}(t - \lambda_1) = P_F(t) \text{ und } r_1 = \mu(P_{G_1}; 0) = \mu(P_F; \lambda_1).$$

Wir können also das Lemma von Fitting (Lemma 5.63) auf  $G_1$  anwenden und erhalten

$$V = \text{Ker } G_1^d \oplus W = \text{Ker } G_1^{r_1} \oplus W = \text{Hau}(F; \lambda_1) \oplus W \quad (5.6)$$

und beide Summanden sind wegen  $F = G_1 + \lambda_1 \text{id}_V$  auch  $F$ -invariant. Ferner erhalten wir aus Lemma 5.63 eine Basis  $\mathcal{B}$  von  $V$  bestehend aus einer Basis  $\mathcal{B}_1$  von  $\text{Ker } G_1^d = V_1$  und einer Basis  $\mathcal{B}'$  von  $W$ . Weiterhin sei  $N_1$  der nilpotente Block aus der zugehörigen Darstellungsmatrix  $\mathcal{M}_{\mathcal{B}}(G)$ .

Aus der  $F$ -Invarianz der beiden Summanden in (5.6) erhalten wir ferner  $P_F = P_{F|_{V_1}} \cdot P_{F|_W}$ . Zusammen mit der Beobachtung, dass  $P_{F|_{V_1}} = \pm(t - \lambda_1)^{r_1}$ , folgt daraus für die Abbildung

$F|_W : W \rightarrow W$ , dass

$$P_{F|_W} = \pm(t - \lambda_2)^{r_2} \cdot \dots \cdot (t - \lambda_k)^{r_k}.$$

Somit können wir nun wieder das Lemma von Fitting auf  $G_2 := F|_W - \lambda_2 \cdot \text{id}_W$  anwenden und dieses Vorgehen iterieren, woraus die Aussagen 1 und 2. folgen. Wir setzen nun

$$D := \begin{pmatrix} \lambda_1 E_{r_1} & & 0 \\ & \ddots & \\ 0 & & \lambda_k E_{r_k} \end{pmatrix} \text{ sowie } N := \begin{pmatrix} N_1 & & 0 \\ & \ddots & \\ 0 & & N_k \end{pmatrix},$$

wobei die Matrizen  $N_1, \dots, N_k$  die (nilpotenten) Darstellungsmatrizen der Abbildungen  $G_i|_{V_i}, i = 1, \dots, n$  zu den Basen  $\mathcal{B}_i$  von  $V_i$  sind. Wir können nun nachrechnen, dass  $D \cdot N = N \cdot D$  gilt. Sei nun  $\mathcal{B}$  die aus den Basen  $\mathcal{B}_i$  zusammengesetzte Basis von  $V$  und  $F_D, F_N \in \text{End}(V)$  die eindeutigen Endomorphismen mit  $\mathcal{M}_{\mathcal{B}}(F_D) = D$  und  $\mathcal{M}_{\mathcal{B}}(F_N) = N$ . Dann folgt (mit der Additivität von  $\mathcal{M}_{\mathcal{B}}$ ), dass  $F = F_D + F_N$ , sowie (mit der Multiplikativität von  $\mathcal{M}_{\mathcal{B}}$ ) aus  $D \cdot N = N \cdot D$ , dass  $F_D \circ F_N = F_N \circ F_D$ .  $\square$

Nun werden wir uns mit Normalformen von nilpotenten Endomorphismen beschäftigen. Die gewonnenen Einsichten werden uns bei der Berechnung der allgemeinen Jordanschen Normalform helfen.

**Definition 5.66.** Eine **Jordanmatrix** (oder **Jordanblock**) der Größe  $k$  zum Eigenwert  $0$ , ist eine Matrix der Gestalt:

$$J_k(0) := \begin{pmatrix} 0 & 1 & & 0 \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ 0 & & & 0 \end{pmatrix}.$$

**Lemma 5.67.** Die Matrixabbildung  $F_{J_k(0)}$  ist nilpotent mit  $F_{J_k(0)}^k = 0$ . Es gilt

$$k = \min\{r \in \mathbb{N} \mid F_{J_k(0)}^r = 0\}.$$

*Beweis.* Siehe Bsp. 5.56.  $\square$

**Satz 5.68.** Sei  $G \in \text{End}(V)$  nilpotent mit  $\dim V = n$  und  $d := \min\{r \in \mathbb{N} \mid G^r = 0\}$ . Dann gibt es eindeutig bestimmte Zahlen  $s_1, \dots, s_d \in \mathbb{N} \cup \{0\}$  mit

$$d \cdot s_d + (d - 1)s_{d-1} + \dots + s_1 = n = \dim V$$



$$3. G(W_d) \subset U_{d-1}.$$

$$4. G(W_d) \cap U_{d-2} = \{0\}.$$

Beweis von 3.: Wende  $G$  jeweils links und rechts von Gleichung 5.7 an:

$$U_{d-1} \stackrel{1.}{\supset} G(U_d) = G(U_{d-1}) + G(W_d)$$

also  $G(W_d) \subset U_{d-1}$ .

Beweis von 4.: Sei  $w \in U_{d-2} \cap G(W_d)$ . Somit existiert  $v \in W_d$  mit  $G(v) = w$ . Gleichzeitig gilt mit  $w \in U_{d-2}$  die Eigenschaft

$$0 = G^{d-2}(w) = G^{d-2}(G(v)) = G^{d-1}(v)$$

woraus  $v \in U_{d-1}$  folgt. Da allerdings wegen  $U_{d-1} \oplus W_d$  die Beziehung  $U_{d-1} \cap W_d = \{0\}$  gilt, folgt  $v = 0$  und somit  $w = 0$ .

Mit den beiden Eigenschaften  $G(W_d) \subset U_{d-1}$  und  $G(W_d) \cap U_{d-2} = \{0\}$  können wir  $U_{d-1}$  weiter zerlegen in

$$U_{d-1} = U_{d-2} \oplus W_{d-1} \text{ mit } G(W_d) \subset W_{d-1}.$$

Wir iterieren dieses Verfahren und erhalten folgende schematische Darstellung, wobei die Pfeile jeweils die Anwendung von  $G$  illustrieren.

$$\begin{array}{ccccccc}
 V = \ker G^d & & & & & & \\
 \downarrow & & & & & & \\
 V = \ker G^{d-1} \oplus W_d & & & & & & \\
 \downarrow & \downarrow & & & & & \\
 V = \ker G^{d-2} \oplus W_{d-1} \oplus W_d & & & & & & \\
 \downarrow & \downarrow & \downarrow & & & & \\
 \vdots & \vdots & \vdots & & & & \\
 \downarrow & \downarrow & \downarrow & \ddots & & & \\
 V = \ker G^1 \oplus W_2 \oplus W_3 \oplus \cdots \oplus W_d & & & & & & \\
 \downarrow & \downarrow & \downarrow & & & \downarrow & \\
 V = \ker G^0 \oplus W_1 \oplus W_2 \oplus \cdots \oplus W_{d-1} \oplus W_d & & & & & & 
 \end{array}$$

Mit  $\ker G^0 = \{0\}$  folgt die gewünschte Zerlegung

$$V = W_1 \oplus W_2 \oplus \cdots \oplus W_{d-1} \oplus W_d.$$

Mit Eigenschaft 2. folgt, dass die Abbildungen

$$G : W_{\ell+1} \rightarrow G(W_\ell), \ell = d-1, \dots, 1$$

allesamt injektiv sind und somit können wir die Bilder einer Basis von  $W_\ell$  zu einer Basis von  $W_{\ell-1}$  ergänzen, vgl. Satz 3.12, 4. Wir können dieses Schema folgendermaßen illustrieren:

$$\begin{array}{cccccccc}
 w_1^{(d)}, & \dots, & w_{s_d}^{(d)}, & & & & & \\
 G(w_1^{(d)}), & \dots, & G(w_{s_d}^{(d)}), & w_1^{(d-1)}, & \dots, & w_{s_{d-1}}^{(d-1)}, & & \\
 \vdots & & \vdots & \vdots & & \vdots & & \\
 G^{d-1}(w_1^{(d)}), & \dots, & G^{d-1}(w_{s_d}^{(d)}), & G^{d-2}(w_1^{(d-1)}), & \dots, & G^{d-2}(w_{s_{d-1}}^{(d-1)}), & \dots, & w_1^{(1)}, \dots, w_{s_1}^{(1)}
 \end{array} \tag{5.8}$$

Hierbei entspricht die erste Zeile einer Basis des Untervektorraums  $W_d$ , die zweite einer Basis des Untervektorraums  $W_{d-1}$  und die letzte Zeile einer Basis von  $W_1 = \text{Ker } G$ . Die gesuchte Basis  $\mathcal{B}$  hat nun die Form

$$\mathcal{B} = (G^{d-1}(w_1^{(d)}), G^{d-2}(w_1^{(d)}), \dots, w_1^{(d)}, G^{d-1}(w_2^{(d)}), \dots, w_2^{(d)}, \dots, w_1^{(1)}, \dots, w_{s_1}^{(1)}),$$

wobei wir die Basis  $\mathcal{B}$  so aufgebaut haben, dass wir das Schema 5.8 von links nach rechts (im Sinne der Spalten) und von unten nach oben (im Sinne der Zeilen) durchlaufen. Wir können nun recht einfach sehen, dass die ersten  $d$  Elemente der Basis  $\mathcal{B}$  (welche der ersten Spalte von Schema 5.8 entsprechen) den ersten Jordan Block  $J_d$  in der darstellenden Matrix  $\mathcal{M}_{\mathcal{B}}(G)$  erzeugen usw.

Nun zeigen wir noch die Eigenschaften der Zahlen  $s_\ell, \ell = 1, \dots, d$ . Es gilt  $\text{Ker } G^{d-1} \subsetneq \text{Ker } G^d$ , aufgrund der Definition von  $d$  (als Minimum). Somit folgt  $s_d = \dim W_d \geq 1$ . Sei  $\mathcal{A}$  irgendeine Basis von  $V$ , so dass die darstellende Matrix  $\mathcal{M}_{\mathcal{A}}(G)$  eine Jordanform hat. Wir werden zeigen, dass die Anzahl  $s_\ell$  der jeweiligen Blöcke allein durch  $G$  bestimmt sind. Dies folgt aus der Überlegung, dass die Dimensionen von  $\text{Ker } G^\ell$  unabhängig von der Wahl der Basis sind und in folgender Beziehung zueinander stehen:

$$\begin{aligned}
 s_d &= \dim \ker G^d - \dim \ker G^{d-1} \\
 s_d + s_{d-1} &= \dim \ker G^{d-1} - \dim \ker G^{d-2} \\
 &\vdots \\
 s_d + s_{d-1} + \dots + s_2 &= \dim \ker G^2 - \dim \ker G \\
 s_d + s_{d-1} + \dots + s_2 + s_1 &= \dim \ker G
 \end{aligned}$$

Wir können von oben nach unten die einzelnen Zahlen einsetzen und somit sind diese eindeutig festgelegt. □

**Beispiel 5.69.** Hier kommt ein Beispiel. Wir betrachten die reelle Matrix

$$B = \begin{pmatrix} 0 & 1 & 3 \\ 0 & 0 & 2 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{mit} \quad B^2 = \begin{pmatrix} 0 & 0 & 2 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \quad \text{und} \quad B^3 = 0.$$

Hier ist also  $d = 3$ , und für die Kerne der Potenzen  $B^\ell$  gilt

$$\{0\} = \text{Ker } B^0 \subset \text{Ker } B = \text{span}(e^1) \subset \text{Ker } B^2 = \text{span}(e^1, e^2) \subset \text{Ker } B^3 = \mathbb{R}^3.$$

Aus der Bedingung

$$\mathbb{R}^3 = \text{Ker } B^2 \oplus W_3$$

folgt, dass wir  $W_3 = \text{span}(e_3)$  wählen können. Somit ist  $s_3 = 1$ . Aus

$$\mathbb{R}^3 = \text{Ker } B \oplus W_2 \oplus W_3$$

folgt dann  $\dim W_2 = 1$ , also  $s_2 = 0$ , und  $B \cdot e^3 = (3, 2, 0)^T$  ist der richtige Basisvektor von  $W_2$ . Wir erhalten

$$\mathbb{R}^3 = W_1 \oplus W_2 \oplus W_3.$$

Der gesuchte Basisvektor von  $W_1$  ist  $B \cdot (3, 2, 0)^T = B^2 \cdot e^3 = 2e^1$ , somit ist auch  $s_1 = 0$ . Werden die gefundenen Basisvektoren in der richtigen Reihenfolge als Spalten in eine Matrix eingetragen, so erhalten wir

$$T^{-1} = \begin{pmatrix} 2 & 3 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad \text{und daraus} \quad T = \frac{1}{4} \begin{pmatrix} 2 & -3 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{pmatrix}.$$

Das Ergebnis ist die Jordan-Matrix

$$TBT^{-1} = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}.$$

In dem wir nun die Hauptraumzerlegung (Satz 5.65) und den Satz 5.68 über Jordanmatrizen für nilpotente lineare Abbildungen kombinieren, erhalten wir den allgemeinen Satz über Jordandarstellungen.

**Satz 5.70.** Sei  $F \in \text{End}(V)$  derart, dass das charakteristische Polynom in Linearfaktoren zerfällt, also

$$P_F = \pm(t - \lambda_1)^{r_1} \cdot \dots \cdot (t - \lambda_k)^{r_k}$$





154 | Kapitel 5. Eigenwerte und Normalformen von Endomorphismen  
gilt dann

$$\mathcal{M}_{\mathcal{B}_i}(F_i) = \left( \begin{array}{cccc} J_{d_i}(\lambda_i) & & & \\ & \ddots & & \\ & & J_{d_i}(\lambda_i) & \\ & & & \ddots \\ & & & & J_1(\lambda_i) \\ & & & & & \ddots \\ & & & & & & J_1(\lambda_i) \end{array} \right) \left. \begin{array}{l} s_{d_i}^{(i)}\text{-mal} \\ \vdots \\ s_1^{(i)}\text{-mal} \end{array} \right\}$$

wobei  $J_k(\lambda)$  einen Jordan Block der Größe  $k \times k$  mit Eigenwert  $\lambda$  bezeichnet.

Zur Eindeutigkeit der Zahlen  $s_i$  gehen wir wie bei der Eindeutigkeit der entsprechenden Zahlen bei nilpotenten Endomorphismen vor. Sei  $\mathcal{A}$  irgendeine Basis von  $V$ , so dass die darstellende Matrix  $\mathcal{M}_{\mathcal{A}}(F)$  eine Jordanform hat. Man zeigt dann, dass die Anzahlen  $s_\ell^{(i)}$  der jeweiligen Blöcke allein durch  $F$  bestimmt sind. Dies folgt wiederum aus der Überlegung, dass die Dimensionen von  $\text{Ker } G_i^\ell$  unabhängig von der Wahl der Basis sind (Übung).  $\square$

## Kapitel 6

# Euklidische und unitäre Vektorräume

In diesem Kapitel werden wir uns mit Vektorräumen über den Körpern  $\mathbb{R}$  (Euklidische Vektorräume) oder  $\mathbb{C}$  (unitäre Vektorräume), die jeweils ein **Skalarprodukt** besitzen, befassen. Skalarprodukte sind Abbildungen der Form  $s : V \times W \rightarrow U$  für Vektorräume  $V, W, U$  und treten als Spezialfall von **Bilinearformen** auf.

### 6.1 Bilinearformen

**Definition 6.1 (Linearform, Dualraum).** Sei  $V$  ein  $K$ -Vektorraum. Eine lineare Abbildung  $F : V \rightarrow K$  heißt **Linearform**. Der Vektorraum aller Linearformen auf  $V$

$$V^* := \text{Hom}(V, K)$$

wird **Dualraum** von  $V$  genannt.

**Beispiel 6.2.** Für  $x = (x_1, \dots, x_n) \in K^n$  ist

$$\phi_x : K^n \rightarrow K, y \mapsto x_1 y_1 + \dots + x_n y_n$$

eine Linearform. Mit Satz 3.27 folgt, dass zu jeder Linearform  $\phi : K^n \rightarrow K$  genau ein  $x \in K$  existiert, mit  $\phi = \phi_x$ . Ferner folgt mit Satz 3.27, dass die Abbildung

$$K^n \rightarrow V^*, x \mapsto \phi_x$$

ein Vektorraumisomorphismus ist.

Nun kommen wir zu allgemeinen Bilinearformen.

**Definition 6.3.** Seien  $V, W$  jeweils  $K$ -Vektorräume. Eine **skalarwertige** Abbildung

$$s : V \times W \rightarrow K$$

heißt **Bilinearform** oder **bilinear**, falls für alle  $v, v' \in V, w, w' \in W$  und  $\lambda \in K$  gilt:

1.  $s(v + v', w) = s(v, w) + s(v', w)$
2.  $s(v, w + w') = s(v, w) + s(v, w')$
3.  $s(\lambda v, w) = \lambda s(v, w) = s(v, \lambda w)$ .

Wir geben einige Beispiele.

**Beispiel 6.4.** • Die Abbildung  $\mu : K \times K \rightarrow K, (x, y) \mapsto x \cdot y$  ist bilinear.

- Sei  $V$  ein  $K$ -Vektorraum. Dann ist  $\alpha : V^* \times V \rightarrow K, (F, v) \mapsto F(v)$  bilinear.

Wie für lineare Abbildungen erhalten wir auch für bilineare Abbildungen Darstellungssätze mit Hilfe von geeigneten Matrizen.

**Definition 6.5.** Sei  $A = (a_{ij}) \in M(m \times n, K)$ . Definiere die induzierte **bilineare Matrix-Abbildung** durch

$$s_A : K^m \times K^n \rightarrow K, s_A(x, y) := \sum_{i=1}^m \sum_{j=1}^n a_{ij} x_i y_j.$$

(Übung:  $s_A$  ist in der Tat eine Bilinearform.) In Matrixproduktschreibweise erhalten wir

$$s_A(x, y) = x^T A y.$$

**Satz 6.6.** Seien  $V, W$  jeweils  $K$ -Vektorräume mit Basen  $\mathcal{A} = (v_1, \dots, v_m)$  und  $\mathcal{B} = (w_1, \dots, w_n)$ . Ferner sei  $s : V \times W \rightarrow K$  eine Bilinearform für die wir mit Hilfe von  $\mathcal{A}$  und  $\mathcal{B}$  folgende Matrix definieren:

$$\mathcal{M}_{\mathcal{B}}^{\mathcal{A}}(s) := A := (a_{ij})_{\substack{i=1, \dots, m \\ j=1, \dots, n}} \in M(m \times n, K).$$

Sei  $v \in V$  mit zugehörigen Koordinatenvektoren  $x$  bzgl.  $\mathcal{A}$  und  $w \in W$  mit zugehörigen Koordinatenvektoren  $y$  bzgl.  $\mathcal{B}$  (d.h.  $x = \Phi_{\mathcal{A}}^{-1}(v)$  sowie  $y = \Phi_{\mathcal{B}}^{-1}(w)$ ). Dann gilt:

$$s(v, w) = s_A(x, y) = x^T A y.$$

Die Matrix  $\mathcal{M}_{\mathcal{B}}^{\mathcal{A}}(s)$  heißt **darstellende Matrix** von  $s$  bzgl.  $\mathcal{A}$  und  $\mathcal{B}$ .

*Beweis.*

$$s(v, w) = s\left(\sum_{i=1}^m x_i v_i, \sum_{j=1}^n y_j w_j\right) = \sum_{i=1}^m \sum_{j=1}^n s(v_i, w_j) x_i y_j = \sum_{i=1}^m \sum_{j=1}^n a_{ij} x_i y_j,$$

wobei der vorletzte Schritt jeweils die Linearität in beiden Komponenten von  $s$  ausnutzt. □

Der folgende Satz ist zentral für die Beschreibung von Bilinearformen nach einem Basiswechsel.

**Satz 6.7** (Änderung der Darstellungsmatrix einer Bilinearform bei Basiswechsel). Seien  $V, W$  jeweils  $K$ -Vektorräume,  $\mathcal{A} = (v_1, \dots, v_m)$  und  $\mathcal{A}' = (v'_1, \dots, v'_m)$  Basen von  $V$ ,  $\mathcal{B} = (w_1, \dots, w_n)$  und  $\mathcal{B}' = (w'_1, \dots, w'_n)$  Basen von  $W$ . Sei  $S = T_{\mathcal{A}'}^{\mathcal{A}}$  die Transformationsmatrix von  $\mathcal{A}'$  nach  $\mathcal{A}$  und  $T := T_{\mathcal{B}}^{\mathcal{B}'}$  die Transformationsmatrix von  $\mathcal{B}'$  nach  $\mathcal{B}$ . Sei  $s : V \times W \rightarrow K$  eine Bilinearform,  $\mathcal{M}_{\mathcal{B}}^{\mathcal{A}}(s) = A = (a_{ij})$  die darstellende Matrix von  $s$  bezüglich  $\mathcal{A}$  und  $\mathcal{B}$  sowie  $\mathcal{M}_{\mathcal{B}'}^{\mathcal{A}'}(s) = A' = (a'_{ij})$  die darstellende Matrix von  $s$  bezüglich  $\mathcal{A}'$  und  $\mathcal{B}'$ . Dann gilt:

$$A' = S^T A T.$$

*Beweis.* Es gilt

$$v'_i = \sum_{k=1}^m s_{ki} v_k, \quad w'_j = \sum_{l=1}^n t_{lj} w_l.$$

Weiter ist

$$a'_{ij} = s(v'_i, w'_j) = \sum_{k=1}^m \sum_{l=1}^n s_{ki} t_{lj} s(v_k, w_l) = \sum_{k=1}^m s_{ki} \left( \sum_{l=1}^n a_{kl} t_{lj} \right).$$

Daher ist  $a'_{ij}$  der  $i, j$ -te Eintrag der Matrix  $S^T A T$  und somit gilt  $A' = S^T A T$ .  $\square$

**Beispiel 6.8.** Sei  $V = W = \mathbb{R}^2$  mit Standardbasis  $\mathcal{A} = \mathcal{B} = (e^1, e^2)$ . Wähle neue Basen  $\mathcal{A}' = (v'_1, v'_2)$  und  $\mathcal{B}' = (w'_1, w'_2)$  mit Transformationsmatrizen  $S = T_{\mathcal{A}'}^{\mathcal{A}}$  (für  $\mathcal{A}' \rightarrow \mathcal{A}$ ) und  $T := T_{\mathcal{B}}^{\mathcal{B}'}$  (für  $\mathcal{B}' \rightarrow \mathcal{B}$ ). Für die Bilinearform  $s(x, y) = x_1 y_1 + x_2 y_2$  ist  $\mathcal{M}_{\mathcal{B}}^{\mathcal{A}}(s) = A = e^2$  die Einheitsmatrix. Die neue darstellende Matrix ist  $\mathcal{M}_{\mathcal{B}'}^{\mathcal{A}'}(s) = A' = S^T A T = S^T T$ .

**Definition 6.9** (Symmetrische Bilinearform, symmetrische Matrix). Sei  $V$  ein  $K$ -Vektorraum.

1. Eine Bilinearform  $s : V \times V \rightarrow K$  heißt **Bilinearform auf  $V$** . Für die Darstellungsmatrix von  $s$  bzgl. einer Basis  $\mathcal{A}$  von  $V$  schreiben wir  $\mathcal{M}_{\mathcal{A}} := \mathcal{M}_{\mathcal{A}}^{\mathcal{A}} \in M(n \times n, K)$ .
2. Eine Bilinearform  $s : V \times V \rightarrow K$  heißt **symmetrisch**, falls für alle  $v, v' \in V$  die Beziehung  $s(v, v') = s(v', v)$  gilt.
3. Eine Matrix  $A \in M(n \times n, K)$  heißt **symmetrisch**, falls  $A^T = A$ .

**Satz 6.10.** Sei  $V$  ein endlich-dimensionaler  $K$ -Vektorraum,  $s$  eine Bilinearform auf  $V$ . Dann ist  $s$  genau dann symmetrisch, wenn die Darstellungsmatrix von  $s$  bzgl. einer beliebigen Basis von  $V$  symmetrisch ist.

*Beweis.* Sei  $\mathcal{A} = (v_1, \dots, v_n)$  eine Basis von  $V$  und  $A = (a_{ij})$  die darstellende Matrix von  $s$  bzgl.  $\mathcal{A}$ .

⇒: Weil  $s$  symmetrisch ist, gilt für alle  $v = \sum_{i=1}^n x_i v_i$  und  $w = \sum_{j=1}^n y_j v_j$

$$s(v, w) = \sum_{i,j=1}^n a_{ij} x_i y_j, \quad s(w, v) = \sum_{i,j=1}^n a_{ij} y_i x_j.$$

Für  $v = v_i$  und  $w = v_j$  folgt deshalb  $a_{ij} = a_{ji}$ , also  $A^T = A$ .

⇐: Der Beweis geht genauso. □

**Definition 6.11 (Quadratische Form).** Sei  $K$  ein Körper mit  $1 + 1 \neq 0$  und sei  $V$  ein  $K$ -Vektorraum. Eine Abbildung  $Q : V \rightarrow K$  heißt **quadratische Form**, falls eine symmetrische Bilinearform  $s$  auf  $V$  existiert mit  $Q(v) = \frac{1}{2}s(v, v)$  für alle  $v \in V$ .

**Lemma 6.12.** Eine symmetrische Bilinearform  $s$  ist durch die quadratische Form  $Q$  eindeutig bestimmt. Genauer gilt

$$s(v, w) = Q(v + w) - Q(v) - Q(w) \text{ für alle } v, w \in V.$$

*Beweis.* Seien  $v, w \in V$ . Es gilt:

$$\begin{aligned} Q(v + w) - Q(v) - Q(w) &= \frac{1}{2}s(v + w, v + w) - \frac{1}{2}s(v, v) - \frac{1}{2}s(w, w) \\ &= \frac{1}{2}(s(v, v + w) + s(w, v + w) - s(v, v) - s(w, w)) \\ &= \frac{1}{2}(s(v, v) + s(v, w) + s(w, v) + s(w, w) - s(v, v) - s(w, w)) \\ &= s(v, w). \end{aligned}$$

□

**Beispiel 6.13.** 1. Die Abbildung  $Q : K^n \rightarrow K$ ,  $Q(x) = x_1^2 + \dots + x_n^2$  ist eine quadratische Form.

2. Sei  $V = C([0, 1])$  der  $\mathbb{R}$ -Vektorraum aller reellen, stetigen Funktionen auf  $[0, 1]$ . Die Abbildung  $Q : V \rightarrow \mathbb{R}$ ,  $Q(f) = \int_0^1 f^2(x) dx$ , ist eine quadratische Form auf  $V$ .

## 6.2 Reelle Skalarprodukte

$V$  wird in diesem Abschnitt stets als ein  $\mathbb{R}$ -Vektorraum aufgefasst.

**Definition 6.14 (Skalarprodukt, Euklidischer Raum).** Sei  $V$  ein  $\mathbb{R}$ -Vektorraum.

1. Eine symmetrische Bilinearform  $s$  auf  $V$  heißt **positiv (negativ) semidefinit**, falls  $s(v, v) \geq 0$  (bzw.  $s(v, v) \leq 0$ ) für alle  $v \in V$ . Gilt zusätzlich  $s(v, v) \neq 0$  für alle  $v \in V \setminus \{0\}$ , so heißt  $s$  **positiv (negativ) definit**. Ist  $s$  weder positiv noch negativ semidefinit, so heißt  $s$  **indefinit**.

2. Eine symmetrische positiv definite Bilinearform  $s$  auf  $V$  heißt ein (reelles) **Skalarprodukt** auf  $V$ . Das Paar  $(V, s)$  heißt **euklidischer Raum**. In diesem Fall wird  $s(v, w)$  oft als  $\langle v, w \rangle$  geschrieben.

**Übung 6.15.** Zeige die folgenden Aussagen:

1. Es gilt

$$s \text{ positiv (semi-)definit} \iff -s \text{ negativ (semi-)definit.}$$

2. Sei  $V = \mathbb{R}^n$  und sei  $\langle x, y \rangle = \sum_{i=1}^n x_i y_i = x^T y$  für  $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}, y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} \in \mathbb{R}^n$ .

Dann ist  $(\mathbb{R}^n, \langle \cdot, \cdot \rangle)$  ein euklidischer Raum.  $\langle \cdot, \cdot \rangle$  heißt das **Standardskalarprodukt** auf  $\mathbb{R}^n$ .

3. Sei  $V = C([0, 1])$  der  $\mathbb{R}$ -Vektorraum aller reellen, stetigen Funktionen auf  $[0, 1]$ . Definiere  $\langle f, g \rangle = \int_0^1 f(t)g(t) dt$ . Dann ist  $(C([0, 1]), \langle \cdot, \cdot \rangle)$  ein euklidischer Raum.
4.  $s: \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}, s(x, y) = x_1 y_1$  für  $x = (x_1, x_2), y = (y_1, y_2) \in \mathbb{R}^2$  ist eine positiv semidefinite Bilinearform, die nicht positiv definit ist.
5.  $s: \mathbb{R}^2 \times \mathbb{R}^2 \rightarrow \mathbb{R}, s(x, y) = x_1 y_1 - x_2 y_2$  ist eine indefinite symmetrische Bilinearform.

**Definition 6.16 (Norm).** Sei  $V$  ein  $\mathbb{R}$ -Vektorraum. Eine Abbildung  $\| \cdot \|: V \rightarrow \mathbb{R}$  heißt eine **Norm** auf  $V$ , wenn für alle  $v, w \in V$  und  $\lambda \in \mathbb{R}$  gilt:

1.  $\|v\| \geq 0$  (Nichtnegativität)
2.  $\|v\| = 0 \iff v = 0$  (Definitheit)
3.  $\|\lambda v\| = |\lambda| \|v\|$  (Homogenität)
4.  $\|v + w\| \leq \|v\| + \|w\|$  (Dreiecksungleichung)

$(V, \| \cdot \|)$  heißt dann ein **normierter Vektorraum**.

**Übung 6.17.** 1. Sei  $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ .

- (a) Durch  $\|x\|_2 := \sqrt{x_1^2 + \dots + x_n^2}$  wird  $\|x\|_2$  zu einer Norm auf  $\mathbb{R}^n$ .
- (b) Durch  $\|x\|_1 := |x_1| + \dots + |x_n|$  wird  $\|x\|_1$  zu einer Norm auf  $\mathbb{R}^n$ .
- (c) Durch  $\|x\|_\infty := \max\{|x_1|, \dots, |x_n|\}$  wird  $\|x\|_\infty$  zu einer Norm auf  $\mathbb{R}^n$ .

2. Für  $f \in C([0, 1])$  definiere  $\|f\| := \int_0^1 |f(t)| dt$ . Dann ist  $\| \cdot \|$  eine Norm auf  $C([0, 1])$ .

**Definition 6.18.** Um Abstände zwischen zwei Punkten im  $V$  zu messen definieren wir die norm-induzierte **Abstandsfunktion**

$$d : V \times V \rightarrow \mathbb{R}, d(x, y) := \|x - y\|.$$

**Übung 6.19.** Sei  $(V, \|\cdot\|)$  ein normierter Vektorraum und  $d : V \times V \rightarrow \mathbb{R}$  die induzierte Norm. Zeige folgende Aussagen für beliebige  $u, v, w \in V$ :

1.  $d(v, w) \geq 0$
2.  $d(v, w) = 0 \Leftrightarrow v = w$ .
3.  $d(v, w) = d(w, v)$
4.  $d(u, w) \leq d(u, v) + d(v, w)$ .

**Satz 6.20.** Sei  $(V, \langle \cdot, \cdot \rangle)$  ein euklidischer Raum. Dann ist die Abbildung

$$\|\cdot\| : V \rightarrow \mathbb{R}, \|v\| = \sqrt{\langle v, v \rangle}$$

eine Norm auf  $V$ . Sie heißt auch die **euklidische Norm** von  $V$ .

Zum Beweis von Satz 6.20 benötigen wir die **Cauchy-Schwarzsche Ungleichung**:

**Lemma 6.21.** Sei  $(V, \langle \cdot, \cdot \rangle)$  ein euklidischer Raum und sei  $\|v\| := \sqrt{\langle v, v \rangle}$  für  $v \in V$ . Dann gilt für alle  $v, w \in V$ :

$$|\langle v, w \rangle| \leq \|v\| \|w\|.$$

*Beweis.* Seien  $v, w \in V$ . O.E. sei  $w \neq 0$ , da sonst die Behauptung trivial ist. Dann ist  $\|w\| > 0$  und es gilt für  $\lambda \in \mathbb{R}$ :

$$\begin{aligned} 0 &\leq \langle v - \lambda w, v - \lambda w \rangle \\ &= \langle v, v - \lambda w \rangle - \lambda \langle w, v - \lambda w \rangle \\ &= \langle v, v \rangle - \lambda \langle v, w \rangle - \lambda \langle w, v \rangle + \lambda^2 \langle w, w \rangle \\ &= \|v\|^2 - 2\lambda \langle v, w \rangle + \lambda^2 \|w\|^2. \end{aligned}$$

Speziell gilt für  $\lambda := \langle v, w \rangle / \|w\|^2$ :

$$0 \leq \|v\|^2 - 2 \frac{\langle v, w \rangle^2}{\|w\|^2} + \frac{\langle v, w \rangle^2}{\|w\|^2} = \|v\|^2 - \frac{\langle v, w \rangle^2}{\|w\|^2},$$

also  $\langle v, w \rangle^2 \leq \|v\|^2 \|w\|^2$  und damit folgt das Lemma.  $\square$

*Beweis von Satz 6.20:* Die Punkte 1.-3. von Definition 6.16 sind klar.

Zu 4.: Seien  $v, w \in V$ . Dann gilt:

$$\begin{aligned} \|v + w\|^2 &= \langle v + w, v + w \rangle \\ &= \langle v, v \rangle + 2\langle v, w \rangle + \langle w, w \rangle \\ &\leq \|v\|^2 + 2|\langle v, w \rangle| + \|w\|^2 \\ &\stackrel{\text{C.S.}}{\leq} \|v\|^2 + 2\|v\|\|w\| + \|w\|^2 = (\|v\| + \|w\|)^2. \end{aligned}$$

□

**Beispiel 6.22.** • Sei  $\mathbb{R}^n$  versehen mit dem Standard-Skalarprodukt  $\langle \cdot, \cdot \rangle$ . Dann ist die zugehörige Norm

$$\|x\| = \sqrt{x_1^2 + \dots + x_n^2}.$$

• Sei  $(C([0, 1]), \langle \cdot, \cdot \rangle)$  wie in Übung 6.17. Die zugehörige Norm ist gegeben durch

$$\|f\| = \sqrt{\int_0^1 f(t)^2 dt}$$

für  $f \in C([0, 1])$ .

Sei  $(V, \langle \cdot, \cdot \rangle)$  ein euklidischer Raum. Aus der Cauchy-Schwarzschen Ungleichung folgt, dass für  $v, w \in V \setminus \{0\}$  gilt:

$$-1 \leq \frac{\langle v, w \rangle}{\|v\|\|w\|} \leq 1.$$

Deshalb existiert genau ein  $\varphi \in [0, \pi]$  mit

$$\cos \varphi = \frac{\langle v, w \rangle}{\|v\|\|w\|}.$$

Die Zahl  $\varphi$  heißt der **Winkel** zwischen  $v$  und  $w$  und wird mit  $\angle(v, w)$  bezeichnet.

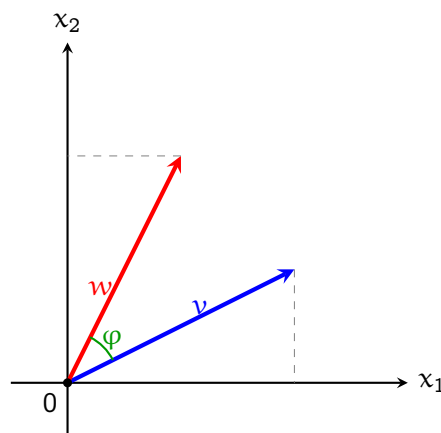


Abbildung 6.1: Veranschaulichung des Winkels  $\varphi = \angle(v, w)$  zwischen zwei Vektoren in  $\mathbb{R}^2$ .

**Definition 6.23 (Orthogonal).** Sei  $(V, \langle \cdot, \cdot \rangle)$  ein euklidischer Raum. Zwei Vektoren  $v, w \in V$  heißen **orthogonal** (in Zeichen  $v \perp w$ ), falls  $\langle v, w \rangle = 0$  gilt.

Sind  $v, w \in V \setminus \{0\}$ , so gilt:

$$v \perp w \iff \sphericalangle(v, w) = \frac{\pi}{2}.$$

**Satz 6.24 (Pythagoras).** Sei  $(V, \langle \cdot, \cdot \rangle)$  ein euklidischer Raum und seien  $v, w \in V$ . Dann gilt:

$$v \perp w \iff \|v + w\|^2 = \|v\|^2 + \|w\|^2.$$

*Beweis.* Es gilt:

$$\begin{aligned} \|v + w\|^2 &= \langle v + w, v + w \rangle \\ &= \langle v, v \rangle + 2\langle v, w \rangle + \langle w, w \rangle \\ &= \|v\|^2 + 2\langle v, w \rangle + \|w\|^2, \end{aligned}$$

und damit die Behauptung. □

**Beispiel 6.25.** In  $\mathbb{R}^2$  mit dem Standardskalarprodukt gilt  $(0, 1) \perp (1, 0)$  und allgemeiner  $(x, y) \perp (y, -x)$ .

### 6.3 Komplexe Skalarprodukte

Wir nehmen an, dass  $V$  in diesem Abschnitt stets ein  $\mathbb{C}$ -Vektorraum ist.

**Definition 6.26 (Komplexes Skalarprodukt, unitärer Raum).** Eine Abbildung

$$\langle \cdot, \cdot \rangle: V \times V \rightarrow \mathbb{C}$$

heißt **(komplexes) Skalarprodukt** auf  $V$ , falls für alle  $v, v', w \in V$  und  $\lambda \in \mathbb{C}$  gilt:

1.  $\langle v, v \rangle \in \mathbb{R}$  und  $\langle v, v \rangle > 0$  für  $v \neq 0$  (positive Definitheit)
2.  $\langle v + v', w \rangle = \langle v, w \rangle + \langle v', w \rangle$  (Additivität)
3.  $\langle \lambda v, w \rangle = \lambda \langle v, w \rangle$  (Homogenität im 1. Argument)
4.  $\langle v, w \rangle = \overline{\langle w, v \rangle}$  (Hermitesche Symmetrie)

Das Paar  $(V, \langle \cdot, \cdot \rangle)$  heißt dann ein **unitärer Vektorraum (oder unitärer Raum)**.

**Beispiel 6.27.** 1. Sei  $V = \mathbb{C}^n$  mit  $\langle x, y \rangle = \sum_{i=1}^n x_i \overline{y_i}$  für  $x = (x_1, \dots, x_n), y = (y_1, \dots, y_n) \in \mathbb{C}^n$ . Dies ist das **Standardskalarprodukt** auf  $\mathbb{C}^n$ .

2. Sei  $C([0, 1], \mathbb{C})$  der  $\mathbb{C}$ -Vektorraum aller stetigen  $\mathbb{C}$ -wertigen Funktionen auf  $[0, 1]$ . Mit

$$\langle f, g \rangle = \int_0^1 f(t) \overline{g(t)} dt$$

ist  $(C([0, 1], \mathbb{C}), \langle \cdot, \cdot \rangle)$  ein unitärer Raum. Hier ist das Integral definiert als

$$\int_0^1 f(t) dt := \int_0^1 \operatorname{Re} f(t) dt + i \cdot \int_0^1 \operatorname{Im} f(t) dt.$$

**Satz 6.28 (Sesquilinearität).** Sei  $(V, \langle \cdot, \cdot \rangle)$  ein unitärer Vektorraum. Dann gilt für alle  $v, w, w' \in V$  und  $\lambda, \mu \in \mathbb{C}$ :

$$\langle v, \lambda w + \mu w' \rangle = \bar{\lambda} \langle v, w \rangle + \bar{\mu} \langle v, w' \rangle.$$

*Beweis.* Es gilt:

$$\begin{aligned} \langle v, \lambda w + \mu w' \rangle &= \overline{\langle \lambda w + \mu w', v \rangle} \quad (\text{nach Eigenschaft 3.}) \\ &= \overline{\lambda \langle w, v \rangle + \mu \langle w', v \rangle} \quad (\text{Linearität im 1. Argument}) \\ &= \bar{\lambda} \overline{\langle w, v \rangle} + \bar{\mu} \overline{\langle w', v \rangle} \\ &= \bar{\lambda} \langle v, w \rangle + \bar{\mu} \langle v, w' \rangle. \end{aligned}$$

□

**Definition 6.29 (Norm auf unitärem Vektorraum).** Sei  $V$  ein  $\mathbb{C}$ -Vektorraum. Wie für reelle Vektorräume heißt eine Abbildung  $\|\cdot\|: V \rightarrow \mathbb{R}$  Norm auf  $V$ , falls für alle  $v, w \in V$  und  $\lambda \in \mathbb{C}$ :

1.  $\|v\| \geq 0$  (Nichtnegativität)
2.  $\|v\| = 0 \Leftrightarrow v = 0$  (Definitheit)
3.  $\|\lambda v\| = |\lambda| \|v\|$  (Homogenität)
4.  $\|v + w\| \leq \|v\| + \|w\|$  (Dreiecksungleichung)

$(V, \|\cdot\|)$  heißt dann **normierter Vektorraum**.

**Satz 6.30 (Cauchy-Schwarz-Ungleichung - komplexe Variante).** Sei  $(V, \langle \cdot, \cdot \rangle)$  ein unitärer Vektorraum. Für alle  $v, w \in V$  gilt:

$$|\langle v, w \rangle| \leq \|v\| \|w\| \quad \text{mit} \quad \|v\| = \sqrt{\langle v, v \rangle}.$$

*Beweis.* Für  $w = 0$  ist die Aussage trivial. Sei  $w \neq 0$  und  $\lambda := \frac{\langle v, w \rangle}{\|w\|^2}$ . Dann gilt:

$$\begin{aligned} 0 &\leq \langle v - \lambda w, v - \lambda w \rangle \\ &= \langle v, v \rangle - \bar{\lambda} \langle v, w \rangle - \lambda \overline{\langle v, w \rangle} + |\lambda|^2 \langle w, w \rangle \\ &= \|v\|^2 - \frac{|\langle v, w \rangle|^2}{\|w\|^2} - \frac{|\langle v, w \rangle|^2}{\|w\|^2} + \frac{|\langle v, w \rangle|^2}{\|w\|^2} \end{aligned}$$

$$= \|v\|^2 - \frac{|\langle v, w \rangle|^2}{\|w\|^2}$$

Umstellen ergibt  $|\langle v, w \rangle|^2 \leq \|v\|^2 \|w\|^2$ .  $\square$

**Satz 6.31 (Induzierte Norm - komplexe Variante).** Sei  $(V, \langle \cdot, \cdot \rangle)$  ein unitärer Vektorraum. Die Abbildung  $\|\cdot\| = \sqrt{\langle \cdot, \cdot \rangle}$  ist eine Norm auf  $V$  und heißt **zugehörige Norm**.

*Beweis.* 1. und 2. folgen direkt aus der positiven Definitheit des Skalarprodukts.

Zu 3.: Sei  $\lambda \in \mathbb{C}$ . Wir erhalten:

$$\|\lambda v\|^2 = \langle \lambda v, \lambda v \rangle = \lambda \bar{\lambda} \langle v, v \rangle = |\lambda|^2 \|v\|^2.$$

Zu 4.: Mit Cauchy-Schwarz erhalten wir:

$$\begin{aligned} \|v + w\|^2 &= \langle v + w, v + w \rangle \\ &= \|v\|^2 + 2\operatorname{Re}\langle v, w \rangle + \|w\|^2 \\ &\leq \|v\|^2 + 2|\langle v, w \rangle| + \|w\|^2 \\ &\leq (\|v\| + \|w\|)^2. \end{aligned}$$

$\square$

**Definition 6.32 (Orthogonalität - komplexe Variante).** Sei  $(V, \langle \cdot, \cdot \rangle)$  ein unitärer Vektorraum. Zwei Vektoren  $v, w \in V$  heißen **orthogonal** ( $v \perp w$ ), falls  $\langle v, w \rangle = 0$ .

**Satz 6.33 (Pythagoras - komplexe Variante).** Sei  $(V, \langle \cdot, \cdot \rangle)$  ein unitärer Vektorraum. Für  $v, w \in V$  gilt:

$$v \perp w \Rightarrow \|v + w\|^2 = \|v\|^2 + \|w\|^2.$$

*Beweis.*

$$\|v + w\|^2 = \langle v + w, v + w \rangle = \|v\|^2 + \langle v, w \rangle + \langle w, v \rangle + \|w\|^2$$

Für  $v \perp w$  ( $\langle v, w \rangle = 0$ ) vereinfacht sich dies zu  $\|v\|^2 + \|w\|^2$ .  $\square$

**Beispiel 6.34.** (1) In  $\mathbb{C}^2$  mit dem Standardskalarprodukt sind  $x = (1, i)$  und  $y = (i, 1)$  orthogonal.

(2) Für  $v \in V$  gilt:  $v \perp v \Rightarrow v = 0$ .

## 6.4 Orthogonalität

In euklidischen oder unitären Vektorräumen können wir mithilfe des Skalarproduktes und der **Cauchy-Schwarz** Ungleichung eine Winkelmessung erklären. Im Folgenden sind wir in erster Linie daran interessiert, das Konzept der Orthogonalität auf Untervektorräume zu erweitern.

**Definition 6.35.** Sei  $V$  ein euklidischer bzw. unitärer Vektorraum.

(a) Zwei Vektoren  $v, w \in V$  heißen **orthogonal**, in Zeichen

$$v \perp w \Leftrightarrow \langle v, w \rangle = 0.$$

(b) Zwei Teilmengen  $U, W \subset V$  heißen **orthogonal**, in Zeichen

$$U \perp W \Leftrightarrow u \perp w \quad \text{für alle } u \in U \text{ und alle } w \in W.$$

(c) Ist  $U \subset V$  eine Teilmenge, so ist ihr **orthogonales Komplement** definiert als

$$U^\perp := \{v \in V : v \perp u \text{ für alle } u \in U\}.$$

(d) Eine Familie  $(v_1, \dots, v_n)$  in  $V$  heißt **orthogonal** oder **Orthogonalsystem**, wenn

$$v_i \perp v_j \quad \text{für alle } i \neq j.$$

Sie heißt **orthonormal** oder **Orthonormalsystem**, falls zusätzlich

$$\|v_i\| = 1 \quad \text{für alle } i,$$

und **Orthonormalbasis**, falls sie auch eine Basis ist, d. h. eine Basis mit

$$\langle v_i, v_j \rangle = \delta_{ij}.$$

Eine Orthonormalbasis wird im Folgenden mit **ONB** abgekürzt.

(e) Ist  $V = V_1 \oplus \dots \oplus V_k$ , so heißt die direkte Summe **orthogonal**, in Zeichen

$$V = V_1 \oplus \dots \oplus V_k, \quad \text{falls } V_i \perp V_j \text{ für alle } i \neq j.$$

**Satz 6.36.** Sei  $U, W \subset V$ . Dann gilt:

1.  $U \perp W \Leftrightarrow \text{span}(U) \perp \text{span}(W)$ .
2.  $U^\perp$  ist ein Untervektorraum von  $V$ .
3.  $(\text{span}(U))^\perp = U^\perp$ .

*Beweis.* Übung. □

**Lemma 6.37.** Ist  $(v_1, \dots, v_n)$  eine orthogonale Familie in  $V$  und  $v_i \neq 0$  für alle  $i$ , so gilt:

- (a) Die Familie  $(\alpha_1 v_1, \dots, \alpha_n v_n)$  mit  $\alpha_i := \|v_i\|^{-1}$  ist orthonormal.

(b)  $(v_1, \dots, v_n)$  ist linear unabhängig.

*Beweis.* (a) Aus

$$\langle \alpha_i v_i, \alpha_j v_j \rangle = \alpha_i \overline{\alpha_j} \langle v_i, v_j \rangle$$

folgt für  $i \neq j$ , dass die Vektoren orthogonal sind, und für  $i = j$  gilt wegen  $\alpha_i \in \mathbb{R}$

$$\langle \alpha_i v_i, \alpha_i v_i \rangle = \alpha_i \overline{\alpha_i} \langle v_i, v_i \rangle = \alpha_i^2 \|v_i\|^2 = 1.$$

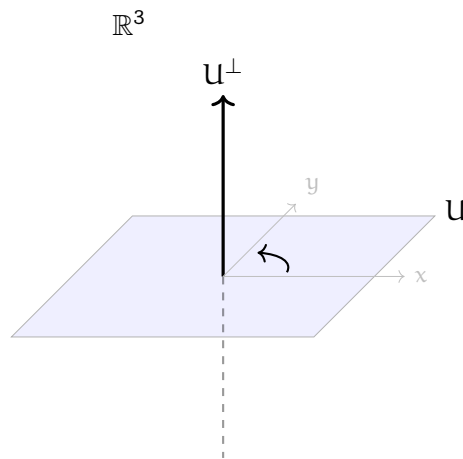
(b) Sei

$$\lambda_1 v_1 + \dots + \lambda_n v_n = 0.$$

Bildet wir das Skalarprodukt von beiden Seiten der Gleichung mit  $v_i$ , so erhalten wir

$$\left\langle \sum_{j=1}^n \lambda_j v_j, v_i \right\rangle = \lambda_i \langle v_i, v_i \rangle = \lambda_i = 0.$$

Also folgt  $\lambda_1 = \dots = \lambda_n = 0$ . Das entspricht ganz der Vorstellung, dass senkrecht zueinander stehende Vektoren linear unabhängig sind. □



Eine ganz einfache, aber wichtige Eigenschaft einer ONB zeigt die folgende Bemerkung.

**Lemma 6.38.** Sei  $(v_1, \dots, v_n)$  eine ONB von  $V$  und  $v \in V$  beliebig. Setzen wir

$$\lambda_i := \langle v, v_i \rangle, \quad \text{so ist} \quad v = \lambda_1 v_1 + \dots + \lambda_n v_n.$$

*Beweis.* Da  $(v_1, \dots, v_n)$  Basis von  $V$  ist, existieren eindeutig bestimmte  $\alpha_1, \dots, \alpha_n \in K$  mit

$$v = \alpha_1 v_1 + \dots + \alpha_n v_n.$$

Wir bilden wieder das Skalarprodukt von beiden Seiten der Gleichung mit  $v_i$ , und erhalten

$$\langle v, v_i \rangle = \left\langle \sum_{j=1}^n \alpha_j v_j, v_i \right\rangle = \alpha_i \langle v_i, v_i \rangle = \alpha_i.$$

Mit  $\lambda_i = \alpha_i$  für alle  $i = 1, \dots, n$  folgt die Behauptung.  $\square$

Wir erhalten folgendes Korollar.

**Korollar 6.39.** Sei  $\mathcal{B} = (v_1, \dots, v_n)$  eine ONB von  $V$  und  $F \in \text{End}(V)$  mit  $\dim V = n$ . Dann gilt

$$\mathcal{M}_{\mathcal{B}}(F) = A \text{ mit } a_{ij} = \langle F(v_j), v_i \rangle \text{ für } i, j = 1, \dots, n.$$

Im  $\mathbb{R}^n$  oder  $\mathbb{C}^n$  mit dem kanonischen Skalarprodukt ist die kanonische Basis orthonormal. Wir geben nun einen konstruktiven Satz an (der auf Gram und Schmidt zurückgeht), welcher eine ONB berechnet.

**Satz 6.40 (Orthonormalisierungssatz).** Sei  $V$  ein endlichdimensionaler euklidischer bzw. unitärer Vektorraum und  $W \subset V$  ein Untervektorraum mit ONB  $(w_1, \dots, w_m)$ . Dann gibt es eine Ergänzung zu einer ONB

$$(w_1, \dots, w_m, w_{m+1}, \dots, w_n) \text{ von } V.$$

*Beweis.* Ist  $W = V$ , so ist nichts mehr zu tun. Andernfalls gibt es einen Vektor  $v \in V \setminus W$ , und wir definieren

$$\tilde{v} := \langle v, w_1 \rangle w_1 + \dots + \langle v, w_m \rangle w_m.$$

Dies nennen wir die **senkrechte Projektion** von  $v$  auf  $W$  (siehe Abb. 6.2), denn setzen wir

$$w := v - \tilde{v}, \text{ so ist } w \perp W.$$

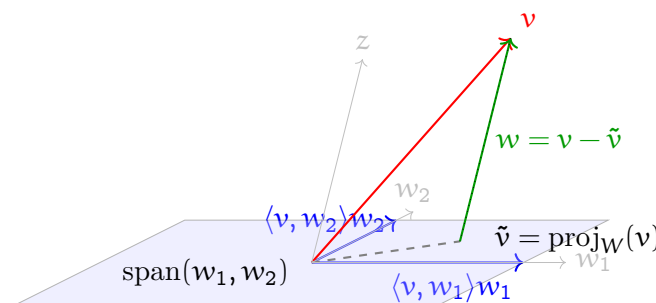


Abbildung 6.2: Konstruktion eines zu  $W = \text{span}(w_1, w_2)$  senkrechten Vektors  $w$  mittels Projektion  $\tilde{v} = \text{proj}_W(v)$  von  $v \notin W$  auf  $W$ .

Dazu genügt es zu zeigen, dass  $w \perp w_i$  für  $i = 1, \dots, m$ , und das folgt aus

$$\langle w, w_i \rangle = \langle v, w_i \rangle - \langle \tilde{v}, w_i \rangle = \langle v, w_i \rangle - \langle v, w_i \rangle = 0,$$

da  $\langle w_i, w_j \rangle = \delta_{ij}$ . Bis auf die Länge ist  $w$  schon für die Ergänzung der Basis geeignet, daher setzen wir

$$w_{m+1} := \frac{1}{\|w\|} \cdot w.$$

Nun hat  $W' := \text{span}(w_1, \dots, w_{m+1})$  eine Orthonormalbasis  $(w_1, \dots, w_{m+1})$ , und es ist  $\dim W' = m + 1$ . Indem wir das obige Verfahren so oft wie nötig wiederholt, erhalten wir die gewünschte Ergänzung.  $\square$

Da der Fall  $W = \{0\}$  erlaubt ist, folgt sofort das

**Korollar 6.41.** Jeder endlichdimensionale euklidische bzw. unitäre Vektorraum besitzt eine ONB.

Die nach der Aussage des Satzes existierenden Vektoren  $w_{m+1}, \dots, w_n$  stehen senkrecht auf  $W$ , also ist

$$W' := \text{span}(w_{m+1}, \dots, w_n) \subset W^\perp.$$

Wir zeigen nun  $W' = W^\perp$ . Denn ist umgekehrt

$$w = \lambda_1 w_1 + \dots + \lambda_m w_m + \lambda_{m+1} w_{m+1} + \dots + \lambda_n w_n \in W^\perp,$$

so folgt durch skalare Multiplikation dieser Gleichung von rechts mit  $w_i, i = 1, \dots, m$

$$0 = \langle w, w_i \rangle = \lambda_i \quad \text{für } i = 1, \dots, m,$$

also  $w \in W'$ , und somit

$$W^\perp = W' = \text{span}(w_{m+1}, \dots, w_n).$$

Wir erhalten folgendes Korollar.

**Korollar 6.42.** Ist  $W$  Untervektorraum eines endlich dimensionalen euklidischen oder unitären Vektorraums  $V$ , so gilt

$$V = W \oplus W^\perp, \quad \dim V = \dim W + \dim W^\perp.$$

## 6.5 Volumina und Haddamard Ungleichung

In einem endlichdimensionalen euklidischen Vektorraum  $V$  seien beliebige Vektoren  $v_1, \dots, v_m \in V$  gegeben, wobei  $m \leq n$ . Dann nennen wir

$$G(v_1, \dots, v_m) := \det \begin{pmatrix} \langle v_1, v_1 \rangle & \cdots & \langle v_1, v_m \rangle \\ \vdots & \ddots & \vdots \\ \langle v_m, v_1 \rangle & \cdots & \langle v_m, v_m \rangle \end{pmatrix}$$

die **Gramsche Determinante** von  $v_1, \dots, v_m$ .

**Lemma 6.43.** Es gilt stets  $G(v_1, \dots, v_m) \geq 0$  und

$$G(v_1, \dots, v_m) > 0 \iff (v_1, \dots, v_m) \text{ linear unabhängig.}$$

*Beweis.* Nach Satz 6.40 gibt es eine Orthogonalbasis  $(w_1, \dots, w_n)$  von  $V$ . Ist

$$v_i = a_{i1}w_1 + \dots + a_{in}w_n$$

und  $A := (a_{ij}) \in M(m \times n; \mathbb{R})$ , so ist

$$G(v_1, \dots, v_m) = \det(A \cdot A^T) = \sum_{1 \leq k_1 \leq \dots \leq k_m \leq n} \det(A^{k_1, \dots, k_m})^2,$$

wobei

$$A^{k_1, \dots, k_m} := (a^{k_1}, \dots, a^{k_m}) \text{ ein } m\text{-reihiger Minor von } A \text{ ist.}$$

Die Darstellung der Determinante mit Hilfe der  $m$ -reihigen Minoren nehmen wir als gegeben hin; sie kann mithilfe der Determinantenregeln nachgerechnet werden. Somit folgt die Aussage des Lemmas, da die rechte Seite genau dann 0 ist, wenn der Rang von  $A$  kleiner als  $m$  ist.  $\square$

Nach dieser Bemerkung können wir aus der Gramschen Determinante die Wurzel ziehen und durch

$$\text{Vol}(v_1, \dots, v_m) := \sqrt{G(v_1, \dots, v_m)}$$

das  $m$ -dimensionale **Volumen** des durch  $v_1, \dots, v_m$  aufgespannten Spates erklären. Das ist die Grundlage der Definition von verallgemeinerten Oberflächenintegralen in der Analysis. Im  $\mathbb{R}^n$  mit dem kanonischen Skalarprodukt können wir eine beliebige Basis  $(v_1, \dots, v_n)$  als Zeilen in eine Matrix  $A$  eintragen. Dann ist

$$\text{Vol}(v_1, \dots, v_n) := \sqrt{(\det A)^2} = |\det A|.$$

Eine natürliche Frage ist nun, wie wir mit  $m$  Vektoren einer festgelegten Länge ein möglichst großes Volumen aufspannen. Diese Frage wird mit der Ungleichung von Hadamard

**Satz 6.44 (Hadamard Ungleichung).** Seien  $m$  Vektoren  $v_1, \dots, v_m \in V$  eines Vektorraums  $V$  mit  $\dim V = n \geq m$  gegeben. Dann gilt:

$$\text{Vol}(v_1, \dots, v_m) \leq \|v_1\| \cdot \dots \cdot \|v_m\|.$$

*Beweis.* Es genügt, die Ungleichung in der quadratischen Form

$$G(v_1, \dots, v_m) \leq \langle v_1, v_1 \rangle \cdots \langle v_m, v_m \rangle$$

zu beweisen, und dazu genügt es zu zeigen, dass für jedes  $r \leq m$  folgendes gilt: Für die eindeutig bestimmte orthogonale Zerlegung

$$v_r = \tilde{v} + v' \quad \text{mit} \quad \tilde{v} = \lambda_1 v_1 + \dots + \lambda_{r-1} v_{r-1}, \quad \text{und} \quad v' \perp \text{span}(v_1, \dots, v_{r-1})$$

ist

$$G(v_1, \dots, v_r) = G(v_1, \dots, v_{r-1}) \cdot \langle v', v' \rangle \leq G(v_1, \dots, v_{r-1}) \cdot \langle v_r, v_r \rangle,$$

und die Ungleichung ist genau dann eine Gleichung, wenn

$$v_r \perp \text{span}(v_1, \dots, v_{r-1}).$$

Die Ungleichung und das Kriterium für die Gleichheit folgen unmittelbar aus der Definition. Zum Beweis der ersten Gleichung verwenden wir, dass

$$\langle v_i, v_r \rangle = \langle v_i, \tilde{v} \rangle \quad \text{für } i = 1, \dots, r-1 \quad \text{und} \quad \langle v_r, v_r \rangle = \langle v', v' \rangle + \langle \tilde{v}, \tilde{v} \rangle.$$

Daher ist

$$G(v_1, \dots, v_r) = \det \begin{pmatrix} \langle v_1, v_1 \rangle & \cdots & \langle v_1, v_{r-1} \rangle & \langle v_1, \tilde{v} \rangle \\ \vdots & \ddots & \vdots & \vdots \\ \langle v_{r-1}, v_1 \rangle & \cdots & \langle v_{r-1}, v_{r-1} \rangle & \langle v_{r-1}, \tilde{v} \rangle \\ \langle \tilde{v}, v_1 \rangle & \cdots & \langle \tilde{v}, v_{r-1} \rangle & \langle \tilde{v}, \tilde{v} \rangle + \langle v', v' \rangle \end{pmatrix}.$$

Indem wir von der letzten Spalte die Spalten 1 bis  $r-1$ , jeweils multipliziert mit  $\lambda_1, \dots, \lambda_{r-1}$ , abziehen, wird daraus die Spalte

$$\begin{pmatrix} 0 \\ \vdots \\ 0 \\ \langle v', v' \rangle \end{pmatrix},$$

und die Behauptung folgt durch Entwicklung der Determinante nach der neuen letzten Spalte.  $\square$

## 6.6 Adjungierte Abbildungen

Seien  $(V, \langle \cdot, \cdot \rangle)$  und  $(W, \langle \cdot, \cdot \rangle)$  beide euklidisch oder beide unitär und entsprechend  $K = \mathbb{R}$  oder  $K = \mathbb{C}$ .

**Definition 6.45 (Adjungierte Abbildung).** Sei  $F : V \rightarrow W$  linear. Eine lineare Abbildung  $F^* : W \rightarrow V$  heißt zu  $F$  **adjungierte Abbildung**, falls für alle  $v \in V, w \in W$  gilt:

$$\langle F(v), w \rangle = \langle v, F^*(w) \rangle.$$

**Lemma 6.46.** Wenn  $F \in \text{Hom}(V, W)$  eine adjungierte Abbildung besitzt, so ist diese eindeutig bestimmt.

*Beweis.* Seien  $G, H \in \text{Hom}(W, V)$  adjungierte Abbildungen zu  $F$ . Dann gilt für alle  $v \in V, w \in W$

$$\langle v, G(w) \rangle = \langle F(v), w \rangle = \langle v, H(w) \rangle, \quad \text{also } \langle v, G(w) - H(w) \rangle = 0.$$

Speziell für  $v := G(w) - H(w)$  gilt

$$\langle G(w) - H(w), G(w) - H(w) \rangle = 0, \quad \text{also } G(w) = H(w).$$

$\square$

**Satz 6.47.** Sei  $V$  ein  $K$ -Vektorraum mit  $\dim V = n$ . Dann besitzt jedes  $F \in \text{Hom}(V, W)$  eine adjungierte Abbildung.

*Beweis.* Ist  $V = \{0\}$ , so ist die Behauptung klar. Sei  $V \neq \{0\}$  und  $(v_1, \dots, v_n)$  eine ONB von  $V$ , die nach Korollar 6.41 existiert. Definiere

$$F^* : W \rightarrow V, \quad w \mapsto F^*(w) := \sum_{j=1}^n \langle w, F(v_j) \rangle v_j.$$

Wir rechnen leicht nach, dass  $F^*$  linear ist. Wir zeigen, dass  $F^*$  zu  $F$  adjungiert ist. Sei dazu  $v \in V$  und  $w \in W$ . Dann gilt:

$$\begin{aligned} \langle v, F^*(w) \rangle &= \left\langle v, \sum_{j=1}^n \langle w, F(v_j) \rangle v_j \right\rangle = \sum_{j=1}^n \overline{\langle w, F(v_j) \rangle} \cdot \langle v, v_j \rangle = \sum_{j=1}^n \langle v, v_j \rangle \cdot \overline{\langle w, F(v_j) \rangle} \\ &= \sum_{j=1}^n \langle v, v_j \rangle \cdot \langle F(v_j), w \rangle = \left\langle \sum_{j=1}^n \langle v, v_j \rangle F(v_j), w \right\rangle = \left\langle F\left(\sum_{j=1}^n \langle v, v_j \rangle v_j\right), w \right\rangle \\ &= \langle F(v), w \rangle. \end{aligned}$$

□

**Definition 6.48 (Adjungierte Matrix).** Sei  $A \in M(m \times n, K)$ . Ist  $A = (a_{ij})$ , so wird mit  $\bar{A}$  die Matrix  $(\bar{a}_{ij})$  bezeichnet. Die Matrix  $A^* := \bar{A}^T \in M(n \times m, K)$  heißt die zu  $A$  adjungierte Matrix.

**Übung 6.49.** Zeige:

- a) Ist  $A \in M(m \times n, \mathbb{R})$ , so ist  $A^* = A^T$ .
- b) Für  $A, B \in M(m \times n, K)$  und  $\lambda \in K$  gilt:
  - (i)  $(A^*)^* = A$ ,
  - (ii)  $(A + B)^* = A^* + B^*$ ,
  - (iii)  $(\lambda A)^* = \bar{\lambda} A^*$ ,
  - (iv)  $(AB)^* = B^* A^*$ ,
  - (v)  $\det(A^*) = \overline{\det(A)}$ .

**Satz 6.50.** Seien  $V$  und  $W$  jeweils  $K$ -Vektorräume mit  $\dim V = n$  und  $\dim W = m$ . Sei  $\mathcal{A} = (v_1, \dots, v_n)$  eine ONB von  $V$  und  $\mathcal{B} = (w_1, \dots, w_m)$  eine ONB von  $W$ . Wird  $F \in \text{Hom}(V, W)$  bezüglich  $\mathcal{A}$  und  $\mathcal{B}$  durch die Matrix  $A := \mathcal{M}_{\mathcal{B}}^{\mathcal{A}}(F) \in M(m \times n, K)$  dargestellt, so wird  $F^*$  bezüglich  $\mathcal{B}$  und  $\mathcal{A}$  durch  $A^*$  dargestellt.

*Beweis.* Nach Lemma 6.38 gilt für  $j \in \{1, \dots, n\}$

$$F(v_j) = \sum_{i=1}^m \langle F(v_j), w_i \rangle w_i, \quad a_{ij} = \langle F(v_j), w_i \rangle.$$

Nach dem Beweis von Satz 6.47 gilt für  $i \in \{1, \dots, m\}$

$$F^*(w_i) = \sum_{j=1}^n \langle w_i, F(v_j) \rangle v_j,$$

Wird  $F^*$  bezüglich  $\mathcal{B}$  und  $\mathcal{A}$  durch Matrix  $B := (b_{ji})$  dargestellt, so gilt deshalb

$$b_{ji} = \langle w_i, F(v_j) \rangle = \overline{\langle F(v_j), w_i \rangle} = \bar{a}_{ij}.$$

Also gilt  $B = \bar{A}^T = A^*$ . □

**Definition 6.51 (Normaler Endomorphismus).** Eine lineare Abbildung  $F \in \text{End}(V)$  heißt **normal**, falls die zu  $F$  adjungierte Abbildung  $F^* \in \text{End}(V)$  existiert und

$$F \circ F^* = F^* \circ F.$$

**Satz 6.52.** Sei  $F \in \text{End}(V)$ . Dann sind äquivalent:

- (i)  $F$  ist normal.
- (ii)  $F$  besitzt eine adjungierte Abbildung  $F^*$  und es gilt für alle  $v, v' \in V$

$$\langle F(v), F(v') \rangle = \langle F^*(v), F^*(v') \rangle.$$

*Beweis.* (i)  $\Rightarrow$  (ii): Seien  $v, v' \in V$ . Dann gilt:

$$\begin{aligned} \langle F(v), F(v') \rangle &= \langle v, (F^* \circ F)(v') \rangle = \langle v, (F \circ F^*)(v') \rangle \\ &= \overline{\langle (F \circ F^*)(v'), v \rangle} = \overline{\langle F^*(v'), F^*(v) \rangle} = \langle F^*(v), F^*(v') \rangle. \end{aligned}$$

(ii)  $\Rightarrow$  (i): Seien  $v, v' \in V$  beliebig. Dann gilt:

$$\begin{aligned} \langle (F \circ F^*)(v), v' \rangle &= \langle F^*(v), F^*(v') \rangle = \langle F(v), F(v') \rangle = \overline{\langle F(v'), F(v) \rangle} \\ &= \overline{\langle v', (F^* \circ F)(v) \rangle} = \langle (F^* \circ F)(v), v' \rangle. \end{aligned}$$

Daraus folgt  $\langle (F \circ F^*)(v) - (F^* \circ F)(v), v' \rangle = 0$ . Setzen wir für  $v'$  den Vektor  $(F \circ F^*)(v) - (F^* \circ F)(v)$  ein, so erhält man

$$(F \circ F^*)(v) - (F^* \circ F)(v) = 0,$$

also wegen  $v \in V$  beliebig  $F \circ F^* = F^* \circ F$ . □

**Satz 6.53.** Sei  $F \in \text{End}(V)$  normal und sei  $v \in V$ . Dann ist  $v$  Eigenvektor von  $F$  zum Eigenwert  $\lambda \in K$  genau dann, wenn  $v$  Eigenvektor von  $F^*$  zum Eigenwert  $\bar{\lambda}$  ist.

*Beweis.* Für  $\lambda \in K$  und  $v \in V$  gilt

$$\begin{aligned} \|F(v) - \lambda v\|^2 &= \langle F(v) - \lambda v, F(v) - \lambda v \rangle \\ &= \langle F(v), F(v) \rangle - \lambda \langle v, F(v) \rangle - \bar{\lambda} \langle F(v), v \rangle + |\lambda|^2 \langle v, v \rangle \\ &= \langle F^*(v), F^*(v) \rangle - \lambda \langle F^*(v), v \rangle - \bar{\lambda} \langle v, F^*(v) \rangle + |\lambda|^2 \langle v, v \rangle \\ &= \langle F^*(v) - \bar{\lambda} v, F^*(v) - \bar{\lambda} v \rangle \\ &= \|F^*(v) - \bar{\lambda} v\|^2. \end{aligned}$$

Also folgt  $F(v) = \lambda v$  genau dann, wenn  $F^*(v) = \bar{\lambda} v$  und damit die Behauptung. □

**Satz 6.54 (Charakterisierung normaler Endomorphismen).** Sei  $(V, \langle \cdot, \cdot \rangle)$  ein euklidischer oder unitärer mit  $\dim(V) = n \geq 1$  und sei  $F \in \text{End}(V)$ . Ist  $(V, \langle \cdot, \cdot \rangle)$  euklidisch, so werde zusätzlich vorausgesetzt, dass das charakteristische Polynom von  $F$  in Linearfaktoren zerfällt. Dann sind äquivalent:

- (i)  $F$  ist normal.

(ii) Es existiert eine ONB von Eigenvektoren von  $F$ .

*Beweis.* (i)  $\Rightarrow$  (ii): Wir beweisen die Aussage durch Induktion nach  $n = \dim(V)$ .

$n = 1$ : Sei  $\lambda_1$  ein Eigenwert von  $F$  und sei  $v_1$  Eigenvektor von  $F$  zum Eigenwert  $\lambda_1$ . O.E. sei  $\|v_1\| = 1$ . Dann erfüllt  $(v_1)$  die Behauptung.

$n - 1 \rightarrow n$ : Sei  $\lambda_n$  ein Eigenwert von  $F$  und  $v_n$  ein zugehöriger Eigenvektor. Sei  $U := \{v_n\}^\perp$ . Nach Satz 6.36 ist  $U$  Untervektorraum von  $V$  und nach Korollar 6.42 gilt  $V = \text{span}(v_n) \oplus U$  und deshalb  $\dim(U) = n - 1$ .

**Behauptung 1:**  $U$  ist  $F$ -invariant und  $F^*$ -invariant.

**Beweis von Behauptung 1:** Sei  $u \in U$ . Dann gilt:

$$\langle F(u), v_n \rangle = \langle u, F^*(v_n) \rangle \stackrel{\text{Satz 6.53}}{=} \langle u, \bar{\lambda}_n v_n \rangle = \lambda_n \langle u, v_n \rangle = 0.$$

Also ist  $F(u) \in \{v_n\}^\perp = U$ . Invarianz von  $U$  bzgl.  $F^*$  folgt analog.

**Behauptung 2:**  $F|_U \in \text{End}(U)$  ist normal mit  $(F|_U)^* = F^*|_U$ .

**Beweis von Behauptung 2:** Für  $u_1, u_2 \in U$  gilt:

$$\langle F(u_1), u_2 \rangle = \langle u_1, F^*(u_2) \rangle.$$

Nach der Induktionsvoraussetzung existiert eine ONB  $(v_1, \dots, v_{n-1})$  von  $U$  aus Eigenvektoren von  $F|_U$ . Offensichtlich ist dann  $(v_1, \dots, v_n)$  eine ONB von  $V$  aus Eigenvektoren von  $F$ .

(i)  $\Leftarrow$  (ii):: Sei  $(v_1, \dots, v_n)$  eine ONB von  $V$  aus Eigenvektoren von  $F$ . Sei  $\lambda_j$  der Eigenwert von  $F$  zum Eigenvektor  $v_j$ . Nach Satz 3.10 existiert genau eine lineare Abbildung  $G: V \rightarrow V$  mit  $G(v_j) = \bar{\lambda}_j v_j$  für  $j \in \{1, \dots, n\}$ . Seien  $v, v' \in V$  und seien  $\mu_j, \alpha_j \in K$  mit  $v = \sum_{j=1}^n \mu_j v_j$  und  $v' = \sum_{j=1}^n \alpha_j v_j$ . Dann gilt:

$$\begin{aligned} \langle F(v), v' \rangle &= \left\langle \sum_{i=1}^n \mu_i F(v_i), \sum_{j=1}^n \alpha_j v_j \right\rangle = \sum_{i,j=1}^n \mu_i \bar{\alpha}_j \langle \lambda_i v_i, v_j \rangle \\ &= \sum_{j=1}^n \mu_j \bar{\alpha}_j \langle v_j, \bar{\lambda}_j v_j \rangle = \sum_{j=1}^n \mu_j \sum_{i=1}^n \bar{\alpha}_i \langle v_j, G(v_i) \rangle = \langle v, G(v') \rangle, \end{aligned}$$

also ist  $G$  die zu  $F$  adjungierte Abbildung, d.h.  $G = F^*$ . Für  $j \in \{1, \dots, n\}$  gilt:

$$\begin{aligned} (F^* \circ F)(v_j) &= F^*(\lambda_j v_j) = \lambda_j F^*(v_j) = \lambda_j \bar{\lambda}_j v_j \\ &= \bar{\lambda}_j F(v_j) = F(\bar{\lambda}_j v_j) = (F \circ F^*)(v_j). \end{aligned}$$

Da  $F^* \circ F$  und  $F \circ F^*$  auf der Basis  $(v_1, \dots, v_n)$  übereinstimmen, folgt aus Satz 3.10, dass  $F \circ F^* = F^* \circ F$ . Also ist  $F$  normal.  $\square$

**Bemerkung 6.55.** a) Sei  $F \in \text{End}(V)$  und  $\mathcal{B} = (v_1, \dots, v_n)$  eine ONB von  $V$ . Wird  $F$  bezüglich  $\mathcal{B}$  durch die Matrix  $A$  dargestellt, so wird  $F^*$  nach Satz 6.50 bezüglich  $\mathcal{B}$  durch  $A^*$  dargestellt. Außerdem wird  $F \circ F^*$  durch  $AA^*$  und  $F^* \circ F$  durch  $A^*A$  dar-

gestellt. Daraus folgt  $AA^* = A^*A$ , falls  $F$  normal ist. Wird umgekehrt  $F$  bezüglich  $B$  durch  $A$  mit  $AA^* = A^*A$  dargestellt, so ist  $F$  normal.

- b) Ist  $A \in M(n \times n, K)$  eine Matrix mit  $AA^* = A^*A$ , deren charakteristisches Polynom in Linearfaktoren zerfällt, so folgt aus Satz 6.54 sofort, dass  $A$  diagonalisierbar ist. Es lässt sich sogar zeigen, dass eine reguläre Matrix  $S \in GL(n, K)$  existiert mit  $S^{-1} = S^*$ , für die  $S^{-1}AS$  Diagonalgestalt hat. Im Fall  $K = \mathbb{R}$  heißt ein solches  $S$  orthogonale Matrix, im Fall  $K = \mathbb{C}$  unitäre Matrix. Wir sagen deshalb, dass  $A$  orthogonal bzw. unitär diagonalisierbar ist.

**Definition 6.56 (Normale Matrix).**

$A \in M(n \times n, K)$  heißt **normal**, falls  $AA^* = A^*A$ .

**Satz 6.57.** Sei  $V$  endlichdimensional. Für  $F \in \text{End}(V)$  sind äquivalent:

- (i)  $F$  ist normal,
- (ii)  $F$  wird bezüglich jeder ONB von  $V$  durch eine normale Matrix dargestellt.

*Beweis.* Siehe Bemerkung 6.55. □

## 6.7 Selbstadjungierte Endomorphismen

Sei  $(V, \langle \cdot, \cdot \rangle)$  ein euklidischer oder unitärer Raum und sei  $K$  entsprechend gleich  $\mathbb{R}$  oder  $\mathbb{C}$ .

**Definition 6.58 (Selbstadjungierter Endomorphismus).** Eine Abbildung  $F \in \text{End}(V)$  heißt **selbstadjungiert**, falls zu  $F$  eine adjungierte Abbildung  $F^*$  existiert mit  $F = F^*$ , d.h.

$$\langle F(v), v' \rangle = \langle v, F(v') \rangle \quad \text{für alle } v, v' \in V.$$

**Beispiel 6.59.** Sei  $\mathbb{C}^2$  versehen mit dem Standardskalarprodukt  $\langle \cdot, \cdot \rangle$ . Sei

$$f : \mathbb{C}^2 \rightarrow \mathbb{C}^2, \quad F(z_1, z_2) = (z_1 + iz_2, -iz_1 + z_2).$$

Dann gilt für  $(z_1, z_2), (w_1, w_2) \in \mathbb{C}^2$

$$\begin{aligned} \langle F(z_1, z_2), (w_1, w_2) \rangle &= (z_1 + iz_2)\overline{w_1} + (-iz_1 + z_2)\overline{w_2} \\ &= z_1\overline{w_1} + iz_2\overline{w_1} - iz_1\overline{w_2} + z_2\overline{w_2}, \\ \langle (z_1, z_2), F(w_1, w_2) \rangle &= z_1\overline{(w_1 + iw_2)} + z_2\overline{(-iw_1 + w_2)} \\ &= z_1(\overline{w_1} - i\overline{w_2}) + z_2(i\overline{w_1} + \overline{w_2}) \\ &= z_1\overline{w_1} + iz_2\overline{w_1} - iz_1\overline{w_2} + z_2\overline{w_2}. \end{aligned}$$

Somit ist  $F$  selbstadjungiert.

**Definition 6.60 (Hermitesche Matrix).** Eine Matrix  $A \in M(n \times n, \mathbb{C})$  heißt **hermitesch**, falls  $A^* = A$ .

**Satz 6.61.** Sei  $V$  endlichdimensional und sei  $\mathcal{B}$  eine ONB von  $V$ . Sei  $F \in \text{End}(V)$ . Dann sind äquivalent:

- (i)  $F$  ist selbstadjungiert.
- (ii) Die Darstellungsmatrix von  $F$  bezüglich  $\mathcal{B}$  ist hermitesch für  $K = \mathbb{C}$  bzw. symmetrisch für  $K = \mathbb{R}$ .

*Beweis.*  $F$  werde bezüglich  $\mathcal{B}$  durch die Matrix  $A$  dargestellt. Nach Satz 6.50 wird dann  $F^*$  bezüglich  $\mathcal{B}$  durch  $A^*$  dargestellt. Deshalb gilt  $F = F^*$  genau dann, wenn  $A = A^*$  ist, und die Aussage des Satzes ist bewiesen.  $\square$

**Satz 6.62.** Jeder selbstadjungierte Endomorphismus ist normal.

*Beweis.* Dies folgt, da für  $F = F^*$  offensichtlich  $F^* \circ F = F \circ F^*$ .  $\square$

**Satz 6.63.** Sei  $(V, \langle \cdot, \cdot \rangle)$  unitär und sei  $F \in \text{End}(V)$  selbstadjungiert. Dann gilt:

- (i) Alle Eigenwerte von  $F$  sind reell.
- (ii) Sind  $v \in V$  und  $w \in V$  Eigenvektoren von  $F$  zu verschiedenen Eigenwerten, so ist  $v \perp w$ .

*Beweis.* (i) Sei  $\lambda \in \mathbb{C}$  Eigenwert von  $F$ . Dann existiert  $v \in V \setminus \{0\}$  mit  $F(v) = \lambda v$ . Es gilt:

$$\begin{aligned} \lambda \langle v, v \rangle &= \langle \lambda v, v \rangle = \langle F(v), v \rangle \\ &= \langle v, F^*(v) \rangle \stackrel{F^*=F}{=} \langle v, F(v) \rangle \\ &= \langle v, \lambda v \rangle = \bar{\lambda} \langle v, v \rangle. \end{aligned}$$

Daraus folgt  $\lambda = \bar{\lambda}$  und deshalb  $\lambda \in \mathbb{R}$ .

(ii) Sei  $v$  Eigenvektor von  $F$  zum Eigenwert  $\lambda$  und sei  $w$  Eigenvektor von  $F$  zum Eigenwert  $\mu \neq \lambda$ . Dann gilt:

$$\begin{aligned} \lambda \langle v, w \rangle &= \langle \lambda v, w \rangle = \langle F(v), w \rangle \\ &= \langle v, F(w) \rangle = \langle v, \mu w \rangle \\ &= \bar{\mu} \langle v, w \rangle \stackrel{(i)}{=} \mu \langle v, w \rangle. \end{aligned}$$

Wegen  $\lambda \neq \mu$  folgt  $\langle v, w \rangle = 0$  und damit  $v \perp w$ .  $\square$

**Bemerkung 6.64.** Analog zeigt man: Sei  $(V, \langle \cdot, \cdot \rangle)$  euklidisch und sei  $F \in \text{End}(V)$  selbstadjungiert. Sind  $v \in V$  und  $w \in V$  Eigenvektoren von  $F$  zu verschiedenen Eigenwerten, so ist  $v \perp w$ .

**Korollar 6.65.** Sei  $A \in M(n \times n, \mathbb{C})$  hermitesch. Dann sind alle Eigenwerte von  $A$  reell.

*Beweis.* Dies folgt aus Satz 6.61 und Satz 6.63(i) angewandt auf  $F_A \in \text{End}(\mathbb{C}^n)$ , wobei  $\mathbb{C}^n$  mit dem Standardskalarprodukt versehen ist.  $\square$

**Korollar 6.66.** Sei  $A \in M(n \times n, \mathbb{R})$  symmetrisch. Dann zerfällt das charakteristische Polynom  $P_A$  von  $A$  über  $\mathbb{R}$  in Linearfaktoren.

*Beweis.* Fassen wir  $A$  als  $n \times n$ -Matrix über  $\mathbb{C}$  auf, so ist  $A$  hermitesch. Das charakteristische Polynom  $P_A$  von  $A$  zerfällt als Polynom über  $\mathbb{C}$  in Linearfaktoren und hat nach Korollar 6.65 lauter reelle Nullstellen. Deshalb zerfällt  $P_A$  als Polynom über  $\mathbb{R}$ .  $\square$

**Satz 6.67.** Sei  $V$  ein endlichdimensionaler euklidischer Raum und sei  $F \in \text{End}(V)$  selbstadjungiert. Dann zerfällt das charakteristische Polynom  $P_F$  von  $F$  in Linearfaktoren.

*Beweis.* Sei  $\mathcal{B} = (v_1, \dots, v_n)$  eine ONB von  $V$  und sei  $A \in M(n \times n, \mathbb{R})$  die Darstellungsmatrix von  $F$  bezüglich  $\mathcal{B}$ . Nach Satz 6.61(ii) ist  $A$  symmetrisch. Wegen  $P_F = P_A$  folgt die Behauptung des Satzes aus Korollar 6.66.  $\square$

**Satz 6.68 (Spektralsatz für selbstadjungierte Endomorphismen).** Sei  $V$  endlichdimensional mit  $\dim(V) \geq 1$  und sei  $F \in \text{End}(V)$  selbstadjungiert. Dann existiert eine ONB von  $V$  aus Eigenvektoren von  $F$  (und  $F$  wird bezüglich dieser Basis durch eine Diagonalmatrix dargestellt).

*Beweis.* Nach Satz 6.67 zerfällt das charakteristische Polynom von  $F$  in Linearfaktoren und nach Satz 6.62 ist  $F$  normal. Somit folgt die zu zeigende Aussage direkt aus Satz 6.54.  $\square$

**Korollar 6.69.** a) Jede symmetrische Matrix  $A \in M(n \times n, \mathbb{R})$  ist diagonalisierbar (über  $\mathbb{R}$ ).

b) Jede hermitesche Matrix  $A \in M(n \times n, \mathbb{C})$  ist diagonalisierbar (über  $\mathbb{C}$ ).

*Beweis.* Dies folgt aus Satz 6.61 und Satz 6.68.  $\square$

**Beispiel 6.70.** Für die symmetrische Matrix

$$A = \begin{pmatrix} 5 & 4 & 2 \\ 4 & 5 & 2 \\ 2 & 2 & 2 \end{pmatrix} \in M(3 \times 3, \mathbb{R})$$

soll eine ONB von Eigenvektoren bestimmt werden.  $A$  hat das charakteristische Polynom

$$P_A(t) = \det \begin{pmatrix} 5-t & 4 & 2 \\ 4 & 5-t & 2 \\ 2 & 2 & 2-t \end{pmatrix} = (t-10)(t-1)^2.$$

Also ist 10 Eigenwert von  $A$  mit algebraischer Vielfachheit 1 und 1 ist Eigenwert von  $A$  mit algebraischer Vielfachheit 2.

#### Eigenraum zum Eigenwert 10:

Lösen wir das lineare Gleichungssystem  $(A - 10E_3)x = 0$  mit Hilfe des Gauss-Algorithmus, sehen wir, dass  $\{(2, 2, 1)^T\}$  Basis des Eigenraums von  $A$  zum Eigenwert 10 ist. Durch Normierung erhalten wir

$$v_1 := \left( \frac{2}{3}, \frac{2}{3}, \frac{1}{3} \right)^T$$

als normierten Eigenvektor zum Eigenwert 10.

#### Eigenraum zum Eigenwert 1:

Lösen wir das lineare Gleichungssystem  $(A - E_3)x = 0$  mit dem Gauss-Algorithmus, sehen wir, dass  $\{(1, 0, -2)^T, (0, 1, -2)^T\}$  Basis von Eigenvektoren des Eigenraums zum Eigenwert 1. Durch Anwendung des Schmidtschen Orthonormalisierungsverfahrens wird diese Basis wie folgt orthonormiert:

Setze

$$v_2 := \left( \frac{1}{\sqrt{5}}, 0, -\frac{2}{\sqrt{5}} \right)^T$$

und

$$\begin{aligned} v_3' &= (0, 1, -2)^T - \left\langle \left( \frac{1}{\sqrt{5}}, 0, -\frac{2}{\sqrt{5}} \right)^T, (0, 1, -2)^T \right\rangle \left( \frac{1}{\sqrt{5}}, 0, -\frac{2}{\sqrt{5}} \right)^T \\ &= (0, 1, -2)^T - \left( \frac{1}{5} \right) (4, 0, -8)^T \\ &= \left( -\frac{4}{5}, 1, -\frac{2}{5} \right)^T. \end{aligned}$$

und somit

$$v_3 := \left( -\frac{4}{3\sqrt{5}}, \frac{5}{3\sqrt{5}}, -\frac{2}{3\sqrt{5}} \right)^T.$$

Also ist

$$\left\{ \left( \frac{1}{\sqrt{5}}, 0, -\frac{2}{\sqrt{5}} \right)^T, \left( -\frac{4}{3\sqrt{5}}, \frac{5}{3\sqrt{5}}, -\frac{2}{3\sqrt{5}} \right)^T \right\}$$

eine ONB des Eigenraums zum Eigenwert 1. Aus Satz 6.63(ii) folgt dann, dass

$$\mathcal{B} = \left( \left( \frac{2}{3}, \frac{2}{3}, \frac{1}{3} \right)^T, \left( \frac{1}{\sqrt{5}}, 0, -\frac{2}{\sqrt{5}} \right)^T, \left( -\frac{4}{3\sqrt{5}}, \frac{5}{3\sqrt{5}}, -\frac{2}{3\sqrt{5}} \right)^T \right)$$

eine ONB von  $\mathbb{R}^3$  ist, die aus Eigenvektoren von  $A$  besteht, und dass  $F_A \in \text{End}(\mathbb{R}^3)$

bezüglich  $\mathcal{B}$  durch die Matrix

$$\begin{pmatrix} 10 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

dargestellt wird.

## 6.8 Unitäre und orthogonale Abbildungen

Sei  $K = \mathbb{R}$  oder  $K = \mathbb{C}$ .

**Definition 6.71 (Orthogonale und unitäre Abbildung).** Seien  $(V, \langle \cdot, \cdot \rangle)$  und  $(W, \langle \cdot, \cdot \rangle)$  beide euklidisch (bzw. beide unitär) und sei  $F : V \rightarrow W$  linear.

$F$  heißt **orthogonal** (bzw. **unitär**), wenn für alle  $v, v' \in V$

$$\langle F(v), F(v') \rangle = \langle v, v' \rangle$$

gilt.

**Satz 6.72.** Seien  $(V, \langle \cdot, \cdot \rangle)$  und  $(W, \langle \cdot, \cdot \rangle)$  beide euklidisch (bzw. beide unitär) und sei  $F : V \rightarrow W$  linear. Dann sind äquivalent:

- (i)  $F$  ist orthogonal (bzw. unitär).
- (ii) Für alle  $v \in V$  ist  $\|F(v)\| = \|v\|$ .
- (iii) Ist  $\{w_1, \dots, w_n\}$  ein Orthonormalsystem in  $V$ , so ist  $\{F(w_1), \dots, F(w_n)\}$  ein Orthonormalsystem in  $W$ .

*Beweis.* (i)  $\Rightarrow$  (ii): Dies folgt aus der Definition, indem wir  $v = v'$  setzen.

(ii)  $\Rightarrow$  (iii): Sei  $\{w_1, \dots, w_n\}$  ein Orthonormalsystem in  $V$ . Für  $k, l \in \{1, \dots, n\}$  mit  $k \neq l$  gilt:

$$\langle F(w_k + w_l), F(w_k + w_l) \rangle = \langle F(w_k), F(w_k) \rangle + \langle F(w_l), F(w_l) \rangle + \langle F(w_l), F(w_k) \rangle + \langle F(w_k), F(w_l) \rangle. \quad (6.1)$$

Ist  $(W, \langle \cdot, \cdot \rangle)$  euklidisch, so folgt hieraus:

$$\|F(w_k + w_l)\|^2 = \|F(w_k)\|^2 + \|F(w_l)\|^2 + 2\langle F(w_k), F(w_l) \rangle.$$

Also wegen (ii):

$$\langle F(w_k), F(w_l) \rangle = \frac{1}{2} (\|w_k + w_l\|^2 - \|w_k\|^2 - \|w_l\|^2) \stackrel{(\text{Pyth.})}{=} 0.$$

Da  $\|F(w_k)\| = \|w_k\| = 1$  für alle  $k \in \{1, \dots, n\}$ , ist also  $\{F(w_1), \dots, F(w_k)\}$  ein Orthonormalsystem in  $W$ .

Ist  $(W, \langle \cdot, \cdot \rangle)$  unitär, so gilt neben (6.1) auch

$$\langle F(w_k + iw_l), F(w_k + iw_l) \rangle = \langle F(w_k), F(w_k) \rangle - i\langle F(w_k), F(w_l) \rangle + i\langle F(w_l), F(w_k) \rangle + \langle F(w_l), F(w_l) \rangle.$$

Daraus folgt:

$$\|F(w_k + iw_l)\|^2 = \|F(w_k)\|^2 - i\left(\langle F(w_k), F(w_l) \rangle - \overline{\langle F(w_k), F(w_l) \rangle}\right) + \|F(w_l)\|^2.$$

und damit:

$$2\operatorname{Im}\langle F(w_k), F(w_l) \rangle \stackrel{\text{(ii)}}{=} \|F(w_k + iw_l)\|^2 - \|F(w_k)\|^2 - \|F(w_l)\|^2 = \|w_k + iw_l\|^2 - \|w_k\|^2 - \|w_l\|^2 \stackrel{\text{Pyth.}}{=} 0.$$

Aus (6.1) folgt

$$2\operatorname{Re}\langle F(w_k), F(w_l) \rangle = \|F(w_k + w_l)\|^2 - \|F(w_k)\|^2 - \|F(w_l)\|^2 \stackrel{\text{(ii)}}{=} \|w_k + w_l\|^2 - \|w_k\|^2 - \|w_l\|^2 \stackrel{\text{Pyth.}}{=} 0.$$

Also folgt  $\langle F(w_k), F(w_l) \rangle = 0$ , und damit wie oben die Behauptung.

(iii)  $\Rightarrow$  (i): Seien  $v, v' \in V$ . Wir zeigen  $\langle v, v' \rangle = \langle F(v), F(v') \rangle$ .

**1. Fall:**  $v$  und  $v'$  sind linear abhängig.

O.E. existiere ein  $\lambda \in K$  mit  $v = \lambda v'$ . Im Fall  $v' = 0$  ist die Behauptung richtig. Sei also  $v' \neq 0$ . Dann ist  $\{v'/\|v'\|\}$  ein Orthonormalsystem in  $V$ . Nach (iii) ist dann  $\{F(v'/\|v'\|)\}$  ein Orthonormalsystem in  $W$ , d.h.

$$1 = \left\langle F\left(\frac{v'}{\|v'\|}\right), F\left(\frac{v'}{\|v'\|}\right) \right\rangle = \frac{1}{\|v'\|^2} \langle F(v'), F(v') \rangle,$$

also  $\langle v', v' \rangle = \langle F(v'), F(v') \rangle$ . Damit folgt

$$\langle F(v), F(v') \rangle = \langle F(\lambda v'), F(v') \rangle = \lambda \langle F(v'), F(v') \rangle = \lambda \langle v', v' \rangle = \langle \lambda v', v' \rangle = \langle v, v' \rangle.$$

**2. Fall:**  $v$  und  $v'$  sind linear unabhängig.

Nach dem Schmidt'schen Orthonormalisierungsverfahren in Satz 6.40 existiert ein Orthonormalsystem  $\{w_1, w_2\} \subset V$  mit  $\operatorname{span}(v, v') = \operatorname{span}(w_1, w_2)$ . Es existieren also  $\lambda_v, \mu_v, \lambda_{v'}, \mu_{v'} \in K$  mit  $v = \lambda_v w_1 + \mu_v w_2$ ,  $v' = \lambda_{v'} w_1 + \mu_{v'} w_2$ . Dann gilt:

$$\begin{aligned} \langle F(v), F(v') \rangle &= \langle \lambda_v F(w_1) + \mu_v F(w_2), \lambda_{v'} F(w_1) + \mu_{v'} F(w_2) \rangle \\ &= \lambda_v \overline{\lambda_{v'}} \langle F(w_1), F(w_1) \rangle + \mu_v \overline{\lambda_{v'}} \langle F(w_2), F(w_1) \rangle \\ &\quad + \lambda_v \overline{\mu_{v'}} \langle F(w_1), F(w_2) \rangle + \mu_v \overline{\mu_{v'}} \langle F(w_2), F(w_2) \rangle \\ &\stackrel{\text{(iii)}}{=} \lambda_v \overline{\lambda_{v'}} \langle w_1, w_1 \rangle + \mu_v \overline{\lambda_{v'}} \langle w_2, w_1 \rangle \\ &\quad + \lambda_v \overline{\mu_{v'}} \langle w_1, w_2 \rangle + \mu_v \overline{\mu_{v'}} \langle w_2, w_2 \rangle \\ &= \langle v, v' \rangle. \end{aligned}$$

□

**Satz 6.73.** Jede orthogonale (bzw. unitäre) lineare Abbildung ist injektiv.

*Beweis.* Sei  $F : V \rightarrow W$  orthogonal bzw. unitär. Sei  $v \in V$  mit  $F(v) = 0$ . Dann gilt nach Satz 6.72,(ii):

$$0 = \|F(v)\| = \|v\| \Rightarrow v = 0.$$

□

**Satz 6.74.** Seien  $(V, \langle \cdot, \cdot \rangle)$  und  $(W, \langle \cdot, \cdot \rangle)$  beide euklidisch bzw. beide unitär und gegeben sei eine bijektive lineare Abbildung  $F : V \rightarrow W$ . Dann sind äquivalent:

- (i)  $F$  ist orthogonal bzw. unitär,
- (ii)  $F$  besitzt eine adjungierte Abbildung  $F^*$  und es gilt  $F^{-1} = F^*$ .

Gegebenenfalls ist dann  $F$  normal.

*Beweis.* (i)  $\Rightarrow$  (ii): Sei  $v \in V$  und  $w \in W$ . Dann gilt

$$\langle F(v), w \rangle = \langle F(v), F(F^{-1}(w)) \rangle \stackrel{(i)}{=} \langle v, F^{-1}(w) \rangle.$$

Also ist  $F^{-1}$  die adjungierte Abbildung zu  $F$ .

(ii)  $\Rightarrow$  (i): Sei nun  $F^{-1}$  die zu  $F$  adjungierte Abbildung. Dann gilt für alle  $v, v' \in V$ :

$$\langle F(v), F(v') \rangle \stackrel{(ii)}{=} \langle v, F^{-1}(F(v')) \rangle = \langle v, v' \rangle,$$

also ist  $F$  orthogonal bzw. unitär.

Gegebenenfalls ist dann wegen  $F^{-1} \circ F = F \circ F^{-1}$  die Abbildung  $F$  normal. □

**Korollar 6.75.** Sei  $(V, \langle \cdot, \cdot \rangle)$  ein endlichdimensionaler euklidischer bzw. unitärer Raum und sei  $F \in \text{End}(V)$ .

Ist  $F$  orthogonal (bzw. unitär), so ist  $F$  bijektiv und normal mit  $F^* = F^{-1}$ .

*Beweis.* Nach Satz 6.73 ist  $F$  injektiv und wegen  $\dim(V) < \infty$  daher bijektiv. Die Behauptung folgt dann mit Satz 6.74. □

**Satz 6.76.** Sei  $(V, \langle \cdot, \cdot \rangle)$  ein endlichdimensionaler euklidischer bzw. unitärer Raum und sei  $F \in \text{End}(V)$ .

Ist  $F$  orthogonal (bzw. unitär), so haben alle Eigenwerte von  $F$  den Betrag 1. Des Weiteren ist  $|\det(F)| = 1$ .

*Beweis.* Sei  $F \in \text{End}(V)$  orthogonal bzw. unitär. Sei  $\lambda \in K$  ein Eigenwert von  $F$ . Dann existiert  $v \in V \setminus \{0\}$  mit  $F(v) = \lambda v$ . Nach Satz 6.72 ist

$$\|v\| = \|F(v)\| = \|\lambda v\| = |\lambda| \cdot \|v\|,$$

also ist  $|\lambda| = 1$ . Weiter gilt:

$$|\det(F)| = \overline{|\det(F)|} \stackrel{\text{Übung 6.49, Satz 6.50}}{=} |\det(F^*)| = |\det(F^{-1})| = |\det(F)|^{-1}.$$

Daraus folgt  $|\det(F)| = 1$ . □

**Satz 6.77.** Sei  $(V, \langle \cdot, \cdot \rangle)$  ein endlichdimensionaler unitärer Raum und sei  $F \in \text{End}(V)$ . Ist  $F$  normal und haben alle Eigenwerte von  $F$  den Betrag 1, so ist  $F$  unitär.

*Beweis.* Nach Satz 6.54 existiert eine ONB  $\mathcal{B} = (w_1, \dots, w_n)$  von  $V$ , bestehend aus Eigenvektoren von  $F$ . Für  $i \in \{1, \dots, n\}$  sei  $\lambda_i$  der Eigenwert zum Eigenvektor  $w_i$ . Dann wird  $F$  bezüglich  $\mathcal{B}$  durch die Matrix

$$\mathcal{M}_{\mathcal{B}}(F) = \begin{pmatrix} \lambda_1 & & \\ & \ddots & \\ & & \lambda_n \end{pmatrix}$$

dargestellt.

Nach Satz 6.50 wird dann  $F^*$  bezüglich  $\mathcal{B}$  durch

$$\begin{pmatrix} \overline{\lambda_1} & & \\ & \ddots & \\ & & \overline{\lambda_n} \end{pmatrix}$$

dargestellt. Wegen  $|\lambda_i| = 1$  gilt  $\lambda_i^{-1} = \overline{\lambda_i}$ . Insbesondere ist  $F$  invertierbar, und  $F^{-1}$  wird bezüglich  $\mathcal{B}$  durch

$$\begin{pmatrix} \lambda_1^{-1} & & \\ & \ddots & \\ & & \lambda_n^{-1} \end{pmatrix}$$

dargestellt. Also ist  $F^* = F^{-1}$  und  $F$  ist unitär. □

**Definition 6.78 (Orthogonale bzw. unitäre Matrix).** a)  $A \in M(n \times n, \mathbb{R})$  heißt **orthogonal**, falls  $A$  invertierbar ist und  $A^{-1} = A^T$ .  
b)  $A \in M(n \times n, \mathbb{C})$  heißt **unitär**, falls  $A$  invertierbar ist und  $A^{-1} = A^*$ .

**Satz 6.79.** Sei  $(V, \langle \cdot, \cdot \rangle)$  ein endlichdimensionaler euklidischer bzw. unitärer Raum und sei  $F \in \text{End}(V)$ . Dann sind äquivalent:

- (i)  $F$  ist orthogonal (bzw. unitär),

- (ii) Bezüglich einer (jeder) ONB von  $V$  wird  $F$  durch eine orthogonale (bzw. unitäre) Matrix dargestellt.

*Beweis.* Sei  $\mathcal{B}$  eine ONB von  $V$ . Wenn  $F$  bezüglich  $\mathcal{B}$  durch  $A$  dargestellt wird, so wird nach Satz 6.50  $F^*$  bezüglich  $\mathcal{B}$  durch  $A^*$  dargestellt. Außerdem wird  $F^{-1}$  bezüglich  $\mathcal{B}$  durch  $A^{-1}$  dargestellt. Deshalb gilt:

$$F^{-1} = F^* \iff A^{-1} = A^*$$

und die Behauptung ist gezeigt.  $\square$

**Definition 6.80 (Orthogonale bzw. unitäre Gruppe).** a) Sei  $(V, \langle \cdot, \cdot \rangle)$  ein endlichdimensionaler euklidischer bzw. unitärer Raum. Die Menge  $O(V)$  aller orthogonalen bzw. die Menge  $U(V)$  aller unitären Endomorphismen von  $V$  bildet mit der Hintereinanderausführung von Abbildungen als Verknüpfung eine Gruppe: die **orthogonale** bzw. **unitäre Gruppe** von  $V$ .

- b) Die Menge  $O(n)$  aller orthogonalen bzw. die Menge  $U(n)$  aller unitären  $n \times n$ -Matrizen mit der Matrizenmultiplikation als Verknüpfung bilden eine Gruppe: die **orthogonale** bzw. **unitäre Gruppe**.

**Satz 6.81.** Für  $A \in M(n \times n, K)$  sind äquivalent:

1.  $A$  ist orthogonal (bzw. unitär).
2. Die Zeilenvektoren von  $A = (a_{ij})$  bilden ein Orthonormalsystem (und damit eine ONB) in  $K^n$  bzgl. dem Standardskalarprodukt, d.h.

$$\sum_{k=1}^n a_{ik} \overline{a_{jk}} = \begin{cases} 0, & i \neq j, \\ 1, & i = j. \end{cases}$$

3. Die Spaltenvektoren von  $A = (a_{ij})$  bilden ein Orthonormalsystem (und damit eine ONB) in  $K^n$  bzgl. dem Standardskalarprodukt, d.h.

$$\sum_{k=1}^n a_{ki} \overline{a_{kj}} = \begin{cases} 0, & i \neq j, \\ 1, & i = j. \end{cases}$$

*Beweis.* 2. ist äquivalent zu  $AA^* = E_n$  und damit zu  $A^* = A^{-1}$ , also zu 1.

3. ist äquivalent zu  $A^*A = E_n$  und damit zu  $A^* = A^{-1}$ , also zu 1.  $\square$

**Definition 6.82 (orthogonal bzw. unitär diagonalisierbar).** Eine Matrix  $A \in M(n \times n, K)$  heißt **orthogonal** (bzw. **unitär**) **diagonalisierbar**, falls eine orthogonale (bzw. unitäre) Matrix  $S \in M(n \times n, K)$  existiert, sodass  $S^{-1}AS (= S^*AS)$  eine Diagonalmatrix ist.

**Satz 6.83.** Jede normale  $n \times n$ -Matrix  $A$ , deren charakteristisches Polynom in Linearfaktoren zerfällt, ist orthogonal (bzw. unitär) diagonalisierbar.

*Beweis.* Nach Satz 6.57 ist  $F_A \in \text{End}(K^n)$  normal, wobei  $K^n$  mit dem Standardskalarprodukt versehen ist. Nach Satz 6.54 existiert eine ONB  $(w_1, \dots, w_n)$  von  $K^n$  aus Eigenvektoren von  $F_A$ . Sei  $S$  die Übergangsmatrix von der kanonischen Basis nach  $(w_1, \dots, w_n)$ , d.h.  $w_i$  ist der  $i$ -te Spaltenvektor von  $S$ . Nach Satz 6.81 ist  $S$  deshalb orthogonal bzw. unitär. Nach Lineare Algebra I, Satz 3.53, wird  $F_A$  bezüglich  $(w_1, \dots, w_n)$  durch  $S^{-1}AS$  dargestellt. Da  $(w_1, \dots, w_n)$  Eigenvektoren von  $F_A$  sind, ist deshalb  $S^{-1}AS$  eine Diagonalmatrix.  $\square$

**Korollar 6.84 (Hauptachsentransformation).** Jede symmetrische bzw. hermitesche  $n \times n$ -Matrix ist orthogonal bzw. unitär diagonalisierbar.

*Beweis.* Offensichtlich ist jede symmetrische bzw. hermitesche Matrix normal. Die Behauptung folgt aus Korollar 6.66 und Satz 6.83.  $\square$

**Definition 6.85 (Drehung).** Sei  $(V, \langle \cdot, \cdot \rangle)$  ein endlichdimensionaler euklidischer Raum. Eine orthogonale Abbildung  $F: V \rightarrow V$  heißt eine **Drehung**, falls  $\det(F) = 1$ .

**Satz 6.86.** Sei  $\mathbb{R}^2$  versehen mit dem Standardskalarprodukt  $\langle \cdot, \cdot \rangle$ . Für  $F \in \text{End}(\mathbb{R}^2)$  sind äquivalent:

1.  $F$  ist eine Drehung.
2.  $F = F_A$ , wobei

$$A = \begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}$$

mit  $\alpha \in [0, 2\pi[$ .

*Beweis.* **1.  $\Rightarrow$  2.:** Sei  $(e^1, e^2)$  die kanonische Basis von  $\mathbb{R}^2$ . Dann wird nach Korollar 6.39  $F$  bezüglich  $(e^1, e^2)$  durch die Matrix

$$A = \begin{pmatrix} \langle F(e^1), e^1 \rangle & \langle F(e^2), e^1 \rangle \\ \langle F(e^1), e^2 \rangle & \langle F(e^2), e^2 \rangle \end{pmatrix}$$

dargestellt. Es ist also  $F = F_A$ . Nach Satz 6.79 und Satz 6.81 folgt:

$$\langle F(e^1), e^1 \rangle^2 + \langle F(e^1), e^2 \rangle^2 = 1.$$

Also gibt es genau ein  $\alpha \in [0, 2\pi[$  mit

$$\langle F(e^1), e^1 \rangle = \cos \alpha, \quad \langle F(e^1), e^2 \rangle = \sin \alpha.$$

Nach Satz 6.81 gilt:

$$(\cos \alpha) \langle F(e^2), e^1 \rangle + (\sin \alpha) \langle F(e^2), e^2 \rangle = 0.$$

Außerdem ist

$$-(\sin \alpha) \langle F(e^2), e^1 \rangle + (\cos \alpha) \langle F(e^2), e^2 \rangle = \det(F) = 1.$$

Durch Lösung dieses Gleichungssystems erhält man:

$$\langle F(e^2), e^1 \rangle = -\sin \alpha, \quad \langle F(e^2), e^2 \rangle = \cos \alpha,$$

und damit 2.

2  $\Rightarrow$  1: Da jede Matrix der Form

$$\begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix}$$

orthogonal mit Determinante 1 ist, folgt die Behauptung aus Satz 6.79.  $\square$

| **Bemerkung 6.87.** Der in Satz 6.86 eingeführte Winkel  $\alpha$  heißt der **Drehwinkel** von  $F$ .

**Satz 6.88.** Sei  $\mathbb{R}^3$  versehen mit dem Standardskalarprodukt und sei  $F \in \text{End}(\mathbb{R}^3)$ . Dann sind äquivalent:

1.  $F$  ist eine Drehung.
2. Es existiert eine ONB  $(w_1, w_2, w_3)$  von  $\mathbb{R}^3$ , bezüglich der  $F$  durch eine Matrix der Form

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \alpha & -\sin \alpha \\ 0 & \sin \alpha & \cos \alpha \end{pmatrix} \quad \text{mit } \alpha \in [0, 2\pi[$$

dargestellt wird.

*Beweis.* 1  $\Rightarrow$  2: Das charakteristische Polynom  $P_F$  von  $F$  ist von der Form  $P_F(t) = t^3 + a_2 t^2 + a_1 t + a_0$  mit  $a_0, a_1, a_2 \in \mathbb{R}$ . Es gilt  $a_0 = (-1)^3 \det(F) = -1$ .

Aus dem Zwischenwertsatz folgt, dass  $P_F$  wenigstens eine reelle Nullstelle  $\lambda_1$  besitzt. Seien  $\lambda_2, \lambda_3$  die weiteren Nullstellen von  $P_F$  in  $\mathbb{C}$  (nicht notwendig voneinander oder von  $\lambda_1$  verschieden). Die Abbildung  $F$  werde bezüglich der kanonischen Basis von  $\mathbb{R}^3$  durch die Matrix  $A$  dargestellt. Nach Satz 6.79 ist  $A$  orthogonal. Es ist  $P_F(t) = P_A(t)$ . Da  $A$  bezüglich der kanonischen Basis von  $\mathbb{C}^3$  eine unitäre Abbildung darstellt, gilt nach Satz 6.76, dass alle komplexen Eigenwerte von  $A$ , d.h. alle komplexen Nullstellen von  $A$ , den Betrag 1

haben. Also folgt:

$$|\lambda_1| = |\lambda_2| = |\lambda_3| = 1,$$

insbesondere  $\lambda_1 = 1$  oder  $\lambda_1 = -1$ . Nun gilt stets  $\lambda_1\lambda_2\lambda_3 = -a_0 = 1$ .

Sind  $\lambda_1, \lambda_2, \lambda_3 \in \mathbb{R}$ , so folgt, dass mindestens ein  $\lambda_i = 1$  ist. O.E. sei dann  $\lambda_1 = 1$ .

Ist  $\lambda_2 \in \mathbb{C} \setminus \mathbb{R}$ , so gilt:

$$P_F(\lambda_2) = \lambda_2^3 + a_2\lambda_2^2 + a_1\lambda_2 + a_0 = \overline{\lambda_2}^3 + a_2\overline{\lambda_2}^2 + a_1\overline{\lambda_2} + a_0 = 0,$$

also ist  $\lambda_3 = \overline{\lambda_2}$ . Damit ist  $\lambda_2\lambda_3 = \lambda_2\overline{\lambda_2} = |\lambda_2|^2 = 1$ , also ebenfalls  $\lambda_1 = 1$ .

Sei nun  $w_1 \in \mathbb{R}^3$  ein normierter Eigenvektor von  $F$  zum Eigenwert  $\lambda_1 = 1$ . Sei  $U := \{w_1\}^\perp$ . Dann gilt  $F(U) \subseteq U$  (direktes Nachrechnen). Sei  $(w'_2, w'_3)$  eine ONB von  $U$ . Dann wird  $F$  bezüglich der ONB  $(w_1, w'_2, w'_3)$  von  $V$  durch eine Matrix der Form

$$\begin{pmatrix} 1 & 0 \\ 0 & A' \end{pmatrix}$$

dargestellt, wobei  $F|_U$  bezüglich  $(w'_2, w'_3)$  durch  $A'$  dargestellt wird. Da  $F|_U$  orthogonal ist und

$$1 = \det(F) = 1 \cdot \det(A') = \det(F|_U),$$

ist  $F|_U$  eine Drehung. Deshalb existiert eine ONB  $(w_2, w_3)$  von  $U$ , sodass  $F|_U$  bezüglich der ONB  $(w_2, w_3)$  durch eine Matrix der Form

$$\begin{pmatrix} \cos \alpha & -\sin \alpha \\ \sin \alpha & \cos \alpha \end{pmatrix} \quad \text{mit } \alpha \in [0, 2\pi[$$

dargestellt wird (Beweis analog zum Beweis von Satz 6.88). Also wird  $F$  bezüglich der ONB  $(w_1, w_2, w_3)$  von  $\mathbb{R}^3$  durch eine Matrix der gewünschten Form dargestellt.

2.  $\Rightarrow$  1.: Offensichtlich ist jede Matrix von der in 2. angegebenen Form orthogonal und hat Determinante 1, stellt also eine Drehung dar.  $\square$

**Bemerkung 6.89.** Der von  $w_1$  aufgespannte Unterraum heißt die **Drehachse** von  $F$  und  $U := \text{span}(w_2, w_3)$  heißt die **Drehebene** von  $F$ . Die Zahl  $\alpha$  heißt der Drehwinkel von  $F$ , welche nur dann eindeutig bestimmt ist, wenn eine Orientierung in der Drehebene festgelegt ist.

## 6.9 Symmetrische Bilinearformen und positiv definite Matrizen

Sei  $(V, \langle \cdot, \cdot \rangle)$  ein endlichdimensionaler euklidischer Raum mit  $\dim(V) = n$ . Wir werden in diesem Kapitel nur am Rande auf unitäre Räume eingehen.

**Satz 6.90** (Charakterisierung der symmetrischen Bilinearformen auf  $V$ ). Sei  $s : V \times V \rightarrow \mathbb{R}$ . Dann sind äquivalent:

1.  $s$  ist eine symmetrische Bilinearform.
2. Es existiert eine selbstadjungierte Abbildung  $F \in \text{End}(V)$ , so dass für alle  $v, w \in V$

$$s(v, w) = \langle F(v), w \rangle.$$

$F$  ist unter der Bedingung 1. (oder 2.) eindeutig bestimmt.

*Beweis.* 1.  $\Rightarrow$  2.: Sei  $(w_1, \dots, w_n)$  eine ONB von  $V$  und sei  $A = (a_{ij})$  die Darstellungsmatrix von  $s$  bzgl. dieser ONB. Da  $s$  symmetrisch ist, ist nach Satz 6.10  $A$  symmetrisch. Nach Satz 6.61 stellt  $A$  bzgl.  $(w_1, \dots, w_n)$  einen selbstadjungierten Endomorphismus  $F$  von  $V$  dar.

Seien  $v, w \in V$  und seien  $\lambda_1, \dots, \lambda_n, \mu_1, \dots, \mu_n \in \mathbb{R}$  mit

$$v = \lambda_1 w_1 + \dots + \lambda_n w_n, \quad w = \mu_1 w_1 + \dots + \mu_n w_n.$$

Dann gilt:

$$\begin{aligned} s(v, w) &= \sum_{i,j=1}^n a_{ij} \lambda_i \mu_j = \sum_{i=1}^n \lambda_i \left( \sum_{j=1}^n a_{ij} \mu_j \right) \\ &= \left\langle \sum_{i=1}^n \lambda_i w_i, \sum_{i=1}^n \left( \sum_{j=1}^n a_{ij} \mu_j \right) w_i \right\rangle \\ &= \langle v, F(w) \rangle \\ &= \langle F(v), w \rangle. \end{aligned}$$

**Übung 6.91.** Beweise die Gültigkeit der 2. Gleichung.

2.  $\Rightarrow$  1.: Direktes Nachprüfen liefert, dass  $(v, w) \mapsto \langle F(v), w \rangle$  eine symmetrische Bilinearform ist.

**Eindeutigkeit:**

Seien  $F, G \in \text{End}(V)$  selbstadjungiert mit  $\langle F(v), w \rangle = \langle G(v), w \rangle$  für alle  $v, w \in V$ . Dann ist  $\langle (F - G)(v), w \rangle = 0$ . Setzen wir speziell  $w := (F - G)(v)$ , erhalten wir

$$\|F(v) - G(v)\|^2 = 0,$$

also  $F(v) = G(v)$  für alle  $v \in V$ . □

**Bemerkung 6.92.** Der obige Beweis zeigt, dass die Darstellungsmatrix von  $s$  bzgl. einer ONB  $\mathcal{B}$  gleich der Darstellungsmatrix des zugehörigen selbstadjungierten Endomorphismus  $F$  bzgl.  $\mathcal{B}$  ist.

**Beispiel 6.93.** Der  $\mathbb{R}^n$  sei mit dem Standardskalarprodukt versehen. Sei  $A \in M(n \times n, \mathbb{R})$  symmetrisch. Dann ist die zur symmetrischen Bilinearform  $s_A$  gehörige selbstadjungierte Abbildung gegeben durch  $F_A$ .

**Satz 6.94.** Sei  $s : V \times V \rightarrow \mathbb{R}$  eine symmetrische Bilinearform. Dann existiert eine ONB von  $V$ , so dass die zugehörige Darstellungsmatrix von  $s$  eine Diagonalmatrix  $D \in M(n \times n, \mathbb{R})$  ist.

*Beweis.* Nach Satz 6.90 existiert ein selbstadjungiertes  $F \in \text{End}(V)$  mit  $s(v, w) = \langle F(v), w \rangle$  für alle  $v, w \in V$ . Nach Satz 6.68 existiert eine ONB  $(w_1, \dots, w_n)$  von Eigenvektoren von  $F$ . Die Darstellungsmatrix  $D \in M(n \times n, \mathbb{R})$  von  $F$  bzgl.  $(w_1, \dots, w_n)$  und damit die Darstellungsmatrix von  $s$  bzgl.  $(w_1, \dots, w_n)$  ist daher eine Diagonalmatrix.  $\square$

**Bemerkung 6.95.** In der Diagonalen von  $D$  stehen gerade die Eigenwerte der Darstellungsmatrix von  $s$  bzgl. einer beliebigen ONB von  $V$ .

**Definition 6.96 (positiv (bzw. negativ) (semi-)definite Matrix).** Eine symmetrische Matrix  $A \in M(n \times n, \mathbb{R})$  (oder hermitesche Matrix  $A \in M(n \times n, \mathbb{C})$ ) heißt **positiv** (bzw. **negativ**) **semidefinit**, falls die symmetrische Bilinearform

$$s_A : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}, \quad s_A(x, y) := x^T A y$$

oder

$$s_A : \mathbb{C}^n \times \mathbb{C}^n \rightarrow \mathbb{C}, \quad s_A(x, y) := x^T A \bar{y}$$

positiv (bzw. negativ) semidefinit ist. Sie heißt **positiv** (bzw. **negativ**) **definit**, falls  $s_A$  positiv (bzw. negativ) definit ist.

**Bemerkung 6.97.**  $A$  ist negativ (semi-)definit genau dann, wenn  $-A$  positiv (semi-)definit ist.

**Satz 6.98 (Eigenwertkriterium für positive Definitheit).** Sei  $A \in M(n \times n, \mathbb{R})$  symmetrisch. Dann gilt:

1.  $A$  ist positiv semidefinit genau dann, wenn jeder Eigenwert von  $F_A$  größer gleich 0 ist.
2.  $A$  ist positiv definit genau dann, wenn jeder Eigenwert von  $F_A$  größer 0 ist.

*Beweis.* Nach Satz 6.94 existiert eine ONB  $(w_1, \dots, w_n)$  von  $\mathbb{R}^n$  mit dem Standardskalarprodukt, so dass die Darstellungsmatrix von  $s_A$  bzgl.  $(w_1, \dots, w_n)$  eine Diagonalmatrix  $D$  ist. Nach Bemerkung 6.95 stehen auf der Diagonalen von  $D$  gerade die Eigenwerte von

$F_A$ . Sei

$$D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}.$$

Es gilt für  $x = \mu_1 w_1 + \dots + \mu_n w_n$

$$\begin{aligned} s_A(x, x) &= \sum_{i=1}^n \sum_{j=1}^n \mu_i \mu_j s_A(w_i, w_j) \\ &= \sum_{i=1}^n \sum_{j=1}^n \mu_i \mu_j w_i^T A w_j \\ &= \sum_{i=1}^n \mu_i^2 \lambda_i. \end{aligned}$$

Es gilt also  $s_A(x, x) \geq 0$  für alle  $x \in \mathbb{R}^n$  genau dann, wenn  $\sum_{i=1}^n \mu_i^2 \lambda_i \geq 0$  für alle  $(\mu_1, \dots, \mu_n) \in \mathbb{R}^n$ , also genau dann, wenn  $\lambda_i \geq 0$  für alle  $i \in \{1, \dots, n\}$ . Analog ist  $s_A(x, x) > 0$  für alle  $x \in \mathbb{R}^n \setminus \{0\}$  genau dann, wenn  $\sum_{i=1}^n \mu_i^2 \lambda_i > 0$  für alle  $(\mu_1, \dots, \mu_n) \in \mathbb{R}^n \setminus \{0\}$ , also genau dann, wenn  $\lambda_i > 0$  für alle  $i \in \{1, \dots, n\}$ .  $\square$

**Lemma 6.99.** Sei  $A \in M(n \times n, \mathbb{R})$  symmetrisch und positiv definit. Dann gilt  $\det(A) > 0$ .

*Beweis.* Da  $A$  symmetrisch und damit nach Korollar 6.69 diagonalisierbar ist, folgt

$$\det(A) = \lambda_1 \cdot \dots \cdot \lambda_n,$$

wobei  $\lambda_1, \dots, \lambda_n$  die (nicht notwendigen verschiedenen) Eigenwerte von  $F_A$  sind. Nach Satz 6.98 gilt  $\lambda_i > 0$  für alle  $i \in \{1, \dots, n\}$ , also  $\det(A) > 0$ .  $\square$

**Satz 6.100 (Determinantenkriterium für positive Definitheit).** Sei  $A = (a_{ij}) \in M(n \times n, \mathbb{R})$  symmetrisch. Dann sind äquivalent:

- (i)  $A$  ist positiv definit.
- (ii) Für alle  $k \in \{1, \dots, n\}$  gilt

$$\det \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{k1} & \cdots & a_{kk} \end{pmatrix} > 0.$$

*Beweis.* (i)  $\Rightarrow$  (ii): Sei  $(e^1, \dots, e^n)$  die kanonische Basis von  $\mathbb{R}^n$ . Für  $k \in \{1, \dots, n\}$  sei

$$s_k : \mathbb{R}^k \times \mathbb{R}^k \rightarrow \mathbb{R}, \quad s_k((x_1, \dots, x_k), (y_1, \dots, y_k)) = s_A \left( \sum_{i=1}^k x_i e^i, \sum_{i=1}^k y_i e^i \right).$$

Dann ist  $s_k$  offensichtlich eine positiv definite symmetrische Bilinearform auf  $\mathbb{R}^k$  und  $s_k = s_{A_k}$  mit

$$A_k := \begin{pmatrix} a_{11} & \cdots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{k1} & \cdots & a_{kk} \end{pmatrix}.$$

Damit ist  $A_k$  positiv definit und hat deshalb nach Lemma 6.99 positive Determinante.

(ii)  $\Rightarrow$  (i): Beweis durch Induktion nach  $n$ .

$n = 1$ : Dies ist klar.

$n \rightarrow n + 1$ : Sei  $A \in M((n + 1) \times (n + 1), \mathbb{R})$  symmetrisch, so dass (ii) erfüllt ist. Sei  $(e^1, \dots, e^n, e^{n+1})$  die kanonische Basis von  $\mathbb{R}^{n+1}$  und  $U := \text{span}(e^1, \dots, e^n)$ .

Wir setzen

$$C := A_n = \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix}.$$

Nach Induktionsvoraussetzung ist  $C = A_n$  positiv definit. Sei  $s_C := s|_{U \times U}$  diejenige positiv definite symmetrische Bilinearform auf  $U \cong \mathbb{R}^n$ , die bzgl.  $(e^1, \dots, e^n)$  die Darstellungsmatrix  $C$  hat. Nach Satz 6.94 existiert eine ONB  $(w_1, \dots, w_n)$  von  $U$ , so dass die Darstellungsmatrix von  $s_C$  bzgl.  $(w_1, \dots, w_n)$  eine reelle Diagonalmatrix ist, deren Diagonalelemente  $\lambda_1, \dots, \lambda_n$  nach Bemerkung 6.95 und Satz 6.98 alle positiv sind.

**Behauptung:** Es existiert  $w_{n+1} \in \mathbb{R}^{n+1} \setminus \{0\}$  mit  $s_A(x, w_{n+1}) = 0$  für alle  $x \in U$ .

**Beweis der Behauptung:** Nach Satz 6.90 existiert ein selbstadjungiertes  $F \in \text{End}(\mathbb{R}^{n+1})$  mit  $s_A(x, y) = \langle F(x), y \rangle$  für alle  $x, y \in \mathbb{R}^{n+1}$ . Da  $\dim(U) \leq n$  existiert ein normierter Vektor  $w_{n+1} \in F(U)^\perp$ . Dann gilt für jedes  $x \in U$

$$s_A(x, w_{n+1}) = \langle F(x), w_{n+1} \rangle = 0,$$

und die Behauptung ist bewiesen.

Setze  $\lambda_{n+1} := s_A(w_{n+1}, w_{n+1})$ . Dann ist die Darstellungsmatrix von  $s_A$  bzgl.  $(w_1, \dots, w_n, w_{n+1})$  gleich

$$D = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_{n+1} \end{pmatrix}.$$

Nach Bemerkung 6.92 sind  $A$  und  $D$  Darstellungsmatrizen von  $F$  und somit sind  $A$  und



Beachte, dass die Diagonaleigenschaft von Diagonalmatrizen bzgl. Transposition sowie Multiplikation invariant ist. Für  $D = \text{diag}(\lambda_1, \dots, \lambda_r, 0, \dots, 0) \in M(m \times n, K)$  gilt z. B.:

$$D^T D = \text{diag}(\lambda_1^2, \dots, \lambda_r^2, 0, \dots, 0) \in M(n \times n, K).$$

Mit diesen Vorüberlegungen können wir den angekündigten Normalformensatz beweisen. Für den Beweis ist es praktisch, zunächst die Matrixform dieses Satzes zu betrachten.

**Satz 6.104 (Singularwertzerlegung - Matrixform).** Zu jeder Matrix  $A \in M(m \times n, K)$  gibt es orthogonale bzw. unitäre Matrizen  $S \in M(m \times m, K)$  und  $T \in M(n \times n, K)$ , so dass

$$S^{-1} A T = \bar{S}^T A T = \text{diag}(\lambda_1, \dots, \lambda_r, 0, \dots, 0) =: D \in M(m \times n, \mathbb{R})$$

eine Diagonalmatrix mit reellen positiven Diagonaleinträgen  $\lambda_1, \dots, \lambda_r$  ist.

Die  $\lambda_1, \dots, \lambda_r$  sind dabei bis auf die Reihenfolge eindeutig bestimmt und heißen die **Singularwerte** der Matrix  $A$ . Die Darstellung

$$A = S D T^{-1} = S D T^T$$

heißt die **Singularwertzerlegung von  $A$** .

*Beweis.* Wir zeigen zunächst die Eindeutigkeit der Singularwerte: Sind  $S$  und  $T$  beliebige orthogonale bzw. unitäre Matrizen der passenden Größen, so dass

$$\bar{S}^T A T = \text{diag}(\lambda_1, \dots, \lambda_r, 0, \dots, 0) =: D,$$

eine Diagonalmatrix mit reellen positiven  $\lambda_1, \dots, \lambda_r$  ist, so folgt durch Transponieren und Konjugieren auch

$$\bar{T}^T \bar{A}^T S = \bar{D}^T = D^T,$$

und damit

$$T^{-1} \bar{A}^T A T = \bar{T}^T \bar{A}^T S \bar{S}^T A T = D^T D = \text{diag}(\lambda_1^2, \dots, \lambda_r^2, 0, \dots, 0) \in M(n \times n, K).$$

Die Matrix  $\bar{A}^T A$  ist also ähnlich zur Diagonalmatrix  $D^T D$  und hat damit dieselben Eigenwerte. Die Eigenwerte  $\lambda_1^2, \dots, \lambda_r^2$  (und damit auch  $\lambda_1, \dots, \lambda_r$ ) sind daher bis auf ihre Reihenfolge eindeutig bestimmt, nämlich als die Eigenwerte von  $\bar{A}^T A \neq 0$ .

Der Existenzbeweis der Singularwertzerlegung ist konstruktiv und erfolgt in zwei Schritten, wobei der erste an den den gerade geführten Eindeutigkeitsbeweis angelehnt ist:

(a) **Bestimmung von  $T$ .**

Die Matrix  $\bar{A}^T A \in M(n \times n, K)$  ist symmetrisch bzw. hermitesch, da

$$\overline{(\bar{A}^T A)}^T = \bar{A}^T A.$$

Nach Korollar 6.84 gibt es also eine orthogonale (oder unitäre) Matrix  $T \in M(n \times n, K)$ , so

dass  $T^T(\overline{A}^T A)T$  eine Diagonalmatrix ist mit den Eigenwerten von  $F_{(\overline{A}^T A)}$  in der Diagonalen. Aus Korollar 6.65 folgt, dass alle diese Eigenwerte reell sind. Nun gilt

$$x^T(\overline{A}^T A)\overline{x} = (\overline{A}x)^T A\overline{x} = (\overline{A}x)^T \overline{A}x = \langle \overline{A}x, \overline{A}x \rangle = \|\overline{A}x\|^2 \geq 0$$

für alle  $x \in K^n$ , wobei  $\|\cdot\|$  die Norm des Standardskalarprodukts ist. Somit folgt, dass  $(\overline{A}^T A)$  positiv semi-definit ist. Die Einträge der Diagonalmatrix  $\overline{T}^T \overline{A}^T A T$  sind also nicht-negativ, und damit können wir (nach geeigneter Anordnung der Spalten von  $T$ )

$$\overline{T}^T \overline{A}^T A T = \text{diag}(\lambda_1^2, \dots, \lambda_r^2, 0, \dots, 0) =: \tilde{D} \quad (6.2)$$

mit reellen  $\lambda_1, \dots, \lambda_r > 0$  für ein  $r \leq n$  schreiben.

(b) **Bestimmung von  $S$ .**

Für alle  $i \leq r$  setzen wir

$$s_i := \frac{1}{\lambda_i} A t^i \in K^m. \quad (6.3)$$

Diese Vektoren sind bezüglich des Standardskalarprodukts orthonormal, denn für alle  $i, j \leq r$  gilt

$$\overline{s_i}^T s_j = \frac{1}{\lambda_i \lambda_j} \overline{A t^i}^T A t^j = \frac{1}{\lambda_i \lambda_j} \overline{t^i}^T \overline{A}^T A t^j \stackrel{(6.2)}{=} \frac{1}{\lambda_i \lambda_j} \tilde{d}_{ij} \delta_{ij} = \delta_{ij}.$$

Wir können die Vektoren  $s^i$  nun zu einer ONB von  $K^m$  ergänzen (es folgt also  $r \leq m$ ) und daraus bilden wir die unitäre Matrix

$$S = (s_1, \dots, s_m).$$

Wir werden zeigen, dass  $\overline{S}^T A T = D$ , also  $A T = S D$  gilt. Die Gleichheit dieser Matrizen überprüfen wir spaltenweise, d.h. wir zeigen  $A t^i = S d^i$  für alle  $i = 1, \dots, n$ .

- Für  $i \leq r$  gilt  $A t^i = s_i \lambda_i = S d^i$  wegen (6.3).
- Für  $i > r$  betrachten wir:

$$\|\overline{A t^i}\|^2 = \overline{A t^i}^T A t^i = \overline{t^i}^T \overline{A}^T A t^i \stackrel{(6.2)}{=} \tilde{d}_{ii} \delta_{ii} = 0$$

und damit  $A t^i = 0 = S d^i$ , da die  $i$ -te Spalte  $d^i$  von  $D$  eine Nullspalte ist.

Mit  $\overline{S}^T A T = D$  sind  $A$  und  $D$  auch äquivalent, so dass  $\text{rang}(A) = \text{rang}(D) = r$ .  $\square$

**Bemerkung 6.105.** Wie am Anfang dieses Abschnitts erwähnt, gilt der Satz über die Singulärwertzerlegung auch allgemein für lineare Abbildungen  $F : V \rightarrow W$  zwischen euklidischen bzw. unitären Räumen. Dabei ist  $S$  eine ONB von  $W$ ,  $T$  eine ONB von  $V$  und  $D = \text{diag}(\lambda_1, \dots, \lambda_r, 0, \dots, 0)$  mit  $\lambda_i > 0$  so dass wir für die Darstellungsmatrix von  $F$  bezüglich  $\mathcal{A}$  und  $\mathcal{B}$  die Form

$$\mathcal{M}_{\mathcal{A}}^{\mathcal{B}}(F) = D$$

erhalten.